



การเปรียบเทียบคุณสมบัติทางจิตมิติของแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิด
หลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนที่ต่างกัน

A COMPARISON OF THE PSYCHOMETRIC PROPERTIES BETWEEN MULTIPLE TRUE-FALSE
AND MODIFIED MULTIPLE TRUE-FALSE TESTS WITH DIFFERENT SCORING PROCEDURES

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อวิเคราะห์เปรียบเทียบคุณสมบัติทางจิตมิติของแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือกกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน 1,150 คน เครื่องมือที่ใช้ในการวิจัย คือ แบบสอบถูกผิดหลายตัวเลือกที่มี 4 ตัวเลือกและ 5 ตัวเลือก วิเคราะห์คุณสมบัติทางจิตมิติของข้อสอบและแบบสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) โมเดลโลจิสติก 3 พารามิเตอร์และโมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า โดยใช้โปรแกรมคอมพิวเตอร์สำเร็จรูป

ผลการวิจัยพบว่า 1) ค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าความยาก ค่าอำนาจจำแนก และค่าโอกาสในการเดาของแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และค่าฟังก์ชันสารสนเทศของข้อสอบที่ระดับความสามารถปานกลาง แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

2) ค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือก ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และค่าฟังก์ชันสารสนเทศของข้อสอบที่ระดับความสามารถต่ำจนถึงระดับความสามารถสูง แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3) ค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือกแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และค่าฟังก์ชันสารสนเทศของข้อสอบที่ระดับความสามารถต่ำและระดับความสามารถปานกลาง แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

Abstract

The purposes of this research were to compare the psychometric properties between multiple true-false test and modified multiple true-false test with different scoring procedures. Participants comprised of 1,150 students in grade-8. Data was collected through multiple true-false tests with four and five options. Data was analyzed using descriptive statistics, the three parameter logistic model and polytomous item response model, all of which employ computer programs.

The research findings were as follows: 1) The psychometric properties of multiple true-false and modified multiple true-false tests with 0-1 scoring procedure were significantly different at .05 level and item information at the mediocre ability were significantly different at .05 level.

2) The discriminations of multiple true-false tests with 0-1 and MTF scoring procedures were not significantly different and item information at the low ability level to the high ability level were significantly different at .05 level.

3) The discriminations of modified multiple true-false tests with 0-1 and MTF scoring procedures were significantly different at .05 level and item information at the low ability level and the mediocre ability level were significantly different at .05 level.

คำสำคัญ: คุณสมบัติทางจิตมิติ แบบสอบถูกผิดหลายตัวเลือก แบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ วิธีการตรวจให้คะแนน

KEYWORDS: PSYCHOMETRIC PROPERTIES/MULTIPLE TRUE-FALSE TEST/MODIFIED MULTIPLE TRUE-FALSE TEST/SCORING PROCEDURE

บทนำ

การวัดผลและการสร้างแบบสอบผลสัมฤทธิ์เป็นกระบวนการที่สำคัญอย่างยิ่งสำหรับการเรียนการสอน ในทุกระดับการศึกษา วิธีการที่สำคัญอย่างหนึ่งสำหรับการวัดผลการเรียนรู้ของผู้เรียนก็คือการทดสอบ ทั้งนี้เป้าหมายที่สำคัญของการทดสอบก็คือ เพื่อที่จะได้คะแนนของผู้สอบที่เป็นคะแนนความสามารถที่แท้จริง ซึ่งในการทดสอบนั้นนิยมใช้แบบสอบ (test) เป็นเครื่องมือในการทดสอบ แบบสอบที่นิยมใช้กันอย่างกว้างขวาง ก็คือแบบสอบประเภทเลือกคำตอบที่เป็นข้อสอบแบบหลายตัวเลือก (multiple-choice item) เพราะมีข้อดีหลายประการ เช่น สามารถสร้างข้อสอบได้ครอบคลุมเนื้อหาและจุดประสงค์การเรียนรู้ มีความเป็นปรนัยในการตรวจสูง (Haladyna & Downing, 1989; ณีภรณ์ หลาวทอง, 2548; โชติกา ภาชีผล, 2556; ศิริชัย กาญจนวาสี, 2552; เยาวดี ราชชัยกุล วิบูลย์ศรี, 2552)

ข้อสอบแบบหลายตัวเลือกมีข้อดีในการวัดผลหลายประการ แต่ก็มีข้อสงสัยว่านักเรียนที่เลือกข้อที่ถูกต้องนั้น มีความรู้ในเรื่องนั้นจริงหรือว่าเขาตอบได้ถูกต้อง และนักเรียนที่ตอบข้อสอบข้อนั้นผิด เป็นเพราะไม่มีความรู้ในเรื่องนั้นเลยหรือไม่ บุญชม ศรีสะอาด (2540) ได้กล่าวว่า วิธีการตอบและให้คะแนนของข้อสอบแบบหลายตัวเลือก นั้น ส่วนใหญ่ให้ผู้สอบเลือกเฉพาะคำตอบที่ถูกที่สุดเพียงคำตอบเดียว ผู้สอบที่ไม่รู้จริงในเรื่องนั้นจึงมักเดาคำตอบ คือเปิดโอกาสให้ผู้สอบเดาข้อสอบได้ ดังนั้นคะแนนที่ผู้สอบทำได้จึงบอกความหมายที่แท้จริงได้ไม่ชัดเจน จากปัญหาดังกล่าวนักวัดผลพยายามปรับปรุงรูปแบบข้อสอบเพื่อแก้จุดอ่อนดังกล่าว โดยการเลือกใช้แบบเลือกตอบ ชนิดหลายคำตอบ (multiple-answer: MA) การใช้แบบสอบเลือกตอบชนิดเชิงซ้อน (complex multiple-choice: CMC) และ การใช้แบบสอบถูกผิดหลายตัวเลือก (multiple true-false: MTF) ทั้งนี้แบบสอบทั้งสามที่

พัฒนาขึ้นมาก็มีจุดเด่นจุดด้อยที่ต่างกัน แต่จากผลการวิจัยพบว่าแบบสอบถูกผิดหลายตัวเลือก (MTF) มีความเที่ยงและความตรงสูงกว่าแบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว (Frisbie & Sweeney, 1982; Leila, 2014)

