

ความเชื่อมั่นของแบบวัดที่มีเป้าหมายของการวัดแตกต่างกัน
กรณีศึกษา: แบบประเมินผลการเรียนการสอน มหาวิทยาลัยกรุงเทพ
The Reliability of Multi-target Assessment Tryouts: A Case Study
for Teaching and Learning Assessment at Bangkok University

วธนา สุนทรชัย¹

Wathna Soonthorndhai

วนิดา คูชัยสิทธิ์²

Wanida Kuchaisit

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาความเชื่อมั่นของการทดลองใช้แบบวัดที่คำนวณจากกลุ่มย่อยและกลุ่มรวม และศึกษาความสัมพันธ์ระหว่างจำนวนผู้ตอบกับค่าความเชื่อมั่นของแบบวัด กลุ่มตัวอย่าง คือ นักศึกษา มหาวิทยาลัยกรุงเทพ จำนวน 6 กลุ่มๆ ละ 5 ถึง 15 คน รวมเป็น 57 คน สถิติที่ใช้คือ ค่าสหสัมพันธ์ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าความเชื่อมั่น สถิติ Fisher's Z และสัมประสิทธิ์สหสัมพันธ์เพียร์สัน

ผลการวิจัยพบว่า ความเชื่อมั่นของแบบวัดที่คำนวณจากกลุ่มย่อยมีค่าเท่ากับค่าความเชื่อมั่นของแบบวัดที่คำนวณจากกลุ่มรวม และจำนวนผู้ตอบกับค่าความเชื่อมั่น ไม่มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ดังนั้น ในขั้นตอนของการทดลองใช้แบบวัดใดๆ กลุ่มตัวอย่างที่เป็นตัวแทนของการศึกษาจะเป็นนักศึกษาที่เรียนกับผู้สอนที่คนก็ได้หรือที่วิชาก็ได้ มาเป็นตัวแทนของการศึกษา

คำสำคัญ: ความเชื่อมั่น การทดลองใช้

ABSTRACT

This study investigates the reliability of the multi-target assessment tryouts calculated from both main group and subgroups, and the relationships between the number of respondents and the reliability of the assessments. The subjects of this study are 57 Bangkok University students, categorized into 6 groups, with 5-15 students in each group. The statistics used in this study are corrected item-total correlation, Cronbach's alpha if item deleted, mean, standard deviation, reliability, Fisher's Z, and Pearson's correlation coefficient.

¹ รองศาสตราจารย์ประจำคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยกรุงเทพ

School of Science and Technology, Bangkok University, e-mail: wathna.s@bu.ac.th

² สำนักวิชาการ มหาวิทยาลัยกรุงเทพ

Academic Affairs Office, Bangkok University, e-mail: wanida.k@bu.ac.th

The results show that the reliability calculated from the subgroups is at the same level with that from the main groups. Moreover, the number of respondents and the reliability are not statistically significant at the level of 0.05. Therefore, to try out any assessment, the subjects may be recruited from students in any class taught by any number of instructors.

Keywords: Reliability, Tryouts

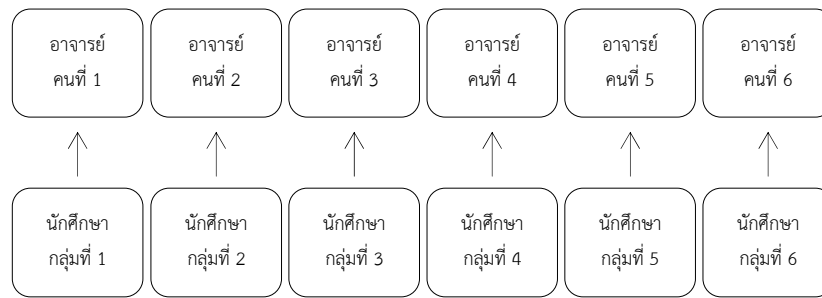
บทนำ

การทดลองใช้ (Tryout) เป็นขั้นตอนหนึ่งของกระบวนการพัฒนาแบบวัด เพื่อตรวจสอบคุณภาพของเครื่องมือวิจัย โดยผู้ตอบแบบวัดต้องเป็นกลุ่มตัวอย่างที่เป็นตัวแทนของประชากรที่ต้องการศึกษาเรื่องนั้นๆ และตอบคำถามที่มีเป้าหมายของการวัดเป็นเรื่องเดียวกัน โดยทั่วไปเป้าหมายของการวัดโดยใช้แบบทดสอบหรือแบบสอบถามเรื่องใดเรื่องหนึ่งนั้น ไม่ว่าจะมีคนตอบกี่กลุ่มก็ตาม ควรมีเป้าหมายของการวัดค่าเป็นเรื่องเดียวกัน เช่น เป้าหมายการวัดของแบบทดสอบวิชาหนึ่ง คือ ความรู้ในบริบทตามที่กำหนด ผู้ตอบทุกคนจะต้องตอบเนื้อหาเดียวกันที่ปรากฏอยู่ในข้อสอบนั้นๆ และถือว่าบริบทที่กำหนด คือ เป้าหมายของการวัดค่าในที่นี้ หรือถ้าเป็นแบบวัดความคิดเห็นหรือแบบวัดเจตคติต่อเรื่องใดเรื่องหนึ่ง เป้าหมายของการวัด คือ ความคิดเห็นหรือเจตคติต่อเรื่องที่ ต้องการวัดตามที่กำหนด ผู้ตอบทุกคนจะต้องตอบเนื้อหาในเรื่องเดียวกันที่ปรากฏอยู่ในแบบวัดนั้นๆ และถือว่าเป้าหมายของการวัดค่าในที่นี้ คือ เนื้อหาในเรื่องที่กำหนด เป็นต้น

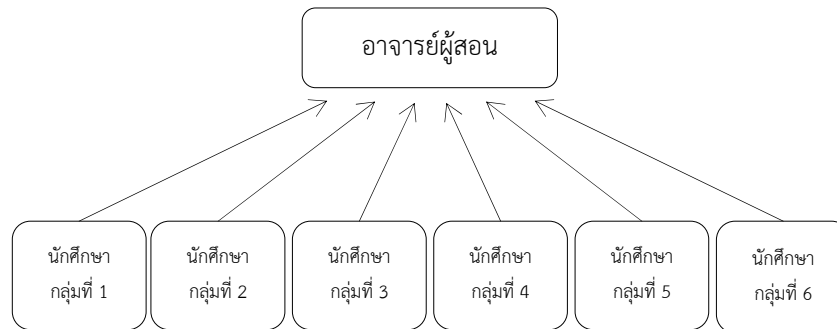
แบบวัดที่ต้องการศึกษาในงานวิจัยนี้ คือ แบบประเมินผลการเรียนการสอน ซึ่งมีข้อมูลหลักสามส่วน คือ ข้อมูลเกี่ยวกับวิชาที่เรียน ข้อมูลเกี่ยวกับอาจารย์ผู้สอน และข้อมูลเกี่ยวกับการประเมินตนเองของนักศึกษาเกี่ยวกับผลลัพธ์ที่ได้จากการเรียนในการนำแบบประเมินผลการเรียนการสอนนี้ไปใช้จริง

