

371.2721

พ ๒๘๖๗.

ร.๓

ผลการเทียบคะแนนแบบทดสอบวิชาการพยาบาลสูติศาสตร์
ตามทฤษฎีการตอบข้อสอบที่ใช้วิธีการวิเคราะห์ต่างกัน

ปริญญานิพนธ์

ของ

พรโสภา พิณใจผล

19 S.A. 2539

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา

ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

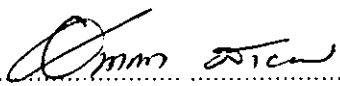
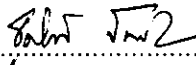
พฤษภาคม 2539

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

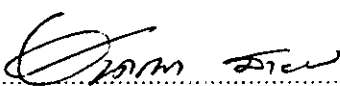
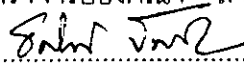
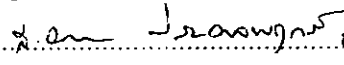
๒ 52325

คณะกรรมการควบคุมและคณะกรรมการสอบได้พิจารณาปริญญาบัตรฉบับนี้แล้ว
เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาการศึกษามหาบัณฑิต
วิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒได้

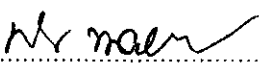
คณะกรรมการควบคุม

 ประธาน
(รองศาสตราจารย์อังคณา สายยศ)
 กรรมการ
(ดร. ชูศักดิ์ ชัมภลิจิต)

คณะกรรมการสอบ

 ประธาน
(รองศาสตราจารย์อังคณา สายยศ)
 กรรมการ
(ดร. ชูศักดิ์ ชัมภลิจิต)
 กรรมการที่แต่งตั้งเพิ่มเติม
(รองศาสตราจารย์ ดร. ส. วาสนา ประवालพฤษ์)

บัณฑิตวิทยาลัยอนุมัติให้รับปริญญาบัตรฉบับนี้ เป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษาของ
มหาวิทยาลัยศรีนครินทรวิโรฒ

 คณบดีบัณฑิตวิทยาลัย
(ดร. ศิริฎา พูลสุวรรณ)

วันที่ 1 เดือน พฤษภาคม พ.ศ. 2539

ประกาศคุณูปการ

ในการทำปริญญาบัตรฉบับนี้ ผู้วิจัยได้รับความกรุณาจากรองศาสตราจารย์ อังคณา สายยศ และ ดร. ชูศักดิ์ ชัมภลลิขิต กรรมการที่ปรึกษาปริญญาบัตร ที่กรุณาให้ คำปรึกษา แนะนำ ตลอดจนข้อคิดเห็นต่างๆ ในการทำปริญญาบัตรมาโดยตลอดตั้งแต่ต้น จนสำเร็จลุล่วงไปด้วยดี ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง

ขอกราบขอบพระคุณ รองศาสตราจารย์ ดร. ส.วาสนา ประवालพฤษ์ ที่กรุณา ให้ข้อเสนอแนะ เพื่อให้ปริญญาบัตรฉบับนี้สมบูรณ์ยิ่งขึ้น

ขอขอบคุณ พี่ๆ และเพื่อนๆ ชาววัดผลที่เป็นแรงสนับสนุนที่สำคัญในการ คิดแก้ปัญหาต่างๆ ที่เกิดขึ้นในการทำปริญญาบัตรฉบับนี้มาโดยตลอด

สุดท้ายผู้วิจัยขอน้อมรำลึกถึงพระคุณของพ่อพันธ์ และแม่แสงนิล พิณีผล ขอขอบคุณ คุณนพดล พิณีผล ที่เชื้อเพื่อในการพิมพ์ด้วยคอมพิวเตอร์ แพทย์หญิงศศิธร พิณีผลและพี่น้องทุกคนที่ให้กำลังใจและให้การสนับสนุนในการศึกษาปริญญาโทของข้าพเจ้า มาโดยตลอด

พรโสภา พิณีผล

สารบัญ

บทที่	หน้า
1 บทนำ.....	1
ภูมิหลัง.....	1
ความมุ่งหมายของการศึกษาค้นคว้า.....	4
ขอบเขตของการศึกษาค้นคว้า.....	4
ขอบเขตของการศึกษาค้นคว้า.....	4
ข้อจำกัดของการวิจัย.....	5
คำนิยามศัพท์เฉพาะ.....	6
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	8
แนวคิดเกี่ยวกับการเทียบคะแนน.....	8
ทฤษฎีที่เกี่ยวข้องกับการเทียบคะแนน.....	14
งานวิจัยที่เกี่ยวข้องกับการเทียบคะแนน.....	24
สมมติฐานของการศึกษาค้นคว้า.....	31
3 วิธีดำเนินการศึกษาค้นคว้า.....	33
ประชากร.....	33
กลุ่มตัวอย่าง.....	33
เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	34
วิธีดำเนินการสร้างแบบทดสอบ.....	35
วิธีดำเนินการเก็บรวบรวมข้อมูล.....	39
การวิเคราะห์ข้อมูลและสถิติที่ใช้.....	39
4 ผลการวิเคราะห์ข้อมูล.....	46
สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล.....	46
การเสนอผลการวิเคราะห์ข้อมูล.....	47

บทที่	หน้า
ผลการวิเคราะห์ข้อมูล.....	47
5 สรุปผล อภิปรายผล และข้อเสนอแนะ.....	64
ความมุ่งหมายของการศึกษาค้นคว้า.....	64
กลุ่มตัวอย่าง.....	64
เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	64
วิธีดำเนินการเก็บรวบรวมข้อมูล.....	65
สรุปผลการวิเคราะห์ข้อมูล.....	66
อภิปรายผล.....	67
ข้อเสนอแนะ.....	69
บรรณานุกรม.....	71
ภาคผนวก.....	76
ประวัติย่อของผู้วิจัย.....	123

บัญชีตาราง

ตาราง	หน้า
1 การแจกแจงคะแนนดิบของผู้สอบที่มีความสามารถ θ สำหรับข้อสอบ 3 ข้อ.....	23
2 รายชื่อวิทยาลัยและจำนวนนักศึกษาในกลุ่มตัวอย่างที่ใช้ในการศึกษา ค้นคว้าครั้งนี้.....	34
3 ค่าสถิติพื้นฐานของแบบทดสอบวิชาการพยาบาลสุติศาสตร์ 1.....	47
4 คะแนนสมมูลของการเทียบคะแนนจากแบบทดสอบฉบับ X ไปสู่ แบบทดสอบฉบับ Y ที่ระดับคะแนนและความสามารถเดียวกันตาม วิธีการเทียบคะแนนจริง.....	50
5 คะแนนสมมูลของการเทียบคะแนนจากแบบทดสอบฉบับ X ไปสู่ แบบทดสอบฉบับ Y ที่ระดับคะแนนและความสามารถเดียวกันตาม วิธีการเทียบคะแนนดิบ.....	53
6 คะแนนสมมูลของการเทียบคะแนนทั้งสองวิธีที่ระดับคะแนนและความ สามารถเดียวกัน.....	55
7 การเปรียบเทียบความแตกต่างของคะแนนที่ได้จากการเทียบคะแนนของ แบบทดสอบทั้งสองวิธี.....	56
8 การเปรียบเทียบความแตกต่างของคะแนนที่ได้จากการเทียบคะแนนของ แบบทดสอบทั้งสองวิธีเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB	57
9 ความคลาดเคลื่อนของการเทียบคะแนนทั้งสองวิธีที่ระดับคะแนนและความ สามารถเดียวกัน.....	57
10 การเปรียบเทียบความแตกต่างของคะแนนความคลาดเคลื่อนที่ได้จากการ เทียบคะแนนของแบบทดสอบทั้งสองวิธี.....	60
11 ความเพียงพอของการเทียบคะแนนทั้งสองวิธีที่ระดับความสามารถเดียวกัน.....	61
12 การเปรียบเทียบความเพียงพอของการเทียบคะแนนทั้งสองวิธี.....	63
13 ค่าไอเกนขององค์ประกอบที่มีค่าเกิน 1.0 ของแบบทดสอบ วิชาการพยาบาลสุติศาสตร์ 1.....	98
14 ผลการวิเคราะห์ค่าพารามิเตอร์จากการทดลองเครื่องมือ.....	100

15	ค่าพารามิเตอร์ของแบบทดสอบเทียบคะแนนและแบบทดสอบรวมทั้งสองฉบับ.....	105
16	ค่าพารามิเตอร์ของแบบทดสอบทั้งสองฉบับ.....	111
17	การคำนวณคะแนนที่ได้จากการเทียบคะแนนจริง.....	114
18	การคำนวณคะแนนที่ได้จากการเทียบคะแนนดิบ.....	118
19	การคำนวณคะแนนที่ได้จากการเทียบคะแนนทั้งสองวิธี.....	121

บัญชีภาพประกอบ

ภาพประกอบ	หน้า
1 แบบแผนการเก็บรวบรวมข้อมูลเพื่อเทียบคะแนนแบบทดสอบ.....	13
2 การเทียบคะแนนตามทฤษฎีการตอบข้อสอบ ระหว่างแบบทดสอบฉบับ X และ ฉบับ Y.....	19
3 แผนผังการสร้างแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา พย. 212.....	35
4 คะแนนจริงของผู้สอบแบบทดสอบทั้งสองฉบับที่ระดับคะแนนและความ สามารถต่างๆ.....	49
5 คะแนนที่คาดหวังของผู้สอบแบบทดสอบทั้งสองฉบับที่คะแนนโอกาส ในการตอบข้อสอบถูก.....	52
6 ความคลาดเคลื่อนของการเทียบคะแนนทั้งสองวิธีที่ระดับคะแนนและความ สามารถเดียวกัน.....	58
7 โค้งลักษณะเฉพาะของแบบทดสอบทั้งสองฉบับ.....	108
8 โค้งอินฟอर्मชันของแบบทดสอบที่ระดับความสามารถต่างๆ.....	109

บทที่ 1

บทนำ

ภูมิหลัง

๐ ผู้เป็นครูย่อมมีภาระหน้าที่มากมายหลายอย่างนอกเหนือไปจากการสอนและการเตรียมการสอน ในบรรดาภาระหน้าที่นั้น การประเมินผลการศึกษาเป็นหน้าที่ที่สำคัญที่สุดอย่างหนึ่งของครูเพราะเป็นเครื่องแสดงให้ทราบว่านักเรียนแต่ละคนมีความเจริญงอกงามเพียงใด และนักเรียนเหล่านั้นได้บรรลุจุดประสงค์ที่เราวางไว้หรือไม่ การวัดผลการเรียนรู้การสอนเป็นงานของครูที่จะหลีกเลี่ยงไม่ได้ เป็นงานที่เปลืองเวลาและต้องกระทำต่อเนื่องกันไป เป็นงานที่จุกจิกต้องการความประณีตละเอียดรอบคอบ ต้องใช้สมองและเทคนิคและคุณธรรมอันสูงส่ง (เสริมศักดิ์ วิชาลาภรณ์ และเอนกกุล กริแสง. 2517 : 1) ✓

ในการเรียนการสอนทางการพยาบาล ต้องใช้การวัดและประเมินผลทางการศึกษาเช่นเดียวกัน เนื่องจากวิชาชีพพยาบาล เป็นวิชาชีพที่เน้นด้านการปฏิบัติงานเป็นสำคัญ (Practical-oriented Discipline) ดังนั้นการพยาบาลศึกษา จึงเป็นการเตรียมพยาบาลที่มีความสามารถทั้งในด้านวิชาการตามลักษณะของวิชาชีพและต้องมีความสามารถในการปฏิบัติการพยาบาล การเรียนการสอนจึงประกอบไปด้วยส่วนสำคัญสองส่วนคือภาคทฤษฎีและภาคปฏิบัติซึ่งต้องสำคัญและสอดคล้องประสมประสานกันไปและเป็นที่ยอมรับในวิชาชีพทางการพยาบาล ดังนั้นการสอนพยาบาลภาคปฏิบัติ มีจุดมุ่งหมายที่สำคัญคือ

1. ให้นักศึกษาได้นำความรู้ภาคทฤษฎีประยุกต์ในการปฏิบัติทางการพยาบาล
2. ส่งเสริมให้นักศึกษาเกิดการเรียนรู้ทางด้านสติปัญญา ทักษะและทัศนคติในวิชาชีพ

3. ส่งเสริมให้นักศึกษา มีทักษะที่จำเป็นในการปฏิบัติทางการพยาบาล เช่นทักษะการแก้ปัญหา

4. แสดงความสัมพันธ์ระหว่างภาคทฤษฎีกับภาคปฏิบัติ

5. สนับสนุนให้นักศึกษาเกิดความคิดอย่างมีอิสระ (จินตนา ยูนิพันธ์. 2527)

จากขบวนการเรียนการสอนดังกล่าว จะเห็นได้ว่า ส่วนใหญ่เน้นการฝึกภาคปฏิบัติควบคู่ไปกับภาคทฤษฎี ดังนั้นนักศึกษาพยาบาลจึงต้องมีการฝึกภาคปฏิบัติดังกล่าว การศึกษาพยาบาลในยุคในตึงเกล้าว่า ผู้ที่จะเป็นพยาบาลได้ต้องได้รับการฝึกหัดในโรงพยาบาล (จินตนา

ยูนิพันธ์. 2527 : 20 ; อ้างอิงมาจาก Griffin and Griffin. 1965) นักศึกษาจึงต้องมีการฝึกปฏิบัติงานบนหอผู้ป่วย แต่ในการฝึกภาคปฏิบัติ นั้น เนื่องจากจำนวนนักศึกษาในแต่ละปีมีจำนวนมาก เพื่อให้ให้การเรียนการสอนมีประสิทธิภาพ จึงต้องมีการแบ่งนักศึกษาออกเป็นกลุ่มๆ ในการฝึกหัดในโรงพยาบาล ทั้งนี้เพื่อประสิทธิภาพของการเรียนการสอนให้เป็นไปด้วยดี หลังจากสิ้นสุดการฝึกปฏิบัติ ครูต้องมีการประเมินผลเพื่อทราบความก้าวหน้าของการฝึกหัด ซึ่งจะหลีกเลี่ยงไม่ได้ก็คือการสอบ ดังนั้นเพื่อป้องกันไม่ให้เกิดปัญหาข้อสอบรั่วหรือเกิดจากการได้เปรียบเสียเปรียบกันในระหว่างกลุ่ม ~~อื่น~~ เนื่องจากมาจากนักศึกษาแต่ละกลุ่มสอบไม่พร้อมกัน จึงมีผู้ประยุกต์วิธีการทางสถิติมาใช้ในการที่จะทำให้คะแนนจากแบบทดสอบที่วัดคุณลักษณะเดียวกันแต่ต่างฉบับกัน สามารถนำผลมาเปรียบเทียบกันได้อย่างมีความหมายมากขึ้น วิธีการดังกล่าวคือ "การเทียบมาตรฐาน" (Equating) ซึ่งซุคคักดี ชัมภลชิต (2527 : 144) ได้สรุปว่า การเทียบมาตรฐานเป็นกระบวนการที่เกี่ยวข้องกับกิจกรรม 2 ประการคือ 1) กระบวนการที่ทำให้แบบทดสอบ 2 ฉบับใด ๆ มีความทัดเทียมกันในเชิงของโครงสร้างและ 2) การใช้วิธีการทางสถิติเพื่อปรับ(Adjust)คะแนนที่ได้จากแบบทดสอบแต่ละฉบับให้อยู่ในมาตรฐานเดียวกันและเทียบกันได้ และข้อตกลงเบื้องต้นของการเทียบมาตรฐานก็คือ "แบบทดสอบ 2 ฉบับ ใดๆ จะเทียบคะแนนกันได้ก็ต่อเมื่อแบบทดสอบทั้ง 2 ฉบับนั้นคู่ขนานกันในด้านต่อไปนี้คือ

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

ในการเทียบมาตรฐาน เป็นการกำหนดเงื่อนไขและสร้างกฎการแปลงคะแนนในแต่ละครั้งที่ทำการเทียบมาตรฐาน ดังนั้นจึงมีรูปแบบการเทียบมาตรฐาน (Equating Model) ที่ต่าง ๆ กัน ออกไปหลายรูปแบบ ซึ่งถ้าใช้เกณฑ์ในเชิงทฤษฎีการวัดและประเมินผลพอจะแบ่งได้ 2 รูปแบบคือ

1. รูปแบบการเทียบคะแนนโดยการใช้ทฤษฎีคลาสสิกอล ซึ่งแองกอฟ (Hambleton and Swaminathan. 1985 : 200 ; citing Angoff.1971) ได้กล่าวว่ามี 3 วิธีคือ วิธีอีควิเปอร์เซ็นไทล์ (Equipercentile Equating) วิธีการเทียบคะแนนแบบเส้นตรง (Linear Equating) และวิธีการ Regression (Regression Method)

2. รูปแบบการเทียบคะแนนโดยการใช้ทฤษฎีการตอบข้อสอบ ซึ่งภายใต้เงื่อนไขบางประการของทฤษฎีการตอบข้อสอบหรือเรียกอีกชื่อว่า "ทฤษฎีประมาณค่าคุณลักษณะภายใน" (Latent Trait Theory) จะสามารถคำนวณค่าความสามารถแท้ θ ของผู้สอบได้ชัดเจนและสามารถเชื่อมโยงความสามารถของผู้สอบเข้ากับโอกาสที่เขาจะทำข้อสอบได้ถูกต้องได้อย่างชัดเจนในลักษณะฟังก์ชัน Logistic ที่แสดงในรูปเส้นโค้งลักษณะเฉพาะของข้อสอบแต่ละข้อ (Item Characteristic Curve) จากการเทียบมาตราโดยการใช้ทฤษฎีคลาสสิคอล โคลเลน (Kolen. 1981) กล่าวว่าไม่มีเหตุผลเพียงพอในการที่จะอธิบายถึงความทัดเทียมกัน (Equity) ความสมมาตร (Symmetry) และความไม่ผันแปรตามกลุ่มตัวอย่าง (Invariance) แต่ในการเทียบมาตราโดยการใช้ทฤษฎีการตอบข้อสอบ สามารถที่จะแก้ปัญหานี้ได้ จากการศึกษาของ ปีเตอร์สันและคณะ (Petersen and others. 1982 : 71 -135) ได้ศึกษาเปรียบเทียบผลการใช้รูปแบบเชิงเส้นตรงรูปแบบอิกวิเปอร์เซ็นไทล์และรูปแบบการใช้ทฤษฎีการตอบข้อสอบ พบว่ารูปแบบการใช้ทฤษฎีการตอบข้อสอบสามารถประมาณค่าของคะแนนการสอบได้แน่นอนและเหมาะสมกว่าทฤษฎีคลาสสิคอล และจากการศึกษาของโคลเลนและวิทนี (Kolen and Whitney. 1982 : 279-293) ได้ศึกษาความเพียงพอของการเทียบคะแนนโดยเปรียบเทียบวิธีการเดียวกันกับปีเตอร์สันและคณะ พบว่ารูปแบบการเทียบคะแนนโดยการใช้ทฤษฎีการตอบข้อสอบให้ความเพียงพอมากที่สุด จากคุณสมบัติดังกล่าวของการเทียบคะแนนโดยการใช้ทฤษฎีการตอบข้อสอบ ได้มีวิธีการเทียบคะแนนหลายวิธีคือ 1) การเทียบคะแนนโดยการใช้คะแนนจริง 2) การเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วม 3) การเทียบคะแนนดิบโดยการใช้แบบทดสอบร่วม และ 4) วิธีการเทียบคะแนนล่วงหน้า (Preequating) (Lord. 1980 : 199-206)

ผลที่ การวิจัยเกี่ยวกับคะแนนดิบ
 ดังนั้นในการศึกษาวิจัยครั้งนี้ ผู้วิจัยมีความสนใจที่จะศึกษาวิธีการเทียบคะแนนตามทฤษฎีการตอบข้อสอบ โดยเปรียบเทียบระหว่างวิธีการเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วม และวิธีการเทียบคะแนนดิบโดยการใช้แบบทดสอบร่วมในแบบทดสอบวิชาการพยาบาลสุติศาสตร์ เพื่อประโยชน์ในการเรียนการสอนวิชาการพยาบาลสุติศาสตร์ ซึ่งมีการวัดและประเมินผลภาคทฤษฎีหลังสิ้นสุดการฝึกภาคปฏิบัติโดยการใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนจำนวน 2 ฉบับในการประเมินผลการสอบ ดังนั้นจึงต้องใช้วิธีการเทียบคะแนนดังกล่าวเพื่อที่จะศึกษาวิธีที่ให้ผลในการเทียบคะแนนเหมาะสมและดีที่สุดในการ

การตัดสินใจที่จะเลือกวิธีการเทียบคะแนนตาม ทฤษฎีการตอบข้อสอบวิธีใดวิธีหนึ่งจากสองวิธี ในการนำไปใช้ให้สมบูรณ์ยิ่งขึ้นไป

ความมุ่งหมายของการศึกษาค้นคว้า

ในการศึกษาครั้งนี้ ผู้วิจัยมีความมุ่งหมายเพื่อศึกษาเปรียบเทียบผลการเทียบคะแนนแบบทดสอบวิชาการพยาบาลสูติศาสตร์ จำนวนสองฉบับด้วยวิธีการเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วมและวิธีการเทียบคะแนนดิบโดยการใช้แบบทดสอบร่วมตามทฤษฎีการตอบข้อสอบ โดยมีความมุ่งหมายเฉพาะ ดังนี้

1. เพื่อเปรียบเทียบคะแนนที่เทียบได้ของแบบทดสอบวิชาการพยาบาลสูติศาสตร์ ระหว่างวิธีการวิเคราะห์ต่างกัน
2. เพื่อเปรียบเทียบความคลาดเคลื่อนของการเทียบคะแนนระหว่างวิธีการวิเคราะห์ต่างกัน
3. เพื่อเปรียบเทียบความเพียงพอของการเทียบคะแนน ระหว่างวิธีการวิเคราะห์ต่างกัน

ความสำคัญของการศึกษาค้นคว้า

ผลจากการศึกษาค้นคว้าครั้งนี้ ทำให้ทราบว่า วิธีการเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วม และวิธีการเทียบคะแนนดิบโดยการใช้แบบทดสอบร่วม ตามทฤษฎีการตอบข้อสอบ วิธีใดทั้ง 2 วิธีที่ให้ผลในการเทียบคะแนนเหมาะสมและดีที่สุด อันจะเป็นประโยชน์ในการตัดสินใจที่จะเลือกวิธีการเทียบคะแนนตามทฤษฎีการตอบข้อสอบ ไปใช้ในการวัดและประเมินผลนักศึกษาพยาบาลได้อย่างเหมาะสม อีกทั้งทำให้ได้แบบทดสอบคู่ขนานที่สามารถนำไปใช้ในการเทียบคะแนนของการสอบได้อย่างมีความหมาย และมีประสิทธิภาพยิ่งขึ้นในการสอบแต่ละครั้งต่อไป

ขอบเขตของการศึกษาค้นคว้า

1. ประชากร ที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักศึกษาพยาบาลหลักสูตรพยาบาลศาสตร์(ระดับต้น) ชั้นปีที่ 2 ปีการศึกษา 2538 ของวิทยาลัยพยาบาลในสังกัดสถาบัน

พัฒนากำลังคนด้านสาธารณสุข กระทรวงสาธารณสุข ในเขตภาคเหนือ ภาคกลางและภาคตะวันออกเฉียงเหนือ จำนวนนักศึกษา 2,050 คน จำนวน 20 วิทยาลัย

2. กลุ่มตัวอย่าง ที่ใช้ในการศึกษาค้างนี้ เป็นนักศึกษาพยาบาลหลักสูตรพยาบาลศาสตร(ระดับต้น) ชั้นปีที่ 2 ปีการศึกษา 2538 ของวิทยาลัยพยาบาลในสังกัดสถาบันพัฒนากำลังคนด้านสาธารณสุข กระทรวงสาธารณสุข ในเขตภาคเหนือ ภาคกลางและภาคตะวันออกเฉียงเหนือ จำนวนนักศึกษา 980 คน จำนวน 10 วิทยาลัย ได้มาโดยการสุ่มแบบอย่างง่ายและมีวิทยาลัยเป็นหน่วยการสุ่ม

3. ตัวแปรที่ใช้ในการศึกษา

3.1 ตัวแปรอิสระ คือวิธีการเทียบคะแนนโดยการใช้ทฤษฎีการตอบข้อสอบ
2 วิธีคือ

3.1.1 การเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วม(True Score Equating with an Anchor Test)

3.1.2 การเทียบคะแนนดิบโดยการใช้แบบทดสอบร่วม(Raw Score Equating with an Anchor Test)

3.2 ตัวแปรตาม คือผลของการเทียบคะแนนซึ่งพิจารณาได้จาก

3.2.1 คะแนนที่เทียบได้แต่ละวิธี

3.2.2 ความคลาดเคลื่อนของการเทียบคะแนน

3.2.3 ความเพียงพอของการเทียบคะแนน

ข้อจำกัดของการวิจัย

1. ในการศึกษาครั้งนี้ ได้จำกัดขอบเขตของแบบทดสอบที่ใช้ในการเทียบคะแนนคือเป็นแบบชนิดเลือกตอบที่ใช้ในการวัดผลภาคทฤษฎีเท่านั้น ซึ่งผู้วิจัยสร้างขึ้นเองและใช้ในการวัดผลภาคปฏิบัติของวิชาการพยาบาลสูติศาสตร์

2. วิธีการวิเคราะห์ที่ใช้ในการเทียบคะแนนคือวิธีการเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วมและวิธีการเทียบคะแนนดิบโดยการใช้แบบทดสอบร่วมตามทฤษฎีการตอบข้อสอบเท่านั้น

คำนิยามศัพท์เฉพาะ

✓ 1. การเทียบคะแนน หมายถึง กระบวนการที่ใช้วิธีการทางสถิติเพื่อแปลงคะแนนที่ได้จากแบบทดสอบวิชาการพยาบาลสาธารณสุข 2 ฉบับ (ฉบับ X และ Y) ให้มีสเกลคะแนนเดียวกันและนำมาเปรียบเทียบกันได้

✓ 2. การเทียบคะแนนโดยการใช้ทฤษฎีการตอบข้อสอบ หมายถึง การเทียบมาตราในเชิงโครงสร้างที่ใช้คะแนนความสามารถของผู้สอบตามแบบจำลองโลจิสติกสามพารามิเตอร์และคำนวณความสามารถแท้จากโค้งลักษณะเฉพาะของข้อสอบซึ่งมีวิธีการเทียบคะแนน 2 วิธีคือ

2.1 การเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วม หมายถึง การเทียบคะแนนโดยการใช้คะแนนจริงจากแบบทดสอบสองฉบับ (ฉบับ X และ Y) ที่ใช้ความสัมพันธ์เชิงคณิตศาสตร์ระหว่างความสามารถ θ และคำนวณคะแนนตามข้อที่ตอบถูกจากสูตร

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

$$\xi_y = \sum_{i=1}^n P_i(\theta_y)$$

✓ 2.2 การเทียบคะแนนดิบโดยการใช้แบบทดสอบร่วม หมายถึง การเทียบคะแนนดิบโดยการคำนวณหาค่าคะแนนดิบจากแบบทดสอบสองฉบับ (ฉบับ X และ ฉบับ Y) ในแต่ละความสามารถ θ แล้วนำมาเทียบคะแนนโดยการใช้สูตร

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

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

✓ 3. ผลของการเทียบคะแนน หมายถึง ค่าที่พิจารณาได้จาก

3.1 คะแนนที่เทียบได้แต่ละวิธี หมายถึง คะแนนที่ได้จากการเทียบมาตราในวิธีการเทียบคะแนนวิธีที่ 2.1 และ 2.2

3.2 ความคลาดเคลื่อนของการเทียบคะแนน หมายถึง ความคลาดเคลื่อนที่เกิดจากการนำคะแนนสอบจากแบบทดสอบฉบับ X ไปแปลงคะแนน คะแนนที่เทียบได้มีค่าเบี่ยงเบนออกจากคะแนนเดิมซึ่งหาได้จากสูตรความคลาดเคลื่อนที่คิดเป็นรากที่สองของค่าเฉลี่ยของความคลาดเคลื่อน

✓ 3.3 ความเพียงพอของการเทียบคะแนน หมายถึง ค่าที่แสดงถึงคุณภาพของการเทียบคะแนนโดยยึดหลักการประเมินของปีเตอร์เซ็นและคนอื่นๆ (Petersen and others. 1982 : 93-94) ซึ่งมีเกณฑ์ในการพิจารณาดังนี้คือ

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

✓ 4. แบบทดสอบเทียบคะแนน หมายถึง แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วิชาการพยาบาลสูติศาสตร์ 1 สำหรับนักศึกษาพยาบาลศาสตร์(ระดับต้น) ชั้นปีที่ 2 ซึ่งผู้วิจัยสร้างขึ้นและวิเคราะห์หาคุณภาพโดยการใช้ทฤษฎีการตอบข้อสอบ แล้วแบ่งออกเป็นสองฉบับ (ฉบับ X และ Y) โดยพิจารณาค่าโค้งลักษณะข้อสอบ(Item Characteristic Curve) ที่เท่าหรือเทียบเคียงกัน เป็นคู่ๆ

✓ 5. แบบทดสอบร่วม หมายถึง ข้อสอบจำนวน 10 ข้อที่มีอยู่ทั้งในฉบับ X และ Y ที่หามาจากแบบทดสอบที่มีคุณภาพแบบโลจิสติกสามพารามิเตอร์ เพื่อนำผลมาใช้ในการปรับระดับคะแนนของแบบทดสอบฉบับ X และ ฉบับ Y ให้อยู่ในสเกลเดียวกัน

✓ 6. กลุ่มเทียบคะแนน หมายถึง กลุ่มตัวอย่างที่เป็นนักศึกษาพยาบาลศาสตร์ (ระดับต้น) ชั้นปีที่ 2 ที่สอบแบบทดสอบเทียบคะแนนและแบบทดสอบร่วมเพื่อนำผลการสอบไปใช้ในการแปลงคะแนนสำหรับสร้างตารางเทียบคะแนน

✓ 7. กลุ่มทานผล หมายถึง กลุ่มตัวอย่างที่เป็นนักศึกษาพยาบาลศาสตร์(ระดับต้น) ชั้นปีที่ 2 ที่สอบแบบทดสอบเทียบคะแนนและแบบทดสอบร่วมเพื่อนำผลการสอบไปใช้ในการตรวจสอบความคลาดเคลื่อนของการเทียบคะแนนในแต่ละวิธี

✓ 8. คะแนนสมมูล หมายถึง คะแนนจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วิชาการพยาบาลสูติศาสตร์ 1 จากแบบทดสอบฉบับ X และ Y ซึ่งมีคะแนนเท่าหรือใกล้เคียงกันจากการเทียบคะแนนในแต่ละวิธี

✓ 9. ตารางเทียบคะแนน หมายถึง ตารางเปรียบเทียบคะแนนสมมูลจากแบบ แบบทดสอบฉบับ X และ Y จากวิธีการเทียบคะแนนในแต่ละวิธี

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการนำเสนอเอกสารและงานวิจัยที่เกี่ยวข้องผู้วิจัยได้นำเสนอโดยแยกกล่าวตามลำดับดังนี้

ตอนที่ 1. แนวคิดเกี่ยวกับการเทียบคะแนน

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

ตอนที่ 2. ทฤษฎีที่เกี่ยวข้องกับการเทียบคะแนน

- 2.1 การเทียบคะแนนตามทฤษฎีคลาสสิกคอลล
 - 2.1.1 วิธีอิกวิเปอร์เซ็นไทล์
 - 2.1.2 วิธีเทียบแบบเส้นตรง
- 2.2 การเทียบคะแนนตามทฤษฎีการตอบข้อสอบ
 - 2.2.1 วิธีการเทียบคะแนนจริง
 - 2.2.2 วิธีการเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วม
 - 2.2.3 วิธีการเทียบคะแนนดิบโดยการใช้แบบทดสอบร่วม

ตอนที่ 3. งานวิจัยที่เกี่ยวข้องกับการเทียบคะแนน

- 3.1 งานวิจัยต่างประเทศ
- 3.2 งานวิจัยในประเทศ

ตอนที่ 1. แนวคิดเกี่ยวกับการเทียบคะแนน

ความหมายของการเทียบคะแนน

กัลลิกเซน (Gulliksen. 1950 : 298-304) ได้ให้ความหมายของการเทียบคะแนนว่าเป็นวิธีการทำคะแนนที่ได้จากแบบทดสอบสองชุดของวิชาการเดียวกัน ให้เป็นคะแนนสมมูลที่เปรียบเทียบกันได้โดยตรง

ฟลานาแกน (Flanagan. 1951 : 747-748) ให้ความหมายของการเทียบคะแนนว่าเป็นวิธีการทำคะแนนจากแบบทดสอบต่างฉบับ ให้สามารถนำมาเทียบกันได้

แองกอฟ (Angoff. 1971 : 562) ให้ความหมายของการเทียบคะแนน เป็นวิธีการเปลี่ยนระบบหน่วยคะแนนของแบบทดสอบฉบับหนึ่งไปสู่ระบบหน่วยคะแนนของแบบทดสอบอีกฉบับหนึ่ง เพื่อให้ได้คะแนนที่แปลงจากแบบทดสอบทั้งสองฉบับมีความเท่าเทียมกัน

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

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

1. กระบวนการที่ทำให้แบบทดสอบ 2 ฉบับใดๆ มีความทัดเทียมกันหรือเท่ากันในเชิงของโครงสร้างและ

2. การใช้วิธีการทางสถิติเพื่อปรับ (Adjust) คะแนนที่ได้จากแบบทดสอบแต่ละฉบับให้อยู่ในมาตรฐานเดียวกันและเทียบกันได้

จากความหมายของการเทียบคะแนนดังกล่าว ผู้วิจัยพอสรุปได้ว่า การเทียบคะแนนเป็นวิธีการเชิงประจักษ์ที่ใช้วิธีการทางสถิติ ที่ทำให้คะแนนจากแบบทดสอบต่างฉบับที่มีลักษณะเดียวกันให้เปรียบเทียบกันได้

เงื่อนไขของการเทียบคะแนน

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

1. ความทัดเทียมกัน (Equity) ลอร์ด (Lord. 1980 : 195) กล่าวว่า "ถ้าการเทียบคะแนนของแบบทดสอบฉบับ X และ Y มีความยุติธรรมต่อผู้สอบแล้ว จะต้องไม่มีความแตกต่างที่ทุกๆระดับความสามารถ θ ไม่ว่าจะสอบด้วยแบบทดสอบ ฉบับ X หรือ Y ก็ตาม" และมีข้อกำหนดดังนี้

1.1 แบบทดสอบที่วัดคุณลักษณะ (Trait) หรือความสามารถ (Ability) ต่างกัน จะไม่สามารถเทียบคะแนนกันได้

1.2 คะแนนแบบทดสอบที่มีค่าความเชื่อมั่นไม่เท่ากันแล้ว ไม่สามารถเทียบกันได้

1.3 คะแนนของแบบทดสอบที่มีระดับความยากต่างกัน ในการเทียบคะแนนตามแนวคิด ไม่สามารถจะเทียบคะแนนกันได้ (เนื่องจากคะแนนจริงของแบบทดสอบจะมีความสัมพันธ์กันแบบไม่เป็นเส้นตรง และทำให้แบบทดสอบจะมีค่าความเชื่อมั่นไม่เท่ากัน ระดับความสามารถ (θ) ที่แตกต่างกัน)

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

1.5 คะแนนที่อาจผิดพลาดจากแบบทดสอบฉบับ X และ Y ไม่สามารถจะเทียบกันได้ นอกจากว่าแบบทดสอบฉบับ X และ Y คู่ขนานกัน

1.6 แบบทดสอบที่มีความเชื่อมั่นที่ดีสามารถเทียบกันได้

2. ความไม่แปรเปลี่ยนไปตามกลุ่ม (Invariance Across Group) คือคะแนนของกลุ่มต่างๆที่ผ่านการเทียบแล้ว จะมีค่าไม่แตกต่างกันไม่ว่าจะเทียบจากกลุ่มตัวอย่างใด

3. ความสมมาตร (Symmetry) คือผลของการเทียบคะแนนจากแบบทดสอบฉบับ X ไปสู่ฉบับ Y หรือจากฉบับ Y ไปสู่ฉบับ X จะต้องเหมือนกัน

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

ประเภทของการเทียบคะแนน

สงบ. ลักษณะ (2525 : 23) ได้แบ่งการเทียบคะแนนออกเป็น 2 กรณีคือ

1. การเทียบคะแนนตามแนวนอน (Horizontal Equating) เป็นกรณีที่มีข้อสอบ 2 ฉบับคือ ฉบับ ก. และ ข. มีความยากง่ายปานๆกัน นำมาทดสอบกับบุคคลประชากรเดียวกัน แล้วต้องหาว่าที่คะแนนของแบบทดสอบชุด ก. จึงจะเทียบเท่ากับที่คะแนนของแบบทดสอบชุด ข.