ข้อสอบถูกผิดหลายตัวเลือก (MTF) เป็นข้อสอบที่มีลักษณะคล้ายกับข้อสอบแบบหลายตัวเลือก (MC) เกือบทุกประการโดยมีส่วนประกอบอยู่ 2 ส่วน ได้แก่ ส่วนที่เป็นข้อคำถามและส่วนที่เป็นตัวเลือก สิ่งที่แตกต่างกันก็คือ ข้อสอบถูกผิดหลายตัวเลือกผู้สอบจะต้องพิจารณาตัวเลือกทุกตัว และจำนวนตัวเลือกที่เป็นคำตอบถูกอาจจะถูก 1, 2, 3 หรือถูกทุกตัวเลือก หรือไม่มีตัวถูกเลยก็ได้ ขึ้นอยู่กับเนื้อหาที่ต้องการวัด นอกจากนี้จากการศึกษาเอกสารและงานวิจัยเกี่ยวกับแบบสอบถูกผิดที่มีการดัดแปลงเป็นแบบสอบถูกผิดปรับปรุงใหม่ (modified true-false) พบว่า มีความยากมากกว่าแบบสอบถูกผิดธรรมดา ทั้งนี้เพราะสามารถลดการเดาของผู้สอบเนื่องจากรูปแบบ (format) ของข้อสอบที่กำหนดให้ผู้สอบต้องแก้ไขข้อความที่ผิด พร้อมแก้ไขข้อความดังกล่าวให้ถูกต้อง และยังพบว่ามีความสัมพันธ์ในการจำแนกผู้เรียน กลุ่มเก่งและกลุ่มอ่อนออกจากกันได้ดี มีความเที่ยงในระดับดีที่จะใช้ในการประเมินผู้สอบ ด้วยเหตุนี้ผู้วิจัยจึงนำหลักการและแนวคิดดังกล่าวมาประยุกต์ใช้กับแบบสอบถูกผิดหลายตัวเลือก ให้เป็นแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ (modified multiple true-false: MMTF) คือข้อสอบจะมีส่วนประกอบเหมือนกับข้อสอบถูกผิดหลายตัวเลือกทุกประการ แต่ต่างกันที่วิธีการตอบข้อสอบ โดยในตัวเลือกที่ผู้สอบพิจารณาว่า “ผิด” ผู้สอบต้องแก้ไขข้อความในตัวเลือกนั้นให้ถูกต้องด้วย ทั้งนี้เพื่อเพิ่มสารสนเทศจากแบบสอบถูกผิดหลายตัวเลือกและเพื่อใช้ในการประเมินความสามารถของผู้สอบได้ตรงตามความสามารถที่แท้จริงหรือใกล้เคียงความสามารถที่แท้จริงของผู้สอบให้มากที่สุด (Frisbie, 1992; Albanese & Sabers, 1988; Frisbie & Sweeney, 1982; Leila, 2014; Howard & Francis, 1926; Francis, Wesley & Nine, 1943; ศิริชัย กาญจนวาสี, 2552; นภา หลิมรัตน์, 2552)

ข้อสอบถูกผิดหลายตัวเลือกมีวิธีการตรวจให้คะแนนหลายวิธีด้วยกัน สามารถใช้วิธีการตรวจให้คะแนนแบบวิธีประเพณีนิยมและการให้คะแนนความรู้บางส่วน (โชติกา ภาษีผล, 2545) จากการศึกษางานวิจัยที่ผ่านมา ยังไม่มีความกระจ่างชัดในเรื่องของวิธีการตรวจให้คะแนนของข้อสอบแบบถูกผิดหลายตัวเลือก ที่ผ่านมามีผู้ศึกษาเกี่ยวกับวิธีการตรวจให้คะแนนที่แตกต่างกันหลายวิธี เช่น วิธีการตรวจให้คะแนนแบบ 0-1 วิธีนี้เป็นวิธีที่ผู้สอบจะได้ 1 คะแนนเมื่อผู้สอบเลือกตัวเลือกที่ตรงกับสภาพจริงของตัวเลือกนั้น ๆ ได้ทุกตัวเลือก และได้ 0 คะแนน เมื่อผู้สอบเลือกตัวเลือกที่ไม่ตรงกับสภาพจริงของตัวเลือกนั้นๆ ตั้งแต่ 1 ตัวเลือกขึ้นไป หรือวิธีการตรวจให้คะแนนรายตัวเลือกวิธีการตรวจให้คะแนนวิธีนี้ Albanese et al (1988) ได้เสนอไว้ โดยวิธีการตรวจให้คะแนนวิธีนี้ใช้กับแบบสอบถูกผิดหลายตัวเลือก โดยในข้อสอบหนึ่งข้อจะมีตัวเลือกที่ถูกและผิดมากกว่าหนึ่งตัวเลือก แล้วให้ผู้สอบพิจารณาตัวเลือกที่ถูกและที่ผิด ซึ่งวิธีการตรวจจะให้ 1 คะแนน ต่อตัวเลือก เมื่อผู้สอบตอบได้ตรงกับสภาพจริงของตัวเลือกนั้น และให้ 0 คะแนนต่อตัวเลือก ที่ผู้สอบตอบไม่ตรงกับสภาพจริงของตัวเลือกนั้น นอกจากนี้ยังมีวิธีการตรวจให้คะแนนแบบอื่น ๆ อีก ซึ่งแต่ละวิธีต่างก็มีเหตุผลที่เหมาะสมในการเลือกใช้ที่แตกต่างกัน และยังไม่มีการสรุปที่ชัดเจนว่าวิธีการตรวจให้คะแนนวิธีใดจะทำให้ได้ค่าสารสนเทศของแบบสอบที่ดีที่สุด

จากงานวิจัยที่กล่าวมาที่เกี่ยวกับข้อสอบถูกผิดหลายตัวเลือกและวิธีการตรวจให้คะแนน ส่วนใหญ่จะอาศัยทฤษฎีการวัดแบบดั้งเดิม (classical test theory: CTT) ในการวิเคราะห์ข้อมูลโดยมีเพียงการวิเคราะห์ความยากอำนาจจำแนก ความเที่ยงและความตรง เช่น งานวิจัยของ Frisbie and Sweeney (1982) ที่ได้ศึกษาข้อได้เปรียบของแบบทดสอบวัดผลสัมฤทธิ์ที่มีตัวเลือกถูกหลายตัวเลือก โดยเทียบกับแบบทดสอบที่มีตัวเลือกถูกตัวเดียว งานวิจัยของ Fu and Hoi (1993) ที่ได้ศึกษาค่าความเที่ยง ค่าความยากและอำนาจจำแนกของข้อสอบถูกผิด

