



รายงานการวิจัย

เรื่อง

ผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบ

และความเชื่อมั่นของข้อสอบทั้งฉบับ: รายวิชาวิทยาการระบาด

The Effects on Test Scores and Test Reliability after Removing Items

with Negative Discrimination Power: Epidemiology Subject

ดร.ศรีสุดา วงศ์วิเศษกุล

นางสาวมนชยา สมจรรย์

นางสาววาสนา จันทร์จ่าย

มหาวิทยาลัยสวนดุสิต

พ.ศ. 2558

ลิขสิทธิ์ของมหาวิทยาลัยสวนดุสิต



รายงานการวิจัย

เรื่อง

ผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบ

และความเชื่อมั่นของข้อสอบทั้งฉบับ: รายวิชาวิทยาการระบาด

The Effects on Test Scores and Test Reliability When Removing Items

with Negative Discrimination Power: Epidemiology Subject

ดร.ศรีสุดา วงศ์วิเศษกุล

(คณะพยาบาลศาสตร์)

นางสาวমনชยา สมจรรย์

(คณะพยาบาลศาสตร์)

นางสาววาสนา จันทร์จ่าย

(สำนักส่งเสริมวิชาการและงานทะเบียน)

มหาวิทยาลัยสวนดุสิต

พ.ศ. 2558

ลิขสิทธิ์ของมหาวิทยาลัยสวนดุสิต

(งานวิจัยนี้ได้รับทุนอุดหนุนจากมหาวิทยาลัยสวนดุสิต ปีงบประมาณ 2556)

หัวข้อวิจัย ผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบและความเชื่อมั่นของข้อสอบทั้งฉบับ: รายวิชาวิทยาการระบาด

ผู้ดำเนินการวิจัย อาจารย์ ดร.ศรีสุดา วงศ์วิเศษกุล
อาจารย์মনชยา สมจรรย์
นางสาววาสนา จันทร์จ่าย

หน่วยงาน คณะพยาบาลศาสตร์ มหาวิทยาลัยสวนดุสิต

ปี พ.ศ. 2558

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาคุณภาพของข้อสอบรายวิชาวิทยาการระบาด การเปลี่ยนแปลงของค่าความเชื่อมั่นของข้อสอบ การเปลี่ยนแปลงของคะแนนการสอบของนักศึกษารายบุคคลก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก และการเปลี่ยนแปลงของลำดับคะแนนจากการสอบของนักศึกษาในวิชาวิทยาการระบาดหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก กลุ่มตัวอย่างคือข้อสอบและคะแนนการสอบรายวิชาวิทยาการระบาด ภาคการศึกษาที่ 1 ปีการศึกษา 2555 เครื่องมือที่ใช้ในการวิจัยคือวิเคราะห์ข้อมูลโดยการวิเคราะห์ข้อสอบรายข้อตามแนวคิดทฤษฎีการวิเคราะห์แบบทดสอบแบบประเพณีนิยม สถิติ Paired t-test และสถิติ Signed Rank Test

ผลการวิจัย พบว่า

1. ข้อสอบกลางภาคร้อยละ 91.70 และข้อสอบปลายภาค ร้อยละ 82.50 มีดัชนีความยากง่าย (P) ระหว่าง .20 - .80
2. ข้อสอบกลางภาคร้อยละ 61.70 และ และข้อสอบปลายภาค ร้อยละ 56.25 มีค่าอำนาจจำแนก (R) \geq .20
3. ภายหลังจากตัดข้อสอบที่มีค่า r ติดลบออก ค่าความเชื่อมั่นของข้อสอบกลางภาคและปลายภาคทั้งฉบับเพิ่มขึ้น (จาก 0.554 เป็น 0.622 และจาก 0.694 เป็น 0.727 ตามลำดับ) ค่าเฉลี่ยคะแนนการสอบกลางภาคและปลายภาคเพิ่มขึ้น (จาก 55.36 เป็น 56.68 และ จาก 49.76 เป็น 51.19 ตามลำดับ) อย่างมีนัยสำคัญทางสถิติที่ $p = .001$ นักศึกษาส่วนใหญ่ (ร้อยละ 70.65 และร้อยละ 81.52) มีลำดับคะแนนการสอบเพิ่มขึ้นอย่างมีนัยสำคัญทางสถิติที่ $p = .001$ เช่นกัน

สรุปได้ว่า เมื่อตัดข้อสอบที่มีค่า r ติดลบออกทำให้มีการเปลี่ยนแปลงค่าความเชื่อมั่นของข้อสอบคะแนนสอบและลำดับคะแนนการสอบ

คำสำคัญ: คุณภาพข้อสอบ อำนาจจำแนกเป็นลบ ความเชื่อมั่นของข้อสอบชนิดหลายตัวเลือก

Research Title The Effects on Test Scores and Test Reliability after Removing Items with Negative Discrimination Power: Epidemiology Subject

Researcher Sresuda Wongwiseskul, Monchaya Somjarit, Wassana Janjai

Organization Faculty of Nursing, Suandusit University

Year 2015

This study aimed to evaluate the quality of test item, explore changes of test score and students' ranking, and changes of mid-term and final test reliability after removing the test items with negative discrimination power. Sample group was test and test items in Epidemiology, semester 1 educational year 2012. Research tool was test item analysis according to Classical test Theory. Difficulty Index, discrimination index, test reliability were calculated using computerized program. Paired t-test and Sign Rank test were applied to test the changes of test score and students' ranking.

Results of the study revealed that:

1. 91.70% of items in mid-term test and 82.50% of items in final test had index of difficulty (P) between .20 - .80

2. 61.70% of items in mid-term test and 56.25% of items in final test had discrimination power (R) \geq .20

3. After removing the test items with negative discrimination power, the reliability of mid-term and final test were increased (from 0.554 to 0.622 and from 0.694 to 0.727 respectively). Mean score of mid-term and final test were increased (from 55.36 to 56.68 and from 49.76 to 51.19 respectively) significantly at $p=.001$. 70.6% and 81.52% of students got higher ranking; the increasing was significantly at $p=.001$.

In conclusion, excluding test items with negative discrimination power affected on test reliability, mean score of the test, and score ranking.

Keywords: quality of test items, negative discrimination power, reliability of MEQ examination

กิตติกรรมประกาศ

การวิจัย เรื่องผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบและความเชื่อมั่นของข้อสอบทั้งฉบับ: รายวิชาวิทยาการระบาด มุ่งศึกษาคุณภาพข้อสอบในด้านความยากง่ายอำนาจจำแนก ความเชื่อมั่น การตัดข้อสอบที่มีค่าอำนาจจำแนกเป็นลบออกจะทำให้การคิดคะแนนมาจากคะแนนที่แท้จริงที่ผู้สอบทำได้ ผลการวิจัยจะนำไปสู่การคัดข้อสอบที่มีคุณภาพดีเพื่อจัดทำเป็นคลังข้อสอบต่อไป

การวิจัยนี้สำเร็จลงได้โดยสมบูรณ์ด้วยความกรุณาของ ดร.เบญจมา เตากล่ำ คณบดีคณะพยาบาลศาสตร์ อาจารย์วิมล เตากล่ำ และผู้ช่วยศาสตราจารย์ ดร.ประกาย จิโรจน์กุลที่ให้ข้อเสนอแนะอันทรงคุณค่า คณาจารย์ในสาขาการพยาบาลชุมชนที่สนับสนุนด้านเวลาและภาระงาน รวมทั้งผู้อำนวยการสถาบันวิจัยและพัฒนา (ดร.ชนะศึก นิขานนท์) และเจ้าหน้าที่ ผู้วิจัยขอขอบอกคุณมา ณ โอกาสนี้ และขอขอบอกคุณมหาวิทยาลัยสวนดุสิตที่ให้การสนับสนุนด้านทุนวิจัย

คณะผู้วิจัย

พ.ศ. 2558

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ก
บทคัดย่อภาษาอังกฤษ	ข
กิตติกรรมประกาศ	ค
สารบัญ	ง
สารบัญตาราง	ฉ
สารบัญภาพ	ช
บทที่ 1 บทนำ	1
ความเป็นมาและความสำคัญ	1
วัตถุประสงค์ของการวิจัย	2
ขอบเขตการวิจัย	2
คำจำกัดความที่ใช้ในงานวิจัย	3
ประโยชน์ที่คาดว่าจะได้รับ	3
บทที่ 2 แนวคิด ทฤษฎี เอกสารและงานวิจัยที่เกี่ยวข้อง	4
ทฤษฎีการวิเคราะห์ข้อสอบ	4
รายวิชาวิทยาการระบาด	8
บทความและงานวิจัยที่เกี่ยวข้อง	11
กรอบแนวคิดในการวิจัย	15
บทที่ 3 วิธีดำเนินการวิจัย	16
ประชากรและการเลือกกลุ่มตัวอย่าง	16
การเก็บรวบรวมข้อมูล	16
เครื่องมือในการวิจัยและการตรวจสอบคุณภาพเครื่องมือ	17
การวิเคราะห์ข้อมูล	17

	หน้า
บทที่ 4 ผลการวิจัย	17
ข้อมูลทั่วไป	17
คุณภาพข้อสอบของรายวิชาวิทยาการระบาด	18
การเปลี่ยนแปลงของคะแนนสอบรายวิชาและลำดับคะแนน	21
บทที่ 5 สรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ	22
สรุปผลการวิจัย	22
อภิปรายผล	22
ข้อเสนอแนะในการนำผลการวิจัยไปใช้	24
ข้อเสนอแนะในการทำวิจัยครั้งต่อไป	24
บรรณานุกรม	25
บรรณานุกรมภาษาไทย	25
บรรณานุกรมภาษาอังกฤษ	25
ภาคผนวก	
ตารางการกรอกข้อมูลวิชาวิทยาการระบาด ชั้นปีที่ 3	27
ประวัติผู้วิจัย	30

สารบัญตาราง

ตารางที่		หน้า
2.1	ข้อเสนอแนะในการเลือกข้อสอบจากค่าความยากง่าย	5
2.2	แผนการประเมินผลสัมฤทธิ์การเรียนรู้รายวิชาวิทยาการระบาด	9
2.3	การกระจายตัวของข้อสอบตามระดับการวัดผลการเรียนรู้	10
4.1	จำนวนนักศึกษาและจำนวนข้อสอบในแต่ละชุดข้อสอบ ที่ใช้ในการศึกษาวิจัย จำแนกตามชุดข้อสอบ รายวิชาวิทยาการระบาด	17
4.2	ค่าความยากง่าย ของข้อสอบกลางภาค วิชาวิทยาการระบาด	18
4.3	ค่าความยากง่าย ของข้อสอบปลายภาค วิชาวิทยาการระบาด	18
4.4	ค่าอำนาจจำแนก ของข้อสอบกลางภาค วิชาวิทยาการระบาด	19
4.5	ค่าอำนาจจำแนก ของข้อสอบปลายภาค วิชาวิทยาการระบาด	19
4.6	ค่าความเชื่อมั่นของข้อสอบวิชาวิทยาการระบาด ก่อน และหลังการตัดข้อสอบที่มี มีค่า r ตีลบออก	20
4.7	ผลการทดสอบการเปลี่ยนแปลงของคะแนนการสอบของนักศึกษาในวิชา วิทยาการระบาด หลังการตัดข้อสอบที่มีค่า r ตีลบออก	20
4.8	ผลการทดสอบการเปลี่ยนแปลงของคะแนนสอบและลำดับคะแนนจากการสอบ ของนักศึกษาในวิชาวิทยาการระบาด หลังการตัดข้อสอบที่มีค่าอำนาจจำแนกตี ลบออก	21

สารบัญภาพ

ภาพที่		หน้า
2.1	กรอบแนวคิดในการวิจัย	15

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญ

การวัดและประเมินผลเป็นองค์ประกอบที่สำคัญของการจัดการศึกษาที่มีคุณภาพและจุดมุ่งหมายของการวัดผลทางการศึกษา เพื่อให้ได้มาซึ่งข้อมูลที่เที่ยงตรง และเชื่อถือได้ ที่แสดงถึงความสามารถที่แท้จริง (True ability) ของผู้เรียน (อุทุมพร จามรมาน: 2535)

ข้อสอบเป็นเครื่องมือสำคัญในการประเมินว่าผู้เรียนมีความรู้ตามที่หลักสูตรคาดหวังหรือไม่ ข้อสอบที่มีคุณภาพ จึงเป็นปัจจัยสำคัญของการจัดการเรียนการสอน คุณสมบัติที่สำคัญของข้อสอบที่ดี คือ ข้อสอบที่มีค่าความยากง่ายเหมาะสม (p) อยู่ระหว่าง 0.2-0.8 และ ข้อสอบที่มีค่าอำนาจจำแนก (r) ที่เหมาะสม คือ ≥ 0.2

