# คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม: การประยุกต์ใช้การปรับเทียบตามทฤษฎีการตอบสนองข้อสอบด้วยวิธีโค้งคุณลักษณะและการปรับค่าพารามิเตอร์พร้อมกัน

นางสาวอัญชลี ศรีกลชาญ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรดุษฎีบัณฑิต
สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2552
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

# THE QUALITY OF EQUATING FOR MIXED-FORMAT TESTS: AN APPLICATION OF IRT EQUATING WITH CHARACTERISTIC CURVE AND CONCURRENT CALIBRATION METHODS

Miss Anchalee Srikolchan

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Educational Measurement and Evaluation
Department of Educational Research and Psychology
Faculty of Education Chulalongkorn University
Academic Year 2009

Copyright of Chulalongkorn University

| | |
|---|---|
| หัวข้อวิทยานิพนธ์ | คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม: การประยุกต์ใช้การปรับเทียบตามทฤษฎีการตอบสนองข้อสอบด้วยวิธีโค้งคุณลักษณะและการปรับค่าพารามิเตอร์พร้อมกัน |
| โดย | นางสาวอัญชลี ศรีกลชาญ |
| สาขาวิชา | การวัดและประเมินผลการศึกษา |
| อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก | รองศาสตราจารย์ ดร.โชติกา ภาษีผล |
| อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม | ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี |

---

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาดุษฎีบัณฑิต

..... คณบดีคณะครุศาสตร์
(ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(รองศาสตราจารย์ ดร.โชติกา ภาษีผล)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี)

..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.ณัฏฐภรณ์ หลาวทอง)

..... กรรมการภายนอกมหาวิทยาลัย
(รองศาสตราจารย์ ดร.สุภมาส อังศุโชติ)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

อัญชลี ศรีกลชาญ : คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม: การประยุกต์ใช้การปรับเทียบตามทฤษฎีการตอบสนองข้อสอบด้วยวิธีโค้งคุณลักษณะและการปรับค่าพารามิเตอร์พร้อมกัน. (THE QUALITY OF EQUATING FOR MIXED-FORMAT TESTS: AN APPLICATION OF IRT EQUATING WITH CHARACTERISTIC CURVE AND CONCURRENT CALIBRATION METHODS) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : รศ.ดร.โชติกา ภาษีผล, อ. ที่ปรึกษาวิทยานิพนธ์ร่วม ศ.ดร.ศริชัย กาญจนวาสี, 340 หน้า.

การศึกษาครั้งนี้มีจุดมุ่งหมายเพื่อเปรียบเทียบคุณภาพการปรับเทียบคะแนนตามแนวทฤษฎีการตอบสนองข้อสอบ สำหรับแบบสอบรูปแบบผสมระหว่างข้อสอบแบบเลือกตอบที่ตรวจให้คะแนนสองค่า และข้อสอบแบบเขียนตอบที่ตรวจให้คะแนนหลายค่า การศึกษาครั้งนี้ใช้การจำลองข้อมูลโดยข้อสอบที่ตรวจให้คะแนนสองค่า จำลองข้อมูลตามโมเดลโลจีสติก 3 พารามิเตอร์ และข้อสอบที่ตรวจให้คะแนนแบบหลายค่าจำลองข้อมูลตามโมเดลการให้คะแนนบางส่วนแบบทั่วไป (Generalized partial credit model) ที่แตกต่างกันตามรูปแบบการเก็บรวบรวมข้อมูล สัดส่วนของข้อสอบทั้ง 2 ชนิดในแบบสอบ สัดส่วนของข้อสอบทั้ง 2 ชนิดในข้อสอบร่วม และวิธีการปรับเทียบ (2×4×6×2) จำนวนทั้งหมด 96 เงื่อนไข การประเมินคุณภาพการปรับเทียบพิจารณาจากดัชนี MSE

ผลการศึกษา สรุปได้ดังนี้

1. คุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะและการปรับค่าพารามิเตอร์พร้อมกัน มีความคลาดเคลื่อนต่ำ และให้ผลที่ใกล้เคียงกันเมื่อมีรูปแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน สำหรับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน วิธีการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกันจะมีความคลาดเคลื่อนในการปรับเทียบต่ำกว่าวิธีโค้งคุณลักษณะ

2. คุณภาพของการปรับเทียบคะแนนทั้งวิธีโค้งคุณลักษณะและการปรับค่าพารามิเตอร์พร้อมกัน มีความคลาดเคลื่อนต่ำลงเมื่อ สัดส่วนของข้อสอบทั้งสองชนิดในแบบสอบมีจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าเพิ่มขึ้น และเมื่อสัดส่วนของข้อสอบทั้งสองชนิดในข้อสอบร่วมมีจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าลดลง

3. ผลการเปรียบเทียบค่าเฉลี่ย MSE สำหรับการปรับเทียบคะแนนที่มีสัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่ากับหลายค่าในแบบสอบที่แตกต่างกัน 4 สัดส่วน พบว่า ค่าเฉลี่ย MSE ที่ได้จากการปรับเทียบคะแนนสำหรับแบบสอบทั้ง 4 สัดส่วนไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และค่าเฉลี่ย MSE ที่ได้จากการปรับเทียบคะแนนสำหรับข้อสอบร่วมทั้ง 6 สัดส่วนไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4. ผลการเปรียบเทียบค่าเฉลี่ย MSE สำหรับการปรับเทียบแต่ละวิธีที่มีรูปแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน พบว่า มีอย่างน้อย 1 คู่ที่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .05 โดยคู่ที่มีความแตกต่างมากที่สุดคือ การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม กับการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบไม่เท่าเทียม

ภาควิชา วิจัยและจิตวิทยาการศึกษา ลายมือชื่อนิสิต .....

สาขาวิชา การวัดและประเมินผลการศึกษา ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก .....

ปีการศึกษา 2552 ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม .....

## 4884661527 : MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION
KEYWORDS : EQUATING/ MIXED – FORMAT TEST/ CHARACTERISTIC CURVE METHOD / CONCURRENT CALIBRATION METHOD

ANCHALEE SRIKOLCHAN : THE QUALITY OF EQUATING FOR MIXED-FORMAT TESTS: AN APPLICATION OF IRT EQUATING WITH CHARACTERISTIC CURVE AND CONCURRENT CALIBRATION METHODS. THESIS ADVISOR : ASSOC.PROF. SHOTIGA PASIPHOL, Ph.D., THESIS CO-ADVISOR: PROF. SIRICHAI KANJANAWASEE, Ph.D , 340 pp.

This study was to investigate the quality of equating method based on Item Response Theory (IRT) equating for mixed-format test consisting in terms of different item formats (MC and CR) where the MC was dichotomous response model and the CR was polytomous response model. The data of MC and CR were simulated by fitting to the model 3PL and GPC, respectively. There were 96 conditions of the data simulation aspect to the 4 variables as follows (2X4X6X2) : 1) data collection designs, 2) proportions between MC and CR in test 3) proportions between MC and CR in common item, and 4) equating methods. The effectiveness, in this study, has been performed by the mean-squared error (MSE) in evaluating the equating methods.

The obtained results were briefly detailed as follows :

1. The quality of equating methods by means of Characteristic curve and Concurrent calibration methods with the equivalent groups design were low MSE. However, design was on non-equivalent groups its MSE by means of Concurrent calibration were lower than Characteristic curve method.

2. The quality of equating methods by means of Characteristic curve and Concurrent calibration methods with the equivalent groups design were decreased when the number of MC was increased in accordance with the proportion between MC and CR in Tests and the number of MC was decreased in accordance with the proportion between MC and CR in common items set.

3. The comparison results for 4 proportions between MC and CR in terms of mixed-format test was found that the obtained average MSE index presents statistically non-significant difference ($p < .05$) and the comparison results for 6 proportions between MC and CR in terms of common items set also.

4. The comparison results for 2 designs nested in 2 equating methods was found that the obtained average MSE index presents significantly difference ($p < .05$), largest difference was on the Characteristic with non-equivalent groups and curve and Concurrent calibration methods with non-equivalent groups.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Department : Educational Research and Psychology  Student's Signature ....

Field of Study : Educational Measurement and Evaluation  Advisor's Signature S. pasiphol

Academic Year : 2009  Co-Advisor's Signature S. Kanjanawasee

# กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ สำเร็จลงได้ด้วยความกรุณาจาก รองศาสตราจารย์ ดร.โชติกา ภาษีผล ที่ปรึกษาวิทยานิพนธ์ ที่กรุณาให้คำปรึกษาแนะนำแนวทางที่ถูกต้อง ตลอดจนแก้ไข ข้อบกพร่องต่างๆ ด้วยความละเอียดถี่ถ้วน และเอาใจใส่ด้วยดีเสมอมา อีกทั้งยังคอยกำกับติดตามอย่างสม่ำเสมอ จนทำให้วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี ผู้วิจัยรู้สึกซาบซึ้งเป็นอย่างยิ่ง จึงขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณ ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี ที่ปรึกษาวิทยานิพนธ์ร่วม ที่สละเวลาอันมีค่า เพื่อให้ข้อเสนอแนะที่เป็นประโยชน์ ต่อการทำวิทยานิพนธ์ฉบับนี้จนสมบูรณ์

ขอขอบพระคุณ รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ ประธานกรรมการสอบวิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.ณัฏฐภรณ์ หลาวทอง และรองศาสตราจารย์ ดร.สุภมาส อังศุโชติ กรรมการสอบวิทยานิพนธ์ ที่ได้สละเวลาอันมีค่ามาร่วมสอบวิทยานิพนธ์และให้ข้อเสนอแนะที่เป็นประโยชน์ในการทำวิทยานิพนธ์ฉบับนี้จนเสร็จสมบูรณ์

ขอขอบพระคุณคณาจารย์ภาควิชาวิจัยและจิตวิทยาทุกท่าน ที่สั่งสอน ให้ความรู้ มอบความรักและความห่วงใยให้กับนิสิตทุกคน ด้วยความเมตตา ผู้วิจัยรู้สึกประทับใจและซาบซึ้งเป็นอย่างมาก

ขอขอบคุณ น.ต.ดร.หญิง ดร.หฤทัย อาจปรุ ที่มีความอดทนในการรับฟังปัญหา และให้กำลังใจด้วยดีเสมอมาโดยตลอด รวมทั้งเพื่อน ๆ ร่วมชั้นเรียนที่คอยถามไถ่ด้วยความห่วงใย และให้กำลังใจกันและกันเสมอมา

ขอขอบคุณ คุณชยุตม์ ภิรมย์สมบัติ ที่ให้ความช่วยเหลือในการค้นคว้าเอกสารจากต่างประเทศ อาจารย์บัญชา ศรีสมบัติ ที่ให้ความช่วยเหลือในการตรวจสอบความถูกต้องของการวิเคราะห์ข้อมูล คุณกีรติ สุขในสิทธิ์ ที่สละเวลาในการพิสูจน์อักษร ตลอดจนเป็นแรงใจในการมุ่งมั่นทำวิทยานิพนธ์จนเสร็จสมบูรณ์

ขอขอบคุณบุคคลทุกท่านที่มิได้ปรากฏชื่อในที่นี้ ที่มีส่วนช่วยเหลือในการทำวิทยานิพนธ์ฉบับนี้ จนสำเร็จลุล่วงไปด้วยดี

ท้ายที่สุด ขอกราบขอบพระคุณ บิดา มารดา สำหรับความรัก ความห่วงใย ความเอาใจใส่และกำลังใจที่มีให้ผู้วิจัยตลอดมา วิทยานินพนธ์ฉบับนี้สำเร็จลุล่วงได้เพราะมีบุคคลเหล่านี้เป็นแรงใจในการขับเคลื่อนจนเสร็จสมบูรณ์

# สารบัญ


หน้า

บทคัดย่อภาษาไทย..... ง

บทคัดย่อภาษาอังกฤษ..... จ

กิตติกรรมประกาศ..... ฉ

สารบัญ..... ช

สารบัญตาราง..... ฌ

สารบัญแผนภาพ..... ฑ

บทที่

1 บทนำ ..... 1

ความเป็นมาและความสำคัญของปัญหา ..... 1

วัตถุประสงค์ของการวิจัย ..... 11

สมมติฐานการวิจัย ..... 12

ขอบเขตของการวิจัย ..... 14

กรอบแนวคิดสำหรับการวิจัย ..... 18

คำนิยามเชิงปฏิบัติการที่ใช้ในการวิจัย ..... 20

ประโยชน์ที่คาดว่าจะได้รับ ..... 22

2 เอกสารและงานวิจัยที่เกี่ยวข้อง ..... 23

ตอนที่ 1 ความสำคัญของแบบสอบรูปแบบผสม ..... 14

ตอนที่ 2 ทฤษฎีการตอบสนองข้อสอบ ..... 29

ตอนที่ 3 การปรับเทียบคะแนน ..... 51

ตอนที่ 4 การปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ ..... 65

ตอนที่ 5 โปรแกรมคอมพิวเตอร์ที่ใช้ในการวิจัย..... 93

ตอนที่ 6 งานวิจัยที่เกี่ยวข้องกับการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม..... 114

3 วิธีดำเนินการวิจัย ..... 143

ตอนที่ 1 เงื่อนไขที่ศึกษาการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม..... 147

# สารบัญ (ต่อ)


| บทที่ | | หน้า |
|---|---|---|
| 3 | ตอนที่ 2 การจำลองข้อมูลเพื่อดำเนินการปรับเทียบคะแนนตามเงื่อนไขที่กำหนด... | 149 |
| | ตอนที่ 3 การตรวจสอบความถูกต้องของข้อมูล ..... | 157 |
| | ตอนที่ 4 การปรับเทียบคะแนน ..... | 168 |
| | ตอนที่ 5 การประเมินคุณภาพของการปรับเทียบคะแนน..... | 170 |
| 4 | ผลการวิเคราะห์ข้อมูล..... | 172 |
| | สัญลักษณ์ที่ใช้ในการนำเสนอผลการวิเคราะห์ข้อมูล..... | 172 |
| | ตอนที่ 1 ผลการวิเคราะห์การปรับเทียบและการเปรียบเทียบคุณภาพของการปรับเทียบ..... ..... | 173 |
| | ตอนที่ 2 การสรุปผลการวิจัยตามวัตถุประสงค์..... | 219 |
| 5 | สรุป อภิปราย และข้อเสนอแนะ..... | 227 |
| | สรุปผลการวิจัย..... | 229 |
| | อภิปรายผลการวิจัย..... | 235 |
| | ข้อเสนอแนะ..... | 246 |
| รายการอ้างอิง..... | | 252 |
| ภาคผนวก..... | | 260 |
| ภาคผนวก ก ..... | | 261 |
| ภาคผนวก ข ..... | | 271 |
| ภาคผนวก ค ..... | | 315 |
| ประวัติผู้เขียนวิทยานิพนธ์..... | | 340 |


ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

# สารบัญตาราง


| ตารางที่ | | หน้า |
|---|---|---|
| 1 | การออกแบบการปรับเทียบคะแนน ..... | 19 |
| 2 | ตัวอย่างพารามิเตอร์ข้อสอบและค่าความสามารถของผู้สอบสำหรับ 2 สเกลของแบบสอบ ..... | 67 |
| 3 | สรุปประเด็นสำคัญของการเทียบคะแนนในแบบสอบรูปแบบผสม ..... | 136 |
| 4 | รูปแบบข้อสอบที่ใช้ในการปรับเทียบคะแนนจำแนกตามสัดส่วนของข้อสอบในแบบสอบ..... | 151 |
| 5 | ค่าเฉลี่ยของพารามิเตอร์ของข้อสอบที่ได้จากโปรแกรม Wingen..... | 158 |
| 6 | ผลการเปรียบเทียบค่าเฉลี่ยพารามิเตอร์ข้อสอบจากโปรแกรม WINGEN และ MULTILOG สำหรับ Test A จากการประมาณค่า 100 ครั้ง..... | 159 |
| 7 | ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test B1..... | 160 |
| 8 | ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test B2..... | 161 |
| 9 | ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของเงื่อนไขสัดส่วนของข้อสอบในแบบสอบฉบับ 50 : 10 ด้วยการทำซ้ำ 100 ครั้ง ..... | 163 |
| 10 | ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของเงื่อนไขสัดส่วนของข้อสอบในแบบสอบฉบับ 60 : 8 ด้วยการทำซ้ำ 100 ครั้ง..... | 164 |
| 11 | ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของเงื่อนไขสัดส่วนของข้อสอบในแบบสอบฉบับ 70 : 6 ด้วยการทำซ้ำ 100 ครั้ง..... | 165 |
| 12 | ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของเงื่อนไขสัดส่วนของข้อสอบในแบบสอบฉบับ 80 : 4 ด้วยการทำซ้ำ 100 ครั้ง..... | 166 |
| 13 | ความเที่ยงของแบบสอบแต่ละฉบับ..... | 167 |
| 14 | ค่าMSEจากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบ DS50 : PS10 | 174 |
| 15 | ค่าMSEจากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบ DS60 : PS8 | 176 |

# สารบัญตาราง (ต่อ)


| ตารางที่ | | หน้า |
|---|---|---|
| 16 | ค่าMSEจากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบ DS70 : PS6.. | 178 |
| 17 | ค่าMSEจากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบ DS80 : PS 4. | 180 |
| 18 | ค่าMSEจากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบร่วม DS25 : PS2. | 183 |
| 19 | ค่าMSEจากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบร่วม DS20 : PS2..... | 185 |
| 20 | ค่าMSEจากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบร่วม DS15 : PS2..... | 187 |
| 21 | ค่าMSEจากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบร่วม DS5 : PS3 | 189 |
| 22 | ค่าMSEจากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบร่วม DS10 : PS2..... | 191 |
| 23 | ค่า MSE จากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบร่วม DS15 : PS1..... | 193 |
| 24 | ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม... | 196 |
| 25 | การวิเคราะห์ความแตกต่างสัดส่วนข้อสอบในแบบสอบจากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับกลุ่มเท่าเทียม... | 208 |
| 26 | การวิเคราะห์ความแตกต่างสัดส่วนข้อสอบในแบบสอบจากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับกลุ่มไม่เท่าเทียม..... | 209 |
| 27 | การวิเคราะห์ความแตกต่างสัดส่วนข้อสอบในแบบสอบจากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีการเก็บรวบรวมข้อมูลกับกลุ่มเท่าเทียม..... | 210 |
| 28 | การวิเคราะห์ความแตกต่างสัดส่วนข้อสอบในแบบสอบจากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีการเก็บรวบรวมข้อมูลกับกลุ่มไม่เท่าเทียม..... | 210 |

# สารบัญตาราง (ต่อ)


| ตารางที่ | | หน้า |
|---|---|---|
| 29 | การวิเคราะห์ความแตกต่างระหว่างข้อสอบร่วมแต่ละสัดส่วน จากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยแบบสอบที่มีสัดส่วน DS 50 : PS 10..... | 211 |
| 30 | การวิเคราะห์ความแตกต่างระหว่างข้อสอบร่วมแต่ละสัดส่วน จากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยแบบสอบที่มีสัดส่วน DS 60 : PS 8..... | 212 |
| 31 | การวิเคราะห์ความแตกต่างระหว่างข้อสอบร่วมแต่ละสัดส่วน จากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยแบบสอบที่มีสัดส่วน DS 70 : PS 6..... | 213 |
| 32 | การวิเคราะห์ความแตกต่างระหว่างข้อสอบร่วมแต่ละสัดส่วน จากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยแบบสอบที่มีสัดส่วน DS 80 : PS 4..... | 213 |
| 33 | การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนี MSE ของการปรับเทียบที่มีรูปแบบการเก็บรวบรวมข้อมูลต่างกัน สำหรับแบบสอบสัดส่วน DS 50 : PS 10..... | 214 |
| 34 | การวิเคราะห์ความแตกต่างของค่า MSE เฉลี่ย ของการปรับเทียบ 4 วิธีที่แตกต่างกัน เมื่อจำแนกตามแบบสอบ DS 60 : PS 8..... | 215 |
| 35 | การวิเคราะห์ความแตกต่างของค่า MSE เฉลี่ย ของการปรับเทียบ 4 วิธีที่แตกต่างกัน เมื่อจำแนกตามแบบสอบ DS 70 : PS 6..... | 216 |
| 36 | การวิเคราะห์ความแตกต่างของค่า MSE เฉลี่ย ของการปรับเทียบ 4 วิธีที่แตกต่างกัน เมื่อจำแนกตามแบบสอบ DS 80 : PS 4..... | 217 |
| 37 | การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนี MSE ของการปรับเทียบแต่ละวิธีที่มีรูปแบบการเก็บรวบรวมข้อมูลแตกต่างกัน (ภาพรวม)..... | 218 |
| 38 | ผลการทดสอบรายคู่ที่มีนัยสำคัญทางสถิติที่ระดับ .05 พิจารณาตามสัดส่วนของแบบสอบ 4 สัดส่วน..... | 223 |
| 39 | ผลการทดสอบรายคู่ที่มีนัยสำคัญทางสถิติที่ระดับ .05 พิจารณาตามสัดส่วนของข้อสอบร่วม 6 สัดส่วน..... | 224 |

# สารบัญตาราง (ต่อ)


| ตารางที่ | | หน้า |
|---|---|---|
| 40 | ผลการทดสอบรายคู่ที่มีนัยสำคัญทางสถิติที่ระดับ .05 พิจารณาตามวิธีการปรับเทียบและการเก็บรวบรวมข้อมูล..... | 225 |
| 41 | ค่าพารามิเตอร์ของข้อสอบที่ใช้ในการจำลองข้อมูลผลการตอบข้อสอบสำหรับแบบสอบฉบับ A..... | 262 |
| 42 | ค่าพารามิเตอร์ของข้อสอบที่ใช้ในการจำลองข้อมูลผลการตอบข้อสอบ สำหรับแบบสอบฉบับ B1..... | 265 |
| 43 | ค่าพารามิเตอร์ของข้อสอบที่ใช้ในการจำลองข้อมูลผลการตอบข้อสอบ สำหรับแบบสอบฉบับ B2..... | 268 |
| 44 | ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test A จากการประมาณค่า 100 ครั้ง..... | 273 |
| 45 | ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test B1 จากการประมาณค่า 100 ครั้ง..... | 277 |
| 46 | ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test B2 จากการประมาณค่า 100 ครั้ง..... | 281 |
| 47 | ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบ สำหรับแบบสอบ 50 : 10 ด้วยการทำซ้ำ 100 ครั้ง..... | 286 |
| 48 | ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบ สำหรับแบบสอบ 60 : 8 ด้วยการทำซ้ำ 100 ครั้ง..... | 293 |
| 49 | ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบ สำหรับแบบสอบ 70 : 6 ด้วยการทำซ้ำ 100 ครั้ง..... | 300 |
| 50 | ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบ สำหรับแบบสอบ 80 : 4 ด้วยการทำซ้ำ 100 ครั้ง..... | 306 |
| 51 | ความเที่ยงของแบสอบแต่ละฉบับ..... | 312 |
| 52 | ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 50 และ PS 10 ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียมโดยใช้ข้อสอบร่วม..... | 316 |

# สารบัญตาราง (ต่อ)


| ตารางที่ | | หน้า |
|---|---|---|
| 53 | ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 60 และ PS 8 ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียมโดยใช้ข้อสอบร่วม..... | 319 |
| 54 | ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 70 และ PS 6 ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียมโดยใช้ข้อสอบร่วม..... | 322 |
| 55 | ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 80 และ PS 4 ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียมโดยใช้ข้อสอบร่วม..... | 325 |
| 56 | ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 50 และ PS 10 ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม..... | 328 |
| 57 | ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 60 และ PS 8 ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม..... | 331 |
| 58 | ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 70 และ PS 6 ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม..... | 334 |
| 59 | ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 80 และ PS 4 ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม..... | 337 |


ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

# สารบัญแผนภาพ


| แผนภาพที่ | | หน้า |
|---|---|---|
| 1 | กรอบแนวคิดในการวิจัย ..... | 19 |
| 2 | หน้าจอแสดงขั้นตอนการจำลองข้อมูลพารามิเตอร์ความสามารถของผู้สอบ..... | 96 |
| 3 | ตัวอย่างการจำลองข้อมูลพารามิเตอร์ความสามารถของผู้สอบ..... | 97 |
| 4 | หน้าจอแสดงขั้นตอนการจำลองข้อมูลพารามิเตอร์ความสามารถของข้อสอบ... | 98 |
| 5 | ตัวอย่างการจำลองข้อมูลพารามิเตอร์ข้อสอบ..... | 99 |
| 6 | ตัวอย่างกราฟแสดงโค้งคุณลักษณะ..... | 99 |
| 7 | หน้าจอแสดงขั้นตอนการจำลองข้อมูลผลการตอบข้อสอบ..... | 100 |
| 8 | ตัวอย่างการจำลองข้อมูลผลการตอบข้อสอบ..... | 101 |
| 9 | เมนูอื่นๆ สำหรับการจำลองข้อมูล..... | 102 |
| 10 | แสดงตัวอย่างค่าพารามิเตอร์ความสามารถของผู้สอบและพารามิเตอร์ข้อสอบ. | 104 |
| 11 | ตัวอย่างไฟล์ Output ค่าสถิติบรรยาย และผลการตอบข้อสอบ..... | 105 |
| 12 | หน้าจอหลักของโปรแกรม IRTEQ..... | 113 |
| 13 | แสดงการนำข้อมูลเข้าสู่โปรแกรม..... | 113 |
| 14 | แสดงการกระจายของพารามิเตอร์ข้อสอบที่เป็นข้อสอบร่วม..... | 113 |
| 15 | โค้งคุณลักษณะของแบบสอบสำหรับแบบสอบแต่ละฉบับ..... | 114 |
| 16 | การจำลองข้อมูลตามเงื่อนไขที่ศึกษา..... | 157 |
| 17 | ขั้นตอนการจำลองข้อมูล..... | 171 |
| 18 | ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่าง DS 50 : PS 10. | 176 |
| 19 | ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่าง DS 60 : PS 8… | 178 |
| 20 | ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่าง DS 70 : PS 6… | 180 |
| 21 | ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่าง DS 80 : PS 4... | 182 |
| 22 | ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วม DS25 : PS2 | 184 |
| 23 | ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วม DS20 : PS2 | 186 |
| 24 | ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วม DS15 : PS2 | 188 |
| 25 | ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วม DS5 : PS3.. | 190 |

# สารบัญแผนภาพ (ต่อ)


| แผนภาพที่ | | หน้า |
|---|---|---|
| 26 | ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วม DS10 : PS2 | 192 |
| 27 | ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วม DS15 : PS1 | 194 |
| 28 | ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบในแบบสอบเท่ากับ DS 50 : PS 10 | 198 |
| 29 | ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบในแบบสอบเท่ากับ DS 60 : PS 8 | 199 |
| 30 | ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบในแบบสอบเท่ากับ DS 70 : PS 6 | 200 |
| 31 | ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบในแบบสอบเท่ากับ DS 80 : PS 4 | 201 |
| 32 | ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบร่วมในแบบสอบเท่ากับ DS 25 : PS 2 | 202 |
| 33 | ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบร่วมในแบบสอบเท่ากับ DS 20 : PS 2 | 203 |
| 34 | ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบร่วมในแบบสอบเท่ากับ DS 15 : PS 2 | 204 |
| 35 | ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบร่วมในแบบสอบเท่ากับ DS 5 : PS 3 | 205 |
| 36 | ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบร่วมในแบบสอบเท่ากับ DS 10 : PS 2 | 206 |
| 37 | ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบร่วมในแบบสอบเท่ากับ DS 15 : PS 1 | 207 |

# บทที่ 1
# บทนำ

## ความเป็นมาและความสำคัญของปัญหา

สถานการณ์การวัดและประเมินผลแบบดั้งเดิมส่วนใหญ่เป็นการทดสอบโดยใช้ข้อสอบแบบหลายตัวเลือก (Multiple Choice: MC) ที่มีการตรวจให้คะแนนแบบสองค่า (Dichotomous) ข้อสอบประเภทนี้มีความน่าเชื่อถือในการตรวจให้คะแนนและมีความยุติธรรม แต่ในขณะเดียวกันข้อสอบลักษณะนี้สามารถสร้างข้อสอบให้วัดความสามารถขั้นสูงได้ค่อนข้างยาก ต้องอาศัยผู้ออกข้อสอบที่มีประสบการณ์และมีความเชี่ยวชาญในเนื้อหาเป็นอย่างมาก หากเป็นการทดสอบโดยใช้ข้อสอบแบบเขียนตอบ ซึ่งเปิดโอกาสให้ผู้สอบได้แสดงความรู้ความสามารถของตนเองมากกว่าการสอบด้วยข้อสอบแบบเลือกตอบ ทำให้พิจารณาระดับความสามารถของผู้สอบในขั้นสูงได้มากยิ่งขึ้น Thorndike และ Hagen (1977), Kubiszyn และ Borich (2003) กล่าวว่า ข้อสอบแบบเขียนตอบหรือข้อสอบอัตนัย สามารถวัดความสามารถของผู้เรียนในการจัดระบบ (organize) การบูรณาการ (integrate) และ การสังเคราะห์ความรู้ (synthesize) เพื่อใช้ข้อมูลในการแก้ปัญหา หรือเป็นการริเริ่มและปรับปรุง วิธีการแก้ปัญหาใหม่ ขณะที่ข้อสอบปรนัยหรือข้อสอบเลือกตอบ ส่วนใหญ่สามารถวัดได้แค่ระดับความจำและวัดได้แคบ ในกรอบของตัวเลือกที่นำเสนอเท่านั้น ไม่สามารถคิดบนพื้นฐานเหตุผลส่วนตัวได้ และในปัจจุบันการประเมินผู้เรียนมีแนวทางที่จะวัดความสามารถของผู้เรียนในระดับที่สูงขึ้น การทดสอบจึงจำเป็นต้องใช้แบบสอบทั้ง 2 รูปแบบควบคู่กันไป ส่วนใหญ่แล้วก็จะใช้รูปแบบของการวัดและประเมินโดยใช้แบบสอบรูปแบบผสม (Mixed-format Tests) ซึ่งประกอบด้วยข้อสอบแบบหลายตัวเลือก (multiple choice : MC) และข้อสอบแบบเขียนตอบ (Constructed Response: CR) เนื่องจากการใช้แบบสอบเพียงรูปแบบใดรูปแบบหนึ่งอาจจะไม่สามารถวัดความรู้ ความสามารถของผู้เรียนได้อย่างครอบคลุม

ปัจจุบันนี้มีการนำแบบสอบรูปแบบผสม (mixed format test) ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่าและหลายค่ารวมในแบบสอบฉบับเดียวกันมาใช้มากขึ้น ข้อสอบเหล่านี้เป็นข้อสอบเลือกตอบชนิดหลายตัวเลือก (MC) และข้อสอบแบบเขียนตอบ (CR) Hambleton (1996) ได้กล่าวว่าการประเมินทางการศึกษาในทศวรรษที่ 20 และในทศวรรษหน้าจะมีความแตกต่างอย่างมากจากการประเมินในสองทศวรรษที่ผ่านมา รูปแบบที่การประเมินมีความเป็นปรนัยดังเช่นแบบสอบเลือกตอบจะ

นำมาใช้ให้สมดุลกับแบบสอบรูปแบบอื่นมากยิ่งขึ้น โดยการวัดร่วมกับการประเมินการปฏิบัติงาน การเขียนโครงงาน และจัดทำพอตโฟลิโอ โดยแบบสอบรูปแบบผสมระหว่างข้อสอบที่มีรูปแบบต่างกัน จะนำไปใช้ทั้งในระดับชั้นเรียน และในระดับมหภาค ซึ่งข้อสอบลักษณะนี้มักใช้วัดทักษะของผู้เรียนในด้านต่าง ๆ มากกว่าที่จะใช้แบบสอบที่ประกอบด้วยข้อสอบเพียงรูปแบบเดียว (Single Format) และแบบสอบรูปแบบผสมส่วนใหญ่เป็นการผสมระหว่างข้อสอบแบบหลายตัวเลือก (MC) และข้อสอบแบบเขียนตอบ (CR) ซึ่งข้อสอบ MC จะมีการตรวจให้คะแนนแบบสองค่า (dichotomous) และข้อสอบ CR จะมีการตรวจให้คะแนนหลายค่า (polytomous)

สำหรับการทดสอบที่มีการใช้แบบสอบรูปแบบผสมที่พบได้ในอดีตจนถึงปัจุบัน เช่น การสอบคัดเลือกบุคคลเข้าศึกษาในสถาบันอุดมศึกษา มีการผสมข้อสอบที่ตรวจให้คะแนนสองค่าและหลายค่าไว้ในแบบสอบฉบับเดียวกัน โดยได้กำหนดสัดส่วนน้ำหนักข้อสอบไว้หลายสัดส่วน ซึ่งในการกำหนดน้ำหนักสัดส่วนระหว่างข้อสอบปรนัยและอัตนัยจะแตกต่างกันไป ขึ้นอยู่กับลักษณะของข้อสอบและลักษณะการตอบ ถ้าข้อสอบอัตนัยเป็นแบบตอบสั้น หรือเติมคำ จะสามารถออกข้อสอบได้ปริมาณมากกว่าข้อสอบอัตนัยที่เขียนตอบแบบความเรียง จากการทดสอบที่ผ่านมามีการนำข้อสอบอัตนัยมาใช้ร่วมกับข้อสอบปรนัยมากยิ่งขึ้น และมีการกำหนดน้ำหนักสัดส่วนของคะแนนในข้อสอบอัตนัยน้อยกว่าข้อสอบปรนัย เนื่องจากธรรมชาติของข้อสอบอัตนัยจะใช้เวลาทำข้อสอบมากกว่าข้อสอบปรนัย และการตรวจให้คะแนนยังมีความคลาดเคลื่อนมากกว่าด้วย เช่น การสอบ O – NET (Ordinary National Education Test) ประจำปีการศึกษา 2548 ได้กำหนดสัดส่วนน้ำหนักของข้อสอบในวิชา ภาษาไทย คณิตศาสตร์ วิทยาศาสตร์ สังคมศึกษา และภาษาอังกฤษ ไว้ 2 สัดส่วนคือ ข้อสอบปรนัย 80 - 90% และข้อสอบอัตนัย 10 – 20 % ของข้อสอบทั้งหมด เป็นต้น หรือการสอบในสถาบันการศึกษาทั่ว ๆ ไป ในบางวิชา เช่น วิชาคณิตศาสตร์ จะมีข้อสอบทั้งที่ตรวจให้คะแนนสองค่า และตรวจให้คะแนนหลายค่ารวมอยู่ด้วย เพื่อให้การสอบมีความครอบคลุมเนื้อหา และสามารถวัดความสามารถขั้นสูงของผู้เรียนได้ หรือแม้แต่การทดสอบความสามารถทางภาษาอังกฤษ TOEFL ในระบบคอมพิวเตอร์ ที่ประกอบด้วยข้อสอบ 4 ส่วนคือ ส่วนการฟัง ส่วนของไวยากรณ์ ส่วนของการอ่าน และส่วนของการเขียน ซึ่งข้อสอบจะมีทั้งที่ตรวจให้คะแนนสองค่า และหลายค่ารวมอยู่ในแบบสอบฉบับเดียวกัน โดยส่วนของการเขียน ผู้สอบจะได้ทำข้อสอบ 1 ข้อ ในประเด็นที่คอมพิวเตอร์สุ่มมาให้ และเขียนตอบเพื่อแสดงความคิดเห็นเกี่ยวกับประเด็นเหล่านั้น และมีการกำหนดการให้คะแนนแบบหลายค่าเป็น 7 ระดับคือ มีระดับคะแนนตั้งแต่ 0 – 6 คะแนน สำหรับการทดสอบความสามารถทางภาษาอังกฤษ TOEFL นี้ ผู้สอบสามารถเข้าสอบได้ตลอดทั้งปี โดยคะแนนผลการสอบ

สามารถเก็บไว้ใช้ได้ 2 ปี ในการสอบผู้สอบจะได้รับข้อสอบที่แตกต่างกันไปในแต่ละครั้ง ซึ่งอาจมีข้อสอบที่มีโครงสร้างเนื้อหา หรือความยากง่ายแตกต่างกันเกิดขึ้นได้

สถานการณ์การทดสอบในปัจจุบันมีความจำเป็นต้องใช้แบบสอบหลายฉบับในการวัดครั้งเดียวกัน หรือใช้แบบสอบหลายฉบับในการประเมินการเรียนรู้ในเวลาที่ต่างกัน ทำให้ต้องมีการเปรียบเทียบผลที่ได้จากการทำแบบสอบต่างฉบับให้อยู่ในมาตรวัดเดียวกัน เพื่อให้คะแนนของแบบสอบต่างฉบับสามารถเปรียบเทียบกันได้ การดำเนินการดังกล่าวต้องอาศัยกระบวนการที่เรียกว่า การปรับเทียบคะแนนระหว่างแบบสอบ (Test equating) ซึ่งเป็นกระบวนการที่เกี่ยวข้องกับกิจกรรม 2 ประการ คือ กระบวนการที่ทำให้แบบสอบสองฉบับใด ๆ มีความทัดเทียมกันหรือเท่ากันในเชิงโครงสร้าง และการใช้วิธีการทางสถิติเพื่อปรับคะแนนที่ได้จากแบบสอบแต่ละฉบับให้อยู่ในมาตรฐานเดียวกันและเทียบกันได้

การปรับเทียบมีบทบาทสำคัญต่อการวัดและการประเมินทางการศึกษาในปัจจุบันมากยิ่งขึ้น และนักวิชาการทางด้านวัดผลการศึกษาชี้ให้เห็นประเด็นว่า แบบสอบที่มีโครงสร้างเนื้อหาและคุณภาพไม่เท่าเทียมกัน เช่น ความยากของข้อสอบ ย่อมทำให้เกิดการได้เปรียบเสียเปรียบในการสอบแต่ละครั้ง เมื่อนำคะแนนสอบไม่เท่าเทียมกันมาเปรียบเทียบกันจึงก่อให้เกิดความไม่ยุติธรรม การปรับเทียบคะแนนระหว่างแบบสอบจึงเป็นกระบวนการที่จำเป็นต่อการนำมาใช้ในสถานการณ์เช่นนี้ เพื่อปรับคะแนนสอบในแต่ละครั้งให้อยู่ในมาตรวัดเดียวกันและเป็นคะแนนที่เปรียบเทียบกันได้ ทำให้การสอบคัดเลือกใช้เกณฑ์ได้ถูกต้องและมีคุณภาพโดยทั่วไปการปรับเทียบจำแนกตามทฤษฎีการทดสอบได้สองวิธีการใหญ่ๆ คือ การปรับเทียบตามทฤษฎีการทดสอบแบบดั้งเดิม และการปรับเทียบตามทฤษฎีการตอบสนองข้อสอบ โดยมีหลักการว่าคะแนนจากแบบสอบต่างฉบับที่วัดในสิ่งเดียวกันถือว่าเทียบเท่ากันได้ ถ้าผู้สอบทำแบบสอบต่างฉบับแล้วได้คะแนนจริง (true score) หรือระดับความสามารถ (ability) เท่ากัน สำหรับวิธีการปรับเทียบตามทฤษฎีการตอบสนองข้อสอบ สามารถจำแนกตามเทคนิควิธีการปรับค่าพารามิเตอร์ได้สองแนวทาง คือ การปรับค่าพารามิเตอร์พร้อมกัน (simultaneous or concurrent calibration) และการปรับค่าพารามิเตอร์แยกกัน (separate calibration) ข้อมูลจากการตอบจะถูกปรับค่าอย่างไรนั้นขึ้นอยู่กับแบบแผนการรวบรวมข้อมูล (Kolen และ Brennan, 2004)

การปรับเทียบคะแนนระหว่างแบบสอบ (Test Equating) เป็นศัพท์เฉพาะที่นักจิตมิติ (Psychometrician) นำมาใช้ในกระบวนการวัดและประเมินผล ในช่วงแรกนักวัดผลโดยส่วนมากไม่ได้ให้ความสนใจเท่าที่ควร จะมีการศึกษากันในเฉพาะกลุ่มนักจิตมิติเท่านั้น (Kolen and Brennan, 2004)

ต่อมาในปี ค.ศ. 1980 ได้มีเอกสารเกี่ยวกับเรื่องนี้เพิ่มมากขึ้นทำให้นักวัดผลเริ่มเห็นความสำคัญและนำมาใช้ให้เป็นประโยชน์กับการทดสอบ 15 ปี ต่อมาได้มีการเพิ่มจำนวนและความหลากหลายของโปรแกรมการสอบที่มีการใช้แบบสอบหลายฉบับ ผู้รับผิดชอบในการทำแบบสอบมีความต้องการที่ใช้เทคนิคและหน้าที่ของการปรับเทียบคะแนนเพื่อค้นหาจุดแข็งจุดอ่อนของการทดสอบ ความเคลื่อนไหวในการตรวจสอบและติดตามผลการจัดการศึกษา และประเด็นความยุติธรรมเกี่ยวกับการทดสอบที่ปรากฏชัดเจนขึ้น ทำให้ศาสตร์ด้านการปรับเทียบคะแนนได้รับการศึกษาและพัฒนาขึ้นมาตามลำดับซึ่งสอดคล้องกับคำกล่าวของ ศิริชัย กาญจนวาสี (2545) ที่กล่าวไว้ว่า การทดสอบในอนาคต จักเป็นระบบการทดสอบที่บูรณาการระหว่างศาสตร์แห่งการวัดแนวใหม่ (Modern Technologies) เพื่อสนองตอบความต้องการใช้สารสนเทศสำหรับการตัดสินใจ เกี่ยวกับการพัฒนาสังคมและทรัพยากรมนุษย์ การวัดและประเมินจะเป็นจักรกลสำคัญของการบริหารแบบมีส่วนร่วมของทุกฝ่ายอย่างสร้างสรรค์ สำหรับใช้ในการจัดกลุ่มความสามารถ คัดเลือก เลือกสรร ติดตามกำกับความก้าวหน้า พร้อมทั้งพัฒนาสังคมและทรัพยากรมนุษย์อย่างเสมอภาคและเที่ยงธรรม โครงสร้าง หลักการ และวิธีดำเนินการทดสอบในอนาคตจักมีความตรงประเด็น กระชับ แต่ยืดหยุ่นได้อย่างสอดคล้องกับความต้องการจำเป็นและสภาวะแวดล้อม ตลอดจนสนองตอบต่อความแตกต่าง (ระหว่างองค์กรและบุคคล) และความต้องการของทุกฝ่ายที่เกี่ยวข้องอย่างมีประสิทธิผลและประสิทธิภาพยิ่งขึ้น การปรับเทียบคะแนนระหว่างแบบสอบเป็นกระบวนการออกแบบการทดสอบและเทคนิคการวิเคราะห์ผลการสอบ เพื่อให้สามารถนำผลการทดสอบจากแบบสอบต่างฉบับและผู้สอบต่างกลุ่มกันมาเปรียบเทียบกันอย่างเที่ยงธรรม การวิจัยเพื่อพัฒนาวิธีการเก็บรวบรวมข้อมูลด้วยการทดสอบ และวิธีการวิเคราะห์แปลงคะแนนที่สะดวกในทางปฏิบัติ เพื่อให้ผลการปรับเทียบที่เที่ยงตรงที่สุดยังคงดำเนินต่อไปอย่างต่อเนื่อง เทคนิคการปรับเทียบคะแนนย่อมมีบทบาทมากขึ้นในอนาคต โดยเข้ามาทดแทนการสร้างแบบสอบคู่ขนานซึ่งยากแก่การปฏิบัติและแทนที่การทดสอบแบบประเพณีนิยมซึ่งขาดความยืดหยุ่น

จากการศึกษางานวิจัยเกี่ยวกับการเชื่อมโยงคะแนนโดยใช้เทคนิคการปรับเทียบคะแนนแบบสอบที่มีการให้คะแนนสองค่า ผลที่ได้แตกต่างกันไปตามองค์ประกอบที่เกี่ยวข้องกับการปรับเทียบ (Bastari, 2000) สรุปได้ว่า การปรับเทียบด้วยวิธีโค้งคุณลักษณะให้ผลที่คงที่มากกว่าวิธี mean/mean, mean/sigma (Baker – Al-karni, 1991; Hung et.al, 1991; Hanson and Beguin 2002) นอกจากนี้ Ogasawara (2002) ยังพบว่า วิธี mean/mean ให้ผลการปรับเทียบที่คงที่มากกว่าวิธี mean/sigma เมื่อ Ogasawara (2002) ได้ประมาณค่าความคลาดเคลื่อนมาตรฐาน (Standard

error) สำหรับพารามิเตอร์ข้อสอบและโค้งคุณลักษณะข้อสอบ (ICC) พบว่า สามารถประมาณค่า ICC ได้ถูกต้อง ในขณะที่พารามิเตอร์ของข้อสอบสามารถประมาณค่าได้อย่างแม่นยำ ข้อค้นพบนี้สนับสนุนข้อค้นพบที่ว่า วิธีการเชื่อมโยงด้วยวิธีโค้งคุณลักษณะของแบบสอบ (Test Characteristic Curve) ให้ความถูกต้องมากกว่าวิธี mean/mean, mean/sigma Kim และ Cohen (1998) ได้เปรียบเทียบวิธีการเชื่อมโยงโดยใช้วิธีโค้งคุณลักษณะแบบสอบตามแบบ Stocking - Lord (1983) กับวิธีการปรับค่าพารามิเตอร์พร้อมกัน โดยใช้โปรแกรม MULTILOG และยังใช้โปรแกรม BILOG ในการตรวจสอบการปรับค่าพารามิเตอร์ด้วย การจำลองข้อมูลเป็นแบบโมเดลโลจีสติก ชนิด 3 พารามิเตอร์ มีจำนวนข้อสอบร่วมเพียงเล็กน้อย ผลการวิจัยพบว่า วิธีการปรับค่าพารามิเตอร์พร้อมกันให้ผลดีกว่าวิธีโค้งคุณลักษณะแบบสอบ เมื่อใช้โปรแกรม MULTILOG ซึ่งสอดคล้องกับงานของ Hanson และ Beguin (2002) ให้ผลการปรับเทียบแบบการปรับค่าพารามิเตอร์พร้อมกันกับจำนวนข้อสอบร่วมเพียงเล็กน้อยถูกต้องมากกว่าโปรแกรม BILOG แต่เมื่อเพิ่มจำนวนข้อสอบร่วมแล้ว การปรับเทียบจะให้ผลใกล้เคียงกัน นอกจากนี้ Beguin และคณะ (2000) และ Beguin และ Hanson (2001 อ้างถึงใน Kim และ Lee, 2006) ได้เปรียบเทียบวิธี Stocking – Lord กับการปรับค่าพร้อมกันโดยการจำลองข้อมูลให้ไม่สอดคล้องกับโมเดล IRT เนื่องจากเป็นแบบพหุมิติ เมื่อกลุ่มผู้สอบเป็นแบบกลุ่มไม่เท่าเทียม และความสามารถของผู้สอบมีความสัมพันธ์กันสูง มาตรวัด (scaling) ที่ใช้ในวิธี Stocking and Lord มีความถูกต้องมากกว่า วิธีปรับค่าพร้อมกัน ซึ่งข้อค้นพบนี้แตกต่างจาก Kim และ Cohen (1998) และ Hanson และ Benguin (2002) ในส่วนที่ข้อมูลจำลองให้สอดคล้องกับโมเดล

จากข้างต้นจะเห็นว่าการปรับค่าพร้อมกันให้ผลที่ดีมากกว่าวิธีอื่นๆ เมื่อข้อมูลสอดคล้องกับโมเดล แต่อาจจะแกร่งน้อยกว่าการปรับค่าแยกกันโดยใช้วิธีโค้งคุณลักษณะ เมื่อมีการฝ่าฝืนข้อตกลงเบื้องต้น ประโยชน์ของการประมาณค่าแยกกันคือ ตรวจสอบค่าประมาณพารามิเตอร์ข้อสอบได้ง่ายขึ้นสำหรับข้อสอบร่วม

จากที่กล่าวมาข้างต้น เป็นการศึกษาการปรับเทียบคะแนนสำหรับแบบสอบที่มีการตรวจให้คะแนน 2 ค่า แต่เมื่อพัฒนาการของการวัดและการประเมินเปลี่ยนแปลงไป รูปแบบของเครื่องมือวัดส่วนใหญ่ที่ใช้ในการประเมินการเรียนรู้ได้พัฒนาอยู่ในรูปแบบของข้อสอบที่ตรวจให้คะแนนหลายค่าหรือข้อสอบผสมระหว่างการตรวจให้คะแนนสองค่ากับการตรวจให้คะแนนหลายค่า ทำให้มีการพัฒนาเทคนิควิธีดำเนินการเชื่อมโยงและการปรับเทียบตามทฤษฎีการตอบสนองข้อสอบ สำหรับใช้กับแบบสอบที่ข้อสอบตรวจให้คะแนนหลายค่า โดยพัฒนาจากวิธีการปรับเทียบสำหรับแบบสอบที่ข้อสอบตรวจ

ให้คะแนนสองค่า จากเดิมกระบวนการปรับเทียบตามวิธีการในทฤษฎีการตอบสนองข้อสอบก่อนดำเนินการปรับเทียบตามวิธีที่กล่าวมาแล้วข้างต้นจะต้องพิจารณาว่าจะเลือกโมเดลการวิเคราะห์ค่าพารามิเตอร์จากโมเดลโลจีสติก 1, 2 หรือ 3 พารามิเตอร์ ซึ่งเป็นโมเดลการวิเคราะห์ค่าพารามิเตอร์สำหรับข้อสอบที่มีการตรวจให้คะแนนสองค่า ดังนั้น เมื่อนำข้อสอบที่ตรวจให้คะแนนหลายค่ามาเทียบคะแนนตามวิธีการในทฤษฎีตอบสนองข้อสอบจึงต้องวิเคราะห์ค่าพารามิเตอร์ในโมเดลของทฤษฎีการตอบสนองข้อสอบที่ตรวจให้คะแนนหลายค่า (Polytomous item response model) ซึ่งโมเดลนี้พัฒนามาจากโมเดล IRT ที่ใช้สำหรับข้อสอบที่ให้คะแนนสองค่า คือ โมเดลโลจีสติก 1, 2 พารามิเตอร์ โมเดล IRT สำหรับแบบสอบที่มีการให้คะแนนหลายค่ามีหลายโมเดล เช่น Grade response model (Samejima, 1997), Nominal response model (Bock, 1972), Partial credit model (Master and Wright, 1997), Generalized partial credit model (Muraki, 1997)

มีการศึกษาเกี่ยวกับการเชื่อมโยงสเกลเพื่อปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมเพียงเล็กน้อย การวิจัยที่ได้ส่วนใหญ่ใช้วิธีการจำลองข้อมูล ในโมเดลเกรด เรสพอนส์ โมเดลพาเชียล เครดิต และโมเดลเจนเนอรัลไลซ์ พาเชียล เครดิต โดยศึกษาผลที่ได้จากวิธีการปรับเทียบคะแนนในองค์ประกอบต่างๆ ที่เกี่ยวข้องกับการปรับเทียบคะแนน เช่น แบบแผนการเชื่อมโยง ความยาวของแบบสอบ ขนาดของกลุ่มตัวอย่าง เป็นต้น โดย เปรียบเทียบระหว่างวิธีการปรับเทียบคะแนนตามทฤษฎีการทดสอบแบบดั้งเดิมกับวิธีการตามทฤษฎีการตอบสนองข้อสอบ ได้แก่ Li, Lissitz และ Yang (1999) ได้เสนอวิธีการเชื่อมโยง Stocking – Lord ในเวอร์ชันที่ขยายสำหรับแบบสอบรูปแบบผสมระหว่างโมเดลการให้คะแนนแบบ 2 ค่า และมากกว่า 2 ค่า คือโมเดล 3PL และ GPCM ตามลำดับ Tate (2000) ได้อธิบายวิธีการปรับเทียบคะแนนแบบ mean and sigma และวิธี Stocking – Lord ในเวอร์ชันที่ขยายสำหรับแบบสอบรูปแบบผสม และมีการศึกษาการใช้ข้อสอบปลายเปิดเพื่อประเมินการปฏิบัติของนักเรียนที่มีความสามารถใกล้เคียงกัน ปรากฏว่าวิธีการเทียบคะแนนอีควิปเปอร์เซ็นไทล์ให้ผลที่เท่าเทียมกับการปรับเทียบคะแนนในโมเดล พาเชียล เครดิต ด้วยวิธีโค้งลักษณะแบบทดสอบ เมื่อข้อสอบมีความยากปานกลาง (Huynh and Ferrara, 1994 : 139) และให้ผลที่คล้ายคลึงกับวิธีการเทียบคะแนนในโมเดลพาเชียล เครดิต ด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน (Hennings, Hirsch and Zhang, 1996 : 6) และได้มีการศึกษาเปรียบเทียบวิธีการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ สำหรับข้อสอบเขียนตอบในโมเดลเกรด เรสพอนส์ ได้แก่ วิธีโค้งลักษณะแบบทดสอบ (TRF) วิธีไคสแควร์ต่ำสุด (MCS) และวิธีการที่ใช้พารามิเตอร์บางค่า หรือทุกค่าไม่พร้อมกันซึ่งปรับสูตรจากข้อมูลที่ได้จากข้อสอบที่ตรวจให้คะแนนสองค่า ได้แก่ วิธีใช้

ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของพารามิเตอร์ความยาก (วิธี mean/sigma : MS) วิธีใช้ค่าเฉลี่ยของพารามิเตอร์ความยากและอำนาจจำแนก (วิธี mean/mean : LH) และวิธีใช้ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของพารามิเตอร์ความยากที่กำหนดน้ำหนัก (WMS) (Linn and others, 1981 อ้างถึงใน Cohen and Kim, 1998 : 118) ปรากฏว่าส่วนใหญ่ทั้ง 5 วิธีให้ผลที่คล้ายคลึงกันเกือบทุกองค์ประกอบที่ศึกษาจากการจำลองข้อมูล (Cohen and Kim. 1998 : 126 – 129) ในการศึกษาส่วนใหญ่ยังคงพิจารณาความสอดคล้องของผลที่ได้เทียบกับผลจากข้อมูลที่ได้จากข้อสอบตรวจให้คะแนนสองค่า

นอกจากการศึกษาเปรียบเทียบการปรับเทียบตามวิธีต่าง ๆ ดังที่ได้กล่าวมาแล้ว ประเด็นสำคัญอย่างหนึ่งสำหรับการปรับเทียบคือ รูปแบบการเก็บรวบรวมข้อมูล จากการศึกษาที่ผ่านมาส่วนใหญ่การเก็บรวบรวมข้อมูลเป็นแบบกลุ่มสุ่ม (Random group design) ซึ่งผู้สอบมีลักษณะเท่าเทียมกัน แบบกลุ่มผู้สอบไม่เท่าเทียมกันโดยใช้ข้อสอบร่วม (Common item non-equivalent group) ผลการวิจัยมีความแตกต่างกันตามเงื่อนไขที่ศึกษา การเก็บรวบรวมข้อมูลแต่ละรูปแบบมีจุดเด่นจุดด้อยแตกต่างกันไป เช่น กลุ่มเท่าเทียมกัน โดยสุ่มกลุ่มผู้สอบทั้ง 2 กลุ่มให้มีลักษณะที่ใกล้เคียงกันมากที่สุด (Equivalent group) ผู้สอบแต่ละกลุ่มทำแบบสอบฉบับเดียวสามารถหลีกเลี่ยงปัญหาเรื่องการเรียนรู้ การฝึกฝนและความเมื่อยล้าจากการทำแบบสอบฉบับแรก ซึ่งเป็นลักษณะที่เกิดขึ้นกับการออกแบบการเก็บรวบรวมข้อมูลแบบผู้สอบกลุ่มเดียว (Single group design) และรูปแบบกลุ่มผู้สอบร่วม (Common person design) ที่ต้องทำแบบสอบสองครั้ง แต่ปัญหาที่เกิดกับกลุ่มเท่าเทียมคือ กลุ่มที่ใช้อาจมีความแตกต่างกันบ้างเพียงเล็กน้อย คือ อาจมีการแจกแจงความสามารถที่แตกต่างกัน และไม่มีข้อมูลที่นำมาใช้ปรับความแตกต่างของกลุ่ม ความแตกต่างที่เกิดขึ้นเพียงเล็กน้อยนี้ย่อมมีผลกระทบต่อความลำเอียงในการปรับเทียบคะแนนได้ แต่สามารถแก้ปัญหาได้ด้วยการสร้างข้อสอบร่วม (common item) ที่มีคุณภาพ คือมีความสัมพันธ์กันในเรื่องเนื้อหาและระดับความยากระหว่างข้อสอบร่วมกับแบบสอบที่ต้องการปรับเทียบคะแนน การเก็บรวบรวมข้อมูลอีกรูปแบบหนึ่งที่ได้รับความนิยมคือ กลุ่มผู้สอบไม่เท่าเทียมกันโดยใช้แบบสอบร่วม (Common item non-equivalent group design) รูปแบบนี้ใช้ผู้สอบมีลักษณะไม่เท่าเทียมกัน ผู้สอบแต่ละกลุ่มทำแบบสอบฉบับเดียว โดยแบบสอบแต่ละฉบับมีข้อสอบร่วมจำนวนหนึ่ง ซึ่งอาจใช้แบบสอบร่วมภายในหรือแบบสอบร่วมภายนอกก็ได้

นอกจากรูปแบบการเก็บรวบรวมข้อมูลแล้ว สิ่งสำคัญอีกประการหนึ่งของการปรับเทียบคะแนนคือ ข้อสอบร่วม (common item) จากการศึกษาการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม ผลการปรับเทียบที่มีข้อสอบที่ตรวจให้คะแนนสองค่า และข้อสอบที่ตรวจให้คะแนนหลายค่าเป็น

ข้อสอบร่วม จะให้ผลการปรับเทียบแม่นยำกว่าการใช้ข้อสอบเพียงรูปแบบใดรูปแบบหนึ่งเป็นข้อสอบร่วม และยังพบอีกว่า ข้อสอบร่วมที่มีจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าลดลง จะมีความแม่นยำในการปรับเทียบมากขึ้นด้วย อาจเนื่องมาจากได้รับอิทธิพลของคะแนนรวมจากข้อสอบที่ตรวจให้คะแนนหลายค่า นั่นคือ ปริมาณข้อสอบร่วมที่เป็นข้อสอบที่ตรวจให้คะแนนสองค่าลดลง จะทำให้คะแนนรวมใกล้เคียงกับคะแนนรวมของข้อสอบที่ตรวจให้คะแนนหลายค่า (Li, Lissitz & Yang, 1999 ; Kim และ Lee, 2006) ดังนั้นการศึกษาวิจัยเกี่ยวกับการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมนอกจากจะพิจารณาสัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่า และข้อสอบที่ตรวจให้คะแนนหลายค่าในแบบสอบทั้งฉบับแล้ว จึงควรพิจารณาสัดส่วนของข้อสอบทั้ง 2 ชนิดที่ประกอบอยู่ในข้อสอบร่วม และคะแนนรวมที่อยู่ในข้อสอบร่วมด้วย เนื่องจากการแบบทดสอบที่มีการนำคะแนนมาปรับเทียบในอดีตนั้น ส่วนใหญ่เป็นข้อสอบที่มีการตรวจให้คะแนนสองค่า (0 และ 1) เมื่อกำหนดจำนวนข้อสอบร่วม 20% ตามแนวคิดของ Angoff (1971) เช่น จำนวนข้อสอบร่วม 20 ข้อ จากข้อสอบทั้งหมด 100 ข้อในแบบสอบ ก็จะมีคะแนนของข้อสอบร่วม 20 คะแนน แต่สำหรับแบบสอบรูปแบบผสม เมื่อมีการนำข้อสอบที่ตรวจให้คะแนน 2 ค่า และมากกว่า 2 ค่า มาเป็นข้อสอบร่วมนั้น ข้อสอบร่วมที่มีคะแนน 20 คะแนน อาจจะไม่ได้มาจากจำนวนข้อสอบ 20 ข้อก็ได้ ถ้ามีการนำข้อสอบที่ตรวจให้คะแนนหลายค่ามาเป็นข้อสอบร่วมด้วย ดังนั้น การศึกษาเกี่ยวกับข้อสอบร่วม นอกจากจะศึกษาสัดส่วนของจำนวนข้อสอบแล้ว จึงควรศึกษาปริมาณคะแนนของข้อสอบแต่ละชนิดที่อยู่ในข้อสอบร่วมด้วย

จากข้อมูลที่กล่าวมาข้างต้นจะเห็นได้ว่าการปรับเทียบคะแนนสำหรับแบบสอบที่ให้คะแนนสองค่าและหลายค่าจะให้ผลที่คล้ายกันอย่างหนึ่งคือ วิธีการปรับเทียบตามแนวทฤษฎีการตอบสนองข้อสอบแบบโค้งคุณลักษณะ (Characteristic Curve) และวิธีการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent Calibration) ให้ผลการปรับเทียบเป็นที่น่าพอใจกว่าวิธีอื่น ซึ่งการปรับเทียบคะแนนสอบสำหรับแบบสอบรูปแบบผสมนั้นยังมีการศึกษาเปรียบเทียบวิธีการปรับเทียบไม่มากนัก

ผู้วิจัยจึงสนใจศึกษาเทคนิคการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบสำหรับแบบสอบรูปแบบผสม 2 วิธีคือ วิธีการปรับเทียบตามแนวทฤษฎีการตอบสนองข้อสอบแบบโค้งคุณลักษณะ (Characteristic Curve) และวิธีการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent Calibration) เพื่อศึกษาคุณภาพของการปรับเทียบคะแนนทั้ง 2 วิธีเปรียบเทียบกันในเงื่อนไขของสัดส่วนของข้อสอบทั้ง 2 แบบในแบบสอบทั้งฉบับ สัดส่วนของข้อสอบทั้ง 2 รูปแบบในข้อสอบร่วมโดยมีข้อสอบร่วมอย่างน้อยร้อยละ 20 ของข้อสอบทั้งหมดตามแนวคิดของ Angoff คะแนนรวมในข้อสอบ

ร่วม รูปแบบการเก็บรวบรวมข้อมูล โดยการจำลองข้อมูลการตอบสำหรับแบบสอบที่มีรูปแบบผสม เนื่องจากเป็นเทคนิคอย่างหนึ่งที่ทำให้ได้ข้อมูลของค่าพารามิเตอร์ที่แท้จริง ทั้งพารามิเตอร์ข้อสอบ และพารามิเตอร์ผู้สอบ เพื่อใช้ในการศึกษาการปรับเทียบคะแนนในเงื่อนไขต่าง ๆ แม้ว่าในการศึกษาที่ผ่านยังไม่มีผลงานวิจัยที่ชัดเจนว่าควรมีการสร้างการจำลองข้อมูลโดยการทำซ้ำ (Replication) ทั้งหมดกี่ครั้ง เพื่อให้ผลที่เกิดขึ้นมีความเที่ยงมากที่สุด แต่จากการศึกษาของ Harwell, Hsu และ Kirisci (1996) พบว่าถ้าศึกษาโดยใช้โมเดล IRT เป็นฐานควรมีการทำซ้ำอย่างน้อย 20 ครั้ง เมื่อจำนวนกลุ่มตัวอย่างมีตั้งแต่ 1000 คนขึ้นไป และจากการศึกษางานวิจัยที่ทำการศึกษาการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมโดยการจำลองข้อมูล จะมีการทำซ้ำหลายค่าแตกต่างกันไป Tate (2000) จำลองข้อมูลการตอบตามโมเดล 3 PL และโมเดล GR ด้วยโปรแกรม RESGEN (Muraki, 1998) โดยมีกลุ่มตัวอย่าง 2000 คน มีการทำซ้ำ 100 ครั้ง และ Kim (2004) และ Kim และ Lee (2006) จำลองข้อมูลการตอบตามโมเดล 3PL และ GPC กับกลุ่มตัวอย่างจำนวน 500 และ 3000 คน มีการทำซ้ำ 100 ครั้ง และพบว่าการปรับเทียบคะแนนเกิดความคลาดเคลื่อนในการปรับเทียบน้อย เมื่อกลุ่มตัวอย่างมีขนาดมากขึ้น ดังนั้นการศึกษาครั้งนี้ ผู้วิจัยได้นำผลการศึกษาจากข้างต้นมาพิจารณาใช้ในการจำลองข้อมูลแบบสอบรูปแบบผสมตามการศึกษาของ Tate (2000), Kim (2004) และ Kim และ Lee (2006) ที่มีการทำซ้ำ 100 ครั้ง และพิจารณาผลการศึกษาของ Harwell, Hsu และ Kirisci (1996) ที่ว่าหากกลุ่มตัวอย่าง 1000 คนขึ้นไป สามารถทำซ้ำได้เพียง 20 ครั้ง ผู้วิจัยจึงศึกษาโดยใช้การจำลองข้อมูลที่มีการทำซ้ำ 100 ครั้ง เพื่อให้สอดคล้องกับการจำลองข้อมูลสำหรับแบบสอบรูปแบบผสมตามการศึกษาของ Tate (2000), Kim (2004) และ Kim และ Lee (2006) และพิจารณาจำลองข้อมูลโดยมีกลุ่มตัวอย่างเพียง 1000 คน

ส่วนโมเดลที่ผู้วิจัยเลือกใช้ในการจำลองข้อมูลสำหรับแบบสอบรูปแบบผสมระหว่าง โมเดลตามทฤษฎีการตอบสนองข้อสอบที่มีการตรวจให้คะแนนแบบ 2 ค่า แบบ 3 พารามิเตอร์ เนื่องจากลักษณะของแบบสอบที่มีการตรวจให้คะแนนแบบสองค่า ส่วนใหญ่เป็นข้อสอบชนิดหลายตัวเลือกที่มีการตรวจให้คะแนนเป็นถูกได้ 1 และผิดได้ 0 ซึ่งข้อสอบชนิดนี้มีคำตอบให้ผู้สอบได้เลือก ฉะนั้นแม้ผู้สอบไม่มีความรู้ก็สามารถที่จะเดาเพื่อเลือกคำตอบเหล่านั้นได้ ดังนั้นโมเดลโลจีสติกแบบ 3 พารามิเตอร์ซึ่งประกอบด้วยพารามิเตอร์ความยาก (b) พารามิเตอร์อำนาจจำแนก (a) และพารามิเตอร์การเดา (c) จึงน่าจะเป็นโมเดล ที่มีความสอดคล้องกับลักษณะของข้อสอบที่ได้รับความนิยมในการทดสอบความรู้ของผู้สอบมากที่สุด ส่วนโมเดลตามทฤษฎีการตอบสนองข้อสอบที่มีการตรวจให้คะแนนแบบหลายค่า ผู้วิจัยเลือกโมเดล Generalized Partial Credit : GPC ที่พัฒนาโดย Muraki (1992, 1993) เพื่อให้สอดคล้องกับ

งานวิจัยของ Kim (2004) และ Kim และ Lee (2006) ที่ได้จำลองข้อมูลเพื่อเปรียบเทียบวิธีการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ ที่มีการปรับค่าพารามิเตอร์แยกกัน 4 วิธี

สำหรับคุณภาพของการปรับเทียบคะแนนสามารถพิจารณาได้หลายแนวทาง ซึ่งในการศึกษาคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่พบส่วนใหญ่พิจารณาจากความคลาดเคลื่อนที่เกิดจากการปรับเทียบทั้งในรูปแบบของรากที่สองของคะแนนเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (Root mean square error : RMSE) ค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (Mean square error) หรือความลำเอียง (Bias) ที่เกิดจากการปรับเทียบ เป็นต้น ซึ่งความคลาดเคลื่อนเหล่านี้จะมีพื้นฐานมากจากความแตกต่างระหว่างค่าที่ประมาณได้กับค่าที่แท้จริง และเกณฑ์อีกลักษณะหนึ่งที่ถูกนำมาใช้ในการประเมินคุณภาพการปรับเทียบคือ เกณฑ์สำหรับเปรียบเทียบคุณภาพการปรับเทียบที่พิจารณาจากเกณฑ์ค่าคงที่ของสเกลการแปลงคะแนน (Scale Transformation Constants : STC) มีพื้นฐานมาจากความแตกต่างระหว่างค่าคงที่ของการแปลงคะแนนที่ประมาณค่าได้ (estimated transformation) และค่าคงที่ของการแปลงสเกลที่แท้จริง (true transformation) และเกณฑ์โค้งคุณลักษณะลำดับขั้นการตอบ (Category Response Curve Criterion : CRC) ซึ่งเกณฑ์โค้งคุณลักษณะลำดับขั้นการตอบ (CRC) มีพื้นฐานมาจากความแตกต่างระหว่างโค้งลำดับขั้นการตอบที่ประมาณค่าได้ (estimated category response curve) กับโค้งลำดับขั้นการตอบที่แท้จริง (true category response curve) บนแบบสอบฉบับเทียบคะแนน (new form) ซึ่งทั้งเกณฑ์ STC และ CRC เป็นค่าที่แสดงถึงค่าเฉลี่ยของค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (Mean square errors) ทั้งหมดที่ได้จากความแตกต่างระหว่างโค้งคุณลักษณะลำดับขั้นที่ประมาณค่าได้ กับโค้งคุณลักษณะลำดับขั้นที่แท้จริงสำหรับข้อสอบข้อที่ j (Hanson และ Beguin, 2002 ; Kim, 2004 ; Kim & Lee, 2006) สำหรับการศึกษาคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมครั้งนี้ จึงพิจารณาจากค่า MSE ที่ได้จากเกณฑ์ CRC เพื่อให้สอดคล้องกับการศึกษาของ Hanson และ Beguin (2002), Kim (2004), และ Kim และ Lee (2006) ผลที่ได้จากการศึกษาครั้งนี้จะเป็นประโยชน์เพื่อประกอบการตัดสินใจในการเลือกใช้วิธีการเชื่อมโยงคะแนนที่เหมาะสมและสอดคล้องกับสถานการณ์ที่ต้องการศึกษามากที่สุด

## คำถามวิจัย

สถานการณ์การทดสอบในปัจจุบันมีการใช้แบบสอบรูปแบบผสมมากยิ่งขึ้น เพื่อประเมินความสามารถของผู้สอบในทักษะที่หลากหลาย และมีการจัดการทดสอบต่างเวลามากยิ่งขึ้น หากจำเป็นต้องนำเทคนิคการปรับเทียบคะแนนมาใช้เพื่อให้เกิดความยุติธรรมแล้ว เทคนิคการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ ทั้ง 2 วิธีคือ วิธีโค้งคุณลักษณะ และวิธีปรับค่าพารามิเตอร์พร้อมกัน คุณภาพของวิธีการปรับเทียบแต่ละวิธีภายใต้แต่ละเงื่อนไขที่ศึกษา คือ สัดส่วนของข้อสอบทั้งสองชนิดในแบบสอบ สัดส่วนของข้อสอบทั้งสองชนิดในข้อสอบร่วม แบบแผนการเก็บรวบรวมข้อมูล มีค่า MSE เป็นเท่าใด และคุณภาพของวิธีการปรับเทียบคะแนนที่แตกต่างกัน 2 วิธี คือ วิธีโค้งคุณลักษณะ และวิธีปรับค่าพารามิเตอร์พร้อมกัน ภายใต้แต่ละเงื่อนไขที่ศึกษา จะมีคุณภาพแตกต่างกันหรือไม่ และเงื่อนไขใดมีคุณภาพดีกว่า

## วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์หลักเพื่อศึกษาผลของการปรับเทียบคะแนนตามแนวทฤษฎีการตอบสนองข้อสอบ สำหรับแบบสอบที่มีรูปแบบผสมระหว่างการตรวจให้คะแนนแบบสองค่าตามโมเดล IRT โลจีสติกแบบ 3 PL และการตรวจให้คะแนนแบบหลายค่าตามโมเดล GPC โดยมีวัตถุประสงค์เฉพาะดังนี้

1. เพื่อวิเคราะห์คุณภาพการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ (Characteristic Curve) และวิธีการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent Calibration) เมื่อแบบแผนการเก็บรวบรวมข้อมูล และสัดส่วนของข้อสอบทั้ง 2 ชนิดในแบบสอบรูปแบบผสม และสัดส่วนของข้อสอบทั้ง 2 ชนิดในข้อสอบร่วมของแบบสอบรูปแบบผสม ต่างกัน

2. เพื่อเปรียบเทียบคุณภาพผลการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ (Characteristic Curve) และวิธีการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent Calibration) เมื่อแบบแผนการเก็บรวบรวมข้อมูล และสัดส่วนของข้อสอบทั้ง 2 ชนิดในแบบสอบรูปแบบผสม และสัดส่วนของข้อสอบทั้ง 2 ชนิดในข้อสอบร่วมของแบบสอบรูปแบบผสม ต่างกันโดยพิจารณาจากดัชนี MSE

## สมมติฐานการวิจัย

การศึกษาวิธีการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมมีผู้ศึกษาทั้งการปรับเทียบตามแนวทฤษฎีดั้งเดิม และการปรับเทียบตามแนวทฤษฎีตอบสนองข้อสอบ Bastari (2000) ศึกษาเปรียบเทียบวิธีการปรับเทียบคะแนนตามแนวทฤฎีตอบสนองข้อสอบระหว่างวิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน โดยศึกษาในเงื่อนไขความยาวของแบบสอบ สัดส่วนของข้อสอบทั้ง 2 รูปแบบในแบบสอบ ขนาดกลุ่มตัวอย่าง และการแจกแจงของกลุ่มตัวอย่าง ผลการวิจัยพบว่าวิธีการปรับค่าพารามิเตอร์พร้อมกันมีประสิทธิภาพมากกว่าวิธีโค้งคุณลักษณะ เมื่อแบบสอบมีความยาวมากขึ้น สัดส่วนข้อสอบที่มีการตรวจให้คะแนนสองค่ามีปริมาณมากขึ้น จำนวนข้อสอบร่วมที่เพิ่มขึ้น (RMSE ของข้อสอบร่วม 20% ต่ำกว่า RMSE ของข้อสอบร่วม 10%) ขนาดกลุ่มตัวอย่างขนาดใหญ่ และกลุ่มผู้สอบมีการแจกแจงความสามารถที่ใกล้เคียงกัน ซึ่งสอดคล้องกับงานวิจัยของ Tate (2000) ได้ศึกษาวิธีการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบด้วยวิธี moment 2 วิธี และวิธีโค้งคุณลักษณะตามแนวของ Stocking – Lord พบว่า จำนวนข้อสอบร่วมที่มีปริมาณมากขึ้นจะทำให้สัมประสิทธิ์การปรับเทียบคะแนนมีความแม่นยำยิ่งขึ้น

จากงานวิจัยของ Kim (2004) และ Kim และ Lee (2006) ที่ได้ศึกษาเปรียบเทียบการเทียบคะแนนตามแนวทฤษฎีการตอบสนองข้อสอบด้วยวิธี moment 2 วิธี และวิธีโค้งคุณลักษณะ 2 วิธี (พิจารณาค่า MSE ที่ได้จากเกณฑ์ CRC ) ได้ผลสอดคล้องกับงานวิจัยของ Tate (2000) ที่พบว่า วิธีโค้งคุณลักษณะ ให้ผลการปรับเทียบคะแนนเป็นที่น่าพอใจกว่าวิธี moment ทั้งสองวิธี และพบว่าจำนวนกลุ่มตัวอย่างขนาดใหญ่จะทำให้ผลการปรับเทียบมีความแม่นยำยิ่งขึ้น สำหรับการปรับเทียบโดยใช้ข้อสอบร่วม จะมีความคลาดเคลื่อนสูงขึ้นเมื่อใช้ข้อสอบเพียงชนิดใดชนิดหนึ่งเป็นข้อสอบร่วม และถ้าข้อสอบร่วมที่ประกอบด้วยข้อสอบแบบเลือกตอบและเขียนตอบจะเกิดความความเคลื่อนสูงขึ้นเมื่อจำนวนข้อสอบเลือกตอบที่อยู่ในข้อสอบร่วมมีจำนวนเพิ่มขึ้น

นอกจากนี้ยังมีข้อพิจารณาที่น่าสนใจเกี่ยวกับประสิทธิภาพสัมพัทธ์ของการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent Calibration) และการปรับค่าพารามิเตอร์แยกกัน (Seperated Calibration) กับวิธีการแปลงสเกลที่แตกต่างกัน โดยทั่วไป การปรับค่าพารามิเตอร์พร้อมกัน จะให้ผลที่น่าพอใจมากกว่าการปรับค่าพารามิเตอร์แยกกัน ภายใต้เงื่อนไขที่ศึกษาหลาย ๆ เงื่อนไข เนื่องมาจากมีความเชื่อเกี่ยวกับการปรับค่าพารามิเตอร์พร้อมกันว่าสามารถทำให้ได้สารสนเทศ

ที่เป็นประโยชน์อย่างสมบูรณ์ และอาจจะขจัดความคลาดเคลื่อนบางอย่างที่เกิดขึ้นในการปรับเทียบได้อย่างมีศักยภาพ ซึ่งความไม่ถูกต้องในกระบวนการแปลงสเกลอาจจะเกิดขึ้นในการปรับค่าพารามิเตอร์แยกกัน Kim และ Cohen (1998) ใช้กระบวนการจำลองข้อมูลสำหรับการตอบที่มีการให้คะแนนแบบสองค่า เพื่อเปรียบเทียบการปรับค่าพารามิเตอร์แยกกันด้วยวิธี Stocking – Lord กับการปรับค่าพารามิเตอร์พร้อมกัน โดยใช้ซอฟต์แวร์ที่ต่างกัน ผลการศึกษาพบว่า ชุดข้อสอบร่วมที่มีจำนวนข้อน้อยๆ ใช้วิธีการปรับค่าพารามิเตอร์พร้อมกันจะให้ผลที่ถูกต้องมากกว่าวิธีการปรับค่าพารามิเตอร์แยกกัน ด้วยวิธี Stocking – Lord แต่ถ้าจำนวนข้อสอบร่วมมีขนาดใหญ่วิธีการปรับค่าพารามิเตอร์แยกกันและการปรับค่าพารามิเตอร์พร้อมกันจะให้ผลของการปรับเทียบสเกลคล้ายคลึงกัน Hanson และ Beguin (2002) ได้จำลองข้อมูลการตอบแบบสองค่าสำหรับเปรียบเทียบประสิทธิภาพสัมพัทธ์ของการปรับค่าพารามิเตอร์แยกกัน (mean/mean, mean/sigma, Heabara, และ Stocking – Lord) กับวิธีการปรับค่าพารามิเตอร์พร้อมกัน ผลการศึกษาพบว่า โดยทั่วไปวิธีการปรับค่าพารามิเตอร์พร้อมกันให้ความคลาดเคลื่อนต่ำกว่าวิธีการปรับค่าพารามิเตอร์แยกกัน แม้ว่าจะไม่ใช้ผลที่ครอบคลุมทั้งหมดก็ตาม เนื่องจากในการปรับค่าพารามิเตอร์แยกกันด้วยวิธีโค้งคุณลักษณะ 2 วิธี ให้ผลการแปลงสเกลที่ถูกต้องมากกว่าวิธี mean/mean และ mean/sigma และ Kim และ Cohen (2002) ได้จำลองข้อมูลผลการตอบที่มีการให้คะแนนแบบหลายค่าภายใต้โมเดล GRM และเปรียบเทียบการปรับค่าพารามิเตอร์แยกกันกับวิธี Stocking – Lord และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ผลการศึกษาพบว่าวิธีการปรับค่าพารามิเตอร์พร้อมกันให้ผลที่คงที่ และให้ที่มีความถูกต้องมากกว่าวิธีการแปลงสเกลแบบแยกกันเพียงเล็กน้อย นอกจากนี้ Kim และ Lee (2006) ได้ขยายวิธีการแปลงสเกล 4 วิธี ได้แก่ mean/mean, mean/sigma, Heabara, และ Stocking – Lord ในการใช้กับแบบสอบรูปแบบผสม พบว่า วิธีโค้งคุณลักษณะโดยทั่วไปจะให้ผลที่มีความถูกต้องในการเชื่อมโยงคะแนนมากกว่าวิธีการเชื่อมโยงแบบจุด (Moment Methods) เมื่อให้ปัจจัยอื่นๆ คงที่ ส่วนการเชื่อมโยงโดยใช้กลุ่มเท่าเทียมจะมีความคลาดเคลื่อนในการเชื่อมโยงน้อยกว่าการปรับเทียบกับกลุ่มไม่เท่าเทียม และเมื่อการเชื่อมโยงนั้นมีข้อสอบทั้งแบบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนหลายค่าเป็นข้อสอบร่วม โดยทั่วไปจะมีผลการปรับเทียบที่ถูกต้องมากกว่าการปรับเทียบโดยใช้ข้อสอบที่ตรวจให้คะแนนสองค่า หรือข้อสอบที่ตรวจให้คะแนนหลายค่าอย่างใดอย่างหนึ่งเป็นข้อสอบร่วมเพียงอย่างเดียว

จากข้อค้นพบที่ได้จึงนำไปสู่การตั้งสมมติฐานเพื่อตอบคำถามวิจัย ดังนี้

1. วิธีการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะที่มีแบบแผนการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียมน่าจะได้ผลการปรับเทียบที่เกิดความคลาดเคลื่อนต่ำกว่าวิธีการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะที่มีแบบแผนการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม

2. วิธีการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกันที่มีแบบแผนการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียมน่าจะได้ผลการปรับเทียบที่มีความคลาดเคลื่อนใกล้เคียงกับวิธีการปรับเทียบคะแนนด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกันที่มีแบบแผนการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม และมีความคลาดเคลื่อนต่ำกว่าการปรับเทียบด้วยวิธีโค้งคุณลักษณะ

3. วิธีการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกันน่าจะได้ผลการปรับเทียบที่เกิดความคลาดเคลื่อนต่ำกว่าวิธีปรับด้วยวิธีโค้งคุณลักษณะเมื่อมีแบบแผนการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม

4. วิธีการปรับเทียบด้วยวิธีโค้งคุณลักษณะและวิธีการปรับค่าพารามิเตอร์พร้อมกัน น่าจะให้ผลการปรับเทียบที่มีความคลาดเคลื่อนต่ำเมื่อสัดส่วนของข้อสอบทั้ง 2 ชนิด มีจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าในแบบสอบทั้งฉบับมีจำนวนมากขึ้น

5. วิธีการปรับเทียบด้วยวิธีโค้งคุณลักษณะและวิธีการปรับค่าพารามิเตอร์พร้อมกัน น่าจะให้ผลการปรับเทียบที่มีความคลาดเคลื่อนต่ำเมื่อข้อสอบร่วมที่ประกอบด้วยข้อสอบทั้ง 2 ชนิด มีข้อสอบที่ตรวจให้คะแนนสองค่าที่อยู่ในข้อสอบร่วมจำนวนลดลง

## ขอบเขตของการวิจัย

1. สถานการณ์การจำลองข้อมูลครั้งนี้ เป็นการจำลองข้อมูลให้สอดคล้องกับโมเดลการตอบข้อสอบที่มีการตรวจให้คะแนนแบบสองค่า ตามโมเดล IRT โลจีสติกแบบ 3PL และโมเดลการตอบข้อสอบที่มีการตรวจให้คะแนนหลายค่า ตามโมเดล GPC 6 ลำดับขั้น

2. ประชากรข้อสอบที่ใช้ในการวิจัยครั้งนี้เป็นค่าพารามิเตอร์คุณลักษณะข้อสอบที่สร้างขึ้นด้วยฟังก์ชั่นคณิตศาสตร์ โดยการจำลองแบบค่าที่เป็นไปได้ของข้อสอบตามแนวทฤษฎีการตอบสนองข้อสอบจากโมเดลโลจิสติก 3 พารามิเตอร์ ซึ่งประกอบด้วย

2.1 ค่าความยากของข้อสอบ (b) ช่วงที่ยอมรับได้ว่าเป็นข้อสอบที่มีคุณภาพ คือ มีค่า b อยู่ระหว่าง -2.5 ถึง +2.5 โดยเฉลี่ยพารามิเตอร์ความยากของข้อสอบอยู่ในระดับปานกลาง คือ ค่าเฉลี่ยเท่ากับ 0 มีการแจกแจงแบบปกติ

2.2 ค่าอำนาจจำแนกของข้อสอบ (a) ที่ช่วงที่ยอมรับได้ว่าเป็นข้อสอบที่มีคุณภาพ คือ มีค่า a มีค่าอยู่ระหว่าง +0.5 ถึง +2.5

2.3 ค่าการเดา (c) ที่ช่วงที่ยอมรับได้ว่าเป็นข้อสอบที่มีคุณภาพ คือ มีค่า c ไม่เกิน 0.3

สำหรับโมเดลการตอบข้อสอบที่มีการตรวจให้คะแนนหลายค่า เป็นโมเดล Generalized Partial Credit ที่มีรายการคำตอบ 6 ลำดับขั้น (คะแนน 0 - 5 คะแนน)

3. ตัวแปรในการวิจัยประกอบด้วย

**ตัวแปรอิสระ ได้แก่**

1. สัดส่วนของคะแนนระหว่างข้อสอบที่ตรวจให้คะแนนแบบสองค่า (Dichotomous : DS) กับข้อสอบที่ตรวจให้คะแนนหลายค่า (Polytomous : PS) ตามโมเดล GPC ที่มีรายการคำตอบ 6 ลำดับขั้น จำนวน 4 สัดส่วน ซึ่งสัดส่วนน้ำหนักคะแนนระหว่างข้อสอบ DS และ PS ได้แก่ 50 : 50, 60 : 40, 70 : 30 และ 80 : 20 ตามลำดับ ข้อสอบชนิด DS ให้คะแนนแบบตอบถูกได้ 1 และตอบผิดได้ 0 เมื่อคิดเป็นจำนวนข้อสอบแล้วจะได้ DS : 50 ข้อ, 60 ข้อ, 70 ข้อ, และ 80 ข้อ ตามลำดับ ส่วนข้อสอบแบบ PS มีรายการคำตอบ 6 ลำดับขั้น (คะแนนเต็ม 5 คะแนน) คิดเป็นจำนวนข้อสอบแล้วจะได้ PS : 10 ข้อ, 8 ข้อ, 6 ข้อ, และ 4 ข้อ ตามลำดับ จากสัดส่วนของน้ำหนักคะแนน และคิดเป็นสัดส่วนจำนวนข้อสอบ แสดงดังต่อไปนี้

| **สัดส่วนน้ำหนักคะแนน** | | **สัดส่วนข้อสอบ** | | **รวมจำนวนข้อสอบ** |
|---|---|---|---|---|
| 1.1 DS 50% : PS 50% | → | DS 50 ข้อ : PS 10 ข้อ | → | 60 ข้อ |
| 1.2 DS 60% : PS 40% | → | DS 60 ข้อ : PS 8 ข้อ | → | 68 ข้อ |
| 1.3 DS 70% : PS 30% | → | DS 70 ข้อ : PS 6 ข้อ | → | 76 ข้อ |
| 1.4 DS 80% : PS 20% | → | DS 80 ข้อ : PS 4 ข้อ | → | 84 ข้อ |

2. ข้อสอบร่วม ผู้วิจัยสนใจศึกษาข้อสอบร่วม 2 มิติด้วยกันคือ จำนวนข้อสอบทั้งสองชนิดในข้อสอบร่วม และจำนวนคะแนนรวมระหว่างข้อสอบทั้งสองชนิดในข้อสอบร่วม โดยทั้ง 2 มิติ ผู้วิจัยใช้ทั้งข้อสอบที่ตรวจให้คะแนนสองค่าและตรวจให้คะแนนหลายค่าเป็นข้อสอบร่วม เพื่อให้มีความเป็นตัวแทนแบบสอบทั้งฉบับ (Kolen และ Brennan, 2004)

2.1 ข้อสอบร่วมในมิติของจำนวนข้อสอบ เป็นการศึกษาสัดส่วนของข้อสอบทั้งสองชนิดในข้อสอบร่วมที่แตกต่างกันไป 3 ระดับ เพื่อศึกษาแนวโน้มของผลการปรับเทียบคะแนนว่าเมื่อจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าในข้อสอบร่วมลดลง จะทำให้ผลการปรับเทียบมีความแม่นยำมากขึ้นหรือไม่ การกำหนดจำนวนข้อสอบร่วมได้แนวทางจากแนวคิดของ Angoff (1971) ที่ว่าข้อสอบร่วมควรมีจำนวนอย่างน้อย 20% ของข้อสอบทั้งหมด ผู้วิจัยจึงกำหนดจำนวนของข้อสอบร่วมโดยพิจารณาปริมาณข้อสอบจากแบบสอบในข้อที่ 1 ทั้ง 4 ฉบับ คือ 60, 68, 76 และ 84 ข้อตามลำดับ แล้วจึงกำหนดสัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่ากับข้อสอบที่ตรวจให้คะแนนหลายค่าในชุดของข้อสอบร่วม แสดงดังนี้

| จำนวนข้อสอบในแบบสอบ | สัดส่วนของข้อสอบ | จำนวนข้อสอบร่วม (20%) |
|---|---|---|
| 60 | DS 50 : PS 10 | 12 (DS 10 : PS 2) |
| 68 | DS 60 : PS 8 | 14 (DS 12 : PS 2) |
| 76 | DS 70 : PS 6 | 15 (DS 14 : PS 1) |
| 84 | DS 80 : PS 4 | 17 (DS 16 : PS 1) |

จากข้อมูลดังกล่าวข้างต้นผู้วิจัยกำหนดจำนวนข้อสอบร่วมขั้นต่ำ (20%) ที่สามารถนำไปใช้ได้กับแบบสอบทั้ง 4 ฉบับ คือ จำนวนข้อสอบร่วม 17 ข้อ (ประกอบด้วยข้อสอบทั้งที่ตรวจให้คะแนนสองค่าและตรวจให้คะแนนหลายค่า) และกำหนดระดับของข้อสอบที่ตรวจให้คะแนนสองค่าที่ประกอบอยู่ในข้อสอบร่วมให้แตกต่างกัน 3 ระดับ โดยให้ข้อสอบที่ตรวจให้คะแนนหลายค่ามีจำนวนคงที่ เพื่อตรวจสอบว่าจำนวนข้อสอบที่ตรวจให้คะแนนสองค่า (DS) ที่แตกต่างกันจะมีผลต่อความคลาดเคลื่อนในการปรับเทียบหรือไม่ สัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่า และข้อสอบที่ตรวจให้คะแนนหลายค่าในข้อสอบร่วมกำหนด 3 ระดับ ดังนี้

2.1.1 DS 15 ข้อ : PS 2 ข้อ

2.1.2 DS 20 ข้อ : PS 2 ข้อ

2.1.3 DS 25 ข้อ : PS 2 ข้อ

2.2 ข้อสอบร่วมในมิติของคะแนน จากการศึกษาการปรับเทียบคะแนนที่ผ่านมาส่วนใหญ่เป็นการปรับเทียบคะแนนสำหรับแบบสอบที่ตรวจให้คะแนนสองค่า Angoff (1971) กล่าวว่า ข้อสอบร่วม

ควรมีจำนวนอย่างน้อย 20% ของข้อสอบทั้งหมด ซึ่งถ้าเป็นข้อสอบชนิดที่ตรวจให้คะแนนสองค่า (0 และ 1) จำนวนข้อสอบร่วมก็คือจำนวนคะแนนนั่นเอง เช่น 20% ของข้อสอบทั้งหมดก็คือคะแนน 20 คะแนน (ข้อสอบ 100 ข้อ 100 คะแนน) และจากผลการวิจัยของ Li, Lissitz และ Yang (1999) ; Kim และ Lee (2006) ที่พบว่ายิ่งจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าในข้อสอบร่วมลดลง จะทำให้ผลการปรับเทียบคะแนนมีความแม่นยำยิ่งขึ้น อาจเนื่องมาจากได้รับอิทธิพลของคะแนนรวมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่ากับข้อสอบที่ตรวจให้คะแนนหลายค่า ดังนั้นผู้วิจัยจึงศึกษาสัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนหลายค่า ในมิติของคะแนนรวมในแบบสอบทั้งฉบับ โดยกำหนดให้คะแนนรวมของข้อสอบร่วมเท่ากับ 20 คะแนน (ร้อยละ 20 ของคะแนนเต็มทั้งฉบับ) และกำหนดสัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนหลายค่าในข้อสอบร่วมแตกต่างกันไป 3 ระดับ นั่นคือ กำหนดให้ข้อสอบที่ตรวจให้คะแนนสองค่าในข้อสอบร่วมมีคะแนนรวมมากกว่า เท่ากับ และน้อยกว่า คะแนนรวมของข้อสอบที่ตรวจให้คะแนนหลายค่าในข้อสอบร่วม เพื่อศึกษาผลของการปรับเทียบคะแนน สัดส่วนของข้อสอบทั้งสองชนิดในข้อสอบร่วมดังกล่าว แสดงดังนี้

2.2.1 DS 15 ข้อ (15 คะแนน) : PS 1 ข้อ (5 คะแนน) รวม 20 คะแนน

2.2.2 DS 10 ข้อ (10 คะแนน) : PS 2 ข้อ (10 คะแนน) รวม 20 คะแนน

2.2.3 DS 5 ข้อ (5 คะแนน) : PS 3 ข้อ (15 คะแนน) รวม 20 คะแนน

3. แบบแผนการเก็บรวบรวมข้อมูล

3.1 กลุ่มเท่าเทียมโดยใช้ข้อสอบร่วม (Equivalent group with common items)

3.2 กลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม (Non – equivalent group with common items)

4. วิธีการปรับเทียบคะแนน 2 วิธีคือ

4.1 วิธีโค้งคุณลักษณะ

4.2 การปรับค่าพารามิเตอร์พร้อมกัน

**ตัวแปรตาม** คือ ค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (Mean Square Error : MSE)

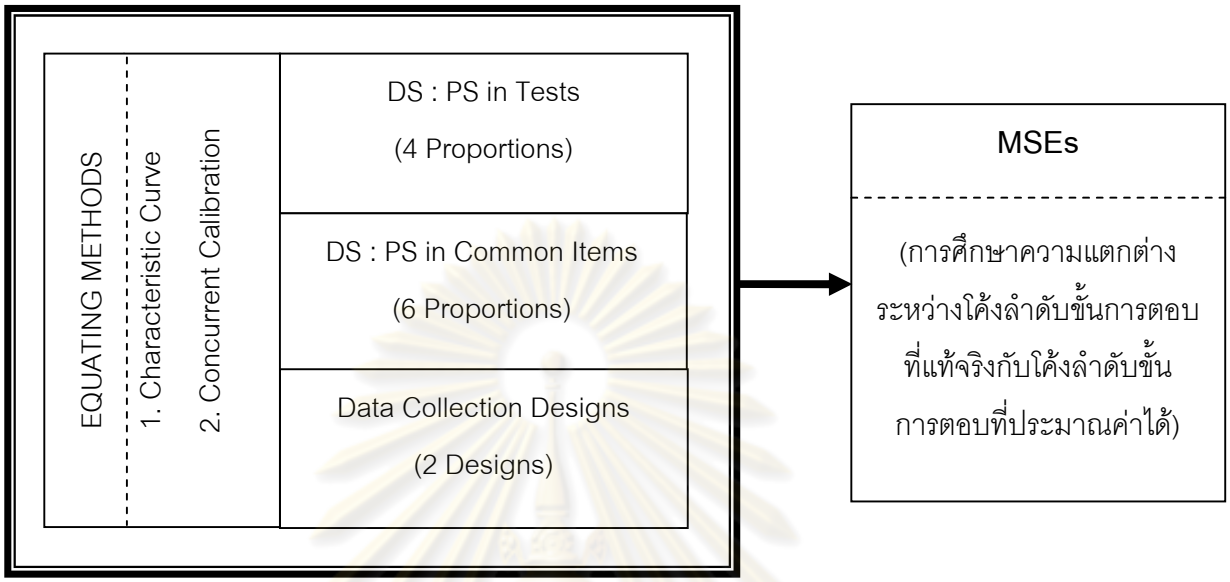
4. สำหรับการเปรียบเทียบคุณภาพของการปรับเทียบคะแนนในการวิจัยครั้งนี้ พิจารณาจากค่า MSE ที่มาจากเกณฑ์โค้งคุณลักษณะลำดับขั้นการตอบ (Category response curve : CRC)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

โดยมีพื้นฐานมาจากความแตกต่างระหว่างโค้งลำดับขั้นการตอบที่ประมาณค่าได้ กับโค้งลำดับขั้นการตอบที่เป็นค่าที่แท้จริง เพื่อให้สอดคล้องกับงานวิจัยของ Hanson และ Beguin (2002), Kim (2004), และ Kim และ Lee (2006) ที่ได้ศึกษาเปรียบเทียบคุณภาพการปรับเทียบคะแนนด้วยวิธีการปรับเทียบตามทฤษฎีการตอบสนองข้อสอบ ระหว่างการปรับค่าพารามิเตอร์แยกกัน และการปรับค่าพารามิเตอร์พร้อมกัน เกณฑ์ที่ใช้นี้เป็นค่า MSE จากการทำซ้ำ 100 ครั้ง โดยเปรียบเทียบค่าที่ได้ ถ้าเงื่อนไขที่ศึกษาใดมีค่า MSE น้อยกว่าแสดงว่ามีคุณภาพในการปรับเทียบคะแนนสูงกว่า หรือความคลาดเคลื่อนในการปรับเทียบคะแนนต่ำกว่า

## กรอบแนวคิดสำหรับการวิจัย

ผู้วิจัยศึกษาคุณภาพของการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบสำหรับแบบสอบรูปแบบผสม 2 วิธีคือ วิธีโค้งคุณลักษณะ และวิธีปรับค่าพารามิเตอร์พร้อมกัน ในเงื่อนไขที่แตกต่างกัน ได้แก่ สัดส่วนของข้อสอบที่ตรวจให้คะแนน 2 ค่าและตรวจให้คะแนนหลายค่าในแบบสอบรูปแบบผสม โดยข้อสอบที่ตรวจให้คะแนน 2 ค่าเมื่อตอบถูกได้ 1 คะแนน ตอบผิดได้ 0 คะแนน และข้อสอบที่ตรวจให้คะแนนหลายค่ามีคะแนนระหว่าง 0 – 5 คะแนน มีจำนวน 4 สัดส่วนคือ 50%: 50%, 60%: 40%, 70% : 30%, และ 80% : 20% ตามลำดับ (คิดเป็นจำนวนสัดส่วนข้อสอบ DS : PS ได้แก่ 50 : 10, 60 : 8, 70: 6, และ 80 : 4 ตามลำดับ) สัดส่วนระหว่าง DS : PS ในข้อสอบร่วม เมื่อกำหนดตามจำนวนข้อ 3 สัดส่วน คือ 15 : 2, 20 : 2 และ 25 : 2 ตามลำดับ ข้อสอบร่วมเมื่อพิจารณาจากคะแนนรวมของข้อสอบร่วมร้อยละ 20 ของคะแนนเต็ม โดยพิจารณาจากข้อสอบที่ตรวจให้คะแนนสองค่าในข้อสอบร่วมมีค่ามากกว่า เท่ากับ และน้อยกว่า คะแนนรวมของข้อสอบที่ตรวจให้คะแนนหลายค่าในข้อสอบร่วม เมื่อใช้คะแนนร้อยละ 20 ของคะแนนทั้งหมดเป็นตัวกำหนดสัดส่วนคะแนนระหว่าง DS : PS จำนวน 3 สัดส่วน คือ 15 : 5 (DS 15 ข้อ : PS 1 ข้อ), 10 : 10 (DS 10 ข้อ : PS 2 ข้อ) และ 5 : 15 (DS 5 ข้อ : PS 3 ข้อ) รูปแบบการเก็บรวบรวมข้อมูล 2 รูปแบบคือ การเก็บรวบรวมข้อมูลสำหรับกลุ่มเท่าเทียมโดยใช้ข้อสอบร่วม และกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม ผลการปรับเทียบคะแนนจะพิจารณาคุณภาพในการปรับเทียบจากค่า MSE ที่ได้จากเกณฑ์ CRC หากการปรับเทียบเงื่อนไขใดมีค่า MSE น้อยกว่า แสดงว่ามีคุณภาพในการปรับเทียบมากกว่า จากกรอบแนวคิดในการวิจัย เขียนเป็นแผนภาพได้แผนภาพที่ 1 และแสดงการออกแบบการเชื่อมโยงคะแนน ดังตารางที่ 1




**แผนภาพที่** 1 กรอบแนวคิดในการวิจัย

ตารางที่ 1 การออกแบบการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม

| แบบสอบ | สัดส่วนน้ำหนักคะแนน | | | | กลุ่มเท่าเทียม | | กลุ่มไม่เท่าเทียม | |
|---|---|---|---|---|---|---|---|---|
| | 50 : 50 | 60:40 | 70:30 | 80:20 | จำนวนผู้สอบ | $\theta$ | จำนวนผู้สอบ | $\theta$ |
| แบบสอบ X | 50 : 10 (ข้อ) | 60:8 (ข้อ) | 70:6 (ข้อ) | 80:4 (ข้อ) | 1000 | $\mu = 0$ $\sigma = 1$ | 1000 | $\mu = 0$ $\sigma = 1$ |
| แบบสอบ Y | 50 : 10 (ข้อ) | 60:8 (ข้อ) | 70:6 (ข้อ) | 80:4 (ข้อ) | 1000 | $\mu = 0$ $\sigma = 1$ | 1000 | $\mu = 1$ $\sigma = 1$ |
| ข้อสอบร่วม | มิติของจำนวนข้อ (DS : PS) 15 : 2 , 20 : 2, 25 : 2 | | | | | | | |
| | มิติของคะแนน (DS : PS) 15 : 5, 10 : 10, 5 : 15 → คิดเป็นจำนวนข้อ (DS : PS) 15 : 1, 10 : 2, 5 : 3 | | | | | | | |

## คำนิยามเชิงปฏิบัติการที่ใช้ในการวิจัย

**แบบสอบรูปแบบผสม** หมายถึง แบบสอบที่ใช้ข้อสอบปรนัยที่ตรวจให้คะแนนสองค่า และข้อสอบอัตนัยที่ตรวจให้คะแนนหลายค่ารวมอยู่ในแบบสอบฉบับเดียวกัน โดยข้อสอบอัตนัยมีลักษณะเป็นข้อสอบที่เขียนตอบไม่ยาวมาก เช่น ข้อสอบแบบตอบสั้น เพื่อให้การให้คะแนนสำหรับข้อสอบมีความน่าเชื่อถือมากขึ้น และเหมาะต่อการนำไปใช้ในสถานการณ์การทดสอบระดับมหภาคได้ สำหรับการวิจัยครั้งนี้ได้ประยุกต์ใช้โมเดลการตรวจให้คะแนนตามทฤษฎีตอบสนองข้อสอบ โดยข้อสอบปรนัยประยุกต์ใช้โมเดลการตอบสนองข้อสอบที่มีการให้คะแนน 2 ค่า และข้อสอบอัตนัยประยุกต์ใช้โมเดลการตอบสนองข้อสอบที่มีการให้คะแนนหลายค่าไว้ในแบบสอบฉบับเดียวกัน

**ข้อสอบปรนัย** หมายถึง ข้อสอบที่มีการตรวจให้คะแนนสองค่า คือถูกได้ 1 คะแนน และผิดได้ 0 คะแนน โดยประยุกต์ใช้โมเดลการตอบสนองข้อสอบที่มีการให้คะแนน 2 ค่า คือโมเดลโลจีสติกแบบ 3 พารามิเตอร์ ประกอบด้วยพารามิเตอร์ความยาก (b) พารามิเตอร์อำนาจจำแนก (a) และพารามิเตอร์การเดา (c)

**ข้อสอบอัตนัย** หมายถึงข้อสอบที่มีการตรวจให้คะแนนมากกว่า 2 ค่า โดยประยุกต์โมเดลการตอบสนองข้อสอบที่มีการให้คะแนนหลายค่า คือโมเดล Generalized Partial Credit ประกอบด้วยค่าพารามิเตอร์ระดับความยากของขั้นการตอบ ($\delta_{ij}$) และค่าพารามิเตอร์ความชันของข้อคำถาม ($\alpha_i$) สำหรับการวิจัยครั้งนี้โมเดล GPC มีลำดับขั้นการให้คะแนน 6 ลำดับขั้น (คะแนนเต็ม 5 คะแนน) เป็นสเกลคะแนนที่ไม่กว้างมาก ใช้ได้กับข้อสอบที่มีลักษณะตอบสั้น ๆ การตรวจให้คะแนนมีความเป็นปรนัยและมีความน่าเชื่อถือ สามารถนำไปใช้ในการทดสอบระดับมหภาคได้