หลายตัวเลือกที่มีการตรวจให้คะแนนต่างกัน 6 วิธี คือวิธี MR วิธี MTF วิธี Count-2 วิธี Count-3 วิธี CFG และวิธี LO หรืองานวิจัยของวัชรภรณ์ จิตรมาศ (2538) ที่ได้เปรียบเทียบค่าความยาก ค่าอำนาจจำแนกและค่าความเที่ยงของแบบทดสอบเลือกตอบชนิดถูกผิดหลายตัวเลือกที่มีจำนวนตัวเลือกและวิธีการตรวจให้คะแนนต่างกัน มีงานวิจัยเพียงไม่กี่เรื่องที่ใช้ทฤษฎีการตอบสนองข้อสอบ (item response theory: IRT) ในการวิเคราะห์ข้อมูล เช่น งานวิจัยของพัชรี มีวรรณ (2540) ที่ได้เปรียบเทียบผลของจำนวนตัวเลือกที่มีต่อค่าฟังก์ชันสารสนเทศของแบบสอบเลือกตอบแบบถูกผิดในระดับการศึกษาที่ต่างกัน ซึ่งทฤษฎีการตอบสนองข้อสอบเป็นทฤษฎีที่มีบทบาทอย่างมากต่อการวัดผลการศึกษาและจิตวิทยาในปัจจุบัน เพราะสามารถแก้จุดอ่อนของทฤษฎีการวัดแบบดั้งเดิมได้ โดยทฤษฎีการตอบสนองข้อสอบ ได้ให้แนวคิดที่ชัดเจนเกี่ยวกับการประมาณค่าความสามารถของผู้สอบ ซึ่งค่าพารามิเตอร์ของผู้สอบ (θ) คือระดับความสามารถของผู้สอบ นอกจากนี้ยังมีค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าโอกาสในการเดา (c) (เฉลิมฤทธิ์ แก้วกล้าหาญ, 2550) ในการวิจัยครั้งนี้ผู้วิจัยจึงใช้ทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นแนวทางในการวิเคราะห์ เพื่อเปรียบเทียบคุณสมบัติทางจิตมิติของแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่มีวิธีการตรวจให้คะแนนที่ต่างกัน เพื่อให้ได้ข้อความรู้ที่สามารถนำไปใช้เป็นแนวทางในการสร้าง และพัฒนาแบบสอบให้มีคุณภาพและสามารถนำไปใช้วัดผลการเรียนรู้ได้อย่างมีประสิทธิภาพ

วัตถุประสงค์การวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์ทั่วไปเพื่อศึกษาเปรียบเทียบคุณสมบัติทางจิตมิติของแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือก โดยมีวัตถุประสงค์เฉพาะ ดังนี้

1. เพื่อเปรียบเทียบคุณสมบัติทางจิตมิติของแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1
2. เพื่อเปรียบเทียบคุณสมบัติทางจิตมิติของแบบสอบถูกผิดหลายตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 และวิธีตรวจให้คะแนนแบบรายตัวเลือก
3. เพื่อเปรียบเทียบคุณสมบัติทางจิตมิติของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 และวิธีตรวจให้คะแนนแบบรายตัวเลือก

สมมติฐานของการวิจัย

เนื่องจากยังไม่มีการศึกษาเกี่ยวกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ ผู้วิจัยจึงได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับแบบสอบถูกผิดที่มีการดัดแปลงเป็นแบบสอบถูกผิดปรับปรุงใหม่ (Modified True-False) ผลการวิจัยพบว่ามีความยากมากกว่าแบบสอบถูกผิดธรรมดา ทั้งนี้เพราะสามารถลดการเดาของผู้สอบ เนื่องจากรูปแบบ (format) ของข้อสอบที่กำหนดให้ผู้สอบต้องแก้ไขข้อความที่ผิด พร้อมแก้ไขข้อความดังกล่าวให้ถูกต้อง นอกจากนี้ยังพบว่ามีความสัมพันธ์ในการจำแนกผู้เรียนกลุ่มเก่งและกลุ่มอ่อนออกจากกันได้ดี มีความเที่ยงในระดับดีที่จะใช้ในการประเมินผู้สอบ (Howard & Francis, 1926; Francis, Wesley & Nine, 1943; ศิริชัย กาญจนวาสี, 2552; นภา หลิมรัตน์, 2552)

เมื่อพิจารณาจากงานวิจัยที่เกี่ยวข้อง ผู้วิจัยจึงตั้งสมมติฐานของการวิจัย ดังนี้

สมมติฐานข้อที่ 1 คุณสมบัติทางจิตมิติของแบบสอบถามหลายตัวเลือกปรับปรุงใหม่将会有ค่าความยาก (b) ค่าอำนาจจำแนก (a) สูงกว่าแบบสอบถามหลายตัวเลือก ในขณะที่ค่าโอกาสในการเดา (c) น่าจะมีค่าต่ำกว่าแบบสอบถามหลายตัวเลือก

จากงานวิจัยที่ศึกษาเกี่ยวกับวิธีการตรวจให้คะแนนของข้อสอบถูกผิดหลายตัวเลือก (MTF) พบว่าวิธีที่ตรวจให้คะแนนส่วนใหญ่ วิเคราะห์คุณภาพแบบสอบและข้อสอบโดยใช้ทฤษฎีการทดสอบแบบดั้งเดิม (CTT) และมีผลการวิเคราะห์ที่แตกต่างกันขึ้นอยู่กับบริบทของเนื้อหาของแบบสอบ แต่วิธีการตรวจให้คะแนนที่มีความเที่ยงเกือบทุกงานวิจัยก็คือ วิธีการตรวจให้คะแนนแบบให้คะแนนทุกตัวเลือกเป็น 1 คะแนน แต่ก็ยังไม่มีย่อสรุปที่ชัดเจนว่าวิธีการตรวจให้คะแนนวิธีใดจะทำให้ได้ค่าสารสนเทศของแบบสอบที่ดีที่สุด (Frisbie & Sweeney, 1982; Albaness & Sabers, 1988; Fu & Hoi, 1993; วัชรารณณ์ จิตรมาศ, 2538; พัชรี มีวรรณ, 2540; สุรชัย มีชาญ, 2545)

เมื่อพิจารณาจากงานวิจัยที่เกี่ยวข้อง ผู้วิจัยจึงตั้งสมมติฐานของการวิจัย ดังนี้

สมมติฐานข้อที่ 2 ค่าอำนาจจำแนก (a) ของแบบสอบถามหลายตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบรายตัวเลือก น่าจะมีค่าสูงกว่าวิธีการตรวจให้คะแนนแบบ 0-1 และค่าฟังก์ชันสารสนเทศของข้อสอบของแบบสอบถามหลายตัวเลือก (MTF) ที่ใช้วิธีการตรวจให้คะแนนแบบรายตัวเลือกน่าจะจำแนกผู้เรียนกลุ่มเก่งและกลุ่มอ่อนออกจากกันได้สูงกว่าวิธีการตรวจให้คะแนนแบบ 0-1

