



รายงานการวิจัย เรื่อง

การศึกษาเปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกและ 4 ตัวเลือก
The Quality Comparison Study of 5 Choice and 4 Choice Tests

โดย
อาจารย์ ดร.สฤณีเทพ สุขแก้ว
อาจารย์ ดร.ศุภมาส ชุ่มแก้ว

การวิจัยครั้งนี้ได้รับทุนอุดหนุนการวิจัยประจำปี 2565
(ประเภทโครงการวิจัยสถาบันเชิงนโยบาย)
มหาวิทยาลัยสุโขทัยธรรมาธิราช

บทสรุปสำหรับผู้บริหาร

การวัดและประเมินผลการศึกษาในระบบทางไกลของมหาวิทยาลัยสุโขทัยธรรมมาธิราช คณะนั้นส่วนหนึ่งจะวัดผลจากการสอบประจำภาคของมหาวิทยาลัยสุโขทัยธรรมมาธิราช ส่วนใหญ่มหาวิทยาลัยใช้แบบทดสอบแบบเลือกตอบ และมีบางชุดวิชาที่มีข้อสอบอัตนัยร่วมด้วย ซึ่งการที่มหาวิทยาลัยสุโขทัยธรรมมาธิราชเลือกใช้แบบทดสอบแบบเลือกตอบเป็นหลัก ก็ด้วยเหตุผลที่ว่า มหาวิทยาลัยจัดการศึกษาในระบบเปิดซึ่งเปิดโอกาสให้นักศึกษาเข้าศึกษาโดยไม่มีชั้นเรียน ให้โอกาสกับทุกคน ทุกเพศ ทุกวัย และทุกสถานที่ ได้มีโอกาสเข้ามาศึกษา ดังนั้นในแต่ละชุดวิชาจึงมีนักศึกษาเลือกลงทะเบียนเป็นจำนวนมาก ข้อสอบที่เหมาะสมสำหรับใช้สอบกับคนจำนวนมากและต้องประกาศผลให้ได้ภายในเวลาอันจำกัดที่ไม่นานจนเกินไป ก็คือการใช้แบบทดสอบแบบเลือกตอบ โดยเลือกใช้แบบ 5 ตัวเลือก เพื่อลดอัตราการเดาให้เหลือน้อยที่สุด

จากมติที่ประชุมคณะกรรมการบริหารสำนักทะเบียนและวัดผลที่มุ่งดำเนินการตามตัวชี้วัดที่เกี่ยวข้องให้บรรลุเป้าหมายตามเกณฑ์ในคำรับรองการปฏิบัติราชการประจำปีงบประมาณ 2564 และสำหรับตัวชี้วัดที่ 3 ในมิติที่ 1 เรื่อง ระดับความสำเร็จของการปรับระบบการวัดและประเมินผลให้สอดคล้องกับทิศทางการพัฒนามหาวิทยาลัยสุโขทัยธรรมมาธิราชในอนาคต ผู้วิจัยจึงสนใจที่จะเป็นส่วนหนึ่งในการพัฒนาระบบการวัดและประเมินผลของมหาวิทยาลัยสุโขทัยธรรมมาธิราชด้วยการพัฒนาข้อสอบรูปแบบใหม่ที่มีคุณภาพ โดยลดตัวเลือกจาก 5 ตัวเลือกเป็น 4 ตัวเลือก เพื่อตอบสนองวิสัยทัศน์ในการพัฒนามหาวิทยาลัยสุโขทัยธรรมมาธิราชและพันธกิจของศูนย์วิจัยและพัฒนาแบบทดสอบ สำนักทะเบียนและวัดผลที่มีภารกิจในการวิจัยและพัฒนาแบบทดสอบของมหาวิทยาลัยสุโขทัยธรรมมาธิราชให้มีคุณภาพเพื่อนำไปใช้วัดและประเมินสมรรถนะของนักศึกษาเป็นลำดับต่อไป

โดยงานวิจัยนี้มีวัตถุประสงค์ดังนี้ 1) เปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก 2) เปรียบเทียบคะแนนและเปรียบเทียบคะแนนของนักศึกษาที่ได้จากการเปรียบเทียบคะแนนจากคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก และ 3) เปรียบเทียบค่าใช้จ่ายในการจัดทำแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก

ขั้นตอนการดำเนินการวิจัยในครั้งนี้มีรายละเอียดดังนี้ คือ การวิจัยครั้งนี้เป็นการวิจัยเชิงปริมาณ โดยใช้ข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ของนักศึกษาที่กำลังศึกษาในหลักสูตรระดับปริญญาตรีของมหาวิทยาลัยสุโขทัยธรรมมาธิราช ที่มีสิทธิ์สอบในภาคการศึกษาที่ 2/2562 ในรายวิชาที่ถูกเลือกมาศึกษาในงานวิจัยนี้เป็นข้อมูลตั้งต้น จากนั้นผู้วิจัยจึงทำการวิเคราะห์ข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก เพื่อคำนวณหาคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ด้วยการคำนวณหาค่าอำนาจจำแนกและใช้การวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบด้วยการทำให้เรียบเคอเนล แล้วใช้ผลที่ได้ในการตัดตัวเลือกที่ไม่มีประสิทธิภาพออก 1 ตัวเลือก เพื่อให้ได้แบบทดสอบ

ฉบับใหม่ที่เหลือตัวเลือกเพียง 4 ตัวเลือก แล้วจึงใช้วิธีการจำลอง (Simulation) สถานการณ์การตอบของผู้เข้าสอบโดยใช้แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก เพื่อจัดหาตัวเลือกสำหรับการแทนที่ตัวเลือกที่ตรงกับตัวเลือกที่ถูกตัดออกโดยใช้เกณฑ์การพิจารณาค่าอำนาจจำแนกและผลการวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบด้วยการทำให้เรียบเคอเนล โดยอาศัยข้อตกลงเบื้องต้นที่ผู้วิจัยกำหนดขึ้นในแต่ละกรณี ทั้งหมด 3 กรณี มาใช้ในการจัดหาตัวเลือกเพื่อการแทนค่าตัวเลือกที่ถูกตัดออกให้กับแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ในแต่ละข้อ จนครบตามจำนวนของผู้เข้าสอบแต่ละคนที่ต้องถูกตัดตัวเลือกนั้นออกไป แล้วจึงนำข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และข้อมูลผลการจำลองสถานการณ์การตอบแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ในแต่ละกรณีของการแทนค่าตัวเลือกที่ถูกตัดออก ทั้ง 3 กรณี มาทำการศึกษาคุณภาพของแบบทดสอบ ทำการปรับเทียบคะแนนเพื่อตอบสนองมาตรฐานที่ผู้วิจัยตั้งไว้ แล้วคำนวณค่าใช้จ่ายหลังจากที่มีการลดจำนวนตัวเลือกของแบบทดสอบ

ผลการวิจัยที่ได้ พบว่า 1. ผลการเปรียบเทียบคุณภาพของแบบทดสอบจากการวิเคราะห์ค่าความเที่ยง ความยาก และอำนาจจำแนก แยกเป็น 4 กรณีตามสมมติฐานการวิจัย กรณีที่ 1 แบบทดสอบแบบเลือกตอบ 5 ตัวเลือก พบว่าค่าความเที่ยงของแบบทดสอบแบบ 5 ตัวเลือกมีค่าเท่ากับ 0.81 กรณีที่ 2 แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) พบว่าค่าความเที่ยงของแบบทดสอบแบบ 4 ตัวเลือกมีค่าเท่ากับ 0.79 กรณีที่ 3 แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) พบว่าค่าความเที่ยงของแบบทดสอบแบบ 4 ตัวเลือกมีค่าเท่ากับ 0.79 กรณีที่ 4 แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (แทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) พบว่าค่าความเที่ยงของแบบทดสอบแบบ 4 ตัวเลือกมีค่าเท่ากับ 0.87 โดยในกรณีที่ 4 ค่าความเที่ยงของแบบทดสอบมีค่าสูงที่สุดในบรรดากรณีอื่นๆ ส่วนค่าความยากกับอำนาจจำแนกไม่แตกต่างกันในทุกกรณี 2. ผลการปรับเทียบคะแนนสอบของนักศึกษาระหว่างคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกและแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (3 กรณี ที่ผู้วิจัยศึกษา) พบว่า คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (ในทุกกรณี) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก 3. ผลการวิเคราะห์ค่าใช้จ่ายในการจัดทำแบบทดสอบ พบว่า ผลการจำลองข้อมูลในการวิเคราะห์ค่าใช้จ่ายของแบบทดสอบที่มี 4 ตัวเลือกเปรียบเทียบกับค่าใช้จ่ายของแบบทดสอบที่มี 5 ตัวเลือก เมื่อทำการลดจำนวนตัวเลือกในแบบทดสอบลง 1 ตัวเลือก ด้วยการตัดตัวเลือกที่ไม่มีคุณภาพออกไปตามสมมติฐานการวิจัยทั้ง 3 กรณี สามารถลดจำนวนหน้าของข้อสอบลงได้เป็นจำนวน 2-8 หน้า โดยคิดเป็นจำนวนหน้าที่ลดลงเฉลี่ยอยู่ที่ 3 หน้า ช่วยประหยัดค่าใช้จ่ายโดยเฉลี่ยอยู่ 11.54% ของค่าใช้จ่ายทั้งหมดจากการจัดทำแบบทดสอบที่มี 5 ตัวเลือก

ข้อเสนอแนะจากการวิจัยครั้งนี้ ประกอบด้วย ข้อเสนอแนะในการนำผลการวิจัยไปใช้ ได้แก่ 1. สถาบันอุดมศึกษาหรือหน่วยงานที่เกี่ยวข้องสามารถนำผลการวิจัยไปศึกษาเพื่อเป็นแนวทางในการกำหนดรูปแบบการวัดและประเมินผล กรณีที่ต้องการพิจารณาปรับลดจำนวนตัวเลือกในข้อสอบ

แบบเลือกตอบที่เหมาะสมสำหรับกลุ่มผู้สอบได้ และ 2. ผลการวิจัยที่ได้จากงานวิจัยเรื่องนี้เป็นสารสนเทศส่วนหนึ่งที่เป็นทางเลือกให้กับมหาวิทยาลัยสุโขทัยธรรมมาธิราชได้พิจารณาเพื่อเป็นแนวทางในการตัดสินใจเลือกปรับรูปแบบของการสอบจากเดิมที่มี 5 ตัวเลือก เป็น 4 ตัวเลือก โดยใช้ข้อมูลสนับสนุนทางวิชาการที่ผ่านกระบวนการวิจัย ทั้งนี้ผลการวิจัยที่ได้มาจากการวิเคราะห์ข้อมูลเพียง 1 ชุมวิชาเท่านั้น ข้อมูลดังกล่าวยังไม่มีความเป็นตัวแทนของทุกชุมวิชาที่เปิดสอนในมหาวิทยาลัยสุโขทัยธรรมมาธิราช การนำผลการวิจัยไปใช้ในการอ้างอิงจึงควรพิจารณาในกรณีนี้ด้วย และขอเสนอแนะในการวิจัยครั้งต่อไป ได้แก่ 1. การวิจัยครั้งต่อไปควรขยายผลไปที่การปรับลดจำนวนข้อสอบเพื่อหาจำนวนข้อสอบที่เหมาะสมในการใช้เป็นตัวแทนเพื่อวัดความรู้ในเนื้อหาแต่ละเรื่อง และ 2. ควรทำการวิจัยเช่นเดียวกันนี้ในชุมวิชาอื่น ๆ ให้ครบทุกสาขาวิชาที่เปิดสอนในมหาวิทยาลัยสุโขทัยธรรมมาธิราช เพื่อให้ได้ผลการวิจัยที่มีความเป็นตัวแทนของทุกชุมวิชา

ชื่อเรื่อง การศึกษาเปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกและ 4 ตัวเลือก

ชื่อผู้วิจัย อาจารย์ ดร.สฤณีเทพ สุขแก้ว และอาจารย์ ดร.ศุภมาส ชุมแก้ว

ปีที่แล้วเสร็จ 2566

บทคัดย่อ

งานวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) เปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก 2) ปรับเทียบคะแนนและเปรียบเทียบคะแนนของนักศึกษาที่ได้จากการปรับเทียบคะแนนจากคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก 3) เปรียบเทียบค่าใช้จ่ายในการจัดทำแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก การวิจัยครั้งนี้เป็นการวิจัยเชิงปริมาณ (Quantitative Research) โดยใช้ข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ของนักศึกษาที่กำลังศึกษาในหลักสูตรระดับปริญญาตรีของมหาวิทยาลัยสุโขทัยธรรมมาธิราช ที่มีสิทธิ์สอบในภาคการศึกษาที่ 2/2562 ข้อสอบเป็นแบบปรนัยหรือแบบเลือกตอบทั้ง 120 ข้อ ซึ่งเป็นผลการตอบแบบทดสอบของนักศึกษาจำนวน 1,966 คน คำนวณหาค่าอำนาจจำแนกและวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบด้วยการทำให้เรียบเคอเนลเพื่อนำค่าสถิติมาใช้พิจารณาประกอบการตัดตัวเลือกที่ไม่มีประสิทธิภาพออก 1 ตัวเลือก ทำให้ได้แบบทดสอบฉบับใหม่ที่เหลือตัวเลือกเพียง 4 ตัวเลือก จากนั้นใช้วิธีการจำลอง (Simulation) สถานการณ์การตอบของผู้เข้าสอบโดยใช้ตรงกับตัวเลือกที่ถูกตัดออกโดยใช้ โดยอาศัยข้อตกลงเบื้องต้นที่ผู้วิจัยกำหนดขึ้นในแต่ละกรณี ทั้งหมด 4 กรณี คือ กรณีที่ 1 แบบทดสอบ 5 แบบ 5 ตัวเลือก ไม่มีการตัดตัวเลือกหรือแทนค่าใด ๆ เป็นแบบทดสอบฉบับดั้งเดิม กรณีที่ 2 : แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย กรณีที่ 3 : แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย และกรณีที่ 4 : แทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ นำข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และข้อมูลผลการจำลองสถานการณ์การตอบแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ทั้ง 3 กรณี มาศึกษาคุณภาพของแบบทดสอบ ทำการปรับเทียบคะแนน และคำนวณค่าใช้จ่ายหลังจากที่มีการลดจำนวนตัวเลือกของแบบทดสอบ ผลการวิจัยพบว่า

1) ผลการเปรียบเทียบคุณภาพของแบบทดสอบจากการวิเคราะห์ค่าความเที่ยง ความยาก และอำนาจจำแนกแยกเป็น 4 กรณีตามสมมติฐานการวิจัย

กรณีที่ 1 แบบทดสอบแบบเลือกตอบ 5 ตัวเลือก พบว่าค่าความเที่ยง (raw_alpha) ของแบบทดสอบแบบ 5 ตัวเลือกมีค่าเท่ากับ 0.81

กรณีที่ 2 แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) พบว่าค่าความเที่ยง (raw_alpha) ของแบบทดสอบแบบ 4 ตัวเลือกมีค่าเท่ากับ 0.79

กรณีที่ 3 แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) พบว่าค่าความเที่ยง (raw_alpha) ของแบบทดสอบแบบ 4 ตัวเลือกมีค่าเท่ากับ 0.79

กรณีที่ 4 แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (แทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) พบว่าค่าความเที่ยง (raw_alpha) ของแบบทดสอบแบบ 4 ตัวเลือกมีค่าเท่ากับ 0.87

โดยในกรณีที่ 4 ค่าความเที่ยงของแบบทดสอบมีค่าสูงที่สุดในบรรดากรณีอื่นๆ ส่วนค่าความยากกับอำนาจจำแนกไม่แตกต่างกันในทุกกรณี

2) ผลการเปรียบเทียบคะแนนเฉลี่ยของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกและแบบทดสอบแบบเลือกตอบ 4 ตัวเลือกในแต่ละกรณี (จากทั้งหมด 3 กรณี คือ กรณีที่ 2-4) พบว่า คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (ในทุกกรณี) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

3) ผลการวิเคราะห์ค่าใช้จ่ายในการจัดทำแบบทดสอบ พบว่า ในกระบวนการจัดทำแบบทดสอบ จำนวน 1 หน้า มีค่าใช้จ่ายในการดำเนินการทำ เท่ากับ 50 บาท และเมื่อจำลองข้อมูลในการวิเคราะห์ค่าใช้จ่ายของแบบทดสอบที่มี 4 ตัวเลือกเปรียบเทียบกับค่าใช้จ่ายของแบบทดสอบที่มี 5 ตัวเลือก พบว่า การลดจำนวนตัวเลือกในแบบทดสอบลง 1 ตัวเลือก ด้วยการตัดตัวเลือกที่ไม่มีคุณภาพออกไปตามสมมติฐานการวิจัยทั้ง 3 กรณี สามารถลดจำนวนหน้าของข้อสอบลงได้เป็นจำนวน 2-8 หน้า โดยคิดเป็นจำนวนหน้าที่ลดลงเฉลี่ยอยู่ที่ 3 หน้า ช่วยประหยัดค่าใช้จ่ายโดยเฉลี่ยอยู่ 11.54% ของค่าใช้จ่ายทั้งหมดจากการจัดทำแบบทดสอบที่ 5 ตัวเลือก

คำสำคัญ แบบทดสอบแบบเลือกตอบ 5 ตัวเลือก แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก การเปรียบเทียบคะแนน คุณภาพแบบทดสอบ

Abstract

The objectives of this research were 1) to compare the quality of the 5-choice multiple choice test and the 4-choice multiple choice test; from a 5-choice multiple choice test and a 4-choice multiple choice test; 3) to compare the cost of preparing a 5-choice multiple choice test and a 4-choice multiple choice test. This research is quantitative by using data from the multiple 5-choice tests of the students who study at Sukhothai Thammathirat Open University. Who are eligible to take the exam in 2019 in semester 2, the exam is a multiple-choice or multiple-choice exam, all 120 questions, which are the results of the test from 1,966 students. The discriminant power was calculated and analyzed using test response theory with kernel smoothing. To take the statistical value into consideration in eliminating one ineffective option, resulting in a new version of the test with only four options remaining. Then, by using the simulation method, the candidate's answer situation is matched to the eliminated option using By relying on the preliminary agreement that the researcher defines in each case, all 4 cases are Case 1: A 5-type, 5-choice multiple choice test with no cut-off options or substitutes any values as the original test. Case 2: replaces the cut-off options at random. With the rest of the options, excluding the solution options. Case 3: Substituting the randomly eliminated option values. With the rest of the options, including the answer options, and Case 4: represents the cut-off options according to the test taker's proficiency. Take the data from the 5-choice multiple choice test and the 4-choice multiple choice test simulation in all 3 cases to study the quality of the test. Make a score comparison and calculate the cost after the number of test options has been reduced. The results showed that

1) The results of the comparison of the quality of the test from the analysis of reliability, difficulty, and discriminant power, separated into 4 cases according to the research hypothesis.

Case 1: 5-choice multiple choice test, found that the reliability (raw_alpha) of the 5-choice test was 0.81.

Case 2: Multiple-choice test with 4 options (instead of the randomly eliminated options with the other remaining options excluding the answer options), the reliability (raw_alpha) of the four-choice test was 0.79.

Case 3: Multiple-choice test with 4 options (instead of the randomly eliminated options With the other remaining options including the answer options), the reliability (raw_alpha) of the four-choice test was 0.79.

Case 4: Four-choice multiple-choice test (instead of the cut-off value according to the test taker's ability level) found that the reliability (raw_alpha) of the four-choice test was 0.87.

In case 4, the reliability of the test was the highest among the other cases. As for the difficulty and power of classification, they were not different in all cases.

2) The results of comparing the mean scores of students obtained from the 5-choice multiple choice test and the 4-choice multiple choice test in each case (out of all 3 cases, i.e. case 2-4). It was found that the scores of students who students score obtained the 4-choice multiple choice test (all cases) were more significant than the student's scores from the 5-choice test.

3) The analysis of the cost of the test preparation found that in the process of the test preparation of 1 page, the cost of conducting the test was 50 baht and when simulating the cost analysis of the 4-choice multiple choice test compared to the cost of the 5-choice multiple choice test, it was found that reducing the number of choices in the test by 1 choice by eliminating the choice. According to the assumptions of research in all 3 cases, the number of exam pages can be reduced by 2-8 pages, representing an average number of pages reduced to 3 pages, saving an average cost of 11.54% of the value. Total spending from the preparation of the 5-choice multiple choice test.

Keyword: 5-choice multiple choice test, 4-choice multiple choice test, Adaptive test, Quality of test

กิตติกรรมประกาศ

งานวิจัยนี้สำเร็จลุล่วงไปได้ด้วยดีเพราะได้รับการสนับสนุนและการส่งเสริมจากอาจารย์ที่ปรึกษาในโครงการวิจัย ผู้ช่วยศาสตราจารย์ ดร.ศศิธร ชูตินันทกุล ที่คอยให้คำแนะนำและผลักดันให้ผู้วิจัยค้นคว้าและใช้องค์ความรู้ทางวิธีวิทยาการวิจัยมาวิเคราะห์งานประจำที่ทำให้กลายเป็นงานวิจัยที่ให้สารสนเทศที่เป็นประโยชน์กับมหาวิทยาลัยสุโขทัยธรรมมาธิราช

ขอขอบพระคุณผู้บริหารของสำนักทะเบียนและวัดผล ตลอดจนบุคลากรของศูนย์วิจัยและพัฒนาแบบทดสอบ สำนักทะเบียนและวัดผล ที่ให้ความอนุเคราะห์กับผู้วิจัยในการเข้าถึงข้อมูลการวิจัยในครั้งนี้

ขอขอบพระคุณอาจารย์ ดร.สุทธิศานต์ ชุ่มวิจารณ์ และผู้ช่วยศาสตราจารย์ ดร.อนุสรณ์ เกิดศรี ที่ให้คำแนะนำและให้ความอนุเคราะห์ในการให้คำปรึกษาเกี่ยวกับการวิเคราะห์ข้อมูลด้วยโปรแกรม R

ขอขอบพระคุณสถานวิจัยสถาบัน มหาวิทยาลัยสุโขทัยธรรมมาธิราช ที่สนับสนุนและอนุมัติเงินอุดหนุนการวิจัยสถาบัน “ทุนอุดหนุนการวิจัย โครงการวิจัยสถาบันประจำปีงบประมาณ พ.ศ. 2565”

สารบัญ

	หน้า
บทสรุปผู้บริหาร	ก
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	ฉ
กิตติกรรมประกาศ	ช
สารบัญ	ฌ
สารบัญตาราง	ฎ
สารบัญภาพ	ฐ
บทที่ 1 บทนำ	1
วัตถุประสงค์ของการวิจัย	3
ขอบเขตของการวิจัย	3
ข้อตกลงเบื้องต้นที่ใช้ในงานวิจัย	5
สมมติฐานการวิจัย	6
นิยามศัพท์	6
บทที่ 2 การทบทวนวรรณกรรมที่เกี่ยวข้อง	7
ตอนที่ 1 แบบทดสอบแบบเลือกตอบ	7
ตอนที่ 2 คุณภาพของแบบทดสอบ	12
ตอนที่ 3 แนวคิดของทฤษฎีการตอบสนองข้อสอบด้วยการทำให้เรียบเคอเนล (Kernel Smoothing in Item Response Theory)	15
ตอนที่ 4 การปรับเทียบคะแนน (Equating)	17
ตอนที่ 5 งานวิจัยที่เกี่ยวข้อง	24
บทที่ 3 วิธีดำเนินการวิจัย	31
ข้อมูลที่ใช้ในการวิจัย	31
ขั้นตอนการดำเนินการวิจัย	33
การวิเคราะห์ข้อมูล	41
บทที่ 4 ผลการวิจัย	43
ตอนที่ 1 ผลการวิเคราะห์เปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก	43

ตอนที่ 2 ผลการวิเคราะห์การปรับเทียบคะแนนและเปรียบเทียบคะแนนของ นักศึกษาที่ได้จากการปรับเทียบคะแนนจากคะแนนที่ได้จากแบบทดสอบ แบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก	54
ตอนที่ 3 ผลการวิเคราะห์เปรียบเทียบค่าใช้จ่ายในการจัดทำแบบทดสอบแบบ เลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก	54
บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ	76
สรุปผลการวิจัย	76
อภิปรายผลการวิจัย	78
ข้อเสนอแนะจากการวิจัย	81
บรรณานุกรม	82
ภาคผนวก	85
ภาคผนวก ก หนังสือขอความอนุเคราะห์ข้อมูลวิจัย	86

สารบัญตาราง

	หน้า
ตารางที่ 2.1 ประเภทของความเที่ยง	14
ตารางที่ 3.1 ค่าอำนาจจำแนกที่ได้จากการตัดตัวเลือกของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก เป็นแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก	35
ตารางที่ 3.2 ข้อตกลงเบื้องต้นและวิธีการแทนค่าคำตอบที่ผู้วิจัยกำหนด	37
ตารางที่ 3.3 วิธีการแทนค่าคำตอบที่ผู้วิจัยกำหนด	38
ตารางที่ 4.1 ค่าความเที่ยงของแบบทดสอบทั้ง 4 กรณี	53
ตารางที่ 4.2 ผลการวิเคราะห์คุณภาพด้านความยากและอำนาจจำแนกด้วย PCA ของ แบบทดสอบทั้ง 4 กรณี	53
ตารางที่ 4.3 ผลการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จาก แบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณี แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวม ตัวเลือกที่เป็นเฉลย)	55
ตารางที่ 4.4 คะแนนสมมูลที่ได้จากการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่ นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัว อื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย)	56
ตารางที่ 4.5 ผลการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จาก แบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณี แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวม ตัวเลือกที่เป็นเฉลย)	58
ตารางที่ 4.6 คะแนนสมมูลที่ได้จากการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่ นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัว อื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย)	59
ตารางที่ 4.7 ผลการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จาก แบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ)	61

ตารางที่ 4.8	คะแนนสมมูลที่ได้จากการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่ นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ แบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถ ของผู้สอบ)	62
ตารางที่ 4.9	ค่าใช้จ่ายในกระบวนการจัดทำแบบทดสอบก่อนส่งเข้าโรงพิมพ์	72
ตารางที่ 4.10	ค่าใช้จ่ายในการจัดทำต้นฉบับแบบทดสอบปีการศึกษา 2562-2564	73
ตารางที่ 4.11	การเปรียบเทียบค่าใช้จ่ายในการจัดทำแบบทดสอบแบบ 5 ตัวเลือกและ 4 ตัวเลือก ในแต่ละภาคการศึกษา	75

สารบัญภาพ

	หน้า
ภาพที่ 2.1 การปรับเทียบคะแนนตามแนวนอน (Horizontal Equating)	18
ภาพที่ 2.2 การปรับเทียบคะแนนตามแนวตั้ง (Vertical Equating)	19
ภาพที่ 2.3 รูปแบบการเก็บข้อมูลกับผู้สอบกลุ่มเดียว	20
ภาพที่ 2.4 รูปแบบการเก็บข้อมูลกับผู้สอบกลุ่มเท่าเทียมกัน	21
ภาพที่ 2.5 รูปแบบการเก็บข้อมูลกับผู้สอบต่างกลุ่มโดยใช้แบบทดสอบร่วม	22
ภาพที่ 2.6 รูปแบบการเก็บข้อมูลกับผู้สอบต่างกลุ่มโดยมีผู้สอบร่วม	22
ภาพที่ 3.1 โค้งคุณลักษณะตัวเลือก (OCC) ของข้อสอบข้อที่ 1	36
ภาพที่ 4.1 ผลการวิเคราะห์ค่าความเที่ยง กรณีที่ 1 แบบทดสอบแบบ 5 ตัวเลือก	44
ภาพที่ 4.2 ผลการวิเคราะห์อำนาจจำแนกและความยากด้วย PCA กรณีที่ 1 แบบทดสอบ แบบ 5 ตัวเลือก	45
ภาพที่ 4.3 ผลการวิเคราะห์ค่าความเที่ยง กรณีที่ 2 แบบทดสอบแบบ 4 ตัวเลือก (กรณีแทน ค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือก ที่เป็นเฉลย)	46
ภาพที่ 4.4 ผลการวิเคราะห์อำนาจจำแนกและความยากด้วย PCA กรณีที่ 2 แบบทดสอบ แบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือ ตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย)	47
ภาพที่ 4.5 ผลการวิเคราะห์ค่าความเที่ยง กรณีที่ 3 แบบทดสอบแบบ 4 ตัวเลือก (กรณีแทน ค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่ เป็นเฉลย)	48
ภาพที่ 4.6 ผลการวิเคราะห์อำนาจจำแนกและความยากด้วย PCA กรณีที่ 3 แบบทดสอบ แบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือ ตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย)	49
ภาพที่ 4.7 ผลการวิเคราะห์ค่าความเที่ยง กรณีที่ 4 แบบทดสอบแบบ 4 ตัวเลือก (กรณีแทน ค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ)	51
ภาพที่ 4.8 ผลการวิเคราะห์อำนาจจำแนกและความยากด้วย PCA กรณีที่ 4 แบบทดสอบ แบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของ ผู้สอบ) ..	52

ภาพที่ 4.9	ผลการวิเคราะห์คะแนนเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของนักศึกษาที่ได้จาก แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่าง สุ่มด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย)	65
ภาพที่ 4.10	ผลการวิเคราะห์คะแนนเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของนักศึกษาที่ได้จาก แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่าง สุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย)	66
ภาพที่ 4.11	ผลการวิเคราะห์เพื่อทดสอบข้อตกลงเบื้องต้นของข้อมูลในเรื่องของการแจกแจง แบบปกติของข้อมูล	67
ภาพที่ 4.12	ผลการวิเคราะห์ข้อมูลด้วย Wilcoxon Signed Rank test	67
ภาพที่ 4.13	ผลการวิเคราะห์ข้อมูลเพื่อทดสอบสมมติฐานว่าคะแนนของนักศึกษาที่ได้จาก แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตาม ระดับความสามารถของผู้สอบ) มากกว่าคะแนนของนักศึกษาที่ได้จาก แบบทดสอบแบบเลือกตอบ 5 ตัวเลือก	68
ภาพที่ 4.14	การทดสอบข้อตกลงเบื้องต้นของข้อมูลในเรื่องของการแจกแจงแบบปกติของ ข้อมูล	68
ภาพที่ 4.15	ผลการวิเคราะห์ข้อมูลด้วยสถิติ Wilcoxon Signed Rank test	69

บทที่ 1

บทนำ

ความสำคัญและที่มาของปัญหา

การวัดและประเมินผลการศึกษาในระบบทางไกลของมหาวิทยาลัยสุโขทัยธรรมาธิราช ผู้สอนไม่สามารถประเมินความก้าวหน้าของผู้เรียนได้โดยตรง ต้องใช้การประเมินทางอ้อม โดยให้ผู้เรียนเป็นผู้ประเมินตนเองจากการทำแบบประเมินตนเองก่อนการศึกษาเอกสารการสอนแต่ละหน่วย เพื่อจะได้ทราบว่าตนเองมีความรู้ในหน่วยที่กำลังจะศึกษามากน้อยเพียงใด และเมื่อศึกษาเอกสารการสอนหน่วยนั้นจบแล้ว ผู้เรียนจะตรวจสอบว่าตนเองมีความรู้มากน้อยเพียงใด ด้วยการทำแบบประเมินตนเองหลังการศึกษา หรือการประเมินโดยการทำกิจกรรมที่อยู่ท้ายเรื่องนั้น ๆ รวมถึงการสอบประจำภาค ซึ่งทางมหาวิทยาลัยจะจัดขึ้นเพื่อประเมินว่านักศึกษาจะผ่านชุดวิชานั้นหรือไม่ การสอบประจำภาคของมหาวิทยาลัยสุโขทัยธรรมาธิราช ส่วนใหญ่ใช้แบบทดสอบแบบเลือกตอบ มีบางชุดวิชาที่มีข้อสอบอัตนัยด้วย (พิศเพลิน เขียวหวาน และคณะ, 2538, น. 12 – 13)

การที่มหาวิทยาลัยสุโขทัยธรรมาธิราชเลือกใช้แบบทดสอบแบบเลือกตอบเป็นหลัก ก็ด้วยเหตุผลที่มหาวิทยาลัยจัดการศึกษาในระบบเปิด เปิดโอกาสให้นักศึกษาเข้าศึกษาโดยไม่มีชั้นเรียน ให้โอกาสกับทุกคน ทุกเพศ ทุกวัย และทุกสถานที่ ได้มีโอกาสเข้ามาศึกษา ดังนั้นในแต่ละชุดวิชาจึงมีนักศึกษาเลือกลงทะเบียนเป็นจำนวนมาก ข้อสอบที่เหมาะสมสำหรับใช้สอบกับคนจำนวนมากและต้องประกาศผลให้ได้ภายในเวลาอันจำกัดที่ไม่นานจนเกินไป ก็คือการใช้แบบทดสอบแบบเลือกตอบ โดยเลือกใช้แบบ 5 ตัวเลือก เพื่อลดอัตราการเดาให้เหลือน้อยที่สุดซึ่งขวาล แพร์ตกุล กล่าวว่าข้อสอบแบบเลือกตอบ 3 ตัวเลือกเหมาะกับเด็กเล็ก ๆ ระดับอนุบาลหรือประถมศึกษาปีที่ 1 ประถมศึกษาปีที่ 2-6 ควรใช้ 4 ตัวเลือก เมื่อถึงขั้นสูงกว่านั้น จึงค่อยใช้ 5 ตัวเลือก และในการสร้างตัวลองที่ดีที่มีประสิทธิภาพจะต้องสร้างให้ดูเิน ๆ แล้ว จะเห็นว่าทุกตัวเลือกถูกหมด แต่ละตัวเลือกมีน้ำหนักที่จะถูกอยู่บ้าง และมีความสม่าเสมอเป็นเอกพันธ์ ไม่ผิดโง่งเพื่อพรางมิให้เด็กเดา ให้เด็กใช้ความคิดในการวินิจฉัยชี้ขาดว่าตัวเลือกใดเป็นตัวถูก ดังนั้นตัวลองที่ดีจะต้องมีโอกาสให้คนที่ไม่รู้จริงมาเลือก หากตัวลองใดสร้างขึ้นมาแล้วมีคนเลือกเลยหรือเลือกน้อย ถือว่าเป็นตัวลองที่ไม่ดี ไม่ได้ทำหน้าที่ของตัวลอง การมี 5 ตัวเลือกก็ไม่มีประโยชน์ ถือเสมือนมีแค่ 4 ตัวเลือกเท่านั้นเอง ถ้าแบบทดสอบใดใช้แบบ 5 ตัวเลือก แต่มี 1 ตัวเลือกที่ไม่มีคนเลือกหรือเลือกน้อยเป็นจำนวนมาก ก็ไม่จำเป็นที่จะต้องสร้างถึง 5 ตัวเลือก ให้เสียเวลาในการคิดตัวลอง และสิ้นเปลืองค่าใช้จ่าย ที่ผ่าน ๆ มาแบบทดสอบของมหาวิทยาลัยมีจำนวนไม่น้อยที่ตัวลองมีคนเลือกน้อยหรือไม่เลือกเลย (ต่ำกว่า 5%) ดังนั้นถ้าหากว่าคุณภาพของแบบทดสอบที่หากลดตัวเลือกลงมาเหลือ 4 ตัวเลือก โดยตัดตัวลองที่มีคนเลือกน้อยออก แล้วคุณภาพของแบบทดสอบไม่แตกต่างกันระหว่าง 4 ตัวเลือก กับ 5 ตัวเลือก มหาวิทยาลัยก็ควรปรับลดตัวเลือกของแบบทดสอบแบบเลือกตอบลงมาเหลือเพียง 4 ตัวเลือก ซึ่งแบบทดสอบมาตรฐานทั่วไปนิยมใช้ 4 หรือ 5 ตัวเลือกทั้งนั้น (ขวาล แพร์ตกุล, 2552) ถ้าตัวลองไม่สามารถจะลงผู้สอบได้แล้ว นอกจากจะทำให้เสียเวลาในการออกข้อสอบ ยังทำให้ค่าความตรง

(validity) และความเที่ยง (reliability) ของแบบทดสอบลดลงด้วย (ณัชชา มาลัย และทิพย์วัน เลิศชุมทรัพย์, 2550 อ้างถึงใน Marshall and Hales, 1972)

แบบทดสอบแบบเลือกตอบที่ใช้กันอยู่ทั่วไปมีทั้ง 4 และ 5 ตัวเลือก การมีจำนวนตัวเลือกน้อยเกินไป จะมีโอกาสเดาได้ถูกมากขึ้น ส่วนการใช้ตัวเลือก 5 ตัวเลือกจะทำให้ความเที่ยง (Reliability) มากที่สุด อย่างไรก็ตามในบางกรณี การหาตัวเลือกที่ 5 เป็นเรื่องที่ยากมากหรือไม่มีโอกาสเป็นไปได้ การเลือกใช้ 4 ตัวเลือกที่มีประสิทธิภาพก็จะเป็นทางเลือกที่ดีกว่า (อาณภาพ เลขกุล, 2559) สอดคล้องกับผลการวิจัยของวารุณี ปิตรวัชชัย (2514) ที่ศึกษาเปรียบเทียบค่าสถิติของแบบทดสอบชนิดเลือกตอบที่มีจำนวนตัวเลือกไม่เท่ากัน พบว่าแบบทดสอบ 5 ตัวเลือกและ 4 ตัวเลือกมีค่าความเที่ยงของแบบทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญ การลดจำนวนตัวเลือกของข้อสอบจาก 5 ตัวเลือกเป็น 4 ตัวเลือกนั้นไม่ทำให้ค่าอำนาจจำแนกของข้อสอบแต่ละข้อแตกต่างกันอย่างมีนัยสำคัญ และการลดจำนวนตัวเลือกของข้อสอบจาก 5 ตัวเลือกเป็น 4 ตัวเลือก ไม่ทำให้ระดับความยากของแบบทดสอบแตกต่างกันอย่างมีนัยสำคัญ

จากผลการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องดังกล่าว พบว่าการลดจำนวนตัวเลือกของข้อสอบจาก 5 ตัวเลือกเป็น 4 ตัวเลือก ไม่ทำให้คุณภาพของแบบทดสอบด้านความเที่ยงและคุณภาพของข้อสอบในด้านอำนาจจำแนกและความยากง่ายแตกต่างกันอย่างมีนัยสำคัญ ผู้วิจัยจึงต้องการจะยืนยันข้อค้นพบนี้อีกครั้งว่าจะเมื่อผู้วิจัยทำการวิจัยซ้ำอีกครั้งข้อค้นพบนี้จะเป็นจริงอยู่หรือไม่ เนื่องจากข้อจำกัดของการทำวิจัยของผู้วิจัย คือ ผู้วิจัยไม่สามารถสร้างแบบทดสอบเพื่อนำแบบทดสอบที่ผู้วิจัยสร้างขึ้นไปทำการสอบกับนักศึกษาได้โดยตรง เนื่องจากผู้วิจัยไม่มีบทบาทและหน้าที่ในส่วนของการออกข้อสอบ ดังนั้นการจะนำข้อสอบไปสอบกับนักศึกษาจริงก็ทำได้ยากมาก เนื่องจากการจัดการเรียนการสอนของมหาวิทยาลัยสุโขทัยธรรมมาธิราชเป็นการจัดการเรียนการสอนทางไกล โอกาสที่อาจารย์จะพบเจอนักศึกษาเพื่อเข้าไปทำการเก็บข้อมูลผลการสอบจากนักศึกษากระทำได้ยากและไม่สะดวกอย่างยิ่ง ดังนั้นผู้วิจัยจึงจำเป็นต้องใช้ข้อมูลสถิติซึ่งเป็นคะแนนผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกของนักศึกษามาเป็นข้อมูลตั้งต้น แล้วทำการจำลองสถานการณ์การสอบขึ้นมาว่า ถ้าตัดตัวเลือกของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกลงให้ได้แบบทดสอบฉบับใหม่ที่มีเพียง 4 ตัวเลือก แล้วคาดการณ์การตอบแบบทดสอบแบบเลือกตอบ 4 ตัวเลือกของนักศึกษา ตามข้อตกลงเบื้องต้นที่ผู้วิจัยกำหนดขึ้น 3 ข้อตกลงเบื้องต้น โดยสาเหตุที่ผู้วิจัยต้องสร้างข้อตกลงเบื้องต้นขึ้นมาเช่นนี้ เพราะผู้วิจัยไม่สามารถนำแบบทดสอบแบบเลือกตอบ 4 ตัวเลือกที่ผู้วิจัยสร้างขึ้นไปทำการสอบจริงกับนักศึกษาได้ดังที่กล่าวไว้แล้วข้างต้น เพื่อหาข้อสรุปว่าคุณภาพของแบบทดสอบที่ถูกตัดตัวเลือกให้เหลือน้อยลง จะยังคงมีคุณภาพแตกต่างกันกับแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกที่มหาวิทยาลัยสุโขทัยธรรมมาธิราชใช้อยู่หรือไม่ นอกจากนั้นยังต้องการยืนยันอีกว่าผลคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ที่ได้จากการจำลองและคาดการณ์สถานการณ์การตอบแบบทดสอบตามข้อตกลงเบื้องต้นของผู้วิจัยจะมากกว่าคะแนนสอบที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกหรือไม่ ผู้วิจัยจึงใช้การเปรียบเทียบคะแนนด้วยวิธีการที่ง่าย เป็นที่นิยมใช้ และมีคำสั่งในโปรแกรมสำเร็จรูปทางสถิติ คือ วิธีการเปรียบเทียบเชิงเส้น มาปรับคะแนนที่ได้จาก

แบบทดสอบแบบเลือกตอบ 5 ตัวเลือกและ 4 ตัวเลือกให้อยู่ในมาตราเดียวกัน แล้วทำการเปรียบเทียบคะแนนที่ได้จากการเปรียบเทียบว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือกมากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกหรือไม่ ตลอดจนอยากได้สารสนเทศอีกว่า หากมีการปรับลดตัวเลือกของแบบทดสอบลงจะทำให้ประหยัดค่าใช้จ่ายในการจัดทำแบบทดสอบมากน้อยเพียงใด