2. การเทียบคะแนนตามแนวตั้ง (Vertical Equating) เป็นกรณีที่มีแบบทดสอบสองชุด หรือมากกว่า 2 ชุด มีความยากต่างกัน นำไปทดสอบประชากรบุคคลต่างระดับ

ความสามารถกัน กล่าวคือต้องการเทียบคะแนนจากแบบทดสอบสองชุดเมื่อความยากของข้อสอบต่างกัน และเมื่อสอบกับบุคคลที่มีความสามารถต่างกัน การเทียบมาตราในลักษณะนี้ในลักษณะนี้จะกระทำโดยมีแบบทดสอบร่วมฉบับหนึ่ง

แบบแผนการเก็บรวบรวมข้อมูล

ข้อมูลจากการสอบโดยใช้แบบทดสอบหลายฉบับและเก็บรวบรวมข้อมูล โดยใช้รูปแบบต่างนั้น ข้อมูลที่ได้จากบางรูปแบบไม่สามารถนำมาเทียบคะแนนได้ เช่น ข้อมูลที่ได้จากการสอบโดยใช้แบบทดสอบ 2 ฉบับ กับผู้สอบที่มีความสามารถแตกต่างกัน 2 กลุ่ม กับผู้สอบที่มีความแตกต่างกัน 2 กลุ่ม จะนำคะแนนที่ได้จากแบบทดสอบต่างฉบับมาเปรียบเทียบกันไม่ได้ โดยทั่วไปคะแนนจากแบบทดสอบต่างฉบับที่จะนำมาเปรียบเทียบกันได้ ต้องได้มาจากผู้สอบกลุ่มเดียวกัน หรือผู้สอบกลุ่มเท่าเทียมกัน หรือได้จากผู้สอบที่สอบข้อสอบร่วมกัน รูปแบบใดรูปแบบหนึ่ง แบบแผนการเก็บรวบรวมข้อมูล ซึ่งสามารถนำคะแนนจากแบบทดสอบต่างฉบับมาเปรียบเทียบกันได้ จำแนกได้ดังนี้ (Petersen and others. 1989 : 244-246)

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

แบบแผนที่ 2 ผู้สอบกลุ่มเดียวที่ได้รับการจัดให้สมดุล(Counterbalanced Random Group Design) ในทางปฏิบัติ ลำดับของการทดสอบก่อนหลัง ย่อมส่งผลต่อการสอบ ดังนั้นแบบแผนนี้จึงปรับมาจากแบบแผนแรกโดยจัดผู้สอบเป็นกลุ่มสุ่ม 2 กลุ่ม แต่ละกลุ่มย่อยได้รับการสอบด้วยแบบทดสอบ 2 ฉบับ ในลักษณะการจัดลำดับก่อนหลัง สลับให้เกิดความสมดุล การจัดแบบนี้เพื่อให้การสอบโดยใช้แบบทดสอบ 2 ฉบับ ได้รับอิทธิพลจากองค์ประกอบในเรื่องการเรียนรู้ ความเมื่อยล้าและการฝึกฝนในลักษณะที่สมดุลกัน

แบบแผนที่ 3 ผู้สอบกลุ่มเท่าเทียมกัน(Equivalent Group Design) ในทางปฏิบัติ การให้ผู้สอบทำแบบทดสอบทั้ง 2 ฉบับ อาจจะทำได้ยาก การใช้ผู้สอบกลุ่มเท่าเทียมกันเป็นรูปแบบที่สามารถนำมาใช้ได้ แบบแผนนี้จำเป็นต้องจัดกลุ่มผู้สอบให้มีความคล้ายคลึงกันมากเท่าที่จะทำได้ ผู้สอบในแต่ละกลุ่มจะทำแบบทดสอบฉบับเดียว ดังนั้นจึงหลีกเลี่ยงปัญหาเรื่อง

การเรียนรู้ ความเมื่อยล้าและการฝึกฝนจากการสอบครั้งแรกได้ แต่จะไม่มีข้อมูลที่จะนำมาใช้ปรับความแตกต่างของกลุ่มตัวอย่างทั้งสอง วิธีการที่จะลดความแตกต่างระหว่างกลุ่มทำได้โดยใช้กลุ่มตัวอย่างขนาดใหญ่

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

แบบแผนที่ 5 กลุ่มผู้สอบที่ไม่เท่าเทียมกันโดยมีแบบทดสอบร่วม(Anchor Test Non-Equivalent Group Design) ในการสอบบางครั้งจำเป็นต้องใช้แบบทดสอบฉบับใหม่กับผู้สอบกลุ่มใหม่ในโปรแกรมการสอบเดียวกัน การสอบโดยใช้แบบทดสอบต่างฉบับและเวลาสอบต่างกันนี้ ผู้สอบไม่ได้มาจากกลุ่มประชากรเดียวกัน จึงจำเป็นต้องใช้แบบทดสอบร่วมในการปรับความแตกต่างระหว่างกลุ่ม ตามแบบแผนนี้ผู้สอบในแต่ละกลุ่มจะได้รับแบบทดสอบ 1 ฉบับและแบบทดสอบร่วม ซึ่งต้องมีเนื้อหาและระดับความยากคล้ายกับแบบทดสอบที่ต้องการเทียบคะแนนให้มากที่สุดเท่าที่จะเป็นไปได้ คะแนนจากแบบทดสอบร่วมจะใช้เพื่อลดความลำเอียงในการเทียบคะแนนจากกลุ่มที่ต่างกัน

จากแบบแผนการเก็บรวบรวมข้อมูลที่กล่าวมาทั้งหมด สามารถอธิบายได้โดยภาพประกอบ 1 ดังนี้

แบบแผนที่ 1

กลุ่มตัวอย่าง	แบบทดสอบฉบับที่ 1	แบบทดสอบฉบับที่ 2
P1	*	*

แบบแผนที่ 2

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

แบบแผนที่ 3

กลุ่มตัวอย่าง	แบบทดสอบฉบับที่ 1	แบบทดสอบฉบับที่ 2
P1	*	
P2		*

แบบแผนที่ 4

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

แบบแผนที่ 5

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

ภาพประกอบ 1 แบบแผนการเก็บรวบรวมข้อมูลเพื่อเทียบคะแนนแบบทดสอบ

ตอนที่ 2. ทฤษฎีที่เกี่ยวข้องกับการเทียบคะแนน

1. การเทียบคะแนนตามทฤษฎีคลาสสิกอล

1.1 วิธีอิกวิเปอร์เซ็นไทล์ การเทียบคะแนนโดยวิธีการนี้อยู่บนพื้นฐานของนิยามที่ว่า คะแนนจากแบบทดสอบฉบับ X และฉบับ Y จะสมมูลกันถ้าการแจกแจงความถี่ของแบบทดสอบฉบับ X และฉบับ Y สำหรับประชากรผู้สอบที่มีลักษณะคล้ายกัน ตามที่แองกอฟ (Angoff, 1982) กล่าวว่า "คะแนนจากแบบทดสอบฉบับ X และฉบับ Y จะเทียบหาคะแนนสมมูลทำได้โดยใช้คะแนน ณ ตำแหน่งเปอร์เซ็นไทล์เดียวกันของคะแนน 2 ฉบับนั้น ผลของการเทียบมาตราแสดงด้วยกราฟ ขั้นตอนการแปลงคะแนนคือ

ขั้นที่ 1 การกระจายความถี่สะสมสัมพัทธ์เป็นตารางหรือเป็นกราฟ สำหรับแบบทดสอบที่นำมาเทียบมาตราคือ

1. นำคะแนนของกลุ่มผู้สอบที่มีความสามารถทั้งเก่ง ปานกลาง และอ่อน ซึ่งสุ่มกลุ่มตัวอย่างออกเป็น 2 กลุ่ม ให้กลุ่มหนึ่งทำแบบทดสอบ X อีกกลุ่มหนึ่งทำแบบทดสอบฉบับ Y มาทำการแจกแจงคะแนน x และ y

2. คำนวณหาจุดกลางเปอร์เซ็นไทล์ของแต่ละการแจกแจง

3. อ่านและทำเครื่องหมายสำหรับค่าคะแนนของแบบทดสอบฉบับ X และ ฉบับ Y ของการแจกแจงความถี่ที่สมมูลกันลงบนกระดาษ Arithmetic Graph Paper โดยแกน นอนเป็นคะแนนดิบ แกนตั้งเป็นตำแหน่งเปอร์เซ็นไทล์ ประมาณ 30 จุด เชื่อมต่อกันเป็นเส้นกราฟ

ขั้นที่ 2 เทียบมาตราคะแนนจากรายละเอียดในข้อ 3 ขั้นตอนที่ 1 นำมาพล็อตกราฟใหม่ลงบนกระดาษ โดยแกนนอนเป็นคะแนนฉบับ Y แกนตั้งเป็นคะแนนฉบับ X ทำการปรับเส้นกราฟให้เรียบ เส้นกราฟนี้จะใช้อ่านค่า X ที่สมมูลกันกับ Y จากนั้นก็สร้างตารางคะแนนสมมูล

1.2 วิธีการเทียบคะแนนแบบเส้นตรง วิธีการนี้ตั้งอยู่ในนิยามที่ว่า คะแนนจากแบบทดสอบ 2 ฉบับใดๆ ต่างก็วัดในสิ่งเดียวกันจะพิจารณาให้เทียบกันได้ ก็ต่อเมื่อคะแนนมาตรฐานทั้งสองชุดมีค่าเท่ากัน

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

เมื่อ X คือ คะแนนจากแบบทดสอบ X

Y คือ คะแนนจากแบบทดสอบ Y

M_x คือ ค่าเฉลี่ยของแบบทดสอบ X

M_y คือ ค่าเฉลี่ยของแบบทดสอบ Y

S_x คือ ค่าเบี่ยงเบนมาตรฐานของแบบทดสอบ X

S_y คือ ค่าเบี่ยงเบนมาตรฐานของแบบทดสอบ Y

กรณีการเทียบคะแนนแบบเส้นตรงโดยการใช้แบบทดสอบร่วมหาได้จากสูตรของ ทักเกอร์ซึ่งเป็นผู้เริ่มสูตรนี้

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

เมื่อ μ_x , σ_x , μ_y , σ_y เป็นค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานของคะแนนจากแบบทดสอบ ฉบับ X และ ฉบับ Y ตามลำดับ

2. การเทียบคะแนนโดยการใช้ทฤษฎีการตอบข้อสอบ

ในการเทียบคะแนนตามทฤษฎีการตอบข้อสอบ อาศัยรูปแบบความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้ตอบข้อสอบใดๆ ในแบบทดสอบว่าขึ้นอยู่กับความสามารถของผู้สอบซึ่งการประมาณค่าได้จากโค้งลักษณะเฉพาะของข้อสอบ (Item Characteristic Curve : ICC) ของแต่ละรูปแบบที่ใช้ ซึ่งอาจจะเป็นรูปแบบหนึ่ง สอง หรือสามพารามิเตอร์ ทั้งนี้ไม่ขึ้นกับการแจกแจงของคะแนนของกลุ่มตัวอย่าง (Lord and Novick. 1968 : 359) ดังนั้นการเทียบคะแนนด้วยวิธีนี้จะสามารถประมาณค่าความสามารถแท้ของผู้สอบแต่ละคนได้ (Harris and Hoover. 1987 : 151)

นิยามการเทียบคะแนนโดยการใช้ทฤษฎีการตอบข้อสอบกล่าวว่า คะแนนแปลง Y และคะแนนดิบ X จะเทียบกันได้ก็ต่อเมื่อ ไม่ว่าผู้สอบจะสอบด้วยแบบทดสอบชุด X หรือชุด Y ก็จะไม่มีความแตกต่างกัน (Lord. 1977 : 128) ดังนั้นนิยามการเทียบคะแนน จึงเน้นความ

สำคัญของความทัดเทียมกัน(Equity) หรือความเป็นธรรมที่บุคคลจะได้รับการทดสอบอย่างเท่าเทียมกัน ถึงแม้จะสอบด้วยแบบทดสอบต่างชุดก็ตาม (Lord. 1982 : 14) ผู้สอบจะไม่ได้เสียเปรียบหรือได้เปรียบแต่อย่างใดเมื่อได้รับการทดสอบด้วยแบบทดสอบชุดใดชุดหนึ่งที่มีความยากหรือง่ายกว่ากัน (Angoff. 1971 : 563)

ต่อมา ลอร์ด (Lord. 1980 : 199) ได้ตั้งข้อกำหนดที่สำคัญ 3 ประการในการเทียบคะแนนจากแบบทดสอบสองชุด ที่วัดคุณลักษณะเดียวกัน ซึ่งเทียบคะแนนที่สมมูลกันที่ระดับความสามารถเดียวกัน ได้ดังนี้

1. ความทัดเทียมกัน หมายถึงทุกความสามารถ θ การแจกแจงความถี่แบบมีเงื่อนไขของคะแนนแปลง $X(Y)$ หรือ Y ที่ θ ที่กำหนดให้ต้องเหมือนกับการแจกแจงความถี่แบบมีเงื่อนไขของคะแนน X

2. ความไม่แปรเปลี่ยนไปตามกลุ่มตัวอย่าง หมายถึง คะแนนแปลง $X(Y)$ จะคงที่โดยไม่ขึ้นกับตัวแปรด้านอื่นๆของประชากรที่นำมาเทียบคะแนน

3. ความสมมาตร หมายถึง การเทียบคะแนนจะต้องเหมือนกันไม่ว่าการเทียบนั้นจะเป็นการเทียบจากแบบทดสอบ X หรือ แบบทดสอบ Y ไปยังแบบทดสอบ X ก็ตาม โมเดลในการเทียบคะแนนแบ่งออกได้เป็น 3 รูปแบบ ซึ่งแตกต่างกันไปตามพารามิเตอร์ที่ใช้ในแต่ละรูปแบบดังนี้

1. รูปแบบที่ใช้พารามิเตอร์ 2 ตัว (Two-Parameters Logistic Model) ฟังก์ชันที่ใช้กับเส้นโค้งลักษณะของข้อสอบ มีรูปแบบดังนี้

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

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นในการตอบข้อสอบที่ i ถูก เมื่อผู้สอบมีความสามารถ θ

a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i

b_i คือ ค่าความยากของข้อสอบข้อที่ i

D คือ ค่าปรับสเกล เท่ากับ 1.702

e คือ ค่าคงที่ เท่ากับ 2.718

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

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

เมื่อ c_i คือ ค่าการเดาของข้อสอบข้อที่ i ทั้งนี้รูปแบบนี้ไม่สามารถประมาณค่าพารามิเตอร์ที่อยู่ต่ำกว่าการเดาได้

3. รูปแบบที่ใช้พารามิเตอร์ 1 ตัว (One - parameter Logistic Model) ซึ่งราสซ์ (Rash) ได้พัฒนาขึ้นมาใช้โดยรูปแบบนี้เป็นกรณีพิเศษของรูปแบบ 2 พารามิเตอร์ที่เบอร์นบอม (Birnbaum) ได้พัฒนาขึ้น ข้อตกลงเบื้องต้นที่เพิ่มขึ้น คือค่าอำนาจจำแนกของข้อสอบ (a) เป็นค่าคงที่ของข้อสอบแต่ละข้อและค่าความยากของข้อสอบแปรเปลี่ยนไป ซึ่งสูตรในการคำนวณมีดังนี้

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

และในการเทียบคะแนนโดยการใช้ทฤษฎีการตอบข้อสอบ มีวิธีการเทียบคะแนน 3 วิธีคือ

1. วิธีการเทียบคะแนนจริง (True Score Equating)
2. วิธีการเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วม (True-Score Equating with an Anchor Test)
3. วิธีการเทียบคะแนนดิบโดยการใช้แบบทดสอบร่วม (Raw-Score Equating with an Anchor Test)

การเทียบคะแนนจริง

ในการเทียบคะแนนจริงนี้ทำได้โดยแปลงค่า θ ไปสู่คะแนนจริงที่สอดคล้องกันได้

ทั้งนี้โดยอาศัยสมการสัมพันธ์เชิงคณิตศาสตร์ระหว่างค่าความสามารถ θ และคะแนนจริงที่ได้จากแบบทดสอบแต่ละฉบับ (ฉบับ X และ ฉบับ Y) ดังนี้ (Hambleton and Swaminathan. 1985 : 212)

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

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

θ_x คือ ระดับความสามารถของผู้สอบในแบบทดสอบชุด X

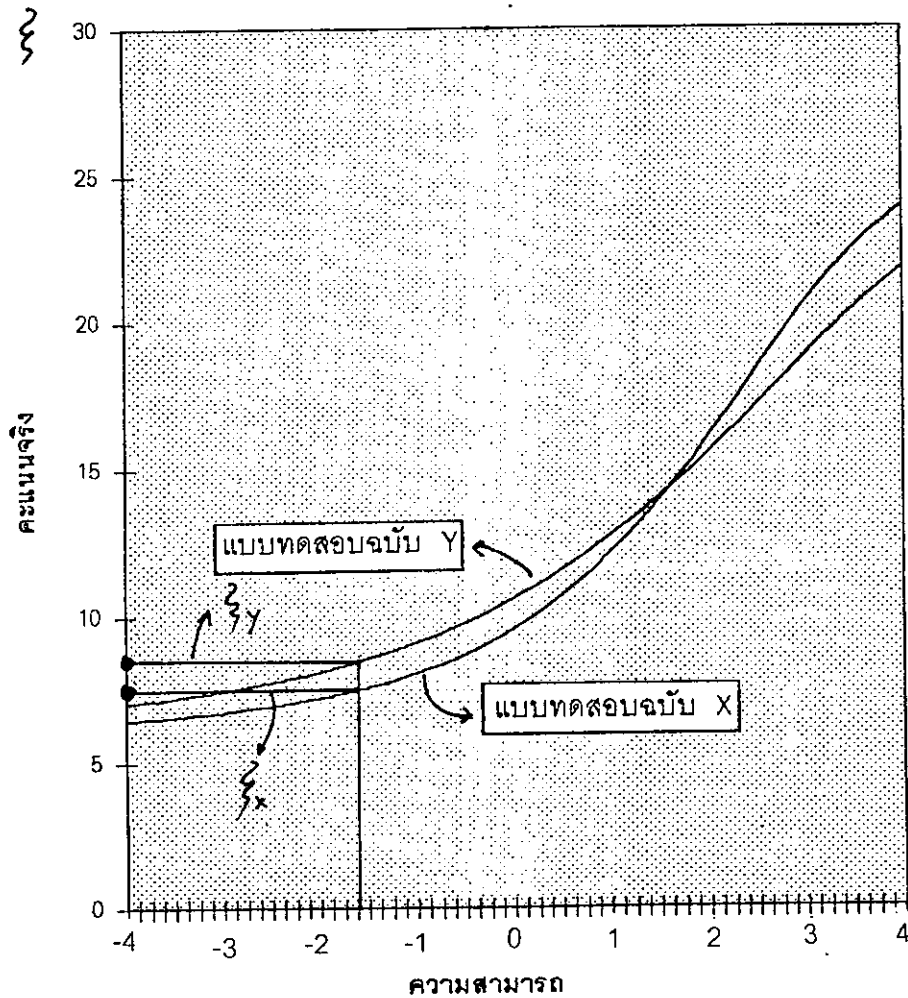
ξ_x คือ คะแนนจริงที่ได้จากแบบทดสอบชุด X

ในทำนองเดียวกัน ถ้า θ_y คือระดับความสามารถของผู้สอบในแบบทดสอบชุด Y แล้ว คะแนนจริงที่สมมูลกันในแบบทดสอบชุด Y คือ

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

เมื่อ $\theta_y = \alpha\theta_x + \beta$ คือ เส้นความสัมพันธ์ระหว่าง θ_y และ θ_x สำหรับให้ค่าคะแนนจริงที่สมมูลกันแต่ละคู่ (ξ_x, ξ_y) ดังนั้นจึงมีความเป็นไปได้ที่จะเทียบคะแนนจริงจากแบบทดสอบต่างชุด โดยคำนวณค่าคงที่ของการเทียบคะแนน α และ β หลังจากนั้นจะสามารถพล็อตกราฟเพื่อเทียบคะแนนจริงระหว่างแบบทดสอบได้ ดังภาพประกอบ 2

(Hambleton and Swaminathan. 1985 : 213)



ภาพประกอบ 2 การเทียบคะแนนจริงตามทฤษฎีการตอบข้อสอบระหว่างแบบทดสอบ
สองฉบับคือฉบับ X และ ฉบับ Y

การเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วม

ในการเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วม มีวิธีการเทียบคือถ้าแบบทดสอบ
ฉบับ X และ ฉบับ Y สอบกับผู้สอบกลุ่มเดียวกันซึ่งในการดำเนินการสอบไม่ได้มีการสอบสอง
ครั้งเนื่องจากอาจเกิดผลกระทบ (Effect) อันเนื่องมาจากความเหนื่อยล้า แต่ถ้าแบบทดสอบ

สองฉบับ สอบกับผู้สอบต่างกลุ่มกัน ซึ่งการเทียบคะแนนจะถูกเทียบจากกลุ่มตัวอย่างที่ต่างกลุ่มกัน เกี่ยวกับความแตกต่างของกลุ่มตัวอย่างนี้จะสามารถสอบและควบคุมการสอบโดยให้ผู้สอบแต่ละกลุ่ม สอบแบบทดสอบร่วม (Anchor Test) ในการวัดคุณลักษณะเดียวกันของฉบับ X และ ฉบับ Y "แบบทดสอบร่วมจะใช้ในการเทียบคะแนนกับกลุ่มผู้สอบ X และกลุ่มผู้สอบ Y ที่ไม่ได้มีความสามารถในระดับเดียวกัน หรือ "ต่างกลุ่มกัน" (Lord, 1980 : 200) แบบทดสอบร่วมอาจจะมีทั้งในฉบับ X และฉบับ Y ซึ่งเรียกว่า "แบบทดสอบร่วมภายใน" (Internal Anchor Test) และถ้าแบบทดสอบร่วมถูกแยกออกจากแบบทดสอบเทียบคะแนนแล้ว เรียกว่า "แบบทดสอบร่วมภายนอก" (External Anchor Test) และในการเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วมก็มีขบวนการเทียบคะแนนเช่นเดียวกันกับการเทียบคะแนนจริงจากแบบทดสอบฉบับ X

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

เมื่อ $P_i(\theta_x)$ คือ ค่าความน่าจะเป็นในการตอบข้อสอบข้อที่ i ได้ถูกต้อง เมื่อผู้สอบมีความสามารถ θ

θ_x คือ ระดับความสามารถของผู้สอบในแบบทดสอบ X

ξ_x คือ คะแนนจริงที่ได้จากแบบทดสอบฉบับ X และสมการคะแนนจริงจากแบบทดสอบฉบับ Y

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

เมื่อ ξ_y คือ คะแนนจริงที่ได้จากแบบทดสอบฉบับ Y

θ_y คือ ระดับความสามารถของผู้สอบในแบบทดสอบฉบับ Y

ดังนั้นคะแนนจริงที่สมมูลกันในแบบทดสอบฉบับ X และ ฉบับ Y คือ

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

เมื่อ α, β คือ ค่าคงที่ของการเทียบคะแนน หาได้โดยวิธีใช้โค้งลักษณะเฉพาะของข้อสอบซึ่งจะกำหนดความสัมพันธ์ระหว่างค่าความยากของข้อสอบ จากสมการ

$$\begin{aligned} & b_y = \alpha b_x + \beta \\ \text{เมื่อ} & b_{y_i} = \alpha b_{x_i} + \beta \\ & a_{y_i} = a_{x_i} / \alpha \\ & c_{y_i} = c_{x_i} \end{aligned}$$

ค่าคงที่ α และ β ควรเลือกให้มีความแตกต่างระหว่าง ξ_{x_a} และ ξ_{y_a} น้อย
ดังสมการ (Stocking and Lord. 1983 : 203-204)

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

และแปลงค่าพารามิเตอร์ของแบบทดสอบทั้ง 2 ฉบับ (ฉบับ X และ ฉบับ Y) ให้อยู่
ในสเกลเดียวกันโดยสมการ

1. พารามิเตอร์ของผู้สอบจาก

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

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

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

ดังนั้นจะได้สมการดังนี้

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

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

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

3. พารามิเตอร์ของค่าสัมประสิทธิ์การเดา ดังนั้น

$$a_y = a_x / \alpha$$

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

การเทียบคะแนนดิบโดยการให้แบบทดสอบร่วม

ในการเทียบคะแนนโดยวิธีนี้ (Hambleton and Swaminathan. 1985 : 214) สามารถที่จะหาได้โดยวิธีที่ง่ายตรงไปตรงมาจากการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ(θ) หรือการประมาณคะแนนจริง (ξ) มีความเหมาะสมนั้นคือ คะแนนจริง (ξ) ของสเกลคะแนนดิบ (r) คือ

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

เมื่อ U_i คือ ผลของการตอบข้อสอบข้อที่ i ของผู้สอบคนที่ $1, 2, 3, \dots, n$ ถ้าตอบถูกได้ 1 คะแนน ($U_i = 1$) และถ้าตอบผิดได้ 0 คะแนน ($U_i = 0$) และหาค่า $E(r) = \xi$ ซึ่งหาค่าได้โดย

1. ใช้ความสัมพันธ์ระหว่างคะแนนจริง ξ_x และ ξ_y ของแบบทดสอบ 2 ฉบับ (ดังกล่าวมา)

2. หาค่าความสัมพันธ์ระหว่างคะแนนดิบ r_x และ r_y แล้วเทียบคะแนนดิบ

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

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

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

การเทียบคะแนนดิบโดยการให้ทฤษฎีการตอบข้อสอบ ซึ่งจะให้การแจกแจงของคะแนนสังเกตของแบบทดสอบฉบับ X และ ฉบับ Y แล้วนำมาเทียบคะแนน(Equate) โดยใช้วิธีอิกคิวเปอร์เซ็นไทล์ ในการแจกแจงของคะแนนสังเกตตามทฤษฎี คือ $f(r/\theta)$ ของแบบทดสอบหาได้โดยสูตร

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

จากสมการดังกล่าว เป็นสมการที่ใช้ในการคำนวณในแต่ละความสามารถ θ จากสูตรจะแสดงวิธีการคำนวณได้ดังนี้ ตัวอย่างเช่น ถ้ามีข้อสอบจำนวน 3 ข้อ ($n = 3$) แล้ว ดังนั้นจะได้สมการขวามือและทางซ้ายที่เท่ากันคือสมการทางขวามือ

$$\begin{aligned} \prod_{i=1}^3 (Q_i + tP_i) &= (Q_1 + tP_1)(Q_2 + tP_2)(Q_3 + tP_3) \\ &= Q_1Q_2Q_3 + t(Q_1Q_2P_3 + Q_1P_2Q_3 + P_1Q_2Q_3) + t^2(Q_1P_2P_3 + P_1Q_2P_3 + P_1P_2Q_3) + t^3P_1P_2P_3 \end{aligned}$$

สมการทางซ้ายมือ

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

จากที่กล่าวมาสามารถสรุปได้ดังตาราง 1 แสดงการแจกแจงคะแนนดิบของความสามารถ θ สำหรับข้อสอบจำนวน 3 ข้อ ดังตาราง 1

ตาราง 1 การแจกแจงคะแนนดิบของผู้สอบที่มีความสามารถ θ สำหรับข้อสอบ 3 ข้อ

คะแนนดิบ (Raw Score)	การแจกแจงความถี่สัมพัทธ์ของคะแนนแบบมีเงื่อนไข (Conditional Relative Frequency : $f(r/\theta)$)
0	$Q_1Q_2Q_3$
1	$Q_1Q_2P_3 + Q_1P_2Q_3 + P_1Q_2Q_3$
2	$Q_1P_2P_3 + P_1Q_2P_3 + P_1P_2Q_3$
3	$P_1P_2P_3$

สัมประสิทธิ์ t , $f(r/\theta)$ คือผลรวมในรูปของ "c" ซึ่งการคำนวณแบบนี้ก็คือการคำนวณโดยใช้ค่า Maximum likelihood Estimation นั่นเอง เมื่อแทนค่า θ ลงไปก็สามารถที่จะคำนวณค่า $f(r/\theta)$ ได้ เช่นถ้ามีผู้สอบที่มีความสามารถ $\theta_1, \theta_2, \dots, \theta_a, \dots, \theta_n$ และคำนวณหาค่า Marginal Frequency Distribution $f(r)$ หากจากฟังก์ชัน

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

ขั้นตอนการเทียบคะแนนทำได้ดังนี้

1. แปลงค่าพารามิเตอร์ของข้อสอบ และพารามิเตอร์ของผู้สอบให้อยู่ในสเกลเดียวกัน ของแบบทดสอบทั้งสองฉบับ
2. หาค่า Conditional Frequency Distribution $f_x(r/\theta)$ สำหรับผู้สอบแต่ละคน ในกลุ่มเทียบคะแนนโดยการประมาณค่าความสามารถของผู้สอบ(θ) และพารามิเตอร์ของแบบทดสอบฉบับ X (a_i, b_i, c_i)
3. หาค่า Marginal Frequency Distribution $f_x(r)$ จากสมการที่กำหนดให้ดังกล่าว
4. ทำเช่นเดียวกับขั้นตอนที่ 1 และขั้นตอนที่ 2 ของแบบทดสอบฉบับ Y
5. เทียบคะแนนโดยใช้วิธีอิควิปเปอร์เซ็นไทล์

ตอนที่ 3. งานวิจัยที่เกี่ยวข้องกับการเทียบคะแนน

งานวิจัยที่เกี่ยวข้องกับการเทียบคะแนนตามทฤษฎีการตอบข้อสอบได้แบ่งแยกศึกษาออกเป็นวิธีการต่างๆอีกมากมาย ในที่นี้พอจะสรุปได้ดังรายละเอียดต่อไปนี้