สมมติฐานข้อที่ 3 ค่าอำนาจจำแนก (a) ของแบบสอบถามหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบรายตัวเลือกน่าจะมีความสูงกว่าวิธีการตรวจให้คะแนนแบบ 0-1 และค่าฟังก์ชันสารสนเทศของข้อสอบของแบบสอบถามหลายตัวเลือกปรับปรุงใหม่ (MMTF) ที่ใช้วิธีการตรวจให้คะแนนแบบรายตัวเลือกน่าจะจำแนกผู้เรียนกลุ่มเก่งและกลุ่มอ่อนออกจากกันได้สูงกว่าวิธีการตรวจให้คะแนนแบบ 0-1

วิธีดำเนินการวิจัย

ประชากรและตัวอย่างที่ใช้ในการวิจัย

ประชากรที่ใช้ในการวิจัยครั้งนี้คือนักเรียนชั้นมัธยมศึกษาปีที่ 2 ของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษา เขต 1 และ 2 (กรุงเทพมหานคร) จำนวน 44,719 คน

ตัวอย่างที่ใช้ในการศึกษาครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ปีการศึกษา 2557 ของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษา เขต 1 และ 2 (กรุงเทพมหานคร) โดยในการวิจัยครั้งนี้ผู้วิจัยใช้การสุ่มแบบหลายขั้นตอน (multi-stage sampling) ดังนี้ ขั้นที่ 1 ผู้วิจัยสุ่มโรงเรียนโดยการสุ่มอย่างง่าย (simple random sampling) ได้ 5 โรงเรียน ได้แก่ โรงเรียนศรีอยุธยาในพระอุปถัมภ์ฯ โรงเรียนสันติราษฎร์วิทยาลัย โรงเรียนยานนาวาเวศวิทยาคม อิสลามวิทยาลัยแห่งประเทศไทย และโรงเรียนสตรีวัดมหาพฤฒารามในพระบรมราชินูปถัมภ์ ขั้นที่ 2 สุ่มห้องเรียนชั้นมัธยมศึกษาปีที่ 2 ทั้งหมดของแต่ละโรงเรียน ได้โรงเรียนละ 4 ห้อง โดยการสุ่มอย่างง่าย (simple random sampling) รวมจำนวนนักเรียนทั้งหมด 1,150 คน

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลครั้งนี้ เป็นแบบสอบถามเลือกตอบ เนื้อหาวิทยาศาสตร์
ชั้นมัธยมศึกษาปีที่ 2 ที่ผู้วิจัยสร้างเอง จำนวน 2 ฉบับ ฉบับละ 25 ข้อ คือ

ฉบับที่ 1 แบบสอบถามชนิดหลายตัวเลือกที่มี 4 ตัวเลือก

ฉบับที่ 2 แบบสอบถามชนิดหลายตัวเลือกที่มี 5 ตัวเลือก

วิธีดำเนินการเก็บรวบรวมข้อมูล

1. ประสานงานและทำจดหมายขออนุญาตและเข้าไปติดต่อโรงเรียนที่เป็นตัวอย่าง จำนวน 5 โรงเรียน
เพื่อกำหนดวัน เวลาและสถานที่ที่ใช้ในการสอบ หลังจากนั้นวางแผนการดำเนินการสอบและจัดเตรียมแบบสอบ
พร้อมทั้งกระดาษคำตอบให้เพียงพอกับจำนวนนักเรียน

2. ก่อนการทดสอบ ผู้วิจัยทำความเข้าใจกับผู้สอบเกี่ยวกับรายละเอียดของแบบสอบ รวมไปถึงวิธีการตอบ
ข้อสอบของแบบสอบถามชนิดหลายตัวเลือก เนื่องจากเป็นแบบสอบที่นักเรียนไม่คุ้น ผู้วิจัยเขียนคำชี้แจงและ
คำแนะนำในการสอบอย่างชัดเจนพิมพ์ไว้ในหน้าแรกของแบบสอบเพื่อให้ผู้สอบได้รับทราบตรงกัน นอกจากนี้ผู้วิจัย
ต้องขอให้ครูผู้สอนแจ้งให้นักเรียนทราบว่า คะแนนที่ได้จากการสอบครั้งนี้ใช้เป็นคะแนนเก็บ เพื่อให้นักเรียนตั้งใจ
ทำข้อสอบ

3. คำชี้แจงในการทำแบบสอบถามชนิดหลายตัวเลือก (MTF) ที่มี 4 และ 5 ตัวเลือก

คำถามแต่ละข้อจะมีตัวเลือกที่ถูกต้องตั้งแต่หนึ่งตัวเลือกจนถึงถูกหมดทุกตัวเลือก ในการตอบนักเรียน
จะต้องพิจารณาทุกตัวเลือก ถ้าตัวเลือกใดเป็นจริงให้กากบาทในช่องคำว่า “ถูก” ถ้าตัวเลือกใดไม่เป็นจริงให้
กากบาทในช่องคำว่า “ผิด” ลงในกระดาษคำตอบ

ตัวอย่างกระดาษคำตอบของข้อสอบชนิดหลายตัวเลือก

ข้อที่	ตัวเลือก	ถูก	ผิด
1	ก	×	
	ข	×	
	ค		×
	ง		×

4. คำชี้แจงในการทำแบบสอบถามชนิดหลายตัวเลือกปรับปรุงใหม่ (MMTF) ที่มี 4 และ 5 ตัวเลือก

คำถามแต่ละข้อจะมีตัวเลือกที่ถูกต้องตั้งแต่หนึ่งตัวเลือกจนถึงถูกทุกตัวเลือก ในการตอบนักเรียน
จะต้องพิจารณาทุกตัวเลือก ถ้าตัวเลือกใดเป็นจริงให้กากบาทในช่องคำว่า “ถูก” ถ้าตัวเลือกใดไม่เป็นจริงให้
กากบาทในช่องคำว่า “ผิด” ลงในกระดาษคำตอบ แล้วแก้ไขข้อความในตัวเลือกที่ผิดนั้นให้ถูกต้อง

ตัวอย่างกระดาษคำตอบของข้อสอบชนิดหลายตัวเลือกปรับปรุงใหม่

ข้อที่	ตัวเลือก	ถูก	ผิด	คำแก้ไขหรือเหตุผล
1	ก	×		
	ข	×		
	ค		×	ดินชั้นบนมีความพรุนมากกว่า
	ง		×	ดินชั้นบนมีขนาดของเม็ดดินใหญ่กว่า