นักศึกษาจะประเมินรายวิชา ประเมินผู้สอนในรายวิชานั้น และประเมินตนเองจากการเรียนวิชานั้น ซึ่งมีค่าความเชื่อมั่นส่วนละ 1 ค่า จะได้ค่าความเชื่อมั่น 3 ค่า คือ ความเชื่อมั่นของวิชาเรียน ความเชื่อมั่นของผู้สอน และความเชื่อมั่นของนักศึกษา แต่ในขั้นตอนการทดลองใช้ของงานวิจัยครั้งนี้ มีกลุ่มตัวอย่างซึ่งสุ่มมาจากนักศึกษาที่เรียนวิชาต่างๆ 6 วิชา จึงมีอาจารย์ผู้สอน 6 คน โดยสอนวิชาที่แตกต่างกัน ทำให้สถานการณ์ของการทดลองใช้และการนำไปใช้จริงมีความแตกต่างกัน ผู้วิจัยจึงแบ่งนักศึกษาตามรายวิชาและผู้สอนที่แตกต่างกันออกเป็น 6 กลุ่ม โดยนักศึกษากลุ่มที่ 1 ประเมินวิชาที่ 1 ผู้สอนคนที่ 1 และประเมินตนเอง ส่วนกลุ่มที่ 2 ประเมินวิชาที่ 2 ผู้สอนคนที่ 2 และประเมินตนเองตามลำดับจนครบทั้ง 6 กลุ่ม โดยผลการคำนวณความเชื่อมั่นของแต่ละกลุ่มจะได้ความเชื่อมั่น 3 ส่วน ซึ่งเป็นเป้าหมายของการประเมินของแต่ละ ดังนั้น จึงมีค่าในการทดลองมี 6 กลุ่มๆ ละ 3 ค่า รวมทั้งสิ้น 18 ค่า

เพื่อให้เกิดความเข้าใจที่ตรงกัน จึงได้นำเสนอแผนภาพจำลองความคิด เปรียบเทียบการประเมินในระหว่างขั้นตอนการทดลองใช้แบบวัดและสถานการณ์ใช้จริง โดยนำเสนอแผนภาพเฉพาะในส่วนของการประเมินผู้สอน ดังแผนภาพที่ 1



แผนภาพจำลองความคิดของการทดลองใช้แบบวัด



แผนภาพจำลองความคิดของสถานการณ์ใช้จริง

ภาพที่ 1 แผนภาพจำลองความคิด

ตามแผนภาพจำลองความคิดข้างต้น ในขั้นตอนของการทดลองใช้นั้น นักศึกษาจำนวน 6 กลุ่ม ประเมินอาจารย์ 6 คน กลุ่มละหนึ่งคน แต่ในสถานการณ์ใช้จริงนั้น นักศึกษาทุกคนประเมินอาจารย์คนเดียวกัน ซึ่งสิ่งที่ชวนให้สงสัย คือการเลือกกลุ่มตัวอย่างมาทดลองใช้แบบในสถานการณ์ที่แตกต่างกัน คือในการทดลองและการนำไปใช้จริงนั้น จะยังทำให้การนำแบบวัดนี้ไปใช้มีความแตกต่างกันหรือไม่อย่างไร การวิจัยในครั้งนี้จึงเป็นการตรวจสอบว่า ค่าความเชื่อมั่นที่คำนวณ 6 ครั้ง จากนักศึกษา 6 กลุ่ม ที่มีเป้าหมายแตกต่างกันตามรายวิชาและผู้สอนนั้น มีค่าเท่ากับความเชื่อมั่นที่คำนวณครั้งเดียวจากกลุ่มรวมหรือไม่ และมีความสัมพันธ์ระหว่างจำนวนผู้ตอบกับค่าความเชื่อมั่นของแบบวัดหรือไม่ อย่างไร ผลการวิจัยจะมีประโยชน์สำหรับนักวิจัยรุ่นใหม่ในการตัดสินใจเลือกกลุ่มตัวอย่างเพื่อคำนวณความเชื่อมั่นของแบบวัดที่สร้างขึ้น ดังนั้น ถ้าพบว่า ความเชื่อมั่นของกลุ่มย่อยไม่แตกต่างจากความเชื่อมั่นของกลุ่มรวมแล้ว ผู้วิจัยก็สามารถเลือกใช้กลุ่มทดลองที่มาจากกลุ่มใดกลุ่มหนึ่งหรือมาจากหลายกลุ่มรวมกันก็ได้ แต่ถ้าความเชื่อมั่นของกลุ่มย่อยแตกต่างจากความเชื่อมั่นของกลุ่มรวมแล้ว ผู้วิจัยจะต้องใช้กลุ่มตัวอย่างที่ผู้ตอบมาจากผู้สอนคนเดียวกัน

เท่านั้น นอกจากนี้ ถ้าผลการวิจัยพบว่า จำนวนผู้ตอบแบบวัดกับค่าความเชื่อมั่นมีความสัมพันธ์ทางบวก ผู้วิจัยสามารถเพิ่มค่าความเชื่อมั่นของแบบวัดให้สูงขึ้นได้โดยการเพิ่มจำนวนของผู้ตอบแบบวัด จึงเป็นที่มาของวัตถุประสงค์และสมมติฐานการวิจัยดังต่อไปนี้

วัตถุประสงค์

การวิจัยนี้มีวัตถุประสงค์ดังนี้

1. เพื่อศึกษาความเชื่อมั่นของแบบวัดที่คำนวณจากการทดลองใช้ในกลุ่มย่อยและกลุ่มรวม
2. เพื่อศึกษาความสัมพันธ์ระหว่างจำนวนผู้ตอบกับค่าความเชื่อมั่นของแบบวัด

สมมติฐานทางวิจัย

การวิจัยนี้มีสมมติฐานทางวิจัยดังต่อไปนี้

1. ความเชื่อมั่นของแบบวัดที่คำนวณจากกลุ่มย่อยมีค่าเท่ากับค่าความเชื่อมั่นของแบบวัดที่คำนวณจากกลุ่มรวม
2. จำนวนผู้ตอบมีความสัมพันธ์ทางบวกกับค่าความเชื่อมั่นของแบบวัด

วรรณกรรมที่เกี่ยวข้อง

คำว่า “Reliability” สามารถแทนด้วยคำในภาษาไทยได้หลายคำ เช่น ความเที่ยง ความเที่ยงตรง ความน่าเชื่อถือ ความเชื่อมั่น เป็นต้น ในงานวิจัยนี้จะใช้ความเชื่อมั่น ซึ่งหมายถึง ความแน่นอนของผลการวัดค่า ไม่ว่าจะวัดค่ากี่ครั้งก็ตาม ภายใต้เงื่อนไขและสถานการณ์เดียวกันผลการวัดจะมีค่าเท่ากัน ความเชื่อมั่นของเครื่องมือวิจัยมีค่าอยู่ระหว่าง 0.00 – 1.00 ค่ายิ่งใกล้ 1.00 มากเพียงใด ความเชื่อมั่นยิ่งสูงมากตามเพียงนั้น ทั้งนี้ แบบวัดที่มีค่าความเชื่อมั่นเท่ากับ 1.00 (ค่าสูงสุด) จะแปลความหมายได้ว่า ผู้ตอบแบบวัดจะ

ได้คะแนนเท่ากันในการวัดทุกครั้ง เมื่อเงื่อนไขและสถานการณ์ไม่เปลี่ยนแปลง เช่น ผู้ตอบแบบวัดคนหนึ่งได้คะแนนร้อยละ 80 ในครั้งแรกเมื่อกลับมาทำซ้ำอีก หากเงื่อนไขและสถานการณ์ (สุขภาพทั้งกาย-ใจ และสิ่งแวดล้อมอื่นๆ) ไม่เปลี่ยนแปลง รวมทั้งยังไม่มีการเรียนรู้เพิ่มเติมแล้ว ก็จะได้คะแนนเท่ากับร้อยละ 80 เช่นเดิม นักวัดผลจึงปรารถนาที่จะทำให้แบบวัดมีค่าความเชื่อมั่นสูง

เกณฑ์การแปลค่าความเชื่อมั่นมีหลากหลาย เช่น George และ Mallery (2003) ได้แปลค่าความเชื่อมั่นตามช่วงคะแนน ดังตารางต่อไปนี้

