การเปรียบเทียบสถิติบูทสแตรปของการออกแบบการปรับเทียบสำหรับแบบสอบรูปแบบผสม
ระหว่างวิธีการที่ต่างกัน

นางสาวพนิดา พานิชวัฒนะ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรมหาบัณฑิต
สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2556
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย
The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the University Graduate School.

A COMPARISION OF BOOTSTRAPPED STATISTICS IN EQUATING DESIGNS FOR MIXED FORMAT TEST AMONG DIFFERENT METHODS

Miss Panida Panidvadtana

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Education Program in Educational Measurement and Evaluation
Department of Educational Research and Psychology
Faculty of Education
Chulalongkorn University
Academic Year 2013

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การเปรียบเทียบสถิติบูทสแตรปของการออกแบบการปรับเทียบสำหรับแบบสอบรูปแบบผสมระหว่างวิธีการที่ต่างกัน
โดย	นางสาวพนิดา พานิชวัฒนะ
สาขาวิชา	การวัดและประเมินผลการศึกษา
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร. ณัฏฐภรณ์ หลาวทอง

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....คณบดีคณะครุศาสตร์

(รองศาสตราจารย์ ดร. ชนิตา รักษ์พลเมือง)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ

(ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(ผู้ช่วยศาสตราจารย์ ดร. ณัฏฐภรณ์ หลาวทอง)

.....กรรมการภายนอกมหาวิทยาลัย

(ดร. ชนะศึก นิชานนท์)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

พนิดา พานิชวัฒนะ : การเปรียบเทียบสถิติบูทสแตรปของการออกแบบการปรับเทียบสำหรับแบบสอบรูปแบบผสมระหว่างวิธีการที่ต่างกัน. (A COMPARISION OF BOOTSTRAPPED STATISTICS IN EQUATING DESIGNS FOR MIXED FORMAT TEST AMONG DIFFERENT METHODS) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ. ดร. ณัฏฐภรณ์ หลาวทอง, 326 หน้า.

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) เพื่อเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจที่แตกต่างกัน 2) เพื่อเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีของเลอวิน (Levine method) และวิธี CLE (Chained Linear Equating) กลุ่มตัวอย่างเป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 สำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต 1 กรุงเทพมหานคร จำนวน 1,609 คน โดยสุ่มกลุ่มตัวอย่างด้วยวิธีการสุ่มแบบหลายขั้นตอน เครื่องมือที่ใช้ในการวิจัยครั้งนี้ ได้แก่ 1) แบบสอบวิชาฟิสิกส์รูปแบบผสม เรื่องงานและพลังงาน ระดับชั้นมัธยมศึกษาปีที่ 4 จำนวน 2 ฉบับ ซึ่งเป็นแบบสอบคู่ขนาน ได้แก่ แบบสอบฉบับใหม่ ประกอบด้วยข้อสอบแบบหลายตัวเลือกจำนวน 16 ข้อ และข้อสอบแบบสร้างคำตอบจำนวน 8 ข้อ ที่มีค่าความเที่ยงเป็น .634 และ .824 ตามลำดับ และแบบสอบฉบับอ้างอิง ประกอบด้วยข้อสอบแบบหลายตัวเลือกจำนวน 16 ข้อ และข้อสอบแบบสร้างคำตอบจำนวน 8 ข้อ ที่มีค่าความเที่ยงเป็น .602 และ .872 ตามลำดับ ซึ่งแบบสอบทั้งสองฉบับมีข้อสอบร่วมจำนวน 12 ข้อ 2) เกณฑ์การให้คะแนนที่ใช้ในการตรวจข้อสอบแบบสร้างคำตอบในแบบสอบฉบับใหม่และฉบับอ้างอิงที่มีความเที่ยงระหว่างผู้ประเมิน มีค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (r) คือ .96 และ .97 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ตามลำดับ 3) ผู้ตรวจให้คะแนน 2 คน ที่มีความสอดคล้องในการให้คะแนนระหว่างผู้ประเมินเป็นรายข้อสำหรับแบบสอบทั้ง 2 ฉบับ มีค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (r) ในแต่ละข้ออยู่ระหว่าง .86 - .99 และ .84 – 1.00 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ตามลำดับ โดยวิธีการตรวจทั้ง 3 วิธี คือ 1) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม 2) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม 3) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม การปรับเทียบคะแนนด้วยวิธีการปรับเทียบเชิงเส้นตรงทั้ง 2 วิธี และการวิเคราะห์ค่าสถิติบูทสแตรปจะใช้โปรแกรม R ส่วนการวิเคราะห์ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสองใช้โปรแกรม Microsoft excel

ผลการวิจัยพบว่า 1) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม เมื่อใช้วิธี CLE (Chained Linear Equating) มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และ ค่าความลำเอียง (bias) ต่ำที่สุด คือ .802, .251 และ .840 ตามลำดับ 2) ในบริบทของการวิจัยนี้วิธี CLE (Chained Linear Equating) มีค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) จากวิธีการตรวจวิธีที่ 1, 2 และ 3 คือ 1.421, 1.481 และ 1.481 ตามลำดับ ซึ่งเป็นค่าที่ต่ำกว่าวิธีของเลอวิน (Levine method) จากวิธีการตรวจวิธีที่ 1, 2 และ 3 คือ 1.717, 1.804 และ 1.793 ตามลำดับ

ภาควิชา วิจัยและจิตวิทยาการศึกษา ลายมือชื่อนิสิต .....

สาขาวิชา การวัดและประเมินผลการศึกษา ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก .....

ปีการศึกษา 2556

5583327827 : MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION
KEYWORDS: EQUATING DESIGNS / MIXED FORMAT TEST / BOOTSTRAPPED STATISTICS

PANIDA PANIDVADTANA: A COMPARISION OF BOOTSTRAPPED STATISTICS IN EQUATING DESIGNS FOR MIXED FORMAT TEST AMONG DIFFERENT METHODS. ADVISOR: ASST. PROF. NUTTAPORN LAWTHONG, Ph.D., 326 pp.

The purposes of this research were 1) to compare bootstrapped statistics of equating mixed-format test of nonequivalent groups with anchor test equating design among scoring methods 2) to compare root mean squared difference (RMSD) of equating mixed-format test of nonequivalent groups with anchor test equating design between Levine method and CLE (Chained linear equating) method. The sample were 1,607 students obtained by multi-stage random sampling from Matayomsuksa 4 student of Office of Bangkok Educational Service Area Zone I. They were used for data collecting with 1) two parallel mixed - format tests of the Physics, the work and energy for Matayomsuksa 4. There were new form and reference form. New form consisted 16 multiple-choice items and 8 construct-response items, their reliability values were .634 and .824. Reference form consisted 16 multiple-choice items and 8 construct-response items, with reliability value of .602 and .872. Those form had 8 multiple-choice items and 4 construct-response items in anchor test. 2) rubric scoring used for constructed-response items in Physics test, new and reference forms, had inter-rater reliability with Pearson's product moment correlation were .96 and .97 at the .05 level of significance. 3) two raters, who had consistency between two raters of two tests with Pearson's product moment correlation were .86 - .99 and .84 – 1.00 at the .05 level of significance. Three scoring methods were 1) constructed-response anchor items scoring methods with no trend scoring method 2) constructed-response anchor items scoring methods with full anchor test trend scoring method 3) constructed-response anchor items scoring methods with half anchor test trend scoring method. Equating score with linear equating method and analyzing bootstrapped statistics employed program R and analyzing Root Mean Squared Difference employed Microsoft excel.

The results of the study revealed that 1) constructed-response anchor items scoring methods with half anchor test trend scoring method with CLE (Chained linear equating) method resulted in smallest Root Mean Squared Error (RMSE), Equating error and Bias were .802, .251 and .840. 2) For this study, CLE (Chained linear equating) method had Root Mean Squared Difference (RMSD) from scoring methods of 1, 2 and 3 method were 1.421, 1.481 and 1.481, which lower than those of equating with Levine method were 1.717, 1.804 and 1.793.

Department: Educational Research and Psychology
Field of Study: Educational Measurement and Evaluation
Academic Year: 2013

Student's Signature .....
Advisor's Signature .....

กิตติกรรมประกาศ

การทำวิทยานิพนธ์ฉบับนี้ ผู้วิจัยได้รับความรู้และคำแนะนำต่างๆด้วยความกรุณาอย่างยิ่งจากอาจารย์ที่ปรึกษาวิทยานิพนธ์ คือ ผู้ช่วยศาสตราจารย์ ดร. ณัฏฐภรณ์ หลาวทอง ที่ได้ให้กรุณาให้คำปรึกษาแนะนำแนวทางที่ถูกต้อง ตลอดจนแก้ไขข้อบกพร่องต่าง ๆ ด้วยความละเอียดถี่ถ้วน อีกทั้งยังคอยกำกับติดตามอย่างสม่ำเสมอ จนวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี ผู้วิจัยรู้สึกซาบซึ้งเป็นอย่างมาก และขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณ ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี ประธานกรรมการสอบวิทยานิพนธ์ ดร. ชนะศึก นิชานนท์ กรรมการสอบวิทยานิพนธ์ ที่ได้สละเวลาอันมีค่ามาร่วมสอบวิทยานิพนธ์ และให้ข้อเสนอแนะที่เป็นประโยชน์ในการทำวิทยานิพนธ์ฉบับนี้จนเสร็จสมบูรณ์

ขอกราบขอบพระคุณ รองศาสตราจารย์ ดร. ศิริเดช สุชีวะ ที่ให้คำแนะนำเกี่ยวกับการออกแบบการปรับเทียบคะแนนและการตรวจให้คะแนน ซึ่งทำให้ผู้วิจัยได้ประเด็นที่ใช้ในการทำวิจัยในครั้งนี้

ขอกราบขอบพระคุณ อาจารย์ ดร. สิวะโชติ ศรีสุทธิยากร ที่ได้สละเวลาให้คำแนะนำเกี่ยวกับสถิติและการใช้งานโปรแกรม R จนผู้วิจัยได้ผลการวิเคราะห์ข้อมูลที่มีความสำคัญยิ่งสำหรับวิทยานิพนธ์ฉบับนี้

ขอกราบขอบพระคุณ รองศาสตราจารย์ ดร. โชติกา ภาษีผล ผู้ช่วยศาสตรจารย์ กมลวรรณ ตังธนกานนท์ อาจารย์ วรรณา นาคศรีอาภรณ์ และอาจารย์อมรรัตน์ บุบผโชติ รวมไปถึงผู้เชี่ยวชาญทุกท่านที่ได้ให้คำแนะนำที่เป็นประโยชน์สำหรับการสร้างเครื่องมือที่ใช้ในการวิจัย

ขอกราบขอบพระคุณคณาจารย์ภาควิชาวิจัยและจิตวิทยาทุกท่าน ที่ได้ให้ความรู้ สั่งสอน และห่วงใยนิสิตทุกคนด้วยความเมตตา ผู้วิจัยรู้สึกประทับใจและซาบซึ้งเป็นอย่างมาก

ขอขอบคุณบุคคลทุกท่านที่ไม่ได้ปรากฏชื่อในที่นี้ ที่มีส่วนช่วยเหลือในการทำวิทยานิพนธ์ฉบับนี้ จนสำเร็จลุล่วงไปด้วยดี และขอขอบใจเพื่อน ๆ สาขาการวัดและประเมินผลการศึกษาทุกคนที่คอยให้ความช่วยเหลือ และเป็นกำลังใจตลอดช่วงเวลาของการเรียนและการทำวิทยานิพนธ์ รวมไปถึงคุณพีชาณิกา เพชรสังข์ ที่ให้คำแนะนำเกี่ยวกับการใช้ระบบ e-thesis และคุณอนุสรณ์ ที่ให้ความช่วยเหลือเกี่ยวกับการใช้โปรแกรม R

ท้ายที่สุด ขอกราบขอบพระคุณ บิดา มารดา สำหรับความรัก ความห่วงใย ความเอาใจใส่และกำลังใจที่มีให้ผู้วิจัยเสมอมา วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้เพราะมีบิดา มารดาเป็นแรงจูงใจในการขับเคลื่อนจนเสร็จสมบูรณ์ คุณค่าและประโยชน์ที่มีต่อวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแก่คุณพ่อขวัญชัย และคุณแม่พิมพ์ทอง

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... จ

กิตติกรรมประกาศ ..... ฉ

สารบัญ ..... ช

สารบัญตาราง ..... ญ

สารบัญภาพ ..... ฒ

บทที่ 1 บทนำ 1

ความเป็นมาและความสำคัญของปัญหา 1

คำถามการวิจัย 6

วัตถุประสงค์ของการวิจัย 7

ขอบเขตของการวิจัย 8

คำจำกัดความเชิงปฏิบัติการ 9

ประโยชน์ที่ได้รับ 14

บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง 15

ตอนที่ 1 แนวคิดเชิงทฤษฎีเกี่ยวกับการปรับเทียบคะแนนระหว่างแบบสอบ 16

1.1 ความหมายของการปรับเทียบคะแนนระหว่างแบบสอบ 16

1.2 เงื่อนไขของการปรับเทียบคะแนนระหว่างแบบสอบ 17

1.3 คำศัพท์ที่เกี่ยวข้องกับการปรับเทียบคะแนนระหว่างแบบสอบ 18

ตอนที่ 2 การออกแบบการปรับเทียบคะแนน (equating designs) 19

2.1 การออกแบบวิธีการเก็บรวบรวมข้อมูล (designs for data collection) 19

2.2 การออกแบบวิธีการปรับเทียบคะแนน (equating methods) 23

2.3 การประเมินคุณภาพของการปรับเทียบคะแนน 53

2.4 การตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) สำหรับการปรับเทียบคะแนน 58

2.5 งานวิจัยที่เกี่ยวข้องกับรูปแบบการเชื่อมต่อ (linking designs) การออกแบบการปรับเทียบคะแนน (equating designs) สัดส่วนของข้อสอบร่วมในการปรับเทียบคะแนน และการตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) 59

กรอบแนวคิดในการวิจัย 69

หน้า

บทที่ 3 วิธีการดำเนินการวิจัย 70
ประชากร 70
กลุ่มตัวอย่าง 70
การออกแบบวิธีการเก็บรวบรวมข้อมูลเพื่อการปรับเทียบคะแนนของการวิจัย 72
วิธีการตรวจสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) 73
เครื่องมือที่ใช้ในการวิจัย 75
วิธีการดำเนินการสร้างแบบสอบ 75
วิธีดำเนินการสร้างเกณฑ์การให้คะแนน (rubric score) สำหรับข้อสอบแบบสร้างคำตอบ 86
วิธีคัดเลือกผู้ตรวจให้คะแนน และการเตรียมความพร้อมให้กับผู้ตรวจให้คะแนน 89
วิธีการจัดการสอบและการตรวจให้คะแนน 92
การเก็บรวบรวมข้อมูล 93
การวิเคราะห์ข้อมูล 94
บทที่ 4 ผลการวิเคราะห์ข้อมูล 100
ตอนที่ 1 ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน 101
ตอนที่ 2 ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ 106
ตอนที่ 3 ผลการวิเคราะห์ความสอดคล้องในการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบระหว่างผู้ตรวจ 1 และ 2 109
ตอนที่ 4 ผลการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) โดยใช้วิธีการปรับเทียบเชิงเส้นตรง 110
ตอนที่ 5 ผลการปรับเทียบคะแนนสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ โดยใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น 116
ตอนที่ 6 ผลการเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี ระหว่างวิธีของเลอวิน (Levine method) และวิธี CLE (Chained linear equating) 125
ตอนที่ 7 ผลการเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของวิธีการตรวจ 3 วิธีสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) 126

หน้า

บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ 132
สรุปผลการวิจัย 133
อภิปรายผล 138
ข้อเสนอแนะ 146
รายการอ้างอิง 148
ภาคผนวก ก 151
ภาคผนวก ข 153
ภาคผนวก ค 175
ภาคผนวก ง 191
ภาคผนวก จ 201
ภาคผนวก ฉ 218
ภาคผนวก ช 235
ภาคผนวก ซ 276
ภาคผนวก ฌ 289
ประวัติผู้เขียนวิทยานิพนธ์ 326

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญตาราง

หน้า

ตาราง 1 รูปแบบการเก็บรวบรวมข้อมูลสำหรับการปรับเทียบ 22
ตาราง 2 สถานการณ์ที่เหมาะสมกับวิธีการปรับเทียบคะแนนระหว่างแบบสอบแบบต่าง ๆ 43
ตาราง 3 สรุปประเด็นสำคัญของการออกแบบวิธีการปรับเทียบคะแนนเชิงเส้นตรง 49
ตาราง 4 จำนวนประชากรโรงเรียน นักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 จากสำนักงานเขตการศึกษามัธยมศึกษา เขต 1 ปีการศึกษา 2556 70
ตาราง 5 กลุ่มตัวอย่างโรงเรียนที่ทำแบบสอบเรื่องงานและพลังงาน รายวิชาฟิสิกส์ ระดับชั้นมัธยมศึกษาปีที่ 4 71
ตาราง 6 จำนวนกลุ่มตัวอย่างโรงเรียนและนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 จำแนกตามขนาดของโรงเรียน สำหรับการวิจัย สำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 1 ปีการศึกษา 2556 72
ตาราง 7 การออกแบบวิธีการเก็บรวบรวมข้อมูลเพื่อการปรับเทียบคะแนน 73
ตาราง 8 ตารางวิเคราะห์หลักสูตร เรื่องงานและพลังงาน รายวิชาฟิสิกส์ ระดับชั้นมัธยมศึกษาปีที่ 4 78
ตาราง 9 ผลการวิเคราะห์ค่าสถิติข้อสอบแบบหลายตัวเลือกในแบบสอบเรื่องงานและพลังงาน ฉบับใหม่และฉบับอ้างอิงจากการทดลองกับกลุ่มตัวอย่างจำนวน 217 คน 80
ตาราง 10 ผลการวิเคราะห์ค่าสถิติข้อสอบแบบสร้างคำตอบในแบบสอบเรื่องงานและพลังงาน ฉบับใหม่และอ้างอิง จากการทดลองกับกลุ่มตัวอย่างจำนวน 217 คน 81
ตาราง 11 การทดสอบความแตกต่างของค่าเฉลี่ยคะแนนในแบบสอบฉบับใหม่และฉบับอ้างอิงของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 82
ตาราง 12 รายละเอียดของข้อสอบร่วมแบบหลายตัวเลือกในแต่ละจุดประสงค์เชิงพฤติกรรมจำนวน 8 ข้อ 84
ตาราง 13 รายละเอียดของข้อสอบเฉพาะของข้อสอบแบบหลายตัวเลือกในแต่ละจุดประสงค์เชิงพฤติกรรมจำนวน 8 ข้อ 84
ตาราง 14 รายละเอียดของข้อสอบร่วมแบบสร้างคำตอบในแต่ละจุดประสงค์เชิงพฤติกรรมจำนวน 8 ข้อ 85
ตาราง 15 รายละเอียดของข้อสอบเฉพาะของข้อสอบแบบหลายตัวเลือกในแต่ละจุดประสงค์ เชิงพฤติกรรมจำนวน 8 ข้อ 85
ตาราง 16 รายละเอียดของแบบสอบย่อยแต่ละฉบับ 86
ตาราง 17 ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันระหว่างคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบของผู้ตรวจ 1 และผู้ตรวจ 2 สำหรับแบบสอบฉบับใหม่ 88

CHULALONGKORN UNIVERSITY

หน้า

ตาราง 18 ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันระหว่างคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบของผู้ตรวจคนที่ 1 และผู้ตรวจคนที่ 2 สำหรับแบบสอบฉบับอ้างอิง..... 89
ตาราง 19 ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันระหว่างคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบของผู้ตรวจ 1 และผู้ตรวจ 2 (แบบรายข้อ) สำหรับแบบสอบฉบับใหม่..... 91
ตาราง 20 ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันระหว่างคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบของผู้ตรวจ 1 และผู้ตรวจ 2 (แบบรายข้อ) สำหรับแบบสอบฉบับอ้างอิง..... 91
ตาราง 21 ผลการทดสอบความแตกต่างของค่าเฉลี่ยคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบของผู้ตรวจคน 1 และผู้ตรวจ 2 สำหรับแบบสอบฉบับใหม่ 91
ตาราง 22 ผลการทดสอบความแตกต่างของค่าเฉลี่ยคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบของผู้ตรวจ 1 และผู้ตรวจ 2 สำหรับแบบสอบฉบับอ้างอิง 92
ตาราง 23 รายชื่อโรงเรียนในกลุ่มใหม่ (new) และกลุ่มอ้างอิง (reference)..... 92
ตาราง 24 ช่วงเวลาในการเก็บข้อมูลกับกลุ่มตัวอย่างทั้ง 3 กลุ่ม..... 94
ตาราง 25 ค่าสถิติของกลุ่มผู้สอบที่ทำแบบสอบฉบับใหม่ (new) และฉบับอ้างอิง (reference) สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)..... 103
ตาราง 26 ค่าสถิติของกลุ่มผู้สอบอ้างอิงที่ใช้เป็นเกณฑ์ ที่ทำแบบสอบฉบับใหม่ (new) และฉบับอ้างอิง (reference) สำหรับการออกแบบการปรับเทียบที่เป็นเกณฑ์ 107
ตาราง 27 ค่าสัมประสิทธิ์สหพันธ์ของเพียร์สัน (r) ระหว่างการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบของผู้ตรวจ 1 ในวิธีการตรวจวิธีที่ 1: Non Trend และการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบของผู้ตรวจ 2 ในวิธีการตรวจวิธีที่ 2: Full Trend..... 109
ตาราง 28 ผลการวิเคราะห์คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกับผู้สอบกลุ่มอ้างอิง สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจทั้ง 3 วิธี ด้วยวิธี CLE (Chained Linear Equating).... 112
ตาราง 29 ผลการวิเคราะห์คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกับผู้สอบกลุ่มอ้างอิง สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจทั้ง 3 วิธี ด้วยวิธีของเลอวิน (Levine method)..... 115
ตาราง 30 ค่าสถิติพื้นฐานของคะแนนจากแบบสอบฉบับใหม่ (new) ที่ระดับ degree ต่างๆ 118
ตาราง 31 ผลการวิเคราะห์ค่า Deviance ของโมเดลต่างๆ 120
ตาราง 32 ค่าสถิติพื้นฐานของคะแนนจากแบบสอบฉบับอ้างอิง (reference) ที่ระดับ degree ต่างๆ 122

หน้า

ตาราง 33 ผลการวิเคราะห์ค่า Deviance ของโมเดลต่างๆ 123
ตาราง 34 ผลการวิเคราะห์คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกับผู้สอบกลุ่มอ้างอิง สำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ ด้วยวิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed)..... 124
ตาราง 35 ผลการวิเคราะห์ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ระหว่างวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)..... 126
ตาราง 36 ผลการวิเคราะห์สถิติบูทสแตรป (bootstrapped statistics) สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี 129
ตาราง 37 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับใหม่ (ฉบับที่ 1) 154
ตาราง 38 ผลการปรับปรุงข้อสอบวิชาฟิสิกส์ เรื่อง งานและพลังงาน ระดับชั้นมัธยมศึกษาปีที่ 4 (ฉบับที่1) ตามคำแนะนำของผู้เชี่ยวชาญ 162
ตาราง 39 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับอ้างอิง (ฉบับที่ 2) 165
ตาราง 40 ผลการปรับปรุงข้อสอบวิชาฟิสิกส์ เรื่อง งานและพลังงาน ระดับชั้นมัธยมศึกษาปีที่ 4 (ฉบับที่2) ตามคำแนะนำของผู้เชี่ยวชาญ 172
ตาราง 41 ผลการวิเคราะห์ข้อสอบแบบหลายตัวเลือกวิชาฟิสิกส์เรื่องงานและพลังงานฉบับใหม่ ... 192
ตาราง 42 ผลการวิเคราะห์ข้อสอบแบบหลายตัวเลือกวิชาฟิสิกส์เรื่องงานและพลังงานฉบับอ้างอิง 194
ตาราง 43 ผลการวิเคราะห์ข้อสอบแบบสร้างคำตอบวิชาฟิสิกส์เรื่องงานและพลังงานฉบับใหม่ 196
ตาราง 44 ผลการวิเคราะห์ข้อสอบแบบสร้างคำตอบวิชาฟิสิกส์เรื่องงานและพลังงานฉบับอ้างอิง.. 196
ตาราง 45 ผลการปรับปรุงข้อสอบวิชาฟิสิกส์ เรื่อง งานและพลังงานระดับชั้นมัธยมศึกษาปีที่ 4 ของแบบสอบฉบับใหม่และฉบับอ้างอิง..... 197
ตาราง 46 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับใหม่ (ครั้งที่ 1)..... 219
ตาราง 47 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับอ้างอิง (ครั้งที่ 1)..... 225

หน้า

ตาราง 48 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับใหม่และฉบับอ้างอิง ครั้งที่ 2 228
ตาราง 49 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบ ที่ 1 ด้วยวิธี CLE (Chained Linear Equating) 277
ตาราง 50 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบ ที่ 2 ด้วยวิธี CLE (Chained Linear Equating) 279
ตาราง 51 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบ ที่ 3 ด้วยวิธี CLE (Chained Linear Equating) 281
ตาราง 52 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธีของเลอวิน (Levine method) 283
ตาราง 53 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธีของเลอวิน (Levine method) 285
ตาราง 54 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธีของเลอวิน (Levine method) 287
ตาราง 55 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธี CLE (Chained Linear Equating) 290
ตาราง 56 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธี CLE (Chained Linear Equating) 292
ตาราง 57 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธี CLE (Chained Linear Equating) 294
ตาราง 58 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธี CLE (Chained Linear Equating) 296
ตาราง 59 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธี CLE (Chained Linear Equating) 298
ตาราง 60 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธี CLE (Chained Linear Equating) 300
ตาราง 61 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธี CLE (Chained Linear Equating) 302
ตาราง 62 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธี CLE (Chained Linear Equating) 304

หน้า

ตาราง 63 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธี CLE (Chained Linear Equating)..... 306
ตาราง 64 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธีของเลอวิน (Levine method)..... 308
ตาราง 65 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธีของเลอวิน (Levine method)..... 310
ตาราง 66 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธีของเลอวิน (Levine method)..... 312
ตาราง 67 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธีของเลอวิน (Levine method)..... 314
ตาราง 68 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธีของเลอวิน (Levine method)..... 316
ตาราง 69 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธีของเลอวิน (Levine method)..... 318
ตาราง 70 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธีของเลอวิน (Levine method)..... 320
ตาราง 71 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธีของเลอวิน (Levine method)..... 322
ตาราง 72 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธีของเลอวิน (Levine method)..... 324

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญภาพ

หน้า

ภาพ 1 การออกแบบการปรับเทียบ (equating designs) สำหรับแบบสอบรูปแบบผสม..... 62
ภาพ 2 การออกแบบการปรับเทียบ (equating designs) สำหรับแบบสอบแบบสร้างคำตอบ..... 63
ภาพ 3 การออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วม 68
ภาพ 4 กรอบแนวคิดในการวิจัย 69
ภาพ 5 กราฟการกระจายของคะแนนแบบสอบฉบับใหม่ทั้งฉบับและข้อสอบร่วมทั้งหมด สำหรับวิธีการตรวจ 1 102
ภาพ 6 กราฟการกระจายของคะแนนแบบสอบฉบับอ้างอิงทั้งฉบับและข้อสอบร่วมทั้งหมด สำหรับวิธีการตรวจ 1 102
ภาพ 7 การกระจายของคะแนนจากแบบสอบฉบับใหม่ทั้งฉบับและข้อสอบร่วมทั้งหมด สำหรับวิธีการตรวจ 2..... 104
ภาพ 8 การกระจายของคะแนนจากแบบสอบฉบับอ้างอิงทั้งฉบับและข้อสอบร่วมทั้งหมด สำหรับวิธีการตรวจ 2 104
ภาพ 9 การกระจายของคะแนนจากแบบสอบฉบับใหม่ทั้งฉบับและข้อสอบร่วมทั้งหมด สำหรับวิธีการตรวจ 3..... 105
ภาพ 10 การกระจายของคะแนนจากแบบสอบฉบับอ้างอิงและข้อสอบร่วมทั้งหมด สำหรับวิธีการตรวจ 3..... 106
ภาพ 11 การกระจายของคะแนนจากแบบสอบฉบับอ้างอิงทั้งฉบับและข้อสอบร่วมทั้งหมด สำหรับรูปแบบเกณฑ์ 108
ภาพ 12 การกระจายของคะแนนจากแบบสอบฉบับใหม่ทั้งฉบับและข้อสอบร่วมทั้งหมด สำหรับรูปแบบเกณฑ์ 108
ภาพ 13 การแจกแจงของคะแนนจากแบบสอบฉบับใหม่ เมื่อ degree =1 xdegree =1 118
ภาพ 14 การแจกแจงของคะแนนจากแบบสอบฉบับใหม่ เมื่อ degree =2 xdegree =1 118
ภาพ 15 การแจกแจงของคะแนนจากแบบสอบฉบับใหม่ เมื่อ degree =3 xdegree =1 118
ภาพ 16 การแจกแจงของคะแนนจากแบบสอบฉบับใหม่ เมื่อ degree =4 xdegree =1 118
ภาพ 17 การแจกแจงของคะแนนจากแบบสอบฉบับอ้างอิง เมื่อ degree =1 xdegree =1 121
ภาพ 18 การแจกแจงของคะแนนจากแบบสอบฉบับอ้างอิง เมื่อ degree =2 xdegree =1 121
ภาพ 19 การแจกแจงของคะแนนจากแบบสอบฉบับอ้างอิง เมื่อ degree =4 xdegree =1 121
ภาพ 20 การแจกแจงของคะแนนจากแบบสอบฉบับอ้างอิง เมื่อ degree =3 xdegree =1 121
ภาพ 21 ค่า $Biasi$ สำหรับวิธีการตรวจ 1 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี 129

CHULALONGKORN UNIVERSITY

หน้า

ภาพ 22 ค่า $SEEi$ สำหรับวิธีการตรวจ 1 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี 130
ภาพ 23 ค่า $RMSEi$ สำหรับวิธีการตรวจ 1 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี 130
ภาพ 24 ค่า $Biasi$ สำหรับวิธีการตรวจ 2 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี 130
ภาพ 25 ค่า $SEEi$ สำหรับวิธีการตรวจ 2 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี 130
ภาพ 26 ค่า $RMSEi$ สำหรับวิธีการตรวจ 2 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี 131
ภาพ 27 ค่า $Biasi$ สำหรับวิธีการตรวจ 3 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี 131
ภาพ 28 ค่า $SEEi$ สำหรับวิธีการตรวจ 3 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี 131
ภาพ 29 ค่า $RMSEi$ สำหรับวิธีการตรวจ 3 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี 131

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 1
บทนำ

ความเป็นมาและความสำคัญของปัญหา

การทดสอบขนาดใหญ่มีการใช้ข้อสอบแบบสร้างคำตอบ (construct-response: CR) ร่วมกับข้อสอบแบบหลายตัวเลือก (multiple-choice: MC) มากขึ้น โดยข้อสอบแบบหลายตัวเลือก (MC) มีจุดเด่น คือ การให้คะแนนมีความเป็นปรนัยและมีความเที่ยงสูง สามารถวัดได้คลอบคลุมเนื้อหา อีกทั้งยังสามารถวัดระดับผลการเรียนรู้ขั้นสูงได้ เช่น ความรู้ ความเข้าใจ การนำไปใช้ และการวิเคราะห์ ขณะที่ข้อสอบแบบสร้างคำตอบ (CR) ขาดความเป็นปรนัยและความเที่ยงในการให้คะแนน สร้างให้คลอบคลุมเนื้อหาทั้งหมดทำได้ยาก (Messick, 1994 cited in Kim et al., 2010) แต่อย่างไรก็ตาม ข้อสอบแบบสร้างคำตอบสามารถวัดระดับผลการเรียนรู้ขั้นสูงได้ดีกว่าข้อสอบแบบอื่น สร้างได้ง่าย อีกทั้งยังส่งเสริมผู้สอบได้แสดงความคิดเห็น ความคิดสร้างสรรค์ และทัศนคติ (ศิริชัย กาญจนวาสี, 2555) เนื่องจากข้อสอบแบบหลายตัวเลือกและแบบสร้างคำตอบต่างก็มีจุดแข็งและจุดอ่อน ดังนั้นในการทดสอบจึงมีการสร้างแบบสอบรูปแบบผสม ที่ประกอบด้วยข้อสอบแบบสร้างคำตอบและข้อสอบแบบหลายตัวเลือก ซึ่งข้อสอบแบบสร้างคำตอบมีการตรวจให้คะแนนแบบหลายค่า (polytomous) ส่วนข้อสอบแบบหลายตัวเลือกมีการตรวจให้คะแนนแบบสองค่า (dichotomous)

การทดสอบที่ใช้แบบสอบรูปแบบผสม เช่น การสอบคัดเลือกบุคคลเข้าระดับอุดมศึกษา โดยมีการกำหนดสัดส่วนข้อสอบปรนัยและอัตนัยที่แตกต่างกัน ซึ่งข้อสอบอัตนัยอาจเป็นแบบตอบสั้น เติมคำ เรียงความ ที่ผ่านมาการกำหนดสัดส่วนของคะแนนในข้อสอบอัตนัยต่ำกว่าข้อสอบปรนัย เนื่องจากข้อสอบอัตนัยใช้เวลาในทำข้อสอบมากกว่าข้อสอบปรนัย และความคลาดเคลื่อนในการให้คะแนนมากกว่า เช่น แบบสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) ที่มีทั้งข้อสอบปรนัยและอัตนัยในอัตราส่วน 70% - 90% : 30% - 10% โดยข้อสอบอัตนัยเป็นแบบเขียนตอบสั้นให้เวลาในการทำข้อสอบวิชาละ 2 ชั่วโมง ข้อสอบแต่ละข้ออาจมีคะแนนไม่เท่ากันขึ้นอยู่กับความยากของข้อสอบ โดยข้อสอบคลอบคลุมสาระและทักษะสำคัญของ 5 กลุ่มสาระการเรียนรู้ รวมไปถึงการทดสอบภาษาอังกฤษ TOEFL ประกอบด้วย 4 ส่วน คือ การฟัง ไวยากรณ์ การอ่าน และการเขียน ซึ่งมีทั้งข้อสอบที่ตรวจให้คะแนนสองค่า และหลายค่าอยู่ในฉบับเดียวกัน ในส่วนของการเขียน ผู้สอบทำข้อสอบ 1 ข้อ จากที่คอมพิวเตอร์สุ่มให้ และเขียนคำตอบเกี่ยวกับประเด็นเหล่านั้น และมีการให้คะแนนแบบหลายค่า 7 ระดับ คือ ระดับคะแนนตั้งแต่ 0 - 6 คะแนน

สถานการณ์ทดสอบปัจจุบัน มีความจำเป็นต้องใช้แบบสอบหลายฉบับในการวัดคุณลักษณะเดียว ซึ่งการพัฒนาแบบสอบแต่ละฉบับต้องมั่นใจว่าแบบสอบนั้น ๆ มีเนื้อหาและคุณลักษณะทางสถิติ (statistical characteristics) คล้ายคลึงกัน และสำหรับแบบสอบที่บรรจุข้อสอบแบบสร้างคำตอบต้องมีเกณฑ์การให้คะแนน (scoring rubrics) เฉพาะข้อสอบแต่ละข้อตลอดการทดสอบที่แตกต่างกัน (Kim et al., 2010) ด้วยเหตุนี้จึงมีความจำเป็นต้องใช้กระบวนการปรับเทียบคะแนน (equating process) เพื่อปรับเทียบผลที่ได้จากการทำแบบสอบต่างฉบับให้อยู่ในมาตรวัดเดียวกัน เพื่อให้คะแนน

ของแบบสอบต่างฉบับสามารถเปรียบกันได้ โดยการปรับเทียบคะแนนระหว่างแบบสอบเป็นกระบวนการที่เกี่ยวข้องกับ 2 กระบวนหลัก คือ กระบวนการที่ทำให้แบบสอบสองฉบับมีความเท่าเทียมกันในเชิงโครงสร้าง และการใช้สถิติเพื่อปรับเทียบคะแนนที่ได้จากแบบสอบแต่ละฉบับให้อยู่ในมาตรเดียวกันและเปรียบเทียบกันได้

การปรับเทียบคะแนนระหว่างแบบสอบต้องพิจารณาถึงการออกแบบการปรับเทียบคะแนน (equating designs) สำหรับแบบสอบรูปแบบผสม ข้อสอบแบบสร้างคำตอบไม่สามารถใช้ซ้ำในการทดสอบสองแบบที่แตกต่างกัน เพราะได้รับผลกระทบจากการจดจำ (Muraki et al., 2000) ดังนั้น การใช้ข้อสอบร่วมแบบสร้างคำตอบ (common constructed-response items) จึงไม่นิยมใช้ในการปรับเทียบคะแนน และในกรณีที่ใช้ข้อสอบร่วมเป็นข้อสอบแบบหลายตัวเลือกเพียงอย่างเดียว พบว่า ทำให้เกิดความลำเอียงในการปรับเทียบคะแนน (Kim & Kolen, 2006; Li, Lissitz, & Yang, 1999 cited in Kim et al., 2010) ซึ่งสอดคล้องกับงานวิจัยของ Kim และคณะ (2010) ที่พบว่าข้อสอบร่วมแบบหลายตัวเลือกเพียงอย่างเดียว มีค่าความความคลาดเคลื่อนในการปรับเทียบสูงกว่าการใช้ข้อสอบร่วมแบบผสม ข้อสอบร่วมที่ใช้จึงควรเป็นข้อสอบแบบหลายตัวเลือกผสมกับข้อสอบแบบสร้างคำตอบ

การใช้ข้อสอบแบบสร้างคำตอบอาจมีปัญหาเกี่ยวกับมาตรฐานการให้คะแนน จึงมีการประยุกต์การตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) โดยผู้ประเมินกลุ่มหนึ่งทำการตรวจข้อสอบร่วมแบบสร้างคำตอบทั้ง 2 กลุ่ม ซึ่ง Kim และคณะ (2010) พบว่ารูปแบบที่ใช้ข้อสอบร่วมรูปแบบผสม และมีการตรวจให้คะแนนข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ และรูปแบบที่ใช้กลุ่มผู้สอบเท่าเทียมกัน ไม่มีข้อสอบร่วม แต่มีตรวจให้คะแนนข้ามกลุ่มในข้อสอบแบบสร้างคำตอบ (ฉบับอ้างอิง (reference)) มีระดับความคลาดเคลื่อนของการปรับเทียบคะแนนระหว่างแบบสอบใกล้เคียงกันมาก โดยการออกแบบการปรับเทียบคะแนนทั้ง 2 รูปแบบมีค่าความลำเอียง (bias) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำกว่ารูปแบบที่ใช้ข้อสอบร่วมแบบหลายตัวเลือก และรูปแบบที่ใช้ข้อสอบร่วมแบบผสมที่ไม่มีการตรวจให้คะแนนข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ

การออกแบบวิธีการเก็บรวบรวมข้อมูล (designs for data collection) เพื่อการปรับเทียบคะแนนมีหลายแบบ ซึ่งแต่ละรูปแบบก็มีจุดเด่นแตกต่างกันไป เช่น รูปแบบผู้สอบกลุ่มเท่าเทียมกัน (equivalent group) สามารถหลีกเลี่ยงปัญหาเรื่องการเรียนรู้ และความเหนื่อยล้าจากการทำแบบสอบฉบับแรก ซึ่งเป็นลักษณะที่เกิดขึ้นกับรูปแบบผู้สอบกลุ่มเดียว (single - group design) และรูปแบบผู้สอบร่วม ที่ผู้สอบต้องทำแบบสอบถึงสองครั้ง แต่รูปแบบผู้สอบกลุ่มเท่าเทียมกันมีจุดอ่อน คือ ถ้ากลุ่มผู้สอบที่ใช้ไม่เท่าเทียมกันจริง อาจมีการแจกแจงความสามารถที่แตกต่างกัน และไม่มีข้อมูลในการนำมาปรับความแตกต่างของกลุ่ม ความแตกต่างที่เกิดขึ้นนี้ย่อมส่งผลกระทบต่อความลำเอียงในการปรับเทียบคะแนนระหว่างแบบสอบ โดยวิธีการที่ช่วยลดความแตกต่างนี้ได้ คือ การใช้กลุ่มตัวอย่างขนาดใหญ่ (ศิริชัย กาญจนวาสี, 2555) ทำให้ยากต่อการนำไปใช้กับสภาพจริง ด้วยเหตุนี้จึงมีการแก้ปัญหาโดยใช้ข้อสอบร่วม (common item/ anchor item) โดยเนื้อหาและระดับความยากของข้อสอบร่วมต้องมีความสัมพันธ์กับแบบสอบที่ต้องการปรับเทียบ การเก็บรวบรวมข้อมูลที่ได้รับความนิยมอีกรูปแบบหนึ่ง คือ รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกัน โดยใช้ข้อสอบร่วม (anchor-

test nonequivalent group design หรือ nonequivalence with anchor test : NEAT) ซึ่งผู้สอบแต่ละกลุ่มทำแบบสอบฉบับเดียว กลุ่มละฉบับ โดยแบบสอบแต่ละฉบับมีข้อสอบร่วมกันจำนวนหนึ่ง มักใช้กับสถานการณ์ที่กลุ่มผู้สอบต่างประชากรกัน เช่น กลุ่มผู้สอบต่างโปรแกรม เวลา หรือระดับ เป็นต้น (ศิริชัย กาญจนวาสี, 2555) เช่นเดียวกับ Kolen และ Brennan (2004) ที่กล่าวว่า รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกัน โดยใช้ข้อสอบร่วมนิยมนำมาใช้อย่างแพร่หลาย เพราะการทดสอบในสภาพจริงต้องการแบบสอบหลายฉบับที่วัดคุณลักษณะเดียวไปใช้ในการทดสอบในเวลาที่ต่างกัน โดยในงานวิจัยของ Kim และคณะ (2010) ได้ศึกษาการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

การปรับเทียบคะแนนต้องอาศัยวิธีการปรับเทียบคะแนน (equating methods) ในการเชื่อมโยงคะแนน ซึ่งมีอยู่หลายวิธี ได้แก่ วิธีการปรับเทียบค่าเฉลี่ย วิธีปรับเทียบเชิงเส้นตรง วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์ และวิธีปรับเทียบโดยใช้ทฤษฎีการตอบสนองข้อสอบ (IRT) ซึ่งแต่ละวิธีมีจุดเด่นและจุดด้อยที่แตกต่างกัน งานวิจัยที่ผ่านมาพบว่าวิธีการปรับเทียบโดยใช้ทฤษฎีการตอบสนองข้อสอบ (IRT) ให้ผลการปรับเทียบที่ดี แต่อย่างไรก็ตามวิธีการดังกล่าวจำเป็นต้องใช้กลุ่มข้อมูลขนาดใหญ่ ดังนั้นการปรับเทียบที่ใช้วิธีดังกล่าวจึงใช้การจำลองข้อมูลมากกว่าการเก็บข้อมูลจากการทดสอบจริง Kamata และ Tate (2005) กล่าวว่า วิธีการเชื่อมโยงโดยใช้ทฤษฎีการตอบสนองข้อสอบ (IRT) ไม่เหมาะสมในการนำมาใช้กับกลุ่มตัวอย่างขนาดเล็ก ในกรณีที่กลุ่มตัวอย่างมีขนาดเล็กนิยมใช้วิธีการปรับเทียบแบบดั้งเดิม เช่น วิธี CLE (Chained Linear Equating), วิธีของ Tucker (Tucker method), วิธีของเลอวิน (Levine method), วิธีการประมาณค่าความถี่ (frequently estimate) หรือวิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ (chained equipercentile equating) ซึ่งวิธีการเหล่านี้มักใช้เพื่อเชื่อมต่อแบบสอบรูปแบบผสม และจากการทบทวนวรรณกรรมที่เกี่ยวข้องยังไม่มีการศึกษาสถิติบูทสแตรป (bootstrapped statistics) ของการปรับเทียบคะแนนโดยใช้ทฤษฎี IRT เมื่อออกแบบให้มีการตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) และใช้ข้อมูลจากการทดสอบจริง

งานวิจัยที่ผ่านมาได้ศึกษาเปรียบเทียบประสิทธิภาพของวิธีการปรับเทียบคะแนนแบบดั้งเดิม โดย Suh และคณะ (2009) พบว่า ถ้ากลุ่มตัวอย่างแต่ละกลุ่มมีความสามารถแตกต่างกัน วิธีของเลอวิน (Levine method) มีค่าความลำเอียง (bias) และค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (Root Mean Squared Difference: RMSD) ต่ำกว่าวิธีอื่นๆ แต่ถ้ากลุ่มตัวอย่างมีความสามารถใกล้เคียงกัน วิธีการปรับเทียบเชิงเส้นทั้ง 5 มีคุณภาพดี ดังนั้นวิธีของเลอวิน (Levine method) จึงมีความเหมาะสมที่สุด เมื่อกลุ่มตัวอย่างแต่ละกลุ่มมีความสามารถแตกต่างกัน และวิธีการดังกล่าวมีความถูกต้องมากขึ้นเมื่อกลุ่มตัวอย่างมีขนาดใหญ่ สอดคล้องกับ Mroch และคณะ (2009) ที่พบว่า วิธีของเลอวิน (Levine method) เป็นตัวเลือกที่เหมาะสมที่สุดในวิธีการปรับเทียบเชิงเส้นตรง เมื่อกลุ่มตัวอย่างมีความแตกต่างกัน นอกจากนี้ Puhan (2010) ได้ศึกษาผลลัพธ์ของวิธีการปรับเทียบคะแนน 3 วิธี ได้แก่ วิธี CLE (Chained Linear Equating) วิธีของ Tucker (Tucker method) และวิธี Levine observed score ภายใต้สถานการณ์ที่กลุ่มตัวอย่างทำแบบสอบฉบับใหม่และฉบับเก่าที่มีความสามารถใกล้เคียงกัน และมีความสามารถแตกต่างกัน ซึ่งพบว่า วิธี CLE (Chained Linear Equating) หรือ วิธี Levine observed score สามารถ

นำมาใช้กับกลุ่มตัวอย่างใหม่และกลุ่มเก่าที่มีความสามารถแตกต่างกัน หรือเมื่อข้อสอบร่วมกับข้อสอบทั้งหมดมีความสัมพันธ์กันสูงได้ โดยวิธี CLE (Chained Linear Equating) ให้ผลการปรับเทียบคะแนนที่ดี เนื่องจากมีค่าความลำเอียง (bias) ต่ำ และมีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำ วิธีการปรับเทียบคะแนนแบบ Levine observed score method มีค่าความลำเอียงต่ำ และมีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำในบางกรณี ซึ่งในงานวิจัยของ Kim และคณะ (2010) ได้ใช้วิธี CLE (Chained Linear Equating) เพื่อปรับเทียบคะแนนแบบสอบรูปแบบผสมและแบบสอบแบบสร้างคำตอบ จากการวิจัยที่ผ่านมาสามารถสรุปได้ว่า วิธีการปรับเทียบที่เหมาะสมกับกลุ่มตัวอย่างขนาดเล็กและมีการเก็บข้อมูลจากการทดสอบจริง คือ วิธีของเลอวิน (Levine method) และวิธี CLE (Chained Linear Equating)

งานวิจัยที่ผ่านมาได้ศึกษาเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของการออกแบบการปรับเทียบคะแนน (equating designs) โดย Kim และคณะ (2010) ที่ได้ศึกษาสถิติบูทสแตรป (bootstrapped statistics) ของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) พบว่า รูปแบบที่ใช้ข้อสอบร่วมแบบหลายตัวเลือกเพียงอย่างเดียว หรือรูปแบบที่ใช้ข้อสอบร่วมแบบผสมระหว่างข้อสอบแบบหลายตัวเลือกและข้อสอบแบบสร้างคำตอบที่ไม่มีการตรวจให้คะแนนข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ มีความลำเอียงในการปรับเทียบคะแนนมากกว่าอีกสองรูปแบบที่เหลือ นอกจากนี้ยังพบว่ารูปแบบที่ใช้ผู้สอบกลุ่มเท่าเทียมกัน (equivalent group) และมีการตรวจให้คะแนนข้ามกลุ่มในข้อสอบแบบสร้างคำตอบมีค่าความลำเอียง (bias) ในการปรับเทียบคะแนนต่ำที่สุด สอดคล้องกับงานวิจัยของ Kim และคณะ (2010) ที่ได้ศึกษาสถิติบูทสแตรป (bootstrapped statistics) ของการออกแบบการปรับเทียบคะแนนแบบสอบแบบสร้างคำตอบโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) พบว่า รูปแบบที่ใช้ข้อสอบร่วมแบบสร้างคำตอบที่มีการตรวจให้คะแนนข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ และรูปแบบที่ใช้กลุ่มผู้สอบเท่าเทียมกัน ไม่มีข้อสอบร่วม แต่มีตรวจให้คะแนนข้ามกลุ่มในข้อสอบแบบสร้างคำตอบ มีระดับความคลาดเคลื่อนของการปรับเทียบคะแนนระหว่างแบบสอบต่ำและใกล้เคียงกันมาก ซึ่งแสดงให้เห็นว่าค่าความลำเอียง (bias) ในการปรับเทียบคะแนนระหว่างแบบสอบที่ทำให้คะแนนเกิดความคลาดเคลื่อน สามารถควบคุมได้โดยใช้การตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) เนื่องจากข้อสอบร่วมแบบสร้างคำตอบที่มีการตรวจให้คะแนนข้ามกลุ่ม (trend construct-response anchor) มีความคลาดเคลื่อนต่ำกว่าข้อสอบร่วมแบบสร้างคำตอบที่ไม่มีการตรวจให้คะแนนแบบข้ามกลุ่ม (no-trend construct-response anchor)

ผู้วิจัยจึงสนใจศึกษาการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ที่มีการตรวจให้คะแนนข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ และไม่มีการตรวจให้คะแนนข้ามกลุ่ม แต่ในการตรวจให้คะแนนข้ามกลุ่มในงานวิจัยที่ผ่านมานั้น ผู้ตรวจกลุ่มหนึ่ง (ผู้ตรวจกลุ่ม B) ต้องตรวจข้อสอบร่วมแบบสร้างคำตอบทั้งสองฉบับ ซึ่งผู้ตรวจกลุ่มดังกล่าวต้องรับหน้าที่ในการตรวจข้อสอบจำนวนมาก ด้วยเหตุนี้ผู้วิจัยจึงต้องการลดภาระของผู้ตรวจกลุ่มหนึ่ง/คนหนึ่งลง แต่ยังคงใช้การตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) ด้วยการออกแบบให้ผู้ตรวจแต่ละกลุ่ม (ผู้ตรวจกลุ่ม A และ B) ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่มอย่างละเท่าๆกัน โดยผู้ตรวจกลุ่ม B ตรวจข้อสอบร่วมแบบสร้างคำตอบจำนวน

ครึ่งหนึ่ง (ครึ่งแรก) ในแบบสอบทั้งสองฉบับ ส่วนผู้ตรวจกลุ่ม A ตรวจข้อสอบร่วมแบบสร้างคำตอบจำนวนครึ่งหนึ่ง (ครึ่งหลัง) ในแบบสอบทั้งสองฉบับ

นอกจากนี้ในงานวิจัยของ Kim และคณะ (2010) ใช้ผู้ตรวจ 2 กลุ่ม คือ ผู้ตรวจกลุ่ม A และผู้ตรวจกลุ่ม B ซึ่งมีวิธีการตรวจ ดังนี้ 1) วิธีการตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการให้คะแนนข้ามกลุ่ม (ผู้ตรวจกลุ่ม A ตรวจแบบสอบฉบับอ้างอิง (reference) และผู้ตรวจกลุ่ม B ตรวจแบบสอบฉบับใหม่ (new)) 2) วิธีการตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการให้คะแนนข้ามกลุ่ม (ผู้ตรวจกลุ่ม B ตรวจข้อสอบร่วมแบบสร้างคำตอบในแบบสอบฉบับอ้างอิง (reference) และฉบับใหม่ (new)) ผู้วิจัยจึงสนใจศึกษากรณีที่ใช้ผู้ตรวจ 2 คน คือ ผู้ตรวจ 1 และ ผู้ตรวจ 2 ซึ่งมีวิธีการตรวจดังนี้ 1) วิธีการตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการให้คะแนนข้ามกลุ่ม (ผู้ตรวจ 1 ตรวจแบบสอบฉบับอ้างอิง (reference) และผู้ตรวจ 2 ตรวจแบบสอบฉบับใหม่ (new)) 2) วิธีการตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจ 2 ตรวจข้อสอบร่วมแบบสร้างคำตอบในแบบสอบฉบับอ้างอิง (reference) และฉบับใหม่ (new)) 3) วิธีการตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม (ผู้ตรวจ 2 ตรวจข้อสอบร่วมแบบสร้างคำตอบ 2 ข้อแรกในแบบสอบฉบับอ้างอิง (reference) และฉบับใหม่ (new) ส่วนผู้ตรวจ 1 ตรวจข้อสอบร่วมแบบสร้างคำตอบ 2 ข้อหลังในแบบสอบฉบับอ้างอิง (reference) และฉบับใหม่ (new)) เพื่อเพิ่มทางเลือกในการนำไปใช้งานจริง และลดภาระงานของผู้ตรวจให้คะแนนลง ดังนั้นวิธีการตรวจสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ที่จะทำการศึกษาในครั้งนี้ ได้แก่ 1) วิธีการตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม ซึ่งผู้ตรวจ 1 ตรวจแบบสอบฉบับใหม่ (new) และผู้ตรวจ 2 ตรวจแบบสอบฉบับอ้างอิง (reference) 2) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจคนหนึ่ง ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม) โดยผู้ตรวจ 2 ตรวจแบบสอบฉบับใหม่ (new) และตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบกลุ่มใหม่และกลุ่มอ้างอิง ส่วนผู้ตรวจ 1 ตรวจแบบสอบฉบับอ้างอิงส่วนที่เหลืออยู่ 3) วิธีการตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม (ผู้ตรวจทั้ง 2 คน ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม โดยแบ่งกันตรวจคนละครึ่งหนึ่งของข้อสอบร่วม) โดยผู้ตรวจ 2 ตรวจแบบสอบฉบับใหม่ (new) และตรวจข้อสอบร่วมแบบสร้างคำตอบ (2 ข้อแรก) ของผู้สอบกลุ่มใหม่และกลุ่มอ้างอิง ส่วนผู้ตรวจ 1 ตรวจแบบสอบฉบับอ้างอิง (reference) และตรวจข้อสอบร่วมแบบสร้างคำตอบ (2 ข้อหลัง) ของผู้สอบกลุ่มใหม่และกลุ่มอ้างอิง นอกจากนี้ยังทำการเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ระหว่างวิธีการปรับเทียบเชิงเส้นตรง 2 วิธีร่วมด้วย ได้แก่ วิธีของเลอวิน (Levine method) และวิธี CLE (Chained Linear Equating) โดยใช้รูปแบบการใช้ผู้สอบกลุ่มเดียวเป็นเกณฑ์ในการเปรียบเทียบ ซึ่งกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์ต้องทำทั้งแบบสอบฉบับใหม่ (new) และฉบับอ้างอิง (reference) ที่มีข้อสอบร่วมระหว่างแบบสอบทั้งสองชุด (Kim et al., 2010; Puhan, 2010) โดยผู้ตรวจ 2 ตรวจแบบสอบฉบับใหม่ (new) และผู้ตรวจ 1 ตรวจแบบสอบฉบับอ้างอิง (reference) ซึ่งการได้มาของผลการวิจัยครั้งนี้ทำให้ได้ข้อสรุปเกี่ยวกับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

ที่เป็นประโยชน์ในทางปฏิบัติ คือ สามารถนำไปใช้ประกอบการตัดสินใจในการเลือกวิธีการปรับเทียบที่เหมาะสม และสอดคล้องกับสถานการณ์มากที่สุด และสามารถใช้เป็นแนวทางการบริหารการจัดการสอบขนาดเล็กที่ใช้แบบสอบรูปแบบผสมที่มีการทดสอบต่างกลุ่มผู้สอบหรือต่างเวลาให้สามารถนำคะแนนมาเทียบกันได้โดยตรง และข้อมูลเชิงเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) และค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ซึ่งเป็นประโยชน์ในทางวิชาการ คือ ได้ขยายองค์ความรู้เกี่ยวกับวิธีการตรวจ และวิธีการปรับเทียบคะแนน (equating method) สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

คำถามการวิจัย

งานวิจัยที่ผ่านมาได้มีการศึกษาการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) พบว่าการตรวจให้คะแนนแบบข้ามกลุ่ม โดยผู้ตรวจให้คะแนนคนเดียว ช่วยลดความลำเอียงในการปรับเทียบคะแนน แต่ในการตรวจให้คะแนนแบบข้ามกลุ่มในงานวิจัยที่ผ่านมานั้น ผู้ตรวจกลุ่มหนึ่ง (ผู้ตรวจกลุ่ม B) ต้องตรวจข้อสอบร่วมแบบสร้างคำตอบทั้งสองฉบับ ซึ่งผู้ตรวจกลุ่มดังกล่าวจึงต้องรับหน้าที่ในการตรวจข้อสอบจำนวนมาก ด้วยเหตุนี้ผู้วิจัยจึงต้องการลดภาระของผู้ตรวจกลุ่มหนึ่ง/คนหนึ่งลง แต่ยังคงใช้การตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) ด้วยการออกแบบให้ผู้ตรวจแต่ละกลุ่ม ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่มอย่างละเท่าๆกัน ซึ่งงานวิจัยที่ผ่านมาก็ยังไม่ได้มีการศึกษาการออกแบบการปรับเทียบในรูปแบบดังกล่าว และงานวิจัยที่ผ่านมาแบ่งผู้ประเมินเป็น 2 กลุ่ม ผู้วิจัยจึงต้องการเพิ่มทางเลือกในการปรับเทียบคะแนน โดยใช้ผู้ประเมินจำนวน 2 คน รวมไปถึงการศึกษาค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีการปรับเทียบเชิงเส้นตรงที่มีค่าต่ำที่สุดในการปรับเทียบคะแนน เมื่อกลุ่มตัวอย่างมีขนาดเล็ก มีความสามารถและช่วงเวลาในการสอบแตกต่างกัน ซึ่งได้แก่ วิธีของเลอวิน (Levine method) และวิธี CLE (Chained Linear Equating) ผู้วิจัยจึงได้กำหนดคำถามวิจัยไว้ดังนี้

1. วิธีการตรวจแบบใดที่สถิติบูทสแตรป (bootstrapped statistics) มีค่าต่ำที่สุด สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

2. วิธีการปรับเทียบเชิงเส้นตรงระหว่างวิธีของเลอวิน (Levine method) และวิธี CLE (Chained Linear Equating) วิธีการใดที่ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) มีค่าต่ำที่สุด สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีจุดมุ่งหมายหลักเพื่อศึกษาเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ระหว่างวิธีการตรวจ 3 วิธี และเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ระหว่างวิธีการปรับเทียบเชิงเส้นตรง 2 วิธีของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

1. เพื่อเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของการออกแบบการ ปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจที่แตกต่างกัน
2. เพื่อเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีของเลอวิน (Levine method) และวิธี CLE (Chained Linear Equating)

สมมติฐานการวิจัย

จากผลการศึกษาของ Kim และคณะ (2010) ได้ศึกษาเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ด้วยวิธี CLE (Chained Linear Equating) ผลการศึกษาพบว่ารูปแบบที่ใช้ข้อสอบร่วมแบบหลายตัวเลือกเพียงอย่างเดียว หรือรูปแบบที่ใช้ข้อสอบร่วมแบบผสมระหว่างข้อสอบแบบหลายตัวเลือกและข้อสอบแบบสร้างคำตอบที่ไม่มีการตรวจให้คะแนนข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบมีค่าความลำเอียง (bias) ในการปรับเทียบคะแนนมากกว่าอีกสองรูปแบบที่เหลือ นอกจากนี้ยังพบว่ารูปแบบที่ใช้ผู้สอบกลุ่มเท่าเทียมกัน (equivalent group) และมีการตรวจให้คะแนนข้ามกลุ่มในข้อสอบแบบสร้างคำตอบมีค่าความลำเอียง (bias) ในการปรับเทียบคะแนนต่ำที่สุด นอกจากนี้ Kim และคณะ (2010) ยังได้ศึกษาสถิติบูทสแตรป (bootstrapped statistics) ของการออกแบบการปรับเทียบคะแนนแบบสอบแบบสร้างคำตอบโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) พบว่า รูปแบบที่ใช้ข้อสอบร่วมแบบสร้างคำตอบที่มีการตรวจให้คะแนนข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ และรูปแบบที่ใช้กลุ่มผู้สอบเท่าเทียมกัน ไม่มีข้อสอบร่วม แต่มีตรวจให้คะแนนข้ามกลุ่มในแบบสอบแบบสร้างคำตอบ มีระดับความคลาดเคลื่อนของการปรับเทียบคะแนนระหว่างแบบสอบต่ำและใกล้เคียงกันมาก ซึ่งแสดงให้เห็นว่าค่าความลำเอียง (bias) ในการปรับเทียบคะแนนระหว่างแบบสอบที่ทำให้คะแนนเกิดความคลาดเคลื่อน สามารถควบคุมได้โดยใช้การตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) เนื่องจากข้อสอบร่วมแบบสร้างคำตอบที่มีการตรวจให้คะแนนข้ามกลุ่ม (trend constructed-response anchor) มีความคลาดเคลื่อนต่ำกว่าข้อสอบร่วมแบบสร้างคำตอบที่ไม่มีการตรวจให้คะแนนข้ามกลุ่ม (no-trend constructed-response anchor) ผู้วิจัยจึงได้กำหนดสมมติฐานของการวิจัยตามวัตถุประสงค์ข้อที่ 1 ดังนี้

1. วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งสองวิธีมีค่าสถิติบูทสแตรป (bootstrapped statistics) ต่ำกว่าวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบโดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

ในการศึกษาเปรียบเทียบวิธีการปรับเทียบ (equating method) โดย Suh และคณะ (2009) ได้เปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีการปรับเทียบคะแนนเชิงเส้นตรง 5 วิธี สำหรับรูปแบบผู้สอบกลุ่มไม่เท่าเทียมกัน โดยใช้ข้อสอบร่วม (NEAT) ถ้าแต่ละกลุ่มตัวอย่างมีความแตกต่างกันวิธีของเลอวิน (Levine method) มีค่าความลำเอียงและค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ต่ำกว่าวิธีอื่น ๆ สอดคล้องกับงานวิจัย Mroch และคณะ (2009) ที่ใช้ผลการศึกษาครั้งที่ผ่านมาของ Suh และคณะ (2009) และจากการทบทวนงานวิจัยที่เกี่ยวข้องกับสถิติบูทสแตรป (bootstrapped statistics) ของวิธีการปรับเทียบเชิงเส้นตรงมาใช้ในการประเมินวิธีการปรับเทียบเชิงเส้นตรงทั้ง 5 วิธี พบว่า วิธีของเลอวิน (Levine method) เป็นตัวเลือกที่เหมาะสมที่สุดในวิธีการปรับเทียบเชิงเส้นตรง เมื่อกลุ่มตัวอย่างมีความแตกต่างกันมาก นอกจากนี้ Puhan (2010) ได้ศึกษาผลลัพธ์ของวิธีการปรับเทียบคะแนน 3 วิธี ได้แก่ วิธี CLE (Chained Linear Equating), วิธีของ Tucker (Tucker method) และวิธี Levine observed score ซึ่งพบว่า วิธีCLE (Chained Linear Equating) ให้ผลการปรับเทียบคะแนนที่ดี เนื่องจากมีค่าความลำเอียง (bias) ต่ำ และมีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำ ขณะที่วิธี Levine observed score มีความลำเอียงต่ำ และมีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำในบางกรณี โดยเขาได้สรุปว่าวิธี CLE (Chained Linear Equating) และ วิธี Levine observed score สามารถนำมาใช้กับกลุ่มตัวอย่างสองกลุ่มที่มีความแตกต่างกัน หรือเมื่อข้อสอบร่วมกับข้อสอบทั้งหมดมีความสัมพันธ์กันสูงได้ ผู้วิจัยจึงได้กำหนดสมมติฐานของการวิจัยตามวัตถุประสงค์ข้อที่ 2 ดังนี้

2. ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีของเลอวิน (Levine method) มีค่าต่ำกว่าวิธี CLE (Chained Linear Equating) สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ขอบเขตของการวิจัย

การออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ในส่วนของวิธีการตรวจ ผู้วิจัยได้ศึกษา 3 วิธี ได้แก่ 1) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม 2) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม และ 3) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม ขณะที่วิธีการปรับเทียบคะแนน (equating method) ใช้วิธีการปรับเทียบแบบดั้งเดิม ซึ่งก็คือ วิธีการปรับเทียบเชิงเส้นตรง 2 วิธี ได้แก่ วิธีของเลอวิน (Levine method) วิธี CLE (Chained Linear Equating)

จากข้อมูลที่กล่าวมาข้างต้น ทำให้ผู้วิจัยกำหนดตัวแปรที่ใช้ในการวิจัยครั้งนี้ ได้ดังนี้

1) ตัวแปรที่ใช้ในการวิจัย

1.1 ตัวแปรอิสระ ได้แก่

1.1.1 วิธีการตรวจ 3 วิธี ดังนี้

- วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม
- วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม
- วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม

1.1.2 วิธีการปรับเทียบคะแนน (equating methods) 2 วิธี

- วิธีของเลอวิน (Levine method)
- วิธี CLE (Chained Linear Equating)

1.2 ตัวแปรตาม คือ

1.2.1 สถิติบูทสแตรป (bootstrapped statistics) ของการปรับเทียบคะแนน ได้แก่ ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ค่าความลำเอียง (bias) และค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) ซึ่งใช้ในการเปรียบเทียบวิธีการตรวจทั้ง 3 วิธี

1.2.2 ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ซึ่งใช้ในการเปรียบเทียบวิธีการปรับเทียบเชิงเส้นตรงทั้ง 2 วิธี

คำจำกัดความเชิงปฏิบัติการ

การออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) คือ การออกแบบการเก็บรวบรวมข้อมูลที่สามารถเชื่อมโยงกันได้ (linking designs) โดยการทดสอบใช้ข้อสอบร่วม (common items) อยู่ในแบบสอบแต่ละฉบับ ซึ่งผู้สอบทั้ง 2 กลุ่มมาจากประชากรต่างกลุ่มกัน ได้แก่ ผู้สอบกลุ่มอ้างอิงจำนวน 400 คน ที่ทำแบบสอบฉบับอ้างอิง (reference) จำนวน 24 ข้อ คะแนนเต็ม 48 คะแนน (MC 16 ข้อ คะแนนเต็มข้อละ 1 คะแนน รวมเป็น 16 คะแนน และ CR 8 ข้อ คะแนนเต็มข้อละ 4 คะแนน รวมเป็น 32 คะแนน) และผู้สอบกลุ่มใหม่จำนวน 1,609 คน ที่ทำแบบสอบฉบับใหม่ (new) จำนวน 24 ข้อ คะแนนเต็ม 48 คะแนน (MC 16 ข้อ คะแนนเต็มข้อละ 1 คะแนน รวมเป็น 16 คะแนน และ CR 8 ข้อ คะแนนเต็มข้อละ 4 คะแนน รวมเป็น 32 คะแนน) โดยแบบสอบทั้งสองฉบับมีข้อสอบร่วมกันอยู่จำนวน 12 ข้อ คะแนนเต็ม 24 คะแนน (MC 8 ข้อ คะแนนเต็มข้อละ1 คะแนน รวมเป็น 8 คะแนน และ CR 4 ข้อ คะแนนเต็มข้อละ 4 คะแนน รวมเป็น 16 คะแนน)

การออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ (criterion design) คือ การออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ในการเปรียบเทียบระหว่างวิธีการปรับเทียบเชิงเส้นตรงทั้ง 2 วิธี โดย ผู้สอบกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์จำนวน 125 คน ทำแบบสอบฉบับใหม่ (new) ก่อน หลังจากนั้นประมาณ 3 สัปดาห์ ผู้สอบกลุ่มดังกล่าวทำแบบสอบฉบับอ้างอิง (reference) ซึ่งแบบสอบทั้งสองฉบับมีข้อสอบร่วมรูปแบบผสมจำนวน 12 ข้อ ซึ่งการออกแบบการปรับเทียบดังกล่าวใช้เชื่อมต่อคะแนนจากแบบสอบฉบับใหม่ (new) ที่ได้จากผู้ตรวจ 2 ไปยังคะแนนจากแบบสอบฉบับอ้างอิง (reference) ที่ได้จากผู้ตรวจ 1 โดยใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed)

แบบสอบรูปแบบผสม คือ แบบสอบรูปแบบผสม 2 ฉบับ ซึ่งเป็นแบบสอบคู่ขนาน ได้แก่ แบบสอบฉบับใหม่ (new) จำนวน 24 ข้อ ประกอบด้วยข้อสอบแบบหลายตัวเลือก 16 ข้อ และข้อสอบแบบสร้างคำตอบ 8 ข้อ และแบบสอบฉบับอ้างอิง (reference) จำนวน 24 ข้อ ประกอบด้วยข้อสอบแบบหลายตัวเลือก 16 ข้อ และข้อสอบแบบสร้างคำตอบ 8 ข้อ โดยคะแนนเต็มของแบบสอบฉบับใหม่และอ้างอิง เท่ากับ 48 คะแนน ซึ่งแบบสอบทั้งสองฉบับมีข้อสอบร่วมรูปแบบผสม จำนวน 12 ข้อ

ข้อสอบร่วมรูปแบบผสม คือ ข้อสอบร่วมที่อยู่ในแบบสอบแต่ละฉบับ จำนวน 12 ข้อ ประกอบด้วยข้อสอบแบบหลายตัวเลือกจำนวน 8 ข้อ และข้อสอบแบบสร้างคำตอบจำนวน 4 ข้อ โดยข้อสอบแบบหลายตัวเลือกมีคะแนนข้อละ 1 คะแนน รวมเป็น 8 คะแนน และข้อสอบแบบสร้างคำตอบมีคะแนนข้อละ 4 คะแนน รวมเป็น 16 คะแนน ดังนั้นคะแนนรวมของข้อสอบร่วมรูปแบบผสมจึงเท่ากับ 24 คะแนน

วิธีการตรวจ คือ วิธีการตรวจให้คะแนนข้อสอบร่วมแบบสร้างคำตอบที่ใช้สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ซึ่งประกอบด้วย 3 วิธี ได้แก่ 1) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม 2) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม และ 3) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม

การตรวจให้คะแนนข้ามกลุ่ม คือ การตรวจให้คะแนนข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ ซึ่งมี 2 รูปแบบ คือ 1) ผู้ตรวจคนหนึ่ง ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม โดยผู้ตรวจ 2 ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบกลุ่มใหม่และอ้างอิง ส่วนผู้ตรวจ 1 ตรวจแบบสอบฉบับอ้างอิง (reference) ส่วนที่เหลืออยู่ของผู้สอบกลุ่มอ้างอิง 2) ผู้ตรวจทั้ง 2 คน ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม โดยแบ่งกันตรวจคนละครึ่งหนึ่งของข้อสอบร่วม โดยผู้ตรวจ 2 ตรวจแบบสอบฉบับใหม่ (new) และตรวจข้อสอบร่วมแบบสร้างคำตอบ (2 ข้อแรก) ของผู้สอบกลุ่มใหม่และกลุ่มอ้างอิง ส่วนผู้ตรวจ 1 ตรวจแบบสอบฉบับอ้างอิง (reference) และตรวจข้อสอบร่วมแบบสร้างคำตอบ (2 ข้อหลัง) ของผู้สอบกลุ่มใหม่และกลุ่มอ้างอิง

วิธีการปรับเทียบคะแนน (equating methods) คือ วิธีการที่ใช้ในการปรับเทียบคะแนนระหว่างแบบสอบรูปแบบผสมวิชาฟิสิกส์ 2 ฉบับ ซึ่งประกอบด้วยข้อสอบหลายตัวเลือกจำนวน 24 ข้อ และข้อสอบแบบสร้างคำตอบจำนวน 12 ข้อ โดยแบบสอบแต่ละฉบับข้อสอบร่วมที่อยู่ในแบบสอบแต่ละฉบับจำนวน 12 ข้อ

วิธีของเลอวิน (Levine method) คือ วิธีการปรับเทียบเชิงเส้นตรงที่ใช้ทฤษฎีการทดสอบแบบดั้งเดิม โดยใช้วิธี Levine observe score method ซึ่งเป็นการปรับเทียบคะแนนจากแบบสอบฉบับใหม่ (new) ไปยังคะแนนจากแบบสอบอ้างอิง (reference) ซึ่งคะแนนที่สังเกตได้ของแบบสอบฉบับใหม่ (new: X) มีความสัมพันธ์กับคะแนนที่สังเกตได้ของแบบสอบฉบับอ้างอิง (reference: Y) ดังสมการด้านล่าง โดยวิเคราะห์ผ่านโปรแกรม R (Albano, 2014)

$$l_{Ys}(x) = \left[\frac{\gamma_2}{\gamma_1}\right][x - \mu_1(X)] + \mu_2(Y) + \gamma_2[\mu_1(V) - \mu_2(V)]$$

วิธี CLE (Chained Linear Equating) คือ วิธีการปรับเทียบเชิงเส้นตรงที่ใช้ทฤษฎีการทดสอบแบบดั้งเดิมอีกวิธีหนึ่ง ซึ่งเป็นการแปลงคะแนนจากคะแนนของแบบสอบฉบับใหม่ (new: X) ไปยังคะแนนของข้อสอบร่วม (A) และแปลงคะแนนจากคะแนนของแบบสอบฉบับอ้างอิง (reference: Y) ไปยังคะแนนของข้อสอบร่วม (A) ดังสมการด้านล่าง โดยวิเคราะห์ผ่านโปรแกรม R (Albano, 2014)

$$y = f(x) = \mu_2(Y) + \frac{\sigma_2(Y)\sigma_1(A)}{\sigma_2(A)\sigma_1(X)}[x - \mu_1(X)] + \frac{\sigma_2(Y)}{\sigma_2(A)}[\mu_1(A) - \mu_2(A)]$$

วิธีปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (loglinear presmoothed method) คือ วิธีการที่ใช้ในการปรับการแจกแจงคะแนนของแบบสอบทั้งสองฉบับให้มีลักษณะเป็นโค้งปกติที่ราบเรียบ โดยการแจกแจงคะแนนที่ถูกปรับให้ราบเรียบ (smoothed distribution) ต้องมีค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับการแจกแจงที่ยังไม่ได้ถูกปรับให้เป็นโค้งราบเรียบ (unsmoothed distribution) โดยใช้วิธีล็อกเชิงเส้น (loglinear) ในการหาโมเดล (nested model) ที่มีความเหมาะสมกับข้อมูล กำหนดให้ degree = 3 xdegree = 1สำหรับแบบสอบทั้งสองฉบับ ซึ่งวิเคราะห์ผ่านโปรแกรม R (Albano, 2014)

วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed) คือ วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ (chained equipercentile) ที่มีการปรับการแจกแจงคะแนนของแบบสอบทั้งสองฉบับให้มีความราบเรียบต่อเนื่อง (presmoothing method) ก่อนทำการปรับเทียบคะแนน โดยอาศัยวิธีล็อกเชิงเส้น (loglinear) ซึ่งวิธีการปรับเทียบดังกล่าววิเคราะห์ผ่านโปรแกรม R (Albano, 2014) โดยฟังก์ชั่นการเชื่อมโยงอีควิเปอร์เซ็นไทล์ (equipercentile linking function) คือ $echain_Y(x) = e_{Y2}[e_{V2}(x)]$ ซึ่งวิธีการปรับเทียบนี้ใช้สำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ (criterion design)

สถิติบูทสแตรป (bootstrapped statistics) คือ ค่าสถิติที่ใช้วัดความแม่นยำ (the measure of accuracy) ของวิธีการตรวจทั้ง 3 วิธี สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ได้แก่ ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง ค่าความลำเอียง และค่าความคลาดเคลื่อนในการปรับเทียบในแต่ละจุดคะแนน (ค่า $RMSE_i$ ค่า $Bias_i$ และค่า SEE_i ตามลำดับ) ซึ่งค่าสถิติที่ได้มาจากกลุ่มตัวอย่างที่ถูกสุ่มโดยใช้คอมพิวเตอร์ และเมื่อคำนวณค่าสถิติได้แล้ว กลุ่มตัวอย่างดังกล่าวถูกส่งกลับไปยังประชากรอีกครั้ง โดยทำการสุ่มซ้ำจำนวน 500 ครั้ง โดยวิเคราะห์ผ่านโปรแกรม R (Albano, 2014) หลังจากนั้นนำค่าสถิติบูทสแตรปดังกล่าวมาคำนวณแบบถ่วงน้ำหนัก เพื่อใช้เป็นค่าในการเปรียบเทียบวิธีการตรวจทั้ง 3 วิธี ซึ่งได้แก่ ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ค่าความลำเอียง (bias) และค่าความคลาดเคลื่อนในการปรับเทียบ (equating error)

ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (Root Mean Squared Error : RMSE) คือ รากที่สองของผลรวมระหว่างค่าความลำเอียง (bias) ยกกำลังสอง และความคลาดเคลื่อนมาตรฐานอย่างมีเงื่อนไขของการปรับเทียบคะแนน (CSEE) ยกกำลังสอง โดยค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสองในแต่ละจุดคะแนน ($RMSE_i$) วิเคราะห์ผ่านโปรแกรม R (Albano, 2014) ดังสมการ

$$RMSE_i = \sqrt{{Bias_i}^2 + {SEE_i}^2}$$

หลังจากนั้นนำค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสองที่ได้ในแต่ละจุดคะแนน ($RMSE_i$) มาหาค่า Weighted average RMSE ดังสมการ $\sqrt{\sum_i w_i RMSE_i^2}$ เพื่อให้ได้ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) สำหรับวิธีการตรวจในแต่ละวิธี

ค่าความลำเอียง (bias) คือ ค่าความลำเอียงในการปรับเทียบคะแนน ซึ่งแปรผันตามค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) โดยค่าความลำเอียง (bias) สามารถหาได้จากผลต่างระหว่างคะแนนปรับเทียบจากแบบสอบฉบับใหม่ (new) ที่อยู่บนสเกลของแบบสอบฉบับอ้างอิง (reference) ในแต่ละรูปแบบของการปรับเทียบกับฟังก์ชั่นการปรับเทียบที่ใช้เป็นเกณฑ์ โดยค่าความลำเอียงในแต่ละจุดคะแนน ($Bias_i$) วิเคราะห์ผ่านโปรแกรม R (Albano, 2014) ดังสมการ

$$Bias_i = \hat{\bar{e}}_Y(x) - e_Y(x)$$

เมื่อ $$\hat{\bar{e}}_Y(x) = \frac{1}{R}\sum_{r=1}^{R} \hat{e}_{Yr}(x)$$

หลังจากนั้นนำค่าความลำเอียง (bias) ที่ได้ในแต่ละจุดคะแนน ($Bias_i$) มาหาค่า Weighted average root mean square bias ดังสมการ $\sqrt{\sum_i w_i Bias_i^2}$ เพื่อให้ได้ค่าความลำเอียง (bias) สำหรับวิธีการตรวจในแต่ละวิธี

ค่าความเคลื่อนในการปรับเทียบ (equating error) คือ ค่าความเคลื่อนมาตรฐานอย่างมีเงื่อนไขของการปรับเทียบได้ (conditional standard error of equating, CSEE) หรือความคลาดเคลื่อนเนื่องจากความผันแปรของการสุ่มตัวอย่าง ซึ่งก็คือส่วนเบี่ยงเบนมาตรฐานของความแตกต่างที่จุดของคะแนนที่สังเกตได้ x_i โดยค่าค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ในแต่ละจุดคะแนน (SEE_i) วิเคราะห์ผ่านโปรแกรม R (Albano, 2014) ดังสมการ

$$SEE_i = \sqrt{\frac{1}{R}\sum_{r=1}^{R}[\hat{e}_Y(x) - e_Y(x)]^2}$$

หลังจากนั้นนำค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ที่ได้ในแต่ละจุดคะแนน (SEE_i) มาหาค่า Weighted average equating error ดังสมการ $\sqrt{\sum_i w_i SEE_i^2}$ เพื่อให้ได้ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) สำหรับวิธีการตรวจในแต่ละวิธี

รากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (Root Mean Squared Difference : RMSD) คือ ค่าน้ำหนักความแตกต่างระหว่างการแปลงคะแนน หาได้จากรากที่สองของผลรวมของผลต่างระหว่างคะแนนปรับเทียบจากแบบสอบฉบับใหม่ (new) ที่อยู่บนสเกลของฉบับอ้างอิง (reference) ด้วยวิธีการปรับเทียบแบบต่าง ๆ ในแต่ละวิธีการตรวจจากการปรับเทียบที่ใช้เป็นเกณฑ์ยกกำลังสองแล้วกับคูณกับสัดส่วนของผู้สอบที่ทำแบบสอบฉบับใหม่ โดยค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ใช้ในการพิจารณาว่าวิธีการปรับเทียบใดที่มีความเหมาะสมมากที่สุดในบริบทที่ผู้วิจัยทำการศึกษา

$$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$$

เมื่อ i คือ จุดของคะแนนที่สังเกตได้

$\hat{e}_i(x_i)$ คือ คะแนนที่ถูกปรับเทียบโดยใช้วิธีการปรับเทียบแบบต่าง ๆ ที่คะแนนที่สังเกตได้ x

$e_i(x_i)$ คือ ฟังก์ชั่นของการปรับเทียบที่เป็นเกณฑ์ที่คะแนนสังเกตได้ x ในที่นี้ คือ คะแนนที่ถูกปรับเทียบภายใต้การออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ (SG design)

w_i คือ สัดส่วนสัมพัทธ์ของผู้สอบที่ทำแบบสอบฉบับใหม่ที่แต่ละจุดของคะแนน

วิธีการปรับเทียบใดที่มีค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ต่ำ ๆ แสดงว่ามีความแม่นยำในการปรับเทียบมากกว่าวิธีการปรับเทียบที่มีค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) สูง ๆ

ประโยชน์ที่ได้รับ

1. ประโยชน์ทางด้านวิชาการ

1.1) ได้ขยายองค์ความรู้เกี่ยวกับวิธีการตรวจสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

1.2) ได้ขยายองค์ความรู้เกี่ยวกับวิธีการปรับเทียบคะแนน (equating methods) สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

2. ประโยชน์ทางด้านการนำไปใช้

2.1) ผลการวิจัยสามารถนำไปใช้ประกอบการตัดสินใจในการเลือกวิธีการปรับเทียบคะแนน และวิธีการตรวจได้อย่างเหมาะสมและสอดคล้องกับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) มากที่สุด

2.2) ผลการวิจัยสามารถใช้เป็นแนวทางการบริหารการจัดการสอบขนาดเล็ก ที่ใช้แบบสอบรูปแบบผสม เมื่อมีการทดสอบต่างกลุ่มผู้สอบ หรือต่างเวลาให้สามารถนำคะแนนมาเทียบกันได้โดยตรง ซึ่งการปรับเทียบคะแนนเป็นเทคนิคอย่างหนึ่งที่ทำให้การวัดและประเมินผลมีความยุติธรรมมากยิ่งขึ้น

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการนำเสนอเอกสารและงานวิจัยที่เกี่ยวข้องของการวิจัยนี้ ผู้วิจัยได้นำเสนอโดยจำแนกออกเป็น 2 ตอนดังนี้

ตอนที่ 1 แนวคิดเชิงทฤษฎีเกี่ยวกับการปรับเทียบคะแนนระหว่างแบบสอบ

- 1.1 ความหมายของการปรับเทียบคะแนน
- 1.2 เงื่อนไขของการปรับเทียบคะแนน
- 1.3 คำศัพท์ที่เกี่ยวข้องการปรับเทียบคะแนน

ตอนที่ 2 การออกแบบการปรับเทียบคะแนน (equating designs)

- 2.1 การออกแบบวิธีการเก็บรวบรวมข้อมูล (designs for data collection)
  - 2.1.1 รูปแบบผู้สอบกลุ่มเดียว (single-group design)
  - 2.1.2 รูปแบบผู้สอบกลุ่มเท่าเทียมกัน (equivalent-group design)
  - 2.1.3 รูปแบบผู้สอบต่างกลุ่มโดยใช้แบบสอบร่วม (anchor-test design)
  - 2.1.4 รูปแบบผู้สอบต่างกลุ่มโดยมีผู้สอบร่วม (common-person design)
- 2.2 การออกแบบวิธีการปรับเทียบคะแนน (designs for equating method)
  - 2.2.1 วิธีการปรับเทียบคะแนนค่าเฉลี่ย
  - 2.2.2 วิธีการปรับเทียบเชิงเส้นตรง
  - 2.2.3 วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์
  - 2.2.4 วิธีการปรับเทียบโดยใช้ทฤษฎีการตอบสนองข้อสอบ
  - 2.2.5 สถานการณ์ที่เหมาะสมกับวิธีการปรับเทียบคะแนนระหว่างแบบสอบต่างๆ
  - 2.2.6 งานวิจัยที่เกี่ยวข้องกับการออกแบบวิธีการปรับเทียบคะแนน (equating method)
- 2.3 การประเมินของการปรับเทียบคะแนน
  - 2.3.1 ความเคลื่อนมาตรฐานของการปรับเทียบ (standard error of equating)
  - 2.3.2 ความเพียงพอของการปรับเทียบ (adequacy of equating models)
  - 2.3.3 สถิติบูทสแตรป (bootstrapped statistics)
- 2.4 การตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) สำหรับการปรับเทียบคะแนน
- 2.5 งานวิจัยที่เกี่ยวข้องกับรูปแบบการเชื่อมต่อ (linking designs) การออกแบบการปรับเทียบคะแนน (equating designs) สัดส่วนของข้อสอบร่วมในการปรับเทียบคะแนน และการตรวจให้คะแนนข้ามกลุ่ม (trend scoring method)

ตอนที่ 1 แนวคิดเชิงทฤษฎีเกี่ยวกับการปรับเทียบคะแนนระหว่างแบบสอบ

แนวคิดเชิงทฤษฎีเกี่ยวกับการปรับเทียบคะแนนระหว่างแบบสอบ ประกอบด้วย ความหมายของการปรับเทียบคะแนน เงื่อนไขของการปรับเทียบคะแนน และคำศัพท์ที่เกี่ยวข้องการปรับเทียบคะแนน

1.1 ความหมายของการปรับเทียบคะแนนระหว่างแบบสอบ

Gulliksen (1950 อ้างอิงใน ศิริชัย กาญจนวาสี, 2555) กล่าวว่า การปรับเทียบคะแนนระหว่างแบบสอบ คือ วิธีการปรับคะแนนจากแบบสอบสองฉบับที่วัดเนื้อหาเดียวกันให้เป็นคะแนนที่สมมูล (Equivalent Score) เพื่อให้คะแนนของแบบสอบทั้งสอบฉบับสามารถนำมาเปรียบเทียบกันได้ โดยมีวิธีการอย่างง่ายในการปรับเทียบคะแนน คือ ให้ผู้สอบกลุ่มเดียวทำแบบสอบทั้งสองฉบับที่พยายามสร้างให้คู่ขนานกัน แล้วคำนวณหาคะแนนมาตรฐานของแบบสอบแต่ละฉบับ คะแนนแปลงที่ตรงกับคะแนนมาตรฐานเดียวกัน ถือว่าเป็นคะแนนที่สมมูลกัน สามารถนำมาเทียบกันได้โดยตรง ซึ่งสอดคล้องกับ Lord (1980 อ้างอิงใน ศิริชัย กาญจนวาสี, 2555) ที่กล่าวว่า การปรับเทียบคะแนนระหว่างแบบสอบ คือ วิธีการแปลงคะแนนของแบบสองฉบับให้มีความหมายที่สับเปลี่ยนกันได้ เพื่อให้เกิดความเสมอภาคระหว่างผู้สอบ โดยที่ Lord (1980 อ้างอิงใน พัชรี จันเพ็ง, 2547) ยังได้กำหนดวัตถุประสงค์ของการปรับเทียบคะแนนระหว่างแบบสอบ คือ เพื่อกำหนดความเท่าเทียมระหว่างคะแนนที่สังเกตได้จากแบบสอบต่างฉบับ โดยการปรับเทียบคะแนนระหว่างแบบสอบเป็นวิธีการในการหาหลักฐานเชิงประจักษ์ จึงต้องมีการกำหนดแบบแผนในการเก็บรวบรวมข้อมูลและกฎเกณฑ์ของการปรับเทียบคะแนนของแบบสอบฉบับหนึ่งไปยังแบบสอบอีกฉบับ โดยสอดคล้องกับ Petersen และคณะ (1989 อ้างถึงใน พัชรี จันทร์เพ็ง, 2547) ที่กล่าวว่า การปรับเทียบคะแนนระหว่างแบบสอบ คือ กระบวนการเชิงประจักษ์ในการสร้างระบบความสัมพันธ์ระหว่างคะแนนของแบบสอบต่างฉบับ เพื่อแปลงคะแนนของแบบฉบับหนึ่งไปสู่คะแนนของแบบสอบอีกฉบับหนึ่ง และในแนวคิดของ Kolen และ Brennan (2004) การปรับเทียบคะแนนระหว่างแบบสอบ คือ กระบวนการทางสถิติที่ใช้ในการปรับเทียบคะแนนระหว่างแบบสอบ เพื่อให้คะแนนที่แปลงได้จากแบบสอบทั้งสองฉบับสามารถใช้แทนกันได้ (interchangeably) โดยแบบสอบที่ใช้ในการปรับเทียบคะแนนนั้น ต้องมีความยากง่ายเท่าเทียมกัน และมุ่งวัดในเนื้อหาเดียวกัน นอกจากนี้ ศิริชัย กาญจนวาสี (2555) ยังได้ให้นิยามว่า การปรับเทียบคะแนนระหว่างแบบสอบ คือ การแปลงระบบคะแนนของแบบสอบต่างฉบับที่วัดในเนื้อหาเดียวกัน โดยแปลงคะแนนของแบบสอบฉบับหนึ่งให้อยู่บนมาตรหรือสเกลของแบบสอบอีกฉบับหนึ่ง ที่ถือเป็นคะแนนที่สมมูลกัน (equivalent score) เพื่อให้คะแนนของแบบสอบต่างฉบับสามารถเปรียบเทียบกันได้โดยตรง

จากแนวคิดและความหมายของดังกล่าวข้างต้น สรุปได้ว่าการปรับเทียบคะแนน คือ กระบวนการทางสถิติที่ใช้ในการปรับเทียบคะแนนระหว่างแบบสอบที่มุ่งวัดในเนื้อหาเดียวกัน โดยแปลงคะแนนจากแบบสอบหนึ่งให้อยู่ในหน่วยหรือสเกลของแบบสอบอีกฉบับหนึ่ง เพื่อให้คะแนนที่แปลงได้จากแบบสอบทั้งสองฉบับสามารถใช้เปรียบเทียบและใช้แทนกันได้

1.2 เงื่อนไขของการปรับเทียบคะแนนระหว่างแบบสอบ

การปรับเทียบคะแนนระหว่างแบบสอบ สามารถกระทำได้ภายใต้เงื่อนบางประการ ซึ่งมีผู้นำเสนอเงื่อนไขต่างๆดังนี้

Gulliksen (1950 อ้างอิงใน ศิริชัย กาญจนวาสี, 2555) กล่าวว่า แบบสอบที่นำมาปรับเทียบคะแนนกันได้ ต้องเป็นแบบสอบคู่ขนานที่สร้างโดยการจับคู่ข้อสอบที่เหมือนกัน แล้วสุ่มแยกออกเป็นแบบสอบคู่ขนานสองฉบับ

Flanagan (1951 cited in Angoff, 1971) กล่าวว่า คะแนนสอบสองชุด โดยคะแนนชุดหนึ่งมาจากแบบสอบ X ส่วนคะแนนอีกชุดมาจากแบบสอบ Y อาจมีความเท่าเทียมกัน ถ้าตำแหน่งเปอร์เซนไทล์ (Percentile ranks) ของคะแนนสอบสองชุดเท่ากัน โดยแบบสอบ X และ Y ต้องวัดในเนื้อหาเดียวกัน และมีระดับความเที่ยงเท่ากัน ในการปรับเทียบคะแนนต้องพิจาณาก่อนว่าการแจกแจงของคะแนนที่สังเกตได้ในประชากรเหมือนกันและมีความเที่ยงเท่ากัน จึงสามารถนำคะแนนที่สังเกตได้มาปรับเทียบกัน ในทางปฏิบัติ Flanagan เสนอว่าให้ตรวจสอบคะแนนเฉลี่ยจากแบบสอบที่ได้จากลุ่มตัวอย่างขนาดใหญ่ ถ้าคะแนนเฉลี่ยแตกต่างอย่างมีนัยสำคัญทางสถิติ ควรสร้างหรือปรับปรุงแบบสอบทั้งสองให้มีความเป็นคู่ขนาน แล้วเลือกวิธีการปรับเทียบคะแนนที่เหมาะสม ซึ่งเสนอไว้ 4 วิธี ดังนี้ (Flanagan, 1951 อ้างอิงใน พัชรี จันทร์เพ็ง, 2547)

1. ใช้ค่าเฉลี่ย โดยคำนวณค่าเฉลี่ยของการแจกแจงคะแนนทั้งสองชุด ถ้าความแตกต่างของค่าเฉลี่ยอยู่ภายใต้ขอบเขตของการแปรปรวนเชิงสุ่ม ถือว่าคะแนนสอบสองชุดสามารถปรับเทียบกันได้ แต่ถ้าค่าเฉลี่ยดังกล่าวมีความแตกต่างอย่างมีนัยสำคัญทางสถิติ ต้องใช้วิธีบวกเข้าหรือลบออกเท่ากับจำนวนที่แตกต่างจากคะแนนชุดที่หนึ่ง เพื่อให้เกิดคะแนนสมมูลกับอีกชุดหนึ่ง

2. ใช้เทคนิคสมการถดถอย โดยหาค่าประมาณที่ดีที่สุดของคะแนนจากแบบสอบชุดหนึ่งโดยใช้ค่าของแบบสอบอีกชุดหนึ่ง

3. ใช้คะแนนมาตรฐานที่เท่ากันเป็นตัวปรับเทียบคะแนน

4. ใช้วิธีอีควิปเปอร์เซนไทล์ โดยคะแนนเทียบเท่ากันได้ เมื่อคะแนนจากแบบสอบต่างฉบับมีตำแหน่งเปอร์เซนไทล์เท่ากัน

Levine (1955 อ้างอิงใน ศิริชัย กาญจนวาสี, 2555) กล่าวว่า แบบสอบที่นำมาปรับเทียบคะแนนกันได้ ต้องอยู่ในมาตรเดียวกัน และต้องเป็นแบบสอบคู่ขนานกัน (parallel forms) ในด้าน 1) เนื้อเรื่อง (subject) 2) โครงสร้าง (structure) 3) รูปแบบ (format) 4) ชนิดของข้อสอบ (item type) และ 5) เวลาที่ใช้สอบ ซึ่งสอดคล้องกับ ศิริชัย กาญจนวาสี (2555) ที่กล่าวว่า แบบสอบที่นำมาปรับเทียบคะแนนกันได้ ต้องมีคุณสมบัติดังนี้ 1) วัดคุณลักษณะเด่นเดียวเหมือนกัน (unidimensionality) 2) มีลักษณะของความเป็นคู่ขนานในด้านเนื้อหา โครงสร้าง รูปแบบ ชนิดของข้อสอบ เวลาที่ใช้สอบ และ 3) มีความเที่ยงสูงทัดเทียมกัน

Lord (1977 cited in Hambleton & Swaminathan, 1985) กล่าวว่า การปรับเทียบคะแนนระหว่างแบบสอบ ต้องไม่ทำให้เกิดความแตกต่างระหว่างผู้สอบ ไม่ว่าผู้สอบจะสอบด้วยแบบสอบฉบับ X หรือฉบับ Y หลังจากนั้น Lord ก็ได้พัฒนาหลักความเสมอภาค (equity) ในการปรับเทียบคะแนนระหว่างแบบสอบ ซึ่งเสนอว่า การปรับเทียบคะแนนระหว่างแบบสอบ X และ Y มีความเสมอภาคกัน เมื่อผู้สอบที่มีความสามารถระดับต่าง ๆ ทำแบบสอบฉบับ X หรือฉบับ Y แล้วไม่ทำให้เกิดผลที่แตกต่างกันระหว่างผู้สอบ ข้อตกลงของความเสมอภาคที่ Lord กำหนดมีดังนี้ 1) แบบสอบแต่ละฉบับต้องวัดในคุณลักษณะเดียวกัน 2) คะแนนที่สังเกตได้จากแบบสอบทั้งสองฉบับต้องมีความเที่ยงเท่ากัน 3) แบบสอบแต่ละฉบับมีความยากเท่าเทียมกัน (เพราะความสัมพันธ์ระหว่างคะแนนจริง (true score) ของแบบสอบทั้งสองฉบับไม่เป็นเส้นตรง (nonlinear) และทำให้แบบสอบมีความเที่ยงไม่เท่ากัน ณ ระดับความสามารถที่ต่างกัน) 4) การแจกแจงความถี่แบบมีเงื่อนไข (The Conditional Frequency Distribution) ณ ระดับความสามารถ $\theta, f[\langle x|\theta\rangle]$ ของคะแนน θ จากแบบสอบฉบับ X ต้องเป็นการแจกความถี่แบบมีเงื่อนไขเช่นเดียวกับการแปลงคะแนน x(y), $f[\langle x(y)|\theta\rangle]$ เมื่อ x(y) เป็นฟังก์ชันแบบหนึ่งต่อหนึ่งของ y นั่นคือ การแจกแจงของคะแนนก่อนแปลงและหลังแปลงเหมือนกัน 5) แบบสอบทั้งสองฉบับต้องมีความเป็นคู่ขนาน (เป็นนัยว่าการแจกแจงของคะแนนจากแบบสอบทั้งฉบับมีความเท่าเทียมกัน) 6) แบบสอบสองฉบับที่มีความเที่ยงสูงมาก หรือมีความเที่ยงแบบสมบูรณ์ สามารถเป็นแบบสอบที่เท่าเทียมกันได้ นอกจากนี้ Angoff และคณะ (1971) ได้เสนอคุณสมบัติของแบบสอบที่นำมาปรับเทียบคะแนนกันได้อย่างสอดคล้องกัน ดังนี้ 1) คะแนนที่ปรับเทียบกันแล้วไม่แปรผันไปตามกลุ่มผู้สอบ 2) ต้องปรับเทียบคะแนนแบบสมมาตร (symmetry) ซึ่ง ศิริชัย กาญจนวาสี (2555) ได้อธิบายเพิ่มเติมว่า ผลการปรับเทียบคะแนนต้องเหมือนกัน ไม่ว่าจะปรับเทียบจากแบบสอบฉบับ X ไปยังฉบับ Y หรือปรับเทียบคะแนนจากแบบสอบฉบับ Y ไปยังฉบับ X จากข้อตกลงที่กล่าวไว้ข้างต้น Hambleton และ Swaminathan (1985) จึงได้สรุปเป็นข้อกำหนดของการปรับเทียบ ดังนี้ 1) ความเท่าเทียม (equity) 2) ความไม่แปรเปลี่ยนไปตามกลุ่ม (invariance across group) 3) ความสมมาตร (symmetry) 4) ความเป็นมิติเดียวกันของแบบสอบ (unidimentional of the test)

Kolen และ Brennan (2004) ยังกล่าวว่า แบบสอบที่นำมาปรับเทียบคะแนนกันได้ ต้องเป็นแบบสอบที่มุ่งวัดในเนื้อหาเดียวกัน มีความยากง่ายมีคุณลักษณะพิเศษทางสถิติเท่าเทียมกัน

1.3 คำศัพท์ที่เกี่ยวข้องกับการปรับเทียบคะแนนระหว่างแบบสอบ

คำศัพท์ที่เกี่ยวข้องกับการปรับเทียบคะแนนระหว่างแบบสอบ ประกอบด้วยด้วย Equating, Calibration, Projection และModeration โดยคำศัพท์ทั้งสี่คำนี้คือรูปแบบหนึ่งของการเชื่อมต่อ (Linking) ที่มีความแตกต่างกันในข้อตกลงเบื้องต้นที่กำหนด และลดต่ำลงตามลำดับ

Kolen และ Brennan (2004) กล่าวว่า Equating เป็นการเชื่อมต่อรูปแบบแรกซึ่งเกี่ยวข้องกับความสัมพันธ์ระหว่างคะแนนของแบบสอบสองฉบับที่มุ่งวัดเนื้อหาเดียวกันโดยข้อกำหนดของการปรับเทียบ (equating) มีความเข้มงวดมากที่สุด (Misley,1992; Linn,1993 อ้างอิงใน พัชรี จันทร์เพ็ง, 2547)

Calibration เป็นการเชื่อมต่อรูปแบบที่สอง ซึ่งเกี่ยวข้องกับความสัมพันธ์ระหว่างแบบสอบสองฉบับที่มีรายละเอียดของเนื้อหาเหมือนกันแต่มีรายละเอียดเชิงสถิติที่แตกต่างกัน กล่าวคือ แบบสอบสองฉบับมีเนื้อหาเหมือนกัน แต่มีค่าความเที่ยงหรือค่าความยากต่างกัน

Projection มีลักษณะที่แตกต่างจาก Equating และCalibration คือ 1) Projection ซึ่งเป็นไปในทิศทางเดียวกัน (unidirectional) 2) ต้องการรูปแบบผู้สอบกลุ่มเดียว 3) ไม่จำเป็นต้องมีเนื้อหาที่วัดเหมือนกัน นอกจากนี้ความสัมพันธ์ของ Projection ยังเกี่ยวข้องกับสมการถดถอย (แบบเส้นตรง หรือไม่ใช่เส้นตรง) ที่มีความสัมพันธ์แบบไม่สมมาตร ดังนั้น Projection จึงเป็นการเชื่อมโยงคะแนนจากแบบสอบที่วัดโครงสร้างต่างกัน โดยในการเชื่อมโยงคะแนนได้ใช้วิธีการทำนาย (Feuer et. Al., 1999; Mislevy, 1992; Lin, 1993 อ้างถึงใน อัญชลี ศรีกลชาญ, 2552)

Moderation หรือบางครั้งเรียกว่า "distribution matching" ซึ่งใช้ข้อมูลจากรูปแบบผู้สอบกลุ่มเดียว (อย่างเช่น ผู้สอบกลุ่มเดียว ที่ทำแบบสอบทั้งสองฉบับ) รูปแบบผู้สอบที่มาจากการสุ่ม และรูปแบบผู้สอบต่างกลุ่ม โดยข้อกำหนด Moderation มีความเข้มงวดต่ำที่สุด (Misley,1992; Linn,1993 อ้างอิงใน พัชรี จันทร์เพ็ง, 2547) (พัชรี จันทรเพ็ง, 2547)

ตอนที่ 2 การออกแบบการปรับเทียบคะแนน (equating designs)

Lord (1975a อ้างถึงในศิริชัย กาญจนวาสี, 2555) กล่าวว่า การปรับเทียบคะแนนระหว่างแบบสอบมีความเป็นไปได้ เมื่อมีการออกแบบการเก็บรวบรวมข้อมูลที่สามารถเชื่อมโยงกันได้ (linking designs) ดังต่อไปนี้

1. ผู้สอบกลุ่มหนึ่ง และผู้สอบอีกกลุ่มหนึ่ง ต้องมีสมาชิกร่วมกันจำนวนหนึ่ง
2. ผู้สอบกลุ่มหนึ่ง และผู้สอบอีกกลุ่มหนึ่ง ต่างเป็นตัวแทนของผู้สอบที่มาจากประชากรเดียวกัน
3. แบบสอบฉบับหนึ่ง และแบบสอบอีกฉบับหนึ่ง มีข้อสอบร่วมกันอยู่จำนวนหนึ่ง

การปรับเทียบคะแนนระหว่างแบบสอบทำได้ก็ต่อเมื่อมีการออกแบบการปรับเทียบคะแนนที่เหมาะสม โดยการทดสอบต้องมีผู้สอบร่วม (common samples) อยู่ในผู้สอบแต่ละกลุ่ม หรือใช้กลุ่มผู้สอบแต่ละกลุ่มที่เท่าเทียมกัน (equivalent or matching groups) หรือใช้ข้อสอบร่วม (common items) อยู่ในแบบสอบแต่ละฉบับ

ดังนั้น การปรับเทียบคะแนน จึงต้องอาศัย 2 กระบวนการที่สำคัญ ได้แก่

1) การออกแบบวิธีการเก็บรวบรวมข้อมูล (designs for data collection)

2) การออกแบบวิธีการปรับเทียบคะแนน (designs for equating method)

2.1 การออกแบบวิธีการเก็บรวบรวมข้อมูล (designs for data collection)

ในการปรับเทียบคะแนนระหว่างแบบสอบต้องมีการออกแบบการเก็บรวบรวมข้อมูล ซึ่งต้องออกแบบให้มีข้อสอบร่วมอยู่ในแบบสอบแต่ละฉบับ หรือมีผู้สอบร่วมอยู่ในผู้สอบแต่ละกลุ่ม หรือใช้กลุ่มผู้สอบแต่ละกลุ่มที่เป็นกลุ่มเท่าเทียมกัน โดย Lord (1975b, Angoff 1984; Petersen et.al, 1989 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555) ได้จำแนกรูปแบบการเก็บรวบรวมข้อมูลออกเป็น 4 แบบ ดังนี้

2.1.1 รูปแบบผู้สอบกลุ่มเดียว (single-group design) แบ่งออกได้ 2 แบบย่อย ได้แก่

2.1.1.1 รูปแบบผู้สอบกลุ่มเดียวที่ไม่ได้รับการจัดให้สมดุล (uncounterbalanced design) รูปแบบนี้ใช้ผู้สอบเพียงกลุ่มเดียวเป็นผู้สอบร่วม ในการเก็บรวบรวมข้อมูลใช้กลุ่มผู้สอบกลุ่มเดียว ซึ่งผู้สอบทำแบบสอบทั้ง 2 ฉบับ โดยผู้สอบทำแบบสอบฉบับหนึ่งแล้วตามด้วยแบบสอบอีกฉบับ และเพื่อหลีกเลี่ยงปัญหาแทรกซ้อน ในการให้ผู้สอบกลุ่มเดียวทำแบบสอบทั้งสองฉบับนั้น ระดับความสามารถของผู้สอบที่ทำแบบสอบทั้งฉบับต้องเท่ากัน และต้องไม่มีปัญหาในเรื่องความยากของแบบสอบต่างฉบับ ดังนั้นคะแนนจากแบบสอบจึงสามารถปรับเทียบกันได้บนพื้นฐานความสามารถที่เท่ากันจากคนกลุ่มเดียวกัน

ข้อจำกัดของรูปแบบนี้ คือ อิทธิพลจากลำดับของการทดสอบก่อน-หลัง กล่าวคือ การสอบฉบับหลัง ผู้สอบอาจได้รับปัจจัยที่ตกทอดมาจากการสอบฉบับแรก เช่น การเรียนรู้ การฝึกฝน ความเมื่อยล้าจากการทำแบบสอบฉบับแรก ซึ่งอาจมีอิทธิผลต่อการปรับเทียบคะแนน โดยสอดคล้องกับตัวอย่างที่ Kolen และ Brennan (2004) นำเสนอ คือ ถ้าผู้สอบกลุ่มเดียวทำแบบสอบฉบับ X แล้วตามด้วยแบบฉบับ Y ผลการสอบอาจพบว่าแบบสอบฉบับ Y ยากกว่าแบบสอบฉบับ X ที่เป็นเช่นนี้เพราะ ผู้สอบเกิดความเหนื่อยล้าจากการทำแบบสอบฉบับแรก จึงทำให้ความสามารถในการทำแบบสอบฉบับหลังลดลง จากข้อจำกัดนี้จึงนำมาปรับปรุงเป็นรูปแบบผู้สอบกลุ่มเดียวที่ได้รับการจัดให้สมดุล (counterbalanced design)

2.1.1.2 รูปแบบผู้สอบกลุ่มเดียวที่ได้รับการจัดให้สมดุล (counterbalanced design) รูปแบบนี้ปรับจากรูปแบบแรก เพื่อขจัดผลของลำดับการทดสอบก่อน-หลังในการเก็บรวบรวมข้อมูลกระทำได้โดยการสุ่มผู้สอบแยกเป็น 2 กลุ่มย่อย แต่ละกลุ่มย่อยได้รับการทดสอบด้วยแบบสอบทั้งสองฉบับ โดยกลุ่มย่อยที่หนึ่งทำแบบสอบฉบับ X แล้วตามด้วยแบบสอบฉบับ Y ส่วนอีกกลุ่มย่อยหนึ่งทำแบบสอบฉบับ Y แล้วตามด้วยแบบสอบฉบับ X

ข้อจำกัดของรูปแบบนี้ คือ อิทธิพลจากการลำดับการทดสอบ การเรียนรู้ การฝึกฝน และความเมื่อยล้าที่สมดุลกันในทั้งสองกลุ่มย่อย

2.1.2 รูปแบบผู้สอบกลุ่มเท่าเทียมกัน (equivalent-group design) รูปแบบนี้ใช้กลุ่มผู้สอบเท่าเทียมกัน (equivalent group) โดยต้องจัดกลุ่มผู้สอบให้มีความคล้ายคลึงกันมากที่สุดเท่าที่ทำได้ แม้ว่ากลุ่มผู้สอบไม่เหมือนกันทุกประการ (identical groups) แต่ก็ควรมีลักษณะทั่วไปคล้ายกันมากที่สุด กลุ่มดังกล่าวอาจได้มาโดยการสุ่ม (random)

ข้อดีของรูปแบบนี้ คือ สามารถหลีกเลี่ยงปัญหาเรื่องการเรียนรู้ การฝึกฝน และความเมื่อยล้าจากการทำแบบสอบฉบับแรก เพราะกลุ่มผู้สอบแต่กลุ่มทำแบบสอบฉบับเดียว แต่ปัญหาที่ตามมา คือ กลุ่มผู้สอบที่ใช้ไม่เหมือนกันอาจมีการแจกแจงความสามารถที่แตกต่างกัน และไม่มีข้อมูลในการนำมาปรับความแตกต่างของกลุ่ม ความแตกต่างที่เกิดขึ้นนี้ย่อมส่งผลกระทบต่อความลำเอียงในการปรับเทียบคะแนนระหว่างแบบสอบ ดังนั้นวิธีการที่ช่วยลดความแตกต่าง คือ การใช้กลุ่มตัวอย่างขนาดใหญ่

2.1.3 รูปแบบผู้สอบต่างกลุ่มโดยใช้แบบสอบร่วม (anchor-test design)

2.1.3.1 รูปแบบผู้สอบกลุ่มสุ่มโดยใช้แบบสอบร่วม (anchor-test design random group design) รูปแบบนี้เป็นการใช้กลุ่มผู้สอบที่ได้จากการสุ่ม แล้วให้ผู้สอบแต่ละกลุ่มทำแบบสอบฉบับเดียว โดยแบบสอบแต่ละฉบับมีข้อสอบร่วมกันจำนวนหนึ่ง (common or anchor items) ซึ่งรูปแบบนี้ช่วยลดความลำเอียงในการปรับเทียบคะแนน ที่มีสาเหตุมากจากการเรียนรู้ การฝึกฝน ความเมื่อยล้า และความแตกต่างเกี่ยวกับระดับความสามารถของกลุ่มผู้สอบ รูปแบบนี้จึงได้เปรียบกว่าสองรูปแบบแรก การใช้แบบสอบร่วมนี้สามารถกระทำได้ 2 แบบ ดังนี้

- แบบสอบร่วมภายใน (internal anchor-test design) เป็นแบบสอบที่มีข้อสอบร่วมที่จัดรวมไว้ในแบบสอบที่ต้องการนำมาปรับเทียบ

- แบบสอบร่วมภายนอก (external anchor-test design) เป็นแบบสอบที่มีข้อสอบร่วมที่จัดแยกออกมาต่างหากจากแบบสอบที่ต้องการนำมาปรับเทียบ ซึ่งกลุ่มผู้สอบต้องทำและจับเวลาแยกออกจากตัวแบบสอบที่ต้องการนำมาปรับเทียบคะแนน

คุณภาพของการปรับเทียบคะแนนสำหรับรูปแบบนี้ ขึ้นอยู่กับปัจจัยบางประการ เช่น คุณภาพของข้อสอบร่วม ซึ่งต้องคำนึงถึงความสัมพันธ์ระหว่างข้อสอบร่วมที่ต้องมีเนื้อหาและระดับความยากเท่าเทียมกับแบบสอบที่ต้องการปรับเทียบ และความยาวของแบบสอบร่วมเป็นสำคัญ

2.1.3.2 รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกันโดยใช้แบบสอบร่วม (anchor-test design nonequivalent groups design) รูปแบบนี้ใช้กลุ่มผู้สอบที่ไม่เท่าเทียมกันแล้วให้ผู้สอบแต่ละกลุ่มทำแบบสอบฉบับเดียว โดยแบบสอบแต่ละฉบับมีข้อสอบร่วมกันจำนวนหนึ่ง (common or anchor items) ซึ่งเรียกว่าแบบสอบร่วม (anchor-test) อาจใช้แบบสอบร่วมภายในหรือแบบสอบร่วมภายนอกก็ได้ ในการทดสอบบางสถานการณ์อาจจำเป็นต้องใช้แบบสอบร่วมกับกลุ่มผู้สอบต่างประชากรกัน เช่น กลุ่มผู้สอบต่างโปรแกรม ต่าง เวลา หรือต่างระดับ เป็นต้น ซึ่งคะแนนจากแบบสอบร่วมใช้เป็นกลไกสำคัญในการปรับเทียบคะแนนระหว่างแบบสอบ

Kolen และ Brennan (2004) ได้อธิบายถึงลักษณะของแบบสอบร่วม/ชุดของข้อสอบร่วมไว้ดังนี้ เมื่อใช้รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกันโดยใช้ข้อสอบร่วม (common-item nonequivalent group design) ชุดของข้อสอบร่วมควรสร้างให้มีรายละเอียดคล้ายคลึงกับแบบสอบ ทั้งฉบับ โดยข้อสอบร่วมควรสร้างให้ยาวพอที่เป็นตัวแทนเนื้อหาในแบบสอบ ซึ่งจำนวนข้อสอบร่วมก็เป็นส่วนสำคัญที่ต้องได้รับการพิจารณา โดยจากวิจัยที่ผ่านมา พบว่า การใช้ข้อสอบสอบร่วมจำนวนมาก ทำให้ความคลาดเคลื่อนในการปรับเทียบคะแนนลดลง (Budescu, 1985; Wingersky et al., 1987 cited in Kolen & Brenan, 2004) เช่นเดียวกับ Petersen et al. (1983 cited in Kolen & Brenan, 2004) ที่กล่าวว่าข้อสอบร่วมจำนวนต่ำทำให้เกิดปัญหาในการปรับเทียบคะแนน แต่อย่างไรก็ตามจากการทบทวนวรรณกรรมของ Harris (1993 cited in Kolen & Brenan, 2004) พบว่า ผลจากงานวิจัยหลายงานแนะนำให้ใช้ข้อสอบร่วมจำนวนต่ำ ที่ยึดตามข้อมูลที่เลียนแบบมาจากโมเดลหนึ่งมิติ (unidimentional model) และเนื่องจากการทดสอบทางการศึกษามีลักษณะที่หลากหลาย/วิวิธพันธ์ การใช้ข้อสอบร่วมจำนวนมากจึงจำเป็นต้องใช้เมื่อการปรับเทียบในเชิงปฏิบัติต้องการความพอเพียง (adequate) โดย Angoff และคณะ (1971) ได้แนะนำว่า ควรใช้ข้อสอบร่วม

อย่างต่ำ 20% เมื่อแบบสอบนั้นมีจำนวนข้อสอบทั้งหมด 40 ข้อ และในกรณีที่แบบสอบนั้นมีข้อสอบจำนวนมาก จำนวนข้อสอบร่วมที่เพียงพอ คือ 30% แต่อย่างไรก็ตามในงานวิจัยของ Kim และคณะ (2010) ที่ศึกษาการออกแบบการปรับเทียบสำหรับแบบสอบรูปแบบผสม เมื่อใช้รูปแบบผู้สอบกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม ซึ่งมีการใช้การตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) ในข้อสอบร่วมแบบสอบสร้างคำตอบ โดยใช้จำนวนข้อสอบร่วม (12 ข้อ) เท่ากับจำนวนข้อสอบเฉพาะทั้งฉบับ (12 ข้อ) โดยแบ่งเป็นข้อสอบแบบหลายตัวเลือกจำนวน 8 ข้อ และข้อสอบแบบสร้างคำตอบจำนวน 4 ข้อ ซึ่งจากการทบทวนเอกสารและวิจัยที่เกี่ยวข้องยังไม่มีงานวิจัยใดที่ศึกษาจำนวนข้อสอบร่วมที่เหมาะสม/พอเพียงสำหรับการออกแบบการปรับเทียบที่มีการประยุกต์ใช้ Trend scoing method ด้วยเหตุนี้ผู้วิจัยจึงกำหนดให้จำนวนข้อสอบร่วมภายในที่ใช้ในแบบสอบเท่ากับจำนวนข้อสอบเฉพาะทั้งหมดเช่นเดียวกับงานวิจัยของ Kim และคณะ (2010)

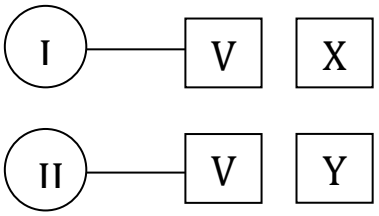
2.1.4 รูปแบบผู้สอบต่างกลุ่มโดยมีผู้สอบร่วม (common-person design) รูปแบบนี้ใช้ผู้สอบต่างกลุ่มกันทำแบบสอบต่างฉบับ แต่มีผู้สอบส่วนหนึ่งจากทั้งสองกลุ่มที่ต้องทำแบบสอบทั้งสองฉบับ ซึ่งเป็นลักษณะเฉพาะของรูปแบบการใช้ผู้สอบร่วม

ตาราง 1 รูปแบบการเก็บรวบรวมข้อมูลสำหรับการปรับเทียบ
(ศิริชัย กาญจนวาสี, 2555)

การออกแบบ (DESIGNS)	กลุ่มผู้สอบ	แบบสอบ	ปัจจัยเกี่ยวข้อง
1. การใช้ผู้สอบกลุ่มเดียว (SINGLE-GROUP DESIGN)	I	X Y	■ ความเมื่อยล้า ■ ลำดับการทำแบบทดสอบ ■ การเรียนรู้จากแบบสอบ
2. การใช้ผู้สอบกลุ่มเท่าเทียมกัน (EQUIVALENT-GROUP DESIGN)	I I′	X Y	■ การสุ่ม ■ ความทัดเทียม
3. การใช้ข้อสอบร่วม (ANCHOR TEST) 3.1 แบบสอบร่วมภายใน (INTERNAL ANCHOR TEST)	I II	X V V Y	■ ความยาวของ V ■ ความสัมพันธ์ระหว่าง V กับ X และ Y

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 1 รูปแบบการเก็บรวบรวมข้อมูลสำหรับการปรับเทียบ (ต่อ) (ศิริชัย กาญจนวาสี, 2555)

การออกแบบ (DESIGNS)	กลุ่มผู้สอบ แบบสอบ	ปัจจัยเกี่ยวข้อง
3. การใช้ข้อสอบร่วม (ANCHOR TEST) 3.2 แบบสอบร่วมภายนอก (EXTERNAL ANCHOR TEST)		▪ ความยาวของ V ความสัมพันธ์ระหว่าง V กับ X และ Y
4. การใช้ผู้สอบร่วม (COMMON-PERSON DESIGN)		▪ ความเมื่อยล้าของกลุ่มผู้สอบ ▪ ลำดับการทำแบบสอบ

2.2 การออกแบบวิธีการปรับเทียบคะแนน (equating methods)

Kolen และ Brennan (2004) ได้นำเสนอวิธีการปรับเทียบคะแนนไว้ดังนี้

2.2.1 วิธีการปรับเทียบค่าเฉลี่ย (mean equating)

วิธีการนี้เป็นการพิจารณาความยากของแบบสอบ X ที่ต่างจากแบบสอบ Y เป็นค่าคงที่ ตั้งแต่ค่าต่ำที่สุดไปจนถึงค่าที่มากที่สุด เช่น แบบสอบ X ง่ายกว่าแบบสอบ Y อยู่ 2 คะแนนในผู้สอบกลุ่มสูง และแบบสอบ X ง่ายกว่าแบบสอบ Y อยู่ 2 คะแนนในผู้สอบกลุ่มต่ำเช่นกัน

กำหนดให้แบบสอบ X เป็นแบบสอบฉบับใหม่ โดย X เป็นคะแนนที่ได้จากแบบสอบ X ทั้งหมด ส่วน x เป็นคะแนนเฉพาะจากแบบสอบ X ที่ต้องการปรับเทียบ และ Y เป็นแบบสอบฉบับเก่า โดย Y เป็นคะแนนที่ได้จากแบบสอบ Y ทั้งหมด ส่วน y เป็นคะแนนเฉพาะจากแบบสอบ Y ที่ต้องการปรับเทียบ ดังนั้น $\mu(X)$ เป็นคะแนนเฉลี่ยของแบบสอบ X และ $\mu(Y)$ เป็นคะแนนเฉลี่ยของแบบสอบ Y ความหมายของการปรับเทียบ คือ คะแนนจากแบบสอบทั้งสองฉบับแตกต่างจากค่าเฉลี่ยของแบบสอบแต่ละฉบับเท่ากัน นั่นคือ

$$x - \mu(X) = y - \mu(y) \qquad (1.1)$$

จะได้ $$m_Y(x) = y = x - \mu(X) + \mu(Y) \qquad (1.2)$$

เมื่อ $m_Y(x)$ คือ คะแนน x จากแบบสอบ X แปลงไปยังสเกลของแบบสอบ Y โดยใช้วิธีการปรับเทียบค่าเฉลี่ย