**ข้อสอบร่วม** หมายถึง ข้อสอบที่ใช้ร่วมกันระหว่างแบบสอบ 2 ฉบับ สำหรับการวิจัยครั้งนี้ ข้อสอบร่วมประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า และข้อสอบที่ตรวจให้คะแนนมากกว่า 2 ค่า โดยพิจารณาลักษณะของข้อสอบร่วม 2 มิติ คือ จำนวนข้อสอบร่วม และคะแนนรวมในข้อสอบร่วม

**จำนวนข้อสอบร่วม** หมายถึง จำนวนข้อสอบที่ตรวจให้คะแนน 2 ค่าและข้อสอบที่ตรวจให้คะแนนมากกว่า 2 ค่า ที่ประกอบอยู่ในข้อสอบร่วม โดยใช้จำนวนข้อสอบร่วมขั้นต่ำร้อยละ 20 ของข้อสอบทั้งหมด

**คะแนนรวมในข้อสอบร่วม** หมายถึง คะแนนจากข้อสอบที่ตรวจให้คะแนนสองค่าและคะแนนจากข้อสอบที่ตรวจให้คะแนนมากกว่า 2 ค่า ที่รวมอยู่ในข้อสอบร่วม สำหรับการวิจัยครั้งนี้พิจารณาคะแนนรวมจากข้อสอบทั้ง 2 ชนิด ในข้อสอบร่วมร้อยละ 20 ของคะแนนเต็มทั้งฉบับ

**การปรับเทียบคะแนน** หมายถึง การแปลงระบบคะแนนของแบบสอบต่างฉบับที่วัดเนื้อหาเดียวกัน โดยแปลงคะแนนของแบบสอบฉบับหนึ่งให้อยู่บนมาตรหรือสเกลของแบบสอบอีกฉบับหนึ่งที่ถือเป็นคะแนนที่สมมูลกัน เพื่อทำให้คะแนนจากแบบสอบต่างฉบับกันสามารถเปรียบเทียบกันได้โดยตรง สำหรับการวิจัยครั้งนี้เป็นการแปลงคะแนนของแบบสอบสองฉบับที่วัดเนื้อหาเดียวกันด้วยวิธีการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ

**วิธีการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ** หมายถึง กระบวนการวิเคราะห์ที่นำค่าพารามิเตอร์ความสามารถ และพารามิเตอร์ข้อสอบมาดำเนินการโดยยึดหลักว่า ถ้าพารามิเตอร์ความสามารถอยู่บนสเกลเดียวกันแล้ว สามารถนำคะแนนจากแบบสอบสองฉบับมาเทียบกันได้ โดยใช้ข้อสอบร่วมเป็นตัวปรับความแตกต่างของกลุ่ม การวิจัยครั้งนี้ปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ 2 วิธี คือ

วิธีโค้งคุณลักษณะแบบสอบ หมายถึง กระบวนการปรับเทียบคะแนนที่แปลงค่าพารามิเตอร์ความสามารถเชิงเส้นจากแบบทดสอบต่างฉบับให้อยู่ในมาตรวัดเดียวกัน โดยประมาณค่าสัมประสิทธิ์การปรับเทียบคะแนนด้วยโค้งคุณลักษณะแบบสอบร่วม

วิธีปรับค่าพารามิเตอร์พร้อมกัน หมายถึง กระบวนการเทียบคะแนนที่วิเคราะห์ค่าพารามิเตอร์ข้อสอบและความสามารถจากข้อสอบทั้งสองฉบับพร้อมกันทำให้ค่าพารามิเตอร์อยู่ในมาตรวัดเดียวกัน

**คุณภาพของการปรับเทียบคะแนน** (Quality of Equating Method) หมายถึง ความถูกต้องแม่นยำของการปรับเทียบคะแนน ซึ่งพิจารณาจากดัชนีที่พิจารณาขนาดความคลาดเคลื่อน ที่เกิดจากวิธีการปรับเทียบคะแนน โดยคำนวณค่าความแตกต่างค่าที่ประมาณได้ กับค่าที่แท้จริง ซึ่งเกณฑ์ที่นำมาใช้ในการเปรียบเทียบคุณภาพของการปรับเทียบคะแนนในครั้งนี้ พิจารณาจากค่าค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (Mean Square Error : MSE) ที่ได้จากเกณฑ์โค้งคุณลักษณะลำดับขั้นการตอบ (Category response curve) มีพื้นฐานมาจากความแตกต่างระหว่างโค้งลำดับขั้นการตอบที่ประมาณค่าได้ กับโค้งลำดับขั้นการตอบที่เป็นค่าที่แท้จริง

## ประโยชน์ที่คาดว่าจะได้รับ

**1. ประโยชน์ทางด้านวิชาการ**

1.1 ได้ขยายองค์ความรู้ในด้านการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบสำหรับแบบสอบรูปแบบผสม ทั้งทางทฤษฎีและการปฏิบัติเพื่อนำไปประยุกต์ใช้ในบริบททางการศึกษา

1.2 ผลการวิจัยให้ข้อมูลสารสนเทศเกี่ยวกับผลการเปรียบเทียบคุณภาพของการปรับเทียบคะแนนที่มีวิธีการแตกต่างกัน 2 วิธี ระหว่างการปรับค่าพารามิเตอร์พร้อมกัน และการปรับค่าพารามิเตอร์แยกกันด้วยวิธีโค้งคุณลักษณะ ในเงื่อนไขของการศึกษา คือ การออกแบบการเก็บรวบรวมข้อมูล สัดส่วนของข้อสอบระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า และข้อสอบที่ตรวจให้คะแนนหลายค่าที่อยู่ในแบบสอบ และสัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่า กับข้อสอบที่ตรวจให้คะแนนหลายค่าที่อยู่ในชุดข้อสอบร่วม ทำให้ได้ผลสรุปที่ชัดเจนขึ้นเกี่ยวกับวิธีการเชื่อมโยงคะแนนในสถานการณ์การศึกษาที่แตกต่างกัน และเป็นพื้นฐานที่สำคัญในการวิจัยเกี่ยวกับการวัดและประเมินผล

**2. ประโยชน์ทางด้านการนำไปใช้**

2.1 ผลการวิจัยสามารถนำไปใช้ประกอบการตัดสินใจในการเลือกใช้วิธีการปรับเทียบคะแนนที่เหมาะสมและสอดคล้องกับสถานการณ์ที่ต้องการศึกษามากที่สุด

2.2 ผลการวิจัยเป็นแนวทางในการนำไปใช้ในการบริหารจัดการ การทดสอบให้มีความยุติธรรมมากยิ่งขึ้น ในการทดสอบระดับมหภาค ที่มีการสอบวัดผลสัมฤทธิ์ของผู้เรียน ในด้านทักษะต่าง ๆ ที่ไม่สามารถใช้ข้อสอบเพียงรูปแบบใดรูปแบบหนึ่งเพียงรูปแบบเดียวได้ เช่น การสอบวัดผลสัมฤทธิ์ทางคณิตศาสตร์ของผู้เรียน ซึ่งทักษะการคำนวณที่ซับซ้อน ต้องการให้ผู้สอบได้แสดงความสามารถเกี่ยวกับลำดับการคิดจนได้คำตอบ การใช้ข้อสอบเลือกตอบเพียงอย่างเดียวอาจไม่สามารถวัดทักษะเหล่านั้นได้ การใช้แบบสอบรูปแบบผสมจึงมีบทบาทที่สำคัญสำหรับการแก้ปัญหาเหล่านี้ หรือการทดสอบคัดเลือกบุคคลเข้าศึกษาต่อ ที่มีข้อสอบรูปแบบผสม หรือการทดสอบในระดับชาติ ระดับสากล เช่น การทดสอบภาษาอังกฤษ TOEFL ด้วยการสอบระบบคอมพิวเตอร์ กับการทดสอบภาษาอังกฤษ TOEFL ที่ใช้รูปแบบเดิม (paper based test) ซึ่งการสอบทั้งสองรูปแบบนี้คะแนนไม่สามารถเทียบกันได้โดยตรง หากมีการทดสอบต่างกลุ่มผู้สอบ ต่างเวลา การปรับเทียบคะแนนก็เป็นเทคนิคอย่างหนึ่งที่ทำให้การวัดและประเมินผลมีความยุติธรรมยิ่งขึ้น

# บทที่ 2
# เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบและประเมินคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างการตรวจให้คะแนนสองค่าและการตรวจให้คะแนนหลายค่าที่เป็นเอกมิติ ตามวิธีการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ ผู้วิจัยได้ศึกษาค้นคว้าตำรา เอกสาร บทความ และงานวิจัยที่เกี่ยวข้องกับการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมทั้งในและต่างประเทศ เพื่อให้ได้โมเดลการตอบที่เหมาะสมกับสภาพการศึกษาและสถานการณ์ที่มีความเป็นไปได้ต่อการนำไปใช้จริง กระบวนการที่เหมาะสมสำหรับการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมรวมทั้งการประเมินผลการปรับคะแนนด้วย โดยแสดงให้เห็นถึงความสำคัญและตัวอย่างของแบบสอบรูปแบบผสม การประยุกต์ใช้ทฤษฎีตอบสนองข้อสอบโมเดลการตอบสนองข้อสอบที่ตรวจให้คะแนนสองค่า และหลายค่า หลักการสำคัญของการปรับเทียบคะแนน การประยุต์ใช้วิธีการปรับเทียบตามทฤษฎีการตอบสนองข้อสอบสำหรับแบบสอบรูปแบบต่าง ๆ และได้นำเสนอผลการวิจัยที่ศึกษาการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม จากที่กล่าวมาข้างต้น ผู้วิจัยได้แสดงรายละเอียดของการค้นคว้าแบ่งเป็น 6 ตอน ดังต่อไปนี้

ตอนที่ 1 ความสำคัญของแบบสอบรูปแบบผสม
ตอนที่ 2 ทฤษฎีการตอบสนองข้อสอบ
ตอนที่ 3 การปรับเทียบคะแนนระหว่างแบบสอบ
ตอนที่ 4 การปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ
ตอนที่ 5 โปรแกรมคอมพิวเตอร์ที่ใช้ในการวิจัย
ตอนที่ 6 งานวิจัยที่เกี่ยวข้องกับการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม

## ตอนที 1 ความสำคัญของแบบสอบรูปแบบผสม

จากพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ.2542 (แก้ไขเพิ่มเติม พ.ศ.2545) มาตรา 26 ได้กล่าวถึงหลักการวัดและประเมินผลการเรียนรู้ว่า *สถานศึกษาจัดการประเมินผู้เรียนโดยพิจารณาจากพัฒนาการ ความประพฤติ การสังเกตพฤติกรรมการเรียน การร่วมกิจกรรมและการทดสอบควบคู่ไปในกระบวนการเรียนการสอน* โดยมีแนวคิดและหลักการของการปฏิรูประบบการประเมินการเรียนรู้ที่สำคัญ (ศิริชัย กาญจนวาสี, 2546) คือ สถานศึกษาจะต้องกำหนดเป้าหมายการเรียนรู้อย่างชัดเจน มีนโยบายในการวัดและประเมินผลการเรียนรู้ที่เป็นกลไกของการกำกับติดตาม สนับสนุนและพัฒนาการเรียนรู้ของผู้เรียน ทำการวัดและประเมินผลการเรียนรู้อย่างเที่ยงธรรมโปร่งใสและมีความต่อเนื่องทั้งก่อน ระหว่างและหลังเรียนเพื่อพัฒนาผู้เรียนเป็นสำคัญ ดังนั้นกระบวนการวัดและประเมินผลการเรียนรู้ตามแนวคิดและหลักการตามเจตนารมณ์สำคัญของการปฏิรูปการศึกษาดังกล่าวจำเป็นจะต้องมีการดำเนินการสอดคล้องกับธรรมชาติการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ อันประกอบด้วยวิธีการประเมินที่หลากหลายไม่ว่าจะเป็นการประเมินตามสภาพจริง (authenthic assessment) การประเมินภาคปฏิบัติ (performance assessment) การประเมินด้วยแฟ้มสะสมงาน (portfolio) การประเมินโดยใช้ศูนย์การประเมิน (assessment center) รวมถึงการประเมินด้วยการทดสอบ (testing)

จากวิธีการประเมินที่กล่าวมานั้นจะเห็นว่าการประเมินด้วยการทดสอบเป็นวิธีการที่ได้รับความนิยมและมีการใช้กันอย่างแพร่หลายไม่เพียงแต่ด้านการศึกษาเท่านั้น ศิริชัย กาญจนวาสี (2548) ได้ให้ความหมายของการทดสอบว่า เป็นกระบวนการใช้แบบสอบสำหรับกำหนดหรือบรรยายคุณลักษณะหรือคุณภาพของบุคคลหรือกลุ่มบุคคลเพื่อใช้เป็นสารสนเทศสำหรับการตัดสินใจ กล่าวคือ การวัดและประเมินด้วยวิธีการทดสอบนั้นจะใช้เครื่องมือที่เรียกว่า แบบสอบ (test) ซึ่งเป็นชุดของข้อคำถามที่ใช้วัดกลุ่มตัวอย่างพฤติกรรมเกี่ยวกับความสามารถทางสมองหรือความรู้สึกนึกคิดทางจิตใจ หรือทักษะการดำเนินงานของบุคคล หรือกลุ่มบุคคลภายใต้สถานการณ์ที่เป็นมาตรฐานและมีการกำหนดหลักเกณฑ์การให้คะแนนที่ชัดเจน สำหรับแบบสอบที่ใช้ในการวัดและประเมินผลของโรงเรียนส่วนใหญ่เป็นแบบสอบผลสัมฤทธิ์ (achievement test) มีจุดมุ่งหมายที่สำคัญ คือ เพื่อใช้วัดผลการเรียนรู้ด้านเนื้อหาวิชาและทักษะต่าง ๆ ของแต่ละวิชา ประกอบด้วยแบบสอบประเภทข้อเขียน (paper and pencil test) และแบบสอบประเภทปฏิบัติ (performance test) (เยาวดี วิบูลย์ศรี, 2548) แบบสอบข้อเขียนที่ใช้ในการวัดผลสัมฤทธิ์ของผู้เรียนจำแนกตามประเภทของวิธีการให้คะแนนได้เป็น 2 ประเภท คือ แบบสอบปรนัย (objective

test) และแบบสอบอัตนัย (subjective/essay test) แบบสอบทั้งสองประเภทนี้มีความแตกต่างกันทั้งด้านลักษณะข้อสอบ ความครอบคลุมของจุดมุ่งหมายการเรียนรู้ ระดับความสามารถที่มุ่งวัด ตลอดจนโอกาสการเดาคำตอบได้ถูก

การทดสอบด้วยแบบสอบอัตนัยมีการใช้กันอย่างแพร่หลายตลอดหลายสิบปีที่ผ่านมา โดยนำมาใช้ประโยชน์หลายด้านไม่ว่าจะเป็นการติดตามพัฒนาการเกี่ยวกับเนื้อหาและทักษะการเขียนของนักเรียน การประเมินผลสัมฤทธิ์ปลายภาค ตลอดจนการคัดเลือกบุคคลเข้าศึกษาต่อระดับที่สูงขึ้น ต่อมาภายหลังการใช้แบบสอบอัตนัยของโรงเรียนเริ่มมีปัญหาเกี่ยวกับความเป็นปรนัยในการตรวจจึงมีการปรับให้มีการตอบสั้นลง ขณะเดียวกันมีงานวิจัยที่ศึกษา เปรียบเทียบผลการสอบจากแบบสอบหลายตัวเลือกและแบบสอบอัตนัยซึ่งพบว่าคะแนนจากการทดสอบทั้งสองมีความสัมพันธ์กันสูง (Walberge and Haertel, 1992) จึงนำไปสู่การใช้แบบสอบแบบหลายตัวเลือกในการวัดและประเมินผลแทนแบบสอบอัตนัยมากขึ้น อย่างไรก็ตามจากผลการประเมินความก้าวหน้าทางการศึกษาระดับชาติ (National Assessment of Education Progress) ของอเมริกา (NAEP, 1981 อ้างถึงใน Walberge and Haertel, 1992) ได้มีข้อเสนอแนะเกี่ยวกับความสอดคล้องของความสามารถของผู้เรียนตามเป้าหมายของหลักสูตร ซึ่งพบข้อบกพร่องที่สืบเนื่องมาจากการใช้แบบสอบปรนัยแบบหลายตัวเลือกในการวัดและประเมินผลการเรียนรู้ของผู้เรียนว่าไม่ได้พัฒนาความสามารถและทักษะการเขียนของผู้เรียน ขณะเดียวกันในปัจจุบันสังคมเริ่มตั้งข้อสังเกตถึงผลข้างเคียงจากการใช้แบบสอบปรนัยว่าทำให้นักเรียนเขียนหนังสือไม่เป็น ข้อสอบปรนัยสามารถวัดได้แค่ระดับความจำและวัดได้แคบในกรอบของตัวเลือกที่นำเสนอเท่านั้น ไม่สามารถคิดบนพื้นฐานเหตุผลส่วนตัวได้ (ทศพล เชี่ยวชาญประพันธ์, 2543 อ้างถึงใน พวงแก้ว ปุณยกนก, 2546) ขณะที่แบบสอบอัตนัยเป็นแบบสอบที่ให้อิสระในการแสดงความคิด สามารถวัดผลการเรียนรู้ตามแนวคิดของบลูม (Bloom) ตั้งแต่ระดับความรู้ความจำจนถึงระดับการวิเคราะห์ สังเคราะห์ และประเมินค่า ซึ่ง Thorndike และ Hagen (1977) และ Kubiszyn และ Borich (2003) กล่าวว่า ข้อสอบอัตนัยสามารถวัดความสามารถของผู้เรียนในการจัดระบบ (organize) การบูรณาการ (integrate) และการสังเคราะห์ความรู้ (synthesize) เพื่อใช้ข้อมูลในการแก้ปัญหาหรือเป็นการริเริ่ม และปรับปรุงวิธีการแก้ปัญหาใหม่ โดยแบบสอบอัตนัยสามารถสร้างและดำเนินการสอบได้ง่าย ไม่สิ้นเปลืองวัสดุอุปกรณ์ในการจัดพิมพ์ สามารถทดสอบได้ทุกเวลาที่ทำการสอน (Coffman, 1971 อ้างถึงใน ไพรัตน์ วงษ์นาม, 2533) ด้วยเหตุผลดังกล่าวจึงมีการผลักดันให้นำแบบสอบอัตนัยกลับมาใช้อีกครั้งเพื่อกระตุ้นให้ผู้เรียนมีการคิดใช้เหตุผลขั้นสูง มีการแก้ปัญหาที่ซับซ้อนตลอดจนการประยุกต์เนื้อหาความรู้ที่มีอยู่

แม้ว่าแบบสอบอัตนัยจะเป็นเครื่องมือสำคัญสำหรับการวัดและประเมินผลการเรียนรู้ของผู้เรียนในยุคปัจจุบัน แต่ด้วยข้อจำกัดของแบบสอบอัตนัย 2 ประการ คือ 1) ข้อจำกัดด้านการตรวจแบบสอบอัตนัยที่ต้องใช้ระยะเวลาและแรงงานมาก รวมทั้งอาจเกิดความคลาดเคลื่อนได้หากไม่มีเกณฑ์การตรวจที่ชัดเจน และ 2) ข้อจำกัดด้านการสร้างแบบสอบอัตนัยที่ไม่สามารถสร้างแบบสอบให้ครอบคลุมเนื้อหาและวัตถุประสงค์การเรียนรู้ที่กำหนดได้ทั้งหมด (Coffman, 1971; Ebel และ Frisble, 1986; Mehrens และ Lehmann, 1984 อ้างถึงใน อรุณี เร้าอรุณ, 2536; Brown, 1983 อ้างถึงใน เยาวดี วิบูลย์ศรี, 2548) ซึ่งข้อจำกัดเหล่านี้ล้วนมีผลต่อประสิทธิภาพของการนำแบบสอบอัตนัยมาใช้เป็นอย่างมาก อย่างไรก็ตามแม้ว่าการใช้แบบสอบอัตนัยจะมีข้อจำกัดดังกล่าวแต่ด้วยคุณประโยชน์ของแบบสอบอัตนัยที่มีต่อการพัฒนาผู้เรียนในยุคการศึกษาปฏิรูปจึงไม่อาจปฏิเสธที่จะนำมาแบบสอบอัตนัยมาใช้ในการวัดและประเมินผลการเรียนรู้ของผู้เรียนได้ ดังนั้น เพื่อให้การวัดและประเมินผลการเรียนรู้ของผู้เรียนมีประสิทธิภาพยิ่งขึ้น การใช้แบบสอบอัตนัยผสมกับแบบสอบปรนัยจึงเป็นอีกทางเลือกหนึ่งที่จะให้การวัดผลการเรียนรู้มีความครอบคลุมในเนื้อหาสาระโดยการวัดด้วยข้อสอบปรนัย และยังสามารถให้ผู้เรียนได้แสดงความสามารถในการจัดระบบความคิดด้วยการวัดจากข้อสอบอัตนัย หากมีการนำแบบสอบรูปแบบผสมไปใช้ในการวัดและประเมินผลการเรียนรู้อย่างกว้างขวาง ก็จะเป็นการส่งเสริมให้ผู้เรียนมีการพัฒนาความสามารถของตนอย่างรอบด้านอีกด้วย ซึ่ง Hambleton (1996) ได้กล่าวว่า การประเมินทางการศึกษาในทศวรรษที่ 20 และในทศวรรษหน้าจะมีความแตกต่างอย่างมากจากการประเมินในสองทศวรรษที่ผ่านมา รูปแบบที่การประเมินมีความเป็นปรนัยดังเช่นแบบสอบเลือกตอบจะนำมาใช้ให้สมดุลกับแบบสอบรูปแบบอื่นมากยิ่งขึ้น โดยการวัดร่วมกับการประเมินการปฏิบัติงาน การเขียนโครงงาน และจัดทำพอตโฟลิโอ โดยแบบสอบรูปแบบผสมระหว่างข้อสอบที่มีรูปแบบต่างกัน จะนำไปใช้ทั้งในระดับชั้นเรียน และในระดับมหภาค ซึ่งข้อสอบลักษณะนี้มักใช้วัดทักษะของผู้เรียนในด้านต่าง ๆ มากกว่าที่จะใช้แบบสอบที่ประกอบด้วยข้อสอบเพียงรูปแบบเดียว (Single Format)

ในปัจจุบันได้มีการนำแบบสอบรูปแบบผสมไปใช้ในหลายระดับ ทั้งที่เป็นการวัดและประเมินในชั้นเรียน หรือในระดับมหภาค ผู้วิจัยได้ยกตัวอย่างการทดสอบที่ใช้แบบสอบรูปแบบผสม ดังนี้

**การทดสอบภาษาอังกฤษ** TOEFL คือ แบบทดสอบความสามารถ ในการใช้ภาษาอังกฤษ ของผู้ที่ไม่ได้ใช้ภาษาอังกฤษ เป็นภาษาประจำชาติ โดยที่ จะจัดการสอบเป็นแบบปรนัย เพื่อวัดความเข้าใจ ภาษาอังกฤษ (แบบอเมริกาเหนือ) ปัจจุบันการสอบ TOEFL ได้เปลี่ยน

มาใช้การทดสอบแบบ Computer - Based Testing แทนการทดสอบแบบ Paper test โดย เชื่อว่าวิธีนี้ จะสามารถวัดระดับความรู้ภาษาอังกฤษของนักเรียน ได้ถูกต้อง ใกล้เคียงความเป็นจริงมากกว่าการสอบแบบเดิม (Paper-Based Test) การทดสอบแบบ Computer - Based Testing จะแบ่งแบบทดสอบออกเป็นสี่ส่วน คือ ส่วนของการฟัง (Listening) ส่วนของไวยากรณ์ (Structure) ส่วนของการอ่าน (Reading) และส่วนของการเขียน (Writing) โดยลักษณะคำถามส่วนใหญ่แล้วเป็นแบบปรนัย หรือ มีคำตอบให้เลือกตอบ แต่ปัจจุบันมีคำถามลักษณะใหม่ๆ ออกมาค่อนข้างมาก ไม่ว่าจะเป็น คำถามที่ให้เลือกภาพในการตอบ หรืออาจเป็นคำถามที่มีหลายคำตอบ หรือ ให้เรียงลำดับสิ่งของ หรือ จับคู่ให้เป็นหมวดหมู่

สำหรับการให้คะแนนในส่วนของการฟังจะมีคะแนนในช่วง 0 – 30 คะแนน ส่วนของไวยากรณ์ และการเขียนมีคะแนนในช่วง 0 – 30 คะแนน ซึ่งส่วนของไวยากรณ์คะแนนจะรวมอยู่กับส่วนการเขียนที่มีการตรวจให้คะแนนหลายค่า 7 ระดับ คือระดับ 0 – 6 ส่วนของการอ่านอยู่ในช่วง 0 – 30 คะแนน

**การทดสอบ GMAT** (Graduate Management Admission Test) เป็นข้อสอบที่ใช้วัดความสามารถของผู้ที่ต้องการเข้าศึกษาในระดับปริญญาโทและปริญญาเอกสาขาบริหารธุรกิจซึ่งรวมทั้ง MBA, M.S. Marketing, M.S. Finance, MIS (สำหรับสาขา MIS ของบางมหาวิทยาลัยอาจต้องใช้คะแนน GRE แทน), DBA และ Ph.D. ด้านบริหารธุรกิจส่วนใหญ่ต้องใช้คะแนน GMAT ในการพิจารณารับนักศึกษา ซึ่งข้อสอบเป็นการวัดความรู้ในการสื่อสารซึ่งรวมทั้งการอ่านและการเขียน, ทักษะการวิเคราะห์และทักษะในการคำนวณ ที่จะสามารถใช้เป็นเครื่องชี้วัดความสำเร็จในการเรียนต่อทางด้านบริหารธุรกิจ

ข้อสอบ GMAT ประกอบด้วยข้อสอบ 3 ส่วนคือ 1. การเขียน (Analytical Writing Assessment) 2. คณิตศาสตร์ (Quantitative) และ 3. ภาษาอังกฤษ (Verbal)

*ข้อสอบการเขียน (AWA)* มีรูปแบบของข้อสอบ 2 ลักษณะได้แก่ 1. การเขียนเพื่อแสดงความคิดเห็น (Issue) และ 2. การเขียนเพื่อแสดงวิจารณ์บทความ (Argument) ผู้เข้าสอบจะมีเวลา 30 นาทีต่อหนึ่ง essay ซึ่งส่วนนี้จะเป็นการตรวจให้คะแนนแบบหลายค่า

*ข้อสอบคณิตศาสตร์ (Quantitative)* ผู้เข้าสอบจะต้องทำโจทย์เลขแบบ multiple-choice จำนวน 37 ข้อโดยมีรูปแบบของข้อสอบสองลักษณะนั้นคือ 1. Problem Solving 24 ข้อ และ 2. Data Sufficiency 13 ข้อ โดยมีเวลาทำข้อสอบทั้งสิ้น 75 นาที ส่วนนี้ข้อสอบจะมีการให้คะแนนแตกต่างกันไปขึ้นอยู่กับระดับความยากของข้อสอบ

*ข้อสอบภาษาอังกฤษ (Verbal)* ข้อสอบส่วนนี้จะเป็นโจทย์แบบ multiple-choice จำนวน 41 ข้อโดยมีรูปแบบของข้อสอบสามลักษณะนั้นคือ 1. การอ่าน (Reading Comprehension) 14 ข้อ 2. การวิเคราะห์ (Critical Reasoning) 14 ข้อและ 3. ไวยกรณ์และการเขียน (Sentence Correction) 13 ข้อ โดยมีเวลาทำข้อสอบทั้งสิ้น 75 นาที

**การทดสอบ** CU-TAD

การทดสอบทักษะด้านการออกแบบ ใช้สำหรับผู้สอบที่ประสงค์เข้าศึกษา ต่อในระดับปริญญาตรี ภาควิชาสถาปัตยกรรมศาสตร์ (หลักสูตรนานาชาติ) คณะสถาปัตยกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

แบบทดสอบ CU-TAD ประกอบด้วยข้อสอบ 3 ส่วน โดยส่วนที่ 1 เป็นข้อสอบวัด Spatial Perception and Logical Analysis ซึ่งส่วนนี้ข้อสอบเป็นแบบหลายตัวเลือกที่มีการตรวจให้คะแนนแบบสองค่า มีจำนวนข้อสอบ 30 ข้อ คะแนน 30 คะแนน ใช้เวลาในการทำส่วนนี้ 45 นาที ส่วนที่ 2 เป็นข้อสอบวัด Generation of Ideas และ Design Ability ส่วนนี้เป็นข้อสอบที่ตรวจให้คะแนนหลายค่า มีจำนวยข้อ 7 ข้อคะแนนรวม 50 คะแนน และใช้เวลาในการทดสอบ 90 นาที และส่วนที่ 3 เป็นข้อสอบวัด Drawing and Painting Skill เป็นข้อสอบที่ตรวจให้คะแนนหลายค่า มีจำนวนข้อ 1 ข้อ คะแนนเต็ม 20 คะแนน ใช้เวลาในการสอบ 45 นาที

**การทดสอบ GRE** (Graduate Record Examination) และเป็นข้อสอบวัดเชาวน์ปัญญาทั่วไป ที่ต้องสอบสำหรับผู้ต้องการเรียนต่อในระดับสูงกว่าปริญญาตรี ในประเทศสหรัฐอเมริกา ยกเว้น Business School ที่ต้องสอบ GMAT และ Law School ที่ต้องสอบความรู้ด้านกฎหมาย การสอบนั้นประกอบไปด้วย multiple-choice สามส่วนด้วยกัน, Quantitative, Analytical, Verbal และอีกส่วนหนึ่งที่สอบแยกต่างหากคือ การเขียน essay ที่เรียกว่า Writing Assessment. ซึ่งGRE เพิ่งนำการสอบแบบ Writing Assessment เมื่อเดือนตุลาคม 1999 การสอบนี้จะสอบต่างวันจากการสอบอื่นๆ ในแต่ละคณะจะต้องการข้อเขียนที่แตกต่างกันไป จึงต้องเลือกว่า จะสอบเพื่อนำไปใช้เรียนต่อ ในสาขาวิชาใด เพื่อจะได้ ทำข้อสอบให้ตรงกับความต้องการของสาขาวิชานั้นๆ ในการสอบแบบนี้จะมี อยู่สองคำถาม ให้ทำ โดย บทความแรกจะเป็นการให้เราออกความเห็นเกี่ยวกับเรื่องที่อ่าน ซึ่งใช้เวลาสี่สิบห้านาที ส่วนที่สองนั้น จะให้เราวิเคราะห์ ความเป็นเหตุและผลในข้อเขียนที่อ่าน เป็นเวลา สามสิบนาที Writing Assessment จะมีการประเมินผลอยู่ในระดับ 0 ถึง 6 คะแนน ในประเทศไทย ปัจจุบันได้เปลี่ยนรูปแบบการสอบ มาเป็นแบบ Computer Adaptive Test (CAT)

การสอบ GRE มีอยู่ 2 รูปแบบ คือ การสอบทั่วไป (General Test ) และการสอบเฉพาะวิชาสาขา (Subject Test) ในวิชาต่างๆ 16 สาขา การสอบทั่วไป (General Test ) เป็นการสอบ

เพื่อวัดทักษะของผู้สอบที่มีอยู่ โดยวัดออกมาในรูปของคะแนนของความสามารถทางภาษา คำนวณและความสามารถในเชิงวิเคราะห์ การสอบใช้เวลา 3 ชั่วโมง 30 นาที

## ตอนที่ 2 ทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) เป็นทฤษฎีการวัดที่อธิบายความสัมพันธ์ระหว่างคุณลักษณะภายใน หรือความสามารถที่มีอยู่ในตัวบุคคลกับพฤติกรรมการตอบข้อสอบของบุคคลนั้น ว่ามีโอกาสตอบข้อสอบถูกมากน้อยเพียงใด ทฤษฎีนี้มีพื้นฐานความเชื่อว่าพฤติกรรมการตอบข้อสอบของผู้สอบถูกกำหนดโดยคุณลักษณะภายใน หรือความสามารถที่มีอยู่ภายในตัวบุคคล (ศิริชัย กาญจนวาสี, 2545) ทฤษฎีนี้มีแนวคิดพื้นฐานที่สำคัญ 2 ประการ (Hambleton, Swaminatan and Roger, 1991) คือ

1. พฤติกรรมการแสดงออกของบุคคลในการตอบข้อสอบสามารถทำนายได้ด้วยกลุ่มขององค์ประกอบที่เรียกว่าคุณลักษณะ (trait) หรือคุณลักษณะแฝง (latent traits) หรือความสามารถ (abilities) ของบุคคลนั้น

2. ความสัมพันธ์ระหว่างการแสดงพฤติกรรมการตอบข้อสอบของผู้สอบกับชุดของคุณลักษณะแฝงที่อยู่ภายใต้คุณลักษณะที่แสดงออกของข้อสอบสารถอธิบายได้โดยฟังก์ชันที่เพิ่มขึ้นในลักษณะทิศทางเดียว (monotonically) ที่เรียกว่าฟังก์ชันคุณลักษณะข้อสอบหรือโค้งลักษณะข้อสอบ (item characteristic curve)

**ลักษณะทั่วไปของทฤษฎีการตอบสนองข้อสอบ**

ทฤษฎีนี้เกิดขึ้นท่ามกลางข้อจำกัดของทฤษฎีการทดสอบแบบดั้งเดิมหลายประการคือ (Hambleton and Swaminatan, 1985; Hambleton and others, 1991)

1. ค่าสถิติของข้อสอบ เช่น ความยาก จะขึ้นอยู่กับลักษณะของกลุ่มผู้สอบ กล่าวคือ ถ้าผู้สอบมีความสามารถสูง ข้อสอบจะกลายเป็นข้อสอบที่ง่าย แต่ถ้าผู้สอบมีความสามารถต่ำข้อสอบดังกล่าวจะกลายเป็นข้อสอบที่ยาก ส่วนอำนาจจำแนกขึ้นอยู่กับความเป็นเอกพันธ์ของความสามารถของผู้สอบ ถ้าผู้สอบมีความสามารถแตกต่างกันมากข้อสอบก็จะมีอำนาจจำแนกสูง ซึ่งมีผลทำให้ความเที่ยงของแบบสอบมีค่าสูงตามไปด้วย เนื่องจากความเที่ยงของแบบสอบมีความสัมพันธ์ทางบวกกับค่าความแปรปรวนของคะแนนจากแบบสอบ

2. การเปรียบเทียบความสามารถของผู้สอบนั้น จะต้องใช้แบบสอบฉบับเดียวกันหรือแบบสอบคู่ขนาน ปัญหาที่เกิดขึ้นคือแบบสอบวัดผลสัมฤทธิ์และแบบสอบวัดความถนัดนั้น ส่วนใหญ่แล้วจะเหมาะสมกับผู้ที่มีความสามารถปานกลาง ดังนั้นความถูกต้องแม่นยำของการวัดผู้สอบที่มีความสามารถสูงและผู้สอบที่มีความสามารถต่ำจึงลดลง

3. ค่าความเที่ยงของแบบสอบถูกนิยามในรูปของผลที่ได้จากการใช้แบบสอบคู่ขนาน ซึ่งในทางปฏิบัติจริงนั้นนับว่าเป็นเรื่องยากที่จะให้การสอบ 2 ครั้งมีสภาพที่เหมือนกัน ถึงแม้ว่าแบบสอบคู่ขนานจะขนานกันจริง แต่ผู้สอบอาจจะมีลักษณะที่แตกต่างกันไปจากการสอบครั้งแรกเกี่ยวกับแรงจูงใจ ความกังวล การลืม หรือการพัฒนาตนเองในบางทักษะ เป็นต้น

4. ทฤษฎีการทดสอบแบบดั้งเดิมไม่สามารถบอกได้ว่าผู้สอบจะตอบข้อสอบอย่างไร ยกเว้นแต่ว่าจะได้ใช้ข้อสอบข้อนั้นกับผู้สอบที่มีลักษณะคล้ายคลึงกันมาแล้ว

5. ทฤษฎีการทดสอบแบบดั้งเดิมใช่ค่าความแปรปรวนของความคลาดเคลื่อนในการวัด (variance of error of measurement) เหมือนกันกับผู้สอบทุกคน ซึ่งตามความเป็นจริงแล้วผู้สอบที่มีความสามารถสูงและต่ำจะมีค่าความแปรปรวนของความคลาดเคลื่อนในการวัดต่างจากผู้สอบที่มีความสามารถปานกลาง

**หลักการของทฤษฎีการตอบสนองข้อสอบ**

ทฤษฎีการตอบสนองข้อสอบ (Item response Theory) มีความเชื่อเกี่ยวกับค่าพารามิเตอร์ของข้อสอบ (item parameter) คือ ค่าความยาก ($b_i$) ค่าอำนาจจำแนก ($a_i$) ค่าการเดา ($c_i$) ของข้อสอบแต่ละข้อว่าเป็นคุณลักษณะที่คงที่ในตัวข้อสอบนั้น เพราะฉะนั้นค่าพารามิเตอร์เหล่านี้จึงไม่ควรแปรเปลี่ยนไปตามกลุ่มผู้สอบและในทำนองเดียวกัน ค่าพารามิเตอร์ของผู้สอบ (person parameter) หรือความสามารถที่แม้จริงของผู้สอบ ก็เป็นคุณลักษณะที่มีอยู่ภายในตัวผู้สอบ จึงไม่ควรแปรเปลี่ยนไปตามชุดข้อสอบที่เลือกใช้ แต่เนื่องจากความสามารถของผู้สอบเป็นคุณลักษณะแฝงไม่สามารถสังเกตหรือวัดได้โดยตรง (unobservable) จึงจำเป็นต้องใช้การทำนาย (predict) หรืออธิบาย (explain) คุณลักษณะดังกล่าว โดยอาศัยผลที่ได้จากการตอบแบบสอบ (test performance) หรือคะแนน (score) ซึ่งเป็นสิ่งที่สามารถสังเกตและวัดได้ (observable) (Lord and Novick, 1968; Hambleton and Cook, 1977; Hambleton and Swaminatan, 1985)

นักวัดผลจึงได้พยายามหาความสัมพันธ์ระหว่างผลที่ได้จากการตอบแบบสอบหรือคะแนน (test performance or score) กับระดับความสามารถ (ability) ของผู้ตอบแต่ละคน เพื่อเขียนเป็น

โมเดลทางคณิตศาสตร์ (mathematical model) (Hambleton and Cook, 1977; Hambleton and Swaminatan, 1985)

ความสัมพันธ์ระหว่างผลที่ได้จากการตอบแบบสอบกับระดับความสามารถของผู้สอบสามารถเขียนในรูปของความสัมพันธ์ ได้ดังนี้

$$P = f(U_i / \theta_1, \theta_2, \theta_3, \ldots \theta_k ; \beta_k)$$

เมื่อ P แทน ผลจากการตอบแบบสอบ ((test performance)

f แทน ฟังก์ชัน (function)

$U_i$ แทน ผลการตอบข้อสอบข้อที่ I (ตอบถูก $U_i = 1$ ตอบผิด $U_i = 0$)

$\theta_1, \theta_2, \theta_3, \ldots \theta_k$ แทน ระดับความสามารถ (ability) ที่ 1, 2, 3,…,k

$\beta_k$ แทน ค่าพารามิเตอร์ของข้อสอบข้อที่ j

เนื่องจากความสัมพันธ์ดังกล่าวเป็นเพียงฟังก์ชันความสัมพันธ์ในลักษณะทั่ว ๆ ไป นักวัดผลการศึกษาจึงต้องหาโมเดลทางคณิตศาสตร์ที่เหมาะสม เพื่อใช้แทนฟังก์ชันความสัมพันธ์ดังกล่าว โดยอาศัยข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

**ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ**

ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบที่สำคัญ ดังนี้ (Lord and Novick, 1968; Hambleton and Swaminathan, 1985; ศิริชัย กาญจนวาสี, 2545)

1. ความเป็นเอกมิติ (unidimentional)

โดยทั่วไปมีข้อสมมุติว่ามีเพียงลักษณะเดียวหรือความสามารถเดียวเท่านั้นที่จำเป็นต่อการอธิบายคะแนนที่ได้จากการตอบของผู้สอบ ตามโมเดลของการตอบสนองรายข้อ กำหนดว่าการที่มีความสามารถเดียวหรือคุณลักษณะเดียวคือความเป็นเอกมิติ โดยทั่วไปแล้วข้อตกลงข้อนี้เป็นไปได้ค่อนข้างยาก เนื่องจากมีปัจจัยที่มีผลต่อคะแนนสอบ เช่น ปัจจัยด้านความรู้ความเข้าใจ (Cognitive) บุคลิกภาพ และปัจจัยเกี่ยวกับการจัดการสอบ ปัจจัยเหล่านี้อาจรวมถึงแรงจูงใจ ความวิตกกังวลในการสอบ ความสามารถในการทำงานได้รวดเร็ว ความรู้เกี่ยวกับการใช้กระดาษคำตอบ เมื่อเป็นเช่นนี้สิ่งที่ทำให้ข้อตกลงนี้เป็นไปได้ คือการพิจารณาว่าแบบสอบฉบับนั้นมีองค์ประกอบใดหรือปัจจัยใดที่เด่นที่สุด ก็ถือว่าแบบสอบได้วัดในสิ่งนั้น

จากฟังก์ชันความน่าจะเป็นในการตอบข้อสอบถูก ณ ระดับความสามารถต่างๆ ($\theta$) ที่แสดงด้วยฟังก์ชันโลจิสติกหรือโอไจฟ์ปกติ ฟังก์ชันทั้งสองยืนยันว่าความน่าจะเป็นในการทำข้อสอบแต่ละข้อถูกขึ้นอยู่กับพารามิเตอร์ของข้อสอบ 3 ตัว ถ้าเราต้องการทราบความสามารถของ

คนเพื่อกำหนดความน่าจะเป็นที่จะทำข้อสอบถูกในข้อใดข้อหนึ่ง ก็จะไม่ช่วยให้ทราบความน่าจะเป็นที่ทำข้อสอบข้ออื่นๆ ถูก ถ้าข้อสอบที่เพิ่มขึ้นเข้าไปไม่ใช่การวัดความสามารถเดียวกับข้อสอบเดิม ผลการสอบจากข้อสอบเหล่านั้นจะขึ้นอยู่กับคุณลักษณะบางอย่างมากกว่าความสามารถของมิติเดียวนี้ หมายถึง ข้อสอบทั้งหลายที่วัดเพียงเนื้อเดียวของความรู้หรือความสามารถ ข้อตกลงเกี่ยวกับความเป็นมิติเดียวนี้เป็นเรื่องที่ซับซ้อนและยุ่งยากมากที่สุด ของทฤษฎีการตอบสนองข้อสอบ ในการทดสอบถึงความเป็นมิติเดียวของแบบสอบส่วนใหญ่แล้วจะใช้วิธีการวิเคราะห์ตัวประกอบ แล้วสังเกตค่าไอเกน (Eigen Value) ของตัวประกอบที่นำมาลงจุดในกราฟกับอันดับของตัวประกอบ ถ้าค่าไอเกนในตัวประกอบแรกมากกว่าตัวประกอบที่ 2 และค่าไอเกนของตัวประกอบที่เหลือมีค่าใกล้เคียงกันจะแสดงถึงความเป็นมิติเดียว

2. ความเป็นอิสระในการตอบข้อสอบ (local independent)

ความเป็นอิสระในการตอบข้อสอบ หมายถึงความน่าจะเป็นในการตอบข้อสอบแต่ละข้อได้ถูกต้องเป็นอิสระจากกัน นั่นคือ การตอบข้อสอบข้อใดข้อหนึ่งถูก หรือผิด จะไม่มีผลกระทบต่อการตอบข้ออื่น ๆ ด้วย หรืออาจจะกล่าวในเชิงคณิตศาสตร์ได้ว่า ความเป็นอิสระในการตอบข้อสอบ หมายถึง ความน่าจะเป็นในการตอบข้อสอบถูกทั้งหมดมีค่าเท่ากับผลคูณของความน่าจะเป็นในการตอบข้อสอบถูกเป็นรายข้อ นั่นคือ ผู้สอบที่มีความสามารถ ($\theta$) จะมีความน่าจะเป็นที่จะตอบข้อสอบทั้งข้อ 1 และข้อ 2 ถูกเท่ากัน ซึ่งได้มาจากความน่าจะเป็นในการตอบข้อสอบข้อที่ 1 ถูก และความน่าจะเป็นในการตอบข้อสอบข้อที่ 2 ถูก คือ ถ้าผู้สอบมีความสามารถ ($\theta$) = 1.5 มีความน่าจะเป็นในการตอบข้อสอบข้อที่ 1 ถูกเท่ากับ 0.5 และมีความน่าจะเป็นในการตอบข้อสอบข้อที่ 2 ถูกเท่ากับ 0.6 ดังนั้นผู้สอบที่มีความสามารถ ($\theta$) = 1.5 มีความน่าจะเป็นในการตอบข้อสอบทั้งสองข้อถูกภายใต้เงื่อนไขความเป็นอิสระมีค่าเท่ากับ 0.3 = (0.5)(0.6)

อย่างไรก็ตาม Hambleton และ Swaminathan, (1985) กล่าวว่า ถ้าแบบสอบมีความเป็นเอกมิติอยู่แล้ว ความเป็นอิสระในการตอบข้อสอบก็จะเกิดขึ้นตามไปด้วย

3. โค้งคุณลักษณะข้อสอบ (item characteristic curve)

โค้งคุณลักษณะข้อสอบเป็นฟังก์ชันทางคณิตศาสตร์ สามารถใช้อธิบายความสัมพันธ์ระหว่างความน่าจะเป็นหรือโอกาสที่ผู้สอบจะตอบข้อสอบถูกกับระดับความสามารถที่วัดได้โดยใช้ชุดของข้อสอบหรือแบบสอบฉบับนั้น ทั้งนี้ความน่าจะเป็นหรือโอกาสในการตอบข้อสอบถูกจะขึ้นอยู่กับโค้งลักษณะข้อสอบในแต่ละโมเดลที่เลือกใช้ โดยที่รูปร่าง (shape) ของโค้งคุณลักษณะข้อสอบในแต่ละข้อมีคุณสมบัติไม่แปรเปลี่ยน (invariant) ไปตามกลุ่มตัวอย่างที่ใช้ ดังนั้นจึงทำให้ความน่าจะเป็นหรือโอกาสในการตอบข้อสอบถูกในแต่ละข้อไม่แปรเปลี่ยนด้วย คุณสมบัตินี้ถือเป็นลักษณะเด่นของ

โมเดลต่าง ๆ ในทฤษฎีการตอบสนองข้อสอบ โค้งคุณลักษณะข้อสอบมีหลายรูปแบบขึ้นอยู่กับว่าเลือกใช้พารามิเตอร์ของข้อสอบกี่พารามิเตอร์

4. ข้อสอบที่ใช้ต้องไม่เป็นข้อสอบประเภทความเร็ว (Speededness)

ผู้สอบทุกคนควรมีโอกาสในการทำข้อสอบทุกข้อ เพื่อให้คะแนนรวมจากการสอบเป็นค่าความสามารถที่แท้จริงของผู้สอบ ไม่มีข้อจำกัดเกี่ยวกับเวลาในการสอบ

**พารามิเตอร์ในทฤษฎีการตอบสนองข้อสอบ**

พารามิเตอร์ในทฤษฎีการตอบสนองข้อสอบ แบ่งออกเป็น 2 ชนิด คือ พารามิเตอร์ข้อสอบ (item parameter) ได้แก่ ความยาก ($b_i$) อำนาจจำแนก ($a_i$) พารามิเตอร์การเดา ($c_i$) และความรอบคอบ ($\gamma_i$) ส่วนพารามิเตอร์ของผู้สอบ (person parameter) ได้แก่ ระดับความสามารถหรือคุณลักษณะของผู้สอบ ($\theta$) ซึ่งพิสัยของค่าพารามิเตอร์ต่าง ๆ มีดังนี้ (Hambleton and Swaminathan, 1985; ศิริชัย กาญจนวาสี, 2545)

1. พารามิเตอร์ความยาก ($b_i$) ในทางทฤษฎีมีค่าตั้งแต่ $-\infty$ ถึง $\infty$ แต่ในทางปฏิบัติจะมีค่าอยู่ระหว่าง -2.5 ถึง +2.5 ค่าที่เป็นลบแสดงว่าข้อสอบง่าย และค่าที่เป็นบวกแสดงว่าข้อสอบยาก

2. พารามิเตอร์อำนาจจำแนก ($a_i$) ในทางทฤษฎีมีค่าตั้งแต่ $-\infty$ ถึง $\infty$ ควรมีค่าเป็นบวก ตามปกติมีค่าไม่เกิน +2.5 ในทางปฏิบัตินิยมใช้ข้อสอบที่มีค่าพารามิเตอร์อำนาจจำแนกอยู่ระหว่าง +0.5 ถึง +2.5

3. พารามิเตอร์การเดา ($c_i$) เป็นค่าแสดงความน่าจะเป็นหรือโอกาสของการตอบข้อสอบได้ถูกต้อง โดยไม่มีความรอบรู้หรือคุณลักษณะในเรื่องนั้นๆ ในทางทฤษฎีพารามิเตอร์การเดามีค่าระหว่าง 0 ถึง 1 โดยทั่วไปนิยมใช้ข้อสอบที่มีค่าพารามิเตอร์การเดาไม่เกิน 0.30

4. ความรอบคอบ ($\gamma_i$) McDonald (1967) และ Barton และ Lord (1981) ได้เสนอพารามิเตอร์ที่แสดงถึงความรอบคอบของผู้สอบ เป็นค่าพารามิเตอร์ที่บ่งชี้ว่าผู้สอบที่มีความสามารถสูงอาจจะตอบข้อสอบได้ไม่ถูกต้องเสมอไป ซึ่งอาจเกิดความไม่รอบคอบในการพิจารณาคำตอบ หรือผู้สอบอาจจะมีสารสนเทศอื่น ๆ เกี่ยวกับผู้ออกข้อสอบทำให้เลือกตอบในตัวเลือกที่ไม่ใช่คำตอบที่ถูกต้อง Barton และ Lord (1981) กล่าวว่า พารามิเตอร์ตัวนี้จะเหมาะสมในการศึกษาทางทฤษฎีเท่านั้น ซึ่งในทางปฏิบัติแล้วไม่สามารถพบพารามิเตอร์นี้ได้ (Hambleton and Swaminatan, 1985)

5. พารามิเตอร์ผู้สอบ เป็นระดับความสามารถของผู้สอบ ($\theta$) ที่ประมาณได้จากโมเดลตามทฤษฎีการตอบสนองข้อสอบ นิยมปรับให้เป็นคะแนนมาตรฐานที่มีค่าเฉลี่ยเป็น 0 และส่วน

เบี่ยงเบนมาตรฐานเป็น 1 ซึ่งพารามิเตอร์ผู้สอบมีค่าระหว่าง -∞ ถึง ∞ แต่ส่วนใหญ่จะมีค่าอยู่ในช่วง -3.0 ถึง +3.0 ค่าที่เป็นลบแสดงว่าผู้สอบมีความสามารถต่ำ และค่าที่เป็นบวกแสดงว่าผู้สอบมีความสามารถสูง

ข้อตกลงเบื้องต้นและพารามิเตอร์ที่กล่าวมานี้ มีความหมายเด่นชัดในกรณีที่ข้อสอบนั้นให้คะแนนแบบสองค่า ในการประยุกต์ทฤษฎีเพื่อใช้กับข้อสอบที่ให้คะแนนแบบหลายค่า ข้อตกลงเบื้องต้นทั้งหมดก็เทียบเคียงในทำนองเดียวกัน แตกต่างกันเพียงรายละเอียดปลีกย่อยเกี่ยวกับเงื่อนไขเฉพาะของแต่ละโมเดลเท่านั้น

**ฟังก์ชันสารสนเทศ** (Information Function)

ฟังก์ชันสารสนเทศจะเกี่ยวข้องกับการคัดเลือกข้อสอบ การพัฒนาแบบสอบ และการประเมินความแม่นยำของการประมาณค่าความสามารถของผู้สอบ ซึ่งในที่นี้จะกล่าวถึงฟังก์ชันสารสนเทศของข้อสอบ และฟังก์ชันสารสนเทศของแบบสอบ

***ฟังก์ชันสารสนเทศของข้อสอบ*** *(item information function)*

การอธิบายข้อสอบและแบบสอบ การเปรียบเทียบประสิทธิภาพของแบบสอบของทฤษฎีการตอบสนองข้อสอบและการเปรียบเทียบประสิทธิภาพของโมเดล ต้องอาศัยคุณสมบัติของฟังก์ชันสารสนเทศของข้อสอบ ฟังก์ชันสารสนเทศของข้อสอบเป็นความสัมพันธ์ของอัตราส่วนระหว่างกำลังสองค่าอนุพันธ์ของโอกาสในการตอบข้อสอบถูกของผู้สอบที่ระดับความสามารถนั้นๆ กับผลคูณของโอกาสในการตอบข้อสอบถูกและผิดของผู้สอบในระดับความสามารถนั้น ๆ เขียนเป็นสมการได้ดังนี้

$$I_i(\theta) = \frac{\left[P_i'(\theta)\right]^2}{P_i(\theta)Q_i(\theta)} \quad , i = 1, 2, 3, \ldots, k$$

เมื่อ $I_i(\theta)$ = ค่าฟังก์ชันสารสนเทศ หรือค่าสารสนเทศที่ได้รับจากข้อสอบข้อที่ i สำหรับผู้ตอบที่มีความสามารถ $\theta$

$P_i'(\theta)$ = $P_i'$ = ความชันของฟังก์ชันการตอบข้อสอบข้อที่ i ณ ตำแหน่งความสามารถ $\theta$

$P_i(\theta)$ = $P_i$ = ความน่าจะเป็นที่ผู้ตอบซึ่งมีความสามารถ $\theta$ จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

$Q_i(\theta)$ = $Q_i = 1 - P_i(\theta)$

ฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อรวมกันเป็นฟังก์ชันสารสนเทศของแบบสอบ ซึ่งค่าฟังก์ชันสารสนเทศของข้อสอบขึ้นอยู่กับค่าความชันของฟังก์ชันการตอบข้อสอบ และค่าความ

แปรปรวนที่เงื่อนไขแต่ละระดับของความสามารถหรือคุณลักษณะ ค่าความชันสูงและค่าความแปรปรวนต่ำทำให้ค่าสารสนเทศของข้อสอบมีค่าสูง และทำให้ค่าของความคลาดเคลื่อนมาตรฐานในการวัดมีค่าต่ำ การแจกแจงของฟังก์ชันสารสนเทศของข้อสอบมีลักษณะเป็นรูประฆังคว่ำ ค่าสารสนเทศที่สูงที่สุดจะอยู่ที่จุด $b_i$ บนสเกลความสามารถสำหรับโมเดลการตอบแบบโลจีสติกแบบ 1 และ 2 พารามิเตอร์ ส่วนโมเดล 3 พารามิเตอร์นั้น ค่าสารสนเทศของข้อสอบข้อที่ i จะสูงสุดที่จุด $\theta_{max}$ เมื่อ

$$\theta_{max} = b_i + \frac{1}{Da_i}\left[\ln\left(\frac{1+\sqrt{(1+8c_i)}}{2}\right)\right]$$

สำหรับโมเดล 1 พารามิเตอร์นั้นค่าสูงสุดของสารสนเทศของข้อสอบจะคงที่ และขณะเดียวกันโมเดล 2 พารามิเตอร์ค่าสูงสุดของสารสนเทศของข้อสอบจะเป็นสัดส่วนโดยตรงกับกำลังสองของค่าอำนาจจำแนก ถ้าค่าอำนาจจำแนกของข้อสอบสูงก็จะทำให้ค่าสารสนเทศของข้อสอบมีค่ามาก ส่วนโมเดล 3 พารามิเตอร์นั้นค่าสารสนเทศของข้อสอบสูงสุดจะมีค่าดังนี้

$$I_i(\theta)_{max} = D^2a^2\frac{\left[1-20c_i-8c_i^2+(1+8c_i)^{3/2}\right]}{8(1-c_i^2)}$$

ถ้าค่า $c_i$ ลดลง ค่าสารสนเทศของข้อสอบก็จะเพิ่มขึ้น

ลักษณะสารสนเทศของข้อสอบ

1. ผลรวมค่าสารสนเทศของข้อสอบทุกข้อคือสารสนเทศของแบบสอบ
2. ค่าฟังก์ชันสารสนเทศขึ้นอยู่กับค่าความชันของฟังก์ชันการตอบข้อสอบ และค่าความแปรปรวนที่มีเงื่อนไขที่แต่ละระดับของความสามารถ ถ้าค่าความชันมากและค่าความแปรปรวนต่ำทำให้ค่าสารสนเทศของข้อสอบมีค่าสูงซึ่งทำให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดมีค่าต่ำ

***ฟังก์ชันสารสนเทศของแบบสอบ (Test Information)***

การวิเคราะห์ตามทฤษฎี IRT จะใช้แบบแผนการตอบสนองแบบสอบเป็นรายข้อในการประมาณค่าความสามารถของผู้สอบ ดังนั้นการประเมินคุณภาพของแบบสอบจึงสามารถพิจาณาความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้ตอบ โดยใช้ดัชนีตัวหนึ่งเรียกว่าสารสนเทศของแบบสอบ (Test Information ; $I(\theta)$) ซึ่งเป็นค่าฟังก์ชันสารสนเทศของแบบสอบอันเกิดจากผลรวมเชิงพีชคณิตของฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อรวมเข้าด้วยกันทั้งฉบับ ณ ตำแหน่ง $\theta$ เดียวกัน (ศิริชัย กาญจนวาสี, 2545) ดังสูตร

$$I(\theta) = \sum_{i=1}^{k} I_i(\theta)$$

ลักษณะของฟังก์ชันสารสนเทศของแบบสอบ

1. ฟังก์ชันสารสนเทศของแบบสอบเป็นสิ่งที่ถูกกำหนดขึ้นสำหรับชุดของข้อสอบที่แต่ละจุดของสเกลความสามารถ

2. ค่าสารสนเทศของแบบสอบเป็นผลมาจากคุณภาพและจำนวนของข้อสอบ

3. ณ ตำแหน่งความสามารถเดียวกัน เส้นถดถอยที่มีความชันมากกว่าจะให้ค่าสารสนเทศของแบบสอบสูงกว่าเส้นถดถอยที่มีความชันน้อยกว่า

4. ข้อสอบที่มีค่าความแปรปรวนต่ำจะส่งผลให้ค่าสารสนเทศของแบบสอบสูง

5. ค่าสารสนเทศของแบบสอบจะไม่ขึ้นอยู่กับการจัดหมวดหมู่เฉพาะของข้อสอบ และข้อสอบแต่ละข้อเป็นอิสระจากกัน

6. ค่าสารสนเทศของแบบสอบมีความสัมพันธ์แบบผกผันกับค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถที่ระดับเดียวกัน ดังสมการ

$$SE(\theta) = \frac{1}{\sqrt{I(\theta)}}$$

## โมเดลการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ (IRT) เสนอฟังก์ชันสำหรับอธิบายความสัมพันธ์ระหว่างความน่าจะเป็นของการตอบข้อสอบได้ถูกต้อง (หรือการเลือกรายการคำตอบ) กับความสามารถของผู้ตอบ (หรือคุณลักษณะภายในของผู้ตอบ) และคุณลักษณะข้อสอบ (หรือรายการคำตอบ) โดยใช้โมเดล 2 ประเภท ตามลักษณะการตรวจให้คะแนนคำตอบ ได้แก่ โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า (Binary or dichotomous IRT Models) และโมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า (polytomous IRT Models) (ศิริชัย กาญจนวาสี, 2545)

**โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า** (ศิริชัย กาญจนวาสี, 2545; Hambleton and Swaminatan, 1985)

เนื่องจากข้อมูลที่ได้จากการสอบมีหลายลักษณะ ได้แก่ ข้อมูลแบบทวิภาค (dichotomous) และข้อมูลแบบพหุภาค (polytomous) ดังนั้นจึงมีการพัฒนารูปแบบเพื่อให้สอดคล้องกับลักษณะของข้อมูลดังกล่าวขึ้นมากมาย แต่สำหรับข้อมูลที่เป็นแบบทวิภาค รูปแบบที่นิยมใช้เป็นรูปแบบโลจีสติก (Logistic model) ซึ่งแตกต่างไปตามจำนวนพารามิเตอร์ที่ใช้ในแต่ละรูปแบบ มีรายละเอียดดังนี้

1. **รูปแบบโลจีสติก 1 พารามิเตอร์** (One-Parameter Logistic Model)

รูปแบบนี้บางครั้งเรียกว่า รูปแบบราสช์ (Rasch Model) เนื่องจากรูปแบบนี้ได้พัฒนาโดยนักวิทยาศาสตร์ชาวเดนมาร์ค ชื่อ Georg Rasch ในปี ค.ศ. 1966 โค้งคุณลักษณะข้อสอบตามรูปแบบนี้คือ

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

ถึงแม้ว่ารูปแบบนี้จะเป็นกรณีเฉพาะของรูปแบบสองพารามิเตอร์ และสามพารามิเตอร์ แต่ก็ยังมีคุณสมบัติพิเศษที่ทำให้นิยมใช้กันคือ ประการแรก เนื่องจากรูปแบบนี้มีจำนวนพารามิเตอร์ไม่มากจึงสะดวกต่อการใช้งาน ประการที่สอง ปัญหาที่เกิดจากการประมาณค่าพารามิเตอร์มีน้อยกว่า การประมาณค่าพารามิเตอร์สำหรับรูปแบบที่มีพารามิเตอร์หลาย ๆ ตัว

2. **รูปแบบโลจีสติก 2 พารามิเตอร์** (Two-Parameter Logistic Model)

รูปแบบนี้เสนอโดย Birnbaum เมื่อปี ค.ศ. 1975 เป็นโค้งคุณลักษณะข้อสอบ และเป็นฟังก์ชันของการแจกแจงที่มี 2 พารามิเตอร์ คือพารามิเตอร์ความยาก และพารามิเตอร์อำนาจจำแนก และโมเดล 2PL มีความเหมาะสมสำหรับการวัดคุณลักษณะแฝงที่แต่ละคนมีไม่เท่ากัน

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

เมื่อ $P_i(\theta)$ เป็นค่าความน่าจะเป็นของผู้สอบที่มีความสามารถ $\theta$ สามารถตอบข้อสอบข้อที่ i ได้ถูกต้อง ส่วน $b_i$ และ $a_i$ เป็นพารามิเตอร์ของข้อสอบข้อที่ i และ D เป็นค่าสเกลองค์ประกอบ (Scaling Factor) ซึ่งถ้ากำหนดให้มีค่าเท่ากับ 1.7 แล้วค่า $P_i(\theta)$ จากโค้งความถี่สะสมกับโค้งโลจีสติกจะมีค่าที่ต่างกันน้อยกว่า 0.01 สำหรับทุกค่าของ $\theta$

จากรูปแบบนี้อยู่บนข้อตกลงที่ว่าการเดาคำตอบจะไม่เกิดขึ้น ซึ่งถ้าจะเป็นเช่นนี้ได้ก็ต่อเมื่อค่าพารามิเตอร์ $a_i > 0$ (ข้อสอบที่มีความสัมพันธ์ด้านบวกระหว่างคะแนนจากการสอบกับความสามารถของผู้สอบที่วัดโดยแบบสอบนั้น) และค่าความน่าจะเป็นในการตอบข้อสอบได้ถูกจะลดลงถึงศูนย์เมื่อความสามารถลดลง

3. **รูปแบบโลจีสติก 3 พารามิเตอร์** (Three-Parameter Logistic Model)

รูปแบบโลจีสติกสามพารามิเตอร์ เป็นการปรับปรุงสองพารามิเตอร์เพียงแต่เพิ่มพารามิเตอร์ตัวที่สามคือ พารามิเตอร์การเดาคำตอบ หรือพารามิเตอร์ $c_i$ เข้าไปในรูปแบบนี้ ดังในข้อสอบแบบหลายตัวเลือก ความน่าจะเป็นของการตอบถูกมากกว่า 0 แม้ว่าผู้สอบจะมีความสามารถต่ำก็ตาม เขียนในรูปแบบสมการเชิงคณิตศาสตร์ได้ดังนี้

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

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นของผู้สอบที่มีความสามารถ $\theta$ จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

$b_i$ คือ พารามิเตอร์ความยาก

$a_i$ คือ พารามิเตอร์อำนาจจำแนก

$c_i$ คือ พารามิเตอร์โอกาสการเดาข้อสอบได้ถูก

D คือ ค่าสเกลองค์ประกอบ (D = 1.7)

พารามิเตอร์ $c_i$ เป็นจุดต่ำสุดที่โค้งคุณลักษณะข้อสอบ ซึ่งพารามิเตอร์นี้จะใช้เมื่อคิดว่าการเดาเป็นองค์ประกอบในการตอบข้อสอบ บางครั้งเรียกพารามิเตอร์นี้ว่าโอกาสที่จะตอบข้อสอบได้ถูกต้องสำหรับคนที่มีความสามารถต่ำ

ในการปรับรูปแบบสามพารามิเตอร์ให้เป็นรูปแบบสองพารามิเตอร์ ต้องอยู่บนข้อตกลงที่ว่าพารามิเตอร์การเดามีค่าเท่ากับศูนย์

**โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า**

โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า เป็นโมเดลความสัมพันธ์ที่ไม่ใช่เชิงเส้นตรงระหว่างความสามารถของผู้ตอบกับโอกาสของการเลือกตอบแต่ละรายการคำตอบที่กำหนดให้ มีผู้พัฒนาโมเดลการตอบสนองข้อสอบสำหรับการตรวจให้คะแนนรายการคำตอบมากกว่า 2 ค่าไว้หลายโมเดล (Hambleton and Swaminatan, 1985; Embretson and Reise, 2000; ศิริชัย กาญจนวาสี, 2545) ดังนี้

**Graded – response Model**

Samejima (1969, 1996) ได้พัฒนา Graded – response model (GRM) มีความเหมาะสมสำหรับการให้คะแนนที่เป็นลำดับขั้น ใช้กับแบบสอบหรือแบบวัดที่แต่ละข้อคำถามมีรายการคำตอบแบบมาตรเรียงลำดับ (ordered categorical response) โดยแบบสอบชุดเดียวกันอาจมีจำนวนรายการคำตอบในแต่ละข้อแตกต่างกันได้ เช่น การตรวจให้คะแนนความรู้บางส่วนที่แต่ละข้อมีจำนวนลำดับขั้นของการตรวจให้คะแนนแตกต่างกัน เป็นต้น โมเดล GRM นี้ ได้พัฒนามาจากโมเดลการตอบข้อสอบแบบให้คะแนนสองค่าแบบ 2 พารามิเตอร์ โดยใช้ค่าพารามิเตอร์ความชัน (slope parameter) ในการอธิบายข้อสอบแต่ละข้อ และค่าพารามิเตอร์ Threshold (Threshold parameter) จำนวน m ค่า โดยที่ m มีค่าเท่ากับ k – 1 เมื่อ k เป็นจำนวนรายการคำตอบในแต่ละข้อ และใช้หลักการคำนวณความน่าจะเป็นของการตอบแต่ละรายการคำตอบแบบ 2 ขั้นตอน (indirect IRT Model) โดยขั้นตอนแรกเป็นการประมาณค่าความชัน (slope parameter: $\alpha$) 1 ค่า แล้วจึงประมาณ

ค่า Threshold (Threshold parameter: $\beta_{ij}$) m ค่า เพื่อให้ได้โค้งลักษณะปฏิบัติการ (operating characteristic curves, OCC) จำนวน m โค้งของข้อคำถามแต่ละข้อ

***โมเดล***

ใน GRM คำถามแต่ละข้อ (i) อธิบายได้ด้วยความชันร่วมของคำถาม (common item slope parameter, $\alpha_i$) และค่า Threshold ของแต่ละรายการคำตอบ (category threshold parameter, $\beta_{ij}$) เมื่อ $j = 1,\ldots,m_i$ เป็นจำนวนของ threshold ของข้อ i และจำนวนรายการคำตอบของข้อ i ($K_i$) = $m_i + 1$ ดังตัวอย่าง

ดังนั้น คำถามข้อนี้มี K = 5 Categories (0,1,2,3,4)

m = 4 Thresholds (1,2,3,4)

การวิเคราะห์ตามโมเดล GRM จึงมีเป้าหมายเพื่อประมาณค่า $\alpha_i$ และตำแหน่งของ $\beta_{ij}$ ของผู้ตอบที่มีค่าคุณลักษณะ ($\theta$) บนสเกลที่ต่อเนื่องกัน โดยใช้สูตรดังนี้

$$P^{*}_{ix}(\theta) = \frac{\exp\left[\alpha_i(\theta - \beta_{ij})\right]}{1 + \exp\left[\alpha_i(\theta - \beta_{ij})\right]}$$

เมื่อ $x = j = 1,\ldots,m$

$P^{*}_{ix}(\theta)$ = ความน่าจะเป็นที่ผู้ตอบซึ่งมีคุณลักษณะระดับ $\theta$ จะตอบข้อ i ด้วยการเลือกรายการคำตอบที่ x เมื่อ x = 1,2,…,m

$\alpha_i$ = ค่าพารามิเตอร์ ความชันร่วม (slope parameter) ของข้อที่ i

$\beta_{ij}$ = ค่าพารามิเตอร์ Threshold ของแต่ละรายการคำตอบ (Threshold parameter) ของข้อที่ i

ค่า $\alpha_i$ คล้ายกับค่าอำนาจจำแนกของข้อสอบตามทฤษฎีการทดสอบแบบดั้งเดิม แต่ไม่ควรพิจารณาโดยตรงว่าเป็นอำนาจจำแนกของข้อสอบ เพราะการประเมินขนาดความสามารถในการจำแนก จำเป็นต้องคำนวณจากค่าสารสนเทศของข้อสอบที่ระดับ $\theta$ ของผู้สอบ ส่วน

ค่าพารามิเตอร์ Threshold ($\beta_{ij}$) คือค่าที่แสดงระดับ $\theta$ ที่จำเป็นต้องมี เพื่อให้มีโอกาสในการตอบเหนือ Thresholds j ด้วยความน่าจะเป็นในการตอบเท่ากับ 0.5

โค้งแสดงฟังก์ชันของ $P_{ix}^{*}(\theta)$ เรียกว่าโค้งลักษณะปฏิบัติการ (operating characteristic curves, OCC) ซึ่งต้องคำนวณแต่ละโค้งที่แยกระหว่างรายการคำตอบ ดังนั้นจึงต้องประมาณค่า $\beta_{ij}$ ตามตัวอย่างข้อคำถาม จำนวน 4 ค่า และ $\alpha_i$ จำนวน 1 ค่าที่ร่วมกันของแต่ละข้อ โดย $\beta_{ij}$ มีความหมายคล้ายเป็นระดับค่า $\theta$ ที่จำเป็นจะต้องมีเพื่อให้มีโอกาสตอบเหนือ Threshold j ด้วยความน่าจะเป็น 0.50 หรือ 50%

ขั้นตอนที่ 2 ในการวิเคราะห์ความสัมพันธ์ระหว่างระดับความสามารถกับความน่าจะเป็นของการเลือกรายการคำตอบในแต่ละข้อ ในโมเดล Grade Response แสดงได้ด้วยโค้งเลือกรายการคำตอบ (Category response curves)

ในการคำนวณค่าพารามิเตอร์ Threshold 4 ค่า ได้แก่ 1) การตอบ 0 เทียบกับ 1,2,3,4 2) การตอบ 0,1 เทียบกับ 2,3,4 3) การตอบ 0,1,2 เทียบกับ 3,4 4) การตอบ 0,1,2,3 เทียบกับ 4 ด้วยการใช้โมเดล 2 พารามิเตอร์ สำหรับแต่ละรายการคำตอบภายใต้เงื่อนไขความชันของโค้ง OCC ที่เท่ากันสำหรับแต่ละข้อ ทำให้สามารถคำนวณความน่าจะเป็นในการตอบแต่ละรายการคำตอบของผู้มีคุณลักษณะ $\theta$ สำหรับ x = 0,1,2,3,4 โดยใช้การลบค่าความน่าจะเป็นดังสมการ

$$P_{ix}(\theta) = P_{ix}^{*}(\theta) - P_{i(x+1)}^{*}(\theta)$$

เมื่อพิจารณาจากข้อเท็จจริง ความน่าจะเป็นของการเลือกตอบ รายการคำตอบต่ำสุดหรือเหนือกว่าจะมีค่าเป็น 1.00 นั่นคือ $P_{i0}^{*}(\theta) = 1.00$ และความน่าจะเป็นของการเลือกตอบเหนือกว่ารายการคำตอบสูงสุด จะมีค่าเป็น 0 ดังนั้น $P_{i5}^{*}(\theta) = 0$ จากตัวอย่างคำถามที่ใช้ ความน่าจะเป็นของการเลือกตอบแต่ละรายการคำตอบของผู้สอบที่มีคุณลักษณะ $\theta$ เป็นดังต่อไปนี้

$$P_{i0}(\theta) = 1.0 - P_{i0}^{*}(\theta)$$
$$P_{i1}(\theta) = P_{i1}^{*} - P_{i2}^{*}(\theta)$$
$$P_{i2}(\theta) = P_{i2}^{*} - P_{i3}^{*}(\theta)$$
$$P_{i3}(\theta) = P_{i3}^{*} - P_{i4}^{*}(\theta)$$
$$P_{i4}(\theta) = P_{i4}^{*} - 0$$

**Modified Graded – response Model**

Muraki (1990) ได้ปรับปรุงโมเดล GRM จึงเรียกว่า Modified Graded – response Model (M-GRM) เพื่อให้สะดวกแก่การใช้กับแบบวัดประเภทมาตรประมาณค่า (Rating scale) ที่นิยมให้มีจำนวนรายการคำตอบที่เท่ากัน หรือมีรูปแบบการตอบที่คงที่สำหรับทุกข้อคำถาม เช่น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปแบบการตอบเป็นมาตรประมาณค่า 5 ระดับของลิเคิร์ท (Likert rating scale) เป็นต้น โดย M-GRM มีลักษณะบางประการเหมือนกับ GRM คือ แต่ละข้อคำถามมีการประมาณค่าความชันและค่าความชันของแต่ละข้อมีค่าไม่เท่ากัน แต่แตกต่างกันตรงที่ค่าพารามิเตอร์ Threshold ($\beta_{ij}$) ใน M-GRM แบ่งเป็น 2 ส่วน คือ location parameter ($b_i$) โดยมีจำนวนเท่ากับจำนวนข้อคำถามคือข้อคำถามแต่ละข้อจะมีค่า $b_i$ จำนวน 1 ค่า และค่าลำดับขั้นของพารามิเตอร์ Threshold (category threshold parameter: $c_j$) มีค่าไม่เท่ากันในแต่ละข้อคำถาม โดยที่ $\beta_{ij} = b_i - c_j$ ทำให้ M-GRM มีความจำกัดในช่วงห่างของรายการคำตอบมากกว่าและเหมือนกันทุกข้อคำถาม ทำให้มีการประมาณค่าพารามิเตอร์น้อยกว่า GRM

M – GRM มีลักษณะเป็นโมเดลเฉพาะของโมเดล GRM โดยสามารถนำไปใช้กับข้อคำถามที่มีรายการคำตอบแบบมาตรเรียงลำดับ ที่มีจำนวนรายการคำตอบเท่ากันทุกข้อคำถาม หรือมีรูปแบบการตอบที่คงที่เหมือนกันทุกข้อ สำหรับการคำนวณความน่าจะเป็นของการตอบแต่ละรายการคำตอบใช้วิธี 2 ขั้นตอน (indirect method) เหมือนโมเดล GRM

***โมเดล***

ใน M – GRM ลักษณะข้อคำถามแต่ละข้อ (i) อธิบายได้ด้วยความชันร่วมของข้อคำถาม 1 ค่า (common item slope parameter, $\alpha_i$) และ threshold parameters ($\beta_{ij}$) ของแต่ละรายการคำตอบ จำนวน $m_i$ แต่ $\beta_{ij}$ สามารถแยกได้เป็น 2 ส่วนคือ ค่าความยากของข้อสอบ (item location parameter, $b_i$) และชุดของ threshold parameter สำหรับแต่ละรายการคำตอบทั้งฉบับ ($c_i$) ดังนี้

$$\beta_{ij} = b_i - c_i$$

โค้งคุณลักษณะปฏิบัติการ (OCC) ของโมเดล M – GRM สามารถเขียนได้ดังสมการ

$$P_{ix}^{*}(\theta) = \frac{\exp\left[\alpha_i(\theta - (b_i - c_j))\right]}{1 + \exp\left[\alpha_i(\theta - (b_i - c_j)\right]}$$

หรือ $$P_{ix}^{*}(\theta) = \frac{\exp\left[\alpha_i(\theta - (b_i + c_j))\right]}{1 + \exp\left[\alpha_i(\theta - (b_i + c_j)\right]}$$

เมื่อ $P_{ix}^{*}(\theta)$ = ความน่าจะเป็นที่ผู้ตอบซึ่งมีคุณลักษณะ $\theta$ จะตอบข้อ i ด้วยการเลือกรายการคำตอบที่ x หรือสูงกว่า เมื่อ x = 1,2,…,$m_i$

$\alpha_i$ = ค่าพารามิเตอร์ความชันร่วมของข้อคำถามข้อที่ i

$b_i$ = ค่าพารามิเตอร์ตำแหน่งโค้ง OCC หรือค่าความยากของข้อคำถามที่ i

$c_i$ = ค่าพารามิเตอร์ Threshold สำหรับแต่ละรายการคำตอบทั้งฉบับ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อพิจารณาจากสมการ M – GRM จึงเป็นโมเดลที่มีความเฉพาะ เพราะมีสมมติฐานว่าขอบเขตของแต่ละรายการคำตอบที่มีระยะห่างจากกันและกันเท่านั้น ซึ่งแตกต่างจาก GRM ที่ระยะดังกล่าวปล่อยให้เป็นค่าอิสระที่แตกต่างกันได้ระหว่างข้อคำถาม ผลที่ตามมาทำให้โมเดล M –GRM มีจำนวนค่าพารามิเตอร์ที่ต้องคำนวณน้อยกว่าโมเดล GRM และมีข้อดีตรงที่การคำนวณค่า $b_i$ แยกออกจาก $\beta_{ij}$ แต่โมเดล GRM จะสะดวกกว่าในกรณีที่แต่ละข้อคำถามมีจำนวนรายการคำตอบแตกต่างกัน ทำให้สามารถนำค่า $b_i$ มาใช้เรียงลำดับความยากของข้อคำถามรวมทั้งค่า $c_j$ ใช้บอกระยะห่างของสเกลทางจิตวิทยาคล้ายกับค่าน้ำหนักสเกลตามวิธีของ Thurstone's method

ค่าพารามิเตอร์ของข้อสอบประกอบด้วย พารามิเตอร์ความชันร่วมของคำถาม ($\alpha$) ค่าพารามิเตอร์ threshold ของแต่ละคำถาม (b) และค่าพารามิเตอร์ threshold ร่วม ของแต่ละรายการคำตอบทั้งฉบับ ($c_i$) ซึ่งรวมกันมีค่าประมาณเท่ากับ 0 และมีค่าเรียงเป็นลำดับ

ค่าพารามิเตอร์ความชันร่วมของคำถาม ($\alpha$) ตาม M – GRM จะมีความผันแปรน้อยกว่าเมื่อเทียบกับ GRM ตามโปรแกรม PARSCALE (Muraki, 1993) จะมีการตรวจสอบความสอดคล้องของโมเดลกับข้อมูลด้วยการทดสอบ $\chi^2$ – test ซึ่งเป็นค่าสถิติที่บ่งชี้ได้ว่าข้อสอบหรือข้อคำถามใดบ้างที่ทำการวิเคราะห์ให้เหมาะสมด้วย M – GRM

M-GRM มีความเหมาะสมในการนำไปใช้ในกรณีที่ต้องการวิเคราะห์ค่าพารามิเตอร์ของเครื่องมือที่มีรายการคำตอบเท่ากันทุกข้อ ในขณะที่ GRM มีความเหมาะสมในการนำไปใช้ในกรณีที่ต้องการวิเคราะห์ค่าพารามิเตอร์ของเครื่องมือที่มีรายการคำตอบไม่เท่ากันทุกข้อ ในกรณีที่เครื่องมือมีรายการคำตอบไม่เท่ากันทุกข้อ สามารถวิเคราะห์โดยใช้ M-GRM ได้โดยต้องจัด block ข้อที่มีรายการคำตอบให้เท่ากัน แต่ก็มีข้อจำกัดในการเปรียบเทียบค่าความชัน ($\alpha_i$) และค่าความยาก ($b_i$) ในข้อที่อยู่ต่าง block เนื่องจากมีค่าลำดับขั้นพารามิเตอร์ threshold ($c_i$) ที่ต่างกัน

**Partial Credit Model**

Master (1982) ได้พัฒนา Partial Credit Model (PCM) สำหรับใช้ในการวิเคราะห์ข้อสอบหรือข้อคำถามที่มีกระบวนการตอบหลายลำดับขั้น ซึ่งจำเป็นต้องมีการตรวจให้คะแนนการตอบถูกต้องหรือการตอบถูกบางส่วนในแต่ละลำดับขั้นของกระบวนการตอบ เช่น ข้อสอบวัดผลสัมฤทธิ์ทางคณิตศาสตร์ ซึ่งมีการให้คะแนนคำตอบถูกบางส่วน แบบวัดเจตคติ บุคลิกภาพ เชาว์ปัญญา ที่มีการให้คะแนนคำตอบเป็นลำดับขั้น เป็นต้น

PCM มีลักษณะเป็นโมเดลที่พัฒนาขยายต่อจากโมเดลการตอบสนองข้อสอบข้อที่มี 1 พารามิเตอร์ (1 – parameter model) จึงมีลักษณะพารามิเตอร์มาตรฐานคล้ายโมเดลของราสช์ (Rasch Model) และใช้หลักการคำนวณความน่าจะเป็นของการตอบแต่ละขั้นการตอบโดยตรงแบบขั้นตอนเดียว (Direct IRT Method)

***โมเดล***

ใน PCM ลักษณะข้อคำถามแต่ละข้อ (i) อธิบายได้ด้วยค่าพารามิเตอร์ระดับความยากระหว่างรายการคำตอบที่อยู่ถัดไป (item step difficulty, $\delta_{ij}$) ถ้ามีค่าสูงแสดงว่ามีความยากมากขึ้นจากระดับรายการคำตอบหนึ่งไปสู่ระดับอื่น โดยทุกข้อมีค่าพารามิเตอร์ความชัน ($\alpha_i$) เท่ากัน (equal slope) ทำให้ค่าความชันไม่ปรากฏในโมเดล สมมติว่าข้อคำถาม i มีคะแนน $x = 0,1,\ldots,m_i$ โดยมีจำนวนรายการคำตอบ $K_i = m_i + 1$

ตัวอย่างคำถาม: ข้อความเจตคติแบบมี 4 รายการคำตอบ

ในการตอบคำถามข้อนี้ ผู้สอบจะเลือกรายการสูงสุดต้องคิด 3 ขั้นตอน โดยทำการตัดสินใจขั้นที่ 1 เลือกระหว่างไม่เคย (0) กับน้อย (1) ขั้นที่สองเลือกระหว่างน้อย (1) กับปานกลาง (2) ขั้นที่สามเลือกระหว่างปานกลาง (2) กับมาก (3) ค่า $\delta_{ij}$ ไม่ได้บอกถึง $\theta$ ของผู้ตอบที่มีโอกาส 0.50 ในการเลือกรายการคำตอบที่อยู่เหนือถัดไปอย่างโมเดล GRM แต่ $\delta_{ij}$ แสดงถึงระดับความยากสัมพัทธ์ของแต่ละขั้นการตอบ ภายในคำถามข้อเดียวกัน แต่ละขั้นตอนอาจง่ายหรือยากกว่าบางขั้นตอน $\delta_{ij}$ ยังแสดงถึงตำแหน่ง $\theta$ ที่โค้งการเลือกรายการคำตอบตัดกัน จึงบ่งบอกถึงผู้สอบที่มีคุณลักษณะ $\theta$ จะมีโอกาสเลือกรายการคำตอบใดมากกว่านั้น

สำหรับ x = j รายการคำตอบ โค้งฟังก์ชันการเลือกรายการคำตอบของโมเดล PCM เขียนได้ดังนี้

$$P_{ix}(\theta) = \frac{\exp\left[\sum_{j=0}^{x}(\theta - \delta_{ij})\right]}{\sum_{r=0}^{m_i}\left[\exp\sum_{j=0}^{r}(\theta - \delta_{ij})\right]}$$

เมื่อ $\sum_{j=0}^{0}(\theta-\delta_{ij})\equiv 0$

$P_{ix}(\theta)$ = ความน่าจะเป็นที่ผู้ตอบซึ่งมีคุณลักษณะ $\theta$ จะตอบข้อ i ด้วยการเลือกหรือสามารถทำรายการคำตอบขั้นที่ x จากจำนวน $m_i$ ขั้น (step)

$\delta_{ij}$ = ค่าพารามิเตอร์ระดับความยากของขั้นการตอบที่ j ในข้อ i (item step difficulty) เมื่อ j = 1,2,…,$m_i$ ค่าที่แสดงถึงขั้นการตอบนั้นมีความยากสัมพัทธ์สูงกว่าขั้นอื่น

ความสัมพันธ์ระหว่างคุณลักษณะ $\theta$ กับระดับขั้นของรายการคำตอบแต่ละข้อคำถามตามโมเดล PCM สามารถนำมาคำนวณคะแนนที่คาดหมาย (Expected score) หรือคะแนนจริง (True score) แต่ละข้อของคำถามได้ตามสูตร ดังนี้

$$E(X)=\sum_{x=0}^{m_i} xp_x(\theta)$$

โค้งของฟังก์ชันดังกล่าวแสดงถึงการแจกแจงคะแนนรายข้อที่คาดหวังของผู้ตอบ ที่มีคุณลักษณะ $\theta$ เมื่อรวมโค้งของทุกข้อเข้าด้วยกันทั้งฉบับ จะทำให้ได้โค้งการแจกแจงคะแนนดิบรวมที่คาดหวังของผู้ตอบที่มีคุณลักษณะ $\theta$ ข้อดีอย่างหนึ่งของโมเดล PCM ซึ่งคล้ายกับโมเดลของราสช์ คือ คะแนนดิบที่คาดหวังเป็นค่าสถิติที่พอเพียงสำหรับการคำนวณค่าของคุณลักษณะ $\theta$ ดังนั้น คะแนนดิบของข้อคำถามชุดเดียวกันที่เป็นไปได้ตามโมเดล PCM จะสอดคล้องหรือสมมูลกับทุกข้อจะต้องมีความสัมพันธ์ที่เท่าเทียมกันกับคุณลักษณะ $\theta$

การแปลความหมายของ $\delta_{ij}$ (category intersection parameter) ใช้สเกลของระดับคุณลักษณะ (latent trait) ณ จุดตัดที่เกิดจากโค้งเลือกรายการคำตอบ (category response curve) 2 โค้งตัดกัน ในกรณีข้อคำถามที่มี x รายการคำตอบ (x = 0, 1, 2, 3) จะมี item step difficulty, $\delta_{ij}$ = m เมื่อ m = x – 1

**Generalized Partial Credit Model**

Muraki (1992, 1993) ได้พัฒนา Generalized Partial Credit Model (G-PCM) ทำให้โมเดล PCM มีลักษณะเป็นโมเดลทั่วไป โดยยอมให้ข้อคำถามแต่ละข้อสามารถมีค่าพารามิเตอร์ความชันแตกต่างกันได้

G-PCM มีลักษณะเป็นโมเดลทั่วไปของ PCM และใช้หลักการคำนวณความน่าจะเป็นของการตอบแต่ละระดับขั้นการตอบโดยตรงแบบขั้นตอนเดียว (Direct IRT Method)

***โมเดล***

ใน G-PCM ลักษณะคำถามแต่ละข้อ (i) อธิบายด้วยค่าพารามิเตอร์ระดับความยากของขั้นการตอบ ($\delta_{ij}$) และค่าพารามิเตอร์ความชันของข้อคำถาม ($\alpha_i$) ฟังก์ชันทางคณิตศาสตร์ของโมเดลนี้ โมเดลนี้วิเคราะห์โดยใช้โปรแกรม PARSCALE ดังสมการ

$$P_{ix}(\theta) = \frac{\exp\left[\sum_{j=0}^{x} \alpha_i(\theta - \delta_{ij})\right]}{\sum_{r=0}^{m}\left[\exp\sum_{j=0}^{r} \alpha_i(\theta - \delta_{ij})\right]}$$

เมื่อ $\sum_{j=0}^{0}(\theta - \delta_{ij}) \equiv 0$

$P_{ix}(\theta)$ = ความน่าจะเป็นที่ผู้ตอบซึ่งมีคุณลักษณะ $\theta$ จะตอบข้อ i ด้วยการเลือกหรือสามารถทำรายการคำตอบขั้นที่ x จากจำนวน $m_i$ ขั้น (step)

$\delta_{ij}$ = ค่าพารามิเตอร์ระดับความยากของขั้นการตอบที่ j ในข้อ i (item step difficulty) เมื่อ $j = 1,2,\ldots,m_i$

$\alpha_i$ = ค่าพารามิเตอร์ความชันของข้อคำถามข้อที่ i

ค่าระดับความยากของขั้นการตอบ ($\delta_{ij}$) ซึ่งเป็นค่าสเกล $\theta$ ตรงตำแหน่งที่ตัดกันของโค้งรายการคำตอบ มีความหมายเหมือนกับโมเดล PCM ส่วนค่าความชัน ($\alpha_i$) มีความหมายต่างจากโมเดลการตอบสนองข้อสอบแบบ 2 ค่า เพราะว่าในโมเดลการตอบสนองข้อสอบแบบมากกว่า 2 ค่า ค่าอำนาจจำแนกของข้อคำถามขึ้นอยู่กับผลรวมกันระหว่างพารามิเตอร์ความชัน และการกระจายของพารามิเตอร์ Threshold ของรายการคำตอบ (สำหรับโมเดล G-PCM) ในโมเดล G-PCM ค่าพารามิเตอร์ความชัน แสดงถึงระดับความผันแปรของรายการคำตอบระหว่างข้อ เมื่อ $\theta$ ของผู้ตอบเปลี่ยนไป ถ้า $\alpha_i$ มีค่ามากกว่า 1.00 แสดงว่าโค้งรายการคำตอบของข้อนั้นมีความชันสูงมากกว่าโค้งรายการคำตอบในโมเดล PCM

**Rating Scale Model**

มีผู้พัฒนาโมเดลที่เรียกว่า Rating Scale Model หลายลักษณะ ซึ่งมีความแตกต่างกันในแง่ของความสลับซับซ้อนอย่างหลากหลายแนวคิด (Anderson, 1995) ในที่นี้จะขอนำเสนอ Rating Scale Model (RSM) ตามแนวคิด Andrich (198a, 1978b) ซึ่งมีที่มาจาก PCM คือมีค่าความชันหรือค่าอำนาจจำแนกของแต่ละข้อคำถามเท่ากัน คือ แต่ละข้อมีค่าพารามิเตอร์ตำแหน่งมาตร (scale location parameter, $\lambda_i$) ซึ่งเป็นค่าความง่ายหรือความยากของแต่ละข้อคำถาม

นอกจากนี้ยังมีค่า $\delta_j$ (category intersection parameter) ซึ่งเป็นค่าพารามิเตอร์จุดตัดระหว่างรายการคำตอบ มีจำนวนเท่ากับ j = k – 1 (โดยที่ k คือรายการคำตอบ)

RSM เป็นโมเดลที่คล้ายกับ PCM อยู่บนพื้นฐานของโมเดลการตอบสนองข้อสอบที่มี 1 พารามิเตอร์ (1-Parameter model) และใช้หลักการคำนวณความน่าจะเป็นของการตอบแต่ละระดับขั้นการตอบโดยตรงแบบขั้นตอนเดียว (Direct IRT Method)

***โมเดล***

RSM มีความคล้ายโมเดล PCM แต่มีความแตกต่างที่สำคัญอยู่หลายประการใน RSM ลักษณะคำถามแต่ละข้อ (i) อธิบายด้วยค่าพารามิเตอร์ตำแหน่งของข้อคำถาม (location parameter; $\lambda_i$) ซึ่งสะท้อนถึงค่าความยากง่ายสัมพัทธ์ของแต่ละข้อ และค่าพารามิเตอร์ Threshold ของรายการคำตอบ ($\delta_j$) ซึ่งเป็นค่าร่วมกันของทุกข้อ โดย j = K-1 รายการคำตอบ โมเดล RSM มีข้อตกลงเบื้องต้นว่าทุกข้อคำถามมีค่าพารามิเตอร์ความชันเท่ากัน หรือมีค่าอำนาจจำแนกเท่ากัน (มีค่าพารามิเตอร์ระดับความยากของขั้นการตอบร่วมกันทั้งฉบับ) และคะแนนดิบรวมเป็นค่าสถิติที่เพียงพอสำหรับการคำนวณค่าคุณลักษณะ θ ของผู้ตอบ การที่เป็นเช่นนี้ เพราะในการวัดทัศนคติที่เป็นมาตรวัดที่มีรายการคำตอบเป็น 1=ไม่เห็นด้วย 2 = เห็นด้วยบ้าง 3 = เห็นด้วยมากที่สุด นั้น ความยากสัมพัทธ์ระหว่างข้อน่าจะมีความแตกต่างกัน ต่างกับแบบสอบที่วัดผลสัมฤทธิ์ เช่น วิชาคณิตศาสตร์ ที่มีบางข้อมีค่าความยากง่ายแตกต่างกันในขณะที่โมเดล PCM ไม่มีข้อตกลงเบื้องต้นเกี่ยวกับระดับความยากสัมพัทธ์ในแต่ละข้อคำถาม ทำให้ค่าพารามิเตอร์จุดตัดระหว่างรายการคำตอบ (category intersection parameter) มีความแตกต่างระหว่างข้อคำถาม ยอมให้ค่าระดับความยากของแต่ละขั้นการตอบมีความผันแปรได้ระหว่างข้อคำถาม

ใน RSM ค่าพารามิเตอร์ Threshold ของแต่ละรายการคำตอบ ถูกจำแนกออกเป็น 2 ส่วน ประกอบด้วย $\lambda_i$ และ $\delta_j$ โดย $\delta_{ij} = (\lambda_i + \delta_j)$ ดังนั้น โค้งฟังก์ชันการเลือกรายการคำตอบของโมเดล RSM เขียนได้ดังนี้

$$P_x(\theta) = \frac{\exp\left\{\sum_{j=0}^{x}\left[\theta - (\lambda_i + \delta_j)\right]\right\}}{\sum_{x=0}^{m}\exp\left\{\sum_{j=0}^{x}\left[\theta - (\lambda_i + \delta_j)\right]\right\}}$$

เมื่อ $\sum_{j=0}^{0}\left[\theta - (\lambda_i + \delta_j)\right] = 0$

$P_x(\theta)$ = ความน่าจะเป็นที่ผู้ตอบมีคุณลักษณะ θ จะเลือก หรือสามารถทำรายการคำตอบขั้นที่ x จากจำนวน M ขั้น (step)

$\lambda_i$ = ค่าพารามิเตอร์ตำแหน่งของข้อคำถาม i (Location parameter)

$\delta_i$ = ค่าพารามิเตอร์ Threshold ของขั้นการตอบที่ j (category intersection parameter) ซึ่งเป็นค่าร่วมกันของทุกข้อคำถาม

RSM ใช้ชุดของค่าคงที่ร่วมกันระหว่างระดับขั้นของคำตอบ ถ้าข้อคำถามมีสเกลการตอบแตกต่างกัน โมเดล RSM ไม่เหมาะที่จะนำมาใช้ แต่บางโปรแกรมที่มีลักษณะพิเศษ เช่น โปรแกรม RUMN (Sheridan, Andrich and Luo, 1996) สามารถใช้ประมาณค่าพารามิเตอร์ของกลุ่มข้อคำถามที่มีรูปแบบของสเกลคำตอบที่เหมือนกันจัดไว้ในกลุ่ม (block) เดียวกันได้ การวิเคราะห์จึงอาจมีกลุ่มข้อคำถามหลายกลุ่มในฉบับเดียวกันได้ แต่ในทางปฏิบัติอาจเกิดปัญหาของการปรับเทียบค่าพารามิเตอร์ เพื่อการปรับเทียบระหว่างกลุ่มข้อคำถาม

โมเดล RSM เหมาะสำหรับมาตรวัดที่มีรายการคำตอบเท่ากันทุกข้อคำถาม ดังนั้นมาตรวัดใดที่มีรายการคำถามไม่เท่ากัน จึงไม่เหมาะสมที่จะใช้โมเดลนี้วิเคราะห์แต่ถ้าต้องการวิเคราะห์ด้วยโมเดลนี้โดยการจัดบล็อคกลุ่มข้อคำถามที่มีรายการคำตอบเหมือนกันไว้ด้วยกันก็สามารถทำได้แต่จะมีปัญหาในการเปรียบเทียบคะแนนระหว่างบล็อค

**Nominal Response Model**

Bock (1972) ได้พัฒนา Nominal response model (NRM) สำหรับใช้วิเคราะห์ข้อสอบหรือข้อคำถามที่รายการคำตอบไม่จำเป็นต้องถูกเรียงลำดับ NRM มีลักษณะเป็นโมเดลทั่วไปที่ใช้หลักการคำนวณความน่าจะเป็นของการตอบแต่ละรายการคำตอบโดยตรงแบบขั้นตอนเดียว (Direst IRT Model) โมเดลที่คำนวณความน่าจะเป็นแบบขั้นตอนเดียวที่กล่าวมาข้างต้นต่างเป็นโมเดลลักษณะเฉพาะ (special cases) ของโมเดล NRM

***โมเดล***

ใน NRM ลักษณะข้อคำถามแต่ละข้อ (i) อธิบายได้ด้วยค่าพารามิเตอร์ความชันของแต่ละรายการคำตอบ (slope of the trace lines; $\alpha_{ix}$) และค่าพารามิเตอร์จุดตัดของแต่ละรายการในโมเดล NRM สามารถเขียนเป็นฟังก์ชันได้ดังนี้

$$P_{ix}(\theta) = \frac{\exp(\alpha_{ix}\theta + c_{ix})}{\sum_{x=0}^{m_i} \exp(\alpha_{ix}\theta + c_{ix})}$$

เมื่อ $P_{ix}(\theta)$ = ความน่าจะเป็นที่ผู้ตอบมีคุณลักษณะ $\theta$ จะเลือกรายการคำตอบขั้นที่ ในเมื่อ x = 0, 1,…,m

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

$\alpha_{ix}$ = ค่าพารามิเตอร์ความชันของแต่ละรายการคำตอบ (slope parameter)

$c_{ix}$ = ค่าพารามิเตอร์จุดตัดของแต่ละรายการคำตอบ (intercept parameter)

เพื่อให้สามารถประมาณค่าพารามิเตอร์ในสมการได้ จึงจำเป็นต้องกำหนดเงื่อนไขให้ $\sum\alpha_{ix} = \sum c_{ix} = 0$ บางกรณีมีการกำหนดให้ค่าพารามิเตอร์ของรายการคำตอบต่ำสุด $\alpha_{i1} = c_{i1} = 0$ ตามโมเดล NRM จะมีการประมาณค่าพารามิเตอร์ $\alpha_{ix}$ และ $c_{ix}$ สำหรับแต่ละรายการคำตอบซึ่งมีจำนวนทั้งหมด $m_i+1$ รายการคำตอบในแต่ละข้อ

Thissen (1993) ได้แสดงตัวอย่างข้อคำถามสำหรับวัดความเสี่ยงต่อโรคเบื่ออาหาร และทำการประมาณค่าพารามิเตอร์ของแต่ละรายการคำตอบโดยใช้โมเดล NRM

ตัวอย่างข้อคำถาม : ท่านชอบบรรยากาศการรับประทานอาหารลักษณะใด

(1) ที่บ้านตามลำพัง (2) ที่บ้านกับครอบครัว (3) ที่ภัตราคาร (4) ที่บ้านเพื่อน (5) แบบใดก็ได้

**ข้อควรพิจารณาเกียวกับโมเดลการตอบสนองข้อสอบทีมีการตรวจให้คะแนนมากกว่า 2 ค่า** (ศิริชัย กาญจนวาสี, 2545) ในประเด็นต่าง ๆ มีดังนี้

**สารสนเทศของข้อสอบและแบบสอบ**

การประมาณค่าสารสนเทศของข้อสอบ (Item Information Function, IIF) หรือ $I_i(\theta)$ และค่าฟังก์ชันสารสนเทศของแบบสอบ (Test Information Function, TIF) หรือ $I(\theta)$

สำหรับโมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า สามารถนำแนวคิดมาประยุกต์ใช้กับโมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่าได้เช่นเดียวกัน กล่าวคือ เมื่อพิจารณาข้อสอบหรือข้อคำถามแต่ละข้อ เราสามารถแปลงโค้งรายการคำตอบให้เป็นโค้งสารสนเทศของข้อสอบหรือข้อคำถามแต่ละข้อ เราสามารถแปลงโค้งรายการนำมารวมกัน (ณ ตำแหน่ง $\theta$ เดียวกัน) ทำให้ได้โค้งสารสนเทศของแบบสอบ ซึ่งสามารถเขียนในรูปของฟังก์ชันสารสนเทศของข้อสอบและแบบสอบได้ (Dodd, DeAyala, and Koch, 1995 อ้างใน ศิริชัย กาญจนวาสี, 2545) ดังสมการ

$$\text{IIF หรือ } I_i(\theta) = \sum_{x=0}^{m}\left[\frac{P_{ix}^{*}(\theta)^2}{P_{ix}(\theta)}\right]$$

จาก IIF สามารถคำนวณ TIF และ $SE(\theta)$ ได้ดังนี้

$$\text{TIF หรือ } I(\theta) = \sum_{i=1}^{k} I_i(\theta)$$

$$SE = \frac{1}{\sqrt{I(\theta)}}$$

**การเลือกใช้โมเดลการตอบข้อสอบที่มีการตรวจให้คะแนนมากกว่า 2 ค่า**

***1. การเปรียบเทียบระหว่างโมเดล***

ในการเลือกใช้โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า สิ่งแรกที่ควรพิจารณาถึง ก็คือ ลักษณะของข้อมูลที่ได้จากการเก็บรวบรวมข้อมูล ถ้ามีข้อมูลพื้นฐานบ่งชี้หรือทำให้เชื่อได้ว่า ข้อสอบหรือข้อคำถามแต่ละข้อมีอำนาจจำแนกเท่า ๆ กัน ก็ควรเลือกใช้โมเดลที่พัฒนาบนพื้นฐานของโมเดลของราสช์ แต่ถ้าข้อสอบหรือข้อคำถามแต่ละข้อน่าจะมีอำนาจจำแนกต่างกัน ซึ่งอาจจะพิจารณาจากผลการวิเคราะห์ตัวประกอบ ก็ควรเลือกใช้ GRM หรือโมเดลที่ปรับปรุงโดยมูราคิ เช่น M-GRM, G-PCM เป็นต้น สำหรับข้อสอบหรือข้อคำถามที่รายการคำตอบไม่ได้จัดเรียงอย่างเป็นลำดับขั้นเหมาะสมที่จะเลือกใช้ NRM แต่ถ้าข้อสอบหรือข้อคำถามถูกกำหนดให้มีจำนวนรายการคำตอบแบบเดียวกัน ก็ควรเลือกใช้ RSM หรือ M-GRM

การเปรียบเทียบความเหมาะสมสำหรับการเลือกใช้โมเดล สามารถพิจารณาได้จากสถิติทดสอบสมการโครงสร้างของโมเดลที่เรียกว่า log – likelihood อันเป็นค่าสถิติที่แสดงถึงความสอดคล้องระหว่างค่าอุดมคติของโมเดลกับค่าที่สังเกตได้จากข้อมูลจริง เพื่อทำการเปรียบเทียบระหว่างโมเดล ว่าโมเดลใดสอดคล้องกับข้อมูลจริงมากกว่ากัน (Rost, 1988; Levine et.al.,1992; Maydeu – Olivares, Drasgow and Mead, 1994 อ้างถึงใน ศิริชัย กาญจนวาสี, 2545)

