



รายงานการวิจัย
เรื่อง

ผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบ และ
ความเชื่อมั่นของข้อสอบทั้งฉบับ: รายวิชาการพยาบาลเด็กและวัยรุ่น

The effects on test scores and the reliability of the test
after excluding the Items with negative discrimination power:
Children and Adolescent Nursing Subject

นางสาวพัชมน อ้นโต
พ.ท.หญิงวชิรา ไกรถิน
นางสาวสุณีย์ ชื่นจันทร์

มหาวิทยาลัยราชภัฏสวนดุสิต

2558

ลิขสิทธิ์ของมหาวิทยาลัยราชภัฏสวนดุสิต



รายงานการวิจัย
เรื่อง

ผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบ และ
ความเชื่อมั่นของข้อสอบทั้งฉบับ: รายวิชาการพยาบาลเด็กและวัยรุ่น
The effects on test scores and the reliability of the test
after excluding the Items with negative discrimination power:
Children and Adolescent Nursing Subject

นางสาวพัชมน อันโต
พ.ท.หญิงวชิรา ไกรถิน
นางสาวสุณีย์ ชื่นจันทร์
คณะพยาบาลศาสตร์

มหาวิทยาลัยราชภัฏสวนดุสิต

2558

ลิขสิทธิ์ของมหาวิทยาลัยราชภัฏสวนดุสิต

(งานวิจัยนี้ได้รับทุนอุดหนุนจากมหาวิทยาลัยราชภัฏสวนดุสิต ปีงบประมาณ 2556)

หัวข้อวิจัย	ผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบ และความเชื่อมั่นของข้อสอบทั้งฉบับ: รายวิชาการพยาบาลเด็กและวัยรุ่น
ผู้ดำเนินการวิจัย	นางสาวพัชมน อินโต พันโทหญิงวชิรา ไกรถิน นางสาวสุณีย์ ชื่นจันทร์
ที่ปรึกษา	ผู้ช่วยศาสตราจารย์จรูไร อภัยจิรรัตน์
หน่วยงาน	คณะพยาบาลศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิต
ปี พ.ศ.	2558

การวิจัยครั้งนี้เป็นการศึกษาย้อนหลังจากชุดข้อสอบและคะแนนการสอบรายวิชาการพยาบาลเด็กและวัยรุ่นภาคการเรียนที่ 1 ปีการศึกษา 2555 มีวัตถุประสงค์เพื่อศึกษา 1) คุณภาพของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่น 2) การเปลี่ยนแปลงคะแนนการสอบของนักศึกษาในรายวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก 3) การเปลี่ยนแปลงลำดับคะแนนการสอบของนักศึกษาในวิชาพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก 4) การเปลี่ยนแปลงค่าความเชื่อมั่นของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นทั้งฉบับหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก ประชากรและกลุ่มตัวอย่างคือ ข้อสอบและคะแนนการสอบของนักศึกษาหลักสูตรพยาบาลศาสตรบัณฑิตชั้นปีที่ 2 ที่ลงทะเบียนเรียนและเข้าสอบวิชาพยาบาลเด็กและวัยรุ่นในภาคการเรียนที่ 1 ปีการศึกษา 2555 จำนวน 88 คน ชุดข้อสอบที่ใช้เป็นแบบเลือกตอบ (Multiple choices) 4 ตัวเลือก เครื่องมือที่ใช้ในการวิจัยประกอบด้วยแบบบันทึกข้อมูลที่ผู้วิจัยใช้สำหรับเก็บข้อมูลจากการสอบแต่ละครั้งและโปรแกรมการวิเคราะห์ข้อสอบรายข้อ (Item analysis) สถิติที่ใช้ในการวิจัยได้แก่ Paired t – test (one - tailed) และ Wilcoxon Signed Ranks test ผลการวิจัยพบว่า

1) เมื่อพิจารณาคุณภาพข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นเป็นรายข้อจากค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) พบว่าข้อสอบในการสอบครั้งที่ 1 จำนวน 75 ข้อมีค่าความยากง่ายในเกณฑ์ที่เหมาะสม (p อยู่ระหว่าง 0.20 - 0.80) จำนวน 62 ข้อ คิดเป็นร้อยละ 82.67 มีค่าอำนาจจำแนกในเกณฑ์ที่เหมาะสม ($r \geq 0.20$) จำนวน 37 ข้อ คิดเป็นร้อยละ 49.33 และมีค่าอำนาจจำแนกติดลบจำนวน 5 ข้อคิดเป็นร้อยละ 6.67 ส่วนคุณภาพของข้อสอบทั้งฉบับมีค่าสัมประสิทธิ์ความเชื่อมั่น 0.60 ข้อสอบในการสอบครั้งที่ 2 จำนวน 70 ข้อมีค่าความยากง่ายในเกณฑ์ที่เหมาะสมจำนวน 60 ข้อ คิดเป็นร้อยละ 85.71 มีค่าอำนาจจำแนกในเกณฑ์ที่เหมาะสมจำนวน 39 ข้อ คิดเป็นร้อยละ 55.71 และมีค่าอำนาจจำแนกติดลบจำนวน 6 ข้อ คิดเป็นร้อยละ 8.57 ส่วนคุณภาพของข้อสอบทั้งฉบับมีค่าสัมประสิทธิ์ความเชื่อมั่น 0.66 ข้อสอบในการสอบครั้งที่ 3 จำนวน 58 ข้อมีค่า

ความยากง่ายในเกณฑ์ที่เหมาะสมจำนวน 47 ข้อ คิดเป็นร้อยละ 81.03 มีค่าอำนาจจำแนกในเกณฑ์ที่เหมาะสมจำนวน 27 ข้อ คิดเป็นร้อยละ 46.55 และมีค่าอำนาจจำแนกติดลบจำนวน 3 ข้อ คิดเป็นร้อยละ 5.17 ส่วนคุณภาพของข้อสอบทั้งฉบับมีค่าสัมประสิทธิ์ความเชื่อมั่น 0.50 และข้อสอบในการสอบปลายภาคจำนวน 75 ข้อ มีค่าความยากง่ายในเกณฑ์ที่เหมาะสมจำนวน 60 ข้อ คิดเป็นร้อยละ 80 มีค่าอำนาจจำแนกในเกณฑ์ที่เหมาะสมจำนวน 36 ข้อ คิดเป็นร้อยละ 48 และมีค่าอำนาจจำแนกติดลบจำนวน 4 ข้อ คิดเป็นร้อยละ 5.33 ส่วนคุณภาพของข้อสอบทั้งฉบับมีค่าสัมประสิทธิ์ความเชื่อมั่น 0.66

2) การเปลี่ยนแปลงคะแนนการสอบของนักศึกษาในวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก โดยใช้สถิติ Paired t - test (one - tailed) พบว่าในการสอบครั้งที่ 1 ครั้งที่ 3 และการสอบปลายภาคนักศึกษามีคะแนนเฉลี่ยสูงขึ้นอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ส่วนการสอบครั้งที่ 2 คะแนนเฉลี่ยของนักศึกษาไม่มีการเปลี่ยนแปลงอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

3) การเปลี่ยนแปลงลำดับคะแนนการสอบของนักศึกษาในวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก โดยใช้สถิติ Wilcoxon Signed Ranks test พบว่าในการสอบครั้งที่ 1 ครั้งที่ 3 และการสอบปลายภาคลำดับคะแนนที่นักศึกษาได้รับมีการเปลี่ยนแปลงอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ส่วนการสอบครั้งที่ 2 พบว่าลำดับคะแนนที่นักศึกษาได้รับไม่มีการเปลี่ยนแปลงอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

4) การเปลี่ยนแปลงค่าความเชื่อมั่นของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นทั้งฉบับหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออกพบว่า ในการสอบครั้งที่ 1 ค่าความเชื่อมั่นเพิ่มขึ้นจาก 0.60 เป็น 0.65 การสอบครั้งที่ 3 ค่าความเชื่อมั่นเพิ่มขึ้นจาก 0.66 เป็น 0.72 และการสอบปลายภาคค่าความเชื่อมั่นเพิ่มขึ้นจาก 0.66 เป็น 0.69 ส่วนการสอบครั้งที่ 2 ค่าความเชื่อมั่นลดลงหลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออกจาก 0.50 เป็น 0.46

คำสำคัญ: คุณภาพข้อสอบ ค่าอำนาจจำแนกติดลบ ค่าความเชื่อมั่นของข้อสอบทั้งฉบับ

Research Title	The effects on test scores and the reliability of the test after excluding the Items with negative discrimination power: Children and Adolescent Nursing Subject
Researcher	Miss Patchamon Onto Lt. Col. Vachira Kraitin Miss Sunee Chuenjan
Research Consultants	Assistant Professor Churai Arpaichiraratana
Organization	Faculty of Nursing, Suan Dusit Rajabhat University
Year	2015

This retrospective study was done with the test items and scores in Children and Adolescent Nursing subject in the academic year 2012. The objectives were to 1) describe the quality of the test 2) change the test scores students obtained after excluding the items with negative discrimination power 3) change student's rank from the scores after excluding the items with negative discrimination power and 4) change the reliability of the tests (Kuder - Richardson 20) after excluding the items with negative discrimination power. Data were collected from test items with 4 choices and the results of item analysis from 4 examinations in Children and Adolescent Nursing subject taken by 88 second year student nurses who enrolled in the first semester of academic year 2012. Data were then analyzed by Paired t-test (one-tailed), and Wilcoxon Signed Ranks test.

The study found that 1) the quality of the first examination (total of 75 items) had the appropriate index of difficulty (p) between 0.20 – 0.80 for 62 items (82.67%), the index of discrimination ($r \geq 0.20$) for 37 items (49.33%), and negative discrimination power for 5 items (6.67%). The reliability of the test was 0.60. The second examination test items (total of 70 items) had the appropriate index of difficulty (p) for 60 items (85.71%), the index of discrimination ($r \geq 0.20$) for 39 items (55.71%), and negative discrimination power for 6 items (8.57%). The reliability of the test was 0.66. The third examination test items (total of 58 items) had the appropriate index of difficulty (p) for 47 items (81.03%), the index of discrimination ($r \geq 0.20$) for 27 items (46.55%), and negative discrimination power for 3 items (5.71%). The reliability of the test was

0.50. The final examination test items (total of 75 items) had the appropriate index of difficulty (p) for 60 items (80%), the index of discrimination ($r \geq 0.20$) for 36 items (48%), and negative discrimination power for 4 items (5.33%). The reliability of the test was 0.66.

2) After excluding the items with negative discrimination power using Paired t-test (one-tailed), the first, the third and the final examination had higher mean score significantly at 0.01 level. The second examination had no significant change of mean score at 0.01 level.

3) After excluding the items with negative discrimination power using Wilcoxon Signed Ranks test, the first, the third and the final examination revealed the change in student's rank significantly at 0.01 level. The second examination had no change in student's rank at 0.01 level.

4) Finally, without the items with negative discrimination power, the reliability of the first, the third, and the final examination increased from 0.60 to 0.65, 0.66 to 0.72, and 0.66 to 0.69 respectively. In contrast, the reliability of the second examination decreased from 0.50 to 0.46.

Keywords: quality of the test, negative discrimination power, reliability of the test

กิตติกรรมประกาศ

งานวิจัยฉบับนี้สำเร็จลุล่วงไปได้ด้วยดีด้วยความกรุณาเป็นอย่างยิ่งจากผู้ช่วยศาสตราจารย์ ไร อภัยจิรรัตน์ หัวหน้าสาขาการพยาบาลเด็กและวัยรุ่น คณะพยาบาลศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิต ที่ได้กรุณาให้คำปรึกษาและให้ข้อเสนอแนะอันเป็นประโยชน์ยิ่ง ตลอดจนช่วยเหลือแก้ไขข้อบกพร่องต่างๆ คณะผู้วิจัยจึงขอขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอขอบพระคุณผู้ช่วยศาสตราจารย์ ดร. ประกาย จิโรจน์กุล รองคณบดีฝ่ายวิชาการ และ ดร. เบ็ญจา เตากกล้า คณบดีคณะพยาบาลศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิต ที่ช่วยส่งเสริมสนับสนุนและให้กำลังใจตลอดระยะเวลาที่ศึกษาวิจัย ขอขอบคุณนักศึกษาหลักสูตรพยาบาลศาสตรบัณฑิต ชั้นปีที่ 2 ปีการศึกษา 2555 ที่ลงทะเบียนเรียนวิชาการพยาบาลเด็กและวัยรุ่น และร่วมมือให้ข้อมูลในการวิจัย รวมทั้งเจ้าหน้าที่สำนักส่งเสริมวิชาการและงานทะเบียนของมหาวิทยาลัย ที่ให้ความร่วมมืออย่างดียิ่งในการช่วยวิเคราะห์ข้อสอบเพื่อใช้เป็นข้อมูลในการวิจัย

ท้ายสุดคณะผู้วิจัยขอขอบคุณมหาวิทยาลัยราชภัฏสวนดุสิต ที่ให้ทุนอุดหนุนการวิจัยในครั้งนี้

คณะผู้วิจัย

2558

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ก
บทคัดย่อภาษาอังกฤษ	ค
กิตติกรรมประกาศ	จ
สารบัญ	ฉ
สารบัญตาราง	ซ
บทที่ 1 บทนำ	1
ความเป็นมาและความสำคัญ	1
วัตถุประสงค์ของการวิจัย	5
ขอบเขตการวิจัย	5
คำจำกัดความที่ใช้ในงานวิจัย	6
ประโยชน์ที่คาดว่าจะได้รับ	6
บทที่ 2 แนวคิด ทฤษฎี เอกสารและงานวิจัยที่เกี่ยวข้อง	7
แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน	8
การวิเคราะห์คุณภาพของข้อสอบ	13
งานวิจัยที่เกี่ยวข้องกับคุณภาพของข้อสอบ	27
กรอบแนวคิดในการวิจัย	31
บทที่ 3 วิธีดำเนินการวิจัย	32
ประชากรและกลุ่มตัวอย่าง	32
เครื่องมือในการวิจัยและการตรวจสอบคุณภาพเครื่องมือ	32
การเก็บรวบรวมข้อมูล	33
การวิเคราะห์ข้อมูล	34

บทที่ 4	ผลการวิจัย	35
	ข้อมูลทั่วไป	35
	ผลการวิเคราะห์คุณภาพข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่น	36
	ผลการวิเคราะห์การเปลี่ยนแปลงคะแนนการสอบของนักศึกษา ในรายวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่า อำนาจจำแนกติดลบออก	42
	ผลการวิเคราะห์การเปลี่ยนแปลงลำดับคะแนนการสอบของนักศึกษา ในรายวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่า อำนาจจำแนกติดลบออก	43
	ผลการวิเคราะห์การเปลี่ยนแปลงค่าความเชื่อมั่นของข้อสอบรายวิชา การพยาบาลเด็กและวัยรุ่นทั้งฉบับหลังการตัดข้อสอบที่มีค่าอำนาจ จำแนกติดลบออก	44
บทที่ 5	สรุปผลการวิจัยอภิปรายผลและข้อเสนอแนะ	45
	สรุปผลการวิจัย	45
	อภิปรายผล	47
	ข้อเสนอแนะในการนำผลการวิจัยไปใช้	50
	ข้อเสนอแนะในการทำวิจัยครั้งต่อไป	51
บรรณานุกรม		52
	บรรณานุกรมภาษาไทย	52
	บรรณานุกรมภาษาต่างประเทศ	53
ภาคผนวก		55
	ภาคผนวก ก. ตารางวิเคราะห์หลักสูตร (Test blueprint)	56
	ภาคผนวก ข. ผลการวิเคราะห์ข้อมูลเพิ่มเติม	59
	ภาคผนวก ค. สรุปผลการวิเคราะห์ข้อสอบ	63
ประวัติผู้วิจัย		72

สารบัญตาราง

ตารางที่		หน้า
1.1	การวิเคราะห์ข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น ปีการศึกษา 2554	3
2.1	การวิเคราะห์คุณภาพข้อสอบจากค่าความยากง่าย	16
2.2	การวิเคราะห์คุณภาพข้อสอบจากค่าอำนาจจำแนก	17
2.3	เกณฑ์การพิจารณาค่าความยากง่ายของข้อสอบแบบอิงเกณฑ์	20
2.4	เกณฑ์การพิจารณาอำนาจจำแนกหรือดัชนีความไวของข้อสอบแบบอิงเกณฑ์	21
2.5	การแปลความหมายของค่าสัมประสิทธิ์ความเชื่อมั่น	25
4.1	จำนวนข้อสอบที่ใช้ในการสอบครั้งที่ 1 ครั้งที่ 2 ครั้งที่ 3 และปลายภาค รายวิชา การพยาบาลเด็กและวัยรุ่น	36
4.2	จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นในการสอบ ครั้งที่ 1 จำแนกตามค่าความยากง่าย (p) รายข้อ	36
4.3	จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่น ในการสอบ ครั้งที่ 2 จำแนกตามค่าความยากง่าย (p) รายข้อ	37
4.4	จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่น ในการสอบ ครั้งที่ 3 จำแนกตามค่าความยากง่าย (p) รายข้อ	37
4.5	จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่น ในการสอบ ปลายภาค จำแนกตามค่าความยากง่าย (p) รายข้อ	38
4.6	จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่น ในการสอบ ครั้งที่ 1 จำแนกตามค่าอำนาจจำแนก (r) รายข้อ	38
4.7	จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่น ในการสอบ ครั้งที่ 2 จำแนกตามค่าอำนาจจำแนก (r) รายข้อ	39
4.8	จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่น ในการสอบ ครั้งที่ 3 จำแนกตามค่าอำนาจจำแนก (r) รายข้อ	40
4.9	จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่น ในการสอบ ปลายภาค จำแนกตามค่าอำนาจจำแนก (r) รายข้อ	40
4.10	ค่าความเชื่อมั่นของข้อสอบทั้งฉบับรายวิชาการพยาบาลเด็กและวัยรุ่น	41
4.11	การเปรียบเทียบคะแนนการสอบของนักศึกษาในรายวิชาการพยาบาล เด็กและวัยรุ่น ก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก	42

สารบัญตาราง (ต่อ)

4.12	การเปลี่ยนแปลงของลำดับคะแนนการสอบของนักศึกษาในรายวิชาการพยาบาล เด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก	43
4.13	ค่าความเชื่อมั่นของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นก่อนและหลังการตัด ข้อสอบที่มีค่าอำนาจจำแนกติดลบออก	44
ก – 1	ตารางวิเคราะห์หลักสูตร (Test blueprint)	57
ค – 1	ผลการวิเคราะห์คะแนนการสอบครั้งที่ 1 วิชาการพยาบาลเด็กและวัยรุ่น ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก	64
ค – 2	ผลการวิเคราะห์ความเที่ยงของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นในการสอบ ครั้งที่ 1 ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก	64
ค – 3	ผลการวิเคราะห์ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับวิชาการ พยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 1 ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนก ติดลบออก	64
ค – 4	ผลการวิเคราะห์คะแนนการสอบครั้งที่ 1 วิชาการพยาบาลเด็กและวัยรุ่น หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก	65
ค – 5	ผลการวิเคราะห์ความเที่ยงของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นในการสอบ ครั้งที่ 1 หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก	65
ค – 6	ผลการวิเคราะห์ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับวิชาการ พยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 1 หลังตัดข้อสอบที่มีค่าอำนาจจำแนก ติดลบออก	65
ค – 7	ผลการวิเคราะห์คะแนนการสอบครั้งที่ 2 วิชาการพยาบาลเด็กและวัยรุ่น ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก	66
ค – 8	ผลการวิเคราะห์ความเที่ยงของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นในการสอบ ครั้งที่ 2 ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก	66
ค – 9	ผลการวิเคราะห์ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับวิชาการ พยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 2 ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนก ติดลบออก	66
ค – 10	ผลการวิเคราะห์คะแนนการสอบครั้งที่ 2 วิชาการพยาบาลเด็กและวัยรุ่น หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก	67

สารบัญตาราง (ต่อ)

ค - 23	ผลการวิเคราะห์ความเที่ยงของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นในการสอบ ปลายภาค หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก	71
ค - 24	ผลการวิเคราะห์ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับวิชาการ พยาบาลเด็กและวัยรุ่นในการสอบปลายภาค หลังตัดข้อสอบที่มีค่าอำนาจจำแนก ติดลบออก	71

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญ

การศึกษาเป็นปัจจัยสำคัญอย่างยิ่งในการสร้างและพัฒนาทรัพยากรมนุษย์ให้มีคุณภาพและสอดคล้องกับความต้องการของสังคม ทั้งยังเป็นการเตรียมความพร้อมผู้เรียนเพื่อก้าวเข้าสู่วิชาชีพต่อไป โดยเฉพาะอย่างยิ่งวิชาชีพพยาบาลซึ่งต้องการบุคลากรที่มีคุณภาพ มีความรู้ความสามารถในอันที่จะให้การบริการที่ดี มีมาตรฐานตรงตามความคาดหวังของประชาชนที่เพิ่มสูงขึ้น การจัดการศึกษาเพื่อให้บรรลุเป้าหมายดังกล่าวจำเป็นต้องมีการกำหนดจุดมุ่งหมายที่ชัดเจนเพื่อช่วยกำหนดคุณลักษณะที่พึงประสงค์ของผู้สำเร็จการศึกษาสำหรับใช้วางแผนการเรียนรู้ มีการจัดประสบการณ์การเรียนรู้หรือกิจกรรมการเรียนการสอนที่เหมาะสม รวมทั้งมีการวัดและประเมินผลเพื่อตัดสินผลการเรียนรู้ของผู้เรียนว่าบรรลุผลสำเร็จตามจุดมุ่งหมาย ซึ่งจะเป็นประโยชน์ต่อการปรับปรุงและพัฒนาการเรียนรู้อของผู้เรียนและการจัดกิจกรรมการเรียนรู้ของผู้สอน (ศิริชัย กาญจนวาสี, 2556, น. 1) การวัดและประเมินผลจึงเป็นองค์ประกอบที่สำคัญของการจัดการศึกษาที่มีคุณภาพเพื่อให้ได้มาซึ่งข้อมูลที่เที่ยงตรงและเชื่อถือได้ที่แสดงถึงความสามารถที่แท้จริงของผู้เรียน (อุทุมพร จามรมาน, 2535, น. 2) โดยขจัดความผิดพลาดต่างๆ ให้หมดไปหรือเกิดขึ้นน้อยที่สุด วิธีการหนึ่งที่จะทำให้ผลของการวัดและประเมินผลมีความถูกต้องแม่นยำและน่าเชื่อถือก็คือ การใช้เครื่องมือที่มีคุณภาพ

เครื่องมือที่ใช้ในการวัดและประเมินผลการศึกษา มีหลายประเภท การเลือกใช้เครื่องมือแต่ละประเภทขึ้นอยู่กับจุดมุ่งหมายหรือพฤติกรรมเรียนที่ต้องการวัด สำหรับวิชาการพยาบาลเด็กและวัยรุ่นนั้นเน้นการวัดผลการเรียนรู้ด้านพุทธิพิสัยเป็นส่วนใหญ่ ซึ่งเป็นการวัดความสามารถด้านสติปัญญาของผู้เรียนทั้งด้านความรู้ – ความจำ ความเข้าใจ การนำไปใช้และการวิเคราะห์ โดยใช้แบบทดสอบหรือข้อสอบซึ่งเป็นเครื่องมือที่ใช้วัดด้านพุทธิพิสัยได้ดีที่สุด เพราะสมอบจะแสดงความสามารถสูงสุดออกมาให้ปรากฏ ข้อสอบที่ถูกสร้างขึ้นจะถูกนำมาใช้เพื่อประโยชน์ทางการสอน ผลการสอบจะช่วยให้ผู้สอนสามารถจัดการสอนให้เหมาะสมกับผู้เรียนทั้งรายบุคคลและรายกลุ่ม ผู้เรียนสามารถนำผลที่ได้จากการสอบข้อสอบที่ดีมาใช้ในการวินิจฉัยความสามารถของตนด้วยตนเอง ช่วยจูงใจให้ผู้เรียนปรับปรุงประสิทธิภาพในการเรียนเพิ่มมากขึ้น (สมบุญ ชาติพงศ์ และเตือนใจ เกตุษา, 2554, น. 9 - 5) ทั้งนี้ข้อสอบต้องวัดความรู้ความสามารถของผู้เรียนได้ถูกต้องตรงกับความสามารถที่แท้จริงจึงจะถือว่าเป็นข้อสอบที่มีคุณภาพและมีประโยชน์ในการจัดการเรียนการสอน ในทางตรงข้ามหากทดสอบด้วยข้อสอบที่ไม่มีคุณภาพ นั่นคือการเก็บข้อมูลโดยใช้เครื่องมือที่ขาดคุณภาพหรือคุณภาพต่ำแล้ว ข้อมูลที่ได้ก็จะเป็นข้อมูลที่ไม่ถูกต้อง ไม่น่าเชื่อถือ (รังสรรค์ มณีเล็ก,

2552, น. 5) ข้อสอบที่สร้างขึ้นสำหรับวัดผลการเรียนรู้ของผู้เรียนจะมีคุณภาพเพียงไรขึ้นอยู่กับ การวางแผนสร้างข้อสอบ ความลึกซึ้งในเนื้อหาวิชาของผู้เขียนข้อสอบและทักษะการเขียนข้อสอบ ในทางปฏิบัติผู้สร้างข้อสอบสามารถบวทวนลักษณะทางกายภาพของข้อสอบและศึกษาความเหมาะสมของ ข้อสอบในเชิงเหตุผลได้ แต่ยังไม่สามารถทราบถึงคุณภาพข้อสอบในเชิงประจักษ์จนกว่าจะนำข้อสอบ นั้นไปใช้หรือทดลองใช้แล้วนำผลการตอบข้อสอบมาทำการวิเคราะห์หาคุณภาพของข้อสอบ (ศิริชัย กาญจนวาสี, 2556, น. 221)

ปัจจุบันมีการตรวจสอบคุณภาพของข้อสอบด้วยการวิเคราะห์คุณภาพของข้อสอบเป็นรายข้อ และทั้งฉบับ ถ้าข้อสอบข้อใดหรือฉบับใดมีคุณภาพดีก็ควรนำไปใช้ต่อไป แต่ถ้าบกพร่องก็ควรปรับปรุง แก้ไข การทำเช่นนี้จะได้ข้อสอบที่มีคุณภาพดีไปทดสอบผู้เรียน ช่วยให้การวัดและประเมินผลมี ประสิทธิภาพดียิ่งขึ้น (สมนึก ภัททิยธนี, 2555, น. 193) ชนิดของข้อสอบที่นิยมนำมาวิเคราะห์หาค่า ที่สุดคือ แบบเลือกตอบ โดยนำผลที่ได้จากการสอบของผู้เรียนมาวิเคราะห์ดัชนีบ่งชี้คุณภาพของ ข้อสอบเป็นรายข้อได้แก่ ค่าความยากง่าย ค่าอำนาจจำแนก และประสิทธิภาพของตัวลง ข้อสอบที่มี ค่าความยากง่ายอยู่ระหว่าง 0.2 - 0.8 ถือว่าเป็นข้อสอบที่ดี ถ้าอยู่นอกขอบเขตนี้ถือว่ายากหรือง่าย เกินไป (นิธิภัทร บาลศิริ, 2557, น. 223) ส่วนค่าอำนาจจำแนกที่ดีจะต้องมีค่าเป็นบวก และควรมีค่า ตั้งแต่ 0.20 ขึ้นไป (ศิริชัย กาญจนวาสี, 2556, น. 223) จึงจะสามารถแยกหรือจำแนกผู้เรียนกลุ่มที่ได้ คะแนนสูงหรือกลุ่มเก่งกับกลุ่มที่ได้คะแนนต่ำหรือกลุ่มอ่อนออกจากกัน ในการวิเคราะห์ประสิทธิภาพ ตัวลงทำโดยการตรวจสอบความถี่ของผู้ที่เลือกตัวลงนั้นๆ ถ้าตัวลงใดมีผู้เลือกในสัดส่วนที่สูง ถือว่า เป็นตัวลงที่ใช้ได้ แต่ถ้าตัวลงใดมีผู้เลือกน้อย แสดงว่าตัวลงนั้นขาดประสิทธิภาพ สมควรที่จะต้องมีการปรับปรุงแก้ไขให้ดีขึ้น (ศิริชัย กาญจนวาสี, 2556, น. 241) ตัวลงที่จัดว่าเป็นตัวลงที่ดีนั้นผู้ที่ เรียนอ่อนควรจะเลือกตอบมากกว่าผู้ที่มีผลการเรียนดี

เมื่อใดก็ตามที่มีการนำข้อสอบไปทดสอบ ข้อสอบจะต้องมีคุณสมบัติที่สำคัญประการหนึ่งคือ การให้ผลการสอบที่ใกล้เคียงของเดิมเมื่อผู้สอบคนนั้นทำการทดสอบซ้ำภายใต้สภาพการสอบที่ เหมือนเดิม คุณสมบัติของความคงเส้นคงวาของคะแนนที่ได้จากการสอบเรียกว่า ความเที่ยงหรือ ความเชื่อมั่น (ศิริชัย กาญจนวาสี, 2556, น. 55) ซึ่งถือเป็นดัชนีบ่งชี้คุณภาพของข้อสอบทั้งฉบับ ถ้าข้อสอบมีค่าความเชื่อมั่นสูง แสดงว่าคะแนนที่ได้จากการสอบใกล้เคียงกับคะแนนที่เป็นจริงของ ผู้สอบข้อสอบควรมีค่าสัมประสิทธิ์ความเที่ยงสูงสุดเท่าที่จะเป็นไปได้ในสถานการณ์นั้นอย่างน้อยที่สุด ควรมีค่าไม่ต่ำกว่า 0.50 (ศิริชัย กาญจนวาสี, 2556, น. 97)