ตารางที่ 1 ระดับความเชื่อมั่น

ความเชื่อมั่น	การแปลค่า
0.90 ขึ้นไป	ดีมาก
0.80 - 0.89	ดี
0.70 - 0.79	ยอมรับได้
0.60 - 0.69	ยังน่าสงสัย
0.50 - 0.59	ค่อนข้างรับไม่ได้
น้อยกว่า 0.50	รับไม่ได้

ที่มา: ปรับปรุงจาก George และ Mallery (2003) และ Wikipedia (2014)

สูตรในการคำนวณหาความเชื่อมั่น มีทั้งที่คำนวณจากค่าสัมประสิทธิ์สหสัมพันธ์ เช่น สัมประสิทธิ์สหสัมพันธ์เพียร์สัน (Pearson's r) ที่ใช้แบบวัดคู่ขนาน (Paralleled Form) การวัดซ้ำ (Test-Retest) การแบ่งครึ่งแบบวัด เป็นต้น และคำนวณจากค่าความแปรปรวน เช่น วิธีของคูเดอร์-ริชาร์ดสัน (Kuder-Richardson) วิธีสัมประสิทธิ์อัลฟาของครอนบาค (The Coefficient of Alpha) เป็นต้น ในงานวิจัยนี้จะหาความเชื่อมั่นด้วยวิธีสัมประสิทธิ์อัลฟาของครอนบาค เพราะเป็นวิธีที่นักวิจัยนิยมใช้

สัมประสิทธิ์อัลฟาของครอนบาค

สัมประสิทธิ์อัลฟาของครอนบาคเป็นความเชื่อมั่นชนิดความสอดคล้องภายใน (Internal Consistency) ที่นักวิจัยนิยมใช้มากที่สุด ผู้เริ่มต้นพัฒนาคือ คูเดอร์-ริชาร์ดสัน (Kuder & Richardson, 1937) โดยใช้กับข้อมูลชนิดทวิภาค (มีค่า 0 หรือ 1) เท่านั้น ต่อมา ครอนบาค (Cronbach, 1951)

ได้พัฒนาสูตรครอนบาค เพื่อให้สามารถใช้ได้ทั้งข้อมูลทวิภาคและข้อมูลที่ไม่ใช่ทวิภาค จึงเรียกค่าความเชื่อมั่นชนิดนี้ว่า **ครอนบาคอัลฟา (α)** มีสูตรดังนี้

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum S_i^2}{S_t^2} \right)$$

เมื่อ α คือ ค่าความเชื่อมั่น หรือครอนบาคอัลฟา

n คือ จำนวนข้อคำถาม

$\sum S_i^2$ คือ ผลรวมของความแปรปรวนของคะแนนแต่ละข้อ

S_t^2 คือ ความแปรปรวนของคะแนนรวม

โดยอาศัยแนวคิดของวิชาแคลคูลัสเรื่องลิมิต เราสามารถวิเคราะห์ค่าความเชื่อมั่นจากสูตรข้างต้นได้ดังนี้

1. ถ้าค่า S_i^2 มีขนาดใหญ่กว่า $\sum S_i^2$ จะทำให้อัตราส่วน $\sum S_i^2 / S_t^2$ มีขนาดเล็ก ส่งผลให้ $(1 - \sum S_i^2 / S_t^2)$ มีค่าเข้าใกล้ 1.00 ซึ่งจะทำให้ความเชื่อมั่นมีค่าเข้าใกล้ 1.00 ด้วย

2. ค่า S_i^2 จะมีขนาดใหญ่เมื่อคะแนนรวมของผู้ตอบแบบวัดแต่ละคนมีค่าแตกต่างกันมาก หมายถึง ถ้าแบบวัดเป็นการวัดความคิดเห็น ก็แสดงว่าผู้ตอบแบบวัดมีความคิดเห็นในเรื่องที่วัดแตกต่างกันมาก หรือถ้าแบบวัดเป็นข้อสอบ ก็แสดงว่าผู้สอบมีระดับความสามารถแตกต่างกันมาก ซึ่งการแจกแจงของข้อมูลที่มีความแตกต่างกันดังกล่าว จะมีแนวโน้มเข้าสู่เส้นโค้งปกติ โดยทั่วไปแล้ว ตัวอย่างขนาด 30 ขึ้นไปจะมีการแจกแจงข้อมูลเข้าสู่เส้นโค้งปกติ

3. ถ้าผู้ตอบแบบวัดมีระดับความสามารถใกล้เคียงกัน จะทำให้ $\sum S_i^2$ และ S_i^2 มีค่าใกล้เคียงกัน ส่งผลให้ $\sum S_i^2 / S_i^2$ มีค่าเข้าใกล้ 1 ทำให้ $(1 - \sum S_i^2 / S_i^2)$ มีค่าน้อย ซึ่งจะทำให้ค่าความเชื่อมั่นมีค่าน้อยตามไปด้วย

4. คำว่า “ผู้ตอบมีระดับความสามารถใกล้เคียงกัน” หมายถึง ถ้าแบบวัดเป็นข้อสอบ ก็แสดงว่าผู้เข้าสอบเป็นกลุ่มอ่อนหรือกลุ่มเก่งกลุ่มใดกลุ่มหนึ่ง แต่ถ้าแบบวัดเป็นแบบสอบถาม ก็แสดงว่าผู้ตอบมีความคิดเห็นในเรื่องที่ถามใกล้เคียงกัน

5. แบบทดสอบหรือแบบสอบถาม ที่ข้อคำถามส่วนใหญ่มีลักษณะยาก (หรือง่าย) เกินไป หรือมีลักษณะคลุมเครือซึ่งผู้ตอบจะตอบด้วยการเดา จะทำให้คะแนนรวมมีค่าใกล้เคียงกัน นั่นคือ ทำให้ S_i^2 มีค่าน้อย ส่งผลให้ $\sum S_i^2 / S_i^2$ มีค่ามาก ผลที่ตามมาคือ $(1 - \sum S_i^2 / S_i^2)$ มีค่าน้อย เป็นเหตุให้ค่าความเชื่อมั่นมีค่าน้อยตามไปด้วย

6. จากข้อ 2 ถึงข้อ 5 แสดงว่า ค่าความเชื่อมั่นจะเปลี่ยนแปลงไปตามกลุ่มตัวอย่าง

7. ถ้าค่า $\sum S_i^2$ มากกว่าค่า S_i^2 แล้ว จะทำให้ $(1 - \sum S_i^2 / S_i^2)$ มีค่าติดลบ ซึ่งจะทำให้ค่าความเชื่อมั่นมีค่าติดลบตามไปด้วย ในกรณีนี้ถือเป็นกรณีผิดปกติของแบบทดสอบหรือแบบสอบถาม

การเพิ่มขนาดของค่าความเชื่อมั่น

องค์ประกอบหนึ่งที่มีผลต่อความเชื่อมั่น คือ จำนวนข้อของแบบวัดโดยหากเพิ่มจำนวนข้อของแบบวัดให้มากขึ้น ความเชื่อมั่นใหม่จะเพิ่มมากขึ้นตามไปด้วย ดังสูตรซึ่งปรับปรุงจากสูตรของสเปียร์แมน-บราวน์ (Wikipedia, 2013) ดังนี้ $r_n = \frac{qr_k}{1 + (q-1)r_k}$ เมื่อ r_n = ความเชื่อมั่นค่าใหม่, r_k = ความเชื่อมั่นค่าเดิม, n = จำนวนข้อของแบบวัดชุดใหม่, k = จำนวนข้อของแบบวัดชุดเดิม และ $q = n/k$

