

## การเปรียบเทียบประสิทธิภาพระหว่างการทดสอบแบบปกติกับการทดสอบ ทางโปรแกรมคอมพิวเตอร์

### A Comparison of Psychometric Properties between Paper-pencil-testing Method and Computer Testing Method

พิศิษฐ์ ตัณฑวณิช

#### บทคัดย่อ

การวิจัยครั้งนี้ มีวัตถุประสงค์เพื่อพัฒนาโปรแกรมคอมพิวเตอร์เพื่อใช้ทดแทนการทดสอบแบบปกติ และเพื่อมุ่งศึกษาเปรียบเทียบประสิทธิภาพระหว่างการทดสอบแบบปกติกับการทดสอบทางโปรแกรมคอมพิวเตอร์ เมื่อใช้ทดสอบกับนักศึกษาระดับปริญญาตรีสาขาการศึกษา โดยได้พัฒนาโปรแกรมคอมพิวเตอร์ขึ้น 2 โปรแกรม เพื่อใช้กับการทดสอบกรณีที่คำถามเป็นข้อความ และกรณีที่คำถามปรากฏอยู่ในลักษณะของรูปภาพ โดยผู้วิจัยได้ พัฒนาแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนขึ้น 2 ฉบับคือ “แบบทดสอบความรู้พื้นฐานด้านการวัดผลประเมินผล” และ “แบบทดสอบความรู้พื้นฐานด้านการวิเคราะห์ข้อสอบ” แบบทดสอบแต่ละฉบับเป็นแบบทดสอบชนิดทำ ตัวเลือก ประกอบด้วยข้อสอบจำนวน 40 และ 30 ข้อตามลำดับ โดยนำไปศึกษาแก่นักศึกษาระดับปริญญาตรีสาขา การศึกษา มหาวิทยาลัยราชภัฏสุรินทร์ ปีการศึกษา 2552 จำนวน 69 คนที่ลงทะเบียนเรียนรายวิชา “การประเมินผล และการสร้างแบบทดสอบ” โดยกลุ่มตัวอย่างทุกคนต้องทำแบบสอบทั้งสองฉบับในการทดสอบแต่ละฉบับจะแบ่ง นักศึกษาเป็นสองกลุ่มย่อยจำนวนเท่าๆ กัน ลักษณะการทดสอบ ผู้วิจัยจะสุ่มนักศึกษาครั้งหนึ่งทำการทดสอบแบบ ปกติ อีกกลุ่มหนึ่งได้รับการทดสอบทางโปรแกรมคอมพิวเตอร์ และเมื่อทำการทดสอบกับแบบทดสอบอีกฉบับหนึ่ง จะมีการสุ่มสลับกลุ่มตัวอย่าง ให้ได้รับการทดสอบในลักษณะที่ต่างไปจากเดิม จากนั้นได้มีการบันทึกผลการ ทดสอบและนำผลมาวิเคราะห์ ซึ่งได้ข้อค้นพบโดยสรุปดังนี้

1. ค่าคะแนนเฉลี่ยของคะแนนในส่วนของ “แบบทดสอบความรู้พื้นฐานด้านการวัดผลประเมินผล” ของ กลุ่มตัวอย่างเมื่อทดสอบโดยวิธีต่างกันคือ การทดสอบแบบปกติ และการทดสอบกับโปรแกรมคอมพิวเตอร์พบว่าไม่ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ และพบว่า ค่าคะแนนเฉลี่ยของคะแนนในส่วนของ “แบบทดสอบ ความรู้พื้นฐานด้านการวิเคราะห์ข้อสอบ” เมื่อทดสอบโดยวิธีต่างกันคือ การทดสอบแบบปกติ และการทดสอบกับ โปรแกรมคอมพิวเตอร์ ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติทั้งสองฉบับ

