

371.26013

๘๘4๑๐

๘.๓.

การศึกษาวิธีการเทียบคะแนนแบบเส้นตรงตามแบบจำลองคะแนนจริงสัมพัทธ์

19 S.A. 2539

ปริญญาโท

ของ

สุนนา โรตติพลอนันต์

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา

ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

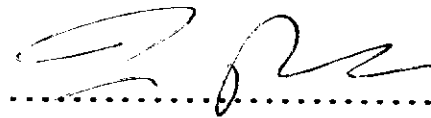
พฤษภาคม 2539

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

๘. 54102

คณะกรรมการควบคุมและคณะกรรมการสอบได้พิจารณาปัญหานี้เรียบร้อยแล้ว เห็น
สมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาบัณฑิต การศึกษามหาบัณฑิต วิชาเอก
การวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒได้

คณะกรรมการควบคุม

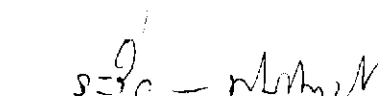

.....ประธาน
(รศ.ดร.บุญเชิด บุญโง่นนันทพงษ์)


.....กรรมการ
(อ.สุพา มานะจิตต์)

คณะกรรมการสอบ


.....ประธาน
(รศ.ดร.บุญเชิด บุญโง่นนันทพงษ์)


.....กรรมการ
(อ.สุพา มานะจิตต์)


.....กรรมการที่แต่งตั้งเพิ่มเติม
(อ.ระวีวรรณ พันธุ์พานิช)

บัณฑิตวิทยาลัยอนุมัติให้รับปัญหานี้เรียบร้อยแล้ว เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษาบัณฑิต วิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒได้


.....คณบดีบัณฑิตวิทยาลัย
(ดร.ศิริสุภา พุฒสุวรรณ)

วันที่..1..เดือน..พฤษภาคม..พ.ศ.2539..

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้ สำเร็จออกมาได้ก็เพราะได้รับความช่วยเหลือ และร่วมมืออย่างดียิ่งจาก รศ.ดร.บุญเชิด ภิญโญอนันตพงษ์ และอ.บุพา มานะจิตต์ รวมทั้ง อ.ระวีวรรณ พันธุ์พานิช ที่ได้กรุณาให้คำแนะนำ ชี้คิดเห็น ตลอดจนการแก้ไขข้อบกพร่องต่าง ๆ ซึ่งผู้วิจัยรู้สึกซาบซึ้งยิ่งนัก และขอกราบขอบพระคุณมา ณ ที่นี้ด้วย

ขอกราบขอบพระคุณ รศ.ชาญวิทย์ เทียมบุญประเสริฐ ผศ.ทองหล่อ วิภาวิน และ ผศ.จันทิมา พรหมโชติกุล จากสำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร ที่กรุณาช่วยตรวจเครื่องมือที่ใช้ในการวิจัย จนสามารถทำให้การศึกษาลุล่วงไปได้ด้วยดี

ขอกราบขอบพระคุณ คณาจารย์ที่เป็นกรรมการในการสอบเค้าโครงปริญญานิพนธ์ ที่ช่วยทำให้การศึกษากลายเป็นไปอย่างราบรื่น และมีคุณค่า

ขอขอบพระคุณ ผู้อำนวยการวิทยาลัยพยาบาลบรมราชชนนีนครทุ่งเทพ ผู้อำนวยการวิทยาลัยพยาบาลบรมราชชนนีสระบุรี ผู้อำนวยการวิทยาลัยพยาบาลบรมราชชนนีนพรัตนมาท และผู้อำนวยการวิทยาลัยพยาบาลบรมราชชนนีสลบุรี ที่กรุณาให้ทำการเก็บข้อมูลจากนักศึกษาพยาบาล

นอกจากนี้ ผู้วิจัยขอขอบคุณทุก ๆ ท่านที่มีส่วนเกี่ยวข้องกับการศึกษาค้นคว้าครั้งนี้ ไม่ว่าจะได้โดยตรง หรือโดยอ้อมก็ตาม จนทำให้การศึกษามีบรรลุผลตามเป้าหมายได้

คุณค่า และประโยชน์ของปริญญานิพนธ์ฉบับนี้ ขอมอบเป็นเครื่องบูชาพระคุณบิดา มารดา ครู อาจารย์ และผู้มีพระคุณทุก ๆ ท่านของผู้วิจัย

สุมนา ใสตถิผลอนันต์

ปริญญานิพนธ์ฉบับนี้

ได้รับทุนอุดหนุนการศึกษาสำหรับอาจารย์พยาบาลจาก

"มูลนิธิหม่อมเจ้าหญิงบุญจิราธร (ทูลพล) จุฑาธุช"

ผู้วิจัยขอกราบขอบพระคุณมา ณ ที่นี้

สารบัญ

บทที่	หน้า
1 บทนำ	1
ภูมิหลัง	1
ความมุ่งหมายของการศึกษาค้นคว้า	5
ความสำคัญของการศึกษาค้นคว้า	5
ขอบเขตของการศึกษาค้นคว้า	6
นิยามศัพท์เฉพาะ	7
2 เอกสารและงานวิจัยที่เกี่ยวข้อง	9
ความหมายของการเทียบคะแนน และวิธีการเทียบคะแนนแบบเส้นตรง	9
เงื่อนไขของการเทียบคะแนน	11
วิธีการเทียบคะแนน	11
แผนแบบการเก็บรวบรวมข้อมูล	13
การพัฒนาวิธีการเทียบคะแนนแบบเส้นตรง กรณีที่ใช้แบบทดสอบหลัก	15
งานวิจัยที่เกี่ยวข้อง	34
สมมุติฐานของการศึกษาค้นคว้า	42
3 วิธีดำเนินการศึกษาค้นคว้า	43
ประชากร	43
กลุ่มตัวอย่าง	43
เครื่องมือที่ใช้ในการศึกษาค้นคว้า	45
วิธีดำเนินการสร้างเครื่องมือ	45
ลักษณะของแบบทดสอบที่ใช้ในการวิจัย	49
แผนแบบการเก็บรวบรวมข้อมูล	50
วิธีดำเนินการเก็บรวบรวมข้อมูล	52

ขั้นตอนในการวิเคราะห์ข้อมูล	53
สถิติที่ใช้ในการวิเคราะห์ข้อมูล	54
 4 ผลการวิเคราะห์ข้อมูล	 62
สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล	62
การนำเสนอผลการวิเคราะห์ข้อมูล	65
- ค่าสถิติพื้นฐานของคะแนนสอบ	65
- คะแนนแปลงเฉลี่ยของผลการสอบ	67
- ความคลาดเคลื่อนของการเทียบคะแนน	69
- ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน	71
- การเปรียบเทียบความแตกต่างของคะแนนแปลงที่ได้มาจากวิธีการคำนวณ ด้วยสูตร สี่วิธี	75
- การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานของ การเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี	87
 5 สรุป อภิปรายผล และข้อเสนอแนะ	 99
ความมุ่งหมายของการศึกษาค้นคว้า	99
สมมุติฐานของการศึกษาค้นคว้า	99
กลุ่มตัวอย่าง	100
เครื่องมือที่ใช้ในการศึกษาค้นคว้า	100
วิธีดำเนินการเก็บรวบรวมข้อมูล	100
สรุปผลการวิเคราะห์ข้อมูล	101
อภิปรายผล	102
ข้อเสนอแนะ	104

บทที่

หน้า

บรรณานุกรม 105

ภาคผนวก 111

ประวัติย่อของผู้วิจัย 139

บัญชีตาราง

ตาราง

หน้า

1	จำนวนนักศึกษาพยาบาลของแต่ละชั้นปีและวิทยาลัยพยาบาลที่เข้าเป็นกลุ่มสำหรับ เทียบคะแนน	44
2	การเก็บรวบรวมข้อมูลที่ใช้แบบทดสอบหลักกับประชากรแตกต่างตามแผนแบบ CINEP 51	
3	ค่าสถิติพื้นฐานของคะแนนจากแบบทดสอบเทียบคะแนน สองพอร์ม เมื่อใช้แบบทดสอบ หลัก ที่มีขนาดความยาวต่างกัน	62
4	คะแนนแปลงเฉลี่ยของแบบทดสอบสำหรับเทียบคะแนน ที่ได้จากวิธีการคำนวณด้วยสูตร สี่วิธี จำนวนตามการใช้แบบทดสอบหลักสองลักษณะ ที่มีขนาดความยาวต่างกัน ..	65
5	ความคลาดเคลื่อนของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี จำนวนตามการใช้แบบทดสอบหลัก สองลักษณะ ที่มีขนาดความยาวต่างกัน	68
6	ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี จำนวนตามการใช้แบบทดสอบหลัก สองลักษณะ ที่มีขนาดความยาวต่างกัน	70
7	ความถี่ของความคลาดเคลื่อนของการเทียบคะแนน แต่ละค่า เมื่อแบ่งตามระดับ ความพอใจ 5 ระดับ	73
8	การเปรียบเทียบคะแนนแปลงของการเทียบคะแนน ที่ได้จากวิธีการคำนวณด้วยสูตร สี่วิธี จำนวนตามการใช้แบบทดสอบหลัก ที่มีขนาดความยาวต่างกัน	75
9	การเปรียบเทียบคะแนนแปลงของการเทียบคะแนน โดยใช้แบบทดสอบเชื่อมโยงที่มี ขนาดความยาวต่างกัน ด้วยการวิเคราะห์ความแปรปรวนแบบทางเดียวแบบวัดซ้ำ	77
10	การเปรียบเทียบความแตกต่างรายคู่ของคะแนนแปลงเฉลี่ย ที่ได้วิธีจากการคำนวณ ด้วยสูตร สี่วิธี เมื่อใช้แบบทดสอบหลักแบบเชื่อมโยง ที่ขนาดความยาว 8 ข้อ .	79
11	การเปรียบเทียบความแตกต่างรายคู่ของคะแนนแปลงเฉลี่ย ที่ได้วิธีจากการคำนวณ ด้วยสูตร สี่วิธี เมื่อใช้แบบทดสอบหลักแบบเชื่อมโยง ที่ขนาดความยาว 12 ข้อ .	80
12	การเปรียบเทียบความแตกต่างรายคู่ของคะแนนแปลงเฉลี่ย ที่ได้จากวิธีการคำนวณ ด้วยสูตร สี่วิธี เมื่อใช้แบบทดสอบหลักแบบเชื่อมโยง ที่ขนาดความยาว 16 ข้อ .	81

13	การเปรียบเทียบคะแนนแปลงของการเทียบคะแนน โดยใช้อ้างอิงรวม ที่มีขนาด ความยาวต่างกัน ด้วยการวิเคราะห์ความแปรปรวนแบบทางเดียวแบบวัดซ้ำ ..	82
14	การเปรียบเทียบความแตกต่างรายคู่ของคะแนนแปลงเฉลี่ย ที่ได้จากวิธีการคำนวณ ด้วยสูตร สี่วิธี เมื่อใช้อ้างอิงรวม ที่ขนาดความยาว 8 ข้อ	84
15	การเปรียบเทียบความแตกต่างรายคู่ของคะแนนแปลงเฉลี่ย ที่ได้จากวิธีการคำนวณ ด้วยสูตร สี่วิธี เมื่อใช้อ้างอิงรวม ที่ขนาดความยาว 12 ข้อ	85
16	การเปรียบเทียบความแตกต่างรายคู่ของคะแนนแปลงเฉลี่ย ที่ได้จากวิธีการคำนวณ ด้วยสูตร สี่วิธี เมื่อใช้อ้างอิงรวม ที่ขนาดความยาว 16 ข้อ	86
17	การเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากวิธี การคำนวณด้วยสูตร สี่วิธี จำแนกตามการใช้แบบทดสอบหลักที่มีขนาดความยาวต่างกัน	87
18	การเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน โดยใช้แบบทดสอบ หลักแบบเชื่อมโยงที่มีขนาดความยาวต่างกัน ด้วยการวิเคราะห์ความแปรปรวน แบบทางเดียว แบบวัดซ้ำ	89
19	การเปรียบเทียบความแตกต่างรายคู่ของความคลาดเคลื่อนมาตรฐาน ที่เกิดจากวิธี การคำนวณด้วยสูตร สี่วิธี เมื่อใช้แบบทดสอบหลักแบบเชื่อมโยงที่ขนาดความยาว 8 ข้อ	91
20	การเปรียบเทียบความแตกต่างรายคู่ของความคลาดเคลื่อนมาตรฐาน ที่เกิดจากวิธี การคำนวณด้วยสูตร สี่วิธี เมื่อใช้แบบทดสอบหลักแบบเชื่อมโยงที่ขนาดความยาว 12	ข้อ 92
21	การเปรียบเทียบความแตกต่างรายคู่ของความคลาดเคลื่อนมาตรฐาน ที่เกิดจากวิธี การคำนวณด้วยสูตร สี่วิธี เมื่อใช้แบบทดสอบหลักแบบเชื่อมโยงที่ขนาดความยาว 16 ข้อ	93
22	การเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน โดยใช้อ้างอิงรวม ที่มีขนาดความยาวต่างกัน ด้วยการวิเคราะห์ความแปรปรวนแบบทางเดียว แบบวัดซ้ำ	94

23	การเปรียบเทียบความแตกต่างรายคู่ ของความคลาดเคลื่อนมาตรฐานของการ เทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี เมื่อใช้ข้อสอบรวม ที่ ขนาดความยาว 8 ข้อ	96
24	การเปรียบเทียบความแตกต่างรายคู่ ของความคลาดเคลื่อนมาตรฐานของการ เทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี เมื่อใช้ข้อสอบรวม ที่ ขนาดความยาว 12 ข้อ	97
25	การเปรียบเทียบความแตกต่างรายคู่ ของความคลาดเคลื่อนมาตรฐานของการ เทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี เมื่อใช้ข้อสอบรวม ที่ ขนาดความยาว 16 ข้อ	98
26	การเทียบคะแนนของคะแนนแปลงที่ได้จากการคำนวณด้วยสูตร สี่วิธี กับคะแนน x	133

บทที่ 1

บทนำ

ภูมิหลัง

การทดสอบที่ใช้นวทางการศึกษาต่าง ๆ ไม่ว่าจะเป็นการสอบเพื่อคัดเลือกเข้าศึกษาต่อ การสอบเพื่อรับรองความสำเร็จของการศึกษา หรือแม้แต่การสอบคัดเลือกเข้าทำงาน การฝึกอบรมในโครงการต่าง ๆ มีเป้าหมายสูงสุดเพื่อทำให้นักเรียนที่มีคุณสมบัติ และความสามารถตามความต้องการของสถาบันการศึกษา หรือหน่วยงานนั้น ๆ ซึ่งในการสอบแต่ละครั้งนั้น อาจมีการใช้แบบทดสอบในลักษณะที่แตกต่างกันไป ทั้งนี้ ก็เพื่อป้องกันการรั่วไหลของข้อสอบ การสอบในกรณีนี้ จึงทำให้เกิดปัญหาว่า คะแนนที่สอบได้จากการใช้แบบทดสอบที่ต่างฉบับกันนั้น สามารถที่จะนำมาเปรียบเทียบกันได้หรือไม่ ซึ่งที่ผ่านมาแล้ว มีนักวัดผลทางการศึกษาได้พยายามแก้ปัญหาโดยใช้วิธีการแปลงคะแนนที่สอบได้ให้เป็นคะแนนมาตรฐาน จากวิธีการแก้ปัญหานี้ ทำให้มีการสร้างแบบทดสอบมาตรฐานขึ้นมา โดยนำคะแนนที่ได้จากการสอบจากกลุ่มคนที่มีจำนวนมาก หรือกลุ่มมาตรฐาน (Norm Groups) เหล่านั้นมาทำการเปรียบเทียบกัน และวิธีนี้เป็นวิธีที่นิยมใช้ในวงการการศึกษาต่าง ๆ เพื่อนำมาใช้เปรียบเทียบวัดความสามารถของผู้เข้าสอบ ในกลุ่มต่าง ๆ แองกอฟ (Angoff, 1984 : 561) เป็นผู้หนึ่งที่ศึกษาวิธีการดังกล่าว ได้กล่าวถึงปัญหาที่เกิดจากการสอบที่มีการใช้กลุ่มคนจำนวนมากนี้ ว่าเป็นการสอบที่มาจากกลุ่มคน ที่มีความแตกต่างกัน แล้วนำคะแนนมาเปรียบเทียบกัน โดยให้เหตุผลว่า คนที่มาจากต่างสถาบันการศึกษากัน จะมีความสามารถที่แตกต่างกัน การจัดดำเนินการสอบในกลุ่มบุคคลต่าง ๆ แต่ละแห่ง แล้วนำผลการสอบมาสร้างเป็นเกณฑ์ปกติขึ้น พบว่า ความแตกต่างกันที่เกิดขึ้น มีตั้งแต่เรื่องของการใช้แบบทดสอบที่ใช้อธิบายสภาพกลุ่มต่าง ๆ ปัญหาเฉพาะอย่างที่เกิดขึ้นแต่ละแห่ง และเหตุผลของความคลาดเคลื่อนที่เกิดจากกลุ่มต่าง ๆ ซึ่งจุดอ่อนที่ควรให้ความสนใจ ก็คือ ถ้าเปรียบเทียบกลุ่มมาตรฐาน สองกลุ่ม กลุ่มหนึ่งมาจากประชากรที่มีความสามารถหลากหลาย แต่อีกกลุ่มหนึ่งมาจากประชากรที่มีความสามารถค่อนข้างสูงที่ใกล้เคียงกัน กรณีเช่นนี้ ค่าเฉลี่ย และความเบี่ยงเบนมาตรฐานของคะแนน สองกลุ่ม จะแตกต่างกัน (Lindquist, 1951: 746) ดังนั้น

แองกอฟ จึงได้เสนอวิธีการเทียบคะแนน (Equating) ขึ้นมาชี้แทนการเปรียบเทียบคะแนน ที่มาจากการใช้กลุ่มมาตรฐาน ในขณะที่เดียวกัน ฟลานาแกน (Flanagan.1951: 751-753) ก็ได้เสนอเกี่ยวกับการใช้คะแนนมาตรฐาน ในการเปรียบเทียบความสามารถของบุคคล ที่อาศัยวิธีการเทียบคะแนนนี้ไว้ว่า ควรเริ่มตั้งแต่การสร้างแบบทดสอบแต่ละฉบับ ให้มีความคู่ขนานกัน (Parallel Forms) ก่อน แล้วจึงเลือกใช้วิธีการเทียบคะแนนทางสถิติที่เหมาะสม โดยเสนอวิธีการไว้ สี่วิธี คือ 1.) วิธีการเทียบค่าเฉลี่ย วิธีนี้ต้องคำนวณหาค่าเฉลี่ยของคะแนนจากแบบทดสอบทั้งสองฉบับ หากค่าเฉลี่ยที่ได้ไม่แตกต่างกันมากนัก ก็สามารถนำมาเปรียบเทียบกันได้ โดยดูจากรูปทรงของการกระจายของคะแนน 2.) วิธีการคำนวณจากสมการถดถอย โดยใช้วิธีการกะประมาณที่ดีที่สุด (Best Estimate) ของคะแนนจากแบบทดสอบฉบับหนึ่ง จากคะแนนของแบบทดสอบอีกฉบับหนึ่ง 3.) วิธีการคำนวณคะแนนมาตรฐาน ซึ่งเป็นวิธีการปรับค่าเฉลี่ย ที่แตกต่างจากแบบทดสอบ สองฉบับ ให้มาอยู่ในคะแนนที่เป็นมาตราเดียวกัน 4.) วิธีการเทียบคะแนนที่ใช้ตำแหน่งที่เท่ากัน (Equipercentile Method) ซึ่งเป็นวิธีการหาค่าของคะแนน โดยวิธีการเปรียบเทียบตามสัดส่วนของการแจกแจงคะแนน

ต่อมา เลนนอน (Lennon.1964 , 1969) ได้เสนอวิธีเทียบคะแนน โดยการใช้แบบทดสอบหลัก (Anchor Test or Item Common Test) ซึ่งเป็นวิธีการเทียบคะแนนโดยการนำข้อสอบจากแบบทดสอบที่ใช้เทียบคะแนนทั้งสองฉบับ มาอย่างละครึ่งฉบับ แล้วรวมกันเป็นแบบทดสอบอีกหนึ่งฉบับ เพื่อใช้เป็นเกณฑ์มาตรฐานเดียวกัน คะแนนที่วัดได้จากแบบทดสอบฉบับใหม่นี้ จะมีลักษณะเป็นคะแนนสมมูล (Equivalent Score) กับแบบทดสอบสองฉบับแรก ซึ่งในกรณีนี้ จะสามารถนำคะแนนมาเปรียบเทียบกันได้โดยตรง เนื่องจากได้มีการปรับความแตกต่างระหว่างแบบทดสอบทั้งสองฉบับแล้ว และต้องนำไปสอบกับกลุ่มผู้สอบกลุ่มต่าง ๆ เพื่อให้เป็นกลุ่มมาตรฐานสำหรับใช้เทียบคะแนน (Standard Norm Groups) กลุ่มผู้สอบดังกล่าวก็จะถือว่าเป็นกลุ่มประชากรที่คล้ายคลึงกัน เช่นเดียวกับที่ แองกอฟ (Angoff.1982 , 1984) ได้เสนอการใช้แบบทดสอบหลัก ภายใต้เงื่อนไขของการใช้แผนแบบที่ III และแผนแบบที่ IV ของการเก็บรวบรวมข้อมูล แต่การใช้แบบทดสอบหลักนี้ ก็ยังเป็นแบบทดสอบที่คงความเป็นคู่ขนานไว้ กับแบบทดสอบที่ต้องการใช้เทียบคะแนนกันทั้งสองฉบับ ซึ่งการสร้างแบบทดสอบที่เป็นคู่ขนานกันนั้น ในทางปฏิบัติ มักจะมีปัญหาการนำไปใช้ เพราะเมื่อระยะเวลาผ่านไป ก็

จะมีการเปลี่ยนแปลงของโครงสร้างหลักสูตรและเนื้อหาวิชา การที่จะเทียบคะแนนจากการนำคะแนนจากแบบทดสอบฉบับใหม่ (New Form) ไปเทียบคะแนนยังแบบทดสอบฉบับเก่า (Old Form) นั้น ก็จะมีปัญหามาก ทั้งนี้ เนื่องจากข้อจำกัดของความเป็นคู่ขนานของแบบทดสอบหลักจะขึ้นอยู่กับเนื้อหาวิชา (Test Content) ที่ใช้ในการสอบ ระดับชั้นเรียน (Grade Levels) และยังไม่มีความที่เป็นมาตรฐานสำหรับเทียบคะแนนในแนวตั้ง (Absence of Vertically Equated Scaled Scores) (Bianchini and Loret.1975 : 212)

จากปัญหาการใช้แบบทดสอบหลักที่ผ่านมา โดยมีข้อจำกัดของความเป็นคู่ขนานกับแบบทดสอบเทียบคะแนน ทำให้ทักเกอร์ (Tucker) (Angoff.1984 : 580) ได้พัฒนาการใช้แบบทดสอบหลัก ที่อาศัยข้อตกลงของการเทียบคะแนนจากการใช้เส้นถดถอย และใช้วิธีการจัดชุดข้อสอบใหม่ของแบบทดสอบหลัก ที่เรียกว่า "แผนแบบของการใช้แบบทดสอบหลักที่มีประชากรแตกต่างกัน" (Common Item Non-equivalent Population Design or CINEP) ซึ่งแบบทดสอบหลักในแผนแบบนี้ไม่จำเป็นต้องเป็นคู่ขนานกัน (Non-parallel Form) แต่ต้องมีโครงสร้างทางทฤษฎี และเนื้อหาเดียวกัน รวมทั้งประชากรที่เป็นกลุ่มผู้เข้าสอบ ก็ต้องเป็นกลุ่มที่มีความสามารถไม่แตกต่างกันมาก

ดังนั้นในการเทียบคะแนน จึงจำเป็นต้องอาศัยสิ่งสำคัญสองประการ คือ ประการแรก ต้องออกแบบการเก็บรวบรวมข้อมูลให้ชัดเจน ซึ่งมีอยู่หลายวิธีด้วยกัน และวิธีที่นิยมใช้กันมาก คือ การใช้แผนแบบของการใช้แบบทดสอบหลักที่มีประชากรแตกต่างกัน (CINEP) ซึ่งมีลักษณะของแผนแบบการเก็บรวบรวมข้อมูลที่ว่า เป็นการนำแบบทดสอบฟอร์ม X มาสอบกับผู้เข้าสอบกลุ่มที่หนึ่ง และแบบทดสอบฟอร์ม Y สอบกับผู้เข้าสอบกลุ่มที่สอง แล้วนำแบบทดสอบหลักสอบกับกลุ่มผู้เข้าสอบทั้งสองกลุ่ม ทั้งนี้ เพื่อใช้เป็นตัวร่วมในการแปลงคะแนนจากแบบทดสอบฟอร์ม X ไปยังคะแนนจากแบบทดสอบฟอร์ม Y แบบทดสอบหลักที่ใช้ในการเก็บรวบรวมข้อมูลด้วยแผนแบบนี้ อาจนำไปรวมอยู่ในแบบทดสอบทั้งสองฟอร์ม หรือแยกเป็นแบบทดสอบต่างหากอีกฉบับหนึ่ง ซึ่งการนำข้อสอบจากแบบทดสอบหลัก ไปรวมไว้กับแบบทดสอบที่จะนำมาเทียบคะแนนกันนั้น เรียกวิธีการเช่นนี้ว่า "การเทียบคะแนนที่ใช้ข้อสอบร่วม" (Internal Anchor Test) ส่วนการใช้แบบทดสอบหลักที่แยกเป็นฉบับต่างหาก จากแบบทดสอบเทียบคะแนนนั้น ก็ จะเรียกว่า "การเทียบคะแนนที่ใช้แบบทดสอบเชื่อมโยง" (External Anchor Test)

ประการที่สอง จะต้องมีการกำหนดวิธีการทางสถิติให้ชัดเจน ซึ่งวิธีการทางสถิติที่นำมาใช้ในการเทียบคะแนนนั้น มีอยู่หลายวิธี ได้แก่ การเทียบคะแนนแบบเส้นตรง (Linear Equating) การเทียบคะแนนแบบใช้ตำแหน่งที่เท่ากัน (Equipercentile Equating) และการเทียบคะแนนตามทฤษฎีการตอบข้อสอบ (IRT Equating)

การเทียบคะแนนแบบเส้นตรง นับว่าเป็นวิธีการทางสถิติที่คนส่วนใหญ่นิยมใช้ และได้รับความนิยมกันมาก ซึ่งวิธีการนี้จะอาศัยความสัมพันธ์ของคะแนนสอบ จากแบบทดสอบฉบับเดิม กับแบบทดสอบฉบับใหม่ที่วัดในสิ่งเดียวกันมาใช้ในการคำนวณ โดยวิธีการเทียบคะแนนแบบเส้นตรงนี้ จำเป็นต้องอาศัยข้อตกลง สองข้อ คือ 1.) ต้องมีความลาดชันของเส้นถดถอยเท่ากัน และ 2.) เส้นถดถอยนั้น ต้องตัดแกนตั้งที่จุดเดียวกัน

วิธีการเทียบคะแนนแบบเส้นตรงนี้ ได้มีผู้เสนอสูตร หรือวิธีการเทียบคะแนนไว้หลายวิธี แต่ละวิธีจะมีข้อตกลงที่แตกต่างกัน วิธีการเทียบคะแนนที่ใช้ข้อตกลงของแบบจำลองคะแนนจริงสัมพันธ์ (Congeneric Model) เป็นวิธีการที่มีความเป็นไปได้ในทางปฏิบัติมากกว่าข้อตกลงอื่น ๆ ซึ่งนักวัดผลที่สนใจในการเทียบคะแนนที่ใช้ข้อตกลงดังกล่าวนี้ คือ วัตดรุฟ (Woodruff, 1986) ได้เสนอวิธีการเทียบคะแนนแบบเส้นตรง ที่ใช้กับแผนแบบของการใช้แบบทดสอบหลักที่มีประชากรแตกต่างกัน (CINEP) โดยเสนอไว้ สองสูตร ที่อาศัยข้อตกลงของแบบจำลองคะแนนจริงสัมพันธ์ ต่อมา แมคแคน (MacCann, 1989) ก็ได้เสนอสูตร หรือวิธีการเทียบคะแนนแบบเส้นตรงที่ใช้แบบทดสอบหลักที่เป็นแบบทดสอบเชื่อมโยง (External Anchor Test) ซึ่งอาศัยข้อตกลงของแบบทดสอบที่นำมาเทียบคะแนนกัน สองฉบับว่า ต้องมีความเชื่อมั่นเท่ากัน และเป็นคะแนนจริงสัมพันธ์

โดยทั่วไปการสอบที่ใช้ขนาดความยาวของแบบทดสอบต่างกัน จะส่งผลต่อคะแนนสอบได้ ดังนั้นการใช้แบบทดสอบหลัก มาเป็นแบบทดสอบที่นำมาใช้ประกอบการเทียบคะแนน ของแต่ละวิธีดังกล่าวไปแล้ว ขนาดความยาวของแบบทดสอบหลักแต่ละขนาด น่าจะส่งผลต่อวิธีการเทียบคะแนนในวิธีทั้งสี่วิธี คือ วิธีของแมคแคน สองวิธี และวิธีของวัตดรุฟ สองวิธี นอกจากนี้ การใช้แบบทดสอบหลักเป็นแบบทดสอบเชื่อมโยง กับเป็นข้อสอบร่วมกัน จะให้ผลการเทียบคะแนนที่แตกต่างกันหรือไม่ และมีวิธีการใดให้ผลดีที่สุด ซึ่งผู้วิจัยได้วางแผนการศึกษาไว้ โดยการใช้แบบทดสอบวัดการคิดเอกลัษณ์แบบสัมพันธ์ ซึ่งวัดความสามารถทางสมอง ตามโครงสร้างทฤษฎี

ของกิลฟอร์ด ที่ผู้วิจัยสร้างขึ้นเอง

ความมุ่งหมายของการศึกษา

ในการวิจัยครั้งนี้ ผู้วิจัยขอเสนอจุดมุ่งหมายของการศึกษากันไว้ดังนี้

1. เพื่อเปรียบเทียบคะแนนแปลง อันเป็นผลจากวิธีการเทียบคะแนนแบบเส้นตรง สี่วิธี ว่าแตกต่างกันหรือไม่
2. เพื่อเปรียบเทียบคะแนนแปลง อันเป็นผลจากวิธีการเทียบคะแนนแบบเส้นตรง ที่ใช้แบบทดสอบเชื่อมเรียงกับการใช้ข้อสอบรวม ว่าแตกต่างกันหรือไม่
3. เพื่อเปรียบเทียบคะแนนแปลง อันเป็นผลจากวิธีการเทียบคะแนนแบบเส้นตรง ที่มีขนาดความยาวของแบบทดสอบหลักแบบเชื่อมเรียง กับแบบข้อสอบรวม ในสัดส่วน .2 , .3 และ
4. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน สี่วิธี ว่าแตกต่างกันหรือไม่

ความสำคัญของการศึกษาค้นคว้า

การศึกษาค้นคว้าครั้งนี้ ทำให้ได้ข้อเท็จจริงเชิงประจักษ์ที่ว่า การเทียบคะแนนแบบเส้นตรงที่อาศัย ข้อตกลงของคะแนนจริงสัมพัทธ์ เป็นวิธีการเทียบคะแนนที่สามารถนำไปใช้ในทางปฏิบัติได้จริงมากที่สุดเพราะมีความผ่อนปรนทางปฏิบัติ สามารถเลือกวิธีการแปลงคะแนนได้จากการพิจารณา ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนที่ให้ผลต่ำที่สุด และเพื่อหาได้รูปแบบของแบบทดสอบหลัก ที่เหมาะสมสำหรับการนำไปใช้ สามารถนำไปทดสอบกับกลุ่มผู้เข้าสอบที่เป็นนักศึกษาที่มาจากต่างสถาบันการศึกษา ซึ่งจะเป็นประโยชน์ต่อการสอบคัดเลือกนักศึกษาเข้าศึกษาต่อ หรือใช้การกำหนดคุณภาพของ ผู้สำเร็จการศึกษาจากหลักสูตรต่าง ๆ ให้ความเท่าเทียมกัน ทำให้เกิดมีความยุติธรรมต่อผู้เข้าสอบในแต่ละปี แต่ละสถาบันการศึกษา

ขอบเขตของการศึกษาค้นคว้า

1. ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักศึกษาพยาบาล ชั้นปีที่ 1 และชั้นปีที่ 2 ของวิทยาลัยพยาบาล สังกัดสถาบันพัฒนากำลังคนด้านสาธารณสุข กระทรวงสาธารณสุข ในเขตภาคกลาง ปีการศึกษา 2538 จำนวน 11 แห่ง

2. กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักศึกษาพยาบาลชั้นปีที่ 1 และชั้นปีที่ 2 ของวิทยาลัยพยาบาล สังกัดสถาบันพัฒนากำลังคนด้านสาธารณสุข กระทรวงสาธารณสุข ในเขตภาคกลาง ปีการศึกษา 2538 ซึ่งได้มาจากการสุ่มอย่างง่าย 2 ขั้นตอน จำนวน 3 แห่ง มีจำนวนนักศึกษาทั้งสิ้น 864 คน

3. ตัวแปรที่ใช้ในการศึกษาค้นคว้าครั้งนี้

3.1. ตัวแปรอิสระ ได้แก่

3.1.1. วิธีการเทียบคะแนนแบบเส้นตรง ซึ่งแบ่งเป็น สี่วิธี คือ

- วิธีที่ 1 ของแมคเคน
- วิธีที่ 2 ของแมคเคน
- วิธีที่ 1 ของวีตดรัฟ
- วิธีที่ 2 ของวีตดรัฟ

3.1.2. สถานการณ์การให้แบบทดสอบหลัก สองลักษณะ คือ

- การให้แบบทดสอบเชื่อมโยง
- การให้ข้อสอบรวม

3.1.3. ขนาดความยาวของแบบทดสอบหลัก สามขนาด คือ

- จำนวน 8 ข้อ
- จำนวน 12 ข้อ
- จำนวน 16 ข้อ

3.2. ตัวแปรตาม ได้แก่

3.2.1. คะแนนแปลงของการเทียบคะแนนแบบเส้นตรง

3.2.2. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน

นิยามศัพท์เฉพาะ

1. การเทียบคะแนน(Equating) หมายถึง การนำคะแนนจากแบบทดสอบสองฉบับที่ข้อสอบกับกลุ่มตัวอย่างคนละกลุ่มมาเปรียบเทียบกัน โดยใช้วิธีการแปลงคะแนนจากแบบทดสอบฉบับหนึ่งไปเป็นคะแนนของแบบทดสอบอีกฉบับหนึ่ง เพื่อให้เป็นคะแนนมาตราเดียวกัน

2. แบบทดสอบสำหรับเทียบคะแนน หมายถึง แบบทดสอบสองฉบับ ที่วัดโครงสร้างทางความคิดและเนื้อหาเดียวกัน แต่แตกต่างกันตรงสัดส่วนของเนื้อหากับจำนวนข้อ และสอบกับกลุ่มตัวอย่างคนละกลุ่ม

ในการวิจัยครั้งนี้ ใช้แบบทดสอบที่วัดความสามารถทางสมองตามโครงสร้างทางทฤษฎีของกิลฟอร์ด ด้านความคิดเอกนัยแบบสัมพันธ์ (Convergent Relation) ที่มีเนื้อหาเป็นภาพกับภาษา ได้แก่ พอร์ม A มีจำนวน 40 ข้อ มีเนื้อหาเป็นภาพ 24 ข้อ เป็นภาษา 16 ข้อ คิดเป็นสัดส่วน 3:2 และพอร์ม B มีจำนวน 50 ข้อ มีเนื้อหาเป็นภาพ 20 ข้อ เป็นภาษา 30 ข้อ คิดเป็นสัดส่วน 2:3

3. แบบทดสอบหลัก (Anchor Test) หมายถึง แบบทดสอบที่จัดทำขึ้นเป็นหลัก (พอร์ม C) โดยแยกต่างหากจากแบบทดสอบพอร์ม A กับพอร์ม B ที่ต้องการนำมาเทียบคะแนนกัน ในการวิจัยครั้งนี้ แบบทดสอบหลักเป็นแบบทดสอบที่สร้างตามโครงสร้างทางทฤษฎี เช่นเดียวกับพอร์ม A และพอร์ม B แต่มีจำนวนข้อ 16 ข้อ มีเนื้อหาเป็นภาพ 8 ข้อ เป็นภาษา 8 ข้อ คิดเป็นสัดส่วน 1:1 แล้วจัดเป็นฉบับย่อย 3 ฉบับ ดังนี้

ฉบับที่ 1 (พอร์ม C₁) มีจำนวน 16 ข้อ (คิดเป็นสัดส่วน .4 ของความยาวฉบับพอร์ม A) มีเนื้อหาที่เป็นภาพ กับภาษา คิดเป็นสัดส่วน 1:1

ฉบับที่ 2 (พอร์ม C₂) มีจำนวน 12 ข้อ (คิดเป็นสัดส่วน .3 ของความยาวฉบับพอร์ม A) มีเนื้อหาที่เป็นภาพ กับภาษา คิดเป็นสัดส่วน 1:1

ฉบับที่ 3 (พอร์ม C₃) มีจำนวน 8 ข้อ (คิดเป็นสัดส่วน .2 ของความยาวฉบับพอร์ม A) มีเนื้อหาที่เป็นภาพ กับภาษา คิดเป็นสัดส่วน 1:1

4. สถานการณ์การใช้แบบทดสอบหลัก หมายถึง การนำแบบทดสอบหลักไปใช้ในสองลักษณะ เพื่อการเทียบคะแนน ได้แก่

4.1. การใช้แบบทดสอบเชื่อมโยง หมายถึง การนำแบบทดสอบหลักไปใช้ควบคู่กับแบบทดสอบสำหรับเทียบคะแนน โดยพิมพ์แยกต่างหากเป็นคนละฉบับกัน ในการสอบผู้สอบกลุ่มหนึ่งสอบพอร์ม A กับพอร์ม C อีกกลุ่มหนึ่งสอบพอร์ม B กับพอร์ม C

4.2. การใช้แบบข้อสอบรวม หมายถึง การนำข้อสอบจากแบบทดสอบหลักไปรวมกับแบบทดสอบสำหรับเทียบคะแนน โดยใช้วิธีสุ่มลำดับที่ของข้อสอบจากแบบทดสอบหลัก แล้วพิมพ์แทรกเข้าไปให้อยู่ในฉบับเดียวกัน กับแบบทดสอบสำหรับเทียบคะแนน ในการสอบผู้สอบกลุ่มหนึ่งสอบพอร์ม A+C อีกกลุ่มหนึ่งสอบพอร์ม B+C

5. แบบทดสอบวัดการคิดเอกลัแบบสัมพันธ์ หมายถึง แบบทดสอบวัดความสามารถทางสมอง ตามโครงสร้างทางทฤษฎีของกิลฟอร์ด ด้านการคิดเอกลัที่มีเนื้อหาเป็นภาพและภาษาแบบสัมพันธ์ ในสัดส่วนที่แตกต่างกัน 3 ฉบับ ได้แก่ พอร์ม A, พอร์ม B และพอร์ม C

6. แบบจำลองคะแนนจริงสัมพันธ์ หมายถึง โครงสร้างทางคณิตศาสตร์ของคะแนนที่วัดด้วยแบบทดสอบแต่ละฉบับ ซึ่งคะแนนจริงของแต่ละฉบับ จะมีกราฟความสัมพันธ์เป็นเส้นตรง หรือมีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับหนึ่ง

7. คะแนนแปลงของการเทียบคะแนน หมายถึง คะแนนที่ได้มาจากการนำคะแนนจากการสอบแบบทดสอบสำหรับเทียบคะแนน พอร์ม B ไปจัดกระทำโดยวิธีการคำนวณจากสูตรการเทียบคะแนนแบบเส้นตรง สี่วิธี คือ 1.) วิธีที่ 1 ของแมคแคน 2.) วิธีที่ 2 ของแมคแคน 3.) วิธีที่ 1 ของวูดเวิร์ฟ 4.) วิธีที่ 2 ของวูดเวิร์ฟ ซึ่งคะแนนที่ได้จากการคำนวณนี้ จะเป็นคะแนนแปลงที่สามารถนำไปเทียบกับคะแนนจากการสอบแบบทดสอบพอร์ม A ได้

8. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน หมายถึง ความคลาดเคลื่อนของคะแนนที่เกิดจากการนำคะแนนสอบจากแบบทดสอบพอร์ม B ไปแปลงคะแนน เพื่อนำมาเทียบกับคะแนนจากการสอบแบบทดสอบพอร์ม A คะแนนที่ได้จากการแปลงคะแนนนั้น มีความคลาดเคลื่อนที่คิดเป็นรากที่สองของค่าเฉลี่ยของความคลาดเคลื่อน (Root Mean Squared Error)

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยขอเสนอรายละเอียดของเอกสาร และงานวิจัยที่เกี่ยวข้องดังนี้

1. ความหมายของการเทียบคะแนน และการเทียบคะแนนแบบเส้นตรง
2. เงื่อนไขของการเทียบคะแนน
3. วิธีการเทียบคะแนน
4. แผนแบบการเก็บรวบรวมข้อมูล
5. การพัฒนาวิธีการเทียบคะแนนแบบเส้นตรง กรณีที่ใช้แบบทดสอบหลัก
6. งานวิจัยที่เกี่ยวข้อง

ความหมายของการเทียบคะแนน และวิธีการเทียบคะแนนแบบเส้นตรง

นักการศึกษาหลายท่านได้ให้ ความหมายของการเทียบคะแนน และวิธีการเทียบคะแนนแบบเส้นตรงไว้ ดังนี้

กัลลิกเสน (Gulliksen.1950 : 298) ให้ความหมายของการเทียบคะแนนว่า เป็นการนำคะแนนจากแบบทดสอบ สองฉบับในวิชาเดียวกัน มาทดสอบกับกลุ่มตัวอย่างที่คล้ายคลึงกัน แล้วนำมาเปรียบเทียบกัน โดยมีข้อแม้ว่าต้องแปลงคะแนนให้เป็นมาตรฐานเดียวกัน มีการกระจายของคะแนนเป็นโค้งปกติ หรือทำเป็นคะแนนเปอร์เซ็นต์ไทล์ แล้วจึงนำมาเปรียบเทียบกันได้