งานวิจัยต่างประเทศ

ลอร์ด (Lord. 1977 : 117-138) มาร์โค(Marco. 1977 : 139-160) ได้เปรียบเทียบวิธีการเทียบมาตรฐานระหว่างวิธีแบบดั้งเดิม (Conventional) และวิธีการใช้ทฤษฎีการตอบข้อสอบ ผลปรากฏว่าผลที่ได้แตกต่างกัน ซึ่งสอดคล้องกับการศึกษาของ มาร์โค ปีเตอร์เซน และสตีเวอร์ท (Kolen. 1981 : Citing Marco, Petersen and Stewart) ที่ได้ศึกษา

เปรียบเทียบวิธีการเทียบมาตรฐานเชิงเส้นตรง วิธีอิกวิเปอร์เซ็นไทล์ และวิธีการใช้ค้ำลักษณะเฉพาะของข้อสอบ โดยใช้แบบทดสอบ SAT แบ่งแบบทดสอบโดยใช้แบบสอบร่วม ผลการศึกษาพบว่า แบบทดสอบที่มีค่าความยากแตกต่างกัน วิธีการเทียบคะแนนโดยใช้ค้ำลักษณะเฉพาะของข้อสอบให้ผลดีที่สุด ส่วนวิธีเทียบแบบเส้นตรงได้ผลน้อยที่สุด

โคเลน และวิทนีย์ (Kolen and Whitney, 1982 : 279-294) ได้เปรียบเทียบความเพียงพอของวิธีการเทียบคะแนน 4 วิธีคือ วิธีอิกวิเปอร์เซ็นไทล์ วิธีเชิงเส้นตรง วิธีการใช้ทฤษฎีการตอบข้อสอบชนิดหนึ่งพารามิเตอร์โลจิสติกโมเดล หรือราสซิมเดลและชนิดสามพารามิเตอร์โลจิสติกโมเดล โดยใช้แบบทดสอบ GED (Test of General Educational Development) ซึ่งเป็นแบบทดสอบมาตรฐานวัดผลสัมฤทธิ์ทางการเรียนจำนวน 5 วิชา แต่ละวิชามีเนื้อหาแตกต่างกันดังนี้ ทักษะการเขียน (80 ข้อ) การศึกษาสังคม (60 ข้อ) วิทยาศาสตร์ (60 ข้อ) ทักษะการอ่าน (40 ข้อ) และคณิตศาสตร์ (50 ข้อ) ซึ่งใช้สำหรับตัดสินให้ประกาศนียบัตรแก่นักเรียนระดับอุดมศึกษาทั่วประเทศที่มีความรู้เทียบเท่าระดับที่แสดงว่าได้เรียนสำเร็จแล้ว แบบทดสอบแต่ละวิชามี 12 ชุด ดำเนินการสอบผู้สอบกลุ่มละ 2 ชุด และทุกกลุ่มต้องสอบแบบสอบชุดที่ 12 ซึ่งแยกออกมาเป็นแบบทดสอบร่วม การจัดลำดับแบบทดสอบให้กับแต่ละกลุ่มใช้วิธีกลุ่มสมดุลย์ (Counterbalance) ทำการเทียบคะแนนแต่ละคู่ของแบบทดสอบในแนวนอนด้วยวิธีการเทียบ 4 วิธีดังกล่าว กลุ่มตัวอย่างสุ่มจากการสอบในปี 1980 ซึ่งมีจำนวนมากกว่า 800,000 คน โดยสุ่มมาขนาดเล็ก กลุ่มละประมาณ 200 คน คิดเป็นสัดส่วนโรงเรียนละ 22 คนด้วยวิธีการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) ตามประเภทของโรงเรียน สภาพทางภูมิศาสตร์และสถานภาพทางสังคม กลุ่มทวนผลก็ใช้วิธีสุ่มเช่นเดียวกัน โดยสุ่มมาจากผู้สอบแต่ละวิชา และแต่ละชุดให้ได้ผู้สอบ 20 คน (ประมาณร้อยละ 10 ของผู้เข้าสอบทั้งหมด) การเทียบคะแนนเชิงเส้นตรงและอิกวิเปอร์เซ็นไทล์ กระทำตามวิธีของแองกอฟฟ์ ส่วนวิธีการใช้ทฤษฎีการตอบข้อสอบประมาณค่าพารามิเตอร์โดยใช้โปรแกรมโลจิสติกเฉพาะแบบทดสอบร่วมก่อน จากนั้นใช้ค่าพารามิเตอร์ความสามารถของผู้สอบที่กำหนดให้คงที่ในการประมาณค่าพารามิเตอร์จากแบบทดสอบอีก 11 ชุด แล้วสร้างตารางคะแนนสมมูลระหว่างแบบสอบทั้งหมดโดยใช้ประมาณค่าคะแนนจริงส่วนคะแนนซึ่งอยู่ต่ำกว่าระดับการเดาตามรูปแบบสามพารามิเตอร์นั้นอาศัยวิธีการเชิงเส้นตรงมาประมาณค่าเพิ่มเติมเกณฑ์ในการตรวจสอบความเพียงพอของวิธีการเทียบคะแนนใช้ค่าสถิติจากกลุ่มทวนผลซึ่งผู้สอบแต่ละคนจะต้องสอบแบบสอบทั้งสองชุดที่นำมาเทียบ ดัชนีเปรียบเทียบเปอร์เซ็นไทล์ซึ่งเป็นค่าเฉลี่ยกำลังสองของความแตกต่างระหว่าง

คะแนน X ที่ได้จากแบบสอบร่วมกับคะแนน Y ที่แปลงไปสู่สกุลของคะแนน X โดยใช้วิธีเทียบแต่ละวิธี และทดสอบความมีนัยสำคัญของค่าอันดับจากการแปลงคะแนนตามปริมาณของค่าดัชนีด้วยการทดสอบฟรายด์แมน

ผลการวิจัยพบว่า วิธีการเทียบคะแนนอิกวิเปอร์เซ็นไทล์ให้ผลการเทียบที่คงเส้นคงวาน้อยกว่าการเทียบคะแนนเชิงเส้นตรง ส่วนการเทียบคะแนนตามทฤษฎีการตอบข้อสอบชนิดหนึ่งพารามิเตอร์หรือราสนิมเดลให้ผลคงที่เหมือนกับวิธีเทียบเชิงเส้นตรง ขณะที่วิธีเทียบตามทฤษฎีการตอบข้อสอบชนิดสามพารามิเตอร์ และวิธีอิกวิเปอร์เซ็นไทล์ให้ผลน้อยกว่า ดังนั้นการศึกษานี้จึงให้ผลการเทียบแตกต่างจากการศึกษาอื่น ได้แก่การศึกษาของโคเลน (Kolen. 1981 : 1-11) ที่พบว่าวิธีการเทียบคะแนนแบบอิกวิเปอร์เซ็นไทล์ให้ผลมากที่สุดเมื่อเปรียบเทียบกับวิธีเชิงเส้นตรงและทฤษฎีการตอบข้อสอบชนิดหนึ่งพารามิเตอร์ ทั้งนี้การศึกษาของโคเลนใช้กลุ่มตัวอย่างขนาดใหญ่ซึ่งแตกต่างจากการศึกษารั้งนี้ที่ใช้กลุ่มตัวอย่างขนาดเล็ก นอกจากนี้แบบทดสอบที่ใช้เทียบในการศึกษารั้งนี้มีความยากใกล้เคียงกันน้อยกว่า ส่วนผลการเปรียบเทียบของคนอื่น ๆ (Kolen and Whitney. 1982 : 291 ; Petersen, Cook and Stocking. 1981) ซึ่งต่างพบว่าเมื่อแบบทดสอบที่เทียบโดยใช้แบบทดสอบร่วมกระทำกับกลุ่มที่ไม่ได้มาจากการสุ่มซึ่งมีขนาดใหญ่ นั้นวิธีการใช้ทฤษฎีการตอบข้อสอบชนิดสามพารามิเตอร์ จะให้ผลการเทียบที่มีความเพียงพอมากที่สุด อย่างไรก็ตามวิธีเชิงเส้นตรงเพียงพอมากที่สุดหากแบบทดสอบที่เทียบมีความคล้ายคลึงกันมาก ผลการศึกษาที่แตกต่างกันเหล่านี้ โคเลนและวิทนีได้สรุปว่า วิธีการเทียบที่มีความเพียงพอขึ้นกับองค์ประกอบหลายประการ ได้แก่คุณลักษณะต่างๆของแบบทดสอบ รูปแบบที่ใช้ในการเทียบ และลักษณะของกลุ่มตัวอย่างที่ใช้เทียบ เป็นต้น

โคเลน (Kolen. 1981 : 1-11) ได้เปรียบเทียบผลของการเทียบคะแนนระหว่างรูปแบบดั้งเดิมสองวิธี กับรูปแบบ IRT เจ็ดวิธี โดยใช้ข้อมูลจากผลการสอบนักเรียนในรัฐไอโอวา ปี 1978 ตามโครงการชื่อ The Iowa Test of Educational Development แบบทดสอบที่ใช้คือ แบบทดสอบที่พิมพ์ใหม่ครั้งที่ 7 เทียบไปสู่แบบทดสอบชุดเก่าที่พิมพ์ครั้งที่ 6 ซึ่งมีหนึ่งชุดคือ X_6 ใช้สอบกับนักเรียนทั้งสองระดับ คือ ระดับ 1 และ ระดับ 2 ซึ่งระดับ 1 เป็นนักเรียนเกรด 9 และ 10 ส่วนระดับ 2 เป็นนักเรียนเกรด 11 และ 12 ส่วนแบบทดสอบที่พิมพ์ครั้งที่ 7 มี 2 ชุดที่คู่ขนานกันคือ X_7 และ Y_7 แต่ละชุดแบ่งออกเป็น 2 ระดับ ซึ่งมีความยากต่างกันเพื่อแยกสอบนักเรียนแต่ละคน มีแบบแผนในการรวบรวมข้อมูลดังนี้ นักเรียน

ระดับ 1 สอบแบบทดสอบ X_6 , X_7 ระดับ 1 และ Y_7 ระดับ 2 ทั้งนี้นักเรียนแต่ละคนจะได้รับแบบทดสอบเพียง 1 ชุด คือ ชุด X_6 , X_7 หรือ Y_7 โดยการสุ่ม ดังนั้นจำนวนนักเรียนที่สอบแบบทดสอบแต่ละชุดจะมี 1 ใน 3 ของจำนวนนักเรียนในแต่ละระดับ ส่วนแบบทดสอบ X_6 จะมีนักเรียนจำนวน 1 ใน 3 ของจำนวนนักเรียนรวมสองระดับ สำหรับเกณฑ์ในการเปรียบเทียบผลจะใช้ค่าดัชนีจากกลุ่มทานผลซึ่งเป็นนักเรียนทุก ๆ คนที่ 3 จากการสอบแบบทดสอบแต่ละชุดและ แต่ละระดับการเทียบคะแนนแบบดั้งเดิมใช้วิธีการเชิงเส้นตรงและวิธีอิคิวเปอร์เซ็นไทล์ซึ่งจัดกระทำตามวิธีของแองกอฟฟ์ ส่วนรูปแบบของทฤษฎีการตอบข้อสอบใช้ หนึ่ง สอง และสามพารามิเตอร์โลจิสติกโมเดล โดยแต่ละโมเดลเทียบ 2 แบบคือ เทียบโดยการประมาณค่าคะแนนจริง วิธีเทียบคะแนนสังเกต และวิธีราสซิมโมเดล ดัชนีการทานผลโคเลนใช้วิธีคำนวณจากค่าเฉลี่ยกำลังสองของความแตกต่าง (สำหรับผู้สอบแบบทดสอบ X_6 ในการทานผล) ระหว่างคะแนนจากแบบทดสอบ X_6 กับคะแนนจากแบบทดสอบ X_7 และ Y_7 ที่แปลงแล้ว ณ ตำแหน่งเปอร์เซ็นไทล์เดียวกัน ทั้งนี้โดยการใช้วิธีการเทียบคะแนนแบบต่างๆทั้งเก่าวิธี หากค่าดัชนีจากกลุ่มทานผลที่ได้จากวิธีการเทียบใดมีค่าน้อยกว่าแสดงว่าวิธีนั้นมีความสอดคล้องระหว่างการแจกแจงคะแนนของแบบทดสอบชุดเก่ากับคะแนนของแบบทดสอบชุดใหม่ที่แปลงไปยังสเกลของชุดเก่ามากกว่า นั่นคือ วิธีนั้นมีความคงเส้นคงวามากกว่า ทั้งนี้โคเลนทดสอบความแตกต่างของดัชนีที่ได้จากการเทียบคะแนนแบบต่างๆ เพื่อเปรียบเทียบผลของแต่ละวิธี โดยใช้สถิติทดสอบฟรายด์แมน (Friedman Test) ผลการวิจัยพบว่า วิธีการเทียบคะแนนแบบต่างๆให้ผลแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และ .05 เมื่อใช้ แบบทดสอบระดับ 1 และ 2 ตามลำดับ ถ้าพิจารณาแบบทดสอบระดับที่ 1 พบว่า วิธีที่ให้ผลคงเส้นคงวามากที่สุดคือ วิธีการตามทฤษฎีการตอบข้อสอบแบบสามพารามิเตอร์ที่ประมาณค่าคะแนนสังเกตส่วนวิธีการที่ให้ผลคงเส้นคงวน้อยที่สุดคือ วิธีการเชิงเส้นตรง ถัดมาคือวิธีการใช้ทฤษฎีการตอบข้อสอบชนิดหนึ่งพารามิเตอร์ สำหรับแบบสอบระดับที่ 2 พบว่าวิธีการที่ให้ผลคงเส้นคงวามากที่สุด คือวิธีการใช้ทฤษฎีการตอบข้อสอบแบบสามพารามิเตอร์ที่ประมาณค่าคะแนนจริง วิธีของราสซิมให้ผลคงเส้นคงวามากกว่าวิธีตามทฤษฎีการตอบข้อสอบชนิดหนึ่งพารามิเตอร์ที่ประมาณค่าคะแนนจริง ซึ่งโคเลนอภิปรายโดยส่วนรวมว่า วิธีตามทฤษฎีการตอบข้อสอบชนิดหนึ่งพารามิเตอร์ให้ผลการเทียบที่ไม่เพียงพอ ทั้งนี้อาจเนื่องมาจากการเดาของผู้สอบมากยกเว้นเมื่อจัดการเดาของผู้สอบโดยใช้ วิธีของไรท์และสโตน(Wright and Stone, 1979) ซึ่งตัดข้อสอบที่พบว่าไม่เหมาะกับราสซิมโมเดลออก

จากผลการสอบผลการวิจัยนี้สนับสนุนงานวิจัยของสไลด์และลินน์ (Slinde and Linn. 1978 : 23-35 ; Slinde and Linn. 1979 : 437 - 452) ที่พบว่า การเทียบคะแนนในแนวตั้งโดยใช้แบบทดสอบร่วมที่มีความยากต่างกันสอบกับนักเรียนต่างระดับชั้นตามวิธีของราสซิมเดลให้ผลที่ไม่เพียงพอ

ลอร์ด (Lord. 1980 : 203) ได้ศึกษาทฤษฎีการเทียบคะแนนดิบของแบบทดสอบสองชุด ด้วยวิธีการเทียบคะแนนจริงและคะแนนสังเกต พบว่ามีข้อได้เปรียบและเสียเปรียบวิธีการเทียบคะแนนจริงให้ความหมายของคะแนนสมมูลที่อยู่เหนือค่าเฉลี่ยการเดา ส่วนวิธีการประมาณค่าคะแนนสังเกต ก็มีจุดอ่อนที่เป็นเพียงการเทียบมาตรงๆอย่างประมาณเท่านั้น ต่อข้อคำถามที่ว่า วิธีใดจะให้ผลการเทียบที่ดีกว่านั้น ลอร์ดกล่าวว่า ทั้งสองวิธีมีความสอดคล้องกันมาก แต่การสรุปอ้างอิงทั่วไปต้องทำอย่างพิถีพิถัน

ลอร์ด และวิงเกอร์สกี (Lord and Wingersky. 1984 : 453-461) ได้เปรียบเทียบการเทียบคะแนนระหว่างรูปแบบการเทียบคะแนนจริงโดยการใช้ทฤษฎีการตอบข้อสอบ และการเทียบคะแนนดิบโดยใช้วิธีอิกวิบเปอร์เซ็นไทล์ของการใช้ทฤษฎีการตอบข้อสอบจากแบบทดสอบ SAT ด้านภาษา จำนวน 6 ชุด เทียบเข้าสู่สเกลเดียวกันแบบสายใช้โดยใช้แบบทดสอบร่วม ทั้งนี้แบบทดสอบชุดแรกและชุดสุดท้ายเป็นชุดเดียวกัน ขั้นตอนการเทียบคะแนนเริ่มจาก แบบทดสอบ V_4 เทียบไปสู่แบบสอบ X_2 โดยใช้แบบสอบร่วม fe ส่วนแบบทดสอบ X_2 เทียบไปสู่ Y_3 โดยใช้แบบสอบร่วม fm ซึ่งจะใช้ผลการเทียบจาก V_4 ไป Y_3 กระทำการเทียบคะแนนในลักษณะเดียวกันไปยังแบบทดสอบอื่น สุดท้ายจะได้ผลการเทียบ V_4 กับ V_4 ซึ่งเป็นการเทียบแบบทดสอบกับตัวเอง และใช้เป็นเกณฑ์ในการเปรียบเทียบความคงเส้นคงวาของวิธีการเทียบคะแนนทั้งสองวิธี ทั้งนี้เกณฑ์นี้แตกต่างจากที่ โคลเนน (Kolen. 1981 1-11) ได้เคยศึกษาความคงเส้นคงวาของวิธีการเทียบคะแนนแบบต่างๆ โคลเนนใช้เกณฑ์จากกลุ่มทานผลซึ่งเป็นกลุ่มอิสระอีกกลุ่มหนึ่ง โดยผู้วิจัยกล่าวไว้ว่าเกณฑ์ที่โคลเนนใช้ไม่ใช่เกณฑ์ที่มีคุณสมบัติสำหรับการเลือกวิธีการเทียบคะแนนที่ดีที่สุด เพราะวิธีการเทียบคะแนนที่ไม่ถูกต้องอาจให้ผลการเทียบคงเส้นคงวามากกว่าวิธีการที่ถูกต้อง เกณฑ์การเทียบแบบทดสอบกับตัวเองที่ใช้ในการศึกษานี้อาจไม่เที่ยงตรงในเชิงทฤษฎี และไม่น่าพอใจในการปฏิบัติสำหรับการเทียบแบบทดสอบสองฉบับที่แตกต่างกัน แต่เป็นเกณฑ์ที่มีประโยชน์ต่อการเปลี่ยนวิธีการเทียบคะแนนใหม่หากพบว่าคะแนนที่แปลงแล้วในขั้นสุดท้ายไม่เป็นอย่างเดียวกัน

กับคะแนนเดิมของตัวเอง ผลการวิจัยยังไม่มี การตัดสินใจทางทฤษฎีที่ชัดเจนในการเลือกใช้วิธีการเทียบคะแนนแบบไฮอาร์ทีที่ใช้คะแนนจริงกับคะแนนสังเกตโดยวิธีอควิบเปอร์เซ็นต์ไคล์ เนื่องจากทั้งสองวิธีมีความสอดคล้องกัน

งานวิจัยในประเทศ

ชูชีพ พงษ์สมบุรณ์ (2528 : 112-118) ได้เปรียบเทียบการเทียบมาตรฐานระหว่างรูปแบบที่ใช้แบบทดสอบร่วมกับรูปแบบที่ใช้ผู้สอบร่วมโดยใช้วิธีการเทียบ 3 วิธี คือ การเทียบเชิงเส้นตรง การเทียบโดยใช้เปอร์เซ็นต์ไคล์และการเทียบโดยใช้โด่งลักษณะข้อสอบเพื่อต้องการเปรียบเทียบประสิทธิภาพและความคงที่ของการเทียบมาตรฐานทั้ง 3 วิธี ใช้กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2526 จำนวน 3,721 คน ที่สอบจากแบบทดสอบของโครงการตรวจสอบคุณภาพการศึกษาทั่วประเทศในวิชาคณิตศาสตร์ ของกรมวิชาการ กระทรวงศึกษาธิการ ในข้อสอบจำนวน 60 ข้อ แบ่งออกเป็น 2 ฉบับ ฉบับละ 38 ข้อ ซึ่งรวมด้วยแบบทดสอบร่วม 20 ข้อ ด้วยผลการศึกษาพบว่า ผลของการเปรียบเทียบประสิทธิภาพของการเทียบมาตรฐานระหว่างรูปแบบทดสอบร่วมกับรูปแบบที่ใช้ผู้สอบร่วมในแต่ละวิธีไม่แตกต่างกัน การเปรียบเทียบความคงที่ของการเทียบมาตรฐานที่ใช้ผู้สอบร่วม พบว่าการเทียบโดยใช้โด่งลักษณะข้อสอบคงที่พอๆ กับการเทียบโดยใช้เปอร์เซ็นต์ไคล์ และการเทียบเชิงเส้นตรงมีความคงที่ต่ำที่สุด ส่วนการเปรียบเทียบความคงที่ของการเทียบมาตรฐานที่ใช้แบบทดสอบร่วม ได้ผลเช่นเดียวกับแบบใช้ผู้สอบร่วม

ภาวณี ศรีสุขวัฒนานันท์ (2529 : 155-170) ศึกษาเปรียบเทียบผลของการใช้รูปแบบการเทียบมาตรฐานที่ต่างกัน 3 รูปแบบ คือ รูปแบบอควิบเปอร์เซ็นต์ไคล์ รูปแบบเชิงเส้นตรง และรูปแบบอิงทฤษฎีการตอบข้อสอบแบบโลจิสสามพารามิเตอร์ ด้วยการออกแบบเทียบมาตรฐานใช้แบบทดสอบร่วมภายในที่มีความยาวต่างกันสามขนาดคิดเป็นร้อยละคือร้อยละ 60 (21 ข้อ) ร้อยละ 40 (14 ข้อ) และร้อยละ 20 (7 ข้อ) โดยใช้ข้อสอบทั้งหมด 35 ข้อ การประเมินผลของการเทียบมาตรฐาน โดยการวิเคราะห์ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐานใช้เป็นเกณฑ์ภายในเพื่อเปรียบเทียบความแม่นยำระหว่างวิธีการเทียบภายในรูปแบบเดียวกัน และวิเคราะห์กลุ่มทานผลใช้เป็นเกณฑ์ภายนอก เพื่อเปรียบเทียบความเพียงพอของการเทียบมาตรฐานโดยตรงกับผลสุดท้ายของการเทียบมาตรฐาน โดยศึกษากับประชากรสองกรณี คือการสอบคัดเลือกจากประชากร 15,875 คน และแบบทดสอบ

วัดผลสัมฤทธิ์ จากประชากร 12,383 คน ในแต่ละกรณีใช้กลุ่มตัวอย่าง กลุ่มละ 1,500 คน แบ่งแบบทดสอบกรณีละ 2 ชุดโดยใช้ชุดละ 35 ข้อ มีแบบทดสอบรวมขนาด 21 14 และ 7 ข้อ โดยยึดหลักแบบทดสอบรวมค่าความเชื่อมั่นเท่าเทียมกัน ในแต่ละคู่อใช้รูปแบบการเทียบมาตรา 3 รูปแบบ สร้างตารางแปลงคะแนนสมมูลจากคะแนนแบบทดสอบชุดที่ 1 ไปยังแบบทดสอบชุดที่ 2 แล้ววิเคราะห์ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรา ค่าดัชนีประสิทธิภาพสัมพัทธ์ (RE) การวิเคราะห์กลุ่มทานผล โดยคำนวณค่าดัชนีเปรียบเทียบความแตกต่าง(c) ผลการศึกษาพบว่า การเทียบมาตราทั้ง 9 รูปแบบได้ผลเป็นที่ยอมรับได้ ความยาวแบบทดสอบร่วมมีผลต่อความแม่นยำ และความเพียงพอของการเทียบมาตรา โดยที่แบบทดสอบร่วมมีความยาวมากและแบบทดสอบร่วมทั้งสองฉบับ มีค่าความเชื่อมั่นสูงและใกล้เคียงกัน จะทำให้การเทียบมาตราแม่นยำมากขึ้น และในแต่ละสถานการณ์การสอบ ให้ผลการเทียบมาตราในรูปแบบต่างกัน คือ กรณีการสอบคัดเลือก รูปแบบที่ให้ผลดีตามลำดับ คือ อิกวิเปอร์เซ็นไทล์ รูปแบบอิงทฤษฎีการตอบข้อสอบและรูปแบบอิงเส้นตรง ส่วนกรณีวัดผลสัมฤทธิ์ รูปแบบที่ให้ผลดีตามลำดับคือ รูปแบบเชิงเส้นตรง รูปแบบอิกวิเปอร์เซ็นไทล์ และ รูปแบบอิงทฤษฎีการตอบข้อสอบ

เววดี อินทะสระระ (2530 : 64-67) ศึกษาเปรียบเทียบความคลาดเคลื่อนมาตรฐาน และความเพียงพอของการเทียบคะแนนตามทฤษฎีการตอบข้อสอบแบบสามพารามิเตอร์กับวิธีการวิเคราะห์องค์ประกอบ กลุ่มตัวอย่างในการศึกษาเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2529 จำนวน 2823 คน แบบทดสอบที่ใช้เป็นแบบทดสอบวัดผลสัมฤทธิ์ วิชาคณิตศาสตร์ 2 ฉบับๆละ 45 ข้อ โดยมีแบบสอบภายในฉบับละ 15 ข้อ ผลการวิจัยพบว่าความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนโดยวิธีสามพารามิเตอร์น้อยกว่าวิธีการใช้เทคนิคการวิเคราะห์องค์ประกอบ ส่วนความเพียงพอของการเทียบคะแนน โดยวิธีสามพารามิเตอร์อยู่ในระดับที่น่าพอใจส่วนวิธีการวิเคราะห์องค์ประกอบอยู่ในระดับปานกลาง

วรรณดี แสงประทีปทอง (2538 : 113-123) ศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย ของนักเรียนชั้นมัธยมต้น โดยใช้วิธีการเทียบคะแนนตามแนวตั้งตามทฤษฎีการตอบข้อสอบ แบบสามพารามิเตอร์โดยศึกษาความก้าวหน้าทั้งในกลุ่มรวมและในกลุ่มย่อยตามระดับความสามารถของนักเรียน กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมต้นของโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษาในจังหวัดสิงห์บุรี ภาคเรียนที่ 1 ปีการศึกษา 2537 กลุ่มตัวอย่างแบ่งออกเป็น 2 กลุ่ม คือ กลุ่มตัวอย่างที่ใช้ในการเทียบคะแนนเป็นนักเรียน

ชั้น ม. 1 จำนวน 1,245 คน ม. 2 จำนวน 1,280 คน และม. 3 จำนวน 1,171 คน กลุ่มทานผลคือ นักเรียนชั้น ม. 1 จำนวน 373 ม. 2 จำนวน 384 และ ม. 3 จำนวน 352 คน เครื่องมือที่ใช้ในการวิจัยคือแบบทดสอบเทียบคะแนน 3 ฉบับ คือ ท 101 ท203 และ ท305 และแบบทดสอบร่วม ซึ่งเป็นแบบทดสอบวัดความสามารถทั่วไป 1 ฉบับ เก็บรวบรวมข้อมูลจากการสอบจริงปลายภาคเรียนที่ 1 ปีการศึกษา 2537 พร้อมกันทุกระดับชั้น

การวิเคราะห์ข้อมูลใช้แบบจำลองสามพารามิเตอร์ทฤษฎีการตอบข้อสอบเพื่อเปรียบเทียบการเทียบคะแนนต่างระดับชั้น และเปรียบเทียบผลสัมฤทธิ์ทางการเรียนแต่ละระดับชั้น ภายหลังจากการเทียบคะแนน โดยเปรียบเทียบในกลุ่มรวมและกลุ่มย่อยตามระดับความสามารถ ผลการศึกษาพบว่า ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม. 3 สูงกว่า ชั้น ม. 1 และ ม. 2 ตามลำดับ แต่ผลสัมฤทธิ์ทางการเรียนของชั้น ม.1 และ ม.2 ไม่แตกต่างกัน ในกลุ่มนักเรียนที่มีความสามารถสูง ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่าชั้น ม.1 แต่ผลสัมฤทธิ์ทางการเรียนชั้น ม.3 และ ชั้น ม.2 ไม่แตกต่างกัน และผลสัมฤทธิ์ทางการเรียนชั้น ม.2 และ ชั้น ม.1 ไม่แตกต่างกัน ในกลุ่มที่มีความสามารถปานกลาง ผลสัมฤทธิ์ทางการเรียนชั้น ม.1 ม.2 และ ม.3 ก้าวหน้าขึ้นตามลำดับ โดยที่ผลสัมฤทธิ์ทางการเรียนชั้น ม.3 สูงกว่า ชั้น ม.2 และผลสัมฤทธิ์ทางการเรียนชั้น ม.2 สูงกว่าชั้น ม.1 ส่วนในกลุ่มที่มีความสามารถต่ำ ผลสัมฤทธิ์ทางการเรียนของชั้น ม.3 สูงกว่าชั้น ม.2 และ ม. 1 ตามลำดับ แต่ผลสัมฤทธิ์ทางการเรียนของชั้นม.1 และ ม.2 ไม่แตกต่างกัน

จากเอกสารและงานวิจัยที่เกี่ยวข้องดังกล่าว แสดงให้เห็นว่า การเทียบคะแนนตามทฤษฎีการตอบข้อสอบเป็นวิธีที่ให้ความคงที่และความเพียงพอของการเทียบคะแนนดีกว่าวิธีการเทียบคะแนนตามทฤษฎีคลาสสิคอลล ดังนั้นผู้วิจัยมีความสนใจที่จะศึกษาว่า วิธีการเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วม และวิธีการเทียบคะแนนดิบโดยการใช้แบบทดสอบร่วมตามทฤษฎีการตอบข้อสอบ ว่าวิธีการใดทั้ง 2 วิธีที่ให้ผลการเทียบคะแนนเหมาะสมและดีที่สุด เพื่อประโยชน์ในการเลือกใช้ให้สมบูรณ์ยิ่งขึ้นไป

สมมติฐานของการศึกษาค้นคว้า

1. คะแนนที่เทียบได้ของแบบทดสอบวิชาการพยาบาลสูติศาสตร์ ระหว่างวิธีวิเคราะห์ต่างกันมีค่าแตกต่างกัน
2. ความคลาดเคลื่อนของการเทียบคะแนนของแบบทดสอบวิชาการพยาบาลสูติศาสตร์ระหว่างวิธีวิเคราะห์ต่างกันมีค่าแตกต่างกัน

3. ความเพียงพอของการเทียบคะแนนของแบบทดสอบวิชาการพยาบาล
สูติศาสตร์ระหว่างวิธีวิเคราะห์ต่างกันมีค่าแตกต่างกัน

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักศึกษาพยาบาลศาสตร์(ระดับต้น) ชั้นปีที่ 2 ปีการศึกษา 2538 ของวิทยาลัยพยาบาลในสังกัดสถาบันพัฒนากำลังคนด้านสาธารณสุข กระทรวงสาธารณสุข ในเขตภาคเหนือ ภาคกลางและภาคตะวันออก จำนวนนักศึกษา 2,050 คน จำนวน 20 วิทยาลัย

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักศึกษาพยาบาลศาสตร์(ระดับต้น) ชั้นปีที่ 2 ปีการศึกษา 2538 ของวิทยาลัยพยาบาลในสังกัดสถาบันพัฒนากำลังคนด้านสาธารณสุข กระทรวงสาธารณสุข ในเขตภาคเหนือ ภาคกลางและภาคตะวันออก จำนวนนักศึกษา 980 คน จำนวน 10 วิทยาลัย ซึ่งได้มาโดยวิธีการสุ่มอย่างง่าย (Simple Random Sampling) โดยใช้ร้อยละ 50 ของจำนวนวิทยาลัยทั้งหมด และมีวิทยาลัยเป็นหน่วยการสุ่ม (Sampling Unit) ดังแสดงในตาราง 2

ตาราง 2 รายชื่อวิทยาลัยและจำนวนนักศึกษาในกลุ่มตัวอย่างที่ใช้ในการศึกษาครั้งนี้

วิทยาลัยพยาบาลบรมราชชนนี	จำนวน (คน)
กรุงเทพฯ	99
สวรรคค์ประชารักษ์	100
ผดุงครรภ์นครสวรรค์	119
ราชบุรี 2	100
บำราชนราดูร	48
ชัยนาท	100
สระบุรี	96
พระพุทธบาท	108
ชลบุรี	100
ลำปาง 2	110
รวม	980