2. ในส่วนของการศึกษาความตรงของผลการทดสอบ เมื่อศึกษาเกี่ยวกับค่าคะแนนที่ได้จากการสอบกับ คะแนนเกณฑ์คือผลการทดสอบของกลุ่มตัวอย่างในการทดสอบปลายภาคเรียน พบว่า ในส่วนของ “แบบทดสอบ ความรู้พื้นฐานด้านการวัดผลประเมินผล” ที่ทดสอบโดยวิธีการต่างกันคือ การทดสอบแบบปกติ กับ การทดสอบกับ โปรแกรมคอมพิวเตอร์ และในส่วนของ “แบบทดสอบความรู้พื้นฐานด้านการวิเคราะห์ข้อสอบ” ที่ทดสอบโดย วิธีการต่างกันคือ การทดสอบแบบปกติ กับ การทดสอบกับโปรแกรมคอมพิวเตอร์ แต่ละฉบับมีค่าสัมประสิทธิ์ สหสัมพันธ์สูงอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

3. ค่าความเที่ยงเมื่อคำนวณโดยสูตร KR20 ของ “แบบทดสอบความรู้พื้นฐานด้านการวัดผลประเมินผล” เมื่อทดสอบโดยวิธีการต่างกัน คือการทดสอบแบบปกติ กับ การทดสอบกับโปรแกรมคอมพิวเตอร์พบว่ามีค่าสูงและ ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ รวมทั้งในส่วนของ “แบบทดสอบความรู้พื้นฐานด้านการวิเคราะห์ข้อสอบ”

ที่ทดสอบโดยวิธีการต่างกัน มีค่าความเที่ยงสูงและไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

4. เมื่อเปรียบเทียบความความยากของข้อสอบรายคู่ของ “แบบทดสอบความรู้พื้นฐานด้านการวัดผลประเมินผล” เมื่อทดสอบโดยวิธีการต่างกัน พบว่ามีข้อสอบจำนวน 36 ข้อจากจำนวนทั้งหมด 40 ข้อที่มีค่าสัดส่วนความยากไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ในส่วนของแบบทดสอบ “แบบทดสอบความรู้พื้นฐานด้านการวิเคราะห์ข้อสอบ” ที่ทดสอบโดยวิธีการต่างกัน มีข้อสอบ 29 ข้อจากจำนวนทั้งหมด 30 ข้อ ที่มีค่าสัดส่วนความยากไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

5. ในการศึกษาค่าอำนาจจำแนก พบว่าในส่วนของ “แบบทดสอบความรู้พื้นฐานด้านการวัดผลประเมินผล” ค่าอำนาจจำแนกของข้อสอบที่ทดสอบโดยวิธีการต่างกันมีค่าสัมประสิทธิ์สหสัมพันธ์สูงอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ในส่วนของ “แบบทดสอบความรู้พื้นฐานด้านการวิเคราะห์ข้อสอบ” ค่าอำนาจจำแนกของข้อสอบที่ทดสอบโดยวิธีการต่างกันมีค่าสัมประสิทธิ์สหสัมพันธ์สูงอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 เช่นกัน

จากผลการวิเคราะห์ข้อมูลทั้งหมด แสดงให้เห็นว่า การทดสอบโดยวิธีการปกติ กับการทดสอบทางโปรแกรมคอมพิวเตอร์ให้ผลต่าง ๆ ในด้านคุณสมบัติของแบบทดสอบที่ดีที่ใกล้เคียงกันอย่างมาก

**คำสำคัญ:** การทดสอบแบบปกติ การทดสอบทางโปรแกรมคอมพิวเตอร์ คุณภาพแบบทดสอบ

### Abstract

The objectives of this research were to construct a computer program for classroom testing and to compare the psychometric properties between paper-pencil-testing method (The PP Method) and computer testing method (The CT Method) applied to students studying in B.Ed. program. The two computer programs were developed. One was for using in the case of text-file-test and another one was used in case of picture-like-test. The two tests which were, The Basic Concept in Measurement and Evaluation Test (The BCME test), and The Concept of Item Analysis Test (The CIA test) were written by the researcher. Each test consisted of 40 and 30 items respectively, the Tests were written in 5 multiple choices form. Subjects consisted of 69 students studying in the Educational Bachelor degree program in 2009 academic year in Surin Rajabhat University. All subjects took the “Educational evaluation and Test Construction course taught by the researcher. All subjects took both The BCME Test and The CIA Test, However subjects were divided into two subgroups and were random to take test by The PP Method or The CT Method. Also each subgroup were swapped for testing by The PP Method and The CT Method. The data was collected and the results from data analysis were as the following details:-

1. Means scores of The BCME Test testing by The PP Method and The CT were not statistical difference. Also means score of The CIA Test testing by The PP Method and The CT Method were not statistical difference.

