

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

การวิจัยครั้งนี้เป็นการศึกษาย้อนหลังจากชุดข้อสอบและคะแนนการสอบรายวิชาการพยาบาลเด็กและวัยรุ่นภาคการเรียนที่ 1 ปีการศึกษา 2555 มีวัตถุประสงค์เพื่อศึกษา 1) คุณภาพของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่น 2) การเปลี่ยนแปลงคะแนนการสอบของนักศึกษาในรายวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก 3) การเปลี่ยนแปลงลำดับคะแนนการสอบของนักศึกษาในวิชาการพยาบาลเด็กและวัยรุ่นหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก 4) การเปลี่ยนแปลงค่าความเชื่อมั่นของข้อสอบรายวิชาการพยาบาลเด็กและวัยรุ่นทั้งฉบับหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก ประชากรและกลุ่มตัวอย่างคือ ข้อสอบและคะแนนการสอบของนักศึกษาหลักสูตรพยาบาลศาสตรบัณฑิตชั้นปีที่ 2 ที่ลงทะเบียนเรียนและเข้าสอบวิชาการพยาบาลเด็กและวัยรุ่นในภาคการเรียนที่ 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

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