การจัดการเรียนการสอนในหลักสูตรพยาบาลศาสตรบัณฑิต ของมหาวิทยาลัยสวนดุสิต ได้ยึดหลักการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ และ ใช้ระเบียบการวัดและประเมินผลของมหาวิทยาลัย ควบคู่กับแนวทางการวัดและประเมินผลที่คณะได้กำหนดขึ้น คือ ใช้การประเมินผลแบบอิงเกณฑ์ในรายวิชาภาคปฏิบัติ ส่วนรายวิชาภาคทฤษฎี ใช้การประเมินผลแบบอิงเกณฑ์ควบคู่กับอิงกลุ่ม กล่าวคือ นักศึกษาต้องมีคะแนนดิบถึงเกณฑ์ 60% จึงจะได้รับเกรด C นอกจากนั้น ยังได้มีการพัฒนามาตรการในการช่วยเหลือผู้เรียนที่เรียนอ่อน เพื่อพัฒนานักศึกษาให้มีความรู้ความสามารถตามเกณฑ์ที่หลักสูตรกำหนด และได้พัฒนาเต็มตามศักยภาพของตน

ประสบการณ์ในการจัดการเรียนการสอนวิชาวิทยาการระบาด ในปีการศึกษา 2555 ที่ผ่านมาพบว่า นักศึกษาบางคนสอบไม่ผ่านเกณฑ์ขั้นต่ำ 60% ตามที่หลักสูตรกำหนด ต้องมีการสอบซ่อมทั้งการสอบกลางภาคและปลายภาค จากผลการวิเคราะห์ข้อสอบในปีการศึกษา 2554 พบว่า มีข้อสอบที่มีค่าอำนาจจำแนกติดลบ ร้อยละ 8.34 ซึ่งแสดงว่า คุณภาพของข้อสอบข้อนั้นๆไม่สามารถจำแนกผู้เรียนเก่งและอ่อนได้ ผู้เรียนอ่อนส่วนมากตอบข้อนั้นถูก แต่ผู้เรียนเก่งส่วนมากไม่ได้คะแนนจากข้อนั้น

ผู้วิจัยจึงเห็นความจำเป็นในการศึกษาคุณภาพของข้อสอบในรายวิชาระบาดวิทยา และเปรียบเทียบการเปลี่ยนแปลงของคะแนนการสอบของนักศึกษารายบุคคล ก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก พร้อมทั้งผลที่เกิดขึ้นต่อการเปลี่ยนแปลงค่าความเชื่อมั่นของข้อสอบทั้งฉบับ

เพื่อให้ได้ข้อมูลที่เป็นหลักฐานเชิงประจักษ์ ที่นำมาสู่ข้อสรุป หรือข้อเสนอแนะเชิงนโยบายในด้านการวัด และประเมินผลการจัดการเรียนการสอนในแต่ละรายวิชา ในหลักสูตรต่อไป

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาคุณภาพของข้อสอบรายวิชาวิทยาการระบาด
2. เพื่อศึกษาการเปลี่ยนแปลงของค่าความเชื่อมั่นของข้อสอบก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก
3. เพื่อศึกษาการเปลี่ยนแปลงของคะแนนการสอบของนักศึกษารายบุคคล ก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก
4. เพื่อศึกษาการเปลี่ยนแปลงของลำดับคะแนนจากการสอบของนักศึกษาในวิชาวิทยาการระบาด ภายหลังจากตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ขอบเขตการวิจัย

1. ขอบเขตด้านประชากร

ประชากรในการวิจัยนี้ ประกอบด้วยข้อสอบกลางภาคและปลายภาครายวิชาวิทยาการระบาด ปีการศึกษา 2555 สำหรับนักศึกษาพยาบาลชั้นปีที่ 2 ที่ลงทะเบียนเรียนจำนวน 92 คน

2. ขอบเขตด้านเวลา

การวิจัยนี้ รวบรวมข้อมูลเกี่ยวกับการวิเคราะห์อำนาจจำแนกของข้อสอบวิชาวิทยาการระบาด ที่จัดสอบกลางภาคและปลายภาคในปีการศึกษา 2555

3. ขอบเขตด้านเนื้อหา

คุณภาพของข้อสอบในการวิจัยนี้ประกอบด้วย

- 3.1 คุณภาพของข้อสอบรายข้อพิจารณาจากค่าความยากง่ายและอำนาจจำแนก
- 3.2 คุณภาพของข้อสอบทั้งฉบับพิจารณาจากค่าความเชื่อมั่น
- 3.3 คุณภาพของข้อสอบด้านความตรง (Validity) คือความตรงเชิงโครงสร้าง (Construct validity) พิจารณาจากความสอดคล้องทั้งจำนวนข้อสอบและประเด็นเนื้อหากับตาราง

วิเคราะห์การออกแบบการสร้างแบบทดสอบ (Test Blueprint) ที่ได้ออกแบบและแสดงไว้ใน มคอ. 3 รายวิชาระบบสุขภาพและการพยาบาล รวมทั้งการวิพากษ์ข้อสอบโดยผู้เชี่ยวชาญก่อนนำข้อสอบไปใช้จริง

คำจำกัดความที่ใช้ในงานวิจัย

คุณภาพของข้อสอบรายข้อ หมายถึง คุณภาพของข้อสอบในวิชาวิทยาการระดับกลางภาคและปลายภาค ที่ได้มาจากการวิเคราะห์ค่าความยากง่ายและอำนาจจำแนกเป็นรายข้อ

คุณภาพของข้อสอบทั้งฉบับ หมายถึง คุณภาพของข้อสอบในวิชาวิทยาการระดับกลางภาคและปลายภาค ที่ได้มาจากการวิเคราะห์ความเชื่อมั่น (KR - 20) ของข้อสอบทั้งฉบับ

ข้อสอบที่มีอำนาจจำแนกดี หมายถึง ข้อสอบรายข้อที่ได้วิเคราะห์ค่าอำนาจจำแนก (r) แล้วได้ค่าน้อยกว่า 0

การตัดข้อสอบที่มีค่าอำนาจจำแนกดีออก หมายถึง การนำข้อสอบที่มีอำนาจจำแนก (r) เป็นลบออกก่อนนำไปวิเคราะห์คุณภาพข้อสอบในครั้งที่สอง

คะแนนการสอบ หมายถึงคะแนนที่ได้จากการตอบแบบทดสอบชนิดเลือกตอบ (Multiple Choices Question) 4 ตัวเลือก และมีคำตอบที่ถูกต้องที่สุดเพียงข้อเดียว ข้อที่ตอบถูกต้องคะแนน 1 ตอบผิดได้คะแนน 0

ประโยชน์ที่คาดว่าจะได้รับ

1. ได้แนวทางในการวิเคราะห์ข้อสอบในรายวิชาอื่นๆ ที่ใช้แบบทดสอบชนิดเลือกตอบ (Multiple Choices Question Examination)
2. ได้แนวทางการคัดเลือกข้อสอบที่มีคุณภาพเพื่อพัฒนาสู่ข้อสอบมาตรฐาน

บทที่ 2

แนวคิด ทฤษฎี เอกสารและงานวิจัยที่เกี่ยวข้อง

วรรณกรรมที่เกี่ยวข้องกับการวิจัยเรื่อง ผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบและความเชื่อมั่นของข้อสอบทั้งฉบับ: รายวิชาวิทยาการระบาด ประกอบด้วย

1. ทฤษฎีการวิเคราะห์ข้อสอบ
 2. รายวิชาวิทยาการระบาด
 3. บทความและงานวิจัยที่เกี่ยวข้อง
- ดังรายละเอียดต่อไปนี้

ทฤษฎีการวิเคราะห์ข้อสอบ

เชิดศักดิ์ ไอรณรัตน์ (2552) กล่าวถึงแนวคิดสำคัญในการวิเคราะห์ข้อสอบ โดยระบุว่าประกอบด้วย 2 ส่วนคือ การวิเคราะห์ข้อสอบรายข้อ (Item analysis) และการวิเคราะห์ข้อสอบโดยรวม (Test analysis) โดยมีรายละเอียดพอสังเขปดังนี้

1. การวิเคราะห์ข้อสอบรายข้อ (Item analysis) การวิเคราะห์ข้อสอบแต่ละข้อพิจารณา 3 ปัจจัย คือ

1.1 ความยากง่ายของข้อสอบ (Item difficulty, p) ความยากง่ายของข้อสอบวัดโดยใช้ค่า p ซึ่งย่อมาจาก Proportion of examinees answering item correctly (สัดส่วนของผู้ที่ตอบข้อสอบนั้นถูก) ซึ่งหาได้จากการนำจำนวนผู้สอบที่ตอบข้อนั้นถูก ต่อดังด้วยจำนวนผู้สอบที่ตอบข้อนั้นทั้งหมด หากข้อสอบข้อนั้นเป็นข้อสอบที่ง่ายผู้สอบทุกคนตอบถูก ค่า p ก็จะเป็น 1 หากไม่มีผู้สอบคนใดตอบถูกเลย ข้อสอบข้อนั้นก็จะมีค่า p เป็น 0 หากมีคนตอบถูก 70% ข้อสอบข้อนั้นก็จะมีค่า p เท่ากับ 0.7

สุมาลี จันทร์ชลอ (2542: 136) ได้ให้ข้อเสนอแนะในการเลือกข้อสอบจากการพิจารณา ค่า P ดังตารางที่ 2.1

ตารางที่ 2.1 ข้อเสนอแนะในการเลือกข้อสอบจากค่าความยากง่าย

ค่าความยากง่าย (P)	ความหมาย	ข้อเสนอแนะ
.81-1.00	ง่ายมาก	ควรตัดทิ้ง
.61-.80	ค่อนข้างง่าย	ดีพอใช้ ควรเก็บไว้ใช้
.41-.60	ความยากง่ายพอเหมาะ	ดีมากเก็บไว้ใช้
.20-.40	ค่อนข้างยาก	ดีพอใช้ ควรเก็บไว้ใช้
.00-.19	ยากมาก	ควรตัดทิ้ง

สูตรในการคำนวณค่าความยากง่าย (พวงรัตน์ ทวีรัตน์ 2540: 129)

$$P = R / N$$

เมื่อ P = ค่าความยากง่ายของคำถามข้อนั้น
 R = จำนวนผู้ตอบถูกในข้อนั้น
 N = จำนวนผู้ตอบทั้งหมด

1.2 ความสามารถในการจำแนกผู้สอบ หมายถึง ความสามารถของข้อสอบข้อหนึ่งๆ ในการแยกผู้สอบที่ทำคะแนนได้ดี ออกจากผู้สอบที่ทำคะแนนได้ไม่ดี ข้อสอบที่มีความสามารถในการแยกแยะได้ดีนั้น ผู้สอบที่ตอบข้อสอบข้อนั้นผิดมักจะได้คะแนนต่ำ ดัชนีที่ใช้วัดความสามารถในการจำแนกผู้สอบที่ใช้กันมากที่สุดในปัจจุบันคือค่า point-biserial correlation ซึ่งนิยมใช้อักษรย่อเป็น r ซึ่งสามารถคำนวณได้จากสูตรต่อไปนี้ (เชิดศักดิ์ ไอรณรัตน์, 2552)

$$r = \frac{M_p - M_q}{SD} \sqrt{pq}$$

เมื่อ M_p = คะแนนรวมเฉลี่ยของผู้สอบที่ตอบข้อสอบถูก
 M_q = คะแนนสอบเฉลี่ยของผู้สอบที่ตอบข้อสอบผิด

- SD = ค่าเบี่ยงเบนมาตรฐานของคะแนนสอบ
 P = สัดส่วนของผู้สอบที่ตอบข้อสอบถูกต้องผู้สอบทั้งหมด
 q = สัดส่วนของผู้สอบที่ตอบข้อสอบผิดต่อผู้สอบทั้งหมด

ค่า point-biserial correlation ที่คำนวณได้มีค่าอยู่ในช่วง -1 ถึง 1 โดยค่าที่ติดลบหมายถึงข้อสอบข้อนั้นผู้ที่ตอบถูกมักสอบได้คะแนนรวมต่ำ แต่ผู้ที่ตอบผิดมักสอบได้คะแนนรวมสูง ในทางตรงกันข้าม หากค่า point-biserial ยิ่งสูงแสดงถึงข้อสอบที่มีความสามารถในการแยกแยะดี ผู้ที่ตอบข้อสอบข้อนั้นถูกมักทำคะแนนรวมได้สูง ข้อสอบที่ดีควรมีค่า point-biserial สูงกว่า 0.2 ข้อสอบที่ดีพอใช้ควรมีค่า point-biserial อยู่ในช่วง 0.1 – 0.19 ข้อสอบที่มีค่า point-biserial ต่ำกว่า 0.1 เป็นข้อสอบที่ไม่ดีนัก โดยเฉพาะข้อสอบที่มีค่า point-biserial ต่ำกว่า 0 ไม่ควรนำมาคิดคะแนน

1.3 ประสิทธิภาพของตัวลวง (distractor functionality) ตัวลวงที่มีประสิทธิภาพนั้นมีคุณสมบัติ 2 ประการคือ

