

การเปรียบเทียบคุณภาพของวิธีการตรวจให้คะแนนความรู้บางส่วน:
การพัฒนาวิธีประยุกต์การให้คะแนนแบบตัดตัวลง-เลือกตัวถูก

A Comparison of Partial Knowledge Scoring Method Quality:
A Development of Modified Number Right-Elimination Scoring Method

ณัฐภรณ์ เลขะวัฒนพงษ์¹

Natthaporn Lekhawattthanapong

natthaporncu@gmail.com

ศิริเดช สุชีวะ²

Siridej Sujiva

Received: 25-03-2019

Revised: 10-06-2019

Accepted: 13-06-2019

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาวิธีการให้คะแนนความรู้บางส่วนวิธีประยุกต์ตัดตัวลง-เลือกตัวถูก และเปรียบเทียบความยาก อำนาจจำแนกของข้อสอบ ความตรงเชิงโครงสร้างของแบบสอบ ความเที่ยงของแบบสอบ ฟังก์ชันสารสนเทศของแบบสอบ และความแม่นยำในการประมาณค่าความสามารถของผู้สอบที่ได้จากวิธีประยุกต์คัมภ์ วิธีตัดตัวลง-เลือกตัวถูก และวิธีประยุกต์ตัดตัวลง-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก ตัวอย่างวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 จำนวน 1,251 คน ในโรงเรียนเขตกรุงเทพมหานคร เครื่องมือวิจัย ได้แก่ แบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาเคมี เรื่อง พันธะโคเวเลนต์ เป็นแบบสอบหลายตัวเลือก จำนวน 40 ข้อ แบ่งเป็นข้อสอบ 4 ตัวเลือกและ 5 ตัวเลือกอย่างละ 20 ข้อ วิเคราะห์ข้อมูลความยาก อำนาจจำแนกของข้อสอบ และฟังก์ชันสารสนเทศของแบบสอบโดยใช้โปรแกรม IRTPRO 4 Student วิเคราะห์ความตรงเชิงโครงสร้างโดยใช้โปรแกรม Mplus วิเคราะห์ความเที่ยงโดยใช้โปรแกรม SPSS และเปรียบเทียบความแตกต่างของความเที่ยงโดยใช้ Feldt test

ผลการวิจัย พบว่า (1) วิธีการให้คะแนนความรู้บางส่วนที่ต่างกันส่งผลให้ความยากของข้อสอบแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยวิธีประยุกต์ของคัมภ์มีค่าความยากสูงสุด นอกจากนี้ วิธีการให้คะแนนความรู้บางส่วนและจำนวนตัวเลือกที่ต่างกันส่งผลต่ออำนาจจำแนกของข้อสอบแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยวิธีประยุกต์คัมภ์มีค่าอำนาจจำแนกสูงสุด ส่วนวิธีตัดตัวลง-เลือกตัวถูกและวิธีประยุกต์ตัดตัวลง-เลือกตัวถูกมีค่าอำนาจจำแนกไม่แตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (2) วิธีประยุกต์คัมภ์และวิธีประยุกต์ตัดตัวลง-เลือกตัว เมื่อข้อสอบมี 4 และ 5 ตัวเลือก มีความตรงเชิงโครงสร้าง ส่วนวิธีตัดตัวลง-เลือกตัวถูก เมื่อข้อสอบมี 5 ตัวเลือกมีความตรงเชิงโครงสร้างแต่ 4 ตัวเลือกไม่มีความตรงเชิง

¹ นิสิตมหาบัณฑิตสาขาวิชาการวัดและประเมินการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
Division of Educational Measurement and Evaluation, Educational Research and Psychology Department, Faculty of Education, Chulalongkorn University

² อาจารย์ประจำสาขาวิชาการวัดและประเมินการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
Division of Educational Measurement and Evaluation, Educational Research and Psychology Department, Faculty of Education, Chulalongkorn University

โครงสร้าง (3) ความเที่ยงของทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก ไม่แตกต่างกัน โดยวิธีประยุกต์ตัดตัวเลือก-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกมีค่าความเที่ยงสูงสุด (4) วิธีตัดตัวเลือก-เลือกตัวถูก เมื่อข้อสอบมี 5 ตัวเลือกให้สารสนเทศสูงที่สุดในช่วงระดับความสามารถต่ำ วิธีประยุกต์ตัดตัวเลือก-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกให้สารสนเทศสูงที่สุดในช่วงระดับความสามารถปานกลางและทั้ง 3 วิธีให้สารสนเทศใกล้เคียงกันในช่วงระดับความสามารถสูง วิธีประยุกต์ตัดตัวเลือก-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกมีความแม่นยำในการประมาณค่าความสามารถของผู้สอบสูงสุด และ (5) วิธีประยุกต์ตัดตัวเลือก-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกมีคุณภาพสูงที่สุด รองลงมา คือ วิธีประยุกต์ตัดตัวเลือก-เลือกตัวถูก เมื่อข้อสอบมี 5 ตัวเลือก

คำสำคัญ: วิธีการให้คะแนนความรู้บางส่วน, วิธีประยุกต์ของคูมพ์ วิธีตัดตัวเลือก-เลือกตัวถูก วิธีประยุกต์ตัดตัวเลือก-เลือกตัวถูก

Abstract

The purposes of this research were to develop a modified number right-elimination scoring method; and to compare difficulty index, discriminant index, construct validity, reliability coefficient, test information function, and accuracy in estimating examinees' ability among using the modified Coomb's method, the number right-elimination method (NRE), and the modified number right-elimination method (M-NRE) with 4 choices and 5 choices. Research subjects were 1,251 tenth grade students from secondary schools in Bangkok. Research instrument was covalent bond tests with 40 multiple-choice items comprising 20 items for 4 choices and 20 items for 5 choices. Data analysis was conducted by using the following programs: IRTPRO 4 Student for analyzing difficulty index, discriminant index and test information function; Mplus for analyzing construct validity; and SPSS for analyzing reliability coefficient which was then compared by using the Feldt test.

Research results revealed that (1) different scoring methods result in different difficulty index of the test at the .05 level of statistical significance. The modified Coomb's method yielded the highest difficulty index. In addition, partial knowledge scoring method and number of choices affect different discriminant index at the .05 level of statistical significance. The discriminant index of the modified Coomb's method was the highest, while there was no statistically significant difference between the NRE method and the M-NRE method; (2) The modified Coomb's method, the M-NRE method with 4 choices and 5 choices and the NRE method with 5 choices had construct validity, whereas the NRE method with 4 choices had no construct validity; (3) Among the 3 methods, there was no difference in reliability when the test had 4-5 choices. However, the M-NRE method with 4 choices had the highest reliability; (4) The NRE method with 5 choices had the highest information function in low ability group. The M-NRE method with 4 choices had the highest information function in moderate ability group, and all of the 3 methods had close information function in high ability group. The M-NRE method with 4 choices had the highest accuracy in estimating examinees' ability, and (5) The M-NRE method with 4 choices had the highest quality followed by the M-NRE method with 5 choices.

Keywords: Partial Knowledge Scoring Method, Modified Coomb's Scoring Method, Number Right-Elimination Scoring Method, Modified Number Right-Elimination Scoring Method

ความเป็นมาและความสำคัญของปัญหา

การวัดและการประเมินผลการเรียนรู้ที่มีบทบาทมากในปัจจุบัน คือ การใช้แบบสอบหลายตัวเลือก สำหรับการวัดความรู้ความสามารถและผลสัมฤทธิ์ทางการเรียนของผู้เรียน (Chang, Lin, & Lin, 2007; รัตนา ไชยตรี, 2546) โดยแบบสอบหลายตัวเลือกมีความสะดวกในการนำไปใช้หลายประการ เช่น วัดพฤติกรรมการเรียนรู้ได้หลากหลายและครอบคลุมเนื้อหา มีความเป็นปรนัยในการตรวจให้คะแนน ใช้เวลาในการตรวจน้อย สามารถนำผลการตอบไปวิเคราะห์ข้อมูลทางสถิติได้ เหมาะสำหรับการสอบที่มีผู้สอบจำนวนมาก เป็นต้น (โชติกา ภาษีผล ณัฐภรณ์ หลาวทอง และกมลวรรณ ตั้งธณานนท์, 2558)

อย่างไรก็ตาม แบบสอบหลายตัวเลือกก็มีจุดอ่อนหลายประการเช่นกัน เช่น ข้อสอบออกยาก ใช้เวลาในการออกข้อสอบนาน ไม่สามารถวัดความคิดสร้างสรรค์ได้ ผู้สอบมีโอกาสดำเนินการเดาคำตอบได้ ซึ่งจะทำให้คะแนนที่ได้ไม่สะท้อนความสามารถที่แท้จริงของผู้เรียนและไม่สามารถให้สารสนเทศเพียงพอที่จำแนกผู้เรียนที่มีความรู้จริง มีความรู้บางส่วน หรือไม่มีความรู้ (เอมอร์ จังศิริพรบูรณ์, 2545b) จากข้อจำกัดของแบบสอบหลายตัวเลือก ดังกล่าว จึงมีนักวิชาการได้ศึกษาวิธีการแก้ปัญหาการเดาคำตอบของผู้สอบและการให้สารสนเทศในการจำแนกผู้เรียน โดยใช้วิธีการให้คะแนนความรู้บางส่วน (Partial Knowledge) (โชติกา ภาษีผล, 2554) ดังนั้น จึงเป็น

ที่มาของการศึกษาเกี่ยวกับวิธีการให้คะแนนความรู้บางส่วนว่าวิธีใดมีคุณภาพมากกว่ากัน วิธีการให้คะแนนความรู้บางส่วนในปัจจุบันมีหลายวิธี ซึ่งในงานวิจัยนี้ผู้วิจัยสนใจศึกษาวิธีประยุกต์คุมบ์ วิธีตัดตัวลง-เลือกตัวถูกและวิธีประยุกต์ตัดตัวลง-เลือกตัวถูก

วิธีประยุกต์คุมบ์เป็นวิธีที่พัฒนามาจากวิธีให้คะแนนแบบคุมบ์ดั้งเดิม เป็นวิธีที่ค่าอำนาจจำแนกเฉลี่ย ฟังก์ชันสารสนเทศของแบบสอบ ประสิทธิภาพสัมพัทธ์เฉลี่ยของแบบสอบ ความตรงตามเกณฑ์สัมพัทธ์และมีคุณภาพสูงสุดเมื่อเปรียบเทียบกับวิธีอื่น ๆ (สุพจน์ เกิดสุวรรณ, 2545; เอมอร์ จังศิริพรภรณ์, 2545a) แต่วิธีประยุกต์การให้คะแนนของคุมบ์เป็นวิธีที่ทำให้นักเรียนมีความสับสนและไม่สอดคล้องกับความเป็นจริง เนื่องจากวิธีนี้เป็นการประเมินความสามารถในการตอบข้อใดข้อหนึ่งที่ขัดแย้งกับในความเป็นจริงที่ผู้เรียนควรจะถูกฝึกให้แก้ปัญหาเพื่อหาคำตอบที่ถูกมากกว่า (Lau, Lau, Hong, & Usop, 2014)