กรณีที่ไม่มีเครื่องมือวิเคราะห์ทางสถิติช่วยตัดสินใจเลือกโมเดล ในทางปฏิบัติอาจพิจารณา 1) จำนวนพารามิเตอร์ที่ต้องการประมาณค่าในโมเดล ถ้ามีทรัพยากรจำกัด ควรเลือกใช้โมเดลที่มีจำนวนพารามิเตอร์ที่ต้องประมาณค่าจำนวนน้อยตัวจะดีกว่า เพราะอาจใช้ขนาดกลุ่มตัวอย่างที่เล็กลง 2) แบบแผนการตอบและการตรวจ ตามปกติควรเลือกใช้แบบแผนการตอบที่สะดวกและสามารถตรวจให้คะแนนได้ง่ายอย่างมีความเป็นปรนัย จะช่วยลดความสลับซับซ้อนของโมเดลการวิเคราะห์และการแปลผล 3) ปรัชญาความเชื่อเกี่ยวกับโมเดลและจุดมุ่งหมายของการนำผลไปใช้ของผู้พัฒนาแบบสอบ เช่น จำนวนสเกลคำตอบควรเท่ากันสำหรับแต่ละข้อคำถามหรือไม่ ความต้องการเกี่ยวกับอำนาจจำแนกของข้อสอบหรือข้อคำถามว่าแต่ละข้อน่าจะมีอำนาจจำแนกเท่ากันหรือไม่ ควรมีค่าความยากร่วมกันหรือเปล่า เป็นต้น

นอกจากนี้ยังมีสิ่งที่ควรพิจารณาเพิ่มเติมเกี่ยวกับลักษณะข้อมูล สำหรับการเลือกใช้โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า ดังนี้ 1) ควรใช้กลุ่มตัวอย่างขนาดใหญ่ที่มีความเป็นวิวิธพันธ์ (heterogeneous sample) เพื่อให้การประมาณค่ามีความน่าเชื่อถือตลอดช่วงความสามารถหรือคุณลักษณะ ($\theta$) และ 2) ข้อมูลที่นำมาวิเคราะห์จะต้องมีการตอบทุก

ข้อ และแต่ละข้อจะต้องมีการตอบทุกรายการ ถ้าไม่มีการตอบหรือความถี่ของรายการคำตอบมีน้อย โปรแกรมคอมพิวเตอร์ที่ใช้ในการวิเคราะห์จะไม่สามารถประมาณค่าพารามิเตอร์หรือประมาณค่าได้โดยมีความคลาดเคลื่อนสูง การแก้ปัญหานี้อาจกระทำได้โดยเขียนข้อคำถามให้มีรายการคำตอบจำนวนน้อย หรืออาจมีความจำเป็นต้องยุบรวมบางรายการคำตอบเข้าด้วยกัน

***2. ขนาดกลุ่มตัวอย่าง***

ขนาดกลุ่มตัวอย่างที่เพียงพอสำหรับการวิเคราะห์ IRT ขึ้นอยู่กับหลายปัจจัย เช่น คุณลักษณะของโมเดลที่เลือกใช้ จำนวนพารามิเตอร์ที่จะต้องประมาณค่าในโมเดล เป็นต้น Reise และ Yu (1990 อ้างถึงใน ศิริชัย กาญจนวาสี, 2545) ทำการศึกษาโดยใช้เทคนิคมอนติคาโล (Monte Carlo simulation) ได้แสดงให้เห็นว่า เราสามารถประมาณค่าพารามิเตอร์ของ GRM โดยใช้โปรแกรม MULTILOG ด้วยขนาดกลุ่มตัวอย่าง 250 คน แต่ถ้าต้องการให้ได้ผลดี ควรใช้ขนาดกลุ่มตัวอย่างไม่น้อยกว่า 500 คน ซึ่งข้อเสนอแนะดังกล่าวอยู่ภายใต้สถานการณ์ของข้อมูลจำลองที่ศึกษา ถ้าพิจารณาตามหลักการแล้ว ควรกำหนดขนาดกลุ่มตัวอย่างให้มีขนาดใหญ่พอที่จะทำให้ความคลาดเคลื่อนมาตรฐานของการประมาณค่าพารามิเตอร์มีขนาดเล็กลง ถึงระดับที่ยอมรับได้ตามเป้าหมายของการนำผลไปใช้ในทางปฏิบัติ

ในปัจจุบันได้มีการใช้แบบสอบรูปแบบผสม(mixed–format test) ซึ่งเป็นการผสมข้อคำถามที่มีรูปแบบการตรวจให้คะแนนต่างกันอยู่ในแบบสอบฉบับเดียวกัน โดยรูปแบบของข้อสอบส่วนใหญ่จะจำแนกเป็น 2 รูปแบบคือ ข้อสอบแบบหลายตัวเลือก (multiple choice: MC) และข้อสอบแบบเขียนคำตอบ (constructed response: CR) ซึ่งข้อสอบแบบ MC มีการตรวจให้คะแนนแบบสองค่า (dichotomous) เป็นวิธีที่นิยมใช้กันมาก เพราะสะดวก รวดเร็ว เหมาะกับข้อสอบประเภทเลือกตอบ ข้อดีของการให้คะแนนแบบนี้คือ สะดวก รวดเร็ว ประหยัดเวลา ข้อเสียคือ ผู้สอบที่ตอบได้ 0 คะแนนไม่ได้หมายความว่าผู้สอบไม่มีคุณลักษณะที่มุ่งวัดนั้น และเราไม่สามารถทราบได้ว่าผู้สอบที่ตอบข้อสอบข้อนั้นได้ 0 คะแนน มีความสามารถหรือ คุณลักษณะจริงๆ อยู่ในระดับใด ซึ่งหากเป็นข้อสอบแบบ CR ที่มีการตรวจให้คะแนนแบบหลายค่า (polytomous) จะมีการให้คะแนนแบบที่กำหนดคะแนนในมาตรวัดแบบมาตรประมาณค่าชนิดต่าง ๆ ที่ใช้วัดคุณลักษณะ ซึ่งในมาตรวัดก็จะกำหนดคะแนนตามลำดับขั้น (category) ของคะแนนในเครื่องมือนั้นๆ โดยถ้าเป็นแบบวัดคุณลักษณะ คะแนนแต่ละค่าจะหมายถึงระดับของคุณลักษณะที่มุ่งวัด แต่ถ้าเป็นแบบสอบวัดผลสัมฤทธิ์การให้คะแนนแต่ละค่าจะแสดงถึงระดับความสามารถของผู้สอบที่ตอบข้อสอบข้อนั้น ข้อดีของการให้คะแนนแบบหลายค่าคือ สามารถแสดงให้เห็นถึงระดับความสามารถของผู้สอบที่ตอบข้อสอบข้อนั้นได้ แต่ข้อจำกัดคือ ในบางสถานการณ์เราไม่สามารถกำหนดคะแนนในแต่ละลำดับขั้นได้อย่างชัดเจน แต่ในการวัดคุณลักษณะบางประการไม่สามารถ

วัดได้จากแบบสอบรูปแบบใดรูปแบบหนึ่ง ดังนั้นแบบสอบรูปแบบผสมจึงเป็นทางเลือกอีกทางหนึ่งในการนำไปใช้วัดคุณลักษณะต่าง ๆ ได้

สำหรับการวิจัยครั้งนี้ผู้วิจัยเลือกใช้โมเดลการตอบตามทฤษฎีการตอบสนองข้อสอบที่ให้คะแนนสองค่า แบบโลจีสติก 3 พารามิเตอร์ เนื่องจากลักษณะของข้อสอบที่ได้รับความนิยมในการนำไปใช้ทดสอบไม่ว่าจะเป็นระดับจุลภาคหรือระดับมหภาค มักจะเป็นข้อสอบแบบเลือกตอบ ซึ่งธรรมชาติของข้อสอบแบบเลือกตอบแต่ละข้อจะมีความยาก และอำนาจจำแนกไม่เท่ากัน อีกทั้งผู้สอบยังสามารถเดาคำตอบได้ ดังนั้นการใช้โมเดลโลจีสติกแบบ 3 พารามิเตอร์ จึงน่าจะเหมาะสมและใกล้เคียงกับสภาพการวัดและประเมินจริง ส่วนโมเดลการตอบสนองข้อสอบตามทฤษฎีการตอบสนองข้อสอบที่ให้คะแนนหลายค่า ผู้วิจัยเลือกใช้โมเดล Generalize Partial Credit : GPC ที่พัฒนาโดย Muraki (1992, 1993) เพราะโมเดลนี้อาจเป็นข้อสอบได้หลายรูปแบบ ทั้ง essay, passage – based หรือ rating scale ซึ่งสามารถนำไปใช้กับข้อสอบหลายรูปแบบตามความเหมาะสม และยังมีการให้คะแนนแบบตอบถูกบางส่วนหรือให้คะแนนตามลำดับขั้นก็ได้ อีกทั้งการศึกษาเอกสารและวรรณกรรมที่เกี่ยวข้องยังพบว่าแบบสอบรูปแบบผสมเป็นการผสมระหว่างโมเดล 3PL กับโมเดล GPC มากที่สุดด้วย

## ตอนที 3 การปรับเทียบคะแนน

### มโนทัศน์ของการปรับเทียบคะแนน

การเปรียบเทียบคะแนนจากแบบสอบต่างฉบับที่วัดความสามารถเดียวกันเป็นประเด็นที่สำคัญที่ควรพิจารณาสำหรับนักพัฒนาแบบสอบ ผู้เชี่ยวชาญทางการวัด และผู้สร้างแบบสอบ ถ้าผู้สอบ 2 คนทำแบบสอบต่างฉบับกัน เราจะเปรียบเทียบคะแนนเหล่านั้นได้อย่างไร ปัญหานี้เป็นปัญหาที่สำคัญเมื่อต้องนำไปใช้ในการออกไปรับรอง (certification) การคัดเลือก หรือการตัดสินใจให้ผ่านหรือตก ซึ่งแบบสอบต่างฉบับที่ใช้ทดสอบควรจะมีเนื้อหาไม่แตกต่างกัน

คะแนนจากแบบสอบเป็นส่วนหนึ่งของสารสนเทศที่ใช้ในการตัดสินใจที่สำคัญ การตัดสินใจบางอย่างมีจุดเน้นที่ระดับบุคคล (individual level) เช่น เมื่อนักเรียนตัดสินใจเข้าศึกษาต่อ สำหรับการตัดสินใจอื่น ๆ จะเน้นที่ระดับสถาบันหรือระดับองค์กรมากขึ้น (institutional level) ตัวอย่างเช่น ตัวแทนหรือหน่วยงานจำเป็นต้องตัดสินคะแนนจากแบบสอบที่ต้องการเพื่อรับรองความเชี่ยวชาญของบุคคล หรือรับนักศึกษาเข้าศึกษาต่อในสถาบัน ส่วนการตัดสินใจอื่นที่ยังมีอยู่ก็จะเป็นการตัดสินใจระดับนโยบายสาธารณะ (public policy level) เช่น กำหนดสิ่งที่ต้องทำให้

เสร็จสิ้นเพื่อปรับปรุงการศึกษาในสหรัฐอเมริกา และเปลี่ยนแปลงระบบการปฏิบัติเพื่อให้สามารถประเมินได้ การตัดสินใจควรจะมีพื้นฐานของสารสนเทศที่มีความถูกต้อง เชื่อถือได้ และมีความเป็นไปได้ เป็นสำคัญ นั่นคือ เมื่อสารสนเทศมีความถูกต้องมากขึ้น การตัดสินใจก็จะดีขึ้นด้วย

การตัดสินใจในหลาย ๆ บริบทต้องทำการทดสอบหลายสถานการณ์ เช่น การสอบคัดเลือกบุคคลเข้าศึกษาต่อ ซึ่งต้องมีการทดสอบหลายปี ถ้าแบบสอบมีคำถามเหมือนกันและผู้สอบบางคนสอบสองครั้งก็จะทำให้เป็นการวัดการจำข้อสอบของผู้สอบ มากกว่าที่จะวัดโครงสร้างเนื้อหา

ปัญหาที่เกิดขึ้นข้างต้นสามารถแก้ไขได้โดยการใช้แบบสอบที่ต่างกัน นั่นคือ มีรูปแบบต่างกัน (test form) เพื่อให้ผู้สอบคนที่ทำแบบสอบสองครั้งหรือมากกว่าได้รับแบบสอบที่ต่างกัน รูปแบบของแบบสอบเป็นชุดของข้อคำถามที่สร้างให้สอดคล้องกับเนื้อหาและแบบแผนของแบบสอบทางสถิติ (Millman and Greene, 1989) แบบแผนของแบบสอบเป็นแนวทางสำหรับพัฒนาแบบสอบ เพื่อให้ผู้พัฒนาแบบสอบมั่นใจได้ว่าแบบสอบที่สร้างขึ้นมีเนื้อหาและลักษณะทางสถิติคล้ายกับแบบสอบฉบับอื่น

การใช้แบบสอบที่ต่างกันและต่างเวลานั้นก็อาจจะเป็นสาเหตุของปัญหาอื่น ๆ อีก ดังสถานการณ์ที่จะกล่าวถึงคือ นักเรียน 2 คนจบจากวิทยาลัยเดียวกันทำแบบสอบต่างรูปแบบและต่างเวลากัน และนักเรียนคนที่ 1 ได้คะแนนสูงกว่านักเรียนคนที่ 2 สามารถอธิบายได้ว่านักเรียนคนที่ 1 มีผลสัมฤทธิ์ทางการเรียนสูงกว่านักเรียนคนที่ 2 แต่อีกนัยหนึ่งคือ นักเรียนคนที่ 1 ได้รับแบบสอบที่ง่ายกว่านักเรียนคนที่ 2 ซึ่งเป็นการสอบที่ทำให้นักเรียนคนที่ 2 เสียเปรียบนักเรียนคนที่ 1 ในกรณีนี้ คะแนนที่ต่างกันอาจจะเป็นความแตกต่างที่เนื่องมาจากความยากของแบบสอบ มากกว่าที่จะเป็นความแตกต่างด้านระดับความสามารถของนักเรียน เพื่อให้หลีกเลี่ยงปัญหานี้ จึงมีการนำวิธีการปรับเทียบเข้ามาใช้ในการรับนักเรียนเข้าศึกษาต่อ ถ้าแบบสอบได้รับการปรับเทียบเรียบร้อยแล้ว ความแตกต่างของคะแนนที่ถูกปรับเทียบแล้วของนักเรียนคนที่ 1 และ คนที่ 2 ก็จะไม่ได้เกิดจากคุณลักษณะที่ว่านักเรียนคนที่ 1 ทำแบบสอบที่ง่ายกว่านักเรียนคนที่ 2

กระบวนการของการปรับเทียบใช้ในสถานการณ์ที่รูปแบบของแบบสอบต่างกันและคะแนนที่ได้มาจากรูปแบบของแบบสอบที่ต่างกันนำไปเปรียบเทียบกับแบบสอบอื่น แม้ว่าผู้พัฒนาแบบสอบจะพยายามสร้างรูปแบบของแบบสอบที่มีเนื้อหาและลักษณะทางสถิติที่คล้ายกันมากที่สุด แต่แบบสอบก็จะยังมีความยากที่ต่างกัน การปรับเทียบนำไปสู่การปรับความยากเหล่านี้ การปรับเทียบแบบสอบที่ความยากต่างกัน ไม่ใช่การปรับเทียบแบบสอบที่เนื้อหาต่างกัน

การปรับเทียบคะแนนเป็นรูปแบบหนึ่งของการเชื่อมโยงคะแนน โดย Kolen และ Brennan (2004) ได้จำแนกรูปแบบการเชื่อมโยงคะแนนตามเกณฑ์ จากการนำแนวคิดของ Feuer et al. (1999) ผนวกกับแนวคิดของ Mislevy (1992) และ Linn (1993) ในการจำแนกรูปแบบการเชื่อมโยงคะแนน ซึ่งมีรายละเอียดในการจำแนกรูปแบบการเชื่อมโยงคะแนนและเกณฑ์ที่ใช้ดังต่อไปนี้

*รูปแบบที่ 1 การปรับเทียบคะแนน (Equating)* เป็นการเชื่อมโยงคะแนนจากแบบสอบที่มีการสอบกันคนละครั้ง แต่แบบสอบถูกสร้างมาด้วยลักษณะของความร่วมกันหรือสัมพันธ์กันในด้านเนื้อหาและค่าสถิติของแบบสอบ หรือเป็นการนำมาจากการมีลักษณะของเนื้อหาเดียวกันและมีค่าสถิติของแบบสอบแบบเดียวกัน (same content and same statistical specification)

ถ้าพิจารณาในกรณีของการจำแนกรูปแบบตามเกณฑ์ของ Feuer et al. โดยใช้เกณฑ์ในการจำแนก คือ การกำหนดโครงร่างของแบบสอบและการกำหนดแบบแผนของแบบสอบ พบว่าแบบสอบทั้ง 2 ฉบับต้องมีการกำหนดโครงร่างของแบบสอบและการกำหนดแบบแผนของแบบสอบที่มีลักษณะเดียวกัน และควรมีคุณสมบัติของความไม่ผันแปรตามกลุ่ม คือ ไม่ว่าจะนำผลของคะแนนที่ได้ไปใช้จะไม่ขึ้นอยู่กับกลุ่มผู้สอบ

*รูปแบบที่ 2 การทำให้เป็นคะแนนมาตรฐาน (Calibration)* เป็นการเชื่อมโยงคะแนนจากแบบสอบที่วัดคุณลักษณะเดียวกันแต่แบบสอบทั้ง 2 ฉบับมีค่าความยากหรือค่าความเที่ยงต่างกัน ถ้าพิจารณาในกรณีของการจำแนกรูปแบบตามเกณฑ์ของ Feuer และคณะ โดยใช้เกณฑ์ในการจำแนก คือ การกำหนดโครงร่างของแบบสอบและการกำหนดแบบแผนของแบบสอบ พบว่าสามารถแบ่งออกได้เป็น 2 กรณี คือ

กรณีที่ 1 แบบสอบทั้ง 2 ฉบับที่จะนำมาทำเป็นคะแนนมาตรฐานได้นั้นจะต้องมีการกำหนดโครงร่างของแบบสอบที่มีลักษณะเดียวกันแต่มีการกำหนดแบบแผนของแบบสอบที่ต่างกัน

กรณีที่ 2 แบบสอบทั้ง 2 ฉบับมีการกำหนดโครงร่างของแบบสอบและการกำหนดแบบแผนแบบสอบที่ต่างกัน ตลอดจนในกรณีที่อาจมีการใช้โครงร่างของแบบสอบที่ร่วมกัน

*รูปแบบที่ 3 การทำนาย (Projection)* เป็นการเชื่อมโยงคะแนนจากแบบสอบที่วัดโครงสร้างต่างกัน โดยใช้วิธีการทำนายเข้ามาช่วยในการเชื่อมโยงคะแนนและมีข้อแตกต่างจากการปรับเทียบ (equating)และการทำให้เป็นคะแนนมาตรฐาน (calibration) คือ

1) การทำนายเป็นการศึกษาความสัมพันธ์แบบทางเดียว (unidirectional) คือขาดคุณสมบัติของความสมมาตร (symmetric relationship) คือสหสัมพันธ์ระหว่าง x กับ y มีค่าไม่เท่ากับ ความสัมพันธ์ระหว่าง y กับ x

2) การทำนาย เป็นวิธีที่เหมาะสมกับกรณีของการออกแบบในการเก็บรวบรวมข้อมูลแบบผู้สอบกลุ่มเดียว (single group design)

3) การทำนายเป็นวิธีที่ไม่คำนึงถึงคุณลักษณะของแบบสอบว่าต้องมีลักษณะเดียวกัน

*รูปแบบที่ 4 การปรับค่าทางสถิติ (Moderation)* หรือในบางครั้งเรียกว่า “การปรับให้การกระจายของคะแนนเท่าเทียมกัน (distribution matching)” โดยใช้กรณีที่มีการออกแบบเก็บรวบรวมข้อมูลแบบผู้สอบกลุ่มเดียว (single group design) ตลอดจนรวมถึงการออกแบบการเก็บรวบรวมข้อมูลแบบใช้ผู้สอบกลุ่มสุ่ม (random groups) และแบบการใช้ผู้สอบกลุ่มที่ไม่ทัดเทียมกัน (nonequivalent groups) ด้วย อย่างไรก็ตามการตัดสินใจที่แตกต่างกันเกิดมาจากการนำวิธีมาพิจารณาความชำนาญในแต่ละกรณีมีความแตกต่างกัน

นอกจากนี้ Dorans (2000) ได้ทำการแบ่งชนิดของการเชื่อมโยงคะแนนออกเป็น 3 ชนิดด้วยกัน คือ

1. การปรับเทียบคะแนน (Equating) เป็นการเชื่อมโยงคะแนนโดยที่ผลคะแนนที่ได้สามารถนำมาทดแทนกันได้อย่างสมบูรณ์ (fully exchangeable) โดยที่แบบสอบทั้ง 2 ฉบับถูกสร้างมาโดยมีการกำหนดแบบแผนข้อสอบ (Item Specifications) ที่มีลักษณะเดียวกัน

2. การสร้างมาตรวัด (Scaling) เป็นการเชื่อมโยงคะแนนโดยเป็นการแปลงคะแนนจากแบบสอบหนึ่งมาอยู่บนมาตรวัดของแบบสอบอีกฉบับหนึ่งโดยอาศัยการจัดตำแหน่งของผู้สอบในแต่ละแบบสอบ โดยที่แบบสอบ 2 ฉบับนี้มีการวัดคุณลักษณะที่คล้ายคลึงกันแต่มีการกำหนดแบบแผนข้อสอบ (Item Specifications) ที่มีลักษณะต่างกัน และผลของคะแนนที่ได้ไม่สามารถนำมาทดแทนกันได้

3. การทำนาย (Prediction) เป็นการเชื่อมโยงคะแนนที่อยู่ภายใต้ข้อตกลงเบื้องต้นของแบบสอบที่มีความผ่อนคลายของข้อตกลงเบื้องต้นเกี่ยวกับแบบสอบมากที่สุด แต่ผลของคะแนนที่ได้ไม่มีคุณลักษณะของความสมมาตร (symmetric)

จากแนวคิดเกี่ยวกับวิธีการของการเชื่อมโยงคะแนนที่กล่าวมาข้างต้น ผู้วิจัยสรุปได้ดังนี้

วิธีการในการเชื่อมโยงคะแนนระหว่างแบบสอบ 2 ฉบับ ประกอบด้วย 4 วิธี คือ

1. การปรับเทียบคะแนน (Equating) เป็นการเชื่อมโยงคะแนนเมื่อรูปแบบของแบบสอบมีเนื้อหาเดียวกันและมีค่าสถิติของแบบสอบมีลักษณะเดียวกัน ทำให้คะแนนที่ได้จากการปรับเทียบสามารถนำคะแนนมาทดแทนกันได้ (Angoff, 1971)

2. การทำคะแนนให้เป็นมาตรฐาน (Calibration) เป็นการเชื่อมโยงคะแนนเมื่อแบบสอบมีโครงสร้างเดียวกันแต่มีความแตกต่างกันในค่าความยากและค่าความเที่ยงของแบบสอบ เช่น การเปรียบเทียบคะแนนจากแบบสอบชนิดสั้นกับยาว (Angoff, 1971)

3. การปรับค่าทางสถิติ (Statistical moderation) เป็นการเชื่อมโยงคะแนนเมื่อแบบสอบมีโครงสร้างต่างกัน (Mislevy, 1992 และ Linn, 1993) มีการกำหนดแบบแผนข้อสอบ (Item Specifications) ที่มีลักษณะต่างกัน และผลของคะแนนที่ได้ไม่สามารถนำมาทดแทนกันได้ (Doran, 2000) ซึ่งคะแนนที่ได้เรียกว่าคะแนนความสอดคล้อง (Concordance Score)

4. การทำนาย (Prediction) เป็นการเชื่อมโยงคะแนนที่อยู่ภายใต้ข้อตกลงเบื้องต้นของแบบสอบที่มีความผ่อนคลายของข้อตกลงเบื้องต้นเกี่ยวกับแบบสอบมากที่สุด คือเป็นการเชื่อมโยงคะแนนจากแบบสอบที่วัดโครงสร้างต่างกัน ใช้การทำนายเข้ามาช่วยในการเชื่อมโยงคะแนน

**ความหมายของการปรับเทียบ**

จากสภาพการศึกษาในปัจจุบันซึ่งมีการใช้แบบสอบและผลการสอบในการวัดและประเมินผลสัมฤทธิ์ทางการเรียนของผู้เรียนอาจมีข้อจำกัดหลายประการในการบริหารการสอบ เช่น มีผู้สอบจำนวนมากเกินไป ซึ่งจำเป็นต้องแบ่งเป็นกลุ่มย่อยหลายกลุ่ม หรือจัดการสอบในเวลาต่างกันและแบบสอบที่ใช้ต่างฉบับกัน ระดับความยากแตกต่างกัน คะแนนจากข้อสอบและแบบสอบย่อยมีค่าไม่เท่ากันและมีความหมายต่างกัน ดังนั้นคะแนนที่ได้จากแบบสอบจึงควรที่จะนำมาปรับเทียบให้อยู่บนสเกลเดียวกัน เพื่อที่จะได้ประเมินหรือตัดสินได้อย่างถูกต้องและยุติธรรม

คำศัพท์ที่เกี่ยวข้องกับคำว่า “การปรับเทียบคะแนนระหว่างแบบสอบ” (Test Equating) มีอยู่หลายคำ Mislevy (1992) และ Lin (1993) (อ้างถึงใน Johnson and Owen, 1998) ได้อธิบายให้เห็นความแตกต่างของคำศัพท์ 4 คำ คือ Equating, Calibration, Projection และ Moderation ต่างเป็นรูปแบบหนึ่งของการเชื่อมโยง (linking) กระบวนการจำเป็นที่จะทำให้การเชื่อมโยงได้รับผลสำเร็จ ขึ้นอยู่กับว่าการประเมินทั้งสองมีความเหมือนกันในด้านเป้าหมาย การครอบคลุมเนื้อหาและคุณสมบัติของการวัดเพียงใด Mislevy และ Lin ได้อธิบายคำทั้ง 4 คำข้างต้นว่ามีความแตกต่างกันในข้อตกลงเบื้องต้นที่กำหนดซึ่งลดน้อยลงตามลำดับ ข้อกำหนดของการปรับเทียบ (Equating) จะมีข้อตกลงที่เข้มงวดกว่าการเชื่อมโยงวิธีการอื่นๆ

Mislevy และ Lin ได้กล่าวถึง การปรับเทียบคะแนนว่าเป็นการเชื่อมโยงการประเมินสองการประเมินด้วยแบบสอบที่สร้างจากตารางวิเคราะห์เดียวกัน ข้อกำหนดคือความเท่าเทียมกันในโครงสร้าง เช่น ความยาก รูปแบบของคำถาม วิธีการจัดการ จำนวนข้อสอบและความแม่นยำในการ

วัดแต่ละระดับของคะแนน ภายใต้สภาพการณ์ที่ควบคุม ทำให้ผลการประเมินสามารถสับเปลี่ยนกันได้ จึงสร้างตารางคะแนนฉบับ X ที่สอดคล้องกับคะแนนฉบับ Y ได้ ดังนั้น ทุก ๆ ข้อคำถามจึงกำหนดคะแนนจากการประเมิน X ซึ่งกำหนดในวิธีเดียวกันกับคะแนนแปลงจากการประเมิน Y เมื่อพิจาณาเทียบกับการปรับ (Calibration) ซึ่งเป็นการเชื่อมโยงที่เงื่อนไขน้อยการว่าเทียบคะแนนคือถ้าการประเมิน Y ถูกสร้างจากตารางวิเคราะห์รายละเอียดเดียวกันกับการประเมิน X แต่มีความแม่นยำแตกต่างกัน หรือระดับของความยากแตกต่างกัน ในกรณีที่ไม่สามารถปรับเทียบคะแนนได้ แต่ผลของการประเมินทั้งสองสามารถถูกปรับให้เหมาะสมได้ และสำหรับ Projection และ Moderation เป็นการเชื่อมโยงมาตรวัดเมื่อข้อสอบวัดในโครงสร้างที่แตกต่างกัน

ในขณะเดียวกันได้มีผู้ให้ความหมายการปรับเทียบคือ Angoff (1984) ได้ให้ความหมายของการปรับเทียบคะแนนระหว่างแบบสอบว่า เป็นการแปลงระบบคะแนนของแบบสอบฉบับหนึ่งให้อยู่บนสเกลของอีกฉบับอื่น คะแนนจากแบบสอบทั้งสองฉบับที่แปลงแล้วถือว่าเท่าเทียมกัน จึงสามารถเปรียบเทียบกันได้ และได้ให้แนวคิดเกี่ยวกับการปรับเทียบไว้ว่า การที่จะสร้างให้แบบสอบสองฉบับมีความยากเท่ากันนั้นเป็นไปได้ยากมาก ดังนั้นจึงต้องมีการปรับเทียบคะแนน การปรับเทียบคะแนนเป็นการปรับระบบคะแนนของแบบสอบฉบับหนึ่งให้อยู่บนหน่วยของแบบสอบฉบับอื่นที่มุ่งวัดในเรื่องเดียวกัน โดยมีจุดประสงค์เพื่อสามารถนำมาเปรียบเทียบกันได้ ซึ่งสอดคล้องกับ ศิริชัย กาญจนวาสี (2545) ที่ให้นิยามการปรับเทียบว่า เป็นกระบวนการแปลงระบบคะแนนของแบบสอบต่างฉบับที่วัดเนื้อหาเดียวกัน โดยแปลงคะแนนของแบบสอบฉบับหนึ่งให้อยู่บนมาตรวัดหรือสเกลของแบบสอบอีกฉบับหนึ่ง และสอดคล้องกับ Petersen และคณะ (1982) ที่ให้นิยามการปรับเทียบคะแนนระหว่างแบบสอบว่าเป็นกระบวนการเชิงประจักษ์ที่ใช้ในการแปลงคะแนนจากแบบสอบฉบับหนึ่งไปยังแบบสอบฉบับหนึ่ง โดยที่แบบสอบทั้งสองนั้นต้องวัดคุณลักษณะเดียวกัน ส่วน Lord (1980) ได้กล่าวถึงการปรับเทียบคะแนนระหว่างแบบสอบว่า การปรับเทียบคะแนนทำให้เกิดความเท่าเทียมกันระหว่างคะแนนดิบจากแบบสอบสองฉบับ และในแนวคิดของ Kolen และ Brennan (2004) การปรับเทียบคะแนนระหว่างแบบสอบ หมายถึงกระบวนการทางสถิติที่ใช้ในการปรับคะแนนจากแบบสอบต่างฉบับ เพื่อให้คะแนนสามารถใช้สับเปลี่ยน (interchangeably) กันได้ โดยที่แบบสอบแต่ละฉบับมีความยากแตกต่างกันได้ แต่ถ้ามีเนื้อหาต่างกัน จะไม่สามารถปรับเทียบกันได้ ดังคำกล่าวที่ว่า “Equating adjusts to differences in difficulty, not for differences in content”

จากแนวคิดของ Angoff ที่กล่าวมา พอจะมองเห็นลักษณะที่เป็นจุดเด่นเกี่ยวกับแบบสอบที่นำมาปรับเทียบกันได้ 2 ประการคือ

1) เครื่องมือทั้งสองฉบับ ต้องวัดคุณลักษณะเดียวกัน เช่น วัดความยาว วัดอุณหภูมิ วัดความสูง แต่ไม่สามารถเปลี่ยนสเกลแบบสองภาษา ให้เป็นแบบสอบทางคำนวณได้ และไม่สามารถเปลี่ยนสเกลอุณหภูมิไปเป็นสเกลความสูงได้ เป็นต้น

2) การปรับเทียบคะแนนต้องมีความคงทน คือ ในการปรับเทียบคะแนนไม่ว่าข้อมูลจะได้มาจากกลุ่มตัวอย่างใดก็ตามจะไม่แปรเปลี่ยนตามกลุ่ม นั่นคือ เป็นอิสระจากผู้ให้ข้อมูลสามารถกระทำได้กรณีแบบสอบสองชุดมีความเที่ยงเท่าเทียมกันและเป็นแบบสอบคู่ขนาน

ในขณะที่ Lord (1980) ได้กำหนดเงื่อนไขของการปรับเทียบคะแนนไว้ดังนี้

1) แบบสอบทั้งสองฉบับจะต้องวัดความสามารถเดียวกัน (same ability) คือวัดคุณลักษณะเดียวกัน อาจเป็นคุณลักษณะแฝง ความสามารถหรือทักษะอย่างใดอย่างหนึ่งก็ได้

2) มีความเสมอภาค (equity) คือ เมื่อทุกกลุ่มมีความสามารถเท่าเทียมกัน การแจกแจงของคะแนนจากแบบสอบ Y หลังจากที่มีการแปลงคะแนนแล้วมีการแจกแจงเหมือนกับการแจกแจงคะแนนจากแบบสอบ X

3) ความไม่แปรเปลี่ยนของประชากร (population invariance) คือการแปลงคะแนนต้องเป็นไปในลักษณะเดียวกันไม่ว่าคะแนนมาจากกลุ่มตัวอย่างใดก็ตาม

4) มีความสมมาตร (symmetry) คือการแปลงคะแนนสามารถเปลี่ยนกลับได้ระหว่างแบบสอบฉบับ X กับ Y

ดังนั้นจึงกล่าวได้ว่า การปรับเทียบคะแนนระหว่างแบบสอบเป็นกระบวนการทางสถิติที่นำมาใช้ในการปรับคะแนนของแบบสอบต่างฉบับที่วัดคุณลักษณะเดียวกันให้อยู่บนสเกลเดียวกัน คะแนนจากแบบสอบทั้งสองฉบับจึงสามารถเปรียบเทียบกันได้อย่างมีความหมาย

แบบสอบที่นำมาปรับเทียบคะแนน นอกจากวัดคุณลักษณะ (trait) เดียวกันแล้ว จะต้องมีความเที่ยงสูงและเป็นแบบสอบคู่ขนาน (parallel) ในด้านโครงสร้าง (structure) เวลาที่ใช้สอบ (timing) ชนิดของข้อสอบ (item types) รูปแบบ (format) และเนื้อหา (subject matter) ซึ่งข้อสอบของแต่ละฉบับอาจมีความยากแตกต่างกันได้ โดยเทคนิคการปรับเทียบคะแนนระหว่างแบบสอบจะช่วยปรับคะแนนจากแบบสอบต่างฉบับให้มีความเท่าเทียมกัน (Levien, 1955 cited in Hollan and Rubin, 1982) นอกจากนี้ Angoff (1971) Lord (1980) และ Hambleton and Swamminatan (1985) ได้กำหนดเงื่อนไขในการปรับเทียบคะแนนไว้ว่า แบบสอบต่างฉบับจะสามารถนำมาเปรียบเทียบกันได้ก็ต่อเมื่อแบบสอบเหล่านั้นมีคุณสมบัติ 4 ประการ ดังนี้

1) ความเป็นเอกมิติของแบบสอบ (unidimentionality of the tests) คือ แบบสอบทั้งสองฉบับวัดคุณลักษณะหรือความสามารถเดียวกัน เพียงคุณลักษณะเดียว การปรับเทียบจำเป็นต้อง

สร้างแบบสอบแต่ละรูปแบบให้มีเนื้อหาและมีแบบแผนทางสถิติเหมือนกัน ถ้าไม่คำนึงถึงกระบวนการทางสถิติที่ใช้ คะแนนจะไม่สามารถใช้แทนที่กันได้ (interchangeable) คุณสมบัตินี้จึงจำเป็นที่สุดถ้าคะแนนบนแบบสอบที่แตกต่างกัน (alternate forms) จะนำมาใช้ในการแทนที่กันได้

2) ความเสมอภาค (equity) คือ เมื่อทุกกลุ่มมีความสามารถเดียวกัน การแจกแจงคะแนนของแบบสอบ Y หลังจากที่มีการแปลงคะแนนแล้วจะมีการแจกแจงเหมือนกับการแจกแจงของคะแนนจากแบบสอบ X อาจกล่าวได้ว่าคุณสมบัตินี้เป็นลักษณะที่ผู้สอบได้คะแนนจริงที่มีลักษณะเหมือนกับคะแนนสังเกตได้ นั่นคือค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และลักษณะการแจกแจงของคะแนนที่แปลงจากแบบสอบ X และคะแนนจากแบบสอบ Y มีลักษณะเดียวกัน โดยเฉพาะการระบุส่วนเบี่ยงเบนมาตรฐานบอกเป็นนัยว่าความคลาดเคลื่อนมาตรฐานของการวัดเป็นเงื่อนไขสำหรับคะแนนจริงทุกค่าที่เท่ากันบนแบบสอบทั้งสองรูปแบบ ตัวอย่างเช่น ถ้าแบบสอบรูปแบบ X วัดบางสิ่งบางอย่างในระดับคะแนนสูงได้แม่นยำมากกว่าแบบสอบรูปแบบ Y ดังนั้นก็จะขาดคุณสมบัติของความเสมอภาค

3) ความไม่แปรผันตามกลุ่ม (invariance across group) คือการแปลงคะแนนต้องเป็นไปในลักษณะเดียวกันไม่ว่าคะแนนจะมาจากกลุ่มตัวอย่างใดก็ตาม

4) ความสมมาตร (symmetry) คือ การแปลงคะแนนจากฉบับ X ไปยังฉบับ Y มีผลเช่นเดียวกันกับการแปลงคะแนนจากแบบสอบฉบับ Y ไปสู่ฉบับ X คุณสมบัตินี้ต้องการให้ฟังก์ชันที่ใช้ในการแปลงคะแนนจากแบบสอบ X ไปยังแบบสอบ Y ตรงข้ามกับฟังก์ชันที่ใช้ในการแปลงคะแนนจากแบบสอบ Y ไปยังแบบสอบ X ตัวอย่างเช่น ถ้าคะแนนดิบจากแบบสอบ X เท่ากับ 26 แปลงไปสู่คะแนนดิบเท่ากับ 27 จากแบบสอบ Y ดังนั้นคะแนนดิบจากแบบสอบ Y ที่เท่ากับ 27 จะต้องเปลี่ยนแปลงไปสู่คะแนนดิบที่เท่ากับ 26 บนแบบสอบ X เช่นเดียวกัน

การปรับเทียบคะแนนสามารถจำแนกตามความแตกต่างของความยากของข้อสอบและความสามารถของผู้สอบ โดยแบ่งได้ 2 ประเภท (Hambleton and Swaminatan, 1985; Kolen and Brennan, 2004; ศิริชัย กาญจนวาสี, 2545) ได้แก่ การปรับเทียบคะแนนตามแนวระดับ (Horizontal Equating) และการปรับเทียบคะแนนตามแนวดิ่ง (Vertical Equating)

**ประเภทของการปรับเทียบคะแนน**

การปรับเทียบคะแนนสามารถจำแนกตามความแตกต่างของความยากของข้อสอบและความสามารถของผู้สอบ โดยแบ่งได้ 2 ประเภท (Hambleton and Swaminatan, 1985; Kolen and Brennan, 2004; ศิริชัย กาญจนวาสี, 2545) ได้แก่ การปรับเทียบคะแนนตามแนวระดับ (Horizontal Equating) และการปรับเทียบคะแนนตามแนวดิ่ง (Vertical Equating)

***การปรับเทียบคะแนนตามแนวระดับ*** *(Horizontal Equating)*

การปรับเทียบคะแนนตามแนวนอนเป็นการปรับเทียบคะแนนระหว่างแบบสอบต่างฉบับกัน เมื่อแต่ละฉบับมุ่งวัดคุณลักษณะเดียวกัน มีระดับความยากใกล้เคียงกัน และกลุ่มผู้สอบมีการแจกแจงความสามารถอยู่ในประชากรเดียวกัน หรือมีความสามารถใกล้เคียงกัน ซึ่งเป็นเทคนิคที่เหมาะสมในสถานการณ์ที่มีความจำเป็นต้องสร้างแบบสอบเนื้อหาเดียวกันขึ้นมาหลาย ๆ ฉบับ เพื่อนำไปใช้ในการทดสอบให้เกิดความยุติธรรมและป้องกันความลับของข้อสอบเมื่อใช้ต่างเวลากันสำหรับกลุ่มผู้สอบขนาดใหญ่ เพื่อปรับเทียบว่าคะแนนที่ได้จากฉบับหนึ่งเทียบเป็นเท่าไรของอีกฉบับหนึ่ง ซึ่งวัดในระดับเดียวกัน จึงเป็นการปรับเทียบคะแนนระหว่างแบบสอบต่างฉบับของวิชาเดียวกัน สำหรับกลุ่มผู้สอบระดับชั้นเดียวกัน

***การปรับเทียบคะแนนตามแนวตั้งหรือแนวดิ่ง*** *(Vertical Equating)*

การปรับเทียบคะแนนตามแนวดิ่ง เป็นการปรับเทียบคะแนนระหว่างแบบสอบต่างฉบับกัน เมื่อแต่ละฉบับมุ่งวัดคุณลักษณะเดียวกัน แต่มีระดับความยากแตกต่างกัน และกลุ่มผู้สอบมีการแจกแจงความสามารถอยู่ต่างประชากรกัน หรือมีความสามารถแตกต่างกันซึ่งเป็นเทคนิคที่เหมาะสมในสถานการณ์ที่มีความจำเป็นต้องสร้างแบบสอบเนื้อหาเดียวกัน แต่ต่างฉบับต่างมุ่งวัดความสามารถของผู้สอบที่ต่างระดับกัน เพื่อปรับเทียบว่าคะแนนที่สอบได้จากฉบับหนึ่ง เทียบเป็นเท่าไรของฉบับอื่นที่วัดต่างระดับกัน จึงเป็นการปรับเทียบคะแนนระหว่างแบบสอบต่างระดับของวิชาเดียวกัน สำหรับกลุ่มผู้สอบที่ต่างระดับชั้นกัน

**การออกแบบการเก็บรวบรวมข้อมูล** (Design for Data Collection)

การเก็บรวบรวมข้อมูลสำหรับนำผลมาใช้ในการปรับเทียบสามารถจำแนกได้ ดังนี้ (Kolen and Brennan, 2004; ศิริชัย กาญจนวาสี, 2545)

1) ***รูปแบบผู้สอบกลุ่มเดียว*** *(Single Group Design)* แบบสอบ 2 ฉบับเชื่อมโยงโดยมีกลุ่มผู้สอบเหมือนกัน ซึ่งเป็นการออกแบบง่าย ๆ แต่อาจจะไม่สามารถปฏิบัติได้ เพราะใช้ระยะเวลาในการทดสอบยาวนาน และก่อให้เกิดความเหนื่อยล้า อาจมีผลต่อการประมาณค่าพารามิเตอร์ การออกแบบรูปแบบผู้สอบกลุ่มเดียว แบ่งเป็น 2 รูปแบบคือ

1.1 ผู้สอบกลุ่มเดียวที่ไม่ได้รับการจัดให้สมดุล (Single group design with uncounterbalancing)

รูปแบบนี้เป็นรูปแบบที่ง่ายที่สุด ใช้ผู้สอบกลุ่มเดียวทำหน้าที่เป็นผู้สอบร่วมวิธีการเก็บรวบรวมข้อมูลทำได้โดยใช้กลุ่มตัวอย่างผู้สอบกลุ่มเดียว แต่ละคนทำแบบสอบทั้งสองฉบับ โดยให้ผู้สอบทำแบบสอบฉบับหนึ่งแล้วตามด้วยแบบสอบอีกฉบับหนึ่ง

การให้ผู้สอบกลุ่มเดียวกันทำแบบสอบทั้ง 2 ฉบับ ระดับความสามารถของผู้สอบทั้ง 2 ฉบับเท่ากัน จึงไม่เป็นปัญหาแทรกซ้อน รวมทั้งไม่มีปัญหาในเรื่องความยากของแบบสอบต่างฉบับ คะแนนจากแบบสอบจึงสามารถปรับเทียบกันได้บนพื้นฐานความสามารถที่เท่ากันจากคนกลุ่มเดียวกัน แต่ในทางปฏิบัติ การสอบฉบับหลัง ผู้สอบอาจได้รับปัจจัยที่ส่งผลกระทบตกทอดจากการสอบฉบับแรก เช่น การเรียนรู้ การฝึกฝน ความเมื่อยล้าจากการทำแบบสอบฉบับแรก เป็นต้น ซึ่งอาจมีอิทธิพลต่อคะแนนการปรับเทียบ

ฉบับแรก

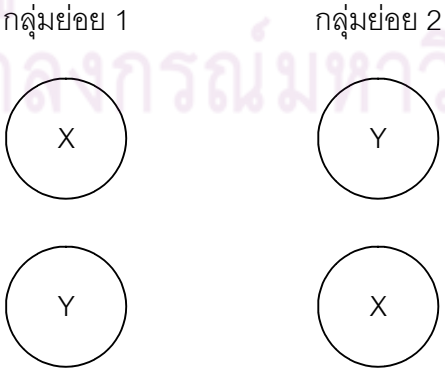
ฉบับหลัง

1.2 ผู้สอบกลุ่มเดียวที่ได้รับการจัดให้สมดุล (Single group design with counterbalancing)

รูปแบบนี้ปรับมาจากรูปแบบแรก เพื่อขจัดผลของลำดับการทดสอบก่อน – หลัง โดยการสุ่มผู้สอบออกเป็น 2 กลุ่มย่อย แต่ละกลุ่มย่อยได้รับการทดสอบด้วยแบบสอบทั้ง 2 ฉบับ ในลักษณะที่กลุ่มย่อยหนึ่งให้ทำแบบสอบฉบับที่ 1 แล้วตามด้วยแบบสอบฉบับที่ 2 ส่วนอีกกลุ่มย่อยหนึ่งให้ทำแบบสอบฉบับที่ 2 แล้วตามด้วยแบบสอบฉบับที่ 1 สลับกันเพื่อให้เกิดความสมดุล

การใช้รูปแบบนี้ กลุ่มผู้สอบจึงได้รับอิทธิพลจากลำดับการทดสอบ การเรียนรู้ การฝึกฝนและความเมื่อยล้าที่สมดุลกันในทั้งสองกลุ่มย่อย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



*2) **รูปแบบผู้สอบกลุ่มเท่าเทียมกัน** (Equivalent – Group Design)*

รูปแบบนี้เป็นการใช้กลุ่มผู้สอบเท่าเทียมกัน แต่ก็ไม่เหมือนกันทุกประการ โดยการจัดกลุ่มผู้สอบให้มีความคล้ายคลึงกันมากที่สุดเท่าที่จะทำได้ (Equivalent group) แล้วให้ผู้สอบในแต่ละกลุ่มทำแบบสอบฉบับเดียว ในขณะที่ Kolen และ Brennan (2004) เรียกการเก็บรวบรวมข้อมูลรูปแบบนี้ว่า เป็นการเก็บรวบรวมข้อมูลแบบผู้สอบกลุ่มสุ่ม (Random Groups)

ถึงแม้การจัดกลุ่มจะไม่เหมือนกันทุกประการ (Identical groups) แต่ก็ควรมีลักษณะทั่วไปคล้ายคลึงกันมากที่สุด (equivalent group) กลุ่มดังกล่าวอาจได้มาโดยการสุ่ม ข้อดีคือ แต่ละกลุ่มทำแบบสอบฉบับเดียว จึงสามารถหลีกเลี่ยงปัญหาเรื่องการเรียนรู้ การฝึกฝน และความเมื่อยล้าจากการทำแบบสอบฉบับแรก แต่ปัญหาที่ตามมาคือ กลุ่มที่ใช้ไม่เหมือนกันอาจมีการแจกแจงความสามารถที่ต่างกัน และไม่มีข้อมูลที่นำมาใช้ปรับความแตกต่างของกลุ่ม ความแตกต่างที่เกิดขึ้นแม้เพียงเล็กน้อยย่อมส่งผลกระทบต่อความลำเอียงในการปรับเทียบคะแนนได้ วิธีการที่ช่วยลดความแตกต่างระหว่างกลุ่มได้โดยการใช้กลุ่มตัวอย่างสุ่มขนาดใหญ่ และไม่ควรเป็นการสุ่มแบบมีระบบ เช่น การจัดที่นั่งให้ผู้สอบในลักษณะของ ชาย – หญิง ชาย – หญิง ชาย – หญิง เป็นต้น ถ้ามีการจัดที่นั่งสอบในลักษณะนี้ผู้สอบเพศชายก็จะได้แบบสอบฉบับ X และผู้สอบเพศหญิงก็จะได้แบบสอบฉบับ Y

*3) **รูปแบบผู้สอบต่างกลุ่มโดยใช้แบบสอบร่วม** (Anchor – Test Design)*

รูปแบบผู้สอบต่างกลุ่มโดยใช้แบบสอบร่วมเป็นการเชื่อมโยงแบบสอบโดยมีกลุ่มผู้สอบ 2 กลุ่มที่แตกต่างกัน แบบสอบแต่ละฉบับจะมีชุดของข้อสอบร่วมที่อาจจะอยู่ภายใน หรือภายนอกแบบสอบ การออกแบบนี้สามารถดำเนินการได้และใช้ได้บ่อย และถ้าเลือกข้อสอบร่วม (anchor items) อย่างเหมาะสม จะสามารถหลีกเลี่ยงปัญหาที่เกิดขึ้นในการออกแบบรูปแบบกลุ่มผู้สอบกลุ่มเดียว (Single group) หรือ รูปแบบกลุ่มผู้สอบเท่าเทียม (Equivalent – group design)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3.1 รูปแบบผู้สอบกลุ่มสุ่มโดยใช้แบบสอบร่วม (Anchor – Test Random Group Design)

รูปแบบนี้เป็นการใช้กลุ่มผู้สอบซึ่งได้จากการสุ่ม แล้วให้ผู้สอบแต่ละกลุ่มทำแบบสอบฉบับเดียว โดยแบบสอบแต่ละฉบับมีข้อสอบร่วมกันจำนวนหนึ่ง (Common or Anchor items) ซึ่งเรียกว่าแบบสอบร่วม (Anchor Test) การใช้แบบสอบร่วมนี้สามารถกระทำได้ 2 ดังนี้

3.1.1 แบบสอบร่วมภายใน (Internal Anchor Test) เป็นแบบสอบที่มีข้อสอบร่วมที่จัดรวมไว้อยู่ในแบบสอบที่ต้องนำมาปรับเทียบคะแนน

3.1.2 แบบสอบร่วมภายนอก (External Anchor Test) เป็นแบบสอบที่มีข้อสอบร่วมที่ถูกจัดแยกออกมาต่างหากจากตัวแบบสอบที่ต้องการนำมาปรับเทียบคะแนน ซึ่งกลุ่มผู้สอบจะต้องทำและจับเวลาแยกออกจากตัวแบบสอบที่ต้องการนำมาปรับเทียบคะแนน

รูปแบบนี้ช่วยลดความลำเอียงในการปรับเทียบคะแนน อันมีสาเหตุมาจากการเรียนรู้ การฝึกฝน ความเมื่อยล้าและความแตกต่างเกี่ยวกับระดับความสามารถของกลุ่มผู้สอบ รูปแบบนี้จึงได้เปรียบเหนือกว่า 2 รูปแบบแรก แต่อย่างไรก็ตามประสิทธิภาพของการปรับเทียบคะแนนขึ้นอยู่กับปัจจัยที่สำคัญบางประการเช่น คุณภาพของแบบสอบร่วม ซึ่งจะต้องคำนึงถึงความสัมพันธ์ระหว่างข้อสอบร่วมที่ต้องมีเนื้อหาและระดับความยากคล้ายกับแบบสอบที่ต้องการปรับเทียบคะแนน และความยาวของแบบสอบร่วมเป็นสำคัญ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3.2 รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกันโดยใช้แบบสอบร่วม (Anchor –Test Non-equivalent Groups Design)

รูปแบบนี้เป็นการใช้กลุ่มผู้สอบที่มีลักษณะไม่เท่าเทียมกัน แล้วให้ผู้สอบแต่ละกลุ่มทำแบบสอบฉบับเดียว กลุ่มละฉบับ โดยแบบสอบแต่ละฉบับมีข้อสอบร่วมกันจำนวนหนึ่ง (Common or Anchor items) ซึ่งเรียกว่าแบบสอบร่วม (Anchor Test) ซึ่งอาจให้แบบสอบร่วมภายในหรือแบบสอบร่วมภายนอกก็ได้

ในการทดสอบบางสถานการณ์ อาจมีความจำเป็นต้องใช้แบบสอบร่วมกับกลุ่มผู้สอบต่างประชากร เช่น กลุ่มผู้สอบต่างโปรแกรม เวลา หรือระดับ เป็นต้น ซึ่งคะแนนจากแบบสอบร่วมจะใช้เป็นกลไกสำคัญสำหรับการปรับเทียบคะแนนระหว่างแบบสอบ

ในการใช้แบบสอบร่วมนี้มีข้อพึงพิจารณา ดังนี้

- เนื้อหาของข้อสอบร่วมควรเป็นตัวแทนเนื้อเรื่องของแบบสอบทั้ง 2 ฉบับ
- ความยากของข้อสอบร่วมควรพอเหมาะกับความสามารถของผู้สอบทั้ง 2 กลุ่ม
- จำนวนข้อสอบร่วมไม่ควรน้อยกว่า 20 ข้อ หรือ 20% ของจำนวนข้อสอบ (Angoff, 1984 อ้างถึงใน ศิริชัย กาญจนวาสี, 2545)
- จำนวนข้อสอบร่วมควรมีประมาณ 20% ถึง 25% ของข้อสอบทั้งฉบับ (Hambleton, Swaminatan และ Roger, 1991 อ้างถึงใน ศิริชัย กาญจนวาสี, 2545)

***4) รูปแบบผู้สอบต่างกลุ่มโดยมีผู้สอบร่วม (Common – Person Design)***

รูปแบบนี้เป็นการใช้กลุ่มผู้สอบต่างกลุ่มกันทำแบบสอบต่างฉบับ แต่มีจำนวนผู้สอบส่วนหนึ่งจากทั้งสองกลุ่มร่วมกันทำแบบสอบทั้ง 2 ฉบับ ซึ่งเป็นรูปแบบของการใช้ผู้สอบร่วม

รูปแบบนี้จึงมีข้อดีในด้านมีผู้สอบร่วมกันทำแบบสอบทั้งสองฉบับ ซึ่งคล้ายรูปแบบกลุ่มเดียวกัน แต่ก็อาจมีปัญหาเกี่ยวกับความเมื่อยล้าของผู้สอบร่วม ลำดับการทำแบบสอบและการเรียนรู้จากแบบสอบได้เช่นเดียวกัน

**การปรับค่าพารามิเตอร์ของแบบสอบต่างฉบับให้อยู่ในมาตรวัดเดียวกัน**

การปรับค่าพารามิเตอร์ของโมเดล IRT เพื่อแปลงหรือเทียบคะแนนมาตรวัดของแบบสอบฉบับอื่นไปยังแบบสอบฉบับพื้นฐาน สามารถจำแนกได้ 2 ประเภทตามเทคนิคการวิเคราะห์ข้อมูลหรือการปรับค่าพารามิเตอร์พร้อมกัน และการปรับค่าพารามิเตอร์แยกกัน มีรายละเอียดดังนี้

*1.1 **การปรับค่าพารามิเตอร์พร้อมกัน** (Simultaneous or Concurrent Calibration)*

เมื่อปรับค่าพารามิเตอร์ของแบบสอบหลายฉบับพร้อม ๆ กัน โดยไม่ทราบค่าประมาณของพารามิเตอร์ข้อสอบร่วม ค่าประมาณทั้งหมดของข้อสอบทั้งที่เป็นข้อสอบร่วมและไม่ได้เป็นข้อสอบร่วมจะถูกเชื่อมโยงโดยอัตโนมัติไปสู่มาตรวัดร่วม (Mislevy and Bock, 1990 cited in Basteri, 2000) เช่น ประมาณค่าพารามิเตอร์แบบสอบฉบับ X และ Y พร้อมกัน โดยการ run เพียง 1 ครั้ง ในโปรแกรม LOGIST สามารถทำโดยใช้ข้อสอบสำหรับฉบับ X และฉบับ Y จากผู้สอบทั้งสองกลุ่ม และข้อสอบที่ผู้สอบกลุ่มหนึ่งไม่ได้ทำจะใช้คำว่า “not reach” เมื่อมีการประมาณค่าในลักษณะนี้ ค่าประมาณที่ได้จะอยู่ในมาตรวัดเดียวกัน

นอกจากนี้ยังมีอีกวิธีหนึ่งเรียกว่าวิธีการกำหนดค่าก่อนปรับพารามิเตอร์ (Fixed Precalibration Item Parameter) (Lord, 1980 cited in Basteri, 2000) ซึ่งเป็นวิธีการที่อยู่ในกลุ่มการปรับค่าพารามิเตอร์พร้อมกัน สองวิธีนี้อาจแตกต่างกันในแนวทางการประมาณค่าพารามิเตอร์ข้อสอบในการดำเนินการ ค่าประมาณพารามิเตอร์ข้อสอบร่วมถูกกำหนดค่าระหว่างการปรับค่าแบบสอบฉบับอื่น ๆ ใหม่ พารามิเตอร์ที่ถูกกำหนดค่าเป็นค่าประมาณที่ถูกจัดการให้เป็นพารามิเตอร์ และค่าเหล่านี้ได้รับจากข้อสอบที่รู้ค่าแล้ว อาจจะได้จากธนาคารข้อสอบ เมื่อแบบสอบหนึ่งฉบับหรือมากกว่าถูกเชื่อมโยงหรือร่วมค่าประมาณพารามิเตอร์ข้อสอบสำหรับข้อสอบใหม่ ถูกนำมาสู่มาตรวัดเดียวกัน วิธีการนี้สมมติว่าค่าประมาณของพารามิเตอร์ที่กำหนดมาเหมาะสมกับข้อมูลและโมเดลเป็นอย่างดี และถูกปรับในแนวทางเดียวกันกับฉบับอื่น

วิธีการกำหนดค่าพารามิเตอร์ และวิธีการปรับค่าพารามิเตอร์พร้อมกันสามารถรวมกันได้ เช่น เมื่อรู้ค่าประมาณพารามิเตอร์ของข้อสอบร่วม ให้กำหนดค่าดังกล่าวในแบบสอบฉบับใหม่ ซึ่งอาจมีหนึ่งฉบับหรือมากกว่าก็ได้ ดังนั้นแบบสอบสองฉบับหรือหลายฉบับจะได้รับการปรับค่าพารามิเตอร์พร้อมกัน ในกรณีนี้เป็นการประยุกต์ใช้วิธีการทั้งสองพร้อมกัน หารวมหรือแยกวิธีทั้งสองขึ้นอยู่กับการประมาณค่าพารามิเตอร์ข้อสอบระหว่างขั้นตอนการปรับค่าพารามิเตอร์

*1.2 **การปรับค่าพารามิเตอร์แยกกัน** (Separate Calibration)*

เมื่อข้อมูลจากการตอบแบบสอบสองฉบับหรือมากกว่า จากกลุ่มผู้สอบที่แตกต่างกัน จะได้รับการปรับค่าพารามิเตอร์แยกกันจึงจำเป็นต้องเชื่อมโยงมาตรวัดโดยใช้สารสนเทศจากข้อสอบร่วม เพื่อแปลค่าพารามิเตอร์ความสามารถเชิงเส้น ซึ่งมาตรวัดความสามารถสำหรับกลุ่มพื้นฐานยังไม่มีการเปลี่ยนแปลงจึงไม่มีความจำเป็นต้องสร้างมาตรวัดขึ้นมาใหม่ ในทางตรงกันข้ามต้องแปลงมาตรวัดความสามารถของกลุ่มเทียบคะแนนไปสู่มาตรวัดที่กำหนดโดยกลุ่มพื้นฐาน ความน่าจะเป็นของการตอบถูกสำหรับข้อสอบข้อใดได้รับจากมาตรวัดค่าความสามารถใหม่ของแต่ละบุคคลสำหรับกลุ่มเทียบคะแนนยังไม่เปลี่ยนแปลง วิธีการประมาณค่าสัมประสิทธิ์การเทียบ

คะแนนได้แก่ ความชัน (A) และจุดตัด (B) มีอยู่หลายวิธี กลุ่มแรกเป็นวิธีการที่อยู่บนพื้นฐานโมเมนต์ (Moment Method) และกลุ่มที่สองเป็นวิธีฟังก์ชันการตอบหรือโค้งคุณลักษณะ (Response Function or Characteristic Curve Method) โดยพื้นฐานแล้ว ทุก ๆ วิธีพยายามหาสมการเชิงเส้นที่ดีที่สุดเพื่อแปลงมาตรวัดจากแบบสอบฉบับหนึ่งไปสู่ฉบับอื่น หรือกล่าวอีกนัยหนึ่งว่าการประมาณค่าความชันและจุดตัด ค่าประมาณพารามิเตอร์ความสามารถและข้อสอบ ได้รับการแปลงจากมาตรวัดของแบบสอบฉบับหนึ่งให้มีประสิทธิภาพมากที่สุดไปสู่มาตรวัดที่ใช้กับฉบับอื่นๆ (Lord, 1980 cited in Macro, 1997)

## ตอนที่ 4 การปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ

การปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ มีการปรับเปลี่ยนค่าความสามารถและพารามิเตอร์ของข้อสอบต่างกลุ่มให้อยู่บนสเกลเดียวกัน ความสัมพันธ์ของการปรับเปลี่ยนเป็นความสัมพันธ์เชิงเส้นตรง ฉะนั้นต้องมีการหาค่าความชัน (A) และค่าคงที่ (B) ในสมการความสัมพันธ์เชิงเส้นตรง วิธีการคำนวณหาค่าความชัน และค่าคงที่ หรือค่าสัมประสิทธิ์ของการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ มีดังต่อไปนี้

**สมการการเปลี่ยนแปลง**

กำหนดสเกล I และสเกล J เป็นสเกลที่แตกต่างกันตามทฤษฎีการตอบสนองข้อสอบ 3 พารามิเตอร์ โดยมีการเปลี่ยนแปลงเชิงเส้นตรง เมื่อ $\theta$ คือ ค่าความสามารถของ 2 สเกลมีความสัมพันธ์กันดังนี้

$$\theta_{Ji} = A\theta_{Ii} + B \quad (1)$$

เมื่อ A และ B เป็นค่าคงที่ในสมการเส้นตรง $\theta_{Ji}$ และ $\theta_{Ii}$ เป็นค่าความสามารถคนที่ i ในสเกล J และสเกล I พารามิเตอร์ข้อสอบของ 2 สเกล มีความสัมพันธ์ดังนี้

$$a_{Jj} = \frac{a_{Ij}}{A} \quad (2)$$

$$b_{Jj} = Ab_{Ij} + B \quad (3)$$

และ $$c_{Jj} = c_{Ij} \quad (4)$$

เมื่อ $a_{Jj}$, $b_{Jj}$ และ $c_{Jj}$ คือพารามิเตอร์ข้อสอบ j ในสเกล J และ $a_{Ij}$, $b_{Ij}$ และ $c_{Ij}$ คือพารามิเตอร์ข้อสอบ j ในสเกล I

การเปลี่ยนแปลงสเกลตามทฤษฎีการตอบสนองข้อสอบจากฟังก์ชันการตอบสนองข้อสอบ

$$P_{ij} = P_{ij}(\theta_i, a_j, b_j, c_j) = c_j + (1-c_j)\frac{\exp\left[Da_j(\theta_i - b_j)\right]}{1+\exp\left[Da_j(\theta_i - b_j)\right]} \quad (5)$$

มีค่า A และ B เป็นผลจากการเปลี่ยนแปลงสเกลจากการเปลี่ยนแปลงพารามิเตอร์จากสเกล I ไปสู่สเกล J พิจารณาทางขวาของฟังก์ชันสำหรับสเกล J เท่ากับ

$$c_{Jj} + (1-c_{Jj})\frac{\exp\left[Da_{Jj}(\theta_{Ji} - b_{Jj})\right]}{1+\exp\left[Da_{Jj}(\theta_{Ji} - b_{Jj})\right]} \quad (6)$$

แทน ($\theta_I$, $a_j$, $b_j$, $c_j$) ในสมการที่ 2 – 4 ได้

$$= c_{Ij} + (1-c_{Ij})\frac{\exp\left\{D\frac{a_{Ij}}{A}\left[A\theta_{Ii} + B - (Ab_{Ij} + B)\right]\right\}}{1+\exp\left\{D\frac{a_{Ij}}{A}\left[A\theta_{Ii} + B - (Ab_{Ij} + B)\right]\right\}}$$

$$= c_{Ij} + (1-c_{Ij})\frac{\exp\left[Da_{Ij}(\theta_{Ij} - b_{Ij})\right]}{1+\exp\left[Da_{Ij}(\theta_{Ij} - b_{Ij})\right]} \quad (7)$$

ผลที่ได้เท่ากับข้างขวาของฟังก์ชันสำหรับสเกล I ที่แสดงว่า A และ B ในสมการที่ 2 – 4 ช่วยในการเปลี่ยนแปลงสเกล

**การหาค่าคงที่ A และ B**

การหาค่า A และ B สำหรับบุคคล i และ $i^*$ หรือข้อสอบ j กับ $j^*$ A และ B ในสมการที่ 2 – 4 สามารถหาค่าได้ดังนี้ (Kolen and Brennan, 2004)

$$A = \frac{\theta_{Ji} - \theta_{Ji^*}}{\theta_{Ii} - \theta_{Ii^*}} = \frac{b_{Jj} - b_{Jj^*}}{b_{Ij} - b_{Ij^*}} = \frac{a_{Ij}}{a_{Jj}} \quad (8)$$

และ $$B = b_{Jj} - Ab_{Ij} = \theta_{Ji} - A\theta_{Ii} \quad (9)$$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 2 ตัวอย่างพารามิเตอร์ข้อสอบและค่าความสามารถของผู้สอบสำหรับ 2 สเกลของแบบสอบ

| | สเกลที่ 1 | | | สเกลที่ 2 | | |
|---|---|---|---|---|---|---|
| พารามิเตอร์ข้อสอบ | | | | | | |
| ข้อที่ | $a_{Ij}$ | $b_{Ij}$ | $c_{Ij}$ | $a_{Jj}$ | $b_{Jj}$ | $c_{Jj}$ |
| j = 1 | 1.30 | -1.30 | .10 | 2.60 | -1.15 | .10 |
| j = 2 | .60 | -.10 | .17 | 1.20 | -.55 | .17 |
| j = 3 | 1.70 | .90 | .18 | 3.40 | -.05 | .18 |
| ความสามารถผู้สอบ | | | | | | |
| คนที่ | $\theta_{Ii}$ | | | $\theta_{Ji}$ | | |
| i = 1 | -2.00 | | | -1.50 | | |
| i = 2 | 1.00 | | | .00 | | |
| ค่าคงที่ที่ใช้ในการเปลี่ยนแปลงสเกล | | | | | | |
| A = .5 | B = -.5 | | | | | |
| ความน่าจะเป็นที่จะตอบข้อสอบได้ถูก | | | | | | |
| | $P_{ij}(\theta_{Ii} ; a_{Ij} , b_{Ij} , c_{Ij})$ | | | $P_{ij}(\theta_{Ji} ; a_{Jj} , b_{Jj} , c_{Jj})$ | | |
| | คนที่ | | | คนที่ | | |
| ข้อที่ | i = 1 | i = 2 | | i = 1 | i = 2 | |
| j = 1 | .26 | .99 | | .26 | .99 | |
| j = 2 | .27 | .80 | | .27 | .80 | |
| j = 3 | .18 | .65 | | .18 | .65 | |

จากตารางแทนค่าได้ดังนี้

$$A = \frac{(-1.15)-(-.55)}{(-1.30)-(-.10)} = \frac{-.6}{-1.2} = .5$$

หรือใช้ความชันจากพารามิเตอร์ของข้อสอบข้อที่ 1 ได้

$$A = \frac{1.3}{2.6} = .5$$

ใช้สมการที่ 10 กับพารามิเตอร์ความยากข้อสอบข้อที่ 1 ได้

$$B = (-1.15) - (.5)(-1.30) = -.5$$

ค่า A และ B ใช้ในการเปลี่ยนแปลงสเกล I ไปสู่สเกล J ได้จากสมการ 1 – 4 ตัวอย่าง เช่น จะเปลี่ยนค่าความสามารถของผู้สอบคนที่ 1 จากสเกล I ไปสู่สเกล J ได้ดังนี้

$$\theta_{J1} = A\theta_{I1} + B = .5(-2.00) + (-.5) = -1.5$$

ค่าความสามารถที่ได้ดังแสดงในตารางที่ 2

เมื่อต้องการเปลี่ยนพารามิเตอร์ของข้อสอบข้อที่ 3 จากสเกล I ไปสู่สเกล J ใช้สมการที่ 2 – 4 จะได้

$$a_{J3} = \frac{a_{I3}}{A} = \frac{1.7}{.5} = 3.4,$$

$$b_{J3} = Ab_{I3} + B = .5(.90) - .5 = -.05,$$

และ $c_{J3} = c_{I3} = .18$

ค่า $P_{ij}$ เป็นความน่าจะเป็นที่ตอบข้อสอบได้ถูก ค่าที่ได้สามารถคำนวณจากค่าความสามารถและพารามิเตอร์ข้อสอบในตารางที่ 2 ค่า $P_{ij}$ เหมือนกันทั้งสเกล I และสเกล J และจะเป็นเอกลักษณ์ซึ่งกันและกัน เมื่อสเกลมีความสัมพันธ์เชิงเส้นตรง

**การหาค่า A และ B ในกลุ่มของข้อสอบหรือกลุ่มบุคคล**

เมื่อต้องการหาความสัมพันธ์ระหว่างสเกลในกลุ่มของข้อสอบหรือกลุ่มผู้สอบ โดยการปรับสมการที่ 9 และ 10 ได้ดังนี้

$$A = \frac{\sigma(b_J)}{\sigma(b_I)} \qquad (10a)$$

$$= \frac{\mu(a_I)}{\mu(a_J)} \qquad (10b)$$

$$= \frac{\sigma(\theta_J)}{\sigma(\theta_I)} \qquad (10c)$$

$$B = \mu(b_J) - A\mu(b_I), \qquad (11a)$$

$$= \mu(\theta_J) - A\mu(\theta_I). \qquad (11b)$$

ตัวอย่างการคำนวณ ใช้ข้อมูลจากตารางที่ 2

$\mu(b_I) = .1667$, $\sigma(b_I) = .8994$, $\mu(a_I) = 1.2$, $\mu(b_J) = -.5833$, $\sigma(b_J) = .4497$ และ $\mu(a_J) = 2.4$ แทนค่าในสมการที่ 10 และ 11 ได้

$$A = \frac{\sigma(b_J)}{\sigma(b_I)} = \frac{\mu(a_I)}{\mu(a_J)} = \frac{.4497}{.8994} = \frac{1.2000}{2.4000} = .50000$$

และ $B = \mu(b_J) - A\mu(b_I) = -.5833 - .5000(-.1667) = -.5000$

ในการปรับเทียบคะแนนในกลุ่มไม่เท่าเทียมกัน การประมาณพารามิเตอร์จากข้อสอบร่วมจะนำมาใช้เป็นประโยชน์สำหรับผู้สอบทั้งสองกลุ่ม การประมาณค่าพารามิเตอร์ของข้อสอบร่วมใช้หาค่าคงตัวของสเกล โดยการแทนลงในสมการเพื่อประมาณหาค่าคงที่

เมื่อนำแบบสอบชุด I ไปสอบกับประชากรกลุ่มที่ 1 และนำแบบสอบชุด J ไปทดสอบกลับกลุ่มที่ 2 ที่มีความแตกต่างจากกลุ่มที่ 1 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของความสามารถไม่สามารถนำไปใช้ในการปรับสเกลได้ในสมการที่ 11c และ 12b เพราะสมการเหล่านี้ใช้กับกลุ่มผู้สอบกลุ่มเดียวที่ทำแบบสอบทั้งสองฉบับ

**วิธีการปรับเทียบตามแนวทฤษฎีการตอบสนองข้อสอบ สำหรับข้อสอบที่ตรวจให้คะแนนสองค่า**

วิธีการเทียบคะแนนตามทฤษฎีดั้งเดิม เป็นการเทียบคะแนนดิบ (Raw scores) ระหว่างแบบสอบ ซึ่งมีจุดอ่อนทั้งในด้านความเสมอภาพ (Equity) ความสมมาตร (Symmetry) และความไม่แปรผันตามกลุ่ม (Invariance) วิธีการเทียบคะแนนตามแนวทฤษฎีการตอบสนองข้อสอบสามารถแก้ปัญหาเหล่านี้ได้ ถ้ามีการเลือกใช้โมเดลการตอบสนองข้อสอบ (Item Response Models) ที่สอดคล้องกับข้อมูล วิธีการเทียบคะแนนตามแนวทฤษฎีการตอบสนองข้อสอบมีหลายวิธี ดังนี้

**วิธีใช้ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าความยาก (Mean and Sigma Method)**

เมื่อต้องการปรับคะแนนจากสเกล I ให้อยู่ในสเกล J หาค่าความชัน (A) และค่าคงที่ (B) จากสมการต่อไปนี้

$$A = \frac{\sigma(b_J)}{\sigma(b_I)}, \qquad B = \mu(b_J) - A\mu(b_I)$$

เมื่อ $\sigma(b_J)$ เป็นส่วนเบี่ยงเบนมาตรฐานของค่าความยากจากแบบสอบชุด J

$\sigma(b_I)$ เป็นส่วนเบี่ยงเบนมาตรฐานของค่าความยากจากแบบสอบชุด I

$\mu(b_J)$ เป็นค่าเฉลี่ยของค่าความยากจากแบบสอบชุด J

$\mu(b_I)$ เป็นค่าเฉลี่ยของค่าความยากจากแบบสอบชุด I

**วิธีใช้ค่าเฉลี่ยของค่าความยากและค่าอำนาจจำแนก (Mean and Mean Method)**

เมื่อต้องการปรับคะแนนจากสเกล I ให้อยู่ในสเกล J หาความชัน (A) และค่าคงที่ (B) จากสมการต่อไปนี้

$$A = \frac{\mu(a_I)}{\mu(a_J)}, \qquad B = \mu(b_J) - A\mu(b_I)$$

เมื่อ $\mu_{(a_J)}$ เป็นค่าเฉลี่ยของค่าอำนาจจำแนกจากแบบสอบชุด J

$\mu_{(a_I)}$ เป็นค่าเฉลี่ยของค่าอำนาจจำแนกจากแบบสอบชุด I

$\mu_{(b_J)}$ เป็นค่าเฉลี่ยของค่าความยากจากแบบสอบชุด J

$\mu_{(b_I)}$ เป็นค่าเฉลี่ยของค่าความยากจากแบบสอบชุด I

**วิธีปรับค่าเฉลี่ยและซิกมา** (Robust Mean and Sigma Method)

การคำนวณตามวิธีค่าเฉลี่ยและซิกมา มิได้พิจารณาถึงระดับความแม่นยำที่เกิดขึ้นจากการประมาณค่าพารามิเตอร์ของข้อสอบแต่ละตัวที่ใช้ในสมการ Linn, et.al ได้เสนอวิธีการปรับแก้ค่าเฉลี่ยและซิกมา โดยนำความคลาดเคลื่อนมาตรฐานของการประมาณค่าพารามิเตอร์ของข้อสอบร่วมเป็นรายข้อมาใช้เป็นตัวถ่วงน้ำหนัก ข้อสอบร่วมข้อที่มีค่าความแปรปรวนระหว่างฉบับสูงจะมีน้ำหนักน้อย ส่วนข้อที่มีค่าแปรปรวนระหว่างฉบับต่ำจะมีน้ำหนักมาก ค่าน้ำหนักเป็นส่วนกลับของความแปรปรวนของค่าประมาณตัวที่มีค่ามาก โดยมีขั้นตอนการคำนวณดังนี้

1) สำหรับแต่ละคู่ของ $(b_{y_{ci}}, b_{x_{ci}})$ คำนวณค่าน้ำหนัก $w_i$

$$w_i = [\text{ค่าตัวมากระหว่าง } \{ v(b_{y_{ci}}), v(b_{x_{ci}}) \}]^{-1}$$

เมื่อ $v(b_{y_{ci}})$ และ $v(b_{x_{ci}})$ เป็นค่าความแปรปรวนของค่าประมาณพารามิเตอร์ b สำหรับ ข้อสอบร่วมฉบับ y และ x

2) กำหนดสเกลของค่าน้ำหนัก

$$W_i' = \frac{W_i}{\sum_{j=1}^{k} W_j}$$

เมื่อ k เป็นจำนวนข้อสอบร่วมสำหรับแบบสอบฉบับ x และ y

3) คำนวณค่าประมาณของพารามิเตอร์ที่ถ่วงน้ำหนัก

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

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

4) คำนวณค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานของค่าประมาณพารามิเตอร์ที่ถ่วงน้ำหนัก

5) คำนวณค่าคงที่ $\alpha$ และ $\beta$ โดยใช้วิธีค่าเฉลี่ยและซิกมา ดังกล่าวข้างต้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

**วิธีโค้งคุณลักษณะข้อสอบ** (The Characteristic Curve Method)

ตามทฤษฎีการตอบสนองข้อสอบ (IRT) เมตริกซ์การเปลี่ยนแปลงเบื้องต้นกำหนดโดย

$$\theta_j^* = A\theta_j + B$$

เมื่อ A เป็นความชัน

B เป็นค่าคงที่

$\theta_j$ เป็นระดับความสามารถของผู้สอบในแบบสอบปัจจุบัน

$\theta_j^*$ เป็นระดับความสามารถของผู้สอบในแบบสอบที่เป็นเป้าหมาย

การใช้สัมประสิทธิ์ A และ B ค่าพารามิเตอร์ของข้อสอบในชุดปัจจุบัน เมื่อ $a_i$ เป็นค่าการกระจายของข้อสอบ และ $b_i$ เป็นค่าความยากของข้อสอบ สามารถเปลี่ยนเป็นค่าพารามิเตอร์ในแบบสอบเป้าหมายได้ดังนี้

$$a_i^* = \frac{a_i}{A} \qquad (12)$$

และ $b_i^* = Ab_i + B$ (13)

เมื่อ i = 1, 2, 3, … แทนจำนวนข้อสอบ

ฮาบารา (Haebara, 1980 cited in Baker, 1997) ได้พัฒนาใช้กับแบบสอบที่มีการให้คะแนนทวิวิภาคหรือแบบ 0 – 1 เป็นคนแรก ต่อมาสต๊อคกิงและลอร์ด (Stocking and Lord, 1983) ปรับสูตรใหม่ในรูปแบบโค้งคุณลักษณะแบบสอบ ผลที่ได้คือค่าคงที่หรือสัมประสิทธิ์การปรับเทียบคะแนนเป็นตัวปรับคะแนนระหว่างแบบสอบ

วิธีโค้งคุณลักษณะข้อสอบนี้เป็นวิธีที่ใช้ทั้งการปรับเทียบคะแนนแนวตั้งและแนวนอน สำหรับการตรวจให้คะแนนแบบทวิวิภาคหรือแบบ 0 – 1 ในกรณีที่จำเป็นต้องประมาณค่าพารามิเตอร์ของแบบสอบแยกกัน เป็นวิธีการที่ใช้โค้งคุณลักษณะข้อสอบเพื่อหาสัมประสิทธิ์การปรับเทียบ (Equating Coefficients) มีรายละเอียดดังนี้

ในกรณีการตอบแบบสอบที่มีการตรวจให้คะแนน 0 – 1 กำหนดให้ $P_i(\theta_j)$ แทนค่าความน่าจะเป็นของการตอบถูกของข้อสอบร่วม (Anchor Test) ในแบบสอบที่เป็นหลัก และ $P^*_i(\theta_j)$ แทนความน่าจะเป็นของการตอบถูกของข้อสอบร่วมในแบบสอบที่ปรับคะแนนแล้ว ถ้าหากว่า $[P_i(\theta_j) - P^*_i(\theta_j)] = 0$ แสดงว่าการปรับเทียบคะแนนมีความสมบูรณ์ ดังนั้นการปรับเทียบคะแนนคือการหาสัมประสิทธิ์การปรับเทียบคะแนน A และ B ได้จากเกณฑ์ที่กำหนดค่าผลต่างจากสมการ

$$F = \frac{1}{N}\sum_{j=1}^{N}\left\{\sum_{i=1}^{n}\left[P_i(\theta_j) - P_i^*(\theta_j)\right]\right\}^2, \qquad (14)$$

เมื่อ N คือจำนวนจุดบน θ ที่กำหนดขึ้นเองและ n คือจำนวนข้อสอบร่วมของทั้งสองแบบสอบ การกำหนด $P_i(\theta_j)$ และ $P^*_i(\theta_j)$ ขึ้นอยู่กับโมเดลที่ใช้ในฟังก์ชันการตอบสนองข้อสอบ ถ้าใช้โมเดล 2 พารามิเตอร์จะได้

$$P_i(\theta_j) = \frac{1}{1+\exp[-a_i'(\theta_j - b_i')]}$$

และ $$P_i(\theta_j) = \frac{1}{1+\exp[-a_i^*(\theta_j - b_i^*)]}$$

เมื่อ $a_i$ และ $b_i$ เป็นพารามิเตอร์ของข้อสอบร่วมจากแบบสอบหลัก

$a^*_i$ และ $b^*_i$ เป็นผลจากการใช้สมการที่ 12 และ 13 เพื่อประมาณค่าพารามิเตอร์ข้อสอบที่ปรับเทียบ

$\theta_j$ แทนเมตริกซ์ร่วมของคะแนนความสามารถ

ส่วนความสัมพันธ์ของผลรวมอันดับ 2 ในสมการที่ 14 กำหนดเป็นคะแนนจริงทั้งสองแบบสอบ สัมประสิทธิ์การปรับเทียบคะแนนสามารถหาได้จาก A Quadratic Loss Function กำหนดในรูปของโค้งคุณลักษณะข้อสอบดังนี้ (Stocking and Lord, 1983)

$$F = \frac{1}{N}\sum_{j=1}^{N}(T_j - T_j^*)^2,$$

เมื่อ $T_j = \sum_{i=1}^{n} P_i(\theta_j),$ และ $T_j^* = \sum_{i=1}^{n} P_i^*(\theta_j)$

ตามการหาค่าสัมประสิทธิ์การปรับเทียบคะแนน A และ B ที่ได้จากการใช้เทคนิคการค้นหาหลายตัวแปรหลายครั้ง (The Iterative Multivariate Search Technique) ซึ่งเดวิดดอน (Davidon, 1959 cited in Baker, 1997) และเฟรซเซอร์กับพาเวลล์ (Fletcher and Powell, 1963 cite in Baker, 1997) พบว่าเมื่อค่า F ลดลงมากที่สุดจะได้ค่าสัมประสิทธิ์การปรับเทียบ ถ้าการปรับเทียบคะแนนสมบูรณ์ค่า F = 0

**วิธีประมาณค่าไคว์สแคว์น้อยที่สุด (The Minimum $\chi^2$ Method)**

วิธีประมาณค่าไคว์สแคว์น้อยที่สุดเสนอโดย ดิฟกิ (Divgi, 1985 cited in Kim and Cohen, 1995) เป็นวิธีที่ใช้ทั้งการปรับเทียบคะแนนแนวตั้งและแนวนอน สำหรับการตรวจให้คะแนนแบบทวิวิภาค หรือแบบ 0 – 1 ในกรณีที่จำเป็นต้องประมาณค่าพารามิเตอร์ของข้อสอบแยกกัน เป็นวิธีที่วิเคราะห์ใช้เกณฑ์ที่ทำให้ค่า $\chi^2$ น้อยที่สุดเพื่อจะให้ได้ค่าสัมประสิทธิ์การปรับเทียบคะแนน มีรายละเอียดดังนี้

ลอร์ด (Lord, 1980 cite in Kim and Cohen, 1995) ได้แสดงให้เห็นว่าใน IRT ความสัมพันธ์ระหว่างเมตริกซ์พารามิเตอร์ข้อสอบของสองข้อสอบที่เปรียบเทียบเป็นเชิงเส้นตรง เมื่อปรับค่าพารามิเตอร์ข้อสอบจากฉบับที่ 2 ให้เทียบกับเมตริกซ์ฉบับที่ 1 การประมาณค่าอำนาจจำแนก (a) และความยากของข้อสอบ (b) ของข้อสอบ i หาได้จาก

$$a^{*}_{i2} = \frac{a_{i2}}{A} \qquad \text{และ} \qquad b^{*}_{i2} = Ab_{i2} + B,$$

เมื่อ * แทนค่าที่ปรับของแบบสอบชุดที่ i

A คือ ค่าความชัน

B คือ ค่าคงที่

คะแนนความสามารถ ($\theta$) ของคนที่ j แสดงได้ดังนี้

$$\theta^{*}_{j2} = A\theta_{j2} + B$$

การปรับเทียบสองเมตริกซ์ต้องหาสัมประสิทธิ์การปรับเทียบ A และ B ที่เหมาะสมโดยวิธีประมาณค่าไคว์สแคว์ต่ำสุด เพื่อให้ได้สัมประสิทธิ์การปรับเทียบคะแนนจากสมการ

$$\chi^2 = \sum_{i=1}^{n} \chi_i^2 = \sum_{i=1}^{n} \xi_i' \sum_i^{-1} \xi_i$$

เมื่อ n แทนจำนวนข้อสอบทั้งหมด

$\xi$ แทนเวกเตอร์ความแตกต่างระหว่างพารามิเตอร์ข้อสอบครั้งแรกและครั้งที่ 2 ที่ปรับแล้ว

และ $\sum_i$ แทนเมตริกซ์ความแปรปรวน – ความแปรปรวนร่วมของ $\xi_i$

สำหรับโมเดล IRT – 2 พารามิเตอร์ มีตัวอย่างดังนี้

$$\xi_i = \xi_{i1} - \xi^{*}_{i2} \quad \text{และ} \quad \sum_i = \sum_{i1} - \sum^{*}_{i2}$$

เมื่อ $\xi_{i1} = (a_{i1}, b_{i1})'$ และ $\xi^{*}_{i2} = (a^{*}_{i2}, b^{*}_{i2})'$

$\sum_{i1}$ แทน เมตริกซ์ความแปรปรวนร่วม มิติ 2x2 ของความคลาดเคลื่อนของการสุ่มตัวอย่างข้อสอบ i จากแบบสอบชุดแรก และ

$\sum^{*}_{i2}$ แทน เมตริกซ์ความแปรปรวนร่วมที่ปรับจากแบบสอบชุดที่ 2

จะได้

$$\chi^2 = \sum_{i=1}^{n} (a_{i1} - a^{*}_{i2},\ b_{i1} - b^{*}_{i2})(\sum_{i1} - \sum^{*}_{i2})(a_{i1} - a^{*}_{i2},\ b_{i1} - b^{*}_{i2})' \qquad (15)$$

$\chi^2$ นี้เป็นฟังก์ชันของสองสัมประสิทธิ์การปรับเทียบที่ยังไม่ทราบค่าคือ A และ B ได้ B จากการได้มาบางส่วน (Partial Derivative) ของสมการที่ 15 กับการพิจารณาที่ $B: \frac{\partial \chi^2}{\partial B} = 0$ การได้มา

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บางส่วนมีความสัมพันธ์เชิงเส้นตรงกับการพิจารณา B และทำให้หาค่า A ได้ง่ายขึ้น แสดงให้เห็นว่า $S_{iab}$ และ $S_{ibb}$ เป็นส่วนประกอบเฉพาะจากเมตริกซ์

$$S_i = \sum_i^{-1} = (\sum_{i1} + \sum_{i2}^{*})^{-1}$$

เมื่อ $$B = \frac{\sum_{i=1}^{n}\left[(\frac{a_{i1} - a_{i2}}{A})S_{iab} + (b_{i1} - Ab_{i2})S_{ibb}\right]}{\sum_{i=1}^{n} S_{ibb}}$$

เมื่อค่า B ได้จากสมการนี้แทนจากที่ได้ $\chi^2$ ทำให้หาค่า A ได้ง่ายขึ้น โดยใช้วิธีของนิวตัน (Newton's Method) จากสมการดังต่อไปนี้

$$A^{(s)} = A^{(s-1)} - \left[H^{(s-1)}\right]^{-1} f^{(s-1)},$$

เมื่อ s เป็นจำนวนรอบที่ทำซ้ำ

โดยที่ $f^{(s-1)} = \frac{\partial \chi^2}{\partial A_{A^{(s-1)}}}$, และ $H^{(s-1)} = \frac{\partial^2 \chi^2}{\partial^2 A^2_{A^{(s-1)}}}$

จากที่กล่าวมาข้างต้นเป็นวิธีการปรับเทียบคะแนนสำหรับแบบสอบที่มีการตรวจให้คะแนน 2 ค่า ผู้วิจัยได้ศึกษางานวิจัยเกี่ยวกับการเทียบคะแนนในแบบทดสอบที่ข้อสอบตรวจให้คะแนนหลายค่า (Baker, 1992, 1993, 1997; Huynh and Ferrara, 1994; Kim and Cohen. 1995; Hennings, Hirsch and Zhang, 1996; Cohen and Kim, 1998; Li, Lissitz and Yang, 1999; Bastari, 2000; Lee and others, 2001) **สรุปวิธีการเทียบคะแนนในแต่ละโมเดลที่ข้อสอบตรวจให้คะแนนหลายค่า ตามลักษณะของโมเดลการตอบ** ดังต่อไปนี้

**วิธีการปรับเทียบตามแนวทฤษฎีการตอบสนองข้อสอบ สำหรับข้อสอบทีตรวจให้คะแนนหลายค่า**

**โมเดลเกรด เรสพอนส์** (Grade Response Model : GRM)

มีการเปรียบเทียบคะแนนในโมเดลนี้หลายวิธี คือ

1. **วิธีที่ใช้พารามิเตอร์ไม่พร้อมกันทุกค่าในการประมาณค่าสัมประสิทธิ์การเทียบคะแนน**

โคเฮนและคิม (Cohen and Kim, 1998) ได้ศึกษาการเทียบคะแนนโมเดลเกรด เรสพอนส์ในกลุ่มวิธีการนี้ 3 วิธี คือ

1.1 วิธีการใช้ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของพารามิเตอร์ความยาก (วิธี mean และ sigma : MS)

มาร์โค (Macro, 1977) ได้เสนอวิธีการนี้โดยสมมุติให้ค่าประมาณพารามิเตอร์ตำแหน่ง (location) จากกลุ่ม 1 คือ $b_{jk1}$ กลุ่ม 2 คือ $b_{jk2}$ โดยที่ $j = 1, \ldots, n_c$ ; $k = 1, \ldots, m_j - 1$ และ $n_c$ เป็นข้อสอบร่วมในการเชื่อมต่อความสัมพันธ์เชิงเส้นของพารามิเตอร์ข้อสอบมีดังนี้

$$b_{jk1} = Ab_{jk2} + B \tag{16}$$

และ

$$a_{j1} = \frac{a_{j2}}{A} \tag{17}$$

สัมประสิทธิ์การเชื่อมต่อวิธี MS คือ A และ B ได้จากความสัมพันธ์ดังนี้

$$\bar{b}_1 = A\bar{b}_2 + B \tag{18}$$

เมื่อ $\bar{b}_1$ และ $\bar{b}_2$ คือ ค่าเฉลี่ยของ $b_{jk1}$ และ $b_{jk2}$ ตามลำดับ และ

$$S(b_1) = AS(b_2) \tag{19}$$

เมื่อ $S(b_1)$ และ $S(b_2)$ คือส่วนเบี่ยงเบนมาตรฐานของ $b_{jk1}$ และ $b_{jk2}$ ตามลำดับ ดังนั้น A และ B ได้จาก

$$A = \frac{S(b_1)}{S(b_2)} \tag{20}$$

และ $$B = \bar{b}_1 - A\bar{b}_2 \tag{21}$$

1.2 วิธีการใช้ค่าเฉลี่ยของพารามิเตอร์ความยากและอำนาจจำแนก (วิธี mean และ mean : LH)

ลอยด์และฮูเวอร์ (Loyd และ Hoover. 1980 cited in Cohen and Kim. 1998 : 118 – 119) ได้เสนอให้ประมาณค่า A โดยใช้อัตราส่วนของค่าเฉลี่ยอำนาจจำแนกที่ได้จากการประมาณค่าแบบทดสอบ 2 ฉบับ วิธีการนี้ใช้กับโมเดลราสช์มาแล้ว ต่อมาเบเกอร์ และ อัล – คานิ (Baker และ Al–Kani, 1991 : 161) ได้ใช้วิธี LH เพื่อเชื่อมต่อการวัดในโมเดล IRT โลจีสติก 3 พารามิเตอร์ ความสัมพันธ์เชิงเส้นของพารามิเตอร์อำนาจจำแนก คือ

$$a_{j1} = \frac{a_{j2}}{A}$$

A สำหรับวิธี LH ในโมเดล GRM จะได้จาก

$$A = \frac{\bar{a}_2}{\bar{a}_1} \quad (22)$$

และ $$B = \bar{b}_1 - A\bar{b}_2 \quad (23)$$

เมื่อ $\bar{a}_1$ และ $\bar{a}_2$ คือ ค่าเฉลี่ยของ $a_{j1}$ และ $a_{j2}$ ตามลำดับ

ข้อจำกัดของวิธี MS และ LH คือ ถ้าประมาณค่าความยากของข้อสอบได้ไม่ดี อาจทำให้ค่า A และ B มีค่าเป็นลบ

1.3 วิธีการใช้ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของพารามิเตอร์ความยากที่กำหนดน้ำหนัก (วิธี Weighted mean และ sigma : WMS)

ลินน์และคณะ (Linn, et al., 1980, 1981 cited in Cohen and Kim, 1998 : 119 – 120) ได้แก้ไขปัญหาที่พบในวิธี MS และ LH โดยปรับปรุงกระบวนการ MS เพื่อรวมน้ำหนักของค่าประมาณความยากของข้อสอบ โดยอินเวอร์สของความคลาดเคลื่อนมาตรฐานกำลังสองมีขนาดใหญ่กว่า ต่อมาสตอกกิงและลอร์ด (Stocking and Lord. 1983 : 203, 208 – 209) กำหนดน้ำหนักโดยใช้ผลรวมของน้ำหนักระหว่างข้อสอบทุกข้อ สำหรับโมเดล GRM กำหนดน้ำหนักสำหรับค่าประมาณพารามิเตอร์ตำแหน่งของข้อสอบ j ลำดับขั้น k คือ $w_{jk}$ ดังนี้

$$W_{jk} = \frac{\left\{\max\left[SE(b_{jk1}), SE(b_{jk2})\right]\right\}^{-2}}{\sum_{j=1}^{n_c}\sum_{k=1}^{m_j-1}\left\{\max\left[SE(b_{jk1}), SE(b_{jk2})\right]\right\}^{-2}} \quad (24)$$

ค่าประมาณพารามิเตอร์ตำแหน่งที่กำหนดน้ำหนักคือ

$$b^{w}_{jk1} = W_{jk} b_{jk1} \quad (25)$$

และ

$$b^{w}_{jk2} = W_{jk} b_{jk2} \quad (26)$$

แล้ว

$$\bar{b}^{w}_{1} = A\bar{b}^{w}_{2} + B \quad (27)$$

และ

$$S(b^{w}_{1}) = AS(b^{w}_{2}) \quad (28)$$

เมื่อ $\bar{b}^{w}_{1}$ และ $S(b^{w}_{1})$ คือค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานที่กำหนดน้ำหนักค่าดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

$$\bar{b}_1^{w} = \frac{\sum_{j=1}^{n_c}\sum_{k=1}^{m_j-1} W_{jk} b_{jk1}}{\sum_{j=1}^{n_c}\sum_{k=1}^{m_j-1} W_{jk}} \quad (29)$$

และ

$$S(b_1^{w}) = \sqrt{\frac{\sum_{j=1}^{n_c}\sum_{k=1}^{m_j-1}(W_{jk} b_{jk1} - \bar{b}_1^{w})^2}{\sum_{j=1}^{n_c}\sum_{k=1}^{m_j-1} W_{jk}}} \quad (30)$$

$\bar{b}_2^{w}$ และ $S(b_2^{w})$ กำหนดค่าไว้เช่นเดียวกัน สำหรับค่า A และ B คำนวณจาก

$$A = \frac{S(b_1^{w})}{S(b_2^{w})} \quad (31)$$

และ

$$B = \bar{b}_1^{w} - A\bar{b}_2^{w} \quad (32)$$

นอกเหนือจากนี้ยังมีวิธีการอื่นในกลุ่มนี้ที่โคเฮนและคิมไม่ได้ศึกษาไว้คือ วิธี robust mean และ sigma ซึ่งอธิบายโดยบีจาและวิงเกอสกี (Bejar and Wingersky. 1981 cited in Cohen and Kim, 1998 : 120) และวิธีการทำซ้ำของ mean และ sigma โดยสตอกกิงและลอร์ด (Stocking and Lord. 1983 : 203, 209)

**2. วิธีฟังก์ชันการตอบหรือโค้งลักษณะ (Response Function Method or Characteristic Curve Method : TRF)**

2.1 เบเกอร์ (Baker. 1992 : 89 – 92) ได้ขยายวิธีฟังก์ชันการตอบแบบทดสอบ (TRF) ของสตอกกิงและลอร์ด (Stocking and Lord, 1983 : 203 – 205) ไปยังโมเดล GRM เทคนิคของเบเกอร์ในการคำนวณสัมประสิทธิ์เชิงเส้น 2 ค่าโดยการทำให้ฟังก์ชันรอส (Loss function) กำลังสองมีค่าน้อยที่สุด

$$F = \frac{1}{N}\sum_{i=1}^{N}(T_{i1} - T_{i2}^{*})^2 \quad (33)$$

เมื่อ N คือ จำนวนของจุดในการวัด $\theta_i$ และ

$T_{i1}$ และ $T^{*}_{i2}$ คือค่าคาดหวังของจำนวนคะแนนที่ตอบถูกสำหรับกลุ่ม 1 และกลุ่ม 2 กำหนดไว้ดังนี้

$$T_{i1} = \sum_{j=1}^{n_c} \sum_{k=1}^{m_j} U_{jk} \hat{P}_{jk1}(\theta_i) \tag{34}$$

$$T^{*}_{i2} = \sum_{j=1}^{n_c} \sum_{k=1}^{m_j} U_{jk} \hat{P}^{*}_{jk2}(\theta_i) \tag{35}$$

เมื่อ $U_{jk}$ คือ น้ำหนักที่กระจายไปตามลำดับขั้นการตอบ k (Response category) ในข้อสอบ j เป็นดัชนีที่เป็นจำนวนเต็มของลำดับขั้นการตอบ (category)

2.2 วิธีไคสแควร์ต่ำสุด (Minimum $\chi^2$ Method : MCS)

คิมและโคเฮน (Kim and Cohen, 1995 : 169 – 171) ขยายวิธีไคสแควร์ต่ำสุด (MCS) ของดีฟจิ (Divgi, 1985 : 413 – 414) ไปยังโมเดล GRM วิธีการ คือ การทำให้ฟังก์ชันกำลังสองมีค่าต่ำสุด ดังสมการต่อไปนี้

$$\chi^2 = \sum_{j=1}^{n_c} \chi^2_{jm_j} = \sum_{j=1}^{n_c} \hat{\xi}'_{jm_j} \hat{\Sigma}^{-1}_{jm_j} \hat{\xi}_{jm_j} \tag{36}$$

เมื่อ

$$\hat{\xi}_{jm_j} = \hat{\xi}_{jm_j 1} - \hat{\xi}^{*}_{jm_j 2} \tag{37}$$

$$\hat{\xi}_{jm_j 1} = (a_{j1}, b_{j11}, \ldots b_{jk1}, \ldots b_{j(m_j-1)1})' \tag{38}$$

$$\hat{\xi}^{*}_{jm_j 2} = (a^{*}_{j2}, b^{*}_{j12}, \ldots b^{*}_{jk2}, \ldots b^{*}_{j(m_j-1)2})' \tag{39}$$

$$\hat{\Sigma}_{jm_j} = \hat{\Sigma}_{jm_j 1} + \hat{\Sigma}^{*}_{jm_j 2} \tag{40}$$

เมื่อ $\hat{\Sigma}_{jm_j 1}$ คือค่าประมาณเมตริกซ์ความแปรปรวน – ความแปรปรวนร่วม (variance – covariance matrix) ของ $\hat{\xi}_{jm_j 1}$

เมื่อ $\hat{\Sigma}^{*}_{jm_j 2}$ คือค่าประมาณเมตริกซ์ความแปรปรวน – ความแปรปรวนร่วม (variance – covariance matrix) ที่แปลงแล้วของ $\hat{\xi}^{*}_{jm_j 2}$

**3. วิธีปรับค่าพารามิเตอร์พร้อมกัน**

เฮนนิง เฮิร์ช และแซง (Hennings, Hirsch and Zhang. 1996: 4 – 5) ได้ใช้วิธีปรับค่าพารามิเตอร์พร้อมกันชนิดกำหนดค่าในโมเดลเกรด เรสพอนส์ สำหรับข้อสอบที่เป็นคำถาม

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ปลายเปิด ต่อมาบาสทาริ (Bastari. 2000: 26) ได้ใช้วิธีการนี้ในแบบทดสอบข้อสอบผสม ซึ่งข้อสอบที่ตรวจให้คะแนนหลายค่าเป็นข้อสอบเขียนตอบถูกวิเคราะห์ค่ารวมกับข้อสอบที่ตรวจให้คะแนนสองค่าที่เป็นข้อสอบเลือกตอบในโมเดลเกรด เรสพอนส์ และโลจีสติก 3 พารามิเตอร์ โดยแบบทดสอบเทียบคะแนนทั้งสองฉบับปรับค่าพารามิเตอร์พร้อมกัน

**โมเดลนอมินอล เรสพอนส์ (Nominal Response Model : NRM)**

เบเกอร์ (Baker, 1993 : 243 – 245) ได้ขยายฟังก์ชันการตอบแบบทดสอบ (TRF) ของสตอกกิงและลอร์ด (Stocking and Lord, 1983 : 203 – 205) ไปยังโมเดลนอมินอลเทคนิคของเบเกอร์ สำหรับการหาสัมประสิทธิ์เชิงเส้น คือ การทำให้ค่าฟังก์ชันรอสกำลังสองมีค่าต่ำสุดโดย

$$F = \frac{1}{NS}\sum_{i=1}^{N}\sum_{j=1}^{n}\sum_{k=1}^{m_j}\left[P_{jk}(\theta_i) - P^{*}_{jk}(\theta_i)\right]^2 \qquad (41)$$

เมื่อ

$$S = \sum_{j=1}^{n} m_j \qquad (42)$$

เมื่อ $k = 1, \ldots, m_j$ เป็นลำดับขั้นการตอบข้อสอบ

$j = 1, \ldots, n$ เป็นจำนวนข้อสอบ

$i = 1, \ldots, N$ เป็นจำนวนผู้สอบ

เบอเกอร์ได้สร้างโปรแกรม EQUATE 2.0 (Baker, 1993b : 20) ไว้สำหรับหาค่าสัมประสิทธิ์การเทียบคะแนนที่ใช้โค้งลักษณะแบบทดสอบในโมเดลเกรด เรสพอนส์ และโมเดลนอมินอล เรสพอนส์

**โมเดลพาเชียล เครดิต (Partial Credit Model : PCM)**

ฮันและเฟอร์รารา (Huynh and Ferrara, 1994 : 134 – 135) ได้เทียบคะแนนในโมเดลพาเชียล เครดิต ด้วยโค้งลักษณะแบบทดสอบ โดยกำหนดค่าเฉลี่ยความยากของข้อสอบเท่ากับศูนย์ทั้งสอบฉบับ แปลงค่าพารามิเตอร์ความสามารถเชิงเส้นแล้วเทียบคะแนนจริง นอกจากนี้ เฮนนิง เฮิร์ช และแซง (Hennings, Hirsch and Zhang, 1996 : 5 – 6) ได้เทียบคะแนนในโมเดลพาเชียลเครดิตวิธีปรับค่าพารามิเตอร์พร้อมกันด้วยการกำหนดค่าพารามิเตอร์

