

หัวข้อวิจัย	ผลของการตัดข้อสอบที่มีอำนาจจำแนกติดลบออกต่อคะแนนการสอบและความเชื่อมั่นของข้อสอบทั้งฉบับ: รายวิชาทฤษฎีและจริยธรรมวิชาชีพ
ผู้ดำเนินการวิจัย	นางนิภา ลีสุคนธ์ นางสาวจิตราพร จันทรวงศ์
ที่ปรึกษา	ผศ.ดร.ประกาย จิโรจน์กุล
หน่วยงาน	คณะพยาบาลศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิต
ปี พ.ศ.	2558

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาคุณภาพของข้อสอบ การเปลี่ยนแปลงของคะแนนเฉลี่ย และอันดับของการสอบของนักศึกษารายบุคคล ก่อน และหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก การเปลี่ยนแปลงของค่าความเชื่อมั่นของข้อสอบ ก่อน และหลังการตัดข้อสอบที่มีค่าอำนาจจำแนกติดลบออก ในรายวิชาทฤษฎีและจริยธรรมวิชาชีพ โดยประชากร และกลุ่มตัวอย่าง ได้แก่ ข้อสอบ และคะแนนผลการสอบรายวิชารายวิชาทฤษฎีและจริยธรรมวิชาชีพ ของนักศึกษาพยาบาลศาสตร์ ชั้นปีที่ 4 จำนวน 60 คน ที่ลงทะเบียนเรียน รายวิชาทฤษฎีและจริยธรรมวิชาชีพ ในภาคการศึกษาที่ 1 ปีการศึกษา 2555 โดยใช้โปรแกรมการวิเคราะห์ข้อสอบรายข้อ (Items Analysis) สถิติที่ใช้ในงานวิจัยประกอบด้วย Paired t-test (one-tailed) และ Sign Ranked test ผลการวิเคราะห์ข้อมูลจากชุดข้อสอบแบบเลือกตอบ (Multiple choice) 4 ตัวเลือก จำนวน 2 ฉบับ เป็นชุดข้อสอบกลางภาค จำนวน 40 ข้อและปลายภาค 40 ข้อ พบว่า

1) ด้านคุณภาพข้อสอบ เมื่อพิจารณาจากดัชนีความยากง่าย (P) พบว่า ชุดข้อสอบกลางภาค และชุดข้อสอบปลายภาค มีข้อสอบที่มีค่าความยากง่ายในเกณฑ์ที่เหมาะสม (P อยู่ระหว่าง 0.20 – 0.80) จำนวนรวม 30, 25 ข้อ คิดเป็นร้อยละ 75, 65 ตามลำดับ ด้านค่าอำนาจจำแนก (r) พบว่า ชุดข้อสอบกลางภาค และชุดข้อสอบปลายภาค มีข้อสอบที่มีค่าอำนาจจำแนกในเกณฑ์ที่เหมาะสม ($r \geq 0.20$) จำนวนรวม 23, 19 ข้อ คิดเป็นร้อยละ 57.5 และร้อยละ 47.5 ตามลำดับ ทั้งนี้ ข้อสอบที่มีค่าอำนาจจำแนกติดลบ ($r = -1.00 - 0.00$) ในชุดข้อสอบกลางภาค และชุดข้อสอบปลายภาค มีจำนวนชุดละ 1 ข้อ คิดเป็นร้อยละ 2.5 ส่วนคุณภาพของข้อสอบทั้งฉบับ พบว่า ข้อสอบกลางภาคมีค่าความเชื่อมั่น 0.48 และเมื่อตัดข้อสอบที่มีค่า r ติดลบออก ทำให้ค่าความเชื่อมั่นเพิ่มขึ้นเป็น 0.51 ส่วนข้อสอบปลายภาค มีค่าความเชื่อมั่น = 0.40 และเมื่อตัดข้อสอบที่มีค่า r ติดลบออก ทำให้ค่าความเชื่อมั่นเพิ่มขึ้นเป็น 0.43

2) ผลการเปรียบเทียบคะแนนเฉลี่ยที่นักศึกษาได้รับ ในรายวิชาทฤษฎีและจริยธรรมวิชาชีพ จากข้อสอบทั้งชุดกลางภาค และชุดปลายภาค ก่อน และหลังการตัดข้อสอบที่มีค่า r ติดลบ

ออก โดยสถิติ Paired t-test (one-tailed) พบว่า นักศึกษามีคะแนนเฉลี่ยสูงขึ้นอย่างมีนัยสำคัญทางสถิติ ที่ $\alpha = 0.05$ ($t=0.00$, $df = 59$, $P \text{ value} < 0.005$)

3) ผลการทดสอบการเปลี่ยนแปลงของลำดับคะแนนที่นักศึกษาได้รับ ในรายวิชากฎหมาย และจริยธรรมวิชาชีพ ก่อน และหลังการตัดข้อสอบที่มีค่า r ติดลบออก โดยสถิติ Signed Rank Test พบว่า ลำดับคะแนนที่นักศึกษาได้รับมีการเปลี่ยนแปลงลำดับที่อย่างมีนัยสำคัญทางสถิติ ที่ $\alpha = 0.05$ ($P \text{ value} < 0.01$) ในข้อสอบทั้ง 2 ชุด

คำสำคัญ: คุณภาพข้อสอบ ค่าอำนาจจำแนกติดลบ คะแนนสอบ ค่าความเชื่อมั่นของข้อสอบทั้งฉบับ

Research Title	The Effects on the Test Scores and Reliability after excluding items with a Negative Discrimination Power: The Laws and Professional Ethics Subject
Researcher	Mrs.Nipa Leesukol Miss. Jittraporn Jandharakhul
Research Consultants	Ass.Prof. Dr. Pragai Jirojjanakul
Organization	Faculty of Nursing, Suan Dusit Rajabhat University
Year	2015

The objectives of this research were 1) to describe the quality of the test in terms of each item's index of difficulty (P), index of discrimination (r) and its reliability. 2) to compare the test scores and students rank before and after excluding the items with negative discrimination power. A 40 item multiple choice midterm and final test were used in this study. The participants were 60 fourth year nursing students who were registered in the Laws and Professional Ethics Subject during the first semester of the academic year 2012. The item analysis was used to analyze the test items and scores. The test items were analyzed for the index of difficulty, index of discrimination and reliability by a computer program. The statistics used in this research included the Paired t-test (one – tailed) and the Sign Ranked test.

The results found that:

1) Quality of the test: there were 30 midterm and 25 final test items with a proper difficulty index ($P = 0.20 - 0.80$). This accounted for 75% and 65 % respectively. Twenty-three midterm test items (57.5%) and 19 final test items (47.5%) had an index of discrimination (r) with a proper range of ($r \geq 0.2$), while 1 item (2.5%) of each test had a negative discrimination power. ($r = -1.00 - 0.00$) The reliability of the midterm test and the final test were 0.48 and 0.40 respectively and it was increased to 0.51 and 0.43 after excluding items with negative discrimination power.

2) Comparing the mean scores before and after excluding the item with negative discrimination power, using a paired t-test (one-tailed) statistic, found that students had obtained significant higher scores at $\alpha = 0.05$ ($t=0.00$, $df=59$, p value < 0.005 in both tests.

3)The Signed Rank Test showed that the student's ranks were significantly changed in both tests with $\alpha = 0.05$, p value < 0.01 after excluding the item with negative discrimination power.

Key words: quality of test, negative discrimination power, test score, reliability of the test.