ตัวอย่างเช่น คะแนนเฉลี่ยจากแบบสอบ X เท่ากับ 72 และค่าเฉลี่ยจากแบบสอบ Y เท่ากับ 77 จากสมการ 1.2 ต้องเพิ่มอีก 5 คะแนนในแบบสอบ X เพื่อแปลงคะแนนจากแบบสอบ X ไปยังแบบสอบ Y ดังนี้

$$m_Y(x) = y = x - 72 + 77 = x + 5$$

คะแนนของแบบสอบ X เท่ากับ 72 เมื่อพิจาณาแล้วอยู่ในระดับความสามารถเท่ากับ77 ในแบบสอบ Y (77=72+5) และคะแนนจากแบบสอบ X เท่ากับ 75 อยู่ในระดับ

ความสามารถเท่ากับ 80 ในแบบสอบ Y (80=75+5) ดังนั้นการปรับเทียบค่าเฉลี่ยจึงเป็นการเพิ่มค่าคงที่ (อาจเป็นค่าติดลบก็ได้) ในคะแนนที่สังเกตได้ของแบบสอบ X เพื่อหาคะแนนปรับเทียบของแบบสอบ Y

2.2.2 วิธีการปรับเทียบเชิงเส้นตรง (liner equating)

วิธีการนี้เป็นการพิจารณาความยากของแบบสอบ X ที่ต่างจากแบบสอบ Y ซึ่งแปรผันไปตามสเกลของคะแนน เช่น ในการปรับเทียบคะแนนของแบบสอบ X ที่ยากกว่าแบบสอบ Y ในกลุ่มผู้สอบกลุ่มต่ำ แต่มีความยากต่ำกว่าในผู้สอบกลุ่มสูง

ในการปรับเทียบคะแนนเชิงเส้นตรง คะแนนจากแบบสองฉบับเท่าเทียมกัน เมื่อคะแนนของแบบสอบแต่ละฉบับแตกต่างจากคะแนนเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของแต่ละฉบับเท่ากัน กำหนดให้ $\sigma(X)$ และ $\sigma(Y)$ คือส่วนเบี่ยงเบนมาตรฐานของคะแนนจากแบบสอบ X และแบบสอบ Y ตามลำดับ ดังนั้นในการแปลงคะแนนเชิงเส้นตรง (linear conversion) คะแนนของแบบสอบสมมูลกัน เมื่อคะแนนมาตรฐาน (standard score) ของแบบสอบแต่ละฉบับเท่ากัน (Kolen & Brennan, 2004)

$$\frac{x-\mu(X)}{\sigma(X)} = \frac{y-\mu(Y)}{\sigma(Y)} \qquad (1.3)$$

ถ้าส่วนเบี่ยงเบนมาตรฐานของแบบสอบทั้งสองฉบับเท่ากัน สามารถใช้สมการปรับเทียบค่าเฉลี่ย ดังสมการ (1.3) และได้ว่า

$$l_Y(x) = y = \sigma(Y)\left[\frac{x-\mu(X)}{\sigma(X)}\right] + \mu(Y) \qquad (1.4)$$

เมื่อ $l_Y(x)$ คือ คะแนนที่สังเกตได้จากแบบสอบ X แปลงไปยังสเกลของแบบสอบ Y ซึ่งสามารถจัดรูปสมการใหม่ได้ ดังนี้

$$l_Y(x) = y = \frac{\sigma(Y)}{\sigma(X)}x + \left[\mu(Y) - \frac{\sigma(Y)}{\sigma(X)}\mu(X)\right] \qquad (1.5)$$

สมการ (1.5) คือสมการเชิงเส้นตรง โดย $slope = \frac{\sigma(Y)}{\sigma(X)}$

และ จุดตัด $= \mu(Y) - \frac{\sigma(Y)}{\sigma(X)}\mu(X)$ (1.6)

อย่างเช่น ถ้าส่วนเบี่ยงเบนมาตรฐานของแบบสอบทั้งสองแบบ ได้แก่ $\sigma(X) = 10$ และ $\sigma(Y) = 9$ ได้ความชันเป็น 9/10 = .9 และจุดตัดแกน X เท่ากับ 77-(9/10)72 = 12.2 ดังนั้นสมการของการแปลงคะแนนจะได้ว่า $l_Y(x) = .9(x) + 12.2$ แล้ว $l_Y(x)$ ที่ $x = 75$ มีค่าเท่าใด

$$l_Y(75) = .9(75) + 12.2 = 79.7$$

และถ้า ที่ $x = 77$หรือที่ $x = 85$ จะได้ว่า

$$l_Y(77) = .9(77) + 12.2 = 81.5$$
$$l_Y(85) = .9(85) + 12.2 = 88.7$$

ตัวอย่างดังกล่าวแสดงให้เห็นถึงความแตกต่างระหว่างความยากของแบบสอบทั้งสองฉบับที่มีอยู่หลายค่าตามระดับของคะแนน สำหรับตัวอย่างนี้ความแตกต่างของความยากระหว่างแบบสอบ X

และแบบสอบ Y เมื่อคะแนนของแบบสอบ X เท่ากับ 75 ได้คะแนนเป็น 4.7 (79.7-75) ขณะที่ความแตกต่างเมื่อคะแนนของแบบสอบ X เท่ากับ 85 ได้คะแนนเป็น 3.7 (88.7-85)

พารามิเตอร์ของประชากรที่สังเคราะห์ 4 พารามิเตอร์สามารถแสดงอยู่ในเทอมของพารามิเตอร์สำหรับประชากร 1 และ2 ดังนี้

$$\mu_s(X) = w_1\mu_1(X) + w_2\mu_2(X) \quad (2.2)$$

$$\mu_s(Y) = w_1\mu_1(Y) + w_2\mu_2(Y) \quad (2.3)$$

$$\sigma_s^2(X) = w_1\sigma_1^2(X) + w_2\sigma_2^2(X) + w_1w_2[\mu_1(X) - \mu_2(X)]^2 \quad (2.4)$$

$$\sigma_s^2(Y) = w_1\sigma_1^2(Y) + w_2\sigma_2^2(Y) + w_1w_2[\mu_1(Y) - \mu_2(Y)]^2 \quad (2.5)$$

โดยเครื่องหมาย 1 และ 2 ข้างใต้ แสดงถึงประชากร 1 และ 2 ตามลำดับ

สำหรับรูปแบบผู้สอบต่างกลุ่มโดยใช้ข้อสอบร่วม (common-item nonequivalent group design) นั้น X ไม่สามารถดำเนินการไปยังผู้สอบในประชากร 2 และ Y ไม่สามารถดำเนินการไปยังผู้สอบในประชากร1 ได้ ดังนั้น $\mu_2(X)$, $\sigma_2^2(X)$, $\mu_1(X)$ และ$\sigma_1^2(Y)$ ในสมการ (2.2)-(2.5) จึงไม่สามารถประมาณได้โดยตรง (estimated directly) ดังนั้นจึงต้องใช้วิธีของ Tucker และวิธีของ Levine ในประมาณค่าพารามิเตอร์ในสมการ (2.2)-(2.5) ที่ไม่สามารถประมาณได้โดยตรง

2.2.2.1 วิธีของ Tucker (Tucker method)

วิธีของ Tucker สร้างข้อตกลงเบื้องต้น 2 แบบ เพื่อประมาณค่าพารามิเตอร์ในสมการ (2.2)-(2.5) ซึ่งไม่สามารถประมาณค่าพารามิเตอร์ได้โดยตรง ข้อตกลงเบื้องต้นประเภทแรกเกี่ยวข้องกับสมการถดถอยของคะแนนรวมของข้อสอบร่วมทั้งหมด ข้อตกเบื้องต้นประเภทที่สองเกี่ยวข้องกับความแปรผันที่เป็นเงื่อนไขของคะแนนทั้งหมด

<u>ข้อตกลงเบื้องต้นสมการถดถอยเชิงเส้นตรง</u> (linear regression assumption)

สมการการถดถอยของแบบสอบ X และแบบสอบร่วมต้องเป็นฟังก์ชั่นเส้นตรง (linear function) เดียวกัน ทั้งประชากรกลุ่ม 1 และ 2 ข้อตกลงเบื้องต้นถูกสร้างสำหรับแบบสอบ Y โดยใช้ข้อสอบร่วม เมื่อ α คือ ความชันของการถดถอย และ β คือ จุดตัดแกน X ของการถดถอย (Kolen & Brennan, 2004)

$$\alpha_1(X|V) = \sigma_1(X,V)/\sigma_1^2(V) \quad (2.6)$$

และ

$$\beta_1(X|V) = \mu_1(X) - \alpha_1(X|V)\mu_1(V) \quad (2.7)$$

สมการ (2.6) และ (2.7) คือ ความชันและจุดตัดแกน X ตามลำดับของสมการการถดถอยของ X โดยใช้ข้อสอบร่วมในประชากร 1 ซึ่งค่าความชันและจุดตัดแกน X สามารถสังเกตได้โดยตรง ส่วนความชันและจุดตัดแกน X ของประชากร 2 คือ

$$\alpha_2(X|V) = \sigma_2(X,V)/\sigma_2^2(V) \quad (2.8)$$

และ

$$\beta_2(X|V) = \mu_2(X) - \alpha_2(X|V)\mu_2(V) \quad (2.9)$$

ค่าทั้งสองค่านี้ ไม่สามารถวัดได้โดยตรง สำหรับ X และ V ข้อตกลงเบื้องต้นของสมการถดถอย คือ

$$\alpha_2(X|V) = \alpha_1(X|V) \quad (2.10)$$

และ

$$\beta_2(X|V) = \beta_1(X|V) \quad (2.11)$$

ขณะที่ค่าทางด้านซ้ายของสมการไม่สามารถวัดได้โดยตรง ข้อตกลงเบื้องต้นของการถดถอย สำหรับ Y และ V คือ

$$\alpha_1(Y|V) = \alpha_2(Y|V) \quad (2.12)$$

และ

$$\beta_1(Y|V) = \beta_2(Y|V) \quad (2.13)$$

<u>ข้อตกลงเบื้องต้นความแปรปรวนอย่างมีเงื่อนไข</u> (conditional variance assumption)

สำหรับวิธีของ Tucker ความแปรปรวนอย่างมีเงื่อนไขของ X ให้ข้อตกลงเบื้องต้นของ V ทั้งในประชากร 1 และ 2 ข้อตกลงเหล่านั้น คือ

$$\sigma_2^2(X)[1-\rho_2^2(X,V)] = \sigma_1^2(X)[1-\rho_1^2(X,V)]$$

และ

$$\sigma_1^2(Y)[1-\rho_1^2(Y,V)] = \sigma_2^2(Y)[1-\rho_2^2(Y,V)] \quad (2.14)$$

โดยที่ ρ คือ ค่าสหสัมพันธ์ ซึ่งค่าดังกล่าวไม่สามารถสังเกตได้โดยตรง

<u>ผลลัพธ์ที่ได้</u>

ค่าเฉลี่ยและความแปรปรวนที่เกิดจากการสังเคราะห์ประชากรในสมการ (2.2)-(2.5) คือ

$$\mu_s(X) = \mu_1(X) - w_2\gamma_1[\mu_1(V) - \mu_2(V)] \quad (2.15)$$

$$\mu_s(Y) = \mu_2(Y) - w_1\gamma_2[\mu_1(V) - \mu_2(V)] \quad (2.16)$$

$$\sigma_s^2(X) = \sigma_1^2(X) - w_2\gamma_1^2[\sigma_1^2(V) - \sigma_2^2(V)] + w_1w_2\gamma_1^2[\mu_1(V) - \mu_2(V)]^2 \quad (2.17)$$

$$\sigma_s^2(Y) = \sigma_2^2(Y) - w_1\gamma_2^2[\sigma_1^2(V) - \sigma_2^2(V)] + w_1w_2\gamma_2^2[\mu_1(V) - \mu_2(V)]^2 \quad (2.18)$$

ถ้าเทอม γ คือ ความชันของสมการการถดถอย ได้ว่า

$$\gamma_1 = \alpha_1(X|V) = \sigma_1(X,V)/\sigma_1^2(V) \quad (2.19)$$

$$\gamma_2 = \alpha_2(X|V) = \sigma_2(X,V)/\sigma_2^2(V) \quad (2.20)$$

2.2.2.2 Levine Observed Score Method

วิธีการนี้เป็นวิธีในการปรับเทียบคะแนนที่สังเกตได้ โดยใช้สมการ 1 เพื่อเชื่อมคะแนนที่สังเกต X ไปยังสเกลของคะแนนที่สังเกตได้ Y แต่อย่างไรก็ตาม ข้อตกลงเบื้องต้นสำหรับวิธีนี้เกี่ยวข้องกับคะแนนจริง T_X, T_Y และT_V ซึ่งข้อตกลงเบื้องต้นเหล่านี้เกี่ยวข้องกับคะแนนที่สังเกตได้ตามทฤษฎีการทดสอบแบบดั้งเดิม (Kolen & Brennan, 2004) คือ

$$X = T_X + E_X \quad (2.21)$$

$$Y = T_Y + E_Y \quad (2.22)$$

$$V = T_V + E_V \quad (2.23)$$

โดย E_X, E_Y และE_Z คือ ความคลาดเคลื่อน ซึ่งคาดหวังว่ามีค่าเป็นศูนย์ และไม่มีความสัมพันธ์กับคะแนนจริง

ข้อตกลงเบื้องต้นสหพันธ์ (correlational assumption)

วิธีของ Levine มีข้อตกลงเบื้องต้นว่า X, Y และ V วัดในสิ่งเดียวกัน ดังนั้น T_X และT_V จึงมีความสัมพันธ์กันอย่างสมบูรณ์ทั้งในประชากร 1 และ 2 นอกจากนี้ T_Y และ T_V ยังมีความสัมพันธ์กันอย่างสมบูรณ์ทั้งในประชากร 1 และ 2 จะได้ว่า

$$\rho_1(T_X, T_V) = \rho_2(T_X, T_V) = 1 \quad (2.24)$$

$$\rho_1(T_Y, T_V) = \rho_2(T_Y, T_V) = 1 \quad (2.25)$$

โดย T_X และ T_Y ในสมการ (2.24) และ (2.25) เป็นฟังก์ชั่นที่สัมพันธ์กับประชากรทั้งสองกลุ่ม

ข้อตกลงเบื้องต้นสมการถดถอยเชิงเส้นตรง (linear regression assumption)

สมการการถดถอยของ T_X ด้วย T_V มีข้อตกลงเบื้องต้นว่าเป็นฟังก์ชั่นเชิงเส้นตรงเหมือนกันทั้งในประชากร 1 และ 2 เช่นเดียวกับข้อตกลงเบื้องต้นของสมการการถดถอยของ T_Y ด้วย T_V โดยความชันของ T_X ด้วย T_V คือ $\alpha_1(T_X|T_V) = \rho_1(T_X, T_V)\sigma_1(T_X)/\sigma_1(T_V)$ เพราะว่า $\rho_1(T_X, T_V) = 1$ ซึ่งมาจากข้อตกลงเบื้องต้นเกี่ยวกับความสัมพันธ์ในสมการ (2.24) และ $\alpha_1(T_X|T_V) = \sigma_1(T_X)/\sigma_1(T_V)$

ในทำนองเดียวกัน $\alpha_2(T_X|T_V) = \sigma_2(T_X)/\sigma_2(T_V)$ ดังนั้นข้อตกลงเบื้องต้นของความชันของสมการการถดถอยของคะแนนจริงที่เท่าเทียมกันสำหรับ T_X ด้วย T_V ในประชากร 1 และ2 คือ

$$\frac{\sigma_2(T_X)}{\sigma_2(T_V)} = \frac{\sigma_1(T_X)}{\sigma_1(T_V)} \quad (2.26)$$

ในทำนองเดียวกัน

$$\frac{\sigma_1(T_Y)}{\sigma_1(T_V)} = \frac{\sigma_2(T_Y)}{\sigma_2(T_V)} \quad (2.27)$$

จากสมการของโมเดลทฤษฎีการตอบสนองข้อสอบแบบดั้งเดิม (2.21)-(2.23) ค่าเฉลี่ยของคะแนนที่สังเกตได้เท่ากับค่าเฉลี่ยของคะแนนจริง ดังนั้นข้อตกลงเบื้องต้นของจุดตัดแกน X (เมื่อ Y=0) ของสมการการถดถอยของคะแนนจริงที่เท่าเทียมกันสำหรับ T_X ด้วย T_V ในประชากร 1 และ 2 คือ

$$\mu_2(X) - \frac{\sigma_2(T_X)}{\sigma_2(T_V)}\mu_2(V) = \mu_1(X) - \frac{\sigma_1(T_X)}{\sigma_1(T_V)}\mu_1(V) \quad (2.28)$$

ในทำนองเดียวกัน จุดตัดของ T_Y ด้วยT_V คือ

$$\mu_1(Y) - \frac{\sigma_1(T_Y)}{\sigma_1(T_V)}\mu_1(V) = \mu_2(Y) - \frac{\sigma_2(T_Y)}{\sigma_2(T_V)}\mu_2(V) \quad (2.29)$$

ข้อตกลงเบื้องต้นความแปรปรวนของความคลาดเคลื่อน (error variance assumptions)

วิธีของ Levine มีข้อตกลงเบื้องต้นว่า ความแปรปรวนของความคลาดเคลื่อนในการวัด (measurement error variance) สำหรับ X เหมือนกันทั้งในประชากร 1 และ 2 เช่นเดียวกับข้อตกลงเบื้องต้นของ Y และ V เพราะว่าในทฤษฎีการทดสอบแบบดั้งเดิมนั้นคะแนนจริงและความคลาดเคลื่อนไม่สัมพันธ์กัน โดยความแปรปรวนของความคลาดเคลื่อน (error

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

variance) คือ ความแตกต่างระหว่างความแปรปรวนของคะแนนที่สังเกตได้กับความแปรปรวนของคะแนนจริง ดังนั้นข้อตกลงเบื้องต้นสำหรับความแปรปรวนของความคลาดเคลื่อน คือ

$$\sigma_2^2(X) - \sigma_2^2(T_X) = \sigma_1^2(X) - \sigma_1^2(T_X) \quad (2.30)$$

$$\sigma_1^2(Y) - \sigma_1^2(T_Y) = \sigma_2^2(Y) - \sigma_2^2(T_Y) \quad (2.31)$$

$$\sigma_1^2(V) - \sigma_1^2(T_V) = \sigma_2^2(V) - \sigma_2^2(T_V) \quad (2.32)$$

ผลลัพธ์ของโมเดลร่วมกันเชิงแบบฉบับ (classical congeneric model result)

Classical Congeneric Model คือ ข้อตกลงเบื้องต้นของ X และ V และของประชากรเพียงกลุ่มเดียว โมเดลดังกล่าวระบุว่า T_Xและ T_V มีความสัมพันธ์กันเป็นเส้นตรง ถ้า $T_X = \lambda_X T + \delta_X$ และ $T_V = \lambda_V T + \delta_V$ โดย λ คือ ความชัน และ δ คือ จุดตัดแกน X ดังนั้นสมการของ X และ V คือ

$$X = T_X + E_X = (\lambda_X T + \delta_X) + E_X \quad (2.33)$$

$$V = T_V + E_V = (\lambda_V T + \delta_V) + E_V \quad (2.34)$$

ถ้าให้เทอม γ คือ $\sigma(T_X)/\sigma(T_V)$ จากสมการ (2.33) และ (2.34) จะได้ว่า

$$\gamma = \frac{\sigma(T_X)}{\sigma(T_V)} = \frac{\lambda_X \sigma(T)}{\lambda_V \sigma(T)} = \frac{\lambda_X}{\lambda_V} \quad (2.35)$$

γ สามารถใช้อธิบายในข้อสอบร่วมได้ 2 กรณี คือ 1) ข้อสอบร่วมแบบภายใน V ซึ่งถูกรวมไว้ใน X 2) ข้อสอบร่วมแบบภายนอก ซึ่ง V และ X แยกกันคนละฉบับ ดังนี้

ข้อสอบร่วมแบบภายใน (internal Anchor)

$$\gamma_1 = \frac{1}{\alpha_1(V|X)} = \frac{\sigma_1^2(X)}{\sigma_1(X,V)} \quad (2.36)$$

$$\gamma_2 = \frac{1}{\alpha_2(V|Y)} = \frac{\sigma_2^2(Y)}{\sigma_2(Y,V)} \quad (2.37)$$

ข้อสอบร่วมแบบภายนอก (external Anchor)

$$\gamma_1 = \frac{\sigma_1^2(X)+\sigma_1(X,V)}{\sigma_1^2(V)+\sigma_1(X,V)} \quad (2.38)$$

และ

$$\gamma_2 = \frac{\sigma_2^2(Y)+\sigma_2(Y,V)}{\sigma_2^2(V)+\sigma_2(Y,V)} \quad (2.39)$$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.2.2.3 Levine True Score Method

ความแตกต่างระหว่างวิธีที่ใช้คะแนนที่สังเกตได้และวิธีที่ใช้คะแนนจริง คือ วิธีที่ใช้คะแนนที่สังเกตได้โดยใช้สมการ (2.1) ในการปรับเทียบคะแนนที่สังเกตได้ X ไปยังสเกลของคะแนนที่สังเกตได้ Y ขณะที่วิธีที่ใช้คะแนนจริงเป็นการปรับเทียบคะแนนจริง โดยสมการที่ใช้ในการแปลงคะแนนจริงของ X ไปยังสเกลของคะแนนจริงของ Y (Kolen & Brennan, 2004) คือ

$$l_{Y_s}(t_x) = \frac{\sigma_s(T_Y)}{\sigma_s(T_X)}[t_x - \mu_S(T_X)] + \mu_S(T_Y)$$

ในทฤษฎีการทดสอบแบบดั้งเดิม ค่าเฉลี่ยของคะแนนที่สังเกตได้เท่ากับค่าเฉลี่ยของคะแนนจริง ดังนั้น

$$l_{Y_s}(t_x) = \frac{\sigma_s(T_Y)}{\sigma_s(T_X)}[t_x - \mu_S(X)] + \mu_S(Y) \qquad (2.40)$$

แต่ในทางปฏิบัติสามารถใช้คะแนนที่สังเกตได้ (observed score) มาใช้แทนที่คะแนนจริง (true score) ได้ ดังนั้นคะแนนที่สังเกตได้จากแบบสอบ X สามารถปรับเทียบไปยังคะแนนที่สังเกตได้จากแบบสอบ Y โดยใช้สมการ

$$l_Y(x) = \frac{\gamma_2}{\gamma_1}[x - \mu_1(X)] + \mu_1(Y) + \gamma_2[\mu_1(V) - \mu_2(V)]$$

2.2.2.4 วิธี CLE (Chained Linear Equating)

Davier (2011) ยังได้เสนอว่าวิธีการปรับเทียบเชิงเส้นตรงอีกวิธีหนึ่ง คือ วิธี CLE (Chained Linear Equating) ซึ่งเป็นการเชื่อมโยงเชิงเส้นตรงแบบสมมาตร (symmetric linear links) จากคะแนน X ไปยังคะแนน A และจากคะแนน Y ไปยังคะแนน A ที่ไม่แปรผันไปตามประชากร ดังนั้นการเชื่อมโยงเชิงเส้นตรงแบบสมมาตรจาก X ไป A ในประชากร 1 คือ

$$a = f(x) = \alpha_P + \beta_1 x \qquad (2.41)$$

เมื่อ a คือ ค่าของ A และ x คือ ค่าของ X และ

$$\beta_1 = \sigma_1(A)/\sigma_1(X) \qquad (2.42)$$

$$\alpha_P = \mu_1(A) - \beta_1\mu_1(X) \qquad (2.43)$$

ภายใต้ข้อตกลงเบื้องต้นของสมการ 2.40 ที่ไม่แปรผันไปตามประชากร ดังนั้นค่าที่ได้จากสมการ 2.41 และ 2.42 ในประชากร 1 และประชากร 2 จึงเท่ากัน นั่นคือ

$$\beta_1 = \frac{\sigma_2(A)}{\sigma_2(X)} = \frac{\sigma_1(A)}{\sigma_1(X)} \qquad (2.44)$$

$$\mu_2(A) - \beta_2\mu_2(X) = \mu_1(A) - \beta_1\mu_1(X) \qquad (2.45)$$

สมการ 2.44 และ 2.45 สามารถใช้พิสูจน์หา $\mu_2(X)$ และ $\sigma_2(X)$ โดยกำหนดให้ $\mu_2(A) = w\mu_1(A) + (1-w)\mu_3(A)$ ภายใต้ข้อตกลงเบื้องต้นของความไม่แปรผันไปตามประชากรของวิธี CLE (Chained Linear Equating) ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของ X และ Y ในประชากร 2 หาได้จากสมการ 2.46

(a) $\mu_2(X) = \mu_1(X) - (1-w)\left[\frac{\sigma_1(X)}{\sigma_1(A)}\right][\mu_1(A) - \mu_3(A)]$

(b) $\mu_2(Y) = \mu_3(Y) + w\left[\frac{\sigma_3(Y)}{\sigma_3(A)}\right][\mu_1(A) - \mu_3(A)]$

(c) $\sigma_2(X) = \sigma_2(A)\left[\frac{\sigma_1(X)}{\sigma_1(A)}\right]$ (2.46)

(d) $\sigma_2(Y) = \sigma_2(A)\left[\frac{\sigma_3(Y)}{\sigma_3(A)}\right]$

แทนที่สมการ 2.46 ในสมการ 2.41 จะได้สมการสำหรับการปรับเทียบด้วยวิธี CLE (Chained Linear Equating) ดังนี้

$$y = f(x) = \mu_2(Y) + \frac{\sigma_2(Y)\sigma_1(A)}{\sigma_2(A)\sigma_1(X)}[x - \mu_1(X)] + \frac{\sigma_2(Y)}{\sigma_2(A)}[\mu_1(A) - \mu_2(A)] \quad (2.47)$$

การตัดเทอมน้ำหนัก w ออกจากสมการ 2.47 และวิธี CLE (Chained Linear Equating) ไม่ขึ้นกับสัดส่วนเชิงสัมพัทธ์ (proportions of populations) ของประชากร 1 และ 3 ในประชากร 2

2.2.3 วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์ (equipercentile equating)

Kolen และ Brennan (2004) กล่าวว่า การปรับเทียบอีควิเปอร์เซ็นไทล์ใช้เส้นโค้งในการอธิบายความแตกต่างของค่าความยากระหว่างแบบสอบสองฉบับ ดังนั้นวิธีการปรับเทียบอีควิเปอร์เซ็นไทล์จึงเป็นการแปลงคะแนนแบบไม่ใช่เส้นตรง ตัวอย่างการใช้วิธีการดังกล่าว เช่น แบบสอบ X มีความยากมากกว่าแบบสอบ Y ที่ระดับคะแนนสูงหรือคะแนนต่ำ แต่ไม่มีความแตกต่างในคะแนนระดับกลาง ๆ

ฟังก์ชั่นการปรับเทียบเป็นฟังก์ชั่นการปรับเทียบอีควิเปอร์เซ็นไทล์ ก็ต่อเมื่อการแจกแจงของคะแนนจากแบบสอบ X ที่แปลงไปยังสเกลของแบบสอบ Y เหมือนกับการแจกแจงของคะแนนจากแบบสอบ Y โดยการปรับเทียบอีควิเปอร์เซ็นไทล์ใช้นิยามการปรับเทียบคะแนนว่า คะแนนจากแบบสอบ X และแบบสอบ Y มีความเท่าเทียมกันก็ต่อเมื่อคะแนนทั้งสองฉบับอยู่ที่ตำแหน่งเปอร์ เซ็นไทล์เดียวกัน

ในการปรับเทียบอีควิเปอร์เซ็นไทล์ ต้องพิจารณาว่าคะแนนจากแบบสอบ X มีตำแหน่งเปอร์เซ็นไทล์เดียวกับคะแนนจากแบบสอบ Y หรือไม่

กำหนดให้ y คือ คะแนนจากแบบสอบ Y K_{Y} คือ จำนวนข้อสอบในแบบสอบ Y

$g(y)$ คือ คะแนนเต็มหน่วยของ y

$G(y)$ คือ การแจกแจงสะสมแบบไม่ต่อเนื่อง

$Q(y)$ คือ ตำแหน่งเปอร์เซ็นไทล์ของ y

y_U^* คือ คะแนนแบบจำนวนเต็มจากแบบสอบ Y

x คือ คะแนนเฉพาะจากแบบสอบ X ที่ต้องการปรับเทียบ

$P(x)$ คือ ตำแหน่งเปอร์เซ็นไทล์ของคะแนนจากแบบสอบ X

Q^{-1} คือ ส่วนกลับของฟังก์ชั่นตำแหน่งเปอร์เซ็นไทล์ของแบบสอบ Y

$e_Y(x)$ คือ ฟังก์ชั่นการปรับเทียบที่ใช้ในการแปลงคะแนนจากแบบสอบ X ไปยังสเกลของแบบสอบ Y

$$e_Y(x) = y = Q^{-1}[P(x)]$$

$$= \frac{\frac{P(x)}{100} - G(y_U^* - 1)}{G(y_U^*) - G(y_U^* - 1)} + (y_U^* - 5), \quad 0 \le P(x) < 100$$

$$= K_{\mathrm{Y}} + .5, \qquad P(x) = 100 \quad (2.47)$$

จากสมการ (2.47) ต้องหา 1) ตำแหน่งเปอร์เซ็นไทล์ของคะแนนเฉพาะจากแบบสอบ X ที่ต้องการปรับเทียบในการแจกแจงของแบบสอบ X 2) หาคะแนนของแบบสอบ Y ที่มีตำแหน่ง

เปอร์เซ็นไทล์ตรงกับการแจกแจงของแบบสอบ Y ถ้าสมการ (2.47) สมมาตรแสดงว่าคะแนนจากแบบสอบทั้งสองฉบับเท่าเทียมกัน (สมการ 2.47 ใช้ได้ก็ต่อเมื่อความน่าจะเป็น (probabilities) ของคะแนนจากแบบสอบ Y เป็นศูนย์)

โดยวิธีการปรับเทียบอีควิเปอร์เซ็นไทล์ในรูปแบบผู้สอบต่างกลุ่มกัน พิจารณาการแจกแจงของคะแนนรวม และคะแนนของแบบสอบร่วมแทนที่จะพิจารณาเพียงค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และความแปรปรวนร่วม การปรับเทียบอีควิเปอร์เซ็นไทล์ คือ กระบวนการปรับเทียบคะแนนที่สังเกตได้ ดังนั้นการใช้วิธีดังกล่าวกับรูปแบบผู้สอบต่างกลุ่มแบบใช้ข้อสอบร่วมจึงต้องการการสังเคราะห์ประชากร (Kolen & Brennan, 2004)

2.2.3.1 วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์ในการประมาณค่าความถี่ (frequency estimation equipercentile equating)

<u>การแจกแจงอย่างมีเงื่อนไข</u> (conditional distributions)

กำหนดให้ $f(x,v)$ คือ ความน่าจะเป็นที่ $X=x$ และ $V=v$ ส่วน $f(x)$ คือ ความน่าจะเป็นที่ $X=x$ ขณะที่ $h(v)$ คือ ความน่าจะเป็นที่ $V=v$และ$f(x|v)$ คือ ความน่าจะเป็นที่ $X=x$ เพื่อให้ $V=v$

$$f(x|v)=\frac{f(x,v)}{h(v)} \qquad (3.1)$$

นั่นคือ

$$f(x,v)=f(x|v)h(v) \qquad (3.2)$$

<u>วิธีการประมาณค่าความถี่</u> (frequency estimated method)

วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์ในการประมาณค่าความถี่จำเป็นต้องแสดงการแจกแจงของประชากรที่เกิดจากการสังเคราะห์ (synthetic population) การแจกแจงนี้พิจารณาผลรวมน้ำหนักของการแจกแจงประชากรในแต่ละกลุ่ม สำหรับแบบสอบ X และแบบสอบ Y คือ

$$f_s(x)=w_1f_1(x)+w_2f_2(x) \qquad (3.3)$$

และ

$$g_s(y)=w_1g_1(y)+w_2g_2(y) \qquad (3.4)$$

โดย เลข 1 แทนประชากรที่สอบด้วยแบบสอบ X ส่วนเลข 2 แทนประชากรที่สอบด้วยแบบสอบ Y เช่นเดียวกับ f และ g แทนการแจกแจงของแบบสอบ X และแบบสอบ Y ตามลำดับ w_1 และ$w_2(w_1+w_2)=1$ แทนน้ำหนักของประชากร 1และ2 ซึ่งมาจากประชากรที่เกิดจากการสังเคราะห์

ข้อตกลงเบื้องต้นที่ใช้ในวิธีการปรับเทียบอีควิเปอร์เซ็นไทล์ทั้งแบบสอบ X และ Y โดยการแจกแจงอย่างมีเงื่อนไขของคะแนนรวมทำให้ $V=v$ ทั้งในประชากร1 และ 2 ข้อตกลงเบื้องต้นคือ

$f_1(x|v)=f_2(x|v)$ สำหรับ v ทั้งหมด

และ $g_1(y|v)=g_2(y|v)$ สำหรับ v ทั้งหมด (3.5)

ดังนั้นการแจกแจงประชากรที่เกิดจากการสังเคราะห์ คือ

$$f_s(x) = w_1 f_1(x) + w_2 \sum_v f_1(x|v)\, h_2(v)$$

และ

$$g_s(y) = w_1 g_1(y|v) h_1(v) + w_2 g_2(y) \qquad (3.6)$$

ถ้า P_s คือ ฟังก์ชั่นของตำแหน่งเปอร์เซ็นไทล์ของแบบสอบ X และ Q_s คือ ฟังก์ชั่นของตำแหน่งเปอร์เซ็นไทล์ของแบบสอบ Y ขณะที่ P_s^{-1} และ Q_s^{-1} คือฟังก์ชั่นของตำแหน่งเปอร์เซ็นไทล์

ฟังก์ชั่นของตำแหน่งเปอร์เซ็นไทล์ของประชากรที่เกิดจากการสังเคราะห์ คือ

$$e_{Ys}(x) = Q_s^{-1}[P_s(x)] \qquad (3.7)$$

2.2.3.2 วิธีการเชิงเส้นตรงของ Braun-Holland (Braun-Holland linear method)

Braun และ Holland (1982 อ้างถึงใน Kolen & Brennan, 2004) นำเสนอวิธีเชิงเส้นตรงที่ใช้ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน ซึ่งเป็นผลมาจากการใช้ข้อตกลงเบื้องต้นในการประมาณค่าความถี่ที่นำไปสู่การปรับเทียบเชิงเส้นตรง วิธีการนี้ใกล้เคียงกับวิธีของ Tucker ภายใต้ข้อตกลงเบื้องต้นในการประมาณค่าความถี่ของสมการ (3.5) ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนจากแบบสอบ X สำหรับประชากรที่เกิดจากการสังเคราะห์ คือ

$$\mu_s(X) = \sum_x x f_s(x) \qquad (3.8)$$

$$\sigma_s^2(X) = \sum_x [x - \mu_s(X)]^2 f_s(x) \qquad (3.9)$$

ค่า $f_s(x)$ ได้มาจากสมการ (3.6) สมการผลลัพธ์นี้ก็คือ วิธีเชิงเส้นตรงของ Braun-Holland โดย Braun และ Holland กล่าวว่า การปรับเทียบคือผลลัพธ์ที่ได้จากการใช้วิธีเชิงเส้นตรงของ Braun-Holland ซึ่งเหมือนกับวิธีเชิงเส้นตรงของ Tucker กรณีที่เหมาะสมในการนำวิธีดังกล่าวไปใช้ ได้แก่

(1) สมการถดถอยของ X ด้วย V และสมการถดถอยของ Y ด้วย V เป็นเส้นตรงทั้งคู่

(2) สมการถดถอยของ X ด้วย V และ Y ด้วย V เป็น Homoscedastistic ค่าเฉลี่ยของคุณลักษณะนี้ คือ ความแปรปรวนของ X ที่ให้ v เป็นเหมือนกับ v ทั้งหมด และความแปรปรวนของ Y ที่ให้ v เป็นเหมือนกับ v ทั้งหมด

วิธีเชิงเส้นตรงของ Braun-Holland เป็นการประยุกต์ใช้วิธีเชิงเส้นตรงของ Tucker เมื่อสมการการถดถอยของคะแนนทั้งหมดจากข้อสอบร่วมไม่เป็นเส้นตรง โดยวิธีเชิงเส้นตรงของ Braun-Holland มีการคำนวณที่ยากกว่าวิธีเชิงเส้นตรงของ Tucker จึงไม่นิยมใช้มากนัก

2.2.3.3 วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ (chained equipercentile equating)

วิธีการนี้คะแนนของแบบสอบ X ถูกแปลงไปยังคะแนนของแบบสอบร่วมที่ใช้ผู้สอบจากประชากร 1 เมื่อคะแนนของแบบสอบร่วมถูกปรับเทียบไปยังคะแนนของแบบสอบ Y ที่ใช้ผู้สอบจากประชากร 2 โดยการแปลงคะแนนทั้งสองแบบถูกโยงไว้ด้วยกันเพื่อใช้แปลงคะแนนของแบบสอบ X ไปยังคะแนนของแบบสอบ Y

ขั้นตอนการปรับเทียบคะแนนของวิธีดังกล่าว มีดังนี้

1. หาความสัมพันธ์ของการปรับเทียบอีควิเปอร์เซ็นไทล์สำหรับคะแนนที่แปลงด้วยแบบสอบ X ไปยังแบบสอบร่วม ซึ่งผู้สอบมาจากประชากร1 โดยใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์ (ผู้สอบทุกคนต้องทำทั้งแบบสอบ X และแบบสอบร่วม) ฟังก์ชั่นของอีควิเปอร์เซ็นไทล์นี้ คือ $e_{V1}(x)$

2. หาความสัมพันธ์ของการปรับเทียบอีควิเปอร์เซ็นไทล์สำหรับคะแนนที่แปลงด้วยแบบสอบร่วมไปยังคะแนนของแบบสอบ Y ซึ่งผู้สอบมาจากประชากร 2 ฟังก์ชั่นผลลัพธ์ที่ได้ คือ $e_{Y2}(v)$

3. ปรับเทียบคะแนนของแบบสอบ X ไปยังแบบสอบ Y เริ่มจากแปลงจากแบบสอบ X ไปยังคะแนนของแบบสอบร่วมโดยใช้ $e_{V1}(x)$ หลังจากนั้นปรับเทียบคะแนนผลลัพธ์ของแบบสอบร่วมไปยังแบบสอบ Y โดยใช้ $e_{Y2}(v)$

ในการหาอีควิเปอร์เซ็นไทล์ของแบบสอบ Y ที่เท่าเทียมกับคะแนนของแบบสอบ X หาได้จาก

$$e_{Y(chain)} = e_{Y2}[e_{V1}(x)]$$

การปรับเทียบด้วยวิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ (chained equipercentile equating) ไม่ได้ต้องการพิจาณาการแจกแจงของรอยต่อระหว่างคะแนนทั้งหมดและคะแนนของแบบสอบร่วม ดังนั้นการคำนวณของวิธีการนี้ จึงละเอียดกว่าการประมาณความถี่และการปรับเทียบอีควิเปอร์เซ็นไทล์ แต่อย่างไรก็ตามวิธีการดังกล่าวก็เพิ่งถูกนำมาใช้ เนื่องจาก 1) เป็นการปรับเทียบคะแนนทั้งหมด (แบบสอบยาว) ไปยังคะแนนของแบบสอบร่วม (แบบสอบสั้น) ซึ่งการปรับเทียบแบบสอบที่มีความยาวไม่เท่ากันดูไม่ค่อยสมเหตุสมผล 2) วิธีการนี้ไม่ได้รวมประชากรที่สังเคราะห์ได้ไว้ด้วยกันโดยตรง

นอกจากนี้ Braun และ Holland (1982 cited in Kolen & Brennan, 2004) ยังกล่าวว่า วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ วิธีการประมาณค่าความถี่ และวิธีการปรับเทียบอีควิเปอร์เซ็นไทล์ให้ผลที่ไม่เหมือนกัน รวมไปถึงข้อตกลงเบื้องต้นของวิธีการทั้งสามไม่เหมือนกันอีกด้วย และ Harris และ Kolen (1990 cited in Kolen & Brennan, 2004) ยังกล่าวว่าวิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ (chained equipercentile equating) ไม่จำเป็นต้องใช้กับประชากรสองกลุ่มที่เหมือนกันมาก ดังนั้นวิธีการนี้อาจสามารถนำไปใช้ในกรณีที่กลุ่มตัวอย่างสองกลุ่มแตกต่างกัน

2.2.3.4 วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์ด้วยการปรับโค้งให้ราบเรียบ (smoothing in equipercentile equating)

Kolen และ Brennan (2004) กล่าวว่า วิธีปรับโค้งให้ราบเรียบ (smoothing method) เป็นวิธีการที่ใช้เพื่อประมาณการแจกแจงเชิงประจักษ์และความสัมพันธ์อีควิเปอร์เซ็นไทล์ที่มีการแจกแจงไม่ราบเรียบซึ่งเป็นลักษณะเฉพาะของประชากร โดยคาดหวังว่าผลจากการประมาณดังกล่าวมีความแม่นยำมากกว่าการแจกแจงที่ยังไม่ได้ปรับให้ราบเรียบ (unsmoothed distribution) แต่อย่างไรก็ตามการใช้วิธีการปรับโค้งให้ราบเรียบอาจได้ผลของการประมาณการแจกแจงของประชากรที่แย่กว่าการประมาณค่าที่ไม่ได้ปรับโค้งให้ราบเรียบ โดยประเภทของการปรับโค้งให้ราบเรียบมี 2 ประเภท คือ การปรับการแจกแจงคะแนน (presmoothing) และการปรับอีควิเปอร์เซ็นไทล์ (postsmoothing) โดยการปรับการแจกแจงคะแนน (presmoothing) เป็นการปรับการแจกแจงของคะแนนให้ราบเรียบ ขณะที่การปรับอีควิเปอร์เซ็นไทล์ (postsmoothing) เป็นการปรับความเท่าเทียมของอีควิเปอร์เซ็นไทล์ (equipercentile equivalents) ในที่นี้ผู้วิจัยเน้นการอธิบายถึงการปรับโค้งให้ราบเรียบด้วยการปรับการแจกแจงคะแนน (presmoothing) เป็นหลัก

การปรับการแจกแจงคะแนน (presmoothing) เป็นวิธีที่ใช้ในการปรับการแจกแจงของคะแนนให้เป็นโค้งราบเรียบ ซึ่งการแจกแจงที่ได้จากการปรับนั้น ต้องมีการประมาณการแจกแจงที่ดี นั่นคือ การรักษาโมเมนต์ไว้ (moment preservertion) ซึ่งในการรักษาโมเมนต์ไว้นั้น การแจกแจงที่ราบเรียบต้องมีโมเมนต์สำคัญ (central moment) เหมือนกับการแจกแจงที่สังเกตได้ อธิบายอย่างง่าย คือ ค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานของการแจกแจงที่ปรับให้เป็นโค้งราบเรียบต้องเหมือนกับค่าเฉลี่ยและค่าส่วนเบี่ยงมาตรฐานของการแจกแจงที่ยังไม่ได้ปรับให้เป็นโค้งราบเรียบ โดยการปรับการแจกแจงคะแนน (presmoothing) อาจใช้โมเดลแบบจำลองล็อกเชิงเส้นพหุนาม (polynomial log-linear model) หรือโมเดลการแจกแจงคะแนนจริง (strong true score distribution model) ซึ่งหลังจากที่ทำการปรับการแจกแจงคะแนน (presmoothing) ข้อมูลด้วยสองวิธีดังกล่าวแล้ว สามารถปรับเทียบคะแนนจากแบบสอบ X ไปยังคะแนนจากแบบสอบ Y ได้โดยใช้การแจงแจงที่ถูกปรับให้เป็นโค้งราบเรียบแล้ว (Kolen & Brennan, 2004) ในที่นี้ผู้วิจัยเน้นการอธิบายไปที่โมเดลแบบจำลองล็อกเชิงเส้นพหุนาม (polynomial log-linear model) ซึ่งถูกนำมาใช้ในปรับเทียบคะแนนสำหรับการออกแบบการปรับเทียบที่เป็นเกณฑ์

โมเดลแบบจำลองล็อกเชิงเส้นพหุนาม (polynomial log-linear model) เป็นวิธีที่มีความยืดหยุ่นในการปรับโค้งการแจกแจงให้มีความราบเรียบด้วยดีกรี (degree) ต่าง ๆ โดยที่ดีกรีของแต่ละเทอมของพหุนาม (polynomial) เป็นการกำหนดโมเมนต์ของการแจกแจงคะแนนดิบที่ต้องรักษา/คงไว้ ตัวอย่างเช่น โมเดลที่ประกอบด้วยกำลัง 1 2 และ 3 ได้สร้างการแจกแจงที่ปรับโค้งให้ราบเรียบ ซึ่งเหมาะสมหรือเข้ากันกับค่าเฉลี่ย ความแปรปรวน และความเบ้ของคะแนนดิบดังสมการ $log(p_i) = \beta_0 + \beta_1 x_i^1 + \beta_2 x_i^2 + \beta_3 x_i^3$ โดยแสดงถึงล็อกของความถี่สัมพัทธ์ที่คาดไว้ (log of expected relative frequency, p_i) ที่จุดคะแนน i ซึ่งแสดงไว้ในเทอมของค่าคงที่ปกติ

(normalizing constant, β_0) และฟังก์ชั่นของคะแนนที่ถูกถ่วงน้ำหนัก (weight score function, x_1, x_2, x_3) ของค่าคะแนนที่เป็นไปได้ของแบบสอบ X (Albado, 2014)

โดยดีกรีที่ใช้ในการปรับโค้งให้ราบเรียบนี้ ถูกนำสร้างโมเดลต่างๆที่มีจำนวนของเทอมพหุนามที่แตกต่างกัน ซึ่งโมเดลเหล่านี้ถูกนำมาเปรียบเทียบกัน เพื่อหาโมเดลที่มีความเหมาะสมกับข้อมูล (Kolen & Brennan, 2004 cited in Albado, 2014) ซึ่งในโปรแกรม Equate package ของ Albado (2014) ค่าล็อกเชิงเส้นฟังก์ชั่น (function loglinear vaules) ได้สร้างค่าที่เหมาะสมสำหรับการแจกแจงตัวแปรเดียว (univariate) และการแจกแจงสองตัวแปร (bivariate) ซึ่งค่าดังกล่าวสามารถใช้ในการเปรียบเทียบโมเดลโดยยึดตามจำนวนของดัชนีที่เหมาะสม โดยค่าล็อกเชิงเส้นฟังก์ชั่นได้ใช้ค่า Newton-Rapson maximum likelihood สร้างโมเดลภายใต้โปรแกรม SAS macro ของ Moses และ Davier (2011) ซึ่งค่าประมาณที่เหมาะสมในการเปรียบเทียบ (comparable fitted estimates) หาได้จากฟังก์ชั่น Fitting Generalized Linear Model (glm)

2.2.4 วิธีการปรับเทียบโดยใช้ทฤษฎีการตอบสนองข้อสอบ (items response theory method)

Hambleton และคณะ (1991) กล่าวว่า ในทฤษฎีการตอบสนองข้อสอบนั้น ค่าพารามิเตอร์ความสามารถ (θ) ของผู้สอบแต่ละคนไม่ได้รับผลจากการใช้ชุดข้อสอบที่แตกต่างกัน นั่นคือการประมาณค่าความสามารถของผู้สอบไม่แปรผันตามความแตกต่างของชุดข้อสอบที่ใช้ ดังนั้นเมื่อทราบค่าพารามิเตอร์ก็สามารถประมาณค่าความสามารถของผู้สอบที่ทำแบบสอบต่างชุดกันให้อยู่บนสเกลเดียวกัน จึงไม่จำเป็นต้องปรับเทียบคะแนนหรือสเกลระหว่างแบบสอบ

กรณีที่ไม่ทราบค่าพามิเตอร์ของข้อสอบ เมื่อนำแบบสอบฉบับเดียวกันไปทดสอบกับผู้สอบ 2 กลุ่ม ฟังก์ชั่นการตอบสนองข้อสอบนั้นไม่แปรผันตามกลุ่ม ถ้าในการคำนวณได้กำหนดค่าความสามารถ (θ) หรือค่าความยากของสอบ (b) ให้อยู่บนสเกลของค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน

กรณีที่ไม่ทราบทั้งค่าพารามิเตอร์ของข้อสอบและพารามิเตอร์ของผู้สอบมาก่อน เมื่อนำแบบสอบร่วมฉบับเดียวกันไปทดสอบกับผู้สอบสองกลุ่ม (แบบสอบร่วม) หรือผู้สอบกลุ่มเดียวกันทำแบบสอบสองฉบับ (ผู้สอบร่วม) แล้วทำการวิเคราะห์ข้อมูลแยกกัน ทฤษฎีการตอบสนองข้อสอบได้ประมาณค่าความสามารถของผู้สอบและค่าพารามิเตอร์ของข้อสอบ ผลที่ได้อาจได้ค่าที่แตกต่างกันระหว่างแบบสอบหรือระหว่างกลุ่มผู้สอบ ความแตกต่างพารามิเตอร์เกิดจากการกำหนดความเฉลี่ยของการแจกแจงความสามารถของผู้สอบ หรือค่าเฉลี่ยของการแจกแจงค่าพารามิเตอร์ความยากของข้อสอบ (b) ในการวิเคราะห์แต่ละครั้งที่ไม่เหมือนกัน แต่อย่างไรก็ตามค่าพารามิเตอร์ระหว่างกลุ่มหรือฉบับมีความสัมพันธ์เชิงเส้นตรง (linear relationship) ซึ่งใช้ในการสร้างสมการแปลงค่าพารามิเตอร์ระหว่างกลุ่มหรือฉบับให้อยู่บนสเกลเดียวกันได้ (ศิริชัย กาญจนวาสี, 2555)

ศิริชัย กาญจนวาสี (2555) ได้นำเสนอวิธีการปรับเทียบโดยใช้ทฤษฎีการตอบสนองข้อสอบไว้ดังนี้

2.2.4.1 วิธีสมการถดถอย (regression method)

การปรับค่าพารามิเตอร์ความยาก (b) ของข้อสอบร่วมจากแบบสอบ X ให้อยู่บนมาตรหรือสเกลเดียวกับแบบสอบ Y ด้วยสมการเส้นตรง (ศิริชัย กาญจนวาสี, 2555) คือ

$$b_{y_c} = \alpha b_{x_c} + \beta$$

โดยที่ b_{y_c}, b_{x_c} คือ ค่าประมาณพารามิเตอร์ความยากของข้อสอบร่วมจากแบบสอบ Y และ X

α, β คือ ค่าคงที่หรือสัมประสิทธิ์การถดถอย

เมื่อ $\alpha = r\frac{s_{yc}}{s_{x_c}}$ และ $\beta = \bar{b}_{y_c} - \alpha\bar{b}_{X_c}$

โดยที่ r คือ สัมประสิทธิ์ความสัมพันธ์ระหว่างค่าประมาณพารามิเตอร์ความยากของข้อสอบร่วมจากแบบสอบ X และ Y

ในการปรับค่าประมาณความสามารถ (θ) ของผู้สอบร่วมที่สอบฉบับ X บนมาตรหรือสเกลเดียวกันกับฉบับ Y ด้วยสมการดังนี้

$$\theta_{y_c} = \alpha\theta_{x_c} + \beta$$

ปัญหาของวิธีสมการถดถอย คือ ความสมมาตรของผลการเปรียบเทียบคะแนน เพราะสัมประสิทธิ์ของการทำนาย b_{y_c} จาก b_{y_c} แตกต่างจากการทำนาย b_{x_c} จาก b_{y_c} นั่นคือ การปรับเทียบคะแนนจากแบบสอบ X และY ให้ผลต่างจากการเปรียบคะแนนจากแบบสอบ Y ไป X

2.2.4.2 วิธีค่าเฉลี่ยและซิกมา (mean and sigma method)

$$b_{y_c} = \alpha b_{x_c} + \beta$$

จะได้ว่า $\bar{b}_{y_c} = \alpha\bar{b}_{x_c} + \beta$

และ $s_{y_c} = \alpha s_{x_c}$

ดังนั้น $\alpha = \frac{s_{y_c}}{s_{x_c}}$

และ $\beta = \bar{b}_{y_c} - \alpha\bar{b}_{x_c}$

เพราะว่า $b_{y_c} = \alpha b_{x_c} + \beta$

การแปลงคะแนนจาก b_{y_c} ไปยัง b_{x_c} จะได้ว่า $b_{x_c} = \frac{b_{y_c} - \beta}{\alpha}$

ดังนั้นวิธีค่าเฉลี่ยและซิกมา จึงมีลักษณะสมมาตร (เมื่อใช้รูปแบบผู้สอบร่วมกัน ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าประมาณความสามารถใช้เพื่อระบุ α และ β) (ศิริชัย กาญจนวาสี, 2555)

- กรณี 1 พารามิเตอร์

$$b^*_{y_c} = b_{x_c} + \beta$$

เนื่องจาก $\alpha = 1$ จะได้ว่า

$$\bar{b}_{y_c} = \bar{b}_{x_c} + \beta$$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ดังนั้น $\beta = \bar{b}_{y_c} - \bar{b}_{x_c}$

- กรณี 2 และ 3 พารามิเตอร์

$$b^*_{y_c} = \alpha b_{x_c} + \beta$$

และ $a^*_y = \frac{a_X}{\alpha}$

2.2.4.3 วิธีปรับแก้ค่าเฉลี่ยและซิกมา (robust mean and sigma method)

วิธีค่าเฉลี่ยและซิกมาไม่ได้พิจารณาถึงระดับความแม่นยำในการประมาณค่าพารามิเตอร์ของข้อสอบ Linn และคณะจึงได้เสนอวิธีการปรับแก้ค่าเฉลี่ยและซิกมา โดยถ่วงน้ำหนักด้วยของความเคลื่อนในการกะประมาณค่าพารามิเตอร์ของข้อสอบร่วมรายข้อ ข้อสอบร่วมที่มีความแปรแปรวนระหว่างฉบับสูง มีน้ำหนักมาก ส่วนข้อสอบร่วมที่มีความแปรแปวนระหว่างฉบับต่ำ มีน้ำหนักต่ำ โดยค่าน้ำหนักเป็นส่วนกลับของความแปรปรวนของค่าประมาณตัวที่มีค่ามาก มีขั้นตอนการคำนวณ ดังนี้ (ศิริชัย กาญจนวาสี, 2555)

(1) สำหรับแต่ละคู่ ($b_{y_{ci}}, b_{x_{ci}}$) คำนวณน้ำหนักได้จาก

$$\mathrm{Wi} = [\text{ค่าตัวมากระหว่าง}\{\mathrm{v(b_{Yci})},\ \mathrm{v(b_{Xci})}\}]^{-1}$$

เมื่อ $v(\mathrm{b_{Yci}}),\ v(\mathrm{b_{Xci}})$ คือ ความแปรวนของค่าประมาณพารามิเตอร์ b ของข้อสอบร่วมฉบับ y และ x

(2) กำหนดค่าประมาณของพารามิเตอร์ที่ถ่วงน้ำหนัก

$$w'_i = w_i / \sum_{j=1}^{k} w_j$$

(3) คำนวณค่าประมาณของพารามิเตอร์ที่ถ่วงน้ำหนัก

$$b'_{y_{ci}} = w'_i b_{y_{ci}}$$

$$b_{x_{ci}} = w'_i b_{x_{ci}}$$

(4) คำนวณค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าประมาณพารามิเตอร์ข้อสอบที่ถ่วงน้ำหนัก

(5) คำนวณค่า α และ β โดยใช้วิธีค่าเฉลี่ยและซิกมา

2.2.4.4 วิธีโค้งลักษณะข้อสอบ (characteristic curve method)

Stocking และ Lord (1983 อ้างถึงใน ศิริชัย กาญจนาวาสี, 2555) ได้เสนอวิธีโค้งลักษณะข้อสอบ โดยนำเสนอสารสนเทศที่ได้จากค่าพารามิเตอร์ความยากของข้อสอบและค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ

ถ้าคะแนนจริงของผู้สอบที่มีความสามารถ θ_a จากการทำแบบสอบ X ที่มีข้อสอบร่วม k ข้อ แทนด้วย $\mathrm{T_{Xa}}$

$$T_{X_a} = \sum_{i=1}^{k} P(\theta_a, b_{x_{ci}}, a_{x_{ci}}, c_{x_{ci}})$$

ในทำนองเดียวกัน คะแนนจริงของผู้สอบที่มีความสามารถเท่ากัน θ_a จากการทำแบบสอบ Y ที่มีข้อสอบร่วม k ข้อ

$$T_{Y_a} = \sum_{i=1}^{k} P(\theta_a, b_{y_{ci}}, a_{y_{ci}}, c_{y_{ci}})$$

สำหรับชุดของข้อสอบร่วม k ข้อ

$$b_{y_{ci}} = \alpha b_{x_{ci}} + \beta$$

$$a_{y_{ci}} = \frac{a_{x_{ci}}}{\alpha}$$

และ $c_{y_{ci}} = c_{x_{ci}}$

การคำนวณค่า α และ β ได้มาจากการทำให้ฟังก์ชั่น F มีค่าต่ำสุด

$$F = \frac{1}{N}\sum_{a=1}^{N}(T_{X_a} - T_{Y_a})^2$$

เมื่อ N คือ จำนวนผู้สอบ ส่วน F คือ ฟังก์ชั่นของ α และ β ที่แสดงความแตกต่างระหว่างคะแนนจริงของแบบสอบ X และ Y ของผู้สอบที่มีความสามารถเดียวกัน การคำนวณค่าคงที่ α และ β ต้องใช้กระบวนการทำซ้ำหลายรอบจนได้ค่าที่ดีที่สุด

รูปแบบการตอบข้อสอบรายข้อ เนื่องจากข้อมูลที่ได้จากการสอบมีหลายลักษณะ ทั้งข้อมูลแบบทวิภาค (dichotomous) และข้อสอบแบบหลายค่า (polytomous) ดังนั้นจึงมีการพัฒนารูปแบบเพื่อให้สอดคล้องกับลักษณะข้อมูลที่แตกต่างกัน โดยข้อมูลแบบทวิภาคนิยมใช้รูปแบบโลจิสติก (logistic model) ซึ่งแตกต่างไปตามจำนวนพารามิเตอร์ที่ใช้ในแต่ละรูปแบบ โดยมีรายละเอียดดังนี้ (ศิริชัย กาญจนวาสี, 2555)

<u>รูปแบบโลจิสติก 2 พารามิเตอร์</u> (two-parameter logistic model)

Birnbaum (1983 อ้างถึงใน ศิริชัย กาญจนาวาสี, 2555) ได้เสนอรูปแบบดังกล่าวในปี ค.ศ. 1957 ซึ่งเป็นโค้งลักษณะข้อสอบและเป็นฟังก์ชั่นของการแจกแจงที่มีสองพารามิเตอร์

$$P_i(\theta) = \frac{1}{1+e^{-Da_i(\theta - b_i)}} \qquad (i = 1, 2, 3, \ldots, n)$$

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นของผู้สอบที่มีความสามารถ θ ตอบข้อสอบข้อที่ i ถูกต้อง

b_i คือ ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i ซึ่งเป็นค่าที่แสดงตำแหน่งของ ICC ณ จุด θ ที่มีโอกาสตอบถูก 0.5

a_i คือ ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i ซึ่งเป็นค่าที่แสดงตำแหน่งของ ICC ณ ตำแหน่ง b_i

D คือ ค่าองค์ประกอบสเกล = 1.70

e = 2.710

จากรูปแบบนี้การเดาไม่เกิดขึ้น เมื่อค่าพารามิเตอร์ $a_i > 0$ และค่าความน่าจะเป็นในการตอบข้อสอบได้ถูกลดลงถึงศูนย์เมื่อความสามารถลดลง

<u>รูปแบบโลจิสติก 3 พารามิเตอร์</u> (three-parameter logistic model)

รูปแบบนี้เป็นการปรับปรุงสองพารามิเตอร์ โดยเพิ่มพารามิเตอร์ตัวที่สาม คือ พารามิเตอร์การเดาคำตอบ คือ พารามิเตอร์ c_i เข้าไปในรูปแบบนี้ ดังนี้

$$P_i(\theta) = c_i + \frac{(1-c_i)}{1+e^{-Da_i(\theta-b_i)}} \qquad (i = 1, 2, 3, \ldots, n)$$

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นของผู้สอบที่มีความสามารถ θ ตอบข้อสอบข้อที่ i ถูกต้อง

b_i คือ ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i ซึ่งเป็นค่าที่แสดงตำแหน่งของ ICC ณ จุด θ ที่มีโอกาสตอบถูก $\frac{(1+c_i)}{2}$

a_i คือ ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i ซึ่งเป็นค่าที่แสดงตำแหน่งของ ICC ณ ตำแหน่ง b_i

c_i คือ ค่าพารามิเตอร์โอกาสเดาข้อสอบได้ถูก

D คือ ค่าองค์ประกอบสเกล = 1.70

e = 2.718

<u>รูปแบบโลจิสติก 1 พารามิเตอร์</u> (one-parameter logistic model)

รูปแบบนี้บางครั้งเรียกว่า Rasch Model โดยโค้งลักษณะข้อสอบที่เขียนด้วยฟังก์ชั่นโลจิส ดังนี้

$$P_i(\theta) = \frac{1}{1+e^{-(\theta-b_i)}}$$

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นของผู้สอบที่มีความสามารถ θ ตอบข้อสอบข้อที่ i ถูกต้อง

b_i คือ ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i ซึ่งเป็นค่าที่แสดงตำแหน่งของ ICC ณ จุด θ ที่มีโอกาสตอบถูก 0.5

e = 2.718

การเลือกใช้การปรับเทียบคะแนนระหว่างแบบสอบโดยทฤษฎีการตอบสนองข้อสอบให้สอดคล้องกับรูปแบบการเก็บรวบรวมข้อมูลและการวิเคราะห์ รูปแบบ 1 พารามิเตอร์เหมาะสมกับกรณีที่กลุ่มตัวอย่างมีขนาดใหญ่ มีความเหมาะสมกับการปรับเทียบคะแนนในแนวนอนรูปแบบ 2 พารามิเตอร์ หรือ 3 พารามิเตอร์ เหมาะสมกับการปรับเทียบแนวตั้ง

2.2.4.5 การกำหนดสเกลสำหรับรายงานผล

การรายงานผลคะแนนที่ปรับเทียบแล้ว สามารถเลือกมาตรสำหรับการรายงานเป็นค่าความสามารถ (θ) คะแนนจริง (ξ) และคะแนนดิบ (r)

ศิริชัย กาญจนวาสี (2555) กล่าวว่า ถ้าต้องการรายงานผลการปรับเทียบคะแนนความสามารถ (θ) จากสมการการปรับเทียบ θ ดังกล่าว เมื่อคำนวณผลตามสมการ จึงสามารถสร้างตารางรายงานผลคะแนนสมมูลเปรียบเทียบกันได้เลย แต่ในกรณีของการรายงานผลการปรับเทียบคะแนนจริง (ξ) ที่สมมูลกับตำแหน่ง θ เดียวกัน จำเป็นต้องสร้างตารางปรับเทียบ หรือพล็อตกราฟที่

ทำให้ได้คะแนนจริงที่สมมูลกัน ส่วนกรณีของการรายงานผลการปรับเทียบด้วยคะแนนดิบหรือคะแนนที่สังเกตได้ (r) จำเป็นต้องมีการคำนวณเพิ่มเติม

โดย Hambleton และคณะ (1985) ได้เสนอการกำหนดมาตรสำหรับรายงานผลดังนี้

1) การปรับเทียบคะแนนจริง (true-score equating)

การปรับเทียบคะแนนจริงเป็นแปลงค่าความสามารถ (θ) ไปยังคะแนนจริง โดยการปรับเทียบแบบดังกล่าว มีวิธีการดำเนินการดังนี้

ให้ θ_x คือ ความสามารถของผู้สอบที่ทำแบบสอบฉบับ X

ξ_x คือ คะแนนจริงของผู้สอบที่ทำแบบสอบฉบับ X

นั่นคือ $\xi_X = \sum_{i=0}^{n} P_i(\alpha\theta_x + \beta)$

ในทำนองเดียวกัน ถ้า θ_y คือ ความสามารถของผู้สอบที่ทำแบบสอบฉบับ Y และ ξ_y คือ คะแนนจริงของผู้สอบที่ทำแบบสอบฉบับ Y จะได้ว่า

$$\xi_y = \sum_{j=1}^{m} P_j(\theta_y) = \sum_{j=1}^{m} P_j(\alpha\theta_x + \beta)$$

จะเห็นว่า $\theta_y = \alpha\theta_x + \beta$ นั่นคือ θ_yและ θ_xมีความสัมพันธ์เชิงเส้นตรง

ค่า θ_x แต่ละค่าสามารถแปลงเป็นคะแนนจริงที่สมมูลกับ θ_y ได้คือ (ξ_y, ξ_X) จากการคำนวณค่าสมมูลเป็นรายคู่ จึงสามารถทำเป็นตารางปรับเทียบคะแนนระหว่างแบบสอบทั้งสองฉบับ หรือพล็อตกราฟ ซึ่งกราฟความสัมพันธ์ระหว่างค่าความสามารถ θ_x และ θ_yที่สมมูลกัน เมื่อทราบค่า θ_xและ θ_y ที่สมมูลกันแล้ว ก็สามารถนำมาแปลงเป็น ξ_x และ ξ_y ได้ โดยกราฟความสัมพันธ์ระหว่าง ξ_x และ ξ_y ไม่เป็นเส้นตรง ซึ่งเป็นผลมาจากความคลาดเคลื่อนในการประมาณค่า θ_x และ θ_y เนื่องจากความคลาดเคลื่อนของการสุ่ม ซึ่งมีขนาดใหญ่มากขึ้นสำหรับค่าที่อยู่สุดขั้วของพารามิเตอร์ข้อสอบ และพารามิเตอร์ความสามารถของผู้สอบ ดังนั้นการปรับเทียบคะแนนจริงตรงคะแนนปลายขั้ว โดยใช้เส้นกราฟควรใช้วิธีการแก้ของ Angoff (1982) หรือหลีกเลี่ยงการหาความสัมพันธ์ระหว่าง ξ_X และ ξ_y

2) การปรับเทียบคะแนนที่สังเกตได้ (observed-score equating)

การปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ สามารถทำโดยการทราบค่าพารามิเตอร์ความสามารถ (θ) หรือคะแนนจริง (ξ) คำถามที่เกิดขึ้น คือ สามารถปรับเทียบคะแนนที่ได้จากการสอบหรือไม่ ในกรณีที่ต้องการปรับเทียบคะแนนระหว่างคะแนนที่สังเกตได้หรือคะแนนดิบ โดยใช้ทฤษฎีการตอบสนองข้อสอบสามารถกระทำได้เช่นกัน (Hambleton และคณะ, 1985)

คะแนนจริง (ξ) อยู่บนสเกลเดียวกับคะแนนที่สังเกตได้ (r) เมื่อ $r = \sum_{i=1}^{n} U_i$

ถ้าทฤษฎีการตอบสนองข้อสอบมีความสมเหตุสมผล จะได้ว่า $E(r) = \xi$ ดังนั้นสิ่งที่ต้องทำมีดังนี้ 1) หาความสัมพันธ์ระหว่างคะแนนจริง ξ_x และ ξ_y จากแบบสอบทั้งสองฉบับ 2) ให้การกระทำกับความสัมพันธ์เหล่านี้เหมือนกับเป็นความสัมพันธ์ระหว่างคะแนนที่สังเกตได้ r_x และ r_y และปรับเทียบคะแนนที่สังเกตได้

Lord (1980 cited in Hambleton & Swaminathan, 1985) กล่าวว่า ความสัมพันธ์ระหว่างคะแนนจริง ξ_x และ ξ_y ไม่จำเป็นต้องเหมือนกับความสัมพันธ์ระหว่างคะแนนที่สังเกตได้ r_x และ r_y เหตุการณ์นี้สามารถอธิบายได้จากคะแนนที่สังเกตได้ในโมเดล 3 พารามิเตอร์ กล่าวคือ $\xi_X \geq \sum_{i=1}^{n} c_i$ และ $\xi_y \geq \sum_{i=1}^{m} c_i$ โดยที่คะแนนที่สังเกตได้ r_x และ r_y อาจเป็นศูนย์ ดังนั้นการปรับเทียบโดยใช้คะแนนจริงจึงไม่เหมาะกับผู้สอบที่มีคะแนนที่สังเกตได้ต่ำกว่าระดับการเดา เพื่อหลีกเลี่ยงปัญหาดังกล่าวอาจนำสูตรการปรับเทียบคะแนนมาใช้ แต่อย่างไรก็ตามถ้ากล่าวกันตามหลักการแล้ว ความสัมพันธ์ระหว่างคะแนนจริง ξ_xและ ξ_y ไม่ควรใช้ในการปรับเทียบคะแนนที่สังเกตได้ r_x และ r_y

ทฤษฎีการตอบสนองข้อสอบให้วิธีในการทำนายการแจกแจงคะแนนที่สังเกตได้ของแบบสอบ เมื่อทราบลักษณะการแจกแจงคะแนนที่สังเกตได้ของแบบสอบ X และ Y ก็สามารถใช้การปรับเทียบอีควิเปอร์เซ็นไทล์ จุดเด่นของการใช้การปรับเทียบโดยใช้ทฤษฎีการตอบสนองข้อสอบ คือ แบบสอบที่ใช้ในการปรับเทียบไม่จำเป็นต้องความยากง่ายใกล้เคียงกัน ในขณะที่การปรับเทียบโดยใช้ทฤษฎีการทดสอบแบบดั้งเดิมจำเป็นต้องใช้แบบสอบที่มีความยากง่ายใกล้เคียงกัน

ตามทฤษฎีการแจกแจงความถี่ของคะแนนที่สังเกตได้ $f(r|\theta)$ ของแบบสอบสามารถกำหนดจากแบบสอบเฉพาะที่ต้องการปรับเทียบ

$$\sum_{r=0}^{n} f(r|\theta) = \prod_{i=1}^{n} [Q_i(\theta) + tP_i(\theta)]$$

ขั้นตอนการปรับเทียบคะแนนที่สังเกตได้โดยใช้ทฤษฎีการตอบสนองข้อสอบ มีดังนี้ ศิริชัย กาญจนวาสี (2555)

1. จากกลุ่มผู้สอบแบบฉบับ X นำข้อมูลมาวิเคราะห์เพื่อประมาณค่าพารามิเตอร์ของข้อสอบและผู้สอบ แล้วนำค่า θ_x ของผู้สอบแต่ละคนมาคำนวณความถี่สัมพัทธ์อย่างมีเงื่อนไขทางทฤษฎีการแจกแจงคะแนนดิบ $f_x(r|\theta_x)$

2. แปลงค่า θ_x ให้สมมูลกับ θ_yด้วยสูตร $\xi_x = \alpha\theta_x + \beta$ แล้วนำ ξ_x ของกลุ่มผู้สอบมาคำนวณความถี่ของคะแนนที่สังเกตได้ทางทฤษฎี $f_x(r)$ จึงได้ตารางการแจกแจงความถี่ของคะแนนที่สังเกตได้ทางทฤษฎีของกลุ่มผู้สอบที่ทำแบบสอบฉบับ X

3. จากผู้สอบที่ทำแบบสอบฉบับ Y นำข้อมูลมาวิเคราะห์เพื่อประมาณค่าพารามิเตอร์ของข้อสอบและผู้สอบ แล้วนำค่า θ_y ของผู้สอบแต่ละคนมาคำนวณความถี่สัมพัทธ์อย่างมีเงื่อนไขทางทฤษฎีของการแจงแจงคะแนนที่สังเกตได้ $f_y(r/\theta_y)$

4. นำค่า θ_y ของผู้สอบมาคำนวณความถี่ของคะแนนที่สังเกตได้ทางทฤษฎี $f_y(r)$ จึงได้ตารางการแจกแจงความถี่ของคะแนนที่สังเกตได้ทางทฤษฎีของผู้สอบที่ทำแบบสอบ Y

5. ปรับเทียบคะแนนระหว่างคะแนนที่สังเกตได้ระหว่างแบบสอบ X และ Y (จากข้อ 2 และ 4) โดยใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์

2.2.4.6 กระบวนการในการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ

Hambleton และคณะ (1985) ได้เสนอกระบวนการในการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ มีขั้นตอนดังนี้

1. เลือกรูปแบบการปรับเทียบคะแนน (equating designs) ที่เหมาะสม โดยพิจารณาจากธรรมชาติของแบบสอบ และกลุ่มของผู้สอบ ซึ่งรูปแบบการปรับเทียบอาจเป็นรูปแบบผู้สอบกลุ่มเดียว (single-group design) รูปแบบผู้สอบกลุ่มเท่าเทียมกัน (equivalent-group design) รูปแบบผู้สอบต่างกลุ่มกันโดยใช้แบบสอบร่วม (anchor-test design) และรูปแบบผู้สอบต่างกลุ่มกันโดยใช้ผู้สอบร่วม (common-person design)

2. กำหนดโมเดลการตอบสนองข้อสอบ (item response model) ในกรณีของการปรับเทียบคะแนนแบบแนวตั้ง โมเดล 1 พารามิเตอร์ไม่เหมาะสมในการนำมาใช้ เพราะว่าโครงสร้างของแบบสอบไม่เหมาะสมกับโมเดล แต่ถ้าเป็นการปรับเทียบคะแนนแบบแนวนอนที่มีขนาดของกลุ่มตัวอย่างไม่ใหญ่มากเหมาะสมกับโมเดล 1 พารามิเตอร์ แต่ถ้ากลุ่มตัวอย่างมีขนาดใหญ่ควรใช้โมเดล 2 พารามิเตอร์ หรือ 3 พารามิเตอร์ ซึ่งมีความเหมาะสมมากกว่า

3. สร้างเมทริกซ์ร่วมของความสามารถและค่าพารามิเตอร์ของข้อสอบ เพราะว่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบมีความสัมพันธ์กันเชิงเส้นตรง ดังนั้นจึงสามารถสร้างเมทริกซ์ร่วมได้ โดยการคำนวณค่าคงที่ในสมการสำหรับการเปรียบเทียบค่าพารามิเตอร์ของข้อสอบให้อยู่บนสเกลเดียวกัน เพื่อปรับเทียบ θ ให้อยู่บนสเกลที่สามารถเปรียบเทียบได้โดยตรง

4. กำหนดสเกลในการรายงานคะแนนสอบ:

4.1 ถ้าต้องการรายงานผลการเปรียบเทียบด้วยคะแนนความสามารถ (θ) จากสมการปรับเทียบความสามารถ เมื่อคำนวณผลตามสมการแล้วก็สามารถสร้างตารางผลการเปรียบเทียบกันได้เลย

4.2 ถ้าต้องการรายงานผลการเปรียบเทียบด้วยคะแนนจริง (ξ) สามารถคำนวณคะแนนจริงที่สมมูลกันตรงตำแหน่งความสามารถต่าง ๆ ได้ เพื่อนำมาสร้างตารางปรับเทียบ หรือพล็อตเป็นกราฟ เพื่อได้ค่าคะแนนที่สมมูลกัน

4.3 ถ้าต้องการรายงานผลการเปรียบเทียบด้วยคะแนนที่สังเกตได้ มีขั้นตอนดังนี้

- คำนวณความถี่ของการแจกแจงคะแนนที่สังเกตได้อย่างมีเงื่อนไขตามทฤษฎีจากกลุ่มตัวอย่าง
- คำนวณความถี่รวมของการแจกแจงคะแนนที่สังเกตได้ตามทฤษฎีของแบบสอบแต่ละฉบับ
- คำนวณค่าเปอร์เซ็นไทล์ของคะแนนจากแบบสอบแต่ละฉบับ โดยปรับเทียบจากหลักการของอีควิเปอร์เซ็นไทล์
- สร้างตารางปรับเทียบ หรือพล็อตกราฟที่ปรับเทียบคะแนนที่สังเกตได้ที่สมมูลกันจากแบบสอบแต่ละฉบับ

CHULALONGKORN UNIVERSITY

2.2.5 สถานการณ์ที่เหมาะสมกับวิธีการปรับเทียบคะแนนระหว่างแบบสอบแบบต่าง ๆ

ตาราง 2 สถานการณ์ที่เหมาะสมกับวิธีการปรับเทียบคะแนนระหว่างแบบสอบแบบต่าง ๆ (Kolen & Brenan, 2004)

วิธีการ	สถานการณ์ที่เหมาะสม
วิธีการปรับเทียบค่าเฉลี่ย (mean)	**รูปแบบผู้สอบที่ได้มาจากการสุ่ม(random groups designs) และรูปแบบผู้สอบต่างกลุ่มโดยใช้แบบสอบร่วม (common-item nonequivalent groups designs)** 1. การควบคุมคุณภาพของการปรับเทียบและสภาพมาตรฐานทำได้ในระดับปานกลาง 2. กลุ่มผู้สอบขนาดเล็กมาก 3. ความยากของแบบสอบแต่ละฉบับใกล้เคียงกัน 4. ตารางการปรับเทียบคะแนนหรือสมการที่ใช้ในการปรับเทียบคะแนนเข้าใจง่ายไม่ซับซ้อนมากนัก 5. ความแม่นยำ (accuracy) ของผลการปรับเทียบต้องเข้าใกล้ค่าเฉลี่ยจึงจะยอมรับได้ **รูปแบบผู้สอบต่างกลุ่มโดยใช้แบบสอบร่วม (common-item nonequivalent groups designs)** 6. ข้อตกลงเบื้องต้นใช้กับกลุ่มผู้สอบที่ไม่มีความหลากหลายมาก กล่าวคือ ข้อสอบร่วมต้องเป็นตัวแทนของแบบสอบ และผู้สอบต้องมีระดับผลสัมฤทธิ์ทางการเรียนไม่แตกต่างกันมาก
2. วิธีการปรับเทียบเชิงเส้นตรง (linear)	**รูปแบบผู้สอบที่ได้มาจากการสุ่ม (random groups designs) และรูปแบบผู้สอบต่างกลุ่มโดยใช้แบบสอบร่วม (common-item nonequivalent groups designs)** 1. การควบคุมคุณภาพของการปรับเทียบและสภาพมาตรฐานทำได้ในระดับปานกลาง แบบสอบทางเลือกที่ใช้ควรสร้างให้มีรายละเอียดเหมือนกัน 2. กลุ่มผู้สอบขนาดเล็ก 3. ความยากของแบบสอบแต่ละฉบับใกล้เคียงกัน 4. ตารางการปรับเทียบคะแนนหรือสมการที่ใช้ในการปรับเทียบคะแนนเข้าใจง่ายไม่ซับซ้อนมากนัก 5. ความแม่นยำ (accuracy) ของผลการปรับเทียบต้องเข้าใกล้ค่าเฉลี่ยจึงจะยอมรับได้ **รูปแบบผู้สอบต่างกลุ่มโดยใช้แบบสอบร่วม (common-item nonequivalent groups designs)** 6. ข้อตกลงเบื้องต้นใช้กับกลุ่มผู้สอบที่ไม่มีความหลากหลายมาก คือ ข้อสอบร่วมต้องเป็นตัวแทนของแบบสอบ และผู้สอบต้องมีระดับผลสัมฤทธิ์ทางการเรียนไม่แตกต่างกันมาก
3. วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์ (equipercentile)	**รูปแบบผู้สอบที่ได้มาจากการสุ่ม (random groups designs) และรูปแบบผู้สอบต่างกลุ่มโดยใช้แบบสอบร่วม (common-item nonequivalent groups designs)** 1. การควบคุมคุณภาพของการปรับเทียบและสภาพมาตรฐานทำได้ในระดับปานกลาง 2. กลุ่มผู้สอบมีขนาดใหญ่ 3. ความยากของแบบสอบแต่ละฉบับแตกต่างกันได้มากกว่าวิธีการปรับเทียบเชิงเส้นตรง 4. ตารางการปรับเทียบคะแนนหรือสมการที่ใช้ในการปรับเทียบคะแนนมีความซับซ้อน 5. ความแม่นยำ (accuracy) ของผลการปรับเทียบมีความสำคัญมากตลอดสเกลคะแนน

ตาราง 2 สถานการณ์ที่เหมาะสมกับวิธีการปรับเทียบคะแนนระหว่างแบบสอบแบบต่าง ๆ (ต่อ) (Kolen & Brenan, 2004)

วิธีการ	สถานการณ์ที่เหมาะสม
3. วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์ (equipercentile) (ต่อ)	**รูปแบบผู้สอบต่างกลุ่มโดยใช้แบบสอบร่วม (common-item nonequivalent groups designs)** 6. ข้อตกลงเบื้องต้นใช้กับกลุ่มผู้สอบที่ไม่มีความหลากหลายมาก คือ ข้อสอบร่วมต้องเป็นตัวแทนของแบบสอบ และผู้สอบต้องมีระดับผลสัมฤทธิ์ทางการเรียนไม่แตกต่างกันมาก
4. วิธีการปรับเทียบโดยใช้ทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (three-parameter IRT)	**รูปแบบผู้สอบที่ได้มาจากการสุ่ม(random groups designs) และรูปแบบผู้สอบต่างกลุ่มโดยใช้แบบสอบร่วม (common-item nonequivalent groups designs)** 1. การควบคุมคุณภาพของการปรับเทียบและสภาพมาตรฐานทำได้ในระดับปานกลาง 2. กลุ่มผู้สอบมีขนาดใหญ่ 3. ความยากของแบบสอบแต่ละฉบับแตกต่างกันได้มากกว่าวิธีการปรับเทียบเชิงเส้นตรง 4. ตารางการปรับเทียบคะแนนหรือสมการที่ใช้ในการปรับเทียบคะแนนมีความซับซ้อน 5. ความแม่นยำ (accuracy) ของผลการปรับเทียบมีความสำคัญมากตลอดสเกลคะแนน 6. ข้อตกลงเบื้องต้นที่ใช้ในวิธีการนี้ต้องมีความสมเหตุสมผล **รูปแบบผู้สอบต่างกลุ่มโดยใช้แบบสอบร่วม (common-item nonequivalent groups designs)** 7. ข้อตกลงเบื้องต้นใช้กับกลุ่มผู้สอบที่ไม่มีความหลากหลายมาก คือ ข้อสอบร่วมต้องเป็นตัวแทนของแบบสอบ และผู้สอบต้องมีระดับผลสัมฤทธิ์ทางการเรียนไม่แตกต่างกันมาก

2.2.6 งานวิจัยที่เกี่ยวข้องกับการออกแบบวิธีการปรับเทียบคะแนน

1) งานวิจัยที่เกี่ยวข้องกับการออกแบบวิธีการปรับเทียบเชิงเส้นตรง (linear equating)

Suh และคณะ (2009) ได้เปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของวิธีการปรับเทียบคะแนนเชิงเส้นตรง 5 วิธี ได้แก่ วิธีของ Tucker (Tucker method) วิธี Tucker-like วิธี Levine observe score วิธี Levine true score และวิธี CLE (Chain Linear Equating) สำหรับรูปแบบผู้สอบกลุ่มไม่เท่าเทียมกัน โดยใช้ข้อสอบร่วม (NEAT) ข้อมูลที่ใช้ในการวิจัยครั้งนี้มาจากการทดสอบ Multistate Bar Examination (MBE) ซึ่งประกอบด้วยข้อสอบจำนวน 200 ข้อ แต่คำตอบที่นำมาใช้จริงมาจากข้อสอบจำนวน 90 ข้อ โดยข้อสอบจำนวน 90 ข้อใช้เป็นตัวแบบเพื่อสร้างข้อสอบจำนวน 50 ข้อและข้อสอบร่วมภายในจำนวน 10 ข้อ สำหรับรูปแบบกลุ่มไม่เท่าเทียมกัน ข้อมูลจากชุดข้อมูลนี้ใช้เพื่อประเมินสถิติบูทสแตรป (bootstrapped statistics) ของวิธีการปรับเทียบเชิงเส้นทั้ง 5 วิธี ส่วนข้อมูลฐานที่ประกอบด้วยคำตอบของผู้สอบจำนวน 40,000 คนจากข้อสอบจำนวน 90 ข้อ ใช้กับรูปแบบผู้สอบกลุ่มเดียว เพื่อใช้เป็นเกณฑ์ในการประเมินวิธีการปรับเทียบ แต่อย่างไรก็ตามผู้สอบทุกคนทำข้อสอบทั้งรูปแบบ X และY (รวมไปถึงข้อสอบร่วม) โดยสถิติบูทสแตรป (bootstrapped statistics) ของวิธีการปรับเทียบพิจารณาจากค่าความลำเอียง (bias) ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) และ ค่า $\mathrm{SE}_{\mathrm{slope}}$

ผลการวิจัยพบว่าถ้าแต่ละกลุ่มตัวอย่างมีความแตกต่างกัน วิธีของเลอวิน (Levine method) มีค่าความลำเอียง (bias) และค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ต่ำกว่า

วิธีอื่น ๆ แต่ถ้ากลุ่มตัวอย่างมีลักษณะใกล้เคียงกัน วิธีการปรับเทียบเชิงเส้นทั้ง 5 มีคุณภาพดี ดังนั้น วิธีของเลอวิน (Levine method) จึงมีความเหมาะสมที่สุด เมื่อแต่ละกลุ่มมีความแตกต่างกัน และวิธีการดังกล่าวมีความถูกต้องมากขึ้นเมื่อกลุ่มตัวอย่างมีขนาดใหญ่

Mroch และคณะ (2009) ได้ใช้ผลการศึกษาครั้งที่ผ่านมาของ Suh และคณะ (2009) และจากการทบทวนงานวิจัยที่เกี่ยวข้องกับวิธีการปรับเทียบเชิงเส้นตรง มาใช้ในการประเมินวิธีการปรับเทียบเชิงเส้นตรงทั้ง 5 วิธี ได้แก่ วิธี Levine observed score วิธี Levine true score วิธีของ Tucker (Tucker method) วิธี Tucker-like และวิธี CLE (Chained Linear Equating) ในหลากหลายมิติ ตามระดับความแตกต่างของกลุ่ม (group difference) รวมไปถึงข้อตกลงเบื้องต้นของวิธีการดังกล่าว และระดับความลำเอียงและค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีการปรับเชิงเส้นในแต่ละวิธี โดยวิธีการทั้งหมดนี้ใช้กับการออกแบบการเก็บข้อมูลรูปแบบกลุ่มไม่เท่าเทียมกัน โดยใช้ข้อสอบร่วม (NEAT)

ผลการวิจัยพบว่า เมื่อแต่ละกลุ่มมีความคล้ายคลึงกัน วิธีการทั้ง 5 วิธีมีความเหมาะสม แต่ถ้าแต่ละกลุ่มมีความสามารถแตกต่างกันมาก วิธี Levine observed score และวิธี Levine true score มีความเหมาะสมมากกว่าวิธีของ Tucker (Tucker method) วิธี Tucker-like และวิธี CLE (Chained Linear Equating) เนื่องจากค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีของ Tucker (Tucker method) วิธี Tucker-like และวิธี CLE (Chained Linear Equating) มีค่าสูงขึ้น เมื่อกลุ่มตัวอย่างมีความแตกต่างกันมากขึ้น ขณะที่ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธี Levine observed score และวิธี Levine true score มีค่าต่ำและมีค่าค่อนข้างคงที่ แม้ว่ากลุ่มตัวอย่างมีความแตกต่างกันมากขึ้น ซึ่งโดยทั่วไป วิธีของเลอวิน (Levine method) มีความลำเอียง (bias) และค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ต่ำกว่า 3 วิธีที่เหลือ กล่าวคือ วิธีของเลอวิน (Levine method) เป็นตัวเลือกที่เหมาะสมที่สุดสำหรับวิธีการปรับเทียบเชิงเส้นตรง เมื่อกลุ่มตัวอย่างมีความแตกต่างกัน แต่อย่างไรก็ตามค่าความแม่นยำ (accuracy) ของวิธีของเลอวิน (Levine method) ไม่ค่อยเสถียร เนื่องจากน้ำหนักความแตกต่างของกลุ่มมีค่ามากเกินไป และมีผลกระทบมากขึ้นเมื่อกลุ่มตัวอย่างมีความสามารถในทำข้อสอบร่วมแตกต่างกันเล็กน้อย นั่นแสดงให้เห็นว่าวิธีของเลอวิน (Levine method) มีความแม่นยำสูง แต่มีความเสถียรต่ำ ข้อเสนอแนะทั่วไป กล่าวว่า การใช้วิธีของ Tucker (Tucker method) เหมาะกับกลุ่มตัวอย่างที่ไม่แตกต่างกัน และวิธีของเลอวิน (Levine method) เหมาะกับกลุ่มตัวอย่างที่แตกต่างกัน ซึ่งการศึกษาครั้งนี้ก็สนับสนุนข้อเสนอแนะดังกล่าว ดังนั้นจึงแนะนำให้ใช้วิธี Levine observed score และ วิธี Levine true score เพราะวิธีของเลอวิน (Levine method) มีความลำเอียงต่ำกว่าเมื่อแต่ละกลุ่มมีความแตกต่างกัน

Puhan (2010) ได้ศึกษาผลลัพธ์ของวิธีการปรับเทียบคะแนน 3 วิธี ได้แก่ วิธี CLE (Chained Linear Equating), วิธีของ Tucker (Tucker method) และวิธี Levine observed score ภายใต้สถานการณ์ที่กลุ่มตัวอย่างทำแบบสอบฉบับใหม่และฉบับเก่าที่มีความสามารถใกล้เคียงกัน และมีความสามารถแตกต่างกัน นอกจากนี้ยังศึกษาถึงผลกระทบของแบบสอบร่วมด้วยวิธีของการปรับเทียบคะแนนแบบต่าง ๆ โดยเกณฑ์ที่ใช้ในการเปรียบเทียบวิธีการปรับเทียบคะแนนทั้ง 3 แบบ ได้แก่ ความคลาดเคลื่อนของการปรับเทียบอย่างสุ่ม (random equating error) ความลำเอียง (bias)

และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ข้อมูลที่ใช้ในการวิจัยครั้งนี้ คือ คำตอบของผู้สอบจากแบบสอบแบบหลายตัวเลือกสองชุด คือ แบบสอบ X และแบบสอบ Y โดยแบบสอบ X วัดความรู้ในระดับประถมศึกษา ส่วนแบบสอบ Y วัดความรู้ทางคณิตศาสตร์ จุดประสงค์ของการวิจัยครั้งเพื่อเปรียบเทียบฟังก์ชั่นการปรับเทียบคะแนนที่แตกต่างกันที่ได้จากแบบสอบแฝง 2 ฉบับของแบบสอบ X (อ้างไปยังแบบสอบ X1 และ X2) และแบบสอบ Y (อ้างไปยังแบบสอบ Y1) ซึ่งถูกปรับเทียบภายในตัวเอง (เช่น แบบสอบ Y1) โดยการปรับเทียบภายในตัวเองช่วยในการเลียนแบบสถานการณ์ที่แบบสอบชุดใหม่และชุดเก่ามีความยากใกล้เคียงกันมาก

ผลการวิจัยพบว่า วิธี CLE (Chained Linear Equating) ให้ผลการปรับเทียบคะแนนที่ดี เนื่องจากมีความลำเอียงต่ำ และมีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำ วิธีการปรับเทียบคะแนนแบบ Levine observed score มีความเอียงต่ำ และมีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำในบางกรณี วิธีของ Tucker (Tucker method) มีความคลาดเคลื่อนของการปรับเทียบอย่างสุ่มต่ำ แต่มีค่าความลำเอียง (bias) สูง และมีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) สูงกว่าวิธีอื่นๆ ดังนั้นวิธี CLE (Chained Linear Equating) หรือวิธี Levine observed score สามารถนำมาใช้กับกลุ่มตัวอย่างใหม่และกลุ่มเก่าที่มีความสามารถแตกต่างกัน หรือเมื่อข้อสอบร่วมกับข้อสอบทั้งหมดมีความสัมพันธ์กันสูงได้

Puhan (2012) ได้ประเมินวิธีการปรับเทียบคะแนนด้วยวิธีของ Tucker (Tucker method) และวิธี CLE (Chained Linear Equating) ใน 2 สถานการณ์ คือ เมื่อมีผู้ตรวจให้คะแนนหลายคน และเมื่อกลุ่มผู้สอบเท่าเทียมกันอย่างสุ่ม และใช้ข้อสอบร่วม รายละเอียดของสถานการณ์ที่ 1 คือ การให้คะแนนโดยผู้ตรวจหลายคนในแบบสอบฉบับใหม่มีค่าสหสัมพันธ์ระหว่างคะแนนข้อสอบร่วมกับคะแนนทั้งหมดสูง และมีค่าสหสัมพันธ์ปานกลางในแบบสอบฉบับอ้างอิง ซึ่งสถานการณ์นี้อาจส่งผลกระทบทางลบต่อการปรับเทียบคะแนนด้วยวิธีของ Tucker (Tucker method) โดยเฉพาะอย่างยิ่งถ้ากลุ่มตัวอย่างกลุ่มใหม่และกลุ่มอ้างอิงมีความสามารถแตกต่างกัน รายละเอียดของสถานการณ์ที่ 2 คือ กลุ่มตัวอย่างกลุ่มใหม่และกลุ่มอ้างอิงมีความเท่าเทียมกันอย่างสุ่ม แต่ความสัมพันธ์ระหว่างคะแนนข้อสอบร่วมกับคะแนนทั้งหมดอยู่ในระดับต่ำ และเมื่อความสัมพันธ์ระหว่างคะแนนข้อสอบร่วมและคะแนนทั้งหมดมีค่าต่ำ วิธีของ Tucker (Tucker method) จึงยอมรับว่ากลุ่มตัวอย่างกลุ่มใหม่และกลุ่มอ้างอิงมีความสามารถใกล้เคียงกัน ดังนั้นวิธีการปรับเทียบดังกล่าวน่าจะให้ผลที่แม่นยำ

ในสถานการณ์ที่ 1 เก็บข้อมูลจากนักเรียนระดับชั้นประถมศึกษา โดยใช้แบบสอบแบบสร้างคำตอบจำนวน 4 ข้อ การปรับเทียบที่ใช้เป็นเกณฑ์ในการเปรียบเทียบ คือ การปรับเทียบเชิงเส้นของกลุ่มตัวอย่างกลุ่มเดียว (single group linear equating) การคำนวณความแม่นยำและความผันแปรในสถานการณ์ที่ 1 มีขั้นตอนดังนี้ 1) ใช้วิธีของ Tucker (Tucker method) และวิธี CLE (Chained Linear Equating) ปรับเทียบข้อสอบร่วมกับกลุ่มตัวอย่างกลุ่มใหม่จำนวน 794 คน และกลุ่มอ้างอิงจำนวน 452 คน 2) คำนวณหาความคลาดเคลื่อนมาตรฐานอย่างมีเงื่อนไขในการปรับเทียบคะแนน (conditional standard error of equating) ค่าเฉลี่ยน้ำหนักของความคลาดเคลื่อนมาตรฐานอย่างมีเงื่อนไขในการปรับเทียบคะแนน (average SEE) ความลำเอียงอย่างมีเงื่อนไข (conditional bias) ค่าเฉลี่ยน้ำหนักของความลำเอียง (weighted average bias) และค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ส่วนในสถานการณ์ที่ 2 ใช้ข้อมูลจำลองจากข้อมูลจริง โดยข้อมูล

จำลองนี้ถูกสร้างจากแบบสอบ A และ B ที่ประกอบด้วยข้อสอบ จำนวน 110 ข้อ และข้อสอบร่วม จำนวน 52 ข้อโดยข้อสอบชุด A ใช้สร้างข้อสอบชุด A1 จำนวน 50 ข้อ และข้อสอบชุด B ใช้สร้างข้อสอบชุด B1 จำนวน 50 ข้อ การปรับเทียบที่ใช้เป็นเกณฑ์ในกรณีนี้ คือ การปรับเทียบเชิงเส้นตรง (linear equating) จากแบบสอบ A1 ไปยังแบบสอบ B1 ภายใต้รูปแบบกลุ่มผู้สอบมีความเท่าเทียมกันอย่างสุ่ม การปรับเทียบด้วยวิธีของ Tucker (Tucker method) และวิธี CLE (Chained Linear Equating) ใช้ข้อสอบร่วมภายนอกจำนวน 20 ข้อ มาเปรียบเทียบกับเกณฑ์ดังกล่าว ในสถานการณ์ที่ 1 มีขั้นตอนดังนี้ 1) ใช้วิธีของ Tucker (Tucker method) และวิธี CLE (Chained Linear Equating) ปรับเทียบข้อสอบร่วมกับกลุ่มตัวอย่างกลุ่มใหม่จำนวน 3,968 คน และกลุ่มอ้างอิงจำนวน 3,909 คน 2) ทำเช่นเดียวกับข้อ 2 ในสถานการณ์ที่ 1

ผลการวิจัยพบว่า ในสถานการณ์ที่ 1 เมื่อมีผู้ตรวจให้คะแนนหลายคน วิธีของ Tucker (Tucker method) มีความแม่นยำต่ำกว่าวิธี CLE (Chained Linear Equating) แต่ในสถานการณ์ที่ 2 กลุ่มตัวอย่างกลุ่มใหม่และกลุ่มอ้างอิงมีความเท่าเทียมกันอย่างสุ่มและใช้ข้อสอบร่วม วิธีของ Tucker (Tucker method) มีความแม่นยำมากกว่าวิธี CLE (Chained Linear Equating)

บทสรุปงานวิจัยเกี่ยวกับวิธีการปรับเทียบคะแนนเชิงเส้นตรงสำหรับแบบสอบรูปแบบผสม

จากการที่ผู้วิจัยได้ทบทวนเอกสารและงานวิจัยที่เกี่ยวข้อง ผู้วิจัยได้สรุปการศึกษาการปรับเทียบคะแนนเชิงเส้นตรงสำหรับแบบสอบรูปแบบผสม ดังนี้

การปรับเทียบคะแนนระหว่างแบบสอบโดยใช้ทฤษฎีดั้งเดิมด้วยวิธีการปรับเทียบคะแนนเชิงเส้นตรง เป็นการแปลงคะแนนจากแบบสอบ X ให้สมมูลกับแบบสอบ Y โดยให้คะแนนแปลงของ X มีค่าเฉลี่ยเป็น $\bar{Y}$ และส่วนเบี่ยงเบนมาตรฐานเป็น S_y โดยวิธีการปรับเทียบที่ผู้วิจัยได้นำเสนอไว้ใน บทที่ 2 ได้แก่ วิธีของ Tucker (Tucker method) วิธี Levine observed method วิธี Levine true method และวิธี CLE (Chained Linear Equating) ซึ่งวิธีการปรับเทียบเชิงเส้นตรงมีความเหมาะสมกับกลุ่มผู้สอบขนาดเล็ก โดยวิธีการปรับเทียบเชิงเส้น สำหรับรูปแบบผู้สอบที่ได้มาจากการสุ่ม (random groups designs) และรูปแบบผู้สอบกลุ่มไม่เท่าเทียมโดยใช้แบบสอบร่วม (common-item nonequivalent groups designs) เหมาะสมกับกลุ่มผู้สอบขนาดเล็ก ความยากของแบบสอบแต่ละฉบับใกล้เคียงกัน และ ความแม่นยำ (accuracy) ของผลการปรับเทียบต้องเข้าใกล้ค่าเฉลี่ยจึงจะยอมรับได้ แต่ถ้านำวิธีการปรับเทียบคะแนนเชิงเส้นตรงไปใช้กับรูปแบบผู้สอบกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม ข้อสอบร่วมต้องเป็นตัวแทนของแบบสอบ และผู้สอบต้องมีระดับผลสัมฤทธิ์ทางการเรียนไม่แตกต่างกันมาก (Kolen & Brennan, 2004)

วิธีของเลอวิน (Levine method) เป็นวิธีในการปรับเทียบคะแนนที่สังเกตได้ เพื่อเชื่อมคะแนนที่สังเกตได้ของแบบสอบ X ไปยังสเกลของคะแนนที่สังเกตได้ของแบบสอบ Y โดยวิธีของ Levine มีข้อตกลงเบื้องต้นว่า แบบสอบ X แบบสอบ Y และแบบสอบร่วม V วัดในสิ่งเดียวกัน ความแปรปรวนของความคลาดเคลื่อนในการวัด (measurement error variance) สำหรับแบบสอบ X เหมือนกันทั้งในประชากร 1 และ 2 นอกจากนี้ T_X และ T_V ยังมีความสัมพันธ์กันเป็นเส้นตรง

(Kolen & Brennan, 2004) ซึ่งวิธีของเลอวิน (Levine method) เหมาะกับกลุ่มตัวอย่างที่แต่ละกลุ่มมีความแตกต่างกัน หรือเมื่อข้อสอบร่วมกับข้อสอบทั้งหมดมีความสัมพันธ์กันสูง โดยวิธีของเลอวิน (Levine method) มีค่าความลำเอียง (bias) และค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ต่ำกว่าวิธีของ Tucker (Tucker method) วิธี Tucker-like และวิธี CLE (Chained Linear Equating) เมื่อแต่ละกลุ่มมีความแตกต่างกัน และถ้าการออกแบบการเก็บข้อมูลเป็นแบบผู้สอบกลุ่มไม่เท่าเทียมกัน โดยใช้ข้อสอบร่วม (NEAT) ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) และค่าความลำเอียง (bias) ของวิธีของเลอวิน (Levine method) มีค่าต่ำ (Mroch et al., 2009; Puhan, 2010; Puhan, 2012) แต่สำหรับรูปแบบผู้สอบกลุ่มเท่าเทียมกันอย่างสุ่ม โดยผู้สอบสองกลุ่มมีความสามารถใกล้เคียงกัน วิธีของ Tucker (Tucker method) ดีกว่าเล็กน้อย (Puhan, 2012) กล่าวคือ วิธีของเลอวิน (Levine method) เป็นตัวเลือกที่เหมาะสมที่สุดในวิธีการปรับเทียบเชิงเส้นตรง เมื่อแต่ละกลุ่มตัวอย่างมีความสามารถแตกต่างกัน ซึ่งในสถานการณ์จริงเรามักเจอกับกลุ่มผู้สอบที่แต่ละกลุ่มมีความแตกต่างกัน (Kim et al., 2010) ดังนั้นวิธีของเลอวิน (Levine method) มีความเหมาะสมเมื่อนำไปใช้ในสภาพจริง

วิธี CLE (Chained Linear Equating) เหมาะสมกับกลุ่มตัวอย่างแต่ละกลุ่มที่มีความแตกต่างกัน หรือเมื่อข้อสอบร่วมกับข้อสอบทั้งหมดมีความสัมพันธ์กันสูง และถ้าการออกแบบการเก็บข้อมูลเป็นแบบผู้สอบกลุ่มไม่เท่าเทียม โดยใช้ข้อสอบร่วม (NEAT) ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) และค่าความลำเอียง (bias) มีค่าต่ำ แต่ไม่เหมาะกับรูปแบบผู้สอบกลุ่มเท่าเทียมกันอย่างสุ่ม โดยผู้สอบสองกลุ่มมีความคล้ายคลึงกัน (Suh et al., 2009; Puhan, 2010; Puhan, 2012) โดยวิธี CLE (Chained Linear Equating) ยังนำมาใช้งานวิจัยของ Kim และคณะ (2010) เพื่อปรับเทียบคะแนนแบบสอบรูปแบบผสมและแบบสอบแบบสร้างคำตอบ โดยผู้วิจัยให้เหตุผลว่า ฟังก์ชั่นที่ใช้เป็นเกณฑ์เป็นแบบเชิงเส้นตรง วิธีการนี้จึงมีความเหมาะสม ซึ่งวิธี CLE (Chained Linear Equating) ประกอบด้วยการปรับเทียบคะแนนแบบสอบฉบับใหม่ไปยังคะแนนของข้อสอบร่วม แล้วทำการปรับคะแนนจากข้อสอบร่วมไปยังแบบสอบฉบับอ้างอิง ซึ่งพบว่า มีการปรับเทียบแบบลูกโซ่เชื่อมต่อจากคะแนนแบบสอบฉบับใหม่ไปยังแบบสอบฉบับอ้างอิง ในส่วนของรูปแบบผู้สอบกลุ่มเท่าเทียม (EG) ก็ใช้วิธี CLE (Chained Linear Equating) ได้ เนื่องจากผลของการปรับเทียบไม่แตกต่างจากรูปแบบกลุ่มไม่เท่าเทียมกัน โดยใช้ข้อสอบร่วม เพราะว่า ฟังก์ชั่นมีความเท่าเทียมทั้งรูปแบบ NEAT และรูปแบบ EG นอกจากนี้ เมื่อกลุ่มตัวอย่างมีความสามารถไม่แตกต่างกันมากนัก วิธี CLE (Chained Linear Equating) และวิธี Poststratification (อย่างเช่น วิธีของ Tucker) ให้ผลที่ใกล้เคียงกัน แต่ถ้ากลุ่มผู้สอบมีความแตกต่างกันมาก วิธี CLE (Chained Linear Equating) มีความลำเอียงต่ำกว่า ซึ่งในสถานการณ์จริงเรามักเจอกับผู้สอบที่มีความแตกต่างกัน ดังนั้นวิธี CLE (Chained Linear Equating) มีความเหมาะสมในการนำมาใช้กับสภาพจริง (Kim et al., 2010)

นอกจากนี้ Kamata และคณะ (2005) กล่าวว่า วิธีการเชื่อมโยงโดยใช้ทฤษฎีการตอบสนองข้อสอบ (IRT) ไม่เหมาะสมในการนำมาใช้กับกลุ่มตัวอย่างขนาดเล็ก ในกรณีที่กลุ่มตัวอย่างมีขนาดเล็กนิยมใช้วิธีการปรับเทียบแบบดั้งเดิม เช่น วิธี CLE (Chained Linear Equating), วิธีของ Tucker (Tucker method), วิธีของเลอวิน (Levine method), วิธีการประมาณค่าความถี่ (frequently

estimate) หรือวิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ (chained equipercentile equating ซึ่งวิธีการเหล่านี้มักใช้เพื่อเชื่อมต่อแบบสอบรูปแบบผสม และจากการทบทวนวรรณกรรมที่เกี่ยวข้องยังไม่มีการศึกษาสถิติบูทสแตรป (bootstrapped statistics) ของการปรับเทียบคะแนนโดยใช้ทฤษฎี IRT เมื่อออกแบบให้มีการตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) และใช้ข้อมูลจากการทดสอบจริง (Kim, et al., 2010)

เนื่องจากงานวิจัยของผู้วิจัยเป็นการใช้ข้อมูลจริงที่ได้จากการทดสอบที่มีกลุ่มผู้สอบขนาดเล็ก ซึ่งกลุ่มตัวอย่างแต่ละกลุ่มมีความสามารถแตกต่างกัน วิธีการปรับเทียบเชิงเส้นตรงนั้นจึงมีความเหมาะสมกับการนำไปใช้ และเมื่อเปรียบเทียบวิธีการปรับเทียบคะแนนเชิงเส้นตรงทั้ง 3 วิธี พบว่าวิธีการที่ความเหมาะสมมากที่สุด 2 วิธี ได้แก่ วิธีของเลอวิน (Levine method) และวิธี CLE (Chained Linear Equating) ด้วยเหตุนี้ผู้วิจัยมีความประสงค์ในการเปรียบเทียบวิธีการปรับเทียบเชิงเส้นตรงระหว่างวิธีของเลอวิน (Levine method) (วิธี Levine observed score method) กับวิธี CLE (Chained Linear Equating) สำหรับแบบสอบรูปแบบผสม ภายใต้รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกัน โดยใช้ข้อสอบร่วมในที่นี้ ผู้วิจัยได้สรุปภาพรวมในการศึกษาวิธีการปรับเทียบคะแนนเชิงเส้นตรงสำหรับแบบสอบรูปแบบผสม ดังตาราง 3

ตาราง 3 สรุปประเด็นสำคัญของการออกแบบวิธีการปรับเทียบคะแนนเชิงเส้นตรง

ผู้วิจัย/ ปีการศึกษา	วิธีการ ปรับเทียบ	การออกแบบ การเก็บข้อมูล	สถานการณ์ที่ศึกษา	วิธีการ ประเมิน	ผลที่ได้จากการศึกษา
Suh et al., (2009)	- Tucker - Tucker-like - Levine observed-score - Levine true-score - Chain linear	Non - equivalence group with anchor test (NEAT)	1. ข้อสอบจำนวน 50 ข้อและข้อสอบร่วมภายในจำนวน 10 ข้อ ใช้เพื่อประเมินคุณภาพของวิธีการทั้ง 5 วิธี 2. ข้อสอบจำนวน 90 ข้อ ใช้กับรูปแบบผู้สอบกลุ่มเดียว เพื่อใช้เป็นเกณฑ์ในการประเมินวิธีการปรับเทียบ	- ค่า Bias - ค่า RMSD - ค่า SE_{slope}	- ถ้ากลุ่มตัวอย่างมีความสามารถใกล้เคียงกัน วิธีการปรับเทียบเชิงเส้นทั้ง 5 มีคุณภาพดี - วิธีของเลอวิน มีความเหมาะสมที่สุด เมื่อกลุ่มตัวอย่างมีความสามารถแตกต่างกัน และวิธีการดังกล่าวมีความถูกต้องมากขึ้นเมื่อกลุ่มตัวอย่างมีขนาดใหญ่
Mroch et al., (2009)	- Tucker - Tucker-like - Levine observed-score - Levine true-score - Chain linear	Non - equivalence group with anchor test (NEAT)	ใช้ผลการศึกษาครั้งที่ผ่านมาของ Suh และคณะ (2009) และจากการทบทวนงานวิจัยที่เกี่ยวข้องกับวิธีการปรับเทียบเชิงเส้นตรงมาใช้ในการประเมินวิธีการปรับเทียบเชิงเส้นตรงทั้ง 5 วิธี ในหลากหลายมิติ	- ค่า Bias - ค่า RMSD	- วิธีของเลอวิน เป็นตัวเลือกที่เหมาะสมที่สุดในวิธีการปรับ เทียบเชิงเส้นตรง เมื่อกลุ่มตัวอย่างมีความแตกต่างกัน - วิธีของ Tucker เหมาะกับกลุ่มตัวอย่างที่ไม่แตกต่างกัน

ตาราง 3 สรุปประเด็นสำคัญของการออกแบบวิธีการปรับเทียบคะแนนเชิงเส้นตรง (ต่อ)

ผู้วิจัย/ ปีการศึกษา	วิธีการ ปรับเทียบ	การออกแบบ การเก็บข้อมูล	สถานการณ์ที่ศึกษา	วิธีการ ประเมิน	ผลที่ได้จากการศึกษา
Puhan (2010)	- Chained linear - Tucker - Levine-observed score	Equivalence group	- กลุ่มตัวอย่างมีทั้งแบบมีความสามารถใกล้เคียงกัน และมีความสามารถแตกต่างกัน - แบบสอบหลายตัวเลือก 2 ชุด กลุ่มตัวอย่างทำแบบสอบทั้งสองชุด	- ค่า Equating error - ค่า Bias - ค่า RMSE	- วิธีของ Tucker มีความคลาดเคลื่อนของการปรับเทียบอย่างสุ่มต่ำ แต่มีความลำเอียงสูง และมีค่า RMSE สูงกว่าวิธีอื่นๆ - วิธี CLE หรือวิธี Levine observed score นำมาใช้กับกลุ่มตัวอย่างใหม่และกลุ่มเก่าที่มีความสามารถแตกต่างกัน หรือเมื่อข้อสอบร่วมกับข้อสอบทั้งหมดมีความสัมพันธ์กันสูงได้
Puhan (2012)	- Chained linear - Tucker		- สถานการณ์ที่ 1: การให้คะแนนโดยผู้ตรวจหลายคน ความสัมพันธ์ระหว่างคะแนนข้อสอบร่วมกับและคะแนนทั้งหมดอยู่ในระดับสูง - สถานการณ์ที่ 2: กลุ่มตัวอย่างกลุ่มใหม่และกลุ่มอ้างอิงมีความเท่าเทียมกันอย่างสุ่ม แต่ความสัมพันธ์ระหว่างคะแนนข้อสอบร่วมกับคะแนนทั้งหมดอยู่ในระดับต่ำ		- สถานการณ์ที่ 1: วิธี Tucker มีความแม่นยำต่ำกว่าวิธี CLE - สถานการณ์ที่ 2: วิธีของTucker มีความแม่นยำมากกว่าวิธี CLE

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2) งานวิจัยที่เกี่ยวข้องกับการออกแบบวิธีการปรับเทียบตามทฤษฎีการตอบสนองข้อสอบ (IRT)

Tate (2000) ได้ศึกษาคุณภาพของวิธีการในการเชื่อมต่อ (linking) แบบสอบรูปแบบผสม ประกอบด้วยข้อสอบแบบสร้างคำตอบ (construct response) และข้อสอบแบบหลายตัวเลือก (multiple choice) การเชื่อมโยงคะแนนใช้รูปแบบการปรับเทียบคะแนน 2 แบบ คือ วิธีค่าเฉลี่ยและซิกมา และโมเดลการตอบแบบ grade response ซึ่งเป็นภาคขยายของกระบวนการ Stocking และ Lord การศึกษาครั้งนี้ใช้การจำลองข้อมูลจากคอมพิวเตอร์เพื่อพิจารณาความคลาดเคลื่อนในการประมาณค่าสัมประสิทธิ์การเชื่อมโยงของข้อคำถาม ประกอบไปด้วย การประมาณค่าความลำเอียงและความคลาดเคลื่อนในการสุ่มสำหรับแบบสอบรูปแบบผสม โดยศึกษาความเคลื่อนในการประมาณค่าในปัจจัยต่าง ได้แก่ อัตราส่วนเชิงสัมพัทธ์ (relative proportion) ของข้อสอบแบบหลายตัวเลือกในแบบสอบ ขนาดของการเปลี่ยนแปลงความสามารถจากปีหนึ่งไปยังอีกปีหนึ่ง จำนวนของข้อสอบร่วม จำนวนของแบบสอบเชื่อมต่อที่ถูกประเมินใหม่ และการเดา

ผลการศึกษาพบว่า ความคลาดเคลื่อนของการเชื่อมต่อเกี่ยวข้องกับวิธีการในการเชื่อมต่อแบบสอบรูปแบบผสมจากปีหนึ่งไปยังอีกปีหนึ่ง ซึ่งส่งผลให้ช่วงของระดับความสามารถ จากกรณีสมมตินี้เปลี่ยนแปลงไปมาก ผลจากวิธีการประมาณค่าข้อสอบที่มีการตรวจให้คะแนนหลายค่า พบว่าวิธีค่าเฉลี่ยและวิธีซิกมา และวิธีการของ Lord และ Stocking มีความคล้ายคลึงกันมาก นอกจากนี้ยังพบว่าการประมาณค่าของสัมประสิทธิ์ของการเชื่อมต่อมีค่าเพิ่มขึ้นเมื่อมีข้อสอบร่วมและผู้สอบจากปี1มากขึ้น ในการศึกษาครั้งนี้ใช้ข้อสอบร่วมแบบหลายตัวเลือกในการเชื่อมต่อ เหมาะสมกับแบบสอบที่วัดเพียงคุณลักษณะเดียวเท่านั้น แต่พบความลำเอียงในการเชื่อมต่อเมื่อนำไปใช้กับแบบสอบที่วัดหลายคุณลักษณะ ซึ่งสามารถพิจารณาได้จากชุดของข้อสอบร่วม สัดส่วนของข้อสอบแบบหลายตัวเลือกและข้อสอบแบบสร้างคำตอบ ดังนั้นสรุปได้ว่าวิธีค่าเฉลี่ยและวิธีซิกมา และวิธีการของ Lord และ Stocking สามารถนำมาใช้ในการเชื่อมต่อแบบสอบรูปแบบผสมที่ประกอบด้วยข้อสอบแบบหลายตัวเลือก และข้อสอบแบบสร้างคำตอบได้

Kim และ Cohen (2002) ได้เปรียบเทียบวิธีการปรับเทียบระหว่างวิธีการแปลงสเกลแบบแยกกันด้วยวิธีโค้งคุณลักษณะ และวิธีการแปลงสเกลแบบพร้อมกัน สำหรับโมเดล GRM ภายใต้ทฤษฎีการตอบสนองข้อสอบ ข้อมูลที่ใช้มาจากการจำลองข้อมูล โดยสร้างข้อสอบจำนวน 30 ข้อ ภายใต้โมเดล GRM ด้วยโปรแกรม GEIRV ในสถานการณ์นี้มีผู้สอบ 2 กลุ่ม คือ กลุ่มฐาน และกลุ่มเป้าหมาย ความยาวของข้อสอบมี 3 ระดับ คือ 5, 10, 30 ในการประมาณค่าพารามิเตอร์แบบแยกกันได้ใช้โปรแกรม Multilog ภายใต้โมเดล GRM ส่วนการแปลงสเกลแบบพร้อมกันได้รวมข้อมูลจากกลุ่มฐานและกลุ่มเป้าหมาย แล้วนำมาวิเคราะห์โดยใช้ Multilog ภายใต้โมเดล GLM และในการเปรียบเทียบวิธีการปรับเทียบทั้ง 2 วิธี ซึ่งพิจารณาจากค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ผลการศึกษาพบว่า วิธีการแปลงสเกลแบบพร้อมกันมีค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ต่ำกว่าการแปลงสเกลแบบแยกจากกันเล็กน้อย กล่าวคือวิธีการปรับค่าพารามิเตอร์แบบพร้อมกันให้ผลคงที่และมีความถูกต้องกว่าวิธีการปรับค่าพารามิเตอร์แบบแยกกันเพียงเล็กน้อย

Kim (2004 อ้างถึงใน อัญชลี ศรีกลชาญ, 2552) ได้ศึกษากระบวนการวางตำแหน่งของพารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบ ที่มาจากการทดสอบที่ใช้แบบสอบรูปแบบผสมในสถานการณ์ที่ต่างกันบนสเกลร่วมกัน วิธีการเชื่อมโยงมี 4 วิธี คือ วิธีการโมเมนต์ 2 วิธี ประกอบด้วย วิธี mean/mean และวิธี mean/sigma และวิธีการโค้งคุณลักษณะ 2 วิธี (Heabara และ Stocking and Lord) โดยจำลองข้อมูลออกเป็น 2 ส่วน ส่วนแรกอยู่ภายใต้เงื่อนไขความเป็นเอกมิติของแบบสอบที่ประกอบด้วยข้อสอบแบบเลือกตอบ และข้อสอบแบบสร้างคำตอบ ส่วนที่สองเป็นแบบพหุมิติของแบบสอบที่สะท้อนให้เห็นถึงอิทธิพลของรูปแบบ (FE) โดยอิทธิพลของรูปแบบ คือ ความสัมพันธ์ระหว่างโครงสร้างที่สำคัญของข้อสอบทั้ง 2 ลักษณะที่วัดได้จากข้อสอบแบบเลือกตอบและข้อสอบแบบสร้างคำตอบ ซึ่งการศึกษาจำลองทั้ง 2 ส่วนใช้การออกแบบผู้สอบกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม และใช้แปรแกรม MULTILOG ในการประมาณค่าพารามิเตอร์

ผลการวิจัยพบว่า วิธีการโค้งคุณลักษณะมีความเหมาะสมมากกว่าวิธีโมเมนต์ เมื่อไม่คำนึงถึงอิทธิพลของรูปแบบ โดยวิธีการโค้งคุณลักษณะทั้ง 2 วิธีให้ผลที่แตกต่างกันเล็กน้อย และวิธีการปรับค่าพร้อมกันให้ผลดีกว่าวิธีการเชื่อมโยงทั้ง 4 วิธี คือ มีความถูกต้องและยังแกร่งมากพอเมื่อใช้กับแบบสอบที่ต้องพิจารณาอิทธิพลของรูปแบบด้วย

Kim และ Kolen (2006) ได้เปรียบเทียบคุณภาพของวิธีการปรับเทียบตามแนวทฤษฎีการตอบสนองข้อสอบกับการปรับค่าพารามิเตอร์พร้อมกัน ภายใต้รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกันโดยใช้ข้อสอบร่วม โดยวิธีการปรับเทียบตามแนวทฤษฎีการตอบสนองข้อสอบ ประกอบด้วย วิธีโมเมนต์ 2 วิธี คือ วิธีค่าเฉลี่ย/ค่าเฉลี่ย (mean/mean) และวิธีค่าเฉลี่ย/ซิกมา (mean/sigma) และวิธีโค้งคุณลักษณะ 2 วิธี คือ วิธี Haebara และวิธี Stocking-Lord แหล่งข้อมูลที่ใช้ในการจำลองมาจากการทดสอบ NAEP ซึ่งเป็นการประเมินทางวิทยาศาสตร์ แล้วนำข้อมูลเหล่านี้มาประมาณค่าพารามิเตอร์ของข้อสอบ โดยข้อมูลที่จำลองได้เป็นแบบพหุมิติ กล่าวคือ รูปแบบของข้อสอบมี 2 รูปแบบ ชนิดของรูปแบบผสมมี 2 รูปแบบ คือ wide-range และ narrow-range และระดับของผู้สอบ 3 ระดับ ข้อสอบร่วมประกอบข้อสอบแบบหลายตัวเลือก ที่มีการตรวจให้คะแนนแบบสองค่า ซึ่งถูกวิเคราะห์ด้วยโมเดล 3PL และข้อสอบแบบสร้างคำตอบที่มีการตรวจให้คะแนนแบบหลายค่า ซึ่งถูกวิเคราะห์ด้วยโมเดล GPC โดยเกณฑ์ที่ใช้ในการเปรียบเทียบวิธีการปรับเทียบ คือ ค่า OSD