เช่น เดิมแบบวัดมีจำนวน 5 ข้อ ($k = 5$) ซึ่ง $r_5 = 0.67$ ถ้าเพิ่มจำนวนข้อเป็น 15 ข้อ ($n = 15$) แล้ว จะได้ความเชื่อมั่นค่าใหม่ (r_{15}) เท่ากับ 0.86 ดังนี้ ($q = n/k = 15/5 = 3$)

$$r_{15} = \frac{qr_5}{1 + (q-1)r_5} = \frac{3(0.67)}{1 + (3-1)(0.67)} = 0.86$$

จากการวิเคราะห์ข้างต้น สรุปได้ว่า สาเหตุที่ทำให้ความเชื่อมั่นมีค่าน้อย คือ ผู้ตอบแบบวัดมีความคิดเห็นหรือมีระดับความสามารถใกล้เคียงกัน หรือข้อคำถามส่วนใหญ่มีลักษณะยาก หรือง่าย หรือมีคลุมเครือ และสาเหตุที่ทำให้ค่าความเชื่อมั่นสูง คือ ผู้ตอบแบบวัดมีความคิดเห็นแตกต่างกันมาก หรือมีระดับความสามารถแตกต่างกันมาก หรือข้อมูลมีการแจกแจงปกติ นอกจากนี้ การเพิ่มจำนวนข้อคำถามของแบบวัดก็ส่งผลให้ค่าความเชื่อมั่นสูงขึ้นได้เช่นกัน สถิติที่ควรกล่าวถึงอีกสองตัว คือ ค่าความเชื่อมั่นปรับ (Cronbach's Alpha if Item Deleted: α) และค่าสหสัมพันธ์ปรับ Corrected Item-Total Correlation: r)

ค่าความเชื่อมั่นปรับ (α)

เป็นค่าความเชื่อมั่นเมื่อไม่รวมคะแนนของข้อที่กำลังศึกษา ใช้สำหรับพิจารณาการเปลี่ยนแปลงของค่าความเชื่อมั่นเมื่อไม่รวมข้อคำถามที่กำลังศึกษานั้น โดยหากค่านี้ต่ำลง แปลว่าข้อคำถามที่ตัดทิ้งไปนั้นเป็นข้อที่ดี ซึ่งเป็นข้อที่มีความสัมพันธ์กับคะแนนรวมอย่างมาก จึงควรพิจารณาเก็บข้อนั้นไว้ ในทางกลับกัน หากค่านี้สูงขึ้น ก็แปลว่าข้อคำถามที่ตัดทิ้งไปนั้นเป็นข้อที่มีความสัมพันธ์กับคะแนนรวมค่อนข้างน้อย จึงอาจพิจารณาตัดข้อคำถามนั้นทิ้ง

ค่าสหสัมพันธ์ปรับ (r)

เป็นค่าสหสัมพันธ์ของคะแนนแต่ละข้อกับคะแนนรวมของข้อคำถาม โดยไม่รวมถึงคะแนนของคำถามข้อนั้น ซึ่งเป็นค่าที่ส่งผลโดยตรงต่อความเชื่อมั่น มีสูตรคำนวณดังนี้

$$r = \frac{n \sum XY - \sum X \sum Y}{\sqrt{(n \sum X^2 - (\sum X)^2)(n \sum Y^2 - (\sum Y)^2)}}$$

เมื่อ r = Corrected Item-Total Correlation ของคำถามข้อ i

X = คะแนนของข้อคำถามข้อ i

Y = คะแนนรวมของข้อคำถามทุกข้อ ยกเว้นข้อ i

n = จำนวนของข้อคำถามทั้งหมด

r ในที่นี้คือ Pearson's r (Wikipedia, 2014) ซึ่งมีค่าอยู่ในช่วง 0.00 ถึง ± 1.00 โดยถ้า X และ Y ซึ่งเป็นข้อมูลเชิงปริมาณ ไม่มีความสัมพันธ์กันแล้ว ค่า r จะต่ำหรือเข้าใกล้ 0.00 ถ้า X และ Y มีความสัมพันธ์ไปในทิศทางเดียวกันแล้ว ค่า r จะสูงหรือเข้าใกล้ $+1.00$ ในทางกลับกัน ถ้า X และ Y มีความสัมพันธ์ไปในทิศตรงกันข้ามแล้ว จะทำให้ r มีค่าเข้าใกล้ -1.00

กรณีที่ Y ไม่รวมคะแนนของข้อ i ก็เพื่อให้คะแนนของ X ไม่ขึ้นอยู่กับการคะแนนของ Y ดังนั้น ถ้า X และ Y มีความสัมพันธ์กันทางบวกแล้ว สามารถแปลความหมายได้ว่า ตัวแปร X อยู่ในองค์ประกอบเดียวกันกับ Y หรือตัวแปร X อยู่ในมิติเดียวกันกับ Y หรือตัวแปร X วัดค่าในเรื่องเดียวกันกับ Y

สาเหตุที่ทำให้ X มีค่าใกล้เคียงกัน ส่งผลให้ r ของข้อใดข้อหนึ่งมีค่าติดลบหรือมีค่าต่ำกล่าวคือ ข้อคำถามยากเกินไป ซึ่งผู้ตอบส่วนใหญ่จะตอบด้วยการเดา ข้อคำถามที่มีลักษณะคลุมเครือซึ่ง สามารถตีความแตกต่างกันออกไป หรือเป็นข้อคำถามที่ผู้ตอบกลุ่มเรียนอ่อนหรือกลุ่มที่มีคะแนนรวมต่ำตอบได้มากกว่ากลุ่มเรียนเก่งหรือกลุ่มที่มีคะแนนรวมสูง (Allen, Reed-Rhoads, Terry, Murphy & Stone, 2008) ซึ่งถือว่าคำถามข้อนั้นไม่มีคุณภาพ เพราะทำให้ความคงเส้นคงวาลดลง และจะทำให้ข้อคำถามนั้นไม่มีคุณภาพเช่นกัน

ขนาดของ r ที่ถือว่า X และ Y มีความสัมพันธ์กันนั้น โดยทั่วไปนิยมใช้ 0.20 ขึ้นไป (Streiner & Norman, 2003) ในขณะที่นักวิจัยบางคนใช้ 0.25 หรือ 0.30 ขึ้นไป ผู้วิจัยเห็นว่า หากข้อคำถามมีจำนวนมาก และต้องการ

ตัดออกให้เหลือข้อคำถามจำนวนน้อยแล้ว ก็ควรเลือกใช้ 0.30 แต่หากข้อคำถามมีจำนวนน้อย ก็ควรเลือกใช้ 0.20 หรือ 0.25 แล้วแต่ความเหมาะสม

ระเบียบวิธีวิจัย

ประชากร คือ นักศึกษาทุกชั้นปีของมหาวิทยาลัยกรุงเทพ และกลุ่มตัวอย่างคือ นักศึกษาที่ตอบแบบสอบถามจำนวน 6 กลุ่มๆ ละ 5 ถึง 15 คน รวมเป็น 57 คน ได้มาจากการสุ่มแบบกลุ่ม ผู้วิจัยได้กำหนดขอบเขตเนื้อหาตามแบบประเมินผลการเรียนการสอน ซึ่งแบ่งเป็นสามส่วน คือ

ส่วนที่ 1 การประเมินความพึงพอใจในการจัดการเรียนการสอน ประกอบด้วย

- ข้อมูลเกี่ยวกับวิชาที่เรียน มีข้อคำถามจำนวน 4 ข้อ (ศึกษาความเชื่อมั่นของเครื่องมือวัดความพึงพอใจต่อวิชาเรียน)

- ข้อมูลเกี่ยวกับอาจารย์ผู้สอน มีข้อคำถามจำนวน 9 ข้อ (ศึกษาความเชื่อมั่นของเครื่องมือวัดความพึงพอใจต่ออาจารย์ผู้สอน)