5. ทดสอบกับนักเรียน โดยสุ่มนักเรียนแต่ละห้องออกเป็น 4 กลุ่ม ด้วยวิธีการสุ่มอย่างง่าย จากนั้นนำแบบสอบที่ผู้วิจัยสร้างเองจำนวน 2 ฉบับมาจับคู่กับกระดาษคำตอบ ถ้าจับคู่กับกระดาษคำตอบที่ไม่ต้องแก้ไขข้อความจะเรียกแบบสอบชุดนั้นว่า “แบบสอบถูกผิดหลายตัวเลือก” ถ้าจับคู่กับกระดาษคำตอบแบบแก้ไขข้อความจะเรียกแบบสอบชุดนั้นว่า “แบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่” ดังนั้นจะได้แบบสอบทั้งหมด 4 ฉบับ ซึ่งแต่ละฉบับมีข้อสอบทั้งหมด 25 ข้อ หลังจากนั้นสุ่มแบบสอบให้กับกลุ่มตัวอย่างโดยการสุ่มอย่างง่าย ซึ่งมีรายละเอียดของแบบสอบ ดังนี้

ฉบับที่ 1 แบบสอบถูกผิดหลายตัวเลือกที่มี 4 ตัวเลือก

ฉบับที่ 2 แบบสอบแบบหลายตัวเลือกที่มี 5 ตัวเลือก

ฉบับที่ 3 แบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่มี 4 ตัวเลือก

ฉบับที่ 4 แบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่มี 5 ตัวเลือก

6. หลังจากดำเนินการสอบเรียบร้อยแล้ว จัดกระดาษคำตอบของข้อสอบแต่ละฉบับไว้ด้วยกัน

7. นำกระดาษคำตอบที่ได้ มาตรวจให้คะแนนเพื่อนำผลมาวิเคราะห์ทางสถิติและทดสอบสมมติฐาน

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลมีขั้นตอนดังต่อไปนี้

1. ตรวจให้คะแนนแบบสอบแต่ละฉบับ ตามเกณฑ์การตรวจให้คะแนนที่ตั้งไว้

2. วิเคราะห์ค่าสถิติบรรยายของคะแนนสอบนักเรียน วิชาวิทยาศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 โดยใช้โปรแกรมคอมพิวเตอร์สำเร็จรูป ดังนี้ คะแนนต่ำสุด (Minimum) คะแนนสูงสุด (Maximum) ความเบ้ (Skewness: Sk) ความโด่ง (Kurtosis: Ku) และคะแนนเฉลี่ย (mean)

3. วิเคราะห์คุณสมบัติทางจิตมิติของข้อสอบและแบบสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) โมเดลโลจิสติก 3 พารามิเตอร์และโมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า ดังนี้

3.1 ตรวจสอบคุณภาพของข้อสอบที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 โดยใช้การวิเคราะห์โมเดลโลจิสติก 3 พารามิเตอร์ ประมาณค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าโอกาสในการเดา (c) และค่าฟังก์ชันสารสนเทศของข้อสอบ (IIF)

3.2 ตรวจสอบคุณภาพของตัวเลือก ข้อสอบและแบบสอบที่ใช้วิธีการตรวจให้คะแนนแบบรายตัวเลือก โดยใช้โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า ประมาณค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนก (a) และค่าฟังก์ชันสารสนเทศของข้อสอบ (IIF)

ผลการวิจัย

1. ผลการเปรียบเทียบคุณสมบัติทางจิตมิติของแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1

กรณีแบบสอบ 4 ตัวเลือก ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าความยาก ค่าอำนาจจำแนกและค่าโอกาสในการเดา พบว่า ค่าความยากของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่มีค่ามากกว่าแบบสอบถูกผิดหลายตัวเลือก ($M = 2.48, SD = 2.15$ และ $M = 1.38, SD = 1.42$ ตามลำดับ) เมื่อพิจารณา ค่าอำนาจจำแนก พบว่าค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่มีค่ามากกว่า

แบบสอบถูกผิดหลายตัวเลือก ($M = 2.20$, $SD = 2.17$ และ $M = 0.93$, $SD = 1.29$ ตามลำดับ) และค่าโอกาสในการเดาของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่มีค่าน้อยกว่าแบบสอบถูกผิดหลายตัวเลือก ($M = 0.06$, $SD = 0.08$ และ $M = 0.15$, $SD = 0.13$ ตามลำดับ) และผลการเปรียบเทียบค่าพารามิเตอร์ของข้อสอบของแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 พบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

กรณีแบบสอบ 5 ตัวเลือก ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าความยาก ค่าอำนาจจำแนกและค่าโอกาสในการเดา พบว่า ค่าความยากของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่มีค่ามากกว่าแบบสอบถูกผิดหลายตัวเลือก ($M = 2.92$, $SD = 2.39$ และ $M = 1.72$, $SD = 1.22$ ตามลำดับ) เมื่อพิจารณาอำนาจจำแนก พบว่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่มีค่ามากกว่าแบบสอบถูกผิดหลายตัวเลือก ($M = 2.06$, $SD = 1.88$ และ $M = 0.86$, $SD = 1.22$ ตามลำดับ) และค่าโอกาสในการเดาของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่มีค่าน้อยกว่าแบบสอบถูกผิดหลายตัวเลือก ($M = 0.06$, $SD = 0.08$ และ $M = 0.16$, $SD = 0.12$ ตามลำดับ) และผลการเปรียบเทียบค่าพารามิเตอร์ของข้อสอบของแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 พบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังแสดงในตารางที่ 1

ตารางที่ 1 ผลการเปรียบเทียบค่าพารามิเตอร์ของข้อสอบของแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่จำนวน 4 และ 5 ตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1

จำนวน ตัวเลือก	แบบสอบ	ความยาก (b)		อำนาจจำแนก (a)		โอกาสในการเดา (c)	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
4	MTF	1.38	1.42	0.93	1.29	0.15	0.13
	MMTF	2.48	2.15	2.20	2.17	0.06	0.08
การทดสอบความแตกต่าง		$t = -2.149^{**}$		$t = -2.522^{**}$		$t = 2.975^{**}$	
5	MTF	1.72	1.22	0.86	1.22	0.16	0.12
	MMTF	2.92	2.39	2.06	1.88	0.06	0.08
การทดสอบความแตกต่าง		$t = -2.226^{**}$		$t = -2.691^{**}$		$t = 3.580^{**}$	

หมายเหตุ: MTF หมายถึง แบบสอบถูกผิดหลายตัวเลือก
 MMTF หมายถึง แบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่
 ** $p < 0.05$

ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของข้อสอบของแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่จำนวน 4 และ 5 ตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 พบว่า ที่ระดับความสามารถปานกลางแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังแสดงในตารางที่ 2

ตารางที่ 2 ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของข้อสอบของแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่จำนวน 4 และ 5 ตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 จำแนกตามระดับความสามารถของผู้สอบ

จำนวน ตัวเลือก	แบบสอบ	ระดับความสามารถของผู้สอบ					
		ต่ำ ($\theta < -1.00$)		ปานกลาง ($-1.00 \leq \theta \leq 1.00$)		สูง ($\theta > 1.00$)	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
4	MTF	0.008	0.003	0.198	0.238	0.841	1.221
	MMTF	0.013	0.005	0.049	0.021	0.413	0.306
การทดสอบความแตกต่าง		t = -0.884		t = 0.865**		t = -0.413	
5	MTF	0.007	0.003	0.171	0.195	0.804	1.074
	MMTF	0.014	0.005	0.049	0.020	0.375	0.260
การทดสอบความแตกต่าง		t = -1.358		t = 0.800**		t = -0.182	

หมายเหตุ: MTF หมายถึง แบบสอบถูกผิดหลายตัวเลือก
 MMTF หมายถึง แบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่
 ** $p < 0.05$

2. ผลการเปรียบเทียบคุณสมบัติทางจิตมิติของแบบสอบถูกผิดหลายตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 และวิธีตรวจให้คะแนนแบบรายตัวเลือก

กรณีแบบสอบ 4 ตัวเลือก ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนก พบว่าค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบรายตัวเลือกมีค่ามากกว่าวิธีการตรวจให้คะแนนแบบ 0-1 ($M = 0.93, SD = 1.29$ และ $M = 0.64, SD = 0.25$ ตามลำดับ) และเมื่อทดสอบความแตกต่าง พบว่าค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือก ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

กรณีแบบสอบ 5 ตัวเลือก ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนก พบว่าค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบรายตัวเลือกมีค่ามากกว่าวิธีการตรวจให้คะแนนแบบ 0-1 ($M = 0.86, SD = 1.22$ และ $M = 0.57, SD = 0.25$ ตามลำดับ) และเมื่อทดสอบความแตกต่าง พบว่าค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือก ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังแสดงในตารางที่ 3

ตารางที่ 3 ผลการเปรียบเทียบค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกจำนวน 4 และ 5 ตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือก

จำนวนตัวเลือก	วิธีการตรวจให้คะแนน	อำนาจจำแนก	
		M	SD
4	วิธีการตรวจให้คะแนนแบบ 0-1	0.64	0.25
	วิธีการตรวจให้คะแนนแบบรายตัวเลือก	0.93	1.29
การทดสอบความแตกต่าง		t = 1.113	
5	วิธีการตรวจให้คะแนนแบบ 0-1	0.57	0.25
	วิธีการตรวจให้คะแนนแบบรายตัวเลือก	0.86	1.22
การทดสอบความแตกต่าง		t = 1.143	

ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของข้อสอบของแบบสอบถูกผิดหลายตัวเลือกจำนวน 4 และ 5 ตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือก พบว่า ที่ระดับความสามารถต่ำจนถึงระดับความสามารถสูง แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังแสดงในตารางที่ 4

ตารางที่ 4 ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของข้อสอบของแบบสอบถูกผิดหลายตัวเลือกจำนวน 4 และ 5 ตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือก จำแนกตามระดับความสามารถของผู้สอบ

จำนวนตัวเลือก	วิธีการตรวจให้คะแนน	ระดับความสามารถของผู้สอบ					
		ต่ำ		ปานกลาง		สูง	
		$(\theta < -1.00)$		$(-1.00 \leq \theta \leq 1.00)$		$(\theta > 1.00)$	
		M	SD	M	SD	M	SD
4	0-1	0.104	0.012	0.132	0.007	0.121	0.014
	รายตัวเลือก	0.008	0.003	0.198	0.238	0.841	1.221
การทดสอบความแตกต่าง		t = -7.917**		t = -1.728**		t = 0.789**	
5	0-1	0.117	0.013	0.131	0.002	0.111	0.013
	รายตัวเลือก	0.007	0.003	0.171	0.195	0.804	1.074
การทดสอบความแตกต่าง		t = -5.185**		t = -1.199**		t = 1.274**	

หมายเหตุ: ** p < 0.05

3. ผลการเปรียบเทียบคุณสมบัติทางจิตมิติของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 และวิธีตรวจให้คะแนนแบบรายตัวเลือก

กรณีแบบสอบ 4 ตัวเลือก ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนก พบว่าค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบรายตัวเลือกมีค่ามากกว่าวิธีการตรวจให้คะแนนแบบ 0-1 ($M = 2.20, SD = 2.17$ และ $M = 0.77, SD = 0.16$ ตามลำดับ) และเมื่อทดสอบความแตกต่าง พบว่าค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือก แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

กรณีแบบสอบ 5 ตัวเลือก ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนก พบว่าค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบรายตัวเลือกมีค่ามากกว่าวิธีการตรวจให้คะแนนแบบ 0-1 ($M = 2.06, SD = 0.18$ และ $M = 0.72, SD = 0.18$ ตามลำดับ) และเมื่อทดสอบความแตกต่าง พบว่าค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือก แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังแสดงในตารางที่ 5

ตารางที่ 5 ผลการเปรียบเทียบค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่จำนวน 4 และ 5 ตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือก

จำนวนตัวเลือก	วิธีการตรวจให้คะแนน	อำนาจจำแนก	
		<i>M</i>	<i>SD</i>
4	วิธีการตรวจให้คะแนนแบบ 0-1	0.77	0.16
	วิธีการตรวจให้คะแนนแบบรายตัวเลือก	2.20	2.17
การทดสอบความแตกต่าง		t = 3.284**	
5	วิธีการตรวจให้คะแนนแบบ 0-1	0.72	0.18
	วิธีการตรวจให้คะแนนแบบรายตัวเลือก	2.06	1.88
การทดสอบความแตกต่าง		t = 3.539**	

หมายเหตุ: ** p < 0.05

ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของข้อสอบของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่จำนวน 4 และ 5 ตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือก พบว่าที่ระดับความสามารถต่ำและระดับความสามารถปานกลาง แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังแสดงในตารางที่ 6

ตารางที่ 6 ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของข้อสอบของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ จำนวน 4 และ 5 ตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือก จำแนกตามระดับความสามารถของผู้สอบ