วิธีตัดตัวลง-เลือกตัวถูก (Number right elimination: NRE) วิธีนี้เป็นรูปแบบที่ส่งเสริมให้นักเรียนได้ฝึกกลยุทธ์ในการทำข้อสอบ และส่งเสริมให้นักเรียนแก้ปัญหาเพื่อหาคำตอบที่ถูกต้อง นอกจากนี้ จากงานวิจัย พบว่า วิธีนี้มีความเที่ยงสูง มีโอกาสในการเดาน้อย ค่าความคลาดเคลื่อนมาตรฐานของการวัดต่ำกว่าวิธีอื่น ๆ และมีประสิทธิภาพในการประมาณค่าความสามารถของผู้เรียนมากกว่าวิธีการตัดตัวลงและวิธีประเพณีนิยม สามารถวัดความเข้าใจที่คลาดเคลื่อนได้และสามารถจำแนกผู้เรียนที่มีความรู้สมบูรณ์ มีความรู้บางส่วน ไม่มีความรู้ มีความรู้คลาดเคลื่อนบางส่วน และมีความรู้คลาดเคลื่อนทั้งหมดได้ (Lau, Lau, Hong, & Usop, 2011; 2014) อย่างไรก็ตาม วิธีตัดตัวลง-เลือกตัวถูก จะมีคะแนนอยู่ในช่วงติดลบ ซึ่งเกิดจากการถูกหัก -3 คะแนน จากการตัดตัวถูกออกโดยเข้าใจว่าเป็นตัวลง ซึ่งผู้วิจัยเห็นว่าไม่เหมาะสมเนื่องจากข้อที่ได้คะแนนติดลบอาจจะถูกนำไปหักกับคะแนนในข้ออื่น ๆ ได้ ซึ่งจะส่งผลกระทบต่อคะแนนรวมได้ (สุพจน์ เกิดสุวรรณ, 2545; เอมอร์ จังศิริพรภรณ์, 2545a) ทำให้ไม่ได้สารสนเทศที่สะท้อนความรู้ที่แท้จริงของผู้เรียนและไม่สามารถจำแนกผู้เรียนได้ ผู้วิจัยจึงพัฒนาวิธีตัดตัวลง-เลือกตัวถูก (NRE) ของ Lau, Lau, Hong, & Usop (2011) โดยการปรับ

ไม่ให้คะแนนมีค่าติดลบจากการตัดตัวถูกโดยเข้าใจว่าเป็นตัวลง ซึ่งเรียกว่า วิธีประยุกต์ตัดตัวลง-เลือกตัวถูก

วิธีประยุกต์ตัดตัวลง-เลือกตัวถูกเป็นวิธีการให้คะแนนที่ผู้วิจัยพัฒนามาจากวิธีตัดตัวลง-เลือกตัวถูก (NRE) ของ Lau, Lau, Hong, & Usop (2014) วิธีการตอบจะเหมือนกับวิธีตัดตัวลง-เลือกตัวถูกแต่จะมีวิธีการให้คะแนนแตกต่างกันโดยวิธีนี้มีการปรับไม่ให้คะแนนมีค่าติดลบโดยมีการให้คะแนนแบบเสริมแรง เพื่อเป็นการหลีกเลี่ยงการให้คะแนนแบบลงโทษ จากที่กล่าวมาข้างต้นผู้วิจัยจึงสนใจศึกษาคุณภาพของวิธีการตรวจให้คะแนนความรู้บางส่วนทั้ง 3 วิธีโดยวัดคุณภาพจากความยากและอำนาจจำแนกของข้อสอบ ความตรงเชิงโครงสร้างของแบบสอบ ความเที่ยงของแบบสอบ ฟังก์ชันสารสนเทศของแบบสอบ และความแม่นยำในการประมาณค่าความสามารถของผู้สอบที่ได้จากวิธีประยุกต์คุมบ์ วิธีตัดตัวลง-เลือกตัวถูก และวิธีประยุกต์ตัดตัวลง-เลือกตัวเพื่อเปรียบเทียบคุณภาพของวิธีการให้คะแนนความรู้บางส่วนทั้ง 3 วิธี

นอกจากนี้ จำนวนตัวเลือกในแบบสอบหลายตัวเลือกก็มีผลต่อคุณภาพของแบบสอบและความสามารถในการจำแนกผู้เรียนได้ (Nitko, 2004) ซึ่งโดยทั่วไปแบบสอบหลายตัวเลือกจะมี 3, 4 และ 5 ตัวเลือก (McMillan, 2014) ซึ่งจากการศึกษาวิจัยเกี่ยวกับจำนวนตัวเลือก พบว่า งานวิจัยส่วนใหญ่แนะนำให้ใช้แบบสอบที่มีจำนวน 4 ตัวเลือกและ 5 ตัวเลือกในการวัดผลการเรียนรู้ (Rodriguez, 2005) การเพิ่มจำนวนตัวเลือกมากขึ้นจะสามารถแก้ปัญหาการเดาคำตอบของผู้สอบได้ ดังนั้น ถ้าใช้ข้อสอบที่มี 3 ตัวเลือกจะสามารถเดาคำตอบได้ง่ายและทำให้ความยากและอำนาจจำแนกลดลง (MacCann, 2004; Rodriguez, 2005; รณิดา เขยขุ่ม, 2551; อานภาพ เลขะกุล, 2559) ซึ่งจะส่งผลกระทบต่อวัดผลการเรียนรู้ที่ไม่แม่นยำ และแบบสอบที่มี 4 ตัวเลือกและ 5 ตัวเลือก ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (วารุณี ปิตธวัชชัย, 2514) แต่ในงานวิจัยดังกล่าวไม่ได้ศึกษาการให้คะแนนความรู้บางส่วน และจากการศึกษาวิจัยยังไม่ค่อยมีผู้ที่ศึกษาเกี่ยวกับวิธีการให้คะแนนความรู้บางส่วนที่มีจำนวนตัวเลือกต่างกัน ผู้วิจัยจึงสนใจศึกษาทั้ง 2 ตัวแปรนี้ ซึ่งวิธีการให้คะแนนความรู้บางส่วนเป็นวิธีที่แก้ปัญหาในด้านการเดาคำตอบของผู้สอบและการให้สารสนเทศในการจำแนกผู้เรียน

ส่วนการเพิ่มจำนวนตัวเลือกก็สามารถลดโอกาสในการเดาได้ (MacCann, 2004) สามารถให้ข้อมูลในการวินิจฉัยความเข้าใจที่คลาดเคลื่อนได้ (Rodriguez, 2005) และสามารถระบุความสามารถของผู้สอบได้ (Bahrami, 2011) ดังนั้น ถ้านำวิธีการให้คะแนนความรู้บางส่วนกับจำนวนตัวเลือกมาประยุกต์ใช้ร่วมกันน่าจะสามารถแก้ปัญหาของการใช้แบบสอบหลายตัวเลือกได้อย่างสมบูรณ์และสามารถใช้แบบสอบหลายตัวเลือกได้อย่างมีประสิทธิภาพ

วัตถุประสงค์

1. เพื่อพัฒนาวิธีการให้คะแนนความรู้บางส่วนวิธีประยุกต์ตัดตัวลง-เลือกตัวถูก
2. เพื่อเปรียบเทียบความยากและอำนาจจำแนกของข้อสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก
3. เพื่อตรวจสอบความตรงเชิงโครงสร้างของแบบสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก
4. เพื่อเปรียบเทียบความเที่ยงของแบบสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก

5. เพื่อเปรียบเทียบฟังก์ชันสารสนเทศของแบบสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก
6. เพื่อเปรียบเทียบความแม่นยำในการประมาณค่าความสามารถของผู้สอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก

นิยามศัพท์

1. วิธีการตรวจให้คะแนนความรู้บางส่วน หมายถึง การให้คะแนนที่ใกล้เคียงกับความสามารถที่แท้จริงของผู้ตอบ ซึ่งในงานวิจัยนี้ศึกษาวิธีการตรวจให้คะแนนความรู้บางส่วน 3 วิธี ได้แก่ วิธีประยุกต์คัมบ์ (Modified Coombs) วิธีตัดตัวลง-เลือกตัวถูก (NRE) และวิธีประยุกต์ตัดตัวลง-เลือกตัวถูก (M-NRE)
2. วิธีประยุกต์ของคัมบ์ (The modified Coomb's method) เป็นวิธีที่มีคำตอบถูกเพียงคำตอบเดียว โดยให้ผู้สอบตัดตัวลงที่ไม่ใช่คำตอบถูกออกไปให้มากที่สุด มีสูตรการให้คะแนน ดังนี้

$$X = nC$$

(เอมอร์ จังศิริพรปรกรณ์, 2545a)

เมื่อ	X	คือ	คะแนนที่ได้จากการตอบข้อสอบ 1 ข้อ
	n	คือ	จำนวนตัวลงที่ถูกตัดออก
	C	คือ	ค่าความถูกต้องของการตัดตัวลง ซึ่งมี 2 ค่า ดังนี้
	C = 1	เมื่อไม่มีตัวเลือกถูกอยู่ในตัวลงที่ตัดออก	
	C = 0	เมื่อมีตัวเลือกถูกอยู่ในตัวลงที่ตัดออก	

ตาราง 1 การให้คะแนนวิธีประยุกต์ของคัมบ์ เมื่อข้อสอบมี 4 ตัวเลือก

รูปแบบการตอบ	จำนวนตัวเลือกที่ผู้สอบเลือก	ลักษณะการตอบ	จำนวนตัวเลือกที่เลือกถูกต้องตรงกับคำตอบ	คะแนน
1	3	FFF	ตัดตัวลงตรงกับจำนวนตัวลงในข้อนั้น 3 ตัว	3
2	2	FF	ตัดตัวลงตรงกับจำนวนตัวลงในข้อนั้น 2 ตัว	2
3	1	F	ตัดตัวลงตรงกับจำนวนตัวลงในข้อนั้น 1 ตัว	1
4	3	TFF	ตัดตัวลงตรงกับจำนวนตัวลงในข้อนั้น 2 ตัว	0
5	2	TF	ตัดตัวลงตรงกับจำนวนตัวลงในข้อนั้น 1 ตัว	0
6	1	T	ตัดตัวลงไม่ตรงกับจำนวนตัวลงในข้อนั้น	0

หมายเหตุ F คือ ตัวเลือกที่เป็นตัวลง , T คือ ตัวเลือกที่เป็นตัวถูก (โชติกา ภาชีผล, 2554)

3. วิธีตัดตัวลง-เลือกตัวถูก (The number right elimination testing: NRET) เป็นวิธีที่ผสมผสานระหว่างวิธีการตัดตัวลงของคูมบ์ และวิธีประเพณีนิยม โดยให้ผู้สอบตัดตัวลงที่แน่ใจว่าไม่ใช่คำตอบออก เลือก 1 ตัวเลือกที่เป็นคำตอบที่ถูกต้อง โดยตัวลงที่ไม่แน่ใจให้ตอบ “ไม่แน่ใจ” ถ้าตัดตัวลงถูก 1 ตัว จะได้ 1 คะแนน ซึ่งตัวลงทั้งหมดจะมี $k-1$ ตัว เมื่อ k เป็นจำนวนตัวเลือก ซึ่งเป็นไปตามหลักการการให้คะแนนของคูมบ์ ถ้าเลือกคำตอบถูกต้อง จะได้ 1 คะแนน ซึ่ง

เป็นไปตามหลักการการให้คะแนนแบบประเพณีนิยม ถ้าตัดตัวถูกโดยเข้าใจว่าเป็นตัวลงจะถูกหักคะแนน $-(k-1)$ เมื่อ k เป็นจำนวนตัวเลือก ซึ่งเป็นไปตามหลักการการให้คะแนนของคูมบ์ แต่ถ้าไม่แน่ใจให้เลือกตอบ “ไม่แน่ใจ” จะได้ 0 คะแนน การที่มีตัวเลือกไม่แน่ใจ เพื่อเป็นการเพิ่มตัวเลือกในการพิจารณาเลือกคำตอบและหลีกเลี่ยงการเดาของนักเรียน รายละเอียดการให้คะแนน มีดังนี้