2. The correlation coefficients for validity study when the final test scores was used as a criterion score for the BCME Test testing by The PP Method and The CT Method was high and

statistical significant at 0.01 level and also the correlation coefficients for validity study for The CIA Test testing by The PP Method and The CT Method was high and statistical significant at 0.01.

3. The KR20 coefficients of The BCME Test testing by The PP Method and The CT Method were high and not statistical difference and Also the KR20 coefficients of The CIA Test testing by The PP Method and The CT Method were and not statistical difference.

4. When Z proportion test was applied to test the difference between the proportion of students who answer each item correctly in THE BCME Test when difference testing methods were used, there were no statistical difference for 36 items from all 40 items. Also for THE CIA Test when different testing methods were used, there were no statistical difference for 29 items from all 30 items.

5. The correlation coefficient between pair of item discrimination powers for all items in The BCME Test when different testing methods were used was medium high and statistical significant at 0.01 level and also correlation coefficients for item discrimination powers between pair of item in The CIA Test when different testing methods were used was medium high and statistical significant at 0.01 level.

From all findings above, it could be concluded that the psychometric properties of The BCMS Test and The CIA Test when The PP method and The CT Method were administered yielded very similar good results.

**Keywords :** paper-pencil-test, computer test, psychometric property.

**E-mail :** pisitz@hotmail.com

## ความนำ

ปัจจุบันนี้ เทคโนโลยีคอมพิวเตอร์มีการพัฒนาการเจริญรุดหน้าไปเป็นอย่างมากในด้านการประยุกต์ใช้ในทางการศึกษา มีการพัฒนาไปถึงขั้นการออกแบบทดสอบทางโปรแกรมคอมพิวเตอร์ขึ้นมาใช้ควบคุมหรือมุ่งใช้เพื่อการทดแทนการทดสอบในระบบปกติที่มาช้านานแล้ว แต่อย่างไรก็ดีในกรณีของประเทศไทย ยังคงมีประเด็นการใช้เทคโนโลยีการทดสอบทางโปรแกรมคอมพิวเตอร์อยู่ 2 ประการ คือ ประการแรก หากคนไทยไม่คิดพัฒนาโปรแกรมในส่วนดังกล่าวนี้ขึ้นมาใช้เอง ก็คงต้องซื้อโปรแกรมจากต่างประเทศและเสียเงินตราออกไปเป็นอย่างมาก ประการต่อมาคือ ขณะนี้แทบจะยังไม่มีกรวิจัยในลักษณะของการยืนยันหรือตรวจสอบว่า ผลการทดสอบโดยใช้โปรแกรมคอมพิวเตอร์ ยังคงให้คุณภาพต่าง ๆ ของข้อสอบและแบบทดสอบที่เหมือนหรือแตกต่างไปจากสภาพการสอบโดยใช้แบบทดสอบที่ตอบลงบนกระดาษเช่นที่มีมาก่อนหรือไม่อย่างไร เหตุนี้ ในการวิจัยครั้งนี้จึงมีวัตถุประสงค์หลัก 2 ประการคือ เพื่อพัฒนาโปรแกรมคอมพิวเตอร์เพื่อใช้ทดแทนการทดสอบแบบปกติ และเพื่อมุ่งศึกษาเปรียบเทียบประสิทธิภาพระหว่างการทดสอบแบบปกติกับการทดสอบทางโปรแกรมคอมพิวเตอร์ขึ้น