ในการจัดการเรียนการสอนหลักสูตรพยาบาลศาสตรบัณฑิตของคณะพยาบาลศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิตยึดหลักการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ มุ่งให้เกิดคุณภาพตามเงื่อนไข ของกระทรวงศึกษาธิการและตามตัวชี้วัดที่สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (สมศ.) กำหนดไว้ จึงต้องดำเนินการให้กระบวนการวัดและทดสอบผลสัมฤทธิ์ในการเรียนมีความ

นำเชื่อถือตามมาตรฐาน โดยใช้ระเบียบการวัดและประเมินผลของมหาวิทยาลัยควบคู่กับแนวทางการวัดและประเมินผลที่คณะได้กำหนดขึ้นคือ ใช้การประเมินผลแบบอิงเกณฑ์ในรายวิชาภาคปฏิบัติ ส่วนรายวิชาภาคทฤษฎี เช่น รายวิชาการพยาบาลเด็กและวัยรุ่นใช้การประเมินผลแบบอิงกลุ่มควบคู่กับอิงเกณฑ์ กล่าวคือนักศึกษาต้องมีคะแนนดิบถึงเกณฑ์ร้อยละ 60 จึงจะได้รับเกรด C นอกจากนั้นยังได้มีการพัฒนามาตรการในการช่วยเหลือผู้ที่เรียนอ่อน เพื่อพัฒนานักศึกษาให้มีความรู้ความสามารถตามเกณฑ์ที่หลักสูตรกำหนดและได้พัฒนาเต็มตามศักยภาพของตน โดยประเมินศักยภาพของนักศึกษาเป็นระยะเพื่อค้นหาปัญหาและวางแผนช่วยเหลือผู้ที่เรียนอ่อน ทบทวนความรู้แก่นักศึกษาที่มีผลคะแนนสอบต่ำกว่าร้อยละ 60 เป็นต้น

คณะพยาบาลศาสตร์ตระหนักถึงความสำคัญของการนำข้อสอบที่มีคุณภาพมาใช้ประเมินผลการศึกษาของนักศึกษาจึงได้แต่งตั้งผู้ทรงคุณวุฒิของคณะ ทำหน้าที่วิพากษ์ข้อสอบร่วมกับผู้ออกข้อสอบเพื่อปรับปรุงแก้ไขข้อสอบก่อนนำไปทดสอบจริง ในทางปฏิบัติไม่สามารถนำข้อสอบไปทดสอบกับนักศึกษาเพื่อตรวจสอบคุณภาพก่อนนำไปใช้จริงได้ เพราะหากนำไปทดสอบก่อนก็จะเกิดปัญหาการจำข้อสอบได้ กว่าที่จะทราบระดับคุณภาพของข้อสอบก็ต่อเมื่อนำข้อสอบเหล่านี้ไปประเมินผลการศึกษาจริงซึ่งอาจจะพบว่า ยังมีข้อสอบบางข้อที่ด้อยคุณภาพ ประเด็นคุณภาพข้อสอบที่คณะพยาบาลศาสตร์ให้ความสำคัญเป็นอย่างมากคือ ค่าอำนาจจำแนกของข้อสอบที่มีค่าติดลบ เพราะข้อสอบเหล่านี้เป็นข้อสอบที่ผู้ที่เรียนเก่งส่วนใหญ่มีโอกาสจะตอบผิด ในขณะที่ผู้ที่เรียนอ่อนส่วนใหญ่มีโอกาสจะตอบถูก ซึ่งจะไม่สามารถแยกผู้ที่เรียนเก่งออกจากผู้ที่เรียนอ่อนได้

จากประสบการณ์ในการจัดการเรียนการสอนวิชาการพยาบาลเด็กและวัยรุ่นในปีการศึกษา 2554 ที่ผ่านมาพบว่า นักศึกษาบางส่วนไม่ผ่านเกณฑ์ขั้นต่ำร้อยละ 60 ตามที่หลักสูตรกำหนดจึงต้องมีการสอบซ่อม เมื่อนำข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นในปีการศึกษา 2554 ซึ่งเป็นข้อสอบเลือกตอบแบบหลายตัวเลือกมาวิเคราะห์คุณภาพข้อสอบ ผลปรากฏดังตารางที่ 1.1

ตารางที่ 1.1 การวิเคราะห์ข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น ปีการศึกษา 2554

คุณภาพข้อสอบ	จำนวนข้อสอบ		
	สอบครั้งที่ 1	สอบครั้งที่ 2	สอบปลายภาค
ข้อสอบที่ดี			
1. ข้อที่ค่อนข้างง่าย อำนาจการจำแนกดี	1	1	4
2. ข้อที่ยากง่ายปานกลาง อำนาจการจำแนกดี	22	29	31
3. ข้อที่ค่อนข้างยาก อำนาจการจำแนกดี	-	1	2

ตารางที่ 1.1 การวิเคราะห์ข้อสอบวิชาการศึกษาพยาบาลเด็กและวัยรุ่น ปีการศึกษา 2554 (ต่อ)

คุณภาพข้อสอบ	จำนวนข้อสอบ		
	สอบครั้งที่ 1	สอบครั้งที่ 2	สอบปลายภาค
ข้อสอบที่ควรปรับปรุง			
1. ข้อที่ง่ายมาก อำนาจการจำแนกดี	-	-	-
2. ข้อที่ยากมาก อำนาจการจำแนกดี	-	-	-
3. ข้อที่ความยากปานกลาง อำนาจการจำแนกไม่ดี	28	34	51
ข้อสอบที่ควรตัดทิ้ง			
1. ข้อที่ยากมากหรือง่ายมาก อำนาจการจำแนกดี	15	9	23
2. ข้อที่ไม่มีอำนาจจำแนก (0)	-	-	-
3. ข้อที่อำนาจจำแนกเป็นลบ (-)	9	6	11
รวม	75	80	122

จากตารางพบว่า ข้อสอบที่มีค่าอำนาจจำแนกติดลบร้อยละ 9.39 ซึ่งแสดงว่าคุณภาพของข้อสอบข้อนั้นไม่สามารถจำแนกผู้เรียนเก่งและอ่อนได้ ผู้เรียนอ่อนส่วนมากตอบข้อนั้นถูก แต่ผู้เรียนเก่งส่วนมากไม่ได้คะแนนจากข้อนั้น การนำคะแนนจากข้อสอบที่มีค่าอำนาจจำแนกติดลบไปประเมินผลการศึกษานักศึกษาย่อมส่งผลกระทบต่อนักศึกษาโดยตรงทั้งด้านการประเมินผู้เรียนระหว่างกระบวนการเรียนรู้ การตัดเกรดภายหลังการสอน จากเหตุผลดังกล่าวทำให้ผู้วิจัยสนใจที่จะทำการศึกษาเกี่ยวกับการประเมินคุณภาพของข้อสอบในรายวิชาการพยาบาลเด็กและวัยรุ่นและเปรียบเทียบการเปลี่ยนแปลงของคะแนนการสอบของนักศึกษารายบุคคล ตลอดจนลำดับคะแนนจากการสอบของนักศึกษาก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออกพร้อมทั้งผลที่เกิดขึ้นต่อการเปลี่ยนแปลงค่าความเชื่อมั่นของข้อสอบทั้งฉบับเพื่อให้ได้ข้อมูลที่เป็นหลักฐานเชิงประจักษ์นำมาสู่ข้อสรุปหรือข้อเสนอแนะเชิงนโยบายในด้านการวัดและประเมินผลการจัดการเรียนการสอนในหลักสูตรต่อไป

วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษา

1. คุณภาพของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่น
2. การเปลี่ยนแปลงคะแนนการสอบของนักศึกษาในรายวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก
3. การเปลี่ยนแปลงของลำดับคะแนนจากการสอบของนักศึกษาในวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก
4. การเปลี่ยนแปลงค่าความเชื่อมั่นของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นทั้งฉบับหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ขอบเขตการวิจัย

ขอบเขตด้านประชากร คือ นักศึกษาหลักสูตรพยาบาลศาสตรบัณฑิตชั้นปีที่ 2 ซึ่งลงทะเบียนเรียนและเข้าสอบในรายวิชาการพยาบาลเด็กและวัยรุ่น ประจำภาคการเรียนที่ 1 ปีการศึกษา 2555 จำนวน 88 คน

ขอบเขตด้านเนื้อหา คือ วิเคราะห์ข้อสอบและคะแนนสอบในรายวิชาการพยาบาลเด็กและวัยรุ่นของนักศึกษาหลักสูตรพยาบาลศาสตรบัณฑิตชั้นปีที่ 2 ประจำภาคการเรียนที่ 1 ปีการศึกษา 2555 ซึ่งเป็นข้อสอบแบบเลือกตอบ 4 ตัวเลือก จำนวน 4 ฉบับ แบ่งเป็นการสอบครั้งที่ 1 ครั้งที่ 2 ครั้งที่ 3 และการสอบปลายภาคจำนวน 75, 70, 58 และ 75 ข้อตามลำดับ โดยคุณภาพของข้อสอบในการวิจัยครั้งนี้ประกอบด้วย

1. คุณภาพข้อสอบรายข้อ พิจารณาจากค่าความยากง่ายและค่าอำนาจจำแนก
2. คุณภาพข้อสอบทั้งฉบับ พิจารณาจากค่าความเชื่อมั่น ส่วนของคุณภาพข้อสอบด้านความตรงนั้นคณะผู้สอนได้ดำเนินการจัดทำตารางวิเคราะห์หลักสูตรประกอบการออกข้อสอบไว้ในรายละเอียดรายวิชา (มคอ.3) เพื่อให้เกิดความตรงเชิงโครงสร้างและมีการวิพากษ์ข้อสอบโดยคณะผู้สอนร่วมกับกรรมการบริหารหลักสูตรก่อนนำข้อสอบไปใช้ เพื่อให้เกิดความตรงตามเนื้อหา

ขอบเขตด้านเวลา คือ ช่วงเวลาของการสอบครั้งที่ 1 ครั้งที่ 2 ครั้งที่ 3 และสอบปลายภาคของภาคการเรียนที่ 1 ปีการศึกษา 2555 ตามวันและเวลาการสอบนอกตารางที่มหาวิทยาลัยราชภัฏสวนดุสิตกำหนดขึ้น

คำจำกัดความที่ใช้ในงานวิจัย

1. คุณภาพของข้อสอบ หมายถึง ผลการวิเคราะห์ข้อสอบเป็นรายข้อและทั้งฉบับในวิชาการพยาบาลเด็กและวัยรุ่นจากการสอบครั้งที่ 1 ครั้งที่ 2 ครั้งที่ 3 และการสอบปลายภาค สำหรับคุณภาพข้อสอบรายข้อพิจารณาจากค่าความยากง่าย (p) และค่าอำนาจจำแนกของข้อสอบ (r) ส่วนคุณภาพข้อสอบทั้งฉบับพิจารณาจากค่าความเชื่อมั่นของข้อสอบดังต่อไปนี้

1.1 ค่าความยากง่าย (Difficulty) หมายถึง สัดส่วนของจำนวนนักศึกษาผู้ตอบข้อสอบข้อนั้นได้ถูกต้องต่อจำนวนผู้ตอบข้อสอบทั้งหมดและถือว่าข้อสอบที่มีค่าความยากง่ายเหมาะสม (p) อยู่ระหว่าง 0.20 - 0.80

1.2 ค่าอำนาจจำแนก (Discrimination power) หมายถึง คุณสมบัติของข้อสอบที่สามารถจำแนกความแตกต่างของนักศึกษาออกเป็นกลุ่มที่เรียนเก่งหรือได้คะแนนสูง และกลุ่มที่เรียนอ่อนหรือได้คะแนนต่ำอย่างถูกต้องและถือว่าข้อสอบที่มีค่าอำนาจจำแนก (r) ที่เหมาะสมคือ ≥ 0.20

ส่วนข้อสอบที่มีค่าอำนาจจำแนกติดลบ (Negative discrimination power) เป็นข้อสอบที่นักศึกษาเรียนเก่งส่วนใหญ่ตอบผิด แต่นักศึกษาที่เรียนอ่อนส่วนใหญ่ตอบถูก จึงไม่สามารถจำแนกนักศึกษากลุ่มที่เรียนเก่งและกลุ่มที่เรียนอ่อนได้

1.3 ค่าความเชื่อมั่นของข้อสอบ (Reliability) หมายถึงคุณสมบัติของข้อสอบที่มีความคงที่ของคะแนนหรือผลที่ได้จากการนำข้อสอบไปวัด ถ้านำข้อสอบไปทดสอบกับนักศึกษาคนเดิมก็ได้คะแนนคงเดิมหรือใกล้เคียงกับคะแนนเดิม คำนวณได้จากสูตร Kuder - Richardson formula 20

2. การตัดข้อสอบที่มีอำนาจจำแนกติดลบออก หมายถึง การคิดคะแนนของข้อสอบใหม่ทั้งฉบับภายหลังจากการวิเคราะห์ข้อสอบ โดยไม่ตรวจให้คะแนนข้อสอบที่มีค่าอำนาจจำแนกติดลบและปรับคะแนนรวมของข้อสอบที่เหลือให้เป็นร้อยละ 100 เช่นเดิม

3. คะแนนการสอบ หมายถึง ค่าหรือตัวเลขที่ใช้แทนปริมาณการตอบข้อสอบได้ถูกต้อง โดยให้ค่าคะแนนตอบถูกเป็น 1 และตอบผิดเป็น 0

ประโยชน์ที่คาดว่าจะได้รับ

1. เป็นแนวทางในการประเมินผลการเรียนการสอนในรายวิชาการพยาบาลเด็กและวัยรุ่นที่มีความเชื่อมั่นสูง สามารถนำไปใช้ในรายวิชาหรือหลักสูตรอื่นๆได้

2. เป็นแนวทางในการปรับปรุงข้อสอบ วิธีการประเมินผล และการจัดการเรียนการสอนในรายวิชาการพยาบาลเด็กและวัยรุ่น

3. ได้ข้อสรุปหรือข้อเสนอแนะเชิงนโยบายในด้านการวัดและประเมินผลการจัดการเรียนการสอนในหลักสูตรพยาบาลศาสตรบัณฑิต

บทที่ 2

แนวคิด ทฤษฎี เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยเรื่อง ผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบและความเชื่อมั่นของข้อสอบทั้งฉบับ: รายวิชาการพยาบาลเด็กและวัยรุ่น ผู้วิจัยได้ศึกษาค้นคว้าจากตำรา เอกสาร หนังสือ วารสารและงานวิจัยที่เกี่ยวข้อง สรุปสาระสำคัญและนำเสนอตามลำดับดังนี้

1. แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน

2. การวิเคราะห์คุณภาพของข้อสอบ

2.1 ประโยชน์ของการวิเคราะห์ข้อสอบ

2.2 วิธีการวิเคราะห์ข้อสอบ

2.3 การวิเคราะห์ข้อสอบเป็นรายข้อ

2.3.1 การวิเคราะห์ข้อสอบเป็นรายข้อแบบอิงกลุ่ม

1) ค่าความยากง่าย

2) ค่าอำนาจจำแนก

3) ประสิทธิภาพของตัวलग

2.3.2 การวิเคราะห์ข้อสอบเป็นรายข้อแบบอิงเกณฑ์

1) ความสอดคล้องระหว่างข้อสอบกับจุดมุ่งหมาย

2) ค่าความยากง่าย

3) ค่าอำนาจจำแนกหรือดัชนีความไว

4) ประสิทธิภาพของตัวलग

2.4 การวิเคราะห์คุณภาพข้อสอบทั้งฉบับ

2.4.1 ความตรงหรือความเที่ยงตรง

2.4.2 ความเชื่อมั่นหรือความเที่ยง

3. งานวิจัยที่เกี่ยวข้อง

แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน

กระบวนการศึกษา เป็นกระบวนการที่จัดขึ้นเพื่อพัฒนาผู้เรียนให้มีพฤติกรรมบรรลุตาม จุดมุ่งหมายที่วางไว้ประกอบด้วยองค์ประกอบหลัก 3 ประการคือ จุดมุ่งหมายทางการศึกษา ประสบการณ์ในการเรียนรู้และการประเมินผล ซึ่งการวัดและประเมินผลนั้นถือว่ามีความสำคัญมาก เพราะช่วยในการค้นหาสมรรถภาพของผู้เรียน นำไปสู่การพัฒนาผู้เรียนให้บรรลุเป้าหมายของ การศึกษา โดยนำไปเป็นข้อมูลในการปรับปรุงการเรียนการสอน ทำให้ผู้เรียนเกิดการพัฒนาตนเอง ในทางที่เหมาะสม โดยพิจารณาจากความสามารถของตน เป็นข้อมูลสำหรับแนะแนวผู้เรียนผู้ปกครอง และผู้บริหารในการนำมาปรับปรุงและพัฒนาการจัดการศึกษา ในการวัดและประเมินผลการศึกษาทุก ครั้งต้องมีการนำผลที่ได้จากการวัดไปใช้อธิบาย เปรียบเทียบและสรุปผลเกี่ยวกับคุณลักษณะที่วัดได้ วิธีการหนึ่งที่จะทำให้ผลการวัดมีความถูกต้อง แม่นยำและน่าเชื่อถือคือ การใช้เครื่องมือที่มีคุณภาพ เครื่องมือที่นิยมใช้มากที่สุดคือ แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน

ชนิดแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน

ชนิดของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนหรือข้อสอบอาจแบ่งเป็น 3 รูปแบบคือ ข้อสอบปากเปล่า ข้อสอบข้อเขียน และข้อสอบภาคปฏิบัติ สำหรับข้อสอบข้อเขียนนั้นเป็นข้อสอบที่ ส่วนใหญ่นำมาใช้ในการวัดผลการศึกษาและใช้ในการวิจัยครั้งนี้ สามารถแบ่งตามลักษณะของคำถาม และวิธีการตอบเป็น 2 ชนิดประกอบด้วย (สุรศักดิ์ อมรรัตนศักดิ์, 2554, น. 191 – 200; เยาวดี ราชชัยกุล วิบูลย์ศรี, 2552, น. 219 - 227)

1. แบบเขียนตอบ

ข้อสอบแบบนี้ไม่มีคำตอบให้มา ผู้ตอบต้องคิดหาคำตอบเอง แบ่งออกเป็น 3 ประเภท ดังนี้

1.1 แบบเติมความให้สมบูรณ์ ข้อสอบประเภทนี้จะมีข้อความหรือที่ไม่สมบูรณ์มา ให้ผู้ตอบคิดหาคำความมาเติมลงในช่องว่างให้ได้ความสมบูรณ์และถูกต้อง

1.2 แบบตอบสั้น เป็นข้อสอบที่ให้ผู้สอบเขียนตอบสั้นๆ มักเป็นคำถามให้ระลึกถึง ความรู้เกี่ยวกับความจริง

1.3 แบบอรรถนัย เป็นข้อสอบที่ต้องการให้ผู้ตอบอธิบาย บรรยาย อภิปราย เสนอ ความคิดเห็นหรือวิพากษ์วิจารณ์ ผู้ตอบต้องใช้วิธีเขียนตอบแบบความเรียง บางครั้งจึงเรียกข้อสอบ แบบนี้ว่า ข้อสอบแบบความเรียง

2. แบบให้ผู้สอบเลือกคำตอบ

ข้อสอบประเภทนี้จะมีคำตอบมาให้ ผู้สอบไม่ต้องคิดคำตอบเองเพียงแต่เลือกกว่าคำตอบใด เป็นคำตอบที่ถูกต้องที่สุดหรือเลือกกว่าคำตอบที่ให้นั้นถูกหรือผิด ซึ่งอาจแบ่งออกเป็น 3 ประเภทคือ

2.1 แบบถูก - ผิด เป็นแบบทดสอบที่เหมาะสมสำหรับใช้วัดความจำ เกี่ยวข้องกับข้อเท็จจริงที่มีความสำคัญสำหรับเนื้อหาต่างๆ ผู้ตอบจะต้องมีความสามารถในการจำแนกข้อความของคำถามว่าถูกหรือผิดออกมาให้ได้

2.2 แบบจับคู่ มักใช้วัดในระดับการเรียนรู้ขั้นต้น เช่น วัดความจำเกี่ยวกับบุคคล สถานที่ หรือเหตุการณ์ต่างๆ ความหมายของคำจำกัดความ ความรู้ที่เป็นข้อเท็จจริงต่างๆ และความสัมพันธ์ระหว่างสิ่งต่างๆ ข้อสอบประเภทนี้จะให้ผู้ตอบจับคู่คำ วลี ข้อความหรือประโยคที่กำหนดให้กลุ่มหนึ่งที่มีความสัมพันธ์เกี่ยวข้องกับอีกกลุ่มหนึ่ง การสร้างแบบสอบประเภทจับคู่ให้มีมาตรฐานจะทำให้โอกาสที่ตอบได้ถูกจากการเดา มีน้อยกว่าแบบสอบประเภทถูก - ผิด

2.3 แบบเลือกตอบหรือหลายตัวเลือก ข้อสอบประเภทนี้จะให้ผู้สอบเลือกคำตอบหรือเหมาะสมที่สุด ข้อสอบแบบเลือกตอบประกอบด้วยคำถามและตัวเลือก ตัวเลือกนี้จะมีตัวถูกอยู่ตัวเดียว ส่วนตัวอื่นๆเป็นตัวลวง

การจะเลือกใช้ข้อสอบแบบใดขึ้นอยู่กับจุดประสงค์ของการสอบเป็นหลัก ในที่นี้จะกล่าวถึงเฉพาะแบบทดสอบชนิดเลือกตอบซึ่งเป็นแบบทดสอบที่ใช้ในการวิจัยครั้งนี้

ข้อสอบชนิดเลือกตอบ (Multiple choice Question, MCQ)

ข้อสอบชนิดเลือกตอบเป็นแบบสอบที่นิยมใช้กันมากในปัจจุบัน เหมาะสำหรับการประเมินด้านพุทธิพิสัยของผู้เรียน (อานุกาพ เลขะกุล, 2556, น. 2) เป็นข้อสอบที่สามารถจำแนกระดับความรู้ต่างๆ ได้ดีกว่าแบบถูก - ผิดหรือข้อสอบประเภทอื่น โดยเฉพาะอย่างยิ่งถ้าสร้างให้มีมาตรฐาน สามารถวัดระดับการเรียนรู้ในขั้นสูง เช่น ระดับการวิเคราะห์ การสังเคราะห์และการประเมินผล ซึ่งข้อสอบถูก- ผิดหรือแบบจับคู่ส่วนใหญ่จะวัดได้เพียงระดับความรู้ ความจำและความเข้าใจ ส่วนขั้นนำไปใช้สามารถวัดได้บ้างแต่เป็นเพียงส่วนน้อย การสร้างข้อสอบประเภทหลายตัวเลือกให้มีคุณภาพดีนั้น ผู้สร้างจำเป็นต้องมีความรู้เข้าใจในหลักการสร้างข้อกระทง เข้าใจกระบวนการคิดของคน ตลอดจนเข้าใจเนื้อหาวิชาที่จะนำมาทดสอบประกอบกันเป็นอย่างดี (เยาวดี รางชัยกุล วิบูลย์ศรี, 2552, น. 224)

ลักษณะของข้อสอบชนิดเลือกตอบ

โครงสร้างของข้อสอบประกอบด้วยข้อกระทง 2 ส่วนใหญ่ๆคือ ส่วนแรกเป็นคำถามนำและส่วนที่เป็นตัวเลือก ข้อกระทงที่เป็นคำถามนำประกอบด้วยส่วนที่เป็นคำถามหลักแต่ละข้อ ส่วนข้อกระทงที่เป็นตัวเลือกประกอบด้วยตัวคำตอบ 1 ตัวและที่เหลือเป็นตัวลวง ลักษณะของคำถามอาจเป็น

คำถามโดยตรงเหมาะที่จะใช้กับผู้สอบที่มีระดับความรู้ต่ำ หรืออาจเป็นข้อความที่ไม่สมบูรณ์ที่มีลักษณะสั้น กะทัดรัด ได้ใจความและใช้ภาษาที่ไม่ฟุ่มเฟือยจึงเป็นที่นิยมใช้กันมาก ลักษณะของตัวคำตอบอาจมีคำตอบที่ถูกต้องเพียงคำตอบเดียว หรืออาจเป็นคำตอบที่ดีที่สุด เมื่อเทียบกับคำตอบอื่นๆ นอกนั้นเป็นตัวลวง (เยาวตี ราชชัยกุล วิบูลย์ศรี, 2552, น. 224 - 225) โดยมีแนวทางในการสร้างข้อสอบ ข้อดีและข้อจำกัดดังนี้

แนวทางในการสร้างข้อสอบชนิดเลือกตอบ

Tavakol M. & Dennick R. (2001, pp. 447 - 458) กล่าวว่าเป้าหมายที่สำคัญประการหนึ่งของการประเมินความรู้ความสามารถของนักศึกษาในสาขาวิทยาศาสตร์สุขภาพ คือการลดความคลาดเคลื่อนในข้อสอบให้มีน้อยที่สุดเท่าที่จะทำได้ เพื่อที่จะได้ผลคะแนนสอบที่สามารถบ่งบอกระดับความรู้ความสามารถที่แท้จริงของนักศึกษามากที่สุด ดังนั้นในการสร้างข้อสอบจะต้องคำนึงองค์ประกอบที่อาจก่อให้เกิดความลำเอียงในการประเมินผล เริ่มตั้งแต่ขั้นตอนของการออกข้อสอบจนถึงการแปลผลคะแนนที่ได้จากการสอบ ดังนี้

1. ผู้สอนจะต้องเลือกประเด็นที่นักศึกษาจำเป็นต้องเรียนรู้หรือสามารถปฏิบัติได้
2. ผู้สอนจะต้องออกแบบการจัดการเรียนการสอนหรือจัดประสบการณ์การเรียนรู้ ทั้งในการสอนภาคทฤษฎีในชั้นเรียน และการเรียนการสอนภาคปฏิบัติในสถานบริการสุขภาพ ซึ่งเป็นแหล่งฝึกเพื่อให้นักศึกษาเกิดการเรียนรู้ในเนื้อหา และทักษะที่ผู้สอนคาดหวังไว้ให้มากที่สุด
3. ผู้สอนจะต้องมีการประเมินระดับความรู้ ความสามารถในทักษะและเนื้อหาที่จำเป็นของนักศึกษาแต่ละราย ส่วนใหญ่จะใช้วิธีการทำข้อสอบ โดยผู้สอนจะต้องมีการทดสอบความตรง ความเที่ยงหรือค่าความเชื่อมั่นของชุดข้อสอบ เพื่อให้เกิดความแน่ใจได้ว่าข้อสอบในชุดนั้นๆสามารถใช้วัดระดับความรู้ ความสามารถ ทักษะของนักศึกษาในเนื้อหาที่ผู้สอนต้องการได้อย่างแท้จริง และควรมีการคำนวณค่าความยากง่ายของตัวเลือกที่ถูกต้องในข้อสอบแต่ละข้อ เพื่อเป็นการประเมินความยากง่ายรายข้อ ซึ่งข้อสอบข้อที่ง่ายคือ ข้อสอบที่มีนักศึกษาส่วนใหญ่ (มากกว่าร้อยละ 60) เลือกตอบตัวเลือกที่ถูกต้อง และการแปลผลข้อสอบที่ยากมากคือข้อสอบที่มีนักศึกษาเลือกตอบในตัวเลือกที่ถูกต้องน้อยกว่าร้อยละ 30 ของจำนวนนักศึกษาที่ทำข้อสอบทั้งหมด หากข้อสอบข้อใดไม่มีความแตกต่างระหว่างคะแนนเฉลี่ยของนักศึกษาที่ตอบถูกและตอบผิด แสดงว่าข้อสอบข้อนั้นไม่สามารถจำแนกนักศึกษากลุ่มอ่อนและกลุ่มเก่งออกจากกันได้ ควรนำข้อสอบนั้นมาปรับปรุงหรือตัดข้อสอบนั้นออกจากชุดข้อสอบ

ในการสร้างข้อสอบชนิดเลือกตอบต้องพิจารณาทั้งการเขียนข้อคำถามและการเขียนตัวเลือก ข้อสอบแบบเลือกตอบที่ดีต้องประกอบด้วยข้อคำถามที่ชัดเจน สื่อสารได้เข้าใจและสอดคล้องกับ

จุดประสงค์การเรียนรู้ที่กำหนด เพื่อให้ข้อสอบนั้นมีความตรงตามเนื้อหา โดยมีแนวทางในการสร้างข้อสอบดังนี้ (โชติกา ภาชีผล, 2554, น. 18 - 21; อานุกาพ เลขะกุล, 2556, น. 7 - 20)