- | | |
|---|-------------------------|
| 1) เลือกคำตอบที่ถูกต้อง | จะได้ 1 คะแนน |
| 2) ตัดตัวลงถูกต้อง 1 ตัว | จะได้ 1 คะแนน |
| 3) ถ้าตัดตัวถูกออกโดยเข้าใจว่าเป็นตัวลง | จะถูกหัก $-(k-1)$ คะแนน |
| 4) ถ้าตอบ “ไม่แน่ใจ” | จะได้ 0 คะแนน |

ตาราง 2 การให้คะแนนแบบตัดตัวลง - เลือกตัว เมื่อข้อสอบมี 4 ตัวเลือก

รูปแบบการตอบ	ลักษณะการตอบและการตรวจให้คะแนน						คะแนนรวม
	จำนวนตัวลงที่ตัด	คะแนนที่ได้	ตัวถูก	คะแนนที่ได้	จำนวนตัวเลือกที่ไม่แน่ใจ	คะแนนที่ได้	
1	3	3	เลือก	1	0	0	4
2	2	2	เลือก	1	1	0	3
3	1	1	เลือก	1	2	0	2
4	0	0	เลือก	1	3	0	1
5	3	3	ตัด	-3	0	0	0
6	2	2	ตัด	-3	1	0	-1
7	1	1	ตัด	-3	2	0	-2
8	0	0	ตัด	-3	3	0	-3
9	3	3	ไม่แน่ใจ	0	0	0	3
10	2	2	ไม่แน่ใจ	0	1	0	2
11	1	1	ไม่แน่ใจ	0	2	0	1
12	0	0	ไม่แน่ใจ	0	3	0	0

(Lau, Lau, Hong, & Usop, 2014)

4. วิธีประยุกต์ตัดตัวลง-เลือกตัวถูก (modified number right-elimination method : M-NRE) เป็นวิธีการให้คะแนนที่ผู้วิจัยพัฒนามาจากวิธีตัดตัวลง-เลือกตัวถูก ของ Lau, Lau, Hong, & Usop (2014) วิธีการตอบจะเหมือนกับวิธีตัดตัวลง-เลือกตัวถูก นั่นคือ ให้ผู้สอบตัดตัวลงที่แน่ใจว่า

ไม่ใช่คำตอบออก เลือก 1 ตัวเลือกที่เป็นคำตอบที่ถูกต้อง โดยตัวลงที่ไม่แน่ใจให้ตอบ “ไม่แน่ใจ” ดังนั้น จึงเป็นวิธีที่ผสมผสานระหว่างวิธีการตัดตัวลงของคูมบ์ และวิธีการให้คะแนนแบบประเพณีนิยม แต่จะมีวิธีการให้คะแนนแตกต่างกันโดยวิธีนี้มีแนวคิดในการพัฒนา คือ การปรับไม่ให้คะแนนมี

ค่าติดลบ ซึ่งเป็นการหลีกเลี่ยงการให้คะแนนแบบลงโทษ เนื่องจากคะแนนติดลบมีผลกระทบต่อคะแนนในข้ออื่นและคะแนนรวม (สุพจน์ เกิดสุวรรณ, 2545; เอมอร์ จังศิริพรภรณ์, 2545a) ทำให้ไม่ได้สารสนเทศที่สะท้อนความรู้ที่แท้จริงของผู้เรียนและไม่สามารถจำแนกผู้เรียนได้ โดยถ้าตัดตัวถูก 1 ตัว จะได้ 1 คะแนน ตัวถูกทั้งหมดจะมี $k-1$ ตัว เมื่อ k เป็นจำนวนตัวเลือก ซึ่งเป็นไปตามหลักการการให้คะแนนของคูมพ์

ถ้าเลือกคำตอบถูกต้อง จะได้ $k-1$ คะแนน เมื่อ k เป็นจำนวนตัวเลือก ซึ่งเป็นการให้รางวัลตรงข้ามกับการหักคะแนนของคูมพ์ ถ้าไม่แน่ใจให้เลือกตอบ “ไม่แน่ใจ” จะได้ 0 คะแนนในตัวเลือกนั้น ถ้าตัดตัวถูกโดยเข้าใจว่าเป็นตัวถูกหรือเลือกตัวถูกไม่ถูกต้องจะได้ 0 คะแนนในข้อนั้นทันที เนื่องจากเป็นความเข้าใจที่คลาดเคลื่อนของนักเรียน รายละเอียดการให้คะแนน มีดังนี้

- | | |
|--|-----------------------------|
| 1) เลือกคำตอบที่ถูกต้อง | จะได้ $k-1$ คะแนน |
| 2) ตัดตัวถูกถูกต้อง 1 ตัว | จะได้ 1 คะแนน |
| 3) ถ้าตอบ “ไม่แน่ใจ” | จะได้ 0 คะแนน |
| 4) ถ้าตัดตัวถูกโดยเข้าใจว่าเป็นตัวถูกหรือเลือกตัวถูกไม่ถูกต้อง | จะได้ 0 คะแนนในข้อนั้นทันที |

ตาราง 3 การประยุกต์การให้คะแนนแบบตัดตัวถูก-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือก

รูปแบบการตอบ	ลักษณะการตอบและการตรวจให้คะแนน						คะแนนรวม
	จำนวนตัวถูกที่ตัด	คะแนนที่ได้	ตัวถูก	คะแนนที่ได้	จำนวนตัวเลือกที่ไม่แน่ใจ	คะแนนที่ได้	
1	3	3	เลือก	3	0	0	6
2	2	2	เลือก	3	1	0	5
3	1	1	เลือก	3	2	0	4
4	0	0	เลือก	3	3	0	3
5	3	-	ตัด	0	0	0	0
6	2	-	ตัด	0	1	0	0
7	1	-	ตัด	0	2	0	0
8	0	-	ตัด	0	3	0	0
9	3	3	ไม่แน่ใจ	0	0	0	3
10	2	2	ไม่แน่ใจ	0	1	0	2
11	1	1	ไม่แน่ใจ	0	2	0	1
12	0	0	ไม่แน่ใจ	0	3	0	0

หมายเหตุ - หมายถึง ไม่มีการพิจารณาคะแนนในตัวเลือกนั้น

ประโยชน์ที่ได้รับจากการวิจัย

1. ทำให้ทราบวิธีการตรวจให้คะแนนความรู้บางส่วนที่มีคุณภาพ รวมถึงขยายองค์ความรู้เกี่ยวกับวิธีการตรวจให้คะแนนความรู้บางส่วน เพื่อแก้ปัญหาการเดาข้อสอบและการให้สารสนเทศในการจำแนกผู้เรียน
2. ทำให้ทราบว่าวิธีการให้คะแนนความรู้บางส่วนวิธีใดสะท้อนความสามารถที่แท้จริงของผู้เรียนและสามารถตัดสินผลการเรียนรู้ของผู้เรียนได้อย่างแม่นยำ

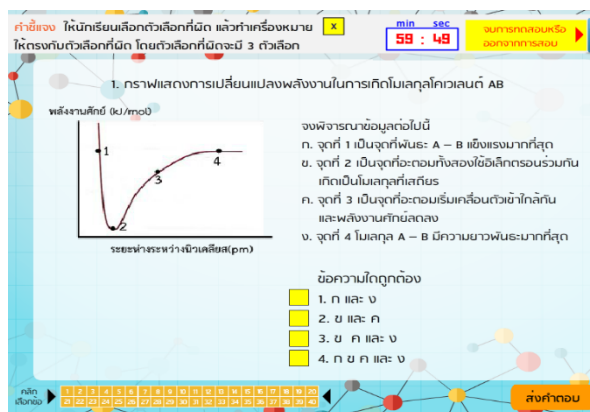
3. เพื่อให้สารสนเทศต่อผู้สอนในการพิจารณาเลือกจำนวนตัวเลือกที่เหมาะสมกับวิธีวัดผลการเรียนรู้ของผู้เรียน

วิธีการวิจัย

ประชากรและตัวอย่างวิจัย

ประชากรที่ใช้ในการวิจัย ได้แก่ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 จำนวน 36,272 คน ในโรงเรียนมัธยมศึกษาสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต 1 และเขต 2 สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (กระทรวงศึกษาธิการ, 2551)

ตัวอย่างวิจัยที่ใช้ในการวิจัยครั้งนี้ได้จากการสุ่มอย่างง่ายจากนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 ในโรงเรียนมัธยมศึกษา ในเขตกรุงเทพมหานคร สังกัดสำนักงานเขตพื้นที่



ภาพ 1 รูปแบบโปรแกรมของวิธีประยุกต์ของคูมบ์

การเก็บรวบรวมข้อมูล

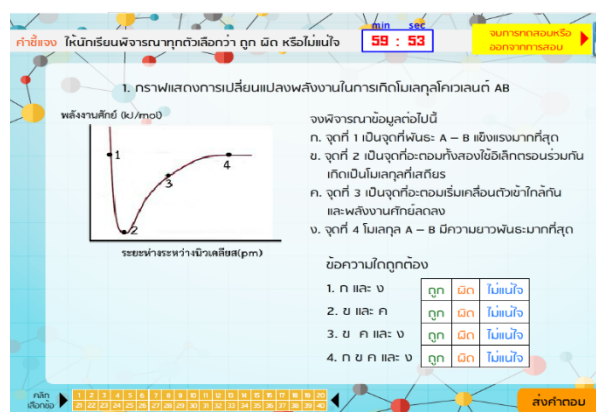
1. การเก็บข้อมูลผลการทดสอบ O-NET วิชาวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2559 เพื่อจำแนกระดับความสามารถของนักเรียนที่มีความสามารถระดับสูง ระดับปานกลาง และระดับต่ำ เพื่อคละระดับความสามารถของตัวอย่างวิจัยแต่ละกลุ่ม

2. การทดสอบด้วยระบบคอมพิวเตอร์ (Computer-based Test) เพื่อเก็บรวบรวมข้อมูลผลการตอบด้วยวิธีการให้คะแนนความรู้บางส่วน โดยนำแบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาเคมี เรื่อง พันธะโคเวเลนต์ จำนวน 40 ข้อ มาทดสอบกับนักเรียนจำนวน 1,251 คน ให้เวลาในการทำข้อสอบ 60

การศึกษามัธยมศึกษาเขต 1 และเขต 2 สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ ได้แก่ แบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาเคมีด้วยระบบคอมพิวเตอร์ (Computer-Based Test) เรื่อง พันธะโคเวเลนต์ ระดับชั้นมัธยมศึกษาปีที่ 4 จำนวน 3 ชุด ได้แก่ ชุดที่ 1 แบบสอบวิธีประยุกต์คูมบ์ ชุดที่ 2 แบบสอบวิธีตัดตัวลง-เลือกตัวถูก และชุดที่ 3 แบบสอบวิธีประยุกต์ตัดตัวลง-เลือกตัวถูก โดยแต่ละชุดจะมีข้อสอบจำนวน 2 ฉบับ ฉบับละ 20 ข้อ โดยแบบสอบฉบับที่ 1 เป็นข้อสอบที่มีจำนวน 4 ตัวเลือก และแบบสอบฉบับที่ 2 เป็นข้อสอบที่มีจำนวน 5 ตัวเลือก ซึ่งเป็นแบบสอบคู่ขนานกัน