ผลการศึกษาพบว่า ผลการปรับเทียบด้วยวิธีโค้งคุณลักษณะให้ผลการปรับเทียบที่มีคุณภาพมากกว่าวิธีโมเมนต์ทั้ง 2 วิธี เมื่อพิจาณาภาพรวมพบว่า การปรับค่าพารามิเตอร์พร้อมกันให้ผลการปรับเทียบที่มีคุณภาพมากกว่าวิธีอื่น และมีความแกร่งต่อโมเดลพหุมิติ แต่อย่างไรก็ตามการปรับสเกลพร้อมกันมีความถูกต้องและแข็งแกร่งมากกว่าวิธีโค้งคุณลักษณะเพียงเล็กน้อย

Li และคณะ (2012) ได้ศึกษาขั้นตอนของการปรับเทียบคะแนนจริงตามทฤษฎีการตอบสนองข้อสอบ (IRT) โดยยึดตามจุดประสงค์การแปลงสเกล (scale transformation) ที่หลากหลาย และประเมินความถูกต้องของการปรับเทียบคะแนน กระบวนการในการปรับเทียบแบบสอบชุดใหม่ไปยังแบบสอบชุดเดิมโดยใช้วิธีการปรับเทียบคะแนนจริงตามทฤษฎีการตอบสนองข้อสอบสามารถทำได้หลายวิธี ในที่นี้พิจารณากระบวนการปรับเทียบคะแนนดังกล่าวภายใต้รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกัน โดยใช้ข้อสอบร่วม (NEAT) และอภิปรายในรายละเอียดของฟังก์ชั่นการปรับเทียบทั้งทางตรงและทางอ้อมที่ได้มาจากแบบสอบแต่ละแบบ ดังนั้นลำดับของรูปแบบของแบบสอบจึงมี 5 แบบ คือ U, V,

X, Y, Z โดยแบบสอบทั้ง 5 รูปแบบเชื่อมต่อกันดังนี้ $U \rightarrow V \rightarrow X \rightarrow Y \rightarrow Z$ (โดย Z คือ แบบสอบที่ใช้เป็นฐาน) ผ่านชุดของข้อสอบร่วม 4 ชุดที่แตกต่างกัน โดยแบบสอบทั้ง 5 ฉบับถูกจำลองขึ้น แต่ละฉบับประกอบด้วยข้อสอบเฉพาะจำนวน 50 ข้อ และชุดของข้อสอบร่วมจำนวน 15 ข้อ ที่ใช้เชื่อมต่อแบบสอบที่อยู่ใกล้เคียงกัน ข้อสอบทั้งหมดมีการตรวจให้คะแนนแบบ 2 ค่า ส่วนการจำลองคำตอบของข้อสอบในแบบสอบแต่ละแบบใช้โมเดล 2PL และโมเดล IRT โดยขนาดของกลุ่มตัวอย่างที่จำลองเท่ากับ 2,000 ซึ่งจำลองมาจากโค้งปกติ ขณะที่การประมาณค่าพารามิเตอร์ข้อสอบใช้โปรแกรม PARSCALE

ผลการศึกษาพบว่า ความลำเอียงและความแปรปรวนของกลุ่มอย่างที่ใช้การปรับเทียบคะแนนจริงตามทฤษฎีการตอบสนองข้อสอบมีขนาดค่อนข้างเล็ก (ยกเว้นวิธีค่าเฉลี่ย/ซิกมา) ความลำเอียงและความแปรปรวนของกลุ่มตัวอย่างที่ใช้ฟังก์ชั่นการปรับเทียบตามหลักวิธีโค้งคุณลักษณะ (characteristic curve method) และการปรับเทียบพร้อมกัน (concurrent calibrations) มีขนาดเล็กว่าความลำเอียงและความแปรปรวนของกลุ่มตัวอย่างที่ใช้ฟังก์ชั่นการปรับเทียบโดยวิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ (chained equipercentile equating) นอกจากนี้เมื่อเปรียบเทียบระหว่างการปรับเทียบคะแนนจริงตามทฤษฎีการตอบสนองข้อสอบกับการปรับเทียบโดยวิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ (chained equipercentile equating) พบว่า กลุ่มตัวอย่างที่ใช้การปรับเทียบด้วยวิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ (chained equipercentile equating) มีความแปรปรวนต่ำกว่าการปรับเทียบคะแนนจริงตามทฤษฎีการตอบสนองข้อสอบ (ยกเว้นในช่วงคะแนนต่ำ)

2.3 การประเมินคุณภาพของการปรับเทียบคะแนน

การตรวจสอบคุณภาพการปรับเทียบ สามารถจัดได้เป็น 2 ประเภท ดังนี้
ศิริชัย กาญจนวาสี (2555)

2.3.1 ความเคลื่อนมาตรฐานของการปรับเทียบ (standard error of equating)

การปรับเทียบคะแนนด้วยวิธีต่าง ๆ ใช้กลุ่มตัวอย่างที่ได้มาจากการสุ่มจากประชากรเดียวกันหรือหลายประชากร ซึ่งมีความผันแปรเชิงสุ่มเกิดขึ้นกับกลุ่มตัวอย่างที่ต้องการศึกษา ดังนั้นจึงสามารถวิเคราะห์ความคลาดเคลื่อนมาตรฐานที่เกิดขึ้นในเชิงทฤษฎี สำหรับการปรับเทียบคะแนนระหว่างแบบสอบจากกลุ่มสอบทานผลได้ ดังนี้

เมื่อ E คือ ความคลาดเคลื่อนของการปรับเทียบคะแนน X บนสเกลของ Y

ดังนั้น $\mathrm{E_i} = |\mathrm{y_i} - \mathrm{y_i^*}|$

โดยที่ $\mathrm{y_i}$ คือ คะแนนที่คนที่ i จากแบบสอบ Y

$\mathrm{y_i^*}$ คือ คะแนนที่คนที่ i จากการปรับเทียบคะแนนแบบสอบ X บนสเกลของแบบสอบ Y

ส่วนเบี่ยงเบนมาตรฐานของคะแนนความคลาดเคลื่อนจากการปรับเทียบ คือ

$$SD_{\mathrm{E}} = \sqrt{\frac{\sum E_i^2}{n} - \left(\frac{\sum E_i}{n}\right)^2}$$

เมื่อ n คือ ขนาดของกลุ่มตัวอย่างในกลุ่มสอบทานผล

ดังนั้นความคลาดเคลื่อนมาตรฐานของการปรับเทียบ (SEE) หาได้จากสูตร

$$SEE = \frac{SD_{\mathrm{E}}}{\sqrt{n}}$$

2.3.2 ความเพียงพอของการปรับเทียบ (adequacy of equating models)

ในการปรับเทียบคะแนนระหว่างแบบสอบแต่ละวิธี ต้องมีเงื่อนไขของความเสมอภาคหรือความเหมาะสมซึ่งเป็นไปตามข้อตกลงเบื้องต้นของแต่รูปแบบหรือโมเดล ดังนั้นการเลือกรูปแบบใดไปใช้ สภาพการทดสอบอาจเป็นไปตามเงื่อนไขเพียงข้อใดข้อหนึ่ง ซึ่ง Braun และ Holland (1982 อ้างถึงใน พัชรี จันทร์เพ็ง, 2547) ได้ยกตัวอย่างโปรแกรมการทดสอบระดับนานาชาติ ซึ่งมีทั้งกฎเกณฑ์และนโยบายเป็นตัวกำหนด จึงไม่สามารถควบคุมการสอบให้เป็นไปตามข้อตกลงเชิงทฤษฎี ดังนั้นข้อมูลที่นำมาใช้ศึกษาจึงไม่สามารถระบุได้ว่าเป็นตัวอย่างของประชากรได้อย่างชัดเจน และความเป็นจริงอาจจัดกระทำกับประชากรมากกว่า (population quantities) ไม่ใช่ค่าประมาณ (sample estimate)

การปรับเทียบคะแนนที่จัดกระทำอยู่ในสภาพที่มีเงื่อนไขต่ำกว่าข้อตกลงที่มีในแต่ละรูปแบบ จึงจำเป็นต้องตรวจสอบความเพียงพอของการปรับเทียบ (adequacy of equating models) โดย Petersen และคณะ (1982 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555) ได้เสนอการตรวจสอบความเพียงพอของการปรับเทียบคะแนนไว้ ดังนี้

1) ดัชนีเปรียบเทียบเปอร์เซ็นไทล์ (percentile comparison index)

Kowen และ Whitney (1982 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555) ได้เสนอการใช้คะแนนจากกลุ่มตัวอย่างสอบทานผล ที่เป็นกลุ่มตัวอย่างที่สุ่มมาจากประชากรกลุ่มเดียวกันกับกลุ่มตัวอย่างเทียบมาตรา โดยไม่มีหน่วยตัวอย่างซ้ำ ในการปรับเทียบคะแนนของ X บนสเกลของ Y (Y^*) ด้วยการปรับเทียบคะแนน y และคะแนนปรับเทียบ y^* ณ ตำแหน่งเปอร์เซ็นไทล์เดียวกัน ดังสมการ

$$c = \frac{\sum(y - y^*)}{nk}$$

เมื่อ c คือ ดัชนีเปรียบเทียบเปอร์เซ็นไทล์

y คือ คะแนนจากแบบสอบ Y ซึ่งเป็นคะแนนเกณฑ์

y^* คือ คะแนนเปรียบเทียบของแบบสอบ X ที่อยู่บนสเกลของ Y ณ ตำแหน่งเปอร์เซ็นไทล์เดียวกัน

n คือ จำนวนกลุ่มสอบทานผล

k คือ จำนวนข้อสอบร่วม

จากสมการข้างต้น ดัชนีเปรียบเทียบเปอร์เซ็นไทล์ คือ ค่าเฉลี่ยกำลังสองของความแตกต่าง (the mean-squared difference) ที่ได้จากการแจกแจงของคะแนนต่าง ๆ ของเกณฑ์ X กับคะแนนแปลง X ที่แปลงมาจาก Y ด้วยวิธีการปรับเทียบคะแนนที่ระบุไว้ ณ ที่ตำแหน่งเปอร์เซ็นไทล์เดียวกัน

ถ้าค่า c มีค่าต่ำ หมายความว่า รูปแบบการปรับเทียบที่นำมาใช้มีความเพียงพอและเหมาะสมในการให้ผลการเปรียบเทียบที่มีความน่าเชื่อถือ หรือให้ผลการปรับเทียบที่คงเส้นคงวา

2) ดัชนีความแตกต่าง (discrepancy index)

Petersen และคณะ (1982 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555) ได้เสนอดัชนีความแตกต่างระหว่างคะแนนเกณฑ์กับคะแนนปรับเทียบ โดยใช้ค่าความแปรปรวนเป็นตัวถ่วงน้ำหนัก เพื่อให้ได้ค่าคงที่เป็นมาตรฐาน ดังนี้

$$c = \frac{\sum_{i=1}^{n}(y_i - y_i^*)^2}{nS_y^2}$$

เมื่อ c คือ ดัชนีความแตกต่างมาตรฐาน

y_i คือ คะแนนจากแบบสอบ Y ของคนที่ i ซึ่งเป็นคะแนนเกณฑ์

y^* คือ คะแนนเปรียบเทียบของแบบสอบ X ที่อยู่บนสเกลของ Y สำหรับคนที่ i

n คือ จำนวนกลุ่มสอบทานผล

S_y^2 คือ ความแปรปรวนของคะแนนเกณฑ์

3) สัมประสิทธิ์การทำนาย (coefficient of prediction)

สัมประสิทธิ์การทำนาย คือ ค่าสหสัมพันธ์กำลังสอง (R^2) สำหรับการทำนายคะแนนจาก y จากคะแนนปรับเทียบ y^* ซึ่งมีสมการในการคำนวณ ดังนี้

$$R = \frac{n\sum(yy^*) - (\sum y)(\sum y^*)}{\sqrt{[n\sum y^2 - (\sum y)^2][n\sum y^{*2} - (\sum y^*)^2]}}$$

ค่า R^2 มีค่าสูงมากเท่าใด แสดงว่าค่าคะแนนปรับเทียบ (y^*) สามารถทำนายค่าคะแนนเกณฑ์ (y) ได้ใกล้เคียงมากเท่านั้น

นอกจากนี้ Kim และคณะ (2010) ยังได้เสนอวิธีการในการวัดความแม่นยำของการปรับเทียบ (the measure of accuracy) ดังนี้

2.3.3 สถิติบูทสแตรป (bootstrapped statistics)

สถิติบูทสแตรป (bootstrapped statistics) เป็นวิธีการประมาณความคลาดเคลื่อนมาตรฐานในทางสถิติที่ใช้กันอย่างกว้างขวาง โดยคำนวณอย่างละเอียดด้วยคอมพิวเตอร์ ซึ่งสถิติบูทสแตรปเกี่ยวข้องกับการสุ่มกลุ่มตัวอย่างโดยใช้เทคนิคการสุ่มซ้ำ (resampling) ด้วยการแทนที่คืน (replacement) กล่าวคือ ใช้คอมพิวเตอร์ทำการสุ่มตัวอย่างหลายชุด ๆ จากประชากร โดยสุ่มตัวอย่างมาหนึ่งกลุ่ม เมื่อวัดค่าได้แล้วกลุ่มตัวอย่างดังกล่าวก็ถูกส่งกลับไปยังประชากร แล้วทำการสุ่มซ้ำจนครบตามจำนวนที่ต้องการ ตัวอย่างอย่างง่ายของการนำสถิติบูทสแตรปไปใช้ คือ การประมาณความเคลื่อนมาตรฐาน (standard error) ซึ่งมีขั้นตอนดังนี้ (Kolen & Brennan, 2004)

การประมาณความเคลื่อนมาตรฐาน (standard errors) โดยใช้สถิติบูทสแตรป

ขั้นตอนการประมาณความเคลื่อนมาตรฐาน (standard errors) โดยใช้สถิติบูทสแตรปจากกลุ่มตัวอย่างกลุ่มหนึ่ง มีดังนี้

1. เริ่มต้นด้วยกลุ่มตัวอย่างขนาด N
2. สร้างกลุ่มตัวอย่างอย่างสุ่ม (random sample) ด้วยการแทนที่จากกลุ่มตัวอย่างขนาด N โดยกลุ่มตัวอย่างที่สุ่มขึ้นมานี้ เรียกว่า กลุ่มตัวอย่างบูทสแตรป (bootstrap sample)
3. ทำตามขั้นตอน 2 และ 3 ซ้ำอีก R ครั้ง
4. คำนวณความคลาดเคลื่อนมาตรฐานของสถิติของกลุ่มตัวอย่างบูทสแตรป R กลุ่ม ซึ่งความคลาดเคลื่อนมาตรฐานนี้ คือ ความคลาดเคลื่อนมาตรฐานที่ถูกประมาณโดยสถิติบูทสแตรป

ในการประยุกต์ใช้สถิติบูทสแตรปมาใช้ในการปรับเทียบแบบสอบทั้งสองฉบับสำหรับรูปแบบผู้สอบกลุ่มสุ่ม (random groups design) และรูปแบบผู้สอบกลุ่มไม่เท่าเทียมกันโดยใช้ข้อสอบร่วม (common-item nonequivalent groups design) ขนาดของตัวอย่างประกอบด้วยขนาดของกลุ่มตัวอย่างคะแนนจากแบบสอบ X คือ N_X และขนาดของกลุ่มตัวอย่างคะแนนจากแบบสอบ Y คือ N_Y

สถิติบูทสแตรปอิงพารามิเตอร์ (parametric bootstrap)

กลุ่มตัวอย่างของสถิติบูทสแตรปอิงพารามิเตอร์ (parametric bootstrap) ได้มาจากการแจกแจงที่มีความเหมาะสมกับข้อมูล (fitted distribution) ขณะที่กลุ่มตัวอย่างของสถิติบูทสแตรปมาจากการแจกแจงของข้อมูลเชิงประจักษ์ (empirical distribution) แต่ถ้าโมเดลอิงพารามิเตอร์ที่เหมาะสมกับข้อมูล (fitted parametric model) ไม่สามารถประมาณการแจกแจงของประชากรได้อย่างถูกต้อง การใช้สถิติบูทสแตรปอิงพารามิเตอร์อาจทำให้เกิดค่าความลำเอียง (bias) ในการประมาณความคลาดเคลื่อนมาตรฐาน (standard errors) มากกว่าการใช้สถิติบูทสแตรป

นอกจากนี้ Kim และคณะ (2010) ได้นำเสนอสถิติบูทสแตรปที่ใช้ในการเปรียบเทียบความแม่นยำ (accuracy) ของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ได้แก่ ค่าความลำเอียง (bias) ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE)

2.3.3.1 ค่าความลำเอียง (Bias)

ค่าความลำเอียง (Bias) สามารถตีความได้ว่าเป็นความคลาดเคลื่อนอย่างเป็นระบบ โดยที่ค่าความลำเอียง (bias) กับค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) แปรผันตามกัน (Kane 1996, cited in Suh et al., 2009) ค่าความลำเอียงในแต่ละจุดคะแนน ($Bias_i$) สามารถคำนวณได้จากสมการด้านล่างผ่านโปรแกรม R (Albano, 2014)

$$Bias_i = \hat{\bar{e}}_Y(x) - e_Y(x)$$

เมื่อ $$\hat{\bar{e}}_Y(x) = \frac{1}{R}\sum_{r=1}^{R} \hat{e}_{Yr}(x)$$

ถ้า $e_Y(x)$ คือ ฟังก์ชั่นการปรับเทียบของประชากร (population equating function)

$\hat{e}_{Yr}(x)$ คือ ฟังก์ชั่นการปรับเทียบที่คำนวณจากการออกแบบการปรับเทียบคะแนนรูปแบบต่างๆ สำหรับกลุ่มตัวอย่าง r = 1, 2,..., R

$\hat{\bar{e}}_Y(x)$ คือ ค่าประมาณเฉลี่ยของ $\hat{e}_{Yr}(x)$ สำหรับกลุ่มตัวอย่าง r = 1, 2,..., R

หลังจากนั้นนำค่าความลำเอียง (bias) ที่ได้ในแต่ละจุดคะแนน มาหาค่าเฉลี่ยน้ำหนักของความลำเอียง (weighted average bias) โดยแทนในสมการ $\sqrt{\sum_i w_i Bias_i^2}$

2.3.3.2 ค่าความเคลื่อนในการปรับเทียบ (standard error of equating)

ค่าความเคลื่อนในการปรับเทียบ (standard error of equating) หรือเรียกอีกอย่างหนึ่งว่า ความคลาดเคลื่อนในการปรับเทียบ (equating error) ซึ่งถือว่าเป็นความคลาดเคลื่อนอย่างสุ่ม โดยค่าความคลาดเคลื่อนในการปรับเทียบในแต่ละจุดคะแนน (SEE_i) สามารถคำนวณได้จากสมการด้านล่างด้วยโปรแกรม R (Albano, 2014)

$$SEE_i = \sqrt{\frac{1}{R}\sum_{r=1}^{R}[\hat{\bar{e}}_Y(x) - e_Y(x)]^2}$$

หลังจากนั้นนำค่า SEE ในแต่ละจุดคะแนน มาหาค่าเฉลี่ยน้ำหนักของความคลาดเคลื่อนในการปรับเทียบ (weighted average equating error) โดยแทนในสมการ $\sqrt{\sum_i w_i SEE_i^2}$

2.3.3.3 ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลัง (Root Mean Squared Error: RMSE)

ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลัง (RMSE) เป็นการรวมค่าความคลาดเคลื่อนอย่างเป็นระบบและความคลาดเคลื่อนอย่างสุ่ม โดยค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสองในแต่ละจุดคะแนน ($RMSE_i$) สามารถคำนวณได้จากสมการด้านล่างด้วยโปรแกรม R (Albano, 2014)

$$RMSE_i = \sqrt{bias^2 + SEE^2}$$

หลังจากนั้นนำค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ที่ได้ในแต่ละจุดมาหาค่าเฉลี่ยน้ำหนักของรากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลัง (weighted average RMSE) โดยแทนในสมการ $\sqrt{\sum_i w_i RMSE_i^2}$

2.3.4 ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (Root Mean Squared Difference: RMSD)

ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (Root Mean Squared Difference: RMSD) ใช้ในการเปรียบเทียบประสิทธิภาพของวิธีการปรับเทียบ (equating method) ซึ่งคำนวณจากโปรแกรม Microsoft excel ตามสมการด้านล่าง

$$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$$

เมื่อ i คือ จุดของคะแนนที่สังเกตได้ (raw score point)

$\hat{e}_i(x_i)$ คือ คะแนนที่ถูกปรับเทียบโดยใช้วิธีการปรับเทียบแบบต่าง ๆ ที่จุดของคะแนนที่สังเกตได้ x

$e_i(x_i)$ คือ ฟังก์ชั่นของการปรับเทียบที่เป็นเกณฑ์ (criterion equating function) ที่จุดของคะแนนที่สังเกตได้ x

w_i คือ สัดส่วนสัมพัทธ์ของผู้สอบที่ทำแบบสอบฉบับใหม่ที่แต่ละจุดของคะแนน

Kane (1996, cited in Suh et al., 2009) กล่าวว่า วิธีการปรับเทียบที่มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำ ๆ มีประสิทธิภาพในการปรับเทียบมากกว่าวิธีการปรับเทียบที่มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) มาก ๆ

2.4 การตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) สำหรับการปรับเทียบคะแนน

การศึกษาการเชื่อมต่อที่เกี่ยวข้องกับการให้คะแนนในข้อสอบแบบสร้างคำตอบมาจากประชากรกลุ่มอ้างอิง (reference) โดยกลุ่มตัวอย่างที่เป็นตัวแทนของข้อสอบร่วมสำหรับผู้สอบปี 1 (reference year) ถูกแทรกเข้าไปในกระบวนการการประเมินสำหรับผู้สอบปี 2 (new year) ดังนั้นคำตอบของผู้สอบกลุ่มอ้างอิง (reference) ถูกให้คะแนนโดยผู้ตรวจกลุ่มเดิมกับข้อสอบร่วมของผู้สอบกลุ่มใหม่ (new) เพราะฉะนั้น Trend paper จึงมีคะแนนสอบสองชุด คือ คะแนนชุดหนึ่งจากผู้ตรวจให้คะแนนฉบับเก่า (old) และคะแนนชุดหนึ่งจากผู้ตรวจให้คะแนนฉบับใหม่ (new) ดังนั้นการตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) คือ การตรวจให้คะแนนข้อสอบร่วมแบบข้ามกลุ่มสำหรับผู้สอบที่มีความแตกต่างกัน (Tate, 1999, 2000 อ้างถึงใน Kim et al., 2010)

เนื่องจากข้อสอบร่วมแบบสร้างคำตอบอาจทำให้เกิดความแตกต่างของคะแนนที่ได้จากผู้ตรวจให้คะแนนทั้งสองกลุ่ม ดังนั้นจึงต้องใช้กระบวนการปรับเทียบ โดยผู้ตรวจให้คะแนนกลุ่มหนึ่งตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบกลุ่มใหม่ (new) และกลุ่มอ้างอิง (reference) ซึ่งวิธีนี้ถือว่าความเข้มงวดของกลุ่มผู้ประเมินเป็นค่าคงที่ และความแตกต่างระหว่างการประมาณค่าพารามิเตอร์ของข้อสอบร่วมของผู้สอบกลุ่มใหม่ (new) และกลุ่มอ้างอิง (reference) สะท้อนถึงความแตกต่างระหว่างผู้สอบกลุ่มใหม่ (new) และกลุ่มอ้างอิง (reference)

2.5 งานวิจัยที่เกี่ยวข้องกับรูปแบบการเชื่อมต่อ (linking designs) การออกแบบการปรับเทียบคะแนน (equating designs) สัดส่วนของข้อสอบร่วมในการปรับเทียบคะแนน และการตรวจให้คะแนนข้ามกลุ่ม (trend scoring method)

Bastari (2000 อ้างถึงใน อัญชลี ศรีกลชาญ, 2552) ได้ศึกษาองค์ประกอบที่มีผลต่อความแม่นยำในการเชื่อมโยงค่าประมาณพารามิเตอร์ข้อสอบรูปแบบผสมระหว่างข้อสอบแบบหลายตัวเลือก (MC) และข้อสอบแบบสร้างคำตอบ (CR) โดยใช้โมเดลโลจิสติก 3 พารามิเตอร์ร่วมกับ Graded response model การศึกษาครั้งนี้เป็นการจำลองข้อมูลจากค่าประมาณพารามิเตอร์ข้อสอบของแบบสอบ MACS ปี 1998 สำหรับวิชาวิทยาศาสตร์ ซึ่งแบบสอบร่วมที่ทำการศึกษามีทั้งข้อสอบแบบหลายตัวเลือกและข้อสอบแบบสร้างคำตอบ โดยปัจจัยที่ใช้ในการศึกษามี 6 องค์ประกอบ คือ 1) ความยาวของแบบสอบมีขนาด คือ 30 ข้อ และ 50 ข้อ 2) สัดส่วนของข้อสอบแต่ละประเภทในแบบสอบ 2 ระดับ คือ ข้อสอบแบบหลายตัวเลือก 80% และข้อสอบแบบสร้างคำตอบ 20% กับข้อสอบแบบหลายตัวเลือก 90% และข้อสอบแบบสร้างคำตอบ 10% 3) ความยาวของข้อสอบร่วมโดยใช้กฎร้อยละ 10 และ 20 ของความยาวแบบสอบทั้งฉบับ 4) ขนาดของหน่วยตัวอย่าง คือ 400, 1,000 และ 3,000 5) การแจกแจงความสามารถโดยมีการแจกแจงปกติและจัดให้มีระดับความแตกต่างในค่าเฉลี่ย 2 ระดับ คือ มีความแตกต่างเป็น 0 และ 1 หน่วยเบี่ยงเบนมาตรฐาน 6) วิธีการเชื่อมโยงคะแนนที่นำมาใช้ 2 วิธี คือ วิธีการปรับค่าพารามิเตอร์พร้อมกัน และวิธีโค้งคุณลักษณะข้อสอบ ซึ่งใช้โปรแกรม MULTILOG ในการวิเคราะห์ โดยปรับค่าพารามิเตอร์พร้อมกันและแยกกัน เกณฑ์ที่ใช้ในการประเมินมี 2 เกณฑ์ เกณฑ์ที่ 1 คือ ความแตกต่างกำลังสองเฉลี่ย (MSD) ประกอบด้วยค่าความลำเอียง (bias) และความแปรปรวนสำหรับแต่ละค่าประมาณพารามิเตอร์ ได้แก่ พารามิเตอร์ความยากและพารามิเตอร์อำนาจจำแนก ส่วนเกณฑ์ที่ 2 คือ ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) และค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) เกณฑ์ที่สองใช้หลักในการพิจารณา คือ ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) และค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ที่มีค่าน้อยกว่าจะมีความแม่นยำมากกว่า

ผลการศึกษาพบว่า ปัจจัยที่มีผลให้ประสิทธิภาพมีมากขึ้น คือ แบบสอบที่ยาวกว่า (RMSE ของข้อสอบ 50 ข้อ ต่ำกว่า RMSE ของข้อสอบ 30 ข้อ) สัดส่วนของข้อสอบแบบสร้างคำตอบที่มากขึ้น (RMSE ของสัดส่วน 90 : 10 ต่ำกว่า RMSE ของสัดส่วน 80 : 20) แบบสอบร่วมที่ยาวกว่า (RMSE ของข้อสอบร่วม 20% ต่ำกว่า RMSE ของข้อสอบร่วม 10%) หน่วยตัวอย่างขนาดใหญ่กว่า (RMSE ของจำนวนกลุ่มตัวอย่าง 3,000 คน ต่ำที่สุด และ RMSE ของกลุ่มตัวอย่าง 400 คนสูงที่สุด) และความไม่แตกต่างระหว่างความสามารถของกลุ่มที่ใช้ในการเชื่อมโยงคะแนน และวิธีการปรับค่าพารามิเตอร์พร้อมกัน

Muraki และคณะ (2000) ได้ศึกษาวิธีการในการเชื่อมต่อ (linking methods) ในการประเมินผลจากการปฏิบัติ (performance assessments) โดยประเด็นที่ศึกษาได้แก่ 1) การเปรียบเทียบรูปแบบการเชื่อมต่อ 3 รูปแบบ ได้แก่ รูปแบบผู้สอบกลุ่มเดียว (single group) รูปแบบผู้สอบกลุ่มเท่าเทียมอย่างสุ่ม (randomly equivalent groups) และผู้สอบกลุ่มไม่เท่าเทียมกัน โดย

ใช้ข้อสอบร่วม (nonequivalent groups with anchor items) 2) วิธีการในการเชื่อมต่อ 2 แบบ ได้แก่ รูปแบบการเชื่อมต่อที่ใช้ทฤษฎีดั้งเดิม และรูปแบบการเชื่อมต่อที่ใช้ทฤษฎีการตอบสนองข้อสอบ โดยรูปแบบการปรับเทียบโดยใช้ทฤษฎีดั้งเดิม ใช้วิธีปรับเทียบคะแนนเชิงเส้นตรง (linear equating) และวิธีการปรับเทียบคะแนนอีควิปเปอร์เซ็นไทล์ (equipercentile equating) ส่วนรูปแบบการปรับเทียบโดยใช้ทฤษฎีการตอบสนองสนองข้อสอบ ใช้วิธี item response function, vertical, common population และ multiple-group

ผลการศึกษาพบว่า การประเมินผลจากปฏิบัติพบความไม่เป็นอิสระของตำแหน่งข้อสอบ (local item dependence) มากกว่าแบบสอบประเภทหลายตัวเลือก และการประเมินผลจากการปฏิบัติยังมีผลกระทบต่อวิธีการปรับเทียบโดยใช้ทฤษฎีการตอบสนองข้อสอบ ในการทดสอบความเข้าใจในการอ่านโดยใช้ข้อสอบแบบหลายตัวเลือกที่มีการให้คะแนนแบบสองค่ามีผลกระทบของความไม่อิสระของตำแหน่งข้อสอบในระดับต่ำกว่าการประเมินผลจากการปฏิบัติที่มีการให้คะแนนแบบหลายค่า และพบว่ายังไม่มีโมเดลของทฤษฎีการตอบสนองข้อสอบที่เหมาะสมในการนำไปใช้ในสถานการณ์ดังกล่าว นอกจากนี้ Muraki และคณะ (2000 cited in Kim et al., 2010) ยังพบว่าข้อสอบแบบสร้างคำตอบไม่สามารถใช้ซ้ำในการทดสอบสองแบบที่แตกต่างกัน เพราะได้รับผลกระทบจากการจดจำ ดังนั้นการใช้ข้อสอบร่วมแบบสร้างคำตอบ (common CR items) จึงไม่นิยมใช้ในการปรับเทียบคะแนน

Kim และ Lee (2006) ได้ศึกษาคุณภาพของวิธีการปรับเทียบคะแนนตามแนวทฤษฎีการตอบสนองข้อสอบ สำหรับแบบสอบรูปแบบผสมที่ประกอบด้วยข้อสอบแบบหลายตัวเลือก และข้อสอบแบบสร้างตัวเลือก ซึ่งประกอบด้วย 4 วิธี ได้แก่ วิธีค่าเฉลี่ย/ค่าเฉลี่ย วิธีค่าเฉลี่ย/ซิกมา วิธี Haebara และวิธี Stocking-Lord โดยข้อสอบที่มีการตรวจให้คะแนนแบบสองค่า (DS) ใช้โมเดล IRT แบบ 3 พารามิเตอร์ ข้อสอบที่มีการตรวจให้คะแนนแบบหลายค่า (PS) ใช้โมเดล Generalized Partial Credit ที่มีการตรวจให้คะแนนเป็น 5 ลำดับ องค์ประกอบที่ใช้ในการศึกษาครั้งนี้ ได้แก่ 1) การออกแบบการเชื่อมต่อ สำหรับกลุ่มผู้สอบที่เท่าเทียมกัน และไม่เท่าเทียมกัน 2) ขนาดของกลุ่มตัวอย่าง มี 2 ขนาดได้แก่ 500 คน และ 3,000 คน 3) สัดส่วนของข้อสอบที่มีการตรวจให้คะแนนสองค่า (DS) 10 ข้อ และข้อสอบที่มีการตรวจให้คะแนนหลายค่า (PS) ได้แก่ DS 10 ข้อ, DS 20 ข้อ และ PS 5 ข้อ และ DS 30 ข้อ และ PS 2 ข้อ 4) ประเภทของการเชื่อมต่อ ได้แก่ เชื่อมโยงข้อสอบร่วมเฉพาะ DS และเชื่อมโยงข้อสอบร่วมเฉพาะ PS และเชื่อมโยงข้อสอบร่วมทั้ง DS และ PS ในการประมาณค่าพารามิเตอร์ของข้อสอบได้ใช้โปรแกรม MULTILOG ส่วนการเชื่อมโยงคะแนนในแต่ละวิธีได้ใช้โปรแกรม STUIRT และเกณฑ์ที่ใช้ในการประเมินวิธีการเชื่อมโยงพิจารณาจากเกณฑ์โค้งลักษณะรายการคำตอบ (category characteristic curve)

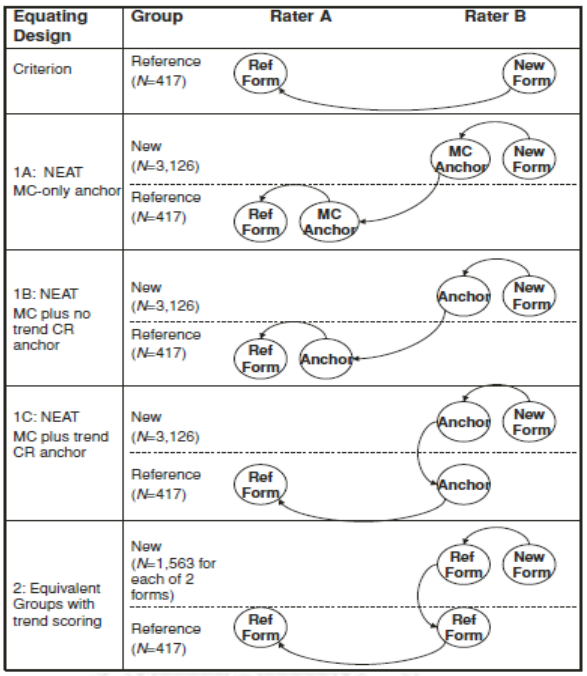
ผลที่ได้จากการศึกษา พบว่า วิธีโค้งลักษณะมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ต่ำกว่าวิธีโมเมนต์ในทุกเงื่อนไข ส่วนอีก 4 วิธีที่เหลือ กล่าวคือ การปรับเทียบค่าวิธีโค้งคุณลักษณะให้ผลการปรับเทียบที่มีคุณภาพมากกว่าการปรับเทียบแบบโมเมนต์ ดังนั้นวิธีโค้งลักษณะจึงมีความเหมาะสมในการนำมาใช้เชื่อมโยงคะแนนของแบบสอบรูปแบบผสมมากกว่าวิธีโมเมนต์ ในภาพรวมวิธีการปรับเทียบคะแนนแบบการปรับค่าพารามิเตอร์พร้อมกันให้ผลดีกว่าวิธีการปรับเทียบวิธีอื่น ๆ นอกจากนี้ยังพบว่าการปรับเทียบมีความคลาดเคลื่อนต่ำเมื่อ 1) กลุ่มผู้สอบเท่าเทียมกัน

2) ข้อสอบร่วมเป็นข้อสอบที่มีการตรวจให้คะแนนสองค่า (DS) ลดลง และ 3) ข้อสอบร่วม ซึ่งประกอบด้วยข้อสอบที่มีการตรวจให้คะแนนสองค่า (DS) และข้อสอบที่มีการตรวจให้คะแนนหลายค่า (PS) แต่อย่างไรก็ตามงานวิจัยยังมีข้อจำกัดในการนำผลการวิจัยไปใช้สรุปยังสถานการณ์อื่น ๆ เพราะงานวิจัยนี้จำลองค่าพารามิเตอร์จากวิธีทางสถิติ ไม่ได้จำลองมาจากข้อมูลจริง และมีการควบคุมคลาดเคลื่อนในการสุ่มอย่างเคร่งครัด ทำให้ข้อมูลที่ได้มีความเหมาะสมในการศึกษาเชิงทฤษฎีมากกว่าสถานการณ์จริง ดังนั้นในการวิจัยครั้งต่อไปจึงควรจำลองข้อมลมาจากข้อมูลจริง

Kim และคณะ (2010) ได้ศึกษาความคลาดเคลื่อนของการปรับเทียบคะแนนระหว่างแบบสอบในกลุ่มผู้สอบผู้สอบกลุ่มไม่เท่าเทียมกัน (nonequivalent groups) สำหรับแบบสอบรูปแบบผสมที่ประกอบด้วยข้อสอบแบบหลายตัวเลือก (multiple-choice) และข้อสอบแบบสร้างคำตอบ (constructed-response) เพื่อพิจาณาว่ารูปแบบการปรับเทียบแบบใดที่มีคุณภาพในการเชื่อมต่อแบบสอบรูปแบบผสม (linking test) มากที่สุด ข้อมูลที่ใช้ในการศึกษาครั้งนี้มาจากการทดสอบระดับใหญ่ (SAT) ด้วยวิธีการปรับเทียบคะแนนระหว่างแบบสอบแบบดั้งเดิม (classical equating methods) ด้วยวิธี CLE (Chained Linear Equating) โดยรูปแบบการเชื่อมต่อที่ศึกษามี 4 รูปแบบ ได้แก่ 1) ใช้ข้อสอบร่วมแบบหลายตัวเลือก 2) ใช้ข้อสอบร่วมแบบผสมระหว่างข้อสอบแบบหลายตัวเลือก และข้อสอบแบบสร้างคำตอบ 3) ใช้ข้อสอบร่วมแบบผสมระหว่างข้อสอบแบบหลายตัวเลือก และข้อสอบแบบสร้างคำตอบ ที่มีการตรวจให้คะแนนแบบข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ 4) ใช้ผู้สอบกลุ่มเท่าเทียมกัน (equivalent group) ที่มีการตรวจให้คะแนนแบบข้ามกลุ่มในข้อสอบแบบสร้างคำตอบ (ฉบับอ้างอิง) โดยการออกแบบการปรับเทียบที่เป็นเกณฑ์ (criterion design) ใช้รูปแบบผู้สอบกลุ่มเดียว (single group) โดยผู้สอบในรูปแบบดังกล่าวทำแบบสอบฉบับใหม่ หลังจากนั้นผ่านไปประมาณ 1 เดือน แล้วให้ผู้สอบกลุ่มเดียวกันนี้ทำแบบสอบฉบับอ้างอิง ซึ่งแบบสอบทั้ง 2 ฉบับนี้เป็นแบบสอบคู่ขนานที่มีข้อสอบร่วมรูปแบบผสม โดยการออกแบบการปรับเทียบที่เป็นเกณฑ์ใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed) ในการเชื่อมต่อคะแนนจากแบบสอบฉบับใหม่โดยผู้ตรวจ 2 ไปยังคะแนนจากแบบสอบฉบับอ้างอิงโดยผู้ตรวจ 1 ดังภาพ 1 กลุ่มตัวอย่างที่ใช้ในการศึกษาครั้งนี้ ประกอบด้วย 2 กลุ่มย่อย ได้แก่ กลุ่มตัวอย่างกลุ่มอ้างอิงจำนวน 417 คน กลุ่มตัวอย่างกลุ่มใหม่จำนวน 3,126 คน และมีผู้ตรวจให้คะแนน 2 กลุ่ม คือ กลุ่ม A และกลุ่ม B วิธีการในการปรับเทียบคะแนนระหว่างแบบสอบ ใช้การปรับเทียบคะแนนแบบเส้นตรง (วิธี CLE (Chained Linear Equating))

ผลการศึกษาพบว่า รูปแบบที่ใช้ข้อสอบร่วมแบบหลายตัวเลือกเพียงอย่างเดียว หรือรูปแบบใช้ข้อสอบร่วมแบบผสมระหว่างข้อสอบแบบหลายตัวเลือกและข้อสอบแบบสร้างคำตอบที่ไม่มีการตรวจให้คะแนนแบบข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ มีความลำเอียงในการปรับเทียบคะแนนมากกว่าอีกสองรูปแบบที่เหลือ นอกจากนี้ยังพบว่ารูปแบบที่ใช้ผู้สอบกลุ่มเท่าเทียมกัน (equivalent group) ที่มีการตรวจให้คะแนนแบบข้ามกลุ่มในข้อสอบแบบสร้างคำตอบ มีค่าความลำเอียง (bias) ในการปรับเทียบคะแนนต่ำที่สุด เนื่องจากมีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำที่สุด (RMSE = .549) ดังนั้นการออกแบบการปรับเทียบคะแนนที่มีความเหมาะสมสำหรับแบบสอบรูปแบบผสม ภายใต้รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกัน โดยใช้

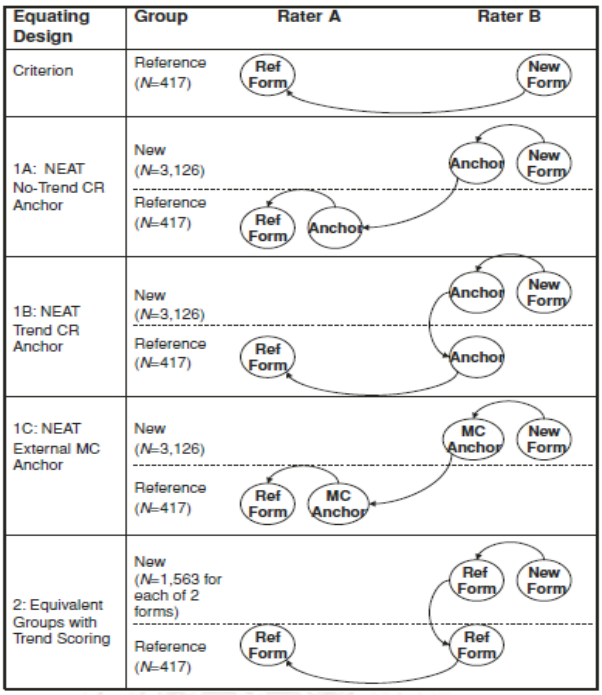
ข้อสอบร่วม (NEAT) คือ รูปแบบที่ใช้ข้อสอบร่วมรูปแบบผสมระหว่างข้อสอบแบบหลายตัวเลือก และข้อสอบแบบสร้างคำตอบ และมีการตรวจให้คะแนนแบบข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ และรูปแบบใช้ผู้สอบกลุ่มเท่าเทียมกัน (equivalent group) และมีการตรวจให้คะแนนแบบข้ามกลุ่มในข้อสอบแบบสร้างคำตอบ



ภาพ 1 การออกแบบการปรับเทียบ (equating designs) สำหรับแบบสอบรูปแบบผสม

Kim และคณะ (2010) ยังได้ศึกษาสถิติบูทสแตรป (bootstrapped statistics) ของการออกแบบการปรับเทียบคะแนนระหว่างแบบสอบแบบสร้างคำตอบ (constructed-response tests) ในการทดสอบขนาดใหญ่ (SAT) โดยเปรียบเทียบการออกการปรับเทียบคะแนนระหว่างแบบสอบประเภทสร้างคำตอบ 4 รูปแบบ ได้แก่ 1) ใช้ข้อสอบร่วมแบบสร้างคำตอบ 2) ใช้ข้อสอบร่วมแบบสร้างคำตอบ ที่มีการตรวจให้คะแนนข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ 3) ใช้ข้อสอบร่วมภายนอกแบบหลายตัวเลือก (multiple-choice) 4) ใช้กลุ่มผู้สอบเท่าเทียมกัน ไม่มีข้อสอบร่วม แต่มีตรวจให้คะแนนข้ามกลุ่มในแบบสอบประเภทสร้างคำตอบชุดอ้างอิง (reference) เครื่องมือที่ใช้ในการศึกษาครั้งนี้ คือ แบบสอบที่ประกอบด้วยข้อสอบแบบหลายตัวเลือกจำนวน 24 ข้อ และข้อสอบแบบสร้างคำตอบจำนวน 12 ข้อ ซึ่งแบ่งออกเป็นแบบสอบ 3 แบบย่อยได้แก่ แบบสอบชุดอ้างอิง (reference form) ที่ประกอบด้วยข้อสอบแบบสร้างคำตอบจำนวน 8 ข้อ (มีข้อสอบร่วมรวมอยู่ในแบบสอบฉบับนี้จำนวน 4 ข้อ) แบบสอบชุดใหม่ (new form) ที่ประกอบด้วยข้อสอบแบบสร้างคำตอบจำนวน 8 ข้อ (มีข้อสอบร่วมรวมอยู่ในแบบสอบฉบับนี้จำนวน 4 ข้อ) ดังภาพ 2 โดยแบบสอบแบบสร้างคำตอบทั้งสองฉบับมีเนื้อหาและระดับความยากที่คู่ขนานกัน และแบบสอบแบบหลายตัวเลือก (MC test form) ที่ประกอบด้วยข้อสอบแบบหลายตัวเลือกจำนวน 24 ข้อ กลุ่มตัวอย่างที่ใช้ในการศึกษาครั้งนี้ ประกอบด้วย 2 กลุ่มย่อย ได้แก่ กลุ่มตัวอย่างกลุ่มอ้างอิงจำนวน 417 คน กลุ่มตัวอย่างกลุ่มใหม่จำนวน 3,126 คน และมีผู้ตรวจให้คะแนน 2 กลุ่ม คือ กลุ่ม A และกลุ่ม B วิธีการใน

การปรับเทียบคะแนนระหว่างแบบสอบใช้การปรับเทียบคะแนนแบบเส้นตรง (วิธี CLE (Chained Linear Equating))



ภาพ 2 การออกแบบการปรับเทียบ (equating designs) สำหรับแบบสอบแบบสร้างคำตอบ

ผลการศึกษาพบว่า การออกแบบการปรับเทียบคะแนนระหว่างแบบสอบประเภทสร้างคำตอบที่มีคลาดเคลื่อนมากที่สุด คือ รูปแบบที่ใช้ข้อสอบร่วมแบบสร้างคำตอบ ที่ไม่มีการตรวจให้คะแนนข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ รองลงมา คือ รูปแบบที่ใช้ข้อสอบร่วมภายนอกแบบหลายตัวเลือก (multiple-choice) ส่วนรูปแบบที่ใช้ข้อสอบร่วมแบบสร้างคำตอบ ที่มีการตรวจให้คะแนนข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ และรูปแบบที่ใช้กลุ่มผู้สอบเท่าเทียมกัน ไม่มีข้อสอบร่วม แต่มีการตรวจให้คะแนนข้ามกลุ่มในแบบสอบประเภทสร้างคำตอบในฉบับอ้างอิง (reference) มีระดับความคลาดเคลื่อนของการปรับเทียบคะแนนระหว่างแบบสอบใกล้เคียงกันมาก ซึ่งแสดงให้เห็นว่า ความลำเอียงในการปรับเทียบคะแนนระหว่างแบบสอบที่ทำให้คะแนนเกิดความคลาดเคลื่อนสามารถควบคุมได้โดยใช้การตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) เนื่องจากข้อสอบร่วมแบบหลายตัวเลือกที่มีการตรวจให้คะแนนแบบข้ามกลุ่ม (trend constructed-response anchor) มีความคลาดเคลื่อนต่ำกว่าข้อสอบร่วมแบบสร้างคำตอบที่ไม่มีการตรวจให้คะแนนแบบข้ามกลุ่ม (no-trend constructed-response anchor) โดยค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ของรูปแบบข้อสอบร่วมแบบหลายตัวเลือกที่มีการตรวจให้คะแนนแบบข้ามกลุ่ม (trend constructed-response anchor) มีค่าเท่ากับ .368 ส่วนค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ของรูปแบบข้อสอบร่วมแบบสร้างคำตอบที่ไม่มีการตรวจให้คะแนนแบบข้ามกลุ่ม (no-trend constructed-response anchor) มีค่าเท่ากับ 1.367

Kim และ Walker (2012) ได้ศึกษาองค์ประกอบที่เหมาะสมของข้อสอบร่วมในแบบสอบรูปแบบผสมที่ประกอบด้วยข้อสอบหลายตัวเลือกและข้อสอบแบบสร้างคำตอบ โดยพิจารณาจากดัชนีความไม่แปรผันของกลุ่มประชากรย่อย (subpopulation invariance indices) การเก็บรวบรวมข้อมูลใช้รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกัน โดยใช้แบบสอบร่วม (NEAT) ซึ่งใช้ข้อสอบร่วม 2 แบบ คือ 1) ข้อสอบแบบหลายตัวเลือกอย่างเดียว 2) ข้อสอบแบบหลายตัวเลือกและข้อสอบแบบสร้างคำตอบ โดยฟังก์ชั่นของการเชื่อมต่อได้รับแยกมาจากเพศชายและเพศหญิง และฟังก์ชั่นของประชากรย่อยนี้ถูกเปรียบเทียบไปยังฟังก์ชั่นของกลุ่มตัวอย่างทั้งหมด ข้อมูลที่ใช้ในการวิจัยมาจากการทดสอบใบอนุญาตระดับชาติขนาดใหญ่ 2 การทดสอบ โดยใช้การเก็บรวบรวมข้อมูลแบบ NEAT แบบสอบประกอบด้วยข้อสอบแบบหลายตัวเลือก 24 ข้อ และข้อสอบแบบสร้างคำตอบ 12 ข้อ โดยที่ข้อสอบแบบหลายตัวเลือก 12 ข้อ และข้อสอบแบบสร้างคำตอบ 6 ข้อ เป็นข้อสอบร่วมระหว่างแบบสอบชุดใหม่ X และแบบสอบชุดเก่า Y โดยข้อสอบแบบสร้างคำตอบมีผู้ตรวจให้คะแนนคนเดียว แต่ละข้อมีช่วงคะแนน 0-2 และคูณน้ำหนักด้วย 2 ดังนั้นคะแนนเต็มของแบบทดสอบนี้เท่ากับ 72 คะแนน

ผลการวิจัยพบว่า ข้อสอบร่วมแบบหลายตัวเลือกเพียงอย่างเดียว มีความแตกต่างระหว่างฟังก์ชั่นของประชากรกลุ่มย่อยและฟังก์ชั่นของกลุ่มตัวอย่างทั้งหมดในขอบเขตของคะแนน (รวมถึงคะแนนจุดตัด) ซึ่งนำไปสู่ความขัดแย้งในการตัดสินใจว่าให้ผู้สอบผ่าน/ไม่ผ่านการทดสอบดังกล่าว ข้อสอบร่วมแบบผสมจึงเป็นทางเลือกที่ดีกว่าข้อสอบร่วมแบบหลายตัวเลือกอย่างเดียว เนื่องจากข้อสอบร่วมแบบผสมไม่มีความแปรผันระหว่างกลุ่มตัวอย่างเพศชายและเพศหญิง นอกจากนี้วิจัยดังกล่าวยังสนับสนุนว่าดัชนีความไม่แปรผันของกลุ่มประชากรย่อย (subpopulation invariance indices) สามารถเป็นตัวกำหนดความพอเพียงของข้อสอบร่วมได้

อัญชลี ศรีชลกาญ (2552) ได้เปรียบเทียบคุณภาพการปรับเทียบคะแนนตามแนวทฤษฎีการตอบสนองข้อสอบ สำหรับแบบสอบรูปแบบผสมที่ประกอบด้วยข้อสอบแบบหลายตัวเลือกที่ตรวจให้คะแนนแบบสองค่า และข้อสอบแบบสร้างคำตอบที่ตรวจให้คะแนนมากกว่า 2 ค่า โดยทำการจำลองข้อมูลตามโมเดลการคะแนนบางส่วนแบบทั่วไป (generalized partial credit model) ที่แตกต่างตามรูปแบบการเก็บรวบรวมข้อมูล สัดส่วนของข้อสอบทั้ง 2 ชนิดในข้อสอบร่วม สัดส่วนของข้อสอบ 2 ชนิดในแบบสอบ และวิธีการปรับเทียบจำนวนทั้งหมด 96 เงื่อนไข $(2 \times 4 \times 6 \times 20)$ และค่าพารามิเตอร์คุณลักษณะของข้อสอบจะสร้างขึ้นด้วยฟังก์ชั่นคณิตศาสตร์ โดยการจำลองค่าความยาก อำนาจจำแนก และการเดาของข้อสอบตามทฤษฎีการตอบสนองข้อสอบจากโมเดลโลจิสติก 3 พารามิเตอร์ โดยจะประเมินคุณภาพการปรับเทียบคะแนนจากค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE)

ผลการศึกษา พบว่า คุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมด้วยวิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกันที่เก็บรวบรวมข้อมูลจากกลุ่มผู้สอบที่มีความเท่าเทียมกัน จะมีความคลาดเคลื่อนในการปรับเทียบคะแนนค่อนข้างต่ำ (< 0.300) แต่ถ้าเก็บรวบรวมข้อมูลจากกลุ่มผู้สอบที่ไม่มีความเท่าเทียมกัน วิธีการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกันจะมีความคลาดเคลื่อนในการปรับเทียบต่ำกว่าวิธีโค้งคุณลักษณะ นอกจากนี้คุณภาพการปรับเทียบคะแนนด้วยด้วยวิธีโค้งคุณลักษณะและวิธีการปรับค่าพารามิเตอร์พร้อมกัน จะมีความคลาดเคลื่อนในการปรับเทียบคะแนนลดลง เมื่อสัดส่วนของข้อสอบในแบบสอบมีจำนวนข้อสอบที่

ตรวจให้คะแนนสองค่าเพิ่มขึ้น และเมื่อสัดส่วนของข้อสอบทั้งสองชนิดในข้อสอบร่วมมีจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าลดลง กล่าวคือ การปรับเทียบที่ใช้ข้อสอบร่วมจำนวนมากกว่าร้อยละ 20 ของข้อสอบทั้งหมด (25/2, 20/2, และ 15/2) จะให้ผลการปรับเทียบที่มีความคลาดเคลื่อนต่ำกว่าการปรับเทียบที่ใช้ข้อสอบร่วมที่มีคะแนนรวมร้อยละ 20 จากคะแนนทั้งหมด (จำนวนข้อต่ำว่าร้อยละ 20 ของข้อสอบร่วมทั้งหมด: 15/1, 10/2, 5/3) ขณะที่ผลการเปรียบเทียบคุณภาพของการปรับเทียบสำหรับแบบสอบรูปแบบผสมที่มีสัดส่วนของข้อสอบทั้งสองชนิดต่างกัน 4 สัดส่วน พบว่า ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ที่ได้จากการปรับเทียบคะแนนสำหรับแบบสอบทั้ง 4 สัดส่วนไม่มีความแตกต่างกัน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ที่ได้จากการปรับเทียบคะแนนสำหรับแบบสอบทั้ง 4 สัดส่วนไม่มีความแตกต่างกัน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ที่ได้จากการปรับเทียบคะแนนสำหรับแบบสอบทั้ง 6 สัดส่วนไม่มีความแตกต่างกัน ส่วนผลการปรับเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) สำหรับการปรับเทียบในแต่ละวิธีที่มีรูปแบบการเก็บรวบรวมข้อมูลแตกต่างกัน พบว่า มีอย่างน้อย 1 คู่ที่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งคู่ที่มีความแตกต่างกันมากที่สุด คือ การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลแบบผู้สอบต่างกลุ่มกับการปรับเทียบด้วยวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบต่างกลุ่ม

บทสรุปงานวิจัยเกี่ยวกับรูปแบบการเชื่อมต่อ (linking design) การออกแบบการปรับเทียบคะแนน (equating designs) สัดส่วนของข้อสอบร่วมในการปรับเทียบคะแนน และการตรวจให้คะแนนข้ามกลุ่ม (trend scoring method)

จากการที่ผู้วิจัยได้ทบทวนเอกสารและงานวิจัยที่เกี่ยวข้อง ผู้วิจัยได้สรุปรูปแบบการเชื่อมต่อ (linking design) และการตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) ดังนี้

การตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) คือ การตรวจให้คะแนนข้ามกลุ่มโดยผู้ตรวจกลุ่มเดิม รูปแบบการเชื่อมต่อนี้ช่วยลดความลำเอียงในการปรับเทียบคะแนนระหว่างแบบสอบที่ทำให้คะแนนเกิดความคลาดเคลื่อน สำหรับรูปแบบผู้สอบกลุ่มไม่เท่าเทียม โดยใช้ข้อสอบร่วม (NEAT) ซึ่งวิธีการตรวจให้คะแนนแบบข้ามกลุ่มโดยผู้ตรวจคนเดิมเหมาะสมในการนำมาใช้กับข้อสอบร่วมที่มีข้อสอบแบบหลายตัวเลือกปนอยู่ด้วย โดยมีการตรวจให้คะแนนแบบข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ (trend construct-response anchor) กล่าวคือ ผู้ตรวจกลุ่ม B ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม (กลุ่มใหม่ และกลุ่มอ้างอิง) ดังนั้นความลำเอียงเนื่องจากการปรับเทียบ จึงถูกควบคุมด้วยการการตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) (Kim et al., 2010)

การออกแบบการปรับเทียบคะแนน (equating designs) ที่มีความเหมาะสมกับแบบสอบรูปแบบผสมระหว่างข้อสอบแบบหลายตัวเลือกและข้อสอบแบบสร้างคำตอบ ภายใต้รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกัน โดยใช้ข้อสอบร่วม (NEAT) คือ รูปแบบที่ใช้ข้อสอบร่วมแบบผสม (ระหว่างข้อสอบแบบสร้างคำตอบ และข้อสอบแบบหลายตัวเลือก) ที่มีการตรวจให้คะแนนข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ และรูปแบบผู้สอบกลุ่มเท่าเทียมกัน (equivalent group) ที่มีการตรวจให้คะแนนข้ามกลุ่มในข้อสอบแบบสร้างคำตอบ เนื่องจากมีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อน

ยกกำลังสอง (RMSE) และค่าความลำเอียง (bias) ต่ำกว่าวิธีอื่น ๆ (Kim et al., 2010) โดยข้อสอบร่วมแบบผสมนั้นไม่ผันแปรไปตามเพศของกลุ่มตัวอย่าง (Kim & Walker, 2012)

แม้ว่ารูปแบบการใช้ผู้สอบกลุ่มเท่าเทียมกัน (equivalent group) ให้ค่าสถิติบูทสแตรป (bootstrapped statistics) ที่ดี แต่ถ้ากลุ่มผู้สอบที่ใช้ไม่เหมือนกันอาจมีการแจกแจงความสามารถที่แตกต่างกัน และไม่มีข้อมูลที่ในการนำมาปรับความแตกต่างของกลุ่ม ความแตกต่างที่เกิดขึ้นนี้ย่อมส่งผลกระทบต่อความลำเอียงในการปรับเทียบคะแนนระหว่างแบบสอบ โดยวิธีการที่ลดความแตกต่างนี้ได้ คือ ต้องใช้กลุ่มตัวอย่างขนาดใหญ่ (ศิริชัย กาญจนวาสี, 2555) ทำให้ยากต่อการนำไปใช้กับสภาพจริง งานวิจัยเกี่ยวกับการปรับเทียบคะแนน จึงมักกระทำภายใต้การออกแบบการเก็บข้อมูลแบบผู้สอบกลุ่มไม่เท่าเทียมกัน โดยใช้ข้อสอบร่วม (NEAT) โดย Angoff (1971 cited in Kim et al., 2010) กล่าวว่า รูปแบบการปรับเทียบคะแนนที่นิยมใช้ คือ รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกัน โดยใช้ข้อสอบร่วม (NEAT) โดยกลุ่มผู้สอบที่แตกต่างกันทำแบบสอบฉบับเดียว โดยแบบสอบแต่ละฉบับมีข้อสอบร่วมกันจำนวนหนึ่ง ดังนั้นข้อสอบร่วมจึงใช้เพื่อการปรับเทียบคะแนนของผู้สอบสองกลุ่มที่มีความสามารถแตกต่างกัน ซึ่งสอดคล้องกับ Kim และคณะ (2010) ที่กล่าวว่า ในสถานการณ์จริงเรามักเจอกับกลุ่มผู้สอบที่มีความแตกต่างกัน เช่น ต่างระดับชั้น ต่างโปรแกรม ต่างเวลา

นอกจากนี้ยังมีงานวิจัยที่ทำการศึกษาเกี่ยวกับสัดส่วนของข้อสอบร่วมในแบบสอบรูปแบบผสม ซึ่งพบว่า เมื่อจำนวนข้อสอบร่วมเพิ่มขึ้น ทำให้ประสิทธิภาพของการปรับเทียบคะแนนเพิ่มขึ้น เนื่องจากค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) มีค่าลดลง (Bastari, 2000 อ้างถึงใน อัญชลี ศรีกลชาญ, 2552) โดยข้อสอบร่วมรูปแบบผสมมีความเหมาะสมในการนำมาใช้ในการปรับเทียบสำหรับแบบสอบรูปแบบผสม เนื่องจากการใช้ข้อสอบร่วมแบบสร้างคำตอบเพียงอย่างเดียว หรือการใช้ข้อสอบร่วมแบบหลายตัวเลือกเพียงอย่างเดียว ทำให้ความคลาดเคลื่อนในการปรับเทียบคะแนนเพิ่มขึ้น ขณะที่การใช้ข้อสอบร่วมรูปแบบผสม ช่วยลดความคลาดเคลื่อนในการปรับเทียบคะแนน (Muraki et al., 2000; Kim et al., 2006) โดยมีข้อค้นพบเพิ่มเติมว่าสัดส่วนของข้อสอบร่วมที่มีการตรวจให้คะแนนสองค่า (DS) เพิ่มขึ้น ช่วยลดความเคลื่อนในการปรับเทียบคะแนน (Kim & Lee, 2006; อัญชลี ศรีกลชาญ, 2552) นอกจากนี้ Angoff และคณะ (1971) ได้แนะนำว่า ควรใช้ข้อสอบร่วมอย่างต่ำ 20% เมื่อแบบสอบมีจำนวนข้อสอบทั้งหมด 40 ข้อ และในกรณีที่แบบสอบมีข้อสอบจำนวนมาก จำนวนข้อสอบร่วมที่เพียงพอ คือ 30% เช่นเดียวกับ Hambleton และคณะ (1991) ที่แนะนำว่าข้อสอบร่วมควรมีจำนวนไม่น้อยกว่า20 % - 25% ของข้อสอบทั้งฉบับแต่อย่างไรก็ตามในงานวิจัยของ Kim และคณะ (2010) ที่ศึกษาการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ซึ่งมีการใช้การตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) ในข้อสอบร่วมแบบสร้างคำตอบ ได้ใช้จำนวนข้อสอบร่วม (12 ข้อ) เท่ากับจำนวนข้อสอบเฉพาะทั้งฉบับ (12 ข้อ) โดยแบ่งเป็นข้อสอบแบบหลายตัวเลือกจำนวน 8 ข้อ และข้อสอบแบบสร้างคำตอบจำนวน 4 ข้อ ซึ่งจากการทบทวนเอกสารและวิจัยที่เกี่ยวข้องยังไม่มีงานวิจัยใดที่ศึกษาจำนวนข้อสอบร่วมที่เหมาะสม/พอเพียงสำหรับการออกแบบการปรับเทียบที่มีการประยุกต์ใช้การตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) ดังนั้นผู้วิจัยจึงกำหนดให้จำนวนข้อสอบร่วมภายในที่ใช้ในแบบสอบเท่ากับจำนวนข้อสอบเฉพาะทั้งหมดเช่นเดียวกับงานวิจัยของ Kim และคณะ (2010)

ด้วยเหตุนี้ผู้วิจัยจึงต้องการศึกษาวิธีการตรวจสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ซึ่งข้อสอบร่วมที่ใช้เป็นข้อสอบร่วมรูปแบบผสมที่มีการตรวจให้คะแนนข้ามกลุ่มและไม่มีการตรวจให้คะแนนข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ แต่ในการตรวจให้คะแนนข้ามกลุ่มในงานวิจัยที่ผ่านมานั้น ผู้ตรวจกลุ่มหนึ่ง (ผู้ตรวจกลุ่ม 1) ต้องตรวจข้อสอบร่วมแบบสร้างคำตอบทั้ง 2 ฉบับ ผู้ตรวจกลุ่มดังกล่าวจึงต้องรับหน้าที่ในการตรวจข้อสอบจำนวนมาก ด้วยเหตุนี้ผู้วิจัยจึงต้องการลดภาระของผู้ตรวจกลุ่มหนึ่ง/คนหนึ่งลง แต่ยังคงใช้การตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) ด้วยการออกแบบให้ผู้ตรวจแต่ละกลุ่มตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่มอย่างละเท่าๆกัน ซึ่งงานวิจัยที่ผ่านมายังไม่ได้มีการศึกษาการออกแบบการปรับเทียบในรูปแบบดังกล่าว และเนื่องจากงานวิจัยที่ผ่านมาใช้ผู้ตรวจ 2 กลุ่ม คือ กลุ่ม A และกลุ่ม B ผู้วิจัยจึงสนใจศึกษาในกรณีที่ใช้ผู้ตรวจ 2 คน คือ ผู้ตรวจ 1 และผู้ตรวจ 2 ซึ่งถ้าผลให้การปรับเทียบที่ดี ถือว่าเป็นการเพิ่มทางเลือกให้กับการนำไปใช้ในสภาพจริง

ดังนั้นวิธีการตรวจสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ที่ทำการศึกษา มีดังนี้

1) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม โดยผู้ตรวจ 1 ตรวจแบบสอบฉบับใหม่ (new) และผู้ตรวจ 2 ตรวจแบบสอบฉบับอ้างอิง (reference)

2) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม โดยผู้ตรวจ 2 ตรวจแบบสอบฉบับใหม่ (new) และตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบกลุ่มใหม่และกลุ่มอ้างอิง

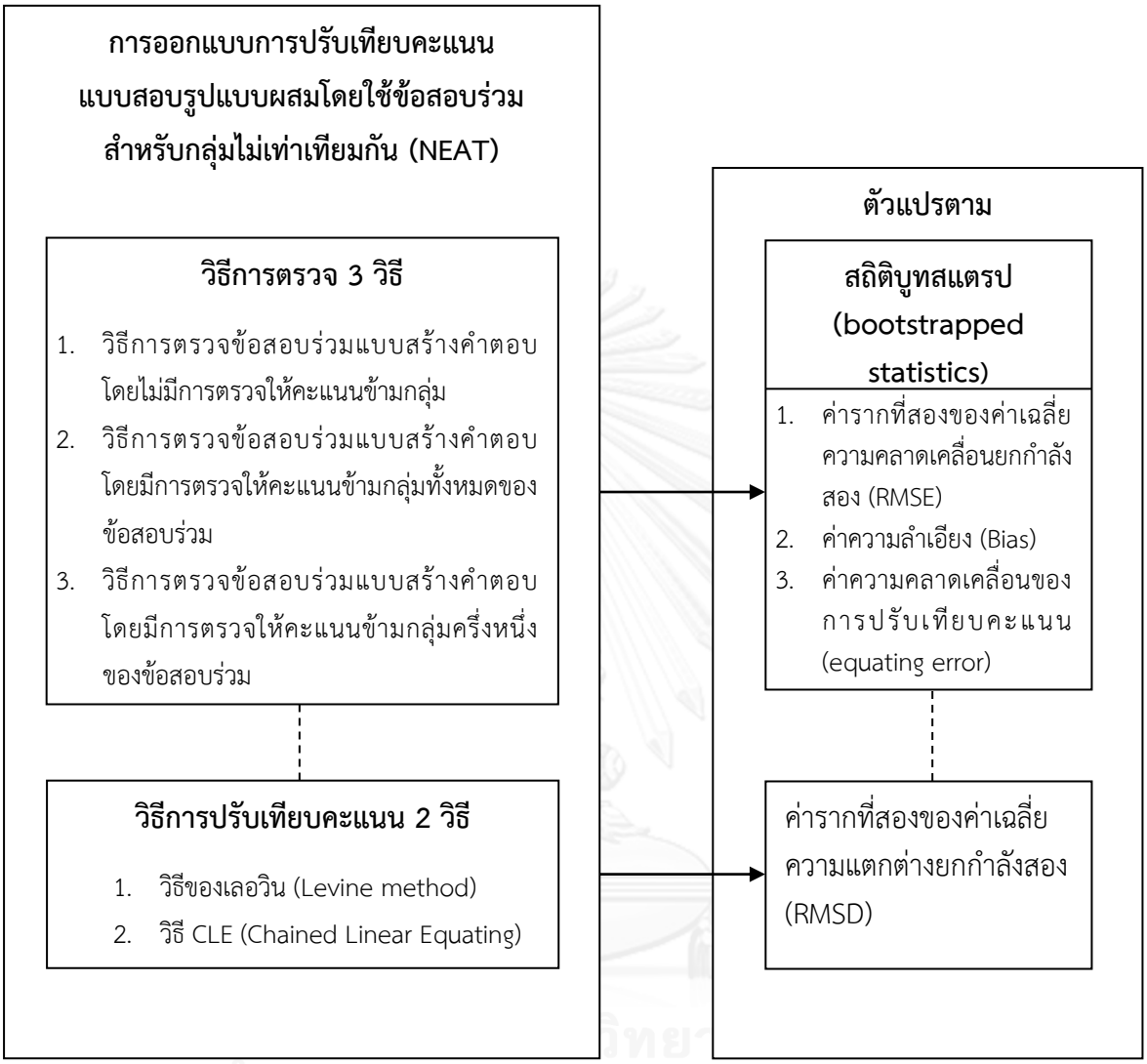
3) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม โดยผู้ตรวจ 2 ตรวจแบบสอบฉบับใหม่ (new) และตรวจข้อสอบร่วมแบบสร้างคำตอบ (2 ข้อแรก) ของผู้สอบกลุ่มใหม่และกลุ่มอ้างอิง ส่วนผู้ตรวจ 1 ตรวจแบบสอบฉบับอ้างอิง (reference) และตรวจข้อสอบร่วมแบบสร้างคำตอบ (2 ข้อหลัง) ของผู้สอบกลุ่มใหม่และกลุ่มอ้างอิง

นอกจากนี้ผู้วิจัยยังทำการเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ระหว่างวิธีการปรับเทียบเชิงเส้นตรง 2 วิธีร่วมด้วย ซึ่งได้แก่ วิธีของเลอวิน (Levine method) และวิธี CLE (Chained Linear Equating) โดยใช้รูปแบบการใช้ผู้สอบกลุ่มเดียวเป็นเกณฑ์ในการเปรียบเทียบ ซึ่งกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์ต้องทำทั้งแบบสอบฉบับใหม่ (new) และฉบับอ้างอิง (reference) ที่มีข้อสอบร่วมระหว่างแบบสอบทั้งสองชุด โดยผู้ตรวจ 2 ตรวจแบบสอบฉบับใหม่ (new) และผู้ตรวจ 1 ตรวจแบบสอบฉบับอ้างอิง (reference) ซึ่งการปรับเทียบคะแนนในรูปแบบการใช้ผู้สอบกลุ่มเดียวได้ใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed) ในการเชื่อมต่อคะแนนจากแบบสอบฉบับใหม่/ผู้ตรวจ 2 ไปยังคะแนนจากแบบสอบฉบับอ้างอิง/ผู้ตรวจ 1 ดังภาพ 3

วิธีการตรวจ	Group	Rater 1	Rater 2
เกณฑ์ (SG DESIGN)	Reference 125 คน	REF	NEW
1 : วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม	New 1,084 คน	ANCHOR	NEW
	Reference 400 คน	REF	ANCHOR
2 : วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม	New 1,084 คน	ANCHOR 4 ข้อ	NEW
	Reference 400 คน	REF	ANCHOR 4 ข้อ
3 : วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม	New 1,084 คน	ANCHOR 2 ข้อหลัง / ANCHOR 2 ข้อแรก	NEW
	Reference 400 คน	REF / ANCHOR 2 ข้อหลัง	ANCHOR 2 ข้อแรก

ภาพ 3 การออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วม สำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจที่แตกต่างกัน

กรอบแนวคิดในการวิจัย



ภาพ 4 กรอบแนวคิดในการวิจัย

CHULALONGKORN UNIVERSITY

บทที่ 3
วิธีการดำเนินการวิจัย

การวิจัยนี้เป็นการวิจัยเชิงทดลอง เพื่อเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจที่แตกต่างกัน และเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีของเลอวิน (Levine method) และวิธี CLE (Chained Linear Equating) รายละเอียดของการวิธีการวิจัย มีดังนี้

ประชากร

ประชากรในการวิจัยครั้งนี้ เป็นนักเรียนช่วงชั้นที่ 3 ระดับชั้นมัธยมศึกษาปีที่ 4 ภาคเรียนที่ 1 ปีการศึกษา 2556 ของโรงเรียนในเขตพื้นที่การศึกษาที่ 1 จังหวัดกรุงเทพมหานคร จำนวน 67 โรงเรียน จำนวน 121,710 คน เป็นนักเรียนระดับมัธยมศึกษาปีที่ 4 จำนวน 20,585 คน (ข้อมูลจากสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต 1, 2556) ดังตาราง 4

ตาราง 4 จำนวนประชากรโรงเรียน นักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 จากสำนักงานเขตการศึกษามัธยมศึกษา เขต 1 ปีการศึกษา 2556

ขนาดของโรงเรียน	จำนวนโรงเรียน	จำนวนนักเรียนชั้น ม. 4
ใหญ่พิเศษ	19	10,855
ใหญ่	20	6,488
กลาง และเล็ก	28	3,242
รวม	**67**	**20,585**

*** หมายเหตุ** โรงเรียนขนาดเล็กมีจำนวนน้อยมากจึงนำมารวมกับโรงเรียนขนาดกลาง

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการวิจัย ได้แก่ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 สำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 1 ปีการศึกษา 2556 ซึ่งได้แบ่งกลุ่มตัวอย่างออกเป็นกลุ่มๆ ดังนี้

1. กลุ่มที่ทดลองใช้เครื่องมือ (try out) จำนวน 217 คน ประกอบด้วย นักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 จากโรงเรียนสายปัญญา ในพระบรมราชินูปถัมภ์ และโรงเรียนสตรีวัดอัปสรสวรรค์ จำนวน 106 คน และ 111 คนตามลำดับ โดยผู้วิจัยแบ่งนักเรียนแต่ละห้องออกเป็น 2 กลุ่ม โดยนักเรียนกลุ่มหนึ่งทำแบบสอบฉบับใหม่ (ฉบับที่ 1) และนักเรียนอีกกลุ่มหนึ่งทำแบบสอบฉบับอ้างอิง (ฉบับที่ 2) ซึ่งนักเรียนที่ทำแบบสอบฉบับใหม่ มีทั้งหมด 108 คน ขณะที่นักเรียนที่ทำแบบสอบฉบับอ้างอิง มีทั้งหมด 109 คน รายละเอียดดังตาราง 5

ตาราง 5 กลุ่มตัวอย่างโรงเรียนที่ทำแบบสอบเรื่องงานและพลังงาน รายวิชาฟิสิกส์ ระดับชั้นมัธยมศึกษาปีที่ 4

แบบสอบ	โรงเรียน	ห้องเรียน	จำนวน	รวม
ฉบับที่ 1	โรงเรียนสายปัญญา ในพระบรมราชินูปถัมภ์	ห้อง 4/1	10	108
		ห้อง 4/2	15	
		ห้อง 4/3	27	
		รวม	52	
	โรงเรียนสตรีวัดอัปสรสวรรค์	ห้อง 4/1	15	
		ห้อง 4/2	20	
		ห้อง 4/3	21	
		รวม	56	
ฉบับที่ 2	โรงเรียนสายปัญญา ในพระบรมราชินูปถัมภ์	ห้อง 4/1	15	109
		ห้อง 4/2	11	
		ห้อง 4/3	28	
		รวม	54	
	โรงเรียนสตรีวัดอัปสรสวรรค์	ห้อง 4/1	22	
		ห้อง 4/2	14	
		ห้อง 4/3	19	
		รวม	55	

2. กลุ่มตัวอย่างที่เก็บข้อมูลจริง ซึ่งได้มาจากการสุ่มแบบหลายขั้นตอน (multi-stage random sampling) โดยมีการคำนวณขนาดของกลุ่มตัวอย่าง และการสุ่มตัวอย่างดังนี้

2.1 คำนวณขนาดกลุ่มตัวอย่างจากนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 จำนวน 20,585 คน โดยใช้จากตารางแสดงขนาดของกลุ่มตัวอย่างที่เหมาะสมสำหรับการวิจัย (ศิริชัย กาญจนวาสี และคณะ, 2537) ในระดับความเชื่อมั่น 95% $\pm$ 5% พบว่าควรใช้ขนาดกลุ่มตัวอย่าง 1,080 คน ซึ่งผู้วิจัยได้วางแผนว่าจะเก็บข้อมูลจากกลุ่มตัวอย่าง 1,600 คน ซึ่งประกอบด้วย กลุ่มใหม่ (new) จำนวน 1,200 คน กลุ่มอ้างอิง (reference) จำนวน 390 คน และกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์ 130 คน โดยในการเก็บข้อมูลจริง ผู้วิจัยได้กลุ่มตัวอย่างที่ใช้ในงานวิจัยจำนวน 1,609 คน แบ่งเป็นกลุ่มใหม่ (new) จำนวน 1,084 คน กลุ่มอ้างอิง (reference) จำนวน 400 คน และกลุ่มอ้างอิง (reference) ที่ใช้เป็นเกณฑ์จำนวน 125 คน

2.2 วิธีการสุ่มตัวอย่าง กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ได้แก่ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 สำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 1 ปีการศึกษา 2556 ที่ได้จากการสุ่มแบบหลายขั้นตอน (multi - stage random sampling) โดยสุ่มโรงเรียนแบบแบ่งชั้น (stratified random sampling) โดยมีขนาดโรงเรียนเป็นตัวแปรสำหรับแบ่งชั้น คือ ขนาดใหญ่พิเศษ ขนาดใหญ่ ขนาดกลาง และขนาดเล็ก โดยใช้จำนวนนักเรียนเป็นเกณฑ์ในการแบ่งกลุ่มขนาดโรงเรียน ซึ่งได้จำนวนประชากรโรงเรียน และนักเรียนที่แบ่งตามขนาดของโรงเรียน ดังตาราง 4 จากนั้นผู้วิจัยสุ่มโรงเรียนอย่างง่าย (simple random sampling) ตามขนาดของโรงเรียน แต่ในการเก็บข้อมูลจริงนั้นพบว่า นักเรียนมัธยมศึกษาปีที่ 4 ที่เรียนสายคณิตศาสตร์-วิทยาศาสตร์ของโรงเรียนขนาดกลาง/ขนาด

เล็กบางแห่งมีเพียงห้องเดียว นอกจากนี้นักเรียนในแต่ละห้องก็มีจำนวนต่ำกว่าที่คาดการณ์ไว้ อย่างเช่น โรงเรียนวัดประดู่ในทางธรรมและโรงเรียนไชยฉิมพลี ที่มีนักเรียนสายคณิตศาสตร์-วิทยาศาสตร์เพียงห้องเดียว โดยมีจำนวนนักเรียน 36 คน และ 17 คนตามลำดับ เป็นต้น ด้วยเหตุนี้ ผู้วิจัยจึงไม่สามารถเก็บข้อมูลจากกลุ่มตัวอย่างสำหรับโรงเรียนขนาดกลาง/ขนาดเล็กได้มากนักเมื่อเทียบกับโรงเรียนขนาดใหญ่พิเศษและขนาดใหญ่ จากการสุ่มกลุ่มตัวอย่างอย่างง่าย ได้กลุ่มตัวอย่างสำหรับกลุ่มใหม่ที่มาจากโรงเรียนขนาดใหญ่พิเศษ 4 โรงเรียน จำนวน 493 คน ขณะที่มาจากโรงเรียนขนาดใหญ่ 7 โรงเรียน จำนวน 454 คน และมาจากโรงเรียนขนาดกลาง/ขนาดเล็ก 4 โรงเรียน จำนวน 147 คน ด้วยเหตุนี้กลุ่มตัวอย่างสำหรับกลุ่มใหม่จึงมีทั้งหมด 1,084 คน จาก 15 โรงเรียน ส่วนกลุ่มตัวอย่างสำหรับกลุ่มอ้างอิงมาจากโรงเรียนขนาดใหญ่พิเศษ 1 โรงเรียน จำนวน 43 คน ขณะที่มาจากโรงเรียนขนาดใหญ่ 3 โรงเรียน จำนวน 295 คน และมาจากโรงเรียนขนาดกลาง/เล็ก 2 โรงเรียน จำนวน 62 คน ดังนั้นกลุ่มตัวอย่างสำหรับกลุ่มอ้างอิง จึงมีทั้งหมด 400 คน แต่อย่างไรก็ตาม กลุ่มตัวอย่างสำหรับกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์ มีโรงเรียนที่อนุญาตและสะดวกให้เข้าไปเก็บข้อมูลกับนักเรียนกลุ่มเดิม 2 ครั้งเพียงโรงเรียนเดียว คือ โรงเรียนสวนกุหลาบ ธนบุรี ซึ่งเป็นโรงเรียนขนาดใหญ่ จำนวน 125 คน ดังนั้นจำนวนกลุ่มตัวอย่างที่ใช้ในงานวิจัยจึงมีทั้งหมด 1,609 คน รายละเอียดดังตาราง 6

ตาราง 6 จำนวนกลุ่มตัวอย่างโรงเรียนและนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 จำแนกตามขนาดของโรงเรียน สำหรับการวิจัย สำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 1 ปีการศึกษา 2556

ขนาดของโรงเรียน	กลุ่ม	จำนวนโรงเรียน	รวมจำนวนห้องเรียน	รวมจำนวนนักเรียน
ใหญ่พิเศษ	ใหม่	4	12	493
(540 คน)	อ้างอิง	1	1	43
ใหญ่	ใหม่	7	15	454
(540 คน)	อ้างอิง	3	7	295
	อ้างอิง (เกณฑ์)	1	3	125
กลาง และเล็ก	ใหม่	4	5	137
(540 คน)	อ้างอิง	2	2	62
รวม	ใหม่	13	35	1,084
จำนวนนักเรียนใน	อ้างอิง	6	10	400
แต่ละกลุ่ม	อ้างอิง (เกณฑ์)	1	3	125
รวม				**1,609**

การออกแบบวิธีการเก็บรวบรวมข้อมูลเพื่อการปรับเทียบคะแนนของการวิจัย

รูปแบบวิธีการเก็บรวบรวมข้อมูลเพื่อการปรับเทียบคะแนนของการวิจัยครั้งนี้ คือ รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกัน โดยใช้ข้อสอบร่วม (anchor-test nonequivalent group design หรือ nonequivalence with anchor test : NEAT) ซึ่งข้อสอบร่วมมีลักษณะเป็นข้อสอบร่วมภายใน (internal anchor test) แบบแผนการปรับเทียบคะแนน ดังตาราง 7

ตาราง 7 การออกแบบวิธีการเก็บรวบรวมข้อมูลเพื่อการปรับเทียบคะแนน

กลุ่มตัวอย่าง	แบบสอบ					
	ฉบับใหม่		ฉบับอ้างอิง		ข้อสอบร่วม	
	MC	CR	MC	CR	MC	CR
กลุ่มใหม่	8 ข้อ	4 ข้อ	-	-	8 ข้อ	4 ข้อ
กลุ่มอ้างอิง	-	-	8 ข้อ	4 ข้อ	8 ข้อ	4 ข้อ

วิธีการตรวจสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

วิธีการตรวจ 3 วิธี สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ได้แก่ 1) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม โดยผู้ตรวจ 1 ตรวจแบบสอบฉบับใหม่ (new) และผู้ตรวจ 2 ตรวจแบบสอบฉบับอ้างอิง (reference) 2) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจคนหนึ่ง ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม) โดยผู้ตรวจ 2 ตรวจแบบสอบฉบับใหม่ (new) และตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบกลุ่มใหม่และกลุ่มอ้างอิง ส่วนผู้ตรวจ 1 ตรวจแบบสอบฉบับอ้างอิง (reference) ส่วนที่เหลืออยู่ 3) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม (ผู้ตรวจทั้ง 2 คน ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม โดยแบ่งกันตรวจคนละครึ่งหนึ่งของข้อสอบร่วม) โดยผู้ตรวจ 2 ตรวจแบบสอบฉบับใหม่ (new) และตรวจข้อสอบร่วมแบบสร้างคำตอบ (2 ข้อแรก) ของผู้สอบกลุ่มใหม่และกลุ่มอ้างอิง ส่วนผู้ตรวจ 1 ตรวจแบบสอบฉบับอ้างอิง (reference) และตรวจข้อสอบร่วมแบบสร้างคำตอบ (2 ข้อหลัง) ของผู้สอบกลุ่มใหม่และกลุ่มอ้างอิง

นอกจากยังมีการออกแบบการปรับเทียบที่เป็นเกณฑ์ (criterion design) ซึ่งใช้รูปแบบผู้สอบกลุ่มเดียว (single group design) โดยผู้สอบในรูปแบบดังกล่าวทำแบบสอบฉบับใหม่ (new) หลังจากนั้นผ่านไปประมาณ 3 สัปดาห์ ผู้สอบกลุ่มเดียวกันนี้ทำแบบสอบฉบับอ้างอิง (reference) ซึ่งแบบสอบทั้ง 2 ฉบับนี้เป็นแบบสอบคู่ขนานที่มีข้อสอบร่วมรูปแบบผสม โดยผู้ตรวจ 1 ตรวจแบบสอบฉบับใหม่ (new) และผู้ตรวจ 2 ตรวจแบบสอบฉบับอ้างอิง (reference) การออกแบบการปรับเทียบที่เป็นเกณฑ์นี้ใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed) ในการเชื่อมต่อคะแนนจากแบบสอบฉบับใหม่/ผู้ตรวจ 2 ไปยังคะแนนจากแบบสอบฉบับอ้างอิง/ผู้ตรวจ 1 ดังภาพ 3

วิธีการตรวจ	Group	Rater 1	Rater 2
เกณฑ์ (SG DESIGN)	Reference 125 คน	REF	NEW
1 : วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม	New 1,084 คน	ANCHOR	NEW
	Reference 400 คน	REF	ANCHOR
2 : วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม	New 1,084 คน	ANCHOR 4 ข้อ	NEW
	Reference 400 คน	REF	ANCHOR 4 ข้อ
3 : วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม	New 1,084 คน	ANCHOR 2 ข้อหลัง / ANCHOR 2 ข้อแรก	NEW
	Reference 400 คน	REF	ANCHOR 2 ข้อหลัง / ANCHOR 2 ข้อแรก

ภาพ 3 การออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วม สำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจที่แตกต่างกัน

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการศึกษาครั้งนี้ ได้แก่

1. แบบสอบวิชาฟิสิกส์รูปแบบผสม ระดับมัธยมศึกษาปีที่ 4 ภาคการศึกษาต้น ปีการศึกษา 2556 เรื่อง "งานและพลังงาน" ที่ผู้วิจัยสร้างขึ้น 2 ฉบับ ซึ่งเป็นแบบสอบคู่ขนาน ได้แก่

1.1 แบบสอบฉบับใหม่ (new) จำนวน 24 ข้อ ประกอบด้วยข้อสอบแบบหลายตัวเลือก 16 ข้อ คะแนนเต็มข้อละ 1 คะแนน และข้อสอบแบบสร้างคำตอบ 8 ข้อ คะแนนเต็มข้อละ 4 คะแนน ดังนั้นคะแนนเต็มของแบบสอบฉบับนี้จึงเท่ากับ 48 คะแนน

1.2 แบบสอบฉบับอ้างอิง (reference) จำนวน 24 ข้อ ประกอบด้วยข้อสอบแบบหลายตัวเลือก 16 ข้อ คะแนนเต็มข้อละ 1 คะแนน และข้อสอบแบบสร้างคำตอบ 8 ข้อ คะแนนเต็มข้อละ 4 คะแนน ดังนั้นคะแนนเต็มของแบบสอบฉบับนี้จึงเท่ากับ 48 คะแนน

โดยแบบสอบทั้งสองฉบับมีความเป็นคู่ขนานกันทั้งด้านเนื้อหา และความยากของแบบสอบ ซึ่งมีข้อสอบร่วมกัน จำนวน 12 ข้อ ประกอบด้วยข้อสอบแบบหลายตัวเลือกจำนวน 8 ข้อ คะแนนเต็มข้อละ 1 คะแนน และข้อสอบแบบสร้างคำตอบจำนวน 4 ข้อคะแนนเต็มข้อละ 4 คะแนน ดังนั้นคะแนนเต็มของข้อสอบร่วมทั้งหมดเท่ากับ 24 คะแนน

2. เกณฑ์การให้คะแนน (rubric score) ที่ใช้ในตรวจข้อสอบแบบสร้างคำตอบในแบบสอบวิชาฟิสิกส์ ระดับมัธยมศึกษาปีที่ 4 เรื่อง "งานและพลังงาน" ทั้ง 2 ฉบับ

3. ผู้ตรวจให้คะแนน 2 คน ได้แก่ ผู้ตรวจ 1 และผู้ตรวจ 2 โดยผู้ตรวจ 2 คือ ผู้วิจัย ส่วนผู้ตรวจ 1 คือ ครูผู้สอนวิชาฟิสิกส์ ที่มีประสบการณ์ในโรงเรียนมาแล้วอย่างต่ำ 3 ปี โดยรายละเอียดเกี่ยวกับผู้ตรวจให้คะแนนอยู่ในส่วนวิธีคัดเลือกผู้ตรวจให้คะแนน และการเตรียมความพร้อมให้กับผู้ตรวจให้คะแนน

วิธีการดำเนินการสร้างแบบสอบ

แบบสอบฉบับใหม่ (new) และอ้างอิง (reference) เป็นแบบสอบรูปแบบผสมที่เป็นคู่ขนานกันทั้งด้านเนื้อหา และความยาก ประกอบด้วยข้อสอบแบบหลายตัวเลือกและข้อสอบแบบสร้างคำตอบ เรื่อง "งานและพลังงาน" ระดับมัธยมศึกษาปีที่ 4 ภาคการศึกษาต้น ปีการศึกษา 2556 โดยมีวิธีดำเนินการสร้างดังนี้

1. ศึกษาหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 ที่กำหนดโดยกระทรวงศึกษาธิการ สาระที่ 4 แรงและการเคลื่อนที่

มาตรฐาน ว 4.2 คือ เข้าใจลักษณะการเคลื่อนที่แบบต่างๆของวัตถุในธรรมชาติ มีกระบวนการเสาะหาความรู้และจิตวิทยาศาสตร์ สื่อสารสิ่งที่เรียนรู้และนำความรู้ไปใช้ประโยชน์

ตัวชี้วัดชั้นปี คือ ทดลองและอธิบายความสัมพันธ์ระหว่างแรงกับการเคลื่อนที่ของวัตถุในสนามโน้มถ่วง และอธิบายการนำความรู้ไปใช้ประโยชน์ (มฐ ว 4.2 ตัวชี้วัดข้อ 1)

2. ศึกษาเนื้อหาในเรื่องงานและพลังงานของวิชาฟิสิกส์ ระดับชั้นมัธยมศึกษาปีที่ 4

งานและพลังงาน ประกอบด้วย 4 ประเด็นหลัก ได้แก่

1. งาน: งานเกิดขึ้นก็ต่อเมื่อมีแรงกระทำต่อวัตถุนั้นแล้ววัตถุเกิดการเคลื่อนที่ งานหาได้จากผลคูณระหว่างแรงที่กระทำกับระยะการเคลื่อนที่ ซึ่งงานเป็นปริมาณสเกลาร์และมีหน่วยเป็นจูล วัตถุที่เคลื่อนที่ไปตามทิศทางของแรงที่กระทำจะได้งานเป็นบวก แต่ถ้าวัตถุที่เคลื่อนที่ในทิศตรงข้ามกับแรงที่กระทำจะได้งานเป็นลบ

2. กำลัง: กำลังหาได้จากอัตราส่วนของงานที่ทำได้กับเวลาที่ใช้ในการทำงาน หน่วยของกำลังเป็นหน่วยของจูล/วินาที หรือวัตต์

3. พลังงาน: พลังงานเป็นผลเนื่องมาจากการทำงาน โดยงานเป็นปริมาณที่บอกค่าพลังงานที่เปลี่ยนจากรูปหนึ่งไปยังอีกรูปหนึ่ง พลังงานสามารถแบ่งออกเป็น 2 ประเภทใหญ่ ๆ คือ พลังงานศักย์โน้มถ่วง และพลังงานจลน์

4. กฎการอนุรักษ์พลังงานกล: พลังงานไม่อาจทำให้สูญหาย หรือเพิ่มได้ เนื่องจากพลังงาน มีเพียงการเปลี่ยนรูปพลังงานจากรูปหนึ่งไปสู่อีกรูปหนึ่ง

3. กำหนดวัตถุประสงค์เชิงพฤติกรรมให้มีความสอดคล้องกับมาตรฐานการเรียนรู้ และตัวชี้วัด ดังนี้

เมื่อเรียนจบบทเรียน นักเรียนสามารถ

1. อธิบายงานและวิเคราะห์งานของแรงต่างๆได้

2. คำนวณงานที่กระทำจากผลคูณระหว่างขนาดของแรงในแนวที่ทำให้วัตถุเคลื่อนที่กับการกระจัดที่วัตถุเคลื่อนที่ หรือหางานจากพื้นที่ใต้กราฟระหว่างแรงกับการกระจัดได้

3. คำนวณหากำลังจากโจทย์ที่กำหนดให้ได้

4. อธิบายพลังงาน พลังงานจลน์ พลังงานศักย์ และความสัมพันธ์ระหว่างงานและพลังงานได้

5. คำนวณหาพลังงานจลน์ พลังงานศักย์โน้มถ่วงของวัตถุที่ระดับต่างๆ และพลังงานศักย์ยืดหยุ่นของสปริง

6. อธิบายกฎการอนุรักษ์พลังงานกลในการวิเคราะห์การเคลื่อนที่ในสถานการณ์ต่าง ๆ

7. ประยุกต์ใช้กฎการอนุรักษ์พลังงานกับการเคลื่อนในสถานการณ์ต่าง ๆ ได้

4. สร้างผังข้อสอบ (test blueprint) วิชาฟิสิกส์ ระดับชั้นมัธยมศึกษาปีที่ 4 เรื่อง งานและพลังงาน ดังตาราง 8

5. สร้างแบบสอบวิชาฟิสิกส์ระดับมัธยมศึกษาปีที่ 4 เรื่อง "งานและพลังงาน" จำนวน 2 ฉบับ ฉบับละ 24 ข้อ ที่มีความเป็นคู่ขนานกัน ซึ่งแบ่งเป็นข้อสอบแบบหลายตัวเลือกจำนวน 16 ข้อ และข้อสอบแบบสร้างคำตอบจำนวน 8 ข้อ

6. นำแบบสอบจากข้อ 5 ไปตรวจสอบคุณภาพขั้นต้น โดยให้ผู้เชี่ยวชาญทางการวัดผล และครูผู้สอนวิชาฟิสิกส์ จำนวน 5 ท่าน (รายชื่ออยู่ในภาคผนวก ก) พิจารณาความสอดคล้องของข้อสอบกับมาตรฐานการเรียนรู้ และผลการเรียนรู้ที่คาดหวัง เพื่อตรวจสอบความตรงเชิงเนื้อหา ด้วยการหาดัชนีความสอดคล้องของข้อสอบ (item objective congruence: IOC) พร้อมทั้งปรับปรุงตามคำแนะนำของผู้เชี่ยวชาญ

จากการตรวจสอบความตรงเนื้อหาของแบบสอบฉบับใหม่ (new) และอ้างอิง (reference) โดยผู้ทรงคุณวุฒิด้านการวัดและประเมินผลการศึกษา และผู้ทรงคุณวุฒิทางด้านการสอนวิชาฟิสิกส์ จำนวน 5 คน สรุปผลกการตรวจสอบได้ดังนี้

ผลการตรวจสอบความสอดคล้องระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรมและระดับพฤติกรรม ของแบบสอบวิชาฟิสิกส์เรื่องงานและพลังงาน ฉบับใหม่ (new) พบว่า ข้อสอบทุกข้อวัดได้ตรงตามจุดประสงค์และระดับพฤติกรรม พิจารณาค่าดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรมและระดับพฤติกรรม ซึ่งพบว่ามีค่าอยู่ระหว่าง .60 - 1.00 รายละเอียดอยู่ในภาคผนวก ข แต่มีผู้เชี่ยวชาญบางท่านได้เสนอแนะให้ปรับปรุงแก้ไขข้อสอบบางข้อ ซึ่งผู้วิจัยได้ปรับปรุงแก้ไขดังปรากฏในภาคผนวก ข

ผลการตรวจสอบความสอดคล้องระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรมและระดับพฤติกรรม ของแบบสอบวิชาฟิสิกส์เรื่องงานและพลังงานฉบับอ้างอิง (reference) พบว่า ข้อสอบทุกข้อวัดได้ตรงตามจุดประสงค์และระดับพฤติกรรม พิจารณาค่าดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรมและระดับพฤติกรรม ซึ่งพบว่ามีค่าอยู่ระหว่าง .60 - 1.00 รายละเอียดอยู่ในภาคผนวก ข แต่มีผู้เชี่ยวชาญบางท่านได้เสนอแนะให้ปรับปรุงแก้ไขข้อสอบบางข้อ ซึ่งผู้วิจัยได้ปรับปรุงแก้ไขดังปรากฏในภาคผนวก ข

จากการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบทั้งสองฉบับโดยผู้เชี่ยวชาญ ทำให้ได้แบบสอบที่นำไปทดลองใช้ (tryout) จำนวน 24 ข้อ ซึ่งประกอบด้วย ข้อสอบแบบเลือกตอบจำนวน 16 ข้อ และข้อสอบแบบสร้างคำตอบจำนวน 8 ซึ่งสาเหตุที่ผู้วิจัยไม่ได้สร้างแบบสอบที่นำไปทดลองใช้ (try out) ให้มีจำนวนมากกว่าที่ต้องการ เนื่องจาก ในสถานการณ์จริงผู้วิจัยมีเวลาในการทดสอบนักเรียนภายในคาบเรียน เพียง 100 นาทีเท่านั้น ซึ่งเมื่อพิจารณาแล้วจำนวนข้อสอบดังกล่าวมีความเหมาะสมสำหรับการนำไปทดสอบจริง

7. นำแบบสอบที่ได้จากข้อ 6 ไปทดลองกับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 ที่ไม่ใช่กลุ่มตัวอย่าง จำนวน 217 คน ซึ่งประกอบด้วย นักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 จากโรงเรียนสายปัญญา ในพระบรมราชินูปถัมภ์ และโรงเรียนสตรีวัดอัปสรสวรรค์ จำนวน 106 คน และ 111 คน ตามลำดับ โดยผู้วิจัยแบ่งนักเรียนแต่ละห้องออกเป็น 2 กลุ่ม โดยนักเรียนกลุ่มหนึ่งทำแบบสอบฉบับใหม่ (ฉบับที่ 1) และนักเรียนอีกกลุ่มหนึ่งทำแบบสอบฉบับอ้างอิง (ฉบับที่ 2) ซึ่งนักเรียนที่ทำแบบสอบฉบับใหม่มีทั้งหมด 108 คน ขณะที่นักเรียนที่ทำแบบสอบฉบับอ้างอิงมีทั้งหมด 109 คน

8. ตรวจให้คะแนนแบบสอบ โดยข้อสอบแบบหลายตัวเลือกมีคะแนนเต็มข้อละ 1 คะแนน ตอบถูกได้ 1 คะแนน ตอบผิดได้ 0 คะแนน ข้อสอบแบบสร้างคำตอบมีคะแนนข้อละ 4 คะแนน

ตาราง 8 ตารางวิเคราะห์หลักสูตร เรื่องงานและพลังงาน รายวิชาฟิสิกส์ ระดับชั้นมัธยมศึกษาปีที่ 4

เนื้อหา	จุดมุ่งหมายเชิงพฤติกรรม	จำ	เข้าใจ	นำไปใช้	วิเคราะห์	สังเคราะห์	ประเมินค่า	รวม	อันดับความสำคัญ
งาน	เมื่อเรียนจบบทเรียนแล้ว นิสิตสามารถ 1. อธิบายงานและวิเคราะห์งานของแรงต่างๆได้		2					2	4
	2. คำนวณหางานในการเคลื่อนที่ของวัตถุได้		(4)	1	1(4)			10	1
กำลัง	3. คำนวณหากำลังจากโจทย์ที่กำหนดให้ได้			2	(4)			6	2
พลังงาน	4. อธิบายพลังงาน พลังงานจลน์ พลังงานศักย์ และความสัมพันธ์ระหว่างงานและพลังงานได้		2					2	4
	5. คำนวณหาพลังงานจลน์ พลังงานศักย์โน้มถ่วงของวัตถุที่ระดับต่างๆ และพลังงานศักย์ยืดหยุ่นของสปริง			2	(4)			6	2
	6. คำนวณโจทย์เกี่ยวกับความสัมพันธ์ระหว่างงานและพลังงานได้			2	(4)			6	2
กฎการอนุรักษ์พลังงานกล	7.อธิบายกฎการอนุรักษ์พลังงานกลในการวิเคราะห์การเคลื่อนที่ในสถานการณ์ต่าง ๆ		1					1	5
	8. ประยุกต์ใช้กฎการอนุรักษ์พลังงานกับการเคลื่อนในสถานการณ์ต่าง ๆ ได้			2 (4)	(4)			10	1
เครื่องกล	9. คำนวณโจทย์เกี่ยวกับเครื่องกลโดยอาศัยหลักการทำงานของเครื่องกลได้			1(4)				5	3
รวม			9	18	21	-	-	48	
อันดับความสำคัญ			3	2	1				

****หมายเหตุ** ข้อสอบอัตนัยในจุดประสงค์เชิงพฤติกรรมที่ 2, 3, 5, 6, 8 และ 9 มีคะแนนเต็มข้อละ 4 คะแนน โดยที่ ( ) คือ คะแนนเต็ม

9. นำข้อสอบ และคะแนนของนักเรียนจากข้อ 8 มาวิเคราะห์หาค่าความยาก (b) และค่าอำนาจจำแนก (a) ความเที่ยงของแบบสอบ โดยข้อสอบแบบหลายตัวเลือกใช้โปรแกรม TAP โดยใช้แบ่งกลุ่มแบบ 25% ขณะที่ข้อสอบแบบสร้างคำตอบใช้โปรแกรม B-INDEX ในวิเคราะห์ข้อสอบโดยใช้แบ่งกลุ่มแบบ 25% ซึ่งในการวิเคราะห์ข้อสอบใช้เกณฑ์การแปลความหมายผลการวิเคราะห์ (ศิริชัย กาญจนวาสี, 2552) ดังนี้

อำนาจจำแนกของข้อสอบ (r)

.60 - 1.00	หมายถึง	จำแนกได้ดีมาก
.40 - .59	หมายถึง	จำแนกได้ดี
.20 - .39	หมายถึง	จำแนกได้พอใช้
.10 - .19	หมายถึง	จำแนกได้ค่อนข้างต่ำ ควรได้รับการปรับปรุง
.0 - .09	หมายถึง	จำแนกได้ต่ำมาก ควรตัดทิ้ง

ความยากง่ายของข้อสอบ (p)

.80 - 1.00	หมายถึง	ง่ายมาก
.60 - .79	หมายถึง	ค่อนข้างง่าย
.40 - .59	หมายถึง	ปานกลาง
.20 - .39	หมายถึง	ค่อนข้างยาก
0 - .19	หมายถึง	ยากมาก

9.1 ผลการวิเคราะห์คุณภาพของข้อสอบแบบหลายตัวเลือกรายข้อ

ผลการวิเคราะห์ข้อสอบแบบหลายตัวเลือกจำนวน 16 ข้อ จากแบบสอบเรื่องงานและพลังงานทั้ง 2 ฉบับ จากกลุ่มตัวอย่างจำนวน 217 คน โดยแบ่งเป็นกลุ่มที่ทำแบบสอบฉบับใหม่จำนวน 108 คน และกลุ่มที่ทำแบบสอบฉบับอ้างอิง จำนวน 109 คน พบว่า ค่าเฉลี่ยของข้อสอบแบบหลายตัวเลือกทั้งหมดในแบบสอบฉบับใหม่และฉบับอ้างอิง คือ 6.713 และ 7.404 คะแนน ส่วนเบี่ยงเบนมาตรฐาน คือ 2.851 และ 2.896 โดยที่ข้อสอบแบบหลายตัวเลือกทั้งหมดในแบบสอบฉบับอ้างอิง มีค่าเฉลี่ยของแบบสอบสูงกว่าแบบสอบฉบับใหม่ เมื่อพิจารณาค่าความเบ้ของแบบสอบ (.092 และ .301) พบว่าข้อสอบแบบหลายตัวเลือกในแบบสอบฉบับใหม่มีลักษณะแจกแจงแบบปกติ แสดงว่า นักเรียนส่วนใหญ่ที่ทำข้อสอบแบบหลายตัวเลือกในแบบสอบฉบับดังกล่าวได้คะแนนใกล้เคียงกับค่าเฉลี่ย ขณะที่ข้อสอบแบบหลายตัวเลือกในแบบสอบฉบับอ้างอิง มีลักษณะการแจกแจงแบบเบ้ขวาเล็กน้อย แสดงว่า นักเรียนส่วนใหญ่ที่ทำข้อสอบแบบหลายตัวเลือกในแบบสอบฉบับดังกล่าวได้คะแนนต่ำกว่าค่าเฉลี่ยเล็กน้อย และเมื่อพิจารณาค่าความโด่ง (-.582 และ -.479) พบว่าข้อสอบแบบหลายตัวเลือกในแบบสอบทั้งสองฉบับมีการกระจายแบนกว่าโค้งปกติ และเมื่อพิจารณาค่าความเที่ยงของข้อสอบแบบหลายตัวเลือกทั้งฉบับ KR20 ในแบบสอบฉบับใหม่และฉบับอ้างอิง มีค่าอยู่ในระดับปานกลาง คือ .634 และ .602 ตามลำดับ แสดงว่าข้อสอบแบบหลายตัวเลือกทั้งหมดในแบบสอบทั้ง 2 ฉบับ มีความคงเส้นคงวาของคะแนนที่ได้จากการสอบเมื่อทำการสอบซ้ำ ๆ ในระดับปานกลาง รายละเอียดดังตาราง 9

ผลการวิเคราะห์ข้อสอบแบบหลายตัวเลือกรายข้อในแบบสอบฉบับใหม่ (ฉบับที่ 1) พบว่า ค่าความยากของข้อสอบแบบหลายตัวเลือกในแบบสอบดังกล่าวมีค่าอยู่ระหว่าง .19 - .80 ค่าความยากเฉลี่ย คือ .420 ค่าอำนาจจำแนกของข้อสอบแบบหลายตัวเลือกในแบบสอบดังกล่าวมีค่าอยู่ระหว่าง .18 - .56 ค่าอำนาจจำแนกเฉลี่ย คือ .400 (ผลการวิเคราะห์ข้อสอบรายข้ออยู่ในภาคผนวก ง) แสดงว่าข้อสอบแบบหลายตัวเลือกทั้งฉบับในแบบสอบฉบับใหม่ มีระดับความยากอยู่ในระดับค่อนข้างง่ายไปจนถึงยากมาก และมีความยากเฉลี่ยของข้อสอบแบบหลายตัวเลือกทั้งฉบับในแบบสอบดังกล่าวอยู่ในระดับปานกลาง รวมทั้งมีอำนาจจำแนกอยู่ในระดับค่อนข้างต่ำจนถึงดี และมีอำนาจจำแนกเฉลี่ยของข้อสอบแบบหลายตัวเลือกทั้งฉบับในแบบสอบดังกล่าวอยู่ในระดับดี

ผลการวิเคราะห์ข้อสอบแบบหลายตัวเลือกรายข้อในแบบสอบฉบับอ้างอิง (ฉบับที่ 2) พบว่า ค่าความยากของข้อสอบแบบหลายตัวเลือกในแบบสอบดังกล่าวมีค่าอยู่ระหว่าง .28 - .75 ค่าความยากเฉลี่ย คือ .463 ค่าอำนาจจำแนกของข้อสอบแบบหลายตัวเลือกในแบบสอบดังกล่าวมีค่าอยู่ระหว่าง -.13 - .65 ค่าอำนาจจำแนกเฉลี่ย คือ .384 (ผลการวิเคราะห์ข้อสอบรายข้ออยู่ในภาคผนวก ง) แสดงว่าข้อสอบแบบหลายตัวเลือกฉบับในแบบสอบฉบับใหม่ (new) มีระดับความยากอยู่ในระดับค่อนข้างง่ายไปจนถึงยากมาก และมีความยากเฉลี่ยของข้อสอบแบบหลายตัวเลือกทั้งฉบับในแบบสอบดังกล่าวอยู่ในระดับปานกลาง รวมทั้งมีอำนาจจำแนกตั้งแต่จำแนกไม่ได้ไปจนถึงจำแนกได้ดีมาก และมีอำนาจจำแนกเฉลี่ยของข้อสอบแบบหลายตัวเลือกทั้งฉบับในแบบสอบดังกล่าวอยู่ในระดับพอใช้

ตาราง 9 ผลการวิเคราะห์ค่าสถิติข้อสอบแบบหลายตัวเลือกในแบบสอบเรื่องงานและพลังงานฉบับใหม่และฉบับอ้างอิงจากการทดลองกับกลุ่มตัวอย่างจำนวน 217 คน

ค่าสถิติ	แบบสอบฉบับใหม่	แบบสอบฉบับอ้างอิง
จำนวนนักเรียน (คน)	108	109
ค่าเฉลี่ย	6.713	7.404
ส่วนเบี่ยงเบนมาตรฐาน	2.851	2.896
ความเบ้	.092	.301
ความโด่ง	-.582	-.479
ความยากเฉลี่ย	.420	.463
อำนาจจำแนกเฉลี่ย	.400	.384
ความเที่ยงของแบบสอบ KR20	.634	.602

9.2 ผลการวิเคราะห์ข้อสอบรายข้อของข้อสอบแบบสร้างคำตอบ

ผลการวิเคราะห์ข้อสอบแบบสร้างคำตอบจำนวน 8 ข้อ จากแบบสอบเรื่องงานและพลังงานทั้ง 2 ฉบับ จากกลุ่มตัวอย่างจำนวน 217 คน โดยแบ่งเป็นกลุ่มที่ทำแบบสอบฉบับใหม่จำนวน 108 คน และกลุ่มที่ทำแบบสอบฉบับอ้างอิง จำนวน 109 คน พบว่า ค่าเฉลี่ยของข้อสอบแบบสร้างคำตอบทั้งหมดในแบบสอบฉบับใหม่และอ้างอิง คือ 6.685 และ 6.615 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐาน คือ 6.462 และ 6.979 โดยที่ข้อสอบแบบสร้างคำตอบทั้งหมดในแบบสอบฉบับอ้างอิงมีค่าเฉลี่ยของแบบสอบสูงกว่าแบบสอบฉบับใหม่ และเมื่อพิจารณาค่าความเชื่อมั่นของข้อสอบแบบ

หลายตัวเลือกทั้งฉบับ (coefficient alpha) ในแบบสอบฉบับใหม่และอ้างอิง มีค่าอยู่ในระดับสูง คือ .824 และ .872 ตามลำดับ แสดงว่าข้อสอบแบบหลายตัวเลือกทั้งหมดในแบบสอบทั้ง 2 ฉบับ มีความคงเส้นคงวาของคะแนนที่ได้จากการสอบเมื่อทำการสอบซ้ำๆในระดับสูง รายละเอียดดังตาราง 10

ผลการวิเคราะห์ข้อสอบแบบสร้างคำตอบในแบบสอบฉบับใหม่ (ฉบับที่ 1) พบว่า ค่าความยากของข้อสอบแบบสร้างคำตอบในแบบสอบดังกล่าวมีค่าอยู่ระหว่าง .07 - .43 ค่าความยากเฉลี่ย คือ .258 ค่าอำนาจจำแนกของข้อสอบแบบสร้างคำตอบในแบบสอบดังกล่าวมีค่าอยู่ระหว่าง .14 - .83 ค่าอำนาจจำแนกเฉลี่ย คือ .454 (ผลการวิเคราะห์ข้อสอบรายข้ออยู่ในภาคผนวก ง) แสดงว่าข้อสอบแบบสร้างคำตอบทั้งฉบับในแบบสอบฉบับใหม่ มีระดับความยากอยู่ในระดับค่อนข้างปานกลางไปจนถึงยากมาก และมีความยากเฉลี่ยของข้อสอบแบบสร้างคำตอบทั้งฉบับในแบบสอบดังกล่าวอยู่ในระดับค่อนข้างยาก รวมทั้งมีอำนาจจำแนกอยู่ในระดับต่ำมากจนถึงดีมาก และมีอำนาจจำแนกเฉลี่ยของข้อสอบแบบสร้างคำตอบทั้งฉบับในแบบสอบดังกล่าวอยู่ในระดับพอใช้

ผลการวิเคราะห์ข้อสอบแบบสร้างคำตอบในแบบสอบฉบับใหม่ (ฉบับที่ 2) พบว่า ค่าความยากของข้อสอบแบบสร้างคำตอบในแบบสอบดังกล่าวมีค่าอยู่ระหว่าง .08 - .45 ค่าความยากเฉลี่ย คือ .268 ค่าอำนาจจำแนกของข้อสอบแบบสร้างคำตอบในแบบสอบดังกล่าวมีค่าอยู่ระหว่าง .16 - .88 ค่าอำนาจจำแนกเฉลี่ย คือ .513 (ผลการวิเคราะห์ข้อสอบรายข้ออยู่ในภาคผนวก ง) แสดงว่า ข้อสอบแบบสร้างคำตอบทั้งฉบับในแบบสอบฉบับใหม่ มีระดับความยากอยู่ในระดับปานกลางไปจนถึงยากมาก และมีความยากเฉลี่ยของข้อสอบแบบสร้างคำตอบทั้งฉบับในแบบสอบดังกล่าวอยู่ในระดับค่อนข้างยาก รวมทั้งมีอำนาจจำแนกอยู่ในระดับต่ำมากจนถึงดีมาก และมีอำนาจจำแนกเฉลี่ยของข้อสอบแบบสร้างคำตอบทั้งฉบับในแบบสอบดังกล่าวอยู่ในระดับพอใช้

ตาราง 10 ผลการวิเคราะห์ค่าสถิติข้อสอบแบบสร้างคำตอบในแบบสอบเรื่องงานและพลังงานฉบับใหม่และอ้างอิง จากการทดลองกับกลุ่มตัวอย่างจำนวน 217 คน

ค่าสถิติ	แบบสอบฉบับใหม่	แบบสอบฉบับอ้างอิง
จำนวนนักเรียน (คน)	108	109
ค่าเฉลี่ย	6.685	6.615
ส่วนเบี่ยงเบนมาตรฐาน	6.462	6.979
ความยากเฉลี่ย	.258	.268
อำนาจจำแนกเฉลี่ย	.454	.513
ความเที่ยงของแบบสอบ KR20	.824	.872

10. นำคะแนนของนักเรียนแต่ละคนที่ทำแบบสอบแต่ละฉบับ มาวิเคราะห์ความแตกต่างระหว่างคะแนนเฉลี่ยของนักเรียนในแบบสอบฉบับใหม่และอ้างอิง เพื่อตรวจสอบความเป็นคู่ขนานของแบบสอบทั้ง 2 ฉบับ โดยใช้ด้วยสถิติ t-test แบบ Paired-sample โดยใช้โปรแกรม SPSS V.17

จากการทดสอบ พบว่า คะแนนของนักเรียนที่ทำแบบสอบแต่ละฉบับไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (p = .06) รายละเอียดดังตาราง 11

ตาราง 11 การทดสอบความแตกต่างของค่าเฉลี่ยคะแนนในแบบสอบฉบับใหม่และฉบับอ้างอิงของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4

แบบสอบ	ค่าสถิติพื้นฐาน		95% Confidence Interval of the Difference		t	p
	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน	Lower	Upper		
ฉบับใหม่	6.71	2.87	-1.39	.02	-1.92	.06
ฉบับอ้างอิง	7.40	2.92				

11. หลังจากนั้นคัดเลือกข้อสอบที่มีความยาก (b) ในช่วง .20 - .80 และค่าอำนาจจำแนกตั้งแต่ 0.20 ขึ้นไป เพื่อจัดทำเป็นแบบสอบฉบับใหม่ และอ้างอิงและเนื่องจากแบบสอบฉบับทดลองที่ผู้วิจัยนำไปใช้จริงมีจำนวนพอดีกับความต้องการ ในกรณีที่ข้อสอบข้อใดมีค่าความยากและค่าอำนาจผู้วิจัยทำการสร้างข้อสอบใหม่เพื่อทดแทนข้อสอบเดิม โดยข้อสอบที่สร้างขึ้นใหม่นี้ยึดตามจุดประสงค์เชิงพฤติกรรมที่ได้สร้างไว้

11.1 การคัดเลือกข้อสอบแบบหลายตัวเลือก และแนวทางในการปรับปรุงข้อสอบ

จากการพิจาณาผลการวิเคราะห์คุณภาพของข้อสอบแบบหลายตัวเลือกในแบบสอบฉบับใหม่ (new) ตามเกณฑ์ที่กำหนดไว้ พบว่า ข้อสอบที่ต้องได้รับการปรับปรุง ได้แก่ ข้อ 1 เนื่องจากข้อสอบข้อนี้ยากเกินไป แต่จำแนกได้พอใช้ (p = .19 และ r = .30) และ ข้อ 11 เนื่องจากข้อสอบข้อนี้ค่อนข้างยาก และจำแนกได้ต่ำ (p = .24 และ r = .18) และเมื่อพิจาณาข้อสอบแบบ หลายตัวเลือกในแบบสอบฉบับอ้างอิง พบว่า ข้อสอบที่ต้องได้รับการปรับปรุงได้แก่ ข้อ 8 เนื่องจากข้อสอบข้อนี้ยากปานกลาง แต่จำแนกไม่ได้ (p = .46 และ r = -.13) และข้อ 11 เนื่องจากข้อสอบข้อนี้ค่อนข้างยาก แต่จำแนกต่ำ (p = .33 และ r = .19)

ผลจากการวิเคราะห์พบว่า ข้อ 11 ต้องได้รับการปรับปรุงทั้งสองฉบับ ซึ่งแสดงถึงความสอดคล้องของผลการวิเคราะห์คุณภาพของข้อสอบ นั่นคือ ข้อสอบข้อนี้ค่อนข้างยาก ส่งผลให้ค่าอำนาจจำแนกต่ำ ผู้วิจัยจึงทำการปรับปรุงข้อสอบให้ง่ายลง แต่ยังยึดตามจุดประสงค์เชิงพฤติกรรมและระดับพฤติกรรมไว้เช่นเดิม (รายละเอียดของข้อสอบที่ทำการแก้ไข อยู่ในภาคผนวก ง) แต่เมื่อพิจาณาข้อ 1 พบว่า ข้อสอบข้อดังกล่าวสามารถใช้ได้ในแบบสอบฉบับใหม่ แต่ไม่สามารถใช้ได้ในแบบสอบฉบับอ้างอิงทั้งที่ข้อ 1 ในแบบสอบทั้งสองฉบับเป็นคำถามเดียวกัน เพียงแต่สลับตำแหน่งของตัวเลือกเท่านั้น ซึ่งจากการทดสอบทางสถิติเกี่ยวกับความเป็นคู่ขนานของแบบสอบ พบว่า แบบสอบทั้งสองฉบับมีความเป็นคู่ขนาน ถ้าหากทำการปรับปรุงข้อสอบข้อนี้ ก็ต้องทำการปรับปรุงทั้งสองฉบับ ซึ่งต้องเสียข้อ 1 ที่ใช้ได้ในแบบสอบฉบับใหม่ไป ผู้วิจัยจึงเห็นว่าควรให้ข้อ 1 ในแบบสอบฉบับใหม่เป็นข้อสอบร่วม ซึ่งทำให้ไม่ต้องสร้างข้อสอบข้อ 1 อีกข้อขึ้นมาทดแทน ขณะที่ข้อ 8 ที่สามารถใช้ได้ในแบบสอบฉบับใหม่ แต่ไม่สามารถใช้ได้แบบสอบฉบับอ้างอิงนั้น ลักษณะของข้อคำถามมีความใกล้เคียงกัน คือ ถามเกี่ยวกับความสัมพันธ์ระหว่างพลังงานกับตัวแปรค่าหนึ่ง ซึ่งเมื่อผู้วิจัยพิจารณาอย่างถี่ถ้วนแล้ว พบว่า ข้อคำถามในฉบับอ้างอิงมีความซับซ้อนมากกว่าข้อคำถามในฉบับใหม่อยู่เล็กน้อย กล่าวคือ ข้อคำถามในฉบับใหม่เป็นการถามถึงความสัมพันธ์ระหว่างพลังงานศักย์โน้มถ่วงกับตัวแปรหนึ่ง (การกระจัด S) ซึ่งเป็นตัวแปรที่อยู่ในสมการของพลังงานศักย์โน้มถ่วงอยู่แล้ว ขณะที่