1. การเขียนข้อคำถามหรือโจทย์

- 1.1 เนื้อหาของโจทย์และคำถามควรสอดคล้องกับวัตถุประสงค์และตารางข้อสอบ
- 1.2 โจทย์ควรคล้ายคลึงกับสถานการณ์จริงที่ผู้สอบจะต้องประสบในการนำไปใช้
- 1.3 ระดับความยากง่ายเหมาะสมกับผู้สอบ เนื้อหาไม่สูงกว่าพื้นฐานความรู้ที่เรียน
- 1.4 ควรเน้นความรู้ที่นำไปใช้มากกว่าแบบท่องจำ ได้แก่ ระดับความเข้าใจขึ้นไป
- 1.5 การสร้างโจทย์ที่เป็นปัญหาในคลินิก หากเป็นวิชาที่สอนก่อนนักศึกษาชั้นฝึกปฏิบัติงานในคลินิกไม่ควรสลับซับซ้อน และควรโยงไปสู่ความรู้พื้นฐานมากกว่าการประยุกต์ใช้ในการดูแลผู้ป่วย เพราะนักศึกษายังไม่เกิดประสบการณ์ในคลินิก หากเป็นวิชาที่ผู้สอบผ่านการเรียนรู้ทางคลินิกมาแล้วโจทย์อาจซับซ้อนมากขึ้นและอาศัยทักษะการปฏิบัติตามกระบวนการพยาบาลมาออกได้ เนื้อหาโจทย์ควรเป็นสาระสำคัญๆ หรือเป็นปัญหาที่พบบ่อย
 - 1.6 โจทย์ควรอยู่ในรูปแบบของประโยคคำถามที่มีใจความหรือมีความหมายในตัว ประโยคเอง การใช้ประโยคที่มีใจความสมบูรณ์จะทำให้ข้อสอบดูง่ายขึ้นและมีผลต่อการเพิ่มค่าความเที่ยงของข้อสอบเล็กน้อย
 - 1.7 ข้อมูลที่จำเป็นทั้งหมดควรปรากฏอยู่ในโจทย์ ไม่ควรนำไปเพิ่มเติมในตัวเลือก เพราะจะเป็นการสร้างโจทย์ใหม่หลายข้อในข้อสอบข้อเดียว
 - 1.8 พยายามเน้นเรื่องที่ต้องการถามให้เด่นชัด ไม่ถามคลุมเครือ อ่านแล้วตีความยาก เข้าใจไม่ตรงกัน ทำให้ขาดความเป็นปรนัย
 - 1.9 ใช้คำที่มีความหมายชัดเจน สั้นกระชับ ไม่ซับซ้อน ไม่ฟุ่มเฟือย แต่มีข้อมูลเพียงพอ สำหรับการตอบคำถามได้ สื่อความหมายด้วยภาษาเขียนที่สละสลวย ทำให้ผู้สอบอ่านโจทย์ และตัวเลือกได้เข้าใจง่าย ชัดเจน ไม่สับสน
 - 1.10 หลีกเลี่ยงโจทย์และคำถามที่เป็นเชิงลบหรือปฏิเสธ เช่น ไม่ ยกเว้น ไม่ถูกต้อง เป็นต้น ถ้าจะใช้ควรขีดเส้นใต้หรือเป็นตัวทึบและเมื่อใช้คำปฏิเสธในโจทย์แล้วไม่ควรใช้คำปฏิเสธในตัวเลือกอีกเพราะจะทำให้ผู้สอบสับสนได้
 - 1.11 หลีกเลี่ยงโจทย์ที่เป็นคำบรรยาย รูป หรือลักษณะที่ทำให้ผู้สอบตีความหมายหรือสร้างจินตนาการเอง เพราะผู้สอบแต่ละคนอาจมีการรับรู้และตีความแตกต่างกัน หากต้องการให้แปลผล อาจใช้รูปภาพประกอบแทน
 - 1.12 ใช้ภาษาให้เหมาะกับระดับของผู้สอบ หลีกเลี่ยงการใช้ภาษาอังกฤษ เพราะจะทำให้การวัดไม่ตรงกับพฤติกรรมการเรียนรู้ที่เป็นวัตถุประสงค์รายวิชา ยกเว้นเป็นคำศัพท์ทางเทคนิคที่ผู้เรียนจำเป็นต้องทราบ

1.13 ไม่ใช่คำถามหรือโจทย์ที่เป็นการแนะนำคำตอบ หรือข้อความในคำถามซ้ำกับข้อความในตัวเลือกที่เป็นคำตอบถูก

1.14 อย่าให้ข้อสอบข้อหนึ่งส่งผลต่อการชี้แนะคำตอบในข้ออื่นๆ

1.15 ไม่ควรลอกประโยคจากตำรามาเขียนเป็นข้อสอบ แต่ควรจัดเรียบเรียงใหม่

2. การเขียนตัวเลือก

2.1 จำนวนตัวเลือกควรมี 4 หรือ 5 ตัวเลือก แต่ 5 ตัวเลือกจะมีความเที่ยงดีที่สุด การใช้ตัวเลือกมากกว่า 5 ข้อ ไม่มีความจำเป็นและไม่ทำให้ค่าอำนาจจำแนกสูงขึ้น

2.2 การเขียนตัวเลือกที่ถูกต้องมีเพียงคำตอบเดียว

2.3 ตัวเลือกแต่ละข้อควรเป็นอิสระต่อกัน มีความเป็นไปได้ที่จะเป็นคำตอบที่ถูกต้อง สมเหตุสมผล และดึงดูดให้ผู้สอบเลือก

2.4 ความยาวของเนื้อหาของตัวเลือกแต่ละข้อควรจะใกล้เคียงกัน และไม่ควรรยาวจนเกินไป เนื่องจากคำตอบที่ถูกต้องมักมีแนวโน้มมีความยาวมากกว่าตัวเลือกที่ผิด ควรเป็นข้อความสั้นๆและไม่ควรเกินหนึ่งบรรทัด

2.5 เนื้อหาของตัวเลือกควรเป็นเรื่องในกลุ่มเดียวกัน ประเด็นที่คล้ายกันหรือเป็นเรื่องเดียวกันหรือเป็นไปในทิศทางเดียวกัน

2.6 หากตัวเลือกเป็นค่าตัวเลข ควรจะคงที่ เช่น ค่าตัวเลขเป็นช่วงก็ต้องเป็นช่วงทุกข้อ ไม่ใช่มีตัวเลือกบางข้อเป็นค่าเดี่ยวและควรเรียงลำดับกันไป

2.7 หลีกเลี่ยงตัวเลือกที่ซับซ้อนหรือทำให้ผู้ตอบสับสน

2.8 หลีกเลี่ยงตัวเลือกที่ขัดแย้งกันเองหรือตรงกันข้าม เช่น ตัวเลือกหนึ่งเป็น “residual volume เพิ่มขึ้น” อีกตัวเลือกหนึ่งเป็น “residual volume ลดลง” หากตัวคำตอบถูกเป็นตัวใดตัวหนึ่งในนี้ ตัวเลือกจะลดลงโดยปริยายไปหนึ่งตัวเลือก

2.9 เนื้อหาในตัวเลือกแต่ละข้อไม่ควรทับซ้อนกัน เช่น คำตอบที่ถูกเป็น “ให้ยาลดความดันโลหิต” ตัวลวงข้อหนึ่งเป็น “ให้ ACE inhibitor” ซึ่งเป็นกลุ่มยาลดความดันโลหิตพวกหนึ่ง

2.10 หลีกเลี่ยงการใช้คำต่อไปนี้ ได้แก่ ทั้งหมด เสมอ ไม่เคย เป็นต้น ซึ่งตัวเลือกเหล่านี้มักจะเป็นคำตอบที่ผิด

2.11 หลีกเลี่ยงตัวเลือกที่เป็นปฏิเสธหรือปฏิเสธซ้อนปฏิเสธ เพราะทำให้ผู้อ่านงงได้

2.12 หลีกเลี่ยงตัวเลือก “ถูกหมดทุกข้อ” จะทำให้เดาได้ง่าย เพราะถ้ามีคำตอบถูก 2 ข้อไม่ทราบตัวเลือกที่สาม ก็จะได้ข้อถูกหมดทุกข้อหรือทราบตัวเลือกผิด 1 ข้อ ก็จะไม่เลือกข้อถูกหมดทุกข้อ

2.13 หลีกเลี่ยงตัวเลือกถูก “ผิดหมดทุกข้อ” เพราะผู้ตอบถูก อาจเลือกเพราะความไม่รู้ ซึ่งเป็นการเดา

2.14 ถ้าตัวเลือกใช้คำซ้ำกัน ควรนำไปเขียนไว้ในประโยคคำถาม หรือแยกเขียนไว้เป็นหัวข้อต่างหาก

2.15 ตำแหน่งของคำตอบที่ถูกไม่ควรซ้ำกัน หรือไม่ควรมีรูปแบบเดียวกันตลอด ควรกระจายให้มีสัดส่วนที่ใกล้เคียงกัน

2.16 การจัดลำดับของตัวเลือก อาจทำได้หลายรูปแบบ ดังนี้

2.16.1 ตัวเลือกเป็นคำควรเรียงตามลำดับตัวอักษร

2.16.2 ตัวเลือกเป็นค่าตัวเลข ควรเรียงจากน้อยไปมาก

2.16.3 จัดเรียงตามลำดับเหตุการณ์ที่เกิดขึ้น

2.16.4 จัดเรียงตามกลุ่มที่มีแนวโน้มเหมือนกัน เช่น ลดลง เพิ่มขึ้น

2.16.5 เรียงตามความยาวของข้อความจากสั้นไปยาว ทำให้ผู้สอบอ่านง่ายขึ้น

2.16.6 อาจใช้ข้อผิดพลาดหรือความเข้าใจผิดที่พบบ่อยของผู้เรียนเป็นตัวลวง

2.16.5 หลีกเลี่ยงตัวลวงที่ผิดชัดเจนเกินไปหรือไม่มีเหตุผลจนไม่มีใครเลือก

2.16.6 หลีกเลี่ยงการสร้างตัวลวงจากความซับซ้อนหรือสับสนของคำตอบ

การวิเคราะห์คุณภาพของข้อสอบ

การวิเคราะห์ข้อสอบหมายถึง การประเมินหรือตรวจสอบคุณภาพของแบบทดสอบว่ามีความเหมาะสมจะนำไปใช้ในการวัดและประเมินผลเพียงใด โดยมีการพิจารณาทั้งแบบเป็นรายข้อและทั้งฉบับ การประเมินคุณภาพของข้อสอบหรือการวิเคราะห์คุณภาพของข้อสอบมีความสำคัญอย่างยิ่งในการจัดการเรียนการสอน เพราะจะช่วยให้การวัดและประเมินผลการศึกษามีประสิทธิภาพดียิ่งขึ้น ข้อสอบที่สร้างขึ้นเพื่อวัดหรือประเมินผลการศึกษานั้นจะมีคุณภาพเพียงใดขึ้นกับการวางแผนการสร้างข้อสอบ ความลึกซึ้งในเนื้อหาวิชาของผู้เขียนข้อสอบและทักษะการเขียนข้อสอบ ซึ่งสามารถทบทวนลักษณะทางกายภาพของข้อสอบ และศึกษาความเหมาะสมของข้อสอบในเชิงเหตุผลได้ แต่ยังไม่สามารถทราบถึงคุณภาพของข้อสอบในเชิงประจักษ์ จนกว่าผู้สร้างข้อสอบจะนำข้อสอบไปใช้หรือทดลองใช้ เพื่อนำผลการตอบข้อสอบมาวิเคราะห์คุณภาพข้อสอบ (ศิริชัย กาญจนวาสี, 2556, น. 221)

ประโยชน์ของการวิเคราะห์ข้อสอบ (พิชิต ฤทธิจรูญ, 2553, น. 139 - 140; ศิริชัย กาญจนวาสี, 2556, น. 222)

1. ทำให้ทราบข้อมูลพื้นฐานของตัวข้อสอบและคำตอบ สามารถนำไปอภิปรายเกี่ยวกับผลสอบได้ว่า ข้อสอบแต่ละข้อได้ทำหน้าที่วัดผลสัมฤทธิ์อย่างมีคุณภาพเพียงใด

2. ทำให้ได้ข้อสอบที่มีคุณภาพ สามารถนำไปวัดและประเมินผลได้อย่างเที่ยงตรง น่าเชื่อถือ
3. ชี้ให้เห็นถึงจุดอ่อนที่ผู้สอนหรือผู้เรียนต้องปรับปรุงแก้ไข เพื่อพัฒนาความสามารถและทักษะของผู้เรียนให้เป็นตามเป้าหมาย
4. เป็นข้อมูลพื้นฐานที่สำคัญในการปรับปรุงพัฒนาการเรียนการสอน โดยเฉพาะอย่างยิ่งด้านพฤติกรรมการเรียนรู้ที่เหมาะสมสำหรับเนื้อหาส่วนนั้นหรือวิชานั้นๆ เพราะการวิเคราะห์ข้อสอบจะช่วยให้ทราบจุดบกพร่องของการเรียนการสอน
5. ช่วยเพิ่มทักษะในการสร้างข้อสอบ ทำให้ทราบถึงอำนาจการจำแนก ระดับความยากง่าย ประสิทธิภาพของตัวลอง ตลอดจนการเขียนข้อสอบในลักษณะใดจึงจะได้ข้อสอบที่ดี เพราะผลการประเมินจะเป็นข้อมูลย้อนกลับไปยังผู้ออกข้อสอบ
6. สามารถคัดเลือกข้อสอบที่มีคุณภาพมารวมเป็นฉบับข้อสอบที่ดี จัดทำแบบสอบคู่ขนานเป็นรากฐานสำคัญในการพัฒนาแบบสอบมาตรฐานและเป็นแนวทางในการสร้างคลังข้อสอบ

วิธีการวิเคราะห์ข้อสอบ

การตรวจสอบคุณภาพของแบบทดสอบเพื่อให้ผลการวัดมีความเที่ยงตรง น่าเชื่อถือ อาจประเมินได้ใน 2 ลักษณะคือ การวิเคราะห์เป็นรายข้อและการวิเคราะห์ข้อสอบทั้งฉบับ ในที่นี้จะกล่าวถึงเฉพาะวิธีในการวิเคราะห์ข้อสอบแบบเลือกตอบซึ่งเป็นชนิดข้อสอบที่ใช้ในการวิจัยครั้งนี้

การวิเคราะห์คุณภาพข้อสอบหรือแบบทดสอบ แบ่งออกเป็น

1. การวิเคราะห์แบบทดสอบรายข้อ (Item analysis)
2. การวิเคราะห์แบบทดสอบทั้งฉบับ (Test analysis)

การวิเคราะห์ข้อสอบเป็นรายข้อ

การวิเคราะห์ข้อสอบเป็นรายข้อ เป็นกระบวนการตรวจสอบคำตอบของผู้สอบในแต่ละข้อ เพื่อพิจารณาว่าข้อสอบแต่ละข้อมีระดับความยากง่าย และค่าอำนาจจำแนกเพียงใด รวมทั้งพิจารณาถึงประสิทธิภาพตัวลองในตัวเลือกของข้อสอบนั้นด้วย ข้อสอบที่จะวิเคราะห์คุณภาพรายข้อนี้จะต้องเป็นข้อสอบแบบปรนัยที่ตอบถูกได้ 1 คะแนน ตอบผิดได้ 0 คะแนน ดังนั้นลักษณะข้อสอบจึงเป็นแบบเลือกตอบ หรือแบบถูก – ผิด หรือแบบเติมคำ แต่ต้องมีรูปแบบเดียวกันทั้งฉบับ การวิเคราะห์ข้อสอบเป็นรายข้อแบ่งเป็น 2 ประเภทดังนี้ (พิชิต ฤทธิจรูญ, 2553, น. 140 - 142; ศิริชัย กาญจนวาสี, 2556, น. 223)

1. การวิเคราะห์ข้อสอบเป็นรายข้อแบบอิงกลุ่ม (Item Analysis Procedure for Norm – Referenced tests)

เป็นการวิเคราะห์โดยการนำผลคะแนนของผู้สอบแต่ละคนไปเปรียบเทียบกับผลคะแนนของผู้เข้าสอบคนอื่นๆในกลุ่มเดียวกัน เพื่อจำแนกผู้ที่เรียนเก่งหรือได้คะแนนสูงออกจากผู้ที่เรียนอ่อนหรือได้คะแนนต่ำ ตัวบ่งชี้คุณภาพของแบบทดสอบเป็นรายข้อแบบอิงกลุ่ม ได้แก่ ระดับความยากง่าย (Level of difficulty) ค่าอำนาจจำแนก (Discriminative power) และประสิทธิภาพของตัวลวง (Distractor analysis) ซึ่งเป็นการวิเคราะห์สัดส่วนของผู้เลือกตัวลวง และค่าอำนาจจำแนกของตัวลวง

1.1 ความยากง่ายของข้อสอบ (Difficulty)

ความยากง่าย เป็นสัดส่วนของผู้ตอบถูกจากผู้ตอบในข้อนั้นทั้งหมด แสดงให้เห็นว่าข้อสอบนั้นมีผู้ตอบถูกมากหรือน้อย ถ้ามีผู้ตอบถูกเป็นจำนวนมากถือว่าข้อสอบง่าย แต่ถ้ามีผู้ตอบถูกน้อยถือว่าเป็นข้อสอบที่ยาก ค่าสัดส่วนนี้เรียกว่า ดัชนีความยากง่าย (Index of difficulty) นิยมเขียนแทนด้วย p การหาค่าความยากง่ายเป็นการวิเคราะห์รายข้อ (Item analysis) ไม่ใช่เป็นการวิเคราะห์ภาพรวมทั้งฉบับ ค่าที่ได้จะอยู่ระหว่าง 0 ถึง 1 คำนวณได้จากสูตรดังนี้ (โชติกา ภาชีผล, 2554, น. 79)

$$p = R / N$$

$$p = \text{ดัชนีความยากง่าย}$$

$$R = \text{จำนวนผู้ตอบถูกทั้งหมด}$$

$$N = \text{จำนวนผู้เข้าสอบทั้งหมด}$$

ในการคำนวณจากคะแนนของผู้สอบกลุ่มสูงและกลุ่มต่ำ สามารถคำนวณจากสูตร ดังนี้

$$p = \frac{H + L}{N}$$

$$p = \text{ดัชนีความยากง่าย}$$

$$N = \text{จำนวนผู้สอบในกลุ่มสูงและกลุ่มต่ำ}$$

$$H = \text{จำนวนผู้ตอบถูกในกลุ่มสูง}$$

$$L = \text{จำนวนผู้ตอบถูกในกลุ่มต่ำ}$$

ข้อคำถามที่มีค่า p มาก แสดงว่าข้อคำถามนั้นมีคนเลือกตอบถูกจำนวนมาก แปลว่าข้อนี้ง่าย ส่วนข้อคำถามที่มีค่า p น้อย แสดงว่าข้อคำถามนั้นมีคนเลือกตอบถูกจำนวนน้อย แปลว่าข้อนี้ยาก

ข้อสอบที่ดีต้องมีระดับความยากเหมาะสมกับสติปัญญาของผู้สอบ ลักษณะวิชา และวัตถุประสงค์ของการวัดผล (บุญใจ ศรีสถิตนรากร, 2555, น. 341) ในการวัดผลทั่วไปนิยมเลือกใช้ข้อสอบที่มีค่าความยากง่ายปานกลาง กล่าวคือไม่ยากเกินระดับความสามารถของผู้เรียนจนผู้เรียนส่วนใหญ่ทำไม่ได้ ขณะเดียวกันต้องไม่ง่ายเกินไปจนผู้เรียนส่วนใหญ่ตอบถูกหมด ข้อสอบที่ง่าย

หรือยากเกินไปจะมีประโยชน์ในการสอบวัดผลน้อยมาก ข้อสอบที่มีค่าความยากง่ายปานกลางคือ ข้อสอบที่มีค่า p อยู่ระหว่าง 0.20 - 0.80 (สุรศักดิ์ อมรรัตนศักดิ์, 2554, น. 139) ข้อคำถามที่มีค่าความยากง่ายพอเหมาะหรือค่า p เข้าใกล้ 0.50 ถือเป็นข้อคำถามที่ดีเพราะจะช่วยให้แปลความหมายคะแนนได้ดี เครื่องมือวัดมีค่าอำนาจจำแนกสูง ทำให้คะแนนมีการกระจายมากหรือมีความแปรปรวนสูง ซึ่งจะมีผลให้เครื่องมือทั้งฉบับมีความเชื่อมั่นสูงตามไปด้วย ข้อคำถามที่ง่ายหรือยากเกินไปจะส่งผลให้มีค่าอำนาจจำแนกต่ำ จึงไม่มีประโยชน์ในการประเมินผลแบบอิงกลุ่มหรืออิงเกณฑ์ (บุญเชิด ภิญญ์ อนันตพงษ์, 2552, น. 77)

โชติกา ภาชีผล (2554, น. 79) และสุมาลี จันทร์ชะลอ (2542, น. 136) ได้ให้ข้อเสนอแนะในการเลือกข้อสอบจากการพิจารณาค่า p ซึ่งใช้วิเคราะห์คุณภาพข้อสอบได้ดังนี้

ตารางที่ 2.1 การวิเคราะห์คุณภาพข้อสอบจากค่าความยากง่าย

ค่าความยากง่าย (p)	ความหมาย	คุณภาพข้อสอบ
0.81 - 1.00	ง่ายมาก	ไม่ดี ควรตัดทิ้งหรือปรับปรุงใหม่
0.61 - 0.80	ค่อนข้างง่าย	พอใช้ได้ ควรเก็บไว้ใช้
0.41 - 0.60	พอเหมาะ	ดีมากเก็บไว้ใช้
0.20 - 0.40	ค่อนข้างยาก	พอใช้ได้ ควรเก็บไว้ใช้
0.00 - 0.19	ยากมาก	ไม่ดี ควรตัดทิ้งหรือปรับปรุงใหม่

แสดงว่าข้อสอบที่เหมาะสมควรมีค่าความยากง่ายอยู่ระหว่าง 0.20 - 0.80 หากพิจารณาภาพรวมของคุณภาพข้อสอบพบว่า ข้อสอบทั้งฉบับควรมีระดับความยากง่ายเฉลี่ยประมาณ 0.50 (ศิริชัย กาญจนวาสี, 2556, น. 223)

1.2 ค่าอำนาจจำแนก (Discriminative power)

ค่าอำนาจจำแนก แสดงถึงความสามารถของข้อสอบในการจำแนก หรือแยกให้เห็นความแตกต่างระหว่างผู้สอบที่มีผลสัมฤทธิ์ต่างกัน เช่น จำแนกคนเก่งกับคนอ่อนออกจากกัน หรือจำแนกคนที่มีความสามารถพิเศษกับคนที่ไม่มีความสามารถออกจากกัน โดยถือว่าคนที่เก่งหรือมีความสามารถควรทำข้อสอบนั้นได้ ส่วนผู้ที่อ่อนหรือไม่มีความสามารถไม่ควรทำข้อสอบนั้นได้ (ศิริชัย กาญจนวาสี, 2556, น. 223) หรืออาจกล่าวได้ว่า ค่าอำนาจจำแนกเป็นคุณสมบัติของแบบทดสอบที่สามารถจำแนกความแตกต่างของกลุ่มผู้เรียนเป็น 2 กลุ่ม ได้แก่ กลุ่มเก่งและกลุ่มอ่อน หรือกลุ่มผู้เรียนที่ได้คะแนนสูงและกลุ่มผู้เรียนที่ได้คะแนนต่ำ (บุญใจ ศรีสถิตยน์รากร, 2555, น. 17)

การคำนวณค่าอำนาจจำแนกของข้อสอบ (r) อย่างง่ายสามารถคำนวณได้จากผลต่างระหว่างสัดส่วนจำนวนคนตอบถูกในกลุ่มเก่งกับสัดส่วนจำนวนคนตอบถูกในกลุ่มอ่อน ดังนั้นอำนาจจำแนกข้อสอบจะมีค่าตั้งแต่ - 1 ถึง +1 (ศิริชัย กาญจนวาสี, 2556, น. 223) ข้อคำถามที่ดีจะต้องมีค่าอำนาจจำแนกเป็นบวก เพราะถ้าเป็นลบแสดงว่ามีอำนาจจำแนกกลับกันคือ คนเก่งทำข้อสอบไม่ได้ แต่คนอ่อนทำข้อสอบนั้นได้ ถือว่าเป็นข้อสอบที่ไม่ดี ควรตัดทิ้ง (ปวรส บุตะเขียว, 2556, น. 2)

ถ้าค่าอำนาจจำแนกเป็น 0 แสดงว่าผู้เรียนในกลุ่มเก่งและกลุ่มอ่อนทำข้อสอบข้อนั้นได้เท่าๆกัน ข้อคำถามที่ดีควรมีค่าอำนาจจำแนก 0.20 ขึ้นไป และยังมีค่าเข้าใกล้ +1 ยิ่งดี (Chase, 1978, p. 140)

ค่าอำนาจจำแนกของแบบทดสอบคำนวณได้จากสูตร ดังนี้ (โชติกา ภาชีผล, 2554, น. 80)

$$r = \frac{H - L}{N}$$

r = ค่าอำนาจจำแนกของคำถามข้อนั้น

H = จำนวนผู้ตอบถูกในข้อนั้น ในกลุ่มเก่ง หรือได้คะแนนสูง

L = จำนวนผู้ตอบถูกในข้อนั้น ในกลุ่มอ่อนหรือได้คะแนนต่ำ

N = จำนวนผู้สอบในกลุ่มสูงและกลุ่มต่ำ

ตารางที่ 2.2 การวิเคราะห์คุณภาพข้อสอบจากค่าอำนาจจำแนก (Chase, 1978, p. 140)

ค่าอำนาจจำแนก	ความหมายของคุณภาพข้อสอบ
0.80 - 1.00	อำนาจจำแนกดีมาก
0.60 - 0.79	อำนาจจำแนกดี
0.40 - 0.59	อำนาจจำแนกปานกลาง
0.20 - 0.39	อำนาจจำแนกพอใช้
0.01 - 0.19	อำนาจจำแนกต่ำ (ควรปรับปรุงหรือตัดทิ้ง)
-1.00 - 0.00	อำนาจจำแนกต่ำมาก (ควรตัดทิ้ง)

โดยทั่วไปข้อสอบที่มีค่าอำนาจจำแนกสูงคือ ข้อสอบที่มีค่า r ตั้งแต่ 0.20 ขึ้นไป ทั้งนี้ไพศาลหวังพานิชได้เสนอแนะวิธีสร้างคำถามให้มีค่าอำนาจจำแนกสูงดังนี้ (สุรศักดิ์ อมรรัตนศักดิ์, 2554, น. 141)

1. คำถามต้องมีความชัดเจน หรือมีความเป็นปรนัยสูง
2. คำถามนั้นต้องถามคุณลักษณะที่ใช้ความคิด หรือต้องใช้ดุลยพินิจในการตอบ
3. เป็นข้อคำถามที่เป็นไปตามหลักวิชาการ มีความสมเหตุสมผล
4. กรณีที่เป็นข้อสอบแบบเลือกตอบ ตัวเลือกที่ใช้ต้องมีความเป็นเอกพันธ์

นิธิภัทร บาลศิริ (2557, น. 220 - 221) กล่าวไว้ว่า ถ้ากลุ่มผู้สอบมีความเป็นเอกพันธ์มาก หรือมีความคล้ายคลึงกันมาก เช่น ความรู้ความสามารถเท่าเทียมกัน จะทำให้ความแปรปรวนของ คะแนนน้อยหรือมีการกระจายตัวของคะแนนต่ำ ทำให้แบบวัดมีค่าอำนาจจำแนกต่ำ

Richard F. Burton (2001) กล่าวว่าค่าอำนาจจำแนกเป็นรายชื่อของข้อสอบแบบเลือกตอบ ที่ให้คะแนนเป็น 0 และ 1 สามารถบอกคุณภาพของข้อสอบได้ โดยจะเปรียบเทียบและจำแนกผลการ ทดสอบระหว่างกลุ่มที่มีคะแนนสูงและกลุ่มที่มีคะแนนต่ำ ทั้งนี้กลุ่มทดสอบต้องมีจำนวนมากพอและมีการกระจายของคะแนนข้อสอบจะมีความเที่ยงหรือความเชื่อมั่นสูงเมื่อนำข้อที่มีค่าอำนาจจำแนกต่ำ หรือไม่ดีออกไป นอกจากนี้ความน่าเชื่อถือของดัชนีอำนาจจำแนกของข้อสอบต้องคำนึงถึงผลกระทบ ของการเดาของผู้ตอบแบบทดสอบและความสมบูรณ์ของคำถามด้วย

หากพิจารณาถึงความสัมพันธ์ระหว่าง ความยากง่ายและ อำนาจจำแนกจะพบว่า ถ้าข้อสอบ มีความยากง่าย 0.50 แสดงว่าผู้เรียนจำนวนครึ่งหนึ่งทำข้อสอบข้อนั้นถูก อำนาจจำแนกมีค่าสูงสุด แต่ หากข้อสอบยากมากหรือง่ายมาก ค่าอำนาจจำแนกจะลดลง (สุมาลี จันทร์ชลอ, 2542, น. 138)

1.3 ประสิทธิภาพของตัวลวง (Distractor analysis)

นอกจากจะประเมินคุณภาพข้อสอบรายชื่อโดยการคำนวณค่าความยากง่าย และค่าอำนาจจำแนกจากตัวเลือกที่ถูกแล้ว ยังสามารถประเมินได้จากค่าความยากง่าย และค่าอำนาจ จำแนกของตัวลวงด้วยเพื่อใช้พิจารณาคูณภาพข้อสอบได้ดังนี้

1.3.1 ค่าสัดส่วนของผู้เลือกตัวลวง (P) หมายถึง สัดส่วนของจำนวนคนที่เลือกตัวลวงนั้นๆ จะมีค่าจาก 0 ถึง 1 (ศิริชัย กาญจนวาสี, 2556, น. 223) วิธีการคำนวณเช่นเดียวกับการคำนวณจาก ตัวเลือกที่ถูก แต่ต่างกันตรงความหมายของคูณภาพข้อสอบคือ ความยากของตัวลวงเป็นสัดส่วนที่ แสดงว่าตัวลวงนั้นมีผู้สอบเลือกมากหรือน้อย เกณฑ์ที่ใช้พิจารณาคูณภาพข้อสอบข้อนั้นไม่ใช่ 0.2 - 0.8 แต่เป็น 0.05 ขึ้นไป (เอมอร จังศิริพรปกรณ์, 2550, น. 74)