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

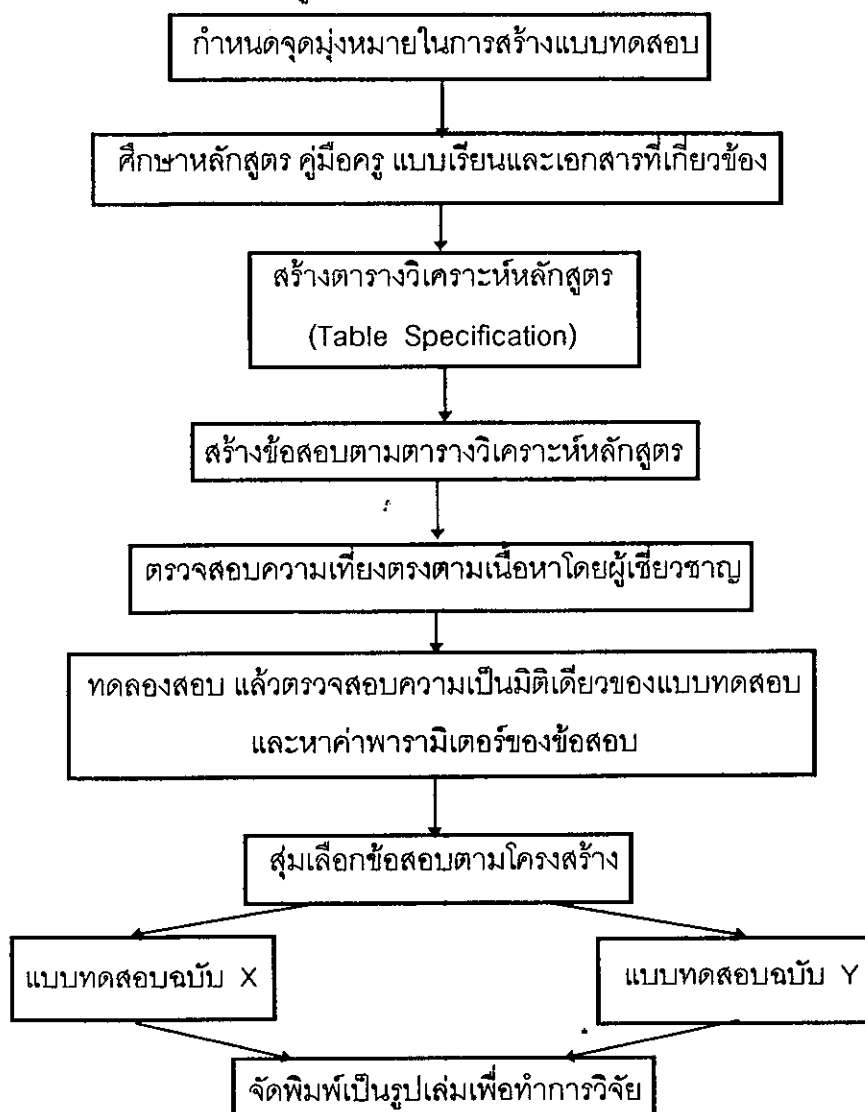
เครื่องมือที่ใช้ในการศึกษาครั้งนี้ เป็นแบบทดสอบเลือกตอบวัดผลสัมฤทธิ์ทางการเรียนวิชาการพยาบาลสูติศาสตร์ 1 (พย.212) เรื่องการพยาบาลในระยะตั้งครรภ์ สำหรับนักศึกษาพยาบาลศาสตร์(ระดับต้น) ชั้นปีที่ 2 ปีการศึกษา 2538 ที่ผู้วิจัยสร้างขึ้นตามประมวลรายวิชาหลักสูตรประกาศนียบัตรพยาบาลศาสตร์(ระดับต้น) จำนวน 100 ข้อ เป็นแบบทดสอบเลือกตอบชนิด 5 ตัวเลือก และแบ่งข้อสอบออกเป็น 2 ฉบับ โดยการพิจารณาค่าความยาก ค่าอำนาจจำแนกและค่าสัมประสิทธิ์การเดาที่เท่าหรือเทียบเคียงกันจะได้แบบทดสอบ 2 ฉบับดังนี้

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

โดยที่แบบทดสอบในแต่ละฉบับ จะมีแบบทดสอบร่วมชนิดแบบทดสอบร่วมภายใน อยู่ด้วย(Internal Anchor Test) จำนวน 10 ข้อรวมอยู่ด้วย โดยมีขั้นตอนการสร้าง แบบทดสอบดังนี้

วิธีดำเนินการสร้างแบบทดสอบ

ในการสร้างแบบทดสอบผู้วิจัยได้ดำเนินการสร้างแบบทดสอบตามแผนผังดังนี้



ภาพประกอบ 3 แผนผังการสร้างแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา พย. 212

รายละเอียดในการดำเนินการสร้างแบบทดสอบมีดังนี้

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบ
 - 1.1 เพื่อสร้างแบบทดสอบเลือกตอบวัดผลสัมฤทธิ์ทางการเรียนวิชาการพยาบาลสูติศาสตร์ ชั้นปีที่ 2 หลักสูตรพยาบาลศาสตร์(ระดับต้น) เพื่อที่จะนำไปใช้ในการวิจัย
 - 1.2 เพื่อเปรียบเทียบผลการเทียบคะแนนของแบบทดสอบ
 2. ศึกษาหลักสูตร คู่มือครู แบบเรียนและเอกสารที่เกี่ยวข้อง ในเนื้อหาวิชาการพยาบาลสูติศาสตร์ 1 (พย.212) จากหลักสูตรสำหรับนักศึกษาพยาบาลศาสตร์(ระดับต้น) ชั้นปีที่ 2 เรื่องการพยาบาลมารดาในระยะตั้งครรภ์
 3. สร้างตารางวิเคราะห์หลักสูตร(Table Specification) โดยกำหนดเนื้อหาที่จะนำมาใช้สร้างข้อสอบและพฤติกรรมการเรียนรู้ที่ต้องการวัด แล้วให้ผู้เชี่ยวชาญจำนวน 5 ท่าน คือ ครูผู้สอนวิชาการพยาบาลสูติศาสตร์ 1 ที่มีประสบการณ์ด้านการสอนอย่างน้อย 3 ปี จำนวน 2 ท่าน และนักวัดผลที่จบปริญญาโททางการวัดผลจำนวน 3 ท่าน เป็นผู้พิจารณาว่าแต่ละเนื้อหาควรวัดพฤติกรรมด้านใดบ้าง
 4. สร้างข้อสอบปรนัยชนิดเลือกตอบ 5 ตัวเลือก จำนวน 100 ข้อ ตามเนื้อหาและพฤติกรรมที่กำหนดไว้ในตารางวิเคราะห์หลักสูตร(Table Specification) ดังแสดงในภาคผนวก
 5. ตรวจสอบความเที่ยงตรงตามเนื้อหา(Content Validity) โดยนำข้อสอบที่สร้างขึ้นไปให้ผู้เชี่ยวชาญตรวจสอบว่าข้อสอบที่สร้างขึ้นวัดเนื้อหาและพฤติกรรมที่ต้องการหรือไม่ โดยพิจารณาค่า IOC ที่มีค่าตั้งแต่ 0.5 ขึ้นไป จะถือว่าเป็นข้อสอบที่มีความเที่ยงตรงตามเนื้อหา
- ตัวอย่าง แบบประเมินความสอดคล้องระหว่างข้อสอบกับพฤติกรรมที่ต้องการวัด
- คำชี้แจง ให้ท่านพิจารณาว่าข้อสอบแต่ละข้อต่อไปนี้ วัดได้ตรงตามเนื้อหาและพฤติกรรมที่กำหนดไว้หรือไม่ โดยเขียนเครื่องหมาย ✓ ลงในช่องความคิดเห็น พิจารณาให้คะแนนตามเกณฑ์ดังนี้
- ให้คะแนน +1 เมื่อท่านพิจารณาแล้วเห็นว่าข้อสอบข้อนั้นวัดได้ตรงตามพฤติกรรมนั้นจริง
 - ให้คะแนน 0 เมื่อท่านพิจารณาแล้วไม่แน่ใจว่าข้อสอบข้อนั้นวัดได้ตรงตามพฤติกรรมนั้นจริง
 - ให้คะแนน -1 เมื่อท่านพิจารณาแล้วเห็นว่าข้อสอบข้อนั้นไม่ได้วัดตรงตามพฤติกรรมนั้นจริง

พฤติกรรมที่ต้องการวัด	ข้อสอบ	คะแนนการพิจารณา		
		+1	0	-1
ความจำเป็นเกี่ยวกับการเปลี่ยนแปลงทางสรีรวิทยาในระยะตั้งครรภ์	00. ขณะตั้งครรภ์ อวัยวะใดที่มีการเปลี่ยนแปลงมากที่สุด ก. ผนังมดลูก ข. เอ็นยึดมดลูก ค. กล้ามเนื้อมดลูก ง. กล้ามเนื้อหน้าท้อง จ. กล้ามเนื้อพื้นเชิงกราน	√		

หลังจากให้ผู้เชี่ยวชาญพิจารณาความสอดคล้อง ระหว่างข้อสอบกับพฤติกรรมที่ต้องการวัด นำคะแนนไปคำนวณหาค่าดัชนีความสอดคล้อง (ล้วน สายยศ และอังคณา สายยศ. 2536) ถ้ามีค่ามากกว่า 0.5 ถือว่าข้อสอบข้อนั้น ๆ วัดตรงตามพฤติกรรมนั้นจริง ผลปรากฏว่า ได้ข้อสอบที่มีค่าดัชนีความสอดคล้องตั้งแต่ 0.5 - 1.0 จำนวน 100 ข้อ

6. นำแบบทดสอบที่สร้างขึ้นทั้งหมดไปทดลองสอบกับนักศึกษาพยาบาลศาสตร์ (ระดับต้น) ชั้นปีที่ 2 ปีการศึกษา 2538 ที่ไม่ใช่กลุ่มตัวอย่าง จำนวน 501 คน เพื่อที่จะนำมาวิเคราะห์ดังนี้

6.1 ตรวจสอบความเป็นมิติเดียวของแบบทดสอบ(Unidimensionality) โดยวิธีการวิเคราะห์องค์ประกอบ(Factor Analysis) โดยการใช้โปรแกรม SPSS/FW ซึ่งเลือกสกัดองค์ประกอบหลัก(Principal Component : PC) แล้วหมุนแกนด้วยวิธีแวนิแมกซ์ (Varimax) ได้ผลดังตาราง 13 ในภาคผนวก ข ซึ่งปรากฏว่า ได้องค์ประกอบที่มีค่าไอเกนเกิน 1.0 จำนวน 30 องค์ประกอบ โดยที่ค่าไอเกน(Eigen Value) ขององค์ประกอบที่ 1 ของแบบทดสอบ มีค่าเท่ากับ 4.485 และค่าไอเกนขององค์ประกอบที่ 2 ถึงองค์ประกอบที่ 30 มีค่าอยู่ระหว่าง 2.409 ถึง 1.048 แสดงให้เห็นว่าองค์ประกอบที่ 1 มีค่าไอเกนสูงกว่าองค์ประกอบอื่นๆเป็นอย่างมาก และองค์ประกอบที่ 2 ถึงองค์ประกอบที่ 30 มีค่าไอเกนที่ใกล้เคียงกัน แสดงว่าแบบทดสอบวิชาการพยาบาลสูติศาสตร์ 1 มีความเป็นมิติเดียวพอประมาณ(Warm. 1978 : 104 ; Lord. 1980 : 19) นั่นคือสามารถนำแบบทดสอบไปวิเคราะห์ตามทฤษฎีการตอบข้อสอบได้

6.2 หาค่าพารามิเตอร์ของข้อสอบจากแบบทดสอบคือค่าความยาก(bi) ค่าอำนาจจำแนก(ai) และค่าสัมประสิทธิ์การเดา(ci) โดยการวิเคราะห์ด้วยโปรแกรม BILOG เวอร์ชัน 3.04 ผลปรากฏว่า มีข้อสอบจำนวน 23 ข้อ ที่ไม่สามารถประมาณค่าพารามิเตอร์ของข้อสอบได้ ดังนั้นจึงเหลือข้อสอบที่วิเคราะห์ทั้งหมดจำนวน 77 ข้อ ซึ่งมีค่าความยากอยู่ระหว่าง -4.959 ถึง 6.311 ค่าอำนาจจำแนกอยู่ระหว่าง .144 ถึง 1.142 และค่าสัมประสิทธิ์การเดาอยู่ระหว่าง 0.074 ถึง 0.500 (รายละเอียดดังในตาราง 14 ในภาคผนวก ค) ผู้วิจัยจึงคัดเลือกข้อสอบที่มีคุณภาพถึงเกณฑ์คือค่าความยากตั้งแต่ -3.00 ถึง 3.00 ค่าอำนาจจำแนกตั้งแต่ 0.3 ถึง 2.00 และค่าสัมประสิทธิ์การเดาต่ำกว่า 0.30 จำนวน 50 ข้อ ได้ผลดังตาราง 15 ในภาคผนวก ง

7. จากข้อสอบที่มีคุณภาพตามข้อ 6 นำข้อสอบแต่ละข้อมาจัดเป็นคู่ตามเนื้อหาและพฤติกรรมเดียวกัน โดยพิจารณาค่าความยาก ค่าอำนาจจำแนก และค่าสัมประสิทธิ์การเดาที่เท่าหรือเทียบเคียงกันมากที่สุด จำนวน 20 คู่ และเลือกข้อสอบที่เหลือในแต่ละเนื้อหา ๆ ละ 2 ข้อ จำนวน 10 ข้อ ใช้เป็นแบบทดสอบร่วม แล้วนำข้อสอบที่เป็นคู่มาแบ่งเป็น 2 ฉบับ กับแบบทดสอบร่วมอีก 10 ข้อ จะได้แบบทดสอบจำนวน 2 ฉบับ คือ ฉบับ X จำนวน 30 ข้อ และฉบับ Y จำนวน 30 ข้อ โดยที่ในแบบทดสอบแต่ละฉบับจะมีแบบทดสอบร่วมชนิด แบบทดสอบร่วมภายใน(Internal Anchor Test) จำนวน 10 ข้อ รวมอยู่ด้วย ซึ่งมีคุณภาพ ดังนี้

แบบทดสอบ ฉบับ X มีค่าพารามิเตอร์ค่าอำนาจจำแนกอยู่ระหว่าง 0.144 ถึง 0.848 ค่าพารามิเตอร์ความยากอยู่ระหว่าง -2.893 ถึง 2.615 และค่าสัมประสิทธิ์การเดาของแบบทดสอบมีค่าอยู่ระหว่าง 0.072 ถึง 0.227 ส่วนอินฟอริเมชันของแบบทดสอบมีค่าอยู่ระหว่าง 0.25 ถึง 2.625 และ ณ ระดับความสามารถที่เพิ่มขึ้น ค่าอินฟอริเมชันของแบบทดสอบจะเพิ่มขึ้น ซึ่งจะให้ค่าที่สูงสุดที่ระดับความสามารถ 1.50

แบบทดสอบ ฉบับ Y มีค่าพารามิเตอร์ค่าอำนาจจำแนกอยู่ระหว่าง 0.144 ถึง 0.893 ค่าพารามิเตอร์ความยากอยู่ระหว่าง -2.893 ถึง 2.716 และค่าสัมประสิทธิ์การเดาของแบบทดสอบมีค่าอยู่ระหว่าง 0.070 ถึง 0.206 ส่วนอินฟอริเมชันของแบบทดสอบมีค่าอยู่ระหว่าง 0.183 ถึง 3.25 และ ณ ระดับความสามารถที่เพิ่มขึ้น ค่าอินฟอริเมชันของแบบทดสอบจะเพิ่มขึ้น ซึ่งจะให้ค่าที่สูงสุดที่ระดับความสามารถ 2.0

จากแบบทดสอบที่ได้จำนวน 2 ฉบับ (ฉบับ X และฉบับ Y) มีคุณภาพดังรายละเอียดในภาคผนวก ง

วิธีดำเนินการเก็บรวบรวมข้อมูล

ในการเก็บรวบรวมข้อมูลในครั้งนี้ ผู้วิจัยได้ดำเนินการตามขั้นตอนดังนี้

1. วางแผนในการดำเนินการสอบ โดยติดต่อขอความร่วมมือในการนำแบบทดสอบไปทดสอบกับนักศึกษาพยาบาลศาสตร์(ระดับต้น) ชั้นปีที่ 2 จากผู้อำนวยการวิทยาลัยพยาบาลที่เป็นกลุ่มตัวอย่าง กำหนดวัน เวลาและสถานที่สอบ
2. นำแบบทดสอบทั้ง 2 ฉบับ(ฉบับ X และ ฉบับ Y) ไปทดสอบกับกลุ่มตัวอย่างจำนวน 980' คน โดยจัดให้ทุกคนทำแบบทดสอบทั้งสองฉบับ แล้วนำมาสุ่มโดยใช้คะแนนจากแบบทดสอบฉบับ X จำนวน 490 ฉบับ และฉบับ Y จำนวน 490 ฉบับ ส่วนคะแนนที่เหลือในแต่ละฉบับใช้เป็นคะแนนทานผลของการเทียบคะแนน
3. นำผลการสอบที่ได้ในข้อ 2 มาเตรียมข้อมูลสำหรับการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบและพารามิเตอร์ของข้อสอบ แล้วนำผลการประมาณค่าไปวิเคราะห์ทางสถิติและทดสอบสมมติฐานของการวิจัยต่อไป

การวิเคราะห์ข้อมูลและสถิติที่ใช้

การวิเคราะห์ข้อมูลดำเนินการโดยแบ่งออกเป็น 4 ตอนได้แก่

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

ตอนที่ 1 การวิเคราะห์แบบทดสอบ

1.1 ตรวจสอบความเป็นมิติเดียว (Unidimensionality) ของแบบทดสอบแต่ละฉบับตามข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อสอบ โดยนำแบบทดสอบแต่ละฉบับไปทดลองสอบกับนักศึกษาพยาบาลศาสตร์(ระดับต้น) ชั้นปีที่ 2 ที่ไม่ใช่กลุ่มตัวอย่าง จำนวน 501 คน แล้วนำผลมาวิเคราะห์องค์ประกอบ (Factor Analysis) โดยการใช้โปรแกรมสำเร็จรูป SPSS/FW ซึ่งเลือกสกัดองค์ประกอบด้วยวิธีการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis : PC) แล้วหมุนแกนด้วยวิธีแวนิแมกซ์ (Varimax)

✓ 1.2 วิเคราะห์ข้อสอบในแบบทดสอบเทียบคะแนน และแบบทดสอบรวมเพื่อประมาณค่าพารามิเตอร์ของข้อสอบคือ ค่าความยาก(b_i) ค่าอำนาจจำแนก(a_i) และค่าสัมประสิทธิ์การเดา(c_i) ด้วยโปรแกรมสำเร็จรูป BILOG เวอร์ชัน 3.04 (Mislevy and Bock. : 1982) ได้ผลดังตาราง 15 ในภาคผนวก ง

✓ 1.3 หาค่าสถิติพื้นฐานของแบบทดสอบแต่ละฉบับก่อนการเทียบคะแนน คือค่าเฉลี่ย ค่าความเบี่ยงเบนมาตรฐาน ด้วยโปรแกรมสำเร็จรูป SPSS/FW

ตอนที่ 2 การสร้างตารางเทียบคะแนนแบบทดสอบ

ผู้วิจัยมีขั้นตอนการสร้างตารางเทียบคะแนนในแต่ละวิธีดังนี้

2.1 การเทียบคะแนนจริงโดยการใช้แบบทดสอบรวม

มีขั้นตอนการสร้างตารางเทียบคะแนนดังนี้

ขั้นที่ 1 ประมาณค่าพารามิเตอร์ของความสามารถของผู้สอบแต่ละคนในแบบทดสอบฉบับ X และ ฉบับ Y โดยการใช้โปรแกรมสำเร็จรูป BILOG เวอร์ชัน 3.04

ขั้นที่ 2 ประมาณค่าพารามิเตอร์ของข้อสอบในแบบทดสอบฉบับ X ฉบับ Y และแบบทดสอบรวม โดยการใช้โปรแกรมสำเร็จรูป BILOG เวอร์ชัน 3.04

✓ ขั้นที่ 3 แปลงค่าพารามิเตอร์อำนาจจำแนก(a_i) ความยาก(b_i) และพารามิเตอร์ความสามารถ(θ) ที่ได้จากการประมาณค่าโดยการใช้โปรแกรมสำเร็จรูป BILOG เวอร์ชัน 3.04 ให้อยู่ในสเกลเดียวกัน จากสูตร (Warm. 1978 : 114-115)

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

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

$$a_1 = \left[\frac{SD_{b_2}}{SD_{b_1}} \right] a_2$$

- เมื่อ θ_1 คือพารามิเตอร์ความสามารถของผู้สอบที่ทำแบบทดสอบฉบับ Y
 θ_2 คือพารามิเตอร์ความสามารถของผู้สอบที่ทำแบบทดสอบฉบับ X
 SD_{b_1} คือค่าเบี่ยงเบนมาตรฐานของพารามิเตอร์ความยากของแบบทดสอบร่วม
 ของกลุ่มผู้สอบที่ทำแบบทดสอบฉบับ Y
 SD_{b_2} คือค่าเบี่ยงเบนมาตรฐานของพารามิเตอร์ความยากของแบบทดสอบร่วม
 ของผู้สอบกลุ่มที่ทำแบบทดสอบฉบับ X
 \bar{b}_1 คือค่าเฉลี่ยของพารามิเตอร์ความยากของแบบทดสอบร่วมของกลุ่มที่ทำ
 แบบทดสอบ ฉบับ Y
 \bar{b}_2 คือค่าเฉลี่ยของพารามิเตอร์ความยากของแบบทดสอบร่วมของกลุ่มที่ทำ
 แบบทดสอบ ฉบับ X
 b_1 คือพารามิเตอร์ความยากของแบบทดสอบฉบับ Y
 b_2 คือพารามิเตอร์ความยากของแบบทดสอบฉบับ X
 a_1 คือพารามิเตอร์อำนาจจำแนกของแบบทดสอบฉบับ Y
 a_2 คือพารามิเตอร์อำนาจจำแนกของแบบทดสอบฉบับ X

ขั้นที่ 4 นำค่าพารามิเตอร์ที่แปลงแล้วมาคำนวณค่าความน่าจะเป็นในการ
 ตอบข้อสอบถูกของผู้สอบทั้งสองกลุ่ม จากสูตร (Hambleton and Swaminathan, 1985 :
 37-38)

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

เมื่อ $P_i(\theta)$ คือความน่าจะเป็นในการตอบข้อสอบข้อที่ i ถูกเมื่อผู้สอบมีความ
 สามารถ (θ)

- a_i คือ พารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i
 b_i คือ พารามิเตอร์ความยากของข้อสอบข้อที่ i
 c_i คือ ค่าพารามิเตอร์การเดาของข้อสอบข้อที่ i
 D คือ ค่าปรับเสถลมีค่าประมาณ 1.702
 e คือ ค่าคงที่มีค่าประมาณ 2.718

ขั้นที่ 5 คำนวณคะแนนจริงของผู้สอบแต่ละคนที่ได้จากการตอบแบบทดสอบแต่ละข้อในแต่ละฉบับ โดยใช้สูตร (Hambleton and Swaminathan, 1985 : 210)

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

$$\xi_y = \sum_{i=1}^m P_i(\theta_y)$$

เมื่อ ξ_x คือ คะแนนจริงของผู้สอบที่มีความสามารถ θ_x ได้จากการตอบแบบทดสอบ ฉบับ X

ξ_y คือ คะแนนจริงของผู้สอบที่มีความสามารถ θ_y ได้จากการตอบแบบทดสอบฉบับ Y

หมายเหตุ การคำนวณหาค่า $P_i(\theta_x)$, $P_i(\theta_y)$, ξ_x และ ξ_y ได้อาศัยโปรแกรมคอมพิวเตอร์ ในการคำนวณได้แยกคำนวณทีละกรณีของแบบทดสอบทีละชุด และใช้วิธีการกำหนดค่า θ ที่ครอบคลุมพิสัยกลุ่มตัวอย่าง

ขั้นที่ 6 นำคะแนนจริงของผลการสอบที่หาได้ในขั้นที่ 5 มาหาความสัมพันธ์กันระหว่าง ξ_x และ ξ_y โดยพล็อตกราฟให้แกนนอน x เป็นค่าความสามารถ θ ของผู้สอบ แกนตั้ง y เป็นคะแนนจริงของแบบทดสอบทั้งสองฉบับ ดังนั้นค่าสมนัยกันของ ξ_x และ ξ_y มีความหมายเป็นอย่างเดียวกัน จึงเป็นการเทียบมาตราของแบบทดสอบสองชุด ณ ที่จุดแต่ละจุดของ θ เดียวกัน

ขั้นที่ 7 สร้างตารางเทียบคะแนนของแบบทดสอบทั้งสองฉบับ ณ ค่า θ เดียวกันเป็นหลัก ปรากฏผลดังตาราง 17 ในภาคผนวก ข

2.2 การเทียบคะแนนดิบโดยการใช้แบบทดสอบร่วม ✓

มีขั้นตอนการสร้างตารางเทียบคะแนนดังนี้คือ

ขั้นที่ 1 ประเมินค่าพารามิเตอร์ของความสามารถของผู้สอบแต่ละคน ในแบบทดสอบฉบับ X และ ฉบับ Y โดยการใช้โปรแกรมสำเร็จรูป BILOG เวอร์ชัน 3.04

ขั้นที่ 2 ประมาณค่าพารามิเตอร์ของข้อสอบในแบบทดสอบฉบับ X ฉบับ Y และแบบทดสอบร่วมโดยการใช้โปรแกรมสำเร็จรูป BILOG เวอร์ชัน 3.04

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

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

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

เมื่อ $f(x/\theta)$ คือ ค่า Conditional Frequency Distribution ของแบบทดสอบ ฉบับ X

$f(y/\theta)$ คือ ค่า Conditional Frequency Distribution ของแบบทดสอบ ฉบับ Y

x คือ จำนวนข้อที่ตอบถูกในแบบทดสอบฉบับ X

y คือ จำนวนข้อที่ตอบถูกในแบบทดสอบฉบับ Y

θ คือ พารามิเตอร์ความสามารถของผู้สอบคนที่ i โดยที่

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

t คือ ความถี่ หรือจำนวนคนที่ได้คะแนน x โดยที่ $x = 0, 1, 2, \dots, n$

และ y โดยที่ $y = 0, 1, 2, \dots, m$

$P_i(\theta)$ คือ โอกาสหรือความน่าจะเป็นในการตอบข้อสอบข้อที่ i ถูก ของ ผู้สอบที่มีความสามารถ θ

$Q_i(\theta)$ คือ ค่าที่หามาจากสมการ $Q \equiv 1 - P$

หมายเหตุ การคิดคำนวณโดยเขียนโปรแกรมคำสั่งลงในคอมพิวเตอร์

ขั้นที่ 4 นำคะแนนดิบของผลการสอบที่หาได้ในขั้นที่ 3 มาเขียนกราฟแสดงความสัมพันธ์ของการเทียบคะแนนดิบของแบบทดสอบฉบับ X และ ฉบับ Y โดยให้แกนตั้ง y

เป็นคะแนนดิบ (Raw Score) และแกนนอน x เป็นคะแนนโอกาสในการตอบถูกต้องตั้งแต่คะแนน 0 ถึง 28 คะแนน

ขั้นที่ 5 สร้างตารางเทียบคะแนนของแบบทดสอบทั้ง 2 ฉบับ ณ θ เดียวกัน เป็นหลัก ปรากฏผลดังตาราง 18 ในภาคผนวก ข

ตอนที่ 3 การศึกษาผลของการเทียบคะแนน ซึ่งพิจารณาได้จาก

3.1 คะแนนที่ได้จากการเทียบคะแนน ✓

คะแนนที่ได้จากการเทียบคะแนนตามวิธีในข้อที่ 2.1 และ 2.2 ซึ่งแสดงรายละเอียดดังตาราง 20 ในภาคผนวก ฉ

3.2 ความคลาดเคลื่อนของการเทียบคะแนน (Standard Error of Equating) ✓

หาได้จากการใช้สูตร (Macro. 1977 : 157)

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

เมื่อ MSE คือ Mean Square Error

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

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

N คือ จำนวนผู้สอบที่เป็นกลุ่มเทียบคะแนน

3.3 ความเพียงพอของการเทียบคะแนน (Discrepancy Index)

ในการหาค่าความเพียงพอของการเทียบคะแนนในแต่ละวิธีหาได้จากสูตร (Petersen and others. 1982 : 91)

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

- เมื่อ C คือค่าความเพียงพอของการเทียบคะแนน
 x_i คือคะแนน x ของผู้สอบคนที่ i จากแบบทดสอบชุด X
 x_i^* คือคะแนน x ของผู้สอบคนที่ i จากการแปลงคะแนน y
 ที่สมมูลกัน
 n คือจำนวนคนในกลุ่มตัวอย่างสอบทานผล
 S^2_x คือค่าความแปรปรวนของคะแนนจากแบบทดสอบ X

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

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

ตอนที่ 4 การทดสอบความแตกต่างของผลการเทียบคะแนน ✓

ในการทดสอบความแตกต่างของผลการเทียบคะแนนทั้ง 2 วิธีซึ่งได้แก่ คะแนนที่ได้จากการเทียบ ความคลาดเคลื่อนและความเพียงพอของการเทียบคะแนน โดยการ ใช้สถิติวิเคราะห์ คือ t -test แบบ t -Dependent (ล้วนและ อังคณา สายยศ. 2536 : 87)

บทที่ 4

ผลการวิเคราะห์ข้อมูล

ในการเสนอผลการวิเคราะห์ข้อมูล ผู้วิจัยได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูลดังนี้

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

- \bar{X} แทน คะแนนเฉลี่ย
- N แทน จำนวนนักศึกษาในกลุ่มตัวอย่าง
- n แทน จำนวนข้อในแบบทดสอบ
- k แทน จำนวนคู่ที่ใช้ในการทดสอบความแตกต่าง
- SD แทน ค่าเบี่ยงเบนมาตรฐาน
- θ แทน ค่าความสามารถของผู้สอบ
- θ_1 แทน ค่าความสามารถของผู้สอบที่ทำแบบทดสอบฉบับ X
- θ_2 แทน ค่าความสามารถของผู้สอบที่ทำแบบทดสอบฉบับ Y
- a แทน ค่าพารามิเตอร์อำนาจจำแนกของแบบทดสอบ
- b แทน ค่าพารามิเตอร์ความยากของแบบทดสอบ
- c แทน ค่าพารามิเตอร์สัมประสิทธิ์การเดาของแบบทดสอบ
- MSE แทน ค่าความคลาดเคลื่อน
- t-test แทน ค่าการทดสอบค่าที่
- df แทน ชั้นของความเป็นอิสระ
- C แทน ดัชนีความเพียงพอของการเทียบคะแนน

การเสนอผลการวิเคราะห์ข้อมูล

ในการศึกษาค้างนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับดังนี้

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

ผลการวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐานของแบบทดสอบ

จากการนำแบบทดสอบวิชาการพยาบาลสุติศาสตร์ 1 จำนวน 2 ฉบับ คือ ฉบับ X และฉบับ Y ฉบับละ 28 ข้อ ไปทดสอบกับนักศึกษาพยาบาลศาสตร์(ระดับต้น) ชั้นปีที่ 2 จำนวน 980 คน ได้ค่าสถิติพื้นฐานของคะแนนดังแสดงในตาราง 3

ตาราง 3 ค่าสถิติพื้นฐานของแบบทดสอบวิชาการพยาบาลสุติศาสตร์ 1

แบบทดสอบ	คะแนนเต็ม	\bar{X}	SD
ฉบับ X	28	13.906	3.944
ฉบับ Y	28	11.835	4.741

จากตาราง 3 แสดงให้เห็นว่า แบบทดสอบทั้งสองฉบับมีคะแนนเฉลี่ยที่ต่ำกว่า 14 คะแนน แสดงว่าแบบทดสอบทั้งสองฉบับค่อนข้างยาก และแบบทดสอบฉบับ X มีคะแนนเฉลี่ยประมาณครึ่งหนึ่งของคะแนนเต็ม แสดงว่ามีความยากระดับปานกลาง แบบทดสอบฉบับ Y มีคะแนนเฉลี่ยที่ต่ำกว่า 14 คะแนน แสดงว่าค่อนข้างยาก ส่วนค่าความเบี่ยงเบนมาตรฐานพบว่าแบบทดสอบฉบับ Y มีค่าการกระจายคะแนนมากกว่าแบบทดสอบ ฉบับ X