ฟลานาแกน (Flanagan.1951 : 745-747) ให้ความหมายของการเทียบคะแนนว่า เป็นกระบวนการเปรียบเทียบคะแนนจากแบบทดสอบ สองฉบับในวิชาเดียวกัน และทดสอบกับกลุ่มประชากรเดียวกัน คะแนนที่ได้มาจะมีการกระจายของคะแนนจริงที่เหมือนกัน

บราน และฮอลแลนด์ (Braun and Holland.1982 : 15) ได้ให้ความหมายของการเทียบคะแนนว่า เป็นการนำคะแนนจากแบบทดสอบ สองฉบับมาเทียบคะแนนกัน ซึ่งคะแนนที่ได้มาจากประชากรเดียวกัน และมีการกระจายของคะแนนที่เท่า ๆ กัน

สงข ลักษณ์ (2525 : 22) ได้กล่าวถึงการเทียบคะแนนไว้ว่า การเทียบคะแนน เป็นการนำคะแนนจากแบบทดสอบ ต่างฉบับกันที่วัดในสิ่งเดียวกัน แต่ไม่จำเป็นต้องเป็นข้อสอบ คู่ขนานกัน จะถือว่าเท่าเทียมกันได้ ถ้านักเรียนคนใดก็ตามทำข้อสอบต่างฉบับแล้ว ได้คะแนนจริง หรือความสามารถที่แท้จริงเท่ากัน

ฐศักดิ์ ชัมภลิจิต (สุรัตน์ ขวัญบุญจันทร์.2531 : 10 ; อ้างอิงมาจาก ฐศักดิ์ ชัมภลิจิต. 2527) ให้ความหมายของการเทียบคะแนนไว้ว่า เป็นกระบวนการที่เกี่ยวข้องกับกิจกรรม สอง ประการ คือ กระบวนการที่ทำให้แบบทดสอบ สองฉบับมีความเท่าเทียมกันในเรื่องโครงสร้าง และ การใช้วิธีการทางสถิติ เพื่อปรับคะแนนที่ได้จากการสอบแต่ละฉบับให้อยู่ในมาตราเดียวกัน แล้ว จึงเปรียบเทียบกันได้

วิรัช วรรณรัตน์ (2530 : 69) ได้ให้ความหมายของการเทียบคะแนนว่า เป็นการนำ เอาคะแนนที่สังเกตได้จากแบบทดสอบ ต่างฉบับ ซึ่งแต่ละฉบับต่างก็วัดในสิ่งเดียวกันมาเทียบกัน หรือกล่าวได้ว่าการเทียบคะแนน เป็นวิธีการเปลี่ยนระบบคะแนนของแบบทดสอบฉบับหนึ่ง ไปสู่ ระบบคะแนนของแบบทดสอบอีกฉบับหนึ่ง

จากความหมายที่กล่าวมานั้น พอสรุปได้ว่า การเทียบคะแนนเป็นกระบวนการที่จัดกระทำ กับข้อมูลที่เป็นคะแนนจากแบบทดสอบ ต่างฉบับที่วัดในเนื้อหาเดียวกัน สามารถนำมาเปรียบเทียบ กันได้ภายใต้คะแนนมาตราเดียวกัน

สำหรับการเทียบคะแนนแบบเส้นตรงนั้น มีข้อตกลงหรือเงื่อนไขเบื้องต้นที่ว่า เป็นการนำ คะแนนจากแบบทดสอบ สองฉบับที่ต่างกันมาเทียบคะแนนกัน โดยแบบทดสอบทั้งสองฉบับนั้น จะต้อง มีความเชื่อมั่นเท่า ๆ กัน และมีลักษณะเป็นคู่ขนานกัน เป็นการทำให้คะแนนทั้งสองมีความสมมูลกัน ความเบี่ยงเบนมาตรฐานของคะแนนแต่ละฉบับมีค่าเท่าเทียมกัน และสามารถเขียนเป็นสมการ เส้นตรงแสดงความสัมพันธ์ของคะแนนทั้งสอง (Angoff.1982 : 56-57 ; 1984 : 564)

วูดดรัฟ (Woodruff.1986 : 246) ให้ความหมายของการเทียบคะแนนแบบเส้นตรง ไว้ว่า เป็นการนำคะแนนที่ได้จากแบบทดสอบต่างฉบับกัน มาแปลงคะแนนให้เป็นไปตามลักษณะแบบ เส้นตรง (Linear Transformation) ดังนั้นคะแนนที่ได้จากการแปลงคะแนนแล้ว จะมีลักษณะ ของค่าเฉลี่ย และความแปรปรวน เช่นเดียวกับกับคะแนนจากแบบทดสอบที่นำมาใช้เทียบคะแนน โดยแบบทดสอบทั้งสองฉบับ ต้องนำมาเปรียบเทียบกับกลุ่มตัวอย่างที่มีลักษณะคล้ายคลึงกัน

แสดงว่าการเทียบคะแนนแบบเส้นตรง เป็นการแสดงความสัมพันธ์ ระหว่างคะแนนจากแบบทดสอบสองฉบับที่ต่างกัน โดยอาศัยวิธีการทางคณิตศาสตร์ และสามารถเขียนเป็นสมการได้ ภายได้เงื่อนไขว่า คะแนนนั้นต้องเป็นคะแนนสมมูลกัน

เงื่อนไขของการเทียบคะแนน

การที่จะนำคะแนนดิบจากแบบทดสอบ สองฉบับที่ต่างกันมาเทียบคะแนนกัน โดยการนำมาแปลงคะแนนจากแบบทดสอบฉบับหนึ่ง ไปสู่คะแนนมาตราของแบบทดสอบอีกฉบับหนึ่ง เพื่อทำให้เกิดความสมมูลกันนั้น ลอร์ด (Petersen and others.1989 : 242 ; citing Lord.1980) ได้กล่าวถึงเงื่อนไข 4 ข้อ ที่จะต้องนำมาใช้ในการเทียบคะแนน ดังนี้

1. แบบทดสอบทั้งสองฉบับ ต้องวัดคุณลักษณะ ความสามารถหรือทักษะของผู้เข้าสอบ ที่เท่าเทียมกัน
2. กลุ่มผู้เข้าสอบ จะต้องมีความสามารถเท่าเทียมกัน โดยดูจากผลการแจกแจงของคะแนนที่ได้หลังจากมีการแปลงคะแนนแล้ว ว่ามีความเท่าเทียมกับคะแนนที่ได้จากแบบทดสอบที่ใช้เป็นฉบับเทียบคะแนน
3. ความไม่แปรผันตามกลุ่มผู้สอบ คะแนนแปลงที่ได้จากการแปลงคะแนน ไม่ว่าจะมาจากกลุ่มใด ๆ ก็ตาม จะมีค่าเท่าเทียมกัน หรือมีค่าไม่เปลี่ยนแปลงไปตามกลุ่มผู้เข้าสอบ
4. ความเท่าเทียมกันระหว่างคะแนนแปลง กล่าวคือ คะแนนที่ได้จากการเทียบคะแนนจากแบบทดสอบฟอร์ม X ไปยังฟอร์ม Y หรือจากแบบทดสอบฟอร์ม Y ไปยังฟอร์ม X ก็ให้ผลการเทียบคะแนนที่เท่าเทียมกัน

วิธีการเทียบคะแนน (Equating Methods)

วิธีการที่ใช้ในการเทียบคะแนน มีอยู่หลายวิธีด้วยกัน โดยที่วิธีการเทียบคะแนนที่ใช้เป็นพื้นฐานนั้น มีอยู่ 3 วิธี ได้แก่

1. วิธีเทียบคะแนนแบบเส้นตรง (Linear Equating)

วิธีนี้ใช้ข้อตกลงว่า คะแนนจากแบบทดสอบต่างฉบับกันนั้นจะต้องอยู่ในคะแนนมาตรฐานเดียวกัน (Scale Score) โดยการแปลงคะแนนดิบจากแบบทดสอบทั้งสองฉบับ ให้เป็นคะแนนมาตรฐาน ซึ่งคะแนนที่ได้มาจากกลุ่มผู้เข้าสอบสองกลุ่มนั้น จะมีความเท่าเทียมกัน และสามารถเขียนเป็นสมการเส้นตรงได้ดังนี้ (Angoff.1984 : 569)

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

เมื่อ X , Y แทน คะแนนที่ได้จากแบบทดสอบฉบับ X และ Y

M_x , M_y แทน ค่าเฉลี่ยของคะแนนจากแบบทดสอบฉบับ X และ Y

S_x , S_y แทน ความเบี่ยงเบนมาตรฐานของจากแบบทดสอบฉบับ X และ Y

วิธีนี้เป็นการเทียบคะแนนแบบเส้นตรงที่มีข้อจำกัดที่ต้องคำนึงถึงที่สำคัญ คือ ประการแรก ได้แก่ ความเป็นคู่ขนานกันของแบบทดสอบที่นำมาใช้เทียบคะแนน ประการที่สอง ได้แก่ ค่าการกระจายของคะแนนจากแบบทดสอบทั้งสองฉบับ ต้องมีรูปทรงที่คล้ายคลึงกัน และเป็นเชิงปกติ ประการที่สาม ได้แก่ ค่าพหุคูณของแบบทดสอบ ซึ่งหมายถึงค่าความยาก และอำนาจจำแนกของแบบทดสอบทั้งสองฉบับ ต้องมีค่าใกล้เคียงกัน จากข้อจำกัดดังกล่าวนี้ ทำให้เป็นปัญหาด้านทางปฏิบัติ จึงมีนักวัดผลทางการศึกษาหลายท่านพยายามคิดค้นวิธีการแก้ข้อจำกัดนี้ และได้เสนอแนวทางต่าง ๆ โดยเฉพาะการปรับเปลี่ยนจากแบบทดสอบที่เป็นคู่ขนานกัน มาสู่แบบทดสอบที่ไม่มีความเป็นคู่ขนานกัน (Non-parallel) จนกลายมาเป็นคะแนนจริงสัมพันธ์ (Congeneric) ดังจะได้กล่าวต่อไป

2. วิธีการเทียบคะแนนแบบใช้ตำแหน่งช่วงที่เท่ากัน (Equipercntile Equating)

วิธีนี้เป็นการเทียบคะแนนจากแบบทดสอบต่างกัน สองฉบับที่วัดคุณลักษณะเดียวกัน มีค่าความเชื่อมั่นเท่า ๆ กัน มาเทียบคะแนนโดยใช้คะแนนช่วงของเปอร์เซ็นต์ไทล์ (Percentile Ranks) ที่ตรงกัน เป็นคะแนนที่เท่าเทียมกัน วิธีนี้ต่างกับวิธีเทียบคะแนนแบบเส้นตรง ที่ว่า การกระจายของคะแนนดิบจากแบบทดสอบ สองฉบับอาจมีรูปร่างที่แตกต่างกัน คือเส้นกราฟจะไม่เป็นเส้นตรงได้

ทั้งวิธีการเทียบคะแนนแบบเส้นตรง และวิธีเทียบคะแนนแบบใช้ตำแหน่งช่วงที่เท่ากันนั้น สามารถให้คะแนนใกล้เคียงกันอย่างมาก เมื่อรูปร่างของการแจกแจงคะแนนดิบจากแบบทดสอบทั้งสองคล้ายคลึงกัน

3. วิธีการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ (IRT Equating)

วิธีนี้อาศัยลักษณะคะแนนสมมูลกันจากแบบทดสอบสองฉบับ โดยใช้การกะประมาณระดับความสามารถเดียวกัน และไม่คำนึงถึงความแตกต่างของผู้เข้าสอบ แต่ใช้ความสามารถของผู้เข้าสอบที่จะกะประมาณค่าได้ จากฟังก์ชันลักษณะการตอบข้อสอบ (Item Characteristic Curve) ซึ่งวิธีการนี้จะแทนที่แปลงได้จะไม่แปรผัน (Invariance) ไปตามกลุ่มผู้เข้าสอบ

แผนแบบการเก็บรวบรวมข้อมูล

การเทียบคะแนนนั้นจะต้องอาศัยลักษณะการจัดกลุ่มของผู้เข้าสอบ เพื่อนำคะแนนที่ได้จากการสอบ ไปจัดกระทำด้วยวิธีการทางสถิติ และแปลงคะแนนให้มาสู่คะแนนมาตราเดียวกัน จึงจะสามารถนำมาเทียบคะแนนกันได้ แองกอฟ (Angoff, 1984 : 569-586) ได้เสนอแนวความคิดเกี่ยวกับแผนแบบการเก็บรวบรวมข้อมูล โดยแบ่งออกเป็นแผนแบบต่าง ๆ ที่เสนอไว้มี 5 แบบด้วยกัน คือ

แบบที่ 1 กลุ่มผู้เข้าสอบกลุ่มเดียว ทำแบบทดสอบฉบับเดียวที่มีคล้ายคลึงกัน กล่าวคือ 1) กลุ่มหนึ่งทำแบบทดสอบฉบับที่ 1 และอีกกลุ่มหนึ่งทำแบบทดสอบฉบับที่ 2 โดยแบบทดสอบทั้งสองต้องมีความคู่ขนานกัน

แบบที่ II กลุ่มผู้เข้าสอบสองกลุ่ม ที่ได้มาจากการสุ่มตัวอย่าง ให้ทำแบบทดสอบทั้งสองฉบับ ซึ่งเป็นการจัดให้มีคุณสมบัติกัน กล่าวคือ ทั้งกลุ่มที่หนึ่ง และกลุ่มที่สอง ได้ทำแบบทดสอบทั้งสองฉบับ

แบบที่ III กลุ่มผู้เข้าสอบสองกลุ่มที่ได้มาจากการสุ่มตัวอย่าง ให้ทำแบบทดสอบคนละฉบับ และมีการใช้แบบทดสอบหลักร่วมด้วย โดยทดสอบกับกลุ่มตัวอย่างทั้งสองกลุ่ม กล่าวคือ ให้กลุ่มที่หนึ่งทำแบบทดสอบฉบับที่ 1 กับแบบทดสอบหลักหรือแบบทดสอบเชื่อมโยง และกลุ่มที่สองทำแบบทดสอบฉบับที่ 2 กับแบบทดสอบหลัก

แบบที่ IV กลุ่มผู้เข้าสอบสองกลุ่ม ที่ไม่ได้มาจากการสุ่มตัวอย่าง ทำแบบทดสอบคนละฉบับ และมีการใช้แบบทดสอบหลักร่วมด้วย โดยทดสอบกับกลุ่มตัวอย่างทั้งสองกลุ่ม กล่าวคือ กลุ่มตัวอย่างทั้งสองกลุ่มมีความแตกต่างกัน กลุ่มที่หนึ่งทำแบบทดสอบฉบับที่ 1 กับแบบทดสอบหลัก และกลุ่มที่สองทำแบบทดสอบฉบับที่ 2 กับแบบทดสอบหลัก

แบบที่ V วิธีการที่นำข้อมูลไปใช้ทางสถิติ ได้แก่

ก. นำคะแนนจากแบบทดสอบสองฉบับ มาเทียบคะแนนกัน โดยใช้แบบทดสอบหลัก มีเงื่อนไขว่า ถ้าฟอร์ม X และฟอร์ม Y เป็นแบบทดสอบคู่ขนานกัน และแบบทดสอบหลักมีลักษณะเหมือนกับแบบทดสอบทั้งสองฉบับด้วย ก็สามารถนำมาเทียบคะแนนกันได้ โดยสุ่มจากประชากรอิสระได้ แต่ถ้าแบบทดสอบหลักไม่คู่ขนานกันกับแบบทดสอบฟอร์ม X และฟอร์ม Y แล้ว ก็จะต้องสุ่มมาจากประชากรเดียวกัน

ข. ใช้การพยากรณ์จากแบบทดสอบหลัก ไปยังแบบทดสอบที่นำมาเทียบคะแนนกัน โดยใช้แบบทดสอบฟอร์ม X สอบกับประชากรกลุ่มหนึ่ง ฟอร์ม Y สอบกับประชากรอีกกลุ่มหนึ่ง คะแนนที่ได้จากการสอบจะเป็นคะแนนสมมูลกัน ดังนั้น จึงสามารถพยากรณ์คะแนนจากฟอร์ม X และฟอร์ม Y จะถูกพยากรณ์ โดยคะแนนที่มาจากคะแนนของแบบทดสอบหลักได้

ค. ใช้คะแนนจากแบบทดสอบที่นำมาเทียบคะแนนกัน พยากรณ์ไปยังแบบทดสอบหลัก โดยใช้แบบทดสอบฟอร์ม X สอบกับประชากรกลุ่มหนึ่ง ฟอร์ม Y สอบกับประชากรอีกกลุ่มหนึ่ง และให้สอบแบบทดสอบหลักทั้งสองกลุ่ม คะแนนที่ได้จากแบบทดสอบฟอร์ม X และฟอร์ม Y จะเป็นคะแนนสมมูลกัน ถ้าต่างก็ใช้พยากรณ์คะแนนของแบบทดสอบหลัก

การพัฒนาวิธีการเทียบคะแนนแบบเส้นตรง กรณีที่ใช้แบบทดสอบหลัก

1. วิธีการเทียบคะแนนของทัคเกอร์ (Tucker Equating)

ทัคเกอร์ (Tucker) เป็นผู้ริเริ่มให้มีการใช้แบบทดสอบหลักที่ไม่จำเป็นต้องเป็นแบบทดสอบคู่ขนานกัน (Non-parallel forms) แต่ใช้แผนแบบการเก็บรวบรวมข้อมูลที่เป็น "แผนแบบของการใช้แบบทดสอบหลักที่มีประชากรแตกต่างกัน" (Common Item Non-equivalent Population Design) ที่ยังยึดเนื้อหาเดียวกันอยู่และกำหนดเนื้อหาในแบบทดสอบหลักนี้ว่า จะต้องเป็นตัวแทนของแบบทดสอบทั้งสองฉบับที่จะนำมาเทียบกัน วิธีการที่ทัคเกอร์ใช้ในการจัดกลุ่มข้อสอบใหม่นี้ เรียกว่า "การสังเคราะห์ประชากร" (The Synthetic Population) ซึ่งเป็นการรวมข้อสอบที่มาจากแบบทดสอบต่างกัน สองฉบับ หรือมากกว่า มาจัดเป็นแบบทดสอบหลักฉบับใหม่ขึ้นมา ภายใต้การกำหนดสัดส่วนของน้ำหนักของแบบทดสอบ สองฉบับที่ต่างกัน โดยน้ำหนักรวม ($w_1 + w_2$) มีค่าเท่ากับ 1 และน้ำหนักของข้อสอบที่มาจากแบบทดสอบทั้งสองฉบับ มีค่ามากกว่าและเท่ากับ 0 ($w_1, w_2 \geq 0$) (Koren.1985 : 210 ; Koren and Brennan.1987 : 264)

ข้อตกลงของวิธีเทียบคะแนนของทัคเกอร์ ที่ได้มาจากแผนแบบการเก็บรวบรวมข้อมูลแบบที่ IV ของแองกอฟ (Angoff.1982, 1984) ที่ว่า คะแนน X ได้มาจากการสอบของประชากรกลุ่มที่หนึ่ง คะแนน Y ได้มาจากการสอบของประชากรกลุ่มที่สอง และคะแนน V ได้มาจากการสอบของประชากรทั้งกลุ่มที่หนึ่ง และกลุ่มที่สอง นำคะแนนที่สอบได้มาหาสมการถดถอยของคะแนน X บนคะแนน V ซึ่งข้อตกลงของทัคเกอร์กำหนดไว้ว่า (Angoff.1982; Koren. 1985 ; Budescu. 1985)

1. ประชากรทั้งสองกลุ่ม ต้องมีระดับความสามารถที่ไม่แตกต่างกันมาก
2. ความลาดชัน และค่าพิกัดแกนตั้ง ของเส้นสมการถดถอยที่ใช้พยากรณ์คะแนน X จากคะแนน V ในประชากรกลุ่มรวม (กลุ่มที่หนึ่ง และกลุ่มที่สอง) กับประชากรกลุ่มที่หนึ่ง ต้องมีค่าเท่ากัน

3. ความคลาดเคลื่อนของความแปรปรวน (Variance Error) ของคะแนน X จากคะแนน V ในประชากรกลุ่มรวมกับประชากรกลุ่มที่หนึ่ง ต้องมีค่าเท่ากัน

ข้อตกลงทั้ง 3 ข้อที่กล่าวมานี้ จะมีความเป็นไปไม่ได้เมื่อ แบบทดสอบที่นำมาเทียบคะแนนกันนั้น ต้องมีความเชื่อมั่นของแบบทดสอบที่เท่าเทียมกัน คะแนน X คะแนน Y และคะแนน V

มีการกระจายเป็นฟังก์ชันปกติ

เบรนนัน และโคเลน (Brennan and Koren.1987 ; Koren.1985) ได้แสดงวิธีการเทียบคะแนนแบบเส้นตรงของทักเกอร์ โดยเขียนสมการของการเทียบคะแนน ที่มาจากการใช้แบบทดสอบหลักที่เป็นการสังเคราะห์ประชากรไว้ว่า

$$l(x) = \frac{\sigma_S(X)}{\sigma_S(Y)} [X - u_S(X)] + u_S(X)$$

เมื่อ $u_S(X)$ แทน ค่าเฉลี่ยของคะแนน X บนคะแนน V

$u_S(Y)$ แทน ค่าเฉลี่ยของคะแนน Y บนคะแนน V

$\sigma_S(X)$ แทน ความเบี่ยงเบนมาตรฐานของคะแนน X บนคะแนน V

$\sigma_S(Y)$ แทน ความเบี่ยงเบนมาตรฐานของคะแนน Y บนคะแนน V

และสามารถหาค่าเหล่านี้ได้โดยใช้สมการต่อไปนี้

$$u_S(X) = u_1(X) - w_2\alpha_1 (X/V)[u_1(V) - u_2(V)]$$

$$u_S(Y) = u_2(X) - w_1\alpha_2 (Y/V)[u_1(V) - u_2(V)]$$

$$\begin{aligned}\sigma_S^2(X) &= \sigma_1^2(X) - w_2\alpha_1^2(X/V)[\sigma_1^2(V) - \sigma_2^2(V)] + \\ &\quad w_1w_2\alpha_1^2(X/V)[u_1(V) - u_2(V)]^2\end{aligned}$$

$$\begin{aligned}\sigma_S^2(Y) &= \sigma_2^2(Y) - w_1\alpha_2^2(Y/V)[\sigma_1^2(V) - \sigma_2^2(V)] + \\ &\quad w_1w_2\alpha_2^2(Y/V)[u_1(V) - u_2(V)]^2\end{aligned}$$

เมื่อ $\alpha_1 (X/V)$ แทน ค่าความลาดชันของเส้นสมการถดถอยของคะแนน X บนคะแนน V

$\alpha_2 (Y/V)$ แทน ค่าความลาดชันของเส้นสมการถดถอยของคะแนน Y บนคะแนน V

$$\alpha_1 (X/V) = \sigma_1 (X,V) / \sigma_1^2(V)$$

$$\alpha_2 (Y/V) = \sigma_2 (Y,V) / \sigma_2^2(V)$$

ดังนั้นเมื่อต้องการเทียบคะแนนโดยวิธีของทักเกอร์ ก็สามารถใช้สมการดังกล่าวข้างต้น และหาค่าน้ำหนักคะแนนได้จากสัดส่วนของขนาดกลุ่มตัวอย่าง นั่นคือ

$$w_1 = n_1 / (n_1 + n_2)$$

$$w_2 = n_2 / (n_1 + n_2)$$

เมื่อ n_1 และ n_2 แทน ขนาดกลุ่มตัวอย่างของผู้สอบ

2. วิธีเทียบคะแนนสอบที่อาศัยแบบทดสอบคะแนนจริงสัมพันธ์

แมคแคน (MacCann. 1989) ได้ศึกษาวิธีการเทียบคะแนนสอบ โดยใช้แบบทดสอบที่เป็นคะแนนจริงสัมพันธ์ คือแบบทดสอบทั้งสองฉบับและแบบทดสอบหลัก ต่างก็เป็นแบบทดสอบที่วัดคุณลักษณะที่เหมือนกันทั้งสามฉบับ แมคแคนได้แนวความคิดมาจากรูปแบบการใช้แบบทดสอบหลัก ที่เป็นแผนแบบของการใช้แบบทดสอบหลักที่มีประชากรแตกต่างกัน ซึ่งออกแบบโดยวูดดรัฟ (Woodruff. 1986) กับโคเรน และเบรแนน (Koren and Brennan. 1987)

วิธีการที่แมคแคนศึกษาไว้ อาศัยข้อตกลงของการเทียบคะแนนสอบ ที่นำมาจากวิธีของแองกอฟ 2 ข้อ ที่ว่า

1. แบบทดสอบสองฉบับ ที่มีหน้าที่ทางจิตวิทยาที่มีคุณลักษณะเดียวกัน หรือมิติเดียวกัน และรูปทรงการกระจายของคะแนนสอบ จากแบบทดสอบทั้งสองฉบับต้องมีความสมมูลกัน ทั้งนี้ เมื่อแบบทดสอบทั้งสอง มีค่าความเชื่อมั่นสูงมากขึ้น ก็จะทำให้รูปร่างการกระจายของคะแนนจริงที่เป็นรูปทรงเดียวกัน (Identical) คือ มีการกระจายของคะแนนจริงที่สมมูลเช่นกัน ซึ่งทั้งนี้

แบบทดสอบทั้งสองฉบับ จะต้องวัดหน้าที่คุณลักษณะเดียวกัน คะแนนจริงของแบบทดสอบจึงจะมีความสัมพันธ์กันอย่างสมบูรณ์ หรือสัมพันธ์กันเป็นเส้นตรง เรียกแบบทดสอบทั้งสองว่า เป็นแบบทดสอบคะแนนจริงสัมพันธ์ ซึ่งในทางปฏิบัติการเทียบคะแนนแบบเส้นตรงจะยอมให้มีการกระจายของคะแนนจริงแตกต่างกันได้บ้าง และยอมให้มีการวัดหน้าที่คุณลักษณะที่แตกต่างกันได้บ้างเล็กน้อย ถืออาศัยข้อตกลงของคะแนนจริงสัมพันธ์ และตามความเป็นจริงในทางปฏิบัติแล้ว ถ้าใช้เงื่อนไขดังกล่าวที่แตกต่างไปจากข้อตกลงการเทียบคะแนนแบบเส้นตรง ก็อาจไม่ได้ผล

2. แบบทดสอบสองฉบับที่นำมาเทียบคะแนนสอบ จะต้องมีความเชื่อมั่นเท่าเทียมกัน การทำให้แบบทดสอบมีความเชื่อมั่นเท่าเทียมกันนั้น จะต้องใช้กลุ่มผู้เข้าสอบที่มีความสามารถเท่าเทียมกัน หากว่ากลุ่มผู้เข้าสอบมีระดับความสามารถสูง ก็จะมีผลทำให้มีความเชื่อมั่นสูงขึ้น (MacCann.1989 : 264 ; citing Lord. 1977) และแอ็งกอฟ (MacCann.1989 : 264 ; citing Angoff. 1977) ได้เสนอแนะเกี่ยวกับการเทียบคะแนน จากแบบทดสอบที่มีความเชื่อมั่นไม่เท่าเทียมกัน โดยการปรับความแตกต่างนี้ จากการทำให้แบบทดสอบมีการวัดในคุณลักษณะเดียวกัน ก็จะได้แบบทดสอบที่มีความเชื่อมั่นเท่าเทียมกัน

จากข้อตกลงทั้งสองข้อดังกล่าวข้างต้น เขียนเป็นสมการได้ ดังนี้

$$(1) \quad T_{x1} = aT_{y1} + b \quad \text{เมื่อ } a \text{ และ } b \text{ เป็นค่าคงที่ และ } a > 0$$

$$(2) \quad r_{xx} = r_{yy}$$

เมื่อ r_{xx} และ r_{yy} เป็นความเชื่อมั่นของแบบทดสอบฉบับ X และ Y ตามลำดับ

วิธีการคำนวณการเทียบคะแนนแบบเส้นตรงของแมคแคน ตามแบบจำลองคะแนนจริงสัมพันธ์นี้เป็นการนำคะแนนสอบมาเทียบคะแนนกัน โดยยึดข้อตกลงดังกล่าวข้างต้น สามารถใช้วิธีการคำนวณได้เป็น สองวิธี ดังนี้ (บุญเชิด วิทยุณันตพงษ์.2538)

การแสดงผลการได้มาของการเทียบคะแนนสอบวิธีที่ 1

การแสดงผลการได้มาของการเทียบคะแนนสอบวิธีที่ 1 จะอาศัยข้อตกลงมาตรฐานของการแสดงที่มาของการเทียบคะแนนสอบที่ใช้กัน คือ อาศัยข้อตกลงของเส้นถดถอยที่สมมูลกัน (Equivalent) ซึ่งส่วนมากมักเรียกว่า เส้นถดถอยที่ขนานกัน

สมมุติให้ X และ Y เป็นแบบทดสอบสองฟอร์ม นำไปสอบนักเรียนกลุ่มที่ 1 และกลุ่มที่ 2 ตามลำดับ และให้ Z เป็นแบบทดสอบหลัก ที่นำไปสอบกับนักเรียนกลุ่มรวมทั้งหมด (กลุ่มที่ 1 และกลุ่มที่ 2 รวมกัน)

การแสดงผลการได้มาของสูตรจะใช้ข้อตกลงเกี่ยวกับ เส้นถดถอยสองเส้นที่ขนานกัน จากการกำหนดข้างต้นจะได้เส้นถดถอยของการพยากรณ์ X ด้วย Z สองเส้น ดังนี้ เส้นถดถอยเส้นหนึ่งได้จากคะแนนของนักเรียนกลุ่มรวม (กลุ่มที่ 1 และกลุ่มที่ 2) และเส้นถดถอยอีกเส้นหนึ่งได้จากคะแนนของนักเรียนกลุ่มที่ 1

จากข้อตกลงที่ว่าเส้นถดถอยสองเส้นขนานกัน สามารถแตกออกมาเป็น สองข้อตกลงย่อย คือ

1. ข้อตกลงที่ว่าเส้นถดถอยสองเส้นต้องมีความลาดชันเท่ากัน (Equal Slope)
2. ข้อตกลงที่ว่าเส้นถดถอยสองเส้นต้องมีพิสัยตัดแกนตั้งเท่ากัน (Equal

Ordinate Intercepts)

ข้อตกลงความลาดชันของเส้นถดถอยสองเส้นจากนักเรียนกลุ่มรวม และนักเรียนกลุ่มที่ 1 เขียนเป็นสมการได้ดังนี้

$$(3) \quad b_{XZ} = b_{XZ1}$$

เมื่อ b_{XZ} และ b_{XZ1} เป็นสัมประสิทธิ์ความลาดชันของคะแนนนักเรียนกลุ่มรวม และกลุ่มที่ 1 ตามลำดับ

และมี

$$(4) \quad b_{xz} = r_{xz} \left[\frac{S_x}{S_z} \right]$$

$$(5) \quad b_{xz1} = r_{xz1} \left[\frac{S_{x1}}{S_{z1}} \right]$$

r_{xz} และ r_{xz1} เป็นสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบแต่ละฟอร์ม ของคะแนนจากนักเรียนกลุ่มรวม และกลุ่มที่ 1 ตามลำดับ

S_x และ S_z เป็นส่วนเบี่ยงเบนมาตรฐานของคะแนนจากนักเรียนกลุ่มรวม ที่สอบแบบทดสอบฟอร์ม X และ ฟอร์ม Z ตามลำดับ

S_{x1} และ S_{z1} เป็นส่วนเบี่ยงเบนมาตรฐานของคะแนนจากนักเรียนกลุ่มที่ 1 ที่สอบแบบทดสอบฟอร์ม X และ ฟอร์ม Z ตามลำดับ

ข้อตกลงของพิกัดแกนตั้งของเส้นถดถอยสองเส้น จากนักเรียนกลุ่มรวมและนักเรียนกลุ่มที่ 1 เขียนเป็นสมการได้ ดังนี้

$$(6) \quad \bar{X} - b_{xz}\bar{Z} = \bar{X}_1 - b_{xz1}\bar{Z}_1$$

นำสมการ (3) แทนค่าลงในสมการ (6) และจัดเรียงค่าใหม่ จะประมาณค่าเฉลี่ยของ X จากนักเรียนกลุ่มรวมได้ ดังนี้

$$(7) \quad \bar{X} = \bar{X}_1 + b_{xz1} (\bar{Z} - \bar{Z}_1)$$

พหุคูณเดียวกัน นำสมการ (3) แทนค่าลงในสมการ (4) และจัดเรียงค่าใหม่ จะประมาณค่า ส่วนเบี่ยงเบนมาตรฐานของคะแนน X จากนักเรียนกลุ่มรวมได้ ดังนี้

$$(8) \quad S_x = \left[\begin{array}{c} b_{xz1} \\ \hline r_{xz} \end{array} \right] S_z$$

พหุคูณเดียวกัน ถ้าใช้เส้นถดถอยของการพยากรณ์ Y ด้วย Z จากนักเรียนกลุ่มรวม เส้นหนึ่ง และจากนักเรียนกลุ่มที่สองอีกเส้นหนึ่ง ซึ่งตามข้อตกลงของเส้นถดถอยที่ขนานกัน จะได้ สมการ ที่คล้ายกับสมการ (7) และ (8) ดังนี้

$$(7.1) \quad \bar{Y} = \bar{Y}_2 + b_{yz2} (\bar{Z} - \bar{Z}_2)$$

และ

$$(8.1) \quad S_y = \left[\begin{array}{c} b_{yz2} \\ \hline r_{yz} \end{array} \right] S_z$$

คะแนนสอบ สองคะแนน สามารถนำมาเทียบคะแนนแบบเส้นตรงได้ โดยถือว่าคะแนน มาตรฐานที่ได้จากคะแนนทั้งสองเท่ากัน จากแนวความคิดนี้ จะสามารถแปลงคะแนนสอบได้จาก สมการเส้นตรง ดังนี้

$$(9) \quad Y = \bar{Y} + \frac{S_y}{S_x} (X - \bar{X})$$

เนื่องจาก X และ Y สามารถหาได้ โดยตรงจากคะแนนสอบตั้งสมการ (7) และ (7.1) แต่ค่าอัตราส่วน S_y และ S_x ไม่สามารถหาได้ เนื่องจากไม่ทราบค่า r_{xz} และ r_{yz} เพราะในทางปฏิบัติเราไม่ได้นำ X และ Y ไปสอบกับนักเรียนกลุ่มรวม ดังนั้น ค่า S_y และ S_x จึงต้องใช้วิธีประมาณค่าจากสมการ (8) และ (8.1) โดยนำสมการ (8.1) ไปหารสมการ (8) ดังนี้

$$(10) \quad \frac{S_y}{S_x} = \frac{b_{yz2}r_{xz}}{b_{xz1}r_{yz}}$$

และใช้สูตรการปรับแก้สัมประสิทธิ์สหสัมพันธ์ที่ลดลง เนื่องจากความคลาดเคลื่อนแล้วจะได้สมการ ดังนี้

$$(11) \quad r_{xz} = r_{TxTz} \sqrt{r_{xx}r_{zz}}$$

ทำนองเดียวกันก็จะได้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง Y และ Z ดังนี้

$$(11.1) \quad r_{yz} = r_{TyTz} \sqrt{r_{yy}r_{zz}}$$

หารสมการ (11) ด้วยสมการ (11.1) และใช้ข้อตกลงที่ว่า ความเชื่อมั่นของแต่ละฟอร์มเท่ากันตามสมการ (2) จะได้

$$(12) \quad \frac{r_{xz}}{r_{yz}} = \frac{r_{TxTz}}{r_{TyTz}}$$

ตามที่สมการ (1) ซึ่งเป็นข้อตกลงของคะแนนจริงสัมพันธ์มีความหมายว่าคะแนนจริงมีความสัมพันธ์เป็นเส้นตรง จากสมการ (1) นี้ จะเห็นได้ว่าสหสัมพันธ์ของ T_Z กับ T_X จะเท่ากับสหสัมพันธ์ของ T_Z กับ T_Y นั่นคือ

$$(13) \quad r_{T_X T_Z} = r_{T_Y T_Z}$$

แทนค่าสมการ (12) และ (13) ลงในสมการ (10) จะได้

$$(14) \quad \frac{S_Y}{S_X} = \frac{b_{Y Z_2}}{b_{X Z_1}}$$

ขั้นสุดท้ายแทนสมการ (7.1) และสมการ (14) ลงในสมการ (9) และเขียนเป็นนิพจน์ของความลาดชันทั้งหมด หรือสูตรสำหรับแปลงคะแนนสอบได้ดังนี้

$$(15) \quad Y = \left[\frac{r_{Y Z_2} S_{Y_2} S_{Z_1}}{r_{Z X_1} S_{X_1} S_{Z_2}} \right] X + \bar{Y}_2 + \left[\frac{r_{Y Z_2} S_{Y_2}}{S_{Z_2}} \right] \left[\frac{(\bar{Z}_1 - \bar{Z}_2)}{r_{Z X_1} S_{X_1}} - \frac{1}{S_{Z_1}} \right] \bar{X}_1$$

การแสดงผลการได้มาของการเทียบคะแนนสอบวิธีที่ 2

การแสดงผลการได้มาของวิธีที่ 1 ดังที่กล่าวไปแล้วนั้นได้ใช้ข้อตกลงของเส้นถดถอยของการพยากรณ์ X ด้วย Z สำหรับการแสดงผลการได้มาของวิธีที่ 2 จะใช้เส้นถดถอยของการพยากรณ์ Z ด้วย X โดยกำหนดข้อตกลงของเส้นถดถอยว่า ความลาดชันของเส้นถดถอยว่า ความลาดชันของเส้นถดถอยที่ได้จากคะแนนกลุ่มที่ 1 เท่ากับความลาดชันที่ได้จากนักเรียนกลุ่มรวม และพิสัยแกนตั้งของเส้นถดถอยที่ได้จากนักเรียนกลุ่มที่ 1

จากข้อตกลงของความลาดชันจะได้

$$(16) \quad b_{zx} = b_{zx1}$$

และมี

$$(17) \quad b_{zx} = r_{zx} \begin{bmatrix} s_z \\ \text{---} \\ s_x \end{bmatrix}$$

$$(17.1) \quad b_{zx1} = r_{zx1} \begin{bmatrix} s_{z1} \\ \text{---} \\ s_{x1} \end{bmatrix}$$

และใช้ข้อตกลงที่ว่า พิกัดแกนตั้งมีค่าเท่ากัน จะเขียนเป็นสมการได้ดังนี้

$$(18) \quad \bar{z} - b_{zx}\bar{x} = \bar{z}_1 - b_{zx1}\bar{x}_1$$

จากสมการที่ (16) ถึง (18) จะประมาณค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานของคะแนน X ที่สอบจากนักเรียนกลุ่มรวม ซึ่งจะได้สมการที่คล้ายคลึงกับในวิธีที่ 1

$$(19) \quad \bar{x} = \bar{x}_1 + \frac{1}{b_{zx1}} (\bar{z} - \bar{z}_1)$$

$$(20) \quad s_x = \begin{bmatrix} 1 \\ \text{---} \\ b_{zx1} \end{bmatrix} r_{xz} s_z$$

ทำนองเดียวกัน ถ้าใช้เส้นถดถอยของการพยากรณ์ Z ด้วย Y จากนักเรียนกลุ่มที่ 2 จะได้ค่าสมการดังนี้

$$(19.1) \quad \bar{Y} = \bar{Y}_2 + \frac{1}{b_{zy2}} (\bar{Z} - \bar{Z}_2)$$

และ

$$(20.1) \quad S_y = \left[\frac{1}{b_{zy2}} \right] r_{zy} S_z$$

ทำนองเดียวกันกับการได้มาของสมการ (14) ในวิธีที่ 1 จะได้

$$(21) \quad \frac{S_y}{S_x} = \frac{b_{zx1}}{b_{zy2}}$$

แทนค่าสมการ (19.1) และ (21) ลงในสมการ (9) จะเขียนเป็นนิพจน์ของความลาดชันทั้งหมด หรือสมการเทียบคะแนนสอบดังนี้

$$(22) \quad Y/ = \begin{bmatrix} r_{xz1} & S_{y2} & S_{z1} \\ \hline & & \\ r_{yz2} & S_{x1} & S_{z2} \end{bmatrix} X + \bar{Y}_2 + \begin{bmatrix} 1 & S_{y2} \\ \hline & \\ r_{yz2} & S_{z2} \end{bmatrix} \left[(\bar{Z}_1 - \bar{Z}_2) - r_{zx1} \frac{S_{z1}}{S_{x1}} \bar{X}_1 \right]$$

ถ้าเปรียบเทียบสมการ (22) กับ (15) จะเห็นว่าวิธีที่ 2 แตกต่างจากวิธีที่ 1 เฉพาะที่ค่าสหสัมพันธ์เท่านั้น คือค่าสหสัมพันธ์ในสมการ (22) จะเขียนเป็นส่วนกลับของค่าสหสัมพันธ์ในสมการ (15)

วูดดรัฟ (Woodruff, 1986: 245-257) แห่งสำนักงานโครงการทดสอบระดับวิทยาลัยของอเมริกา (The American College Testing Program) ได้เสนอวิธีการเทียบคะแนนโดยอาศัยแบบจำลองคะแนนจริงสัมพันธ์ไว้ 2 วิธี ซึ่งแสดงที่มาของสูตรโดยอาศัยแผนแบบการเก็บรวบรวมข้อมูลที่เป็นการใช้แบบทดสอบหลักที่มีประชากรต่างกัน (CINEP) โดยมีรายละเอียด ดังนี้ (บุญเชิด วิทยุอนันตพงษ์, 2538)