ข้อคำถามในฉบับอ้างอิงเป็นการถามถึงความสัมพันธ์ระหว่างพลังงานจลน์กับตัวแปรหนึ่ง (การกระจัด S) ซึ่งไม่ได้เป็นตัวแปรที่อยู่ในสมการของพลังงานจลน์โดยตรง แต่มีความสัมพันธ์กับอัตราเร็วในสมการดังกล่าว ซึ่งนักเรียนต้องใช้ความรู้เดิมเกี่ยวกับการเคลื่อนที่ในหนึ่งมิติเข้ามาพิจารณาด้วย ดังนั้นผู้วิจัยจึงเห็นว่าควรปรับปรุงข้อคำถามของข้อ 8 ในแบบสอบฉบับอ้างอิงให้ง่ายลงด้วยการถามเกี่ยวกับความสัมพันธ์ระหว่างพลังงานศักย์โน้มถ่วงกับตัวแปรหนึ่ง (การกระจัด S) เช่นเดียวกับข้อ 8 ในฉบับใหม่ แต่เปลี่ยนสถานการณ์จากปล่อยวัตถุลงเป็นโยนวัตถุขึ้น นอกจากนี้ผู้วิจัยทำการปรับปรุงข้อสอบข้อ 11 แล้ว ผู้วิจัยได้ทำการปรับปรุงตัวลวงในตัวลวง 3 ในข้อ 11 และตัวลวง 1 ในข้อ 14 ของแบบสอบฉบับใหม่ รวมไปถึง ตัวลวง 4 ในข้อ 2 และตัวลวง 4 ในข้อ 8 ของแบบสอบฉบับอ้างอิง (รายละเอียดอยู่ในภาคผนวก ง)

11.2 การคัดเลือกข้อสอบแบบสร้างคำตอบ และแนวทางในการปรับปรุงข้อสอบ

จากการพิจาณาผลการวิเคราะห์คุณภาพของข้อสอบแบบสร้างคำตอบในแบบสอบฉบับใหม่ (new) ตามเกณฑ์ที่กำหนดไว้ พบว่า ข้อสอบที่ต้องได้รับการปรับปรุง ได้แก่ ข้อ 4 เนื่องจากข้อสอบข้อนี้ยากมาก แต่จำแนกได้พอใช้ ($p = .13$, $r = .07$) และ ข้อ 5 เนื่องจากข้อสอบข้อนี้ยากมาก และจำแนกได้ต่ำ ($p = .07$, $r = .14$) และเมื่อพิจาณาข้อสอบแบบหลายตัวเลือกในแบบสอบฉบับอ้างอิง (reference) พบว่า ข้อสอบที่ต้องได้รับการปรับปรุงได้แก่ ข้อ 4 เนื่องจากข้อสอบข้อนี้ยากมาก แต่จำแนกได้พอใช้ ($p = .18$, $r = .31$) และข้อ 5 เนื่องจากข้อสอบข้อนี้ยาก และจำแนกได้ต่ำ ($p = .08$, $r = .16$)

ผลการวิเคราะห์พบว่า ทั้งข้อ 4 และข้อ 5 ต้องได้รับการปรับปรุงทั้งสองฉบับ เพราะข้อสอบทั้งสองข้อยากมาก จึงอาจส่งผลให้ค่าอำนาจจำแนกต่ำ ผู้วิจัยจึงปรับปรุงข้อสอบดังกล่าวด้วยการลดความยากของข้อสอบ ซึ่งผู้วิจัยกลับไปใช้ข้อสอบที่เคยสร้างไว้ให้ผู้เชี่ยวชาญทั้ง 5 คนได้ตรวจสอบความตรงเชิงเนื้อหาก่อนทำการปรับปรุงเป็นแบบสอบฉบับทดลอง ซึ่งข้อสอบข้อ 4 และข้อ 5 ที่ผู้วิจัยได้ออกไว้นั้น ผู้เชี่ยวชาญจำนวน 3 ท่าน ได้ให้ค่า IOC เป็น +1 ขณะที่ผู้เชี่ยวชาญจำนวน 2 ท่านเห็นว่าควรเพิ่มความซับซ้อนของโจทย์ แต่เมื่อนำไปในการทดสอบจริง พบว่า ข้อสอบทั้งสองข้อยากมาก ผู้วิจัยจึงนำโจทย์ทั้งสองที่เคยออกไว้ดังกล่าวมาใช้ทดแทนข้อสอบทั้ง 2 ข้อที่ยากเกินไป ซึ่งวัดในจุดประสงค์เชิงพฤติกรรมและระดับพฤติกรรมเดิม โดยข้อสอบข้อ 4 และ ข้อ 5 นี้มีค่า IOC เป็น 0.6

12. คัดเลือกข้อสอบเพื่อจัดทำแบบสอบคู่ขนาน โดยพิจารณาความคลอบคลุมเนื้อหา และค่าสถิติต่าง ๆ คือ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และความแปรปรวน นอกจากนี้ในการคัดเลือกข้อสอบคู่ขนานกันพิจาณาจากค่าความยากง่าย และอำนาจจำแนกเท่ากัน หรือใกล้เคียงกัน ในกรณีที่วัตถุประสงค์เชิงพฤติกรรมข้อใดมีจำนวนข้อสอบมากกว่า 1 ข้อ พิจารณาให้วัตถุประสงค์นั้นมีข้อสอบร่วม แล้วทำการทบทวนและปรับปรุงข้อสอบที่ได้รับการคัดเลือกอีกครั้ง

ในสถานการณ์จริง ผู้วิจัยได้ทำการสร้างแบบสอบคู่ขนานจำนวน 2 ฉบับที่ใช้ในการทดลองมีจำนวนพอดีกับความต้องการ เนื่องจากเวลาในการทำการทดสอบกับนักเรียนมีจำกัดเพียง 100 นาที ทำให้ไม่มีการคัดเลือกข้อสอบเพื่อจัดทำแบบสอบคู่ขนาน แต่มีการปรับปรุงข้อสอบที่ไม่ผ่านเกณฑ์จากการวิเคราะห์คุณภาพข้อสอบมาปรับปรุงตามที่ได้อธิบายไว้ในข้อ 11

ในส่วนของการคัดเลือกข้อสอบร่วมนั้น ผู้วิจัยคัดเลือกข้อสอบร่วมให้กระจายครอบคลุมจุดประสงค์เชิงพฤติกรรมที่สำคัญ แบ่งออกเป็น 2 ส่วน คือ ข้อสอบร่วมแบบหลายตัวเลือก และข้อสอบร่วมแบบสร้างคำตอบ ซึ่งผู้วิจัยมีแนวทางในการคัดเลือกดังนี้

12.1 การคัดเลือกข้อสอบร่วมแบบหลายตัวเลือกจากแบบสอบทั้งสองฉบับ

ผู้วิจัยทำการข้อสอบร่วมแบบหลายตัวเลือกจำนวน 8 ข้อ โดยพิจารณาจากจำนวนข้อสอบของจุดประสงค์เชิงพฤติกรรมในตารางวิเคราะห์หลักสูตรที่ผู้วิจัยได้สร้างไว้ กล่าวคือ จุดประสงค์เชิงพฤติกรรมข้อใดที่มีจำนวนข้อสอบแบบหลายตัวเลือกมากกว่า 1 ข้อ ผู้วิจัยเลือก 1 ข้อมาเป็นข้อร่วม ซึ่งมีอยู่ 7 จุดประสงค์ ได้แก่จุดประสงค์เชิงพฤติกรรมข้อที่ 1, 2, 3, 4, 5, 6 และ 8 โดยผู้วิจัยเลือกข้อสอบร่วมในจุดประสงค์เชิงพฤติกรรมดังกล่าวที่มีค่าคุณภาพของข้อสอบที่ดี อย่างเช่น ข้อสอบที่มีค่าอำนาจแจกสูง จากแบบสอบทั้ง 2 ฉบับ ยกเว้นในจุดประสงค์ข้อที่ 1 ที่ผู้วิจัยได้เลือกข้อสอบร่วมเป็นข้อสอบข้อ 1 จากแบบสอบฉบับใหม่ (ฉบับที่1) ซึ่งได้อธิบายเหตุไว้ในข้อ 11.1 และผู้วิจัยได้ทำการเลือกข้อสอบร่วมอีก 1 จากจุดประสงค์ข้อที่ 9 เพราะข้อสอบข้อนี้มีค่าความยากและค่าอำนาจจำแนกที่ดี รายละดังตาราง 12 และตาราง 13

ตาราง 12 รายละเอียดของข้อสอบร่วมแบบหลายตัวเลือกในแต่ละจุดประสงค์เชิงพฤติกรรมจำนวน 8 ข้อ

จุดประสงค์	ข้อสอบร่วม	คุณภาพของข้อสอบ	
		p	r
ข้อ 1	ข้อ 1 จากแบบสอบฉบับอ้างอิง (ฉบับที่ 2)	.36	.55
ข้อ 2	ข้อ 3 จากแบบสอบฉบับใหม่ (ฉบับที่ 1)	.70	.34
ข้อ 3	ข้อ 6 จากแบบสอบฉบับใหม่ (ฉบับที่ 1)	.49	.56
ข้อ 4	ข้อ 7 จากแบบสอบฉบับอ้างอิง (ฉบับที่ 2)	.56	.43
ข้อ 5	ข้อ 10 จากแบบสอบฉบับอ้างอิง (ฉบับที่ 2)	.34	.65
ข้อ 6	ข้อ 12 จากแบบสอบฉบับใหม่ (ฉบับที่ 1)	.44	.49
ข้อ 8	ข้อ 14 จากแบบสอบฉบับอ้างอิง (ฉบับที่ 2)	.38	.52
ข้อ 9	ข้อ 16 จากแบบสอบฉบับใหม่ (ฉบับที่ 1)	.56	.43

ตาราง 13 รายละเอียดของข้อสอบเฉพาะของข้อสอบแบบหลายตัวเลือกในแต่ละจุดประสงค์เชิงพฤติกรรมจำนวน 8 ข้อ

จุดประสงค์	ข้อสอบเฉพาะ	จุดประสงค์	ข้อสอบเฉพาะ
ข้อ 1	ข้อ 2	ข้อ 5	ข้อ 9
ข้อ 2	ข้อ 4	ข้อ 6	ข้อ 11
ข้อ 3	ข้อ 5	ข้อ 7	ข้อ 13
ข้อ 4	ข้อ 8	ข้อ 8	ข้อ 15

12.2 การคัดเลือกข้อสอบร่วมแบบสร้างคำตอบจากแบบสอบทั้งสองฉบับ

ผู้วิจัยทำการข้อสอบร่วมแบบสร้างคำตอบจำนวน 4 ข้อ โดยพิจารณาจากจำนวนข้อสอบของจุดประสงค์เชิงพฤติกรรมในตารางวิเคราะห์หลักสูตรที่ผู้วิจัยได้สร้างไว้ กล่าวคือ จุดประสงค์เชิงพฤติกรรมข้อใดที่มีจำนวนข้อสอบแบบสร้างคำตอบมากกว่า 1 ข้อ ผู้วิจัยเลือก 1 ข้อ มาเป็นข้อร่วม ซึ่งมีอยู่ 2 จุดประสงค์ ได้แก่จุดประสงค์เชิงพฤติกรรมข้อที่ 2 และ 8 โดยผู้วิจัยเลือกข้อสอบร่วมในจุดประสงค์เชิงพฤติกรรมดังกล่าวที่มีค่าคุณภาพของข้อสอบที่ดี อย่างเช่น ข้อสอบที่มีค่าอำนาจแจกสูงจากแบบสอบทั้ง 2 ฉบับ และผู้วิจัยเลือกข้อสอบร่วมอีก 2 ข้อ จากจุดประสงค์ข้อที่ 3 และข้อ 9 ส่วนข้อสอบที่เหลือในแต่ละจุดประสงค์ของแบบสอบแต่ละฉบับเป็นข้อสอบเฉพาะในแบบสอบฉบับนั้นๆ รายละเอียดของข้อสอบร่วมและข้อสอบเฉพาะ ดังตาราง 14 และตาราง 15

ตาราง 14 รายละเอียดของข้อสอบร่วมแบบสร้างคำตอบในแต่ละจุดประสงค์เชิงพฤติกรรมจำนวน 8 ข้อ

จุดประสงค์	ข้อสอบร่วม	คุณภาพของข้อสอบ	
		p	r
ข้อ 2	ข้อ 1 จากแบบสอบฉบับอ้างอิง (ฉบับที่ 2)	.45	.88
ข้อ 3	ข้อ 3 จากแบบสอบฉบับอ้างอิง (ฉบับที่ 2)	.33	.65
ข้อ 8	ข้อ 6 จากแบบสอบฉบับใหม่ (ฉบับที่ 1)	.31	.63
ข้อ 9	ข้อ 8 จากแบบสอบฉบับอ้างอิง (ฉบับที่ 2)	.25	.50

ตาราง 15 รายละเอียดของข้อสอบเฉพาะของข้อสอบแบบหลายตัวเลือกในแต่ละจุดประสงค์เชิงพฤติกรรมจำนวน 8 ข้อ

จุดประสงค์	ข้อสอบเฉพาะ	จุดประสงค์	ข้อสอบเฉพาะ
ข้อ 2	ข้อ 2	ข้อ 6	ข้อ 5
ข้อ 5	ข้อ 4	ข้อ 8	ข้อ 7

13. นำข้อสอบที่คัดเลือก และปรับปรุงจากข้อ 11 มาจัดเป็นแบบสอบเรียงตามลำดับ ดังตาราง 16 ได้ดังนี้

13.1 แบบสอบฉบับใหม่ (new) ประกอบด้วย ข้อสอบเฉพาะจำนวน 12 ข้อ ซึ่งเป็นข้อสอบแบบหลายตัวเลือก 8 ข้อ ตั้งแต่ข้อ 1-8 และข้อสอบแบบสร้างคำตอบ 4 ข้อ ตั้งแต่ข้อ 17-20 และข้อสอบร่วมจำนวน 12 ข้อ ซึ่งเป็นข้อสอบแบบหลายตัวเลือก 8 ข้อ ตั้งแต่ข้อ 9-16 และข้อสอบแบบสร้างคำตอบ 4 ข้อ ตั้งแต่ข้อ 21 -24 (แบบสอบฉบับใหม่ อยู่ในภาคผนวก จ)

13.2 แบบสอบฉบับอ้างอิง (reference) ประกอบด้วย ข้อสอบร่วมจำนวน 12 ข้อ ซึ่งเป็นข้อสอบแบบหลายตัวเลือก 8 ข้อ ตั้งแต่ข้อ 1-8 และข้อสอบแบบสร้างคำตอบ 4 ข้อ ตั้งแต่ข้อ 17-20 และข้อสอบเฉพาะจำนวน 12 ข้อ ซึ่งเป็นข้อสอบแบบหลายตัวเลือก 8 ข้อ ตั้งแต่ข้อ 9-16 และข้อสอบแบบสร้างคำตอบ 4 ข้อ ตั้งแต่ข้อ 21 -24 (แบบสอบฉบับอ้างอิง อยู่ในภาคผนวก จ)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 16 รายละเอียดของแบบสอบย่อยแต่ละฉบับ

ประเภทข้อสอบ	แบบสอบฉบับใหม่ 24 ข้อ			แบบสอบฉบับอ้างอิง 24 ข้อ		
	ข้อสอบเฉพาะ	ข้อสอบร่วม	รวม	ข้อสอบร่วม	ข้อสอบเฉพาะ	รวม
แบบ MC	8 ข้อ (ข้อ 1-8)	8 ข้อ (ข้อ 9-16)	16 ข้อ	8 ข้อ (ข้อ 1-8)	8 ข้อ (ข้อ 9-16)	16 ข้อ
แบบ CR	4 ข้อ (ข้อ 17-20)	4 ข้อ (ข้อ 21-24)	8 ข้อ	4 ข้อ (ข้อ 17-20)	4 ข้อ (ข้อ 21-24)	8 ข้อ
รวม	**12 ข้อ**	**12 ข้อ**	**24 ข้อ**	**12 ข้อ**	**12 ข้อ**	**24 ข้อ**

วิธีดำเนินการสร้างเกณฑ์การให้คะแนน (rubric score) สำหรับข้อสอบแบบสร้างคำตอบ

ข้อสอบแบบสร้างคำตอบในแบบสอบแต่ละฉบับ มีคะแนนเต็มข้อละ 4 คะแนน โดยเกณฑ์การให้คะแนนที่ใช้ในการตรวจข้อสอบแบบสร้างคำตอบแต่ละข้อในแบบสอบวิชาฟิสิกส์ ระดับมัธยมศึกษาปีที่ 4 มีวิธีดำเนินการสร้าง ดังนี้

1. ตรวจสอบวัตถุประสงค์เชิงพฤติกรมว่ามีความสอดคล้องกับมาตรฐานการเรียนรู้ในหลักสูตรแกนกลาง พ.ศ. 2551

2. กำหนดงานให้นักเรียนปฏิบัติให้สอดคล้องกับวัตถุประสงค์เชิงพฤติกรรม ซึ่งในที่นี้งานที่นักเรียนต้องปฏิบัติ คือ การแสดงวิธีทำในข้อสอบแบบสร้างคำตอบจำนวน 8 ข้อจากแบบสอบวิชาฟิสิกส์ ระดับมัธยมศึกษาปีที่ 4

3. กำหนดคุณลักษณะที่ต้องการประเมินให้สอดคล้องกับวัตถุประสงค์เชิงพฤติกรรมของข้อสอบแบบสร้างคำตอบแต่ละข้อ โดยข้อสอบทั้ง 8 ข้อในแบบสอบวิชาฟิสิกส์ ระดับมัธยมศึกษาปีที่ 4 มีคุณลักษณะร่วมกัน คือ การแสดงวิธีทำ และระบุคำตอบจากโจทย์ที่กำหนดให้ได้อย่างถูกต้อง

4. เลือกประเภทของรูบริคที่ต้องการใช้ ในที่นี้ผู้วิจัยเลือกใช้รูบริคแบบภาพรวม (holistic rubric) ในการประเมินวิธีทำ และคำตอบของนักเรียนจากข้อสอบแบบสร้างคำตอบทั้ง 8 ข้อ

5. เขียนเฉลยวิธีทำ และคำตอบของข้อสอบแบบสร้างคำตอบแต่ละข้ออย่างละเอียด เพื่อใช้เป็นแนวทางในการกำหนดระดับคะแนน และคำอธิบายในแต่ละระดับคะแนน

6. กำหนดจำนวนระดับของคะแนนของข้อสอบแบบสร้างคำตอบ โดยระดับคะแนนแบ่งออกเป็น 5 ระดับ คือ 0 1 2 3 และ 4

7. ตรวจสอบความชัดเจนและความถูกต้องของเกณฑ์การให้คะแนนของข้อสอบแบบสร้างคำตอบในแต่ละข้อ

8. ตรวจสอบคุณภาพด้านความตรงของรูบริคโดยให้ผู้เชี่ยวชาญช่วยตรวจสอบคุณภาพเพื่อคำนวณหาค่าดัชนี IOC โดยผู้วิจัยต้องการค่า IOC เท่ากับหรือมากกว่า 0.6

8.1 ผลการตรวจสอบสอดคล้องของเกณฑ์การให้คะแนนของข้อสอบแบบสร้างคำตอบสำหรับแบบสอบฉบับใหม่ (new) กับข้อสอบและเฉลยของข้อสอบ พบว่า เกณฑ์การให้คะแนนของข้อสอบแบบสร้างคำตอบจำนวน 3 ข้อ วัดได้สอดคล้องกับข้อสอบและเฉลยของข้อสอบ ได้แก่ ข้อ 3 5 และ 8 โดยมีค่าดัชนีความสอดคล้องระหว่างข้อสอบแบบสร้างคำตอบกับข้อสอบและ

เฉลยของข้อสอบอยู่ระหว่าง .67 - 1.00 ขณะที่เกณฑ์การให้คะแนนของข้อสอบแบบสร้างคำตอบจำนวน 5 ข้อ วัดได้ไม่สอดคล้องกับข้อสอบและเฉลยของข้อสอบ ได้แก่ ข้อ 1 2 4 6 และ 7 โดยมีค่าดัชนีความสอดคล้องระหว่างข้อสอบแบบสร้างคำตอบกับข้อสอบและเฉลยของข้อสอบคือ 0 (รายละเอียดอยู่ในภาคผนวก ฉ) ซึ่งผู้เชี่ยวชาญเห็นว่าองค์ประกอบของเกณฑ์การให้คะแนนถูกต้องตามหลักของการแสดงวิธีทำในวิชาฟิสิกส์ แต่เกณฑ์การให้คะแนนในแต่ละองค์ประกอบขาดความชัดเจน ซึ่งต้องมีการปรับการเขียนคำอธิบายในแต่ละระดับคะแนนให้เป็นตามหลักของเกณฑ์การให้คะแนนแบบ rubric score ซึ่งผู้วิจัยได้ปรับปรุงแก้ไขแล้วนำไปให้ผู้เชี่ยวชาญพิจารณาอีกครั้งหนึ่ง

ผลการตรวจสอบสอดคล้องของเกณฑ์การให้คะแนนของข้อสอบแบบสร้างคำตอบสำหรับแบบสอบฉบับอ้างอิง (reference) กับข้อสอบและเฉลยของข้อสอบ พบว่า เกณฑ์การให้คะแนนของข้อสอบแบบสร้างคำตอบจำนวน 1 ข้อ วัดได้สอดคล้องกับข้อสอบและเฉลยของข้อสอบได้แก่ ข้อ 3 โดยมีค่าดัชนีความสอดคล้องระหว่างข้อสอบแบบสร้างคำตอบกับข้อสอบและเฉลยของคือ 1.00 ขณะที่เกณฑ์การให้คะแนนของข้อสอบแบบสร้างคำตอบจำนวน 3 ข้อ วัดได้ไม่สอดคล้องกับข้อสอบและเฉลยของข้อสอบได้แก่ 1 2 และ 4 โดยมีค่าดัชนีความสอดคล้องระหว่างข้อสอบแบบสร้างคำตอบกับข้อสอบและเฉลยของข้อสอบ คือ 0 (รายละเอียดอยู่ในภาคผนวก ฉ) ซึ่งผู้เชี่ยวชาญเห็นว่าองค์ประกอบของเกณฑ์การให้คะแนนถูกต้องตามหลักของการแสดงวิธีทำในวิชาฟิสิกส์ แต่เกณฑ์การให้คะแนนในแต่ละองค์ประกอบขาดความชัดเจน โดยต้องมีการปรับการเขียนคำอธิบายในแต่ละระดับคะแนนให้เป็นตามหลักของเกณฑ์การให้คะแนนแบบ rubric score เช่นเดียวกับแบบสอบฉบับใหม่ ซึ่งผู้วิจัยได้ปรับปรุงแก้ไขแล้วนำไปให้ผู้เชี่ยวชาญพิจารณาอีกครั้งหนึ่ง

8.2 การตรวจสอบความตรงตามเนื้อหาของเกณฑ์การให้คะแนนของข้อสอบแบบสร้างคำตอบสำหรับแบบสอบวิชาฟิสิกส์เรื่องงานและพลังงาน ในครั้งที่ 2 โดยผู้ทรงคุณวุฒิด้านการวัดและประเมินผลการศึกษา และผู้ทรงคุณวุฒิทางด้านการสอนวิชาฟิสิกส์ จำนวน 3 คน ซึ่งในครั้งนี้ผู้วิจัยได้ให้ผู้เชี่ยวชาญตรวจสอบความตรงตามเนื้อหาดังกล่าวเฉพาะแบบสอบฉบับใหม่เท่านั้น เนื่องจากแบบสอบทั้งสองฉบับมีเกณฑ์การให้คะแนนเดียวกัน และผู้เชี่ยวชาญได้แนะนำให้ผู้วิจัยปรับเกณฑ์การให้คะแนนของข้อสอบแบบสร้างคำตอบในแบบสอบฉบับอ้างอิง (reference) ตามแบบสอบฉบับใหม่ (new) และผลการตรวจสอบได้ดังนี้

ผลการตรวจสอบสอดคล้องของเกณฑ์การให้คะแนนของข้อสอบแบบสร้างคำตอบสำหรับแบบสอบวิชาฟิสิกส์เรื่องงานและพลังงานกับข้อสอบและเฉลยของข้อสอบ พบว่า เกณฑ์การให้คะแนนของข้อสอบแบบสร้างคำตอบทุกข้อวัดได้สอดคล้องกับข้อสอบและเฉลยของข้อสอบ โดยข้อสอบแบบสร้างคำตอบทุกข้อมีค่าดัชนีความสอดคล้องระหว่างข้อสอบแบบสร้างคำตอบกับข้อสอบและเฉลยของข้อสอบเป็น 1.00 โดยผู้เชี่ยวชาญได้เสนอแนะให้มีการปรับคำอธิบายของระดับคะแนน 1 เพียงเล็กน้อย (รายละเอียดอยู่ในภาคผนวก ฉ)

9. ทำการปรับปรุงรูบริคของข้อสอบแบบสร้างคำตอบสำหรับแบบสอบรายวิชาฟิสิกส์ทั้ง 2 ฉบับ ตามคำแนะนำของผู้เชี่ยวชาญ

10. ให้ผู้ตรวจ 1 และ B ทดลองให้คะแนนข้อสอบแบบสร้างคำตอบจากแบบสอบฉบับใหม่ (new) และฉบับอ้างอิงคนละ 23 ฉบับ

11. ทำการหาสัมประสิทธิ์สหพันธ์แบบเพียร์สัน (r) ระหว่างคะแนนจากผู้ตรวจ 1 และผู้ตรวจ 2 เพื่อความเที่ยงระหว่างผู้ประเมิน (inter-rater reliability) ซึ่งเป็นการตรวจสอบว่าเกณฑ์การให้คะแนนมีความเที่ยงหรือไม่ และผู้ตรวจทั้งสองคนมีความเข้าใจเกณฑ์การให้คะแนนตรงกันหรือไม่

11.1 การตรวจสอบความเที่ยงระหว่างผู้ประเมิน สำหรับแบบสอบฉบับใหม่

ในการตรวจสอบความเที่ยงระหว่างผู้ประเมิน (inter-rater reliability) โดยใช้ค่าสัมประสิทธิ์สหพันธ์แบบเพียร์สัน (r) พบว่าคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบในแบบสอบเรื่องงานและพลังงาน ของผู้ตรวจ 1 และผู้ตรวจ 2 มีความสัมพันธ์กันทางบวกในระดับสูง (r = .96) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (p = .00) ดังนั้นสรุปได้ว่า เกณฑ์การให้คะแนนสำหรับข้อสอบแบบสร้างคำตอบในแบบสอบเรื่องงานและพลังงานฉบับใหม่ (new) วิชาฟิสิกส์ ม.4 มีความเที่ยง รายละเอียดดังตาราง 17

ตาราง 17 ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันระหว่างคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบของผู้ตรวจ 1 และผู้ตรวจ 2 สำหรับแบบสอบฉบับใหม่

สัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน		
	ผู้ตรวจ 1	ผู้ตรวจ 2
ผู้ตรวจ 1	1	
ผู้ตรวจ 2	.96**	1
ค่าเฉลี่ย	1.55	1.57
ส่วนเบี่ยงเบนมาตรฐาน	1.70	1.68

**p< .01

11.2 การตรวจสอบความเที่ยงระหว่างผู้ประเมิน สำหรับแบบสอบฉบับอ้างอิง

ในการตรวจสอบความเที่ยงระหว่างผู้ประเมิน (inter-rater reliability) โดยใช้ค่าสัมประสิทธิ์สหพันธ์แบบเพียร์สัน (r) พบว่าคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบในแบบสอบเรื่องงานและพลังงานของผู้ตรวจ 1 และผู้ตรวจ 2 มีความสัมพันธ์กันทางบวกในระดับสูง (r = .97) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (p = .00) ดังนั้นสรุปได้ว่าเกณฑ์การให้คะแนนสำหรับข้อสอบแบบสร้างคำตอบในแบบสอบเรื่องงานและพลังงานฉบับที่ 2 วิชาฟิสิกส์ ม.4 มีความเที่ยง รายละเอียดดังตาราง 18

ตาราง 18 ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันระหว่างคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบของผู้ตรวจคนที่ 1 และผู้ตรวจคนที่ 2 สำหรับแบบสอบฉบับอ้างอิง

สัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน		
	ผู้ตรวจ 1	ผู้ตรวจ 2
ผู้ตรวจ 1	1	
ผู้ตรวจ 2	.97**	1
ค่าเฉลี่ย	1.64	1.70
ส่วนเบี่ยงเบนมาตรฐาน	1.72	1.70

**p< .01

12. ทำการปรับปรุงเกณฑ์การให้คะแนนถ้ามีค่าสัมประสิทธิ์สหพันธ์แบบเพียร์สันต่ำกว่า 0.7 แต่เนื่องจากเกณฑ์การให้คะแนนมีค่าสัมประสิทธิ์สหพันธ์แบบเพียร์สันมากกว่า .07 ผู้วิจัยจึงไม่ได้มีการปรับปรุงเกณฑ์การให้คะแนนเพิ่มเติม

13. จัดทำเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบฉบับสมบูรณ์ สำหรับแบบสอบเรื่องงานและพลังงานฉบับใหม่และฉบับใหม่ รายวิชาฟิสิกส์ ระดับชั้นมัธยมศึกษาปีที่ 4 แล้วส่งเกณฑ์ดังกล่าวไปยังผู้ตรวจทั้งสองคนเพื่อใช้ในการตรวจให้คะแนนในแบบสอบฉบับจริง รายละเอียดอยู่ในภาคผนวก ช

วิธีคัดเลือกผู้ตรวจให้คะแนน และการเตรียมความพร้อมให้กับผู้ตรวจให้คะแนน

1. ผู้วิจัยคัดเลือกผู้ตรวจให้คะแนน โดยเลือกจากครูที่สอนวิชาฟิสิกส์ในระดับมัธยมศึกษาปีที่ 4 ที่มีประสบการณ์ใกล้เคียงในโรงเรียนมาเป็นเวลา 3-5 ปี

2. ผู้วิจัยทำการติดต่อไปยังผู้ตรวจคะแนนที่ได้รับการคัดเลือก เพื่อสอบถามความสะดวกของผู้ตรวจให้คะแนน โดยผู้วิจัยชี้แจงขั้นตอนต่าง ๆ เกี่ยวกับตรวจให้คะแนนให้กับผู้ตรวจทั้งสองคน

แต่เนื่องในสถานการณ์จริงครูที่สอนวิชาฟิสิกส์ในโรงเรียนต่างมีภาระหน้าที่ประจำอยู่ในโรงเรียนค่อนข้างเยอะ ทำให้ครูหลายท่านที่ผู้วิจัยได้ติดต่อไม่สะดวกเป็นผู้ตรวจให้คะแนน ผู้วิจัยจึงสามารถสรรหาผู้ตรวจให้คะแนนที่มีคุณสมบัติตรงตามที่กำหนดไว้ และมีความสะดวกเป็นผู้ตรวจให้คะแนนได้จนเสร็จสิ้นได้เพียง 1 ท่าน โดยทำหน้าที่เป็นผู้ตรวจ 1 ซึ่งก็คือ

อาจารย์โชคชัย พวงสมบัติ ปัจจุบันเป็นครูผู้สอนวิชาฟิสิกส์โรงเรียนวัดราชาธิวาส แต่ได้เกษียณอายุราชการแล้ว โดยบรรจุเป็นข้าราชการครูตั้งแต่ปี 2518 นอกจากนี้ปัจจุบันยังมีตำแหน่งเป็นผู้ช่วยรองผู้อำนวยการกลุ่มบริหารงานทั่วไปของโรงเรียนอีกด้วย

ด้วยเหตุนี้ ผู้วิจัยจึงตัดสินใจให้ตนเองเป็นผู้ตรวจให้คะแนนอีกคนหนึ่ง แม้ว่าผู้วิจัยไม่ได้มีประสบการณ์สอนในโรงเรียนมาเป็นเวลา 3-5 ปี แต่ผู้วิจัยเป็นผู้สร้างแบบสอบเรื่องงานและพลังงานทั้ง 2 ฉบับ และเกณฑ์การให้คะแนนของข้อสอบแบบสร้างคำตอบในแบบสอบดังกล่าวด้วยตนเอง ซึ่งได้ผ่านการตรวจสอบคุณภาพด้านความตรงตามเนื้อหา และปรับปรุงตามคำแนะนำเรียบร้อยแล้ว ผู้วิจัยจึงมีความเข้าใจในเกณฑ์การให้คะแนน นอกจากนี้ผู้วิจัยยังมีกระบวนการตรวจสอบความเที่ยงระหว่างผู้ประเมิน (inter-rater reliability) รวมไปถึงการตรวจสอบความสอดคล้องระหว่างการให้

คะแนนของผู้ตรวจให้คะแนนทั้งสองคนเป็นรายข้อ เพื่อให้มั่นใจว่าทั้งผู้วิจัย และผู้ตรวจให้คะแนนอีกท่านหนึ่งมีความเข้าใจร่วมกันเกี่ยวกับเกณฑ์การให้คะแนนดังกล่าว ซึ่งรายละเอียดของการตรวจสอบดังกล่าวอยู่ในข้อ 6 ในที่นี้ผู้วิจัยทำหน้าที่เป็นผู้ตรวจ 2

3. หลังจากที่ครูท่านดังกล่าวได้ตกลงรับหน้าที่เป็นผู้ตรวจให้คะแนนแบบสอบวิชาฟิสิกส์ ระดับมัธยมศึกษาปีที่ 4 ผู้วิจัยได้ทำการส่งเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบทั้ง 12 ข้อไปยังผู้ตรวจท่านดังกล่าวก่อนทำการส่งแบบสอบฉบับทดลองไปให้ผู้ตรวจทำการตรวจให้คะแนนเป็นเวลา 1 สัปดาห์ เพื่อให้ผู้ตรวจทำความเข้าใจกับเกณฑ์การให้คะแนน

4. ผู้วิจัยส่งข้อสอบแบบสร้างคำตอบของแบบสอบฉบับใหม่ (new) และฉบับอ้างอิง (reference) ไปให้ผู้ตรวจ 2 จำนวน 23 ฉบับ เพื่อตรวจให้คะแนน ในขณะเดียวกันผู้วิจัยก็ทำการตรวจให้คะแนนข้อสอบดังกล่าวเช่นกัน

5. หลังจากนั้นผู้วิจัยนำคะแนนจากข้อสอบแบบสร้างคำตอบในแต่ละข้อของผู้ตรวจ 1 และ 2 มาทำการวิเคราะห์ความสอดคล้องภายในระหว่างผู้ตรวจให้คะแนน โดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (r) แล้วรายงานผลการวิเคราะห์ไปยังผู้ตรวจ 1 อีกครั้งหนึ่ง โดยผู้วิจัยต้องการค่าความสอดคล้องภายในระหว่างผู้ตรวจให้คะแนนในแต่ละข้อเท่ากับหรือมากกว่า .7 ในกรณีที่ได้ค่า r ต่ำกว่า .7 ผู้วิจัยต้องทำตรวจสอบการให้คะแนนข้อสอบแบบสร้างคำตอบในข้อนั้น ๆ เพื่อหาความแตกต่างระหว่างการประเมินของผู้ตรวจทั้งสองคน หลังจากนั้นดำเนินการสร้างความเข้าใจเกี่ยวกับการใช้เกณฑ์การให้คะแนนอีกครั้ง ผลจากการวิเคราะห์ความสอดคล้องภายในระหว่างผู้ตรวจให้คะแนน ประกอบด้วย

5.1 ความสอดคล้องในการให้คะแนนระหว่างผู้ประเมินเป็นรายข้อสำหรับแบบสอบฉบับใหม่

จากการตรวจสอบความสอดคล้องในการให้คะแนนระหว่างผู้ประเมินเป็นรายข้อ โดยใช้ค่าสัมประสิทธิ์สหพันธ์แบบเพียร์สัน (r) พบว่า คะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบทุกข้อในแบบสอบเรื่องงานและพลังงานฉบับใหม่ (new) ของผู้ตรวจ 1 และผู้ตรวจ 2 มีความสัมพันธ์กันทางบวกในระดับสูงอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($p = .00$) โดยมีค่าสัมประสิทธิ์สหพันธ์แบบเพียร์สัน (r) อยู่ระหว่าง .86 - .99 รายละเอียดดังตาราง 19

5.2 ความสอดคล้องในการให้คะแนนระหว่างผู้ประเมินเป็นรายข้อสำหรับแบบสอบฉบับอ้างอิง

จากการตรวจสอบความสอดคล้องในการให้คะแนนระหว่างผู้ประเมินเป็นรายข้อ โดยใช้ค่าสัมประสิทธิ์สหพันธ์แบบเพียร์สัน (r) พบว่า คะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบทุกข้อในแบบสอบเรื่องงานและพลังงานฉบับอ้างอิง (reference) ของผู้ตรวจ 1 และผู้ตรวจ 2 มีความสัมพันธ์กันทางบวกในระดับสูง อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($p = .00$) โดยมีค่าสัมประสิทธิ์สหพันธ์แบบเพียร์สัน (r) อยู่ระหว่าง .84 -1.00 รายละเอียดดังตาราง 20

ตาราง 19 ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันระหว่างคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบของผู้ตรวจ 1 และผู้ตรวจ 2 (แบบรายข้อ) สำหรับแบบสอบฉบับใหม่

ข้อ	ผู้ตรวจ 1		ผู้ตรวจ 2		ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน
	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน	
1	1.48	1.30	1.59	1.52	.86**
2	1.78	1.65	1.78	1.70	.98**
3	1.17	1.70	1.13	1.66	.96**
4	.96	1.58	1.22	1.68	.91**
5	3.17	1.44	3.13	1.49	.95**
6	.70	1.42	.83	1.37	.95**
7	2.70	1.36	2.65	1.37	.99**
8	.48	.95	.52	.95	.98**

****p<. 01**

ตาราง 20 ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันระหว่างคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบของผู้ตรวจ 1 และผู้ตรวจ 2 (แบบรายข้อ) สำหรับแบบสอบฉบับอ้างอิง

ข้อ	ผู้ตรวจ 1		ผู้ตรวจ 2		ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน
	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน	
1	3.39	.98	3.35	1.03	.98**
2	1.26	1.71	1.17	1.75	.99**
3	2.70	1.58	2.70	1.58	1.00*
4	.130	.344	.174	.39	.84**
5	1.48	1.62	1.61	1.53	.98**
6	1.70	1.72	1.70	1.72	1.00**
7	1.40	1.80	1.57	1.70	.84**
8	3.39	.99	3.35	1.03	.96**

****p< .01**

นอกจากนี้ผู้วิจัยยังได้ทำการทดสอบความแตกต่างค่าเฉลี่ยของคะแนนที่ได้จากการตรวจให้คะแนนของผู้ตรวจทั้ง 2 คน สำหรับแบบสอบฉบับใหม่ (new) และฉบับอ้างอิง (reference) พบว่าการตรวจให้คะแนนของนักเรียนในการแสดงวิธีทำจากผู้ตรวจทั้ง 2 คน สำหรับแบบสอบฉบับใหม่และฉบับอ้างอิงไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (p = .86, p =.80 ตามลำดับ) รายละเอียดดังตาราง 21 และ 22

ตาราง 21 ผลการทดสอบความแตกต่างของค่าเฉลี่ยคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบของผู้ตรวจคน 1 และผู้ตรวจ 2 สำหรับแบบสอบฉบับใหม่

ผู้ตรวจ	**จำนวน**	**ค่าเฉลี่ย**	**ส่วนเบี่ยงเบนมาตรฐาน**	**F**	**p**
ผู้ตรวจคนที่ 1	23	12.43	5.59	.23	.86
ผู้ตรวจคนที่ 2	23	12.74	5.91		

ตาราง 22 ผลการทดสอบความแตกต่างของค่าเฉลี่ยคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบของผู้ตรวจ 1 และผู้ตรวจ 2 สำหรับแบบสอบฉบับอ้างอิง

ผู้ตรวจ	จำนวน	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน	F	p
ผู้ตรวจคนที่ 1	23	13.13	5.68	.03	.80
ผู้ตรวจคนที่ 2	23	12.74	5.91		

วิธีการจัดการสอบและการตรวจให้คะแนน

1. การทดสอบช่วงที่ 1 มีขั้นตอนในการจัดการสอบและการตรวจให้คะแนนดังนี้

1.1 ผู้วิจัยดำเนินการจัดการสอบกับ 1) กลุ่มอ้างอิงที่ใช้เป็นเกณฑ์ โดยให้ผู้สอบกลุ่มดังกล่าวทำแบบสอบฉบับใหม่ (new) 2) กลุ่มใหม่ โดยให้ผู้สอบกลุ่มดังกล่าวทำแบบสอบฉบับใหม่ (new)

1.2 ผู้ตรวจ 1 และ B (ผู้วิจัย) ทำการตรวจแบบสอบของกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์และกลุ่มใหม่ ตามวิธีการตรวจที่ได้กำหนดไว้ ดังภาพ 3

2. การทดสอบช่วงที่ 2 ซึ่งมีขั้นตอนในการจัดการสอบและการตรวจให้คะแนนดังนี้

2.1 ผู้วิจัยดำเนินการจัดการสอบกับ 1) กลุ่มอ้างอิงที่ใช้เป็นเกณฑ์ โดยให้ผู้สอบกลุ่มดังกล่าวทำแบบสอบฉบับอ้างอิง ซึ่งทำการทดสอบหลังจากการทดสอบครั้งแรกประมาณ 3 สัปดาห์ 2) กลุ่มอ้างอิง โดยให้ผู้สอบกลุ่มดังกล่าวทำแบบสอบฉบับอ้างอิง (reference)

2.2 ผู้ตรวจ 1 และ B (ผู้วิจัย) ทำการตรวจแบบสอบของกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์และกลุ่มอ้างอิง ตามวิธีการตรวจที่ได้กำหนดไว้ ดังภาพ 3 โดยรายชื่อโรงเรียนในกลุ่มใหม่ กลุ่มอ้างอิงและกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์ เป็นดังตาราง 23

ตาราง 23 รายชื่อโรงเรียนในกลุ่มใหม่ (new) และกลุ่มอ้างอิง (reference)

กลุ่ม	โรงเรียน	จำนวน	รวม
1) กลุ่มใหม่ (new)	**โรงเรียนขนาดใหญ่พิเศษ**		
	1.1 โรงเรียนรัตนโกสินทร์สมโภช บางขุนเทียน 2 ห้อง	97	
	1.2 โรงเรียนมัธยมวัดหนองแขม 4 ห้อง	179	
	1.3 โรงเรียนวัดนวลนรดิศ 3 ห้อง	126	
	1.4 โรงเรียนจันทร์ประดิษฐารามวิทยาคม 3 ห้อง	91	
	โรงเรียนขนาดใหญ่		
	1.5 โรงเรียนราชวินิต มัธยม 2 ห้อง	46	
	1.6 โรงเรียนอิสลามวิทยาลัยแห่งประเทศไทย 2 ห้อง	87	1,084
	1.7 โรงเรียนสันติราษฎร์วิทยาลัย 2 ห้อง	50	
	1.8 โรงเรียนมัธยมวัดมกุฎกษัตริย์ 3 ห้อง	99	
	1.9 โรงเรียนมัธยมวัดดุสิตาราม 2 ห้อง	56	
	1.10 โรงเรียนราชวินิตบางแคปานขำ 2 ห้อง	53	
	โรงเรียนขนาดกลาง/เล็ก		
	1.11 โรงเรียนวัดราชาธิวาส 2 ห้อง	26	
	1.12 โรงเรียนมัธยมวัดนายโรง 2 ห้อง	63	

ตาราง 23 รายชื่อโรงเรียนในกลุ่มใหม่ (new) และกลุ่มอ้างอิง (reference) (ต่อ)

กลุ่ม	โรงเรียน	จำนวน	รวม
1) กลุ่มใหม่ (new) (ต่อ)	**โรงเรียนขนาดกลาง/เล็ก (ต่อ)**		
	1.13 โรงเรียนทวีธาภิเศก 2 2 ห้อง	58	1,084
	1.14 โรงเรียนวัดประดู่ในทรงธรรม 1 ห้อง	36	
	1.15 โรงเรียนไชยฉิมพลี 1 ห้อง	17	
2) กลุ่มอ้างอิง (reference)	**โรงเรียนขนาดใหญ่พิเศษ**		
	2.1 โรงเรียนศึกษานารี 1 ห้อง	43	
	โรงเรียนขนาดใหญ่		
	2.2 โรงเรียนโพธิสารพิทยากร 2 ห้อง	87	400
	2.3 โรงเรียนสตรีวัดระฆัง 2 ห้อง	83	
	โรงเรียนขนาดกลาง/เล็ก		
	2.4 โรงเรียนฤทธิณรงค์รอน 1 ห้อง	36	
	2.5 โรงเรียนวัดรางบัว 1 ห้อง	27	
3) กลุ่มอ้างอิง (reference) ที่ใช้เป็นเกณฑ์	**โรงเรียนขนาดใหญ่**		
	3.1 โรงเรียนสวนกุหลาบวิทยาลัย ธนบุรี 3 ห้อง	125	125
	รวม		**1,609**

การเก็บรวบรวมข้อมูล

1. ขอข้อมูลรายชื่อโรงเรียนและจำนวนนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 จากสำนักงานเขตพื้นที่มัธยมศึกษากรุงเทพมหานคร เขต 1

2. ขอหนังสือขอความร่วมมือในการเก็บข้อมูลวิจัยจากงานหลักสูตรและการจัดการเรียนการสอน ฝ่ายวิชาการ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

3.ดำเนินการเก็บรวบรวมข้อมูลจากโรงเรียนที่เป็นกลุ่มตัวอย่าง (กลุ่ม try out) หลังจากนั้นนำข้อสอบที่ได้จากกลุ่มดังกล่าวมาทำการตรวจให้คะแนน และวิเคราะห์ข้อสอบเพื่อปรับปรุงคุณภาพของข้อสอบ รวมไปถึงใช้ในการตรวจสอบเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบ และความสอดคล้องระหว่างผู้ตรวจให้คะแนนทั้งสองคน

4. นำแบบสอบฉบับสมบูรณ์ที่ผ่านการแก้ไขแล้วไปใช้สอบกับกลุ่มตัวอย่างที่ได้กำหนดไว้ตามรูปแบบการออกแบบการปรับเทียบคะแนน โดยมีช่วงเวลาในการเก็บข้อมูลกับกลุ่มอ้างอิงใช้เป็นเกณฑ์ กลุ่มใหม่ และกลุ่มอ้างอิง ดังตาราง 24

5. นำแบบสอบที่เก็บรวบรวมได้มาตรวจให้คะแนนโดยผู้ตรวจ 1 และ B ตามตามวิธีการตรวจทั้ง 3 วิธีที่ได้กำหนดไว้ หลังจากนั้นนำข้อมูลที่ได้มาวิเคราะห์

ตาราง 24 ช่วงเวลาในการเก็บข้อมูลกับกลุ่มตัวอย่างทั้ง 3 กลุ่ม

กลุ่ม	โรงเรียน	ธ.ค.		ม.ค				ก.พ	
		3	4	1	2	3	4	1	2
เกณฑ์	โรงเรียนสวนกุหลาบวิทยาลัย ธนบุรี	(1)			(2)				
ใหม่	โรงเรียนวัดนวลนรดิศ, โรงเรียนมัธยมวัดนายโรง								
	โรงเรียนวัดราชาธิวาส, โรงเรียนวัดประดู่ในทรงธรรม								
	โรงเรียนจันทร์ประดิษฐารามวิทยาคม , โรงเรียนมัธยมวัดมกุฎกษัตริย์, โรงเรียนทวีธาภิเศก								
	โรงเรียนอิสลามวิทยาลัยแห่งประเทศไทย โรงเรียนมัธยมวัดมกุฎกษัตริย์ โรงเรียนราชวินิตบางแคปานขำ								
	โรงเรียนรัตนโกสินทร์สมโภช บางขุนเทียน โรงเรียนราชวินิต มัธยม, โรงเรียนไชยฉิมพลี								
	โรงเรียนมัธยมวัดหนองแขม, โรงเรียนมัธยมวัดดุสิตาราม								
อ้างอิง	โรงเรียนศึกษานารี, โรงเรียนสตรีวัดระฆัง, โรงเรียนโพธิสารพิทยากร								
	โรงเรียนฤทธิณรงค์รอน, โรงเรียนวัดรางบัว								

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในการวิจัยครั้งนี้ แบ่งออกเป็น 7 ตอนได้แก่

1. การวิเคราะห์ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

2. การวิเคราะห์ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์

3. การวิเคราะห์ความสอดคล้องในการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบระหว่างผู้ตรวจ 1 และ 2

4. การปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) โดยใช้วิธีการปรับเทียบเชิงเส้นตรง

5. การปรับเทียบคะแนนสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ โดยใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed)

6. การเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี ระหว่างวิธีของเลอวิน (Levine method) และวิธี CLE (Chained linear equating)

7. การเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของวิธีการตรวจ 3 วิธี สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

ตอนที่ 1 การวิเคราะห์ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

วิเคราะห์ค่าสถิติพื้นฐานของคะแนนของผู้สอบกลุ่มอ้างอิงและกลุ่มใหม่ แบ่งตามวิธีการตรวจ 3 วิธี ได้แก่ จำนวนผู้สอบแต่ละกลุ่ม จำนวนข้อสอบแบบหลายตัวเลือกและแบบสร้างคำตอบในแบบสอบของผู้สอบแต่ละกลุ่ม ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนจากแบบสอบทั้งฉบับของผู้สอบทั้งสองกลุ่ม ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนจากข้อสอบร่วมของผู้สอบทั้งสองกลุ่ม จำนวนข้อสอบร่วมทั้งหมดในแบบสอบของผู้สอบแต่ละกลุ่ม ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนจากข้อสอบร่วมของผู้สอบทั้งสองกลุ่ม และค่าสหสัมพันธ์ระหว่างคะแนนรวมของข้อสอบทั้งฉบับกับคะแนนของข้อสอบร่วม เป็นต้น โดยใช้โปรแกรม SPSS for Window Version 17.0

ตอนที่ 2 การวิเคราะห์ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์

วิเคราะห์ค่าสถิติพื้นฐานของคะแนนของผู้สอบกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์ ได้แก่ จำนวนผู้สอบแต่ละกลุ่ม จำนวนข้อสอบแบบหลายตัวเลือกและแบบสร้างคำตอบในแบบสอบของผู้สอบแต่ละกลุ่ม ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนจากแบบสอบทั้งฉบับของผู้สอบทั้งสองกลุ่ม ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนจากข้อสอบร่วมของผู้สอบทั้งสองกลุ่ม จำนวนข้อสอบร่วมทั้งหมดในแบบสอบของผู้สอบแต่ละกลุ่ม ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนจากข้อสอบร่วมของผู้สอบทั้งสองกลุ่ม และค่าสหสัมพันธ์ระหว่างคะแนนรวมของข้อสอบทั้งฉบับกับคะแนนของข้อสอบร่วม เป็นต้น โดยใช้โปรแกรม SPSS for Window Version 17.0

ตอนที่ 3 การวิเคราะห์ความสอดคล้องในการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบระหว่างผู้ตรวจ 1 และ 2

วิเคราะห์ค่าสัมประสิทธิ์สหพันธ์ของเพียร์สัน (r) ระหว่างการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบของผู้ตรวจ 1 ในวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม และการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบของผู้ตรวจ 2 ในวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจคนหนึ่ง ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม) โดยใช้โปรแกรม SPSS for Window Version 17.0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตอนที่ 4 การปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) โดยใช้วิธีการปรับเทียบเชิงเส้นตรง

1) การปรับเทียบคะแนนโดยใช้วิธีของเลอวิน (Levine method) ซึ่งเป็นการแปลงคะแนนของแบบสอบฉบับใหม่ (new) ไปยังคะแนนของแบบสอบฉบับอ้างอิง (reference: Y) ในแต่ละคะแนน x ในแต่ละวิธีการตรวจ โดยใช้โปรแกรม R (Albano, 2014) ซึ่ง Albano ได้อ้างอิงสมการที่ใช้ในการปรับเทียบสำหรับวิธีดังกล่าว ตามสมการที่อ้างไว้ใน Kolen และ Brennan (2004) ดังสมการ

$$l_{Ys}(x) = \left[\frac{\gamma_2}{\gamma_1}\right][x - \mu_1(X)] + \mu_2(Y) + \gamma_2[\mu_1(V) - \mu_2(V)]$$

2) การปรับเทียบคะแนนโดยใช้วิธี CLE (Chained Linear Equating) ซึ่งเป็นการแปลงคะแนนจากคะแนนของแบบสอบฉบับใหม่ (new: X) ไปยังคะแนนของข้อสอบร่วม (A) และแปลงคะแนนจากคะแนนของแบบสอบฉบับอ้างอิง (reference: Y) ไปยังคะแนนของข้อสอบร่วม (A) ในแต่ละวิธีการตรวจ โดยใช้โปรแกรม R (Albano, 2014) ซึ่ง Albano ได้อ้างอิงสมการที่ใช้ในการปรับเทียบสำหรับวิธีดังกล่าว ตามสมการที่อ้างไว้ใน Livingston, Dorans และ Wright (1990) ซึ่งก็เป็นสมการเดียวกับที่อ้างไว้ใน Davier (2011) ดังสมการ

$$y = f(x) = \mu_2(Y) + \frac{\sigma_2(Y)\sigma_1(A)}{\sigma_2(A)\sigma_1(X)}[x - \mu_1(X)] + \frac{\sigma_2(Y)}{\sigma_2(A)}[\mu_1(A) - \mu_2(A)]$$

ตอนที่ 5 การปรับเทียบคะแนนสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ โดยใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed)

ในการปรับเทียบคะแนนสำหรับรูปแบบที่ใช้เป็นเกณฑ์ (criterion) จากคะแนนจากแบบสอบฉบับใหม่โดยผู้ตรวจ 2 ไปยังคะแนนจากแบบสอบฉบับอ้างอิงโดยผู้ตรวจ 1 นั้น โดยใช้โปรแกรม R (Albano, 2014) มีขั้นตอนดังนี้

1. ใช้วิธีล็อกเชิงเส้น (loglinear) ในการแปลงคะแนนแบบสอบทั้งสองฉบับที่มีการแจกแจงเป็นโค้งไม่ราบเรียบให้เป็นคะแนนที่มีการแจกแจงเป็นโค้งราบเรียบ ซึ่งต้องกำหนดค่า degree สำหรับการแจกแจงของตัวแปรเดียว และ xdegree สำหรับการแจกแจงสองตัวแปร โดย Albano แนะนำว่า อาจกำหนดค่า xdegree = 1 เพื่อให้ง่ายในการแปลความหมายและอธิบายคะแนนที่ได้จากโมเดล แต่ให้ปรับค่า degree ให้มีค่าเหมาะสมกับข้อมูล โดยค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานที่ได้จากการปรับโค้งให้มีความราบเรียบควรมีค่าเท่ากับข้อมูลเดิมที่ยังไม่ได้ทำการปรับ

2. หลังจากนั้นโปรแกรมทำการสร้างโมเดล (nested model) ที่หลากหลาย ตั้งแต่โมเดลอย่างง่ายไปจนถึงโมเดลที่ซับซ้อน พร้อมทั้งตารางวิเคราะห์ค่า Deviance

3. ผู้วิจัยทำการเลือกโมเดลที่มีความเหมาะสมกับข้อมูล โดย Albano แนะนำว่าโมเดลที่มีความเหมาะสมกับข้อมูลต้องมีค่า Akaike's Information Criterion (AIC) และค่า Bayesian Information Criterion (BIC) ต่ำๆ

4. หลังจากนั้นจึงนำค่าต่อเนื่องที่เหมาะสมจากวิธีล็อกเชิงเส้น (loglinear) ในข้อ 3 ไปใช้ในการปรับเทียบคะแนนของกลุ่มตัวอย่างดังกล่าว โดยใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed) ซึ่งผู้วิจัยใช้โปรแกรม R (Albano, 2014) ในการปรับเทียบคะแนนในรูปแบบที่ใช้เป็นเกณฑ์นี้ โดยฟังก์ชั่นการปรับเทียบที่ใช้ในการแปลงคะแนนจากแบบสอบ X ไปยังสเกลของแบบสอบ Y ก็คือ $e_{Y(chain)}$ โดยที่ $e_{Y(chain)} = e_{Y2}[e_{V1}(x)]$

ตอนที่ 6 การเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี ระหว่างวิธีของเลอวิน (Levine method) และวิธี CLE (Chained linear equating)

การเปรียบเทียบวิธีการปรับเทียบเชิงเส้นตรง (equating method) 2 วิธี สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) คือ วิธีของเลอวิน (Levine method) และวิธี CLE (Chained Linear Equating) โดยพิจารณาจากค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (Root Mean Squared Difference: RMSD) ซึ่งคำนวณจากโปรแกรม Microsoft excel ตามสมการด้านล่าง

$$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$$

เมื่อ i คือ จุดของคะแนนที่สังเกตได้ (raw score point)

$\hat{e}_i(x_i)$ คือ คะแนนที่ถูกปรับเทียบโดยใช้วิธีการปรับเทียบแบบต่าง ๆ ที่จุดของคะแนนที่สังเกตได้ x

$e_i(x_i)$ คือ ฟังก์ชั่นของการปรับเทียบที่เป็นเกณฑ์ (criterion equating function) ที่จุดของคะแนนที่สังเกตได้ x ในที่นี้ คือ คะแนนที่ถูกปรับเทียบภายใต้การออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์

w_i คือ สัดส่วนสัมพัทธ์ของผู้สอบที่ทำแบบสอบฉบับใหม่ที่แต่ละจุดของคะแนน

ตอนที่ 7 การเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของวิธีการตรวจ 3 วิธีสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

การเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของวิธีการตรวจ 3 วิธีสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ได้แก่ 1) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม 2) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจคนหนึ่ง ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม) 3) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม (ผู้ตรวจทั้ง 2 คน ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม โดยแบ่งกันตรวจคนละครึ่งหนึ่งของข้อสอบร่วม) โดยวิเคราะห์ผ่านโปรแกรม R (Albano, 2014) ซึ่งพิจารณาจากค่า

1. ค่าเฉลี่ยน้ำหนักของรากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (weighted average RMSE) ซึ่งก็คือ $\sqrt{\sum_i w_i RMSE_i^2}$ หรือเรียกสั้น ๆ ว่าค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE)

โดยผู้วิจัยต้องคำนวณค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง ในแต่ละจุดคะแนน ($RMSE_i$) ก่อนด้วยโปรแกรม R (Albano, 2014)

เมื่อ $RMSE_i = \sqrt{{Bias_i}^2 + {SEE_i}^2}$

ถ้า $Bias_i$ คือ ค่าความลำเอียงในแต่ละจุดคะแนน

SEE_i คือ ค่าความคลาดเคลื่อนในการปรับเทียบในแต่ละจุดคะแนน

หลังจากนั้นนำค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสองที่ได้ในแต่ละจุดคะแนนแทนในสมการ $\sqrt{\sum_i w_i RMSE_i^2}$ เพื่อหาค่าเฉลี่ยน้ำหนักของรากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (weighted average RMSE) โดยใช้โปรแกรม Microsoft excel

2. ค่าเฉลี่ยน้ำหนักของค่าความลำเอียง (weighted average root mean square bias) ซึ่งก็คือ $\sqrt{\sum_i w_i Bias_i^2}$ หรือเรียกสั้น ๆ ว่าค่าความลำเอียง (bias)

โดยผู้วิจัยต้องคำนวณค่าความลำเอียง (bias) ในแต่ละจุดคะแนน ($Bias_i$) ก่อนด้วยโปรแกรม R (Albano, 2014) ซึ่งค่า $Bias_i$ เป็นความคลาดเคลื่อนเชิงระบบ (systematic error)

$$Bias_i = \hat{\bar{e}}_Y(x) - e_Y(x)$$

เมื่อ $$\hat{\bar{e}}_Y(x) = \frac{1}{R}\sum_{r=1}^{R} \hat{e}_{Yr}(x)$$

ถ้า $e_Y(x)$ คือ ฟังก์ชั่นการปรับเทียบของประชากร

$\hat{e}_{Yr}(x)$ คือ ค่าประมาณของฟังก์ชั่นการปรับเทียบของกลุ่มตัวอย่าง

r = 1, 2,..., R

หลังจากนั้นนำค่าความลำเอียง (bias) ที่ได้ในแต่ละจุดคะแนนแทนในสมการ $\sqrt{\sum_i w_i Bias_i^2}$ เพื่อหาค่าเฉลี่ยน้ำหนักของค่าความลำเอียง (weighted average root mean square bias) โดยใช้โปรแกรม Microsoft excel

3. ค่าเฉลี่ยน้ำหนักของความเคลื่อนในการปรับเทียบ (weighted average equating error) ซึ่งก็คือ $\sqrt{\sum_i w_i SEE_i^2}$ หรือเรียกสั้น ๆ ว่าค่าความคลาดเคลื่อนในการปรับเทียบ (equating error)

โดยผู้วิจัยต้องคำนวณค่าความคลาดเคลื่อนในการปรับเทียบในแต่ละจุดคะแนน (SEE_i) ก่อนด้วยโปรแกรม R (Albano, 2014) ซึ่งค่า SEE_i เป็นความคลาดเคลื่อนแบบสุ่ม (random error)

$$SEE_i = \sqrt{\frac{1}{R}\sum_{r=1}^{R}[\hat{e}_Y(x) - e_Y(x)]^2}$$

หลังจากนั้นนำค่า SEE ในแต่ละจุดคะแนนแทนในสมการ $\sqrt{\sum_i w_i SEE_i^2}$ เพื่อหาค่าเฉลี่ยน้ำหนักของความเคลื่อนในการปรับเทียบ (weighted average equating error) โดยใช้โปรแกรม Microsoft excel

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 4
ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีจุดประสงค์ เพื่อเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจที่แตกต่างกัน และเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ระหว่างวิธีของเลอวิน (Levine method) และวิธี CLE (Chained Linear Equating) สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ของนักเรียนระดับชั้นมัธยมศึกษาชั้นปีที่ 4

แบบสอบที่ใช้เป็นแบบสอบรายวิชาฟิสิกส์ เรื่องงานและพลังงาน ระดับชั้นมัธยมศึกษาปีที่ 4 จำนวน 2 ฉบับที่มีความเป็นคู่ขนานกัน มีข้อสอบร่วมจำนวน 12 ข้อ ซึ่งประกอบ ข้อสอบแบบหลายตัวเลือก 8 ข้อ และข้อสอบแบบสร้างคำตอบ 4 ข้อ โดยนำแบบสอบดังกล่าวไปใช้กลุ่มผู้สอบ 3 กลุ่ม ได้แก่ กลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) กลุ่มอ้างอิงที่ทำแบบสอบฉบับอ้างอิง (reference) ขณะที่กลุ่มอ้างอิงที่ใช้เป็นเกณฑ์ ทำทั้งแบบสอบฉบับใหม่ (new) และฉบับอ้างอิง (reference)

ผู้วิจัยนำผลจากการสอบของผู้สอบกลุ่มใหม่ และกลุ่มอ้างอิงทั้ง 2 กลุ่ม มาวิเคราะห์และปรับเทียบโดยใช้วิธีการปรับเทียบเชิงเส้นตรง 2 วิธี ได้แก่ วิธีของเลอวิน (Levine method) และวิธี CLE (Chained Linear Equating) ขณะที่นำผลจากการสอบของผู้สอบกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์มาวิเคราะห์ด้วยวิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed) จากโปรแกรม R (Albano, 2014) เพื่อปรับคะแนนของผู้สอบทั้ง 3 กลุ่มให้อยู่ในมาตรวัดเดียวกัน หลังจากนั้นทำการเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจที่แตกต่างกัน และเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ระหว่างวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี โดยเสนอผลการวิเคราะห์ดังนี้

ตอนที่ 1 ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

ตอนที่ 2 ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์

ตอนที่ 3 ผลการวิเคราะห์ความสอดคล้องในการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบระหว่างผู้ตรวจ 1 และ 2

ตอนที่ 4 ผลการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) โดยใช้วิธีการปรับเทียบเชิงเส้นตรง

ตอนที่ 5 ผลการปรับเทียบคะแนนสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ โดยใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed)

ตอนที่ 6 ผลการเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี ระหว่างวิธีของเลอวิน (Levine method) และวิธี CLE (Chained Linear Equating)

ตอนที่ 7 ผลการเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของวิธีการตรวจ 3 วิธี สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

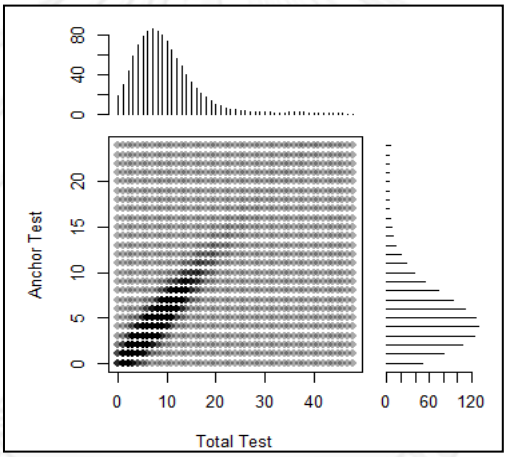
ตอนที่ 1 ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

ค่าสถิติพื้นฐานของคะแนนของผู้สอบกลุ่มอ้างอิงและกลุ่มใหม่ แบ่งตามวิธีการตรวจ 3 วิธี ได้แก่ จำนวนผู้สอบแต่ละกลุ่ม จำนวนข้อสอบแบบหลายตัวเลือกและแบบสร้างคำตอบในแบบสอบของผู้สอบแต่ละกลุ่ม ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนจากแบบสอบทั้งฉบับของผู้สอบทั้งสองกลุ่ม ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนจากข้อสอบร่วมของผู้สอบทั้งสองกลุ่ม จำนวนข้อสอบร่วมทั้งหมดในแบบสอบของผู้สอบแต่ละกลุ่ม ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนจากข้อสอบร่วมของผู้สอบทั้งสองกลุ่ม และค่าสหสัมพันธ์ระหว่างคะแนนรวมของข้อสอบทั้งฉบับกับคะแนนของข้อสอบร่วม เป็นต้น โดยใช้โปรแกรม SPSS for Window Version 17.0

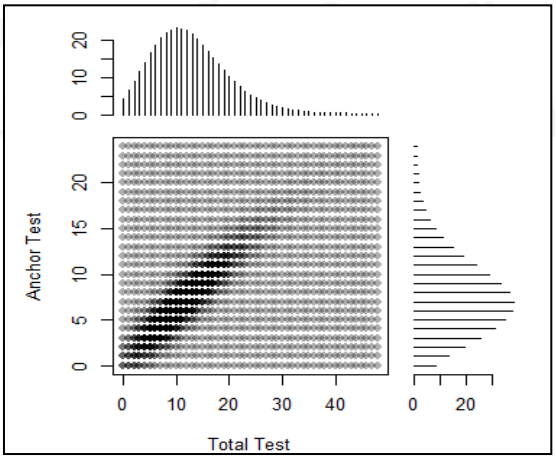
กำหนดให้วิธีการตรวจวิธีที่ 1 คือ วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม วิธีการตรวจวิธีที่ 2 คือ วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจ 2 ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม) วิธีการตรวจวิธีที่ 3 คือ วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม (ผู้ตรวจทั้ง 2 คน ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม โดยแบ่งกันตรวจคนละครึ่งหนึ่งของข้อสอบร่วม) เพื่อความสะดวกในการอธิบายผู้วิจัยกำหนดให้วิธีการตรวจวิธีที่ 1, 2 และ 3 เรียกสั้นๆ ว่า Non Trend, Full Trend และ Half Trend ตามลำดับ

วิธีการตรวจวิธีที่ 1: Non Trend เมื่อพิจารณาคะแนนของผู้สอบกลุ่มใหม่จากแบบสอบฉบับใหม่ (new) ทั้งฉบับ จำนวน 1,084 คน พบว่า คะแนนดังกล่าวมีค่าสูงสุดและต่ำสุด คือ 43 และ 1 ตามลำดับ มีค่าเฉลี่ย คือ 9.38 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐาน คือ 6.18 โดยการแจกแจงของคะแนนจากแบบสอบทั้งฉบับมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มใหม่ส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ย และมีความโด่งมากกว่าโค้งปกติเล็กน้อย คือ มีการกระจายต่ำ โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 1 คะแนน ถึง 20 คะแนน ดังภาพ 5 และเมื่อพิจารณาเฉพาะคะแนนจากข้อสอบร่วมจำนวน 12 ข้อ พบว่า คะแนนจากข้อสอบร่วมทั้งหมดมีค่าสูงสุดและต่ำสุด คือ 22 และ 1 ตามลำดับ มีค่าเฉลี่ย คือ 5.56 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐาน คือ 3.96 โดยการแจกแจงของคะแนนจากข้อสอบร่วมมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มใหม่ส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ย และมีความโด่งที่แบนมากกว่าโค้งปกติเล็กน้อย คือ มีการกระจายสูง โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 1 คะแนน ถึง 10 คะแนน ดังภาพ 5 นอกจากนี้คะแนนของแบบสอบทั้งฉบับและคะแนนของข้อสอบร่วมมีความสัมพันธ์กันในระดับสูง (r = .93) อย่างมีนัยสำคัญทางสถิติ

ที่ระดับ .01 และเมื่อพิจารณาคะแนนของผู้สอบกลุ่มอ้างอิงจากแบบสอบฉบับอ้างอิง (reference) ทั้งฉบับจำนวน 400 คน พบว่า คะแนนดังกล่าวมีค่าสูงสุดและต่ำสุด คือ 40 และ 2 ตามลำดับ มีค่าเฉลี่ย คือ 12.83 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐาน คือ 7.51 โดยการแจกแจงของคะแนนจากแบบสอบทั้งฉบับมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มอ้างอิงส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ย และมีความโด่งที่แบนมากกว่าโค้งปกติเล็กน้อย โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 1 คะแนน ถึง 25 คะแนน ดังภาพ 6 และเมื่อพิจารณาเฉพาะคะแนนจากข้อสอบร่วมจำนวน 12 ข้อ พบว่า คะแนนจากข้อสอบร่วมทั้งหมด มีค่าสูงสุดและต่ำสุด คือ 22 และ 1 ตามลำดับ มีค่าเฉลี่ย คือ 7.74 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐาน คือ 4.32 โดยการแจกแจงของคะแนนจากข้อสอบร่วมมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มใหม่ส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ย และมีความโด่งที่แบนมากกว่าโค้งปกติ โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 0 คะแนน ถึง 13 คะแนน ดังภาพ 6 นอกจากนี้คะแนนของแบบสอบทั้งฉบับและคะแนนของข้อสอบร่วมมีความสัมพันธ์กันในระดับสูง (r = .91) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 รายละเอียดดังตาราง 25



ภาพ 5 กราฟการกระจายของคะแนนแบบสอบฉบับใหม่ทั้งฉบับและข้อสอบร่วมทั้งหมด สำหรับวิธีการตรวจ 1



ภาพ 6 กราฟการกระจายของคะแนนแบบสอบฉบับอ้างอิงทั้งฉบับและข้อสอบร่วมทั้งหมด สำหรับวิธีการตรวจ 1

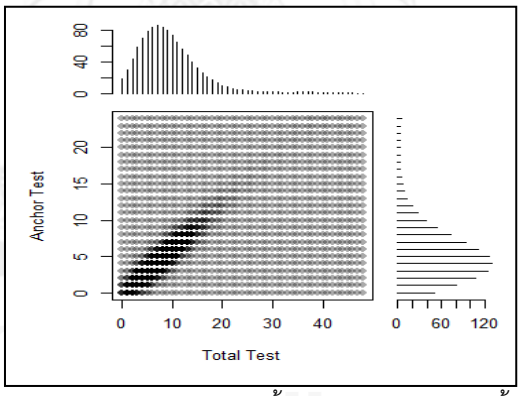
ตาราง 25 ค่าสถิติของกลุ่มผู้สอบที่ทำแบบสอบฉบับใหม่ (new) และฉบับอ้างอิง (reference) สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

	Non Trend		Full Trend		Half Trend	
	New	Ref	New	Ref	New	Ref
จำนวนผู้สอบ	1084	400	1084	400	1084	400
จำนวนข้อสอบ MC/CR	MC16/ CR 8	MC16/ CR 8	MC16/ CR 8	MC16/ CR 8	MC16/ CR 8	MC16/ CR 8
ผู้ตรวจข้อสอบ CR	B	A	B	A	B	A
คะแนนเต็มของแบบสอบทั้งฉบับ	48	48	48	48	48	48
คะแนนสูงสุดของแบบสอบทั้งฉบับ	43	40	43	39	47	40
คะแนนต่ำสุดของแบบสอบทั้งฉบับ	1	2	1	2	1	2
ค่าเฉลี่ยของคะแนนสอบทั้งฉบับ	9.38	12.83	9.38	12.98	9.38	12.96
S.D ของคะแนนสอบทั้งฉบับ	6.18	7.51	6.18	7.57	6.22	7.55
มัธยฐานของคะแนนสอบทั้งฉบับ	7.00	12.00	7.00	12.00	7.00	12.00
ฐานนิยมของคะแนนสอบทั้งฉบับ	6.00	13.00	6.00	10.00	6.00	10.00
ความเบ้ของคะแนนสอบทั้งฉบับ	1.60	.97	1.60	.93	1.64	.94
ความโด่งของคะแนนสอบทั้งฉบับ	3.21	.75	3.21	.62	3.58	.65
จำนวนข้อสอบร่วมทั้งหมด	MC 8/ CR 4	MC 8/ CR 4	MC 8/ CR 4	MC 8/ CR 4	MC 8/ CR 4	MC 8/ CR 4
ผู้ตรวจข้อสอบร่วม CR	A	B	B	B	A, B	A, B
คะแนนเต็มของข้อสอบร่วมทั้งหมด	24	24	24	24	24	24
คะแนนสูงสุดของแบบสอบทั้งฉบับ	22	22	22	22	24	22
คะแนนต่ำสุดของแบบสอบทั้งฉบับ	1	1	1	1	1	1
ค่าเฉลี่ยของคะแนนจากข้อสอบร่วมทั้งหมด	5.56	7.74	5.56	7.89	5.57	7.86
S.D ของคะแนนจากข้อสอบร่วมทั้งหมด	3.96	4.32	3.96	4.39	3.99	4.36
มัธยฐานของข้อสอบร่วมทั้งหมด	4.00	7.00	4.00	7.00	4.00	7.00
ฐานนิยมของข้อสอบร่วมทั้งหมด	3.00	3.00	3.00	3.00	3.00	3.00
ความเบ้ของข้อสอบร่วมทั้งหมด	1.46	.65	1.46	.60	1.49	.62
ความโด่งของข้อสอบร่วมทั้งหมด	2.22	-.17	2.22	-.26	2.40	-.21
ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนสอบทั้งฉบับและคะแนนของข้อสอบร่วม	.93**	.91**	.93**	.91**	.94**	.91**

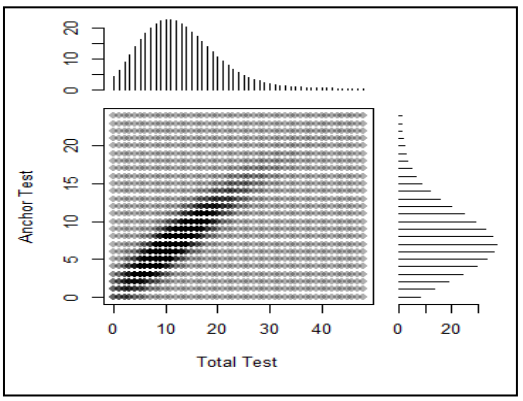
**p< .01

วิธีการตรวจวิธีที่ 2: Full Trend เมื่อพิจารณาคะแนนของผู้สอบกลุ่มใหม่จากแบบสอบฉบับใหม่ (new) ทั้งฉบับ จำนวน 1,084 คน พบว่าคะแนนดังกล่าวมีค่าสูงสุดและต่ำสุด คือ 43 และ 1 ตามลำดับ มีค่าเฉลี่ย คือ 9.38 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐาน คือ 6.18 โดยการแจกแจงของคะแนนจากแบบสอบทั้งฉบับมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มใหม่ส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ย และมีความโด่งมากกว่าโค้งปกติเล็กน้อย โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 1 คะแนน ถึง 20 คะแนน ดังภาพ 7 และเมื่อพิจารณาเฉพาะคะแนนจากข้อสอบร่วมจำนวน 12 ข้อ พบว่า คะแนนจากข้อสอบร่วมทั้งหมด มีค่าสูงสุดและต่ำสุด คือ 22 และ 1 ตามลำดับ มีค่าเฉลี่ย คือ

5.56 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐาน คือ 3.96 โดยการแจกแจงของคะแนนจากข้อสอบร่วมมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มใหม่ส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ย และมีความโด่งที่แบนมากกว่าโค้งปกติ โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 1 คะแนน ถึง 10 คะแนน ดังภาพ 7 นอกจากนี้คะแนนของแบบสอบทั้งฉบับและคะแนนของข้อสอบร่วมมีความสัมพันธ์กันในระดับสูง (r = .93) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และเมื่อพิจารณาคะแนนของผู้สอบกลุ่มอ้างอิงจากแบบสอบฉบับอ้างอิง (reference) ทั้งฉบับ จำนวน 400 คน พบว่าคะแนนดังกล่าวมีค่าสูงสุดและต่ำสุด คือ 39 และ 2 ตามลำดับมีค่าเฉลี่ย คือ 12.98 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐาน คือ 7.57 โดยการแจกแจงของคะแนนจากแบบสอบทั้งฉบับมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มอ้างอิงส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ย และมีความโด่งที่แบนมากกว่าโค้งปกติเล็กน้อย โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 1 คะแนน ถึง 20 คะแนน ดังภาพ 8 และเมื่อพิจารณาเฉพาะคะแนนจากข้อสอบร่วมจำนวน 12 ข้อ พบว่า คะแนนจากข้อสอบร่วมทั้งหมดมีค่าสูงสุดและต่ำสุด คือ 22 และ 1 ตามลำดับ มีค่าเฉลี่ย คือ 7.89 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐาน คือ 4.39 โดยการแจกแจงของคะแนนจากข้อสอบร่วมมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มอ้างอิงส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ย และมีความโด่งที่แบนมากกว่าโค้งปกติเล็กน้อย โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 1 คะแนน ถึง 13 คะแนน ดังภาพ 8 นอกจากนี้คะแนนของแบบสอบทั้งฉบับและคะแนนของข้อสอบร่วมมีความสัมพันธ์กันในระดับสูง (r = .91) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 รายละเอียดดังตาราง 25

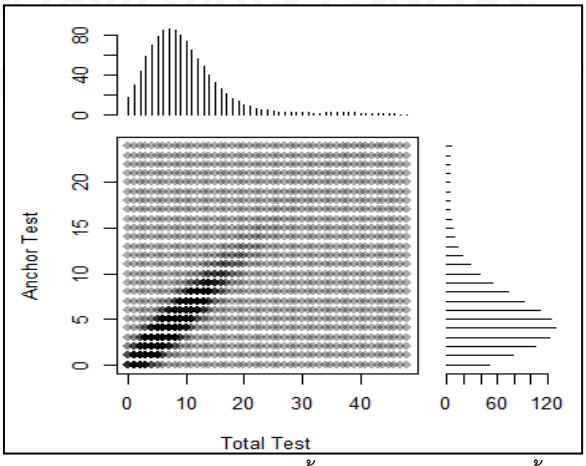


ภาพ 7 การกระจายของคะแนนจากแบบสอบฉบับใหม่ทั้งฉบับและข้อสอบร่วมทั้งหมด สำหรับวิธีการตรวจ 2

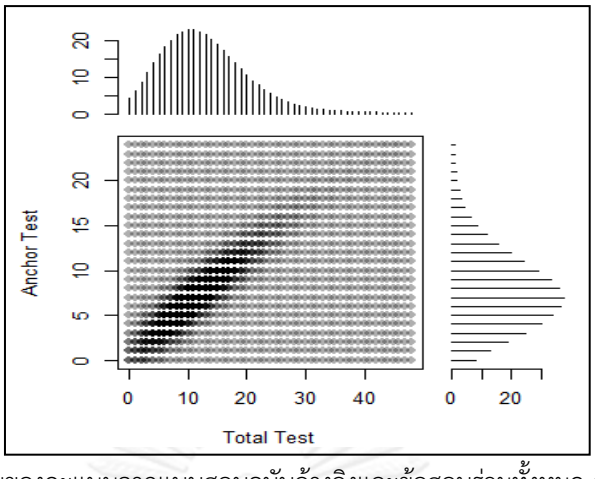


ภาพ 8 การกระจายของคะแนนจากแบบสอบฉบับอ้างอิงทั้งฉบับและข้อสอบร่วมทั้งหมด สำหรับวิธีการตรวจ 2

วิธีการตรวจวิธีที่ 3: Half Trend เมื่อพิจารณาคะแนนของผู้สอบกลุ่มใหม่จากแบบสอบฉบับใหม่ (new) ทั้งฉบับ จำนวน 1,084 คน พบว่าคะแนนดังกล่าวมีค่าสูงสุดและต่ำสุด คือ 47 และ 1 ตามลำดับมีค่าเฉลี่ย คือ 9.38 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐาน คือ 6.22 โดยการแจกแจงของคะแนนจากแบบสอบทั้งฉบับมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มใหม่ส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ย และมีความโด่งมากกว่าโค้งปกติ โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 1 คะแนน ถึง 20 คะแนน ดังภาพ 9 และเมื่อพิจารณาเฉพาะคะแนนจากข้อสอบร่วมจำนวน 12 ข้อ พบว่าคะแนนจากข้อสอบร่วมทั้งหมดมีค่าสูงสุดและต่ำสุด คือ 24 และ 1 ตามลำดับ มีค่าเฉลี่ย คือ 5.57 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐาน คือ 3.99 โดยการแจกแจงของคะแนนจากข้อสอบร่วมมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มใหม่ส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ย และมีความโด่งมากกว่าโค้งปกติ โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 1 คะแนน ถึง 10 คะแนน ดังภาพ 9 นอกจากนี้คะแนนของแบบสอบทั้งฉบับและคะแนนของข้อสอบร่วมมีความสัมพันธ์กันในระดับสูง ($r = .94$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และเมื่อพิจารณาคะแนนของผู้สอบกลุ่มอ้างอิงจากแบบสอบฉบับอ้างอิงทั้งฉบับ จำนวน 400 คน พบว่า คะแนนดังกล่าวมีค่าสูงสุดและต่ำสุด คือ 40 และ 2 ตามลำดับมีค่าเฉลี่ย คือ 12.96 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐาน คือ 7.55 โดยการแจกแจงของคะแนนจากแบบสอบทั้งฉบับมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มใหม่ส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ย และมีความโด่งมากกว่าโค้งปกติเล็กน้อย โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 1 คะแนน ถึง 20 คะแนน ดังภาพ 10 และเมื่อพิจารณาเฉพาะคะแนนจากข้อสอบร่วมจำนวน 12 ข้อ พบว่า คะแนนจากข้อสอบร่วมทั้งหมด มีค่าสูงสุดและต่ำสุด คือ 22 และ 1 ตามลำดับ มีค่าเฉลี่ย คือ 7.86 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐาน คือ 4.36 โดยการแจกแจงของคะแนนจากข้อสอบร่วมมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มอ้างอิงส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ยและมีความโด่งที่แบนมากกว่าโค้งปกติเล็กน้อย โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 1 คะแนน ถึง 13 คะแนน ดังภาพ 10 นอกจากนี้คะแนนของแบบสอบทั้งฉบับและคะแนนของข้อสอบร่วมมีความสัมพันธ์กันในระดับสูงอย่างมีนัยสำคัญทางสถิติ .01 ($r = .91$) รายละเอียดดังตาราง 25



ภาพ 9 การกระจายของคะแนนจากแบบสอบฉบับใหม่ทั้งฉบับและข้อสอบร่วมทั้งหมด สำหรับวิธีการตรวจ 3



ภาพ 10 การกระจายของคะแนนจากแบบสอบฉบับอ้างอิงและข้อสอบร่วมทั้งหมด สำหรับวิธีการตรวจ 3

เมื่อพิจารณาคะแนนแบบสอบทั้งฉบับของผู้สอบกลุ่มใหม่ในวิธีการตรวจทั้ง 3 วิธี พบว่า คะแนนดังกล่าวมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานที่ใกล้เคียงกันมากเช่นเดียวกับคะแนนแบบสอบทั้งฉบับของผู้สอบกลุ่มอ้างอิงในการออกแบบการปรับเทียบคะแนนทั้ง 3 แบบที่มีค่าเฉลี่ยและส่วนเบี่ยงเบนใกล้เคียงกันมาก และเมื่อพิจารณาคะแนนข้อสอบร่วมทั้งหมดของผู้สอบกลุ่มใหม่ในวิธีการตรวจทั้ง 3 วิธี พบว่า คะแนนดังกล่าวมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานที่ใกล้เคียงกันเช่นเดียวกับคะแนนข้อสอบร่วมทั้งหมดของผู้สอบกลุ่มอ้างอิงในวิธีการตรวจทั้ง 3 วิธีมีค่าเฉลี่ย และส่วนเบี่ยงเบนใกล้เคียงกัน โดยค่าสถิติของผู้สอบกลุ่มใหม่ในวิธีการตรวจวิธีที่ 1: Non Trend และวิธีการตรวจวิธีที่ 2: Full Trend มีค่าเท่ากัน

แต่อย่างไรก็ตาม เมื่อเปรียบเทียบคะแนนแบบสอบทั้งฉบับของผู้สอบกลุ่มใหม่และกลุ่มอ้างอิง พบว่า คะแนนแบบสอบทั้งฉบับของผู้สอบทั้งสองกลุ่มแตกต่างกัน โดยคะแนนแบบสอบทั้งฉบับของผู้สอบกลุ่มอ้างอิงมีค่าเฉลี่ยสูงกว่าผู้สอบกลุ่มใหม่ เช่นเดียวกับคะแนนข้อสอบร่วมทั้งหมดของผู้สอบทั้งสองกลุ่มที่แตกต่างกัน โดยคะแนนข้อสอบร่วมทั้งหมดของผู้สอบกลุ่มอ้างอิงสูงกว่าผู้สอบกลุ่มใหม่ แต่อย่างไรก็ตามค่าส่วนเบี่ยงเบนมาตรฐานของผู้สอบทั้งสองกลุ่มก็ไม่แตกต่างกันมากนัก ซึ่งผู้วิจัยได้อภิปรายข้อค้นพบนี้ในบทที่ 5 ต่อไป

ตอนที่ 2 ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์

ในการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ (SG design) เมื่อพิจารณาคะแนนของผู้สอบกลุ่มใหม่จากแบบสอบฉบับใหม่ (new) ทั้งฉบับ จำนวน 125 คน พบว่าคะแนนดังกล่าวมีค่าสูงสุดและต่ำสุด คือ 38 และ 3 ตามลำดับมีค่าเฉลี่ย คือ 13.08 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐาน คือ 7.27 โดยการแจกแจงของคะแนนจากแบบสอบทั้งฉบับมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มใหม่ส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ย และมีความโด่งมากกว่าโค้งปกติเล็กน้อย โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 3 คะแนน ถึง 25 คะแนน ดังภาพ 11 และเมื่อพิจารณาเฉพาะคะแนนจากข้อสอบร่วมจำนวน 12 ข้อ พบว่า คะแนนจากข้อสอบร่วมทั้งหมดมีค่าสูงสุดและต่ำสุด คือ 23 และ 1 ตามลำดับ มีค่าเฉลี่ย คือ 7.68 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐาน คือ 4.27 โดยการแจกแจง

ของคะแนนจากข้อสอบร่วมมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มใหม่ส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ย และมีความโด่งมากกว่าโค้งปกติเล็กน้อย โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 1 คะแนน ถึง 15 คะแนน ดังภาพ 11 นอกจากนี้คะแนนของแบบสอบทั้งฉบับและคะแนนของข้อสอบร่วมมีความสัมพันธ์กันในระดับสูง (r = .91) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และเมื่อพิจารณาคะแนนของผู้สอบกลุ่มอ้างอิงจากแบบสอบฉบับอ้างอิง (reference) ทั้งฉบับ จำนวน 125 คน พบว่าคะแนนดังกล่าวมีค่าสูงสุดและต่ำสุด คือ 36 และ 3 ตามลำดับมีค่าเฉลี่ย คือ 15.80 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐาน คือ 8.03 โดยการแจกแจงของคะแนนจากแบบสอบทั้งฉบับมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มใหม่ส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ย และมีความโด่งที่แบนมากกว่าโค้งปกติเล็กน้อย โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 3 คะแนน ถึง 25 คะแนน ดังภาพ 12 และเมื่อพิจารณาเฉพาะคะแนนจากข้อสอบร่วมจำนวน 12 ข้อ พบว่า คะแนนจากข้อสอบร่วมทั้งหมด มีค่าสูงสุดและต่ำสุด คือ 20 และ 2 ตามลำดับ มีค่าเฉลี่ย คือ 9.12 คะแนน และมีค่า ส่วนเบี่ยงเบนมาตรฐาน คือ 4.56 โดยการแจกแจงของคะแนนจากข้อสอบร่วมมีลักษณะเบ้ขวา แสดงว่า ผู้สอบกลุ่มอ้างอิงส่วนใหญ่ได้คะแนนต่ำกว่าค่าเฉลี่ย และมีความโด่งที่แบนมากกว่าโค้งปกติเล็กน้อย โดยผู้สอบส่วนใหญ่ได้คะแนนอยู่ในช่วงประมาณ 2 คะแนน ถึง 15 คะแนน ดังภาพ 12 นอกจากนี้คะแนนของแบบสอบฉบับทั้งฉบับและคะแนนของข้อสอบร่วมมีความสัมพันธ์กันในระดับสูง (r = .89) อย่างมีนัยสำคัญทางสถิติ .01 รายละเอียดดังตาราง 26

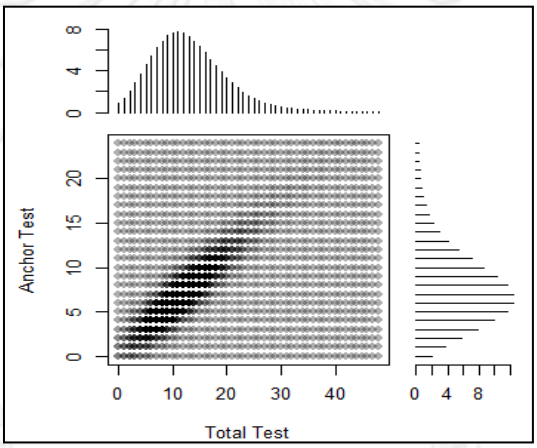
ตาราง 26 ค่าสถิติของกลุ่มผู้สอบอ้างอิงที่ใช้เป็นเกณฑ์ ที่ทำแบบสอบฉบับใหม่ (new) และฉบับอ้างอิง (reference) สำหรับการออกแบบการปรับเทียบที่เป็นเกณฑ์

	แบบสอบ	
	ฉบับใหม่	**ฉบับอ้างอิง**
จำนวนผู้สอบ	125	125
จำนวนข้อสอบ MC/CR	MC16/ CR 8	MC16/CR 8
ผู้ตรวจข้อสอบ CR	B	A
คะแนนเต็มของแบบสอบทั้งฉบับ	48	48
คะแนนสูงสุดของแบบสอบทั้งฉบับ	38	36
คะแนนต่ำสุดของแบบสอบทั้งฉบับ	3	3
ค่าเฉลี่ยของคะแนนสอบทั้งฉบับ	13.08	15.80
S.D ของคะแนนสอบทั้งฉบับ	7.27	8.03
มัธยฐานของคะแนนสอบทั้งฉบับ	12.00	14.00
ฐานนิยมของคะแนนสอบทั้งฉบับ	7.00	13.00
ความเบ้ของคะแนนสอบทั้งฉบับ	.97	.69
ความโด่งของคะแนนสอบทั้งฉบับ	3.56	2.78
จำนวนข้อสอบร่วมทั้งหมด	MC 8/CR 4	MC 8/CR 4
ผู้ตรวจข้อสอบร่วม CR	B	A
คะแนนเต็มของข้อสอบร่วมทั้งหมด	24	24
คะแนนสูงสุดของข้อสอบร่วมทั้งหมด	23	20
คะแนนต่ำสุดของข้อสอบร่วมทั้งหมด	1	2
ค่าเฉลี่ยของคะแนนจากข้อสอบร่วมทั้งหมด	7.68	9.12

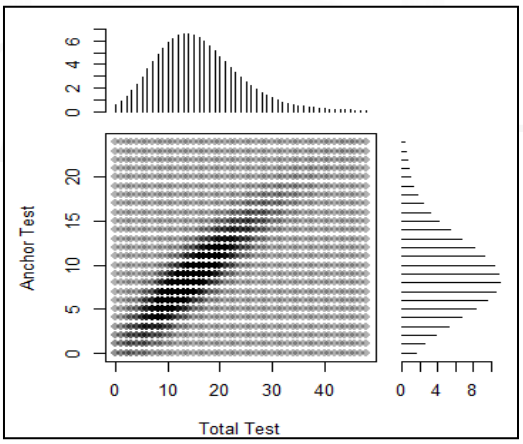
ตาราง 26 ค่าสถิติของกลุ่มผู้สอบอ้างอิงที่ใช้เป็นเกณฑ์ ที่ทำแบบสอบฉบับใหม่ (new) และฉบับอ้างอิง (reference) สำหรับการออกแบบการปรับเทียบที่เป็นเกณฑ์ (ต่อ)

	แบบสอบ	
	ฉบับใหม่	**ฉบับอ้างอิง**
S.D ของคะแนนจากข้อสอบร่วมทั้งหมด	4.27	4.56
มัธยฐานของข้อสอบร่วมทั้งหมด	7.00	8.00
ฐานนิยมของข้อสอบร่วมทั้งหมด	3.00	6.00
ความเบ้ของข้อสอบร่วมทั้งหมด	.77	.42
ความโด่งของข้อสอบร่วมทั้งหมด	3.25	2.32
ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนสอบทั้งฉบับและคะแนนของข้อสอบร่วม	.91**	.89**

**p< .01



ภาพ 11 การกระจายของคะแนนจากแบบสอบฉบับอ้างอิงทั้งฉบับและข้อสอบร่วมทั้งหมด สำหรับรูปแบบเกณฑ์



ภาพ 12 การกระจายของคะแนนจากแบบสอบฉบับใหม่ทั้งฉบับและข้อสอบร่วมทั้งหมด สำหรับรูปแบบเกณฑ์

ตอนที่ 3 ผลการวิเคราะห์ความสอดคล้องในการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบระหว่างผู้ตรวจ 1 และ 2

วิเคราะห์ค่าสัมประสิทธิ์สหพันธ์ของเพียร์สัน (r) ระหว่างการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบของผู้ตรวจ 1 ในวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม และการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบของผู้ตรวจ 2 ในวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจคนหนึ่ง ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม) โดยใช้โปรแกรม SPSS for Window Version 17.0

ผลการวิเคราะห์ พบว่า คะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบในแบบสอบฉบับอ้างอิง (reference) ในวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่มของผู้ตรวจ 1 และคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบในแบบสอบฉบับอ้างอิง (reference) ในวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจ 2 ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม) มีความสัมพันธ์กันทางบวก (r = .99) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (p = .00) ดังนั้นจึงกล่าวได้ว่าการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบระหว่างผู้ตรวจ 1 และ 2 มีความสอดคล้องกัน รายละเอียดดังตาราง 27

ตาราง 27 ค่าสัมประสิทธิ์สหพันธ์ของเพียร์สัน (r) ระหว่างการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบของผู้ตรวจ 1 ในวิธีการตรวจวิธีที่ 1: Non Trend และการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบของผู้ตรวจ 2 ในวิธีการตรวจวิธีที่ 2: Full Trend

สัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน		
	ผู้ตรวจ 1	ผู้ตรวจ 2
ผู้ตรวจ 1	1	
ผู้ตรวจ 2	.99**	1
ค่าเฉลี่ย	7.74	7.89
ส่วนเบี่ยงเบนมาตรฐาน	4.32	4.39

**p< .01

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตอนที่ 4 ผลการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) โดยใช้วิธีการปรับเทียบเชิงเส้นตรง

4.1) การปรับเทียบคะแนนโดยใช้วิธี CLE (Chained Linear Equating)

ผู้วิจัยได้นำแบบสอบเรื่องงานและพลังงานฉบับใหม่ (new) ไปทดสอบกับผู้สอบกลุ่มใหม่ จำนวน 1,084 คน และนำแบบสอบเรื่องงานและพลังงานฉบับอ้างอิง (reference) ไปทดสอบกับผู้สอบกลุ่มอ้างอิง จำนวน 400 คน แล้วนำไปปรับเทียบคะแนนด้วยวิธี CLE (Chained Linear Equating) โดยใช้โปรแกรม R (Albano, 2014) ซึ่ง Albano ได้อ้างอิงสมการที่ใช้ในการปรับเทียบสำหรับวิธีดังกล่าว ตามสมการที่อ้างไว้ใน Livingston, Dorans และ Wright (1990) ซึ่งก็เป็นสมการเดียวกับที่อ้างไว้ใน Davier (2011) ดังสมการด้านล่าง และได้ผลดังสมการ

$$y = f(x) = \mu_3(Y) + \frac{\sigma_3(Y)\sigma_1(A)}{\sigma_3(A)\sigma_1(X)}[x - \mu_1(X)] + \frac{\sigma_3(Y)}{\sigma_3(A)}[\mu_1(A) - \mu_3(A)]$$

จากตาราง 28 ผลการวิเคราะห์คะแนนของผู้สอบกลุ่มใหม่ไปหาคะแนนของผู้สอบกลุ่มอ้างอิง ซึ่งทำให้ได้คะแนนอยู่ในสเกลเดียวกัน พบว่า

วิธีการตรวจวิธีที่ 1 คือ วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม พบว่า คะแนนของผู้สอบกลุ่มใหม่จากแบบสอบฉบับใหม่ (new) ที่นำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิงจากแบบสอบฉบับอ้างอิง (reference) มีค่าลดลง 1 คะแนน ในช่วง 0 คะแนน ถึง 7 คะแนน เช่น ผู้สอบกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนนจริง 4 คะแนน ถ้านำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิงได้ 3 คะแนน เป็นต้น ขณะที่มีค่าเท่าเดิมในช่วง 8 คะแนน ถึง 16 คะแนน และมีค่าเพิ่มขึ้น 1 คะแนน ในช่วง 17 คะแนน ถึง 25 คะแนน อีกทั้งยังมีค่าเพิ่มขึ้น 2 คะแนน ในช่วง 26 คะแนน ถึง 34 คะแนน และมีค่าเพิ่มขึ้น 3 คะแนน ในช่วง 35 คะแนน ถึง 43 คะแนน นอกจากนี้ยังมีค่าเพิ่มขึ้น 4 คะแนน ในช่วง 44 คะแนน ถึง 48 คะแนน เช่น ผู้สอบกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนนจริง 44 คะแนน ถ้านำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิงได้ 48 คะแนน เป็นต้น

วิธีการตรวจวิธีที่ 2 คือ วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจคนหนึ่ง ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม) พบว่า คะแนนของผู้สอบกลุ่มใหม่จากแบบสอบฉบับใหม่ (new) ที่นำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิงจากแบบสอบฉบับอ้างอิง (reference) มีค่าลดลง 1 คะแนน ในช่วง 0 คะแนน ถึง 8 คะแนน เช่น ผู้สอบกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนนจริง 4 คะแนน ถ้านำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิงได้ 3 คะแนน เป็นต้น ขณะที่มีค่าเท่าเดิมในช่วง 9 คะแนน ถึง18 คะแนน และมีค่าเพิ่มขึ้น 1 คะแนน ในช่วง 19 คะแนน ถึง 27 คะแนน อีกทั้งยังมีค่าเพิ่มขึ้น2 คะแนน ในช่วง 28 คะแนน ถึง 37 คะแนน และมีค่าเพิ่มขึ้น 3 คะแนน ในช่วง 38 คะแนน ถึง 47 คะแนน นอกจากนี้ในช่วง 44 คะแนน ถึง 48 คะแนน มีค่าเพิ่มขึ้น4 คะแนน เช่น ผู้สอบกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนนจริง 44 คะแนน ถ้านำมาปรับเทียบกับคะแนนของนักเรียนกลุ่มอ้างอิงได้ 48 คะแนน เป็นต้น

วิธีการตรวจวิธีที่ 3 คือ วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม (ผู้ตรวจทั้ง 2 คน ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม โดยแบ่งกันตรวจคนละครึ่งหนึ่งของข้อสอบร่วม) พบว่า คะแนนของผู้สอบกลุ่มใหม่จากแบบสอบฉบับใหม่ (new) ที่นำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิงจากแบบสอบฉบับอ้างอิง (reference) มีค่าลดลง 1 คะแนนในช่วง 0 คะแนน ถึง 8 คะแนน เช่น ผู้สอบกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนนจริง 4 คะแนน ถ้านำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิงได้ 3 คะแนน เป็นต้น ขณะที่มีค่าเท่าเดิมในช่วง 9 คะแนน ถึง 17 คะแนน และมีค่าเพิ่มขึ้น 1 คะแนน ในช่วง 18 คะแนน ถึง 26 คะแนน อีกทั้งยังมีค่าเพิ่มขึ้น 2 คะแนน ในช่วง 27 คะแนน ถึง 35 คะแนน และมีค่าเพิ่มขึ้น 3 คะแนน ในช่วง 36 คะแนน ถึง 44 คะแนน เช่น ผู้สอบกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนนจริง 40 คะแนน ถ้านำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิงได้ 43 คะแนน เป็นต้นนอกจากนี้ยังมีค่าเพิ่มขึ้น 4 คะแนน ในช่วง 45 คะแนน ถึง 48 คะแนน

จะสังเกตเห็นว่า ในการปรับเทียบคะแนนสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจทั้ง 3 วิธี ด้วยวิธี CLE (Chained Linear Equating) ได้คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกันกับผู้สอบกลุ่มอ้างอิงเท่าเดิม ในช่วง 9 คะแนน ถึง 16 คะแนน และคะแนนของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกันกับผู้สอบกลุ่มอ้างอิง มีค่าต่างกันมากขึ้นในคะแนนส่วนท้าย อย่างเช่น ในการวิธีการตรวจทั้ง 3 วิธี ได้คะแนนจริงของผู้สอบกลุ่มใหม่ (new) ที่แปลงไปสู่สเกลเดียวกันกับผู้สอบกลุ่มอ้างอิงในช่วง 45 คะแนน ถึง 48 คะแนน มีค่าเพิ่มขึ้น 4 คะแนน เป็นต้น

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 28 ผลการวิเคราะห์คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกับผู้สอบกลุ่มอ้างอิง สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจทั้ง 3 วิธี ด้วยวิธี CLE (Chained Linear Equating)

คะแนนจริงของ	**คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกับผู้สอบกลุ่มอ้างอิง**		
ผู้สอบกลุ่มใหม่	Non Trend	Full Trend	Half Trend
0	-1	-1	-1
1	0	0	0
2	1	1	1
3	2	2	2
4	3	3	3
5	4	4	4
6	5	5	5
7	6	6	6
8	8	7	7
9	9	9	9
10	10	10	10
11	11	11	11
12	12	12	12
13	13	13	13
14	14	14	14
15	15	15	15
16	16	16	16
17	18	17	17
18	19	18	19
19	20	20	20
20	21	21	21
21	22	22	22
22	23	23	23
23	24	24	24
24	25	25	25
25	26	26	26
26	28	27	27
27	29	28	29
28	30	30	30
29	31	31	31
30	32	32	32
31	33	33	33
32	34	34	34
33	35	35	35
34	36	36	36

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 28 ผลการวิเคราะห์คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกับผู้สอบกลุ่มอ้างอิง สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจทั้ง 3 วิธี ด้วยวิธี CLE (Chained Linear Equating) (ต่อ)

คะแนนจริงของผู้สอบกลุ่มใหม่	คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกับผู้สอบกลุ่มอ้างอิง		
	Non Trend	Full Trend	Half Trend
35	38	37	37
36	39	38	39
37	40	39	40
38	41	41	41
39	42	42	42
40	43	43	43
41	44	44	44
42	45	45	45
43	46	46	46
44	48	47	47
45	49	48	49
46	50	49	50
47	51	50	51
48	52	52	52

4. 2) การปรับเทียบคะแนนโดยใช้วิธีของเลอวิน (Levine method)

ผู้วิจัยได้นำแบบสอบเรื่องงานและพลังงานฉบับใหม่ (new) ไปทดสอบกับผู้สอบกลุ่มใหม่ จำนวน 1,084 คน และนำแบบสอบเรื่องงานและพลังงานฉบับอ้างอิง (reference) ไปทดสอบกับผู้สอบกลุ่มอ้างอิง จำนวน 400 คน แล้วนำไปปรับเทียบคะแนนด้วยวิธีของเลอวิน (Levine method) โดยใช้โปรแกรม R (Albano, 2014) ซึ่ง Albano ได้อ้างอิงสมการที่ใช้ในการปรับเทียบสำหรับวิธีดังกล่าว ตามสมการที่อ้างไว้ใน Kolen และ Brennan (2004) ดังสมการด้านล่าง

$$l_{Ys}(x) = \left[\frac{\gamma_2}{\gamma_1}\right][x - \mu_1(X)] + \mu_2(Y) + \gamma_2[\mu_1(V) - \mu_2(V)]$$

จากตาราง 29 ผลการวิเคราะห์คะแนนของผู้สอบกลุ่มใหม่ไปหาคะแนนของผู้สอบกลุ่มอ้างอิงซึ่งทำให้ได้คะแนนอยู่ในสเกลเดียวกัน พบว่า

วิธีการตรวจวิธีที่ 1 คือ วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม พบว่า คะแนนของผู้สอบกลุ่มใหม่จากแบบสอบฉบับใหม่ (new) ที่นำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิงจากแบบสอบฉบับอ้างอิง (reference) มีค่าลดลง 2 คะแนน ที่ 0 คะแนน มีค่าลดลง 1 คะแนน ในช่วง 1 คะแนน ถึง 11 คะแนน เช่น นักเรียนกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนนจริง 4 คะแนน ถ้านำมาปรับเทียบกับคะแนนของนักเรียนกลุ่มอ้างอิงได้ 3 คะแนน เป็นต้น ขณะที่มีค่าเท่าเดิม ในช่วง 12 คะแนน ถึง 22 คะแนน และมีค่าเพิ่มขึ้น 1 คะแนน

ในช่วง 23 คะแนน ถึง 33 คะแนน อีกทั้งยังมีค่าเพิ่มขึ้น 2 คะแนน ในช่วง 34 คะแนน ถึง 46 คะแนน เช่น นักเรียนกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนน 36 คะแนน ถ้านำมาปรับเทียบกับคะแนนของนักเรียนกลุ่มอ้างอิงได้ 38 คะแนน เป็นต้น และมีค่าเพิ่มขึ้น 3 คะแนน ในช่วง 45 ถึง 48 คะแนน

วิธีการตรวจวิธีที่ 2 คือ วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจคนหนึ่ง ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม) พบว่า คะแนนของผู้สอบกลุ่มใหม่จากแบบสอบฉบับใหม่ (new) ที่นำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิงจากแบบสอบฉบับอ้างอิง (reference) มีค่าลดลง 2 คะแนน ที่ 0 คะแนน มีค่าลดลง 1 คะแนนในช่วง 1 คะแนน ถึง 13 คะแนน เช่น นักเรียนกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนนจริง 13 คะแนน ถ้านำมาปรับเทียบกับคะแนนของนักเรียนกลุ่มอ้างอิง ได้ 12 คะแนน เป็นต้น ขณะที่มีค่าเท่าเดิม ในช่วง 14 คะแนน ถึง 26 คะแนน และมีค่าเพิ่มขึ้น 1 คะแนน ในช่วง 27 คะแนน ถึง 39 คะแนน อีกทั้งยังมีค่าเพิ่มขึ้น 2 คะแนน ในช่วง 40 คะแนน ถึง 48 คะแนน เช่น นักเรียนกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนนจริง 42 คะแนน ถ้านำมาปรับเทียบกับคะแนนของนักเรียนกลุ่มอ้างอิงได้ 44 คะแนน เป็นต้น

วิธีการตรวจวิธีที่ 3 คือ วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม (ผู้ตรวจทั้ง 2 คน ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม โดยแบ่งกันตรวจคนละครึ่งหนึ่งของข้อสอบร่วม) พบว่า คะแนนของผู้สอบกลุ่มใหม่จากแบบสอบฉบับใหม่ (new) ที่นำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิงจากแบบสอบฉบับอ้างอิง (reference) มีค่าลดลง 2 คะแนน ในช่วง 0 คะแนน ถึง 1 คะแนน เช่น นักเรียนกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ ที่มีคะแนนจริง 1 คะแนน ถ้านำมาปรับเทียบกับคะแนนของนักเรียนกลุ่มอ้างอิงได้ -1 คะแนน และมีค่าลดลง 1 คะแนน ในช่วง 2 คะแนน ถึง 12 คะแนน เช่น นักเรียนกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนนจริง 11 คะแนน ถ้านำมาปรับเทียบกับคะแนนของนักเรียนกลุ่มอ้างอิงได้ 10 คะแนน เป็นต้น ขณะที่มีค่าเท่าเดิม ในช่วง 13 คะแนน ถึง 23 คะแนน และมีค่าเพิ่มขึ้น 1 คะแนนในช่วง 24 คะแนน ถึง 34 คะแนน เช่น นักเรียนกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนนจริง 34 คะแนน ถ้านำมาปรับเทียบกับคะแนนของนักเรียนกลุ่มอ้างอิงได้ 35 คะแนน เป็นต้น อีกทั้งยังมีค่าเพิ่มขึ้น 2 คะแนน ในช่วง 35 คะแนน ถึง 45 คะแนน เช่น นักเรียนกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนนจริง 35 คะแนน ถ้านำมาปรับเทียบกับคะแนนของนักเรียนกลุ่มอ้างอิงได้ 37 คะแนน และมีค่าเพิ่มขึ้น 3 คะแนน ในช่วง 46 คะแนน ถึง 48 คะแนน คะแนน เช่น นักเรียนกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนนจริง 46 คะแนน ถ้านำมาปรับเทียบกับคะแนนของนักเรียนกลุ่มอ้างอิงได้ 49 คะแนน

จะสังเกตเห็นว่าในการปรับเทียบคะแนนสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจทั้ง 3 วิธี ด้วยวิธีของเลอวิน (Levine method) ได้คะแนนของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกันกับผู้สอบกลุ่มอ้างอิงเท่าเดิมในช่วง 14 คะแนน ถึง 22 คะแนน และคะแนนของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกันกับผู้สอบกลุ่มอ้างอิง มีค่าต่างกันมากขึ้นในคะแนนส่วนท้าย เช่น ในวิธีการ

ตรวจวิธีที่ 1 และ 3 ได้คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกันกับผู้สอบกลุ่มอ้างอิง ในช่วง 46 คะแนน ถึง 48 คะแนน มีค่าเพิ่มขึ้น 3 คะแนน เป็นต้น

ตาราง 29 ผลการวิเคราะห์คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกับผู้สอบกลุ่มอ้างอิง สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจทั้ง 3 วิธี ด้วยวิธีของเลอวิน (Levine method)

คะแนนจริงจากของ	คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกับผู้สอบกลุ่มอ้างอิง		
ผู้สอบกลุ่มใหม่	Non Trend	Full Trend	Half Trend
0	-2	-2	-2
1	0	0	-1
2	1	1	1
3	2	2	2
4	3	3	3
5	4	4	4
6	5	5	5
7	6	6	6
8	7	7	7
9	8	8	8
10	9	9	9
11	10	10	10
12	12	11	11
13	13	12	13
14	14	14	14
15	15	15	15
16	16	16	16
17	17	17	17
18	18	18	18
19	19	19	19
20	20	20	20
21	21	21	21
22	22	22	22
23	24	23	23
24	25	24	25
25	26	25	26
26	27	26	27
27	28	28	28
28	29	29	29
29	30	30	30
30	31	31	31
31	32	32	32

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 29 ผลการวิเคราะห์คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกับผู้สอบกลุ่มอ้างอิง สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจทั้ง 3 วิธี (ต่อ) ด้วยวิธีของเลอวิน (Levine method)

คะแนนจริงจากของ	คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกับผู้สอบกลุ่มอ้างอิง		
ผู้สอบกลุ่มใหม่	Non Trend	Full Trend	Half Trend
32	33	33	33
33	34	34	34
34	36	35	35
35	37	36	37
36	38	37	38
37	39	38	39
38	40	39	40
39	41	40	41
40	42	42	42
41	43	43	43
42	44	44	44
43	45	45	45
44	46	46	46
45	48	47	47
46	49	48	49
47	50	49	50
48	51	50	51

ตอนที่ 5 ผลการปรับเทียบคะแนนสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ โดยใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed)

ในการปรับเทียบคะแนนสำหรับรูปแบบที่ใช้เป็นเกณฑ์ (criterion) จากคะแนนจากแบบสอบฉบับใหม่ (new) โดยผู้ตรวจ 2 ไปยังคะแนนจากแบบสอบฉบับอ้างอิงโดยผู้ตรวจ 1 นั้น ใช้โปรแกรม R (Albano, 2014)

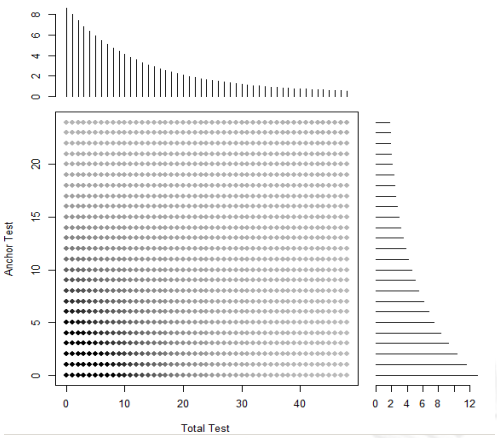
5.1 การปรับการแจกแจงของคะแนนจากแบบสอบฉบับใหม่ (new) ให้เป็นโค้งราบเรียบ (presmooth) ด้วยวิธีล็อกเชิงเส้น (loglinear)

ผลการปรับการแจกแจงของคะแนนจากแบบสอบฉบับใหม่ (new) ให้เป็นโค้งราบเรียบ ประกอบด้วย การวิเคราะห์ค่า degree สำหรับการแจกแจงของตัวแปรเดียว (univariate) และ xdegree สำหรับการแจกแจงของสองตัวแปร (bivariate) การวิเคราะห์โมเดล (nested model) ที่หลากหลาย หลังจากนั้นจึงนำข้อมูลที่ได้รับการปรับแล้วไปใช้ในการปรับเทียบคะแนน

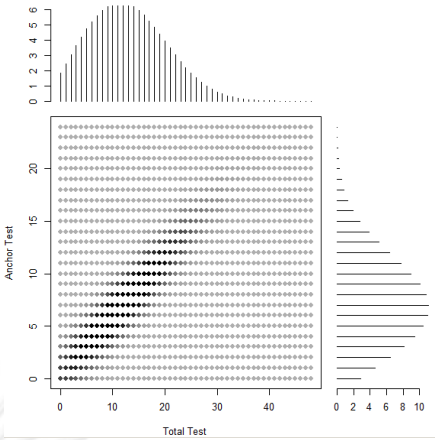
5.1.1 การวิเคราะห์ค่า degree และ xdegree สำหรับการแจกแจงของตัวแปรเดียว (univariate) และ การแจกแจงของสองตัวแปร (bivariate) ตามลำดับ

การใช้วิธีล็อกเชิงเส้น (loglinear) ในการแปลงคะแนนแบบสอบทั้งสองฉบับที่มีการแจกแจงเป็นโค้งไม่ราบเรียบให้เป็นคะแนนที่มีการแจกแจงเป็นโค้งราบเรียบ ต้องกำหนดค่า degree สำหรับการแจกแจงของตัวแปรเดียว และ xdegree สำหรับการแจกแจงสองตัวแปร โดย Albano (2014) แนะนำว่า อาจกำหนดค่า xdegree =1 เพื่อให้ง่ายในการแปลความหมายและอธิบายคะแนนที่ได้จากโมเดล แต่ให้ปรับค่า degree ให้มีค่าเหมาะสมกับข้อมูล โดยค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานที่ได้จากการปรับโค้งให้มีความราบเรียบควรมีค่าเท่ากับข้อมูลเดิมที่ยังไม่ได้ทำการปรับ

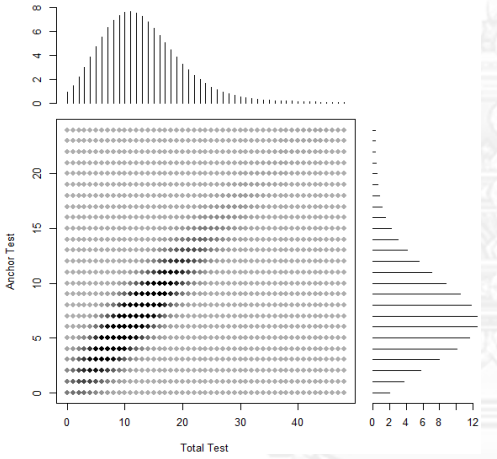
เมื่อพิจารณาการปรับการแจกแจงของคะแนนจากแบบสอบฉบับใหม่ (new) ให้เป็นโค้งที่ราบเรียบ เมื่อกำหนดให้ degree =1 และ xdegree=1 พบว่า ภาพการแจกแจงที่ได้จากการปรับโค้งให้ราบเรียบยังไม่มีลักษณะเป็นโค้งราบเรียบ ดังภาพ 13 ค่าเฉลี่ยของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดของข้อมูลเดิมเท่ากับค่าเฉลี่ยของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดที่ปรับโค้งให้ราบเรียบแล้ว ขณะที่ส่วนเบี่ยงเบนมาตรฐาน ความเบ้ และความโด่งของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดจากข้อมูลเดิมต่างจากส่วนเบี่ยงเบนมาตรฐาน ความเบ้ และความโด่งของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดที่ปรับโค้งให้ราบเรียบแล้ว รายละเอียดดังตาราง 30 แสดงให้เห็นว่าค่า degree = 1 ที่ใช้ยังไม่มีความเหมาะสม เมื่อกำหนดให้ degree = 2 และ xdegree = 1 พบว่า การแจกแจงที่ได้จากการปรับโค้งให้ราบเรียบมีลักษณะเป็นโค้งมากขึ้น ดังภาพ 14 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดของข้อมูลเดิมเท่ากับค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดที่ปรับโค้งให้ราบเรียบแล้ว ขณะที่ค่าความเบ้และความโด่งของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดจากข้อมูลเดิมต่างจากค่าความเบ้และความโด่งของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดที่ปรับโค้งให้ราบเรียบ รายละเอียดดังตาราง 30 เมื่อกำหนดให้ degree=3 และ xdegree=1 พบว่า การแจกแจงที่ได้จากการปรับโค้งให้ราบเรียบมีลักษณะเป็นโค้งราบเรียบ ดังภาพ 15 ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดของข้อมูลเดิมเท่ากับค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดที่ปรับโค้งให้ราบเรียบแล้ว ส่วนค่าความเบ้ของคะแนนข้อสอบร่วมทั้งหมดจากข้อมูลเดิมเท่ากับค่าความเบ้ของคะแนนข้อสอบร่วมทั้งหมดที่ปรับโค้งให้ราบเรียบ แต่ค่าความเบ้ของคะแนนแบบสอบทั้งฉบับจากข้อมูลเดิมต่างจากค่าความเบ้ของคะแนนแบบสอบทั้งฉบับที่ปรับโค้งให้ราบเรียบเพียงเล็กน้อย ขณะที่ค่าความโด่งของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดจากข้อมูลเดิมต่างจากค่าความเบ้ของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดที่ปรับโค้งให้ราบเรียบ รายละเอียดดังตาราง 30 แต่เมื่อกำหนดให้ degree = 4 และ xdegree = 1 พบว่า ภาพการแจกแจงที่ได้จากการปรับโค้งให้ราบเรียบมีลักษณะเป็นโค้งที่แย่กว่า degree = 3 และ 2 ดังภาพ 16 ดังนั้นควรเลือกใช้ degree = 3 และ xdegree = 1 ในการปรับการแจกแจงของคะแนนจากแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดให้เป็นโค้งราบเรียบ



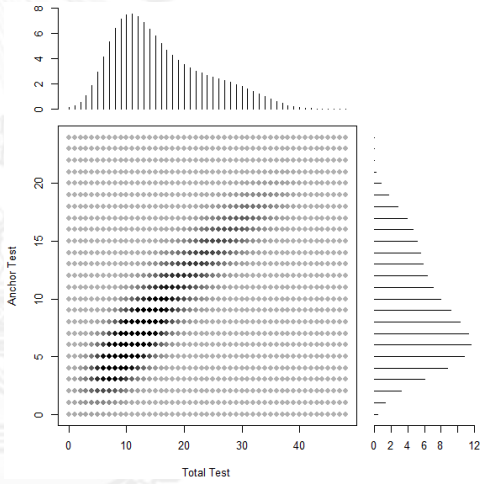
ภาพ 13 การแจกแจงของคะแนนจากแบบสอบฉบับใหม่ เมื่อ degree =1 xdegree =1



ภาพ 14 การแจกแจงของคะแนนจากแบบสอบฉบับใหม่ เมื่อ degree =2 xdegree =1



ภาพ 15 การแจกแจงของคะแนนจากแบบสอบฉบับใหม่ เมื่อ degree =3 xdegree =1



ภาพ 16 การแจกแจงของคะแนนจากแบบสอบฉบับใหม่ เมื่อ degree =4 xdegree =1

ตาราง 30 ค่าสถิติพื้นฐานของคะแนนจากแบบสอบฉบับใหม่ (new) ที่ระดับ degree ต่างๆ

ค่า degree และ xdegree	คะแนนแบบสอบฉบับใหม่	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน	ความเบ้	ความโด่ง
degree = 1	ทั้งฉบับ	13.08	7.271908	.9722094	3.564064
xdegree = 1	ข้อสอบร่วม	7.68	4.270303	.7685213	3.245522
	ทั้งฉบับ(ปรับ)	13.08	11.993124	1.0505716	3.253706
	ข้อสอบร่วม(ปรับ)	7.68	6.644759	.7835232	2.563244
degree = 2	ทั้งฉบับ	13.08	7.271908	.9722094	3.564064
xdegree = 1	ข้อสอบร่วม	7.68	4.270303	.7685213	3.245522
	ทั้งฉบับ(ปรับ)	13.08	7.271908	.4262703	2.797168
	ข้อสอบร่วม(ปรับ)	7.68	4.270303	.4124942	2.788039