จำนวน ตัวเลือก	วิธีการตรวจให้คะแนน	ระดับความสามารถของผู้สอบ					
		ต่ำ		ปานกลาง		สูง	
		$(\theta < -1.00)$		$(-1.00 \leq \theta \leq 1.00)$		$(\theta > 1.00)$	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
4	0-1	0.013	0.005	0.049	0.021	0.413	0.306
	รายตัวเลือก	0.147	0.011	0.177	0.003	0.131	0.027
	การทดสอบความแตกต่าง	t = -11.294**		t = -6.679**		t = 0.915	
5	0-1	0.014	0.005	0.049	0.020	0.375	0.260
	รายตัวเลือก	0.192	0.004	0.195	0.003	0.165	0.017
	การทดสอบความแตกต่าง	t = -10.895**		t = -7.367**		t = 0.545	

หมายเหตุ: ** p < 0.05

อภิปรายผล

จากผลการวิจัย ผู้วิจัยนำเสนอการอภิปรายผลในประเด็นดังต่อไปนี้

1. การเปรียบเทียบคุณสมบัติทางจิตมิติของแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1

ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าความยาก ค่าอำนาจจำแนก และค่าโอกาสในการเดาของแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 พบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งสอดคล้องกับสมมติฐานข้อที่ 1 คือคุณสมบัติทางจิตมิติของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่มีค่าความยาก (b) ค่าอำนาจจำแนก (a) สูงกว่าแบบสอบถูกผิดหลายตัวเลือก ในขณะที่ค่าโอกาสในการเดา (c) มีค่าต่ำกว่าแบบสอบถูกผิดหลายตัวเลือก ทั้งนี้เพราะรูปแบบของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ ผู้สอบต้องแก้ไขข้อความที่ผิดให้ถูกต้อง จึงทำให้ข้อสอบยากกว่าสามารถจำแนกผู้เรียนได้ดีกว่าและโอกาสในการเดาคำตอบจึงมีน้อยลง สอดคล้องกับงานวิจัยของ Howard and Francis (1926) ที่ได้ศึกษารูปแบบของแบบสอบถูกผิดปรับปรุงใหม่ และงานวิจัยของ นภา หลิมธรัตน์ (2552) ที่ได้เปรียบเทียบผลการวิเคราะห์แบบทดสอบแบบถูกผิดกับแบบสอบถูกผิดที่ต้องแก้ไขข้อความ

2. การเปรียบเทียบคุณสมบัติทางจิตมิติของแบบสอบถูกผิดหลายตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 และวิธีตรวจให้คะแนนแบบรายตัวเลือก

ผลการวิเคราะห์ค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือกไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมติฐานข้อที่ 2 ทั้งนี้อาจจะเนื่องมาจากรูปแบบของแบบสอบถูกผิดหลายตัวเลือกที่ไม่ต้องแก้ไขข้อความที่ผิดให้ถูกต้อง รวมทั้งเนื้อหาของข้อสอบอาจจะเป็นเนื้อหาที่ไม่ยากจนเกินไป จึงทำให้คะแนนของผู้สอบที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือกใกล้เคียงกัน แต่เมื่อพิจารณาค่าฟังก์ชันสารสนเทศของข้อสอบเป็นไปตามสมมติฐาน คือ ค่าฟังก์ชันสารสนเทศของข้อสอบของแบบสอบถูกผิดหลายตัวเลือกที่ใช้วิธีการตรวจให้คะแนนแบบรายตัวเลือกจำแนกผู้เรียนกลุ่มเก่งและกลุ่มอ่อนออกจากกันได้สูงกว่าวิธีการตรวจให้คะแนนแบบ 0-1 ซึ่งที่ระดับความสามารถต่ำจนถึงระดับความสามารถสูง แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3. ผลการเปรียบเทียบคุณสมบัติทางจิตมิติของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 และวิธีตรวจให้คะแนนแบบรายตัวเลือก

ผลการวิเคราะห์ค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบ 0-1 กับวิธีการตรวจให้คะแนนแบบรายตัวเลือกแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งสอดคล้องกับสมมติฐานข้อที่ 3 คือ ค่าอำนาจจำแนกของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบรายตัวเลือกมีค่าสูงกว่าวิธีการตรวจให้คะแนนแบบ 0-1 ทั้งนี้เพราะรูปแบบของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ผู้สอบต้องแก้ไขข้อความที่ผิดให้ถูกต้อง ในการให้คะแนนต้องพิจารณาทั้งในส่วนของการเลือกตัวเลือกนั้น ๆ ว่าตอบถูกหรือผิด รวมทั้งในกรณีที่ผู้สอบเลือกตอบตัวเลือกนั้นว่าผิด ผู้สอบต้องมีการแก้ไขข้อความ ในการให้คะแนนจึงต้องพิจารณาข้อความที่แก้ไขอีกส่วนหนึ่งด้วย ถ้าตรวจให้คะแนนแบบ 0-1 ถึงแม้ผู้สอบพิจารณาตัวเลือกได้ถูกต้อง แต่แก้ไขข้อความผิดก็จะได้คะแนนในข้อนั้นเลย ซึ่งแท้จริงแล้วผู้สอบอาจจะมีความรู้บางส่วน เช่น อาจพิจารณาตัวเลือกถูกแต่อาจจะแก้ไขข้อความได้ไม่สมบูรณ์ ด้วยเหตุนี้วิธีการตรวจให้คะแนนแบบรายตัวเลือกจึงมีค่าอำนาจจำแนกมากกว่า และค่าฟังก์ชันสารสนเทศของข้อสอบของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ที่ใช้วิธีการตรวจให้คะแนนแบบรายตัวเลือกจำแนกผู้เรียนกลุ่มเก่งและกลุ่มอ่อนออกจากกันได้สูงกว่าวิธีการตรวจให้คะแนนแบบ 0-1 ซึ่งค่าฟังก์ชันสารสนเทศของข้อสอบที่ระดับความสามารถต่ำและระดับความสามารถปานกลาง แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ข้อเสนอแนะ

ข้อเสนอแนะเพื่อการนำผลการวิจัยไปใช้

1. ในบริบทของเนื้อหาที่สามารถสร้างข้อสอบแบบถูกผิดหลายตัวเลือกได้ ควรสร้างแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ เพราะรูปแบบของแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ ผู้สอบต้องแก้ไขข้อความที่ผิดให้ถูกต้อง จึงทำให้ข้อสอบยากกว่า สามารถจำแนกผู้เรียนได้ดีกว่าและทำให้โอกาสในการเดาคำตอบน้อยลง