ให้ X แทน แบบทดสอบฟอร์มใหม่ที่สอบกับนักเรียนกลุ่มที่ 1

Y แทน แบบทดสอบฟอร์มเดิมที่สอบกับนักเรียนกลุ่มที่ 2

Z แทน แบบทดสอบหลัก (Anchor Test) ซึ่งมีจำนวนข้อหรือความยาว

น้อยกว่าฉบับ X และ Y โดยวัดเนื้อหาเดียวกันทั้งสามฉบับ

การเทียบคะแนนจะแปลงคะแนนฟอร์ม X ให้เป็นคะแนนฟอร์ม Y โดยอาศัยคะแนนจากแบบทดสอบหลัก Z ซึ่งทำได้ 2 วิธี ถ้าใช้แบบทดสอบ Z สอบแยกฉบับจะเรียกแบบทดสอบหลักนี้ว่าแบบทดสอบเชื่อมโยง แต่ถ้าใช้แบบทดสอบ Z รวมกับ X และ Y จะเรียกว่า การใช้แบบทดสอบหลัก (เช่นเดียวกับที่กล่าวไปตอนต้น) คะแนนที่ได้จากการนำ Z มารวมไว้จะกลายเป็นฉบับ A และฉบับ B นั่นคือ $A = X + Z$ และ $B = Y + Z$ ในที่นี้จะนำเสนอเฉพาะการเทียบคะแนนแบบเส้นตรง โดยใช้แบบทดสอบเชื่อมโยง (External Anchor Test) ซึ่งจะอาศัยสมการหลักดังนี้

$$(1) \quad \frac{Y - u(Y)}{s(Y)} = \frac{X - u(X)}{s(X)}$$

ถ้าแปลงคะแนน X ให้เป็นคะแนน Y จะได้สมการ

$$(2) \quad Y/ = \frac{\sigma(Y)}{\sigma(X)} X + \left[u(Y) - \frac{\sigma(Y)}{\sigma(X)} u(X) \right]$$

ในกรณีที่เทียบคะแนนจากแบบทดสอบที่มีค่าความเชื่อมั่นไม่เท่ากัน ลีวีน (Levine) และแองกอล์ฟ (Angolf) ใช้ความเบี่ยงเบนมาตรฐานของคะแนนจริงแทน ความเบี่ยงเบนมาตรฐานของคะแนนรวม ดังนี้

$$(3) \quad \frac{Y - u(Y)}{\sigma(T_Y)} = \frac{X - u(X)}{\sigma(T_X)}$$

แล้วตกลงให้ $u(T_Y) = u(Y)$, และ $u(T_X) = u(X)$
จากสมการ (2) สมการ แปลงคะแนน X ที่สอบจากกลุ่มที่ 1 ให้เป็นคะแนน Y ของกลุ่มที่ 1
ซึ่งเป็นคะแนนเทียบมา หรือคะแนนแปลง เขียนแทนด้วย Y_1 ได้ดังนี้

$$(4) \quad Y_1 = \tilde{\pi}_{EX} X_1 + \tilde{w}_{EX}$$

$$(5) \quad \text{เมื่อ } \tilde{\pi}_{EX} = \tilde{\sigma}(Y_1) / \sigma(X_1)$$

$$(6) \quad \tilde{w}_{EX} = \tilde{u}(Y_1) - \tilde{\pi}_{EX} u(X_1)$$

ตัวห้อย EX เป็นสัญลักษณ์ที่ใช้แสดงถึงการเทียบคะแนน โดยใช้แบบทดสอบเชื่อมโยง (External Anchor) ในกรณีที่ใช้แบบทดสอบหลัก (Internal Anchor) จะได้สมการ ดังนี้

$$(7) \quad B = \tilde{\pi}_{IX} A_1 + \tilde{w}_{IX}$$

$$(8) \quad \text{เมื่อ} \quad \tilde{\pi}_{IX} = \tilde{\sigma}(B_1) / \sigma(A_1)$$

ตัวห้อย IX เป็นสัญลักษณ์แทนการเทียบคะแนนที่ใช้แบบทดสอบหลัก (Internal Anchor Test)

การประมาณค่าพารามิเตอร์ $\tilde{u}(Y_1)$, $\tilde{u}(B_1)$, $\tilde{\sigma}(Y_1)$, และ $\tilde{\sigma}(B_1)$ จากการแปลงคะแนนของแบบทดสอบ หามาได้โดยอาศัยแบบจำลองของคะแนนสองแบบจำลองดังนี้

วิธีที่ 1 การเทียบคะแนนตามแบบจำลองคะแนนจริงสัมพันธ์ (Congeneric Model)

การเทียบคะแนนสอบของแบบทดสอบหลาย ๆ ฉบับ ตามแบบจำลองคะแนนจริงสัมพันธ์นี้มีลักษณะคล้ายกับองค์ประกอบที่มีองค์ประกอบร่วม องค์ประกอบเดียว โดยไม่มีองค์ประกอบเฉพาะ (Unique factors) สามารถเขียนเป็นรูปสัญลักษณ์ได้ดังนี้

$$X_1 = \lambda_x T_1 + \delta_x + E_x$$

$$Y_2 = \lambda_y T_2 + \delta_y + E_y$$

$$\text{และ} \quad Z_i = \lambda_z T_i + \delta_z + E_z \quad (i = 1, 2)$$

ตัวห้อย 1 แสดงถึงคะแนนที่ได้มาจากกลุ่มที่ 1 และตัวห้อย 2 แสดงถึง คะแนนที่ได้มาจากกลุ่มที่ 2 T แทนคะแนนจริง ประชากรนักเรียนกลุ่มที่ 1 และกลุ่มที่ 2 อาจมีค่าเฉลี่ยและค่าความแปรปรวนต่างกันก็ได้ δ และ λ เป็นค่าพารามิเตอร์เฉพาะของแบบทดสอบ ที่เรียกว่า Location and Scale Parameter ตามลำดับ และมีข้อตกลงว่า จะมีค่าไม่แปรเปลี่ยน (Invariant) ไปตามกลุ่มประชากร ส่วน E แทนความคลาดเคลื่อนสุ่มของแบบทดสอบแต่ละฉบับมีค่าไม่เท่ากันได้ แต่ค่าความแปรปรวนของความคลาดเคลื่อนเฉพาะของแบบทดสอบ (σ)

ตกลงว่ามีค่าไม่แปรเปลี่ยนไปตามกลุ่มประชากรที่สอบ และค่าเฉลี่ยของคะแนน E ของแบบทดสอบทุกฉบับที่สอบจากประชากรทั้งสองกลุ่มมีค่าเป็นศูนย์ และไม่มีสหสัมพันธ์ซึ่งกันและกัน ตลอดจนไม่มีความสัมพันธ์กับคะแนนจริง ตามแบบจำลองคะแนนจริงสัมพันธ์นี้ บังคับว่า (Constrains)

$$\lambda_x + \lambda_y + \lambda_z = 1.0 \quad \text{และ} \quad \delta_x + \delta_y + \delta_z = 0$$

การประมาณค่า $\tilde{u}(Y_1)$ จะอาศัยค่าเฉลี่ยจากสมการหลักข้างต้นดังนี้

$$\begin{aligned} \tilde{u}(Y_1) - u(Y_2) &= \lambda_y [u(T_1) - u(T_2)] \\ &= (\lambda_y / \lambda_z) [u(Z_1) - u(Z_2)] \end{aligned}$$

ซึ่งจะได้ว่า

$$(10) \quad \tilde{u}(Y_1) = u(Y_2) + (\lambda_y / \lambda_z) [u(Z_1) - u(Z_2)]$$

หาอนุพันธ์ร่วมกันสามารถประมาณค่า $\tilde{\sigma}^2(Y_1)$ ได้จากสมการดังนี้

$$\begin{aligned} \tilde{\sigma}^2(Y_1) - \sigma^2(Y_2) &= \lambda_y^2 [\sigma^2(T_1) - \sigma^2(T_2)] \\ &= (\lambda_y^2 / \lambda_z^2) [\sigma^2(Z_1) - \sigma^2(Z_2)] \end{aligned}$$

ซึ่งจะได้ว่า

$$(11) \quad \tilde{\sigma}^2(Y_1) = \sigma^2(Y_2) + (\lambda_y / \lambda_z)^2 [\sigma^2(Z_1) - \sigma^2(Z_2)]$$

จากสมการ (10) และ (11) มีตัวแปรที่สำคัญ คือ λ_y และ λ_z ดังนั้น จึงควรเขียนตัวแปรค่าที่อยู่ในรูปของคะแนนที่สอบได้ดังนี้

ให้ b_{y2z2} เป็นความลาดชันของเส้นถดถอยของคะแนน Y จากคะแนน Z ที่สอบจากนักเรียนกลุ่มที่ 2

$$\begin{aligned} b_{y2z2} &= \frac{\sigma(Y_2, Z_2)}{\sigma^2(Z_2)} = \frac{\lambda_y \lambda_z \sigma^2(T_2)}{\lambda_z^2 \sigma^2(T_2) + \sigma^2(E_z)} \\ &= \frac{\lambda_y}{\lambda_z} \cdot \frac{\lambda_z^2 \sigma^2(T_2)}{\lambda_z^2 \sigma^2(T_2) + \sigma^2(E_z)} \\ &= (\lambda_y / \lambda_z) \rho(Z_2) \end{aligned}$$

เมื่อ $\sigma(Z_2)$ แทน ค่าความเชื่อมั่นของแบบทดสอบ Z ซึ่งได้มาจากข้อตกลงและข้อบังคับ (Constraints) ตามแบบจำลองคะแนนจริงสัมพันธ์ที่ จอเรสคอก (Joreskog, 1971) นิยาม ดังนั้นจะได้ว่า

$$(12) \quad \lambda_y / \lambda_z = b_{y2z2} / \rho(Z_2)$$

จากสมการ (12) นำไปแทนค่าลงในสมการ (10) และ (11) แล้วนำพจน์แทนลงในสมการ (5) และ (6) แล้วจะสามารถแปลงคะแนน X เป็นคะแนน Y ได้ตามสมการ (4)

วิธีที่ 2 การเทียบคะแนนแบบเส้นตรงตามแบบจำลองคะแนนสัมพันธ์แบบมาตรฐานเดิม
(Congeneric-Parallel Model)

แบบจำลองที่นำมาเสนอในแบบที่สองนี้ เรียกว่าการเทียบคะแนนสอบที่อาศัยข้อตกลงว่าแบบทดสอบแต่ละฟอร์ม มีคะแนนจริงสัมพันธ์และคู่ขนาน หรือเรียกว่า มีคะแนนจริงสัมพันธ์แบบมาตรฐานเดิม เพื่อนำมาใช้อ้างอิง แผนแบบการเก็บรวบรวมข้อมูลที่ใช้แบบทดสอบหลักที่มีประชากรแตกต่างกัน (CINEP) โดยนำแบบทดสอบที่จะนำมาเทียบคะแนนกันสองฟอร์ม คือ ฟอร์ม X กับ ฟอร์ม Y ไปสอบกับนักเรียนกลุ่มที่ 1 และกลุ่มที่ 2 ตามลำดับ และนำแบบทดสอบเชื่อมโยง (Anchor Test) หรือฟอร์ม Z สอบนักเรียนทั้งสองกลุ่ม ซึ่งจะมีแบบจำลองของนักเรียน ดังนี้

$$X_1 = \lambda_x T_1 + \delta_x + \lambda_x^{1/2} E_x$$

$$Y_1 = \lambda_y T_1 + \delta_y + \lambda_y^{1/2} E_y$$

$$\text{และ} \quad Z_1 = \lambda_z T_i + \delta_z + \lambda_z^{1/2} E_z$$

ถ้าแทนค่า λ ด้วยจำนวนข้อสอบของแบบทดสอบแต่ละฟอร์มแล้ว แบบจำลองนี้จะมีลักษณะคล้ายคลึงกับแบบจำลองคะแนนจริงแบบมาตรฐานเดิม (Classical Parallel Model) โดยที่มีค่านิยามของ T และ δ เหมือนกับที่นิยามไว้ใน แบบจำลองคะแนนจริงสัมพันธ์ กล่าวคือ T_i กับ E ไม่มีความสัมพันธ์กัน และค่าเฉลี่ยของคะแนน E จากแบบทดสอบทุกฟอร์ม มีค่าเป็นศูนย์ แต่ค่าความแปรปรวนของคะแนน E จะเท่ากันทุกฉบับ (Homogeneous Variance) จากนักเรียนทั้งสองกลุ่ม ดังนั้น ความแปรปรวนของคะแนน E_x , E_y และ E_z มีค่าคงที่ และเขียนแทนด้วย $\sigma^2(E)$ โดยที่คะแนน E ของแต่ละฟอร์มมีความสัมพันธ์เป็นศูนย์ ดังนั้น λ จึงเป็นหน่วยวัด (Scale) ของทั้ง T และ E ตามแนวคิดของแองกอฟ (Angof, 1953) λ ก็คือ ค่าพารามิเตอร์ความยาวของแบบทดสอบ ที่เกิดจากคะแนนสอบ (Effective Test Length Parameters) นอกจากนั้นแบบจำลองคะแนนจริงสัมพันธ์ แบบมาตรฐานเดิมจะมีข้อบังคับ (Constraints) สองข้อเช่นเดียวกับแบบจำลองคะแนนจริงสัมพันธ์ ดังนั้น แบบจำลอง

คะแนนจริงสัมพันธ์แบบมาตรฐานเดิม จะมีข้อจำกัดมากกว่าแบบจำลองคะแนนจริงสัมพันธ์ ยิ่งไปกว่านั้น แบบจำลองคะแนนจริงสัมพันธ์แบบมาตรฐานเดิม จะมีความเหมาะสมพอดีจริงๆ (Perfect Fit) เฉพาะเมื่อมีแบบทดสอบเพียงสองฟอร์มเท่านั้น

การแสดงให้เห็นของสูตรตามแบบจำลองคะแนนจริงสัมพันธ์แบบมาตรฐานเดิม ก็เป็นไปตามสมการที่ (10), (11) และ (12) แต่ที่ λ_y / λ_z สามารถประมาณได้โดยไม่ต้องอาศัยการตอบข้อสอบจากแบบทดสอบ z ดังนี้

$$\begin{aligned} \frac{\sigma(Y_2, Z_2) + \sigma^2(Y_2)}{\sigma(Y_2, Z_2) + \sigma^2(Z_2)} &= \frac{\lambda_y \lambda_z \sigma^2(T_2) + \lambda_y^2 \sigma^2(T_2) + \lambda_y \sigma^2(E)}{\lambda_y \lambda_z \sigma^2(T_2) + \lambda_z^2 \sigma^2(T_2) + \lambda_z \sigma^2(E)} \\ &= \frac{\lambda_y (\lambda_y + \lambda_z) \sigma^2(T_2) + \sigma^2(E)}{\lambda_z (\lambda_y + \lambda_z) \sigma^2(T_2) + \sigma^2(E)} \\ &= \lambda_y / \lambda_z \end{aligned}$$

ดังนั้น

$$(13) \quad \lambda_y / \lambda_z = \frac{\sigma(Y_2, Z_2) + \sigma^2(Y_2)}{\sigma(Y_2, Z_2) + \sigma^2(Z_2)} = \frac{1 + b_{z2y2} \sigma^2(Y_2)}{1 + b_{y2z2} \sigma^2(Z_2)}$$

จากสมการ (13) กับสมการ (12) แสดงว่าตามแบบจำลองคะแนนจริงสัมพันธ์แบบมาตรฐานเดิม จะได้

$$(14) \quad p(Z_2) = \frac{1 + \sigma(Y_2, Z_2) / \sigma^2(Z_2)}{1 + \sigma^2(Y_2) / \sigma(Y_2, Z_2)} = \frac{1 + b_{y2z2}}{1 + 1/b_{z2y2}}$$

นำสมการ (13) แทนในสมการ (10) และ (11) แล้วนำผลแทนลงในสมการ (5) และ (6) แล้วสามารถแปลงคะแนน X เป็นคะแนน Y ได้ตามสมการ (4)

งานวิจัยที่เกี่ยวข้อง

งานวิจัยต่างประเทศ

งานวิจัยเกี่ยวกับการใช้แบบทดสอบหลัก ที่ใช้วิธีการเทียบคะแนนแบบเส้นตรง พบว่า มีจำนวนมาก โดยมีการศึกษานานงุ่มที่แตกต่างกันไป ตั้งแต่การเปรียบเทียบวิธีการเทียบคะแนนต่าง ๆ การศึกษาลักษณะของกลุ่มตัวอย่าง การเปรียบเทียบขนาดของกลุ่มตัวอย่างที่แตกต่างกัน การศึกษาความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน นอกจากนี้ยังมีการศึกษาเกี่ยวกับการใช้แบบทดสอบหลัก ของวิธีการเทียบคะแนนแบบเส้นตรง ที่ใช้แบบทดสอบหลักเป็นคู่ขนานกับแบบทดสอบที่นำมาเทียบคะแนนกัน และมีการพัฒนามาเป็นแบบทดสอบหลักมาเป็นแบบทดสอบที่ไม่จำเป็นต้องใช้ความเป็นคู่ขนานกัน คือถ้ามาสู่แบบทดสอบที่เป็นคะแนนจริงสัมพันธ์ มีแผนแบบการเก็บรวบรวมข้อมูลที่ใช้ แผนแบบของการใช้แบบทดสอบหลักที่มีประชากรแตกต่างกัน (CINEP) ซึ่งผู้วิจัยพอที่จะสรุปได้เป็นเรื่อง ๆ ดังนี้

ปีเตอร์เสน, มาร์โร และสจ๊วต (Petersen, Macro and Stewart.1982 : 71-185) ได้ศึกษาความเพียงพอของการเทียบคะแนน ที่เป็นการเทียบคะแนนแบบเส้นตรง จากวิธีต่าง ๆ กับวิธีแบบวัดตำแหน่งที่เท่ากันในสถานการณ์ต่าง ๆ โดยใช้ข้อมูลที่เป็นผลการสอบระดับชาติ 2 ครั้ง คือ ครั้งที่ 1 สอบในเดือนเมษายน ปี ค.ศ.1975 เป็นแบบทดสอบวัดความถนัด (Scholastic Aptitude Test or SAT) ซึ่งมี 3 ฉบับ คือ ee, fe และ fm ส่วนครั้งที่ 2 สอบในเดือนพฤศจิกายน ปี ค.ศ.1975 เป็นแบบทดสอบระดับมาตรฐานของการเขียนภาษาอังกฤษ (Test of Standard Written English or TSWE) ซึ่งมี 3 ฉบับ คือ eg, fg และ fo การศึกษาเป็นการศึกษาย้อนหลัง กลุ่มตัวอย่างจึงเป็นกลุ่มที่ถูกกำหนดขึ้นตามที่ต้องการ แล้วทำการสุ่มมาจำนวน 6 กลุ่ม แต่ละกลุ่มมีจำนวนผู้สอบ 4,731 คน แบ่งแต่ละกลุ่มออกเป็น 3 กลุ่มย่อยที่มาให้เข้ากันเลย ทั้งนี้ก็มีลักษณะต่างกัน ดังนี้ กลุ่มที่มีความสามารถคล้ายคลึงกัน (ใช้ค่าเฉลี่ยความสามารถทางภาษาเป็นเกณฑ์) และกลุ่มที่มีความสามารถแตกต่างกัน ดังนั้น จึงมีกลุ่มย่อยทั้งหมด 54 กลุ่ม การวางแผนการเทียบคะแนนใช้แบบทดสอบหลักในลักษณะต่าง ๆ กันกับแบบทดสอบที่ต้องการเทียบคะแนน ได้แก่ 1.) การใช้แบบทดสอบหลัก ที่เป็นแบบทดสอบเชื่อมโยงกับการใช้เป็นข้อสอบรวม 2.) ความแตกต่างในเนื้อหาที่มีลักษณะคล้ายคลึงกันกับไม่คล้ายคลึง 3.) ความแตกต่างเกี่ยวกับค่าความยาก ที่เป็นความยากใกล้เคียงกันกับที่แตกต่าง

กัน แล้วทำการเทียบคะแนนจากแบบทดสอบฉบับ SAT ส่วนที่เป็นภาษากับตัวมันเอง โดยใช้แบบทดสอบหลักในลักษณะต่าง ๆ เพื่อศึกษาระดับความคล้ายคลึงระหว่างแบบทดสอบ ที่ต้องการเทียบคะแนนกับแบบทดสอบหลัก และระดับความคล้ายคลึงระหว่างกลุ่มตัวอย่าง 2 กลุ่ม ที่ใช้ในการเทียบคะแนน ที่มีผลต่อการเทียบคะแนน และได้มีการเทียบคะแนนในระหว่างแบบทดสอบ SAT กับ TSWE โดยใช้ข้อสอบร่วม เพื่อศึกษาระดับความคล้ายคลึงระหว่างแบบทดสอบที่ต้องการเทียบคะแนนทั้งสองฉบับ และระดับความคล้ายคลึงระหว่างกลุ่มตัวอย่าง ที่ใช้เทียบคะแนนสองกลุ่มที่มีผลต่อการเทียบคะแนน โดยแบบทดสอบที่ใช้เทียบคะแนนมีลักษณะแตกต่างกันในด้านต่าง ๆ คือ ทั้งด้านเนื้อหา ความยาก และความยาวของแบบทดสอบ

ผลการวิจัยสรุปได้ว่าเมื่อพิจารณาที่ค่าความยากของแบบทดสอบ จะให้ผลดังนี้

1. ค่าความยากระหว่างแบบทดสอบหลักและแบบทดสอบ ที่ใช้เทียบคะแนนแตกต่างกันจะทำให้ความคลาดเคลื่อนในการเทียบคะแนน มากกว่าเมื่อมีความแตกต่างในเนื้อหา
2. การใช้แบบทดสอบหลัก พบว่า แบบทดสอบหลักที่สร้างให้มีลักษณะคล้ายคลึง กับแบบทดสอบที่จะใช้เทียบคะแนน เมื่อมีความคล้ายคลึงกันมาก หรืออาจเรียกว่าเป็นฉบับย่อส่วนแล้ว การเทียบคะแนนจะให้ผลดีที่สุด

โคป (Cope, 1987) ได้เปรียบเทียบการเทียบคะแนนแบบเส้นตรง 5 วิธี ระหว่างวิธีของแองกอฟ (Angoff, 1971) ตามแผนแบบการเก็บรวบรวมข้อมูลที่ V ซึ่งมี 3 วิธีย่อยกับวิธีของทักเกอร์ (Tucker) และลีวีน (Levine) ในกรณีที่แบบทดสอบมีความเชื่อมั่นเท่ากันและใช้แบบทดสอบหลัก ซึ่งแบบทดสอบที่ใช้เทียบคะแนน เป็นแบบทดสอบที่ใช้ในประกาศนียบัตรวิชาชีพ จำนวน 5 ฉบับ ทดสอบกับผู้สอบคนละกลุ่ม 5 กลุ่ม กลุ่มละ 1 ฉบับ ปีละ 2 ครั้ง คือแบบทดสอบฉบับ B , D และ E สอบในฤดูใบไม้ผลิ ส่วนแบบทดสอบฉบับ A กับฉบับ C สอบในฤดูฝน การเปรียบเทียบวิธีการเทียบคะแนนมี 2 แบบ คือ การเปรียบเทียบผลการเทียบคะแนนระหว่างแบบทดสอบแต่ละคู่ ภายใต้วีธีการแบบเส้นตรงที่แตกต่างกัน 5 วิธี ทั้งนี้โดยยังไม่เปรียบเทียบความถูกต้องของวิธีการต่าง ๆ เพราะว่ามีมาตราของคะแนนดิบของแบบทดสอบ 2 ฉบับ ถูกคาดว่าจะต่างกัน อีกแบบหนึ่ง คือ การเปรียบเทียบผลการเทียบคะแนนแบบสายโซ่ ซึ่งเป็นการเปรียบเทียบคะแนนแบบทดสอบแต่ละฉบับกับตัวเอง โดยผ่านแบบทดสอบฉบับอื่นที่เลือกไว้ ทั้งนี้เพื่อทำการเปรียบเทียบความถูกต้องของวิธีการเหล่านั้นด้วย การวัดความคลาดเคลื่อนของการเทียบคะแนน

าซ์สถิติ 2 ตัว คือ ค่าประมาณความลำเอียง (Estimated Bias) และค่าประมาณที่ถ่วงน้ำหนักแล้ว ของรากที่สองของค่าเฉลี่ยยกกำลังสองของความคลาดเคลื่อน ในการเทียบคะแนน (Estimated Weighted Root Squared Error or RMSE) ซึ่งใช้เปรียบเทียบระหว่างวิธีที่แตกต่างกัน

ผลการวิจัยสรุปได้ ว่าในสถานการณ์ที่ศึกษาภายใต้แผนแบบที่ V ตามวิธีของแองกอฟาห้คะแนนดิบที่สมมูลกันใกล้เคียงกับวิธีการของทักเกอร์และลีไวน์ โดยการเทียบคะแนนแบบสายรั้งทั้ง 5 วิธี ให้ค่าประมาณความลำเอียง และรากที่สองของค่าเฉลี่ยยกกำลังสองของความคลาดเคลื่อนที่ใกล้เคียงกัน ส่วนการเทียบคะแนนแต่ละคู่ของแบบทดสอบ พบว่า คะแนนดิบที่เท่าเทียมกันจากแผนแบบที่ V มีค่าอยู่ระหว่างคะแนนจากวิธีการของทักเกอร์และลีไวน์ ซึ่งผู้วิจัยสรุปว่า ผลการวิจัยนี้สนับสนุนวิธีการเทียบคะแนนแบบเส้นตรง ตามแผนแบบที่ V โดยมีข้อตกลงที่มีข้อจำกัดน้อยกว่าวิธีของทักเกอร์และลีไวน์ ดังนั้น ในทางปฏิบัติควรใช้วิธีการตามแผนแบบที่ V ซึ่งมีวิธีการย่อย 3 วิธี โดยเลือกวิธีการใดวิธีการหนึ่งใน 3 วิธีก็ได้

โคเลน และวิทนี (Kolen and Whitney, 1982) ได้เปรียบเทียบความเพียงพอของวิธีเทียบคะแนน สี่วิธี คือ วิธีแบบเส้นตรง วิธีแบบใช้ตำแหน่งที่เท่ากัน วิธีแบบใช้ทฤษฎีการตอบข้อสอบชนิดหนึ่งพารามิเตอร์ (ราสซ์โมเดล) และวิธีใช้ทฤษฎีการตอบข้อสอบชนิดสามพารามิเตอร์ (โรจิสติก โมเดล) โดยชักกลุ่มตัวอย่างขนาดเล็ก กลุ่มละประมาณ 200 คน ที่สุ่มมาจากประชากรที่มีงานมากถึง 800,000 คน ทำการทดสอบโดยใช้แบบทดสอบ GED (Tests of General Educational Development) ซึ่งเป็นแบบทดสอบมาตรฐานวัดผลสัมฤทธิ์ทางการเรียน 5 วิชา ได้แก่ ทักษะการเขียน (80 ข้อ) สังคมศึกษา (60 ข้อ) วิทยาศาสตร์ (60 ข้อ) ทักษะการอ่าน (40 ข้อ) และคณิตศาสตร์ (50 ข้อ) โดยแต่ละวิชานี้มีเนื้อหาแตกต่างกันแบ่งเป็น 12 ฉบับ การทดสอบนี้เพื่อนำไปใช้ตัดสินสำหรับการให้ประกาศนียบัตรแก่ผู้สำเร็จการศึกษา และเทียบคะแนนระหว่างกลุ่มผู้เข้าสอบในแต่ละวิชา โดยใช้แบบทดสอบหลัก 11 ฉบับ ที่ได้มาจากแบบทดสอบแต่ละวิชา ผู้เข้าสอบจะต้องสอบวิชาเฉพาะ 1 วิชา กับแบบทดสอบหลัก 1 ฉบับ และทุกกลุ่ม จะต้องสอบแบบทดสอบฉบับที่ 12 ซึ่งนำมาจัดทำเป็นแบบทดสอบหลัก การทดสอบนี้ใช้วิธีจัดให้เป็นกลุ่มสมดุล (Counterbalance) และเทียบคะแนนแต่ละคู่ของแบบทดสอบในแนวนอนด้วยวิธีเทียบคะแนน สี่วิธี และตรวจสอบโดยชักกลุ่มตามผลจำนวน 20 คน ที่สุ่มมาจากผู้เข้าสอบแต่ละ

วิชา แล้วจึงหาดัชนีเปรียบเทียบเบอร์เซนต์ไทล์

ผลการวิจัยพบว่า วิธีการเทียบคะแนนแบบเส้นตรงให้ความเที่ยงพอมากที่สุด โดยดูจาก ช่วงค่าเฉลี่ย และผู้วิจัยสรุปว่า ปัจจัยที่ช่วยให้การเทียบคะแนนมีความเที่ยงพอนั้น ขึ้นอยู่กับปัจจัย หลายประการ ได้แก่ คุณลักษณะของแบบทดสอบ การใช้แผนแบบรวบรวมข้อมูล (Equated Design) และขนาดกลุ่มตัวอย่าง เป็นต้น

คลิน และจาโจรา (Klein and Jarjoura.1985) ได้ศึกษาถึง ความสำคัญของ เนื้อหาแบบทดสอบหลัก ที่ใช้เป็นตัวแทนของแบบทดสอบที่นำมาเทียบคะแนน โดยใช้ข้อสอบร่วมกับ แบบทดสอบที่นำมาเทียบคะแนนให้มีเนื้อหาที่เป็นตัวแทน และเนื้อหาที่ไม่เป็นตัวแทนรวมเข้าใบ ใน แบบทดสอบที่นำมาเทียบคะแนน แต่ให้มีความยาวของข้อสอบรวมยาวกว่า และจำนวนข้อเท่ากัน ทั้ง สองฉบับ คือ 60 ข้อ และศึกษาเฉพาะความยาวของส่วนที่เพิ่มเข้าใบ ในแบบทดสอบฉบับที่ 1ให้มีเนื้อหาข้อสอบรวมไม่เป็นตัวแทน คือ ให้มีจำนวนเพิ่มจาก 60 ข้อ มาเป็น 105 กับ 101 ข้อ และแบ่งเนื้อหาเป็น 5 ส่วน ในลักษณะจำนวนข้อแตกต่างกัน นำไปทดสอบกับกลุ่มตัวอย่างเดียวกัน 3 กลุ่ม แล้วเปรียบเทียบเป็นรายคู่ระหว่างแบบทดสอบฉบับที่มีข้อสอบรวมที่เป็นตัวแทนของเนื้อหา กับฉบับที่ไม่เป็นตัวแทนของเนื้อหา

ผลการศึกษาพบว่า เมื่อเขียนเป็นกราฟแสดงโปรไฟล์ของค่าเฉลี่ยจากคะแนนในแต่ละ เนื้อหา 5 ส่วน พบว่า มีค่าเฉลี่ยของคะแนนแตกต่างกันในแต่ละคู่ ผู้วิจัยสรุปว่าเป็นความคลาดเคลื่อนของคะแนน ที่เกิดจากความลำเอียง

คุก, อิกนอร์ และทัฟท์ (Cook and Pertersen.1987 : 229 ; citting Cook, Eignor and Taft.1985) ได้ศึกษาการใช้แบบทดสอบหลัก ที่เป็นข้อสอบรวม 3 พอร์ม ที่แตกต่างกัน 3 พอร์ม โดยใช้วิชาชีววิทยามาจัดทำเป็นแบบทดสอบที่มีข้อสอบรวมจำนวน 58 ข้อ และข้อสอบที่นำมาเทียบคะแนนกันได้มาจากพอร์มเก่า (สอบในฤดูฝน) กับข้อสอบพอร์มใหม่ (สอบ ในฤดูใบไม้ผลิ กับฤดูฝน) เป็นจำนวน 3 ฉบับ ทำการทดสอบกับกลุ่มตัวอย่างที่มาจากระชากร เดียวกัน แต่สอบในเวลาที่แตกต่างกัน กลุ่มที่สอบในฤดูใบไม้ผลิ จะมีความรู้ในวิชาชีววิทยาครบถ้วนเพราะเรียนมาครบเทอม จึงเป็นกลุ่มย่อยที่มีความสามารถแตกต่างกัน นำผลการสอบมา เทียบคะแนนกัน สามวิธี พบว่า ค่าเฉลี่ยของกลุ่มที่สอบในฤดูใบไม้ผลิ แตกต่างจากค่าเฉลี่ยของ กลุ่มที่สอบในฤดูฝนทั้ง สองกลุ่ม โดยแบบทดสอบทั้ง 3 ฉบับ มีลักษณะเนื้อหาของแบบทดสอบหลักกับ

แบบทดสอบ ที่นำมาเทียบคะแนนกันเป็นคู่ขนานกัน ที่มีค่าสหสัมพันธ์กันสูงและเท่าเทียมกัน ผลที่เกิดขึ้นนี้ มาจากความสัมพันธ์กับระยะเวลาของการฝึกปฏิบัติ ซึ่งมีผลต่อการสอบในวิชาเดียวกัน แต่วัดคุณลักษณะที่ต่างโครงสร้างกัน ฉะนั้น การเทียบคะแนนสอบจึงต้องขึ้นกับกลุ่มผู้สอบด้วย

วูดดรัฟ (Woodruff.1986) ได้ศึกษาวิธีเทียบคะแนนแบบเส้นตรง โดยใช้แผนแบบของ CINEP และแบบจำลองคะแนนจริงสัมพันธ์ โดยใช้กลุ่มตัวอย่างจำนวน 4,000 คน ทดสอบกับแบบทดสอบ 5 ฉบับ ที่เป็นฟอร์มใหม่ มีความยาวที่แตกต่างกันคือตั้งแต่จำนวน 25 ถึง 115 ข้อ และใช้แบบทดสอบที่เป็นฟอร์มเก่า ร่วมกับแบบทดสอบเชื่อมโยง มาทดสอบกับกลุ่มตัวอย่าง จำนวน 4,000 คน นำคะแนนที่ได้มาศึกษาวิธีเทียบคะแนน สองวิธี

ผลการศึกษาพบว่า มีความแตกต่างของกลุ่มผู้สอบเล็กน้อย โดยดูจากความเบี่ยงเบนมาตรฐานจากแบบทดสอบ 5 ฉบับ ที่ใช้แบบทดสอบหลักและใช้วิธีการเทียบคะแนน สองวิธี ให้ความเที่ยงพอในการปฏิบัติ แม้ว่าแบบทดสอบทั้ง 5 ฉบับ จะวัดคุณลักษณะที่ต่างกัน

แมคแคน (MacCann.1989) ได้ศึกษาการเปรียบเทียบวิธีเทียบคะแนน สองวิธีที่ใช้ข้อตกลงว่ามีความเชื่อมั่นเท่าเทียมกัน และใช้แบบทดสอบที่เป็นคะแนนจริงสัมพันธ์ ทำการทดสอบกับกลุ่มตัวอย่างทั้งที่เป็นกลุ่มคล้ายคลึงกัน และกลุ่มที่มีความแตกต่างกัน โดยมีขนาดกลุ่มตัวอย่างที่แตกต่างกันในแต่ละวิชาที่สอบ และแต่ละกลุ่มที่ใช้เป็นกลุ่มเทียบคะแนน ตั้งแต่จำนวน 326 คน ถึง 5,610 คน ใช้แบบทดสอบ 5 วิชา แบ่งเป็นส่วยย่อยที่แตกต่างกันในแต่ละวิชา และใช้แบบทดสอบเชื่อมโยงที่มีความยาว คิดเป็นสัดส่วนแตกต่างกัน ทั้งที่มากกว่าและน้อยกว่าแบบทดสอบที่นำมาใช้เทียบคะแนน แล้วจึงนำมาเปรียบเทียบความคลาดเคลื่อนมาตรฐานของแต่ละวิธี (สี่วิธี) คือ วิธีของแมคแคน สองวิธี และวิธีของทักเกอร์กับวิธีของลิโวน์ พบว่าวิธีที่ 1 ให้ความคลาดเคลื่อนกับวิธีของทักเกอร์ ส่วนวิธีที่ 2 ให้ความคลาดเคลื่อนกับวิธีของลิโวน์ และทั้ง สี่วิธี ให้ความคลาดเคลื่อนมาตรฐานในระดับที่น่าพอใจเป็นส่วนมาก ผู้วิจัยสรุปว่า ข้อดีของวิธีที่ 1 กับวิธีที่ 2 นั้นไม่จำเป็นต้องให้ข้อตกลงที่ว่า การกระจายของคะแนนจากแบบทดสอบ สองฉบับ ที่นำมาเทียบคะแนนกัน จะต้องมีการแจกแจงเป็นโค้งปกติ

งานวิจัยในประเทศ

1. งานวิจัยที่ศึกษาวิธีการเทียบคะแนนที่ไม่ได้ใช้แบบทดสอบหลัก ได้แก่

เยาวดี ราชชัยกุล (2518) ได้ศึกษาการเทียบคะแนนแบบเส้นตรง โดยเปรียบเทียบคะแนนที่ได้จากผลการสอบในประเทศไทย ที่ เป็นนักเรียนในระบบโรงเรียนกับนักศึกษาที่มาจากระบบการศึกษานอกระบบเรียน ซึ่งเป็นกลุ่มตัวอย่างที่แตกต่างกันทั้ง 2 กลุ่ม แล้วใช้แผนแบบการเก็บรวบรวมข้อมูล แบบที่ I ของแองกอฟ และทำการเทียบคะแนนจากแบบทดสอบแต่ละฉบับไปสู่มาตราส่วน (Common Scale) ที่มีช่วงคะแนน 1-175 มีค่าเฉลี่ยเท่ากับ 100 ความเบี่ยงเบนมาตรฐานเท่ากับ 25 ผลการวิจัยพบว่า เกิดความไม่ยุติธรรมแก่ผู้เข้าสอบทั้งสองกลุ่ม ผู้วิจัยจึงได้ให้ข้อเสนอแนะไว้ว่า ควรเทียบคะแนนโดยใช้แผนแบบการเก็บรวบรวมข้อมูล แบบที่ IV ตามที่แองกอฟเสนอไว้ คือ ให้ใช้แบบทดสอบหลัก (สุจินดา ผ่องอักษร.2533 : 61 ; อ้างอิงมาจากเยาวดี ราชชัยกุล.2518 : 87-94)

สุจินดา ผ่องอักษร (2533 : บทคัดย่อ) ได้ศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนในกลุ่มทักษะ ของนักเรียนชั้นประถมศึกษาปีที่ 6 ในช่วงระยะเวลา 3 ปี โดยใช้แผนแบบการเทียบคะแนนของราสส์ และเทียบคะแนนจากการนำคะแนนจาก แบบทดสอบต่างฉบับในแต่ละวิชามาอยู่ขนาดเดียวกัน แล้วจึงเปรียบเทียบกันโดยตรง รวมทั้งได้กำหนดคะแนนจุดตัดที่ใช้เป็นเกณฑ์ผ่าน-ไม่ผ่าน

2. งานวิจัยที่ใช้แบบทดสอบหลักมาเทียบคะแนน ได้แก่

ฐิทิพ พงษ์สมบูรณ์ (2528 : 112-118) ได้ทำการเปรียบเทียบประสิทธิภาพและความคงที่ของการเทียบคะแนน 3 วิธี คือ 1.) การเทียบคะแนนแบบเส้นตรง 2.) การเทียบคะแนนแบบใช้ตำแหน่งที่เท่ากัน 3.) การเทียบคะแนนโดยใช้ทั้งลักษณะการตอบข้อสอบ เมื่อใช้แผนแบบที่เป็นแบบทดสอบหลักกับผู้สอบร่วม เปรียบเทียบประสิทธิภาพของการเทียบคะแนน โดยใช้การพิจารณาจากความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ส่วนการเปรียบเทียบความคงที่ของวิธีการเทียบคะแนน (ความไม่แปรผันไปตามกลุ่มตัวอย่าง ความสมมาตร และความเสมอภาค) นั้น ทำการทดสอบนัยสำคัญของความแตกต่าง โดยใช้สถิติทดสอบโคลมโกรอฟ-สมิรโนฟ (Kolmogorov-Smirnov Two Sample Test) กลุ่มตัวอย่างที่ใช้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 ทั่วประเทศ ที่เข้าสอบตามโครงการตรวจสอบคุณภาพการศึกษาของกรมวิชาการ กระทรวง

ศึกษาธิการ ซึ่งสุ่มมาจำนวน 3,721 คน แบ่งเป็นกลุ่มย่อย 6 กลุ่ม กลุ่มละ 620 คน แต่ละกลุ่มมีความสามารถใกล้เคียงกัน แบบทดสอบที่มีข้อสอบทั้งหมด 60 ข้อ จัดแบ่งเป็น 2 ฉบับ ฉบับละ 38 ข้อ ที่มีข้อสอบรวมจำนวน 20 ข้อ แบบทดสอบทั้งสองฉบับนี้มีความยากใกล้เคียงกัน ผลการวิจัยพบว่า ประสิทธิภาพของการเทียบคะแนนระหว่างการใช้ที่เป็นข้อสอบรวม กับการใช้ข้อสอบรวมด้วย วิธีการที่ต่างกัน 3 วิธี ไม่แตกต่างกัน ส่วนผลการเปรียบเทียบความคงที่ของวิธีการเทียบคะแนน โดยการใช้ข้อสอบรวม ให้ผลเหมือนกับการใช้แผนแบบที่ใช้ข้อสอบรวมทั้งสามวิธี

ภาวณิ ศรีสุววัฒนานันท์ (2528: 155-170) ได้เปรียบเทียบผลของการใช้ข้อสอบรวม กับวิธีการเทียบคะแนนที่ต่างกัน 3 แบบ คือ แบบใช้ตำแหน่งที่เท่าเทียมกัน แบบเส้นตรง และแบบอิงทฤษฎีการตอบข้อสอบชนิดสามพารามิเตอร์ โดยใช้ข้อสอบรวมที่ต่างกัน 3 ขนาด คือ ขนาดร้อยละ 20 (7 ข้อ) ขนาดร้อยละ 40 (14 ข้อ) และขนาดร้อยละ 60 (21 ข้อ) โดยมีข้อสอบทั้งหมด 100 ข้อ จัดแบ่งเป็น 2 ฉบับ ฉบับละ 35 ข้อ ซึ่งทั้งสองฉบับมีความยากใกล้เคียงกัน ที่เหลือจัดเป็นข้อสอบรวมขนาดต่าง ๆ สำหรับกลุ่มตัวอย่างแยกเป็น 2 กรณี คือ กรณีใช้แบบทดสอบคัดเลือก 2 กลุ่ม ๆ ละ 1,500 คน และกลุ่มตรวจสอบผลอีก 1 กลุ่ม จำนวน 1,500 คน อีกกรณีหนึ่ง คือ กรณีใช้แบบวัดผลสัมฤทธิ์ ซึ่งมีจำนวนกลุ่มตัวอย่างเช่นเดียวกับกรณีแรก ใช้การประเมินผลการเทียบคะแนน 2 ลักษณะ คือ

1. เปรียบเทียบปริมาณความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ด้วยวิธีการที่ต่างกันของแบบทดสอบแต่ละกรณี

2. ตรวจสอบความเพียงพอ ของวิธีเทียบคะแนนแต่ละวิธี โดยใช้ค่าดัชนีเปรียบเทียบความแตกต่าง (index C) ที่ได้จากกลุ่มตรวจสอบผล

ผลการวิจัยพบว่า การใช้ข้อสอบรวมที่ยาวกว่า ให้ประสิทธิภาพของการเทียบคะแนนที่สูงกว่า (หรือมีความคลาดเคลื่อนมาตรฐานในการการเทียบคะแนนน้อยกว่า) และให้ผลที่น่าพอใจมากกว่าทั้ง สองกรณีของการใช้แบบทดสอบเทียบคะแนน

อาภรณ์ กาญจนกิจโรจน์ (2531 : 66-67) ได้ศึกษาการสร้างแบบทดสอบและตารางเทียบระดับคะแนนตามแนวตั้ง ในวิชาคณิตศาสตร์ เรื่องสมการและอสมการโดยใช้วิธีเทียบคะแนนแบบราชสีห์ สำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น กลุ่มตัวอย่างที่ใช้เทียบคะแนน เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 จำนวนระดับชั้นปีละ 490, 482 และ 415 คน ตามลำดับ ผู้วิจัย

ได้นำแบบทดสอบที่สร้างขึ้นระดับชั้นละ 90 ข้อ ไปทดลองกับกลุ่มตัวอย่างอื่น แล้วคัดเลือกข้อสอบที่มีคุณภาพ เรียงตามค่าความยากที่ต้องการระดับชั้นละ 40, 30 และ 30 ข้อ โดยแต่ละชั้นจะมีข้อสอบรวมระหว่างชั้นอยู่ 10 ข้อ หลังจากนั้น นำไปทดสอบกับกลุ่มตัวอย่างที่ใช้เทียบคะแนน และสร้างตารางเทียบระดับคะแนน เพื่อใช้ศึกษาพัฒนาการของนักเรียนกลุ่มเดิมที่ผ่านชั้นไปเรียนในชั้นสูงขึ้นว่า นักเรียนมีความก้าวหน้าเป็นไปตามปกติ หรือเบี่ยงเบนไปจากปกติ

สุรัตน์ ขวัญบุญจันทร์ (2531 : บทคัดย่อ) ได้ศึกษาการสร้างแบบทดสอบ และตารางเทียบคะแนนแนวนอน วิชาวิทยาศาสตร์ เรื่องพลังงานและการเปลี่ยนแปลง ที่วิเคราะห์ด้วยราชส์โรมเดล กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ปีการศึกษา 2531 ในจังหวัดชัยภูมิ จำนวน 1,391 คน โดยใช้วิธีการสุ่มหลายขั้นตอน ใช้แบบทดสอบ 2 ฉบับ ฉบับละ 60 ข้อ มีข้อสอบรวม 10 ข้อ ผลการศึกษาพบว่า แบบทดสอบ 2 ฉบับ มีค่าความยากตั้งแต่ -1.38 ถึง 2.25 และ -1.63 ถึง 1.17 มีค่า t -test ตั้งแต่ -1.70 ถึง 1.71 และ -1.40 ถึง 2.09 ความเชื่อมั่นของแบบทดสอบที่คำนวณจาก KR-20 เท่ากับ $.76$ และ $.78$ ตามลำดับ เมื่อนำคะแนนดิบจากแบบทดสอบ 2 ฉบับ มาสร้างตารางเทียบคะแนน ภายใต้อำนาจที่สามารถที่เท่ากัน ได้คะแนนดิบแต่ละค่าเท่ากันทุกคะแนน และได้ค่าดัชนีความเพียงพอของความคลาดเคลื่อนในการเทียบคะแนนเท่ากับ $.1019$ ซึ่งอยู่ในระดับที่น่าพอใจมาก

เรวดี อินทสระ (2530 : บทคัดย่อ) ได้ศึกษาการเปรียบเทียบความคลาดเคลื่อน และระดับความเพียงพอของการเทียบคะแนน ระหว่างรูปแบบอิงทฤษฎีการตอบข้อสอบ กับรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ โดยใช้กลุ่มทนายผล เครื่องมือที่ใช้ในการวิจัย คือแบบทดสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 เรื่องอัตราส่วนร้อยละ ที่ผู้วิจัยสร้างขึ้นจากการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบข้อสอบ โดยใช้โปรแกรมโรจิส 5 แล้ว จัดข้อสอบออกเป็น 2 ฉบับ ๆ ละ 30 ข้อ และแบบทดสอบหลักอีกฉบับละ 15 ข้อ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2529 โรงเรียนสังกัดกรมสามัญศึกษา จังหวัดพัทลุง จำนวน 2,823 คน โดยให้กลุ่มตัวอย่างทำแบบทดสอบทั้งสองฉบับ แล้วสุ่มมากลุ่มละ 1 ฉบับ เพื่อนำคะแนนมาเทียบมาตรฐานทั้งสองรูปแบบ และคะแนนอีกฉบับจะเป็นคะแนนแทนผล ผลการศึกษา พบว่า ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ของรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ มีคะแนนความคลาดเคลื่อนแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ $.01$

ความเพียงพอของความคลาดเคลื่อนการเทียบคะแนน รูปแบบอิงทฤษฎีการตอบข้อสอบ อยู่ในระดับ
น่าพอใจ และการเทียบคะแนนรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ อยู่ในระดับ
ปานกลาง ซึ่งมีค่าเท่ากับ 0.2356 และ 0.8851 ตามลำดับ

สมมติฐานในการศึกษาค้นคว้า

1. คะแนนแปลง ที่ได้จากการสอบด้วยวิธีเทียบคะแนนแบบเส้นตรง สี่วิธี แตกต่างกัน
2. คะแนนแปลง ที่ได้จากการสอบด้วยวิธีเทียบคะแนนแบบเส้นตรง ที่ใช้แบบทดสอบ
เชื่อมโยง กับที่ข้อสอบรวม แตกต่างกัน
3. คะแนนแปลง ของแบบทดสอบสำหรับเทียบคะแนน ที่มีแบบทดสอบหลักขนาดความยาว
ต่างกัน แตกต่างกัน
 - 3.1. คะแนนแปลง ของแบบทดสอบสำหรับเทียบคะแนน ที่มีแบบทดสอบหลักแบบ
เชื่อมโยงขนาดความยาวต่างกัน แตกต่างกัน
 - 3.2. คะแนนแปลง ของแบบทดสอบสำหรับเทียบคะแนน ที่มีข้อสอบรวมขนาดความยาว
ต่างกัน แตกต่างกัน
4. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากรีการเทียบคะแนน
แบบเส้นตรง สี่วิธี แตกต่างกัน

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักศึกษาพยาบาล ชั้นปีที่ 1 และ ชั้นปีที่ 2 หลักสูตรพยาบาลศาสตร์ และพยาบาลศาสตรระดับต้น จากวิทยาลัยพยาบาลต่าง ๆ สังกัดสถาบันพัฒนากำลังคนด้านสาธารณสุข สำนักงานปลัดกระทรวงสาธารณสุข กระทรวงสาธารณสุข ในเขตภาคกลาง ปีการศึกษา 2538 จำนวน 11 แห่ง

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักศึกษาพยาบาล ชั้นปีที่ 1 และชั้นปีที่ 2 หลักสูตรพยาบาลศาสตร์ และพยาบาลศาสตรระดับต้น จากวิทยาลัยพยาบาล สังกัดสถาบันพัฒนากำลังคนด้านสาธารณสุข สำนักงานปลัดกระทรวงสาธารณสุข กระทรวงสาธารณสุข ในเขตภาคกลาง ปีการศึกษา 2538 จำนวน 3 แห่ง มีจำนวนนักศึกษาที่เป็นกลุ่มตัวอย่าง ทั้งหมด 864 คน โดยคัดเลือกกลุ่มตัวอย่างด้วยวิธีสุ่มอย่างง่าย (Simple Random Sampling) สองครั้ง คือ

1. สุ่มวิทยาลัยพยาบาล โดยใช้สถาบันการศึกษาเป็นหน่วยการสุ่ม ได้วิทยาลัยพยาบาลบรมราชชนนีกทม. วิทยาลัยพยาบาลบรมราชชนนีนครบุรี วิทยาลัยพยาบาลบรมราชชนนีสระบุรี มีจำนวนนักศึกษารวมทั้งสามสถาบัน 1,183 คน
2. สุ่มนักศึกษาพยาบาล จากสถาบันการศึกษาดังกล่าวแต่ละแห่ง ให้มีจำนวนนักศึกษาเท่า ๆ กัน ในแต่ละชั้นปีและของแต่ละหลักสูตร เพื่อใช้เป็นกลุ่มสำหรับเทียบคะแนน ได้กลุ่มตัวอย่าง 12 กลุ่ม กลุ่มละ 72 คน รวมสถาบันละ 288 คน เป็นกลุ่มตัวอย่างทั้งหมด 864 คน ดังตาราง 1

ตาราง 1 จำนวนนักศึกษาพยาบาลของแต่ละชั้นปี และแต่ละหลักสูตรของวิทยาลัยพยาบาลที่เข้า
เป็นกลุ่มตัวอย่างสำหรับการสอบเทียบคะแนน

วิทยาลัยพยาบาล บรมราชชนนี	จำนวนนักศึกษา				รวม
	หลักสูตรพยาบาลศาสตร์		พยาบาลศาสตรระดับต้น		
	ปีที่ 1	ปีที่ 2	ปีที่ 1	ปีที่ 2	
1. สระบุรี	72	72	72	72	288
2. ชลบุรี	72	72	72	72	288
3. กรุงเทพ	72	72	72	72	288
รวม					864

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบวัดการคิดเอกลักษ์แบบสัมพันธ์ ที่มีเนื้อหาเป็นภาพกับภาษาละกัน ซึ่งสร้างโดยอาศัยโครงสร้างทางทฤษฎีวัดความสามารถทางสมองของกิลฟอร์ด แบ่งเป็นแบบทดสอบ 2 ประเภท คือ

1. แบบทดสอบสำหรับเทียบคะแนน มี 2 พอร์ม ได้แก่

1.1 แบบทดสอบพอร์ม A จำนวน 40 ข้อ ประกอบด้วยข้อสอบที่มีเนื้อหาเป็นภาพ 24 ข้อ และเนื้อหาที่เป็นภาษา 16 ข้อ

1.2 แบบทดสอบพอร์ม B จำนวน 50 ข้อ ประกอบด้วยข้อสอบที่มีเนื้อหาเป็นภาพ 20 ข้อ และเนื้อหาที่เป็นภาษา 30 ข้อ

2. แบบทดสอบหลัก (Anchor test) หรือแบบทดสอบพอร์ม C มีจำนวน 3 ฉบับ ได้แก่

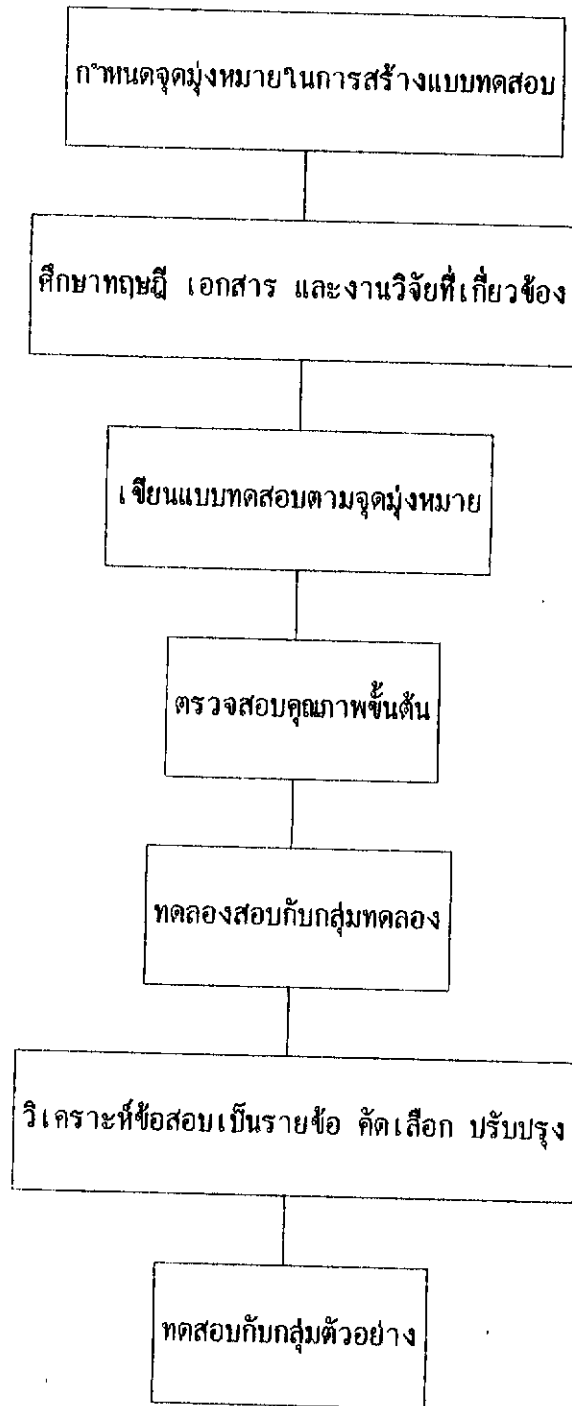
ฉบับที่ 1 (พอร์ม C₁) จำนวน 16 ข้อ ประกอบด้วยข้อสอบที่มีเนื้อหาเป็นภาพ 8 ข้อ และเนื้อหาที่เป็นภาษา 8 ข้อ

ฉบับที่ 2 (พอร์ม C₂) จำนวน 12 ข้อ ประกอบด้วยข้อสอบที่มีเนื้อหาเป็นภาพ 6 ข้อ และเนื้อหาที่เป็นภาษา 6 ข้อ

ฉบับที่ 3 (พอร์ม C₃) จำนวน 8 ข้อ ประกอบด้วยข้อสอบที่มีเนื้อหาเป็นภาพ 4 ข้อ และเนื้อหาที่เป็นภาษา 4 ข้อ

วิธีดำเนินการสร้างเครื่องมือ

การดำเนินการสร้างเครื่องมือสำหรับใช้ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยได้ดำเนินการสร้างตามขั้นตอน ดังภาพประกอบ 1



ภาพประกอบ 1 แสดงลำดับขั้นตอนดำเนินการสร้างเครื่องมือ

จากภาพประกอบ 1 เป็นลำดับขั้นตอนการสร้างเครื่องมือ ซึ่งมีรายละเอียดในการปฏิบัติ ดังนี้

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบ การสร้างแบบทดสอบครั้งนี้มีจุดมุ่งหมาย วัดความสามารถทางสติปัญญา โดยใช้ทฤษฎีโครงสร้างทางสติปัญญาของกิลฟอร์ด ด้านการคิด เอกนัยแบบสัมพันธ์ (Convergent Relation) ที่มีเนื้อหาเป็นภาพกับภาษามาสอบวัดกับนักศึกษาพยาบาลชั้นปีที่ 1 และชั้นปีที่ 2 ของวิทยาลัยพยาบาล สังกัดสถาบันพัฒนากำลังคนด้านสาธารณสุข กระทรวงสาธารณสุข ในเขตภาคกลาง
2. ศึกษาทฤษฎี เอกสาร และงานวิจัยที่เกี่ยวข้องกับ การสร้างแบบทดสอบวัดการคิด เอกนัยแบบสัมพันธ์ ซึ่งเป็นการวัดความสามารถของสติปัญญาด้านหนึ่งของกิลฟอร์ด
3. เขียนข้อสอบตามลักษณะของการคิดเอกนัยแบบสัมพันธ์ ที่มีเนื้อหาเป็นภาพกับภาษา จัดเป็นแบบทดสอบวัดการคิดเอกนัยแบบสัมพันธ์ฉบับทดลองจำนวน 3 ฉบับ ได้แก่
 - ฉบับที่ 1 จำนวน 60 ข้อ ที่มีเนื้อหาเป็นภาพต่อภาษา คิดเป็นสัดส่วน 3:2
 - ฉบับที่ 2 จำนวน 70 ข้อ ที่มีเนื้อหาเป็นภาพต่อภาษา คิดเป็นสัดส่วน 2:3
 - ฉบับที่ 3 จำนวน 30 ข้อ ที่มีเนื้อหาเป็นภาพต่อภาษา คิดเป็นสัดส่วน 1:1
 นำแบบทดสอบฉบับทดลองนี้ ไปทดลองสอบกับกลุ่มตัวอย่าง ที่นำชกลุ่มตัวอย่างที่ใช้ในการวิจัย แต่มีลักษณะคล้ายคลึงกับกลุ่มตัวอย่างที่ใช้ในการวิจัย จำนวน 110 คน
4. นำผลการสอบมาตรวจนับคะแนน โดยให้คะแนนเป็น 0 - 1 ดังนี้
 - ให้ 0 คะแนน สำหรับคำตอบที่ผิด ไม่ตอบ หรือตอบมากกว่า 1 ตัวเลือก
 - ให้ 1 คะแนน สำหรับคำตอบที่ถูกต้อง
5. นำข้อมูลที่ได้จากการตรวจนับคะแนน มาวิเคราะห์ข้อสอบรายข้อ หาค่าความยาก และค่าอำนาจจำแนก โดยใช้สูตรอย่างง่ายคำนวณ แล้วคัดเลือกข้อสอบที่มีค่าความยาก ตั้งแต่
6. จัดทำเป็นแบบทดสอบวัดการคิดเอกนัยแบบสัมพันธ์ จำนวน 3 ฉบับ ตามขั้นตอนดังนี้
 - 6.1. คัดเลือกข้อสอบที่มีคุณภาพตามเกณฑ์ จากแบบทดสอบฉบับทดลอง ฉบับที่ 1 มาจำนวน 40 ข้อ จัดทำเป็นแบบทดสอบวัดการคิดเอกนัยแบบสัมพันธ์ พอร์ม A ซึ่งจัดทำเป็นแบบทดสอบสำหรับเทียบคะแนน ที่มีเนื้อหาเป็นภาพกับภาษา จำนวน 24 ข้อกับ 16 ข้อ ตาม

ลำดับ (คิดเป็นสัดส่วน 3:2)

6.2. คัดเลือกข้อสอบที่มีคุณภาพตามเกณฑ์ จากแบบทดสอบฉบับทดลอง ฉบับที่ 2 มาจำนวน 50 ข้อ จัดทำเป็นแบบทดสอบวัดการคิดเอกลัคนัยแบบสัมพันธ์ พอร์ม B ซึ่งจัดทำเป็นแบบทดสอบสำหรับเทียบคะแนน เช่นกัน มีเนื้อหาเป็นภาพกับภาษา จำนวน 20 ข้อกับ 30 ข้อตามลำดับ (คิดเป็นสัดส่วน 2:3)

6.3. คัดเลือกข้อสอบที่มีคุณภาพตามเกณฑ์ จากแบบทดสอบฉบับทดลอง ฉบับที่ 3 มาจำนวน 16 ข้อ จัดทำเป็นแบบทดสอบวัดการคิดเอกลัคนัยแบบสัมพันธ์ พอร์ม C ซึ่งจัดทำเป็นแบบทดสอบหลัก ให้มีเนื้อหาเป็นภาพกับภาษา จำนวน 8 ข้อกับ 8 ข้อ ตามลำดับ คิดเป็นสัดส่วน 1:1 แล้วจึงจัดเป็นฉบับย่อยอีก 3 ฉบับ ที่แบ่งตามสัดส่วนของขนาดความยาวจากแบบทดสอบพอร์ม A ดังนี้

6.3.1. แบบทดสอบหลัก ฉบับที่ 1 (พอร์ม C₁) จำนวน 16 ข้อ (คิดเป็นสัดส่วน .4 ของความยาวจากพอร์ม A) มีสัดส่วนของเนื้อหาที่เป็นภาพกับภาษา คิดเป็น 1:1

6.3.2. แบบทดสอบหลัก ฉบับที่ 2 (พอร์ม C₂) จำนวน 12 ข้อ (คิดเป็นสัดส่วน .3 ของความยาวจากพอร์ม A) ซึ่งได้มาจากการสุ่มข้อสอบจากแบบทดสอบพอร์ม C ฉบับที่ 1 ออกไปจำนวน 4 ข้อ และคงสัดส่วนของเนื้อหาที่เป็นภาพกับภาษา คิดเป็น 1:1

6.3.3. แบบทดสอบหลัก ฉบับที่ 3 (พอร์ม C₃) จำนวน 8 ข้อ (คิดเป็นสัดส่วน .2 ของความยาวจากพอร์ม A) ซึ่งได้มาจากการสุ่มข้อสอบจากแบบทดสอบพอร์ม C ฉบับที่ 2 ออกไปอีกจำนวน 4 ข้อ และคงสัดส่วนของเนื้อหาที่เป็นภาพกับภาษา คิดเป็น 1:1

7. หาความเชื่อมั่นของแบบทดสอบที่ได้คัดเลือกไว้แล้วทั้งสามพอร์ม โดยวิธีสูตรของเฟลด์ต์ (ดังหน้าที่ 54)

8. เตรียมแบบทดสอบสำหรับเทียบคะแนน เพื่อสอบกับกลุ่มตัวอย่าง 12 กลุ่ม ดังนี้

ก. แบบทดสอบเทียบคะแนนที่ใช้แบบทดสอบหลักแบบเชื่อมโยง ที่มีขนาดความยาวต่างกัน ได้แก่

พอร์ม A จำนวน 40 ข้อ พอร์ม C₁ จำนวน 16 ข้อ
 พอร์ม A จำนวน 40 ข้อ พอร์ม C₂ จำนวน 12 ข้อ
 พอร์ม A จำนวน 40 ข้อ พอร์ม C₃ จำนวน 8 ข้อ
 พอร์ม B จำนวน 50 ข้อ พอร์ม C₁ จำนวน 16 ข้อ
 พอร์ม B จำนวน 50 ข้อ พอร์ม C₂ จำนวน 12 ข้อ
 พอร์ม B จำนวน 50 ข้อ พอร์ม C₃ จำนวน 8 ข้อ

ข. แบบทดสอบเทียบคะแนนที่ใช้แบบทดสอบหลักแบบข้อสอบรวม ที่มีขนาดความยาวต่างกัน ได้แก่

พอร์ม A จำนวน 40 ข้อ + พอร์ม C₁ จำนวน 16 ข้อ
 พอร์ม A จำนวน 40 ข้อ + พอร์ม C₂ จำนวน 12 ข้อ
 พอร์ม A จำนวน 40 ข้อ + พอร์ม C₃ จำนวน 8 ข้อ
 พอร์ม B จำนวน 50 ข้อ + พอร์ม C₁ จำนวน 16 ข้อ
 พอร์ม B จำนวน 50 ข้อ + พอร์ม C₂ จำนวน 12 ข้อ
 พอร์ม B จำนวน 50 ข้อ + พอร์ม C₃ จำนวน 8 ข้อ

ลักษณะของแบบทดสอบที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ ประกอบด้วยแบบทดสอบวัดการคิดเอกลัแบบสัมพัทธ์ที่มีเนื้อหาเป็นภาพกับภาษา ถละกันในส่วนที่แตกต่างกันทั้งสามพอร์ม (ดังภาคผนวก ก)

แผนแบบการเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลครั้งนี้ จะใช้แบบทดสอบหลัก (Anchor Test) ที่เป็นแผนแบบของการใช้แบบทดสอบหลักที่มีประชากรแตกต่างกัน (CINEP) โดยมีแผนแบบดังนี้

1. การใช้แบบทดสอบหลักแบบเชื่อมโยง

1.1. ใช้แบบทดสอบเชื่อมโยงที่มีความยาว 16 ข้อ โดยใช้แบบทดสอบฟอร์ม A สอดกับกับกลุ่มเทียบคะแนนกลุ่มที่ 1 ฟอร์ม B สอดกับกับกลุ่มเทียบคะแนนกลุ่มที่ 2 ฟอร์ม C₁ สอดกับกับกลุ่มเทียบคะแนนทั้งสองกลุ่ม ดังตาราง 2

1.2. ใช้แบบทดสอบเชื่อมโยงที่มีความยาว 12 ข้อ โดยใช้แบบทดสอบฟอร์ม A สอดกับกับกลุ่มเทียบคะแนนกลุ่มที่ 3 ฟอร์ม B สอดกับกับกลุ่มเทียบคะแนนกลุ่มที่ 4 ฟอร์ม C₂ สอดกับกับกลุ่มเทียบคะแนนทั้งสองกลุ่ม ดังตาราง 2

1.3. ใช้แบบทดสอบเชื่อมโยงที่มีความยาว 8 ข้อ โดยใช้แบบทดสอบฟอร์ม A สอดกับกับกลุ่มเทียบคะแนนกลุ่มที่ 5 ฟอร์ม B สอดกับกับกลุ่มเทียบคะแนนกลุ่มที่ 6 ฟอร์ม C₃ สอดกับกับกลุ่มเทียบคะแนนทั้งสองกลุ่ม ดังตาราง 2

2. การใช้แบบทดสอบหลักแบบข้อสอบรวม เป็นการนำข้อสอบจากแบบทดสอบฟอร์ม C ไปรวมกับแบบทดสอบฟอร์ม A และฟอร์ม B แล้วนำไปสอดกับกับกลุ่มเทียบคะแนนทั้งสองกลุ่ม ดังนี้

2.1. ใช้ข้อสอบรวมที่มีความยาว 16 ข้อ รวมกับแบบทดสอบฟอร์ม A เป็นแบบทดสอบ 1 ฉบับ (ฟอร์ม A+C₁) สอดกับกับกลุ่มเทียบคะแนนกลุ่มที่ 7 และรวมกับแบบทดสอบฟอร์ม B เป็นแบบทดสอบอีก 1 ฉบับ (ฟอร์ม B+C₁) สอดกับกับกลุ่มเทียบคะแนนกลุ่มที่ 8 ดังตาราง 2

2.2. ใช้ข้อสอบรวมที่มีความยาว 12 ข้อ รวมกับแบบทดสอบฟอร์ม A เป็นแบบทดสอบ 1 ฉบับ (ฟอร์ม A+C₂) สอดกับกับกลุ่มเทียบคะแนนกลุ่มที่ 9 และรวมกับแบบทดสอบฟอร์ม B เป็นแบบทดสอบอีก 1 ฉบับ (ฟอร์ม B+C₂) สอดกับกับกลุ่มเทียบคะแนนกลุ่มที่ 10 ดังตาราง 2

2.3. ใช้ข้อสอบรวมที่มีความยาว 8 ข้อ รวมกับแบบทดสอบฟอร์ม A เป็นแบบทดสอบ 1 ฉบับ (ฟอร์ม A+C₃) สอบกับกับกลุ่มเทียบคะแนนกลุ่มที่ 9 และรวมกับแบบทดสอบฟอร์ม B เป็นแบบทดสอบอีก 1 ฉบับ (ฟอร์ม B+C₃) สอบกับกลุ่มเทียบคะแนนกลุ่มที่ 12 ดังตาราง 2

ตาราง 2 การเก็บรวบรวมข้อมูลที่ใช้แบบทดสอบหลักกับประชากรต่างกัน ตามแผนแบบ CINEP

สถานการณ์การใช้แบบทดสอบหลัก			
การใช้แบบเชื่อมโยง		การใช้แบบข้อสอบรวม	
กลุ่มที่ 1 ฟอร์ม A ฟอร์ม C ₁	กลุ่มที่ 2 ฟอร์ม B ฟอร์ม C ₁	กลุ่มที่ 7 ฟอร์ม A+C ₁	กลุ่มที่ 8 ฟอร์ม B+C ₁
กลุ่มที่ 3 ฟอร์ม A ฟอร์ม C ₂	กลุ่มที่ 4 ฟอร์ม B ฟอร์ม C ₂	กลุ่มที่ 9 ฟอร์ม A+C ₂	กลุ่มที่ 10 ฟอร์ม B+C ₂
กลุ่มที่ 5 ฟอร์ม A ฟอร์ม C ₃	กลุ่มที่ 6 ฟอร์ม B ฟอร์ม C ₃	กลุ่มที่ 11 ฟอร์ม A+C ₃	กลุ่มที่ 12 ฟอร์ม B+C ₃

วิธีดำเนินการเก็บรวบรวมข้อมูล

วิธีดำเนินการเก็บรวบรวมข้อมูล สำหรับการศึกษาครั้งนี้ ผู้วิจัยได้ดำเนินการเป็นขั้นตอน ดังนี้

1. ติดต่อกับวิทยาลัยพยาบาลต่าง ๆ ที่ได้รับคัดเลือกให้เป็นกลุ่มเทียบคะแนน กำหนดวัน เวลา และสถานที่ที่ใช้เป็นห้องสอบ โดยขออนุญาตจากผู้อำนวยการวิทยาลัยพยาบาล
2. จัดเตรียมแบบทดสอบตามแผนที่ได้วางไว้ โดยจัดทำเป็นแบบทดสอบที่มีลักษณะตามการใช้แบบทดสอบของแบบทดสอบหลัก กับขนาดความยาวของแบบทดสอบหลัก เพื่อใช้ทดสอบกับกลุ่มเทียบคะแนนที่วางแผนไว้ และจัดจำนวนของแบบทดสอบให้เพียงพอในแต่ละกลุ่ม
3. นำแบบทดสอบที่ได้จัดเตรียมไว้จากข้อ 2 ไปทำการทดสอบกับกลุ่มตัวอย่างสำหรับเทียบคะแนนทั้งสองกลุ่ม ตามแผนที่ได้กำหนดไว้ ดังนี้
 - 3.1. นักศึกษาพยาบาลจากวิทยาลัยพยาบาลบรมราชชนนีนครชุม
 - 3.1.1. หลักสูตรพยาบาลศาสตรบัณฑิตชั้นปีที่ 1 เป็นกลุ่มตัวอย่างสำหรับเทียบคะแนนกลุ่มที่ 1
 - 3.1.2. หลักสูตรพยาบาลศาสตรบัณฑิตชั้นปีที่ 2 เป็นกลุ่มตัวอย่างสำหรับเทียบคะแนนกลุ่มที่ 2
 - 3.1.3. หลักสูตรพยาบาลศาสตรระดับต้นชั้นปีที่ 1 เป็นกลุ่มตัวอย่างสำหรับเทียบคะแนนกลุ่มที่ 7
 - 3.1.4. หลักสูตรพยาบาลศาสตรระดับต้นชั้นปีที่ 2 เป็นกลุ่มตัวอย่างสำหรับเทียบคะแนนกลุ่มที่ 8
 - 3.2. นักศึกษาพยาบาลจากวิทยาลัยพยาบาลบรมราชชนนีนครบุรี
 - 3.2.1. หลักสูตรพยาบาลศาสตรบัณฑิตชั้นปีที่ 1 เป็นกลุ่มตัวอย่างสำหรับเทียบคะแนนกลุ่มที่ 3
 - 3.2.2. หลักสูตรพยาบาลศาสตรบัณฑิตชั้นปีที่ 2 เป็นกลุ่มตัวอย่างสำหรับเทียบคะแนนกลุ่มที่ 4
 - 3.2.3. หลักสูตรพยาบาลศาสตรระดับต้นชั้นปีที่ 1 เป็นกลุ่มตัวอย่างสำหรับเทียบคะแนนกลุ่มที่ 9

3.2.4. หลักสูตรพยาบาลศาสตรระดับต้นชั้นปีที่ 2 เป็นกลุ่มตัวอย่างสำหรับเทียบคะแนนกลุ่มที่ 10

3.3. นักศึกษาพยาบาลจากวิทยาลัยพยาบาลบรมราชชนนีสระบุรี

3.3.1. หลักสูตรพยาบาลศาสตรชั้นปีที่ 1 เป็นกลุ่มตัวอย่างสำหรับเทียบคะแนนกลุ่มที่ 5

3.3.2. หลักสูตรพยาบาลศาสตรชั้นปีที่ 2 เป็นกลุ่มตัวอย่างสำหรับเทียบคะแนนกลุ่มที่ 6

3.3.3. หลักสูตรพยาบาลศาสตรระดับต้นชั้นปีที่ 1 เป็นกลุ่มตัวอย่างสำหรับเทียบคะแนนกลุ่มที่ 11

3.3.4. หลักสูตรพยาบาลศาสตรระดับต้นชั้นปีที่ 2 เป็นกลุ่มตัวอย่างสำหรับเทียบคะแนนกลุ่มที่ 12

4. ดำเนินการสอบตามวันเวลาที่ได้นัดหมายไว้ในเดือนกรกฎาคม พ.ศ. 2538 โดยชี้แจงถึงวัตถุประสงค์ วิธีการสอบ และขอความร่วมมือในการทำแบบทดสอบจากกลุ่มตัวอย่าง ให้ออกมาในการตอบแบบทดสอบ แล้วจึงรวบรวมแบบทดสอบในเดือนสิงหาคม พ.ศ. 2538

5. นำผลการสอบมาตรวจนับคะแนน โดยใช้เครื่องคอมพิวเตอร์ช่วย และจัดกระทำกับข้อมูลด้วยการวิเคราะห์ข้อมูล โดยวิธีการทางสถิติด้วยโปรแกรม SPSS-PC

ขั้นตอนในการวิเคราะห์ข้อมูล

ขั้นตอนในการวิเคราะห์ข้อมูล ผู้วิจัยได้ดำเนินการเป็นลำดับ ดังต่อไปนี้

1. วิเคราะห์ข้อมูลที่ได้จากการทดลองใช้เครื่องมือ (Try out) โดยดำเนินการดังนี้

1.1. หาค่าความยาก และค่าอำนาจจำแนก โดยใช้สูตรอย่างง่ายคำนวณ

1.2. หาค่าความเชื่อมั่นของแบบทดสอบที่คัดเลือกไว้แล้ว จากข้อ 1.1. โดยใช้สูตรความเชื่อมั่นของเฟลด์ต์ (Feldt)

2. วิเคราะห์ข้อมูลที่ได้จากการทาวิจัย โดยดำเนินการ ดังนี้
 - 2.1. หาค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ($\bar{X}$) และความเบี่ยงเบนมาตรฐาน (SD)
 - 2.2. หาคะแนนแปลง ด้วยวิธีการแปลงคะแนนจากคะแนนดิบ (Y) ให้เป็นคะแนนแปลง (Y') โดยใช้วิธีการคำนวณด้วยสูตร สี่วิธี คือ วิธีที่ 1 ของแมคเคน วิธีที่ 2 ของแมคเคน วิธีที่ 1 ของวู้ดคริฟ และวิธีที่ 2 ของวู้ดคริฟ
 - 2.3. หาค่าความคลาดเคลื่อนของการเทียบคะแนน โดยใช้สูตรความคลาดเคลื่อนของการเทียบคะแนน (Total Error) และสูตรความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน (RMSE or Root Mean Squared Error)
 - 2.4. หาค่าความถี่ของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน จากการใช้เกณฑ์พิจารณาจากดัชนีความคลาดเคลื่อนมาตรฐาน โดยใช้วิธีของปีเตอร์สัน และคนอื่น ๆ (Petersen and others. 1982)
3. วิเคราะห์ความแปรปรวน (Analysis of Variance) ของคะแนนแปลง และความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน โดยวิเคราะห์องค์ประกอบแบบวัดซ้ำ มิติเดียว (Repeated Measure on The A Dimention) และเมื่อพบว่าปฏิสัมพันธ์มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ จึงวิเคราะห์ความแปรปรวนแบบทางเดียว (Oneway MANOVA) ที่เป็นแบบวัดซ้ำ
4. ทดสอบความแตกต่างรายคู่ด้วยวิธีของเชฟเฟ (Scheffe/)

สถิติสำหรับการวิเคราะห์ข้อมูล

1. หาค่าสถิติพื้นฐานของคะแนน ได้แก่ ค่าเฉลี่ย ($\bar{X}$) และความเบี่ยงเบนมาตรฐาน (SD)
2. วิเคราะห์ข้อสอบรายข้อ ได้แก่ หาค่าความยาก (P) และค่าอำนาจจำแนก (r) โดยใช้สูตรอย่างง่าย (บุญเชิด ภิญญอนันตพงษ์. 2519 : 145-148)
3. หาค่าความเชื่อมั่นของแบบทดสอบแต่ละฉบับ โดยใช้สูตรความเชื่อมั่นของเฟลด์ต์ (Feldt.1975) (บุญเชิด ภิญญอนันตพงษ์. 2533 : 30)

$$r_{tt} = \frac{s_{x1x2} / \lambda_1 \lambda_2}{s_x^2}$$

$$\lambda_1 = \left[s_{x1}^2 + s_{x1x2} \right] / s_x^2$$

$$\lambda_2 = \left[s_{x2}^2 + s_{x1x2} \right] / s_x^2$$

เมื่อ r_{tt}	แทน	ความเชื่อมั่นของแบบทดสอบแต่ละฉบับ
λ_1, λ_2	แทน	สัดส่วนของจำนวนข้อสอบแต่ละส่วน
s_{x1}^2	แทน	ความแปรปรวนของคะแนนจากส่วนที่ 1
s_{x2}^2	แทน	ความแปรปรวนของคะแนนจากส่วนที่ 2
s_x^2	แทน	ความแปรปรวนของคะแนนทั้งฉบับ
s_{x1x2}	แทน	ความแปรปรวนร่วมของคะแนน

4. หากคะแนนแปลงจากแบบทดสอบที่ใช้เทียบคะแนน โดยใช่วิธีการคำนวณด้วยสูตรของ แมคแคน 2 วิธี คือ (MacCann.1989: 266)

4.1. วิธีที่ 1 ของแมกแคน ใช้สูตร

$$Y/ = \begin{bmatrix} r_{yz2} & s_{y2} & s_{z1} \\ r_{xz1} & s_{x1} & s_{z2} \end{bmatrix} X + \bar{Y}_2 + \begin{bmatrix} s_{y2} \\ r_{yz2} & - \\ s_{z2} \end{bmatrix} \begin{bmatrix} 1 & s_{z1} \\ (\bar{Z}_1 - \bar{Z}_2) & - \\ r_{xz1} s_{x1} & \bar{X}_1 \end{bmatrix}$$

เมื่อ r_{xz1} , r_{yz2} แทน สัมประสิทธิ์สหสัมพันธ์ของคะแนน ระหว่างคะแนน X กับคะแนน Z_1 และคะแนน Y กับคะแนน Z_2 ตามลำดับ

s_x , s_y แทน ความเบี่ยงเบนมาตรฐานของคะแนน X และคะแนน Y ตามลำดับ

X , Y แทน คะแนนจากแบบทดสอบที่นำมาเทียบคะแนนกัน 2 ฉบับ

Z แทน คะแนนจากแบบทดสอบหลัก

4.2. วิธีที่ 2 ของแมกแคน ใช้สูตร

$$Y/ = \begin{bmatrix} r_{xz1} & s_{y2} & s_{z1} \\ r_{yz2} & s_{x1} & s_{z2} \end{bmatrix} X + \bar{Y}_2 + \begin{bmatrix} 1 & s_{y2} \\ - & - \\ r_{yz2} & s_{z2} \end{bmatrix} \begin{bmatrix} s_{z1} \\ (\bar{Z}_1 - \bar{Z}_2) - r_{xz1} & - \\ s_{x1} & \bar{X}_1 \end{bmatrix}$$

5. หาคะแนนแปลงจากแบบทดสอบที่ใช้เทียบคะแนน โดยใช้วิธีการคำนวณด้วยสูตรของ
 วัตต์รฟ์ 2 วิธี คือ (บุญเชิด ภิญโญนันตพงษ์. 2538)

5.1. วิธีที่ 1 ของวัตต์รฟ์ ใช้สูตร

$$Y' = \frac{s'(Y_1)}{s(X_1)} (X_1 - \bar{X}_1) + \bar{Y}_1'$$

$$\text{เมื่อ } s'(Y_1) = \left[s^2(Y_2) + \left[\frac{b_{y2z2}}{r(Z_2)} \right]^2 [s^2(Z_1) - s^2(Z_2)] \right]^{1/2}$$

$$Y_1' = \bar{Y}_2 + \left[\frac{b_{y2z2}}{r(Z_2)} \right] [\bar{Z}_1 - \bar{Z}_2]$$

$r(Z_2)$ แทน ค่าความเชื่อมั่นของคะแนนจริงสัมพันธ์

b_{y2z2} แทน ความลาดชันของเส้นถดถอย

5.2. วิธีที่ 2 ของวัตต์รฟ์ ใช้สูตร

$$Y' = \frac{s'(Y_1)}{s(X_1)} (X_1 - \bar{X}_1) + \bar{Y}_1'$$

$$\text{เมื่อ } s^2(Y_1) = \left[s^2(Y_2) + \frac{1 + b_{z_2y_2}}{1 + b_{y_2z_2}} \frac{s^2(Y_2)}{s^2(Z_2)} [s^2(Z_1) - s^2(Z_2)] \right]^{1/2}$$

$$\bar{Y}_1' = \bar{Y}_2 + \frac{1 + b_{z_2y_2}}{1 + b_{y_2z_2}} \frac{s^2(Y_2)}{s^2(Z_2)} [\bar{Z}_1 - \bar{Z}_2]$$

$b_{z_2y_2} \cdot b_{y_2z_2}$ แทน ความลาดชันของเส้นถดถอย

6. หาความคลาดเคลื่อนของการเทียบคะแนน จากผลต่างระหว่างคะแนนดิบกับคะแนนแปลง โดยใช้สูตร (MacCann, 1989 : 270)

$$\text{Total Error} = \frac{\sum f_i d_i^2}{n_i}$$

เมื่อ $d_i = Y_i - Y_i'$ แทน ค่าความแตกต่างระหว่างคะแนนดิบกับคะแนนแปลง

Y_i แทน คะแนนดิบ Y ของผู้เข้าสอบคนที่ i

Y_i' แทน คะแนนแปลง Y' ของผู้เข้าสอบคนที่ i

n_i แทน จำนวนผู้เข้าสอบ

7. หาความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน (Root Mean Squared Error) โดยใช้สูตรของโคป (Cope.1987 : 146 : MacCann.1989 : 270)

$$RMSE = \left[\frac{\sum f_i (Y_i - \bar{Y})^2}{ni} \right]^{1/2}$$

8. หาความถี่ของความคลาดเคลื่อน จากการใช้เกณฑ์พิจารณาจากดัชนีความคลาดเคลื่อนมาตรฐาน โดยใช้วิธีของปีเตอร์สันและคนอื่น ๆ (Petersen and others. 1982 : 91-93) ซึ่งได้แบ่งความคลาดเคลื่อนมาตรฐานตามระดับความพอใจเป็น 5 ระดับ ดังนี้ (MacCann.1989 : 270) (ภาวิณี ศรีสุขวัฒนานนท์.2528 : 102-103)