1.3.1 มีผู้สอบเลือกตัวลวงนั้นไม่ต่ำกว่า ร้อยละ 5 ของจำนวนผู้สอบทั้งหมด

1.3.2 มีค่า point-biserial correlation ของตัวลวงนั้นเป็นลบ กล่าวคือตัวลวงที่ดีจะลวงให้ผู้สอบที่มีความรู้ไม่ดี (มีคะแนนต่ำ) มาเลือก หากตัวลวงใดมีค่า point-biserial correlation เป็นบวก ให้ทบทวนข้อสอบข้อนั้นดูว่าอาจเฉลยผิด หรือมีคำตอบที่ถูกต้องมากกว่า 1 ตัวเลือก ตัวลวงใดที่มีผู้สอบเลือกน้อย หรือลวงให้ผู้ที่มีความรู้ดีมาเลือกจัดเป็นตัวลวงที่ไม่ดี สมควรพิจารณาตัดทิ้งหรือปรับเปลี่ยน

2. การวิเคราะห์ข้อสอบโดยรวม (Test-analysis)

การวิเคราะห์ข้อสอบโดยรวมเป็นการพิจารณาว่าเมื่อข้อสอบทั้งชุดทำงานร่วมกันแล้วผลสอบที่ได้ออกมาเป็นอย่างไร มีระดับความยากง่ายเป็นอย่างไร มีการกระจายตัวของคะแนนเป็นอย่างไร มีความน่าเชื่อถือของคะแนนสอบมากน้อยเพียงใด ดัชนีต่างๆที่ต้องพิจารณาได้แก่

2.1 ความเที่ยงตรงของคะแนนสอบ (Internal consistency reliability) เป็นการตรวจสอบว่าคะแนนที่ได้ออกมานั้นมีความน่าเชื่อถือเพียงใด เป็นการตอบคำถามว่าหากนำผู้สอบมาสอบใหม่ในสภาวะการเดิม ด้วยข้อสอบที่มีระดับความยากง่ายเท่าเดิม และผู้สอบมีความรู้เท่าเดิมไม่ได้ไปศึกษาเพิ่มเติม จะได้คะแนนสอบเท่าเดิมหรือไม่

ดัชนีชี้วัดความเที่ยงตรงของคะแนนสอบที่นิยมใช้ในการรายงานผลสอบด้วยข้อสอบปรนัยคือค่าสัมประสิทธิ์อัลฟา (Coefficient Alpha) ซึ่งสามารถคำนวณได้จากสูตร

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum Q_{xi}^2}{Q_x^2} \right)$$

- เมื่อ α = สัมประสิทธิ์ อัลฟา (Coefficient Alpha)
 n = จำนวนชุดย่อยของข้อสอบที่ทำการแบ่งออกเพื่อหาความเที่ยง
 Q_x^2 = การกระจายตัว (variance) ของคะแนนรวม
 Q_{xi}^2 = การกระจายตัว (variance) ของคะแนนข้อสอบย่อยชุดที่ i

ค่าสัมประสิทธิ์อัลฟานี้มีค่าอยู่ในช่วง 0 - 1 ค่าต่ำแสดงว่าคะแนนที่ได้มีความเชื่อถือได้น้อย ไม่แตกต่างไปจากการเดาสุ่ม ค่าสูงแสดงว่าคะแนนที่ได้นั้นมีความน่าเชื่อถือมาก หากทำการทดสอบซ้ำคะแนนที่ได้ก็จะใกล้เคียงเดิม โดยทั่วไประดับของความเที่ยงตรงของคะแนนสอบที่ยอมรับได้นั้นขึ้นกับว่าต้องการนำคะแนนสอบไปใช้ทำอะไร หากการตัดสินผลสอบนั้นมีความสำคัญมากต้องการคะแนนสอบที่มีค่าสัมประสิทธิ์อัลฟาไม่ต่ำกว่า 0.9 หากการตัดสินผลสอบนั้นมีความสำคัญปานกลางต้องการคะแนนสอบที่มีค่าสัมประสิทธิ์อัลฟาในช่วง 0.8 - 0.89 หากการตัดสินผลสอบนั้นมีความสำคัญน้อย ต้องการคะแนนสอบที่มีค่าสัมประสิทธิ์อัลฟาในช่วง 0.7 - 0.79

ประเด็นสำคัญที่ต้องพิจารณาคือเมื่อได้คะแนนสอบที่มีค่าสัมประสิทธิ์อัลฟาต่ำ จะต้องดำเนินการอย่างไรเพื่อพัฒนาให้การสอบครั้งต่อไปไม่ประสบปัญหาเรื่องความไม่น่าเชื่อถือของคะแนนสอบ ปัจจัยหลักที่จะช่วยความเที่ยงตรงของคะแนนสอบปรนัยมี 3 ปัจจัยคือ

1. เพิ่มจำนวนข้อสอบให้มากขึ้น ยิ่งมีข้อสอบมากข้อคะแนนที่ได้จะมีความเที่ยงตรงเพิ่มมากขึ้น
2. ปรับให้ข้อสอบมีการคละกันของข้อสอบที่ยากและง่ายอย่างเหมาะสม เพื่อปรับให้คะแนนมีการกระจายตัวมากขึ้น หากข้อสอบทั้งชุดประกอบไปด้วยข้อสอบที่ง่ายหมด ผู้สอบเกือบทั้งหมดได้คะแนนสูงมาก จะทำให้มีความแตกต่างของคะแนนน้อย โอกาสที่แยกแยะผู้สอบที่มีความรู้ดีออกจากผู้ที่มีความรู้ปานกลาง หรือไม่ดีได้อย่างมั่นใจเป็นไปได้น้อย ดังนั้นหากอาจารย์ปรับให้มีการคละกันของข้อสอบยากและง่ายอย่างเหมาะสม จะทำให้ผู้สอบมีระดับคะแนนแตกต่างกันมาก ค่าสัมประสิทธิ์อัลฟาก็จะสูงมากขึ้น

3. ปรับสภาวะแวดล้อมของการสอบให้เหมาะสม กำจัดสิ่งรบกวนสมาธิของผู้สอบให้เหมาะสม เช่น เสียงรบกวน แสงไฟที่ไม่เพียงพอ หรือไฟติดๆดับๆ เป็นต้น

2.2 การกระจายตัวของคะแนน และคะแนนเฉลี่ย (Standard deviation and mean Score) การตรวจสอบดูลักษณะพื้นฐานของคะแนนสอบนี้ จะช่วยบอกได้คร่าวๆว่าการเรียนการสอนมีประสิทธิภาพเพียงใด หากนักเรียนทั้งชั้นเรียนเข้าใจเนื้อหาดี คะแนนสอบที่ได้ออกมาควรจะไม่มีกระจายตัวมากนัก (คะแนนเกาะกลุ่มกัน) และคะแนนเฉลี่ยก็ควรจะค่อนข้างสูง หากคะแนนสอบของนักเรียนควรมีการกระจายตัวมากผิดปกติ แสดงว่าอาจมีปัญหาบางประการในการจัดการเรียนการสอน ทำให้นักเรียนบางคนมีความรู้ความเข้าใจดี แต่มีนักเรียนบางกลุ่มที่ไม่ค่อยเข้าใจ

2.3 ค่าความยากง่ายเฉลี่ยของข้อสอบ (Average difficult) จากการวิเคราะห์ข้อสอบราย ข้อ หลังจากได้ค่าความยากง่ายของข้อสอบแต่ละข้อ (p) เมื่อนำค่า p ของข้อสอบทุกข้อมาหาค่าเฉลี่ย จะได้ค่าความยากง่ายของข้อสอบทั้งหมด ค่าที่ได้จะใช้เป็นค่าดัชนีชี้วัดว่าข้อสอบทั้งหมดโดยรวมนั้นมีระดับความยากง่ายอย่างไร หากผู้สอบเป็นนักศึกษากลุ่มใหญ่พอที่จะตั้งสมมติฐานว่าระดับความสามารถมีการกระจายตัวอย่างเหมาะสม และไม่ต่างจากระดับความสามารถเฉลี่ยของกลุ่มผู้สอบปีก่อนๆ เราสามารถนำค่าความยากง่ายของข้อสอบทั้งหมดนี้มาเทียบได้ว่าข้อสอบที่ใช้มีความยากง่ายแตกต่างจากปีก่อนๆหรือไม่ ซึ่งอาจารย์อาจนำข้อมูลนี้มาใช้พิจารณาปรับเกณฑ์การตัดเกรดว่าต้องมีการปรับระดับคะแนนที่ได้เกรดต่างๆหรือไม่ อย่างไร

2.4 ค่าความสามารถในการแยกแยะผู้สอบเฉลี่ย (Average discrimination) การนำค่า point - biserial correlation ของข้อสอบทั้งหมดมาหาค่าเฉลี่ย เป็นการบอกคร่าวๆว่าโดยรวมแล้วข้อสอบชุดนี้มีความสามารถในการแยกแยะผู้สอบตามระดับความสามารถเพียงใด ยิ่งได้ค่าสูงก็ยิ่งดี แต่มีข้อควรระวังในการแปลผลในกรณีที่การเรียนการสอนเป็นไปได้ดี และผู้สอบทั้งหมด หรือเกือบทั้งหมดทำคะแนนได้สูง ค่า point - biserial correlation เฉลี่ยของข้อสอบทั้งหมดจะไม่สูงแต่ไม่ได้แปลว่าข้อสอบที่ใช้มีคุณภาพไม่ดี

รายวิชาวิทยาการระบาด

รายวิชานี้จัดการเรียนการสอนในชั้นปีที่ 2 เป็นวิชาภาคทฤษฎี ครอบคลุมวิชาพื้นฐานทางวิชาชีพและเป็นวิชาภาคบังคับ มีสาระเนื้อหาเกี่ยวกับแนวคิด หลักการที่เกี่ยวข้องกับวิทยาการระบาด รวมถึงธรรมชาติของการเกิดโรค การกระจายโรคในชุมชน การคำนวณและวิเคราะห์ สถิติชีพ ดัชนีชี้วัดภาวะสุขภาพของชุมชน บทบาทของพยาบาลในการคัดกรอง การเฝ้าระวัง การป้องกัน การควบคุมโรคและความพิการในชุมชน กิจกรรมการเรียนการสอนประกอบด้วย 1) การบรรยายประกอบสไลด์ PowerPoint

2) การมอบหมายให้นักศึกษาจัดทำรายงาน 2 ฉบับคือรายงานการวิเคราะห์สถานการณ์ด้านวิทยาการระบาดและรายงานการวิเคราะห์บทความด้านวิทยาการระบาด แผนการประเมินผลแสดงดังตารางที่ 2.2

ตารางที่ 2.2 แผนการประเมินผลลัพธ์การเรียนรู้รายวิชาวิทยาการระบาด

วิธีการประเมินผล/ผลการเรียนรู้ที่คาดหวัง	สัปดาห์ที่กำหนด	สัดส่วนของการประเมินผล
1. สอบ	สัปดาห์ที่ 9 และ 15	ร้อยละ 70
2. ประเมินพฤติกรรมด้านคุณธรรมและวินัย	ตลอดภาคการศึกษา	ร้อยละ 5
3. รายงาน 2 ฉบับ (ประกอบด้วย		
3.1 รายงานการวิเคราะห์สถานการณ์ด้านวิทยาการระบาด	สัปดาห์ที่ 1 - 3	ร้อยละ 10
3.2 รายงานการวิเคราะห์บทความด้านวิทยาการระบาด	สัปดาห์ที่ 3 - 6	ร้อยละ 10
4. การนำเสนอรายงาน อภิปราย และแสดงความคิดเห็นในชั้นเรียน	ตลอดภาคการศึกษา	ร้อยละ 5
	รวม	ร้อยละ 100

การประเมินผล

การประเมินผลการเรียนใช้แบบอิงกลุ่ม และอิงเกณฑ์ โดยใช้เกณฑ์ 60% ขึ้นไปสำหรับระดับคะแนน C กำหนดการกระจายตัวของข้อสอบตามระดับการวัดผลการเรียนรู้แสดงดังตารางที่ 2.3