ภาพ 2 รูปแบบโปรแกรมของวิธีตัดตัวลง-เลือกตัวถูก และวิธีประยุกต์ตัดตัวลง-เลือกตัวถูก

โดยตัวอย่างวิจัยแบ่งเป็น 3 กลุ่ม คือ กลุ่มที่ 1 ผู้สอบจะต้องตอบข้อสอบแบบวิธีประยุกต์คูมบ์ กลุ่มที่ 2 ผู้สอบจะต้องตอบข้อสอบโดยวิธีตัดตัวลง-เลือกตัวถูก และกลุ่มที่ 3 ผู้สอบจะต้องตอบข้อสอบโดยวิธีประยุกต์ตัดตัวลง-เลือกตัวถูก

การวิเคราะห์ข้อมูล

1. การวิเคราะห์ความยากและอำนาจจำแนกของข้อสอบ ที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก โดยวิเคราะห์ด้วยโปรแกรม IRTPRO4 Student และวิเคราะห์ด้วยโมเดล G-PCM

2. การวิเคราะห์ความตรงเชิงโครงสร้างของแบบสอบ ซึ่งในงานวิจัยนี้เลือกใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน ใน

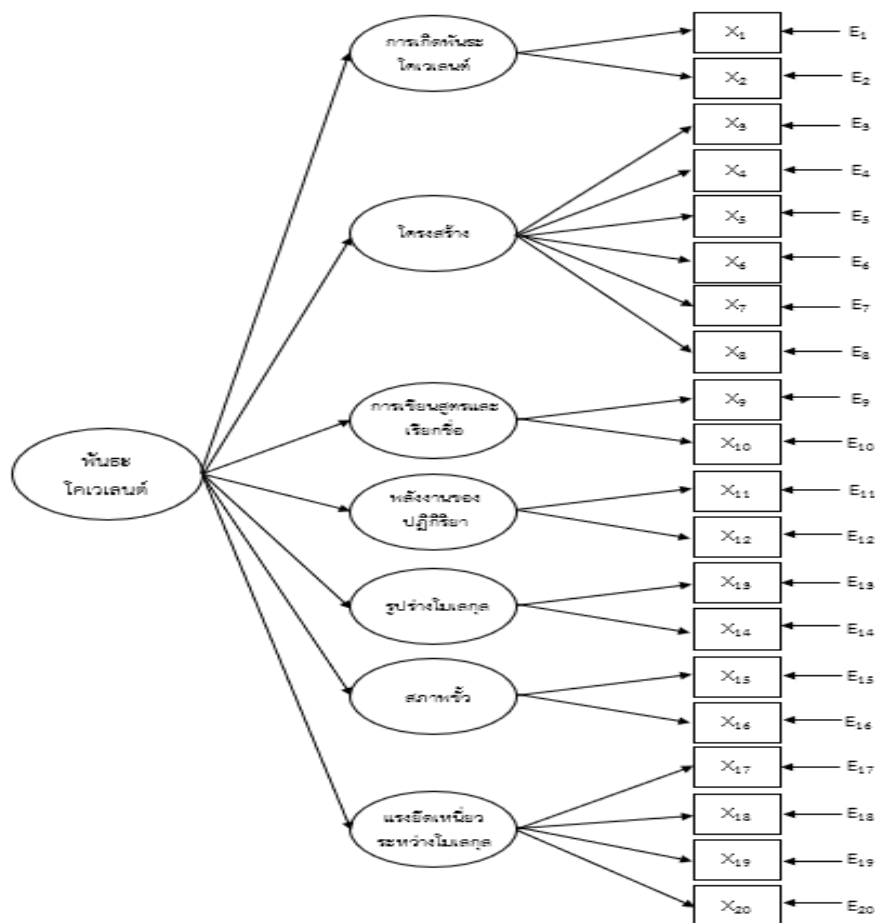
การพิจารณาความสอดคล้องระหว่างแบบสอบถามกับคุณลักษณะที่ต้องการวัดโดยการประมาณค่าพารามิเตอร์ของโมเดลใช้โปรแกรม Mplus และเป็นการประมาณค่าความน่าจะเป็นสูงสุด การตรวจสอบความสอดคล้องระหว่างโมเดลและข้อมูลทำได้โดยการพิจารณา χ^2 ต้องมีค่า p-value มากกว่า .05, χ^2/df ต้องไม่เกิน 2, RMSEA และ SRMR ต้องไม่เกิน .05 และ CFI ต้องมากกว่า .97 ซึ่งโครงสร้างของแบบสอบถามสำหรับการวิเคราะห์องค์ประกอบเชิงยืนยันของเนื้อหา เรื่อง พันธะโคเวเลนต์ ประกอบด้วย 7 คุณลักษณะที่ต้องการวัดและตัวแปรสังเกตได้เป็นข้อสอบจำนวน 20 ข้อ (X_1-X_{20}) แสดงโครงสร้างโมเดล ดังภาพ 3

3. การวิเคราะห์ความเที่ยงของแบบสอบ ที่ได้จากวิธีประยุกต์คูมบ์ วิธีตัดตัวลง-เลือกตัวถูก และวิธีประยุกต์ตัดตัวลง-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก โดยใช้ความเที่ยงแบบความสอดคล้องภายใน วิธีสัมประสิทธิ์

แอลฟาครอนบาค (Cronbach's Alpha) และเปรียบเทียบความแตกต่างของความเที่ยงของคะแนนที่ได้จากทั้ง 3 วิธี ทำได้โดยใช้โปรแกรมสำเร็จรูป Feldt test (For large samples, $N > 99$) (Feldt, 1969)

4. การวิเคราะห์ฟังก์ชันสารสนเทศของแบบสอบ ที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก โดยวิเคราะห์ด้วยโปรแกรม IRTPRO4 Student เพื่ออธิบายความสัมพันธ์ระหว่างความสามารถของผู้ตอบและโอกาสในการเลือกตอบแต่ละรายการที่กำหนด ซึ่งการวิเคราะห์ฟังก์ชันสารสนเทศของแบบสอบวิเคราะห์ด้วยโมเดล G-PCM (Muraki, 1992)

5. การวิเคราะห์ความแม่นยำในการประมาณค่าความสามารถของผู้สอบ ที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก โดยวิเคราะห์จากอัตราส่วนของค่าฟังก์ชันสารสนเทศเฉลี่ยของแต่ละวิธี



ภาพ 3 โครงสร้างของแบบสอบถามสำหรับการวิเคราะห์องค์ประกอบเชิงยืนยัน

ผลการวิจัยและอภิปรายผล

ตอนที่ 1 ผลการวิเคราะห์ความยากและอำนาจจำแนกของข้อสอบ

การวิเคราะห์ความแตกต่างระหว่างค่าความยากเฉลี่ยของข้อสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือก และ 5 ตัวเลือกโดยใช้การวิเคราะห์ 2-way ANOVA พบว่า ปฏิสัมพันธ์ระหว่างวิธีการให้คะแนนและจำนวนตัวเลือก ไม่มีนัยสำคัญทางสถิติที่ระดับ .05, $F(2, 114) = 0.584, p = .559$,

$\eta_p^2 = .010$ โดยเมื่อพิจารณาอิทธิพลหลัก พบว่า วิธีการให้คะแนนความรู้บางส่วนที่ต่างกันส่งผลต่อค่าความยากของข้อสอบแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05, $F(2, 114) = 83.154, p < .001, \eta_p^2 = .593$ ส่วนจำนวนตัวเลือกที่ต่างกันส่งผลต่อค่าความยากของข้อสอบ แต่ไม่มีนัยสำคัญทางสถิติที่ระดับ .05, $F(1, 114) = 1.185, p = .279, \eta_p^2 = .010$ ดังตาราง 4

ตาราง 4 ผลการวิเคราะห์ความแปรปรวนแบบสองทางของความยากของข้อสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก

แหล่งของความแปรปรวน	SS	df	MS	F	p	η_p^2
วิธีการให้คะแนน	167.323	2	83.661	83.154	< .001	.593
จำนวนตัวเลือก	1.192	1	1.192	1.185	.279	.010
ปฏิสัมพันธ์ระหว่างวิธีการให้คะแนนและจำนวนตัวเลือก	1.175	2	.587	.584	.559	.010
ความคลาดเคลื่อน	114.695	114	1.006			
รวม	296.191	120				

Levene's test $F(5, 114) = 0.816, p = .541$

$R^2 = .597$, Adjusted $R^2 = .579$

จากการวิเคราะห์อิทธิพลหลัก พบว่า วิธีการให้คะแนนความรู้บางส่วนที่ต่างกันส่งผลต่อค่าความยากของข้อสอบ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05, $F(2, 114) = 83.154, p < .001, \eta_p^2 = .593$ ซึ่งความแปรปรวนของความยากของข้อสอบที่ได้จากวิธีการให้คะแนนความรู้บางส่วนแต่ละวิธีแตกต่างกัน แต่ไม่มีนัยสำคัญทางสถิติที่ระดับ .05, $F(5, 114) = 0.816, p = .541$ จึงทำการทดสอบภายหลังโดยใช้สถิติ Bonferroni เมื่อพิจารณาความยากของข้อสอบที่ได้จากวิธีการให้คะแนนความรู้ที่แตกต่างกัน พบว่า วิธีประยุกต์ของคัมภ์

($M = 0.597, SD = 0.159$) มีค่าความยากสูงกว่าวิธีตัดตัวลง-เลือกตัวถูก ($M = -1.981, SD = 0.159$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05, $F = 2.578, p < .001$ วิธีประยุกต์ตัดตัวลง-เลือกตัวถูก ($M = 0.443, SD = 0.159$) มีค่าความยากสูงกว่าวิธีตัดตัวลง-เลือกตัวถูก ($M = -1.981, SD = 0.159$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05, $F = -2.578, p < .001$ ส่วนวิธีประยุกต์ของคัมภ์และวิธีประยุกต์ตัดตัวลง-เลือกตัวถูกแตกต่างกัน แต่ไม่มีนัยสำคัญทางสถิติที่ระดับ .05, $F = 0.154, p = 1.000$ ดังตาราง 5

ตาราง 5 ผลการทดสอบภายหลังของค่าความยากเฉลี่ยของข้อสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก

วิธีการให้คะแนน	ประยุกต์คุมภ์	ตัดตัวลง-เลือกตัวถูก	ประยุกต์ตัดตัวลง-เลือกตัวถูก
ประยุกต์คุมภ์	-	2.578 ($p < .001$)	0.154 ($p = 1.000$)
ตัดตัวลง-เลือกตัวถูก	-	-	-2.578 ($p < .001$)
ประยุกต์ตัดตัวลง-เลือกตัวถูก	-	-	-

การวิเคราะห์ความแตกต่างระหว่างค่าอำนาจจำแนกเฉลี่ยของข้อสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือกโดยใช้การวิเคราะห์ 2-way ANOVA พบว่าปฏิสัมพันธ์ระหว่างวิธีการให้คะแนนและจำนวนตัวเลือก ไม่มีนัยสำคัญทางสถิติที่ระดับ .05, $F(2, 114) = 1.231, p = .296$,

$\eta_p^2 = .021$ โดยเมื่อพิจารณาอิทธิพลหลัก พบว่า วิธีการให้คะแนนความรู้บางส่วนและจำนวนตัวเลือกที่ต่างกันส่งผลต่ออำนาจจำแนกของข้อสอบแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05, $F(2, 114) = 10.216, p < .001, \eta_p^2 = .152$, $F(1, 114) = 11.418, p = .001, \eta_p^2 = .091$ ดังตาราง 6