ผนวกกับมติที่ประชุมคณะกรรมการบริหารสำนักทะเบียนและวัดผลที่มุ่งดำเนินการตามตัวชี้วัดที่เกี่ยวข้องให้บรรลุเป้าหมายตามเกณฑ์ในคำรับรองการปฏิบัติราชการประจำปีงบประมาณ 2564 และสำหรับตัวชี้วัดที่ 3 ในมิติที่ 1 เรื่อง ระดับความสำเร็จของการปรับระบบการวัดและประเมินผลของ มสธ. ให้สอดคล้องกับทิศทางการพัฒนามหาวิทยาลัยสุโขทัยธรรมาธิราชในอนาคต ผู้วิจัยในฐานะอาจารย์นักวัดผลการศึกษาจึงสนใจที่จะเป็นส่วนหนึ่งในการพัฒนาระบบการวัดและประเมินผลของมหาวิทยาลัยสุโขทัยธรรมาธิราชด้วยการพัฒนาข้อสอบรูปแบบใหม่ที่มีคุณภาพ โดยลดตัวเลือกจาก 5 ตัวเลือกเป็น 4 ตัวเลือก เพื่อตอบสนองวิสัยทัศน์ในการพัฒนามหาวิทยาลัยสุโขทัยธรรมาธิราชและพันธกิจของศูนย์วิจัยและพัฒนาแบบทดสอบ สำนักทะเบียนและวัดผลที่มีภารกิจในการวิจัยและพัฒนาแบบทดสอบของ มสธ. ให้มีคุณภาพซึ่งเมื่อนำไปใช้วัดและประเมินสมรรถนะของนักศึกษาแล้ว จะทำให้ได้สารสนเทศสำหรับนำไปใช้ในการพัฒนาและส่งเสริมนักศึกษาให้สำเร็จการศึกษา กลายเป็นบัณฑิตที่มีความรู้ ความสามารถ และคุณธรรม จริยธรรม รับผิดชอบต่อสังคมและประเทศชาติต่อไป

วัตถุประสงค์ของการวิจัย

- 1) เพื่อเปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก
- 2) เพื่อเปรียบเทียบคะแนนและเปรียบเทียบคะแนนของนักศึกษาที่ได้จากการเปรียบเทียบคะแนนจากคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก
- 3) เพื่อเปรียบเทียบค่าใช้จ่ายในการจัดทำแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก

ขอบเขตของการวิจัย

ข้อมูลที่ใช้ในการวิจัยครั้งนี้แบ่งออกเป็น 2 ส่วน คือ ข้อมูลทุติยภูมิของผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และข้อมูลที่ได้จากการจำลองสถานการณ์การตอบแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก 3 กรณี ดังนี้

ส่วนที่ 1 เป็นข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ของนักศึกษาที่กำลังศึกษาในหลักสูตรระดับปริญญาตรีของมหาวิทยาลัยสุโขทัยธรรมาธิราช ที่มีสิทธิ์สอบในภาคการศึกษาที่ 2/2562 ในรายวิชาที่ถูกเลือกมาศึกษาในงานวิจัยนี้เป็นข้อมูลตั้งต้น ซึ่งเป็นข้อมูลทุติยภูมิ (Secondary Data) โดยแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก มีจำนวนข้อสอบทั้งหมด 120 ข้อ และเป็นข้อสอบแบบปรนัยหรือ

แบบเลือกตอบทั้ง 120 ข้อ ซึ่งเป็นผลการตอบแบบทดสอบของนักศึกษา จำนวน 1,966 คน ข้อมูลชุดนี้ ผู้วิจัยได้มาจากการทำหนังสือขอความอนุเคราะห์ผ่านไปยังท่านผู้อำนวยการสำนักทะเบียนและวัดผล ซึ่งได้รับความอนุเคราะห์ข้อมูลจากหัวหน้าหน่วยวิเคราะห์แบบทดสอบ ศูนย์วิจัยและพัฒนาแบบทดสอบ สำนักทะเบียนและวัดผล มหาวิทยาลัยสุโขทัยธรรมาธิราช

ส่วนที่ 2 เป็นข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ของนักศึกษาที่กำลังศึกษา ในหลักสูตรระดับปริญญาตรีของมหาวิทยาลัยสุโขทัยธรรมาธิราช ที่มีสิทธิ์สอบในภาคการศึกษาที่ 2/2562 ในรายวิชาที่ถูกเลือกมาศึกษาในงานวิจัยส่วนหนึ่งสำหรับข้อสอบบางข้อของนักศึกษาบางคนที่ไม่ได้ถูกตัด ตัวเลือก ผนวกเข้ากับข้อมูลอีกส่วนหนึ่งซึ่งเป็นข้อมูลที่ได้จากการจำลองสถานการณ์การตอบแบบทดสอบ แบบเลือกตอบ 4 เลือก ของนักศึกษา ซึ่งแบบทดสอบแบบเลือกตอบ 4 ตัวเลือกนี้ เกิดจากการตัดตัวเลือก จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกออก โดยอาศัยเกณฑ์การพิจารณาค่าอำนาจจำแนกและผลการ วิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบด้วยการทำให้เรียบเคอเนล โดยผู้วิจัยจะใช้วิธีการจำลองการ ตอบแบบทดสอบด้วยการแทนที่ตัวเลือกที่ถูกตัดออกตามข้อตกลงเบื้องต้นที่ผู้วิจัยตั้งขึ้น 3 กรณี โดยจะ แทนที่ตัวเลือกที่ถูกตัดออกให้กับข้อสอบแต่ละข้อที่ตรงกับเกณฑ์การพิจารณาตัดตัวเลือกข้างต้น ให้กับ ผู้เข้าสอบแต่ละคนที่ถูกตัดตัวเลือกทั้งไปเท่านั้น สำหรับผู้เข้าสอบที่ในข้อสอบข้อนั้น ๆ ไม่ถูกตัดตัวเลือก ผู้วิจัยก็ยังจะใช้ผลการตอบแบบทดสอบคงเดิม เหมือนกับคำตอบที่ผู้เข้าสอบตอบในแบบทดสอบแบบ เลือกตอบ 5 ตัวเลือก ตามข้อมูลทฤษฎีที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกที่เป็นข้อมูลตั้งต้น เนื่องจากผู้วิจัยจำลองวิธีการเก็บรวบรวมข้อมูลด้วยรูปแบบผู้สอบกลุ่มเดียวที่ไม่ได้รับการจัดให้สมดุล (Single – Group Uncounterbalanced Design) โดยใช้กลุ่มผู้สอบกลุ่มเดียวทำหน้าที่เป็นผู้สอบรวม กล่าวคือ ผู้สอบคนเดียวกันทำแบบสอบทั้ง 2 ฉบับ ดังนั้นระดับความสามารถของผู้สอบทั้ง 2 ฉบับ เท่ากัน แสดงว่าคำตอบของผู้สอบที่สอบด้วยแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก จะเหมือนกับคำตอบของผู้ สอบที่สอบด้วยแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

ข้อมูลในส่วนที่ 2 นี้ จะแบ่งออกเป็น 3 ชุดข้อมูล

ชุดที่ 1 เป็นข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ตามข้อมูลตั้งต้น สำหรับผู้เข้าสอบที่ไม่ถูกตัดตัวเลือกในข้อสอบข้อนั้น ๆ ออก ผนวกกับข้อมูลจากการจำลองผลการตอบ แบบทดสอบ 4 ตัวเลือก สำหรับข้อสอบข้อนั้น ๆ ที่ตรงกับตัวเลือกที่ถูกตัดออก ของผู้เข้าสอบคนนั้น ๆ ที่มี ข้อสอบที่ตรงกับตัวเลือกที่ถูกตัดออก การแทนที่ตัวเลือกที่ถูกตัดออกจะเป็นไปตาม ข้อตกลงเบื้องต้นของ ผู้วิจัยที่ว่า หากผู้สอบคนนั้นมีความสามารถสูงจะต้องตอบตัวเลือกที่เป็นเฉลยตั้งแต่ครั้งแรกที่ลงมือทำ ข้อสอบ

ชุดที่ 2 เป็นข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ตามข้อมูลตั้งต้น สำหรับผู้เข้าสอบที่ไม่ถูกตัดตัวเลือกในข้อสอบข้อนั้น ๆ ออก ผนวกกับข้อมูลจากการจำลองผลการตอบ แบบทดสอบ 4 ตัวเลือก สำหรับข้อสอบข้อนั้น ๆ ที่ตรงกับตัวเลือกที่ถูกตัดออก ของผู้เข้าสอบคนนั้น ๆ ที่มี ข้อสอบที่ตรงกับตัวเลือกที่ถูกตัดออก การแทนที่ตัวเลือกที่ถูกตัดออกจะเป็นไปตาม ข้อตกลงเบื้องต้นของ ผู้วิจัยที่ว่า ผู้สอบทุกคนมีสิทธิ์ที่จะเลือกตอบตัวเลือกใด ๆ ก็ได้ นั่นหมายความว่า ตัวเลือกแต่ละตัวมีโอกาส

ที่ผู้สอบจะเลือกตอบเท่า ๆ กัน ผู้สอบที่มีความสามารถสูงก็มีสิทธิ์จะเลือกตัวเลือกที่เป็นได้ทั้งตัวลงและเฉลย เช่นเดียวกับผู้สอบที่มีความสามารถต่ำ

ชุดที่ 3 เป็นข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ตามข้อมูลตั้งต้นสำหรับผู้เข้าสอบที่ไม่ถูกตัดตัวเลือกในข้อสอบข้อนั้น ๆ ออก ผนวกกับข้อมูลจากการจำลองผลการตอบแบบทดสอบ 4 ตัวเลือก สำหรับข้อสอบข้อนั้น ๆ ที่ตรงกับตัวเลือกที่ถูกตัดออก ของผู้เข้าสอบคนนั้น ๆ ที่มีข้อสอบที่ตรงกับตัวเลือกที่ถูกตัดออก การแทนที่ตัวเลือกที่ถูกตัดออกจะเป็นไปตาม ข้อตกลงเบื้องต้นของผู้วิจัยที่ว่า ผู้สอบที่มีความสามารถสูงย่อมต้องทำข้อสอบถูกทุกครั้ง ทุกกรณี แม้ว่าจะมีการปรับลดตัวเลือกของแบบทดสอบลงจาก 5 ตัวเลือก เหลือเพียง 4 ตัวเลือก ผู้สอบที่มีความสามารถสูงจะต้องทำข้อสอบข้อนั้น ๆ ถูกต้องเสมอ และสำหรับผู้ที่มีความสามารถต่ำก็มีโอกาสที่จะตอบตัวเลือกอื่น ๆ ที่ไม่ใช่ตัวเลือกที่เป็นเฉลย ด้วยความน่าจะเป็นในการเลือกตอบแต่ละตัวเลือกเท่า ๆ กัน เพราะคนที่มีความสามารถต่ำ แม้ว่าจะมีการปรับลดตัวเลือกของแบบทดสอบลงจาก 5 ตัวเลือก เหลือเพียง 4 ตัวเลือก ผู้สอบที่มีความสามารถต่ำ ก็ไม่มีโอกาสที่จะตอบข้อสอบข้อนั้นได้ถูกต้อง

โดยข้อมูลทั้ง 3 ชุดนี้ ถือว่าเป็นข้อมูลที่ได้จากการจำลองสถานการณ์การตอบแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ซึ่งมีจำนวนข้อสอบเท่ากับ 120 ข้อ และได้มาจากการตอบแบบทดสอบของผู้เข้าสอบ จำนวน 1,966 คน

ข้อจำกัดของการวิจัย

ในการวิจัยครั้งนี้ใช้แบบทดสอบจำนวน 1 ชุดวิชาเท่านั้นมาใช้ในการวิเคราะห์และจำลองข้อมูล เนื่องจากงานวิจัยนี้มีระยะเวลาในการวิจัยภายใต้กรอบของทุนวิจัยสถาบัน ในระยะเวลา 6 เดือน ดังนั้นผลการวิจัยที่ได้จากข้อมูลเพียง 1 ชุดวิชาจึงยังไม่มีความเป็นตัวแทนของชุดวิชาที่อยู่ในศาสตร์อื่น ๆ

ข้อตกลงเบื้องต้นที่ใช้ในงานวิจัย

ผู้วิจัยกำหนดข้อตกลงเบื้องต้นสำหรับการแทนค่าตัวเลือกที่ถูกตัดออก เพื่อให้ได้ข้อมูลจากการจำลองสถานการณ์การตอบแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ในกรณีต่าง ๆ 3 กรณี ตามข้อตกลงเบื้องต้นที่ผู้วิจัยกำหนดขึ้นดังต่อไปนี้

ข้อตกลงเบื้องต้นข้อที่ 1 : ผู้สอบที่มีความสามารถสูงจะต้องตอบตัวเลือกที่เป็นเฉลยตั้งแต่ครั้งแรกที่ลงมือทำข้อสอบ

ข้อตกลงเบื้องต้นข้อที่ 2 : ผู้สอบทุกคนมีสิทธิ์ที่จะเลือกตอบตัวเลือกใด ๆ ก็ได้ นั่นหมายความว่าตัวเลือกแต่ละตัวมีโอกาสที่ผู้สอบจะเลือกตอบเท่า ๆ กัน ผู้สอบที่มีความสามารถสูงก็มีสิทธิ์จะเลือกตัวเลือกที่เป็นได้ทั้งตัวลงและเฉลย เช่นเดียวกับผู้สอบที่มีความสามารถต่ำ

ข้อตกลงเบื้องต้นข้อที่ 3 : ผู้สอบที่มีความสามารถสูงย่อมต้องทำข้อสอบถูกทุกครั้ง ทุกกรณี แม้ว่าจะมีการปรับลดตัวเลือกของแบบทดสอบลงจาก 5 ตัวเลือก เหลือเพียง 4 ตัวเลือก ผู้สอบที่มีความสามารถสูงจะต้องทำข้อสอบข้อนั้น ๆ ถูกต้องเสมอ และสำหรับผู้ที่มีความสามารถต่ำก็มีโอกาสที่จะตอบตัวเลือก

อื่น ๆ ที่ไม่ใช่ตัวเลือกที่เป็นเฉลย ด้วยความน่าจะเป็นในการเลือกตอบแต่ละตัวเลือกเท่า ๆ กัน เพราะคนที่มีความสามารถต่ำ แม้ว่าจะมีการปรับลดตัวเลือกของแบบทดสอบลงจาก 5 ตัวเลือก เหลือเพียง 4 ตัวเลือก ผู้สอบที่มีความสามารถต่ำ ก็ไม่มีโอกาส

สมมติฐานการวิจัย

ผู้วิจัยกำหนดสมมติฐานการวิจัยดังนี้

สมมติฐานข้อที่ 1: คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

สมมติฐานข้อที่ 2 : คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

สมมติฐานข้อที่ 3 : คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

นิยามศัพท์

แบบทดสอบแบบเลือกตอบ หมายถึง แบบทดสอบที่จัดสร้างขึ้นตามรูปแบบของมหาวิทยาลัยสุโขทัย ธรรมาธิราช โดยให้ผู้สอบเลือกตัวเลือกสำหรับการตัดสินใจตอบข้อสอบแต่ละข้อ แบบทดสอบแบบเลือกตอบฉบับนี้จะมีตัวเลือก 5 ตัวเลือก ผู้วิจัยจะใช้แบบทดสอบแบบเลือกตอบ 5 ตัวเลือก เป็นแบบทดสอบตั้งต้น สำหรับนำมาคำนวณหาคุณภาพด้านอำนาจจำแนก เพื่อตัดตัวเลือกที่ไม่มีประสิทธิภาพออก 1 ตัวเลือก เพื่อให้ได้แบบทดสอบฉบับใหม่ที่เป็นแบบทดสอบแบบเลือกตอบ เช่นเดิม แต่มีจำนวนตัวเลือกลดลง 1 ตัวเลือก เหลือเพียง 4 ตัวเลือก

คุณภาพของแบบทดสอบ หมายถึง คุณลักษณะที่สะท้อนให้เห็นถึงความสามารถของแบบทดสอบในการวัดและประเมินความรู้ ความสามารถของนักศึกษามหาวิทยาลัยสุโขทัยธรรมาธิราช ประกอบด้วย ความเที่ยง ความยากง่าย และอำนาจจำแนก

ประโยชน์ที่คาดว่าจะได้รับ

มหาวิทยาลัยสุโขทัยธรรมาธิราชได้สารสนเทศเพื่อการตัดสินใจที่จะนำไปใช้เป็นแนวทางในการพัฒนาระบบการวัดและประเมินผล และได้แนวทางในการปฏิบัติงานรูปแบบใหม่ที่มีประสิทธิภาพและประสิทธิผล

บทที่ 2

การทบทวนวรรณกรรมที่เกี่ยวข้อง

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการศึกษาเปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกและแบบทดสอบแบบเลือกตอบ 4 ตัวเลือกที่เป็นประโยชน์ สามารถนำมาใช้ในการออกแบบการวิจัย การอภิปรายผลในประเด็นข้อค้นพบให้กว้างขวางและลึกซึ้งมากยิ่งขึ้น การนำเสนอการทบทวนวรรณกรรมในบทนี้ ได้เสนอรายละเอียด ประกอบด้วย 4 ตอน คือ ตอนที่ 1 แบบทดสอบแบบเลือกตอบ ตอนที่ 2 คุณภาพของแบบทดสอบ ตอนที่ 3 แนวคิดของทฤษฎีการตอบสนองข้อสอบด้วยการทำให้เรียบเคอเนล (Kernel Smoothing in Item Response Theory) ตอนที่ 4 การปรับเทียบคะแนน (Equating) และ ตอนที่ 5 งานวิจัยที่เกี่ยวข้อง ดังรายละเอียดต่อไปนี้

ตอนที่ 1 แบบทดสอบแบบเลือกตอบ

1.1 ความหมายของแบบทดสอบ

แบบทดสอบ เป็นเครื่องมือที่ใช้ในการวัดผลการศึกษาตามโรงเรียนหรือสถาบันต่าง ๆ จะใช้แบบทดสอบเป็นหลักในการวัดการเรียนรู้ หรือความงอกงามของผู้เรียน โดยมีผู้ให้ความหมายไว้หลายท่านด้วยกัน ดังนี้

ชวาล แพร์ตกุล (2507: 104) ให้ความหมายของแบบทดสอบไว้ว่า แบบทดสอบ หมายถึง ชุดของคำถามหรือกลุ่มงานใด ๆ ที่สร้างขึ้น เพื่อที่จะชักนำให้ผู้ถูกสอบแสดงพฤติกรรมอย่างใดอย่างหนึ่งออกมาให้ผู้สอบสังเกตหรือวัดได้

बारोलและเบอร์นาท (Baron and Bernard. 1967: 286) กล่าวว่า แบบทดสอบ คือ เครื่องมือที่เราออกแบบสร้างขึ้น ประกอบด้วย 2 ภาค คือ ภาคเร้า (Stimulus) ได้แก่ ข้อคำถามในแบบทดสอบ และภาคสนอง (Response) ได้แก่ การตอบสนองของเด็กรออกมาเป็นคำพูด ข้อความ หรือตัวเลข สามารถสังเกตและวัดได้

กล่าวโดยสรุป แบบทดสอบ หมายถึง ชุดของข้อคำถามที่กระตุ้นหรือเร้าให้ผู้ถูกสอบแสดงพฤติกรรมอย่างใดอย่างหนึ่งซึ่งสามารถวัดหรือสังเกตได้

1.2 ลักษณะของแบบทดสอบแบบเลือกตอบ

เมเรนส์และเลแมน (Mehrens and Lehmann. 1969: 206 – 277) กล่าวว่าแบบทดสอบแบบเลือกตอบ ประกอบด้วย 2 ส่วน คือ ส่วนที่เป็นปัญหา (Stem) และส่วนที่เป็นคำตอบ (Response or Option) สำหรับส่วนที่เป็นคำตอบมีถูกมีผิด ส่วนที่เป็นปัญหาอาจจะเป็นคำถามตรง ๆ หรือประโยคที่ไม่สมบูรณ์ก็ได้ หน้าที่ของผู้ตอบต้องเลือกตอบคำตอบที่ถูกเพียง 1 คำตอบเท่านั้น

ธอร์นไคต์ และเฮเกน (Thorndike and Hagen. 1960: 35 – 58) กล่าวว่าแบบทดสอบแบบเลือกตอบ ประกอบด้วย 2 ส่วน คือ ส่วนที่เป็นคำถาม (Stem) ซึ่งเป็นส่วนที่เป็นปัญหา และส่วนที่เป็นคำตอบที่เป็นไปได้ (Options) ส่วนที่เป็นคำถาม อาจเป็นประโยคที่ไม่สมบูรณ์หรือเป็นคำถามหนึ่งก็ได้

โกวิท ประวาลพุกษ์และสมศักดิ์ สันธุระเวช (2523: 60) และวิทยา วิชาสารณ (2530: 136 – 137) กล่าวว่าแบบทดสอบแบบเลือกตอบ (Multiple Choice) มีส่วนประกอบที่สำคัญ 2 ส่วน คือ 1) ส่วนที่เป็นตัวปัญหาหรือคำถามหรือตอนนำ (Problem or Question or Stem) หมายถึง ข้อความที่เขียนเป็นประโยคคำถามที่สมบูรณ์หรือไม่สมบูรณ์ ทำหน้าที่เป็นสิ่งที่เร้าเพื่อให้ผู้สอบตอบสนอง 2) ส่วนที่เป็นตอนตอบหรือตัวเลือก (Alternative หรือ Choice) ซึ่งตัวเลือกที่ถูก เรียกว่า ตัวถูก (Correct Choice) และตัวเลือกที่ไม่ถูกต้อง เรียกว่า ตัวลวงหรือตัวผิด (Fails หรือ Distractors) โดยที่ตัวเลือกนี้มีไว้เพื่อให้ผู้สอบพิจารณาเลือกเป็นคำตอบ

อานุกาพ เลชะกุล (2559: 3) กล่าวว่าแบบทดสอบแบบเลือกตอบมีส่วนประกอบที่สำคัญ 4 ส่วน ได้แก่ 1. โจทย์ (Stem) และคำถาม (Lead – in) เป็นส่วนที่ให้ข้อมูลเพื่อใช้ในการตอบคำถาม อาจมีรูปภาพ กราฟหรือแผนภูมิประกอบ 2. ตัวเลือก (Option หรือ Alternative) เป็นคำตอบที่ให้ผู้สอบเลือกตอบ โดยทั่วไปอาจมี 4 หรือ 5 ตัวเลือก การมีจำนวนตัวเลือกน้อยเกินไป จะมีโอกาสเดาได้ถูกมากขึ้น ตัวเลือก 5 ข้อ จะมีความเที่ยง (Reliability) มากที่สุด อย่างไรก็ตามในบางกรณี การหาตัวเลือกที่ 5 เป็นเรื่องที่ยากมากหรือไม่มีโอกาส เป็นไปได้ ก็อาจใช้เพียง 4 ตัวเลือก ส่วนการใช้ตัวเลือกมากกว่า 5 ข้อ ก็ไม่มีความจำเป็นและไม่ทำให้ อำนาจการจำแนกสูงขึ้น 3. คำตอบที่ถูกต้อง (Correct answer หรือ Key) เป็นตัวเลือกที่ถูกต้องที่สุด 4. ตัวลวง (Distractor) เป็นตัวเลือกที่เป็นคำตอบที่ไม่ถูกต้อง เพื่อลวงให้ผู้สอบเลือกตอบ การสร้างตัวลวงที่ดี จะทำให้ความเที่ยงของข้อสอบข้อนั้นดีขึ้น

ดังนั้นลักษณะของแบบทดสอบแบบเลือกตอบจะมีส่วนประกอบ 2 ส่วน คือ 1) ส่วนที่เป็นปัญหาหรือคำถาม และ 2) ส่วนที่เป็นคำตอบ หรือ ตัวเลือก ซึ่งสามารถแบ่งออกเป็น 2 ประเภท ได้แก่ ตัวเลือกที่ถูกต้อง เรียกว่า ตัวถูก และตัวเลือกที่ไม่ถูกต้อง เรียกว่า ตัวลวง

1.3 หลักการสร้างแบบทดสอบแบบเลือกตอบ

หลักการสร้างแบบทดสอบแบบเลือกตอบนั้น ผู้เชี่ยวชาญในการสร้างแบบทดสอบหลายท่าน เช่น ชวาล แพรัตกุล (2516: 116 – 190) อนันต์ ศรีโสภา (2515: 75 – 82) ล้วน สายยศและอังคณา สายยศ (2539: 94 – 107) เมเรนส์และเลห์แมน (Mehrens and Lehmann. 1984: 119 – 155) และไอเคน (Aiken. 1987: 44 – 58) ได้เสนอหลักการสร้างแบบทดสอบแบบเลือกตอบไว้ ดังนี้

- 1) ข้อสอบแต่ละข้อควรวัดการเรียนรู้ที่สำคัญ เป็นศูนย์รวมของปัญหา และถามได้ครอบคลุม
- 2) ใช้ประโยคคำถามหรือประโยคเติมคำสำหรับเป็นตัวคำถามนำ แต่รูปแบบของคำถามต้องวางเอาไว้ตอนท้ายของข้อความ

- 3) การใช้ข้อความในประโยคคำถาม ควรชัดเจนเหมาะสมอยู่ในระดับการอ่านของผู้สอบ แต่ควรหลีกเลี่ยงการสร้างข้อคำถามที่ใช้คำถามต่าง ๆ จากคำถามที่ใช้คำถามตรง ๆ จากตำราเรียน
- 4) ตัวข้อสอบควรอ่านได้ใจความ มีระเบียบ เข้าใจง่าย และชัดเจนที่สุด
- 5) ควรหลีกเลี่ยงข้อความที่เป็นปฏิเสธในตัวคำถาม
- 6) ไม่ควรใช้คำกำกวม หรือคำแปลง ในประโยคตอนนำและในตัวเลือก
- 7) ควรเขียนตอนนำและตัวเลือกให้อยู่คนละบรรทัด และเขียนหมายเลขหน้าข้อให้เหมาะสม
- 8) ควรหลีกเลี่ยงข้อสอบให้แสดงความคิดเห็น
- 9) ข้อสอบแต่ละข้อควรเป็นอิสระจากกัน
- 10) ข้อสอบแต่ละข้อควรมีตัวเลือกที่ถูกที่สุดเพียงตัวเดียว
- 11) เขียนตัวเลือกให้มีเนื้อหา แบบฟอร์ม และโครงการเป็นเอกพันธ์
- 12) หลีกเลี่ยงตัวเลือกที่เหลื่อมล้ำกัน
- 13) ตัวลวงทุกตัวมีเหตุผลที่อาจเป็นไปได้ ดึงดูดความสนใจของผู้ตอบที่ไม่รู้จริง
- 14) ใช้ตัวเลือก “ถูกต้องทั้งหมด” หรือ “ผิดทั้งหมด” ให้เหมาะสม
- 15) ควรใช้ 4 หรือ 5 ตัวเลือก แต่ก็สามารถสร้างให้มี 2 หรือ 3 ตัวเลือกก็ได้ เพราะข้อสอบ 3 ตัวเลือกใช้ได้เหมาะสมกว่าแบบ 4 หรือ 5 ตัวเลือก สำหรับเด็กในระดับชั้นต่ำ
- 16) ควรเรียงตัวเลือกที่เป็นจำนวน โดยเรียงจากน้อยไปหามาก
- 17) ควรกระจายตำแหน่งตัวถูกในลักษณะสุ่ม
- 18) พิจารณาจำนวนข้อให้เหมาะสมกับเวลา
- 19) หลีกเลี่ยงการใช้ข้อสอบแบบเลือกตอบร่วมกับข้อสอบรูปแบบอื่น ๆ ในแบบทดสอบฉบับเดียวกัน
- 20) ควรใช้แบบฟอร์มข้อสอบที่มีประสิทธิภาพ

1.4 หลักการสร้างตัวเลือกแบบทดสอบแบบเลือกตอบ

ในการสร้างแบบทดสอบแบบเลือกตอบถือว่าการสร้างตัวเลือกเป็นสิ่งสำคัญที่สุดอย่างหนึ่ง เนื่องจากคุณภาพของข้อสอบชนิดนี้จะมากหรือน้อยขึ้นอยู่กับวิธีการเขียนตัวเลือกเป็นประการสำคัญ การเขียนตัวเลือกที่ประกอบด้วยตัวถูก และตัวลวงนั้น ตัวลวงที่ดีจะช่วยให้ข้อสอบมีคุณภาพ (Adams. 1964: 340) ถ้าตัวลวงไม่สามารถจะลวงผู้สอบได้แล้ว นอกจากจะทำให้เสียเวลาในการออกข้อสอบแล้ว ยังทำให้ความตรงและความเที่ยงของแบบทดสอบลดลงด้วย (Marshall and Hales. 1972: 62) ดังนั้นในการสร้างข้อสอบชนิดนี้จำเป็นต้องใช้ตัวลวงให้ใกล้เคียงกับคำตอบมากที่สุด เพื่อให้นักเรียนที่ไม่มีความรู้จริงเลือกตอบตัวลวงเหล่านี้กระจายออกไปทุกตัว ซึ่งในการสร้างตัวเลือกของแบบทดสอบแบบเลือกตอบนี้ มีผู้ให้ข้อเสนอแนะไว้ ดังนี้

ธอร์นไดค์ (Thorndike. 1971: 113 – 114) ได้กล่าวถึงตัวเลือกของแบบทดสอบแบบเลือกตอบไว้ ดังนี้

- 1) ตัวเลือกที่ใช้ควรเหมาะสม สมเหตุสมผล และสอดคล้องกัน
- 2) ตัวเลือกทุกตัวต้องสอดคล้องกับคำตอบ
- 3) ให้ตัวลวงทุกตัวดึงดูดความสนใจของผู้ตอบที่ไม่รู้จริง
- 4) หลีกเลี่ยงตัวเลือกที่หลอกล้ำกัน
- 5) ไม่ควรใช้คำว่า “ถูกทุกข้อ” หรือ “ผิดทุกข้อ”
- 6) เรียงลำดับตัวเลือกตามลำดับตรรกวิทยา (Logical order)
- 7) ไม่เขียนตัวเลือกเป็นลักษณะคล้ายแบบถูกผิด

ชวาล แพร์ตกุล (2516: 75 – 93) ได้ให้ข้อคิดเห็นเกี่ยวกับตัวเลือกของแบบทดสอบแบบเลือกตอบ ดังนี้

- 1) แบบทดสอบชนิดนี้สร้างได้ง่ายและใช้ได้กับทุกระดับชั้น เป็นรูปแบบคำถามสำหรับวัดความแม่นยำในหลักวิชา และข้อเท็จจริงต่าง ๆ
- 2) ข้อสำคัญของการเขียนข้อสอบชนิดนี้มีอยู่ว่าจะต้องพยายามให้ตัวเลือกทุกตัวเป็นเอกพันธ์กับตัวคำตอบถูกเสมอ
- 3) ควรระวังเรื่องการใช้ปูจณาสรพนามให้ถูกต้อง เหมาะสม เช่นคำว่า ใคร อะไร ที่ไหน เมื่อใด เท่าใด ฯลฯ สิ่งเหล่านี้ควรต้องใช้ถามเกี่ยวกับ คน สิ่งของ สถานที่ เวลา จำนวน ฯลฯ
- 4) ชนิดตัวเลือกที่ดีที่สุด จะต้องกำหนดไว้ในตัวคำถามเลยว่า “ดีที่สุด ในแง่ใด ด้านใด” โดยหลักวิชาหรือตามคติและความเห็นของใคร ของตำราเล่มใดไว้ด้วย
- 5) ควรระวังสิ่งที่มีผลล่อกันบ่อย ๆ คือ อย่าให้ตัวถูกมีความยากมากกว่าตัวอื่น ๆ จนสะดุดตา แต่ถ้ามีความจำเป็นจะต้องยาวเช่นนั้นจริง ๆ ก็ควรเรียงตัวเลือกตามลำดับความสั้นยาวเสียให้หมดทุกข้อ
- 6) แบบตัวเลือกที่ดีที่สุด ห้ามใช้ตัวเลือกปลายปิดประเภท “ที่กล่าวมาถูกหมด” หรือ “ถูกหมดทุกข้อ” เป็นอันขาด เพราะตัวเลือกปลายเปิดนี้ จะกลายเป็นตัวคำตอบที่ถูกต้องที่สุดของคำถามชนิดนี้ไปทันที
- 7) ชนิดตัวเลือกใกล้เคียงมีข้อที่ต้องระวังอยู่เพียงประการเดียว คือ ห้ามลงท้ายด้วยตัวเลือกปลายเปิด ประเภทไม่มีคำตอบถูก ที่กล่าวมาผิดหมด หรือผิดหมดทุกตัวเลือก เพราะตัวเลือกปลายเปิดนี้จะกลายเป็นคำตอบถูกไปเสียเอง เนื่องจากตัวเลือกทุกตัวเลือกที่ให้ไว้นั้น ต่างก็ยังเป็นคำตอบที่ไม่ถูกต้องแท้ ยังมีความคลาดเคลื่อนปนแทรกอยู่ด้วยทั้งนั้น

ชูศักดิ์ ศิริรุ่งพันธ์ (2559) กล่าวว่า การเขียนตัวเลือกส่วนใหญ่จะมีปัญหาเกี่ยวกับการเขียนตัวเลือกไม่ครบตามจำนวนที่ต้องการ หรือเขียนตัวเลือกได้ไม่ดีพอ หรือในบางครั้งเกิดความผิดพลาด ทำให้ข้อสอบไม่มีตัวถูก หรือมีตัวถูกมากกว่าหนึ่งตัว สำหรับหลักและวิธีการเขียนตัวเลือกของข้อสอบเลือกตอบ ได้แก่

1) ตัวเลือกต้องมีความเป็นเอกพันธ์ (Homogeneous) โดยเขียนตัวเลือกให้มีคุณสมบัติ
ดังนี้

1.1 เป็นเรื่องราวเดียวกัน เป็นพวกเดียวกันหรือประเภทเดียวกัน ถ้ากล่าวถึงเรื่องใดก็ให้เป็นเรื่องทำนองนั้นทุกตัว เช่น ถ้าเป็นช็อกก็เป็นช็อกให้หมด เป็นประโยชน์ก็ต้องเป็นประโยชน์ทั้งหมด ทั้งนี้ให้ยึดตัวถูกเป็นหลักก่อน แล้วจึงเขียนตัวลวงให้สอดคล้องใกล้เคียงกับตัวถูกนั้น

1.2 มีทิศทางเดียวกัน คำถามที่ถามอย่างมีทิศทาง ได้แก่ คำถามที่ถามเกี่ยวกับประโยชน์ คุณค่า ข้อดี (ทิศทางบวก) หรือ โทษ ข้อเสีย จุดอ่อน (ทิศทางเป็นลบ) ถ้าคำถามถามในทางบวก ตัวเลือกทุกตัวก็ต้องเป็นในทางบวก ตรงข้ามถ้าถามโทษหรือทางลบก็ต้องเขียนตัวเลือกให้เป็นโทษให้หมด

1.3 มีโครงสร้างสอดคล้องกัน หมายถึง ตัวเลือกทุกตัวต้องมีโครงสร้างในลักษณะเดียวกัน ถ้าเป็นคำหรือวลี ก็ต้องเป็นทุกตัวเลือก ถ้าเป็นประโยคก็ควรเป็นประโยคทุกตัวเลือก

2) ใช้ตัวเลือกที่เป็นไปได้ (Possible Choice) ตัวเลือกทุกตัวโดยเฉพาะอย่างยิ่งตัวลวง ควรเขียนให้มีทางเป็นไปได้จริงโดยทั่วไป นั่นคือต้องไม่ใช่ตัวลวงที่ผิดอย่างชัดเจน

3) เขียนตัวเลือกในกรณีทีโจทย์เป็น คำถามปฏิเสธไม่ควรมีตัวเลือก “ไม่มีข้อใดถูก” “ถูกทุกข้อ” “ข้อ 1 และ 2 ถูก”

4) เขียนตัวเลือกให้อิสระขาดจากกัน พยายามอย่าให้ตัวเลือกตัวใดตัวหนึ่งเป็นส่วนหนึ่งหรือส่วนประกอบของตัวอื่น ต้องให้แต่ละตัวเป็นอิสระจากกันและกันอย่างแท้จริง

5) ใช้ภาษาที่รัดกุมชัดเจน ไม่ควรใช้ถ้อยคำสำนวนที่ยืดยาว หรือใช้ข้อความซ้ำๆ กันในแต่ละตัวเลือก ทั้งนี้ยกเว้นคำที่แสดงเหตุผลหรือคำที่บ่งบอกอาการนาม เช่น เพราะ การ ความ ไม่ถือว่าซ้ำซ้อน รวมทั้งหน่วยต่างๆ ในทางคณิตศาสตร์หรือวิทยาศาสตร์ เป็นต้น ถ้ามีข้อความซ้ำๆ ในตัวเลือกให้นำมาไว้ในโจทย์จะทำให้คำถามชัดเจน และประหยัดเวลาในการอ่านตัวเลือก

6) ควรเรียงลำดับตัวเลือกที่เป็นตัวเลข ในกรณีที่ตัวเลือกเป็นตัวเลข โดยเฉพาะในทางคณิตศาสตร์ หรือวิทยาศาสตร์ หรือที่เกี่ยวกับวันเดือนปี ควรเรียงลำดับตัวเลือกจากเลขน้อยไปมาก หรือจากมากไปน้อย ทั้งนี้เพื่อช่วยให้นักศึกษาสะดวกในการเปรียบเทียบค่าของตัวเลขเหล่านั้น

7) ควรใช้ตัวเลือกที่นักศึกษารู้จักและเข้าใจ นอกจากตัวคำถามจะถามในสิ่งที่ใกล้ตัวเกี่ยวข้องกับชีวิตประจำวันแล้ว ตัวเลือกที่ใช้ก็เช่นกัน ต้องใช้สิ่งที่นักศึกษาเข้าใจและรู้จัก

8) หลีกเลี่ยงการแนะนำคำตอบ ไม่ควรใช้ข้อสอบที่นักศึกษาสามารถใช้ไหวพริบหรือใช้เพียงการสังเกตก็สามารถหาคำตอบได้ ทั้งๆ ที่มีความรู้ไม่มากนัก ทั้งนี้เพราะตัวข้อสอบนั้นชี้แนะคำตอบถูกให้นักศึกษาสังเกตเห็นได้ การแนะนำคำตอบอาจเกิดขึ้นหลายกรณี ซึ่งจะต้องพึงระวังในเรื่องต่อไปนี้

(1) อย่าให้คำถามข้อแรกๆ แนะนำคำตอบข้อหลังๆ หรือข้อหลังๆ แนะนำคำตอบข้อแรกๆ

(2) อย่าให้ตัวถูกมีคำซ้ำกับคำถาม

(3) หลีกเลียงคำแนะนำ เช่น บางครั้ง บ้าง อาจจะ เสมอๆ ไม่เคย ทั้งหมด ไม่มี
เท่านั้น

(4) อย่าถามเรื่องที่นักศึกษาคล่องปากหรือเคยชินเกินไป

(5) หลีกเลียงการถาม พ.ศ. จำนวนขั้นตอน จำนวนประเภท

(6) อย่าใช้ตัวเลือกที่สั้นยาวต่างกันมาก โดยปกติข้อสอบเลือกตอบ ถ้าเขียน
ตัวเลือกโดยไม่พิจารณาให้ถี่ถ้วนแล้ว ตัวถูกมักจะยาวกว่าตัวลวง นักศึกษาสามารถคาดเดาโดยเลือก
ตัวเลือกที่ยาว จึงเป็นสิ่งที่ควรระวังและพยายามแก้ไขให้ตัวเลือกมีความสั้น ยาวพอๆ กัน โดยการตัดคำ
บางคำที่ไม่จำเป็นออก หรือเพิ่มคำขยายเพื่อปรับให้ยาวพอๆ กัน ในกรณีที่ตัวเลือกใช้คำเฉพาะ หรือมี
ความรัดกุมแจ่มชัดดีแล้ว ไม่สามารถปรับแก้ให้แต่ละตัวมีความยาวพอๆ กันได้ ให้ใช้วิธีเรียงตัวเลือกจากตัว
ที่สั้นไปหาตัวที่ยาว

(7) ควรกระจายตำแหน่งตัวถูก ข้อสอบที่มีตัวถูกซ้ำตำแหน่งกันหรือเวียนกันเป็น
ตัวถูกอย่างมีระบบ ทำให้ผู้สอบคาดเดาได้ว่า คำตอบข้อต่อไปควรเป็นตัวเลือกใด วิธีการวางตำแหน่งตัวถูก
โยกย้ายสลับเปลี่ยนตำแหน่งตัวถูกของบางข้อ และต้องไม่เรียงตำแหน่งตัวถูกให้เป็นระบบที่แน่นอนตายตัว

(8) ตัวเลือกที่เป็นตัวถูกไม่ควรใช้ข้อความที่ลอกมาจากเอกสารการสอนเพราะ
ทำให้ตัวเลือกนั้นมองดูดีกว่าตัวลวง ทำให้ผู้สอบจำได้ว่าเคยเห็นมาก่อน

สรุป แบบทดสอบที่ผู้วิจัยเลือกนำมาใช้ในงานวิจัยครั้งนี้จะมีลักษณะเป็นแบบทดสอบแบบ
เลือกตอบ 5 ตัวเลือก ที่มีการตัดเลือกที่ไม่มีประสิทธิภาพออกในทุกข้อของแบบทดสอบแบบเลือกตอบ 5
ตัวเลือก ให้เหลือแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก แล้วจึงใช้วิธีการแทนค่าตามข้อตกลงที่ผู้วิจัย
สร้างขึ้น เพื่อศึกษาคุณภาพของแบบทดสอบเป็นลำดับต่อไป

ตอนที่ 2 คุณภาพของแบบทดสอบ

สำหรับในงานวิจัยนี้จะขอพิจารณาคุณภาพของแบบทดสอบและข้อสอบแบบเลือกตอบ โดยอาศัย
เกณฑ์ ได้แก่ ความยาก อำนาจจำแนก และความเที่ยง ซึ่งศิริชัย กาญจนวาสี (2556) ได้กล่าวถึง
รายละเอียดของเกณฑ์ต่าง ๆ ดังนี้

2.1 ความยาก

ความยาก (p) (Level of difficulty of the items) หมายถึง สัดส่วนของจำนวนคนที่ตอบ
ข้อสอบข้อนั้นถูก เช่น ข้อสอบข้อหนึ่งมีคนตอบ 100 คน ปรากฏว่าตอบถูกเพียง 30 คน แสดงว่าข้อสอบ
ข้อนั้นมีระดับความยาก (p) เท่ากับ 0.3 (หรือ 30%) เป็นต้น ดังนั้นระดับความยากของข้อสอบจึงมีค่า
ตั้งแต่ 0 ถึง 1.0 ถ้าข้อสอบข้อใดมีคนตอบถูกมาก p จะมีค่าสูง (เข้าใกล้ 1) แสดงว่าข้อนั้นง่าย ในทาง
ตรงกันข้ามถ้าข้อสอบข้อใดมีคนตอบถูกน้อย p จะมีค่าต่ำ (เข้าใกล้ 0) แสดงว่าข้อนั้นยาก โดยทั่วไป
ข้อสอบที่มีค่า p ระหว่าง 0.2 – 0.8 ถือว่าเป็นข้อสอบที่มีความยากพอเหมาะ และข้อสอบทั้งฉบับควรมี
ระดับความยากเฉลี่ยประมาณ 0.5

$$p = \frac{(\text{จำนวนคนที่ตอบถูกในกลุ่มสูง} + \text{จำนวนคนที่ตอบถูกในกลุ่มต่ำ})}{(\text{จำนวนคนทั้งหมดในกลุ่มสูง} + \text{จำนวนคนทั้งหมดในกลุ่มต่ำ})}$$

$$\text{เกณฑ์} \quad 0.20 \leq p \leq 0.80$$

ความยาก (p)	ความหมาย
0.81 – 1.00	ง่ายมาก
0.60 – 0.80	ค่อนข้างง่าย
0.40 – 0.59	ปานกลาง
0.20 – 0.39	ค่อนข้างยาก
0 – 0.19	ยากมาก