2. การสร้างตารางเทียบคะแนนของแบบทดสอบ

2.1 ผลการสร้างตารางเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วม

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

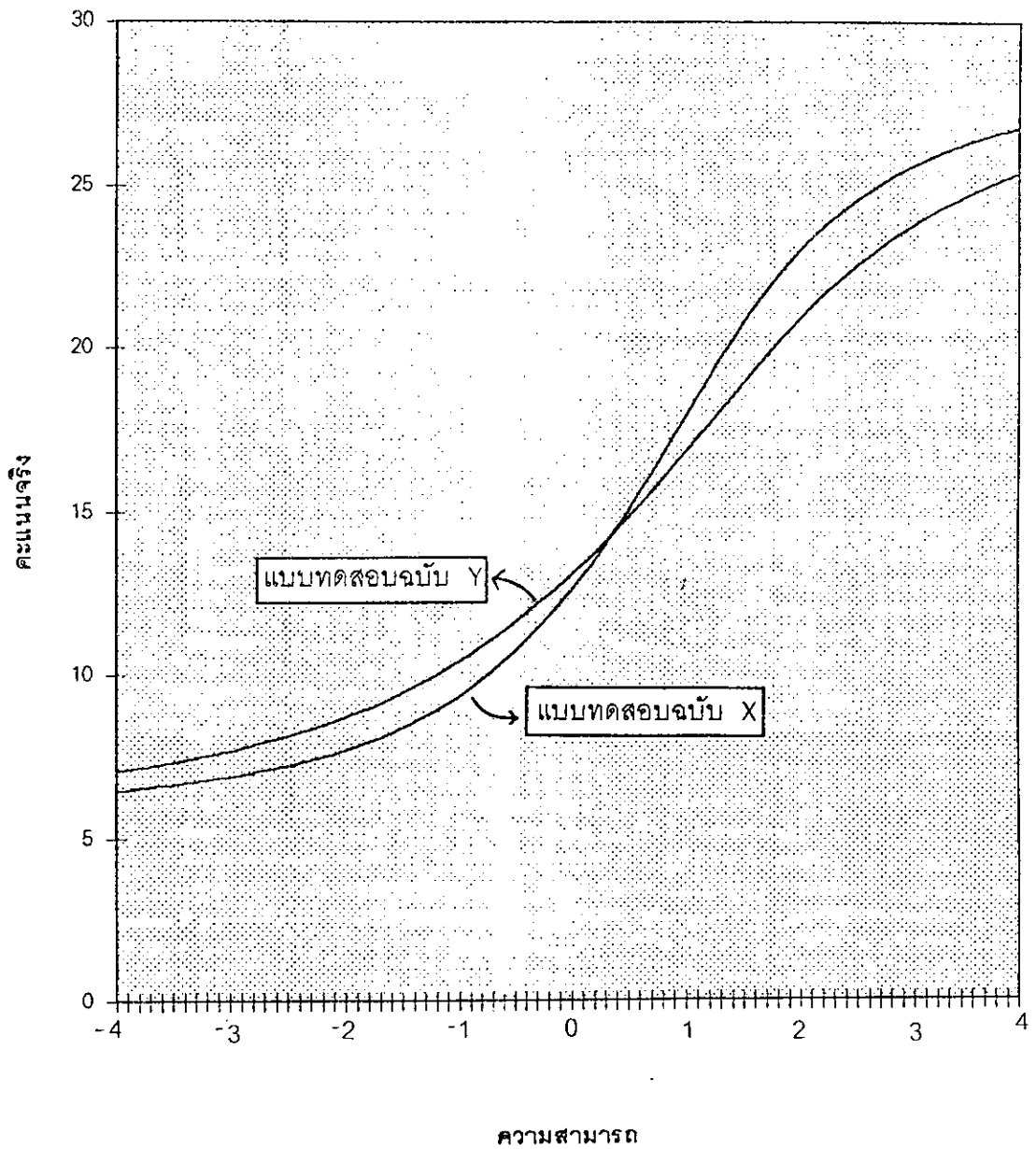
$$\theta_1 = 0.7075\theta_2 + 0.1384$$

$$b_1 = 0.7075b_2 + 0.1384$$

$$a_1 = 1.413a_2$$

จากนั้นผู้วิจัยได้นำค่าพารามิเตอร์ที่แปลงให้อยู่ในสเกลเดียวกัน
ค่าคะแนนจริงของผู้สอบที่ระดับความสามารถต่างๆ แล้วนำมาเขียนกราฟ
ภาพประกอบ 4

ไปคำนวณหา
ดังแสดงใน



ภาพประกอบ 4 คะแนนจริงของผู้สอบแบบทดสอบทั้งสองฉบับที่ระดับความสามารถต่างๆ

จากกราฟแสดงคะแนนจริงของผู้สอบแบบทดสอบทั้งสองฉบับที่ระดับความสามารถต่างๆ สามารถอ่านค่าได้เป็นคะแนนสมมูลของการเทียบคะแนน โดยเทียบคะแนนจากแบบทดสอบฉบับ X ไปสู่อันดับ Y ที่ระดับความสามารถเดียวกัน และเนื่องจากคะแนนที่ได้มีค่าคะแนนใกล้เคียงกัน จึงต้องขยายคะแนนด้วยวิธีการแปลงเชิงเส้นตรง (Linear Transformation) ให้เป็นคะแนนมาตรฐาน CEEB เพื่อที่จะให้เห็นคะแนนที่มีค่าแตกต่างกัน จะได้ตารางเทียบคะแนนดังนี้

ตาราง 4 คะแนนสมมูลของการเทียบคะแนน จากแบบทดสอบฉบับ X ไปสู่แบบทดสอบฉบับ Y ที่ระดับคะแนนและความสามารถเดียวกันตามวิธีการเทียบคะแนนจริง

คะแนน	θ	x	y	CEEB _x	CEEB _y
5	-4.0	7.02	6.45	382	390
6	-1.9	8.68	7.69	408	406
7	-1.0	10.16	9.07	432	425
8	-0.5	11.31	10.37	450	442
9	-0.2	12.12	11.37	463	455
10	0.1	13.03	12.55	477	471
11	0.3	13.70	13.46	488	483
12	0.4	14.05	13.94	493	490
13	0.6	14.78	17.98	505	504
14	0.7	15.16	15.53	511	511
15	0.8	15.56	16.10	517	519
16	1.0	16.37	17.28	530	534
17	1.1	16.78	17.88	537	542
16	1.2	17.20	17.47	543	550
17	1.3	17.63	19.06	550	558
18	1.4	18.05	19.64	557	566

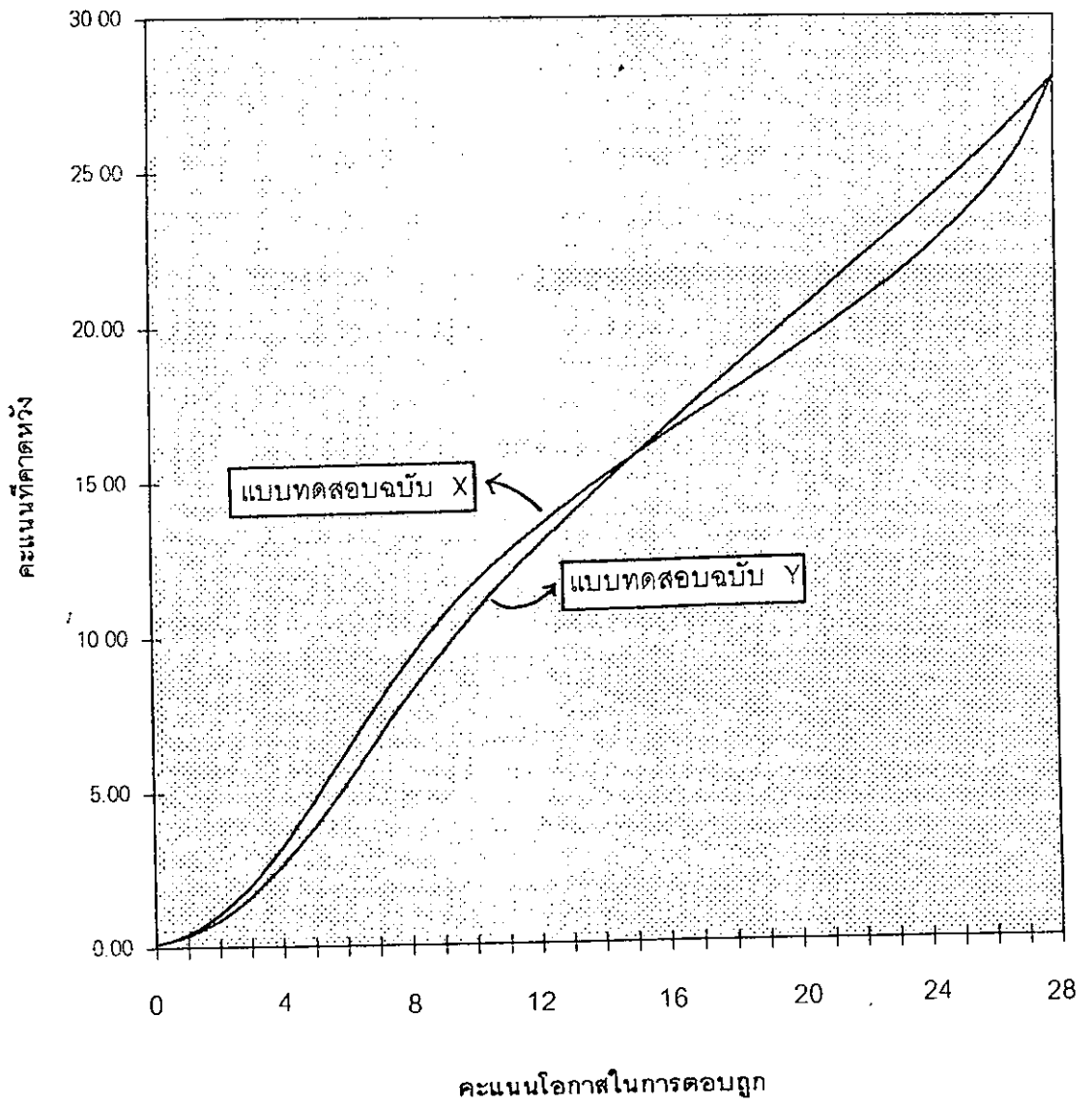
ตาราง 4 (ต่อ)

คะแนน	θ	x	y	CEEB _x	CEEB _y
19	1.3	17.63	19.06	550	558
20	1.4	18.05	19.64	557	566
21	1.5	18.47	20.19	564	573
22	1.6	18.89	20.73	570	580
23	1.8	19.71	21.71	583	593
24	2.0	20.48	22.57	596	605
25	2.3	21.55	23.65	612	619
26	2.6	22.48	24.51	627	631
27	3.4	26.07	26.07	657	652
28	4.0	26.76	26.76	673	661

จากตาราง 4 แสดงให้เห็นว่า ที่ระดับความสามารถตั้งแต่ -4.0 ถึง 4.0 คะแนนจริงของแบบทดสอบฉบับ X มีค่าอยู่ระหว่าง 7.02 ถึง 26.76 และเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้วพบว่า มีค่าอยู่ระหว่าง 382 ถึง 673 ส่วนคะแนนจริงของแบบทดสอบฉบับ Y มีค่าอยู่ระหว่าง 6.45 ถึง 26.76 และเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้วพบว่า มีค่าอยู่ระหว่าง 390 ถึง 661

2.2 ผลการสร้างตารางเทียบคะแนนดิบโดยการใช้แบบทดสอบร่วม

ในการสร้างตารางเทียบคะแนนของแบบทดสอบได้ดำเนินการโดยนำผลการวิเคราะห์ค่าพารามิเตอร์ต่างๆ ไปใช้กับสูตรของลอร์ด แฮมบลิตันและสวามินาทาน (Lord, 1980 : 196-197 ; Hambleton and Swaminathan, 1985 : 215 : 216) เพื่อที่จะคำนวณหาค่าคะแนนดิบของผู้สอบที่ระดับคะแนนโอกาสในการตอบข้อสอบถูก แล้วนำมาเขียนกราฟ ดังแสดงในภาพประกอบ 5



ภาพประกอบ 5 คะแนนที่คาดหวังของผู้สอบแบบทดสอบทั้งสองฉบับที่คะแนนโอกาสในการตอบถูก

จากกราฟแสดงคะแนนดิบที่ได้จากผู้สอบแบบทดสอบทั้งสองฉบับ ที่ระดับคะแนนต่างๆ สามารถอ่านค่าได้คะแนนสมมูลของการเทียบคะแนน โดยเทียบจากแบบทดสอบฉบับ X ไปสู่อันดับ Y ที่ระดับคะแนนเดียวกัน ที่ระดับความสามารถเดียวกัน และเนื่องจากคะแนนที่ได้มีค่าคะแนนใกล้เคียงกัน จึงต้องขยายคะแนนด้วยวิธีการแปลงเชิงเส้นตรง (Linear Transformation) ให้เป็นคะแนนมาตรฐาน CEEB เพื่อที่จะให้เห็นคะแนนที่มีค่าแตกต่างกัน จะได้ตารางเทียบคะแนนดังนี้

ตาราง 5 คะแนนสมมูลของการเทียบคะแนน จากแบบทดสอบฉบับ X ไปสู่แบบทดสอบฉบับ Y ที่ระดับคะแนนและความสามารถเดียวกันตามวิธีการเทียบคะแนนดิบ

คะแนน	θ	x	y	CEEB _x	CEEB _y
5	-4.0	3.81	4.65	384	386
6	-1.9	5.20	6.63	400	406
7	-1.0	6.67	7.81	416	425
8	-0.5	8.11	9.27	433	443
9	-0.2	9.47	10.56	448	459
10	0.1	10.71	11.67	462	472
11	0.3	11.83	12.63	475	484
12	0.4	12.86	13.48	487	495
13	0.6	13.84	14.27	498	504
14	0.7	14.79	15.02	509	514
15	0.8	15.73	15.75	519	523
16	1.0	16.67	16.47	530	531
17	1.1	17.60	17.16	540	540
18	1.2	18.52	17.85	551	548
19	1.3	19.43	18.54	561	557
20	1.4	20.34	19.23	571	565

ตาราง 5 (ต่อ)

คะแนน	θ	x	y	CEEB _x	CEEB _y
21	1.5	21.23	19.95	582	574
22	1.6	22.13	20.70	592	583
23	1.8	23.03	21.50	602	593
24	2.0	23.94	22.37	612	604
25	2.3	24.88	23.33	623	616
26	2.6	25.86	24.47	634	630
27	3.4	26.89	25.89	646	647
28	4.0	28.00	28.00	658	673

จากตาราง 5 แสดงให้เห็นว่า ที่ระดับความสามารถตั้งแต่ -4.0 ถึง 4.0 คะแนนดิบของแบบทดสอบฉบับ X มีค่าอยู่ระหว่าง 3.81 ถึง 28.00 และเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้วพบว่า มีค่าอยู่ระหว่าง 384 ถึง 658 ส่วนคะแนนดิบของแบบทดสอบฉบับ Y มีค่าอยู่ระหว่าง 4.65 ถึง 28.00 และเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้วพบว่า มีค่าอยู่ระหว่าง 386 ถึง 673

3. คะแนนที่ได้จากการเทียบคะแนน

3.1 จากคะแนนสมมูลที่ได้จากการสร้างตารางเทียบคะแนนทั้งสองวิธี ผู้วิจัยได้นำเสนอในระดับคะแนนและความสามารถเดียวกันดังตาราง 6

ตาราง 6 คะแนนสมมูลของการเทียบคะแนนทั้งสองวิธีที่ระดับคะแนนและ
ความสามารถเดียวกัน

คะแนน	θ	การเทียบคะแนนจริง		การเทียบคะแนนดิบ	
		x	CEEB	x	CEEB
5	-4.0	6.45	390	4.65	386
6	-1.9	7.69	406	6.63	406
7	-1.0	9.07	425	7.81	425
8	-0.5	10.37	442	9.27	443
9	-0.2	11.37	455	10.56	459
10	0.1	12.55	471	11.67	472
11	0.3	13.46	483	12.63	484
12	0.4	13.94	490	13.48	495
13	0.6	14.98	504	14.27	504
14	0.7	15.53	511	15.02	514
15	0.8	16.10	519	15.75	523
16	1.0	17.28	534	16.47	531
17	1.1	17.88	542	17.16	540
18	1.2	17.47	550	17.85	548
19	1.3	19.06	558	18.54	557
20	1.4	19.64	566	19.23	565
21	1.5	20.19	573	19.95	574
22	1.6	20.73	580	20.70	583
23	1.8	21.71	593	21.50	593
24	2.0	22.57	605	22.37	604
25	2.3	23.65	619	23.33	616
26	2.6	24.51	631	24.47	630
27	3.4	26.07	652	25.89	647
28	4.0	26.76	661	28.00	673

จากตาราง 6 แสดงให้เห็นว่าคะแนนสมมูลที่ได้จากการเทียบคะแนน จากแบบทดสอบฉบับ X ไปสู่แบบทดสอบ ฉบับ Y ในช่วงคะแนน 5 - 11 วิธีการเทียบคะแนนจริงให้คะแนนสมมูลสูงกว่าวิธีการเทียบคะแนนดิบ และช่วงคะแนน 12 - 28 คะแนน ทั้งสองวิธีมีคะแนนสมมูลที่ใกล้เคียงกัน

3.2 การเปรียบเทียบความแตกต่างของคะแนนที่ได้จากการเทียบคะแนนทั้งสองวิธี

จากคะแนนที่ได้จากการเทียบทั้งสองวิธี สามารถนำมาทดสอบความแตกต่างของคะแนนสมมูล ดังแสดงในตาราง 7

ตาราง 7 การเปรียบเทียบความแตกต่างของคะแนนที่ได้จากการเทียบคะแนนทั้งสองวิธี

วิธีการเทียบคะแนน	K	\bar{D}	$S_{\bar{D}}$	t-test
คะแนนจริง	24	-0.4929	0.1213	-4.06**
คะแนนดิบ	24			

** มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง 7 แสดงให้เห็นว่า คะแนนที่ได้จากการเทียบคะแนนทั้งสองวิธีแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 นั่นก็คือ คะแนนที่ได้จากการเทียบคะแนนทั้งสองวิธีมีความแตกต่างกัน และการนำผลการเทียบคะแนนไปใช้ในเชิงปฏิบัติ นั้น จะใช้ในรูปแบบคะแนนมาตรฐาน ดังนั้นจึงแปลงเป็นคะแนนที่ได้จากการเทียบคะแนนเป็นคะแนนมาตรฐาน (CEEB) ผู้วิจัยจึงนำคะแนนมาตรฐานที่แปลงได้ทั้งสองวิธีมาเปรียบเทียบความแตกต่างของคะแนนดังแสดงในตาราง 8 ดังนี้

ตาราง 8 การเปรียบเทียบความแตกต่างของคะแนนที่ได้จากการเทียบคะแนนทั้งสองวิธี
เมื่อแปลงเป็นคะแนนมาตรฐาน CEEB

วิธีการเทียบคะแนน	K	\bar{D}	$S_{\bar{D}}$	t-test
คะแนนจริง	24	-0.50	0.7223	-0.69220
คะแนนดิบ	24			

จากตาราง 8 แสดงให้เห็นว่า คะแนนที่ได้จากการเทียบคะแนนทั้งสองวิธีแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นก็คือ คะแนนที่ได้จากการเทียบคะแนนทั้งสองวิธีไม่แตกต่างกัน

4. ความคลาดเคลื่อนของการเทียบคะแนน

4.1 จากคะแนนที่ได้จากการเทียบคะแนน เมื่อนำมาเปรียบเทียบกับคะแนนที่ได้จากการสอบ แล้วนำมาคำนวณหาค่าความคลาดเคลื่อนในแต่ละระดับคะแนนและความสามารถเดียวกันดังตาราง 9

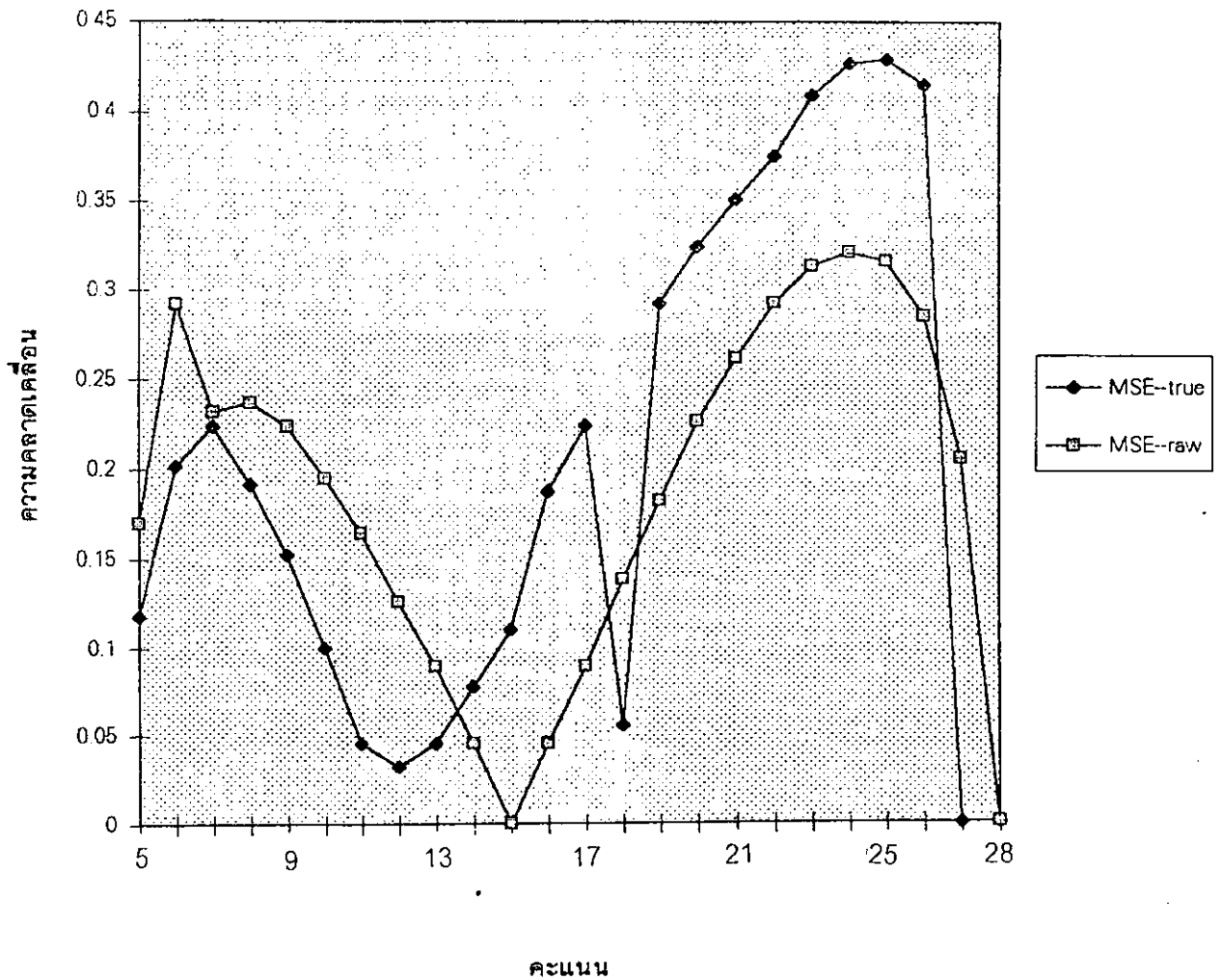
ตาราง 9 ความคลาดเคลื่อนของการเทียบคะแนนของทั้งสองวิธีที่ระดับความสามารถเดียวกัน

คะแนน	θ	MSE ของคะแนนจริง	MSE ของคะแนนดิบ
5	-4.0	0.188	0.170
6	-1.9	0.202	0.292
7	-1.0	0.224	0.232
8	-0.5	0.192	0.237
9	-0.2	0.152	0.224
10	0.1	0.100	0.195
11	0.3	0.045	0.104

ตาราง 9 (ต่อ)

คะแนน	θ	MSE ของคะแนนจริง	MSE ของคะแนนดิบ
12	0.4	0.032	0.126
13	0.6	0.045	0.089
14	0.7	0.077	0.045
15	0.8	0.110	0.000
16	1.0	0.187	0.045
17	1.1	0.224	0.089
18	1.2	0.055	0.138
19	1.3	0.292	0.182
20	1.4	0.324	0.226
21	1.5	0.351	0.261
22	1.6	0.375	0.292
23	1.8	0.409	0.313
24	2.0	0.427	0.321
25	2.3	0.429	0.316
26	2.6	0.415	0.285
27	3.4	0.000	0.205
28	4.0	0.000	0.000
\bar{X}		0.199375	0.18529
SD		0.146270	0.10040

จากตาราง 9 แสดงให้เห็นว่า ค่าความคลาดเคลื่อนของการเทียบคะแนนจริง มีค่าอยู่ระหว่าง 0.000 ถึง 0.429 ส่วนวิธีการเทียบคะแนนดิบ มีค่าอยู่ระหว่าง 0.000 ถึง 0.316 และเมื่อพิจารณาค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐาน พบว่า วิธีการเทียบคะแนนจริงมีการกระจายคะแนนที่มากกว่าวิธีการเทียบคะแนนดิบ และเพื่อให้เห็นภาพที่ชัดเจนผู้วิจัยได้นำค่าความคลาดเคลื่อนที่ได้จากการเทียบคะแนนไปเขียนกราฟได้ดังภาพประกอบ 6



ภาพประกอบ 6 ความคลาดเคลื่อนของการเทียบคะแนนทั้งสองวิธีที่ระดับคะแนนและความสามารถเดียวกัน

จากภาพประกอบ 6 แสดงให้เห็นว่า ที่ระดับคะแนนตั้งแต่ 5 ถึง 13 วิธีการเทียบคะแนนดิบให้ค่าความคลาดเคลื่อนที่สูงกว่าวิธีการเทียบคะแนนจริง ส่วนระดับคะแนน 14 ถึง 17 วิธีการเทียบคะแนนจริงให้ค่าความคลาดเคลื่อนที่สูงกว่าวิธีการเทียบคะแนนดิบ และระดับคะแนนตั้งแต่ 18 ถึง 28 วิธีการเทียบคะแนนจริงให้ค่าความคลาดเคลื่อนที่สูงกว่าวิธีการเทียบคะแนนดิบ

4.2 การเปรียบเทียบความคลาดเคลื่อนที่ได้จากการเทียบคะแนนทั้งสองวิธี

จากความคลาดเคลื่อนที่ได้จากการเทียบคะแนนทั้งสองวิธี สามารถนำมาทดสอบความแตกต่างของคะแนนความคลาดเคลื่อน ดังแสดงในตาราง 10

ตาราง 10 การเปรียบเทียบความแตกต่างของคะแนนความคลาดเคลื่อนที่ได้จากการเทียบคะแนนทั้งสองวิธี

วิธีการเทียบคะแนน	K	\bar{D}	$S_{\bar{D}}$	t-test
คะแนนจริง	24	0.009375	0.020767	0.451437
คะแนนดิบ	24			

จากตาราง 10 แสดงให้เห็นว่า คะแนนความคลาดเคลื่อนที่ได้จากการเทียบคะแนนทั้งสองวิธี แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นก็คือ คะแนนความคลาดเคลื่อนที่ได้จากการเทียบคะแนนทั้งสองวิธีไม่แตกต่างกัน

5. ความเพียงพอของการเทียบคะแนน

5.1 ความเพียงพอของการเทียบคะแนนทั้งสองวิธี ในการหาความเพียงพอของการเทียบคะแนน ผู้วิจัยได้นำแบบทดสอบทั้งสองฉบับ ไปทดสอบกับกลุ่มตัวอย่างที่ใช้ในการทานผลการเทียบคะแนน แล้วมาคำนวณหาความคลาดเคลื่อนที่ได้จากการเทียบคะแนน เพื่อที่จะสามารถนำไปประเมินความเพียงพอของการเทียบคะแนน ดังแสดงในตาราง 11

ตาราง 11 ความเพียงพอของการเทียบคะแนนทั้งสองวิธีที่ระดับความสามารถเดียวกัน

คะแนน	θ	$(x - x^*)^2$ ของคะแนนจริง	$(x - x^*)^2$ ของคะแนนดิบ
5	-4.0	0.152	0.168
6	-1.9	0.422	0.250
7	-1.0	0.476	0.336
8	-0.5	0.152	0.372
9	-0.2	0.336	0.360
10	0.1	0.212	0.302
11	0.3	0.137	0.230
12	0.4	0.040	0.137
13	0.6	0.017	0.068
14	0.7	0.000	0.022
15	0.8	0.022	0.002
16	1.0	0.048	0.006
17	1.1	1.877	0.040
18	1.2	0.202	0.096
19	1.3	0.334	0.176
20	1.4	0.462	0.314
21	1.5	0.548	0.372
22	1.6	0.672	0.436
23	1.8	0.774	0.449
24	2.0	0.884	0.410
25	2.3	0.960	0.325
26	2.6	1.000	0.203

ตาราง 11 (ต่อ)

คะแนน	θ	$(x - x^*)^2$ ของคะแนนจริง	$(x - x^*)^2$ ของคะแนนดิบ
27	3.4	0.884	0.068
28	4.0	0.689	0.123
k		24	24
SD		5.45750	4.10720
c		0.01581	0.00434
$(.05SD)^2$		0.07450	0.12630

จากตาราง 11 แสดงให้เห็นว่า ค่าคะแนนความคลาดเคลื่อนของการเทียบคะแนนจริงมีค่าอยู่ระหว่าง -1.00 ถึง 0.69 ส่วนค่าคะแนนความคลาดเคลื่อนของการเทียบคะแนนดิบมีค่าอยู่ระหว่าง -0.58 ถึง 0.67 แต่เมื่อดำเนินหาค่าความเพียงพอของการเทียบคะแนนแล้วพบว่า วิธีการเทียบคะแนนจริงให้ค่าความเพียงพอที่สูงกว่าวิธีการเทียบคะแนนดิบ แต่เมื่อเปรียบเทียบกับเกณฑ์แล้ว ทั้งสองวิธีให้ค่าความเพียงพอของการเทียบคะแนนอยู่ในระดับที่น่าพอใจอย่างยิ่ง

5.2 การเปรียบเทียบความเพียงพอของการเทียบคะแนนทั้งสองวิธี

จากการหาค่าความเพียงพอของการเทียบคะแนนที่ได้จากการเทียบคะแนนทั้งสองวิธี สามารถนำเสนอเปรียบเทียบให้เห็นชัดเจนดังตาราง 12

ตาราง 12 การเปรียบเทียบความเพียงพอของการเทียบคะแนนทั้งสองวิธี

วิธีการเทียบคะแนน	K	$\sum(x-x^*)^2$	s^2	C
คะแนนจริง	24	11.3023	29.7844	0.01581**
คะแนนดิบ	24	5.264	50.5117	0.00434**

$$C < (.05SD)^2 \quad ** \text{ระดับน่าพอใจอย่างยิ่ง}$$

จากตาราง 12 แสดงให้เห็นว่า ความเพียงพอของการเทียบคะแนนของวิธีการเทียบคะแนนจริงให้ค่าสูงกว่าวิธีการเทียบคะแนนดิบเพียงเล็กน้อย แต่เมื่อเปรียบเทียบกับเกณฑ์ของความเพียงพอในการเทียบคะแนนแล้วพบว่า ทั้งสองวิธีอยู่ในเกณฑ์ที่น่าพอใจอย่างยิ่ง

บทที่ 5

สรุปผล อภิปรายผล และข้อเสนอแนะ

ความมุ่งหมายของการศึกษาค้นคว้า

ในการศึกษาครั้งนี้ มีความมุ่งหมายเพื่อศึกษาเปรียบเทียบผลการเทียบคะแนนแบบทดสอบวิชาการพยาบาลสูติศาสตร์จำนวน 2 ฉบับด้วยวิธีการเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วมและวิธีการเทียบคะแนนดิบโดยการใช้แบบทดสอบร่วมตามทฤษฎีการตอบข้อสอบ โดยมีความมุ่งหมายเฉพาะดังนี้

1. เพื่อเปรียบเทียบคะแนนที่เทียบได้ของแบบทดสอบวิชาการพยาบาลสูติศาสตร์ระหว่างวิธีการวิเคราะห์ต่างกัน
2. เพื่อเปรียบเทียบความคลาดเคลื่อนของการเทียบคะแนนระหว่างวิธีการวิเคราะห์ต่างกัน
3. เพื่อเปรียบเทียบความเพียงพอของการเทียบคะแนน ระหว่างวิธีการวิเคราะห์ต่างกัน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาครั้งนี้เป็นนักศึกษาพยาบาลศาสตร์(ระดับต้น) ชั้นปีที่ 2 ปีการศึกษา 2538 ของวิทยาลัยพยาบาลในสังกัดสถาบันพัฒนากำลังคนด้านสาธารณสุข กระทรวงสาธารณสุข ในเขตภาคเหนือ ภาคกลางและภาคตะวันออก จำนวนนักศึกษา 980 คน จำนวน 10 วิทยาลัย ซึ่งได้มาโดยวิธีการสุ่มอย่างง่าย

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาครั้งนี้ เป็นแบบทดสอบเลือกตอบวัดผลสัมฤทธิ์ทางการเรียนวิชาการพยาบาลสูติศาสตร์ 1 (พย.212) เรื่องการพยาบาลในระยะตั้งครรภ์ สำหรับนักศึกษาพยาบาลศาสตร์(ระดับต้น) ชั้นปีที่ 2 ปีการศึกษา 2538 ที่ผู้วิจัยสร้างขึ้นตามประมวลรายวิชาหลักสูตรประโภยบัตรพยาบาลศาสตร์(ระดับต้น) จำนวน 100 ข้อ เป็นแบบทดสอบเลือกตอบชนิด 5 ตัวเลือก และแบ่งข้อสอบออกเป็น 2 ฉบับ โดยการพิจารณาค่าความยาก ค่าอำนาจจำแนกและค่าสัมประสิทธิ์การเดาที่เท่าหรือเทียบเคียงกันจะได้แบบทดสอบ 2 ฉบับ ดังนี้

1. แบบทดสอบวิชาการพยาบาลสูติศาสตร์ 1 ฉบับ X จำนวน 30 ข้อ

2. แบบทดสอบวิชาการพยาบาลสูติศาสตร์ 1 ฉบับ Y จำนวน 30 ข้อ

โดยที่แบบทดสอบในแต่ละฉบับ จะมีแบบทดสอบร่วมชนิดแบบทดสอบร่วม ภายในอยู่ด้วย (Internal Anchor Test) จำนวน 10 ข้อ รวมอยู่ด้วย

วิธีดำเนินการเก็บรวบรวมข้อมูล

ในการเก็บรวบรวมข้อมูลในครั้งนี้ ผู้วิจัยได้ดำเนินการตามขั้นตอนดังนี้

1. วางแผนในการดำเนินการสอบ โดยติดต่อขอความร่วมมือในการนำแบบทดสอบไปทดสอบกับนักศึกษาพยาบาลศาสตร์(ระดับต้น) ชั้นปีที่ 2 จากผู้อำนวยการวิทยาลัยพยาบาลที่เป็นกลุ่มตัวอย่าง กำหนดวัน เวลาและสถานที่สอบ

2. นำแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนที่สร้างขึ้นจำนวน 100 ข้อ ไปทดลองสอบกับนักศึกษาพยาบาลศาสตร์(ระดับต้น) ชั้นปีที่ 2 ปีการศึกษา 2538 จำนวนนักศึกษา 501 คนที่ไม่ใช่กลุ่มตัวอย่าง

3. นำผลการสอบจากข้อ 2. มาตรวจให้คะแนน และเตรียมข้อมูลสำหรับการวิเคราะห์ข้อสอบรายข้อ แล้วตรวจสอบความเป็นมิติเดียวของแบบทดสอบ(Unidimensionality) โดยวิธีการวิเคราะห์องค์ประกอบ (Factor Analysis) ก่อนที่จะนำไปวิเคราะห์ตามทฤษฎีการตอบข้อสอบ โดยการใช้โปรแกรม SPSS/FW และคำนวณหาค่าพารามิเตอร์ต่างๆ โดยการใช้โปรแกรมสำเร็จรูป BILOG เวอร์ชัน 3.04 (Mislevy and Bock. 1982)

4. นำข้อสอบที่มีคุณภาพจากการวิเคราะห์ในข้อ 3. จัดเป็นแบบทดสอบ 2 ฉบับ คือ ฉบับ X และฉบับ Y ฉบับละ 30 ข้อ

5. นำแบบทดสอบทั้งสองฉบับ (ฉบับ X และฉบับ Y) ไปทดสอบกับกลุ่มตัวอย่างจำนวน 980 คน โดยจัดให้ทุกคนทำแบบทดสอบทั้งสองฉบับ แล้วนำมาสุ่ม โดยใช้คะแนนจากแบบทดสอบฉบับ X จำนวน 490 ฉบับและ ฉบับ Y จำนวน 490 ฉบับ ส่วนคะแนนที่เหลือในแต่ละฉบับใช้เป็นคะแนนแทนผลของการเทียบคะแนนในแต่ละวิธี

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลดำเนินการโดยแบ่งออกเป็น 4 ตอนได้แก่

1. ค่าสถิติพื้นฐานของแบบทดสอบ

2. การสร้างตารางเทียบคะแนนแบบทดสอบ

3. คะแนนที่ได้จากการเทียบคะแนน
4. ความคลาดเคลื่อนของการเทียบคะแนน
5. ความเพียงพอของการเทียบคะแนน

สรุปผลการวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐานของแบบทดสอบ

ค่าสถิติพื้นฐานของแบบทดสอบ พบว่าแบบทดสอบทั้งสองฉบับมีคะแนนเฉลี่ยที่ต่ำกว่า 14 คะแนน แสดงว่าแบบทดสอบทั้งสองฉบับค่อนข้างยาก และแบบทดสอบฉบับ X มีคะแนนเฉลี่ยประมาณครึ่งหนึ่งของคะแนนเต็ม แสดงว่ามีความยากอยู่ในระดับปานกลาง แบบทดสอบฉบับ Y มีคะแนนเฉลี่ยต่ำกว่า 14 คะแนน แสดงว่าค่อนข้างยาก จากคะแนนเฉลี่ยของแบบทดสอบทั้งสองฉบับ แสดงให้เห็นถึงคะแนนเฉลี่ยที่ไม่เท่ากันเนื่องมาจากความยากของแบบทดสอบต่างกัน ดังนั้นจึงมีความจำเป็นในการที่จะปรับเทียบคะแนนแบบทดสอบทั้งสองฉบับให้อยู่ในสเกลเดียวกัน ส่วนค่าความเบี่ยงเบนมาตรฐาน พบว่าแบบทดสอบฉบับ Y มีการกระจายของคะแนนมากกว่าแบบทดสอบฉบับ X

2. การสร้างตารางเทียบคะแนนแบบทดสอบ

2.1 การเทียบคะแนนจริงโดยการใช้แบบทดสอบร่วม พบว่าที่ระดับความสามารถ ตั้งแต่ -4.0 ถึง 4.0 คะแนนจริงของแบบทดสอบฉบับ X มีค่าอยู่ระหว่าง 7.02 ถึง 26.76 และเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้วพบว่า มีค่าอยู่ระหว่าง 382 ถึง 673 ส่วนคะแนนจริงของแบบทดสอบฉบับ Y มีค่าอยู่ระหว่าง 6.45 ถึง 26.76 และเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้วพบว่า มีค่าอยู่ระหว่าง 390 ถึง 661

2.2 การเทียบคะแนนดิบโดยการใช้แบบทดสอบร่วม พบว่าที่ระดับความสามารถ ตั้งแต่ -4.0 ถึง 4.0 คะแนนดิบของแบบทดสอบฉบับ X มีค่าอยู่ระหว่าง 3.81 ถึง 28.00 และเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้วพบว่า มีค่าอยู่ระหว่าง 384 ถึง 658 ส่วนคะแนนดิบของแบบทดสอบฉบับ Y มีค่าอยู่ระหว่าง 4.65 ถึง 28.00 และเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้วพบว่า มีค่าอยู่ระหว่าง 386 ถึง 673

3. คะแนนที่ได้จากการเทียบคะแนน

3.1 คะแนนสมมูลที่ได้จากการเทียบคะแนนจากแบบทดสอบฉบับ X ไปสู่แบบทดสอบฉบับ Y พบว่า ในช่วง 5 - 11 คะแนน วิธีการเทียบคะแนนจริงให้คะแนนสมมูล

สูงกว่าวิธีการเทียบคะแนนดิบ และช่วงคะแนน 12 - 28 คะแนน ทั้งสองวิธีมีคะแนนสมมูลที่ใกล้เคียงกัน

3.2 การเปรียบเทียบความแตกต่างของคะแนนที่ได้จากการเทียบ พบว่าวิธีการเทียบคะแนนจริงกับวิธีการเทียบคะแนนดิบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 นั่นก็คือคะแนนที่ได้จากการเทียบคะแนนทั้งสองวิธีมีความแตกต่างกัน แต่เมื่อเปรียบเทียบความแตกต่างของคะแนนที่ได้จากการเทียบโดยใช้คะแนนที่แปลงเป็นคะแนนมาตรฐาน CEEB แล้วพบว่า คะแนนที่ได้จากการเทียบทั้งสองวิธีแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นก็คือคะแนนที่ได้จากการเทียบคะแนนทั้งสองวิธีไม่แตกต่างกัน

4. ความคลาดเคลื่อนของการเทียบคะแนน

4.1 ความคลาดเคลื่อนของการเทียบคะแนน พบว่าวิธีการเทียบคะแนนจริงมีค่าคะแนนความคลาดเคลื่อนอยู่ระหว่าง 0.000 ถึง 0.316 และเมื่อพิจารณาค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐาน พบว่าวิธีการเทียบคะแนนจริงมีการกระจายของคะแนนที่มากกว่าวิธีการเทียบคะแนนดิบ

4.2 การเปรียบเทียบความแตกต่างของคะแนนความคลาดเคลื่อนที่ได้จากการเทียบคะแนน พบว่าวิธีการเทียบคะแนนจริงกับวิธีการเทียบคะแนนดิบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นก็คือคะแนนความคลาดเคลื่อนที่ได้จากการเทียบคะแนนทั้งสองวิธีไม่แตกต่างกัน

5. ความเพียงพอของการเทียบคะแนน

ความเพียงพอของการเทียบคะแนน พบว่าพบว่าวิธีการเทียบคะแนนจริงมีค่าความเพียงพอของการเทียบคะแนนที่สูงกว่าวิธีการเทียบคะแนนดิบ แต่เมื่อเปรียบเทียบกับเกณฑ์ของการประเมินความเพียงพอแล้วพบว่า ทั้งสองวิธีอยู่ในเกณฑ์ที่น่าพอใจอย่างยิ่ง

อภิปรายผล

1. การเปรียบเทียบความแตกต่างของคะแนนที่ได้จากการเทียบตามวิธีการเทียบคะแนนจริงกับวิธีการเทียบคะแนนดิบมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ ทั้งนี้อาจเป็นเพราะว่า คะแนนที่ได้จากการเทียบทั้งสองวิธีมาจากขั้นตอนการคิดคะแนนที่ซับซ้อนกว่ากัน โดยที่วิธีการเทียบคะแนนจริงมีการคำนวณคะแนนที่ได้ค่อนข้างละเอียดน้อยกว่าวิธีการเทียบคะแนนดิบ ดังนั้นคะแนนที่ได้จากการเทียบทั้งสองวิธีจึงมีความแตกต่างกัน แต่เมื่อเปรียบเทียบความแตกต่างของคะแนนที่ได้จากการ

เทียบเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้วพบว่า วิธีการเทียบคะแนนทั้งสองวิธีมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ ทั้งนี้อาจเป็นเพราะว่า การที่แปลงคะแนนที่ได้จากการเทียบไปสู่คะแนนมาตรฐาน CEEB นั้นเป็นการปรับคะแนนที่มีอยู่เดิมให้เป็นคะแนนมาตรฐานที่เข้าสูการกระจายเป็นโค้งปกติ จึงทำให้ช่วงการกระจายของคะแนนที่ได้จากการเทียบทั้งสองวิธีมีค่าเท่ากันและคะแนนมาตรฐาน CEEB ที่ได้จากการเทียบทั้งสองวิธีจะมีค่าเข้าสู่คะแนนเฉลี่ยของคะแนนมาตรฐาน CEEB ซึ่งมีค่าเข้าใกล้ 500 คะแนน ดังนั้นคะแนนมาตรฐานที่ได้จากการเทียบคะแนนทั้งสองวิธีจึงมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. การเปรียบเทียบความแตกต่างของคะแนนความคลาดเคลื่อนที่ได้จากการเทียบตามวิธีการเทียบคะแนนจริงกับวิธีการเทียบคะแนนดิบมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ เนื่องจากความคลาดเคลื่อนที่ได้จากการเทียบคะแนนทั้งสองวิธีไม่แตกต่างกัน ทั้งนี้อาจเป็นเพราะว่า ทั้งสองวิธีเป็นการเทียบคะแนนตามทฤษฎีการตอบข้อสอบ โดยที่วิธีการเทียบคะแนนจริง คะแนนที่ได้มาจากการคำนวณหาผลรวมของความน่าจะเป็นในการตอบข้อสอบถูกของแบบสอบ (Hambleton and Swaminathan, 1982 : 212) ซึ่งมาจากการประมาณค่าพารามิเตอร์ของข้อสอบ ส่วนวิธีการเทียบคะแนนดิบนั้นคะแนนที่ได้มาจากการประมาณค่าคะแนนโอกาสในการตอบข้อสอบถูกกับคะแนนที่คาดหวังซึ่งมีค่าตั้งแต่คะแนนศูนย์ถึงคะแนนเต็ม จะเห็นได้ว่าวิธีการคำนวณคะแนนที่ได้ทั้งสองวิธีเป็นวิธีการที่ค่อนข้างละเอียด คะแนนความคลาดเคลื่อนที่ได้จากการเทียบคะแนนจึงมีค่าไม่แตกต่างจากคะแนนเดิมมากนัก ดังนั้นเมื่อนำมาทดสอบสมมติฐานจึงไม่มีความแตกต่างกัน และทั้งสองวิธีดังกล่าวต่างก็เป็นวิธีการเทียบคะแนนตามทฤษฎีการตอบข้อสอบ ที่จะให้ค่าความคลาดเคลื่อนของการเทียบคะแนนน้อยกว่าวิธีอื่น ซึ่งสอดคล้องกับงานวิจัยของโคเลน (Kolen, 1981 : 1-11) และเวรตี อินทะสระระ (2530)

3. การเปรียบเทียบความแตกต่างของความเพียงพอที่ได้จากการเทียบคะแนนตามวิธีการเทียบคะแนนจริงกับวิธีการเทียบคะแนนดิบมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ ทั้งนี้อาจเป็นเพราะว่า ในการประเมินความเพียงพอนั้น การเทียบคะแนนตามทฤษฎีการตอบข้อสอบจะให้ค่าความเพียงพอของการเทียบคะแนนอยู่ในระดับที่น่าพอใจ แต่จากการศึกษาครั้งนี้พบว่า ความเพียงพอของการเทียบคะแนนอยู่ในเกณฑ์ที่น่าพอใจอย่างยิ่ง ซึ่งก็เป็นไปตามทฤษฎีการตอบข้อสอบที่กล่าวไว้ และสอดคล้องคล้อยกับการศึกษาของลอร์ด (Lord, 1980 : 203) โคเลนและวิทนีย์ (Kolen and

Whitney. 1982 : 279 -239) และเวดี อินทะสระ (2530) ที่ศึกษาการเทียบคะแนนตาม ทฤษฎีการตอบข้อสอบที่ให้ค่าความเพียงพอของการเทียบคะแนนอยู่ในระดับที่น่าพอใจ

จะเห็นได้ว่า วิธีการเทียบคะแนนทั้งสองวิธีให้ผลการเทียบคะแนนความคลาดเคลื่อน และความเพียงพอของคะแนนที่ได้จากการเทียบไม่แตกต่างกัน ดังนั้นในการที่จะนำผลการ ศึกษาครั้งนี้ไปใช้ ควรใช้วิธีการเทียบคะแนนจริง เพราะมีขั้นตอนในการคิดคำนวณที่ซับซ้อน น้อยกว่าวิธีการเทียบคะแนนดิบ

ข้อเสนอแนะ

จากการศึกษาค้นคว้าครั้งนี้ ทำให้ผู้วิจัยได้แนวคิดที่จะเสนอแนะดังนี้

1. จากผลการเทียบคะแนนทั้งสองวิธี พบว่าไม่แตกต่างกัน ดังนั้นในการที่จะนำผล การศึกษาครั้งนี้ไปใช้อ้างอิงต้องทำอย่างพิถีพิถัน เพราะการเทียบคะแนนทั้งสองวิธีมีข้อได้ เปรียบเสียเปรียบกันคือ วิธีการเทียบคะแนนจริงไม่สามารถอธิบายคะแนนที่อยู่ต่ำกว่าระดับ การเดาได้ โดยจะให้ความหมายของคะแนนสมมูลเฉพาะคะแนนที่อยู่เหนือค่าเฉลี่ยของ การเดา ส่วนวิธีการเทียบคะแนนดิบ เป็นการเทียบคะแนนโดยประมาณ แต่สามารถอธิบาย คะแนนสมมูลจากแบบทดสอบทั้งสองฉบับ ได้ครอบคลุมพิสัยของคะแนนดิบ อย่างไรก็ตาม วิธีการเทียบคะแนนดิบนี้นำไปใช้ในการวิจัยน้อยมาก ซึ่งอาจเป็นเพราะว่า วิธีการนี้มีความ ซับซ้อนและลงทุนสูงกว่าวิธีการเทียบคะแนนจริง (Lord. 1980 : 202 - 203 ; Lord and Wingersky. 1984 : 453)

2. เนื่องจากการวิเคราะห์ข้อสอบโดยการใช้โมเดลโลจิสติกแบบสามพารามิเตอร์ ไม่มีการวิเคราะห์ตัวเลือกและตัวลวงของข้อสอบ ดังนั้นเพื่อให้ได้ข้อสอบที่ดีควรจะวิเคราะห์ ตามทฤษฎีคลาสสิก(The Classical Test Theory) เพื่อหาคุณภาพตัวเลือกก่อนที่จะนำไป ประมาณค่าพารามิเตอร์ของข้อสอบที่จะใช้ในงานวิจัยต่อไป

3. ในการศึกษาผลการเทียบคะแนนของแบบทดสอบนั้นขึ้นอยู่กับองค์ประกอบที่ สำคัญคือ วิธีการเทียบคะแนน(Method of Equating) และแบบแผนการเก็บรวบรวม ข้อมูล(Design of Equating) โดยเฉพาะแบบแผนการเก็บรวบรวมข้อมูล ซึ่งในการวิจัยครั้งนี้ เก็บข้อมูลในสถานการณ์จริง ทำให้ผู้วิจัยไม่สามารถควบคุมตัวแปรแทรกซ้อนที่เกิดจากอิทธิพล ของการสอบได้คือ ช่วงเวลาในการเรียนภาคทฤษฎีแตกต่างกันตามลักษณะภูมิภาคของแต่ละ วิทยาลัย ครูผู้สอนมีประสบการณ์ที่แตกต่างกัน จึงทำให้กลุ่มตัวอย่างตอบแบบทดสอบไม่เป็น

ที่น่าสนใจ ดังนั้นการศึกษามูลการเทียบคะแนนที่ดี ควรจะศึกษาในสถานการณ์จำลอง(Simulation)

4. ควรศึกษามูลการเทียบคะแนนโดยการเปรียบเทียบกับแบบแผนการเก็บรวบรวมข้อมูลที่แตกต่างกัน เพื่อที่จะดูว่าให้ผลการเทียบคะแนนที่แตกต่างกันหรือไม่

5. ควรศึกษาการเทียบคะแนนตามทฤษฎีการตอบข้อสอบโดยการเปรียบเทียบกับวิธีการ Pre-equate

6. ควรนำวิธีการเทียบคะแนนไปใช้ในจุดมุ่งหมายอื่นๆ เช่นโครงการคัดเลือกบุคคลเข้าทำงาน โครงการทดสอบเพื่อปรับตำแหน่งในการปฏิบัติงาน เป็นต้น ทั้งนี้เพื่อให้มาตรฐานในการตัดสินผลการสอบและก่อให้เกิดความยุติธรรมแก่ผู้สอบมากขึ้น

บรรณานุกรม

บรรณานุกรม

- จินตนา ยูนิพันธ์. การเรียนการสอนทางพยาบาลศาสตร์. ภาควิชาพยาบาลศึกษา คณะครุศาสตร์จุฬาลงกรณ์มหาวิทยาลัย, 2527.
- ชูชีพ พงษ์สมบุรณ์. การเปรียบเทียบการเทียบมาตรฐานระหว่างรูปแบบที่ใช้ผู้สอบร่วมกับรูปแบบที่ใช้แบบสอบร่วม. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2528. ถ่ายเอกสาร.
- ชูศักดิ์ ชัมภลิจิต. Test equating procedures. เอกสารประกอบการบรรยายในการสัมมนา ระดับชาติครั้งที่ 4 โครงการพัฒนาอาเซียน, 2527. 13 หน้า.
- ภาวินี ศรีสุขวัฒน์นนท์. การเปรียบเทียบผลจากการใช้รูปแบบการเทียบมาตรฐานที่ต่างกัน เมื่อแบบสอบร่วมมีความยาวต่างกัน. วิทยานิพนธ์ ค.ด. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2529. ถ่ายเอกสาร.
- เรวดี อินทะสระระ. การเปรียบเทียบความคลาดเคลื่อนของการเทียบมาตรฐานระหว่างรูปแบบอิงทฤษฎีการตอบข้อสอบกับรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ. ปริญญาโท กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2530. อัดสำเนา.
- วรรณดี แสงประทีปทอง. การศึกษาคำวามก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้นมัธยมศึกษาตอนต้น โดยการเทียบคะแนนตามแนวตั้งตามทฤษฎีการตอบข้อสอบ แบบจำลองโลจิสติก สามพารามิเตอร์. ปริญญาโท กศ.ด. กรุงเทพฯ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2538. อัดสำเนา.
- ล้วน และอังคณา สายยศ. เทคนิคการวิจัยทางการศึกษา. พิมพ์ครั้งที่ 3. สำนักพิมพ์ศูนย์ส่งเสริมวิชาการ คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2536.
- สงบ ลักษณะ. "การเทียบระดับคะแนนระหว่างแบบทดสอบ," วารสารวัดผลการศึกษา. 4(2) : 21-32 ; 2525.
- เสริมศักดิ์ วิชาลาภรณ์ และเอนกกุล กรีแสง. หลักเบื้องต้นของการวัดผลการศึกษา. มหาวิทยาลัยศรีนครินทรวิโรฒ พิษณุโลก, 2517.

- Angoff, W.H. "Scale, Norm and Equivalent Score," In R. L. Thorndike(Ed.).
Educational Measurement. 2nd ed. Washington : D.C. American
Council on Educational, 508-600 ; 1971.
- Flanagan, T.C. "Unit , Score and Norm ," In E.F. Lindquist(Ed.) Educational
Measurement. 2nd ed. Washington : D.C. American Council on
Educational, 695 - 673 ; 1951.
- Gulliksen , Harold. Theory of Mental Tests. New York : John Wiley and Sons ,
Inc . , 1950.
- Hambleton, R.K. and H. Swaminathan. Item Response Theory. Kluwer. Nijhoff,
Boston, 312 ; 1985.
- Harris, D.J. and H.D.Hoover "An Application of The Three-Parameter IRT Model
to Vertical Equating," Journal of Applied Psychological Measurement.
11(2) : 151 - 159 ; June, 1987.
- Kolen, M.J. "Comparison of Traditional and Item Response Theory Methods for
Equating Tests," Journal of Educational Measurement. 18(1) ; 1-11
Spring, 1981.
- Kolen, M.J. and D.R. Whitney "Comparison of Four Procedure for Equating The
Test of General Educational Development ," Journal of Educational
Measurement. 19 : 279-294 ; 1982.

- Lord, F.M. Application of Item Response Theory to Practical Testing Problem. Hillsdale, N.J. : Erlbaum , 1980.
- Lord, F.M. "Practical Applications of Item Characteristic Curve Theory," Journal of Educational Measurement. 14(2) : 117-138 ; 1977.
- Lord, F.M. "The Standard Error of Equipercentile Equating ," Journal of Educational Statistic. 14(2) : 117-138 ; 1977.
- Lord, F.M. "Item Response Theory and Equating-A Technical Summary," In P.W. Holland and D.B. Rubin (Eds). Test equating. New York : Academic press . 1982.
- Lord, F.M. and M.R. Novick Statistical Theory of Mental Test Score. MA : Annison - Wesley, 1968.
- Lord, F.M. and M.S. Wingersky. "Comparison of IRT True Score and Equipercentile Observed-score Equating," Journal of Applied Psychological Measurement. 4(8) : 453-461 ; 1984.
- Marco, Gary L. "Item Characteristic Curve Solution to Three Intractable Testing Problems," Journal of Educational Measurement . 14 : 139-160 ; 1977.
- Petersen, N.S., L. Cook and L. Stocking IRT Versus Conventional Equating Method : A Comparative Study of Scale Stability. Los Angeles, 1981.

- Petersen, N.S. and others "A Test of Adequacy of Linear Score Equating Method," In P.W. Holland and D.B. Rubin (Eds). Test equating. New York : Academic press. 1982.
- Petersen, N.S., M.J. Kolen and H.D. Hoover "Scaling, Norming, and Equating," in R.L. Linn (Ed.) Educational Measurement. p.221- 262 3rd. New York : Macmillan, 1989.
- Rubin, B.Donald. and D.W. Holland Test Equating. Princeton, New Jersey : Education Testing Serving , 1982.
- Slinde, J.A. and R.L. Linn "Vertical Equating Test : Fact or Phantom," Journal of Educational Measurement. 14 : 23-35 ; Spring ; 1978.
- Slinde, J.A. and R.L. Linn "The Rash Model , Objective Measurement Equating and Robustness," Journal of Applied Psychological Measurement. 3 : 437-452 ; 1979.
- Wright , B.D. and M.H. Stone. Best Test Design. Chicago : Mesa Press , 1979.
- Warm , Thomas A. A Primer of Item Response Theory. 1978.

ภาคผนวก

ภาคผนวก ก

- ตารางวิเคราะห์หลักสูตรวิชาการพยาบาลศาสตร 1
- แบบทดสอบวิชาการพยาบาลศาสตร 1 ฉบับ X
- แบบทดสอบวิชาการพยาบาลศาสตร 1 ฉบับ Y

ตารางวิเคราะห์หลักสูตร

วิชาการพยาบาลสูติศาสตร์ 1 ชั้นปีที่ 2 หลักสูตรประกาศนียบัตรพยาบาลศาสตร์(ระดับต้น)

เนื้อหา พฤติกรรม	ความ จำ	ความ เข้าใจ	นำไปใช้	วิเคราะห์	สังเคราะห์	ประเมิน ค่า	รวม
1 การดำเนินการตั้งครรภ์ และการวินิจฉัย	10	6	4	-	-	-	20
2 อาการและการเปลี่ยนแปลง ของร่างกายในระยะตั้ง ครรภ์	10	4	4	2	-	-	20
3 การดูแลหญิงในระยะตั้ง ครรภ์	10	8	6	4	2	2	32
4 การให้สุขศึกษาแก่หญิง ตั้งครรภ์และครอบครัว	2	2	4	2	-	-	10
5 การประเมินปัญหาที่พบ ในหญิงตั้งครรภ์ ด้านร่าง กาย จิตใจ อารมณ์ และ สังคม	4	2	4	4	2	2	18
รวม	36	22	26	12	4	4	100

ข้อสอบวิชาการพยาบาลสูติศาสตร์ (ฉบับ X)

คำชี้แจง

1. แบบทดสอบฉบับนี้มีคำถามทั้งหมด 30 ข้อ และมีเวลาให้ทำ 30 นาที
2. คำถามทั้งหมดเป็นแบบทดสอบปรนัยชนิดเลือกตอบทั้งสิ้น คือคำถามแต่ละข้อ จะให้นักศึกษาเลือกตอบข้อที่เห็นว่าถูกต้องที่สุด ดีที่สุด หรือเหมาะสมที่สุด เพียงคำตอบเดียวจาก ก. , ข. , ค. , ง. และ จ. ที่ให้ไว้ เมื่อนักศึกษาเลือกคำตอบใดให้ทำเครื่องหมาย ทับให้ตรงตัวเลือกที่ต้องการ ดังตัวอย่างการตอบดังต่อไปนี้

1.) ก. ข. ค. ง. จ.

ถ้านักศึกษาขีดตอบไปแล้วแต่ต้องการเปลี่ยนคำตอบใหม่ ให้ทำเครื่องหมายขีดทับข้อที่ไม่ต้องการให้ชัดเจนทุกครั้ง แล้วค่อยขีดคำตอบใหม่ ดังนี้

1.) ก. ข. ค. ง. จ.

3. จงระวังขีดคำตอบให้ตรงกับข้อคำถามเสมอ
4. แบบทดสอบมีทั้งหมด 4 หน้า
5. อย่าขีดเขียนหรือทำเครื่องหมายใดๆ ในแบบทดสอบนี้ เพราะข้อสอบนี้จะต้องนำไปใช้อีก

1. Amnion ต่างกับ Chorion อย่างไร
 - ก. นานกว่า ใส เรียบ
 - ข. ไม่นาน ฆาดง่าย ไม่เรียบ
 - ค. เป็นมัน ฆาดง่าย ไม่เรียบ
 - ง. เป็นมัน เหนียว สีขาวขุ่น
 - จ. นานกว่า ไม่นาน ไม่เรียบ

2. นางสมศรีมาฝากครรภ์ ให้ประวัติว่า จำ L.M.P. ไม่ได้แต่หน้าท้องโตขึ้น คลำ Outline ของทารกได้ ท่านคิดว่านางสมศรีมีอาการของอะไร
 - ก. Probable sign
 - ข. Positive sign
 - ค. Presumptive sign
 - ง. Palpation of the fetal outline
 - จ. Enlargement of the abdomen

3. Gonadotropin Hormone จะพบในปัสสาวะเมื่อตั้งครรภ์ได้ที่สัปดาห์
 - ก. 4 - 6 สัปดาห์
 - ข. 2 - 6 สัปดาห์
 - ค. 2 - 10 สัปดาห์
 - ง. 4 - 8 สัปดาห์
 - จ. 4 - 10 สัปดาห์

4. นางดอกจัน อายุครรภ์ได้ 12 สัปดาห์ รับประทานอาหารไม่ค่อยได้ รู้สึกเบื่ออาหาร และเมื่อได้กลิ่นอาหารแปลกๆ ก็จะมีอาการคลื่นไส้ อาเจียน บางครั้งรับประทานอาหารก็เกิดอาการขย้อนและแสบร้อนในอก ท่านจะให้การรักษาช่วยเหลืออย่างไร
 - ก. ให้รับประทานอาหารที่ชอบมากๆ
 - ข. ให้รับประทานอาหารเป็นเวลา
 - ค. ให้รับประทานอาหารที่ปรุงสุกใหม่ๆ
 - ง. ให้หลีกเลี่ยงการรับประทานอาหารที่มีรสชาดแปลก
 - จ. ให้รับประทานอาหารครั้งละน้อยๆ แต่บ่อยครั้ง

5. นางทองกวาว ตั้งครรภ์ได้ 6 เดือน มีเส้นดำที่แนวกลางท้อง เห็นได้ชัด รู้สึกวิตกกังวลมาก ท่านจะให้การช่วยเหลืออย่างไร
- ให้ซื้อยาเสตียรอยด์มาทา
 - ให้ใช้น้ำมันมะกอกทาทุกวัน
 - ให้ไปปรึกษาแพทย์เพื่อวินิจฉัยรักษา
 - ให้อาบน้ำและทำความสะอาดร่างกาย
 - ไม่ต้องให้การช่วยเหลือเพราะจะหายไปเองภายหลังคลอด
6. ตรวจครรภ์หญิงรายหนึ่ง พบว่าได้อายุครรภ์ 24 สัปดาห์ วัดยอดมดลูกสูง 1/4 มากกว่าระดับสะดือ เด็กท่า ROA แสดงว่าเกิดภาวะใด
- Hegar 's sign
 - Chadwick 's sign
 - Braxton Hick 's sign
 - Ballotement 's sign
 - Goodell 's sign
7. Posterior Fontanel ประกอบด้วยรอยต่อในข้อใด
- Sagital Suture กับ Coronal Suture
 - Lamboid Suture กับ Coronal Suture
 - Sagital Suture กับ Lamboid Suture
 - Coronal Suture กับ Frontal Suture
 - Sagital Suture กับ Frontal Suture
8. คำว่า presentation หมายถึงอะไร
- ท่าของเด็ก
 - ส่วนนำของเด็ก
 - รูปทรงของเด็ก
 - แนวลำตัวของเด็ก
 - ความยาวของเด็ก

9. คำว่า lie หมายถึงข้อใด

- ก. ท่าของเด็ก
- ข. ส่วนหน้าของเด็ก
- ค. รูปทรงของเด็ก
- ง. แนวลำตัวของเด็ก
- จ. ขนาดของเด็ก

10. จากการตรวจจรรยาเพื่อทราบท่าของเด็ก พบว่า ที่ยอดมดลูกคล้ำได้กลม แข็ง หลังของเด็กอยู่ด้านซ้าย ด้านขวาคล้ำได้เป็นปุ่มๆ แสดงว่าเด็กอยู่ในท่าใด

- ก. RSA.
- ข. LSA.
- ค. RST.
- ง. LOA.
- จ. LOP

11. ถ้ามี Brow เป็น presentation แล้ว Denominator คืออะไร

- ก. Occiput
- ข. Face
- ค. Frontal
- ง. Sacrum
- จ. Scapular

12. จากการซักประวัติของนางดาวราย พบว่า โดยปกติประจำเดือนจะมาสม่ำเสมอทุกวันที่ 10 - 14 ของเดือน เมื่อวันที่ 15 - 19 กันยายน 2538 ประจำเดือนมาแล้วขาดหายไปและมาอีกครั้งวันที่ 10 ธันวาคม 2538 มีปริมาณเลือดชุ่มผ้าอนามัย 1 แผ่นแล้วหยุดหายไป อีก ข้อใดเป็น L.M.P. ของหญิงรายนี้

- ก. 10 กันยายน 2538
- ข. 15 กันยายน 2538
- ค. 17 กันยายน 2538
- ง. 10 ธันวาคม 2538
- จ. 14 ธันวาคม 2538

13. กำหนดคลอดจะเป็นเมื่อใด ถ้า L.M.P. คือวันที่ 26 มีนาคม 2538
- 2 ธันวาคม 2538
 - 25 ธันวาคม 2538
 - 2 มกราคม 2539
 - 26 มกราคม 2539
 - 30 มกราคม 2539
14. วิธีการใดสำคัญที่สุดในการลดปัจจัยต่างๆ ที่มีผลกระทบต่อมารดาและทารกในครรภ์
- ให้ความรู้เรื่องการดูแลสุขภาพระหว่างตั้งครรภ์
 - เยี่ยมบ้านหญิงตั้งครรภ์รายที่มีภาวะเสี่ยงสูงทุกราย
 - หลีกเลี่ยงปัจจัยเสี่ยงที่มีผลต่อมารดาและทารกในครรภ์
 - กระตุ้นให้หญิงตั้งครรภ์ เห็นความสำคัญของการมาฝากครรภ์อย่างสม่ำเสมอ
 - บำรุงสุขภาพหญิงตั้งครรภ์ด้วยอาหารให้ครบ 5 หมู่ เป็นการส่งเสริมสุขภาพและป้องกันการเกิดโรคแทรกซ้อน
15. หญิงตั้งครรภ์เป็นครั้งแรกมาฝากครรภ์ เมื่อวันที่ 21 กันยายน 2538 ให้ประวัติ L.M.P. วันที่ 28 พฤษภาคม 2538 แล้วหญิงรายนี้ควรปฏิบัติตัวอย่างไร
- ห้ามเดินทางไกล
 - งดเว้นการอยู่ร่วมกับสามี
 - งดเว้นนมสำหรับให้นมบุตร
 - ท้องผูกให้รับประทานผลไม้มากๆ
 - รับประทานอาหารพวกแคลเซียมฟอสฟอรัส
16. ข้อใดไม่ใช่อาการผิดปกติที่หญิงตั้งครรภ์ต้องมาพบแพทย์ก่อนกำหนดนัด
- เด็กดิ้นน้อยลง
 - มีเลือดออกทางช่องคลอด
 - มีน้ำใสๆ ไหลออกทางช่องคลอด
 - บวมตามขา น้ำหนักขึ้นหนึ่งกิโลกรัมต่อสัปดาห์
 - มีตกขาวลักษณะใส คล้ายแป้งเปียกออกมา