ตาราง 6 ผลการวิเคราะห์ความแปรปรวนแบบสองทางของอำนาจจำแนกของข้อสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก

แหล่งของความแปรปรวน	SS	df	MS	F	p	η_p^2
วิธีการให้คะแนน	.400	2	.200	10.216	< .001	.152
จำนวนตัวเลือก	.224	1	.224	11.418	.001	.091
ปฏิสัมพันธ์ระหว่างวิธีการให้คะแนนและจำนวนตัวเลือก	.048	2	.024	1.231	.296	.021
ความคลาดเคลื่อน	2.232	114	.020			
รวม	11.353	120				

Levene's test $F(5, 114) = 3.487, p = .006$

$R^2 = .231$, Adjusted $R^2 = .198$

จากการวิเคราะห์อิทธิพลหลัก พบว่า วิธีการให้คะแนนความรู้บางส่วนและจำนวนตัวเลือกที่ต่างกันส่งผลต่ออำนาจจำแนกของข้อสอบแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05, $F(2, 114) = 10.216, p < .001, \eta_p^2 = .152$, $F(1, 114) = 11.418, p = .001, \eta_p^2 = .091$ โดยความแปรปรวนของอำนาจจำแนกของข้อสอบที่ได้จากวิธีการให้คะแนนความรู้บางส่วนแต่ละวิธีแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05, $F(5, 114) = 3.487, p = .006$ จึงทำการทดสอบภายหลังโดยใช้สถิติ Bonferroni พบว่า วิธีประยุกต์

คุมภ์ ($M = 0.346, SD = 0.022$) มีค่าอำนาจจำแนกสูงกว่าวิธีตัดตัวลง-เลือกตัวถูก ($M = 0.236, SD = 0.022$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05, $F = 0.110, p = .002$ วิธีประยุกต์คุมภ์ ($M = 0.346, SD = 0.022$) มีค่าอำนาจจำแนกสูงกว่าวิธีประยุกต์ตัดตัวลง-เลือกตัวถูก ($M = 0.214, SD = 0.022$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05, $F = 0.132, p < .001$ ส่วนวิธีตัดตัวลง-เลือกตัวถูกและวิธีประยุกต์ตัดตัวลง-เลือกตัวถูกมีอำนาจจำแนกแตกต่างกัน แต่ไม่มีนัยสำคัญทางสถิติที่ระดับ .05, $F = 0.022, p = 1.000$ ดังตาราง 7

ตาราง 7 ผลการทดสอบภายหลังของอำนาจจำแนกของข้อสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก

วิธีการให้คะแนน	ประยุต์คัมภ์	ตัดตัวลง-เลือกตัวถูก	ประยุต์ตัดตัวลง-เลือกตัวถูก
ประยุต์คัมภ์	-	0.110 ($p = .002$)	0.132 ($p < .001$)
ตัดตัวลง-เลือกตัวถูก	-	-	0.022 ($p = 1.000$)
ประยุต์ตัดตัวลง-เลือกตัวถูก	-	-	-

ตอนที่ 2 ผลการวิเคราะห์ความตรงเชิงโครงสร้างของแบบสอบ

การวิเคราะห์ความตรงเชิงโครงสร้างของแบบสอบมีการวิเคราะห์ทั้งหมด 6 โมเดล ผลการตรวจสอบความสอดคล้องระหว่างโมเดลและข้อมูลทั้ง 6 โมเดล พบว่า โมเดลที่ได้จากวิธีประยุต์คัมภ์ เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก โมเดลที่ได้จากวิธีตัดตัวลง-เลือกตัวถูก เมื่อข้อสอบมี

5 ตัวเลือก และโมเดลที่ได้จากวิธีประยุต์ตัดตัวลง-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก มีความตรงเชิงโครงสร้าง ส่วนโมเดลที่ได้จากวิธีตัดตัวลง-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกไม่มีความตรงเชิงโครงสร้าง โดยโมเดลที่ดีที่สุด ได้แก่ วิธีประยุต์คัมภ์ เมื่อข้อสอบมี 4 ตัวเลือก เนื่องจากมีค่า AIC และ BIC ต่ำสุด รายละเอียดแสดงดังตาราง 8

ตาราง 8 ผลการวิเคราะห์ความสอดคล้องระหว่างโมเดลและข้อมูลที่ได้จากวิธีการให้คะแนนความรู้บางส่วน เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก

Model Tested	χ^2	p	df	χ^2/df	RMSEA	CFI	SRMR	AIC	BIC
ประยุต์คัมภ์									
4 ตัวเลือก	202.382	0.020	163	1.242	.024	.961	.038	27209.034	27480.524
5 ตัวเลือก	219.230	0.002	163	1.345	.028	.936	.039	31743.421	32014.911
ตัดตัวลง-เลือกตัวถูก									
4 ตัวเลือก	397.972	<.001	167	2.380	.058	.821	.064	35179.890	35433.520
5 ตัวเลือก	269.171	<.001	163	1.651	.040	.914	.043	38082.661	38352.394
ประยุต์ตัดตัวลง-เลือกตัวถูก									
4 ตัวเลือก	244.047	<.001	163	1.497	.035	.951	.039	38355.752	38625.161
5 ตัวเลือก	270.895	<.001	163	1.662	.040	.926	.043	43082.153	43351.562

ตอนที่ 3 ผลการวิเคราะห์ความเที่ยงของแบบสอบ

การวิเคราะห์ความเที่ยงของแบบสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก โดยใช้ความเที่ยงแบบความสอดคล้องภายใน วิธีสัมประสิทธิ์แอลฟาครอนบาค (Cronbach's Alpha) พบว่า วิธีประยุต์ตัดตัวลง-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกมีค่าความเที่ยงสูงสุด, $\alpha = .860$ รองลงมา คือ วิธีประยุต์ตัดตัวลง-เลือกตัวถูก เมื่อข้อสอบมี

5 ตัวเลือก, $\alpha = .843$ และวิธีตัดตัวลง-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือก, $\alpha = .812$ ตามลำดับ เมื่อวิเคราะห์ความแตกต่างของความเที่ยงของแบบสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก โดยใช้โปรแกรมสำเร็จรูป Feldt test พบว่า ความเที่ยงของทั้ง 3 วิธี เมื่อข้อสอบมีจำนวนตัวเลือกต่างกัน แตกต่างกัน แต่ไม่มีนัยสำคัญทางสถิติที่ระดับ .05 ดังตาราง 9

ตาราง 9 ผลการวิเคราะห์ความเที่ยงของแบบสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก

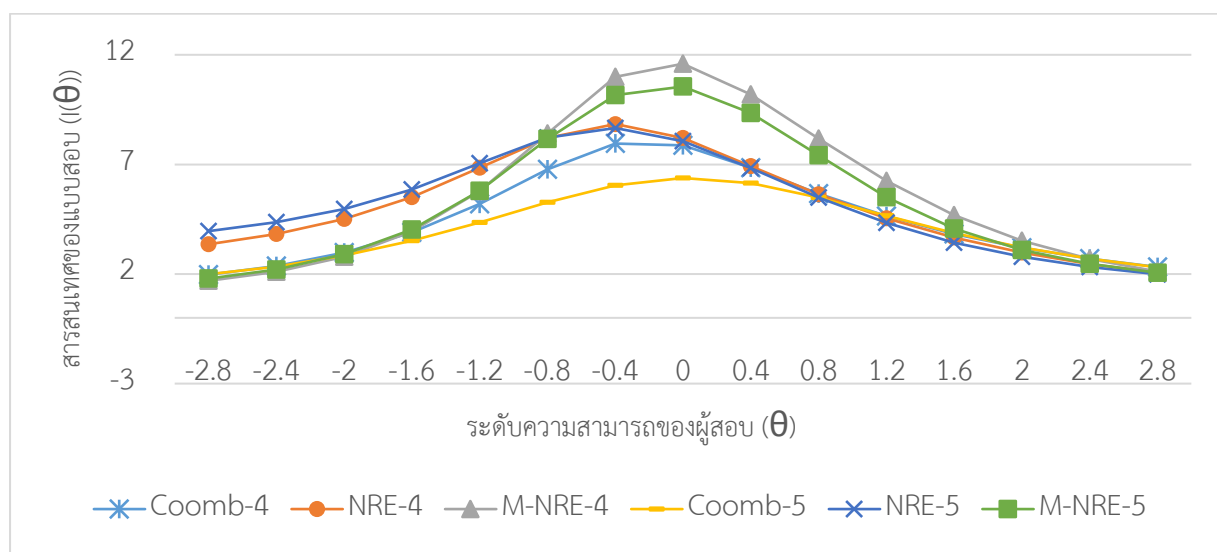
วิธีการให้คะแนน	จำนวนตัวเลือก	ประยุกต์คัมป์		ตัดตัวलग-เลือกตัวถูก		ประยุกต์ตัดตัวलग-เลือกตัวถูก	
		4 ตัวเลือก	5 ตัวเลือก	4 ตัวเลือก	5 ตัวเลือก	4 ตัวเลือก	5 ตัวเลือก
ประยุกต์คัมป์	4 ตัวเลือก	.787	1.000	1.133	1.121	1.521	1.357
			($p = .500$)	($p = .899$)	($p = .878$)	($p = 1.000$)	($p = .999$)
	5 ตัวเลือก		.787	1.293	1.279	1.736	1.548
				($p = .996$)	($p = .994$)	($p = 1.000$)	($p = 1.000$)
ตัดตัวलग-เลือกตัวถูก	4 ตัวเลือก			.812	0.990	1.343	1.198
					($p = .457$)	($p = .999$)	($p = .966$)
	5 ตัวเลือก				.810	1.357	1.210
						($p = .999$)	($p = .973$)
ประยุกต์ตัดตัวलग-เลือกตัวถูก	4 ตัวเลือก					.860	1.121
							($p = .877$)
	5 ตัวเลือก						.843

หมายเหตุ ค่าความเที่ยงของแต่ละวิธีแสดงตามแนวทแยง และผลวิเคราะห์ความแตกต่างของความเที่ยงของแบบสอบแสดงค่าการทดสอบซี (Z-Statistic) อยู่เหนือแนวทแยงและตัวเลขในวงเล็บแสดงค่า p

ตอนที่ 4 ผลการวิเคราะห์ฟังก์ชันสารสนเทศของแบบสอบ

การวิเคราะห์ฟังก์ชันสารสนเทศของแบบสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก พบว่าวิธีตัดตัวलग-เลือกตัวถูก เมื่อข้อสอบมี 5 ตัวเลือกให้สารสนเทศสูงสุดในช่วงระดับความสามารถต่ำ ($\theta = -2.8$ ถึง

-0.8) วิธีประยุกต์ตัดตัวलग-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกให้สารสนเทศสูงสุดในช่วงระดับความสามารถปานกลาง ($\theta = -0.8$ ถึง 0.8) และทั้ง 3 วิธีให้สารสนเทศใกล้เคียงกันในช่วงระดับความสามารถสูง ($\theta = 0.8$ ถึง 2.8) ดังภาพ 4



ภาพ 4 โค้งฟังก์ชันสารสนเทศของแบบสอบที่ได้จากวิธีประยุกต์ของคัมป์ วิธีตัดตัวलग-เลือกตัวถูก และวิธีประยุกต์ตัดตัวलग-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก

การวิเคราะห์ความแตกต่างระหว่างฟังก์ชันสารสนเทศของแบบสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือกโดยใช้การวิเคราะห์ 2-way ANOVA พบว่า ปฏิสัมพันธ์ระหว่างวิธีการให้คะแนนและจำนวนตัวเลือก ไม่มีนัยสำคัญทางสถิติที่ระดับ .05, $F(2, 84) = 1.788, p =$