## ขั้นตอนและวิธีดำเนินการ

ขั้นตอนของการวิจัยเริ่มต้นขึ้น โดยผู้วิจัยได้พัฒนาโปรแกรมคอมพิวเตอร์ขึ้น 2 โปรแกรม เพื่อใช้กับการทดสอบกรณีคำถามเป็นข้อความ และกรณีคำถามปรากฏอยู่ในลักษณะของรูปภาพ โดยใช้โปรแกรม Visual Basic 6 จากนั้นผู้วิจัยได้พัฒนาแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนขึ้น 2 ฉบับคือ “แบบทดสอบความรู้พื้นฐานด้านการวัดผลประเมินผล” ใช้สอบกับโปรแกรมคอมพิวเตอร์ที่มีลักษณะคำถามเป็นข้อความ และ “แบบทดสอบความรู้พื้นฐานด้านการวิเคราะห์ข้อสอบ” ใช้สอบกับโปรแกรมคอมพิวเตอร์ที่มีลักษณะเป็นรูปภาพ แบบทดสอบแต่ละฉบับเป็นแบบทดสอบชนิด หัวตัวเลือกมีจำนวน 40 และ 30 ข้อตามลำดับ เพื่อเตรียมนำไปใช้กับนักศึกษาระดับปริญญาตรีสาขาการศึกษา มหาวิทยาลัยราชภัฏสุรินทร์ ภาคปลาย ปีการศึกษา 2552 จำนวน 69 คนที่ลงทะเบียนเรียนรายวิชา “การประเมินผลและการสร้างแบบทดสอบ” นักศึกษาทั้งหมดกำลังศึกษาในระดับปริญญาตรี ประกอบด้วยนักศึกษารายวิชาเอกต่าง ๆ คือ คณิตศาสตร์ พลศึกษา ดนตรี และนาฏศิลป์ โดยกลุ่มตัวอย่างทุกคนต้องทำแบบสอบทั้งสองฉบับ ในการทดสอบแต่ละฉบับจะแบ่งนักศึกษาเป็นสองกลุ่มย่อยจำนวนเท่า ๆ กัน ลักษณะการทดสอบ ผู้วิจัยจะสุ่มนักศึกษาครึ่งหนึ่งทำการทดสอบแบบปกติ อีกกลุ่มหนึ่งได้รับการทดสอบทางโปรแกรมคอมพิวเตอร์ และเมื่อทำการทดสอบกับแบบทดสอบอีกฉบับหนึ่งจะมีการสุ่มสลับกลุ่ม ให้ได้รับการทดสอบในลักษณะที่ต่างไปจากเดิม การทดสอบใช้เวลาในช่วงฉบับละไม่เกิน 50 นาที จากนั้นได้มีการบันทึกผล การทดสอบและนำผลมาวิเคราะห์ รวมทั้งได้ใช้แบบสอบถามปลายเปิดเพื่อสอบถามความคิดเห็นของนักศึกษาในด้านการทดสอบทางโปรแกรมคอมพิวเตอร์ ในช่วงท้ายของการทดสอบนักศึกษาแต่ละคน

## ผล การวิจารณ์ และข้อเสนอแนะ

### ผลการวิจัยที่สำคัญมีดังนี้

1. ค่าคะแนนเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และจำนวนนักศึกษาที่ได้รับการทดสอบในแบบทดสอบ “แบบทดสอบความรู้พื้นฐานด้านการวัดผลประเมินผล” เมื่อทดสอบโดยวิธีต่างกันคือ การทดสอบแบบปกติ และการทดสอบกับโปรแกรมคอมพิวเตอร์ ปรากฏค่าเป็น 22.63 6.27 33 และ 20.28 6.74 และ 36 ตามลำดับ และคะแนนเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และจำนวนนักเรียนที่ได้รับการทดสอบในแบบทดสอบ “แบบทดสอบความรู้พื้นฐานด้านการวิเคราะห์ข้อสอบ” เมื่อทดสอบโดยวิธีต่างกันคือ การทดสอบแบบปกติ และการทดสอบกับโปรแกรมคอมพิวเตอร์ ปรากฏค่าเป็น 17.57 6.56 35 และ 18.30 7.08 และ 33 ตามลำดับ การทดสอบ t independent test เพื่อทดสอบความแตกต่างของค่าเฉลี่ยของแบบทดสอบแต่ละฉบับที่ทำการทดสอบโดยวิธีการแตกต่างกัน พบว่า ค่าเฉลี่ยดังกล่าวไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