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 30 ค่าสถิติพื้นฐานของคะแนนจากแบบสอบฉบับใหม่ (new) ที่ระดับ degree ต่างๆ (ต่อ)

ค่า degree และ xdegree	คะแนนแบบสอบฉบับใหม่	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน	ความเบ้	ความโด่ง
degree = 3	ทั้งฉบับ	13.08	7.271908	.9722094	3.564064
xdegree = 1	ข้อสอบร่วม	7.68	4.270303	.7685213	3.245522
	ทั้งฉบับ(ปรับ)	13.08	7.271908	.9722095	4.597208
	ข้อสอบร่วม(ปรับ)	7.68	4.270303	.7685213	3.8432899

5.1.2 การวิเคราะห์โมเดล (nested model) ที่หลากหลาย

หลังจากที่ได้ค่า degree และ xdegree ที่ต้องการแล้ว ก็นำค่าดังกล่าวมาสร้างโมเดล (nested model) ที่หลากหลาย ตั้งแต่โมเดลอย่างง่ายไปจนถึงโมเดลที่ซับซ้อน พร้อมทั้งตารางวิเคราะห์ค่า Deviance ซึ่งผลจากการวิเคราะห์พบว่า ค่า AIC และค่า BIC ของโมเดล 7 มีค่าต่ำที่สุด คือ 466.89 และ 507.78 ซึ่ง Albano (2014) แนะนำว่าโมเดลที่มีความเหมาะสมกับข้อมูลต้องมีค่า AIC และค่า BIC ต่ำ ๆ ดังนั้นโมเดลที่มีความเหมาะสมกับข้อมูลมากที่สุด คือ โมเดล 7 นอกจากนี้จะสังเกตเห็นว่าโมเดล 7 ที่มีค่า AIC และ BIC ต่ำที่สุด มีค่า Deviance มากที่สุดอีกด้วย รายละเอียดดังตาราง 31 โดยโมเดล 7 มีฟังก์ชั่นของคะแนน (score function) คือ

$$x^1 + x^2 + x^3 + v^1 + v^2 + v^3 + xv^1$$

โดยที่ x คือ คะแนนของแบบสอบฉบับใหม่ (new) ทั้งฉบับ

v คือ คะแนนของข้อสอบร่วมจากแบบสอบฉบับใหม่ (new)

จะเห็นว่าเลขกำลังของ x และ v ที่มากที่สุด ก็คือ ค่า degree = 3 สำหรับการแจกแจงของตัวแปรเดียว ขณะที่กำลัง xv มากที่สุด ก็คือ ค่า xdegree = 1 สำหรับการแจกแจงของสองตัวแปร โดยที่โมเดลอื่นๆ มีฟังก์ชั่นของคะแนน ดังนี้

โมเดล 1 : x^1

โมเดล 2 : $x^1 + x^2$

โมเดล 3 : $x^1 + x^2 + x^3$

โมเดล 4 : $x^1 + x^2 + x^3 + v^1$

โมเดล 5 : $x^1 + x^2 + x^3 + v^1 + v^2$

โมเดล 6 : $x^1 + x^2 + x^3 + v^1 + v^2 + v^3$

โมเดล 7 : $x^1 + x^2 + x^3 + v^1 + v^2 + v^3 + xv^1$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 31 ผลการวิเคราะห์ค่า Deviance ของโมเดลต่างๆ

โมเดล	AIC	BIC	Df	Deviance
1	833.22	843.44		
2	780.93	796.26	1	54.289
3	774.69	795.13	1	8.240
4	730.07	755.63	1	46.611
5	681.28	711.94	1	50.795
6	677.62	713.39	1	5.663
7	466.89	507.78	1	212.725

5.2 การปรับการแจกแจงของคะแนนจากแบบสอบฉบับอ้างอิง (reference) ให้เป็นโค้งราบเรียบ (presmooth) ด้วยวิธีล็อกเชิงเส้น (loglinear)

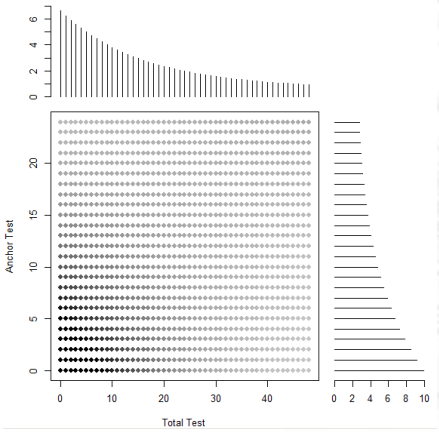
ผลการปรับการแจกแจงของคะแนนจากแบบสอบฉบับอ้างอิง (reference) ให้เป็นโค้งราบเรียบ ประกอบด้วย การวิเคราะห์ค่า degree และ xdegree สำหรับการแจกแจงของตัวแปรเดียว (univariate) และ การแจกแจงของสองตัวแปร (bivariate) และการวิเคราะห์โมเดล (nested model) ที่หลากหลาย หลังจากนั้นจึงนำข้อมูลได้รับการปรับแล้วไปใช้ในการปรับเทียบคะแนน

5.2.1 การวิเคราะห์ค่า degree และ xdegree สำหรับการแจกแจงของตัวแปรเดียว (univariate) และ การแจกแจงของสองตัวแปร (bivariate)

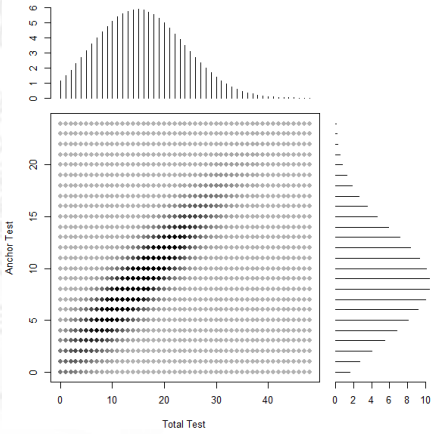
การใช้วิธีล็อกเชิงเส้น (loglinear) ในการแปลงคะแนนแบบสอบทั้งสองฉบับที่มีการแจกแจงเป็นโค้งไม่ราบเรียบให้เป็นคะแนนที่มีการแจกแจงเป็นโค้งราบเรียบ ต้องกำหนดค่า degree สำหรับการแจกแจงของตัวแปรเดียว และ xdegree สำหรับการแจกแจงสองตัวแปร โดย Albano (2014) แนะนำว่า อาจกำหนดค่า xdegree =1 เพื่อให้ง่ายในการแปลความหมายและอธิบายคะแนนที่ได้จากโมเดล แต่ให้ปรับค่า degree ให้มีค่าเหมาะสมกับข้อมูล โดยค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานที่ได้จากการปรับโค้งให้มีความราบเรียบควรมีค่าเท่ากับข้อมูลเดิมที่ยังไม่ได้ทำการปรับ

เมื่อพิจารณาการปรับการแจกแจงของคะแนนจากแบบสอบฉบับใหม่ (new) ให้เป็นโค้งที่ราบเรียบ เมื่อกำหนดให้ degree =1 และ xdegree =1 พบว่า ภาพการแจกแจงที่ได้จากการปรับโค้งให้ราบเรียบยังไม่มีลักษณะเป็นโค้งราบเรียบ ดังภาพ 17 ค่าเฉลี่ยของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดจากข้อมูลเดิมเท่ากับค่าเฉลี่ยของคะแนนจากแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดที่ปรับโค้งให้ราบเรียบแล้ว ขณะที่ส่วนเบี่ยงเบนมาตรฐาน ความเบ้ และความโด่งของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดจากข้อมูลเดิมต่างจากส่วนเบี่ยงเบนมาตรฐาน ความเบ้ และความโด่งของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดที่ปรับโค้งให้ราบเรียบแล้ว รายละเอียดดังตาราง 32 แสดงให้เห็นค่า degree =1 ที่ใช้ยังไม่มีความเหมาะสม เมื่อกำหนดให้ degree =2 และ xdegree =1 พบว่า การแจกแจงที่ได้จากการปรับโค้งให้ราบเรียบมีลักษณะเป็นโค้งมากขึ้น ดังภาพ 18 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนแบบสอบทั้งฉบับ

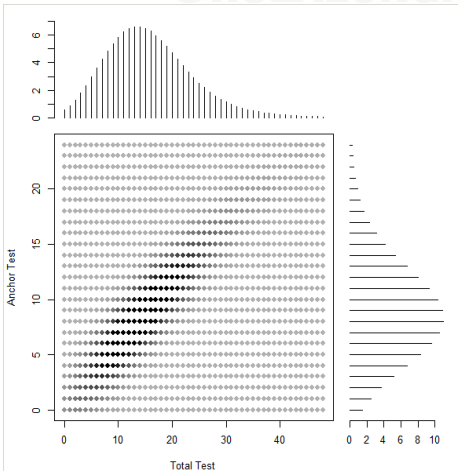
และข้อสอบร่วมทั้งหมดของข้อมูลเดิมเท่ากับค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดที่ปรับโค้งให้ราบเรียบแล้ว ขณะที่ค่าความเบ้และความโด่งของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดจากข้อมูลเดิมต่างจากค่าความเบ้และความโด่งของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดที่ปรับโค้งให้ราบเรียบ รายละเอียดดังตาราง 32 เมื่อกำหนดให้ degree =3 และ xdegree =1 พบว่า การแจกแจงที่ได้จากการปรับโค้งให้ราบเรียบมีลักษณะเป็นโค้งราบเรียบ ดังภาพ 20 ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และความเบ้ของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดของข้อมูลเดิมเท่ากับค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และความเบ้ของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดที่ปรับโค้งให้ราบเรียบแล้ว ขณะที่ค่าความโด่งของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดจากข้อมูลเดิมต่างจากค่าความโด่งของคะแนนแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดที่ปรับโค้งให้ราบเรียบ รายละเอียดดังตาราง 32 แต่เมื่อกำหนดให้ degree =4 และ xdegree =1 พบว่า การแจกแจงที่ได้จากการปรับโค้งให้ราบเรียบมีลักษณะเป็นโค้งที่แย่กว่า degree =3 และ 2 ดังภาพ 19 ดังนั้นควรเลือกใช้ degree =3 และ xdegree =1 ในการปรับการแจกแจงของคะแนนจากแบบสอบทั้งฉบับและข้อสอบร่วมทั้งหมดให้เป็นโค้งราบเรียบ



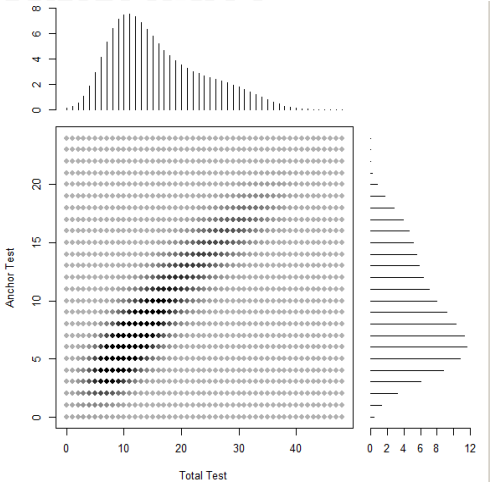
ภาพ 17 การแจกแจงของคะแนนจากแบบสอบฉบับอ้างอิง เมื่อ degree =1 xdegree =1



ภาพ 18 การแจกแจงของคะแนนจากแบบสอบฉบับอ้างอิง เมื่อ degree =2 xdegree =1



ภาพ 20 การแจกแจงของคะแนนจากแบบสอบฉบับอ้างอิง เมื่อ degree =3 xdegree =1



ภาพ 19 การแจกแจงของคะแนนจากแบบสอบฉบับอ้างอิง เมื่อ degree =4 xdegree =1

ตาราง 32 ค่าสถิติพื้นฐานของคะแนนจากแบบสอบฉบับอ้างอิง (reference) ที่ระดับ degree ต่างๆ

ค่า degree และ xdegree	คะแนนแบบสอบฉบับอ้างอิง	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน	ความเบ้	ความโด่ง
degree = 1	ทั้งฉบับ	15.08	8.031189	.6948529	2.778868
xdegree = 1	ข้อสอบร่วม	9.12	4.558933	.4186632	2.321760
	ทั้งฉบับ(ปรับ)	15.08	8.031189	.7497764	2.495328
	ข้อสอบร่วม(ปรับ)	9.12	4.558933	.5036466	2.070749
degree = 2	ทั้งฉบับ	15.08	8.031189	.6948529	2.778868
xdegree = 1	ข้อสอบร่วม	9.12	4.558933	.4186632	2.321760
	ทั้งฉบับ(ปรับ)	15.08	8.031189	.3044613	2.693444
	ข้อสอบร่วม(ปรับ)	9.12	4.558933	..2811067	2.675319
degree = 3	ทั้งฉบับ	15.08	8.031189	.6948529	2.778868
xdegree = 1	ข้อสอบร่วม	9.12	4.558933	.4186632	2.321760
	ทั้งฉบับ(ปรับ)	15.08	8.031189	.6948529	3.617570
	ข้อสอบร่วม(ปรับ)	9.12	4.558933	.4186632	2.986618

5.1.2 การวิเคราะห์โมเดล (nested model) ที่หลากหลาย

หลังจากที่ได้ค่า degree และ xdegree ที่ต้องการแล้ว ก็นำค่าดังกล่าวมาสร้างโมเดล (nested model) ที่หลากหลาย ตั้งแต่โมเดลอย่างง่ายไปจนถึงโมเดลที่ซับซ้อน พร้อมทั้งตารางวิเคราะห์ค่า Deviance ซึ่งผลจากการวิเคราะห์พบว่า ค่า AIC และค่า BIC ของโมเดล 7 มีค่าต่ำที่สุด คือ 512.28 และ 553.16 ซึ่ง Albano (2014) แนะนำว่าโมเดลที่มีความเหมาะสมกับข้อมูลต้องมีค่า AIC และค่า BIC ต่ำ ๆ ดังนั้นโมเดลที่มีความเหมาะสมกับข้อมูลมากที่สุด คือ โมเดล 7 นอกจากนี้จะสังเกตเห็นว่าโมเดล 7 ที่มีค่า AIC และ BIC ต่ำที่สุด มีค่า Deviance มากที่สุดอีกด้วย รายละเอียดดังตาราง 33 โดยโมเดล 7 มีฟังก์ชั่นของคะแนน (score function) คือ

$$x^1 + x^2 + x^3 + v^1 + v^2 + v^3 + xv^1$$

โดยที่ x คือ คะแนนของแบบสอบฉบับอ้างอิง (reference) ทั้งฉบับ

v คือ คะแนนของข้อสอบร่วมจากแบบสอบฉบับอ้างอิง (reference)

จะเห็นว่าเลขกำลังของ x และ v ที่มากที่สุด ก็คือ ค่า degree = 3 สำหรับการแจกแจงของตัวแปรเดียว ขณะที่กำลัง xv มากที่สุด ก็คือ ค่า xdgree = 1 สำหรับการแจกแจงของสองตัวแปร โดยที่โมเดลอื่นๆ มีฟังก์ชั่นของคะแนน ดังนี้

โมเดล 1 : x^1
โมเดล 2 : $x^1 + x^2$
โมเดล 3 : $x^1 + x^2 + x^3$
โมเดล 4 : $x^1 + x^2 + x^3 + v^1$
โมเดล 5 : $x^1 + x^2 + x^3 + v^1 + v^2$
โมเดล 6 : $x^1 + x^2 + x^3 + v^1 + v^2 + v^3$
โมเดล 7 : $x^1 + x^2 + x^3 + v^1 + v^2 + v^3 + xv^1$

ตาราง 33 ผลการวิเคราะห์ค่า Deviance ของโมเดลต่างๆ

โมเดล	AIC	BIC	Df	Deviance
1	841.02	851.25		
2	781.37	796.71	1	61.651
3	778.36	798.81	1	5.010
4	760.09	785.65	1	20.269
5	705.13	735.79	1	56.969
6	705.72	741.49	1	1.408
7	512.28	553.16	1	212.725

5.3 วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ (chained equipercentile equating)

หลังจากนั้นนำคะแนนที่ได้รับการปรับจาก model 7 ไปใช้ในการปรับเทียบคะแนนในการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ (SG design) โดยใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ (chained equipercentile equating) ด้วยโปรแกรม R (Albano, 2014) โดยฟังก์ชั่นการปรับเทียบที่ใช้ในการแปลงคะแนนจากแบบสอบ X ไปยังสเกลของแบบสอบ Y ก็คือ $e_{Y(chain)}$ โดยที่ $e_{Y(chain)} = e_{Y2}[e_{V1}(x)]$

จากตาราง 34 ผลการวิเคราะห์คะแนนของผู้สอบกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์จากแบบสอบฉบับใหม่ (new) ไปหาคะแนนจากแบบสอบฉบับอ้างอิง (reference) สำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ ซึ่งทำให้ได้คะแนนอยู่ในสเกลเดียวกัน พบว่า คะแนนของของผู้สอบกลุ่มใหม่จากแบบสอบฉบับใหม่ (new) ที่นำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิงจากแบบสอบฉบับอ้างอิง (reference) มีค่าเพิ่มขึ้น 1 คะแนน ในช่วง 1 คะแนน ถึง 23 คะแนน และ ช่วง 40 คะแนน ถึง 45 คะแนน เช่น ผู้สอบกลุ่มที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนน 40 คะแนน ถ้านำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิง ได้ 41 คะแนน เป็นต้น ขณะที่มีค่าเท่าเดิมที่ 0 คะแนน และในช่วง 46 คะแนน ถึง 48 คะแนน อีกทั้งยังมีค่าเพิ่มขึ้น 2 คะแนน ในช่วง 24 คะแนน ถึง 39 คะแนน เช่น ผู้สอบกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ที่มีคะแนน 32 คะแนน ถ้านำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิงได้ 34 คะแนน เป็นต้น

ตาราง 34 ผลการวิเคราะห์คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกับผู้สอบกลุ่มอ้างอิง สำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ ด้วยวิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed)

คะแนนจริงจากของผู้สอบกลุ่มใหม่	คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกับผู้สอบกลุ่มอ้างอิง
0	0
1	2
2	3
3	4
4	5
5	6
6	7
7	8
8	9
9	10
10	11
11	12
12	13
13	14
14	15
15	16
16	17
17	18
18	19
19	20
20	21
21	22
22	23
23	24
24	26
25	27
26	28
27	29
28	30
29	31
30	32
31	34
32	34
33	35

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 24 ผลการวิเคราะห์คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกับผู้สอบกลุ่มอ้างอิง สำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ ด้วยวิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed) (ต่อ)

คะแนนจริงจากของผู้สอบกลุ่มใหม่	คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกับผู้สอบกลุ่มอ้างอิง
34	36
35	37
36	38
37	39
38	40
39	41
40	41
41	42
42	43
43	44
44	45
45	46
46	46
47	47
48	48

ตอนที่ 6 ผลการเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี ระหว่างวิธีของเลอวิน (Levine method) และวิธี CLE (Chained linear equating)

การเปรียบเทียบวิธีการปรับเทียบเชิงเส้นตรง (equating method) 2 วิธี ระหว่างวิธีของเลอวิน (Levine method) และวิธี CLE (Chained Linear Equating) ภายใต้การออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ซึ่งพิจารณาจากค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (Root Mean Squared Difference: RMSD) ซึ่งคำนวณผ่านโปรแกรม Microsoft excel ด้วยสมการ

$$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$$

โดยเปรียบเทียบกับคะแนนที่ได้จากการปรับเทียบภายใต้การออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์

ผลการวิเคราะห์จากตาราง 35 พบว่า ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธี CLE (Chained Linear Equating) มีค่าต่ำกว่าวิธีของเลอวิน (Levine method) ในทุกวิธีการตรวจของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) โดยค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD)

ของวิธี CLE (Chained Linear Equating) มีค่ามากที่สุดอยู่ที่วิธีการตรวจวิธีที่ 2: Full Trend และ 3: Half Trend คือ 1.481 และมีค่าต่ำที่สุดอยู่ที่วิธีการตรวจวิธีที่ 1: Non Trend คือ 1.421 ขณะที่ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีของเลอวิน (Levine method) มีค่ามากที่สุดอยู่ที่วิธีการตรวจวิธีที่ 2: Full Trend รองลงมา คือ วิธีการตรวจวิธีที่ 3: Half Trend และ 1: Non Trend ตามลำดับ โดยมีค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) เท่ากับ 1.804, 1.793 และ 1.717 ตามลำดับ นอกจากนี้ยังพบว่า ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในแต่ละวิธีการตรวจ จากวิธีการปรับเทียบวิธีเดียวกันมีค่าใกล้เคียงกัน โดยรายละเอียดของการคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ที่แต่ละจุดคะแนนอยู่ที่ภาคผนวก ซ

ตาราง 35 ผลการวิเคราะห์ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ระหว่างวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

วิธีการตรวจ	วิธีการปรับเทียบคะแนน	
	Chain Linear	Levine
วิธีที่ 1: Non Trend	1.421	1.717
วิธีที่ 2: Full Trend	1.481	1.804
วิธีที่ 3: Half Trend	1.481	1.793

ตอนที่ 7 ผลการเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของวิธีการตรวจ 3 วิธี สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

ในการเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของวิธีการตรวจ 3 วิธี สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ได้แก่ 1) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม 2) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจคนหนึ่ง ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม) 3) วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม (ผู้ตรวจทั้ง 2 คน ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม โดยแบ่งกันตรวจคนละครึ่งหนึ่งของข้อสอบร่วม) ผู้วิจัยใช้โปรแกรม R (Albano, 2014) ในการวิเคราะห์หาค่าความลำเอียง (bias) ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสองในแต่ละจุดคะแนน หลังจากนั้นผู้วิจัยนำค่าดังกล่าวมาหาเป็นค่าสถิติสรุปในแต่ละรูปแบบของการออกแบบการปรับเทียบคะแนน ได้แก่ 1) ค่า Weighted average root mean square bias ซึ่งก็คือ $\sqrt{\sum_i w_i Bias_i^2}$ 2) ค่า Weighted average standard error of equating

ซึ่งคือ $\sqrt{\sum_i w_i SEE_i^2}$ และ 3) ค่า Weighted average RMSE ซึ่งก็คือ $\sqrt{\sum_i w_i RMSE_i^2}$ หรืออาจเรียกอย่างง่ายว่าเป็นค่าความลำเอียง (bias) ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ตามลำดับ โดยผู้วิจัยคำนวณค่าดังกล่าวผ่านโปรแกรม Excel (ดูรายละเอียดของการคำนวณได้จากภาคผนวก ฌ)

7.1) ผลการวิเคราะห์สถิติบูทสแตรป (bootstrapped statistics) สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ด้วยวิธี CLE (Chained Linear Equating)

ผลการวิเคราะห์สถิติบูทสแตรป (bootstrapped statistics) ที่แสดงไว้ในตาราง 36 ได้แก่ ค่า Weighted average root mean square bias ค่า Weighted average standard error of และค่า Weighted average RMSE หรืออาจเรียกอย่างง่ายว่าเป็นค่าความลำเอียง (bias) ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ตามลำดับ ซึ่งเป็นค่าสรุปของค่าความลำเอียง (bias) ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ในแต่ละจุดคะแนนจากเทคนิคการสุ่มแบบ Bootstrap resampling (สุ่มซ้ำจำนวน 500 ครั้ง) ในแต่ละวิธีการตรวจสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ด้วยวิธี CLE (Chained Linear Equating)

เมื่อพิจารณาที่วิธี CLE (Chained Linear Equating) พบว่า วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม มีค่าความลำเอียง (bias) ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) เท่ากับ .818 .260 และ .858 ตามลำดับ ขณะที่วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบโดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจคนหนึ่ง ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม) มีค่าความลำเอียง (bias) ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) เท่ากับ .803 .261 และ .844 ตามลำดับ ส่วนวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม (ผู้ตรวจทั้ง 2 คน ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม โดยแบ่งกันตรวจคนละครึ่งหนึ่งของข้อสอบร่วม) มีค่าความลำเอียง (bias) ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) เท่ากับ .802 .251 และ .840 ตามลำดับ ซึ่งเมื่อเปรียบเทียบค่าสถิติบูทสแตรปในแต่ละวิธีการตรวจทั้ง 3 วิธี พบว่า วิธีการตรวจที่มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำที่สุด คือ วิธีการตรวจวิธีที่ 3: Half Trend, 2: Full Trend และ 1: Non Trend ตามลำดับ ส่วนวิธีการตรวจที่มีค่าความลำเอียง (bias) ต่ำที่สุด คือ วิธีการตรวจวิธีที่ 3: Half Trend, 2: Full Trend และ 1: Non Trend ตามลำดับ ส่วนวิธีการตรวจที่มีค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) ต่ำที่สุด ได้แก่ วิธีการตรวจวิธีที่ 3: Half Trend, 1: Non Trend และ

2: Full Trend ตามลำดับ กล่าวคือ วิธีการตรวจวิธีที่ 3: Half Trend มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ค่าความลำเอียง (bias) และค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) ต่ำที่สุด วิธีการตรวจวิธีที่ 2: Full Trend มีค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) มากที่สุด ส่วนวิธีการตรวจวิธีที่ 1: Non Trend มีค่าความลำเอียง (bias) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) มากที่สุด

7.2) ผลการวิเคราะห์สถิติบูทสแตรป (bootstrapped statistics) สำหรับสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ด้วยวิธีของเลอวิน (Levine method)

ผลการวิเคราะห์สถิติบูทสแตรป (bootstrapped statistics) ที่แสดงไว้ในตาราง 36 ได้แก่ ค่า Weighted average root mean square bias ค่า Weighted average standard error of equating และค่า Weighted average RMSE จากเทคนิคการสุ่มแบบ Bootstrap resampling (สุ่มซ้ำจำนวน 500 ครั้ง) ในแต่ละวิธีการตรวจสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ด้วยวิธีของเลอวิน (Levine method)

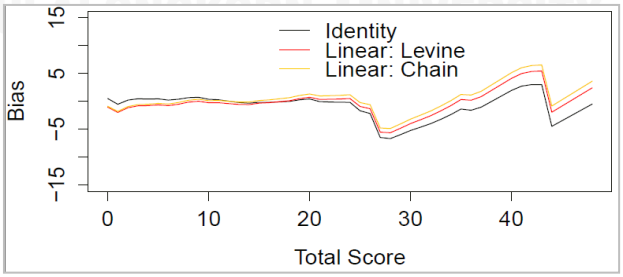
เมื่อพิจารณาที่วิธีของเลอวิน (Levine method) พบว่า วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม มีค่าความลำเอียง (bias) ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) เท่ากับ .976 .274 และ 1.014 ตามลำดับ ขณะที่วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจคนหนึ่งตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม) มีค่าความลำเอียง (bias) ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) เท่ากับ .975 .282 และ 1.015 ตามลำดับ ส่วนวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม (ผู้ตรวจทั้ง 2 คน ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม โดยแบ่งกันตรวจคนละครึ่งหนึ่งของข้อสอบร่วม) มีค่าความลำเอียง (bias) ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) เท่ากับ .988 .271 และ 1.024 ตามลำดับ ซึ่งเมื่อเปรียบค่าสถิติบูทสแตรปในแต่ละวิธีการตรวจทั้ง 3 วิธี ด้วยวิธีของเลอวิน (Levine method) พบว่า วิธีการตรวจที่มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำที่สุด คือ วิธีการตรวจวิธีที่ 1: Non Trend, 2: Full Trend และ 3: Half Trend ตามลำดับ ส่วนวิธีการตรวจวิธีที่มีค่าความลำเอียง (bias) ต่ำที่สุด คือ วิธีการตรวจวิธีที่ 2: Full Trend, 1: Non Trend และ 3: Half Trend ตามลำดับ ส่วนวิธีการตรวจที่มีค่าความคลาดเคลื่อน ในการปรับเทียบ (equating error) ต่ำที่สุด ได้แก่ วิธีการตรวจวิธีที่ 3: Half Trend, 1: Non Trend และ 2: Full Trend ตามลำดับ กล่าวคือ วิธีการตรวจวิธีที่ 2: Full Trend มีค่าความลำเอียง (bias) ต่ำที่สุด แต่มีค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) มากที่สุด ส่วนวิธีการตรวจวิธีที่ 3: Half Trend มีค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) ต่ำที่สุด แต่มีค่าความลำเอียง (bias) และค่ารากที่

สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) มากที่สุด วิธีการตรวจวิธีที่ 1: Non Trend มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำที่สุด แต่ถ้าเปรียบเทียบค่าสถิติบูทสแตรปทั้ง 3 ค่า ที่ได้จากวิธีการปรับเทียบเชิงเส้นตรงทั้ง 2 วิธี พบว่า ค่าสถิติบูทสแตรปทั้ง 3 ค่า ที่ได้จากวิธีของเลอวิน (Levine method) ในแต่ละวิธีการตรวจมีค่ามากกว่าวิธี CLE (Chained Linear Equating)

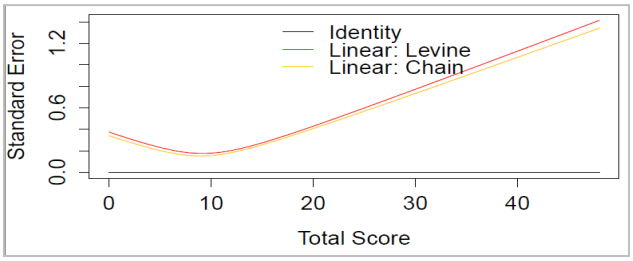
นอกจากนี้เมื่อพิจารณาค่าความลำเอียง ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสองในแต่ละจุดคะแนน ($Bias_i$, SEE_i และ $RMSE_i$ ตามลำดับ) ที่ได้จากวิธีการปรับเทียบเชิงเส้นตรงทั้ง 2 วิธี (รายละเอียดอยู่ในภาคผนวก ฌ) สำหรับวิธีการตรวจวิธีที่ 1: Non Trend พบว่า ค่า $Bias_i$ และค่า $RMSE_i$ มีค่าค่อนข้างสม่ำเสมอ และมีค่าเข้าใกล้ 0 ในช่วง 0 คะแนน ถึง 20 คะแนน และมีค่าไม่สม่ำเสมอในช่วงคะแนนปลาย ๆ ดังภาพ 21 และ 23 ตามลำดับ ขณะที่ค่า SEE_i มีค่าเข้าใกล้ 0 ในช่วง 10 คะแนน และค่อยๆ มีค่าเพิ่มขึ้นอย่างสม่ำเสมอ เนื่องจากกราฟที่ได้เป็นเส้นตรง ดังภาพ 22 เช่นเดียวกับวิธีการตรวจวิธีที่ 2: Full Trend และ 3: Half Trend ดังภาพ 24 - 29 ซึ่งลักษณะกราฟของค่า $Bias_i$ ค่า SEE_iและค่า $RMSE_i$ ที่ได้ในแต่ละวิธีการตรวจจากวิธีการปรับเทียบเดียวกันไม่แตกต่างกันมากนัก

ตาราง 36 ผลการวิเคราะห์สถิติบูทสแตรป (bootstrapped statistics) สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี

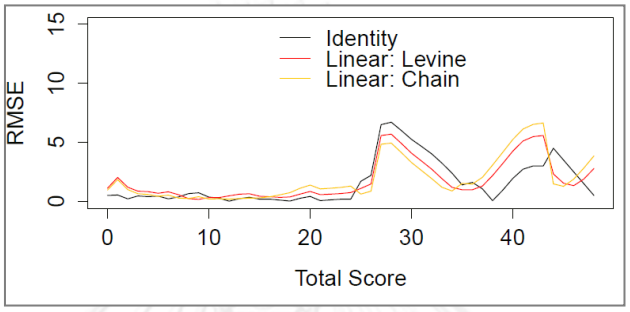
วิธีการตรวจ	วิธี CLE (Chained Linear Equating)			วิธีของเลอวิน (Levine method)		
	BIAS	EQUATING ERROR	RMSE	BIAS	EQUATING ERROR	RMSE
Non Trend	0.818	0.260	0.858	0.976	0.274	1.014
Full Trend	0.803	0.261	0.844	0.975	0.282	1.015
Half Trend	0.802	0.251	0.840	0.988	0.271	1.024



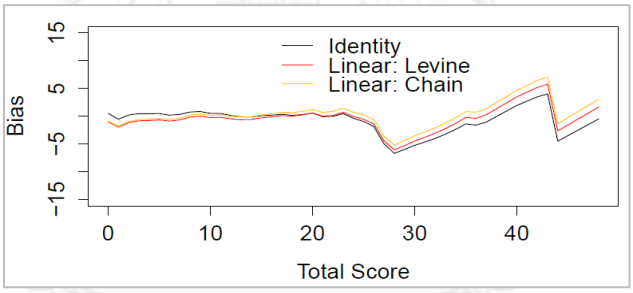
ภาพ 21 ค่า $Bias_i$ สำหรับวิธีการตรวจ 1 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี



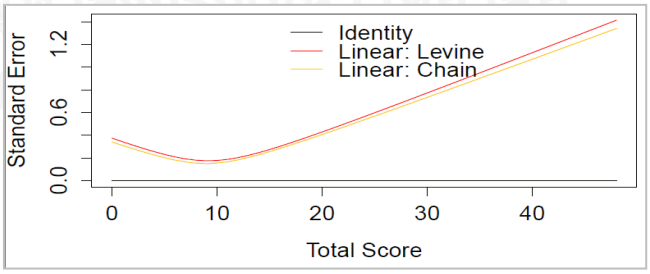
ภาพ 22 ค่า $\boldsymbol{SEE_i}$ สำหรับวิธีการตรวจ 1 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี



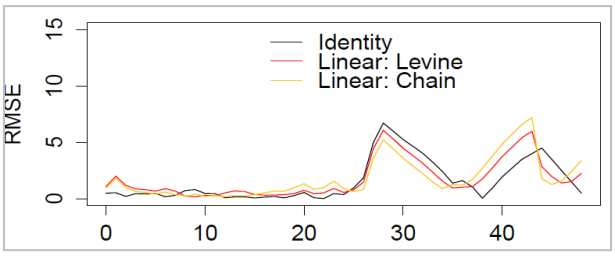
ภาพ 23 ค่า $\boldsymbol{RMSE_i}$ สำหรับวิธีการตรวจ 1 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี



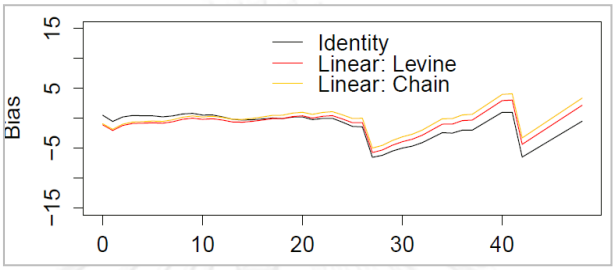
ภาพ 24 ค่า $\boldsymbol{Bias_i}$ สำหรับวิธีการตรวจ 2 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี



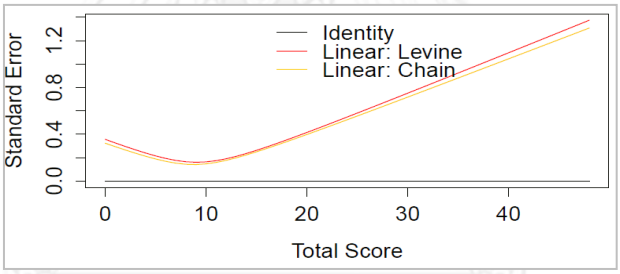
ภาพ 25 ค่า $\boldsymbol{SEE_i}$ สำหรับวิธีการตรวจ 2 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี



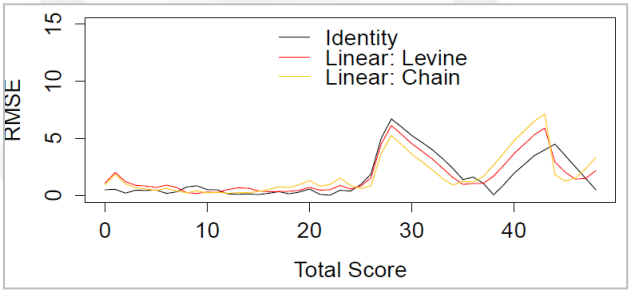
ภาพ 26 ค่า $\boldsymbol{RMSE_i}$ สำหรับวิธีการตรวจ 2 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี



ภาพ 27 ค่า $\boldsymbol{Bias_i}$ สำหรับวิธีการตรวจ 3 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี



ภาพ 28 ค่า $\boldsymbol{SEE_i}$ สำหรับวิธีการตรวจ 3 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี



ภาพ 29 ค่า $\boldsymbol{RMSE_i}$ สำหรับวิธีการตรวจ 3 ด้วยวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี

บทที่ 5
สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งมีจุดมุ่งหมาย เพื่อเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจที่แตกต่างกัน และเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ระหว่างวิธีของเลอวิน (Levine method) และวิธี CLE (Chained Linear Equating) สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ของนักเรียนระดับชั้นมัธยมศึกษาชั้นปีที่ 4

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 ของโรงเรียนในเขตพื้นที่การศึกษาสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 1 ภาคเรียนที่ 2 ปีการศึกษา 2556 จำนวน 1,609 คน โดยสุ่มกลุ่มตัวอย่างด้วยวิธีสุ่มแบบหลายขั้นตอน (multi - stage random sampling) โดยสุ่มโรงเรียนแบบแบ่งชั้น (stratified random sampling) โดยมีขนาดโรงเรียนเป็นตัวแปรสำหรับแบ่งชั้น คือ ขนาดใหญ่พิเศษ ขนาดใหญ่ ขนาดกลาง และขนาดเล็ก จากนั้นผู้วิจัยสุ่มโรงเรียนด้วยการสุ่มอย่างง่าย (simple random sampling) ในแต่ละขนาดของโรงเรียน ได้กลุ่มตัวอย่างสำหรับกลุ่มใหม่ จำนวน 1,084 คน กลุ่มอ้างอิง จำนวน 400 คน และกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์จำนวน 125 คน รวมเป็น 1,609 คน

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ ประกอบด้วย 1) แบบสอบวิชาฟิสิกส์รูปแบบผสม ระดับมัธยมศึกษาปีที่ 4 ภาคการศึกษาต้น ปีการศึกษา 2556 เรื่อง "งานและพลังงาน" ที่ผู้วิจัยสร้างขึ้น จำนวน 2 ฉบับ ได้แก่ แบบสอบฉบับใหม่ (new) และแบบสอบฉบับอ้างอิง (reference) โดยแต่ละฉบับมีข้อสอบจำนวน 24 ข้อ ประกอบด้วยข้อสอบแบบหลายตัวเลือก 16 ข้อ คะแนนเต็มข้อละ 1 คะแนน และข้อสอบแบบสร้างคำตอบ 8 ข้อ คะแนนเต็มข้อละ 4 คะแนน ดังนั้นคะแนนเต็มของแบบสอบฉบับนี้จึงเท่ากับ 48 คะแนน ซึ่งแบบสอบทั้งสองฉบับมีความเป็นคู่ขนานกันทั้งด้านเนื้อหาและความยากของแบบสอบ โดยมีข้อสอบร่วมกัน จำนวน 12 ข้อ ประกอบด้วยข้อสอบแบบหลายตัวเลือกจำนวน 8 ข้อ และข้อสอบแบบสร้างคำตอบจำนวน 4 ข้อ 2) เกณฑ์การให้คะแนน (rubric score) ที่ใช้ในตรวจข้อสอบแบบสร้างคำตอบในแบบสอบวิชาฟิสิกส์ระดับมัธยมศึกษาปีที่ 4 เรื่องงานและพลังงาน 3) ผู้ตรวจให้คะแนน 2 คน ได้แก่ ผู้ตรวจ 1 และผู้ตรวจ 2 โดยผู้ตรวจ 2 คือ ผู้วิจัย ส่วนผู้ตรวจ 1 คือ ครูผู้สอนวิชาฟิสิกส์ที่มีประสบการณ์ในโรงเรียนมาแล้วอย่างต่ำ 3 ปี

การเก็บรวบรวมข้อมูล มี 2 ช่วง คือ ช่วงที่ 1 นำแบบสอบฉบับใหม่ (new) ที่มีข้อสอบร่วมไปสอบกับผู้สอบกลุ่มใหม่ และผู้สอบกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์ หลังจากนั้นผู้ตรวจ 1 และ 2 (ผู้วิจัย) ทำการตรวจแบบสอบของกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์และกลุ่มใหม่ ตามวิธีการตรวจที่ได้กำหนดไว้ ส่วนช่วงที่ 2 นำแบบสอบฉบับอ้างอิง (reference) ที่มีข้อสอบร่วมไปสอบกับผู้สอบกลุ่มอ้างอิงและกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์ หลังจากนั้นผู้ตรวจ 1 และ 2 (ผู้วิจัย) ทำการตรวจแบบสอบของกลุ่มอ้างอิงที่ใช้เป็นเกณฑ์และกลุ่มอ้างอิง ตามวิธีการตรวจที่ได้กำหนดไว้

การวิเคราะห์ข้อมูล ใช้วิธีการปรับเทียบเชิงเส้นตรง 2 วิธี สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ขณะที่การออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ ใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed) โดยแบ่งการวิเคราะห์ออกเป็น 7 ตอน ตอนแรกเป็นการวิเคราะห์ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ตอนที่สองเป็นการวิเคราะห์ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ ตอนที่สามเป็นการวิเคราะห์ความสอดคล้องในการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบระหว่างผู้ตรวจ 1 และ 2 ตอนที่สี่เป็นการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) โดยใช้วิธีการปรับเทียบเชิงเส้นตรง ตอนที่ห้าเป็นการปรับเทียบคะแนนสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ โดยใช้วิธีการปรับเทียบ อีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed) ตอนที่หกเป็นการเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี คือ วิธีของเลอวิน (Levine method) และวิธี CLE (Chained linear equating) และตอนที่เจ็ดเป็นการเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของวิธีการตรวจ 3 วิธีสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

สรุปผลการวิจัย

1. ผลการวิเคราะห์ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

ผลการวิเคราะห์ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) สำหรับวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม ซึ่งผู้ตรวจ 2 ตรวจแบบสอบฉบับใหม่ (new) ทั้งหมด ขณะที่ผู้ตรวจ 1 ตรวจแบบสอบฉบับอ้างอิง (reference) ทั้งหมด โดยผู้สอบกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ได้คะแนนจากแบบสอบทั้งฉบับอยู่ระหว่าง 1-43 คะแนน ซึ่งมีค่าเฉลี่ย คือ 9.38 คะแนน และส่วนเบี่ยงเบนมาตรฐาน คือ 6.18 ขณะที่ได้คะแนนจากข้อสอบร่วมทั้งหมดอยู่ระหว่าง 1-22 คะแนน ซึ่งมีค่าเฉลี่ย 5.56 คะแนน และมีส่วนเบี่ยงเบนมาตรฐาน คือ 3.96 โดยคะแนนของแบบสอบใหม่ทั้งฉบับกับคะแนนของข้อสอบร่วมทั้งหมดจากแบบสอบดังกล่าวมีความสัมพันธ์กันในระดับสูง ($r = .93$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนผู้สอบกลุ่มอ้างอิงที่ทำแบบสอบฉบับอ้างอิง (reference) ได้คะแนนจากแบบสอบทั้งฉบับอยู่ระหว่าง 2-40 คะแนน ซึ่งมีค่าเฉลี่ย 12.83 คะแนน และมีส่วนเบี่ยงเบนมาตรฐาน คือ 7.51 ขณะที่คะแนนจากข้อสอบร่วมทั้งหมดอยู่ระหว่าง 1-22 คะแนน ซึ่งมีค่าเฉลี่ย 7.74 คะแนน และส่วนเบี่ยงเบนมาตรฐาน คือ 4.32 โดยที่คะแนนของแบบสอบฉบับอ้างอิงทั้งฉบับกับคะแนนของข้อสอบร่วมทั้งหมดจากแบบสอบดังกล่าวมีความสัมพันธ์กันในระดับสูง ($r = .91$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ผลการวิเคราะห์ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) สำหรับวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจ 2 ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม) โดยผู้สอบกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ได้คะแนนจากแบบสอบทั้งฉบับอยู่ระหว่าง 1-43 คะแนน ซึ่งมีค่าเฉลี่ย คือ 9.38 คะแนน และส่วนเบี่ยงเบนมาตรฐาน คือ 6.18 ขณะที่ได้คะแนนจากข้อสอบร่วมทั้งหมดอยู่ระหว่าง 1-22 คะแนน ซึ่งมีค่าเฉลี่ย 5.6 คะแนน และมีส่วนเบี่ยงเบนมาตรฐาน คือ 3.96 โดยที่คะแนนของแบบสอบฉบับใหม่ทั้งฉบับกับคะแนนของข้อสอบร่วมทั้งหมดจากแบบสอบดังกล่าวมีความสัมพันธ์กันในระดับสูง (r = .93) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนผู้สอบกลุ่มอ้างอิงที่ทำแบบสอบฉบับอ้างอิง (reference) ได้คะแนนจากแบบสอบทั้งฉบับอยู่ระหว่าง 2-39 คะแนน ซึ่งมีค่าเฉลี่ย 12.98 คะแนน และมีส่วนเบี่ยงเบนมาตรฐาน คือ 7.57 ขณะที่คะแนนจากข้อสอบร่วมทั้งหมดอยู่ระหว่าง 1-22 คะแนน ซึ่งมีค่าเฉลี่ย 7.89 คะแนน และส่วนเบี่ยงเบนมาตรฐาน คือ 4.39 โดยที่คะแนนของแบบสอบฉบับอ้างอิงทั้งฉบับกับคะแนนของข้อสอบร่วมทั้งหมดจากแบบสอบดังกล่าวมีความสัมพันธ์กันในระดับสูง (r = .91) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ผลการวิเคราะห์ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) สำหรับวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม (ผู้ตรวจทั้ง 2 คน ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม โดยแบ่งกันตรวจคนละครึ่งหนึ่งของข้อสอบร่วม) โดยผู้สอบกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ได้คะแนนจากแบบสอบทั้งฉบับอยู่ระหว่าง 1-47 คะแนน ซึ่งมีค่าเฉลี่ย คือ 9.38 คะแนน และส่วนเบี่ยงเบนมาตรฐาน คือ 6.22 ขณะที่ได้คะแนนจากข้อสอบร่วมทั้งหมดอยู่ระหว่าง 1-24 คะแนน ซึ่งมีค่าเฉลี่ย 5.57 คะแนน และมีส่วนเบี่ยงเบนมาตรฐาน คือ 3.99 โดยที่คะแนนของแบบสอบใหม่ทั้งฉบับกับคะแนนของข้อสอบร่วมทั้งหมดจากแบบสอบดังกล่าวมีความสัมพันธ์กันในระดับสูง (r = .94) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนผู้สอบกลุ่มอ้างอิงที่ทำแบบสอบฉบับอ้างอิง (reference) ได้คะแนนจากแบบสอบทั้งฉบับอยู่ระหว่าง 2 - 40 คะแนน ซึ่งมีค่าเฉลี่ย 12.96 คะแนน และมีส่วนเบี่ยงเบนมาตรฐาน คือ 7.55 ขณะที่คะแนนจากข้อสอบร่วมทั้งหมดอยู่ระหว่าง 1-22 คะแนน ซึ่งมีค่าเฉลี่ย 7.86 คะแนน และส่วนเบี่ยงเบนมาตรฐาน คือ 4.36 โดยที่คะแนนของแบบสอบฉบับอ้างอิงทั้งฉบับกับคะแนนของข้อสอบร่วมทั้งหมดจากแบบสอบดังกล่าวมีความสัมพันธ์กันในระดับสูง (r = .91) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2. ผลการวิเคราะห์ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์

ผลการวิเคราะห์ค่าสถิติที่เกี่ยวข้องกับแบบสอบรูปแบบผสมสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ (SG design) โดยผู้สอบกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ได้คะแนนจากแบบสอบทั้งฉบับอยู่ระหว่าง 3-38 คะแนน ซึ่งมีค่าเฉลี่ย คือ 13.08 คะแนน และส่วนเบี่ยงเบนมาตรฐาน คือ 7.27 ขณะที่ได้คะแนนจากข้อสอบร่วมทั้งหมดอยู่ระหว่าง 1-23 คะแนน ซึ่งมีค่าเฉลี่ย 7.68 คะแนน และมีส่วนเบี่ยงเบนมาตรฐาน คือ 4.27 โดยที่คะแนนของแบบสอบฉบับ

ใหม่ทั้งฉบับกับคะแนนของข้อสอบทั้งหมดจากแบบสอบดังกล่าวมีความสัมพันธ์กันในระดับสูง (r = .91) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนผู้สอบกลุ่มอ้างอิงที่ทำแบบสอบฉบับอ้างอิง (reference) ได้คะแนนจากแบบสอบทั้งฉบับอยู่ระหว่าง 3-36 คะแนน ซึ่งมีค่าเฉลี่ย 15.80 คะแนน และมีส่วนเบี่ยงเบนมาตรฐาน คือ 8.03 ขณะที่คะแนนจากข้อสอบร่วมทั้งหมดอยู่ระหว่าง 2-20 คะแนน ซึ่งมีค่าเฉลี่ย 9.12 คะแนน และส่วนเบี่ยงเบนมาตรฐาน คือ 4.56 โดยที่คะแนนของแบบสอบฉบับอ้างอิงทั้งฉบับกับคะแนนของข้อสอบทั้งหมดจากแบบสอบดังกล่าวมีความสัมพันธ์กันในระดับสูง (r = .89) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

3. ผลการวิเคราะห์ความสอดคล้องในการให้คะแนนข้อสอบร่วมแบบสร้างคำตอบระหว่างผู้ตรวจ 1 และ 2

คะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบในแบบสอบฉบับอ้างอิง (reference) ในวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่มของผู้ตรวจ 1 และคะแนนจากการตรวจให้คะแนนข้อสอบแบบสร้างคำตอบในแบบสอบฉบับอ้างอิง (reference) ในวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจ 2 ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม) มีความสัมพันธ์กันทางบวก (r = .99) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (p = .00) ดังนั้นจึงกล่าวได้ว่า การให้คะแนนข้อสอบร่วมแบบสร้างคำตอบระหว่างผู้ตรวจ 1 และ 2 มีความสอดคล้องกัน นั่นคือ ผู้ตรวจทั้งสองคนมีความเที่ยงระหว่างผู้ประเมิน

4. ผลการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) โดยใช้วิธีการปรับเทียบเชิงเส้นตรง

4.1 การปรับเทียบคะแนนโดยใช้วิธี CLE (Chained Linear Equating)

ในการปรับเทียบคะแนนสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจทั้ง 3 วิธี ได้คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกันกับผู้สอบกลุ่มอ้างอิงเท่าเดิมในช่วง 9 คะแนน ถึง 16 คะแนน และคะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกันกับผู้สอบกลุ่มอ้างอิง มีค่าต่างกันมากขึ้นในคะแนนส่วนท้าย อย่างเช่น ในวิธีการตรวจทั้ง 3 วิธี ได้คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกันกับผู้สอบกลุ่มอ้างอิงในช่วง 45 คะแนน ถึง 48 คะแนน มีค่าเพิ่มขึ้น 4 คะแนน ขณะที่คะแนนส่วนต้นได้คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกันกับผู้สอบกลุ่มอ้างอิงต่างกัน 1 คะแนน เป็นต้น

4.2 การปรับเทียบคะแนนโดยใช้วิธีของเลอวิน (Levine method)

ในการปรับเทียบคะแนนสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ระหว่างวิธีการตรวจทั้ง 3 วิธี ได้คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกันกับผู้สอบกลุ่มอ้างอิงเท่าเดิม ในช่วง

14 คะแนน ถึง 22 คะแนน และคะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกันกับผู้สอบกลุ่มอ้างอิง มีค่าต่างกันมากขึ้นในคะแนนส่วนท้าย อย่างเช่น วิธีการตรวจวิธีที่ 1: Non Trend และ 3: Half Trend ได้คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกันกับผู้สอบกลุ่มอ้างอิงในช่วง 46 คะแนน ถึง 48 คะแนน มีค่าเพิ่มขึ้น 3 คะแนน เป็นต้น ขณะที่คะแนนส่วนต้น ได้คะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกันกับผู้สอบกลุ่มอ้างอิงต่างกัน 2 คะแนน

5. ผลการปรับเทียบคะแนนสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ โดยใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed)

ในการปรับเทียบคะแนนสำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ โดยใช้วิธีการปรับเทียบอีควิเปอร์เซ็นไทล์แบบลูกโซ่ ที่ปรับการแจกแจงคะแนนด้วยล็อกเชิงเส้น (chained equipercentile equating with loglinear presmoothed) คะแนนของของผู้สอบกลุ่มใหม่จากแบบสอบฉบับใหม่ (new) ที่นำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิงจากแบบสอบฉบับอ้างอิง (reference) มีค่าเพิ่มขึ้น 1 คะแนน ในช่วง 1 คะแนน ถึง 23 คะแนน และช่วง 40 คะแนน ถึง 45 คะแนน ขณะที่มีค่าเท่าเดิมที่ 0 คะแนน และในช่วง 46 คะแนน ถึง 48 คะแนน อีกทั้งยังมีค่าเพิ่มขึ้น 2 คะแนน ในช่วง 24 คะแนน ถึง 39 คะแนน เช่น ผู้สอบกลุ่มใหม่ที่ทำแบบสอบฉบับใหม่ (new) ที่มีคะแนน 32 คะแนน ถ้านำมาปรับเทียบกับคะแนนของผู้สอบกลุ่มอ้างอิงได้ 34 คะแนน เป็นต้น ซึ่งคะแนนที่ได้จากปรับเทียบนี้นำไปใช้เป็นเกณฑ์ในการหาค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD)

6. ผลการเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีการปรับเทียบเชิงเส้นตรง 2 วิธี คือ วิธีของเลอวิน (Levine method) และวิธี CLE (Chained linear equating)

ผลการวิเคราะห์ พบว่า ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธี CLE (Chain Linear Equating) มีค่าต่ำกว่าวิธีของเลอวิน (Levine method) ในทุกวิธีการตรวจของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) โดยค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธี CLE (Chain Linear Equating) มีค่ามากที่สุดอยู่ที่วิธีการตรวจวิธีที่ 2: Full Trend และ 3: Half Trend คือ 1.481 และมีค่าต่ำที่สุดอยู่ที่วิธีการตรวจวิธีที่ 1: Non Trend คือ 1.421 ขณะที่ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีของเลอวิน (Levine method) มีค่ามากที่สุดอยู่ที่วิธีการตรวจวิธีที่ 2: Full Trend รองลงมา คือ วิธีการตรวจวิธีที่ 3: Half Trend และ 1: Non Trend ตามลำดับ โดยมีค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) เท่ากับ 1.804 1.793 และ 1.717 ตามลำดับ ดังนั้นวิธีการปรับเทียบคะแนนที่เหมาะสมการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) คือ วิธี CLE (Chain Linear Equating)

7. ผลการเปรียบเทียบสถิติบูทสแตรป (bootstrapped statistics) ของวิธีการตรวจ 3 วิธีสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT)

การปรับเทียบคะแนนด้วยวิธี CLE (Chained Linear Equating) สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) พบว่า วิธีการตรวจที่มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำที่สุด คือ วิธีการตรวจวิธีที่ 3: Half Trend, 2: Full Trend และ 1: Non Trend ตามลำดับ ส่วนวิธีการตรวจที่มีค่าความลำเอียง (bias) ต่ำที่สุด คือ วิธีการตรวจวิธีที่ 3: Half Trend, 2: Full Trend และ1: Non Trend ตามลำดับ ส่วนรูวิธีการตรวจที่มีค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) ต่ำที่สุด ได้แก่ วิธีการตรวจวิธีที่ 3: Half Trend, 1: Non Trend และ 2: Full Trend ตามลำดับ กล่าวคือ วิธีการตรวจวิธีที่ 3: Half Trend มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และค่าความลำเอียง (bias) ต่ำที่สุด วิธีการตรวจวิธีที่ 2: Full Trend มีค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) มากที่สุด ส่วนวิธีการตรวจวิธีที่ 1: Non Trend มีค่าความลำเอียง (bias) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) มากที่สุด

การปรับเทียบคะแนนด้วยวิธีของเลอวิน (Levine method) สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) พบว่า วิธีการตรวจที่มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำที่สุด คือ วิธีการตรวจที่มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำที่สุด คือ วิธีการตรวจวิธีที่ 1: Non Trend, 2: Full Trend และ 3: Half Trend ตามลำดับ ส่วนรูปแบบที่มีค่าความลำเอียง (bias) ต่ำที่สุด คือ วิธีการตรวจวิธีที่ 2: Full Trend, 1: Non Trend และ 3: Half Trend ตามลำดับ ส่วนรูปแบบที่มีค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) ต่ำที่สุด ได้แก่ วิธีการตรวจวิธีที่ 3: Half Trend, 1: Non Trend และ 2: Full Trend ตามลำดับ กล่าวคือ วิธีการตรวจวิธีที่ 2: Full Trend มีค่าความลำเอียง (bias) ต่ำที่สุด แต่มีค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) มากที่สุด ส่วนวิธีการตรวจวิธีที่ 3: Half Trend มีค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) ต่ำที่สุด แต่มีค่าความลำเอียง (bias) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) มากที่สุด วิธีการตรวจวิธีที่ 1: Non Trend มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำที่สุด แต่ถ้าเปรียบเทียบค่าสถิติบูทสแตรปทั้ง 3 ค่า ที่ได้จากวิธีการปรับเทียบเชิงเส้นตรงทั้ง 2 วิธี พบว่า ค่าสถิติบูทสแตรปทั้ง 3 ค่า ที่ได้จากวิธีของเลอวิน (Levine method) ในแต่ละวิธีการตรวจมีค่ามากกว่าวิธี CLE (Chained Linear Equating)

เมื่อพิจารณาค่าความลำเอียง ค่าความคลาดเคลื่อนในการปรับเทียบ และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสองในแต่ละจุดคะแนนที่ได้จากวิธีการปรับเทียบเชิงเส้นตรงทั้ง 2 วิธี ($Bias_i$, SEE_i และ $RMSE_i$ ตามลำดับ) สำหรับวิธีการตรวจวิธีที่ 1: Non Trend 2: Full Trend และ 3: Half Trend พบว่า ค่า $Bias_i$ และค่า $RMSE_i$ มีค่าค่อนข้างสม่ำเสมอ และมีค่าเข้าใกล้ 0 ในช่วง 0 คะแนน ถึง 20 คะแนน และมีค่าไม่สม่ำเสมอในช่วงคะแนน

ปลายๆ ขณะที่ค่า SEE_i มีค่าเข้าใกล้ 0 ในช่วง 10 คะแนน และค่อย ๆ มีค่าเพิ่มขึ้นอย่างสม่ำเสมอ เนื่องจากกราฟที่ได้เป็นเส้นตรง โดยลักษณะกราฟของค่า $Bias_i$ ค่า SEE_i และค่า $RMSE_i$ ที่ได้ในแต่ละวิธีการตรวจจากวิธีการปรับเทียบวิธีเดียวกันไม่แตกต่างกันมากนัก

อภิปรายผล

การอภิปรายผลในงานวิจัยนี้นำเสนอ 4 ประเด็น คือ การอภิปรายผลเกี่ยวกับค่าความเที่ยงในข้อสอบแบบหลายตัวเลือก การอภิปรายเกี่ยวกับการจัดเรียงข้อสอบร่วมในแบบสอบ การอภิปรายผลที่ได้จากการวิเคราะห์คะแนนที่ได้จากแบบสอบแต่ละฉบับในแต่ละวิธีการตรวจ และการอภิปรายตามสมมติฐานการวิจัย มีรายละเอียดดังนี้

1. การอภิปรายผลเกี่ยวกับค่าความเที่ยงในข้อสอบแบบหลายตัวเลือก

เมื่อพิจาณาผลการวิเคราะห์ข้อสอบแบบหลายตัวเลือกและข้อสอบแบบสร้างคำตอบ พบว่า ค่าความเที่ยงในข้อสอบแบบหลายตัวเลือกในแบบสอบฉบับใหม่และฉบับอ้างอิง (.634 และ .602 ตามลำดับ) มีค่าต่ำกว่าค่าความเที่ยงในข้อสอบแบบสร้างคำตอบในแบบสอบฉบับใหม่และฉบับอ้างอิง (.824 และ .872 ตามลำดับ) ทั้งที่จำนวนข้อสอบแบบหลายตัวเลือกมีจำนวนมากกว่าจำนวนข้อสอบแบบสร้างคำตอบ โดยข้อสอบแบบหลายตัวเลือกในแบบสอบแต่ละฉบับมีจำนวน 16 ข้อ ขณะที่ข้อสอบแบบสร้างคำตอบในแบบสอบแต่ละฉบับมีจำนวน 8 ข้อ ซึ่งสาเหตุที่ทำให้ค่าความเที่ยงในข้อสอบแบบหลายตัวเลือกมีค่าต่ำกว่าค่าความเที่ยงในข้อสอบแบบสร้างคำตอบ เป็นเพราะค่าอำนาจจำแนกเฉลี่ยในข้อสอบแบบหลายตัวเลือกในแบบสอบฉบับใหม่และฉบับอ้างอิง (.400 และ .384 ตามลำดับ) มีค่าต่ำกว่าค่าอำนาจจำแนกเฉลี่ยในข้อสอบแบบสร้างคำตอบในแบบสอบฉบับใหม่และฉบับอ้างอิง (.454 และ .513 ตามลำดับ) ดังนั้นถ้าต้องการนำแบบสอบเรื่องงานและพลังงานทั้งสองฉบับไปใช้ในการทดสอบครั้งต่อไป ควรปรับปรุงข้อสอบแบบหลายตัวเลือกให้มีความสามารถในการจำแนกผู้สอบมากขึ้น โดยช่วงของค่าอำนาจจำแนกของข้อสอบขึ้นอยู่กับค่าความยากของข้อสอบ (ศิริชัย กาญจนวาสี, 2552) และเมื่อพิจารณาข้อสอบแบบสร้างคำตอบ พบว่า ความยากของข้อสอบแบบสร้างคำตอบในแบบสอบฉบับใหม่และฉบับอ้างอิงอยู่ในระดับค่อนข้างยาก (.258 และ .268 ตามลำดับ) ขณะที่ความยากของข้อสอบแบบหลายตัวเลือกในแบบสอบฉบับใหม่และฉบับอ้างอิงอยู่ในระดับปานกลาง (.400 และ .384 ตามลำดับ) ดังนั้นในการปรับปรุงครั้งต่อไปอาจสร้างข้อสอบแบบหลายตัวเลือกให้มีความยากเพิ่มขึ้น เพื่อเพิ่มความสามารถในการจำแนกผู้สอบของข้อสอบแบบหลายตัวเลือก ซึ่งส่งผลให้ค่าความเที่ยงของข้อสอบแบบหลายตัวเลือกเพิ่มขึ้น

2. การอภิปรายเกี่ยวกับการจัดเรียงข้อสอบร่วมในแบบสอบ

แบบสอบเรื่องงานและพลังงานฉบับใหม่และฉบับอ้างอิง ประกอบด้วยข้อสอบเฉพาะจำนวน 12 ข้อ แบ่งเป็น ข้อสอบแบบหลายตัวเลือกจำนวน 8 ข้อ และข้อสอบแบบสร้างคำตอบจำนวน 4 ข้อ และข้อสอบร่วมจำนวน 12 ข้อ แบ่งเป็น ข้อสอบแบบหลายตัวเลือกจำนวน 8 ข้อ และข้อสอบแบบสร้างคำตอบจำนวน 4 ข้อ ในแบบสอบฉบับใหม่มีการเรียงข้อสอบโดยเริ่มจากข้อสอบเฉพาะแบบ

หลายตัวเลือก ข้อสอบร่วมแบบหลายตัวเลือก ข้อสอบเฉพาะแบบสร้างคำตอบ และข้อสอบร่วมแบบสร้างคำตอบ ตามลำดับ กล่าวคือ ข้อสอบร่วมเรียงไว้หลังจากข้อสอบเฉพาะ ขณะที่แบบสอบฉบับอ้างอิงมีการเรียงข้อสอบโดยเริ่มจากข้อสอบร่วมแบบหลายตัวเลือก ข้อสอบเฉพาะแบบหลายตัวเลือก ข้อสอบร่วมแบบสร้างคำตอบ และข้อสอบเฉพาะแบบสร้างคำตอบ ตามลำดับ กล่าวคือ ข้อสอบร่วมเรียงไว้ก่อนหน้าข้อสอบเฉพาะ ในกรณีที่ข้อสอบร่วมเรียงไว้หลังจากข้อสอบเฉพาะ ถ้าผู้สอบเกิดความเมื่อยล้าในการทำข้อสอบอาจทำให้ผู้สอบไม่ใส่ใจในการข้อสอบข้อหลัง ๆ ซึ่งมีแต่เฉพาะเป็นข้อสอบร่วม อาจทำให้ผลการปรับเทียบคะแนนเกิดความคลาดเคลื่อนขึ้นได้ เพราะคะแนนจากข้อสอบร่วมต้องใช้เป็นคะแนนเชื่อมต่อจากแบบฉบับใหม่ไปยังแบบสอบฉบับอ้างอิง ในกรณีที่ข้อสอบร่วมเรียงไว้ก่อนหน้าข้อสอบเฉพาะ ถ้าผู้สอบเกิดความเมื่อยล้าในการทำข้อสอบอาจทำให้ผู้สอบไม่ใส่ใจในการข้อสอบข้อหลัง ๆ ซึ่งมีแต่เฉพาะเป็นข้อสอบเฉพาะ ดังนั้นนักเรียนอาจได้คะแนนข้อสอบเฉพาะน้อยลงเนื่องจากความเมื่อยล้าแต่ไม่ได้มาจากความสามารถของผู้สอบ และเมื่อพิจารณาคะแนนเฉลี่ยของข้อสอบร่วมในแบบสอบฉบับใหม่และฉบับอ้างอิงในวิธีการตรวจทั้ง 3 วิธี สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) พบว่า คะแนนเฉลี่ยของข้อสอบร่วมในแบบสอบฉบับอ้างอิงมีค่ามากกว่าคะแนนเฉลี่ยของข้อสอบร่วมในแบบสอบฉบับอ้างอิงในทุกวิธีการตรวจ เช่นเดียวกับคะแนนเฉลี่ยของข้อสอบร่วมในแบบสอบฉบับใหม่และฉบับอ้างอิง สำหรับการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์ (SG design) ที่พบว่า คะแนนเฉลี่ยของข้อสอบร่วมในแบบสอบฉบับอ้างอิงมีค่ามากกว่าคะแนนเฉลี่ยของข้อสอบร่วมในแบบสอบฉบับอ้างอิง ดังนั้นในการสร้างแบบสอบทั้งสองฉบับในครั้งต่อไปควรเรียงข้อสอบร่วมอย่างสุ่ม โดยเรียงตามเนื้อหา ระดับความยาก หลังจากนั้นจึงเรียงข้อสอบแบบสุ่ม เพื่อลดความคลาดเคลื่อนในการปรับเทียบ และให้คะแนนจากข้อสอบเฉพาะมาจากความสามารถของผู้สอบไม่ได้เกิดจากความเมื่อยล้าของผู้สอบ

3. การอภิปรายผลที่ได้จากการวิเคราะห์คะแนนที่ได้จากแบบสอบ

เมื่อเปรียบเทียบคะแนนของผู้สอบของผู้สอบกลุ่มใหม่และกลุ่มอ้างอิง ภายใต้การออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) พบว่า คะแนนของผู้สอบทั้งสองกลุ่มแตกต่างกันทั้งที่แบบสอบที่ใช้ได้ผ่านการตรวจสอบความเป็นคู่ขนานแล้ว โดยคะแนนของผู้สอบกลุ่มใหม่มากกว่ากลุ่มอ้างอิงทั้งจากแบบสอบทั้งฉบับและเฉพาะข้อสอบร่วม สาเหตุอาจเป็นเพราะในช่วงแรกของการเก็บข้อมูลนั้นเกิดเหตุการณ์ทางการเมืองขึ้นในช่วงเดือนธันวาคม พ.ศ. 2555 ถึง เดือนกุมภาพันธ์ 2556 ทำให้ผู้สอบกลุ่มใหม่บางส่วนที่มาจากโรงเรียนที่อยู่ใกล้พื้นที่ชุมนุม จำเป็นต้องหยุดการเรียนการสอน ทำให้ความต่อเนื่องในการเรียนลดลง อีกทั้งครูผู้สอนจำเป็นต้องสอนเนื้อหาจำนวนมากในเวลาจำกัด จึงไม่ได้ให้นักเรียนฝึกทำโจทย์ในห้องเรียนมากนัก ซึ่งในการจัดการสอนในรายวิชาฟิสิกส์นั้น มีความจำเป็นที่ต้องให้นักเรียนได้ฝึกทำโจทย์ต่าง ๆ ในขณะที่ผู้สอบกลุ่มอ้างอิงที่ทำแบบทดสอบในช่วงที่ 2 นั้น อยู่เขตฝั่งธนบุรีทั้งหมด จึงไม่ได้รับผลกระทบจากเหตุการณ์ดังกล่าว และได้ดำเนินการเรียนการสอนตามปกติ จึงมีความเป็นไปได้ที่คะแนนของผู้สอบกลุ่มใหม่ได้ต่ำกว่าคะแนนของผู้สอบกลุ่มอ้างอิง ดังเห็นได้ในตาราง 25 หน้า 103

ขณะที่ผู้สอบในการออกแบบการปรับเทียบที่ใช้เป็นเกณฑ์นั้น ได้คะแนนเฉลี่ยจากแบบสอบฉบับอ้างอิง (reference) มากกว่าคะแนนเฉลี่ยจากแบบสอบฉบับใหม่ (new) แต่ความแตกต่างของคะแนนดังกล่าวยังต่ำกว่าในการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) สาเหตุของความแตกต่างของดังกล่าว อาจเกิดจากช่วงเวลาของการจัดการสอบ โดยการสอบในครั้งแรกเกิดขึ้นหลังจากที่ผู้สอบเรียนเรื่องงานและพลังงานจบไปแล้วประมาณ 1 สัปดาห์ และใช้คะแนนสอบครั้งแรกเป็นคะแนนเก็บสำหรับผู้สอบ ขณะที่การสอบครั้งที่สองเกิดขึ้นหลังจากที่ผู้สอบสอบกลางภาคไป ประมาณ 1 สัปดาห์ ซึ่งอาจทำให้ผลจากการเตรียมตัวสอบกลางภาคของนักเรียนส่งผลต่อคะแนนของการสอบครั้งที่สองมากกว่าการสอบครั้งแรกที่อาจมีเวลาในการเตรียมตัวน้อยกว่า แต่อย่างไรตามความแตกต่างดังกล่าวยังต่ำกว่าในรูปแบบ NEAT นอกจากนี้ยังพบว่า คะแนนเฉลี่ยของผู้สอบในรูปแบบที่ใช้เป็นเกณฑ์จากแบบสอบทั้งฉบับมีค่าสูงกว่าคะแนนเฉลี่ยของผู้สอบในรูปแบบ NEAT ซึ่งอาจเกิดจากการที่ผู้สอบในรูปแบบดังกล่าวมาจากโรงเรียนเดียวกันที่มีการเตรียมตัวในการสอบทั้ง 2 ครั้ง และยังเป็นโรงเรียนที่ไม่ได้รับผลกระทบจากเหตุการณ์ทางการเมือง ขณะที่ในกลุ่ม NEAT มาจากหลากหลายโรงเรียน ซึ่งคะแนนในการจัดสอบครั้งนี้อาจมีผลหรือไม่มีผลกับคะแนนเก็บผู้สอบบางคน จึงมีผู้สอบบางคนเตรียมตัวมาอย่างดีในการสอบ ในขณะเดียวกันอาจมีผู้สอบบางส่วนที่ไม่ได้เตรียมตัวในการสอบ ประกอบกับกลุ่มผู้สอบจากบางโรงเรียนในรูปแบบ NEAT ได้รับผลกระทบจากการเหตุการณ์ทางการเมือง ทำให้ความพร้อมในการสอบลดลง เมื่อเปรียบเทียบกับโรงเรียนที่ใช้เป็นเกณฑ์ ซึ่งไม่ได้รับผลกระทบจากเหตุการณ์ทางการเมืองที่เกิดขึ้น

เมื่อพิจารณาคะแนนของผู้สอบกลุ่มใหม่จากวิธีการตรวจทั้ง 3 วิธี ของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) จากแบบสอบทั้งฉบับ พบว่า คะแนนดังกล่าวมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานที่ใกล้เคียงกันมาก เช่นเดียวกับคะแนนของผู้สอบกลุ่มอ้างอิงจากวิธีการตรวจทั้ง 3 วิธีที่มีค่าเฉลี่ย และส่วนเบี่ยงเบนใกล้เคียงกันมาก และเมื่อพิจารณาคะแนนข้อสอบร่วมทั้งหมดของผู้สอบกลุ่มใหม่จากวิธีการตรวจทั้ง 3 วิธี ของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) พบว่า คะแนนดังกล่าวมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานที่ใกล้เคียงกัน เช่นเดียวกับคะแนนข้อสอบร่วมทั้งหมดของผู้สอบกลุ่มอ้างอิงจากวิธีการตรวจทั้ง 3 วิธีที่มีค่าเฉลี่ยและส่วนเบี่ยงเบนใกล้เคียงกัน โดยค่าสถิติของผู้สอบกลุ่มใหม่ในวิธีการตรวจวิธีที่ 1: Non Trend และ 2: Full Trend มีค่าเท่ากัน เนื่องจากวิธีการตรวจทั้ง 2 วิธีนั้น ผู้ตรวจ 2 ทำหน้าที่ตรวจแบบสอบฉบับใหม่ (new) ทั้งหมดเพียงคนเดียว

แต่อย่างไรก็ตาม เมื่อพิจารณาโดยภาพรวม ค่าเฉลี่ย ค่ามัธยฐาน และฐานนิยมของคะแนนจากแบบสอบทั้ง 2 ฉบับในงานวิจัยนี้ มีค่าต่ำกว่าครึ่งหนึ่งของคะแนนเต็ม แสดงให้เห็นว่าแบบสอบทั้ง 2 ฉบับมีระดับความยากค่อนมาทางยาก เมื่อพิจารณาค่าความเบ้และความโด่งของแบบสอบทั้ง 2 ฉบับ พบว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนต่ำ และเนื่องจากผู้วิจัยมีโอกาสได้คุมสอบและตรวจข้อสอบส่วนใหญ่ของงานวิจัยนี้ จึงได้เห็นพฤติกรรมการตอบข้อสอบอัตนัยของนักเรียนที่ค่อนข้างคล้ายกัน คือ นักเรียนส่วนใหญ่มักเลือกทำข้อสอบปรนัยก่อน ขณะที่ข้อสอบอัตนัยนักเรียนเลือกทำในภายหลัง หรือเลือกที่ไม่ทำข้อสอบอัตนัย ซึ่งมีนักเรียนจำนวนไม่น้อยเลือกทำเพียง

เฉพาะข้อสอบปรนัยเท่านั้น และไม่มีการเขียนสมการหรือข้อความใด ๆ ในข้อสอบแบบอัตนัย ทำให้คะแนนของนักเรียนในภาพรวมได้คะแนนต่ำกว่าครึ่งหนึ่งของแบบสอบ และเมื่อพิจารณาถึงพฤติกรรมของการเขียนแสดงวิธีทำของผู้สอบพบว่า ผู้สอบบางส่วนมักไม่เขียนสูตรหรือสมการที่ใช้ในการแก้โจทย์ปัญหา โดยเริ่มต้นที่การแทนค่า ดังนั้นเมื่อผู้สอบแทนค่าผิดและหาคำตอบผิด ทำให้ผู้สอบไม่ได้คะแนนสำหรับโจทย์ข้อนั้น นอกจากนี้ยังพบว่า นักเรียนส่วนใหญ่ที่ไม่เขียนสูตร/สมการที่ใช้ มักแทนค่าผิดและหาคำตอบผิด ดังนั้นในการวิจัยครั้งต่อไป จึงควรทดลองใช้กับวิชาอื่นๆที่เป็นการเขียนบรรยาย เช่น วิชาภาษาไทย วิชาภาษาอังกฤษ หรือวิชาชีววิทยา เป็นต้น

4. การอภิปรายตามสมมติฐานการวิจัย

4.1 การอภิปรายตามสมมติฐานการวิจัยข้อที่ 1 การออกแบบการปรับเทียบคะแนน (equating designs) ที่มีการตรวจให้คะแนนแตกต่างกัน

จากการวิจัยในครั้งนี้ พบว่า วิธีปรับเทียบคะแนนด้วยวิธี CLE (Chained Linear Equating) ให้ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ต่ำกว่าวิธีของเลอวิน (Levine method) ด้วยเหตุนี้ผู้วิจัยอภิปรายการออกแบบการปรับเทียบด้วยวิธี CLE (Chained Linear Equating) เป็นหลัก จากการเปรียบเทียบสถิติบูทสแตรปของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) พบว่าสอดคล้องกับสมมติฐานข้อที่ 1 นั่นคือ วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มทั้ง 2 วิธีมีค่าสถิติบูทสแตรปส่วนใหญ่ต่ำกว่าวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบโดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม ถ้าพิจารณาค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ซึ่งเป็นค่าความคลาดเคลื่อนที่มาจากค่าความลำเอียง (bias) และค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) พบว่า วิธีการตรวจวิธีที่ 3: Half Trend มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำที่สุด ซึ่งเป็นผลมาจากค่าความลำเอียง (bias) และค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) มีค่าต่ำ ขณะที่วิธีการตรวจวิธีที่ 2: Full Trend มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำเป็นลำดับที่ 2 ซึ่งเป็นผลจากค่าความลำเอียง (bias) มีค่าต่ำเป็นลำดับที่ 2 ส่วนวิธีการตรวจวิธีที่ 1: Non Trend มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) มากที่สุด เนื่องจากมีค่าความลำเอียง (bias) มีค่ามากที่สุด แต่อย่างไรก็ตามค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) และค่าความลำเอียง (bias) ที่ได้จากวิธีการตรวจวิธีที่ 2: Full Trend และ 3: Half Trend มีค่าใกล้เคียงกัน จึงได้กล่าวว่า ในการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ด้วยวิธี CLE (Chained Linear Equating) ที่ใช้วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่ม ทำให้ประสิทธิภาพในการปรับเทียบคะแนนมากกว่าวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม ซึ่งสอดคล้องกับงานวิจัยของ Kim และคณะ (2010) ที่ได้ศึกษาเปรียบเทียบสถิติบูทสแตรปของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ด้วยวิธี CLE (Chained Linear Equating) ซึ่งพบว่า การปรับเทียบคะแนนที่ใช้วิธีตรวจ

ข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่ม มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำที่สุด ขณะที่การปรับเทียบคะแนนที่ใช้วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยไม่มีการตรวจให้คะแนนข้ามกลุ่ม ให้ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) และค่าความลำเอียง (bias) มากที่สุด อีกทั้งยังสอดคล้องกับงานวิจัยของ Kim และคณะ (2010) ได้ศึกษาสถิติบูทสแตรปของการออกแบบการปรับเทียบคะแนนแบบสอบแบบสร้างคำตอบโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) พบว่า การปรับเทียบคะแนนที่ใช้วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่ม มีระดับความคลาดเคลื่อนของการปรับเทียบคะแนนระหว่างแบบสอบต่ำ แสดงให้เห็นว่าความคลาดเคลื่อนในการปรับเทียบคะแนนและความลำเอียงในการปรับเทียบคะแนนสามารถควบคุมได้ด้วยการตรวจให้คะแนนแบบข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ (trend construct-response anchor) ซึ่งทำให้ค่าความลำเอียง (bias) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำกว่าข้อสอบร่วมแบบสร้างคำตอบที่ไม่มีการตรวจให้คะแนนแบบข้ามกลุ่ม (no-trend construct-response anchor) และเมื่อเปรียบเทียบระหว่างวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่ม ทั้ง 2 วิธี พบว่า วิธีการตรวจวิธีที่ 3: Half Trend มีค่าความลำเอียง (bias) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำกว่าวิธีการตรวจวิธีที่ 2: Full Trend อาจเป็นเพราะการที่ผู้ตรวจให้คะแนนทั้ง 2 คนร่วมกันตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม โดยแบ่งกันตรวจคนละครึ่งหนึ่งของข้อสอบร่วมเป็นการลดความลำเอียง และความเมื่อยล้าในการตรวจให้คะแนนเมื่อเทียบกับอีกวิธีหนึ่งที่ผู้ตรวจให้คะแนนเพียงคนเดียวต้องทำหน้าที่ในการตรวจให้คะแนนข้อสอบร่วมสร้างคำตอบทั้ง 4 ข้อของผู้สอบทั้ง 2 กลุ่มจากแบบสอบจำนวน 1,609 ฉบับ ซึ่งสอดคล้องกับ Johnson และคณะ (2009) ที่กล่าวว่า ในการประเมินข้อสอบแบบอัตนัย (essay test) และทักษะปฏิบัติต่างๆ ควรมีผู้ประเมิน (rater) หลายคน เพื่อให้คะแนนที่ได้จากการประเมินมีความเที่ยงมากขึ้น ซึ่งหากมีผู้ประเมินคนเดียวและผู้ประเมินเป็นผู้ออกข้อสอบหรือกำหนดงานที่ใช้ในการประเมิน ผู้ประเมินอาจนึกถึงคำตอบที่ตัวเองเฉลยไว้แล้วละเลยคำตอบอื่นๆที่เป็นไปได้ ดังนั้นควรมีผู้ประเมินอีกหนึ่งคนเพื่อให้ได้มุมมองคำตอบอื่น ๆ ที่เป็นไปได้ และสามารถใช้ข้อมูลในการตรวจสอบความเที่ยงระหว่างผู้ประเมินได้ด้วย (Lenel, 1990 cited in Johnson et al., 2009)

ข้อสังเกตที่ได้จากวิจัย พบว่า ในการแปลงคะแนนของนักเรียนระดับชั้นมัธยมศึกษาชั้นปีที่ 4 ด้วยวิธี CLE (Chained Linear Equating) ค่าความลำเอียง (bias) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสองในแต่ละจุดคะแนน ($Bias_i$ และ $RMSE_i$) ในช่วงคะแนนต้นๆ ประมาณช่วง 1–20 คะแนน มีค่าเข้าใกล้ศูนย์และมีค่าสม่ำเสมอ ซึ่งเมื่อพิจารณาความถี่ของผู้สอบในแต่ละจุดคะแนน พบว่า ผู้สอบกลุ่มใหม่และผู้สอบกลุ่มอ้างอิงมีความถี่ส่วนใหญ่อยู่ในช่วง 1-20 คะแนน โดยคะแนนจริงของผู้สอบกลุ่มใหม่ที่แปลงไปสู่สเกลเดียวกันกับผู้สอบกลุ่มอ้างอิงในช่วง 1-20 คะแนน มีช่วงความต่างของคะแนนคงที่มากกว่าในช่วงคะแนนปลาย ๆ กล่าวคือ ในช่วง 1-20 คะแนน คะแนนจริงที่ได้จากการแปลงไปสู่สเกลที่ต้องการมีช่วงต่างของคะแนน 3 ช่วง คือ ลดลง 1 คะแนน เท่าเดิม และเพิ่มขึ้น 1 คะแนน ขณะที่ตั้งแต่ 20 คะแนนขึ้นไป คะแนนจริงที่ได้จากการแปลงไปสู่สเกลที่ต้องการมีช่วงต่างของคะแนน 4 ช่วง คือ เพิ่มขึ้น 1 คะแนน, 2 คะแนน,

3 คะแนน และเพิ่มขึ้น 4 คะแนน โดยค่าความลำเอียง (bias) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ในช่วงตั้งแต่ 20 คะแนนขึ้นไปมีค่าไม่สม่ำเสมอ กล่าวคือ บางช่วงคะแนนมีค่าดังกล่าวเข้าใกล้ศูนย์ ในขณะที่บางช่วงคะแนนมีค่าดังกล่าวมากขึ้น นั่นแสดงให้เห็นว่าความถี่ของผู้สอบในแต่ละจุดคะแนนน่าจะมีผลต่อการแปลงคะแนนให้อยู่ในสเกลเดียวกัน รวมไปถึงค่าสถิติบูทสแตรปที่ได้จากการปรับเทียบคะแนน

แต่อย่างไรก็ตาม เมื่อพิจารณาสถิติบูทสแตรปของการปรับเทียบคะแนนที่ได้จากวิธีของเลอวิน (Levine method) มีความสอดคล้องกับวิธี CLE (Chained Linear Equating) กล่าวคือ วิธีการตรวจวิธีที่ 3: Half Trend มีค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) ต่ำที่สุด ในขณะเดียวกันก็มีความขัดแย้งกันระหว่างผลได้จากทั้ง 2 วิธี กล่าวคือ ในวิธีของเลอวิน (Levine method) วิธีตรวจที่มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำที่สุด คือ วิธีการตรวจวิธีที่ 1: Non Trend ส่วนรูปแบบที่มีค่าความลำเอียง (bias) ต่ำที่สุด คือ วิธีการตรวจวิธีที่ 2: Trend ขณะที่วิธี CLE (Chained Linear Equating) วิธีการตรวจที่มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) และค่าความลำเอียง (bias) ต่ำที่สุด คือ วิธีการตรวจวิธีที่ 3: Half Trend แสดงให้เห็นว่า นอกจากการข้อสอบร่วมแบบสร้างคำตอบที่มีผลต่อค่าสถิติบูทสแตรปแล้ว วิธีการปรับเทียบที่ใช้ย่อมมีผลต่อค่าสถิติดังกล่าวเช่นกัน ดังนั้นในการปรับเทียบคะแนนจึงต้องใช้วิธีการตรวจและวิธีการปรับเทียบคะแนนให้มีความเหมาะสม ซึ่งสำหรับบริบทของการวิจัยในครั้งนี้ วิธี CLE (Chained Linear Equating) มีความเหมาะสมในการปรับเทียบคะแนนมากกว่าวิธีของเลอวิน (Levine method)

ด้วยเหตุนี้ ในการปรับเทียบคะแนนสำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ด้วยวิธี CLE (Chained Linear Equating) ควรใช้วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วม (ผู้ตรวจทั้ง 2 คน ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม โดยแบ่งกันตรวจคนละครึ่งหนึ่งของข้อสอบร่วม) หรือวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบโดยมีการตรวจให้คะแนนข้ามกลุ่มทั้งหมดของข้อสอบร่วม (ผู้ตรวจคนหนึ่ง ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้ง 2 กลุ่ม) เนื่องจากการตรวจให้คะแนนข้ามกลุ่ม (trend scoring method) ในข้อสอบร่วมแบบสร้างคำตอบช่วยลดความลำเอียงในการปรับเทียบคะแนนได้ (Kim et al., 2010) และทำให้การปรับเทียบคะแนนในบริบทดังกล่าวมีประสิทธิภาพ

4.2 การอภิปรายตามสมมติฐานการวิจัยข้อที่ 2 วิธีการปรับเทียบเชิงเส้นตรง (equating methods)

จากการวิเคราะห์ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) พบว่า ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธี CLE (Chained Linear Equating) มีค่าต่ำกว่าวิธีของเลอวิน (Levine method) ในทุกวิธีการตรวจของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ดังนั้นวิธีการปรับเทียบคะแนนที่เหมาะสมในการนำมาใช้ในการปรับเทียบแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ที่มีการทดสอบจริงในชั้นเรียน คือ วิธี CLE

(Chained Linear Equating) นอกจากนี้เมื่อพิจารณาค่าสถิติบูทสแตรปของวิธีการตรวจทั้ง 3 วิธี ของการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) พบว่า ค่าความลำเอียง (bias) ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และ ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ที่ได้จากวิธี CLE (Chained Linear Equating) มีค่าต่ำกว่าสถิติบูทสแตรปดังกล่าวที่ได้จากวิธีของเลอวิน (Levine method) ซึ่งพบว่า ขัดแย้งกับสมมติฐานข้อที่ 2 ที่ผู้วิจัยตั้งไว้ว่า ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีการปรับเทียบเชิงเส้นตรงด้วยวิธีของเลอวิน (Levine method) มีค่าต่ำกว่าวิธี CLE (Chained Linear Equating) สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ที่ใช้ข้อมูลจากการทดสอบจริง ซึ่งเมื่อพิจาณางานวิจัยของ Suh และคณะ (2009) ที่ได้เปรียบเทียบวิธีการปรับเทียบเชิงเส้นตรงภายใต้รูปแบบ NEAT ซึ่งพบว่า วิธีของเลอวิน (Levine method) มีค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ต่ำกว่าวิธี CLE (Chained Liner Equating) แต่อย่างไรก็ตามงานวิจัยดังกล่าวใช้กลุ่มตัวอย่างขนาดใหญ่ คือ 40,000 คน จากการทดสอบ MBE ที่ใช้แบบสอบแบบหลายตัวเลือก ในขณะที่งานวิจัยในครั้งนี้ใช้กลุ่มตัวอย่างขนาดเล็กเพียง 1,609 คน และแบบสอบที่ใช้เป็นแบบสอบรูปแบบผสม ซึ่งบริบทของการทดสอบดังกล่าวไม่ว่าจะเป็นจำนวนกลุ่มตัวอย่าง และรูปแบบของแบบสอบนั้นต่างจากงานวิจัยในครั้งนี้ จึงเป็นไปได้ว่าผลการปรับเทียบที่ได้อาจแตกต่างกัน นอกจากนี้งานวิจัยดังกล่าวยังได้อภิปรายเพิ่มเติมว่า ถึงแม้ว่าค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ของวิธีของเลอวิน (Levine method) มีค่าต่ำกว่าวิธีอื่น ๆ แต่วิธี CLE (Chained Linear Equating) ก็สามารถนำมาใช้ได้เมื่อมีกลุ่มตัวอย่างมีความแตกต่างกัน หรือเมื่อคะแนนของข้อสอบร่วมทั้งหมดกับข้อสอบทั้งหมดมีความสัมพันธ์กันสูง และเมื่อพิจาณาค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนสอบทั้งฉบับและคะแนนของข้อสอบร่วมสำหรับวิธีการตรวจทั้ง 3 วิธี พบว่า คะแนนสอบทั้งฉบับและคะแนนของข้อสอบร่วมมีความสัมพันธ์กันสูง โดยค่า r มีค่าอยู่ระหว่าง .91 - .94 จึงเป็นไปได้ว่าวิธี CLE (Chained Linear Equating) มีความเหมาะสมในการนำมาใช้กับบริบทของงานวิจัยนี้ที่ใช้กับแบบสอบรูปแบบผสม และกลุ่มตัวอย่างมีขนาดเล็ก อีกทั้งในงานวิจัยของ Suh และคณะ (2009) ไม่ได้ใช้การตรวจข้อสอบแบบข้ามกลุ่ม เช่นเดียวกับงานวิจัยของ Mroch และคณะ (2009) ที่ได้ทำการประเมินวิธีการปรับเทียบเชิงเส้นตรงด้วยการวิเคราะห์จากเอกสารและงานวิจัยที่เกี่ยวข้อง ซึ่งไม่ได้มีการเก็บข้อมูลจริง งานวิจัยดังกล่าวพบว่า วิธีของเลอวิน (Levine method) มีค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ต่ำกว่าวิธี CLE (Chained Linear Equating) และวิธีของ Tucker (Tucker method) แต่อย่างไรก็ตามเมื่อพิจารณางานวิจัยของ Puhan (2010) ที่ได้ศึกษาผลลัพธ์ของวิธีการปรับเทียบเชิงเส้นตรง 3 วิธีกับกลุ่มตัวอย่างขนาดใหญ่ โดยใช้แบบสอบแบบหลายตัวเลือก ภายใต้กลุ่มตัวอย่างที่มีความสามารถแตกต่างกัน โดยใช้เกณฑ์ในการเปรียบเทียบ คือ ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ค่าความลำเอียง (bias) และค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) ซึ่งผลการวิจัยพบว่า ในกรณีที่กลุ่มตัวอย่างมีความสามารถแตกต่างกันอย่างมาก วิธีของเลอวิน (Levine method) ให้การปรับเทียบที่ดีกว่าเนื่องจากมีค่าความลำเอียง (bias) และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ต่ำ ในกรณีที่กลุ่มตัวอย่างมีความสามารถแตกต่างกันระดับ

ปานกลางไปจนถึงแตกต่างกันมาก วิธีของเลอวิน (Levine method) ให้ผลการปรับเทียบที่ดี แต่ถ้าในกรณีที่กลุ่มตัวอย่างมีความสามารถแตกต่างไม่มากนัก วิธี CLE (Chained Linear Equating) มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ค่าความลำเอียง (bias) และค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) ต่ำกว่าทั้งสองวิธีดังกล่าว นอกจากนี้ Puhan ยังได้อธิบายเพิ่มเติมว่า ถ้าผู้สอบสองกลุ่มมีความสามารถแตกต่างกัน โดยที่คะแนนจากแบบสอบทั้งฉบับและคะแนนจากข้อสอบร่วมทั้งหมดมีความสัมพันธ์กันสูง วิธี CLE (Chained Linear Equating) และวิธีของเลอวิน (Levine method) ต่างเป็นวิธีการที่เหมาะสมที่ในการนำมาใช้ในการปรับเทียบ และเมื่อพิจารณาวิจัยของ Puhan (2012) ที่ได้ประเมินวิธีการปรับเทียบคะแนน 2 วิธี ระหว่างวิธีของ Tucker (Tucker method) และ วิธี CLE (Chained Linear Equating) ภายใต้สถานการณ์ที่สมมติขึ้น คือ เมื่อการทดสอบมีผู้ตรวจหลายคน และคะแนนของแบบสอบทั้งฉบับกับคะแนนจากข้อสอบร่วมทั้งหมดมีความสัมพันธ์กันสูงมาก ผลการวิจัยพบว่า วิธีของ Tucker (Tucker method) มีความแม่นยำในการปรับเทียบคะแนนต่ำกว่าวิธี CLE (Chained Linear Equating)

จากงานวิจัยที่ผ่านมาที่ได้ทำการศึกษาถึงประสิทธิภาพของวิธีการปรับเทียบคะแนนเชิงเส้นตรง พบว่า เมื่อกลุ่มตัวอย่างมีความสามารถแตกต่างกัน และคะแนนของแบบสอบทั้งฉบับกับคะแนนของข้อสอบร่วมทั้งหมดมีความสัมพันธ์กันสูง วิธี CLE (Chained Linear Equating) และวิธีของเลอวิน (Levine method) ต่างมีความเหมาะสมในการนำมาใช้ในการปรับเทียบคะแนน แต่อย่างไรก็ตามงานวิจัยที่ผ่านมาส่วนใหญ่ทำการศึกษากับกลุ่มตัวอย่างขนาดใหญ่ อีกทั้งยังใช้แบบสอบแบบหลายตัวเลือก และในบางงานวิจัยก็ทำการศึกษาจากการจำลองข้อมูลสถานการณ์สมมติ และจากวิเคราะห์เอกสารและงานวิจัยที่เกี่ยวข้อง แต่ในการวิจัยครั้งนี้เป็นการศึกษากับกลุ่มตัวอย่างขนาดเล็ก ในสภาพการทดสอบจริงที่ใช้แบบสอบรูปแบบผสมที่มีข้อสอบร่วมรูปแบบผสมร่วมกับการตรวจให้คะแนนแบบข้ามกลุ่มในข้อสอบร่วมแบบสร้างคำตอบ ซึ่งยังไม่มีงานวิจัยใดที่ทำการเปรียบเทียบวิธีการปรับเทียบเชิงเส้นตรงทั้ง 2 วิธี ในบริบทดังกล่าว และเมื่อพิจารณาผลของการวิจัยในครั้งนี้ที่พบว่าวิธี CLE (Chain Linear Equating) ให้ค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ต่ำกว่าวิธีของเลอวิน (Levine method) ซึ่งมีความเป็นไปได้ที่อาจเกิดขึ้น เนื่องจากงานวิจัยที่ผ่านมาพบว่าวิธี CLE (Chain linear equating) มีความเหมาะสมในการนำมาใช้กับกลุ่มตัวอย่างที่มีความสามารถแตกต่างกัน โดยที่คะแนนของแบบสอบทั้งฉบับและคะแนนของข้อสอบร่วมทั้งหมดมีความสัมพันธ์กันสูง ซึ่งเมื่อพิจารณาคะแนนของกลุ่มใหม่และกลุ่มอ้างอิงที่ได้จากงานวิจัยนี้ในแต่ละวิธีการตรวจ พบว่า คะแนนเฉลี่ยของผู้สอบทั้งสองกลุ่มมีความแตกต่างกัน และค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนของแบบสอบทั้งฉบับและคะแนนของข้อสอบร่วมทั้งหมดก็มีค่าสูง ดังนั้นจึงสามารถสรุปได้ว่าวิธี CLE (Chained Linear Equating) มีความเหมาะสมมากกว่าวิธีของเลอวิน (Levine method) ในการนำมาใช้ปรับเทียบคะแนนภายใต้การออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) โดยกลุ่มตัวอย่างมีขนาดเล็กและมีความสามารถแตกต่างกัน เนื่องจากวิธี CLE (Chained Linear Equating) มีค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ต่ำกว่าวิธีของเลอวิน (Levine method) อีกทั้งค่าสถิติบูทสแตรปของวิธีการตรวจทั้ง 3 วิธี ได้แก่ ค่าความลำเอียง (bias) ค่าความคลาดเคลื่อนในการปรับเทียบ (equating error) และค่ารากที่สองของ

ค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ที่ได้จากวิธี CLE (Chained Linear Equating) มีค่าต่ำกว่าค่าสถิติบูทสแตรปดังกล่าวที่ได้จากวิธีของเลอวิน (Levine method) อีกด้วย

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำไปปฏิบัติ

1.1 ในการแปลความหมายคะแนนที่ปรับเทียบแล้ว ควรระวังในการสื่อความหมายหรือการนำไปใช้อ้างอิง เนื่องจากความหมายของคะแนนเหล่านี้ไม่ได้แปลความหมายเหมือนจากการสอบ เพราะคะแนนที่ได้เป็นค่าความสัมพันธ์เมื่อเปรียบกับผู้สอบอีกกลุ่มหนึ่ง ที่มีการสอบคนละช่วงเวลากัน แต่มีการใช้ข้อสอบร่วมกัน

1.2 ผลการวิจัยในครั้งนี้ พบว่า ค่าสถิติบูทสแตรป (bootstrapped statistics) ของวิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่มครึ่งหนึ่งของข้อสอบร่วมมีค่าต่ำที่สุดในการแปลงคะแนนจากผู้สอบกลุ่มใหม่ไปยังกลุ่มอ้างอิง สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ดังนั้นในการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ควรให้ผู้ตรวจทั้งสองคนทำหน้าที่ตรวจข้อสอบร่วมแบบสร้างคำตอบของผู้สอบทั้งสองกลุ่ม โดยแบ่งกันตรวจคนละครึ่งหนึ่ง และถ้าต้องการให้การปรับเทียบคะแนนในรูปแบบดังกล่าวมีค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ต่ำ ๆ ควรเลือกใช้วิธี CLE (Chained Linear Equating) ในการแปลงคะแนนของผู้สอบให้อยู่ในสเกลเดียวกันสำหรับบริบทที่กลุ่มตัวอย่างมีขนาดเล็ก ภายใต้การออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ที่ทำการทดสอบจริงในชั้นเรียน

1.3 ผลการวิจัยครั้งนี้ พบว่า คะแนนเฉลี่ยของผู้สอบทั้ง 2 กลุ่มจากแบบสอบเรื่องงานและพลังงานมีคะแนนเฉลี่ยต่ำกว่าครึ่งหนึ่ง และผู้สอบจำนวนไม่น้อยที่ไม่ทำข้อสอบแบบสร้างคำตอบ โดยเลือกทำเฉพาะข้อสอบแบบหลายตัวเลือกเท่านั้น ดังนั้นครูควรมีการจัดการเรียนการสอน หรือการทดสอบในห้องเรียนที่ส่งเสริมให้นักเรียนได้แสดงวิธีทำ หรือเขียนคำตอบของตนเองมากขึ้น เนื่องจากสารสนเทศที่ได้จากการเขียนตอบของนักเรียนทำให้เห็นถึงมโนทัศน์ของนักเรียนที่ไม่ถูกต้องเกี่ยวกับเรื่องนั้น ๆ ซึ่งครูสามารถนำสารสนเทศดังกล่าวไปใช้ในการปรับปรุงการจัดการเรียนการสอนให้ประสิทธิภาพมากยิ่งขึ้น

2. ข้อเสนอแนะสำหรับการวิจัยในครั้งต่อไป

2.1 ควรมีการเปรียบเทียบสถิติบูทสแตรปของวิธีการตรวจทั้ง 3 วิธี สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) ในรายวิชาอื่น ๆ อย่างเช่น วิชาภาษาไทย วิชาชีววิทยา และวิชาภาษาอังกฤษ เป็นต้น เพื่อที่จะได้ทราบว่าวิธีการตรวจวิธีใดที่เหมาะสมในการนำไปใช้ในรายวิชาต่าง ๆ

2.2 ควรมีการออกแบบวิธีการตรวจ และวิธีการปรับเทียบคะแนน สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียม

กัน (NEAT) ในแบบสอบที่มีลักษณะหลายมิติ (multidimentional) เนื่องจากในปัจจุบันมีการใช้ข้อสอบที่มีลักษณะเนื้อหาหลายมิติมากขึ้น หากมีความต้องการปรับเทียบคะแนนในแบบสอบประเภทดังกล่าว จะได้มีวิธีการตรวจและวิธีการปรับเทียบคะแนนที่มีความเหมาะสมและมีประสิทธิภาพ

2.3 ควรมีการศึกษาสัดส่วน/จำนวนของข้อสอบร่วมรูปแบบผสมที่ใช้ในการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) เมื่อใช้วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่ม ซึ่งจากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องลักษณะและสัดส่วนข้อสอบร่วม พบว่า ข้อสอบร่วมควรสร้างให้มีรายละเอียดคล้ายคลึงกับแบบสอบทั้งฉบับ โดยข้อสอบร่วมควรมีความยาวพอที่จะเป็นตัวแทนเนื้อหาในแบบสอบ ซึ่ง Hambleton และคณะ (1991) แนะนำว่า ข้อสอบร่วมควรมีจำนวนไม่น้อยกว่า 20 % - 25% ของข้อสอบทั้งฉบับ ขณะที่ Angoff และคณะ (1971) แนะนำว่า ถ้าแบบสอบมีจำนวนข้อสอบทั้งหมด 40 ข้อ ควรใช้ข้อสอบร่วมอย่างต่ำ 20% และในกรณีที่แบบสอบมีข้อสอบจำนวนมาก จำนวนข้อสอบร่วมที่เพียงพอ คือ 30% ซึ่งการวิจัยในครั้งนี้ได้สร้างข้อสอบร่วมให้กระจายตามจุดประสงค์เชิงพฤติกรรม แต่อย่างไรก็ตามแบบสอบทั้งสองฉบับมีจำนวนข้อสอบเฉพาะและจำนวนข้อสอบร่วมเท่ากันตามงานวิจัยของ Kim และคณะ (2010) ซึ่งอาจทำให้ผู้สอบใช้เวลาในการทำข้อสอบร่วมมากเกินไป ดังนั้นในงานวิจัยครั้งต่อไปอาจศึกษาสัดส่วนของข้อสอบร่วมสำหรับบริบทดังกล่าวในสัดส่วนต่างๆ อย่างเช่น 20% 30% 40% และ 50% สำหรับการออกแบบการปรับเทียบคะแนนแบบสอบรูปแบบผสมโดยใช้ข้อสอบร่วมสำหรับกลุ่มไม่เท่าเทียมกัน (NEAT) เมื่อใช้วิธีตรวจข้อสอบร่วมแบบสร้างคำตอบ โดยมีการตรวจให้คะแนนข้ามกลุ่ม

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

รายการอ้างอิง

ภาษาไทย

พัชรี จันทรเพ็ง. (2547). การปรับเทียบคุณภาพของวิธีการเชื่อมโยงคะแนนตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติภายใต้การหมุนแกน โครงสร้างเชิงมิติและระดับความสัมพันธ์ที่แตกต่างกัน (ปริญญามหาบัณฑิต), จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2552). ทฤษฎีการทดสอบแบบดั้งเดิม (6 ed.). กรุงเทพ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย

ศิริชัย กาญจนวาสี. (2555). ทฤษฎีการทดสอบแนวใหม่ (4 ed.). กรุงเทพ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย

ศิริชัย กาญจนวาสี, ดิเรก ศรีสุโข, & ทวีวัฒน์ ปิตยานนท์. (2537). การเลือกใช้สถิติที่เหมาะสมสำหรับการวิจัย (3 ed.). กรุงเทพฯ: บุญศิริการพิมพ์.

อัญชลี ศรีชลกาญ. (2552). คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม : การประยุกต์ใช้การปรับเทียบตามทฤษฎีการตอบสนองข้อสอบ ด้วยวิธีโค้งคุณลักษณะและการปรับค่าพารามิเตอร์พร้อมกัน (ปริญญาดุษฎีบัณฑิต), จุฬาลงกรณ์มหาวิทยาลัย.

ภาษาอังกฤษ

Albano, A. (2014). An R Package for observed-score linking and equating. Retrieved 22 March 2014, from http://cran.r-project.org/web/packages/equate/index.html

Angoff, W. H., Thorndike, R. L., & Lindquist, E. F. (1971). *Educational measurement* (2nd ed.). Washington: American Council on Education.

Davier, A. A. (2011). *Statistical models for test equating, scaling, and linking* A. A. Davier (Ed.)

Hambleton, R. K., & Swaminathan, H. (1985). *Item respons theory : principles and application* Boston: Nighoff Pub.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory* (2 ed.). Newbury Park, Calif: Sage Publications.

Johnson, R. L., Penny, J. A., & Gordon, B. (2009). *Assessing performance : designing, scoring, and validating performance tasks*. New York: The Guilford Press.

Kamata, A., & Tate, R. (2005). The Performance of a Method for the Long-term Equating of Mixed-Format Assessment. *Journal of Educational Measurement, 42*(2), 193-213.

Kim, S.-H., & Cohen, A. S. (2002). A Comparison of Linking and Concurrent Calibration Under the Graded Response Model. *Applied Psychological Measurement, 26*(1), 25-41. doi: 10.1177/0146621602026001002

Kim, S., & Kolen, M. J. (2006). Robustness to Format Effects of IRT Linking Methods for Mixed-Format Tests. *Applied Measurement in Education, 19*(4), 357-381. doi: 10.1207/s15324818ame1904_7

Kim, S., & Lee, W.-C. (2006). An Extension of Four IRT Linking Methods forMixed-Format Tests. *Journal of Educational Measurement, 43*(1), 53-76.

Kim, S., & Walker, M. (2012). Determining the Anchor Composition for a Mixed-Format Test: Evaluation of Subpopulation Invariance of Linking Functions. *Applied Measurement in Education, 25*(2), 178-195. doi: 10.1080/08957347.2010.524720

Kim, S., Walker, M. E., & McHale, F. (2010). Comparisons among Designs for Equating Mixed-Format Tests in Large-Scale Assessments. *Journal of Educational Measurement, 47*(1), 36-53.

Kim, S., Walker, M. E., & McHale, F. (2010). Investigating the Effectiveness of Equating Designs for Constructed-Response Tests in Large-Scale Assessments. *Journal of Educational Measurement, 47*(2), 186-201.

Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking : methods and practices* (2 ed.). New York Springer Science Business Media.

Li, D., Jiang, Y., & Davier, A. A. v. (2012). The Accuracy and Consistency of a Series of IRT True Score Equatings. *Journal of Educational Measurement, 40*(2), 167-180.

Mroch, A. A., Suh, Y., Kane, M. T., & Ripkey, D. R. (2009). An Evaluation of Five Linear Equating Methods for the NEAT Design. *Measurement: Interdisciplinary Research & Perspective, 7*(3-4), 174-193. doi: 10.1080/15366360903418063

Muraki, E., Hombo, C. M., & Lee, Y. W. (2000). Equating and Linking of Performance Assessments. *Applied Psychological Measurement, 24*(4), 325-337. doi: 10.1177/01466210022031787

Puhan, G. (2010). A Comparison of Chained Linear and Poststratification Linear Equating Under Different Testing Conditions. *Journal of Educational Measurement, 47*(1), 54-75.

Puhan, G. (2012). Choosing Among Tucker or Chained Linear Equating in Two Testing Situations: Rater Comparability Scoring and Randomly Equivalent Groups With an Anchor. *Journal of Educational Measurement, 49*(3), 312-329.

Suh, Y., Mroch, A. A., Kane, M. T., & Ripkey, D. R. (2009). An Empirical Comparison of Five Linear Equating Methods for the NEAT Design. *Measurement: Interdisciplinary Research & Perspective, 7*(3-4), 147-173. doi: 10.1080/15366360903418048

Tate, R. (2000). Performance of a Proposed Method for the Linking of Mixed Format Tests With Constructed Response and Multiple Choice Items. *Journal of Educational Measurement, 37*(4), 329-346.