.174, $\eta_p^2 = .002$ โดยเมื่อพิจารณาอิทธิพลหลัก พบว่า วิธีการให้คะแนนความรู้บางส่วนและจำนวนตัวเลือกส่งผลต่อฟังก์ชันสารสนเทศของแบบสอบ แต่ไม่มีนัยสำคัญทางสถิติที่ระดับ .05, $F(2, 84) = 1.788, p = .174, \eta_p^2 = .041, F(1, 84) = 0.242, p = .624, \eta_p^2 = .003$ ตามลำดับ ดังตาราง 10

ตาราง 10 ผลการวิเคราะห์ความแตกต่างระหว่างฟังก์ชันสารสนเทศของแบบสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก

แหล่งของความแปรปรวน	SS	df	MS	F	p	η_p^2
วิธีการให้คะแนน	22.394	2	11.197	1.788	.174	.041
จำนวนตัวเลือก	1.516	1	1.516	0.242	.624	.003
ปฏิสัมพันธ์ระหว่างวิธีการให้คะแนนและจำนวนตัวเลือก	1.138	2	0.569	0.091	.913	.002
ความคลาดเคลื่อน	526.014	84	6.262			
รวม	551.062	89				

Levene's test $F(5, 84) = 3.717, p = .004$

$R^2 = .045$, Adjusted $R^2 = .011$

ตอนที่ 5 ผลการวิเคราะห์ความแม่นยำในการประมาณค่าความสามารถของผู้สอบ

จากการวิเคราะห์ค่าสารสนเทศเฉลี่ยของแบบสอบที่ได้จากวิธีการให้คะแนนความรู้บางส่วน เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก พบว่า วิธีประยุกต์ตัดตัวลง-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือก มีค่าสารสนเทศเฉลี่ยของแบบ

สอบสูงสุด รองลงมา คือ วิธีประยุกต์ตัดตัวลง-เลือกตัวถูกเมื่อข้อสอบมี 5 ตัวเลือก และวิธีตัดตัวลง-เลือกตัวถูกเมื่อข้อสอบมี 5 ตัวเลือก ตามลำดับ และวิธีที่มีค่าสารสนเทศเฉลี่ยของแบบสอบต่ำสุด คือ วิธีประยุกต์คุมบ์ เมื่อข้อสอบมี 5 ตัวเลือก ดังตาราง 11

ตาราง 11 ผลการวิเคราะห์ค่าสารสนเทศเฉลี่ยของแบบสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก

วิธีการให้คะแนน	จำนวนตัวเลือก	ค่าสารสนเทศเฉลี่ยของแบบสอบ
ประยุกต์คุมบ์	4 ตัวเลือก	4.547
	5 ตัวเลือก	4.074
ตัดตัวลง-เลือกตัวถูก	4 ตัวเลือก	5.169
	5 ตัวเลือก	5.221
ประยุกต์ตัดตัวลง-เลือกตัวถูก	4 ตัวเลือก	5.661
	5 ตัวเลือก	5.305

อภิปรายผล

1. การวิเคราะห์ความยากและอำนาจจำแนกของข้อสอบ

การวิเคราะห์ความแตกต่างระหว่างค่าความยากของข้อสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก พบว่า วิธีการให้คะแนนความรู้บางส่วนที่ต่างกันส่งผลต่อความยากของข้อสอบแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งสอดคล้องกับงานวิจัยของ รัตนา ไชยตรี (2546); เอมอร์ จังศิริพรภรณ์ (2545) เนื่องจากแต่ละวิธีมีวิธีการตอบและการให้คะแนนที่แตกต่างกัน จึงส่งผลต่อการพิจารณาเลือกคำตอบ เช่น วิธีประยุกต์คัมภ์และวิธีประยุกต์ตัดตัวลง-เลือกตัวถูก เป็นวิธีที่มีการให้คะแนนในช่วงที่เป็นบวกทำให้ผู้สอบกล้าตอบข้อสอบมากกว่าการทำข้อสอบโดยวิธีตัดตัวลง-เลือกตัวถูกซึ่งมีการให้คะแนนในช่วงที่ติดลบ นอกจากนี้ พบว่าข้อสอบที่ได้จากวิธีประยุกต์คัมภ์ เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือกเป็นข้อสอบที่มีค่าความยากสูงสุด เมื่อเปรียบเทียบกับวิธีอื่น ๆ ซึ่งสอดคล้องกับงานวิจัยของรัตนา ไชยตรี (2546) ทั้งนี้เนื่องจากวิธีประยุกต์ของคัมภ์เป็นวิธีที่ให้ผู้สอบเลือกตัวเลือกที่ผิดซึ่งตรงข้ามกับวิธีการทำข้อสอบที่นักเรียนคุ้นเคยในปัจจุบัน นั่นคือ เลือกตัวเลือกที่ถูก ทำให้ผู้สอบเกิดความสับสนในการทำข้อสอบ นอกจากนี้ วิธีประยุกต์ของคัมภ์ให้ผู้สอบเลือกตัวเลือกที่มั่นใจว่าผิดก็ตัวเลือกก็ได้ นั่นคือ ผู้สอบอาจจะตอบ 1, 2 หรือ 3 ตัวเลือกก็ได้ตามความมั่นใจของผู้สอบ แต่ผู้สอบส่วนใหญ่มักจะตอบทั้ง 3 ตัวเลือก ในกรณีที่มี 4 ตัวเลือก และมักจะตอบทั้ง 4 ตัวเลือก ในกรณีที่มี 5 ตัวเลือก ซึ่งทำให้โอกาสได้คะแนนแต่ละข้อเป็น 0 คะแนนมากยิ่งขึ้นและส่งผลให้วิธีประยุกต์ของคัมภ์มีค่าความยากสูงสุด ส่วนวิธีตัดตัวลง-เลือกตัวถูกส่วนใหญ่มีค่าความยากต่ำสุดเนื่องจากวิธีนี้เป็นวิธีที่มีโอกาสในการเดาน้อย ค่าความคลาดเคลื่อนมาตรฐานของการวัดต่ำกว่าวิธีอื่น ๆ และสามารถจำแนกผู้เรียนที่มีความรู้สมบูรณ์ มีความรู้บางส่วน ไม่มีความรู้ มีความรู้คลาดเคลื่อนบางส่วนและมีความรู้คลาดเคลื่อนทั้งหมดได้ (Lau, Lau, Hong, & Usop, 2011; 2014) อีกทั้งมีการให้คะแนนแบบติดลบ ทำให้ผู้สอบมีการคิดพิจารณาคำตอบมากกว่าเดิม และไม่กล้าเสี่ยงที่จะเดาคำตอบจึงทำให้โอกาสในการได้คะแนนในแต่ละข้อมากขึ้น และวิธีประยุกต์ตัดตัวลง-เลือกตัวถูกส่วนใหญ่มีค่าความยากปานกลาง เนื่องจากถ้าพิจารณาจากสารสนเทศของแบบสอบที่ได้จากวิธีประยุกต์ตัดตัวลง-

เลือกตัวถูกจะเห็นว่าวิธีนี้จะให้สารสนเทศมากสำหรับผู้สอบที่มีความสามารถระดับปานกลางซึ่งค่าความยากและความสามารถของผู้สอบแสดงบนสเกลเดียวกันทำให้วิธีนี้มีค่าความยากเข้าใกล้ 0 มากกว่าวิธีอื่น ๆ

ส่วนจำนวนตัวเลือกที่ต่างกันไม่ส่งผลต่อความยากของข้อสอบ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งไม่สอดคล้องกับงานวิจัยของ Bahrami (2011); Rodriguez (2005) เนื่องจากในงานวิจัยของ Bahrami (2011) เป็นงานวิจัยที่ศึกษาข้อสอบที่มี 4 ตัวเลือกและ 6 ตัวเลือก ซึ่งมีช่วงห่างค่อนข้างมากอาจจะทำให้จำนวนตัวเลือกที่ต่างกันส่งผลต่อความยากของข้อสอบอย่างชัดเจน และจากงานวิจัยของ Rodriguez (2005) พบว่า ข้อสอบที่มีจำนวนตัวเลือก 4 ตัวเลือกและ 5 ตัวเลือกมีค่าความยากแตกต่างกันเพียงเล็กน้อยอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 แต่ถ้าเป็นข้อสอบที่มีจำนวนตัวเลือก 3 ตัวเลือกและ 4 ตัวเลือกจะมีค่าความยากแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

การวิเคราะห์ความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก พบว่า วิธีการให้คะแนนความรู้บางส่วนและจำนวนตัวเลือกที่ต่างกันส่งผลต่อค่าอำนาจจำแนกของข้อสอบแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งสอดคล้องกับงานวิจัยของ Bahrami (2011); Rodriguez (2005); รัตนา ไชยตรี (2546); เอมอร์ จังศิริพรภรณ์ (2545a) โดยข้อสอบที่ได้จากวิธีประยุกต์คัมภ์ เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก มีค่าอำนาจจำแนกสูงกว่าวิธีอื่น ๆ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เนื่องจากวิธีประยุกต์ของคัมภ์เป็นวิธีที่มีความยากที่สุด ดังนั้น ผู้สอบที่เก่งก็จะสามารถทำข้อสอบได้ดี ส่วนผู้สอบที่อ่อนก็จะไม่สามารถทำข้อสอบได้หรืออาจเกิดการเดาข้อสอบและทำได้ 0 คะแนน ส่วนการเปรียบเทียบวิธีตัดตัวลง-เลือกตัวถูกและวิธีประยุกต์ตัดตัวลง-เลือกตัวถูกมีค่าอำนาจจำแนกใกล้เคียงกัน เนื่องจากมีช่วงการให้คะแนนใกล้เคียงกันทำให้ทั้ง 2 วิธีนี้จำแนกผู้สอบที่มีความสามารถต่างกันในทุกระดับความสามารถได้ใกล้เคียงกัน

นอกจากนี้ เมื่อวิเคราะห์ความสัมพันธ์ระหว่างค่าความยากและค่าอำนาจจำแนกเฉลี่ย พบว่า วิธีประยุกต์คัมภ์และวิธีประยุกต์ตัดตัวลง-เลือกตัวถูกมีความความสัมพันธ์ระหว่างค่าความยากและค่าอำนาจจำแนกติดลบ เนื่องจากวิธี

ประยุกต์คุณและวิธีประยุกต์ตัดตัวลง-เลือกตัวถูกเป็นข้อสอบก่อนข้างยาก จึงทำให้ค่าอำนาจจำแนกต่ำ ดังนั้น ความสัมพันธ์จึงติดลบ (Pande et al., 2013; Sim & Rasiah, 2006; ศิริชัย กาญจนาวาสี, 2556)

2. การวิเคราะห์ความตรงเชิงโครงสร้างของแบบสอบ

การตรวจสอบความสอดคล้องระหว่างโมเดลและข้อมูล พบว่า โมเดลทั้ง 5 โมเดลมีความตรงเชิงโครงสร้าง ซึ่งแต่ละวิธีมีค่าความตรงเชิงโครงสร้างต่างกันซึ่งสอดคล้องกับสุพจน์ เกิดสุวรรณ (2545) ประกอบกับเครื่องมือวิจัยมีการตรวจสอบความตรงเชิงเนื้อหา ดังนั้น จึงทำให้โมเดลที่วิเคราะห์มีความตรงเชิงโครงสร้างด้วยเช่นกัน ส่วนโมเดลที่ได้จากวิธีตัดตัวลง-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกไม่มีความตรงเชิงโครงสร้าง เนื่องจากวิธีตัดตัวลง-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกเป็นวิธีที่มีการให้คะแนนแบบติดลบจึงอาจทำให้กระทบต่อคะแนนรวมและผลการวิเคราะห์ได้