ระดับน่าพอใจอย่างมาก	หมายถึง	$C < (.05 SD_X)^2$
ระดับน่าพอใจ	หมายถึง	$(.05 SD_X)^2 < C \leq (.10 SD_X)^2$
ระดับพอใจปานกลาง	หมายถึง	$(.10 SD_X)^2 < C \leq (.15 SD_X)^2$
ระดับไม่น่าพอใจ	หมายถึง	$(.15 SD_X)^2 < C \leq (.20 SD_X)^2$
ระดับไม่น่าพอใจอย่างมาก	หมายถึง	$(.20 SD_X)^2 < C$

9. วิเคราะห์ความแปรปรวน (Analysis of Variance) ของคะแนนแปลง และความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน โดยใช้แผนแบบ $4 \times 2 \times 3$ องค์ประกอบแบบวัดซ้ำที่มีมิติเดียว ($4 \times 2 \times 3$ Factorial Design with Repeated Measure on The A Dimention) (Dayton.1970 : 271-273)

Source of Variation	DF	E (MS)
Among Subjects	<u>$nr-1$</u>	
B	$q-1$	$\sigma^2 + p\sigma_B \tau^2 + prn\sigma_B^2$
C	$r-1$	$\sigma^2 + pq\sigma \tau^2 + pqn\sigma r^2$
BC	$(q-1)(r-1)$	$\sigma^2 + p\sigma_B \tau^2 + pn\sigma_B r^2$
S	$r(n-1)$	$\sigma^2 + pq\sigma \tau^2$
Within Subjects	<u>$nr(pq-1)$</u>	
A	$p-1$	$\sigma^2 + q\sigma \alpha \tau^2 + prn\sigma \alpha^2$
AB	$(p-1)(q-1)$	$\sigma^2 + \sigma \alpha_B \tau^2 + rn\sigma \alpha_B^2$
AC	$(p-1)(r-1)$	$\sigma^2 + q\sigma \alpha \tau^2 + qn\sigma \alpha r^2$
ABC	$(p-1)(q-1)(r-1)$	$\sigma^2 + \sigma \alpha_B \tau^2 + n\sigma \alpha_B r^2$
AS	$r(p-1)(n-1)$	$\sigma^2 + q\sigma \alpha \tau^2$

10. ทดสอบนัยสำคัญของความแตกต่างรายคู่ โดยใช้สูตรของเชฟเฟ (Scheffe/)
(Byrkit.1975 : 276-277)

$$CV_d = \sqrt{(k-1)(F^*) (MS_{within})(2/n)}$$

เมื่อ k แทน จำนวนกลุ่มตัวอย่าง

F* แทน ค่า F ที่เปิดจากตาราง (critical value)

MS_{within} แทน ค่า mean squar within group ที่คำนวณได้แล้วใน
การวิเคราะห์ความแปรปรวน

n แทน จำนวนคนในกลุ่มตัวอย่างแต่ละกลุ่ม

บทที่ 4

ผลการวิเคราะห์ข้อมูล

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

ในการเสนอผลการวิเคราะห์ข้อมูล เพื่อความสะดวกและความเข้าใจตรงกัน ผู้วิจัยได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูลดังนี้

N	แทน	จำนวนนักศึกษาในกลุ่มเทียบคะแนน
k	แทน	จำนวนข้อสอบของแบบทดสอบ
X	แทน	คะแนนดิบจากแบบทดสอบฟอร์ม A
Y	แทน	คะแนนดิบจากแบบทดสอบฟอร์ม B
$\bar{X}$	แทน	ค่าเฉลี่ยของคะแนนจากแบบทดสอบ
SD	แทน	ความเบี่ยงเบนมาตรฐาน
$\bar{Y}$	แทน	คะแนนเฉลี่ยของคะแนนดิบจากแบบทดสอบฟอร์ม B
Y'	แทน	คะแนนแปลงของคะแนนดิบจากแบบทดสอบฟอร์ม B
$\bar{Y}'$	แทน	คะแนนแปลงเฉลี่ยของคะแนนดิบจากแบบทดสอบฟอร์ม B
RMSE	แทน	ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน

การนำเสนอผลการวิเคราะห์ข้อมูล

ผู้วิจัยขอรายงานผลการวิจัย โดยขอเสนอผลการวิเคราะห์ ดังนี้

ก. การวิเคราะห์ค่าความเชื่อมั่นของแบบทดสอบ ที่ได้จากการคำนวณด้วยสูตรหาค่าความเชื่อมั่นของเฟลด์ต์ ได้ค่าความเชื่อมั่นของแบบทดสอบทั้งสามพอร์ม ดังนี้

1. แบบทดสอบพอร์ม A มีความเชื่อมั่นเท่ากับ .6932 (แอลฟา = .7904)
2. แบบทดสอบพอร์ม B มีความเชื่อมั่นเท่ากับ .5983 (แอลฟา = .7621)
3. แบบทดสอบพอร์ม C มีความเชื่อมั่นเท่ากับ .5068 (แอลฟา = .6368)

ข. การวิเคราะห์ข้อมูลตามแผนแบบการวิจัย

ผลการวิเคราะห์ข้อมูล จากการนำคะแนนดิบ (Y) มาแปลงเป็นคะแนนแปลง (Y') เพื่อศึกษาคะแนนแปลงที่ได้จากวิธีการคำนวณด้วยสูตรสัวิธิ และความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สัวิธิ ผลจากการใช้แบบทดสอบหลัก สองลักษณะ คือ การใช้แบบเชื่อมโยงกับการใช้แบบข้อสอบรวม ผลจากการใช้ขนาดความยาวของแบบทดสอบหลัก สามขนาด คือ 8 ข้อ 12 ข้อ และ 16 ข้อ (คิดเป็นสัดส่วน .2, .3 และ .4) แล้วจึงเปรียบเทียบคะแนนแปลงที่ได้จากวิธีการคำนวณด้วยสูตร สัวิธิ จำแนกตามการใช้แบบทดสอบหลักที่มีขนาดความยาวต่างกัน และเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนที่เกิดจากวิธีการคำนวณด้วยสูตร สัวิธิ จำแนกตามการใช้แบบทดสอบหลักที่มีขนาดความยาวต่างกันเช่นกัน ดังจะได้นำเสนอตามลำดับ ต่อไปนี้

1. ค่าสถิติพื้นฐานของคะแนนจากผลการสอบ (คะแนนดิบ) ที่ได้จากแบบทดสอบสำหรับเทียบคะแนน สองฉบับ คือ พอร์ม A กับพอร์ม B จำแนกตามการใช้แบบทดสอบหลักสองลักษณะ คือ การใช้แบบทดสอบเชื่อมโยงเป็นแบบทดสอบหลัก กับการใช้ข้อสอบรวมเป็นแบบทดสอบหลัก ที่มีขนาดความยาวของแบบทดสอบหลัก สามขนาด คือ 8 ข้อ, 12 ข้อ และ 16 ข้อ

2. คะแนนแปลงเฉลี่ย ($\bar{Y}'$) ซึ่งได้มาจากผลการสอบจากแบบทดสอบพอร์ม B แล้วนำมาแปลงเป็นคะแนนแปลง (Y') เพื่อเทียบคะแนนไปยังคะแนนจากแบบทดสอบพอร์ม A โดยใช้วิธีการคำนวณด้วยสูตร สัวิธิ จำแนกตามการใช้แบบทดสอบหลัก สองลักษณะ ที่มีขนาดความยาวต่างกัน สามขนาด

3. ความคลาดเคลื่อนของการเทียบคะแนน (Total Error) ระหว่างคะแนนเฉลี่ย ($\bar{Y}$) กับคะแนนแปลงเฉลี่ย ($\bar{Y}'$) ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี จําแนกตามการraise แบบทดสอบหลัก สองลักษณะ ที่มีขนาดความยาวต่างกัน สามขนาด

4. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน (RMSE) ระหว่างคะแนนเฉลี่ย ($\bar{Y}$) กับคะแนนแปลงเฉลี่ย ($\bar{Y}'$) ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี จําแนกตามการraise แบบทดสอบหลัก สองลักษณะ ที่มีขนาดความยาวต่างกัน สามขนาด

5. การเปรียบเทียบคะแนนแปลง ที่ได้มาจากการคำนวณด้วยสูตร สี่วิธี จําแนกตามการraise แบบทดสอบหลัก สองลักษณะ ที่มีขนาดความยาวต่างกัน สามขนาด แล้วจึงเปรียบเทียบคะแนนแปลง ที่ได้มาจากการคำนวณด้วยสูตร สี่วิธี เมื่อใช้แบบทดสอบเชื่อมโยงเป็นแบบทดสอบหลัก ในแต่ละขนาดความยาว ด้วยการวิเคราะห์ความแปรปรวนแบบทางเดียว แบบวัดซ้ำ และทดสอบความแตกต่างรายคู่ของคะแนนแปลงเฉลี่ย ด้วยวิธีการทดสอบของเชฟเฟ

ในทางนองเดียวกัน เปรียบเทียบคะแนนแปลง ที่ได้มาจากการคำนวณด้วยสูตร สี่วิธี เมื่อใช้ข้อสอบร่วมเป็นแบบทดสอบหลักในแต่ละขนาดความยาว ด้วยการวิเคราะห์ความแปรปรวนแบบทางเดียว แบบวัดซ้ำ และทดสอบความแตกต่างรายคู่ของคะแนนแปลงเฉลี่ย ด้วยวิธีการทดสอบของเชฟเฟ

6. การเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี จําแนกตามการraise แบบทดสอบหลัก สองลักษณะ ที่มีขนาดความยาวต่างกัน สามขนาด แล้วจึงเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี เมื่อใช้แบบทดสอบเชื่อมโยงเป็นแบบทดสอบหลักในแต่ละขนาดความยาว ด้วยการวิเคราะห์ความแปรปรวนแบบทางเดียว แบบวัดซ้ำ และทดสอบความแตกต่างรายคู่ของค่าเฉลี่ย ด้วยวิธีการทดสอบของเชฟเฟ

ในทางนองเดียวกัน ก็เปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี เมื่อใช้ข้อสอบร่วมเป็นแบบทดสอบหลักในแต่ละขนาดความยาว ด้วยการวิเคราะห์ความแปรปรวนแบบทางเดียว แบบวัดซ้ำ และทดสอบความแตกต่างรายคู่ของค่าเฉลี่ย ด้วยวิธีการทดสอบของเชฟเฟ

จากการเสนอลำดับขั้นการนำเสนอผลการวิเคราะห์ข้อมูล ดังกล่าวข้างต้นนี้ ผู้วิจัย
ขอเสนอผลการวิเคราะห์ข้อมูลตามลำดับ โดยแสดงเป็นตาราง และมีรายละเอียดดังต่อไปนี้

ตาราง 3 ค่าสถิติพื้นฐานของคะแนนจากแบบทดสอบที่ใช้เทียบคะแนน สองฟอร์ม เมื่อใช้
แบบทดสอบหลัก ที่มีขนาดความยาวต่างกัน

ขนาดความยาวของ แบบทดสอบหลัก	แบบทดสอบสำหรับเทียบคะแนน			
	ฟอร์ม A (40 ข้อ)		ฟอร์ม B (50 ข้อ)	
<u>แบบทดสอบหลักแบบเชื่อมโยง</u>	$\bar{X}$	SD	$\bar{Y}$	SD
8 ข้อ	25.4028	5.0509	34.2500	4.1478
12 ข้อ	28.7639	4.8134	30.5139	5.1534
16 ข้อ	27.7083	3.1908	35.2639	3.1038
<u>แบบทดสอบหลักแบบข้อสอบรวม</u>				
8 ข้อ	25.1250	5.2109	32.4861	4.6147
12 ข้อ	25.2083	5.3627	34.3197	5.1178
16 ข้อ	27.3750	3.6326	30.7500	5.5430

จากตาราง 3 คะแนนดิบที่ได้จากแบบทดสอบฟอร์ม A และฟอร์ม B (คะแนน X และ
คะแนน Y) เมื่อพิจารณาตามการใช้แบบทดสอบหลักที่มีขนาดความยาวต่างกัน ผลปรากฏ ดังนี้

1. เมื่อใช้แบบทดสอบหลักเป็นแบบเชื่อมโยง ที่มีขนาดความยาว 8 ข้อ, 12 ข้อ
และ 16 ข้อ พบว่า มีค่าเฉลี่ยของแบบทดสอบฟอร์ม A และฟอร์ม B ตั้งแต่ 25.4028
ถึง 28.7639 และตั้งแต่ 30.5139 ถึง 35.2639 ตามลำดับ ความเบี่ยงเบนมาตรฐาน
ของคะแนน ตั้งแต่ 3.1908 ถึง 5.0509 และตั้งแต่ 3.1038 ถึง 5.1534 ตามลำดับ

2. เมื่อใช้แบบทดสอบหลักแบบข้อสอบรวม ที่มีขนาดความยาว 8 ข้อ, 12 ข้อ

และ 16 ข้อ พบว่า มีค่าเฉลี่ยของแบบทดสอบฟอร์ม A และฟอร์ม B ตั้งแต่ 25.1250 ถึง 27.3750 และตั้งแต่ 30.7500 ถึง 34.3197 ตามลำดับ ความเบี่ยงเบนมาตรฐานของคะแนน ตั้งแต่ 3.6326 ถึง 5.3627 และตั้งแต่ 4.6147 ถึง 5.5430 ตามลำดับ

เมื่อพิจารณาค่าเฉลี่ย จำแนกตามการใช้แบบทดสอบหลัก สองลักษณะ โดยพิจารณาในแต่ละขนาดความยาวของแบบทดสอบหลักแล้ว ปรากฏผล ดังนี้

1. เมื่อใช้แบบทดสอบหลักแบบเชื่อมโยง พบว่า

1.1. ที่ขนาดความยาว 8 ข้อ ของแบบทดสอบหลักแบบเชื่อมโยง ค่าเฉลี่ยที่ได้จากแบบทดสอบฟอร์ม A และฟอร์ม B มีค่าเท่ากับ 25.4028 และ 34.2500 ตามลำดับ ความเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 5.0509 และ 4.1478 ตามลำดับ

1.2. ที่ขนาดความยาว 12 ข้อ ของแบบทดสอบหลักแบบเชื่อมโยง ค่าเฉลี่ยที่ได้จากแบบทดสอบฟอร์ม A และฟอร์ม B มีค่าเท่ากับ 28.7639 และ 30.5139 ตามลำดับ ความเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 4.8134 และ 5.1534 ตามลำดับ

1.3. ที่ขนาดความยาว 16 ข้อ ของแบบทดสอบหลักแบบเชื่อมโยง ค่าเฉลี่ยที่ได้จากแบบทดสอบฟอร์ม A และฟอร์ม B มีค่าเท่ากับ 27.7083 และ 35.2639 ตามลำดับ ความเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 3.1908 และ 3.1038 ตามลำดับ

2. เมื่อใช้แบบทดสอบหลักแบบข้อสอบรวม พบว่า

2.1. ที่ขนาดความยาว 8 ข้อ ของการใช้ข้อสอบรวม ค่าเฉลี่ยที่ได้จากแบบทดสอบฟอร์ม A และฟอร์ม B มีค่าเท่ากับ 25.1250 และ 32.4861 ตามลำดับ ความเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 5.2109 และ 4.6147 ตามลำดับ

2.2. ที่ขนาดความยาว 12 ข้อ ของการใช้ข้อสอบรวม ค่าเฉลี่ยที่ได้จากแบบทดสอบฟอร์ม A และฟอร์ม B มีค่าเท่ากับ 25.2083 และ 34.3197 ตามลำดับ ความเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 5.3627 และ 5.1178 ตามลำดับ

2.3. ที่ขนาดความยาว 16 ข้อ ของการใช้ข้อสอบรวม ค่าเฉลี่ยที่ได้จากแบบทดสอบฟอร์ม A และฟอร์ม B มีค่าเท่ากับ 27.3750 และ 30.7500 ตามลำดับ ความเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 3.6326 และ 5.5430 ตามลำดับ

ตาราง 4 คะแนนแปลงเฉลี่ยของแบบทดสอบสำหรับเทียบคะแนน ที่ได้มาจากวิธีการคำนวณด้วยสูตร สี่วิธี จำแนกตามการใช้แบบทดสอบหลัก สองลักษณะ ที่มีขนาดความยาวต่างกัน

ขนาดความยาวของ แบบทดสอบหลัก	คะแนนแปลงเฉลี่ย ($\bar{Y}$)			
	แมคแคนน1	แมคแคนน2	รูดอล์ฟ1	รูดอล์ฟ2
<u>แบบทดสอบหลักแบบเชื่อมโยง</u>				
8 ข้อ	34.38	42.64	33.08	36.82
12 ข้อ	31.99	39.66	35.99	35.96
16 ข้อ	34.46	26.49	33.38	31.85
<u>แบบทดสอบหลักแบบข้อสอบรวม</u>				
8 ข้อ	32.09	25.11	33.94	28.86
12 ข้อ	33.36	33.05	32.63	32.56
16 ข้อ	31.70	37.36	37.10	31.60

จากตาราง 4 คะแนนแปลงเฉลี่ย ($\bar{Y}$) ที่ได้มาจากวิธีการคำนวณด้วยสูตร สี่วิธี คือ วิธีที่ 1 ของแมคแคนน, วิธีที่ 2 ของแมคแคนน, วิธีที่ 1 ของรูดอล์ฟ และวิธีที่ 2 ของรูดอล์ฟ เมื่อพิจารณาตามการใช้แบบทดสอบหลัก สองลักษณะ ที่มีขนาดความยาวต่างกัน สามขนาด ให้ผลปรากฏ ดังนี้

1. เมื่อใช้แบบทดสอบหลักแบบเชื่อมโยง ที่มีขนาดความยาวต่างกัน สามขนาด พบว่า คะแนนแปลงของการเทียบคะแนน ที่ได้มาจากวิธีการคำนวณด้วยสูตร สามวิธี มีค่าที่ใกล้เคียงกัน ได้แก่ วิธีที่ 1 ของแมคแคนน มีค่าตั้งแต่ 31.99 ถึง 34.46, วิธีที่ 1 ของรูดอล์ฟ มีค่าตั้งแต่ 33.08 ถึง 35.99 และวิธีที่ 2 ของรูดอล์ฟ มีค่าตั้งแต่ 31.85 ถึง 36.82 ยกเว้นวิธีที่ 2 ของแมคแคนน ที่มีช่วงคะแนนห่างกันมาก คือ ตั้งแต่ 26.49 ถึง 42.64

2. เมื่อใช้แบบทดสอบหลักแบบข้อสอบรวม ที่มีขนาดความยาวต่างกัน สามขนาด พบว่าคะแนนแปลงของการเทียบคะแนน ที่ได้มาจากวิธีการคำนวณด้วยสูตร สามวิธี มีค่าที่ใกล้เคียงกัน ได้แก่ วิธีที่ 1 ของแมคเคน มีค่าตั้งแต่ 31.70 ถึง 33.36, วิธีที่ 1 ของวีตดริฟ มีค่าตั้งแต่ 32.63 ถึง 37.10 และวิธีที่ 2 ของวีตดริฟ มีค่าตั้งแต่ 28.86 ถึง 32.56 ยกเว้นวิธีที่ 2 ของแมคเคน ที่มีช่วงคะแนนห่างกันมาก คือ มีค่าตั้งแต่ 25.11 ถึง 37.36,

เมื่อพิจารณางานแนวทางการใช้ของแบบทดสอบหลักในแต่ละขนาดความยาว ได้ผลดังนี้

1. เมื่อใช้แบบทดสอบหลักแบบเชื่อมโยง พบว่า

1.1. ที่ขนาดความยาว 8 ข้อ มีคะแนนแปลงเฉลี่ย ดังนี้

วิธีที่ 1 ของแมคเคน มีค่าเท่ากับ 34.38 , วิธีที่ 2 ของแมคเคน มีค่าเท่ากับ 42.64 ,
วิธีที่ 1 ของวีตดริฟ มีค่าเท่ากับ 33.08 และวิธีที่ 2 ของวีตดริฟ มีค่าเท่ากับ 36.82

1.2. ที่ขนาดความยาว 12 ข้อ มีคะแนนแปลงเฉลี่ย ดังนี้

วิธีที่ 1 ของแมคเคน มีค่าเท่ากับ 31.99 , วิธีที่ 2 ของแมคเคน มีค่าเท่ากับ 39.66 ,
วิธีที่ 1 ของวีตดริฟ มีค่าเท่ากับ 35.99 และวิธีที่ 2 ของวีตดริฟ มีค่าเท่ากับ 35.96

1.3. ที่ขนาดความยาว 16 ข้อ มีคะแนนแปลงเฉลี่ย ดังนี้

วิธีที่ 1 ของแมคเคน มีค่าเท่ากับ 34.46 , วิธีที่ 2 ของแมคเคน มีค่าเท่ากับ 26.49 ,
วิธีที่ 1 ของวีตดริฟ มีค่าเท่ากับ 33.38 และวิธีที่ 2 ของวีตดริฟ มีค่าเท่ากับ 31.85

2. เมื่อใช้ข้อสอบรวม เป็นแบบทดสอบหลัก พบว่า

2.1. ที่ขนาดความยาว 8 ข้อ มีคะแนนแปลงเฉลี่ย ดังนี้

วิธีที่ 1 ของแมคเคน มีค่าเท่ากับ 32.09 , วิธีที่ 2 ของแมคเคน มีค่าเท่ากับ 25.11 ,
วิธีที่ 1 ของวีตดริฟ มีค่าเท่ากับ 33.94 และวิธีที่ 2 ของวีตดริฟ มีค่าเท่ากับ 28.86

2.2. ที่ขนาดความยาว 12 ข้อ มีคะแนนแปลงเฉลี่ย ดังนี้

วิธีที่ 1 ของแมคเคน มีค่าเท่ากับ 33.36 , วิธีที่ 2 ของแมคเคน มีค่าเท่ากับ 33.05 ,
วิธีที่ 1 ของวีตดริฟ มีค่าเท่ากับ 32.63 และวิธีที่ 2 ของวีตดริฟ มีค่าเท่ากับ 32.56

2.3. ที่ขนาดความยาว 16 ข้อ มีคะแนนแปลงเฉลี่ย ดังนี้

วิธีที่ 1 ของแมคเคน มีค่าเท่ากับ 31.70 , วิธีที่ 2 ของแมคเคน มีค่าเท่ากับ 37.36 ,
วิธีที่ 1 ของวีตดริฟ มีค่าเท่ากับ 37.10 และวิธีที่ 2 ของวีตดริฟ มีค่าเท่ากับ 31.60

ตาราง 5 ความคลาดเคลื่อนของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร ส่วริ
จำแนกตามการใช้แบบทดสอบหลัก สองลักษณะ ที่มีขนาดความยาวต่างกัน

ขนาดความยาวของ แบบทดสอบหลัก	ความคลาดเคลื่อนของการเทียบคะแนน (Total Error)			
	แมคเคน1	แมคเคน2	รูตครีฟ1	รูตครีฟ2
<u>แบบทดสอบหลักแบบเชื่อมโยง</u>				
8 ข้อ	.32	2.40	.27	.34
12 ข้อ	.67	1.77	.76	.76
16 ข้อ	.19	1.59	.18	.29
<u>แบบทดสอบหลักแบบข้อสอบรวม</u>				
8 ข้อ	.44	2.44	.32	.48
12 ข้อ	.98	.77	.43	.44
16 ข้อ	.96	1.20	.99	.44

จากตาราง 5 พบว่า ความคลาดเคลื่อนของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร ส่วริ เมื่อพิจารณาตามการใช้แบบทดสอบหลัก สองลักษณะ ที่มีขนาดความยาวต่างกัน สามขนาด โดยพิจารณาในแต่ละขนาดความยาว ปรากฏผล ดังนี้

1. เมื่อใช้แบบทดสอบหลักแบบเชื่อมโยง พบว่า

1.1. ที่ขนาดความยาว 8 ข้อ ความคลาดเคลื่อนของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สามวิธี มีค่าใกล้เคียงกัน ได้แก่ วิธีที่ 1 ของแมคเคน มีค่าเท่ากับ .32 , วิธีที่ 1 ของรูตครีฟ มีค่าเท่ากับ .27 และวิธีที่ 2 ของรูตครีฟ มีค่าเท่ากับ .34 ยกเว้นวิธีที่ 2 ของแมคเคน ที่มีค่าค่อนข้างมาก คือ มีค่าเท่ากับ 2.40

1.2. ที่ขนาดความยาว 12 ข้อ (.3) ความคลาดเคลื่อนของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สามวิธี มีค่าใกล้เคียงกัน ได้แก่ วิธีที่ 1 ของแมคเคน มีค่าเท่ากับ .67 , วิธีที่ 1 ของวู้ดดรัฟ มีค่าเท่ากับ .76 และวิธีที่ 2 ของวู้ดดรัฟ มีค่าเท่ากับ .76 ยกเว้นวิธีที่ 2 ของแมคเคน ที่มีค่าค่อนข้างมาก คือ มีค่าเท่ากับ 1.77

1.3. ที่ขนาดความยาว 16 ข้อ (.4) ความคลาดเคลื่อนของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สามวิธี มีค่าใกล้เคียงกัน ได้แก่ วิธีที่ 1 ของแมคเคน มีค่าเท่ากับ .19 , วิธีที่ 1 ของวู้ดดรัฟ มีค่าเท่ากับ .18 และวิธีที่ 2 ของวู้ดดรัฟ มีค่าเท่ากับ .29 ยกเว้นวิธีที่ 2 ของแมคเคน ที่มีค่าค่อนข้างมาก คือ มีค่าเท่ากับ 1.59

2. เมื่อใช้ข้อสอบรวม เป็นแบบทดสอบหลัก พบว่า

2.1. ที่ขนาดความยาว 8 ข้อ ความคลาดเคลื่อนของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สามวิธี มีค่าใกล้เคียงกัน ได้แก่ วิธีที่ 1 ของแมคเคน มีค่าเท่ากับ .44 , วิธีที่ 1 ของวู้ดดรัฟ มีค่าเท่ากับ .32 และวิธีที่ 2 ของวู้ดดรัฟ มีค่าเท่ากับ .48 ยกเว้นวิธีที่ 2 ของแมคเคน ที่มีค่าค่อนข้างมาก คือ มีค่าเท่ากับ 2.44

2.2. ที่ขนาดความยาว 12 ข้อ ความคลาดเคลื่อนของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สองวิธี มีค่าใกล้เคียงกันเป็นคู่ ได้แก่ วิธีที่ 1 ของวู้ดดรัฟ มีค่าเท่ากับ .43 กับวิธีที่ 2 ของวู้ดดรัฟ มีค่าเท่ากับ .44 และวิธีที่ 1 ของแมคเคน มีค่าเท่ากับ .98 กับวิธีที่ 2 ของแมคเคน มีค่าเท่ากับ .77

2.3. ที่ขนาดความยาว 16 ข้อ ความคลาดเคลื่อนของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สองวิธี มีค่าใกล้เคียงกัน ได้แก่ วิธีที่ 1 ของแมคเคน มีค่าเท่ากับ .96 กับวิธีที่ 1 ของวู้ดดรัฟ มีค่าเท่ากับ .99 ส่วนวิธีที่ 2 ของแมคเคน มีค่ามากกว่า คือ มีค่าเท่ากับ 1.20 และวิธีที่ 2 ของวู้ดดรัฟ มีค่าน้อยกว่า คือ มีค่าเท่ากับ .44

ตาราง 6 ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี จำแนกตามการใช้แบบทดสอบหลัก สองลักษณะ ที่มีขนาดความยาวต่างกัน

ขนาดความยาวของ แบบทดสอบหลัก	ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน (RMSE)							
	แมคเคน1		แมคเคน2		รูคดร์ฟ1		รูคดร์ฟ2	
แบบทดสอบหลักแบบเชื่อมโยง	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD
8 ข้อ	.45	.35	1.26	.90	.44	.28	.43	.40
12 ข้อ	.63	.52	1.14	.69	.71	.52	.71	.51
16 ข้อ	.34	.27	1.06	.69	.32	.28	.43	.33
แบบทดสอบหลักแบบข้อสอบรวม	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD
8 ข้อ	.51	.42	1.20	1.01	.42	.39	.59	.36
12 ข้อ	.76	.63	.71	.52	.54	.38	.54	.38
16 ข้อ	.79	.59	.91	.62	.80	.60	.50	.44

จากตาราง 6 ค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน (RMSE) ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี เมื่อพิจารณาตามการใช้แบบทดสอบหลัก สองลักษณะ ที่มีขนาดความยาวต่างกัน สามขนาด โดยพิจารณาในแต่ละขนาดความยาว ปรากฏผล ดังนี้

1. เมื่อใช้แบบทดสอบหลักแบบเชื่อมโยง พบว่า

1.1. ที่ขนาดความยาว 8 ข้อ ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สามวิธี มีค่าใกล้เคียงกัน ได้แก่ วิธีที่ 1 ของแมคเคน มีค่าเท่ากับ .45 , วิธีที่ 1 ของรูคดร์ฟ มีค่าเท่ากับ .44 และวิธีที่ 2 ของรูคดร์ฟ มีค่าเท่ากับ .43 ยกเว้นวิธีที่ 2 ของแมคเคน ที่มีค่าเฉลี่ยค่อนข้างมาก คือ มีค่าเท่ากับ 1.26 และมีการกระจายของคะแนน ที่ต่างจากวิธีการคำนวณด้วยสูตร สามวิธีข้างต้น คือ มีความเบี่ยงเบนมาตรฐานเท่ากับ .90

1.2. ที่ขนาดความยาว 12 ข้อ ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนที่เกิดจากวิธีการคำนวณด้วยสูตร สามวิธี มีค่าใกล้เคียงกัน ได้แก่ วิธีที่ 1 ของแมคแคน มีค่าเท่ากับ .63 , วิธีที่ 1 ของรูตดริฟ มีค่าเท่ากับ .71 และวิธีที่ 2 ของรูตดริฟ มีค่าเท่ากับ .71 ยกเว้นวิธีที่ 2 ของแมคแคน ที่มีค่าเฉลี่ยค่อนข้างมาก คือ มีค่าเท่ากับ 1.14 ส่วนการกระจายของคะแนน จากวิธีการคำนวณด้วยสูตรทั้ง สี่วิธี มีค่าใกล้เคียงกัน คือ มีความเบี่ยงเบนมาตรฐาน ตั้งแต่ .51 ถึง .69

1.3. ที่ขนาดความยาว 16 ข้อ ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนที่เกิดจากวิธีการคำนวณด้วยสูตร สามวิธี มีค่าใกล้เคียงกัน ได้แก่ วิธีที่ 1 ของแมคแคน มีค่าเท่ากับ .34 , วิธีที่ 1 ของรูตดริฟ มีค่าเท่ากับ .32 และวิธีที่ 2 ของรูตดริฟ มีค่าเท่ากับ .33 ยกเว้นวิธีที่ 2 ของแมคแคน ที่มีค่าเฉลี่ยค่อนข้างมาก คือ มีค่าเท่ากับ 1.06 และมีการกระจายของคะแนน ที่ต่างจากวิธีการคำนวณด้วยสูตร สามวิธีข้างต้น คือ มีความเบี่ยงเบนมาตรฐานเท่ากับ .69

2. เมื่อใช้ข้อสอบรวม เป็นแบบทดสอบหลัก พบว่า

2.1. ที่ขนาดความยาว 8 ข้อ ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนที่เกิดจากวิธีการคำนวณด้วยสูตร สามวิธี มีค่าใกล้เคียงกัน ได้แก่ วิธีที่ 1 ของแมคแคน มีค่าเท่ากับ .51 , วิธีที่ 1 ของรูตดริฟ มีค่าเท่ากับ .42 และวิธีที่ 2 ของรูตดริฟ มีค่าเท่ากับ .59 ยกเว้นวิธีที่ 2 ของแมคแคน ที่มีค่าเฉลี่ยค่อนข้างมาก คือ มีค่าเท่ากับ 1.20 และมีการกระจายของคะแนน ที่ต่างจากวิธีการคำนวณด้วยสูตร สามวิธีข้างต้น คือ มีความเบี่ยงเบนมาตรฐานเท่ากับ 1.01

2.2. ที่ขนาดความยาว 12 ข้อ ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนที่เกิดจากวิธีการคำนวณด้วยสูตรทั้ง สี่วิธี มีค่าใกล้เคียงกัน ได้แก่ วิธีที่ 1 ของแมคแคน มีค่าเท่ากับ .76, วิธีที่ 2 ของแมคแคน มีค่าเท่ากับ .71, วิธีที่ 1 ของรูตดริฟ มีค่าเท่ากับ .54 และวิธีที่ 2 ของรูตดริฟ มีค่าเท่ากับ .54 ส่วนการกระจายของคะแนน จากวิธีการคำนวณด้วยสูตรทั้ง สี่วิธี มีค่าใกล้เคียงกัน คือ มีความเบี่ยงเบนมาตรฐาน ตั้งแต่ .38 ถึง .63

2.3. ที่ขนาดความยาว 16 ข้อ ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนที่เกิดจากวิธีการคำนวณด้วยสูตรทั้ง สี่วิธี มีค่าใกล้เคียงกัน ได้แก่ วิธีที่ 1 ของแมคแคน มีค่าเท่ากับ .79 , วิธีที่ 2 ของแมคแคน มีค่าเท่ากับ .91, วิธีที่ 1 ของรูตดริฟ มีค่าเท่ากับ .80 และวิธีที่ 2 ของรูตดริฟ มีค่าเท่ากับ .50 ส่วนการกระจายของคะแนน จากวิธีการคำนวณด้วยสูตรทั้ง สี่วิธี มีค่าใกล้เคียงกัน คือ มีความเบี่ยงเบนมาตรฐาน ตั้งแต่ .44 ถึง .62

เมื่อศึกษาความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธีแล้ว นำคะแนนที่ได้มาเข้าเกณฑ์ดัชนีความคลาดเคลื่อนของการเทียบคะแนน ที่แมคแคน ได้ศึกษาไว้ โดยแบ่งเป็นช่วงของความคลาดเคลื่อนมาตรฐาน หรือเป็นระดับความพอใจ 5 ระดับ ซึ่งค่าที่ได้จากการแจกแจงนั้น เป็นความถี่ของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน แต่ละค่า ดังแสดงไว้ใน ตาราง 7

ตาราง 7 ความถี่ของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนแต่ละค่า เมื่อแบ่งตามระดับความพอใจ 5 ระดับ

ขนาดความยาวของ แบบทดสอบหลัก	N	ความถี่ของความคลาดเคลื่อนของการเทียบคะแนน 5 ระดับ																			
		แมคแคน1					แมคแคน2					รูคดรัฟ1					รูคดรัฟ2				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
<u>แบบเชื่อมโยง</u>																					
8 ข้อ	72	5	2	7	4	54	1	5	-	1	65	1	5	4	4	58	5	10	6	9	42
12 ข้อ	72	3	3	5	4	57	1	1	2	2	66	3	3	6	4	56	2	4	6	4	56
16 ข้อ	72	10	6	6	9	41	1	-	1	2	68	9	10	10	2	41	4	4	9	7	48
<u>แบบข้อสอบรวม</u>																					
8 ข้อ	72	5	5	3	7	52	2	1	2	-	67	7	3	8	9	45	4	-	6	3	59
12 ข้อ	72	3	7	6	5	51	3	3	1	2	63	3	5	2	5	57	4	5	2	4	57
16 ข้อ	72	2	4	5	2	59	1	3	4	-	64	4	3	2	3	60	8	5	5	3	51

จากตาราง 7 พบว่า ความถี่ของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี เมื่อพิจารณาตามการใช้แบบทดสอบหลัก สองลักษณะ ที่มี ีขนาดความยาวต่างกัน สามขนาดให้ผลปรากฏ ดังนี้

1. เมื่อใช้แบบทดสอบหลักแบบเชื่อมโยง ที่มีขนาดความยาว 8 ข้อ, 12 ข้อ และ 16 ข้อ พบว่า ความถี่ของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี มีการกระจายของคะแนนความคลาดเคลื่อนของการเทียบคะแนน จากระดับความพอใจ ตั้งแต่ระดับที่ 1 ถึงระดับที่ 5 โดยมีความถี่ของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนในแต่ละระดับที่ใกล้เคียงกัน และส่วนใหญ่มักอยู่ในระดับที่ 5 คือ วิธีที่ 1 ของแมคแคน มีค่าความถี่ของความคลาดเคลื่อนมาตรฐาน ตั้งแต่ 41 ถึง 57, วิธีที่ 2 ของแมคแคน มีค่าความถี่ของความคลาดเคลื่อนมาตรฐาน ตั้งแต่ 65 ถึง 68 , วิธีที่ 1 ของ ู๊ดครีฟ มีค่าความถี่ของความคลาดเคลื่อนมาตรฐาน ตั้งแต่ 41 ถึง 58 และวิธีที่ 2 ของ ู๊ดครีฟ มีค่าความถี่ของความคลาดเคลื่อนมาตรฐาน ตั้งแต่ 42 ถึง 56
2. เมื่อใช้แบบทดสอบหลักแบบข้อสอบรวม ที่มีขนาดความยาว 8 ข้อ, 12 ข้อ และ 16 ข้อ พบว่า ความถี่ของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี มีการกระจายของคะแนนความคลาดเคลื่อนของการเทียบคะแนน จากระดับความพอใจ ตั้งแต่ระดับที่ 1 ถึงระดับที่ 5 โดยมีความถี่ของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนในแต่ละระดับที่ใกล้เคียงกัน และส่วนใหญ่มักอยู่ในระดับที่ 5 คือ วิธีที่ 1 ของแมคแคน มีค่าความถี่ของความคลาดเคลื่อนมาตรฐาน ตั้งแต่ 51 ถึง 59, วิธีที่ 2 ของแมคแคน มีค่าความถี่ของความคลาดเคลื่อนมาตรฐาน ตั้งแต่ 63 ถึง 67 , วิธีที่ 1 ของ ู๊ดครีฟ มีค่าความถี่ของความคลาดเคลื่อนมาตรฐาน ตั้งแต่ 45 ถึง 60 และวิธีที่ 2 ของ ู๊ดครีฟ มีค่าความถี่ของความคลาดเคลื่อนมาตรฐาน ตั้งแต่ 51 ถึง 59

ตาราง 8 การเปรียบเทียบคะแนนแปลงของการเทียบคะแนน ที่ได้มาจากวิธีการคำนวณด้วยสูตร สี่วิธี จำนวนตามการใช้แบบทดสอบหลัก ที่มีขนาดความยาว ต่างกัน

แหล่งความแปรปรวน	SS	DF	MS	F
<u>ระหว่างกลุ่ม</u>				
ภายในเซลล์	17402.56	426	40.85	
ค่าคงที่	1947667.62	1	1947667.60	47677.27
ขนาดความยาว	7277.03	2	3638.51	89.07**
การใช้แบบทดสอบหลัก	2296.17	1	2296.17	56.21**
การใช้แบบทดสอบหลัก กับขนาดความยาว	220.10	2	110.05	2.69**
<u>ภายในกลุ่ม</u>				
ภายในเซลล์	12442.67	1278	9.74	
วิธีการคำนวณ	653.55	3	217.85	22.38**
วิธีการคำนวณ กับขนาดความยาว	10144.50	6	1690.75	173.66**
วิธีการคำนวณ กับการใช้แบบทดสอบหลัก	1721.65	3	573.88	58.94**
วิธีการคำนวณ กับการใช้แบบทดสอบหลัก กับขนาดความยาว	1938.59	6	323.10	33.19**

** มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง 8 คะแนนแปลงที่ได้มาจากวิธีการคำนวณด้วยสูตรต่างกัน สี่วิธี กับการใช้แบบทดสอบหลักต่างกัน สองลักษณะ ที่มีขนาดความยาวต่างกัน สามขนาด พบว่า ปฏิสัมพันธ์ของคะแนนแปลงมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อพิจารณาคะแนนแปลงที่ได้มาจากวิธีการคำนวณด้วยสูตรต่างกัน สี่วิธี กับขนาดความยาวของแบบทดสอบหลักต่างกัน สามขนาด ก็พบว่า ปฏิสัมพันธ์ของคะแนนแปลงมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อพิจารณาคะแนนแปลงที่ได้มาจากวิธีการคำนวณด้วยสูตรต่างกัน สี่วิธี กับการใช้แบบทดสอบหลักต่างกัน สองลักษณะ ก็พบว่า ปฏิสัมพันธ์ของคะแนนแปลงมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

นอกจากนี้ เมื่อพิจารณาคะแนนแปลงที่ได้มาจากการใช้แบบทดสอบหลักต่างกัน สองลักษณะ กับขนาดความยาวต่างกัน สามขนาด พบว่า ปฏิสัมพันธ์ของคะแนนแปลงมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เช่นกัน

ตาราง 9 การเปรียบเทียบคะแนนแปลงของการเทียบคะแนน โดยการใช้แบบทดสอบหลักแบบ
เชื่อมเรียงที่มีขนาดความยาวต่างกัน ด้วยการวิเคราะห์ความแปรปรวนแบบทางเดียว แบบวัดซ้ำ