2. ในส่วนของการศึกษาความตรงของผลการทดสอบ เมื่อศึกษาเกี่ยวกับค่าคะแนนที่ได้จากการสอบ กับคะแนนเกณฑ์คือผลการทดสอบของกลุ่มตัวอย่างในการทดสอบตอนปลายเรียน พบว่า ในส่วนของ “แบบทดสอบความรู้พื้นฐานด้านการวัดผลประเมินผล” ที่ทดสอบโดยวิธีการต่างกันคือ การทดสอบแบบปกติ กับ การทดสอบกับโปรแกรมคอมพิวเตอร์ มีค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.89 และ 0.71 ตามลำดับ ในส่วนของ “แบบทดสอบความรู้พื้นฐานด้านการวิเคราะห์ข้อสอบ” ที่ทดสอบโดยวิธีการต่างกันคือ การทดสอบแบบปกติ กับ การทดสอบกับโปรแกรมคอมพิวเตอร์ มีค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.79 และ 0.84 ตามลำดับ ซึ่งค่าสัมประสิทธิ์สหสัมพันธ์ทุกค่ามีนัยสำคัญทางสถิติที่ระดับ 0.01

3 ค่าความเที่ยงเมื่อคำนวณโดย KR20 ของ “แบบทดสอบความรู้พื้นฐานด้านการวัดผลประเมินผล” เมื่อทดสอบโดยวิธีการต่างกัน คือการทดสอบแบบปกติ กับ การทดสอบกับโปรแกรมคอมพิวเตอร์ มีค่าเป็น 0.80 และ 0.82 ในส่วนของทดสอบ “แบบทดสอบความรู้พื้นฐานด้านการวิเคราะห์ข้อสอบ” ที่ทดสอบโดยวิธีการต่างกัน มีค่าเป็น 0.87 และ 0.82 ซึ่งค่า KR20 แต่ละคู่ดังกล่าวเมื่อทดสอบความแตกต่างของค่าดังกล่าวโดยใช้วิธีการที่ Feldt (Feldt 1969) ได้พัฒนาขึ้นพบว่าไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

4. ค่าความยากของ “แบบทดสอบความรู้พื้นฐานด้านการวัดผลประเมินผล” เมื่อทดสอบโดยการทดสอบแบบปกติมีจำนวนข้อสอบที่มีค่าความยากอยู่ในช่วง 0.20 – 0.80 เป็นจำนวน 35 ข้อ จากจำนวนข้อสอบทั้งหมด 40 ข้อ เมื่อทดสอบกับโปรแกรมคอมพิวเตอร์มีค่าความยากอยู่ในช่วง 0.20 - 0.80 เป็นจำนวน 36 ข้อ ในส่วนของ “แบบทดสอบความรู้พื้นฐานด้านการวิเคราะห์ข้อสอบ” เมื่อทดสอบโดยการทดสอบแบบปกติมีค่าความยากอยู่ในช่วง 0.20 – 0.80 เป็นจำนวน 29 ข้อจากจำนวนข้อสอบทั้งหมด 30 ข้อ เมื่อทดสอบกับโปรแกรมคอมพิวเตอร์มีค่าอยู่ในช่วง 0.20 - 0.80 เป็นจำนวน 28 ข้อ

5. ในส่วนของค่าอำนาจจำแนก “แบบทดสอบความรู้พื้นฐานด้านการวัดผลประเมินผล” เมื่อทดสอบโดยการทดสอบแบบปกติมีจำนวนข้อสอบที่มีค่าอำนาจจำแนก 0.20 – 1.00 เป็นจำนวน 25 ข้อ จากจำนวนข้อสอบทั้งหมด 40 ข้อ เมื่อทดสอบกับโปรแกรมคอมพิวเตอร์มีค่าอำนาจจำแนกอยู่ในช่วง 0.20 – 1.00 เป็นจำนวน 30 ข้อ ในส่วนของ “แบบทดสอบความรู้พื้นฐานด้านการวิเคราะห์ข้อสอบ” เมื่อทดสอบโดยการทดสอบแบบปกติมีค่าอำนาจจำแนกอยู่ในช่วง 0.20 - 1.00 เป็นจำนวน 25 ข้อจากจำนวนข้อสอบทั้งหมด 30 ข้อ เมื่อทดสอบกับโปรแกรมคอมพิวเตอร์มีค่าอำนาจจำแนกอยู่ในช่วง 0.20 – 1.00 เป็นจำนวน 25 ข้อ (ทั้งนี้เมื่อคำนวณโดยใช้สูตรการคำนวณค่าอำนาจจำแนกในรูปของความแตกต่างของค่าสัดส่วนระหว่างกลุ่มสูงกับกลุ่มต่ำ)