2.2 อำนาจจำแนก

อำนาจจำแนก (r) (Discrimination power of the items) หมายถึง ความสามารถของข้อสอบในการจำแนก หรือแยกให้เห็นความแตกต่างระหว่างผู้สอบที่มีผลสัมฤทธิ์ต่างกัน เช่น จำแนกคนเก่งกับคนอ่อนออกจากกันได้ หรือจำแนกคนที่มีความสามารถพิเศษกับคนที่ไม่มีความสามารถออกจากกันได้ โดยถือว่าคนที่เก่งหรือมีความสามารถควรทำข้อสอบนั้นได้ ส่วนผู้ที่อ่อนหรือไม่มีความสามารถไม่ควรทำข้อสอบข้อนั้นได้

$$r = \frac{(\text{จำนวนคนที่ตอบถูกในกลุ่มสูง} - \text{จำนวนคนที่ตอบถูกในกลุ่มต่ำ})}{(\text{จำนวนคนทั้งหมดในกลุ่มสูงหรือจำนวนคนทั้งหมดในกลุ่มต่ำ})}$$

$$\text{เกณฑ์} \quad +0.20 \leq r$$

อำนาจจำแนก (r)	ความหมาย
0.60 – 1.00	ดีมาก
0.40 – 0.59	ดี
0.20 – 0.39	พอใช้ได้
0.10 – 0.19	ค่อนข้างต่ำ ควรปรับปรุง
0 – 0.09	ต่ำมาก ควรปรับปรุง

2.3 ความเที่ยง

ความเที่ยง หมายถึง ความคงที่หรือความคงเส้นคงวาของผลที่ได้จากการวัดซ้ำ ถ้าการวัดสิ่งเดียวกันหลาย ๆ ครั้งได้ค่าที่ค่อนข้างคงเส้นคงวามากขึ้นเพียงใดก็ถือว่าการวัดมีความเที่ยงมากขึ้นเพียงนั้น แบบทดสอบที่มีความเที่ยงสูงจะเป็นเครื่องมือที่ใช้วัดคุณลักษณะที่ต้องการได้ผลอย่างคงเส้นคงวา

การประมาณค่าความเที่ยงสามารถกระทำได้หลายวิธี แต่ละวิธีมีความคล้ายคลึงกันในการหาความสัมพันธ์ระหว่างข้อมูล 2 ชุด ซึ่งได้มาจากเครื่องมือเดียวกัน หรือเครื่องมือที่คู่ขนานกัน เนื่องจากการใช้สัมประสิทธิ์สหสัมพันธ์ (Correlation coefficient) สำหรับประมาณค่าความเที่ยง บางครั้งจึงเรียกผลที่ได้ว่า สัมประสิทธิ์ความเที่ยง (Reliability coefficient)

ความเที่ยงสามารถจำแนกออกเป็น 4 ประเภท ตามตารางที่ 2.1 ดังนี้

ตารางที่ 2.1 ประเภทของความเที่ยง

ประเภท	ความหมาย	วิธีประมาณค่า
1. ความเที่ยงแบบคงที่	ความคงเส้นคงวาของคะแนนจากการวัดในช่วงเวลาที่ต่างกัน โดยวิธีสอบซ้ำด้วยแบบทดสอบเดิม	คำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่วัดได้จากคนกลุ่มเดียวกัน โดยทำการวัดซ้ำสองครั้งในเวลาที่แตกต่างกัน
2. ความเที่ยงแบบสมมูล	ความสอดคล้องกันของคะแนนจากการวัดในช่วงเวลาเดียวกัน โดยใช้แบบทดสอบที่สมมูลกัน	คำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่วัดได้ในเวลาเดียวกันจากคนกลุ่มเดียวกัน โดยใช้เครื่องมือ 2 ฉบับที่ทดสอบกัน
3. ความเที่ยงแบบความคงที่และสมมูล	ความสอดคล้องกันของคะแนนจากการวัดในช่วงเวลาที่ต่างกัน โดยวิธีสอบซ้ำด้วยแบบทดสอบที่สมมูลกัน	คำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่วัดได้ในช่วงเวลาที่ต่างกันจากคนกลุ่มเดียวกัน โดยใช้เครื่องมือ 2 ฉบับ ที่ทดสอบกัน
4. ความเที่ยงแบบความสอดคล้องภายใน	ความสอดคล้องกันระหว่างคะแนนรายข้อหรือความเป็นเอกพันธ์ของเนื้อหา รายข้ออันเป็นตัวแทนของคุณลักษณะเด่นเดียวกันที่ต้องการวัด โดยใช้วิธีต่าง ๆ ดังนี้	คำนวณค่าสัมประสิทธิ์ความเป็นเอกพันธ์ระหว่างคะแนนของกลุ่มข้อสอบ 2 กลุ่ม จากการวัดด้วยแบบทดสอบเดียวกัน
	4.1 วิธีแบ่งครึ่งข้อสอบ	คำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่วัดได้จากการแบ่งครึ่งข้อสอบที่สมมูลกัน เช่น แบ่งเป็นข้อคู่

ตารางที่ 2.1 (ต่อ)

ประเภท	ความหมาย	วิธีประมาณค่า
		และข้อคือ เป็นต้น จากนั้นจึงใช้สูตรของสเปียร์แมน บราวน์
	4.2 วิธีของคูเดอร์ – ริชาร์ดสัน	คำนวณค่าสถิติของคะแนนรายข้อ (ซึ่งให้คะแนนแบบ 0,1) และคะแนนรวม จากนั้นจึงใช้สูตรของคูเดอร์ – ริชาร์ดสัน
	4.3 วิธีสัมประสิทธิ์แอลฟาของครอนบาค	คำนวณค่าสถิติของคะแนนรายข้อและคะแนนรวม จากนั้นจึงใช้สูตรคำนวณสัมประสิทธิ์แอลฟาของครอนบาค
	4.4 วิธีวิเคราะห์ความแปรปรวนของฮอยท์	วิเคราะห์ความแปรปรวนแบบสองทาง จากนั้นจึงใช้สูตรของฮอยท์

ผู้วิจัยใช้คุณภาพด้านความเที่ยง ความยาก และอำนาจจำแนกในการหาคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก โดยผู้วิจัยใช้เกณฑ์ค่าแบบทดสอบที่มีความเที่ยงในการวัดจะต้องสูงกว่าเกณฑ์ขั้นต่ำที่กำหนด คือ 0.50 (ศิริชัย กาญจนวาสี, 2556) และเกณฑ์อำนาจจำแนกเป็นตัวช่วยร่วมกับการพิจารณาไค์คุณลักษณะตัวเลือก OCCs สำหรับใช้ในการตัดตัวเลือกที่ไม่มีประสิทธิภาพ ก่อนที่จะทำการแทนค่าตัวเลือกที่ถูกตัดออก และใช้คุณสมบัติของความเที่ยงในการบอกถึงคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ใน 3 กรณีที่ผู้วิจัยได้ทำการแทนค่าตัวเลือกที่ถูกตัดออก

ตอนที่ 3 แนวคิดของทฤษฎีการตอบสนองข้อสอบด้วยการทำให้เรียบเคอเนล (Kernel Smoothing in Item Response Theory)

Mcguire (2012) ได้กล่าวถึงแนวคิดของทฤษฎีการตอบสนองข้อสอบด้วยการทำให้เรียบเคอเนลว่า การวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรแฝงและตัวแปรสังเกตทวิภาคและพหุภาคในการทดสอบทางการศึกษาและจิตวิทยาจะใช้ทฤษฎีการตอบสนองข้อสอบ ตัวแปรสังเกตได้มาจากแบบทดสอบหรือแบบสอบถามหลาย ๆ ข้อ จาก 2 ประเภทดังนี้ คือ ข้อสอบแบบเลือกตอบหลายตัวเลือก จะมีหนึ่งตัวเลือกที่ออกแบบให้เป็นตัวเลือกที่ถูกต้อง และข้อคำถามที่เป็นแบบมาตราประมาณค่าซึ่งในแต่ละตัวเลือกจะมีน้ำหนักที่แตกต่างกัน (การให้น้ำหนักแบบพหุภาค) ข้อสอบแบบเลือกตอบหลายตัวเลือกอาจมองได้ว่า

ตัวเลือกตัวใดตัวหนึ่งมีน้ำหนักเท่ากับ 1 และตัวเลือกตัวอื่น ๆ ที่เหลือ มีน้ำหนักเท่ากับ 0 (การให้น้ำหนักแบบทวิภาค) โดยธรรมชาติกลุ่มของข้อสอบหรือข้อคำถามจะถูกรวมอยู่ใน 2 ประเภทนี้

ปัญหาของทฤษฎีการตอบสนองข้อสอบทวิภาค (Polytomous IRT) คือ ข้อจำกัดของตัวแบบทางคณิตศาสตร์ในการพรรณาน่าจะเป็นของการเลือกตัวเลือกต่าง ๆ ที่อธิบายอยู่ในรูปของตัวแบบของข้อคำถามหรือข้อสอบที่ใช้วัดตัวแปรแฝงต่อเนื่อง ฟังก์ชันหรือโค้งที่กล่าวถึง คือ โค้งคุณลักษณะตัวเลือก (Option Characteristic Curve (OCC)) ยกตัวอย่างเช่น การวิเคราะห์ข้อสอบแบบเลือกตอบหลายตัวเลือกมักพึ่งพาตัวเลขและสถิติ เช่น ค่า p (หรือสัดส่วนของผู้สอบที่เลือกในแต่ละตัวเลือก) และสหสัมพันธ์แบบ Point Biserial (คุณสมบัติด้านอำนาจจำแนกรายข้อ) ซึ่งในทางปฏิบัติสารสนเทศจากค่า p และสหสัมพันธ์แบบ Point Biserial อาจให้สารสนเทศได้ไม่เพียงพอ อาจต้องการสารสนเทศมากกว่านั้น คือ สารสนเทศจาก OCCs รวมด้วย มากไปกว่านั้น OCCs ยังถือว่าเป็นจุดเริ่มต้นของการวิเคราะห์ทฤษฎีการตอบสนองข้อสอบ

เป้าหมายของการประมาณ OCCs เปรียบเสมือนกับการสร้างตัวแบบทางสถิติแบบดั้งเดิม มี 2 วิธีที่สามารถทำได้ วิธีที่ 1 เป็นวิธีพบเห็นกันมากที่สุด คือ วิธีทฤษฎีการตอบสนองข้อสอบแบบใช้พารามิเตอร์ (Parametric IRT) โครงสร้างพารามิเตอร์ถูกตั้งข้อตกลงว่าการประมาณ OCC ถูกลดลงเพื่อที่จะประมาณเวกเตอร์พารามิเตอร์ ในแง่ของการเปลี่ยนแปลงของมิติจากตัวแบบไปสู่ตัวแบบ ในแต่ละข้อ เวกเตอร์จะถูกพิจารณาโดยตรง และการประมาณเวกเตอร์จะใช้เพื่อสรุปสถิติเพื่อพรรณนา เช่น ความยาก และอำนาจจำแนกของแต่ละข้อ วิธีที่ 2 คือ วิธีทฤษฎีการตอบสนองข้อสอบแบบไม่ใช้พารามิเตอร์ (Nonparametric IRT) เป็นวิธีการที่ใช้ประมาณโดยตรง โดยปราศจากการตั้งข้อตกลงเกี่ยวกับรูปแบบทางคณิตศาสตร์ของ OCCs เป็นการประมาณที่มีความยืดหยุ่น อ้างอิงได้ว่าการประมาณด้วยวิธีการนี้ประมาณได้ใกล้เคียงกับค่าจริงของ OCCs มากกว่าใช้วิธีการประมาณด้วยวิธีทฤษฎีการตอบสนองข้อสอบแบบใช้พารามิเตอร์ แต่ก็มีข้อโต้แย้งว่าวิธีทฤษฎีการตอบสนองข้อสอบแบบไม่ใช้พารามิเตอร์อาจเป็นวิธีการที่มีการนำไปอ้างอิงน้อยกว่าวิธีทฤษฎีการตอบสนองข้อสอบแบบใช้พารามิเตอร์ โดยทั่วไปข้อโต้แย้งหลักของวิธีทฤษฎีการตอบสนองข้อสอบแบบไม่ใช้พารามิเตอร์คือมีความยืดหยุ่นและสะดวกในการคำนวณ มากไปกว่านั้น แม้ว่าตัวแบบไม่ใช้พารามิเตอร์อาจไม่เป็นไปตามคุณลักษณะของการใช้พารามิเตอร์ แต่มีหลายคนที่สนับสนุนการนำกราฟของวิธีการไม่ใช้พารามิเตอร์มาใช้ประโยชน์

สำหรับการวิจัยนี้ ผู้วิจัยได้นำวิธีทฤษฎีการตอบสนองข้อสอบแบบไม่ใช้พารามิเตอร์ (Nonparametric IRT) มาวิเคราะห์ข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก เหตุผลที่ผู้วิจัยใช้วิธีการดังกล่าวเนื่องจากผู้วิจัยไม่ได้เป็นผู้ออกแบบและสร้าง และนำแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกฉบับนี้ไปใช้ตั้งแต่เริ่มต้น ผู้วิจัยไม่สามารถควบคุมและจัดการข้อมูลที่ได้จากแบบทดสอบฉบับนี้ได้ ดังนั้นผู้วิจัยจึงไม่สามารถแน่ใจได้ว่าข้อมูลที่ได้จากแบบทดสอบฉบับนี้จะเป็นไปตามข้อตกลงเบื้องต้นของวิธีทฤษฎีการตอบสนองข้อสอบแบบใช้พารามิเตอร์ (Parametric IRT) หรือไม่ ดังนั้นผู้วิจัยจึงใช้วิธีทฤษฎีการตอบสนองข้อสอบแบบไม่ใช้พารามิเตอร์ (Nonparametric IRT) มาใช้ในการวิเคราะห์ข้อมูล

ผู้วิจัยใช้ผลการวิเคราะห์ข้อมูลจากวิธีทฤษฎีการตอบสนองข้อสอบแบบไม่ใช้พารามิเตอร์ (Nonparametric IRT) ที่ให้ผลลัพธ์ออกมาอยู่ในรูปแบบของโค้งคุณลักษณะตัวเลือก หรือ OCCs สำหรับใช้เป็นข้อมูลอ้างอิงในการตัดตัวเลือกที่ไม่มีประสิทธิภาพออก ก่อนการแทนที่ตัวเลือกที่ถูกตัดออกเหล่านั้น และยังใช้ผลการวิเคราะห์องค์ประกอบหลัก หรือ PCA ในการช่วยวิเคราะห์หาคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และ 4 ตัวเลือก

ตอนที่ 4 การปรับเทียบคะแนน (Equating)

4.1 ความหมายของการปรับเทียบคะแนน

ความหมายและนิยามของการปรับเทียบคะแนน ที่นักการศึกษาหลายท่านได้กล่าวไว้ เช่น กัลลิกเสน (ภาวิณี ศรีสุขพัฒนานันท์, 2529 : 17 อ้างอิงจาก Gulliksen, 1950 : 298) ได้ให้ ความหมายของการปรับเทียบคะแนนระหว่างแบบทดสอบว่าเป็นวิธีการทำคะแนนจากแบบทดสอบสองชุดที่วัดในวิชาเดียวกันให้เป็นคะแนนสมมูล (Equivalent scores) ที่เปรียบเทียบกันโดยตรง โดยเสนอวิธีการให้ผู้สอบกลุ่มเดียวทำแบบทดสอบสองชุด และใช้วิธีง่าย ๆ คือ แปลงคะแนนแต่ละชุดให้เป็นคะแนนมาตรฐานแล้วนำคะแนนที่แปลงแล้วมาเทียบกันโดยตรง แต่ต้องตรวจสอบให้แน่ใจก่อนว่า คะแนนสองชุดนั้นมีการกระจายของคะแนนเป็นโค้งปกติ ซึ่งสอดคล้องกับ ปีเตอร์เซนและคณะ (Petersen & other, 1982 : 242) ที่ให้นิยามของการปรับเทียบคะแนนระหว่างแบบทดสอบว่าเป็นกระบวนการเชิงประจักษ์ที่ใช้ในการแปลงคะแนนจากแบบทดสอบฉบับหนึ่งไปยังแบบทดสอบอีกฉบับหนึ่ง โดยที่แบบทดสอบทั้งสองนั้นต้องวัดคุณลักษณะเดียวกัน และสอดคล้องกับศิริชัย กาญจนวาสี (2541 : 2) ซึ่งได้ให้นิยามว่า การปรับเทียบคะแนนระหว่างแบบทดสอบเป็นกระบวนการแปลงคะแนนของแบบทดสอบต่างฉบับที่มุ่งวัดคุณลักษณะเดียวกันให้เป็นคะแนนที่สมมูลกันเพื่อทำให้คะแนนสามารถเปรียบเทียบกันได้โดยตรง ส่วนลอร์ด (Lord, 1980 : 195) ได้กล่าวถึงการปรับเทียบคะแนน ระหว่างแบบทดสอบว่าเป็นการแปลงคะแนนจากแบบทดสอบต่างฉบับให้มีความหมายสลับเปลี่ยนกันได้ และเพื่อความเสมอภาคของบุคคลผู้รับการทดสอบ ในขณะที่แองกอฟ (Angoff, 1984 : 85) ให้ความหมายของการปรับเทียบคะแนนระหว่างแบบทดสอบว่า หมายถึง กระบวนการแปลงระบบคะแนนของแบบทดสอบฉบับหนึ่งไปสู่ระบบคะแนนของแบบทดสอบอีกฉบับหนึ่งซึ่งวัดสิ่งเดียวกัน เพื่อให้คะแนนแปลงจากแบบทดสอบทั้งสองฉบับนั้น สามารถเทียบเท่ากันและใช้แทนกันได้ จากแนวคิดและการให้ความหมายดังกล่าวข้างต้นพอจะสรุปได้ว่า การปรับเทียบคะแนนเป็นวิธีการทางสถิติในการนำคะแนนจากแบบทดสอบต่างฉบับกันแต่วัดในสิ่งเดียวกันไปแปลงให้มีหน่วยวัดเดียวกัน เพื่อให้คะแนนจากแบบสอบต่างฉบับกันสามารถเปรียบเทียบกันได้โดยตรง

4.2 เงื่อนไขของการปรับเทียบคะแนน

เนื่องจากวัตถุประสงค์ของการปรับเทียบคะแนนระหว่างแบบทดสอบ คือการทำให้คะแนนที่ได้จากแบบทดสอบ 2 ฉบับใด ๆ มีความเท่าเทียมกันมากที่สุด ซึ่งต้องเกี่ยวข้องกับแบบแผนในการเก็บรวบรวมข้อมูล และกฎเกณฑ์การแปลงคะแนนจากแบบทดสอบฉบับหนึ่ง ไปยังแบบทดสอบอีกฉบับหนึ่ง

จึงต้องกำหนดเงื่อนไขบางประการในการปรับเทียบคะแนน ซึ่งลอร์ด (Petersen and others, 1982 : 242 citing Lord, 1980) ได้กล่าวถึงเงื่อนไข 4 ข้อ ที่จะต้องนำมาใช้ในการปรับเทียบคะแนน ดังนี้

1. แบบทดสอบทั้งสองฉบับจะต้องวัดความสามารถเดียวกัน คือ แบบทดสอบทั้งสองต้องวัดคุณลักษณะเดียวกัน คุณลักษณะนี้อาจเป็นคุณลักษณะแฝง หรือความสามารถ หรือทักษะอย่างใดอย่างหนึ่งก็ได้

2. ความเสมอภาค (Equity) คือ เมื่อทุกกลุ่มมีความสามารถเดียวกัน การแจกแจงคะแนนของแบบทดสอบที่ได้ หลังจากที่มีการแปลงคะแนนแล้ว จะมีการแจกแจงเหมือนกับการแจกแจงของคะแนนจากแบบทดสอบที่ใช้เป็นฉบับเทียบคะแนน

3. ความไม่แปรผันตามกลุ่ม (Invariance Across Groups) คือ คะแนนที่ได้จากการแปลงคะแนนไม่ว่าจะมาจากกลุ่มใด ๆ ก็ตามจะมีความเท่าเทียมกันหรือมีความไม่เปลี่ยนแปลงไปตามกลุ่มผู้เข้าสอบ

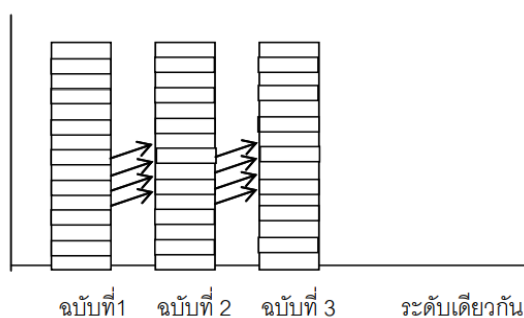
4. มีความสมมาตร (Symmetry) คือ ผลของการปรับเทียบคะแนนต้องเหมือนกันไม่ว่าจะเทียบจากแบบทดสอบฉบับ X ไปยังฉบับ Y หรือเทียบจากแบบทดสอบฉบับ Y ไปยังฉบับ X

4.3 ประเภทของการปรับเทียบคะแนน

การทดสอบที่มีความจำเป็นต้องใช้เทคนิคการปรับเทียบคะแนนระหว่างแบบทดสอบ สามารถจำแนกได้เป็น 2 สถานการณ์ (ศิริชัย กาญจนวาสี, 2541 : 5 อ้างอิงมาจาก Hambleton and Swaminathan, 1985) ซึ่งต้องใช้รูปแบบการปรับเทียบคะแนน (Equating Forms) 2 รูปแบบ ดังนี้

4.3.1 การปรับเทียบคะแนนตามแนวนอน (Horizontal Equating) การปรับเทียบคะแนนตามแนวนอนเป็นการปรับเทียบคะแนนระหว่างแบบทดสอบต่างฉบับกัน เมื่อแต่ละฉบับมุ่งวัดคุณลักษณะเดียวกันมีระดับความยากใกล้เคียงกัน และกลุ่มผู้สอบมีการแจกแจงความสามารถอยู่ในประชากรเดียวกัน และมีความสามารถใกล้เคียงกัน การปรับเทียบคะแนนตามแนวนอนเป็นเทคนิคที่เหมาะสมในสถานการณ์ที่มีความจำเป็นต้อง สร้างแบบทดสอบเนื้อหาเดียวกันขึ้นมาหลายๆ ฉบับ ด้วยเหตุผลของการนำไปใช้ในการทดสอบให้ เกิดความยุติธรรมและป้องกันความลับของข้อสอบเมื่อใช้ในเวลาที่แตกต่างกัน สำหรับผู้สอบกลุ่มขนาดใหญ่เพื่อเทียบว่าคะแนนที่ได้จากฉบับหนึ่งเทียบเป็นเท่าไรของอีกฉบับหนึ่ง ซึ่งวัดในระดับเดียวกันจึงเป็นการปรับเทียบคะแนนระหว่างแบบทดสอบต่างฉบับของวิชาเดียวกัน สำหรับกลุ่มผู้สอบระดับชั้นเดียวกัน เช่น การปรับเทียบคะแนนระหว่างแบบทดสอบวิชาคณิตศาสตร์ 2 ฉบับ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 เป็นต้น

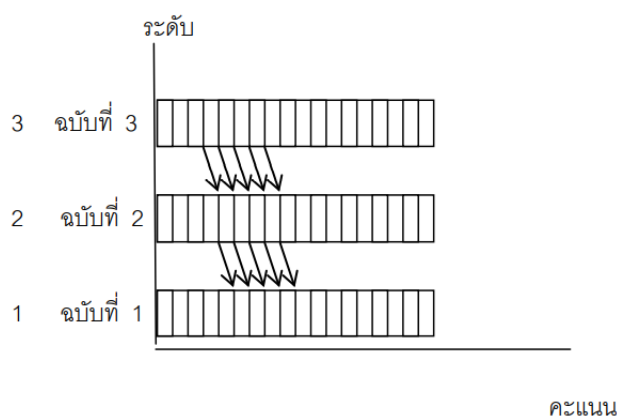
คะแนน



ภาพที่ 2.1 การปรับเทียบคะแนนตามแนวนอน (Horizontal Equating)

แบบทดสอบต่างฉบับที่ต้องการนำมาเทียบมาตรากันตามแนวนอนนั้น ถึงแม้ว่าแบบทดสอบเหล่านั้นยากที่จะมีความเท่าเทียมกันในทุกด้าน แต่ก็ควรเป็นแบบทดสอบที่พยายามสร้างหรือมุ่งให้คู่ขนานกัน (Alternate Forms) นอกจากนี้การแจกแจงความสามารถของผู้สอบที่ทำการสอบด้วยแบบทดสอบแต่ละฉบับควรมีการแจกแจงที่ประมาณได้ว่ามีระดับความสามารถที่ใกล้เคียงกัน

4.3.2 การปรับเทียบคะแนนตามแนวตั้ง (Vertical Equating) การปรับเทียบคะแนนแนวตั้งเป็นเทคนิคที่เหมาะสมในสถานการณ์ที่มีความจำเป็นต้องสร้างแบบทดสอบเนื้อหาเดียวกัน แต่ต่างฉบับต่างมุ่งวัดความสามารถของผู้สอบที่ต่างระดับกัน เพื่อเทียบว่าคะแนนที่สอบได้จากฉบับหนึ่งเทียบเป็นเท่าไรของฉบับอื่นที่วัดต่างระดับกัน จึงเป็นการปรับเทียบคะแนนระหว่างแบบทดสอบต่างระดับของวิชาเดียวกันสำหรับกลุ่ม ผู้สอบต่างระดับชั้นกัน เช่น การปรับเทียบคะแนนระหว่างแบบทดสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 สำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น เป็นต้น



ภาพที่ 2.2 การปรับเทียบคะแนนตามแนวตั้ง (Vertical Equating)

แบบทดสอบต่างฉบับที่วัดเนื้อหาเรื่องเดียวกัน แต่ต่างระดับกันที่ต้องการนำมาเทียบคะแนนกันนั้น แบบทดสอบแต่ละฉบับจึงมีความยากแตกต่างกันไป รวมทั้งกลุ่มผู้สอบแต่ละฉบับมีการแจกแจงความสามารถอยู่ต่างประชากรหรือมีการแจกแจงความสามารถที่อยู่ระดับต่างประชากรหรือมีการแจกแจงความสามารถที่อยู่ในระดับที่ต่างกัน ดังนั้นการปรับเทียบคะแนนตามแนวตั้งจึงมีความสลับซับซ้อนกว่าการปรับเทียบคะแนนตามแนวนอนทั้งในแง่ทฤษฎีและการปฏิบัติ

รูปแบบของการปรับเทียบคะแนนสามารถแบ่งได้ 2 รูปแบบ คือ การเทียบมาตรากันตามแนวนอน เป็นการปรับเทียบคะแนนระหว่างแบบทดสอบที่มีความยากใกล้เคียงกัน และกลุ่มผู้สอบมีความสามารถใกล้เคียงกันด้วย การปรับเทียบคะแนนตามแนวตั้งเป็นการปรับเทียบคะแนนของแบบทดสอบที่มีความยากต่างกัน และกลุ่มผู้สอบก็มีความสามารถต่างกันด้วย

4.4 รูปแบบวิธีการเก็บข้อมูล

ข้อมูลจากการใช้แบบทดสอบหลายฉบับและเก็บรวบรวมข้อมูลสำหรับนำมาใช้ในการเทียบมาตรา สามารถใช้จำแนกเป็นแบบแผนการเก็บรวบรวมข้อมูลได้ดังนี้ (ศิริชัย กาญจนวาสี, 2541 : 8 – 11 อ้างอิงมาจาก Petersen and others, 1989 : 244 – 246)

4.4.1 รูปแบบผู้สอบกลุ่มเดียว (Single Group Design)

ก. รูปแบบผู้สอบกลุ่มเดียวที่ไม่ได้รับการจัดให้สมดุล (Unconterbalanced Design) รูปแบบนี้เป็น รูปแบบที่ง่ายที่สุด ใช้ผู้สอบกลุ่มเดียวทำหน้าที่เป็นผู้สอบร่วม วิธีการเก็บรวบรวมข้อมูลกระทำโดยใช้กลุ่มตัวอย่างผู้สอบกลุ่มเดียว แต่ละคนทำแบบทดสอบทั้ง 2 ฉบับ ให้ผู้สอบทำแบบทดสอบฉบับหนึ่งแล้วตามด้วยแบบทดสอบอีกฉบับหนึ่ง การให้ผู้สอบกลุ่มเดียวกันทำแบบทดสอบทั้ง 2 ฉบับ ระดับความสามารถของผู้สอบทั้ง 2 ฉบับ เท่ากัน จึงไม่เป็นปัญหาแทรกซ้อนรวมทั้งไม่เป็นปัญหาในเรื่องความยากของแบบทดสอบต่างฉบับ คะแนนจากแบบทดสอบจึงสามารถเปรียบเทียบกันได้บนพื้นฐานความสามารถที่เท่ากันจากคนกลุ่มเดียวกัน แต่ในทางปฏิบัติ การทดสอบฉบับหลังผู้สอบอาจได้รับปัจจัยที่ส่งผลกระทบตอกทอดจากการทดสอบฉบับแรก เช่น การเรียนรู้ การฝึกฝน ความเมื่อยล้าจากการทำแบบทดสอบฉบับแรก เป็นต้น ซึ่งอาจมีอิทธิพลต่อการเปรียบเทียบคะแนน

ข. รูปแบบกลุ่มเดียวที่ได้รับการจัดให้สมดุล (Counterbalanced Design) รูปแบบนี้ปรับจากแบบแรก เพื่อขจัดผลของลำดับการทดสอบ ก่อน – หลัง โดยการสุ่มผู้สอบแยกเป็น 2 กลุ่มย่อย แต่ละกลุ่มได้รับการทดสอบด้วยแบบทดสอบทั้ง 2 ฉบับ ในลักษณะที่กลุ่มย่อยหนึ่งให้ทำแบบทดสอบฉบับ 1 แล้วตามด้วยแบบทดสอบฉบับ 2 ส่วนอีกกลุ่มย่อยหนึ่งให้ทำแบบทดสอบฉบับ 2 แล้วตามด้วยฉบับ 1 สลับกัน เพื่อให้เกิดความสมดุล การใช้รูปแบบนี้กลุ่มผู้สอบจึงได้รับอิทธิพลจากลำดับการทดสอบ การเรียนรู้ การฝึกฝนและความ เมื่อยล้าที่สมดุลกันในทั้งสองกลุ่มย่อย

การเก็บรวบรวมข้อมูล	กลุ่มตัวอย่าง	แบบทดสอบ		
		ฉบับที่ 1	ฉบับที่ 2	แบบทดสอบรวม
รูปแบบผู้สอบกลุ่มเดียว				
ก. รูปแบบผู้สอบกลุ่มเดียวที่ไม่ได้รับการจัดให้สมดุล	P_1	X	Y	
ข. รูปแบบผู้สอบกลุ่มเดียวที่ได้รับการจัดให้สมดุล	P_1	X	Y	
	P_2	Y	X	

ภาพที่ 2.3 รูปแบบการเก็บข้อมูลกับผู้สอบกลุ่มเดียว

4.4.2 รูปแบบผู้สอบกลุ่มเท่าเทียมกัน (Equivalent – Group Design) รูปแบบนี้เป็นการใช้กลุ่มเท่าเทียมกัน โดยการจัดกลุ่มผู้สอบให้มีความคล้ายคลึงกันมากที่สุดเท่าที่จะทำได้ แล้วให้ผู้สอบในแต่ละ

กลุ่มทำแบบทดสอบฉบับเดียวกัน ถึงแม้การจัดกลุ่มจะไม่เหมือนกันทุกประการ (Identical Groups) แต่ก็ควรมีลักษณะทั่วไป คล้ายกันมากที่สุดอาจจะได้มาโดยวิธีการสุ่ม (Random) ข้อดีของรูปแบบนี้คือ แต่ละกลุ่มทำ แบบทดสอบฉบับเดียว จึงสามารถหลีกเลี่ยงปัญหา การเรียนรู้ การฝึกฝนและความเมื่อยล้าจากการทำแบบทดสอบ แต่ปัญหาที่ตามมาคือ กลุ่มที่ใช้ไม่เหมือนกันอาจมีการแจกแจงความสามารถที่แตกต่างกัน และไม่มีข้อมูลที่นำมาใช้ปรับความแตกต่างของกลุ่ม ความแตกต่างที่เกิดขึ้นแม้เล็กน้อยย่อมส่งผลกระทบต่อความลำเอียงในการเทียบมาตราได้ วิธีการที่จะช่วยลดความแตกต่างระหว่างกลุ่มได้โดยใช้กลุ่มตัวอย่างสุ่มขนาดใหญ่

การเก็บรวบรวมข้อมูล	กลุ่มตัวอย่าง	แบบทดสอบ		
		ฉบับที่ 1	ฉบับที่ 2	แบบทดสอบรวม
รูปแบบผู้สอบกลุ่มเท่าเทียมกัน	P_1	X		
	P_2		Y	

ภาพที่ 2.4 รูปแบบการเก็บข้อมูลกับผู้สอบกลุ่มเท่าเทียมกัน

4.4.3 รูปแบบผู้สอบต่างกลุ่มโดยใช้แบบทดสอบรวม (Anchor – Test Design)

ก. รูปแบบผู้สอบกลุ่มสุ่มโดยใช้แบบทดสอบรวม (Anchor – Test Random Group Design) รูปแบบนี้เป็นการใช้กลุ่มผู้สอบซึ่งได้จากการสุ่มแล้วให้ผู้สอบแต่ละกลุ่มสอบแบบทดสอบฉบับเดียวกัน โดยแบบทดสอบแต่ละฉบับมีข้อสอบร่วมกันจำนวนหนึ่ง ซึ่งเรียกว่า แบบทดสอบรวม (Anchor Test) การใช้แบบทดสอบรวมนี้สามารถกระทำได้ 2 ลักษณะดังนี้

1. แบบทดสอบรวมภายใน (Internal Anchor Test) เป็นแบบทดสอบที่มีข้อสอบรวมที่จัดไว้อยู่ในแบบสอบที่ต้องการนำมาปรับเทียบคะแนน
2. แบบทดสอบรวมภายนอก (External Anchor Test) เป็นแบบทดสอบที่มีข้อสอบรวมที่ถูกจัดแยกออกมาต่างหากจากแบบทดสอบที่ต้องการนำมาปรับเทียบคะแนน ซึ่งกลุ่มผู้สอบจะต้องทำและจับเวลาแยกออกจากตัวแบบทดสอบที่ต้องการนำมาเทียบคะแนน

รูปแบบนี้จะช่วยลดความลำเอียงในการเทียบคะแนนมีสาเหตุมาจากการเรียนรู้ ฝึกฝน ความเมื่อยล้าและความแตกต่างของระดับความสามารถของกลุ่มผู้สอบ รูปแบบนี้จึงได้เปรียบเหนือกว่า 2 รูปแบบแรก แต่อย่างไรก็ตามประสิทธิภาพของการเทียบคะแนนขึ้นอยู่กับปัจจัยที่สำคัญบางประการ เช่น คุณภาพของแบบทดสอบรวม ซึ่งจะต้องคำนึงถึงความสัมพันธ์ระหว่างข้อสอบรวมที่ต้องมีเนื้อหาและระดับความยากคล้ายกับแบบทดสอบที่ต้องการเทียบคะแนนและความยากของแบบทดสอบรวมเป็นสำคัญ

ข. รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกันโดยใช้แบบทดสอบรวม (Anchor – Test Nonequivalent Groups Design) รูปแบบนี้เป็นการใช้กลุ่มผู้สอบซึ่งมีลักษณะไม่เท่าเทียมกัน แล้วให้ผู้สอบแต่ละกลุ่มทำแบบทดสอบฉบับเดียวกันกลุ่มละฉบับ โดยแบบทดสอบแต่ละฉบับมีข้อสอบร่วมกันจำนวนหนึ่ง ซึ่งเรียกว่า ข้อสอบรวม (Anchor Test) ซึ่งอาจใช้แบบทดสอบรวมภายในหรือแบบทดสอบรวมภายนอกก็ได้ในการ

ทดสอบบางสถานการณ์อาจมีความจำเป็นต้องใช้แบบทดสอบร่วมกับกลุ่มผู้สอบต่างประชากรกัน เช่น กลุ่มผู้สอบต่างโปรแกรม เวลา หรือระดับ เป็นต้น ซึ่งคะแนนจากแบบทดสอบร่วมจะเป็นกลไกสำคัญสำหรับการเทียบคะแนนระหว่างแบบทดสอบ

การเก็บรวบรวมข้อมูล	กลุ่มตัวอย่าง	แบบทดสอบ		
		ฉบับที่ 1	ฉบับที่ 2	แบบทดสอบร่วม
รูปแบบผู้สอบต่างกลุ่มโดยใช้แบบทดสอบร่วม ก.รูปแบบผู้สอบกลุ่มสุมโดยใช้แบบทดสอบร่วม (1) แบบทดสอบร่วมภายใน	P ₁	XV		
	P ₂		YV	
(2) แบบทดสอบร่วมภายนอก	P ₁	X		V
ข. รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกันโดยใช้แบบทดสอบร่วม (1) แบบทดสอบร่วมภายใน	P ₂		Y	V
	P ₁	XV		
(2) แบบทดสอบร่วมภายนอก	Q ₂		YV	V
	P ₁	X		
	Q ₂		Y	V

ภาพที่ 2.5 รูปแบบการเก็บข้อมูลกับผู้สอบต่างกลุ่มโดยใช้แบบทดสอบร่วม

4.4.4 รูปแบบผู้สอบต่างกลุ่มโดยใช้ผู้สอบร่วม (Common – Person Design) รูปแบบนี้เป็นการใช้กลุ่มผู้สอบต่างกลุ่มกันทำแบบทดสอบต่างฉบับแต่มีจำนวนผู้สอบส่วนหนึ่ง จากทั้งสองกลุ่มทำแบบทดสอบทั้ง 2 ฉบับ ซึ่งเป็นรูปแบบของการใช้ผู้สอบร่วม

การเก็บรวบรวมข้อมูล	กลุ่มตัวอย่าง	แบบทดสอบ		
		ฉบับที่ 1	ฉบับที่ 2	แบบทดสอบร่วม
รูปแบบผู้สอบต่างกลุ่มโดยมีผู้สอบร่วม	P _{1c}	X		
	P	X	Y	
	Q ₁		Y	
	Q _{1c}	X	Y	

ภาพที่ 2.6 รูปแบบการเก็บข้อมูลกับผู้สอบต่างกลุ่มโดยมีผู้สอบร่วม

ในการเก็บรวบรวมข้อมูลที่ใช้ในแบบทดสอบหลายฉบับและการเก็บข้อมูลเพื่อการเปรียบเทียบคะแนนแบบทดสอบ สามารถจำแนกการเก็บรวบรวมข้อมูลได้ 4 รูปแบบ คือ รูปแบบผู้สอบกลุ่มเดียว รูปแบบผู้สอบกลุ่มเท่าเทียมกัน รูปแบบผู้สอบต่างกลุ่มโดยใช้แบบทดสอบร่วม และรูปแบบผู้สอบต่างกลุ่มโดยมีผู้สอบร่วม

4.5 รูปแบบวิธีการปรับเทียบคะแนน (Designs for Equating Method)

วิธีการที่ใช้ในการปรับเทียบคะแนนมีหลายวิธี แต่สามารถจำแนกตามแนวคิดพื้นฐานของการปรับเทียบได้เป็นรูปแบบหลัก 2 รูปแบบ (ศิริชัย กาญจนวาสี, 2541 : 11 – 42) ได้แก่

4.5.1 รูปแบบการปรับเทียบคะแนนโดยใช้ทฤษฎีดั้งเดิม (Classical Models of Equating)

4.5.2 รูปแบบการปรับเทียบโดยใช้ทฤษฎีการตอบสนองข้อสอบ (Equating through Item Response Theory)

ในที่นี้ผู้วิจัยขอกล่าวถึงเฉพาะรูปแบบและวิธีการที่ผู้วิจัยนำมาใช้ในการวิจัยเรื่องนี้เท่านั้น คือ วิธีการปรับเทียบคะแนนเชิงเส้น (Linear Equating) ซึ่งเป็นวิธีการหนึ่งที่อยู่ในรูปแบบการปรับเทียบคะแนนโดยใช้ทฤษฎีดั้งเดิม ดังรายละเอียดต่อไปนี้

วิธีการปรับเทียบคะแนนเชิงเส้น (Linear Equating)

การเทียบคะแนนเชิงเส้นมีนิยามว่า สำหรับกลุ่มผู้สอบที่กำหนดคะแนนจากแบบทดสอบสองฉบับ จะถือว่าเท่าเทียมกันถ้าต่างก็ตรงกับคะแนนมาตรฐานเดียวกัน (Angoff, 1971; Petersen and Other, 1982 อ้างถึงใน สุนิสา จัยม่วงศรี, 2537 : 13) การเทียบคะแนนเชิงเส้นเป็นวิธีการเทียบคะแนนที่ง่ายและสะดวกที่สุดในการนำไปใช้ (Hills, Subhiyah and Hirsch, 1988 : 229)

แองกอฟฟ์ (Angoff, 1984 อ้างถึงใน สุนิสา จัยม่วงศรี, 2537 : 13) ได้เสนอแบบแผนการรวบรวมและการจัดกระทำข้อมูลในทางสถิติของการเทียบคะแนนเชิงเส้นไว้หลายรูปแบบ ซึ่งแต่ละรูปแบบจะมีวิธีการประมาณค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของแบบสอบแตกต่างกันออกไป ตามเงื่อนไขของรูปแบบการรวบรวมข้อมูล แต่ทุกรูปแบบจะตัดสินคะแนนสมมูลจากค่าคะแนนมาตรฐานเดียวกันคือ

$$\frac{Y - M_Y}{S_Y} = \frac{X - M_X}{S_X} \quad \text{-----}(1)$$

จากสมการที่ (1) สามารถเขียนใหม่ได้ดังนี้

$$(Y - M_Y)/S_Y = (X - M_X)/S_X$$

เมื่อ X, Y คือ คะแนนจากแบบทดสอบฉบับ X และฉบับ Y
 M_X, M_Y คือ ค่าเฉลี่ยจากคะแนนแบบทดสอบฉบับ X และฉบับ Y
 S_X, S_Y คือ ส่วนเบี่ยงเบนมาตรฐานของคะแนนจากแบบทดสอบฉบับ X และฉบับ