**โมเดลเจเนอรอลไลซ์ พาเชียล เครดิต (Generalize Partial Credit Model : GPCM)**

ไล ลิซซี่ และแยง (Li, Lissitz and Yang, 1991 : 7 – 8) ได้กล่าวถึงการแปลงค่าประมาณพารามิเตอร์ข้อสอบของโมเดล GPCM คือ

$$b^{*}_{B} = Ab_{E} + B \qquad (43)$$

$$d^{*}_{kB} = Ad_{kE} \qquad (44)$$

$$b^{*}_{kB} = Ab_{kE} + B \text{ , โดยที่ } (b_k = b - d_k) \quad (45)$$

$$a^{*}_{B} = \frac{a_E}{A} \quad (46)$$

เมื่อ $a^{*}_{B}$ แทนค่ามาตรวัดที่แปลงไปสู่มาตรวัดพื้นฐาน

A คือ ความชัน และ B คือจุดตัด

สัญลักษณ์ B, E แทนกลุ่มพื้นฐานและกลุ่มเทียบคะแนน

หลังจากการแปลงข้างต้น ค่าประมาณพารามิเตอร์ข้อสอบซึ่งขึ้นอยู่กับมาตรวัดความสามารถของกลุ่มเทียบคะแนนเป็นการอ้างอิงใหม่ไปยังมาตรวัดความสามารถของกลุ่มพื้นฐาน

$$\hat{\theta}^{*}_{B} = A\hat{\theta}_{E} + B \quad (47)$$

สำหรับการเทียบคะแนนในโมเดล GPCM จากการศึกษางานวิจัยปรากฏว่ามีการเทียบคะแนนโดยใช้วิธีโค้งลักษณะ (Characteristic Curve Method : CCM) ซึ่งเป็นการศึกษาในแบบทดสอบที่มีรูปแบบผสม (Li, Lissitz and Yang. 1999 : 8 – 11) สัมประสิทธิ์การเทียบคะแนนของ A และ B ได้จากการวิเคราะห์ค่าเฉลี่ยความแตกต่างกำลังสองระหว่างคะแนนจริงที่ถูกคาดหวัง 2 ค่า จาก 2 ชุดของค่าประมาณพารามิเตอร์ข้อสอบร่วมสำหรับ N ระดับความสามารถที่กำหนดให้ค่า N ต่างกันน้อยที่สุดเท่าที่เป็นไปได้ ค่าประมาณพารามิเตอร์การเทียบคะแนนค่อนข้างจะคงที่เมื่อ N มีค่ามากกว่า 100 ฟังก์ชันต่อไปนี้ถูกทำให้มีค่าต่ำที่สุด (Hambleton, Swaminathan และ Rogers. 1991) คือ

$$f(A,B) = \frac{1}{N}\sum_{j=1}^{N}\left[t(\hat{\theta}_{jB}) - t^{*}(\hat{\theta}_{jE})\right]^2 \quad (48)$$

เมื่อ

$$t(\hat{\theta}_{jB}) = \sum_{i=1}^{L}\sum_{k=1}^{m} x_{ik} P\left(X_i = x_{ik} \middle| \hat{\theta}_B\right) \quad (49)$$

และ

$$t^{*}(\hat{\theta}_{jE}) = \sum_{i=1}^{L}\sum_{k=1}^{m} x_{ik} P\left(X_i = x_{ik} \middle| \hat{\theta}_E\right) \quad (50)$$

โดยที่ $X_{ik}$ เป็นคะแนนสังเกตของลำดับขั้น k ในข้อสอบ i อยู่ในช่วง 1 – 4 (0 หรือ 3) สำหรับข้อสอบที่มีการให้คะแนน 4 ลำดับขั้น

L เป็นจำนวนของข้อสอบร่วม

M เป็นจำนวนของลำดับขั้นคะแนน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

$P(\theta)$ เป็นความน่าจะเป็นของลำดับขั้นการตอบ k สามารถคำนวณได้จากโมเดล GPCM คะแนนจริงที่คาดหวังของผู้สอบที่มีความสามารถ $\theta_j$ บนกลุ่มของข้อสอบร่วม ในแบบทดสอบพื้นฐาน (B) และแบบทดสอบเทียบคะแนน (E) คำนวณในสมการ (49) และ (50) ตามลำดับ

สัมประสิทธิ์ A และ B คำนวณจากการทำให้ฟังก์ชัน (สมการ 48) มีค่าต่ำสุดด้วยการหาอนุพันธ์ของฟังก์ชันเทียบกับ A และ B โดยกำหนดสมการของอนุพันธ์ย่อย (partial derivative) ทั้งสองสมการให้มีค่าเท่ากับ 0 โดยการกำหนดค่า B เป็นค่าคงที่ $B_0$ และพิจารณาฟังก์ชัน $f(A, B_0)$ เป็นฟังก์ชันที่มีคำตอบเดียวในกรณีของฟังก์ชัน f(A, B) ที่มี 2 คำตอบ (สมการ 48) การประมาณค่าความแตกต่างสำหรับการหาอนุพันธ์ย่อยเทียบกับ A ที่ $A_0$ สามารถเขียนสมการดังนี้ (Nakamura. 1996 cited in Li, Lissitz and Yang. 1999 : 9)

$$\frac{\partial f(A,B)}{\partial A} \approx \frac{f((A_0+\Delta A),B_0)-f(A_0,B_0)}{\Delta A}=0 \qquad (51)$$

เมื่อ $\Delta A$ เป็นช่วงระหว่างจุดที่ต่อเนื่องกัน 2 จุดบนเส้นจำนวนในการประมาณค่าของ A การกำหนดโดยทั่วไปจะกำหนดค่าที่เล็กมากๆ เช่น 0.001 ด้วยหลักการเดียวกันประยุกต์ใช้การหาอนุพันธ์ย่อยเทียบกับ B ณ ค่า $B_0$ จะได้

$$\frac{\partial f(A,B)}{\partial B} \approx \frac{f(A_0,(B_0+\Delta B))-f(A_0,B_0)}{\Delta B}=0 \qquad (52)$$

เมื่อ $\Delta B$ เป็นช่วงระหว่างจุดที่ต่อเนื่องกัน 2 จุดบนเส้นจำนวนในการประมาณค่าของ B

สมการทั้ง 2 สมการ (62 และ 63) ไม่ใช่สมการเชิงเส้น สามารถแก้สมการเพื่อหาค่า A และ B ได้โดยใช้กระบวนการ Newton – Raphson ด้วยการทำซ้ำ t ครั้ง (Baker. 1992 cited in Li, Lissitz and Yang. 1999 : 10) ค่าประมาณพารามิเตอร์ของข้อสอบร่วมในแบบทดสอบเทียบคะแนนนำมาสร้างมาตรวัดใหม่ โดยการหาค่า A และ B ที่ทำซ้ำ t ครั้ง ขณะที่คำนวณค่าที่คาดหวังของคะแนนจริง (สมการ 50) สมการที่ทำซ้ำ คือ

$$\begin{bmatrix}\hat{A}\\ \hat{B}\end{bmatrix}_{t+1}=\begin{bmatrix}\hat{A}\\ \hat{B}\end{bmatrix}_{t}-\begin{bmatrix}\frac{\partial^2 f}{\partial A^2} & \frac{\partial^2 f}{\partial A\partial B}\\ \frac{\partial^2 f}{\partial B\partial A} & \frac{\partial^2 f}{\partial B^2}\end{bmatrix}^{-1}\begin{bmatrix}\frac{\partial f}{\partial A}\\ \frac{\partial f}{\partial B}\end{bmatrix}_{t} \qquad (53)$$

การประมาณค่าความแตกต่างสำหรับอนุพันธ์ลำดับสองของฟังก์ชัน f(A, B) ที่ $(A_0, B_0)$ อธิบายได้ดังนี้ (Nakamura. 1996 cited in Li, Lissitz and Yang. 1999 : 10)

$$\frac{\partial^2 f(A,B)}{\partial A^2} \approx \frac{f((A_0+\Delta A),B_0)-2f(A_0,B_0)+f((A_0-\Delta A),B_0)}{\Delta A^2} \qquad (54)$$

$$\frac{\partial^2 f(A,B)}{\partial B^2} \approx \frac{f(A_0,(B_0+\Delta B))-2f(A_0,B_0)+f(A_0,(B_0-\Delta B))}{\Delta B^2} \qquad (55)$$

$$\frac{\partial^2 f(A,B)}{\partial A \partial B} \approx \frac{f((A_0+\Delta A),(B_0+\Delta B)) - f((A_0-\Delta A),(B_0+\Delta B))}{\Delta A \Delta B} \\ \frac{-f((A_0+\Delta A),(B_0-\Delta B)) + f((A_0-\Delta A),(B_0-\Delta B))}{\Delta A \Delta B} \quad (56)$$

$$\frac{\partial^2 f(A,B)}{\partial B \partial A} \approx \frac{\partial^2 f(A,B)}{\Delta A \Delta B} \quad (57)$$

ค่าเริ่มต้นของ A และ B เป็นสิ่งสำคัญสำหรับการแก้สมการทำซ้ำ (64) ค่าเริ่มต้นสำหรับ A และ B คำนวณโดยสมการข้างล่างนี้ จะต้องเป็นค่าที่ทำให้มีการลู่เข้าเป็นส่วนใหญ่

$$A_{starValue} = \frac{\sum_{i=1}^{L} a_{iE}}{\sum_{i=1}^{L} a_{iB}} \quad (58)$$

$$B_{starValue} = \left[\frac{\sum_{i=1}^{L} b_{iB}}{L}\right] - A_{starValue}\left[\frac{\sum_{i=1}^{L} b_{iE}}{L}\right] \quad (59)$$

จากนั้นจะได้พารามิเตอร์การเทียบคะแนน ค่าประมาณความสามารถและข้อสอบจากแบบทดสอบต่างฉบับซึ่งแปลงไปสู่มาตรวัดเดียวกัน

**ขั้นตอนการปรับเทียบคะแนน**

การปรับเทียบคะแนนจะเป็นประโยชน์กับการจัดการศึกษา ต้องมีขั้นตอนดำเนินการดังต่อไปนี้ (Kolen and Brennan, 2004)

1. กำหนดจุดมุ่งหมายในการปรับเทียบคะแนนว่าจะนำผลจากการปรับเทียบคะแนนเพื่อดูพัฒนาการเปลี่ยนแปลงทางการศึกษา หรือเพื่อเทียบความสามารถผู้สอบเป็นข้อมูลใช้ตัดสินผลร่วมกัน ใช้แทนกันได้

2. สร้างแบบสอบหลายฉบับ แต่ละฉบับวัดเนื้อหาเดียวกัน และสร้างตามแบบแผนการออกข้อสอบเดียวกัน (Item Specification) เป็นแบบสอบที่มีลักษณะของความเป็นคู่ขนานในด้านเนื้อเรื่อง โครงสร้าง รูปแบบ ชนิดของข้อสอบ และเวลาที่ใช้สอบ

3. เลือกวิธีเก็บรวบรวมข้อมูล มีรูปแบบกลุ่มสุ่ม (Random Groups Design) โดยสุ่มกลุ่มตัวอย่างจากประชากรกลุ่มเดียวกัน แต่ละกลุ่มทำแบบสอบคนละชุด รูปแบบผู้สอบกลุ่มเดียว (Single-Group Design) ผู้สอบกลุ่มเดียวทำแบบสอบทั้งสองชุด รูปแบบกลุ่มเดียวที่ได้รับการจัดให้สมดุล (Single-Group Design with Counterbalancing) แบ่งผู้สอบแต่ละกลุ่ม แบบสอบแต่ละ

ชุดเป็น 2 ส่วน ให้ผู้สอบกลุ่มย่อยแรกสอบแบบสอบชุดที่ 1 ตอนแรก ตามด้วยแบบสอบชุดที่ 2 ตอนหลัง และผู้สอบกลุ่มย่อยที่ 2 ทำแบบสอบชุดที่ 2 ตอนแรก ตามด้วยแบบสอบชุดที่ 1 ตอนหลัง และรูปแบบผู้สอบกลุ่มไม่เท่าเทียมกันโดยใช้แบบสอบร่วม (Common-Item Nonequivalent Groups Design) ผู้สอบต่างกลุ่มประชากรทำแบบสอบคนละชุด และผู้สอบทุกคนทำแบบสอบร่วมที่อาจจะแทรกภายในแบบสอบทั้ง 2 ฉบับ (Internal Common-Item) หรือแยกออกจากแบบสอบ (External Common-Item)

4. เก็บรวบรวมข้อมูลตามรูปแบบที่กำหนดไว้

5. เลือกนิยามเชิงปฏิบัติการของการปรับเทียบคะแนน เพื่อตัดสินใจว่าจะใช้วิธีการปรับเทียบคะแนนเชิงเส้นตรง (Linear Equating Methods) หรือวิธีการปรับเทียบที่ไม่เป็นเชิงเส้นตรง (Nonlinear Equating Methods)

6. เลือกวิธีประมาณค่าสถิติที่ใช้วิเคราะห์ เลือกให้สอดคล้องกับนิยามเชิงปฏิบัติการที่กำหนด มีวิธีการปรับจากค่าเฉลี่ย (Mean Equating) โดยพิจารณาคะแนนสมมูลกันเมื่อคะแนนจากแบบสอบต่างฉบับเบี่ยงเบนไปจากคะแนนเฉลี่ยเท่ากัน วิธีการปรับเทียบเชิงเส้นตรง (Linear Equating) พิจารณาคะแนนสมมูลกันเมื่อคะแนนจากแบบสอบต่างฉบับมีคะแนนมาตรฐานเดียวกัน วิธีการปรับเทียบอีควิปเปอร์เซนไทล์ (Equipercentile Equating) ที่คะแนนสมมูลกันเมื่อคะแนนจากแบบสอบต่างฉบับมีตำแหน่งเปอร์เซนไทล์เท่ากัน และวิธีการปรับเทียบคะแนนโดยใช้สมการถดถอย (Regression Equating) เป็นการสร้างสมการทำนายคะแนนจากแบบสอบชุดหนึ่งไปยังอีกชุดหนึ่ง หรือได้คะแนนสมมูลกันเมื่อคะแนนของแบบสอบแต่ละฉบับทำนายคะแนนเกณฑ์ได้เท่ากัน ทั้ง 4 วิธีนี้ เป็นวิธีการปรับเทียบคะแนนตามแนวทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) ส่วนวิธีการปรับเทียบคะแนนตามแนวทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) เป็นการหาสัมประสิทธิ์การปรับเทียบ หรือค่าความชัน (Slope) และค่าคงที่ ของฟังก์ชันเชิงเส้นตรงที่เป็นความสัมพันธ์ของการปรับเทียบคะแนน วิธีการหาสัมประสิทธิ์การปรับเทียบ มีวิธีใช้ค่าเฉลี่ยของค่าอำนาจจำแนกและค่าเฉลี่ยของความยากของข้อสอบ (Mean and Mean Method) วิธีใช้ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าความยากของข้อสอบ (Mean and Sigma Method) วิธีการทำให้ความแตกต่างระหว่างคะแนนจริงเดิมกับคะแนนจริงที่ปรับแล้วมีค่าน้อยที่สุด โดยใช้สถิติ F- test (Characteristic Curve Method) และวิธีทำให้ความแตกต่างระหว่างพารามิเตอร์ข้อสอบเดิมกับที่ปรับแล้วมีค่าน้อยที่สุด โดยใช้สถิติ $\chi^2$ – test (Minimum $\chi^2$)

7. ประเมินผลการปรับเทียบคะแนน มีเกณฑ์ที่ใช้ดังนี้ (Harris and Crouse, 1993) ความเสมอภาค (Weak Equity) ของ Divgi และ Yen ที่พิจารณาจากความเท่าเทียมกันของการแจกแจงตามเงื่อนไขของคะแนนที่ได้จากแบบสอบต่างฉบับหลังจากปรับเทียบแล้ว ดัชนีสำหรับการเปลี่ยนแปลงคะแนน (Indices) ของ Angoff ความคลาดเคลื่อนมาตรฐาน (Standard Error) ของ Angoff เป็นการวิเคราะห์เพื่อประมาณความคลาดเคลื่อนของการปรับเทียบจากการสุ่มตัวอย่าง ข้อมูลที่จำลองขึ้น (Generated Data) ของ Lord เพื่อใช้สำหรับการปรับเทียบคะแนน การปรับเทียบคะแนนจากแบบสอบกลับสู่แบบสอบเดิม (Equating a Test to Itself) ของ Lord เป็นการปรับเทียบคะแนนกลับสู่แบบสอบเดิมโดยตรง หรือปรับผ่านแบบสอบอื่นก่อนปรับกลับสู่แบบสอบเดิม กลุ่มตัวอย่างขนาดใหญ่ (Large Sample) ของ Angoff เป็นการใช้กลุ่มตัวอย่างขนาดใหญ่ในการปรับเทียบคะแนนที่มีขนาดกลุ่มตัวอย่างน้อยกว่า ความคงเส้นคงวา (Consistency) เป็นการประเมินผลการปรับเทียบข้ามวิธี เพื่อหาความคงเส้นคงวา ความคงที่ (Stability) ของ Angoff เป็นการปรับเทียบคะแนนซ้ำ เพื่อตรวจความคงที่ของผลการปรับเทียบคะแนน

เงื่อนไขที่ส่งเสริมให้การปรับเทียบคะแนนเป็นที่น่าพอใจ มีรายละเอียด ดังนี้ (Kolen and Brennan, 2004)

1. ลักษณะทั่วไป

1.1 เป้าหมายของการที่นำคะแนนมาปรับเทียบมีการระบุอย่างชัดเจน

1.2 การออกแบบการเก็บรวบรวมข้อมูล การวางแผนการเชื่อมโยงการปรับเทียบคะแนน วิธีการทางสถิติที่ใช้ และการเลือกผลที่ได้จากการปรับเทียบหลายวิธี มีความเหมาะสมกับความสำเร็จของการปรับเทียบคะแนน

1.3 มีการติดตามและควบคุมคุณภาพของกระบวนการปรับเทียบคะแนนอย่างเพียงพอ

2. การพัฒนาแบบสอบสำหรับการเก็บรวบรวมข้อมูลทุกรูปแบบ

2.1 กำหนดเนื้อหาแบบสอบและกำหนดแบบแผนทางสถิติข้อสอบเป็นอย่างดี มีความคงที่

2.2 การสร้างแบบสอบเพื่อนำมาปรับเทียบพิจารณาจากสถิติข้อสอบ เช่น ค่าความยาก ค่าอำนาจจำแนก จากการสอบครั้งก่อน

3. การพัฒนาแบบสอบสำหรับรูปแบบผู้สอบกลุ่มไม่เท่าเทียมกันใช้แบบสอบร่วม

3.1 ข้อสอบร่วมต้องเป็นตัวแทนของแบบสอบที่ใช้ในการปรับเทียบ ทั้งคุณลักษณะของเนื้อหาและค่าสถิติของข้อสอบ

3.2 ข้อสอบร่วมมีจำนวนพอเหมาะ อย่างน้อย 20% ของแบบสอบที่นำมาปรับเทียบที่มีจำนวน 40 ข้อ หรือมากกว่า และอย่างน้อย 30 ข้อ เมื่อแบบสอบที่นำมาปรับเทียบมีขนาดใหญ่

3.3 ข้อสอบร่วมแต่ละข้อสร้างให้มีลักษณะที่ใกล้เคียงกับแบบสอบที่นำมาปรับเทียบทั้งคำถามและตัวเลือก

4. กลุ่มตัวอย่าง

4.1 กลุ่มตัวอย่างมีลักษณะเป็นตัวแทน

4.2 กลุ่มตัวอย่างมีความคงที่

4.3 กลุ่มตัวอย่างมีขนาดใหญ่เพียงพอ

4.4 กลุ่มตัวอย่าง 2 กลุ่ม สำหรับกลุ่มผู้สอบที่ไม่เท่าเทียมกันใช้แบบสอบร่วม มีความสามารถไม่แตกต่างกันมากนัก

5. การบริหารการสอบ

5.1 ดำเนินการทดสอบอย่างปลอดภัย

5.2 การสอบในแต่ละครั้งดำเนินการอย่างระมัดระวัง และจัดสภาพการสอบให้เหมือนกัน

6. หลักสูตร เนื้อหาการปฏิบัติ หรือสาระที่นำมาใช้ในการปรับเทียบคะแนนมีความคงที่

- ผลของการปรับเทียบคะแนนมีความเชื่อถือได้ หรือมีความถูกต้องแม่นยำขึ้นอยู่กับการเลือกสถิติที่ใช้ในการปรับคะแนนให้มีความเหมาะสมกับแบบแผนการออกข้อสอบ การเก็บรวบรวมข้อมูล และการดูแลการสอบที่เป็นไปอย่างมีมาตรฐาน

**การประเมินคุณภาพของการปรับเทียบคะแนน**

ในการประเมินคุณภาพของการปรับเทียบคะแนนระหว่างแบบสอบ นิยมศึกษาสมการปรับเทียบจากกลุ่มตัวอย่างเทียบมาตราและนำสมการปรับเทียบที่ได้ไปใช้ตรวจสอบคุณภาพกับกลุ่มตัวอย่างสอบทานผล มีผู้เสอนวิธีการตรวจสอบคุณภาพของการปรับเทียบไว้ ดังนี้

1. **ความเสมอภาค** (“Weak Equity)

Lord (1980) ได้เสนอคุณลักษณะที่ดีของการปรับเทียบคะแนนด้านความเสมอภาค (Equity) คือ เมื่อทุกกลุ่มผู้สอบมีความสามารถเดียวกัน การแจกแจงคะแนนจากแบบสอบชุดที่ 1 ที่ปรับสเกลเข้าสู่แบบสอบชุดที่ 2 จะมีการแจกแจงคะแนนเหมือนกับแบบสอบชุดที่ 2 Divgi (1981) และ Yen (1983) ได้นำคุณลักษณะความเสมอภาคของลอร์ดไปประยุกต์เป็นเกณฑ์ โดยกำหนดคุณภาพที่ดีของการปรับเทียบคะแนนนั้นพิจารณาจากความเท่ากันของการแจกแจงตาม

เงื่อนไขของคะแนนของแบบสอบแต่ละฉบับหลังจากการปรับเทียบแล้ว Hanson (1991b) ได้เสนอเกณฑ์ในลักษณะเดียวกันนี้ ใช้ชื่อว่าคะแนนความคาดหวังที่เท่าเทียมกัน (Equivalent Expected Score) โดยมีข้อตกลงเบื้องต้นว่าคะแนนจริงของแบบสอบชุดเดิมที่ไม่ได้ปรับ (X) และคะแนนจริงของแบบสอบชุดใหม่ที่ต้องการปรับไปสู่แบบสอบชุด X มีความสัมพันธ์กันเชิงฟังก์ชัน เกณฑ์จะพิจารณาคะแนนจากแบบสอบชุด X ที่ปรับคะแนนมาจากแบบสอบชุด Y จะมีความเท่าเทียมกันกับคะแนนจากแบบสอบ X แสดงได้ดังสมการ

$$E\left[e(Y)\middle|\gamma(\tau)\right] = E(X|\tau) = \tau$$

เมื่อ $\tau$ แทน คะแนนจริงจากแบบสอบเดิมชุด X

$E\left[e(Y)\middle|\gamma(\tau)\right]$ แทน คะแนนจากแบบสอบใหม่ (Y) ที่ปรับเทียบแล้ว

$E(X|\tau)$ แทน คะแนนจากแบบสอบชุดเดิม X

2. **ดัชนี** (Indices)

ใช้เพื่อเปรียบเทียบคะแนนแปลงที่ได้จากการเทียบคะแนน 2 วิธี ดัชนีนี้ใช้สรุปภาพรวมของความแตกต่าง ดัชนีเหล่านี้เดิมใช้ประเมินความคลาดเคลื่อนทางสถิติ แต่นำมาใช้กับการเทียบคะแนนอาจจะไม่ชัดเจนนักแต่สามารถประยุกต์ใช้ดัชนีเหล่านี้ได้โดยให้ค่าที่แท้จริง (true value) เป็นเกณฑ์ ดัชนีกลุ่มนี้จำแนกเป็นดัชนีถ่วงน้ำหนักและดัชนีไม่ถ่วงน้ำหนัก

2.1 ดัชนีถ่วงน้ำหนัก ด้วยความถี่ของคะแนนดิบของแบบสอบฉบับเป้าหมาย ประกอบด้วย

2.1.1 $$RMSD = \left( \frac{\sum_i f_i (A_i - B_i)^2}{\sum_i f_i} \right)^{1/2}$$

2.1.2 $$MAD = \left( \frac{\sum_i f_i \left| A_i - B_i \right|}{\sum_i f_i} \right)$$

2.1.3 $$MSD = \frac{\sum_i f_i (A_i - B_i)}{\sum_i f_i}$$

เมื่อ $A_i$ คือ คะแนนสมมูล i จากแบบสอบฉบับใหม่ที่เทียบคะแนนไปยังสเกลของฉบับเก่าด้วยการเทียบคะแนนวิธีที่หนึ่ง

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

$B_i$ คือ คะแนนเกณฑ์ ซึ่งเป็นคะแนนที่เป็นจริง หรือคะแนนสมมูล i จากแบบสอบฉบับใหม่ที่เทียบคะแนนไปยังสเกลของฉบับเก่าด้วยการเทียบคะแนนวิธีอื่น

2.2 ดัชนีไม่ถ่วงน้ำหนัก เป็นการพิจารณาความแตกต่างที่เกิดขึ้นตลอดช่วงคะแนนของมาตรวัด โดยไม่กำหนดน้ำหนักในแต่ละระดับคะแนน ประกอบด้วย

2.2.1 $$RMSD = \left( \frac{\sum_i (A_i - B_i)^2}{k} \right)^{1/2}$$

2.2.2 $$MAD_u = \left( \frac{\sum_i |A_i - B_i|}{k} \right)$$

2.1.3 $$MSD_u = \frac{\sum_i (A_i - B_i)}{k}$$

โดยที่ k คือ จำนวนของระดับคะแนน (Score Point)

นอกจากนี้ ดัชนี RMSD ยังสามารถนำมาแยกเป็นความแปรปรวนของความแตกต่างและความลำเอียงยกกำลังสอง คือ

$$\sum_i f_i (d_i)^2 / n = \sum_i f_i (d_i - \bar{d})^2 / n + \bar{d}^2$$

เมื่อ $d_i$ คือ $(A_i - B_i)$

$\bar{d}$ คือ ค่าความแตกต่างเฉลี่ย

ความลำเอียง (bias) ในการเทียบคะแนน คือ ความคลาดเคลื่อนเชิงระบบ (systematic error)

การใช้ดัชนีเหล่านี้เป็นเกณฑ์ในการเทียบคะแนนความสามารถแปลความหมายตามขนาดดัชนี ซึ่งเป็นค่าที่เรียงลำดับได้ว่าดัชนีที่มีค่าน้อยแสดงว่าความแตกต่างนั้นมีน้อย แต่ดัชนีขนาดเท่าใดที่ทำให้ผลการเทียบคะแนนเป็นที่น่าพอใจยังไม่ได้มีการระบุไว้

Petersen และคณะ (1982) ได้เสนอดัชนีความแตกต่าง (Discrepancy Index) มีสูตรการคำนวณคือ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

$$\text{total error} = \sum_{i} f_j d_j^2 \Big/ nS_t^2$$

เมื่อ $d_j = (t_j - t_j')$

$t_j$ = คะแนนเกณฑ์ ซึ่งเป็นคะแนนแปลงจากการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ

$t_j'$ = คะแนนแปลงจากการเทียบคะแนนวิธีอื่น

$S_t^2$ = ความแปรปรวนของคะแนน $t_j$

ค่าดัชนีที่ได้มีลักษณะเป็นค่ามาตรฐาน เพราะถ่วงน้ำหนักด้วยความแปรปรวนของคะแนนเกณฑ์ สามารถนำมาเปรียบเทียบกันได้โดยตรง ถึงแม้ในสถานการณ์ที่ได้ข้อมูลต่างกันก็ตาม นอกจากนี้ปีเตอร์เซนและคณะยังได้กำหนดหลักเกณฑ์การประเมินความเพียงพอในการเทียบคะแนนซึ่งเป็นการประเมินประสิทธิภาพของการเทียบคะแนนตามระดับการยอมรับ 5 ระดับ โดยใช้ร้อยละของส่วนเบี่ยงเบนมาตรฐานของคะแนนเกณฑ์ยกกำลังสองเป็นเกณฑ์ในการเปรียบเทียบแต่ละระดับ

นอกจากนี้เจเกอร์ (Jaeger, 1981) ยังได้เสนอดัชนีที่ใช้ตรวจสอบความพอเพียงของกระบวนการเทียบคะแนนเชิงเส้นตรง เป็นการตรวจสอบความคล้ายคลึงกันของการแจกแจงคะแนนสะสม (Similarity of Cumulative Score Distributions) ระหว่างคะแนนแปลงกับคะแนนของแบบสอบฉบับเป้าหมายด้วยสถิติทดสอบ Kolmogorov – Sminov two sample test

**3. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน** (Standard error of equating)

ความคลาดเคลื่อนโดยทั่วไปจำแนกเป็นความคลาดเคลื่อนเชิงสุ่มและความคลาดเคลื่อนเชิงระบบ ความคลาดเคลื่อนเชิงสุ่มเป็นความคลาดเคลื่อนที่เกิดจากการใช้หน่วยตัวอย่างจากประชากรมาประมาณค่าความสัมพันธ์ของการเทียบคะแนน ซึ่งอาจเกิดความแตกต่างจากความสัมพันธ์ของการเทียบคะแนนในประชากร สำหรับความคลาดเคลื่อนเชิงระบบได้แก่ วิธีการประมาณค่าทำให้เกิดความลำเอียงในการประมาณค่าความสัมพันธ์ของการเทียบคะแนน เช่น เทคนิคการปรับเทียบ มีประโยชน์คือลดความคลาดเคลื่อนเชิงสุ่ม แต่อาจทำให้เกิดความคลาดเคลื่อนเชิงระบบคือ ขนาดของความคลาดเคลื่อนเชิงสุ่มจะลดลงเมื่อหน่วยตัวอย่างมีขนาดเพิ่มขึ้น แต่ไม่ทำให้ความคลาดเคลื่อนเชิงระบบลดลง ดังนั้นความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนก็คือ ค่าประมาณของความคลาดเคลื่อนอันเนื่องจากการสุ่มตัวอย่าง คือ ส่วนเบี่ยงเบนมาตรฐานของคะแนนแปลงที่เกิดจากการทำซ้ำในกระบวนการเทียบคะแนน

ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนเป็นเกณฑ์สำคัญในการเทียบคะแนน แต่ในการเทียบคะแนนบางวิธี เช่น การเทียบคะแนนสังเกตตามแนวทฤษฎีการตอบสนองข้อสอบยังไม่มีสูตรเบื้องต้นในการคำนวณ ทำให้การปรับเทียบคะแนนบางวิธีไม่สามารถคำนวณความคลาดเคลื่อนมาตรฐานได้ การใช้เกณฑ์ความคลาดเคลื่อนมาตรฐานยังคงใช้ขนาดความคลาดเคลื่อนที่เล็กกว่าแสดงถึงประสิทธิภาพที่ดีกว่า แต่ยังไม่สามารถระบุได้ว่าขนาดความคลาดเคลื่อนเท่าใดจึงจะยอมรับได้

4. **การสอบทานผล** (Replication Samples / Cross – Validation Studies Stability)

เป็นการนำผลที่ได้รับในกลุ่มตัวอย่างหนึ่งไปใช้กับกลุ่มตัวอย่างอื่น เป็นการตรวจสอบความคงที่ของการเทียบคะแนน ซึ่งเป็นเกณฑ์ภายนอกที่ใช้เพื่อเปรียบเทียบความแตกต่างของการเทียบคะแนนด้วยวิธีที่แตกต่างกัน คุณภาพของการวัดด้วยการสอบทานผลแสดงความคงที่ของการเทียบคะแนน กลุ่มตัวอย่างที่ใช้ตรวจสอบผลเรียกว่า กลุ่มสอบทานผล มีลักษณะเช่นเดียวกับกลุ่มตัวอย่างเทียบคะแนน กลุ่มสอบทานผลจะทำแบบสอบทั้งสองฉบับ แล้วนำคะแนนจากแบบสอบฉบับหนึ่งไปแปลงจากตารางเทียบคะแนนที่สร้างไว้ แล้วพิจารณาผลต่างระหว่างคะแนนที่ทำให้ได้คะแนนจริงกับคะแนนแปลงเป็นการตรวจสอบความคลาดเคลื่อนจากการเทียบคะแนน

โคเลนและวิทนีย์ (Kolen and Whitney, 1982) ได้ใช้เกณฑ์สอบทานผลในการเทียบความแตกต่างระหว่างวิธีการเทียบคะแนน 4 วิธี โดยนำคะแนนของกลุ่มสอบทานผลที่ทำแบบสอบทั้งสองฉบับ ไปตรวจสอบด้วยดัชนีความแตกต่างระหว่างการแจกแจงคะแนนแปลงกับคะแนนที่ทำได้จริงเรียกว่าดัชนีเปรียบเทียบเปอร์เซนไทล์ เป็นดัชนีความแตกต่างกำลังสองเฉลี่ยระหว่างคะแนนที่ได้กับคะแนนแปลง มีสูตรในการคำนวณดังนี้

$$C = \sum_{i} \left| \left( X_i - Y_i' \right) \right| nk$$

เมื่อ n คือ จำนวนคะแนนดิบของกลุ่มสอบทานผล

K คือ จำนวนข้อสอบในแบบสอบร่วมที่ใช้

ถ้า C มีค่าน้อย หมายความว่า วิธีการเทียบคะแนนที่ใช้สร้างคะแนนแปลงนั้น มีความเหมาะสมและเพียงพอที่ทำให้ผลการแปลงคะแนนมีความคงเส้นคงวา ดัชนีนี้เหมาะสมที่จะนำไปใช้กับกลุ่มสมมูลแบบสุ่ม 2 กลุ่ม ที่ทำแบบสอบคนละฉบับหรือกลุ่มเดี่ยวที่ทำแบบสอบทั้งสองฉบับ ซึ่งโคเลนและวิทนีย์ได้กล่าวถึงเกณฑ์การสอบทานผลว่า เป็นเกณฑ์ภายนอกที่เป็นองค์ประกอบสำคัญของการเปรียบเทียบความแตกต่างระหว่างการเทียบคะแนนด้วยวิธีการที่แตกต่างกัน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

**การประเมินคุณภาพการปรับเทียบคะแนนสำหรับการวิจัยครั้งนี้**

ในแต่ละเงื่อนไขของการจำลองข้อมูล จะมีค่าประมาณพารามิเตอร์ข้อสอบ 100 ชุด สำหรับแบบสอบแต่ละฟอร์ม ทั้งเก่าและใหม่ เนื่องจากมีการทำซ้ำ 100 ครั้งของการจัดกระทำข้อมูล เป็นสิ่งที่คาดหวังว่าแต่ละเงื่อนไขการจำลองข้อมูลพารามิเตอร์ข้อสอบสำหรับกลุ่มใหม่ควรจะอยู่บนสเกลเดียวกันเหมือนกับว่าเป็นประชากรของค่าพารามิเตอร์จากกลุ่มเดียวกัน หลังจากการแปลงค่าประมาณพารามิเตอร์ของฟอร์มใหม่ เกณฑ์การประเมิน 2 กลุ่มที่ใช้เพื่อประเมินความแม่นยำของกระบวนการจำลองข้อมูลทั้ง 4 วิธี ในการเปรียบเทียบกับวิธีการปรับค่าพารามิเตอร์พร้อมกัน (concurrent calibration) ซึ่งอยู่ภายใต้สถานการณ์ในอุดมติที่โมเดล IRT เป็นไปตามข้อตกลงเบื้องต้นทุกประการ ซึ่งเกณฑ์การประเมินที่ใช้คือ เกณฑ์ค่าคงที่ของการแปลงสเกล (scale transformation constants : STC) และเกณฑ์โค้งคุณลักษณธการตอบ (category response curve : CRC) ซึ่งเกณฑ์ STC มีพื้นฐานมาจากความแตกต่างระหว่างค่าคงที่ของการแปลงคะแนนที่ประมาณค่าได้กับค่าคงที่ของการแปลงคะแนนจริง (estimated and true transformation constant) ดังนั้น เกณฑ์ STC จึงไม่ประยุกต์ไปสู่วิธีการปรับค่าพารามิเตอร์พร้อมกัน ส่วนเกณฑ์ CRC มีพื้นฐานมาจากความแตกต่างระหว่างโค้งคุณลักษณะที่ประมาณค่าได้ กับโค้งคุณลักษณะจริง (estimated and true category response curve) ซึ่งเกณฑ์ของ CRC มี 2 เกณฑ์ คือ เกณฑ์ Conditional CRC และ เกณฑ์ Overall CRC สำหรับเกณฑ์ทั้ง STC และ CRC มีเพียงข้อสอบของฟอร์มใหม่เท่านั้นที่ถูกนำมาคิด โดยประเมินประสิทธิภาพของการปรับเทียบคะแนนทั้ง 4 วิธี สามารถประเมินโดยการตรวจสอบค่าประมาณพารามิเตอร์ข้อสอบของฟอร์มใหม่วางอย่างเหมาะสมบนสเกลเก่าหรือไม่ อย่างไร ถ้าค่าประมาณพารามิเตอร์ข้อสอบของฟอร์มใหม่วางอย่างเหมาะสมบนสเกลเก่าแล้วดังนั้นก็จะเป็นฟังก์ชันการตอบข้อสอบของฟอร์มใหม่ด้วย

การพิจารณาเกณฑ์ STC สำหรับการจำลองข้อมูลแต่ละเงื่อนไข สุ่มกลุ่มตัวอย่างกลุ่มผู้สอบสำหับกลุ่มเก่าและกลุ่มใหม่ถูกเลือกตามลำดับจาก N(0,1) และ N(B,A) เมื่อ A และ B ถูกออกแบบเป็นระดับของกลุ่มไม่เท่าเทียม 9 ระดับระหว่างกลุ่มเก่าและกลุ่มใหม่ สังเกตว่าค่า A และ B เป็นค่าที่แสดงบนสเกลเก่า (old form scale) เริ่มต้นโดยสเกลของฟอร์มใหม่ใช้เมทริกซ์ 0,1 เพื่อขจัดสเกลที่ไม่สามารถหาค่าได้ มันคือสิ่งที่คาดหวังว่า A และ B ควรจะเป็น true slope และ intercept ของการแปลงค่าเชิงเส้น เพื่อเชื่อมโยง 2 ฟอร์ม ดังนั้นค่าประมาณของ A และ B, $\hat{A}$ และ $\hat{B}$ ในแต่ละครั้งของการทำซ้ำ สามารถประเมินค่าจริงโดยลำดับ สำหรับเกณฑ์ STC มี 2 ชนิดของสถิติที่คำนวณ อันแรกคือ ค่าเฉลี่ยทั้งหมดของการประมาณค่า $\hat{A}$ และ $\hat{B}$ ทั้ง100 ครั้ง ซึ่งเป็นส่วนที่คาดหวังว่าจะใกล้กับค่า A และ B อันที่สองคือ ค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง

(Mean Square Errors : MSEs) จากการประมาณค่า 100 ครั้ง ซึ่งเป็นความแตกต่างระหว่างค่าที่ถูกประมาณ และค่าจริงสำหรับแต่ละค่าของ A และ B คำนวณได้จากสูตร

$$\frac{1}{100}\sum_{r=1}^{100}(\mathrm{A}-\hat{\mathrm{A}}_{\mathrm{r}})^2 \qquad \text{และ} \quad \frac{1}{100}\sum_{r=1}^{100}(B-\hat{B}_r)^2$$

สถิติทั้ง 2 ตัวนี้ คำนวณสำหรับการปรับเทียบแต่ละวิธีภายใต้การจำลองข้อมูลแต่ละเงื่อนไข เนื่องจากวิธีการปรับค่าพารามิเตอร์พร้อมกันไม่จำเป็นต้องประมาณค่าความชันและจุดตัดของการแปลงคะแนนเชิงเส้น เพื่อพัฒนาเมทริกร่วม (common metric) ฉะนั้นสถิตินี้จึงใช้เพื่อการปรับเทียบทั้ง 4 วิธีเท่านั้น

ให้ $P_{jk}(\theta)$ เป็นฟังก์ชันลำดับขั้นการตอบที่คำนวณกับประชากรพารามิเตอร์ข้อสอบสำหรับลำดับขั้นที่ k ของข้อที่ j แสดงบนสเกล $\theta$ เก่า เมื่อให้ค่าของ $\theta$ ก็จะเป็นการพิจารณาการวัดความแตกต่างสำหรับข้อที่ j ด้วย $K_j$ ลำดับขั้น ไม่ว่าจะเป็น MC หรือ CR ก็ได้ จากการทำซ้ำ 100 ครั้ง ดังสมการ

$$\frac{1}{K_j}\sum_{k=1}^{K_j}\left[P_{ij}(\theta)-\hat{P}_{(jk|r)(\theta)}\right]^2$$

เมื่อ $\hat{P}_{(jk|r)(\theta)}$ ถูกคำนวณกับพารามิเตอร์ข้อสอบที่ถูกแปลงสเกลแล้วในลำดับขั้นที่ k ของข้อที่ j บนฟอร์มใหม่จากการทำซ้ำครั้งที่ r ในค่า $\overline{P}_{(jk|r)(\theta)}$ มีเครื่องหมาย | อ่านเหมือนกับสมการข้างบน แสดงค่าเฉลี่ยทั้งหมดของลำดับขั้น ของกำลังสองของความแตกต่างระหว่างฟังก์ชันลำดับขั้นการตอบที่คาดหวังกับฟังก์ชันลำดับขั้นการตอบค่าจริง เงื่อนไขบนค่า $\theta$ ค่าเฉลี่ยของการทำซ้ำทั้งหมด 100 ครั้งจากสมการข้างต้นสำหรับข้อที่ j แสดงดังสมการ

$$\frac{1}{\mathrm{K_j}}\sum_{\mathrm{k=1}}^{\mathrm{K_j}}\left\{\frac{1}{100}\sum_{\mathrm{r=1}}^{100}\left[\mathrm{P_{ij}}(\theta)-\hat{\mathrm{P}}_{(\mathrm{jk|r})(\theta)}\right]^2\right\}$$

สมการข้างบนเป็นค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน (MSE) บนเงื่อนไขของ $\theta$ ของการทำซ้ำทั้ง 100 ครั้ง ของค่าเฉลี่ยระหว่างโค้งลำดับขั้นการตอบที่ประมาณค่าได้กับโค้งลำดับขั้นการตอบค่าจริงสำหรับข้อที่ j

ดังนั้น เกณฑ์ Conditional CRC เป็นค่าเฉลี่ยของข้อสอบจากฟอร์มใหม่ทั้งหมด ของค่า MSE จากสมการข้างบน

$$\frac{1}{n}\sum_{j=1}^{n}\frac{1}{K_j}\sum_{k=1}^{K_j}\left\{\frac{1}{100}\sum_{r=1}^{100}\left[P_{ij}(\theta)-\hat{P}_{(jk|r)}(\theta)\right]^2\right\}$$

เกณฑ์ Conditional CRC เป็นการเตรียมสารสนเทศเพื่อใช้ในการเปรียบเทียบกระบวนการปรับเทียบ (รวมถึงวิธีการปรับค่าพารามิเตอร์พร้อมกัน) ด้วยการดำเนินการเมื่อระบุค่าของ $\theta$ ซึ่งเป็นลักษณะของค่าต่อเนื่อง เพื่อเปรียบเทียบประสิทธิภาพของการปรับเทียบคะแนน รวมถึงวิธีการปรับค่าพารามิเตอร์พร้อมกันด้วย ค่าของเกณฑ์ Conditional CRC โดยกระบวนการเชื่อมโยงคะแนนเป็นกราฟของเงื่อนไขค่า $\theta$

การสรุปดัชนีทั้งหมดเกี่ยวกับความคลาดเคลื่อนในการเชื่อมโยงครบคลุมค่า $\theta$ ที่ต่อเนื่องทั้งหมด สามารถเป็นเครื่องมือที่เป็นประโยชน์สำหรับประเมินความถูกต้องของกระบวนการเชื่อมโยง เกณฑ์ Overall CRC จึงถูกพัฒนาตามมา ให้ค่า $g(\theta)$ เป็นฟังก์ชันการแจกแจงของค่า $\theta$ ซึ่งเป็นส่วนที่เป็นความหนาแน่น (density) ของการแจกแจงปกติมาตรฐานในการจำลองข้อมูลเงื่อนไขแรก ค่าคาดหวังของค่าเฉลี่ย MSEs ในสมการข้างบน กับความสัมพันธ์กับ $\theta$ สามารถแสดงได้ดังสมการ

$$\int\frac{1}{n}\sum_{j=1}^{n}\frac{1}{K_j}\sum_{k=1}^{K_j}\left\{\frac{1}{100}\sum_{r=1}^{100}\left[P_{ij}(\theta)-\hat{P}_{(jk|r)}(\theta)\right]^2\right\}g(\theta)d(\theta)$$

สมการนี้เป็นเกณฑ์ Over all CRC เป็นจุดที่สำคัญ สามารถทบทวนได้ใน Hanson และ Beguin (2002)

เพื่อการคำนวณทางตัวเลขการอินทิเกรทในสมการข้างบน เป็น Gauss-Hermite quadrature หรือ Monte Carlo integration สามารถใช้ได้อย่างเหมาะสมโดยเริ่มจากการอินทิเกรทในสมการ แล้วรวมเข้าด้วยกันเป็นพื้นที่ปกติ (normal density) ซึ่งเป็นส่วนที่ทำให้รูปร่างปรับเรียบ การศึกษานี้ใช้ Gauss-Hermite quadrature เพื่อคำนวณการอินทิเกรทสำหรับกระบวนการนี้จะมี quadrature points and weights 100 ค่า ที่ใช้เพื่อคำนวณให้ได้ค่าความแม่นยำที่สูงที่สุด

**สำหรับการวิจัยครั้งนี้ ผู้วิจัยใช้ค่า** MSE **ที่ได้จากเกณฑ์** Overall CRC **ในการประเมินคุณภาพการปรับเทียบสำหรับแบบสอบรูปแบบผสม เพื่อให้สามารถเปรียบเทียบคุณภาพของการปรับเทียบได้ทั้งวิธีการปรับเทียบด้วยการปรับค่าพารามิเตอร์แยกกัน** (Separated Calibration) **และการปรับค่าพารามิเตอร์พร้อมกัน** (Concurrent Calibration)

**เพื่อให้สอดคล้องกับงานวิจัยของ** Hanson **และ** Beguin (2002), Kim (2004) **และ** Kim **และ** Lee (2006) **ที่ได้ศึกษาเปรียบเทียบคุณภาพการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบด้วยการปรับค่าพารามิเตอร์แยกกัน** (Separated Calibration) **และการปรับค่าพารามิเตอร์พร้อมกัน** (Concurrent Calibration)

## ตอนที่ 5 โปรแกรมคอมพิวเตอร์ที่ใช้ในการวิจัย

สำหรับการศึกษาครั้งนี้ เป็นการศึกษาโดยการจำลองข้อมูล เพื่อให้ได้รูปแบบการตอบที่เป็นไปตามโมเดลตามทฤษฎีการตอบสนองข้อสอบ ผู้วิจัยได้เสนอโปรแกรม WinGen (Windows Software that Generates IRT Model Parameters and Item Responses) ที่พัฒนาโดย Han และ Hambleton (2007) เมื่อได้รูปแบบการตอบจากโปรแกรมดังกล่าวแล้ว ผู้วิจัยใช้โปรแกรม MULTILOG ในการประมาณค่าพารามิเตอร์ และการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ส่วนการปรับเทียบด้วยวิธีโค้งคุณลักษณะผู้วิจัยใช้โปรแกรม IRT EQ ที่พัฒนาโดย Han (2007) รายละเอียดของโปรแกรมดังกล่าว ดังนี้

**โปรแกรม** WinGen

ทฤษฎีการตอบสนองข้อสอบ (Item response theory: IRT) ได้รับความนิยมในการทดสอบทางการศึกษาและจิตวิทยา เนื่องจากคุณสมบัติที่น่าสนใจ เช่น ความไม่แปรเปลี่ยนของพารามิเตอร์ข้อสอบและผู้สอบ (เมื่อโมเดล IRT สอดคล้องกับข้อมูล) และพารามิเตอร์ข้อสอบและผู้สอบยังรายงานบนสเกลร่วมกัน อย่างไรก็ตาม โมเดลส่วนใหญ่ยังอยู่บนพื้นฐานของข้อตกลงเบื้องต้นที่สำคัญ 2 ข้อ คือ ความเป็นเอกมิติ (Unidimentional) และรูปแบบทางคณิตศาสตร์ของฟังก์ชันโค้งคุณลักษณะข้อสอบ (Mathematical form of the item characteristic functions) ซึ่งการละเมิดข้อตกลงเบื้องต้นเหล่านี้จะทำให้เกิดผลในทางลบเป็นอย่างมากต่อกระบวนการวัด แนวทางหนึ่งในการประเมินผลกระทบของการละเมิดข้อตกลงเบื้องต้นทั้ง 2 ข้อนี้ จำเป็นต้องมีการศึกษาในหลาย ๆ ปัจจัย เช่น ศึกษาองค์ประกอบของตัวเลือกในโมเดล ขนาดของกลุ่มผู้สอบ ลักษณะการแจกแจงความสามารถและความยาวของแบบสอบ และปัจจัยอื่น ๆ การศึกษาเหล่านี้จะใช้แนวทางในการศึกษาจากการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล

ส่วนสำคัญของการศึกษาด้วยเทคนิคมอนติคาร์โล คือ การจำลองข้อมูลการตอบของผู้สอบเพื่อให้ได้ค่าพารามิเตอร์ที่แท้จริงตามโมเดลที่ต้องการ ซึ่งโปรแกรมคอมพิวเตอร์สำหรับจำลองข้อมูลมีการพัฒนามาตั้งแต่ปี 1970 แต่ส่วนใหญ่เป็นโปรแกรมที่พัฒนาด้วยระบบ

ปฏิบัติการ DOS (เช่น DATAGEN, Hambleton และ Rovinelli, 1973; GenIRV, Baker, 1989; RESGEN, Muraki, 1992) โปรแกรมเหล่านี้มีข้อจำกัดเนื่องจากปัญหาต่าง ๆ ที่เกิดขึ้นกับระบบปฏิบัติการ DOS นั่นคือ มีระบบปฏิบัติการที่ช้า (16-bit) ไม่สะดวกต่อการนำไปใช้ เข้ากันไม่ได้กับระบบปฏิบัติการ Windows-based OSs 32-bit และใช้ค่อนข้างยาก ปัจจุบันมีบางโปรแกรม เช่น WINIRT (Fang และ Johanson, 2005) และโปรแกรม PARDSIM (Yoes, 1997) ที่พัฒนาบนระบบปฏิบัติการ Windows-based OSs แต่ก็มีตัวเลือกในการจำลองข้อมูลไม่มากนัก ดังนั้น โปรแกรม WinGen ซึ่งพัฒนาโดย Kyung T. Han และ Ronald K. Hambleton (Center for Educational Assessment.University of Massachusetts Amherst) ที่ได้พัฒนาขึ้นมาใหม่นี้จึงสามารถใช้งานได้ง่ายและมีตัวเลือกในการจำลองข้อมูลให้เหมาะสมกับข้อมูลที่ได้จากวิธีการวัดทางจิตมากยิ่งขึ้น

**คุณสมบัติของโปรแกรม**

1. **โมเดลการตอบสนองข้อสอบ**

โปรแกรมสามารถจำลองข้อมูลตามโมเดล IRT ได้ทั้ง แบบเอกมิติ และแบบพหุมิติ ดังนี้

- โมเดลการตอบสนองข้อสอบแบบ 2 ค่า ชนิด 1, 2, และ 3 พารามิเตอร์
- โมเดลการตอบสนองข้อสอบแบบมากกว่าสองค่า ด้วยโมเดล Partial credit, Generalized partial credit, graded response, Rating scale, และ Nominal response
- โมเดลแบบ Non – parametric
- โมเดลพหุมิติแบบชดเชย (Multidimentional compensatory model)

โปรแกรม WinGen สามารถจำลองข้อมูลได้มากกว่า 1 โมเดลในชุดของข้อสอบ 1 ชุด เช่น ข้อสอบ 10 ข้อแรกเป็นแบบ 2 PL อีก 30 ข้อถัดไปเป็น 3 PL และ 10 ข้อสุดท้ายเป็น Graded response model การจำลองข้อมูลลักษณะนี้จะเป็นประโยชน์ในทางปฏิบัติมากขึ้น

2. **ลักษณะการแจกแจง**

โปรแกรมสามารถจำลองชุดของพารามิเตอร์ข้อสอบ และพารามิเตอร์ความสามารถของผู้สอบ โดยการจำลองพารามิเตอร์ให้มีลักษณะการแจกแจงได้หลายลักษณะเพื่อให้สอดคล้องกับสภาพข้อมูลที่เกิดขึ้นจริง โปรแกรมการจำลองข้อมูลส่วนใหญ่จะมีลักษณะการแจกแจงให้เลือกเพียงไม่กี่ลักษณะ ได้แก่ การแจกแจงแบบ Normal และ /หรือ Uniform แม้ว่าการแจกแจงทั้ง 2 แบบนี้อาจจะเป็นไปตามทฤษฎี หรือง่ายต่อการเข้าใจ แต่การแจกแจงลักษณะดังกล่าวนั้นอาจจะไม่สอดคล้องกับข้อมูลจริงที่เกิดขึ้น หรือไม่สอดคล้องกับลักษณะที่นักวิจัยต้องการศึกษา สำหรับโปรแกรม WinGen ในการจำลองข้อมูลพารามิเตอร์ความสามารถของผู้สอบ ผู้ใช้สามารถเลือก

การแจกแจงแบบ Normal, Uniform หรือ Beta (ในส่วนที่สามารถเลือกได้สำหรับการจำลองข้อมูลให้การแจกแจงมีความเบ้ จากที่เป็นบวกสูงสุด ไปหาเป็นลบสูงสุด) สำหรับพารามิเตอร์ข้อสอบผู้ใช้สามารถเลือกการแจกแจงแบบ Normal, Uniform, Beta หรือ Log-Normal ขึ้นอยู่กับสิ่งที่ผู้ใช้ต้องการศึกษาและ/ หรือ ให้มีความสอดคล้องกับข้อมูลจริงมากยิ่งขึ้น โปรแกรม WinGen ยังสามารถแสดงภาพกราฟในลักษณะที่แสดงฮิสโตแกรม ในการแจกแจงความสามารถของผู้สอบและแสดงลักษณะโค้งของพารามิเตอร์ข้อสอบและฟังก์ชันต่าง ๆ ที่ต้องการได้

3. **การใช้งาน**

โปรแกรม WinGen เป็นโปรแกรมที่เขียนขึ้นบนฐานของ Windows จะทำให้ง่ายต่อการใช้งาน ผู้ใช้สามารถจำลองข้อมูลการตอบโดยการคลิกที่ปุ่มปฏิบัติการเท่านั้น ซึ่งการจำลองข้อมูลมี 3 ขั้นตอนหลัก คือ 1) การจำลองข้อมูล หรือการอ่านข้อมูลที่เป็นพารามิเตอร์ความสามารถของผู้สอบ 2) การจำลองข้อมูลหรือการอ่านค่าข้อมูลพารามิเตอร์ข้อสอบ และ 3) การจำลองข้อมูลผลการตอบ ซึ่งทั้ง 3 ขั้นตอนสามารถทำแยกกันได้ ทำให้สามารถตรวจสอบความผิดพลาดของค่าพารามิเตอร์ในแต่ละขั้นตอนได้ง่ายยิ่งขึ้น นอกจากนี้โปรแกรมยังสามารถแสดงกราฟต่าง ๆ ได้ เช่น Item characteristic curves (ICC), Test characteristic curves (TCC), Item information curves (IFC), Test information curves (TFC), ทำให้ผู้ใช้มั่นใจได้ว่าพารามิเตอร์ที่จำลองขึ้นมาสอดคล้องกับเป้าหมายที่ต้องการศึกษาหรือไม่

4. **คุณสมบัติของคอมพิวเตอร์**

โปรแกรม WinGen พัฒนาจาก Microsoft. NET frameworks 2.0 ซึ่งโปรแกรมส่วนใหญ่จะพัฒนามาจากโปรแกรมนี้ โปรแกรม WinGen สามารถรันบน Windows series 32 bit (เช่น Windows XP) หรือ Windows series 64 bit (เช่น Windows vista) ก็ได้ ซึ่งก็เป็นสิ่งที่ดีที่โปรแกรมสามารถทำงานได้อย่างมีประสิทธิภาพคือ จัดการกับข้อมูลขนาดใหญ่ได้ภายในเวลาเพียงเล็กน้อย

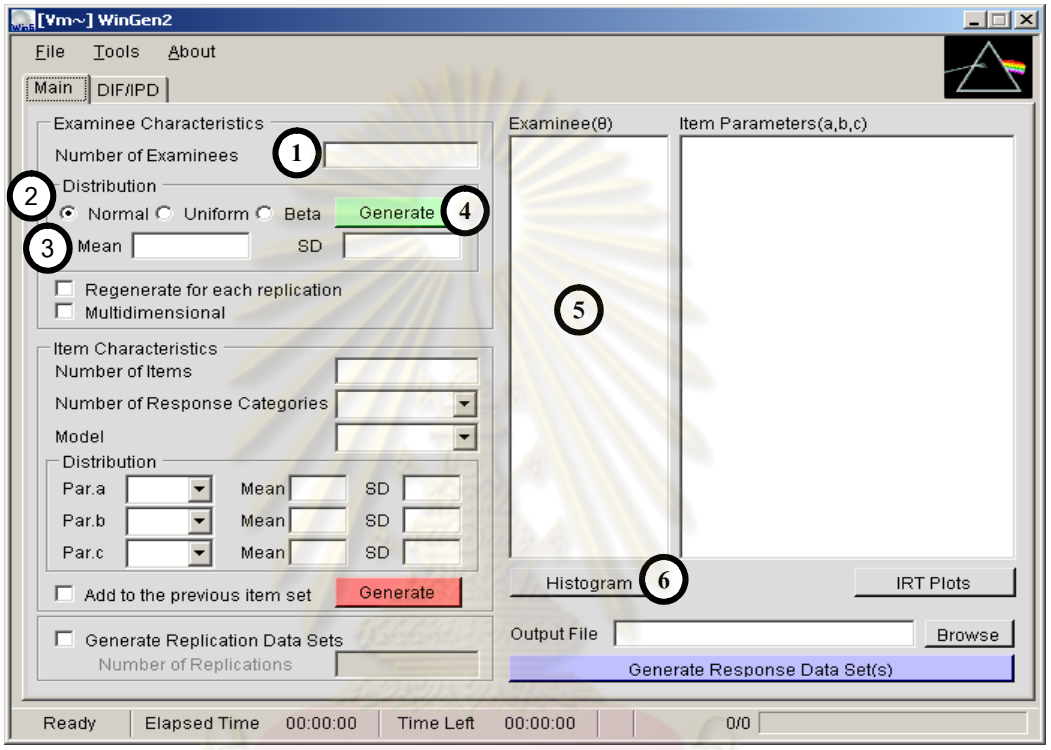
5. **การประยุกต์ข้อมูลไปใช้กับโปรแกรมอื่น ๆ**

โปรแกรม WinGen เป็นเครื่องมือที่ช่วยให้นักวิจัยบรรลุเป้าหมายได้หลาย ๆ เป้าหมาย โปรแกรมสามารถทำซ้ำได้ถึง 1 ล้านครั้ง และยังสามารถสร้าง syntax เพื่อไปใช้ในการวิเคราะห์ด้วยโปรแกรมอื่น ๆ ได้ เช่น PARSCALE (Muraki และ Bock, 2003), BILOG (Zimowski, Muraki, Mislevy, และ Bock, 2003) และ MULTILOG (Thissen, 2003) หรือผู้ใช้สามารถระบุตัวอย่าง syntax file แล้วให้โปรแกรมทำการจำลองข้อมูลตามลักษณะของ syntax ที่ต้องการได้อีกด้วย

**ขั้นตอนการใช้โปรแกรมในการจำลองข้อมูล**

ขั้นตอนการจำลองข้อมูลด้วยโปรแกรม Wingen ประกอบด้วย 3 ขั้นตอนหลัก ได้แก่ การจำลองพารามิเตอร์ผู้สอบ การจำลองพารามิเตอร์ข้อสอบ และการจำลองข้อมูลการตอบ

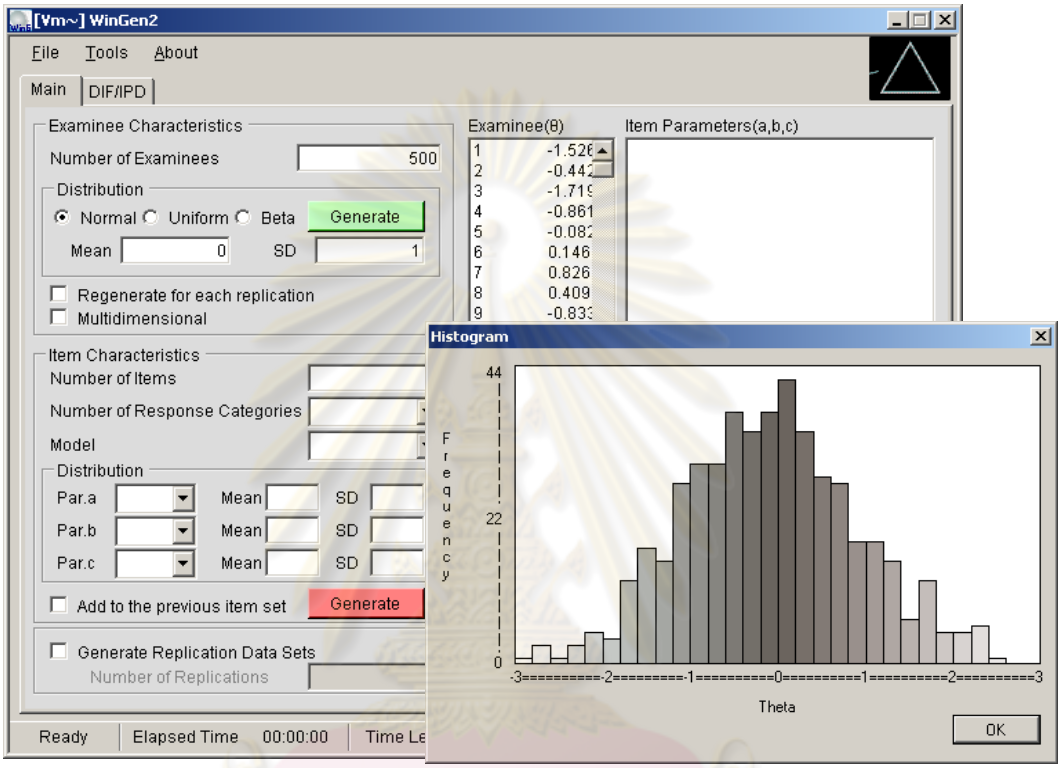
**การจำลองพารามิเตอร์ผู้สอบ**




**แผนภาพที่ 2** หน้าจอแสดงขั้นตอนการจำลองข้อมูลพารามิเตอร์ความสามารถของผู้สอบ

1. ระบุจำนวนผู้สอบ
2. ระบุลักษณะการแจงแจงความสามารถของผู้สอบ ซึ่งโปรแกรมมีการแจกแจงให้เลือก 3 แบบ คือ Normal, Uniform และ Beta
3. ระบุค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน เมื่อเลือกการแจกแจงเป็นแบบ Normal ระบุค่าสูงสุด และค่าต่ำสุดเมื่อเลือกการแจงแจงแบบ Uniform และระบุค่า a, b เมื่อเลือกการแจกแจงแบบ Beta
4. คลิกปุ่ม Generate เพื่อจำลองข้อมูลพารามิเตอร์ความสามารถของผู้สอบ
5. ค่าพารามิเตอร์ความสามารถของผู้สอบจะปรากฏในส่วน Examinee และสามารถบันทึก (save) พารามิเตอร์ความสามารถ (Examinee data) โดยคลิกที่ File เลือก save และเลือก Examinee data บันทึกลงใน Folder ที่ต้องการจัดเก็บ

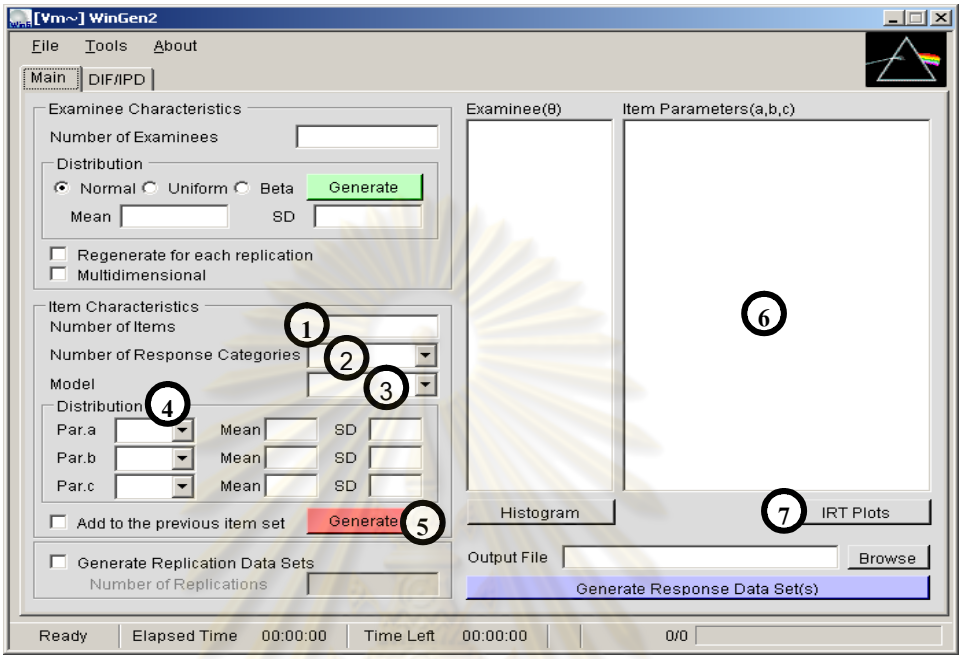
6. คลิกที่ปุ่ม Histogram เพื่อดูลักษณะการแจงแจงความสามารถของผู้สอบจากการจำลองข้อมูล




**แผนภาพที่** 3 ตัวอย่างการจำลองข้อมูลพารามิเตอร์ความสามารถของผู้สอบ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

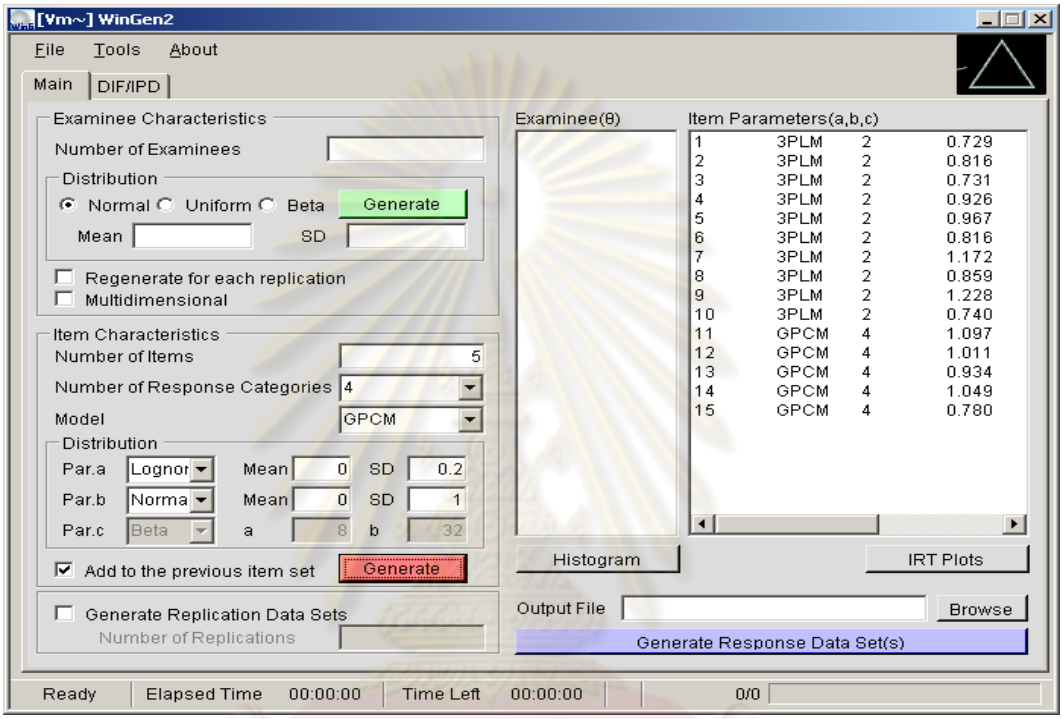
**การจำลองพารามิเตอร์ข้อสอบ**



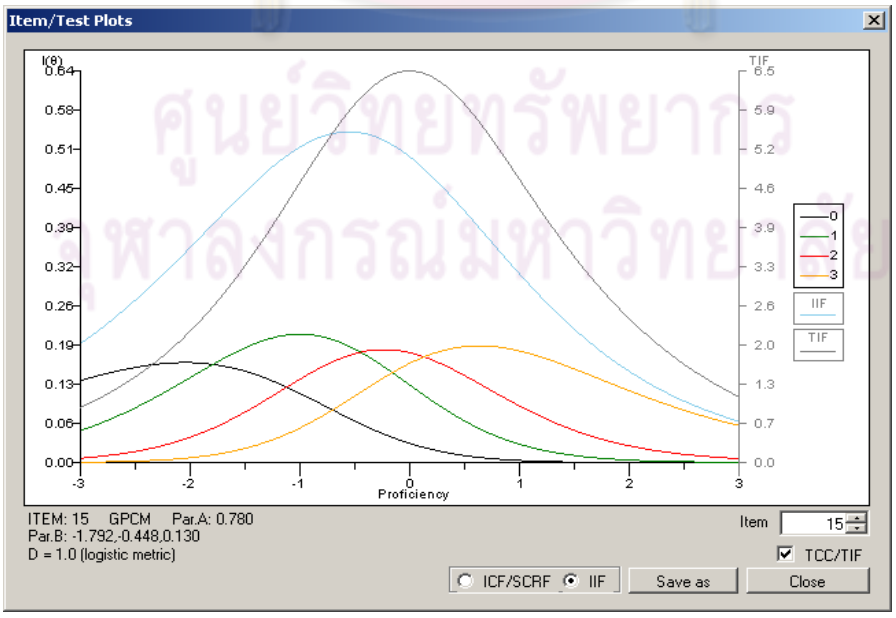

**แผนภาพที่** 4 หน้าจอแสดงขั้นตอนการจำลองข้อมูลพารามิเตอร์ความสามารถของข้อสอบ

1. ระบุจำนวนข้อสอบ
2. ระบุจำนวนรายการคำตอบ
3. เลือกโมเดลการตอบ ถ้าระบุรายการคำตอบในข้อ 2 เป็น 2 โมเดลการตอบจะมีให้เลือก 4 โมเดล คือ 1PL, 2PL, 3PL และ Non-Parametric ถ้าระบุรายการคำตอบในข้อ 2 มากกว่า 2 โมเดลการตอบจะมีให้เลือก 4 โมเดล คือ GRM, PRM, RSM,GPCM, และ NRM
4. เลือกลักษณะการแจกแจงค่าพารามิเตอร์ข้อสอบ
5. คลิกปุ่ม Generate
6. พารามิเตอร์ข้อสอบจะปรากฏในส่วนของ Item Parameter และสามารถบันทึก (save) พารามิเตอร์ข้อสอบ โดยคลิกที่ File เลือก save และเลือก Item Parameter Data บันทึกลงใน Folder ที่ต้องการจัดเก็บ
7. โปรแกรมสามารถแสดงภาพ ICC, TCC, IIF, TIF โดยการคลิกที่ปุ่ม IRT Plots

8. หากแบบสอบมีลักษณะโมเดลการตอบที่แตกต่างกันในแบบสอบฉบับเดียวกัน เมื่อจำลองพารามิเตอร์ข้อสอบในส่วนแรกเรียบร้อยแล้ว ให้ Mark ที่กล่อง “Add to the previous item set” และเลือกรายการคำตอบ และโมเดลการตอบรูปแบบอื่น ๆ ตามที่ต้องการจำนวนข้อสอบชุดใหม่ จะต่อท้ายจากข้อสอบชุดแรกที่ได้จำลองไว้แล้ว

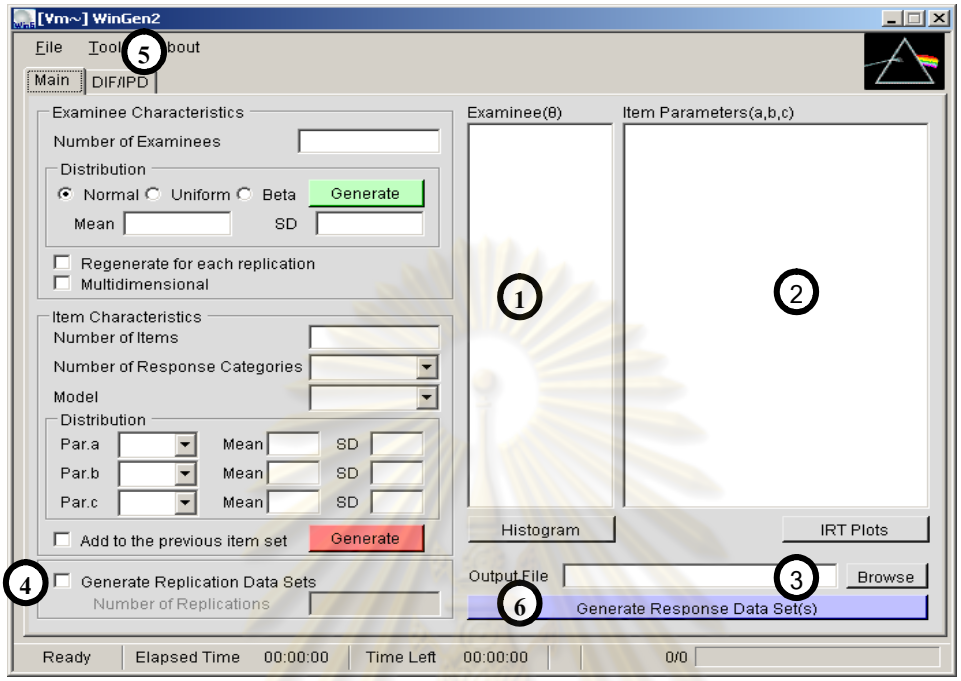



**แผนภาพที่** 5 ตัวอย่างการจำลองข้อมูลพารามิเตอร์ข้อสอบ



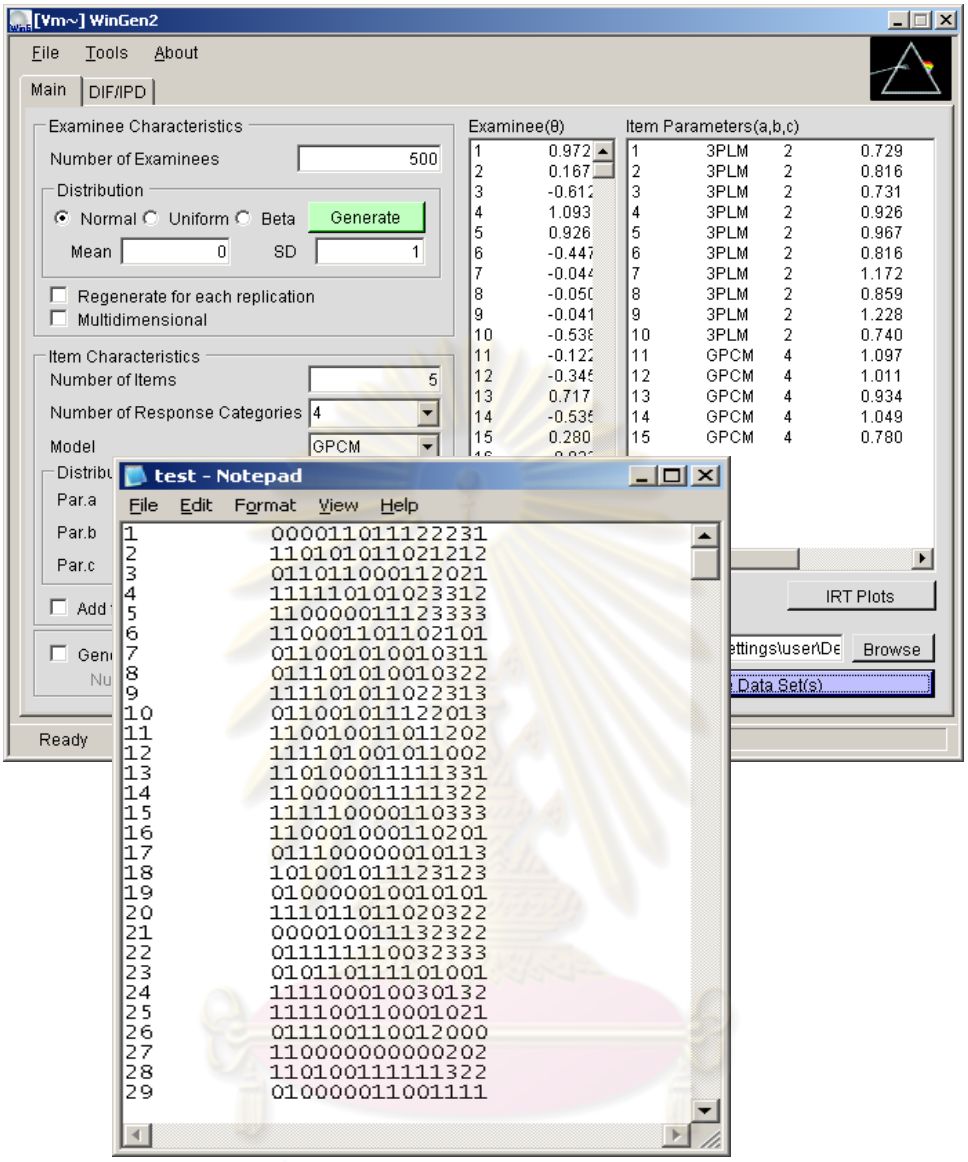

**แผนภาพที่** 6 ตัวอย่างกราฟแสดงโค้งคุณลักษณะ

**การจำลองข้อมูลการตอบ**




**แผนภาพที่** 7 หน้าจอแสดงขั้นตอนการจำลองข้อมูลผลการตอบข้อสอบ

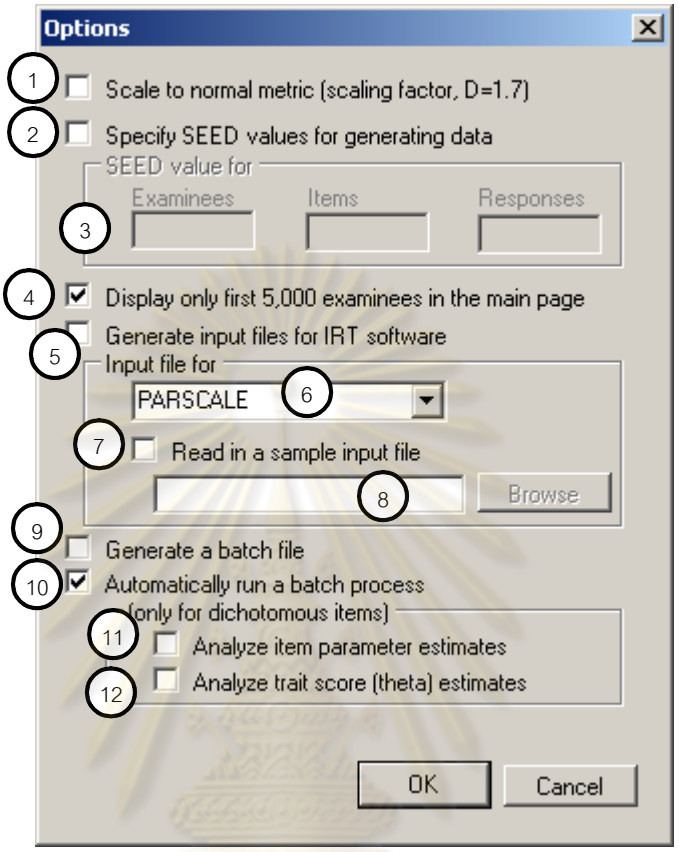
1. เปิดไฟล์ที่บันทึก Examinee Data
2. เปิดไฟล์ที่บันทึก Item Parameter Data
3. ระบุชื่อไฟล์ และโฟล์เดอร์ ที่ต้องการเก็บข้อมูลผลการตอบโดยคลิกที่ Browse
4. ถ้าต้องการทำซ้ำหลายครั้งให้ mark ที่กล่อง Generate Replication Data Sets
5. คลิกปุ่ม Option เมื่อต้องการสารสนเทศอื่น ๆ
6. คลิกปุ่ม Generate Response Data Set(s) โปรแกรมจะทำการจำลองข้อมูลการตอบตามจำนวนครั้งของการทำซ้ำที่ต้องการ ไฟล์ output จะเป็นไฟล์ที่มีนามสกุล *.wgr




**แผนภาพที่** 8 ตัวอย่างการจำลองข้อมูลผลการตอบข้อสอบ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

**เมนู** Option




**แผนภาพที** 9 เมนูอื่นๆ สำหรับการจำลองข้อมูล

1. ‘Scale to normal metric (scaling factor, D = 1.7)’ โปรแกรมกำหนดค่าตั้งต้นให้ค่า D = 1.7 ถ้าเลือกตัวเลือกนี้ โปรแกรมจะจำลองข้อมูลการตอบเป็น Normal - ogive model.
2. ‘Specify SEED values for generating data’ โปรแกรมกำหนด : ไม่เลือก โปรแกรมจะกำหนดค่า SEED ให้สอดคล้องกับเวลาในการประมวลผลโดยอัตโนมัติ
3. ถ้าเลือกตัวเลือกในข้อ 2 (SEED) ต้องระบุค่าของ SEED จำนวน 3 ค่า สำหรับประมวลผลพารามิเตอร์ผู้สอบ พารามิเตอร์ข้อสอบและข้อมูลการตอบข้อสอบตามลำดับ
4. ‘Display only the first 5,000 examinees in the main page’ โปรแกรมกำหนด : เลือก โปรแกรมจะแสดงผลการจำลองพารามิเตอร์ผู้สอบ 5,000 คนแรก ถ้าหากมี

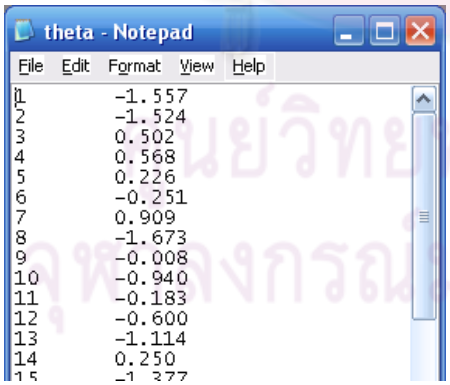
จำนวนคนมากกว่า 5,000 คน และผู้ใช้ไม่เลือกตัวเลือกนี้ จะทำให้โปรแกรมใช้เวลาในการประมวลผลนานขึ้น

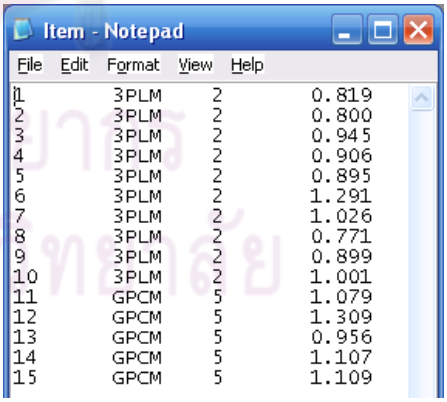
5. ‘Generate input files for estimating program’ โปรแกรมกำหนด : ไม่เลือก ตัวเลือกนี้โปรแกรม WinGen จะสร้าง syntax สำหรับวิเคราะห์ด้วยโปรแกรมอื่น ๆ โดยอัตโนมัติ เช่น โปแกรม PARSCALE, BILOG, MULTILOG เป็นต้น
6. ถ้าตัวเลือกในข้อ 5 ถูกเลือก โปรแกรมจะสร้าง syntax โดยอัตโนมัติ สำหรับโปรแกรม PARSCALE และ BILOG โปรแกรมจะสร้าง syntax บนพื้นฐานของโมเดลการตอบแบบ 2 ค่า แบบ 3 PL ส่วนโปรแกรม MULTILOG โปรแกรมจะสร้าง syntax บนพื้นฐานของโมเดลการตอบแบบ 2 ค่า แบบ 2 PL
7. ‘Read in a sample input file’ โปรแกรมกำหนด : ไม่เลือก
8. ถ้าตัวเลือกในข้อ 7 ถูกเลือก และอยู่นอกเหนือ syntax ที่สามารถสร้างได้ในข้อ 6 ผู้ใช้ต้องมีตัวอย่าง syntax ให้โปรแกรมได้สร้าง syntax ใหม่ตามที่ต้องการ เช่น syntax สำหรับประมาณค่าพารามิเตอร์ข้อสอบที่ตรวจให้คะแนนหลายค่า ด้วยโปรแกรม MULTILOG
9. ‘Generate batch file’ โปรแกรมกำหนด : ไม่เลือก ถ้าเลือกตัวเลือกนี้ โปรแกรมจะสร้าง batch file โดยอัตโนมัติ
10. ‘Automatically run a batch process’ โปรแกรมกำหนด : ไม่เลือก ถ้าเลือกตัวเลือกนี้ โปรแกรมจะรัน batch file โดยอัตโนมัติ เพื่อที่จะรันค่าประมาณ หลังจากที่จำลองข้อมูลแล้ว
11. ‘Analyze item parameter estimates’ โปรแกรมกำหนด : ไม่เลือก ถ้าเลือกตัวเลือกนี้ โปรแกรมจะนำส่งข้อมูลพารามิเตอร์ข้อสอบไปคำนวณค่าสหสัมพันธ์, RMSE (RMSD), MAD และ BIAS ข้อมูลที่ได้จะบันทึกในนามสกุล *.wgz
12. ‘Analyze trait (theta) estimates’ โปรแกรมกำหนด : ไม่เลือก ถ้าเลือกตัวเลือกนี้ โปรแกรมจะนำส่งข้อมูลคะแนนของผู้สอบไปคำนวณค่าสหสัมพันธ์, RMSE (RMSD), MAD และ BIAS ข้อมูลที่ได้จะบันทึกในนามสกุล *.wgz

โปรแกรม WinGen สามารถสร้างข้อมูล input และ output ได้หลายชนิด สรุปได้ดังนี้

| ไฟล์ | | ชนิด |
|---|---|---|
| * . wgc | WinGen cue file for executing sets of syntax files | Input |
| * . wgd | WinGen descriptive statistics for simulated data (examinee, item) | Output |
| * . wge | WinGen data file for examinees | Input and Output |
| * . wgi | WinGen data file for item parameters | Input and Output |
| * . wgr | WinGen data file for generated responses | Output |
| * . wgs | WinGen syntax file | Input |
| * . wgz | WinGen analysis summary file | Output |
| * . blm | Bilog – MG syntax file | Input and Output |
| * . psl | PARSCALE syntax file | Input and Output |
| * . mlg | MULTILOG syntax file | Input and Output |
| * . par | Item parameter data file for Bilog-MG, PARSCALE, or MULTILOG | Output |
| * . sco | Examinee score data file for Bilog-MG, PARSCALE, or MULTILOG | Output |

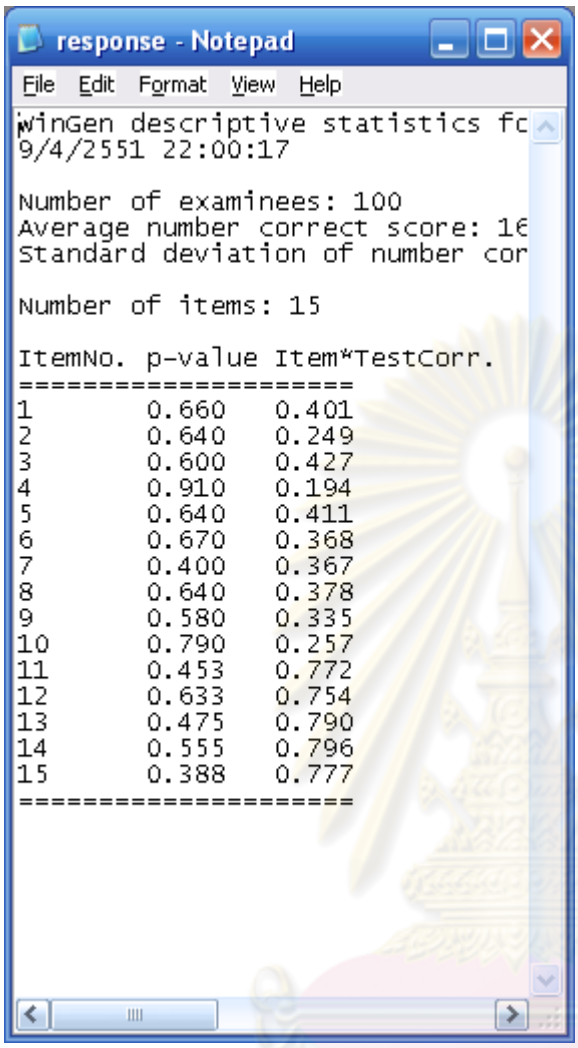
**ตัวอย่างไฟล์ Output ค่าพารามิเตอร์ความสามารถ และค่าพารามิเตอร์ข้อสอบ**




| ค่าพารามิเตอร์ความสามารถของผู้สอบ ชื่อไฟล์ “theta.**wge**” | ผลการจำลองข้อมูลการตอบ ชื่อไฟล์ “item.**wgi**” |
|---|---|

**แผนภาพที่** 10 แสดงตัวอย่างค่าพารามิเตอร์ความสามารถของผู้สอบและพารามิเตอร์ข้อสอบ



response - Notepad

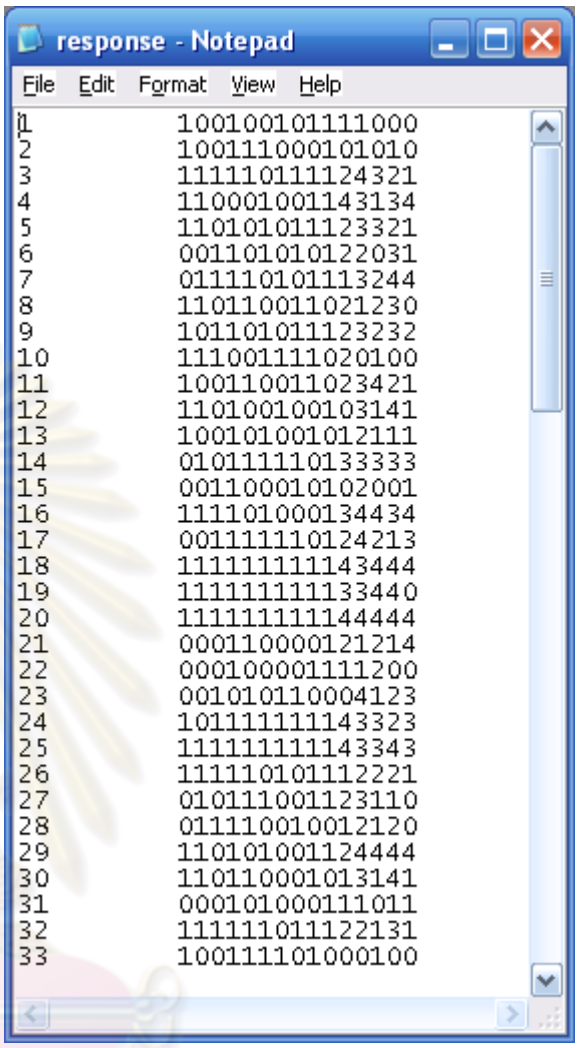
File Edit Format View Help

```
WinGen descriptive statistics fc
9/4/2551 22:00:17

Number of examinees: 100
Average number correct score: 16
Standard deviation of number cor

Number of items: 15

ItemNo. p-value Item*TestCorr.
=====================
1        0.660    0.401
2        0.640    0.249
3        0.600    0.427
4        0.910    0.194
5        0.640    0.411
6        0.670    0.368
7        0.400    0.367
8        0.640    0.378
9        0.580    0.335
10       0.790    0.257
11       0.453    0.772
12       0.633    0.754
13       0.475    0.790
14       0.555    0.796
15       0.388    0.777
=====================
```



response - Notepad

File Edit Format View Help

```
1          100100101111000
2          100111000101010
3          111110111124321
4          110001001143134
5          110101011123321
6          001101010122031
7          011110101113244
8          110110011021230
9          101101011123232
10         111001111020100
11         100110011023421
12         110100100103141
13         100101001012111
14         010111110133333
15         001100010102001
16         111101000134434
17         001111110124213
18         111111111143444
19         111111111133440
20         111111111144444
21         000110000121214
22         000100001111200
23         001010110004123
24         101111111143323
25         111111111143343
26         111110101112221
27         010111001123110
28         011110010012120
29         110101001124444
30         110110001013141
31         000101000111011
32         111111011122131
33         100111101000100
```

| สถิติบรรยายที่ได้จากการจำลองข้อมูล<br>ชื่อไฟล์ “response.**wgd**” | ผลการจำลองข้อมูลการตอบ<br>ชื่อไฟล์ “response.**wgr**” |
|---|---|

**แผนภาพที่** 11 ตัวอย่างไฟล์ Output ค่าสถิติบรรยาย และผลการตอบข้อสอบ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

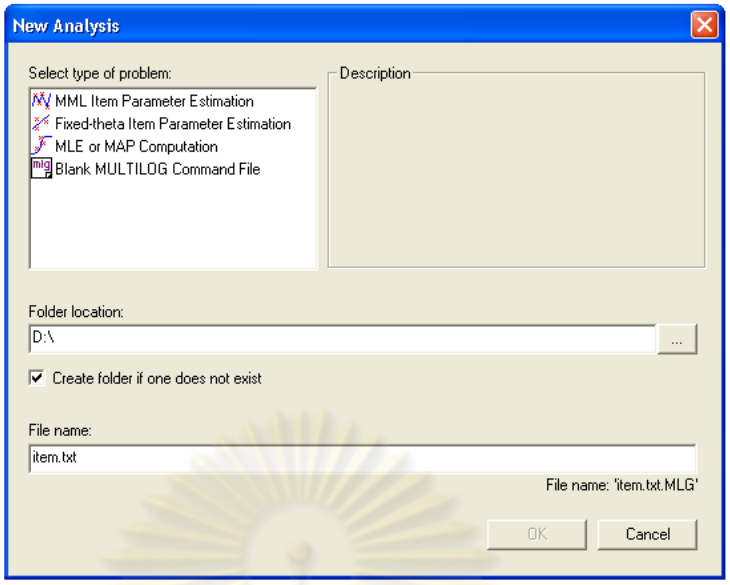
## การประมาณค่าพารามิเตอร์โดยใช้โปรแกรม MULTILOG

### การประมาณค่าพารามิเตอร์

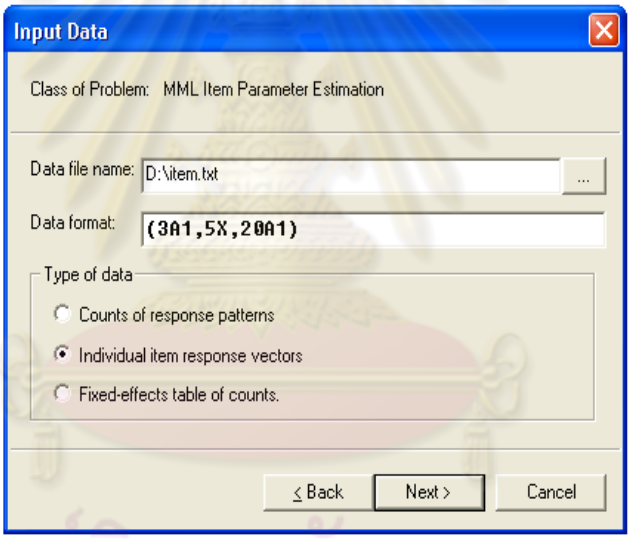
โปรแกรม MULTILOG เป็นโปรแกรมใช้สำหรับ วิเคราะห์ค่าพารามิเตอร์ของข้อสอบ และคะแนนความสามารถของผู้สอบ ภายใต้ทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นโปรแกรมที่ใช้วิเคราะห์ข้อสอบที่ตรวจให้คะแนนแบบหลายค่า ซึ่งโมเดลการตรวจให้คะแนนแบบหลายค่า ได้แก่ โมเดลเกรดเรสพอนด์ (Graded respons ) โมเดลนอมินอลเรสพอนส์ (Nominal response) และโมเดลหลายตัวเลือก (Multiple choice ite ) และสามารถวิเคราะห์ข้อสอบที่ตรวจให้คะแนนสองค่าในโมเดลโลจีสติกส์ได้อีกด้วย โปรแกรม MULTILOG จะประมาณค่ามิเตอร์ข้อสอบโดยใช้ Marginal Maximum Likellihood (MML) และใช้ Maximum Likellihood ( ML) ในการประมาณค่ากรณี Fixed – effect ค่าไคสแควร์แสดงถึงการ fit ของข้อมูลกับโมเดล นอกจากนี้การประมาณค่าพารามิเตอร์ความสามารถของผู้สอบใช้วิธี Maximum A Posteriori (MAP) ขั้นตอนการใช้ง่ายอย่างงาย มีดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ใช้คำสั่ง File – New เพื่อเริ่มต้นสร้างข้อมูลการวิเคราะห์คะแนน ซึ่งผู้ใช้ต้องทำการคลิกเลือกรูปแบบการวิเคราะห์ MML Item Parameter Estimation สำหรับวิเคราะห์ค่าพารามิเตอร์ข้อสอบ




ตั้งชื่อ File name ที่มีข้อมูลอยู่ ด้วยนามสกุล .txt แล้วกด OK



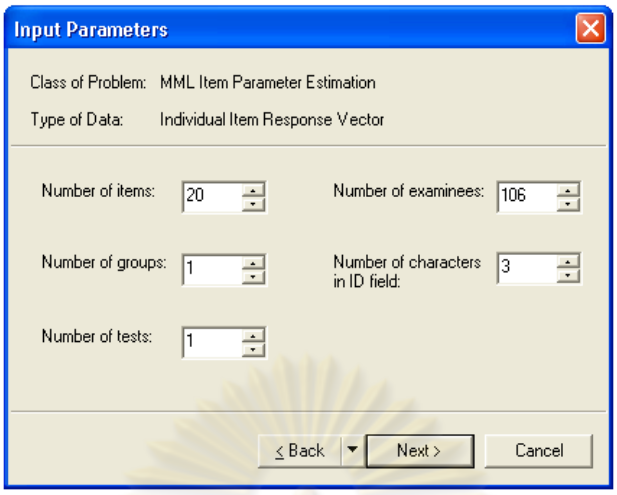

เขียนคำสั่งลงใน data format เช่นกรณี ที่ข้อสอบมี 35 ข้อ มีการให้คะแนนแบบ 2 ค่า ผู้สอบ 30 คน ก็จะเขียนเป็น ( 2A1,6x,35A1 )

2A1 คือ ตำแหน่งเลขที่ของผู้สอบที่มี 2 หลัก เช่น 30 คน

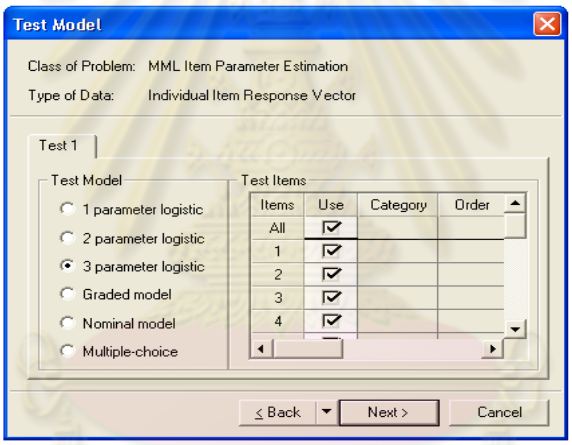
6x คือ การเว้นวรรคหรือระยะห่างของ ตำแหน่งเลขที่ของผู้สอบ กับ ผลการตอบรายข้อ

35A1 คือ จำนวนข้อของข้อสอบ

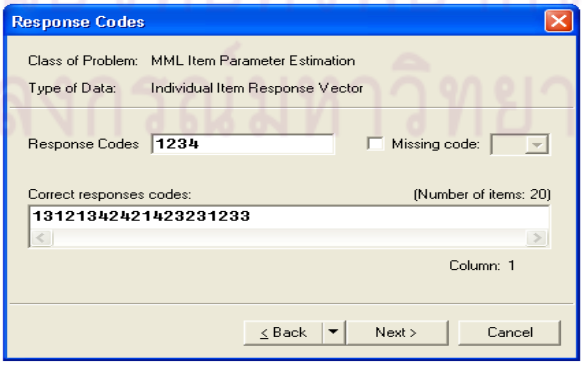
เลือก type of data คือ Individual item response vectors คือ การเลือกวิเคราะห์เป็นรายบุคคล จากนั้นกด Next



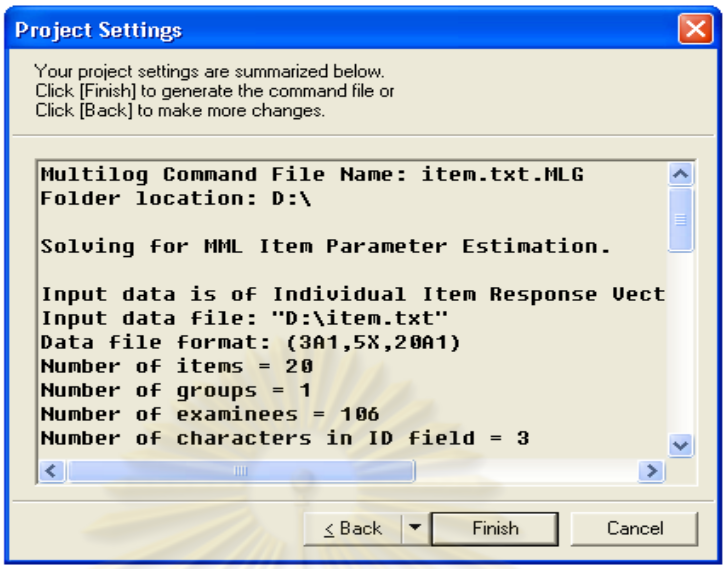

กำหนดรายละเอียดการวิเคราะห์ ใน Input Parameter จากนั้นกด Next



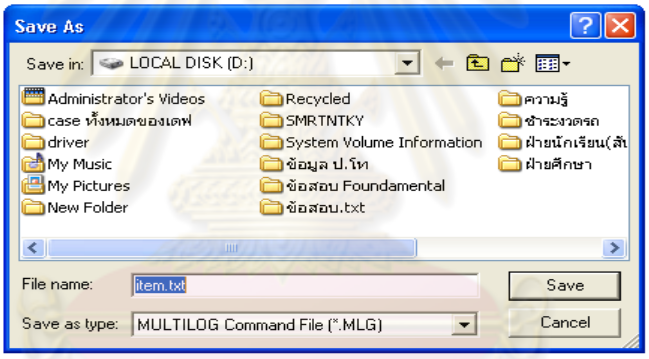

เลือกการวิเคราะห์ว่า จะวิเคราะห์กีพารามิเตอร์ ตามโมเดลที่ต้องการจากนั้นกด Next



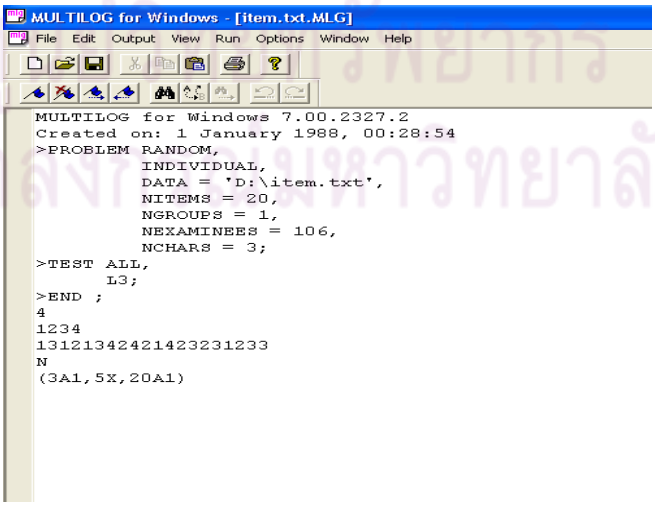

ใส่ Response Codes คือจำนวนตัวเลือกนั้นเอง จากนั้นก็ใส่ Correct Response Codes หมายถึงตัวเลือกที่ถูกในแต่ละข้อ จากนั้นกด Next