3. การวิเคราะห์ความเที่ยงของแบบสอบ

การวิเคราะห์ความเที่ยงของแบบสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก พบว่า ค่าความเที่ยงของทุกวิธีมีค่าสูงกว่า .70 ซึ่งอยู่เกณฑ์ที่ยอมรับได้ (ศิริชัย กาญจนาวาสี, 2556) โดยวิธีประยุกต์ตัดตัวลง-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือกมีความเที่ยงสูงกว่าวิธีอื่นๆ ซึ่งสอดคล้องกับสมมติฐาน เนื่องจาก วิธีนี้มีช่วงการให้คะแนนอยู่ในช่วงบวกจึงมีความเที่ยงสูงกว่าวิธีที่มีการให้คะแนนอยู่ในช่วงติดลบ (สุพจน์ เกิดสุวรรณ, 2545) และเนื่องจากวิธีตัดตัวลง-เลือกตัวถูกเป็นวิธีที่มีคะแนนอยู่ในช่วงติดลบซึ่งเป็นการให้คะแนนแบบการลงโทษจะทำให้พฤติกรรมในการทำข้อสอบลดลง (สุรางค์ ไคว์ตระกูล, 2548) นั่นคือผู้สอบอาจจะคิดพิจารณาคำตอบมากกว่าเดิม มีความกังวลในการทำข้อสอบมากกว่าเดิม และไม่กล้าเสี่ยงในการตอบข้อสอบและส่งผลให้ความแปรปรวนของคะแนนสอบต่ำ นอกจากนี้ วิธีประยุกต์ตัดตัวลง-เลือกตัวถูกมีช่วงการให้คะแนนกว้างกว่าวิธีประยุกต์คุมบ์ ทำให้ความแปรปรวนของคะแนนที่ได้จากวิธีประยุกต์ตัดตัวลง-เลือกตัวถูกเพิ่มขึ้นและทำให้ความเที่ยงก็เพิ่มขึ้นด้วยเช่นกัน (Frey, 1989) ซึ่งถ้าช่วงการให้คะแนนน้อยจะทำให้ข้อมูลที่เป็นประโยชน์ขาดหายไปและความแม่นยำในการวัดต่ำลง (Ma, 2004) ดังนั้น วิธีประยุกต์ตัดตัวลง-เลือกตัวถูกจึงมีความเที่ยงสูงกว่าวิธีประยุกต์คุมบ์ แต่เมื่อวิเคราะห์

ความแตกต่างของความเที่ยงของแบบสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก พบว่า ความเที่ยงของทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก ไม่แตกต่างกัน เนื่องจากแต่ละวิธีให้ผู้สอบได้แสดงสารสนเทศในการตอบข้อสอบอย่างละเอียดและเป็นวิธีที่ป้องกันการเดาคำตอบ จึงทำให้ความเที่ยงของแต่ละวิธีไม่แตกต่างกัน

4. การวิเคราะห์ฟังก์ชันสารสนเทศของแบบสอบ

การวิเคราะห์ฟังก์ชันสารสนเทศของแบบสอบที่ได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก พบว่า วิธีตัดตัวลง-เลือกตัวถูก เมื่อข้อสอบมี 5 ตัวเลือกให้สารสนเทศสูงที่สุดในช่วงระดับความสามารถต่ำ ($\theta = -2.8$ ถึง -0.8) เนื่องจากวิธีตัดตัวลง-เลือกตัวถูกเป็นวิธีที่มีค่าความยากต่ำที่สุดทำให้ผู้สอบที่มีระดับความสามารถต่ำสามารถทำข้อสอบได้และให้สารสนเทศสูงในช่วงนี้ ส่วนวิธีประยุกต์ตัดตัวลง-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกให้สารสนเทศสูงที่สุดในช่วงระดับความสามารถปานกลาง ($\theta = -0.8$ ถึง 0.8) เนื่องจากเป็นวิธีที่ผสมผสานระหว่างวิธีการตัดตัวลงของคุมบ์และวิธีประเพณีนิยมซึ่งเป็นวิธีที่ลดโอกาสในการเดาและมีการปรับให้ค่าคะแนนไม่ติดลบ และมีการเสริมแรงทางบวกในกรณี que ผู้สอบเลือกตัวถูกได้ถูกต้องซึ่งจะทำให้ได้สารสนเทศที่สะท้อนความสามารถที่แท้จริงของผู้เรียนและสามารถจำแนกผู้เรียนได้ (สุพจน์ เกิดสุวรรณ, 2545) วิธีประยุกต์การให้คะแนนแบบตัดตัวลง-เลือกตัวถูกทำให้ผู้สอบมีการคิดพิจารณาก่อนตอบข้อสอบมากกว่าเดิม รองลงมา คือ วิธีตัดตัวลง-เลือกตัวถูกซึ่งจะมีฟังก์ชันสารสนเทศของแบบสอบสูงกว่าวิธีประยุกต์ของคุมบ์เล็กน้อย เนื่องจากวิธีตัดตัวลง-เลือกตัวถูกมีช่วงการให้คะแนนกว้างกว่าทำให้มีความแปรปรวนของคะแนนสอบมากกว่า มีความคลาดเคลื่อนมาตรฐานในการวัดต่ำ สามารถตรวจสอบความเข้าใจที่คลาดเคลื่อนได้และลดการเดาได้ (Lau, Lau, Hong, & Usop, 2014) จึงให้สารสนเทศของแบบสอบสูงกว่าวิธีประยุกต์ของคุมบ์เล็กน้อย ส่วนวิธีประยุกต์ของคุมบ์มีช่วงการให้คะแนนน้อยกว่าวิธีตัดตัวลง-เลือกตัวถูกและวิธีประยุกต์แบบตัดตัวลง-เลือกตัวถูกจะทำให้ข้อมูลที่เป็นประโยชน์ขาดหายไปและความแม่นยำในการวัดต่ำลง (Ma, 2004) แต่ในทางตรงข้าม วิธีตัดตัวลง-เลือกตัวถูกมีการให้คะแนนแบบติดลบซึ่งเป็นการให้คะแนนแบบการลงโทษจะทำให้พฤติกรรมในการทำข้อสอบลดลง (สุรางค์ ไคว์ตระกูล,

2548) นั่นคือ ผู้สอบอาจจะคิดพิจารณาคำตอบมากกว่าเดิม มีความกังวลในการทำข้อสอบมากกว่าเดิม และไม่กล้าเสี่ยงในการตอบข้อสอบ และทั้ง 3 วิธีให้สารสนเทศใกล้เคียงกันในช่วงระดับความสามารถสูง ($\theta = 0.8$ ถึง 2.8) เนื่องจากผู้สอบที่มีความสามารถระดับสูงมีแนวโน้มที่จะทำข้อสอบได้ถูกต้องอย่างสมบูรณ์โดยไม่ขึ้นอยู่กับวิธีการตอบและการให้คะแนนความรู้บางส่วน

นอกจากนี้ เมื่อพิจารณาทั้ง 3 วิธี ในทุกช่วงระดับความสามารถ พบว่า สารสนเทศของแบบสอบแตกต่างกันมาก เมื่อผู้สอบมีความสามารถระดับปานกลาง และสารสนเทศของแบบสอบไม่แตกต่างกันเมื่อผู้สอบมีความสามารถระดับต่ำและสูง เนื่องจากผู้สอบที่มีความสามารถระดับต่ำส่วนใหญ่จะเดา ซึ่งทั้ง 3 วิธีเป็นวิธีที่ลดโอกาสในการเดาทำให้ผู้สอบมีโอกาสได้ 0 คะแนนหรือคะแนนติดลบมากขึ้น เช่น ทั้ง 3 วิธีให้ผู้สอบตัดตัวลวงออกก็ได้แต่ผู้สอบที่มีความสามารถระดับต่ำมักจะตัดตัวลวงออก 3 ตัวทำให้โอกาสที่จะได้ 0 คะแนนมากขึ้นจึงทำให้สารสนเทศที่ได้ไม่แตกต่างกัน ส่วนผู้สอบที่มีความสามารถระดับสูงส่วนใหญ่จะทำข้อสอบได้ถูกต้องแบบสมบูรณ์โดยไม่ขึ้นอยู่กับวิธีการตอบและการให้คะแนน เนื่องจากผู้ที่มีความสามารถระดับสูงมีความรู้มาก ไม่ว่าจะให้ตอบวิธีใดก็จะสามารถตอบได้อย่างมั่นใจ จึงทำให้สารสนเทศที่ได้ไม่แตกต่างกัน

5. การวิเคราะห์ความแม่นยำในการประมาณค่าความสามารถของผู้สอบ

จากการวิเคราะห์ค่าฟังก์ชันสารสนเทศเฉลี่ยของแบบสอบได้จากทั้ง 3 วิธี เมื่อข้อสอบมี 4 ตัวเลือกและ 5

ตัวเลือก พบว่า วิธีประยุกต์ตัดตัวลวง-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือก มีความแม่นยำในการประมาณค่าความสามารถของผู้สอบสูงสุด รองลงมา คือ วิธีประยุกต์ตัดตัวลวง-เลือกตัวถูก เมื่อข้อสอบมี 5 ตัวเลือก และวิธีตัดตัวลวง-เลือกตัวถูกเมื่อข้อสอบมี 5 ตัวเลือก ตามลำดับ เนื่องจากวิธีประยุกต์ตัดตัวลวง-เลือกตัวถูกเป็นวิธีที่พัฒนามาจากวิธีตัดตัวลวง-เลือกตัวถูกแต่มีการปรับช่วงการให้คะแนนไม่ให้ติดลบ ดังนั้น วิธีตัดตัวลวง-เลือกตัวถูกและวิธีประยุกต์ตัดตัวลวง-เลือกตัวถูกจึงมีช่วงในการให้คะแนนใกล้เคียงกันทำให้สามารถจำแนกความสามารถที่แท้จริงของผู้เรียนได้อย่างแม่นยำกว่าวิธีประยุกต์คุมภ์ ซึ่งสอดคล้องกับงานวิจัยของ Lau, Lau, Hong, & Usop (2014) พบว่า วิธีตัดตัวลวง-เลือกตัวถูกมีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบและให้ข้อมูลที่เป็นประโยชน์ในการจำแนกผู้สอบตามระดับความสามารถของแต่ละบุคคลสูงกว่าวิธีประยุกต์คุมภ์

บทสรุป

การวิเคราะห์คุณภาพของวิธีการตรวจให้คะแนนความรู้บางส่วน เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก พิจารณาจากความตรงเชิงโครงสร้าง ความเที่ยงและฟังก์ชันสารสนเทศของแบบสอบเฉลี่ยและนำผลการวิเคราะห์มาจัดอันดับคุณภาพ พบว่า วิธีประยุกต์ตัดตัวลวง-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกมีคุณภาพสูงที่สุด รองลงมา คือ วิธีประยุกต์ตัดตัวลวง-เลือกตัวถูก เมื่อข้อสอบมี 5 ตัวเลือก ส่วนวิธีประยุกต์คุมภ์ เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือกและวิธีตัดตัวลวง-เลือกตัวถูก เมื่อข้อสอบมี 4 ตัวเลือกเป็นวิธีที่มีคุณภาพต่ำ รายละเอียดดังตาราง 12