Y

โดยที่คะแนนจากแบบทดสอบฉบับ X และแบบทดสอบฉบับ Y สามารถนำมาปรับเทียบเชิงเส้นตรง ได้ดังสมการ

$$Y = A_{YX}X + B_{YX}$$

$$\text{เมื่อ } A_{YX} = \frac{S_Y}{S_X} \text{ และ } B_{YX} = M_Y - A_{YX} M_X$$

ดังนั้น $Y = A_{YX}X + B_{YX}$ อาจเขียนในรูป

$$Y = \left(\frac{S_Y}{S_X} \right) X + \left(M_Y - \frac{S_Y}{S_X} M_X \right) \text{ -----(2)}$$

$$\begin{aligned} (Y - M_Y) / \frac{S_X}{S_Y} &= X - M_X \\ X &= M_X + (Y - M_Y) \frac{S_X}{S_Y} \\ &= M_X + Y \frac{S_X}{S_Y} - M_Y \frac{S_X}{S_Y} \\ &= \frac{S_X}{S_Y} Y + \left[M_X - M_Y \frac{S_X}{S_Y} \right] \text{ -----(3)} \end{aligned}$$

สำหรับการวิจัยนี้ ผู้วิจัยจำลองวิธีการเก็บรวบรวมข้อมูลด้วยรูปแบบผู้สอบกลุ่มเดียวที่ไม่ได้รับการจัดให้สมดุล (Single – Group Uncounterbalanced Design) โดยใช้กลุ่มผู้สอบกลุ่มเดียวทำหน้าที่เป็นผู้สอบรวม กล่าวคือ ผู้สอบคนเดียวกันทำแบบสอบทั้ง 2 ฉบับ หรือเรียกอีกอย่างหนึ่งว่า Single Group Design และใช้วิธีการปรับเทียบคะแนนเชิงเส้น (Linear Equating) ซึ่งเป็นวิธีการหนึ่งที่อยู่ในรูปแบบการปรับเทียบคะแนนโดยใช้ทฤษฎีดั้งเดิม (Classical Models of Equating)

ตอนที่ 5 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการศึกษาคุณภาพของแบบทดสอบ

วารุณี ปิตธวัชชัย (2514) ได้ทำการวิจัยเรื่อง การศึกษาเปรียบเทียบค่าสถิติของแบบทดสอบชนิดเลือกตอบที่มีจำนวนตัวเลือกไม่เท่ากัน มีจุดมุ่งหมายเพื่อที่จะศึกษาเปรียบเทียบค่าความเชื่อมั่น ค่าอำนาจจำแนก และระดับความยากง่ายของแบบทดสอบชนิดเลือกตอบชุดเดียวกัน เพื่อลดตัวเลือกลงจาก 5 ตัวเลือก เป็น 4 ตัวเลือก และ 3 ตัวเลือก ว่าจะแตกต่างกันหรือไม่ ในการเก็บรวบรวมข้อมูลใช้แบบทดสอบความถนัดทางการเรียนประเภทอุปมาอุปไมยซึ่งสร้างโดยโครงการวิจัยการเลือกสรร วิทยาลัยวิชาการศึกษาประสานมิตร เป็นแบบทดสอบแบบเลือกตอบชนิด 5 ตัวเลือก จำนวน 48 ข้อ ใช้เวลาในการทำตอบ 14 นาที ในการทดสอบแยกแบบทดสอบออกเป็น 3 ฉบับ คือ ชนิด 5 ตัวเลือก 4 ตัวเลือก และ 3 ตัวเลือก โดยให้แบบทดสอบแต่ละฉบับมีข้อความคำถามเหมือนกัน และมีจำนวนข้อเท่ากัน ในการลดหรือตัดทอนตัวเลือก ใช้วิธีสุ่มตัดตัวลงของข้อสอบแต่ละข้อทีละตัว นำแบบทดสอบทั้ง 3 ฉบับ ไปทดสอบนักเรียนชั้นประถมศึกษาปีที่ 7 โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย โดยทำการทดสอบด้วยกัน 3 ครั้ง แต่ละครั้งเว้นช่วงในการทดสอบ 1 เดือน การทดสอบแต่ละครั้ง แบ่งนักเรียนออกเป็น 3 กลุ่ม กลุ่มละ 60 คน ให้เด็กทั้ง 3 กลุ่ม มีโอกาสตอบแบบทดสอบทั้ง 3 ฉบับ นำข้อมูลที่ได้มาคำนวณหาค่าความเชื่อมั่นของแบบทดสอบโดยใช้สูตรคูเดอร์ ริชาร์ดสัน 21 และหาค่าอำนาจจำแนกของข้อสอบแต่ละข้อ และระดับความยากของแบบทดสอบ โดยวิธีของไฟนดเลย์ และจอห์นสัน แล้วเปรียบเทียบความแตกต่างระหว่างค่าความเชื่อมั่น ค่าอำนาจจำแนก และระดับความยากของแบบทดสอบทั้ง 3 ชนิด

ผลการวิจัยปรากฏว่า 1. การลดจำนวนตัวเลือกของข้อสอบจาก 5 ตัวเลือก เป็น 4 ตัวเลือก และ 3 ตัวเลือก จะทำให้ค่าความเชื่อมั่นของแบบทดสอบแตกต่างกัน โดยค่าความเชื่อมั่นของแบบทดสอบ 5 ตัวเลือก มีค่าสูงสุด คือ 0.8048 แล้วค่อย ๆ ลดต่ำลงตามลำดับ เมื่อจำนวนตัวเลือกของข้อสอบลดลงเป็น 4 ตัวเลือก และ 3 ตัวเลือก (0.7784 และ 0.6582) และค่าความเชื่อมั่นของแบบทดสอบ ชนิดเลือกตอบระหว่าง 5 ตัวเลือก กับ 3 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก กับ 3 ตัวเลือก มีค่าแตกต่างกันอย่างมีนัยสำคัญ ส่วนแบบทดสอบ 5 ตัวเลือก กับ 4 ตัวเลือก มีค่าความเชื่อมั่นของแบบทดสอบไม่แตกต่างกัน 2. การลดจำนวนตัวเลือกของข้อสอบจาก 5 ตัวเลือก เป็น 4 ตัวเลือก และ 3 ตัวเลือกนั้น ไม่ทำให้ค่าอำนาจจำแนกของข้อสอบแต่ละข้อแตกต่างกันอย่างมีนัยสำคัญ และ 3. การลดจำนวนตัวเลือกของข้อสอบจาก 5 ตัวเลือก เป็น 4 ตัวเลือก และ 3 ตัวเลือก ไม่ทำให้ระดับความยากของแบบทดสอบแตกต่างกันอย่างมีนัยสำคัญ

ณัชชา มาลัยและทิพย์วัน เลิศชุมทรัพย์ (2550) ทำวิจัยเพื่อศึกษาคุณภาพรายข้อของข้อสอบแบบเลือกตอบที่มีลักษณะการเขียนตัวเลือกแบบต่างกัน โดยข้อสอบแบบเลือกตอบของมหาวิทยาลัยสุโขทัยธรรมมาธิราชมีรูปแบบการเขียนตัวเลือกแตกต่างกัน จำแนกได้เป็น 7 ประเภท คือ 1) ข้อสอบที่มีตัวเลือก “ถูกทุกข้อ” เป็นตัวถูก 2) ข้อสอบที่มีตัวเลือก “ถูกทุกข้อ” เป็นตัวลวง 3) ข้อสอบที่มีตัวเลือก “ข้อ 1 และ 2 ถูก” หรือ “ข้อ 2 และ 3 ถูก” หรือ “ข้อ 1 และ 3 ถูก” หรือ “ข้อ 1 2 และ 3 ถูก” เป็นตัวถูก 4) ข้อสอบที่มีตัวเลือก “ข้อ 1 และ 2 ถูก” หรือ “ข้อ 2 และ 3 ถูก” หรือ “ข้อ 1 และ 3

ถูก” หรือ “ข้อ 1 2 และ 3 ถูก” เป็นตัวลวง 5) ข้อสอบที่มีตัวเลือก “ไม่มีข้อใดถูก” เป็นตัวถูก 6) ข้อสอบที่มีตัวเลือก “ไม่มีข้อใดถูก” เป็นตัวลวง 7) ข้อสอบที่มีตัวเลือกซ้อน การวิจัยนี้เป็นการวิจัยเชิงสำรวจ ประชากรการวิจัยคือ แบบทดสอบชุดวิชาที่เปิดสอบ ในการสอบไล่ ภาคการศึกษาที่ 1 ปีการศึกษา 2547 2548 และ 2549 รวม 3 ภาคการศึกษา การเลือกกลุ่มตัวอย่างได้จากการเลือกแบบเจาะจงชุดวิชาที่มีการเขียนตัวเลือกที่ต้องการศึกษาที่มีจำนวนข้อมากที่สุด จากภาคการศึกษาที่ 1 ปีการศึกษา 2547 เป็นเกณฑ์ ในการเลือก และศึกษาต่อเนื่อง ชุดวิชาที่เลือกอีก 2 ภาคการศึกษา คือ 2548 และ 2549 ได้กลุ่มตัวอย่างทั้งหมด 24 ชุดวิชา จำนวนแบบทดสอบรวม 72 ฉบับ การวิเคราะห์ข้อสอบรายชื่อของข้อสอบที่มีลักษณะ การเขียนตัวเลือกที่ศึกษากระทำโดยการวิเคราะห์ค่าความยากและอำนาจจำแนก เพื่อจำแนกประเภท ข้อสอบตามระดับความยากและระดับอำนาจจำแนก และหาค่าร้อยละของข้อสอบแต่ละประเภท

ผลการวิจัยพบว่า 1. ข้อสอบแบบเลือกตอบที่มีตัวเลือกถูกทุกข้อเป็นตัวถูก มีจำนวนทั้งสิ้น 499 ข้อ จากข้อสอบที่นำมาวิเคราะห์ทั้งหมด 8,370 ข้อ คิดเป็นร้อยละ 5.4 เป็นข้อสอบที่มีคุณภาพอยู่ในเกณฑ์ ใช้ได้เพียงร้อยละ 13.6 ข้อสอบที่มีคุณภาพไม่ถึงเกณฑ์ส่วนใหญ่เป็นข้อสอบที่ง่ายมาก (ค่าความยาก ระหว่าง .81 – 1.00) คิดเป็นร้อยละ 59 และค่อนข้างง่าย (ค่าความยากระหว่าง .61 – .80) ร้อยละ 29.0 และร้อยละ 68.8 มีอำนาจจำแนกต่ำกว่า .20 สาเหตุอาจเนื่องจากผู้สอบพิจารณาตัวเลือกแล้วเมื่อมี คำตอบถูกมากกว่า 1 ข้อ ผู้สอบก็สามารถเดาได้ว่าตัวเลือก ถูกทุกข้อเป็นคำตอบถูก 2. ข้อสอบแบบ เลือกตอบที่มีตัวเลือกถูกทุกข้อเป็นตัวลวง มีจำนวนทั้งสิ้น 400 ข้อ จากข้อสอบที่นำมาวิเคราะห์ทั้งหมด 8,370 ข้อ คิดเป็นร้อยละ 4.8 เป็นข้อสอบที่มีคุณภาพอยู่ในเกณฑ์ใช้ได้ ร้อยละ 45.3 ในจำนวน 400 ข้อนี้ เป็นข้อสอบที่ค่อนข้างยากและจำแนกได้น้อย ร้อยละ 15.3 ง่ายพอเหมาะ จำแนกได้น้อย ร้อยละ 11.3 ค่อนข้างง่ายจำแนกได้น้อย ร้อยละ 7.8 ข้อสอบที่มีคุณภาพไม่ถึงเกณฑ์ส่วนใหญ่เป็นข้อสอบที่ยากมาก (ค่า ความยากระหว่าง 0 – .19) คิดเป็น ร้อยละ 37.5 และร้อยละ 52.8 มีอำนาจจำแนกต่ำกว่า .20 เนื่องจาก ผู้สอบที่เดาตอบเมื่อเห็นตัวเลือกถูกทุกข้อก็จะเลือกทันที ซึ่งจะเห็นได้จากข้อสอบประเภทนี้ 202 ข้อ (ร้อย ละ 50.5) มีนักศึกษาเลือกตัวเลือกถูกทุกข้อเป็นคำตอบ ร้อยละ 50 – 100 3. ข้อสอบแบบเลือกตอบที่มี ตัวเลือก ข้อ 1 และ 2 ถูก ข้อ 2 และ 3 ถูก ข้อ 1 และ 3 ถูก ข้อ 1 2 และ 3 ถูก เป็นตัวถูก มีจำนวนทั้งสิ้น 724 ข้อ จากข้อสอบที่นำมาวิเคราะห์ทั้งหมด 8,370 ข้อ คิดเป็นร้อยละ 8.6 เป็นข้อสอบที่มีคุณภาพอยู่ใน เกณฑ์ใช้ได้ร้อยละ 28.9 ในจำนวน 724 ข้อนี้เป็นข้อสอบที่ค่อนข้างง่ายและจำแนกได้น้อย ร้อยละ 11.7 ง่ายพอเหมาะจำแนกได้น้อย ร้อยละ 9.7 ค่อนข้างยาก จำแนกได้น้อย ร้อยละ 5.6 ข้อสอบที่มีคุณภาพไม่ ถึงเกณฑ์ส่วนใหญ่เป็นข้อสอบที่มีอำนาจจำแนกต่ำกว่า .20 คิดเป็นร้อยละ 62.7 เนื่องจากผู้สอบจะคิดว่า ตัวเลือกประเภทนี้เป็นตัวถูก 4. ข้อสอบแบบเลือกตอบที่มีตัวเลือก ข้อ 1 และ 2 ถูก ข้อ 2 และ 3 ถูก ข้อ 1 และ 3 ถูก ข้อ 1 2 และ 3 ถูก เป็นตัวลวง มีจำนวนทั้งสิ้น 340 ข้อ จากข้อสอบที่นำมาวิเคราะห์ทั้งหมด 8,370 ข้อ คิดเป็นร้อยละ 4.0 เป็นข้อสอบที่มีคุณภาพอยู่ในเกณฑ์ใช้ได้ร้อยละ 40.0 ในจำนวน 340 ข้อ นี้ เป็นข้อสอบที่ค่อนข้างยากและจำแนกได้น้อย ร้อยละ 15.9 ง่ายพอเหมาะ จำแนกได้น้อย ร้อยละ 13.5 ค่อนข้างง่าย จำแนกได้น้อย ร้อยละ 4.7 ข้อสอบที่มีคุณภาพไม่ถึงเกณฑ์ส่วนใหญ่เป็นข้อสอบที่ยากมาก (ค่าความยากระหว่าง 0 – .19) คิดเป็นร้อยละ 33.5 และร้อยละ 50.3 มีอำนาจจำแนกต่ำกว่า .20

เนื่องจากผู้สอบที่เดาตอบเมื่อเห็นตัวเลือกประเภทนี้จะคิดว่าเป็นคำตอบที่ถูก ซึ่งจะเห็นได้จากข้อสอบประเภทนี้ 100 ข้อ (ร้อยละ 29.4) มีนักศึกษาเลือกตัวเลือกนี้เป็นคำตอบ ร้อยละ 30 – 49 และอีก 98 ข้อ (ร้อยละ 28.8) มีนักศึกษาเลือกตัวเลือกนี้เป็นคำตอบ ร้อยละ 50 – 100 5. ข้อสอบแบบเลือกตอบที่มีตัวเลือก ไม่มีข้อใดถูกเป็นตัวเลือก มีจำนวนทั้งสิ้น 37 ข้อ จากข้อสอบที่นำมาวิเคราะห์ทั้งหมด 8,370 ข้อ คิดเป็นร้อยละ 0.4 เป็นข้อสอบที่มีคุณภาพอยู่ในเกณฑ์ใช้ได้ ร้อยละ 51.4 ในจำนวน 37 ข้อนี้เป็นข้อสอบที่ค่อนข้างง่ายและจำแนกได้น้อย ร้อยละ 18.9 ค่อนข้างยาก จำแนกได้น้อย ร้อยละ 13.5 ง่ายพอเหมาะ จำแนกได้น้อย ร้อยละ 10.8 ข้อสอบที่มีคุณภาพไม่ถึงเกณฑ์ ส่วนใหญ่เป็นข้อสอบที่ยากมาก (ค่าความยากระหว่าง 0 – .19) คิดเป็นร้อยละ 20 และมีอำนาจจำแนกต่ำกว่า .20 ร้อยละ 37.8 เนื่องจากผู้สอบพิจารณาตัวเลือกแล้วเมื่อมีคำตอบถูกมากกว่า 1 ข้อ ผู้สอบก็สามารถเดาได้ว่าตัวเลือก ถูกทุกข้อเป็นคำตอบถูก เนื่องจากผู้สอบคาดว่าตัวเลือกประเภทนี้ไม่น่าจะเป็นตัวเลือก 6. ข้อสอบแบบเลือกตอบที่มีตัวเลือก ไม่มีข้อใดถูกเป็นตัวเลือก มีจำนวนทั้งสิ้น 236 ข้อ จากข้อสอบที่นำมาวิเคราะห์ทั้งหมด 8,370 ข้อ คิดเป็นร้อยละ 2.8 เป็นข้อสอบที่มีคุณภาพอยู่ในเกณฑ์ใช้ได้ ร้อยละ 48.3 ในจำนวน 236 ข้อนี้เป็นข้อสอบที่ง่ายพอเหมาะ จำแนกได้น้อย ร้อยละ 14.8 ค่อนข้างยาก และจำแนกได้น้อย ร้อยละ 13.1 ค่อนข้างง่าย จำแนกได้น้อย ร้อยละ 12.7 ข้อสอบที่มีคุณภาพไม่ถึงเกณฑ์ส่วนใหญ่เป็นข้อสอบที่ยากมาก (ค่าความยากระหว่าง 0 – .19) คิดเป็นร้อยละ 19.5 ค่อนข้างยาก (ค่าความยากระหว่าง .20 – .40) ร้อยละ 28.8 และมีอำนาจจำแนกต่ำกว่า .20 ร้อยละ 43.3 และ 7. ข้อสอบแบบเลือกตอบที่มีตัวเลือกซ้อน มีจำนวนทั้งสิ้น 108 ข้อจากข้อสอบที่นำมาวิเคราะห์ทั้งหมด 8,370 ข้อ คิดเป็นร้อยละ 1.3 เป็นข้อสอบที่มีคุณภาพอยู่ในเกณฑ์ใช้ได้ร้อยละ 43.5 ในจำนวน 108 ข้อนี้เป็นข้อสอบที่ค่อนข้างยากและจำแนกได้น้อย ร้อยละ 17.6 ง่ายพอเหมาะ จำแนกได้น้อย ร้อยละ 13.0 ค่อนข้างง่าย จำแนกได้น้อย ร้อยละ 8.3 ข้อสอบที่มีคุณภาพไม่ถึงเกณฑ์ส่วนใหญ่เป็นข้อสอบที่ยากมาก (ค่าความยากระหว่าง 0 – .19) คิดเป็นร้อยละ 17.6 ค่อนข้างยาก (ค่าความยากระหว่าง .20 – .40) ร้อยละ 32.4 และมีอำนาจจำแนกต่ำกว่า .20 ร้อยละ 48.1 เนื่องจากตัวเลือกซ้อนเป็นแนวทางหนึ่งที่ใช้สำหรับหลีกเลี่ยงแนวทางการเขียนข้อสอบแบบ “ถูกทุกข้อ” หรือ “ข้อ 1 และ 3 ถูก” หรือ “ข้อ 2 และ 3 ถูก” หรือ “ข้อ 1 2 และ 3 ถูก” จึงเป็นข้อสอบที่ค่อนข้างยากถึงยากมาก คิดเป็นร้อยละ 50.0

ข้อเสนอแนะในการนำผลการวิจัยไปใช้พบว่า 1. ข้อสอบที่มีตัวเลือก “ถูกทุกข้อ” ในกรณีที่ “ถูกทุกข้อ” ข้อสอบส่วนใหญ่จะเป็นข้อสอบที่ง่ายมาก และจำแนกไม่ได้ และในกรณีที่ “ถูกทุกข้อ” เป็นตัวเลือก ข้อสอบส่วนใหญ่จะเป็นข้อสอบที่ยากมาก และจำแนกไม่ได้ เนื่องจากผู้สอบเมื่อพบตัวเลือก “ถูกทุกข้อ” ก็จะตัดสินใจเลือก ทำให้การตอบถูกหรือผิดไม่ได้สะท้อนความสามารถที่แท้จริง ผู้ออกข้อสอบจึงควรหลีกเลี่ยง หากจำเป็นต้องออกข้อสอบที่ตัวเลือกถูกมีหลายตัว ควรเลี่ยงไปออกข้อสอบแบบตัวเลือกซ้อน เนื่องจากข้อสอบแบบตัวเลือกซ้อนร้อยละ 43.5 เป็นข้อสอบที่มีคุณภาพอยู่ในเกณฑ์ใช้ได้ 2. ข้อสอบที่มีตัวเลือก ข้อ 1 และ 2 ถูก ข้อ 2 และ 3 ถูก ข้อ 1 และ 3 ถูก ข้อ 1 2 และ 3 ถูก เป็นตัวเลือก ส่วนใหญ่เป็นข้อสอบที่มีอำนาจจำแนกต่ำ เนื่องจากผู้สอบจะเดาว่าข้อนี้เป็นคำตอบถูก และในกรณีที่ เป็นตัวเลือก ก็จะ เป็นข้อสอบที่ยากมาก และจำแนกไม่ได้ ผู้ออกข้อสอบจึงควรหลีกเลี่ยงตัวเลือกประเภทนี้ นอกจากนั้นตัว

ลวงประเภทนี้ ยังทำให้ยากต่อการปรับปรุงคุณภาพข้อสอบ เนื่องจากตัวเลือกแต่ละตัวไม่เป็นอิสระกัน 3. ข้อสอบที่มีตัวเลือก “ไม่มีข้อใดถูก” เป็นตัวถูก เป็นข้อสอบที่มีคุณภาพอยู่ในเกณฑ์ใช้ได้ถึงร้อยละ 51.4 และในกรณีที่ข้อสอบที่มีตัวเลือก “ไม่มีข้อใดถูก” เป็นตัวลวง เป็นข้อสอบที่อยู่ในเกณฑ์ใช้ได้ถึงร้อยละ 48.3 ผู้ออกข้อสอบจึงใช้ตัวเลือกนี้ได้ในโอกาสที่เหมาะสม เช่น สำหรับข้อสอบที่มีการคำนวณ ซึ่งผู้สอบคิดคำนวณออกมาแล้วไม่ตรงกับตัวเลือกใดเลย ก็จะเลือกตัวเลือกนี้ได้ และ 4. ควรนำผลจากการวิจัยนี้มาใช้ในการพิจารณาการเขียนตัวเลือกทั้ง 7 ประเภท เพื่อเป็นแนวทางในการพัฒนาคุณภาพของข้อสอบรายข้อ และกำหนดสัดส่วนของระดับความยากง่ายของแบบทดสอบทั้งฉบับ ว่าควรมีสัดส่วนของข้อสอบที่มีระดับความยากระดับปานกลาง และระดับง่าย ที่ร้อยละเท่าใดจึงจะมีความพอเหมาะ

งานวิจัยที่เกี่ยวข้องกับการปรับเทียบคะแนน

อุทัยวรรณ พงศ์อร่าม (2545) ศึกษาขนาดตัวอย่างที่เหมาะสมสำหรับการปรับเทียบคะแนน ด้วยวิธีอิกวิเปอร์เซ็นไทล์ และวิธีเส้นตรงตามแบบจำลองคะแนนจริงสัมพันธ์ โดยใช้กลุ่มตัวอย่างขนาด 7 ขนาด (30 50 100 200 300 400 และ 500 คน) โดยใช้ข้อสอบรวม พบว่าการปรับเทียบคะแนนด้วยวิธีอิกวิเปอร์เซ็นไทล์แบบแผนกลุ่มสมมูล ที่ใช้ความยาวของแบบสอบ 60 ข้อ ขนาดตัวอย่างที่เหมาะสม คือ 200 คน ความยาวของแบบสอบ 40, 50 ข้อ ขนาดตัวอย่างที่เหมาะสม คือ 300 คน ความยาวของแบบสอบ 30 ข้อ ขนาดตัวอย่างที่เหมาะสม คือ 400 คน ส่วนการปรับเทียบคะแนนด้วยวิธีอิกวิเปอร์เซ็นไทล์แบบแผนข้อสอบรวม ที่ใช้ความยาวของแบบสอบ 50, 60 ข้อ ขนาดตัวอย่างที่เหมาะสม คือ 200 คน ความยาวของแบบสอบ 30, 40 ข้อ ขนาดตัวอย่างที่เหมาะสม คือ 500 คน และการปรับเทียบคะแนนด้วยวิธีเชิงเส้นตรงทั้งแบบแผนกลุ่มสมมูลและแบบแผนข้อสอบรวม ใช้กลุ่มตัวอย่างน้อยกว่าวิธีอิกวิเปอร์เซ็นไทล์

สร้อยญา จันทรชูสกุล ไชยยศ ไพวิทยศิริธรรม ยุวรี ผลพันธิน และพิทักษ์ สุพรรณโณภาพ (2560) ศึกษาการเปรียบเทียบคุณภาพของการปรับเทียบคะแนนสอบภาษาอังกฤษระหว่างวิธีอิกวิเปอร์เซ็นไทล์ วิธีเชิงเส้นตรง และวิธีสมการถดถอยกับนักศึกษาปริญญาตรี มหาวิทยาลัยศิลปากร จำนวน 40 คน โดยใช้แบบสอบ Silpakorn English Proficiency Test (SEPT) และแบบสอบ Test of English for International Communication (TOEIC) พบว่าทั้ง 3 วิธีมีค่าเฉลี่ยของความคลาดเคลื่อนไม่แตกต่างกัน และคะแนนจาก SEPT ที่ปรับเทียบด้วยวิธีอิกวิเปอร์เซ็นไทล์ วิธีเชิงเส้นตรง และวิธีสมการถดถอย สามารถทำนายค่าคะแนน TOEIC ได้ร้อยละ 62.6, 62.4 และ 60.2 ตามลำดับ โดยที่ทั้ง 3 วิธีมีประสิทธิภาพในการทำนายไม่แตกต่างกัน

Purya Baghaei & Nazila Amrahi ได้ศึกษาเกี่ยวกับผลกระทบของจำนวนตัวเลือกต่อคุณลักษณะทางจิตมิติของข้อสอบปรนัยแบบเลือกตอบ การศึกษานี้มีวัตถุประสงค์เพื่อกำหนดจำนวนตัวเลือกที่เหมาะสมที่สุดสำหรับข้อสอบแบบปรนัย ในการวิจัยครั้งนี้ผู้วิจัยได้ศึกษาจากข้อสอบในรายวิชาภาษาอังกฤษที่วัดความรู้เกี่ยวกับคำศัพท์จำนวน 3 ฟอรัม คือ ฟอรัม A ฟอรัม B และฟอรัม C โดยศึกษากับนักศึกษามหาวิทยาลัยที่เรียนในสาขาวิชาเอกภาษาอังกฤษทั้งที่เป็นนักศึกษาในระดับบัณฑิตศึกษาและ

นักศึกษาระดับปริญญาตรีจำนวน 180 คน ซึ่งข้อสอบทั้ง 3 ฟอรมนี้มีความแตกต่างกันตรงที่จำนวนตัวเลือกที่แตกต่างกันในข้อสอบแต่ละข้อ ข้อสอบทั้ง 3 ฟอรมจะมีจำนวน 30 ข้อเท่ากันและข้อสอบทั้ง 3 ฟอรมจะมีข้อรวมที่เชื่อมกันจำนวน 10 -ข้อเพื่อนำไปสู่แนวคิดของการวิเคราะห์การเปรียบเทียบข้อสอบ โดยข้อสอบฟอรม A จะเป็นข้อสอบที่มี 5 ตัวเลือก ข้อสอบฟอรม B จะเป็นข้อสอบที่มี 4 ตัวเลือก และ ข้อสอบฟอรม C จะเป็นข้อสอบที่มี 3 ตัวเลือก การวิจัยครั้งนี้ใช้การวิเคราะห์แบบจำลอง Rasch เพื่อเปรียบเทียบความค่าความยากของข้อสอบ ค่าสถิติ เช่นค่าเฉลี่ย ค่าสัมประสิทธิ์สหสัมพันธ์สัมพันธ์ระหว่างเฉลี่ยกับตัวलग พฤติกรรมการตอบสนองของบุคคล ยกเว้นอำนาจจำแนกของตัวलगเพราะจำนวนตัวलगในข้อสอบแต่ละฟอรมไม่เท่ากัน ผลการวิจัยพบว่า เปิดเผยว่าจำนวนตัวเลือกในข้อสอบที่แตกต่างกันไม่พบการเปลี่ยนแปลงที่มีนัยสำคัญทางสถิติทั้งในมิติของการพิจารณาค่าความยาก ค่าสถิติอื่น พฤติกรรมการตอบสนองของบุคคล และค่าความเที่ยงระหว่างข้อสอบทั้ง 3 ฟอรม และเมื่อพิจารณาถึงระยะเวลาที่ใช้ในการสอบและพลังงานที่จำเป็นสำหรับการออกข้อสอบในการสร้างตัวलगจำนวนมาก ๆ นั้น งานวิจัยเรื่องนี้พบว่าข้อสอบปรนัยแบบเลือกตอบจำนวน 3 ตัวเลือก เป็นรูปแบบของข้อสอบที่เหมาะสมที่สุด

Brian McGuire (2012) ได้ศึกษาและนำเสนอเกี่ยวกับแบบจำลองทฤษฎีการตอบสนองรายการ (IRT) เพื่อช่วยให้นักวิจัยสามารถประเมินแบบทดสอบหรือประเมินคุณลักษณะทางจิตมิติของข้อสอบเพื่อการพิจารณาตัดสินเกี่ยวกับค่าความยากและคุณภาพของได้แม่นยำยิ่งขึ้น การวิเคราะห์ IRT ส่วนใหญ่ใช้ตัวแบบพาราเมตริก ซึ่งมักจะไม่เป็นไปตามสมมติฐานที่จำเป็นของแบบจำลองเหล่านี้ แพ็คเกจ KernSmoothIRT ใช้เคอร์เนลที่ปรับให้เรียบจาก Ramsay (1991) จึงถูกนำมาใช้เพื่อประเมินเส้นโค้งลักษณะรายการและตัวเลือก รวมทั้งสร้างการทดสอบหลาย ๆ แบบและแปลงตามหัวเรื่อง การวิเคราะห์ IRT แบบไมอิงพาราเมตริกไม่ได้อาศัยสมมติฐานของวิธีพาราเมตริกทั่วไป แพ็คเกจนี้ใช้งานง่ายและเป็นมิตรกับผู้ใช้ สามารถใช้ได้กับข้อสอบที่มีลักษณะแบบหนึ่งตัวเลือก หลายตัวเลือก หรือการตอบสนองตามมาตราส่วนก็ได้

Hongwen Guo, Jiyun Zu and Patrick Kyllonen (2018) ได้ทำการวิจัยเกี่ยวกับการสร้างแบบจำลองเพื่อศึกษาความเหมาะสมของจำนวนตัวเลือกในข้อสอบแบบเลือกตอบ โดยในงานวิจัยนี้มุ่งชี้ประเด็นว่า สำหรับการทดสอบแบบปรนัยภายใต้การพัฒนาหรือการออกแบบใหม่ สิ่งสำคัญคือต้องเลือกจำนวนตัวเลือกที่เหมาะสมที่สุดต่อรายการ เพื่อให้การทดสอบมีคุณสมบัติทางจิตมิติของข้อสอบได้ตามที่ต้องการ ในการวิจัยครั้งนี้ได้ทำการจำลองข้อมูลบนพื้นฐานของข้อมูลที่มีอยู่สำหรับการประเมินแบบปรนัยที่มี โดยเริ่มต้นจากแบบทดสอบที่มี 8 ตัวเลือก และจะตัดทอนจำนวนตัวเลือกลงไปเรื่อย ๆ เพื่อศึกษาผลกระทบของการเปลี่ยนแปลงจำนวนตัวเลือกที่มีต่อคุณสมบัติทางจิตมิติของข้อสอบ ซึ่งประกอบไปด้วยค่าความยาก ความเที่ยง และการเปรียบเทียบคะแนน โดยใช้การจำลอง ข้อมูลด้วยการกำหนดเกณฑ์ขึ้นมา 2 เกณฑ์ (ตัดตัวलगที่มีจำนวนคนเลือกน้อยออกไปจากข้อสอบแต่ละข้อและตัดตัวलगที่มีค่าอำนาจจำแนกไม่ต้อออก) และทดลองแทนค่าตัวलगที่ตัดออกไปด้วยวิธีการ 2 รูปแบบคือ (การแทนค่าแบบสุ่มและการแทนค่าตามระดับความสามารถของผู้ตอบ) เพื่อจำลองพฤติกรรมการตอบสนองตามสมมติฐานสำหรับตัวเลือกที่ตัดออก งานวิจัยนี้พบว่า การลดจำนวนตัวเลือกจาก 8 ตัวเลือกให้มีจำนวนน้อยลง ทำให้รูปแบบ

การทดสอบง่ายขึ้น แต่รูปแบบการทดสอบที่มีตัวเลือกที่ลดลงอาจมีความเที่ยงมากขึ้นหากตัวเลือกนั้นลงคนมาตอบได้น้อยหรือไม่มีคุณภาพด้านอำนาจจำแนกที่ติดตามเกณฑ์ที่ถูกตัดออก ซึ่งผลจากการใช้วิธีการแทนค่าตัวลงที่ถูกตัดออกตามระดับความสามารถของผู้ตอบ พบว่า จำนวนตัวเลือกที่เหมาะสมที่สุดสำหรับการทดสอบนี้โดยเฉลี่ยจะอยู่ที่ประมาณ 5 ตัวเลือก ถึงแม้ว่าตัวลงในข้อสอบจะมีจำนวนน้อยลงแต่คุณสมบัติทางจิตมิติของข้อสอบยังคงได้คุณภาพครบถ้วนในทุกมิติ

จากการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้องจึงเป็นเหตุที่ผู้วิจัยสนใจจะทำงานวิจัยเรื่องนี้ คือ มุ่งที่จะศึกษาว่าหาต้องการข้อเสนอแนะสำหรับนำไปใช้เป็นแนวทางในการประหยัดค่าใช้จ่ายในการจัดทำแบบทดสอบของมหาวิทยาลัยสุโขทัยธรรมมาธิราช ในแง่ของการปรับเปลี่ยนรูปแบบของการจัดทำแบบทดสอบแบบเลือกตอบ จากเดิมที่ต้องจัดทำข้อสอบโดยสร้างแบบทดสอบแบบเลือกตอบให้มีตัวเลือกถึง 5 ตัวเลือก แต่จากการศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องพบว่า การลดตัวเลือกลงไม่ส่งผลต่อคุณภาพของแบบทดสอบ ผู้วิจัยจึงต้องการจะยืนยันข้อค้นพบนี้อีกครั้งว่าจะเมื่อผู้วิจัยทำการวิจัยซ้ำอีกครั้งข้อค้นพบนี้จะยังเป็นจริงอยู่หรือไม่

ด้วยข้อจำกัดของการทำวิจัยของผู้วิจัย คือ ผู้วิจัยไม่สามารถสร้างแบบทดสอบเพื่อนำแบบทดสอบที่ผู้วิจัยสร้างขึ้นไปทำการสอบกับนักศึกษาได้โดยตรง เนื่องจากผู้วิจัยไม่มีบทบาทและหน้าที่ในส่วนของการออกข้อสอบ ดังนั้นผู้วิจัยจึงจำเป็นต้องใช้ข้อมูลทฤษฎีภูมิซึ่งเป็นคะแนนผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกของนักศึกษามาเป็นข้อมูลตั้งต้น แล้วทำการจำลองสถานการณ์การสอบขึ้นมาว่า ถ้าตัดตัวเลือกของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกลง ให้ได้แบบทดสอบฉบับใหม่ที่มีเพียง 4 ตัวเลือก แล้วคาดการณ์การตอบแบบทดสอบแบบเลือกตอบ 4 ตัวเลือกของนักศึกษา ตามข้อตกลงเบื้องต้นที่ผู้วิจัยกำหนดขึ้น 3 ข้อตกลงเบื้องต้นด้วยกัน เหตุที่ต้องทำเช่นนี้เพราะผู้วิจัยไม่สามารถนำแบบทดสอบแบบ 4 ตัวเลือกที่ผู้วิจัยสร้างขึ้นไปทำการสอบจริงกับนักศึกษาได้ เพื่อหาข้อสรุปว่าคุณภาพของแบบทดสอบที่ถูกตัดตัวเลือกให้เหลือน้อยลง จะยังคงมีคุณภาพแตกต่างกันกับแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกที่มหาวิทยาลัยสุโขทัยธรรมมาธิราชใช้อยู่หรือไม่ นอกจากนั้นยังต้องการยืนยันอีกว่าผลคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ที่ได้จากการจำลองและคาดการณ์สถานการณ์การตอบแบบทดสอบตามข้อตกลงเบื้องต้นของผู้วิจัยจะมากกว่าคะแนนสอบที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกหรือไม่ ผู้วิจัยจึงใช้การเปรียบเทียบคะแนนด้วยวิธีการที่ง่าย เป็นที่นิยมใช้ และมีคำสั่งในโปรแกรมสำเร็จรูปทางสถิติ คือ วิธีการเปรียบเทียบเชิงเส้น มาปรับคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกและ 4 ตัวเลือกให้อยู่ในมาตราเดียวกัน แล้วทำการเปรียบเทียบคะแนนที่ได้จากการเปรียบเทียบว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือกมากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกหรือไม่ ตลอดจนอยากได้สารสนเทศอีกว่า หากมีการปรับลดตัวเลือกของแบบทดสอบลงจะทำให้ประหยัดค่าใช้จ่ายในการจัดทำแบบทดสอบมากน้อยเพียงใด

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงปริมาณ (Quantitative Research) โดยใช้ข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ของนักศึกษาที่กำลังศึกษาในหลักสูตรระดับปริญญาตรีของมหาวิทยาลัยสุโขทัยธรรมมาธิราช ที่มีสิทธิ์สอบในภาคการศึกษาที่ 2/2562 ในรายวิชาที่ถูกเลือกมาศึกษาในงานวิจัยนี้เป็นข้อมูลตั้งต้น จากนั้นผู้วิจัยจึงทำการวิเคราะห์ข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกเพื่อคำนวณหาคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ด้วยการคำนวณหาค่าอำนาจจำแนกและผลการวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบด้วยการทำให้เรียบเคอเนล แล้วใช้ผลที่ได้ในการตัดตัวเลือกที่ไม่มีประสิทธิภาพออก 1 ตัวเลือก เพื่อให้ได้แบบทดสอบฉบับใหม่ที่เหลือตัวเลือกเพียง 4 ตัวเลือก แล้วจึงใช้วิธีการจำลอง (Simulation) สถานการณ์การตอบของผู้เข้าสอบโดยใช้แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก เพื่อจัดหาตัวเลือกสำหรับการแทนที่ตัวเลือกที่ตรงกับตัวเลือกที่ถูกตัดออกโดยใช้เกณฑ์การพิจารณาค่าอำนาจจำแนกและผลการวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบด้วยการทำให้เรียบเคอเนล โดยอาศัยข้อตกลงเบื้องต้นที่ผู้วิจัยกำหนดขึ้นในแต่ละกรณี ทั้งหมด 3 กรณี มาใช้ในการจัดหาตัวเลือกเพื่อการแทนค่าตัวเลือกที่ถูกตัดออกให้กับแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ในแต่ละข้อ จนครบตามจำนวนของผู้เข้าสอบแต่ละคนที่ต้องถูกตัดตัวเลือกนั้นออกไป แล้วจึงนำข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และข้อมูลผลการจำลองสถานการณ์การตอบแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ในแต่ละกรณีของการแทนค่าตัวเลือกที่ถูกตัดออก ทั้ง 3 กรณี มาทำการศึกษาคูณภาพของแบบทดสอบ ทำการเปรียบเทียบคะแนนเพื่อตอบสมมติฐานที่ผู้วิจัยตั้งไว้ แล้วคำนวณค่าใช้จ่ายหลังจากที่มีการลดจำนวนตัวเลือกของแบบทดสอบ

ข้อมูลที่ใช้ในการวิจัย

ข้อมูลที่ใช้ในการวิจัยครั้งนี้แบ่งออกเป็น 2 ส่วน คือ ข้อมูลหัตถุติภูมิของผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และข้อมูลที่ได้จากการจำลองสถานการณ์การตอบแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก 3 กรณี ดังนี้

ส่วนที่ 1 เป็นข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ของนักศึกษาที่กำลังศึกษาในหลักสูตรระดับปริญญาตรีของมหาวิทยาลัยสุโขทัยธรรมมาธิราช ที่มีสิทธิ์สอบในภาคการศึกษาที่ 2/2562 ในรายวิชาที่ถูกเลือกมาศึกษาในงานวิจัยนี้เป็นข้อมูลตั้งต้น ซึ่งเป็นข้อมูลหัตถุติภูมิ (Secondary Data) โดยแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก มีจำนวนข้อสอบทั้งหมด 120 ข้อ และเป็นข้อสอบแบบปรนัยหรือแบบเลือกตอบทั้ง 120 ข้อ ซึ่งเป็นผลการตอบแบบทดสอบของนักศึกษา จำนวน 1,966 คน ข้อมูลชุดนี้ผู้วิจัยได้มาจากการทำหนังสือขอความอนุเคราะห์ผ่านไปยังท่านผู้อำนวยการสำนักทะเบียนและวัดผล ซึ่งได้รับความอนุเคราะห์ข้อมูลจากหัวหน้าหน่วยวิเคราะห์แบบทดสอบ ศูนย์วิจัยและพัฒนาแบบทดสอบ สำนักทะเบียนและวัดผล มหาวิทยาลัยสุโขทัยธรรมมาธิราช