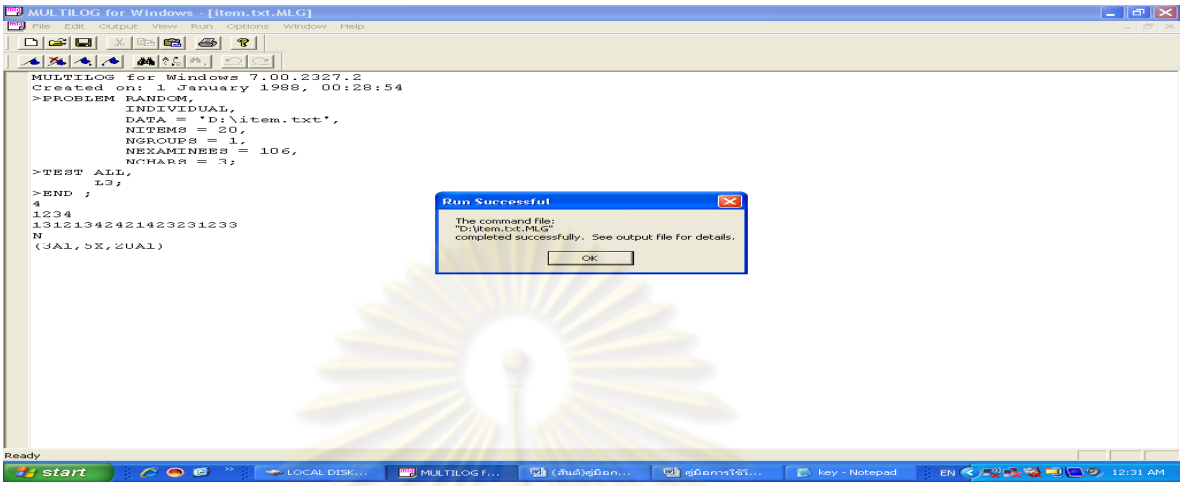

Project setting ที่เป็นรายละเอียดต่างๆที่เราได้เลือก และเสนอในรูปของ syntax จากนั้นกด finish




เลือกตำแหน่งที่จะ save File ของคำสั่ง ซึ่งจะถูกบันทึกเป็นนามสกุล *.mlg




คลิก RUN จะได้ Print out ออกมา




จะพบหน้าต่างนี้ กด OK จากนั้นคลิก Out put จะได้ Print out ที่ต้องการ

**การใช้โปรแกรม MULTILOG ในการปรับเทียบคะแนนด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน**

ในแต่ละเงื่อนไขของการจำลองข้อมูล ทั้งการปรับค่าพารามิเตอร์แยกกันเพื่อประมาณค่าพารามิเตอร์และปรับค่าพารามิเตอร์พร้อมกันเพื่อการปรับเทียบคะแนน ได้ดำเนินการโดยใช้โปรแกรม MULTILOG เพื่อให้ได้ค่าประมาณพารามิเตอร์ข้อสอบจากการจำลองข้อมูลการตอบสำหรับกลุ่มเก่าและกลุ่มใหม่ โปรแกรม MULTILOG ได้เตรียมการประมาณค่าแบบ (Marginal maximum likelihood : MML) เพื่อประมาณค่าพารามิเตอร์ข้อสอบสำหรับข้อมูลในส่วนที่เป็นตัวแปรแฝง ในการบ่งชี้ความสามารถที่เป็นอิทธิพลสุ่ม (Random effect) โปรแกรม MULTILOG สามารถใช้ได้เหมาะสมกับโมเดลหลายโมเดล และสามารถวิเคราะห์ค่าพารามิเตอร์ข้อสอบได้พร้อมกัน (simultaneous) ข้ามโมเดลการตอบได้

สำหรับการปรับค่าพารามิเตอร์พร้อมกัน (concurrent) และแยกกัน (separate) ในการวิเคราะห์ข้อสอบแบบเลือกตอบ (Multiple Choices : MC)ได้วิเคราะห์โดยใช้โมเดล 3PL และข้อสอบแบบเขียนตอบ (Constructed response : CR) วิเคราะห์โดยใช้โมเดล GPCM ในการวิเคราะห์ค่าพารามิเตอร์แต่ละโมเดลแยกกัน พารามิเตอร์ของข้อสอบ MC และ CR จะประมาณค่าแยกกัน สำหรับการดำเนินการวิเคราะห์พร้อมกัน ข้อสอบแบบ MC และ CR จะดำเนินการวิเคราะห์ด้วยกันภายในการ run โปรแกรมเพียง 1 ครั้ง สำหรับการดำเนินการวิเคราะห์ทั้งแยกกัน

และพร้อมกัน การแจกแจงของประชากรของกลุ่มเก่าและกลุ่มใหม่ถูกสันนิษฐานว่าเป็นการแจกแจงแบบปกติมาตรฐาน ภายใต้ข้อตกลงเบื้องต้นนี้ จะใช้การประมาณค่าแบบ MML เพื่อวิเคราะห์ข้อสอบ

สำหรับการปรับค่าพารามิเตอร์พร้อมกัน (concurrent) ชุดข้อมูล 2 ชุดจากกลุ่มเก่าและกลุ่มใหม่ ที่รวมไว้ในชุดข้อมูลชุดเดียว โดยระบุให้ข้อสอบที่อีกกลุ่มหนึ่งไม่ได้ตอบตอบเป็น “not reach” ในขณะที่ระบุสมาชิกในกลุ่มเป็น 1 หรือ 2 ในคอลัมภ์แรกของการรวมชุดข้อมูล ดังนั้นข้อสอบทั้งหมดจากทั้ง 2 ฉบับ จะวิเคราะห์โดยการรันโปรแกรม MULTILOG โดยการเขียนคำสั่งโดยใช้ syntax สำหรับการปรับเทียบกับกลุ่มผู้สอบหลายกลุ่ม (multiple groups form equating) เมื่อเตรียมคำสั่งเสร็จแล้ว เพื่อขจัดสเกลที่ไม่สามารถหาค่าได้ในโมเดล IRT การแจกแจงประชากรของกลุ่มเก่าและใหม่ จำเป็นต้องระบุอย่างเหมาะสม โปรแกรม MULTILOG จะค่าตั้งต้น (default) การประมาณค่าแบบ MML ซึ่งจะกำหนดว่าการแจกแจงของประชากรเป็นโค้งปกติอยู่แล้ว ถ้าวางให้กลุ่มเก่าเป็นกลุ่มสุดท้ายโปรแกรมก็จะกำหนดให้ค่าเฉลี่ยเป็น 0 สำหรับกลุ่มเก่า โดยที่มีส่วนเบี่ยงเบนมาตรฐานเป็น 1 ทั้ง 2 กลุ่ม การกำหนดแบบนี้สามารถใช้คำสั่ง FIX เพื่อสนับสนุนผู้ใช้ในการกำหนดค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน ซึ่งการกระทำเช่นนี้บอกเป็นนัยว่า พารามิเตอร์ข้อสอบ (และความสามารถ) ของกลุ่มใหม่ถูกประมาณค่าและวางบนสเกลของกลุ่มเก่าโดยปราศจากการระบุค่าเฉลี่ยของการแจกแจงของกลุ่มใหม่ แต่ค่าส่วนเบี่ยงเบนมาตรฐานต้องระบุโดยการใช้ค่าตั้งต้นของโปรแกรม หรือให้ผู้ใช้กำหนดเองก็ได้ การพิจารณาความต้องการเหล่านี้ใช้สำหรับการจำลองข้อมูลแต่ละเงื่อนไขในการวิเคราะห์ปรับเทียบแบบปรับค่าพารามิเตอร์พร้อมกัน ซึ่งส่วนเบี่ยงเบนมาตรฐานของการแจกแจงประชากรกลุ่มใหม่ สามารถใช้พารามิเตอร์จากการจำลองข้อมูลที่ปรากฏอยู่ได้

**โปรแกรม IRT EQ**

การเปรียบเทียบคะแนนจากแบบสอบที่ต่างฉบับกัน หรือจากการบริหารการทดสอบที่แตกต่างกันนั้นเป็นความท้าทายหลักของนักทดสอบการวัดทางจิต (Psychometricians) ปัจจุบันจุดเน้นที่สำคัญของโปรแกรมการทดสอบมีความก้าวหน้ามากขึ้น ความท้าทายหลักที่ยังคงอยู่คือการเปรียบเทียบคะแนนก็เป็นสิ่งที่สำคัญมากเช่นกัน ซึ่งมีการพัฒนาวิธีการปรับเทียบ/การแปลงสเกลบนพื้นฐานทฤษฎีการตอบสนองข้อสอบขึ้นจำนวนมากเพื่อแก้ไขปัญหาดังกล่าว (เช่น Lord, 1980; Hambleton, Swaminathan, and Roger, 1991; Kolen และ Brennan, 2004)

ในขณะที่มีงานวิจัยเกี่ยวกับการปรับเทียบเกิดขึ้นจำนวนมาก เนื่องจากมีความจำเป็นในทางเทคนิคการวัดผลเพื่อออกแบบและวิเคราะห์ข้อมูล แต่การพัฒนาทางด้านโปรแกรม

คอมพิวเตอร์ก็ยังมีข้อจำกัด เนื่องจากโปรแกรมที่พัฒนาขึ้นมาในอดีตมักจะมีข้อจำกัดหลายอย่าง เช่น โปรแกรมไม่สามารถประยุกต์ใช้กับโมเดลการตอบสนองข้อสอบ หรือไม่สอดคล้องกับรูปแบบการเก็บรวบรวมข้อมูล เป็นต้น ซึ่งความต้องการเหล่านี้จำเป็นต้องได้รับการพัฒนา ดังนั้น Han (2007) จึงได้พัฒนาโปรแกรม IRT EQ เพื่อใช้ในการปรับเทียบคะแนนตามความต้องการดังกล่าว รายละเอียดของโปรแกรม มีดังนี้

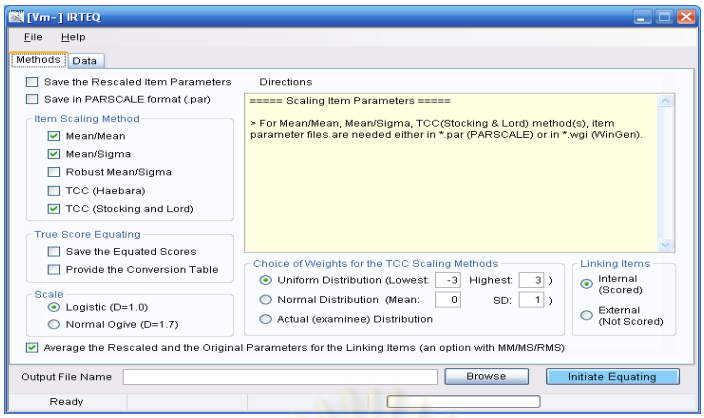
***การแปลงสเกลภายใต้การออกแบบการเก็บรวบรวมข้อมูลโดยใช้ข้อสอบร่วม***

โปรแกรม IRT EQ สามารถขยายแนวคิดไปสู่การปรับเทียบคะแนนสำหรับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกันโดยใช้ข้อสอบร่วมได้ โดยมีวิธีการปรับเทียบตามทฤษฎีการตอบสนองข้อสอบหลายวิธี ได้แก่ mean/mean (Loyd และ Hoover, 1980), mean/sigma (Marco, 1977), Roburt mean/sigma (Linn, Levine, Hastings, และ Wardrop, 1981), และวิธีการโค้งคุณลักษณะ ได้แก่ Heabara (1980) และ Stocking-Lord (1983) โปรแกรม IRT EQ ที่พัฒนาขึ้นมานี้สามารถประยุกต์ใช้ได้หลายแนวทาง เช่น วิธีการปรับเทียบแบบโค้งคุณลักษณะ ได้มีตัวเลือกให้ผู้ใช้เลือกการแจกแจงได้หลายรูปแบบ เป็นต้น

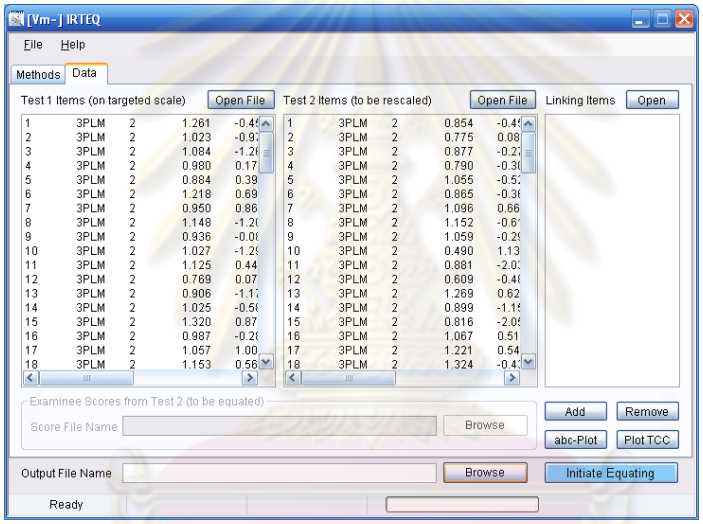
***โมเดลที่สามารถประยุกต์ใช้กับโปรแกรม IRTEQ***

โปรแกรม IRTEQ สามารถใช้ได้กับโมเดลการตอบข้อสอบตามทฤษฎี IRT ที่มีความเป็นเอกมิติ ได้แก่ โมเดลที่มีการตรวจให้คะแนนแบบสอบค่า โลจีสติกแบบ 1, 2 และ 3 พารามิเตอร์ โมเดลที่มีการตรวจให้คะแนนแบบหลายค่า ได้แก่ GPCM, PCM, GRM หรือการผสมโมเดลการตอบระหว่างโมเดลที่ตรวจให้คะแนนแบบสองค่า และโมเดลที่ตรวจให้คะแนนแบบหลายค่าภายใต้รูปแบบของโมเดลเหล่านี้ และในทางปฏิบัติไม่มีข้อจำกัดในเรื่องของจำนวนข้อสอบในแบบสอบแต่ละฉบับ และจำนวนข้อสอบร่วมด้วย ซึ่งโปรแกรม IRT EQ สามารถนำเข้าข้อมูลพารามิเตอร์ข้อสอบ และคะแนนที่มาจากโปรแกรม Wingen (Han, 2007) และ/หรือโปรแกรม PARSCALE (Muraki และ Bock, 2003) ได้โดยตรง

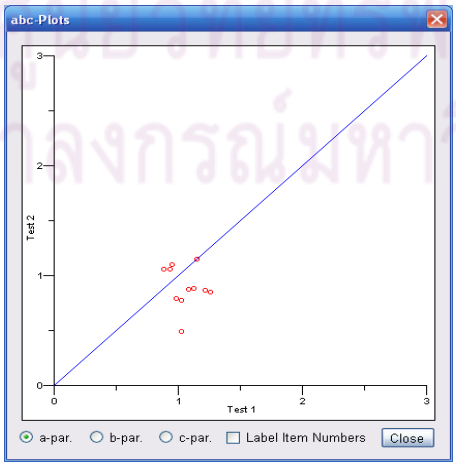
โปรแกรม IRTEQ เป็นโปรแกรมที่ง่ายต่อการใช้งาน สามารถแสดงกราฟคุณลักษณะของแบบสอบแต่ละฉบับและโค้งคุณลักษณะของสเกลที่มีการปรับเทียบแล้ว ลักษณะของโปรแกรมแสดงดังภาพต่อไปนี้



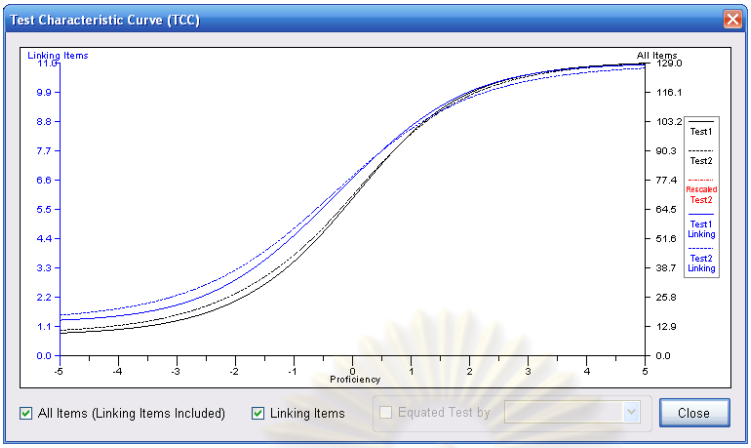

**แผนภาพที่** 12 หน้าจอหลักของโปรแกรม IRTEQ




**แผนภาพที่** 13 แสดงการนำข้อมูลเข้าสู่โปรแกรม




**แผนภาพที่** 14 แสดงการกระจายของพารามิเตอร์ข้อสอบที่เป็นข้อสอบร่วม




**แผนภาพที** 15 โค้งคุณลักษณะของแบบสอบสำหรับแบบสอบแต่ละฉบับ

## ตอนที 6 งานวิจัยที่เกี่ยวข้องกับการปรับเทียบคะแนน

เนื่องจากการทดสอบที่ผ่านมาส่วนใหญ่จะใช้แบบสอบที่มีรูปแบบข้อคำถามเพียงรูปแบบเดียว ในการวัดคุณลักษณะด้านใดด้านหนึ่งของนักเรียน การทดสอบลักษณะดังกล่าวอาจทำให้การประมาณค่าความสามารถของผู้เรียนมีความคลาดเคลื่อนเกิดขึ้น เนื่องจากการวัดคุณลักษณะบางอย่างไม่สามารถที่จะวัดได้จากการตอบข้อสอบเพียงรูปแบบเดียว และในปัจจุบันได้มีการใช้แบบสอบรูปแบบผสม(mixed–format test) ซึ่งเป็นการผสมข้อคำถามที่มีรูปแบบการตรวจให้คะแนนต่างกันอยู่ในแบบสอบฉบับเดียวกัน ส่วนใหญ่แล้วจะใช้ทำสอบวัดทักษะด้านต่าง ๆ ของนักเรียน โดยรูปแบบของข้อสอบส่วนใหญ่จะจำแนกเป็น 2 รูปแบบคือ ข้อสอบแบบหลายตัวเลือก (multiple choice: MC) และข้อสอบแบบเขียนคำตอบ (constructed response: CR) ซึ่งข้อสอบแบบ MC มีการตรวจให้คะแนนแบบสองค่า (dichotomous) และข้อสอบแบบ CR มีการตรวจให้คะแนนแบบหลายค่า (polytomous)

การปรับเทียบในแบบสอบรูปแบบผสมได้มีการศึกษาไม่มากนัก (Kolen และ Brennan, 2004; Kim และ Lee, 2006) โดยที่วิธีการศึกษาเป็นวิธีการปรับเทียบทั้งตามทฤษฎีการตอบสนองข้อสอบและตามทฤษฎีการทดสอบแบบดั้งเดิม สำหรับวิธีการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ ได้แก่ วิธีโค้งลักษณะแบบสอบ และวิธีไคสแควร์ต่ำสุด ซึ่งขยายมาจากวิธีการที่ใช้กับข้อมูลที่มีการตรวจให้คะแนนแบบสองค่า นอกจากนี้ยังมีวิธีการแปลงเชิงเส้นโดยใช้ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของพารามิเตอร์ความยาก (วิธี mean and sigma) วิธีใช้ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของพารามิเตอร์ความยากที่กำหนดน้ำหนัก (วิธี weighted mean and

sigma) และวิธีใช้ค่าเฉลี่ยของพารามิเตอร์ความยากและอำนาจจำแนก (วิธี mean and mean) ซึ่งปรับสูตรจากการเชื่อมต่อในข้อมูลที่ข้อสอบตรวจให้คะแนนสองค่า และวิธีการปรับค่าพารามิเตอร์พร้อมกัน โดยการศึกษาตามทฤษฎีการตอบสนองข้อสอบมีการศึกษาทั้งใน Graded response model, Nominal response model, Partial Credit model, และ Generalize Partial Credit Model สำหรับวิธีการปรับเทียบตามทฤษฎีการทดสอบแบบดั้งเดิม ได้มีการนำวิธีการปรับเทียบเชิงเส้นตรง วิธีการปรับเทียบแบบค่าเฉลี่ยและวิธีอีควิปเปอร์เซ็นไทล์มาศึกษา

การศึกษาการปรับเทียบในโมเดลที่มีการตรวจให้คะแนนหลายค่า (Polytomous) ส่วนใหญ่เป็นการปรับเทียบโดยการจำลองข้อมูล Cohen และ Kim (1998) ได้เปรียบเทียบวิธี mean and mean, mean and sigma, weighted mean and sigma, Stocking and Lord (1983) และวิธีที่ขยายมาจาก Divgi (1985) ที่ Kim และ Cohen พัฒนาสำหรับการเชื่อมโยงสเกลภายใต้โมเดล Graded Response ผลการศึกษาสรุปว่าทุกวิธีให้ผลคล้ายคลึงกัน Baker (1997) ได้ศึกษาการแจกแจงการสุ่มตัวอย่างเชิงประจักษ์ของสัมประสิทธิ์การเชื่อมโยง ภายใต้โมเดล Graded Response , Kim และ Cohen (2002) ได้เปรียบเทียบการเชื่อมโยงโดยใช้วิธีการปรับเทียบแบบ Stocking and Lord และวิธีการปรับค่าพร้อมกันสำหรับข้อมูลที่จำลองให้สอดคล้องกับ โมเดล Graded Response พบว่าวิธีการปรับค่าพร้อมกันให้ผลถูกต้องมากกว่าเล็กน้อย Kolen และ Brennan (2004) เสนอแนะว่า เพื่อให้เกิดความชัดเจนยิ่งขึ้นควรมีการศึกษาเกี่ยวกับวิธีการเชื่อมโยงและเปรียบเทียบระหว่างวิธีการเชื่อมโยง และการปรับค่าพร้อมกันหรือแยกกันสำหรับโมเดลการตอบที่มีการให้คะแนนแบบหลายค่า

ส่วนการศึกษาการปรับเทียบแบบสอบรูปแบบผสมนั้นยังมีไม่มากนัก ซึ่งส่วนใหญ่เป็นการเปรียบเทียบวิธีการปรับเทียบแบบสอบรูปแบบผสม ในเงื่อนไขต่าง ๆ โดยพิจารณาถึงประสิทธิภาพของแต่ละวิธีการ รวมถึงความคลาดเคลื่อนที่เกิดขึ้นจากการปรับเทียบด้วย **การศึกษาการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม** มีดังนี้

Li, Lissitz และ Yang (1999) ได้ศึกษาวิธีการปรับเทียบคะแนนโดยใช้โค้งลักษณะข้อสอบกับแบบสอบรูปแบบผสม คือข้อสอบที่มีการตรวจให้คะแนน 2 ค่าและมากกว่า 2 ค่า ใน Generalized Partial Credit Model การศึกษาจำลองข้อมูลจากพารามิเตอร์ข้อสอบของแบบสอบการอ่านและการเขียนที่ใช้กับนักเรียนเกรด 4 ในโรงเรียนระดับตำบล องค์ประกอบต่างๆ ที่ใช้ในการศึกษา คือ 1) สัดส่วนของข้อสอบร่วมที่ให้คะแนน 2 ค่าและมากกว่า 2 ค่า โดยข้อสอบที่ให้คะแนนมากกว่า 2 ค่ามี 5 ข้อ แต่ละข้อมี 4 ลำดับขั้น รวมอยู่ในข้อสอบร่วมที่มีการให้คะแนนแบบ

2 ค่า 3 เงื่อนไข คือ 10 15 และ 20 ข้อตามลำดับ ซึ่งเงื่อนไขดังกล่าวใช้ผลจากการศึกษาของ Donoghue (1994) 2) ขนาดของหน่วยตัวอย่างมี 3 ขนาดคือ 1000 2000 และ 3000 คน 3) สถานการณ์ในการปรับเทียบมี 2 สถานการณ์ คือ สถานการณ์ที่ศึกษาค่าของพารามิเตอร์ใหม่ (parameter recovery) เมื่อใช้เฉพาะความคลาดเคลื่อนของของค่าประมาณพารามิเตอร์ในแบบสอบเทียบคะแนน สำหรับสถานการณ์ที่สองเป็นสถานการณ์เทียบคะแนนในกลุ่มข้อมูลจริง เมื่อความคลาดเคลื่อนของค่าประมาณพารามิเตอร์ที่มีอยู่ทั้งแบบสอบเทียบคะแนน และแบบสอบพื้นฐาน 4) ประเภทของการปรับเทียบ มี 2 ประเภทคือ การปรับเทียบตามแนวนอนและแนวตั้ง จะได้เงื่อนไขต่างๆ ที่ใช้ในการศึกษา 36 เงื่อนไข (3 สัดส่วนของข้อสอบ 2 ชนิด × 3 ขนาดหน่วยตัวอย่าง × 2 สถานการณ์การศึกษา × 2 สถานการณ์การปรับเทียบ) การประเมินความแม่นยำของสัมประสิทธิ์การปรับเทียบคะแนนโดยใช้ค่าความลำเอียง (BIAS คือ ค่าเฉลี่ยความแตกต่างพารามิเตอร์จริงและค่าประมาณที่สอดคล้องกัน) และค่ารากที่สองของความคลาดเคลื่อนเฉลี่ย (RMSE) ใช้วิธีการวิเคราะห์การถดถอยพหุคูณในการสรุปองค์ประกอบต่าง ๆ ที่ใช้ในการจำลองข้อมูลว่ามีผลกระทบต่อความแม่นยำของสัมประสิทธิ์การปรับเทียบหรือไม่ โดยเมื่อพิจารณาองค์ประกอบใดก็ให้องค์ประกอบอื่นๆ เป็นค่าคงที่

ผลจากการศึกษาปรากฏว่า วิธีการปรับเทียบคะแนนโดยใช้โค้งลักษณะข้อสอบในโมเดล GPCM สามารถประยุกต์ใช้กับแบบสอบรูปแบบผสมภายใต้เงื่อนไขต่าง ๆ ตามที่กำหนดได้ ความลำเอียงของสัมประสิทธิ์การปรับเทียบมีค่าใกล้ศูนย์และสอดคล้องกับค่า RMSE ที่มีค่าน้อย ซึ่งผู้วิจัยได้อภิปรายไว้ว่าค่าสองค่านี้ อาจมีค่ามากกว่านี้ถ้าประมาณค่าจากการทดสอบจริง นอกจากรูปร่างการแจกแจงสุ่มเชิงประจักษ์ของสัมประสิทธิ์การปรับเทียบคะแนนที่ได้รับการประมาณค่าตามเงื่อนไขต่าง ๆ มีรูปร่างเป็นโค้งระฆังคว่ำที่สมมาตรด้วยความแปรปรวนน้อยและไม่มี outlier สิ่งหนึ่งที่ปรากฏอยู่ในผลการวิจัยครั้งนี้ก็คือ ภายใต้สถานการณ์ที่ข้อสอบตรวจให้คะแนนหลายค่ารวมอยู่ในข้อสอบร่วมที่ให้คะแนนสองค่าที่มีจำนวนข้อมากกว่าไม่ได้ให้ค่าความลำเอียง และค่า RMSE ของสัมประสิทธิ์การปรับเทียบที่ต่ำกว่า ซึ่งคณะผู้วิจัยคาดว่าเกิดจากผลกระทบของการกำหนดน้ำหนักสัดส่วนของรูปแบบข้อสอบที่แตกต่างกัน จึงอาจเกิดผลต่อความแม่นยำของสัมประสิทธิ์การปรับเทียบที่ได้ จึงควรจะได้มีการศึกษาเงื่อนไขดังกล่าวให้ชัดเจนต่อไป โดยใช้แบบแผนที่มีเงื่อนไขที่ดีที่สุดของแบบสอบรูปแบบผสม (ซึ่งไม่ใช่หัวข้อของการศึกษาในครั้งนี้) ผลที่ได้ข้างต้นอาจจะเป็นการเตือนนักพัฒนาแบบสอบ ว่าลักษณะพารามิเตอร์ข้อสอบของกลุ่มข้อสอบร่วมมีความสำคัญต่อความแม่นยำของสัมประสิทธิ์การปรับเทียบคะแนนมากกว่าองค์ประกอบของจำนวนข้อสอบร่วม และสำหรับการวิเคราะห์ผลการถดถอยของ log[RMSE] ในค่าประมาณ A

หรือ B จากองค์ประกอบต่าง ๆ ที่จำลองขึ้นมา พบว่ามีความไวต่อการผันแปรและความแม่นยำของการประมาณค่าในแต่ละการประมาณค่าพารามิเตอร์ของการแปลง แสดงว่าองค์ประกอบต่างๆ ที่จำลองขึ้นมา ส่งผลต่อการประมาณค่าสัมประสิทธิ์การปรับเทียบ จุดหนึ่งที่น่าสนใจจากข้อค้นพบนี้คือ คณะผู้วิจัยได้ตั้งคำถามว่า ข้อค้นพบนี้สามารถที่จะใช้อ้างอิงไปยังการปรับเทียบที่ข้อมูลได้รับจากการทดสอบจากแบบสอบ 2 ฉบับใด ๆ ในสถานการณ์จริง แล้วให้ผลที่ตรงกันหรือไม่ในทุก ๆ เงื่อนไขที่เหมือนกัน

Bastari (2000) ได้ศึกษาผลขององค์ประกอบที่มีต่อความแม่นยำของการเชื่อมโยงค่าประมาณพารามิเตอร์ข้อสอบผสมระหว่างข้อสอบแบบเลือกตอบ (MC) และข้อสอบแบบเขียนตอบ (CR) โดยใช้โมเดลโลจีสติก 3 พารามิเตอร์ร่วมกับ Graded response model การศึกษาครั้งนี้ใช้แบบแผนแบบสอบร่วมที่มีทั้งข้อสอบแบบเลือกตอบและข้อสอบแบบเขียนตอบ องค์ประกอบที่ใช้ในการศึกษามี 6 องค์ประกอบ คือ 1) ความยาวของแบบสอบมี 2 ขนาดคือ 30 ข้อ และ 50 ข้อ 2) สัดส่วนของข้อสอบแต่ละชนิดในแบบสอบมี 2 ระดับ คือ ข้อสอบ MC 80% และ CR 20% และข้อสอบ MC 90% และ CR 10% 3) ความยาวของข้อสอบร่วมโดยใช้กฎร้อยละ 10 และร้อยละ 20 ของความยาวแบบสอบทั้งฉบับ 4) ขนาดของหน่วยตัวอย่าง 3 ขนาด คือ 400, 1000 และ 3000 คน 5) การแจกแจงความสามารถโดยเลือกการแจกแจงที่เป็นการแจกแจงปกติและจัดให้มีระดับความแตกต่างในค่าเฉลี่ย 2 ระดับ คือมีความแตกต่างเป็น 0 และ 1 หน่วยเบี่ยงเบนมาตรฐาน 6) วิธีการเชื่อมโยงคะแนนที่นำมาใช้ในการศึกษามี 2 วิธีคือใช้วิธีการปรับค่าพารามิเตอร์พร้อมกัน (Simultaneous calibration) และวิธีโค้งลักษณะข้อสอบ การศึกษาครั้งนี้เป็นการจำลองข้อมูลจากค่าประมาณพารามิเตอร์ข้อสอบของแบบสอบ MCAS ปี 1998 สำหรับวิชาวิทยาศาสตร์ ในนักเรียนเกรด 10 สุ่มข้อมูลการตอบข้อสอบทั้งข้อสอบ MC และ CR เพื่อให้ได้แบบสอบฉบับพื้นฐาน 2 ฉบับ แต่ละฉบับมีข้อสอบ MC 40 ข้อ และ CR 9 ข้อ ตามลำดับ แบบสอบร่วมประกอบด้วยข้อสอบ MC 9 ข้อ และข้อสอบ CR 9 ข้อ ตามลำดับ แบบสอบทั้งสองฉบับถูกสร้างจนเกือบคู่ขนานกัน โดยเฉพาะในค่าเฉลี่ยความยากของข้อสอบ MC และดำเนินการตามเงื่อนไขในองค์ประกอบได้ทั้งหมด 96 สถานการณ์ มีการทำซ้ำข้อมูล 50 ครั้ง โปรแกรมที่ใช้ในการวิเคราะห์คือ MULTILOG โดยการปรับค่าพารามิเตอร์พร้อมกันและแยกกัน ค่าประมาณพารามิเตอร์เมื่อปรับข้อมูลพร้อมกันจะนำมาใช้ได้โดยตรงสำหรับค่าประมาณที่ปรับค่าพารามิเตอร์แยกกันถูกนำมาเชื่อมโยงตามวิธีโค้งลักษณะข้อสอบ เกณฑ์ที่ใช้ในการประเมินมี 2 เกณฑ์ เกณฑ์ที่ 1 คือ ความแตกต่างกำลังสองเฉลี่ย (MSD) ประกอบด้วยความลำเอียง และความ

แปรปรวนสำหรับแต่ละค่าประมาณพารามิเตอร์ ได้แก่ พารามิเตอร์ความยากและพารามิเตอร์อำนาจจำแนก และเกณฑ์ที่สองคือ รากของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) เป็นการพิจารณาความแตกต่างระหว่างโค้งลักษณะข้อสอบ (ICC) จากพารามิเตอร์ข้อสอบกับค่าประมาณ เป็นการประเมินโดยใช้ค่าพารามิเตอร์ข้อสอบพร้อม ๆ กันทั้งสองเกณฑ์ จะถูกเปรียบเทียบจากข้อสอบที่ไม่ได้ร่วมในแต่ละฉบับเท่านั้น เกณฑ์ทั้งสองใช้หลักการที่ว่าค่า RMSD และ RMSE ที่มีค่าน้อยกว่าจะมีความแม่นยำมากกว่า นอกจากนี้ยังมีการวิเคราะห์ความแปรปรวน (ANOVA) เพื่อสำรวจข้อค้นพบทั่ว ๆ ไป ก่อนที่จะมีการศึกษาในรายละเอียดขั้นต่อไป

ผลการวิจัยปรากฏว่า รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) และรากของความแตกต่างกำลังสองเฉลี่ย (RMSD) สำหรับพารามิเตอร์ความยากให้ผลการประเมินที่สอดคล้องกัน เมื่อพิจารณาในภาพรวมแล้วพบว่าเงื่อนไขขององค์ประกอบต่าง ๆ ที่ทำให้ได้ผลมีประสิทธิภาพมากกว่าคือ แบบสอบที่ยาวกว่า สัดส่วนขนาดใหญ่ของข้อสอบเลือกตอบในแบบสอบ แบบสอบร่วมที่ยาวกว่า หน่วยตัวอย่างขนาดใหญ่กว่า ความไม่แตกต่างระหว่างความสามารถของกลุ่มที่ใช้ในการเชื่อมโยงคะแนนบนแบบสอบ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน โดยขนาดของกลุ่มตัวอย่างและวิธีการแปลงคะแนนจะมีอิทธิพลต่อผลที่ได้ หน่วยตัวอย่างขนาดใหญ่ก่อให้เกิดค่า RMSE RMSD ความลำเอียง ความแปรปรวนน้อยที่สุด สำหรับแต่ละชนิดของข้อสอบ ในสถานการณ์เดียวกันนี้เกิดจากวิธีการปรับค่าพารามิเตอร์พร้อมกัน องค์ประกอบอื่น ๆ ได้แก่ ความยาวของแบบสอบ สัดส่วนของข้อสอบ ความยาวของข้อสอบร่วม และความแตกต่างระหว่างความสามารถของหน่วยตัวอย่างมีอิทธิพลที่ไม่ชัดเจนนัก และปรากฏในสถานการณ์เฉพาะคือ ในองค์ประกอบความยาวของแบบสอบที่เป็นข้อสอบ MC แบบสอบที่ยาวกว่าจะให้ค่า RMSE RMSD ความลำเอียงและความแปรปรวนน้อยกว่าแบบสอบที่สั้นกว่า แต่ให้ผลตรงข้ามในข้อสอบ CR ซึ่งผู้วิจัยได้อภิปรายว่าน่าจะมาจากพารามิเตอร์ข้อสอบที่ใช้ในการสร้างข้อมูลพารามิเตอร์ความยากของข้อสอบ CR ที่ไม่ได้ร่วมมีขนาดใหญ่กว่าข้อสอบ MC มาก (ค่าเฉลี่ยพารามิเตอร์ความยากสำหรับข้อสอบ MC มีค่า 0.097 ขณะที่ข้อสอบ CR มีค่า 1.546) แสดงว่าแบบสอบ CR มีความยากกว่า จึงส่งผลให้เกิดความคลาดเคลื่อนขนาดใหญ่ในการเชื่อมโยง สำหรับองค์ประกอบความแตกต่างในความสามารถ ผลที่ผิดปกติจนเห็นได้ชัดเจนคือความคลาดเคลื่อนในการเชื่อมโยงน้อยกว่าเมื่อความแตกต่างของความสามารถในกลุ่มเพิ่มขึ้น ที่เป็นเช่นนี้เนื่องจากความยากของข้อสอบ CR มีค่าเฉลี่ยถึง 1.546 เมื่อใช้กับกลุ่มผู้สอบที่มีความสามารถเฉลี่ยเท่ากับ 0 ข้อสอบ CR จึงยากเกินไป เมื่อเป็นเช่นนี้ การนำมาใช้กับผู้สอบที่มีความสามารถเฉลี่ยเป็น 1 ข้อสอบ CR จึงประมาณค่าอย่างแม่นยำมากกว่า ทำให้เกิดความ

คลาดเคลื่อนของการเชื่อมโยงที่เล็กลง สำหรับข้อสอบ MC ที่มีค่าเฉลี่ยความยากเป็น 0.096 จึงให้ค่าประมาณที่แม่นยำเมื่อการแจกแจงความสามารถของกลุ่มผู้สอบสองกลุ่มเท่าเทียมกัน

Tate (2000) ได้ศึกษาวิธีการเชื่อมโยง 2 วิธีสำหรับแบบสอบรูปแบบผสมโดยการศึกษานี้มีจุดมุ่งหมายเพื่อตรวจสอบประสิทธิภาพของวิธีการเชื่อมโยง โดยใช้การประมาณค่าสัมประสิทธิ์การเชื่อมโยงคะแนนโดยใช้รูปแบบการปรับเทียบ 2 แนวทาง คือ แนวทาง mean and sigma (Cohen และ Kim, 1998) และโมเดลการตอบแบบ graded response เป็นภาคขยายของกระบวนการ Stocking and Lord ที่เสนอโดย Baker (1992) การศึกษานี้ใช้การจำลองข้อมูลโดยคอมพิวเตอร์เพื่อพิจารณาข้อคำถามเกี่ยวกับ ความคลาดเคลื่อนในการประมาณค่าสัมประสิทธิ์การเชื่อมโยง ประกอบไปด้วยทั้งการประมาณค่าความลำเอียง และความคลาดเคลื่อนในการสุ่ม สำหรับแบบสอบแบบผสมจะประกอบด้วยข้อสอบแบบ MC และ CR ซึ่ง Tate ต้องการศึกษาการเปลี่ยนแปลงของความคลาดเคลื่อนในการประมาณค่าในปัจจัยต่าง ๆ คือ เมื่อสัดส่วนของข้อสอบ MC และ CR เปลี่ยนไป ศึกษาความคลาดเคลื่อนที่เปลี่ยนไป เมื่อจำนวนของข้อสอบหลัก และจำนวนของผู้สอบในปีที่ 1 ที่ใช้ในการศึกษาการเชื่อมโยงในปีที่ 2 เปลี่ยนไป เมื่อใช้เพียงข้อสอบ MC เป็นข้อสอบร่วม เมื่อมีการเดาเป็นปัจจัยในการตอบของนักเรียนในการตอบข้อสอบ MC และเมื่อมีการละเมิดข้อตกลงเบื้องต้นของความเป็นเอกมิติ (unidimentional) ในข้อสอบร่วมที่ประกอบด้วยข้อสอบ CR และ MC การออกแบบการเชื่อมโยงเป็นแบบ non equivalent group และใช้ข้อสอบร่วมในการประมาณค่าสัมประสิทธ์การเชื่อมโยง โดยใช้วิธีการเชื่อมโยง 2 วิธี คือ mean and sigma ที่เสนอโดย Kohen และ Kim (1998) เพื่อคำนวณค่าสัมประสิทธิ์การเชื่อมโยง A และ B และวิธี Stocking-Lord (1983) เพื่อกำหนดสัมประสิทธิ์ของการเชื่อมโยงโดยใช้โปรแกรม EQUATE (version 2.1; Baker, 1993)

การศึกษาความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ข้อสอบร่วมครั้งนี้ ไม่ได้นำเสนอผลการวิจัยทั้งหมด ซึ่งส่วนของผลการวิจัยที่นำเสนอ คือ ความคลาดเคลื่อนในการเชื่อมโยงที่เกี่ยวกับกระบวนการที่เสนอสำหรับการเชื่อมโยงจากปีหนึ่งสู่อีกปีหนึ่งของแบบสอบรูปแบบผสม ผลจากกระบวนการประมาณค่าทั้ง 2 วิธีสำหรับข้อสอบที่ตรวจให้คะแนนหลายค่าทั้งวิธี mean and sigma และ Stocking and Lord จะให้ผลคล้าย ๆ กัน คือความแม่นยำของการประมาณค่าของสัมประสิทธิ์การเชื่อมโยงจะเพิ่มขึ้นเมื่อจำนวนข้อสอบร่วมมากขึ้น และผู้สอบจากปีที่ 1 เพิ่มขึ้นสำหรับการศึกษาเชื่อมโยง การใช้เพียงข้อสอบ MC อย่างเดียวเป็นข้อสอบร่วมจะแสดงถึงความน่าเชื่อถือ เมื่อแบบสอบมีความเป็นเอกมิติอย่างมาก แต่เป็นผลในเรื่องของความ

ลำเอียงในการเชื่อมโยงจะมีมากถ้าข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกมิติถูกละเลย หากมีการละเลยข้อตกลงเบื้องต้นเรื่องความเป็นเอกมิติการเชื่อมโยงจะมีความแกร่งพอสมควรเมื่อชุดของข้อสอบร่วมประกอบด้วยสัดส่วนที่เหมาะสมของ MC และ CR

Lee และคณะ (2001) ได้เปรียบเทียบผลการปรับเทียบ เมื่อใช้ทฤษฎีการตอบสนองข้อสอบที่มีการตรวจให้คะแนนสองค่าแบบโมเดลโลจีสติก 3 พารามิเตอร์ และโมเดล IRT ที่มีการตรวจให้คะแนนหลายค่า แบบ Graded Response Model และ Nominal Response Model ในบริบทของการปรับเทียบที่แบบสอบประกอบด้วยแบบสอบย่อย (Testlets) โดยใช้ข้อมูลจาก IOWA Tests of Basic Skills (ITBS) ปี 1995 ฉบับ M และฉบับ K หน่วยตัวอย่างเป็นนักเรียนในเกรด 8 แบบสอบ 3 ฉบับประกอบด้วยแบบสอบด้านการอ่าน แผนที่และไดอะแกรม การแก้ปัญหาและตีความทางคณิตศาสตร์ ในแบบสอบฉบับ K ข้อสอบที่สัดเกี่ยวกับการอ่านมีจำนวน 49 ข้อ 8 ตอน แต่ละตอนประกอบด้วยข้อคำถามจำนวน 9 4 7 5 5 6 3 10 ข้อ ตามลำดับ ข้อสอบที่วัดแผนที่และไดอะแกรมมีจำนวน 33 ข้อ 5 ตอน แต่ละตอนประกอบด้วยข้อสอบจำนวน 7 7 6 6 7 ข้อ ตามลำดับ ข้อสอบที่วัดด้านคณิตศาสตร์จำนวน 36 ข้อ 8 ตอน แต่ละตอนประกอบด้วยข้อสอบจำนวน 8 4 4 4 4 4 4 4 ข้อ ตามลำดับ ในแบบสอบฉบับ M ข้อสอบที่สัดเกี่ยวกับการอ่านมีจำนวน 49 ข้อ 7 ตอน แต่ละตอนประกอบด้วยข้อคำถามจำนวน 8 4 9 4 5 8 11 ข้อ ตามลำดับ ข้อสอบที่วัดแผนที่และไดอะแกรมมีจำนวน 33 ข้อ 5 ตอน แต่ละตอนประกอบด้วยข้อสอบจำนวน 7 7 6 6 7 ข้อ ตามลำดับ ข้อสอบที่วัดด้านคณิตศาสตร์จำนวน 36 ข้อ 7 ตอน แต่ละตอนประกอบด้วยข้อสอบจำนวน 8 6 6 4 4 4 4 ข้อ ตามลำดับ ข้อสอบแต่ละข้อในแต่ละตอนให้คะแนนสองค่า จำนวนหน่วยตัวอย่างในแบบสอบฉบับ K จำแนกตามองค์ประกอบการวัดจำนวน 663 632 และ 537 คน ตามลำดับ และจำนวนหน่วยตัวอย่างในแบบสอบฉบับ M จำแนกตามองค์ประกอบของการวัดมีจำนวน 680 653 และ 561 คน ตามลำดับ

แบบแผนการปรับเทียบเป็นแผนกลุ่มสมมูลโดยใช้การปรับเทียบตามทฤษฎีการทดสอบแบบดั้งเดิมเป็นเกณฑ์ ได้แก่ วิธีการปรับเทียบแบบค่าเฉลี่ยเชิงเส้นตรง และอีควิปเปอร์เซ็นไทล์ เหตุที่ใช้วิธีการปรับเทียบในทฤษฎีการทดสอบแบบดั้งเดิมเป็นเกณฑ์เนื่องจากวิธีการในทฤษฎีนี้คะแนนรวมไม่ได้รับอิทธิพลจากการฝ่าฝืนข้อตกลงเบื้องต้นของความเป็นอิสระของข้อสอบ IOWA ที่ปฏิบัติกันอยู่นั้นใช้การปรับเทียบวิธีอีควิปเปอร์เซ็นไทล์อยู่แล้ว การวิเคราะห์ข้อมูลที่นำมาใช้ในการศึกษาครั้งนี้ใช้โปรแกรม RAGE ในการปรับเทียบโดยวิธีค่าเฉลี่ยเชิงเส้นตรงและอีควิปเปอร์เซ็นไทล์ สำหรับการปรับเทียบ IRT ข้อสอบที่ตรวจให้คะแนนสองค่าใช้โปรแกรม BILOG ในการ

ประมาณค่าพารามิเตอร์ แล้วใช้โปรแกรมเทียบคะแนนจริงและคะแนนสังเกต สำหรับการปรับเทียบตามทฤษฎี IRT ข้อสอบที่ตรวจให้คะแนนหลายค่าทั้ง 2 โมเดลใช้โปรแกรม MULTILOG ในการประมาณค่าพารามิเตอร์ แล้วจึงวิเคราะห์โค้งลักษณะข้อสอบทั้งสองฉบับและเทียบคะแนนจริงและคะแนนสังเกต โดยใช้โปรแกรมฟอร์แทนที่เขียนขึ้น โดยเปรียบเทียบโมเมนต์ของการแจกแจงคะแนนแปลงจากการปรับเทียบแต่ละวิธีกับโมเมนต์ของการแจกแจงจากแบบสอบฉบับเป้าหมาย และเปรียบเทียบความแตกต่างระหว่างคะแนนแปลงจากการปรับเทียบ IRT แต่ละวิธีกับคะแนนแปลงจากการปรับเทียบตามทฤษฎีการทดสอบแบบดั้งเดิมแต่ละวิธี โดยใช้ดัชนีรากของค่าเฉลี่ยกำลังสองถ่วงน้ำหนัก (RSM) และไม่ได้ถ่วงน้ำหนัก (URMS)

ผลที่ได้จากการวิจัย ปรากฏว่า ข้อตกลงเบื้องต้นเกี่ยวกับความเป็นอิสระและความเป็นเอกมิติสำหรับโมเดล IRT ถูกฝ่าฝืนสำหรับแบบสอบที่ประกอบด้วยแบบสอบฉบับย่อย เมื่อข้อสอบแต่ละข้อถูกใช้เป็นหน่วยการวิเคราะห์ โดยแบบสอบด้านคณิตศาสตร์มีการฝ่าฝืนน้อยที่สุด แต่เมื่อใช้แบบสอบย่อยเป็นหน่วยของการวิเคราะห์ กลับพบว่าเป็นไปตามข้อตกลงเบื้องต้น ดังนั้นจึงเป็นเหตุผลสนับสนุนว่าการปรับเทียบแบบสอบที่ประกอบด้วยแบบสอยย่อยควรใช้โมเดลการตอบแบบหลายค่าที่เป็นเอกมิติ แทนที่โมเดล IRT ที่มีการตรวจให้คะแนนแบบสองค่า โดยเทียบคะแนนจริงและคะแนนสังเกตได้ในโมเดล IRT ที่ข้อสอบให้คะแนนหลายค่าให้ผลการปรับเทียบเหมือนกับการปรับเทียบตามทฤษฎีการทดสอบแบบดั้งเดิมมากกว่าการปรับเทียบจริงหรือคะแนนสังเกตได้ของโมเดล IRT ที่ข้อสอบตรวจให้คะแนนสองค่า โดยเฉพาะในแบบสอบการอ่าน และแผนที่ แต่ให้ผลไม่ชัดเจนสำหรับแบบสอบคณิตศาสตร์ อาจจะเนื่องด้วยมีการฝ่าฝืนข้อตกลงเบื้องต้นของโมเดล IRT น้อยกว่าแบบสอบฉบับอื่นๆ (ในการเทียบฟังก์ชันการปรับเทียบกับฟังก์ชันการปรับเทียบในทฤษฎีการทดสอบแบบดั้งเดิม) ในโมเดล IRT ที่ตรวจให้คะแนนหลายค่าแบบสอบเป็นไปตามข้อตกลงเบื้องต้นของโมเดล ดังนั้น ความสัมพันธ์ของการปรับเทียบในโมเดล IRT ที่ข้อสอบตรวจให้คะแนนหลายค่าจึงดีกว่าการปรับเทียบในโมเดล IRT ที่ข้อสอบตรวจให้คะแนนสองค่า โดยเฉพาะ Nominal Response Model และGraded Response Model จะเป็นโมเดลทางเลือกที่ดีที่สุดสำหรับการปรับเทียบแบบสอบที่ประกอบด้วยแบบสอบฉบับย่อย

Kim (2004) ได้ศึกษากระบวนการเชื่อมโยงสเกลตามทฤษฎีการตอบสนองข้อสอบที่เป็นเอกมิติสำหรับแบบสอบรูปแบบผสมและความแกร่งของการเชื่อมโยงสู่ความเป็นพหุมิติ จุดมุ่งหมายเบื้องต้นของการศึกษานี้ เพื่อค้นหากระบวนการวางตำแหน่งของพารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบ ที่ได้มาจากการจัดการทดสอบโดยใช้แบบสอบรูปแบบผสมใน

สถานการณ์ที่ต่างกันบนสเกลร่วมกัน การบวนการเชื่อมโยง 4 วิธีที่ศึกษาคือ วิธีการโมเมนต์ 2 วิธี (mean/mean และ mean/sigma) และวิธีการโค้งคุณลักษณะ 2 วิธี (Heabara และ Stocking and Lord) การวิจัยนี้เป็นการศึกษาจำลองข้อมูล 2 ส่วนเพื่อประเมินความถูกต้องสัมพัทธ์ของการเชื่อมโยงทั้ง 4 วิธี สำหรับแบบสอบที่เป็นเอกมิติและตรวจสอบความแกร่งของการเชื่อมโยงไปสู่แบบสอบที่เป็นพหุมิติ ในการเปรียบเทียบกับการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent calibration)

การจำลองข้อมูลส่วนแรกอยู่ภายใต้เงื่อนไขความความเป็นเอกมิติของแบบสอบที่ประกอบด้วยข้อสอบ 2 ลักษณะที่ต่างกันคือ ข้อสอบแบบเลือกตอบ (Multiple choice : MC) และข้อสอบแบบเขียนตอบ (Construct response : CR) โดยพื้นฐานแล้วอิทธิพลของรูปแบบ (format effect : FE) ก็คือความสัมพันธ์ระหว่างโค้งสร้างที่สำคัญของข้อสอบทั้ง 2 ลักษณะที่วัดได้จาก MC และ CR การจำลองข้อมูลส่วนที่ 2 เป็นแบบพหุมิติที่สะท้อนถึงอิทธิพลของรูปแบบ (FE) ปัจจัยที่ศึกษามี 3 ปัจจัยคือ การเชื่อมโยงกับกลุ่มผู้สอบไม่เท่าเทียม ชนิดของการเชื่อมโยงในการจำลองข้อมูลจากส่วนแรก และอิทธิพลของรูปแบบที่ศึกษาเพิ่มเติมในการจำลองข้อมูลส่วนที่สอง การศึกษาจำลองข้อมูลทั้ง 2 ส่วน ใช้การออกแบบกลุ่มไม่เท่าเทียมใช้ข้อสอบร่วม และใช้โปรแกรม MULTILOG ในการประมาณค่าพารามิเตอร์

ผลการวิจัย พบว่าวิธีการโค้งคุณลักษณะให้ผลน่าพอใจกว่าวิธีโมเมนต์ โดยไม่คำนึงถึงอิทธิพลของรูปแบบ สำหรับวิธีการโค้งคุณลักษณะ 2 วิธี ให้ผลที่แตกต่างกันเพียงเล็กน้อย โดยทั่วไปวิธีการปรับค่าพร้อมกันจะให้ผลดีกว่าวิธีการเชื่อมโยงทั้ง 4 วิธี คือมีความถูกต้องและยังแกร่งมากพอที่จะใช้กับแบบสอบที่ต้องพิจารณาอิทธิของรูปแบบด้วย

Robert (2005) ได้ศึกษาอิทธิพลของรูปแบบข้อสอบในการปรับเทียบที่ใช้ข้อสอบสอบร่วม งานวิจัยที่ผ่านมาแสดงให้เห็นว่ามีการใช้ข้อสอบเลือกตอบเพียงรูปแบบเดียวเป็นข้อสอบร่วมซึ่งเป็นตัวแทนของแบบสอบทั้งฉบับ แต่ปัจจุบันแบบสอบส่วนใหญ่ประกอบด้วยข้อสอบที่มีลักษณะต่างกันภายในแบบสอบฉบับเดียว จึงมีคำถามว่าข้อสอบร่วมควรจะบรรจุข้อสอบรูปแบบที่ต่างกันไว้หรือไม่ การวิจัยครั้งนี้จึงสำรวจอิทธิพลของรูปแบบข้อสอบในการปรับเทียบที่ใช้ข้อสอบร่วม ข้อมูลสำหรับการวิจัยได้มาจากข้อมูลจริงในแบบสอบวัดการอ่าน และคณิตศาสตร์ แต่ยังจำลองข้อมูลเพื่อเป็นการยืนยันผล ค่าประมาณพารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบจากการปรับเทียบข้อมูลจริงใช้เพื่อหาโครงสร้างไว้สำหรับการจำลองข้อมูล ตัวแปรที่ควบคุมในการจำลองข้อมูล ประกอบด้วยความสัมพันธ์ระหว่างมิติของรูปแบบข้อสอบ และการเปลี่ยนแปลงใน

ความสามารถของประชากรจากปีที่ 1 ไปสู่ปีที่ 2 โดยระดับความสัมพันธ์ระหว่างมิติของรูปแบบข้อสอบมี 3 ระดับ คือ ความสัมพันธ์เป็นเอกมิติ, $\rho = 0.5$ และ $\rho = 0.8$ ความเปลี่ยนแปลงของความสามารถของประชากรแปรเปลี่ยนไปตามมิติ การวิเคราะห์ข้อมูลการปรับเทียบที่ใช้ข้อสอบร่วมตามแนวทฤษฎีการตอบสนองข้อสอบใช้วิธีการปรับเทียบคะแนนจริง คะแนนสังเกตได้ และปรับเทียบความสามารถโดยใช้ข้อสอบร่วม 2 ชุดที่ถูกเลือกมาเป็นอย่างดี การเชื่อมโยงชุดแรกข้อสอบร่วมมีข้อสอบหลายตัวเลือกหลายข้อ และข้อสอบเขียนตอบข้อเดียว ในขณะที่การเชื่อมโยงชุดอื่นข้อสอบร่วมมีเพียงข้อสอบหลายตัวเลือกเพียงอย่างเดียว ซึ่งข้อสอบร่วมเลือกมาให้มีความเป็นตัวแทนของแบบสอบทั้งฉบับ ผลของการวิจัยแสดงให้เห็นว่า ความสามารถที่แตกต่างกันมีอิทธิพลต่อการปรับเทียบมากกว่าระดับมิติความสัมพันธ์ของรูปแบบข้อสอบ นอกจากนี้ผลการปรับเทียบแสดงให้เห็นว่าถ้ามีข้อสอบแบบเขียนตอบ หรือใช้ข้อสอบแบบเขียนตอบเพียงอย่างเดียวเป็นข้อสอบร่วม จะให้ผลการปรับเทียบที่ดีกว่าเมื่อความสามารถของประชากรที่แตกต่างกันใน 2 มิติที่ศึกษา นั่นคือ ข้อสอบแบบเขียนตอบที่อยู่ในการเชื่อมโยงจะให้ผลของคะแนนเท่าเทียมได้ดีกว่าการเชื่อมโยงที่มีข้อสอบเลือกตอบเพียงอย่างเดียวความสามารถของประชากรในมิติของข้อสอบเขียนตอบจะดีกว่าความสามารถในมิติของข้อสอบแบบเลือกตอบ ข้อค้นพบนี้สนับสนุนงานวิจัยที่ผ่านมา ซึ่งเป็นสิ่งสำคัญสำหรับการออกแบบข้อสอบร่วม

Kim และ Lee (2006) ได้ศึกษาเปรียบเทียบการเชื่อมโยงคะแนนโดยใช้เทคนิคการปรับเทียบคะแนน 4 วิธีที่ใช้สำหรับแบบสอบรูปแบบผสม โดยข้อสอบที่มีการตรวจให้คะแนนแบบ 2 ค่าใช้โมเดล IRT แบบ 3 พารามิเตอร์ ข้อสอบที่มีการตรวจให้คะแนนแบบหลายค่าใช้โมเดล Generalized Partial Credit ที่มีการตรวจให้คะแนนเป็น 5 ลำดับขั้น วิธีการปรับเทียบคะแนนที่ใช้ในการศึกษาครั้งนี้ได้แก่ mean and mean, mean and sigma, Heabara, และ Stocking and Lord องค์ประกอบที่ใช้ในการศึกษาได้แก่ 1) การออกแบบการเชื่อมโยง (equivalent และ non equivalent group) 2) ขนาดของกลุ่มตัวอย่าง 2 ขนาด ได้แก่ 500 และ 3000 3) สัดส่วนของข้อสอบที่มีการตรวจให้คะแนนสองค่า (DS) และข้อสอบที่มีการตรวจให้คะแนนหลายค่า (PS) 3 เงื่อนไข ได้แก่ DS 10 ข้อ PS 10 ข้อ, DS 20 ข้อ PS 5 ข้อ และ DS 30 ข้อ PS 2 ข้อ โดยแบบสอบรูปแบบผสมทั้ง 3 รูปแบบนี้ มีจำนวนข้อสอบร่วมเป็น DS 10 ข้อ และ PS 2 ข้อ 4) รูปแบบของการเชื่อมโยง 3 รูปแบบ คือ เชื่อมโยงข้อสอบร่วมทั้ง DS และ PS (ปรับค่าพร้อมกัน: simultaneous linking), เชื่อมโยงข้อสอบร่วมเฉพาะ DS และ เชื่อมโยงข้อสอบร่วมเฉพาะ PS จะได้เงื่อนไขต่างๆ ที่ใช้ในการศึกษา 36 เงื่อนไข (2 การออกแบบการเชื่อมโยง 2 รูปแบบ × 2 ขนาดหน่วยตัวอย่าง × 3

สัดส่วนของข้อสอบ DS และ PS ในแบบสอบ × 3 สถานการณ์การเชื่อมโยง) ซึ่งเงื่อนไขทั้ง 36 เงื่อนไขนี้ใช้ในการเชื่อมโยงคะแนนทั้ง 4 วิธี การประมาณค่าพารามิเตอร์ของข้อสอบใช้โปรแกรม MULTILOG (Thissen, 1991) ส่วนการเชื่อมโยงคะแนนแต่ละวิธีใช้โปรแกรม STUIRT (Kim, Kolen, 2004b) สำหรับการเชื่อมโยง 4 วิธีสำหรับแบบสอบรูปแบบผสม เกณฑ์ที่ใช้ในการประเมินวิธีการเชื่อมโยงพิจารณาจากเกณฑ์โค้งลักษณะรายการคำตอบ (category characteristic curve)

ผลที่ได้จากการศึกษาปรากฏว่า วิธีโค้งคุณลักษณะมีค่า MSE (mean square error) ต่ำกว่าวิธีโมเมนต์ในทุก ๆ เงื่อนไข คือการเชื่อมโยงด้วยวิธี Heabara มีค่า MSE ต่ำที่สุดเมื่อเทียบกับ 3 วิธีที่เหลือ ซึ่งวิธี Heabara, และ Stocking and Lord มีค่า MSE น้อยกว่าวิธีอื่นเป็นผลเนื่องมาจากมีความแปรปรวนต่ำซึ่งสอดคล้องกับผลการวิจัยของ Hanson และ Beguin (2002) และค่าเฉลี่ยของความลำเอียงกำลังสอง (square bias: SB) และค่า MSE ในการเชื่อมโยงกลุ่มที่ไม่เท่าเทียมกันมีค่าสูงกว่าการเชื่อมโยงกลุ่มเท่าเทียมกัน ค่า MSE จะลดลงเมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้นซึ่งใน SB ก็เช่นเดียวกัน ในเงื่อนไขการเชื่อมโยงพร้อมกัน (DS และ PS) จะมีค่า MSE มากขึ้นเมื่อจำนวนข้อสอบ DS เพิ่มขึ้น แต่จะมีค่าน้อยกว่าการเชื่อมโยงข้อสอบร่วมเพียงรูปแบบใดรูปแบบหนึ่ง (DS หรือ PS) แต่ถ้าเป็นการเชื่อมโยงด้วยข้อสอบรูปแบบใดรูปแบบหนึ่งค่า MSE จะลดลงเมื่อจำนวนข้อสอบเพิ่มขึ้น ภายใต้เงื่อนไขของแบบสอบที่มี DS 30 และ PS 2 นั้นจะให้ค่า MSE โดยเฉลี่ยแล้วสูงที่สุดเมื่อใช้วิธีการเชื่อมโยงแบบ mean and mean

เนื่องจากการศึกษาครั้งนี้เป็นการจำลองข้อมูลจึงควรระวังในการนำผลการศึกษาหรือข้อสรุปไปใช้ในการสรุปอ้างอิงสู่สถานการณ์อื่น ๆ โดยเฉพาะอย่างยิ่งพารามิเตอร์ของข้อสอบที่จำลองขึ้นมาเป็นการจำลองโดยใช้สถิติมากกว่าที่จะจำลองมาจากข้อมูลจริง และมีการควบคุมความคลาดเคลื่อนในการสุ่มเป็นอย่างดีซึ่งข้อมูลเหล่านี้เป็นข้อมูลที่เหมาะสมในการศึกษาทางทฤษฎีมากกว่าข้อมูลจากสถานการณ์จริง ดังนั้นหากมีการศึกษาต่อไปควรนำข้อมูลจากสถานการณ์จริงเข้ามาศึกษา ซึ่งอาจจะพบข้อค้นพบใหม่เกี่ยวกับวิธีการเชื่อมโยงทั้ง 4 วิธีนี้

จากผลการศึกษาข้างต้นแสดงให้เห็นว่าวิธีเชื่อมโยงโค้งคุณลักษณะให้ผลที่น่าพอใจมากกว่าวิธีโมเมนต์ ซึ่งการเชื่อมโยงข้อสอบทั้ง 2 รูปแบบพร้อมกันให้ผลที่ถูกต้องมากกว่าการเชื่อมโยงเพียงข้อสอบรูปแบบใดรูปแบบหนึ่งอย่างเดียว ซึ่งจากผลนี้ก็อาจกล่าวได้ว่าความถูกต้องในการเชื่อมโยงจะสูงขึ้นเมื่อจำนวนข้อสอบร่วมเพิ่มขึ้น และเมื่อเปรียบเทียบระหว่างวิธีการเชื่อมโยงรูปแบบ DS อย่างเดียว กับการเชื่อมโยงรูปแบบ PS อย่างเดียว พบว่าการเชื่อมโยงรูปแบบใดรูปแบบหนึ่งเกิดความคลาดเคลื่อนต่ำกว่าและเกิดความผันแปรของ MSE ในกลุ่มของวิธีการเชื่อมโยงน้อยที่สุด นั่นคือความแตกต่างเพียงเล็กน้อยในความถูกต้องของการเชื่อมโยงระหว่างการเชื่อมโยง

ข้อสอบ 2 รูปแบบพร้อมกัน การเชื่อมโยงรูปแบบใดรูปแบบหนึ่ง ข้อค้นพบนี้บอกได้ว่า ถ้าสถานการณ์ในทางปฏิบัติแล้วไม่ยินยอมให้ใช้การเชื่อมโยงข้อสอบ 2 รูปแบบพร้อมกัน ทางเลือกคืออาจจะเลือกใช้วิธีการเชื่อมโยงข้อสอบรูปแบบใดแบบหนึ่งหลังจากพิจารณาทั้งจำนวนของลำดับขั้นการตอบและความคงที่ของการประมาณค่าพารามิเตอร์ของข้อสอบแต่ละแบบอย่างรอบคอบ

Kim และ Kolen (2006) ได้ทำการศึกษาวิธีการเชื่อมโยงต่อจาก Kim และ Lee (2006) เพื่อตรวจสอบว่าวิธีการเหล่านั้นมีความแกร่งในการนำไปใช้กับข้อมูลที่มีความเป็นพหุมิติอันเนื่องมาจากรูปแบบข้อสอบหรือไม่ แล้วเปรียบเทียบวิธีการทั้งสี่กับการแปลงสเกลพร้อมกันภายใต้การออกแบบการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม แหล่งของข้อมูลที่ใช้ในการจำลองข้อมูลเพื่อการศึกษาครั้งนี้มาจาก การประเมินทางวิทยาศาสตร์ในปี 1996 ของการทดสอบ NAEP และนำข้อมูลเหล่านี้มาประมาณค่าพารามิเตอร์ของข้อสอบ ข้อมูลที่จำลองให้เป็นพหุมิติได้มีการกำหนดรูปแบบข้อสอบ 2 รูปแบบมีความสัมพันธ์กัน ส่วนปัจจัยอื่น ๆ คือกำหนดระดับความสัมพันธ์เป็น 3 ระดับคือ 1, 0.8 และ 0.5 และชนิดของแบบสอบรูปแบบผสมกำหนดเป็น 2 รูปแบบคือ wide-range และ narrow-range และกำหนดระดับความไม่เท่าเทียมของกลุ่มผู้สอบเป็น 3 ระดับ คือ ค่าเฉลี่ยของกลุ่มเป็น 0, 0.5 และ 1 ข้อสอบร่วมประกอบด้วยข้อสอบทั้งแบบตรวจให้คะแนนสองค่าและตรวจให้คะแนนมากกว่าสองค่า โมเดลที่ใช้ในการศึกษาครั้งนี้คือ 3PL/GPC การแปลงสเกลประกอบด้วยแบบพร้อมกันและแบบแยกกันทั้ง 4 วิธี เพื่อใช้ในการวางตำแหน่งแบบสอบฉบับใหม่ บนแบบสอบฉบับเก่า ผลการปรับเทียบประเมินโดยใช้เกณฑ์การแจกแจงคะแนนสังเกตได้ (observed score distribution: OSD) ซึ่งมีฐานแนวคิดมาจากความแตกต่างระหว่างค่าที่ประมาณได้กับค่าจริง ผลการศึกษาพบว่า ผลการแปลงสเกลแบบพร้อมกันให้ผลที่น่าพอใจมากกว่าการแปลงสเกลแบบแยกกันกับวิธีการเชื่อมโยงหลาย ๆ วิธี ในรูปของความถูกต้องในการเชื่อมโยงคะแนนและความแกร่งในการนำไปใช้กับข้อมูลพหุมิติจากรูปแบบของแบบสอบ ในกลุ่มของวิธีการเชื่อมโยงคะแนน วิธีโค้งคุณลักษณะให้ผลที่มีความถูกต้องในการเชื่อมโยงมากกว่าวิธีการเชื่อมโยงแบบจุด โดยไม่คำนึงถึงระดับของความเป็นพหุมิติ อย่างไรก็ตามมีความแตกต่างเพียงเล็กน้อยในการเชื่อมโยงคะแนน ระหว่างวิธีการปรับสเกลพร้อมกันกับวิธีการปรับสเกลแยกกันด้วยวิธีโค้งคุณลักษณะ

Kim, Walker และ McHale (2007) ได้ศึกษาการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมในการประเมินระดับมหภาค ออกแบบเก็บรวบรวมข้อมูลโดยใช้กลุ่มไม่เท่าเทียมใช้ข้อสอบร่วม ซึ่งข้อสอบร่วมชนิดข้อสอบแบบเขียนตอบที่มีการให้คะแนนใหม่จากการศึกษาวิจัยของ Tate (1999; 2000) ในบริบทของการปรับเทียบตามทฤษฎีแบบดั้งเดิม การวิจัยครั้งนี้มีแบบแผนการเชื่อมโยง 4 รูปแบบคือ 1) ใช้เพียงข้อสอบที่ตรวจให้คะแนนสองค่าเป็นข้อสอบร่วม 2) ใช้ทั้งข้อสอบที่ให้คะแนนสองค่าและหลายค่าเป็นแบบสอบร่วม 3) ใช้ทั้งข้อสอบที่ให้คะแนนสองค่าและหลายค่าที่มีการให้คะแนนใหม่เป็นแบบสอบร่วม 4) การผสมผสานระหว่างการรวมการออกแบบกลุ่มเดียวกับกลุ่มไม่เท่าเทียม ผลการวิจัยพบว่า รูปแบบการเชื่อมโยงคะแนนที่มีข้อสอบที่ตรวจให้คะแนนสองค่าเป็นข้อสอบร่วมเพียงอย่างเดียว (รูปแบบที่ 1) หรือใช้ทั้งข้อสอบที่ให้คะแนนสองค่าและหลายค่าที่ไม่ได้ให้คะแนนใหม่เป็นแบบสอบร่วม (รูปแบบที่ 2) ทำให้เกิดความลำเอียงอย่างมาก และมีความคลาดเคลื่อนในการปรับเทียบมากกว่าอีก 2 รูปแบบที่เหลือเพียงเล็กน้อย การออกแบบในรูปแบบที่ 4 เกิดความลำเอียงน้อยที่สุด แต่เกิดความคลาดเคลื่อนในการปรับเทียบคะแนนมากกว่ารูปแบบที่ 3 เพียงเล็กน้อย ในภาพรวมแล้วรูปแบบที่ 4 ให้ค่า RMSE ต่ำที่สุด จึงสรุปได้ว่าการออกแบบแบบลูกผสมที่ไม่มีการออกแบบข้อสอบร่วมสามารถนำผลมาปรับปรุงกับวิธีการในปัจจุบันสำหรับการปรับเทียบแบบสอบที่มีการบรรจุ CR อยู่ด้วย

**บทสรุปงานวิจัยที่เกี่ยวข้องกับการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม**

จากการทบทวนเอกสารและงานวิจัยที่กล่าวมาข้างต้น ผู้วิจัยได้สรุปการศึกษาวิจัยการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมมีรายละเอียด ดังนี้

**กระบวนการปรับเทียบคะแนน**

การวิจัยครั้งนี้พิจารณาเฉพาะกระบวนการปรับเทียบคะแนนตามแนวคิดของ IRT เท่านั้น โดยปกติกระบวนการปรับเทียบคะแนนประกอบไปด้วย 3 ขั้นตอน คือ ขั้นตอนแรก เป็นการประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลที่เหมาะสมหรือโมเดลที่เลือกใช้ ขั้นที่ 2 เป็นการแปลงสเกลเพื่อวางตำแหน่งพารามิเตอร์ที่ถูกประมาณค่าใหม่จากแบบสอบฉบับหนึ่งไปสู่แบบสอบฉบับฐาน (Baseline) และขั้นที่ 3 เป็นการรายงานค่าคะแนนดิบที่แปลี่ยนแปลงไป แต่ในการศึกษาครั้งนี้ทำการศึกษาเพียง 2 ขั้นตอนแรกเท่านั้น

ในขั้นตอนของการประมาณค่าพารามิเตอร์ของข้อสอบ ดังที่กล่าวมาแล้วว่าโมเดลแบบ 3PL ใช้ในการประมาณค่าพารามิเตอร์ของข้อสอบที่มีการตรวจให้คะแนนสองค่า และโมเดล GPCM ใช้ในการประมาณค่าพารามิเตอร์ของข้อสอบที่มีการตรวจให้คะแนนหลายค่า ในทาง

ทฤษฎี พารามิเตอร์ของข้อสอบของแบบสอบที่มีรูปแบบผสมสามารถประมาณค่าแยกกันตามรูปแบบของข้อสอบหรือโมเดล (a.k.a., format-wise calibration) หรือประมาณค่าพร้อมกัน ๆ ข้ามรูปแบบของแบบสอบ (a.k.a., simultaneous calibration) เมื่อจำเป็นต้องมีสเกลคะแนนร่วมและต้องรายงานคะแนนรวม การประมาณค่าแบบพร้อมกันจะเหมาะสมกว่าการประมาณค่าแบบแยกกัน เนื่องจากได้เตรียมแนวทางที่เหมาะสมทางสถิติในการแก้ปัญหาเกี่ยวกับการให้น้ำหนักการประมาณค่าแบบแยกกัน ดังนั้น รูปแบบของข้อสอบที่แตกต่างกันสามารถวางบนสเกลเดียวกันได้และประสิทธิภาพของข้อสอบต่างรูปแบบจะสามารถเปรียบเทียบกันได้โดยตรง (Ercikan et al., 1998; Sykes and Yen, 2000)

ในขั้นตอนของการแปลงสเกล มีกระบวนการที่เป็นทางเลือก 2 กระบวนการ คือ การปรับค่าพารามิเตอร์แยกกัน (Separate calibration) และการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent calibration) ทั้งสองกระบวนการนี้สามารถประยุกต์ใช้สำหรับการแทนที่พารามิเตอร์ที่ประมาณค่าตามทฤษฎีการตอบสนองข้อสอบบนสเกลร่วม ในการปรับค่าพารามิเตอร์แยกกัน (Separate calibration) สำหรับแบบสอบรูปแบบผสม 2 ฉบับ ที่ประมาณค่าแยกกันภายใต้การวิเคราะห์โดยคอมพิวเตอร์ 2 ครั้ง วิธีการแปลงสเกลนี้เป็นส่วนที่สามารถแปลงค่าประมาณพารามิเตอร์ย้อนกลับจากแบบสอบฉบับหนึ่งไปยังสเกลของแบบสอบฉบับอื่นตลอดจนชุดของข้อสอบร่วม การแปลงคะแนนทั่ว ๆ ไป ที่ใช้กับโมเดลการตอบข้อสอบแบบสองค่ามี 2 วิธีที่เป็นการแปลงคะแนนแบบจุด คือ mean/mean (Loyd และ Hoover, 1980) และวิธี Mean/sigma (Macro, 1977) ส่วนวิธีการแปลงสเกลด้วยการใช้โค้งคุณลักษณะมี 2 วิธี คือ Heabara (Heabara, 1980) และ Stocking – Lord (Stockinh และ Lord, 1983) และได้มีการขยายแนวคิดการแปลงคะแนนไปสู่โมเดล IRT ที่มีการตรวจให้คะแนนแบบหลายค่า (Baker, 1992, 1993; Cohen and Kim, 1998; Kim and Cohen, 1995) ต่อมา Kim และ Lee (2006) ได้ขยายแนวคิดการแปลงคะแนนทั้ง 4 แบบนี้ไปสู่แบบสอบบรูปแบบผสมด้วย โดยเป็นการรวมกันระหว่างโมเดลที่มีการตรวจให้คะแนน 2 ค่า แบบ 3PL กับโมเดลที่มีการตรวจให้คะแนนหลายค่าแบบ GRM, GPCM, NRM และ MC ส่วนการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent calibration) จะเป็นการพารามิเตอร์จากแบบสอบทั้ง 2 ฉบับพร้อมกัน ในการวิเคราะห์โดยคอมพิวเตอร์เพียง 1 ครั้ง ซึ่งเป็นส่วนที่รับประกันได้ว่าค่าประมาณพารามิเตอร์จะอยู่บนสเกลเดียวกัน นั่นคือการรวมข้อมูลการตอบจากผู้สอบทั้ง 2 กลุ่มและกำหนดให้ข้อสอบที่อีกกลุ่มหนึ่งไม่ได้ทำเป็นค่า “not reach” หรือค่าขาดหาย (missing)

ข้อพิจารณาที่น่าสนใจเกี่ยวกับประสิทธิภาพสัมพัทธ์ของการปรับค่าพารามิเตอร์พร้อมกันและแบบแยกกันกับวิธีการแปลงสเกลที่แตกต่างกัน โดยทั่วไป การปรับค่าพารามิเตอร์พร้อมกันจะให้ผลที่น่าพอใจกว่าการปรับค่าพารามิเตอร์แยกกัน ภายใต้เงื่อนไขที่ศึกษาหลาย ๆ เงื่อนไข เนื่องมาจากมีความเชื่อเกี่ยวกับการปรับค่าพารามิเตอร์พร้อมกันว่าสามารถทำให้ได้สารสนเทศที่เป็นประโยชน์อย่างสมบูรณ์ และอาจจะขจัดความคลาดเคลื่อนบางอย่างที่เกิดขึ้นในการปรับเทียบได้อย่างมีศักยภาพ ซึ่งความไม่ถูกต้องในกระบวนการแปลงสเกลอาจจะเกิดขึ้นในการปรับค่าพารามิเตอร์แยกกัน Kim และ Cohen (1998) ใช้กระบวนการจำลองข้อมูลในข้อมูลที่มีการให้คะแนนแบบสองค่า สำหรับเปรียบเทียบการปรับค่าพารามิเตอร์แยกกันกับวิธี Stocking – Lord และการปรับค่าพารามิเตอร์พร้อมกัน โดยใช้ซอฟต์แวร์ที่ต่างกัน ผลการศึกษาพบว่า ชุดข้อสอบร่วมที่มีจำนวนข้อน้อยๆ ใช้วิธีการปรับค่าพารามิเตอร์พร้อมกันจะให้ผลที่ถูกต้องมากกว่าวิธีการปรับค่าพารามิเตอร์แยกกัน กับวิธี Stocking – Lord แต่ถ้าจำนวนข้อสอบร่วมมีขนาดใหญ่วิธีการปรับค่าพารามิเตอร์แยกกันและแบบพร้อมกันจะให้ผลของการแปลงสเกลคล้ายคลึงกัน Hanson และ Beguin (2002) ได้จำลองข้อมูลการตอบแบบ 2 ค่าสำหรับเปรียบเทียบประสิทธิภาพสัมพัทธ์ของการปรับค่าพารามิเตอร์แยกกัน (mean/mean, mean/sigma, Heabara, และ Stocking – Lord) กับวิธีการปรับค่าพารามิเตอร์พร้อมกัน ผลการศึกษาพบว่า โดยทั่วไปวิธีการปรับค่าพารามิเตอร์พร้อมกันให้ความคลาดเคลื่อนต่ำกว่าวิธีการปรับค่าพารามิเตอร์แยกกัน แม้ว่าจะไม่ใช้ผลที่ครอบคลุมทั้งหมดก็ตาม เนื่องจากในการปรับค่าพารามิเตอร์แยกกันด้วยวิธีโค้งคุณลักษณะ 2 วิธี ให้ผลการแปลงสเกลที่ถูกต้องมากกว่าวิธี mean/mean และ mean/sigma Kim และ Cohen (2002) ได้จำลองข้อมูลผลการตอบที่มีการให้คะแนนแบบหลายค่าภายใต้โมเดล GRM และเปรียบเทียบการแปลงเกลแบบแยกกันกับวิธี Stocking – Lord และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ผลการศึกษาพบว่าวิธีการปรับค่าพารามิเตอร์พร้อมกันให้ผลที่คงที่ และให้ที่มีความถูกต้องมากกว่าวิธีการปรับค่าพารามิเตอร์แยกกันเพียงเล็กน้อย

จากที่กล่าวมาข้างต้น เป็นการศึกษาการแปลงสเกลกับโมเดลการตอบข้อสอบที่มีความเป็นเอกมิติ และเป็นข้อมูลที่จำลองขึ้นเพื่อการศึกษา Beguin, Hanson, และ Glas (2000) และ Beguin และ Hanson (2001) ได้เสนอการจำลองข้อมูลแบบพหุมิติซึ่งไม่สอดคล้องกับโมเดลการตอบข้อสอบแบบสองค่าที่เป็นเอกมิติ และทำการเปรียบเทียบการปรับค่าพารามิเตอร์แยกกัน กับวิธี Stocking – Lord และวิธี การปรับค่าพารามิเตอร์พร้อมกัน ผลการศึกษาพบว่า ความเป็นพหุมิติมีผลต่อประสิทธิภาพสัมพัทธ์ของการแปลงคะแนนแบบแยกกันและแบบพร้อมกัน โดยค่าประมาณจากโมเดลแบบพหุมิติที่ระบุอย่างถูกต้องผลที่ได้โดยทั่วไปจะมีความคลาดเคลื่อนน้อย

กว่าโมเดลแบบเอกมิติ ปกติการปรับค่าพารามิเตอร์พร้อมกันสำหรับโมเดลที่เป็นเอกมิติแล้ว ผลที่ได้จะเกิดความคลาดเคลื่อนโดยรวมน้อยกว่าหรือเท่ากับวิธีการปรับค่าพารามิเตอร์แยกกัน Kim และ Kolen (2006) ได้จำลองข้อมูลผลการตอบที่เป็นพหุมิติที่เป็นอิทธิพลของรูปแบบของข้อสอบในแบบสอบรูปแบบผสม และได้เปรียบเทียบประสิทธิภาพเชิงสัมพัทธ์วิธีการปรับค่าพารามิเตอร์พร้อมกันกับวิธีการแปลงสเกลแบบแยกกัน 4 วิธี ภายใต้โมเดลที่เป็นเอกมิติ ผลการศึกษาพบว่าการปรับค่าพารามิเตอร์พร้อมกันโดยทั่วไปให้ผลน่าพอใจกว่าการปรับค่าพารามิเตอร์แยกกันในรูปของความถูกต้องในการเชื่อมโยง และความแกร่งของการปรับเทียบไปสู่ข้อมูลที่เป็นพหุมิติ

การศึกษาวิธีการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมมีผู้ศึกษาทั้งการปรับเทียบตามแนวทฤษฎีดั้งเดิม และการปรับเทียบตามแนวทฤษฎีตอบสนองข้อสอบ ได้แก่ Bastari (2000) ศึกษาเปรียบเทียบวิธีการปรับเทียบคะแนนตามแนวทฤฎีตอบสนองข้อสอบระหว่างวิธี โค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน โดยศึกษาในเงื่อนไขความยาวของแบบสอบ 2 ระดับคือ 30 ข้อ และ 50 ข้อ โดยสัดส่วนของข้อสอบ MC และ CR มี 2 สัดส่วนคือ 80 : 20 และ 90 : 10 ตามลำดับ ข้อสอบร่วมมี 2 ระดับ คือ 10% และ 20% ของจำนวนข้อสอบทั้งหมด ขนาดกลุ่มตัวอย่าง 400 1000 และ 3000 คน และการแจกแจงของกลุ่มตัวอย่าง ผลการวิจัยพบว่าวิธีปรับค่าพารามิเตอร์พร้อมกันมีประสิทธิภาพมากกว่าวิธีโค้งคุณลักษณะ เมื่อแบบสอบมีความยาวมากกว่า (RMSE ของข้อสอบ 50 ข้อ ต่ำกว่า RMSE ของข้อสอบ 30 ข้อ) สัดส่วนข้อสอบที่มีการตรวจให้คะแนนสองค่ามีปริมาณมากขึ้น (RMSE ของสัดส่วน 90 : 20 ต่ำกว่า RMSE ของสัดส่วน 80 : 20) จำนวนข้อสอบร่วมที่เพิ่มขึ้น (RMSE ของข้อสอบร่วม 20% ต่ำกว่า RMSE ของข้อสอบร่วม 10%) ขนาดกลุ่มตัวออย่างขนาดใหญ่ (RMSE ของจำนวนกลุ่มตัวอย่าง 3000 คน ต่ำที่สุด และ RMSE ของกลุ่มตัวอย่างจำวน 400 คน สูงที่สุด) และกลุ่มผู้สอบมีการแจกแจงความสามารถที่ใกล้เคียงกัน ซึ่งสอดคล้องกับงานวิจัยของ Tate (2000) ได้ศึกษาวิธีการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบด้วยวิธี moment 2 วิธี และวิธีโค้งคุณลักษณะตามแนวของ Stocking และ Lord พบว่าจำนวนข้อสอบร่วมที่มีปริมาณมากขึ้นจะทำให้สัมประสิทธิ์การปรับเทียบคะแนนมีความแม่นยำยิ่งขึ้น ต่อมา Lee (2001) ได้ศึกษาการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมโดยใช้การปรับเทียบตามทฤษฎีแบบดั้งเดิม ในบริบทของแบบสอบชุดย่อย (Testlet) ผลการวิจัยพบว่า หากแบบสอบมีลักษณะเป็นแบบสอบชุดย่อย ควรมีการกำหนดการให้คะแนนเป็นแบบหลายค่าจะให้ผลการปรับเทียบคะแนนที่น่าพอใจกว่าการกำหนดคะแนนเป็นสองค่าแล้วจัด Block ให้กับแบบสอบชุดย่อย

นอกจากนี้งานวิจัยของ Kim (2004) และ Kim และ Lee (2006) ที่ได้ศึกษาเปรียบเทียบการเทียบคะแนนตามแนวทฤษฎีการตอบสนองข้อสอบด้วยวิธี moment 2 วิธี และวิธีโค้ง

คุณลักษณะ 2 วิธี ซึ่งสอดคล้องกับงานวิจัยของ Tate (2000) ซึ่งพบว่าวิธี โค้งคุณลักษณะ ให้ผลการปรับเทียบคะแนนเป็นที่น่าพอใจกว่าวิธี moment ทั้งสองวิธี โดยเงื่อนไขของกลุ่มตัวอย่างจำนวน 500 และ 3000 คน พบว่าจำนวนกลุ่มตัวอย่าง 3000 คน จะทำให้ผลการปรับเทียบมีความแม่นยำมากกว่ากลุ่มตัวอย่าง 5000 คน การปรับเทียบในกลุ่มเท่าเทียมจะเกิดความคลาดเคลื่อนต่ำกว่ากลุ่มไม่เท่าเทียม สำหรับการปรับเทียบโดยใช้ข้อสอบ DS และ PS เป็นข้อสอบร่วม จะมีค่าความคลาดเคลื่อนสูงขึ้นเมื่อจำนวนข้อสอบ DS เพิ่มขึ้นเมื่อจำนวนข้อสอบ DS เพิ่มขึ้น ซึ่งสอดคล้องกับงานวิจัยของ Basteri (2000) นอกจากนี้ยังพบอีกว่าหากนำแบบสอบที่มีการตรวจให้คะแนนสองค่าเป็นข้อสอบร่วมเพียงอย่างเดียวจะทำให้เกิดความคลาดเคลื่อนเพิ่มมากขึ้นด้วย

การพิจารณาในมิติของการเปรียบเทียบกระบวนการปรับเทียบหลังจากที่แปลงสเกลแยกกัน สำหรับโมเดลการตอบที่มีการให้คะแนนสองค่า งานวิจัยทั่ว ๆ ไป พบว่า วิธีการ Heabara และ Stocking – Lord ให้ผลที่คงที่มากกว่าวิธี mean/mean และ mean/sigma (Baker and Al-Karni, 1991; Kim and Cohen, 1992; Ogasawara, 2001; Hanson and Beguin, 2002) ต่อมา Ogasawara (2002) ได้สนับสนุนข้อค้นพบนี้ โดยการแสดงตัวอย่างเกี่ยวกับตัวเลขค่าประมาณฟังก์ชันการตอบข้อสอบ/แบบสอบที่มีความคงที่มากกว่าแม้ว่าค่าประมาณพารามิเตอร์ของข้อสอบจะไม่คงที่ก็ตาม Kaskowwitz และ De Ayala (2001) พบว่า กระบวนการโค้งคุณลักษณะข้อสอบมีความแกร่งมากที่สุด

การปรับเทียบกลุ่มผู้สอบที่ไม่เท่าเทียมกันโดยใช้ข้อสอบร่วม นั่นคือ ชุดพารามิเตอร์ของข้อสอบ 2 ชุด จากแบบสอบฉบับเก่าและฉบับใหม่สามารถวางบนสเกลร่วมกันโดยใช้การปรับค่าพารามิเตอร์พร้อมกัน (Concurrent calibration) ซึ่งโปรแกรมที่สามารถใช้ในการปรับสเกลพร้อมกันได้ คือ BILOG-MG (Zimowski, Muraki, Mislevy, and Bock, 1996), ICL (Hanson, 2002) และ MULTILOG (Thissen, 1991) โปรแกรมเหล่านี้มีคุณสมบัติที่สามารถใช้ประมาณค่ากับกลุ่มผู้สอบที่มีระดับความสามารถแตกต่างกันได้ งานวิจัยที่เปรียบเทียบการปรับค่าพารามิเตอร์พร้อมกัน กับการเชื่อมโยงแบบอื่นๆ โดยทั่วไปพบว่า การปรับสเกลพร้อมกันมีความถูกต้องมากกว่าวิธี การปรับเทียบแบบโค้งคุณลักษณะเมื่อศึกษาจากข้อมูลที่จำลองขึ้นโดยให้ข้อมูลสอดคล้องกับโมเดล (Kim and Cohen, 1998, 2002; Hanson and Beguin, 2002) อย่างไรก็ตาม Beguin, Hanson และ Glas (2000) พบว่า เมื่อกลุ่มผู้สอบ 2 กลุ่มมีระดับความสามารถไม่เท่าเทียมกัน และความ สามารถเป็นแบบพหุมิติแต่มีความสัมพันธ์กันสูง การเชื่อมโยงดยใช้วิธี Stocking – Lord จะทำให้เกิดความถูกต้องมากกว่าวิธีการปรับสเกลพร้อมกัน ข้อค้นพบเหล่านี้ว่าเสนอแนะให้เห็นว่า การปรับสเกลพร้อมกันอาจจะมีความแกร่งน้อยเมื่อมีการฝ่าฝืนข้อตกลงเบื้องต้นเมื่อเทียบกับวิธีโค้งคุณลักษณะที่

เป็นการปรับสเกลแยกกัน Kolen และ Brennan (in preparation) เสนอแนะว่า การใช้วิธีการเชื่อมโยงแบบคุณลักษณะ ในทางปฏิบัติแล้วดูเหมือนว่าจะปลอดภัยที่สุดและวิธีการปรับสเกลพร้อมกันสามารถใช้เหมือนกับเป็นส่วนเสริมของการปรับค่าพารามิเตอร์แยกกัน

**โมเดลในแบบสอบรูปแบบผสม และสัดส่วนของข้อสอบทีตรวจให้คะแนนสองค่าและตรวจให้คะแนนหลายค่า**

สำหรับแบบสอบรูปแบบผสมเป็นการรวมกันระหว่างข้อสอบแบบหลายตัวเลือกกับข้อสอบแบบเขียนตอบบรรจุในแบบสอบฉบับเดียว ถ้าจะพิจารณาตามโมเดลการตอบข้อสอบโดยอาศัยพื้นฐานของทฤษฎีการตอบสนองข้อสอบ ข้อสอบแบบหลายตัวเลือกก็คือข้อสอบที่มีการตรวจให้คะแนนแบบสองค่า ส่วนข้อสอบแบบหลายตัวเลือกก็คือข้อสอบที่มีการตรวจให้คะแนนมากกว่าสองค่า ซึ่งโมเดลการตอบข้อสอบทั้ง 2 นี้จะมีการประมาณค่าพารามิเตอร์ตามแนวคิดของทฤษฎีการตอบสนองข้อสอบ สำหรับข้อสอบที่มีการตรวจให้คะแนนแบบสองค่าสามารถประยุกต์ใช้โมเดลตามทฤษฎีการตอบสนองข้อสอบได้หลายโมเดล เช่น โมเดลโลจีสติกแบบ 1, 2 และ 3 พารามิเตอร์ (1PL, 2PL และ 3PL) ส่วนข้อสอบที่มีการตรวจให้คะแนนแบบหลายค่ามีการประยุกต์ใช้โมเดลหลายโมเดล เช่น Graded response model (GRM), Generalized partial credit model (GPCM) และ Nominal response model (NRM) เป็นต้น

Baker และ Kim (2004) ชี้ให้เห็นว่า การรวมกันของโมเดลที่มีการตรวจให้คะแนนสองค่าและหลายค่าตามแนวคิดการตอบสนองข้อสอบ จะมีการวิเคราะห์ผลการตอบข้อสอบในแบบสอบรูปแบบผสม ในการวิจัยครั้งนี้ โมเดลที่ใช้คือ 3PL/GPCM และประมาณค่าพารามิเตอร์ ซึ่งโมเดลในลักษณะเดียวกันนี้มีการประยุกต์ในการวิจัยหลายเรื่อง

เหตุผลในการเลือกโมเดล 3PL และ GPCM ในการรวมกันเป็นแบบสอบรูปแบบผสม นั่นคือ ข้อสอบแบบหลายตัวเลือกนั้น งานวิจัยที่ผ่านมาแสดงให้เห็นว่าโมเดลการตอบข้อสอบแบบสองค่าที่นิยมใช้คือแบบ 3 พารามิเตอร์ เนื่องจากธรรมชาติของข้อสอบแบบหลายตัวเลือก ผู้สอบสามารถเดาคำตอบจากตัวเลือกได้ และโมเดลการตอบข้อสอบแบบ 1 และ 2 พารามิเตอร์ จะไม่มีการประมาณค่าการเดา ดังนั้น โมเดลที่เหมาะสมและสามารถประยุกต์ใช้ได้คือโมเดลแบบ 3 พารามิเตอร์ และนอกจากนี้ยังมีการประเมินระดับมหภาคที่ประยุกต์ใช้โมเดลการตอบข้อสอบที่มีการตรวจให้คะแนนแบบสองค่าแบบ 3 พารามิเตอร์จำนวนมาก เช่น National Assessment of Educational Progress (NAEP), The Comprehensive Tests of Basic Skills (CTBS), The Armed Services Vocational Aptitude Battery (ASVAB), The Law School Admissions Test (LSAT), The Graduate Management Admissions Test (GMAT), The Scholastic

Assessment Test (SAT), The Graduate Record Exam (GRE), และ The Test of English as a Foreign Language (TOEFL) เป็นต้น สำหรับโมเดลที่มีการตรวจให้คะแนนแบบหลายค่ามีโมเดลที่ได้รับความนิยมในการใช้คือ GRM และ GPCM เช่น Dodd (1984) ได้เปรียบเทียบโมเดล GRM และ PCM ในรูปแบบของราชช์โมเดล และพบว่าโมเดลทั้ง 2 มีความแตกต่างในเชิงแนวคิดและฟังก์ชันทางคณิตศาสตร์ แต่มีแนวทางในการประยุกต์ใช้คล้ายคลึงกัน Maydeu-Olivares, Drasgow, และ Mead (1994) ได้เปรียบเทียบโมเดล GRM และ GPCM และพบว่า ทั้งสองโมเดลมีความสอดคล้องของข้อมูลกับโมเดลเป็นอย่างดี และในงานวิจัยอื่น ๆ มีแนวทางคำตอบคล้ายคลึงกัน (เช่น Cao, Yin, and Gao, 2007; Tang and Eignor, 1997)

การวิจัยที่ศึกษาการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมโดยศึกษาเงื่อนไขเกี่ยวกับสัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่า และตรวจให้คะแนนหลายค่า ให้ผลที่คล้ายกันคือ เมื่อจำนวนข้อสอบที่ตรวจให้คะแนนสองค่ามีมากขึ้น จะทำให้ผลการปรับเทียบคะแนนแม่นยำยิ่งขึ้น (Li, Lissitz and Yang, 1999; Bastari, 2002) และยังมีการพิจารณาเกี่ยวกับสัดส่วนของข้อสอบร่วม ซึ่งผลการศึกษาของ Li, Lissitz และ Yang (1999) พบว่าข้อสอบร่วมที่มีทั้งข้อสอบที่ตรวจให้คะแนนสองค่าและหลายค่ารวมอยู่ด้วยกัน จะมีผลการปรับเทียบที่แม่นยำยิ่งขึ้นเมื่อจำนวนข้อสอบที่มีการตรวจให้คะแนนสองค่ามีปริมาณน้อยลง ซึ่งสอดคล้องกับงานวิจัยของ Kim และ Lee (2006) คือเมื่อ ข้อสอบร่วมที่มีข้อสอบที่ตรวจให้คะแนนสองค่าน้อยลงจะทำให้ผลการปรับเทียบแม่นยำยิ่งขึ้น นอกจากนี้ Kim และ Lee (2006) ยังพบอีกว่า การใช้ข้อสอบร่วมเป็นข้อสอบที่ตรวจให้คะแนนสองค่า หรือหลายค่า เพียงรูปแบบใดรูปแบบหนึ่ง จะทำให้ผลการปรับเทียบเกิดความคลาดเคลื่อนเพิ่มมากขึ้นกว่าการใช้ข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า และตรวจให้คะแนนหลายค่า และ Li, Lissitz และ Yang (1999) ได้ข้อเสนอแนะในการวิจัยต่อไปว่า ควรจะมีการพิจารณาเกี่ยวกับสัดส่วนของข้อสอบร่วมมากกว่าสัดส่วนของแบบสอบทั้งฉบับ

**ข้อสอบร่วม**

Petersen, Macro และ Stewart (1982) ได้ตรวจสอบความเป็นตัวแทนของข้อสอบร่วมด้านสถิติและด้านเนื้อหาที่ส่งผลต่อการปรับเทียบเชิงเส้นแบบดั้งเดิม ผลการวิจัยพบว่า ความยากที่แตกต่างกันระหว่างข้อสอบร่วมกับข้อสอบทั้งหมดทำให้เกิดความคลาดเคลื่อนในการปรับเทียบมากกว่าข้อสอบร่วมที่มีความแตกต่างด้านเนื้อหาระดับปานกลางกับแบบสอบทั้งฉบับ ต่อมา Klein และ Jarjour (1985) ได้เปรียบเทียบในทำนองเดียวกันแต่ความยาวของข้อสอบร่วมนั้น

มากกว่า และชุดของข้อสอบร่วมไม่เป็นตัวแทนด้านเนื้อหาของแบบสอบทั้งฉบับ ผลการวิจัยพบว่า ข้อสอบร่วมที่ยาวขึ้นและไม่เป็นตัวแทนด้านเนื้อหาทำให้เกิดความคลาดเคลื่อนในการปรับเทียบคะแนนต่ำกว่าข้อสอบร่วมจำนวนน้อยข้อ นั่นแสดงให้เห็นว่าความเป็นตัวแทนด้านเนื้อหาเป็นสิ่งสำคัญมากต่อความถูกต้องในการปรับเทียบ Cook และ Petrerson (1987) พบว่า ชุดของข้อสอบร่วมที่เป็นตัวแทนด้านเนื้อหาเพียงพอ จะเกิดปัญหาหลาย ๆ ประการ โดยเฉพาะอย่างยิ่งเมื่อกลุ่มผู้สอบ 2 กลุ่ม ที่มีระดับต่างกันและมีการกระจายของระดับความสามารถมากๆ Harris (1991) และ Yang (2000) ได้ศึกษาในลักษณะเดียวกันและให้ข้อสนับสนุนผลการวิจัยนี้ อย่างไรก็ตาม Gao, Hanson และ Harris (1999) ได้เสนอข้อค้นพบบางสิ่งที่แตกต่างออกไป นั่นคือ พวกเขาได้ตรวจสอบความเป็นตัวแทนด้านเนื้อหาและสถิติของข้อสอบรวมกับกลุ่มผู้สอบไม่เท่าเทียมและพบว่าถ้าข้อสอบร่วมไม่เป็นตัวแทน จะไม่มีอิทธิพลมากนักต่อผลการปรับเทียบ แต่ถ้าพิจารณาปฏิสัมพันธ์ระหว่างเนื้อหาและสถิติของข้อสอบร่วมจะส่งผลต่อการปรับเทียบมากกว่า โดยที่ชุดของข้อสอบร่วมที่ไม่เป็นตัวแทนด้านเนื้อหาจะส่งผลต่อความคลาดเคลื่อนในการปรับเทียบสูงกว่าข้อสอบร่วมที่เป็นตัวแทนด้านเนื้อหา ในทางตรงกันข้ามถ้าไม่กล่าวถึงความเป็นตัวแทนของเนื้อหาในชุดของข้อสอบร่วม และพิจารณาเพียงสถิติไม่เป็นตัวแทน จะเกิดความคลาดเคลื่อนในการปรับเทียบมากกว่าข้อสอบร่วมที่ค่าสถิติมีความเป็นตัวแทนของข้อสอบทั้งหมด ต่อมา Hanick และ Huang (2002) ได้สนับสนุนข้อค้นพบของ Gao และคณะ นั่นคือ ชุดข้อสอบร่วมที่มีเนื้อหาไม่เป็นตัวแทนจะมีผลเพียงเล็กน้อยต่อความถูกต้องในการปรับเทียบ ยิ่งไปกว่านั้นยังพบอีกว่าถ้าการปรับเทียบมีการออกแบบที่ดี (well designed) ค่าสถิติที่ไม่เป็นตัวแทนก็ไม่มีผลมากนักต่อผลการปรับเทียบ ดังนั้น จึงไม่มีข้อสรุปที่เป็นสากลเกี่ยวกับคุณลักษณะของชุดข้อสอบร่วมที่มีผลต่อการปรับเทียบ แต่ก็ยังมีความคิดเห็นที่ได้รับการยอมรับอย่างกว้างขวางเกี่ยวกับคุณลักษณะของชุดข้อสอบร่วมที่มีบทบาทที่สำคัญต่อการปรับเทียบคะแนนแบบดั้งเดิมกับแบบสอบแบบเลือกตอบ ดังนั้น ชุดของข้อสอบร่วมควรจะมีเป็นลักษณะการลดรูป (mini version) ของแบบสอบทั้งฉบับ โดยเฉพาะอย่างยิ่งเมื่อกลุ่มผู้สอบมีระดับความสามารถที่แตกต่างกัน

ส่วนประกอบของชุดข้อสอบร่วม มีส่วนสำคัญมากต่อการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมภายใต้การเก็บรวบรวมข้อมูลกับกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม ซึ่งต้องพิจารณาความเป็นตัวแทนด้านเนื้อหาและสถิติของข้อสอบร่วม ในการปรับเทียบสำหรับแบบสอบรูปแบบผสมจะกล่าวถึงอิทธิพลของรูปแบบข้อสอบด้วย ซึ่งเป็นความซับซ้อน ยุ่งยากในการพัฒนาชุดของข้อสอบร่วมให้มีความเหมาะสมสำหรับแบบสอบรูปแบบผสม ในทางปฏิบัติแล้ว มีนักปฏิบัติหลายคนแนะนำว่า ให้ใช้เพียงข้อสอบแบบเลือกตอบเป็นข้อสอบร่วม (Livingston, 1994;

Baghi, DeLain, and Hennings, 1995) การใช้ข้อสอบแบบหลายตัวเลือกเพียงอย่างเดียวเป็นข้อสอบร่วม สามารถทำได้เมื่อข้อสอบแบบหลายตัวเลือกและข้อสอบแบบเขียนตอบวัดในโครงสร้างเนื้อหาเดียวกัน แต่ถ้าข้อสอบแบบหลายนตัวเลือกและข้อสอบแบบเขียนตอบมีบางอย่างที่วัดโครงสร้างแตกต่างกัน แล้วใช้เพียงข้อสอบแบบหลายตัวเลือกเป็นข้อสอบร่วมจะเป็นสิ่งที่อันตรายเป็นอย่างยิ่ง และจะนำไปสู่ความลำเอียงในการเชื่อมโยง (linking bias) เป็นอย่างมาก เพราะฉะนั้นจึงมีนักวิจัยหลายท่าน เสนอแนะว่าควรใช้ทั้งข้อสอบแบบเลือกหลายตัวเลือกและแบบเขียนตอบเป็นข้อสอบร่วมในสัดส่วนที่เหมาะสม เพื่อให้การปรับเทียบมีความแกร่งเพียงพอต่อการหลีกเลี่ยงข้อตกลงเบื้องต้น (Tate, 2000; Sykes et al., 2002; Kirkpatrick, 2005; Kim and Kolen, 2006;)