2. แบบสอบถูกผิดหลายตัวเลือก ซึ่งเป็นแบบสอบที่ต้องพิจารณาตัวเลือกทุกตัวแต่ไม่ต้องมีการแก้ไข ข้อความที่ผิดให้ถูกต้อง สามารถตรวจให้คะแนนได้ทั้งแบบ 0-1 หรือแบบรายตัวเลือก แต่ทั้งนี้ขึ้นอยู่กับความเหมาะสมของเนื้อหา จุดประสงค์ของการวัดผลหรือระดับความสามารถของผู้เรียนด้วย

3. แบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ ซึ่งเป็นแบบสอบที่ต้องพิจารณาตัวเลือกทุกตัวและมีการแก้ไขข้อความที่ผิดให้ถูกต้อง ควรใช้วิธีการตรวจให้คะแนนแบบรายตัวเลือก เพราะวิธีการตรวจให้คะแนนแบบรายตัวเลือกสามารถอ่านจำแนกผู้เรียนได้ดีกว่า

ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

1. ควรมีการศึกษาเปรียบเทียบแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่กับกลุ่มตัวอย่างที่มีระดับความสามารถแตกต่างกันโดยจำแนกออกเป็นกลุ่มเก่ง ปานกลางและอ่อน เพื่อเปรียบเทียบว่าแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่เหมาะสมหรือไม่เหมาะสมกับผู้เรียนกลุ่มใด

2. ควรนำแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ไปทดลองใช้กับวิธีการตรวจให้คะแนนแบบอื่น เช่น วิธี Pass-3 วิธี Count-2 วิธี Count-3

3. ควรพัฒนาแบบสอบถูกผิดหลายตัวเลือกกับแบบสอบถูกผิดหลายตัวเลือกปรับปรุงใหม่ในรายวิชาทางด้านสังคมศาสตร์ ทั้งนี้เนื่องจากส่วนใหญ่แบบสอบถูกผิดหลายตัวเลือกจะใช้ในรายวิชาทางด้านวิทยาศาสตร์

รายการอ้างอิง

ภาษาไทย.

เฉลิมฤทธิ์ แก้วกล้าหาญ. (2550). การวิเคราะห์ข้อถกเถียงงานวิจัยที่เปรียบเทียบผลการวิเคราะห์ข้อสอบระหว่าง ทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ. *OJED* 2(1), 207-216.

โชติกา ภาษิมล. (2545). การเปรียบเทียบคุณภาพของแบบสอบเลือกตอบเมื่อตรวจด้วยวิธีการให้คะแนน ความรู้บางส่วนกับวิธีประเพณีนิยม. *วารสารครุศาสตร์* 31(2), 24-40.

โชติกา ภาษิมล. (2556). *การวัดและประเมินผลการศึกษา*. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
ณัฐภรณ์ หลาวทอง. (2548). *การวัดและประเมินผลการศึกษา*. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์ มหาวิทยาลัย.

นภา หลิมรัตน์. (2552). *เปรียบเทียบผลการวิเคราะห์แบบทดสอบ True-False Question กับ True False Correction Question*. สืบค้นจาก <http://teachingresources.psu.ac.th/document/2551/limrat/4True-False.pdf>

บุญชม ศรีสะอาด. (2540). *การวิจัยทางการวัดผลและประเมินผล*. กรุงเทพมหานคร: สุวีริยาสาส์น.

พัชรี มีวรรณ. (2540). *ผลของจำนวนตัวเลือกที่มีต่อค่าฟังก์ชันสารสนเทศของแบบสอบเลือกตอบแบบ ถูกผิดในระดับการศึกษาที่ต่างกัน* (วิทยานิพนธ์ปริญญาโทบริหารศึกษาศาสตร์). สาขาการวัดและประเมินผล การศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

เยาวดี ราชชัยกุล วิบูลย์ศรี. (2552). *การวัดผลและการสร้างแบบสอบผลสัมฤทธิ์* (พิมพ์ครั้งที่ 8). กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

- วัชรภรณ์ จิตรมาศ. (2538). *การเปรียบเทียบคุณภาพแบบทดสอบเลือกตอบชนิดถูกผิดหลายตัวเลือกที่มีจำนวนตัวเลือกและวิธีการให้คะแนนต่างกัน* (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). สาขาการวัดผลการศึกษา บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- ศิริชัย กาญจนวาสี. (2552). *ทฤษฎีการทดสอบแบบดั้งเดิม* (พิมพ์ครั้งที่ 5). กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สุรชัย มีชาญ. (2545). การให้คะแนนสำหรับความรู้บางส่วนโดยใช้ข้อสอบถูกผิดหลายตัวเลือกและข้อสอบถูกผิดรายขั้นตอน. *วารสารศึกษาศาสตร์*, 14(1), 1-9.
- สุรชัย มีชาญ. (2545). การให้คะแนนสำหรับความรู้บางส่วน (Partial Knowledge) โดยใช้ข้อสอบเลือกตอบถูกผิดหลายตัวเลือก (Multiple True-False Item) และข้อสอบเลือกตอบถูกผิดทุกขั้นตอน (Sequential True-False Item). *วารสารสงขลานครินทร์ ฉบับสังคมศาสตร์และมนุษยศาสตร์*, 8(2), 133-146.

ภาษาอังกฤษ

- Albanese, M. A., & Sabers, D. L. (1988). Multiple true-false item: A study of inter-Item correlations, scoring alternatives, and reliability estimation. *Journal of Education Measurement*, 25, 111-123.
- Francis, D. C., Wesley C. D., & Nine, H. S. (1943). A study of the relative values of two modifications of the true-false test. *The Journal of Educational Research*, 36(7), 517-527.
- Frisbie, D. A. (1992). The multiple true-false item format: A status review. *Educational Measurement: Issues and Practice*, 11, 21-26.
- Frisbie, D. A., & Sweeney, D. C. (1982). "The relative merits of multiple true-false tests". *Journal of Education Measurement*, 19(1), 29-35.
- Fu-Ju,T, & Hoi K. S. (1993). A brief report on a comparison of six scoring methods for multiple true-false items. *Educational and Psychological Measurement*, 53, 399-405.
- Haladyna, T. M., & Downing, S. M. (1989). A taxonomy of multiple-choice item writing rules. *Applied Measurement in Education*, 2(1), 37-50.
- Howard Y. M, & Francis D. C. (1926). A modified form of the true-false test. *The Journal of Educational Research*, 14(3), 213-224.
- Leila J. (2014). The Comparison between multiple-choice (MC) and multiple true-false (MTF) test formats in Iranian intermediate EFL learners' vocabulary learning. *Procedia - Social and Behavioral Sciences*, 98, 84-788.