ตาราง 12 ผลการเปรียบเทียบอันดับคุณภาพของวิธีการตรวจให้คะแนนความรู้บางส่วน เมื่อข้อสอบมี 4 ตัวเลือกและ 5 ตัวเลือก

วิธีการให้คะแนน	จำนวนตัวเลือก	ความตรงเชิงโครงสร้าง	ความเที่ยง	ฟังก์ชันสารสนเทศของแบบสอบเฉลี่ย
ประยุกต์คุมภ์	4 ตัวเลือก	ตรง	5	5
	5 ตัวเลือก	ตรง	5	6
ตัดตัวลวง-เลือกตัวถูก	4 ตัวเลือก	ไม่ตรง	3	4
	5 ตัวเลือก	ตรง	4	3
ประยุกต์ตัดตัวลวง-เลือกตัวถูก	4 ตัวเลือก	ตรง	1	1
	5 ตัวเลือก	ตรง	2	2

ข้อเสนอแนะ**ข้อเสนอแนะในการนำผลวิจัยไปใช้**

1. เมื่อเปรียบเทียบคุณภาพของแต่ละวิธี ซึ่งให้สารสนเทศในการจำแนกผู้เรียนแตกต่างกันตามจำนวนตัวเลือกที่แตกต่างกัน ดังนั้น วิธีประยุกต์คุมภ์และวิธีประยุกต์ตัดตัวลง-เลือกเหมาะกับข้อสอบที่มี 4 ตัวเลือก และวิธีตัดตัวลง-เลือกตัวถูกเหมาะกับข้อสอบที่มี 5 ตัวเลือก ดังนั้น ผู้สอนควรพิจารณาจำนวนตัวเลือกที่เหมาะสมกับแต่ละวิธีการให้คะแนนความรู้บางส่วนด้วย

2. วิธีตัดตัวลง-เลือกตัวถูกและวิธีประยุกต์ตัดตัวลง-เลือกตัวถูกมีคุณภาพใกล้เคียงกันหลายด้านแต่ในทางปฏิบัติ การนำวิธีตัดตัวลง-เลือกตัวถูกอาจจะไม่เหมาะสมสำหรับการนำไปใช้ในการวัดและประเมินผลการเรียนรู้ของผู้เรียนเพื่อตัดสินผล เนื่องจากเป็นวิธีที่ให้คะแนนแบบติดลบซึ่งจะส่งผลกระทบต่อคะแนนรวม อีกทั้งยังมีความซับซ้อนในการตรวจให้คะแนนแต่วิธีตัดตัวลง-เลือกตัวถูกอาจเหมาะสำหรับการวินิจฉัยปัญหาการเรียนรู้ของผู้เรียน

3. จากข้อค้นพบในงานวิจัยครั้งนี้ พบว่า วิธีประยุกต์ตัดตัวลง-เลือกตัวถูกเป็นวิธีที่มีคุณภาพดีกว่าวิธีอื่น ๆ โดยจะให้สารสนเทศสูงสำหรับผู้สอบที่มีความสามารถระดับปานกลาง ดังนั้น วิธีประยุกต์ตัดตัวลง-เลือกตัวถูกจึงเหมาะสำหรับใช้กับผู้สอบที่มีระดับความสามารถปานกลาง

4. วิธีประยุกต์ของคุมภ์ วิธีตัดตัวลง-เลือกตัวถูกและวิธีประยุกต์ตัดตัวลง-เลือกตัวถูกเป็นวิธีที่มีความซับซ้อนในการตอบข้อสอบและผู้สอบไม่คุ้นเคย ดังนั้น การนำวิธีทั้ง 3

วิธีไปใช้ควรระมัดระวังเรื่องการตอบ โดยครูผู้สอนควรอธิบายคำชี้แจงให้ผู้สอบเข้าใจก่อนลงมือทำข้อสอบอย่างละเอียด

5. วิธีตัดตัวลง-เลือกตัวถูกและวิธีประยุกต์ตัดตัวลง-เลือกตัวถูกเป็นวิธีที่ให้ผู้สอบพิจารณาทุกตัวเลือกว่าถูกผิด หรือไม่แน่ใจ ดังนั้น ผู้วิจัยจึงควรพิจารณาความเหมาะสมของเวลาในการทำแบบสอบให้ดี

6. วิธีประยุกต์ของคุมภ์ วิธีตัดตัวลง-เลือกตัวถูกและวิธีประยุกต์ตัดตัวลง-เลือกตัวถูกเป็นวิธีที่มีความซับซ้อนในการตอบข้อสอบ ดังนั้น จึงควรให้ผู้สอบได้มีโอกาสทำความเข้าใจกับความคุ้นเคยกับวิธีการทำข้อสอบ ก่อนทำแบบสอบจริง เพื่อสร้างความคุ้นเคยในการทำข้อสอบและประหยัดเวลาในการอธิบายคำชี้แจงในการทำข้อสอบ

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. ข้อสอบที่ใช้ในงานวิจัยนี้เป็นข้อสอบที่มี 4 ตัวเลือก โดยมีคำตอบถูกเพียง 1 ตัวเลือก ดังนั้น จึงควรมีการศึกษาต่อในกรณีที่ข้อสอบมีคำตอบถูกมากกว่า 1 ตัวเลือก เพื่อให้ได้วิธีที่ให้สารสนเทศจากผลการตอบข้อสอบมากที่สุด และเหมาะสมกับการวัดและประเมินผลการเรียนรู้ของผู้เรียนอย่างแม่นยำ

2. ควรทำการศึกษาเพื่อตรวจสอบคุณภาพของวิธีการให้คะแนนความรู้บางส่วนในวิชาอื่น ๆ เช่น วิชาภาษาไทย วิชาภาษาอังกฤษ วิชาสังคมศึกษา เป็นต้น เพราะบริบทของแต่ละวิชาไม่เหมือนกัน ซึ่งวิธีการให้คะแนนความรู้บางส่วนแต่ละวิธีอาจจะเหมาะสมกับแต่ละวิชาแตกต่างกัน

บรรณานุกรม

- กระทรวงศึกษาธิการ. (2551). แนวปฏิบัติการวัดและประเมินผลการเรียนรู้ ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551. กรุงเทพฯ: สำนักวิชาการและมาตรฐานการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน.
- กระทรวงศึกษาธิการ. (2551). หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551. กรุงเทพฯ: สำนักวิชาการและมาตรฐานการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน.
- โชติกา ภาชีผล. (2554). การสร้างและพัฒนาเครื่องมือในการวัดและประเมินผลการศึกษา. กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย.
- โชติกา ภาชีผล ญัฐภรณ์ หลาวทอง และกมลวรรณ ตั้งชนกานนท์. (2558). การวัดและประเมินผลการเรียนรู้. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ธนิดา เขยขุ่ม. (2551). การเปรียบเทียบคุณภาพของข้อสอบและแบบสอบหลายตัวเลือก ที่มีรูปแบบตัวเลือกแตกต่างกัน. กรุงเทพฯ: คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

- รัตน ไชยตรี. (2546). การเปรียบเทียบคุณภาพและอัตราความคลาดเคลื่อนของการกำหนดเกรดแบบอิงกลุ่ม และอิงเกณฑ์ เมื่อใช้แบบสอบเลือกตอบที่มีการตอบ และการตรวจให้คะแนนความรู้บางส่วนด้วยวิธีแตกต่างกัน. (วิทยานิพนธ์ปริญญาคุชฎบัณฑิต). บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. กรุงเทพฯ
- ศิริชัย กาญจนวาสี. (2556). *ทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory)*. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สุพจน์ เกิดสุวรรณ. (2545). การพัฒนาการวัดความรู้บางส่วนของผู้ตอบแบบสอบเลือกตอบ. (วิทยานิพนธ์ปริญญาคุชฎบัณฑิต). บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. กรุงเทพฯ.
- สุรงค์ ไคว้ตระกูล. (2548). *จิตวิทยาการศึกษา*. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- อานภาพ เลขะกุล. (2559). *การสร้างข้อสอบแบบเลือกตอบ (Multiple choice question)*.
- เอมอร์ จังศิริพรกรณ์. (2545a). การเปรียบเทียบคุณภาพของแบบสอบเลือกตอบเมื่อตรวจด้วยวิธีการให้คะแนนความรู้บางส่วนกับวิธีประเพณีนิยม. กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย.
- เอมอร์ จังศิริพรกรณ์. (2545b). การให้คะแนนความรู้บางส่วนทำได้อย่างไร. *วารสารครุศาสตร์*, 23, 21-32.
- เอมอร์ จังศิริพรกรณ์. (2546). การเปรียบเทียบคุณภาพของแบบสอบเลือกตอบเมื่อตรวจด้วยวิธีการให้คะแนนความรู้บางส่วนกับวิธีประเพณีนิยม. *วารสารครุศาสตร์*, 31, 24-40.
- Bahrami, G. G. (2011). The Role of Number of choices on the Parameters, Item's Goodness of Fit and Information Function of the Items in Different Levels of Ability in Multiple Choice Tests. *Elsevier Ltd.*, 1305-1311.
- Chang, S.-H., Lin, P.-C., & Lin, Z.-C. (2007). Measures of Partial Knowledge and Unexpected Responses in Multiple-Choice Tests. *Educational Technology & Society*, 10(4), 95-109.
- Frey, A. S. (1989). *A test validation study for capturing partial information in multiple-choice questions using polychotomous scoring*. (PH.D Dissertation). University of San Francisco.
- Lau, P. N. K., Lau, S. H., Hong, K. S., & Usop, H. (2011). Guessing, Partial Knowledge, and Misconceptions in Multiple-Choice Tests. *Educational Technology & Society*, 14(4), 99-110.
- Lau, S. H., Hong, K. S., Lau, P. N. K., & Usop, H. (2014). Robustness of number right elimination testing (NRET) scoring method for multiple-choice items in computer adaptive assessment system (CAAS). *Research and Practice in Technology Enhanced Learning*, 9(2), 283-300.
- Ma, X. (2004). *An investigation of alternative approaches to scoring multiple response items on a certification examination*. (PH.D Dissertation). University of Massachusetts.
- MacCann, R. G. (2004). RELIABILITY AS A FUNCTION OF THE NUMBER OF ITEM OPTIONS DERIVED FROM THE "KNOWLEDGE OR RANDOM GUESSING" MODEL. *The Psychometric Society*, 69, 147-157.
- McMillan, J. H. (2014). *Classroom Assessment*. Pearson Education, Inc.
- Muraki, E. (1992). A Generalized Partial Credit Model : Application of an EM Algorithm. *APPLIED PSYCHOLOGICAL MEASUREMENT*, 16, 159-176.
- Nitko, A. J. (2004). *Educational Assessment Of Students*. Pearson Education, Inc.: Kevin M. Davis.
- Pande, S. S., Santosh, P. R., Vrushali, P. R., Archana, N. P., & Sushil, A. H. (2013). Correlation between difficulty & discrimination indices of MCQs in formative exam in Physiology. *Original Research Paper*, 7.
- Rodriguez, M. C. (2005). Three Options Are Optimal for Multiple-Choice Items: A Meta-Analysis of 80 Years of Research. *Educational Measurement: Issues and Practice*.

Sim, S. M., & Rasiah, R. I. (2006). Relationship between item difficulty and discrimination indices in true/false-type multiple choice questions of a para-clinical multidisciplinary paper. *ResearchGate*, 35, 67-71.