ส่วนที่ 2 เป็นข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ของนักศึกษาที่กำลังศึกษาในหลักสูตรระดับปริญญาตรีของมหาวิทยาลัยสุโขทัยธรรมมาธิราช ที่มีสิทธิ์สอบในภาคการศึกษาที่ 2/2562 ในรายวิชาที่ถูกเลือกมาศึกษาในงานวิจัยส่วนหนึ่งสำหรับข้อสอบบางข้อของนักศึกษาบางคนที่ไม่ได้ถูกตัดตัวเลือก ผวนวเข้ากับข้อมูลอีกส่วนหนึ่งซึ่งเป็นข้อมูลที่ได้จากการจำลองสถานการณ์การตอบแบบทดสอบแบบเลือกตอบ 4 เลือก ของนักศึกษา ซึ่งแบบทดสอบแบบเลือกตอบ 4 ตัวเลือกนี้ เกิดจากการตัดตัวเลือกจากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกออก โดยอาศัยเกณฑ์การพิจารณาค่าอำนาจจำแนกและผลการวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบด้วยการทำให้เรียบเคอเนล โดยผู้วิจัยจะใช้วิธีการจำลองการตอบแบบทดสอบด้วยการแทนที่ตัวเลือกที่ถูกตัดออกตามข้อตกลงเบื้องต้นที่ผู้วิจัยตั้งขึ้น 3 กรณี โดยจะแทนที่ตัวเลือกที่ถูกตัดออกให้กับข้อสอบแต่ละข้อที่ตรงกับเกณฑ์การพิจารณาตัดตัวเลือกข้างต้น ให้กับผู้เข้าสอบแต่ละคนที่ถูกตัดตัวเลือกทั้งไปเท่านั้น สำหรับผู้เข้าสอบที่ในข้อสอบข้อนั้น ๆ ไม่ถูกตัดตัวเลือก ผู้วิจัยก็ยังจะใช้ผลการตอบแบบทดสอบคงเดิม เหมือนกับคำตอบที่ผู้เข้าสอบตอบในแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ตามข้อมูลทฤษฎีที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกที่เป็นข้อมูลตั้งต้น เนื่องจากผู้วิจัยจำลองวิธีการเก็บรวบรวมข้อมูลด้วยรูปแบบผู้สอบกลุ่มเดียวที่ไม่ได้รับการจัดให้สมดุล (Single – Group Uncounterbalanced Design) โดยใช้กลุ่มผู้สอบกลุ่มเดียวทำหน้าที่เป็นผู้สอบรวม กล่าวคือ ผู้สอบคนเดียวกันทำแบบสอบทั้ง 2 ฉบับ ดังนั้นระดับความสามารถของผู้สอบทั้ง 2 ฉบับ เท่ากัน แสดงว่าคำตอบของผู้สอบที่สอบด้วยแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก จะเหมือนกับคำตอบของผู้สอบที่สอบด้วยแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

ข้อมูลในส่วนที่ 2 นี้ จะแบ่งออกเป็น 3 ชุดข้อมูล

ชุดที่ 1 เป็นข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ตามข้อมูลตั้งต้น สำหรับผู้เข้าสอบที่ไม่ถูกตัดตัวเลือกในข้อสอบข้อนั้น ๆ ออก ผวนวกับข้อมูลจากการจำลองผลการตอบแบบทดสอบ 4 ตัวเลือก สำหรับข้อสอบข้อนั้น ๆ ที่ตรงกับตัวเลือกที่ถูกตัดออก ของผู้เข้าสอบคนนั้น ๆ ที่มีข้อสอบที่ตรงกับตัวเลือกที่ถูกตัดออก การแทนที่ตัวเลือกที่ถูกตัดออกจะเป็นไปตาม **ข้อตกลงเบื้องต้นของผู้วิจัยที่ว่า หากผู้สอบคนนั้นมีความสามารถสูงจะต้องตอบตัวเลือกที่เป็นเฉลยตั้งแต่ครั้งแรกที่ลงมือทำข้อสอบ**

ชุดที่ 2 เป็นข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ตามข้อมูลตั้งต้น สำหรับผู้เข้าสอบที่ไม่ถูกตัดตัวเลือกในข้อสอบข้อนั้น ๆ ออก ผวนวกับข้อมูลจากการจำลองผลการตอบแบบทดสอบ 4 ตัวเลือก สำหรับข้อสอบข้อนั้น ๆ ที่ตรงกับตัวเลือกที่ถูกตัดออก ของผู้เข้าสอบคนนั้น ๆ ที่มีข้อสอบที่ตรงกับตัวเลือกที่ถูกตัดออก การแทนที่ตัวเลือกที่ถูกตัดออกจะเป็นไปตาม **ข้อตกลงเบื้องต้นของผู้วิจัยที่ว่า ผู้สอบทุกคนมีสิทธิ์ที่จะเลือกตอบตัวเลือกใด ๆ ก็ได้ นั่นหมายความว่า ตัวเลือกแต่ละตัวมีโอกาสที่ผู้สอบจะเลือกตอบเท่า ๆ กัน ผู้สอบที่มีความสามารถสูงก็มีสิทธิ์จะเลือกตัวเลือกที่เป็นได้ทั้งตัวลงและเฉลย เช่นเดียวกับผู้สอบที่มีความสามารถต่ำ**

ชุดที่ 3 เป็นข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ตามข้อมูลตั้งต้น สำหรับผู้เข้าสอบที่ไม่ถูกตัดตัวเลือกในข้อสอบข้อนั้น ๆ ออก ผวนวกับข้อมูลจากการจำลองผลการตอบแบบทดสอบ 4 ตัวเลือก สำหรับข้อสอบข้อนั้น ๆ ที่ตรงกับตัวเลือกที่ถูกตัดออก ของผู้เข้าสอบคนนั้น ๆ ที่มีข้อสอบที่ตรงกับ

กับตัวเลือกที่ถูกตัดออก การแทนที่ตัวเลือกที่ถูกตัดออกจะเป็นไปตาม ข้อตกลงเบื้องต้นของผู้วิจัยที่ว่า ผู้สอบที่มีความสามารถสูงจะต้องทำข้อสอบถูกทุกครั้ง ทุกกรณี แม้ว่าจะมีการปรับลดตัวเลือกของ แบบทดสอบลงจาก 5 ตัวเลือก เหลือเพียง 4 ตัวเลือก ผู้สอบที่มีความสามารถสูงจะต้องทำข้อสอบข้อ นั้น ๆ ถูกต้องเสมอ และสำหรับผู้ที่มีความสามารถต่ำก็มีโอกาสที่จะตอบตัวเลือกอื่น ๆ ที่ไม่ใช่ตัวเลือก ที่เป็นเฉลย ด้วยความน่าจะเป็นในการเลือกตอบแต่ละตัวเลือกเท่า ๆ กัน เพราะคนที่มีความสามารถ ต่ำ แม้ว่าจะมีการปรับลดตัวเลือกของแบบทดสอบลงจาก 5 ตัวเลือก เหลือเพียง 4 ตัวเลือก ผู้สอบที่มี ความสามารถต่ำ ก็ไม่มีโอกาสที่จะตอบข้อสอบข้อนั้นได้ถูกต้อง

โดยข้อมูลทั้ง 3 ชุดนี้ ถือว่าเป็นข้อมูลที่ได้จากการจำลองสถานการณ์การตอบแบบทดสอบแบบ เลือกตอบ 4 ตัวเลือก ซึ่งมีจำนวนข้อสอบเท่ากับ 120 ข้อ และได้มาจากการตอบแบบทดสอบของผู้เข้า สอบ จำนวน 1,966 คน

ขั้นตอนการดำเนินการวิจัย

1) ผู้วิจัยทำหนังสือบันทึกข้อความเพื่อขอความอนุเคราะห์ผู้อำนวยการสำนักทะเบียนและวัดผล ในการขอข้อมูลชุดข้อมูลผลการตอบแบบทดสอบของนักศึกษาที่กำลังศึกษาในหลักสูตรระดับปริญญาตรี ของมหาวิทยาลัยสุโขทัยธรรมาธิราชที่สอบในปีการศึกษา 2562 ถึงปีการศึกษา 2564 และข้อมูลค่าใช้จ่าย ในการจัดทำแบบทดสอบ

2) ผู้อำนวยการสำนักทะเบียนและวัดผลอนุมัติให้ผู้วิจัยสามารถขอข้อมูลผลการตอบแบบทดสอบ ได้ และผู้วิจัยได้ทำการติดต่อหน่วยงานที่รับผิดชอบในการเก็บข้อมูลของนักศึกษา คือ หน่วยวิเคราะห์ แบบทดสอบ ศูนย์วิจัยและพัฒนาแบบทดสอบ สำนักทะเบียนและวัดผล จากนั้นผู้วิจัยจึงติดต่อและ ประสานงานหัวหน้าหน่วยวิเคราะห์แบบทดสอบเพื่อขอข้อมูลผลการตอบแบบทดสอบ และข้อมูลที่ เกี่ยวข้องอื่น ๆ เช่น ชื่อชุดวิชา รหัสชุดวิชา เฉลยในแต่ละชุดวิชาที่มีการเปิดสอบในปีการศึกษา 2562 ถึงปี การศึกษา 2564 โดยผู้วิจัยได้เล่าให้หัวหน้าหน่วยวิเคราะห์แบบทดสอบฟังว่าผู้วิจัยต้องการข้อมูลอะไร ข้อมูล ในลักษณะใดบ้าง และนัดวันเพื่อขอเข้ารับข้อมูลจากหัวหน้าหน่วยวิเคราะห์แบบทดสอบ สำหรับข้อมูล เกี่ยวกับค่าใช้จ่ายในการจัดทำแบบทดสอบ ผู้วิจัยต้องติดต่อประสานงานกับหัวหน้าหน่วยจัดทำ แบบทดสอบ ศูนย์วิจัยและพัฒนาแบบทดสอบ สำนักทะเบียนและวัดผล เพื่อขอข้อมูล โดยผู้วิจัยได้บอก เล่าแก่หัวหน้าหน่วยจัดทำแบบทดสอบว่าผู้วิจัยต้องการข้อมูลอะไร ข้อมูลลักษณะใดบ้าง และนัดวันเพื่อ ขอเข้ารับข้อมูลจากหัวหน้าหน่วยจัดทำแบบทดสอบ

3) ผู้วิจัยได้รับข้อมูลผลการตอบแบบทดสอบของนักศึกษาที่กำลังศึกษาในหลักสูตรระดับปริญญา ตรีของมหาวิทยาลัยสุโขทัยธรรมาธิราชที่สอบในปีการศึกษา 2562 ถึงปีการศึกษา 2564 และข้อมูล ค่าใช้จ่ายในการจัดทำแบบทดสอบ แล้วจึงทำการตรวจสอบว่าข้อมูลที่ได้รับมีความครบถ้วน ถูกต้อง สมบูรณ์ หรือไม่

4) ผู้วิจัยทำการพิจารณาคัดเลือกข้อมูลผลการตอบแบบทดสอบของนักศึกษาที่สอบในปีการศึกษา 2562 ถึงปีการศึกษา 2564 โดยพิจารณาเฉพาะชุดวิชาที่เป็นผลการตอบแบบทดสอบแบบเลือกตอบ

เท่านั้น ผู้วิจัยมีเกณฑ์ในการคัดเลือกชุดวิชา คือ ต้องการได้ชุดวิชาที่มีนักศึกษาเข้าสอบเป็นจำนวนมาก ถ้ามากกว่า 1,000 คน ได้จะเป็นการดีในการที่จะสามารถใช้วิธีการวิเคราะห์ที่ให้ผลลัพธ์เป็นที่น่าเชื่อถือ และต้องเป็นข้อมูลที่ครบถ้วน สมบูรณ์ มีความคลาดเคลื่อนน้อย จากการสอบถามหัวหน้าหน่วยวิเคราะห์แบบทดสอบ ท่านให้ข้อมูลว่าข้อมูลผลการตอบแบบทดสอบในช่วงที่เกิดการแพร่ระบาดของโรคติดเชื้อโคโรนาไวรัสโควิด 19 ในช่วงปีการศึกษา 2563 ถึงปีการศึกษา 2564 ไม่สมบูรณ์นัก อาจทำให้ข้อมูลที่เกิดขึ้นมีความคลาดเคลื่อน ดังนั้นผู้วิจัยจึงตัดสินใจเลือกใช้ข้อมูลผลการตอบแบบทดสอบในปีการศึกษา 2562 และเลือกใช้ของภาคการศึกษาที่ 2 ในชุดวิชาที่มีนักศึกษาเข้าสอบจำนวนมากกว่า 1,000 คน ซึ่งก็คือ ชุดวิชาที่ผู้วิจัยเลือกมาใช้เป็นข้อมูลตั้งต้น ซึ่งเป็นชุดวิชาที่มีนักศึกษาเข้าสอบจำนวน 1,966 คน และมีจำนวนข้อสอบ 120 ข้อ

5) ผู้วิจัยนำข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ในภาคการศึกษาที่ 2 ปีการศึกษา 2562 ซึ่งมีข้อสอบจำนวน 120 ข้อ และเป็นผลการตอบแบบทดสอบของนักศึกษา จำนวน 1,966 คน มาทำการตรวจสอบความถูกต้องของข้อมูลว่าข้อมูลมีความครบถ้วน สมบูรณ์หรือไม่ มีข้อมูลสูญหายหรือไม่ หรือมีข้อมูลที่เป็นค่าผิดปกติหรือไม่ ซึ่งจะพบว่าในข้อมูลจะมีนักศึกษาบางคนไม่เลือกตอบแบบทดสอบ โดยจะมีการลงรหัสเป็นเลข 6 และมีนักศึกษาส่วนหนึ่งที่ถูกลงรหัสเป็นเลข 7 สำหรับในส่วนนี้จะกระทบต่อการตัดและแทนค่าตัวเลือก ดังนั้นผู้วิจัยจึงจำเป็นต้องสุ่มและแทนที่ตัวเลือกให้ใหม่ในตัวเลือก 1, 2, 3, 4, และ 5 ให้แก่นักศึกษาที่มีรหัสเป็น 6 และ 7 ก่อนที่จะนำข้อมูลไปทำการวิเคราะห์ในลำดับถัดไป

6) ผู้วิจัยนำข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ที่ผ่านการตรวจสอบความถูกต้องของข้อมูล การแทนค่าข้อมูลที่สูญหายหรือผิดปกติ มาทำการหาค่าอำนาจจำแนกของข้อสอบจำนวน 120 ข้อ โดยใช้แพ็คเกจ psych ในโปรแกรม R

7) ผู้วิจัยนำข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ที่ผ่านการตรวจสอบความถูกต้องของข้อมูล การแทนค่าข้อมูลที่สูญหายหรือผิดปกติ มาทำการสร้างโค้งคุณลักษณะตัวเลือก (Option Characteristic Curves : OCC) ของข้อสอบ จำนวน 120 ข้อ โดยใช้แพ็คเกจ KernSmoothIRT ในโปรแกรม R

8) ผู้วิจัยนำผลการคำนวณหาค่าอำนาจจำแนกของข้อสอบแต่ละข้อ มาพิจารณาร่วมกับโค้งคุณลักษณะตัวเลือก (OCC) ของข้อสอบแต่ละข้อ เพื่อทำการตัดตัวเลือก (ตัวลง) ที่ไม่มีประสิทธิภาพออกจนครบ 120 ข้อ วิธีการดังกล่าวจะทำให้ผู้วิจัยได้แบบทดสอบฉบับใหม่ 1 ฉบับ จากเดิมมี 5 ตัวเลือกกลายเป็น 4 ตัวเลือก โดยมีเกณฑ์การพิจารณาตัดตัวเลือกออก โดยใช้ค่าอำนาจจำแนก (Guo, Zu and Kyllonen, 2018) และโค้งคุณลักษณะตัวเลือก ดังนี้

เกณฑ์การพิจารณาตัดตัวเลือกจากค่าอำนาจจำแนก (ผลลัพธ์ที่แสดงในโปรแกรม R จะแทนค่าอำนาจจำแนกด้วย discrim) คือ

- ถ้าข้อสอบข้อใดมีตัวเลือก (ตัวลวง) ที่มีเฉพาะค่า discrim ที่เป็นค่าบวก ให้พิจารณาตัดตัวลวงที่มีค่าบวกมากที่สุดออก
- ถ้าข้อสอบข้อใดมีตัวเลือก (ตัวลวง) ที่มีเฉพาะค่า discrim ที่เป็นค่าลบ ให้พิจารณาตัดตัวลวงที่ติดลบค่าน้อย ๆ ออก
- ถ้าข้อสอบข้อใดมีตัวเลือก (ตัวลวง) ที่มีค่า discrim ที่เป็นค่าลบ และค่าบวก ให้พิจารณาตัดตัวลวงที่มีค่าบวกออก

หมายเหตุ : ในโปรแกรม R จะใช้สูตรในการคำนวณค่าอำนาจจำแนก ดังนี้

distractorAnalysis {CTT}	R Documentation
Function for item distractor analysis	
Description	
This function provides a distractor analysis.	
Usage	
<code>distractorAnalysis(items, key, scores, nGroups=4, defineGroups, multiKeySep="none", multiKeyScore=c("or","poly"), validResp, csvReport, pTable=TRUE, digits)</code>	
Arguments	
items	The unscored item response from a multiple-choice test
key	The answer key for the items
scores	An optional set of person scores associated with the item data. If scores are not provided (default) the scores are calculated using the item data and key.
nGroups	Determines the number of groups into which scores are discretized. For example, nGroups=4 (default) performs and analysis based on quartiles.
defineGroups	If provided, determines the quantile breakpoints for groups into which scores are discretized. For example, defineGroups=c(.27,.46,.27) performs and analysis with 3 quantiles and 27 percent of examinees in the top and the bottom groups.
multiKeySep	If a value other than "none" is provided (e.g., ", "), the key and the raw items will be reviewed for the provided delimiter. Using this option allows for multiple correct responses.
multiKeyScore	The first value controls how multiple keys are handled. If "or" any correct response results in score of 1. If "and" all responses must be correct. If the second value is "poly" the returned score is the sum of correct responses. If the second value is "dich" a maximum score of 1 is returned. If the respondent can only provide one response, use "or". If the respondent can provide multiple responses and you use c("and", "poly") the score will be 0 or max score.
validResp	A list of vectors providing valid responses for the distractor tables. If no value is provided, valid responses are determined from the data and assumed to be the same across items. If "fromItem" is provided, values are determined from item responses and NOT assumed to be the same across items.
csvReport	If an optional file name is provided the function will save a .csv file with the results.
pTable	If pTable=FALSE the function returns the counts of examinees who provide each answer. If pTable=TRUE (default) the function returns the proportion of examinees who provide each answer.
digits	If digits (an integer) is provided, it specifies the number of decimals to which results will be rounded.

Details

The scores are used to split respondents into groups, with number determined by nGroups. The proportion (or number if pTable=FALSE) of examinees in each group giving each response is reported. The correct answer is indicated with an "**". Additional item statistics are provided. Descriptors of each item are returned as separate elements in a list.

Value

correct	An "*" indicates the correct response
key	The response option being described
n	The number of respondents choosing that option
rspP	The proportion of respondents with that response
pBis	The point-biserial correlation between that response and the total score with that item removed
discrim	The upper proportion minus the lower proportion
lower	The proportion of respondents choosing that response that are from the lowest score group
upper	The proportion of respondents choosing that response that are from the highest score group

Author(s)

John T. Willse

References

Allen, M. J. & Yen, W. M. (1979). Introduction to Measurement Theory. Lon Grove, Illinois: Waveland Press, INC.

Examples

```
# Example data provided with package
data(CTTdata)
data(CTTkey)

distractorAnalysis(CTTdata, CTTkey)

# Results provided in a .csv file.
distractorAnalysis(CTTdata, CTTkey, csvReport="Hello.csv")
```

[Package CTT version 2.3.3 [Index](#)]

อ้างอิง : <https://search.r-project.org/CRAN/refmans/CTT/html/distractorAnalysis.html>

เกณฑ์การพิจารณาตัดตัวเลือกจากโค้งคุณลักษณะตัวเลือก (OCC) ดูจากเส้นกราฟที่แสดงความน่าจะเป็นในการที่ผู้สอบจะเลือกตอบตัวเลือกนั้นมีเส้นกราฟอยู่ใกล้หรือประชิดแกนนอนมากที่สุด แสดงว่าความน่าจะเป็นที่ผู้สอบจะเลือกตอบตัวเลือกที่เป็นตัวลวงนั้นค่อนข้างน้อย หรือพิจารณาจากการไม่ปรากฏเส้นกราฟของตัวลวงนั้น แสดงว่าความน่าจะเป็นที่ผู้สอบจะเลือกตอบตัวลวงนั้นน้อยจนแทบจะไม่มีผู้สอบคนใดเลือกตัวลวงนั้นเลย สะท้อนให้เห็นว่าตัวลวงนั้นไม่มีประสิทธิภาพ เพราะลวงให้คนที่มีความสามารถต่ำมาตอบได้น้อย

ตัวอย่างการตัดตัวเลือกของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก เป็นแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก

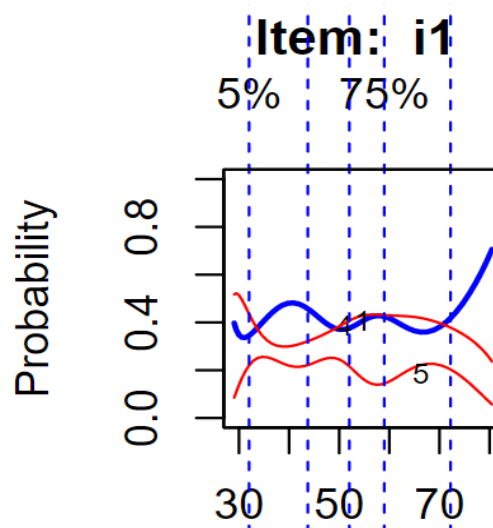
ข้อสอบข้อที่ 1 มีผลการวิเคราะห์ข้อมูล ดังนี้

ส่วนแรก คือ ผลการหาค่าอำนาจจำแนก

ตารางที่ 3.1 ค่าอำนาจจำแนกที่ได้จากการตัดตัวเลือกของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก เป็นแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก

ข้อที่	i1.	i1.	i1.	i1.	i1.	i1.	i1.	i1.	i1.	i1.
1	correct	key	n	rspP	pBis	discrim	lower	mid50	mid75	upper
1		1	902	0.46	-0.04	0.01	0.44	0.47	0.49	0.45
2		2	34	0.02	-0.08	-0.02	0.02	0.02	0.02	0.00
3		3	23	0.01	-0.07	-0.01	0.02	0.01	0.00	0.01
4	*	4	716	0.36	0.01	0.05	0.36	0.33	0.35	0.42
5		5	291	0.15	-0.07	-0.03	0.16	0.17	0.13	0.13

ส่วนที่ 2 คือ โค้งคุณลักษณะตัวเลือก (OCC)



ภาพที่ 3.1 โค้งคุณลักษณะตัวเลือก (OCC) ของข้อสอบข้อที่ 1

จากตารางผลการวิเคราะห์ค่าอำนาจจำแนกและโค้งคุณลักษณะตัวเลือก OCC โดยยกเว้นการตัดตัวเลือกที่ 4 ซึ่งเป็นตัวเลือกที่เป็นเฉลย เมื่อพิจารณาจากค่า discrim พบว่า ตัวเลือกที่อยู่ในขอบข่ายที่สามารถตัดออกได้ คือ ตัวเลือกที่ 1, ตัวเลือกที่ 2, ตัวเลือกที่ 3 และตัวเลือกที่ 5 ถ้าพิจารณาตัดตัวเลือกโดยอาศัยเกณฑ์การพิจารณาตัดตัวเลือกจากค่าอำนาจจำแนกเกณฑ์เดียว จะต้องตัดตัวเลือกที่ 1 ออก

แต่ถ้าพิจารณาให้ลึกซึ้ง จะพบว่าตัวเลือกนี้ลวงคนให้มาเลือกตอบได้ถึง 902 คน หากพิจารณาจากโค้ง OCC ร่วมด้วย พบว่ามีคนจำนวนหนึ่งซึ่งยังคงเลือกตอบในตัวเลือกที่ 1 ดังจะเห็นได้จากกราฟเส้นสีแดงที่กำกับด้วยตัวเลข 1 โดยผู้เข้าสอบในระดับคะแนนที่คาดหมายต่าง ๆ ยังคงเลือกตอบตัวเลือกที่ 1 ด้วยความน่าจะเป็นในการเลือกตอบในตัวเลือกที่ 1 อยู่ระหว่าง 0.3 – 0.5 ซึ่งใกล้เคียงกับความน่าจะเป็นที่ผู้สอบที่มีคะแนนที่คาดหมายในช่วง 0 – 70 คะแนน จะเลือกตอบในตัวเลือกที่ 4 ซึ่งเป็นตัวเลือกที่เป็นเฉลย ดังนั้นผู้วิจัยจึงไม่เลือกตัดตัวเลือกที่ 1 นี้ออก เพราะยังถือว่าเป็นตัวลวงที่มีประสิทธิภาพ สามารถลวงคนที่มีความสามารถต่ำ หรือผู้สอบที่มีคะแนนที่คาดหมายตั้งแต่ 0 – 60 คะแนน มาตอบได้เป็นจำนวนมาก แต่เมื่อพิจารณาจากโค้ง OCC พบว่า ไม่ปรากฏเส้นกราฟแสดงโอกาสในการที่ผู้มีความสามารถในระดับต่าง ๆ เลือกตอบในตัวเลือกที่ 2 และตัวเลือกที่ 3 เลย ดังนั้นผู้วิจัยจึงสนใจที่จะทำการตัดตัวเลือกที่ 2 หรือตัวเลือกที่ 3 ออก เมื่อนำไปพิจารณาร่วมกับค่า discrim พบว่าตัวเลือกที่ 3 มีค่า discrim ตีลบค่าน้อยกว่า ค่า discrim ของตัวเลือกที่ 2 สอดคล้องตามเกณฑ์การพิจารณาตัดตัวเลือกจากค่าอำนาจจำแนก ดังนั้นข้อสอบข้อที่ 1 จึงต้องทำการตัดตัวเลือกที่ 3 ออก

9) ผู้วิจัยทำการแทนที่ตัวเลือกที่ถูกตัดออกในแต่ละข้อ ของผู้สอบแต่ละคนจะครบทุกคน โดยผู้วิจัยจะแทนที่ตัวเลือกที่ถูกตัดออกเฉพาะผู้สอบที่ถูกตัดตัวเลือกเท่านั้น สำหรับผู้เข้าสอบที่ไม่ถูกตัดตัวเลือก ผู้วิจัยจะถือว่าคำตอบที่ผู้เข้าสอบตอบในแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก เป็นคำตอบเดียวกันกับแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก สำหรับวิธีการแทนที่ตัวเลือกที่ถูกตัดออก จะใช้การแทนที่ตัวเลือกตามข้อตกลงเบื้องต้นที่ผู้วิจัยได้ตั้งไว้ 3 ข้อ ได้แก่ ข้อตกลงเบื้องต้นข้อที่ 1 ผู้สอบที่มีความสามารถสูงจะต้องตอบตัวเลือกที่เป็นเฉลยตั้งแต่ครั้งแรกที่ลงมือทำข้อสอบ ข้อตกลงเบื้องต้นข้อที่ 2 ผู้สอบทุกคนมีสิทธิ์ที่จะเลือกตอบตัวเลือกใด ๆ ก็ได้ และข้อตกลงเบื้องต้นข้อที่ 3 ผู้สอบที่มีความสามารถสูงจะมีโอกาสสูงในการทำข้อสอบข้อนั้น ๆ ถูกต้อง และผู้สอบที่มีความสามารถต่ำ ก็จะมีโอกาสน้อยที่จะตอบข้อสอบข้อนั้นได้ถูกต้อง

เมื่อผู้วิจัยกำหนดข้อตกลงเบื้องต้นที่แตกต่างกัน จึงส่งผลทำให้การแทนที่ตัวเลือกที่ถูกตัดออกก็ต้องสอดคล้องกับข้อตกลงเบื้องต้นแต่ละข้อ และใช้หลักการแทนที่ที่แตกต่างกันออกไปด้วย สำหรับข้อตกลงเบื้องต้นข้อที่ 1 จะใช้วิธีการแทนค่า กรณีที่ 1 แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย ข้อตกลงเบื้องต้นข้อที่ 2 จะใช้วิธีการแทนค่า กรณีที่ 2 แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย และข้อตกลงเบื้องต้นข้อที่ 3 จะใช้วิธีการแทนค่า กรณีที่ 3 แทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ เพื่อแยกแยะการทำความเข้าใจผู้วิจัยจำแนกข้อตกลงเบื้องต้นและวิธีการแทนค่าดังตารางที่ 3.2

ตารางที่ 3.2 ข้อตกลงเบื้องต้นและวิธีการแทนค่าคำตอบที่ผู้วิจัยกำหนด

ข้อตกลงเบื้องต้น	วิธีการแทนค่า
ข้อตกลงเบื้องต้นข้อที่ 1 : ผู้สอบที่มีความสามารถสูงจะต้องตอบตัวเลือกที่เป็นเฉลยตั้งแต่ครั้งแรกที่ลงมือทำข้อสอบ	กรณีที่ 1 : แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่มด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย
ข้อตกลงเบื้องต้นข้อที่ 2 : ผู้สอบทุกคนมีสิทธิ์ที่จะเลือกตอบตัวเลือกใด ๆ ก็ได้	กรณีที่ 2 : แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่มด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย
ข้อตกลงเบื้องต้นข้อที่ 3 : ผู้สอบที่มีความสามารถสูงจะมีโอกาสสูงในการทำข้อสอบข้อนั้น ๆ ถูกต้อง และผู้สอบที่มีความสามารถต่ำ ก็จะมีโอกาสน้อยที่จะตอบข้อสอบข้อนั้นได้ถูกต้อง	กรณีที่ 3 : แทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ

ผู้วิจัยขออธิบายวิธีการแทนค่าตัวเลือกที่ถูกตัดออก ด้วยตัวอย่างการตัดตัวเลือกออกของข้อสอบข้อที่ 1 ข้างต้น โดยจะอธิบายแยกเป็นกรณี ๆ ดังรายละเอียดต่อไปนี้

ตารางที่ 3.3 วิธีการแทนค่าคำตอบที่ผู้วิจัยกำหนด

ข้อที่	i1. correc t	i1. key	i1. n	i1. rspP	i1. pBis	i1. discri m	i1. lower	i1. mid50	i1. mid75	i1. upper
1		1	902	0.46	-0.04	0.01	0.44	0.47	0.49	0.45
2		2	34	0.02	-0.08	-0.02	0.02	0.02	0.02	0.00
3		3	23	0.01	-0.07	-0.01	0.02	0.01	0.00	0.01
4	*	4	716	0.36	0.01	0.05	0.36	0.33	0.35	0.42
5		5	291	0.15	-0.07	-0.03	0.16	0.17	0.13	0.13
6		6	0	0.00	NA	0.00	0.00	0.00	0.00	0.00
7		7	0	0.00	NA	0.00	0.00	0.00	0.00	0.00

จากตารางแสดงผลการวิเคราะห์ข้อสอบข้อที่ 1 พบว่าในข้อนี้ผู้วิจัยเลือกตัดตัวเลือกที่ 3 ออก และเมื่อตัดตัวเลือกครบทั้ง 120 ข้อ จะทำให้ได้แบบทดสอบฉบับใหม่ ที่เหลือตัวเลือกเพียง 4 ตัวเลือก เนื่องจากเป็นตัวเลือก (ตัวลวง) ที่ไม่มีประสิทธิภาพ จากข้อมูลในตารางพบว่าผู้สอบจำนวน 23 คน ที่ตอบตัวเลือกที่ 3 ดังนั้นผู้วิจัยจึงต้องใช้วิธีการแทนค่าที่สอดคล้องกับข้อตกลงเบื้องต้นที่ผู้วิจัยกำหนดขึ้น เพื่อหาตัวเลือกตัวใหม่แทนที่ตัวเลือกตัวเดิมที่ถูกตัดออก โดยมีวิธีการแทนค่าที่แตกต่างกัน 3 กรณี ดังตัวอย่างต่อไปนี้

กรณีที่ 1 : แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย

สำหรับกรณีนี้ผู้วิจัยเลือกใช้ตัวเลือกใดก็ได้ ที่ไม่ใช่ตัวเลือกที่ถูกตัดออก คือ ตัวเลือกที่ 3 และตัวเลือกที่เป็นเฉลย คือ ตัวเลือกที่ 4 ในการแทนที่ตัวเลือก 3 ที่ถูกตัดออก ดังนั้นตัวเลือกที่สามารถถูกสุ่มเข้าไปแทนที่ตัวเลือกที่ถูกตัดออก คือ ตัวเลือกที่ 1, ตัวเลือกที่ 2 และตัวเลือกที่ 5

กรณีที่ 2 : แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย (Guo, Zu and Kyllonen, 2018)

สำหรับกรณีนี้ผู้วิจัยเลือกใช้ตัวเลือกใดก็ได้ ที่ไม่ใช่ตัวเลือกที่เป็นเฉลย คือ ตัวเลือกที่ 4 ในการแทนที่ตัวเลือก 3 ที่ถูกตัดออก ดังนั้นตัวเลือกที่สามารถถูกสุ่มเข้าไปแทนที่ตัวเลือกที่ถูกตัดออก คือ ตัวเลือกที่ 1, ตัวเลือกที่ 2, ตัวเลือกที่ 4 และตัวเลือกที่ 5

กรณีที่ 3 : แทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ (Guo, Zu and Kyllonen, 2018)

สำหรับวิธีการแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบนี้ ก่อนทำการแทนค่าตัวเลือกที่ถูกตัดออก ผู้วิจัยจะต้องทำการจัดแบ่งกลุ่มนักศึกษาผู้เข้าสอบออกเป็น 2 กลุ่ม ตามระดับความสามารถของผู้สอบ ผู้วิจัยกำหนดเกณฑ์ที่ใช้ในการแบ่งกลุ่ม คือ หากผู้สอบได้คะแนนรวมจากการสอบตั้งแต่ 60 คะแนนขึ้นไป จากคะแนนเต็ม 120 คะแนน จะถูกจัดอยู่ในกลุ่มผู้สอบที่มีความสามารถอยู่ในระดับสูง แต่ถ้าหากผู้สอบได้คะแนนรวมจากการสอบต่ำกว่า 60 คะแนนลงมา จากคะแนนเต็ม 120 คะแนน จะถูกจัดอยู่ในกลุ่มผู้สอบที่มีความสามารถต่ำ

ลำดับต่อไปจะเป็นการแทนที่ตัวเลือกที่ถูกตัดออกจากจำแนกตามกลุ่มที่แสดงระดับความสามารถของผู้สอบ คือ กลุ่มผู้สอบที่มีความสามารถอยู่ในระดับสูง จะแทนด้วยตัวเลือกที่เป็นเฉลย คือ ตัวเลือกที่ 4 สำหรับกลุ่มผู้สอบที่มีความสามารถอยู่ในระดับต่ำ จะถูกสุ่มแทนที่ตัวเลือกที่ถูกตัดออกด้วยตัวเลือกตัวอื่น ๆ ที่ไม่ใช่ตัวเลือกที่เป็นเฉลย คือ ตัวเลือกที่ 4 และตัวเลือก (ตัวลวง) ที่ไม่มีประสิทธิภาพ คือ ตัวเลือกที่ 3 นั่นคือ ตัวเลือกที่สามารถแทนที่ตัวเลือกที่ถูกตัดออก คือ ตัวเลือกที่ 1, ตัวเลือกที่ 2 และตัวเลือกที่ 3

สำหรับการแทนที่ตัวเลือกที่ถูกตัดออกจะต้องทำให้ครบทุกข้อ ทั้ง 120 ข้อ เพื่อจะทำให้ได้ข้อมูลที่ได้จากการจำลองสถานการณ์การตอบข้อสอบแบบเลือกตอบ 4 ตัวเลือก โดยจะต้องทำให้ครบตามจำนวนของผู้สอบที่ถูกตัดตัวเลือกออก การแทนที่ตัวเลือกที่ถูกตัดออก ผู้วิจัยใช้การเขียนโปรแกรมภาษา Go

เมื่อแทนที่ตัวเลือกที่ถูกตัดออกครบทุกกรณีแล้ว ผู้วิจัยจะได้ข้อมูลทั้งหมด 4 ชุด ได้แก่ ข้อมูลชุดที่ 1 เป็นข้อมูลที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ข้อมูลชุดที่ 2 เป็นข้อมูลที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (ข้อมูลที่ได้ในชุดนี้จะมีการแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) ข้อมูลชุดที่ 3 เป็นข้อมูลที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (ข้อมูลที่ได้ในชุดนี้จะมีการแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) และข้อมูลชุดที่ 4 เป็นข้อมูลที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (ข้อมูลที่ได้ในชุดนี้แทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ)

10) ผู้วิจัยทำการวิเคราะห์คุณภาพของแบบทดสอบ โดยทำการวิเคราะห์คุณภาพของแบบทดสอบที่ได้จากข้อมูล 4 ชุด ข้างต้น โดยใช้โปรแกรม R แพคเกจ psych ในการคำนวณหาค่าความเที่ยงของแบบทดสอบ โดยใช้เกณฑ์ในการพิจารณาความเที่ยงของแบบทดสอบคือ แบบทดสอบที่มีความเที่ยงในการวัดจะต้องสูงกว่าเกณฑ์ขั้นต่ำที่กำหนด คือ 0.50 (ศิริชัย กาญจนวาสี, 2556)

11) ผู้วิจัยทำการวิเคราะห์คุณภาพของแบบทดสอบ ด้านความยาก และอำนาจจำแนกของข้อสอบแต่ละข้อที่ได้จากแบบทดสอบแบบซึ่งจำแนกออกเป็น 4 ชุดข้อมูล โดยใช้แพคเกจ KernsmoothIRT ในโปรแกรม R การพิจารณาความยากและอำนาจจำแนก จะพิจารณาจากการวิเคราะห์องค์ประกอบหลัก (Principle component analysis : PCA) ของโค้งคุณลักษณะข้อสอบ (Item Characteristic Curves : ICCs)

หลักในการพิจารณาความยากและอำนาจจำแนกจากการวิเคราะห์องค์ประกอบหลักของโค้งคุณลักษณะข้อสอบ (Mcguire, 2012)

- ข้อสอบข้อใดที่ตกอยู่ในแนวแกนนอนทางด้านขวาสุดจะถือว่าเป็นข้อสอบที่ยาก
 - ข้อสอบข้อใดที่ตกอยู่ในแนวแกนนอนทางด้านซ้ายสุดจะถือว่าเป็นข้อสอบที่ง่าย
 - ข้อสอบข้อใดอยู่ในแนวเส้นประและอยู่กึ่งกลางกราฟ ข้อสอบข้อนั้นเป็นข้อสอบที่ยากพอเหมาะ
- การพิจารณาความยากของข้อสอบ สามารถสังเกตทิศทางการพิจารณาได้ โดยสังเกตจากโค้ง ICC หากโค้ง ICC มีเส้นกราฟที่ประชิดขอบด้านบนในทิศทางใด แสดงว่าข้อสอบข้อนั้นมุ่งไปในทางยาก และหากโค้ง ICC มีเส้นกราฟที่ประชิดขอบด้านล่างในทิศทางใด แสดงว่าข้อสอบข้อนั้นมุ่งไปในทางง่าย
- ข้อสอบข้อใดที่ตกอยู่ในแนวแกนตั้งทางด้านบนจะถือว่าเป็นข้อสอบที่จำแนกไม่ดี
 - ข้อสอบข้อใดที่ตกอยู่ในแนวแกนตั้งทางด้านล่างสุดจะถือว่าเป็นข้อสอบที่จำแนกดี
 - ข้อสอบข้อใดอยู่ในแนวเส้นประและอยู่กึ่งกลางกราฟ ข้อสอบข้อนั้นเป็นข้อสอบที่จำแนกใช้ได้

การพิจารณาอำนาจจำแนกของข้อสอบ สามารถสลับทิศทางการพิจารณาได้ โดยสังเกตจากโค้ง ICC หากโค้ง ICC มีเส้นกราฟที่มีความชันเป็นลบ (Negatively Discriminates) หรือเป็นฟังก์ชันลด แสดงว่าข้อสอบข้อนั้นมุ่งไปในทางจำแนกไม่ดี และหากโค้ง ICC มีเส้นกราฟที่มีความชันเป็นบวก (Positively Discriminates) หรือเป็นฟังก์ชันเพิ่ม แสดงว่าข้อสอบข้อนั้นมุ่งไปในทางจำแนกได้ดี

12) ผู้วิจัยทำการจำลองสถานการณ์การออกแบบการเก็บรวบรวมข้อมูลให้อยู่ในรูปแบบผู้สอบกลุ่มเดียวที่ไม่ได้รับการจัดสมดุลย์ (Single – Group Uncounterbalanced Design) เป็นวิธีการที่ใช้ผู้สอบกลุ่มเดียวทำหน้าที่เป็นผู้สอบร่วม กล่าวคือ ผู้สอบคนเดียวกันทำแบบทดสอบทั้ง 2 ฉบับ คือ ทำทั้งแบบทดสอบ 5 ตัวเลือกและแบบทดสอบ 4 ตัวเลือก หรือที่เรียกอีกชื่อหนึ่งว่า The Single Group Design (ศิริชัย กาญจนวาสี, 2550; Dorans, Moses and Eignor, 2010; Guo, Zu and Kyllonen, 2018) แล้วจึงนำข้อมูลมาทำการปรับเทียบคะแนนจากแบบทดสอบสองฉบับเป็นรายกรณีดังต่อไปนี้

1. การปรับเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย)

2. การปรับเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย)

3. การปรับเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ)

สำหรับงานวิจัยนี้จะใช้รูปแบบการปรับเทียบโดยใช้ทฤษฎีดั้งเดิม (Classical Model of Equating) เพราะเป็นรูปแบบที่มีข้อดีคือ เป็นการใช้การคำนวณทางคณิตศาสตร์เพื่อแสดงความสัมพันธ์และง่ายต่อการทำความเข้าใจ ประกอบกับเป็นวิธีที่สามารถใช้กลุ่มตัวอย่างขนาดเล็กได้เมื่อเปรียบเทียบกับทฤษฎีการตอบสนองข้อสอบ (Ryan and Brockmann, 2011) ในงานวิจัยนี้ใช้การปรับเทียบคะแนนเชิงเส้นตรง (Linear Equating) โดยใช้โปรแกรม Jamovi ช่วยในการวิเคราะห์ข้อมูล

13) ผู้วิจัยทำการศึกษาเปรียบเทียบคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก กับแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ในแต่ละกรณี ทั้ง 3 กรณี เพื่อทดสอบสมมติฐานการวิจัยของผู้วิจัยที่ว่า คะแนนสอบของนักศึกษาที่ได้จากการทำแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ในแต่ละกรณีสูงกว่าคะแนนสอบของนักศึกษาที่ได้จากการทำแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

โดยผู้วิจัยตั้งสมมติฐานการวิจัยที่ต้องทำการทดสอบ ดังนี้

สมมติฐานที่ 1 : คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

สมมติฐานที่ 2 : คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

สมมติฐานที่ 3 : คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

การทดสอบสมมติฐานเหล่านี้จะใช้วิธีการวิเคราะห์ทางสถิติ โดยใช้การพิจารณาจากค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และการทดสอบ Wilcoxon Signed Rank test ด้วยโปรแกรม Jamovi