**ขนาดกลุ่มตัวอย่าง** ผลการวิจัยส่วนใหญ่พบว่าจำนวนกลุ่มตัวอย่างที่เพิ่มปริมาณมากขึ้นจะทำให้ผลการปรับเทียบคะแนนมีความแม่นยำยิ่งขึ้นตามไปด้วย (Li, Lissitz and Yang, 1999; Bastari, 2000; Kim, 2004; Kim and Lee, 2006)

**การเก็บรวบรวมข้อมูล** การวิจัยส่วนใหญ่ศึกษาเปรียบเทียบการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมโดยใช้การรวบรวมข้อมูลเพียงรูปแบบใดรูปแบบหนึ่ง (Tate 2000, Lee, 2001; Bastari, 2002) หรือหากเป็นการเก็บข้อมูลหลายรูปแบบก็จะใช้วิธีการปรับเทียบคะแนนเพียง 1 หรือ 2 วิธี (Li, Lissitz and Yang, 1999;, Kim, 2004; Kim and Lee, 2006) ซึ่งผลการ วิจัยยังไม่ครอบคลุม ดังนั้นสำหรับการวิจัยครั้งนี้ ผู้วิจัยจึงศึกษาการเก็บรวบรวมข้อมูล 2 รูปแบบคือกลุ่มสมมูลและกลุ่มไม่เท่าเทียม โดยใช้วิธีการปรับเทียบที่ได้รับการสนับสนุนจากผลการวิจัยที่ผ่านมาว่าเป็นวิธีการที่มีคุณภาพคือ วิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน

การออกแบบการเก็บรวบรวมข้อมูลที่นิยมใช้ในการปรับเทียบและการ Scaling มี 3 รูปแบบ คือ แบบกลุ่มเดียว (SG) แบบกลุ่มสุ่มหรือกลุ่มเท่าเทียม (RG) และแบบกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม (CINEG) (Kolen และ Brennan, 2004) ในการออกแบบการเก็บข้อมูลแบบกลุ่มเดี่ยวผู้สอบจะได้ทำแบบสอบทั้งฉบับ X และ Y โดยปกติจะมีการจัดให้สมดุลด้วย การออกแบบแบบกลุ่มเท่าเทียม จะสุ่มผู้สอบให้ทำแบบสอบฉบับ X หรือ Y เท่านั้น ส่วนการออกแบบการเก็บข้อมูลกับกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วมนั้น โดยปกติกลุ่มผู้สอบ 2 กลุ่มจะมีความสามารถไม่เท่าเทียมกัน กลุ่มหนึ่งทำแบบสอบฉบับ X และอีกกลุ่มหนึ่งทำแบบสอบฉบับ Y โดยที่แบบสอบฉบับ X และ Y จะมีข้อสอบบางข้อที่เหมือนกัน ซึ่งเรียกว่าข้อสอบร่วม ถ้าคะแนนจากการทำข้อสอบร่วมนับรวมกับคะแนนจากการทำข้อสอบข้ออื่น ๆ ทั้งหมดในแบบสอบแต่ละฉบับนั้น จะ เป็นลักษณะของข้อสอบ

ร่วมภายใน แต่ถ้าหากคะแนนของข้อสอบร่วมนับแยกส่วนออกจากแบบสอบจะเป็นลักษณะของข้อสอบร่วมภายนอก การออกแบบการเก็บข้อมูลแบบกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม มีความยืดหยุ่นมากกว่าการออกแบบแบบกลุ่มเท่าเทียมคือยอมให้ผู้สอบที่มีความสามารถไม่เท่ากันทำแบบสอบฉบับ X และ Y ได้ ในขณะที่การออกแบบกลุ่มเท่าเทียมนั้นดีกว่าการออกแบบแบบกลุ่มเดี่ยวคือไม่ต้องการให้ผู้สอบทำแบบสอบทั้ง 2 ฉบับ จากที่กล่าวมาจะเห็นได้ว่าการออกแบบการเก็บข้อมูลแบบกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วมจะมีความยืดหยุ่นมากที่สุด

โดยปกติแล้วแบบทดสอบจะมีลักษณะเป็นแบบเลือกตัวชนิดหลายตัวเลือก ขั้นตอนที่สำคัญในการเก็บรวบรวมข้อมูลกับกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม คือ การการกำหนดสัดส่วนชุดของข้อสอบร่วมที่เป็นตัวแทนแบบสอบทั้งฉบับได้โดยพิจารณาทั้งด้านสถิติและเนื้อหา (Kolen and Brennan, 2004) นักวิจัยได้ตรวจสอบผลจากคุณลักษณะของข้อสอบร่วมในหลาย ๆ ลักษณะ เช่น ความยาวของข้อสอบร่วม ความเป็นตัวแทนด้านเนื้อหาและสถิติของชุดข้อสอบร่วมที่ส่งผลต่อความถูกต้องของการปรับเทียบ และส่วนประกอบของข้อสอบร่วมที่มีผลต่อการปรับเทียบ

จากข้อสรุปข้างต้น ผู้วิจัยได้แสดงภาพรวมในการศึกษาวิจัยการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม ดังตารางที่ 3

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3 สรุปประเด็นสำคัญของการเทียบคะแนนในแบบสอบรูปแบบผสม

| ผู้วิจัย/ ปีที่ศึกษา | โมเดลที่ใช้ | วิธีการปรับเทียบ | การออกแบบการเก็บข้อมูล | สถานการณ์ที่ศึกษา | วิธีการประเมิน | ผลที่ได้จากการศึกษา |
|---|---|---|---|---|---|---|
| Li, Lissitz and Yang (1999) | DS : 3PL<br>PS : GPCM | โค้งลักษณะข้อสอบ | common – item | 1. สัดส่วนของข้อสอบแบบ DS และ PS ในข้อสอบร่วม ระดับ คือ DS : 10, 15, 20 และ PS : 5 ข้อ ข้อละ 4 รายการคำตอบ<br>2. ขนาดของหน่วยตัวอย่าง 3 ระดับ 1000, 2000, 3000 | Bias<br>RMSE | - ในภาพรวมวิธีการปรับเทียบแบบโค้งคุณลักษณะมีความคลาดเคลื่อนต่ำกว่าวิธีโมเมนต์ทุกเงื่อนไขที่ศึกษา<br>- ผลการปรับเทียบมีความคลาดเคลื่อนต่ำลงเมื่อข้อสอบ DS ในข้อสอบร่วมมีจำนวนเพิ่มขึ้น และจำนวนกลุ่มตัวอย่างเพิ่มขึ้น |
| Basteri (2000) | DS : 3PL<br>PS : GRM | - โค้งลักษณะข้อสอบ<br>- การปรับค่าพารามิเตอร์พร้อมกัน | common – item | 1. ความยาวของแบบสอบ 30 ข้อ 50 ข้อ<br>2. สัดส่วนของข้อสอบแบบ DS: 80%, 90% และ PS 20%, 10%<br>3. ความยาวของข้อสอบร่วม 10%, 20%<br>4. ขนาดของหน่วย 400, 1000, 3000<br>5. การแจกแจงความสามารถแบบโค้งปกติ 2 ระดับ SD = 0,1 | MSD<br>RMSE<br>RMSD<br>ANOVA | - ผลการปรับเทียบมีความคลาดเคลื่อนต่ำลงเมื่อ แบบสอบยาวขึ้น สัดส่วนของข้อสอบ DS : PS ในแบบสอบมีข้อสอบ DS เพิ่มขึ้น ข้อสอบร่วมมีความยาวมากขึ้น กลุ่มตัวอย่างขนาดใหญ่ขึ้น กลุ่มผู้สอบสองกลุ่มมีความสามารถเท่าเทียมกัน |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3 สรุปประเด็นสำคัญของการเทียบคะแนนในแบบสอบรูปแบบผสม

| ผู้วิจัย/ ปีที่ศึกษา | โมเดลที่ใช้ | วิธีการปรับเทียบ | การออกแบบการเก็บข้อมูล | สถานการณ์ที่ศึกษา | วิธีการประเมิน | ผลที่ได้จากการศึกษา |
|---|---|---|---|---|---|---|
| Tate (2000) | DS : 3PL<br>PS : GRM | - mean/sigma<br>- Stocking - Lord | Common item<br>non – equivalent group | 1. การผ่าผืนข้อตกลงเบื้องต้นของความเป็นเอกมิติ<br>2. แหล่งความคลาดเคลื่อน<br>3. สัดส่วนของข้อสอบแบบ DS และ PS ในแบบสอบ<br>4. ข้อสอบร่วม : Only MC, Both MC and CR | ความแตกต่างระหว่างความคลาดเคลื่อนจากการปรับเทียบจากปีหนึ่งสู่อีกปีหนึ่ง | - ผลการปรับเทียบมีความคลาดเคลื่อนต่ำลงแบบสอบมีความเป็นเอกมิติ สัดส่วนของข้อสอบ DS ในแบบสอบเพิ่มขึ้น ข้อสอบร่วมที่ประกอบด้วย MC เพียงอย่างเดียว |
| Lee (2001) | DS : 3PL<br>PS : GRM, NRM | ค่าเฉลี่ยเชิงเส้นตรง<br>อีควิปเปอร์เซนต์ไทล์ | Common item<br>equivalent group | 1. การละเลยข้อตกลงเบื้องต้นโดยการใช้แบบสอบย่อย (Testlets)<br>2. สัดส่วนข้อสอบย่อยในแต่ละตอนของแบบสอบ 2 ฉบับ | เปรียบเทียบความแตกต่างระหว่างคะแนนแปลงจากการปรับเทียบแต่ละวิธีโดยใช้ RSM และ URSM | - การใช้แบบสอบย่อยในการปรับเทียบคะแนนให้พิจารณาแบบสอบย่อยเป็นข้อสอบที่ตรวจให้คะแนนแบบหลายค่า เพื่อให้แบบสอบมีความเป็นเอกมิติ |
| Kim (2004) | DS : 3PL<br>PS : GPCM | - mean/mean<br>- mean/sigma<br>- Heabara<br>- Stocking – Lord<br>- Concurrent calibration | - Common item non – equivalent group | 1. อิทธิพลของรูปแบบข้อสอบ (Format Effect)<br>2. ระดับความสัมพันธ์ของข้อสอบ 2 แบบ 2 ระดับ<br>3. ระดับกลุ่มผู้สอบไม่เท่าเทียม 3 ระดับ | MSE (mean square error) จากเกณฑ์โค้งลักษณะรายการคำตอบ | - ในภาพรวมวิธีการปรับเทียบแบบโค้งคุณลักษณะมีความคลาดเคลื่อนต่ำกว่าวิธีโมเมนต์ทุกเงื่อนไขที่ศึกษา<br>- วิธีการปรับเทียบคะแนนแบบการปรับค่าพารามิเตอร์พร้อมกันให้ผลดีกว่าวิธีการปรับเทียบวิธีอื่นๆ |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3 สรุปประเด็นสำคัญของการเทียบคะแนนในแบบสอบรูปแบบผสม (ต่อ)

| ผู้วิจัย/ ปีที่ศึกษา | โมเดลที่ใช้ | วิธีการปรับเทียบ | การออกแบบการเก็บข้อมูล | สถานการณ์ที่ศึกษา | วิธีการประเมิน | ผลที่ได้จากการศึกษา |
|---|---|---|---|---|---|---|
| Robert (2005) | ข้อมูลจริง | - True- score equating<br>- Observe - score equating | - Common item non – equivalent group | 1. อิทธิพลของข้อสอบร่วม<br>2. โมเดลพหุมิติ<br>3. การแปรเปลี่ยนพารามิเตอร์ความสามารถในแต่ละมิติ<br>4. ลักษณะของข้อสอบร่วม | MSE | - การปรับเทียบเกิดความคลาดเคลื่อนต่ำเมื่อกลุ่มผู้สอบเท่าเทียมกัน, ข้อสอบ DS ในข้อสอบร่วมลดลง, การใช้ข้อสอบทั้ง DS และ PS เป็นข้อสอบร่วม<br>- ความแตกต่างของความสามารถผู้สอบมีอิทธิพลต่อการปรับเทียบมากกว่ามิติของแบบสอบที่สัมพันธ์กัน |
| Kim and Lee (2006) | DS : 3PL<br>PS : GPCM | - mean/mean<br>- mean/sigma<br>- Heabara<br>- Stocking – Lord<br>- Concurrent Calibration | - Common item non – equivalent group<br>- Common item equivalent group | 1. ขนาดของกลุ่มตัวอย่าง<br>2. สัดส่วนของ DS และ PS ในแบบสอบ<br>3. จำนวนข้อสอบร่วม | - MSE (mean square error) จากเกณฑ์โค้งลักษณะรายการคำตอบ<br>- square bias | - ในภาพรวมวิธีการปรับเทียบแบบโค้งคุณลักษณะมีความคลาดเคลื่อนต่ำกว่าวิธีโมเมนต์ทุกเงื่อนไขที่ศึกษา<br>- วิธีการปรับเทียบคะแนนแบบการปรับค่าพารามิเตอร์พร้อมกันให้ผลดีกว่าวิธีการปรับเทียบวิธีอื่นๆ<br>- การปรับเทียบเกิดความคลาดเคลื่อนต่ำเมื่อกลุ่มผู้สอบเท่าเทียมกัน, ข้อสอบ DS ในข้อสอบร่วมลดลง, การใช้ข้อสอบทั้ง DS และ PS เป็นข้อสอบร่วม |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3 สรุปประเด็นสำคัญของการเทียบคะแนนในแบบสอบรูปแบบผสม (ต่อ)

| ผู้วิจัย/ ปีที่ศึกษา | โมเดลที่ใช้ | วิธีการปรับเทียบ | การออกแบบการเก็บข้อมูล | สถานการณ์ที่ศึกษา | วิธีการประเมิน | ผลที่ได้จากการศึกษา |
|---|---|---|---|---|---|---|
| Kim และ Kolen (2006) | DS : 3PL<br>PS : GPCM | - mean/mean<br>- mean/sigma<br>- Heabara<br>- Stocking – Lord<br>- Concurrent Calibration | Non-equivalent groups | 1. ความสามารถของผู้สอบที่ต่างกัน 3 ระดับ<br>2. โมเดลพหุมิติ | OSD | - ในภาพรวมการปรับค่าพารามิเตอร์พร้อมกันให้ผลการปรับเทียบที่มีคุณภาพมากกว่าวิธีอื่นๆ และมีความแกร่งต่อโมเดลพหุมิติ<br>- การปรับค่าด้วยวิธีโค้งคุณลักษณะให้ผลการปรับเทียบที่มีคุณภาพมากกว่าการปรับเทียบแบบโมเมนต์ |
| Kim, Walker, McHale (2007) | ข้อมูลจริง | ทฤษฎีดั้งเดิม | - Common item non – equivalent group | 1) ใช้เพียง DS เป็นข้อสอบร่วม<br>2) ใช้ทั้ง DS และ PS เป็นข้อสอบร่วม<br>3) ใช้ทั้ง DS และ PS ที่มีการให้คะแนนใหม่เป็นแบบสอบร่วม<br>4)การออกแบบกลุ่มเดียวกับกลุ่มไม่เท่าเทียม | RMSE | - รูปแบบที่ 1 หรือรูปแบบที่ 2 ทำให้เกิดความลำเอียงอย่างมาก และมีความคลาดเคลื่อนในการปรับเทียบมากกว่าอีก 2 รูปแบบที่เหลือเพียงเล็กน้อย |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3 สรุปประเด็นสำคัญของการเทียบคะแนนในแบบสอบรูปแบบผสม (ต่อ)

| ผู้วิจัย/ ปีที่ศึกษา | โมเดลที่ใช้ | วิธีการปรับเทียบ | การออกแบบการเก็บข้อมูล | สถานการณ์ที่ศึกษา | วิธีการประเมิน | ผลที่ได้จากการศึกษา |
|---|---|---|---|---|---|---|
| Kim, Walker, McHale (2007) (ต่อ) | ข้อมูลจริง | ทฤษฎีดั้งเดิม | - Common item non – equivalent group | 1) ใช้เพียง DS เป็นข้อสอบร่วม<br>2) ใช้ทั้ง DS และ PS เป็นข้อสอบร่วม<br>3) ใช้ทั้ง DS และ PS ที่มีการให้คะแนนใหม่เป็นแบบสอบร่วม<br>4)การออกแบบกลุ่มเดียวกับกลุ่มไม่เท่าเทียม | RMSE | - รูปแบบที่ 4 เกิดความลำเอียงน้อยที่สุด แต่เกิดความคลาดเคลื่อนในการปรับเทียบคะแนนมากกว่ารูปแบบที่ 3 เพียงเล็กน้อย ในภาพรวมแล้วรูปแบบที่ 4 ให้ค่า RMSE ต่ำที่สุด จึงสรุปได้ว่าการออกแบบแบบลูกผสมที่ไม่มีการออกแบบข้อสอบร่วมสามารถนำผลมาปรับปรุงกับวิธีการในปัจจุบันสำหรับการปรับเทียบแบบสอบที่มีการบรรจุ CR อยู่ด้วย |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

การศึกษาการเชื่อมโยงคะแนนโดยใช้เทคนิคการปรับเทียบคะแนนในแบบสอบรูปแบบผสม ส่วนใหญ่เป็นการศึกษาในเชิงของผลที่ได้จากวิธีการปรับเทียบแต่ละวิธี และเปรียบเทียบคุณภาพของวิธีการปรับเทียบ ซึ่งวิธีการปรับเทียบที่ได้รับความนิยมในการนำไปใช้คือ วิธีการปรับเทียบตามแนวทฤษฎีการตอบสนองข้อสอบ คือ วิธีโค้งลักษณะข้อสอบ วิธี mean/sigma วิธี Stocking - Lord วิธีการปรับค่าพารามิเตอร์พร้อมกัน มีเป็นส่วนน้อยที่ทำการศึกษาการปรับเทียบแบบสอบรูปแบบผสมตามแนวทฤษฎีการทดสอบแบบดั้งเดิม

การศึกษาดังกล่าวมุ่งเน้นไปที่ปัจจัยต่าง ๆ ของการปรับเทียบแบบแผนของการเก็บรวบรวมข้อมูล สัดส่วนของข้อสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า กับข้อสอบที่ตรวจให้คะแนนหลายค่าในแบบสอบ ความยาวของแบบสอบ ขนาดหน่วยตัวอย่าง จำนวนของข้อสอบร่วม ซึ่ง Li, Lissitz และ Yang (1999) เสนอแนะว่าควรจะได้มีการศึกษาเงื่อนไขของข้อสอบร่วม ควรพิจารณาการกำหนดน้ำหนักสัดส่วนของรูปแบบข้อสอบที่แตกต่างกัน (MC และ CR) ซึ่งอาจเกิดผลต่อความแม่นยำของสัมประสิทธิ์การปรับเทียบ โดยใช้แบบแผนที่มีเงื่อนไขที่ดีที่สุดของแบบสอบรูปแบบผสม

การศึกษาส่วนใหญ่เป็นการจำลองข้อมูลเพื่อให้ได้ข้อมูลที่สอดคล้องกับโมเดลที่ต้องการศึกษา ซึ่งโมเดลที่นำมาใช้ในแบบสอบรูปแบบผสมระหว่างข้อสอบที่มีการตรวจให้คะแนนแบบ 2 ค่า กับข้อสอบที่มีการตรวจให้คะแนนแบบหลายค่า ปรากฏว่า โมเดลการตอบที่มีการตรวจให้คะแนนแบบสองค่าส่วนใหญ่เป็นโมเดลโลจีสติก IRT แบบ 3 พารามิเตอร์ ส่วนโมเดลการตอบที่มีการตรวจให้คะแนนแบบหลายค่าส่วนใหญ่ใช้โมเดล Graded Response และ Generalize Partial Credit โดยการประเมินการปรับเทียบนั้นจะพิจาณาถึงความคลาดเคลื่อนที่เกิดขึ้นจากการประมาณค่า ดัชนีของความคลาดเคลื่อนที่ศึกษาส่วนใหญ่ได้แก่ ค่าความลำเอียง ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ย (RMSE) ค่าความแตกต่างกำลังสองเฉลี่ย (MSD) และค่า RMSD (Li, Lissitz and Yang, 1999; Bastari, 2000; Kim, 2004; Kim and Lee, 2006)

สำหรับการวิจัยครั้งนี้ผู้วิจัยเลือกใช้โมเดลการตอบตามทฤษฎีการตอบสนองข้อสอบที่ให้คะแนนสองค่า แบบโลจีสติก 3 พารามิเตอร์ เนื่องจากลักษณะของข้อสอบที่ได้รับความนิยมในการนำไปใช้ทดสอบไม่ว่าจะเป็นระดับจุลภาคหรือระดับมหภาค มักจะเป็นข้อสอบแบบเลือกตอบ ซึ่งธรรมชาติของข้อสอบแบบเลือกตอบแต่ละข้อจะมีความยาก และอำนาจจำแนกไม่เท่ากัน อีกทั้งผู้สอบยังสามารถเดาคำตอบได้ ดังนั้นการใช้โมเดลโลจีสติกแบบ 3 พารามิเตอร์ จึงน่าจะเหมาะสมและใกล้เคียงกับสภาพการวัดและประเมินจริง ส่วนโมเดลการตอบสนองข้อสอบตามทฤษฎีการตอบสนองข้อสอบที่ให้คะแนนหลายค่า ผู้วิจัยเลือกใช้โมเดล Generalize Partial Credit : GPC ที่พัฒนาโดย Muraki (1992, 1993) เพราะโมเดลนี้อาจเป็นข้อสอบได้หลายรูปแบบ

ทั้ง essay, passage – based หรือ rating scale ซึ่งสามารถนำไปใช้กับข้อสอบหลายรูปแบบตามความเหมาะสม และยังมีการให้คะแนนแบบตอบถูกบางส่วนหรือให้คะแนนตามลำดับขั้นก็ได้ อีกทั้งการศึกษาเอกสารและวรรณกรรมที่เกี่ยวข้อง ยังพบว่าแบบสอบรูปแบบผสมเป็นการผสมระหว่างโมเดล 3PL กับโมเดล GPC มากที่สุดด้วย

การออกแบบการเก็บรวบรวมข้อมูลที่ผู้วิจัยเลือกศึกษาในครั้งนี้มี 2 แบบแผนคือ กลุ่มสมมูลหรือกลุ่มสมมูลใช้ข้อสอบร่วม และกลุ่มไม่เท่าเทียมใช้ข้อสอบร่วม เนื่องจากการวิจัยครั้งนี้เป็นการศึกษาการจำลองข้อมูล ผู้วิจัยจึงต้องการให้ผลการการปรับเทียบครอบคลุมเงื่อนไขต่างๆ ที่มีโอกาสถูกนำไปใช้ในสถานการณ์จริงมากที่สุด

ส่วนข้อสอบร่วมมีทั้งข้อสอบที่ให้คะแนนสองค่าและหลายค่ารวมอยู่ด้วยกัน เนื่องจากมีหลักฐานสนับสนุนว่า ให้ผลการปรับเทียบที่ดีที่สุด เพราะลักษณะของข้อสอบรวมควรมีความเป็นตัวแทนแบบสอบทั้งฉบับให้มากที่สุด รวมศึกษาเกี่ยวกับสัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่า และหลายค่าที่รวมอยู่ในแบบสอบ จำนวนกลุ่มตัวอย่าง เพื่อศึกษาผลของการปรับในเงื่อนไขต่าง ๆ

สำหรับวิธีการปรับเทียบ ผู้วิจัยพิจารณาว่าควรจะศึกษาการปรับเทียบคะแนนวิธีโค้งคุณลักษณะแบบสอบ ซึ่งเป็นการปรับเทียบคะแนนที่แปลงค่าพารามิเตอร์ความสามารถเชิงเส้นและวิธีปรับค่าพารามิเตอร์พร้อมกัน เนื่องจากผลการวิจัยที่ผ่านมา มีหลักฐานสนับสนุนว่า วิธีการปรับเทียบแบบโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน จะให้ผลการปรับเทียบที่มีความคลาดเคลื่อนน้อยที่สุด การปรับเทียบคะแนนตามวิธีการทฤษฎีการตอบสนองข้อสอบเป็นการปรับเทียบทั้งคะแนนจริงและคะแนนสังเกตและโมเดล IRT ที่ใช้ในการศึกษาครั้งนี้เลือกใช้โมเดลที่ผสมระหว่างโมเดลการให้คะแนนสองค่าแบบ 3 พารามิเตอร์ และโมเดลการให้คะแนนหลายค่าแบบ GPC ที่สามารถใช้วิเคราะห์ข้อมูลดังกล่าวได้ รวมทั้งมีโปรแกรมการเทียบคะแนนด้วยโค้งลักษณะแบบทดสอบในโมเดลนี้ด้วย จึงเป็นสิ่งสนับสนุนการวิจัยครั้งนี้ เพื่อให้เกิดสารสนเทศและข้อความรู้ในการนำไปใช้ต่อไป

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

# บทที่ 3
# วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาคุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างการตรวจให้คะแนนแบบสองค่าและการตรวจให้คะแนนแบบหลายค่าที่เป็นเอกมิติ ตามแนวทฤษฎีการตอบสนองข้อสอบด้วยวิธีโค้งคุณลักษณะ (Characteristic Curve) และวิธีการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent Calibration) โดยการศึกษาจำลองข้อมูล (Simulation Study) โมเดลตามทฤษฎีการตอบสนองข้อสอบที่ตรวจให้คะแนนสองค่า แบบโลจีสติก 3 พารามิเตอร์ และโมเดลตามทฤษฎีการตอบสนองข้อสอบที่ตรวจให้คะแนนหลายค่าแบบ Generalized Partial Credit เกณฑ์ที่ใช้ในการประเมินเพื่อเปรียบเทียบวิธีการปรับเทียบทั้ง 2 วิธี คือ ค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (Mean Square Error : MSE) ผลการวิจัยจะได้สารสนเทศเกี่ยวกับการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมในเงื่อนไขต่าง ๆ ที่สามารถเลือกไปใช้กับสถานการณ์ต่าง ๆ ให้เหมาะสมกับข้อมูลจริงต่อไป ขั้นตอนการดำเนินการวิจัย มีดังนี้

1. ศึกษาค้นคว้าเอกสารและงานวิจัย เกี่ยวกับการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างการให้คะแนนสองค่า และหลายค่าอยู่ในแบบสอบฉบับเดียวกัน จากการค้นคว้างานวิจัยที่เกี่ยวข้อง เอกสาร ตำรา วารสารทั้งในและต่างประเทศที่ผ่านมาพบว่าส่วนใหญ่ศึกษาแบบสอบที่มีการศึกษากับโมเดลการตอบตามทฤษฎีการตอบสนองข้อสอบ และโมเดลที่ได้รับความนิยมในการศึกษาคือ โมเดลโลจีสติกแบบ 3 PL โมเดลเกรดเรสพอนส์ โมเดลพาเชียลเครดิต และโมเดลเจนเนอรัลไลซ์พาเชียลเครดิต สำหรับการวิจัยครั้งนี้ผู้วิจัยพิจารณาลักษณะของข้อสอบที่ใช้สอบในบริบทของการศึกษาไทย และตัดสินใจเลือกใช้โมเดลการตอบที่ให้คะแนนแบบสองค่า แบบโลจีสติก 3 พารามิเตอร์ และโมเดลการตอบที่มีการให้คะแนนแบบหลายค่า แบบโมเดลเจนเนอรัลไลซ์พาเชียลเครดิต (Generalized Partial Credit)

2. ศึกษาวิธีการปรับเทียบรูปแบบต่าง ๆ และเงื่อนไขการทดสอบที่เกิดขึ้นในการศึกษาวิจัย จากเอกสารและงานวิจัยที่เกี่ยวข้อง รวมถึงรูปแบบการเก็บรวบรวมข้อมูลในการปรับเทียบคะแนนด้วย

3. ศึกษาแนวทางในการจำลองข้อมูลสำหรับใช้ในการวิจัย เพื่อให้ได้ค่าพารามิเตอร์ที่แท้จริงมาใช้ในการวิจัย

4. จำลองข้อมูลการตอบของผู้สอบตามแบบแผนและตัวแปรที่กำหนดด้วยโปรแกรม Wingen

5. ประมาณค่าพารามิเตอร์ข้อสอบจากข้อมูลผลการตอบ ด้วยโปรแกรม MULTILOG

6. ตรวจสอบความถูกต้องของข้อมูลที่ได้จากการจำลองข้อมูลด้วยโปรแกรม WINGEN และโปรแกรม MULTILOG โดยการวิเคราะห์เปรียบเทียบค่าพารามิเตอร์ข้อสอบที่ได้จากทั้งสองโปรแกรม โดยการวิเคราะห์ t-test

7. ตรวจสอบความเป็นเอกมิติของข้อมูลด้วยการวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis) ด้วยโปรแกรม SPSS

8. วิเคราะห์ปรับเทียบแบบสอบ 2 ฉบับตามแต่ละเงื่อนไข ด้วยวิธีโค้งคุณลักษณะและการปรับค่าพารามิเตอร์พร้อมกัน

9. เปรียบเทียบคุณภาพของผลการปรับเทียบคะแนนแต่ละเงื่อนไขจากค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง

10. สรุปผลการวิเคราะห์ข้อมูล เพื่อเปรียบเทียบคุณภาพการปรับเทียบคะแนนในแต่ละเงื่อนไขที่ทำการศึกษา

จากที่กล่าวมาข้างต้นเป็นการสรุปขั้นตอนการดำเนินงานวิจัยในภาพรวม รายละเอียดต่อไปจะเสนอวิธีดำเนินการวิจัย ดังนี้

**ตัวแปรที่ใช้ในการวิจัย**

เนื่องจากการวิจัยครั้งนี้มุ่งเปรียบเทียบคุณภาพของวิธีการปรับเทียบคะแนนตามแนวทฤษฎีการตอบสนองข้อสอบ โดยวิธีโค้งคุณลักษณะ และการปรับค่าพารามิเตอร์พร้อมกัน สำหรับแบบสอบรูปแบบผสมระหว่างข้อสอบที่มีการตรวจให้คะแนนแบบสองค่า กับข้อสอบที่มีการตรวจให้คะแนนแบบหลายค่า ศึกษาเปรียบเทียบคุณภาพวิธีการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมในเงื่อนไขต่าง ๆ ดังนั้นจึงประกอบด้วยตัวแปรที่ใช้ในการวิจัย ดังต่อไปนี้

**ตัวแปรอิสระ มีดังนี้**

1. สัดส่วนข้อสอบที่ตรวจให้คะแนนแบบสองค่า (Dichotomous 3 PL: DS) กับข้อสอบที่ตรวจให้คะแนนหลายค่า (Polytomous GPCM: PS 6 ลำดับขั้นการตอบ คะแนน 0 – 5 คะแนน) ในแบบสอบ จำนวน 4 สัดส่วน ซึ่งสัดส่วนน้ำหนักการให้คะแนนระหว่างข้อสอบ DS และ PS ได้แก่ 50 : 50, 60 : 40, 70 : 30, และ 80 : 20 ตามลำดับ จากสัดส่วนของน้ำหนักคะแนน และคิดเป็นสัดส่วนจำนวนข้อสอบ แสดงดังต่อไปนี้

| สัดส่วนน้ำหนักคะแนน | | สัดส่วนข้อสอบ | | รวมจำนวนข้อสอบ |
|---|---|---|---|---|
| 1.1 DS 50% : PS 50% | → | DS 50 ข้อ : PS 10 ข้อ | → | 60 ข้อ |
| 1.2 DS 60% : PS 40% | → | DS 60 ข้อ : PS 8 ข้อ | → | 68 ข้อ |
| 1.3 DS 70% : PS 30% | → | DS 70 ข้อ : PS 6 ข้อ | → | 76 ข้อ |
| 1.4 DS 80% : PS 20% | → | DS 80 ข้อ : PS 4 ข้อ | → | 84 ข้อ |

2. ข้อสอบร่วม ผู้วิจัยสนใจศึกษาข้อสอบร่วม 2 มิติด้วยกันคือ จำนวนข้อสอบ และจำนวนคะแนนของรวมของข้อสอบร่วม โดยทั้ง 2 มิติ ผู้วิจัยใช้ข้อสอบทั้งที่ตรวจให้คะแนนสองค่าและตรวจให้คะแนนหลายค่าเป็นข้อสอบร่วม เพื่อให้ข้อสอบร่วมมีความเป็นตัวแทน (Kolen and Brennan, 2004) ข้อสอบร่วมในมิติของจำนวนข้อสอบ โดยศึกษาสัดส่วนของข้อสอบร่วมเมื่อคิดตามจำนวนข้อ ผู้วิจัยใช้ข้อสอบร่วมที่ประกอบด้วยข้อสอบทั้ง 2 ชนิด ที่มีจำนวนข้อตั้งแต่ 20% ขึ้นไป เพื่อให้มีความเป็นตัวแทนของแบบสอบทั้งฉบับ (Kolen and Brennan, 2004); Angoff, 1971) ผู้วิจัยพิจารณาปริมาณข้อสอบทั้ง 4 สัดส่วน คือ 60, 68, 76 และ 84 และศึกษาว่าจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าที่แตกต่างกัน จะมีผลต่อความคลาดเคลื่อนในการปรับเทียบหรือไม่ จึงกำหนดให้จำนวนข้อสอบที่ตรวจให้คะแนนหลายค่ามีจำนวนคงที่ คือจำนวน 2 ข้อ และให้ข้อสอบที่ตรวจให้คะแนนสองค่าแปรเปลี่ยนไป 3 ระดับ เพื่อศึกษาแนวโน้มของผลการปรับเทียบคะแนนว่าเมื่อจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าในข้อสอบร่วมลดลง จะทำให้ผลการปรับเทียบมีความแม่นยำมากขึ้น ดังนี้

2.1.1 DS 15 ข้อ : PS 2 ข้อ

2.1.2 DS 20 ข้อ : PS 2 ข้อ

2.1.3 DS 25 ข้อ : PS 2 ข้อ

2.2 ข้อสอบร่วมในมิติของคะแนน ผู้วิจัยกำหนดให้คะแนนรวมจากข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนหลายค่าในข้อสอบร่วมที่ศึกษาครั้งนี้มีค่าเท่ากับ 20% ของคะแนนจากแบบสอบทั้งฉบับ และศึกษาสัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนหลายค่าในข้อสอบร่วม นั่นคือ หากคะแนนรวมของข้อสอบที่ตรวจให้คะแนนสองค่าในข้อสอบร่วมมีค่ามากกว่า เท่ากับ และน้อยกว่า คะแนนรวมของข้อสอบที่ตรวจให้คะแนนหลายค่าในข้อสอบร่วม ผลการวิจัยจะเป็นอย่างไร จึงกำหนดสัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนหลายค่าในข้อสอบร่วมที่แตกต่างกัน 3 สัดส่วน ดังนี้

2.2.1 DS 15 ข้อ (15 คะแนน) : PS 1 ข้อ (5 คะแนน) รวม 20 คะแนน

2.2.2 DS 10 ข้อ (10 คะแนน) : PS 2 ข้อ (10 คะแนน) รวม 20 คะแนน

2.2.3 DS 5 ข้อ (5 คะแนน) : PS 3 ข้อ (15 คะแนน) รวม 20 คะแนน

3. แบบแผนการเก็บรวบรวมข้อมูล

3.1 กลุ่มเท่าเทียมโดยใช้ข้อสอบร่วม (Equivalent group with common items)

3.2 กลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม (Non – equivalent group with common items)

4. วิธีการปรับเทียบคะแนน 2 วิธีคือ

4.1 วิธีโค้งคุณลักษณะ

4.2 การปรับค่าพารามิเตอร์พร้อมกัน

**ตัวแปรตาม** คือ ค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (Mean Square Error : MSE)

สำหรับการเปรียบเทียบคุณภาพการปรับเทียบคะแนนในการวิจัยครั้งนี้ พิจารณาจากค่า MSE ที่มาจากเกณฑ์โค้งคุณลักษณะลำดับขั้นการตอบ (Category response curve) โดยมีพื้นฐานมาจากความแตกต่างระหว่างโค้งลำดับขั้นการตอบที่ประมาณค่าได้ กับโค้งลำดับขั้นการตอบที่เป็นค่าที่แท้จริง เพื่อให้สอดคล้องกับงานวิจัยของ Hanson และ Beguin (2002), Kim (2004), และ Kim และ Lee (2006) ที่ได้ศึกษาเปรียบเทียบคุณภาพการปรับเทียบคะแนนระหว่างวิธีการปรับเทียบตามทฤษฎีการตอบสนองข้อสอบ เกณฑ์ที่ใช้นี้เป็นค่าของ MSE สำหรับประมาณค่าจากการทำซ้ำ 100 ครั้ง โดยเปรียบเทียบค่าที่ได้ ถ้าเงื่อนไขที่ศึกษาใดมีค่า MSE ที่น้อยกว่าแสดงว่ามีคุณภาพในการปรับเทียบคะแนนได้ดีกว่า นั่นหมายความว่าค่า MSE ที่ลดลง แสดงถึงความคลาดเคลื่อนที่ลดลงไปด้วย

จากที่กล่าวมาข้างต้นเป็นการสรุปขั้นตอนดำเนินงานวิจัยโดยรวมตั้งแต่การศึกษาค้นคว้าเอกสารและงานวิจัยที่เกี่ยวข้องจนถึงการวิเคราะห์และสรุปผลการวิจัย รวมถึงตัวแปรที่ใช้ในการวิจัย ในรายละเอียดต่อไป จะนำเสนอวิธีดำเนินการวิจัยโดยละเอียด ซึ่งในที่นี้แบ่งออกเป็น 5 ตอน ได้แก่

ตอนที่ 1 เงื่อนไขที่ศึกษาการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม

ตอนที่ 2 การจำลองข้อมูลเพื่อดำเนินการปรับเทียบคะแนนตามแบบแผนที่กำหนด

ตอนที่ 3 การตรวจสอบความถูกต้องของข้อมูล

ตอนที่ 4 การปรับเทียบคะแนน

ตอนที่ 5 การเปรียบเทียบคุณภาพการปรับเทียบคะแนน

## ตอนที 1 เงื่อนไขทีศึกษาการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องที่มีการศึกษาการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมส่วนใหญ่เป็นโมเดลการตอบข้อสอบตามทฤษฎีการตอบสนองข้อสอบที่มีการตรวจให้คะแนนสองค่าแบบ 3 พารามิเตอร์ โมเดลที่มีการตรวจให้คะแนนหลายค่าแบบ Grade Response, Partial Credit, และ Generalized Partial Credit สำหรับการวิจัยครั้งนี้แบบสอบรูปแบบผสมผู้วิจัยเลือกโมเดลการให้คะแนน 2 ค่า แบบ 3 พารามิเตอร์ ซึ่งลักษณะของแบบสอบที่มีการตรวจให้คะแนนแบบสองค่าส่วนใหญ่เป็นข้อสอบชนิดเลือกตอบแบบหลายตัวเลือก (Multiple Choices) ธรรมชาติของข้อสอบชนิดนี้มีลักษณะที่ข้อสอบแต่ละข้อมีความยากและอำนาจจำแนกไม่เท่ากัน อีกทั้งยังสามารถเดาคำตอบได้ (Hambleton and Swaminathan, 1984; Embretson and Reise, 2000) ดังนั้นการเลือกใช้โมเดลการตอบที่มีการตรวจให้คะแนนแบบ 2 ค่า ชนิด 3 พารามิเตอร์จึงน่าจะเหมาะสมที่สุด ส่วนโมเดลที่มีการตรวจให้คะแนนหลายค่า ผู้วิจัยเลือกใช้โมเดล Generalize partial Credit ที่มีรายการคำตอบ 6 ลำดับขั้น เนื่องจากโมเดลลักษณะนี้จะมีความเหมาะสมกับข้อสอบที่มีลักษณะการให้คะแนนเป็นลำดับขั้น (Muraki,1992,1993)

จากการวิจัยที่ผ่านมา ส่วนใหญ่แสดงหลักฐานว่าการปรับเทียบคะแนนสำหรับแบบสอบไม่ว่าจะเป็นโมเดลที่มีการตรวจให้คะแนนสองค่า หรือโมเดลที่มีการตรวจให้คะแนนหลายค่า วิธีที่ให้ผลการปรับเทียบเป็นที่น่าพอใจคือวิธีโค้งคุณลักษณะ (Baker and Al-Karni, 1991; Kim and Cohen, 1992; Ogasawara, 2001; Hanson and Beguin, 2002; Kim and Lee, 2006) และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ดังนั้น การวิจัยครั้งนี้ ผู้วิจัยได้ศึกษาเปรียบเทียบการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมตามวิธีทางทฤษฎีการตอบสนองข้อสอบ 2 วิธี คือ วิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน โดยมีเงื่อนไขที่ศึกษาดังนี้

1. แบบแผนการเก็บรวบรวมข้อมูล 2 แบบแผน ได้แก่ กลุ่มเท่าเทียมโดยใช้ข้อสอบร่วม และกลุ่มไม่เท่าเทียมโดยใช้ข้อสอบร่วม เนื่องจากการวิจัยที่ผ่านมามีการศึกษาการปรับเทียบคะแนนสำหรับรูปแบบผสมโดยใช้การออกแบบการเก็บรวบรวมข้อมูลเพียงรูปแบบหนึ่ง การศึกษาเปรียบเทียบแบบแผนการเก็บข้อมูลข้างต้นยังปรากฏไม่มากนัก (Kim and Lee, 2006) ผู้วิจัยจึงสนใจศึกษาผลของการปรับเทียบคะแนนของแบบสอบที่มีแบบแผนการเก็บข้อมูลที่ต่างกัน โดยการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความเท่าเทียมนั้น ผู้สอบทั้ง 2 กลุ่มมีค่าเฉลี่ยความสามารถเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 ในขณะที่การเก็บรวบรวมข้อมูลกับผู้สอบที่ไม่เท่าเทียมกัน ผู้สอบกลุ่มหนึ่งมีค่าเฉลี่ยความสามารถเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 และอีกกลุ่มหนึ่งมีค่าเฉลี่ยความสามารถเท่ากับ 1 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1

2. จำนวนข้อสอบ สำหรับข้อสอบที่ให้คะแนนสองค่า (Dichotomous : DS) เป็นการจำลองข้อมูลให้สอดคล้องกับโมเดลตามทฤษฎีการตอบสนองข้อสอบ โลจีสติกแบบ 3 พารามิเตอร์ และข้อสอบที่มีการตรวจให้คะแนนหลายค่า (Polytomous : PS) จำลองข้อมูลให้สอดคล้องกับโมเดลตามทฤษฎีการตอบสนองข้อสอบตามโมเดล Generalized Partial Credit ที่มีรายการคำตอบ 6 ลำดับขั้น (มีคะแนน 0 – 5 คะแนน) ซึ่งสัดส่วนน้ำหนักของข้อสอบทั้ง 2 ชนิด มี 4 สัดส่วนด้วยกัน ผู้วิจัยจึงทำการจำลองข้อมูลให้สอดคล้องกับโมเดลตามทฤษฎีการตอบสนองข้อสอบและเป็นไปตามข้อตกลงเบื้องต้น กำหนดสัดส่วนน้ำหนักของข้อสอบที่มีการตรวจให้คะแนนสองค่า และข้อสอบที่ตรวจให้คะแนนหลายค่า ดังต่อไปนี้

1) DS 50% : PS 50% คิดเป็นจำนวนข้อเท่ากับ DS 50 : PS 10

2) DS 60% : PS 40% คิดเป็นจำนวนข้อเท่ากับ DS 60 : PS 8

3) DS 70% : PS 30% คิดเป็นจำนวนข้อเท่ากับ DS 70 : PS 6

4) DS 80% : PS 20% คิดเป็นจำนวนข้อเท่ากับ DS 80 : PS 4

3. สัดส่วนของข้อสอบร่วม ศึกษาใน 2 มิติด้วยกันคือ มิติของจำนวนข้อสอบทั้งสอง ชนิดในข้อสอบร่วม และมิติของคะแนนรวมที่ได้จากข้อสอบทั้งสองชนิดในข้อสอบร่วม (รายละเอียดดังที่กล่าวไว้ในบทที่ 1)

3.1 ข้อสอบร่วมในมิติของจำนวนข้อสอบ มี 3 สัดส่วน ดังนี้

3.1.1 DS 15 ข้อ : PS 2 ข้อ

3.1.2 DS 20 ข้อ : PS 2 ข้อ

3.1.3 DS 25 ข้อ : PS 2 ข้อ

3.2 ข้อสอบร่วมในมิติของคะแนน มี 3 สัดส่วน ดังนี้

3.2.1 DS 15 ข้อ (15 คะแนน) : PS 1 ข้อ (5 คะแนน) รวม 20 คะแนน

3.2.2 DS 10 ข้อ (10 คะแนน) : PS 2 ข้อ (10 คะแนน) รวม 20 คะแนน

3.2.3 DS 5 ข้อ (5 คะแนน) : PS 3 ข้อ (15 คะแนน) รวม 20 คะแนน

จากเงื่อนไข แบบแผนการเก็บข้อมูล สัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่ากับข้อสอบที่ตรวจให้คะแนนหลายค่าของแบบสอบทั้งฉบับ สัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่ากับข้อสอบที่ตรวจให้คะแนนหลายค่าในข้อสอบร่วม (มิติของจำนวนข้อสอบร่วม และมิติของคะแนนในข้อสอบร่วม) และวิธีการปรับเทียบ กล่าวได้ว่า การวิจัยครั้งนี้มีเงื่อนไขที่ต้องการศึกษาทั้งหมด 96 เงื่อนไข (2 แบบแผนการเก็บข้อมูล × 4 สัดส่วนของข้อสอบทั้งสองชนิดในแบบสอบ × 6 สัดส่วนของข้อสอบทั้งสองชนิดในข้อสอบร่วม (3 สัดส่วนในมิติของจำนวนข้อ และ 3 สัดส่วนในมิติของคะแนน) × 2 วิธีการปรับเทียบ)

## ตอนที 2 การจำลองข้อมูลเพื่อดำเนินการปรับเทียบคะแนนตามเงื่อนไขทีกำหนด

การจำลองข้อมูลตามเงื่อนไขสัดส่วนของข้อสอบและการออกแบบการเก็บรวบรวมข้อมูล 2 รูปแบบ นั่นคือ กลุ่มผู้สอบที่มีความเท่าเทียมและกลุ่มผู้สอบที่มีความไม่เท่าเทียมกัน ผู้วิจัยจึงกำหนดให้กลุ่มผู้สอบกลุ่มหนึ่งเป็นกลุ่มฐาน คือกลุ่มผู้สอบที่มีค่าเฉลี่ยความสามารถเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 โดยกำหนดให้ทำแบบสอบฉบับ A (Test A) สำหรับกลุ่มปรับเทียบที่กลุ่มผู้สอบที่มีความเท่าเทียมกับกลุ่มผู้สอบฐานจะมีค่าเฉลี่ยความสามารถเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 เช่นเดียวกัน กำหนดให้ทำแบบทดสอบฉบับ B1 (Test B1) ส่วนกลุ่มผู้สอบที่เป็นกลุ่มปรับเทียบอีกกลุ่มหนึ่งที่มีความสามารถไม่เท่าเทียมกับกลุ่มผู้สอบฐานจะมีค่าเฉลี่ยความสามารถเท่ากับ 1 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 กำหนดให้ทำแบบสอบฉบับ B2 (Test B2)

จากแบบทดสอบทั้ง 3 ฉบับข้างต้น (Test A, Test B1, Test B2) ผู้วิจัยพิจารณาจำนวนข้อสอบในแบบสอบแต่ละสัดส่วน แล้วพิจารณาการจำลองข้อมูลโดยใช้จำนวนสูงสุดของข้อสอบแต่ละรูปแบบกำหนดไว้คงที่ (Kim and Lee, 2006) นั่นคือข้อสอบแบบแบบเลือกตอบสูงสุด จำนวน 80 ข้อ ข้อสอบแบบเขียนตอบสูงสุด จำนวน 10 ข้อ จำลองข้อมูลเพื่อใช้ในการปรับเทียบคะแนนตามรูปแบบการเก็บรวบรวมข้อมูลข้างต้น ได้แบบสอบ 3 ฉบับ (Test A, Test B1, Test B2) แล้วจึงกำหนดสัดส่วนข้อสอบตามเงื่อนไข 4 เงื่อนไข ที่กำหนดไว้ในการศึกษา รูปแบบการกำหนดข้อสอบสำหรับสัดส่วนข้อสอบแต่ละเงื่อนไข แสดงดังตารางที่ 4

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4 รูปแบบข้อสอบที่ใช้ในการปรับเทียบคะแนนจำแนกตามสัดส่วนของข้อสอบในแบบสอบ

| Test A | Test B1 | Test B2 | เงื่อนไขสัดส่วนข้อสอบในแบบสอบ | | | |
|---|---|---|---|---|---|---|
| DS | DS | DS | 50/10 | 60/8 | 70/3 | 80/4 |
| 1 | 1 | 1 | X | X | X | X |
| 2 | 2 | 2 | X | X | X | X |
| 3 | 3 | 3 | X | X | X | X |
| 4 | 4 | 4 | X | X | X | X |
| 5 | 5 | 5 | X | X | X | X |
| 6 | 6 | 6 | X | X | X | X |
| 7 | 7 | 7 | X | X | X | X |
| 8 | 8 | 8 | X | X | X | X |
| 9 | 9 | 9 | X | X | X | X |
| 10 | 10 | 10 | X | X | X | X |
| 11 | 11 | 11 | X | X | X | X |
| 12 | 12 | 12 | X | X | X | X |
| 13 | 13 | 13 | X | X | X | X |
| 14 | 14 | 14 | X | X | X | X |
| 15 | 15 | 15 | X | X | X | X |
| 16 | 16 | 16 | X | X | X | X |
| 17 | 17 | 17 | X | X | X | X |
| 18 | 18 | 18 | X | X | X | X |
| 19 | 19 | 19 | X | X | X | X |
| 20 | 20 | 20 | X | X | X | X |
| 21 | 21 | 21 | X | X | X | X |
| 22 | 22 | 22 | X | X | X | X |
| 23 | 23 | 23 | X | X | X | X |
| 24 | 24 | 24 | X | X | X | X |
| 25 | 25 | 25 | X | X | X | X |
| 26 | 26 | 26 | X | X | X | X |
| 27 | 27 | 27 | X | X | X | X |
| 28 | 28 | 28 | X | X | X | X |
| 29 | 29 | 29 | X | X | X | X |
| 30 | 30 | 30 | X | X | X | X |
| 31 | 31 | 31 | X | X | X | X |
| 32 | 32 | 32 | X | X | X | X |
| 33 | 33 | 33 | X | X | X | X |
| 34 | 34 | 34 | X | X | X | X |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4 รูปแบบข้อสอบที่ใช้ในการปรับเทียบคะแนนจำแนกตามสัดส่วนของข้อสอบในแบบสอบ (ต่อ)

| Test A | Test B1 | Test B2 | เงื่อนไขสัดส่วนข้อสอบในแบบสอบ | | | |
|---|---|---|---|---|---|---|
| DS | DS | DS | 50/10 | 60/8 | 70/3 | 80/4 |
| 35 | 35 | 35 | X | X | X | X |
| 36 | 36 | 36 | X | X | X | X |
| 37 | 37 | 37 | X | X | X | X |
| 38 | 38 | 38 | X | X | X | X |
| 39 | 39 | 39 | X | X | X | X |
| 40 | 40 | 40 | X | X | X | X |
| 41 | 41 | 41 | X | X | X | X |
| 42 | 42 | 42 | X | X | X | X |
| 43 | 43 | 43 | X | X | X | X |
| 44 | 44 | 44 | X | X | X | X |
| 45 | 45 | 45 | X | X | X | X |
| 46 | 46 | 46 | X | X | X | X |
| 47 | 47 | 47 | X | X | X | X |
| 48 | 48 | 48 | X | X | X | X |
| 49 | 49 | 49 | X | X | X | X |
| 50 | 50 | 50 | X | X | X | X |
| 51 | 51 | 51 | | X | X | X |
| 52 | 52 | 52 | | X | X | X |
| 53 | 53 | 53 | | X | X | X |
| 54 | 54 | 54 | | X | X | X |
| 55 | 55 | 55 | | X | X | X |
| 56 | 56 | 56 | | X | X | X |
| 57 | 57 | 57 | | X | X | X |
| 58 | 58 | 58 | | X | X | X |
| 59 | 59 | 59 | | X | X | X |
| 60 | 60 | 60 | | X | X | X |
| 61 | 61 | 61 | | | X | X |
| 62 | 62 | 62 | | | X | X |
| 63 | 63 | 63 | | | X | X |
| 64 | 64 | 64 | | | X | X |
| 65 | 65 | 65 | | | X | X |
| 66 | 66 | 66 | | | X | X |
| 67 | 67 | 67 | | | X | X |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4 รูปแบบข้อสอบที่ใช้ในการปรับเทียบคะแนนจำแนกตามสัดส่วนของข้อสอบในแบบสอบ (ต่อ)

| Test A | Test B1 | Test B2 | เงื่อนไขสัดส่วนข้อสอบในแบบสอบ | | | |
|---|---|---|---|---|---|---|
| DS | DS | DS | 50/10 | 60/8 | 70/3 | 80/4 |
| 68 | 68 | 68 | | | X | X |
| 69 | 69 | 69 | | | X | X |
| 70 | 70 | 70 | | | X | X |
| 71 | 71 | 71 | | | | X |
| 72 | 72 | 72 | | | | X |
| 73 | 73 | 73 | | | | X |
| 74 | 74 | 74 | | | | X |
| 75 | 75 | 75 | | | | X |
| 76 | 76 | 76 | | | | X |
| 77 | 77 | 77 | | | | X |
| 78 | 78 | 78 | | | | X |
| 79 | 79 | 79 | | | | X |
| 80 | 80 | 80 | | | | X |
| PS | PS | PS | | | | |
| 1 | 1 | 1 | X | X | X | X |
| 2 | 2 | 2 | X | X | X | X |
| 3 | 3 | 3 | X | X | X | X |
| 4 | 4 | 4 | X | X | X | X |
| 5 | 5 | 5 | X | X | X | |
| 6 | 6 | 6 | X | X | X | |
| 7 | 7 | 7 | X | X | | |
| 8 | 8 | 8 | X | X | | |
| 9 | 9 | 9 | X | | | |
| 10 | 10 | 10 | X | | | |

จากรูปแบบการจำลองข้อมูลข้างต้น ผู้วิจัยได้ทำการจำลองข้อมูลโดยกำหนดรูปแบบการแจกแจงค่าพารามิเตอร์ข้อสอบและพารามิเตอร์ผู้สอบ ดังกล่าวในหัวข้อต่อไป

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

**2.1 การกำหนดพารามิเตอร์ของข้อสอบในการจำลองข้อมูล**

ในการสร้างพารามิเตอร์ข้อสอบต้องเป็นค่าพารามิเตอร์ที่แท้จริง จากงานวิจัยที่ศึกษาเกี่ยวกับการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมมีทั้งการใช้ข้อมูลจริง (Li, Lissit and Yang, 1999; Bastari, 2000; Lee, 2001; Kim, Walker and McHale, 2007) และข้อมูลที่จำลองขึ้น (Tate, 2000; Kim, 2004; Kirkpatrick, 2005; Kim and Lee, 2006) ซึ่งการวิจัยครั้งนี้ผู้วิจัยได้ศึกษาการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมโดยการจำลองข้อมูลเพื่อให้ได้ค่าพารามิเตอร์ที่แท้จริง และเป็นข้อมูลที่สอดคล้องกับโมเดลการตอบข้อสอบ ดังนั้นผู้วิจัยจึงพิจารณาการจำลองข้อมูลให้สอดคล้องกับงานวิจัยของ Kim (2004) และ Kim และ Lee (2006) ที่ได้ศึกษาการปรับเทียบตามทฤษฎีการตอบสนองข้อสอบสำหรับแบบสอบรูปแบบผสม4 วิธี คือ mean/mean, mean/sigma, Heabara, และ Stocking and Lord โดยจำลองข้อมูลพารามิเตอร์ของข้อสอบตามโมเดลการตอบข้อสอบที่ตรวจให้คะแนนแบบสองค่า 3 PL ดังนี้ พารามิเตอร์ความยาก มีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 โดยมีการแจกแจงของพารามิเตอร์แบบ Normal (b∼N(0,1)) พารามิเตอร์อำนาจจำแนก มีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.2 โดยมีการแจกแจงของพารามิเตอร์แบบ Log normal (a∼LN(0,0.2)) และพารามิเตอร์การเดา มีการแจกแจงแบบ Beta โดยมี 2 พารามิเตอร์คือ $\alpha$ และ $\beta$ (BETA($\alpha$,$\beta$)) ส่วนการจำลองข้อมูลสำหรับโมเดลการตรวจให้คะแนนแบบหลายค่าตามโมเดล GPC ค่าพารามิเตอร์ความชันของข้อคำถาม ($\alpha_i$) มีการจำลองเหมือนกับค่าอำนาจจำแนกในโมเดล 3PL ค่าพารามิเตอร์ระดับความยากของขั้นการตอบ ($\delta_{ij}$) 6 ลำดับขั้น โดยมีการแจกแจงแบบปกติ

สำหรับความน่าจะเป็นในการตอบถูกของผู้สอบสามารถคำนวณได้จากผู้สอบแต่ละคนโดยการตอบข้อสอบของผู้สอบคนที่ i ที่มีความสามารถ $\theta$ ตอบข้อสอบข้อที่ j เขียนเป็น $U_{ij}$ จำลองข้อมูลโดยเปรียบเทียบกับการสุ่มแบบเอกรูป (Uniform) เขียนเป็น R ที่อยู่ในช่วง (0,1) ไปจนถึงความน่าจะเป็นในการตอบข้อสอบถูก $P_{ij}$ ตามโมเดล 3PL คือ

$$U_{ij} = \begin{cases} 0, & R > P_{ij} \\ 1, & R \leq P_{ij} \end{cases}$$

ในทำนองเดียวกัน การตอบข้อสอบที่มีการตรวจให้คะแนนหลายค่าของคนที่ i ในรายการคำตอบที่ k ของข้อที่ j เขียนเป็น $U_{ijk}$ มีความน่าจะเป็นในการตอบถูกตามลักษณะของโมเดล GPC คือ

$$U_{ijk} = \begin{cases} 0, & 0 < R \leq \breve{P}_{ij1} \\ 1, & \breve{P}_{ij1} < R \leq \breve{P}_{ij2} \\ . & \\ . & \\ K-1 & \breve{P}_{ij(K-1)} < R \leq \breve{P}_{ijK} \\ K & \breve{P}_{ijK} < R < 1 \end{cases}$$

เมื่อ $\breve{P}_{ijk} = \sum_{h=1}^{k} P_{ijh}$ และ K คือจำนวนรายการคำตอบลบด้วย 1

**2.2 แบบแผนการจำลองข้อมูลตามลักษณะการเก็บรวบรวมข้อมูล**

การปรับเทียบคะแนนจำเป็นต้องมีสเกลร่วม เนื่องจากผู้สอบสองกลุ่มที่ทำแบบสอบ 2 ฉบับที่ต่างกัน จะมีการแจงความความสามารถของผู้สอบ 2 กลุ่มต่างกัน ซึ่งสถานการณ์ที่การวิจัยครั้งนี้ศึกษาคือ การปรับเทียบคะแนนโดยใช้กลุ่มเท่าเทียม และการปรับเทียบคะแนนโดยใช้กลุ่มไม่เท่าเทียม ความแตกต่างระหว่างความสามารถของผู้สอบทั้ง 2 กลุ่มที่เป็นกลุ่มเก่าและกลุ่มใหม่สามารถเขียนในลักษณะสมการแปลงรูปเชิงเส้น ดังนี้

$$\theta^{*}_{N} = A\theta_{N} + B \qquad (1)$$

เมื่อ $\theta^{*}_{N}$ เทียบเท่ากับ $\theta_{O}$

ถ้า A = 1 และ B = 0 จะมีลักษณะเป็นกลุ่มเท่าเทียม และถ้า A = 1 และ B = 1 จะมีลักษณะเป็นกลุ่มไม่เท่าเทียม ในการเชื่อมคะแนนสำหรับ 2 สถานการณ์นี้ จะมีการแจกแจงความสามารถของกลุ่มผู้สอบเป็นการแจกแจงแบบปกติ โดยมีค่าเฉลี่ยคือ $\mu$ และส่วนเบี่ยงเบนมาตรฐานคือ $\sigma$ นั่นคือ $N(\mu,\sigma)$ สำหรับการจำลองข้อมูล กลุ่มผู้สอบที่เป็นกลุ่มเก่า จะสุ่มค่าความสามารถจากการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น 0 และส่วนเบี่ยงเบนมาตรฐานเป็น 1 นั่นคือ N(0,1) ส่วนผู้สอบกลุ่มใหม่ ก็จะสุ่มค่าความสามารถจากการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น 0 และส่วนเบี่ยงเบนมาตรฐานเป็น 1 นั่นคือ N(0,1) เช่นเดียวกัน เนื่องจากเป็นการเก็บข้อมูลที่เป็นกลุ่มเท่าเทียม ส่วนการเก็บข้อมูลที่เป็นกลุ่มไม่เท่าเทียม กลุ่มผู้สอบที่เป็นกลุ่มเก่าจะได้มาจาก สุ่มค่าความสามารถจากการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น 0 และส่วนเบี่ยงเบนมาตรฐานเป็น 1 นั่นคือ N(0,1) และกลุ่มผู้สอบกลุ่มใหม่จะได้มาจาก สุ่มค่าความสามารถจากการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น 1 และส่วนเบี่ยงเบนมาตรฐานเป็น 1 นั่นคือ N(1,1) นั่นเอง

**จากการจำลองข้อมูลเพื่อให้ได้แบบแผนการตอบข้อสอบ ผู้วิจัยแสดงเป็นลำดับขั้นตอนการจำลองข้อมูลในภาพรวม** ดังนี้

1. จำลองข้อมูลพารามิเตอร์ข้อสอบตามโมเดลที่กำหนด นั่นคือ โมเดลการตอบข้อสอบที่ตรวจให้คะแนนแบบสองค่า 3 PL กำหนดค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และลักษณะของการแจกแจง สำหรับพารามิเตอร์ข้อสอบ ทั้ง 3 พารามิเตอร์ คือ พารามิเตอร์อำนาจจำแนก (a), พารามิเตอร์ความยาก (b) และพารามิเตอร์การเดา (c) และจำลองข้อมูลพารามิเตอร์ข้อสอบตามโมเดลการตอบข้อสอบที่ตรวจให้คะแนนหลายค่า โมเดล GPC โดยกำหนดค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และลักษณะการแจกแจงของค่าพารามิเตอร์ระดับความยากของขั้นการตอบ ($\delta_{ij}$) 6 ลำดับขั้น และค่าพารามิเตอร์ความชันของข้อคำถาม ($\alpha_i$)

2. จำลองข้อมูลพารามิเตอร์ผู้สอบตามเงื่อนไขของการรวบรวมข้อมูล นั่นคือ การเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียม กำหนดให้ค่าเฉลี่ยพารามิเตอร์ผู้สอบเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 สำหรับกลุ่มผู้สอบทั้ง 2 กลุ่ม ส่วนการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียมกำหนดให้กลุ่มผู้สอบกลุ่มหนึ่งมีค่าเฉลี่ยพารามิเตอร์ผู้สอบเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 และผู้สอบอีกกลุ่มหนึ่งมีค่าเฉลี่ยพารามิเตอร์ผู้สอบเท่ากับ 1 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1

3. ใช้ข้อมูลในข้อที่ 1 และ 2 คำนวณหาความน่าจะเป็นในการตอบข้อสอบตามลักษณะของโมเดลการตอบข้อสอบ

4. หาผลการตอบของผู้สอบคนที่ i ที่มีความสามารถ $\theta$ ตอบข้อสอบข้อที่ j เขียนเป็น $U_{ij}$ จำลองข้อมูลโดยเปรียบเทียบกับการสุ่มแบบเอกกรูป (Uniform) เขียนเป็น R ที่อยู่ในช่วง (0,1) ไปจนถึงความน่าจะเป็นในการตอบข้อสอบถูก $P_{ij}$ ตามโมเดล 3PL คือ

$$U_{ij} = \begin{cases} 0, & R > P_{ij} \\ 1, & R \leq P_{ij} \end{cases}$$

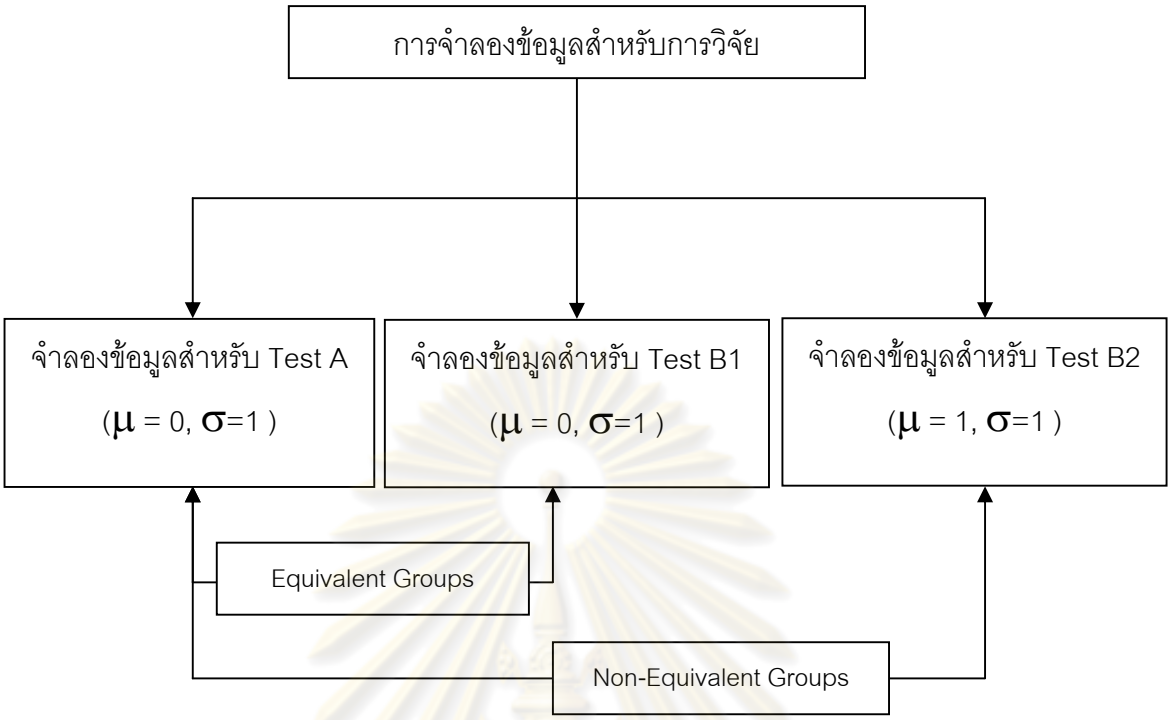
ในทำนองเดียวกัน การตอบข้อสอบที่มีการตรวจให้คะแนนหลายค่าของคนที่ i ในรายการคำตอบที่ k ของข้อที่ j เขียนเป็น $U_{ijk}$ มีความน่าจะเป็นในการตอบถูกตามลักษณะของโมเดล GPC คือ

$$U_{ijk} = \begin{cases} 0, & 0 < R \leq \breve{P}_{ij1} \\ 1, & \breve{P}_{ij1} < R \leq \breve{P}_{ij2} \\ . & \\ . & \\ K-1 & \breve{P}_{ij(K-1)} < R \leq \breve{P}_{ijK} \\ K & \breve{P}_{ijK} < R < 1 \end{cases}$$

เมื่อ $\breve{P}_{ijk} = \sum_{h=1}^{k} P_{ijh}$ และ K คือจำนวนรายการคำตอบลบด้วย 1

5. นำข้อมูลการตอบที่ได้ไปตรวจสอบค่าพารามิเตอร์ข้อสอบด้วยโปรแกรม MULTILOG แบบแผนการจำลองข้อมูลแสดงดังแผนภาพที่ 16

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย




**แผนภาพที** 16 การจำลองข้อมูลตามเงื่อนไขที่ศึกษา

**ตอนที่** 3 **การตรวจสอบความถูกต้องของข้อมูล**

**การตรวจสอบความถูกต้องของข้อมูลที่ได้จากการจำลองด้วยโปรแกรม** WINGEN

ตรวจสอบความถูกต้องของข้อมูลจากการจำลองข้อมูลด้วยโปรแกรม WINGEN (Han and Hambleton, 2007) โดยจำลองข้อมูลในแต่ละเงื่อนไขจากแบบสอบที่ผสมระหว่างโมเดล 3 PL และ โมเดล GPC จำนวนผู้สอบ 1000 คน เนื่องจาก การศึกษาของ Kim (2004) และ Kim และ Lee (2006) ได้จำลองข้อมูลการตอบตามโมเดล 3PL และ GPC กับกลุ่มตัวอย่างจำนวน 500 และ 3000 คน มีการทำซ้ำ 100 ครั้ง และพบว่าการปรับเทียบคะแนนเกิดความคลาดเคลื่อนในการปรับเทียบน้อย เมื่อกลุ่มตัวอย่างมีขนาด 3000 คน อีกทั้งการศึกษาของ Harwell, Hsu และ Kirisci (1996) พบว่าถ้าศึกษาโดยใช้โมเดล IRT เป็นฐานควรมีการทำซ้ำอย่างน้อย 20 ครั้ง เมื่อจำนวนกลุ่มตัวอย่างมีตั้งแต่ 1000 คนขึ้นไป ดังนั้นการศึกษาครั้งนี้ ผู้วิจัยได้นำผลการศึกษาจากข้างต้นมาพิจารณาใช้ในการจำลองข้อมูลแบบสอบรูปแบบผสมตามการศึกษาของ Tate (2000), Kim (2004) และ Kim และ Lee (2006) ที่มีการทำซ้ำ 100 ครั้ง และพิจารณาร่วมกับผลการศึกษาของ Harwell, Hsu และ Kirisci (1996) ผู้วิจัยจึงศึกษาโดยใช้การจำลองข้อมูลที่มีการทำซ้ำ 100 ครั้ง เพื่อให้สอดคล้องกับการจำลองข้อมูลสำหรับแบบสอบรูปแบบผสมตามการศึกษาของ Tate (2000), Kim (2004) และ Kim และ Lee (2006) โดยมีกลุ่มตัวอย่างจำนวน 1000 คน และนำคะแนนที่ได้จากการ

จำลองข้อมูลซึ่งเป็นรูปแบบการตอบข้อสอบแต่ละข้อจากการตอบของผู้สอบแต่ละคนไปวิเคราะห์โดยใช้โปรแกรม MULTILOG พิจารณาค่าพารามิเตอร์ข้อสอบ ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าความน่าจะเป็นในการเดาถูก (c) ในโมเดล 3 PL และค่าพารามิเตอร์ความชันของข้อคำถาม ($\alpha_i$) ในโมเดล GPC นำไปเปรียบเทียบกับพารามิเตอร์ตั้งต้นที่ใช้ในการจำลองข้อมูลรูปแบบการตอบ จากโปรแกรม WINGEN และทดสอบความแตกต่างระหว่างค่าเฉลี่ยพารามิเตอร์ข้อสอบที่ได้จากโปรแกรมทั้งสองโดยใช้ t – test สำหรับทดสอบกลุ่มตัวอย่าง 2 กลุ่มที่เกี่ยวข้องกัน (Two dependent samples t-test)

จากการจำลองข้อมูลโดยพิจารณาจากจำนวนข้อสูงสุดของแต่ละโมเดลในแบบสอบแต่ละฉบับประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า (Dichotomous : DS) จำนวน 80 ข้อ และข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า (Polytomous : PS) จำนวน 10 ข้อ **(ค่าพารามิเตอร์ข้อสอบในแบบสอบแต่ละฉบับที่จำลองได้แสดงในภาคผนวก ก)** ซึ่งพารามิเตอร์ข้อสอบได้จากโปรแกรม Wingen สำหรับแบบสอบแต่ละฉบับ ผู้วิจัยได้สำเสนอค่าเฉลี่ยพารามิเตอร์ข้อสอบที่ใช้ในการเปรียบ 4 พารามิเตอร์ คือ a, b, c และ $\alpha$ แสดงดังตารางที่ 5

ตารางที่ 5 ค่าเฉลี่ยของพารามิเตอร์ที่แท้จริงของข้อสอบที่โปรแกรม Wingen ใช้ในการจำลองข้อมูล

| แบบสอบ | ค่าเฉลี่ยพารามิเตอร์ข้อสอบ | | | |
|---|---|---|---|---|
| | a | b | c | $\alpha$ |
| Test A | 1.049 | -0.035 | 0.126 | 1.000 |
| Test B1 | 1.035 | 0.073 | 0.123 | 1.024 |
| Test B2 | 0.991 | 0.034 | 0.124 | 1.053 |

จากตารางที่ 5 แสดงให้เห็นว่าพารามิเตอร์อำนาจจำแนกของแบบสอบฉบับ A มีค่าเฉลี่ยเท่ากับ 1.049 แบบสอบฉบับ B1 มีค่าพารามิเตอร์อำนาจจำแนกเฉลี่ยเท่ากับ 1.035 และแบบสอบฉบับ B2 มีค่าพารามิเตอร์อำนาจจำแนกเฉลี่ยเท่ากับ 0.991

พารามิเตอร์ความยากของแบบสอบฉบับ A มีค่าเฉลี่ยเท่ากับ -0.035 แบบสอบฉบับ B1 มีค่าเฉลี่ยเท่ากับ 0.073 และแบบสอบฉบับ B2 มีค่าเฉลี่ยเท่ากับ 0.034

พารามิเตอร์การเดาของแบบสอบฉบับ A มีค่าเฉลี่ยเท่ากับ 0.126 แบบสอบฉบับ B1 มีค่าเฉลี่ยเท่ากับ 0.123 และแบบสอบฉบับ B2 มีค่าเฉลี่ยเท่ากับ 0.124

พารามิเตอร์ความชันของข้อสอบของแบบสอบฉบับ A มีค่าเฉลี่ยเท่ากับ 1.000แบบสอบฉบับ B1 มีค่าเฉลี่ยเท่ากับ 1.024 และแบบสอบฉบับ B2 มีค่าเฉลี่ยเท่ากับ 1.053

ข้อมูลข้างต้นเป็นค่าเฉลี่ยพารามิเตอร์ข้อสอบที่แท้จริงสำหรับแบบสอบแต่ละฉบับที่ได้จากโปรแกรม Wingen ซึ่งผู้วิจัยใช้เป็นข้อมูลในการวิเคราะห์เปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยพารามิเตอร์ข้อสอบที่ได้จากโปรแกรม Wingen และพารามิเตอร์ข้อสอบที่ได้จากการวิเคราะห์ด้วยโปรแกรม MULTILOG ผู้วิจัยใช้พารามิเตอร์ข้อสอบจากโปรแกรม Wingen จำลองข้อมูลรูปแบบการตอบของแบบสอบแต่ละฉบับ (Test A, Test B1, Test B2) ฉบับละ 100 ชุด (replication 100 ครั้ง) ทำการประมาณค่าพารามิเตอร์ข้อสอบจากรูปแบบการตอบข้อสอบแต่ละชุดโดยใช้โปรแกรม MULTLOG แล้วจึงตรวจสอบว่าพารามิเตอร์ที่ได้จากการประมาณค่าด้วยโปรแกรม MULTILOG แตกต่างจากพารามิเตอร์ข้อสอบจากโปรแกรม Wingen โดยวิเคราะห์ทดสอบความแตกต่างระหว่างค่าเฉลี่ยพารามิเตอร์ข้อสอบที่ได้จากโปรแกรมทั้งสอง โดยใช้ t–test สำหรับทดสอบกลุ่มตัวอย่างสองกลุ่มที่เกี่ยวข้องกัน (Two dependent sample t-test) ที่ระดับนัยสำคัญ .05 (ผลการวิเคราะห์เปรียบเทียบค่าเฉลี่ยจากการทำซ้ำ 100 ครั้ง แสดงในภาคผนวก ข) ตัวอย่างผลการเปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยดังกล่าว แสดงดังตารางที่ 6 – 8

ตารางที่ 6 ผลการเปรียบเทียบค่าเฉลี่ยพารามิเตอร์ข้อสอบจากโปรแกรม WINGEN และ MULTILOG สำหรับ Test A จากการประมาณค่า 100 ครั้ง

| ครั้งที่ | a | | | b | | | c | | | alpha | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD | t | sig | MD | t | sig | MD | t | sig | MD | t | sig |
| 1 | 0.53 | 0.53 | 0.60 | 0.03 | 0.23 | 0.82 | 0.00 | 0.30 | 0.76 | -0.05 | -0.99 | 0.35 |
| 2 | -0.19 | -0.19 | 0.85 | -0.01 | -0.06 | 0.95 | 0.00 | 0.43 | 0.67 | 0.10 | 1.18 | 0.27 |
| 3 | -0.14 | -0.14 | 0.89 | -0.01 | -0.03 | 0.97 | 0.00 | 0.22 | 0.82 | 0.06 | 0.56 | 0.59 |
| 4 | 0.91 | 0.91 | 0.37 | 0.03 | 0.19 | 0.85 | 0.00 | 0.04 | 0.97 | -0.02 | -0.29 | 0.78 |
| 5 | 0.91 | 0.91 | 0.37 | 0.03 | 0.19 | 0.85 | 0.00 | 0.04 | 0.97 | 0.11 | 1.31 | 0.22 |
| 6 | -0.58 | -0.58 | 0.56 | -0.02 | -0.13 | 0.90 | 0.00 | -0.28 | 0.78 | -0.06 | -0.84 | 0.42 |
| 7 | -0.58 | -0.58 | 0.56 | -0.02 | -0.13 | 0.90 | 0.00 | -0.28 | 0.78 | 0.00 | 0.01 | 0.99 |
| 8 | -0.98 | -0.98 | 0.33 | 0.04 | 0.30 | 0.77 | -0.01 | -1.02 | 0.31 | -0.02 | -0.44 | 0.67 |
| 9 | 1.04 | 1.04 | 0.30 | 0.06 | 0.38 | 0.70 | 0.00 | 0.33 | 0.74 | -0.04 | -0.47 | 0.65 |
| 10 | -1.68 | -1.68 | 0.10 | 0.20 | 1.34 | 0.18 | 0.02 | 1.42 | 0.16 | 0.08 | 0.88 | 0.40 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 95 | -1.69 | -1.69 | 0.09 | 0.11 | 0.71 | 0.48 | 0.00 | -0.51 | 0.61 | 0.09 | 0.76 | 0.46 |
| 96 | -1.91 | -1.91 | 0.06 | -0.05 | -0.33 | 0.75 | 0.01 | 1.00 | 0.32 | 0.00 | -0.03 | 0.97 |
| 97 | -1.91 | -1.91 | 0.06 | -0.05 | -0.33 | 0.75 | 0.01 | 1.00 | 0.32 | 0.01 | 0.11 | 0.92 |
| 98 | -0.26 | -0.26 | 0.80 | -0.03 | -0.18 | 0.86 | 0.00 | 0.28 | 0.78 | 0.01 | 0.15 | 0.88 |
| 99 | -1.14 | -1.14 | 0.26 | 0.08 | 0.50 | 0.62 | 0.01 | 0.49 | 0.63 | 0.00 | 0.00 | 1.00 |
| 100 | -1.02 | -1.02 | 0.31 | 0.10 | 0.72 | 0.47 | 0.00 | 0.34 | 0.74 | 0.08 | 1.00 | 0.34 |

หมายเหตุ MD หมายถึง Mean Difference

ตารางที่ 7 ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test B1 จากการประมาณค่า 100 ครั้ง

| ครั้งที่ | a | | | b | | | c | | | alpha | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD | t | sig | MD | t | sig | MD | t | sig | MD | t | sig |
| 1 | -0.01 | -0.18 | 0.86 | -0.12 | -0.82 | 0.42 | 0.00 | -0.46 | 0.65 | 0.00 | 0.02 | 0.98 |
| 2 | -0.01 | -0.44 | 0.66 | -0.06 | -0.37 | 0.71 | 0.00 | 0.24 | 0.81 | 0.11 | 0.12 | 0.11 |
| 3 | 0.01 | 0.43 | 0.67 | -0.05 | -0.32 | 0.75 | 0.01 | 1.13 | 0.26 | 0.03 | 0.33 | 0.75 |
| 4 | -0.05 | -1.46 | 0.15 | 0.04 | 0.25 | 0.80 | -0.01 | -1.28 | 0.20 | 0.02 | 0.22 | 0.83 |
| 5 | -0.04 | -1.26 | 0.21 | -0.10 | -0.64 | 0.52 | 0.00 | -0.01 | 0.99 | 0.09 | 1.05 | 0.32 |
| 6 | 0.00 | 0.02 | 0.99 | -0.05 | -0.34 | 0.74 | -0.01 | -1.40 | 0.16 | 0.00 | -0.03 | 0.98 |
| 7 | -0.05 | -1.65 | 0.10 | -0.25 | -1.52 | 0.13 | 0.01 | 1.31 | 0.20 | -0.01 | -0.21 | 0.84 |
| 8 | -0.04 | -1.24 | 0.22 | -0.13 | -0.80 | 0.43 | 0.00 | 0.49 | 0.63 | 0.04 | 0.61 | 0.56 |
| 9 | -0.04 | -1.15 | 0.25 | -0.17 | -1.01 | 0.31 | 0.00 | 0.23 | 0.82 | -0.06 | -0.69 | 0.51 |
| 10 | -0.02 | -0.64 | 0.53 | -0.30 | -2.04 | 0.04 | 0.01 | 0.64 | 0.52 | 0.00 | -0.05 | 0.96 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 95 | 0.04 | 1.13 | 0.26 | -0.10 | -0.65 | 0.52 | 0.00 | 0.44 | 0.66 | -0.18 | -2.02 | 0.07 |
| 96 | -0.01 | -0.16 | 0.87 | -0.08 | -0.52 | 0.60 | 0.01 | 0.62 | 0.54 | -0.03 | -0.38 | 0.71 |
| 97 | 0.00 | -0.12 | 0.90 | -0.23 | -1.58 | 0.12 | 0.00 | 0.17 | 0.87 | 0.06 | 0.52 | 0.62 |
| 98 | 0.02 | 0.66 | 0.51 | -0.03 | -0.18 | 0.86 | -0.01 | -0.84 | 0.40 | 0.03 | 0.35 | 0.73 |
| 99 | -0.02 | -0.55 | 0.58 | -0.10 | -0.67 | 0.51 | 0.01 | 1.03 | 0.31 | 0.02 | 0.02 | 0.79 |
| 100 | -0.05 | -1.31 | 0.20 | 0.05 | 0.31 | 0.76 | 0.01 | 0.82 | 0.41 | 0.03 | 0.24 | 0.81 |

หมายเหตุ MD หมายถึง Mean Difference

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 8 ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test B2 จากการประมาณค่า 100 ครั้ง

| ครั้งที่ | a | | | b | | | c | | | alpha | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD | t | Sig | MD | t | sig | MD | t | sig | MD | t | sig |
| 1 | 0.02 | 0.50 | 0.62 | -0.02 | -0.13 | 0.90 | 0.01 | 0.53 | 0.60 | -0.09 | -1.38 | 0.20 |
| 2 | 0.04 | 1.29 | 0.20 | -0.08 | -0.57 | 0.57 | 0.01 | 0.62 | 0.54 | -0.04 | -0.57 | 0.58 |
| 3 | 0.03 | 0.88 | 0.38 | -0.02 | -0.12 | 0.90 | 0.00 | -0.14 | 0.89 | -0.05 | -0.48 | 0.65 |
| 4 | 0.02 | 0.55 | 0.59 | -0.08 | -0.46 | 0.65 | 0.00 | -0.16 | 0.87 | -0.13 | -2.04 | 0.07 |
| 5 | 0.03 | 0.80 | 0.42 | 0.09 | 0.53 | 0.60 | 0.00 | -0.13 | 0.90 | -0.04 | -0.47 | 0.65 |
| 6 | 0.04 | 1.20 | 0.23 | -0.06 | -0.32 | 0.75 | 0.00 | -0.21 | 0.83 | -0.03 | -0.32 | 0.76 |
| 7 | 0.02 | 0.55 | 0.59 | 0.10 | 0.65 | 0.52 | -0.01 | -0.53 | 0.60 | -0.07 | -0.87 | 0.41 |
| 8 | 0.02 | 0.50 | 0.62 | 0.03 | 0.20 | 0.84 | -0.01 | -0.66 | 0.51 | 0.02 | 0.19 | 0.85 |
| 9 | 0.03 | 1.07 | 0.29 | -0.06 | -0.35 | 0.73 | -0.01 | -0.52 | 0.61 | -0.04 | -0.36 | 0.73 |
| 10 | 0.03 | 0.99 | 0.33 | 0.06 | 0.35 | 0.73 | 0.00 | 0.10 | 0.92 | -0.06 | -1.34 | 0.21 |
| . | . | . | . | . | . | . | . | . | . | . | . | . |
| 95 | 0.04 | 1.29 | 0.20 | 0.10 | 0.57 | 0.57 | 0.00 | 0.46 | 0.65 | 0.03 | 0.29 | 0.78 |
| 96 | 0.03 | 1.03 | 0.31 | -0.07 | -0.41 | 0.68 | 0.00 | -0.17 | 0.87 | -0.02 | -0.35 | 0.73 |
| 97 | 0.05 | 1.53 | 0.13 | -0.08 | -0.50 | 0.62 | 0.00 | 0.45 | 0.65 | -0.03 | -0.41 | 0.69 |
| 98 | 0.06 | 1.87 | 0.07 | 0.05 | 0.34 | 0.74 | 0.01 | 0.62 | 0.54 | -0.05 | -0.49 | 0.63 |
| 99 | 0.03 | 0.84 | 0.41 | -0.02 | -0.11 | 0.91 | 0.00 | -0.24 | 0.81 | 0.10 | 1.07 | 0.31 |
| 100 | 0.02 | 0.76 | 0.45 | -0.22 | -1.31 | 0.19 | 0.00 | 0.21 | 0.84 | -0.03 | -0.45 | 0.66 |

หมายเหตุ MD หมายถึง Mean Difference

ผลการวิเคราะห์เปรียบเทียบค่าเฉลี่ยพารามิเตอร์ของข้อสอบจากการโปรแกรม Wingen และค่าพารามิเตอร์ของข้อสอบในแบบสอบฉบับ A (Test A) ฉบับ B1 (Test B1) ฉบับ B2 (Test B2) ที่ประมาณค่าได้จากโปรแกรม MULTILOG พบว่า ในการทำซ้ำข้อมูลทั้ง 100 ครั้ง ผลการวิเคราะห์ไม่พบความแตกต่างระหว่างค่าพารามิเตอร์ของข้อสอบจากโปรแกรม Wingen และค่าพารามิเตอร์ของข้อสอบที่ประมาณค่าได้จากโปรแกรม MULTILOG (ผลการวิเคราะห์เปรียบเทียบค่าเฉลี่ยพารามิเตอร์ข้อสอบระหว่างโปรแกรม Wingen และพารามิเตอร์ข้อสอบที่ได้จากโปรแกรม MULTILOG Test A, Test B1 และ Test B2 แสดงในภาคผนวก ข) แสดงให้เห็นว่าผลจากการจำลองข้อมูลรูปแบบการตอบข้อสอบจากโปรแกรม Wingen มีคุณภาพเพียงพอต่อการนำไปศึกษาได้

**การตรวจสอบความเป็นเอกมิติของข้อมูล**

การตรวจสอบความเป็นเอกมิติของข้อมูล ตามทฤษฎีการตอบสนองข้อสอบ โดยการนำข้อมูลผลการตอบข้อสอบรายข้อที่จำลองจากโปรแกรม WINGEN สำหรับแบบสอบที่ผสมระหว่างโมเดล 3 PL กับโมเดล GPC จำนวนผู้สอบ 1000 ไปวิเคราะห์องค์ประกอบ (Principle component analysis) และหมุนแกนแบบมุมแหลม เพื่อตรวจสอบความเป็นเอกมิติของข้อมูล โดยพิจารณาว่าร้อยละของความแปรปรวนของตัวประกอบแรกพิจารณาแล้วเห็นว่ามีค่ามากพอที่จะสรุปว่ามีความเป็นเอกมิติหรือไม่ ซึ่งเกณฑ์ของ Rackase (อ้างถึงใน Raju, 1993) เสนอแนะการตรวจสอบความเป็นเอกมิติของข้อมูลตามทฤษฎีการตอบสนองข้อสอบ ค่าความแปรปรวนจากตัวประกอบหลักตัวแรกของแบบสอบควรมีค่าไม่ต่ำกว่าร้อยละ 20.00 และ/หรือ ตามเกณฑ์ของวรนุช แหยมแสง (2536) เสนอแนะว่าอัตราส่วนระหว่างตัวประกอบที่ 1 และตัวประกอบที่ 2 ไม่ควรต่ำกว่า 5.739 สามารถสรุปได้ว่า ข้อมูลมีความเป็นเอกมิติเป็นไปตามข้อตกลงของทฤษฎีการตอบสนองข้อสอบ

การตรวจสอบความเป็นเอกมิติของข้อมูลผู้วิจัยได้นำแบบทดสอบทั้ง 3 ฉบับข้างต้น จัดกระทำข้อมูลตามเงื่อนไขสัดส่วนของข้อสอบทั้งสองชนิดในแบบสอบแต่ละฉบับ (Test A, Test B1, Test B2) จำแนกตามเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบเป็น 4 สัดส่วนได้แก่ 1) DS 50 : PS 10, 2) DS 60 : PS 8, 3) DS 70 : PS 6, และ 4) DS 80 : PS 4 แล้วทำการวิเคราะห์ตรวจสอบความเป็นเอกมิติของแบบสอบแต่ละสัดส่วน โดยการวิเคราะห์ตัวประกอบ (Factor Analysis) ด้วยวิธีวิเคราะห์ตัวประกอบสำคัญ (Principal Component Analysis) และหมุนแกนแบบมุมแหลม ตามที่ Lord (1980) เสนอแนะ ใช้โปรแกรมสำเร็จรูป SPSS ในการวิเคราะห์ ปรากฏผล ดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 9 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของเงื่อนไขสัดส่วนของข้อสอบในแบบสอบฉบับ 50 : 10 ด้วยการทำซ้ำ 100 ครั้ง

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 1 | 1 | 19.585 | 29.309 | 20.010 | 29.267 | 19.921 | 29.348 |
| | 2 | 4.178 | 5.473 | 4.431 | 5.748 | 3.812 | 4.991 |
| 2 | 1 | 19.476 | 30.033 | 19.610 | 29.798 | 19.969 | 29.298 |
| | 2 | 5.371 | 6.888 | 5.088 | 6.544 | 3.236 | 4.299 |
| 3 | 1 | 19.252 | 28.862 | 18.913 | 29.448 | 19.201 | 28.897 |
| | 2 | 4.289 | 5.597 | 4.352 | 5.652 | 3.606 | 4.694 |
| 4 | 1 | 19.517 | 30.082 | 19.359 | 29.083 | 20.003 | 29.238 |
| | 2 | 3.754 | 4.886 | 3.985 | 5.231 | 5.124 | 6.568 |
| 5 | 1 | 19.536 | 29.577 | 19.909 | 29.086 | 19.057 | 28.689 |
| | 2 | 2.982 | 4.005 | 4.252 | 5.566 | 4.784 | 6.214 |
| 6 | 1 | 19.591 | 29.574 | 19.103 | 29.029 | 20.069 | 28.815 |
| | 2 | 4.517 | 5.849 | 3.213 | 4.268 | 3.325 | 4.370 |
| 7 | 1 | 19.333 | 29.604 | 19.236 | 29.644 | 20.119 | 30.079 |
| | 2 | 5.323 | 6.824 | 4.950 | 6.370 | 5.299 | 6.786 |
| 8 | 1 | 19.646 | 30.151 | 19.673 | 28.785 | 19.196 | 30.113 |
| | 2 | 3.297 | 4.337 | 5.010 | 6.427 | 5.131 | 6.589 |
| 9 | 1 | 19.678 | 29.319 | 19.401 | 29.508 | 19.918 | 29.295 |
| | 2 | 4.808 | 6.252 | 3.700 | 4.854 | 3.059 | 4.077 |
| 10 | 1 | 19.223 | 29.290 | 19.659 | 29.581 | 19.858 | 28.971 |
| | 2 | 3.486 | 4.570 | 3.113 | 4.135 | 3.929 | 5.159 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 97 | 1 | 19.018 | 29.254 | 19.374 | 28.664 | 19.674 | 29.248 |
| | 2 | 3.384 | 4.433 | 4.574 | 5.929 | 3.409 | 4.458 |
| 98 | 1 | 20.187 | 28.937 | 19.332 | 29.689 | 19.307 | 29.443 |
| | 2 | 5.326 | 6.836 | 5.056 | 6.524 | 3.796 | 4.973 |
| 99 | 1 | 20.095 | 29.490 | 19.714 | 29.078 | 20.124 | 28.924 |
| | 2 | 3.953 | 5.198 | 3.251 | 4.300 | 3.876 | 5.087 |
| 100 | 1 | 19.931 | 29.727 | 19.694 | 29.058 | 19.106 | 29.380 |
| | 2 | 3.586 | 4.663 | 3.128 | 4.152 | 3.623 | 4.725 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 10 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของเงื่อนไขสัดส่วนของข้อสอบในแบบสอบฉบับ 60 : 8 ด้วยการทำซ้ำ 100 ครั้ง

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 1 | 1 | 19.442 | 28.615 | 18.902 | 27.368 | 19.611 | 28.947 |
| | 2 | 3.848 | 5.038 | 3.831 | 5.022 | 4.061 | 5.280 |
| 2 | 1 | 19.307 | 29.069 | 19.746 | 28.750 | 19.489 | 28.705 |
| | 2 | 3.887 | 5.106 | 3.024 | 4.036 | 4.658 | 6.062 |
| 3 | 1 | 19.249 | 28.002 | 19.172 | 28.019 | 19.934 | 28.162 |
| | 2 | 4.540 | 5.882 | 5.353 | 6.873 | 4.455 | 5.787 |
| 4 | 1 | 20.074 | 27.276 | 19.546 | 28.591 | 18.939 | 28.624 |
| | 2 | 3.989 | 5.232 | 4.936 | 6.363 | 4.006 | 5.238 |
| 5 | 1 | 19.187 | 28.970 | 19.420 | 27.243 | 19.174 | 28.190 |
| | 2 | 3.192 | 4.220 | 3.744 | 4.874 | 5.136 | 6.597 |
| 6 | 1 | 19.635 | 29.111 | 19.364 | 28.114 | 20.042 | 28.457 |
| | 2 | 5.028 | 6.462 | 4.162 | 5.461 | 4.819 | 6.261 |
| 7 | 1 | 19.909 | 28.930 | 19.767 | 27.618 | 19.560 | 27.240 |
| | 2 | 4.247 | 5.552 | 5.122 | 6.567 | 4.786 | 6.225 |
| 8 | 1 | 20.076 | 27.654 | 19.752 | 27.498 | 19.172 | 29.100 |
| | 2 | 3.861 | 5.064 | 4.471 | 5.797 | 3.806 | 4.987 |
| 9 | 1 | 20.149 | 28.557 | 19.757 | 28.187 | 20.154 | 28.289 |
| | 2 | 4.653 | 6.052 | 4.089 | 5.337 | 4.014 | 5.242 |
| 10 | 1 | 19.446 | 28.671 | 19.778 | 28.894 | 18.986 | 27.952 |
| | 2 | 3.215 | 4.277 | 4.333 | 5.647 | 3.674 | 4.818 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 97 | 1 | 19.502 | 27.870 | 19.881 | 28.443 | 20.021 | 27.779 |
| | 2 | 4.089 | 5.334 | 3.306 | 4.357 | 2.941 | 3.969 |
| 98 | 1 | 19.247 | 28.384 | 19.030 | 28.620 | 19.287 | 28.684 |
| | 2 | 3.483 | 4.563 | 5.158 | 6.625 | 4.946 | 6.367 |
| 99 | 1 | 19.064 | 29.097 | 19.928 | 28.847 | 20.096 | 29.086 |
| | 2 | 4.796 | 6.238 | 5.017 | 6.432 | 5.060 | 6.529 |
| 100 | 1 | 19.080 | 28.815 | 20.137 | 28.685 | 19.789 | 29.096 |
| | 2 | 3.498 | 4.588 | 3.235 | 4.298 | 2.955 | 3.990 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 11 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของเงื่อนไขสัดส่วนของข้อสอบในแบบสอบฉบับ 70 : 6 ด้วยการทำซ้ำ 100 ครั้ง

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 1 | 1 | 18.268 | 25.422 | 19.683 | 25.596 | 18.182 | 27.887 |
| | 2 | 3.614 | 4.712 | 3.257 | 4.306 | 5.006 | 6.419 |
| 2 | 1 | 19.883 | 26.912 | 17.59 | 28.861 | 18.113 | 28.26 |
| | 2 | 3.197 | 4.233 | 3.648 | 4.757 | 5.051 | 6.516 |
| 3 | 1 | 17.824 | 26.411 | 19.502 | 26.298 | 17.325 | 26.461 |
| | 2 | 3.681 | 4.826 | 4.840 | 6.279 | 4.567 | 5.919 |
| 4 | 1 | 19.437 | 25.862 | 17.453 | 26.181 | 18.546 | 27.586 |
| | 2 | 3.082 | 4.098 | 3.040 | 4.053 | 3.032 | 4.047 |
| 5 | 1 | 19.539 | 28.363 | 19.637 | 28.474 | 17.488 | 27.68 |
| | 2 | 4.098 | 5.355 | 4.885 | 6.312 | 2.878 | 3.935 |
| 6 | 1 | 17.45 | 26.123 | 19.532 | 27.644 | 17.966 | 28.863 |
| | 2 | 3.227 | 4.290 | 4.692 | 6.104 | 4.545 | 5.882 |
| 7 | 1 | 19 | 25.899 | 19.492 | 28.914 | 19.837 | 25.718 |
| | 2 | 4.959 | 6.381 | 5.224 | 6.691 | 4.164 | 5.463 |
| 8 | 1 | 19.09 | 28.964 | 19.189 | 26.485 | 18.668 | 27.623 |
| | 2 | 3.538 | 4.622 | 4.899 | 6.322 | 2.904 | 3.940 |
| 9 | 1 | 19.622 | 27.49 | 17.897 | 28.37 | 18.579 | 27.063 |
| | 2 | 3.418 | 4.474 | 3.839 | 5.023 | 5.341 | 6.870 |
| 10 | 1 | 18.095 | 28.708 | 17.651 | 26.548 | 18.91 | 29.194 |
| | 2 | 3.853 | 5.052 | 5.216 | 6.676 | 4.474 | 5.799 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 97 | 1 | 19.355 | 28.39 | 19.874 | 25.852 | 17.816 | 27.625 |
| | 2 | 3.984 | 5.229 | 3.036 | 4.048 | 4.991 | 6.404 |
| 98 | 1 | 17.606 | 27.221 | 18.25 | 27.578 | 19.37 | 28.911 |
| | 2 | 4.263 | 5.568 | 2.852 | 3.916 | 3.879 | 5.089 |
| 99 | 1 | 19.056 | 26.14 | 19.108 | 27.645 | 18.357 | 28.95 |
| | 2 | 4.292 | 5.604 | 3.855 | 5.053 | 3.981 | 5.225 |
| 100 | 1 | 19.326 | 28.751 | 19.167 | 25.606 | 19.568 | 26.994 |
| | 2 | 5.318 | 6.811 | 3.207 | 4.266 | 4.180 | 5.474 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 12 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของเงื่อนไขสัดส่วนของข้อสอบในแบบสอบฉบับ 80 : 4 ด้วยการทำซ้ำ 100 ครั้ง

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 1 | 1 | 17.542 | 24.151 | 18.094 | 25.552 | 17.682 | 25.011 |
| | 2 | 3.425 | 4.492 | 3.375 | 4.426 | 3.612 | 4.711 |
| 2 | 1 | 18.11 | 25.047 | 18.335 | 24.282 | 17.78 | 24.53 |
| | 2 | 3.896 | 5.121 | 3.644 | 4.750 | 4.320 | 5.627 |
| 3 | 1 | 17.887 | 25.312 | 18.434 | 25.095 | 17.955 | 25.262 |
| | 2 | 5.067 | 6.532 | 5.087 | 6.542 | 4.715 | 6.125 |
| 4 | 1 | 17.296 | 25.644 | 17.548 | 25.101 | 17.902 | 24.162 |
| | 2 | 5.255 | 6.717 | 2.884 | 3.937 | 4.846 | 6.282 |
| 5 | 1 | 18.524 | 24.363 | 17.645 | 24.519 | 17.545 | 25.44 |
| | 2 | 3.112 | 4.133 | 4.895 | 6.321 | 4.862 | 6.301 |
| 6 | 1 | 17.699 | 25.27 | 18.425 | 24.173 | 17.47 | 23.887 |
| | 2 | 3.794 | 4.959 | 4.751 | 6.167 | 3.422 | 4.476 |
| 7 | 1 | 18.373 | 24.357 | 18.339 | 24.489 | 17.454 | 23.884 |
| | 2 | 3.979 | 5.222 | 3.487 | 4.575 | 4.662 | 6.074 |
| 8 | 1 | 18.407 | 24.306 | 17.631 | 24.679 | 17.495 | 23.996 |
| | 2 | 3.027 | 4.039 | 3.170 | 4.195 | 3.565 | 4.636 |
| 9 | 1 | 18.263 | 24.914 | 18.33 | 25.629 | 18.297 | 24.61 |
| | 2 | 3.683 | 4.827 | 5.319 | 6.811 | 3.341 | 4.388 |
| 10 | 1 | 17.429 | 25.287 | 17.814 | 25.446 | 18.255 | 25.306 |
| | 2 | 2.987 | 4.007 | 2.863 | 3.924 | 4.068 | 5.290 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 97 | 1 | 17.386 | 24.133 | 18.171 | 24.779 | 17.682 | 24.495 |
| | 2 | 5.330 | 6.847 | 4.723 | 6.143 | 3.625 | 4.726 |
| 98 | 1 | 17.473 | 24.295 | 17.773 | 24.128 | 17.575 | 24.697 |
| | 2 | 3.098 | 4.111 | 3.468 | 4.542 | 3.226 | 4.283 |
| 99 | 1 | 17.887 | 24.479 | 18.495 | 25.626 | 17.883 | 25.233 |
| | 2 | 3.040 | 4.057 | 4.341 | 5.650 | 3.412 | 4.459 |
| 100 | 1 | 17.803 | 24.553 | 18.32 | 25.142 | 17.564 | 24.677 |
| | 2 | 3.948 | 5.190 | 3.787 | 4.946 | 4.434 | 5.756 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

จากตารางที่ 9 - 12 พบว่า แบบสอบทั้ง 3 ฉบับ แต่ละฉบับมีสัดส่วนของข้อสอบทั้ง 2 ชนิด 4 สัดส่วน เมื่อทำการวิเคราะห์องค์ประกอบเพื่อตรวจสอบความเป็นเอกมิติ พบว่า แบบสอบทุกเงื่อนไข มีร้อยละของความแปรปรวนตั้งแต่ 23.862 พิจารณาแล้วเห็นว่ามีค่ามากพอที่จะสรุปว่ามีความเป็นเอกมิติประเภทที่มีตัวประกอบหลักตัวแรกเด่นกว่าตัวประกอบอื่น ๆ (Essential Unidimentionality) ตามเกณฑ์ของ Rackase (อ้างถึงใน Raju, 1993) และมีค่าค่าไอเกนสูงสุดมากกว่า 17 โดยมีอัตราส่วนระหว่างตัวประกอบที่ 1 และตัวที่ 2 มากกว่า 5.739 (วรนุช แหยมแสง, 2536) ที่ด้วยเหตุนี้จึงสรุปได้ว่า ข้อมูลสำหรับแบบสอบทุกเงื่อนไขมีความเป็นเอกมิติ เป็นไปตามข้อตกลงของทฤษฎีการตอบสนองข้อสอบ สามารถนำไปใช้ในการศึกษาคุณภาพการปรับเทียบคะแนนต่อไปได้

**การตรวจสอบคุณภาพด้านความเที่ยง**

การตรวจสอบคุณภาพของแบบสอบอีกด้านหนึ่งคือ การตรวจสอบคุภาพด้านความเที่ยงของแบบสอบ เพื่อรับประกันว่าข้อมูลที่ได้มาจากการจำลองข้อมูลนั้น มีคุณภาพเพียงพอต่อการนำไปศึกษาปรับเทียบ ผู้วิจัยทำการวิเคราะห์ประมาณค่าความเที่ยงของแบบสอบทั้งฉบับ ด้วยโปรแกรม MULTILOG ปรากฏผลดังนี้