ตารางที่ 2.3 การกระจายตัวของข้อสอบตามระดับการวัดผลการเรียนรู้

ขอบเขตเนื้อหา	จำนวนข้อสอบตามระดับการเรียนรู้				รวม (ข้อ)
	รู้-จำ	เข้าใจ	นำไปใช้	วิเคราะห์	
หน่วยที่ 1 ความรู้เบื้องต้นเกี่ยวกับวิทยาการระบาด (Introduction to Epidemiology)	1	5	1	3	10
หน่วยที่ 2 ปัจจัยสามทางระบาดวิทยา (Epidemiologic triad)	2	2	2	4	10
หน่วยที่ 3 ธรรมชาติของการเกิดโรค	2	2	2	4	10
หน่วยที่ 4 ปัจจัยด้านประชากร สถานที่ และเวลา กกับการเกิดและการกระจายโรคในชุมชน	2	3	3	2	10
หน่วยที่ 5 ดัชนีอนามัย	4	6	6	4	20
หน่วยที่ 6 การศึกษาทางระบาดวิทยา : นิยาม วัตถุประสงค์ ประโยชน์และการวัดอัตราเสี่ยงของการเกิดโรค	4	12	14	-	30
หน่วยที่ 7 กลวิธีระบาดวิทยา	-	10	10	-	20
หน่วยที่ 8 บทบาทของพยาบาลในการคัดกรองโรค	-	1	4	5	10
หน่วยที่ 9 ระบาดวิทยากับการบริการพยาบาลระดับปฐมภูมิ	2	1	4	3	10
หน่วยที่ 10 วิทยาการระบาดกับการบริการพยาบาลระดับทุติยภูมิและตติยภูมิ	2	1	4	3	10
รวม	19	43	50	28	140

การประเมินผลด้วยแบบสอบเป็นข้อสอบปรนัยชนิดหลายตัวเลือก (Multiple choice test) 4 ตัวเลือกและมีคำตอบที่ถูกต้องที่สุดเพียงข้อเดียว เนื้อหาข้อสอบเป็นไปตามคำอธิบายรายวิชาที่กำหนดไว้ใน

หลักสูตร กำหนดวัตถุประสงค์การเรียนรู้เชิงพุทธิพิสัยไว้ 4 ระดับ คือ รู้-จำ เข้าใจ นำไปใช้และวิเคราะห์ การกระจายระดับของการวัดผลการเรียนรู้ขึ้นกับธรรมชาติของเนื้อหาและสัปดาห์ของการเรียน

บทความและงานวิจัยที่เกี่ยวข้อง

Burton (2004) ได้เขียนบทความเรื่องการหาดัชนีอำนาจจำแนกของข้อสอบช่วยปรับปรุงคุณภาพของข้อสอบได้จริงหรือไม่ สามารถสรุปบทความได้ว่าแบบทดสอบที่มีข้อคำตอบให้เลือกเป็นรายข้อและให้คะแนนเป็น 0 (ศูนย์) และ 1 จะมีคุณภาพหรือไม่สามารถบอกได้จากการหาดัชนีอำนาจจำแนกของข้อสอบแต่ละข้อ และเปรียบเทียบดัชนีกลุ่มสูงและกลุ่มต่ำได้ (U-L Index) เพราะว่าจำนวนผู้ที่ตอบแบบสอบถามถูกแต่ละข้อจะถูกแบ่งเป็น 2 กลุ่มคือกลุ่มที่มีคะแนนสูงสุดและกลุ่มที่มีคะแนนต่ำสุด U-L Index จะไม่มีความเที่ยงถ้าไม่ใช้กับกลุ่มทดสอบจำนวนมากและการกระจายของคะแนนรวม และไม่สามารถนำไปเปรียบเทียบกับแบบทดสอบแต่ละชุดได้ การหาค่าสหสัมพันธ์ของคะแนนแต่ละข้อและคะแนนรวม จะมีความตรงหรือน่าเชื่อถือมากกว่าและยอมรับในปัจจุบันแต่ก็ยังมีข้อบกพร่องที่คล้ายกัน เพราะว่าไม่สามารถแทนที่ในเรื่องของการระมัดระวังการใช้ถ้อยคำในแต่ละข้อคำถาม ข้อสอบจะมีความเที่ยงสูงเมื่อนำข้อที่มีค่าอำนาจจำแนกต่ำหรือไม่ดีออกไป นอกจากนี้ความน่าเชื่อถือของดัชนีอำนาจจำแนกของข้อสอบ ต้องคำนึงถึงผลกระทบของการเดาของผู้ตอบแบบทดสอบและความสมบูรณ์ของคำถาม ซึ่งสามารถค้นหาได้อย่างง่ายโดยใช้รูปแบบการประมวลผลทางคอมพิวเตอร์เป็นเครื่องมือในการสืบค้นและเรียนรู้

Phipps et al (2009) ได้ศึกษาความสัมพันธ์ระหว่างข้อคำถามชนิดกรณีศึกษา (Case based) กับ ไม่ใช่กรณีศึกษา (non-case based) ในแบบทดสอบชนิดเลือกตอบ (Multiple choice items) กับ ผลการวิเคราะห์ค่าความยากง่าย และดัชนีอำนาจจำแนก โดยแบบทดสอบที่นำมาวิเคราะห์เป็นชุดข้อสอบชนิดเลือกตอบ 4 – 5 ตัวเลือก วิชา การรักษา (Therapeutics) ที่ใช้สอนนักศึกษาหลักสูตรปริญญาเอกทางเภสัชศาสตร์ 4 ปี (PharmD) คณะเภสัชศาสตร์ ชั้นปีที่ 2 และปีที่ 3 ในช่วงปี ค.ศ. 2004 - 2005 ถึง 2007 - 2008 คณะเภสัชศาสตร์ มหาวิทยาลัยเซเนนโดห์ (Shenandoah university) ผลการศึกษาพบว่า การวิเคราะห์ความยากง่าย ในคำถามชนิดกรณีศึกษา ($p=76.51$, $SD=19.2$) และชนิดไม่ใช่กรณีศึกษา ($p=76.86$, $SD=18.5$) ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($p=0.75$) แต่พบว่า ดัชนีอำนาจจำแนก ในคำถามชนิดไม่ใช่กรณีศึกษา ($r=0.250$, $SD = 0.1$) สูงกว่า คำถามชนิดกรณีศึกษา ($r=0.227$, $SD=0.1$) อย่างมีนัยสำคัญทางสถิติ ($p < 0.01$) และพบว่า ในคำถามชนิดไม่ใช่กรณีศึกษากลุ่ม K-type multiple choiceมีค่าความยากง่าย ($p=72.01$, $SD=20.4$) สูงกว่า กลุ่มคำถามชนิดเลือกตอบมาตรฐาน

($p=76.28$, $SD=18.4$) อย่างมีนัยสำคัญทางสถิติ ($p < 0.01$) แต่ดัชนีอำนาจจำแนกทั้ง 2 กลุ่ม ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ (k -type $r = 0.231$, $SD=0.1$, standard $r=0.246$, $SD = 0.1$) ที่ $p=0.21$ นอกจากนี้เมื่อวิเคราะห์ข้อสอบชนิดที่มี 4 ตัวเลือก และ 5 ตัวเลือก เปรียบเทียบกัน พบว่า ข้อสอบชนิด 5 ตัวเลือก ($p=75.04$, $SD = 19.0$, $r=0.262$, $SD=0.1$) ยากกว่า และมีค่าดัชนีอำนาจจำแนกสูงกว่า ชนิด 4 ตัวเลือก ($p=79.16$, $SD=17.9$, $r=0.221$, $SD=0.1$) อย่างมีนัยสำคัญทางสถิติ ($p < 0.001$)

Tasdemir (2010) ได้ทำการศึกษาเพื่อเปรียบเทียบความยากง่าย อำนาจการจำแนกในการประเมินผลสัมฤทธิ์ทางการเรียนระหว่างข้อสอบแบบหลายตัวเลือกกับแบบถูกผิด รูปแบบการวิจัยเป็นแบบเชิงพรรณนา กลุ่มตัวอย่างคือ นักศึกษาระดับปริญญาตรีชั้นปีที่สี่จำนวน 252 คน ในปีการศึกษา 2007 – 2008 เครื่องมือที่ใช้เก็บข้อมูลเป็นข้อสอบปลายภาครวม 100 ข้อ โดยแบ่งเป็นข้อสอบแบบหลายตัวเลือก 50 ข้อ และแบบถูกผิดจำนวนเท่ากันคือ 50 ข้อ ซึ่งมีโครงสร้างเนื้อหาแบบคู่ขนานกันและผ่านการตรวจสอบจากผู้ทรงคุณวุฒิในด้านความตรงตามเนื้อหา หาค่าความเที่ยงจากความสอดคล้องภายในโดยใช้สูตร Kuder - Richardson (KR 20) ได้ .779 และ Spearman - Brown ได้ .782 ซึ่งถือว่าเป็นแบบทดสอบผลสัมฤทธิ์ทางการเรียนที่มีค่าความน่าเชื่อถือสูง วิเคราะห์ข้อมูลด้วยโปรแกรมสำเร็จรูป SPSS สถิติที่ใช้คือ Paired - t test เพื่อเปรียบเทียบความแตกต่างระหว่างข้อสอบทั้งสองประเภท โดยถือว่าข้อสอบที่ดีควรมีค่าอำนาจการจำแนกมากกว่า 0.19 ขึ้นไป ส่วนข้อสอบที่มีความยากง่ายพอเหมาะมีค่าความยากง่ายประมาณ 0.5 ผลการวิจัยพบว่า

1. คะแนนเฉลี่ยที่ได้จากการทดสอบด้วยข้อสอบแบบหลายตัวเลือกและแบบถูกผิดในเพศชายและหญิงไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ
2. คะแนนเฉลี่ยที่ได้จากการทดสอบแบบหลายตัวเลือกและแบบถูกผิด 28 หัวข้อที่มีโครงสร้างเนื้อหาแบบคู่ขนานกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ กล่าวคือ มีข้อสอบ 13 ข้อที่ผู้เข้าสอบทำคะแนนจากแบบทดสอบแบบหลายตัวเลือกได้สูงกว่าแบบถูกผิด ส่วนอีก 15 ข้อนั้นผู้เข้าสอบทำคะแนนจากแบบทดสอบแบบถูกผิดได้สูงกว่า ในขณะที่อีก 22 ข้อไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ
3. ความยากง่ายของแบบทดสอบแบบหลายตัวเลือกและแบบถูกผิดไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ (ข้อสอบที่มีความยากคือ $1 - P \geq 0.5$) โดยพบว่าข้อสอบแบบหลายตัวเลือกมีข้อสอบที่ค่อนข้างยาก 10 ข้อ ส่วนอีก 40 ข้อค่อนข้างง่าย ส่วนข้อสอบแบบถูกผิดมี 5 ข้อที่ค่อนข้างยาก 45 ข้อค่อนข้างง่าย
4. อำนาจการจำแนกของแบบทดสอบแบบหลายตัวเลือกและแบบถูกผิด ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ โดยข้อสอบแบบหลายตัวเลือกมีค่าอำนาจการจำแนกต่ำ มี 10 ข้อที่มีค่าอำนาจ

การจำแนกอยู่ในเกณฑ์ปกติหรือใช้ได้ 40 ข้อ ส่วนข้อสอบแบบถูกผิดมีค่าอำนาจการจำแนกต่ำ 14 ข้อที่มีค่าอำนาจการจำแนกอยู่ในเกณฑ์ปกติใช้ได้ 36 ข้อ

Oluseyi & Ajeigbe (2012) ได้ทำการวิจัยเรื่องการวิเคราะห์ข้อสอบชนิดหลายตัวเลือกในวิชาเคมีพื้นฐานของมหาวิทยาลัยในไนจีเรียการวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์ความยาก - ง่ายและอำนาจจำแนกของข้อสอบชนิดหลายตัวเลือกในวิชาเคมีพื้นฐาน รูปแบบการวิจัยเป็นแบบการวิเคราะห์ย้อนหลังเพื่อเปรียบเทียบหาสาเหตุ กลุ่มตัวอย่างคือนักศึกษาจำนวน 800 คน ที่ได้มาจากการสุ่มตัวอย่างแบบง่ายจากจำนวนประชากรที่เป็นนักศึกษาที่เรียนวิชาเคมี 102 ในปีการศึกษา 2008 - 2009 ข้อสอบชนิด 4 ตัวเลือกจำนวน 40 ข้อสร้างขึ้นจากเนื้อหาในรายวิชาดังกล่าว โดยอาจารย์ผู้ร่วมสอนในแต่ละหัวข้อในรายวิชานั้น ผู้วิจัยคัดเลือกนักศึกษาจำนวน 442 คนที่ได้คะแนนในเปอร์เซ็นต์ที่ 30 ทั้งจากคะแนนที่น้อยที่สุดและมากที่สุดมาเป็นหน่วยในการวิเคราะห์ สถิติที่ใช้คือค่าความถี่ ค่าความยาก - ง่าย และค่าอำนาจจำแนกผลการศึกษาพบว่า ค่าความยาก - ง่ายรวมทั้งฉบับเท่ากับ 0.54 ค่าอำนาจจำแนกเท่ากับ 0.32 ตัวเลือกที่นักศึกษาเลือกตอบมากที่สุด คือ ตัวเลือกที่ 4 ซึ่งพบถึง 16 ข้อ ในจำนวนข้อสอบที่ 40 ข้อ พบข้อสอบที่มีความผิดพลาดในการวัด 3 ข้อ เนื่องจากอำนาจจำแนกเท่ากับ 0 ซึ่งผู้วิจัยอภิปรายว่าอาจจะเกิดจากการสร้างข้อสอบไม่ถูกต้องหรือผู้ตรวจเฉลยผิด นอกจากนี้ พบด้วยว่า ข้อสอบข้อที่ยากที่สุดจำนวน 1 ข้อ มีค่าความยาก - ง่ายเท่ากับ 0.245 และมีค่าอำนาจจำแนกเท่ากับ 0.23 ซึ่งมีนักเรียนจำนวน 76 จาก 442 คนที่ตอบได้ถูกต้อง การที่ค่าความยากง่ายของข้อสอบเท่ากับ 0.54 แสดงว่าข้อสอบไม่ยากหรือง่ายจนเกินไป ค่าอำนาจจำแนกเท่ากับ 0.32 แสดงว่า นักศึกษาที่ได้คะแนนสูงในวิชานี้เลือกคำตอบได้ถูกต้อง ผู้วิจัยสรุปในตอนท้ายว่าการสร้างข้อสอบมีความสำคัญอย่างยิ่งและควรทำตารางแสดงผลการวิเคราะห์ข้อสอบไว้ท้ายข้อเพื่อเลือกข้อที่ดีไว้ในธนาคารข้อสอบรวมทั้งควรวิพากษ์ข้อสอบก่อนนำมาใช้กับนักศึกษาด้วย