14) ผู้วิจัยทำการประมาณการงบประมาณในการจัดทำแบบทดสอบ เพื่อเปรียบเทียบค่าใช้จ่ายในการจัดทำแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ว่าค่าใช้จ่ายในการจัดทำแบบทดสอบแบบเลือกตอบแบบ 4 เลือกจะลดค่าใช้จ่ายลงได้หรือไม่ ถ้าลดค่าใช้จ่ายลงได้ ค่าใช้จ่ายโดยการประมาณการจะลดลงเป็นจำนวนเท่าใด

ข้อตกลงที่ใช้ในการวิจัยครั้งนี้เกี่ยวกับการวิเคราะห์ค่าใช้จ่าย ผู้วิจัยจะทำการคำนวณค่าใช้จ่ายในการจัดทำแบบทดสอบเฉพาะค่าใช้จ่ายที่เป็นตัวเงินที่เกิดขึ้นจากกระบวนการทำงานของสำนักทะเบียนและวัดผลเท่านั้น ไม่ได้คิดค่าใช้จ่ายที่เกี่ยวข้องกับสำนักพิมพ์และไม่คิดค่าใช้จ่ายสำหรับอาจารย์ที่ออกข้อสอบ ทั้งนี้ค่าใช้จ่ายที่ไม่เป็นตัวเงินหรือปัจจัยอื่นที่ไม่ใช่ตัวเงิน เช่น เวลาที่ใช้ในการตรวจร่างแบบทดสอบ ความเหนื่อยล้าในการจัดทำแบบทดสอบ จะไม่ถูกนำมาคำนวณในส่วน of ค่าใช้จ่ายในงานวิจัยนี้

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในการวิจัยนี้เป็นการวิเคราะห์ข้อมูลเชิงปริมาณ ได้แก่ ค่าเฉลี่ย, ส่วนเบี่ยงเบนมาตรฐาน, การทดสอบ Wilcoxon Signed Rank test, การวิเคราะห์แบบทดสอบเพื่อหาความเที่ยงของแบบทดสอบ โดยใช้ทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory), การวิเคราะห์ข้อสอบเพื่อหาค่าความยากและอำนาจจำแนกของข้อสอบ โดยใช้ทฤษฎีการตอบสนองข้อสอบด้วยการทำให้เรียบเคอร์เนล (Kernel Smoothing in Item Response Theory) และการปรับเทียบคะแนนสอบโดยใช้การปรับเทียบคะแนนเชิงเส้น (Linear Equating) ตามรายละเอียดดังนี้

1. การวิเคราะห์คุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก, แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย), แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) เพื่อหาความเที่ยงของแบบทดสอบ โดยใช้ทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory)

2. การวิเคราะห์ข้อสอบเพื่อหาค่าความยากและอำนาจจำแนกของข้อสอบแต่ละข้อของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก, แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูก

ตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย), แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) โดยใช้ทฤษฎีการตอบสนองข้อสอบด้วยการทำให้เรียบเคอร์เนล (Kernel Smoothing in Item Response Theory)

3. การเปรียบเทียบคะแนนจากแบบทดสอบสองฉบับเป็นรายกรณีดังต่อไปนี้ 1. การเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) 2. การเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) 3. การเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) สำหรับงานวิจัยนี้จะใช้รูปแบบการเปรียบเทียบโดยใช้ทฤษฎีดั้งเดิม (Classical Model of Equating) ด้วยการปรับเทียบคะแนนเชิงเส้น (Linear Equating)

4. การทดสอบสมมติฐานที่ 1 : คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก โดยใช้ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน

5. การทดสอบสมมติฐานที่ 2 : คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และสมมติฐานที่ 3 : คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ด้วยการทดสอบ Wilcoxon Signed Rank test

บทที่ 4

ผลการวิจัย

การนำเสนอผลการวิเคราะห์ข้อมูลเกี่ยวกับการเปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ผู้วิจัยนำเสนอผลการวิเคราะห์ออกเป็น 3 ตอน ได้แก่ ตอนที่ 1 ผลการวิเคราะห์เปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ตอนที่ 2 ผลการวิเคราะห์การปรับเทียบคะแนนและเปรียบเทียบคะแนนของนักศึกษาที่ได้จากการปรับเทียบคะแนนจากคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก และตอนที่ 3 ผลการวิเคราะห์เปรียบเทียบค่าใช้จ่ายในการจัดทำแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก

ตอนที่ 1 ผลการวิเคราะห์เปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก

ผลการวิเคราะห์ในตอนที่ 1 เป็นการนำเสนอผลการวิเคราะห์เปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก เพื่อตอบวัตถุประสงค์การวิจัยข้อที่ 1 โดยผู้วิจัยแบ่งการนำเสนอออกเป็น 4 ส่วนสำคัญ ได้แก่ ผลการวิเคราะห์ข้อมูลกรณีที่ 1 แบบทดสอบแบบเลือกตอบ 5 ตัวเลือก กรณีที่ 2 แบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) กรณีที่ 3 แบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) และกรณีที่ 4 แบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) มีรายละเอียดดังนี้

1.1 ผลการวิเคราะห์เปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก กรณีที่ 1 แบบทดสอบแบบ 5 ตัวเลือก

ผลการวิเคราะห์ความเที่ยง หาโดยใช้แพ็คเกจ psych จะให้ค่าความเที่ยงที่คำนวณโดยใช้สัมประสิทธิ์แอลฟาของครอนบาค

```

Reliability analysis
Call: alpha(x = score$scored, check.keys = FALSE)

raw_alpha std.alpha G6(smc) average_r S/N ase mean sd median_r
0.81      0.8      0.82      0.033   4 0.006 0.43 0.092      0.03

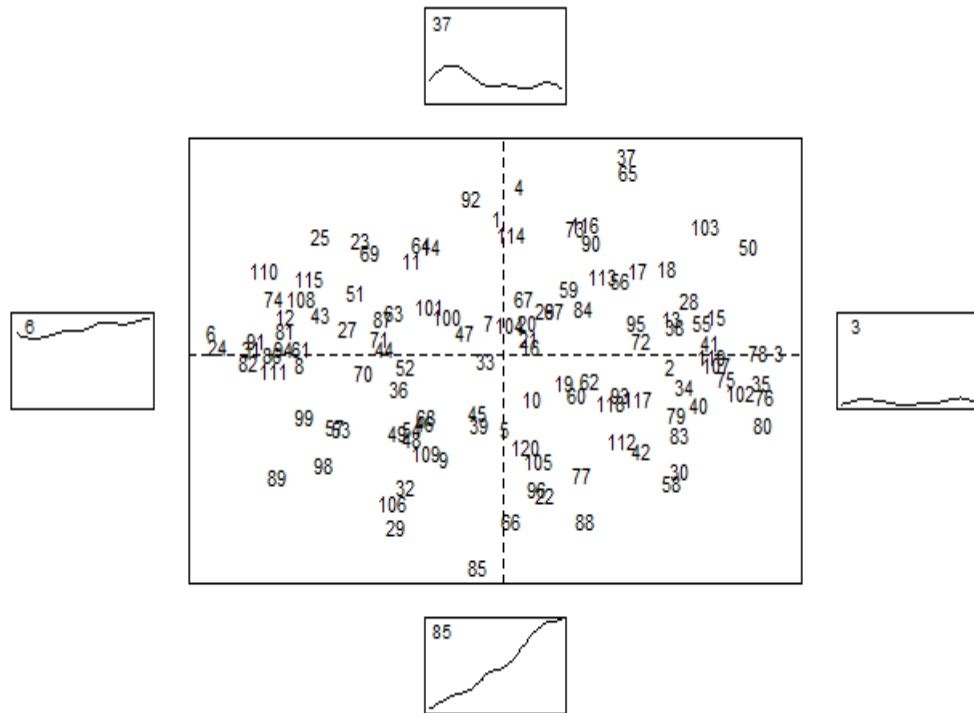
95% confidence boundaries
      lower alpha upper
Feldt    0.8  0.81  0.82
Duhachek 0.8  0.81  0.82

```

ภาพที่ 4.1 ผลการวิเคราะห์ค่าความเที่ยง กรณีที่ 1 แบบทดสอบแบบ 5 ตัวเลือก

จากการวิเคราะห์ความเที่ยงของแบบทดสอบ พบว่าค่าความเที่ยง (raw_alpha) ของแบบทดสอบแบบ 5 ตัวเลือกมีค่าเท่ากับ 0.81 ซึ่งพบว่าแบบทดสอบแบบ 5 ตัวเลือกที่ถูกเลือกมาใช้เป็นข้อมูลตั้งต้นในการวิเคราะห์มีค่าความเที่ยงค่อนข้างสูง

ความยากและอำนาจจำแนก จะพิจารณาการวิเคราะห์องค์ประกอบหลัก (Principle component analysis : PCA) ของโค้งคุณลักษณะข้อสอบ (Item Characteristic Curves : ICCs) ซึ่งได้มาจากแพ็คเกจ KernsmoothIRT ในโปรแกรม R โดย Kernel smoothing in IRT เป็นวิธีการประมาณโค้งคุณลักษณะตัวเลือก (Option Characteristic Curves : OCC) ตามแนวทางการวิเคราะห์นอนพาราเมตริก IRT (Nonparametric IRT) ซึ่งจะใช้วิเคราะห์ตัวแบบ IRT ในกรณีที่ตัวแบบไม่เป็นไปตามข้อตกลงเบื้องต้นของการวิเคราะห์ตัวแบบตามแนวทางพาราเมตริก (Parametric IRT) จากข้อมูลการตอบแบบทดสอบแบบ 5 ตัวเลือกที่ถูกเลือกมาเป็นข้อมูลตั้งต้นจากงานวิจัยนี้เป็นข้อมูลทุติยภูมิ ทางผู้วิจัยไม่ได้ทำการออกข้อสอบเอง ดังนั้นผู้วิจัยจึงไม่สามารถบอกได้แน่ชัดว่าข้อมูลที่ได้จากแบบทดสอบแบบ 5 ตัวเลือกนี้ เป็นไปตามข้อตกลงเบื้องต้นของ Parametric IRT หรือไม่ ดังนั้นผู้วิจัยจึงขอเลี่ยงไปใช้ Nonparametric IRT อีกประการหนึ่งที่ผู้วิจัยเลือกใช้วิธีการนี้เนื่องจากวิธีการดังกล่าวสามารถใช้ได้กับผู้สอบที่มีจำนวนไม่มากนัก



ภาพที่ 4.2 ผลการวิเคราะห์อำนาจจำแนกและความยากด้วย PCA
กรณีที่ 1 แบบทดสอบแบบ 5 ตัวเลือก

จาก PCA พบว่าข้อสอบทั้งหมดจำนวน 120 ข้อ มีข้อสอบข้อ 85 ที่สามารถจำแนกได้ดี เช่นเดียวกับข้อ 29, 66, 88, 106, 32, 22 และ 96 และข้ออื่น ๆ ที่อยู่ใกล้แนวเส้นประ ในแกนตั้ง ซึ่งอยู่ได้แนวเส้นประ ในแกนนอน แต่เมื่ออยู่เหนือแนวเส้นประ ในแกนนอนจะเป็นข้อสอบที่จำแนกได้ไม่ค่อยดีนัก เช่น ข้อ 67, 114 และ 1 เป็นต้น โดยเฉพาะข้อ 4, 92 37 และ 65 ถือว่าเป็นข้อสอบที่จำแนกได้ไม่ดี หากพิจารณาข้อสอบข้อ 103, 50, 3, 78, 35, 76, 80, 102, 75, 107, 110, 41 และ 15 เป็นต้น ถือว่าเป็นข้อสอบที่ค่อนข้างยาก โดยเฉพาะข้อที่อยู่ทางด้านปลายเส้นประ แกนนอน ด้านขวาของเส้นประ แกนตั้ง จะยิ่งสะท้อนว่าข้อสอบข้อนี้ยากมาก เช่น ข้อ 78, 3, 35, 76 และ 80 เป็นต้น สำหรับข้อที่อยู่ทางด้านเส้นประ แกนนอน ที่อยู่ทางด้านซ้ายของเส้นประ แกนตั้ง เช่น ข้อ 6, 24, 82, 31, 91, 88, 110, 74 และ 89 ถือเป็นข้อสอบที่ค่อนข้างง่าย โดยเฉพาะข้อ 6 ถือเป็นข้อสอบข้อที่ง่ายมาก สำหรับข้อสอบข้ออื่น ๆ ถือว่าอยู่ในเกณฑ์ใช้ได้

1.2 ผลการวิเคราะห์เปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก กรณีที่ 2 แบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย)

ค่าความเที่ยงของแบบทดสอบแบบ 4 ตัวเลือก ที่เกิดจากการตัดตัวเลือกที่จำแนกได้ไม่ได้ออก โดยไม่มีการตัดตัวเลือกที่เป็นเฉลยออก แล้วใช้การพิจารณาตัดตัวเลือกที่เหลือโดยดูจากค่า discrim ที่ได้จากแพ็คเกจ CTT โดยใช้เกณฑ์การพิจารณาค่า discrim คือ ถ้าข้อใดมีตัวเลือก (ตัวลวง) ที่มีเฉพาะค่า discrim ที่เป็นค่าบวก ให้พิจารณาตัดตัวลวงที่มีค่าบวกมากที่สุดออก ถ้าข้อใดมีตัวเลือก (ตัวลวง) ที่มีเฉพาะค่า discrim ที่เป็นค่าลบ ให้พิจารณาตัดตัวลวงที่ติดลบค่าน้อย ๆ ออก แต่ถ้าข้อใดมีตัวเลือก (ตัวลวง) ที่มีค่า discrim ที่เป็นค่าลบ และค่าบวก ให้พิจารณาตัดตัวลวงที่มีค่าบวกออก พร้อมกับพิจารณาจากกราฟ OCC โดยเกณฑ์การพิจารณาตัดตัวเลือกจากกราฟ OCC ดูจากเส้นกราฟที่แสดงความน่าจะเป็นในการที่ผู้สอบจะเลือกตอบตัวเลือกนั้นมีเส้นกราฟอยู่ใกล้หรือประชิดแกนนอนมากที่สุด แสดงว่าความน่าจะเป็นที่ผู้สอบจะเลือกตอบตัวเลือกที่เป็นตัวลวงนั้นค่อนข้างน้อย หรือพิจารณาจากการไม่ปรากฏเส้นกราฟของตัวลวงนั้น แสดงว่าความน่าจะเป็นที่ผู้สอบจะเลือกตอบตัวลวงนั้นน้อยจนแทบจะไม่มีผู้สอบคนใดเลือกตัวลวงนั้นเลย สะท้อนให้เห็นว่าตัวลวงนั้นไม่มีประสิทธิภาพ เพราะลวงให้คนที่มีความสามารถต่ำมาตอบได้น้อย จากนั้นจึงทำการแทนที่ตัวลวงที่หายไปอย่างสุ่ม ด้วยตัวเลือกตัวอื่น ๆ ที่เหลืออยู่แต่จะไม่แทนที่ตัวลวงที่ถูกตัดออกไปนั้นด้วยตัวเลือกที่เป็นเฉลย เพราะผู้วิจัยมีข้อตกลงเบื้องต้นว่า หากผู้สอบคนนั้นมีความสามารถสูงจะต้องตอบตัวเลือกที่เป็นเฉลยตั้งแต่ครั้งแรกที่ลงมือทำข้อสอบ

```
Reliability analysis
Call: alpha(x = Score$scored, check.keys = FALSE)

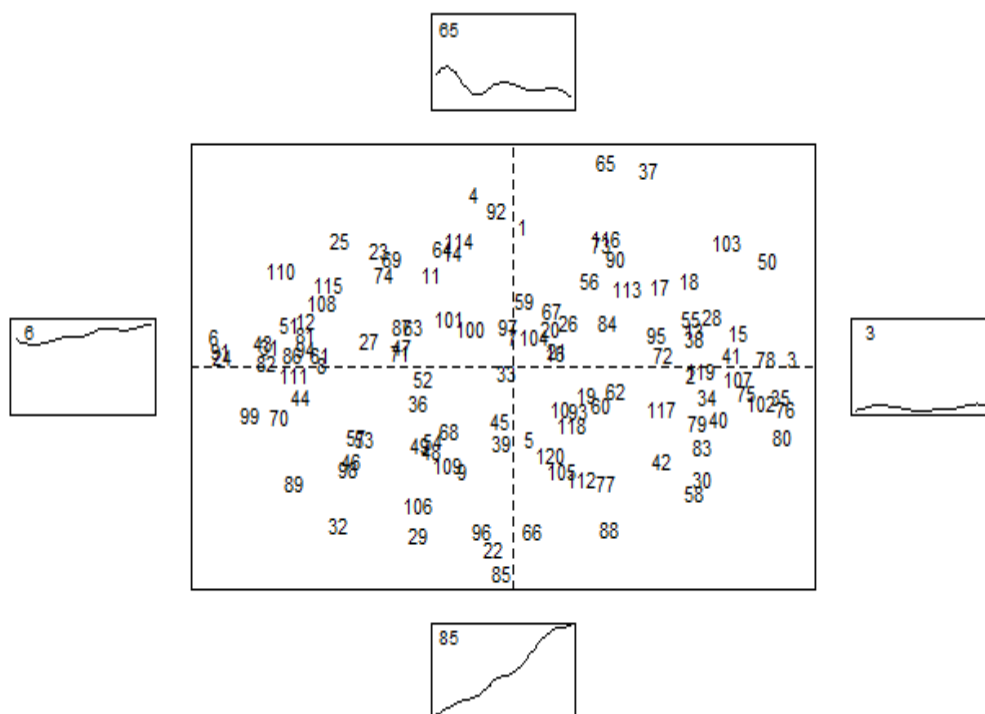
raw_alpha std.alpha G6(smc) average_r S/N ase mean sd median_r
0.79      0.78      0.8      0.029 3.6 0.0066 0.45 0.089 0.027

95% confidence boundaries
      lower alpha upper
Feldt  0.78  0.79  0.8
Duhachek 0.78  0.79  0.8
```

ภาพที่ 4.3 ผลการวิเคราะห์ค่าความเที่ยง กรณีที่ 2 แบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย)

จากการวิเคราะห์ความเที่ยงของแบบทดสอบ พบว่าค่าความเที่ยง (raw_alpha) ของแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) มีค่าเท่ากับ 0.79 ซึ่งมีค่าค่อนข้างสูง และมีค่าไม่แตกต่างจากความเที่ยงของแบบทดสอบแบบ 5 ตัวเลือกมากนัก โดยค่าความเที่ยงที่ได้มีค่าลดลงหลังจากการตัดตัวเลือกออกให้เหลือเพียง 4 ตัวเลือก

ความยากและอำนาจจำแนก จะพิจารณาการวิเคราะห์องค์ประกอบหลัก (Principle component analysis : PCA)



ภาพที่ 4.4 ผลการวิเคราะห์อำนาจจำแนกและความยากด้วย PCA กรณีที่ 2 แบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย)

จาก PCA พบว่าข้อสอบทั้งหมดจำนวน 120 ข้อ มีข้อสอบข้อ 85 ที่สามารถจำแนกได้ดี เช่นเดียวกับข้อ 29, 66, 88, 106, 32, 22 และ 96 และข้ออื่น ๆ ที่อยู่ใกล้แนวเส้นประ ในแกนตั้ง ซึ่งอยู่ใต้แนวเส้นประ ในแกนนอน แต่เมื่ออยู่เหนือแนวเส้นประ ในแกนนอนจะเป็นข้อสอบที่จำแนกได้ไม่ค่อยดีนัก เช่น ข้อ 65, 37, 4, 92 และ 1 เป็นต้น โดยเฉพาะข้อ 4, 37 และ 65 ถือว่าเป็นข้อสอบที่จำแนกได้ไม่ดี หากพิจารณาข้อสอบข้อ 103, 50, 3, 78, 35, 76, 80, 102, 75, 107, 41 และ 15 เป็นต้น ถือว่าเป็นข้อสอบที่ค่อนข้างยาก โดยเฉพาะข้อที่อยู่ทางด้านปลายเส้นประ แกนนอน ด้านขวาของเส้นประ แกนตั้ง จะยิ่งสะท้อนว่าข้อสอบข้อนี้ยากมาก เช่น ข้อ 78, 3, 35, 76 และ 80 เป็นต้น สำหรับข้อที่อยู่ทางด้านเส้นประ แกนนอน ที่อยู่ทางด้านซ้ายของเส้นประ แกนตั้ง เช่น ข้อ 6, 91, 24, 99, 48 และ 82 ถือเป็นข้อสอบที่ค่อนข้างง่าย โดยเฉพาะข้อ 6 ถือเป็นข้อสอบข้อที่ง่ายมาก สำหรับข้อสอบข้ออื่น ๆ ถือว่าอยู่ในเกณฑ์ใช้ได้ เมื่อพิจารณาจากข้อค้นพบนี้แล้ว พบว่าข้อสอบส่วนใหญ่จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) มีคุณลักษณะด้านความยากง่ายและอำนาจจำแนกเป็นไปในทิศทางเดียวกัน แสดงว่าแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่า

ตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) มีความยากและอำนาจจำแนกไม่แตกต่างกัน

1.3 ผลการวิเคราะห์เปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก กรณีที่ 3 แบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย)

ค่าความเที่ยงของแบบทดสอบแบบ 4 ตัวเลือก ที่เกิดจากการตัดตัวเลือกที่จำแนกได้ไม่ได้ออก โดยไม่มีการตัดตัวเลือกที่เป็นเฉลยออก แล้วใช้การพิจารณาตัดตัวเลือกที่เหลือโดยดูจากค่า discrim ที่ได้จากแพคเกจ CTT โดยใช้เกณฑ์การพิจารณาค่า discrim คือ ถ้าข้อใดมีตัวเลือก (ตัวลวง) ที่มีเฉพาะค่า discrim ที่เป็นค่าบวก ให้พิจารณาตัดตัวลวงที่มีค่าบวกมากที่สุดออก ถ้าข้อใดมีตัวเลือก (ตัวลวง) ที่มีเฉพาะค่า discrim ที่เป็นค่าลบ ให้พิจารณาตัดตัวลวงที่ติดลบค่าน้อย ๆ ออก แต่ถ้าข้อใดมีตัวเลือก (ตัวลวง) ที่มีค่า discrim ที่เป็นค่าลบ และค่าบวก ให้พิจารณาตัดตัวลวงที่มีค่าบวกออก พร้อมกับพิจารณาจากกราฟ OCC โดยเกณฑ์การพิจารณาตัดตัวเลือกจากกราฟ OCC ดูจากเส้นกราฟที่แสดงความน่าจะเป็นในการที่ผู้สอบจะเลือกตอบตัวเลือกนั้นที่มีเส้นกราฟอยู่ใกล้หรือประชิดแกนนอนมากที่สุด แสดงว่าความน่าจะเป็นที่ผู้สอบจะเลือกตอบตัวเลือกที่เป็นตัวลวงนั้นค่อนข้างน้อย หรือพิจารณาจากการไม่ปรากฏเส้นกราฟของตัวลวงนั้น แสดงว่าความน่าจะเป็นที่ผู้สอบจะเลือกตอบตัวลวงนั้นน้อยจนแทบจะไม่มีผู้สอบคนใดเลือกตัวลวงนั้นเลย สะท้อนให้เห็นว่าตัวลวงนั้นไม่มีประสิทธิภาพ เพราะลวงให้คนที่มีความสามารถต่ำมาตอบได้น้อย จากนั้นจึงทำการแทนที่ตัวลวงที่หายไปอย่างสุ่ม โดยแทนที่ตัวลวงที่ถูกตัดออกไปนั้นด้วยตัวเลือกตัวอื่น ๆ ที่เหลืออยู่ โดยรวมตัวเลือกที่เป็นเฉลยเข้าไปด้วย เพราะผู้วิจัยมีข้อตกลงเบื้องต้นว่า ผู้สอบทุกคนมีสิทธิ์ที่จะเลือกตอบตัวเลือกใด ๆ ก็ได้ นั่นหมายความว่า ตัวเลือกแต่ละตัวมีโอกาสที่ผู้สอบจะเลือกตอบเท่า ๆ กัน ผู้สอบที่มีความสามารถสูงก็มีสิทธิ์จะเลือกตัวเลือกที่เป็นได้ทั้งตัวลวงและเฉลย เช่นเดียวกับผู้สอบที่มีความสามารถต่ำ

```
Reliability analysis
Call: alpha(x = Score$scored, check.keys = FALSE)

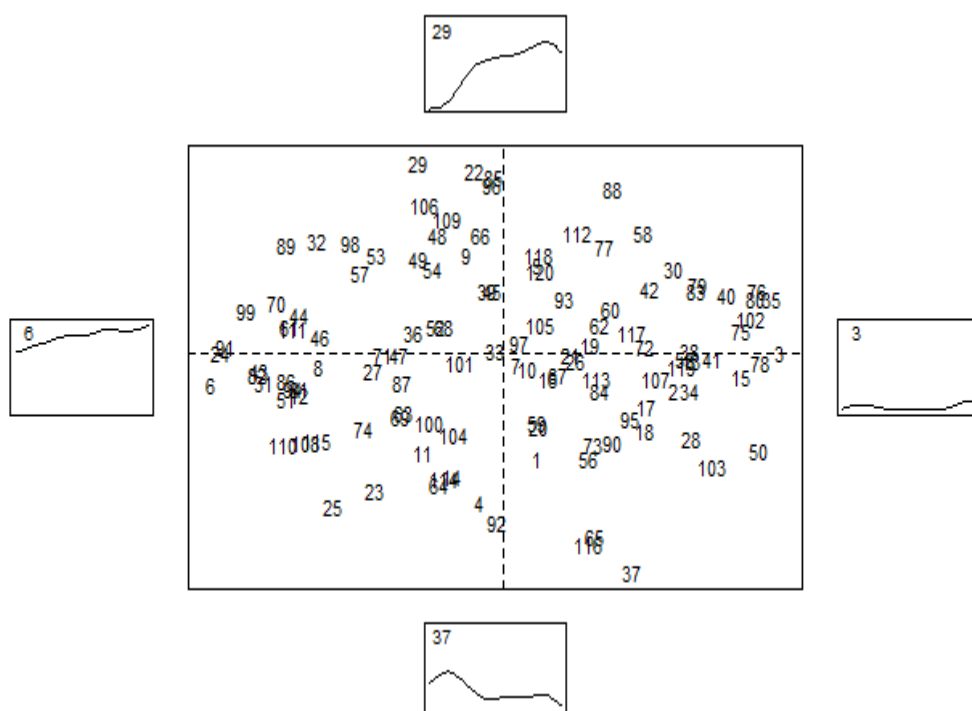
raw_alpha std.alpha G6(smc) average_r S/N ase mean sd median_r
0.79      0.78      0.8      0.029 3.6 0.0066 0.45 0.089 0.027

95% confidence boundaries
      lower alpha upper
Feldt  0.78  0.79  0.8
Duhachek 0.78  0.79  0.8
```

ภาพที่ 4.5 ผลการวิเคราะห์ค่าความเที่ยง กรณีที่ 3 แบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย)

จากการวิเคราะห์ความเที่ยงของแบบทดสอบ พบว่าค่าความเที่ยง (raw_alpha) ของแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) มีค่าเท่ากับ 0.79 ซึ่งมีค่าค่อนข้างสูง และมีค่าเท่ากับค่าความเที่ยงของแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) และมีค่าไม่แตกต่างจากความเที่ยงของแบบทดสอบแบบ 5 ตัวเลือกมากนัก โดยค่าความเที่ยงที่ได้มีค่าลดลงหลังจากการตัดตัวเลือกออกให้เหลือเพียง 4 ตัวเลือก

ความยากและอำนาจจำแนก จะพิจารณาการวิเคราะห์องค์ประกอบหลัก (Principle component analysis : PCA)



ภาพที่ 4.6 ผลการวิเคราะห์อำนาจจำแนกและความยากด้วย PCA กรณีที่ 3 แบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย)

จาก PCA พบว่าข้อสอบทั้งหมดจำนวน 120 ข้อ มีข้อสอบข้อ 29 ที่สามารถจำแนกได้ดี เช่นเดียวกับข้อ 22, 85, 88 และ 96 และข้ออื่น ๆ ที่อยู่ใกล้แนวเส้นประ ในแกนตั้ง ซึ่งอยู่เหนือแนวเส้นประ ในแกนนอน แต่เมื่ออยู่ใต้แนวเส้นประ ในแกนนอนจะเป็นข้อสอบที่จำแนกได้ไม่ค่อยดีนัก เช่น ข้อ 65, 37 และ 116 เป็นต้น โดยเฉพาะข้อ 37 ถือว่าเป็นข้อสอบที่จำแนกได้ไม่ดี หากพิจารณาข้อสอบข้อ 3, 78, 50, 76, 35, 80, 102, 75, 40 และ 15 เป็นต้น ถือว่าเป็นข้อสอบที่ค่อนข้างยาก โดยเฉพาะข้อที่อยู่ทางด้านปลายเส้นประ แกนนอน ด้านขวาของเส้นประ แกนตั้ง จะยิ่งสะท้อนว่าข้อสอบข้อนี้ยากมาก เช่น ข้อ 3, 78, 50, 102, 35, 76, 80 เป็นต้น สำหรับข้อที่อยู่ทางด้านเส้นประ แกนนอน ที่อยู่ทางด้านซ้ายของเส้นประ แกนตั้ง เช่น ข้อ 6, 94, 24 และ 99 ถือเป็นข้อสอบที่ค่อนข้างง่าย โดยเฉพาะข้อ 6 ถือเป็นข้อสอบ

ข้อที่ยากมาก สำหรับข้อสอบข้ออื่น ๆ ถือว่าอยู่ในเกณฑ์ใช้ได้ เมื่อพิจารณาจากข้อค้นพบนี้แล้ว พบว่า ข้อสอบส่วนใหญ่จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก แบบทดสอบ 4 ตัวเลือก (กรณีแทนค่า ตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) และ แบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวม ตัวเลือกที่เป็นเฉลย) มีคุณลักษณะด้านความยากง่ายและอำนาจจำแนกเป็นไปในทิศทางเดียวกัน แสดงว่า แบบทดสอบแบบเลือกตอบ 5 ตัวเลือก แบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) และ แบบทดสอบ 4 ตัวเลือก (กรณีแทน ค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) มีความยากและ อำนาจจำแนกไม่แตกต่างกัน

1.4 ผลการวิเคราะห์เปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และ แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก กรณีที่ 4 แบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ)

ค่าความเที่ยงของแบบทดสอบแบบ 4 ตัวเลือก ที่เกิดจากการตัดตัวเลือกที่จำแนกได้ไม่ได้ออก โดย ไม่มีการตัดตัวเลือกที่เป็นเฉลยออก แล้วใช้การพิจารณาตัดตัวเลือกที่เหลือโดยดูจากค่า discrim ที่ได้จาก แพกเกต CTT โดยใช้เกณฑ์การพิจารณาค่า discrim คือ ถ้าข้อใดมีตัวเลือก (ตัวลวง) ที่มีเฉพาะค่า discrim ที่เป็นค่าบวก ให้พิจารณาตัดตัวลวงที่มีค่าบวกมากที่สุดออก ถ้าข้อใดมีตัวเลือก (ตัวลวง) ที่มีเฉพาะค่า discrim ที่เป็นค่าลบ ให้พิจารณาตัดตัวลวงที่ติดลบค่าน้อย ๆ ออก แต่ถ้าข้อใดมีตัวเลือก (ตัวลวง) ที่มีค่า discrim ที่เป็นค่าลบ และค่าบวก ให้พิจารณาตัดตัวลวงที่มีค่าบวกออก พร้อมกับพิจารณาจากกราฟ OCC โดยเกณฑ์การพิจารณาตัดตัวเลือกจากกราฟ OCC ดูจากเส้นกราฟที่แสดงความน่าจะเป็นในการที่ผู้สอบ จะเลือกตอบตัวเลือกนั้นที่มีเส้นกราฟอยู่ใกล้หรือประชิดแกนนอนมากที่สุด แสดงว่าความน่าจะเป็นที่ ผู้สอบจะเลือกตอบตัวเลือกที่เป็นตัวลวงนั้นค่อนข้างน้อย หรือพิจารณาจากการไม่ปรากฏเส้นกราฟของตัว ลวงนั้น แสดงว่าความน่าจะเป็นที่ผู้สอบจะเลือกตอบตัวลวงนั้นน้อยจนแทบจะไม่มีผู้สอบคนใดเลือกตัวลวง นั้นเลย สะท้อนให้เห็นว่าตัวลวงนั้นไม่มีประสิทธิภาพ เพราะลวงให้คนที่มีความสามารถต่ำมาตอบได้น้อย จากนั้นจึงทำการแทนที่ตัวลวงที่หายไปโดยคำนึงถึงความสามารถของผู้สอบ เกณฑ์การพิจารณาแทนค่าตัว ลวงที่ถูกตัดออกไป เริ่มต้นจากการแบ่งกลุ่มผู้สอบออกเป็น 2 กลุ่ม คือ กลุ่มที่ 1 กลุ่มผู้สอบที่มี ความสามารถสูง พิจารณาจากผู้สอบที่ได้ผลคะแนนรวมเกินกว่าครึ่งหนึ่งของคะแนนเต็ม นั่นคือ คะแนน เต็มของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกที่นำมาใช้ในการวิจัยครั้งนี้ อยู่ที่ 120 คะแนน ครึ่งหนึ่ง คือ 60 หมายความว่าผู้สอบคนใดที่ได้คะแนนตั้งแต่ 60 คะแนนขึ้นไปจะถูกจัดอยู่ในกลุ่มผู้สอบที่มี ความสามารถสูง และกลุ่มที่ 2 คือ กลุ่มผู้สอบที่มีความสามารถต่ำ คือ กลุ่มผู้สอบที่ได้คะแนนต่ำกว่า 60 คะแนน ในกลุ่มที่ 1 จะถูกแทนค่าตัวลวงที่ตัดออกไป ด้วยตัวเลือกที่เป็นเฉลย สำหรับในกลุ่มที่ 2 จะถูก แทนค่าตัวลวงที่ถูกตัดออกไปด้วยตัวเลือกตัวอื่น ๆ ที่เหลือ โดยแทนค่าอย่างสุ่ม เหตุที่ผู้วิจัยทำเช่นนี้ เนื่องจากพิจารณาจากกราฟ OCC ซึ่งเมื่อพิจารณาจากกราฟ OCC สำหรับข้อสอบที่มีความยากง่ายและ

อำนาจจำแนกอยู่ในระดับเหมาะสม จะพบว่าผู้สอบที่ได้คะแนนสูง หรือกลุ่มคนที่มีความสามารถสูงมักมี โอกาสสูงที่จะเลือกตอบข้อสอบแต่ละข้อด้วยตัวเลือกที่เป็นเฉลย เส้นกราฟ OCC ของตัวเลือกที่เป็นเฉลย จะอยู่ในระดับความน่าจะเป็นที่สูง ส่วนผู้สอบที่มีความสามารถต่ำมักจะเลือกตัวเลือกตัวอื่น ๆ ที่ไม่ใช่เฉลย โดยกราฟ OCC ของตัวเลือกที่ไม่ใช่เฉลยจะอยู่เกาะกลุ่มกัน ด้วยระดับความน่าจะเป็นที่ไม่สูงมากนัก เพราะผู้วิจัยมีข้อตกลงเบื้องต้นว่า ผู้สอบที่มีความสามารถสูงย่อมต้องทำข้อสอบถูกทุกครั้ง ทุกกรณี แม้ว่าจะมีการปรับลดตัวเลือกของแบบทดสอบลงจาก 5 ตัวเลือก เหลือเพียง 4 ตัวเลือก ผู้สอบที่มีความสามารถสูงจะต้องทำข้อสอบข้อนั้น ๆ ถูกต้องเสมอ และสำหรับผู้ที่มีความสามารถต่ำก็มีโอกาสที่จะตอบตัวเลือกอื่น ๆ ที่ไม่ใช่ตัวเลือกที่เป็นเฉลย ด้วยความน่าจะเป็นในการเลือกตอบแต่ละตัวเลือกเท่า ๆ กัน เพราะคนที่มีความสามารถต่ำ แม้ว่าจะมีการปรับลดตัวเลือกของแบบทดสอบลงจาก 5 ตัวเลือก เหลือเพียง 4 ตัวเลือก ผู้สอบที่มีความสามารถต่ำ ก็ไม่มีโอกาสที่จะตอบข้อสอบข้อนั้นได้ถูกต้อง

```
Reliability analysis
Call: alpha(x = Score$scored, check.keys = FALSE)

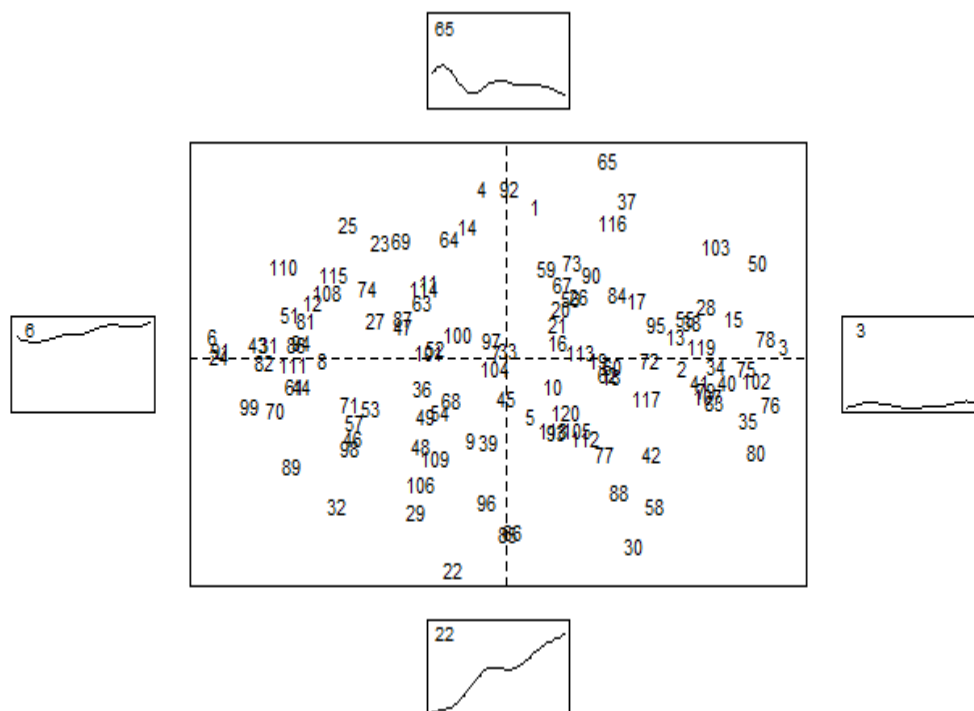
raw_alpha std.alpha G6(smc) average_r S/N ase mean sd median_r
0.87      0.87      0.88      0.053 6.7 0.004 0.44 0.11 0.05

95% confidence boundaries
      lower alpha upper
Feldt 0.86 0.87 0.88
Duhachek 0.86 0.87 0.88
```

ภาพที่ 4.7 ผลการวิเคราะห์ค่าความเที่ยง กรณีที่ 4 แบบทดสอบแบบ 4 ตัวเลือก
(กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ)

จากการวิเคราะห์ความเที่ยงของแบบทดสอบ พบว่าค่าความเที่ยง (raw_alpha) ของแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) มีค่าเท่ากับ 0.87 ซึ่งมีค่าค่อนข้างสูงเมื่อเทียบกับค่าความเที่ยงของแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) และค่าความเที่ยงของแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) ซึ่งมีค่าเท่ากับ 0.79 และมีค่าสูงกว่าค่าความเที่ยงของแบบทดสอบแบบ 5 ตัวเลือก ซึ่งมีค่าเท่ากับ 0.81

ความยากและอำนาจจำแนก จะพิจารณาการวิเคราะห์องค์ประกอบหลัก (Principle component analysis : PCA)



ภาพที่ 4.8 ผลการวิเคราะห์อำนาจจำแนกและความยากด้วย PCA กรณีที่ 4 แบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ)

จาก PCA พบว่าข้อสอบทั้งหมดจำนวน 120 ข้อ มีข้อสอบข้อ 22 ที่สามารถจำแนกได้ดี เช่นเดียวกับข้อ 30 และ 88 และข้ออื่น ๆ ที่อยู่ใกล้แนวเส้นประ ในแกนตั้ง ซึ่งอยู่ได้แนวเส้นประ ในแกนนอน แต่เมื่ออยู่เหนือแนวเส้นประ ในแกนนอนจะเป็นข้อสอบที่จำแนกได้ไม่ค่อยดีนัก เช่น ข้อ 65, 4, 37 และ 92 เป็นต้น โดยเฉพาะข้อ 65 ถือว่าเป็นข้อสอบที่จำแนกได้ไม่ดี หากพิจารณาข้อสอบข้อ 3, 78, 50, 76, 35, 80, 102, 75 และ 15 เป็นต้น ถือว่าเป็นข้อสอบที่ค่อนข้างยาก โดยเฉพาะข้อที่อยู่ทางด้านปลายเส้นประ แกนนอน ด้านขวาของเส้นประ แกนตั้ง จะยิ่งสะท้อนว่าข้อสอบข้อนี้ยากมาก เช่น ข้อ 3, 78, 50 76, 80 และ 102 เป็นต้น สำหรับข้อที่อยู่ทางด้านเส้นประ แกนนอน ที่อยู่ทางด้านซ้ายของเส้นประ แกนตั้ง เช่น ข้อ 6, 82, 24 และ 99 ถือเป็นข้อสอบที่ค่อนข้างง่าย โดยเฉพาะข้อ 6 ถือเป็นข้อสอบข้อที่ง่ายมาก สำหรับข้อสอบข้ออื่น ๆ ถือว่าอยู่ในเกณฑ์ใช้ได้ เมื่อพิจารณาจากข้อค้นพบนี้แล้ว พบว่าข้อสอบส่วนใหญ่จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก แบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) แบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) และแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) มีคุณลักษณะด้านความง่ายและอำนาจจำแนกเป็นไปในทิศทางเดียวกัน แสดงว่าแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก แบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) แบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) และแบบทดสอบแบบ 4 ตัวเลือก

(กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) มีความยากและอำนาจจำแนกไม่แตกต่างกัน

จากผลการวิเคราะห์คุณภาพแบบทดสอบ 4 ฉบับ ในด้านความเที่ยงของแบบทดสอบ 4 กรณี ประกอบด้วย กรณีที่ 1. แบบทดสอบแบบเลือกตอบ 5 ตัวเลือก กรณีที่ 2 . แบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) กรณีที่ 3. แบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) และ กรณีที่ 4. แบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) แสดงรายละเอียดการเปรียบเทียบดังตารางต่อไปนี้

ตารางที่ 4.1 ค่าความเที่ยงของแบบทดสอบทั้ง 4 กรณี

ประเภท แบบทดสอบ	กรณีที่ 1	กรณีที่ 2	กรณีที่ 3	กรณีที่ 4
ค่าความเที่ยง	0.81	0.79	0.79	0.87