ตารางที่ 13 ความเที่ยงของแบบสอบแต่ละฉบับ

| ครั้งที่ | 50_10 | | | 60_8 | | | 70_3 | | | 80_4 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B1 | B2 | A | B1 | B2 | A | B1 | B2 | A | B1 | B2 |
| 1 | 0.968 | 0.968 | 0.982 | 0.983 | 0.972 | 0.983 | 0.973 | 0.971 | 0.970 | 0.981 | 0.971 | 0.971 |
| 2 | 0.972 | 0.975 | 0.980 | 0.977 | 0.983 | 0.981 | 0.982 | 0.972 | 0.981 | 0.984 | 0.979 | 0.982 |
| 3 | 0.971 | 0.969 | 0.969 | 0.967 | 0.970 | 0.978 | 0.983 | 0.977 | 0.969 | 0.979 | 0.982 | 0.978 |
| 4 | 0.974 | 0.967 | 0.980 | 0.975 | 0.980 | 0.968 | 0.983 | 0.969 | 0.976 | 0.976 | 0.980 | 0.970 |
| 5 | 0.976 | 0.983 | 0.970 | 0.969 | 0.974 | 0.969 | 0.985 | 0.966 | 0.976 | 0.972 | 0.980 | 0.970 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 97 | 0.972 | 0.969 | 0.979 | 0.983 | 0.978 | 0.974 | 0.981 | 0.981 | 0.967 | 0.982 | 0.976 | 0.966 |
| 98 | 0.982 | 0.973 | 0.985 | 0.977 | 0.969 | 0.973 | 0.967 | 0.971 | 0.966 | 0.977 | 0.976 | 0.970 |
| 99 | 0.975 | 0.968 | 0.966 | 0.974 | 0.984 | 0.967 | 0.983 | 0.968 | 0.968 | 0.971 | 0.980 | 0.981 |
| 100 | 0.971 | 0.972 | 0.981 | 0.978 | 0.967 | 0.978 | 0.983 | 0.984 | 0.973 | 0.967 | 0.982 | 0.970 |

จากตารางที่ 13 พบว่า แบบสอบทุกฉบับมีความเที่ยงที่ได้จากการประมาณค่าโดยโปรแกรม MULTILOG ตั้งแต่ 0.96 แสดงว่าแบบสอบที่นำมาใช้ในการศึกษาการปรับเทียบครั้งนี้มีคุณภาพสูง

## ตอนที 4 การปรับเทียบคะแนน

การศึกษาประสิทธิภาพการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบสำหรับแบบสอบรูปแบบผสม 2 วิธีคือ วิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน มีรายละเอียดดังนี้

**การเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบวิธีโค้งคุณลักษณะ**

การปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะครั้งนี้ เป็นวิธีการปรับเทียบตามแนวคิดของ Heabara ที่ Kim (2004) ได้ขยายวิธีการปรับเทียบสู่แบบสอบรูปแบบผสมระหว่างโมเดลการตอบข้อสอบ 3PL และ GPC ซึ่งเป็นการแปลงค่าพารามิเตอร์ข้อสอบบนสเกลใหม่ไปสู่สเกลเก่าและ/หรือ แปลงค่าพารามิเตอร์ข้อสอบบนสเกลเก่าไปสู่สเกลใหม่ โดยเป็นการสรุปอ้างอิงจากเกณฑ์ฟังก์ชัน (criterion function) ตามแนวคิดของ Heabara ภายใต้โมเดลการตอบข้อสอบรูปแบบผสม ดังสมการ

$$Q = Q_1 + Q_2 \tag{1}$$

เมื่อ

$$Q_1 = \frac{1}{L_1}\sum_{i=1}^{N_1} W_1(\theta_{iO}) \sum_{m=1}^{M}\sum_{j=1}^{J_m}\sum_{k=1}^{K_{j(m)}} \left(\hat{P}_{ijkO(m)} - P^{*}_{ijkN(m)}\right)^2 \tag{2}$$

$$Q_2 = \frac{1}{L_2}\sum_{i=1}^{N_2} W_2(\theta_{iN}) \sum_{m=1}^{M}\sum_{j=1}^{J_m}\sum_{k=1}^{K_{j(m)}} \left(\hat{P}_{ijkN(m)} - \hat{P}^{\#}_{ijkO(m)}\right)^2 \tag{3}$$

โดยที่ M หมายถึง จำนวนโมเดลการตอบในแบบสอบรูปแบบผสม

$J_m$ หมายถึง จำนวนข้อสอบที่อยู่ภายในโมเดลการตอบที่ m

$K_{j(m)}$ หมายถึง จำนวนลำดับขั้นการตอบในข้อสอบข้อที่ j ภายใต้โมเดลที่ m

$$L_1 = \left(\sum_{i=1}^{N_1} W_1(\theta_{iO})\right) \times \left(\sum_{m=1}^{M}\sum_{j=1}^{J_m} K_{j(m)}\right), \quad \text{และ} \quad L_2 = \left(\sum_{i=1}^{N_2} W_2(\theta_{iN})\right) \times \left(\sum_{m=1}^{M}\sum_{j=1}^{J_m} K_{j(m)}\right)$$

ค่า $L_1$ และ $L_2$ เป็นองค์ประกอบในการปรับค่า $Q_1$ และ $Q_2$ ของเกณฑ์ฟังก์ชัน.ให้เป็นมาตรฐาน จากสมการที่ 1 สามารถใช้กับแบบสอบที่ตรวจให้คะแนนสองค่าได้ โดยให้พิจารณาโมเดลการตอบข้อสอบที่มีการให้คะแนนแบบสองค่าในมิติของโมเดลการตอบข้อสอบที่มีการให้คะแนนมากกว่าสองค่าที่มี 2 ลำดับขั้นการตอบ ซึ่งจากแนวคิดของ Heabara เป็นการคำนวณค่า

ความชันและค่าจุดตัด โดยการลดเกณฑ์ฟังก์ชัน Q ให้มีค่าต่ำที่สุด (minimizing criterion function, Q)

ในการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมด้วยวิธีโค้งคุณลักษณะในการวิจัยครั้งนี้ ผู้วิจัยปรับเทียบคะแนนด้วยโปรแกรม IRTEQ ที่พัฒนาโดย Han (2007) ซึ่งสามารถประมาณค่าสัมประสิทธิ์การปรับเทียบคะแนน และฟังก์ชันเกณฑ์ต่ำสุดได้

**การปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบด้วยวิธีปรับค่าพารามิเตอร์พร้อมกัน**

การปรับค่าพารามิเตอร์ของแบบสอบหลายฉบับพร้อม ๆ กัน โดยไม่ทราบค่าประมาณของพารามิเตอร์ข้อสอบร่วม ค่าประมาณทั้งหมดของข้อสอบทั้งที่เป็นข้อสอบร่วมและไม่ได้เป็นข้อสอบร่วมจะถูกเชื่อมโยงโดยอัตโนมัติไปสู่มาตรวัดร่วม (Mislevy and Bock, 1990 cited in Basteri, 2000) เช่น ประมาณค่าพารามิเตอร์แบบสอบฉบับ X และ Y พร้อมกัน โดยการ run เพียง 1 ครั้ง ทำโดยใช้ข้อสอบสำหรับฉบับ X และฉบับ Y จากผู้สอบทั้งสองกลุ่ม และข้อสอบที่ผู้สอบกลุ่มหนึ่งไม่ได้ทำจะใช้คำว่า “not reach” เมื่อมีการประมาณค่าในลักษณะนี้ ค่าประมาณที่ได้จะอยู่ในมาตรวัดเดียวกัน

การปรับค่าพารามิเตอร์พร้อมกัน เป็นการนำเข้าค่าประมาณพารามิเตอร์ 2 ชุดบนสเกลร่วม สำหรับการปรับค่าพารามิเตอร์พร้อมกัน ขั้นตอนแรกคือ การรวมข้อมูลการตอบจากกลุ่มผู้สอบทั้ง 2 กลุ่ม แล้วกำหนดให้ข้อสอบที่อีกกลุ่มหนึ่งไม่ได้ทำให้เป็นค่า missing หรือ not reach ในระหว่างนั้น ให้ระบุกลุ่มสมาชิก 1 หรือ 2 ในคอลัมภ์แรกของผลการตอบข้อสอบที่นำมารวมกัน ขั้นตอนที่ 2 คือ การเตรียมการเลือกสเกลก่อนที่จะทำการวิเคราะห์ (pre-select a scale before running the calibration) (ในที่นี้คือ สเกลความสามารถของผู้สอบกลุ่มที่ 1) โดยใช้ข้อสอบร่วมเป็นหลักเชื่อมโยง (anchor) ซึ่งโมเดล IRT ที่เป็นเอกมิติ ชนิด 3 PL เป็นข้อสอบที่ตรวจให้คะแนน 2 ค่า (MC) และโมเดล GPCM เป็นข้อสอบที่ตรวจให้คะแนนมากกว่า 2 ค่า (CR) ส่วนโปรแกรมคอมพิวเตอร์ MULTILOG (Thissen, 1991) ใช้เพื่อประมาณค่าพารามิเตอร์สำหรับข้อสอบทั้ง 2 รูปแบบคือ MC และ CR พร้อม ๆ กันในการรันโปรแกรมเพียงครั้งเดียว ซึ่งเป็นส่วนที่การันตีได้ว่าค่าประมาณพารามิเตอร์ทั้งหมดนั้นอยู่บนสเกลเดียวกัน เหมือนกัน การวิเคราะห์ใช้วิธีการประมาณค่าพารามิเตอร์แบบ Marginal Maximum likelihood (MML) และการประมาณค่าแบบ Maximum a Posteriori (MAP) ใช้เพื่อประมาณค่าพารามิเตอร์ความสามารถ ซึ่งเป็นการประมาณค่าที่โปรแกรม MULTILOG กำหนดตั้งต้นไว้แล้ว

## ตอนที 5 การประเมินคุณภาพของการปรับเทียบคะแนน

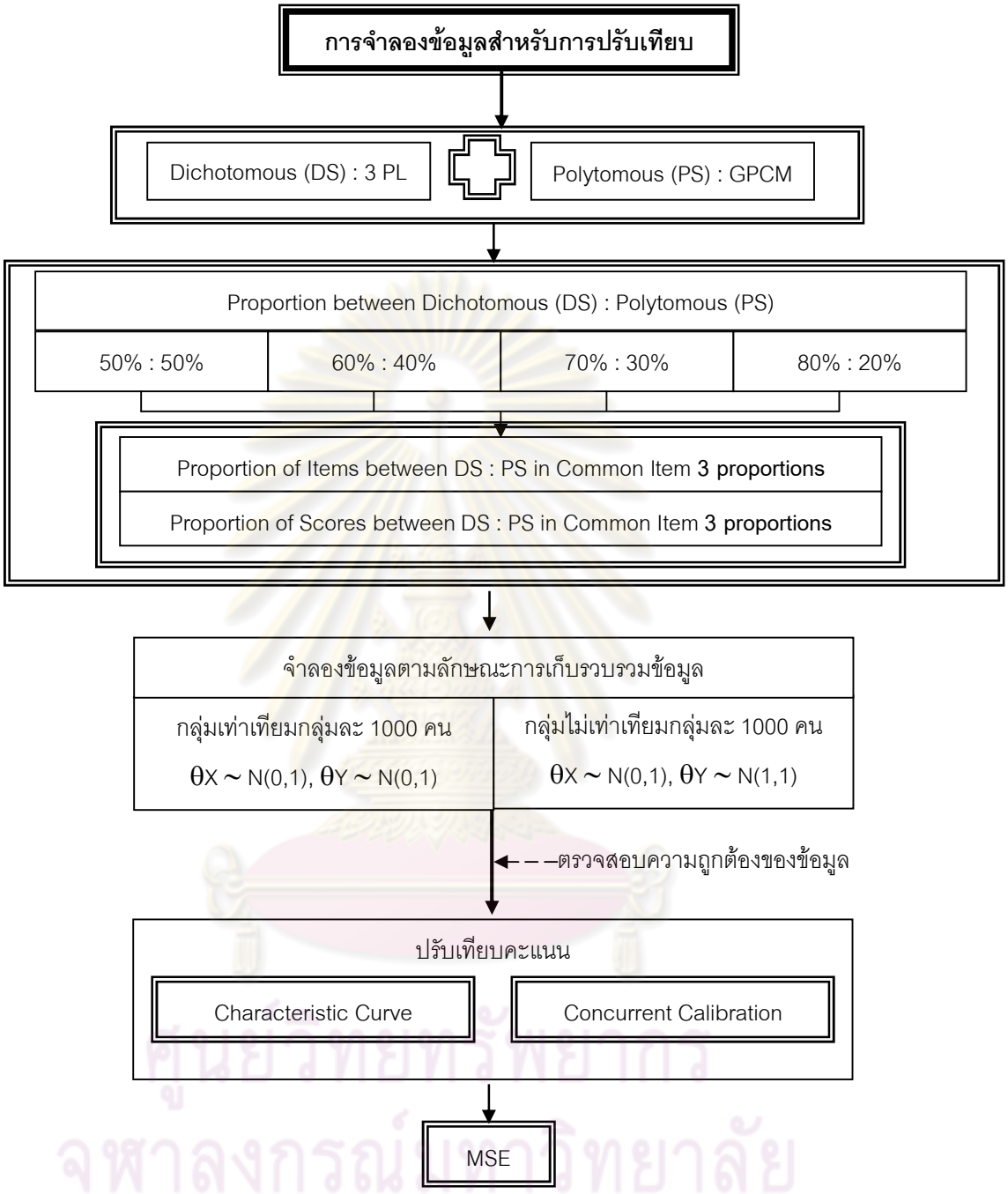
ในการศึกษาครั้งนี้เป็นการจำลองข้อมูลและมีการทำซ้ำ 100 ครั้ง สำหรับการเปรียบเทียบคุณภาพของการปรับเทียบคะแนนในการวิจัยครั้งนี้ พิจารณาจากเกณฑ์การประเมินจากเกณฑ์โค้งคุณลักษณะลำดับขั้นการตอบ (Category response curve : CRC) ซึ่งมีพื้นฐานมาจากความแตกต่างระหว่างโค้งลำดับขั้นการตอบที่ประมาณค่าได้ กับโค้งลำดับขั้นการตอบที่เป็นค่าที่แท้จริง เกณฑ์ CRC ที่นำมาใช้ในการเปรียบเทียบคุณภาพของการปรับเทียบนี้ ให้สารสนเทศที่เป็นประโยชน์และก่อให้เกิดความเข้าใจเป็นอย่างมากเกี่ยวกับประสิทธิภาพของกระบวนการเชื่อมโยงคะแนน ซึ่งเกณฑ์ CRC ที่ใช้นี้เป็นค่าของ MSE สำหรับประมาณค่าจากกาทำซ้ำ 100 ครั้ง

เกณฑ์ CRC คำนวณได้จาก สมการ

$$\int \frac{1}{n}\sum_{j=1}^{n}\frac{1}{K_j}\sum_{k=1}^{K_j}\left\{\frac{1}{100}\sum_{r=1}^{100}\left[P_{ij}(\theta)-\hat{P}_{(jk|r)}(\theta)\right]^2\right\}g(\theta)d(\theta)$$

ผู้วิจัยสรุปแผนภาพขั้นตอนการดำเนินการจำลองข้อมูลสำหรับปรับเทียบคะแนนคะแนนตามแนวทฤษฎีการตอบสนองข้อสอบสำหรับแบบสอบรูปแบบผสมดังแผนภาพที่ 17

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย




**แผนภาพที** 17 ขั้นตอนการจำลองข้อมูล

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

# บทที่ 4
# ผลการวิเคราะห์ข้อมูล

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อวิเคราะห์คุณภาพและเปรียบเทียบผลการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ (Characteristic Curve) และวิธีการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent Calibration) เมื่อแบบแผนการเก็บรวบรวมข้อมูล และสัดส่วนของข้อสอบทั้ง 2 ชนิดในแบบสอบรูปแบบผสม และสัดส่วนของข้อสอบทั้ง 2 ชนิดที่เป็นข้อสอบร่วมของแบบสอบรูปแบบผสมต่างกัน การนำเสนอผลการวิเคราะห์ข้อมูลแบ่งเป็น 2 ตอน ตอนที่ 1 เป็นผลการวิเคราะห์การปรับเทียบคะแนน และผลการเปรียบเทียบคุณภาพของการปรับเทียบคะแนน และตอนที่ 2 เป็นการสรุปผลการวิจัยตามวัตถุประสงค์

## สัญลักษณ์ที่ใช้ในการนำเสนอผลการวิเคราะห์ข้อมูล

| | | |
|---|---|---|
| a | หมายถึง | พารามิเตอร์อำนาจจำแนกของข้อสอบที่ตรวจให้คะแนนสองค่า |
| b | หมายถึง | พารามิเตอร์ความยากของข้อสอบที่ตรวจให้คะแนนสองค่า |
| c | หมายถึง | พารามิเตอร์การเดาของข้อสอบที่ตรวจให้คะแนนสองค่า |
| $\alpha$ | หมายถึง | พารามิเตอร์ความชันของข้อสอบที่มีการตรวจให้คะแนนหลายค่าตามโมเดล GPC |
| $\delta_{jk}$ | หมายถึง | ลำดับขั้นการตอบ |
| EQ | หมายถึง | การออกแบบการเก็บรวบรวมข้อมูลกลุ่มเท่าเทียม |
| NEQ | หมายถึง | การออกแบบการเก็บรวบรวมข้อมูลกลุ่มไม่เท่าเทียม |
| CC | หมายถึง | วิธีการปรับเทียบแบบโค้งคุณลักษณะด้วยวิธี Heabara |
| CO | หมายถึง | วิธีการปรับเทียบแบบปรับค่าพารามิเตอร์พร้อมกัน |
| MD | หมายถึง | ผลต่างของค่าเฉลี่ย (Mean different) |
| MSE | หมายถึง | ค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (Mean Square Error) |
| DS | หมายถึง | ข้อสอบที่มีการตรวจให้คะแนนสองค่า |
| PS | หมายถึง | ข้อสอบที่มีการตรวจให้คะแนนหลายค่า |
| Test A | หมายถึง | แบบสอบฐานสำหรับกลุ่มผู้สอบที่มีค่าเฉลี่ยความสามารถเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 |

Test B1 หมายถึง แบบสอบปรับเทียบสำหรับกลุ่มผู้สอบที่มีค่าเฉลี่ยความสามารถเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1

Test B2 หมายถึง แบบสอบปรับเทียบสำหรับกลุ่มผู้สอบที่มีค่าเฉลี่ยความสามารถเท่ากับ 1 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1

## ตอนที่ 1 ผลการวิเคราะห์การปรับเทียบและการเปรียบเทียบคุณภาพของการปรับเทียบ

คุณภาพของการปรับเทียบคะแนนตามแนวทฤษฎีการตอบสนองข้อสอบสำหรับแบบสอบรูปแบบผสมด้วยวิธีการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ (Heabara) และวิธีการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent Calibration) โดยพิจารณาจากเกณฑ์โค้งคุณลักษณะลำดับขั้นการตอบ (Category Response Curve Criterion : CRC) ที่มีพื้นฐานมาจากความแตกต่างระหว่างโค้งลำดับขั้นการตอบที่ประมาณค่าได้ (estimated category response curve) กับโค้งลำดับขั้นการตอบที่แท้จริง (true category response curve) ค่าที่ได้จากเกณฑ์ CRC ก็คือค่าดัชนี MSE นั่นเอง (Hanson and Beguin, 2002; Kim, 2004; Lim and Lee, 2006)

การศึกษาคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม **ผู้วิจัยนำเสนอผลการวิเคราะห์ข้อมูลเป็น 2 ส่วน คือ ส่วนที่ 1 ผลการวิเคราะห์คุณภาพการปรับเทียบคะแนน และส่วนที่ 2 ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนระหว่างแบบสอบ** มีรายละเอียดดังนี้

**ส่วนที่ 1 ผลการวิเคราะห์คุณภาพการปรับเทียบคะแนน**

ส่วนนี้ผู้วิจัยนำเสนอผลการวิเคราะห์คุณภาพการปรับเทียบคะแนนแบ่งเป็น 2 ตอน

1.1 เสนอผลการวิเคราะห์คุณภาพการปรับเทียบคะแนน**จำแนกตามเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิด** 4 เงื่อนไข ผู้วิจัยรายงานคุณภาพของการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ ด้วยวิธีโค้งคุณลักษณะและวิธีการปรับค่าพารามิเตอร์พร้อมกัน ภายใต้เงื่อนไขการออกแบบการเก็บรวบรวมข้อมูลแต่ละรูปแบบ และเงื่อนไขสัดส่วนข้อสอบร่วมแต่ละเงื่อนไข

1.2 เสนอผลการวิเคราะห์คุณภาพการปรับเทียบคะแนน**จำแนกตามเงื่อนไขสัดส่วนข้อสอบร่วม**ทั้งสองชนิด 6 เงื่อนไข ผู้วิจัยรายงานคุณภาพของการปรับเทียบคะแนนตามทฤษฎี

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

การตอบสนองข้อสอบ ด้วยวิธีโค้งคุณลักษณะและวิธีการปรับค่าพารามิเตอร์พร้อมกัน ภายใต้เงื่อนไขการออกแบบการเก็บรวบรวมข้อมูลแต่ละรูปแบบ และเงื่อนไขสัดส่วนข้อสอบแต่ละเงื่อนไข

1.1 เสนอผลการวิเคราะห์คุณภาพการปรับเทียบคะแนน**จำแนกตามเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิด** 4 เงื่อนไข ผู้วิจัยรายงานคุณภาพของการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ ด้วยวิธีโค้งคุณลักษณะและวิธีการปรับค่าพารามิเตอร์พร้อมกัน ภายใต้เงื่อนไขการออกแบบการเก็บรวบรวมข้อมูลแต่ละรูปแบบ และเงื่อนไขสัดส่วนข้อสอบร่วมแต่ละเงื่อนไข ผลการวิเคราะห์คุณภาพการปรับเทียบคะแนนทั้งสองวิธีจำแนกตามเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิด จำนวน 4 เงื่อนไข ปรากฏดังนี้

ตารางที่ 14 ค่า MSE จากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบ DS 50 : PS 10

| สัดส่วนข้อสอบร่วม | | วิธีการปรับเทียบ | | | |
|---|---|---|---|---|---|
| | | โค้งคุณลักษณะ | | การปรับค่าพารามิเตอร์พร้อมกัน | |
| | | EQ | NEQ | EQ | NEQ |
| ข้อสอบร่วมชนิดที่ 1 | 25/2 | 0.180 | 0.247 | 0.122 | 0.115 |
| | 20/2 | 0.172 | 0.228 | 0.127 | 0.099 |
| | 15/2 | 0.151 | 0.236 | 0.092 | 0.086 |
| ข้อสอบร่วมชนิดที่ 2 หลายค่า | 5/3 | 0.175 | 0.240 | 0.119 | 0.123 |
| | 10/2 | 0.198 | 0.253 | 0.106 | 0.125 |
| | 15/1 | 0.207 | 0.265 | 0.102 | 0.113 |

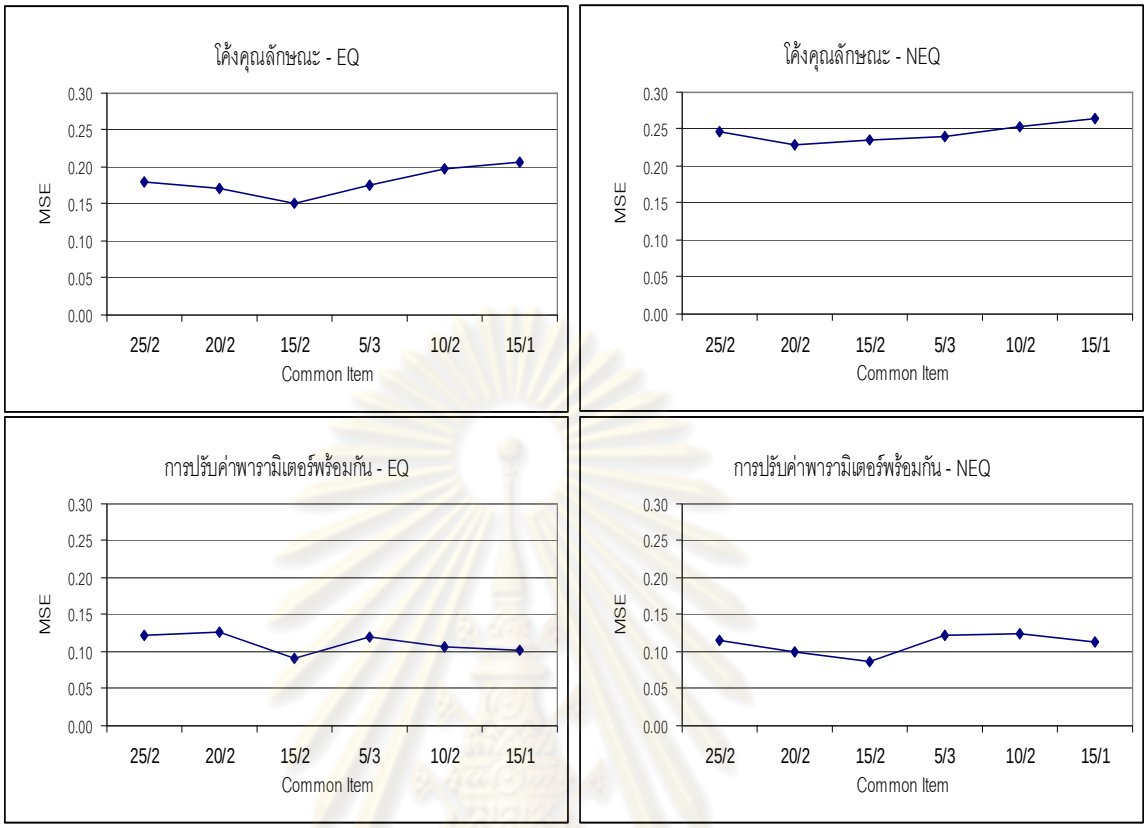
จากตารางที่ 14 พบว่า คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบทีตรวจให้คะแนนสองค่าจำนวน** 50 **ข้อ และข้อสอบทีตรวจให้คะแนนหลายค่าจำนวน** 10 **ข้อ ด้วยวิธีโค้งคุณลักษณะ** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.151 – 0.207 โดยเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 1 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่ไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.228 – 0.265 โดยเงื่อนไข

สัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 1 ข้อ

สำหรับคุณภาพของ**การปรับเทียบคะแนนด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีเท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.092 - 0.127 โดยเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 20 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 2 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.086 - 0.125 โดยเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 10 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 2 ข้อ

คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 10 ข้อ ด้วยวิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีรูปแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน 2 วิธี คือ กลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน และกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน ภายใต้เงื่อนไขสัดส่วนข้อสอบร่วมที่ต่างกัน 6 สัดส่วน แสดงดังแผนภาพที่ 18

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย




แผนภาพที่ 18 ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่าง DS 50 : PS 10

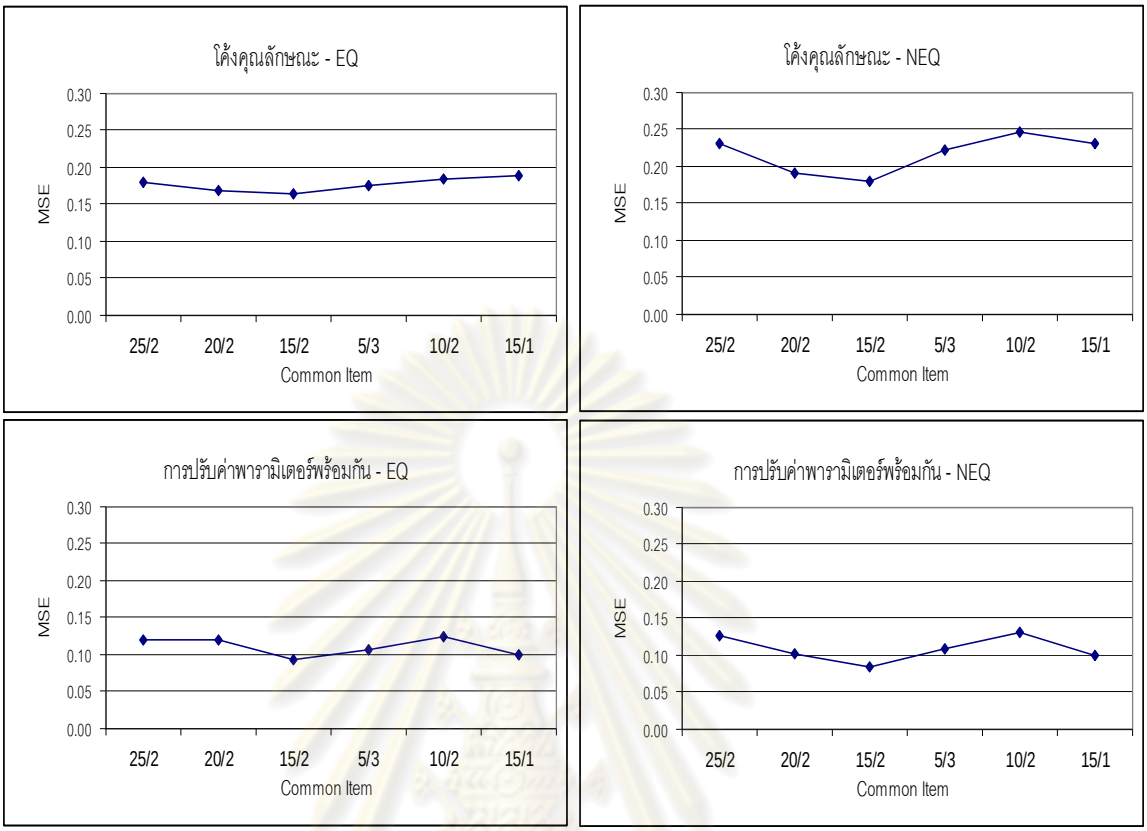
ตารางที่ 15 ค่า MSE จากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบ DS 60 : PS 8

| สัดส่วนข้อสอบร่วม | | วิธีการปรับเทียบ | | | |
|---|---|---|---|---|---|
| | | โค้งคุณลักษณะ | | การปรับค่าพารามิเตอร์พร้อมกัน | |
| | | EQ | NEQ | EQ | NEQ |
| ยูนิเขียง | 25/2 | 0.180 | 0.231 | 0.119 | 0.127 |
| | 20/2 | 0.169 | 0.192 | 0.119 | 0.103 |
| | 15/2 | 0.165 | 0.181 | 0.093 | 0.084 |
| หนหนา | 5/3 | 0.175 | 0.222 | 0.107 | 0.109 |
| | 10/2 | 0.185 | 0.246 | 0.125 | 0.132 |
| | 15/1 | 0.190 | 0.232 | 0.099 | 0.100 |

จากตารางที่ 15 พบว่า คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบทีตรวจให้คะแนนสองค่าจำนวน** 60 **ข้อ และข้อสอบทีตรวจให้คะแนนหลายค่าจำนวน** 8 **ข้อ ด้วยวิธีโค้งคุณลักษณะ** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.165 - 0.190 โดยเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 1 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่ไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.181 - 0.246 โดยเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 10 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 2 ข้อ

สำหรับคุณภาพของ**การปรับเทียบคะแนนด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีเท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.093 - 0.125 โดยเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 10 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 2 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.103 - 0.132 โดยเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 10 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 2 ข้อ

คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 60 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 8 ข้อ ด้วยวิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีรูปแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน 2 วิธี คือ กลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน และกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน ภายใต้เงื่อนไขสัดส่วนข้อสอบร่วมที่ต่างกัน 6 สัดส่วน แสดงดังแผนภาพที่ 19




แผนภาพที่ 19 ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่าง DS 60 : PS 8

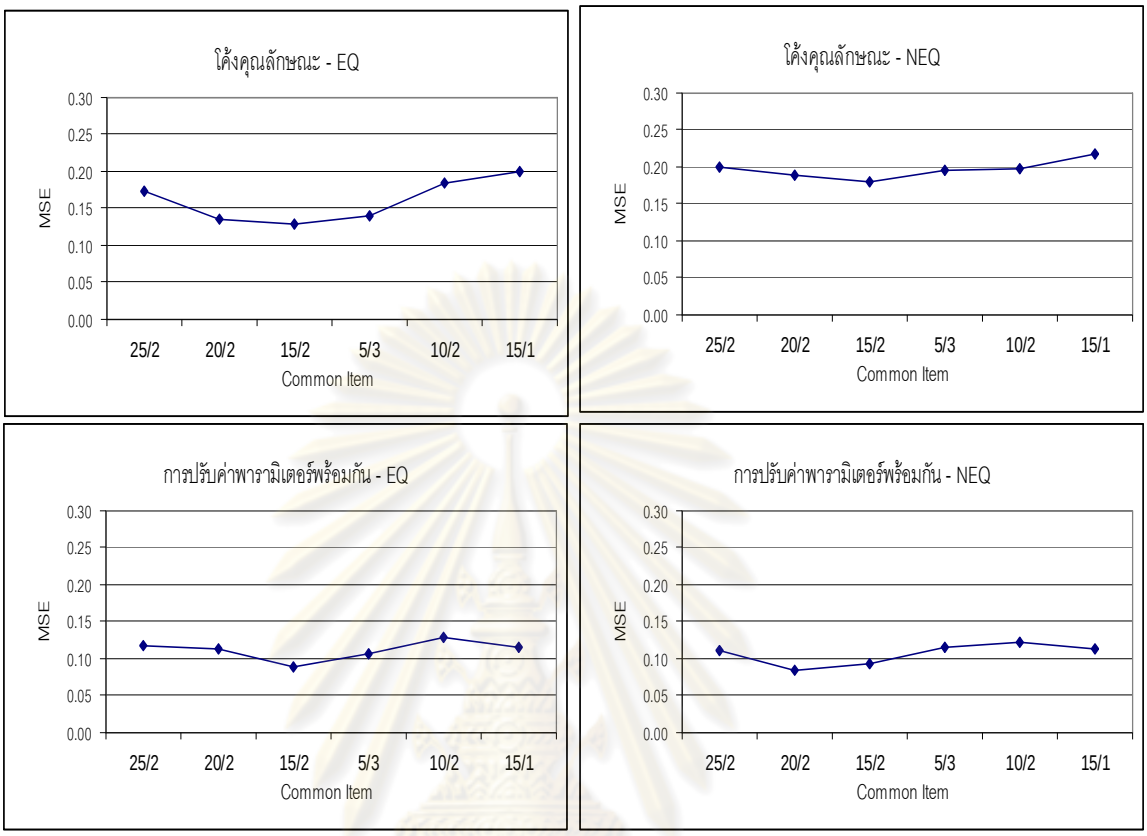
ตารางที่ 16 ค่า MSE จากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบ DS 70 : PS 6

| สัดส่วนข้อสอบร่วม | | วิธีการปรับเทียบ | | | |
|---|---|---|---|---|---|
| | | โค้งคุณลักษณะ | | การปรับค่าพารามิเตอร์พร้อมกัน | |
| | | EQ | NEQ | EQ | NEQ |
| ยูนิโพลาร์ | 25/2 | 0.174 | 0.201 | 0.118 | 0.111 |
| | 20/2 | 0.136 | 0.189 | 0.114 | 0.084 |
| | 15/2 | 0.129 | 0.179 | 0.089 | 0.094 |
| หลายมิติ | 5/3 | 0.139 | 0.195 | 0.106 | 0.115 |
| | 10/2 | 0.185 | 0.198 | 0.128 | 0.123 |
| | 15/1 | 0.199 | 0.217 | 0.115 | 0.114 |

จากตารางที่ 16 พบว่า คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบทีตรวจให้คะแนนสองค่าจำนวน** 70 **ข้อ และข้อสอบทีตรวจให้คะแนนหลายค่าจำนวน** 6 **ข้อ ด้วยวิธีโค้งคุณลักษณะ** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.129 - 0.199 โดยเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 1 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่ไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.179 - 0.217 โดยเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 1 ข้อ

สำหรับคุณภาพของ**การปรับเทียบคะแนนด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีเท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.089 - 0.128 โดยเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 10 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 2 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.084 - 0.123 โดยเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 10 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 2 ข้อ

คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 70 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 6 ข้อ ด้วยวิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีรูปแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน 2 วิธี คือ กลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน และกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน ภายใต้เงื่อนไขสัดส่วนข้อสอบร่วมที่ต่างกัน 6 สัดส่วน แสดงดังแผนภาพที่ 20




แผนภาพที่ 20 ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่าง DS 70 : PS 6

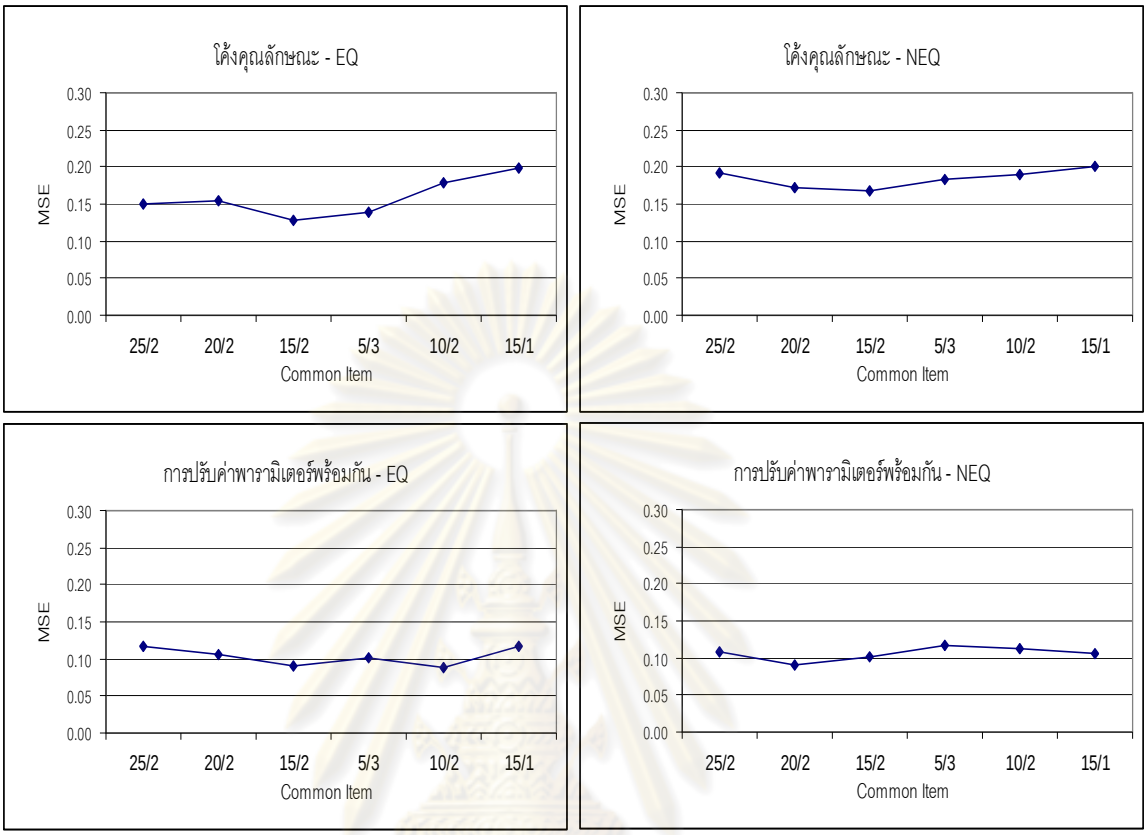
ตารางที่ 17 ค่า MSE จากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบ DS 80 : PS 4

| สัดส่วนข้อสอบร่วม | | วิธีการปรับเทียบ | | | |
|---|---|---|---|---|---|
| | | โค้งคุณลักษณะ | | การปรับค่าพารามิเตอร์พร้อมกัน | |
| | | EQ | NEQ | EQ | NEQ |
| กลุ่มเดียว | 25/2 | 0.150 | 0.191 | 0.116 | 0.109 |
| | 20/2 | 0.155 | 0.172 | 0.105 | 0.091 |
| | 15/2 | 0.129 | 0.168 | 0.090 | 0.101 |
| กลุ่มสอง | 5/3 | 0.139 | 0.184 | 0.102 | 0.117 |
| | 10/2 | 0.178 | 0.189 | 0.089 | 0.113 |
| | 15/1 | 0.198 | 0.200 | 0.118 | 0.106 |

จากตารางที่ 17 พบว่า คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบทีตรวจให้คะแนนสองค่าจำนวน** 80 **ข้อ และข้อสอบทีตรวจให้คะแนนหลายค่าจำนวน** 4 **ข้อ ด้วยวิธีโค้งคุณลักษณะ** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.129 - 0.198 โดยเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 1 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่ไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.168 - 0.200 โดยเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 1 ข้อ

สำหรับคุณภาพของ**การปรับเทียบคะแนนด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีเท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.090 - 0.118 โดยเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 1 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.091 - 0.117 โดยเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 5 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 3 ข้อ

คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 80 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 4 ข้อ ด้วยวิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีรูปแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน 2 วิธี คือ กลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน และกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน ภายใต้เงื่อนไขสัดส่วนข้อสอบร่วมที่ต่างกัน 6 สัดส่วน แสดงดังแผนภาพที่ 21




แผนภาพที่ 21 ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่าง DS 80 : PS 4

1.2 เสนอผลการวิเคราะห์คุณภาพการปรับเทียบคะแนน**จำแนกตามเงื่อนไขสัดส่วนข้อสอบร่วม**ทั้งสองชนิด จำนวน 6 เงื่อนไข ผู้วิจัยรายงานคุณภาพของการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ ด้วยวิธีโค้งคุณลักษณะและวิธีการปรับค่าพารามิเตอร์พร้อมกัน ภายใต้เงื่อนไขการออกแบบการเก็บรวบรวมข้อมูลแต่ละรูปแบบ และเงื่อนไขสัดส่วนข้อสอบแต่ละเงื่อนไข ผลการวิเคราะห์คุณภาพการปรับเทียบคะแนนทั้งสองวิธีจำแนกตามเงื่อนไขสัดส่วนข้อสอบร่วมทั้งสองชนิด จำนวน 6 เงื่อนไข ปรากฏดังตารางที่ 18-23

ตารางที่ 18 ค่า MSE จากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบร่วม DS 25 : PS 2

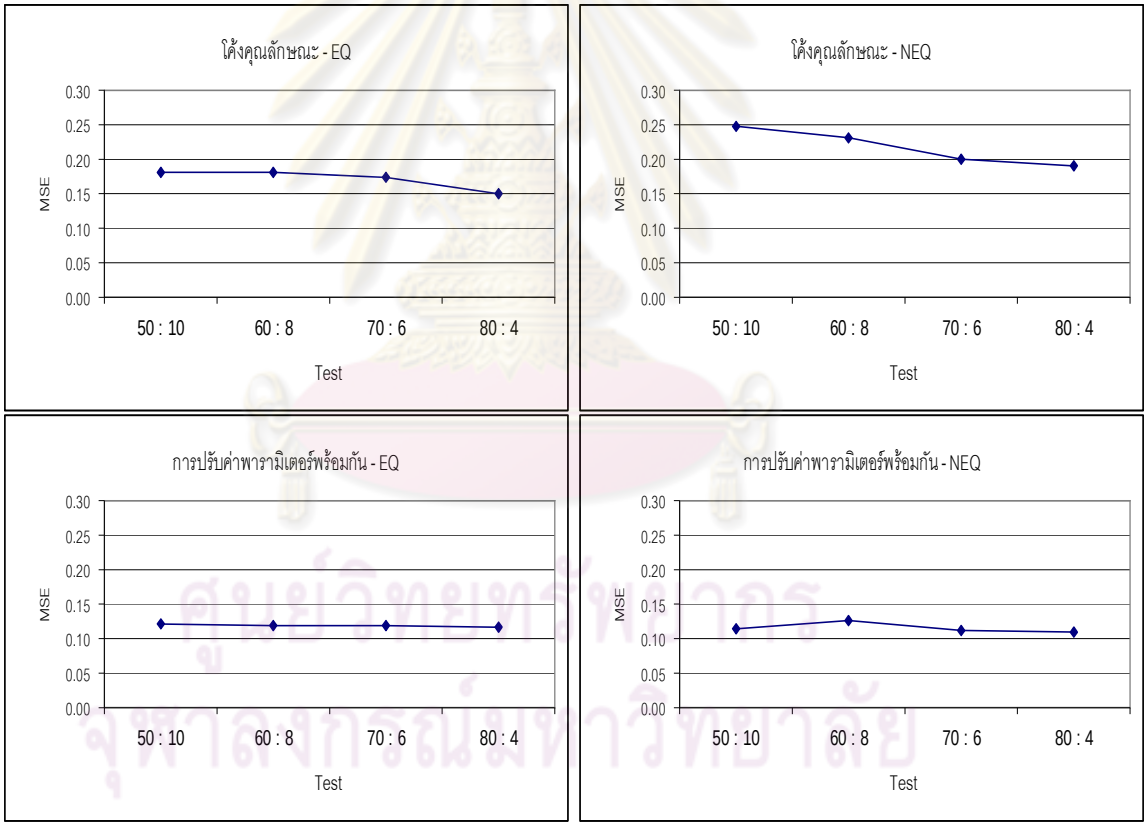
| สัดส่วนข้อสอบในแบบสอบ | วิธีการปรับเทียบ | | | |
|---|---|---|---|---|
| | โค้งคุณลักษณะ | | การปรับค่าพารามิเตอร์พร้อมกัน | |
| | EQ | NEQ | EQ | NEQ |
| DS 50 : PS 10 | 0.180 | 0.247 | 0.122 | 0.115 |
| DS 60 : PS 8 | 0.180 | 0.231 | 0.119 | 0.127 |
| DS 70 : PS 6 | 0.174 | 0.201 | 0.118 | 0.111 |
| DS 80 : PS 4 | 0.150 | 0.191 | 0.116 | 0.109 |

จากตารางที่ 18 พบว่า คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบร่วมในสัดส่วน** 25 : 2 **ด้วยวิธีโค้งคุณลักษณะ** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีเท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.150 - 0.180 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.191 - 0.247 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ

สำหรับคุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบร่วมในสัดส่วน** 25 : 2 **ด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.116 - 0.122 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่ม

ผู้สอบสอง**กลุ่มที่ไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.109 - 0.127 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 60 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 8 ข้อ

คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีสัดส่วนข้อสอบร่วมเท่ากับ 25 : 2 ด้วยวิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีรูปแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน 2 วิธี คือ กลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน และกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน ภายใต้เงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ต่างกัน 4 เงื่อนไข แสดงดังแผนภาพที่ 22




แผนภาพที่ 22 ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วม DS 25 : PS 2

ตารางที่ 19 ค่า MSE จากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบร่วม DS 20 : PS 2

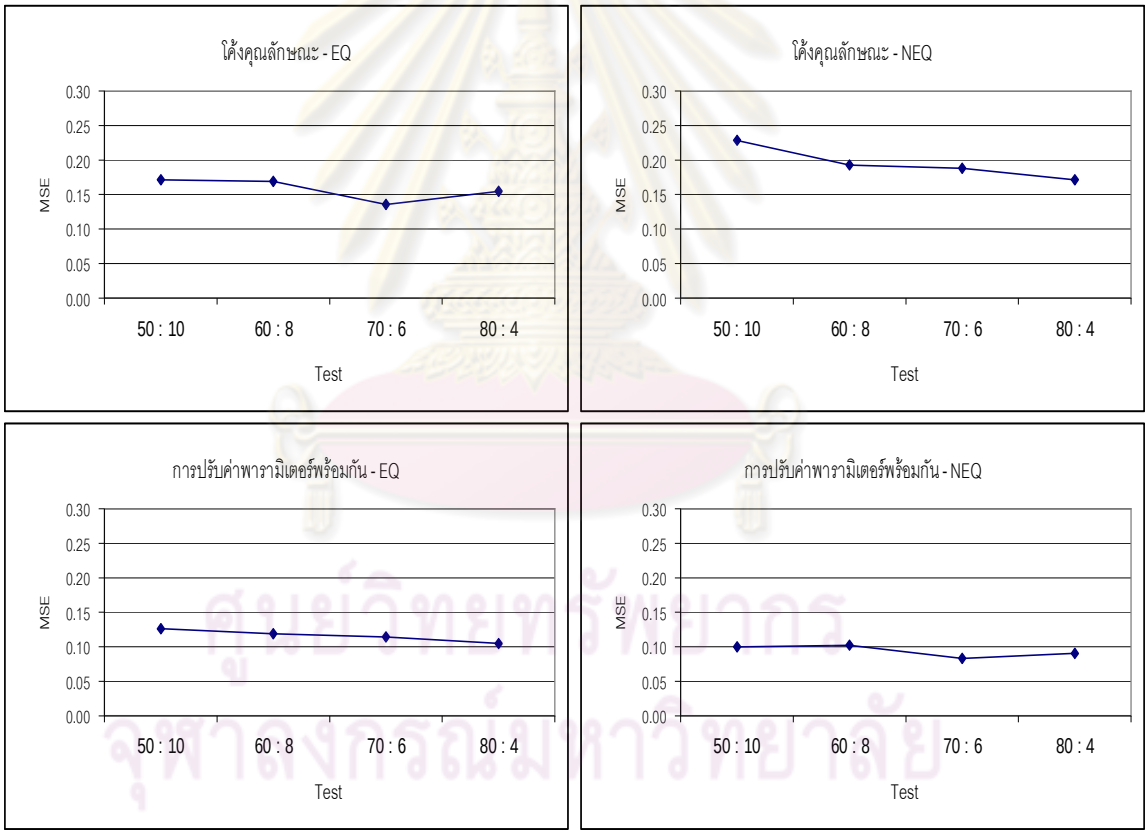
| สัดส่วนข้อสอบในแบบสอบ | วิธีการปรับเทียบ | | | |
|---|---|---|---|---|
| | โค้งคุณลักษณะ | | การปรับค่าพารามิเตอร์พร้อมกัน | |
| | EQ | NEQ | EQ | NEQ |
| DS 50 : PS 10 | 0.172 | 0.228 | 0.127 | 0.099 |
| DS 60 : PS 8 | 0.169 | 0.192 | 0.119 | 0.103 |
| DS 70 : PS 6 | 0.136 | 0.189 | 0.114 | 0.084 |
| DS 80 : PS 4 | 0.155 | 0.172 | 0.105 | 0.091 |

จากตารางที่ 19 พบว่า คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบร่วมในสัดส่วน** 20 : 2 **ด้วยวิธีโค้งคุณลักษณะ** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีเท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.136 - 0.172 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.172 - 0.228 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ

สำหรับคุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบร่วมในสัดส่วน** 20 : 2 **ด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.105 - 0.127 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่ม

ผู้สอบสอง**กลุ่มที่ไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.084 - 0.103 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 60 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 8 ข้อ

คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีสัดส่วนข้อสอบร่วมเท่ากับ 20 : 2 ด้วยวิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีรูปแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน 2 วิธี คือ กลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน และกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน ภายใต้เงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ต่างกัน 4 เงื่อนไข แสดงดังแผนภาพที่ 23




แผนภาพที่ 23 ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วม DS 20 : PS 2

ตารางที่ 20 ค่า MSE จากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบร่วม DS 15 : PS 2

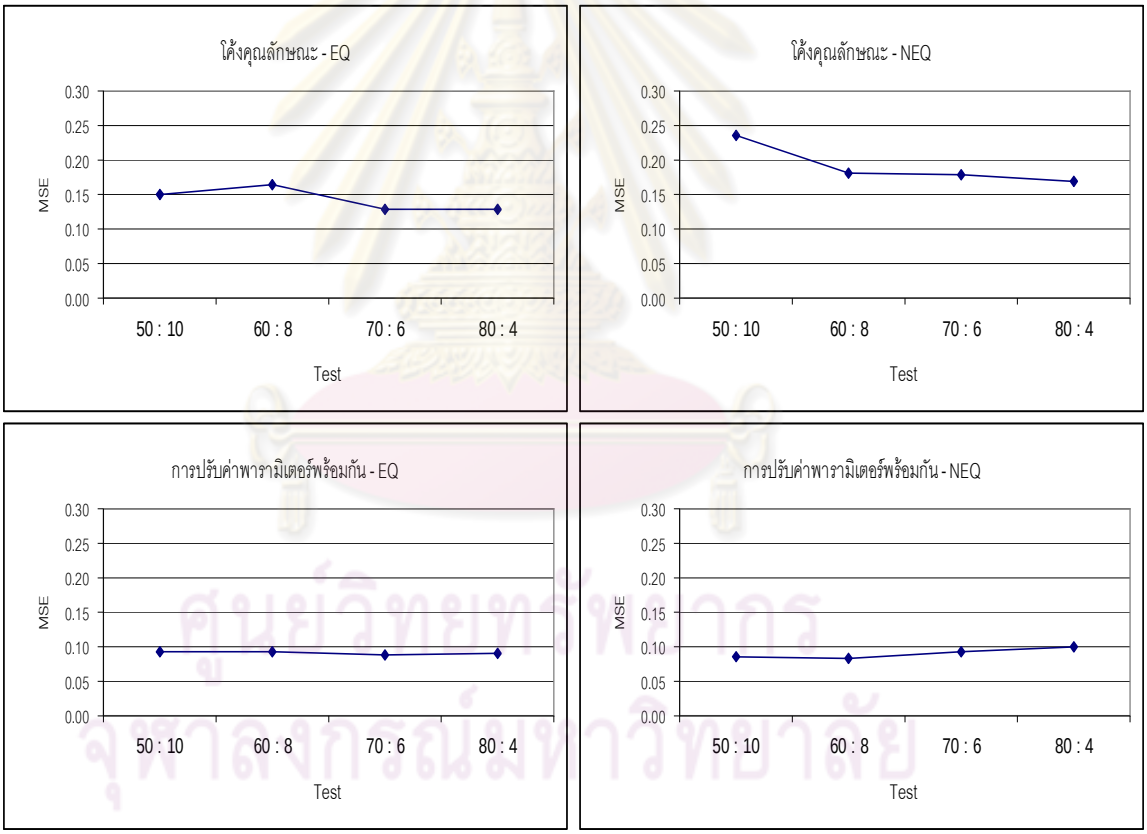
| สัดส่วนข้อสอบในแบบสอบ | วิธีการปรับเทียบ | | | |
|---|---|---|---|---|
| | โค้งคุณลักษณะ | | การปรับค่าพารามิเตอร์พร้อมกัน | |
| | EQ | NEQ | EQ | NEQ |
| DS 50 : PS 10 | 0.151 | 0.236 | 0.092 | 0.086 |
| DS 60 : PS 8 | 0.165 | 0.181 | 0.093 | 0.084 |
| DS 70 : PS 6 | 0.129 | 0.179 | 0.089 | 0.094 |
| DS 80 : PS 4 | 0.129 | 0.168 | 0.090 | 0.101 |

จากตารางที่ 20 พบว่า คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบร่วมในสัดส่วน** 15 : 2 **ด้วยวิธีโค้งคุณลักษณะ** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีเท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.129 - 0.165 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 60 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 8 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.168 - 0.236 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ

สำหรับคุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบร่วมในสัดส่วน** 15 : 2 **ด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.089 - 0.093 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 60 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 8 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่ม

ผู้สอบสอง**กลุ่มที่ไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.084 - 0.101 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 80 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 4 ข้อ

คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีสัดส่วนข้อสอบร่วมเท่ากับ 15 : 2 ด้วยวิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีรูปแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน 2 วิธี คือ กลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน และกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน ภายใต้เงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ต่างกัน 4 เงื่อนไข แสดงดังแผนภาพที่ 24




แผนภาพที่ 24 ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วม DS 15 : PS 2

ตารางที่ 21 ค่า MSE จากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบร่วม DS 5 : PS 3

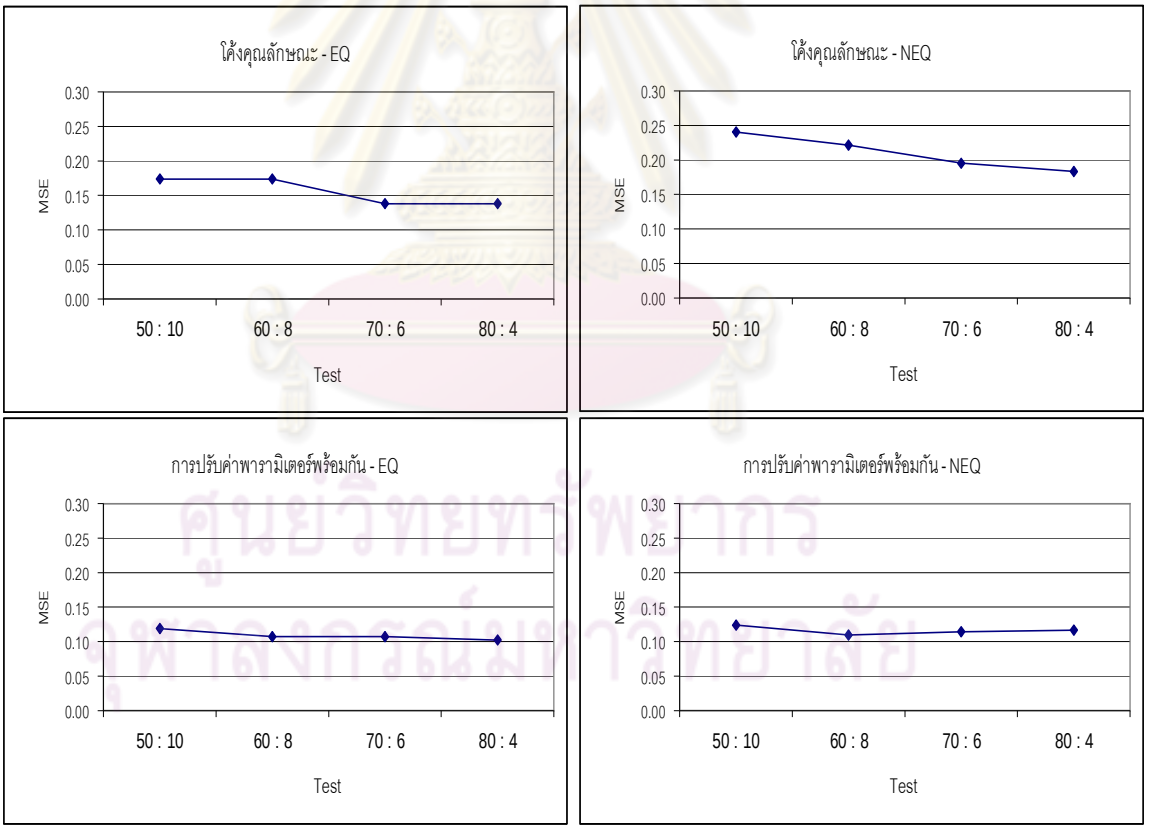
| สัดส่วนข้อสอบในแบบสอบ | วิธีการปรับเทียบ | | | |
|---|---|---|---|---|
| | โค้งคุณลักษณะ | | การปรับค่าพารามิเตอร์พร้อมกัน | |
| | EQ | NEQ | EQ | NEQ |
| DS 50 : PS 10 | 0.175 | 0.240 | 0.119 | 0.123 |
| DS 60 : PS 8 | 0.175 | 0.222 | 0.107 | 0.109 |
| DS 70 : PS 6 | 0.139 | 0.195 | 0.106 | 0.115 |
| DS 80 : PS 4 | 0.139 | 0.184 | 0.102 | 0.117 |

จากตารางที่ 21 พบว่า คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบร่วมในสัดส่วน** 5 : 3 **ด้วยวิธีโค้งคุณลักษณะ** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีเท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.139 - 0.175 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุดมี 2 สัดส่วน คือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ และสัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 60 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 8 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่ไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.184 - 0.240 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ

สำหรับคุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบร่วมในสัดส่วน** 5 : 3 **ด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีเท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.102 - 0.119 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วย

ข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่ไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.109 - 0.123 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ

คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีสัดส่วนข้อสอบร่วมเท่ากับ 5 : 3 ด้วยวิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีรูปแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน 2 วิธี คือ กลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน และกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน ภายใต้เงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ต่างกัน 4 เงื่อนไข แสดงดังแผนภาพที่ 25




แผนภาพที่ 25 ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วม DS 5 : PS 3

ตารางที่ 22 ค่า MSE จากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบร่วม DS 10 : PS 2

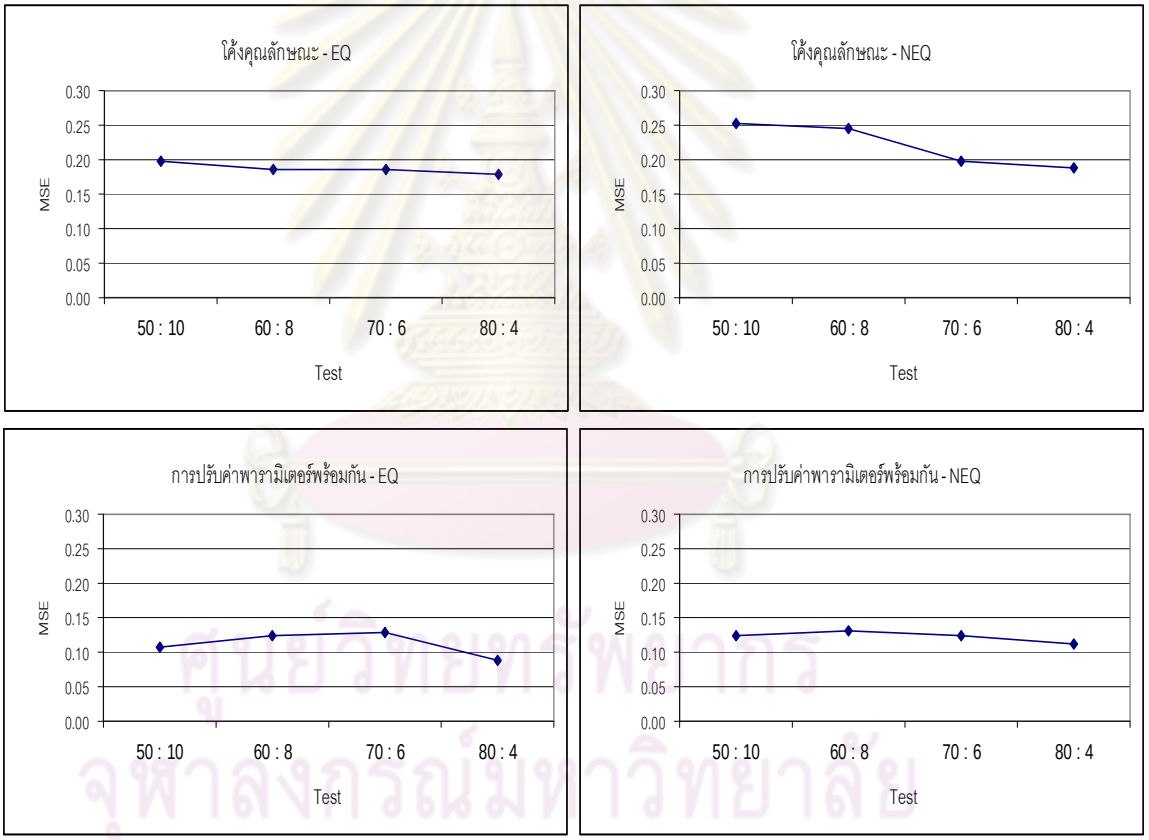
| สัดส่วนข้อสอบในแบบสอบ | วิธีการปรับเทียบ | | | |
|---|---|---|---|---|
| | โค้งคุณลักษณะ | | การปรับค่าพารามิเตอร์พร้อมกัน | |
| | EQ | NEQ | EQ | NEQ |
| DS 50 : PS 10 | 0.198 | 0.253 | 0.106 | 0.125 |
| DS 60 : PS 8 | 0.185 | 0.246 | 0.125 | 0.132 |
| DS 70 : PS 6 | 0.185 | 0.198 | 0.128 | 0.123 |
| DS 80 : PS 4 | 0.178 | 0.189 | 0.089 | 0.113 |

จากตารางที่ 22 พบว่า คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบร่วมในสัดส่วน** 10 : 2 **ด้วยวิธีโค้งคุณลักษณะ** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีเท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.178 - 0.198 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุด คือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่ไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.189 - 0.253 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ

สำหรับคุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบร่วมในสัดส่วน** 10 : 2 **ด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.089 - 0.128 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 70 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 6 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่ม

ผู้สอบสอง**กลุ่มที่ไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.113 - 0.132 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 60 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 8 ข้อ

คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีสัดส่วนข้อสอบร่วมเท่ากับ 10 : 2 ด้วยวิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีรูปแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน 2 วิธี คือ กลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน และกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน ภายใต้เงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ต่างกัน 4 เงื่อนไข แสดงดังแผนภาพที่ 26




แผนภาพที่ 26 ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วม DS 10 : PS 2

ตารางที่ 23 ค่า MSE จากการปรับเทียบด้วยแบบสอบตามเงื่อนไขสัดส่วนข้อสอบร่วม DS 15 : PS 1

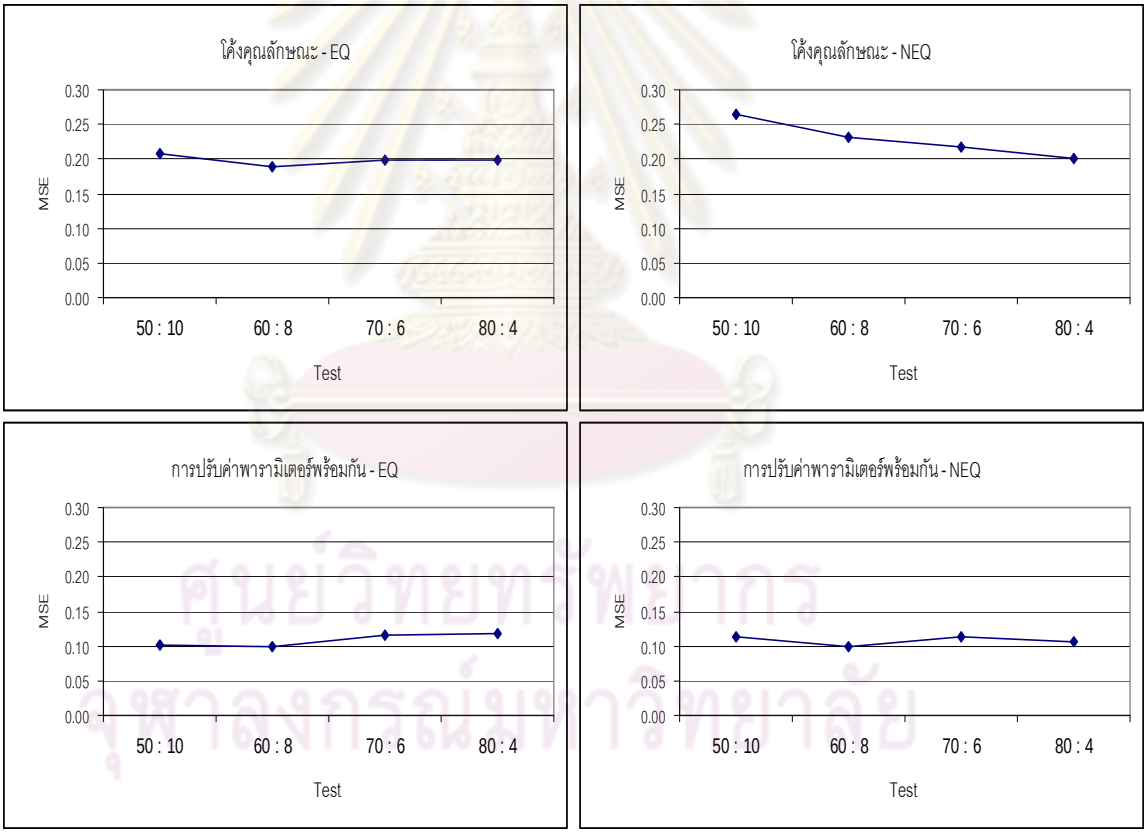
| สัดส่วนข้อสอบในแบบสอบ | วิธีการปรับเทียบ | | | |
|---|---|---|---|---|
| | โค้งคุณลักษณะ | | การปรับค่าพารามิเตอร์พร้อมกัน | |
| | EQ | NEQ | EQ | NEQ |
| DS 50 : PS 10 | 0.207 | 0.265 | 0.102 | 0.113 |
| DS 60 : PS 8 | 0.190 | 0.232 | 0.099 | 0.100 |
| DS 70 : PS 6 | 0.199 | 0.217 | 0.115 | 0.114 |
| DS 80 : PS 4 | 0.198 | 0.200 | 0.118 | 0.106 |

จากตารางที่ 23 พบว่า คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบร่วมในสัดส่วน** 15 : 1 **ด้วยวิธีโค้งคุณลักษณะ** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีเท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.190 - 0.207 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุด คือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่ไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.200 - 0.265 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ

สำหรับคุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบร่วมในสัดส่วน** 15 : 1 **ด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.099 - 0.118 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 80 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 4 ข้อ ส่วนคุณภาพของการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่เก็บรวบรวมข้อมูลกับกลุ่ม

ผู้สอบสอง**กลุ่มทีไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบร่วมแตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนอยู่ระหว่าง 0.100 - 0.114 โดยเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 70 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 6 ข้อ

คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีสัดส่วนข้อสอบร่วมเท่ากับ 15 : 1 ด้วยวิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีรูปแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน 2 วิธี คือ กลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน และกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน ภายใต้เงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ต่างกัน 4 เงื่อนไข แสดงดังแผนภาพที่ 27




แผนภาพที่ 27 ผลการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วม DS 15 : PS 1

**ส่วนที่ 2 ผลการเปรียบเทียบคุณภาพของการปรับเทียบระหว่างแบบสอบ**

การเปรียบเทียบคุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม 2 วิธี คือ การปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ (Characteristic Curve) และวิธีการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent Calibration) เมื่อแบบแผนการเก็บรวบรวมข้อมูล และสัดส่วนของข้อสอบทั้ง 2 ชนิดในแบบสอบรูปแบบผสม และสัดส่วนของข้อสอบทั้ง 2 ชนิดที่เป็นข้อสอบร่วมต่างกัน ผู้วิจัยพิจารณาจากค่าดัชนี MSE ที่ได้มาจากความแตกต่างระหว่างโค้งลำดับขั้นการตอบที่ประมาณค่าได้ (estimated category response curve) กับโค้งลำดับขั้นการตอบที่แท้จริง (true category response curve) บนแบบสอบฉบับเทียบคะแนน (new form) ซึ่งเป็นลักษณะของเกณฑ์คุณลักษณะลำดับขั้นการตอบ (Characteristic response curve : CRC) (Hanson and Beguin, 2002; Kim, 2004, Kim and Lee, 2006) ผลการเปรียบเทียบคุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม ผู้วิจัยนำเสนอเป็น 2 ส่วน คือ ส่วนแรกเป็นผลการเปรียบเทียบคุณภาพการปรับเทียบทั้ง 2 วิธี โดยพิจารณาจากค่า MSE และนำเสนอในรูปของกราฟแสดงผลการเปรียบเทียบ และส่วนที่ 2 เสนอผลการปรับเทียบทั้ง 2 วิธี โดยการวิเคราะห์เปรียบเทียบค่าเฉลี่ยระหว่างกลุ่ม รายละเอียดดังนี้

**2.1 ผลการเปรียบเทียบคุณภาพการปรับเทียบทั้ง 2 วิธี**

การเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนหลายค่า ผู้วิจัยเสนอค่า MSE ที่แสดงถึงคุณภาพของการปรับเทียบคะแนนแต่ละวิธีภายใต้เงื่อนไขที่ศึกษาได้แก่ การออกแบบการเก็บรวบรวมข้อมูล 2 รูปแบบคือ กลุ่มเท่าเทียม และกลุ่มไม่เท่าเทียม สัดส่วนของข้อสอบทั้งสองชนิดในแบบสอบ 4 สัดส่วน และสัดส่วนของข้อสอบทั้งสองชนิดในข้อสอบร่วม 6 สัดส่วน ปรากฏดังตารางที่ 24

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 24 ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม

| สัดส่วนข้อสอบในแบบสอบ | สัดส่วนข้อสอบร่วม | โค้งคุณลักษณะ | | การปรับค่าพารามิเตอร์พร้อมกัน | |
|---|---|---|---|---|---|
| | | EQ | NEQ | EQ | NEQ |
| DS 50 : PS 10 | 25/2 | 0.180 | 0.247 | 0.122 | 0.115 |
| | 20/2 | 0.172 | 0.228 | 0.127 | 0.099 |
| | 15/2 | 0.151 | 0.236 | 0.092 | 0.086 |
| | 5/3 | 0.175 | 0.240 | 0.119 | 0.123 |
| | 10/2 | 0.198 | 0.253 | 0.106 | 0.125 |
| | 15/1 | 0.207 | 0.265 | 0.102 | 0.113 |
| DS 60 : PS 8 | 25/2 | 0.180 | 0.231 | 0.119 | 0.127 |
| | 20/2 | 0.169 | 0.192 | 0.119 | 0.103 |
| | 15/2 | 0.165 | 0.181 | 0.093 | 0.084 |
| | 5/3 | 0.175 | 0.222 | 0.107 | 0.109 |
| | 10/2 | 0.185 | 0.246 | 0.125 | 0.132 |
| | 15/1 | 0.190 | 0.232 | 0.099 | 0.100 |
| DS 70 : PS 6 | 25/2 | 0.174 | 0.201 | 0.118 | 0.111 |
| | 20/2 | 0.136 | 0.189 | 0.114 | 0.084 |
| | 15/2 | 0.129 | 0.179 | 0.089 | 0.094 |
| | 5/3 | 0.139 | 0.195 | 0.106 | 0.115 |
| | 10/2 | 0.185 | 0.198 | 0.128 | 0.123 |
| | 15/1 | 0.199 | 0.217 | 0.115 | 0.114 |
| DS 80 : PS 4 | 25/2 | 0.150 | 0.191 | 0.116 | 0.109 |
| | 20/2 | 0.155 | 0.172 | 0.105 | 0.091 |
| | 15/2 | 0.129 | 0.168 | 0.090 | 0.101 |
| | 5/3 | 0.139 | 0.184 | 0.102 | 0.117 |
| | 10/2 | 0.178 | 0.189 | 0.089 | 0.113 |
| | 15/1 | 0.198 | 0.200 | 0.118 | 0.106 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

จากตารางที่ 24 ผู้วิจัยเสนอค่า MSE ที่แสดงถึงความคลาดเคลื่อนในการปรับเทียบคะแนน จำแนกตามเงื่อนไขที่ทำการศึกษา คือ วิธีการปรับเทียบ 2 วิธี คือ วิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน การออกแบบการเก็บรวบรวมข้อมูล 2 รูปแบบ คือ กลุ่มเท่าเทียม และกลุ่มไม่เท่าเทียม สัดส่วนข้อสอบทั้ง 2 ชนิดในแบบสอบ และสัดส่วนข้อสอบทั้ง 2 ชนิดในข้อสอบร่วม ผู้วิจัยจึงเสนอผลการเปรียบเทียบคุณภาพของการปรับเทียบตามเงื่อนไขที่ศึกษา ซึ่งสัญลักษณ์ที่ใช้ในการนำเสนอ และรายละเอียดของผลการเปรียบเทียบคุณภาพการปรับเทียบมีดังนี้

CC – EQ หมายถึง ผลการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่มีการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่ม ที่มีความสามารถเท่าเทียมกัน

CC – NEQ หมายถึง ผลการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ ที่มีการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่ม ที่มีความสามารถไม่เท่าเทียมกัน

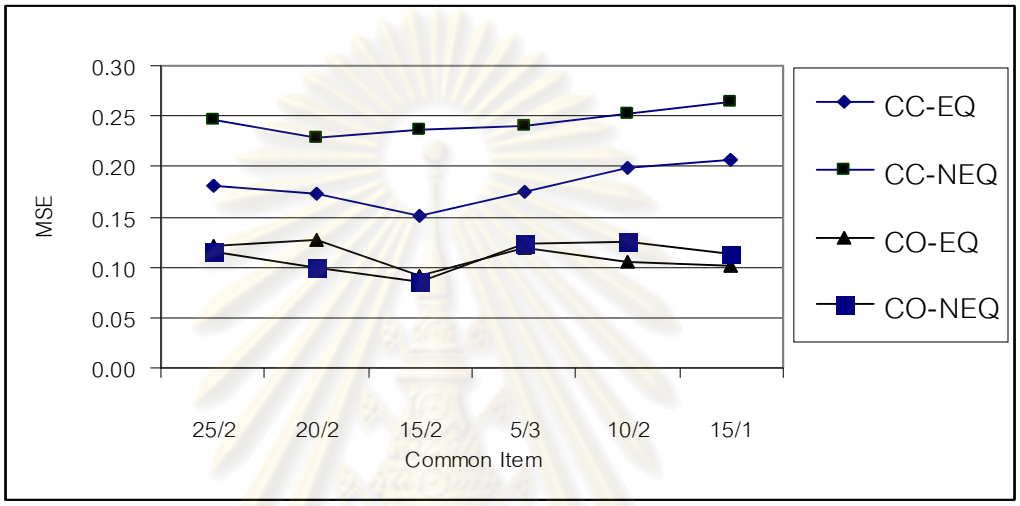
CO – EQ หมายถึง ผลการปรับเทียบคะแนนด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่ม ที่มีความสามารถเท่าเทียมกัน

CO – NEQ หมายถึง ผลการปรับเทียบคะแนนด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่ม ที่มีความสามารถไม่เท่าเทียมกัน

**ผลการเปรียบเทียบคุณภาพการปรับเทียบจำแนกตามสัดส่วนข้อสอบในแบบสอบจำนวน 4 เงื่อนไข** ผู้วิจัยเสนอกราฟแสดงผลการเปรียบเทียบ มีรายละเอียดดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

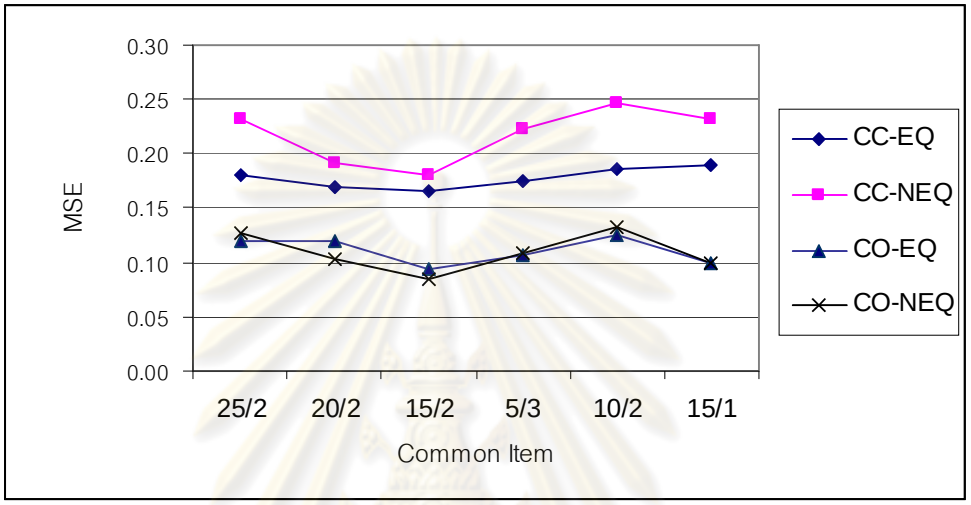
ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 10 ข้อ (DS 50 : PS 10) พิจารณาเปรียบเทียบตามวิธีการปรับเทียบคะแนนทั้ง 2 วิธี รูปแบบการเก็บรวบรวมข้อมูล 2 รูปแบบ ตามสัดส่วนของข้อสอบร่วมแต่ละสัดส่วน ปรากฏดังแผนภาพที่ 28




แผนภาพที่ 28 ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบในแบบสอบเท่ากับ DS 50 : PS 10

จากแผนภาพที่ 28 แสดงผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 10 ข้อ ด้วยวิธีโค้งคุณลักษณะ (CC) วิธีการปรับค่าพารามิเตอร์พร้อมกัน (CO) ที่มีรูปแบบการเก็บรวบรวมข้อมูลต่างกันระหว่างกลุ่มเท่าเทียม (EQ) และกลุ่มไม่เท่าเทียม (NEQ) และสัดส่วนข้อสอบร่วมต่างกัน 6 สัดส่วน พบว่า การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CC – NEQ) เกิดความคลาดเคลื่อนสูงที่สุด ในเงื่อนไขสัดส่วนของข้อสอบร่วมทุกเงื่อนไข และวิธีการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียม (CO – EQ) เกิดความคลาดเคลื่อนในการปรับเทียบต่ำ และมีค่าใกล้เคียงกับวิธีการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม (CO – NEQ)

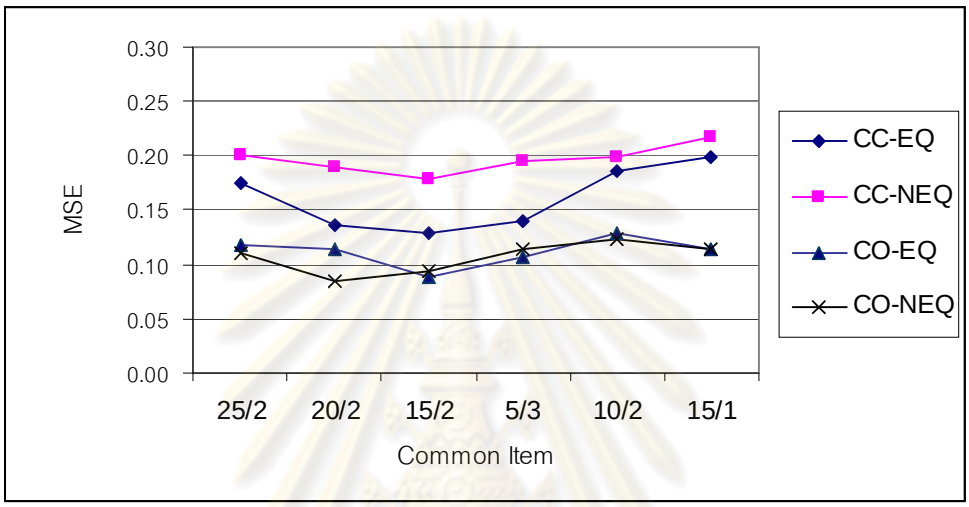
ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 60 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 8 ข้อ (DS 60 : PS 8) พิจารณาเปรียบเทียบตามวิธีการปรับเทียบคะแนนทั้ง 2 วิธี รูปแบบการเก็บรวบรวมข้อมูล 2 รูปแบบ ตามสัดส่วนของข้อสอบร่วมแต่ละสัดส่วน ปรากฏดังแผนภาพที่ 29




แผนภาพที่ 29 ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบในแบบสอบเท่ากับ DS 60 : PS 8

จากแผนภาพที่ 29 แสดงผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 60 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 8 ข้อ ด้วยวิธีโค้งคุณลักษณะ (CC) วิธีการปรับค่าพารามิเตอร์พร้อมกัน (CO) ที่มีรูปแบบการเก็บรวบรวมข้อมูลต่างกันระหว่างกลุ่มเท่าเทียม (EQ) และกลุ่มไม่เท่าเทียม (NEQ) และสัดส่วนข้อสอบร่วมต่างกัน 6 สัดส่วน พบว่า การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CC – NEQ) เกิดความคลาดเคลื่อนสูงที่สุด ในเงื่อนไขสัดส่วนของข้อสอบร่วมทุกเงื่อนไข และวิธีการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียม (CO – EQ) เกิดความคลาดเคลื่อนในการปรับเทียบต่ำ และมีค่าใกล้เคียงกับวิธีการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม (CO – NEQ)

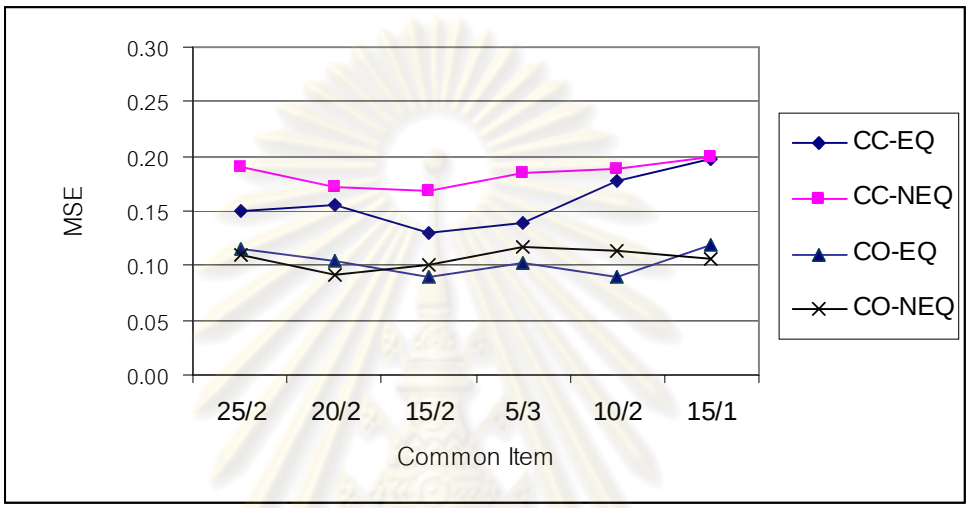
ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 70 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 6 ข้อ (DS 70 : PS 6) พิจารณาเปรียบเทียบตามวิธีการปรับเทียบคะแนนทั้ง 2 วิธี รูปแบบการเก็บรวบรวมข้อมูล 2 รูปแบบ ตามสัดส่วนของข้อสอบร่วมแต่ละสัดส่วน ปรากฏดังแผนภาพที่ 30




แผนภาพที่ 30 ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบในแบบสอบเท่ากับ DS 70 : PS 6

จากแผนภาพที่ 30 แสดงผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 70 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 6 ข้อ ด้วยวิธีโค้งคุณลักษณะ (CC) วิธีการปรับค่าพารามิเตอร์พร้อมกัน (CO) ที่มีรูปแบบการเก็บรวบรวมข้อมูลต่างกันระหว่างกลุ่มเท่าเทียม (EQ) และกลุ่มไม่เท่าเทียม (NEQ) และสัดส่วนข้อสอบร่วมต่างกัน 6 สัดส่วน พบว่า การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CC – NEQ) เกิดความคลาดเคลื่อนสูงที่สุด ในเงื่อนไขสัดส่วนของข้อสอบร่วมทุกเงื่อนไข และวิธีการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียม (CO-EQ) เกิดความคลาดเคลื่อนในการปรับเทียบต่ำ และมีค่าใกล้เคียงกับวิธีการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม (CO – NEQ)

ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 80 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 4 ข้อ (DS 80 : PS 4) พิจารณาเปรียบเทียบตามวิธีการปรับเทียบคะแนนทั้ง 2 วิธี รูปแบบการเก็บรวบรวมข้อมูล 2 รูปแบบ ตามสัดส่วนของข้อสอบร่วมแต่ละสัดส่วน ปรากฏดังแผนภาพที่ 31

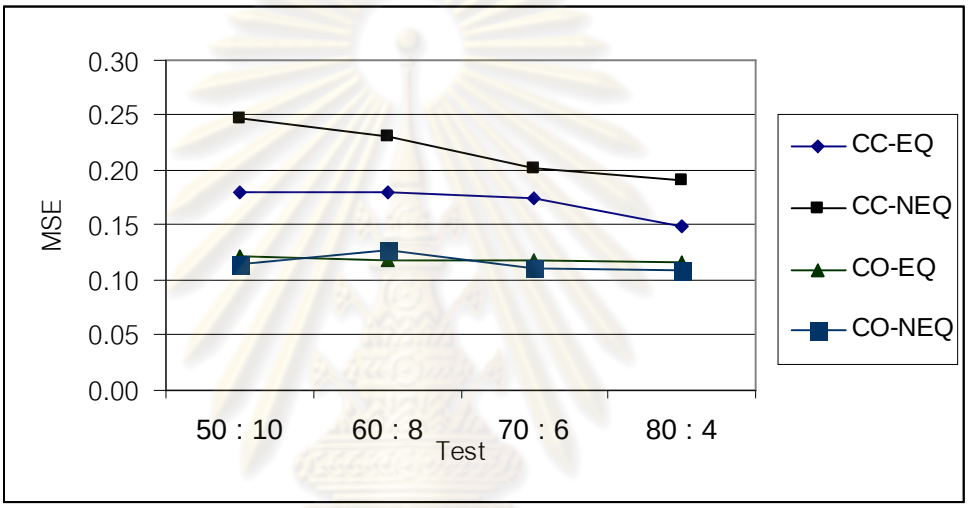



แผนภาพที่ 31 ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบในแบบสอบเท่ากับ DS 80 : PS 4

จากแผนภาพที่ 31 แสดงผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 80 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 4 ข้อ ด้วยวิธีโค้งคุณลักษณะ (CC) วิธีการปรับค่าพารามิเตอร์พร้อมกัน (CO) ที่มีรูปแบบการเก็บรวบรวมข้อมูลต่างกันระหว่างกลุ่มเท่าเทียม (EQ) และกลุ่มไม่เท่าเทียม (NEQ) และสัดส่วนข้อสอบร่วมต่างกัน 6 สัดส่วน พบว่า การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CC – NEQ) เกิดความคลาดเคลื่อนสูงที่สุด ในเงื่อนไขสัดส่วนของข้อสอบร่วมทุกเงื่อนไข และวิธีการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียม (CO-EQ) เกิดความคลาดเคลื่อนในการปรับเทียบต่ำ และมีค่าใกล้เคียงกับวิธีการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม (CO – NEQ)

**ผลการเปรียบเทียบคุณภาพการปรับเทียบจำแนกตามสัดส่วนข้อสอบร่วมในแบบสอบจำนวน 6 เงื่อนไข** ผู้วิจัยเสนอกราฟแสดงผลการเปรียบเทียบ มีรายละเอียดดังนี้

ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 25 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 2 ข้อ (DS 25 : PS 2) พิจารณาเปรียบเทียบตามวิธีการปรับเทียบคะแนนทั้ง 2 วิธี รูปแบบการเก็บรวบรวมข้อมูล 2 รูปแบบ ตามสัดส่วนของข้อสอบร่วมแต่ละสัดส่วน ปรากฏดังแผนภาพที่ 32

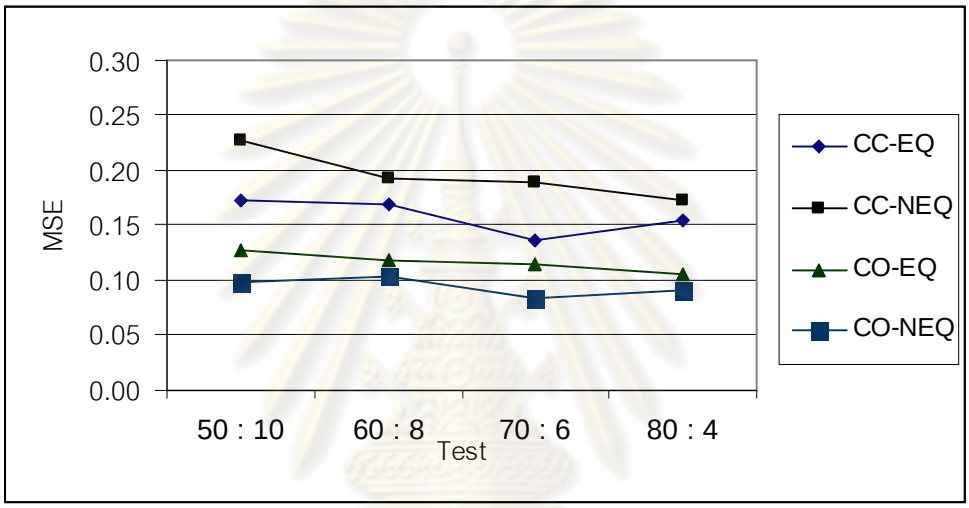



แผนภาพที่ 32 ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบร่วมในแบบสอบเท่ากับ DS 25 : PS 2

จากแผนภาพที่ 32 แสดงผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างที่มีสัดส่วนข้อสอบร่วมที่ตรวจให้คะแนนสองค่า 25 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 2 ข้อ ด้วยวิธีโค้งคุณลักษณะ (CC) วิธีการปรับค่าพารามิเตอร์พร้อมกัน (CO) ที่มีรูปแบบการเก็บรวบรวมข้อมูลต่างกันระหว่างกลุ่มเท่าเทียม (EQ) และกลุ่มไม่เท่าเทียม (NEQ) และสัดส่วนข้อสอบในแบบสอบที่แตกต่างกัน 4 สัดส่วน พบว่า การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CC – NEQ) เกิดความคลาดเคลื่อนสูงที่สุด ในเงื่อนไขสัดส่วนข้อสอบในแบบสอบทุกเงื่อนไข และวิธีการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียม (CO-EQ) เกิดความคลาดเคลื่อนในการปรับเทียบต่ำ และมีค่าใกล้เคียงกับวิธีการปรับเทียบ

ด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม (CO – NEQ)

ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 20 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 2 ข้อ (DS 20 : PS 2) พิจารณาเปรียบเทียบตามวิธีการปรับเทียบคะแนนทั้ง 2 วิธี รูปแบบการเก็บรวบรวมข้อมูล 2 รูปแบบ ตามสัดส่วนของข้อสอบร่วมแต่ละสัดส่วน ปรากฏดังแผนภาพที่ 33

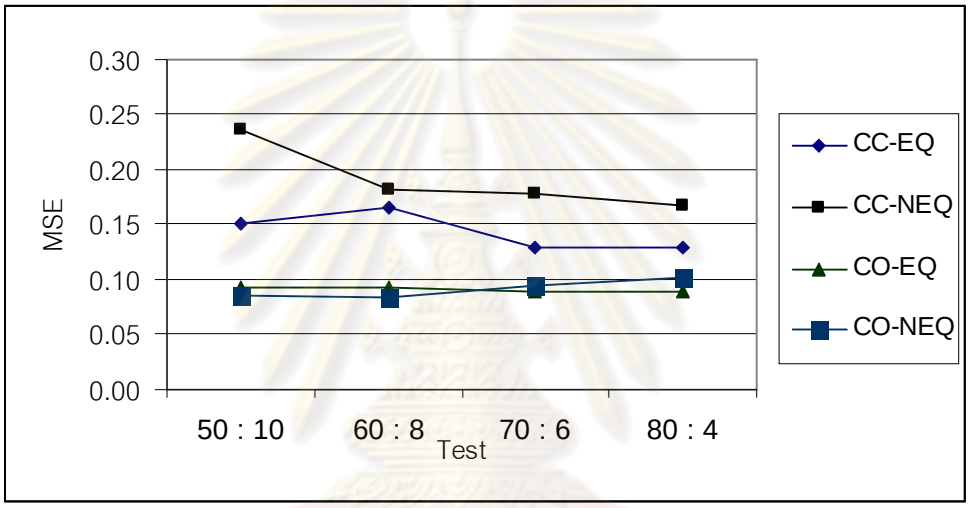



แผนภาพที่ 33 ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบร่วมในแบบสอบเท่ากับ DS 20 : PS 2

จากแผนภาพที่ 33 แสดงผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างที่มีสัดส่วนข้อสอบร่วมที่ตรวจให้คะแนนสองค่า 20 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 2 ข้อ ด้วยวิธีโค้งคุณลักษณะ (CC) วิธีการปรับค่าพารามิเตอร์พร้อมกัน (CO) ที่มีรูปแบบการเก็บรวบรวมข้อมูลต่างกันระหว่างกลุ่มเท่าเทียม (EQ) และกลุ่มไม่เท่าเทียม (NEQ) และสัดส่วนข้อสอบในแบบสอบที่แตกต่างกัน 4 สัดส่วน พบว่า การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CC – NEQ) เกิดความคลาดเคลื่อนสูงที่สุด ในเงื่อนไขสัดส่วนข้อสอบในแบบสอบทุกเงื่อนไข และวิธีการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียม (CO-EQ) เกิดความคลาดเคลื่อนในการปรับเทียบต่ำ และมีค่าใกล้เคียงกับวิธีการปรับเทียบ

ด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม (CO – NEQ)

ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 2 ข้อ (DS 15 : PS 2) พิจารณาเปรียบเทียบตามวิธีการปรับเทียบคะแนนทั้ง 2 วิธี รูปแบบการเก็บรวบรวมข้อมูล 2 รูปแบบ ตามสัดส่วนของข้อสอบร่วมแต่ละสัดส่วน ปรากฏดังแผนภาพที่ 34

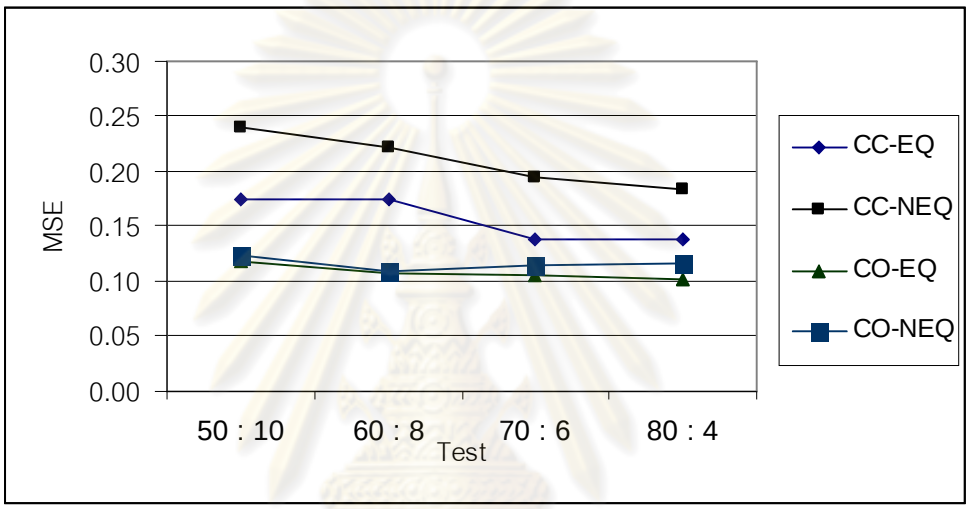



แผนภาพที่ 34 ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบร่วมในแบบสอบเท่ากับ DS 15 : PS 2

จากแผนภาพที่ 34 แสดงผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างที่มีสัดส่วนข้อสอบร่วมที่ตรวจให้คะแนนสองค่า 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 2 ข้อ ด้วยวิธีโค้งคุณลักษณะ (CC) วิธีการปรับค่าพารามิเตอร์พร้อมกัน (CO) ที่มีรูปแบบการเก็บรวบรวมข้อมูลต่างกันระหว่างกลุ่มเท่าเทียม (EQ) และกลุ่มไม่เท่าเทียม (NEQ) และสัดส่วนข้อสอบในแบบสอบที่แตกต่างกัน 4 สัดส่วน พบว่า การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CC – NEQ) เกิดความคลาดเคลื่อนสูงที่สุด ในเงื่อนไขสัดส่วนข้อสอบในแบบสอบทุกเงื่อนไข และวิธีการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียม (CO-EQ) เกิดความคลาดเคลื่อนในการปรับเทียบต่ำ และมีค่าใกล้เคียงกับวิธีการปรับเทียบ

ด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม (CO – NEQ)

ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 5 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 3 ข้อ (DS 5 : PS 3) พิจารณาเปรียบเทียบตามวิธีการปรับเทียบคะแนนทั้ง 2 วิธี รูปแบบการเก็บรวบรวมข้อมูล 2 รูปแบบ ตามสัดส่วนของข้อสอบร่วมแต่ละสัดส่วน ปรากฏดังแผนภาพที่ 35

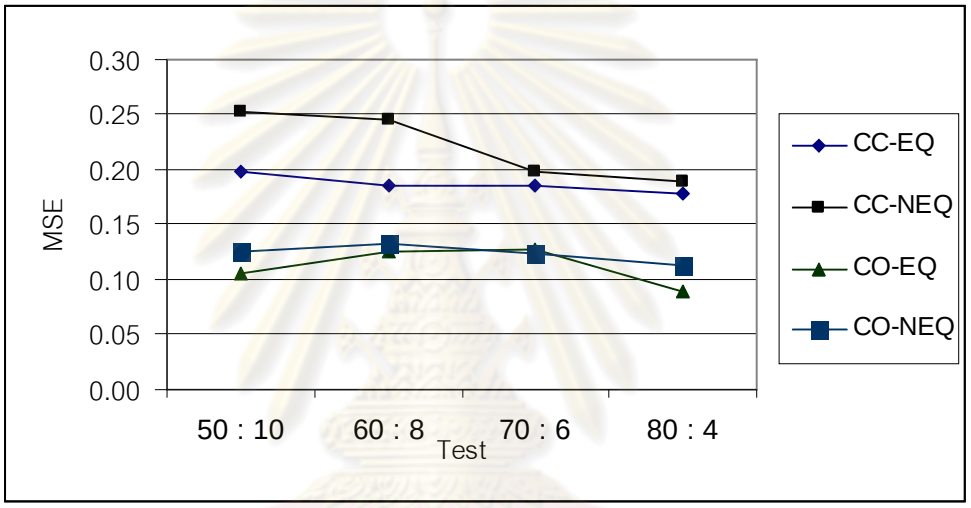



แผนภาพที่ 35 ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบร่วมในแบบสอบเท่ากับ DS 5 : PS 3

จากแผนภาพที่ 35 แสดงผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างที่มีสัดส่วนข้อสอบร่วมที่ตรวจให้คะแนนสองค่า 5 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 3 ข้อ ด้วยวิธีโค้งคุณลักษณะ (CC) วิธีการปรับค่าพารามิเตอร์พร้อมกัน (CO) ที่มีรูปแบบการเก็บรวบรวมข้อมูลต่างกันระหว่างกลุ่มเท่าเทียม (EQ) และกลุ่มไม่เท่าเทียม (NEQ) และสัดส่วนข้อสอบในแบบสอบที่แตกต่างกัน 4 สัดส่วน พบว่า การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CC – NEQ) เกิดความคลาดเคลื่อนสูงที่สุด ในเงื่อนไขสัดส่วนข้อสอบในแบบสอบทุกเงื่อนไข และวิธีการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียม (CO-EQ) เกิดความคลาดเคลื่อนในการปรับเทียบต่ำ และมีค่าใกล้เคียงกับวิธีการปรับเทียบ

ด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม (CO – NEQ)

ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 10 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 2 ข้อ (DS 10 : PS 2) พิจารณาเปรียบเทียบตามวิธีการปรับเทียบคะแนนทั้ง 2 วิธี รูปแบบการเก็บรวบรวมข้อมูล 2 รูปแบบ ตามสัดส่วนของข้อสอบร่วมแต่ละสัดส่วน ปรากฏดังแผนภาพที่ 36

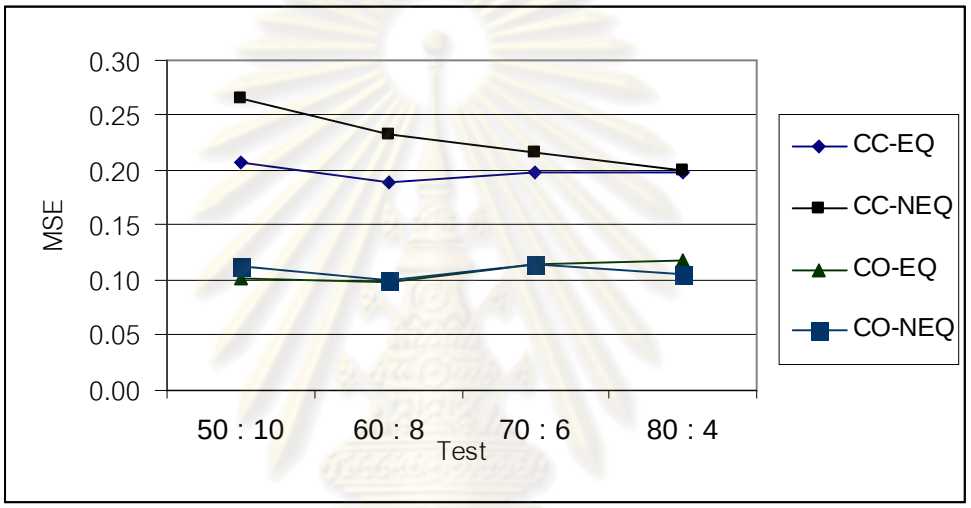



แผนภาพที่ 36 ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบร่วมในแบบสอบเท่ากับ DS 10 : PS 2