แหล่งความแปรปรวน	SS	DF	MS	F
------------------	----	----	----	---

แบบทดสอบหลักแบบเชื่อมเรียง ขนาดความยาว 8 ข้อ

ภายในกลุ่ม

ภายในเซลล์	1172.76	213	5.51	
วิธีการคำนวณ	2696.75	3	898.92	163.26**

แบบทดสอบหลักแบบเชื่อมเรียง ขนาดความยาว 12 ข้อ

ภายในกลุ่ม

ภายในเซลล์	1368.13	213	6.42	
วิธีการคำนวณ	2117.93	3	705.98	109.91**

แบบทดสอบหลักแบบเชื่อมเรียง ขนาดความยาว 16 ข้อ

ภายในกลุ่ม

ภายในเซลล์	3257.94	213	15.30	
วิธีการคำนวณ	3868.07	3	1289.36	84.30**

** มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง 9 พบว่า คะแนนแปลงของการเทียบคะแนน เมื่อใช้แบบทดสอบหลักแบบ
เชื่อมเรียงที่มีขนาดความยาวต่างกัน สามขนาด โดยในแต่ละขนาดความยาว ให้ผลที่แตกต่างกัน
อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อพิจารณาคะแนนแปลงของการเทียบคะแนน ตามขนาดความยาวในแต่ละขนาด พบว่า

1. ที่ขนาดความยาว 8 ข้อ คะแนนแปลงของการเทียบคะแนนที่ได้มาจากวิธีการคำนวณด้วยสูตร สี่วิธี มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
2. ที่ขนาดความยาว 12 ข้อ คะแนนแปลงของการเทียบคะแนนที่ได้มาจากวิธีการคำนวณด้วยสูตร สี่วิธี มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
3. ที่ขนาดความยาว 16 ข้อ คะแนนแปลงของการเทียบคะแนนที่ได้มาจากวิธีการคำนวณด้วยสูตร สี่วิธี มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อทดสอบความแตกต่างรายคู่ของค่าเฉลี่ยของคะแนนแปลง ด้วยวิธีการทดสอบของ เซฟเฟ ได้ผลดังตาราง 10 ,11 และ 12 ตามลำดับ ดังนี้

ตาราง 10 การเปรียบเทียบความแตกต่างรายคู่ของคะแนนแปลงเฉลี่ย ที่ได้จากการคำนวณ ด้วยสูตร สี่วิธี ของการใช้แบบทดสอบหลักแบบเชื่อมโยง ที่ขนาดความยาว 8 ซ่อ

วิธีการคำนวณ	คะแนนแปลงเฉลี่ย ($\bar{Y}/$)	แมคแคน1 $\bar{Y}/= 34.45$	แมคแคน2 $\bar{Y}/= 26.49$	วีตครัพ1 $\bar{Y}/=33.38$	วีตครัพ2 $\bar{Y}/=31.85$
แมคแคน1	34.45	-	7.96*	1.07	2.60*
แมคแคน2	26.49		-	6.89*	5.36*
วีตครัพ1	33.38			-	1.53
วีตครัพ2	31.85				-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 10 พบว่า คะแนนแปลงเฉลี่ยของวิธีการคำนวณด้วยสูตร สี่วิธี คู่ที่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ วิธีที่ 1 ของแมคแคน กับวิธีที่ 2 ของแมคแคน, วิธีที่ 1 ของแมคแคน กับวิธีที่ 2 ของวีตครัพ, วิธีที่ 2 ของแมคแคน กับวิธีที่ 1 ของวีตครัพ และวิธีที่ 2 ของแมคแคน กับวิธีที่ 2 ของวีตครัพ นอกเหนือจากนี้ เป็นคู่ที่มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตาราง 11 การเปรียบเทียบความแตกต่างรายคู่ของคะแนนแปลงเฉลี่ย ที่ได้มาจากการคำนวณด้วยสูตร สี่วิธี ของการไร่แบบทดสอบหลักแบบเชื่อมโยง ที่ขนาดความยาว 12 ซ่อ

วิธีการคำนวณ	คะแนนแปลงเฉลี่ย ($\bar{Y}/$)	แมคแคน1 $\bar{Y}/= 31.99$	แมคแคน2 $\bar{Y}/= 39.65$	วีตดร์ฟ1 $\bar{Y}/=35.98$	วีตดร์ฟ2 $\bar{Y}/=35.96$
แมคแคน1	31.99	-	7.66*	3.99*	3.97*
แมคแคน2	39.65		-	3.67*	3.66*
วีตดร์ฟ1	35.98			-	0.02
วีตดร์ฟ2	35.96				-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 11 พบว่า คะแนนแปลงเฉลี่ยของวิธีการคำนวณด้วยสูตร สี่วิธี คู่ที่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ วิธีที่ 1 ของแมคแคน กับวิธีที่ 2 ของแมคแคน, วิธีที่ 1 ของแมคแคน กับวิธีที่ 1 ของวีตดร์ฟ, วิธีที่ 1 ของแมคแคน กับวิธีที่ 2 ของวีตดร์ฟ, วิธีที่ 2 ของแมคแคน กับวิธีที่ 1 ของวีตดร์ฟ และวิธีที่ 2 ของแมคแคน กับวิธีที่ 2 ของวีตดร์ฟ นอกเหนือจากนี้ เป็นคู่ที่มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

รูคดรฟ1	33.08			-	3.74*
รูคดรฟ2	36.82				-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 12 พบว่า คะแนนแปลงเฉลี่ยของวิธีการคำนวณด้วยสูตร ส่ววิธี คู่ที่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ วิธีที่ 1 ของแมคแคน กับวิธีที่ 2 ของแมคแคน, วิธีที่ 1 ของแมคแคน กับวิธีที่ 2 ของรูคดรฟ, วิธีที่ 2 ของแมคแคน กับวิธีที่ 1 ของรูคดรฟ, วิธีที่ 2 ของแมคแคน กับวิธีที่ 2 ของรูคดรฟ และวิธีที่ 1 ของรูคดรฟ กับวิธีที่ 2 ของรูคดรฟ นอกเหนือจากนี้ เป็นคู่ที่มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตาราง 13 การเปรียบเทียบคะแนนแปลงของการเทียบคะแนน โดยการใช้ข้อสอบร่วม
ที่มีขนาดความยาวต่างกัน ด้วยการวิเคราะห์ความแปรปรวนแบบทางเดียว แบบวัดซ้ำ

แหล่งความแปรปรวน	SS	DF	MS	F
การใช้ข้อสอบร่วม ขนาดความยาว 8 ข้อ				
<u>ภายในกลุ่ม</u>				
ภายในเซลล์	4049.93	213	19.01	
วิธีการคำนวณ	3202.55	3	1067.52	56.14**
การใช้ข้อสอบร่วม ขนาดความยาว 12 ข้อ				
<u>ภายในกลุ่ม</u>				
ภายในเซลล์	1264.41	213	5.94	
วิธีการคำนวณ	62.50	3	20.28	3.51**
การใช้ข้อสอบร่วม ขนาดความยาว 16 ข้อ				
<u>ภายในกลุ่ม</u>				
ภายในเซลล์	1329.51	213	6.24	
วิธีการคำนวณ	2510.48	3	836.83	134.07**

** มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง 13 พบว่า คะแนนแปลงของการเทียบคะแนน เมื่อใช้ข้อสอบร่วมที่มี
ขนาดความยาวต่างกัน สามขนาด โดยในแต่ละขนาดความยาว ให้ผลที่แตกต่างกันอย่าง
มีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อพิจารณาคะแนนแปลงของการเทียบคะแนน ตามขนาดความยาวในแต่ละขนาดพบว่า

1. ที่ขนาดความยาว 8 ข้อ คะแนนแปลงของการเทียบคะแนนที่ได้มาจากวิธีการคำนวณด้วยสูตร สี่วิธี มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
2. ที่ขนาดความยาว 12 ข้อ คะแนนแปลงของการเทียบคะแนนที่ได้มาจากวิธีการคำนวณด้วยสูตร สี่วิธี มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
3. ที่ขนาดความยาว 16 ข้อ คะแนนแปลงของการเทียบคะแนนที่ได้มาจากวิธีการคำนวณด้วยสูตร สี่วิธี มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อทดสอบความแตกต่างรายคู่ของค่าเฉลี่ยของคะแนนแปลง ด้วยวิธีการทดสอบของ
เซฟเฟ ได้ผลดังตาราง 14 ,15 และ 16 ตามลำดับ ดังนี้

ตาราง 14 การเปรียบเทียบความแตกต่างรายคู่ของคะแนนแปลงเฉลี่ย ที่ได้มาจากการคำนวณ
ด้วยสูตร สี่วิธี ของการใช้ข้อสอบรวม ที่ขนาดความยาว 8 ข้อ

วิธีการคำนวณ	คะแนนแปลงเฉลี่ย ($\bar{Y}/$)	แมคแคนน1 $\bar{Y}/= 32.09$	แมคแคนน2 $\bar{Y}/= 25.22$	รูตดรัฟ1 $\bar{Y}/=33.94$	รูตดรัฟ2 $\bar{Y}/=28.86$
แมคแคนน1	32.09	-	6.87*	1.85	3.23*
แมคแคนน2	25.22		-	8.72*	3.64*
รูตดรัฟ1	33.94			-	5.08*
รูตดรัฟ2	28.86				-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 14 พบว่า คะแนนแปลงเฉลี่ยของวิธีการคำนวณด้วยสูตร สี่วิธี คู่ที่มี
แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ วิธีที่ 1 ของแมคแคนน กับวิธีที่ 2 ของ
แมคแคนน, วิธีที่ 1 ของแมคแคนน กับวิธีที่ 2 ของรูตดรัฟ, วิธีที่ 2 ของวิธีของแมคแคนน กับวิธีที่ 1
ของรูตดรัฟ, วิธีที่ 2 ของแมคแคนน กับวิธีที่ 2 ของรูตดรัฟ และวิธีที่ 1 ของรูตดรัฟ กับวิธีที่ 2
ของรูตดรัฟ นอกเหนือจากนี้ เป็นคู่ที่มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตาราง 15 การเปรียบเทียบความแตกต่างรายคู่ของคะแนนแปลงเฉลี่ย ที่ได้มาจากการคำนวณ ด้วยสูตร สี่วิธี ของการใช้ข้อสอบรวม ที่ขนาดความยาว 12 ข้อ

วิธีการคำนวณ	คะแนนแปลงเฉลี่ย ($\bar{Y}/$)	แมคเคนน1 $\bar{Y}/= 33.35$	แมคเคนน2 $\bar{Y}/= 32.04$	รูตครีฟ1 $\bar{Y}/=32.63$	รูตครีฟ2 $\bar{Y}/=32.56$
แมคเคนน1	33.35	-	1.31	0.72	0.79
แมคเคนน2	32.04		-	0.59	0.52
รูตครีฟ1	32.63			-	0.07
รูตครีฟ2	32.56				-

จากตาราง 15 พบว่า คะแนนแปลงเฉลี่ยทุกคู่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตาราง 16 การเปรียบเทียบความแตกต่างรายคู่ของคะแนนแปลงเฉลี่ย ที่ได้มาจากการคำนวณด้วยสูตร สี่วิธี ของการใช้ข้อสอบรวม ที่ขนาดความยาว 16 ข้อ

วิธีการคำนวณ	คะแนนแปลงเฉลี่ย ($\bar{Y}/$)	แมคแคน1 $\bar{Y}/= 34.37$	แมคแคน2 $\bar{Y}/= 42.63$	รูดครัฟ1 $\bar{Y}/=33.08$	รูดครัฟ2 $\bar{Y}/=36.82$
แมคแคน1	34.37	-	8.26*	1.29	2.45*
แมคแคน2	42.63		-	9.55*	5.81*
รูดครัฟ1	33.08			-	3.74*
รูดครัฟ2	36.82				-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 16 พบว่า คะแนนแปลงเฉลี่ยของวิธีการคำนวณด้วยสูตร สี่วิธี คู่ที่มีแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ วิธีที่ 1 ของแมคแคน กับวิธีที่ 2 ของแมคแคน, วิธีที่ 1 ของวิธีของแมคแคน กับวิธีที่ 2 ของรูดครัฟ, วิธีที่ 2 ของแมคแคน กับวิธีที่ 1 ของรูดครัฟ, วิธีที่ 2 ของวิธีของแมคแคน กับวิธีที่ 2 ของรูดครัฟ และวิธีที่ 1 ของรูดครัฟ กับวิธีที่ 2 ของรูดครัฟ นอกเหนือจากนี้ เป็นคู่ที่มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตาราง 17 การเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี จำแนกตามการใช้แบบทดสอบหลัก ที่มีขนาดความยาวต่างกัน

แหล่งความแปรปรวน	SS	DF	MS	F
<u>ระหว่างกลุ่ม</u>				
ภายในเซลล์	305.84	426	.72	
ค่าคงที่	785.84	1	785.84	1094.57
ขนาดความยาว	1.71	2	.86	1.19
การใช้แบบทดสอบหลัก	.37	1	.37	.52
การใช้แบบทดสอบหลัก กับขนาดความยาว	9.90	2	4.95	6.90**
<u>ภายในกลุ่ม</u>				
ภายในเซลล์	185.02	1278	.14	
วิธีการคำนวณ	80.03	3	26.68	184.26**
วิธีการคำนวณ กับขนาดความยาว	14.05	6	2.34	16.17**
วิธีการคำนวณ กับการใช้แบบทดสอบหลัก	10.53	3	3.51	24.24**
วิธีการคำนวณ กับการใช้แบบทดสอบหลัก กับขนาดความยาว	6.11	6	1.02	7.04**

** มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง 17 การเปรียบเทียบความคลาดเคลื่อนมาตรฐาน ที่เกิดจากวิธีการคำนวณด้วยสูตรต่างกัน สี่วิธี กับการใช้แบบทดสอบหลักต่างกัน สองลักษณะ ที่มีขนาดความยาวต่างกัน สามขนาด พบว่า บัญชีสัมพันธ์ของความคลาดเคลื่อนมาตรฐานมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อพิจารณาความคลาดเคลื่อนมาตรฐาน ที่เกิดจากวิธีการคำนวณด้วยสูตรต่างกัน
สี่วิธี กับขนาดความยาวของแบบทดสอบหลักต่างกัน สามขนาด ก็พบว่า ปฏิสัมพันธ์ของความ
คลาดเคลื่อนมาตรฐานมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อพิจารณาความคลาดเคลื่อนมาตรฐาน ที่เกิดจากวิธีการคำนวณด้วยสูตรต่างกัน
สี่วิธี กับการใช้แบบทดสอบหลักต่างกัน สองลักษณะ ก็พบว่า ปฏิสัมพันธ์ของความคลาดเคลื่อน
มาตรฐานมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

นอกจากนี้ เมื่อพิจารณาความคลาดเคลื่อนมาตรฐาน ที่เกิดจากการใช้แบบทดสอบ
หลักต่างกัน สองลักษณะ กับขนาดความยาวต่างกัน สามขนาด พบว่า ปฏิสัมพันธ์ของความ
คลาดเคลื่อนมาตรฐาน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เช่นกัน

ตาราง 18 การเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน โดยการใช้
แบบทดสอบหลักแบบเชื่อมโยงที่มีขนาดความยาวต่างกัน ด้วยการวิเคราะห์ความแปรปรวน
แบบทางเดียว แบบวัดซ้ำ

แหล่งความแปรปรวน	SS	DF	MS	F
แบบทดสอบหลักแบบเชื่อมโยง ขนาดความยาว 8 ข้อ				
<u>ภายในกลุ่ม</u>				
ภายในเซลล์	39.02	213	.18	
วิธีการคำนวณ	36.49	3	12.16	66.40**
แบบทดสอบหลักแบบเชื่อมโยง ขนาดความยาว 12 ข้อ				
<u>ภายในกลุ่ม</u>				
ภายในเซลล์	23.56	213	.11	
วิธีการคำนวณ	11.46	3	3.82	34.52**
แบบทดสอบหลักแบบเชื่อมโยง ขนาดความยาว 16 ข้อ				
<u>ภายในกลุ่ม</u>				
ภายในเซลล์	20.46	213	.10	
วิธีการคำนวณ	26.59	3	8.86	92.27**

** มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง 18 พบว่า ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน เมื่อใช้แบบทดสอบหลักแบบเชื่อมโยงที่มีขนาดความยาวต่างกัน สามขนาด โดยในแต่ละขนาดความยาวให้ผลที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อพิจารณาความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ตามขนาดความยาว ในแต่ละขนาด พบว่า

1. ที่ขนาดความยาว 8 ข้อ ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
2. ที่ขนาดความยาว 12 ข้อ ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
3. ที่ขนาดความยาว 16 ข้อ ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อทดสอบความแตกต่างรายคู่ของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากการคำนวณด้วยสูตร สี่วิธี โดยวิธีการทดสอบของเซฟเฟ ได้ผลดังตาราง 19 ,20 และ 21 ตามลำดับ ดังนี้

ตาราง 19 การเปรียบเทียบความแตกต่างรายคู่ ของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากการคำนวณด้วยสูตร สี่วิธี ของการใช้แบบทดสอบหลักแบบเชื่อมโยง ที่ขนาดความยาว 8 ข้อ

วิธีการคำนวณ	ค่าเฉลี่ย ($\bar{X}$)	แมคแคนน1 $\bar{X} = .45$	แมคแคนน2 $\bar{X} = 1.26$	วีตดร์ฟ1 $\bar{X} = .43$	วีตดร์ฟ2 $\bar{X} = .43$
แมคแคนน1	.45	-	.81*	.02	.02
แมคแคนน2	1.26		-	.83*	.83*
วีตดร์ฟ1	.43			-	0
วีตดร์ฟ2	.43				-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 19 พบว่า ค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากการคำนวณด้วยสูตร สี่วิธี คู่ที่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ วิธีที่ 1 ของแมคแคนน กับวิธีที่ 2 ของแมคแคนน, วิธีที่ 2 ของแมคแคนน กับวิธีที่ 1 ของวีตดร์ฟ และวิธีที่ 2 ของแมคแคนน กับวิธีที่ 2 ของวีตดร์ฟ นอกเหนือจากนี้ เป็นคู่ที่มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตาราง 20 การเปรียบเทียบความแตกต่างรายคู่ ของความคลาดเคลื่อนมาตรฐานของการ
เทียบคะแนน ที่เกิดจากการคำนวณด้วยสูตร สี่วิธี ของการraiseแบบทดสอบหลักแบบเชื่อมโยง
ที่ขนาดความยาว 12 ข้อ

วิธีการคำนวณ	ค่าเฉลี่ย ($\bar{X}$)	แมคแคน1 $\bar{X} = .63$	แมคแคน2 $\bar{X} = 1.13$	รูดอล์ฟ1 $\bar{X} = .70$	รูดอล์ฟ2 $\bar{X} = .70$
แมคแคน1	.63	-	.50*	.07	.07
แมคแคน2	1.13		-	.43*	.43*
รูดอล์ฟ1	.70			-	0
รูดอล์ฟ2	.70				-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 20 พบว่า ค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน
ที่เกิดจากการคำนวณด้วยสูตร สี่วิธี คู่ที่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05
ได้แก่ วิธีที่ 1 ของแมคแคน กับวิธีที่ 2 ของแมคแคน, วิธีที่ 2 ของแมคแคน กับวิธีที่ 1 ของ
รูดอล์ฟ และวิธีที่ 2 ของแมคแคน กับวิธีที่ 2 ของรูดอล์ฟ นอกเหนือจากนี้ เป็นคู่ที่มีความ
แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตาราง 21 การเปรียบเทียบความแตกต่างรายคู่ ของความคลาดเคลื่อนมาตรฐานของการ
เทียบคะแนน ที่เกิดจากการคำนวณด้วยสูตร สี่วิธี ของการraiseแบบทดสอบหลักแบบเชื่อมเรียง
ที่ขนาดความยาว 16 ข้อ

วิธีการคำนวณ	ค่าเฉลี่ย ($\bar{X}$)	แมคแคนน1 $\bar{X} = .34$	แมคแคนน2 $\bar{X} = 1.05$	รูดอล์ฟ1 $\bar{X} = .31$	รูดอล์ฟ2 $\bar{X} = .42$
แมคแคนน1	.34	-	.71*	.03	.08
แมคแคนน2	1.05		-	.31*	.63*
รูดอล์ฟ1	.31			-	.11
รูดอล์ฟ2	.42				-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 21 พบว่า ค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน
ที่เกิดจากการคำนวณด้วยสูตร สี่วิธี คู่ที่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05
ได้แก่ วิธีที่ 1 ของแมคแคนน กับวิธีที่ 2 ของแมคแคนน, วิธีที่ 2 ของแมคแคนน กับวิธีที่ 1 ของ
รูดอล์ฟ และวิธีที่ 2 ของแมคแคนน กับวิธีที่ 2 ของรูดอล์ฟ นอกเหนือจากนี้ เป็นคู่ที่มีความ
แตกต่างกันอย่างไม่มีความนัยสำคัญทางสถิติ เช่นเดียวกันกับตาราง 19 และตาราง 20

ตาราง 22 การเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน โดยการนำข้อสอบรวมที่มีขนาดความยาวต่างกัน ด้วยการวิเคราะห์ความแปรปรวนแบบทางเดียว แบบวัดซ้ำ

แหล่งความแปรปรวน	SS	DF	MS	F
การใช้ข้อสอบรวม ขนาดความยาว 8 ข้อ				
<u>ภายในกลุ่ม</u>				
ภายในเซลล์	58.59	213	.28	
วิธีการคำนวณ	26.79	3	8.93	32.47**
การใช้ข้อสอบรวม ขนาดความยาว 12 ข้อ				
<u>ภายในกลุ่ม</u>				
ภายในเซลล์	13.62	213	.06	
วิธีการคำนวณ	2.90	3	.97	15.14**
การใช้ข้อสอบรวม ขนาดความยาว 16 ข้อ				
<u>ภายในกลุ่ม</u>				
ภายในเซลล์	29.78	213	.14	
วิธีการคำนวณ	6.48	3	2.16	15.46**

** มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง 22 พบว่า ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน เมื่อใช้ข้อสอบรวมที่มีขนาดความยาวต่างกัน สามขนาด ครอบคลุมในแต่ละขนาดความยาว ให้ผลที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อพิจารณาความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ตามขนาดความยาว ในแต่ละขนาด พบว่า

1. ที่ขนาดความยาว 8 ข้อ ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
2. ที่ขนาดความยาว 12 ข้อ ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
3. ที่ขนาดความยาว 16 ข้อ ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อทดสอบความแตกต่างรายคู่ของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากวิธีการคำนวณด้วยสูตร สี่วิธี โดยวิธีการทดสอบของเซฟเฟ ได้ผลดังตาราง 23 , 24 และ 25 ตามลำดับ ดังนี้

ตาราง 23 การเปรียบเทียบความแตกต่างรายคู่ ของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากการคำนวณด้วยสูตร สี่วิธี ของการวิจัยข้อสอบรวม ที่ขนาดความยาว 8 ข้อ

วิธีการคำนวณ	ค่าเฉลี่ย ($\bar{X}$)	แมคแคนน1 $\bar{X} = .51$	แมคแคนน2 $\bar{X} = 1.19$	วู้ดครัฟ1 $\bar{X} = .41$	วู้ดครัฟ2 $\bar{X} = .59$
แมคแคนน1	.51	-	.68*	.10	.08
แมคแคนน2	1.19		-	.78*	.60*
วู้ดครัฟ1	.41			-	.18
วู้ดครัฟ2	.59				-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 23 พบว่า ค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากการคำนวณด้วยสูตร สี่วิธี คู่ที่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ วิธีที่ 1 ของแมคแคนน กับวิธีที่ 2 ของแมคแคนน, วิธีที่ 2 ของแมคแคนน กับวิธีที่ 1 ของวู้ดครัฟ และวิธีที่ 2 ของแมคแคนน กับวิธีที่ 2 ของวู้ดครัฟ นอกเหนือจากนี้ เป็นคู่ที่มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตาราง 24 การเปรียบเทียบความแตกต่างรายคู่ ของความคลาดเคลื่อนมาตรฐานของการ
เทียบคะแนน ที่เกิดจากการคำนวณด้วยสูตร สี่วิธี ของการนำข้อสอบรวม ที่ขนาดความยาว
12 ข้อ

วิธีการคำนวณ	ค่าเฉลี่ย ($\bar{X}$)	แมคเคนน1 $\bar{X} = .76$	แมคเคนน2 $\bar{X} = .71$	รูตครีฟ1 $\bar{X} = .53$	รูตครีฟ2 $\bar{X} = .54$
แมคเคนน1	.76	—	.05	.23	.22
แมคเคนน2	.71		—	.18	.17
รูตครีฟ1	.53			—	.01
รูตครีฟ2	.54				—

จากตาราง 24 พบว่า ค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน
ที่เกิดจากการคำนวณด้วยสูตร สี่วิธี ทุกคู่มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตาราง 25 การเปรียบเทียบความแตกต่างรายคู่ ของความคลาดเคลื่อนมาตรฐานของการ
เทียบคะแนน ที่เกิดจากการคำนวณด้วยสูตร สี่วิธี ของการใช้อ้างอิงข้อสอบรวม ที่ขนาดความยาว
16 ข้อ

วิธีการคำนวณ	ค่าเฉลี่ย ($\bar{X}$)	แมคแคนน1 $\bar{X} = .78$	แมคแคนน2 $\bar{X} = .90$	วีคตอร์ฟ1 $\bar{X} = .79$	วีคตอร์ฟ2 $\bar{X} = .50$
แมคแคนน1	.78	-	.12	.01	.28*
แมคแคนน2	.90		-	.11	.40*
วีคตอร์ฟ1	.79			-	.29*
วีคตอร์ฟ2	.50				-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 25 พบว่า ค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน
ที่เกิดจากการคำนวณด้วยสูตร สี่วิธี คู่ที่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05
ได้แก่ วิธีที่ 1 ของแมคแคนน กับวิธีที่ 2 ของวีคตอร์ฟ, วิธีที่ 2 ของแมคแคนน กับวิธีที่ 2 ของ
วีคตอร์ฟ และวิธีที่ 1 ของวีคตอร์ฟ กับวิธีที่ 2 ของวีคตอร์ฟ นอกเหนือจากนี้ เป็นคู่ที่มีความ
แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

สรุป อภิปรายผล และข้อเสนอแนะ

ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อเปรียบเทียบคะแนนแปลง อันเป็นผลจากการแปลงคะแนน จากวิธีการเทียบคะแนนแบบเส้นตรง สี่วิธี ว่าแตกต่างกันหรือไม่
2. เพื่อเปรียบเทียบคะแนนแปลง อันเป็นผลจากการแปลงคะแนน จากวิธีการเทียบคะแนนแบบเส้นตรง ที่ใช้แบบทดสอบหลักแบบเชื่อมโยง กับแบบข้อสอบรวม ว่าแตกต่างกันหรือไม่
3. เพื่อเปรียบเทียบคะแนนแปลง อันเป็นผลจากการแปลงคะแนน จากวิธีการเทียบคะแนนแบบเส้นตรง ที่มีขนาดความยาวของแบบทดสอบหลักแบบเชื่อมโยง กับแบบข้อสอบรวม ในสัดส่วน .2 ,.3 และ .4 ว่าแตกต่างกันหรือไม่
4. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐาน ของการเทียบคะแนนแบบเส้นตรง จากวิธีการคำนวณ สี่วิธีว่าแตกต่างกันหรือไม่

สมมุติฐานของการศึกษาค้นคว้า

1. คะแนนแปลง ที่ได้จากการสอบด้วยวิธีเทียบคะแนนแบบเส้นตรง สี่วิธี แตกต่างกัน
2. คะแนนแปลง ที่ได้จากการสอบด้วยวิธีเทียบคะแนนแบบเส้นตรง ที่ใช้แบบทดสอบหลักแบบเชื่อมโยง กับแบบข้อสอบรวม แตกต่างกัน
3. คะแนนแปลง ของแบบทดสอบสำหรับเทียบคะแนน ที่มีแบบทดสอบหลักขนาดความยาวต่างกัน แตกต่างกัน
 - 3.1. คะแนนแปลง ของแบบทดสอบสำหรับเทียบคะแนน ที่มีแบบทดสอบหลักแบบเชื่อมโยงขนาดความยาวต่างกัน แตกต่างกัน
 - 3.2. คะแนนแปลง ของแบบทดสอบสำหรับใช้เทียบคะแนน ที่มีแบบทดสอบหลักแบบข้อสอบรวมขนาดความยาวต่างกัน แตกต่างกัน

4. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่เกิดจากวิธีการเทียบคะแนนแบบเส้นตรง สี่วิธี แตกต่างกัน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาครั้งนี้ เป็นนักเรียนพยาบาลชั้นปีที่ 1 และชั้นปีที่ 2 ของหลักสูตรพยาบาลศาสตร์ และพยาบาลศาสตรระดับต้น จากวิทยาลัยพยาบาลบรมราชชนนีกทม. วิทยาลัยพยาบาลบรมราชชนนีนครบุรี และวิทยาลัยพยาบาลบรมราชชนนีสระบุรี จำนวน 3 แห่ง สังกัดสำนักงานปลัดกระทรวงสาธารณสุข กระทรวงสาธารณสุข ในเขตภาคกลาง ปีการศึกษา 2538 ได้มาจากการสุ่มอย่างง่าย มีนักศึกษาจำนวนทั้งสิ้น 864 คน จัดเป็นกลุ่มสำหรับเทียบคะแนน 12 กลุ่ม

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบวัดการคิดเอกลัแบบสัมพันธ์ แบ่งเป็น 2 ตอน ตอนที่ 1 มีเนื้อหาเป็นภาพ และตอนที่ 2 มีเนื้อหาเป็นภาษา จัดทำเป็นแบบทดสอบทั้งหมด 3 พอร์ม คือ พอร์ม A พอร์ม B และพอร์ม C จัดพอร์ม A กับพอร์ม B ให้เป็นแบบทดสอบสำหรับเทียบคะแนน และจัดพอร์ม C ให้เป็นแบบทดสอบหลัก โดยการใช้เป็นแบบเชื่อมโยง กับเป็นข้อสอบรวมที่มีขนาดความยาวต่างกัน สามขนาด คือ คิดเป็นสัดส่วน

วิธีดำเนินการเก็บรวบรวมข้อมูล

1. ใช้แผนแบบของการใช้แบบทดสอบหลักที่มีประชากรแตกต่างกัน (CINEP) โดยจัดแบบทดสอบสำหรับเทียบคะแนน สองฟอร์ม คือ ฟอร์ม A กับฟอร์ม B สอนกับกลุ่มตัวอย่างตามแผนแบบที่วางไว้ โดยให้มีการใช้แบบทดสอบหลักกับกลุ่มตัวอย่างสำหรับเทียบคะแนนทั้งสิบสองกลุ่ม
2. นำไปทดสอบกับกลุ่มเทียบคะแนน โดยใช้แบบทดสอบฟอร์ม A สอนกับนักศึกษาพยาบาลชั้นปีที่ 1 จำนวน 6 กลุ่ม และใช้แบบทดสอบฟอร์ม B สอนกับนักศึกษาพยาบาลชั้นปีที่ 2 จำนวน 6 กลุ่ม และแต่ละกลุ่มของกลุ่มเทียบคะแนนใช้แบบทดสอบหลักที่มีลักษณะต่างกัน คือ เป็นแบบทดสอบเชื่อมโยง กับเป็นข้อสอบรวม และใช้ขนาดความยาวของแบบทดสอบหลักต่างกัน
3. นำคะแนนดิบที่ได้จากการสอบมาวิเคราะห์ โดยการแปลงคะแนนจากแบบทดสอบฟอร์ม B (คะแนน Y) ให้เป็นคะแนนแปลง (Y/) ด้วยวิธีการคำนวณด้วยสูตร สี่วิธี คือ วิธีของแมคแคน สองวิธี และวิธีของรูดอล์ฟ สองวิธี แล้วจึงนำคะแนนแปลงมาวิเคราะห์หาความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน จากนั้นจึงหาความแปรปรวนของคะแนนแปลงของการเทียบคะแนน โดยวิเคราะห์หองศ์ประกอบแบบวัดซ้ำ และหาความแปรปรวนของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน โดยวิเคราะห์หองศ์ประกอบแบบวัดซ้ำ

สรุปผลการวิเคราะห์ข้อมูล

1. คะแนนแปลงจากการเทียบคะแนน ที่ได้จากการคำนวณด้วยสูตร สี่วิธี พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
2. คะแนนแปลงที่ได้จากการใช้แบบทดสอบหลักแบบเชื่อมโยง กับการใช้ข้อสอบรวม พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
3. คะแนนแปลงที่ได้จากการใช้ขนาดความยาว สามขนาด ของแบบทดสอบหลักแบบเชื่อมโยง พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
4. คะแนนแปลงที่ได้จากการใช้ขนาดความยาว สามขนาด ของการใช้ข้อสอบรวม พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ยกเว้นที่ขนาดความยาว 12 ข้อ
5. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่มาจากการคำนวณด้วยสูตร สี่วิธี พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

อภิปรายผล

1. ผลการเปรียบเทียบการเทียบคะแนน จากการใช้วิธีคำนวณด้วยสูตร สี่วิธี คือ วิธีที่ 1 ของแมคแคน, วิธีที่ 2 ของแมคแคน, วิธีที่ 1 ของวู้ดดรัฟ และวิธีที่ 2 ของวู้ดดรัฟ ให้ผลการเทียบคะแนนที่แตกต่างกัน โดยจากผลการเปรียบเทียบรายคู่ พบว่า วิธีที่ 1 ของแมคแคน กับวิธีที่ 2 ของแมคแคน ส่วนใหญ่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ทั้งการใช้แบบทดสอบหลักสองลักษณะ และที่ขนาดความยาว 8 ข้อ กับ 16 ข้อ ซึ่งขัดแย้งกับผลการศึกษาของแมคแคน โดยแมคแคน (MacCann.1989 : 276) ได้กล่าวถึงผลการใช้วิธีการคำนวณด้วยสูตรของแมคแคนไว้ว่า ไม่ว่าจะใช้วิธีที่ 1 หรือวิธีที่ 2 ของแมคแคน ก็ให้ผลไม่ต่างกันมาก เมื่อใช้กับกลุ่มเทียบคะแนนที่คล้ายคลึงกัน โดยที่วิธีที่ 1 ของแมคแคนเป็นการใช้สมการถดถอยเพื่อทำนายคะแนน Y จากคะแนนของแบบทดสอบหลัก (Z) และวิธีที่ 2 ของแมคแคนเป็นการใช้สมการถดถอย เพื่อทำนายคะแนนของแบบทดสอบหลัก (Z) จากคะแนน Y ซึ่งผลการศึกษาที่ให้ผลแตกต่างกัน อาจเพราะว่ากลุ่มเทียบคะแนนที่แมคแคนใช้นั้น แต่ละกลุ่มมีจำนวนมาก คือ มีจำนวนตั้งแต่ 322 ถึง 2,479 คน และผลของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนที่เกิดจากการคำนวณของแมคแคนเองนั้น ก็พบว่า ส่วนใหญ่ตกอยู่ในระดับที่น่าพอใจอย่างยิ่ง

ส่วนวิธีการคำนวณด้วยสูตรของวู้ดดรัฟ 2 วิธี เมื่อทดสอบความแตกต่างรายคู่ พบว่า ส่วนใหญ่ให้ผลแตกต่างกันอย่างไม่มีความสำคัญทางสถิติที่ระดับ .05 โดยวู้ดดรัฟ (Woodruff. 1986 : 256) ได้สรุปถึงวิธีการคำนวณด้วยสูตรของวู้ดดรัฟสองวิธีว่า เป็นวิธีการที่ค่อนข้างเหมือนกัน เพราะเป็นวิธีการเทียบคะแนนแบบเส้นตรง ที่มีข้อตกลงของการใช้แบบทดสอบเป็นแบบจำลองคะแนนจริงสัมพันธ์เช่นเดียวกัน มีค่าความเชื่อมั่นของแบบทดสอบที่เท่าเทียมกัน (MacCann.1989 : 275 ; citing Woodruff.1986)

2. เมื่อใช้แบบทดสอบหลักที่ต่างกัน คือ ใช้เป็นแบบทดสอบหลักแบบเชื่อมโยงกับใช้เป็นข้อสอบรวม ผลการเทียบคะแนนต่างกัน ซึ่งสอดคล้องกับคำกล่าวของแมคแคน (MacCann.1989: 263) ที่กล่าวว่า การเทียบคะแนนแบบเส้นตรง มักนิยมใช้ในลักษณะของแบบทดสอบเชื่อมโยง โดยให้เหตุผลว่า หากแบบทดสอบฟอร์ม A ยากกว่า แบบทดสอบฟอร์ม B เมื่อกลุ่มที่ 1 ได้ทำแบบทดสอบฟอร์ม A และกลุ่มที่ 2 ได้ทำแบบทดสอบฟอร์ม B จะทำให้กลุ่มที่ 1 เกิดการเสียผลประโยชน์ เมื่อนำคะแนนเดิมมาเปรียบเทียบกัน

3. เมื่อใช้ขนาดความยาวของแบบทดสอบหลักที่ต่างกัน คือ สัดส่วน .2 ,.3 และ .4 ให้ผลแตกต่างกัน ซึ่งจากการศึกษาของแมคแคน (MacCann.1989 : 273) ได้ใช้ขนาดความยาวของแบบทดสอบหลัก ตั้งแต่สัดส่วน .25 ถึง 4.0 พบว่า การเทียบคะแนนมีความแตกต่างกัน เมื่อใช้ขนาดความยาวของแบบทดสอบหลักต่างกัน นอกจากนี้ แมคแคน (MacCann.1989 : 275) ยังพบว่า ขนาดความยาวของแบบทดสอบหลักมีความสัมพันธ์กับค่าความเชื่อมั่น เนื่องจากขนาดความยาวยิ่งมากจะให้ผลค่าความเชื่อมั่นที่สูงกว่า วิธีที่ 1 ของแมคแคนจะใช้ได้ดีเมื่อแบบทดสอบพอร์ม A (คะแนน X) มีความยาวมากกว่าแบบทดสอบหลัก กล่าวคือ แบบทดสอบหลักมีจำนวนข้อน้อย จะให้ผลดีกว่าเพราะเป็นการทำนายคะแนน X จากคะแนน Z

4. การประเมินความถี่ของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน จากการนำความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน มาจัดเป็นกลุ่มตามระดับความพอใจ พบว่า ส่วนใหญ่ให้ผลอยู่ในระดับที่ไม่พอใจอย่างยิ่ง (ระดับที่ 5) ของแต่ละกลุ่มเทียบคะแนน ซึ่งให้ผลต่างจากการศึกษาของแมคแคน (MacCann.1989:270) ที่ให้ผลของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ส่วนใหญ่ตกอยู่ในระดับที่ไม่พอใจอย่างยิ่ง จากการสอบทั้งหมด 26 ครั้ง โดยพบว่า เมื่อใช้วิธีที่ 1 ของแมคแคน ให้ผลในระดับที่ 1 เท่ากับ 17 ครั้ง, ระดับที่ 2 เท่ากับ 8 ครั้ง และระดับที่ 3 เท่ากับ 1 ครั้ง เมื่อใช้วิธีที่ 2 ของแมคแคน ให้ผลในระดับที่ 1 เท่ากับ 14 ครั้ง, ระดับที่ 2 เท่ากับ 7 ครั้ง, ระดับที่ 3 เท่ากับ 2 ครั้ง และระดับที่ 4 เท่ากับ 3 ครั้ง ทั้งนี้เนื่องจากว่า กลุ่มตัวอย่างที่แมคแคนใช้มีจำนวนมาก และแตกต่างกัน คือ ตั้งแต่ 322 คน ถึง 2,479 คน (MacCann.1989 :271) และโคป (Cope. 1987 : 146) ก็ได้กล่าวไว้ในงานวิจัยของเขา ซึ่งแสดงให้เห็นถึงความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน (RMSE) กับความคลาดเคลื่อนของการเทียบคะแนน (BIAS) โดยเปรียบเทียบระหว่างวิธีการเทียบคะแนนที่เป็นแบบถ่วงน้ำหนัก 5 วิธี กับการใช้สายโยง (Chain) และบอกว่า กลุ่มตัวอย่างจำนวนมากจะให้ผลของความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ที่ง่ายต่อการแสดงให้ดู หรือมีค่าน้อย กลุ่มตัวอย่างที่ใช้มีจำนวนตั้งแต่ 10,450 ถึง 22,562 คน เช่นเดียวกับการสรุปของโคเลนและวิทนี (Kolen and Whitney.1982 : 291) ที่อภิปรายผลว่า การศึกษาจะให้ผลของความเพียงพอ (Adequacy) ได้ มักต้องอาศัยปัจจัยหลายอย่างร่วมกัน ทั้งลักษณะของแบบทดสอบ แผนแบบของการเก็บรวบรวมข้อมูล และขนาดของกลุ่มตัวอย่าง

ข้อเสนอแนะ

ในการศึกษาค้นคว้าครั้งนี้ เป็นการศึกษาเพื่อเปรียบเทียบการเทียบคะแนนแบบเส้นตรงที่ใช้วิธีการคำนวณต่างกัน และศึกษาสถานการณ์การใช้แบบทดสอบหลัก กับขนาดความยาวของแบบทดสอบหลักต่างกัน โดยมีข้อตกลงของการเทียบคะแนนนี้ว่า จะต้องใช้แบบทดสอบที่เป็นแบบจำลองของคะแนนจริงสัมพันธ์ มีค่าความเชื่อมั่นของแบบทดสอบเท่าเทียมกัน ดังนั้น แนวทางของการศึกษานี้ครั้งต่อไปที่ผู้วิจัยคิดว่าน่าจะได้มีการศึกษา จึงขอเสนอแนะไว้ดังนี้

1. หาความเพียงพอ จากการใช้นิยามของกุ่มตัวอย่างที่มีจำนวนไม่มาก โดยนำกลุ่มตัวอย่างที่ใช้เป็นกลุ่มเทียบคะแนน ที่มีขนาดต่างกัน เพื่อความสะดวกในการศึกษากครั้งต่อไป และลดความสิ้นเปลืองทั้งหลายที่จำเป็น
2. เปรียบเทียบกลุ่มเทียบคะแนนที่มีความสามารถแตกต่างกัน เพื่อศึกษาคะแนนแปลงที่ได้จากการสอบที่ไม่เท่าเทียมกัน โดยดูว่าการใช้วิธีการคำนวณวิธีใด จะให้ผลที่เป็นไปได้มากที่สุด มีค่าความคลาดเคลื่อนมาตรฐานอยู่ในเกณฑ์ที่น่าพอใจมากที่สุด
3. เปรียบเทียบวิธีการเทียบคะแนนแบบเส้นตรง โดยใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนหลาย ๆ วิชา และเปรียบเทียบในลักษณะเนื้อหาที่เป็นตัวแทนกับเนื้อหาที่ไม่เป็นตัวแทน เพื่อยืนยันวิธีการเทียบคะแนนแบบเส้นตรง

บรรณานุกรม

บรรณานุกรม

- ชาณุวิทย์ เทียมบุญประเสริฐ. การวัดความถนัด. กรุงเทพฯ : สำนักทดสอบทางการศึกษา และจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2528.
- รุจีพ พงษ์สมบูรณ์. การเปรียบเทียบการเทียบมาตรฐานระหว่างรูปแบบที่ใช้ผู้สอบร่วมกับรูปแบบที่ใช้แบบสอบรวม. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2528. อัดสำเนา.
- ทองหล่อ วิภาวิน. การวัดความถนัด. กรุงเทพฯ : สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2524.
- บุญเชิด ภิญโญอนันตพงษ์. การศึกษาเปรียบเทียบองค์ประกอบของสมรรถภาพด้านการคิด เอกนัยทางภาษาตามทฤษฎีของกิลฟอร์ดกับผลสัมฤทธิ์ทางการเรียน. วิทยานิพนธ์ คศ.ม. กรุงเทพฯ : วิทยาลัยวิชาการศึกษา ประสานมิตร, 2517. อัดสำเนา.
- _____. การวัดและประเมินผลการศึกษา. กรุงเทพฯ : คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2519.
- _____. "Congeneric Part Reliability," วารสารการวัดผลการศึกษา. 12 (34): 28-32 ; พฤษภาคม-สิงหาคม 2533.
- _____. "เทคนิคการประมาณค่าความเชื่อถือได้ของแบบทดสอบที่แบ่งส่วนย่อยตามแบบจำลองคะแนนจริงสัมพันธ์ (Congeneric Model)," วารสารการวัดผลการศึกษา. 14(42): 8-20 ; มกราคม-เมษายน 2536.
- _____. "แบบทดสอบคะแนนจริงสัมพันธ์ : วิธีเทียบคะแนนสอบ," เอกสารทางวิชาการ การวัดผลการศึกษา. กรุงเทพฯ : คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2538.
- _____. "การเทียบคะแนนสอบแบบเส้นตรง ตามแบบจำลองคะแนนจริงสัมพันธ์สองวิธี," เอกสารทางวิชาการการวัดผลการศึกษา. กรุงเทพฯ : คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2538.