ภาคผนวก

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ภาคผนวก ก
รายชื่อผู้เชี่ยวชาญ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

รายชื่อผู้เชี่ยวชาญ

1. รองศาตราจารย์ ดร. โชติกา ภาษีผล	อาจารย์ประจำภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
2. ผู้ช่วยศาสตราจารย์ ดร. กมลวรรณ ตังธนกานนท์	อาจารย์ประจำภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
3. อาจารย์อมรรัตน์ บุบผโชติ	อาจารย์กลุ่มสาระการเรียนรู้วิทยาศาสตร์ โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม
4. อาจารย์วรรณา นาคศรีอาภรณ์	อาจารย์กลุ่มสาระการเรียนรู้วิทยาศาสตร์ โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม
5. ว่าที่ร้อยตรีปกรณ์ ปานรอด	อาจารย์กลุ่มสาระการเรียนรู้วิทยาศาสตร์ โรงเรียนโรงเรียนสตรีศรีสุริโยทัย
6. นางสาวธันยากานต์ ยืนตระกูล	นักวิชาการ สาขาประเมินมาตรฐาน สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ข

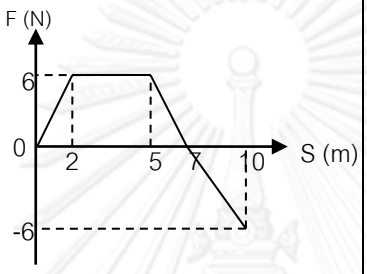
ผลการตรวจสอบความสอดคล้องระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรม
และระดับพฤติกรรมของแบบสอบวิชาฟิสิกส์เรื่องงานและพลังงานฉบับใหม่และฉบับอ้างอิง
พร้อมทั้งผลการปรับปรุงข้อสอบ

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

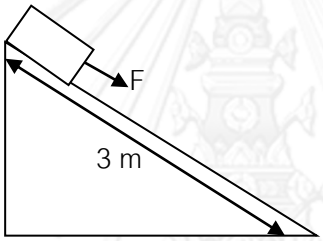
ตาราง 37 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับใหม่ (ฉบับที่ 1)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
1) งาน	1) อธิบายงานและวิเคราะห์งานของแรงต่างๆได้	1) แรง 10 นิวตัน ทำให้วัตถุเคลื่อนที่เป็นเส้นตรงบนพื้นราบลื่นในระยะทาง S ถ้าต้องการหาค่างานของแรง 20 N จะต้องพิจารณาสิ่งใดบ้าง (เข้าใจ) 1. อัตราเร็วของวัตถุ 2. แรงเสียดทานระหว่างวัตถุกับผิวสัมผัส 3. มุมระหว่างทิศของแรงกับแนวการเคลื่อนที่ของวัตถุ ก. 1 ข. 2 ค. 3 ง. 2 และ 3 <u>คำตอบ</u> คือ ข้อ ค	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม แต่ควรปรับหน่วยให้เป็นภาษาไทย
		2) กรณีใดที่ถือว่ามีการทำงานเนื่องจากการออกแรงกระทำต่อหนังสือ (เข้าใจ) 1. แจนยืนถือหนังสือ 2. แจนนั่งอ่านหนังสือเตรียมสอบ 3. แจนยกหนังสือจากพื้นไปวางบนโต๊ะในแนวดิ่ง 4. แจนถือหนังสือแล้วเดินไปตามพื้นราบ ก. 3 ข. 3 และ 4 ค. 1 3 และ 4 ง. 1 2 3 และ 4 <u>คำตอบ</u> คือ ข้อ ก	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม

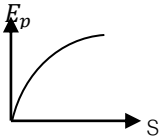
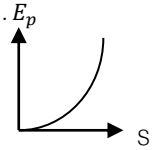
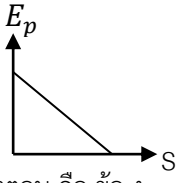
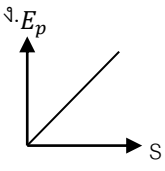
ตาราง 37 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับใหม่ (ฉบับที่ 1) (ต่อ)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
1) งาน (ต่อ)	2) คำนวณหางานในการเคลื่อนที่ของวัตถุ	3) งานทั้งหมดที่กระทำในช่วงระยะทางการเคลื่อนที่ 0 เมตรถึง 10 เมตร จะมีค่ากี่จูล ถ้ากราฟระหว่างแรงกับการกระจัดเป็นดังรูป (เข้าใจ)  คำตอบ คือ 21 จูล	1	1	1	-1	1	0.8	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
		4) มินออกแรง 100 นิวตันลากกล่องในแนวทำมุม 60 องศา กับแนวระดับ ถ้าลากกล่องไปได้ไกล 10 เมตรในแนวระดับ งานของแรงที่ใช้ลากกล่องเป็นเท่าใด (นำไปใช้) ก. 500 จูล ข. $500\sqrt{3}$ จูล ค. 1,000 จูล ง. $1{,}000\sqrt{3}$ จูล คำตอบ คือ ข้อ ก	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
		5) ลังมวล 40 กิโลกรัมเคลื่อนที่ลงไปตามพื้นเอียงผิวขรุขระเป็นระยะทาง 10 เมตร ด้วยความเร็วคงที่ ถ้างานในการเคลื่อนที่ของลังดังกล่าวเป็น 3,600 จูล แรงเสียดทานจลน์จะมีค่าเท่าใด ถ้าพื้นเอียงอยู่สูงจากพื้น 6 เมตร (วิเคราะห์) ก. - 40 N ข. 40 N ค. -120 N ง.120 N คำตอบ คือ ข้อ ค	1	-1	1	1	1	0.8	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม แต่ควรปรับหน่วยให้เป็นภาษาไทย

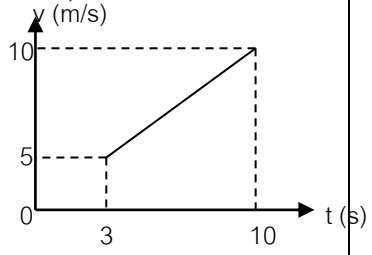
ตาราง 37 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับใหม่ (ฉบับที่ 1) (ต่อ)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
1) งาน (ต่อ)	2) คำนวณหางานในการเคลื่อนที่ของวัตถุ	6) ช้างตัวหนึ่งออกแรงดึงซุงหนัก 140 กิโลกรัม ลงเนินซึ่งเอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ สัมประสิทธิ์ความเสียดทานระหว่างเนินและน้ำหนักเป็น $\sqrt{3}/2$ ถ้าแรงที่ใช้ดึงซุงขนานกับการเคลื่อนที่ดังรูป ช้างต้องใช้งานเท่าใดในการดึงซุงลงมาได้ 3 เมตร (วิเคราะห์)  คำตอบ คือ 1,050 จูล	1	-1	1	1	1	0.8	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
2) กำลัง	3) คำนวณหากำลังจากโจทย์ที่กำหนดให้ได้	7) นาย ก. ออกแรงคงที่ขนาด 100 นิวตัน ผลักวัตถุก้อนหนึ่งจากหยุดนิ่งให้เคลื่อนที่ไปตามแนวราบได้ระยะทาง 3 เมตร ภายในเวลา 2 วินาที จงหากำลังของนาย ก. ที่ใช้ในการผลักวัตถุนี้ (นำไปใช้) ก. 20 วัตต์ ข. 150 วัตต์ ค. 300 วัตต์ ง. 600 วัตต์ คำตอบ คือ ข้อ ข	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม แต่ควรปรับตัวเลือก ก
		8) ลิฟท์ตัวหนึ่งมวล 200 กิโลกรัม ใช้มอเตอร์ 8 กิโลวัตต์ ขณะที่บรรทุกคนให้เคลื่อนที่ขึ้นด้วยความเร็วคงที่ 2 เมตร/วินาที จะต้องใช้แรงเท่าใด (นำไปใช้) ก. 4 นิวตัน ข. 25 นิวตัน ค. 100 นิวตัน ง. 4,000 นิวตัน คำตอบ คือ ง	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม

ตาราง 37 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับใหม่ (ฉบับที่ 1) (ต่อ)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
2) กำลัง	3) คำนวณหากำลังจากโจทย์ที่กำหนดให้ได้	9) รถยนต์คันหนึ่งมีมวล 400 กิโลกรัม แล่นขึ้นพื้นเอียงทำมุม 30 องศากับแนวระดับ ด้วยอัตราเร็วคงที่ 90 เมตร/วินาที จงหากำลังงานของเครื่องยนต์ในหน่วยกิโลวัตต์ ถ้าแรงเสียดทานของพื้นเอียงเท่ากับ 100 นิวตัน (วิเคราะห์) คำตอบ คือ 189 กิโลวัตต์	1	-1	1	1	1	0.8	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
		10) พลังงานจลน์ของวัตถุในกรณีใดมีค่ามากที่สุด (เข้าใจ) ก. เมื่อเพิ่มความเร็วเป็น 2 เท่า ข. เมื่อเพิ่มมวลเป็น 2 เท่า ค. เมื่อลดความเร็วลง 1/2 เท่า ง. เมื่อลดมวลลง 1/2 เท่า คำตอบ คือ ข้อ ก	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
		11) กราฟในตัวเลือกใดแสดงความสัมพันธ์ระหว่างพลังงานศักย์โน้มถ่วง (E_p) ของวัตถุกับการกระจัด (S) ได้ถูกต้อง เมื่อวัตถุนี้เคลื่อนที่จากจุดหยุดนิ่งออกไปในแนวเส้นตรงด้วยอัตราเร็วคงที่ (เข้าใจ) ก.  ข.  ค  ง.  คำตอบ คือ ข้อ ง	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม โดยผู้เชี่ยวชาญแนะนำให้ปรับตัวเลือก ค และ ง และเพิ่มความชัดเจนของโจทย์

ตาราง 37 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับใหม่ (ฉบับที่ 1) (ต่อ)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
3) พลังงานและความสัมพันธ์ระหว่างงานและพลังงาน	5) คำนวณหาพลังงานจลน์ พลังงานศักย์โน้มถ่วงของวัตถุที่ระดับต่างๆ และพลังงานศักย์ยืดหยุ่นของสปริง	12) กล่องมวล 2 กิโลกรัม ตกจากหน้าผาสูง 40 เมตรจากระดับพื้นดิน เมื่อกล่องหล่นมาได้ระยะทาง 15 เมตร พลังงานศักย์มีค่าเท่าใด (นำไปใช้) ก. 300 จูล ข. 500 จูล ค. 800 จูล ง. 1,000 จูล คำตอบ คือ ข้อ ข	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม แต่ควรปรับตัวเลือกข้อ ง
		13) ลูกมะพร้าวมวล 1 กิโลกรัม อยู่บนต้นมะพร้าว จงหาพลังงานจลน์ของลูกมะพร้าวเมื่อหลุดออกจากต้นมะพร้าวที่เวลา 1 วินาที (นำไปใช้) ก. 5 จูล ข. 10 นิวตัน ค . 20 นิวตัน ง. 50 นิวตัน คำตอบ คือ ข้อ ง	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
		14) สปริงถูกออกแรงดึง 20 นิวตัน จะทำให้ยืดออก 10 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริง เมื่อสปริงยืดออกมา 20 เซนติเมตร ในหน่วยจูล (วิเคราะห์) คำตอบ คือ 4 จูล	1	-1	1	0	1	0.6	วัดได้ตรงตามวัตถุประสงค์ ผู้เชี่ยวชาญแนะนำให้เพิ่มความซับซ้อนของโจทย์
	6) คำนวณโจทย์เกี่ยวกับความสัมพันธ์ระหว่างงานและพลังงานได้	15) รถยนต์มวล 2,000 กิโลกรัม เคลื่อนที่ไปตามพื้นราบ งานที่เกิดขึ้นในช่วง 3-10 วินาที มีค่ากี่จูล ถ้ากราฟระหว่างความเร็ว และเวลาของรถยนต์ขณะเคลื่อนที่เป็นดังรูป (นำไปใช้)  ก. 75 จูล ข. 750 จูล ค. 7,500 จูล ง. 75,000 จูล คำตอบ คือ ข้อ ง	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม

ตาราง 37 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับใหม่ (ฉบับที่ 1) (ต่อ)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
3) พลังงานและความสัมพันธ์ระหว่างงานและพลังงาน	6) คำนวณโจทย์เกี่ยวกับความสัมพันธ์ระหว่างงานและพลังงานได้	16) ยิงลูกปืนมวล 150 กรัมไปในแนวระดับ ลูกปืนมีความเร็ว 100 เมตร/วินาที ถ้าลูกปืนฝังเข้าไปฝังในเนื้อไม้ งานที่ไม้ต้านการเคลื่อนที่ของลูกปืนมีค่าเท่าใด (นำไปใช้) ก. 150 จูล ข. 750 จูล ค. 15,000 จูล ง. 75,000 จูล <u>คำตอบ</u> คือ ข้อ ข	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
		17) เคนนำเครื่องสูบน้ำเครื่องหนึ่งไปสูบน้ำมวล 400 กิโลกรัม ขึ้นไปตามท่อสูง 10 เมตร ถ้าน้ำพุ่งออกจากปลายท่อด้วยความเร็ว 10 เมตรต่อวินาที งานที่ได้จากเครื่องสูบมีค่ากี่กิโลจูล (วิเคราะห์) <u>คำตอบ</u> คือ 60 กิโลจูล	1	-1	1	0	1	0.6	วัดได้ตรงตามวัตถุประสงค์ แต่ควรเพิ่มความซับซ้อนของโจทย์
4) กฎการอนุรักษ์พลังงานกล	7) อธิบายกฎการอนุรักษ์พลังงานกลในการวิเคราะห์การเคลื่อนที่ในสถานการณ์ต่าง ๆ	18) เมื่อวัตถุก้อนหนึ่งเคลื่อนที่ลงจากพื้นเอียงลื่นมายังพื้นราบลื่น ข้อใดกล่าวถูกต้อง (เข้าใจ) ก. พลังงานรวมของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน ข. พลังงานจลน์ของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน ค. พลังงานศักย์โน้มถ่วงของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน ง. พลังงานรวม พลังงานจลน์ พลังงานศักย์โน้มถ่วง ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน <u>คำตอบ</u> คือ ข้อ ก	0	1	1	1	1	0.8	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม

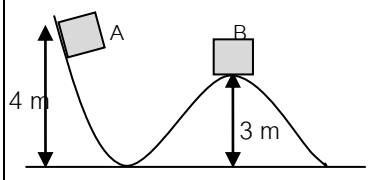
จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ตาราง 37 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับใหม่ (ฉบับที่ 1) (ต่อ)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
3) พลังงานและความสัมพันธ์ระหว่างงานและพลังงาน	6) คำนวณโจทย์เกี่ยวกับความสัมพันธ์ระหว่างงานและพลังงานได้	17) เคนนำเครื่องสูบน้ำเครื่องหนึ่งไปสูบน้ำมวล 400 กิโลกรัม ขึ้นไปตามท่อสูง 10 เมตร ถ้าน้ำพุ่งออกจากปลายท่อด้วยความเร็ว 10 เมตรต่อวินาที งานที่ได้จากเครื่องสูบมีค่ากี่กิโลจูล (วิเคราะห์) คำตอบ คือ 60 กิโลจูล	1	-1	1	0	1	0.6	วัดได้ตรงตามวัตถุประสงค์ แต่ควรเพิ่มความซับซ้อนของโจทย์
4) กฎการอนุรักษ์พลังงานกล	7) อธิบายกฎการอนุรักษ์พลังงานกลในการวิเคราะห์การเคลื่อนที่ในสถานการณ์ต่าง ๆ	18) เมื่อวัตถุก้อนหนึ่งเคลื่อนที่ลงจากพื้นเอียงลื่นมายังพื้นราบลื่น ข้อใดกล่าวถูกต้อง (เข้าใจ) ก. พลังงานรวมของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน ข. พลังงานจลน์ของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน ค. พลังงานศักย์โน้มถ่วงของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน ง. พลังงานรวม พลังงานจลน์ พลังงานศักย์โน้มถ่วง ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน คำตอบ คือ ข้อ ก	0	1	1	1	1	0.8	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
	8) ประยุกต์ใช้กฎการอนุรักษ์พลังงานกับการเคลื่อนในสถานการณ์ต่าง ๆ ได้	19) ลูกตาลมวล 2 กิโลกรัม ตกจากต้นตาลสูง 6 เมตร ขณะที่ลูกตาลเคลื่อนที่ลงมาได้ระยะทาง 2 เมตรจากต้นตาล ความเร็วของลูกตาลมีค่าเท่าใด (นำไปใช้) ก. $\sqrt{40}$ เมตร/วินาที ข. $\sqrt{80}$ เมตร/วินาที ค. 40 เมตร/วินาที ง. 80 เมตร/วินาที คำตอบ คือ ข้อ ข	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 37 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับใหม่ (ฉบับที่ 1) (ต่อ)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
4) กฎการอนุรักษ์พลังงานกล	8) ประยุกต์ใช้กฎการอนุรักษ์พลังงานกับการเคลื่อนในสถานการณ์ต่าง ๆ ได้	20) กล่องมวล 25 กิโลกรัม เคลื่อนที่ด้วยความเร็ว 10 เมตร/วินาที เข้าชนสปริงที่มีค่าคงที่ 100 นิวตัน/เมตร เมื่อกล่องชนสปริง สปริงจะหดสั้นที่สุดเท่าไหร่ (นำไปใช้) ก. 5 ซม. ข. 500 ซม. ค. 550 ซม. ง. 1,000 ซม. <u>คำตอบ</u> คือ ข้อ ข	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม แต่ควรปรับตัวเลือก ค และ ง
		21) ดาวต้องการยิงลูกหินมวล 100 กรัม ด้วยหนังสติ๊กให้ขึ้นไปได้สูง 10 เมตรในแนวดิ่ง ถ้าการยืดของหนังสติ๊กเป็นแบบสปริงและยืดออก 50 เซนติเมตร หนังสติ๊กจะต้องมีค่านิจสปริงกี่นิวตัน/เมตร (นำไปใช้) <u>คำตอบ</u> คือ 80 นิวตัน/เมตร	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
		22) วัตถุมวล 8 กิโลกรัม ที่จุด A มีอัตราเร็ว 0.5 เมตร/วินาที และที่จุด B มีอัตราเร็ว 9 เมตร/วินาที ถ้าระยะทางจากจุด A ไปยังจุด B เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (วิเคราะห์)  <u>คำตอบ</u> คือ -81 นิวตัน	1	-1	1	1	1	0.8	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม

ตาราง 37 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับใหม่ (ฉบับที่ 1) (ต่อ)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
5) เครื่องกล	9) คำนวณโจทย์เกี่ยวกับเครื่องกลโดยอาศัยหลักการทำงานของเครื่องกลได้	23) มอเตอร์เครื่องหนึ่งที่มีประสิทธิภาพ 90% จะให้พลังงานงานที่เป็นประโยชน์ออกมาเท่าใด เมื่อใส่กำลังให้แก่มอเตอร์เครื่องนี้ 0.5 กำลังม้า (1 กำลังม้า = 750 วัตต์) (นำไปใช้) ก. 45.5 วัตต์ ข. 237.5 วัตต์ ค. 337.5 วัตต์ ง. 675 วัตต์ <u>คำตอบ</u> คือ ข้อ ค	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม แต่ควรตรวจสอบค่ากำลังม้า
		24) ลิฟท์อันหนึ่งมวล 650 กิโลกรัม สามารถยกของมวล 350 กิโลกรัม ขึ้นไปได้สูง 15 เมตร ในเวลา 5 วินาที โดยใช้กำลังไฟฟ้าไป 40 กิโลวัตต์ ลิฟท์นี้ต้องมีประสิทธิภาพกี่เปอร์เซ็นต์ (นำไปใช้) <u>คำตอบ</u> คือ 75%	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม

ตาราง 38 ผลการปรับปรุงข้อสอบวิชาฟิสิกส์ เรื่อง งานและพลังงาน ระดับชั้นมัธยมศึกษาปีที่ 4 (ฉบับที่1) ตามคำแนะนำของผู้เชี่ยวชาญ

<u>จุดประสงค์ที่ 1</u> อธิบายงานและวิเคราะห์งานของแรงต่างๆได้

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
1) แรง 10 นิวตัน ทำให้วัตถุเคลื่อนที่เป็นเส้นตรงบนพื้นราบลื่นในระยะทาง S ถ้าต้องการหาค่างานของแรง 20 N จะต้องพิจารณาสิ่งใดบ้าง 1. อัตราเร็วของวัตถุ 2. แรงเสียดทานระหว่างวัตถุกับผิวสัมผัส 3. มุมระหว่างทิศของแรงกับแนวการ เคลื่อนที่ของวัตถุ ก. 1 ข. 2 ค. 3 ง. 3 และ 4	1) แรง 10 นิวตัน ทำให้วัตถุเคลื่อนที่เป็นเส้นตรงบนพื้นราบลื่นในระยะทาง S ถ้าต้องการหาค่างานของแรง 20 นิวตัน จะต้องพิจารณาสิ่งใดบ้าง ก. อัตราเร็วของวัตถุ ข. แรงเสียดทานระหว่างวัตถุกับผิวสัมผัส ค. มุมระหว่างทิศของแรงกับแนวการเคลื่อนที่ของวัตถุ 1. ก 2. ข 3. ค 4. ข และ ค

ตาราง38 ผลการปรับปรุงข้อสอบวิชาฟิสิกส์ เรื่อง งานและพลังงาน ระดับชั้นมัธยมศึกษาปีที่ 4 (ฉบับที่1) ตามคำแนะนำของผู้เชี่ยวชาญ (ต่อ)

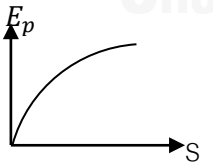
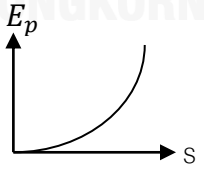
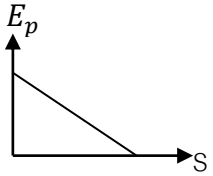
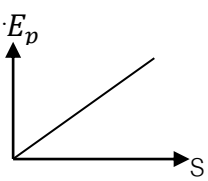
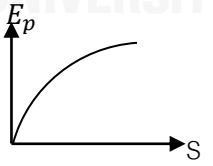
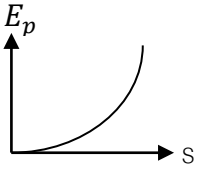
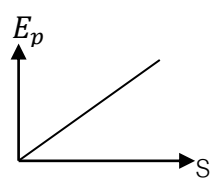
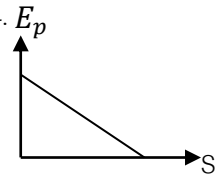
จุดประสงค์ที่ 2 คำนวณหางานในการเคลื่อนที่ของวัตถุ

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
5) ลังมวล 40 กิโลกรัมเคลื่อนที่ลงไปตามพื้นเอียงผิวขรุขระเป็นระยะทาง 10 เมตร ด้วยความเร็วคงที่ ถ้างานในการเคลื่อนที่ของลังดังกล่าวเป็น 3,600 จูล แรงเสียดทานจลน์จะมีค่าเท่าใด ถ้าพื้นเอียงอยู่สูงจากพื้น 6 เมตร ก. - 40 N ข. 40 N ค. -120 N ง.120 N	5) ลังมวล 40 กิโลกรัมเคลื่อนที่ลงไปตามพื้นเอียงผิวขรุขระเป็นระยะทาง 10 เมตร ด้วยความเร็วคงที่ ถ้างานในการเคลื่อนที่ของลังดังกล่าวเป็น 3,600 จูล แรงเสียดทานจลน์จะมีค่าเท่าใด ถ้าพื้นเอียงอยู่สูงจากพื้น 6 เมตร 1. - 40 นิวตัน 2. 40 นิวตัน 3. -120 นิวตัน 4.120 นิวตัน

จุดประสงค์ที่ 3 คำนวณหากำลังจากโจทย์ที่กำหนดให้ได้

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
7) นาย ก. ออกแรงคงที่ขนาด 100 นิวตัน ผลักวัตถุก้อนหนึ่งจากหยุดนิ่งให้เคลื่อนที่ไปตามแนวราบได้ระยะทาง 3 เมตร ภายในเวลา 2 วินาที จงหากำลังของนาย ก. ที่ใช้ในการผลักวัตถุนี้ (นำไปใช้) ก. 20 วัตต์ ข. 150 วัตต์ ค. 300 วัตต์ ง. 600 วัตต์	7) นาย ก. ออกแรงคงที่ขนาด 100 นิวตัน ผลักวัตถุก้อนหนึ่งจากหยุดนิ่งให้เคลื่อนที่ไปตามแนวราบได้ระยะทาง 3 เมตร ภายในเวลา 2 วินาที จงหากำลังของนาย ก. ที่ใช้ในการผลักวัตถุนี้ (นำไปใช้) 1. 17 วัตต์ 2. 150 วัตต์ 3. 300 วัตต์ 4. 600 วัตต์

จุดประสงค์ที่ 4 อธิบายพลังงาน พลังงานจลน์ พลังงานศักย์ และความสัมพันธ์ระหว่างงานและพลังงานได้

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
11) กราฟในตัวเลือกใดแสดงความสัมพันธ์ระหว่างพลังงานศักย์โน้มถ่วง (E_p) ของวัตถุกับการกระจัด (F) ได้ถูกต้อง เมื่อวัตถุนี้เคลื่อนที่จากจุดหยุดนิ่งออกไปในแนวเส้นตรงด้วยอัตราเร็วคงที่ ก.  ข.  ค.  ง. 	11) กราฟในตัวเลือกใดแสดงความสัมพันธ์ระหว่างพลังงานศักย์โน้มถ่วง (E_p) ของวัตถุกับการกระจัด (S) ได้ถูกต้อง เมื่อวัตถุนี้ถูกปล่อยจากจุดหยุดนิ่งลงมาในแนวดิ่ง ด้วยอัตราเร็วคงที่ 1.  2.  3.  4. 

ตาราง38 ผลการปรับปรุงข้อสอบวิชาฟิสิกส์ เรื่อง งานและพลังงาน ระดับชั้นมัธยมศึกษาปีที่ 4 (ฉบับที่1) ตามคำแนะนำของผู้เชี่ยวชาญ (ต่อ)

จุดประสงค์ที่ 5 คำนวณหาพลังงานจลน์ พลังงานศักย์โน้มถ่วงของวัตถุที่ระดับต่างๆ และพลังงานศักย์ยืดหยุ่นของสปริง

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
12) กล่องมวล 2 กิโลกรัม ตกจากหน้าผาสูง 40 เมตร จากระดับพื้นดิน เมื่อกล่องหล่นมาได้ระยะทาง 15 เมตร พลังงานศักย์มีค่าเท่าใด (นำไปใช้) ก. 300 จูล ข. 500 จูล ค. 800 จูล ง. 1,000 จูล	12) กล่องมวล 2 กิโลกรัม ตกจากหน้าผาสูง 40 เมตร จากระดับพื้นดิน เมื่อกล่องหล่นมาได้ระยะทาง 15 เมตร พลังงานศักย์มีค่าเท่าใด (นำไปใช้) 1. 300 จูล 2. 500 จูล 3. 800 จูล 4. 1,100 จูล
14) สปริงถูกออกแรงดึง 20 นิวตัน จะทำให้ยืดออก 10 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 20 เซนติเมตรในหน่วยจูล	14) ออกแรงดึงสปริง 50 นิวตัน ปรากฏว่าสปริงยาว 70 เซนติเมตร แต่ถ้าออกแรงดึงสปริง 60 นิวตัน สปริงจะมีความยาว 80 เซนติเมตร พลังงานศักย์ยืดหยุ่นขณะที่ออกแรงดึงสปริง 60 นิวตันมีค่าเท่าใด

จุดประสงค์ที่ 6 คำนวณโจทย์เกี่ยวกับความสัมพันธ์ระหว่างงานและพลังงานได้

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
17) เคนนำเครื่องสูบน้ำเครื่องหนึ่งไปสูบน้ำมวล 400 กิโลกรัม ขึ้นไปตามท่อสูง 10 เมตร ถ้าน้ำพุ่งออกจากปลายท่อด้วยความเร็ว 10 เมตรต่อวินาที งานที่ได้จากเครื่องสูบมีค่ากี่กิโลจูล	17) เคนเข็นลังใส่ผลไม้มวล 50 กิโลกรัม ให้เคลื่อนที่ในแนวเส้นตรงด้วยแรงคงที่ค่าหนึ่งไปบนพื้นลื่น เมื่อผ่านตำแหน่ง ก เคนต้องใช้กำลังเป็น $\frac{2}{3}$ เท่าของกำลังที่ใช้เมื่อผ่านตำแหน่ง ข ถ้าความเร็วของลังนี้เมื่อผ่านตำแหน่ง ข คือ 3 เมตร/วินาที งานที่ใช้ในการเข็นลังนี้จากตำแหน่ง ก ไปยังตำแหน่ง ข มีค่ากี่จูล

จุดประสงค์ที่ 8 ประยุกต์ใช้กฎการอนุรักษ์พลังงานกับการเคลื่อนในสถานการณ์ต่าง ๆ ได้

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
20) กล่องมวล 25 กิโลกรัม เคลื่อนที่ด้วยความเร็ว 10 เมตร/วินาที เข้าชนสปริงที่มีค่าคงที่ 100 นิวตัน/เมตร เมื่อกล่องชนสปริง สปริงจะหดสั้นที่สุดเท่าไหร่ ก. 5 ซม. ข. 500 ซม. ค. 550 ซม. ง. 1,000 ซม.	20) กล่องมวล 25 กิโลกรัม เคลื่อนที่ด้วยความเร็ว 10 เมตร/วินาที เข้าชนสปริงที่มีค่าคงที่ 100 นิวตัน/เมตร เมื่อกล่องชนสปริง สปริงจะหดสั้นที่สุดเท่าไหร่ 1. 5 ซม. 2. 25 ซม. 3. 500 ซม. 4. 2,500 ซม.

จุดประสงค์ที่ 9 คำนวณโจทย์เกี่ยวกับเครื่องกลโดยอาศัยหลักการทำงานของเครื่องกลได้

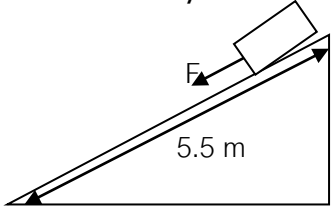
ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
23) มอเตอร์เครื่องหนึ่งที่มีประสิทธิภาพ 90% จะให้พลังงานงานที่เป็นประโยชน์ออกมาเท่าใด เมื่อใส่กำลังให้แก่มอเตอร์เครื่องนี้ 0.5 กำลังม้า (1 กำลังม้า = 750 วัตต์) ก. 45.5 วัตต์ ข. 237.5 วัตต์ ค. 337.5 วัตต์ ง. 675 วัตต์	23) มอเตอร์เครื่องหนึ่งที่มีประสิทธิภาพ 90% จะให้พลังงานงานที่เป็นประโยชน์ออกมาเท่าใด เมื่อใส่กำลังให้แก่มอเตอร์เครื่องนี้ 0.5 กำลังม้า (1 กำลังม้า = 746 วัตต์) 1.0.45 วัตต์ 2.33.57 วัตต์ 3.45.00 วัตต์ 4.335.70 วัตต์

ตาราง 39 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับอ้างอิง (ฉบับที่ 2)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
1) งาน	1) อธิบายงานและวิเคราะห์งานของแรงต่างๆได้	1) กล่องเคลื่อนที่เป็นเส้นตรงบนพื้นราบลื่นด้วยแรง 20 N เป็นระยะทาง S ถ้าต้องการหาค่างานของแรง 10 N จะต้องพิจารณาสิ่งใดบ้าง (เข้าใจ) 1. มุมระหว่างทิศของแรงกับแนวการเคลื่อนที่ของวัตถุ 2. แรงเสียดทานระหว่างวัตถุกับพื้นราบ 3. อัตราเร็วของวัตถุ ก. 1 ข. 2 ค. 3 ง. 2 และ 3 <u>คำตอบ</u> คือ ข้อ ก	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม แต่ควรปรับหน่วยให้เป็นภาษาไทย
		2) กรณีใดที่ถือว่ามีการทำงานเนื่องจากการออกแรงกระทำต่อกล่อง (เข้าใจ) 1. นุชยืนถือกล่อง 2. นุชยกกล่องจากโต๊ะไปวางบนพื้นในแนวดิ่ง 3. นุชยกกล่องจากพื้นไปวางบนโต๊ะในแนวดิ่ง 4. นุชถือกล่องแล้วเดินไปตามพื้นราบ ก. 1 ข. 2 และ 3 ค. 3 และ 4 ง. 1 2 3 และ 4 <u>คำตอบ</u> คือ ข้อ ข	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

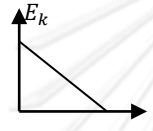
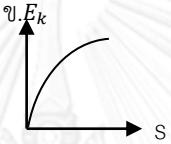
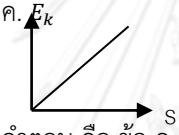
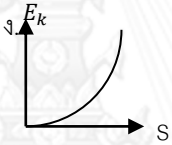
ตาราง 39 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับอ้างอิง (ฉบับที่ 2) (ต่อ)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
1) งาน (ต่อ)	2) คำนวณหางานในการเคลื่อนที่ของวัตถุ	4) นันลากกล่องไปตามพื้นราบเป็นระยะ 20 เมตร ด้วยแรง 70 นิวตัน ในแนวทำมุม 60 องศากับแนวระดับ ถ้างานของแรงที่ใช้ลากกล่องเป็นเท่าใด (นำไปใช้) ก. 700 จูล ข. $700\sqrt{3}$ จูล ค. 1400 จูล ง. $1400\sqrt{3}$ จูล คำตอบ คือ ข้อ ก	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
		5) กล่องมวล 20 กิโลกรัมเคลื่อนที่ลงไปตามพื้นเอียงผิวขรุขระด้วยความเร็วคงที่ เป็นระยะทาง 5 เมตร แรงเสียดทานระหว่างพื้นและลังจะมีค่าเท่าใด เมื่องานในการเคลื่อนที่ของลังดังกล่าวเป็น 1,400 จูล ถ้าพื้นเอียงอยู่สูงจากพื้น 3 เมตร (วิเคราะห์) ก. - 80 นิวตัน ข. -160 นิวตัน ค. 80 นิวตัน ง.160 นิวตัน คำตอบ คือ ข้อ ข	1	-1	1	1	1	0.8	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
		6) ชายคนหนึ่งออกแรงดึงกล่องหนัก 80 กิโลกรัมลงเนินที่เอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ ถ้าเขาออกแรงดึงกล่องในแนวขนานกับการเคลื่อนที่ดังรูป เขาต้องใช้งานเท่าใดในการดึงกล่องลงมาได้ 5.5 เมตร เมื่อสัมประสิทธิ์ความเสียดทานระหว่างเนินและกล่องเท่ากับ $\sqrt{3}/2$ (วิเคราะห์)  คำตอบ คือ 1,100 จูล	1	-1	1	1	1	0.8	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม

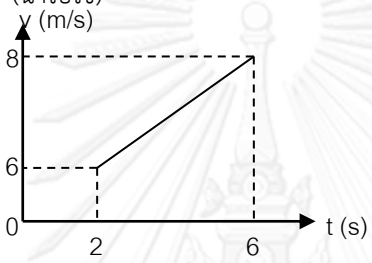
ตาราง 39 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับอ้างอิง (ฉบับที่ 2) (ต่อ)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
2) กำลัง	3) คำนวณหากำลังจากโจทย์ที่กำหนดให้ได้	7) จ็อบออกแรงคงที่ขนาด 200 นิวตัน ผลักกล่องใบหนึ่งจากหยุดนิ่งให้เคลื่อนที่ไปตามแนวราบได้ระยะทาง 2 เมตร ภายในเวลา 4 วินาที จงหากำลังของจ็อบที่ใช้ในการผลักกล่องใบนี้ (นำไปใช้) ก. 100 วัตต์ ข. 400 วัตต์ ค. 500 วัตต์ ง. 600 วัตต์ คำตอบ คือ ข้อ ก	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม แต่ควรปรับตัวเลือก ค และ ง
		8) ลิฟท์ตัวหนึ่งมวล 100 กิโลกรัม ขณะที่บรรทุกคนให้เคลื่อนที่ขึ้นด้วยความเร็วคงที่ 3 เมตร/วินาที โดยใช้มอเตอร์ 9 กิโลวัตต์ จะต้องใช้แรงเท่าใด (นำไปใช้) ก. 3 นิวตัน ข. 30 นิวตัน ค. 300 นิวตัน ง. 3,000 นิวตัน คำตอบ คือ ง	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
		9) มอเตอร์ไซต์มวล 300 กิโลกรัม แล่นขึ้นพื้นเอียงทำมุม 30 องศากับแนวระดับ ด้วยอัตราเร็วคงที่ 120 เมตร/วินาที ถ้าแรงเสียดทานระหว่างพื้นเอียงกับมอเตอร์ไซต์เท่ากับ 200 นิวตัน จงหากำลังงานของเครื่องยนต์ในหน่วยกิโลวัตต์ (วิเคราะห์) คำตอบ คือ 204 กิโลวัตต์	1	-1	1	1	1	0.8	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
3) พลังงานและความสัมพันธ์ระหว่างงานและพลังงาน	4) อธิบายพลังงาน พลังงานจลน์ พลังงานศักย์ และความสัมพันธ์ระหว่างงานและพลังงานได้	10) พลังงานจลน์ของวัตถุในกรณีใดมีค่าต่ำที่สุด (เข้าใจ) ก. เมื่อเพิ่มมวลเป็น 2 เท่า ข. เมื่อเพิ่มความเร็วเป็น 2 เท่า ค. เมื่อลดมวลลง 1/2 เท่า ง. เมื่อลดความเร็วลง 1/2 เท่า คำตอบ คือ ข้อ ง	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม

ตาราง 39 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับอ้างอิง (ฉบับที่ 2) (ต่อ)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
3) พลังงานและความสัมพันธ์ระหว่างงานและพลังงาน	4) อธิบายพลังงานจลน์ พลังงานศักย์ และความสัมพันธ์ระหว่างงานและพลังงานได้	11) กราฟในตัวเลือกใดแสดงความสัมพันธ์ระหว่างพลังงานจลน์ของวัตถุ (E_k) กับการกระจัด (S) ได้ถูกต้อง เมื่อวัตถุนี้เคลื่อนที่จากจุดหยุดนิ่งออกไปในแนวเส้นตรงด้วยอัตราเร็วคงที่ (เข้าใจ) ก.  ข.  ค.  ง.  คำตอบ คือ ข้อ ค	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม โดยผู้เชี่ยวชาญแนะนำให้ปรับตัวเลือก ค ให้สอดคล้องกับโจทย์
	5) คำนวณหาพลังงานจลน์ พลังงานศักย์โน้มถ่วงของวัตถุที่ระดับต่างๆ และพลังงานศักย์ยืดหยุ่นของสปริง	12) ท็อปปล่อยวัตถุมวล 3 กิโลกรัม จากตึกสูง 45 เมตรจากระดับพื้นดิน เมื่อกล่องหล่นมาได้ระยะทาง 25 เมตร พลังงานศักย์มีค่าเท่าใด(นำไปใช้) ก. 450 จูล ข. 600 จูล ค. 750 จูล ง. 1,350 จูล คำตอบ คือ ข้อ ข	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม แต่ควรปรับข้อ ก โดยควรเป็น 2,100
		13) กันต์ปล่อยวัตถุมวล 0.5 กิโลกรัมจากตึกแห่งหนึ่ง จงหาพลังงานจลน์ของวัตถุ เมื่อวัตถุถูกปล่อยลงมาเป็นเวลา 2 วินาที (นำไปใช้) ก. 5 นิวตัน ข. 10 นิวตัน ค. 100 นิวตัน ง. 200 นิวตัน คำตอบ คือ ข้อ ค	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
		14) สปริงถูกดึงด้วยแรง 40 นิวตัน จะทำให้สปริงยืดออก 20 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 30 เซนติเมตรในหน่วยจูล (วิเคราะห์) คำตอบ คือ 9 จูล	1	-1	1	-1	1	0.6	วัดได้ตรงตามวัตถุประสงค์ โดยผู้เชี่ยวชาญแนะนำให้เพิ่มความซับซ้อนของโจทย์

ตาราง 39 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับอ้างอิง (ฉบับที่ 2) (ต่อ)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
3) พลังงานและความสัมพันธ์ระหว่างงานและพลังงาน	6) คำนวณโจทย์เกี่ยวกับความสัมพันธ์ระหว่างงานและพลังงานได้	15) รถยนต์มวล 1,000 กิโลกรัม เคลื่อนที่ไปตามพื้นราบ งานที่เกิดขึ้นในช่วง 2-6 วินาทีมีค่ากี่จูล ถ้ากราฟระหว่างความเร็ว และเวลาของรถยนต์ขณะเคลื่อนที่เป็นดังรูป (นำไปใช้)  ก. 140 จูล ข. 280 จูล ค. 14,000 จูล ง. 28,000 จูล คำตอบ คือ ข้อ ค	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
		16) ยิงลูกปืนมวล 100 กรัมไปในแนวระดับ ลูกปืนมีความเร็ว 200 เมตร/วินาที ถ้าลูกปืนฝังเข้าไปฝังในเนื้อไม้ งานที่ไม้ต้านการเคลื่อนที่ของลูกปืนมีค่าเท่าใด (นำไปใช้) ก. 2×10^3 จูล ข. 4×10^3 จูล ค. 2×10^6 จูล ง. 4×10^3 จูล คำตอบ คือ ข้อ ก	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม ควรปรับตัวเลือก ง เป็น 4×10^6 จูล
		17) เครื่องสูบน้ำใช้สูบน้ำมวล 500 กิโลกรัม ขึ้นจากบ่อลึก 6 เมตร แล้วฉีดน้ำออกจากปลายท่อด้วยความเร็ว 6 เมตรต่อวินาที งานที่ได้จากเครื่องสูบน้ำนี้มีค่ากี่กิโลจูล (วิเคราะห์) คำตอบ คือ 39 กิโลจูล	1	-1	1	-1	1	0.6	วัดได้ตรงตามวัตถุประสงค์แต่ควรเพิ่มความซับซ้อนของโจทย์

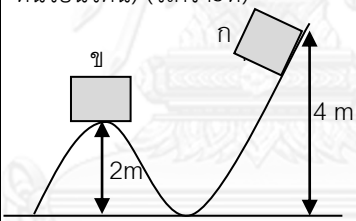
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 39 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับอ้างอิง (ฉบับที่ 2) (ต่อ)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
4) กฎการอนุรักษ์พลังงานกล	7) อธิบายกฎการอนุรักษ์พลังงานกลในการวิเคราะห์การเคลื่อนที่ในสถานการณ์ต่าง ๆ	18) เมื่อกล่องใบหนึ่งถูกปล่อยจากตึกสูงมายังพื้นราบ ข้อใดกล่าวถูกต้อง (เข้าใจ) ก. พลังงานรวม พลังงานจลน์ พลังงานศักย์โน้มถ่วง ณ จุดต่างๆ ของการเคลื่อนที่มีค่าเท่ากัน ข. พลังงานศักย์โน้มถ่วงของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน ค. พลังงานจลน์ของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน ง. พลังงานรวมของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน คำตอบ คือ ข้อ ง	0	1	1	1	1	0.8	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
	8) ประยุกต์ใช้กฎการอนุรักษ์พลังงานกับการเคลื่อนในสถานการณ์ต่าง ๆ ได้	19) ตั้มปล่อยถุงทราย 5 กิโลกรัม ตกจากหอคอยสูง 7 เมตร ขณะที่ถุงทรายเคลื่อนที่ลงมาได้ระยะทาง 4 เมตรจากหอคอย ความเร็วของถุงทรายมีค่าเท่าใด (นำไปใช้) ก. $\sqrt{20}$ เมตร/วินาที ข. 20 เมตร/วินาที ค. $\sqrt{60}$ เมตร/วินาที ง. 60 เมตร/วินาที คำตอบ คือ ข้อ ค	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
		20) วัตถุมวล 3.6 กิโลกรัม เคลื่อนที่ด้วยความเร็ว 20 เมตร/วินาที เข้าชนสปริงที่มีค่าคงที่ 400 นิวตัน/เมตร สปริงจะหดสั้นที่สุดเท่าไหร่เมื่อกล่องชนสปริง (นำไปใช้) ก. 0.6 ซม. ข. 6 ซม. ค. 60 ซม. ง. 600 ซม. คำตอบ คือ ข้อ ง	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม แต่ควรตรวจสอบตัวเลขในโจทย์และตัวเลือก

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 39 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับแบบสอบเรื่องงานและพลังงาน ฉบับอ้างอิง (ฉบับที่ 2) (ต่อ)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	ค่า IOC ของผู้เชี่ยวชาญ (คนที่)						ความหมาย
			1	2	3	4	5	IOC	
4) กฎการอนุรักษ์พลังงานกล	8) ประยุกต์ใช้กฎการอนุรักษ์พลังงานกับการเคลื่อนในสถานการณ์ต่าง ๆ ได้	21) ถ้าแทนต้องการใช้สปริงในการยิงวัตถุมวล 200 กรัมให้ขึ้นไปในแนวดิ่งเป็นระยะทาง 8 เมตร ด้วยการยืดสปริงออก 40 เซนติเมตร สปริงนี้จะต้องมีค่านิจสปริงกี่นิวตัน/เมตร (นำไปใช้) คำตอบ คือ 200 นิวตัน/เมตร	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
		22) กล่องมวล 6 กิโลกรัม ที่จุด ก มีอัตราเร็ว 1 เมตร/วินาที และที่จุด ข มีอัตราเร็ว 7 เมตร/วินาที ถ้าระยะทางจากจุด ก ไปยังจุด ข เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (วิเคราะห์)  คำตอบ คือ -8 นิวตัน	1	-1	1	1	1	0.8	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม
5) เครื่องกล	9) คำนวณโจทย์เกี่ยวกับเครื่องกลโดยอาศัยหลักการทำงานของเครื่องกลได้	23) ถ้าเราให้กำลังแก่มอเตอร์ 0.3 กำลังม้า มอเตอร์จะให้พลังงานงานที่เป็นประโยชน์ออกมาเท่าใด ถ้ามอเตอร์เครื่องนี้มีประสิทธิภาพ 70% (1 กำลังม้า = 750 วัตต์) (นำไปใช้) ก. 15.75 วัตต์ ข. 67.5 วัตต์ ค. 157.5 วัตต์ ง. 675.5 วัตต์ คำตอบ คือ ข้อ ค	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม แต่ควรตรวจสอบค่ากำลังม้า
		24) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัม ขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (นำไปใช้) คำตอบ คือ 80%	1	1	1	1	1	1	วัดได้ตรงตามวัตถุประสงค์และระดับพฤติกรรม

ตาราง 40 ผลการปรับปรุงข้อสอบวิชาฟิสิกส์ เรื่อง งานและพลังงาน ระดับชั้นมัธยมศึกษาปีที่ 4 (ฉบับที่2) ตามคำแนะนำของผู้เชี่ยวชาญ

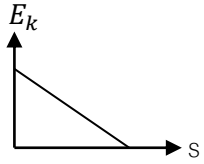
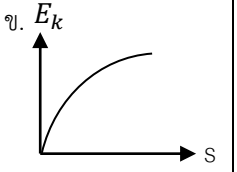
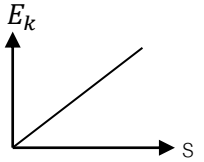
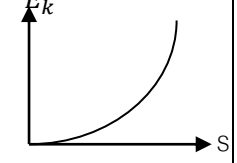
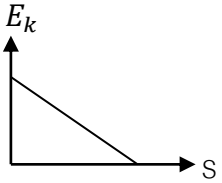
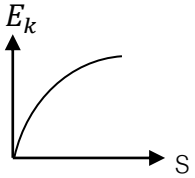
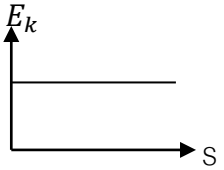
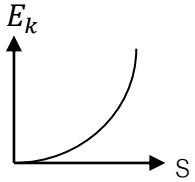
จุดประสงค์ที่ 1 อธิบายงานและวิเคราะห์งานของแรงต่างๆได้

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
1) กล่องเคลื่อนที่เป็นเส้นตรงบนพื้นราบลื่นด้วยแรง 20 N เป็นระยะทาง S ถ้าต้องการหาค่างานของแรง 10 N จะต้องพิจารณาสิ่งใดบ้าง (เข้าใจ) 1. มุมระหว่างทิศของแรงกับแนวการเคลื่อนที่ของวัตถุ 2. แรงเสียดทานระหว่างวัตถุกับพื้นราบ 3. อัตราเร็วของวัตถุ ก. 1 ข. 2 ค. 3 ง. 2 และ 3	1) กล่องเคลื่อนที่เป็นเส้นตรงบนพื้นราบลื่นด้วยแรง 20 นิวตัน เป็นระยะทาง S ถ้าต้องการหาค่างานของแรง 10 นิวตัน จะต้องพิจารณาสิ่งใดบ้าง 1. มุมระหว่างทิศของแรงกับแนวการเคลื่อนที่ของวัตถุ 2. แรงเสียดทานระหว่างวัตถุกับพื้นราบ 3. อัตราเร็วของวัตถุ 1. ก 2. ข 3. ค 4. ข และ ค

จุดประสงค์ที่ 3 คำนวณหากำลังจากโจทย์ที่กำหนดให้ได้

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
7) จ็อบออกแรงคงที่ขนาด 200 นิวตัน ผลักกล่องใบหนึ่งจากหยุดนิ่งให้เคลื่อนที่ไปตามแนวราบได้ระยะทาง 2 เมตร ภายในเวลา 4 วินาที จงหากำลังของจ็อบที่ใช้ในการผลักกล่องใบนี้ ก. 100 วัตต์ ข. 400 วัตต์ ค. 500 วัตต์ ง. 600 วัตต์	7) จ็อบออกแรงคงที่ขนาด 200 นิวตัน ผลักกล่องใบหนึ่งจากหยุดนิ่งให้เคลื่อนที่ไปตามแนวราบได้ระยะทาง 2 เมตร ภายในเวลา 4 วินาที จงหากำลังของจ็อบที่ใช้ในการผลักกล่องใบนี้ ก. 100 วัตต์ ข. 400 วัตต์ ค. 1,200 วัตต์ ง. 1,600 วัตต์

จุดประสงค์ที่ 4 อธิบายพลังงาน พลังงานจลน์ พลังงานศักย์ และความสัมพันธ์ระหว่างงานและพลังงานได้

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
11) กราฟในตัวเลือกใดแสดงความสัมพันธ์ระหว่างพลังงานจลน์ของวัตถุ (E_k) กับการกระจัด (S) ได้ถูกต้อง เมื่อวัตถุนี้เคลื่อนที่จากจุดหยุดนิ่งออกไปในแนวเส้นตรงด้วยอัตราเร็วคงที่ ก.  ข.  ค.  ง. 	11) กราฟในตัวเลือกใดแสดงความสัมพันธ์ระหว่างพลังงานจลน์ของวัตถุ (E_k) กับการกระจัด (S) ได้ถูกต้อง เมื่อวัตถุนี้เคลื่อนที่จากจุดหยุดนิ่งออกไปในแนวเส้นตรงด้วยอัตราเร็วคงที่ 1.  2.  3.  4. 

ตาราง 40 ผลการปรับปรุงข้อสอบวิชาฟิสิกส์ เรื่อง งานและพลังงาน ระดับชั้นมัธยมศึกษาปีที่ 4 (ฉบับที่2) ตามคำแนะนำของผู้เชี่ยวชาญ (ต่อ)

จุดประสงค์ที่ 5 คำนวณหาพลังงานจลน์ พลังงานศักย์โน้มถ่วงของวัตถุที่ระดับต่างๆ และพลังงานศักย์ยืดหยุ่นของสปริง

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
12) ท็อปปล่อยวัตถุมวล 3 กิโลกรัม จากตึกสูง 45 เมตร จากระดับพื้นดิน เมื่อกล่องหล่นมาได้ระยะทาง 25 เมตร พลังงานศักย์มีค่าเท่าใด ก. 450 จูล ข. 600 จูล ค. 750 จูล ง. 1,350 จูล	12) ท็อปปล่อยวัตถุมวล 3 กิโลกรัม จากตึกสูง 45 เมตรจากระดับพื้นดิน เมื่อกล่องหล่นมาได้ระยะทาง 25 เมตร พลังงานศักย์มีค่าเท่าใด 1. 600 จูล ข. 750 จูล ค. 1,350 จูล ง. 2,100 จูล
14) สปริงถูกดึงด้วยแรง 40 นิวตัน จะทำให้สปริงยืดออก 20 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 30 เซนติเมตรในหน่วยจูล	14) ออกแรงดึงสปริง 30 นิวตัน ปรากฏว่าสปริงยาว 50 เซนติเมตร แต่ถ้าออกแรงดึงสปริง 40 นิวตัน สปริงจะมีความยาว 60 เซนติเมตร พลังงานศักย์ยืดหยุ่นขณะที่ออกแรงดึงสปริง 40 นิวตัน มีค่าเท่าใด

จุดประสงค์ที่ 6 คำนวณโจทย์เกี่ยวกับความสัมพันธ์ระหว่างงานและพลังงานได้

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
16) ยิงลูกปืนมวล 100 กรัมไปในแนวระดับ ลูกปืนมีความเร็ว 200 เมตร/วินาที ถ้าลูกปืนฝังเข้าไปฝังในเนื้อไม้ งานที่ไม้ต้านการเคลื่อนที่ของลูกปืนมีค่าเท่าใด (นำไปใช้) ก. 2×10^3 จูล ข. 4×10^3 จูล ค. 2×10^6 จูล ง. 4×10^3 จูล	16) ยิงลูกปืนมวล 100 กรัมไปในแนวระดับ ลูกปืนมีความเร็ว 200 เมตร/วินาที ถ้าลูกปืนฝังเข้าไปฝังในเนื้อไม้ งานที่ไม้ต้านการเคลื่อนที่ของลูกปืนมีค่าเท่าใด (นำไปใช้) 1. 2×10^3 จูล 2. 4×10^3 จูล 3. 2×10^6 จูล 4. 4×10^6 จูล
17) เครื่องสูบน้ำใช้สูบน้ำมวล 500 กิโลกรัม ขึ้นจากบ่อลึก 6 เมตร แล้วฉีดน้ำออกจากปลายท่อด้วยความเร็ว 6 เมตรต่อวินาที งานที่ได้จากเครื่องสูบน้ำนี้มีค่ากี่กิโลจูล	17) แมนเข็นลังใส่ของมวล 40 กิโลกรัม ให้เคลื่อนที่บนพื้นลื่นในแนวเส้นตรงด้วยแรงคงที่ค่าหนึ่ง ซึ่งเมื่อผ่านตำแหน่ง A ลังดังกล่าวมีความเร็ว 4 เมตร/วินาที งานที่แมนต้องใช้ในการเข็นลังนี้จากตำแหน่ง A ไปยังตำแหน่ง B มีค่ากี่จูล ถ้าขณะที่ผ่านตำแหน่ง B เคนต้องใช้กำลังเป็น $\frac{5}{4}$ เท่าของกำลังที่ใช้เมื่อผ่านตำแหน่ง A

จุดประสงค์ที่ 8 ประยุกต์ใช้กฎการอนุรักษ์พลังงานกับการเคลื่อนในสถานการณ์ต่าง ๆ ได้

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
20) วัตถุมวล 3.6 กิโลกรัม เคลื่อนที่ด้วยความเร็ว 20 เมตร/วินาที เข้าชนสปริงที่มีค่าคงที่ 400 นิวตัน/เมตร สปริงจะหดสั้นที่สุดเท่าไหร่เมื่อกล่องชนสปริง ก. 0.6 ซม. ข. 6 ซม. ค. 60 ซม. ง. 600 ซม.	20) วัตถุมวล 36 กิโลกรัม เคลื่อนที่ด้วยความเร็ว 20 เมตร/วินาที เข้าชนสปริงที่มีค่าคงที่ 400 นิวตัน/เมตร สปริงจะหดสั้นที่สุดเท่าไหร่เมื่อกล่องชนสปริง 1. 6 ซม. 2. 36 ซม. 3. 600 ซม. 4. 3,600 ซม.

ตาราง 40 ผลการปรับปรุงข้อสอบวิชาฟิสิกส์ เรื่อง งานและพลังงาน ระดับชั้นมัธยมศึกษาปีที่ 4 (ฉบับที่2) ตามคำแนะนำของผู้เชี่ยวชาญ (ต่อ)

<u>จุดประสงค์ที่ 9</u> คำนวณโจทย์เกี่ยวกับเครื่องกลโดยอาศัยหลักการทำงานของเครื่องกลได้

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
23) ถ้าเราให้กำลังแก่มอเตอร์ 0.3 กำลังม้า มอเตอร์จะให้พลังงานงานที่เป็นประโยชน์ออกมาเท่าใด ถ้ามอเตอร์เครื่องนี้มีประสิทธิภาพ 70% (1 กำลังม้า = 750 วัตต์) (นำไปใช้) ก. 15.75 วัตต์ ข. 67.5 วัตต์ ค. 157.5 วัตต์ ง. 675.5 วัตต์	23) ถ้าเราให้กำลังแก่มอเตอร์ 0.3 กำลังม้า มอเตอร์จะให้พลังงานงานที่เป็นประโยชน์ออกมาเท่าใด ถ้ามอเตอร์เครื่องนี้มีประสิทธิภาพ 70% (1 กำลังม้า = 746 วัตต์) 1. 0.21 วัตต์ 2. 15.67 วัตต์ 3. 21.0 วัตต์ 4. 156.7 วัตต์

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ค
แบบสอบเรื่องงานและพลังงาน รายวิชาฟิสิกส์ สำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 ฉบับใหม่ (ชุด 1) และฉบับอ้างอิง (ชุด 2) สำหรับการทดลองใช้

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

แบบสอบ เรื่องงานและพลังงาน (ชุด 1)

การประเมินผลระหว่างเรียน ระดับชั้นมัธยมศึกษาปีที่ 4 ปีการศึกษา 2556
แบบสอบรายวิชา วิทยาศาสตร์ คะแนนเต็ม 48 คะแนน เวลา 90 นาที

.....

คำชี้แจง 1. แบบสอบฉบับนี้มีทั้งหมด 7 หน้า จำนวน 2 ตอน
ตอนที่ 1 เป็นข้อสอบปรนัย จำนวน 16 ข้อ รวม 16 คะแนน
ตอนที่ 2 เป็นข้อสอบอัตนัย จำนวน 8 ข้อ รวม 32 คะแนน
2. สามารถทดเลขลงในแบบสอบได้
3. ส่งกระดาษคำตอบคืนพร้อมกับแบบสอบ

กำหนดให้ g = 10 เมตร/วินาที^{2}

<u>ตอนที่ 1</u> ให้นักเรียนเลือกคำตอบที่ถูกต้องที่สุด แล้วทำเครื่องหมาย X ลงในกระดาษคำตอบ

1) แรง 10 นิวตัน ทำให้วัตถุเคลื่อนที่เป็นเส้นตรงบนพื้นราบลื่นในระยะทาง S ถ้าต้องการหาค่างานของแรง 10 นิวตัน จะต้องพิจารณาสิ่งใดบ้าง
ก. อัตราเร็วของวัตถุ
ข. แรงเสียดทานระหว่างวัตถุกับผิวสัมผัส
ค. มุมระหว่างทิศของแรงกับแนวการเคลื่อนที่ของวัตถุ
1. ก 2. ข 3. ค 4. ข และ ค

2) กรณีใดที่ถือว่ามีการทำงาน เนื่องจากการออกแรงกระทำต่อหนังสือ
ก. แจนยืนถือหนังสือ ข. แจนนั่งอ่านหนังสือเตรียมสอบ
ค. แจนยกหนังสือจากพื้นไปวางบนโต๊ะในแนวดิ่ง
ง. แจนถือหนังสือแล้วเดินไปตามพื้นราบ
1. ค 2. ค และ ง 3. ก ค และ ง 4. ก ข ค และ ง

3) มินออกแรง 100 นิวตันลากกล่องในแนวทำมุม 60 องศากับแนวระดับ ถ้าลากกล่องไปได้ไกล 10 เมตรในแนวระดับ งานของแรงที่ใช้ลากกล่องเป็นเท่าใด
1. 500 จูล 2. $500\sqrt{3}$ จูล 3. 1,000 จูล 4. $1{,}000\sqrt{3}$ จูล

4) มวล 40 กิโลกรัมเคลื่อนที่ลงไปตามพื้นเอียงผิวขรุขระเป็นระยะทาง 10 เมตร ด้วยความเร็วคงที่ ถ้างานในการเคลื่อนที่ของลังดังกล่าวเป็น 3,600 จูล แรงเสียดทานระหว่างพื้นและลังจะมีค่าเท่าใด ถ้าพื้นเอียงอยู่สูงจากพื้น 6 เมตร
1. - 40 นิวตัน 2. 40 นิวตัน 3. -120 นิวตัน 4.120 นิวตัน

5) นาย ก. ออกแรงคงที่ขนาด 100 นิวตัน ผลักวัตถุก้อนหนึ่งจากหยุดนิ่งให้เคลื่อนที่ไปตามแนวราบได้ระยะทาง 3 เมตร ภายในเวลา 2 วินาที จงหากำลังของนาย ก. ที่ใช้ในการผลักวัตถุนี้
1. 17 วัตต์ 2. 150 วัตต์ 3. 300 วัตต์ 4. 600 วัตต์

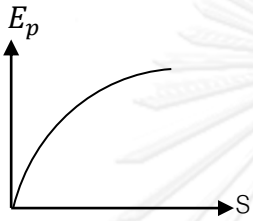
6) ลิฟท์ตัวหนึ่งมวล 200 กิโลกรัม ใช้มอเตอร์ 8 กิโลวัตต์ ขณะที่บรรทุกคนให้เคลื่อนที่ขึ้นด้วยความเร็วคงที่ 2 เมตร/วินาที จะต้องใช้แรงเท่าใด

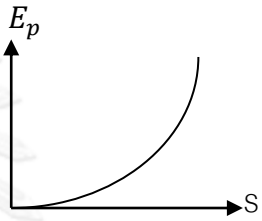
1. 4 นิวตัน 2. 25 นิวตัน 3. 100 นิวตัน 4. 4,000 นิวตัน

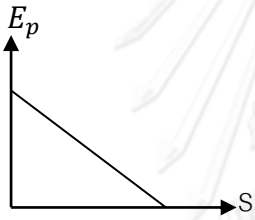
7) พลังงานจลน์ของวัตถุในกรณีใดมีค่ามากที่สุด

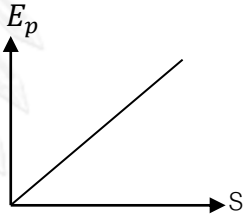
1. เมื่อเพิ่มความเร็วเป็น 2 เท่า
2. เมื่อเพิ่มมวลเป็น 2 เท่า
3. เมื่อลดความเร็วลง 1/2 เท่า
4. เมื่อลดมวลลง 1/2 เท่า

8) กราฟในตัวเลือกใดแสดงความสัมพันธ์ระหว่างพลังงานศักย์โน้มถ่วง (E_p) ของวัตถุกับการกระจัด (S) ได้ถูกต้อง เมื่อวัตถุนี้ถูกปล่อยจากจุดหยุดนิ่งลงมาในแนวดิ่งด้วยอัตราเร็วคงที่

1.


2.


3.


4.


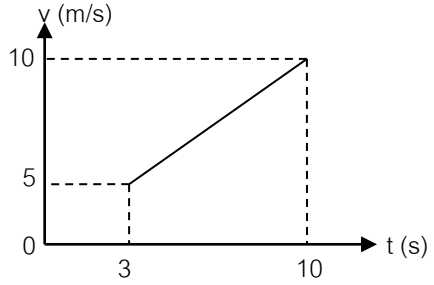
9) กล่องมวล 2 กิโลกรัม ตกจากหน้าผาสูง 40 เมตรจากระดับพื้นดิน เมื่อกล่องหล่นมาได้ระยะทาง 15 เมตร พลังงานศักย์มีค่าเท่าใด

1. 300 จูล 2. 500 จูล 3. 800 จูล 4. 1,100 จูล

10) ลูกมะพร้าวมวล 1 กิโลกรัม อยู่บนต้นมะพร้าว จงหาพลังงานจลน์ของลูกมะพร้าวเมื่อหลุดออกจากต้นมะพร้าวที่เวลา 1 วินาที

1. 5 จูล 2. 10 จูล 3. 20 จูล 4. 50 จูล

11) รถยนต์มวล 2,000 กิโลกรัม เคลื่อนที่ไปตามพื้นราบงานที่เกิดขึ้นในช่วง 3-10 วินาทีมีค่ากี่จูล ถ้ากราฟระหว่างความเร็ว และเวลาของรถยนต์ขณะเคลื่อนที่เป็นดังรูป



1. 75 จูล 2. 750 จูล 3. 7,500 จูล 4. 75,000 จูล

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

12) ลูกปืนมวล 150 กรัมไปในแนวระดับ ลูกปืนมีความเร็ว 100 เมตร/วินาที ถ้าลูกปืนฝังเข้าไปฝังในเนื้อไม้ งานที่ไม้ต้านการเคลื่อนที่ของลูกปืนมีค่าเท่าใด

1. 150 จูล 2. 750 จูล 3. 15,000 จูล 4. 75,000 จูล

13) เมื่อวัตถุก้อนหนึ่งเคลื่อนที่ลงจากพื้นเอียงลื่นมายังพื้นราบลื่น ข้อใดกล่าวถูกต้อง

1. พลังงานรวมของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน
2. พลังงานจลน์ของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน
3. พลังงานศักย์โน้มถ่วงของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน
4. พลังงานรวม พลังงานจลน์ พลังงานศักย์โน้มถ่วง ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน

14) ลูกตาลมวล 2 กิโลกรัม ตกจากต้นตาลสูง 6 เมตร ขณะที่ลูกตาลเคลื่อนที่ลงมาได้ระยะทาง 2 เมตรจากต้นตาล ความเร็วของลูกตาลมีค่าเท่าใด

1. $\sqrt{40}$ เมตร/วินาที 2. $\sqrt{80}$ เมตร/วินาที 3. 40 เมตร/วินาที 4. 80 เมตร/วินาที

15) กล่องมวล 25 กิโลกรัม เคลื่อนที่ด้วยความเร็ว 10 เมตร/วินาที เข้าชนสปริงที่มีค่าคงที่ 100 นิวตัน/เมตร เมื่อกล่องชนสปริง สปริงจะหดสั้นที่สุดเท่าไหร่

1. 5 ซม. 2. 25 ซม. 3. 500 ซม. 4. 2,500 ซม.

16) มอเตอร์เครื่องหนึ่งที่มีประสิทธิภาพ 90% จะให้พลังงานงานที่เป็นประโยชน์ออกมาเท่าใด เมื่อใส่กำลังให้แก่มอเตอร์เครื่องนี้ 0.5 กำลังม้า (1 กำลังม้า = 746 วัตต์)

1. 0.45 วัตต์ 2. 33.57 วัตต์ 3. 45.00 วัตต์ 4. 335.70 วัตต์

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ชื่อ-นามสกุล .....ชั้น ม.4/..... เลขที่ ..... ชุดที่ 1

กระดาษคำตอบ

ตอนที่ 1 (แบบสอบปรนัยจำนวน 16 ข้อ รวม 16 คะแนน)

คำสั่ง ให้นักเรียนเลือกคำตอบที่ถูกต้องที่สุด แล้วทำเครื่องหมาย X ลงในกระดาษคำตอบ

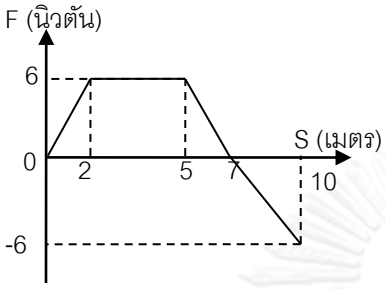
ข้อ	ตัวเลือก			
	1	2	3	4
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ชื่อ-นามสกุล .....ชั้น ม.4/..... เลขที่ ..... ชุดที่ 1

ตอนที่ 2 จงแสดงวิธีทำให้ถูกต้อง โดยแสดงวิธีทำอย่างละเอียด **กำหนดให้** g = 10 เมตร/วินาที^{2}

1) F (นิวตัน)



งานทั้งหมดที่กระทำในช่วงระยะทางการเคลื่อนที่ 0 เมตรถึง 10 เมตร จะมีค่ากี่จูล ถ้ากราฟระหว่างแรงกับการกระจัดเป็นดังรูป (4 คะแนน)

.....

.....

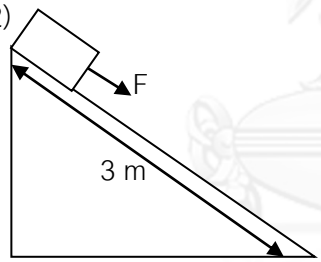
.....

.....

.....

.....

2)



ช้างตัวหนึ่งออกแรงดึงซุงหนัก 140 กิโลกรัม ลงเนินซึ่งเอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ สัมประสิทธิ์ความเสียดทานระหว่างเนินและซุงเป็น $\sqrt{3}/2$ ถ้าแรงที่ใช้ดึงซุงขนานกับการเคลื่อนที่ดังรูป ช้างต้องใช้งานเท่าใดในการดึงซุงลงมาได้ 3 เมตร (4 คะแนน)

.....

.....

.....

.....

.....

.....

.....

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3) รถยนต์คันหนึ่งมีมวล 400 กิโลกรัม แล่นขึ้นพื้นเอียงทำมุม 30 องศากับแนวระดับ ด้วยอัตราเร็วคงที่ 90 เมตร/วินาที จงหากำลังงานของเครื่องยนต์ในหน่วยกิโลวัตต์ ถ้าแรงเสียดทานระหว่างพื้นเอียงกับรถยนต์เท่ากับ 100 นิวตัน (4 คะแนน)

.....

.....

.....

.....

.....

.....

4) ออกแรงดึงสปริง 50 นิวตัน ปรากฏว่าสปริงยาว 70 เซนติเมตร แต่ถ้าออกแรงดึงสปริง 60 นิวตัน สปริงจะมีความยาว 80 เซนติเมตร พลังงานศักย์ยืดหยุ่นขณะที่ออกแรงดึงสปริง 60 นิวตันมีค่าเท่าใด (4 คะแนน)

.....

.....

.....

.....

.....

.....

5) เคนเข็นลังใส่ผลไม้มวล 50 กิโลกรัม ให้เคลื่อนที่ในแนวเส้นตรงด้วยแรงคงที่ค่าหนึ่งไปบนพื้นลื่น เมื่อผ่านตำแหน่ง ก เคนต้องใช้กำลังเป็น $\frac{2}{3}$ เท่าของกำลังที่ใช้เมื่อผ่านตำแหน่ง ข ถ้าความเร็วของลังนี้เมื่อผ่านตำแหน่ง ข คือ 3 เมตร/วินาที งานที่ใช้ในการเข็นลังนี้จากตำแหน่ง ก ไปยังตำแหน่ง ข มีค่ากี่จูล (4 คะแนน)

.....

.....

.....

.....

.....

.....

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

6) ดาวต้องการยิงลูกหินมวล 100 กรัม ด้วยหนังสติ๊กให้ขึ้นไปได้สูง 10 เมตรในแนวดิ่ง ถ้าการยืดของหนังสติ๊กเป็นแบบสปริงและยืดออก 50 เซนติเมตร หนังสติ๊กจะต้องมีค่านิจสปริงกี่นิวตัน/เมตร (4 คะแนน)

.....

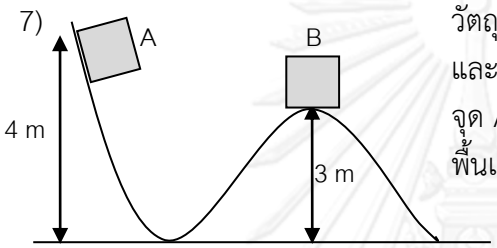
.....

.....

.....

.....

.....

7) 

วัตถุมวล 8 กิโลกรัม ที่จุด A มีอัตราเร็ว 0.5 เมตร/วินาที และที่จุด B มีอัตราเร็ว 9 เมตร/วินาที ถ้าระยะทางจากจุด A ไปยังจุด B เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (4 คะแนน)

.....

.....

.....

.....

.....

.....

8) ลิฟท์อันหนึ่งมวล 650 กิโลกรัม สามารถยกของมวล 350 กิโลกรัม ขึ้นไปได้สูง 15 เมตร ในเวลา 5 วินาที โดยใช้กำลังไฟฟ้าไป 40 กิโลวัตต์ ลิฟท์นี้ต้องมีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)

Chulalongkorn University

.....

.....

.....

.....

.....

.....

แบบสอบ เรื่องงานและพลังงาน (ชุด 2)

การประเมินผลระหว่างเรียน ระดับชั้นมัธยมศึกษาปีที่ 4 ปีการศึกษา 2556
แบบสอบรายวิชา วิทยาศาสตร์ คะแนนเต็ม 48 คะแนน เวลา 90 นาที

คำชี้แจง 1. แบบสอบฉบับนี้มีทั้งหมด 7 หน้า จำนวน 2 ตอน
ตอนที่ 1 เป็นข้อสอบปรนัย จำนวน 16 ข้อ รวม 16 คะแนน
ตอนที่ 2 เป็นข้อสอบอัตนัย จำนวน 8 ข้อ รวม 32 คะแนน
2. สามารถทดเลขลงในแบบสอบได้
3. ส่งกระดาษคำตอบคืนพร้อมกับแบบสอบ

กำหนดให้ g = 10 เมตร/วินาที^{2}

<u>ตอนที่ 1</u> ให้นักเรียนเลือกคำตอบที่ถูกต้องที่สุด แล้วทำเครื่องหมาย X ลงในกระดาษคำตอบ

1) กล่องเคลื่อนที่เป็นเส้นตรงบนพื้นราบลื่นด้วยแรง 20 นิวตัน เป็นระยะทาง S ถ้าต้องการหาค่างานของแรง 20 นิวตัน จะต้องพิจารณาสิ่งใดบ้าง
ก. มุมระหว่างทิศของแรงกับแนวการเคลื่อนที่ของวัตถุ
ข. แรงเสียดทานระหว่างวัตถุกับพื้นราบ
ค. อัตราเร็วของวัตถุ
1. ก 2. ข 3. ค 4. ข และ ค

2) กรณีใดที่ถือว่ามีการทำงาน เนื่องจากการออกแรงกระทำต่อกล่อง
ก. นุชยืนถือกล่อง ข. นุชยกกล่องจากโต๊ะไปวางบนพื้นในแนวดิ่ง
ค. นุชยกกล่องจากพื้นไปวางบนโต๊ะในแนวดิ่ง
ง. นุชถือกล่องแล้วเดินไปตามพื้นราบ
1. ง 2. ข และ ค 3. ข ค และ ง 4. ก ข ค และ ง

3) นันลากกล่องไปตามพื้นราบเป็นระยะ 20 เมตร ด้วยแรง 70 นิวตันในแนวทำมุม 60 องศากับแนวระดับ ถ้า งานของแรงที่ใช้ลากกล่องเป็นเท่าใด
1. 700 จูล 2. $700\sqrt{3}$ จูล 3. 1400 จูล 4. $1400\sqrt{3}$ จูล

4) กล่องมวล 20 กิโลกรัมเคลื่อนที่ลงไปตามพื้นเอียงผิวขรุขระด้วยความเร็วคงที่ เป็นระยะทาง 5 เมตร แรงเสียดทานระหว่างพื้นและลังจะมีค่าเท่าใด เมื่องานในการเคลื่อนที่ของลังดังกล่าวเป็น 1,400 จูล ถ้าพื้นเอียงอยู่สูงจากพื้น 3 เมตร
1. - 80 นิวตัน 2. -160 นิวตัน 3. 80 นิวตัน 4.160 นิวตัน

5) จ็อบออกแรงคงที่ขนาด 200 นิวตัน ผลักกล่องใบหนึ่งจากหยุดนิ่งให้เคลื่อนที่ไปตามแนวราบได้ระยะทาง 2 เมตร ภายในเวลา 4 วินาที จงหากำลังของจ็อบที่ใช้ในการผลักกล่องใบนี้
1. 100 วัตต์ 2. 400 วัตต์ 3. 1,200 วัตต์ 4.1,600 วัตต์

CHULALONGKORN UNIVERSITY

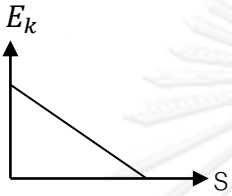
6) ลิฟท์ตัวหนึ่งมวล 100 กิโลกรัม ขณะที่บรรทุกคนให้เคลื่อนที่ขึ้นด้วยความเร็วคงที่ 3 เมตร/วินาที โดยใช้มอเตอร์ 9 กิโลวัตต์ จะต้องใช้แรงเท่าใด

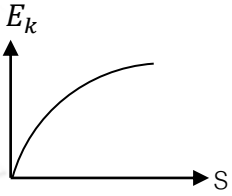
1. 3 นิวตัน 2. 30 นิวตัน 3. 300 นิวตัน 4. 3,000 นิวตัน

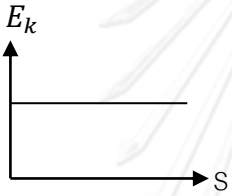
7) พลังงานจลน์ของวัตถุในกรณีใดมีต่ำที่สุด

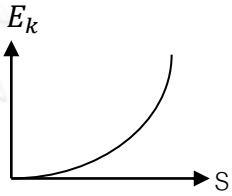
1. เมื่อเพิ่มมวลเป็น 2 เท่า
2. เมื่อเพิ่มความเร็วเป็น 2 เท่า
3. เมื่อลดมวลลง 1/2 เท่า
4. เมื่อลดความเร็วลง 1/2 เท่า

8) กราฟในตัวเลือกใดแสดงความสัมพันธ์ระหว่างพลังงานจลน์ของวัตถุ (E_k) กับการกระจัด (S) ได้ถูกต้อง เมื่อวัตถุนี้เคลื่อนที่จากจุดหยุดนิ่งออกไปในแนวเส้นตรงด้วยอัตราเร็วคงที่

1. 

2. 

3. 

4. 

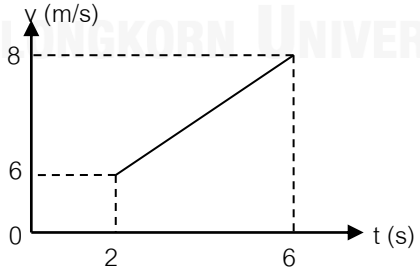
9) ท็อปปล่อยวัตถุมวล 3 กิโลกรัม จากตึกสูง 45 เมตรจากระดับพื้นดิน เมื่อกล่องหล่นมาได้ระยะทาง 25 เมตร พลังงานศักย์มีค่าเท่าใด

1. 600 จูล 2. 750 จูล 3. 1,350 จูล 4. 2,100 จูล

10) กันต์ปล่อยวัตถุมวล 0.5 กิโลกรัมจากตึกแห่งหนึ่ง จงหาพลังงานจลน์ของวัตถุ เมื่อวัตถุถูกปล่อยลงมาเป็นเวลา 2 วินาที

1. 5 จูล 2. 10 จูล 3. 100 จูล 4. 200 จูล

11) รถยนต์มวล 1,000 กิโลกรัม เคลื่อนที่ไปตามพื้นราบงานที่เกิดขึ้นในช่วง 2-6 วินาทีมีค่ากี่จูล ถ้ากราฟระหว่างความเร็ว และเวลาของรถยนต์ขณะเคลื่อนที่ เป็นดังรูป



1. 140 จูล 2. 280 จูล 3. 14,000 จูล 4. 28,000 จูล

12) ยิงลูกปืนมวล 100 กรัมไปในแนวระดับ ลูกปืนมีความเร็ว 200 เมตร/วินาที ถ้าลูกปืนฝังเข้าไปฝังในเนื้อไม้ งานที่ไม้ต้านการเคลื่อนที่ของลูกปืนมีค่าเท่าใด

1. 2×10^3 จูล 2. 4×10^3 จูล 3. 2×10^6 จูล 4. 4×10^6 จูล

13) เมื่อกล่องใบหนึ่งถูกปล่อยจากตึกสูงมายังพื้นราบ ข้อใดกล่าวถูกต้อง
   1. พลังงานรวม พลังงานจลน์ พลังงานศักย์โน้มถ่วง ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน
   2. พลังงานศักย์โน้มถ่วงของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน
   3. พลังงานจลน์ของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน
   4. พลังงานรวมของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน

14) ตั้มปล่อยถุงทราย 5 กิโลกรัม ตกจากหอคอยสูง 7 เมตร ขณะที่ถุงทรายเคลื่อนที่ลงมาได้ระยะทาง 4 เมตรจากหอคอย ความเร็วของถุงทรายมีค่าเท่าใด
   1. $\sqrt{20}$ เมตร/วินาที 2. 20 เมตร/วินาที 3. $\sqrt{60}$ เมตร/วินาที 4.60 เมตร/วินาที

15) วัตถุมวล 36 กิโลกรัม เคลื่อนที่ด้วยความเร็ว 20 เมตร/วินาที เข้าชนสปริงที่มีค่าคงที่ 400 นิวตัน/เมตร สปริงจะหดสั้นที่สุดเท่าไหร่เมื่อกล่องชนสปริง
   1. 6 ซม. 2. 36 ซม. 3. 600 ซม. 4. 3,600 ซม.

16. ถ้าเราให้กำลังแก่มอเตอร์ 0.3 กำลังม้า มอเตอร์จะให้พลังงานงานที่เป็นประโยชน์ออกมาเท่าใด ถ้ามอเตอร์เครื่องนี้มีประสิทธิภาพ 70% (1 กำลังม้า = 746 วัตต์)
   1. 0.21 วัตต์ 2. 15.67 วัตต์ 3. 21.0 วัตต์ 4. 156.7 วัตต์

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ชื่อ-นามสกุล .....ชั้น ม.4/..... เลขที่ ..... ชุดที่ 2

กระดาษคำตอบ

ตอนที่ 1 (แบบสอบปรนัยจำนวน 16 ข้อ รวม 16 คะแนน)

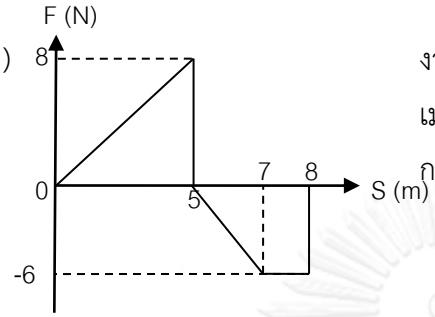
คำสั่ง ให้นักเรียนเลือกคำตอบที่ถูกต้องที่สุด แล้วทำเครื่องหมาย X ลงในกระดาษคำตอบ

ข้อ	ตัวเลือก			
	1	2	3	4
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ชื่อ-นามสกุล .....ชั้น ม.4/..... เลขที่ ..... ชุดที่ 2

ตอนที่ 2 จงแสดงวิธีทำให้ถูกต้อง โดยแสดงวิธีทำอย่างละเอียด **กำหนดให้** g = 10 เมตร/วินาที^{2}

1) 

งานทั้งหมดที่กระทำในช่วงระยะทางการเคลื่อนที่ 0 เมตรถึง 8 เมตร จะมีค่ากี่จูล ถ้ากราฟระหว่างแรงกับการกระจัดเป็นดังรูป (4 คะแนน)

.....

.....

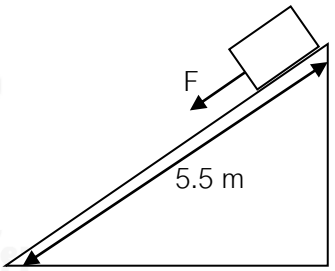
.....

.....

.....

.....

2) ชายคนหนึ่งออกแรงดึงกล่องหนัก 80 กิโลกรัมลงเนินที่เอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ ถ้าเขาออกแรงดึงกล่องในแนวขนานกับการเคลื่อนที่ดังรูป เขาต้องใช้งานเท่าใดในการดึงกล่องลงมาได้ 5.5 เมตร เมื่อสัมประสิทธิ์ความเสียดทานระหว่างเนินและกล่องเท่ากับ $\sqrt{3}/2$ (4 คะแนน)



จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

.....

.....

.....

.....

.....

.....

.....

3) มอเตอร์ไซต์มวล 300 กิโลกรัม แล่นขึ้นพื้นเอียงทำมุม 30 องศากับแนวระดับ ด้วยอัตราเร็วคงที่ 120 เมตร/วินาที ถ้าแรงเสียดทานระหว่างพื้นเอียงกับมอเตอร์ไซต์เท่ากับ 200 นิวตัน จงหากำลังงานของเครื่องยนต์ในหน่วยกิโลวัตต์ (4 คะแนน)

.....

.....

.....

.....

.....

.....

.....

.....

.....

4) ออกแรงดึงสปริง 30 นิวตัน ปรากฏว่าสปริงยาว 50 เซนติเมตร แต่ถ้าออกแรงดึงสปริง 40 นิวตัน สปริงจะมีความยาว 60 เซนติเมตร พลังงานศักย์ยืดหยุ่นขณะที่ออกแรงดึงสปริง 40 นิวตันมีค่าเท่าใด (4 คะแนน)

.....

.....

.....

.....

.....

.....

.....

.....

.....

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

5) แมนเข็นลังใส่ของมวล 40 กิโลกรัม ให้เคลื่อนที่บนพื้นลื่นในแนวเส้นตรงด้วยแรงคงที่ค่าหนึ่ง ซึ่งเมื่อผ่านตำแหน่ง A ลังดังกล่าวมีความเร็ว 4 เมตร/วินาที งานที่แมนต้องใช้ในการเข็นลังนี้จากตำแหน่ง A ไปยังตำแหน่ง B มีค่ากี่จูล ถ้าขณะที่ผ่านตำแหน่ง B เคนต้องใช้กำลังเป็น $\frac{5}{4}$ เท่าของกำลังที่ใช้เมื่อผ่านตำแหน่ง A (4 คะแนน)

.....

.....

.....

.....

.....

.....

.....

.....

.....

6) ถ้าแทนต้องการใช้สปริงในการยิงวัตถุมวล 200 กรัมให้ขึ้นไปในแนวดิ่งเป็นระยะทาง 8 เมตร ด้วยการยืดสปริงออก 40 เซนติเมตร สปริงนี้จะต้องมีค่านิจสปริงกี่นิวตัน/เมตร (4คะแนน)

.....

.....

.....

.....

.....

.....

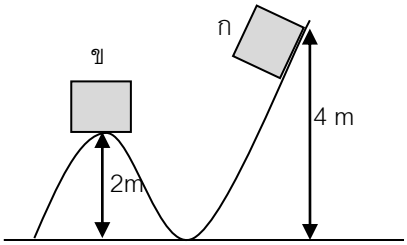
.....

.....

.....

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

7) กล่องมวล 6 กิโลกรัม ที่จุด ก มีอัตราเร็ว 1 เมตร/วินาที และที่จุด ข มีอัตราเร็ว 7 เมตร/วินาที ถ้าระยะทางจากจุด ก ไปยังจุด ข เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (4 คะแนน)



.....

.....

.....

.....

.....

.....

.....

.....

.....

8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัม ขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)

.....

.....

.....

.....

.....

.....

.....

.....

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ง

ผลการวิเคราะห์และผลการปรับปรุง
แบบสอบเรื่องงานและพลังงานทั้ง 2 ฉบับ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 41 ผลการวิเคราะห์ข้อสอบแบบหลายตัวเลือกวิชาฟิสิกส์เรื่องงานและพลังงานฉบับใหม่ (ฉบับทดลอง)

ข้อที	ตัวเลือก	กลุ่มสูง	กลุ่มต่ำ	รวม	P	r	การแปลความหมาย
1	1	4	8	21	.19	-.08	ตัวลวงที่ดี
	2	2	11	18	.117	-.21	ตัวลวงที่ดี
	(3)	12	3	20	.19	.30	ข้อสอบควรปรับปรุง
	4	14	18	49	.45	-.01	ตัวลวงที่ดี
2	(1)	24	9	54	.50	.52	ยากปานกลาง จำแนกดีมาก
	2	5	23	34	.31	-.42	ตัวลวงที่ดี
	3	3	7	18	.17	-.08	ตัวลวงที่ดี
	4	0	1	2	.02	-.03	ตัวลวงที่ดี
3	(1)	27	20	76	.70	.34	ค่อนข้างง่าย จำแนกได้พอใช้
	2	5	7	14	.13	-.02	ตัวลวงที่ดี
	3	0	2	4	.04	-.05	ตัวลวงที่ดี
	4	0	10	13	.12	-.25	ตัวลวงที่ดี
4	1	9	12	29	.27	-.02	ตัวลวงที่ดี
	2	2	8	17	.16	-.14	ตัวลวงที่ดี
	(3)	17	8	9	.39	.33	ค่อนข้างยาก จำแนกได้พอใช้
	4	3	11	18	.17	-.18	ตัวลวงที่ดี
5	1	0	2	3	.03	-.05	ตัวลวงที่ดี
	(2)	31	23	86	.80	.40	ค่อนข้างง่าย จำแนกได้ดี
	3	1	10	9	.13	-.22	ตัวลวงที่ดี
	4	0	5	5	.05	-.13	ตัวลวงที่ดี
6	1	3	9	16	.15	-.13	ตัวลวงที่ดี
	2	1	8	14	.13	-.17	ตัวลวงที่ดี
	3	0	12	22	.20	-.30	ตัวลวงที่ดี
	(4)	26	10	53	.50	.56	ยากปานกลาง จำแนกได้ดี
7	(1)	31	21	80	.74	.44	ค่อนข้างง่าย จำแนกได้ดี
	2	0	8	13	.12	-.20	ตัวลวงที่ดี
	3	0	5	5	.05	-.13	ตัวลวงที่ดี
	4	1	6	10	.09	-.12	ตัวลวงที่ดี
8	1	2	7	13	.12	-.11	ตัวลวงที่ดี
	2	3	13	22	.20	-.23	ตัวลวงที่ดี
	(3)	11	3	24	.22	.27	ค่อนข้างยาก จำแนกพอใช้
	4	16	17	49	.45	.075	ควรปรับปรุงตัวลวง
9	1	9	18	39	.36	-.17	ตัวลวงที่ดี
	(2)	19	3	33	.31	.51	ค่อนข้างยาก จำแนกได้ดี
	3	4	18	34	.31	-.33	ตัวลวงที่ดี
	4	0	1	2	.02	-03	ตัวลวงที่ดี

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 41 ผลการวิเคราะห์ข้อสอบแบบหลายตัวเลือกวิชาฟิสิกส์เรื่องงานและพลังงานฉบับใหม่ (ฉบับทดลอง) (ต่อ)

ข้อที	ตัวเลือก	กลุ่มสูง	กลุ่มต่ำ	รวม	P	r	การแปลความหมาย
10	1	9	18	39	.36	0.17	ตัวลวงที่ดี
	2	19	3	33	.32	-.28	ตัวลวงที่ดี
	3	4	18	34	.32	-.33	ตัวลวงที่ดี
	(4)	14	3	11	.22	.36	ตัวลวงที่ดี
11	1	2	5	10	.09	-.06	ตัวลวงที่ดี
	2	4	12	25	.23	-.18	ตัวลวงที่ดี
	3	13	12	1	.42	.11	ควรปรับปรุงตัวลวง
	(4)	13	9	26	.24	.18	ค่อนข้างยาก จำแนกค่อนข้างต่ำ
12	1	2	5	10	.09	-.06	ตัวลวงที่ดี
	(2)	24	8	46	.43	.55	ยากปานกลาง จำแนกได้ดี
	3	5	22	35	.32	-.39	ตัวลวงที่ดี
	4	2	6	15	.14	-.09	ตัวลวงที่ดี
13	(1)	23	9	48	.44	.49	ยากปานกลาง จำแนกได้ดี
	2	1	10	16	.15	-.22	ตัวลวงที่ดี
	3	0	9	15	.14	-.23	ตัวลวงที่ดี
	4	8	12	29	.27	-.05	ตัวลวงที่ดี
14	1	11	7	29	.27	.17	ตัวลวงที่ดี
	(2)	14	4	23	.21	.33	ค่อนข้างยาก จำแนกได้พอใช้
	3	5	21	42	.32	-.39	ตัวลวงที่ดี
	4	.2	8	14	.13	-14	ตัวลวงที่ดี
15	1	7	8	26	.24	.02	ตัวลวงที่ดี
	2	7	16	36	.33	-.18	ตัวลวงที่ดี
	(3)	16	6	30	.28	.35	ยากปานกลาง จำแนกได้พอใช้
	4	2	10	16	.15	-.19	ตัวลวงที่ดี
16	1	1	4	6	.06	-.07	ตัวลวงที่ดี
	2	5	9	23	.21	-.07	ตัวลวงที่ดี
	3	1	13	19	.18	-.29	ตัวลวงที่ดี
	(4)	25	14	60	.56	.43	ยากปานกลาง จำแนกได้ดี

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 42 ผลการวิเคราะห์ข้อสอบแบบหลายตัวเลือกวิชาฟิสิกส์เรื่องงานและพลังงานฉบับอ้างอิง (ฉบับทดลอง)

ข้อที	ตัวเลือก	กลุ่มสูง	กลุ่มต่ำ	รวม	P	r	การแปลความหมาย
1	(1)	24	7	39	.36	.55	ค่อนข้างยาก จำแนกได้ดี
	2	7	13	23	.22	-.08	ตัวลวงที่ดี
	3	0	4	9	.08	-.09	ตัวลวงที่ดี
	4	14	18	49	.45	-.01	ตัวลวงที่ดี
2	1	1	5	11	.10	-.08	ตัวลวงที่ดี
	(2)	27	21	65	.60	.33	ค่อนข้างง่าย จำแนกได้พอใช้
	3	4	18	29	.27	-.27	ตัวลวงที่ดี
	4	2	2	4	.04	.02	ควรปรับปรุงตัวลวง
3	(1)	26	29	79	.73	.29	ค่อนข้างง่าย จำแนกได้พอใช้
	2	1	4	6	.06	-.6	ตัวลวงที่ดี
	3	3	10	16	.15	-.13	ตัวลวงที่ดี
	4	1	6	8	.07	-.10	ตัวลวงที่ดี
4	1	6	13	26	.24	-.11	ตัวลวงที่ดี
	(2)	19	16	46	.42	.21	ยากปานกลาง จำแนกได้พอใช้
	3	4	9	22	.20	-.08	ตัวลวงที่ดี
	4	5	7	14	.13	-.01	ตัวลวงที่ดี
5	(1)	32	26	82	.75	.38	ค่อนข้างง่าย จำแนกได้พอใช้
	2	1	6	10	.09	-.10	ตัวลวงที่ดี
	3	1	3	5	.05	-.04	ตัวลวงที่ดี
	4	0	11	12	.11	-.24	ตัวลวงที่ดี
6	1	2	4	9	.08	-.03	ตัวลวงที่ดี
	2	2	8	17	.16	-.12	ตัวลวงที่ดี
	3	5	26	39	.34	-.42	ตัวลวงที่ดี
	(4)	25	7	43	.39	.58	ค่อนข้างยาก จำแนกได้ดี
7	1	5	9	19	.17	-.05	ตัวลวงที่ดี
	2	1	5	9	.08	-.08	ตัวลวงที่ดี
	3	1	14	19	.17	-.28	ตัวลวงที่ดี
	(4)	27	17	61	.56	.43	ยากปานกลาง จำแนกได้ดี
8	1	2	4	6	.55	-.03	ตัวลวงที่ดี
	2	11	12	31	.28	.06	ตัวลวงที่ดี
	(3)	11	21	50	.46	-.13	ข้อสอบควรปรับปรุง
	4	10	9	22	.20	.10	ควรปรับปรุงตัวลวง
9	(1)	24	3	34	.31	.64	ค่อนข้างยาก จำแนกได้ดีมาก
	2	8	24	43	.39	-.29	ตัวลวงที่ดี
	3	1	14	24	.22	-.28	ตัวลวงที่ดี
	4	1	4	7	.06	-.06	ตัวลวงที่ดี

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 42 ผลการวิเคราะห์ข้อสอบแบบหลายตัวเลือกวิชาฟิสิกส์เรื่องงานและพลังงานฉบับอ้างอิง (ฉบับทดลอง) (ต่อ)

ข้อที	ตัวเลือก	กลุ่มสูง	กลุ่มต่ำ	รวม	P	r	การแปลความหมาย
10	1	4	13	22	.20	-.17	ตัวลวงที่ดี
	2	5	25	44	.40	-.40	ตัวลวงที่ดี
	(3)	25	4	37	.34	.65	ค่อนข้างยาก จำแนกได้ดีมาก
	4	0	3	5	.05	-.07	ตัวลวงที่ดี
11	1	1	4	7	.06	-.06	ตัวลวงที่ดี
	2	8	15	33	.30	-.09	ตัวลวงที่ดี
	(3)	16	13	36	.33	.19	ข้อสอบควรปรับปรุง
	4	9	13	32	.29	-.02	ตัวลวงที่ดี
12	(1)	25	21	64	.59	.28	ยากปานกลาง จำแนกได้พอใช้
	2	3	18	30	.28	-.30	ตัวลวงที่ดี
	3	7	10	21	.19	-.01	ตัวลวงที่ดี
	4	0	3	5	.05	-.07	ตัวลวงที่ดี
13	1	2	13	20	.19	-.22	ตัวลวงที่ดี
	2	3	18	30	.28	-.30	ตัวลวงที่ดี
	3	4	9	18	.17	-.08	ตัวลวงที่ดี
	(4)	25	6	41	.38	.52	ค่อนข้างยาก จำแนกได้ดี
14	1	5	8	20	.18	-.03	ตัวลวงที่ดี
	2	4	8	32	.29	-.06	ตัวลวงที่ดี
	(3)	23	7	41	.38	.52	ค่อนข้างยาก จำแนกได้ดี
	4	5	8	20	.18	-.03	ตัวลวงที่ดี
15	1	11	15	36	.33	-.01	ตัวลวงที่ดี
	2	7	12	25	.23	-.06	ตัวลวงที่ดี
	(3)	14	8	30	.28	.24	ยากปานกลาง จำแนกได้พอใช้
	4	2	10	17	.16	-.16	ตัวลวงที่ดี
16	1	2	8	11	.10	-.12	ตัวลวงที่ดี
	2	6	7	17	.12	.02	ตัวลวงที่ดี
	3	2	15	21	.19	-.27	ตัวลวงที่ดี
	(4)	24	15	59	.54	.38	ยากปานกลาง จำแนกได้พอใช้

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 43 ผลการวิเคราะห์ข้อสอบแบบสร้างคำตอบวิชาฟิสิกส์เรื่องงานและพลังงานฉบับใหม่ (ฉบับทดลอง)

ข้อที	กลุ่มสูง	กลุ่มต่ำ	P	r	การแปลความหมาย
1	27	27	.43	.83	ยากปานกลางจำแนกได้ดีมาก
2	27	27	.30	.32	ค่อนข้างยาก จำแนกได้พอใช้
3	27	27	.32	.65	ค่อนข้างยาก จำแนกได้ดีมาก
4	27	27	.13	.07	ควรปรับปรุงข้อสอบ
5	27	27	.07	.14	ควรปรับปรุงข้อสอบ
6	27	27	.31	.63	ค่อนข้างยาก จำแนกได้ดีมาก
7	27	27	.29	.58	ค่อนข้างยาก จำแนกได้ดี
8	27	27	.21	.41	ค่อนข้างยาก จำแนกได้ดี

ตาราง 44 ผลการวิเคราะห์ข้อสอบแบบสร้างคำตอบวิชาฟิสิกส์เรื่องงานและพลังงานฉบับอ้างอิง (ฉบับทดลอง)

ข้อที	กลุ่มสูง	กลุ่มต่ำ	P	r	การแปลความหมาย
1	27	27	.45	.88	ยากปานกลางจำแนกได้ดีมาก
2	27	27	.36	.61	ค่อนข้างยาก จำแนกได้ดีมาก
3	27	27	.33	.65	ค่อนข้างยาก จำแนกได้ดีมาก
4	27	27	.18	.31	ควรปรับปรุงข้อสอบ
5	27	27	.08	.16	ควรปรับปรุงข้อสอบ
6	27	27	.25	.51	ค่อนข้างยาก จำแนกได้ดี
7	27	27	.24	.48	ค่อนข้างยาก จำแนกได้ดี
8	27	27	.25	.50	ค่อนข้างยาก จำแนกได้ดี

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

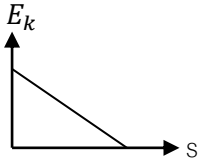
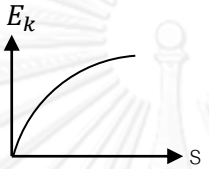
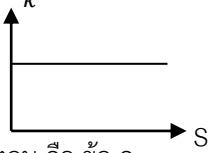
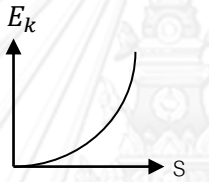
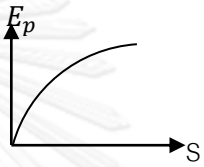
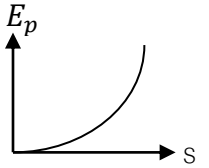
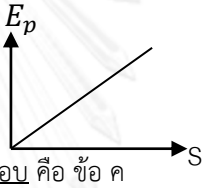
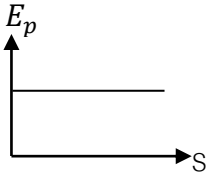
ตาราง 45 ผลการปรับปรุงข้อสอบวิชาฟิสิกส์ เรื่อง งานและพลังงานระดับชั้นมัธยมศึกษาปีที่ 4 ของแบบสอบฉบับใหม่และฉบับอ้างอิง

<u>จุดประสงค์ที่ 1</u> อธิบายงานและวิเคราะห์งานของแรงต่างๆได้

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
<u>แบบสอบฉบับใหม่ (ฉบับที่ 1)</u> 1) แรง 10 นิวตัน ทำให้วัตถุเคลื่อนที่เป็นเส้นตรงบนพื้นราบลื่นในระยะทาง S ถ้าต้องการหาค่างานของแรง 20 N จะต้องพิจารณาสิ่งใดบ้าง (เข้าใจ) 1. อัตราเร็วของวัตถุ 2. แรงเสียดทานระหว่างวัตถุกับผิวสัมผัส 3. มุมระหว่างทิศของแรงกับแนวการเคลื่อนที่ของวัตถุ ก. 1 ข. 2 ค. 3 ง. 2 และ 3 <u>คำตอบ</u> คือ ข้อ ค	<u>แบบสอบฉบับใหม่ (ฉบับที่ 1)</u> 1) งานในข้อใดมีค่าเป็นลบ ก. งานที่เกิดจากแรงทำมุม 90 องศากับแนวระดับ ข. งานที่เกิดจากแรงทำมุม 180 องศากับแนวระดับ ค. งานที่เกิดจากแรงทำมุม 210 องศากับแนวระดับ ง. งานที่เกิดจากแรงทำมุม 300 องศากับแนวระดับ ก. ก ข. ข ค. ข และ ค ง. ข ค และ ง <u>คำตอบ</u> คือ ข้อ ค
<u>แบบสอบฉบับอ้างอิง (ฉบับที่ 2)</u> 1) กล่องเคลื่อนที่เป็นเส้นตรงบนพื้นราบลื่นด้วยแรง 20 N เป็นระยะทาง S ถ้าต้องการหาค่างานของแรง 10 N จะต้องพิจารณาสิ่งใดบ้าง (เข้าใจ) 1. มุมระหว่างทิศของแรงกับแนวการเคลื่อนที่ของวัตถุ 2. แรงเสียดทานระหว่างวัตถุกับพื้นราบ 3. อัตราเร็วของวัตถุ ก. 1 ข. 2 ค. 3 ง. 2 และ 3 <u>คำตอบ</u> คือ ข้อ ก	<u>แบบสอบฉบับอ้างอิง (ฉบับที่ 2)</u> 1) งานในข้อใดมีค่าเป็นบวก ก. งานที่เกิดจากแรงทำมุม 90 องศากับแนวระดับ ข. งานที่เกิดจากแรงทำมุม 180 องศากับแนวระดับ ค. งานที่เกิดจากแรงทำมุม 210 องศากับแนวระดับ ง. งานที่เกิดจากแรงทำมุม 300 องศากับแนวระดับ ก. ก ข. ง ค. ก และ ง ง. ก ค และ ง <u>คำตอบ</u> คือ ข้อ ข
<u>แบบสอบฉบับอ้างอิง (ฉบับที่ 2)</u> 2) กรณีใดที่ถือว่ามีการทำงาน เนื่องจากการออกแรงกระทำต่อกล่อง ก. นุชยืนถือกล่อง ข. นุชยกกล่องจากโต๊ะไปวางบนพื้นในแนวดิ่ง ค. นุชยกกล่องจากพื้นไปวางบนโต๊ะในแนวดิ่ง ง. นุชถือกล่องแล้วเดินไปตามพื้นราบ 1. ง 2. ข และ ค 3. ข ค และ ง 4. ก ข ค และ ง	<u>แบบสอบฉบับอ้างอิง (ฉบับที่ 2)</u> 2) กรณีใดที่ถือว่ามีการทำงาน เนื่องจากการออกแรงกระทำต่อกล่อง ก. นุชยืนถือกล่อง ข. นุชยกกล่องจากโต๊ะไปวางบนพื้นในแนวดิ่ง ค. นุชยกกล่องจากพื้นไปวางบนโต๊ะในแนวดิ่ง ง. นุชถือกล่องแล้วเดินไปตามพื้นราบ 1. ง 2. ข และ ค 3. ข ค และ ง 4. ก ค และ ง

ตาราง 45 ผลการปรับปรุงข้อสอบวิชาฟิสิกส์ เรื่อง งานและพลังงานระดับชั้นมัธยมศึกษาปีที่ 4 ของแบบสอบฉบับใหม่และฉบับอ้างอิง (ต่อ)

<u>จุดประสงค์ที่ 4</u> อธิบายพลังงาน พลังงานจลน์ พลังงานศักย์ และความสัมพันธ์ระหว่างงานและพลังงานได้

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
<u>แบบสอบฉบับอ้างอิง (ฉบับที่ 2)</u> 8) กราฟในตัวเลือกใดแสดงความสัมพันธ์ระหว่างพลังงานจลน์ของวัตถุ (E_k) กับการกระจัด (S) ได้ถูกต้อง เมื่อวัตถุนี้เคลื่อนที่จากจุดหยุดนิ่งออกไปในแนวเส้นตรงด้วยอัตราเร็วคงที่ (เข้าใจ) ก.  ข.  ค.  ง.  <u>คำตอบ</u> คือ ข้อ ค	<u>แบบสอบฉบับอ้างอิง (ฉบับที่ 2)</u> 8) กราฟในตัวเลือกใดแสดงความสัมพันธ์ระหว่างพลังงานศักย์โน้มถ่วง (E_p) ของวัตถุกับการกระจัด (S) ได้ถูกต้อง เมื่อวัตถุนี้ถูกโยนขึ้นไปในแนวดิ่งจากจุดหยุดนิ่งด้วยอัตราเร็วคงที่ (เข้าใจ) ก.  ข.  ค.  ง.  <u>คำตอบ</u> คือ ข้อ ค

<u>จุดประสงค์ที่ 5</u> คำนวณหาพลังงานจลน์ พลังงานศักย์โน้มถ่วงของวัตถุที่ระดับต่างๆ และพลังงานศักย์ยืดหยุ่นของสปริง

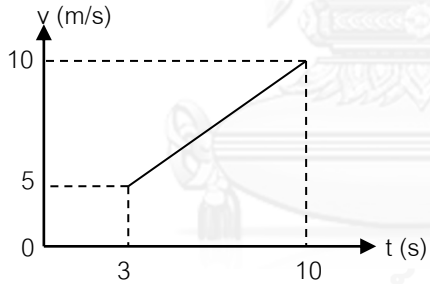
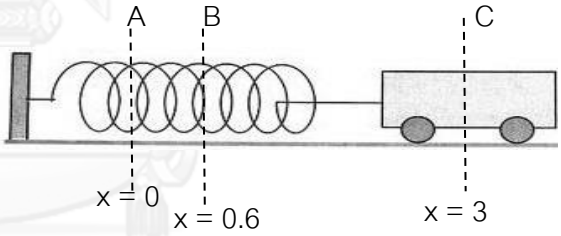
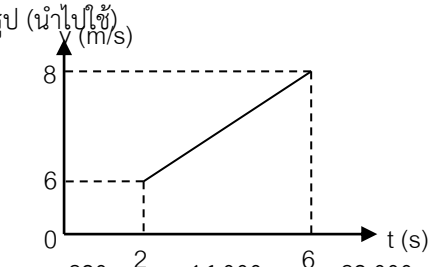
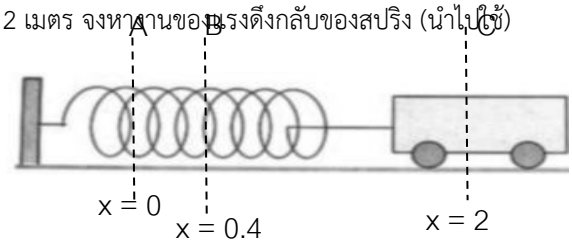
ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
<u>แบบสอบฉบับใหม่ (ฉบับที่ 1)</u> 10) ลูกมะพร้าวมวล 1 กิโลกรัม อยู่บนต้นมะพร้าว จงหาพลังงานจลน์ของลูกมะพร้าวเมื่อหลุดออกจากต้นมะพร้าวที่เวลา 1 วินาที 1. 5 จูล 2. 10 จูล 3. 20 จูล 4. 50 จูล <u>คำตอบ</u> คือ ข้อ ง หมายเหตุ : ปรับเฉพาะตัวลวงข้อ ก	<u>แบบสอบฉบับใหม่ (ฉบับที่ 1)</u> 10) ลูกมะพร้าวมวล 1 กิโลกรัม อยู่บนต้นมะพร้าว จงหาพลังงานจลน์ของลูกมะพร้าวเมื่อหลุดออกจากต้นมะพร้าวที่เวลา 1 วินาที 1. 0.5 จูล 2. 10 จูล 3. 20 จูล 4. 50 จูล <u>คำตอบ</u> คือ ข้อ ง หมายเหตุ : ปรับเฉพาะตัวลวงข้อ ก
<u>แบบสอบฉบับใหม่ (ฉบับที่ 1)</u> 4) (อัตนัย) ออกแรงดึงสปริง 50 นิวตัน ปรากฏว่าสปริงยาว 70 เซนติเมตร แต่ถ้าออกแรงดึงสปริง 60 นิวตัน สปริงจะมีความยาว 80 เซนติเมตร พลังงานศักย์ยืดหยุ่นขณะที่ออกแรงดึงสปริง 60 นิวตันมีค่าเท่าใด (วิเคราะห์) <u>คำตอบ</u> คือ 18 จูล	<u>แบบสอบฉบับใหม่ (ฉบับที่ 1)</u> 4) (อัตนัย) สปริงถูกออกแรงดึง 20 นิวตัน จะทำให้ยืดออก 10 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 20 เซนติเมตรในหน่วยจูล (วิเคราะห์) <u>คำตอบ</u> คือ 4 จูล

ตาราง 45 ผลการปรับปรุงข้อสอบวิชาฟิสิกส์ เรื่อง งานและพลังงานระดับชั้นมัธยมศึกษาปีที่ 4 ของแบบสอบฉบับใหม่และฉบับอ้างอิง (ต่อ)

<u>จุดประสงค์ที่ 5</u> คำนวณหาพลังงานจลน์ พลังงานศักย์โน้มถ่วงของวัตถุที่ระดับต่างๆ และพลังงานศักย์ยืดหยุ่นของสปริง

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
<u>แบบสอบฉบับอ้างอิง (ฉบับที่ 2)</u> 4) (อัตนัย) ออกแรงดึงสปริง 30 นิวตัน ปรากฏว่าสปริงยาว 50 เซนติเมตร แต่ถ้าออกแรงดึงสปริง 40 นิวตัน สปริงจะมีความยาว 60 เซนติเมตร พลังงานศักย์ยืดหยุ่นขณะที่ออกแรงดึงสปริง 40 นิวตันมีค่าเท่าใด วิเคราะห์ <u>คำตอบ</u> คือ 8 จูล	<u>แบบสอบฉบับอ้างอิง(ฉบับที่ 2)</u> 4) (อัตนัย) สปริงถูกดึงด้วยแรง 40 นิวตัน จะทำให้สปริงยืดออก 20 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 30 เซนติเมตรในหน่วยจูล (วิเคราะห์) <u>คำตอบ</u> คือ 9 จูล

<u>จุดประสงค์ที่ 6</u> คำนวณโจทย์เกี่ยวกับความสัมพันธ์ระหว่างงานและพลังงานได้

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
<u>แบบสอบฉบับใหม่ (ฉบับที่ 1)</u> 11) รถยนต์มวล 2,000 กิโลกรัม เคลื่อนที่ไปตามพื้นราบ งานที่เกิดขึ้นในช่วง 3-10 วินาที มีค่ากี่จูล ถ้ากราฟระหว่างความเร็ว และเวลาของรถยนต์ขณะเคลื่อนที่เป็นดังรูป (นำไปใช้)  ก. 75 จูล ข. 750 จูล ค. 7,500 จูล ง. 75,000 จูล <u>คำตอบ</u> คือ ข้อ ง	<u>แบบสอบฉบับใหม่ (ฉบับที่ 1)</u> 11) เมื่อนำสปริงที่มีค่าคงที่ 10 N/m มาผูกปลายข้าง หนึ่งกับรถทดลองมวล m ที่วางอยู่บนพื้นราบลื่นดังรูป สปริงยืดอยู่ก่อนแล้ว 0.6 เมตร ถ้าดึงสปริงให้ยืดเป็น 3 เมตร จงหางานของแรงดึงกลับของสปริง (นำไปใช้)  ก. -45.0 จูล ข. -43.2จูล ค. +43.2 จูล ง. +45.0 จูล <u>คำตอบ</u> คือ ข้อ ข
<u>แบบสอบฉบับอ้างอิง (ฉบับที่ 2)</u> 11) รถยนต์มวล 1,000 กิโลกรัม เคลื่อนที่ไปตามพื้นราบ งานที่เกิดขึ้นในช่วง 2-6 วินาทีมีค่ากี่จูล ถ้ากราฟระหว่างความเร็ว และเวลาของรถยนต์ขณะเคลื่อนที่เป็นดังรูป (นำไปใช้)  ก. 140 จูล ข. 280 จูล ค. 14,000 จูล ง. 28,000 จูล <u>คำตอบ</u> คือ ข้อ ค	<u>แบบสอบฉบับอ้างอิง (ฉบับที่ 2)</u> 11) แจ็คนำสปริงที่มีค่าคงที่ 20 N/m มาผูกปลายข้างหนึ่งกับรถทดลองมวล m ที่วางอยู่บนพื้นราบลื่นดังรูป สปริงยืดอยู่ก่อนแล้ว 0.4 เมตร ถ้าแจ็คดึงสปริงให้ยืดเป็น 2 เมตร จงหางานของแรงดึงกลับของสปริง (นำไปใช้)  ก. +38.4 จูล ข. -38.4จูล ค. -40.0 จูล ง. +40.0 จูล <u>คำตอบ</u> คือ ข้อ ข

ตาราง 45 ผลการปรับปรุงข้อสอบวิชาฟิสิกส์ เรื่อง งานและพลังงานระดับชั้นมัธยมศึกษาปีที่ 4 ของแบบสอบฉบับใหม่และฉบับอ้างอิง (ต่อ)

<u>จุดประสงค์ที่ 6</u> คำนวณโจทย์เกี่ยวกับความสัมพันธ์ระหว่างงานและพลังงานได้

ข้อสอบก่อนปรับปรุง	ข้อสอบหลังปรับปรุง
<u>แบบสอบฉบับใหม่ (ฉบับที่ 1)</u> 5) (อัตนัย) เคนเข็นลังใส่ผลไม้มวล 50 กิโลกรัม ให้เคลื่อนที่ในแนวเส้นตรงด้วยแรงคงที่ค่าหนึ่งไปบนพื้นลื่น เมื่อผ่านตำแหน่ง ก เคนต้องใช้กำลังเป็น $\frac{2}{3}$ เท่าของกำลังที่ใช้เมื่อผ่านตำแหน่ง ข ถ้าความเร็วของลังนี้เมื่อผ่านตำแหน่ง ข คือ 3 เมตร/วินาที งานที่ใช้ในการเข็นลังนี้จากตำแหน่ง ก ไปยังตำแหน่ง ข มีค่ากี่จูล (วิเคราะห์) <u>คำตอบ</u> คือ 125 จูล	<u>แบบสอบฉบับใหม่ (ฉบับที่ 1)</u> 5) (อัตนัย) เคนนำเครื่องสูบน้ำเครื่องหนึ่งไปสูบน้ำมวล 400 กิโลกรัม ขึ้นไปตามท่อสูง 10 เมตร ถ้าน้ำพุ่งออกจากปลายท่อด้วยความเร็ว 10 เมตรต่อวินาที งานที่ได้จากเครื่องสูบมีค่ากี่กิโลจูล (วิเคราะห์) <u>คำตอบ</u> คือ 60 กิโลจูล
<u>แบบสอบฉบับอ้างอิง (ฉบับที่ 2)</u> 5) (อัตนัย) แมนเข็นลังใส่ของมวล 40 กิโลกรัม ให้เคลื่อนที่บนพื้นลื่นในแนวเส้นตรงด้วยแรงคงที่ค่าหนึ่ง ซึ่งเมื่อผ่านตำแหน่ง A ลังดังกล่าวมีความเร็ว 4 เมตร/วินาที งานที่แมนต้องใช้ในการเข็นลังนี้จากตำแหน่ง A ไปยังตำแหน่ง B มีค่ากี่จูล ถ้าขณะที่ผ่านตำแหน่ง B เคนต้องใช้กำลังเป็น $\frac{5}{4}$ เท่าของกำลังที่ใช้เมื่อผ่านตำแหน่ง A (วิเคราะห์) <u>คำตอบ</u> คือ 180 จูล	<u>แบบสอบฉบับอ้างอิง (ฉบับที่ 2)</u> 5) (อัตนัย) เครื่องสูบน้ำใช้สูบน้ำมวล 500 กิโลกรัม ขึ้นจากบ่อลึก 6 เมตร แล้วฉีดน้ำออกจากปลายท่อด้วยความเร็ว 6 เมตรต่อวินาที งานที่ได้จากเครื่องสูบน้ำนี้มีค่ากี่กิโลจูล (วิเคราะห์) <u>คำตอบ</u> คือ 39 กิโลจูล

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก จ

แบบสอบเรื่องงานและพลังงาน รายวิชาฟิสิกส์
สำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4
ฉบับใหม่ (ชุด 1) และฉบับอ้างอิง (ชุด 2)
ฉบับสมบูรณ์

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

แบบสอบ เรื่องงานและพลังงาน (ชุด 1)

การประเมินผลระหว่างเรียน ระดับชั้นมัธยมศึกษาปีที่ 4 ปีการศึกษา 2556
แบบสอบรายวิชา วิทยาศาสตร์ คะแนนเต็ม 48 คะแนน เวลา 90 นาที

คำชี้แจง 1. แบบสอบฉบับนี้มีทั้งหมด 8 หน้า จำนวน 2 ตอน
ตอนที่ 1 เป็นข้อสอบปรนัย จำนวน 16 ข้อ รวม 16 คะแนน
ตอนที่ 2 เป็นข้อสอบอัตนัย จำนวน 8 ข้อ รวม 32 คะแนน
2. สามารถทดเลขลงในแบบสอบได้ **(ไม่อนุญาตให้ใช้เครื่องคิดเลข)**
3. ส่งกระดาษคำตอบคืนพร้อมกับแบบสอบ

กำหนดให้ g = 10 เมตร/วินาที^{2}

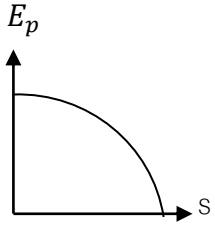
ตอนที่ 1 ให้นักเรียนเลือกคำตอบที่ถูกต้องที่สุด แล้วทำเครื่องหมาย X ลงในกระดาษคำตอบ

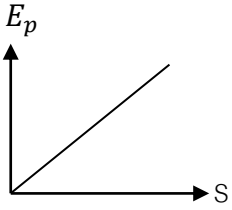
1) กรณีใดที่ถือว่ามีการทำงาน เนื่องจากการออกแรงกระทำต่อหนังสือ
ก. แจนยืนถือหนังสือ ข. แจนนั่งอ่านหนังสือเตรียมสอบ
ค. แจนยกหนังสือจากพื้นไปวางบนโต๊ะในแนวดิ่ง
ง. แจนถือหนังสือแล้วเดินไปตามพื้นราบ
1. ค 2. ค และ ง 3. ก ค และ ง 4. ก ข ค และ ง

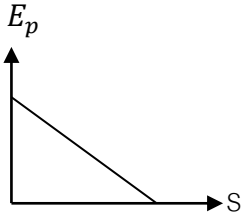
2) ลังมวล 40 กิโลกรัมเคลื่อนที่ลงไปตามพื้นเอียงผิวขรุขระเป็นระยะทาง 10 เมตร ด้วยความเร็วคงที่ ถ้างานในการเคลื่อนที่ของลังดังกล่าวเป็น 3,600 จูล แรงเสียดทานระหว่างพื้นและลังจะมีค่าเท่าใด ถ้าพื้นเอียงอยู่สูงจากพื้น 6 เมตร
1. - 40 นิวตัน 2. 40 นิวตัน 3. -120 นิวตัน 4.120 นิวตัน

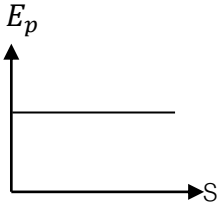
3) นาย ก. ออกแรงคงที่ขนาด 100 นิวตัน ผลักวัตถุก้อนหนึ่งจากหยุดนิ่งให้เคลื่อนที่ไปตามแนวราบได้ระยะทาง 3 เมตร ภายในเวลา 2 วินาที จงหากำลังของนาย ก. ที่ใช้ในการผลักวัตถุนี้
1. 17 วัตต์ 2. 150 วัตต์ 3. 300 วัตต์ 4. 600 วัตต์

4) กราฟในตัวเลือกใดแสดงความสัมพันธ์ระหว่างพลังงานศักย์โน้มถ่วง (E_p) ของวัตถุกับการกระจัด (S) ได้ถูกต้อง เมื่อวัตถุนี้ถูกปล่อยจากจุดหยุดนิ่งลงมาในแนวดิ่งด้วยอัตราเร็วคงที่

1. 

2. 

3. 

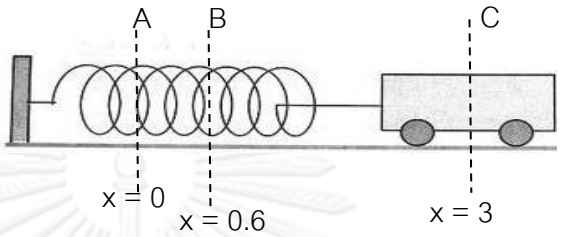
4. 

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

5) กล่องมวล 2 กิโลกรัม ตกจากหน้าผาสูง 40 เมตรจากระดับพื้นดิน เมื่อกล่องหล่นมาได้ระยะทาง 15 เมตร พลังงานศักย์โน้มถ่วงมีค่าเท่าใด

1. 300 จูล 2. 500 จูล 3. 800 จูล 4. 1,100 จูล

6) เมื่อนำสปริงที่มีค่าคงที่ 10 นิวตัน/เมตร มาผูกปลายข้างหนึ่งกับรถทดลองมวล m ที่วางอยู่บนพื้นราบลื่นดังรูป สปริงยืดอยู่ก่อนแล้ว 0.6 เมตร ถ้าดึงสปริงให้ยืดเป็น 3 เมตร จงหางานของแรงดึงกลับของสปริง (นำไปใช้)



1. -45.0 จูล 2. -43.2 จูล 3. +43.2 จูล 4. +45.0 จูล

7) เมื่อวัตถุก้อนหนึ่งเคลื่อนที่ลงจากพื้นเอียงลื่นมายังพื้นราบลื่น ข้อใดกล่าวถูกต้อง

1. พลังงานรวมของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน
2. พลังงานจลน์ของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน
3. พลังงานศักย์โน้มถ่วงของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน
4. พลังงานรวม พลังงานจลน์ พลังงานศักย์โน้มถ่วง ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน

8) กล่องมวล 25 กิโลกรัมเคลื่อนที่ด้วยความเร็ว 10 เมตร/วินาที เข้าชนสปริงที่มีค่าคงที่ 100 นิวตัน/เมตร เมื่อกล่องชนสปริง สปริงจะหดสั้นที่สุดเท่าไหร่

1. 5 ซม. 2. 25 ซม. 3. 500 ซม. 4. 2,500 ซม.

9) กล่องเคลื่อนที่เป็นเส้นตรงบนพื้นราบลื่นด้วยแรง 20 นิวตัน เป็นระยะทาง S ถ้าต้องการหาค่างานของแรง 20 นิวตัน จะต้องพิจารณาสิ่งใดบ้าง

ก. มุมระหว่างทิศของแรงกับแนวการเคลื่อนที่ของวัตถุ
ข. แรงเสียดทานระหว่างวัตถุกับพื้นราบ
ค. อัตราเร็วของวัตถุ

1. ก 2. ข 3. ค 4. ข และ ค

10) มินออกแรง 100 นิวตันลากกล่องในแนวทำมุม 60 องศากับแนวระดับ ถ้าลากกล่องไปได้ไกล 10 เมตรในแนวระดับ งานของแรงที่ใช้ลากกล่องเป็นเท่าใด

1. 500 จูล 2. $500\sqrt{3}$ จูล 3. 1,000 จูล 4. $1{,}000\sqrt{3}$ จูล

11) ลิฟท์ตัวหนึ่งมวล 200 กิโลกรัม ใช้มอเตอร์ 8 กิโลวัตต์ ขณะที่บรรทุกคนให้เคลื่อนที่ขึ้นด้วยความเร็วคงที่ 2 เมตร/วินาที จะต้องใช้แรงเท่าใด

1. 4 นิวตัน 2. 25 นิวตัน 3. 100 นิวตัน 4. 4,000 นิวตัน

12) พลังงานจลน์ของวัตถุในกรณีใดมีค่าต่ำที่สุด

1. เมื่อเพิ่มมวลเป็น 2 เท่า
2. เมื่อเพิ่มความเร็วเป็น 2 เท่า
3. เมื่อลดมวลลง 1/2 เท่า
4. เมื่อลดความเร็วลง 1/2 เท่า

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

13) กันต์ปล่อยวัตถุมวล 0.5 กิโลกรัมจากตึกแห่งหนึ่ง จงหาพลังงานจลน์ของวัตถุ เมื่อวัตถุถูกปล่อยลงมาเป็นเวลา 2 วินาที

1. 5 จูล 2. 10 จูล 3. 100 จูล 4. 200 จูล

14) ยิงลูกปืนมวล 150 กรัมไปในแนวระดับ ลูกปืนมีความเร็ว 100 เมตร/วินาที ถ้าลูกปืนฝังเข้าไปในเนื้อไม้ งานที่ไม้ต้านการเคลื่อนที่ของลูกปืนมีค่าเท่าใด

1. 150 จูล 2. 750 จูล 3. 15,000 จูล 4. 75,000 จูล

15) ตั้มปล่อยถุงทราย 5 กิโลกรัมตกจากหอคอยสูง 7 เมตร ขณะที่ถุงทรายเคลื่อนที่ลงมาได้ระยะทาง 4 เมตรจากหอคอย ความเร็วของถุงทรายมีค่าเท่าใด

1. $\sqrt{20}$ เมตร/วินาที 2. 20 เมตร/วินาที 3. $\sqrt{60}$ เมตร/วินาที 4.60 เมตร/วินาที

16) มอเตอร์เครื่องหนึ่งที่มีประสิทธิภาพ 90% จะให้พลังงานงานที่เป็นประโยชน์ออกมาเท่าใด เมื่อใส่กำลังให้แก่มอเตอร์เครื่องนี้ 0.5 กำลังม้า (1 กำลังม้า = 746 วัตต์)

1. 0.45 วัตต์ 2. 33.57 วัตต์ 3. 45.00 วัตต์ 4. 335.70 วัตต์

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ชื่อ-นามสกุล .....ชั้น ม.4/..... เลขที่ ..... ชุดที่ 1

กระดาษคำตอบ

ตอนที่ 1 (แบบสอบปรนัยจำนวน 16 ข้อ รวม 16 คะแนน)

คำชี้แจง ให้นักเรียนเลือกคำตอบที่ถูกต้องที่สุด แล้วทำเครื่องหมาย X ลงในกระดาษคำตอบ

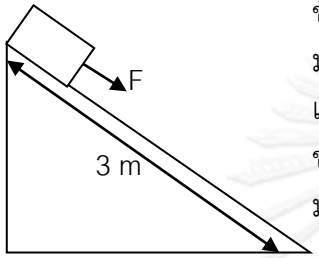
ข้อ	ตัวเลือก			
	1	2	3	4
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ชื่อ-นามสกุล .....ชั้น ม.4/..... เลขที่ ..... ชุดที่ 1

ตอนที่ 2 (แบบสอบอัตนัยจำนวน 8 ข้อ รวม 32 คะแนน) **กำหนดให้** g = 10 เมตร/วินาที^{2}

คำชี้แจง จงแสดงวิธีทำให้ถูกต้อง โดยแสดงวิธีทำอย่างละเอียด (แสดงให้เห็นถึงสมการ การแทนค่า และคำตอบพร้อมระบุหน่วย)

1) 

ช้างตัวหนึ่งออกแรงดึงซุงหนัก 140 กิโลกรัม ลงเนินซึ่งเอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ สัมประสิทธิ์ความเสียดทานระหว่างเนินและซุงเป็น $\sqrt{3}/2$ ถ้าแรงที่ใช้ดึงซุงขนานกับการเคลื่อนที่ดังรูป ช้างต้องใช้งานเท่าใดในการดึงซุงลงมาได้ 3 เมตร (4 คะแนน)

.....

.....

.....

.....

.....

.....

.....

2) สปริงถูกออกแรงดึง 20 นิวตัน จะทำให้ยืดออก 10 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 20 เซนติเมตรในหน่วยจูล (4 คะแนน)

.....

.....

.....

.....

.....

.....

.....

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3) เคนนำเครื่องสูบน้ำเครื่องหนึ่งไปสูบน้ำมวล 400 กิโลกรัม ขึ้นไปตามท่อสูง 10 เมตร ถ้าน้ำพุ่งออกจากปลายท่อด้วยความเร็ว 10 เมตรต่อวินาที งานที่ได้จากเครื่องสูบมีค่ากี่กิโลจูล (4 คะแนน)

.....

.....

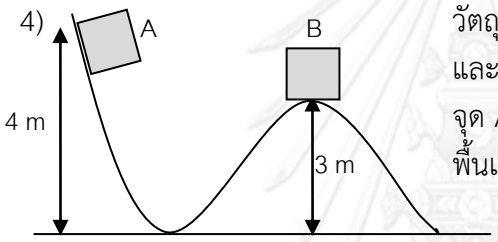
.....

.....

.....

.....

.....

4) 

วัตถุมวล 8 กิโลกรัม ที่จุด A มีอัตราเร็ว 0.5 เมตร/วินาที และที่จุด B มีอัตราเร็ว 9 เมตร/วินาที ถ้าระยะทางจากจุด A ไปยังจุด B เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (4 คะแนน)

.....

.....

.....

.....

.....

.....

.....

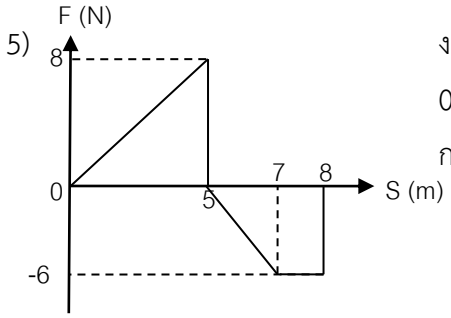
.....

.....

.....

.....

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

5) 

งานทั้งหมดที่กระทำในช่วงระยะทางการเคลื่อนที่ 0 เมตรถึง 8 เมตร จะมีค่ากี่จูล ถ้ากราฟระหว่างแรงกับการกระจัดเป็นดังรูป (4 คะแนน)

.....

.....

.....

.....

.....

.....

6) มอเตอร์ไซต์มวล 300 กิโลกรัม แล่นขึ้นพื้นเอียงทำมุม 30 องศากับแนวระดับ ด้วยอัตราเร็วคงที่ 120 เมตร/วินาที ถ้าแรงเสียดทานระหว่างพื้นเอียงกับมอเตอร์ไซต์เท่ากับ 200 นิวตัน จงหากำลังของเครื่องยนต์ในหน่วย<u>กิโลวัตต์</u> (4 คะแนน)

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

7) ดาวต้องการยิงลูกหินมวล 100 กรัม ด้วยหนังสติ๊กให้ขึ้นไปได้สูง 10 เมตรในแนวดิ่ง ถ้าการยืดของหนังสติ๊กเป็นแบบสปริงและยืดออก 50 เซนติเมตร หนังสติ๊กจะต้องมีค่านิจสปริงกี่นิวตัน/เมตร
(4 คะแนน)

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัมขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)

.....

.....

.....

.....

.....

.....

.....

.....

.....

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

แบบสอบ เรื่องงานและพลังงาน (ชุด 2)

การประเมินผลระหว่างเรียน ระดับชั้นมัธยมศึกษาปีที่ 4 ปีการศึกษา 2556
แบบสอบรายวิชา วิทยาศาสตร์ คะแนนเต็ม 48 คะแนน เวลา 90 นาที

.....