1.3.2 อำนาจจำแนกของตัวลวง (r) หมายถึง ผลต่างระหว่างสัดส่วนของ คนในกลุ่มอ่อนที่เลือกตัวลวงกับสัดส่วนของคนในกลุ่มเก่งที่เลือกตัวลวงนั้นๆ ค่าอำนาจจำแนกของ ตัวลวงมีค่าตั้งแต่ - 1 ถึง + 1 (ศิริชัย กาญจนวาสี, 2556, น. 224) วิธีการคำนวณเช่นเดียวกับการ คำนวณจากตัวเลือกที่ถูก แต่ต่างกันตรงความหมายของคูณภาพข้อสอบคือ ค่าอำนาจจำแนกของตัว ลวง หมายถึงความสามารถของตัวลวงในการจำแนกบุคคลออกเป็น 2 กลุ่มที่ต่างกัน และต่างกันที่ เครื่องหมาย ถ้าเป็นตัวลวงที่ดีต้องมีผู้สอบกลุ่มต่ำเลือกมากกว่ากลุ่มสูง ดังนั้นค่าอำนาจจำแนกต้องมี ค่าเป็นลบ เกณฑ์ในการพิจารณาไม่ใช่ 0.2 ขึ้นไป แต่เป็น 0.05 ขึ้นไป (เอมอร จังศิริพรปกรณ์, 2550, น. 74)

ในปัจจุบันเทคโนโลยีเจริญรุดหน้าเป็นอันมาก จึงได้มีการพัฒนาโปรแกรมคอมพิวเตอร์ สำหรับการวิเคราะห์ข้อสอบเพื่อให้เกิดความสะดวกและมีความแม่นยำมากขึ้น ช่วยประหยัดเวลาในการวิเคราะห์ ซึ่งในงานวิจัยครั้งนี้ได้รับความร่วมมือจากสำนักส่งเสริมวิชาการและงานทะเบียน มหาวิทยาลัยราชภัฏสวนดุสิตในการวิเคราะห์ข้อสอบด้วยโปรแกรมวิเคราะห์ข้อสอบรายข้อที่พัฒนา โดยบริษัท Application Support & Development Department Control Data (Thailand) Ltd. ด้วยเทคนิคการวิเคราะห์ข้อสอบแบบ 27 เปอร์เซ็นต์ เนื่องจากนักศึกษาผู้เข้าสอบมีจำนวนมากถึง 88 คน และการกระจายของคะแนนสอบเป็นแบบโค้งปกติ วิธีนี้เชื่อว่ากลุ่มขนาดตัวอย่างร้อยละ 27 จากผู้สอบที่ได้คะแนนสูงหรือผู้ที่สอบได้คะแนนต่ำสามารถเป็นตัวแทนประชากรของผู้เข้าสอบทั้งหมด

ในการสอบวัดความรู้เพื่อประเมินผลสัมฤทธิ์ทางการเรียน โดยทั่วไปมักนิยมให้มีข้อสอบที่มีความยากง่ายในระดับต่างๆปะปนกันไป โดยจัดให้มีข้อสอบมีค่าความยากง่ายพอเหมาะ (p มีค่าใกล้เคียง 0.5) เป็นส่วนใหญ่รวมทั้งให้มีข้อสอบที่ค่อนข้างยากและค่อนข้างง่ายอีกจำนวนหนึ่งแต่ถ้าเป็นการสอบแข่งขันเพื่อคัดเลือกผู้ที่มีความรู้ความสามารถควรมีสัดส่วนของข้อสอบที่ยากสูงขึ้นทั้งนี้ ข้อสอบที่ดีควรมีค่าความยากง่ายระหว่าง 0.20 – 0.80 ในข้อสอบประเภท 4 ตัวเลือกส่วนข้อสอบประเภทถูก – ผิดค่าความยากง่ายควรอยู่ระหว่าง 0.60 - 0.95 (Nunnally, 1967 อ้างถึงในเยาวดี ราชชัยกุล วิบูลย์ศรี, 2552, น. 144 -145)

2. การวิเคราะห์ข้อสอบเป็นรายข้อแบบอิงเกณฑ์ (Criterion reference)

ข้อสอบแบบอิงเกณฑ์ เป็นข้อสอบที่สร้างขึ้นตามจุดประสงค์การเรียนรู้ โดยมุ่งเอาผลการวัดของผู้เรียนแต่ละคนไปเปรียบเทียบกับเกณฑ์ที่กำหนดไว้ล่วงหน้า ไม่ได้เปรียบเทียบกับผลการวัดของคนอื่นๆในกลุ่มเดียวกัน (สมนึก ภัททิยธนี, 2555, น. 212) การวิเคราะห์แบบอิงเกณฑ์จึงเกี่ยวข้องกับประสิทธิภาพในการวัดผลของการจัดการเรียนการสอนโดยตรง โดยใช้ข้อสอบชุดเดียวกันทั้งการทดสอบก่อนเรียนและหลังเรียน ข้อสอบอิงเกณฑ์ที่มีประสิทธิภาพควรตั้งอยู่บนสมมติฐานที่ว่า กระบวนการเรียนการสอนจะส่งผลโดยตรงต่อการเรียนรู้ของผู้เรียนเมื่อการเรียนการสอนสิ้นสุดลงควรมีผู้ตอบถูกมากกว่าก่อนเรียน (ปวรส บุตะเขี้ยว, 2556, น. 4)

ดัชนีบ่งชี้คุณภาพของข้อสอบสำหรับแบบทดสอบอิงเกณฑ์

คุณภาพข้อสอบของแบบทดสอบอิงเกณฑ์การพิจารณาได้จากความสอดคล้องระหว่างข้อสอบกับจุดมุ่งหมาย ความยากง่ายของข้อสอบ อำนาจจำแนกหรือดัชนีความไว และประสิทธิภาพของตัวลวง (ศิริชัย กาญจนวาสี, 2556, น. 237 - 238) ดังมีรายละเอียดต่อไปนี้

2.1 ความสอดคล้องระหว่างข้อสอบกับจุดมุ่งหมาย (Item – Objective Congruence: IOC) หมายถึง ลักษณะของข้อสอบที่เป็นตัวแทนความรู้ที่กำหนดไว้เป็นเป้าหมายอย่างชัดเจน หรือเป็นตัวแทนจุดมุ่งหมายของการเรียนรู้ที่กำหนดไว้ วิธีการวัดความสอดคล้องสามารถทำได้โดยอาศัย การตัดสินของผู้ทรงคุณวุฒิไม่ต่ำกว่า 3 คน ประเมินค่าความสอดคล้อง 5 ระดับ จากนั้นจึงคำนวณ ค่าเฉลี่ยของผลการตัดสิน เพื่อเปรียบเทียบกับเกณฑ์สอดคล้องซึ่งส่วนใหญ่ \geq ร้อยละ 80 นั่นคือถ้า ค่าเฉลี่ย (\bar{X}) \geq 4.00 ถือว่าสอดคล้อง

2.2 ความยากง่ายของข้อสอบ (Item Difficulty: P_i)

P_i หมายถึง สัดส่วนของผู้ตอบข้อสอบข้อนั้นๆ ได้ถูกต้อง ความยากของข้อสอบอิงเกณฑ์ พิจารณาจากกลุ่มเป้าหมาย ซึ่งประกอบด้วยกลุ่มที่ได้เรียนเรื่องนั้นมาแล้ว หรือกลุ่มที่รอบรู้เรื่องนั้น แล้ว กับกลุ่มที่ยังไม่ได้เรียนเรื่องนั้นหรือกลุ่มที่ยังไม่รอบรู้เรื่องนั้น ข้อสอบที่ง่ายสำหรับกลุ่มที่เคยเรียน สามารถเป็นข้อสอบอิงเกณฑ์ที่ดีได้ จึงทำให้ผู้เรียนส่วนใหญ่เกิดการเรียนรู้และทำข้อสอบได้หรืออาจ เป็นเพราะข้อสอบง่ายเกินไป จึงทำให้ผู้เรียนตอบถูกจากความรู้เดิมโดยไม่ได้เรียนรู้เพิ่มเติม ดังนั้นการ วิเคราะห์ข้อสอบอิงเกณฑ์จึงจำเป็นต้องตรวจสอบว่าก่อนเรียนหรือกลุ่มที่ยังไม่เคยเรียนมีสัดส่วนของ คำตอบถูกมากน้อยเพียงใดเพื่อนำผลมาเปรียบเทียบกับตอนหลังเรียน P_i มีค่าระหว่าง 0 ถึง 1

การวิเคราะห์ข้อสอบแบบอิงเกณฑ์ จะพิจารณาค่าความยากง่ายของข้อสอบแตกต่างจากแบบ อิงกลุ่ม เนื่องจากไม่ได้เน้นที่จะนำค่าความยากง่ายมาเพื่อเลือกข้อสอบ แต่จะเน้นที่คุณภาพในการ สอน กล่าวคือถ้าอาจารย์ยังไม่ได้สอนเนื้อหา นั้น ข้อสอบควรจะยาก คือมีค่า P_i ต่ำ แต่ถ้าได้สอนแล้ว และสอนอย่างมีประสิทธิภาพส่งผลให้ผู้เรียนเกิดการเรียนรู้ข้อสอบควรจะง่าย คือ มีค่า P_i สูง (ปวรส บุตะเขี้ยว, 2556, น. 4) โดยมีเกณฑ์การประเมินความยากง่ายดังนี้

ตารางที่ 2.3 เกณฑ์การพิจารณาค่าความยากง่ายของข้อสอบแบบอิงเกณฑ์ (ศิริชัย กาญจนวาสี, 2556, น. 243)

กลุ่มผู้เรียน	ความยากง่ายของข้อสอบ (P_i)
1. กลุ่มที่เคยเรียนมาแล้ว	<ul style="list-style-type: none"> • ควรมีค่า P_i สูง (0.70-1.00) • ถ้า $P_i < 0.70$ ควรปรับปรุงหรือตัดทิ้ง
2. กลุ่มที่ยังไม่เคยเรียน	<ul style="list-style-type: none"> • ควรมีค่า P_i ต่ำ (0-0.5) • ถ้า $P_i > 0.5$ แสดงว่าข้อสอบง่ายเกินไป (ควรปรับปรุงหรือตัดทิ้ง)

2.3 อำนาจจำแนกหรือดัชนีความไว (Sensitivity Index: S_i)

S_i หมายถึงผลต่างระหว่างสัดส่วนของผู้ที่ตอบถูกหลังเรียน (กลุ่มที่รอบรู้) กับสัดส่วนผู้ที่ตอบถูกก่อนเรียน (กลุ่มที่ยังไม่รอบรู้)

S_i มีค่าระหว่าง - 1 ถึง 1 และควรมีค่าเป็นบวก ยิ่งเข้าใกล้ 1 ยิ่งมีความไวเท่านั้น แสดงว่าข้อสอบข้อนั้นสามารถจำแนกผู้รอบรู้จากผู้ไม่รอบรู้ได้ดี ดังนั้นอำนาจจำแนกหรือดัชนีความไวของข้อสอบ (S_i) ควรมีค่า > 0 โดยมีเกณฑ์การพิจารณาคุณภาพข้อสอบดังตารางที่ 2.4

ตารางที่ 2.4 เกณฑ์การพิจารณาอำนาจจำแนกหรือดัชนีความไวของข้อสอบแบบอิงเกณฑ์ (โชติกา ภาชีผล, 2554, น. 85)

ดัชนีความไว	ความหมาย
$> .50$	ข้อสอบมีคุณภาพดี ผู้เรียนตอบถูกหลังเรียนมากกว่าก่อนเรียนจำนวนมาก
$.01 - .50$	ข้อสอบมีคุณภาพน้อยถึงปานกลาง คือผู้เรียนตอบถูกหลังเรียนมากกว่าก่อนเรียนจำนวนน้อยถึงปานกลาง
$.00$	ข้อสอบไม่มีคุณภาพ ไม่มีความแตกต่างระหว่างผู้เรียนตอบถูกก่อนและหลังเรียน

ค่า S_i นี้จะแตกต่างจากค่าอำนาจจำแนก (r) ของข้อสอบแบบอิงกลุ่มคือ ค่าอำนาจจำแนก (r) ของข้อสอบแบบอิงกลุ่มเป็นดัชนีที่ชี้ให้เห็นว่าสามารถจำแนกคนที่มีผลสัมฤทธิ์ทางการเรียนสูงและต่ำ

2.4 ประสิทธิภาพของตัวลวง

ในการใช้ข้อสอบอิงเกณฑ์แบบหลายตัวเลือก ผู้ที่เลือกตัวลวงถือว่าตอบผิด ซึ่งแสดงให้เห็นว่าผู้นั้นยังไม่สัมฤทธิ์ผลตามเป้าหมายของการวัด ในการวิเคราะห์ประสิทธิภาพตัวลวงทำได้โดยการตรวจสอบความถี่ของผู้ที่เลือกตัวลวงนั้นๆ และควรดูการเลือกตัวลวงจากผลการสอบก่อนการเรียนการสอน ถ้าตัวลวงใดมีผู้เลือกในสัดส่วนที่สูงถือว่าเป็นตัวลวงที่ใช้ได้ แต่ถ้าตัวลวงใดไม่มีผู้เลือกเลยหรือเลือกในสัดส่วนต่ำ แสดงว่าตัวลวงนั้นขาดประสิทธิภาพ สมควรที่จะต้องมีการปรับปรุงแก้ไขหรือเปลี่ยนแปลงให้ดีขึ้น (ศิริชัย กาญจนวาสี, 2556, น. 238 - 239) และตัวลวงที่จัดว่าเป็นตัวลวงที่ดีนั้นผู้ที่เรียนอ่อนควรจะเลือกตอบมากกว่าผู้ที่มีผลการเรียนดี

การวิเคราะห์คุณภาพข้อสอบทั้งฉบับ

การวิเคราะห์ข้อสอบทั้งฉบับเป็นการตรวจสอบคุณภาพของเครื่องมือวัดผลสัมฤทธิ์ทางการเรียนโดยมีดัชนีชี้วัดคุณภาพที่สำคัญได้แก่ ค่าความตรงและความเที่ยงหรือความเชื่อมั่น ซึ่งมีรายละเอียดและวิธีการตรวจสอบดังนี้

1. ความตรงหรือความเที่ยงตรง (Validity)

ความตรงหรือความเที่ยงตรง เป็นคุณสมบัติของเครื่องมือที่สามารถวัดได้ตามวัตถุประสงค์ที่ต้องการวัด ความตรงที่ใช้ในการทดสอบจำแนกเป็น 3 ประเภทดังนี้ (พิชิต ฤทธิ์จรูญ, 2553, น. 136 - 137; โขติกา ภาชีผล, 2554, น. 60 - 70)

1.1 ความตรงตามเนื้อหา (Content Validity) หมายถึงคุณสมบัติของข้อคำถามที่สามารถวัดได้ตรงตามเนื้อหาและพฤติกรรมที่ต้องการวัดหรือวัตถุประสงค์ของวิชาที่สอน การตรวจสอบความตรงตามเนื้อหาทำได้โดยการวิเคราะห์เนื้อหาเชิงเหตุผลของข้อสอบทั้งฉบับ อาศัยดุลยพินิจของผู้เชี่ยวชาญพิจารณาว่า ข้อคำถามหรือข้อความแต่ละข้อนั้นถามได้ตรง ครอบคลุม ครบถ้วน ตลอดจนจำนวนข้อสอบตรงตามตารางวิเคราะห์หลักสูตร ซึ่งจะจำแนกตามเนื้อหาและพฤติกรรมที่ต้องการวัด โดยกำหนดขอบเขตและองค์ประกอบของเนื้อหาให้ชัดเจน

1.2 ความตรงเชิงโครงสร้าง (Construct Validity) หมายถึง คุณสมบัติของข้อคำถามที่สามารถวัดคุณลักษณะหรือพฤติกรรมตรงตามทฤษฎีหรือแนวคิดของเรื่องนั้น การตรวจสอบความตรงตามโครงสร้างมีหลายวิธี ได้แก่

1.2.1 การหาความสัมพันธ์กับเครื่องมือที่มีโครงสร้างเหมือนกัน เป็นการศึกษาความสัมพันธ์ระหว่างผลการวัดที่ได้จากเครื่องมือที่สร้างขึ้นกับผลของเครื่องมือมาตรฐานที่มีโครงสร้างเหมือนกัน โดยคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่าย (Pearson product moment correlation coefficient) ถ้าอยู่ในเกณฑ์สูง แสดงว่าข้อสอบที่สร้างขึ้นมีความตรงตามโครงสร้าง

1.2.2 การวิเคราะห์องค์ประกอบ (Factor analysis) เป็นการอาศัยวิธีทางสถิติสำหรับจับกลุ่มหรือรวมตัวแปรที่มีความสัมพันธ์ไว้เป็นกลุ่มๆหรือเป็นองค์ประกอบที่สำคัญ แล้ววิเคราะห์ความสัมพันธ์ระหว่างข้อสอบแต่ละข้อว่าข้อสอบทั้งหมดนั้นวัดองค์ประกอบอะไรบ้าง ถ้าองค์ประกอบที่วัดตรงตามทฤษฎีหรือสมมติฐานที่ตั้งไว้ แสดงว่าข้อสอบมีความตรงตามโครงสร้าง

1.2.3 การเทียบกลุ่มรู้ชุด (Known Group) วิธีนี้จะนำข้อสอบที่สร้างขึ้นไปเปรียบเทียบกับกลุ่มที่มีคุณลักษณะเดียวกับสิ่งที่ต้องการวัดแต่มีลักษณะตรงข้ามกัน แล้วนำผลมาเทียบกันด้วยการทดสอบที (t - test) ถ้าค่าที่วัดได้จากกลุ่มทั้งสองมีความแตกต่างกันอย่างมีนัยสำคัญเชิงสถิติ แสดงว่าข้อสอบนั้นมีความตรงตามโครงสร้าง

1.3 ความตรงเชิงเกณฑ์สัมพันธ์ (Criterion-Related Validity) เป็นคุณสมบัติของข้อสอบที่สามารถวัดได้สอดคล้องกับเกณฑ์ภายนอกบางอย่าง แบ่งเป็น 2 ประเภท ดังนี้

1.3.1 ความตรงตามสภาพ (Concurrent Validity) เป็นความสามารถของข้อสอบที่วัดได้ตรงกับสภาพความเป็นจริงที่เกิดขึ้นในปัจจุบันของบุคคลในขณะนั้น เมื่อนำไปให้ผู้เรียนคนหนึ่งสอบ ปรากฏว่าได้คะแนนสูง ซึ่งในสภาพความเป็นจริง ผู้เรียนผู้นั้นมีความสามารถสูงจริง แสดงว่าข้อสอบวัดได้ตรงตามสภาพ

1.3.2 ความตรงเชิงพยากรณ์ (Predictive Validity) เป็นความสามารถของข้อสอบที่วัดผลได้ตรงกับสภาพความเป็นจริงที่เกิดขึ้นในอนาคต เช่น ข้อสอบคัดเลือกบุคคลเข้าศึกษาต่อ ถ้าผู้ที่สอบคัดเลือกได้คะแนนสูงเมื่อเข้าศึกษาต่อปรากฏว่าบุคคลนั้นมีผลการเรียนอยู่ในเกณฑ์ที่ดี แสดงว่าข้อสอบมีความตรงเชิงพยากรณ์

2. ค่าความเชื่อมั่นของข้อสอบ (Reliability)

ค่าความเชื่อมั่น หรือความเที่ยง หมายถึง ความคงที่ของคะแนนที่ได้จากการวัดแต่ละครั้ง ถือเป็นคุณสมบัติที่สำคัญของข้อสอบทุกชนิด แบบสอบที่ดีต้องมีค่าความเชื่อมั่นสูง ซึ่งจะทำให้การวัดมีความคลาดเคลื่อนต่ำ การหาค่าความเที่ยงของข้อสอบมีหลายวิธี ได้แก่

2.1 วิธีการทดสอบซ้ำ

เป็นการหาความสัมพันธ์ของคะแนนในการทำข้อสอบฉบับเดียวกันสองครั้ง เพื่อตรวจสอบความคงที่ในการแสดงออกของผู้สอบ โดยเว้นระยะเวลาให้ห่างกันไม่น้อยกว่า 1 สัปดาห์

2.2 วิธีแบบทดสอบคู่ขนาน

เป็นการหาความสัมพันธ์ของคะแนนในการทำข้อสอบ 2 ชุดในเวลาใกล้เคียงกันโดยข้อสอบ 2 ชุดนี้มีลักษณะเป็นคู่ขนาน วัดในเรื่องเดียวกัน จำนวนข้อเท่ากัน ความยากง่ายเท่ากัน คะแนนเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานเท่ากัน

2.3 วิธีหาความคงที่ภายใน หรือเชิงสอดคล้องภายใน

วิธีนี้ใช้ผู้สอบกลุ่มเดียวกันตอบข้อสอบเพียงครั้งเดียว มีวิธีคำนวณดังนี้

2.3.1 วิธีแบ่งครึ่ง

วิธีนี้แบ่งตรวจข้อสอบเป็น 2 ส่วนที่เท่าเทียมกัน โดยอาจแบ่งเป็นข้อคู่ - ข้อคี่ หรือครึ่งฉบับแรกกับครึ่งฉบับหลัง แล้วนำมาหาสัมประสิทธิ์ความเชื่อมั่นของข้อสอบครึ่งฉบับก่อนแล้วจึงนำไปปรับขยายเป็นสัมประสิทธิ์สหสัมพันธ์ของข้อสอบทั้งฉบับ ถ้าค่าความเชื่อมั่นทั้งฉบับมีค่าสูงหรือเข้าใกล้ 1.0 แสดงว่าข้อสอบทั้ง 2 ส่วนมีความสอดคล้องภายใน

2.3.2 วิธีของคูเดอร์และริชาร์ดสัน (Kuder – Richardson: KR)

เป็นวิธีที่นิยมมาก ใช้ได้กับข้อสอบที่ให้คะแนน 0 - 1 เช่น ข้อสอบแบบหลายตัวเลือกโดยมีข้อตกลงว่า ข้อสอบฉบับนั้นต้องมีวัดลักษณะเดียวหรือวัดองค์ประกอบร่วมกัน แล้วนำคะแนนมาวิเคราะห์โดยใช้สูตรของคูเดอร์และริชาร์ดสัน ซึ่งมี 2 สูตร คือ KR - 20 กับ KR - 21

1) สูตร KR - 20 การคำนวณสูตรนี้ต้องทราบความยากง่าย (p) ของข้อสอบแต่ละข้อ และควรมีข้อสอบอย่างน้อย 20 ข้อ โดยข้อสอบแต่ละข้อจะต้องวัดคุณลักษณะเดียวกัน และมีค่าความยากง่ายใกล้เคียงกัน

$$\text{สูตร KR - 20} = \frac{k}{k-1} \left[1 - \frac{\sum pq}{S_x^2} \right]$$

เมื่อ k แทน จำนวนข้อของแบบทดสอบ
 S_x^2 แทน ความแปรปรวนของคะแนนรวม
 p แทน สัดส่วนของผู้ตอบข้อสอบถูกในแต่ละข้อ
 q แทน สัดส่วนของผู้ตอบข้อสอบผิดในแต่ละข้อ หรือมีค่าเป็น 1 - p

2) สูตร KR - 21 การคำนวณด้วยสูตรนี้ต้องทราบค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ยและค่าความแปรปรวน

$$\text{สูตร KR - 21} = \frac{k}{k-1} \left[1 - \frac{\bar{x}(k - \bar{x})}{k S_x^2} \right]$$

เมื่อ k แทน จำนวนของแบบทดสอบ
 \bar{x} แทน ค่าเฉลี่ยของคะแนนรวม
 S_x^2 แทน ความแปรปรวนของคะแนนรวม

สูตร KR - 20 และ KR - 21 นี้ใช้ได้เฉพาะการหาค่าความเชื่อมั่นของข้อสอบที่ให้คะแนนแต่ละข้อเป็นแบบ 0 กับ 1 เท่านั้น ค่าความเชื่อมั่นที่ได้จากสูตร KR - 21 จะต่ำกว่าจากสูตร KR - 20 เพราะสูตร KR - 21 ใช้ค่าเฉลี่ยของข้อสอบทั้งฉบับแทนค่า p, q ของแต่ละข้อ โดยถือว่าข้อสอบทุกข้อมีระดับความยากง่ายเท่ากัน หรือกำหนดให้ค่า p คงที่ ในทางปฏิบัติแล้วข้อสอบมีความยากง่ายไม่เท่ากัน ทำให้มีความคลาดเคลื่อนมากกว่าสูตร KR - 20 แต่สูตร KR - 21 มีขั้นตอนการคำนวณที่น้อยกว่า สะดวกและรวดเร็วกว่าการใช้สูตร KR - 20

2.3.3 วิธีหาจากสูตรสัมประสิทธิ์แอลฟา

การหาค่าความเชื่อมั่นโดยใช้สูตรของครอนบาค (Cronbach) ปรับมาจากสูตร KR - 20 และเป็นวิธีที่ใช้อย่างแพร่หลาย เนื่องจากทดสอบเพียงครั้งเดียว ใช้ได้ทั้งข้อสอบที่ให้คะแนนแบบเรียงอันดับ แบบมาตราส่วนประมาณค่าหรือให้คะแนนแบบ 0,1 รวมทั้งแบบอัตนัย สูตรการคำนวณเป็นดังนี้

$$\alpha = \frac{k}{k-1} \left[1 - \frac{\sum S_i^2}{S_x^2} \right]$$

เมื่อ	α	แทน ค่าความเที่ยงหรือค่าความเชื่อมั่น
	k	แทน จำนวนข้อของแบบทดสอบ
	S_i^2	แทน ผลรวมความแปรปรวนของคะแนนรายข้อ
	S_x^2	แทน ความแปรปรวนของคะแนนรวม

การแปลความหมายของค่าสัมประสิทธิ์ความเชื่อมั่น

ข้อสอบใช้ในการทดสอบควรมีค่าสัมประสิทธิ์ความเที่ยงหรือความเชื่อมั่นสูงที่สุดเท่าที่จะเป็นไปได้ในสภาพการณ์นั้น อย่างน้อยที่สุดควรมีค่าไม่ต่ำกว่า 0.50 (ศิริชัย กาญจนวาสี, 2556, น. 97) ทั้งนี้ Office of Educational Assessment, University of Washington (2005) ได้เสนอแนวทางในการแปลความหมายของค่าสัมประสิทธิ์ความเชื่อมั่นดังตารางที่ 2.5

ตารางที่ 2.5 เกณฑ์การแปลความหมายของค่าสัมประสิทธิ์ความเชื่อมั่น

ค่าสัมประสิทธิ์ความเชื่อมั่น	การแปลความหมาย
$\geq .90$	ยอดเยี่ยม เทียบได้กับข้อสอบมาตรฐาน
.80 - .90	ดีมากเหมาะสำหรับใช้เป็นแบบทดสอบรายวิชาในชั้นเรียน
.70 - .80	ดีเหมาะสำหรับเป็นแบบทดสอบในชั้นเรียนอาจมีข้อสอบบางข้อที่ต้องปรับปรุง
.60 - .70	ค่อนข้างต่ำควรวัดด้วยเครื่องมืออื่นเสริมก่อนตัดเกรด และมีข้อสอบบางข้อที่ต้องปรับปรุงคุณภาพ
.50 - .60	ควรปรับปรุง จำเป็นต้องมีการวัดด้วยเครื่องมืออื่นเสริมก่อนตัดเกรด
$\leq .50$	ไม่น่าเชื่อถือ จำเป็นต้องปรับปรุง ไม่ควรใช้เป็นหลักในการตัดเกรด

ที่มา : Office of Educational Assessment, University of Washington (2005)

ปัจจัยที่มีผลต่อค่าความเชื่อมั่นหรือความเที่ยงของข้อสอบ

ค่าความเชื่อมั่นหรือความเที่ยงที่คำนวณได้จากข้อสอบที่มีความเกี่ยวข้องทั้งลักษณะของข้อสอบ กลุ่มผู้ตอบหรือผู้เข้าสอบ และปัจจัยอื่นๆดังนี้

1. **ความเป็นเอกพันธ์หรือมีคุณลักษณะที่ต้องการวัดคล้ายคลึงกันของกลุ่มผู้สอบ** เช่น ระดับความรู้ความสามารถเท่าเทียมกัน ทำให้การกระจายของคะแนนมีน้อย ซึ่งจะส่งผลให้ความแปรปรวนของคะแนนต่ำ ทำให้ข้อสอบมีค่าอำนาจจำแนกต่ำ ไม่สามารถจำแนกบุคคลออกเป็นกลุ่มตามจุดมุ่งหมายของแบบวัดได้ชัดเจน ค่าความเชื่อมั่นของข้อสอบจึงต่ำ

2. **จำนวนข้อสอบ** ถ้าข้อสอบมีข้อคำถามมากจะวัดได้ครอบคลุมเนื้อหาและพฤติกรรมที่ต้องการวัดจะทำให้ค่าความเชื่อมั่นสูงขึ้น

3. **เวลาที่ใช้ในการทำข้อสอบ** ถ้ามีการกำหนดเวลาที่เหมาะสม ไม่มากหรือน้อยเกินไป ผู้สอบมีเวลาทำข้อสอบได้ครบทุกข้อ ค่าความเชื่อมั่นของข้อสอบจะสูงขึ้น

4. **ความสัมพันธ์ระหว่างข้อสอบ** ถ้าข้อสอบมีความสัมพันธ์กันสูงหรือวัดคุณลักษณะเดียวกัน จะส่งผลให้มีค่าความเชื่อมั่นสูง เพราะผู้สอบตอบคำถามในทิศทางเดียวกันมีความสอดคล้องภายใน

5. **ความยากง่ายของข้อสอบ** ถ้าข้อสอบยากเกินไปหรือง่ายเกินไป หรือมีค่าอำนาจจำแนกต่ำ จะไม่สามารถจำแนกบุคคลออกเป็นกลุ่มสูงหรือต่ำได้ ส่งผลให้ความแปรปรวนของคะแนนต่ำและค่าความเชื่อมั่นต่ำด้วย ในทำนองเดียวกันหากข้อสอบมีความยากง่ายปานกลางหรือมีค่าอำนาจจำแนกสูง จะส่งผลให้ความแปรปรวนของคะแนนสูงและค่าความเชื่อมั่นสูงด้วยเช่นกัน