จากตารางข้างต้น พบว่าแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) ให้ค่าความเที่ยงสูงสุด รองลงมาคือ แบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) มีค่าความเที่ยงเท่ากัน และให้ค่าความเที่ยงต่ำสุด เมื่อพิจารณาเปรียบเทียบทั้ง 4 กรณี

สำหรับคุณภาพด้านความยากและอำนาจจำแนกจะพิจารณาเปรียบเทียบความคล้ายคลึงกันของข้อที่มีความยากง่ายและอำนาจจำแนกในลักษณะที่เป็นไปในทิศทางเดียวกัน โดยจะพิจารณาในภาพรวมซึ่งจะใช้ PCA มาช่วยพิจารณาดังตารางเปรียบเทียบด้านล่างนี้

ตารางที่ 4.2 ผลการวิเคราะห์คุณภาพด้านความยากและอำนาจจำแนกด้วย PCA ของแบบทดสอบทั้ง 4 กรณี

ประเภทแบบทดสอบ		กรณีที่ 1	กรณีที่ 2	กรณีที่ 3	กรณีที่ 4
ความยาก	ง่าย	6, 24, 82, 31, 91, 88, 110, 74, 89	6, 91, 24, 99, 48, 82	6, 94, 24, 99	6, 82, 24, 99
	ยาก	103, 50, 3, 78, 35, 76,	103, 50, 3, 78, 35, 76,	3, 78, 50, 76, 35, 80, 102,	3, 78, 50, 76, 35, 80, 102,

ตารางที่ 4.2 (ต่อ)

ประเภทแบบทดสอบ		กรณีที่ 1	กรณีที่ 2	กรณีที่ 3	กรณีที่ 4
ความยาก	ยาก	80, 102, 75, 107, 110, 41, 15	80, 102, 75, 107, 41, 15	75, 40, 15	75, 15
อำนาจจำแนก	ดี	85 29, 66, 88, 106, 32, 22 96	85, 29, 66, 88, 106, 32, 22, 96	29, 22, 85, 88, 96	22, 30, 88
	ไม่ดี	65, 37, 4, 92 1, 4, 37	65, 37, 4, 92, 1	65, 37, 116	65, 4, 37, 92

จากตารางพบว่าข้อสอบส่วนใหญ่จากแบบทดสอบ 4 กรณีที่ได้จากการพิจารณาโดยอาศัย PCA มีคุณลักษณะด้านความยากและอำนาจจำแนกเป็นไปในทิศทางเดียวกัน หมายความว่า แบบทดสอบทั้ง 4 กรณีมีคุณสมบัติด้านความยากและอำนาจจำแนกไม่แตกต่างกัน

ตอนที่ 2 ผลการเปรียบเทียบคะแนนสอบของนักศึกษาระหว่างคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกและแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (3 กรณี ที่ผู้วิจัยศึกษา) และเปรียบเทียบคะแนนเฉลี่ยของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกและแบบทดสอบแบบเลือกตอบ 4 ตัวเลือกในแต่ละกรณี (จากทั้งหมด 3 กรณี ที่ผู้วิจัยศึกษา)

ผลการวิเคราะห์ในตอนต้นที่ 2 เป็นการนำเสนอข้อมูลใน 2 ส่วน ส่วนที่ 1 คือ ผลการวิเคราะห์ข้อมูลผลการเปรียบเทียบคะแนนจากแบบทดสอบสองฉบับเป็นรายกรณีดังต่อไปนี้ 1) ผลการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) 2) ผลการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) 3) ผลการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) สำหรับงานวิจัยนี้จะใช้รูปแบบการเปรียบเทียบโดยใช้ทฤษฎีดั้งเดิม (Classical Model of Equating) ด้วยการเปรียบเทียบคะแนนเชิงเส้นตรง (Linear Equating) ซึ่งเป็นการเปรียบเทียบแบบทดสอบสองฉบับ คือ แบบทดสอบฉบับ X และแบบทดสอบฉบับ Y ที่วัดในสิ่งเดียวกัน และมีความเที่ยงสูงพอ ๆ กัน อีกทั้งยังเป็นวิธีการที่นิยมใช้และมีคำสั่งให้วิเคราะห์ข้อมูลในโปรแกรมสำเร็จรูป โดยผลการวิเคราะห์ข้อมูลที่ได้จากงานวิจัยนี้ มาจากโปรแกรม Jamovi

สำหรับผลการวิเคราะห์ข้อมูลในส่วนที่ 2 เป็นผลการศึกษาเปรียบเทียบคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก กับแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ในแต่ละกรณี ทั้ง 3 กรณี เพื่อทดสอบสมมติฐานการวิจัยของผู้วิจัยที่ว่า คะแนนสอบของนักศึกษาที่ได้จากการทำแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ในแต่ละกรณีสูงกว่าคะแนนสอบของนักศึกษาที่ได้จากการทำแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก เพื่อเป็นการทดสอบและยืนยันทฤษฎีและผลการวิจัยที่ผ่านมาที่ว่าเมื่อลดจำนวนตัวเลือกของแบบทดสอบแบบเลือกตอบลงจะทำให้คุณสมบัติของแบบทดสอบด้านความยากและอำนาจจำแนกลดลงไปด้วย เมื่อความยากของแบบทดสอบแบบเลือกตอบมีค่าลดลง นั้นหมายความว่าแบบทดสอบแบบเลือกตอบจะมีความง่ายเพิ่มขึ้น ดังนั้นนักศึกษาควรจะต้องทำแบบทดสอบแล้วได้คะแนนเพิ่มมากขึ้น จึงเป็นที่มาของการตั้งสมมติฐานการวิจัยในครั้งนี้ว่า คะแนนสอบของนักศึกษาที่ได้จากการทำแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ในแต่ละกรณีสูงกว่าคะแนนสอบของนักศึกษาที่ได้จากการทำแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก สำหรับวิธีการวิเคราะห์ข้อมูลที่จะใช้ในการตอบวัตถุประสงค์การวิจัยข้อนี้ จะใช้การพิจารณาจากค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และการทดสอบ Wilcoxon Signed Rank test เนื่องจากผู้วิจัยสร้างข้อตกลงเบื้องต้นขึ้นมาว่า จะใช้ผู้สอบเพียง 1 กลุ่ม ในการทำแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ในแต่ละกรณี ทั้ง 3 กรณี

ผู้วิจัยขอแสดงผลการวิเคราะห์ข้อมูลส่วนแรก คือ ผลการเปรียบเทียบคะแนนจากแบบทดสอบสองฉบับเป็นรายกรณีดังต่อไปนี้

1. ผลการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) แสดงได้ดังตารางที่ 4.3

ตารางที่ 4.3 ผลการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย)

สมการเชิงเส้น $y = 0 + 1x$

จุดตัด (ค่าคงที่)	ความชัน
0.00	1.00

จากตารางที่ 4.3 พบว่า สมการเชิงเส้นที่ใช้ในการแปลงคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกให้เป็นคะแนนของแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก คือ สมการเชิงเส้น $y = 0 + 1x$ ที่มีค่าความชันเป็น 1 และค่าคงที่เป็น 0

ผลการเปรียบเทียบคะแนน ได้คะแนนสมมูลดังตารางที่ 4.4

ตารางที่ 4.4 คะแนนสมมูลที่ได้จากการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย)

5 ตัวเลือก	4 ตัวเลือก
5	5
18	18
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46

ตารางที่ 4.4 (ต่อ)

5 ตัวเลือก	4 ตัวเลือก
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76

ตารางที่ 4.4 (ต่อ)

5 ตัวเลือก	4 ตัวเลือก
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
91	91
93	93
94	94

2. ผลการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) แสดงได้ดังตารางที่ 4.5

ตารางที่ 4.5 ผลการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย)

$$\text{สมการเชิงเส้น } y = 3.999 + 0.969x$$

จุดตัด (ค่าคงที่)	ความชัน
4.00	0.969

จากตารางที่ 4.5 พบว่า สมการเชิงเส้นที่ใช้ในการแปลงคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกให้เป็นคะแนนของแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก คือ สมการเชิงเส้น $y = 3.999 + 0.969x$ ที่มีค่าความชันเป็น 0.969 และค่าคงที่เป็น 4.00

ผลการเปรียบเทียบคะแนน ได้คะแนนสมมูลดังตารางที่ 4.6

ตารางที่ 4.6 คะแนนสมมูลที่ได้จากการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย)

5 ตัวเลือก	4 ตัวเลือก
5	8.843475
18	21.43981
22	25.31561
23	26.28456
24	27.25351
25	28.22246
26	29.19141
27	30.16036
28	31.12931
29	32.09826
30	33.06721
31	34.03615
32	35.0051
33	35.97405
34	36.943
35	37.91195
36	38.8809
37	39.84985
38	40.8188
39	41.78775
40	42.7567
41	43.72565
42	44.6946
43	45.66354
44	46.63249
45	47.60144
46	48.57039
47	49.53934

ตารางที่ 4.6 (ต่อ)

5 ตัวเลือก	4 ตัวเลือก
48	50.50829
49	51.47724
50	52.44619
51	53.41514
52	54.38409
53	55.35304
54	56.32199
55	57.29094
56	58.25988
57	59.22883
58	60.19778
59	61.16673
60	62.13568
61	63.10463
62	64.07358
63	65.04253
64	66.01148
65	66.98043
66	67.94938
67	68.91833
68	69.88727
69	70.85622
70	71.82517
71	72.79412
72	73.76307
73	74.73202
74	75.70097
75	76.66992
76	77.63887
77	78.60782

ตารางที่ 4.6 (ต่อ)

5 ตัวเลือก	4 ตัวเลือก
78	79.57677
79	80.54572
80	81.51467
81	82.48361
82	83.45256
83	84.42151
84	85.39046
85	86.35941
86	87.32836
91	92.17311
93	94.11101
94	95.07995

3. ผลการปรับเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) แสดงได้ดังตารางที่ 4.7

ตารางที่ 4.7 ผลการปรับเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ)

$$\text{สมการเชิงเส้น } y = -9.722 + 1.222x$$

จุดตัด (ค่าคงที่)	ความชัน
- 9.72	1.22

จากตารางที่ 4.7 พบว่า สมการเชิงเส้นที่ใช้ในการแปลงคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกให้เป็นคะแนนของแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก คือ สมการเชิงเส้น $y = -9.722 + 1.222x$ ที่มีค่าความชันเป็น 1.22 และค่าคงที่เป็น - 9.72

ผลการปรับเทียบคะแนน ได้คะแนนสมมูลดังตารางที่ 4.8

ตารางที่ 4.8 คะแนนสมมูลที่ได้จากการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ)

5 ตัวเลือก	4 ตัวเลือก
5	-3.61447
18	12.26615
22	17.1525
23	18.37409
24	19.59567
25	20.81726
26	22.03884
27	23.26043
28	24.48202
29	25.7036
30	26.92519
31	28.14678
32	29.36836
33	30.58995
34	31.81154
35	33.03312
36	34.25471
37	35.4763
38	36.69788
39	37.91947
40	39.14106
41	40.36264
42	41.58423
43	42.80582
44	44.0274
45	45.24899
46	46.47057
47	47.69216

ตารางที่ 4.8 (ต่อ)

5 ตัวเลือก	4 ตัวเลือก
48	48.91375
49	50.13533
50	51.35692
51	52.57851
52	53.80009
53	55.02168
54	56.24327
55	57.46485
56	58.68644
57	59.90803
58	61.12961
59	62.3512
60	63.57279
61	64.79437
62	66.01596
63	67.23755
64	68.45913
65	69.68072
66	70.9023
67	72.12389
68	73.34548
69	74.56706
70	75.78865
71	77.01024
72	78.23182
73	79.45341
74	80.675
75	81.89658
76	83.11817
77	84.33976

ตารางที่ 4.8 (ต่อ)

5 ตัวเลือก	4 ตัวเลือก
78	85.56134
79	86.78293
80	88.00452
81	89.2261
82	90.44769
83	91.66928
84	92.89086
85	94.11245
86	95.33403
91	101.442
93	103.8851
94	105.1067

สำหรับผลการวิเคราะห์ข้อมูลในส่วนที่ 2 เป็นผลการศึกษาเปรียบเทียบคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก กับแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ในแต่ละกรณี ทั้ง 3 กรณี เพื่อทดสอบสมมติฐานการวิจัยของผู้วิจัยที่ว่า คะแนนสอบของนักศึกษาที่ได้จากการทำแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ในแต่ละกรณีสูงกว่าคะแนนสอบของนักศึกษาที่ได้จากการทำแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก โดยมีสมมติฐานที่ต้องทดสอบ ดังนี้ 1) คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก 2) คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และ 3) คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก การทดสอบสมมติฐานเหล่านี้จะใช้วิธีการวิเคราะห์ทางสถิติ โดยใช้การพิจารณาจากค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และการทดสอบ Wilcoxon Signed Rank test ด้วยโปรแกรม Jamovi ซึ่งผลการวิเคราะห์ข้อมูลจะประกอบด้วยผลการตรวจสอบข้อตกลงเบื้องต้นของข้อมูล และผลการทดสอบสมมติฐานตามรายละเอียดดังนี้

1. คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) มากกว่าคะแนนของ

นักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก โดยรายละเอียดของผลการวิเคราะห์ข้อมูลเพื่อทดสอบสมมติฐานข้างต้น เป็นดังนี้

Descriptives					
	N	Mean	Median	SD	SE
5ch	1966	51.5	51.0	11.1	0.250
4cht1	1966	51.5	51.0	11.1	0.250

ภาพที่ 4.9 ผลการวิเคราะห์คะแนนเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของนักศึกษา

ที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก

(กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่มด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย)

จากภาพที่ 4.9 พบว่า จากนักศึกษาทั้งหมดที่เข้าสอบในชุดวิชาตั้งต้นที่ถูกนำมาใช้เป็นกรณีศึกษาในงานวิจัยนี้ จำนวน 1,966 คน คะแนนโดยเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบ 5 ตัวเลือก (5ch) และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (4cht1) (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) มีค่าเท่ากัน นั่นหมายความว่าไม่ว่านักศึกษาจะสอบด้วยแบบทดสอบแบบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) ก็จะมีคะแนนโดยเฉลี่ยเท่ากัน เท่ากับ 51.5 คะแนน และเมื่อพิจารณาจากส่วนเบี่ยงเบนมาตรฐานของคะแนนของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) พบว่ามีค่าเท่ากัน เท่ากับ 11.1 คะแนน แสดงว่าการกระจายตัวหรือความแตกต่างของคะแนนสอบของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) พอๆ กัน ในกรณีนี้จึงไม่จำเป็นต้องใช้การทดสอบสมมติฐานทางสถิติเพื่อสรุปผลการทดสอบสมมติฐาน เพราะค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานให้สารสนเทศที่ครบถ้วน จนสามารถสรุปผลการทดสอบสมมติฐานโดยไม่จำเป็นต้องทำการทดสอบสมมติฐาน จากข้อสรุปข้างต้นสะท้อนให้เห็นว่าไม่จำเป็นต้องสร้างและจัดทำแบบทดสอบถึง 5 ตัวเลือก เพราะการใช้แบบทดสอบเพียงแค่ 4 ตัวเลือก ก็ทำให้นักศึกษาได้คะแนนเท่ากับในกรณีที่นักศึกษาลงมือทำแบบทดสอบ 5 ตัวเลือก

2. คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก โดยรายละเอียดของผลการวิเคราะห์ข้อมูลเพื่อทดสอบสมมติฐานข้างต้น แสดงดังภาพที่ 4.10

Descriptives					
	N	Mean	Median	SD	SE
5ch	1966	51.5	51.0	11.1	0.250
4cht2	1966	53.9	54.0	10.7	0.242

ภาพที่ 4.10 ผลการวิเคราะห์คะแนนเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของนักศึกษา
ที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก

(กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย)

จากภาพที่ 4.10 พบว่า จากนักศึกษาทั้งหมดที่เข้าสอบในชุดวิชาตั้งต้นที่ถูกนำมาใช้เป็นกรณีศึกษาในงานวิจัยนี้ จำนวน 1,966 คน คะแนนโดยเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบ 5 ตัวเลือก (5ch) เท่ากับ 51.5 คะแนน และ 11.1 คะแนน และคะแนนโดยเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (4cht2) (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) เท่ากับ 53.9 คะแนน และ 10.7 คะแนน ซึ่งคะแนนโดยเฉลี่ยของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบเลือกตอบ 4 ตัวเลือกสูงกว่าคะแนนโดยเฉลี่ยของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบ 5 ตัวเลือก และส่วนเบี่ยงเบนมาตรฐานของคะแนนของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) น้อยกว่าส่วนเบี่ยงเบนมาตรฐานของคะแนนของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก แสดงว่าคะแนนของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) มีการเกาะกลุ่มของคะแนนมากกว่าคะแนนของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

ผลการทดสอบข้อตกลงเบื้องต้นของข้อมูลในเรื่องของการแจกแจงแบบปกติของข้อมูล แสดงดังภาพที่ 4.11

Normality Test (Shapiro-Wilk)				
			W	p
5ch	-	4cht2	0.916	< .001

Note. A low p-value suggests a violation of the assumption of normality

ภาพที่ 4.11 ผลการวิเคราะห์เพื่อทดสอบข้อตกลงเบื้องต้นของข้อมูล
ในเรื่องของการแจกแจงแบบปกติของข้อมูล

จากภาพที่ 4.11 ค่า p – value ของ Shapiro – Wilk ซึ่งมีค่า < .001 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญ .05 พบว่า ข้อมูลไม่มีการแจกแจงแบบปกติ ดังนั้นไม่สามารถใช้ Paired t – test ในการวิเคราะห์ข้อมูลได้ จึงต้องใช้ Wilcoxon Signed Rank test ในการวิเคราะห์ข้อมูล ได้ผลการวิเคราะห์ดังภาพที่ 4.12

Paired Samples T-Test					
			Statistic	df	p
5ch	4cht2	Student's t	-64.3	1965	< .001
		Wilcoxon W	0.00 ^a		< .001

Note. H₀: μ Measure 1 - Measure 2 < 0
^a 197 pair(s) of values were tied

ภาพที่ 4.12 ผลการวิเคราะห์ข้อมูลด้วย Wilcoxon Signed Rank test

จากภาพที่ 4.12 ค่า p – value ของ Wilcoxon W มีค่า < .001 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญทางสถิติที่ระดับ .05 พบว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

3. คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก โดยรายละเอียดของผลการวิเคราะห์ข้อมูลเพื่อทดสอบสมมติฐานข้างต้น แสดงดังภาพที่ 4.13

Descriptives					
	N	Mean	Median	SD	SE
5ch	1966	51.5	51.0	11.1	0.250
4cht3	1966	53.2	51.0	13.5	0.305

ภาพที่ 4.13 ผลการวิเคราะห์ข้อมูลเพื่อทดสอบสมมติฐานว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ)

จากภาพที่ 4.13 พบว่า จากนักศึกษาทั้งหมดที่เข้าสอบในชุดวิชาตั้งต้นที่ถูกนำมาใช้เป็นกรณีศึกษาในงานวิจัยนี้ จำนวน 1,966 คน คะแนนโดยเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนของ

นักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบ 5 ตัวเลือก (5ch) เท่ากับ 51.5 คะแนน และ 11.1 คะแนน และคะแนนโดยเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (4cht3) (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) เท่ากับ 53.2 คะแนน และ 13.5 คะแนน ซึ่งคะแนนโดยเฉลี่ยของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบเลือกตอบ 4 ตัวเลือกสูงกว่าคะแนนโดยเฉลี่ยของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบ 5 ตัวเลือก และส่วนเบี่ยงเบนมาตรฐานของคะแนนของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) มากกว่าส่วนเบี่ยงเบนมาตรฐานของคะแนนของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก แสดงว่าคะแนนของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) มีการกระจายหรือมีความแตกต่างของคะแนนมากกว่าคะแนนของนักศึกษาที่ได้จากการสอบด้วยแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

ผลการทดสอบข้อตกลงเบื้องต้นของข้อมูลในเรื่องของการแจกแจงแบบปกติของข้อมูล แสดงดังภาพที่ 4.14

Normality Test (Shapiro-Wilk)				
			W	p
5ch	-	4cht3	0.560	< .001

Note. A low p-value suggests a violation of the assumption of normality

ภาพที่ 4.14 การทดสอบข้อตกลงเบื้องต้นของข้อมูลในเรื่องของการแจกแจงแบบปกติของข้อมูล

จากภาพที่ 4.14 ค่า p – value ของ Shapiro – Wilk ซึ่งมีค่า < .001 ซึ่งมีความน้อยกว่าระดับนัยสำคัญ .05 พบว่า ข้อมูลไม่มีการแจกแจงแบบปกติ ดังนั้นไม่สามารถใช้ Paired t – test ในการวิเคราะห์ข้อมูลได้ จึงต้องใช้ Wilcoxon Signed Rank test ในการวิเคราะห์ข้อมูล ได้ผลการวิเคราะห์ดังภาพที่ 4.15

Paired Samples T-Test					
			Statistic	df	p
5ch	4cht3	Student's t	-22.2	1965	< .001
		Wilcoxon W	0.00*		< .001

Note. H_a: μ Measure 1 - Measure 2 < 0

* 1526 pair(s) of values were tied

ภาพที่ 4.15 ผลการวิเคราะห์ข้อมูลด้วยสถิติ Wilcoxon Signed Rank test

จากภาพที่ 4.15 พบว่า ค่า p - value ของ Wilcoxon W มีค่า $< .001$ ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญ .05 พบว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

จากผลการทดสอบสมมติฐานเพื่อการศึกษาเปรียบเทียบคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก กับแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ในแต่ละกรณี ทั้ง 3 กรณี เพื่อทดสอบสมมติฐานการวิจัยของผู้วิจัยที่ว่า คะแนนสอบของนักศึกษาที่ได้จากการทำแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ในแต่ละกรณีสูงกว่าคะแนนสอบของนักศึกษาที่ได้จากการทำแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก โดยมีสมมติฐานที่ต้องทดสอบ ดังนี้ 1. คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก 2. คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และ 3. คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ได้ข้อค้นพบว่า คะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก ทั้ง 3 กรณีมากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

ตอนที่ 3 ผลการวิเคราะห์เปรียบเทียบค่าใช้จ่ายในการจัดทำแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก

ผลการวิเคราะห์ข้อมูลในตอนที่ 3 เป็นการนำเสนอผลการวิเคราะห์ค่าใช้จ่ายในการจัดทำแบบทดสอบและนำเสนอผลของการจำลองตัดแบบทดสอบจาก 5 ตัวเลือกให้เหลือเพียง 4 ตัวเลือก เพื่อเปรียบเทียบค่าใช้จ่ายในการจัดทำแบบทดสอบ ในที่นี้จะหมายถึงร่างต้นฉบับแบบทดสอบทั้งกระบวนการก่อนที่จะดำเนินการนำร่างแบบทดสอบฉบับสมบูรณ์ส่งเข้าไปดำเนินการจัดพิมพ์เป็นแบบทดสอบฉบับจริงที่สำนักพิมพ์เพื่อพร้อมใช้ในการดำเนินการสอบภาคสนาม โดยมีข้อตกลงที่ใช้ในการวิจัยครั้งนี้เกี่ยวกับการวิเคราะห์ค่าใช้จ่าย คือ ผู้วิจัยจะทำการคำนวณค่าใช้จ่ายในการจัดทำแบบทดสอบเฉพาะค่าใช้จ่ายที่เป็นตัวเงินที่เกิดขึ้นจากกระบวนการทำงานของสำนักทะเบียนและวัดผลเท่านั้น ไม่ได้คิดค่าใช้จ่ายที่เกี่ยวข้องกับสำนักพิมพ์และไม่คิดค่าใช้จ่ายสำหรับอาจารย์ที่ออกข้อสอบ ทั้งนี้ค่าใช้จ่ายที่ไม่เป็นตัวเงินหรือปัจจัยอื่นที่ไม่ใช่ตัวเงิน เช่น เวลาที่ใช้ในการตรวจร่างแบบทดสอบ ความเหนื่อยล้าในการจัดทำแบบทดสอบ จะไม่ถูกนำมาคำนวณในส่วน of ค่าใช้จ่ายในงานวิจัยนี้ มีรายละเอียดดังนี้

3.1 ค่าตอบแทนผู้ที่มีส่วนเกี่ยวข้องในกระบวนการจัดทำแบบทดสอบก่อนส่งเข้าโรงพิมพ์

ในกระบวนการดำเนินการจัดเตรียมแบบทดสอบของศูนย์วิจัยและพัฒนาแบบทดสอบ สำนักทะเบียนและวัดผล มหาวิทยาลัยสุโขทัยธรรมาธิราช มีการดำเนินการจ่ายค่าตอบแทนให้กับผู้ที่มีหน้าที่รับผิดชอบในกระบวนการจัดทำแบบทดสอบ ประกอบด้วย ผู้พิมพ์แบบทดสอบกับผู้ตรวจร่างแบบทดสอบ โดยผู้ตรวจร่างแบบทดสอบ ประกอบไปบุคคลดังต่อไปนี้

3.1.1 อาจารย์นักวัดผล (Reviewer) หมายถึง อาจารย์นักวัดผลประจำชุดวิชาที่ได้รับมอบหมาย ทำหน้าที่ในการอ่านและประเมินคุณภาพของแบบทดสอบแต่ละข้อก่อนที่จะมีการดำเนินการจัดพิมพ์ โดยจะอ่านและตรวจสอบในกระบวนการนี้เฉพาะชุดวิชาที่มีการออกข้อสอบใหม่เท่านั้น หน้าที่หลักของตำแหน่งนี้จะเป็นการอ่านเพื่อตรวจสอบโครงสร้างของแบบทดสอบว่าข้อสอบแต่ละข้อนั้นออกสอบได้ตรงตามวัตถุประสงค์ของเนื้อหาในแต่ละหน่วย ตอน ของชุดวิชาที่ระบุไว้ที่กำหนดหรือไม่ หากมีข้อสอบข้อใดที่มีข้อคำถามไม่ชัดเจน ข้อสอบซ้ำ ขาดเฉลย เฉลยไม่ถูกต้อง คำถามกำกวม หรือข้อคำถามไม่มีความชัดเจนว่าวัดในเนื้อหาใด หน่วย ตอน และพฤติกรรมการเรียนรู้ใด Reviewer จะเรียงแจ้งไปยังประธานชุดวิชาให้พิจารณาทบทวนและแก้ไขปัญหาข้อสอบก่อนส่งไปให้ผู้พิมพ์แบบทดสอบดำเนินการจัดพิมพ์แบบทดสอบ หน่วย ตอน โดยแต่ละกระบวนการดำเนินงานมีรายละเอียดและรายละเอียดของค่าใช้จ่ายตามตารางที่ 4 ดังนี้

ผู้ที่อ่านข้อสอบร่าง 1 หมายถึง นักวิชาการศึกษาหรืออาจารย์นักวัดผล ทำหน้าที่อ่านและตรวจสอบแบบทดสอบในรอบที่ 1 ที่ผู้พิมพ์แบบทดสอบได้พิมพ์ส่ง โดยจะตรวจสอบความถูกต้องของการจัดพิมพ์ ข้อมูลในแบบทดสอบ การระบุเฉลย ความซ้ำซ้อนของตัวเลือก รูปแบบของข้อสอบในแต่ละข้อที่จัดพิมพ์เข้าฉบับตามแบบฟอร์ม ซึ่งเป็นการตรวจสอบในรอบแรกต่อการคัดกรองและตรวจสอบที่ผ่านมาจาก Reviewer สำหรับแบบทดสอบที่จัดเป็นชุดวิชายาก (ชุดวิชาคำนวณ ชุดวิชาทางภาษาศาสตร์) ผู้ที่อ่านข้อสอบร่าง 1 จะเป็นอาจารย์นักวัดผลเท่านั้นที่มีหน้าที่รับผิดชอบ สำหรับชุดวิชาที่ไม่ใช่ชุดยาก ผู้ที่อ่านข้อสอบร่าง 1 จะเป็นนักวิชาการศึกษา หากมีประเด็นแก้ไขที่พบในขั้นตอนนี้ ผู้ที่อ่านข้อสอบร่าง 1 จะระบุรายละเอียดของการแก้ไขเพื่อแจ้งและดำเนินการส่งให้ผู้พิมพ์แบบทดสอบที่ได้รับมอบหมายให้พิมพ์แบบทดสอบในชุดวิชานั้น ๆ เป็นผู้รับผิดชอบในการแก้ไข และผู้ที่อ่านข้อสอบร่าง 1 จะทำหน้าที่ในการตรวจสอบความเรียบร้อยทั้งหมดของแบบทดสอบที่ผู้พิมพ์แบบทดสอบได้แก้ไขตามที่ผู้ที่อ่านข้อสอบร่าง 4 ระบุไว้ และทำหน้าที่ลงนามเพื่อจัดส่งข้อสอบเข้าสู่กระบวนการจัดพิมพ์แบบทดสอบที่โรงพิมพ์ต่อไป

ผู้ที่อ่านข้อสอบร่าง 2 หมายถึง อาจารย์นักวัดผล ทำหน้าที่ในการอ่านและตรวจสอบความเรียบร้อยของแบบทดสอบต่อจากผู้ที่อ่านข้อสอบร่าง 1 การทำงานในขั้นตอนนี้ จะเป็นการตรวจสอบซ้ำเพื่อดูความถูกต้องเรียบร้อยของแบบทดสอบ และตรวจสอบประเด็นแก้ไข หรือประเด็นข้อคำถามที่ผู้ที่อ่านข้อสอบร่าง 1 ได้บันทึกรายละเอียดของปัญหาไว้ อาจารย์นักวัดผลที่ทำหน้าที่อ่านข้อสอบร่างที่ 2 จะต้องคอยตรวจสอบและแก้ไขปัญหาข้อสอบที่เกิดขึ้นในขั้นตอนนี้ ตลอดจนเป็นการตรวจสอบซ้ำในประเด็นของข้อมูลในแบบทดสอบ การระบุเฉลย ความซ้ำซ้อนของตัวเลือก รูปแบบของข้อสอบในแต่ละข้อที่จัดพิมพ์เข้าฉบับตามแบบฟอร์ม หากมีประเด็นแก้ไขที่พบในขั้นตอนนี้ ผู้ที่อ่านข้อสอบร่าง 2 จะระบุ

รายละเอียดของการแก้ไขเพื่อแจ้งและดำเนินการส่งให้ผู้พิมพ์แบบทดสอบที่ได้รับมอบหมายให้พิมพ์แบบทดสอบในชุดวิชานั้น ๆ เป็นผู้รับผิดชอบในการแก้ไข

ผู้ที่อ่านข้อสอบร่าง 3 หมายถึง นักวิชาการศึกษาหรืออาจารย์นักวัดผล ทำหน้าที่อ่านและตรวจสอบแบบทดสอบในรอบที่ 3 โดยจะตรวจสอบความถูกต้องของการจัดพิมพ์ ข้อมูลในแบบทดสอบ การระบุเฉลย ความซ้ำซ้อนของตัวเลือก รูปแบบของข้อสอบในแต่ละข้อที่จัดพิมพ์เข้าฉบับตามแบบฟอร์มอีกครั้งซึ่งเป็นการตรวจสอบในรอบแรกต่อการคัดกรองและตรวจสอบที่ผ่านมาจาก Reviewer สำหรับแบบทดสอบที่จัดเป็นชุดวิทยายก (ชุดวิชาคำนวณ ชุดวิชาทางภาษาศาสตร์) ผู้ที่อ่านข้อสอบร่าง 3 จะเป็นอาจารย์นักวัดผลเท่านั้นที่มีหน้าที่รับผิดชอบ สำหรับชุดวิชาที่ไม่ใช่ชุดยก ผู้ที่อ่านข้อสอบร่าง 3 จะเป็นนักวิชาการศึกษา หากมีประเด็นแก้ไขจะดำเนินการส่งให้ผู้พิมพ์แบบทดสอบที่ได้รับมอบหมายให้พิมพ์แบบทดสอบในชุดวิชานั้น ๆ เป็นผู้รับผิดชอบในการแก้ไข ขั้นตอนนี้เป็นเตรียมแบบทดสอบให้มีหน้าตาเหมือนแบบทดสอบฉบับจริงที่นักศึกษาจะใช้สอบมากที่สุด เพื่อดำเนินการส่งให้ประธานชุดวิชาทำหน้าที่ตรวจสอบอีกครั้ง ก่อนดำเนินการส่งเข้าโรงพิมพ์เพื่อเข้าสู่กระบวนการจัดพิมพ์แบบทดสอบสำหรับการสอบต่อไป หากมีประเด็นแก้ไขที่พบในขั้นตอนนี้ ผู้ที่อ่านข้อสอบร่าง 3 จะระบุรายละเอียดของการแก้ไขเพื่อแจ้งและดำเนินการส่งให้ผู้พิมพ์แบบทดสอบที่ได้รับมอบหมายให้พิมพ์แบบทดสอบในชุดวิชานั้น ๆ เป็นผู้รับผิดชอบในการแก้ไข

ผู้ที่อ่านข้อสอบร่าง 4 หมายถึง ประธานชุดวิชา ทำหน้าที่ตรวจสอบภาพรวมของแบบทดสอบและตรวจสอบความเรียบร้อยของแบบทดสอบว่าเป็นไปตามที่อาจารย์ผู้ออกข้อสอบแต่ละคนกำหนดไว้หรือไม่ ทั้งนี้เพราะในบางชุดวิชา รูปแบบของการถามอาจอยู่ในรูปศัพท์เฉพาะ ตาราง สูตรในการคำนวณ ซึ่งจำเป็นต้องให้ประธานชุดวิชาเป็นผู้พิจารณาอีกครั้ง หากมีประเด็นแก้ไขที่พบในขั้นตอนนี้ ผู้ที่อ่านข้อสอบร่าง 4 จะระบุรายละเอียดของการแก้ไขเพื่อแจ้งและดำเนินการส่งให้ผู้พิมพ์แบบทดสอบที่ได้รับมอบหมายให้พิมพ์แบบทดสอบในชุดวิชานั้น ๆ เป็นผู้รับผิดชอบในการแก้ไข

ผู้พิมพ์แบบทดสอบ หมายถึง ผู้ที่ได้รับมอบหมายให้มีหน้าที่ในการจัดพิมพ์แบบทดสอบในชุดวิชานั้น ๆ โดยผู้พิมพ์แบบทดสอบจะรับข้อสอบที่ประธานชุดวิชาส่งมาโดยจะอยู่ในรูปแบบของการออกข้อสอบที่เขียนด้วยลายมือหรือพิมพ์ลงในการ์ดข้อสอบ หรือไฟล์ข้อสอบ ไม่ว่าข้อสอบที่ได้รับมาจะอยู่ในรูปแบบใดผู้พิมพ์แบบทดสอบจะต้องทำหน้าที่ในการพิมพ์และนำเข้าไฟล์ข้อสอบทีละข้อ โดยแยกเป็นข้อคำถามและตัวเลือกลงไปในระบบการจัดพิมพ์แบบทดสอบของศูนย์วิจัยฯ และทำหน้าที่ในการแก้ไขข้อสอบตามผู้ที่อ่านข้อสอบร่างที่ 1-4 ระบุประเด็นแก้ไข

ตารางที่ 4.9 ค่าใช้จ่ายในกระบวนการจัดทำแบบทดสอบก่อนส่งเข้าโรงพิมพ์

ผู้ที่มีส่วนเกี่ยวข้องในกระบวนการจัดทำแบบทดสอบ	ค่าใช้จ่ายในราคาต่อหน้า (บาท)
ผู้พิมพ์แบบทดสอบ	15
อาจารย์นักวัดผล (Reviewer ข้อสอบออกใหม่)	6
ผู้ที่อ่านข้อสอบร่าง 1	8

ตารางที่ 4.9 (ต่อ) ค่าใช้จ่ายในกระบวนการจัดทำแบบทดสอบก่อนส่งเข้าโรงพิมพ์

ผู้ที่มีส่วนเกี่ยวข้องในกระบวนการจัดทำแบบทดสอบ	ค่าใช้จ่ายในราคาต่อหน้า (บาท)
ผู้อ่านข้อสอบร่าง 2	8
ผู้อ่านข้อสอบร่าง 3	8
ผู้อ่านข้อสอบร่าง 4	5
รวม	50

3.2 ค่าใช้จ่ายการจัดทำต้นฉบับแบบทดสอบ ปีการศึกษา 2562 ถึงปีการศึกษา 2564

ผลการวิเคราะห์ค่าใช้จ่ายในการจัดทำต้นฉบับแบบทดสอบที่ได้รับข้อมูลจากศูนย์วิจัยฯ เป็นการรวบรวมข้อมูลจากการดำเนินการจัดสอบในปีการศึกษา 2562-2564 มีรายละเอียดของจำนวนชุดวิชา และค่าใช้จ่ายที่เกิดขึ้น โดยในปีการศึกษา 2563 ภาคการศึกษาที่ 2 การสอบไล่-ซ่อม และภาคการศึกษาพิเศษ และปีการศึกษา 2564 ภาคการศึกษาที่ 1 การสอบไล่-ซ่อม อยู่ในช่วงสถานการณ์ที่ไม่ปกติ กล่าวคือเกิดสถานการณ์การแพร่ระบาดของโรคโควิด-19 ทำให้ไม่สามารถจัดสอบภาคสนามได้ตามปกติ จึงเกิดรูปแบบการสอบออนไลน์เกิดขึ้น

ดังนั้น ในปีการศึกษา 2563 ภาคการศึกษาที่ 1 ข้อสอบที่จัดเตรียมไว้ทั้งหมดในรูปแบบกระดาษ จึงถูกนำมาใช้ในปีการศึกษา 2564 ภาคการศึกษาที่ 1 จะเห็นได้ว่ารายการที่มีจำนวนน้อยกว่าปีอื่น ๆ นั้น เป็นรายการค่าใช้จ่ายที่เกิดขึ้นเฉพาะกระบวนการจัดทำแบบสอบชุดใหม่ที่เป็นชุดวิชาที่เปิดใหม่ที่มีตารางสอบนอกเหนือจากข้อสอบเดิมที่ได้จัดทำไว้ในปีการศึกษา 2563 ภาคการศึกษาที่ 1

สำหรับรายการที่เกิดขึ้นในปีการศึกษา 2563 ภาคการศึกษาที่ 2 เป็นรายการค่าใช้จ่ายที่เกิดขึ้นเฉพาะกระบวนการจัดทำแบบสอบชุดใหม่ที่เป็นชุดวิชาที่เปิดใหม่ที่มีตารางสอบนอกเหนือจากข้อสอบเดิมที่ได้จัดทำไว้ในปีการศึกษา 2561 ภาคการศึกษาที่ 2 ค่าใช้จ่ายที่เกิดขึ้นเป็นค่าใช้จ่ายของกระบวนการจัดทำข้อสอบที่ออกใหม่ในครั้งที่ 1 และ 2 เท่านั้น

รายละเอียดของค่าใช้จ่ายในการจัดทำต้นฉบับแบบทดสอบปีการศึกษา 2562-2564 แสดงไว้ดังตารางที่ 4.10

ตารางที่ 4.10 ค่าใช้จ่ายในการจัดทำต้นฉบับแบบทดสอบปีการศึกษา 2562-2564

ข้อมูล	ปีการศึกษา 2562		ปีการศึกษา 2563		ปีการศึกษา 2564	
	ภาคการศึกษาที่ 1	ภาคการศึกษาที่ 2	ภาคการศึกษาที่ 1	ภาคการศึกษาที่ 2	ภาคการศึกษาที่ 1	ภาคการศึกษาที่ 2
	สอบไล่-ซ่อม และ ซ่อมแบบออนไลน์	สอบไล่-ซ่อมและ ภาคพิเศษ	สอบไล่-ซ่อม	สอบไล่-ซ่อมและ ภาคพิเศษ	สอบไล่-ซ่อม	สอบไล่-ซ่อม
จำนวนข้อสอบ ออกใหม่ (ชุดวิชา)	150	159	167	163	167	185
จำนวนข้อสอบ ปรับปรุง (ชุดวิชา)	98	95	81	97	81	90
จำนวนข้อสอบเก่า (ชุดวิชา)	114	119	108	113	108	109
จำนวนข้อสอบสุ่ม ห้องคลัง (ชุดวิชา)	70	88	88	94	88	87
จำนวนข้อสอบCopy กับชุดอื่น (ชุดวิชา)	16 ชุดวิชา (ไม่มีค่าตอบแทน)	12 (ไม่มีค่าตอบแทน)	14 (ไม่มีค่าตอบแทน)	11 (ไม่มีค่าตอบแทน)	14 (ไม่มีค่าตอบแทน)	10 (ไม่มีค่าตอบแทน)
รวมค่าตอบแทน ผู้พิมพ์ ผู้ตรวจร่าง (บาท)	722,839	633,654	455,532	172,987	70,913	697,600

3.3 ผลของการจำลองตัดแบบทดสอบจาก 5 ตัวเลือกให้เหลือเพียง 4 ตัวเลือก เพื่อเปรียบเทียบค่าใช้จ่ายในการจัดทำแบบทดสอบ