คำชี้แจง 1. แบบสอบฉบับนี้มีทั้งหมด 8 หน้า จำนวน 2 ตอน
ตอนที่ 1 เป็นข้อสอบปรนัย จำนวน 16 ข้อ รวม 16 คะแนน
ตอนที่ 2 เป็นข้อสอบอัตนัย จำนวน 8 ข้อ รวม 32 คะแนน
2. สามารถทดเลขลงในแบบสอบได้ **(ไม่อนุญาติให้ใช้เครื่องคิดเลข)**
3. ส่งกระดาษคำตอบคืนพร้อมกับแบบสอบ

กำหนดให้ g = 10 เมตร/วินาที^{2}

<u>ตอนที่ 1</u> ให้นักเรียนเลือกคำตอบที่ถูกต้องที่สุด แล้วทำเครื่องหมาย X ลงในกระดาษคำตอบ

1) กล่องเคลื่อนที่เป็นเส้นตรงบนพื้นราบลื่นด้วยแรง 20 นิวตัน เป็นระยะทาง S ถ้าต้องการหาค่างานของแรง 20 นิวตัน จะต้องพิจารณาสิ่งใดบ้าง
ก. มุมระหว่างทิศของแรงกับแนวการเคลื่อนที่ของวัตถุ
ข. แรงเสียดทานระหว่างวัตถุกับพื้นราบ
ค. อัตราเร็วของวัตถุ
1. ก 2. ข 3. ค 4. ข และ ค

2) มินออกแรง 100 นิวตันลากกล่องในแนวทำมุม 60 องศากับแนวระดับ ถ้าลากกล่องไปได้ไกล 10 เมตรในแนวระดับ งานของแรงที่ใช้ลากกล่องเป็นเท่าใด
1. 500 จูล 2. $500\sqrt{3}$ จูล 3. 1,000 จูล 4. $1{,}000\sqrt{3}$ จูล

3) ลิฟท์ตัวหนึ่งมวล 200 กิโลกรัม ใช้มอเตอร์ 8 กิโลวัตต์ ขณะที่บรรทุกคนให้เคลื่อนที่ขึ้นด้วยความเร็วคงที่ 2 เมตร/วินาที จะต้องใช้แรงเท่าใด
1. 4 นิวตัน 2. 25 นิวตัน 3. 100 นิวตัน 4. 4,000 นิวตัน

4) พลังงานจลน์ของวัตถุในกรณีใดมีค่าต่ำที่สุด
1. เมื่อเพิ่มมวลเป็น 2 เท่า 2. เมื่อเพิ่มความเร็วเป็น 2 เท่า
3. เมื่อลดมวลลง 1/2 เท่า 4. เมื่อลดความเร็วลง 1/2 เท่า

5) กันต์ปล่อยวัตถุมวล 0.5 กิโลกรัมจากตึกแห่งหนึ่ง จงหาพลังงานจลน์ของวัตถุ เมื่อวัตถุถูกปล่อยลงมาเป็นเวลา 2 วินาที
1. 5 จูล 2. 10 จูล 3. 100 จูล 4. 200 จูล

6) ยิงลูกปืนมวล 150 กรัมไปในแนวระดับ ลูกปืนมีความเร็ว 100 เมตร/วินาที ถ้าลูกปืนฝังเข้าไปในเนื้อไม้ งานที่ไม้ต้านการเคลื่อนที่ของลูกปืนมีค่าเท่าใด
1. 150 จูล 2. 750 จูล 3. 15,000 จูล 4. 75,000 จูล

7) ตั้มปล่อยถุงทราย 5 กิโลกรัมตกจากหอคอยสูง 7 เมตร ขณะที่ถุงทรายเคลื่อนที่ลงมาได้ระยะทาง 4 เมตรจากหอคอย ความเร็วของถุงทรายมีค่าเท่าใด
1. $\sqrt{20}$ เมตร/วินาที 2. 20 เมตร/วินาที 3. $\sqrt{60}$ เมตร/วินาที 4.60 เมตร/วินาที

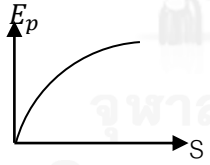
8) มอเตอร์เครื่องหนึ่งที่มีประสิทธิภาพ 90% จะให้พลังงานงานที่เป็นประโยชน์ออกมาเท่าใด เมื่อใส่กำลังให้แก่มอเตอร์เครื่องนี้ 0.5 กำลังม้า (1 กำลังม้า = 746 วัตต์)
1. 0.45 วัตต์ 2. 33.57 วัตต์ 3. 45.00 วัตต์ 4. 335.70 วัตต์

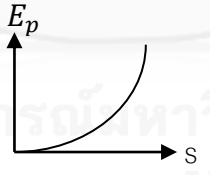
9) กรณีใดที่ถือว่ามีการทำงาน เนื่องจากการออกแรงกระทำต่อกล่อง
ก. นุชยืนถือกล่อง ข. นุชยกกล่องจากโต๊ะไปวางบนพื้นในแนวดิ่ง
ค. นุชยกกล่องจากพื้นไปวางบนโต๊ะในแนวดิ่ง
ง. นุชถือกล่องแล้วเดินไปตามพื้นราบ
1. ง 2. ข และ ค 3. ข ค และ ง 4. ก ค และ ง

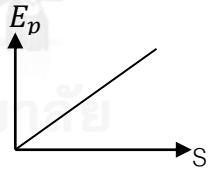
10) กล่องมวล 20 กิโลกรัมเคลื่อนที่ลงไปตามพื้นเอียงผิวขรุขระด้วยความเร็วคงที่ เป็นระยะทาง 5 เมตร แรงเสียดทานระหว่างพื้นและลังจะมีค่าเท่าใด เมื่องานในการเคลื่อนที่ของลังดังกล่าวเป็น 1,400 จูล ถ้าพื้นเอียงอยู่สูงจากพื้น 3 เมตร
1. - 80 นิวตัน 2. -160 นิวตัน 3. 80 นิวตัน 4.160 นิวตัน

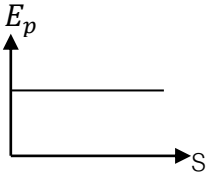
11) จ็อบออกแรงคงที่ขนาด 200 นิวตัน ผลักกล่องใบหนึ่งจากหยุดนิ่งให้เคลื่อนที่ไปตามแนวราบได้ระยะทาง 2 เมตร ภายในเวลา 4 วินาที จงหากำลังของจ็อบที่ใช้ในการผลักกล่องใบนี้
1. 100 วัตต์ 2. 400 วัตต์ 3. 1,200 วัตต์ 4.1,600 วัตต์

12) กราฟในตัวเลือกใดแสดงความสัมพันธ์ระหว่างพลังงานศักย์โน้มถ่วง (E_p) ของวัตถุกับการกระจัด (S) ได้ถูกต้อง เมื่อวัตถุนี้ถูกโยนขึ้นไปในแนวดิ่งจากจุดหยุดนิ่งด้วยอัตราเร็วคงที่

1. 

2. 

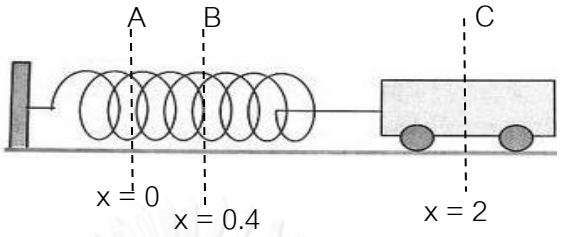
3. 

4. 

13) ท็อปปล่อยวัตถุมวล 3 กิโลกรัม จากตึกสูง 45 เมตรจากระดับพื้นดิน เมื่อกล่องหล่นมาได้ระยะทาง 25 เมตร พลังงานศักย์โน้มถ่วงมีค่าเท่าใด
1. 600 จูล 2. 750 จูล 3. 1,350 จูล 4. 2,100 จูล

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

14) แจ็คนำสปริงที่มีค่าคงที่ 20 นิวตัน/เมตร มาผูกปลายข้างหนึ่งกับรถทดลองมวล m ที่วางอยู่บนพื้นราบลื่นดังรูป สปริงยืดอยู่ก่อนแล้ว 0.4 เมตร ถ้าแจ็คดึงสปริงให้ยืดเป็น 2 เมตร จงหางานของแรงดึงกลับของสปริง



1. +38.4 จูล 2. -38.4จูล 3. -40.0 จูล 4. +40.0 จูล

15) เมื่อกล่องใบหนึ่งถูกปล่อยจากตึกสูงมายังพื้นราบ ข้อใดกล่าวถูกต้อง

1. พลังงานรวม พลังงานจลน์ พลังงานศักย์โน้มถ่วง ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน
2. พลังงานศักย์โน้มถ่วงของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน
3. พลังงานจลน์ของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน
4. พลังงานรวมของมวลก้อนนี้ ณ จุดต่างๆของการเคลื่อนที่มีค่าเท่ากัน

16) วัตถุมวล 36 กิโลกรัม เคลื่อนที่ด้วยความเร็ว 20 เมตร/วินาที เข้าชนสปริงที่มีค่าคงที่ 400 นิวตัน/เมตร สปริงจะหดสั้นที่สุดเท่าไหร่เมื่อกล่องชนสปริง

1. 6 ซม. 2. 36 ซม. 3. 600 ซม. 4. 3,600 ซม.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ชื่อ-นามสกุล .....ชั้น ม.4/..... เลขที่ ..... ชุดที่ 2

กระดาษคำตอบ

ตอนที่ 1 (แบบสอบปรนัยจำนวน 16 ข้อ รวม 16 คะแนน)

คำชี้แจง ให้นักเรียนเลือกคำตอบที่ถูกต้องที่สุด แล้วทำเครื่องหมาย X ลงในกระดาษคำตอบ

ข้อ	ตัวเลือก			
	1	2	3	4
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				

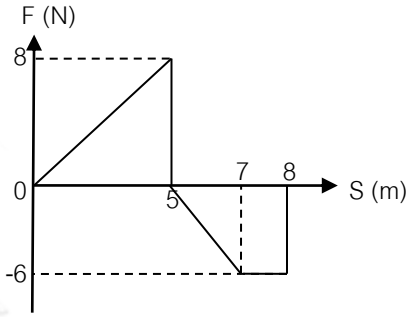
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ชื่อ-นามสกุล .....ชั้น ม.4/..... เลขที่ ..... ชุดที่ 2

ตอนที่ 2 (แบบสอบอัตนัยจำนวน 8 ข้อ รวม 32 คะแนน) **กำหนดให้** g = 10 เมตร/วินาที^{2}

คำชี้แจง จงแสดงวิธีทำให้ถูกต้อง โดยแสดงวิธีทำอย่างละเอียด (แสดงให้เห็นถึงสมการ การแทนค่า และคำตอบพร้อมระบุหน่วย)

1) งานทั้งหมดที่กระทำในช่วงระยะทางการเคลื่อนที่ 0 เมตรถึง 8 เมตร จะมีค่ากี่จูล ถ้ากราฟระหว่างแรงกับการกระจัดเป็นดังรูป (4 คะแนน)



.....

.....

.....

.....

.....

.....

2) มอเตอร์ไซต์มวล 300 กิโลกรัม แล่นขึ้นพื้นเอียงทำมุม 30 องศากับแนวระดับ ด้วยอัตราเร็วคงที่ 120 เมตร/วินาที ถ้าแรงเสียดทานระหว่างพื้นเอียงกับมอเตอร์ไซต์เท่ากับ 200 นิวตัน จงหากำลังของเครื่องยนต์ในหน่วยกิโลวัตต์ (4 คะแนน)

.....

.....

.....

.....

.....

.....

.....

.....

.....

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3) ดาวต้องการยิงลูกหินมวล 100 กรัม ด้วยหนังสติ๊กให้ขึ้นไปได้สูง 10 เมตรในแนวดิ่ง ถ้าการยืดของหนังสติ๊กเป็นแบบสปริงและยืดออก 50 เซนติเมตร หนังสติ๊กจะต้องมีค่านิจสปริงกี่นิวตัน/เมตร
(4 คะแนน)

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

4) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัมขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)

.....

.....

.....

.....

.....

.....

.....

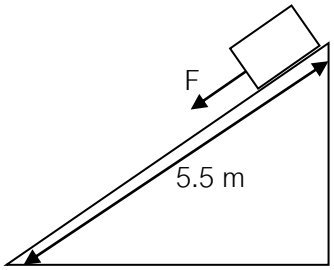
.....

.....

.....

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

5) ชายคนหนึ่งออกแรงดึงกล่องหนัก 80 กิโลกรัมลงเนินที่เอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ ถ้าเขาออกแรงดึงกล่องในแนวขนานกับการเคลื่อนที่ดังรูป เขาต้องใช้งานเท่าใดในการดึงกล่องลงมาได้ 5.5 เมตร เมื่อสัมประสิทธิ์ความเสียดทานระหว่างเนินและกล่องเท่ากับ $\sqrt{3}/2$ (4 คะแนน)



.....

.....

.....

.....

.....

.....

.....

.....

6) สปริงถูกดึงด้วยแรง 40 นิวตัน จะทำให้สปริงยืดออก 20 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 30 เซนติเมตรในหน่วยจูล (4 คะแนน)

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

7) เครื่องสูบน้ำใช้สูบน้ำมวล 500 กิโลกรัม ขึ้นจากบ่อลึก 6 เมตร แล้วฉีดน้ำออกจากปลายท่อด้วยความเร็ว 6 เมตรต่อวินาที งานที่ได้จากเครื่องสูบน้ำนี้มีค่ากี่กิโลจูล (4 คะแนน)

.....

.....

.....

.....

.....

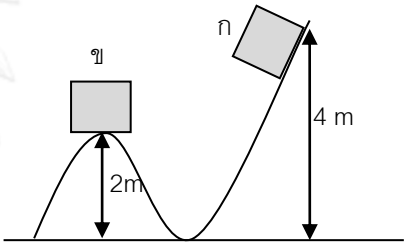
.....

.....

.....

.....

8) กล่องมวล 6 กิโลกรัม ที่จุด ก มีอัตราเร็ว 1 เมตร/วินาที และที่จุด ข มีอัตราเร็ว 7 เมตร/วินาที ถ้าระยะทางจากจุด ก ไปยังจุด ข เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (4 คะแนน)



.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

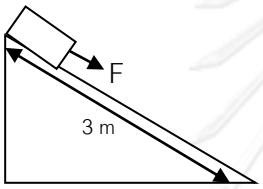
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ฉ

ผลการตรวจสอบสอดคล้องของเกณฑ์การให้คะแนนของข้อสอบแบบสร้างคำตอบ สำหรับแบบสอบฉบับใหม่/ฉบับอ้างอิงกับข้อสอบและเฉลยของข้อสอบ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 46 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับใหม่ (ครั้งที่ 1)

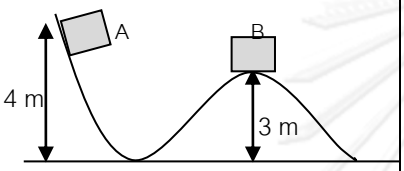
ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
1) ช้างตัวหนึ่งออกแรงดึงซุงหนัก 140 กิโลกรัม ลงเนินซึ่งเอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ สัมประสิทธิ์ความเสียดทานระหว่างเนินและซุงเป็น $\sqrt{3}/2$ ถ้าแรงที่ใช้ดึงซุงขนานกับการเคลื่อนที่ดังรูป ช้างต้องใช้งานเท่าใดในการดึงซุงลงมาได้ 3 เมตร  <u>วิธีทำ</u> เนื่องจากความเร็วคงที่ แรงลัพธ์ในการเคลื่อนที่จึงเท่ากับศูนย์ นั่นคือ $\sum \bar{F} = 0$ แรงซ้าย = แรงขวา $F + mgsin\theta = f$ $F + mgsin\theta = \mu mgcos\theta$ จะได้ว่า $F = \mu mgcos30° - mgsin30°$ $F = \frac{\sqrt{3}}{2}(140)(10)\left(\frac{\sqrt{3}}{2}\right) - (140)(10)\left(\frac{1}{2}\right)$ เพราะฉะนั้น $F = 1050 - 700 = 350$ จะได้ว่า $W = Fscos\theta$ $W = 350(3)cos0° = 1{,}050$	4 คะแนน : หาค่างานพร้อมทั้งระบุหน่วยได้ถูกต้อง 3 คะแนน : หาค่าแรงได้ถูกต้อง 2 คะแนน : แทนค่าสมการของแรงลัพธ์ได้ถูกต้องทั้งหมด 1 คะแนน : ระบุสมการของงานของแรงลัพธ์ได้ถูกต้องทั้งหมด	1	-1	0	0	ระดับการประเมินประเมินได้ไม่สอดคล้องกับข้อสอบและเฉลยของข้อสอบ โดยผู้เชี่ยวชาญให้คำแนะนำว่า เกณฑ์การให้คะแนนมีความถูกต้องในส่วนขององค์ประกอบต้องมีทั้งการกำหนดสูตร การแทนค่า และการหาคำตอบ แต่คำอธิบายในแต่ละระดับคะแนนยังกว้าง ไม่ชัดเจน โดยต้องปรับการเขียนคำอธิบายในระดับคะแนนต่างๆให้เป็นไปตามหลักของ rubric score กล่าวคือ 4 คะแนนต้องมีครบคุณสมบัติครบทั้งหมดแล้วค่อยๆลดหลั่นคุณสมบัติไปตามระดับคะแนน อย่างเช่น 4 คะแนนต้องมีทั้ง 4 คุณสมบัติ ขณะที่ 3 คะแนน ต้องมีคุณสมบัติ 1,2,3 แต่ขาดคุณสมบัติ 4 ส่วนระดับคะแนน 2 ต้องมีคุณสมบัติ 1,2 แต่ขาดคุณสมบัติ 3 และ 4ขณะที่ 1 คะแนน ต้องมีคุณสมบัติ 1 แต่ขาดคุณสมบัติ 2,3,4 เป็นต้น

ตาราง 46 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับใหม่ (ครั้งที่ 1) (ต่อ)

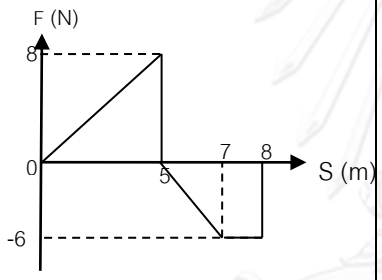
ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
2) สปริงถูกออกแรงดึง 20 นิวตัน จะทำให้ยืดออก 10 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 20 เซนติเมตรในหน่วยจูล <u>วิธีทำ</u> หาค่านิจสปริงจาก $F = kx$ แทนค่า $20 = k(0.1)$ จะได้ว่า $k = \frac{20}{0.1} = 200$ N/m หาพลังงานศักย์ยืดหยุ่น (E_P) จาก ; $E_P = \frac{1}{2}kx^2$ แทนค่า $= \frac{1}{2}(200)(0.2)^2$ $E_P = \frac{1}{2}(200)(0.04) = 4$ จูล **ดังนั้น** พลังงานศักย์ยืดหยุ่นของสปริงจะเท่ากับ 4 จูล	4 คะแนน : หาค่าสมการพลังงานศักย์ยืดหยุ่นพร้อมระบุหน่วยได้ ถูกต้อง 3 คะแนน : แทนค่าสมการพลังงานศักย์ยืดหยุ่นได้ถูกต้อง 2 คะแนน : คำนวณค่านิจของสปริงได้ถูกต้อง 1 คะแนน : แทนค่าสมการของแรงในสปริงได้ถูกต้อง	1	-1	0	0	ระดับการประเมินประเมินได้ไม่สอดคล้องกับข้อสอบและเฉลยของข้อสอบ โดยผู้เชี่ยวชาญให้คำแนะนำว่า เกณฑ์การให้คะแนนมีความถูกต้องในส่วนขององค์ประกอบต้องมีทั้งการกำหนดสูตร การแทนค่า และการหาคำตอบ แต่คำอธิบายในแต่ละระดับคะแนนยังกว้าง ไม่ชัดเจน โดยต้องปรับการเขียนคำอธิบายในระดับคะแนนต่างๆให้เป็นไปตามหลักของ rubric score เช่นเดียวกับคำแนะนำในข้อ 1
3) เคนนำเครื่องสูบน้ำเครื่องหนึ่งไปสูบน้ำมวล 400 กิโลกรัม ขึ้นไปตามท่อสูง 10 เมตร ถ้าน้ำพุ่งออกจากปลายท่อด้วยความเร็ว 10 เมตรต่อวินาที งานที่ได้จากเครื่องสูบมีค่ากี่กิโลจูล (4 คะแนน) <u>วิธีทำ</u> เนื่องจากงานในการเคลื่อนที่สามารถหาได้จากพลังงานที่เปลี่ยนแปลงไป จะได้ว่า $W = mgh + \frac{1}{2}mv^2$ แทนค่า $W = (400)(10)(10) + \frac{1}{2}(400)(10)^2$ $= 40{,}000 + 20{,}000$ $= 60{,}000$ จูล **ดังนั้น**งานมีค่า 60 กิโลจูล	4 คะแนน : หาค่างานในหน่วยกิโลจูลได้ถูกต้อง 3 คะแนน : แทนค่าพลังงานศักย์โน้มถ่วง<u>และ</u>พลังงานจลน์ได้ถูกต้อง 2 คะแนน : แทนค่าพลังงานศักย์โน้มถ่วง<u>หรือ</u>พลังงานจลน์ ได้ถูกต้อง 1 คะแนน : ระบุสมการของงานได้ถูกต้องทั้ง 2 เทอม	1	1	0	0.67	ระดับการประเมินประเมินได้**สอดคล้อง**กับข้อสอบและเฉลยของข้อสอบ แต่ต้องมีการปรับการเขียนคำอธิบายในระดับคะแนนต่างๆให้เป็นไปตามหลักของ rubric score เช่นเดียวกับคำแนะนำในข้อ 1

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 46 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับใหม่ (ครั้งที่ 1) (ต่อ)

ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
4) วัตถุมวล 8 กิโลกรัม ที่จุด A มีอัตราเร็ว 0.5 เมตร/วินาที และที่จุด B มีอัตราเร็ว 9 เมตร/วินาที ถ้าระยะทางจากจุด A ไปยังจุด B เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน)  <u>วิธีทำ</u> (แบบที่ 1) เมื่อกำหนดให้เนินที่จุด B เป็นจุดอ้างอิง จะได้ว่า $E_{P1} + E_{k1} - W_f = E_{k2}$ จาก $mgh + \frac{1}{2}mv^2 - fs = \frac{1}{2}mv^2$ แทนค่า $(8)(10)(1) + \frac{1}{2}(8)(0.5)^2 - f(3) = \frac{1}{2}(8)(9)^2$ จะได้ว่า $80 + 1 - f(3) = 3$ $-f(3) = 324 - 8$ $f = -\frac{243}{3} = -81$ นิวตัน	4 คะแนน : หาค่าแรงเสียดทานพร้อมทั้งระบุหน่วยได้ถูกต้อง 3 คะแนน : แทนค่าสมการของพลังงานได้ถูกต้องทั้งหมด 2 คะแนน : แทนค่าสมการของพลังงานได้ถูกต้องอย่างต่ำ 2 เทอม 1 คะแนน : ระบุสมการพลังงานได้ถูกต้องทั้ง 4 เทอม	1	-1	0	0	ระดับการประเมินประเมินได้ไม่สอดคล้องกับข้อสอบและเฉลยของข้อสอบ โดยผู้เชี่ยวชาญให้คำแนะนำว่า เกณฑ์การให้คะแนนมีความถูกต้องในส่วนขององค์ประกอบที่ต้องมีทั้งการกำหนดสูตร การแทนค่า และการหาคำตอบ แต่คำอธิบายในแต่ละระดับคะแนนยังกว้าง ไม่ชัดเจน โดยต้องปรับการเขียนคำอธิบายในระดับคะแนนต่างๆให้เป็นไปตามหลักของ rubric score เช่นเดียวกับคำแนะนำในข้อ 1
<u>วิธีทำ</u> (แบบที่ 2) เมื่อกำหนดให้เนินที่จุด B เป็นจุดอ้างอิง จะได้ว่า $E_{P1} + E_{k1} - W_f = E_{k2} + E_{P2}$ จาก $mgh + \frac{1}{2}mv^2 - fs = \frac{1}{2}mv^2 + mgh$ แทนค่า $(8)(10)(4) + \frac{1}{2}(8)(0.5)^2 - f(3) = \frac{1}{2}(8)(9)^2 + (8)(10)(3)$	4 คะแนน : หาค่าแรงเสียดทานพร้อมทั้งระบุหน่วยได้ถูกต้อง 3 คะแนน : แทนค่าสมการของพลังงานได้ถูกต้องทั้งหมด 2 คะแนน : แทนค่าสมการของพลังงานได้ถูกต้องอย่างต่ำ 2 เทอม 1 คะแนน : ระบุสมการ	1	-1	**0**	**0**	ระดับการประเมินประเมินได้ไม่สอดคล้องกับข้อสอบและเฉลยของข้อสอบ โดยผู้เชี่ยวชาญให้คำแนะนำว่า เกณฑ์การให้คะแนนมีความถูกต้องในส่วนขององค์ประกอบที่ต้องมีทั้งการกำหนดสูตร การแทนค่า และการหาคำตอบ แต่คำอธิบายในแต่ละระดับคะแนนยัง

ตาราง 46 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับใหม่ (ครั้งที่ 1) (ต่อ)

ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
จะได้ว่า $320 + 1 - f(3) = 324 + 240$ $-f(3) = 243$ $f = -\frac{243}{3} = -81$ นิวตัน	พลังงานได้ถูกต้องทั้ง 4 เทอม					กว้าง ไม่ชัดเจน โดยต้องปรับการเขียนคำอธิบายในระดับคะแนนต่างๆให้เป็นไปตามหลักของ rubric score เช่นเดียวกับคำแนะนำในข้อ 1
5) งานทั้งหมดที่กระทำในช่วงระยะทางการเคลื่อนที่ 0 เมตรถึง 10 เมตร จะมีค่ากี่จูล ถ้ากราฟระหว่างแรงกับการกระจัดเป็นดังรูป  <u>วิธีทำ</u> เนื่องจากงานสามารถหาได้จากแรงคูณกับระยะทางในแนวเดียวกับการเคลื่อนที่ นั่นคือ $W = Fs$ โจทย์ให้กราฟระหว่างแรง (F) กับระยะทางการเคลื่อนที่ (S) ซึ่งพื้นที่ใต้กราฟระหว่างแรงกับระยะทางการเคลื่อนที่ก็คืองาน นั่นเอง จะได้ว่า งานทั้งหมด = พื้นที่ใต้กราฟระหว่างแรงกับระยะทางการเคลื่อนที่ $= \frac{1}{2}(8)5 + \frac{1}{2}(-6)(3+1)$ = 20 - 12 = 8 จูล **ดังนั้น** งานทั้งหมดที่กระทำในช่วงระยะทางการเคลื่อนที่ 0 เมตรถึง 10 เมตร จะเท่ากับ 8 จูล	4 คะแนน : หาค่างานได้ถูกต้อง 3 คะแนน : แทนค่าพื้นที่ใต้กราฟส่วนบน<u>และ</u>ส่วนล่างได้ถูกต้อง 2 คะแนน : แทนค่าพื้นที่ใต้กราฟส่วนบน<u>หรือ</u>ส่วนล่างได้ถูกต้อง 1 คะแนน : ระบุได้ว่างานของการเคลื่อนที่สามารถหาได้จากพื้นที่ใต้ กราฟระหว่างแรงกับการกระจัด <u>หรือ</u>ระบุได้ว่างานทั้งหมด= พื้นที่ใต้กราฟระหว่าง F-S ด้านบน - พื้นที่ใต้กราฟระหว่าง F-S ด้านลบ	1	1	1	1	ระดับการประเมินประเมินได้**สอดคล้อง**กับข้อสอบและเฉลยของข้อสอบ

ตาราง 46 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับใหม่ (ครั้งที่ 1) (ต่อ)

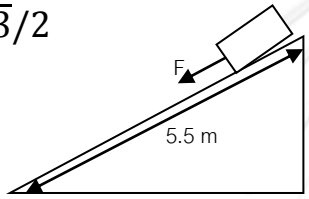
ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
6) มอเตอร์ไซต์มวล 300 กิโลกรัม แล่นขึ้นพื้นเอียงทำมุม 30 องศากับแนวระดับ ด้วยอัตราเร็วคงที่ 120 เมตร/วินาที ถ้าแรงเสียดทานระหว่างพื้นเอียงกับรถยนต์เท่ากับ 200 นิวตัน จงหากำลังงานของเครื่องยนต์ในหน่วยกิโลวัตต์ วิธีทำ ความเร็วคงที่ แรงลัพธ์ในการเคลื่อนที่จึงเท่ากับศูนย์ นั่นคือ $\sum \bar{F} = mgsin30° + f$ $= (300)(10)\left(\frac{1}{2}\right) + 200$ $= 1500 + 100 = 1{,}700$ นิวตัน หากำลังได้จาก $P = Fv = (1{,}700)(120)$ $= 204{,}000$ **ดังนั้น** กำลังเท่ากับ 204 กิโลวัตต์	4 คะแนน : แทนค่าสมการของกำลังได้ถูกต้อง และหาค่าแรงลัพธ์และค่ากำลังพร้อมทั้งระบุหน่วยได้ถูกต้อง 3 คะแนน : หาค่าแรงลัพธ์ได้ถูกต้อง 2 คะแนน : แทนค่าสมการของแรงลัพธ์ได้ถูกต้อง 1 คะแนน : ระบุสมการของแรงลัพธ์ได้ถูกต้องทั้งหมด	1	1	0	0.67	ระดับการประเมินประเมินได้**สอดคล้อง**กับข้อสอบและเฉลยของข้อสอบ แต่ต้องมีการปรับการเขียนคำอธิบายในระดับคะแนนต่างๆให้เป็นตามหลักของ rubric score เช่นเดียวกับคำแนะนำในข้อ 1
7) ดาวต้องการยิงลูกหินมวล 100 กรัม ด้วยหนังสติ๊กให้ขึ้นไปได้สูง 10 เมตรในแนวดิ่ง ถ้าการยืดของหนังสติ๊กเป็นแบบสปริงและยืดออก 50 เซนติเมตร หนังสติ๊กจะต้องมีค่านิจสปริงกี่นิวตัน/เมตร วิธีทำ จากกฎการอนุรักษ์พลังงานกล่าวว่า พลังงานรวมของวัตถุ ณ ตำแหน่งใดๆย่อมมีความค่าคงตัว นั่นคือ $\sum E_{\text{ตอนแรก}} = \sum E_{\text{ตอนหลัง}}$ จะได้ว่า $mgh = \frac{1}{2}kx^2$ แทนค่า $(0.1)(10)(10) = \frac{1}{2}k(0.5)^2$ $10 = \frac{1}{2}k(0.25)$ เพราะฉะนั้น $k = \frac{10(2)}{0.25} = 80$ N/m	4 คะแนน : คำนวณค่านิจสปริงพร้อมระบุหน่วยได้ถูกต้อง 3 คะแนน : แทนค่าพลังงานศักย์โน้มถ่วง<u>และ</u>พลังงานจลน์ได้ถูกต้อง 2 คะแนน : แทนค่าพลังงานศักย์โน้มถ่วง<u>หรือ</u>พลังงานจลน์ ได้ถูกต้อง 1 คะแนน : ระบุสมการของพลังงานได้ถูกต้องทั้ง 2 เทอม (พลังงานศักย์โน้มถ่วงและศักย์ยืดหยุ่น)	1	-1	0	0	ระดับการประเมินประเมินได้ไม่สอดคล้องกับข้อสอบและเฉลยของข้อสอบ โดยผู้เชี่ยวชาญให้คำแนะนำว่า เกณฑ์การให้คะแนนมีความถูกต้องในส่วนขององค์ประกอบต้องมีทั้งการกำหนดสูตร การแทนค่า และการหาคำตอบ แต่คำอธิบายในแต่ละระดับคะแนนยังกว้าง ไม่ชัดเจน โดยต้องปรับการเขียนคำอธิบายในระดับคะแนนต่างๆให้เป็นตามหลักของ rubric score เช่นเดียวกับคำแนะนำในข้อ 1

ตาราง 46 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับใหม่ (ครั้งที่ 1) (ต่อ)

ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัม ขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน) วิธีทำ ประสิทธิภาพของเครื่องกล (E_{ff}) หาได้จาก $E_{ff} = \frac{\text{กำลังที่ได้จริง}}{\text{กำลังที่ได้รับ}} \times 100$ $= \frac{\frac{mgh}{t}}{\text{กำลังที่ได้รับ}} \times 100$ $= \frac{\frac{(750+450)(10)(8)}{4}}{30{,}000} \times 100$ $= \frac{24{,}000}{30{,}000} \times 100$ $= 0.8 \times 100 = 80\%$ ดังนั้น ลิฟท์นี้ต้องมีประสิทธิภาพ 80%	4 คะแนน : หาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง 3 คะแนน แทนค่าในสมการประสิทธิภาพเครื่องกลได้ถูกต้อง 2 คะแนน : ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้องและแทนค่ากำลังที่ได้จริงถูกต้องทั้งหมด 1 คะแนน : ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้องหรือแทนค่ากำลังที่ได้จริงถูกต้อง	1	1	0	0.67	ระดับการประเมินประเมินได้**สอดคล้อง**กับข้อสอบและเฉลยของข้อสอบ แต่ต้องมีการปรับการเขียนคำอธิบายในระดับคะแนนต่างๆให้เป็นเป็นตามหลักของ rubric score เช่นเดียวกับคำแนะนำในข้อ 1

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

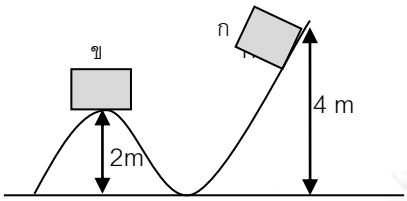
ตาราง 47 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับอ้างอิง (ครั้งที่ 1)

ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
1) ชายคนหนึ่งออกแรงดึงกล่องหนัก 80 กิโลกรัมลงเนินที่เอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ ถ้าเขาออกแรงดึงกล่องในแนวขนานกับการเคลื่อนที่ดังรูป เขาต้องใช้งานเท่าใดในการดึงกล่องลงมาได้ 5.5 เมตร เมื่อสัมประสิทธิ์ความเสียดทานระหว่างเนินและกล่องเท่ากับ $\sqrt{3}/2$  <u>วิธีทำ</u> เนื่องจากความเร็วคงที่ แรงลัพธ์ในการเคลื่อนที่จึงเท่ากับศูนย์ นั่นคือ $\sum \bar{F} = 0$ แรงซ้าย = แรงขวา $F + mgsin\theta = f$ $F + mgsin\theta = \mu mgcos\theta$ จะได้ว่า $F = \mu mgcos30° - mgsin30°$ $= \frac{\sqrt{3}}{2}(80)(10)\left(\frac{\sqrt{3}}{2}\right) - (80)(10)(\frac{1}{2})$ เพราะฉะนั้น $F = 600 - 400 = 200$ $W = 200(5.5)cos0°$ $W = 1{,}100$ จูล	4 คะแนน : หาค่างานพร้อมทั้งระบุหน่วยได้ถูกต้อง 3 คะแนน : หาค่าแรงได้ถูกต้อง 2 คะแนน : แทนค่าสมการของแรงลัพธ์ได้ถูกต้องทั้งหมด 1 คะแนน : ระบุสมการของงานของแรงลัพธ์ได้ถูกต้องทั้งหมด	1	-1	0	0	ระดับการประเมิน ประเมินได้ไม่สอดคล้องกับข้อสอบและเฉลยของข้อสอบ โดยผู้เชี่ยวชาญให้คำแนะนำว่าเกณฑ์การให้คะแนนมีความถูกต้องในส่วนขององค์ประกอบต้องมีทั้งการกำหนดสูตร การแทนค่าและการหาคำตอบ ซึ่งสอดคล้องกับวิธีทำ แต่ต้องมีการปรับการเขียนคำอธิบายในระดับคะแนนต่างๆให้เป็นเป็นตามหลักของ rubric score เช่นเดียวกับคำแนะนำในข้อ 1 ในฉบับใหม่
2) สปริงถูกดึงด้วยแรง 40 นิวตัน จะทำให้สปริงยืดออก 20 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 30 เซนติเมตรในหน่วยจูล <u>วิธีทำ</u> หาค่านิจสปริงจาก $F = kx$ แทนค่า $40 = k(0.2)$	4 คะแนน : หาค่าสมการพลังงานศักย์ยืดหยุ่นพร้อมระบุหน่วยได้ ถูกต้อง 3 คะแนน : แทนค่าสมการพลังงานศักย์ยืดหยุ่นได้ถูกต้อง 2 คะแนน : คำนวณค่านิจของสปริงได้ถูกต้อง	1	-1	0	0	ระดับการประเมิน ประเมินได้ไม่สอดคล้องกับข้อสอบและเฉลยของข้อสอบ โดยผู้เชี่ยวชาญให้คำแนะนำว่าเกณฑ์การให้คะแนนมีความถูกต้องในส่วนขององค์ประกอบต้องมีทั้งการกำหนดสูตร การแทนค่าและการหาคำตอบ ซึ่ง

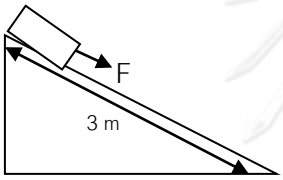
ตาราง 47 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับอ้างอิง (ครั้งที่ 1) (ต่อ)

ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
จะได้ว่า $k = \frac{40}{0.2} = 200$ N/m หาพลังงานศักย์ยืดหยุ่น (E_P) จาก $E_P = \frac{1}{2}kx^2$ แทนค่า $= \frac{1}{2}(200)(0.3)^2$ $= \frac{1}{2}(200)(0.09) = 9$ จูล **ดังนั้น** พลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 20 เซนติเมตรจะเท่ากับ 9 จูล	1 คะแนน : แทนค่าสมการของแรงในสปริงได้ถูกต้อง	1	-1	0	0	สอดคล้องกับวิธีทำ แต่ต้องมีการปรับการเขียนคำอธิบายในระดับคะแนนต่างๆให้เป็นเป็นตามหลักของ rubric score เช่นเดียวกับคำแนะนำในข้อ 1 ในฉบับใหม่
3) เครื่องสูบน้ำใช้สูบน้ำมวล 500 กิโลกรัม ขึ้นจากบ่อลึก 6 เมตร แล้วฉีดน้ำออกจากปลายท่อด้วยความเร็ว 6 เมตรต่อวินาที งานที่ได้จากเครื่องสูบน้ำนี้มีค่ากี่กิโลจูล <u>วิธีทำ</u>เนื่องจากงานในการเคลื่อนที่สามารถหาได้จากพลังงานที่เปลี่ยนแปลง จะได้ว่า $W = mgh + \frac{1}{2}mv^2$ แทนค่า $= (500)(10)(6) + \frac{1}{2}(500)(6)^2$ $= 30{,}000 + 9{,}000$ $= 39{,}000$ จูล **ดังนั้น** งานที่ได้จากเครื่องสูบมีค่า 39,000 จูล หรือ 39 กิโลจูล	4 คะแนน : หาค่างานในหน่วยกิโลจูลได้ถูกต้อง 3 คะแนน : แทนค่าพลังงานศักย์โน้มถ่วง <u>และ</u>พลังงานจลน์ได้ถูกต้อง 2 คะแนน : แทนค่าพลังงานศักย์โน้มถ่วง <u>หรือ</u>พลังงานจลน์ ได้ถูกต้อง 1 คะแนน : ระบุสมการของงานได้ถูกต้องทั้ง 2 เทอม	1	1	1	1	ระดับการประเมิน ประเมินได้**สอดคล้อง**กับข้อสอบและเฉลยของข้อสอบ
4) กล่องมวล 6 กิโลกรัม ที่จุด ก มีอัตราเร็ว 1 เมตร/วินาที และที่จุด ข มีอัตราเร็ว 7 เมตร/วินาที ถ้าระยะทางจากจุด ก ไปยังจุด ข เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน)	4 คะแนน : หาค่าแรงเสียดทานพร้อมทั้งระบุหน่วยได้ถูกต้อง 3 คะแนน : แทนค่าสมการของพลังงานได้ถูกต้องทั้งหมด 2 คะแนน : แทนค่าสมการของพลังงานได้ถูกต้องอย่างต่ำ 2เทอม	1	-1	0	0	ระดับการประเมิน ประเมินได้ไม่สอดคล้องกับข้อสอบและเฉลยของข้อสอบ โดยผู้เชี่ยวชาญให้คำแนะนำว่าเกณฑ์การให้คะแนนมีความถูกต้องในส่วนขององค์ประกอบต้องมีทั้งการกำหนดสูตร การแทนค่าและการหาคำตอบ ซึ่ง

ตาราง 47 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับอ้างอิง (ครั้งที่ 1) (ต่อ)

ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
 วิธีทำ (แบบที่ 1) กำหนดจุด ข เป็นจุดอ้างอิง จะได้ว่า $E_{P1} + E_{k1} - W_f = E_{k2}$ จาก $mgh + \frac{1}{2}mv^2 - fs = \frac{1}{2}mv^2$ แทนค่า $(6)(10)(2) + \frac{1}{2}(6)(1)^2 - f(3) = \frac{1}{2}(6)(7)^2$ จะได้ว่า $120 + 3 - f(3) = 147$ $-f(3) = 147 - 123$ $f = -\frac{24}{3} = -8$ นิวตัน ดังนั้นแรงเสียดทานมีค่า - 8 นิวตัน	1 คะแนน : ระบุสมการพลังงานได้ถูกต้องทั้ง 4 เทอม	1	-1	0	0	สอดคล้องกับวิธีทำ แต่ต้องมีการปรับการเขียนคำอธิบายในระดับคะแนนต่างๆให้เป็นเป็นตามหลักของ rubric score เช่นเดียวกับคำแนะนำในข้อ 1 ในฉบับใหม่
วิธีทำ (แบบที่ 2) เมื่อกำหนดให้เนินที่จุด B เป็นจุดอ้างอิง จะได้ว่า $E_{P1} + E_{k1} - W_f = E_{k2} + E_{P2}$ จาก $mgh + \frac{1}{2}mv^2 - fs = \frac{1}{2}mv^2 + mgh$ แทนค่า $(6)(10)(4) + \frac{1}{2}(6)(1)^2 - f(3) = \frac{1}{2}(6)(7)^2 + (6)(10)(2)$ จะได้ว่า $240 + 3 - f(3) = 147 + 120$ $-f(3) = 2$ $f = -\frac{24}{3} = -8$ นิวตัน **ดังนั้น** แรงเสียดทานระหว่างพื้นและกล่องมีค่า - 8 นิวตัน	4 คะแนน : หาค่าแรงเสียดทานพร้อมทั้งระบุหน่วยได้ถูกต้อง 3 คะแนน : แทนค่าสมการของพลังงานได้ถูกต้องทั้งหมด 2 คะแนน : แทนค่าสมการของพลังงานได้ถูกต้องอย่างต่ำ 2 เทอม 1 คะแนน : ระบุสมการพลังงานได้ถูกต้องทั้ง 4 เทอม	1	-1	0	0	ระดับการประเมิน ประเมินได้ไม่สอดคล้องกับข้อสอบและเฉลยของข้อสอบ โดยผู้เชี่ยวชาญให้คำแนะนำว่าเกณฑ์การให้คะแนนมีความถูกต้องในส่วนขององค์ประกอบต้องมีทั้งการกำหนดสูตร การแทนค่า และการหาคำตอบ ซึ่งสอดคล้องกับวิธีทำ แต่ต้องมีการปรับการเขียนคำอธิบายในระดับคะแนนต่างๆให้เป็นเป็นตามหลักของ rubric score เช่นเดียวกับคำแนะนำในข้อ 1 ในฉบับใหม่

ตาราง 48 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับใหม่และฉบับอ้างอิง ครั้งที่ 2

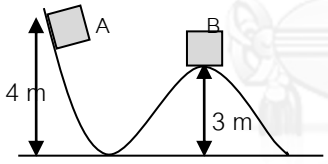
ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
1) ช้างตัวหนึ่งออกแรงดึงซุงหนัก 140 กิโลกรัม ลงเนินซึ่งเอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ สัมประสิทธิ์ความเสียดทานระหว่างเนินและซุงเป็น $\sqrt{3}/2$ ถ้าแรงที่ใช้ดึงซุงขนานกับการเคลื่อนที่ดังรูป ช้างต้องใช้งานเท่าใดในการดึงซุงลงมาได้ 3 เมตร  <u>วิธีทำ</u> เนื่องจากความเร็วคงที่ แรงลัพธ์ในการเคลื่อนที่จึงเท่ากับศูนย์ นั่นคือ $\sum \bar{F} = 0$ แรงซ้าย = แรงขวา $F + mgsin\theta = f$ $F + mgsin\theta = \mu mgcos\theta$ จะได้ว่า $F = \mu mgcos30° - mgsin30°$ $= \frac{\sqrt{3}}{2}(140)(10)\left(\frac{\sqrt{3}}{2}\right) - (140)(10)(\frac{1}{2})$ เพราะฉะนั้น $F = 1050 - 700 = 350$ จะได้ว่า $W = Fscos\theta$ $= 350(3)cos0° = 1{,}050$	4 คะแนน : ระบุสมการในการหางานและสมการของแรงลัพธ์ได้ถูกต้อง แทนค่าสมการของแรงลัพธ์และสมการของงานได้ถูกต้อง และหาค่างานพร้อมทั้งระบุหน่วยได้ถูกต้อง 3 คะแนน : ระบุสมการในการหางานและสมการของแรงลัพธ์ได้ถูกต้อง แทนค่าสมการของแรงลัพธ์และสมการของงานได้ถูกต้อง แต่ไม่สามารถหาค่างานได้ถูกต้อง 2 คะแนน : ระบุสมการในการหางานและสมการของแรงลัพธ์ได้ถูกต้อง และแทนค่าสมการของแรงลัพธ์ได้ถูกต้อง แต่ไม่สามารถแทนค่าสมการของงานได้ถูกต้อง และไม่สามารถหาค่างานได้ถูกต้อง 1 คะแนน : ระบุสมการในการหางานและสมการของแรงลัพธ์ได้ถูกต้อง แต่ไม่สามารถแทนค่าสมการของแรงลัพธ์และสมการของงานได้ถูกต้อง และไม่สามารถหาค่าแรงลัพธ์งานได้ถูกต้อง	1	1	1	1	ระดับการประเมินประเมินได้**สอดคล้อง**กับข้อสอบและเฉลยของข้อสอบ แต่อย่างไรก็ตามผู้เชี่ยวชาญได้แนะนำว่าในระดับคะแนน 1 นักเรียนควรระบุสมการของงานได้ถูกต้อง แต่ยังมาต้องถึงการระบุสมการของแรงลัพธ์ได้ถูกต้อง เพราะบทนี้เน้นไปที่เรื่องงานและพลังงานเป็นหลัก ซึ่งผู้วิจัยเห็นด้วยและปรับตามคำแนะนำดังกล่าว

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

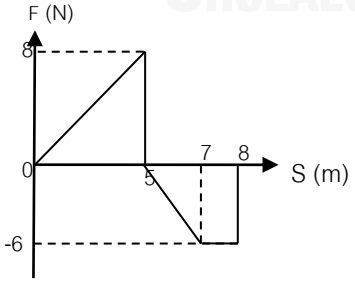
ตาราง 48 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับใหม่และฉบับอ้างอิง ครั้งที่ 2 (ต่อ)

ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
2) สปริงถูกออกแรงดึง 20 นิวตัน จะทำให้ยืดออก 10 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 20 เซนติเมตรในหน่วยจูล วิธีทำ หาค่านิจสปริงจาก $F = kx$ แทนค่า $20 = k(0.1)$ จะได้ว่า $k = \frac{20}{0.1} = 200$ N/m หาพลังงานศักย์ยืดหยุ่น (E_P) จาก ; $E_P = \frac{1}{2}kx^2$ แทนค่า $= \frac{1}{2}(200)(0.2)^2$ $E_P = \frac{1}{2}(200)(0.04) = 4$ จูล **ดังนั้น** พลังงานศักย์ยืดหยุ่นของสปริงจะเท่ากับ 4 จูล	4 คะแนน : ระบุสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง แทนค่าสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง และหาค่าพลังงานศักย์ยืดหยุ่นได้ถูกต้อง 3 คะแนน : ระบุสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง แทนค่าสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง แต่ไม่สามารถหาค่าพลังงานศักย์ยืดหยุ่นได้ถูกต้อง 2 คะแนน : ระบุสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง แทนค่าสมการตามกฎของฮุคได้ถูกต้อง แต่ไม่สามารถแทนค่าสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง และไม่สามารถหาค่าพลังงานศักย์ยืดหยุ่นได้ถูกต้อง 1 คะแนน : ระบุสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง แต่ไม่สามารถแทนค่าสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง และไม่สามารถหาค่างานได้ถูกต้อง	1	1	1	1	ระดับการประเมินประเมินได้**สอดคล้อง**กับข้อสอบและเฉลยของข้อสอบ
3) เคนนำเครื่องสูบน้ำเครื่องหนึ่งไปสูบน้ำมวล 400 กิโลกรัม ขึ้นไปตามท่อสูง 10 เมตร ถ้าน้ำพุ่งออกจากปลายท่อด้วยความเร็ว 10 เมตรต่อวินาที งานที่ได้จากเครื่องสูบมีค่ากี่กิโลจูล (4 คะแนน) วิธีทำ เนื่องจากงานในการเคลื่อนที่สามารถหาได้จากพลังงานที่เปลี่ยนแปลงไป จะได้ $W = mgh + \frac{1}{2}mv^2$	4 คะแนน : ระบุสมการของงานได้ถูกต้องทั้ง 2 เทอม แทนค่าสมการของพลังงานศักย์โน้มถ่วงและสมการของพลังงานจลน์ ได้ถูกต้อง และหาค่างานได้ถูกต้อง 3 คะแนน : ระบุสมการของงานได้ถูกต้องทั้ง 2 เทอม แทนค่าสมการของพลังงานศักย์โน้มถ่วงและสมการของพลังงานจลน์ ได้ถูกต้อง แต่ไม่สามารถหางานได้ถูกต้อง	1	1	1	1	ระดับการประเมินประเมินได้**สอดคล้อง**กับข้อสอบและเฉลยของข้อสอบ แต่ผู้วิจัยพิมพ์ผิดในคำอธิบายของ 1 คะแนน ต้องแก้เป็น ไม่สามารถแทนค่าสมการ

ตาราง 48 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับใหม่และฉบับอ้างอิง ครั้งที่ 2 (ต่อ)

ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
แทนค่า $= (400)(10)(10)$ $+\frac{1}{2}(400)(10)^2$ $= 40{,}000 + 20{,}000$ $= 60{,}000$ **ดังนั้น**งานที่ได้จากเครื่องสูบมีค่า 60,000 จูล หรือ 60 กิโลจูล	2 คะแนน : ระบุสมการของงานได้ถูกต้องทั้ง 2 เทอม แทนค่าพลังงานศักย์โน้มถ่วงหรือพลังงานจลน์ ได้ถูกต้อง แต่ไม่สามารถหางานได้ถูกต้อง 1 คะแนน : ระบุสมการของงานได้ถูกต้องทั้ง 2 เทอม แต่ไม่สามารถแทนค่าสมการของพลังงานศักย์โน้มถ่วงและสมการของพลังงานจลน์ ได้ถูกต้อง และไม่สามารถหางานได้ถูกต้อง					ของพลังงานศักย์โน้มถ่วงหรือสมการของพลังงานจลน์ ได้ถูกต้อง
4) วัตถุมวล 8 กิโลกรัม ที่จุด A มีอัตราเร็ว 0.5 เมตร/วินาที และที่จุด B มีอัตราเร็ว 9 เมตร/วินาที ถ้าระยะทางจากจุด A ไปยังจุด B เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน)  วิธีทำ (แบบที่ 1) เมื่อกำหนดให้เนินที่จุด B เป็นจุดอ้างอิง จะได้ว่า $E_{P1} + E_{k1} - W_f = E_{k2}$ จาก $mgh + \frac{1}{2}mv^2 - fs = \frac{1}{2}mv^2$ แทนค่า $(8)(10)(1) + \frac{1}{2}(8)(0.5)^2 - f(3) = \frac{1}{2}(8)(9)^2$ จะได้ $80 + 1 - f(3) = 3$ $-f(3) = 324 - 8$ $f = -\frac{243}{3} = -81$ นิวตัน	4 คะแนน : ระบุสมการพลังงานได้ถูกต้องทั้งหมด แทนค่าสมการของพลังงานได้ถูกต้องทั้งหมด และหาค่าแรงเสียดทานได้ถูกต้อง 3 คะแนน : ระบุสมการพลังงานได้ถูกต้องทั้งหมด และแทนค่าสมการของพลังงานได้ถูกต้องทั้งหมด แต่ไม่สามารถหาค่าแรงเสียดทานพร้อมได้ถูกต้อง 2 คะแนน : ระบุสมการพลังงานได้ถูกต้องทั้งหมด และแทนค่าสมการของพลังงานได้ถูกต้องอย่างต่ำ 2 เทอม แต่ไม่สามารถหาค่าแรงเสียดทานพร้อมได้ถูกต้อง 1 คะแนน : ระบุสมการพลังงานได้ถูกต้องทั้งหมด แต่ไม่สามารถแทนค่าสมการของพลังงานได้ถูกต้องอย่างต่ำ 2 เทอม และไม่สามารถหาค่าแรงเสียดทานพร้อมได้ถูกต้อง	1	1	1	1	ระดับการประเมินประเมินได้**สอดคล้อง**กับข้อสอบและเฉลยของข้อสอบ

ตาราง 48 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับใหม่และฉบับอ้างอิง ครั้งที่ 2 (ต่อ)

ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
<u>วิธีทำ</u> (แบบที่ 2) เมื่อกำหนดให้เนินที่จุด B เป็นจุดอ้างอิง จะได้ว่า $E_{P1} + E_{k1} - W_f = E_{k2} + E_{P2}$ จาก $mgh + \frac{1}{2}mv^2 - fs = \frac{1}{2}mv^2 + mgh$ แทนค่า $(8)(10)(4) + \frac{1}{2}(8)(0.5)^2 - f(3)$ $= \frac{1}{2}(8)(9)^2 + (8)(10)(3)$ จะได้ว่า $320 + 1 - f(3) = 324 + 240$ $-f(3) = 243$ $f = -\frac{243}{3} = -81$ นิวตัน	4 คะแนน : ระบุสมการพลังงานได้ถูกต้องทั้งหมด แทนค่าสมการของพลังงานได้ถูกต้องทั้งหมด และหาค่าแรงเสียดทานได้ถูกต้อง 3 คะแนน : ระบุสมการพลังงานได้ถูกต้องทั้งหมด และแทนค่าสมการของพลังงานได้ถูกต้องทั้งหมด แต่ไม่สามารถหาค่าแรงเสียดทานพร้อมได้ถูกต้อง 2 คะแนน : ระบุสมการพลังงานได้ถูกต้องทั้งหมด และแทนค่าสมการของพลังงานได้ถูกต้องอย่างต่ำ 2 เทอม แต่ไม่สามารถหาค่าแรงเสียดทานพร้อมได้ถูกต้อง 1 คะแนน : ระบุสมการพลังงานได้ถูกต้องทั้งหมด แต่ไม่สามารถแทนค่าสมการของพลังงานได้ถูกต้องอย่างต่ำ 2 เทอม และไม่สามารถหาค่าแรงเสียดทานพร้อมได้ถูกต้อง	1	1	1	1	ระดับการประเมินประเมินได้**สอดคล้อง**กับข้อสอบและเฉลยของข้อสอบ
5) งานทั้งหมดที่กระทำในช่วงระยะทางการเคลื่อนที่ 0 เมตรถึง 10 เมตร จะมีค่ากี่จูล ถ้ากราฟระหว่างแรงกับการกระจัดเป็นดังรูป  <u>วิธีทำ</u> เนื่องจากงานสามารถหาได้จากแรงคูณกับระยะทางในแนวเดียวกับการเคลื่อนที่ นั่นคือ $W = Fs$	4 คะแนน : ระบุได้ว่างานของการเคลื่อนที่สามารถหาได้จากพื้นที่ใต้กราฟระหว่างแรงกับการกระจัด <u>หรือ</u> ระบุได้ว่างานทั้งหมด=พื้นที่ใต้กราฟระหว่าง F-S ด้านบน - พื้นที่ใต้กราฟระหว่าง F-S ด้านลบ แทนค่าพื้นที่ใต้กราฟส่วนบน<u>และ</u>ส่วนล่างได้ถูกต้อง และหาค่างานได้ถูกต้อง 3 คะแนน : ระบุได้ว่างานของการเคลื่อนที่สามารถหาได้จากพื้นที่ใต้กราฟระหว่างแรงกับการกระจัด <u>หรือ</u> ระบุได้ว่างานทั้งหมด= พื้นที่ใต้กราฟระหว่าง F-S ด้านบน - พื้นที่ใต้กราฟระหว่าง F-S ด้านลบ และแทนค่าพื้นที่ใต้กราฟส่วนบน<u>และ</u>ส่วนล่างได้	1	1	1	1	ระดับการประเมินประเมินได้**สอดคล้อง**กับข้อสอบและเฉลยของข้อสอบ

ตาราง 48 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับใหม่และฉบับอ้างอิง ครั้งที่ 2 (ต่อ)

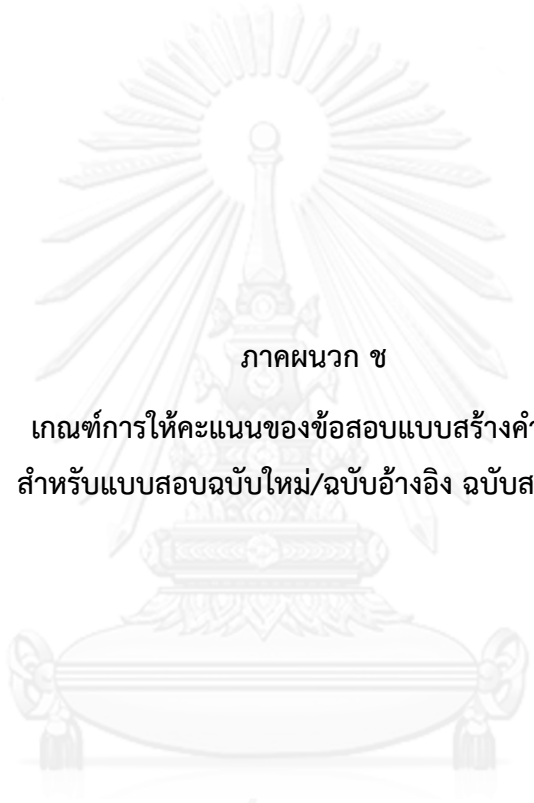
ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
โจทย์ให้กราฟระหว่างแรง (F) กับระยะทางการเคลื่อนที่ (S) ซึ่งพื้นที่ใต้กราฟระหว่างแรงกับระยะทางการเคลื่อนที่ก็คืองาน นั่นเองจะได้ว่า งานทั้งหมด = พื้นที่ใต้กราฟระหว่างแรงกับระยะทางการเคลื่อนที่ = $\frac{1}{2}(8)5 + \frac{1}{2}(-6)(3+1)$ = 20 - 12 = 8 จูล **ดังนั้น** งานทั้งหมดที่กระทำในช่วงระยะทางการเคลื่อนที่ 0 เมตรถึง 10 เมตร จะเท่ากับ 8 จูล	ถูกต้อง แต่ไม่สามารถหาค่างานได้ถูกต้อง 2 คะแนน : ระบุได้ว่างานของการเคลื่อนที่สามารถหาได้จากพื้นที่ใต้กราฟระหว่างแรงกับการกระจัด <u>หรือ</u> ระบุได้ว่างานทั้งหมด= พื้นที่ใต้กราฟระหว่าง F-S ด้านบน - พื้นที่ใต้กราฟระหว่าง F-S ด้านลบ และแทนค่าพื้นที่ใต้กราฟส่วนบน<u>หรือ</u>ส่วนล่างได้ถูกต้อง แต่ไม่สามารถหาค่างานได้ถูกต้อง 1 คะแนน : ระบุได้ว่างานของการเคลื่อนที่สามารถหาได้จากพื้นที่ใต้กราฟระหว่างแรงกับการกระจัด <u>หรือ</u> ระบุได้ว่างานทั้งหมด= พื้นที่ใต้กราฟระหว่าง F-S ด้านบน - พื้นที่ใต้กราฟระหว่าง F-S ด้านลบ แต่ไม่สามารถแทนค่าพื้นที่ใต้กราฟส่วนบนและส่วนล่างได้ถูกต้อง และไม่สามารถหาค่างานได้ถูกต้อง					
6) มอเตอร์ไซต์มวล 300 กิโลกรัม แล่นขึ้นพื้นเอียงทำมุม 30 องศากับแนวระดับ ด้วยอัตราเร็วคงที่ 120 เมตร/วินาที ถ้าแรงเสียดทานระหว่างพื้นเอียงกับรถยนต์เท่ากับ 200 นิวตัน จงหากำลังงานของเครื่องยนต์ในหน่วยกิโลวัตต์ <u>วิธีทำ</u> เนื่องจากความเร็วคงที่ แรงลัพธ์ในการเคลื่อนที่จึงเท่ากับศูนย์ นั่นคือ $\sum \bar{F} = mgsin30° + f$ $= (300)(10)\left(\frac{1}{2}\right) + 200$ $= 1500 + 100 = 1{,}700$ หากำลังได้จาก	4 คะแนน : ระบุสมการของแรงลัพธ์และสมการของกำลังได้ถูกต้อง แทนค่าสมการของแรงลัพธ์และสมการของกำลังได้ถูกต้อง และหาค่าแรงลัพธ์และค่ากำลังได้ถูกต้อง 3 คะแนน : ระบุสมการของแรงลัพธ์และสมการของกำลังได้ถูกต้อง และแทนค่าสมการของแรงลัพธ์และสมการของกำลังได้ถูกต้อง แต่ไม่สามารถหาค่ากำลังได้ถูกต้อง 2 คะแนน : ระบุสมการของแรงลัพธ์และสมการของกำลังได้ถูกต้อง และแทนค่าสมการของแรงลัพธ์ได้ถูกต้อง แต่ไม่สามารถแทนค่าสมการของกำลังได้ถูกต้อง และไม่สามารถหาค่ากำลัง	1	1	1	1	ระดับการประเมินประเมินได้**สอดคล้อง**กับข้อสอบและเฉลยของข้อสอบ

ตาราง 48 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับใหม่และฉบับอ้างอิง ครั้งที่ 2 (ต่อ)

ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
หากำลังได้จาก $P = Fv = (1{,}700)(120) = 204{,}000$ วัตต์ **ดังนั้น** กำลังงานของเครื่องยนต์เท่ากับ 204,000 วัตต์ หรือ 204 กิโลวัตต์	ได้ถูกต้อง 1 คะแนน : ระบุสมการของแรงลัพธ์และสมการของกำลังได้ถูกต้อง แต่ไม่สามารถแทนค่าสมการของแรงลัพธ์และสมการของกำลังได้ถูกต้อง และไม่สามารถหาค่ากำลังได้ถูกต้อง					
7) ดาวต้องการยิงลูกหินมวล 100 กรัม ด้วยหนังสติ๊กให้ขึ้นไปได้สูง 10 เมตรในแนวดิ่ง ถ้าการยืดของหนังสติ๊กเป็นแบบสปริงและยืดออก 50 เซนติเมตร หนังสติ๊กจะต้องมีค่านิจสปริงกี่นิวตัน/เมตร วิธีทำ จากกฎการอนุรักษ์พลังงาน กล่าวว่า พลังงานรวมของวัตถุ ณ ตำแหน่งใดๆย่อมมีความค่าคงตัว นั่นคือ $\sum E_{\text{ตอนแรก}} = \sum E_{\text{ตอนหลัง}}$ จะได้ว่า $mgh = \frac{1}{2}kx^2$ แทนค่า $(0.1)(10)(10) = \frac{1}{2}k(0.5)^2$ $10 = \frac{1}{2}k(0.25)$ เพราะฉะนั้น $k = \frac{10(2)}{0.25} = 80$ นิวตัน/เมตร	4 คะแนน : ระบุสมการของพลังงานได้ถูกต้องทั้ง 2 เทอม (พลังงานศักย์โน้มถ่วงและศักย์ยืดหยุ่น) แทนค่าพลังงานศักย์โน้มถ่วงและพลังงานจลน์ได้ถูกต้อง และคำนวณค่านิจสปริงได้ถูกต้อง 3 คะแนน : ระบุสมการของพลังงานได้ถูกต้องทั้ง 2 เทอม (พลังงานศักย์โน้มถ่วงและศักย์ยืดหยุ่น) แทนค่าพลังงานศักย์โน้มถ่วงและพลังงานจลน์ได้ถูกต้อง แต่ไม่สามารถหาค่าสปริงได้ถูกต้อง 2 คะแนน : ระบุสมการของพลังงานได้ถูกต้องทั้ง 2 เทอม (พลังงานศักย์โน้มถ่วงและศักย์ยืดหยุ่น) และแทนค่าพลังงานศักย์โน้มถ่วงหรือพลังงานจลน์ได้ถูกต้อง แต่ไม่สามารถหาค่าสปริงได้ถูกต้อง 1 คะแนน : ระบุสมการของพลังงานได้ถูกต้องทั้ง 2 เทอม (พลังงานศักย์โน้มถ่วงและศักย์ยืดหยุ่น) แต่ไม่สามารถแทนค่าสมการของแรงลัพธ์และสมการของกำลังได้ถูกต้อง และไม่สามารถหาค่ากำลังได้ถูกต้อง	1	1	1	1	ระดับการประเมินประเมินได้**สอดคล้อง**กับข้อสอบและเฉลยของข้อสอบแต่ผู้วิจัยพิมพ์ผิดในคำอธิบายของ 1 คะแนน ต้องแก้เป็น ไม่สามารถแทนค่าพลังงานศักย์โน้มถ่วงหรือพลังงานจลน์ได้ถูกต้อง และไม่สามารถหาค่าสปริงได้ถูกต้อง

ตาราง 48 สรุปค่า IOC ของผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับเกณฑ์การให้คะแนนข้อสอบแบบสร้างคำตอบของแบบสอบเรื่องงานและพลังงานฉบับใหม่และฉบับอ้างอิง ครั้งที่ 2 (ต่อ)

ข้อ	ระดับคะแนน	ค่า IOC ของผู้เชี่ยวชาญ				ความหมาย
		1	2	3	IOC	
8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัม ขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน) วิธีทำ ประสิทธิภาพของเครื่องกล (E_{ff}) หาได้จาก E_{ff} $= \frac{\text{กำลังที่ได้จริง}}{\text{กำลังที่ได้รับ}} \times 100$ $= \frac{\frac{\text{พลังงาน}}{\text{เวลา}}}{\text{กำลังที่ได้รับ}} \times 100$ $= \frac{\frac{mgh}{t}}{\text{กำลังที่ได้รับ}} \times 100$ $= \frac{\frac{(750+450)(10)(8)}{4}}{30{,}000} \times 100$ $= \frac{24{,}000}{30{,}000} \times 100$ $= 0.8 \times 100 = 80\%$ ดังนั้น ลิฟท์นี้ต้องมีประสิทธิภาพ 80%	4 คะแนน : ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง แทนค่ากำลังที่ได้จริงถูกต้อง แทนค่าในสมการประสิทธิภาพเครื่องกลได้ถูกต้อง และหาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง 3 คะแนน : ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง แทนค่ากำลังที่ได้จริงถูกต้อง และแทนค่าในสมการประสิทธิภาพเครื่องกลได้ถูกต้อง แต่ไม่สามารถหาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง 2 คะแนน : ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง และแทนค่ากำลังที่ได้จริงถูกต้องทั้งหมด แต่ไม่สามารถแทนค่าในสมการประสิทธิภาพเครื่องกลได้ถูกต้อง และไม่สามารถหาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง 1 คะแนน : ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง หรือแทนค่ากำลังที่ได้จริงถูกต้อง แต่ไม่สามารถแทนค่าในสมการประสิทธิภาพเครื่องกลได้ถูกต้อง และไม่สามารถหาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง	1	1	1	1	ระดับการประเมินประเมินได้**สอดคล้อง**กับข้อสอบและเฉลยของข้อสอบ



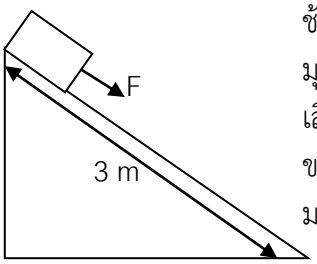
ภาคผนวก ช

เกณฑ์การให้คะแนนของข้อสอบแบบสร้างคำตอบ

สำหรับแบบสอบฉบับใหม่/ฉบับอ้างอิง ฉบับสมบูรณ์

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

เฉลยและเกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงาน (ฉบับใหม่)

1) 

ช้างตัวหนึ่งออกแรงดึงซุงหนัก 140 กิโลกรัม ลงเนินซึ่งเอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ สัมประสิทธิ์ความเสียดทานระหว่างเนินและซุงเป็น $\sqrt{3}/2$ ถ้าแรงที่ใช้ดึงซุงขนานกับการเคลื่อนที่ดังรูป ช้างต้องใช้งานเท่าใดในการดึงซุงลงมาได้ 3 เมตร (4 คะแนน)

<u>วิธีทำ</u> เนื่องจากความเร็วคงที่ แรงลัพธ์ในการเคลื่อนที่จึงเท่ากับศูนย์

นั่นคือ $\sum \bar{F} = 0$

แรงซ้าย = แรงขวา

$$F + mgsin\theta = f$$

$$F + mgsin\theta = \mu mgcos\theta$$

จะได้ว่า $$F = \mu mgcos30° - mgsin30°$$

$$= \frac{\sqrt{3}}{2}(140)(10)\left(\frac{\sqrt{3}}{2}\right) - (140)(10)(\frac{1}{2})$$

เพราะฉะนั้น $$F = 1050 - 700 = 350$$

จาก $$W = Fscos\theta = 350(3)cos0° = 1{,}050$$

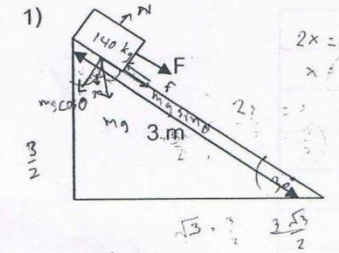
ดังนั้น ช้างต้องใช้งาน 1,050 จูลในการดึงซุงลงมาได้ 3 เมตร

เกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงาน ข้อ 1

ระดับคะแนน	คำอธิบาย
4	☐ระบุสมการในการหางานและสมการของแรงลัพธ์ได้ถูกต้อง ☐แทนค่าสมการของแรงลัพธ์และสมการของงานได้ถูกต้อง ☐หาค่างานพร้อมทั้งระบุหน่วยได้ถูกต้อง
3	☐ระบุสมการในการหางานและสมการของแรงลัพธ์ได้ถูกต้อง ☐แทนค่าสมการของแรงลัพธ์และสมการของงานได้ถูกต้อง ☐แต่ไม่สามารถหาค่างานได้ถูกต้อง
2	☐ระบุสมการในการหางานและสมการของแรงลัพธ์ได้ถูกต้อง ☐แทนค่าสมการของแรงลัพธ์ได้ถูกต้อง ☐แต่ไม่สามารถแทนค่าสมการของงานได้ถูกต้อง และไม่สามารถหาค่างานได้ถูกต้อง
1	☐ระบุสมการของงานของแรงลัพธ์ได้ถูกต้อง ☐แต่ไม่สามารถแทนค่าสมการของแรงลัพธ์หรือสมการของงานได้ถูกต้อง และไม่สามารถหาค่าแรงลัพธ์งานได้ถูกต้อง

ตัวอย่างการแสดงวิธีทำสำหรับระดับคะแนนต่างๆ

ระดับคะแนน 1

1) 

ช้างตัวหนึ่งออกแรงดึงซุงหนัก 140 กิโลกรัม ลงเนินซึ่งเอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ สัมประสิทธิ์ความเสียดทานระหว่างเนินและซุงเป็น $\sqrt{3}/2$ ถ้าแรงที่ใช้ดึงซุงขนานกับการเคลื่อนที่ดังรูป ช้างต้องใช้งานเท่าใดในการดึงซุงลงมาได้ 3 เมตร (4 คะแนน)

$\Sigma F = 0$

$F + mg\sin\theta - f = 0$

$F = f - mg\sin\theta$

$= \mu mg\cos\theta - mg\sin\theta$

$= \frac{\sqrt{3}}{2} \cdot 1400(\cos\theta - \sin\theta)$

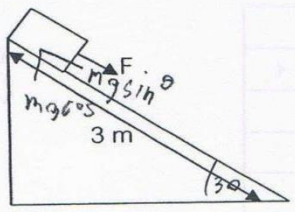
$= \frac{\sqrt{3}}{2} \cdot 1400\left(\frac{\sqrt{3}}{2} - \frac{1}{2}\right)$

$= \frac{\sqrt{3}}{2} \cdot 1400(\sqrt{3} - 1)$

$W = F \cdot s$

$= 3[700\sqrt{3}(\sqrt{3} - 1)]$ J

ระดับคะแนน 2

1) 

ช้างตัวหนึ่งออกแรงดึงซุงหนัก 140 กิโลกรัม ลงเนินซึ่งเอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ สัมประสิทธิ์ความเสียดทานระหว่างเนินและซุงเป็น $\sqrt{3}/2$ ถ้าแรงที่ใช้ดึงซุงขนานกับการเคลื่อนที่ดังรูป ช้างต้องใช้งานเท่าใดในการดึงซุงลงมาได้ 3 เมตร (4 คะแนน)

∴ ช้างต้องใช้งาน $W = Fs$ $s = 3$

$\therefore F + mg\sin\theta - \mu mg\cos\theta = 0$

$F = \mu mg\cos\theta - mg\sin\theta$

$= \frac{\sqrt{3}}{2}(1400)\frac{\sqrt{3}}{2} - 700\frac{1}{2}$

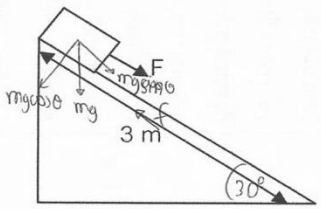
$= 2100 - 700$

$= 1400$ N

∴ ช้างต้องใช้งาน $W = 1400(3) = 4200$ J

ระดับคะแนน 3

1)



ช้างตัวหนึ่งออกแรงดึงซุงหนัก 140 กิโลกรัม ลงเนินซึ่งเอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ สัมประสิทธิ์ความเสียดทานระหว่างเนินและซุงเป็น $\sqrt{3}/2$ ถ้าแรงที่ใช้ดึงซุงขนานกับการเคลื่อนที่ดังรูป ช้างต้องใช้งานเท่าใดในการดึงซุงลงมาได้ 3 เมตร (4 คะแนน)

$\Sigma F = 0$

$F + mg\sin\theta - f = 0$

$F = f - mg\sin\theta$

$= \frac{\sqrt{3}}{2}(140)(10)\cos 30° - (140)(10)\left(\frac{1}{2}\right)$

$= (\sqrt{3})(70)(10)\left(\frac{\sqrt{3}}{2}\right) - 700$

$F = 1050 - 700 = 350$

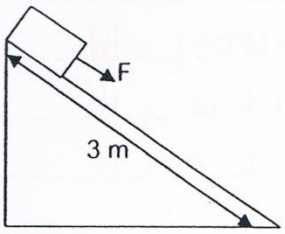
$W = FS\cos\theta$

$= (350)(3)\cos 0°$

$W = 0$ J

ระดับคะแนน 4

1)



ช้างตัวหนึ่งออกแรงดึงซุงหนัก 140 กิโลกรัม ลงเนินซึ่งเอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ สัมประสิทธิ์ความเสียดทานระหว่างเนินและซุงเป็น $\sqrt{3}/2$ ถ้าแรงที่ใช้ดึงซุงขนานกับการเคลื่อนที่ดังรูป ช้างต้องใช้งานเท่าใดในการดึงซุงลงมาได้ 3 เมตร (4 คะแนน)

$\Sigma F = 0$

$F + mg\sin 30° - \mu mg\cos 30 = 0$

$F + mg\sin 30 = \mu mg\cos 30$

$F + (1400)\left(\frac{1}{2}\right) = \frac{\sqrt{3}}{2}(1400)\left(\frac{\sqrt{3}}{2}\right)$

$F + 700 = 1050$

$F = 350$ N

$W = F \cdot S = 350 \cdot 3 = 1050$ J

2) สปริงถูกออกแรงดึง 20 นิวตัน จะทำให้ยืดออก 10 เซนตเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 20 เซนติเมตรในหน่วยจูล (4 คะแนน)

<u>วิธีทำ</u> หาค่านิจสปริงจาก $F = kx$

แทนค่า $20 = k(0.1)$

จะได้ว่า $k = \frac{20}{0.1} = 200$ นิวตัน/เมตร

หาพลังงานศักย์ยืดหยุ่น (E_P) จาก ;

$$E_P = \frac{1}{2}kx^2$$

แทนค่า $= \frac{1}{2}(200)(0.2)^2 = \frac{1}{2}(200)(0.04) = 4$ จูล

ดังนั้น พลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 20 เซนติเมตรจะเท่ากับ 4 จูล

เกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงาน ข้อ 2

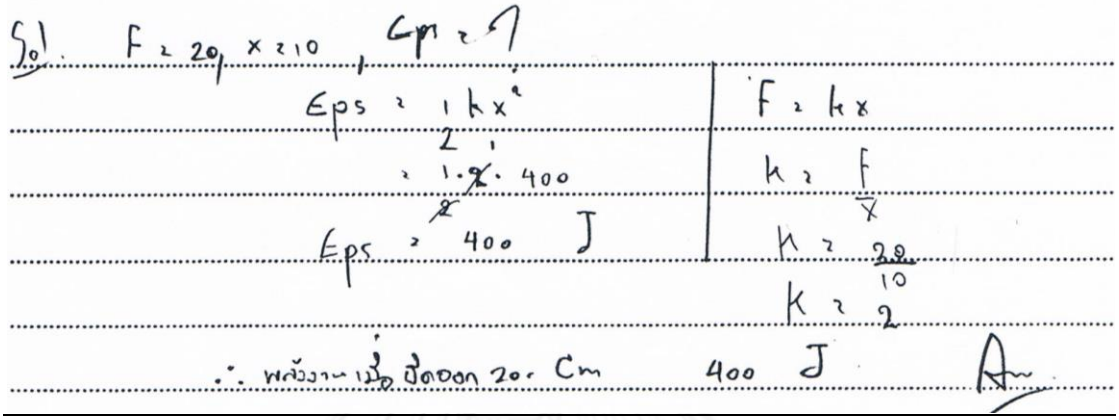
ระดับคะแนน	คำอธิบาย
4	☐ระบุสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง ☐แทนค่าสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง ☐หาค่าพลังงานศักย์ยืดหยุ่นได้ถูกต้อง
3	☐ระบุสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง ☐แทนค่าสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง ☐ แต่ไม่สามารถหาค่าพลังงานศักย์ยืดหยุ่นได้ถูกต้อง
2	☐ระบุสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง ☐แทนค่าสมการตามกฎของฮุคได้ถูกต้อง ☐แต่ไม่สามารถแทนค่าสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง และไม่สามารถหาค่าพลังงานศักย์ยืดหยุ่นได้ถูกต้อง
1	☐ระบุสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง ☐แต่ไม่สามารถแทนค่าสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง และไม่สามารถหาค่างานได้ถูกต้อง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตัวอย่างการแสดงวิธีทำสำหรับระดับคะแนนต่างๆ

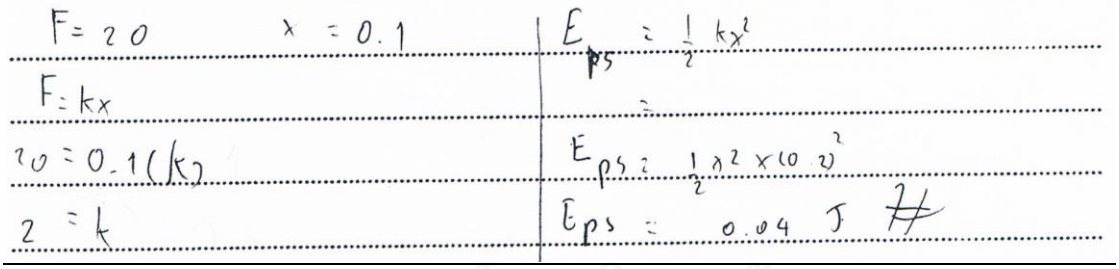
ระดับคะแนน 1

2) สปริงถูกออกแรงดึง 20 นิวตัน จะทำให้ยืดออก 10 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริง เมื่อสปริงยืดออกมา 20 เซนติเมตรในหน่วยจูล (4 คะแนน)

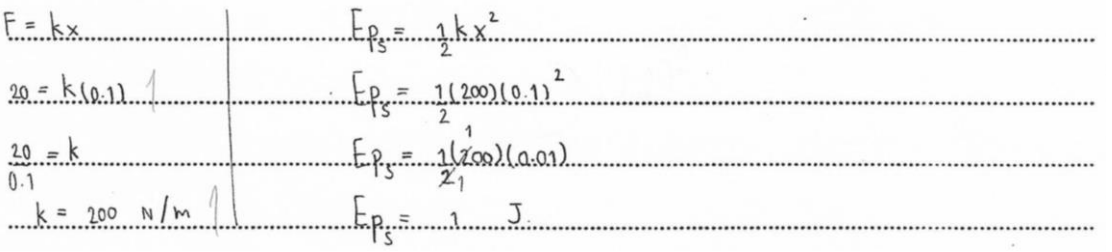


ระดับคะแนน 2

2) สปริงถูกออกแรงดึง 20 นิวตัน จะทำให้ยืดออก 10 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริง เมื่อสปริงยืดออกมา 20 เซนติเมตรในหน่วยจูล (4 คะแนน)



2) สปริงถูกออกแรงดึง 20 นิวตัน จะทำให้ยืดออก 10 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริง เมื่อสปริงยืดออกมา 20 เซนติเมตรในหน่วยจูล (4 คะแนน)



ระดับคะแนน 3

2) สปริงถูกออกแรงดึง 20 นิวตัน จะทำให้ยืดออก 10 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริง เมื่อสปริงยืดออกมา 20 เซนติเมตรในหน่วยจูล (4 คะแนน)

0.2 m

$k = \frac{F}{S}$

$k = \frac{20}{0.1}$

$k = 200\ N/m.$

$E_{ps} = \frac{1}{2}ks^2$

$= \frac{1}{2}200(0.2)^2$

$= 100(0.4)$

$= 40\ J$

ระดับคะแนน 4

2) สปริงถูกออกแรงดึง 20 นิวตัน จะทำให้ยืดออก 10 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริง เมื่อสปริงยืดออกมา 20 เซนติเมตรในหน่วยจูล (4 คะแนน)

$F = ks$

$20 = k\left(\frac{10}{100}\right)$ 1

$k = 200\ N/m$ 1

$E_{ps} = \frac{1}{2}ks^2$

$= \frac{1}{2}(200)\left(\frac{20}{100}\right)^2$

$= 100\left(\frac{400}{10000}\right)$

$= 4\ J$ 1

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

3) เคนนำเครื่องสูบน้ำเครื่องหนึ่งไปสูบน้ำมวล 400 กิโลกรัม ขึ้นไปตามท่อสูง 10 เมตร ถ้าน้ำพุ่งออกจากปลายท่อด้วยความเร็ว 10 เมตรต่อวินาที งานที่ได้จากเครื่องสูบมีค่ากี่กิโลจูล (4 คะแนน)

<u>วิธีทำ</u> เนื่องจากงานในการเคลื่อนที่สามารถหาได้จากพลังงานที่เปลี่ยนแปลงไป

จะได้ว่า $W = mgh + \frac{1}{2}mv^2$

แทนค่า $= (400)(10)(10) + \frac{1}{2}(400)(10)^2 = 40{,}000 + 20{,}000$

$= 60{,}000$ จูล

ดังนั้น งานที่ได้จากเครื่องสูบมีค่า 60,000 จูล หรือ 60 กิโลจูล

เกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงาน ข้อ 3

ระดับคะแนน	คำอธิบาย
4	☐ระบุสมการของงานได้ถูกต้องทั้ง 2 เทอม ☐แทนค่าสมการของพลังงานศักย์โน้มถ่วง<u>และ</u>สมการของพลังงานจลน์ ได้ถูกต้อง ☐หาค่างานได้ถูกต้อง
3	☐ระบุสมการของงานได้ถูกต้องทั้ง 2 เทอม ☐แทนค่าสมการของพลังงานศักย์โน้มถ่วง<u>และ</u>สมการของพลังงานจลน์ ได้ถูกต้อง ☐แต่ไม่สามารถหางานได้ถูกต้อง
2	☐ระบุสมการของงานได้ถูกต้องทั้ง 2 เทอม ☐แทนค่าพลังงานศักย์โน้มถ่วง<u>หรือ</u>พลังงานจลน์ ได้ถูกต้อง ☐แต่ไม่สามารถหางานได้ถูกต้อง
1	☐ระบุสมการของงานได้ถูกต้องทั้ง 2 เทอม ☐แต่ไม่สามารถแทนค่าสมการของพลังงานศักย์โน้มถ่วงหรือสมการของพลังงานจลน์ ได้ถูกต้อง และไม่สามารถหางานได้ถูกต้อง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตัวอย่างการแสดงวิธีทำสำหรับระดับคะแนนต่างๆ

ระดับคะแนน 1

3) เคนนำเครื่องสูบน้ำเครื่องหนึ่งไปสูบน้ำมวล 400 กิโลกรัม ขึ้นไปตามท่อสูง 10 เมตร ถ้าน้ำพุ่งออกจากปลายท่อด้วยความเร็ว 10 เมตรต่อวินาที งานที่ได้จากเครื่องสูบมีค่ากี่กิโลจูล (4 คะแนน)

$E_p + E_k = W$

$mgh + \frac{1}{2}mv^2 = W$

$10(10) + \frac{1}{2}(10)^2 = W$

$100 + 50 = W$

$W = 150 \times 10^3 = 150000$ กิโลจูล (KJ)

ระดับคะแนน 2

3) เคนนำเครื่องสูบน้ำเครื่องหนึ่งไปสูบน้ำมวล 400 กิโลกรัม ขึ้นไปตามท่อสูง 10 เมตร ถ้าน้ำพุ่งออกจากปลายท่อด้วยความเร็ว 10 เมตรต่อวินาที งานที่ได้จากเครื่องสูบมีค่ากี่กิโลจูล (4 คะแนน)

$E_p = E_k$

$\frac{1}{2}mv^2 = mgh$

$\frac{1}{2}(400)v^2 = (400)(10)(10)$

$v^2 = 200$

$v = 10\sqrt{2}$

$v = 14.14$ m/s

$E_p + E_k = W$

$\frac{1}{2}(400)(14.14) + (400)(10)(10) = W$

$2828 + 40000 = W$

$W = 42828$ J

$W = 42.828$ KJ

ระดับคะแนน 3

3) เคนนำเครื่องสูบน้ำเครื่องหนึ่งไปสูบน้ำมวล 400 กิโลกรัม ขึ้นไปตามท่อสูง 10 เมตร ถ้าน้ำพุ่งออกจากปลายท่อด้วยความเร็ว 10 เมตรต่อวินาที งานที่ได้จากเครื่องสูบมีค่ากี่กิโลจูล (4 คะแนน)

$W_{1-2} = E_k = E_p$

$= \frac{1}{2}mv^2 + mgh$

$= \frac{1}{2}(400)(10)^2 + (400)(10)(10)$

$= \frac{40000}{2} + 40000$

$= \frac{80000}{2}$

$\therefore W_{1-2} = 40000$ กิโลจูล

ระดับคะแนน 4

3) เคนนำเครื่องสูบน้ำเครื่องหนึ่งไปสูบน้ำมวล 400 กิโลกรัม ขึ้นไปตามท่อสูง 10 เมตร ถ้าน้ำพุ่งออกจากปลายท่อด้วยความเร็ว 10 เมตรต่อวินาที งานที่ได้จากเครื่องสูบมีค่ากี่กิโลจูล (4 คะแนน)

h = 10 m m = 400 kg v = 10 m/s W = ?

W = Ep + Ek

$W = mgh + \frac{1}{2}mv^2$

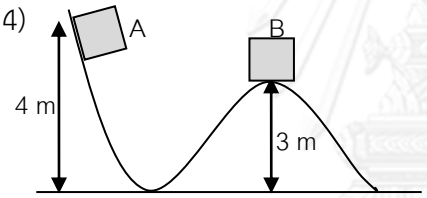
$W = 400(10)(10) + \frac{1}{2}(400)(10)^2$

W = 40000 + 20000

W = 60000 J

∴ W = 60 kJ.

---

4) 

วัตถุมวล 8 กิโลกรัม ที่จุด A มีอัตราเร็ว 0.5 เมตร/วินาที และที่จุด B มีอัตราเร็ว 9 เมตร/วินาที ถ้าระยะทางจากจุด A ไปยังจุด B เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (4 คะแนน)

วิธีทำ (แบบที่ 1)

เมื่อกำหนดให้เนินที่จุด B เป็นจุดอ้างอิง จะได้ว่า

$$E_{P1} + E_{k1} - W_f = E_{k2}$$

จาก $$mgh + \frac{1}{2}mv^2 - fs = \frac{1}{2}mv^2$$

แทนค่า $$(8)(10)(1) + \frac{1}{2}(8)(0.5)^2 - f(3) = \frac{1}{2}(8)(9)^2$$

จะได้ว่า $$80 + 1 - f(3) = 324$$

$$-f(3) = 324 - 81$$

$$f = -\frac{243}{3} = -81 \text{ นิวตัน}$$

ดังนั้นแรงเสียดทานระหว่างพื้นและกล่องมีค่า - 81 นิวตัน

<u>วิธีทำ</u> (แบบที่ 2)

เมื่อกำหนดพื้นราบเป็นจุดอ้างอิง จะได้ว่า

$$E_{P1} + E_{k1} - W_f = E_{k2} + E_{P2}$$

จาก $$mgh + \frac{1}{2}mv^2 - fs = \frac{1}{2}mv^2 + mgh$$

แทนค่า $(8)(10)(4) + \frac{1}{2}(8)(0.5)^2 - f(3) = \frac{1}{2}(8)(9)^2 + (8)(10)(3)$

จะได้ว่า $$320 + 1 - f(3) = 324 + 240$$

$$-f(3) = 243$$

$$f = -\frac{243}{3} = -81 \text{ นิวตัน}$$

ดังนั้นแรงเสียดทานระหว่างพื้นและกล่องมีค่า - 81 นิวตัน

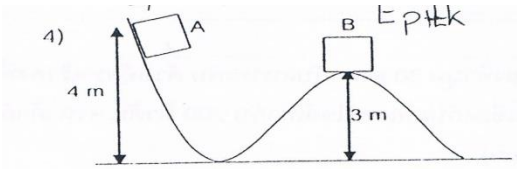
เกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงาน ข้อ 4

ระดับคะแนน	คำอธิบาย
4	☐ระบุสมการพลังงานได้ถูกต้องทั้งหมด ☐แทนค่าสมการของพลังงานได้ถูกต้องทั้งหมด ☐หาค่าแรงเสียดทานได้ถูกต้อง
3	☐ระบุสมการพลังงานได้ถูกต้องทั้งหมด ☐แทนค่าสมการของพลังงานได้ถูกต้องทั้งหมด ☐แต่ไม่สามารถหาค่าแรงเสียดทานพร้อมได้ถูกต้อง
2	☐ระบุสมการพลังงานได้ถูกต้องทั้งหมด ☐แทนค่าสมการของพลังงานได้ถูกต้องอย่างต่ำ 2 เทอม ☐แต่ไม่สามารถหาค่าแรงเสียดทานพร้อมได้ถูกต้อง
1	☐ระบุสมการพลังงานได้ถูกต้องทั้งหมด ☐แต่ไม่สามารถแทนค่าสมการของพลังงานได้ถูกต้องอย่างต่ำ 2 เทอม และไม่สามารถหาค่าแรงเสียดทานพร้อมได้ถูกต้อง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตัวอย่างการแสดงวิธีทำสำหรับระดับคะแนนต่างๆ

ระดับคะแนน 1



วัตถุมวล 8 กิโลกรัม ที่จุด A มีอัตราเร็ว 0.5 เมตร/วินาที และที่จุด B มีอัตราเร็ว 9 เมตร/วินาที ถ้าระยะทางจากจุด A ไปยังจุด B เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (4 คะแนน)

$A = B$

$E_p + E_k = E_p + E_k$

$mgh + \frac{1}{2}mv^2 = mgh + \frac{1}{2}mv^2 + W$

$10 \times 4 + \frac{1}{2}(0.5)^2 = 10 \times 3 + \frac{1}{2}(9)^2 + f(3)$

$40 + \frac{0.25}{2} = 30 + \frac{81}{2} + 3f$

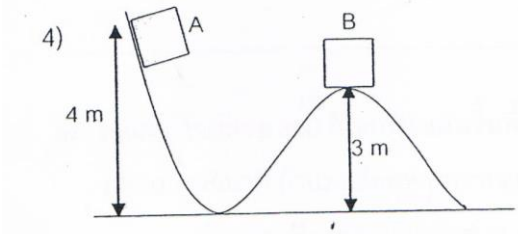
$10 + \frac{0.25}{2} - \frac{81}{2} = 3f$

$\frac{10 + 0.125 - 40.5}{3} = f$

$f = \frac{30.375}{3}$

$f = 10.125$ N

ระดับคะแนน 2



วัตถุมวล 8 กิโลกรัม ที่จุด A มีอัตราเร็ว 0.5 เมตร/วินาที และที่จุด B มีอัตราเร็ว 9 เมตร/วินาที ถ้าระยะทางจากจุด A ไปยังจุด B เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (4 คะแนน)

$E_{K_1} + E_{P_1} = E_{K_2} + E_{P_2} + W_f$

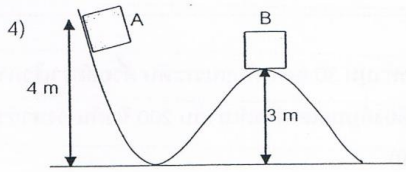
$\frac{1}{2}(8)(0.25) + 8(10)(4) = \frac{1}{2}(8)(81) + 8(10)(3) + W_f$

$1 + 320 = 324 + 240 + W_f$

$321 = 564 + W_f$

$W_f = -243$ J.

ระดับคะแนน 3

4) 

วัตถุมวล 8 กิโลกรัม ที่จุด A มีอัตราเร็ว 0.5 เมตร/วินาที และที่จุด B มีอัตราเร็ว 9 เมตร/วินาที ถ้าระยะทางจากจุด A ไปยังจุด B เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (4 คะแนน)

$W+E_p+E_k = E_p+E_k$

$W+mgh+\frac{1}{2}mv^2 = mgh+\frac{1}{2}mv^2$

$W+8(10)4+\frac{1}{2}(8)0.5^2 = 8(10)3+\frac{1}{2}(8)9^2$

$W+320+1 = 240+324$

$W = 564-321$

$W = 243$ J

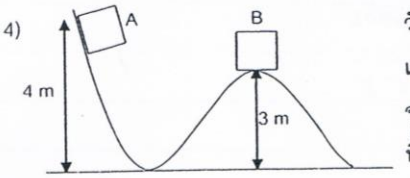
$W = fs$

$243 = -f(3)$

$-f = \frac{243}{3}$

$-f = 81$ N.

ระดับคะแนน 4

4) 

วัตถุมวล 8 กิโลกรัม ที่จุด A มีอัตราเร็ว 0.5 เมตร/วินาที และที่จุด B มีอัตราเร็ว 9 เมตร/วินาที ถ้าระยะทางจากจุด A ไปยังจุด B เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (4 คะแนน)

$E_k+E_p = E_k+E_p+W$

$\left[\frac{1}{2}\times 8\times(0.5)^2\right]+[8\times 10\times 4] = \left[\frac{1}{2}\times 8\times(9)^2\right]+[8\times 10\times 3]+W$

$[4\times 0.25]+[320] = [324+240]+W$

$321 = 564+W$

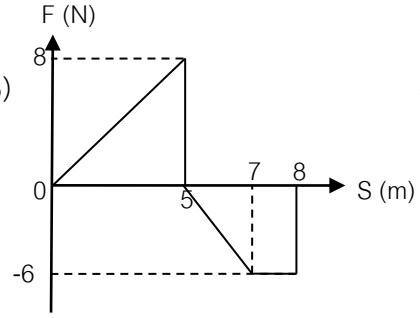
$W = -243$

$W = Fs$

$-243 = F\times 3$

$F = \frac{-243}{3}$

$F = -81$ แรงเสียดทาน $= -81$ N

5) 

งานทั้งหมดที่กระทำในช่วงระยะทางการเคลื่อนที่ 0 เมตรถึง 10 เมตร จะมีค่ากี่จูล ถ้ากราฟระหว่างแรงกับการกระจัดเป็นดังรูป (4 คะแนน)

<u>วิธีทำ</u> เนื่องจากกงานสามารถหาได้จากแรงคูณกับระยะทางในแนวเดียวกับการเคลื่อนที่ นั่นคือ $W = Fs$

โจทย์ให้กราฟระหว่างแรง (F) กับระยะทางการเคลื่อนที่ (S) ซึ่งพื้นที่ใต้กราฟระหว่างแรงกับระยะทางการเคลื่อนที่ ก็คืองานนั่นเอง

จะได้ว่า งานทั้งหมด = พื้นที่ใต้กราฟระหว่างแรงกับระยะทางการเคลื่อนที่

$= \frac{1}{2}(8)5 + \frac{1}{2}(-6)(3+1) = 20 - 12 = 8$ จูล

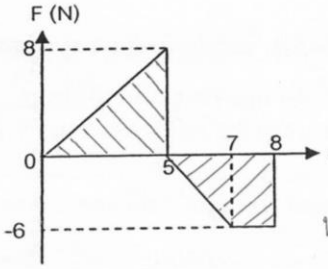
ดังนั้น งานทั้งหมดที่กระทำในช่วงระยะทางการดังกล่าว จะเท่ากับ 8 จูล

เกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงาน ข้อ 5

ระดับคะแนน	คำอธิบาย
4	☐ระบุได้ว่างานของการเคลื่อนที่สามารถหาได้จากพื้นที่ใต้กราฟระหว่างแรงกับการกระจัด <u>หรือ</u>ระบุได้ว่างานทั้งหมด=พื้นที่ใต้กราฟระหว่าง F-S ด้านบน - พื้นที่ใต้กราฟระหว่าง F-S ด้านลบ ☐แทนค่าพื้นที่ใต้กราฟส่วนบน<u>และ</u>ส่วนล่างได้ถูกต้อง ☐หาค่างานได้ถูกต้อง
3	☐ระบุได้ว่างานของการเคลื่อนที่สามารถหาได้จากพื้นที่ใต้กราฟระหว่างแรงกับการกระจัด <u>หรือ</u>ระบุได้ว่างานทั้งหมด=พื้นที่ใต้กราฟระหว่าง F-S ด้านบน - พื้นที่ใต้กราฟระหว่าง F-S ด้านลบ ☐แทนค่าพื้นที่ใต้กราฟส่วนบน<u>และ</u>ส่วนล่างได้ถูกต้อง ☐แต่ไม่สามารถหาค่างานได้ถูกต้อง
2	☐ระบุได้ว่างานของการเคลื่อนที่สามารถหาได้จากพื้นที่ใต้กราฟระหว่างแรงกับการกระจัด <u>หรือ</u>ระบุได้ว่างานทั้งหมด=พื้นที่ใต้กราฟระหว่าง F-S ด้านบน - พื้นที่ใต้กราฟระหว่าง F-S ด้านลบ ☐แทนค่าพื้นที่ใต้กราฟส่วนบน<u>หรือ</u>ส่วนล่างได้ถูกต้อง ☐แต่ไม่สามารถหาค่างานได้ถูกต้อง
1	☐ระบุได้ว่างานของการเคลื่อนที่สามารถหาได้จากพื้นที่ใต้กราฟระหว่างแรงกับการกระจัด <u>หรือ</u>ระบุได้ว่างานทั้งหมด=พื้นที่ใต้กราฟระหว่าง F-S ด้านบน - พื้นที่ใต้กราฟระหว่าง F-S ด้านลบ ☐แต่ไม่สามารถแทนค่าพื้นที่ใต้กราฟส่วนบนและส่วนล่างได้ถูกต้อง และไม่สามารถหาค่างานได้ถูกต้อง

ตัวอย่างการแสดงวิธีทำสำหรับระดับคะแนนต่างๆ

ระดับคะแนน 1



5) งานทั้งหมดที่กระทำในช่วงระยะทางการเคลื่อนที่ 0 เมตรถึง 8 เมตร จะมีค่ากี่จูล ถ้ากราฟระหว่างแรงกับการกระจัดเป็นดังรูป (4 คะแนน)

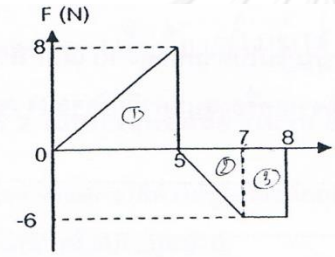
พ.ท. ใต้กราฟ.

$= \frac{1}{2} \times (-6) \times 8 \times 5$

$= -120$ J. Ans.

∴ มีค่า 120 J ถ้ากราฟระหว่างแรงกับการกระจัดเป็นดังรูป

ระดับคะแนน 2



5) งานทั้งหมดที่กระทำในช่วงระยะทางการเคลื่อนที่ 0 เมตรถึง 8 เมตร จะมีค่ากี่จูล ถ้ากราฟระหว่างแรงกับการกระจัดเป็นดังรูป (4 คะแนน)

① $= \frac{1}{2} \times$ ฐาน $\times$ สูง

$= \frac{1}{2} \times 8 \times 5$

$= 20$

② $\frac{1}{2} \times$ ฐาน $\times$ สูง

$\frac{1}{2} \times -6 \times 5$

$= -9$

③ กว้าง $\times$ ยาว

$= 1 \times -6$

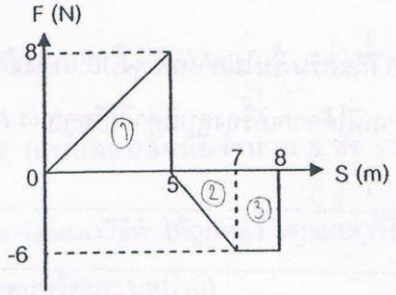
$= -6$

งานทั้งหมด $= 20 + (-9) + (-6)$

$= 20 - 15$

$= 5$ J

ระดับคะแนน 3

5) 

งานทั้งหมดที่กระทำในช่วงระยะทางการเคลื่อนที่ 0 เมตรถึง 8 เมตร จะมีค่ากี่จูล ถ้ากราฟระหว่างแรงกับการกระจัดเป็นดังรูป (4 คะแนน)

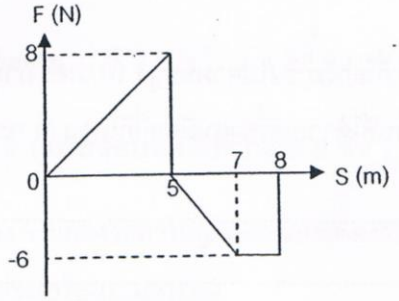
① = $\frac{1}{2} \times 5 \times 8$
$= \frac{1}{2} \times 8 \times 5$
$= 20$

② = $\frac{1}{2} \times 2 \times -6$
$= 1 \times -6 \times 1$
$= -6$

③ = 2×-3
$= -6 \times 1$
$= -6$

W = ① + ② + ③
$= 20 + (-6) + (-6)$
$= 20$ J

ระดับคะแนน 4

5) 

งานทั้งหมดที่กระทำในช่วงระยะทางการเคลื่อนที่ 0 เมตรถึง 8 เมตร จะมีค่ากี่จูล ถ้ากราฟระหว่างแรงกับการกระจัดเป็นดังรูป (4 คะแนน)

$W = (\frac{1}{2} \times 5 \times 8) - (\frac{1}{2}(1+3)(6))$
$= 20 - 12$
$= 8$ J

6) มอเตอร์ไซต์มวล 300 กิโลกรัม แล่นขึ้นพื้นเอียงทำมุม 30 องศากับแนวระดับ ด้วยอัตราเร็วคงที่ 120 เมตร/วินาที ถ้าแรงเสียดทานระหว่างพื้นเอียงกับรถยนต์เท่ากับ 200 นิวตัน จงหากำลังงานของเครื่องยนต์ในหน่วยกิโลวัตต์ (4 คะแนน)

<u>วิธีทำ</u> เนื่องจากความเร็วคงที่ แรงลัพธ์ในการเคลื่อนที่จึงเท่ากับศูนย์

นั่นคือ $$\sum \bar{F} = mgsin30° + f$$
$$= (300)(10)\left(\frac{1}{2}\right) + 200$$
$$= 1500 + 100 = 1{,}700 \text{ นิวตัน}$$

หากำลังได้จาก $$P = Fv = (1{,}700)(120) = 204{,}000 \text{ วัตต์}$$

ดังนั้น กำลังงานของเครื่องยนต์เท่ากับ 204,000 วัตต์ หรือ 204 กิโลวัตต์

เกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงาน ข้อ 6

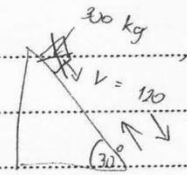
ระดับคะแนน	คำอธิบาย
4	☐ระบุสมการของแรงลัพธ์และสมการของกำลังได้ถูกต้อง ☐แทนค่าสมการของแรงลัพธ์และสมการของกำลังได้ถูกต้อง ☐หาค่าแรงลัพธ์และค่ากำลังได้ถูกต้อง
3	☐ระบุสมการของแรงลัพธ์และสมการของกำลังได้ถูกต้อง ☐แทนค่าสมการของแรงลัพธ์และสมการของกำลังได้ถูกต้อง ☐ แต่ไม่สามารถหาค่ากำลังได้ถูกต้อง
2	☐ระบุสมการของแรงลัพธ์และสมการของกำลังได้ถูกต้อง ☐แทนค่าสมการของแรงลัพธ์ได้ถูกต้อง ☐แต่ไม่สามารถแทนค่าสมการของกำลังได้ถูกต้อง และไม่สามารถหาค่ากำลังได้ถูกต้อง
1	☐ระบุสมการของแรงลัพธ์และสมการของกำลังได้ถูกต้อง ☐แต่ไม่สามารถแทนค่าสมการของแรงลัพธ์และสมการของกำลังได้ถูกต้องและไม่สามารถหาค่ากำลังได้ถูกต้อง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตัวอย่างการแสดงวิธีทำสำหรับระดับคะแนนต่างๆ

ระดับคะแนน 1

6) มอเตอร์ไซต์มวล 300 กิโลกรัม แล่นขึ้นพื้นเอียงทำมุม 30 องศากับแนวระดับ ด้วยอัตราเร็วคงที่ 120 เมตร/วินาที ถ้าแรงเสียดทานระหว่างพื้นเอียงกับมอเตอร์ไซต์เท่ากับ 200 นิวตัน จงหากำลังของเครื่องยนต์ในหน่วยกิโลวัตต์ (4 คะแนน) $P = \frac{W}{t}$



$a = 0$

จากกฎ $\Sigma F = ma$

$F - \mu mg - mg\sin\gamma = ma$

$F = 200 \times \cos 30 - 300 \times 10 \times \sin 30^\circ$

$F = 200 \times \frac{\sqrt{3}}{2} - 3000 \times \frac{1}{2}$

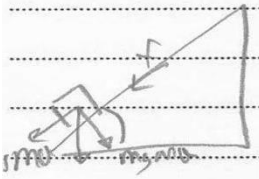
$F = 200\sqrt{3} - 1500 = -1300$

$\therefore P = Fv = -1300 \times 120 = -156000$ W

แต่ KW $= -156 \times 10^3 = -156$ KW

ระดับคะแนน 2

6) มอเตอร์ไซต์มวล 300 กิโลกรัม แล่นขึ้นพื้นเอียงทำมุม 30 องศากับแนวระดับ ด้วยอัตราเร็วคงที่ 120 เมตร/วินาที ถ้าแรงเสียดทานระหว่างพื้นเอียงกับมอเตอร์ไซต์เท่ากับ 200 นิวตัน จงหากำลังของเครื่องยนต์ในหน่วยกิโลวัตต์ (4 คะแนน)



$\Sigma F = mg\sin\theta + f$

$= (300)(10)\sin 30^\circ + 200$

$= (300)(10)\left(\frac{1}{2}\right) + 200$

$= 1500 + 200$

$P = Fv$

$= 1,520 \times 12$

$= 18240$ watt

$= 18.24$ kilowatt

152 × 12 = 304 + 152 = 1824

ระดับคะแนน 3

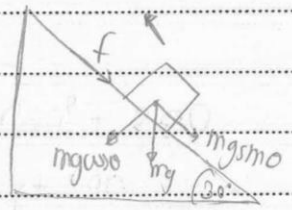
6) มอเตอร์ไซต์มวล 300 กิโลกรัม แล่นขึ้นพื้นเอียงทำมุม 30 องศากับแนวระดับ ด้วยอัตราเร็วคงที่ 120 เมตร/วินาที ถ้าแรงเสียดทานระหว่างพื้นเอียงกับมอเตอร์ไซต์เท่ากับ 200 นิวตัน จงหากำลังของเครื่องยนต์ในหน่วยกิโลวัตต์ (4 คะแนน)

$\Sigma F = mg\sin 30^\circ + f$

$= (300)(10)\left(\frac{1}{2}\right) + 200$

$= 1,500 + 200$

$\Sigma F = 1,700$



$P = FV$

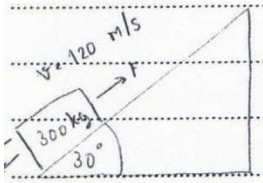
$= (1,700)(120)$

$= 20400$ วัตต์

$P = 20.4$ กิโลวัตต์

ระดับคะแนน 4

6) มอเตอร์ไซต์มวล 300 กิโลกรัม แล่นขึ้นพื้นเอียงทำมุม 30 องศากับแนวระดับ ด้วยอัตราเร็วคงที่ 120 เมตร/วินาที ถ้าแรงเสียดทานระหว่างพื้นเอียงกับมอเตอร์ไซต์เท่ากับ 200 นิวตัน จงหากำลังของเครื่องยนต์ในหน่วยกิโลวัตต์ (4 คะแนน)



$\Sigma F = 0$

$F - (f + mg\sin 30^\circ) = 0$

$F - (200 + (300)(10)\left(\frac{1}{2}\right)) = 0$

$F - (200 + 1500) = 0$

$F = 1700$ N

$P = \frac{W}{t}$

$P = \frac{Fs}{t}$

$P = Fv$

$P = 1,700 \times 120$

$P = 204,000$ W

$P = \boxed{204 \text{ KW}}$ #

7) ดาวต้องการยิงลูกหินมวล 100 กรัม ด้วยหนังสติ๊กให้ขึ้นไปได้สูง 10 เมตรในแนวดิ่ง ถ้าการยืดของหนังสติ๊กเป็นแบบสปริงและยืดออก 50 เซนติเมตร หนังสติ๊กจะต้องมีค่านิจสปริงกี่นิวตัน/เมตร

(4 คะแนน)

<u>วิธีทำ</u> จากกฎการอนุรักษ์พลังงาน กล่าวว่า พลังงานรวมของวัตถุ ณ ตำแหน่งใดๆย่อมมีความค่าคงตัว นั่นคือ $\sum E_{\text{ตอนแรก}} = \sum E_{\text{ตอนหลัง}}$

จะได้ว่า $mgh = \frac{1}{2}kx^2$

แทนค่า $(0.1)(10)(10) = \frac{1}{2}k(0.5)^2$

$10 = \frac{1}{2}k(0.25)$

เพราะฉะนั้น $k = \frac{10(2)}{0.25} = 80$ นิวตัน/เมตร

ดังนั้น หนังสติ๊กจะต้องมีค่านิจสปริง 80 นิวตัน/เมตร

เกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงาน ข้อ 7

ระดับคะแนน	คำอธิบาย
4	☐ระบุสมการของพลังงานได้ถูกต้องทั้ง 2 เทอม (พลังงานศักย์โน้มถ่วงและศักย์ยืดหยุ่น) ☐แทนค่าพลังงานศักย์โน้มถ่วง<u>และ</u>พลังงานจลน์ได้ถูกต้อง ☐หาค่าค่านิจสปริงได้ถูกต้อง
3	☐ระบุสมการของพลังงานได้ถูกต้องทั้ง 2 เทอม (พลังงานศักย์โน้มถ่วงและศักย์ยืดหยุ่น) ☐แทนค่าพลังงานศักย์โน้มถ่วง<u>และ</u>พลังงานจลน์ได้ถูกต้อง ☐แต่ไม่สามารถหาค่าสปริงได้ถูกต้อง
2	☐ระบุสมการของพลังงานได้ถูกต้องทั้ง 2 เทอม (พลังงานศักย์โน้มถ่วงและศักย์ยืดหยุ่น) ☐แทนค่าพลังงานศักย์โน้มถ่วง<u>หรือ</u>พลังงานจลน์ได้ถูกต้อง ☐แต่ไม่สามารถหาค่าสปริงได้ถูกต้อง
1	☐ระบุสมการของพลังงานได้ถูกต้องทั้ง 2 เทอม (พลังงานศักย์โน้มถ่วงและศักย์ยืดหยุ่น) ☐แต่ไม่สามารถแทนค่าพลังงานศักย์โน้มถ่วง<u>หรือ</u>พลังงานจลน์ได้ถูกต้อง และไม่สามารถหาค่าสปริงได้ถูกต้อง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตัวอย่างการแสดงวิธีทำสำหรับระดับคะแนนต่างๆ

ระดับคะแนน 1

7) ดาวต้องการยิงลูกหินมวล 100 กรัม ด้วยหนังสติ๊กให้ขึ้นไปได้สูง 10 เมตรในแนวดิ่ง ถ้าการยืดของหนังสติ๊กเป็นแบบสปริงและยืดออก 50 เซนติเมตร หนังสติ๊กจะต้องมีค่านิจสปริงกี่นิวตัน/เมตร (4 คะแนน)

$E_p = E_k$

$mgh = \frac{1}{2}kx^2$

$100 \times 10 \times 10 = \frac{1}{2} k \times 50^2$

$10000 = \frac{1}{2} k \times 2500$

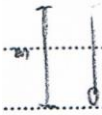
$\frac{10000}{1250} = k$

$k = 8$ N/m

∴ หนังสติ๊กจะต้องมีค่านิจสปริง 8 N/m

ระดับคะแนน 2

7) ดาวต้องการยิงลูกหินมวล 100 กรัม ด้วยหนังสติ๊กให้ขึ้นไปได้สูง 10 เมตรในแนวดิ่ง ถ้าการยืดของหนังสติ๊กเป็นแบบสปริงและยืดออก 50 เซนติเมตร หนังสติ๊กจะต้องมีค่านิจสปริงกี่นิวตัน/เมตร (4 คะแนน)



$E_{ps} = E_p$

$\frac{1}{2}kx^2 = mgh$

$\frac{1}{2}k(50)^2 = 0.1 \cdot 10 \cdot 10$

$\frac{1}{2} \cdot 2500\, k = 10$

$k = \frac{10}{1250}$

$k = \frac{1}{125} = 0.008$

ระดับคะแนน 3

7) ดาวต้องการยิงลูกหินมวล 100 กรัม ด้วยหนังสติ๊กให้ขึ้นไปได้สูง 10 เมตรในแนวดิ่ง ถ้าการยืดของหนังสติ๊กเป็นแบบสปริงและยืดออก 50 เซนติเมตร หนังสติ๊กจะต้องมีค่านิจสปริงกี่นิวตัน/เมตร (4 คะแนน)

$E_p = mgh$

$E_p = 0.1(10)(10)$

$E_p = 10$ J

$10 = \frac{1}{2}kx^2$

$10 = \frac{1}{2}(k)(0.5 \times 0.5)$

$10 = \frac{1}{2}(k)(0.25)$

$k = \frac{20}{0.25}$

$k = 5$ N/m

ระดับคะแนน 4

7) ดาวต้องการยิงลูกหินมวล 100 กรัม ด้วยหนังสติ๊กให้ขึ้นไปได้สูง 10 เมตรในแนวดิ่ง ถ้าการยืดของหนังสติ๊กเป็นแบบสปริงและยืดออก 50 เซนติเมตร หนังสติ๊กจะต้องมีค่านิจสปริงกี่นิวตัน/เมตร (4 คะแนน)

$mgh = \frac{1}{2}k(0.5)^2$

$0.1(10)(10) = \frac{1}{2}k(0.25)$

$10 = \frac{1}{2}k(0.25)$

$\frac{20}{0.25} = k$

$k = 80$ N/m.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัม ขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)

<u>วิธีทำ</u> (แบบที่ 1) ประสิทธิภาพของเครื่องกล (E_{ff}) หาได้จาก

$$E_{ff} = \frac{\text{กำลังที่ได้จริง}}{\text{กำลังที่ได้รับ}} \times 100 = \frac{\frac{\text{พลังงาน}}{\text{เวลา}}}{\text{กำลังที่ได้รับ}} \times 100 = \frac{\frac{mgh}{t}}{\text{กำลังที่ได้รับ}} \times 100$$

$$= \frac{\frac{(750+450)(10)(8)}{4}}{30{,}000} \times 100 = \frac{24{,}000}{30{,}000} \times 100 = 0.8 \times 100 = 80\%$$

ดังนั้น ลิฟท์นี้ต้องมีประสิทธิภาพ 80%

เกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงาน ข้อ 8

ระดับคะแนน	คำอธิบาย
4	☐ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง ☐แทนค่า<u>กำลังที่ได้จริง</u>ถูกต้อง ☐แทนค่าสมการประสิทธิภาพเครื่องกลได้ถูกต้อง ☐หาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง
3	☐ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง ☐แทนค่า<u>กำลังที่ได้จริง</u>ถูกต้อง ☐แทนค่าสมการประสิทธิภาพเครื่องกลได้ถูกต้อง ☐แต่ไม่สามารถหาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง
2	☐ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง ☐แทนค่า<u>กำลังที่ได้จริง</u>ถูกต้อง ☐แต่ไม่สามารถแทนค่าในสมการประสิทธิภาพเครื่องกลได้ถูกต้อง และไม่สามารถหาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง
1	☐ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง <u>หรือ</u>แทนค่า<u>กำลังที่ได้จริง</u>ถูกต้อง ☐แต่ไม่สามารถแทนค่าในสมการประสิทธิภาพเครื่องกลได้ถูกต้อง และไม่สามารถหาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตัวอย่างการแสดงวิธีทำสำหรับระดับคะแนนต่างๆ

ระดับคะแนน 1

8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัม ขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)

Sol ประสิทธิภาพ = กำลังที่ได้ / กำลังที่ใช้ × 100 %

3000 = 7500 / x

x = 7500 / 3000

x = 2.5

P = 4500(8) / 4

= 900 W

ตอบ 30 %

ระดับคะแนน 2

8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัมขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)

ประสิทธิภาพ = $P_{out}/P_{in} \times 100$

$= \frac{24000}{300000} \times 100$

$= 0.08 \times 100$

$= 8\ \%$ ตอบ

$P_{out} = W/t$

$= mgh/t$

$= \frac{(1200)(10)(8)}{4}$

$= 24000$

ระดับคะแนน 3

8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัมขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)

$m = 750\ kg + 450 = 1200\ kg$

$Eff = \frac{\text{กำลัง Out}}{\text{กำลัง in}} \times 100$

$= \frac{(1200)(10)(8)^2}{4} \times 100$

$30,000$

$= 0.8\ \%$ ตอบ

ระดับคะแนน 4

8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัมขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)

ประสิทธิภาพ $= \frac{\text{กำลังที่ได้}}{\text{กำลังที่ใช้}} \times 100 = \frac{24,000}{30,000} \times 100 = 0.8 \times 100 = 80\ \%$

กำลังที่ได้ $= \frac{mgh}{t} = \frac{(750+450)(10)(8)}{4} = 1200(10)(2) = 24,000$

กำลังที่ใช้ $= 30,000$

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

<u>วิธีทำ</u> (แบบที่ 2) ประสิทธิภาพของเครื่องกล (E_{ff}) หาได้จาก

$$E_{ff} = \frac{\text{งานที่ได้จริง}}{\text{งานที่ได้รับ}} \times 100 = \frac{mgh}{w \times t} \times 100 = \frac{(750+450)(10)(8)}{(30{,}000)(4)} \times 100$$

$$= \frac{(1{,}200)(10)(8)}{120{,}000} \times 100 = \frac{96{,}000}{120{,}000} \times 100 = 0.8 \times 100 = 80\%$$