6. เมื่อเปรียบเทียบความความยากของข้อสอบรายคู่ของ “แบบทดสอบความรู้พื้นฐานด้านการวัดผลประเมินผล” เมื่อทดสอบโดยวิธีการต่างกัน โดยใช้วิธีการทดสอบความแตกต่างของค่าสัดส่วน พบว่า มีข้อสอบจำนวน 36 ข้อจากจำนวนทั้งหมด 40 ข้อที่มีค่าสัดส่วนความยากไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ในส่วนของแบบทดสอบ “แบบทดสอบความรู้พื้นฐานด้านการวิเคราะห์ข้อสอบ” ที่ทดสอบโดยวิธีการต่างกัน มีข้อสอบ 29 ข้อจากจำนวนทั้งหมด 30 ข้อ ที่มีค่าสัดส่วนความยากไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

7. ในการศึกษาค่าอำนาจจำแนก พบว่าในส่วนของ “แบบทดสอบความรู้พื้นฐานด้านการวัดผลประเมินผล” ค่าอำนาจจำแนกของข้อสอบที่ทดสอบโดยวิธีการต่างกันมีค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.44 ในส่วนของ “แบบทดสอบความรู้พื้นฐานด้านการวิเคราะห์ข้อสอบ” ค่าอำนาจจำแนกของข้อสอบที่ทดสอบโดยวิธีการต่างกันมีค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.48 ซึ่งค่าสัมประสิทธิ์สหสัมพันธ์ทั้งสองค่ามีนัยสำคัญทางสถิติที่ระดับ 0.01

จากผลการวิเคราะห์ข้อมูลทั้งหมด แสดงให้เห็นว่า การทดสอบโดยวิธีการปกติ กับ การทดสอบทางโปรแกรมคอมพิวเตอร์ให้ผลต่าง ๆ ในด้านคุณสมบัติของแบบทดสอบที่ดีที่สุดใกล้เคียงกันอย่างมาก

ในด้านความคิดเห็นต่อระบบการทดสอบทางโปรแกรมคอมพิวเตอร์ นักศึกษาเห็นว่าข้อดีคือ เป็นระบบการทดสอบที่ทันสมัย (ร้อยละ 34.78) มีการแจ้งผลการทดสอบให้ทราบในทันทีที่สอบเสร็จ (ร้อยละ 46.38) และเห็นว่าคะแนนที่ได้จากการทดสอบทางโปรแกรมคอมพิวเตอร์มีความเป็นธรรมเมื่อเทียบกับการทดสอบแบบปกติ (ร้อยละ 72.46)

ในขณะเดียวกัน การทดสอบทางโปรแกรมคอมพิวเตอร์ยังคงมีความไม่สมบูรณ์คือนักศึกษาตอบว่าตนเองมีภาวะความตื่นเต้น ไม่คุ้นเคยและกังวลต่อระบบการทดสอบ (ร้อยละ 50.72) มีความต้องการให้ปรับโปรแกรมให้ผู้ตอบสามารถทำข้อสอบแบบย้อนข้อกลับไปกลับมาได้ (ร้อยละ 30.43) มีความกังวลและรู้สึกกดดันด้านเวลา (ร้อยละ 23.19) และนักศึกษาส่วนหนึ่งรายงานว่าอุปกรณ์และระบบคอมพิวเตอร์มีปัญหาขณะทำการทดสอบ (ร้อยละ 11.59)