Lee & Winke (2012) ได้ทำการศึกษาความแตกต่างของการใช้ข้อคำตอบแบบ 3, 4 และ 5 ตัวเลือกในการทดสอบการฟังภาษาอังกฤษ จากการศึกษาโดยทำการปรับแบบทดสอบการฟังภาษาอังกฤษของ College Scholastic Ability Test (CSAT) จาก 5 ตัวเลือก เป็น 4 ตัวเลือก และ 3 ตัวเลือก โดยผู้ทรงคุณวุฒิชาวเกาหลีจำนวน 73 คน ที่มีความเชี่ยวชาญภาษาอังกฤษในด้านการพูดและการเรียนรู้ทางภาษา โดยทำการตัดตัวลวงที่เหมาะสมในการคัดออกออกจำนวน 2 รอบ และได้แบ่งกลุ่มผู้รับการทดสอบคือ นักเรียนระดับชั้นมัธยมปลายที่ศึกษาอยู่ในกรุงโซล ประเทศเกาหลี จำนวน 264 คน เป็นจำนวน 3 กลุ่ม โดยให้แต่ละกลุ่มได้ทำแบบทดสอบจำนวน 3 ชุดคือ แบบทดสอบแบบใช้ข้อคำตอบ 5, 4 และ 3 ตัวเลือกอย่างละ 1 ชุด จากแบบทดสอบทั้งหมด 9 ชุด จากผลการศึกษาพบว่า แบบทดสอบ

แบบใช้ข้อคำตอบ 3 ตัวเลือกจะมีค่าคะแนนเฉลี่ยสูงกว่า แบบทดสอบที่มี 4 ตัวเลือก และ 5 ตัวเลือก อย่างมีนัยสำคัญทางสถิติ แต่ไม่พบความแตกต่างของค่าเฉลี่ยความสามารถในการจำแนกผู้สอบของ ข้อสอบทั้ง 3 แบบ นอกจากนี้ ค่าความเชื่อมั่นของแบบทดสอบได้แสดงให้เห็นถึงความไม่สอดคล้องกัน ระหว่างจำนวนตัวเลือกของแบบทดสอบ กับชุดแบบทดสอบ อาจแปลผลได้ว่าคะแนนจากการทดสอบ ของแบบทดสอบแต่ละแบบ มีความสัมพันธ์ในระดับต่ำกับจำนวนตัวเลือกข้อคำตอบของแบบทดสอบใน การวัดทักษะทางการฟัง จากผลการศึกษาทำให้พบว่าในการออกแบบทดสอบโดยการใช้หรือไม่ใช้ข้อ คำตอบแบบ 3 ตัวเลือก ขึ้นอยู่กับมุมมองของผู้ออกข้อสอบ หรือผู้มีส่วนได้ส่วนเสีย และในการพัฒนา แบบทดสอบ ผู้ทำการพัฒนาแบบทดสอบควรพิจารณาใช้สถิติในการทดสอบที่หลากหลาย และมองปัจจัย จากสภาพแวดล้อมรอบด้านในการตัดสินใจสร้างแบบทดสอบว่าควรสร้างข้อคำตอบแบบตัวเลือกจำนวน เท่าใด

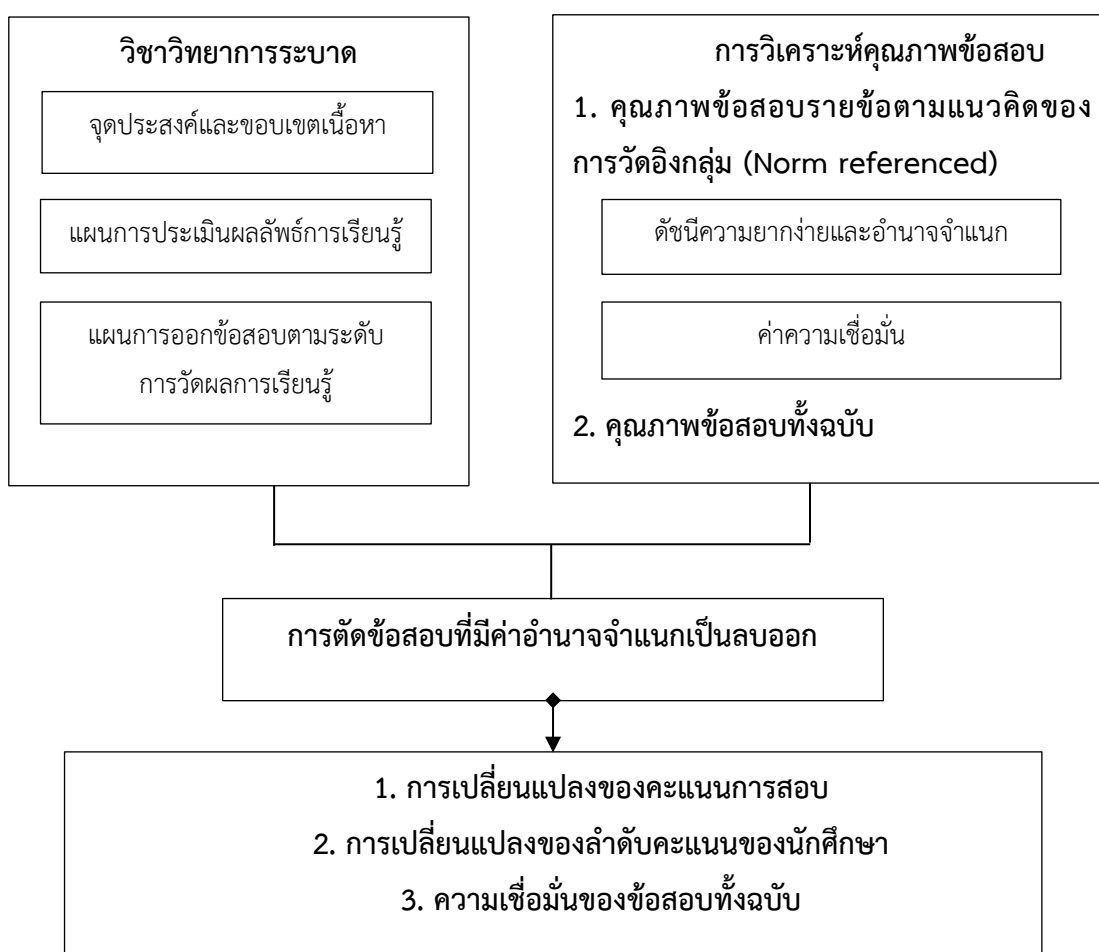
Caldwell & Pate (2013) ได้ทำการศึกษาเพื่อเปรียบเทียบการใช้รูปแบบคำถามตามแนวปฏิบัติ ที่เป็นมาตรฐานและไม่เป็นมาตรฐาน ข้อสอบที่มีแนวปฏิบัติที่เป็นมาตรฐาน ภายในข้อคำถามจะมีหลักการ ในการสร้างคำถาม ดังนี้ 1) การใช้คำถามเชิงบวก หลีกเลี่ยง คำถามเชิงลบ เช่น ยกเว้น ไม่ใช่ 2) ในแต่ละ คำถามควรใช้ตัวเลือกเพียง 3 ตัวเลือก 3) การใช้ตัวเลือก “ ไม่มีข้อใดถูกต้อง ” ไม่แนะนำให้ใช้บ่อย รูปแบบการวิจัยเป็นข้อสอบปรนัย 15 ข้อ คำถาม ทดสอบกับนักศึกษาเภสัชศาสตร์ โดยนักศึกษาจำนวน 55 คน ตอบแบบสอบถามที่สร้างตามแนวปฏิบัติที่เป็นมาตรฐาน เรียกว่า standard scale, และนักศึกษา 54 คน ตอบแบบสอบถามที่ไม่ได้สร้างจากแนวปฏิบัติ เรียกว่า nonstandard scale สัดส่วนของคำถาม เป็นการวัดด้านความรู้ - ความจำ ร้อยละ 60 เข้าใจ ร้อยละ 20 และ นำไปใช้ ร้อยละ 20 ผลการวิจัย พบว่า ข้อสอบที่ถูกสร้างตามแนวปฏิบัติที่เป็นมาตรฐาน นักศึกษาจะได้คะแนนสูงกว่ากลุ่มที่ทำข้อสอบที่ ไม่ได้มาตรฐาน โดยค่าคะแนนของผู้สอบ ข้อสอบ nonstandard scale มีความยากกว่าข้อสอบ standard scale ถึง 12.7 คะแนน แนวปฏิบัติที่แนะนำให้หลีกเลี่ยงการใช้ตัวลวง ไม่มีข้อใดถูกต้อง แสดง ให้เห็นความแตกต่างอย่างมีนัยสำคัญทางสถิติ ของคะแนนเฉลี่ย ระหว่างกลุ่ม ที่ทำ ข้อสอบ standard scale และกลุ่ม ที่ทำ ข้อสอบ nonstandard scale ร้อยละ 53.6 และ 41.3 ตามลำดับ ($p < 0.001$) จากผลการวิจัยสรุปได้ว่า ข้อสอบที่ไม่ได้มาตรฐานจะมีความยาก ทำให้นักศึกษาได้คะแนนน้อยกว่า ข้อสอบที่เป็นมาตรฐาน แต่ไม่สามารถแยกแยะกลุ่ม ถูก/ผิด (เก่ง/อ่อน) ได้ และนักศึกษาที่เรียนอ่อน มักจะตอบผิด ดังนั้นการใช้รูปแบบการสร้างข้อคำถามและตัวลวงควรมีการทดสอบเชิงโครงสร้างของ คำถาม และความสอดคล้องระหว่างข้อถูกและผิด

สรุปว่า องค์ประกอบที่สำคัญสำหรับการวิเคราะห์ข้อสอบส่วนใหญ่คือการวิเคราะห์ดัชนีความยากง่ายและค่าอำนาจจำแนก ซึ่งสอดคล้องกับวิธีการวิเคราะห์ในการวิจัยนี้

กรอบแนวคิดในการวิจัย

แสดงดังภาพที่ 2.1

ภาพที่ 2.1 กรอบแนวคิดในการวิจัย



บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยเรื่องผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบ และความเชื่อมั่นของข้อสอบทั้งฉบับ: รายวิชาวิทยาการระบาด มีวิธีดำเนินการวิจัยดังนี้

1. ประชากรและการเลือกกลุ่มตัวอย่าง
2. เครื่องมือการวิจัย และการตรวจสอบคุณภาพของเครื่องมือ
3. การเก็บรวบรวมข้อมูล
4. การวิเคราะห์ข้อมูล

ประชากรและการเลือกกลุ่มตัวอย่าง

ประชากร คือ ข้อสอบ และคะแนนการสอบรายวิชาวิทยาการระบาดของนักศึกษา หลักสูตรพยาบาลศาสตรบัณฑิตที่มีการวัดและประเมินผลโดยใช้ข้อสอบแบบเลือกตอบ (Multiple choices) 4 ตัวเลือก หมวดวิชาพื้นฐานวิชาชีพ รายวิชาวิทยาการระบาด ของนักศึกษาชั้นปีที่ 2 มีนักศึกษาที่ลงทะเบียนเรียน จำนวน 92 คน ประชากรทั้งหมดที่เข้าถึงได้ (Accessible population) คือ ข้อสอบ และคะแนนการสอบรายวิชาวิทยาการระบาดของนักศึกษา หลักสูตรพยาบาลศาสตรบัณฑิต ชั้นปีที่ 1-4 ที่ลงทะเบียนเรียนในภาคการศึกษาที่ 1 ปีการศึกษา 2555