17. ข้อใดบ่งบอกลักษณะสำคัญของภาวะพิษแห่งครรภ์ในระยะแรกได้ถูกต้อง
- ความดันโลหิตสูง อากาการบวม ไข่ขาวในปัสสาวะ
 - ความดันโลหิตสูง อากาการบวม ปัสสาวะน้อย
 - ความดันโลหิตสูง อากาการตาพร่ามัวมองไม่เห็น ไข่ขาวในปัสสาวะ
 - ความดันโลหิตสูง อากาการตาพร่ามัวมองไม่เห็น ปัสสาวะน้อย
 - ความดันโลหิตสูง ปวดเจ็บบริเวณลิ้นปี่หรือชายโครงขวา อากาการบวม
18. หญิงตั้งครรภ์รายหนึ่งมีอายุครรภ์ 36 สัปดาห์ ตรวจร่างกายและตรวจครรภ์แล้วมีอาการปกติ ให้ประวัติว่าหลังคลอดครรภ์แรกมีไข้และได้รับการรักษาอยู่นาน ครรภ์ที่สองแท้งเมื่ออายุครรภ์ได้ 3 เดือน ได้รับการขูดมดลูก ภาวะใดที่อาจเกิดขึ้นได้ในหญิงตั้งครรภ์รายนี้
- ทารกอยู่ท่าขวาง
 - รกฝังตัวแน่นกว่าปกติ
 - ตกเลือดก่อนคลอด
 - ตกเลือดหลังคลอด
 - ติดเชื้อหลังคลอด
19. หญิงครรภ์แรกรายหนึ่ง อายุครรภ์ 38 สัปดาห์ ตรวจหน้าท้องพบว้ายอดมดลูกสูง 3/4 มากกว่าระดับสะดือ ซึ่งเท่ากับการตรวจครรภ์ครั้งที่แล้ว HPE มีอาการปัสสาวะบ่อยและอาการปวดหลัง หญิงตั้งครรภ์รายนี้มีอาการของอะไร
- Cystitis
 - Quickening
 - Lightening
 - Pyelonephritis
 - Intra-uterine growth retardation
20. ภาวะใดที่เป็นสัญญาณอันตรายมากที่สุดต่อชีวิตของทารกในครรภ์ของหญิงที่มีอายุครรภ์ 32 สัปดาห์
- ส่วนน้ำมีลักษณะกลมนิ่ม
 - ทารกดิ้น 3 - 4 ครั้งใน 1 ชั่วโมง
 - หาความสูงของมดลูกเพิ่มใน 1 ชั่วโมง
 - เสียงหัวใจทารก 164 ครั้งใน 1 นาที
 - ระดับยอดมดลูกสูง 3/4 มากกว่าระดับสะดือ

21. นางกลอยใจคลอดรกออกมา ถ้าสายสะดือยาวมากกว่า 50 เซนติเมตรแล้วภาวะใดที่อาจเกิดขึ้นได้

- ก. Prolapsed cord
- ข. False jelly knot
- ค. True knot of cord
- ง. Abruptio placenta
- จ. Placenta previa

22. นางรัตน์ มีรอบประจำเดือนมาวันที่ 5 - 9 มกราคม 2539 และในวันที่ 15 มกราคม 2539 เริ่มมีมูกเหนียวใส นางรัตน์ควรงดเว้นการมีเพศสัมพันธ์ในวันใดจึงจะไม่เกิดการตั้งครรภ์

- ก. 18 มกราคม 2539
- ข. 20 มกราคม 2539
- ค. 22 มกราคม 2539
- ง. 24 มกราคม 2539
- จ. 26 มกราคม 2539

23. ข้อใดไม่เกี่ยวกับการเปลี่ยนแปลงของร่างกายอันเนื่องมาจากการตั้งครรภ์

- ก. Chloasma ที่เต้านม
- ข. Chadwick ' s sign ที่เต้านม
- ค. Montgomery ' s tubercle ที่ areola
- ง. Striae gravidarum ที่หน้าท้อง
- จ. Braxton Hick ' s contraction

24. หญิงตั้งครรภ์รายใดที่มีการเปลี่ยนแปลงของน้ำหนักตัวที่ผิดปกติ

- ก. อายุครรภ์ 8 สัปดาห์ คลื่นไส้-อาเจียน รับประทานอาหารได้น้อย น้ำหนักลดลงจากก่อนตั้งครรภ์ 1 กิโลกรัม
- ข. อายุครรภ์ 12 สัปดาห์ คลื่นไส้-อาเจียนเล็กน้อย รับประทานอาหารได้ตามปกติ น้ำหนักคงเดิม
- ค. อายุครรภ์ 20 สัปดาห์ รับประทานอาหารได้มากกว่าปกติ น้ำหนักเพิ่มขึ้นจากก่อนตั้งครรภ์ 6 กิโลกรัม
- ง. อายุครรภ์ 32 สัปดาห์ น้ำหนักเพิ่มขึ้นครั้งก่อน 2.5 กิโลกรัม

- จ. อายุครรภ์ 36 สัปดาห์ รับประทานอาหารได้น้อย น้ำหนักเพิ่มจากครั้งก่อน 3 กิโลกรัม
25. หญิงครรภ์แรกมาฝากครรภ์วันที่ 8 พฤศจิกายน 2538 ให้ประวัติ L.M.P. วันที่ 6 กันยายน 2538 หญิงคนนี้มีอายุครรภ์เท่าใด
- ก. 8 สัปดาห์
ข. 9 สัปดาห์
ค. 10 สัปดาห์
ง. 11 สัปดาห์
จ. 12 สัปดาห์
26. หญิงตั้งครรภ์รายใด ที่ควรส่งต่อให้แพทย์ตรวจเพื่อวินิจฉัย
- ก. นางจัน อายุครรภ์ 20 สัปดาห์ เด็กไม่ดิ้นมา 1 วัน
ข. นางจอย อายุครรภ์ 30 สัปดาห์ ปัสสาวะบ่อย ปวดศีรษะเสมอ
ค. นางแจน อายุครรภ์ 36 สัปดาห์ ข้ำที่ข้อมือและข้อเท้า ปวดก้นกบ
ง. นางแจ้ว อายุครรภ์ 36 สัปดาห์ ปวดหลัง นอนไม่หลับ วิดกกังวล
จ. นางแจง อายุครรภ์ 16 สัปดาห์ มีอาการคลื่นไส้ อาเจียน เวียนศีรษะ
27. ถ้านางสายใจ มีฐานะยากจนมาก ตั้งครรภ์แรก รับประทานไม่เพียงพอกับความต้องการของร่างกาย ท่านคิดว่าอาการใดที่จะเกิดกับนางสายใจ
- ก. การแท้งบุตร รกเกาะต่ำ
ข. การแท้งบุตร ทารกน้ำหนักน้อย
ค. การแท้งบุตร คลอดก่อนกำหนด
ง. การแท้งบุตร คลอดก่อนกำหนด ทารกน้ำหนักน้อย
จ. การแท้งบุตร คลอดก่อนกำหนด ทารกน้ำหนักน้อย รกเกาะต่ำ
28. อาหารชนิดใดที่เหมาะสมกับหญิงตั้งครรภ์อายุครรภ์ 30 สัปดาห์ ซึ่งให้ประวัติว่าขณะนอนแล้วมีอาการเกร็งและปวดบริเวณน่องและต้นขาบ่อยขึ้น โดยเฉพาะเวลากลางคืน และเมื่อตรวจสภาพร่างกายทั่วไปมีอาการปกติ
- ก. ข้าวผัดไข่ดาว
ข. ข้าวสวยผัดพริกตับ
ค. ข้าวสวยกับผัดกระเพราไก่
ง. ข้าวสวยกับแกงจืดตำลึงหมูสับ

จ. ข้าวต้มกับปลาไส้ตันทอดกรอบ

29. ข้อใดไม่ใช่การวางแผนเพื่อป้องกันหญิงตั้งครรภ์ที่มีภาวะครรภ์เป็นพิษ

ก. ให้นอนพักมากๆ

ข. ให้อยู่ใกล้เคาเตอร์พยาบาล

ค. ดูแลให้ได้รับสารอาหารให้เพียงพอ

ง. ให้อยู่ในห้องที่เงียบสงบ

จ. ให้สังเกตอาการที่จะทำให้เกิดชัก

30. หญิงวัยรุ่น อายุ 14 ปี มีฐานะยากจน ตั้งครรภ์นอกสมรส อายุครรภ์ 12 สัปดาห์ ต้องการจะทำแท้ง ท่านคิดว่าควรจะทำหรือไม่

ก. ทำ เพราะบุตรที่เกิดมาจะมีปัญหาจากการขาดการยอมรับมารดา

ข. ทำ เพราะหญิงคนนี้จะขาดความพร้อมในการจะเป็นมารดาของบุตรที่เกิดมา

ค. ทำ เพราะหญิงคนนี้จะเกิดความอับอายต่อตนเองและเพื่อนบ้าน

ง. ทำ เพราะเป็นการช่วยเหลือทางด้านสภาพจิตใจของแม่

จ. ทำ เพราะไม่สามารถพูดและอธิบายให้เกิดการยอมรับการตั้งครรภ์ได้

1. นางกลอยใจคลอดรกออกมา ถ้าสายสะดือยาวมากกว่า 50 เซนติเมตรแล้วภาวะใดที่อาจเกิดขึ้นได้
 - ก. Prolapsed cord
 - ข. False jelly knot
 - ค. True knot of cord
 - ง. Abruptio placenta
 - จ. Placenta previa
2. นางรัตน์ มีรอบระดู 28 วัน ประจำเดือนมาวันที่ 5 - 9 มกราคม 2539 และในวันที่ 15 มกราคม 2539 เริ่มมีมูกเหนียวใส นางรัตน์ควรงดเว้นการมีเพศสัมพันธ์ในวันใดจึงจะไม่เกิดการตั้งครรภ์
 - ก. 18 มกราคม 2539
 - ข. 20 มกราคม 2539
 - ค. 22 มกราคม 2539
 - ง. 24 มกราคม 2539
 - จ. 26 มกราคม 2539
3. ข้อใดไม่เกี่ยวกับการเปลี่ยนแปลงของร่างกายอันเนื่องมาจากการตั้งครรภ์
 - ก. Chloasma ที่เต้านม
 - ข. Chadwick 's sign ที่เต้านม
 - ค. Montgomery 's tubercle ที่ areola
 - ง. Striae gravidarum ที่หน้าท้อง
 - จ. Braxton Hick 's contraction
4. หญิงตั้งครรภ์รายใดที่มีการเปลี่ยนแปลงของน้ำหนักตัวที่ผิดปกติ
 - ก. อายุครรภ์ 8 สัปดาห์ คลื่นไส้-อาเจียน รับประทานอาหารได้น้อย น้ำหนักลดลงจากก่อนตั้งครรภ์ 1 กิโลกรัม
 - ข. อายุครรภ์ 12 สัปดาห์ คลื่นไส้-อาเจียนเล็กน้อยรับประทานอาหารได้ตามปกติ น้ำหนักคงเดิม
 - ค. อายุครรภ์ 20 สัปดาห์ รับประทานอาหารได้มากกว่าปกติ น้ำหนักเพิ่มขึ้นจากก่อนตั้งครรภ์ 6 กิโลกรัม
 - ง. อายุครรภ์ 32 สัปดาห์ น้ำหนักเพิ่มขึ้นครั้งก่อน 2.5 กิโลกรัม

- จ. อายุครรภ์ 36 สัปดาห์ รับประทานอาหารได้น้อย น้ำหนัก เพิ่มจากครั้งก่อน 3 กิโลกรัม
5. หญิงครรภ์แรกมาฝากครรภ์วันที่ 8 พฤศจิกายน 2538 ให้ประวัติ L.M.P. วันที่ 6 กันยายน 2538 หญิงคนนี้มีอายุครรภ์เท่าใด
- 8 สัปดาห์
 - 9 สัปดาห์
 - 10 สัปดาห์
 - 11 สัปดาห์
 - 12 สัปดาห์
6. หญิงตั้งครรภ์รายใด ที่ควรส่งต่อให้แพทย์ตรวจเพื่อวินิจฉัย
- นางจัน อายุครรภ์ 20 สัปดาห์ เด็กไม่เดินมา 1 วัน
 - นางจอย อายุครรภ์ 30 สัปดาห์ บัสสาวะบ่อย ปวดศีรษะเสมอ
 - นางแจน อายุครรภ์ 36 สัปดาห์ ซ้ำที่ข้อมือและข้อเท้า ปวดก้นกบ
 - นางแจ้ว อายุครรภ์ 36 สัปดาห์ ปวดหลัง นอนไม่หลับ วิตกกังวล
 - นางแจจ อายุครรภ์ 16 สัปดาห์ มีอาการคลื่นไส้ อาเจียน เวียนศีรษะ
7. ถ้านางสายใจ มีฐานะยากจนมาก ตั้งครรภ์แรก รับประทานไม่เพียงพอกับความต้องการของร่างกาย ท่านคิดว่าอาการใดที่จะเกิดกับนางสายใจ
- การแท้งบุตร รกเกาะต่ำ
 - การแท้งบุตร ทารกน้ำหนักน้อย
 - การแท้งบุตร คลอดก่อนกำหนด
 - การแท้งบุตร คลอดก่อนกำหนด ทารกน้ำหนักน้อย
 - การแท้งบุตร คลอดก่อนกำหนด ทารกน้ำหนักน้อย รกเกาะต่ำ
8. อาหารชนิดใดที่เหมาะสมกับหญิงตั้งครรภ์อายุครรภ์ 30 สัปดาห์ ซึ่งให้ประวัติว่าขณะนอนแล้วมีอาการเกร็งและปวดบริเวณท้องและต้นขาบ่อยขึ้น โดยเฉพาะเวลากลางคืน และเมื่อตรวจสภาพร่างกายทั่วไปมีอาการปกติ
- ข้าวผัดไข่ดาว
 - ข้าวสวยผัดพริกตำ
 - ข้าวสวยกับผัดกระเพราไก่
 - ข้าวสวยกับแกงจืดตำลึงหมูสับ

- จ. ข้าวต้มกับปลาไส้ตันทอดกรอบ
9. ข้อใดไม่ใช่การวางแผนเพื่อป้องกันหญิงตั้งครรภ์ที่มีภาวะครรภ์เป็นพิษ
- ให้นอนพักมากๆ
 - ให้อยู่ใกล้เคาเตอร์พยาบาล
 - ดูแลให้ได้รับสารอาหารให้เพียงพอ
 - ให้อยู่ในห้องที่เงียบสงบ
 - ให้สังเกตอาการที่จะทำให้เกิดชัก
10. หญิงวัยรุ่น อายุ 14 ปี มีฐานะยากจน ตั้งครรภ์นอกสมรส อายุครรภ์ 12 สัปดาห์ ต้องการจะทำแท้ง ท่านคิดว่าควรจะทำหรือไม่
- ทำเพราะบุตรที่เกิดมาจะมีปัญหาจากการขาดการยอมรับจากมารดา
 - ทำเพราะหญิงคนนี้ขาดพร้อมในการจะเป็นมารดาของบุตรที่เกิดมา
 - ทำเพราะหญิงคนนี้จะเกิดความอับอายต่อตนเองและเพื่อนบ้าน
 - ทำเพราะเป็นการช่วยเหลือทางด้านสภาพจิตใจของแม่
 - ทำเพราะไม่สามารถพูดและอธิบายให้เกิดการยอมรับการตั้งครรภ์ได้
11. การวินิจฉัยการตั้งครรภ์ที่ได้ผลแน่นอนที่สุดคือข้อใด
- Probable sign
 - Positive sign
 - Presumptive sign
 - Urine Pregnancy Test
 - Post - Pregnancy sign
12. หญิงรายหนึ่ง อยู่ในระยะใกล้คลอด มดลูกหดรัดตัว 10 - 20 นาที ต่อ ครั้ง แสดงว่าเกิดอาการอะไร
- Hegar ' s sign
 - Chadwick ' s sign
 - Uterine contraction
 - Ballotement ' s sign
 - Braxton Hick ' s contraction
13. คำกล่าวใด ไม่ถูกต้อง ในเรื่องน้ำหนักของหญิงที่อยู่ในระหว่างตั้งครรภ์
- ในระยะหนึ่งสัปดาห์ น้ำหนักควรเพิ่มมากกว่า 1 กิโลกรัม

- ข. ตลอดระยะเวลาของการตั้งครรภ์น้ำหนักจะเพิ่มขึ้นประมาณ 10 - 12 กิโลกรัม
- ค. ในช่วงสามเดือนแรก น้ำหนักจะเพิ่มขึ้น ประมาณ 1 - 1/2 กิโลกรัม
- ง. ในช่วงไตรมาสที่สอง น้ำหนักจะเพิ่มขึ้น ประมาณ 5 - 6 กิโลกรัม
- จ. ในช่วงไตรมาสที่สาม น้ำหนักจะเพิ่มขึ้น ประมาณ 4 - 5 กิโลกรัม
14. ถ้าท่านซักประวัติพบว่า นางซิมมีครรภ์ครั้งแรก เมื่ออายุครรภ์ครบ 24 สัปดาห์ มีตกขาว สีเหลือง คันช่องคลอด ท่านจะปฏิบัติอย่างไร
- ก. แนะนำให้รักษาความสะอาดอวัยวะสืบพันธุ์
- ข. ตรวจครรภ์แล้วส่งให้แพทย์ตรวจภายในพร้อมทั้งแนะนำให้รักษาความสะอาด
- ค. แนะนำให้รักษาความสะอาดทางช่องคลอด โดยนำไปตากแดดฆ่าเชื้อโรค
- ง. แนะนำให้งดร่วมประเวณี จนกว่าจะหายจากการคันทางช่องคลอด
- จ. ไม่ต้องแนะนำเนื่องจากเป็นอาการปกติของหญิงตั้งครรภ์ที่จะมีมูกออกมา
15. นางน้อยมาฝากครรภ์ที่หน่วยฝากครรภ์ ขณะนอนหงายตรวจครรภ์ มีอาการเป็นลม หน้าซีด เหงื่อออก ตัวเย็น อาการนี้หมายถึงอะไร
- ก. Supine Hypertensive Syndrome
- ข. Supine Hypotensive Syndrome
- ค. Faintness
- ง. Dizziness
- จ. Syncope
16. นางดารายาอายุครรภ์ 36 สัปดาห์ มาฝากครรภ์ที่โรงพยาบาลชลบุรี วัดความดันโลหิตได้ 130/60 มม.ปรอท ผลปัสสาวะ +1 Hb 11 gm% ขณะที่ยืนเตียงเพื่อตรวจครรภ์ มีอาการวิงเวียนศีรษะ หน้าซีด เป็นลม ท่านจะให้การรักษาช่วยเหลืออย่างไร
- ก. ให้ดื่มน้ำหวานหรือน้ำอัดลม
- ข. ให้นอนราบศีรษะต่ำ ดมแอมโมเนีย
- ค. ให้นอนศีรษะสูง 45 องศา และให้ลมโกรก
- ง. วัดความดันโลหิต แล้วส่งต่อแพทย์เพื่อทำการรักษา
- จ. ให้นอนตะแคงซ้าย สูดลมหายใจลึกๆ ดมแอมโมเนีย

17. การตรวจโดยวิธี Pawlik grip(ท่าที่ 3) เพื่อหาอะไร

- ก. Attitude
- ข. Position
- ค. Presentation
- ง. Hight of Fundus
- จ. Large part and Small part

18. E.D.C. (Expected Date of Confinement) หมายถึงอะไร

- ก. ระยะเวลาการตั้งครรภ์
- ข. วันแรกที่ทารกในครรภ์คลอด
- ค. การคาดคะเนท่าของทารก
- ง. การตรวจดูการเจริญของทารก
- จ. การตรวจพิเศษทางหน้าท้อง

19. ข้อใดเป็นความหมายที่ถูกต้องที่สุดของ Ante Natal Care

- ก. การดูแลสุขภาพอนามัยของมารดาในระยะตั้งครรภ์ ระยะคลอด และระยะหลังคลอด
- ข. การดูแลสุขภาพของทารกในครรภ์ เพื่อลดอัตราการตายของเด็กในระยะแรกคลอด
- ค. การดูแลช่วยเหลือมารดาให้รู้จักวิธีปฏิบัติตนในระยะคลอดและหลังคลอด
- ง. การดูแลมารดาให้มีสุขภาพสมบูรณ์ทั้งร่างกาย จิตใจ เพื่อลดภาวะแทรกซ้อนที่อาจเกิดขึ้นในระยะตั้งครรภ์
- จ. การดูแลสุขภาพร่างกาย จิตใจ ของมารดาขณะตั้งครรภ์ รวมทั้งการดูแลทารกในครรภ์ เพื่อตรวจค้นและหาแนวทางแก้ไขภาวะแทรกซ้อนที่อาจเป็นอันตรายต่อแม่และเด็ก

20. หญิงอายุครรภ์ 20 สัปดาห์ ให้ประวัติว่าเคยทำแท้ง 1 ครั้ง ตั้งครรภ์ไขปลาอุกเมื่ออายุครรภ์ 22 สัปดาห์ 1 ครั้ง คลอดบุตรเมื่ออายุครรภ์ 34 สัปดาห์ 1 ครั้ง และตั้งครรภ์แฝดคลอดเมื่ออายุครรภ์ 38 สัปดาห์ บุตรทุกคนมีชีวิตอยู่ สัญลักษณ์ เกี่ยวกับการตั้งครรภ์ของหญิงคนนี้คือข้อใด

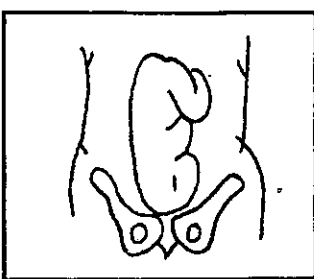
- ก. PARA 0 - 2 - 2 - 3
- ข. PARA 1 - 1 - 2 - 3

ค. PARA 1 - 1 - 1 - 3

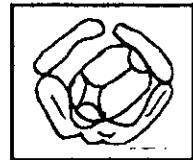
ง. PARA 1 - 2 - 1 - 3

จ. PARA 2 - 1 - 2 - 3

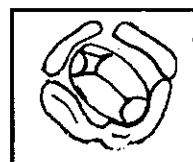
21. ถ้ามองเข้ามาจากช่องเชิงกรานจะเห็นส่วนนำลักษณะใด



ก.



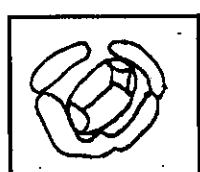
ข.



ค.



ง.

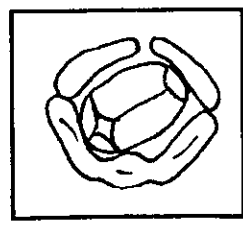


จ.



22. จากภาพที่กำหนดให้เป็นภาพที่มองเข้ามาจากช่องออกเชิงกราน ทารกอยู่ในท่าใด

- ก. LOA
- ข. ROA
- ค. LOP
- ง. ROP
- จ. RSA



23. นางฉายมาฝากครรภ์ ชักประวัติ L.M.P. คือวันที่ 24 มิถุนายน 2538 ถ้ามาตรวจครรภ์ วันที่ 11 มีนาคม 2539 ผู้ตรวจครรภ์ ควรนัดให้นางฉายมาตรวจครรภ์ครั้งต่อไปเมื่อใด
- ก. 1 เมษายน 2539
 - ข. 8 เมษายน 2539
 - ค. 18 มีนาคม 2539
 - ง. 25 มีนาคม 2539
 - จ. 31 มีนาคม 2539
24. ข้อใดที่อาจมีผลต่อการตั้งครรภ์และการคลอดมากที่สุด
- ก. เคยผ่าตัดไส้ติ่งมาแล้ว 1 ปี
 - ข. ครรภ์ก่อนคลอดลำบาก เด็กอยู่ในท่าก้น
 - ค. ครรภ์ก่อนไม่มีแรงเบ่ง เจ็บครรภ์หลายวัน
 - ง. แท้ง ต้องชูดมดลูกเมื่อ 5 เดือนก่อนตั้งครรภ์
 - จ. เคยใส่ห่วงอนามัยก่อน และมีเลือดออกกระปริบกระปรอยก่อนตั้งครรภ์
25. ควรให้คำแนะนำอย่างไรในการที่จะให้หญิงตั้งครรภ์ได้อาหารที่มีประโยชน์ต่อตนเองและทารกในครรภ์
- ก. แนะนำอาหารจำพวกโปรตีน เหลือแร่ และวิตามิน
 - ข. แนะนำอาหารที่มีอยู่ในท้องถิ่นและมีประโยชน์ต่อร่างกาย
 - ค. แนะนำให้หลีกเลี่ยงอาหารที่จะเป็นอันตรายต่อทารกในครรภ์
 - ง. แนะนำรายชื่ออาหารที่ควรและไม่ควรรับประทาน
 - จ. แนะนำให้รับประทานอาหารที่หญิงตั้งครรภ์ชอบ
26. Hypremesis Gravidarum ในหญิงตั้งครรภ์มักเกิดเวลาใด
- ก. เวลาเช้า
 - ข. เวลาเย็น
 - ค. ไม่เป็นเวลา
 - ง. ตลอดเวลา
 - จ. เช้าและเย็น
27. ข้อใดที่แสดงว่าทารกในครรภ์มีภาวะของ Fetal Distress
- ก. F.H.S. 100 ครั้งต่อนาที
 - ข. F.H.S. 110 ครั้งต่อนาที

ค. F.H.S. 120 ครั้งต่อนาที

ง. F.H.S. 130 ครั้งต่อนาที

จ. F.H.S. 160 ครั้งต่อนาที

28. นางจอย ตั้งครรภ์ได้ 20 สัปดาห์แล้วเกิดอาการตกขาว มีลักษณะเหนียวใส ไม่มีกลิ่น ไม่มีอาการคัน และปัสสาวะไม่แสบ การเกิดอาการดังกล่าวตรงกับข้อใด

ก. ปากมดลูกเริ่มขยาย

ข. ต่อมที่ปากมดลูกทำงานมากขึ้น

ค. ฮอร์โมนโปรเจสเตอโรนทำงานมากขึ้น

ง. ความเปียกชื้นของปากช่องคลอดมากขึ้น

จ. การรักษาความสะอาดบริเวณอวัยวะสืบพันธุ์ไม่เพียงพอ

29. หญิงรายหนึ่งมาโรงพยาบาลด้วยอาการประจำเดือนขาดไป 2 เดือน คลื่นไส้ อาเจียนมาก ปัสสาวะบ่อยตรวจหน้าท้องพบว่า ยอดมดลูกสูง 2/3 มากกว่าระดับหัวเหน่า จากข้อมูลดังกล่าวหญิงรายนี้มีภาวะเสี่ยงต่ออะไรน้อยที่สุด

ก. ครรภ์แฝด

ข. ครรภ์เป็นพิษ

ค. ครรภ์แฝดน้ำ

ง. ครรภ์ไขปลาคูก

จ. ครรภ์เล็กกว่าอายุครรภ์

30. อาการของหญิงตั้งครรภ์รายใดควรส่งปรึกษาแพทย์

ก. อายุครรภ์ 10 สัปดาห์ คลื่นไส้-อาเจียนในตอนเช้า อ่อนเพลีย รับประทานอาหารได้น้อยกว่าปกติ น้ำหนักไม่เพิ่ม

ข. อายุครรภ์ 16 สัปดาห์ มีอาการเป็นลมถ้ายืนนานๆ ความดันโลหิต 100/60 มม.ปรอท มีตกขาว ไม่มีกลิ่นเหม็น

ค. อายุครรภ์ 34 สัปดาห์ น้ำหนักเพิ่มจากครั้งก่อน 4 กิโลกรัม หนึ่งอย่าง บวมที่ใบหน้า ความดันโลหิต 140/90 มม.ปรอท

ง. อายุครรภ์ 36 สัปดาห์ ระดับยอดมดลูก 3/4 มากกว่าระดับสะดือ มดลูกแข็งตัวเป็นบางครั้ง ไม่สม่ำเสมอ

จ. อายุครรภ์ 38 สัปดาห์ น้ำหนักเพิ่มจากครั้งก่อน 0.5 กิโลกรัม มีอาการบวมที่หลังเท้า ทารกในครรภ์ดิ้นดี ความดันโลหิต 130/90 มม.ปรอท

ภาคผนวก ข
ค่าไอเทินขององค์ประกอบที่มีค่าเกิน 1.0

ตาราง 13 ค่าไอเทินขององค์ประกอบที่มีค่าเกิน 1.0 ของแบบทดสอบวิชาการพยาบาล
สูติศาสตร์ 1

องค์ประกอบ	ค่าไอเทิน	องค์ประกอบ	ค่าไอเทิน
1	4.485	16	1.441
2	2.409	17	1.410
3	2.078	18	1.365
4	2.051	19	1.347
5	2.017	20	1.304
6	1.917	21	1.302
7	1.865	22	1.270
8	1.807	23	1.225
9	1.747	24	1.184
10	1.711	25	1.154
11	1.649	26	1.149
12	1.597	27	1.127
13	1.571	28	1.085
14	1.563	29	1.051
15	1.496	30	1.048

ภาคผนวก ค

ค่าพารามิเตอร์ของแบบทดสอบจากการทดลองเครื่องมือ

ตาราง 14 ผลการวิเคราะห์ค่าพหามิตอร์จากการทดลองเครื่องมือ

เนื้อหา	พฤติกรรม	ข้อ	a	b	c	CHISQUARE (PROP)	หมายเหตุ
1	ความจำ	1	.185	-4.959	.154	.0095	ใช้ได้
		2	.644	3.755	.076	.9432	
		3	.207	4.735	.500	.0070	
		4	.433	0.800	.156	.8313	
		5	.437	4.570	.110	.9456	
		6	.893	2.188	.061	.0625	
	ความเข้าใจ	7	.467	2.716	.103	.5211	ใช้ได้
		8	.434	1.157	.099	.2033	ใช้ได้
		9	.666	2.555	.120	.3271	ใช้ได้
		10	.384	4.159	.192	.0227	
		11	.453	5.600	.077	.4666	
	นำไปใช้	12	.273	2.375	.157	.5297	ใช้ได้
		13	.407	6.023	.221	.9410	
2	ความจำ	14	.384	3.821	.277	.0423	ใช้ได้
		15	.441	2.530	.074	.0006	
		16	.423	4.131	.283	.2853	ใช้ได้
		17	.424	-.434	.133	.0150	
		18	.432	.300	.108	.9489	
	ความเข้าใจ	19	.613	3.610	.080	.0390	ใช้ได้
		20	.426	-.120	.156	.0563	ใช้ได้
		21	1.057	3.313	.091	.4598	ใช้ได้
		22	.503	2.326	.127	.1381	ใช้ได้
	นำไปใช้	23	.643	.292	.082	.0203	ใช้ได้
		24	.376	-.297	.144	.0552	ใช้ได้
		26	.223	-.2.465	.160	.0445	ใช้ได้

ตาราง 14 (ต่อ)

เนื้อหา	พฤติกรรม	ข้อ	a	b	c	CHISQUARE (PROP)	หมายเหตุ
		27	.387	-1.410	.149	.1244	ใช้ได้
	วิเคราะห์	25	.364	3.404	.239	.8168	ใช้ได้
		28	.578	.598	.094	.0248	ใช้ได้
3	ความจำ	29	.337	-3.939	.152	.0324	ใช้ได้
		30	.557	-.718	.161	.1483	ใช้ได้
		31	.397	3.180	.125	.0932	ใช้ได้
		32	.299	-2.373	.156	.0238	ใช้ได้
		33	.407	7.308	.226	.2110	
		34	.848	1.323	.072	.1390	ใช้ได้
		35	.455	1.913	.135	.3744	ใช้ได้
		36	.441	-2.228	.145	.2678	ใช้ได้
		37	.556	-1.238	.136	.1007	ใช้ได้
	38	.410	6.311	.222	.4470		
	ความเข้าใจ	39	1.073	2.401	.136	.7791	ใช้ได้
		40	.694	1.680	.091	.2506	ใช้ได้
		41	.741	2.347	.156	.2897	ใช้ได้
		42	1.142	1.715	.122	.0001	
		43	.536	.958	.137	.9736	ใช้ได้
		44	.665	1.837	.070	.0172	ใช้ได้
		45	.367	.793	.111	.0003	
	นำไปใช้	46	.297	4.258	.152	.7149	
47		.253	2.615	.176	.1279	ใช้ได้	
48		.404	4.243	.142	.7527		
49		.625	-1.260	.137	.6587		
51		.359	-.485	.139	.5427		

ตาราง 14 (ต่อ)