6. **ความเป็นปรนัยของข้อสอบ** ลักษณะของข้อสอบที่มีความเป็นปรนัยสูงได้แก่ ข้อสอบที่มีความชัดเจนทั้งตัวคำถาม คำตอบและคำอธิบายหรือคำชี้แจงในการตอบใช้ภาษาไม่กำกวม เป็นไปตามจุดมุ่งหมายของการสอบ การตรวจให้คะแนนและการแปลความหมายคะแนนมีเกณฑ์ชัดเจน เป็นมาตรฐานเดียวกัน ทำให้เกิดความคลาดเคลื่อนจากการวัดน้อย คะแนนที่ได้จะมีค่าใกล้เคียงกับคะแนนจริง ส่งผลให้ข้อสอบนั้นมีค่าความเชื่อมั่นสูง

7. **วิธีวิเคราะห์ค่าความเชื่อมั่น** ข้อสอบชุดเดียวกันถ้าวิเคราะห์ค่าความเชื่อมั่นด้วยสูตรสัมประสิทธิ์แอลฟาจะเท่ากับสูตร KR - 20 แต่ไม่เท่ากับการวิเคราะห์โดยวิธีแบ่งครึ่ง ค่าความเชื่อมั่นของข้อสอบที่มีค่าความยากง่ายต่างกันถ้าคำนวณโดยใช้สูตร KR - 20 จะสูงกว่าใช้สูตร KR - 21

นอกจากนี้พรสันต์ เลิศวิทยาวีวัฒน์ (2556, น. 7) ให้ความคิดเห็นในบทความวิชาการเรื่องกระบวนการสร้างข้อสอบมาตรฐาน ในส่วนที่เกี่ยวกับเทคนิคในการปรับค่าความเชื่อมั่นของข้อสอบให้สูงขึ้นว่า เมื่อวิเคราะห์ข้อสอบทั้งฉบับแล้วพบว่าค่าความเชื่อมั่นต่ำกว่าที่ต้องการและผู้สร้างข้อสอบต้องการปรับให้ค่าความเชื่อมั่นของข้อสอบให้สูงขึ้น สามารถทำได้ด้วยการตัดข้อสอบที่มีค่าอำนาจจำแนกน้อยมากหรือติดลบออก ค่าความเชื่อมั่นของข้อสอบทั้งฉบับก็จะสูงขึ้น

งานวิจัยที่เกี่ยวข้องกับคุณภาพของข้อสอบ

พิมพ์อุษา เจริญยศ (2539) ศึกษาเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบเลือกตอบเมื่อลดจำนวนข้อสอบโดยใช้เกณฑ์ค่าสถิติกับใช้ดุลยพินิจของครู พบว่า 1) แบบทดสอบที่มีจำนวน 40 ข้อ มีค่าความเชื่อมั่นโดยใช้เกณฑ์ค่าสถิติมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 กับค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลงเหลือ 35 และ 15 ข้อ 2) ค่าความเชื่อมั่นของแบบทดสอบที่ใช้ดุลยพินิจของครูลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 เมื่อนำไปทดสอบความแตกต่างเป็นรายคู่พบว่า แตกต่างกันเป็นส่วนใหญ่ ยกเว้นค่าความเชื่อมั่นของแบบทดสอบฉบับที่มีจำนวน 40 ข้อกับฉบับที่ลดจำนวนข้อเหลือ 25 ข้อเพียงฉบับเดียวที่มีค่าความเชื่อมั่นไม่แตกต่างกัน 3) ค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติและดุลยพินิจของครูแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ 0.05 ทุกคู่ที่มีจำนวนข้อเท่ากัน

ศักดิ์สิทธิ์ ฤทธิลัน (2541 อ้างถึงใน เอมอร จังศิริพรปรกรณ์, 2550, น. 79 - 80) ศึกษาเปรียบเทียบค่าความยากง่าย ค่าอำนาจจำแนกและค่าความเชื่อมั่นของแบบทดสอบวิชาคณิตศาสตร์ที่สร้างตัวลวงโดยวิธีต่างกัน 4 วิธีคือ สร้างตัวลวงโดยวิธีการใช้คำตอบผิดของนักเรียนที่ซ้ากันมากที่สุด สร้างตัวลวงจากการวิเคราะห์หาคุณภาพจากแบบทดสอบเพื่อสำรวจ สร้างตัวลวงโดยวิธีสุ่มคำตอบผิดของนักเรียน และใช้ตัวลวงจากครูผู้สอน พบว่าแบบทดสอบที่สร้างตัวลวงโดยวิธีต่างกันทั้ง 4 วิธีมีค่าความยากง่ายเฉลี่ย .469, .476, .548 และ .508 ตามลำดับ และมีค่าอำนาจจำแนกเฉลี่ย 0.820, 0.927, 0.860 และ 0.887 ตามลำดับ แสดงว่าแบบทดสอบสร้างตัวลวงโดยวิธีที่ต่างกันทั้ง 4 วิธีมีค่าความยากง่ายเฉลี่ย ค่าอำนาจจำแนกเฉลี่ยและค่าความเที่ยงของแบบทดสอบไม่แตกต่างกัน

Ascalon, Meyers, Davis & Smits (2007, pp. 153 - 170) ศึกษาผลของความคล้ายคลึงของตัวลวงและโครงสร้างของข้อคำถามต่อความยากง่ายของข้อสอบพบว่า ตัวลวงที่มีความสั้นยาวของประโยคและเนื้อหาคล้ายคลึงกับตัวเลือกหรือคำตอบที่ถูกต้อง จะทำให้ข้อสอบข้อนั้นมีความยากมากกว่าข้อสอบที่มีตัวลวงที่มีเนื้อหาของคำตอบแตกต่างกับคำตอบที่ถูกต้องอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ($p < 0.01$) เพราะตัวลวงที่มีเนื้อหาแตกต่างกันจะง่ายต่อการคาดเดา ส่วนความยากง่ายของข้อสอบที่มีการตั้งโจทย์ข้อคำถามปลายปิด ไม่พบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติกับโจทย์ที่มีลักษณะเป็นคำถามปลายเปิด

Caldwell, D. J. & Pate, A. N. (2013, pp. 1 - 5) ได้ศึกษารูปแบบของคำถามต่อคะแนนการสอบของนักศึกษาและคุณภาพของข้อสอบ โดยมีวัตถุประสงค์เพื่อเปรียบเทียบการใช้รูปแบบคำถามของข้อสอบที่สร้างตามแนวปฏิบัติที่เป็นมาตรฐานและไม่มีมาตรฐาน ข้อสอบที่สร้างตามแนวปฏิบัติที่เป็นมาตรฐานจะมีเกณฑ์พิจารณา ดังนี้ 1) ใช้คำถามเชิงบวก และหลีกเลี่ยงคำถามเชิงลบ เช่น การตั้งคำถามด้วยคำว่า “ยกเว้น” “ไม่ใช่” หากจำเป็นต้องใช้ให้ขีดเส้นใต้ หรือใช้ตัวหนา เพื่อให้เห็นคำที่เป็นเชิงลบอย่างชัดเจน 2) ในแต่ละคำถามสามารถใช้ตัวเลือกได้มากกว่าที่ต้องการ แต่ผลการวิจัยพบว่าการใช้ตัวเลือกเพียง 3 ตัวเลือกก็เพียงพอแล้ว 3) ตัวเลือก “ไม่มีข้อใดถูกต้อง” ควรใช้ด้วยความระมัดระวัง ผลการวิจัยพบว่า ข้อสอบที่ไม่ได้สร้างตามแนวปฏิบัติที่เป็นมาตรฐานมีระดับความยากกว่าข้อสอบที่สร้างตามแนวปฏิบัติที่เป็นมาตรฐาน (p value = 0.03) ตามแนวปฏิบัติทั้ง 3 ข้อ มีเพียงข้อแนะนำที่ให้หลีกเลี่ยงการใช้ตัวลวง “ไม่มีข้อใดถูกต้อง” เท่านั้น ที่ทำให้คะแนนของนักศึกษาทั้งสองกลุ่มแตกต่างกันอย่างมีนัยสำคัญทางสถิติ (p value < 0.001) โดยคะแนนเฉลี่ยของกลุ่มที่ทำข้อสอบสร้างตามแนวปฏิบัติที่เป็นมาตรฐานคิดเป็นร้อยละ 53.6 ส่วนกลุ่มที่ทำข้อสอบที่ไม่ได้สร้างตามแนวปฏิบัติที่เป็นมาตรฐานมีคะแนนเฉลี่ยคิดเป็นร้อยละ 41.3

จากผลการวิจัยสรุปได้ว่า ข้อสอบที่ไม่ได้สร้างตามแนวปฏิบัติที่เป็นมาตรฐานจะทำให้ยากต่อการตอบของนักศึกษามากกว่าข้อสอบที่สร้างตามแนวปฏิบัติที่เป็นมาตรฐาน จึงไม่ช่วยในการแยกแยะกลุ่มที่เก่งกับกลุ่มที่อ่อนจากกันได้

Oluseyi & Olufemi (2012, pp. 237 - 246) ได้ศึกษาวิเคราะห์ข้อสอบชนิดหลายตัวเลือกวิชาเคมีพื้นฐานของมหาวิทยาลัยไนจีเรียพบว่า ค่าความยากง่ายของข้อสอบทั้งฉบับเท่ากับ 0.54 ค่าอำนาจจำแนกเท่ากับ 0.32 ตัวเลือกที่นักศึกษาเลือกตอบมากที่สุดคือตัวเลือกที่ 4 ซึ่งพบถึง 16 ข้อ จากจำนวนข้อสอบทั้งสิ้น 40 ข้อ พบข้อสอบที่มีความผิดพลาดในการวัด 3 ข้อ เนื่องจากอำนาจจำแนกเท่ากับ 0 ซึ่งผู้วิจัยอธิบายว่า อาจเกิดจากการสร้างข้อสอบไม่ถูกต้อง หรือเฉลยข้อสอบผิด นอกจากนี้พบด้วยว่า ข้อสอบที่ยากที่สุดมีค่าความยากง่ายเท่ากับ 0.24 และมีค่าอำนาจจำแนกเท่ากับ 0.32 ซึ่งมีนักศึกษาตอบได้ถูกต้องจำนวน 76 คน จาก 442 คน การที่ค่าความยากง่ายของข้อสอบเท่ากับ 0.54 แสดงว่าข้อสอบไม่ยากหรือง่ายจนเกินไป ค่าอำนาจจำแนกเท่ากับ 0.32 แสดงว่าสามารถจำแนกได้พอใช้ ผู้วิจัยให้ข้อเสนอแนะว่า การสร้างข้อสอบมีความสำคัญอย่างยิ่ง และควรทำตารางแสดงผลการวิเคราะห์ข้อสอบไว้ท้ายข้อ เพื่อเลือกข้อที่ดีไว้ในคลังข้อสอบ และควรมีการวิพากษ์ข้อสอบก่อนนำมาใช้ในการทดสอบด้วย

Tasdemir, M. (2010, pp. 258 - 266) ได้ศึกษาเปรียบเทียบข้อสอบแบบหลายตัวเลือกและแบบถูก - ผิด ซึ่งใช้ในการประเมินผลสัมฤทธิ์ทางการเรียนของนักศึกษาในระดับปริญญาตรี พบว่า 1) คะแนนเฉลี่ยของนักศึกษาเพศชายและหญิงที่ได้จากการทดสอบด้วยข้อสอบแบบหลายตัวเลือกและแบบถูก - ผิดไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ 2) คะแนนเฉลี่ยที่ได้จากการทดสอบแบบหลายตัวเลือกและแบบถูกผิดที่มีโครงสร้างเนื้อหาแบบคู่ขนานกัน จำนวน 28 หัวข้อมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ กล่าวคือมีข้อสอบ 13 ข้อที่ผู้เข้าสอบทำคะแนนจากแบบทดสอบแบบหลายตัวเลือกได้สูงกว่าแบบถูก - ผิด ส่วนอีก 15 ข้อที่ผู้เข้าสอบทำคะแนนจากแบบทดสอบแบบถูก - ผิดได้สูงกว่า ในขณะที่อีก 22 ข้อไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ 3) ความยากง่ายของแบบทดสอบแบบหลายตัวเลือกและแบบถูก - ผิด ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ โดยพบว่าข้อสอบแบบหลายตัวเลือกมีข้อสอบที่ค่อนข้างยาก 10 ข้อ ส่วนอีก 40 ข้อค่อนข้างง่าย ส่วนข้อสอบแบบถูก - ผิดมี 5 ข้อที่ค่อนข้างยากและอีก 45 ข้อค่อนข้างง่าย 4) อำนาจการจำแนกของแบบทดสอบแบบหลายตัวเลือกและแบบถูก - ผิด ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ โดยข้อสอบแบบหลายตัวเลือกมีค่าอำนาจการจำแนกต่ำ มี 10 ข้อที่มีค่าอำนาจการจำแนกอยู่ในเกณฑ์ปกติหรือใช้ได้ 40 ข้อ ส่วนข้อสอบแบบถูก - ผิดมีค่าอำนาจการจำแนกต่ำ 14 ข้อที่มีค่าอำนาจการจำแนกอยู่ในเกณฑ์ปกติใช้ได้ 36 ข้อ

ผลการวิจัยสรุปว่าข้อสอบแบบหลายตัวเลือกและข้อสอบแบบถูก - ผิดที่มีโครงสร้างเนื้อหาแบบคู่ขนานกันสามารถใช้ในการวัดผลสัมฤทธิ์ทางการเรียนได้ไม่แตกต่างกัน

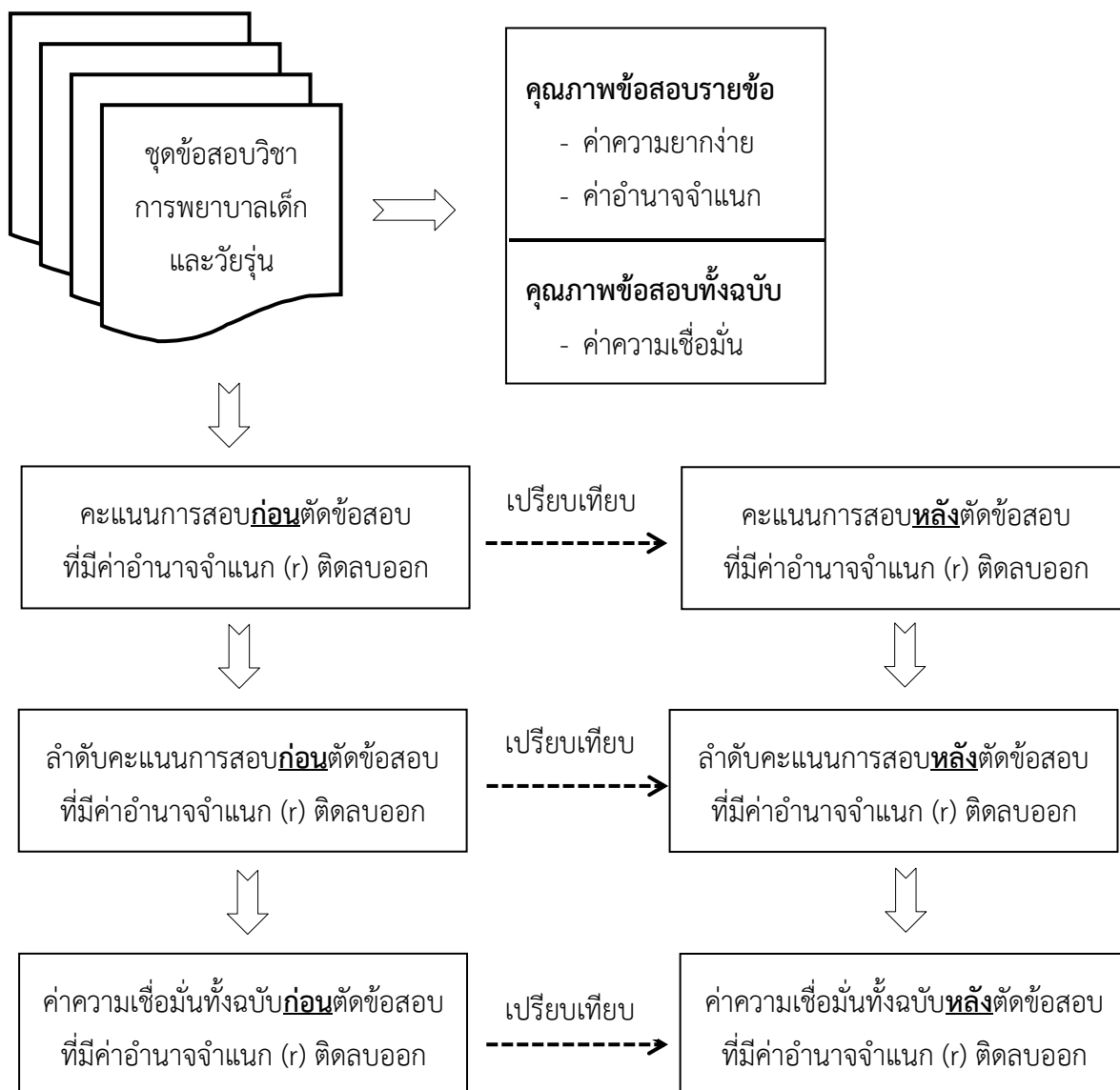
Usova, G M. (1997, pp. 100 - 110) ได้ศึกษาประสิทธิผลของการใช้ข้อสอบเก่าในคลังข้อสอบร่วมกับข้อสอบใหม่ในการจำแนกผู้สอบ เพื่อขอรับใบอนุญาตเป็นผู้ควบคุมโรงไฟฟ้าพลังงานนิวเคลียร์ทั่วประเทศ ผู้เข้าสอบต้องทำคะแนนอย่างน้อยร้อยละ 80 จึงจะสอบผ่านและเข้าสอบภาคปฏิบัติต่อไปการคัดเลือกข้อสอบจาก 3 แบบคือ 1) Test bank items เป็นข้อสอบที่เคยถูกนำมาใช้สอบมาก่อนและผ่านการพิจารณาความตรง และอำนาจจำแนกแล้ว 2) Modified items เป็นข้อสอบที่ปรับปรุงจากข้อสอบเดิม 3) New items เป็นข้อสอบใหม่ตามสัดส่วนดังนี้ 50: 40: 10 พบว่าข้อสอบที่ได้จากการรวบรวมข้อสอบ 3 แบบเข้าด้วยกันในสัดส่วนที่เหมาะสมสามารถจำแนกหรือแยกแยะผู้สอบได้ดี การใช้ข้อสอบเก่าจากคลังข้อสอบจะเปิดโอกาสให้ผู้เข้าสอบได้ฝึกฝนทบทวนตัวเอง เป็นการส่งเสริมการเรียนรู้ แต่ทำให้ผู้เข้าสอบอาจคาดเดาได้หากเคยเห็นข้อสอบข้อนั้นมาก่อน แต่การใช้ข้อสอบที่ปรับปรุงจากข้อสอบเดิมและข้อสอบใหม่ร่วมกับข้อสอบเดิม จะช่วยเพิ่มประสิทธิภาพในการจำแนกผู้สอบได้

สรุป

จากการทบทวนวรรณกรรมข้างต้น จะเห็นว่าข้อสอบที่สร้างขึ้นนั้นจะมีคุณภาพเพียงใดขึ้นกับการวางแผนการสร้างข้อสอบ ความลึกซึ้งในเนื้อหาวิชาของผู้เขียนข้อสอบ และทักษะการเขียนข้อสอบดัชนีบ่งชี้คุณภาพข้อสอบในการทดสอบแบบอิงกลุ่มด้วยข้อสอบแบบเลือกตอบเป็นรายข้อที่สำคัญคือระดับความยากง่ายและอำนาจจำแนกของข้อสอบ ส่วนดัชนีบ่งชี้คุณภาพข้อสอบทั้งฉบับคือความตรงและความเชื่อมั่นหรือความเที่ยง และจากการทบทวนงานวิจัยเกี่ยวกับคุณภาพข้อสอบพบว่า ตัวลงที่มีความสั้นยาวของประโยคและเนื้อหาคล้ายคลึงกับตัวเลือกหรือคำตอบที่ถูกต้อง จะทำให้ข้อสอบข้อนั้นมีควมยากมากกว่าข้อสอบที่มีตัวลงที่มีเนื้อหาของคำตอบแตกต่างกับคำตอบที่ถูกต้อง ข้อสอบที่ไม่ได้สร้างตามแนวปฏิบัติที่เป็นมาตรฐานมีระดับความยากกว่าข้อสอบที่สร้างตามแนวปฏิบัติที่เป็นมาตรฐาน การใช้ข้อสอบเก่าจากคลังข้อสอบจะเปิดโอกาสให้ผู้เข้าสอบได้ฝึกฝนทบทวนตัวเอง เป็นการส่งเสริมการเรียนรู้ แต่ทำให้ผู้เข้าสอบอาจคาดเดาได้หากเคยเห็นข้อสอบข้อนั้นมาก่อน แต่การใช้ข้อสอบที่ปรับปรุงจากข้อสอบเดิมและข้อสอบใหม่ร่วมกับข้อสอบเดิมจะช่วยเพิ่มประสิทธิภาพในการจำแนกผู้สอบได้ ส่วนปัจจัยที่มีผลต่อความเชื่อมั่นของข้อสอบ คือความเป็นเอกพันธ์ของกลุ่มผู้สอบ จำนวนข้อสอบเวลาที่ใช้ทำข้อสอบ ความสัมพันธ์ระหว่างข้อสอบ ความยากง่ายและความเป็นปรนัยของแบบทดสอบ วิธีวิเคราะห์ค่าความเชื่อมั่น ทั้งนี้หากมีการตัดข้อสอบข้อที่มีค่าอำนาจจำแนกน้อยหรือติดลบออก ค่าความเชื่อมั่นของข้อสอบทั้งฉบับจะสูงขึ้น

อย่างไรก็ตามผู้วิจัยยังไม่พบการวิจัยเกี่ยวกับการประเมินผลผู้เรียนเมื่อใช้แบบทดสอบที่มีค่าอำนาจจำแนกติดลบ ผู้วิจัยจึงสนใจศึกษาผลของการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบต่อคะแนนการสอบลำดับคะแนนของนักศึกษาและค่าความเชื่อมั่นของข้อสอบทั้งฉบับในรายวิชาการพยาบาลเด็กและวัยรุ่น

กรอบแนวคิดในการวิจัย



กรอบแนวคิดในการทำวิจัย แสดงตัวแปรที่ใช้ในการศึกษา คือคุณภาพของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น จำนวน 4 ฉบับ ประกอบด้วยคุณภาพของข้อสอบรายข้อและคุณภาพของข้อสอบทั้งฉบับ โดยใช้ค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) เป็นดัชนีชี้วัดคุณภาพของข้อสอบรายข้อ ใช้ค่าความเชื่อมั่นเป็นดัชนีชี้วัดคุณภาพของข้อสอบทั้งฉบับ นอกจากนี้ยังมีการศึกษาการเปลี่ยนแปลงของคะแนนการสอบรายวิชา ลำดับคะแนนการสอบและค่าความเชื่อมั่นของข้อสอบทั้งฉบับหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกตีลบออก

บทที่ 3 วิธีดำเนินการวิจัย

การวิจัยเรื่อง ผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบ และความเชื่อมั่นของข้อสอบทั้งฉบับ: รายวิชาการพยาบาลเด็กและวัยรุ่นนี้เป็นการศึกษาย้อนหลังจากชุดข้อสอบและคะแนนการสอบรายวิชาการพยาบาลเด็กและวัยรุ่น ซึ่งเป็นส่วนหนึ่งของการศึกษาภายใต้งานวิจัยกรณีศึกษาคณะพยาบาลศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิตจำนวน 11 รายวิชาที่เปิดสอนในหลักสูตรพยาบาลศาสตรบัณฑิต ภาคการเรียนที่ 1 ปีการศึกษา 2555 มีวัตถุประสงค์เพื่อศึกษาคุณภาพของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่น การเปลี่ยนแปลงคะแนนการสอบ และลำดับคะแนนการสอบของนักศึกษาในรายวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีอำนาจจำแนกติดลบออก การเปลี่ยนแปลงค่าความเชื่อมั่นของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นทั้งฉบับหลังการตัดข้อสอบที่มีอำนาจจำแนกติดลบออก โดยมีวิธีการดำเนินการวิจัยดังนี้

ประชากรและการเลือกกลุ่มตัวอย่าง

ประชากรคือ ข้อสอบและคะแนนการสอบของนักศึกษาหลักสูตรพยาบาลศาสตรบัณฑิตชั้นปีที่ 2 ที่ลงทะเบียนเรียนและเข้าสอบวิชาการพยาบาลเด็กและวัยรุ่นในภาคการเรียนที่ 1 ปีการศึกษา 2555 จำนวน 88 คน ชุดข้อสอบที่ใช้ในการทดสอบเป็นแบบเลือกตอบ 4 ตัวเลือก มีจำนวน 4 ฉบับ

กลุ่มตัวอย่างในการวิจัยครั้งนี้เป็นการศึกษาจากประชากรทั้งหมดที่เข้าถึงได้ซึ่งเป็นชุดข้อสอบและคะแนนการสอบวิชาการพยาบาลเด็กและวัยรุ่นในภาคการเรียนที่ 1 ปีการศึกษา 2555 จำนวนผู้ลงทะเบียนและเข้าสอบ 88 คน การสอบทั้งสิ้น 4 ครั้งได้แก่ การสอบครั้งที่ 1 จำนวน 75 ข้อ การสอบครั้งที่ 2 จำนวน 70 ข้อ การสอบครั้งที่ 3 จำนวน 58 ข้อและสอบปลายภาคจำนวน 75 ข้อ

เครื่องมือในการวิจัยและการตรวจสอบคุณภาพเครื่องมือ

เครื่องมือที่ใช้ในการวิจัยประกอบด้วยแบบบันทึกข้อมูลที่ผู้วิจัยใช้สำหรับเก็บข้อมูลจากการสอบแต่ละครั้งในรายวิชาการพยาบาลเด็กและวัยรุ่น และโปรแกรมการวิเคราะห์ข้อสอบรายข้อที่พัฒนาโดยบริษัท Application Support & Development Department Control Data (Thailand) Ltd. อาศัยหลักการวิเคราะห์ข้อสอบโดยใช้ทฤษฎีการทดสอบแบบดั้งเดิมหรือประเพณีนิยม (Classical Test Theory) เพื่อศึกษาคุณภาพข้อสอบทั้งแบบรายข้อ และแบบทั้งฉบับ ซึ่งเป็นทฤษฎีที่เหมาะสมกับข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นที่เป็นแบบทดสอบแบบเลือกตอบและ

เป็นการทดสอบแบบอิงกลุ่ม ในการวิจัยนี้ใช้เทคนิคการวิเคราะห์ข้อสอบแบบ 27 เปอร์เซ็นต์เนื่องจาก นักศึกษาผู้เข้าสอบมีจำนวนมาก มีการกระจายของคะแนนสอบเป็นแบบโค้งปกติ วิธีนี้เชื่อว่ากลุ่ม ขนาดตัวอย่าง 27 เปอร์เซ็นต์ จากผู้สอบที่ได้คะแนนสูงหรือผู้ที่สอบได้คะแนนต่ำสามารถเป็นตัวแทน ประชากรของผู้ที่เข้าสอบทั้งหมด โดยได้รับความร่วมมือจากสำนักส่งเสริมวิชาการและงานทะเบียน มหาวิทยาลัยราชภัฏสวนดุสิตในการวิเคราะห์ข้อสอบ

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลสำหรับการวิจัยครั้งนี้ ผู้วิจัยได้ดำเนินการโดยมีขั้นตอนตามลำดับ ดังนี้

1. วางแผนการจัดการเรียนการสอนด้วยการประชุมคณะผู้สอนจำนวน 5 คน จัดทำรายละเอียด วิชาการพยาบาลเด็กและวัยรุ่น (มคอ.3) ให้สอดคล้องตามที่หลักสูตรกำหนด กำหนดตารางวิเคราะห์ หลักสูตร เกณฑ์การวัดและประเมินผล

2. ดำเนินการจัดการเรียนการสอนตามที่กำหนดในรายละเอียดรายวิชา (มคอ.3) อาจารย์ ผู้สอนแต่ละหน่วยการเรียนออกข้อสอบตามตารางวิเคราะห์หลักสูตร (Test blueprint) ที่มีผู้สอน และคณะกรรมการบริหารหลักสูตรของคณะร่วมกันวิพากษ์ข้อสอบ ผู้สอนนำข้อสอบที่ผ่านการ วิพากษ์ไปปรับปรุงและจัดทำเป็นข้อสอบที่สมบูรณ์แล้วจึงดำเนินการสอบครั้งที่ 1 ครั้งที่ 2 ครั้งที่ 3 และสอบปลายภาคตามระเบียบการวัดและประเมินผลของมหาวิทยาลัย

3. ส่งข้อสอบตรวจและวิเคราะห์ที่สำนักส่งเสริมวิชาการและงานทะเบียนของมหาวิทยาลัย ราชภัฏสวนดุสิต เพื่อหาดัชนีคุณภาพของข้อสอบรายข้อ ได้แก่ ความยากง่าย และค่าอำนาจจำแนก และดัชนีคุณภาพของข้อสอบทั้งฉบับ ได้แก่ ค่าความเชื่อมั่นของข้อสอบ