- ภาวิณี ศรีสุขวัฒนานนท์. การเปรียบเทียบผลจากการใช้รูปแบบการเทียบมาตรฐานที่ต่างกัน
เมื่อแบบสอบรวมมีความยาวต่างกัน. วิทยานิพนธ์ ค.ศ. กรุงเทพฯ : จุฬาลงกรณ์
มหาวิทยาลัย, 2529. อัดสำเนา.
- เรวดี อินทสระ. การเปรียบเทียบความคลาดเคลื่อนของการเทียบมาตรฐานระหว่างรูปแบบ
อิงทฤษฎีการตอบข้อสอบกับรูปแบบการใช้เทคนิคการวิเคราะห์ห้องประกอบ.
ปริญญาณิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร,
2530. อัดสำเนา.
- วิรัช วรรณรัตน์. "การเทียบมาตรฐาน (Equating)," วารสารการวัดผลการศึกษา. 11
(29) : 69-73 ; กันยายน-ธันวาคม 2530.
- สงบ ลักษณะ. "การเทียบระดับคะแนนระหว่างแบบทดสอบ," วารสารการวัดผลการศึกษา.
4(2) : 21-32 ; กันยายน-ธันวาคม 2525.
- สุจินดา ผ่องอักษร. การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนในกลุ่มทักษะของ
นักเรียนชั้นประถมศึกษาปีที่ 6 ที่จบตามหลักสูตรประถมศึกษา พุทธศักราช 2521
ในช่วงระยะเวลา 3 ปีการศึกษาโดยใช้การเทียบคะแนนรูปแบบราสส์.
ปริญญาณิพนธ์ กศ.ค. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร,
2533. อัดสำเนา.
- สุรัตน์ ขวัญบุญจันทร์. การสร้างแบบทดสอบและตารางเทียบคะแนนในแนวนอนวิชาวิทยา
ศาสตร์ที่วิเคราะห์ด้วยราสส์โมเดล. ปริญญาณิพนธ์ กศ.ม. กรุงเทพฯ :
มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2531. อัดสำเนา.
- อาภรณ์ กาญจนกิจโรจน. การสร้างแบบทดสอบและตารางเทียบคะแนนตามแนวตั้งในวิชา
คณิตศาสตร์. ปริญญาณิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ
ประสานมิตร, 2531. อัดสำเนา.

- Angoff, W.H. "Technical Problems of Obtaining Equivalent Scores on Test," Principle of educational and psychology measurement. p. 84-87. 2ed. New York : Rand Mc Nally and company, 1969.
- Angoff, W.H. "Scale, Norm and Equivalent Scores," Education Measurement. p. 508-600. 2nd ed. Washington D.C. : American Council on Educational, 1984.
- Angoff, W.H. "Summary and Derivation of Equating Methods Used at ETS," Test Equating. p. 55-70. New York : Academic Press, 1982.
- Angoff, W.H. "Technical and Practical Issues in Equating : A Discussion of Four Papers," Applied Psychological Measurement. 11(3) : 291-300 ; 1987.
- Bianchini, J.C. and P.C.Loret. "Anchor test study : the long and the short of it," Journal of Education Measurement. 12(3): 201-213 ; 1975.
- Braun, Henry I. and P.W.Holland. "Observed-Score Test Equating : A Mathematical Analysis of Some ETS Equating Procedures," Test Equating. p.9-50. New York : Academic Press, 1982.
- Budescu, David."Efficiency of Linear Equating as a Function of the Length of the Anchor Test,"Journal of Education Measurement. 22(1) : 13-20 ; 1985.
- Bykit, D.L. Element of Statistics. 2nded. NewYork : Litton Educational Publishing Inc.,1975.
- Cope, Ronald T. "How Well Do the Angoff Design V Linear Equating Methods Compare With the Tucker and Levine Methods," Applied Psychological Measurement. 11(2) : 143-149 ; 1987.

- Cook, L.L. and N.S. Petersen. "Problems 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.
- Dayton, C.M. The design of education experiments. New York : McGraw-Hill, Inc., 1970.
- Flanagan, T.C. "Unit, Scores and Norm," Education Measurement. p. 695-763. Washington D.C. : American Council on Education, 1951.
- Guildford, J.P. The Nature of Human Intelligence. New York : McGraw-Hill, Inc., 1967.
- Gulliksen, H. Theory of Mental Tests. New York : John Wiley and Sons, 1950.
- Klein, L.W. and D.Jarjoura. "The Importance of Content Representation for Common-Item Equating with Non-random Group," Journal of Education Measurement. 22(3) : 197-206 ; 1985.
- Kolen, M.J. "Standard Errors of Tucker Equating," Applied Psychological Measurement. 9(2) : 209-22 ; 1985.
- Kolen, M.J. and R.L.Brennan. "Linear Equating Model for the Common-Item Nonequivalent-populations Design," Applied Psychological Measurement. 11(3) : 263-277 ; 1987.
- Kolen, M.J. and D.R. Whitney. "Comparison of Four Procedures for Equating the tests of General Educational Development," Journal of Educational Measurement. 19(4) : 279-294 ; 1982.

- Lindquist, E.F. "Equating Scores on Non-Parallel Test," Principle of Education and Psychological Measurement. p. 74-84. 2ed.
New York : Rand McNally and company, 1969.
- Lennon, R.T. "Equating Non-Parallel Tests," Principle of education and psychological Measurement. p. 84-90. 2ed. New York :
Rand Mc Nally and company, 1969.
- MacCann, R.A. "A comparison of two observed-scores equating methods that assume equally reliable, congeneric test," Applied Psychological Measurement. 13(3) : 236-276 ; 1989.
- Petersen, N.S., M.J.Kolen and H.D. Hoover. "Scaling, Norming and Equting," Educational measurement. p. 221-262. 3rd ed.
New York : Macmillan, 1981.
- Petersen, N.S., G.L. Macro and E.E. Stewart. "A test of the Adequacy of Linear score equating Methods," Test Equating. p. 71-135.
New York : Academic Press, 1982.
- Woodruff, D. "Deriviations of observed score linear equating methods base on test scores models for the common item nonequivalent populations design," Journal of Education Statistic. 11(4) : 245-257 ; 1986.

ภาคผนวก

ภาคผนวก ก

เครื่องมือที่ใช้ในการศึกษา

แบบทดสอบวัดการคิดเอกลัแบบสัมพันธ์ ฉบับที่ 1 (ฟอร์ม A)

คำชี้แจง แบบทดสอบฉบับนี้ มีทั้งหมด 40 ข้อ แบ่งเนื้อหาออกเป็น 2 ตอน คือ

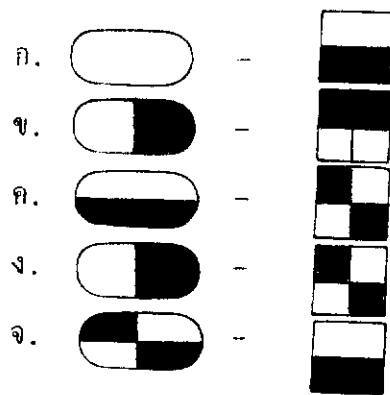
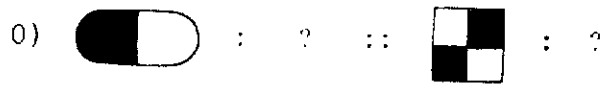
ตอนที่ 1 มีเนื้อหาเป็นภาพ

ตอนที่ 2 มีเนื้อหาเป็นภาษา

ให้นักศึกษาทำแบบทดสอบทั้งหมด โดยเลือกคำตอบที่ถูกต้องที่สุดเพียงคำตอบเดียวลงใน

กระดาษคำตอบที่จัดเตรียมไว้ ดังตัวอย่าง

ตอนที่ 1 จงเลือกคำตอบที่เป็นภาพตรงข้ามกับภาพที่กำหนดให้ 2 ภาพ



คำตอบที่ถูกต้อง คือ ข้อ ง.

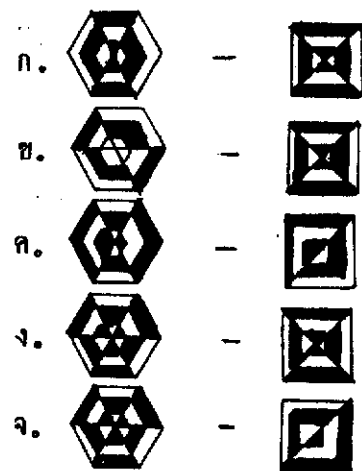
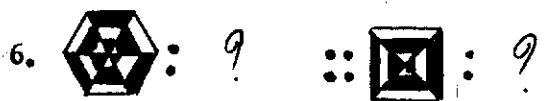
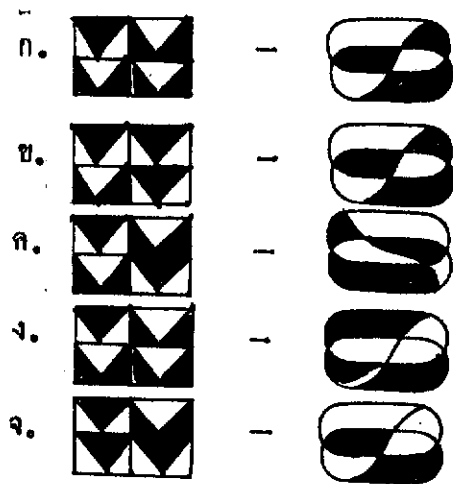
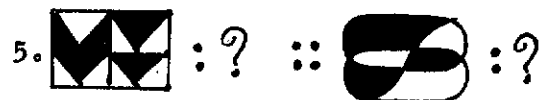
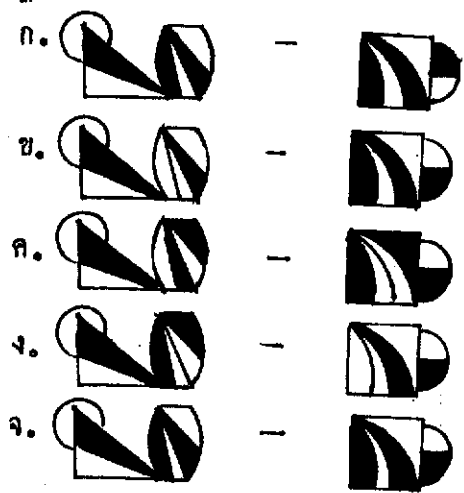
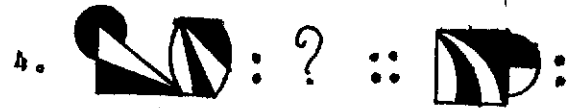
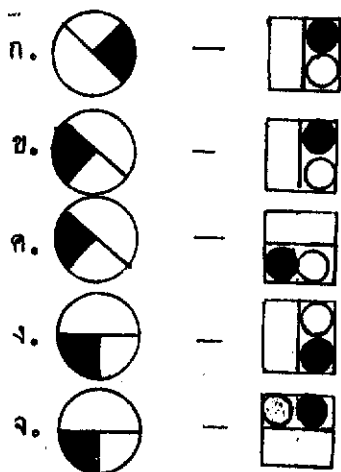
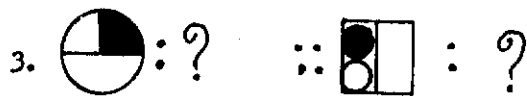
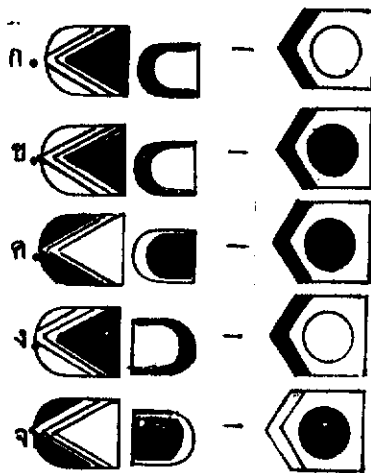
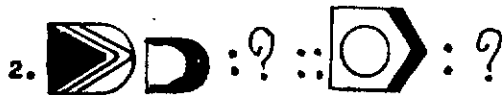
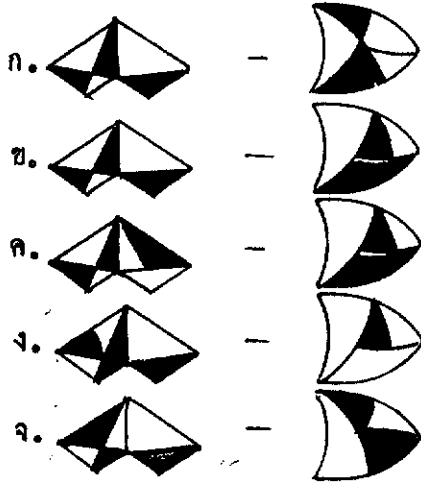
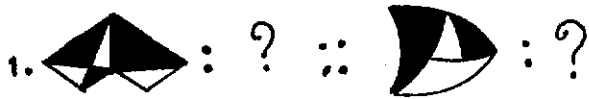
ตอนที่ 2 จงเลือกคำตอบที่เป็นคำตรงข้ามกับคำที่กำหนดให้ 2 คำ

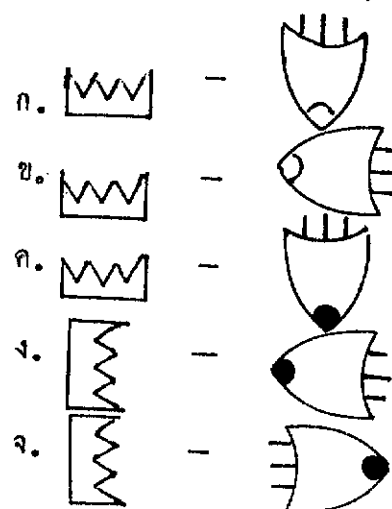
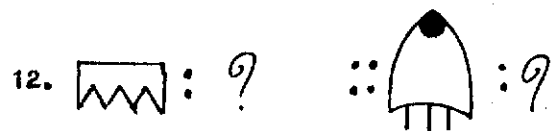
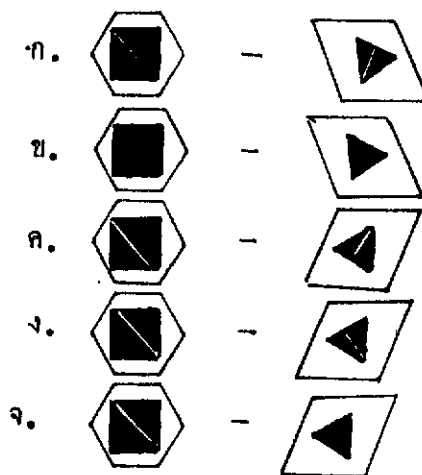
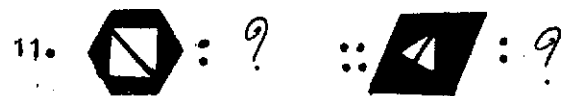
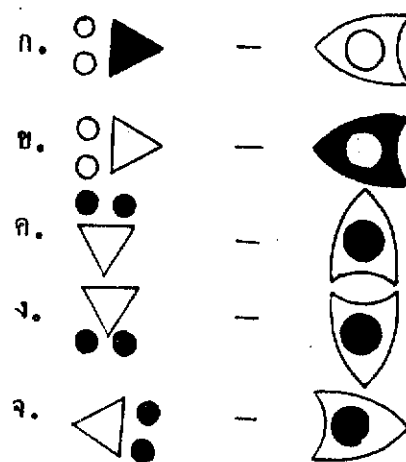
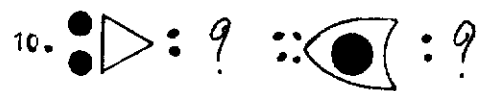
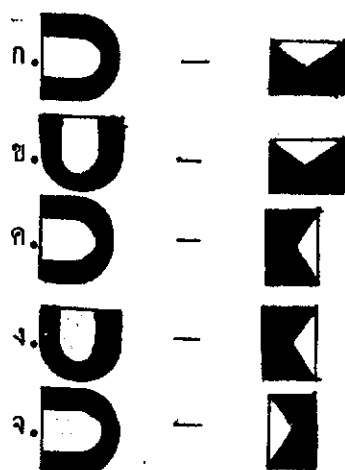
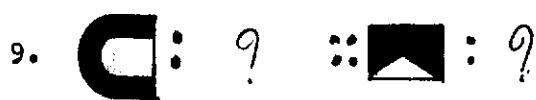
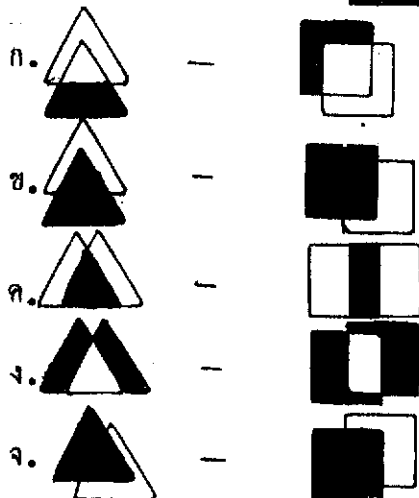
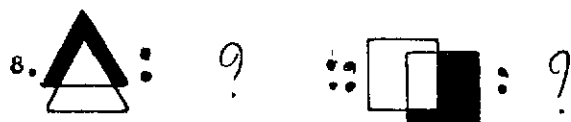
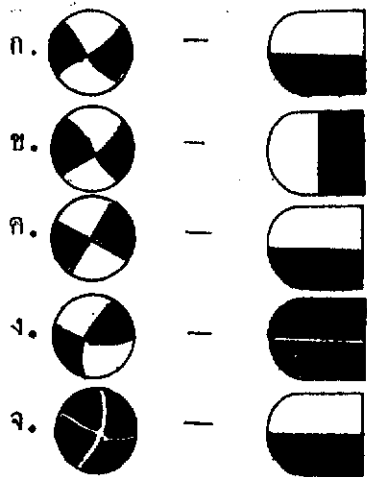
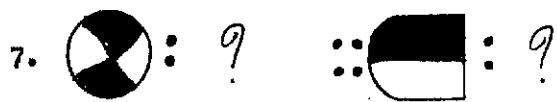
00) สั้น : ? :: กลม : ?

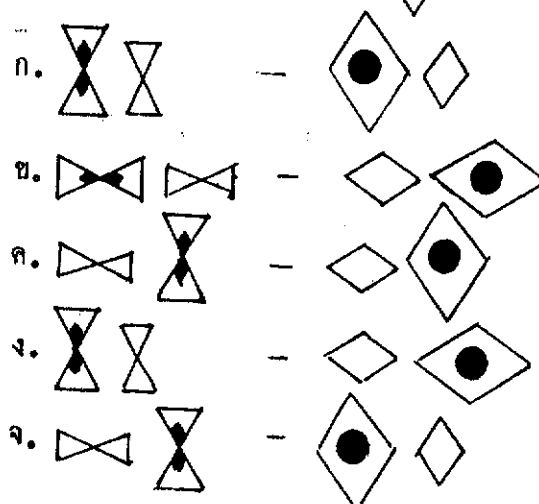
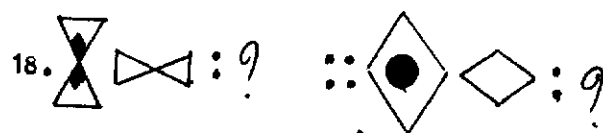
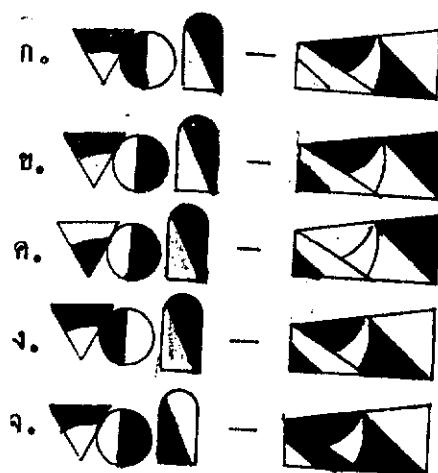
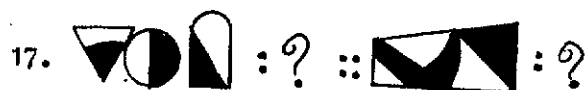
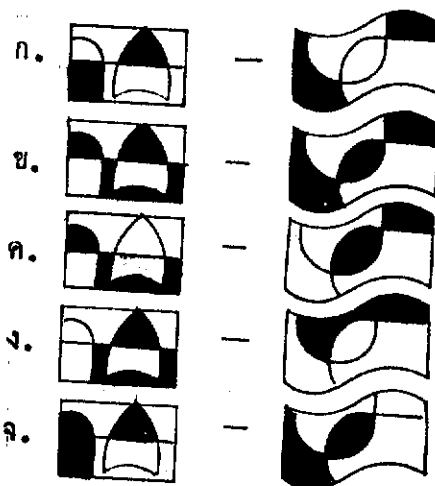
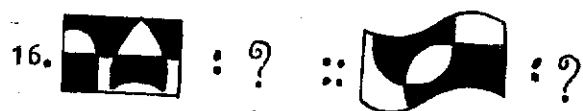
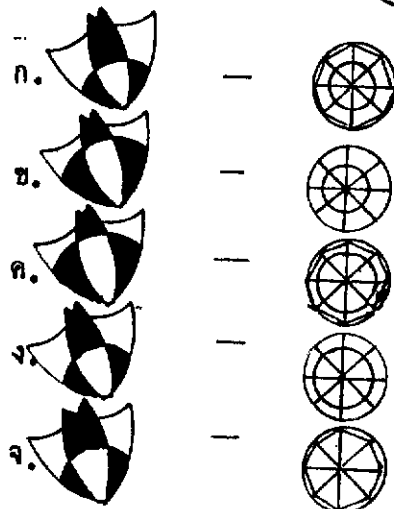
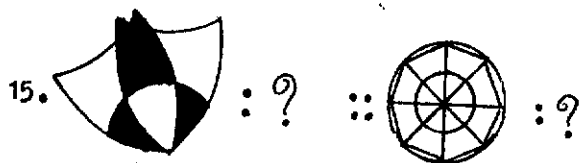
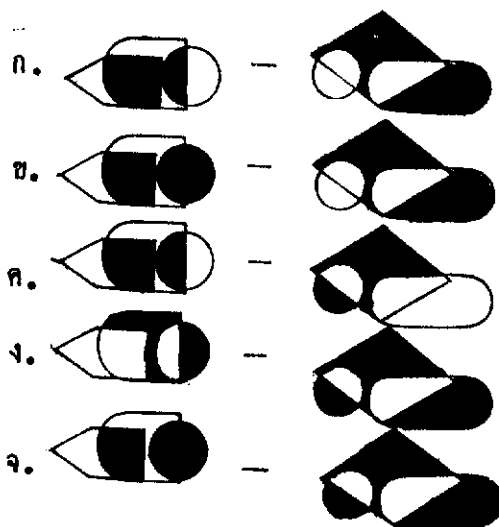
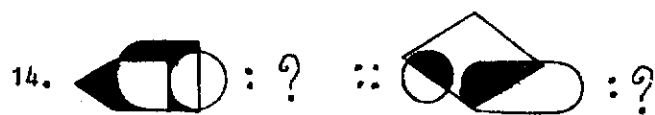
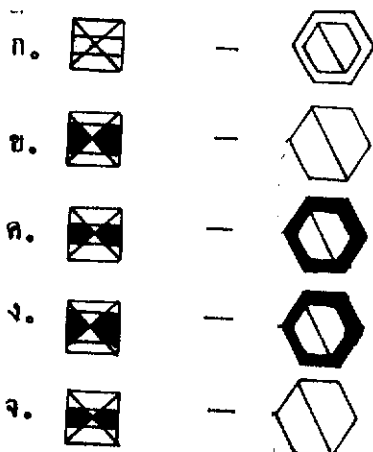
- ก. สูง - หักมุม
 ข. ยาว - เหลื่อม
 ค. แอบ - กว้าง
 ง. ต่ำ - เร็ว
 จ. เล็ก - รั

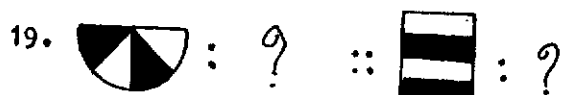
คำตอบที่ถูกต้อง คือ ข้อ ข.

ตอนที่ 1 จงเลือกคำตอบที่เป็นภาพตรงข้ามกับภาพที่กำหนดให้ 2 ภาพ









-



-



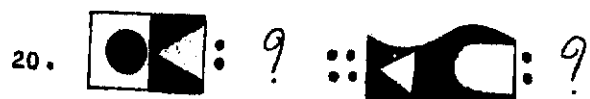
-



-



-



-



-



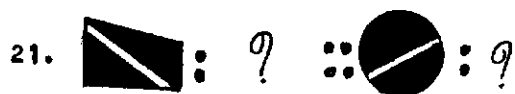
-



-



-



-



-



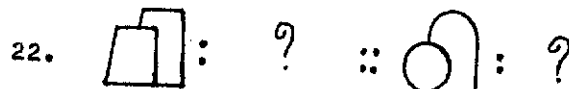
-



-



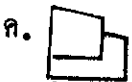
-



-



-



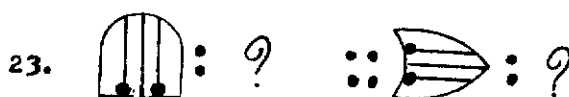
-



-



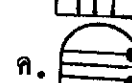
-



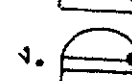
-



-



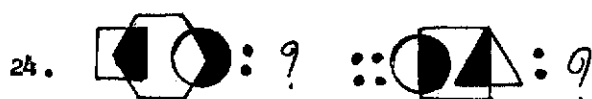
-



-



-



-



-



-



-



-



ตอนที่ 2 จงเลือกคำตอบที่เป็นคำตรงข้ามกับคำที่กำหนดให้ 2 คำ

25. ขับซ้อน : ? :: คลุมเครือ : ?
 ก. ราบเรียบ - เห็นชัด
 ข. เรียบร้อย - เต็มชัด
 ค. เรียบง่าย - กระจ่าง
 ง. ง่ายด้าย - ชัดเจน
 จ. เรียบง่าย - ชัดแจ้ง
26. ทะเล้น : ? :: ล้อเลียน : ?
 ก. สุภาพ - นอบน้อม
 ข. สงบเสงี่ยม - เกรงใจ
 ค. เครื่องคิด - เยาะเย้ย
 ง. อ่อนน้อม - เกรงขาม
 จ. เครื่องขรีม - เคารพ
27. จักัด : ? :: มีดมีด : ?
 ก. เหลือใช้ - สาดส่อง
 ข. ขัดสน - ร่องแจ้ง
 ค. ประปราย - สลัวสลัว
 ง. มากมาย - สว่างไสว
 จ. ถิ่นหลาม - แหว้วว
28. ติดขัด : ? :: สะเพร่า : ?
 ก. ราบรื่น - สุขุม
 ข. ปลอดโปร่ง - ปราณีต
 ค. สะดวก - รอบคอบ
 ง. คล่องแคล่ว - ละเอียด
 จ. ผ่านพ้น - พึงพิถัน
29. ไร่ร้าง : ? :: ช่างเจรจา : ?
 ก. เจียบขรึม - พูดน้อย
 ข. สุขุม - เฉยเมย
 ค. หงอยเหงาม - นิ่งเฉย
 ง. เศร้าสร้อย - นิ่งเจียบ
 จ. อมทุกข์ - ขวานผ่าซาก
30. สามัคคี : ? :: ร่วมใจ : ?
 ก. แดกแยก - หมางเมิน
 ข. ไร่ฉาน - ผินใจ
 ค. รุกราน - ช้องใจ
 ง. ประภักดิ์ - ผินใจ
 จ. อาฆาต - แคลงใจ
31. มลทิน : ? :: อิดโรย : ?
 ก. สะอาด - แข็งแรง
 ข. บริสุทธิ์ - สดชื่น
 ค. ขาวผ่อง - เป่งปลั่ง
 ง. สะอาด - สดใส
 จ. บริสุทธิ์ - เข้มแข็ง
32. ปลอดโยน : ? :: ำห้อย : ?
 ก. ำาเติม - ำาห้อย
 ข. ำาเติม - ำาห้อย
 ค. ำาเติม - ำาห้อย
 ง. ำาเติม - ำาห้อย
 จ. ำาเติม - ำาห้อย

33. ขุ่มขื่น : ? :: ล้นเหลือ : ?
 ก. อับเฉา - เล็กน้อย
 ข. แท้งแล้ง - ขาดแคลน
 ค. เหี่ยวเฉา - กันดาร
 ง. เหือดแห้ง - นิดหน่อย
 จ. เหี่ยวแห้ง - ยากไร้
34. รุ่มร้อน : ? :: ประหม่า : ?
 ก. สุขุม - เชื่อถือ
 ข. ใจเย็น - เชื่อใจ
 ค. เฉยเมย - ไร้ใจ
 ง. สงบเสงี่ยม - มั่นใจ
 จ. เยือกเย็น - เชื่อมั่น
35. สรทธา : ? :: รักรักร : ?
 ก. อุแคลน - รังเกียจ
 ข. อุหมั่น - โกรธแค้น
 ค. เย้ยหยัน - เกสียดขัด
 ง. เหยียดหยาม - ชิงชัง
 จ. สบประหม่า - พยาบาล
36. สำเร็จ : ? :: สมหวัง : ?
 ก. ล้มละลาย - ล้นหวัง
 ข. ล้มเหลว - ผิดหวัง
 ค. พลาดพลั้ง - หมดหวัง
 ง. ล้มเหลว - พลาดหวัง
 จ. พลาดท่า - ไม่สมหวัง

37. เปลี่ยนแปลง : ? :: ตื่นเต้น : ?
 ก. จ้าใจ - เฉยเมย
 ข. ช้าชาก - ตกใจ
 ค. เหมือนเดิม - ขาชิน
 ง. คงเดิม - ชิมเศร้า
 จ. อนุรักษ์ - เปื้อนนาย
38. ใจลอย : ? :: นำเปื้อ : ?
 ก. จดจ่อ - สนุกสนาน
 ข. สนใจ - ตื่นเต้น
 ค. ยั้งใจ - ไร้ใจ
 ง. เฟื่องพิศ - วิต
 จ. จ้องมอง - แสนสุข
39. เข้มข้น : ? :: รุนแรง : ?
 ก. เจือจาง - เบาบาง
 ข. จางใส - อ่อนลง
 ค. เหลวใส - ทุเลา
 ง. เจือจาง - น้อยลง
 จ. จางใส - ทุเลา
40. ปราบปราม : ? :: ผูกมัด : ?
 ก. ร่วมแรง - หึงขวาง
 ข. สนับสนุน - ให้อิสระ
 ค. ติดตาม - ปลอypละ
 ง. ช่วยเหลือ - ให้ออกาส
 จ. หนุนหลัง - ละทิ้ง

แบบทดสอบวัดการคิดเอกลัคนัยแบบสัมพันธ์ ฉบับที่ 2 (ฟอร์ม B)

คำชี้แจง แบบทดสอบฉบับนี้มี ทั้งหมด 50 ข้อ แบ่งเนื้อหาออกเป็น 2 ตอน คือ

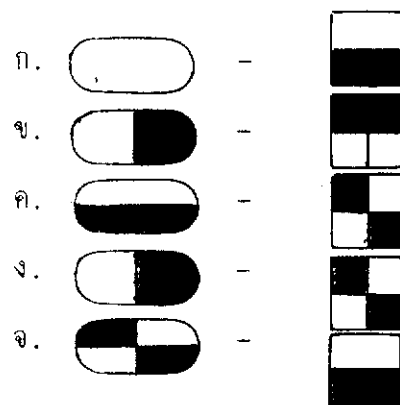
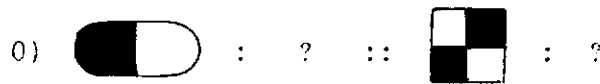
ตอนที่ 1 มีเนื้อหาเป็นภาพ

ตอนที่ 2 มีเนื้อหาเป็นภาษา

ให้นักศึกษาหาแบบทดสอบทั้งหมด โดยเลือกคำตอบที่ถูกต้องที่สุดเพียงคำตอบเดียวลงใน

กระดาษคำตอบที่จัดเตรียมไว้ ดังตัวอย่าง

ตอนที่ 1 จงเลือกคำตอบที่เป็นภาพตรงข้ามกับภาพที่กำหนดให้ 2 ภาพ



คำตอบที่ถูกต้อง คือ ข้อ ง.

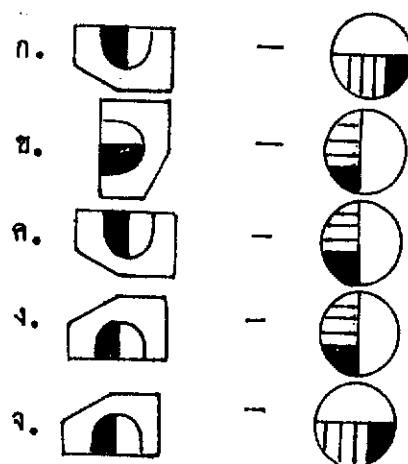
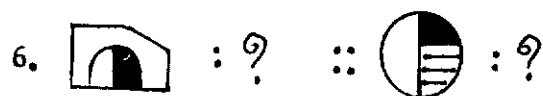
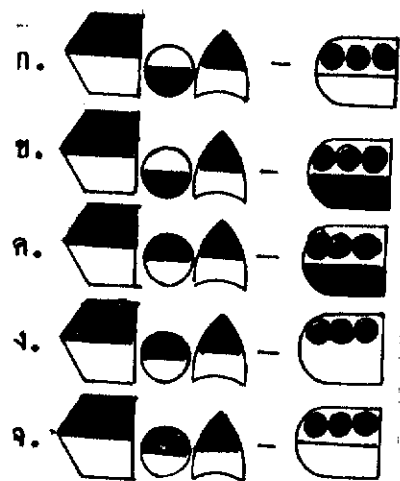
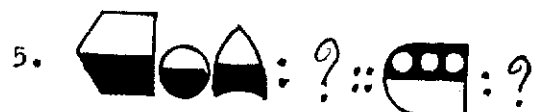
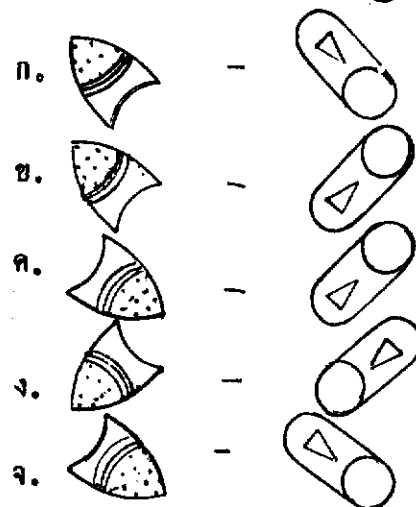
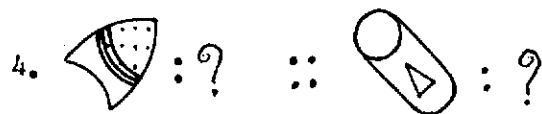
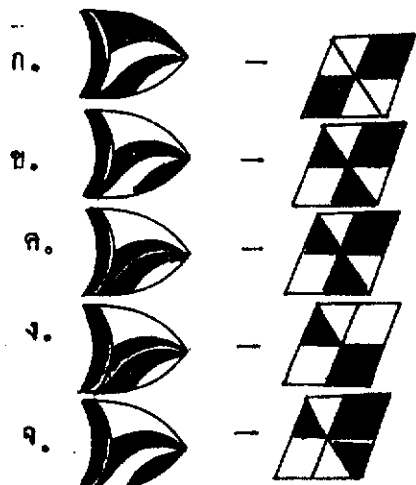
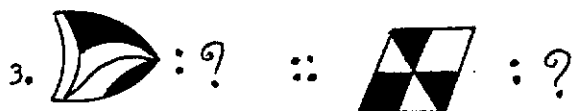
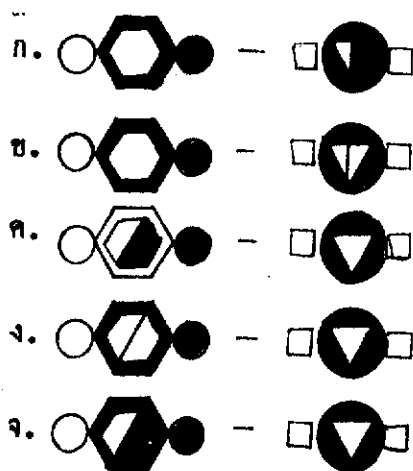
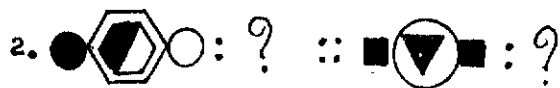
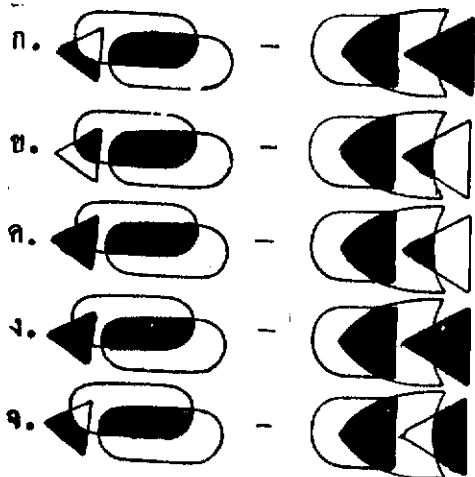
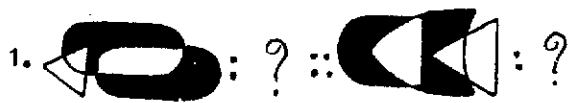
ตอนที่ 2 จงเลือกคำตอบที่เป็นคำตรงข้ามกับคำที่กำหนดให้ 2 คำ

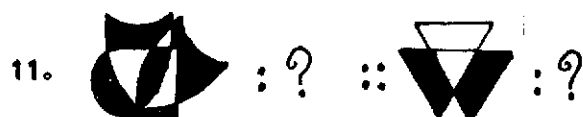
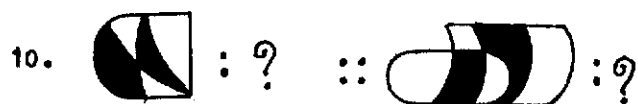
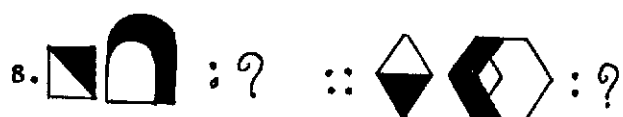
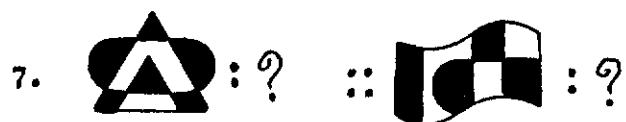
00) สั้น : ? :: กลม : ?