ส่วนที่ 2 การประเมินตนเองของนักศึกษาเกี่ยวกับผลลัพธ์ที่ได้จากการเรียนในวิชานี้ มีข้อคำถามจำนวน 13 ข้อ

(ศึกษาความเชื่อมั่นของเครื่องมือวัดการประเมินตนเองของนักศึกษา)

ส่วนที่ 3 ข้อเสนอแนะอื่นๆ ที่เห็นว่าจะมีประโยชน์ต่อการพัฒนาคุณภาพของการเรียนในวิชานี้

ระยะเวลาในการวิจัย คือ ภาคการศึกษาที่ 2 / 2556

สถิติที่ใช้ คือ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ความเชื่อมั่น สัมประสิทธิ์สหสัมพันธ์เพียร์สัน และ สถิติ Fisher's Z

ผลการวิเคราะห์ข้อมูล

ตารางที่ 2 ค่าสหสัมพันธ์ปรับกับค่าความเชื่อมั่นปรับของแบบวัดความพึงพอใจต่อวิชาเรียนจำแนกตามกลุ่มที่ทดลองใช้

ข้อ	กลุ่มที่ 1 (นศ. 10 คน)		กลุ่มที่ 2 (นศ. 10 คน)		กลุ่มที่ 3 (นศ. 15 คน)		กลุ่มที่ 4 (นศ. 5 คน)		กลุ่มที่ 5 (นศ. 5 คน)		กลุ่มที่ 6 (นศ. 12 คน)		รวม (นศ. 57 คน)	
	r	α	r	α	r	α	r	α	r	α	r	α	r	α
1	.436	.450	.747	.866	.890	.813	.764	.750	.825	.798	.550	.647	.730	.756
2	.408	.469	.747	.866	.682	.886	.671	.780	.945	.686	.485	.645	.637	.797
3	.203	.620	.667	.903	.660	.893	.963	.618	.500	.867	.316	.742	.529	.842
4	.408	.469	.939	.809	.856	.828	.349	.892	.722	.833	.722	.463	.758	.739
Alpha		.580		.891		.890		.830		.841		.704		.831

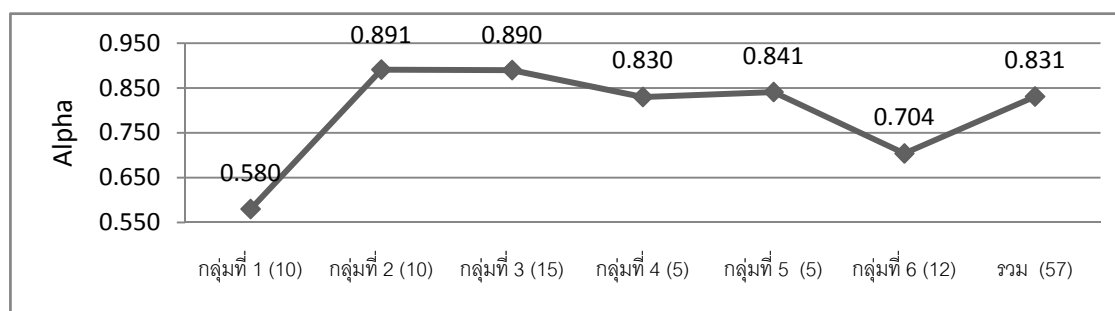
หมายเหตุ - r = Corrected Item-Total Correlation (ค่าสหสัมพันธ์ปรับ) เป็นค่าสหสัมพันธ์ของคะแนนแต่ละข้อกับคะแนนรวมของข้อคำถาม โดยไม่รวมถึงคะแนนของคำถามข้อนั้น

- α = Cronbach's Alpha if Item Deleted (ค่าความเชื่อมั่นปรับ) เป็นค่าความเชื่อมั่นเมื่อไม่รวมคะแนนของข้อที่กำลังศึกษา

- ตัวอย่างการแปลความจากตาราง จากข้อมูลของกลุ่มที่ 1 คำถามข้อที่ 1 มีค่า r = 0.436 ซึ่งเป็นค่าที่สูง เมื่อเทียบกับข้อที่ 2 - 4 แต่หากตัดคำถามข้อที่ 1 ออกจากแบบวัด ค่า Alpha ของกลุ่มที่ 1 จะลดลงจาก 0.580 เป็น 0.450 นั่นคือ มีค่า Alpha ลดลง ในทางตรงข้าม หากพิจารณาตัดคำถามข้อ 3 ออก ซึ่งมีค่า r = 0.203 จะทำให้ค่า Alpha เพิ่มขึ้นจาก 0.580 เป็น 0.620 นั่นคือ มีค่า Alpha สูงขึ้น

จากตารางที่ 2 พบว่า r มีค่าสูงกว่า 0.20 ทุกข้อและทุกกลุ่ม โดยมีค่าจาก .203 ถึง .963 และ Alpha มีค่าสูงที่สุด คือ 0.891 ของกลุ่มที่ 2 รองลงมาซึ่งมีค่าใกล้เคียงกันมากคือ 0.890 ของกลุ่มที่ 3 และมีค่าต่ำที่สุดคือ 0.580 ของกลุ่มที่ 1 โดยค่ารวมคือ 0.831

ข้อสังเกตจากตารางนี้คือ กลุ่มขนาด 10 คนมีค่า Alpha ทั้งต่ำสุด (0.580) และสูงสุด (0.891) ในขณะเดียวกันกลุ่มขนาด 5 คน มีค่า Alpha 0.830 - 0.841 ซึ่งใกล้เคียงกับค่า Alpha โดยรวม (0.831) ซึ่งหมายถึงว่า Alpha จะมีค่ามากหรือน้อย ไม่สามารถคาดการณ์จากขนาดของกลุ่มตัวอย่างได้



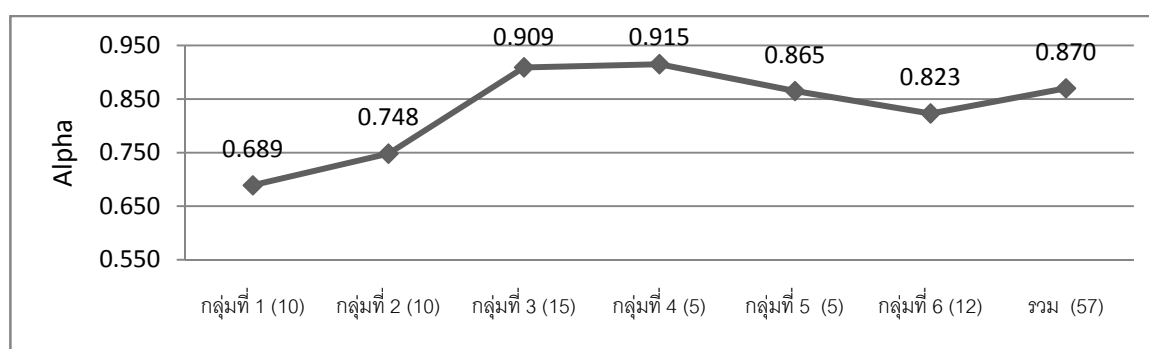
ภาพที่ 2 ค่าความเชื่อมั่นของแบบวัดความพึงพอใจต่อวิชาเรียน จำแนกตามกลุ่มที่ทดลองใช้

ตารางที่ 3 ค่าสหสัมพันธ์ปรับกับค่าความเชื่อมั่นปรับของแบบวัดความพึงพอใจต่ออาจารย์ผู้สอนจำแนกตามกลุ่มที่ทดลองใช้