4. นำผลการวิเคราะห์ข้อสอบมาพิจารณาและตัดข้อสอบที่มีอำนาจจำแนกติดลบออก จากนั้นจึงส่งข้อสอบตรวจให้คะแนนพร้อมทั้งวิเคราะห์ข้อสอบใหม่

5. รวบรวมผลการวิเคราะห์ข้อสอบใหม่ที่ตัดข้อสอบที่มีอำนาจจำแนกติดลบออกมาวิเคราะห์ ข้อมูลตามวัตถุประสงค์ของการวิจัย

6. สรุปและจัดทำรายงานการวิจัย

การวิเคราะห์ข้อมูล

1. ข้อมูลทั่วไป วิเคราะห์ด้วยค่าร้อยละ
2. การศึกษาคุณภาพของข้อสอบเป็นรายข้อจากค่าความยากง่ายและค่าอำนาจจำแนกของข้อสอบ วิเคราะห์ด้วยค่าร้อยละ
3. การเปลี่ยนแปลงของคะแนนสอบรายวิชาหลังการตัดข้อสอบที่มีอำนาจจำแนกติดลบออก วิเคราะห์โดยใช้สถิติ Paired t - test (one - tailed)
4. การเปลี่ยนแปลงของอันดับคะแนนสอบรายวิชาหลังการตัดข้อสอบที่มีอำนาจจำแนกติดลบออก วิเคราะห์โดยใช้สถิติ Wilcoxon Signed Ranks test

บทที่ 4 ผลการวิจัย

ในการศึกษาวิจัยเรื่องผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบและความเชื่อมั่นของข้อสอบทั้งฉบับ : รายวิชาการพยาบาลเด็กและวัยรุ่นเก็บรวบรวมข้อมูลจากชุดข้อสอบและคะแนนการสอบของนักศึกษาหลักสูตรพยาบาลศาสตรบัณฑิตชั้นปีที่ 2 ภาคการเรียนที่ 1 ปีการศึกษา 2555 จำนวน 88 คนที่ผ่านการสอบ 4 ครั้ง ได้แก่ การสอบครั้งที่ 1 ครั้งที่ 2 ครั้งที่ 3 และการสอบปลายภาคแล้วนำไปวิเคราะห์ทางสถิติ ผลการวิเคราะห์ข้อมูลนำเสนอเป็น 5 ส่วนดังนี้

ส่วนที่ 1 ข้อมูลทั่วไป

ส่วนที่ 2 ผลการวิเคราะห์คุณภาพข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น

2.1 การวิเคราะห์คุณภาพข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นเป็นรายข้อ

2.1.1 ค่าความยากง่ายของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น

2.1.2 ค่าอำนาจจำแนกของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น

2.2 การวิเคราะห์คุณภาพข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นทั้งฉบับ จากค่าความเชื่อมั่นของข้อสอบ

ส่วนที่ 3 ผลการวิเคราะห์การเปลี่ยนแปลงของคะแนนการสอบของนักศึกษาในรายวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ส่วนที่ 4 ผลการวิเคราะห์การเปลี่ยนแปลงของลำดับคะแนนการสอบของนักศึกษารายวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ส่วนที่ 5 ผลการวิเคราะห์การเปลี่ยนแปลงค่าความเชื่อมั่นของข้อสอบทั้งฉบับรายวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ส่วนที่ 1 ข้อมูลทั่วไป

การวิจัยครั้งนี้ เป็นการศึกษาย้อนหลังจากชุดข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นที่ใช้ในการสอบของนักศึกษาหลักสูตรพยาบาลศาสตรบัณฑิตชั้นปีที่ 2 ภาคการเรียนที่ 1 ปีการศึกษา 2555 จำนวน 88 คน การสอบทั้งสิ้น 4 ครั้ง ได้แก่ การสอบครั้งที่ 1 ครั้งที่ 2 ครั้งที่ 3 และการสอบปลายภาค จำนวนข้อสอบดังแสดงรายละเอียดตามตารางที่ 4.1

ตารางที่ 4.1 จำนวนข้อสอบที่ใช้ในการสอบครั้งที่ 1 ครั้งที่ 2 ครั้งที่ 3 และปลายภาค รายวิชาการพยาบาลเด็กและวัยรุ่น

การสอบ	จำนวนข้อสอบ (ข้อ)
สอบครั้งที่ 1	75
สอบครั้งที่ 2	70
สอบครั้งที่ 3	58
สอบปลายภาค	75
รวม	278

จากตารางที่ 4.1 พบว่า ในการสอบครั้งที่ 1 และการสอบปลายภาคมีจำนวนข้อสอบมากที่สุดโดยมีจำนวนข้อสอบเท่ากันคือ 75 ข้อ รองลงมาเป็นการสอบครั้งที่ 2 มีจำนวนข้อสอบ 70 ข้อ ส่วนครั้งที่ 3 มีจำนวนข้อสอบน้อยที่สุดคือ 58 ข้อ

ส่วนที่ 2 ผลการวิเคราะห์คุณภาพข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่น

2.1 การวิเคราะห์คุณภาพข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นเป็นรายข้อ

2.1.1 ค่าความยากง่ายของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น

ตารางที่ 4.2 จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 1 จำแนกตามค่าความยากง่าย (p) รายข้อ

ค่าความยากง่าย (p)	ความหมาย	จำนวนข้อสอบ	ร้อยละ
0.81 - 1.00	ง่ายมาก	9	12.00
0.61 - 0.80	ค่อนข้างง่าย	19	25.34
0.41 - 0.60	พอเหมาะ	28	37.33
0.20 - 0.40	ค่อนข้างยาก	15	20.00
0.00 - 0.19	ยากมาก	4	5.33
รวมข้อสอบ		75	100

จากตารางที่ 4.2 พบว่า การสอบครั้งที่ 1 รายวิชาการพยาบาลเด็กและวัยรุ่น ส่วนใหญ่เป็นข้อสอบที่มีความยากง่ายพอเหมาะคิดเป็นร้อยละ 37.33 รองลงมาเป็นข้อสอบค่อนข้างง่ายคิดเป็นร้อยละ 25.34 ส่วนข้อสอบที่ยากมากมีจำนวนน้อยที่สุดคิดเป็นร้อยละ 5.33

หากพิจารณาข้อสอบที่มีค่าความยากง่ายในเกณฑ์ที่เหมาะสมควรเก็บไว้ใช้คือ อยู่ระหว่าง 0.20 - 0.80 มีจำนวน 62 ข้อ คิดเป็นร้อยละ 82.67

ตารางที่ 4.3 จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 2 จำแนกตามค่าความยากง่าย (p) รายข้อ

ค่าความยากง่าย (p)	ความหมาย	จำนวนข้อสอบ	ร้อยละ
0.81 - 1.00	ง่ายมาก	7	10.00
0.61 - 0.80	ค่อนข้างง่าย	24	34.29
0.41 - 0.60	พอเหมาะ	18	25.71
0.20 - 0.40	ค่อนข้างยาก	18	25.71
0.00 - 0.19	ยากมาก	3	4.29
รวมข้อสอบ		75	70

จากตารางที่ 4.3 พบว่า การสอบครั้งที่ 2 รายวิชาการพยาบาลเด็กและวัยรุ่น ส่วนใหญ่เป็นข้อสอบที่ค่อนข้างง่ายคิดเป็นร้อยละ 34.29 รองลงมาเป็นข้อสอบที่มีความยากง่ายพอเหมาะและค่อนข้างยากจำนวนเท่ากันคิดเป็นร้อยละ 25.71 ส่วนข้อสอบที่ยากมากมีจำนวนน้อยที่สุดคิดเป็นร้อยละ 4.29

หากพิจารณาข้อสอบที่มีค่าความยากง่ายในเกณฑ์ที่เหมาะสมควรเก็บไว้ใช้คือ อยู่ระหว่าง 0.20 - 0.80 มีจำนวน 60 ข้อ คิดเป็นร้อยละ 85.71

ตารางที่ 4.4 จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 3 จำแนกตามค่าความยากง่าย (p) รายข้อ

ค่าความยากง่าย (p)	ความหมาย	จำนวนข้อสอบ	ร้อยละ
0.81 - 1.00	ง่ายมาก	4	6.90
0.61 - 0.80	ค่อนข้างง่าย	23	39.66
0.41 - 0.60	พอเหมาะ	13	22.41
0.20 - 0.40	ค่อนข้างยาก	11	18.96
0.00 - 0.19	ยากมาก	7	12.07
รวมข้อสอบ		58	100

จากตารางที่ 4.4 พบว่า การสอบครั้งที่ 3 รายวิชาการพยาบาลเด็กและวัยรุ่น ส่วนใหญ่เป็นข้อสอบที่ค่อนข้างง่ายคิดเป็นร้อยละ 39.66 รองลงมาเป็นข้อสอบที่มีความยากง่ายพอเหมาะคิดเป็นร้อยละ 22.41 ส่วนข้อสอบที่ง่ายมากมีจำนวนน้อยที่สุดคิดเป็นร้อยละ 6.90

หากพิจารณาข้อสอบที่มีค่าความยากง่ายในเกณฑ์ที่เหมาะสมควรเก็บไว้ใช้คือ อยู่ระหว่าง 0.20 - 0.80 มีจำนวน 47 ข้อ คิดเป็นร้อยละ 81.03

ตารางที่ 4.5 จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นในการสอบปลายภาค จำแนกตามค่าความยากง่าย (p) รายข้อ

ค่าความยากง่าย (p)	ความหมาย	จำนวนข้อสอบ	ร้อยละ
0.81 - 1.00	ง่ายมาก	10	13.33
0.61 - 0.80	ค่อนข้างง่าย	25	33.33
0.41 - 0.60	พอเหมาะ	21	28.00
0.20 - 0.40	ค่อนข้างยาก	14	18.67
0.00 - 0.19	ยากมาก	5	6.67
รวมข้อสอบ		75	100

จากตารางที่ 4.5 พบว่า การสอบปลายภาครายวิชาการพยาบาลเด็กและวัยรุ่น ส่วนใหญ่เป็นข้อสอบที่ค่อนข้างง่ายคิดเป็นร้อยละ 33.33 รองลงมาเป็นข้อสอบที่มีค่าความยากง่ายพอเหมาะคิดเป็นร้อยละ 28 ส่วนข้อสอบที่ยากมากมีจำนวนน้อยที่สุดคิดเป็นร้อยละ 6.67

หากพิจารณาข้อสอบที่มีค่าความยากง่ายในเกณฑ์ที่เหมาะสมควรเก็บไว้ใช้คือ อยู่ระหว่าง 0.20 - 0.80 มีจำนวน 60 ข้อ คิดเป็นร้อยละ 80

2.1.2 ค่าอำนาจจำแนกของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่น

ตารางที่ 4.6 จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 1 จำแนกตามค่าอำนาจจำแนก (r) รายข้อ

ค่าอำนาจจำแนก(r)	ความหมาย	จำนวนข้อสอบ	ร้อยละ
0.80 - 1.00	อำนาจจำแนกดีมาก	0	0
0.60 - 0.79	อำนาจจำแนกดี	1	1.33
0.40 - 0.59	อำนาจจำแนกปานกลาง	8	10.67
0.20 - 0.39	อำนาจจำแนกพอใช้	28	37.33
0.01 - 0.19	อำนาจจำแนกต่ำ (ควรปรับปรุงหรือตัดทิ้ง)	25	33.34
-1.00 - 0.00	อำนาจจำแนกต่ำมาก (ควรตัดทิ้ง)	13	17.33
รวมข้อสอบ		75	100

จากตารางที่ 4.6 พบว่า การสอบครั้งที่ 1 รายวิชาการพยาบาลเด็กและวัยรุ่น ส่วนใหญ่เป็นข้อสอบที่ค่าอำนาจจำแนกพอใช้ คิดเป็นร้อยละ 37.33 รองลงมาคือมีอำนาจจำแนกต่ำ (ควรปรับปรุงหรือตัดทิ้ง) คิดเป็นร้อยละ 33.34 และน้อยสุดเป็นข้อสอบที่มีอำนาจจำแนกดี คิดเป็นร้อยละ 1.34

หากพิจารณาข้อสอบที่มีค่าอำนาจจำแนกในเกณฑ์ที่เหมาะสมคือ $r \geq 0.20$ พบว่า มีจำนวน 37 ข้อ คิดเป็นร้อยละ 49.33

ตารางที่ 4.7 จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 2 จำแนกตามค่าอำนาจจำแนก (r) รายข้อ

ค่าอำนาจจำแนก(r)	ความหมาย	จำนวนข้อสอบ	ร้อยละ
0.80 - 1.00	อำนาจจำแนกดีมาก	0	0
0.60 - 0.79	อำนาจจำแนกดี	2	2.86
0.40 - 0.59	อำนาจจำแนกปานกลาง	10	14.28
0.20 - 0.39	อำนาจจำแนกพอใช้	27	38.57
0.01 - 0.19	อำนาจจำแนกต่ำ (ควรปรับปรุงหรือตัดทิ้ง)	23	32.86
-1.00 - 0.00	อำนาจจำแนกต่ำมาก (ควรตัดทิ้ง)	8	11.43
รวมข้อสอบ		70	100

จากตารางที่ 4.7 พบว่า การสอบครั้งที่ 2 รายวิชาการพยาบาลเด็กและวัยรุ่น ส่วนใหญ่เป็นข้อสอบที่ค่าอำนาจจำแนกพอใช้ คิดเป็นร้อยละ 38.57 รองลงมาคือมีค่าอำนาจจำแนกต่ำ (ควรปรับปรุงหรือตัดทิ้ง) คิดเป็นร้อยละ 32.86 และน้อยสุดเป็นข้อสอบที่มีอำนาจจำแนกดี คิดเป็นร้อยละ 2.86

หากพิจารณาข้อสอบที่มีค่าอำนาจจำแนกในเกณฑ์ที่เหมาะสมคือ $r \geq 0.20$ พบว่า มีจำนวน 39 ข้อ คิดเป็นร้อยละ 55.71

ตารางที่ 4.8 จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 3 จำแนกตามค่าอำนาจจำแนก (r) รายข้อ

ค่าอำนาจจำแนก(r)	ความหมาย	จำนวนข้อสอบ	ร้อยละ
0.80 - 1.00	อำนาจจำแนกดีมาก	0	0
0.60 - 0.79	อำนาจจำแนกดี	0	0
0.40 - 0.59	อำนาจจำแนกปานกลาง	8	13.79
0.20 - 0.39	อำนาจจำแนกพอใช้	19	32.76
0.01 - 0.19	อำนาจจำแนกต่ำ (ควรปรับปรุงหรือตัดทิ้ง)	23	39.66
-1.00 - 0.00	อำนาจจำแนกต่ำมาก (ควรตัดทิ้ง)	8	13.79
รวมข้อสอบ		58	100

จากตารางที่ 4.8 พบว่า การสอบครั้งที่ 3 รายวิชาการพยาบาลเด็กและวัยรุ่น ส่วนใหญ่เป็นข้อสอบที่ค่าอำนาจจำแนกต่ำ (ควรปรับปรุงหรือตัดทิ้ง) คิดเป็นร้อยละ 39.66 รองลงมาคือมีค่าอำนาจจำแนกพอใช้ คิดเป็นร้อยละ 32.76 และน้อยที่สุดเป็นข้อสอบที่มีอำนาจจำแนกปานกลางและอำนาจจำแนกต่ำมาก (ควรตัดทิ้ง) คิดเป็นร้อยละ 13.79 เท่ากัน

หากพิจารณาข้อสอบที่มีค่าอำนาจจำแนกในเกณฑ์ที่เหมาะสมคือ $r \geq 0.20$ พบว่า มีจำนวน 27 ข้อ คิดเป็นร้อยละ 46.55

ตารางที่ 4.9 จำนวนและร้อยละของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นในการสอบปลายภาค จำแนกตามค่าอำนาจจำแนก (r) รายข้อ

ค่าอำนาจจำแนก(r)	ความหมาย	จำนวนข้อสอบ	ร้อยละ
0.80 - 1.00	อำนาจจำแนกดีมาก	0	0
0.60 - 0.79	อำนาจจำแนกดี	2	2.66
0.40 - 0.59	อำนาจจำแนกปานกลาง	5	6.67
0.20 - 0.39	อำนาจจำแนกพอใช้	29	38.67
0.01 - 0.19	อำนาจจำแนกต่ำ (ควรปรับปรุงหรือตัดทิ้ง)	27	36.00
-1.00 - 0.00	อำนาจจำแนกต่ำมาก (ควรตัดทิ้ง)	12	16.00
รวมข้อสอบ		75	100

จากตารางที่ 4.9 พบว่า การสอบปลายภาครายวิชาการพยาบาลเด็กและวัยรุ่น ส่วนใหญ่เป็นข้อสอบที่ค่าอำนาจจำแนกพอใช้คิดเป็นร้อยละ 38.67 รองลงมาคือค่าอำนาจจำแนกต่ำ (ควรปรับปรุงหรือตัดทิ้ง) คิดเป็นร้อยละ 36 และน้อยสุดเป็นข้อสอบที่มีอำนาจจำแนกดีคิดเป็นร้อยละ 2.66

หากพิจารณาข้อสอบที่มีค่าอำนาจจำแนกในเกณฑ์ที่เหมาะสมคือ $r \geq 0.20$ พบว่า มีจำนวน 36 ข้อ หรือคิดเป็นร้อยละ 48

1.2 การวิเคราะห์คุณภาพข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นทั้งฉบับ

ตารางที่ 4.10 ค่าความเชื่อมั่นของข้อสอบทั้งฉบับรายวิชาการพยาบาลเด็กและวัยรุ่น

ชุดข้อสอบ	ค่าความเชื่อมั่นของข้อสอบ(KR - 20)	การแปลความหมาย
สอบครั้งที่ 1	0.60	ค่อนข้างต่ำ ควรวัดด้วยเครื่องมืออื่นเสริมก่อนตัดเกรด และมีข้อสอบบางข้อที่ต้องปรับปรุงคุณภาพ
สอบครั้งที่ 2	0.66	ค่อนข้างต่ำ ควรวัดด้วยเครื่องมืออื่นเสริมก่อนตัดเกรด และมีข้อสอบบางข้อที่ต้องปรับปรุงคุณภาพ
สอบครั้งที่ 3	0.50	ควรปรับปรุงจำเป็นต้องการวัดด้วยเครื่องมือเสริมก่อนจะตัดสินเกรด
สอบปลายภาค	0.66	ค่อนข้างต่ำ ควรวัดด้วยเครื่องมืออื่นเสริมก่อนตัดเกรด และมีข้อสอบบางข้อที่ต้องปรับปรุงคุณภาพ

จากตารางที่ 4.10 พบว่า ค่าความเชื่อมั่นของข้อสอบทั้งฉบับรายวิชาการพยาบาลเด็กและวัยรุ่น ส่วนใหญ่มีค่าความเชื่อมั่นค่อนข้างต่ำ มีข้อสอบบางข้อที่ต้องปรับปรุงคุณภาพ โดยมีค่าความเชื่อมั่นของข้อสอบทั้งฉบับ ในการสอบครั้งที่ 2 และการสอบปลายภาคมากที่สุด ซึ่งมีค่าเท่ากันคือ 0.66 ส่วนข้อสอบในการสอบครั้งที่ 3 มีค่าความเชื่อมั่นทั้งฉบับต่ำสุดคือ 0.50 ซึ่งถือว่าควรปรับปรุง

ส่วนที่ 3 ผลการวิเคราะห์การเปลี่ยนแปลงคะแนนการสอบของนักศึกษาในรายวิชาการ
พยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบ (r) ออก

ตารางที่ 4.11 การเปรียบเทียบค่าคะแนนการสอบของนักศึกษาในรายวิชาการพยาบาลเด็กและวัยรุ่น
ก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ข้อสอบ	คะแนนก่อนตัด ข้อสอบที่มีค่า r ติดลบออก		คะแนนหลังตัด ข้อสอบที่มีค่า r ติดลบออก		คะแนนที่ เพิ่มขึ้น	จำนวน ข้อสอบที่มี ค่า r ติดลบ (ร้อยละ)	Paired t - test	p-value
	Mean	S.D.	Mean	S.D.				
สอบครั้งที่ 1	53.14	8.29	54.50	9.09	1.36	6.67	8.714	0.000*
สอบครั้งที่ 2	54.42	9.20	54.35	10.56	-0.07	8.57	-.306	0.380
สอบครั้งที่ 3	53.08	7.80	55.25	8.28	2.17	5.17	17.995	0.000*
สอบปลายภาค	56.83	8.74	58.05	9.35	1.22	5.33	8.016	0.000*

* P < 0.01

จากตารางที่ 4.11 พบว่า หลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก นักศึกษามี
คะแนนสอบเฉลี่ยในการสอบครั้งที่ 1 ครั้งที่ 3 และสอบปลายภาควิชาการพยาบาลเด็กและวัยรุ่น
สูงขึ้นอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ส่วนการสอบครั้งที่ 2 คะแนนสอบเฉลี่ยของนักศึกษาไม่
มีการเปลี่ยนแปลงอย่างมีนัยสำคัญทางสถิติ ($\alpha=0.01$)

ส่วนที่ 4 ผลการวิเคราะห์การเปลี่ยนแปลงลำดับคะแนนการสอบของนักศึกษาในรายวิชาการพยาบาลเด็กและวัยรุ่น หลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ตารางที่ 4.12 การเปลี่ยนแปลงของลำดับคะแนนการสอบของนักศึกษาในรายวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ข้อสอบ	นักศึกษาที่มีลำดับคะแนนเพิ่มขึ้น		นักศึกษาที่มีลำดับคะแนนคงเดิม		นักศึกษาที่มีลำดับคะแนนลดลง		Wilcoxon Signed Ranks Test	p-value
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ		
สอบครั้งที่ 1	71	80.68	3	3.41	14	15.91	6.504	0.000*
สอบครั้งที่ 2	41	46.59	2	2.27	45	51.14	.306	0.380
สอบครั้งที่ 3	84	95.45	0	0	4	4.55	7.934	0.000*
สอบปลายภาค	73	82.95	0	0	15	17.05	6.315	0.000*

* $P < 0.01$

จากตารางที่ 4.12 พบว่า หลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก นักศึกษามีการเปลี่ยนแปลงลำดับคะแนนการสอบวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 1 ครั้งที่ 3 และการสอบปลายภาคอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ส่วนในการสอบครั้งที่ 2 นักศึกษาไม่มีการเปลี่ยนแปลงลำดับคะแนนการสอบอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

ส่วนที่ 5 ผลการวิเคราะห์การเปลี่ยนแปลงค่าความเชื่อมั่นของข้อสอบทั้งฉบับในรายวิชาการพยาบาลเด็กและวัยรุ่น หลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ตารางที่ 4.13 ค่าความเชื่อมั่นของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นก่อนและหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ข้อสอบ	จำนวนข้อสอบทั้งหมด	จำนวนข้อสอบที่มีค่า r ติดลบ	ร้อยละข้อสอบที่มีค่า r ติดลบ	ค่าความเชื่อมั่นของชุดข้อสอบ (KR - 20)		
				ก่อนตัดข้อที่มีค่า r ติดลบ	หลังตัดข้อที่มีค่า r ติดลบ	เพิ่มขึ้น/ลดลง
สอบครั้งที่ 1	75	5	6.67	0.60	0.65	0.05
สอบครั้งที่ 2	70	6	8.57	0.66	0.72	0.06
สอบครั้งที่ 3	58	3	5.17	0.50	0.46	- 0.04
สอบปลายภาค	75	4	5.33	0.66	0.69	0.03
รวม	278	18				

จากตารางที่ 4.13 พบว่า ข้อสอบส่วนใหญ่มีค่าความเชื่อมั่นเพิ่มขึ้นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก โดยการสอบครั้งที่ 1 เพิ่มขึ้นจาก 0.60 เป็น 0.65 การสอบครั้งที่ 2 เพิ่มขึ้นจาก 0.66 เป็น 0.72 และการสอบปลายภาคเพิ่มขึ้นจาก 0.66 เป็น 0.69 มีเพียงชุดเดียวคือ การสอบครั้งที่ 3 ที่มีค่าความเชื่อมั่นลดลงจาก 0.50 เป็น 0.46

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

สรุปผลการวิจัย

การวิจัยเรื่องผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบและความเชื่อมั่นของข้อสอบทั้งฉบับ: รายวิชาการพยาบาลเด็กและวัยรุ่น เป็นการศึกษาย้อนหลังจากชุดข้อสอบและคะแนนการสอบรายวิชาการพยาบาลเด็กและวัยรุ่นภาคการเรียนที่ 1 ปีการศึกษา 2555 โดยมีวัตถุประสงค์เพื่อศึกษา 1) คุณภาพของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่น 2) การเปลี่ยนแปลงคะแนนการสอบของนักศึกษาในรายวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก 3) การเปลี่ยนแปลงลำดับคะแนนการสอบของนักศึกษาในรายวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก 4) การเปลี่ยนแปลงค่าความเชื่อมั่นของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นทั้งฉบับหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก เก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างซึ่งเป็นประชากรทั้งหมดที่เข้าถึงได้คือ ชุดข้อสอบและคะแนนการสอบของนักศึกษาหลักสูตรพยาบาลศาสตรบัณฑิตชั้นปีที่ 2 ภาคการเรียนที่ 1 ปีการศึกษา 2555 จำนวน 88 คน

ชุดข้อสอบที่ใช้ในการทดสอบเป็นแบบเลือกตอบ 4 ตัวเลือก มีจำนวน 4 ฉบับ โดยการสอบครั้งที่ 1 และการสอบปลายภาคมีจำนวนข้อสอบมากที่สุดและมีจำนวนข้อสอบเท่ากันคือ 75 ข้อ รองลงมาเป็นการสอบครั้งที่ 2 มีจำนวนข้อสอบ 70 ข้อ ส่วนครั้งที่ 3 มีจำนวนข้อสอบน้อยที่สุดคือ 58 ข้อ เครื่องมือที่ใช้ในการวิจัยประกอบด้วยแบบบันทึกข้อมูลที่ผู้วิจัยใช้สำหรับเก็บข้อมูลจากการสอบแต่ละครั้ง และโปรแกรมการวิเคราะห์ข้อสอบรายข้อ (Item analysis) ที่พัฒนาโดยบริษัท Application Support & Development Department: Control Data (Thailand) Ltd. สถิติที่ใช้ในการวิจัยได้แก่ Paired t - test (one - tailed) และ Wilcoxon Signed Ranks test

ผลการวิจัยสรุปได้ดังนี้

1. การวิเคราะห์คุณภาพของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น

เมื่อพิจารณาคูณภาพข้อสอบเป็นรายข้อโดยใช้ค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) ของข้อสอบเป็นดัชนีแสดงถึงคุณภาพ และใช้ค่าความเชื่อมั่นเป็นดัชนีแสดงถึงคุณภาพของข้อสอบทั้งฉบับ พบว่า ในการสอบครั้งที่ 1 จำนวนข้อสอบทั้งหมด 75 ข้อ เป็นข้อสอบที่มีค่าความยากง่ายในเกณฑ์ที่เหมาะสม ควรเก็บไว้ใช้คืออยู่ระหว่าง 0.20 - 0.80 จำนวน 62 ข้อ คิดเป็นร้อยละ 82.67 มีค่าอำนาจจำแนกในเกณฑ์ที่เหมาะสม ($r \geq 0.20$) จำนวน 37 ข้อคิดเป็นร้อยละ 49.33 และมีค่าอำนาจจำแนกติดลบจำนวน 5 ข้อคิดเป็นร้อยละ 6.67 ส่วนด้านคุณภาพของข้อสอบทั้งฉบับมีค่า

สัมประสิทธิ์ความเชื่อมั่น 0.60 ซึ่งถือว่าค่อนข้างต่ำ ในการสอบครั้งที่ 2 จำนวนข้อสอบทั้งหมด 70 ข้อ เป็นข้อสอบที่มีค่าความยากง่ายในเกณฑ์ที่เหมาะสม ควรเก็บไว้ใช้คืออยู่ระหว่าง 0.20 - 0.80 จำนวน 60 ข้อ คิดเป็นร้อยละ 85.71 มีค่าอำนาจจำแนกในเกณฑ์ที่เหมาะสม ($r \geq 0.20$) จำนวน 39 ข้อ คิดเป็นร้อยละ 55.71 และมีค่าอำนาจจำแนกติดลบจำนวน 6 ข้อ คิดเป็นร้อยละ 8.57 ส่วนด้านคุณภาพของข้อสอบทั้งฉบับมีค่าสัมประสิทธิ์ความเชื่อมั่น 0.66 ซึ่งถือว่าค่อนข้างต่ำ ในการสอบครั้งที่ 3 จำนวนข้อสอบทั้งหมด 58 ข้อ เป็นข้อสอบที่มีค่าความยากง่ายในเกณฑ์ที่เหมาะสม ควรเก็บไว้ใช้คืออยู่ระหว่าง 0.20 - 0.80 จำนวน 47 ข้อ คิดเป็นร้อยละ 81.03 มีค่าอำนาจจำแนกในเกณฑ์ที่เหมาะสม ($r \geq 0.20$) มีจำนวน 27 ข้อ คิดเป็นร้อยละ 46.55 มีค่าอำนาจจำแนกติดลบจำนวน 3 ข้อ คิดเป็นร้อยละ 5.17 ส่วนด้านคุณภาพของข้อสอบทั้งฉบับมีค่าสัมประสิทธิ์ความเชื่อมั่น 0.50 ซึ่งถือว่าควรมีการปรับปรุง ในการสอบปลายภาคจำนวนข้อสอบทั้งหมด 75 ข้อ เป็นข้อสอบที่มีค่าความยากง่ายในเกณฑ์ที่เหมาะสมควรเก็บไว้ใช้คืออยู่ระหว่าง 0.20 - 0.80 จำนวน 60 ข้อ คิดเป็นร้อยละ 80 มีค่าอำนาจจำแนกในเกณฑ์ที่เหมาะสม ($r \geq 0.20$) จำนวน 36 ข้อ คิดเป็นร้อยละ 48 มีค่าอำนาจจำแนกติดลบจำนวน 4 ข้อ คิดเป็นร้อยละ 5.33 ส่วนด้านคุณภาพของข้อสอบทั้งฉบับมีค่าสัมประสิทธิ์ความเชื่อมั่น 0.66 ซึ่งถือว่าค่อนข้างต่ำ