การเก็บรวบรวมข้อมูล

1. จัดเรียงการสอนตามที่กำหนดในการออกแบบการเรียนการสอนรายวิชา (มคอ.3) การออกข้อสอบเป็นไปตามตารางวิเคราะห์หลักสูตร (Test blueprint) วิพากษ์ข้อสอบโดยคณะกรรมการบริหารหลักสูตรของคณะและดำเนินการสอบกลางภาค และปลายภาค และตัดสินเกรด ตามระเบียบการวัดและประเมินผลของมหาวิทยาลัย

2. ส่งข้อสอบไปวิเคราะห์โดยโปรแกรมวิเคราะห์ข้อสอบสำเร็จรูปคอมพิวเตอร์ที่สำนักส่งเสริมวิชาการ และงานทะเบียนของมหาวิทยาลัยราชภัฏสวนดุสิต กำหนดให้วิเคราะห์ดัชนีความยากง่าย ค่าอำนาจ

จำแนก ค่าความเชื่อมั่น โดยส่งข้อสอบวิเคราะห์ 2 ครั้ง ครั้งแรกวิเคราะห์ข้อสอบตามปกติ ครั้งที่สองวิเคราะห์โดยตัดข้อสอบที่มีค่าอำนาจจำแนกเป็นลบออก

3. รวบรวมผลการวิเคราะห์ข้อสอบรายวิชาวิเคราะห์ข้อมูลตามวัตถุประสงค์ของการวิจัย

เครื่องมือในการวิจัยและการตรวจสอบคุณภาพเครื่องมือ

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ คะแนนสอบวิชาวิทยาการระบาดและผลของการวิเคราะห์ข้อสอบรายข้อ ด้วยโปรแกรมการวิเคราะห์ข้อสอบรายข้อ (Items Analysis) ที่พัฒนาโดยบริษัท ApplicationSupport & Development Department: Control Data (Thailand) Ltd. อาศัยหลักในการคำนวณและวิเคราะห์ข้อสอบโดยใช้ทฤษฎีการวิเคราะห์แบบทดสอบแบบประเพณีนิยม (Classical Test Theory) เพื่อศึกษาคุณภาพของข้อสอบทั้งรายข้อ และรวมทั้งฉบับ ทฤษฎีการวิเคราะห์แบบทดสอบแบบประเพณีนิยมนี้ เหมาะสำหรับแบบทดสอบแบบเลือกตอบ (Multiple choice test) ของแบบทดสอบแบบอิงกลุ่ม (Norm referenced test) โดยทฤษฎีการวิเคราะห์แบบทดสอบนั้นจะมีหลายเทคนิคที่ใช้กันอยู่ แต่ในการวิจัยครั้งนี้จะใช้เทคนิค 27% วิธีนี้นิยมใช้เมื่อผู้สอบแบบทดสอบมีจำนวนมากหรือการกระจายของคะแนนสอบเป็นปกติ วิธีนี้เชื่อว่ากลุ่มขนาดตัวอย่าง 27% จากผู้สอบที่ได้คะแนนสูงหรือผู้ที่สอบได้คะแนนต่ำสามารถเป็นตัวแทนประชากรของผู้ที่สอบทั้งหมด ส่วนคะแนนที่เหลืออีก 46% ซึ่งเป็นของคนที่มีความสามารถปานกลางไม่ต้องนำมาวิเคราะห์การใช้ 27% นี้จะได้ผลเหมือนกับการนำเอาผู้เรียนทั้งหมด (100%) มาวิเคราะห์

การวิเคราะห์ข้อมูล

1. เปรียบเทียบค่าความเชื่อมั่นของข้อสอบวิชาวิทยาการระบาด ทั้งฉบับก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก โดยใช้โปรแกรมการวิเคราะห์ข้อสอบรายข้อ (Items Analysis)
2. การเปลี่ยนแปลงของคะแนนสอบของนักศึกษาวิชาวิทยาการระบาดก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออกใช้สถิติ Paired t-test (one-tailed)
3. การเปลี่ยนแปลงลำดับของคะแนนสอบของนักศึกษาวิชาวิทยาการระบาดก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออกใช้สถิติ Sign Ranked test

บทที่ 4

ผลการวิจัย

การนำเสนอผลการวิเคราะห์ข้อมูลในการวิจัยเรื่องผลของการตัดข้อสอบที่มีอำนาจจำแนก ติดลบออกต่อคะแนนการสอบและความเชื่อมั่นของข้อสอบทั้งฉบับ: รายวิชาวิทยาการระบาด ของ คณะพยาบาลศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิต ฉบับนี้ เป็นผลการศึกษาวิจัย จากโครงการวิจัย เกี่ยวกับผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบ และความเชื่อมั่นของ ข้อสอบทั้งฉบับ ซึ่งประกอบด้วยข้อสอบกลางภาค และข้อสอบปลายภาค ในหมวดวิชาพื้นฐานชีพ วิทยาการระบาด ในประเด็นต่างๆตามวัตถุประสงค์ของการวิจัย ดังต่อไปนี้

1. ข้อมูลทั่วไป
2. คุณภาพของข้อสอบของรายวิชาวิทยาการระบาด
3. การเปลี่ยนแปลงของคะแนนสอบรายวิชาและอันดับก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจ จำแนก ติดลบออกการเปลี่ยนแปลงค่าความเชื่อมั่นของข้อสอบทั้งฉบับก่อนและหลังการตัดข้อสอบที่ มีค่าอำนาจจำแนกติดลบออก

ข้อมูลทั่วไป

การวิจัยครั้งนี้ เป็นการศึกษาจากชุดข้อสอบที่ใช้ในการสอบนักศึกษาชั้นปีที่ 3 ราย วิชา วิทยาการระบาด ประกอบด้วยข้อสอบกลางภาคและข้อสอบปลายภาค ในปีการศึกษา 2555 ภาค การศึกษาที่ 1

ตารางที่ 4.1 จำนวนนักศึกษาและจำนวนข้อสอบในแต่ละชุดข้อสอบ ที่ใช้ในการศึกษาวิจัย จำแนกตามชุดข้อสอบ รายวิชาวิทยาการระบาด

ชุดข้อสอบ	จำนวนนักศึกษาที่เข้าสอบ	จำนวนข้อสอบในแต่ละชุด
กลางภาค	92	60
ปลายภาค	92	80

ตารางที่ 4.1 แสดงให้เห็นว่า ทั้งกลางภาคและปลายภาค มีจำนวนข้อสอบ 92 ชุดเท่ากัน จำนวนข้อสอบกลางภาคน้อยกว่าปลายภาค

คุณภาพข้อสอบของรายวิชาวิทยาการระบาด

ผลการวิเคราะห์ข้อสอบกลางภาค และปลายภาค ของข้อสอบที่ใช้วัดผลการเรียนการสอนในรายวิชาวิทยาการระบาดในปีการศึกษา 2555 ภาคการศึกษาที่ 1 โดยใช้ค่าความยากง่ายและค่าอำนาจจำแนกของข้อสอบเป็นดัชนีแสดงถึงคุณภาพของข้อสอบ ดังแสดงตามตารางต่อไปนี้

ตารางที่ 4.2 ค่าความยากง่ายของข้อสอบกลางภาค วิชาวิทยาการระบาด (N = 92 คน)

ดัชนีความยากง่าย (P)	ความหมาย	จำนวนข้อสอบ	ร้อยละ
0.80-1.00	ง่ายมาก	3	5.0
0.60-0.79	ค่อนข้างง่าย	27	45.0
0.40-0.59	ความยากง่ายพอเหมาะ	18	30.0
0.20-0.39	ค่อนข้างยาก	10	16.7
0.00-0.19	ยากมาก	2	3.3

ตารางที่ 4.2 แสดงให้เห็นว่าข้อสอบกลางภาควิชาวิทยาการระบาดร้อยละ 45 ค่อนข้างง่าย ร้อยละ 30 มีความยากง่ายพอเหมาะ ในขณะที่เดียวกันพบว่าข้อสอบร้อยละ 3.3 ยากมาก

ตารางที่ 4.3 ค่าความยากง่ายของข้อสอบปลายภาค วิชาวิทยาการระบาด (N = 92 คน)

ดัชนีความยากง่าย (P)	ความหมาย	จำนวนข้อสอบ	ร้อยละ
0.80-1.00	ง่ายมาก	4	5.0
0.60-0.79	ค่อนข้างง่าย	29	36.2
0.40-0.59	ความยากง่ายพอเหมาะ	18	22.5
0.20-0.39	ค่อนข้างยาก	19	23.8
0.00-0.19	ยากมาก	10	12.5

ตารางที่ 4.3 แสดงให้เห็นว่าข้อสอบปลายภาควิชาวิทยาการระบาดร้อยละ 36.2 ค่อนข้างง่าย ร้อยละ 22.5 มีความยากง่ายพอเหมาะ ในขณะที่เดียวกันพบว่าข้อสอบร้อยละ 12.5 ยากมาก

ตารางที่ 4.4 ค่าอำนาจจำแนกของข้อสอบกลางภาค วิชาวิทยาการระบาด (N = 92 คน)

ค่าอำนาจจำแนก (r)	ความหมาย	จำนวนข้อสอบ	ร้อยละ
0.80-1.00	จำแนกได้ดีมาก	0	0
0.60-0.79	จำแนกได้ดี	1	1.7
0.40-0.59	จำแนกได้ปานกลาง	9	15.0
0.20-0.39	จำแนกได้พอใช้	27	45.0
0.00-0.19	จำแนกได้น้อยมาก	17	28.3
-1.00-0.00	ใช้ไม่ได้	6	10

ตารางที่ 4.4 แสดงให้เห็นว่าข้อสอบกลางภาควิชาวิทยาการระบาดร้อยละ 1.7 มีค่าอำนาจจำแนกดี ร้อยละ 45 มีค่าอำนาจจำแนกพอใช้ ร้อยละ 28.3 มีค่าอำนาจจำแนกน้อยมาก และร้อยละ 10 เป็นข้อสอบที่มีค่าอำนาจจำแนกเป็นลบ

ตารางที่ 4.5 ค่าอำนาจจำแนกของข้อสอบปลายภาค วิชาวิทยาการระบาด (N = 92 คน)

ค่าอำนาจจำแนก (r)	ความหมาย	จำนวนข้อสอบ	ร้อยละ
0.80-1.00	จำแนกได้ดีมาก	0	0
0.60-0.79	จำแนกได้ดี	3	3.75
0.40-0.59	จำแนกได้ปานกลาง	11	13.75
0.20-0.39	จำแนกได้พอใช้	31	38.75
0.00-0.19	จำแนกได้น้อยมาก	28	35.0
-1.00-0.00	ใช้ไม่ได้	7	8.75

ตารางที่ 4.5 แสดงให้เห็นว่าข้อสอบปลายภาควิชาวิทยาการระบาดร้อยละ 3.75 มีค่าอำนาจจำแนกดี ร้อยละ 38.75 มีค่าอำนาจจำแนกพอใช้ ร้อยละ 35 มีค่าอำนาจจำแนกน้อยมาก และร้อยละ 8.75 เป็นข้อสอบที่มีค่าอำนาจจำแนกเป็นลบ

ตารางที่ 4.6 ค่าความเชื่อมั่นของข้อสอบวิชาวิทยาการระบาด ก่อนและหลัง การตัดข้อสอบที่มีค่า r ตีลบออก (N = 92 คน)

ข้อสอบ	จำนวนข้อสอบที่ตัดออก (ข้อ)	ร้อยละของข้อสอบที่ตัดออก (%)	ค่าความเชื่อมั่นเดิมของข้อสอบ	ค่าความเชื่อมั่นหลังการตัดข้อที่มีค่า r ตีลบออก
กลางภาค	6	10	0.554	0.622
ปลายภาค	7	8.75	0.694	0.727

ตารางที่ 4.6 แสดงให้เห็นว่า หลังการตัดข้อสอบที่มีค่าอำนาจจำแนกตีลบออกจากชุดข้อสอบกลางภาค และข้อสอบปลายภาคพบว่าข้อสอบกลางภาค และข้อสอบปลายภาคมีค่าความเชื่อมั่นทั้งฉบับเพิ่มขึ้นจาก 0.554 เป็น 0.622 และจาก 0.694 เป็น 0.727 ตามลำดับ

ตารางที่ 4.7 ผลการทดสอบการเปลี่ยนแปลงของคะแนนการสอบของนักศึกษาในวิชาวิทยาการระบาดหลังการตัดข้อสอบที่มีค่า r ตีลบออก (N= 92 คน)

ข้อสอบ	ก่อน		หลังการตัดข้อที่มีค่า r ตีลบออก		Paired t-test
	พิสัยของคะแนน	คะแนนเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน	พิสัยของคะแนน	คะแนนเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน	
กลางภาค	33.33-78.33	55.36 (8.79)	33.33-81.48	56.68 (10.08)	-4.75*
ปลายภาค	33.75-73.75	49.76 (8.99)	32.88-75.34	51.19 (10.04)	-7.17*