ดังนั้น ลิฟท์นี้ต้องมีประสิทธิภาพ 80%

เกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงาน ข้อ 8

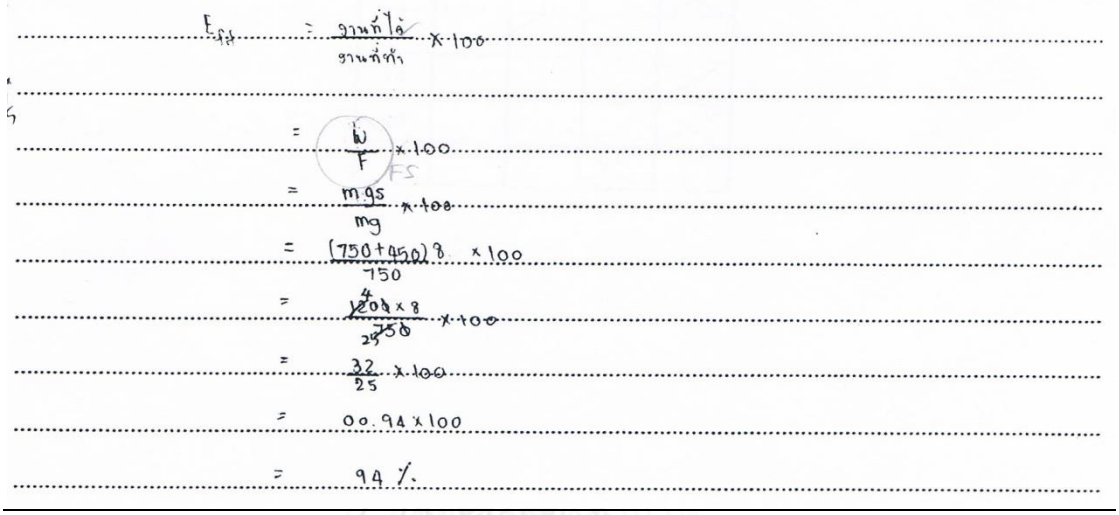
ระดับคะแนน	คำอธิบาย
4	☐ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง ☐แทนค่า<u>งานที่ได้จริง</u>ถูกต้อง ☐แทนค่าสมการประสิทธิภาพเครื่องกลได้ถูกต้อง ☐หาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง
3	☐ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง ☐แทนค่า<u>งานที่ได้จริง</u>ถูกต้อง ☐แทนค่าสมการประสิทธิภาพเครื่องกลได้ถูกต้อง ☐แต่ไม่สามารถหาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง
2	☐ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง ☐แทนค่า<u>งานที่ได้จริง</u>ถูกต้อง ☐แต่ไม่สามารถแทนค่าในสมการประสิทธิภาพเครื่องกลได้ถูกต้อง และไม่สามารถหาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง
1	☐ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง <u>หรือ</u>แทนค่า<u>งานที่ได้จริง</u>ถูกต้อง ☐แต่ไม่สามารถแทนค่าในสมการประสิทธิภาพเครื่องกลได้ถูกต้อง และไม่สามารถหาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตัวอย่างการแสดงวิธีทำสำหรับระดับคะแนนต่างๆ

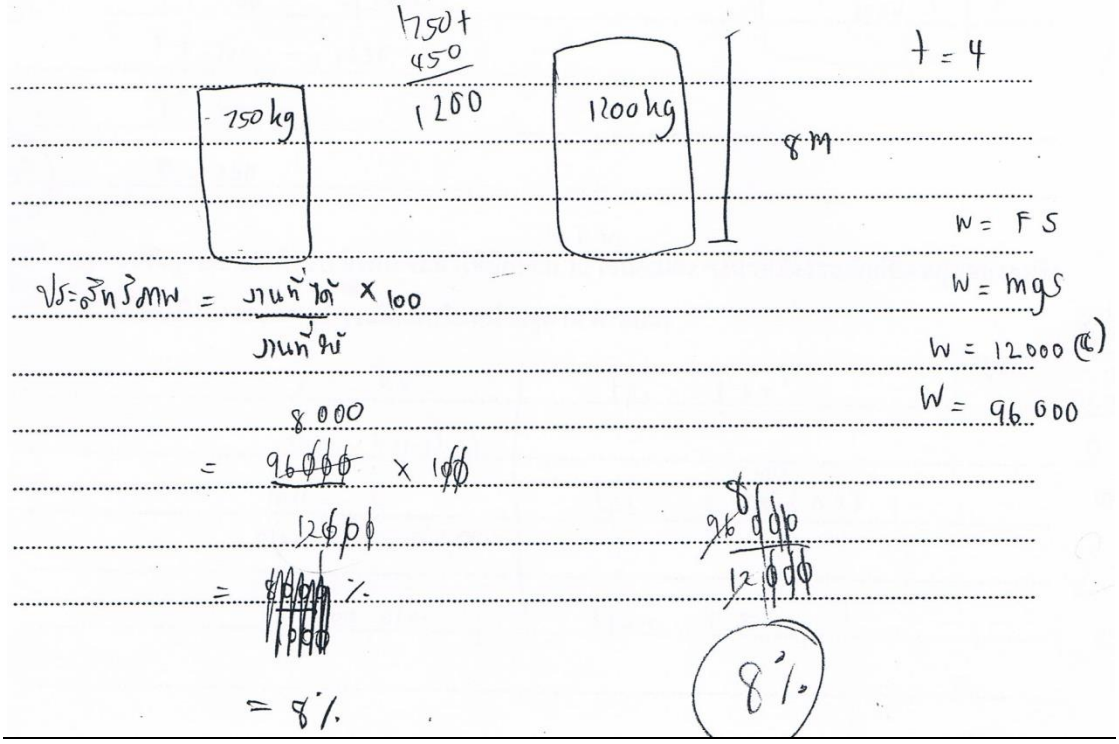
ระดับคะแนน 1

8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัมขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)



ระดับคะแนน 2

8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัมขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)



ระดับคะแนน 3

8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัมขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)

$\eta = \frac{P_{out}}{P_{in}} \times 100 = \frac{W_{out}}{W_{in}} \times 100$

$E_{ff} = \frac{W_{out}}{W_{in}} \times 100$

$= \frac{(1200)(10)(8)}{(30\times10^3)(4)} \times 100$

$= 0.8\ \%$

ประสิทธิภาพ = 0.8 %

ระดับคะแนน 4

8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัมขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)

ประสิทธิภาพ $= \frac{W_{out}}{W_{in}} \times 100$

$W_{in} = Pt = (30\times10^3)(4) = 120\times10^3$

$W_{out} = mgh = (750+450)(10)(8) = (1200)(10)(8) = 96000$

ประสิทธิภาพ $= \frac{96000}{120\times10^3} \times 100$

$= \frac{96}{12} \times 10 = 8\times10 = 80\ \%$

<u>วิธีทำ</u> (แบบที่ 3) ประสิทธิภาพของเครื่องกล (E_{ff}) หาได้จาก

$$E_{ff} = \frac{\text{แรงที่ได้จริง}}{\text{แรงที่ได้รับ}} \times 100 = \frac{mg}{\frac{P \times t}{s}} \times 100 = \frac{(750+450)(10)}{\frac{(30{,}000)(4)}{8}} \times 100$$

$$= \frac{(1{,}200)(10)}{\frac{120{,}000}{8}} \times 100 = \frac{12{,}000}{15{,}000} \times 100 = 0.8 \times 100 = 80\%$$

ดังนั้น ลิฟท์นี้ต้องมีประสิทธิภาพ 80%

เกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงาน ข้อ 8

ระดับคะแนน	คำอธิบาย
4	☐ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง ☐แทนค่า<u>แรงที่ได้จริง</u>ถูกต้อง ☐แทนค่าสมการประสิทธิภาพเครื่องกลได้ถูกต้อง ☐หาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง
3	☐ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง ☐แทนค่า<u>แรงที่ได้จริง</u>ถูกต้อง ☐แทนค่าสมการประสิทธิภาพเครื่องกลได้ถูกต้อง ☐แต่ไม่สามารถหาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง
2	☐ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง ☐แทนค่า<u>แรงที่ได้จริง</u>ถูกต้อง ☐แต่ไม่สามารถแทนค่าในสมการประสิทธิภาพเครื่องกลได้ถูกต้อง และไม่สามารถหาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง
1	☐ระบุสมการประสิทธิภาพของเครื่องกลได้ถูกต้อง <u>หรือ</u>แทนค่า<u>แรงที่ได้จริง</u>ถูกต้อง ☐แต่ไม่สามารถแทนค่าในสมการประสิทธิภาพเครื่องกลได้ถูกต้อง และไม่สามารถหาค่าประสิทธิภาพของเครื่องกลในหน่วยเปอร์เซ็นต์ได้ถูกต้อง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตัวอย่างการแสดงวิธีทำสำหรับระดับคะแนนต่างๆ

ระดับคะแนน 1

8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัมขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)

ประสิทธิภาพ = แรงที่ได้ / แรงที่ให้ × 100

= (750)(10) / 30,000 × 100

ประสิทธิภาพ = 75/3 = 25 %

ระดับคะแนน 2

8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัมขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)

eff = $\frac{F_{out}}{F_{in}} \times 100\%$ → $F_{out} = mg$

= (750 + 450)(10)

= (1200)(10)

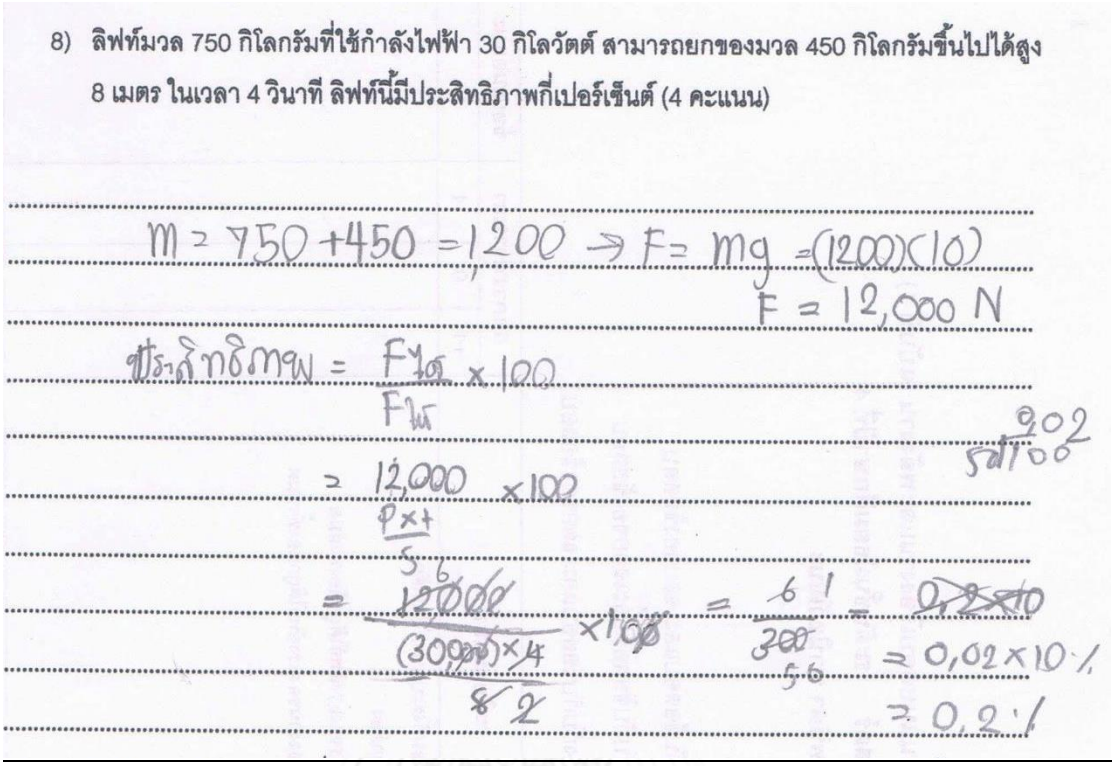
= 12000 N

= 12000/30000 × 100 %

ตอบ eff = 40 %

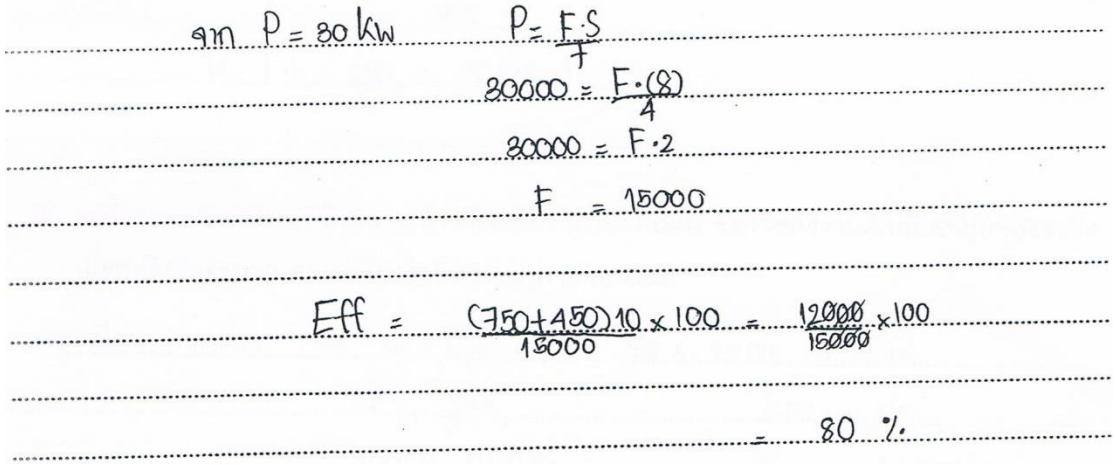
ระดับคะแนน 3

8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัมขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)



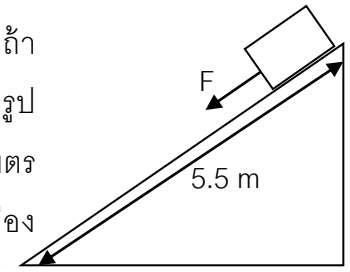
ระดับคะแนน 4

8) ลิฟท์มวล 750 กิโลกรัมที่ใช้กำลังไฟฟ้า 30 กิโลวัตต์ สามารถยกของมวล 450 กิโลกรัมขึ้นไปได้สูง 8 เมตร ในเวลา 4 วินาที ลิฟท์นี้มีประสิทธิภาพกี่เปอร์เซ็นต์ (4 คะแนน)



เฉลยและเกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงานฉบับอ้างอิง

5) ชายคนหนึ่งออกแรงดึงกล่องหนัก 80 กิโลกรัมลงเนินที่เอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ ถ้าเขาออกแรงดึงกล่องในแนวขนานกับการเคลื่อนที่ดังรูป เขาต้องใช้งานเท่าใดในการดึงกล่องลงมาได้ 5.5 เมตร เมื่อสัมประสิทธิ์ความเสียดทานระหว่างเนินและกล่องเท่ากับ $\sqrt{3}/2$ (4 คะแนน)



<u>วิธีทำ</u> เนื่องจากความเร็วคงที่ แรงลัพธ์ในการเคลื่อนที่จึงเท่ากับศูนย์

นั่นคือ $$\sum \bar{F} = 0$$

แรงซ้าย = แรงขวา

$$F + mgsin\theta = f$$

$$F + mgsin\theta = \mu mgcos\theta$$

จะได้ว่า $$F = \mu mgcos30° - mgsin30°$$

$$= \frac{\sqrt{3}}{2}(80)(10)\left(\frac{\sqrt{3}}{2}\right) - (80)(10)(\tfrac{1}{2}) = 200$$

จาก $$W = Fscos\theta = 200(5.5)cos0° = 1{,}100 \text{ จูล}$$

ดังนั้น ช้างต้องใช้งาน 1,100 จูลในการดึงกล่องลงมาได้ 5.5 เมตร

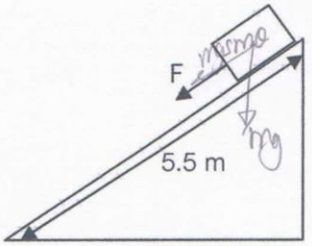
เกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงาน ข้อ 5

ระดับคะแนน	คำอธิบาย
4	☐ระบุสมการในการหางานและสมการของแรงลัพธ์ได้ถูกต้อง ☐แทนค่าสมการของแรงลัพธ์และสมการของงานได้ถูกต้อง ☐หาค่างานพร้อมทั้งระบุหน่วยได้ถูกต้อง
3	☐ระบุสมการในการหางานและสมการของแรงลัพธ์ได้ถูกต้อง ☐แทนค่าสมการของแรงลัพธ์และสมการของงานได้ถูกต้อง ☐แต่ไม่สามารถหาค่างานได้ถูกต้อง
2	☐ระบุสมการในการหางานและสมการของแรงลัพธ์ได้ถูกต้อง ☐แทนค่าสมการของแรงลัพธ์ได้ถูกต้อง ☐แต่ไม่สามารถแทนค่าสมการของงานได้ถูกต้อง และไม่สามารถหาค่างานได้ถูกต้อง
1	☐ระบุสมการของงานของแรงลัพธ์ได้ถูกต้อง ☐แต่ไม่สามารถแทนค่าสมการของแรงลัพธ์และสมการของงานได้ถูกต้อง และไม่สามารถหาค่าแรงลัพธ์งานได้ถูกต้อง

ตัวอย่างการแสดงวิธีทำสำหรับระดับคะแนนต่างๆ

<u>ระดับคะแนน 1</u>

5) ชายคนหนึ่งออกแรงดึงกล่องหนัก 80 กิโลกรัมลงเนินที่เอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ ถ้าเขาออกแรงดึงกล่องในแนวขนานกับการเคลื่อนที่ดังรูป เขาต้องใช้งานเท่าใดในการดึงกล่องลงมาได้ 5.5 เมตร เมื่อสัมประสิทธิ์ความเสียดทานระหว่างเนินและกล่องเท่ากับ $\sqrt{3}/2$ (4 คะแนน)



$F + mg\sin\theta - f = 0$

$F = f - mg\sin\theta$

$= \mu mg\cos\theta - mg\sin\theta$

$= \mu mg(\cos\theta - \sin\theta)$

$= \frac{\sqrt{3}}{2}(80)(10)\left(\frac{\sqrt{3}}{2} - \frac{1}{2}\right)$

$W = FS\cos\theta^{\circ}$

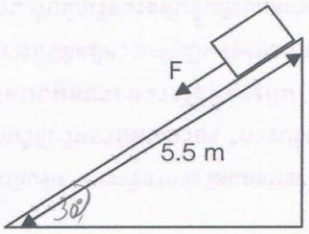
$= FS$

$= 400\sqrt{3}(\sqrt{3}-1)$

$= 200\sqrt{3}(\sqrt{3}-1)$ J

<u>ระดับคะแนน 2</u>

5) ชายคนหนึ่งออกแรงดึงกล่องหนัก 80 กิโลกรัมลงเนินที่เอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ ถ้าเขาออกแรงดึงกล่องในแนวขนานกับการเคลื่อนที่ดังรูป เขาต้องใช้งานเท่าใดในการดึงกล่องลงมาได้ 5.5 เมตร เมื่อสัมประสิทธิ์ความเสียดทานระหว่างเนินและกล่องเท่ากับ $\sqrt{3}/2$ (4 คะแนน)



$m = 80$ kg $\quad \mu = \frac{\sqrt{3}}{2} \quad S = 5.5$ m

$F = f - mg\sin\theta$

$= \mu mg\cos\theta - mg\sin\theta$

$= \frac{\sqrt{3}}{2}(80)(10)\left(\frac{\sqrt{3}}{2}\right) - (80)(10)\left(\frac{1}{2}\right)$

$= 2400 - 400 = 2000$

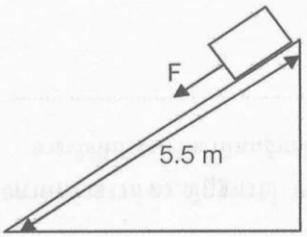
$W = FS$

$= (2000)(5.5)$

$W = 11,000$ J

ระดับคะแนน 3

5) ชายคนหนึ่งออกแรงดึงกล่องหนัก 80 กิโลกรัมลงเนินที่เอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ ถ้าเขาออกแรงดึงกล่องในแนวขนานกับการเคลื่อนที่ดังรูป เขาต้องใช้งานเท่าใดในการดึงกล่องลงมาได้ 5.5 เมตร เมื่อสัมประสิทธิ์ความเสียดทานระหว่างเนินและกล่องเท่ากับ $\sqrt{3}/2$ (4 คะแนน)



$\Sigma F = 0$

$F + mg\sin\theta = f$

$F = f - mg\sin\theta$

$= \mu mg\cos\theta - mg\sin\theta$

$= \frac{\sqrt{3}}{2}(80)(10)\left(\frac{\sqrt{3}}{2}\right) - (80)(10)\left(\frac{1}{2}\right)$

$F = 600 - 400 = 200$

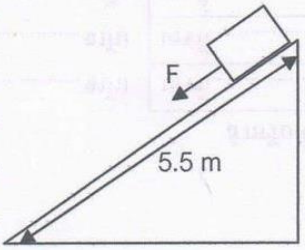
$W = FS$

$= (200)(5.5)$

$W = 11{,}000$ J

ระดับคะแนน 4

5) ชายคนหนึ่งออกแรงดึงกล่องหนัก 80 กิโลกรัมลงเนินที่เอียงทำมุม 30 องศากับแนวระดับด้วยความเร็วคงที่ ถ้าเขาออกแรงดึงกล่องในแนวขนานกับการเคลื่อนที่ดังรูป เขาต้องใช้งานเท่าใดในการดึงกล่องลงมาได้ 5.5 เมตร เมื่อสัมประสิทธิ์ความเสียดทานระหว่างเนินและกล่องเท่ากับ $\sqrt{3}/2$ (4 คะแนน)



$\Sigma F = 0$

$F + mg\sin 30^\circ - \mu mg\cos 30^\circ = 0$

$F = \mu mg\cos 30^\circ - mg\sin 30^\circ$

$= \frac{\sqrt{3}}{2}(80)(10)\left(\frac{\sqrt{3}}{2}\right) - (80)(10)\left(\frac{1}{2}\right)$

$= 600 - 400$

$F = 200$

$W = FS\cos 0^\circ$

$= (200)(5.5)$

$W = 1{,}100$ J

6) สปริงถูกดึงด้วยแรง 40 นิวตัน จะทำให้สปริงยืดออก 20 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 30 เซนติเมตรในหน่วยจูล (4 คะแนน)

<u>วิธีทำ</u> หาค่านิจสปริงจาก $F = kx$

แทนค่า $40 = k(0.2)$

จะได้ว่า $k = \frac{40}{0.2} = 200$ นิวตัน/เมตร

หาพลังงานศักย์ยืดหยุ่น (E_P) จาก ;

$$E_P = \frac{1}{2}kx^2$$

แทนค่า $= \frac{1}{2}(200)(0.3)^2 = \frac{1}{2}(200)(0.09) = 9$ จูล

ดังนั้น พลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 20 เซนติเมตรจะเท่ากับ 9 จูล

เกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงาน ข้อ 6

ระดับคะแนน	คำอธิบาย
4	☐ระบุสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง ☐แทนค่าสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง ☐หาค่าพลังงานศักย์ยืดหยุ่นได้ถูกต้อง
3	☐ระบุสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง ☐แทนค่าสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง ☐ แต่ไม่สามารถหาค่าพลังงานศักย์ยืดหยุ่นได้ถูกต้อง
2	☐ระบุสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง ☐แทนค่าสมการตามกฎของฮุคได้ถูกต้อง ☐แต่ไม่สามารถแทนค่าสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง และไม่สามารถหาค่าพลังงานศักย์ยืดหยุ่นได้ถูกต้อง
1	☐ระบุสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง ☐แต่ไม่สามารถแทนค่าสมการตามกฎของฮุคและสมการของพลังงานศักย์ยืดหยุ่นได้ถูกต้อง และไม่สามารถหาค่างานได้ถูกต้อง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตัวอย่างการแสดงวิธีทำสำหรับระดับคะแนนต่างๆ

ระดับคะแนน 1

6) สปริงถูกดึงด้วยแรง 40 นิวตัน จะทำให้สปริงยืดออก 20 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 30 เซนติเมตรในหน่วยจูล (4 คะแนน)

$F = 40\ N,\ x = 20$ | $x = 30$

$F = kx$ | $E_p = \frac{1}{2}kx^2$

$40 = k(20)$

$k = 2$ | $E_p = \frac{1}{2}(2)(30)^2 = 900\ J$

ระดับคะแนน 2

6) สปริงถูกดึงด้วยแรง 40 นิวตัน จะทำให้สปริงยืดออก 20 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 30 เซนติเมตรในหน่วยจูล (4 คะแนน)

$F = kx$ | $E_p = \frac{1}{2}kx^2$

$40 = k(0.2)$ | $= \frac{1}{2}(2.0)(0.3)^2$

$k = \frac{40}{0.2} = 20$ | $E_p = (10)(0.09) = 0.9\ J$

ระดับคะแนน 3

6) สปริงถูกดึงด้วยแรง 40 นิวตัน จะทำให้สปริงยืดออก 20 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 30 เซนติเมตรในหน่วยจูล (4 คะแนน)

$k = \frac{F}{s}$

$k = \frac{20}{0.1} = 200\ N/m$

$E_{ps} = \frac{1}{2}kx^2$

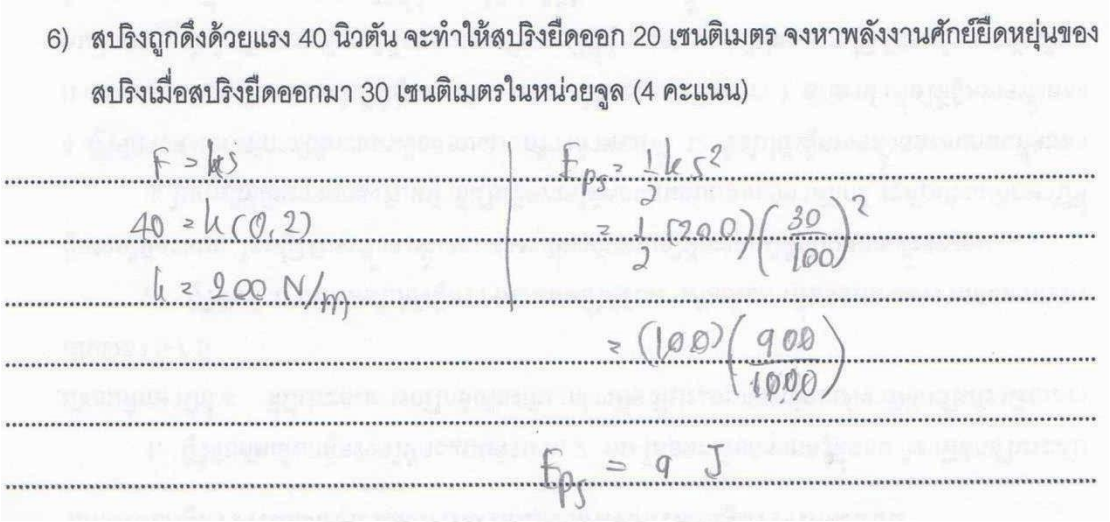
$= \frac{1}{2}(200)(0.3)^2$

$= (100)(0.9)$

$E_{ps} = 90$

ระดับคะแนน 4

6) สปริงถูกดึงด้วยแรง 40 นิวตัน จะทำให้สปริงยืดออก 20 เซนติเมตร จงหาพลังงานศักย์ยืดหยุ่นของสปริงเมื่อสปริงยืดออกมา 30 เซนติเมตรในหน่วยจูล (4 คะแนน)



7) เครื่องสูบน้ำใช้สูบน้ำมวล 500 กิโลกรัม ขึ้นจากบ่อลึก 6 เมตร แล้วฉีดน้ำออกจากปลายท่อด้วยความเร็ว 6 เมตรต่อวินาที งานที่ได้จากเครื่องสูบน้ำนี้มีค่ากี่กิโลจูล (4 คะแนน)

วิธีทำ เนื่องจากงานในการเคลื่อนที่สามารถหาได้จากพลังงานที่เปลี่ยนแปลงไป

จะได้ว่า $W = mgh + \frac{1}{2}mv^2$

แทนค่า $= (500)(10)(6) + \frac{1}{2}(500)(6)^2 = 30{,}000 + 9{,}000$

$= 39{,}000$ จูล

ดังนั้น งานที่ได้จากเครื่องสูบมีค่า 39,000 จูล หรือ 39 กิโลจูล

เกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงาน ข้อ 7

ระดับคะแนน	คำอธิบาย
4	☐ระบุสมการของงานได้ถูกต้องทั้ง 2 เทอม ☐แทนค่าสมการของพลังงานศักย์โน้มถ่วงและสมการของพลังงานจลน์ ได้ถูกต้อง ☐หาค่างานได้ถูกต้อง
3	☐ระบุสมการของงานได้ถูกต้องทั้ง 2 เทอม ☐แทนค่าสมการของพลังงานศักย์โน้มถ่วงและสมการของพลังงานจลน์ ได้ถูกต้อง ☐แต่ไม่สามารถหางานได้ถูกต้อง
2	☐ระบุสมการของงานได้ถูกต้องทั้ง 2 เทอม ☐แทนค่าพลังงานศักย์โน้มถ่วงหรือพลังงานจลน์ ได้ถูกต้อง ☐แต่ไม่สามารถหางานได้ถูกต้อง
1	☐ระบุสมการของงานได้ถูกต้องทั้ง 2 เทอม ☐แต่ไม่สามารถแทนค่าสมการของพลังงานศักย์โน้มถ่วงหรือสมการของพลังงานจลน์ ได้ถูกต้อง และไม่สามารถหางานได้ถูกต้อง

ตัวอย่างการแสดงวิธีทำสำหรับระดับคะแนนต่างๆ

ระดับคะแนน 1

7) เครื่องสูบน้ำใช้สูบน้ำมวล 500 กิโลกรัม ขึ้นจากบ่อลึก 6 เมตร แล้วฉีดน้ำออกจากปลายท่อด้วยความเร็ว 6 เมตรต่อวินาที งานที่ได้จากเครื่องสูบน้ำนี้มีค่ากี่กิโลจูล (4 คะแนน)

$m = 500$ kg, $h = 6$ m, $v = 6$ m/s

$W = E_p + E_k$

$= mgh + \frac{1}{2}mv^2$

$= gh + \frac{1}{2}v^2$

$= (10)(6) + \frac{1}{2}(6)^2$

$= 60 + 18$

$W = 78$ J

ระดับคะแนน 2

7) เครื่องสูบน้ำใช้สูบน้ำมวล 500 กิโลกรัม ขึ้นจากบ่อลึก 6 เมตร แล้วฉีดน้ำออกจากปลายท่อด้วยความเร็ว 6 เมตรต่อวินาที งานที่ได้จากเครื่องสูบน้ำนี้มีค่ากี่กิโลจูล (4 คะแนน)

$E_p + E_k = W$

$(500)(10)(6) + \frac{1}{2}(500)(120) = W$

$30,000 + 30,000 = W$

$60,000 = W$

$W = 60,000 = 60$ kJ

$\frac{1}{2}mv^2 = mgh$

$\frac{1}{2}v^2 = (10)(6)$

$v = 120$

$v = \sqrt{120}$

ระดับคะแนน 3

7) เครื่องสูบน้ำใช้สูบน้ำมวล 500 กิโลกรัม ขึ้นจากบ่อลึก 6 เมตร แล้วฉีดน้ำออกจากปลายท่อด้วยความเร็ว 6 เมตรต่อวินาที งานที่ได้จากเครื่องสูบน้ำนี้มีค่ากี่กิโลจูล (4 คะแนน)

$W_{1\to 2} = E_k + E_p$

$= \frac{1}{2}mv^2 + mgh$

$= \frac{1}{2}(500)(6^2) + (500)(10)(6)$

$= \frac{3000 + 30,000}{2}$

$= 1,500 + 30,000$

$= 31,500 = 31.5$ kJ

ระดับคะแนน 4

7) เครื่องสูบน้ำใช้สูบน้ำมวล 500 กิโลกรัม ขึ้นจากบ่อลึก 6 เมตร แล้วฉีดน้ำออกจากปลายท่อด้วยความเร็ว 6 เมตรต่อวินาที งานที่ได้จากเครื่องสูบน้ำนี้มีค่ากี่กิโลจูล (4 คะแนน)

$m = 500$, $h = 6$ m, $v = 6$ m/s

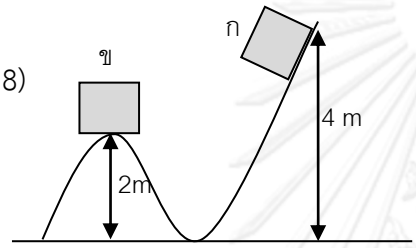
$W = E_p + E_k$

$= mgh + \frac{1}{2}mv^2$

$= (500)(10)(6) + \frac{1}{2}(500)(6)^2$

$= 30{,}000 + 9{,}000$

$W = 39{,}000 = 39$ kJ

8) 

กล่องมวล 6 กิโลกรัม ที่จุด ก มีอัตราเร็ว 1 เมตร/วินาที และที่จุด ข มีอัตราเร็ว 7 เมตร/วินาที ถ้าระยะทางจากจุด ก ไปยังจุด ข เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (4 คะแนน)

วิธีทำ (แบบที่ 1)

เมื่อกำหนดให้เนินที่จุด ข เป็นจุดอ้างอิง จะได้ว่า

$$E_{P1} + E_{k1} - W_f = E_{k2}$$

จาก $$mgh + \frac{1}{2}mv^2 - fs = \frac{1}{2}mv^2$$

แทนค่า $$(6)(10)(2) + \frac{1}{2}(6)(1)^2 - f(3) = \frac{1}{2}(6)(7)^2$$

จะได้ว่า $$120 + 3 - f(3) = 147$$

$$-f(3) = 147 - 123$$

$$f = -\frac{24}{3} = -8 \text{ นิวตัน}$$

ดังนั้นแรงเสียดทานระหว่างพื้นและกล่องมีค่า - 8 นิวตัน

วิธีทำ (แบบที่ 2)

เมื่อกำหนดให้พื้นราบ เป็นจุดอ้างอิง จะได้ว่า

$$E_{P1} + E_{k1} - W_f = E_{k2} + E_{P2}$$

จาก $$mgh + \frac{1}{2}mv^2 - fs = \frac{1}{2}mv^2 + mgh$$

แทนค่า $$(6)(10)(4) + \frac{1}{2}(6)(1)^2 - f(3) = \frac{1}{2}(6)(7)^2 + (6)(10)(2)$$

จะได้ว่า $$240 + 3 - f(3) = 147 + 120$$

$$-f(3) = 24$$

$$f = -\frac{24}{3} = -8 \text{ นิวตัน}$$

ดังนั้นแรงเสียดทานระหว่างพื้นและกล่องมีค่า - 8 นิวตัน

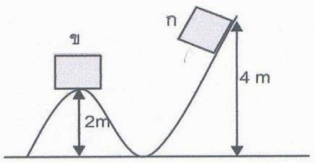
เกณฑ์การให้คะแนนสำหรับโจทย์ปัญหาเรื่องงานและพลังงาน ข้อ 8

ระดับคะแนน	คำอธิบาย
4	☐ระบุสมการพลังงานได้ถูกต้องทั้งหมด ☐แทนค่าสมการของพลังงานได้ถูกต้องทั้งหมด ☐หาค่าแรงเสียดทานได้ถูกต้อง
3	☐ระบุสมการพลังงานได้ถูกต้องทั้งหมด ☐แทนค่าสมการของพลังงานได้ถูกต้องทั้งหมด ☐แต่ไม่สามารถหาค่าแรงเสียดทานพร้อมได้ถูกต้อง
2	☐ระบุสมการพลังงานได้ถูกต้องทั้งหมด ☐แทนค่าสมการของพลังงานได้ถูกต้องอย่างต่ำ 2 เทอม ☐แต่ไม่สามารถหาค่าแรงเสียดทานพร้อมได้ถูกต้อง
1	☐ระบุสมการพลังงานได้ถูกต้องทั้งหมด ☐แต่ไม่สามารถแทนค่าสมการของพลังงานได้ถูกต้องอย่างต่ำ 2 เทอม และไม่สามารถหาค่าแรงเสียดทานพร้อมได้ถูกต้อง

ตัวอย่างการแสดงวิธีทำสำหรับระดับคะแนนต่างๆ

<u>ระดับคะแนน 1</u>

8) กล่องมวล 6 กิโลกรัม ที่จุด ก มีอัตราเร็ว 1 เมตร/วินาที และที่จุด ข มีอัตราเร็ว 7 เมตร/วินาที ถ้าระยะทางจากจุด ก ไปยังจุด ข เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (4 คะแนน)



$\Sigma E_{ก} = \Sigma E_{ข}$

$E_{p_ก} + E_{k_ก} = E_{p_ข} + E_{k_ข} + W_f$

$mgh_ก + \frac{1}{2}mv_ก^2 = mgh_ข + \frac{1}{2}mv_ข^2 + W_f$

$(10)(4) + \frac{1}{2}(6)(1^2) = (10)(2) + \frac{1}{2}(7^2) + 3(f)$

$40 + 3 = 20 + \frac{49}{2} + 3f$

$43 = \frac{89}{2} + 3f$

$\frac{43 - 89}{2} = 3f$

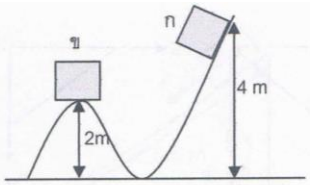
$\frac{86 - 89}{2} = 3f$

$\frac{-3}{2} = 3f$

$f = -\frac{1}{2} = -0.5$ N

ระดับคะแนน 2

8) กล่องมวล 6 กิโลกรัม ที่จุด ก มีอัตราเร็ว 1 เมตร/วินาที และที่จุด ข มีอัตราเร็ว 7 เมตร/วินาที ถ้าระยะทางจากจุด ก ไปยังจุด ข เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (4 คะแนน)



$E_{p_ก} + E_{k_ก} + W = E_{p_ข} + E_{k_ข}$

$mgh + \frac{1}{2}mv^2 + W = mgh + \frac{1}{2}mv^2$

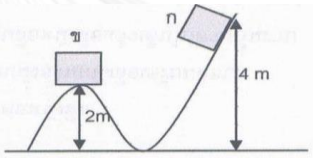
$(6)(10)(4) + \frac{1}{2}(6)(1^2) - W_f = (6)(10)(2) + \frac{1}{2}(6)(7^2)$

$240 + 3 - W_f = 120 + 147$

$W_f = 24$

ระดับคะแนน 3

8) กล่องมวล 6 กิโลกรัม ที่จุด ก มีอัตราเร็ว 1 เมตร/วินาที และที่จุด ข มีอัตราเร็ว 7 เมตร/วินาที ถ้าระยะทางจากจุด ก ไปยังจุด ข เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (4 คะแนน)



$W + mgh + \frac{1}{2}mv^2 = mgh + \frac{1}{2}mv^2$

$-3f + (6)(10)(4) + \frac{1}{2}(6)(1^2) = (6)(10)(2) + \frac{1}{2}(6)(7^2)$

$-3f + 240 + 3 = 120 + 84$

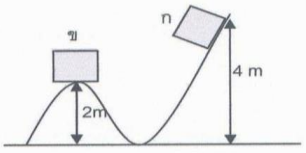
$-3f = 204 - 243$

$-3f = 39$

$f = -13$ N.

ระดับคะแนน 4

8) กล่องมวล 6 กิโลกรัม ที่จุด ก มีอัตราเร็ว 1 เมตร/วินาที และที่จุด ข มีอัตราเร็ว 7 เมตร/วินาที ถ้าระยะทางจากจุด ก ไปยังจุด ข เท่ากับ 3 เมตร แรงเสียดทานระหว่างพื้นและกล่องมีค่าเท่าใด (ในหน่วยนิวตัน) (4 คะแนน)



$E_{p_ก} + E_{k_ก} + W = E_{p_ข} + E_{k_ข}$

$(mgh)_ก + (\frac{1}{2}mv^2)_ก - fs = (mgh)_ข + (\frac{1}{2}mv^2)_ข$

$(6)(10)(4) + \frac{1}{2}(6)(1^2) - 3f = (6)(10)(2) + \frac{1}{2}(6)(7^2)$

$240 + 3 - 3f = 120 + 147$

$-3f = 267 - 243$

$-3f = 24$

$f = \frac{-24}{3} = -8$ N

ภาคผนวก ซ

การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (Root mean squared difference: RMSD)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 49 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบ ที่ 1 ด้วยวิธี CLE (Chained Linear Equating)

คะแนน	$e_i(x_i)$	$\hat{e}_i(x_i)$	$\hat{e}_i(x_i) - e_i(x_i)$	$W_i = f_i / 1084$	$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$
0	0.200063	-1.382849	-1.582912	0.000000	0.000000
1	1.503058	-0.270165	-1.773223	0.004613	0.014503
2	2.670360	0.842518	-1.827842	0.026753	0.089381
3	3.808869	1.955201	-1.853668	0.067343	0.231397
4	4.892185	3.067884	-1.824301	0.097786	0.325439
5	5.973298	4.180568	-1.792730	0.108856	0.349851
6	7.031998	5.293251	-1.738747	0.114391	0.345832
7	8.061973	6.405934	-1.656039	0.083026	0.227695
8	9.058244	7.518617	-1.539627	0.056273	0.133392
9	10.013506	8.631301	-1.382205	0.052583	0.100459
10	10.939139	9.743984	-1.195155	0.050738	0.072474
11	11.849230	10.856667	-0.992562	0.053506	0.052713
12	12.778435	11.969350	-0.809084	0.047970	0.031402
13	13.692005	13.082034	-0.609971	0.037823	0.014073
14	14.675540	14.194717	-0.480823	0.034133	0.007891
15	15.636421	15.307400	-0.329021	0.029520	0.003196
16	16.669097	16.420083	-0.249014	0.021218	0.001316
17	17.689501	17.532767	-0.156734	0.020295	0.000499
18	18.766604	18.645450	-0.121154	0.011993	0.000176
19	19.837862	19.758133	-0.079729	0.012915	0.000082
20	20.949317	20.870817	-0.078500	0.009225	0.000057
21	22.065847	21.983500	-0.082347	0.009225	0.000063
22	23.211694	23.096183	-0.115511	0.000923	0.000012
23	24.375050	24.208866	-0.166184	0.006458	0.000178
24	25.564593	25.321550	-0.243044	0.004613	0.000272
25	26.787601	26.434233	-0.353368	0.001845	0.000230
26	27.995020	27.546916	-0.448104	0.008303	0.001667
27	29.183145	28.659599	-0.523545	0.007380	0.002023
28	30.330229	29.772283	-0.557947	0.004613	0.001436
29	31.438102	30.884966	-0.553136	0.001845	0.000565
30	32.495797	31.997649	-0.498148	0.003690	0.000916
31	33.512015	33.110332	-0.401682	0.003690	0.000595
32	34.484269	34.223016	-0.261253	0.000923	0.000063

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 49 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบ ที่ 1 ด้วยวิธี CLE (Chained Linear Equating) (ต่อ)

คะแนน	$e_i(x_i)$	$\hat{e}_i(x_i)$	$\hat{e}_i(x_i) - e_i(x_i)$	$W_i = f_i / 1084$	$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$
33	35.423308	35.335699	-0.087609	0.001845	0.000014
34	36.333454	36.448382	0.114928	0.000000	0.000000
35	37.220831	37.561065	0.340235	0.000000	0.000000
36	38.092672	38.673749	0.581077	0.001845	0.000623
37	38.950154	39.786432	0.836278	0.000000	0.000000
38	39.801108	40.899115	1.098007	0.000000	0.000000
39	40.643852	42.011798	1.367946	0.000000	0.000000
40	41.482610	43.124482	1.641872	0.000000	0.000000
41	42.317032	44.237165	1.920133	0.000923	0.003401
42	43.148400	45.349848	2.201449	0.000000	0.000000
43	43.978188	46.462532	2.484343	0.000923	0.005694
44	44.804653	47.575215	2.770562	0.000000	0.000000
45	45.630415	48.687898	3.057483	0.000000	0.000000
46	46.452612	49.800581	3.347969	0.000000	0.000000
47	47.273948	50.913265	3.639316	0.000000	0.000000
48	48.092248	52.025948	3.933700	0.000000	0.000000
รวม					2.01958
ROOT					1.421119

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 50 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบ ที่ 2 ด้วยวิธี CLE (Chained Linear Equating)

คะแนน	$e_i(x_i)$	$\hat{e}_i(x_i)$	$\hat{e}_i(x_i) - e_i(x_i)$	$W_i = f_i / 1084$	$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$
0	0.200063	-1.381612	-1.581675	0.000000	0.000000
1	1.503058	-0.278620	-1.781678	0.004613	0.014642
2	2.670360	0.824373	-1.845987	0.026753	0.091165
3	3.808869	1.927365	-1.881504	0.067343	0.238399
4	4.892185	3.030358	-1.861827	0.097786	0.338965
5	5.973298	4.133350	-1.839948	0.108856	0.368522
6	7.031998	5.236343	-1.795655	0.114391	0.368840
7	8.061973	6.339335	-1.722638	0.083026	0.246377
8	9.058244	7.442328	-1.615916	0.056273	0.146939
9	10.013506	8.545320	-1.468185	0.052583	0.113346
10	10.939139	9.648313	-1.290826	0.050738	0.084541
11	11.849230	10.751305	-1.097924	0.053506	0.064498
12	12.778435	11.854298	-0.924137	0.047970	0.040968
13	13.692005	12.957290	-0.734715	0.037823	0.020417
14	14.675540	14.060283	-0.615257	0.034133	0.012921
15	15.636421	15.163275	-0.473146	0.029520	0.006609
16	16.669097	16.266268	-0.402829	0.021218	0.003443
17	17.689501	17.369260	-0.320240	0.020295	0.002081
18	18.766604	18.472253	-0.294351	0.011993	0.001039
19	19.837862	19.575245	-0.262617	0.012915	0.000891
20	20.949317	20.678238	-0.271079	0.009225	0.000678
21	22.065847	21.781230	-0.284616	0.009225	0.000747
22	23.211694	22.884223	-0.327471	0.000923	0.000099
23	24.375050	23.987215	-0.387835	0.006458	0.000971
24	25.564593	25.090208	-0.474385	0.004613	0.001038
25	26.787601	26.193200	-0.594400	0.001845	0.000652
26	27.995020	27.296193	-0.698827	0.008303	0.004055
27	29.183145	28.399185	-0.783959	0.007380	0.004536
28	30.330229	29.502178	-0.828051	0.004613	0.003163
29	31.438102	30.605170	-0.832931	0.001845	0.001280
30	32.495797	31.708163	-0.787634	0.003690	0.002289
31	33.512015	32.811155	-0.700859	0.003690	0.001813
32	34.484269	33.914148	-0.570121	0.000923	0.000300

ตาราง 50 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบ ที่ 2 ด้วยวิธี CLE (Chained Linear Equating) (ต่อ)

คะแนน	$e_i(x_i)$	$\hat{e}_i(x_i)$	$\hat{e}_i(x_i) - e_i(x_i)$	$W_i = f_i / 1084$	$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$
33	35.423308	35.017140	-0.406168	0.001845	0.000304
34	36.333454	36.120133	-0.213321	0.000000	0.000000
35	37.220831	37.223126	0.002295	0.000000	0.000000
36	38.092672	38.326118	0.233446	0.001845	0.000101
37	38.950154	39.429111	0.478957	0.000000	0.000000
38	39.801108	40.532103	0.730995	0.000000	0.000000
39	40.643852	41.635096	0.991243	0.000000	0.000000
40	41.482610	42.738088	1.255478	0.000000	0.000000
41	42.317032	43.841081	1.524048	0.000923	0.002143
42	43.148400	44.944073	1.795674	0.000000	0.000000
43	43.978188	46.047066	2.068877	0.000923	0.003949
44	44.804653	47.150058	2.345405	0.000000	0.000000
45	45.630415	48.253051	2.622635	0.000000	0.000000
46	46.452612	49.356043	2.903431	0.000000	0.000000
47	47.273948	50.459036	3.185088	0.000000	0.000000
48	48.092248	51.562028	3.469780	0.000000	0.000000
รวม					2.192720
ROOT					1.480784

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 51 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบ ที่ 3 ด้วยวิธี CLE (Chained Linear Equating)

คะแนน	$e_i(x_i)$	$\hat{e}_i(x_i)$	$\hat{e}_i(x_i) - e_i(x_i)$	$W_i = f_i / 1084$	$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$
0	0.200063	-1.439893	-1.639956	0.000000	0.000000
1	1.503058	-0.328145	-1.831202	0.004613	0.015467
2	2.670360	0.783604	-1.886756	0.026753	0.095236
3	3.808869	1.895352	-1.913517	0.067343	0.246580
4	4.892185	3.007100	-1.885085	0.097786	0.347487
5	5.973298	4.118848	-1.854450	0.108856	0.374354
6	7.031998	5.230597	-1.801401	0.115314	0.374198
7	8.061973	6.342345	-1.719628	0.082103	0.242789
8	9.058244	7.454093	-1.604151	0.058118	0.149555
9	10.013506	8.565841	-1.447664	0.050738	0.106333
10	10.939139	9.677590	-1.261549	0.051661	0.082218
11	11.849230	10.789338	-1.059892	0.051661	0.058034
12	12.778435	11.901086	-0.877349	0.048893	0.037635
13	13.692005	13.012834	-0.679171	0.038745	0.017872
14	14.675540	14.124582	-0.550957	0.034133	0.010361
15	15.636421	15.236331	-0.400091	0.027675	0.004430
16	16.669097	16.348079	-0.321018	0.019373	0.001996
17	17.689501	17.459827	-0.229674	0.022140	0.001168
18	18.766604	18.571575	-0.195028	0.012915	0.000491
19	19.837862	19.683324	-0.154539	0.013838	0.000330
20	20.949317	20.795072	-0.154245	0.008303	0.000198
21	22.065847	21.906820	-0.159027	0.008303	0.000210
22	23.211694	23.018568	-0.193126	0.000923	0.000034
23	24.375050	24.130317	-0.244734	0.007380	0.000442
24	25.564593	25.242065	-0.322528	0.004613	0.000480
25	26.787601	26.353813	-0.433788	0.001845	0.000347
26	27.995020	27.465561	-0.529459	0.007380	0.002069
27	29.183145	28.577309	-0.605835	0.008303	0.003047
28	30.330229	29.689058	-0.641171	0.004613	0.001896
29	31.438102	30.800806	-0.637296	0.001845	0.000749
30	32.495797	31.912554	-0.583243	0.003690	0.001255
31	33.512015	33.024302	-0.487712	0.003690	0.000878
32	34.484269	34.136051	-0.348218	0.000923	0.000112

ตาราง 51 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบ ที่ 3 ด้วยวิธี CLE (Chained Linear Equating) (ต่อ)

คะแนน	$e_i(x_i)$	$\hat{e}_i(x_i)$	$\hat{e}_i(x_i) - e_i(x_i)$	$W_i = f_i / 1084$	$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$
33	35.423308	35.247799	-0.175509	0.000923	0.000028
34	36.333454	36.359547	0.026093	0.000923	0.000001
35	37.220831	37.471295	0.250465	0.000000	0.000000
36	38.092672	38.583044	0.490372	0.001845	0.000444
37	38.950154	39.694792	0.744638	0.000000	0.000000
38	39.801108	40.806540	1.005432	0.000000	0.000000
39	40.643852	41.918288	1.274436	0.000000	0.000000
40	41.482610	43.030036	1.547426	0.000000	0.000000
41	42.317032	44.141785	1.824752	0.000923	0.003072
42	43.148400	45.253533	2.105133	0.000000	0.000000
43	43.978188	46.365281	2.387093	0.000000	0.000000
44	44.804653	47.477029	2.672377	0.000000	0.000000
45	45.630415	48.588778	2.958362	0.000000	0.000000
46	46.452612	49.700526	3.247914	0.000000	0.000000
47	47.273948	50.812274	3.538326	0.000923	0.011550
48	48.092248	51.924022	3.831774	0.000000	0.000000
รวม				1.000000	2.193349
ROOT					1.480996

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 52 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธีของเลอวิน (Levine method)

คะแนน	$e_i(x_i)$	$\hat{e}_i(x_i)$	$\hat{e}_i(x_i) - e_i(x_i)$	$W_i = f_i / 1084$	$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$
0	0.200063	-1.542587	-1.742650	0.000000	0.000000
1	1.503058	-0.451603	-1.954661	0.004613	0.017623
2	2.670360	0.639381	-2.030979	0.026753	0.110352
3	3.808869	1.730364	-2.078505	0.067343	0.290935
4	4.892185	2.821348	-2.070837	0.097786	0.419342
5	5.973298	3.912332	-2.060966	0.108856	0.462375
6	7.031998	5.003316	-2.028682	0.114391	0.470783
7	8.061973	6.094300	-1.967673	0.083026	0.321454
8	9.058244	7.185283	-1.872961	0.056273	0.197405
9	10.013506	8.276267	-1.737238	0.052583	0.158695
10	10.939139	9.367251	-1.571887	0.050738	0.125365
11	11.849230	10.458235	-1.390995	0.053506	0.103526
12	12.778435	11.549219	-1.229216	0.047970	0.072482
13	13.692005	12.640202	-1.051803	0.037823	0.041843
14	14.675540	13.731186	-0.944354	0.034133	0.030440
15	15.636421	14.822170	-0.814251	0.029520	0.019572
16	16.669097	15.913154	-0.755943	0.021218	0.012125
17	17.689501	17.004138	-0.685363	0.020295	0.009533
18	18.766604	18.095121	-0.671482	0.011993	0.005407
19	19.837862	19.186105	-0.651757	0.012915	0.005486
20	20.949317	20.277089	-0.672228	0.009225	0.004169
21	22.065847	21.368073	-0.697774	0.009225	0.004492
22	23.211694	22.459057	-0.752637	0.000923	0.000523
23	24.375050	23.550040	-0.825010	0.006458	0.004395
24	25.564593	24.641024	-0.923569	0.004613	0.003934
25	26.787601	25.732008	-1.055593	0.001845	0.002056
26	27.995020	26.822992	-1.172028	0.008303	0.011405
27	29.183145	27.913976	-1.269169	0.007380	0.011888
28	30.330229	29.004959	-1.325270	0.004613	0.008101
29	31.438102	30.095943	-1.342158	0.001845	0.003324
30	32.495797	31.186927	-1.308870	0.003690	0.006322
31	33.512015	32.277911	-1.234104	0.003690	0.005620
32	34.484269	33.368895	-1.115374	0.000923	0.001148

ตาราง 52 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธีของเลอวิน (Levine method) (ต่อ)

คะแนน	$e_i(x_i)$	$\hat{e}_i(x_i)$	$\hat{e}_i(x_i) - e_i(x_i)$	$W_i = f_i / 1084$	$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$
33	35.423308	34.459878	-0.963430	0.001845	0.001713
34	36.333454	35.550862	-0.782592	0.000000	0.000000
35	37.220831	36.641846	-0.578984	0.000000	0.000000
36	38.092672	37.732830	-0.359842	0.001845	0.000239
37	38.950154	38.823814	-0.126340	0.000000	0.000000
38	39.801108	39.914797	0.113689	0.000000	0.000000
39	40.643852	41.005781	0.361929	0.000000	0.000000
40	41.482610	42.096765	0.614155	0.000000	0.000000
41	42.317032	43.187749	0.870717	0.000923	0.000699
42	43.148400	44.278733	1.130333	0.000000	0.000000
43	43.978188	45.369716	1.391528	0.000923	0.001786
44	44.804653	46.460700	1.656047	0.000000	0.000000
45	45.630415	47.551684	1.921269	0.000000	0.000000
46	46.452612	48.642668	2.190056	0.000000	0.000000
47	47.273948	49.733652	2.459704	0.000000	0.000000
48	48.092248	50.824635	2.732388	0.000000	0.000000
รวม				1.000000	2.946555
ROOT					1.716555

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 53 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธีของเลอวิน (Levine method)

คะแนน	$e_i(x_i)$	$\hat{e}_i(x_i)$	$\hat{e}_i(x_i) - e_i(x_i)$	$W_i = f_i / 1084$	$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$
0	0.200063	-1.524761	-1.724824	0.000000	0.000000
1	1.503058	-0.447451	-1.950509	0.004613	0.017548
2	2.670360	0.629858	-2.040502	0.026753	0.111389
3	3.808869	1.707168	-2.101701	0.067343	0.297465
4	4.892185	2.784477	-2.107707	0.097786	0.434407
5	5.973298	3.861787	-2.111511	0.108856	0.485332
6	7.031998	4.939097	-2.092901	0.114391	0.501060
7	8.061973	6.016406	-2.045566	0.083026	0.347408
8	9.058244	7.093716	-1.964528	0.056273	0.217179
9	10.013506	8.171026	-1.842480	0.052583	0.178505
10	10.939139	9.248335	-1.690803	0.050738	0.145051
11	11.849230	10.325645	-1.523585	0.053506	0.124203
12	12.778435	11.402955	-1.375480	0.047970	0.090758
13	13.692005	12.480264	-1.211741	0.037823	0.055536
14	14.675540	13.557574	-1.117966	0.034133	0.042661
15	15.636421	14.634883	-1.001538	0.029520	0.029611
16	16.669097	15.712193	-0.956904	0.021218	0.019428
17	17.689501	16.789503	-0.899998	0.020295	0.016439
18	18.766604	17.866812	-0.899791	0.011993	0.009710
19	19.837862	18.944122	-0.893740	0.012915	0.010316
20	20.949317	20.021432	-0.927885	0.009225	0.007943
21	22.065847	21.098741	-0.967106	0.009225	0.008628
22	23.211694	22.176051	-1.035643	0.000923	0.000989
23	24.375050	23.253360	-1.121690	0.006458	0.008125
24	25.564593	24.330670	-1.233923	0.004613	0.007023
25	26.787601	25.407980	-1.379621	0.001845	0.003512
26	27.995020	26.485289	-1.509731	0.008303	0.018924
27	29.183145	27.562599	-1.620546	0.007380	0.019381
28	30.330229	28.639909	-1.690321	0.004613	0.013179
29	31.438102	29.717218	-1.720883	0.001845	0.005464
30	32.495797	30.794528	-1.701269	0.003690	0.010680
31	33.512015	31.871838	-1.640177	0.003690	0.009927
32	34.484269	32.949147	-1.535122	0.000923	0.002174

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 53 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธีของเลอวิน (Levine method) (ต่อ)

คะแนน	$e_i(x_i)$	$\hat{e}_i(x_i)$	$\hat{e}_i(x_i) - e_i(x_i)$	$W_i = f_i / 1084$	$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$
33	35.423308	34.026457	-1.396851	0.001845	0.003600
34	36.333454	35.103766	-1.229688	0.000000	0.000000
35	37.220831	36.181076	-1.039755	0.000000	0.000000
36	38.092672	37.258386	-0.834286	0.001845	0.001284
37	38.950154	38.335695	-0.614458	0.000000	0.000000
38	39.801108	39.413005	-0.388103	0.000000	0.000000
39	40.643852	40.490315	-0.153538	0.000000	0.000000
40	41.482610	41.567624	0.085014	0.000000	0.000000
41	42.317032	42.644934	0.327901	0.000923	0.000099
42	43.148400	43.722243	0.573844	0.000000	0.000000
43	43.978188	44.799553	0.821365	0.000923	0.000622
44	44.804653	45.876863	1.072210	0.000000	0.000000
45	45.630415	46.954172	1.323757	0.000000	0.000000
46	46.452612	48.031482	1.578870	0.000000	0.000000
47	47.273948	49.108792	1.834844	0.000000	0.000000
48	48.092248	50.186101	2.093853	0.000000	0.000000
รวม				1.000000	3.255561
ROOT					1.804317

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 54 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธีของเลอวิน (Levine method)

คะแนน	$e_i(x_i)$	$\hat{e}_i(x_i)$	$\hat{e}_i(x_i) - e_i(x_i)$	$W_i = f_i / 1084$	$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$
0	0.200063	-1.614030	-1.814093	0.000000	0.000000
1	1.503058	-0.523912	-2.026970	0.004613	0.018951
2	2.670360	0.566206	-2.104154	0.026753	0.118447
3	3.808869	1.656324	-2.152546	0.067343	0.312031
4	4.892185	2.746441	-2.145743	0.097786	0.450228
5	5.973298	3.836559	-2.136739	0.108856	0.496999
6	7.031998	4.926677	-2.105321	0.115314	0.511113
7	8.061973	6.016795	-2.045178	0.082103	0.343418
8	9.058244	7.106913	-1.951331	0.058118	0.221296
9	10.013506	8.197031	-1.816475	0.050738	0.167414
10	10.939139	9.287149	-1.651990	0.051661	0.140985
11	11.849230	10.377267	-1.471963	0.051661	0.111932
12	12.778435	11.467385	-1.311050	0.048893	0.084040
13	13.692005	12.557503	-1.134502	0.038745	0.049869
14	14.675540	13.647621	-1.027919	0.034133	0.036065
15	15.636421	14.737739	-0.898683	0.027675	0.022351
16	16.669097	15.827856	-0.841241	0.019373	0.013710
17	17.689501	16.917974	-0.771526	0.022140	0.013179
18	18.766604	18.008092	-0.758511	0.012915	0.007431
19	19.837862	19.098210	-0.739652	0.013838	0.007570
20	20.949317	20.188328	-0.760989	0.008303	0.004808
21	22.065847	21.278446	-0.787401	0.008303	0.005148
22	23.211694	22.368564	-0.843130	0.000923	0.000656
23	24.375050	23.458682	-0.916368	0.007380	0.006197
24	25.564593	24.548800	-1.015793	0.004613	0.004759
25	26.787601	25.638918	-1.148683	0.001845	0.002434
26	27.995020	26.729036	-1.265985	0.007380	0.011828
27	29.183145	27.819154	-1.363991	0.008303	0.015447
28	30.330229	28.909271	-1.420958	0.004613	0.009313
29	31.438102	29.999389	-1.438712	0.001845	0.003819
30	32.495797	31.089507	-1.406289	0.003690	0.007298
31	33.512015	32.179625	-1.332390	0.003690	0.006551
32	34.484269	33.269743	-1.214526	0.000923	0.001361

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 54 การคำนวณค่ารากที่สองของค่าเฉลี่ยความแตกต่างยกกำลังสอง (RMSD) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธีของเลอวิน (Levine method) (ต่อ)

คะแนน	$e_i(x_i)$	$\hat{e}_i(x_i)$	$\hat{e}_i(x_i) - e_i(x_i)$	$W_i = f_i / 1084$	$RMSD = \sqrt{\sum_{i=o}^{48} w_i[\hat{e}_i(x_i) - e_i(x_i)]^2}$
33	35.423308	34.359861	-1.063447	0.000923	0.001043
34	36.333454	35.449979	-0.883475	0.000923	0.000720
35	37.220831	36.540097	-0.680734	0.000000	0.000000
36	38.092672	37.630215	-0.462457	0.001845	0.000395
37	38.950154	38.720333	-0.229821	0.000000	0.000000
38	39.801108	39.810451	0.009342	0.000000	0.000000
39	40.643852	40.900569	0.256716	0.000000	0.000000
40	41.482610	41.990687	0.508076	0.000000	0.000000
41	42.317032	43.080804	0.763772	0.000923	0.000538
42	43.148400	44.170922	1.022523	0.000000	0.000000
43	43.978188	45.261040	1.282852	0.000000	0.000000
44	44.804653	46.351158	1.546505	0.000000	0.000000
45	45.630415	47.441276	1.810861	0.000000	0.000000
46	46.452612	48.531394	2.078782	0.000000	0.000000
47	47.273948	49.621512	2.347564	0.000923	0.005084
48	48.092248	50.711630	2.619382	0.000000	0.000000
				1.000000	3.214428
					1.792883

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ฌ

การคำนวณค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (RMSE) ค่าความลำเอียง (bias) และค่า Equating ในการปรับเทียบคะแนน ภายใต้รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกันโดยใช้ข้อสอบร่วม เมื่อใช้วิธี CLE (Chained Linear Equating) และวิธี Levine

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 55 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธี CLE (Chained Linear Equating)

คะแนน	$Bias_i$	${Bias_i}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = $f_i/1084$	$Bias$
0	-0.892658	0.796839	0	0.000000	0.000000
1	-1.815382	3.295611	5	0.004613	0.015201
2	-0.954523	0.911114	29	0.026753	0.024375
3	-0.592726	0.351325	73	0.067343	0.023659
4	-0.537328	0.288722	106	0.097786	0.028233
5	-0.377549	0.142543	118	0.108856	0.015517
6	-0.494011	0.244047	124	0.114391	0.027917
7	-0.221938	0.049256	90	0.083026	0.004090
8	0.181223	0.032842	61	0.056273	0.001848
9	0.363572	0.132184	57	0.052583	0.006951
10	0.114354	0.013077	55	0.050738	0.000663
11	0.138141	0.019083	58	0.053506	0.001021
12	-0.002871	0.000008	52	0.047970	0.000000
13	-0.124449	0.015488	41	0.037823	0.000586
14	-0.133635	0.017858	37	0.034133	0.000610
15	0.143158	0.020494	32	0.029520	0.000605
16	0.251330	0.063167	23	0.021218	0.001340
17	0.450384	0.202846	22	0.020295	0.004117
18	0.647851	0.419711	13	0.011993	0.005033
19	1.043390	1.088662	14	0.012915	0.014060
20	1.315811	1.731358	10	0.009225	0.015972
21	0.955293	0.912584	10	0.009225	0.008419
22	1.003324	1.006660	1	0.000923	0.000929
23	1.064343	1.132826	7	0.006458	0.007315
24	1.174469	1.379377	5	0.004613	0.006362
25	-0.227905	0.051941	2	0.001845	0.000096
26	-0.613315	0.376156	9	0.008303	0.003123
27	-4.798725	23.027762	8	0.007380	0.169947
28	-4.892468	23.936246	5	0.004613	0.110407
29	-4.069545	16.561195	2	0.001845	0.030556
30	-3.204955	10.271735	4	0.003690	0.037903
31	-2.488513	6.192696	4	0.003690	0.022851
32	-1.744293	3.042558	1	0.000923	0.002807
33	-0.851925	0.725776	2	0.001845	0.001339

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 55 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธี CLE (Chained Linear Equating) (ต่อ)

คะแนน	$Bias_i$	${Bias_i}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = $f_i/1084$	$Bias$
34	0.114517	0.013114	0	0.000000	0.000000
35	1.229107	1.510704	0	0.000000	0.000000
36	1.121475	1.257706	2	0.001845	0.002320
37	1.791621	3.209904	0	0.000000	0.000000
38	2.906211	8.446061	0	0.000000	0.000000
39	4.020801	16.166840	0	0.000000	0.000000
40	5.135391	26.372241	0	0.000000	0.000000
41	6.027759	36.333877	1	0.000923	0.033518
42	6.420127	41.218028	0	0.000000	0.000000
43	6.534717	42.702525	1	0.000923	0.039393
44	-0.850693	0.723678	0	0.000000	0.000000
45	0.263897	0.069642	0	0.000000	0.000000
46	1.378487	1.900227	0	0.000000	0.000000
47	2.493077	6.215435	0	0.000000	0.000000
48	3.607668	13.015265	0	0.000000	0.000000
รวม				1.000000	0.669084
ROOT					0.817976

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 56 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธี CLE (Chained Linear Equating)

คะแนน	SEE_i	$SEE_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	SEE
0	0.3387985	0.1147844	0	0.000000	0.000000
1	0.3086472	0.0952631	5	0.004613	0.000439
2	0.2794129	0.0780716	29	0.026753	0.002089
3	0.2514157	0.0632098	73	0.067343	0.004257
4	0.2251176	0.0506779	106	0.097786	0.004956
5	0.2011861	0.0404758	118	0.108856	0.004406
6	0.1805645	0.0326035	124	0.114391	0.003730
7	0.1645025	0.0270611	90	0.083026	0.002247
8	0.1544293	0.0238484	61	0.056273	0.001342
9	0.1515439	0.0229656	57	0.052583	0.001208
10	0.1562450	0.0244125	55	0.050738	0.001239
11	0.1678967	0.0281893	58	0.053506	0.001508
12	0.1851915	0.0342959	52	0.047970	0.001645
13	0.2067179	0.0427323	41	0.037823	0.001616
14	0.2312974	0.0534985	37	0.034133	0.001826
15	0.2580591	0.0665945	32	0.029520	0.001966
16	0.2863919	0.0820203	23	0.021218	0.001740
17	0.3158734	0.0997760	22	0.020295	0.002025
18	0.3462101	0.1198615	13	0.011993	0.001437
19	0.3771959	0.1422767	14	0.012915	0.001838
20	0.4086830	0.1670218	10	0.009225	0.001541
21	0.4405641	0.1940967	10	0.009225	0.001791
22	0.4727594	0.2235014	1	0.000923	0.000206
23	0.5052088	0.2552359	7	0.006458	0.001648
24	0.5378664	0.2893003	5	0.004613	0.001334
25	0.5706964	0.3256944	2	0.001845	0.000601
26	0.6036708	0.3644184	9	0.008303	0.003026
27	0.6367670	0.4054721	8	0.007380	0.002992
28	0.6699670	0.4488557	5	0.004613	0.002070
29	0.7032561	0.4945691	2	0.001845	0.000912
30	0.7366222	0.5426123	4	0.003690	0.002002
31	0.7700554	0.5929854	4	0.003690	0.002188
32	0.8035473	0.6456882	1	0.000923	0.000596
33	0.8370907	0.7007208	2	0.001845	0.001293

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 56 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธี CLE (Chained Linear Equating) (ต่อ)

คะแนน	SEE_i	SEE_i^2	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	SEE
34	0.8706798	0.7580833	0	0.000000	0.000000
35	0.9043094	0.8177756	0	0.000000	0.000000
36	0.9379753	0.8797976	2	0.001845	0.001623
37	0.9716736	0.9441495	0	0.000000	0.000000
38	1.0054010	1.0108312	0	0.000000	0.000000
39	1.0391548	1.0798428	0	0.000000	0.000000
40	1.0729325	1.1511841	0	0.000000	0.000000
41	1.1067318	1.2248552	1	0.000923	0.001130
42	1.1405508	1.3008562	0	0.000000	0.000000
43	1.1743879	1.3791870	1	0.000923	0.001272
44	1.2082415	1.4598475	0	0.000000	0.000000
45	1.2421103	1.5428379	0	0.000000	0.000000
46	1.2759930	1.6281581	0	0.000000	0.000000
47	1.3098886	1.7158081	0	0.000000	0.000000
48	1.3437961	1.8057880	0	0.000000	0.000000
รวม				1.0000000	0.067739
ROOT					0.260267

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 57 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธี CLE (Chained Linear Equating)

คะแนน	$RMSE_i$	$RMSE_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	w = $f_i/1084$	$RMSE$
0	0.954790	0.911624	0	0.000000	0.000000
1	1.841433	3.390874	5	0.004613	0.015641
2	0.994578	0.989186	29	0.026753	0.026463
3	0.643844	0.414534	73	0.067343	0.027916
4	0.582580	0.339400	106	0.097786	0.033189
5	0.427807	0.183019	118	0.108856	0.019923
6	0.525975	0.276650	124	0.114391	0.031646
7	0.276256	0.076317	90	0.083026	0.006336
8	0.238097	0.056690	61	0.056273	0.003190
9	0.393891	0.155150	57	0.052583	0.008158
10	0.193622	0.037489	55	0.050738	0.001902
11	0.217422	0.047272	58	0.053506	0.002529
12	0.185214	0.034304	52	0.047970	0.001646
13	0.241288	0.058220	41	0.037823	0.002202
14	0.267127	0.071357	37	0.034133	0.002436
15	0.295108	0.087089	32	0.029520	0.002571
16	0.381035	0.145187	23	0.021218	0.003081
17	0.550111	0.302622	22	0.020295	0.006142
18	0.734556	0.539572	13	0.011993	0.006471
19	1.109477	1.230939	14	0.012915	0.015898
20	1.377817	1.898380	10	0.009225	0.017513
21	1.051989	1.106681	10	0.009225	0.010209
22	1.109126	1.230161	1	0.000923	0.001135
23	1.178160	1.388062	7	0.006458	0.008963
24	1.291773	1.668677	5	0.004613	0.007697
25	0.614520	0.377635	2	0.001845	0.000697
26	0.860566	0.740574	9	0.008303	0.006149
27	4.840789	23.433234	8	0.007380	0.172939
28	4.938127	24.385102	5	0.004613	0.112477
29	4.129862	17.055764	2	0.001845	0.031468
30	3.288517	10.814347	4	0.003690	0.039905
31	2.604934	6.785681	4	0.003690	0.025039
32	1.920481	3.688246	1	0.000923	0.003402
33	1.194361	1.426497	2	0.001845	0.002632

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 57 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธี CLE (Chained Linear Equating) (ต่อ)

คะแนน	$RMSE_i$	$RMSE_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	$RMSE$
34	0.878178	0.771197	0	0.000000	0.000000
35	1.525936	2.328480	0	0.000000	0.000000
36	1.462020	2.137504	2	0.001845	0.003944
37	2.038150	4.154054	0	0.000000	0.000000
38	3.075206	9.456892	0	0.000000	0.000000
39	4.152913	17.246682	0	0.000000	0.000000
40	5.246277	27.523425	0	0.000000	0.000000
41	6.128518	37.558733	1	0.000923	0.034648
42	6.520651	42.518884	0	0.000000	0.000000
43	6.639406	44.081712	1	0.000923	0.040666
44	1.477676	2.183526	0	0.000000	0.000000
45	1.269834	1.612480	0	0.000000	0.000000
46	1.878400	3.528385	0	0.000000	0.000000
47	2.816246	7.931243	0	0.000000	0.000000
48	3.849812	14.821053	0	0.000000	0.000000
รวม				1.000000	0.736823
ROOT					0.858384

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 58 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธี CLE (Chained Linear Equating)

คะแนน	$Bias_i$	${Bias_i}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	$Bias$
0	-0.874774	0.765230	0	0.000000	0.000000
1	-1.808203	3.269598	5	0.004613	0.015081
2	-0.958050	0.917859	29	0.026753	0.024555
3	-0.604130	0.364973	73	0.067343	0.024578
4	-0.517733	0.268048	106	0.097786	0.026211
5	-0.375545	0.141034	118	0.108856	0.015352
6	-0.568743	0.323469	124	0.114391	0.037002
7	-0.335615	0.112637	90	0.083026	0.009352
8	0.190572	0.036318	61	0.056273	0.002044
9	0.374429	0.140197	57	0.052583	0.007372
10	0.142546	0.020319	55	0.050738	0.001031
11	0.233032	0.054304	58	0.053506	0.002906
12	-0.028892	0.000835	52	0.047970	0.000040
13	-0.193661	0.037504	41	0.037823	0.001419
14	-0.093998	0.008836	37	0.034133	0.000302
15	0.262202	0.068750	32	0.029520	0.002030
16	0.435820	0.189939	23	0.021218	0.004030
17	0.604781	0.365760	22	0.020295	0.007423
18	0.579237	0.335515	13	0.011993	0.004024
19	0.890028	0.792150	14	0.012915	0.010231
20	1.269589	1.611856	10	0.009225	0.014870
21	0.709404	0.503255	10	0.009225	0.004643
22	0.885692	0.784450	1	0.000923	0.000724
23	1.478862	2.187034	7	0.006458	0.014123
24	0.739890	0.547437	5	0.004613	0.002525
25	0.265203	0.070333	2	0.001845	0.000130
26	-0.540435	0.292071	9	0.008303	0.002425
27	-3.569884	12.744072	8	0.007380	0.094052
28	-5.174332	26.773717	5	0.004613	0.123495
29	-4.362114	19.028041	2	0.001845	0.035107
30	-3.508229	12.307674	4	0.003690	0.045416
31	-2.802493	7.853966	4	0.003690	0.028981
32	-2.068978	4.280671	1	0.000923	0.003949
33	-1.187316	1.409719	2	0.001845	0.002601

ตาราง 58 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธี CLE (Chained Linear Equating) (ต่อ)

คะแนน	$Bias_i$	$Bias_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = $f_i/1084$	$Bias$
34	-0.231579	0.053629	0	0.000000	0.000000
35	0.872306	0.760917	0	0.000000	0.000000
36	0.753968	0.568468	2	0.001845	0.001049
37	1.413409	1.997724	0	0.000000	0.000000
38	2.517294	6.336767	0	0.000000	0.000000
39	3.621178	13.112933	0	0.000000	0.000000
40	4.725063	22.326223	0	0.000000	0.000000
41	5.606726	31.435375	1	0.000923	0.028999
42	6.488389	42.099186	0	0.000000	0.000000
43	7.092273	50.300341	1	0.000923	0.046403
44	-1.303842	1.700003	0	0.000000	0.000000
45	-0.199957	0.039983	0	0.000000	0.000000
46	0.903928	0.817086	0	0.000000	0.000000
47	2.007813	4.031312	0	0.000000	0.000000
48	3.111698	9.682661	0	0.000000	0.000000
รวม				1.0000000	0.644473
ROOT					0.802791

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 59 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธี CLE (Chained Linear Equating)

คะแนน	SEE_i	$SEE_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	SEE
0	0.336454	0.113201	0	0.000000	0.000000
1	0.306001	0.093637	5	0.004613	0.000432
2	0.276479	0.076441	29	0.026753	0.002045
3	0.248220	0.061613	73	0.067343	0.004149
4	0.221707	0.049154	106	0.097786	0.004807
5	0.197645	0.039064	118	0.108856	0.004252
6	0.177037	0.031342	124	0.114391	0.003585
7	0.161210	0.025989	90	0.083026	0.002158
8	0.151671	0.023004	61	0.056273	0.001295
9	0.149626	0.022388	57	0.052583	0.001177
10	0.155372	0.024141	55	0.050738	0.001225
11	0.168112	0.028262	58	0.053506	0.001512
12	0.186417	0.034751	52	0.047970	0.001667
13	0.208829	0.043609	41	0.037823	0.001649
14	0.234171	0.054836	37	0.034133	0.001872
15	0.261594	0.068432	32	0.029520	0.002020
16	0.290509	0.084396	23	0.021218	0.001791
17	0.320512	0.102728	22	0.020295	0.002085
18	0.351325	0.123429	13	0.011993	0.001480
19	0.382752	0.146499	14	0.012915	0.001892
20	0.414653	0.171937	10	0.009225	0.001586
21	0.446927	0.199744	10	0.009225	0.001843
22	0.479499	0.229919	1	0.000923	0.000212
23	0.512311	0.262463	7	0.006458	0.001695
24	0.545322	0.297376	5	0.004613	0.001372
25	0.578495	0.334657	2	0.001845	0.000617
26	0.611806	0.374306	9	0.008303	0.003108
27	0.645232	0.416325	8	0.007380	0.003073
28	0.678757	0.460711	5	0.004613	0.002125
29	0.712367	0.507467	2	0.001845	0.000936
30	0.746050	0.556591	4	0.003690	0.002054
31	0.779797	0.608083	4	0.003690	0.002244
32	0.813600	0.661944	1	0.000923	0.000611
33	0.847451	0.718174	2	0.001845	0.001325

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 59 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธี CLE (Chained Linear Equating) (ต่อ)

คะแนน	SEE_i	SEE_i^2	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	SEE
34	0.881347	0.776772	0	0.000000	0.000000
35	0.915281	0.837739	0	0.000000	0.000000
36	0.949249	0.901074	2	0.001845	0.001662
37	0.983249	0.966778	0	0.000000	0.000000
38	1.017276	1.034851	0	0.000000	0.000000
39	1.051329	1.105292	0	0.000000	0.000000
40	1.085404	1.178101	0	0.000000	0.000000
41	1.119500	1.253280	1	0.000923	0.001156
42	1.153615	1.330826	0	0.000000	0.000000
43	1.187747	1.410742	1	0.000923	0.001301
44	1.221894	1.493026	0	0.000000	0.000000
45	1.256057	1.577678	0	0.000000	0.000000
46	1.290232	1.664699	0	0.000000	0.000000
47	1.324420	1.754089	0	0.000000	0.000000
48	1.358620	1.845847	0	0.000000	0.000000
รวม				1.000000	0.068013
ROOT					0.260793

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 60 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธี CLE (Chained Linear Equating)

คะแนน	$RMSE_i$	${RMSE_i}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	w = f_i/ 1084	$RMSE$
0	0.937247	0.878432	0	0.000000	0.000000
1	1.833912	3.363235	5	0.004613	0.015513
2	0.997146	0.994300	29	0.026753	0.026600
3	0.653136	0.426586	73	0.067343	0.028728
4	0.563207	0.317202	106	0.097786	0.031018
5	0.424379	0.180098	118	0.108856	0.019605
6	0.595660	0.354811	124	0.114391	0.040587
7	0.372325	0.138626	90	0.083026	0.011510
8	0.243561	0.059322	61	0.056273	0.003338
9	0.403218	0.162585	57	0.052583	0.008549
10	0.210855	0.044460	55	0.050738	0.002256
11	0.287342	0.082566	58	0.053506	0.004418
12	0.188642	0.035586	52	0.047970	0.001707
13	0.284805	0.081114	41	0.037823	0.003068
14	0.252333	0.063672	37	0.034133	0.002173
15	0.370380	0.137181	32	0.029520	0.004050
16	0.523770	0.274335	23	0.021218	0.005821
17	0.684462	0.468488	22	0.020295	0.009508
18	0.677454	0.458944	13	0.011993	0.005504
19	0.968839	0.938649	14	0.012915	0.012123
20	1.335587	1.783793	10	0.009225	0.016456
21	0.838450	0.702998	10	0.009225	0.006485
22	1.007159	1.014369	1	0.000923	0.000936
23	1.565087	2.449497	7	0.006458	0.015818
24	0.919137	0.844813	5	0.004613	0.003897
25	0.636388	0.404990	2	0.001845	0.000747
26	0.816319	0.666377	9	0.008303	0.005533
27	3.627726	13.160396	8	0.007380	0.097125
28	5.218662	27.234428	5	0.004613	0.125620
29	4.419899	19.535508	2	0.001845	0.036043
30	3.586679	12.864265	4	0.003690	0.047470
31	2.908960	8.462049	4	0.003690	0.031225
32	2.223199	4.942616	1	0.000923	0.004560
33	1.458730	2.127892	2	0.001845	0.003926

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 60 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธี CLE (Chained Linear Equating) (ต่อ)

คะแนน	$RMSE_i$	${RMSE_i}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = $f_i/1084$	$RMSE$
34	0.911263	0.830401	0	0.000000	0.000000
35	1.264380	1.598656	0	0.000000	0.000000
36	1.212247	1.469543	2	0.001845	0.002711
37	1.721773	2.964503	0	0.000000	0.000000
38	2.715072	7.371618	0	0.000000	0.000000
39	3.770706	14.218225	0	0.000000	0.000000
40	4.848126	23.504325	0	0.000000	0.000000
41	5.717399	32.688655	1	0.000923	0.030156
42	6.590145	43.430012	0	0.000000	0.000000
43	7.191042	51.711083	1	0.000923	0.047704
44	1.786905	3.193029	0	0.000000	0.000000
45	1.271873	1.617661	0	0.000000	0.000000
46	1.575368	2.481785	0	0.000000	0.000000
47	2.405286	5.785401	0	0.000000	0.000000
48	3.395366	11.528509	0	0.000000	0.000000
รวม				1.000000	0.712485
ROOT					0.844088

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 61 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธี CLE (Chained Linear Equating)

คะแนน	$Bias_i$	${Bias_i}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	$Bias$
0	-0.927725	0.860674	0	0.000000	0.000000
1	-1.854666	3.439785	5	0.004613	0.015866
2	-0.998024	0.996052	29	0.026753	0.026647
3	-0.639030	0.408360	73	0.067343	0.027500
4	-0.573010	0.328340	106	0.097786	0.032107
5	-0.472149	0.222925	118	0.108856	0.024267
6	-0.532308	0.283352	125	0.115314	0.032674
7	-0.268313	0.071992	89	0.082103	0.005911
8	0.165413	0.027362	63	0.058118	0.001590
9	0.385165	0.148352	55	0.050738	0.007527
10	0.200312	0.040125	56	0.051661	0.002073
11	0.331141	0.109654	56	0.051661	0.005665
12	0.120902	0.014617	53	0.048893	0.000715
13	-0.171887	0.029545	42	0.038745	0.001145
14	-0.188581	0.035563	37	0.034133	0.001214
15	-0.019152	0.000367	30	0.027675	0.000010
16	0.211642	0.044792	21	0.019373	0.000868
17	0.477189	0.227710	24	0.022140	0.005042
18	0.500550	0.250550	14	0.012915	0.003236
19	0.857936	0.736054	15	0.013838	0.010185
20	1.007302	1.014656	9	0.008303	0.008424
21	0.646521	0.417989	9	0.008303	0.003470
22	0.967150	0.935380	1	0.000923	0.000863
23	1.099746	1.209441	8	0.007380	0.008926
24	0.565675	0.319988	5	0.004613	0.001476
25	-0.035063	0.001229	2	0.001845	0.000002
26	-0.002468	0.000006	8	0.007380	0.000000
27	-4.997651	24.976511	9	0.008303	0.207370
28	-4.545611	20.662577	5	0.004613	0.095307
29	-3.726904	13.889815	2	0.001845	0.025627
30	-3.116531	9.712766	4	0.003690	0.035840
31	-2.672825	7.143991	4	0.003690	0.026362
32	-1.979118	3.916909	1	0.000923	0.003613
33	-1.035412	1.072077	1	0.000923	0.000989

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 61 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธี CLE (Chained Linear Equating) (ต่อ)

คะแนน	$Bias_i$	$Bias_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = $f_i/1084$	$Bias$
34	-0.091705	0.008410	1	0.000923	0.000008
35	-0.064665	0.004182	0	0.000000	0.000000
36	0.545708	0.297797	2	0.001845	0.000549
37	0.656081	0.430442	0	0.000000	0.000000
38	1.766454	3.120360	0	0.000000	0.000000
39	2.876827	8.276135	0	0.000000	0.000000
40	3.987200	15.897767	0	0.000000	0.000000
41	4.097574	16.790109	1	0.000923	0.015489
42	-3.292053	10.837615	0	0.000000	0.000000
43	-2.181680	4.759729	0	0.000000	0.000000
44	-1.071307	1.147699	0	0.000000	0.000000
45	0.039066	0.001526	0	0.000000	0.000000
46	1.149439	1.321211	0	0.000000	0.000000
47	2.259812	5.106752	1	0.000923	0.004711
48	3.370186	11.358150	0	0.000000	0.000000
รวม				1.000000	0.643268
ROOT					0.802040

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 62 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธี CLE (Chained Linear Equating)

คะแนน	SEE_i	$SEE_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	SEE
0	0.323626	0.104734	0	0.000000	0.000000
1	0.294144	0.086521	5	0.004613	0.000399
2	0.265533	0.070508	29	0.026753	0.001886
3	0.238107	0.056695	73	0.067343	0.003818
4	0.212325	0.045082	106	0.097786	0.004408
5	0.188863	0.035669	118	0.108856	0.003883
6	0.168691	0.028456	125	0.115314	0.003281
7	0.153114	0.023444	89	0.082103	0.001925
8	0.143635	0.020631	63	0.058118	0.001199
9	0.141487	0.020018	55	0.050738	0.001016
10	0.146989	0.021606	56	0.051661	0.001116
11	0.159353	0.025393	56	0.051661	0.001312
12	0.177146	0.031381	53	0.048893	0.001534
13	0.198918	0.039568	42	0.038745	0.001533
14	0.223508	0.049956	37	0.034133	0.001705
15	0.250087	0.062543	30	0.027675	0.001731
16	0.278085	0.077331	21	0.019373	0.001498
17	0.307113	0.094319	24	0.022140	0.002088
18	0.336907	0.113506	14	0.012915	0.001466
19	0.367279	0.134894	15	0.013838	0.001867
20	0.398098	0.158482	9	0.008303	0.001316
21	0.429266	0.184270	9	0.008303	0.001530
22	0.460714	0.212257	1	0.000923	0.000196
23	0.492387	0.242445	8	0.007380	0.001789
24	0.524245	0.274833	5	0.004613	0.001268
25	0.556256	0.309421	2	0.001845	0.000571
26	0.588395	0.346209	8	0.007380	0.002555
27	0.620643	0.385197	9	0.008303	0.003198
28	0.652982	0.426385	5	0.004613	0.001967
29	0.685400	0.469773	2	0.001845	0.000867
30	0.717887	0.515361	4	0.003690	0.001902
31	0.750433	0.563150	4	0.003690	0.002078
32	0.783031	0.613138	1	0.000923	0.000566
33	0.815675	0.665326	1	0.000923	0.000614

ตาราง 62 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธี CLE (Chained Linear Equating) (ต่อ)

คะแนน	SEE_i	SEE_i^2	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	SEE
34	0.848360	0.719714	1	0.000923	0.000664
35	0.881080	0.776303	0	0.000000	0.000000
36	0.913833	0.835091	2	0.001845	0.001541
37	0.946615	0.896079	0	0.000000	0.000000
38	0.979422	0.959268	0	0.000000	0.000000
39	1.012253	1.024656	0	0.000000	0.000000
40	1.045105	1.092245	0	0.000000	0.000000
41	1.077976	1.162033	1	0.000923	0.001072
42	1.110865	1.234022	0	0.000000	0.000000
43	1.143770	1.308210	0	0.000000	0.000000
44	1.176690	1.384599	0	0.000000	0.000000
45	1.209623	1.463188	0	0.000000	0.000000
46	1.242569	1.543977	0	0.000000	0.000000
47	1.275525	1.626965	1	0.000923	0.001501
48	1.308493	1.712154	0	0.000000	0.000000
รวม				1.000000	0.062859
ROOT					0.250717

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 63 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธี CLE (Chained Linear Equating)

คะแนน	$RMSE_i$	$RMSE_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	$RMSE$
0	0.982552	0.965408	0	0.000000	0.000000
1	1.877846	3.526305	5	0.004613	0.016265
2	1.032744	1.066559	29	0.026753	0.028533
3	0.681949	0.465054	73	0.067343	0.031318
4	0.611083	0.373422	106	0.097786	0.036515
5	0.508522	0.258594	118	0.108856	0.028150
6	0.558398	0.311809	125	0.115314	0.035956
7	0.308926	0.095435	89	0.082103	0.007836
8	0.219072	0.047993	63	0.058118	0.002789
9	0.410329	0.168370	55	0.050738	0.008543
10	0.248456	0.061731	56	0.051661	0.003189
11	0.367488	0.135048	56	0.051661	0.006977
12	0.214472	0.045998	53	0.048893	0.002249
13	0.262894	0.069113	42	0.038745	0.002678
14	0.292436	0.085519	37	0.034133	0.002919
15	0.250819	0.062910	30	0.027675	0.001741
16	0.349462	0.122124	21	0.019373	0.002366
17	0.567476	0.322028	24	0.022140	0.007130
18	0.603371	0.364056	14	0.012915	0.004702
19	0.933246	0.870948	15	0.013838	0.012052
20	1.083115	1.173138	9	0.008303	0.009740
21	0.776053	0.602259	9	0.008303	0.005000
22	1.071278	1.147637	1	0.000923	0.001059
23	1.204942	1.451886	8	0.007380	0.010715
24	0.771246	0.594821	5	0.004613	0.002744
25	0.557360	0.310651	2	0.001845	0.000573
26	0.588401	0.346215	8	0.007380	0.002555
27	5.036041	25.361708	9	0.008303	0.210568
28	4.592272	21.088962	5	0.004613	0.097274
29	3.789405	14.359588	2	0.001845	0.026494
30	3.198144	10.228127	4	0.003690	0.037742
31	2.776174	7.707141	4	0.003690	0.028440
32	2.128391	4.530046	1	0.000923	0.004179
33	1.318106	1.737403	1	0.000923	0.001603

ตาราง 63 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธี CLE (Chained Linear Equating) (ต่อ)

คะแนน	$RMSE_i$	${RMSE_i}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	$RMSE$
34	0.853302	0.728124	1	0.000923	0.000672
35	0.883450	0.780484	0	0.000000	0.000000
36	1.064372	1.132888	2	0.001845	0.002090
37	1.151747	1.326522	0	0.000000	0.000000
38	2.019809	4.079628	0	0.000000	0.000000
39	3.049720	9.300791	0	0.000000	0.000000
40	4.121894	16.990012	0	0.000000	0.000000
41	4.236997	17.952142	1	0.000923	0.016561
42	3.474426	12.071637	0	0.000000	0.000000
43	2.463319	6.067939	0	0.000000	0.000000
44	1.591320	2.532298	0	0.000000	0.000000
45	1.210254	1.464714	0	0.000000	0.000000
46	1.692686	2.865187	0	0.000000	0.000000
47	2.594941	6.733717	1	0.000923	0.006212
48	3.615288	13.070304	0	0.000000	0.000000
รวม				1.000000	0.706127
ROOT					0.840314

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 64 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธีของเลอวิน (Levine method)

คะแนน	$Bias_i$	${Bias_i}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	$Bias$
0	-1.054065	1.111053	0	0.000000	0.000000
1	-1.998274	3.993098	5	0.004613	0.018418
2	-1.158901	1.343051	29	0.026753	0.035930
3	-0.818590	0.670089	73	0.067343	0.045126
4	-0.784677	0.615718	106	0.097786	0.060209
5	-0.646383	0.417811	118	0.108856	0.045481
6	-0.784331	0.615175	124	0.114391	0.070371
7	-0.533744	0.284882	90	0.083026	0.023653
8	-0.152069	0.023125	61	0.056273	0.001301
9	0.008795	0.000077	57	0.052583	0.000004
10	-0.261909	0.068596	55	0.050738	0.003480
11	-0.259607	0.067396	58	0.053506	0.003606
12	-0.422105	0.178173	52	0.047970	0.008547
13	-0.565169	0.319416	41	0.037823	0.012081
14	-0.595840	0.355025	37	0.034133	0.012118
15	-0.340533	0.115963	32	0.029520	0.003423
16	-0.253846	0.064438	23	0.021218	0.001367
17	-0.076278	0.005818	22	0.020295	0.000118
18	0.099703	0.009941	13	0.011993	0.000119
19	0.473756	0.224445	14	0.012915	0.002899
20	0.724692	0.525178	10	0.009225	0.004845
21	0.342688	0.117435	10	0.009225	0.001083
22	0.369234	0.136334	1	0.000923	0.000126
23	0.408767	0.167091	7	0.006458	0.001079
24	0.497407	0.247414	5	0.004613	0.001141
25	-0.926452	0.858314	2	0.001845	0.001584
26	-1.333348	1.777817	9	0.008303	0.014760
27	-5.540243	30.694298	8	0.007380	0.226526
28	-5.655472	31.984367	5	0.004613	0.147529
29	-4.854034	23.561651	2	0.001845	0.043472
30	-4.010930	16.087559	4	0.003690	0.059364
31	-3.315974	10.995681	4	0.003690	0.040574
32	-2.593240	6.724891	1	0.000923	0.006204
33	-1.722357	2.966514	2	0.001845	0.005473

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 64 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธีของเลอวิน (Levine method) (ต่อ)

คะแนน	$Bias_i$	${Bias_i}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = $f_i/1084$	$Bias$
34	-0.777401	0.604352	0	0.000000	0.000000
35	0.315704	0.099669	0	0.000000	0.000000
36	0.186586	0.034814	2	0.001845	0.000064
37	0.835246	0.697636	0	0.000000	0.000000
38	1.928350	3.718535	0	0.000000	0.000000
39	3.021455	9.129190	0	0.000000	0.000000
40	4.114559	16.929599	0	0.000000	0.000000
41	4.985442	24.854629	1	0.000923	0.022929
42	5.356324	28.690207	0	0.000000	0.000000
43	5.449428	29.696271	1	0.000923	0.027395
44	-1.957467	3.831677	0	0.000000	0.000000
45	-0.864363	0.747123	0	0.000000	0.000000
46	0.228742	0.052323	0	0.000000	0.000000
47	1.321846	1.747278	0	0.000000	0.000000
48	2.414951	5.831988	0	0.000000	0.000000
รวม				1.000000	0.952401
ROOT					0.975910

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ตาราง 65 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธีของเลอวิน (Levine method)

คะแนน	SEE_i	$SEE_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	SEE
0	0.357628	0.127898	0	0.000000	0.000000
1	0.327030	0.106948	5	0.004613	0.000493
2	0.297422	0.088460	29	0.026753	0.002367
3	0.269131	0.072431	73	0.067343	0.004878
4	0.242618	0.058864	106	0.097786	0.005756
5	0.218532	0.047756	118	0.108856	0.005199
6	0.197762	0.039110	124	0.114391	0.004474
7	0.181449	0.032924	90	0.083026	0.002734
8	0.170875	0.029198	61	0.056273	0.001643
9	0.167132	0.027933	57	0.052583	0.001469
10	0.170672	0.029129	55	0.050738	0.001478
11	0.181066	0.032785	58	0.053506	0.001754
12	0.197235	0.038902	52	0.047970	0.001866
13	0.217896	0.047479	41	0.037823	0.001796
14	0.241902	0.058516	37	0.034133	0.001997
15	0.268356	0.072015	32	0.029520	0.002126
16	0.296603	0.087974	23	0.021218	0.001867
17	0.326179	0.106393	22	0.020295	0.002159
18	0.356753	0.127273	13	0.011993	0.001526
19	0.388089	0.150613	14	0.012915	0.001945
20	0.420017	0.176414	10	0.009225	0.001627
21	0.452411	0.204676	10	0.009225	0.001888
22	0.485178	0.235398	1	0.000923	0.000217
23	0.518248	0.268581	7	0.006458	0.001734
24	0.551565	0.304224	5	0.004613	0.001403
25	0.585088	0.342328	2	0.001845	0.000632
26	0.618783	0.382892	9	0.008303	0.003179
27	0.652623	0.425917	8	0.007380	0.003143
28	0.686587	0.471402	5	0.004613	0.002174
29	0.720658	0.519348	2	0.001845	0.000958
30	0.754821	0.569755	4	0.003690	0.002102
31	0.789064	0.622622	4	0.003690	0.002297
32	0.823377	0.677949	1	0.000923	0.000625
33	0.857751	0.735737	2	0.001845	0.001357

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 65 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธีของเลอวิน (Levine method) (ต่อ)

คะแนน	SEE_i	SEE_i^2	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	SEE
34	0.892180	0.795986	0	0.000000	0.000000
35	0.926658	0.858695	0	0.000000	0.000000
36	0.961179	0.923865	2	0.001845	0.001705
37	0.995738	0.991495	0	0.000000	0.000000
38	1.030333	1.061586	0	0.000000	0.000000
39	1.064959	1.134137	0	0.000000	0.000000
40	1.099613	1.209149	0	0.000000	0.000000
41	1.134293	1.286622	1	0.000923	0.001187
42	1.168997	1.366555	0	0.000000	0.000000
43	1.203723	1.448948	1	0.000923	0.001337
44	1.238468	1.533802	0	0.000000	0.000000
45	1.273231	1.621117	0	0.000000	0.000000
46	1.308011	1.710892	0	0.000000	0.000000
47	1.342806	1.803128	0	0.000000	0.000000
48	1.377615	1.897824	0	0.000000	0.000000
รวม				1.000000	0.075094
ROOT					0.274033

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ตาราง 66 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธีของเลอวิน (Levine method)

คะแนน	$RMSE_i$	${RMSE_i}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	w = f_i/ 1084	$RMSE$
0	1.113082	1.238951	0	0.000000	0.000000
1	2.024857	4.100047	5	0.004613	0.018912
2	1.196457	1.431510	29	0.026753	0.038297
3	0.861696	0.742521	73	0.067343	0.050004
4	0.821329	0.674582	106	0.097786	0.065965
5	0.682325	0.465568	118	0.108856	0.050680
6	0.808879	0.654285	124	0.114391	0.074844
7	0.563743	0.317806	90	0.083026	0.026386
8	0.228742	0.052323	61	0.056273	0.002944
9	0.167364	0.028011	57	0.052583	0.001473
10	0.312610	0.097725	55	0.050738	0.004958
11	0.316514	0.100181	58	0.053506	0.005360
12	0.465912	0.217074	52	0.047970	0.010413
13	0.605718	0.366895	41	0.037823	0.013877
14	0.643072	0.413542	37	0.034133	0.014115
15	0.433564	0.187978	32	0.029520	0.005549
16	0.390399	0.152411	23	0.021218	0.003234
17	0.334980	0.112211	22	0.020295	0.002277
18	0.370424	0.137214	13	0.011993	0.001646
19	0.612420	0.375058	14	0.012915	0.004844
20	0.837611	0.701592	10	0.009225	0.006472
21	0.567548	0.322111	10	0.009225	0.002972
22	0.609698	0.371732	1	0.000923	0.000343
23	0.660054	0.435671	7	0.006458	0.002813
24	0.742723	0.551638	5	0.004613	0.002544
25	1.095738	1.200642	2	0.001845	0.002215
26	1.469935	2.160709	9	0.008303	0.017939
27	5.578550	31.120215	8	0.007380	0.229669
28	5.696997	32.455769	5	0.004613	0.149704
29	4.907239	24.080999	2	0.001845	0.044430
30	4.081337	16.657314	4	0.003690	0.061466
31	3.408563	11.618303	4	0.003690	0.042872
32	2.720816	7.402840	1	0.000923	0.006829
33	1.924124	3.702252	2	0.001845	0.006831

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 67 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 1 ด้วยวิธีของเลอวิน (Levine method) (ต่อ)

คะแนน	$RMSE_i$	${RMSE_i}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	$RMSE$
34	1.183359	1.400338	0	0.000000	0.000000
35	0.978961	0.958364	0	0.000000	0.000000
36	0.979122	0.958679	2	0.001845	0.001769
37	1.299666	1.689131	0	0.000000	0.000000
38	2.186349	4.780121	0	0.000000	0.000000
39	3.203643	10.263327	0	0.000000	0.000000
40	4.258961	18.138748	0	0.000000	0.000000
41	5.112852	26.141251	1	0.000923	0.024116
42	5.482405	30.056761	0	0.000000	0.000000
43	5.580790	31.145219	1	0.000923	0.028732
44	2.316350	5.365479	0	0.000000	0.000000
45	1.538909	2.368240	0	0.000000	0.000000
46	1.327861	1.763215	0	0.000000	0.000000
47	1.884252	3.550406	0	0.000000	0.000000
48	2.780254	7.729812	0	0.000000	0.000000
รวม				1.000000	1.027495
ROOT					1.013654

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 67 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธีของเลอวิน (Levine method)

คะแนน	$Bias_i$	${Bias_i}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	$Bias$
0	-1.023922	1.048415	0	0.000000	0.000000
1	-1.982580	3.930625	5	0.004613	0.018130
2	-1.157657	1.340170	29	0.026753	0.035853
3	-0.828968	0.687188	73	0.067343	0.046277
4	-0.767802	0.589519	106	0.097786	0.057647
5	-0.650843	0.423597	118	0.108856	0.046111
6	-0.869272	0.755634	124	0.114391	0.086438
7	-0.661374	0.437416	90	0.083026	0.036317
8	-0.160417	0.025734	61	0.056273	0.001448
9	-0.001791	0.000003	57	0.052583	0.000000
10	-0.258904	0.067031	55	0.050738	0.003401
11	-0.193648	0.037500	58	0.053506	0.002006
12	-0.480803	0.231171	52	0.047970	0.011089
13	-0.670802	0.449975	41	0.037823	0.017019
14	-0.596370	0.355657	37	0.034133	0.012140
15	-0.265400	0.070437	32	0.029520	0.002079
16	-0.117012	0.013692	23	0.021218	0.000291
17	0.026719	0.000714	22	0.020295	0.000014
18	-0.024056	0.000579	13	0.011993	0.000007
19	0.261505	0.068385	14	0.012915	0.000883
20	0.615835	0.379253	10	0.009225	0.003499
21	0.030421	0.000925	10	0.009225	0.000009
22	0.181478	0.032934	1	0.000923	0.000030
23	0.749418	0.561627	7	0.006458	0.003627
24	-0.014785	0.000219	5	0.004613	0.000001
25	-0.514702	0.264918	2	0.001845	0.000489
26	-1.345571	1.810561	9	0.008303	0.015032
27	-4.400250	19.362198	8	0.007380	0.142894
28	-6.029929	36.360039	5	0.004613	0.167712
29	-5.242941	27.488427	2	0.001845	0.050717
30	-4.414286	19.485923	4	0.003690	0.071904
31	-3.733780	13.941112	4	0.003690	0.051443
32	-3.025496	9.153624	1	0.000923	0.008444
33	-2.169063	4.704836	2	0.001845	0.008681

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 67 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธีของเลอวิน (Levine method) (ต่อ)

คะแนน	$Bias_i$	$Bias_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	$Bias$
34	-1.238557	1.534024	0	0.000000	0.000000
35	-0.159903	0.025569	0	0.000000	0.000000
36	-0.303470	0.092094	2	0.001845	0.000170
37	0.330740	0.109389	0	0.000000	0.000000
38	1.409394	1.986392	0	0.000000	0.000000
39	2.488049	6.190387	0	0.000000	0.000000
40	3.566703	12.721373	0	0.000000	0.000000
41	4.423136	19.564129	1	0.000923	0.018048
42	5.279568	27.873837	0	0.000000	0.000000
43	5.858222	34.318770	1	0.000923	0.031659
44	-2.563123	6.569600	0	0.000000	0.000000
45	-1.484469	2.203647	0	0.000000	0.000000
46	-0.405814	0.164685	0	0.000000	0.000000
47	0.672840	0.452714	0	0.000000	0.000000
48	1.751495	3.067735	0	0.000000	0.000000
รวม				1.000000	0.951511
ROOT					0.975453

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ตาราง 68 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธีของเลอวิน (Levine method)

คะแนน	SEE_i	$SEE_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	SEE
0	0.370157	0.137016	0	0.000000	0.000000
1	0.338592	0.114644	5	0.004613	0.000529
2	0.308024	0.094879	29	0.026753	0.002538
3	0.278781	0.077719	73	0.067343	0.005234
4	0.251328	0.063166	106	0.097786	0.006177
5	0.226314	0.051218	118	0.108856	0.005575
6	0.204639	0.041877	124	0.114391	0.004790
7	0.187462	0.035142	90	0.083026	0.002918
8	0.176105	0.031013	61	0.056273	0.001745
9	0.171727	0.029490	57	0.052583	0.001551
10	0.174853	0.030574	55	0.050738	0.001551
11	0.185103	0.034263	58	0.053506	0.001833
12	0.201392	0.040559	52	0.047970	0.001946
13	0.222397	0.049460	41	0.037823	0.001871
14	0.246918	0.060968	37	0.034133	0.002081
15	0.274012	0.075082	32	0.029520	0.002216
16	0.302989	0.091803	23	0.021218	0.001948
17	0.333360	0.111129	22	0.020295	0.002255
18	0.364776	0.133061	13	0.011993	0.001596
19	0.396989	0.157600	14	0.012915	0.002035
20	0.429820	0.184745	10	0.009225	0.001704
21	0.463137	0.214496	10	0.009225	0.001979
22	0.496843	0.246853	1	0.000923	0.000228
23	0.530863	0.281816	7	0.006458	0.001820
24	0.565142	0.319385	5	0.004613	0.001473
25	0.599634	0.359561	2	0.001845	0.000663
26	0.634305	0.402342	9	0.008303	0.003340
27	0.669126	0.447730	8	0.007380	0.003304
28	0.704077	0.495724	5	0.004613	0.002287
29	0.739137	0.546324	2	0.001845	0.001008
30	0.774293	0.599530	4	0.003690	0.002212
31	0.809532	0.655342	4	0.003690	0.002418
32	0.844844	0.713761	1	0.000923	0.000658
33	0.880219	0.774785	2	0.001845	0.001429

ตาราง 68 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธีของเลอวิน (Levine method) (ต่อ)

คะแนน	SEE_i	SEE_i^2	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	SEE
34	0.915651	0.838416	0	0.000000	0.000000
35	0.951133	0.904653	0	0.000000	0.000000
36	0.986659	0.973496	2	0.001845	0.001796
37	1.022226	1.044945	0	0.000000	0.000000
38	1.057828	1.119000	0	0.000000	0.000000
39	1.093463	1.195662	0	0.000000	0.000000
40	1.129128	1.274929	0	0.000000	0.000000
41	1.164819	1.356803	1	0.000923	0.001252
42	1.200534	1.441283	0	0.000000	0.000000
43	1.236272	1.528369	1	0.000923	0.001410
44	1.272030	1.618061	0	0.000000	0.000000
45	1.307807	1.710359	0	0.000000	0.000000
46	1.343601	1.805264	0	0.000000	0.000000
47	1.379411	1.902774	0	0.000000	0.000000
48	1.415235	2.002891	0	0.000000	0.000000
รวม				1.000000	0.079372
ROOT					0.281730

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ตาราง 69 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธีของเลอวิน (Levine method)

คะแนน	$RMSE_i$	$RMSE_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	w = f_i/ 1084	$RMSE$
0	1.088775	1.185432	0	0.000000	0.000000
1	2.011286	4.045270	5	0.004613	0.018659
2	1.197935	1.435049	29	0.026753	0.038392
3	0.874590	0.764907	73	0.067343	0.051511
4	0.807889	0.652685	106	0.097786	0.063823
5	0.689069	0.474815	118	0.108856	0.051687
6	0.893035	0.797511	124	0.114391	0.091228
7	0.687428	0.472558	90	0.083026	0.039234
8	0.238216	0.056747	61	0.056273	0.003193
9	0.171736	0.029493	57	0.052583	0.001551
10	0.312418	0.097605	55	0.050738	0.004952
11	0.267886	0.071763	58	0.053506	0.003840
12	0.521277	0.271730	52	0.047970	0.013035
13	0.706708	0.499436	41	0.037823	0.018890
14	0.645465	0.416625	37	0.034133	0.014221
15	0.381470	0.145520	32	0.029520	0.004296
16	0.324799	0.105494	23	0.021218	0.002238
17	0.334429	0.111843	22	0.020295	0.002270
18	0.365568	0.133640	13	0.011993	0.001603
19	0.475379	0.225985	14	0.012915	0.002919
20	0.750998	0.563998	10	0.009225	0.005203
21	0.464135	0.215421	10	0.009225	0.001987
22	0.528949	0.279787	1	0.000923	0.000258
23	0.918392	0.843443	7	0.006458	0.005447
24	0.565335	0.319604	5	0.004613	0.001474
25	0.790240	0.624478	2	0.001845	0.001152
26	1.487583	2.212903	9	0.008303	0.018373
27	4.450835	19.809928	8	0.007380	0.146199
28	6.070895	36.855763	5	0.004613	0.169999
29	5.294785	28.034751	2	0.001845	0.051725
30	4.481680	20.085453	4	0.003690	0.074116
31	3.820531	14.596454	4	0.003690	0.053861
32	3.141239	9.867385	1	0.000923	0.009103
33	2.340859	5.479622	2	0.001845	0.010110

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ตาราง 69 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 2 ด้วยวิธีของเลอวิน (Levine method) (ต่อ)

คะแนน	$RMSE_i$	$RMSE_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	$RMSE$
34	1.540273	2.372440	0	0.000000	0.000000
35	0.964480	0.930222	0	0.000000	0.000000
36	1.032274	1.065590	2	0.001845	0.001966
37	1.074399	1.154334	0	0.000000	0.000000
38	1.762212	3.105393	0	0.000000	0.000000
39	2.717729	7.386049	0	0.000000	0.000000
40	3.741163	13.996302	0	0.000000	0.000000
41	4.573941	20.920932	1	0.000923	0.019300
42	5.414344	29.315121	0	0.000000	0.000000
43	5.987248	35.847139	1	0.000923	0.033069
44	2.861409	8.187661	0	0.000000	0.000000
45	1.978385	3.914006	0	0.000000	0.000000
46	1.403549	1.969949	0	0.000000	0.000000
47	1.534760	2.355489	0	0.000000	0.000000
48	2.251805	5.070626	0	0.000000	0.000000
รวม				1.000000	1.030883
ROOT					1.015324

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 70 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธีของเลอวิน (Levine method)

คะแนน	$Bias_i$	${Bias_i}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	$Bias$
0	-1.103827	1.218433	0	0.000000	0.000000
1	-2.052155	4.211339	5	0.004613	0.019425
2	-1.216901	1.480847	29	0.026753	0.039617
3	-0.879295	0.773159	73	0.067343	0.052067
4	-0.834662	0.696661	106	0.097786	0.068124
5	-0.755189	0.570311	118	0.108856	0.062082
6	-0.836736	0.700127	125	0.115314	0.080734
7	-0.594128	0.352988	89	0.082103	0.028982
8	-0.181790	0.033048	63	0.058118	0.001921
9	0.016573	0.000275	55	0.050738	0.000014
10	-0.189667	0.035974	56	0.051661	0.001858
11	-0.080226	0.006436	56	0.051661	0.000332
12	-0.311852	0.097252	53	0.048893	0.004755
13	-0.626029	0.391912	42	0.038745	0.015185
14	-0.664111	0.441043	37	0.034133	0.015054
15	-0.516070	0.266328	30	0.027675	0.007371
16	-0.306663	0.094042	21	0.019373	0.001822
17	-0.062504	0.003907	24	0.022140	0.000086
18	-0.060531	0.003664	14	0.012915	0.000047
19	0.275467	0.075882	15	0.013838	0.001050
20	0.403445	0.162768	9	0.008303	0.001351
21	0.021277	0.000453	9	0.008303	0.000004
22	0.320519	0.102732	1	0.000923	0.000095
23	0.431726	0.186387	8	0.007380	0.001376
24	-0.123733	0.015310	5	0.004613	0.000071
25	-0.745859	0.556305	2	0.001845	0.001026
26	-0.734651	0.539712	8	0.007380	0.003983
27	-5.751221	33.076545	9	0.008303	0.274621
28	-5.320569	28.308456	5	0.004613	0.130574
29	-4.523250	20.459794	2	0.001845	0.037749
30	-3.934265	15.478441	4	0.003690	0.057116
31	-3.511946	12.333767	4	0.003690	0.045512
32	-2.839628	8.063485	1	0.000923	0.007439
33	-1.917309	3.676073	1	0.000923	0.003391

ตาราง 70 การคำนวณค่า Weighted average root mean square bias (Bias) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธีของเลอวิน (Levine method) (ต่อ)

คะแนน	$Bias_i$	$Bias_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = $f_i/1084$	$Bias$
34	-0.994990	0.990005	1	0.000923	0.000913
35	-0.989338	0.978790	0	0.000000	0.000000
36	-0.400353	0.160282	2	0.001845	0.000296
37	-0.311367	0.096950	0	0.000000	0.000000
38	0.777618	0.604690	0	0.000000	0.000000
39	1.866604	3.484209	0	0.000000	0.000000
40	2.955589	8.735506	0	0.000000	0.000000
41	3.044574	9.269433	1	0.000923	0.008551
42	-4.366440	19.065801	0	0.000000	0.000000
43	-3.277455	10.741711	0	0.000000	0.000000
44	-2.188470	4.789399	0	0.000000	0.000000
45	-1.099484	1.208865	0	0.000000	0.000000
46	-0.010499	0.000110	0	0.000000	0.000000
47	1.078487	1.163133	1	0.000923	0.001073
48	2.167472	4.697935	0	0.000000	0.000000
รวม				1.000000	0.975666
ROOT					0.987758

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ตาราง 71 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธีของเลอวิน (Levine method)

คะแนน	SEE_i	$SEE_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	SEE
0	0.356686	0.127225	0	0.000000	0.000000
1	0.325751	0.106114	5	0.004613	0.000489
2	0.295744	0.087464	29	0.026753	0.002340
3	0.266977	0.071277	73	0.067343	0.004800
4	0.239897	0.057551	106	0.097786	0.005628
5	0.215143	0.046287	118	0.108856	0.005039
6	0.193609	0.037484	125	0.115314	0.004322
7	0.176477	0.031144	89	0.082103	0.002557
8	0.165123	0.027266	63	0.058118	0.001585
9	0.160776	0.025849	55	0.050738	0.001312
10	0.163995	0.026894	56	0.051661	0.001389
11	0.174360	0.030401	56	0.051661	0.001571
12	0.190710	0.036370	53	0.048893	0.001778
13	0.211663	0.044801	42	0.038745	0.001736
14	0.235995	0.055694	37	0.034133	0.001901
15	0.262770	0.069048	30	0.027675	0.001911
16	0.291315	0.084864	21	0.019373	0.001644
17	0.321158	0.103143	24	0.022140	0.002284
18	0.351970	0.123883	14	0.012915	0.001600
19	0.383516	0.147084	15	0.013838	0.002035
20	0.415630	0.172748	9	0.008303	0.001434
21	0.448189	0.200874	9	0.008303	0.001668
22	0.481104	0.231461	1	0.000923	0.000214
23	0.514306	0.264510	8	0.007380	0.001952
24	0.547742	0.300021	5	0.004613	0.001384
25	0.581373	0.337994	2	0.001845	0.000624
26	0.615166	0.378429	8	0.007380	0.002793
27	0.649096	0.421326	9	0.008303	0.003498
28	0.683143	0.466684	5	0.004613	0.002153
29	0.717290	0.514505	2	0.001845	0.000949
30	0.751523	0.564787	4	0.003690	0.002084
31	0.785831	0.617531	4	0.003690	0.002279
32	0.820205	0.672737	1	0.000923	0.000621
33	0.854637	0.730404	1	0.000923	0.000674

ตาราง 71 การคำนวณค่า Weighted average standard error of equating (SEE) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธีของเลอวิน (Levine method) (ต่อ)

คะแนน	SEE_i	SEE_i^2	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	SEE
34	0.889120	0.790534	1	0.000923	0.000729
35	0.923648	0.853125	0	0.000000	0.000000
36	0.958216	0.918179	2	0.001845	0.001694
37	0.992821	0.985694	0	0.000000	0.000000
38	1.027458	1.055671	0	0.000000	0.000000
39	1.062125	1.128110	0	0.000000	0.000000
40	1.096818	1.203010	0	0.000000	0.000000
41	1.131536	1.280373	1	0.000923	0.001181
42	1.166275	1.360197	0	0.000000	0.000000
43	1.201034	1.442483	0	0.000000	0.000000
44	1.235812	1.527231	0	0.000000	0.000000
45	1.270607	1.614441	0	0.000000	0.000000
46	1.305417	1.704113	0	0.000000	0.000000
47	1.340241	1.796247	1	0.000923	0.001657
48	1.375079	1.890842	0	0.000000	0.000000
รวม				1.000000	0.073507
ROOT					0.271122

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 72 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธีของเลอวิน (Levine method)

คะแนน	$RMSE_i$	$RMSE_i{}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = f_i/ 1084	$RMSE$
0	1.160025	1.345659	0	0.000000	0.000000
1	2.077848	4.317453	5	0.004613	0.019914
2	1.252322	1.568312	29	0.026753	0.041957
3	0.918932	0.844436	73	0.067343	0.056867
4	0.868453	0.754211	106	0.097786	0.073751
5	0.785237	0.616598	118	0.108856	0.067120
6	0.858843	0.737612	125	0.115314	0.085057
7	0.619784	0.384133	89	0.082103	0.031539
8	0.245588	0.060313	63	0.058118	0.003505
9	0.161628	0.026124	55	0.050738	0.001325
10	0.250735	0.062868	56	0.051661	0.003248
11	0.191931	0.036837	56	0.051661	0.001903
12	0.365544	0.133622	53	0.048893	0.006533
13	0.660843	0.436713	42	0.038745	0.016921
14	0.704796	0.496737	37	0.034133	0.016955
15	0.579117	0.335376	30	0.027675	0.009282
16	0.422974	0.178907	21	0.019373	0.003466
17	0.327184	0.107049	24	0.022140	0.002370
18	0.357137	0.127547	14	0.012915	0.001647
19	0.472193	0.222967	15	0.013838	0.003085
20	0.579238	0.335516	9	0.008303	0.002786
21	0.448694	0.201326	9	0.008303	0.001672
22	0.578094	0.334193	1	0.000923	0.000308
23	0.671489	0.450898	8	0.007380	0.003328
24	0.561544	0.315331	5	0.004613	0.001454
25	0.945674	0.894300	2	0.001845	0.001650
26	0.958197	0.918141	8	0.007380	0.006776
27	5.787735	33.497871	9	0.008303	0.278119
28	5.364246	28.775140	5	0.004613	0.132727
29	4.579771	20.974299	2	0.001845	0.038698
30	4.005400	16.043228	4	0.003690	0.059200
31	3.598791	12.951298	4	0.003690	0.047791
32	2.955710	8.736221	1	0.000923	0.008059
33	2.099161	4.406478	1	0.000923	0.004065

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตาราง 72 การคำนวณค่า Weighted average RMSE (RMSE) ในการออกแบบการปรับเทียบรูปแบบที่ 3 ด้วยวิธีของเลอวิน (Levine method) (ต่อ)

คะแนน	$RMSE_i$	${RMSE_i}^2$	ความถี่ของกลุ่ม New ในแต่ละคะแนน, f_i	W = $f_i/1084$	$RMSE$
34	1.334369	1.780539	1	0.000923	0.001643
35	1.353483	1.831915	0	0.000000	0.000000
36	1.038490	1.078461	2	0.001845	0.001990
37	1.040502	1.082643	0	0.000000	0.000000
38	1.288550	1.660361	0	0.000000	0.000000
39	2.147631	4.612318	0	0.000000	0.000000
40	3.152541	9.938516	0	0.000000	0.000000
41	3.248046	10.549805	1	0.000923	0.009732
42	4.519513	20.425998	0	0.000000	0.000000
43	3.490587	12.184194	0	0.000000	0.000000
44	2.513291	6.316630	0	0.000000	0.000000
45	1.680270	2.823307	0	0.000000	0.000000
46	1.305459	1.704223	0	0.000000	0.000000
47	1.720285	2.959380	1	0.000923	0.002730
48	2.566861	6.588777	0	0.000000	0.000000
รวม				1.000000	1.049173
ROOT					1.024291

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ประวัติผู้เขียนวิทยานิพนธ์

ชื่อ พนิดา พานิชวัฒนะ เกิดเมื่อวันที่ 22 มีนาคม พ.ศ. 2532 ที่จังหวัดพิษณุโลก สำเร็จการศึกษาปริญญาครุศาสตรบัณฑิต สาขามัธยมศึกษา-วิทยาศาสตร์ วิชาเอก ฟิสิกส์ จากจุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2555 และเข้าศึกษาต่อในหลักสูตรครุศาสตรมหาบัณฑิต สาขาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University