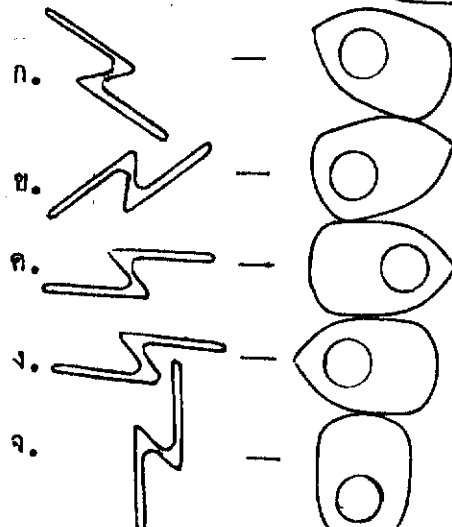
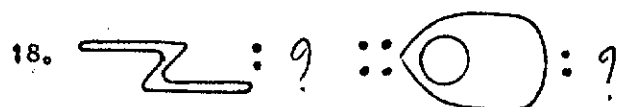
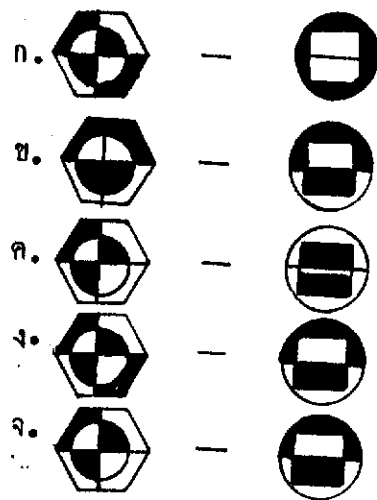
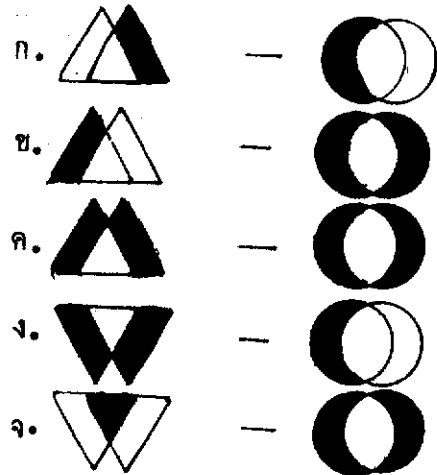
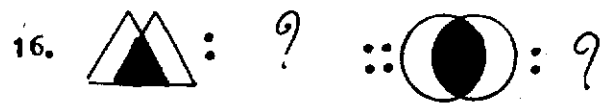
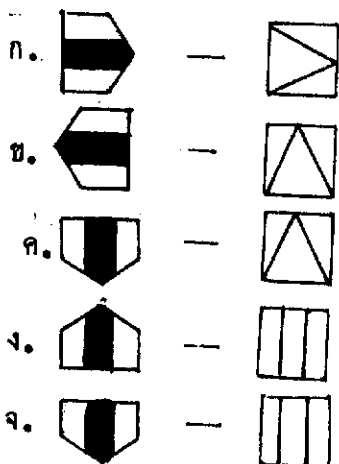
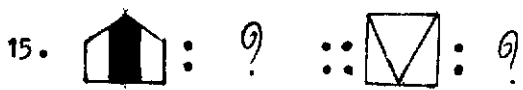
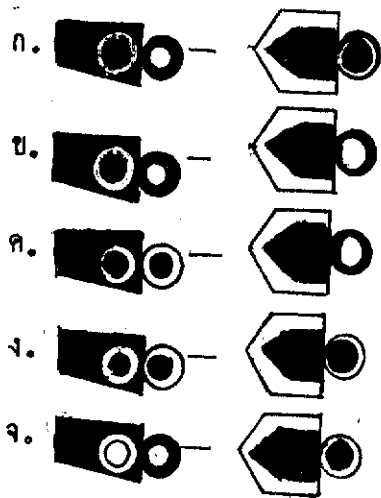
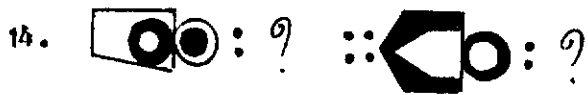
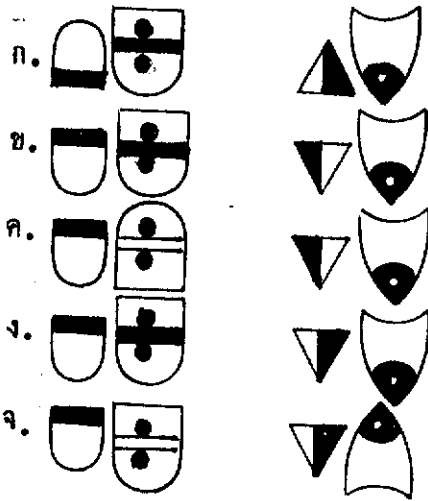
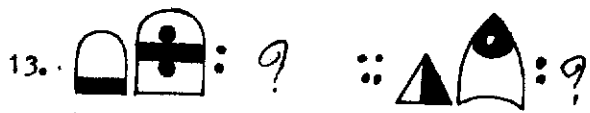
- | | | | |
|----|------|---|---------|
| ก. | สูง | - | หักมุม |
| ข. | ยาว | - | เหลี่ยม |
| ค. | แคบ | - | กว้าง |
| ง. | ดำ | - | ขาว |
| จ. | เล็ก | - | รี |

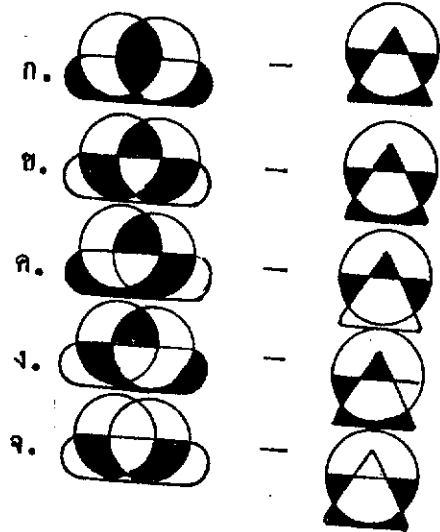
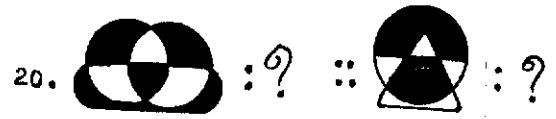
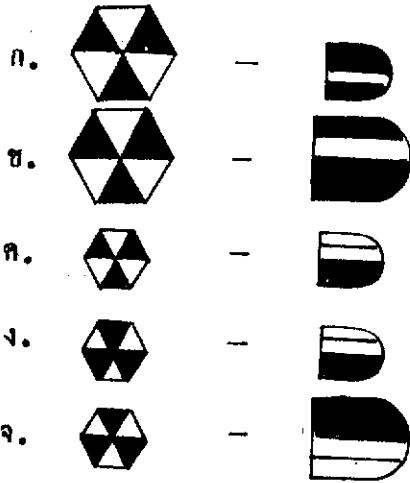
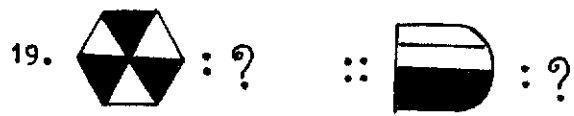
คำตอบที่ถูกต้อง คือ ข้อ ข.

ตอนที่ 1 จงเลือกคำตอบที่เป็นภาพตรงข้ามกับภาพที่กำหนดให้ 2 ภาพ









ตอนที่ 2 จงเลือกคำตอบที่เป็นคำตรงข้ามกับคำที่กำหนดให้ 2 คำ

- | | |
|--|---|
| <p>21. เพ็ญหน้า : ? :: แสดงตัว : ?</p> <p>ก. ถอยหลัง - หนีหาย</p> <p>ข. หลบหนี - ชุกซ่อน</p> <p>ค. หันหลัง - ปิดบัง</p> <p>ง. หลบหลีก - หลบซ่อน</p> <p>จ. หนีหน้า - เก็บตัว</p> | <p>25. วุ่นวาย : ? :: ห้วนเหว : ?</p> <p>ก. สงบสุข - มั่นใจ</p> <p>ข. ร่มรื่น - สบายใจ</p> <p>ค. ร่มเย็น - เชื่อใจ</p> <p>ง. ผาสุก - เบาใจ</p> <p>จ. แสนสุข - สุขใจ</p> |
| <p>22. คู่ครอง : ? :: ยินดี : ?</p> <p>ก. ำหิร้าย - รียยา</p> <p>ข. ทับถม - อวามาต</p> <p>ค. ยุแหย - พยาบาท</p> <p>ง. ระราน - อิจฉา</p> <p>จ. โจมตี - เกลียดชัง</p> | <p>26. แหลม : ? :: คล่องแคล่ว : ?</p> <p>ก. ทื่อ - งุ่มง่าม</p> <p>ข. ชู่ - ชุ่มช้ำ</p> <p>ค. มน - ชักช้า</p> <p>ง. ผิด - ลูกลี้ลูกลอน</p> <p>จ. หล - เรือยเลื่อย</p> |
| <p>23. จัดแย้ง : ? :: สั่นคลอน : ?</p> <p>ก. สมยอม - แน่วแน่</p> <p>ข. สอดคล้อง - มั่นคง</p> <p>ค. ชมชอบ - มั่งคั่ง</p> <p>ง. ชมเชย - มุ่งมั่น</p> <p>จ. เห็นใจ - ยินหยัด</p> | <p>27. ทุกข์ยาก : ? :: อุปสรรค : ?</p> <p>ก. ไร้กังวล - สะดวก</p> <p>ข. หมดห่วง - ปลอดภัย</p> <p>ค. แสนสุข - ราบเรียบ</p> <p>ง. สุขสบาย - ราบรื่น</p> <p>จ. สบายใจ - ง่ายดาย</p> |
| <p>24. ยะโส : ? :: เสสสร้าง : ?</p> <p>ก. สุภาพ - ซื่อสัตย์</p> <p>ข. อ่อนโยน - จริงจัง</p> <p>ค. นอบน้อม - จริงใจ</p> <p>ง. เรียบร้อย - บริสุทธิ์ใจ</p> <p>จ. อ่อนน้อม - กันเอง</p> | <p>28. อึกทิก : ? :: ครึกครื้น : ?</p> <p>ก. สงบ - ขบเซา</p> <p>ข. เจียบสัจ - วังเวง</p> <p>ค. เจียบเสียง - โหวกเหวก</p> <p>ง. เจียบเหงา - เบล่าเบลี่ยว</p> <p>จ. เจียบสงบ - หงอยเหงา</p> |

29. กษัตริย์ : ? :: มังมี : ?

- ก. ราชนี - ต้าด้อย
ข. ยาก - ตกต่ำ
ค. บ่าวไพร่ - แร้นแค้น
ง. พสกนิกร - ยากไร้
จ. บริวาร - ขาดแคลน

30. รุ่งเรือง : ? :: พัฒนา : ?

- ก. ตกต่ำ - เสื่อมถอย
ข. ขบเซา - ลำช้า
ค. ตกอับ - ลำสมัย
ง. ส่อแผ่ - ถอยหลัง
จ. ย่ำแย่ - ลำหลัง

31. มักง่าย : ? :: เลอะเทอะ : ?

- ก. ละเอียด - สวยงาม
ข. รอบคอบ - มีระเบียบ
ค. ขยันขันแข็ง - เรียบร้อย
ง. พิถีพิถัน - สะอาด
จ. สงบเสงี่ยม - เกลี้ยงเกล้า

32. มหิมา : ? :: ปราสาท : ?

- ก. กะทัดรัด - กระต๊อบ
ข. ขนาดย่อม - กระท่อม
ค. เล็กจิ๋ว - สลัม
ง. นิดหน่อย - รังหนู
จ. ขนาดจิ๋ว - กระท่อม

33. กระตือรือร้น : ? :: เข้มแข็ง : ?

- ก. รุ่มง่าม - อ่อนช้อย
ข. เชื่องซึม - อ่อนโยน
ค. เกียจคร้าน - อ่อนไหว
ง. เนือยชา - อ่อนแอ
จ. จี้เกียจ - อ่อนน้อม

34. ผู้นำ : ? :: อาสา : ?

- ก. ผู้ตาม - ถูกบังคับ
ข. ลูกน้อง - เชิดชวน
ค. ลูกทีม - ถูกเลือก
ง. ชาวบ้าน - จำยอม
จ. ประชาชน - สมยอม

35. ยิ้มแย้ม : ? :: เบิกบาน : ?

- ก. บุคปึ้ง - หม่นหมอง
ข. บึ้งตึง - ห่อเหี่ยว
ค. เฉยเมย - เม่อลอย
ง. หน้างอ - ทुरนทुरาย
จ. เศร้าสร้อย - กระสับกระส่าย

36. ขมขื่น : ? :: ซึ้งใจ : ?

- ก. เบาใจ - ฐู่ใจ
ข. สบายใจ - เข้าใจ
ค. เบินใจ - ชอบใจ
ง. สุขใจ - สมใจ
จ. ชื่นใจ - ติดใจ

37. ตอก : ? :: พลักไส : ?

- ก. ย้ำ - ตันเข้า
ข. สิ่ง - ปัดป้อง
ค. ไร่ - สิ่งร้าง
ง. แกะ - จุดกระชาก
จ. แงะ - ยื้อแย่ง

38. กล่าวหาญ : ? :: แข็งแกร่ง : ?

- ก. เหลาะแหละ - เมื่อปล้ำ
ข. หวาดกลัว - อ่อนแรง
ค. อ่อนแอ - บวกเปียก
ง. ขลาดกลัว - อ่อนล้า
จ. อ่อนไหว - บวดเมื่อย

39. งดงาม : ? :: ละเอียดลออ : ?

- ก. ขยับเขย่ง - หยาบคาย
ข. ไรศรกร - เหลาะแหละ
ค. นาซัง - สุ่มส่าม
ง. นาเกลียด - หยาบ
จ. นารังเกียจ - ขากระเปียบ

40. รมเย็น : ? :: ปลอดภัย : ?

- ก. มีทุกข์ - เกิดเหตุ
ข. ไร่ฉาน - เสียภัย
ค. เตือดร้อน - อันตราย
ง. แตกแยก - ตั้งเครียด
จ. กระวนกระวาย - มีภัย

41. เชิดชู : ? :: กล่าวถึง : ?

- ก. สาบแช่ง - ล้มหาย
ข. เพียบอย่า - ถูกลิ้ม
ค. คุแกลน - เวียบหาย
ง. คุหมื่น - ถูกกลืน
จ. ลมหลู่ - ทอดทิ้งจ

42. สรรเสริญ : ? :: อ่อนโยน : ?

- ก. ตาหนี - แข็งกร้าว
ข. นินทา - เข้มแข็ง
ค. กล่าวหา - อ่อนไหว
ง. สาบแช่ง - แข็งกระด้าง
จ. สาบส่ง - แข็งแกร่ง

43. สรุบทความ : ? :: กระชับ : ?

- ก. ท้าวความ - ยืด
ข. เติมความ - คลาย
ค. แต่งความ - บาน
ง. ขยายความ - หลวม
จ. เรียงความ - คลาย

44. ดำเนิน : ? :: เคลื่อนไหว : ?

- ก. ยกเว้น - นิ่งเฉย
ข. วางเว้น - เจียบเจียบ
ค. ละเว้น - แน่นิ่ง
ง. ยกเลิก - หยุดนิ่ง
จ. ถอนตัว - สงบนิ่ง

45. หยั่งเห็น : ? :: ชัดเจน : ?

- ก. มีคมน - เลื่อนลาง
ข. มีดมีด - ลบเลือน
ค. ปิดบัง - คลุมเครือ
ง. บอดสนิท - ยุ่งเหยิง
จ. บังตา - ลายตา

46. สูงส่ง : ? :: มีสัจจะ : ?

- ก. ตกต่ำ - หลอกลวง
ข. ต่ำต้อย - คดโกง
ค. ต่ำต้อย - สับลับ
ง. ต่ำช้า - เหลวไหล
จ. ตกต่ำ - โกงหก

47. อุปการะ : ? :: เกื้อกูล : ?

- ก. ละเลย - ฟังพา
ข. หอดทิ้ง - เบียดเบียน
ค. พลักใส - แก่งแย่ง
ง. ละทิ้ง - เอาเปรียบ
จ. หักขว้าง - แย่งชิง

48. ปลานปล้ม : ? :: ชื่นชม : ?

- ก. ห่อเพียว - กล่าวโทษ
ข. สั่งเวช - ำหิร้าย
ค. เวทนา - สาบแช่ง
ง. หดหู่ - ตำหนิ
จ. เศร้าใจ - ว่ากล่าว

49. ห่วงใย : ? :: เข้มงวด : ?

- ก. ว่างใจ - ตามใจ
ข. ร้องใจ - ละเลย
ค. เชื่อใจ - หักขว้าง
ง. วางใจ - ละเลย
จ. ไร้ง่วง - หักขว้าง

50. โพล้เพล้ : ? :: อับแสง : ?

- ก. สนธยา - แดดจ้า
ข. รุ่งอรุณ - เบิกฟ้า
ค. เข้าตรู่ - กลางแจ้ง
ง. ฟ้าสว่าง - แดดส่อง
จ. พลบค่ำ - ฟ้าสว่าง

แบบทดสอบวัดการคิดเอกลัคนัยแบบสัมพันธ์ ฉบับที่ 3 (ฟอร์ม C)

คำชี้แจง แบบทดสอบฉบับนี้ มีทั้งหมด 16 ข้อ แบ่งเนื้อหาออกเป็น 2 ตอน คือ

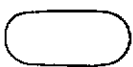
ตอนที่ 1 มีเนื้อหาเป็นภาพ

ตอนที่ 2 มีเนื้อหาเป็นภาษา

ให้นักศึกษาทำแบบทดสอบทั้งหมด โดยเลือกคำตอบที่ถูกต้องที่สุดเพียงคำตอบเดียวลงในกระดาษคำตอบที่จัดเตรียมไว้ ดังตัวอย่าง

ตอนที่ 1 จงเลือกคำตอบที่เป็นภาพตรงข้ามกับภาพที่กำหนดให้ 2 ภาพ

0)  : ? ::  : ?

- | | | | |
|----|---|---|---|
| ก. |  | - |  |
| ข. |  | - |  |
| ค. |  | - |  |
| ง. |  | - |  |
| จ. |  | - |  |

คำตอบที่ถูกต้อง คือ ข้อ ง.

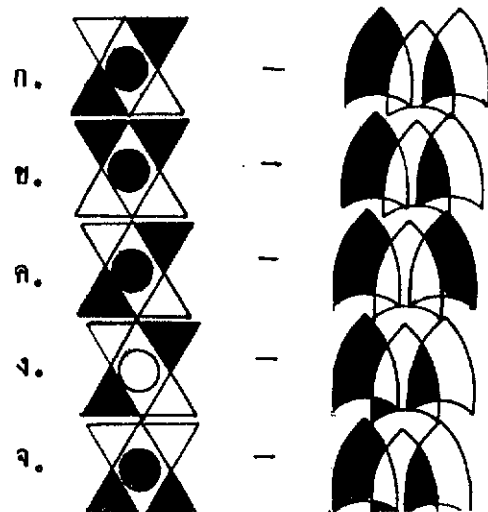
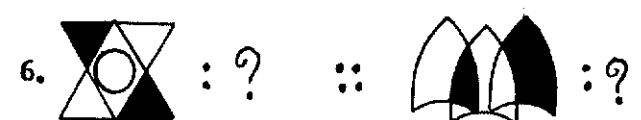
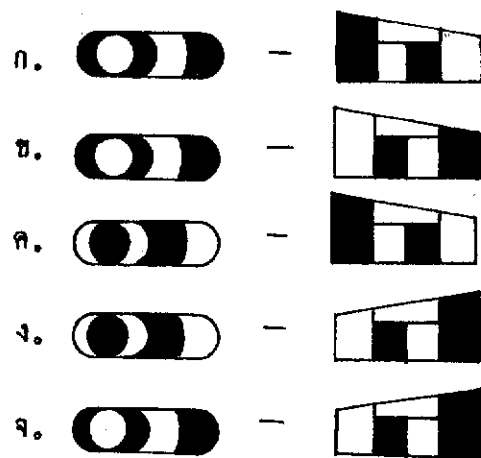
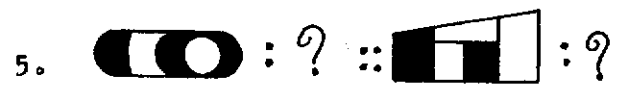
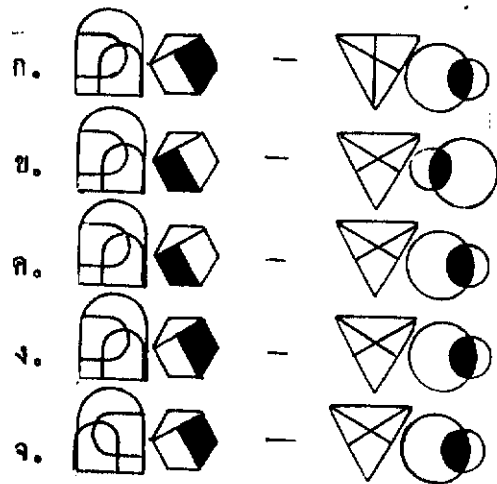
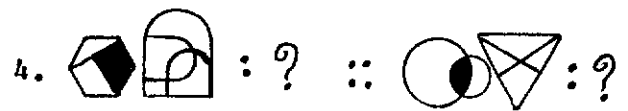
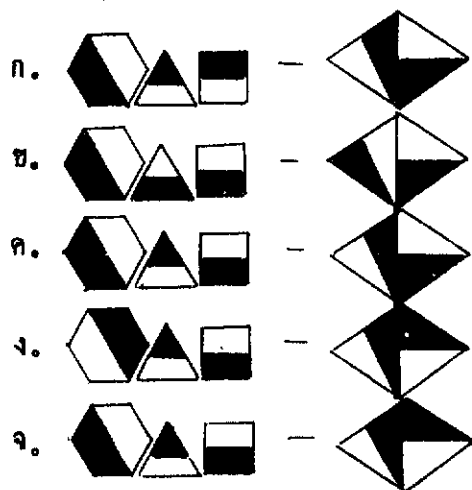
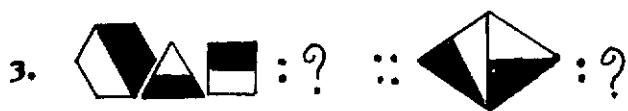
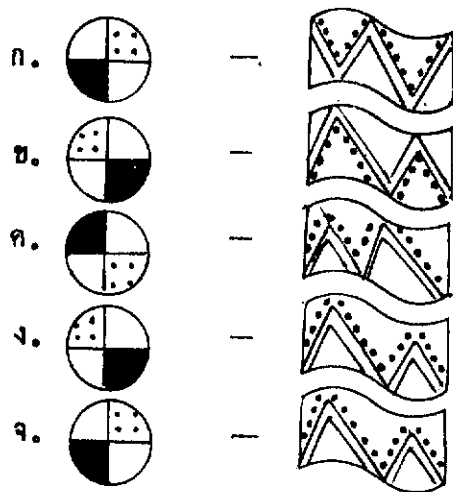
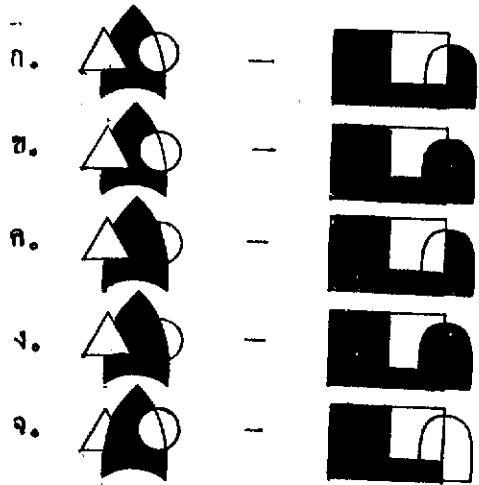
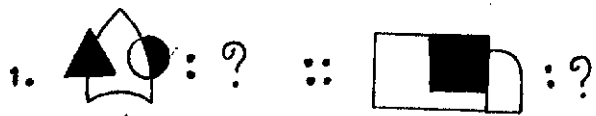
ตอนที่ 2 จงเลือกคำตอบที่เป็นคำตรงข้ามกับคำที่กำหนดให้ 2 คำ

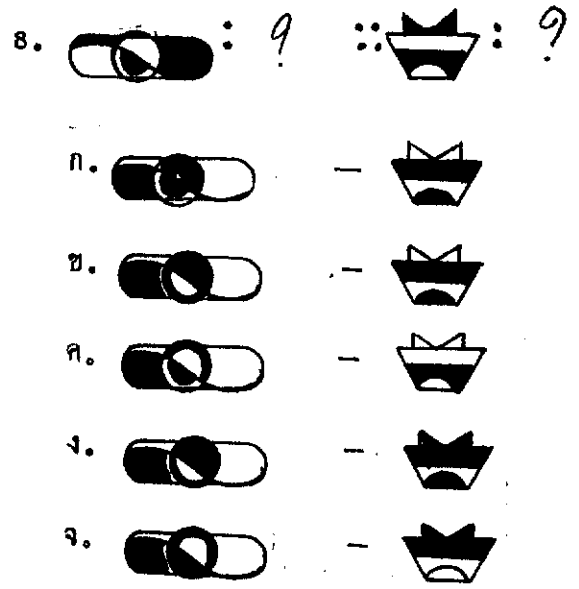
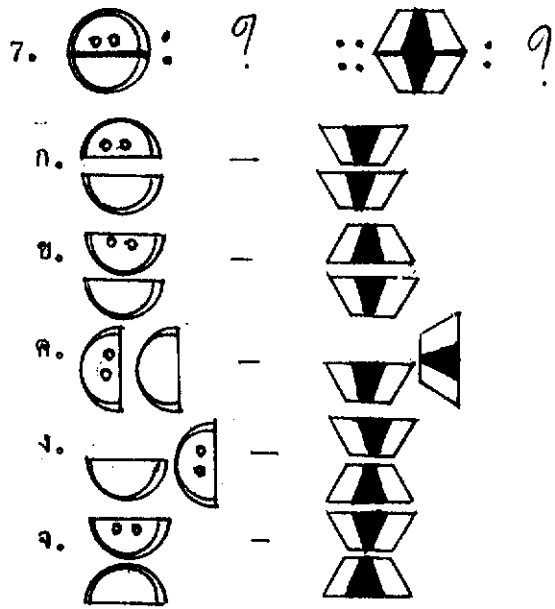
00) - สั้น : ? :: กลม : ?

- | | | | |
|----|------|---|---------|
| ก. | สูง | - | หักมุม |
| ข. | ยาว | - | เหลี่ยม |
| ค. | แคบ | - | กว้าง |
| ง. | ต่ำ | - | เรียว |
| จ. | เล็ก | - | รี |

คำตอบที่ถูกต้อง คือ ข้อ ข.

ตอนที่ 1 จงเลือกคำตอบที่เป็นภาพตรงข้ามกับภาพที่กำหนดให้ 2 ภาพ





ตอนที่ 2 จงเลือกคำตอบที่เป็นคำตรงข้ามกับคำที่กำหนดให้ 2 คำ

- | | |
|---|--|
| <p>9. เกียจคร้าน : ? :: กล่าวโทษ : ?</p> <p>ก. พยายาม - ให้รางวัล</p> <p>ข. บากบั่น - ชื่นชม</p> <p>ค. ขวนขวาย - ชมเชย</p> <p>ง. อุดหนุน - ชื่นชอบ</p> <p>จ. กระตือรือร้น - ยินดี</p> | <p>13. ปิดหุ่ย : ? :: กลอกกลิ้ง : ?</p> <p>ก. เครื่องคิด - เทรตรง</p> <p>ข. เครื่องขรีม - จริงจัง</p> <p>ค. จู้จี้ - มุ่งมั่น</p> <p>ง. บังคับ - แน่วแน่</p> <p>จ. เจ้าระเบียบ - เชื้อมั่ว</p> |
| <p>10. หนักแน่น : ? :: ส้ารวม : ?</p> <p>ก. ใจลอย - ชุ่มช้ำ</p> <p>ข. ทุเบา - ตื่นตระหนก</p> <p>ค. โกลเล - หลุกหลิก</p> <p>ง. ใจอ่อน - แข็งกระด้าง</p> <p>จ. อ่อนไหว - กระสับกระส่าย</p> | <p>14. อ้ามพิต : ? :: ก่อกรรม : ?</p> <p>ก. ใจดี - ทาดี</p> <p>ข. เมตตา - สร้างบุญ</p> <p>ค. สงสาร - ใสบาตร</p> <p>ง. กรุณา - สะสมบาป</p> <p>จ. เวทนา - บริจาคทาน</p> |
| <p>11. สุภาพ : ? :: สุขุม : ?</p> <p>ก. ก้าวร้าว - กระฉับกระเฉง</p> <p>ข. ส่อเสียด - ทะมัดทะแมง</p> <p>ค. แข็งกระด้าง - กระสับกระส่าย</p> <p>ง. เยี้ยหยัน - กระตือรือร้น</p> <p>จ. พยาบคาย - ลูกลี้ลูกกลน</p> | <p>15. อบอุ่น : ? :: สดชื่น : ?</p> <p>ก. ฟ้าเหว่ - หมองคล้ำ</p> <p>ข. หงอยเหงาม - ห่อเหี่ยว</p> <p>ค. อ่างว้าง - หม่นหมอง</p> <p>ง. เจียบเหงาม - ซีดเซียว</p> <p>จ. ชีมเศร้า - เศร้าสร้อย</p> |
| <p>12. อารุโธ : ? :: ร่วงโรย : ?</p> <p>ก. อ่อนหัด - แรกเริ่ม</p> <p>ข. ผู้เยาว์ - แรกแย้ม</p> <p>ค. รุ่งเรือง - เบ่งบาน</p> <p>ง. วัยรุ่น - สดใส</p> <p>จ. รุ่งหลัง - เปล่งปลั่ง</p> | <p>16. กตัญญู : ? :: ยกย่อง : ?</p> <p>ก. เนรคุณ - ดูหมิ่น</p> <p>ข. ทรมาน - ดูถูก</p> <p>ค. ทรมาน - ดูแคลน</p> <p>ง. เนรคุณ - สาปแช่ง</p> <p>จ. ทรมาน - ลบหลู่</p> |

ภาคผนวก ข

การเทียบคะแนนแปลง

ตาราง 26 การเทียบคะแนนของคะแนนแปลงที่ได้มาจากการคำนวณด้วยสูตร สวีซี
กับคะแนน X

ความยาวของ แบบทดสอบหลัก	คะแนนดิบ (X)	คะแนนแปลง (Y/)			
		แมคเคน1	แมคเคน2	สูตรพ1	สูตรพ2
แบบทดสอบเชื่อมโยง	19.00	27.00	9.88	30.64	28.51
8 ข้อ (.2)	20.00	27.86	11.79	30.95	28.90
	21.00	28.71	13.69	31.27	29.28
	22.00	29.57	15.60	31.58	29.66
	23.00	30.42	17.51	31.90	30.05
	24.00	31.28	19.42	32.21	30.43
	25.00	32.14	21.33	32.53	30.81
	26.00	32.99	23.23	32.84	31.20
	27.00	33.85	25.14	33.16	31.58
	28.00	34.71	27.05	33.47	31.97
	29.00	35.56	28.96	33.79	32.35
	30.00	36.42	30.87	34.10	32.73
	31.00	37.27	32.77	34.42	33.12
	32.00	38.13	34.68	34.73	33.50
	34.00	39.84	38.50	35.36	34.27

ความยาวของ แบบทดสอบหลัก	คะแนนดิบ (X)	คะแนนแปลง (Y/)			
		แมคแกน1	แมคแกน2	ผู้ตรวจพ1	ผู้ตรวจพ2
แบบทดสอบเชื่อมโยง 12 ข้อ (.3)	12.00	12.44	21.04	32.23	32.11
	15.00	15.94	24.37	32.90	32.80
	18.00	19.44	27.70	33.58	33.49
	19.00	20.60	28.81	33.80	33.72
	21.00	22.94	31.03	34.25	34.18
	22.00	24.10	32.15	34.47	34.41
	24.00	26.44	34.37	34.92	34.87
	25.00	27.60	35.48	35.14	35.10
	26.00	28.77	36.59	35.37	35.33
	27.00	29.93	37.70	35.59	35.56
	28.00	31.10	38.81	35.82	35.79
	29.00	32.27	39.92	36.04	36.02
	30.00	33.43	41.03	36.26	36.25
	31.00	34.60	42.14	36.49	36.48
	32.00	35.76	43.25	36.71	36.71
	33.00	36.93	44.36	36.94	36.94
	34.00	38.10	45.47	37.16	37.17
	35.00	39.26	46.58	37.39	37.40
	45.00	50.92	57.68	39.63	39.70

ความยาวของ แบบทดสอบหลัก	คะแนนดิบ (X)	คะแนนแปลง (Y/)			
		แมกแคน1	แมกแคน2	รูตตรัฟ1	รูตตรัฟ2
แบบทดสอบเชื่อมโยง 16 ข้อ(.4)	11.00	28.00	16.91	30.74	34.34
	13.00	28.88	20.48	31.06	34.69
	17.00	30.66	27.63	31.71	35.38
	18.00	31.10	29.41	31.88	35.55
	19.00	31.54	31.20	32.04	35.72
	20.00	31.99	32.99	32.20	35.89
	21.00	32.43	34.77	32.36	36.07
	22.00	32.87	36.56	32.53	36.24
	23.00	33.31	38.34	32.69	36.41
	24.00	33.76	40.13	32.85	36.58
	25.00	34.20	41.92	33.02	36.75
	26.00	34.64	43.70	33.18	36.93
	27.00	35.09	45.49	33.34	37.10
	28.00	35.53	47.27	33.50	37.27
	29.00	35.97	49.06	33.67	37.44
	30.00	36.42	50.85	33.83	37.62
	31.00	36.86	52.63	33.99	37.79
	32.00	37.30	54.42	34.16	37.96
	34.00	38.19	57.99	34.48	38.30
	36.00	39.07	61.56	34.81	38.65

ความยาวของ แบบทดสอบหลัก	คะแนนดิบ (X)	คะแนนแปลง (Y/)			
		แมคแคน1	แมคแคน2	ู๊ดคริฟ1	ู๊ดคริฟ2
ข้อสอบรวม	10.00	21.66	-5.10	31.32	25.79
8 ข้อ (.2)	11.00	22.35	-3.10	31.49	25.99
	13.00	23.73	.89	31.84	26.40
	14.00	24.42	2.89	32.01	26.60
	15.00	25.11	4.89	32.19	26.81
	16.00	25.80	6.88	32.36	27.01
	17.00	26.49	8.88	32.54	27.21
	20.00	28.56	14.87	33.06	27.82
	21.00	29.25	16.87	33.23	28.03
	22.00	29.94	18.87	33.40	28.23
	23.00	30.63	20.86	33.58	28.43
	24.00	31.32	22.86	33.75	28.64
	25.00	32.01	24.86	33.92	28.84
	26.00	32.70	26.85	34.10	29.04
	27.00	33.38	28.85	34.27	29.25
	28.00	34.07	30.85	34.44	29.45
	29.00	34.76	32.85	34.62	29.65
	30.00	35.45	34.84	34.79	29.86
	31.00	36.14	36.84	34.96	30.06
	32.00	36.83	38.84	35.14	30.26
	35.00	38.90	44.83	35.66	30.87

ความยาวของ แบบทดสอบหลัก	คะแนนดิบ (X)	คะแนนแปลง (Y/)			
		แมคเคน1	แมคเคน2	วีตอร์ฟ1	วีตอร์ฟ2
ข้อสอบรวม	9.00	15.80	19.37	29.79	29.73
12 ข้อ (.3)	13.00	20.13	22.50	30.49	30.43
	15.00	22.30	24.06	30.84	30.78
	17.00	24.46	25.63	31.19	31.13
	18.00	25.55	26.41	31.37	31.30
	20.00	27.71	27.97	31.72	31.65
	21.00	28.80	28.76	31.89	31.83
	22.00	29.88	29.54	32.07	32.00
	23.00	30.96	30.32	32.24	32.18
	24.00	32.05	31.10	32.42	32.35
	25.00	33.13	31.88	32.59	32.53
	26.00	34.21	32.67	32.77	32.70
	27.00	35.30	33.45	32.94	32.88
	28.00	36.38	34.23	33.12	33.05
	29.00	37.46	35.01	33.30	33.23
	30.00	38.55	35.80	33.47	33.40
	31.00	39.63	36.58	33.65	33.58
	32.00	40.71	37.36	33.82	33.75
	35.00	43.96	39.71	34.35	34.27

ความยาวของ แบบทดสอบหลัก	คะแนนดิบ (X)	คะแนนแปลง (Y/)			
		แมคแคน1	แมคแคน2	รูคดร์พ1	รูคดร์พ2
ข้อสอบรวม	18.00	14.95	31.83	34.00	28.36
16 ข้อ(.4)	19.00	16.74	32.49	34.33	28.71
	20.00	18.53	33.14	34.66	29.05
	21.00	20.31	33.79	34.99	29.40
	22.00	22.10	34.45	35.32	29.74
	23.00	23.89	35.10	35.65	30.09
	24.00	25.67	35.75	35.98	30.44
	25.00	27.46	36.41	36.31	30.78
	26.00	29.25	37.06	36.64	31.13
	27.00	31.03	37.71	36.97	31.48
	28.00	32.82	38.37	37.30	31.82
	29.00	34.61	39.02	37.63	32.17
	30.00	36.39	39.67	37.96	32.51
	31.00	38.18	40.33	38.29	32.86
	32.00	39.96	40.98	38.63	33.21
	33.00	41.75	41.63	38.96	33.55
	34.00	43.54	42.29	39.29	33.90

ประวัติย่อของผู้วิจัย

เกิดวันที่ 25 เดือนกรกฎาคม พ.ศ. 2502 จังหวัด กรุงเทพมหานคร

การศึกษา

พ.ศ. 2517 จบการศึกษาชั้นประถมศึกษา จาก ร.ร.สายน้ำทิพย์

พ.ศ. 2522 จบการศึกษาชั้นมัธยมศึกษาตอนปลาย จาก ร.ร.สายน้ำผึ้ง

พ.ศ. 2526 จบประกาศนียบัตรพยาบาลศาสตร์และผดุงครรภ์ จาก วิทยาลัยพยาบาลกรุงเทพ

การทำงาน

พ.ศ. 2526 วิทยาลัยพยาบาลพะเยา จังหวัด พะเยา ตำแหน่ง วิทยากรที่ 3

พ.ศ. 2532 วิทยาลัยพยาบาลกรุงเทพ จังหวัดกรุงเทพฯ ตำแหน่ง วิทยากรที่ 5

ปัจจุบัน วิทยาลัยพยาบาลบรมราชชนนีกทม ตำแหน่ง วิทยากรที่ 6

การศึกษาวิธีการเทียบคะแนนแบบเส้นตรงตามแบบจำลองคะแนนจริงสัมพัทธ์

บทคัดย่อ

ของ

สุนนา โสคติพลอนันต์

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา

ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

พฤษภาคม 2539

การศึกษาค้นคว้าครั้งนี้ เป็นการเปรียบเทียบวิธีการเทียบคะแนนแบบเส้นตรงที่ได้จากคะแนนสอบมาคำนวณด้วยสูตร สี่วิธี โดยมีข้อตกลงว่า แบบทดสอบที่นำมาเทียบคะแนนนั้น ต้องเป็นไปตามแบบจำลองคะแนนจริงสัมพันธ์ และมีความเชื่อมั่นเท่าเทียมกัน กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักศึกษาพยาบาลชั้นปีที่ 1 และชั้นปีที่ 2 หลักสูตรพยาบาลศาสตร์ และพยาบาลศาสตรระดับต้น ปีการศึกษา 2538 ของวิทยาลัยพยาบาล สังกัดสถาบันพัฒนากำลังคนด้านสาธารณสุข สำนักงานปลัดกระทรวงสาธารณสุข กระทรวงสาธารณสุข ในเขตภาคกลาง จำนวน 864 คน ได้มาจากการสุ่มอย่างง่าย มาจัดเป็นกลุ่มเทียบคะแนน 12 กลุ่ม ที่มีความสามารถทางสมองเท่าเทียมกัน ทำการทดสอบด้วยแบบทดสอบวัดการคิดเอกนิยมแบบสัมพันธ์ที่มีฟอร์มต่างกัน สามฟอร์ม คือ ฟอร์ม A ฟอร์ม B และฟอร์ม C โดยใช้ฟอร์ม C จัดทำเป็นแบบทดสอบหลักที่มีลักษณะการใช้ของแบบทดสอบหลัก ที่เป็นแบบทดสอบเชื่อมโยงและเป็นข้อสอบรวม และมีขนาดความยาวสามขนาด คือ คิดเป็นสัดส่วน .2, .3 และ .4 ของแบบทดสอบเทียบคะแนนฟอร์ม A แล้วนำคะแนนสอบมาวิเคราะห์ทางสถิติ โดยการแปลงคะแนนดิบให้เป็นคะแนนแปลง ด้วยวิธีการคำนวณจากสูตร สี่วิธี แล้วจึงเปรียบเทียบคะแนนแปลง และความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน ด้วยวิธีการหาความแปรปรวนแบบการวัดซ้ำ

ผลการศึกษา พบว่า คะแนนแปลงที่ได้จากวิธีการคำนวณด้วยสูตร สี่วิธี ให้ผลแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เมื่อเปรียบเทียบผลของการใช้แบบทดสอบเชื่อมโยง กับข้อสอบรวม พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 รวมทั้งผลของขนาดความยาวของแบบทดสอบหลัก ก็พบว่า มีความแตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ในตนเองเดียวกัน ผลของการเปรียบเทียบความคลาดเคลื่อนมาตรฐาน ที่เกิดจากการใช้วิธีการคำนวณด้วยสูตร สี่วิธี ก็พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เช่นกัน

A STUDY OF LINEAR EQUATING METHODS UNDER CONGENERIC MODEL

AN ABSTRACT

BY

SUMANA SOTTIPOLANUN

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

May 1996

This study was the comparison of the four observed-score equating methods. The assumptions of the tests to be equated are congeneric and equally reliable. The samples of this study were the first and second year student nurses of Diploma in Nursing Science (B.Sc.status) and Certificate in Nursing Science (Technical Level) in 1995 of Nursing College of Public Health Ministry in the central area of Thailand. The random sampling was 864 student nurses divided into 12 equated-groups. They had the same ability and were tested by the Convergent-Relation Test. The tests were three forms (Form A, Form B and Form C). Form C called the Anchor Test was used in the external and internal situations and the proportion of its length was .2 , .3 and .4 of Form A. Statistical analysis was done on the raw scores and then was compared on the equating-scores derived from four methods by analysis of variance with a repeated measure on A dimension. Again analysis was done and compared with the value of the total error in terms Root Mean Squared Error.

The result of the study showed that the four observed-score equating methods were differed by the significance at the level .01. The external and internal situations of the Anchor Test were differed by the significance at the level .01. The three sizes of the length of the Anchor Test were differed by the significance at the level .01. And in the same way the Root Mean Squared Error from the four observed-scoreequating methods were differed by the significance at the level .01.