ข้อ	กลุ่มที่ 1 (นศ. 10 คน)		กลุ่มที่ 2 (นศ. 10 คน)		กลุ่มที่ 3 (นศ. 15 คน)		กลุ่มที่ 4 (นศ. 5 คน)		กลุ่มที่ 5 (นศ. 5 คน)		กลุ่มที่ 6 (นศ. 12 คน)		รวม (นศ. 57 คน)	
	r	α	r	α	r	α	r	α	r	α	r	α	r	α
1	.000	.726	.600	.700	.563	.907	.516	.918	.588	.852	.359	.822	.512	.864
2	.486	.648	.429	.725	.608	.906	-.256	.943	.790	.839	.633	.807	.520	.863
3	.674	.618	.340	.740	.727	.896	.446	.920	.980	.807	.724	.783	.624	.854
4	.749	.559	.829	.698	.787	.893	.991	.882	.622	.848	.590	.800	.720	.845
5	.491	.634	.829	.698	.715	.900	.887	.896	.814	.837	.623	.792	.593	.858
6	.581	.617	.512	.711	.721	.896	.887	.896	.791	.832	.418	.817	.672	.849
7	-.068	.733	.343	.754	.848	.889	.944	.893	.000	.909	.730	.776	.596	.859
8	.341	.683	.340	.740	.629	.904	.991	.882	.433	.865	.206	.835	.585	.858
9	.159	.694	.191	.758	.709	.898	.901	.894	.663	.844	.573	.799	.658	.851
Alpha		.689		.748		.909		.915		.865		.823		.870

จากตารางที่ 3 ค่า r ของกลุ่มที่ 1 ต่ำกว่า 0.20 จำนวน 3 ข้อ และกลุ่มที่ 2, 4, 5 ต่ำกว่า 0.20 กลุ่มละ 1 ข้อ โดยกลุ่มที่ 3, 6, กลุ่มรวม ซึ่งจำนวนผู้ตอบมีขนาด 15, 12, 57 ตามลำดับ มีค่า r สูงกว่า 0.20 ทุกข้อ และ Alpha มีค่าสูงที่สุดคือ 0.915 เป็นของกลุ่มที่ 4 รองลงมาคือ 0.909 เป็น

ของกลุ่มที่ 3 และมีค่าต่ำที่สุดคือ 0.689 เป็นของกลุ่มที่ 1 โดยค่ารวมคือ 0.870 ข้อสังเกตจากตารางนี้ คือ จำนวนผู้ตอบขนาด 10 คน มีค่า Alpha ต่ำสุด (0.689) ในขณะที่จำนวนผู้ตอบขนาด 5 คนมีค่า Alpha สูงสุด (0.915)



ภาพที่ 3 ค่าความเชื่อมั่นของแบบวัดความพึงพอใจต่ออาจารย์ผู้สอน จำแนกตามกลุ่มที่ทดลองใช้

ตารางที่ 4 ค่าสหสัมพันธ์ปรับกับค่าความเชื่อมั่นปรับของแบบวัดการประเมินตนเองของนักศึกษาจำแนกตามกลุ่มที่ทดลองใช้

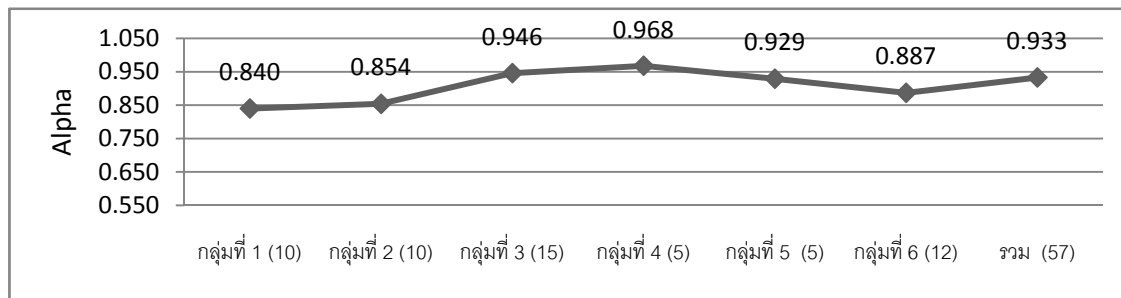
ข้อ	กลุ่มที่ 1 (นศ. 10 คน)		กลุ่มที่ 2 (นศ. 10 คน)		กลุ่มที่ 3 (นศ. 15 คน)		กลุ่มที่ 4 (นศ. 5 คน)		กลุ่มที่ 5 (นศ. 5 คน)		กลุ่มที่ 6 (นศ. 12 คน)		รวม (นศ. 57 คน)	
	r	α	r	α	r	α	r	α	r	α	r	α	r	α
1	.712	.818	.556	.841	.730	.942	.825	.967	.505	.929	.852	.863	.733	.927
2	.259	.858	.479	.845	.615	.946	.935	.963	.420	.931	.575	.879	.593	.931
3	.597	.821	.597	.839	.698	.943	.825	.967	.722	.925	.673	.874	.645	.929
4	.793	.812	.608	.840	.635	.945	-.286	.982	.929	.919	.745	.869	.629	.930
5	.762	.806	.549	.842	.862	.938	.774	.968	.151	.941	.398	.886	.638	.930
6	.741	.813	.748	.829	.917	.936	.977	.962	.552	.928	.780	.867	.804	.924
7	.635	.819	.597	.839	.714	.942	.980	.962	.918	.913	.405	.886	.729	.927
8	.494	.829	.500	.844	.827	.939	.980	.962	.642	.925	.331	.890	.690	.928
9	.252	.842	.324	.854	.547	.947	.980	.962	.903	.916	.389	.888	.653	.929
10	.414	.834	.408	.851	.749	.941	.977	.962	.948	.912	.554	.881	.710	.927
11	.470	.831	.324	.854	.790	.940	.977	.962	.810	.919	.545	.880	.738	.926
12	.258	.845	.676	.837	.765	.941	.977	.962	.899	.916	.723	.871	.754	.926
13	.256	.842	.493	.845	.738	.942	.833	.965	.789	.923	.574	.881	.716	.927
Alpha		.840		.854		.946		.968		.929		.887		.933

จากตารางที่ 4 ค่า r ของกลุ่มที่ 4 และ 5 ต่ำกว่า 0.20 กลุ่มละ 1 ข้อ และ Alpha มีค่าสูงที่สุดคือ 0.968 ของกลุ่มที่ 4 รองลงมาคือ 0.946 ของกลุ่มที่ 3 และมีค่าต่ำที่สุดคือ 0.840 ของกลุ่มที่ 1 โดยค่ารวมคือ 0.933 จากตารางนี้มีข้อที่น่าสังเกตสองประการ คือ

ประการที่หนึ่ง เฉพาะจำนวนผู้ตอบขนาด 5 คนเท่านั้นที่มีค่า r ต่ำกว่า 0.20 หากตัดสินใจตัดข้อที่ด้อยคุณภาพออกตาม

ตัวบ่งชี้ (r) จะต้องตัดข้อ 4. (กลุ่มที่ 4) และข้อ 5. (กลุ่มที่ 5) จะทำให้ค่า Alpha สูงขึ้น หรือถ้าไม่ตัดสองข้อนี้ออกอาจต้องเพิ่มจำนวนขนาดของผู้ตอบให้มากกว่า 5 คน

ประการที่สอง ในส่วนของนักศึกษาที่มีจำนวนข้อคำถามมากที่สุดนี้ มีค่า Alpha ใกล้เคียงกัน ระหว่าง 0.840 – 0.968 ซึ่งใกล้เคียงกับค่า Alpha โดยรวม (0.933)



ภาพที่ 4 ค่าความเชื่อมั่นของแบบวัดการประเมินตนเองของนักศึกษา จำแนกตามกลุ่มที่ทดลองใช้

ผลการทดสอบสมมติฐาน

ตารางที่ 5 ค่าความเชื่อมั่น (r) สถิติ Fisher's Z และสถิติทดสอบ (z*) ของแบบประเมินผลการเรียนการสอนที่คำนวณจากการทดลองใช้กลุ่มย่อยและกลุ่มรวมทั้งในส่วนของวิชาเรียน อาจารย์ผู้สอน และนักศึกษา (สถิติทดสอบสมมติฐานที่ 1)