### วิจารณ์

จากผลการวิจัยทั้งหมดสะท้อนว่า คุณสมบัติต่าง ๆ ที่แบบสอบซึ่งมีคุณภาพจำเป็นต้องมีนั้น ในแบบทดสอบทั้งสองฉบับที่ผู้วิจัยได้พัฒนาขึ้นและนำมาทดสอบโดยวิธีการที่แตกต่างกันสองวิธี คือการทดสอบในระบบปกติ และการทดสอบทางโปรแกรมคอมพิวเตอร์ที่ผู้วิจัยพัฒนาขึ้น ให้คุณสมบัติที่ดีพอกัน กล่าวคือแบบทดสอบแต่ละฉบับที่ดำเนินการทดสอบโดยวิธีการต่างกัน มีค่าเฉลี่ยไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ มีค่าความตรงกับคะแนนเกณฑ์ที่กำหนดสูงและมีนัยสำคัญทางสถิติ ให้ค่าความเที่ยงสูงพอกันและไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ จำนวนข้อสอบในแบบสอบมีอัตราสูงถึงร้อยละ 90 ที่มีค่าสัดส่วนความยากของข้อสอบรายคู่ที่ทดสอบโดยวิธีการต่างกัน ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ รวมทั้งมีค่าอำนาจจำแนกของแบบทดสอบในฉบับเดียวกันที่ทดสอบโดยวิธีการต่างกันสูงอย่างมีนัยสำคัญทางสถิติ (การที่ค่าสถิติในส่วนนี้ไม่สูงมากนักนั้น เป็นเพราะว่า เป็นการหาค่าสัมประสิทธิ์ของค่าอำนาจจำแนกข้อสอบในแบบทดสอบฉบับหนึ่ง ๆ ซึ่งทดสอบโดยวิธีการต่างกัน ผู้เข้ารับการทดสอบเป็นคนละกลุ่ม ค่าสัมประสิทธิ์สหสัมพันธ์ที่คำนวณได้ในระดับ 0.44 และ 0.48 ที่คำนวณได้จึงนับได้ว่าสูงพอ)

### ข้อเสนอแนะ

จากผลการวิจัยครั้งนี้ถือได้ว่ามีข้อมูลเชิงประจักษ์ที่ยืนยันว่า การทดสอบโดยโปรแกรมคอมพิวเตอร์นั้น ให้คุณสมบัติต่าง ๆ ที่จำเป็นของแบบทดสอบและข้อสอบรายข้อที่ดีพอ เมื่อเทียบกับการทดสอบแบบปกติ จึงเป็นเครื่องยืนยันความมั่นใจในการนำเอาการทดสอบในระบบใหม่คือระบบการทดสอบทางโปรแกรมคอมพิวเตอร์มาใช้ในการจัดการศึกษาระดับปริญญาตรี ซึ่งในความรู้สึกของผู้ตอบระบบใหม่นี้มีข้อดี คือ ทำให้ทุกฝ่ายรับรู้ตัวตนเองตามความเปลี่ยนแปลงของเทคโนโลยีได้ทัน ทราบผลการทดสอบได้รวดเร็ว และมีความรู้สึกที่ดีคือยอมรับได้มีความเป็นธรรมของคะแนนที่ได้จากระบบการทดสอบดังกล่าว อย่างไรก็ตาม หากเป็นไปได้ ก็ควรจะมีการพัฒนาระบบการทดสอบทางโปรแกรมคอมพิวเตอร์ต่อไปในรายละเอียดต่าง ๆ ให้สมบูรณ์ยิ่งขึ้น เช่นการออกแบบหน้าจอให้สะดวกต่อการตอบของผู้เข้ารับการทดสอบสอบ รวมทั้งทำให้ผู้รับการทดสอบเกิดความรู้สึกเกิดความคุ้นเคย ไม่เกิดภาวะความเครียด หรือรู้สึกว่าตนเองถูกกดดันจากภาวะการทดสอบ เป็นต้น

### บรรณานุกรม

- Baker, F.B. (1984). "Technology and Testing : State of the Art and Trends for the Futere." Journal of Educational Measurement. Vol. 21, No. 4. : 399 – 406.
- Feldt, L.S. (1969). "A Test of the Hypothesis that Cronbach's Alpha or Kuder-Richardson Coefficient is the Same for Two Tests." Psychometrica. Vol. 34, No. 3 : 363 – 373.
- Luecht, R.M. and Nungester, R.J. (2000). Computer-Adaptive Sequential Testing. In Van der Linden, W.J. and Glas, C.A.W. (Eds). Computerized Adaptive Testing : Theory and Practice. 117 – 128. Dordrecht : Kluwer Academic Publishers.
- McArthur, D.L. and Choppin, B.H. (1984). "Computerized Diagnostic Testing." Journal of Educational Measurement. Vol. 21, No. 4 : 391-397.
- Umar, J. (1997). "Item Banking." In Keeves, J.P. (Ed.) "Educational Research, Methodology, and Measurement : An International Handbook" 2<sup>nd</sup> ed. 923 – 930. New York : Pergamon Press.
- Wright, B.D. and Bell, S.R. (1984). "Item Banks : What , Why, How." Journal of Educational Measurement. Vol. 21, No. 4 : 331- 345.