เนื้อหา	พฤติกรรม	ข้อ	a	b	c	CHISQUARE (PROP)	หมายเหตุ
	วิเคราะห์	50	.293	.284	.169	.0683	ใช้ได้
		52	.225	-1.077	.166	.2788	ใช้ได้
		53	.286	-.432	.148	.5565	ใช้ได้
	สังเคราะห์	54	.257	-1.316	.160	.0272	ใช้ได้
	ประเมินค่า	55	.250	-3.028	.150	.1311	ใช้ได้
4	ความจำ	56	.544	2.325	.049	.0672	ใช้ได้
	ความเข้าใจ	57	.267	.482	.165	.5253	ใช้ได้
	นำไปใช้	58	.150	1.951	.227	.1085	ใช้ได้
		59	.364	-.450	.150	.0001	ใช้ได้
		60	.477	.646	.199	.8812	ใช้ได้
		61	.322	-2.166	.143	.6640	
	วิเคราะห์	62	.306	.070	.143	.0322	ใช้ได้
		63	.189	-4.095	.156	.4638	
5	ความจำ	64	.418	-3.789	.147	.6667	ใช้ได้
		65	.400	-.141	.136	.0535	ใช้ได้
		66	.848	-.864	.123	.2346	ใช้ได้
		67	.272	1.539	.188	.4317	ใช้ได้
		68	.252	2.834	.206	.2085	ใช้ได้
	ความเข้าใจ	69	.755	1.772	.145	.0049	
		70	.368	.699	.143	.0968	ใช้ได้
	นำไปใช้	71	.360	1.562	.178	.3018	ใช้ได้
	วิเคราะห์	72	.312	1.230	.124	.4530	ใช้ได้
		73	.424	3.780	.104	.7287	ใช้ได้
		74	.607	-1.344	.145	.4178	ใช้ได้
		75	.677	-.261	.145	.5319	ใช้ได้

ตาราง 14 (ต่อ)

เนื้อหา	พฤติกรรม	ข้อ	a	b	c	CHISQUARE (PROP)	หมายเหตุ
	สังเคราะห์	76	.325	1.934	.190	.8586	ใช้ได้
	ประเมินค่า	77	.144	-2.893	.170	.0184	ใช้ได้

ภาคผนวก ง
ค่าพารามิเตอร์ของแบบทดสอบเทียบคะแนนทั้งสองฉบับ
และแบบทดสอบร่วม

ตาราง 15 ค่าพารามิเตอร์ของแบบทดสอบเทียบคะแนนและแบบทดสอบร่วมทั้งสองฉบับ

ข้อที่	ข้อเดิม	ฉบับ X			ข้อเดิม	ฉบับ Y		
		a	b	c		a	b	c
1	4	0.433	0.800	0.156	6	0.893	2.188	0.061
2	9	0.666	2.555	0.120	7	0.467	2.716	0.103
3	18	0.432	0.300	0.108	17	0.424	-0.434	0.133
4	26	0.223	-2.465	0.160	27	0.387	-1.410	0.149
5	24	0.760	-0.297	0.144	20	0.426	-0.120	0.156
6	22	0.503	2.326	0.127	23	0.643	0.292	0.082
7	34	0.848	1.323	0.072	35	0.455	1.913	0.135
8	36	0.441	-2.228	0.145	37	0.556	-1.238	0.136
9	32	0.299	-2.373	0.156	30	0.557	-0.718	0.161
10	41	0.741	2.347	0.156	39	1.073	2.401	0.136
11	40	0.649	1.680	0.091	44	0.665	1.837	0.070
12	45	0.367	0.793	0.111	43	0.536	0.958	0.137
13	51	0.359	-0.485	0.139	49	0.625	-1.260	0.137
14	53	0.286	-0.432	0.148	52	0.225	-1.077	0.166
15	58	0.150	1.951	0.227	60	0.477	0.646	0.199
16	66	0.848	-0.864	0.123	65	0.400	-0.141	0.136
17	67	0.272	1.539	0.188	68	0.252	2.384	0.206
18	70	0.368	0.699	0.143	69	0.755	1.772	0.145
19	71	0.360	1.562	0.178	72	0.312	1.230	0.124
20	75	0.677	-0.261	0.145	74	0.607	-1.344	0.145

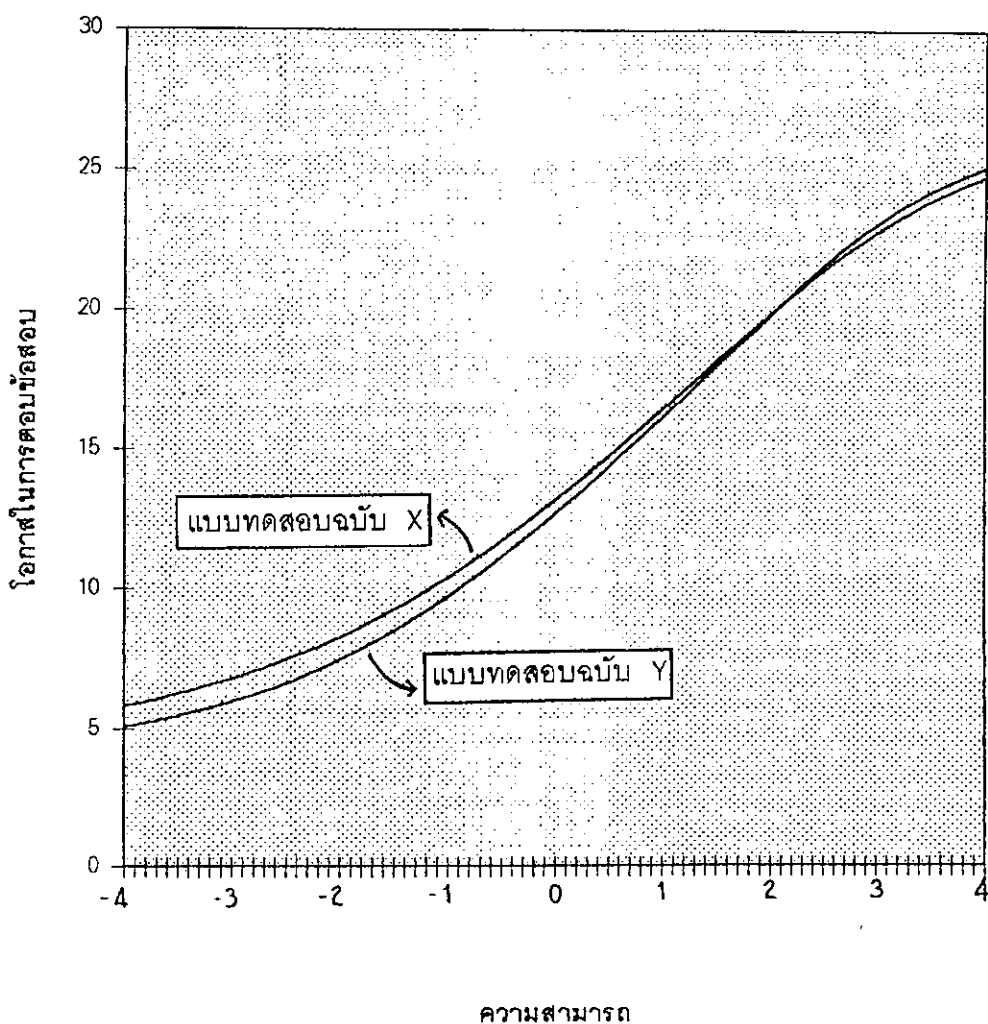
ตาราง 15 (ต่อ)

ข้อที่	ข้อเดิม	เนื้อหาที่	a	b	c
21	8	1	0.434	1.157	0.099
22	12	1	0.273	2.375	0.157
23	15	2	0.441	2.530	0.074
24	28	2	0.578	0.598	0.194
25	47	3	0.253	2.615	0.176
26	50	3	0.293	0.284	0.169
27	57	4	0.267	0.482	0.165
28	62	4	0.306	0.070	0.143
29	76	5	0.325	1.934	0.190
30	77	5	0.144	-2.893	0.170

ภาคผนวก จ
คุณภาพของแบบทดสอบทั้งสองฉบับ

ในการหาคุณภาพของแบบทดสอบ ผู้วิจัยได้นำค่าพารามิเตอร์ประจำข้อแต่ละข้อของแบบทดสอบทั้งสองฉบับ มาหาค่าโอกาสในการตอบถูก และค่าอินฟอร์เมชันของแบบทดสอบที่ระดับความสามารถต่างๆ ดังนี้

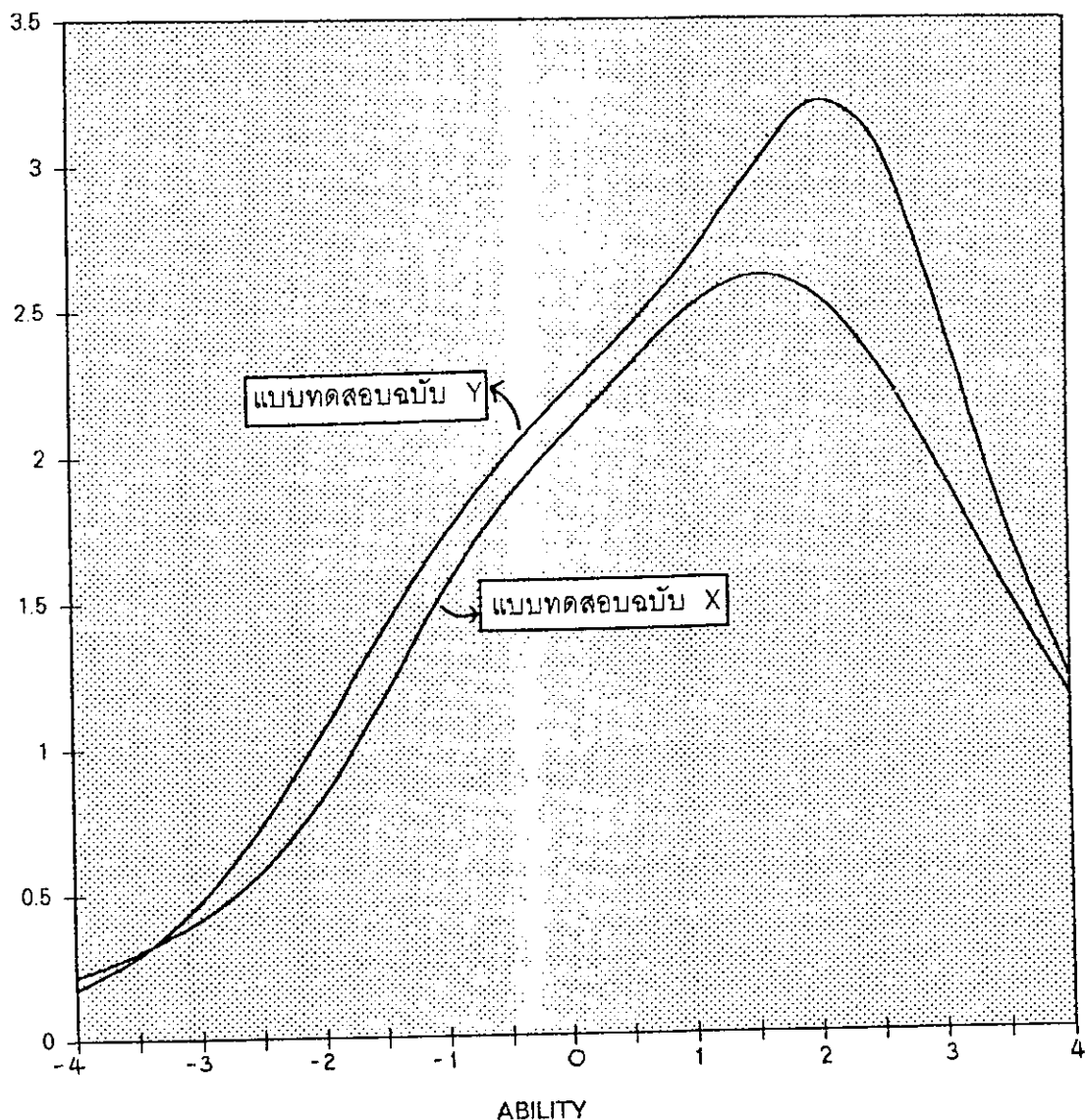
1. โอกาสในการตอบข้อสอบถูก ณ ระดับความสามารถต่างๆ แล้วนำมาเขียนโค้งลักษณะเฉพาะของแบบทดสอบ ซึ่งแสดงได้ดังภาพประกอบ 7



ภาพประกอบ 7 แสดงโค้งลักษณะเฉพาะของแบบทดสอบทั้งสองฉบับ

จากภาพประกอบ 7 แสดงให้เห็นว่า ที่ระดับความสามารถ -4.0 ถึง 1.0 โอกาสในการตอบข้อสอบถูกของแบบทดสอบ ฉบับ X สูงกว่า ฉบับ Y และที่ระดับความสามารถ 1.0 ถึง 4.0 แบบทดสอบทั้งสองฉบับมีโอกาสในการตอบข้อสอบถูกที่ใกล้เคียงกัน

2. ค่าสูงสุดของอินฟอร์เมชันของแบบทดสอบ ณ ระดับความสามารถต่างๆ
 แล้วนำมาเขียนโค้งอินฟอร์เมชันของแบบทดสอบที่ระดับความสามารถต่างๆ จะได้ดัง
 ภาพประกอบ 8



ภาพประกอบ 8 โค้งอินฟอร์เมชันของแบบทดสอบทั้งสองฉบับที่ระดับความสามารถต่างๆ

จากภาพประกอบ 8 แสดงให้เห็นว่า ที่ระดับความสามารถ -4.0 ถึง -3.0 แบบทดสอบทั้งสองฉบับให้ค่าอินฟอร์เมชันของแบบทดสอบที่ใกล้เคียงกัน แต่เมื่อพิจารณาที่ระดับความสามารถตั้งแต่ -3.0 ถึง 4.0 แล้วพบว่าแบบทดสอบฉบับ Y ให้ค่าอินฟอร์เมชันของแบบทดสอบสูงกว่า ฉบับ X

ภาคผนวก ฉ

ผลการวิเคราะห์ค่าพารามิเตอร์ของแบบทดสอบทั้งสองฉบับ

ผู้วิจัยได้นำผลการตอบแบบทดสอบเทียบคะแนนและแบบทดสอบรวม มาวิเคราะห์ เพื่อประมาณค่าพารามิเตอร์ของข้อสอบเป็นรายข้อ ปรากฏว่ามีข้อสอบบางข้อที่ไม่สามารถ ประมาณค่าพารามิเตอร์ความยากได้ คือแบบทดสอบเทียบคะแนนฉบับ X ข้อที่ 3 และแบบทดสอบรวมในฉบับ Y ข้อที่ 3 ดังนั้นในการวิเคราะห์ข้อมูลจึงต้องตัดข้อสอบที่ไม่สามารถ ประมาณค่าพารามิเตอร์ความยากดังกล่าวออกไป และเพื่อให้การแปลความหมายของผลการเทียบคะแนนชัดเจนและเข้าใจง่าย ผู้วิจัยจึงตัดข้อสอบในแบบทดสอบเทียบคะแนนฉบับ Y และแบบทดสอบรวมในฉบับ X โดยตัดข้อ 10 และข้อ 6 ซึ่งเป็นข้อสอบที่ง่ายมากและ อำนาจจำแนกต่ำกว่าข้ออื่นๆออก ดังนั้นในการวิจัยครั้งนี้ จึงวิเคราะห์แบบทดสอบเทียบ คะแนนฉบับละ 19 ข้อ และแบบทดสอบรวมฉบับละ 9 ข้อ ซึ่งได้ค่าพารามิเตอร์ของข้อสอบ ในแบบทดสอบแต่ละฉบับ ดังตาราง

ตาราง 16 ค่าพารามิเตอร์ของแบบทดสอบทั้งสองฉบับ

ข้อ	ฉบับ X			ฉบับ Y		
	a	b	c	a	b	c
1	1.083	1.720	0.252	0.350	3.577	0.094
2	1.204	1.439	0.252	0.588	1.415	0.023
3	0.304	-2.926	0.244	0.871	0.489	0.026
4	0.373	0.663	0.264	0.885	-0.155	0.038
5	0.486	3.049	0.188	0.149	0.056	0.096
6	0.982	1.710	0.116	0.623	0.767	0.044
7	0.811	-0.563	0.214	0.737	1.139	0.023
8	0.328	-1.176	0.244	0.270	0.173	0.078
9	0.550	2.718	0.118	0.443	-0.980	0.077
10	0.685	1.801	0.157	0.404	3.401	0.152

ตาราง 16 (ต่อ)

ข้อ	ฉบับ X			ฉบับ Y		
	a	b	c	a	b	c
11	0.455	1.757	0.237	0.238	1.456	0.092
12	0.395	0.656	0.248	0.295	-0.018	0.080
13	0.531	0.808	0.256	0.626	0.086	0.036
14	0.315	7.271	0.500	0.232	1.349	0.097
15	0.679	0.257	0.228	0.231	-0.453	0.082
16	0.424	1.707	0.279	0.252	3.004	0.085
17	0.552	1.312	0.244	0.352	4.140	0.121
18	0.325	0.771	0.245	0.414	1.467	0.054
19	0.644	0.537	0.210	0.998	0.081	0.024
20	0.337	2.278	0.162	0.659	0.549	0.027
21	0.389	1.375	0.159	0.772	0.523	0.023
22	1.084	0.853	0.065	1.034	0.529	0.020
23	0.313	0.787	0.16	0.349	3.276	0.500
24	0.507	-0.041	0.155	0.709	0.587	0.030
25	0.359	-0.074	0.155	0.954	0.173	0.029
26	0.525	1.647	0.124	0.906	0.701	0.018
27	0.273	2.878	0.207	0.673	0.785	0.026
28	0.116	-1.871	0.195	0.680	-0.336	0.048

ภาคผนวก ข

ผลการสร้างตารางเทียบคะแนนโดยวิธีการเทียบคะแนนจริง

ตาราง 17 การคำนวณคะแนนที่ได้จากการเทียบคะแนนจริง

θ	X	X*	Zx-Score	Zy-Score	CEEBx	CEEBy
-4.0	7.0246	6.4545	-1.1798	-1.1019	382	390
-3.9	7.0710	6.4838	-1.1724	-1.0980	383	390
-3.8	7.1195	6.5149	-1.1647	-1.0938	384	391
-3.7	7.1704	6.5479	-1.1566	-1.0894	384	391
-3.6	7.2236	6.5830	-1.1482	-1.0847	385	392
-3.5	7.2795	6.6204	-1.1393	-1.0798	386	392
-3.4	7.3381	6.6601	-1.1300	-1.0745	387	393
-3.3	7.3996	6.7023	-1.1203	-1.0688	388	393
-3.2	7.4641	6.7472	-1.1101	-1.0628	389	394
-3.1	7.5318	6.7951	-1.0993	-1.0564	390	394
-3.0	7.6030	6.8460	-1.0880	-1.0496	391	395
-2.9	7.6778	6.9002	-1.0761	-1.0424	392	396
-2.8	7.7565	6.9581	-1.0637	-1.0347	394	397
-2.7	7.8393	7.0197	-1.0505	-1.0264	395	397
-2.6	7.9264	7.0855	-1.0367	-1.0177	396	398
-2.5	8.0181	7.1558	-1.0222	-1.0083	398	399
-2.4	8.1148	7.2308	-1.0068	-0.9983	399	400
-2.3	8.2166	7.3110	-0.9907	-0.9875	401	401
-2.2	8.3239	7.3969	-0.9737	-0.9761	403	402
-2.1	8.4371	7.4888	-0.9557	-0.9638	404	404
-2.0	8.5564	7.5873	-0.9368	-0.9507	406	405
-1.9	8.6823	7.6930	-0.9168	-0.9365	408	406
-1.8	8.8150	7.8065	-0.8958	-0.9214	410	408
-1.7	8.9550	7.9285	-0.8736	-0.9051	413	409
-1.6	9.1025	8.0597	-0.8502	-0.8876	415	411
-1.5	9.2579	8.2009	-0.8255	-0.8687	417	413
-1.4	9.4212	8.3529	-0.7996	-0.8484	420	415

ตาราง 17 (ต่อ)

θ	X	X*	Zx-Score	Zy-Score	CEEBx	CEEBy
-1.3	9.5939	8.5170	-0.7722	-0.5265	423	417
-1.2	9.7750	8.6938	-0.7435	-0.8029	426	420
-1.1	9.9653	8.8843	-0.7133	-0.7775	429	422
-1.0	10.1649	9.0896	-0.6816	-0.7501	432	425
-0.9	10.3741	9.3107	-0.6484	-0.7206	435	428
-0.8	10.5932	9.5484	-0.6137	-0.6888	439	431
-0.7	10.8222	9.8037	-0.5774	-0.6547	442	435
-0.6	11.0613	10.0773	-0.5394	-0.6182	446	438
-0.5	11.3108	10.3699	-0.4998	-0.5791	450	442
-0.4	11.5708	10.6821	-0.4586	-0.5375	454	446
-0.3	11.8413	11.0143	-0.4157	-0.4931	458	451
-0.2	12.1227	11.3670	-0.3711	-0.4460	463	455
-0.1	12.4151	11.7407	-0.3247	-0.3961	468	460
0.0	12.7187	12.1359	-0.2765	-0.3433	472	466
0.1	13.0336	12.5531	-0.2266	-0.2876	477	471
0.2	13.3599	12.9927	-0.1748	-0.2289	483	477
0.3	13.6978	13.4552	-0.1212	-0.1672	488	483
0.4	14.0473	13.9407	-0.0658	-0.1024	493	490
0.5	14.4081	14.4491	-0.0085	-0.0345	499	497
0.6	14.7801	14.9796	0.0505	0.0363	505	504
0.7	15.1626	15.5305	0.1112	0.1099	511	511
0.8	15.5551	16.0995	0.1734	0.1859	517	519
0.9	15.9566	16.6832	0.2371	0.2638	524	526
1.0	16.3659	17.2771	0.3020	0.3431	530	534
1.1	16.7817	17.8759	0.3680	0.4230	537	542
1.2	17.2021	18.4733	0.4347	0.5028	543	550
1.3	17.6254	19.0631	0.5018	0.5816	550	558

ตาราง 17 (ต่อ)

θ	X	X*	Zx-Score	Zy-Score	CEEBx	CEEBy
1.4	18.0494	19.6388	0.5691	0.6584	557	566
1.5	18.4718	20.1949	0.6361	0.7327	564	573
1.6	18.8903	20.7267	0.7025	0.8037	570	580
1.7	19.3028	21.2310	0.7679	0.8710	577	587
1.8	19.7029	21.7061	0.8320	0.9344	583	593
1.9	20.1009	22.1512	0.8945	0.9939	589	599
2.0	20.4830	22.5667	0.9551	1.0493	596	605
2.1	20.8520	22.9537	1.0137	1.1010	601	610
2.2	21.2070	23.3135	1.0700	1.1491	607	615
2.3	21.5473	23.6478	1.1240	1.1937	612	619
2.4	21.8727	23.9584	1.1756	1.2352	618	624
2.5	22.1830	24.2468	1.2248	1.2737	622	627
2.6	22.4785	24.5147	1.2717	1.3094	627	631
2.7	22.7596	24.7636	1.3163	1.3427	632	634
2.8	23.0267	24.9947	1.3586	1.3735	636	637
2.9	23.2804	25.2094	1.3989	1.4022	640	640
3.0	23.5212	24.4087	1.4371	1.4288	644	643
3.1	23.7498	25.5937	1.4733	1.4535	647	645
3.2	23.9667	25.7654	1.5077	1.4764	651	648
3.3	24.1726	25.9246	1.5404	1.4977	654	650
3.4	24.3680	26.0723	1.5714	1.5174	657	652
3.5	24.5535	26.2092	1.6008	1.5357	660	654
3.6	24.7295	26.3361	1.6287	1.5526	663	655
3.7	24.8966	26.4536	1.6553	1.5683	666	657
3.8	25.0551	26.5625	1.6804	1.5828	668	658
3.9	25.2056	26.6633	1.7043	1.5963	670	660
4.0	25.3483	26.7566	1.7269	1.6088	673	661

ภาคผนวก ข
ผลการสร้างตารางเทียบคะแนนโดยวิธีการเทียบคะแนนดิบ

ตาราง 18 การคำนวณคะแนนที่ได้จากการเทียบคะแนนดิบ

คะแนน	X	X*	Zx-Score	Zy-Score	CEEBx	CEEBy
0	0.0864	0.1058	-1.5817	-1.6953	342	330
1	0.3561	0.4361	-1.5512	-1.6547	345	335
2	0.8567	1.0543	-1.4944	-1.5788	351	342
3	1.6037	1.9443	-1.4097	-1.4696	359	353
4	2.5971	3.2014	-1.2971	-1.3153	370	368
5	3.8145	4.6527	-1.1590	-1.1371	384	386
6	5.1993	6.2298	-1.0020	-0.9435	400	406
7	6.6637	7.8061	-0.8360	-0.7500	416	425
8	8.1118	9.2715	-0.6718	-0.5701	433	443
9	9.4705	10.5625	-0.5177	-0.4116	448	459
10	10.7167	11.6705	-0.3776	-0.2756	462	472
11	11.8262	12.6269	-0.2506	-0.1582	475	484
12	12.8581	13.4786	-0.1336	-0.0536	487	495
13	13.8370	14.2677	-0.0227	0.0433	498	504
14	14.7904	15.0215	0.0854	0.1358	509	514
15	15.7337	15.7528	0.1924	0.2256	519	523
16	16.6714	16.4658	0.2987	0.3131	530	531
17	17.6020	17.1631	0.4042	0.3987	540	540
18	18.5232	17.8499	0.5087	0.4830	551	548
19	19.4341	18.5357	0.6120	0.5672	561	557
20	20.3361	19.2321	0.7142	0.6527	571	565

ตาราง 18 (ต่อ)

คะแนน	X	X*	Zx-Score	Zy-Score	CEEBx	CEEBy
21	21.2323	19.9508	0.8158	0.7409	582	574
22	22.1280	20.7030	0.9174	0.8333	592	583
23	23.0303	21.5011	1.0197	0.9313	602	593
24	23.9476	22.3657	1.1237	1.0374	612	604
25	24.8886	23.3342	1.2304	1.1563	623	616
26	25.8650	24.4703	1.3411	1.2958	634	630
27	26.8970	25.8877	1.4581	1.4698	646	647
28	28.0000	28.0000	1.5832	1.7291	658	673

ภาคผนวก ฉ

ผลการสร้างตารางเทียบคะแนนโดยวิธีการเทียบคะแนนทั้งสองวิธี

ตาราง 19 การคำนวณคะแนนที่ได้จากการเทียบคะแนนทั้งสองวิธี

คะแนน	Z- SCORE คะแนน ดิบ	Z- SCORE คะแนน จริง	θ	คะแนน ดิบ	คะแนน จริง	CEEB คะแนน จริง	CEEB คะแนน ดิบ
0	-1.6953	-	-	0.1058	-	-	-
1	-1.6547	-	-	0.4361	-	-	-
2	-1.5788	-	-	1.0543	-	-	-
3	-1.4696	-	-	1.9443	-	-	-
4	-1.3153	-	-	3.2014	-	-	-
5	-1.1371	-1.1019	-4.0	4.6527	6.4545	390	386
6	-0.9435	-0.9365	-1.9	6.2298	7.6930	406	406
7	-0.7500	-0.7501	-1.0	7.8061	9.0863	425	425
8	-0.5701	-0.5791	-0.5	9.2715	10.3669	442	443
9	-0.4116	-0.4460	-0.2	10.5625	11.3670	455	459
10	-0.2756	-0.2876	0.1	11.6705	12.5531	471	472
11	-0.1582	-0.1672	0.3	12.6269	13.4552	483	484
12	-0.0536	-0.1024	0.4	13.4786	13.94075	490	495
13	0.0433	0.0363	0.6	14.2667	14.97957	504	504
14	0.1358	0.1099	0.7	15.0215	15.53051	511	514
15	0.2256	0.1856	0.8	15.7528	16.0995	519	523
16	0.3131	0.3431	1.0	16.4658	17.2770	534	531
17	0.3987	0.4230	1.1	17.1631	17.8759	542	540
18	0.4830	0.5028	1.2	17.8499	18.4734	550	548
19	0.5672	0.5816	1.3	18.5357	19.0631	558	557
20	0.6527	0.6584	1.4	19.2321	19.6388	566	565
21	0.7409	0.7327	1.5	19.9508	20.1949	573	574
22	0.8333	0.8037	1.6	20.7030	20.7267	580	583

ตาราง 19 (ต่อ)

คะแนน	Z- SCORE คะแนน ดีบ	Z- SCORE คะแนน จริง	θ	คะแนน ดีบ	คะแนน จริง	CEEB คะแนน จริง	CEEB คะแนน ดีบ
23	0.9313	0.9344	1.8	21.5011	21.7061	593	593
24	1.0374	1.0493	2.0	22.3657	22.5668	605	604
25	1.1563	1.1937	2.3	23.3342	23.6478	619	616
26	1.2958	1.3094	2.6	24.4703	24.5147	631	630
27	1.4698	1.5174	3.4	25.8877	26.0723	652	647
28	1.7291	1.6088	4.0	28.00	26.7567	661	673

ประวัติย่อของผู้วิจัย

ชื่อ นางสาวพรโสภา พิณีจผล

เกิดวันที่ 7 เดือนธันวาคม 2512

สถานที่เกิด อำเภอวังทอง จังหวัดพิษณุโลก

สถานที่อยู่ปัจจุบัน บ้านเลขที่ 5/95 หมู่ 2 ตำบลเนินพระ
อำเภอเมือง จังหวัด ระยอง

ตำแหน่งหน้าที่การงาน วิทยาจารย์ 4

สถานที่ทำงานปัจจุบัน วิทยาลัยพยาบาลบรมราชชนนี ชลบุรี
อำเภอเมือง จังหวัด ชลบุรี

ประวัติการศึกษา

- พ.ศ. 2531 มัธยมศึกษาตอนปลาย(สายวิทย์ - คณิต)
จากโรงเรียนวัดมโนทัยพาศย์ อำเภอเมือง จังหวัดเชียงใหม่
- พ.ศ. 2535 ประกาศนียบัตรพยาบาลศาสตร์
จากวิทยาลัยพยาบาลบรมราชชนนีลำปาง 1 จังหวัดลำปาง
- พ.ศ. 2539 กศ.ม. (การวัดผลการศึกษา)
จากมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

ผลการเทียบคะแนนในแบบทดสอบวิชาการพยาบาลสูติศาสตร์
ตามทฤษฎีการตอบข้อสอบที่ใช้วิธีการวิเคราะห์ต่างกัน

บทคัดย่อ
ของ
พรโสภา พิณิจผล

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
พฤษภาคม 2539

การศึกษาค้างนี้มีจุดมุ่งหมายเพื่อเปรียบเทียบผลการเทียบคะแนนในแบบทดสอบ
วิชาการพยาบาลสูติศาสตร์ตามทฤษฎีการตอบข้อสอบที่ใช้วิธีการวิเคราะห์ 2 วิธี คือ วิธีการ
เทียบคะแนนจริงโดยใช้แบบทดสอบร่วม และวิธีการเทียบคะแนนดิบโดยใช้
แบบทดสอบร่วม กลุ่มตัวอย่างเป็นนักศึกษาพยาบาลศาสตร์(ระดับต้น) ชั้นปีที่ 2 ปีการศึกษา
2538 ของวิทยาลัยพยาบาลในสังกัดสถาบันพัฒนากำลังคนด้านสาธารณสุข กระทรวง
สาธารณสุข ในเขตภาคเหนือ ภาคกลางและภาคตะวันออกเฉียงใต้ จำนวน 980 คน ซึ่งได้มาโดย
วิธีการสุ่มแบบอย่างง่าย เครื่องมือที่ใช้เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา
การพยาบาลสูติศาสตร์ 1 ที่ผู้วิจัยสร้างขึ้นจากการประมาณค่าพารามิเตอร์ของข้อสอบตาม
ทฤษฎีการตอบข้อสอบ โดยใช้โปรแกรม BILOG เวอร์ชัน 3.04 แล้วแบ่งออกเป็น 2
ฉบับ ๆ ละ 10 ข้อ

ผลการวิจัยสรุปได้ดังนี้

1. คะแนนที่ได้จากการเทียบคะแนนตามวิธีการเทียบคะแนนจริงกับวิธีการเทียบ
คะแนนดิบ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ
2. คะแนนความคลาดเคลื่อนที่ได้จากการเทียบคะแนนตามวิธีการเทียบคะแนนจริง
กับวิธีการเทียบคะแนนดิบ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ
3. ความเพียงพอของการเทียบคะแนนตามวิธีการเทียบคะแนนจริงกับวิธีการเที
ยคะแนนดิบ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ผู้วิจัยเสนอแนะวิธีการเทียบคะแนนโดยวิธีการเทียบคะแนนจริง เพราะให้คะแนน
ความคลาดเคลื่อนและความเพียงพอของการเทียบคะแนนไม่แตกต่างกับ วิธีการเทียบ
คะแนนดิบ ซึ่งเป็นวิธีการที่มีข้อคำนวณที่ซับซ้อนและลงทุนมากกว่า

A COMPARISON OF TWO ITEM RESPONSE THEORY EQUATING
PROCEDURES ON AN OBSTETRIC NURSING TEST SCORES

AN ABSTRACT

BY

PONSOPA PINIJPHON

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

May 1996

The purpose of this study was to compare the equating scores of obstetric nursing tests with different analysis methods according to two Item Response Theory Equating procedures ; three parameters logistic equivalent true scores and three parameters logistic raw scores methods. The cross-validation technique was used in the study to compare three-external criterion; scores of equating, standard error of equatings and discrepancy index. Total of 30 obstetric nursing test items as ante natal care were used to estimate item parameters by BILOG Version 3.04 computer program.

The two equivalent test forms were constructed each of 30 items including 10 anchor items. The subjects, who studying in the second year of College of nursing; in technical nursing level curriculum belonging to B. I. ,Ministry of Health, in the Northern, Eastern and the Middle of Thailand, were 980 student nurses. The major findings were 1) there were evidences of the requirement of Equating for two test score forms. 2) the procedures produced statistical insignificant equating score by the criterion of equating and discrepancy index. 3) less complexity in calculation, suggested more practical implication for True score IRT Equating procedure than Raw score IRT Equating procedure.