2. การวิเคราะห์การเปลี่ยนแปลงคะแนนการสอบของนักศึกษาในรายวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออกด้วยสถิติ Paired t – test (one - tailed) พบว่า หลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก นักศึกษามีคะแนนสอบเฉลี่ยในการสอบครั้งที่ 1 ครั้งที่ 3 และสอบปลายภาคสูงขึ้นอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ส่วนการสอบครั้งที่ 2 คะแนนสอบเฉลี่ยของนักศึกษาไม่มีการเปลี่ยนแปลงอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

3. การวิเคราะห์การเปลี่ยนแปลงของลำดับคะแนนการสอบของนักศึกษา ในรายวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออกด้วยสถิติ Wilcoxon Signed Ranks test พบว่า หลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก นักศึกษามีการเปลี่ยนแปลงลำดับคะแนนการสอบวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 1 ครั้งที่ 3 และการสอบปลายภาคอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ส่วนในการสอบครั้งที่ 2 นักศึกษาไม่มีการเปลี่ยนแปลงลำดับคะแนนการสอบอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

4. การวิเคราะห์การเปลี่ยนแปลงค่าความเชื่อมั่นของชุดข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นทั้งหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก พบว่า ส่วนใหญ่มีค่าความเชื่อมั่นเพิ่มขึ้น หลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบ กล่าวคือ การสอบครั้งที่ 1 เพิ่มขึ้นจาก 0.60 เป็น 0.65 การสอบครั้งที่ 2 เพิ่มขึ้นจาก 0.66 เป็น 0.72 และการสอบปลายภาคเพิ่มขึ้นจาก 0.66 เป็น 0.69 มีเพียงชุดเดียวคือ การสอบครั้งที่ 3 ที่มีค่าความเชื่อมั่นลดลงจาก 0.50 เป็น 0.46

อภิปรายผล

1. คุณภาพข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น

1.1 ผลการศึกษาคุณภาพข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นเป็นรายข้อ โดยใช้ค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) เป็นดัชนีแสดงถึงคุณภาพของข้อสอบพบว่า ในการสอบทั้ง 4 ครั้ง ได้แก่ สอบครั้งที่ 1 ครั้งที่ 2 ครั้งที่ 3 และการสอบปลายภาค ข้อสอบส่วนใหญ่มีค่าความยากง่ายในเกณฑ์ที่เหมาะสม ควรเก็บไว้คืออยู่ระหว่าง 0.20 - 0.80 โดยมีจำนวนร้อยละ 82.67, 85.71, 81.03 และ 80 ตามลำดับ อธิบายได้ว่า การจัดทำข้อสอบให้มีค่าความยากง่ายเหมาะสมนั้น เกิดจากคณะผู้สอนได้จัดทำข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นทั้ง 4 ครั้งให้สอดคล้องกับวัตถุประสงค์การเรียนรู้ตามที่กำหนดไว้ในตารางวิเคราะห์หลักสูตร (Test blueprint) ในมคอ.3 นับเป็นการสร้างข้อสอบให้มีความตรงตามโครงสร้าง อีกทั้งทุกครั้งที่การจัดทำข้อสอบจะต้องผ่านการวิพากษ์ในส่วนของข้อคำถาม ตัวเลือกและตัวลวงจากคณะผู้สอนจำนวน 5 คนซึ่งมีประสบการณ์การสอนตั้งแต่ 3 ปีขึ้นไปร่วมกับคณะกรรมการบริหารหลักสูตร หลังจากนั้นคณะผู้สอนได้นำข้อสอบไปปรับปรุงแล้วจัดทำเป็นข้อสอบฉบับสมบูรณ์ ทำให้ข้อสอบมีความตรงตามเนื้อหา สอดคล้องกับที่บุญใจ ศรีสถิตนรากร (2555, น. 341) กล่าวไว้ว่า แบบทดสอบที่ดีต้องมีระดับความยากเหมาะสมกับสติปัญญาของผู้สอบ ลักษณะวิชาและวัตถุประสงค์ของการวัดผล

เมื่อพิจารณาคุณภาพข้อสอบเป็นรายข้อจากค่าอำนาจจำแนก (r) นั้นพบว่า การสอบทั้ง 4 ครั้งมีข้อสอบที่มีค่าอำนาจจำแนกในเกณฑ์ที่เหมาะสม ($r \geq 0.20$) คิดเป็นร้อยละ 49.33, 55.71, 46.55 และ 48 ตามลำดับ ซึ่งถือว่าอยู่ในระดับปานกลาง อธิบายได้ว่า ในการออกข้อสอบนั้นคณะผู้สอนใช้ข้อสอบที่ปรับปรุงจากข้อสอบเดิมและข้อสอบใหม่ร่วมกับข้อสอบเดิม เพื่อช่วยเพิ่มประสิทธิภาพในการจำแนกผู้สอบได้ส่วนหนึ่ง สอดคล้องกับการศึกษาของ Usova, G M. (1997, pp. 100 -110) ที่ศึกษาประสิทธิผลของการใช้ข้อสอบเก่าในคลังข้อสอบร่วมกับข้อสอบใหม่ในการจำแนกผู้สอบ พบว่าข้อสอบที่ได้จากการรวบรวมข้อสอบ 3 แบบเข้าด้วยกันในสัดส่วนที่เหมาะสมจะสามารถจำแนกหรือแยกแยะผู้สอบได้ดี ทั้งนี้ค่าอำนาจจำแนกในการสอบแต่ละครั้งถือว่าไม่สูงมากนัก อาจเนื่องมาจากกระบวนการเรียนการสอนในชั้นเรียนทำให้ความรู้ของนักศึกษาใกล้เคียงกัน โดยผู้สอนมีการสรุปประเด็นสำคัญในตอนท้ายชั่วโมง มีการอภิปรายแลกเปลี่ยนความคิดเห็นในชั้นเรียน มีการมอบหมายให้สืบค้นความรู้เพิ่มเติมจากตำราภาษาอังกฤษ และสรุปใจความสำคัญ รวมทั้งมีการนำกรณีศึกษาปัญหาโรคเด็กและวัยรุ่นให้นักศึกษาวิเคราะห์ วางแผนทางการพยาบาล และนำเสนอหน้าห้องเรียน ทำให้นักศึกษามีความรู้ใกล้เคียงกัน ในการทดสอบจึงจำแนกผู้เรียนเก่งและเรียนอ่อนได้ไม่ชัดเจนนัก ซึ่งสอดคล้องกับที่นิธิภัทร บาลศิริ (2557, น. 220 - 221) กล่าวไว้ว่า ถ้ากลุ่มผู้สอบมีความเป็นเอกพันธ์มากหรือมีความคล้ายคลึงกันมาก เช่น ความรู้ความสามารถเท่าเทียมกัน จะทำให้ความแปรปรวนของคะแนนน้อยหรือมีการกระจายตัวของคะแนนต่ำ ทำให้แบบวัดมีค่าอำนาจจำแนกต่ำ ประกอบกับ

จำนวนข้อสอบมีน้อย ข้อสอบในแต่ละชุดมีจำนวนข้อสอบ 58 -75 ข้อ เพราะยิ่งจำนวนข้อสอบมากขึ้นตั้งแต่ 80 ข้อขึ้นไปจะทำให้ค่าอำนาจจำแนกสูงขึ้นด้วย (อรรวรรณ สุวรรณรัตน์, 2552, น. 387)

1.2 หากพิจารณาคุณภาพของข้อสอบทั้งฉบับด้วยค่าความเชื่อมั่นที่คำนวณได้จากสูตร KR - 20 พบว่า ชุดข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นที่ใช้ในการสอบส่วนใหญ่คือข้อสอบครั้งที่ 1 ครั้งที่ 2 และการสอบปลายภาคมีค่าความเชื่อมั่นอยู่ระหว่าง 0.60 - 0.66 ซึ่งถือว่ามีความเชื่อมั่นค่อนข้างต่ำ ที่เหลือคือข้อสอบครั้งที่ 3 มีค่าความเชื่อมั่น 0.5 ซึ่งถือว่าควรมีการปรับปรุง ทั้งนี้แบบทดสอบที่ใช้ในการทดสอบควรมีค่าสัมประสิทธิ์ความเที่ยงหรือความเชื่อมั่นสูงที่สุดเท่าที่จะเป็นไปได้ในสภาพการณ์ อย่างน้อยที่สุดควรมีค่าไม่ต่ำกว่า 0.50 (ศิริชัย กาญจนวาสี, 2556, น. 97) ซึ่งชุดข้อสอบที่ใช้วัดผลสัมฤทธิ์ในการเรียนวิชาการพยาบาลเด็กและวัยรุ่นทั้ง 4 ฉบับมีค่าความเชื่อมั่นไม่ต่ำกว่า 0.50 การที่ค่าความเชื่อมั่นของแบบทดสอบทั้ง 4 ชุด มีค่าไม่สูงมาก อธิบายว่าอาจเนื่องมาจากกระบวนการเรียนการสอนในชั้นเรียนทำให้ความรู้ของนักศึกษาใกล้เคียงกันในการทดสอบ จึงจำแนกผู้เรียนเก่งและเรียนอ่อนได้ไม่ชัดเจนนัก ซึ่งสอดคล้องกับที่นิธิภัทร บาลศิริ (2557, น. 220 - 221) กล่าวไว้ว่าถ้ากลุ่มผู้สอบมีความเป็นเอกพันธ์มากหรือมีความคล้ายคลึงกันมาก เช่น ความรู้ความสามารถเท่าเทียมกัน ความเชื่อมั่นของแบบทดสอบจะต่ำ อีกทั้งจำนวนข้อสอบในการสอบแต่ละครั้งมีน้อยคือ 58 -75 ข้อ ไม่ถึง 80 ข้อทำให้มีผลต่อค่าความเชื่อมั่นด้วย โดยเฉพาะในการสอบครั้งที่ 3 ที่มีข้อสอบน้อยที่สุดคือ 58 ข้อซึ่งเป็นข้อสอบส่วนที่เป็น Summative จำนวนข้อสอบต่อเนื่องหาที่ต้องการประเมินน้อยกว่าฉบับอื่น จึงมีค่าความเชื่อมั่นน้อยที่สุดเพียง 0.50 สอดคล้องกับที่นิธิภัทร บาลศิริ (2557, น. 220) กล่าวว่าถ้ามีข้อคำถามมากจะวัดได้ครอบคลุมเนื้อหาและพฤติกรรมที่ต้องการวัดก็จะทำให้ค่าความเชื่อมั่นมีค่าเพิ่มสูงขึ้น และสอดคล้องกับผลการศึกษาของพิมพ์อุษา เจริญยศ (2539) ที่ศึกษาเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบเลือกตอบเมื่อลดจำนวนข้อสอบโดยใช้เกณฑ์ค่าสถิติกับใช้ดุลยพินิจของครู พบว่าแบบทดสอบที่มีจำนวน 40 ข้อมีค่าความเชื่อมั่นโดยใช้เกณฑ์ค่าสถิติมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 กับค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลงเหลือ 35 และ 15 ข้อ

2. การเปลี่ยนแปลงคะแนนการสอบของนักศึกษาในรายวิชาการพยาบาลเด็กและวัยรุ่น หลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ผลการศึกษาพบว่า หลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก นักศึกษามีคะแนนสอบเฉลี่ยในการสอบครั้งที่ 1 ครั้งที่ 3 และการสอบปลายภาคสูงขึ้นอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ส่วนการสอบครั้งที่ 2 คะแนนสอบเฉลี่ยของนักศึกษาไม่มีการเปลี่ยนแปลงอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 อธิบายได้ว่า ผลการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก ทำให้เกิดการเปลี่ยนแปลงของคะแนนเพิ่มขึ้นเป็นส่วนใหญ่ อาจเนื่องมาจากค่าอำนาจจำแนกรายข้อของเครื่องมือ

วัดผลการเรียนรู้เป็นสัมประสิทธิ์สหสัมพันธ์ระหว่างความรู้ความสามารถของผู้เรียนกับการตอบถูก (บุญเชิด ภิญญอนันตพงษ์, 2552, น. 73) อำนาจจำแนกของข้อสอบมีความสัมพันธ์กับความสามารถของผู้เข้าสอบ ข้อคำถามที่ดีจะต้องมีค่าอำนาจจำแนกเป็นบวก เพราะถ้าเป็นลบแสดงว่ามีอำนาจจำแนกกลับกันคือ คนเก่งทำข้อสอบไม่ได้ แต่คนอ่อนทำข้อสอบนั้นได้ ถือเป็นข้อสอบที่ไม่ดีควรตัดทิ้ง (ปวรส บุตะเชียว, 2556, น. 2) ข้อสอบที่เหลือจากการตัดข้อที่มีอำนาจจำแนกติดลบออกสามารถวัดความรู้ได้ใกล้เคียงกับความรู้ความสามารถที่แท้จริงของนักศึกษาได้ดีกว่าข้อสอบชุดเดิม ซึ่งมีความคลาดเคลื่อนมากกว่า ดังที่ Tavakol & Dennick, (2011) กล่าวไว้ว่าการลดความคลาดเคลื่อนของการทดสอบให้มีน้อยที่สุดเท่าที่จะทำได้ จะทำให้ผลคะแนนการสอบที่วัดได้นั้นใกล้เคียงกับความรู้ความสามารถที่แท้จริงของนักศึกษามากที่สุด และจากการวิเคราะห์ข้อสอบที่มีอำนาจจำแนกติดลบพบว่า ส่วนใหญ่จะเป็นข้อสอบที่ยากเกินไป ซึ่งผู้เรียนส่วนใหญ่จะทำข้อสอบนั้นไม่ได้ ดังนั้นหากตัดข้อสอบที่มีค่าอำนาจติดลบออก จึงส่งผลให้คะแนนการสอบเพิ่มขึ้น หลังการวิเคราะห์ข้อสอบคณะผู้สอนควรนำข้อสอบที่มีค่าอำนาจจำแนกติดลบมาพิจารณาหาความบกพร่อง ได้แก่ ข้อสอบสามารถวัดความรู้ที่ต้องการวัดได้จริงหรือไม่ ข้อสอบทั้งข้อคำถาม ตัวเลือกและตัวลวงมีความเป็นปรนัยหรือไม่ หรือข้อสอบมีความเป็นตัวแทนของความรู้ที่ต้องการหรือไม่ เพื่อสามารถปรับปรุงข้อสอบได้ตรงเนื้อหาสาระและวัตถุประสงค์การเรียนรู้นำไปสู่การปรับปรุงกระบวนการเรียนการสอนให้มีประสิทธิภาพ

3. การเปลี่ยนแปลงของลำดับคะแนนการสอบของนักศึกษาในรายวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ผลการศึกษาพบว่า หลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบนักศึกษามีการเปลี่ยนแปลงลำดับคะแนนการสอบวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 1 ครั้งที่ 3 และการสอบปลายภาคอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ส่วนในการสอบครั้งที่ 2 นักศึกษาไม่มีการเปลี่ยนแปลงลำดับคะแนนการสอบอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 อธิบายได้ว่า ผลการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก ส่วนใหญ่ทำให้เกิดการเปลี่ยนแปลงของคะแนนเพิ่มขึ้นได้ ดังอภิปรายในข้อ 2 ซึ่งการเพิ่มของคะแนนย่อมมีผลทำให้ลำดับคะแนนเพิ่มขึ้นได้ด้วย

4. การเปลี่ยนแปลงค่าความเชื่อมั่นของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ผลการศึกษาพบว่า ข้อสอบครั้งที่ 1, 2 และปลายภาคมีค่าความเชื่อมั่นเพิ่มขึ้นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก แต่ข้อสอบครั้งที่ 3 มีค่าความเชื่อมั่นลดลง อธิบายได้ว่า ข้อสอบครั้งที่ 1, 2 และปลายภาคมีจำนวนข้อสอบที่ใกล้เคียงกันคือ 75, 70 และ 75 ข้อตามลำดับ ซึ่ง

มากกว่าข้อสอบครั้งที่ 3 ที่มีเพียง 58 ข้อ และข้อสอบที่มีค่าอำนาจจำแนกติดลบที่ถูกตัดออกของข้อสอบครั้งที่ 1, 2 และปลายภาคมีมากกว่าข้อสอบครั้งที่ 3 กล่าวคือข้อสอบครั้งที่ 1, 2 และปลายภาคมีข้อสอบที่มีค่าอำนาจจำแนกติดลบที่ถูกตัดออกคิดเป็นร้อยละ 6.67, 8.57 และ 5.33 ซึ่งมากกว่าจำนวนข้อสอบที่มีค่าอำนาจจำแนกติดลบของข้อสอบครั้งที่ 3 ที่มีเพียงร้อยละ 5.17 การปรับค่าความเชื่อมั่นของข้อสอบทั้งฉบับให้สูงขึ้นทำได้ด้วยการตัดข้อสอบที่มีค่าอำนาจจำแนกน้อยหรืออำนาจจำแนกติดลบออก ค่าความเชื่อมั่นของทั้งฉบับก็จะสูงตามลำดับ ซึ่งสอดคล้องกับพรสันต์ เลิศวิทยาวิวัฒน์ (2556) และ Richard F. Burton (2001)

จากการศึกษาครั้งนี้แสดงให้เห็นถึงความสำคัญของคุณภาพของข้อสอบในด้านอำนาจการจำแนก โดยเฉพาะค่าอำนาจจำแนกติดลบที่มีผลต่อคะแนนการสอบและลำดับคะแนนการสอบในรายวิชาการพยาบาลเด็กและวัยรุ่น รวมทั้งยังส่งผลต่อค่าความเชื่อมั่นของข้อสอบ ซึ่งในการวัดผลทางการศึกษาจำเป็นต้องอาศัยเครื่องมือที่มีคุณภาพเพื่อให้ได้ข้อมูลที่ถูกต้องตรงตามความเป็นจริง และมีความน่าเชื่อถือ สามารถนำไปใช้ในการประเมินผลและตัดสินใจได้

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. ในการพิจารณาคุณภาพข้อสอบวัดผลสัมฤทธิ์การเรียนรู้ ควรพิจารณาถึงคุณภาพข้อสอบรายข้อ ได้แก่ ค่าความยากง่าย ค่าอำนาจจำแนก และคุณภาพข้อสอบทั้งฉบับจากค่าความเชื่อมั่น เพื่อใช้พิจารณาคัดเลือกข้อสอบที่มีคุณภาพ หรือปรับปรุงข้อสอบให้มีคุณภาพเพื่อเก็บไว้ในคลังข้อสอบต่อไป โดยมีการประชุมทีมผู้สอนเพื่อพิจารณาปรับข้อคำถาม ตัวเลือกหรือตัวลวง เพื่อให้มีจำนวนข้อสอบตามระดับความรู้ที่คาดหวังในแต่ละหัวข้อเนื้อหาให้ครบตาม Test blueprint และมั่นใจว่าความตรงเชิงโครงสร้างของข้อสอบไม่เสียไป

2. จากผลการวิจัยพบว่า หากมีการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออกส่วนใหญ่จะทำให้คะแนนสอบเฉลี่ยและลำดับคะแนนของนักศึกษาเปลี่ยนแปลงอย่างมีนัยสำคัญทางสถิติและทำให้ค่าความเชื่อมั่นของข้อสอบสูงขึ้น ดังนั้นจึงควรให้ความสำคัญกับกระบวนการที่จะได้ข้อสอบที่มีคุณภาพทุกขั้นตอน ผู้ออกข้อสอบจึงควรพัฒนาการออกข้อสอบให้มีคุณภาพมากยิ่งขึ้น เพื่อลดจำนวนข้อสอบที่มีค่าอำนาจจำแนกติดลบ ซึ่งไม่สามารถจำแนกกลุ่มนักศึกษาที่เรียนเก่งและเรียนอ่อนออกจากกันได้

3. ควรกำหนดเป็นนโยบายเกี่ยวกับการวัดและประเมินผลสัมฤทธิ์ทางการเรียนในหลักสูตรพยาบาลศาสตรบัณฑิตว่า ในการวิเคราะห์ข้อสอบหลังการสอบทุกครั้ง หากพบข้อสอบที่มีค่าอำนาจจำแนกติดลบควรพิจารณาตัดข้อสอบข้อนั้นออก แล้วจึงประมวลผลคะแนนใหม่ เพื่อให้ค่าความ

เชื่อมั่นของข้อสอบสูงขึ้น ในขณะที่เดียวกันควรพิจารณาด้วยว่าเป็นข้อสอบที่วัดครอบคลุมจุดประสงค์ และเนื้อหาที่ต้องการ ซึ่งในบางครั้งอาจมีความจำเป็นที่จะต้องใช้ข้อสอบที่มีอำนาจการจำแนกไม่สูงนัก เพื่อให้มีข้อสอบที่วัดครอบคลุมเนื้อหาที่ต้องการ ไม่กระทบต่อความตรงตามเนื้อหา

4. ในการสร้างข้อสอบใหม่โดยอาจารย์เขียนขึ้นเอง (Teacher - made test) ควรให้ความสำคัญกับการวิพากษ์ข้อสอบ พิจารณาใช้ข้อสอบที่ผ่านการวิเคราะห์แล้วว่ามีคุณภาพ ร่วมกับข้อสอบที่ปรับปรุงและนำกลับมาใช้อีก โดยอาศัยตารางวิเคราะห์หลักสูตรเป็นแนวทางในการคัดเลือกข้อสอบ และสร้างข้อสอบใหม่

ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

1. ควรศึกษาคุณภาพข้อสอบในด้านอื่นๆที่จะส่งผลต่อค่าความเชื่อมั่นของข้อสอบ นอกเหนือจากค่าความยากง่าย และอำนาจการจำแนก เช่น ประสิทธิภาพตัวलग เป็นต้น

2. ควรศึกษาเปรียบเทียบผลของการตัดข้อสอบที่มีอำนาจจำแนกระดับต่างๆต่อคะแนนการสอบและค่าความเชื่อมั่นของข้อสอบทั้งฉบับ

บรรณานุกรม

บรรณานุกรมภาษาไทย

- โชติกา ภาชีผล. (2554). *การสร้างและพัฒนาเครื่องมือในการวัดและประเมินผลการศึกษา*. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- นิธิภัทร บาลศิริ. (2557). *หลักการวัดและประเมินผลการศึกษา*. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร: โอเอสพริ้นติ้งเฮาส์.
- บุญใจ ศรีสถิตนรากร. (2555). *การพัฒนาและตรวจสอบคุณภาพเครื่องมือวิจัย: คุณสมบัติการวัดเชิงจิตวิทยา*. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- บุญเชิด ภิญโญนนตพงษ์. (2552). หน่วยที่ 3 คุณภาพเครื่องมือวัด. ใน *การพัฒนาเครื่องมือสำหรับประเมินการศึกษา*. (หน้า 73 - 77) พิมพ์ครั้งที่ 4. นนทบุรี: โรงพิมพ์มหาวิทยาลัยสุโขทัยธรรมมาธิราช.
- ปวรส บุตะเขี้ยว. (2556). *การวิเคราะห์ข้อสอบ (Item analysis)*. 8 มกราคม 2556.
www.rta.mi.th/630a0u/file/item_analysis.doc
- พิชิต ฤทธิจรรณ. (2553). *หลักการวัดและประเมินผลการศึกษา*. พิมพ์ครั้งที่ 6. กรุงเทพมหานคร: แฮาส์ออฟเคอร์มีส์.
- พิมพ์อุษา เจริญยศ. (2539). *เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบเลือกตอบเมื่อลดจำนวนข้อสอบโดยใช้เกณฑ์ค่าสถิติกับใช้ดุลยพินิจของครู*. วิทยานิพนธ์หลักสูตรปริญญาการศึกษา มหบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.
- พรสันต์ เลิศวิทยาวิวัฒน์. (2556). *กระบวนการสร้างข้อสอบมาตรฐาน*. 20 มกราคม 2558.
www.ntvc.ac.th./index.php/
- เยาวดี รางชัยกุล วิบูลย์ศรี. (2552). *การวัดผลและการสร้างแบบสอบผลสัมฤทธิ์ (Measurement and achievement test construction)*. พิมพ์ครั้งที่ 8. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- รังสรรค์ มณีเล็ก. (2552). หน่วยที่ 1 แนวคิดพื้นฐานเกี่ยวกับการพัฒนาเครื่องมือสำหรับการประเมินทางการศึกษา. ใน *การพัฒนาเครื่องมือสำหรับประเมินการศึกษา*. (หน้า 5) พิมพ์ครั้งที่ 4. นนทบุรี: โรงพิมพ์มหาวิทยาลัยสุโขทัยธรรมมาธิราช.
- ศิริชัย กาญจนวาสี. (2556). *ทฤษฎีการทดสอบแบบดั้งเดิม (Classical test theory)*. พิมพ์ครั้งที่ 7 ฉบับปรับปรุงเพิ่มเติม. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

- สุรศักดิ์ อมรรัตนศักดิ์. (2554). *การประเมินผลการศึกษา (Education evaluation)*. พิมพ์ครั้งที่ 2 กรุงเทพมหานคร: สำนักพิมพ์มหาวิทยาลัยรามคำแหง.
- สุมาลี จันทร์ชลอ. (2542). *การวัดและประเมินผล*. กรุงเทพมหานคร: ศูนย์สื่อเสริมกรุงเทพ.
- สมนึก ภัททิยธนี. (2555). การวัดผลการศึกษา. พิมพ์ครั้งที่ 8. กทม. สนิช: ประสานการพิมพ์.
- สมบูรณ์ ชิตพงศ์และเตือนใจ เกตุษา. (2554). หน่วยที่ 9 การสร้างและตรวจสอบเครื่องมือวัดด้านพุทธิพิสัย. ใน *เอกสารการสอนชุดวิชา สถิติ วิจัยและประเมินผลการศึกษา หน่วยที่ 9 - 15. ฉบับปรับปรุงครั้งที่ 2*. (หน้า 9 - 5). นนทบุรี: โรงพิมพ์มหาวิทยาลัยสุโขทัยธรรมมาธิราช.
- อานุกาฬ เลขะกุล. *การสร้างข้อสอบแบบเลือกตอบ (Multiple choice question)*. 20 กันยายน 2556. <http://www.med.cmu.ac.th/secret/meded/AOMJAI9A.pdf>
- เอมอร จังศิริพรภรณ์. (2550). *การวัดและประเมินผลการศึกษา*. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- อุทุมพร จามรมาน. (2535). *ข้อสอบ: การสร้างและการพัฒนา*. กรุงเทพมหานคร: ฟีนนี่พับบลิชซิง.
- อรรวรรณ สุวรรณรัตน์. (2552). การหาคุณภาพข้อสอบเพื่อบริหารจัดการด้านการเรียนการสอน. *สงขลานครินทร์เวชสาร*. 27 (5) (กันยายน - ตุลาคม): 381 - 388. 28 พฤศจิกายน 2557 http://medinfo.psu.ac.th/smj2/27_5_09/27_5pdf/03-arawan.pdf

บรรณานุกรมภาษาต่างประเทศ

- Ascalon, M. E., Meyers L. S., Davis, B. W. & Smits, N. (2007). Distractor similarity and Item - stem structure: Effects on item difficulty. *Applied Measurement in Education*, 20 (2), 153 - 170.
- Burton, R. F. (2001). Do item - discrimination indices really help us to improve our tests. *Assessment & Evaluation in Higher Education*, 26 (3), 213 - 220.
- 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.
- Chase, C. I. (1978). *Measurement for Educational Evaluation*. 2nd ed. Reading, M. A.: Addison - Wesley Publishing Company.
- Office of Education Assessment, University of Washington. (2005). *SCOREPAK®: Item analysis*. 8 January 2013. <http://www.Washingtons.edu/oea/score1.htm>.

- Oluseyi, A. E. & Olufemi, A. T. (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), 237 - 246.
- 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.
- Tavakol, M. & Dennick, R. (2011). Post- examination analysis of objective tests. *Medical Teacher*, 33, 447- 458.
- Usova, G. M. (1997). Effective test item discrimination using Bloom's taxonomy. *Education*, 118 (1), 100 -110.

ภาคผนวก

ภาคผนวก ก.