จากข้อมูลที่ปรากฏในตอนที 3 ข้อที่ 3.1-3.2 เป็นข้อมูลที่ได้มาจากการวิเคราะห์ด้วยแบบทดสอบที่มีจำนวน 5 ตัวเลือก ในที่นี้ผู้วิจัยได้นำข้อมูลแบบทดสอบจำนวน 5 ตัวเลือกที่ได้ตัดตัวลงที่ไม่มีคุณภาพออกไปจำนวน 1 ตัวเลือก ตามเกณฑ์การตัดข้อสอบที่ผู้วิจัยกำหนดขึ้นทั้ง 3 เกณฑ์ และด้วยข้อจำกัดของการเข้าถึงข้อสอบซึ่งเป็นเรื่องของความลับในการปฏิบัติงานด้านการสอบของมหาวิทยาลัย สุโขทัยธรรมาธิราช ผู้วิจัยจึงต้องขออนุญาตจากหัวหน้าหน่วยคลังข้อสอบในการเข้าถึงคลังข้อสอบ และเข้าถึงข้อสอบจำนวน 30 ชุดวิชา ซึ่งครอบคลุมในทุกชุดวิชาที่เปิดสอบ และเป็นไปตามธรรมชาติของข้อสอบของมหาวิทยาลัยสุโขทัยธรรมาธิราช ที่ในข้อสอบจะมีทั้งตาราง หรือรูปภาพประกอบอยู่ในโจทย์ และเป็นข้อสอบที่มีลักษณะโจทย์ และโจทย์ร่วมผสมผสานอยู่ในฉบับเดียวกันพบว่า ข้อสอบแบบ 5 ตัวเลือกที่มีใช้อยู่ในปัจจุบันนั้น จะมีจำนวนหน้าของข้อสอบอยู่ระหว่าง 13-46 หน้า ทั้งนี้ในรูปแบบของการพิมพ์หรือจัดหน้ากระดาษของ มหาวิทยาลัยสุโขทัยธรรมาธิราช ใน 1 หน้ากระดาษ จะไม่สามารถใช้มาเป็นตัวกำหนดจำนวนข้อโดยพิจารณาจากจำนวนบรรทัดในหนึ่งหน้ากระดาษได้ เนื่องจากโจทย์บางข้อจะมีความยาวของข้อคำถามมากกว่า 1 บรรทัด และมีตัวเลือกที่ยาวมากกว่า 1 บรรทัดเช่นกัน ในกรณีที่ตัวเลือกไม่ยาว จะใช้แบบฟอร์มในการจัดหน้ากระดาษของตัวเลือกเป็น 2 คอลัมน์ แต่หากตัวเลือกยาว จะใช้การจัดแบบฟอร์มของหน้ากระดาษเป็น 1 คอลัมน์ ทั้งนี้ใน 1 หน้ากระดาษ จะพบรูปแบบการจัดพิมพ์ทั้งสองรูปแบบ คือ มีตัวเลือกแบบ 1 คอลัมน์ และ 2 คอลัมน์ ผสมผสานกันในแต่ละหน้า และในบางชุดวิชาที่มีรูปหรือตารางประกอบจะเข้ากรณีของการจัดแบบฟอร์มในลักษณะที่โจทย์กับตัวเลือกต้องอยู่ในหน้าเดียวกัน ดังนั้นจึงไม่สามารถคำนวณจากการนับบรรทัดที่มีในหนึ่งหน้ากระดาษว่า หากลดจำนวนตัวเลือกลง 1 ตัวเลือกในแต่ละข้อ จะทำให้จำนวนบรรทัดลดลงเท่าใด

จากการสุ่มข้อสอบออกมา 30 ชุดวิชาที่ครอบคลุมทุกสาขาวิชาที่เปิดสอบ พบว่า ข้อสอบแบบ 5 ตัวเลือกที่มีใช้ในปัจจุบันนั้น จะมีจำนวนหน้าของข้อสอบอยู่ระหว่าง 13-46 หน้า คิดเป็นจำนวนหน้าเฉลี่ยอยู่ที่ 26 หน้าต่อหนึ่งชุดวิชา ดังนั้นเมื่อผู้วิจัยได้ทำการทดลองตัดตัวเลือกที่ไม่มีคุณภาพในข้อสอบแต่ละข้อออกไปจำนวน 1 ตัวเลือก จะส่งผลให้ข้อสอบแต่ละข้อเหลือตัวเลือกที่มีคุณภาพจำนวน 4 ตัวเลือกพบว่า จำนวนหน้าของข้อสอบอยู่ระหว่าง 11-40 หน้า สามารถลดจำนวนหน้าของข้อสอบลงได้เป็นจำนวน 2-8 หน้า โดยคิดเป็นจำนวนหน้าที่ลดลงเฉลี่ยอยู่ที่ 3 หน้า ค่าใช้จ่ายที่เกิดขึ้นในการจัดทำแบบทดสอบ 4 ตัวเลือกจะลดลงโดยเฉลี่ยอยู่ 11.54% ของค่าใช้จ่ายทั้งหมดจากการจัดทำแบบทดสอบที่ 5 ตัวเลือก โดยนำเสนอผลการเปรียบเทียบค่าใช้จ่ายในการจัดทำแบบทดสอบแบบ 5 ตัวเลือกและ 4 ตัวเลือก ในแต่ละภาคการศึกษาและภาพรวมของค่าใช้จ่ายที่ลดลงตลอดทั้งปีการศึกษาดังภาพ ดังตารางที่ 4.11

ตารางที่ 4.11 การเปรียบเทียบค่าใช้จ่ายในการจัดทำแบบทดสอบแบบ 5 ตัวเลือกและ 4 ตัวเลือก ในแต่ละภาคการศึกษา

ข้อมูล	ปีการศึกษา 2562		ปีการศึกษา 2563		ปีการศึกษา 2564	
	ภาคการศึกษาที่ 1	ภาคการศึกษาที่ 2	ภาคการศึกษาที่ 1	ภาคการศึกษาที่ 2	ภาคการศึกษาที่ 1	ภาคการศึกษาที่ 2
	สอบไล่-ซ่อม และ ซ่อมแบบออนไลน์	สอบไล่-ซ่อมและ ภาคพิเศษ	สอบไล่-ซ่อม	สอบไล่-ซ่อมและ ภาคพิเศษ	สอบไล่-ซ่อม	สอบไล่-ซ่อม
กรณี แบบทดสอบ 5 ตัวเลือก รวมค่าตอบแทน ผู้พิมพ์ ผู้ตรวจร่าง (บาท)	722,839	633,654	455,532	172,987	70,913	697,600
กรณี แบบทดสอบ 4 ตัวเลือก รวมค่าตอบแทน ผู้พิมพ์ ผู้ตรวจร่าง (บาท)	639,423	560,530	402,963	153,024	62,729	617,096
ค่าใช้จ่ายที่ลดลงใน แต่ละภาคการศึกษา	83,416	73,124	52,569	19,963	8,184	80,504
ค่าใช้จ่ายที่ลดลงใน แต่ละปีการศึกษา	156,540		72,532		88,688	

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยเรื่อง “การศึกษาเปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกและ 4 ตัวเลือก” มีวัตถุประสงค์เพื่อ 1) เพื่อเปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก 2) เพื่อเปรียบเทียบคะแนนและเปรียบเทียบคะแนนของนักศึกษาที่ได้จากการเปรียบเทียบคะแนนจากคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก และ 3) เพื่อเปรียบเทียบค่าใช้จ่ายในการจัดทำแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก มีการดำเนินการวิจัยโดยใช้ข้อมูลผลการตอบแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ของนักศึกษาที่กำลังศึกษาในหลักสูตรระดับปริญญาตรีของมหาวิทยาลัยสุโขทัยธรรมมาธิราช ที่มีสิทธิ์สอบในภาคการศึกษาที่ 2/2562 ในรายวิชาที่ถูกเลือกมาศึกษาในงานวิจัยนี้เป็นข้อมูลตั้งต้น วิเคราะห์ข้อมูลเพื่อคำนวณหาคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก ด้วยการคำนวณหาค่าอำนาจจำแนกและวิเคราะห์ผลการตอบด้วยทฤษฎีการตอบสนองข้อสอบด้วยการทำให้เรียบเคอเนล เพื่อตัดตัวเลือกที่ไม่มีประสิทธิภาพออก 1 ตัวเลือก สำหรับการจัดแบบทดสอบฉบับใหม่ที่เหลือตัวเลือกเพียง 4 ตัวเลือก โดยกำหนดกรณีของความเป็นไปได้ในการตัดตัวเลือกออกเป็น 3 กรณี คือ กรณีที่ 1 แบบทดสอบแบบเลือกตอบ 5 ตัวเลือก กรณีที่ 2 แบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) กรณีที่ 3 แบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) และกรณีที่ 4 แบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) เพื่อศึกษาคุณภาพของแบบทดสอบ 4 ตัวเลือกและทำการเปรียบเทียบคะแนนเพื่อตอบสมมติฐานที่ผู้วิจัยตั้งไว้ แล้วนำผลการวิจัยที่ได้ทั้งหมดมาคำนวณค่าใช้จ่ายหลังจากที่มีการลดจำนวนตัวเลือกของแบบทดสอบ

สรุปผลการวิจัย

1. ผลการเปรียบเทียบคุณภาพของแบบทดสอบจากการวิเคราะห์ค่าความเที่ยง ความยาก และอำนาจจำแนก แยกเป็น 4 กรณีตามสมมติฐานการวิจัย

กรณีที่ 1 แบบทดสอบแบบเลือกตอบ 5 ตัวเลือก พบว่าค่าความเที่ยง (raw_alpha) ของแบบทดสอบแบบ 5 ตัวเลือกมีค่าเท่ากับ 0.81

กรณีที่ 2 แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) พบว่าค่าความเที่ยง (raw_alpha) ของแบบทดสอบแบบ 4 ตัวเลือกมีค่าเท่ากับ 0.79

กรณีที่ 3 แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (แทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) พบว่าค่าความเที่ยง (raw_alpha) ของแบบทดสอบแบบ 4 ตัวเลือกมีค่าเท่ากับ 0.79

กรณีที่ 4 แบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (แทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) พบว่าค่าความเที่ยง (raw_alpha) ของแบบทดสอบแบบ 4 ตัวเลือกมีค่าเท่ากับ 0.87

โดยในกรณีที่ 4 ค่าความเที่ยงของแบบทดสอบมีค่าสูงที่สุดในบรรดากรณีอื่นๆ ส่วนค่าความยากกับอำนาจจำแนกไม่แตกต่างกันในทุกกรณี

2. ผลการเปรียบเทียบคะแนนสอบของนักศึกษาระหว่างคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกและแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (3 กรณี ที่ผู้วิจัยศึกษา) พบว่า

2.1 ผลการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยไม่รวมตัวเลือกที่เป็นเฉลย) ได้สมการเชิงเส้น $y = 0 + 1x$

2.2 ผลการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกอย่างสุ่ม ด้วยตัวเลือกที่เหลือตัวอื่น ๆ โดยรวมตัวเลือกที่เป็นเฉลย) ได้สมการเชิงเส้น $y = 3.999 + 0.969x$

2.3 ผลการเปรียบเทียบคะแนนของนักศึกษาระหว่างคะแนนที่นักศึกษาสอบได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบ 4 ตัวเลือก (กรณีแทนค่าตัวเลือกที่ถูกตัดออกตามระดับความสามารถของผู้สอบ) ได้สมการเชิงเส้น $y = -9.722 + 1.222x$

3. ผลการเปรียบเทียบคะแนนเฉลี่ยของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และแบบทดสอบแบบเลือกตอบ 4 ตัวเลือกในแต่ละกรณี (จากทั้งหมด 3 กรณี ที่ผู้วิจัยศึกษา) พบว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก (ในทุกกรณี) มากกว่าคะแนนของนักศึกษาที่ได้จากแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก

4. ผลการวิเคราะห์ค่าใช้จ่ายในการจัดทำแบบทดสอบ พบว่า มีผู้ที่มีส่วนเกี่ยวข้องในกระบวนการจัดทำแบบทดสอบ ประกอบด้วย ผู้พิมพ์แบบทดสอบกับผู้ตรวจร่างแบบทดสอบ โดยมีค่าตอบแทนให้กับผู้ทำงานในแต่ละตำแหน่ง ดังนี้ ผู้พิมพ์แบบทดสอบ ได้รับค่าตอบแทนในการพิมพ์แบบทดสอบหน้าละ 15 บาท อาจารย์นักวัดผล (Reviewer ข้อสอบออกใหม่) ได้รับค่าตอบแทนในการอ่านข้อสอบชุดวิชาที่ออกใหม่ หน้าละ 6 บาท ผู้ที่อ่านข้อสอบร่าง 1 ร่าง 2 และ ร่าง 3 ได้รับค่าตอบแทนในการอ่านข้อสอบ หน้าละ 8 บาท และผู้ที่อ่านข้อสอบร่าง 4 ได้รับค่าตอบแทนในการอ่านข้อสอบ หน้าละ 5 บาท รวมค่าใช้จ่ายในการดำเนินการจัดพิมพ์ข้อสอบจำนวน 1 หน้า เท่ากับ 50 บาท และเมื่อจำลองข้อมูลในการวิเคราะห์ค่าใช้จ่ายของแบบทดสอบที่มี 4 ตัวเลือกเปรียบเทียบกับค่าใช้จ่ายของแบบทดสอบที่มี 5 ตัวเลือก พบว่าการลดจำนวนตัวเลือกในแบบทดสอบลง 1 ตัวเลือก ด้วยการตัดตัวเลือกที่ไม่มีคุณภาพออกไปตามสมมติฐานการวิจัยทั้ง 3 กรณี สามารถลดจำนวนหน้าของข้อสอบลงได้เป็นจำนวน 2-8 หน้า โดยคิดเป็น

จำนวนหน้าที่ลดลงเหลืออยู่ที่ 3 หน้า ช่วยประหยัดค่าใช้จ่ายโดยเฉลี่ยอยู่ 11.54% ของค่าใช้จ่ายทั้งหมดจากการจัดทำแบบทดสอบที่ 5 ตัวเลือก

อภิปรายผลการวิจัย

การอภิปรายผลการวิจัยในเบื้องต้น แบ่งออกเป็น 2 ประเด็น ดังนี้ 1) คุณภาพของแบบทดสอบ 5 ตัวเลือกและ 4 ตัวเลือก 2) ผลกระทบจากการลดจำนวนตัวเลือกในแบบทดสอบ รายละเอียดดังนี้

คุณภาพของแบบทดสอบ 5 ตัวเลือกและ 4 ตัวเลือก

ในกระบวนการจัดทำแบบทดสอบ คุณภาพของแบบทดสอบควรเริ่มตั้งแต่การกำหนดวัตถุประสงค์ของการวัด การออกแบบผังข้อสอบ และการออกข้อสอบให้มีความตรงตามเนื้อหา (content validity) ซึ่งการกำหนดว่าข้อสอบที่จะใช้นั้น ควรมีจำนวนตัวเลือกเท่าไร ขึ้นอยู่กับว่ากลุ่มผู้ตอบเป็นใคร มีช่วงอายุอยู่ในระดับใด หนึ่งในที่ใช้ในการพิจารณาสำหรับการกำหนดจำนวนตัวเลือกที่เหมาะสมคือ จำนวนตัวเลือกควรมี 4 หรือ 5 ตัวเลือก เพื่อการโอกาสในการเดาของผู้ตอบ แต่หากผู้ตอบเป็นเด็กเล็ก จำนวนตัวเลือกควรมีเพียง 2-3 ตัวเลือกเท่านั้นเพื่อป้องกันความสับสนและมีความเหมาะสมกับวัย ซึ่งสอดคล้องกับที่ Haladyna, Downing, and Rodriguez (2002) เสนอว่า โดยทั่วไปแล้วแนะนำให้ใช้สี่หรือห้าตัวเลือกต่อข้อคำถามหนึ่งรายการ สำหรับแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนในห้องเรียนส่วนใหญ่และแบบทดสอบมาตรฐานสากล (เช่น TOEFL) มักจะใช้ตามกฎสี่ตัวเลือกต่อหนึ่งข้อ แม้จะมีการใช้ตัวเลือกสี่หรือห้าตัวเลือกอย่างแพร่หลายโดยผู้เขียนและผู้พัฒนาแบบทดสอบจำนวนมาก การเขียนตัวเลือกควรเขียนสิ่งที่เบี่ยงเบนความสนใจที่เป็นไปได้มากที่สุด (Aamodt & McShane, 1992; Crehan, Haladyna & Brewer, 1993; Delgado & Prieto, 1998; Haladyna & Downing, 1993; Landrum, Cashin & Theis, 1993; Rodriguez, 2005; Shizuka, Takeuchi, Yashima, & Yoshizawa, 2006; Straton & Catts, 1980)

ในการสร้างแบบทดสอบแบบเลือกตอบ การสร้างตัวเลือกเป็นสิ่งสำคัญที่สุดอย่างหนึ่ง เนื่องจากคุณภาพของข้อสอบชนิดนี้จะมากหรือน้อยขึ้นอยู่กับการเขียนตัวเลือกเป็นประการสำคัญ การเขียนตัวเลือกที่ประกอบด้วยตัวถูก และตัวลวงนั้น ตัวลวงที่ดีจะช่วยให้ข้อสอบมีคุณภาพ (Adams, 1964: 340) ถ้าตัวลวงไม่สามารถจะลวงผู้สอบได้แล้ว นอกจากจะทำให้เสียเวลาในการออกข้อสอบแล้วยังทำให้ความตรงและความเที่ยงของแบบทดสอบลดลงด้วย (Marshall and Halses, 1972: 62) ดังนั้นในการสร้างข้อสอบชนิดนี้จำเป็นต้องใช้ตัวลวงให้ใกล้เคียงกับคำตอบมากที่สุด เพื่อให้ผู้ตอบที่ไม่มีความรู้จริงเลือกตอบตัวลวงเหล่านี้กระจายออกไปทุกตัว สอดคล้องกับ อาณาภพ เลขะกุล (2559: 3) ที่กล่าวว่าแบบทดสอบแบบเลือกตอบมีส่วนประกอบที่สำคัญ ได้แก่ โจทย์ (Stem) คำถาม (Lead – in) และตัวเลือก (Option หรือ Alternative) ซึ่งทั้งโจทย์ คำถาม และตัวเลือกอาจมี รูปภาพ กราฟหรือแผนภูมิประกอบ โดยทั่วไปอาจมี 4 หรือ 5 ตัวเลือก แต่การมีจำนวนตัวลวงน้อยเกินไปจะทำให้ผู้ตอบมีโอกาสเดาได้ถูกมากขึ้น ตัวเลือก 5 ข้อ จะมีความเที่ยง (Reliability) มากที่สุด อย่างไรก็ตามในบางกรณี การหา

หรือการสร้างตัวเลือกที่ 5 ในการออกข้อสอบเป็นเรื่องที่ยากมากหรือไม่มีโอกาสเป็นไปได้ ก็อาจใช้เพียง 4 ตัวเลือกในการสอบได้ ส่วนการใช้ตัวเลือกมากกว่า 5 ข้อ ก็ไม่มีความจำเป็นและไม่ได้ทำให้ค่าอำนาจจำแนกสูงขึ้น สอดคล้องกับงานวิจัยในต่างประเทศที่มีการศึกษาเกี่ยวกับจำนวนตัวเลือกที่เหมาะสมในข้อสอบ สอดคล้องกับงานวิจัยของชูศักดิ์ ศิริรังษ (2559) ที่พบว่า การเขียนตัวเลือกส่วนใหญ่จะมีปัญหาเกี่ยวกับการเขียนตัวเลือกไม่ครบตามจำนวนที่ต้องการ หรือเขียนตัวเลือกได้ไม่ดีพอ ในบางครั้งเกิดความผิดพลาดทำให้ข้อสอบไม่มีตัวถูก หรือมีตัวถูกมากกว่าหนึ่งตัว

นอกจากนี้การพิจารณาปรับลดจำนวนตัวเลือกที่มีจำนวนมากเกินไปหรือการพิจารณาปรับลดตัวเลือกที่ไม่มีคุณภาพออกไปจะส่งผลให้ข้อสอบข้อนั้นมีคุณภาพที่สูงขึ้น สอดคล้องกับงานวิจัยของ ศศิธร ชูตินันท์กุล, ไพรัตน์ วงษ์นาม และสมพงษ์ บัณฑิต (2561) ที่อธิบายว่า การตัดข้อสอบที่ไม่มีคุณภาพทิ้งก่อนเปรียบเทียบคะแนนจะมีคุณภาพของการเปรียบเทียบคะแนนมากกว่าไม่ตัดข้อสอบที่ไม่มีคุณภาพทิ้ง หมายความว่า สามารถใช้แบบทดสอบแบบเลือกตอบ 4 ตัวเลือกในการวัดและประเมินผลนักศึกษาแทนแบบทดสอบแบบเลือกตอบ 5 ตัวเลือกก็ได้ โดยไม่จำเป็นต้องใช้แบบทดสอบแบบเลือกตอบ 5 ตัวเลือกในการวัดและประเมินผลนักศึกษา อีกทั้งยังไม่จำเป็นต้องลงแรง และเสียเวลาในการสร้างและผลิตแบบทดสอบ 5 ตัวเลือก ให้ได้คุณภาพและมาตรฐาน

จากประเด็นที่กล่าวไปข้างต้น การกำหนดจำนวนตัวเลือกที่เหมาะสมหรือการพิจารณาลดจำนวนตัวเลือกจาก 5 ตัวเลือกให้เหลือเพียง 4 ตามที่แบบทดสอบมาตรฐานนิยมใช้กัน เป็นอีกหนึ่งประเด็นที่ทางผู้วิจัยตั้งใจนำเสนอให้ทางมหาวิทยาลัยมีสารสนเทศในการนำไปพิจารณาเพื่อทบทวนมีแนวทางในการกำหนดจำนวนตัวเลือกของแบบทดสอบของมหาวิทยาลัยสุโขทัยธรรมาธิราชต่อไป

ผลกระทบจากการลดจำนวนตัวเลือกในแบบทดสอบ

มหาวิทยาลัยสุโขทัยธรรมาธิราชได้กำหนดรูปแบบการสอบปลายภาคโดยกำหนดเวลาในการสอบไว้ที่ 180 นาที รูปแบบของข้อสอบจะเป็นข้อสอบแบบเลือกตอบที่มี 5 ตัวเลือก มีจำนวนของข้อคำถามอยู่ระหว่าง 90-120 ข้อ ขึ้นอยู่กับลักษณะของชุดวิชา ในการสอบปลายภาคมีจำนวนข้อสอบในแต่ละชุดวิชาที่นักศึกษาต้องทำจำนวนมากข้อ ประกอบกับจำนวนตัวเลือกในแต่ละข้อมี 5 ตัวเลือก ทั้งความยาวของจำนวนข้อสอบและความมากของจำนวนตัวเลือกอาจเป็นเหตุผลสำคัญประการหนึ่งที่ทำให้นักศึกษาสอบไม่ผ่านกันเป็นจำนวนมาก ส่งผลให้เกิดภาวะของการหมดไฟ (burn out) ในการศึกษาต่อให้สำเร็จ การศึกษา ทั้งนี้อาจเป็นเพราะนักศึกษาสอบตกในชุดวิชาเดิมอยู่บ่อยครั้งจนเกิดความท้อแท้ ประเด็นดังกล่าวสอดคล้องกับงานวิจัยของในประเทศอเมริกาเหนือ พบว่า นักศึกษาตัดสินใจลาพักการศึกษาหรือตัดสินใจออกจากชั้น เพราะนักศึกษาแต่ละคนมีความสามารถในทางวิชาการที่ไม่เท่ากัน โดยเฉพาะอย่างยิ่งความสามารถในการอ่านและการเขียน (Aydin, Ozturk, Buyukkose, Era, & Sonmez , 2019; Budiman, 2018; Lin, 2015; Paepe, Zhu, & DePryck, 2016) สำหรับในประเทศไทยพบว่า สาเหตุที่นักศึกษาออกกลางคันมีหลายปัจจัยที่ตรงกับนักศึกษาต่างประเทศ เช่น มีแรงจูงใจในการเรียนต่ำ ไม่มีความรู้ความสามารถในสาขาวิชาที่จะเรียน มีปัญหาในเรื่องทักษะการอ่านและเขียน การสอบตก (Aydin,

Ozturk, Buyukkose, Era, & Sonmez , 2019; Budiman, 2018; Lin, 2015; Paepe, Zhu, & DePryck, 2016; ชนิดาภา บุญประสมและ จริญญา แสนราช; 2562; บุญศรี พรหมมาพันธุ์ และคณะ, 2554; เพียว ดีใจ และคณะ, 2557) ผลการเรียนรู้ต่ำ ขาดการวางแผนการเรียนรู้ที่ดี (Aydin, Ozturk, Buyukkose, Era, & Sonmez , 2019; Budiman, 2018; จารุณี มงคลแก้ว และคณะ, 2553; ชนิดาภา บุญประสม และจริญญา แสนราช; 2562; เพียว ดีใจ และคณะ, 2555) ล้วนเป็นสาเหตุที่ทำให้นักศึกษาตัดสินใจออกกลางคัน ทั้งสิ้น ประเด็นสำคัญประการหนึ่งที่จะช่วยให้นักศึกษายังคงอยู่ระบบการศึกษาและสำเร็จการศึกษาได้ตามระยะเวลาของหลักสูตร คือ การให้ความสำคัญกับกระบวนการวัดและประเมินผลการศึกษา โดยเฉพาะเรื่องรูปแบบทดสอบที่เป็นเครื่องมือสำคัญที่มหาวิทยาลัยสุโขทัยธรรมมาธิราชใช้เป็นเครื่องมือหลักในการตัดสินผลการเรียนของผู้เรียน ทั้งนี้การพิจารณาในประเด็นดังกล่าวอาจย้อนกลับเข้ามาสู่คุณภาพของแบบทดสอบ ด้วยการทบทวนและพิจารณาปรับลดตัวเลือกในแบบทดสอบ โดยคงไว้เฉพาะตัวเลือกที่มีคุณภาพและมีจำนวนตัวเลือกที่พอเหมาะกับการทำหาคำถามความสามารถของผู้เรียน โดยจำนวนตัวเลือกที่คงไว้จะต้องมีคุณสมบัติทางจิตมิติครบ เป็นไปตามเกณฑ์ของคุณภาพข้อสอบ และสามารถจำแนกผู้เรียนที่รอบรู้และไม่รอบรู้ออกจากกันได้ งานวิจัยหลายเรื่องได้ข้อสรุปตรงกันว่า การมีตัวเลือกเพียง 3-4 ตัวเลือกนั้น เพียงพอต่อการจำแนกผู้เรียนและมีคุณสมบัติทางจิตมิติของคุณภาพข้อสอบอย่างครบถ้วนไม่ต่างไปจากข้อสอบที่มีจำนวนตัวเลือกตั้งแต่ 5 ขึ้นไป เมื่อนำแบบทดสอบที่มี 3 4 และ 5 ตัวเลือกมาวิเคราะห์เปรียบเทียบคุณภาพกัน (Esmaeeli, Shandiz, Norooziasl, Shojaei, Pasandideh, Khoshkholgh, & Ahmadi; (2021), Nguyen, Bui, Fujita, Hong, Loc, Snasel, & Vo; 2021) ดังนั้น การพิจารณาเลือกข้อสอบ 4 ตัวเลือกจึงเป็นอีกทางเลือกหนึ่งที่น่าสนใจทั้งในแง่ของการใช้แบบทดสอบมีจำนวนตัวเลือกน้อยลงจะช่วยสร้างแรงจูงใจในการทำแบบทดสอบของนักศึกษา นักศึกษาจะไม่เกิดความเหนื่อยล้าในการทำแบบทดสอบ และผลคะแนนของนักศึกษาที่ได้ก็ยิ่งสูงกว่าแบบทดสอบที่มีตัวเลือกเป็นจำนวนมาก ตลอดจนช่วยลดต้นทุนเพื่อจัดซื้อปัจจัยในการผลิต เช่น กระดาษ หมึกพิมพ์ ค่าไฟฟ้า ค่าตอบแทนของผู้ที่มีส่วนเกี่ยวข้องในกระบวนการจัดทำแบบสอบ (ศศิธร ชูตินันท์กุล, ไพรัตน์ วงษ์นาม และสมพงษ์ ปันหุ่น; 2561) ดังนั้น การพิจารณาปรับลดจำนวนตัวเลือกให้เหลือเพียง 4 ตัวเลือกจากของเดิมที่มี 5 ตัวเลือก จึงเป็นแนวทางในการควบคุมคุณภาพของแบบทดสอบให้คงไว้ซึ่งคุณสมบัติทางจิตมิติของแบบทดสอบที่ครบถ้วน และเป็นอีกแนวทางที่ช่วยลดความเหนื่อยล้าและความเครียดจากการทำแบบทดสอบของผู้สอบ มีโอกาสให้นักศึกษาสามารถทำคะแนนสอบได้สูงขึ้นเพราะไม่ต้องเสียเวลาไปกับการพิจารณาตัดตัวเลือกที่มีจำนวนมากเกินไปออกในแต่ละข้อ นักศึกษาจะได้มีเวลาในการคิดหาคำตอบของข้อคำถามแต่ละข้อได้อย่างผ่อนคลายมากขึ้น และการลดจำนวนตัวเลือกลงจะช่วยประหยัดทรัพยากรในการจัดทำแบบทดสอบลงได้อีกด้วย

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. สถาบันอุดมศึกษาหรือหน่วยงานที่เกี่ยวข้องสามารถนำผลการวิจัยไปศึกษาเพื่อเป็นแนวทางในการกำหนดรูปแบบการวัดและประเมินผลกรณีที่ต้องการพิจารณาปรับลดจำนวนตัวเลือกในข้อสอบแบบเลือกตอบที่เหมาะสมสำหรับกลุ่มผู้สอบได้
2. ผลการวิจัยที่ได้จากงานวิจัยเรื่องนี้เป็นสารสนเทศส่วนหนึ่งที่เป็นทางเลือกให้กับมหาวิทยาลัยสุโขทัยธรรมาธิราชได้พิจารณาเพื่อเป็นแนวทางในการตัดสินใจเลือกปรับรูปแบบของการสอบจากเดิมที่มี 5 ตัวเลือก เป็น 4 ตัวเลือก โดยใช้ข้อมูลสนับสนุนทางวิชาการที่ผ่านกระบวนการวิจัย ทั้งนี้ ผลการวิจัยที่ได้มาจากการวิเคราะห์ข้อมูลเพียง 1 ชุดวิชาเท่านั้น ข้อมูลดังกล่าวยังไม่มีความเป็นตัวแทนของทุกชุดวิชาที่เปิดสอนในมหาวิทยาลัยสุโขทัยธรรมาธิราช การนำผลการวิจัยไปใช้ในการอ้างอิงจึงควรพิจารณาในกรณีนี้ด้วย

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. การวิจัยครั้งต่อไปควรขยายผลไปที่การปรับลดจำนวนข้อสอบเพื่อหาจำนวนข้อสอบที่เหมาะสมในการใช้เป็นตัวแทนเพื่อวัดความรู้ในเนื้อหาแต่ละเรื่อง
2. ควรทำการวิจัยเช่นเดียวกันนี้ในชุดวิชาอื่น ๆ ให้ครบทุกสาขาวิชาที่เปิดสอนในมหาวิทยาลัยสุโขทัยธรรมาธิราช เพื่อให้ได้ผลการวิจัยที่มีความเป็นตัวแทนของทุกชุดวิชา

บรรณานุกรม

- โกวิท ประวาลพุกษ์ และสมศักดิ์ สันธะเวชญ์. (2523). *การประเมินในชั้นเรียน*. กรุงเทพฯ: วัฒนาพานิช.
- ชวาล แพรัตกุล. (2507). *เทคนิคการวัดผล*. กรุงเทพฯ: วัฒนาพานิช.
- ชูศักดิ์ ศิรินุพันธ์. (2559). เทคนิคการเขียนข้อสอบแบบเลือกตอบ. ใน ณรงค์ ทีปประชัย (บรรณาธิการ.), *คู่มือการวัดและประเมินผลในระบบการศึกษาทางไกล สำหรับหลักสูตรระดับปริญญาตรี มหาวิทยาลัยสุโขทัยธรรมาธิราช* (หน้า. 78-84). โรงพิมพ์มหาวิทยาลัยสุโขทัยธรรมาธิราช.
- ณัชชา มาลัย และทิพย์วัน เลิศชุมทรัพย์. (2550). *การศึกษาคุณภาพรายข้อของข้อสอบแบบเลือกตอบที่มีลักษณะการเขียนตัวเลือกแบบต่างกัน*. (รายงานผลการวิจัย) นนทบุรี: สถาบันวิจัยและพัฒนา มหาวิทยาลัยสุโขทัยธรรมาธิราช.
- บุญศรี พรหมมาพันธุ์ และคณะ. (2555). การลาออกกลางคัน การคงอยู่และการเพิ่มจำนวนนักศึกษาปริญญาตรีของ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร. คณะเทคโนโลยีคหกรรมศาสตร์ เปรียบเทียบคุณภาพของการเปรียบเทียบคะแนนสอบภาษาอังกฤษระหว่างวิธีอีคิวเปอร์เซ็นต์ไทล์ วิธีเชิงเส้นตรง และวิธีสมการถดถอย. *วารสารมหาวิทยาลัยศิลปากรฉบับภาษาไทย สาขาสังคมศาสตร์ มนุษยศาสตร์ และศิลปะ*. 10(2) : 2444 – 2455.
- พิศเพลิน เขียวหวาน และคณะ. (2538). *ความยากและความยาวของแบบทดสอบที่เหมาะสมกับเวลาในการสอบของนักศึกษามหาวิทยาลัยสุโขทัยธรรมาธิราช*. (รายงานผลการวิจัย) นนทบุรี: สถาบันวิจัยและพัฒนา มหาวิทยาลัยสุโขทัยธรรมาธิราช.
- ภาวณิ ศรีสุขวัฒนานันท์. (2529). *การเปรียบเทียบผลจากการใช้รูปแบบการเทียบมาตรฐานที่แตกต่างกันเมื่อแบบทดสอบพร้อมมีความยาวต่างกัน*. วิทยานิพนธ์ครุศาสตรดุษฎีบัณฑิต, จุฬาลงกรณ์มหาวิทยาลัย.
- ล้วน สายยศ และอังคณา สายยศ. (2539). *เทคนิคการวัดผลการเรียนรู้*. กรุงเทพฯ: ชมรมเด็ก.
- วารุณี ปิตธวัชชัย. (2514). *การศึกษาเปรียบเทียบค่าสถิติของแบบทดสอบชนิดเลือกตอบที่จำนวนตัวเลือกไม่เท่ากัน*. ปริญญาโท ค.ม. (วิจัยการศึกษา) กรุงเทพฯ: บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.
- ศศิธร ชูตินันทกุล, ไพรัตน์ วงษ์นาม และสมพงษ์ ปั่นหุ่น. (2561). ผลการปรับเทียบคะแนนด้วยวิธีเคอเนล และวิธี IRT ภายใต้เงื่อนไขที่แตกต่างกัน. *วารสารศึกษาศาสตร์ มสธ*, 11(1), 294-306.
- ศิริชัย กาญจนวาสี. (2541). *การปรับเทียบคะแนนระหว่างแบบสอบ (Test Equating)*. กรุงเทพฯ: ศูนย์ตำราและเอกสารทางวิชาการ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- _____. (2550). *ทฤษฎีการทดสอบแนวใหม่*. กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- _____. (2556). *ทฤษฎีการทดสอบแบบดั้งเดิม*. กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สร้อยญา จันทรชุกุล ไชยยศ ไพวิทยศิริธรรม ยური ผลพันธิน และพิทักษ์ สุพรรณโณภาพ. (2560). *การสำเร็จการศึกษาตามแผนของนักศึกษา มหาวิทยาลัยราชภัฏนครราชสีมา*. มหาวิทยาลัยราชภัฏนครราชสีมา.

- สุนิสา จั๋ยม่วงศรี. (2537). *ผลของความยาวของแบบสอบร่วมที่มีต่อคุณภาพของวิธีการเทียบมาตราเชิงเส้นตรง*. วิทยานิพนธ์ครุศาสตรมหาบัณฑิต กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย.
- อุทัยวรรณ พงศ์อร่าม.(2545).*การศึกษาขนาดกลุ่มตัวอย่างที่เหมาะสมสำหรับการเปรียบเทียบคะแนนด้วยวิธีอิกวิเปอร์เซ็นไทล์ และวิธีเชิงเส้นตรงตามแบบจำลองคะแนนจริงสัมพันธ์ ที่มีแบบแผนการเปรียบเทียบและความยาวของแบบสอบแตกต่างกัน*.(ปริญญาณิพนธ์ดุขฎิบัณฑิตไม่ได้ตีพิมพ์). มหาวิทยาลัยศรีนครินทรวิโรฒประสานมิตร, กรุงเทพมหานคร.
- อานุกาพ เลขะกุล. (2559). *การสร้างข้อสอบแบบเลือกตอบ*. สงขลา: มหาวิทยาลัยสงขลานครินทร์.
- Aamodt, M. G., & McShane, T. (1992). A meta-analytic investigation of the effect of various test item characteristics on test scores and test completion times. *Public Personnel Man-agement*, 21(2), 151-160.
- Angoff W.H. (1971). *Scales, Norms, and Equivalent Scores*. Educational Measurement. 2 nd _____. (1984). *Scales, Norms, and Equivalent Scores*, Princeton, NJ : Educational.
- Aydin, Ozturk, Buyukkose, Era, & Sonmez. (2019). An Investivation of Drop - out in Open and Distance Education. *EDUCATIONAL SCIENCES: THEORY & PRACTICE*. April 2019, 19(2), p40 – 57.
- Measurement. 14-16, 2016. p635 – 647.
- Baghaei, P., & Amrahi, N. (2011). The effects of the number of options on the psychometric characteristics of multiple choice items. *Psychological Test and Assessment Modeling*, 53(2), 192.
- Baron, Denis, and Bernard, Harold W. (1967). *Education Technique for classroom Teacher*, New York: Mc Graw – Hill Book Company.
- Budiman, R. (2018). Factors Related to Students' Drop out of a Distance Language Learning Programme. *Journal of Curriculum and Teaching*, 7(2), p12-19.
- Dorans,N. , Moses, T. and Eignor, D. (2010) . *Principles and practices of testscore equating*. [Online]. Available from : <https://www.ets.org/Media/Research/pdf/RR-10-29.pdf>. [2017, May 19].
- Esmaeeli, B., Shandiz, E. E., Norooziasl, S., Shojaei, H., Pasandideh, A., Khoshkholgh, R., ... & Ahmadi71, F. B. (2021). The Optimal Number of Choices in Multiple-Choice Tests: A Systematic Review. *Int J Pediatr*, 2(3).
- Guo, H., Zu, J., & Kyllonen, P. (2018). A SimulationBased Method for Finding the Optimal Number of Options for Multiple Choice Items on a Test. *ETS Research Report Series*, 2018(1), 1-17.
- Hambleton, R.K. and H. Swaminathan. (1985). *Item Response Theory*. Kluwer. Nijhoff, Boston.

- Lin, S. P. (2015). Using EDM for Developing EWS to Predict University Students Drop Out. *International Journal of Intelligent Technologies & Applied Statistics*, 8(4).
- McGuire, B. (2012). *KernSmoothIRT: An R Package allowing for Kernel Smoothing in Item Response Theory* (Doctoral dissertation, Montana State University).
- Mehrens, W.A. and I.J. Lehmann. (1969). *Standard Tests in Education*. New York: Holt, Rinchart and Winton.
- Nguyen, T., Bui, T., Fujita, H., Hong, T. P., Loc, H. D., Snasel, V., & Vo, B. (2021). Multiple-objective optimization applied in extracting multiple-choice tests. *Engineering Applications of Artificial Intelligence*, 105, 104439.
- Paepe, Zhu, & DePryck. (2016). Drop-out, Retention, Satisfaction and Attainment of Online Learners.
- Petersen, N.S., Marco, C.L., and Stewart, E.E. (1982). *A test of the adequacy of linear Score Problems*. Hillssdale, N.J. : Erlbaun.
- Ryan, J. and Brockmann, F. (2011). *A practitioner's introduction to equating with primers on classical test theory and item response theory. [Online]*. Available from : [http://www.ccsso.org/Documents/Equating%20Handbook CoverANDinterior.pdf](http://www.ccsso.org/Documents/Equating%20Handbook%20CoverANDinterior.pdf)2017.
- Thorndike, Robert L. and Hagen, E.P.Elizabeth. (1960). *Measurement and Evaluation in Psychology and Education*. 3rd ed. New York: John Willey.

ภาคผนวก

ภาคผนวก ก
หนังสือขอความอนุเคราะห์ข้อมูลวิจัย



บันทึกข้อความ

ส่วนราชการ ศูนย์วิชาการประเมินผล สำนักทะเบียนและวัดผล โทร 7261-3

ที่ สว 0602.02/คว 625

วันที่ 13 กันยายน 2565

เรื่อง ขอความอนุเคราะห์ข้อมูลเพื่อดำเนินการวิจัย

เรียน ผู้อำนวยการสำนักทะเบียนและวัดผล

ตามที่ อาจารย์ ดร.สุญญ์เทพ สุขแก้ว หัวหน้าโครงการวิจัย และ อาจารย์ ดร.ศุภมาส ชุมแก้ว ผู้ร่วมวิจัย ได้รับทุนอุดหนุนการวิจัยจากสถานวิจัยสถาบัน ประจำปีงบประมาณ 2565 เรื่อง "การศึกษาเปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบ 5 ตัวเลือก และ 4 ตัวเลือก" โดยงานวิจัยเรื่องนี้มีจุดประสงค์ของทางมหาวิทยาลัยด้วยการใช้ข้อมูลสถิติจากที่ทางมหาวิทยาลัยได้ดำเนินการจัดเก็บไว้ซึ่งเป็นระบบมาทำการศึกษาวิจัยเพื่อวิเคราะห์ และนำเสนอผลการวิจัยเพื่อเสนอเป็นสารสนเทศสำหรับการตัดสินใจที่จะนำไปใช้เป็นแนวทางในการพัฒนาระบบการวัดและประเมินผลของมหาวิทยาลัยสู่ทศวรรษที่ 20 นั้น ในการนี้จึงมีความจำเป็นต้องขอความอนุเคราะห์ในการเข้าถึงข้อมูลเพื่อดำเนินการวิจัย ดังนี้

1. ข้อมูลผลการตอบแบบทดสอบของนักศึกษาระดับปริญญาตรีเป็นรายบุคคล จำแนกเป็นรายข้อในชุดวิชานั้น ๆ หรือเฉลย โดยใช้ข้อมูลของชุดวิชาที่เปิดสอนในปีการศึกษา 2562 - 2564 ทั้งในภาคการศึกษาที่ 1 และภาคการศึกษาที่ 2 ที่มีนักศึกษาเข้าสอบจำนวน 1,000 คนขึ้นไป โดยผู้วิจัยจะวิเคราะห์ข้อมูลและสรุปผลการวิเคราะห์ข้อมูลในภาพรวมเท่านั้น ข้อมูลชุดวิชาและรายชื่อนักศึกษาที่ลงทะเบียนในชุดวิชาที่ถูกเลือกมาใช้ในการวิจัยครั้งนี้จะถูกลบทิ้งเป็นความลับและไม่นำมาเผยแพร่ (ขอข้อมูลในรูปแบบไฟล์ excel)

2. ข้อมูลงบประมาณในการจัดทำแบบทดสอบ ตั้งแต่ปีการศึกษา 2562 ถึงปีการศึกษา 2564 ทุกภาคการศึกษา (ขอข้อมูลในรูปแบบไฟล์ word หรือ excel)

ทั้งนี้ไฟล์ข้อมูลทั้ง 2 รายการตามที่เรียนแจ้ง อาจารย์ ดร.สุญญ์เทพ สุขแก้ว หัวหน้าโครงการวิจัย จะเป็นผู้ประสานงานในลำดับต่อไป (โทร 085-142-3169 อีเมล sarhlstthep.sukg@con.ac.th)

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์ จะขอบพระคุณยิ่ง

เรียน ผู้อำนวยการสำนักทะเบียนและวัดผล
ศูนย์วิชาการประเมินผล
สว 0602.02/คว 625
13 ก.ย. 65

สุญญ์เทพ สุขแก้ว
(อาจารย์ ดร.สุญญ์เทพ สุขแก้ว)
หัวหน้าโครงการวิจัย

13