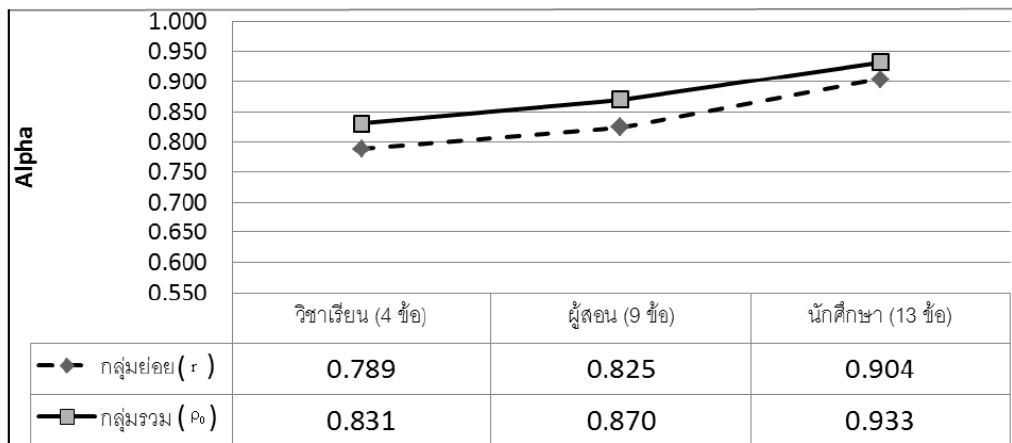
N	วิชาเรียน (4 ข้อ): $\rho_0 = 0.831$			อาจารย์ผู้สอน (9 ข้อ): $\rho_0 = 0.870$			นักศึกษา (13 ข้อ): $\rho_0 = 0.933$		
	r	Z	z	r	Z	z	r	Z	z
10	0.580	0.662	-1.40	0.689	0.846	-1.29	0.840	1.221	-1.22
10	0.891	1.427	0.62	0.748	0.968	-0.96	0.854	1.271	-1.09
15	0.890	1.422	0.80	0.909	1.522	0.65	0.946	1.792	0.39
5	0.830	1.188	0.00	0.915	1.557	0.32	0.968	2.060	0.54
5	0.841	1.225	0.05	0.865	1.313	-0.03	0.929	1.651	-0.04
12	0.704	0.875	-0.95	0.823	1.166	-0.50	0.887	1.408	-0.82

$$z = \frac{Z - \mu_z}{\sigma_z} \text{ เมื่อ } Z = \frac{1}{2} \log_e \left(\frac{1+r}{1-r} \right), \mu_z = \frac{1}{2} \log_e \left(\frac{1+\rho_0}{1-\rho_0} \right), \sigma_z = \frac{1}{\sqrt{N-3}}$$

(Spiegel & Stephens, 1999)

จากตารางที่ 5 ผลการทดสอบสมมติฐานทางวิจัยโดยใช้ระดับนัยสำคัญ 0.05 พบว่า ค่าสัมบูรณ์ของ z อยู่ระหว่าง 0.80 – 1.40 ซึ่งน้อยกว่า 1.96 จึงไม่ปฏิเสธสมมติฐานทาง

วิจัยที่ว่า ค่าความเชื่อมั่นที่คำนวณจากกลุ่มย่อยมีค่าเท่ากับค่าความเชื่อมั่นที่คำนวณจากกลุ่มรวม ทั้งนี้แบบวัดยังมีจำนวนข้อคำถามมาก ยิ่งทำให้ค่าความเชื่อมั่นสูงขึ้นด้วย ดังภาพที่ 5



ภาพที่ 5 เปรียบเทียบค่าความเชื่อมั่นเฉลี่ยที่คำนวณจากกลุ่มย่อยและกลุ่มรวม

ตารางที่ 6 สัมประสิทธิ์สหสัมพันธ์เพียร์สัน (r) ระหว่างตัวแปรจำนวนผู้ตอบกับค่าความเชื่อมั่น (สถิติทดสอบสมมติฐานที่ 2)

ตัวแปร	ค่าความเชื่อมั่น
จำนวนผู้ตอบ	-0.100

หมายเหตุ $p > 0.05$

จากตารางที่ 6 ผลการทดสอบสมมติฐานที่ระดับนัยสำคัญ 0.05 พบว่า ตัวแปรจำนวนผู้ตอบและค่าความเชื่อมั่น ไม่มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 จึงสรุปได้ว่า ไม่ว่าจำนวนผู้ตอบจะมากหรือน้อยก็ตาม ไม่ทำให้ค่าความเชื่อมั่นมีค่ามากหรือน้อยตามไปด้วย

สรุปและอภิปรายผล

สรุปตามวัตถุประสงค์ของการวิจัยดังต่อไปนี้

1. ความเชื่อมั่นของแบบวัดที่คำนวณจากการทดลองใช้ในกลุ่มย่อยและกลุ่มรวม

ข้อมูลเกี่ยวกับ **วิชาเรียน** มีข้อคำถามจำนวน 4 ข้อ จำนวนผู้ตอบขนาด 10 คนมีความเชื่อมั่นทั้งต่ำสุดและสูงสุด โดยจำนวนผู้ตอบขนาด 5 มีความเชื่อมั่นใกล้เคียงกับความเชื่อมั่นโดยรวม และความเชื่อมั่นโดยรวมคือ 0.831 ส่วนข้อมูลเกี่ยวกับ **ผู้สอน** มีข้อคำถามจำนวน 9 ข้อ จำนวนผู้ตอบขนาด 10 คนมีความเชื่อมั่นต่ำสุด และจำนวนผู้ตอบขนาด 5 คนมีความเชื่อมั่นสูงสุด และความเชื่อมั่นโดยรวมคือ 0.870 ส่วนข้อมูลการประเมินตนเองของ **นักศึกษา** มีข้อคำถามจำนวน 13 ข้อ กลุ่มย่อยและกลุ่มใหญ่มีความเชื่อมั่นใกล้เคียงกัน และความเชื่อมั่นโดยรวมคือ 0.933

ผลการทดสอบ **สมมติฐานที่ 1** สรุปว่า ความเชื่อมั่นของแบบวัดที่คำนวณจากกลุ่มย่อยมีค่าเท่ากับค่าความเชื่อมั่นของแบบวัดที่คำนวณจากกลุ่มรวม ซึ่งค่าเฉลี่ยความเชื่อมั่นที่คำนวณจากกลุ่มย่อย 6 กลุ่ม เทียบได้กับค่าเฉลี่ยของกลุ่มตัวอย่าง ซึ่งกลุ่มย่อยแต่ละกลุ่มจะเรียนรายวิชาที่ต่างกัน ผู้สอนก็แตกต่างกัน ดังนั้น เป้าหมายของการประเมินจึงแตกต่างกันไปด้วย และค่าเฉลี่ยความเชื่อมั่นที่คำนวณจากกลุ่มรวม เทียบได้กับค่าเฉลี่ยของประชากร ผลการทดสอบสมมติฐานที่ 1 เป็นคำตอบของวัตถุประสงค์ข้อแรก คือ ผลการศึกษาความเชื่อมั่นของแบบวัดที่คำนวณจากการทดลองใช้ในกลุ่มย่อยและกลุ่มรวม

2. ความสัมพันธ์ระหว่างจำนวนผู้ตอบกับค่าความเชื่อมั่นของแบบวัด

ผลการทดสอบ **สมมติฐานที่ 2** สรุปว่า จำนวนผู้ตอบกับค่าความเชื่อมั่น ไม่มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 สรุปว่า ไม่ว่าจำนวนผู้ตอบจะมากหรือน้อยก็ตาม ไม่ทำให้ค่าความเชื่อมั่นมีค่ามากหรือน้อยตามไปด้วย