ตารางวิเคราะห์หลักสูตร (Test blueprint)

ตารางที่ ก – 1 ตารางวิเคราะห์หลักสูตร (Test Blueprint) ประกอบการออกข้อสอบวิชาการ
พยาบาลเด็กและวัยรุ่นน ภาคเรียนที่ 1 ปีการศึกษา 2555

หัวข้อสอน	ชั่วโมง	Formative					Summative				
		รู้จำ	เข้าใจ	นำไปใช้	วิเคราะห์	รวม	รู้จำ	เข้าใจ	นำไปใช้	วิเคราะห์	รวม
หน่วยที่ 1 แนวคิดการพยาบาลเด็ก	3	3	4	5	3	15	1	2	2	1	6
หน่วยที่ 2 การส่งเสริมสุขภาพเด็ก	2	2	2	4	2	10	1	1	1	1	4
หน่วยที่ 3 การเจริญเติบโตและ พัฒนาการเด็ก	1	1	1	2	1	5	1	1	-	-	2
หน่วยที่ 4 การพยาบาลเด็กที่มีปัญหา โภชนาการ	2	2	2	4	2	10	1	1	1	1	4
หน่วยที่ 5 การพยาบาลทารกแรกเกิด	1	1	2	2	-	5	-	1	1	-	2
หน่วยที่ 6 การพยาบาลทารกแรกเกิด ที่มีภาวะเสี่ยง	3	3	4	5	3	15	1	2	2	1	6
หน่วยที่ 7 การพยาบาลเด็กที่มีปัญหา ระบบทางเดินหายใจ	3	3	5	5	2	15	1	2	2	1	6
หน่วยที่ 8 การพยาบาลเด็กที่ได้รับ สารพิษและอุบัติเหตุ	3	3	4	5	3	15	1	2	2	1	6
หน่วยที่ 9 การพยาบาลเด็กที่มีปัญหา ติดเชื้อและโรคติดต่อ	3	6	3	4	2	15	1	1	3	1	6
หน่วยที่ 10 การพยาบาลเด็กที่มีปัญหา ระบบทางเดินปัสสาวะ	3	3	4	5	3	15	1	2	2	1	6
หน่วยที่ 11 การพยาบาลเด็กที่มีปัญหา ระบบทางเดินอาหาร	3	3	4	5	3	15	1	2	2	1	6
หน่วยที่ 12 การพยาบาลเด็กที่มีปัญหา ระบบต่อมไร้ท่อ	2	2	3	4	1	10	-	2	1	1	4

หัวข้อสอน	ชั่วโมง	Formative					Summative				
		รู้จัก	เข้าใจ	นำไปใช้	วิเคราะห์	รวม	รู้จัก	เข้าใจ	นำไปใช้	วิเคราะห์	รวม
หน่วยที่ 13 การพยาบาลเด็กที่มีปัญหา ระบบเลือดและการแข็งตัวของเลือด	3	3	4	5	3	15	1	1	3	1	6
หน่วยที่ 14 การพยาบาลเด็กที่มีปัญหา ระบบหัวใจและหลอดเลือด	3	5	4	4	2	15	-	3	3	-	6
หน่วยที่ 15 การพยาบาลเด็กที่มีปัญหา มะเร็ง	3	3	5	5	2	15	-	-	-	-	-
หน่วยที่ 16 การพยาบาลเด็กที่มีปัญหา ระบบประสาทและกล้ามเนื้อ	3	3	4	5	3	15	-	-	-	-	-
หน่วยที่ 17 การพยาบาลเด็กที่มีปัญหา สุขภาพจิตและจิตเวช	3	3	4	6	2	15	-	-	-	-	-
รวม	44	49	59	75	37	220	11	23	25	11	70

ภาคผนวก ข.
ผลการวิเคราะห์ข้อมูลเพิ่มเติม

ผลการวิเคราะห์ข้อมูลคะแนนการสอบวิชา การพยาบาลเด็กและวัยรุ่น
ภาคการเรียนที่ 1 ปีการศึกษา 2555

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Before1	88	33.33	74.67	53.1366	8.28992
After1	88	31.43	78.57	54.4964	9.09363
Before2	88	37.14	74.29	54.4161	9.19637
After2	88	34.38	78.13	54.3517	10.55857
Before3	88	39.66	79.31	53.0761	7.79619
After3	88	40.00	81.82	55.2478	8.28154
Before4	88	30.67	77.33	56.8332	8.73579
After4	88	29.58	81.69	58.0510	9.34930
Valid N (listwise)	88				

T-TEST PAIRS = Before1 Before2 Before3 Before4 WITH After1 After2 After3 After4
(PAIRED)

/ MISSING = ANALYSIS

T-Test

Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 Before1	53.1366	88	8.28992	.88371
After1	54.4964	88	9.09363	.96938
Pair 2 Before2	54.4161	88	9.19637	.98034
After2	54.3517	88	10.55857	1.12555
Pair 3 Before3	53.0761	88	7.79619	.83108
After3	55.2478	88	8.28154	.88281
Pair 4 Before4	56.8332	88	8.73579	.93124
After	58.0510	88	9.34930	.99664

Paired Samples Correlations

	N	Correlation	Sig.
Pair 1 Before1 & After1	88	.990	.000
Pair 2 Before2 & After2	88	.989	.000
Pair 3 Before3 & After3	88	.992	.000
Pair 4 Before4 & After4	88	.990	.000

Paired Samples Test

		Paired Differences		
		Mean	Std. Deviation	Std. Error Mean
Pair 1	Before1 - After1	-1.35977	1.46377	.15604
Pair 2	Before2- After2	.06443	1.97506	.21054
Pair 3	Before3 - After3	-2.17170	1.13214	.12069
Pair 4	Before4 - After4	-1.21784	1.42515	.15192

Paired Samples Test

		Paired Differences		t	df	Sig. (2-tailed)
		95% Confidence Interval of the Difference				
		Lower	Upper			
Pair 1	Before1 - After1	-1.66992	-1.04963	-8.714	87	.000
Pair 2	Before2- After2	-.35404	.48291	.306	87	.760
Pair 3	Before3 - After3	-2.41158	-1.93183	-17.995	87	.000
Pair 4	Before4 - After4	-1.51980	-.91588	-8.016	87	.000

Descriptives

COMPUTE difference1 = After1 – Before1.

EXECUTE.

COMPUTE difference2 = After2 – Before2.

EXECUTE.

COMPUTE difference3 = After3 – Before3.

EXECUTE.

COMPUTE difference4 = After4 – Before4.

EXECUTE.

DESCRIPTIVES VARIABLES = difference1 difference2 difference3 difference4

/ STATISTICS = MEAN STDDEV MIN MAX.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
difference1	88	-2.57	5.04	1.3598	1.46377
difference2	88	-4.19	4.33	-.0644	1.97506
difference3	88	-1.10	3.76	2.1717	1.13214
difference4	88	-2.78	4.36	1.2178	1.42515
Valid N (listwise)	88				

NPAR TESTS

/WILCOXON = Before1 Before2 Before3 Before4 WITH After1 After2 After3 After4 (PAIRED)

/MISSING = ANALYSIS

Wilcoxon Signed Ranks Test

		Ranks		
		N	Mean Rank	Sum of Ranks
After1 - Before1	Negative Ranks	14 ^a	24.54	343.50
	Positive Ranks	71 ^b	46.64	3311.50
	Ties	3 ^c		
	Total	88		
After2 - Before2	Negative Ranks	45 ^d	43.14	1941.50
	Positive Ranks	41 ^e	43.89	1799.50
	Ties	2 ^f		
	Total	88		
After3 - Before3	Negative Ranks	4 ^g	12.88	51.50
	Positive Ranks	84 ^h	46.01	3864.50
	Ties	0 ⁱ		
	Total	88		
After4 - Before4	Negative Ranks	15 ^j	29.37	440.50
	Positive Ranks	73 ^k	47.61	3475.50
	Ties	0 ^l		
	Total	88		

- a. After1 < Before1
- b. After1 > Before1
- c. After1 = Before1
- d. After2 < Before2
- e. After2 > Before2
- f. After2 = Before2
- g. After3 < Before3
- h. After3 > Before3
- i. After3 = Before3
- j. After4 < Before4
- k. After4 > Before4
- l. After4 = Before4

Test Statistics^c

	After1 - Before1	After2 - Before2	After3 - Before3	After4 - Before4
Z	-6.504 ^a	-.306 ^b	-7.934 ^a	-6.315 ^a
Asymp. Sig (2-tailed)	.000	.760	.000	.000

- a. Based on negative ranks.
- b. Based on positive ranks.
- c. Wilcoxon Signed Ranks Test.

ภาคผนวก ค.
สรุปผลการวิเคราะห์ข้อสอบ

สรุปผลการวิเคราะห์ข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น ปีการศึกษา 2555
การสอบครั้งที่ 1 ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ตารางที่ ค - 1 ผลการวิเคราะห์คะแนนการสอบครั้งที่ 1 วิชาการพยาบาลเด็กและวัยรุ่น ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

จำนวนข้อสอบ	Max	Min	\bar{X}	S.D.	Median	Mode	Sk	Ku
75	56	25	39.852	6.182	40.000	40.296	0.041	2.932

Normal Curve : Skewness (ความเบ้) = 0 Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน < \bar{x} Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน > \bar{x}
 Kurtosis (ความโค้ง) = 3 Ku > 3 คือคะแนนกระจายแบบโค้งสูง (Leptokurtic) Ku < 3 คือคะแนนกระจายแบบลาด

ตารางที่ ค - 2 ผลการวิเคราะห์ความเที่ยงของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 1 ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความเที่ยงของข้อสอบ				Standard Error of Measurement			
KR20	KR21	Alpha	Split-half (Rtt)	SEM 20	SEM 21	SEM - Alp	SEMTT
0.603	0.518	0.603	0.637	3.895	4.291	3.895	3.727

KR20 เหมาะสำหรับข้อสอบ Heterogeneous

Cronbach เหมาะสำหรับข้อสอบที่มีคะแนน 1,0

KR21, Rtt เหมาะสำหรับข้อสอบ Homogeneous

ค่า KR21 จะน้อยกว่า KR20 เสมอและ Rtt จะได้ค่าความเที่ยงสูงสุด

ตารางที่ ค - 3 ผลการวิเคราะห์ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 1 ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับ				
ความยากง่ายของข้อสอบ		อำนาจจำแนกของข้อสอบ (r)		
Difficulty Index (p)	Delta	Discrimination Index	Biserial RBIS	Point - biserial RBIS
0.534	11.628	0.202	0.219	0.181

Delta 6 - 20 คือ ยากง่ายพอเหมาะ
 Delta ≥ 21 คือ ยากมาก

ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Biserial RBIS ใช้กับ Normal Curve
 ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Point - biserial RBIS ใช้กับ Curve เบ้

Delta ≤ 5 คือ ง่ายมาก

สรุปผลการวิเคราะห์ข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น ปีการศึกษา 2555
การสอบครั้งที่ 1 หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ตารางที่ ค - 4 ผลการวิเคราะห์คะแนนการสอบครั้งที่ 1 วิชาการพยาบาลเด็กและวัยรุ่น หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

จำนวนข้อสอบ	Max	Min	\bar{X}	S.D.	Median	Mode	Sk	Ku
70	55.000	22.000	38.148	6.329	38.000	37.704	0.035	2.910

Normal Curve : Skewness (ความเบ้) = 0 Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน < \bar{x} Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน > \bar{x}
Kurtosis (ความโค้ง) = 3 Ku > 3 คือคะแนนกระจายแบบโค้งสูง (Leptokurtic) Ku < 3 คือคะแนนกระจายแบบลาด

ตารางที่ ค - 5 ผลการวิเคราะห์ความเที่ยงของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 1 หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความเที่ยงของข้อสอบ				Standard Error of Measurement			
KR20	KR21	Alpha	Split-half (Rtt)	SEM 20	SEM 21	SEM - Alp	SEMTT
0.648	0.575	0.648	0.645	3.754	4.127	3.754	3.770

KR20 เหมาะสำหรับข้อสอบ Heterogeneous

Cronbach เหมาะสำหรับข้อสอบที่มีคะแนน 1,0

KR21, Rtt เหมาะสำหรับข้อสอบ Homogeneous

ค่า KR21 จะน้อยกว่า KR20 เสมอและ Rtt จะได้ค่าความเที่ยงสูงสุด

ตารางที่ ค - 6 ผลการวิเคราะห์ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 1 หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับ				
ความยากง่ายของข้อสอบ		อำนาจจำแนกของข้อสอบ (r)		
Difficulty Index (p)	Delta	Discrimination Index	Biserial RBIS	Point - biserial RBIS
0.547	11.801	0.223	0.240	0.198

Delta 6 - 20 คือ ยากพอเหมาะ
Delta ≥ 21 คือ ยากมาก

ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Biserial RBIS ใช้กับ Normal Curve
ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Point - biserial RBIS ใช้กับ Curve เบ้

Delta ≤ 5 คือ ง่ายมาก

สรุปผลการวิเคราะห์ข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น ปีการศึกษา 2555
การสอบครั้งที่ 2 ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ตารางที่ ค - 7 ผลการวิเคราะห์คะแนนการสอบครั้งที่ 2 วิชาการพยาบาลเด็กและวัยรุ่น ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

จำนวนข้อสอบ	Max	Min	\bar{X}	S.D.	Median	Mode	Sk	Ku
70	52.000	26.000	38.091	6.401	38.000	37.818	0.030	2.287

Normal Curve : Skewness (ความเบ้) = 0 Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน < \bar{x} Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน > \bar{x}
Kurtosis (ความโค้ง) = 3 Ku > 3 คือคะแนนกระจายแบบโค้งสูง (Leptokurtic) Ku < 3 คือคะแนนกระจายแบบลาด

ตารางที่ ค - 8 ผลการวิเคราะห์ความเที่ยงของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 2 ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความเที่ยงของข้อสอบ				Standard Error of Measurement			
KR20	KR21	Alpha	Split-half (Rtt)	SEM 20	SEM 21	SEM - Alp	SEMTT
0.657	0.585	0.657	0.716	3.750	4.126	3.750	3.411

KR20 เหมาะสำหรับข้อสอบ Heterogeneous

Cronbach เหมาะสำหรับข้อสอบที่มีคะแนน 1,0

KR21, Rtt เหมาะสำหรับข้อสอบ Homogeneous

ค่า KR21 จะน้อยกว่า KR20 เสมอและ Rtt จะได้ค่าความเที่ยงสูงสุด

ตารางที่ ค - 9 ผลการวิเคราะห์ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 2 ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับ				
ความยากง่ายของข้อสอบ		อำนาจจำแนกของข้อสอบ (r)		
Difficulty Index (p)	Delta	Discrimination Index	Biserial RBIS	Point - biserial RBIS
0.547	11.319	0.230	0.245	0.198

Delta 6 - 20 คือ ยากพอเหมาะ
Delta \geq 21 คือ ยากมาก

ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Biserial RBIS ใช้กับ Normal Curve
ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Point - biserial RBIS ใช้กับ Curve เบ้

Delta \leq 5 คือ ง่ายมาก

สรุปผลการวิเคราะห์ข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น ปีการศึกษา 2555
การสอบครั้งที่ 2 หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ตารางที่ ค - 10 ผลการวิเคราะห์คะแนนการสอบครั้งที่ 2 วิชาการพยาบาลเด็กและวัยรุ่น หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

จำนวนข้อสอบ	Max	Min	\bar{X}	S.D.	Median	Mode	Sk	Ku
64	50.000	22.000	34.784	6.719	35.000	35.432	0.018	2.403

Normal Curve : Skewness (ความเบ้) = 0 Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน < \bar{x} Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน > \bar{x}
Kurtosis (ความโค้ง) = 3 Ku > 3 คือคะแนนกระจายแบบโค้งสูง (Leptokurtic) Ku < 3 คือคะแนนกระจายแบบลาด

ตารางที่ ค - 11 ผลการวิเคราะห์ความเที่ยงของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 2 หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความเที่ยงของข้อสอบ				Standard Error of Measurement			
KR20	KR21	Alpha	Split-half (Rtt)	SEM 20	SEM 21	SEM - Alp	SEMTT
0.717	0.659	0.717	0.756	3.572	3.926	3.572	3.320

KR20 เหมาะสำหรับข้อสอบ Heterogeneous

Cronbach เหมาะสำหรับข้อสอบที่มีคะแนน 1,0

KR21, Rtt เหมาะสำหรับข้อสอบ Homogeneous

ค่า KR21 จะน้อยกว่า KR20 เสมอและ Rtt จะได้ค่าความเที่ยงสูงสุด

ตารางที่ ค - 12 ผลการวิเคราะห์ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 2 หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับ				
ความยากง่ายของข้อสอบ		อำนาจจำแนกของข้อสอบ (r)		
Difficulty Index (p)	Delta	Discrimination Index	Biserial RBIS	Point - biserial RBIS
0.545	11.997	0.268	0.282	0.227

Delta 6 - 20 คือ ยากง่ายพอเหมาะ
Delta ≥ 21 คือ ยากมาก

ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Biserial RBIS ใช้กับ Normal Curve
ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Point - biserial RBIS ใช้กับ Curve เบ้

Delta ≤ 5 คือ ง่ายมาก

สรุปผลการวิเคราะห์ข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น ปีการศึกษา 2555
การสอบครั้งที่ 3 ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ตารางที่ ค - 13 ผลการวิเคราะห์คะแนนการสอบครั้งที่ 3 วิชาการพยาบาลเด็กและวัยรุ่น ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

จำนวนข้อสอบ	Max	Min	\bar{X}	S.D.	Median	Mode	Sk	Ku
58	47.000	22.000	30.784	4.759	30.500	29.932	0.259	3.527

Normal Curve : Skewness (ความเบ้) = 0 Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน < \bar{x} Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน > \bar{x}
 Kurtosis (ความโค้ง) = 3 Ku > 3 คือคะแนนกระจายแบบโค้งสูง (Leptokurtic) Ku < 3 คือคะแนนกระจายแบบลาด

ตารางที่ ค - 14 ผลการวิเคราะห์ความเที่ยงของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 3 ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความเที่ยงของข้อสอบ				Standard Error of Measurement			
KR20	KR21	Alpha	Split-half (Rtt)	SEM 20	SEM 21	SEM - Alp	SEMTT
0.498	0.369	0.498	0.605	3.371	3.782	3.371	2.990

KR20 เหมาะสำหรับข้อสอบ Heterogeneous

Cronbach เหมาะสำหรับข้อสอบที่มีคะแนน 1,0

KR21, Rtt เหมาะสำหรับข้อสอบ Homogeneous

ค่า KR21 จะน้อยกว่า KR20 เสมอและ Rtt จะได้ค่าความเที่ยงสูงสุด

ตารางที่ ค - 15 ผลการวิเคราะห์ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 3 ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับ				
ความยากง่ายของข้อสอบ		อำนาจจำแนกของข้อสอบ (r)		
Difficulty Index (p)	Delta	Discrimination Index	Biserial RBIS	Point - biserial RBIS
0.534	11.567	0.203	0.214	0.179

Delta 6 - 20 คือ ยากง่ายพอเหมาะ
 Delta ≥ 21 คือ ยากมาก

ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Biserial RBIS ใช้กับ Normal Curve
 ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Point - biserial RBIS ใช้กับ Curve เบ้

Delta ≤ 5 คือ ง่ายมาก

สรุปผลการวิเคราะห์ข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น ปีการศึกษา 2555
การสอบครั้งที่ 3 หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ตารางที่ ค - 16 ผลการวิเคราะห์คะแนนการสอบครั้งที่ 3 วิชาการพยาบาลเด็กและวัยรุ่น หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

จำนวนข้อสอบ	Max	Min	\bar{X}	S.D.	Median	Mode	Sk	Ku
55	45.000	22.000	30.307	4.521	30.000	29.386	0.261	3.105

Normal Curve : Skewness (ความเบ้) = 0 Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน < \bar{x} Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน > \bar{x}
Kurtosis (ความโค้ง) = 3 Ku > 3 คือคะแนนกระจายแบบโค้งสูง (Leptokurtic) Ku < 3 คือคะแนนกระจายแบบลาด

ตารางที่ ค - 17 ผลการวิเคราะห์ความเที่ยงของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 3 หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความเที่ยงของข้อสอบ				Standard Error of Measurement			
KR20	KR21	Alpha	Split-half (Rtt)	SEM 20	SEM 21	SEM - Alp	SEMTT
0.460	0.340	0.460	0.566	3.322	3.672	3.322	2.977

KR20 เหมาะสำหรับข้อสอบ Heterogeneous

Cronbach เหมาะสำหรับข้อสอบที่มีคะแนน 1,0

KR21, Rtt เหมาะสำหรับข้อสอบ Homogeneous

ค่า KR21 จะน้อยกว่า KR20 เสมอและ Rtt จะได้ค่าความเที่ยงสูงสุด

ตารางที่ ค - 18 ผลการวิเคราะห์ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับวิชาการพยาบาลเด็กและวัยรุ่นในการสอบครั้งที่ 3 หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับ				
ความยากง่ายของข้อสอบ		อำนาจจำแนกของข้อสอบ (r)		
Difficulty Index (p)	Delta	Discrimination Index	Biserial RBIS	Point - biserial RBIS
0.556	11.573	0.203	0.214	0.180

Delta 6 - 20 คือ ยากง่ายพอเหมาะ
Delta ≥ 21 คือ ยากมาก

ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Biserial RBIS ใช้กับ Normal Curve
ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Point - biserial RBIS ใช้กับ Curve เบ้

Delta ≤ 5 คือ ง่ายมาก

สรุปผลการวิเคราะห์ข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น ปีการศึกษา 2555

การสอบปลายภาค ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ตารางที่ ค - 19 ผลการวิเคราะห์คะแนนการสอบปลายภาควิชาการพยาบาลเด็กและวัยรุ่น ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

จำนวนข้อสอบ	Max	Min	\bar{X}	S.D.	Median	Mode	Sk	Ku
75	58.000	23.000	42.625	6.515	43.000	43.750	0.322	3.949

Normal Curve : Skewness (ความเบ้) = 0 Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน < \bar{x} Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน > \bar{x}
 Kurtosis (ความโค้ง) = 3 Ku > 3 คือคะแนนกระจายแบบโค้งสูง (Leptokurtic) Ku < 3 คือคะแนนกระจายแบบลาด

ตารางที่ ค - 20 ผลการวิเคราะห์ความเที่ยงของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นในการสอบปลายภาค ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความเที่ยงของข้อสอบ				Standard Error of Measurement			
KR20	KR21	Alpha	Split-half (Rtt)	SEM 20	SEM 21	SEM - Alp	SEMTT
0.662	0.574	0.662	0.745	3.787	4.251	3.787	3.290

KR20 เหมาะสำหรับข้อสอบ Heterogeneous

Cronbach เหมาะสำหรับข้อสอบที่มีคะแนน 1,0

KR21, Rtt เหมาะสำหรับข้อสอบ Homogeneous

ค่า KR21 จะน้อยกว่า KR20 เสมอและ Rtt จะได้ค่าความเที่ยงสูงสุด

ตารางที่ ค - 21 ผลการวิเคราะห์ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับวิชาการพยาบาลเด็กและวัยรุ่นในการสอบปลายภาค ก่อนตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับ				
ความยากง่ายของข้อสอบ		อำนาจจำแนกของข้อสอบ (r)		
Difficulty Index (p)	Delta	Discrimination Index	Biserial RBIS	Point - biserial RBIS
0.564	11.205	0.206	0.231	0.191

Delta 6 - 20 คือ ยากพอเหมาะ
Delta ≥ 21 คือ ยากมาก

ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Biserial RBIS ใช้กับ Normal Curve
ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Point - biserial RBIS ใช้กับ Curve เบ้

Delta ≤ 5 คือ ง่ายมาก

สรุปผลการวิเคราะห์ข้อสอบวิชาการพยาบาลเด็กและวัยรุ่น ปีการศึกษา 2555

การสอบปลายภาค หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ตารางที่ ค - 22 ผลการวิเคราะห์คะแนนการสอบปลายภาควิชาการพยาบาลเด็กและวัยรุ่น หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

จำนวนข้อสอบ	Max	Min	\bar{X}	S.D.	Median	Mode	Sk	Ku
71	58.000	21.000	41.216	6.600	41.000	40.568	0.245	3.789

Normal Curve : Skewness (ความเบ้) = 0 Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน < \bar{x} Sk = (เบ้ขวา) คือ นศ. > 50 % ได้คะแนน > \bar{x}
 Kurtosis (ความโค้ง) = 3 Ku > 3 คือคะแนนกระจายแบบโค้งสูง (Leptokurtic) Ku < 3 คือคะแนนกระจายแบบลาด

ตารางที่ ค - 23 ผลการวิเคราะห์ความเที่ยงของข้อสอบวิชาการพยาบาลเด็กและวัยรุ่นในการสอบปลายภาค หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความเที่ยงของข้อสอบ				Standard Error of Measurement			
KR20	KR21	Alpha	Split-half (Rtt)	SEM 20	SEM 21	SEM - Alp	SEMTT
0.691	0.612	0.691	0.642	3.667	4.113	3.667	3.948

KR20 เหมาะสำหรับข้อสอบ Heterogeneous

Cronbach เหมาะสำหรับข้อสอบที่มีคะแนน 1,0

KR21, Rtt เหมาะสำหรับข้อสอบ Homogeneous

ค่า KR21 จะน้อยกว่า KR20 เสมอและ Rtt จะได้ค่าความเที่ยงสูงสุด

ตารางที่ ค - 24 ผลการวิเคราะห์ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับวิชาการพยาบาลเด็กและวัยรุ่นในการสอบปลายภาค หลังตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก

ความยากง่ายและอำนาจจำแนกของข้อสอบทั้งฉบับ				
ความยากง่ายของข้อสอบ		อำนาจจำแนกของข้อสอบ (r)		
Difficulty Index (p)	Delta	Discrimination Index	Biserial RBIS	Point - biserial RBIS
0.575	11.544	0.223	0.247	0.205

Delta 6 - 20 คือ ยากง่ายพอเหมาะ

ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Biserial RBIS ใช้กับ Normal Curve

Delta ≥ 21 คือ ยากมาก

ค่าอำนาจจำแนก ส.ป.ส.สหสัมพันธ์ Point - biserial RBIS ใช้กับ Curve เบ้

Delta ≤ 5 คือ ง่ายมาก

ประวัติผู้วิจัย

หัวหน้าโครงการวิจัย

1. ชื่อ – นามสกุล นางสาวพัชมน อันโต

Miss Patchamon Onto

2. ตำแหน่งปัจจุบัน อาจารย์ประจำสาขาการพยาบาลเด็กและวัยรุ่น คณะพยาบาลศาสตร์

มหาวิทยาลัยราชภัฏสวนดุสิต

3. หน่วยงานและสถานที่อยู่ที่ติดต่อได้สะดวก

คณะพยาบาลศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิต

204/3 อาคารเฉลิมพระเกียรติ 50 พรรษามหาชัฎราลงกรณ ถนนสีรินธร แขวงบางบำหรุ

เขตบางพลัด กทม. 10700

หมายเลขโทรศัพท์ที่ทำงาน 02 – 4239467 โทรสาร 02 – 4239466

มือถือ 081 – 6517476

จดหมายอิเล็กทรอนิกส์ (e-mail): aunchula@gmail.com

4. ประวัติการศึกษา

คุณวุฒิ	ปีที่จบ	สถานศึกษา
พยาบาลศาสตรมหาบัณฑิต (บริหารการพยาบาล)	พ.ศ. 2547	คณะพยาบาลศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
วิทยาศาสตรบัณฑิต (พยาบาลและผดุงครรภ์) เกียรตินิยมอันดับสอง	พ.ศ. 2529	คณะพยาบาลศาสตร์ มหาวิทยาลัยสงขลานครินทร์
หลักสูตร MINI MBA. IN HEALTH	พ.ศ. 2542	คณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

5. สาขาวิชาการที่มีความชำนาญพิเศษ (อาจจะแตกต่างจากวุฒิการศึกษา)

การพยาบาลเด็กและวัยรุ่น การบริหารการพยาบาล

ผู้ร่วมวิจัยคนที่ 1

1. ชื่อ - นามสกุล พ.ท.หญิง วชิรา ไกรถิ่น

Lt. Col. VachiraKraitin

2. ตำแหน่งปัจจุบัน อาจารย์ประจำสาขาการพยาบาลเด็กและวัยรุ่น คณะพยาบาลศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิต

3. หน่วยงานและสถานที่อยู่ที่ติดต่อได้สะดวก

คณะพยาบาลศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิต

204/3 อาคารเฉลิมพระเกียรติ 50 พรรษามหาชัราลงกรณ ถนนสีรินธร แขวงบางบำหรุ เขตบางพลัด กทม. 10700

หมายเลขโทรศัพท์ที่ทำงาน 02 – 4239467 โทรสาร 02 – 4239466

มือถือ 081 – 8177228

จดหมายอิเล็กทรอนิกส์ (e-mail): kraitin@yahoo.com

4. ประวัติการศึกษา

คุณวุฒิ	ปีที่จบ	สถานศึกษา
Master of Science in Nursing (Nursing of Children)	พ.ศ. 2532	St. Louis University, Missouri, USA
Bachelor of Science in Nursing	พ.ศ. 2530	St. Louis University, Missouri, USA
ประกาศนียบัตรพยาบาล ฉุกเฉิน และอนามัย	พ.ศ. 2524	วิทยาลัยพยาบาลกองทัพบก

5. สาขาวิชาการที่มีความชำนาญพิเศษ (อาจจะแตกต่างจากวุฒิการศึกษา)

การพยาบาลเด็กและวัยรุ่น

ผู้ร่วมวิจัยคนที่ 2

1. ชื่อ - นามสกุล นางสาวสุนีย์ ชื่นจันทร์

Miss Sunee Chuenjan

2. ตำแหน่งปัจจุบัน อาจารย์ประจำสาขาการพยาบาลเด็กและวัยรุ่น คณะพยาบาลศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิต

3. หน่วยงานและสถานที่อยู่ที่ติดต่อได้สะดวก

คณะพยาบาลศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิต

204/3 อาคารเฉลิมพระเกียรติ 50 พรรษามหาวิจิตรถาวร ถนนสีรินธร แขวงบางบำหรุ เขตบางพลัด กทม. 10700

หมายเลขโทรศัพท์ที่ทำงาน 02 – 4239467 โทรสาร 02 – 4239466

มือถือ 082 – 4402983

จดหมายอิเล็กทรอนิกส์ (e-mail): schuenjan@live.com

4. ประวัติการศึกษา

คุณวุฒิ	ปีที่จบ	สถานศึกษา
พยาบาลศาสตรมหาบัณฑิต (บริหารการพยาบาล)	พ.ศ. 2536	คณะพยาบาลศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
พยาบาลศาสตรบัณฑิต	พ.ศ. 2529	วิทยาลัยพยาบาลสภากาชาดไทย

5. สาขาวิชาการที่มีความชำนาญพิเศษ (อาจจะแตกต่างจากวุฒิการศึกษา)

การพยาบาลเด็กและวัยรุ่น การบริหารการพยาบาล