* p-value < .01

ตารางที่ 4.7 แสดงให้เห็นว่าคะแนนเฉลี่ยทั้งฉบับของข้อสอบวิชาวิทยาการระบาดกลางภาค และปลายภาค ก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกตีลบออก มีการเปลี่ยนแปลงของ

คะแนนการสอบของนักศึกษา คะแนนเฉลี่ยเพิ่มขึ้นจาก 55.36 (SD = 8.79) คะแนน เป็น 56.68 (SD = 10.08) คะแนน และจาก 49.76 (SD = 8.99) คะแนน เป็น 51.19 (SD = 10.04) คะแนนอย่างมีนัยสำคัญทางสถิติ ($p\text{-value} < .01$)

การเปลี่ยนแปลงของคะแนนสอบรายวิชาและลำดับคะแนน

ตารางต่อไปนี้ แสดงการเปลี่ยนแปลงของคะแนนสอบรายวิชาและลำดับคะแนนภายหลังตัดข้อสอบที่มีค่าอำนาจจำแนกเป็นลบออก

ตารางที่ 4.8 ผลการทดสอบการเปลี่ยนแปลงของคะแนนสอบและลำดับคะแนนจากการสอบของนักศึกษาในวิชาวิทยาการระบาด หลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก (N= 92 คน)

ข้อสอบ	จำนวนนักศึกษาที่ คะแนนเพิ่มขึ้น (ร้อยละ)	จำนวนนักศึกษาที่ คะแนนลดลง (ร้อยละ)	จำนวนนักศึกษา ที่คะแนนคงเดิม (ร้อยละ)	Signed Rank Test
กลางภาค	65 คน (70.65)	25 คน (27.17)	2 คน (2.18)	-4.313*
ปลายภาค	75 คน (81.52)	17 คน (18.48)	0 คน (0)	-6.038*

* $p\text{-value} < .01$

ตารางที่ 4.8 แสดงให้เห็นว่าผลการทดสอบการเปลี่ยนแปลงของลำดับคะแนนจากการสอบของนักศึกษาในวิชาวิทยาการระบาด หลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออกพบว่าในการสอบกลางภาคจำนวนนักศึกษาที่คะแนนเพิ่มขึ้น 65 คน (ร้อยละ 70.65) ลดลง 25 คน (ร้อยละ 27.17) และคงเดิม 2 คน (ร้อยละ 2.18) ในการสอบปลายภาค จำนวนนักศึกษาที่คะแนนข้อสอบเพิ่มขึ้น 75 คน (ร้อยละ 81.52) และลดลง 17 คน (ร้อยละ 18.48) การเปลี่ยนแปลงดังกล่าวมีนัยสำคัญทางสถิติ ($p\text{-value} < .01$)

บทที่ 5

สรุปผลการวิจัยอภิปรายผลและข้อเสนอแนะ

สรุปผลการวิจัย

งานวิจัยเรื่องผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบ และความเชื่อมั่นของข้อสอบทั้งฉบับรายวิชาวิทยาการระดับ : ฉบับนี้เป็นการศึกษาคุณภาพของข้อสอบ การเปลี่ยนแปลงของค่าความเชื่อมั่นของข้อสอบ คะแนนการสอบของนักศึกษารายบุคคล และลำดับคะแนนหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออกในรายวิชาวิทยาการระดับ ปีการศึกษา 2555 ทำการรวบรวมข้อมูลจากผลของการวิเคราะห์ข้อสอบรายข้อ ด้วยโปรแกรมการวิเคราะห์ข้อสอบรายข้อ (Items Analysis) แล้วนำมาทดสอบทางสถิติโดยการเปรียบเทียบค่าความเชื่อมั่นของข้อสอบวิชาวิทยาการระดับ ทั้งฉบับก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก โดยใช้โปรแกรมการวิเคราะห์ข้อสอบรายข้อ (Items Analysis) ทดสอบการเปลี่ยนแปลงของคะแนนสอบของนักศึกษาวិชาวิทยาการระดับก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออกโดยใช้สถิติ Paired t-test(one-tailed) และศึกษาการเปลี่ยนแปลงลำดับของคะแนนสอบของนักศึกษาวิชาวิทยาการระดับก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออกโดยใช้สถิติ Sign Ranked test

ผลการศึกษาพบว่าข้อสอบที่มีค่าอำนาจจำแนกติดลบออก มีการเปลี่ยนแปลงของคะแนนการสอบของนักศึกษา โดยการสอบกลางภาคมีคะแนนเฉลี่ยเพิ่มขึ้นจาก 55.36 (SD = 8.79) คะแนน เป็น 56.68 (SD = 10.08) คะแนน และผลการทดสอบการเปลี่ยนแปลงของลำดับคะแนนจากการสอบของนักศึกษาในวิชาวิทยาการระดับหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออกพบว่าจำนวนนักศึกษาที่คะแนนข้อสอบกลางภาคเพิ่มขึ้น 65 คน (ร้อยละ 70.65) ลดลง 25 คน (ร้อยละ 27.17) และคงเดิม 2 คน (ร้อยละ 2.18) จำนวนนักศึกษาที่คะแนนข้อสอบปลายภาคเพิ่มขึ้น 75 คน (ร้อยละ 81.52) และลดลง 17 คน (ร้อยละ 18.48) การเปลี่ยนแปลงนี้มีนัยสำคัญทางสถิติที่ $p\text{-value} < .01$

อภิปรายผล

1. ผลการศึกษาคุณภาพของข้อสอบวิชาวิทยาการระดับโดยใช้โปรแกรมวิเคราะห์ข้อสอบเพื่อหาค่าความยากง่ายและค่าอำนาจจำแนกพบว่า ข้อสอบกลางภาคมีข้อสอบที่ยากมาก 2 ข้อ (ร้อยละ 3.3)

และง่ายมาก 3 ข้อ (ร้อยละ 5) ข้อสอบปลายภาคมีข้อสอบที่ยากมาก 10 ข้อ (ร้อยละ 12.5) และข้อสอบที่ง่ายมาก 4 ข้อ (ร้อยละ 5) ข้อสอบกลางภาคที่มีค่าอำนาจจำแนกเป็นลบ 6 ข้อ (ร้อยละ 10) ในขณะที่ข้อสอบปลายภาคมี 7 ข้อ (ร้อยละ 8.75) ผลการวิจัยนี้อธิบายตามแนวคิดของ เชิดศักดิ์ ไอรมณีรัตน์ (2552) ได้ว่าการปรับให้ข้อสอบมีการคละกันของข้อสอบที่ยากและง่ายอย่างเหมาะสม จะทำให้ผู้สอบมีระดับคะแนนแตกต่างกันมากขึ้น และมีการกระจายตัวของคะแนนที่เหมาะสม สำหรับข้อสอบที่มีค่าอำนาจจำแนกใช้ไม่ได้ อาจารย์ผู้สอนควรมีการพิจารณาตัดข้อสอบออก หรือปรับเปลี่ยนโจทย์คำถาม หรือตัวเลือกตอบ ให้มีความชัดเจนและเหมาะสม

2. ผลการศึกษาพบการเปลี่ยนแปลงของค่าความเชื่อมั่นของข้อสอบก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก กล่าวคือส่งผลให้ข้อสอบมีค่าความเชื่อมั่นเพิ่มขึ้นทั้งข้อสอบกลางภาคและข้อสอบปลายภาคอย่างมีนัยสำคัญทางสถิติ อภิปรายได้ว่าการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบจะมีผลทำให้ข้อสอบมีค่าความคงที่ของผลการวัดเพิ่มขึ้น สามารถวัดความรู้ความเข้าใจเนื้อหาวิชาที่เรียนได้ดีเพิ่มขึ้น มีความเชื่อมั่นสูงขึ้น ความเชื่อมั่นเป็นคุณสมบัติที่สำคัญประการหนึ่งของข้อสอบ ข้อสอบที่มีความเชื่อมั่นต่ำ เมื่อนำไปใช้ทดสอบ คะแนนที่ได้จากการสอบจะเปลี่ยนไปมาไม่คงที่ ซึ่งไม่สามารถนำมาตีความว่านักศึกษามีความรู้ความเข้าใจวิชาที่เรียนมากน้อยเพียงใดได้

3. ผลการศึกษากการเปลี่ยนแปลงของคะแนนการสอบของนักศึกษารายบุคคล ก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออกและผลการทดสอบการเปลี่ยนแปลงของลำดับคะแนนจากการสอบของนักศึกษาในวิชาวิทยาการระบาดหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออกพบว่า ในการสอบกลางภาค มีนักศึกษาที่คะแนนสอบเพิ่มขึ้น จำนวน 65 คน (ร้อยละ 70.65) ลดลง 25 คน (ร้อยละ 27.17) และคงเดิม 2 คน (ร้อยละ 2.18) สำหรับการสอบปลายภาค มีนักศึกษาที่คะแนนสอบเพิ่มขึ้น จำนวน 75 คน (ร้อยละ 81.52) ลดลง 17 คน (ร้อยละ 18.48) การเปลี่ยนแปลงนี้มีนัยสำคัญทางสถิติ ($p\text{-value} < .01$) อภิปรายได้ว่าการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก ส่งผลให้นักศึกษาบางรายมีคะแนนเพิ่มขึ้น ลดลงหรือคงเดิมได้ และมีลำดับคะแนนที่เปลี่ยนแปลงหรือคงเดิมได้ เนื่องจากการวิเคราะห์ข้อสอบในงานวิจัยชิ้นนี้ เป็นการวิเคราะห์ข้อสอบโดยใช้โปรแกรมสำเร็จรูปด้วยเทคนิคที่มีพื้นฐานจากทฤษฎีการสอบแบบดั้งเดิม ซึ่งค่าต่างๆที่ได้มาจากวิเคราะห์นั้นขึ้นอยู่กับกลุ่มตัวอย่างที่ใช้ในการเก็บข้อมูล ซึ่งกลุ่มตัวอย่างที่ดีควรจะต้องมีขนาดใหญ่พอ และมีการกระจายตัวของระดับความสามารถที่เหมาะสม ค่า p , r , Cronbach's coefficient, mean และ standard deviation อาจเปลี่ยนแปลงไปตามการกระจายตัวของระดับความสามารถในแต่ละกลุ่มของนักศึกษา (เชิดศักดิ์ ไอรมณีรัตน์, 2552: 36) การพิจารณาตัดข้อสอบที่มี

ค่าอำนาจจำแนกติดลบออก อาจารย์ผู้สอนควรพิจารณาอย่างรอบด้าน เช่นจำนวนนักศึกษาที่จะมีลำดับคะแนนสูงขึ้นหรือต่ำลง ข้อสอบที่จะตัดออกมีการครอบคลุมวัตถุประสงค์มากน้อยเพียงใด และเนื่องจากข้อสอบนี้เป็นข้อสอบที่ใช้ในการวัดนักศึกษาวิชาชีพพยาบาล ซึ่งในบางมิติจะเป็นข้อสอบที่ใช้วัดคุณธรรมจริยธรรมของวิชาชีพ ดังนั้นอาจารย์ผู้สอนจึงควรคำนึงถึงการตัดข้อสอบที่ใช้วัดมิติในด้านนี้ด้วย

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. สามารถนำผลการวิเคราะห์ข้อสอบไปใช้ในการปรับปรุงคุณภาพของข้อสอบได้ เช่น ปรับปรุงความคลุมเครือของภาษาที่ใช้ในคำถาม ความชัดเจนหรือความซ้ำซ้อนของตัวเลือกตอบ ข้อสอบที่มีคำตอบมากกว่า 1 ตัวเลือก
2. สามารถนำผลการวิเคราะห์ข้อสอบมาใช้ในการบริหารคลังข้อสอบ เช่น ใช้ในการพิจารณาเลือกเก็บข้อสอบที่มีคุณภาพเพื่อนำข้อสอบกลับมาใช้ใหม่
3. ในรายวิชาที่ยังไม่มีข้อสอบมาตรฐาน ในการสอบครั้งต่อไปควรพิจารณาเลือกใช้ข้อสอบที่ผ่านการวิเคราะห์แล้วว่าอยู่ในเกณฑ์มีคุณภาพ ภายใต้กรอบการสร้างข้อสอบตามผังการกระจายตัวของข้อสอบตามระดับการวัดผลการเรียนรู้ที่ได้ออกแบบไว้ในรายวิชาแล้วเพื่อควบคุมความตรงตามเนื้อหา (Content Validity) ของแบบสอบ
4. ควรเพิ่มพูนสมรรถนะด้านการสร้างข้อสอบให้กับอาจารย์ผู้สอนเพื่อลดโอกาสของการเกิดข้อสอบที่มีค่าอำนาจจำแนกเป็นลบ

ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

1. ควรนำข้อสอบไปใช้กับนักศึกษาในหลายๆกลุ่มเพื่อการพัฒนาสู่ข้อสอบมาตรฐานในคลังข้อสอบได้
2. ควรวิเคราะห์คุณภาพข้อสอบตามแนวคิดการวิเคราะห์ข้อสอบแบบอิงเกณฑ์เพื่อเพิ่มความเที่ยงของข้อสอบ โดยเฉพาะหากข้อสอบที่ค่า r ติดลบเป็นวัตถุประสงค์การเรียนรู้ที่สำคัญของรายวิชา
3. ควรวิเคราะห์คุณภาพข้อสอบที่มีจำนวนข้อใกล้เคียงกัน เพื่อลดความคลาดเคลื่อนจากวิธีการทางสถิติที่ใช้ในการวิจัย

บรรณานุกรม

บรรณานุกรมภาษาไทย

- เชิดศักดิ์ ไอรณณรัตน์. (2552). การวิเคราะห์ข้อสอบปรนัย. *เวชบันทึกศิริราช*, 2(1), 31-37.
- พวงรัตน์ ทวีรัตน์. (2540). *วิธีการวิจัยทางพฤติกรรมศาสตร์และสังคมศาสตร์*. พิมพ์ครั้งที่ 7. กรุงเทพฯ: มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- สุมาลี จันทร์ชลอ. (2542). *การวัดและประเมินผล*. กรุงเทพฯ: ศูนย์สื่อเสริมกรุงเทพ.
- อุทุมพร จามรมาน. (2535). *ข้อสอบ: การสร้างและการพัฒนา*. กรุงเทพฯ: ฟันนี่พับลิชชิ่ง.