อย่างไรก็ตาม การที่จำนวนผู้ตอบมีน้อยจะทำให้ค่าสหสัมพันธ์ปรับมีค่าอ่อนไหว กล่าวคือ ค่าสหสัมพันธ์ปรับจะมีค่าสูงมากๆ หรือต่ำมากๆ หรือติดลบ ในกรณีนี้ นักวิจัย

อาจต้องตัดข้อคำถามที่มีปัญหาทิ้งเพราะมีค่าต่ำกว่า 0.20 ทั้งๆ ที่เมื่อเพิ่มจำนวนของผู้ตอบแล้ว ข้อคำถามดังกล่าวอาจไม่มีปัญหาก็ได้

ข้อเสนอแนะในการนำไปใช้

การทดลองใช้ของแบบวัด

ผลการวิจัยเพื่อตอบวัตถุประสงค์ข้อ 1 พบว่า ความเชื่อมั่นของแบบวัดที่คำนวณจากกลุ่มย่อยไม่แตกต่างจากความเชื่อมั่นของแบบวัดที่คำนวณจากกลุ่มรวม ดังนั้น ในขั้นตอนการทดลองใช้แบบวัด ผู้วิจัยสามารถเลือกใช้กลุ่มทดลองที่มาจากกลุ่มใดกลุ่มหนึ่งที่มีผู้สอนเป็นคนเดียวกันหรือมาจากผู้เรียนหลายกลุ่มที่มีผู้สอนแตกต่างกันก็ได้ ขอเพียงให้กลุ่มทดลองเป็นตัวแทนที่ดีของประชากร

ผลการวิจัยเพื่อตอบวัตถุประสงค์ข้อ 2 พบว่า จำนวนผู้ตอบแบบวัดกับค่าความเชื่อมั่นของแบบวัดไม่มีความสัมพันธ์กัน แสดงว่า การเพิ่มจำนวนผู้ตอบแบบวัด ไม่ทำให้ค่าความเชื่อมั่นของแบบวัดสูงตามไปด้วย อย่างไรก็ตาม ในทางทฤษฎีสถิติ ถือว่า ขนาดของกลุ่ม 30 คนเป็นขนาดที่มีการแจกแจงข้อมูลเข้าสู่โค้งปกติ ดังนั้น กลุ่มทดลองจึงควรมีขนาด 30 หรือมากกว่านี้ แต่หากประสบปัญหาในการเลือกกลุ่มตัวอย่างหรือทำให้กลุ่มตัวอย่างมีขนาดน้อยกว่านี้บ้างก็ยังสามารถใช้ได้ แต่ต้องใช้ความระมัดระวังในการพิจารณาตัดข้อคำถามบางข้อออก

การนำเสนอค่าความเชื่อมั่น

ข้อสังเกตจากงานวิจัยในครั้งนี้ คือ เนื้อหาตามแบบประเมินผลการเรียนการสอนโดยมีส่วนที่ต้องหาค่าความเชื่อมั่นสามส่วน คือ ส่วนวิชาที่เรียน ส่วนอาจารย์ผู้สอน และส่วนของนักศึกษา ซึ่งมีความหมายที่ต่างกัน ดังนั้น ในการนำเสนอค่าความเชื่อมั่นควรแยกนำเสนอเป็นส่วนๆ เพื่อให้ผู้อ่านทราบว่าแต่ละส่วนมีค่าแตกต่างกันอย่างไร ในการนี้ ส่วนของวิชาเรียน ส่วนของผู้สอน และส่วนของนักศึกษา มีค่าความเชื่อมั่นเท่ากับ 0.83, 0.87 และ 0.93 ตามลำดับ ดังนั้น ถ้ามีความจำเป็นที่จะต้องนำเสนอเพียงค่าเดียวแล้ว ก็ควรใช้ค่าเฉลี่ยของสามค่าดังกล่าว คือ $(0.83 + 0.87 + 0.93) / 3 = 0.88$ โดยไม่ควรใช้คำนวณจากข้อคำถามทั้งหมด เพราะแบบวัดทั้งสามส่วนดังกล่าวมาจากตัวแปรที่มีความหมายแตกต่างกันจึงไม่สามารถนำคะแนนมารวมกันได้

ข้อเสนอแนะในการวิจัยครั้งต่อไป

แบบประเมินผลการเรียนการสอนที่ศึกษาในครั้งนี้ มีข้อมูลหลัก 3 ส่วน คือ ข้อมูลเกี่ยวกับวิชาที่เรียน ข้อมูลเกี่ยวกับอาจารย์ผู้สอน และข้อมูลการประเมินตนเองของนักศึกษาเกี่ยวกับผลลัพธ์ที่ได้จากการเรียน ในการวิจัยครั้งต่อไป ควรศึกษามิติหรือองค์ประกอบของข้อคำถามของแบบประเมินที่เกี่ยวกับอาจารย์ผู้สอน และข้อมูลการประเมินตนเองของนักศึกษา เช่น ส่วนของการประเมินตนเองของนักศึกษาที่มีข้อคำถาม 13 ข้อ อยากทราบว่า ข้อคำถามทั้ง 13 ข้อนี้ สามารถสะท้อนผลลัพธ์เกี่ยวกับการเรียนรู้ของนักศึกษาในด้านใดบ้าง เป็นต้น

บรรณานุกรม

- Allen, K., Reed-Rhoads, T., Terry, R. A., Murphy, T. J., & Stone, A. D. (2008). Coefficient alpha: An engineer's interpretation of test reliability. *Journal of Engineering Education*, 97(1), 87-94.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference*. 11.0 update (4th ed.). Boston: Allyn & Bacon.
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2(3), 151-160.
- Spiegel M. R., & Stephens L. J. (1999). *Theory and Problems of Statistics* (3rd ed.). New York: McGraw-Hill.
- Streiner, D. L., & Norman, G. R. (2003). *Health measurement scales: A practical guide to their development and use*. Oxford: Oxford University Press.
- Wikipedia. (2013). Retrieved April 13, 2014, from http://en.wikipedia.org/wiki/Spearman%E2%80%900%9393Brown_prediction_formula.
- Wikipedia. (2014). Retrieved April 10, 2014, from http://en.wikipedia.org/wiki/Pearson_product-moment_correlation_coefficient



รองศาสตราจารย์วัฒนา สุนทรชัย สำเร็จการศึกษาสูงสุดคือ พบ.ม. (สถิติประยุกต์) สถาบันบัณฑิตพัฒนบริหารศาสตร์ ปัจจุบันเป็นรองศาสตราจารย์คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยกรุงเทพ เขียนหนังสือเกี่ยวกับโปรแกรม SPSS จำนวน 4 เล่ม ปี 2556-58 มีงานวิจัยจำนวน 8 เรื่อง (ไม่นับปีก่อนนั้น) และบทความวิชาการจำนวน 3 เรื่อง



อาจารย์วนิดา คุชชัยสิทธิ์ สำเร็จการศึกษา วท.ม. (เคมีอินทรีย์), มหาวิทยาลัยเกษตรศาสตร์ และ กศ.บ. (เคมี), (เกียรตินิยมอันดับ2) มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เคยเป็นอาจารย์ประจำคณะวิทยาศาสตร์และเทคโนโลยี และอาจารย์ประจำสำนักพัฒนาคุณภาพการศึกษา ปัจจุบันดำรงตำแหน่งหัวหน้าศูนย์บริหารจัดการหลักสูตรและวิชาศึกษาทั่วไป สำนักวิชาการ มหาวิทยาลัยกรุงเทพ