บรรณานุกรมภาษาต่างประเทศ

- Burton, Richard F. (2004). Multiple choice and true/false tests: reliability measures and some implications of negative marking. *Assessment & Evaluation in Higher Education*, 29: 5, 585 — 595.
- Chase, C. I. (1978). *Measurement for Educational Evaluation*. 2nd ed. Reading, MA: Addison-Wesley Publishing Company.
- Caldwell, D. J. & Pate, A. M. (2013). Effects of Question Formats on Student and Item Performance. *American Journal of Pharmaceutical Education*, 77 (4), 1-5.
- Lee, HS. & Winke, P. (2012). The differences among three-, four-, and five-option formats in the context of a high-stakes English-language listening test. *Language Testing*, 30(1), 99-123.
- Oluseyi, A. D. & Ajeigbe, TO. (2012). The Analysis of Multiple choice Item of the Test of an introductory course in Chemistry in a Nigerian University. *The International Journal of Learning*, 18(4).
- Phipps, S. D. and et all. (2009). Relationship Between Assessment Item Format And Items Performance Characteristics. *American Journal of Pharmaceutical Education*, 73(8), 1-7.
- Tasdemir, M. (2010). A Comparison of Multiple-Choice Tests and True-False Tests Used in Evaluating Student Progress. *Journal of Instructional Psychology*, 37(3), 258 -266.

ภาคผนวก

ข้อมูลคะแนนการสอบกลางภาคและปลายภาค

วิชาวิทยาการระบาด ชั้นปีที่ 3

ลำดับที่	คะแนนการสอบกลางภาค			คะแนนการสอบปลายภาค		
	ก่อนตัดข้อ r ติดลบออก	หลังตัดข้อ r ติดลบออก	ค่าความ แตกต่าง	ก่อนตัดข้อ r ติดลบ ออก	หลังตัดข้อ r ติดลบ ออก	ค่าความ แตกต่าง
1	33.33	33.33	0.00	38.75	35.62	-3.13
2	38.33	37.04	-1.30	36.25	38.36	2.11
3	71.67	74.07	2.41	52.5	53.42	0.92
4	45.00	42.59	-2.41	52.5	53.42	0.92
5	56.67	53.70	-2.96	56.25	56.16	-0.09
6	55.00	53.70	-1.30	57.5	60.27	2.77
7	48.33	48.15	-0.19	40	42.47	2.47
8	55.00	55.56	0.56	51.25	53.42	2.17
9	60.00	64.81	4.81	48.75	52.05	3.30
10	60.00	61.11	1.11	61.25	64.38	3.13
11	50.00	50.00	0.00	58.75	61.64	2.89
12	63.33	66.67	3.33	45	45.21	0.21
13	58.33	53.70	-4.63	35	32.88	-2.12
14	48.33	53.70	5.37	57.5	58.90	1.40
15	55.00	59.26	4.26	51.25	52.05	0.80
16	48.33	55.56	7.22	46.25	47.95	1.70
17	58.33	59.26	0.93	57.5	60.27	2.77
18	53.33	53.70	0.37	42.5	45.21	2.71
19	55.00	57.41	2.41	47.5	47.95	0.45
20	60.00	57.41	-2.59	56.25	56.16	-0.09
21	60.00	61.11	1.11	46.25	47.95	1.70
22	41.67	37.04	-4.63	35	36.99	1.99
23	40.00	44.44	4.44	46.25	49.32	3.07

ลำดับที่	คะแนนการสอบกลางภาค			คะแนนการสอบปลายภาค		
	ก่อนตัดข้อ r ติดลบออก	หลังตัดข้อ r ติดลบออก	ค่าความ แตกต่าง	ก่อนตัดข้อ r ติดลบ ออก	หลังตัดข้อ r ติดลบ ออก	ค่าความ แตกต่าง
24	68.33	70.37	2.04	61.25	65.75	4.50
25	45.00	48.15	3.15	53.75	54.79	1.04
26	56.67	61.11	4.44	45	45.21	0.21
27	46.67	44.44	-2.22	52.5	56.16	3.66
28	58.33	61.11	2.78	55	58.90	3.90
29	55.00	55.56	0.56	43.75	53.42	9.67
30	55.00	53.70	-1.30	41.25	42.47	1.22
31	61.67	64.81	3.15	61.25	65.75	4.50
32	55.00	57.41	2.41	47.5	47.95	0.45
33	53.33	50.00	-3.33	33.75	32.88	-0.87
34	48.33	46.30	-2.04	52.5	56.16	3.66
35	55.00	55.56	0.56	46.25	46.58	0.33
36	51.67	55.56	3.89	36.25	36.99	0.74
37	50.00	50.00	0.00	50	52.05	2.05
38	75.00	79.63	4.63	71.25	75.34	4.09
39	56.67	57.41	0.74	38.75	36.99	-1.76
40	43.33	46.30	2.96	45	46.58	1.58
41	60.00	59.26	-0.74	43.75	45.21	1.46
42	58.33	61.11	2.78	56.25	60.27	4.02
43	40.00	35.19	-4.81	40	41.10	1.10
44	46.67	48.15	1.48	50	50.68	0.68
45	66.67	68.52	1.85	68.75	72.60	3.85
46	60.00	62.96	2.96	50	52.05	2.05
47	56.67	59.26	2.59	47.5	49.32	1.82
48	46.67	46.30	-0.37	37.5	38.36	0.86
49	43.33	42.59	-0.74	38.75	39.73	0.98

ลำดับที่	คะแนนการสอบกลางภาค			คะแนนการสอบปลายภาค		
	ก่อนตัดข้อ r ติดลบออก	หลังตัดข้อ r ติดลบออก	ค่าความ แตกต่าง	ก่อนตัดข้อ r ติดลบ ออก	หลังตัดข้อ r ติดลบ ออก	ค่าความ แตกต่าง
50	56.67	66.67	10.00	66.25	68.49	2.24
51	51.67	53.70	2.04	63.75	65.75	2.00
52	68.33	72.22	3.89	66.25	68.49	2.24
53	58.33	59.26	0.93	60	58.90	-1.10
54	58.33	62.96	4.63	40	36.99	-3.01
55	45.00	44.44	-0.56	42.5	41.10	-1.40
56	60.00	59.26	-0.74	43.75	45.21	1.46
57	53.33	55.56	2.22	61.25	63.01	1.76
58	55.00	55.56	0.56	57.5	61.64	4.14
59	45.00	44.44	-0.56	40	39.73	-0.27
60	50.00	51.85	1.85	47.5	50.68	3.18
61	51.67	57.41	5.74	51.25	54.79	3.54
62	63.33	66.67	3.33	52.5	54.79	2.29
63	48.33	46.30	-2.04	61.25	64.38	3.13
64	78.33	81.48	3.15	58.75	60.27	1.52
65	61.67	59.26	-2.41	43.75	43.84	0.09
66	70.00	70.37	0.37	46.25	46.58	0.33
67	65.00	68.52	3.52	57.5	60.27	2.77
68	55.00	55.56	0.56	53.75	57.53	3.78
69	53.33	53.70	0.37	42.5	41.10	-1.40
70	60.00	62.96	2.96	46.25	47.95	1.70
71	36.67	35.19	-1.48	37.5	35.62	-1.88
72	71.67	75.93	4.26	61.25	63.01	1.76
73	58.33	61.11	2.78	40	38.36	-1.64
74	53.33	57.41	4.07	40	41.10	1.10
75	56.67	57.41	0.74	38.75	38.36	-0.39

ลำดับที่	คะแนนการสอบกลางภาค			คะแนนการสอบปลายภาค		
	ก่อนตัดข้อ r ติดลบออก	หลังตัดข้อ r ติดลบออก	ค่าความ แตกต่าง	ก่อนตัดข้อ r ติดลบ ออก	หลังตัดข้อ r ติดลบ ออก	ค่าความ แตกต่าง
76	68.33	72.22	3.89	73.75	75.34	1.59
77	68.33	70.37	2.04	52.5	53.42	0.92
78	63.33	64.81	1.48	55	56.16	1.16
79	46.67	46.30	-0.37	50	50.68	0.68
80	65.00	68.52	3.52	45	46.58	1.58
81	61.67	62.96	1.30	61.25	64.38	3.13
82	63.33	66.67	3.33	60	60.27	0.27
83	63.33	68.52	5.19	47.5	49.32	1.82
84	43.33	44.44	1.11	56.25	58.90	2.65
85	58.33	57.41	-0.93	45	43.84	-1.16
86	55.00	55.56	0.56	50	52.05	2.05
87	66.67	70.37	3.70	48.75	49.32	0.57
88	48.33	48.15	-0.19	55	57.53	2.53
89	55.00	53.70	-1.30	41.25	39.73	-1.52
90	48.33	51.85	3.52	42.5	42.47	-0.03
91	60.00	61.11	1.11	48.75	49.32	0.57
92	45.00	42.59	-2.41	37.5	36.99	-0.51
	ค่าความยากง่าย		0.551	ค่าความยากง่าย		0.502
	อำนาจจำแนก		0.228	อำนาจจำแนก		0.233

ประวัติผู้วิจัย

นางศรีสุดา วงศ์วิเศษกุล

เกิดวันที่ 28 มกราคม 2500

ประวัติการศึกษา

พ.ศ. 2520 อนุปริญญาพยาบาลอนามัยและประกาศนียบัตรผดุงครรภ์

คณะพยาบาลศาสตร์ มหาวิทยาลัยมหิดล

พ.ศ. 2525 ศิลปศาสตรบัณฑิต (ภาษาอังกฤษ)

คณะมนุษยศาสตร์ มหาวิทยาลัยรามคำแหง

พ.ศ. 2530 พยาบาลศาสตรบัณฑิต (หลักสูตรต่อเนื่อง)

มหาวิทยาลัยสุโขทัยธรรมมาธิราช

พ.ศ. 2540 สาธารณสุขศาสตรมหาบัณฑิต (การพยาบาลสาธารณสุข)

คณะสาธารณสุขศาสตร์ มหาวิทยาลัยมหิดล

พ.ศ. 2554 ครุศาสตรดุษฎีบัณฑิต (อุดมศึกษา) คณะครุศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย

ตำแหน่งงาน

หัวหน้าสาขาการพยาบาลชุมชน

คณะพยาบาลศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิต

ประวัติผู้วิจัย

นางสาวมนชยา สมจريت

เกิดวันที่ 29 มีนาคม 2526

การศึกษา

ระดับปริญญาตรี

พยาบาลศาสตรบัณฑิต มหาวิทยาลัยมหิดล (พ.ศ.2548)

ระดับปริญญาโท

วิทยาศาสตร์มหาบัณฑิต (สาธารณสุขศาสตร์) มหาวิทยาลัยมหิดล
(พ.ศ.2554)

ประสบการณ์ทำงาน

พยาบาลปฏิบัติการ

หอผู้ป่วยพักรอดูอาการก่อนและหลังผ่าตัด สยามินทร์ 5
งานการพยาบาลผ่าตัด โรงพยาบาลศิริราช (พ.ศ.2548 - พ.ศ.255

อาจารย์ประจำ

สาขาการพยาบาลอนามัยชุมชน คณะพยาบาลศาสตร์
มหาวิทยาลัยราชภัฏสวนดุสิต (พ.ศ.2556 - พ.ศ. 2557)

อาจารย์ประจำ

คณะพยาบาลศาสตร์ สาขาการพยาบาลอนามัยชุมชน
มหาวิทยาลัยราชภัฏเพชรบุรี (พ.ศ. 2557 – ปัจจุบัน)