จากแผนภาพที่ 36 แสดงผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างที่มีสัดส่วนข้อสอบร่วมที่ตรวจให้คะแนนสองค่า 10 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 2 ข้อ ด้วยวิธีโค้งคุณลักษณะ (CC) วิธีการปรับค่าพารามิเตอร์พร้อมกัน (CO) ที่มีรูปแบบการเก็บรวบรวมข้อมูลต่างกันระหว่างกลุ่มเท่าเทียม (EQ) และกลุ่มไม่เท่าเทียม (NEQ) และสัดส่วนข้อสอบในแบบสอบที่แตกต่างกัน 4 สัดส่วน พบว่า การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CC – NEQ) เกิดความคลาดเคลื่อนสูงที่สุด ในเงื่อนไขสัดส่วนข้อสอบในแบบสอบทุกเงื่อนไข และวิธีการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียม (CO-EQ) เกิดความคลาดเคลื่อนในการปรับเทียบต่ำ และมีค่าใกล้เคียงกับวิธีการปรับเทียบ

ด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม (CO – NEQ)

ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีข้อสอบร่วมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 1 ข้อ (DS 15 : PS 1) พิจารณาเปรียบเทียบตามวิธีการปรับเทียบคะแนนทั้ง 2 วิธี รูปแบบการเก็บรวบรวมข้อมูล 2 รูปแบบ ตามสัดส่วนของข้อสอบร่วมแต่ละสัดส่วน ปรากฏดังแผนภาพที่ 37




แผนภาพที่ 37 ผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนทั้ง 2 วิธี เมื่อสัดส่วนข้อสอบร่วมในแบบสอบเท่ากับ DS 15 : PS 1

จากแผนภาพที่ 37 แสดงผลการเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างที่มีสัดส่วนข้อสอบร่วมที่ตรวจให้คะแนนสองค่า 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่า 1 ข้อ ด้วยวิธีโค้งคุณลักษณะ (CC) วิธีการปรับค่าพารามิเตอร์พร้อมกัน (CO) ที่มีรูปแบบการเก็บรวบรวมข้อมูลต่างกันระหว่างกลุ่มเท่าเทียม (EQ) และกลุ่มไม่เท่าเทียม (NEQ) และสัดส่วนข้อสอบในแบบสอบที่แตกต่างกัน 4 สัดส่วน พบว่า การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CC – NEQ) เกิดความคลาดเคลื่อนสูงที่สุด ในเงื่อนไขสัดส่วนข้อสอบในแบบสอบทุกเงื่อนไข และวิธีการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียม (CO-EQ) เกิดความคลาดเคลื่อนในการปรับเทียบต่ำ และมีค่าใกล้เคียงกับวิธีการปรับเทียบ

ด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม (CO – NEQ)

**2.2 ผลการวิเคราะห์เปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม**

การเปรียบเทียบคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมระหว่างวิธีโค้งคุณลักษณะ (CC) และวิธีการปรับค่าพารามิเตอร์พร้อมกัน (CO) ผู้วิจัยเสนอการเปรียบเทียบวิธีการปรับเทียบสองวิธีที่มีรูปแบบการเก็บรวบรวมข้อมูลที่ต่างกัน สัดส่วนข้อสอบทั้งสองชนิดในแบบสอบ และสัดส่วนข้อสอบร่วม ผลการเปรียบเทียบเสนอดังนี้

**2.2.1 ผลการวิเคราะห์เปรียบเทียบค่าเฉลียของดัชนี MSE ของเงื่อนไขสัดส่วนข้อสอบในแบบสอบที่แตกต่างกัน 4 เงื่อนไข**

การนำเสนอในส่วนนี้ผู้วิจัยได้วิเคราะห์เปรียบเทียบค่าเฉลี่ยดัชนี MSE ของแบบสอบแต่ละสัดส่วน (4 สัดส่วนตามเงื่อนไขที่กำหนด) โดยผู้วิจัยคำนวณค่าเฉลี่ยดัชนี MSE ของแบบสอบแต่ละเงื่อนไข จากดัชนี MSE ของสัดส่วนข้อสอบร่วมทั้ง 6 สัดส่วน สำหรับแบบสอบแต่ละฉบับ และทำการวิเคราะห์เปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยดัชนี MSE ของเงื่อนไขสัดส่วนข้อสอบในแบบสอบทั้ง 4 เงื่อนไข และนำเสนอจำแนกตามวิธีการปรับเทียบคะแนนแต่ละวิธีที่มีเงื่อนไขรูปแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน ผลปรากฏดังนี้

ตารางที่ 25 การวิเคราะห์ความแตกต่างสัดส่วนข้อสอบในแบบสอบจากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับกลุ่มเท่าเทียม

| การปรับเทียบ CC-EQ | | ความแตกต่างระหว่างกลุ่ม | | | |
|---|---|---|---|---|---|
| สัดส่วนข้อสอบในแบบสอบ | ค่าเฉลี่ย | 50/10 | 60/8 | 70/6 | 80/4 |
| DS 50 : PS 10 | .181 | - | .004 | .021 | .023 |
| DS 60 : PS 8 | .177 | | - | .017 | .019 |
| DS 70 : PS 6 | .160 | | | - | .002 |
| DS 80 : PS 4 | .158 | | | | - |
| F = 1.572 | Sig = .227 | | | | |

จากตารางที่ 25 พบว่า ผลการเปรียบเทียบค่าเฉลี่ยดัชนี MSE ของแบบสอบแต่ละสัดส่วนที่ปรับเทียบด้วยวิธีโค้งคุณลักษณะ ด้วยการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน ค่าเฉลี่ยดัชนี MSE ของแบบสอบที่มีสัดส่วน DS 50 : PS 10 มีค่าเฉลี่ยดัชนี MSE สูงที่สุด รองลงมาได้แก่ แบบสอบ DS 60 : PS 8, DS 70 : PS 6, และ DS 80 : PS 4 ตามลำดับ เมื่อทำการเปรียบเทียบค่าเฉลี่ยรายคู่ของแบบสอบแต่ละสัดส่วน พบว่า แบบสอบทุกสัดส่วนมีค่าดัชนี MSE ไม่แตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ .05

ตารางที่ 26 การวิเคราะห์ความแตกต่างสัดส่วนข้อสอบในแบบสอบจากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับกลุ่มไม่เท่าเทียม

| การปรับเทียบ CC-NEQ | | ความแตกต่างระหว่างกลุ่ม | | | |
|---|---|---|---|---|---|
| สัดส่วนข้อสอบในแบบสอบ | ค่าเฉลี่ย | 50/10 | 60/8 | 70/6 | 80/4 |
| DS 50 : PS 10 | .245 | - | .028 | .048** | .061** |
| DS 60 : PS 8 | .217 | | - | .020 | .033* |
| DS 70 : PS 6 | .197 | | | - | .013 |
| DS 80 : PS 4 | .184 | | | | - |
| F = 15.148 | Sig = .000 | | | | |

จากตารางที่ 26 พบว่า ผลการเปรียบเทียบค่าเฉลี่ยดัชนี MSE ของแบบสอบแต่ละสัดส่วนที่ปรับเทียบด้วยวิธีโค้งคุณลักษณะ ด้วยการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน ค่าเฉลี่ยดัชนี MSE ของแบบสอบที่มีสัดส่วน DS 50 : PS 10 มีค่าเฉลี่ยดัชนี MSE สูงที่สุด รองลงมาได้แก่ แบบสอบ DS 60 : PS 8, DS 70 : PS 6, และ DS 80 : PS 4 ตามลำดับ เมื่อทำการเปรียบเทียบค่าเฉลี่ยรายคู่ของแบบสอบแต่ละสัดส่วน พบว่า มีค่าดัชนี MSE แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ .05 ระหว่างแบบสอบสัดส่วน DS 50 : PS 10 กับ DS 70 : PS 6, DS 50 : PS 10 กับ DS 80 : PS 4, และ DS 60 : PS 8 กับ DS 80 : PS 4

ตารางที่ 27 การวิเคราะห์ความแตกต่างสัดส่วนข้อสอบในแบบสอบจากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีการเก็บรวบรวมข้อมูลกับกลุ่มเท่าเทียม

| การปรับเทียบ CO-EQ | | ความแตกต่างระหว่างกลุ่ม | | | |
|---|---|---|---|---|---|
| สัดส่วนข้อสอบในแบบสอบ | ค่าเฉลี่ย | 50/10 | 60/8 | 70/6 | 80/4 |
| DS 50 : PS 10 | .111 | - | .001 | -.001 | .008 |
| DS 60 : PS 8 | .110 | | - | -.002 | .007 |
| DS 70 : PS 6 | .112 | | | - | .009 |
| DS 80 : PS 4 | .103 | | | | - |
| F = .554 | Sig = .651 | | | | |

จากตารางที่ 27 พบว่า ผลการเปรียบเทียบค่าเฉลี่ยดัชนี MSE ของแบบสอบแต่ละสัดส่วนที่ปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน ค่าเฉลี่ยดัชนี MSE ของแบบสอบที่มีสัดส่วน DS 50 : PS 10 มีค่าเฉลี่ยดัชนี MSE สูงที่สุด รองลงมาได้แก่ แบบสอบ DS 70 : PS 6, DS 60 : PS 8, และ DS 80 : PS 4 ตามลำดับ เมื่อทำการเปรียบเทียบค่าเฉลี่ยรายคู่ของแบบสอบแต่ละสัดส่วน พบว่าแบบสอบทุกสัดส่วนมีค่าดัชนี MSE ไม่แตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ .05

ตารางที่ 28 การวิเคราะห์ความแตกต่างสัดส่วนข้อสอบในแบบสอบจากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีการเก็บรวบรวมข้อมูลกับกลุ่มไม่เท่าเทียม

| การปรับเทียบ CO-NEQ | | ความแตกต่างระหว่างกลุ่ม | | | |
|---|---|---|---|---|---|
| สัดส่วนข้อสอบในแบบสอบ | ค่าเฉลี่ย | 50/10 | 60/8 | 70/6 | 80/4 |
| DS 50 : PS 10 | .110 | - | .001 | .003 | .004 |
| DS 60 : PS 8 | .109 | | - | .002 | .003 |
| DS 70 : PS 6 | .107 | | | - | .001 |
| DS 80 : PS 4 | .106 | | | | - |
| F = .102 | Sig = .958 | | | | |

จากตารางที่ 28 พบว่า ผลการเปรียบเทียบค่าเฉลี่ยดัชนี MSE ของแบบสอบแต่ละสัดส่วนที่ปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน ค่าเฉลี่ยดัชนี MSE ของแบบสอบที่มีสัดส่วน DS 50 : PS 10 มีค่าเฉลี่ยดัชนี MSE สูงที่สุด รองลงมาได้แก่ แบบสอบ DS 60 : PS 8, DS 70 : PS 6, และ DS 80 : PS 4 ตามลำดับ เมื่อทำการเปรียบเทียบค่าเฉลี่ยรายคู่ของแบบสอบแต่ละสัดส่วน พบว่าแบบสอบทุกสัดส่วนมีค่าดัชนี MSE ไม่แตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ .05

### 2.2.2 ผลการวิเคราะห์เปรียบเทียบค่าเฉลียของดัชนี MSE ของเงือนไขสัดส่วนข้อสอบร่วมในสัดส่วนข้อสอบร่วมในแบบสอบแต่ละฉบับ

การนำเสนอในส่วนนี้ผู้วิจัยได้วิเคราะห์เปรียบเทียบค่าเฉลี่ยดัชนี MSE ของสัดส่วนข้อสอบร่วมแต่ละสัดส่วน (6 สัดส่วนตามเงื่อนไขที่กำหนด) โดยผู้วิจัยคำนวณค่าเฉลี่ยดัชนี MSE ของสัดส่วนข้อสอบร่วมแต่ละเงื่อนไข จากดัชนี MSE ของการปรับเทียบแต่ละวิธี ที่ทำการเก็บรวบรวมข้อมูลแต่ละรูปแบบ และทำการวิเคราะห์เปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยดัชนี MSE ของเงื่อนไขสัดส่วนข้อสอบร่วมทั้ง 6 เงื่อนไข และนำเสนอจำแนกตามสัดส่วนข้อสอบทั้ง 2 ชนิดในแบบสอบแต่ละเงื่อนไข (4 เงื่อนไข) ผลปรากฏดังนี้

ตารางที่ 29 การวิเคราะห์ความแตกต่างระหว่างข้อสอบร่วมแต่ละสัดส่วน จากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยแบบสอบที่มีสัดส่วน DS 50 : PS 10

| แบบสอบ 50/10 | | ความแตกต่างระหว่างกลุ่ม | | | | | |
|---|---|---|---|---|---|---|---|
| Common item | ค่าเฉลี่ย | 15/2 | 20/2 | 25/2 | 15/1 | 10/2 | 5/3 |
| 15/2 | 0.129 | - | -0.012 | -0.027 | -0.032 | -0.032 | -0.019 |
| 20/2 | 0.141 | | - | -0.015 | -0.020 | -0.020 | -0.007 |
| 25/2 | 0.156 | | | - | -0.005 | -0.005 | 0.008 |
| 15/1 | 0.161 | | | | - | 0.000 | 0.013 |
| 10/2 | 0.161 | | | | | - | 0.013 |
| 5/3 | 0.148 | | | | | | - |
| F = .121 | Sig = .986 | | | | | | |

จากตารางที่ 29 พบว่า ผลการเปรียบเทียบค่าเฉลี่ยดัชนี MSE ของข้อสอบร่วมแต่ละสัดส่วน ที่ปรับเทียบด้วยแบบสอบสัดส่วน DS 50 : PS10 ค่าเฉลี่ยดัชนี MSE ของสัดส่วนข้อสอบร่วม 15/1 และ 10/2 มีค่าสูงที่สุด เมื่อทำการเปรียบเทียบค่าเฉลี่ยรายคู่ของสัดส่วนข้อสอบร่วมแต่ละสัดส่วน พบว่า ข้อสอบร่วมทุกสัดส่วนมีค่าดัชนี MSE ไม่แตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ .05

ตารางที่ 30 การวิเคราะห์ความแตกต่างระหว่างข้อสอบร่วมแต่ละสัดส่วน จากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยแบบสอบที่มีสัดส่วน DS 60 : PS 8

| แบบสอบ 60/8 | | ความแตกต่างระหว่างกลุ่ม | | | | | |
|---|---|---|---|---|---|---|---|
| Common item | ค่าเฉลี่ย | 15/2 | 20/2 | 25/2 | 15/1 | 10/2 | 5/3 |
| 15/2 | 0.131 | - | -0.015 | -0.034 | -0.025 | -0.041 | -0.023 |
| 20/2 | 0.146 | | - | -0.019 | -0.010 | -0.026 | -0.008 |
| 25/2 | 0.164 | | | - | 0.009 | -0.008 | 0.011 |
| 15/1 | 0.155 | | | | - | -0.017 | 0.002 |
| 10/2 | 0.172 | | | | | - | 0.019 |
| 5/3 | 0.153 | | | | | | - |
| F = .284 | Sig = .916 | | | | | | |

จากตารางที่ 30 พบว่า ผลการเปรียบเทียบค่าเฉลี่ยดัชนี MSE ของข้อสอบร่วมแต่ละสัดส่วน ที่ปรับเทียบด้วยแบบสอบสัดส่วน DS 60 : PS 8 ค่าเฉลี่ยดัชนี MSE ของสัดส่วนข้อสอบร่วม 10/2 มีค่าสูงที่สุด เมื่อทำการเปรียบเทียบค่าเฉลี่ยรายคู่ของสัดส่วนข้อสอบร่วมแต่ละสัดส่วน พบว่า ข้อสอบร่วมทุกสัดส่วนมีค่าดัชนี MSE ไม่แตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ .05

ตารางที่ 31 การวิเคราะห์ความแตกต่างระหว่างข้อสอบร่วมแต่ละสัดส่วน จากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยแบบสอบที่มีสัดส่วน DS 70 : PS 6

| แบบสอบ 70/6 | | ความแตกต่างระหว่างกลุ่ม | | | | | |
|---|---|---|---|---|---|---|---|
| Common item | ค่าเฉลี่ย | 15/2 | 20/2 | 25/2 | 15/1 | 10/2 | 5/3 |
| 15/2 | 0.123 | - | -0.008 | -0.028 | -0.039 | -0.036 | -0.016 |
| 20/2 | 0.131 | | - | -0.020 | -0.031 | -0.028 | -0.008 |
| 25/2 | 0.151 | | | - | -0.010 | -0.008 | 0.012 |
| 15/1 | 0.161 | | | | - | 0.003 | 0.023 |
| 10/2 | 0.159 | | | | | - | 0.020 |
| 5/3 | 0.139 | | | | | | - |
| F = .499 | Sig = .773 | | | | | | |

จากตารางที่ 31 พบว่า ผลการเปรียบเทียบค่าเฉลี่ยดัชนี MSE ของข้อสอบร่วมแต่ละสัดส่วน ที่ปรับเทียบด้วยแบบสอบสัดส่วน DS 70 : PS 6 ค่าเฉลี่ยดัชนี MSE ของสัดส่วนข้อสอบร่วม 15/1มีค่าสูงที่สุด เมื่อทำการเปรียบเทียบค่าเฉลี่ยรายคู่ของสัดส่วนข้อสอบร่วมแต่ละสัดส่วน พบว่า ข้อสอบร่วมทุกสัดส่วนมีค่าดัชนี MSE ไม่แตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ .05

ตารางที่ 32 การวิเคราะห์ความแตกต่างระหว่างข้อสอบร่วมแต่ละสัดส่วน จากค่าเฉลี่ยดัชนี MSE เมื่อปรับเทียบด้วยแบบสอบที่มีสัดส่วน DS 80 : PS 4

| แบบสอบ 80/4 | | ความแตกต่างระหว่างกลุ่ม | | | | | |
|---|---|---|---|---|---|---|---|
| Common item | ค่าเฉลี่ย | 15/2 | 20/2 | 25/2 | 15/1 | 10/2 | 5/3 |
| 15/2 | 0.122 | - | -0.009 | -0.020 | -0.034 | -0.020 | -0.014 |
| 20/2 | 0.131 | | - | -0.011 | -0.025 | -0.012 | -0.005 |
| 25/2 | 0.142 | | | - | -0.014 | -0.001 | 0.006 |
| 15/1 | 0.156 | | | | - | 0.013 | 0.020 |
| 10/2 | 0.142 | | | | | - | 0.007 |
| 5/3 | 0.136 | | | | | | - |
| F = .302 | Sig = .905 | | | | | | |

จากตารางที่ 32 พบว่า ผลการเปรียบเทียบค่าเฉลี่ยดัชนี MSE ของข้อสอบร่วมแต่ละสัดส่วน ที่ปรับเทียบด้วยแบบสอบสัดส่วน DS 80 : PS 4 ค่าเฉลี่ยดัชนี MSE ของสัดส่วนข้อสอบร่วม 15/1 มีค่าสูงที่สุด เมื่อทำการเปรียบเทียบค่าเฉลี่ยรายคู่ของสัดส่วนข้อสอบร่วมแต่ละสัดส่วน พบว่า ข้อสอบร่วมทุกสัดส่วนมีค่าดัชนี MSE ไม่แตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ .05

**2.2.3 ผลการวิเคราะห์เปรียบเทียบค่าเฉลียของดัชนี MSE ของปรับเทียบคะแนนแต่ละวิธี ทีมีเงือนไขรูปแบบการเก็บรวบรวมข้อมูลทีแตกต่างกัน เมือพิจารณาจำแนกตามสัดส่วนของข้อสอบในแบบสอบ**

การนำเสนอในส่วนนี้ผู้วิจัยได้วิเคราะห์เปรียบเทียบค่าเฉลี่ยดัชนี MSE ของวิธีการปรับเทียบแต่ละวิธีที่มีรูปแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน (CC – EQ, CC – NEQ, CO – EQ, CO - NEQ) จำแนกตามสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบ (4 เงื่อนไข) โดยผู้วิจัยคำนวณค่าเฉลี่ยดัชนี MSE ของวิธีการปรับเทียบแต่ละวิธีที่มีรูปแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน จากดัชนี MSE ของสัดส่วนข้อสอบร่วมที่อยู่ภายในแบบสอบแต่ละฉบับ และนำเสนอจำแนกตามสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบแต่ละเงื่อนไข (4 เงื่อนไข) ผลปรากฏดังนี้

ตารางที่ 33 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนี MSE ของการปรับเทียบที่มีรูปแบบการเก็บรวบรวมข้อมูลต่างกัน สำหรับแบบสอบสัดส่วน DS 50 : PS 10

| การปรับเทียบ | ค่าเฉลี่ย | ความแตกต่างระหว่างกลุ่ม | | | |
|---|---|---|---|---|---|
| | | CC-EQ | CC-NEQ | CO-EQ | CO-NEQ |
| CC-EQ | .181 | - | -.064** | .069** | .070** |
| CC-NEQ | .245 | | - | .134** | .135** |
| CO-EQ | .111 | | | - | .001 |
| CO-NEQ | .110 | | | | - |
| F = 163.780 | Sig = .000 | | | | |

จากตารางที่ 33 พบว่า ค่าเฉลี่ยดัชนี MSE วิธีการปรับเทียบสำหรับแบบสอบสัดส่วน DS 50 : PS 10 ด้วยวิธีการปรับเทียบแบบโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน (CC-NEQ) มีค่าสูงที่สุด เมื่อทำการเปรียบเทียบค่าเฉลี่ยรายคู่ พบว่า ค่าเฉลี่ยดัชนี MSE ของการปรับเทียบด้วยวิธีโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่ม

ผู้สอบที่มีความสามารถไม่เท่าเทียมกัน (CC-NEQ) มีค่า**สูงกว่า**การปรับเทียบแบบโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน (CC-EQ) การปรับเทียบด้วยวิธีปรับค่าพารามิเตอร์พร้อมกัน กับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CO-EQ) และวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CO-NEQ) อย่างมีนัยสำคัญทางสถิติที่ .01 และ การปรับเทียบด้วยวิธีโค้งคุณลักษณะกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CC-EQ) มีค่าเฉลี่ยดัชนี MSE **สูงกว่า** การปรับเทียบวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CO-EQ) และวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CO-NEQ) อย่างมีนัยสำคัญทางสถิติที่ .01 แต่พบว่า ค่าเฉลี่ยดัชนี MSE ของการปรับเทียบวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CO-EQ) และวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CO-NEQ) มีค่าไม่แตกต่างกัน

ตารางที่ 34 การวิเคราะห์ความแตกต่างของค่า MSE เฉลี่ย ของการปรับเทียบ 4 วิธีที่แตกต่างกัน เมื่อจำแนกตามแบบสอบ DS 60 : PS 8

| การปรับเทียบ | ค่าเฉลี่ย | ความแตกต่างระหว่างกลุ่ม | | | |
|---|---|---|---|---|---|
| | | CC-EQ | CC-NEQ | CO-EQ | CO-NEQ |
| CC-EQ | .177 | - | -.040** | .067** | .068** |
| CC-NEQ | .217 | | - | .107** | .108** |
| CO-EQ | .110 | | | - | .001 |
| CO-NEQ | .109 | | | | - |
| F = 132.296 | Sig = .000 | | | | |

จากตารางที่ 34 พบว่า ค่าเฉลี่ยดัชนี MSE วิธีการปรับเทียบสำหรับแบบสอบสัดส่วน DS 60 : PS 8 ด้วยวิธีการปรับเทียบแบบโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน (CC-NEQ) มีค่าสูงที่สุด เมื่อทำการเปรียบเทียบค่าเฉลี่ยรายคู่ พบว่า ค่าเฉลี่ยดัชนี MSE ของการปรับเทียบด้วยวิธีโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน (CC-NEQ) มีค่า**สูงกว่า**การปรับเทียบแบบโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน (CC-EQ) การ

ปรับเทียบด้วยวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CO-EQ) และวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CO-NEQ) อย่างมีนัยสำคัญทางสถิติที่ .01 และ การปรับเทียบด้วยวิธีโค้งคุณลักษณะกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CC-EQ) มีค่าเฉลี่ยดัชนี MSE **สูงกว่า** การปรับเทียบวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CO-EQ) และวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CO-NEQ) อย่างมีนัยสำคัญทางสถิติที่ .01 แต่พบว่า ค่าเฉลี่ยดัชนี MSE ของการปรับเทียบวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CO-EQ) และวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CO-NEQ) มีค่าไม่แตกต่างกัน

ตารางที่ 35 การวิเคราะห์ความแตกต่างของค่า MSE เฉลี่ย ของการปรับเทียบ 4 วิธีที่แตกต่างกัน เมื่อจำแนกตามแบบสอบ DS 70 : PS 6

| การปรับเทียบ | ค่าเฉลี่ย | ความแตกต่างระหว่างกลุ่ม | | | |
|---|---|---|---|---|---|
| | | CC-EQ | CC-NEQ | CO-EQ | CO-NEQ |
| CC-EQ | .160 | - | -.036** | .049** | .054** |
| CC-NEQ | .197 | | - | .085** | .090** |
| CO-EQ | .112 | | | - | .005 |
| CO-NEQ | .107 | | | | - |
| F = 73.194 | Sig = .000 | | | | |

จากตารางที่ 35 พบว่า ค่าเฉลี่ยดัชนี MSE วิธีการปรับเทียบสำหรับแบบสอบสัดส่วน DS 70 : PS 6 ด้วยวิธีการปรับเทียบแบบโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน (CC-NEQ) มีค่าสูงที่สุด เมื่อทำการเปรียบเทียบค่าเฉลี่ยรายคู่ พบว่า ค่าเฉลี่ยดัชนี MSE ของการปรับเทียบด้วยวิธีโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน (CC-NEQ) มีค่า**สูงกว่า**การปรับเทียบแบบโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน (CC-EQ) การปรับเทียบด้วยวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CO-EQ) และวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่า

เทียมกัน (CO-NEQ) อย่างมีนัยสำคัญทางสถิติที่ .01 และ การปรับเทียบด้วยวิธีโค้งคุณลักษณะกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CC-EQ) มีค่าเฉลี่ยดัชนี MSE **สูงกว่า** การปรับเทียบวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CO-EQ) และวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CO-NEQ) อย่างมีนัยสำคัญทางสถิติที่ .01 แต่พบว่า ค่าเฉลี่ยดัชนี MSE ของการปรับเทียบวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CO-EQ) และวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CO-NEQ) มีค่าไม่แตกต่างกัน

ตารางที่ 36 การวิเคราะห์ความแตกต่างของค่า MSE เฉลี่ย ของการปรับเทียบ 4 วิธีที่แตกต่างกัน เมื่อจำแนกตามแบบสอบ DS 80 : PS 4

| การปรับเทียบ | ค่าเฉลี่ย | ความแตกต่างระหว่างกลุ่ม | | | |
|---|---|---|---|---|---|
| | | CC-EQ | CC-NEQ | CO-EQ | CO-NEQ |
| CC-EQ | .158 | - | -.026* | .055** | .052** |
| CC-NEQ | .184 | | - | .081** | .078** |
| CO-EQ | .103 | | | - | -.003 |
| CO-NEQ | .106 | | | | - |
| F = 54.843 | Sig = .000 | | | | |

จากตารางที่ 36 พบว่า ค่าเฉลี่ยดัชนี MSE วิธีการปรับเทียบสำหรับแบบสอบสัดส่วน DS 60 : PS 4 ด้วยวิธีการปรับเทียบแบบโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน (CC-NEQ) มีค่าสูงที่สุด เมื่อทำการเปรียบเทียบค่าเฉลี่ยรายคู่ พบว่า ค่าเฉลี่ยดัชนี MSE ของการปรับเทียบด้วยวิธีโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน (CC-NEQ) มีค่า**สูงกว่า**การปรับเทียบแบบโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน (CC-EQ) อย่างมีนัยสำคัญทางสถิติที่ .01 และพบว่ามีค่า**สูงกว่า**การปรับเทียบด้วยวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CO-EQ) และวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CO-NEQ) อย่างมีนัยสำคัญทางสถิติที่ .05 และ การปรับเทียบด้วยวิธีโค้งคุณลักษณะกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่า

เทียมกัน (CC-EQ) มีค่าเฉลี่ยดัชนี MSE **สูงกว่า** การปรับเทียบวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CO-EQ) และวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CO-NEQ) อย่างมีนัยสำคัญทางสถิติที่ .01 แต่พบว่า ค่าเฉลี่ยดัชนี MSE ของการปรับเทียบวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CO-EQ) และวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CO-NEQ) มีค่าไม่แตกต่างกัน

ตารางที่ 37 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนี MSE ของการปรับเทียบแต่ละวิธีที่มีรูปแบบการเก็บรวบรวมข้อมูลแตกต่างกัน (ภาพรวม)

| แบบสอบทุกฉบับ | | การปรับเทียบ | | | |
|---|---|---|---|---|---|
| การปรับเทียบ | ค่าเฉลี่ย | CC-EQ | CC-NEQ | CO-EQ | CO-NEQ |
| CC-EQ | .169 | - | -.042** | .060** | .061** |
| CC-NEQ | .211 | | - | .102** | .103** |
| CO-EQ | .109 | | | - | .001 |
| CO-NEQ | .108 | | | | - |
| F = 142.141 | Sig = .000 | | | | |

จากตารางที่ 37 พบว่า ค่าเฉลี่ยดัชนี MSE วิธีการปรับเทียบสำหรับแบบสอบในภาพรวม พบว่า วิธีการปรับเทียบแบบโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน (CC-NEQ) มีค่าสูงที่สุด เมื่อทำการเปรียบเทียบค่าเฉลี่ยรายคู่ พบว่า ค่าเฉลี่ยดัชนี MSE ของการปรับเทียบด้วยวิธีโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน (CC-NEQ) มีค่า**สูงกว่า**การปรับเทียบแบบโค้งคุณลักษณะที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน (CC-EQ) การปรับเทียบด้วยวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CO-EQ) และวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CO-NEQ) อย่างมีนัยสำคัญทางสถิติที่ .01 และ การปรับเทียบด้วยวิธีโค้งคุณลักษณะกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CC-EQ) มีค่าเฉลี่ยดัชนี MSE **สูงกว่า** การปรับเทียบวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CO-EQ) และวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CO-NEQ)

อย่างมีนัยสำคัญทางสถิติที่ .01 แต่พบว่า ค่าเฉลี่ยดัชนี MSE ของการปรับเทียบวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน (CO-EQ) และวิธีปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน (CO-NEQ) มีค่าไม่แตกต่างกัน

## ตอนที 2 การสรุปผลการวิจัยตามวัตถุประสงค์

การศึกษาคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมโดยการประยุกต์ใช้วิธีการปรับเทียบตามแนวทฤษฎีการตอบสนองข้อสอบ 2 วิธีคือ วิธีโค้งคุณลักษณะ และการปรับค่าพารามิเตอร์พร้อมกัน ภายใต้เงื่อนไขของการเก็บรวบรวมข้อมูลที่แตกต่างกัน 2 วิธี สัดส่วนของข้อสอบทั้ง 2 ชนิดในแบบสอบ 4 สัดส่วน และสัดส่วนของข้อสอบร่วม 6 สัดส่วน ผู้วิจัยได้สรุปผลการวิจัยตามวัตถุประสงค์การวิจัย มีรายละเอียดดังนี้

2.1 **สรุปผลการวิจัยตามวัตถุประสงค์ข้อที** 1 เพื่อวิเคราะห์คุณภาพการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ (Characteristic Curve) และวิธีการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent Calibration) เมื่อแบบแผนการเก็บรวบรวมข้อมูล และสัดส่วนของข้อสอบทั้ง 2 ชนิดในแบบสอบรูปแบบผสม และสัดส่วนของข้อสอบทั้ง 2 ชนิดที่เป็นข้อสอบร่วมของแบบสอบรูปแบบผสม ต่างกัน

ผลการวิจัยสรุปได้ดังนี้

**2.1.1 ผลการปรับเทียบคะแนนแต่ละวิธี ด้วยการเก็บรวบรวมข้อมูลด้วยรูปแบบทีแตกต่างกัน เมื่อพิจารณาตามสัดส่วนข้อสอบร่วมที่ต่างกัน**

2.1.1.1 คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบทีตรวจให้คะแนนสองค่า และข้อสอบทีตรวจให้คะแนนหลายค่าแตกต่างกัน** 4 **สัดส่วน ทีปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ** และเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่เท่าเทียมกัน** ซึ่งแบบสอบแต่ละสัดส่วน มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบค่อนข้างต่ำ (< 0.300) ซึ่งในภาพรวมเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุด ส่วนใหญ่คือสัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 1 ข้อ

2.1.1.2 คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบทีตรวจให้คะแนนสองค่า และข้อสอบทีตรวจให้คะแนนหลายค่าแตกต่างกัน** 4 **สัดส่วน ทีปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ** และเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีไม่เท่าเทียมกัน** ซึ่งแบบสอบแต่ละสัดส่วน มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบค่อนข้างต่ำ (< 0.300) ซึ่งในภาพรวมเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุด ส่วนใหญ่คือสัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 1 ข้อ

2.1.1.3 คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบทีตรวจให้คะแนนสองค่า และข้อสอบทีตรวจให้คะแนนหลายค่าแตกต่างกัน** 4 **สัดส่วน ทีปรับเทียบคะแนนด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน** และเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีเท่าเทียมกัน** ซึ่งแบบสอบแต่ละสัดส่วน มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบค่อนข้างต่ำ (< 0.300) ซึ่งในภาพรวมเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุด ส่วนใหญ่คือสัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 1 ข้อ

2.1.1.4 คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบทีมีจำนวนข้อสอบทีตรวจให้คะแนนสองค่า และข้อสอบทีตรวจให้คะแนนหลายค่าแตกต่างกัน** 4 **สัดส่วน ทีปรับเทียบคะแนนด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน** และเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีไม่เท่าเทียมกัน** ซึ่งแบบสอบแต่ละสัดส่วน มีสัดส่วนข้อสอบร่วมแตกต่างกัน 6 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบค่อนข้างต่ำ (< 0.300) ซึ่งในภาพรวมเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุด ส่วนใหญ่คือสัดส่วนข้อสอบร่วมที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 1 ข้อ

**2.1.2 ผลการปรับเทียบคะแนนแต่ละวิธี ด้วยการเก็บรวบรวมข้อมูลด้วยรูปแบบทีแตกต่างกัน เมื่อพิจารณาตามสัดส่วนข้อสอบในแบบสอบที่ต่างกัน**

2.1.2.1 คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบรูปแบบผสมด้วยวิธีโค้งคุณลักษณะ** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีเท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนค่อนข้างต่ำ (< 0.300) โดยภาพรวมเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิด

ความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่าจำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ

2.1.2.2.คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบรูปแบบผสมด้วยวิธีโค้งคุณลักษณะ** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนค่อนข้างต่ำ (< 0.300) โดยภาพรวมเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่าจำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ

2.1.2.3 คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบรูปแบบผสมด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มทีเท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนค่อนข้างต่ำ (< 0.300) โดยภาพรวมเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ

2.1.2.4 คุณภาพของ**การปรับเทียบคะแนนกับแบบสอบรูปแบบผสมด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน** ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสอง**กลุ่มที่ไม่เท่าเทียมกัน** ที่มีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 เงื่อนไข มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนค่อนข้างต่ำ (< 0.300) โดยภาพรวมเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือ สัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ

**2.1.3 ผลการปรับเทียบคะแนนแต่ละวิธี ด้วยการเก็บรวบรวมข้อมูลด้วยรูปแบบทีแตกต่างกัน เมื่อพิจารณาในภาพรวม**

2.1.3.1 คุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมด้วยวิธีโค้งคุณลักษณะ ที่การเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความเท่าเทียมกันมีความคลาดเคลื่อนในการปรับเทียบค่อนข้างต่ำ (< 0.300) และในทำนองเดียวกัน เมื่อปรับเทียบด้วยวิธีโค้งคุณลักษณะกับกลุ่มผู้สอบที่มีความไม่เท่าเทียมกัน ก็ให้ผลการปรับเทียบที่มีความคลาดเคลื่อนค่อนข้างต่ำเช่นเดียวกัน

2.1.3.2 คุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่การเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความเท่าเทียมกันมีความ

คลาดเคลื่อนในการปรับเทียบค่อนข้างต่ำ (< 0.300) และในทำนองเดียวกัน เมื่อปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน กับกลุ่มผู้สอบที่มีความไม่เท่าเทียมกัน ก็ให้ผลการปรับเทียบที่มีความคลาดเคลื่อนค่อนข้างต่ำเช่นเดียวกัน

จากข้างต้นสรุปได้ว่าวิธีการปรับเทียบด้วยวิธีโค้งคุณลักษณะ และการปรับค่าพารามิเตอร์พร้อมกัน ที่มีรูปแบบการเก็บรวบรวมข้อมูลที่แตกต่างกันแต่ละวิธี สำหรับแบบสอบรูปแบบผสมที่มีสัดส่วนข้อสอบในแบบสอบที่แตกต่างกัน และสัดส่วนของข้อสอบร่วมที่แตกต่างกัน เกิดความคลาดเคลื่อนในการปรับเทียบค่อนข้างต่ำในทุกๆ เงื่อนไข

2.2 **สรุปผลการวิจัยตามวัตถุประสงค์ข้อที่ 2** เปรียบเทียบคุณภาพผลการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ (Characteristic Curve) และวิธีการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent Calibration) เมื่อแบบแผนการเก็บรวบรวมข้อมูล และสัดส่วนของข้อสอบทั้ง 2 ชนิดในแบบสอบรูปแบบผสม และสัดส่วนของข้อสอบทั้ง 2 ชนิดที่เป็นข้อสอบร่วมของแบบสอบรูปแบบผสม ต่างกัน

ผลการวิจัยสรุปได้ดังนี้

**2.2.1 ผลการเปรียบเทียบคุณภาพของการปรับเทียบคะแนนเมื่อแบบสอบมีสัดส่วนข้อสอบทั้งสองชนิดต่างกัน**

การทดสอบความแตกต่างระหว่างค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (MSE) ของแบบสอบ 4 สัดส่วนที่ปรับเทียบด้วยวิธีการปรับเทียบแตกต่างกัน 2 วิธี คือ วิธีโค้งคุณลักษณะ (Characteristic curve : CC) และการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent calibration : CO) เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบแบบกลุ่มเท่าเทียม (Equivalent group : EQ) และกลุ่มผู้สอบไม่เท่าเทียม (Non-Equivalent group : NEQ) ผลการทดสอบรายคู่แสดงดังตารางที่ 38

ตารางที่ 38 ผลการทดสอบรายคู่ที่มีนัยสำคัญทางสถิติที่ระดับ .05 พิจารณาตามสัดส่วนของแบบสอบ 4 สัดส่วน

| เงื่อนไข | | MSE | | | |
|---|---|---|---|---|---|
| | | DS 50 : PS | DS 60 : PS 8 | DS 70 : PS 6 | DS 80 : PS 4 |
| CC - EQ | DS 50 : PS 10 | - | × | × | × |
| | DS 60 : PS 8 | | - | × | × |
| | DS 70 : PS 6 | | | - | × |
| | DS 80 : PS 4 | | | | - |
| CC - NEQ | DS 50 : PS 10 | - | × | 50 : 10 > 70 : 6 | 50 : 10 > 80 : 4 |
| | DS 60 : PS 8 | | - | × | 60 : 8 > 80 : 4 |
| | DS 70 : PS 6 | | | - | × |
| | DS 80 : PS 4 | | | | - |
| CO - EQ | DS 50 : PS 10 | - | × | × | × |
| | DS 60 : PS 8 | | - | × | × |
| | DS 70 : PS 6 | | | - | × |
| | DS 80 : PS 4 | | | | - |
| CO - NEQ | DS 50 : PS 10 | - | × | × | × |
| | DS 60 : PS 8 | | - | × | × |
| | DS 70 : PS 6 | | | - | × |
| | DS 80 : PS 4 | | | | - |

× ไม่พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

จากตารางที่ 38 สรุปได้ว่า ผลการทดสอบรายคู่พบความแตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เพียง 3 คู่ ได้แก่ แบบสอบที่มีสัดส่วน DS 50 : PS 10 มีค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสองต่างจากแบบสอบสัดส่วน DS 70 : PS 6, แบบสอบที่มีสัดส่วน DS 50 : PS 10 มีค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสองต่างจากแบบสอบสัดส่วน DS 80 : PS 4, และแบบสอบที่มีสัดส่วน DS 60 : PS 8 มีค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสองต่างจากแบบสอบสัดส่วน DS 80 : PS 4 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

**2.2.2 ผลการเปรียบเทียบคุณภาพของการปรับเทียบคะแนนเมือแบบสอบมีสัดส่วนข้อสอบร่วมต่างกัน**

การทดสอบความแตกต่างระหว่างค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (MSE) ของข้อสอบร่วม 6 สัดส่วนที่ปรับเทียบด้วยวิธีการปรับเทียบแตกต่างกัน 2 วิธี คือ วิธีโค้ง

คุณลักษณะ (Characteristic curve : CC) และการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent calibration : CO) เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบแบบกลุ่มเท่าเทียม (Equivalent group : EQ) และกลุ่มผู้สอบไม่เท่าเทียม (Non-Equivalent group : NEQ) ผลการทดสอบรายคู่แสดงดังตารางที่ 39

ตารางที่ 39 ผลการทดสอบรายคู่ที่มีนัยสำคัญทางสถิติที่ระดับ .05 พิจารณาตามสัดส่วนของข้อสอบร่วม 6 สัดส่วน

| เงื่อนไข | | MSE | | | | | |
|---|---|---|---|---|---|---|---|
| | | 15/2 | 20/2 | 25/2 | 15/1 | 10/2 | 5/3 |
| CC - EQ | 15/2 | - | × | × | × | × | × |
| | 20/2 | | - | × | × | × | × |
| | 25/2 | | | - | × | × | × |
| | 15/1 | | | | - | × | × |
| | 10/2 | | | | | - | × |
| | 5/3 | | | | | | - |
| CC - NEQ | 15/2 | - | × | × | × | × | × |
| | 20/2 | | - | × | × | × | × |
| | 25/2 | | | - | × | × | × |
| | 15/1 | | | | - | × | × |
| | 10/2 | | | | | - | × |
| | 5/3 | | | | | | - |
| CO - EQ | 15/2 | - | × | × | × | × | × |
| | 20/2 | | - | × | × | × | × |
| | 25/2 | | | - | × | × | × |
| | 15/1 | | | | - | × | × |
| | 10/2 | | | | | - | × |
| | 5/3 | | | | | | - |
| CO - NEQ | 15/2 | - | × | × | × | × | × |
| | 20/2 | | - | × | × | × | × |
| | 25/2 | | | - | × | × | × |
| | 15/1 | | | | - | × | × |
| | 10/2 | | | | | - | × |
| | 5/3 | | | | | | - |

× ไม่พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

จากตารางที่ 39 สรุปได้ว่า การวิเคราะห์ทดสอบคุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีสัดส่วนของข้อสอบร่วมทั้งสองชนิดต่างกัน 6 สัดส่วน ด้วยการวิเคราะห์ความแตกต่างระหว่างค่าเฉลี่ย MSE ของข้อสอบร่วมแต่ละสัดส่วน ไม่พบความแตกต่างระหว่างค่าเฉลี่ย MSE ของข้อสอบร่วมแต่ละสัดส่วนอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

**2.2.3 ผลการเปรียบเทียบคุณภาพของการปรับเทียบคะแนนเมือรูปแบบการเก็บรวบรวมข้อมูลต่างกัน**

การทดสอบความแตกต่างระหว่างค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (MSE) ของแบบสอบที่ปรับเทียบด้วยวิธีการปรับเทียบแตกต่างกัน 2 วิธี คือ วิธีโค้งคุณลักษณะ (Characteristic curve : CC) และการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent calibration : CO) เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบแบบกลุ่มเท่าเทียม (Equivalent group : EQ) และกลุ่มผู้สอบไม่เท่าเทียม (Non-Equivalent group : NEQ) ผลการทดสอบรายคู่แสดงดังตารางที่ 40

ตารางที่ 40 ผลการทดสอบรายคู่ที่มีนัยสำคัญทางสถิติที่ระดับ .05 พิจารณาตามวิธีการปรับเทียบและการเก็บรวบรวมข้อมูล

| **เงื่อนไข** | MSE | | | |
|---|---|---|---|---|
| | CC – EQ | CC – NEQ | CO – EQ | CO - NEQ |
| CC – EQ | - | CC-NEQ > CC-EQ | CC-EQ > CO-EQ | CC-EQ > CO-NEQ |
| CC – NEQ | | - | CC-NEQ > CO-EQ | CC-NEQ > CO-NEQ |
| CO – EQ | | | - | × |
| CO - NEQ | | | | - |

× ไม่พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

จากตารางที่ 40 สรุปได้ว่า พบความแตกต่างระหว่างค่าเฉลี่ยดัชนี MSE ของการปรับเทียบด้วยวิธีโค้งคุณลักษณะ (CC) กับวิธีการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกัน (CO) ที่มีการออกแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน 2 รูปแบบ คือกลุ่มเท่าเทียม (EQ) และกลุ่มไม่เท่าเทียม (NEQ) ดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

- การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการออกแบบการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียมมีความคลาดเคลื่อนในการปรับเทียบสูงกว่ากลุ่มเท่าเทียมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

- การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการออกแบบการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียมมีความคลาดเคลื่อนในการปรับเทียบสูงกว่าการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกันที่มีการออกแบบการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

- การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการออกแบบการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียมมีความคลาดเคลื่อนในการปรับเทียบสูงกว่าการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกันที่มีการออกแบบการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

- การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการออกแบบการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียมมีความคลาดเคลื่อนในการปรับเทียบสูงกว่าการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกันที่มีการออกแบบการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

- การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการออกแบบการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียมมีความคลาดเคลื่อนในการปรับเทียบสูงกว่าการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกันที่มีการออกแบบการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

# บทที่ 5
# สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้เป็นการศึกษาคุณภาพการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ โดยใช้วิธีการจำลองข้อมูล มีวัตถุประสงค์หลัก คือ เพื่อเปรียบเทียบคุณภาพของวิธีการปรับเทียบ 2 วิธี คือ วิธีโค้งคุณลักษณะ (Characteristic curve) และการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent calibration) เงื่อนไขที่ทำการศึกษามีทั้งหมด 96 เงื่อนไข (2×6×4×2) ประกอบด้วย วิธีการปรับเทียบคะแนน 2 วิธี สัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่ากับข้อสอบที่ตรวจให้คะแนนหลายค่าในแบบสอบ จำนวน 4 สัดส่วน สัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่ากับข้อสอบที่ตรวจให้คะแนนหลายค่าในข้อสอบร่วม จำนวน 6 สัดส่วน และรูปแบบการเก็บรวบรวมข้อมูล 2 รูปแบบ คือ กลุ่มเท่าเทียมและกลุ่มไม่เท่าเทียม แต่ละเงื่อนไขมีการทำซ้ำเงื่อนไขละ 100 ครั้ง

แบบสอบรูปแบบผสมเป็นการผสมระหว่างข้อสอบที่มีการตรวจให้คะแนนสองค่า จำลองข้อมูลตามโมเดลโลจีสติกแบบ 3 พารามิเตอร์ และข้อสอบที่ตรวจให้คะแนนหลายค่า จำลองข้อมูลตามโมเดล GPC ที่มีการให้คะแนน 6 ระดับ (0 – 5 คะแนน) การออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน กำหนดให้มีค่าเฉลี่ยความสามารถเท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 และสำหรับการออกแบบการเก้บรวบรวมข้อมูลกับผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียม กำหนดให้กลุ่มหนึ่งมีค่าเฉลี่ยความสามารถเท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 อีกกลุ่มหนึ่งมีค่าเฉลี่ยความสามารถเท่ากับ 1 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 จำนวนผู้สอบในแต่ละกลุ่มมีจำนวนกลุ่มละ 1000 คน การออกแบบการเก็บข้อมูลเป็นแบบสอบร่วมภายในซึ่งศึกษาใน 2 มิติ คือ มิติที่มีจำนวนข้อสอบร่วมตั้งแต่ร้อยละ 20 ของแบบสอบทั้งฉบับ และมิติคะแนนรวมระหว่างข้อสอบ DS และ PS ในข้อสอบร่วมรวมร้อยละ 20 ซึ่งแต่ละระดับจะมีสัดส่วนของข้อสอบร่วมจะประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่ากับข้อสอบที่ตรวจให้คะแนนหลายค่าในสัดส่วนแตกต่างต่างกันไป

การจำลองข้อมูลมี 3 ขั้นตอนคือ 1) การกำหนดลักษณะของผู้สอบ ได้แก่ จำนวนผู้สอบ และการแจกแจงความสามารถของกลุ่มผู้สอบ 2) การกำหนดข้อมูลเกี่ยวกับข้อสอบ ได้แก่ จำนวนข้อสอบ โมเดลที่ต้องการ และการแจกแจงพารามิเตอร์ของข้อสอบ ซึ่งเป็นการสร้างพารามิเตอร์ที่แท้จริงของข้อสอบ 3) สร้างรูปแบบของการตอบข้อสอบรายข้อของผู้สอบแต่ละคน ในการวิจัยครั้งนี้ทำการศึกษาคุณภาพการปรับเทียบคะแนนตามแนวทฤษฎีการตอบสนองข้อสอบที่มีการออกแบบการเก็บรวบรวมข้อมูล 2 รูปแบบ คือ การเก็บรวบรวมข้อมูลกับกลุ่มสอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน และ

การเก็บรวบรวมข้อมูลกับกลุ่มสอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน ดังนั้นในการจำลองข้อมูล ผู้วิจัยต้องจำลองข้อมูลรูปแบบการตอบจำนวน 3 ชุด ชุดที่ 1 เป็นผลการตอบข้อสอบสำหรับกลุ่มผู้สอบที่ทำแบบสอบฐาน (Test A) ผู้สอบกลุ่มนี้มีค่าความสามารถเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 รูปแบบการตอบชุดที่ 2 เป็นผลการตอบข้อสอบสำหรับกลุ่มผู้สอบที่ทำแบบสอบปรับเทียบ (Test B1) ตามรูปแบบการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียมจึงจำลองข้อมูลให้ผู้สอบกลุ่มนี้มีค่าความสามารถเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 และรูปแบบการตอบชุดที่ 3 เป็นผลการตอบข้อสอบสำหรับกลุ่มผู้สอบที่ทำแบบสอบปรับเทียบ (B2) ตามรูปแบบการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียมจึงจำลองข้อมูลให้ผู้สอบกลุ่มนี้มีค่าความสามารถเฉลี่ยเท่ากับ 1 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 ในกระบวนการจำลองข้อมูลทั้งหมดนับตั้งแต่การกำหนดค่าพารามิเตอร์ที่แท้จริงจนกระทั่งได้รูปแบบการตอบข้อสอบแต่ละข้อของผู้สอบแต่ละคน ผู้วิจัยใช้โปรแกรม Wingen ที่พัฒนาโดย Han (2007) ซึ่งเป็นโปรแกรมที่พัฒนาขึ้นเพื่อจำลองข้อมูลให้เป็นไปตามโมเดลตามทฤษฎีการตอบสนองข้อสอบ เมื่อได้ข้อมูลรูปแบบการตอบที่จำลองขึ้นแล้ว ผู้วิจัยได้วิเคราะห์ประมาณค่าพารามิเตอร์ข้อสอบโดยใช้โปรแกรม MULTILOG และเพื่อให้มั่นใจว่าข้อมูลที่จำลองขึ้นมานั้นมีความถูกต้อง ผู้วิจัยทำการตรวจสอบความถูกต้องของข้อมูลที่ได้จากการจำลองข้อมูลด้วยโปรแกรม Wingen (Han, 2007) โดยการวิเคราะห์เปรียบเทียบค่าพารามิเตอร์ที่แท้จริงจากโปรแกรม Wingen และค่าพารามิเตอร์ที่ได้จากการประมาณค่าด้วยโปรแกรม MULTILOG ด้วยการวิเคราะห์ t-test แล้ววิเคราะห์ตรวจสอบความเป็นเอกมิติของข้อมูลตามเงื่อนไขสัดส่วนของข้อสอบทั้ง 2 ชนิดในแบบสอบแต่ละฉบับโดยการวิเคราะห์ปัจจัย (Factor analysis) และวิเคราะห์ค่าความเที่ยงของแบบสอบ เมื่อได้ข้อมูลที่มีคุณภาพและเป็นไปตามข้อตกลงเบื้องต้นแล้ว จึงวิเคราะห์ปรับเทียบคะแนนตามวิธีโค้งคุณลักษณะ (Characteristic) โดยใช้โปรแกรม IRTEQ (Han, 2007) และวิธีการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent Calibration) โดยใช้โปรแกรม MULTILOG

เกณฑ์สำหรับเปรียบเทียบคุณภาพการปรับเทียบพิจารณาจากเกณฑ์โค้งคุณลักษณะลำดับขั้นการตอบ (Category Response Curve Criterion : CRC) ซึ่งเกณฑ์โค้งคุณลักษณะลำดับขั้นการตอบ (CRC) มีพื้นฐานมาจากความแตกต่างระหว่างโค้งลำดับขั้นการตอบที่ประมาณค่าได้ (estimated category response curve) กับโค้งลำดับขั้นการตอบที่แท้จริง (true category response curve) บนแบบสอบฉบับเทียบคะแนน (new form) ซึ่งเกณฑ์ CRC เรียกอย่างง่ายว่าค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (mean square error : MSE) (Hanson and Beguin, 2002; Kim, 2004, Kim and Lee, 2006) ซึ่งเป็นเกณฑ์ที่ถูกนำมาใช้ในการเปรียบเทียบ

ประสิทธิภาพสัมพัทธ์ของการปรับเทียบตามแนวทฤษฎีการตอบสนองข้อสอบสำหรับแบบสอบรูปแบบผสม

ในบทนี้ผู้วิจัยนำเสนอสรุป อภิปรายผลการวิจัย และข้อเสนอแนะ ดังต่อไปนี้

## สรุปผลการวิจัย

สำหรับผลการเปรียบเทียบคุณภาพการปรับเทียบตามทฤษฎีการตอบสนองข้อสอบ 2 วิธี คือ วิธีการโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ภายใต้เงื่อนไขแบบแผนการเก็บรวบรวมข้อมูลที่ต่างกัน 2 รูปแบบคือ กลุ่มเท่าเทียม และกลุ่มไม่เท่าเทียม สัดส่วนข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนหลายค่าในแบบสอบที่แตกต่างกัน 4 ระดับ สัดส่วนข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนหลายค่าในข้อสอบร่วมที่แตกต่างกัน 6 ระดับ ซึ่ง 3 ระดับแรกพิจารณาในมิติของจำนวนข้อที่มากกว่าร้อยละ 20 ของแบบสอบทั้งฉบับ และมิติที่ 2 พิจารณาคะแนนรวมเท่ากับร้อยละ 20 ของคะแนนจากแบบสอบที่มีเต็ม 100 คะแนน โดยแต่ละระดับมีสัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนหลายค่าแตกต่างกันไป ผลการวิจัยสามารถสรุปได้ดังนี้

### 1. คุณภาพของการปรับเทียบ

#### 1.1 ผลการปรับเทียบคะแนนแต่ละวิธี ด้วยการเก็บรวบรวมข้อมูลด้วยรูปแบบที่แตกต่างกัน เมื่อพิจารณาตามสัดส่วนข้อสอบทั้งสองชนิดในข้อสอบร่วม

1.1.1 คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม ที่ปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่มที่เท่าเทียมกันและกลุ่มไม่เท่าเทียม มีค่าความคลาดเคลื่อนในการปรับเทียบค่อนข้างต่ำ (<0.300) ซึ่งในภาพรวมเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุด ส่วนใหญ่คือ ข้อสอบร่วมที่ถูกพิจารณาในมิติที่ 2 ซึ่งประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 1 ข้อ จำนวนข้อของข้อสอบร่วมมีจำนวนต่ำกว่าร้อยละ 20 ของจำนวนข้อสอบในแบบสอบทั้งฉบับ

1.1.2 คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม ที่ปรับเทียบคะแนนด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่มที่เท่าเทียมกัน และกลุ่มไม่เท่าเทียม มีค่าความคลาดเคลื่อนในการปรับเทียบค่อนข้างต่ำ

(<0.300) ซึ่งในภาพรวมเงื่อนไขสัดส่วนข้อสอบร่วมที่ทำให้เกิดความคลาดเคลื่อนสูงสุด ส่วนใหญ่คือข้อสอบร่วมที่ถูกพิจารณาในมิติที่ 2 ซึ่งประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 15 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 1 ข้อ จำนวนข้อของข้อสอบร่วมมีจำนวนต่ำกว่าร้อยละ 20 ของจำนวนข้อสอบในแบบสอบทั้งฉบับ

**1.2 ผลการปรับเทียบคะแนนแต่ละวิธี ทีมีการเก็บรวบรวมข้อมูลด้วยรูปแบบที่แตกต่างกัน เมื่อพิจารณาตามสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบ**

1.2.1 คุณภาพของการปรับเทียบคะแนนกับแบบสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนหลายค่า แตกต่างกัน 4 สัดส่วน ปรับเทียบด้วยวิธีโค้งคุณลักษณะ เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่มที่เท่าเทียมกัน และกลุ่มไม่เท่าเทียม ผลการปรับเทียบมีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนค่อนข้างต่ำ (< 0.300) โดยภาพรวมเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือสัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ ซึ่งเป็นเงื่อนไขที่มีจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าน้อยที่สุดในแบบสอบ

1.2.2 คุณภาพของการปรับเทียบคะแนนกับแบบสอบรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนหลายค่า แตกต่างกัน 4 สัดส่วน ปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่มที่เท่าเทียมกันและไม่เท่าเทียมกัน มีค่าความคลาดเคลื่อนในการปรับเทียบคะแนนค่อนข้างต่ำ (< 0.300) โดยภาพรวมเงื่อนไขสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบที่ทำให้เกิดความคลาดเคลื่อนสูงสุดคือสัดส่วนข้อสอบที่ประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่า จำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ ซึ่งเป็นเงื่อนไขที่มีจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าน้อยที่สุดในแบบสอบ

**1.3 ผลการปรับเทียบคะแนนแต่ละวิธี ทีเก็บรวบรวมข้อมูลด้วยรูปแบบทีแตกต่างกัน เมื่อพิจารณาในภาพรวม**

1.3.1 คุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่ปรับเทียบด้วยวิธีโค้งคุณลักษณะ เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความเท่าเทียมกันมีความคลาดเคลื่อนในการปรับเทียบค่อนข้างต่ำ (<0.300) และเมื่อปรับเทียบด้วยวิธีโค้งคุณลักษณะกับกลุ่มผู้สอบที่มีความไม่เท่าเทียมกัน ก็ให้ผลการปรับเทียบที่มีความคลาดเคลื่อนค่อนข้างต่ำเช่นเดียวกัน

1.3.2 คุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่ปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน เก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความเท่าเทียมกันมีความคลาดเคลื่อนในการปรับเทียบค่อนข้างต่ำ (<0.300) และเมื่อปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบที่มีความไม่เท่าเทียมกัน ก็ให้ผลการปรับเทียบที่มีความคลาดเคลื่อนค่อนข้างต่ำเช่นเดียวกัน

จากข้างต้นสรุปได้ว่าวิธีการปรับเทียบด้วยวิธีโค้งคุณลักษณะ และการปรับค่าพารามิเตอร์พร้อมกัน ที่มีเงื่อนไขที่ทำการศึกษาแตกต่างกันไป (การออกแบบการเก็บรวบรวมข้อมูล, สัดส่วนของข้อสอบทั้งสองชนิดในแบบสอบ,สัดส่วนของข้อสอบทั้ง 2 ชนิดในข้อสอบร่วม) ผลการปรับเทียบแสดงให้เห็นว่าเกิดความคลาดเคลื่อนในการปรับเทียบค่อนข้างต่ำในทุกๆ เงื่อนไข

**2. การเปรียบเทียบคุณภาพการปรับเทียบ**

**2.1 ผลการเปรียบเทียบคุณภาพของการปรับเทียบคะแนนเมือแบบสอบมีสัดส่วนข้อสอบทั้งสองชนิดในแบบสอบต่างกัน**

2.1.1 คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม ที่มีสัดส่วนของข้อสอบทั้งสองชนิดในแบบสอบที่แตกต่างกัน 4 สัดส่วน ปรับเทียบแบบโค้งคุณลักษณะ ที่มีการเก็บรวบรวมข้อมูลทั้งแบบกลุ่มเท่าเทียม และกลุ่มไม่เท่าเทียม แสดงให้เห็นว่าคุณภาพของการปรับเทียบมีความคลาดเคลื่อนต่ำลงเมื่อทำการปรับเทียบกับแบบสอบที่ประกอบด้วยสัดส่วนข้อสอบที่มีการตรวจให้คะแนนสองค่าในแบบสอบจำนวนเพิ่มขึ้น นั่นคือ แบบสอบที่มีสัดส่วนข้อสอบที่ตรวจให้คะแนนสองค่าจำนวน 80 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 4 ข้อ จะทำให้ผลการปรับเทียบมีความคลาดเคลื่อนต่ำกว่า แบบสอบที่มีสัดส่วนข้อสอบที่ตรวจให้คะแนนสองค่าจำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ

2.1.2 คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม ที่มีสัดส่วนของข้อสอบในแบบสอบที่แตกต่างกัน 4 สัดส่วน ด้วยวิธีการปรับเทียบแบบการปรับค่าพารามิเตอร์พร้อมกัน ที่มีการเก็บรวบรวมข้อมูลทั้งแบบกลุ่มเท่าเทียม และกลุ่มไม่เท่าเทียม แสดงให้เห็นว่าคุณภาพของการปรับเทียบมีความคลาดเคลื่อนต่ำลงเมื่อทำการปรับเทียบกับแบบสอบที่ประกอบด้วยสัดส่วนข้อสอบที่มีการตรวจให้คะแนนสองค่าในแบบสอบจำนวนเพิ่มขึ้น นั่นคือ แบบสอบที่มีสัดส่วนข้อสอบที่ตรวจให้คะแนนสองค่าจำนวน 80 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 4 ข้อ จะทำให้ผลการปรับเทียบมีความคลาดเคลื่อนต่ำกว่า แบบสอบที่มี

สัดส่วนข้อสอบที่ตรวจให้คะแนนสองค่าจำนวน 50 ข้อ และข้อสอบที่ตรวจให้คะแนนหลายค่าจำนวน 10 ข้อ

2.1.3 การวิเคราะห์เปรียบเทียบคุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีสัดส่วนของข้อสอบทั้งสองชนิดต่างกัน 4 สัดส่วน ด้วยการวิเคราะห์ความแตกต่างระหว่างค่าเฉลี่ย MSE ของแบบสอบแต่ละสัดส่วน ไม่พบความแตกต่างระหว่างค่าเฉลี่ย MSE ของแบบสอบแต่ละสัดส่วนอย่างมีนัยสำคัญทางสถิติ

**2.2 ผลการเปรียบเทียบคุณภาพของการปรับเทียบคะแนนเมือแบบสอบมีสัดส่วนข้อสอบทั้งสองชนิดในข้อสอบร่วมต่างกัน**

2.2.1 คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม ที่มีสัดส่วนของข้อสอบร่วมในแบบสอบที่แตกต่างกัน 6 สัดส่วน ด้วยวิธีการปรับเทียบแบบโค้งคุณลักษณะ ที่มีการเก็บรวบรวมข้อมูลทั้งแบบกลุ่มเท่าเทียม และกลุ่มไม่เท่าเทียม แสดงให้เห็นว่าคุณภาพของการปรับเทียบมีความคลาดเคลื่อนสูงขึ้นเมื่อทำการปรับเทียบกับแบบสอบที่ประกอบด้วยสัดส่วนข้อสอบร่วมที่มีการตรวจให้คะแนนสองค่าในแบบสอบจำนวนเพิ่มขึ้น และผลการปรับเทียบที่ใช้ข้อสอบร่วมจำนวนมากกว่าร้อยละ 20 ของข้อสอบทั้งหมด (25/2, 20/2 และ 15/2) จะให้ผลการปรับเทียบที่มีความคลาดเคลื่อนต่ำกว่าการปรับเทียบที่ใช้ข้อสอบร่วมที่มีคะแนนรวมร้อยละ 20 จากคะแนนทั้งหมด (จำนวนข้อต่ำกว่าร้อยละ 20 ของข้อสอบทั้งหมด ; 15/1, 10/2, 5/3)

2.2.2 คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม ที่มีสัดส่วนของข้อสอบร่วมในแบบสอบที่แตกต่างกัน 6 สัดส่วน ด้วยวิธีการปรับเทียบแบบการปรับค่าพารามิเตอร์พร้อมกัน ที่มีการเก็บรวบรวมข้อมูลทั้งแบบกลุ่มเท่าเทียม และกลุ่มไม่เท่าเทียม แสดงให้เห็นว่าคุณภาพของการปรับเทียบมีความคลาดเคลื่อนสูงขึ้นเมื่อทำการปรับเทียบกับแบบสอบที่ประกอบด้วยสัดส่วนข้อสอบร่วมที่มีการตรวจให้คะแนนสองค่าในแบบสอบจำนวนเพิ่มขึ้น และผลการปรับเทียบที่ใช้ข้อสอบร่วมจำนวนมากกว่าร้อยละ 20 ของข้อสอบทั้งหมด (25/2, 20/2 และ 15/2) จะให้ผลการปรับเทียบที่มีความคลาดเคลื่อนต่ำกว่าการปรับเทียบที่ใช้ข้อสอบร่วมที่มีคะแนนรวมร้อยละ 20 จากคะแนนทั้งหมด (จำนวนข้อต่ำกว่าร้อยละ 20 ของข้อสอบทั้งหมด ; 15/1, 10/2, 5/3)

2.2.3 การวิเคราะห์เปรียบเทียบคุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีสัดส่วนของข้อสอบร่วมทั้งสองชนิดต่างกัน 6 สัดส่วน ด้วยการวิเคราะห์ความแตกต่างระหว่างค่าเฉลี่ย MSE ของข้อสอบร่วมแต่ละสัดส่วน ไม่พบความแตกต่างระหว่างค่าเฉลี่ย MSE ของข้อสอบร่วมแต่ละสัดส่วนอย่างมีนัยสำคัญทางสถิติ

**2.3 ผลการเปรียบเทียบคุณภาพของการปรับเทียบคะแนนเมื่อรูปแบบการเก็บรวบรวมข้อมูลต่างกัน**

2.3.1 คุณภาพของการปรับเทียบคะแนนกับแบบสอบรูปแบบผสมด้วยวิธีโค้งคุณลักษณะ ที่ออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน เกิดความคลาดเคลื่อนในการปรับเทียบต่ำกว่าการปรับเทียบด้วยวิธีโค้งคุณลักษณะที่ออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน เมื่อทำการทดสอบความแตกต่างทางสถิติ พบว่า เกิดความแตกต่างระหว่างค่าเฉลี่ยดัชนี MSE ของการปรับเทียบด้วยวิธีโค้งคุณลักษณะ ที่มีการออกแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน 2 รูปแบบ โดยที่ การออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มไม่เท่าเทียมมีความคลาดเคลื่อนในการปรับเทียบสูงกว่ากลุ่มเท่าเทียม

2.3.2 คุณภาพของการปรับเทียบคะแนนกับแบบสอบรูปแบบผสมด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่ออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน เกิดความคลาดเคลื่อนในการปรับเทียบใกล้เคียงกับการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกัน ที่ออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน เมื่อทำการทดสอบความแตกต่างทางสถิติ พบว่า วิธีการปรับค่าพารามิเตอร์พร้อมกันที่มีการออกแบบการเก็บรวบรวมข้อมูลที่แตกต่างกัน 2 รูปแบบ มีค่าเฉลี่ยดัชนี MSE ไม่แตกต่างกัน

2.3.3 คุณภาพของการปรับเทียบคะแนนกับแบบสอบรูปแบบผสมด้วยวิธีโค้งคุณลักษณะ ที่ออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน เกิดความคลาดเคลื่อนในการปรับเทียบสูงกว่าการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกันที่ออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน เมื่อทำการทดสอบความแตกต่างทางสถิติ พบว่า เกิดความแตกต่างระหว่างค่าเฉลี่ยดัชนี MSE ของการปรับเทียบด้วยวิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่มีการออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน โดยที่ การปรับเทียบด้วยวิธีโค้งคุณลักษณะมีความคลาดเคลื่อนในการปรับเทียบคะแนนสูงกว่าการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน

2.3.4 คุณภาพของการปรับเทียบคะแนนกับแบบสอบรูปแบบผสมด้วยวิธีโค้งคุณลักษณะ ที่ออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน เกิดความคลาดเคลื่อนในการปรับเทียบสูงกว่าการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกันที่ออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน เมื่อทำการทดสอบความแตกต่างทางสถิติ พบว่า เกิดความแตกต่างระหว่างค่าเฉลี่ยดัชนี MSE ของการปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน และ

วิธีการปรับค่าพารามิเตอร์พร้อมกันที่มีการออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน โดยที่ การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกันมีความคลาดเคลื่อนในการปรับเทียบคะแนนสูงกว่าการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกันที่มีการออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน

2.3.5 คุณภาพของการปรับเทียบคะแนนกับแบบสอบรูปแบบผสมด้วยวิธีโค้งคุณลักษณะ ที่ออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน เกิดความคลาดเคลื่อนในการปรับเทียบสูงกว่าการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกันที่ออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน เมื่อทำการทดสอบความแตกต่างทางสถิติ พบว่า เกิดความแตกต่างระหว่างค่าเฉลี่ยดัชนี MSE ของการปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน และวิธีการปรับค่าพารามิเตอร์พร้อมกันที่มีการออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน โดยที่ การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกันมีความคลาดเคลื่อนในการปรับเทียบคะแนนสูงกว่าการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกันที่มีการออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน

2.3.6 คุณภาพของการปรับเทียบคะแนนกับแบบสอบรูปแบบผสมด้วยวิธีโค้งคุณลักษณะ ที่ออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน เกิดความคลาดเคลื่อนในการปรับเทียบสูงกว่าการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกันที่ออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน เมื่อทำการทดสอบความแตกต่างทางสถิติ พบว่า เกิดความแตกต่างระหว่างค่าเฉลี่ยดัชนี MSE ของการปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน และวิธีการปรับค่าพารามิเตอร์พร้อมกันที่มีการออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน โดยที่ การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกันมีความคลาดเคลื่อนในการปรับเทียบคะแนนสูงกว่าการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกันที่มีการออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน

จากข้างต้นสรุปได้ว่า วิธีการปรับเทียบแต่ละวิธีที่มีรูปแบบการเก็บรวบรวมข้อมูลต่างกัน เมื่อทำการปรับเทียบกับแบบสอบรูปแบบผสมที่มีสัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่า

เพิ่มขึ้น ทำให้ความคลาดเคลื่อนในการปรับเทียบคะแนนต่ำลง ในทางตรงกันข้าม ถ้าทำการปรับเทียบกับแบบสอบรูปแบบผสมที่สัดส่วนของข้อสอบร่วมมีจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าเพิ่มขึ้น จะทำให้ความคลาดเคลื่อนในการปรับเทียบคะแนนสูงขึ้น

เมื่อพิจารณาตามวิธีการปรับเทียบคะแนนและรูปแบบการเก็บรวบรวมข้อมูล พบว่า การปรับเทียบด้วยวิธีโค้งคุณลักษณะที่มีการเก็บรวบรวมข้อมูลที่แตกต่างกันทั้ง 2 วิธี จะมีความคลาดเคลื่อนสูงกว่าการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกันที่มีการเก็บรวบรวมข้อมูลที่แตกต่างกันทั้ง 2 วิธี และภายใต้วิธีการปรับเทียบคะแนนแบบโค้งคุณลักษณะ ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียมจะมีความคลาดเคลื่อนในการปรับเทียบคะแนนต่ำกว่าวิธีการปรับเทียบคะแนนแบบโค้งคุณลักษณะ ด้วยการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม ในขณะที่การปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกัน ความคลาดเคลื่อนที่เกิดจากการปรับเทียบคะแนนกับกลุ่มเท่าเทียม จะมีค่าใกล้เคียงกับการปรับเทียบด้วยกลุ่มไม่เท่าเทียม

## อภิปรายผลการวิจัย

การอภิปรายผลในการวิจัยนี้นำเสนอ 2 ประเด็น ประเด็นหลักคือ การอภิปรายผลตามสมมติฐานการวิจัย และการอภิปรายผลจากการจำลองข้อมูล โปรแกรม และเกณฑ์ที่ใช้ในการวิจัย มีรายละเอียดดังนี้

1. การอภิปรายผลตามสมมติฐานการวิจัย

1.1 การอภิปรายผลตามสมมติฐานการวิจัยข้อที่ 1 วิธีการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะที่มีแบบแผนการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียมน่าจะได้ผลการปรับเทียบที่เกิดความคลาดเคลื่อนต่ำกว่าวิธีการปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะที่มีแบบแผนการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม

ผลการวิจัยพบว่า ในภาพรวมผลการปรับเทียบด้วยวิธีโค้งคุณลักษณะที่ดำเนินการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน เกิดความคลาดเคลื่อนในการปรับเทียบสูงกว่าการปรับเทียบด้วยวิธีโค้งคุณลักษณะ ที่ดำเนินการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน เมื่อกำหนดให้เงื่อนไขอื่น ๆ นั้นคงที่ (สัดส่วนของข้อสอบทั้งสองชนิดในแบบสอบทั้งฉบับ และสัดส่วนของข้อสอบทั้งสองชนิดในข้อสอบร่วม) ผลที่ได้จากการ

ปรับเทียบพบว่าการปรับเทียบด้วยวิธีโค้งคุณลักษณะเกิดความคลาดเคลื่อนในการปรับเทียบต่ำ เมื่อมีการออกแบบการเก็บรวบรวมข้อมูลกับกลุมผู้สอบที่มีความเท่าเทียมกัน ซึ่งข้อค้นพบนี้เป็นการยืนยันผลการวิจัยที่ผ่านมา นั่นคือ ผลการศึกษาเปรียบเทียบคุณภาพการปรับเทียบคะแนนระหว่างวิธีการปรับค่าพารามิเตอร์แยกกัน และการปรับค่าพารามิเตอร์พร้อมกันของ Kim (2004) ซึ่งผลการวิจัยพบว่า การปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์แยกกันด้วยวิธีโค้งคุณลักษณะเหมาะสมกับการออกแบบการเก็บรวบรวมข้อมูลสำหรับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน ซึ่ง Kim และ Lee (2006) ได้ยืนยันข้อค้นพบนี้เช่นเดียวกัน จากการศึกษาเปรียบเทียบวิธีการปรับเทียบแบบการปรับค่าพารามิเตอร์แยกกัน 4 วิธี ผลการศึกษาพบว่าการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์แยกกันด้วยวิธีโค้งคุณลักษณะ เหมาะสมกับการออกแบบการเก็บรวบรวมข้อมูลสำหรับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน (Li, Lissitz, and Yang, 1999; Beguin, Hanson, and Glas, 2000; Tate, 2000; Kirkpatrick, 2005; Kim and Kolen, 2006; Kim and Lee, 2006)

1.2 การอภิปรายผลตามสมมติฐานข้อที่ 2 วิธีการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกันที่มีแบบแผนการเก็บรวบรวมข้อมูลแบบกลุ่มเท่าเทียมน่าจะได้ผลการปรับเทียบที่มีคลาดเคลื่อนใกล้เคียงกับวิธีการปรับเทียบคะแนนด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกันที่มีแบบแผนการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม และมีความคลาดเคลื่อนต่ำกว่าการปรับเทียบด้วยวิธีโค้งคุณลักษณะ

ผลการวิจัยพบว่า การปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกันมีความคลาดเคลื่อนต่ำ ไม่ว่าจะมีรูปแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถเท่าเทียมกัน หรือกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกันก็ตาม ซึ่งเห็นได้จากค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสองที่มีค่าต่ำกว่า 0.3 นอกจากนี้ ยังพบว่าในภาพรวมแล้ว การปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกันยังมีความคลาดเคลื่อนในการปรับเทียบคะแนนต่ำกว่าการปรับเทียบด้วยวิธีโค้งคุณลักษณะด้วย ข้อพิจารณาที่น่าสนใจเกี่ยวกับประสิทธิภาพสัมพัทธ์ของการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent Calibration) และการปรับค่าพารามิเตอร์แยกกัน (Seperated Calibration) คือ โดยทั่วไปแล้วการปรับค่าพารามิเตอร์พร้อมกันจะให้ผลที่น่าพอใจมากกว่าการปรับค่าพารามิเตอร์แยกกัน ภายใต้เงื่อนไขที่ศึกษาหลาย ๆ เงื่อนไข เนื่องมาจากมีความเชื่อเกี่ยวกับการปรับค่าพารามิเตอร์พร้อมกันว่าสามารถทำให้ได้สารสนเทศที่เป็นประโยชน์อย่างสมบูรณ์ และอาจจะขจัดความคลาดเคลื่อนบางอย่างที่เกิดขึ้นในการปรับเทียบได้อย่างมีศักยภาพ ซึ่งความไม่ถูกต้องในกระบวนการแปลงสเกลอาจจะเกิดขึ้นในการปรับ

ค่าพารามิเตอร์แยกกัน Kim และ Cohen (1998) ใช้กระบวนการจำลองข้อมูลสำหรับการตอบที่มีการให้คะแนนแบบสองค่า เพื่อเปรียบเทียบการปรับค่าพารามิเตอร์แยกกันด้วยวิธี Stocking – Lord กับการปรับค่าพารามิเตอร์พร้อมกัน โดยใช้ซอฟต์แวร์ที่ต่างกัน ผลการศึกษาพบว่า ในภาพรวมวิธีการปรับค่าพารามิเตอร์พร้อมกันให้ผลที่ถูกต้องมากกว่าวิธีการปรับค่าพารามิเตอร์แยกกัน Hanson และ Beguin (2002) ได้จำลองข้อมูลการตอบแบบ 2 ค่าสำหรับเปรียบเทียบประสิทธิภาพสัมพัทธ์ของการปรับค่าพารามิเตอร์แยกกัน (mean/mean, mean/sigma, Heabara, และ Stocking – Lord) กับวิธีการปรับค่าพารามิเตอร์พร้อมกัน ผลการศึกษาพบว่า โดยทั่วไปวิธีการปรับค่าพารามิเตอร์พร้อมกันให้ความคลาดเคลื่อนต่ำกว่าวิธีการปรับค่าพารามิเตอร์แยกกัน แม้ว่าจะไม่ใช้ผลที่ครอบคลุมทั้งหมดก็ตาม เนื่องจากในการปรับค่าพารามิเตอร์แยกกันด้วยวิธีโค้งคุณลักษณะ 2 วิธี ให้ผลการแปลงสเกลที่ถูกต้องมากกว่าวิธี mean/mean และ mean/sigma และ Kim และ Cohen (2002) ได้จำลองข้อมูลผลการตอบที่มีการให้คะแนนแบบหลายค่าภายใต้โมเดล GRM และเปรียบเทียบการปรับค่าพารามิเตอร์แยกกันด้วยวิธี Stocking – Lord และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ผลการศึกษาพบว่าวิธีการปรับค่าพารามิเตอร์พร้อมกันให้ผลที่คงที่ และมีความถูกต้องมากกว่าวิธีการปรับค่าพารามิเตอร์แบบแยกกันเพียงเล็กน้อย

1.3 การอภิปรายผลตามสมมติฐานข้อที่ 3 วิธีการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกันน่าจะได้ผลการปรับเทียบที่เกิดความคลาดเคลื่อนต่ำกว่าวิธีปรับด้วยวิธีโค้งคุณลักษณะเมื่อมีแบบแผนการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียม

ผลการวิจัยพบว่า ในภาพรวมผลการปรับเทียบด้วยวิธีการปรับค่าพารามิเตอร์พร้อมกันกับกลุ่มผู้สอบที่มีความไม่เท่าเทียมมีความคลาดเคลื่อนในการปรับเทียบต่ำกว่าการปรับเทียบด้วยวิธีโค้งคุณลักษณะ กับกลุ่มผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกัน เมื่อเงื่อนไขอื่น ๆ คงที่ (สัดส่วนของข้อสอบทั้งสองชนิดในแบบสอบทั้งฉบับ และสัดส่วนของข้อสอบทั้งสองชนิดในข้อสอบร่วม) ผลที่ได้จากการปรับเทียบพบว่าการปรับเทียบด้วยวิธีโค้งคุณลักษณะเกิดความคลาดเคลื่อนในการปรับเทียบต่ำ เมื่อมีการออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน แต่ถ้ากลุ่มผู้สอบสองกลุ่มมีความสามารถไม่เท่าเทียมกัน การปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกันจะมีความคลาดเคลื่อนต่ำกว่าวิธีโค้งคุณลักษณะ ทั้งนี้เนื่องมาจากการปรับค่าพารามิเตอร์พร้อมกัน (Concurrent calibration) เป็นการประมาณค่าพารามิเตอร์จากแบบสอบทั้ง 2 ฉบับพร้อมกัน ในการประมวลผลโดยคอมพิวเตอร์เพียง 1 ครั้ง ซึ่งเป็นส่วนที่รับประกันได้ว่าค่าประมาณพารามิเตอร์จะอยู่บนสเกลเดียวกัน นั่นคือการรวมข้อมูลการตอบจากผู้สอบทั้ง 2 กลุ่มและกำหนดให้ข้อสอบที่อีกกลุ่มหนึ่งไม่ได้ทำเป็นค่า “not reach” หรือค่าขาดหาย (missing)

ซึ่งการปรับค่าพารามิเตอร์พร้อมกันเป็นการนำชุดพารามิเตอร์ของข้อสอบ 2 ชุด จากแบบสอบ 2 ฉบับ วางบนสเกลร่วมกันโดยใช้การปรับค่าพารามิเตอร์พร้อมกัน (Concurrent calibration) ซึ่งโปรแกรมที่สามารถใช้ในการปรับสเกลพร้อมกันได้ คือ BILOG-MG (Zimowski, Muraki, Mislevy, and Bock, 1996), ICL (Hanson, 2002) และ MULTILOG (Thissen, 1991) โปรแกรมเหล่านี้มีคุณสมบัติที่สามารถใช้ประมาณค่ากับกลุ่มผู้สอบที่มีระดับความสามารถแตกต่างกันได้ งานวิจัยที่เปรียบเทียบการปรับค่าพารามิเตอร์พร้อมกัน กับการเชื่อมโยงแบบอื่นๆ โดยทั่วไปพบว่า การปรับสเกลพร้อมกันมีความถูกต้องมากกว่าวิธีการปรับเทียบแบบโค้งคุณลักษณะเมื่อศึกษาจากข้อมูลที่จำลองขึ้นโดยให้ข้อมูลสอดคล้องกับโมเดล (Kim and Cohen, 1998, 2002; Hanson and Beguin, 2002) อย่างไรก็ตาม Beguin, Hanson และ Glas (2000) พบว่า เมื่อกลุ่มผู้สอบ 2 กลุ่มมีระดับความสามารถไม่เท่าเทียมกัน และความสามารถเป็นแบบพหุมิติแต่มีความสัมพันธ์กันสูง การเชื่อมโยงโดยใช้วิธี Stocking – Lord จะทำให้เกิดความถูกต้องมากกว่าวิธีการปรับสเกลพร้อมกัน ข้อค้นพบเหล่านี้ว่าเสนอแนะให้เห็นว่า การปรับสเกลพร้อมกันอาจจะมีความแกร่งน้อยเมื่อมีการฝ่าฝืนข้อตกลงเบื้องต้นเมื่อเทียบกับวิธีโค้งคุณลักษณะที่เป็นการปรับสเกลแยกกัน Kolen and Brennan (in preparation) เสนอแนะว่า การใช้วิธีการเชื่อมโยงแบบโค้งคุณลักษณะในทางปฏิบัติแล้วดูเหมือนว่าจะปลอดภัยที่สุดและวิธีการปรับสเกลพร้อมกันสามารถใช้เหมือนกับเป็นส่วนเสริมของการปรับค่าพารามิเตอร์แยกกัน

ข้อพิจารณาที่น่าสนใจเกี่ยวกับประสิทธิภาพสัมพัทธ์ของการปรับค่าพารามิเตอร์พร้อมกันและแบบแยกกันกับวิธีการปรับเทียบที่แตกต่างกัน โดยทั่วไป การปรับค่าพารามิเตอร์พร้อมกันจะ ให้ผลที่น่าพอใจกว่าการปรับค่าพารามิเตอร์แยกกัน ภายใต้เงื่อนไขที่ศึกษาหลาย ๆ เงื่อนไข เนื่องมาจากมีความเชื่อเกี่ยวกับการปรับค่าพารามิเตอร์พร้อมกันว่าสามารถทำให้ได้สารสนเทศที่เป็นประโยชน์อย่างสมบูรณ์ และอาจจะขจัดความคลาดเคลื่อนบางอย่างที่เกิดขึ้นในการปรับเทียบได้อย่างมีศักยภาพ ซึ่งความไม่ถูกต้องในกระบวนการแปลงสเกลอาจจะเกิดขึ้นในการปรับค่าพารามิเตอร์แยกกัน

1.4 การอภิปรายผลตามสมมติฐานข้อที่ 4 วิธีการปรับเทียบด้วยวิธีโค้งคุณลักษณะและวิธีการปรับค่าพารามิเตอร์พร้อมกัน น่าจะให้ผลการปรับเทียบที่มีความคลาดเคลื่อนต่ำเมื่อสัดส่วนของข้อสอบทั้ง 2 ชนิด มีจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าในแบบสอบมีจำนวนมากขึ้น

ผลการปรับเทียบวิธีโค้งคุณลักษณะและวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่ดำเนินการเก็บรวบรวมข้อมูลแต่ละวิธีคือ การออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มี

ความสามารถเท่าเทียมกัน และกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน พบว่า ผลการปรับเทียบจะมีความคลาดเคลื่อนต่ำลงเมื่อสัดส่วนของข้อสอบทั้ง 2 ชนิดในแบบสอบประกอบด้วยข้อสอบที่ตรวจให้คะแนนสองค่าเพิ่มมากขึ้น ซึ่งผลของการศึกษานี้สอดคล้องกับการศึกษาของ Li, Lissitz และ Yang (1999) Bastari (2000), Tate (2000) และ Bastari (2002) ที่พบว่า เมื่อสัดส่วนของข้อสอบทที่ตรวจให้คะแนนสองค่าและตรวจให้คะแนนหลายค่าประกอบอยู่ในแบบสอบ และมีสัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่าเพิ่มมากขึ้น จะทำให้ผลการปรับเทียบมีความคลาดเคลื่อนต่ำลง ทั้งนี้อาจเนื่องมาจากมีผลการวิจัยสนับสนุนว่าเมื่อความยาวของแบบสอบเพิ่มมากขึ้นจะทำให้ประสิทธิภาพในการปรับเทียบคะแนนสูงตามไปด้วย ซึ่งความยาวของแบบสอบที่มากขึ้นนั้น ส่วนหนึ่งมาจากจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าเพิ่มมากขึ้นด้วย

1.5 การอภิปรายผลตามสมมติฐานข้อที่ 5 วิธีการปรับเทียบด้วยวิธีโค้งคุณลักษณะและวิธีการปรับค่าพารามิเตอร์พร้อมกัน น่าจะให้ผลการปรับเทียบที่มีความคลาดเคลื่อนต่ำเมื่อข้อสอบร่วมที่ประกอบด้วยข้อสอบทั้ง 2 ชนิด มีจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าที่อยู่ในข้อสอบร่วมจำนวนลดลง

ผลการปรับเทียบวิธีโค้งคุณลักษณะและวิธีการปรับค่าพารามิเตอร์พร้อมกัน ที่ดำเนินการเก็บรวบรวมข้อมูลแต่ละวิธีคือ การออกแบบการเก็บรวบรวมข้อมูลกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน และกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน พบว่า ผลการปรับเทียบจะมีความคลาดเคลื่อนต่ำลงเมื่อข้อสอบร่วมประกอบด้วยข้อสอบทั้ง 2 ชนิด และมีข้อสอบที่ตรวจให้คะแนนสองค่าจำนวนลดลง ซึ่งผลการวิจัยนี้สอดคล้องกับ Bastari (2000) ศึกษาเปรียบเทียบวิธีการปรับเทียบคะแนนตามแนวทฤฎีตอบสนองข้อสอบระหว่างวิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ตามเงื่อนไขต่าง ๆ ผลการวิจัยพบว่าวิธีการปรับค่าพารามิเตอร์พร้อมกันมีประสิทธิภาพมากกว่าวิธีโค้งคุณลักษณะ เมื่อแบบสอบมีจำนวนข้อสอบร่วมที่เพิ่มขึ้น (RMSE ของข้อสอบร่วม 20% ต่ำกว่า RMSE ของข้อสอบร่วม 10%) สอดคล้องกับงานวิจัยของ Tate (2000) ได้ศึกษาวิธีการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบด้วยวิธี moment 2 วิธี และวิธีโค้งคุณลักษณะตามแนวของ Stocking และ Lord พบว่า จำนวนข้อสอบร่วมที่มีปริมาณมากขึ้นจะทำให้สัมประสิทธิ์การปรับเทียบคะแนนมีความแม่นยำยิ่งขึ้น โดยได้รับการสนับสนุนผลการวิจัยจากการศึกษาของ Kim (2004) และ Kim และ Lee (2006) ที่ได้ศึกษาเปรียบเทียบการเทียบคะแนนตามแนวทฤษฎีการตอบสนองข้อสอบด้วยวิธี moment 2 วิธี และวิธีโค้งคุณลักษณะ 2 วิธี สอดคล้องกับงานวิจัยของ Tate (2000) ที่พบว่าวิธี โค้งคุณลักษณะ ให้ผลการปรับเทียบคะแนน

เป็นที่น่าพอใจกว่าวิธี moment ทั้งสองวิธี และการปรับเทียบโดยใช้ข้อสอบร่วม จะมีความคลาดเคลื่อนสูงขึ้นเมื่อใช้ข้อสอบเพียงชนิดใดชนิดหนึ่งเป็นข้อสอบร่วม

ผลการปรับเทียบคะแนนสำหรับแบบสอบที่มีรูปแบบผสมครั้งนี้ ผู้วิจัยใช้ข้อสอบร่วมที่ประกอบด้วยข้อสอบทั้ง 2 ชนิด คือข้อสอบที่ตรวจให้คะแนนสองค่า และข้อสอบที่ตรวจให้คะแนนหลายค่า ซึ่งเป็นปัจจัยหนึ่งที่ทำให้เกิดความคลาดเคลื่อนในการปรับเทียบต่ำลง ตามข้อเสนอแนะของ Kim (2004) และ Kim และ Lee (2006) นอกจากนี้ผู้วิจัยยังได้ศึกษาเกี่ยวกับสัดส่วนของข้อสอบร่วมตามข้อเสนอแนะของ Kolen และ Brennan (2004) ว่าควรศึกษาถึงอิทธิพลของคะแนนรวม หรือสัดส่วนของคะแนนระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า และข้อสอบที่ตรวจให้คะแนนหลายค่าในข้อสอบร่วมด้วย ซึ่งผลการวิจัยได้ขยายองค์ความรู้เกี่ยวกับผลการปรับทียบคะแนน นั่นคือ การพิจารณาคะแนนรวมของข้อสอบร่วมที่มีคะแนนร้อยละ 20 ของคะแนนเต็ม โดยภายใต้คะแนนรวมร้อยละ 20 นั้นจะประกอบด้วยข้อสอบทั้ง 2 ชนิด ในสัดส่วนที่ต่างกัน ซึ่งผลการวิจัยพบว่าเกิดความคลาดเคลื่อนในการปรับเทียบต่ำ และไม่แตกต่างจากการพิจารณาในมิติของจำนวนข้อสอบร่วมในแบบสอบที่มีจำนวนข้อร้อยละ 20 ขึ้นไป นอกจากนี้ ยังมีข้อค้นพบว่าคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบที่พิจารณาข้อสอบร่วมในมิติของคะแนนรวมร้อยละ 20 ซึ่งจะมีจำนวนข้อสอบร่วมน้อยกว่าร้อยละ 20 ของจำนวนข้อในแบบสอบทั้งฉบับ จะให้ผลการปรับเทียบที่มีความคลาดเคลื่อนสูงกว่าการพิจารณาข้อสอบร่วมในมิติของจำนวนข้อที่มากกว่าร้อยละ 20 ทั้งนี้อาจเนื่องมาจาก อิทธิพลของจำนวนข้อสอบร่วมจะมีผลต่อการปรับเทียบคะแนนโดยเฉพาะอย่างยิ่งเมื่อทำการปรับเทียบคะแนนด้วยวิธีการปรับค่าพารามิเตอร์แยกกัน (Kim and Kolen, 2006; Kirkpatrick, 2005)

จากผลการวิจัยครั้งนี้จะเห็นได้ว่าเป็นไปตามสมมติฐานทุกข้อ ทั้งนี้อาจเนื่องมาจากผู้วิจัยมีการควบคุมคุณภาพของการพารามิเตอร์ข้อสอบทุกข้อให้เป็นไปตามคุณลักษณะของพารามิเตอร์ข้อสอบที่มีคุณภาพ นั่นคือ ข้อสอบที่ตรวจให้คะแนนสองค่าตามโมเดลโลจีสติก 3 พารามิเตอร์ มีค่าพารามิเตอร์อำนาจจำแนกมากกว่า 0.5 พารามิเตอร์ความยากมีค่าระหว่าง – 2.5 ถึง +2.5 และพารามิเตอร์การเดามีค่าต่ำกว่า 0.3 และพารามิเตอร์ของข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่ามีโค้งลำดับขั้นการตอบที่เหมาะสม นอกจากนี้ยังมีการวิเคราะห์ตรวจสอบความถูกต้องของพารามิเตอร์ที่ต้องใช้สำหรับการวิเคราะห์เปรียบเทียบคุณภาพการปรับเทียบคะแนนโดยการเปรียบระหว่างพารามิเตอร์ที่ใช้สำหรับวิเคราะห์ปรับเทียบ (วิเคราะห์จากโปรแกรม MULTILOG) กับพารามิเตอร์จากการจำลองข้อมูล (พารามิเตอร์จากโปรแกรม Wingen) ซึ่งผลแสดงให้เห็นถึงความไม่แตกต่างระหว่างพารามิเตอร์ข้อสอบจากทั้ง 2 โปรแกรม รวมทั้งตรวจสอบความเป็นเอกมิติของ

แบบสอบที่ผสมโมเดลการตอบระหว่างโมเดลการตอบที่ตรวจให้คะแนนแบบ 2 ค่า และโมเดลการตอบที่ตรวจให้คะแนนมากกว่าสองค่า หากผลการตรวจสอบพบว่าข้อมูลขาดความเป็นเอกมิติผู้วิจัยจะจำลองข้อมูลชุดใหม่เพื่อทดแทน จากการควบคุมคุณภาพของข้อมูลทุกขั้นตอน จึงเป็นผลทำให้คุณภาพการปรับเทียบคะแนนทุกเงื่อนไขมีความคลาดเคลื่อนต่ำ (MSE < 0.3)

2. การอภิปรายผลข้อมูลที่ได้จากการจำลองข้อมูล

2.1 การสร้างค่าพารามิเตอร์ที่แท้จริง และรูปแบบการตอบข้อสอบ

สำหรับขั้นตอนในการจำลองข้อมูลพารามิเตอร์ที่แท้จริง ผู้วิจัยจำลองข้อมูลพารามิเตอร์ ข้อสอบ a, b, และ c สำหรับโมเดลการตอบข้อสอบที่มีการตรวจให้คะแนนสองค่าด้วยโมเดล 3PL และพารามิเตอร์ข้อสอบ $\alpha$ และ $\delta$ สำหรับโมเดลที่มีการตอบข้อสอบที่มีการตรวจให้คะแนนแบบหลายค่าด้วยโมเดล GPCM ในการจำลองพารามิเตอร์ข้อสอบแต่ละโมเดล จำเป็นต้องกำหนดค่าการแจกแจงตามรูปแบบที่เหมาะสมกับค่าพารามิเตอร์แต่ละค่า โดยในการจำลองข้อมูลครั้งนี้ ผู้วิจัยได้กำหนดรูปแบบการแจกแจงค่าพารามิเตอร์แต่ละค่าให้สอดคล้องกับ Kim (2004) และ Kim และ Lee (2006) เนื่องจากเป็นการศึกษาการจำลองข้อมูลกับแบบสอบรูปแบบผสมเช่นเดียวกัน โดยพารามิเตอร์ข้อสอบที่ตรวจให้คะแนน 2 ค่า มีการกำหนดให้ค่าพารามิเตอร์อำนาจจำแนก (a) มีการแจกแจงแบบ Lognormal มีค่าเฉลี่ยเป็น 0 และความแปรปรวนเป็น 0.2 ซึ่งการกำหนดลักษณะการแจกแจงตามค่าเฉลี่ยและความแปรปรวนที่กำหนดในรูปแบบนี้ จะทำให้ได้ค่าพารามิเตอร์อำนาจจำแนกมีการแจกแจงแบบเบ้ขวา และข้อสอบมีอำนาจจำแนกอยู่ในเกณฑ์ที่มีคุณภาพระหว่าง 0.5 ถึง 2.5 ส่วนพารามิเตอร์ความยากของข้อสอบ (b) มีการแจกแจงแบบ Normal ที่มีค่าเฉลี่ยเป็น 0 และส่วนเบี่ยงเบนมาตรฐานเป็น 1 ซึ่งเป็นการแจงแจงตามลักษณะของโค้งปกติมาตรฐาน ผลการจำลองข้อมูลมีค่าพารามิเตอร์ความยากระหว่าง +3 ถึง -3 ส่วนพารามิเตอร์การเดาของข้อสอบ (c) เป็นการแจกแจงแบบ BETA มีค่า $\alpha$ เท่ากับ 8 ค่า $\beta$ เท่ากับ 32 การแจกแจงลักษณะ BETA นี้จะทำให้ข้อมูลเกิดความเบ้ และค่า $\alpha$ และ $\beta$ ที่กำหนดจะทำให้การแจกแจงมีลักษณะเบ้ขวา ทำให้พารามิเตอร์การเดาที่ได้มีต่ำ ส่วนการจำลองข้อมูลพารามิเตอร์ $\alpha$ และ $\delta$ สำหรับโมเดล GPCM มีรูปแบบการแจกแจงเหมือนกับค่าพารามิเตอร์อำนาจจำแนกและพารามิเตอร์ความยากในโมเดล 3PL ตามลำดับ Kim (2004) และ Kim และ Lee (2006) กล่าวว่าการกำหนดการแจกแจงพารามิเตอร์ดังกล่าว จะทำให้ได้พารามิเตอร์ข้อสอบที่มีค่าใกล้เคียงกับความเป็นจริงมากที่สุด

จากลักษณะการแจกแจงพารามิเตอร์ข้างต้น ผู้วิจัยพบว่า ค่าพารามิเตอร์การเดาของข้อสอบบางข้อที่ได้จากการจำลองข้อมูลมีค่าสูงกว่า 0.3 ซึ่งสูงกว่าเกณฑ์ของข้อสอบที่มีคุณภาพ (ข้อสอบที่มีคุณภาพ ค่าพารามิเตอร์การเดาควรต่ำกว่า 0.3) ผู้วิจัยทำการปรับการแจกแจงพารามิเตอร์การเดาแบบ BETA ให้มีค่า $\alpha$ เป็น 4 และค่า $\delta$ เป็น 32 จึงทำให้ค่าพารามิเตอร์การเดามีค่าต่ำกว่า 0.3 ในการจำลองข้อมูลครั้งนี้ผู้วิจัยได้กำหนดค่าพารามิเตอร์ที่แท้จริงด้วยการแจกแจงดังกล่าวข้างต้น ด้วยโปรแกรม WINGEN ที่พัฒนาโดย Han (2007) เพื่อให้ได้รูปแบบการตอบข้อสอบของผู้สอบที่ทำแบบทดสอบแต่ละฉบับ (Test A, Test B1 และ Test B2) ลักษณะของโปรแกรม Wingen มีการกำหนดค่าตั้งต้นของ Scaling factor มีค่า D = 1.0 และผู้ใช้สามารถเลือก Scaling factor เป็นค่า D = 1.7 ได้ ในการจำลองข้อมูลครั้งนี้ ผู้วิจัยเลือก Scaling factor เป็นค่า D = 1.7 เนื่องจากผู้วิจัยพบว่า หากเลือก Scaling factor เป็นค่า D = 1.0 ตามที่โปรแกรมกำหนด (default) ผลการจำลองข้อมูลจะได้รูปแบบการตอบข้อสอบ เมื่อนำไปวิเคราะห์องค์ประกอบแล้ว ข้อมูลที่ได้จะมีองค์ประกอบจำนวนมาก นั่นแสดงถึงข้อมูลไม่เป็นเอกมิติ ซึ่งไม่เป็นไปตามข้อตกลงเบื้องต้น แต่เมื่อผู้วิจัยใช้การจำลองข้อมูลด้วยการกำหนด Scaling factor เป็นค่า D = 1.7 จะทำให้ข้อมูลมีความเป็นเอกมิติและเป็นไปตามข้อตกลงเบื้องต้นตามแนวทฤษฎีการตอบสนองข้อสอบ โดยการพิจารณาความเป็นเอกมิติของข้อมูล ผู้วิจัยได้พิจารณาจากร้อยละของความแปรปรวนของตัวประกอบแรกพิจารณาแล้วเห็นว่ามีค่ามากพอที่จะสรุปว่ามีความเป็นเอกมิติหรือไม่ โดยผู้วิจัยใช้เกณฑ์ของ Rackase (อ้างใน Raju, 1993) เสนอแนะว่าการใช้ทฤษฎีการตอบสนองข้อสอบวิเคราะห์ข้อสอบนั้น ค่าความแปรปรวนจากตัวประกอบหลักตัวแรกของแบบสอบควรมีค่าไม่ต่ำกว่าร้อยละ 20.00 และ/หรืออัตราส่วนระหว่างตัวประกอบที่ 1 และตัวประกอบที่ 2 ตามเกณฑ์ของวรนุช แหยมแสง (2536) เสนอแนะว่าอัตราส่วนระหว่างตัวประกอบที่ 1 และตัวประกอบที่ 2 ไม่ควรต่ำกว่า 5.739 สามารถสรุปได้ว่า ข้อมูลมีความเป็นเอกมิติเป็นไปตามข้อตกลงของทฤษฎีการตอบสนองข้อสอบ ซึ่งจากการที่ผู้วิจัยทำการจำลองข้อมูลรูปแบบการตอบโดยการกำหนด Scaling factor เป็นค่า D = 1.7 ทำให้ข้อมูลเป็นไปตามเกณฑ์ดังกล่าวข้างต้น แสดงว่าข้อมูลที่นำมาศึกษาครั้งนี้มีความเป็นเอกมิติและเป็นไปตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

2.2 การประมาณค่าพารามิเตอร์ด้วยโปรแกรม MULTILOG

ในการวิจัยครั้งนี้ผู้วิจัยใช้โปรแกรม MULTILOG ในการประมาณค่าพารามิเตอร์ข้อสอบด้วยการประมาณค่าพารามิเตอร์ข้อสอบทั้ง 2 โมเดล (3PL และ GPCM) พร้อมกัน และใช้

ในการปรับเทียบคะแนนด้วยการปรับค่าพารามิเตอร์พร้อมกัน โปรแกรม MULTILOG สามารถประมาณค่าพารามิเตอร์ข้อสอบได้ทั้งข้อสอบที่มีการตรวจให้คะแนนแบบสองค่า และข้อสอบที่มีการตรวจให้คะแนนแบบหลายค่า ผู้ใช้สามารถเลือกวิเคราะห์ทีละโมเดล หรือวิเคราะห์พร้อมกันหลายโมเดลก็ได้ หลักการประมาณค่าใช้หลักการ marginal maximum likelihood (MML) ในการประมาณค่าพารามิเตอร์ข้อสอบ และหลักการประมาณค่า Maximum A Posteriori (MAP) ใช้สำหรับประมาณค่าพารามิเตอร์ความสามารถผู้สอบ รูปแบบการนำเข้าข้อมูลสามารถทำได้หลายลักษณะ เช่น การนำเข้าโดยใช้ข้อมูลดิบผลการตอบข้อสอบรายข้อของผู้สอบแต่ละคน หรือการนำเข้าโดยใช้แบบแผนการตอบข้อสอบ เป็นต้น ในการวิจัยครั้งนี้ ข้อมูลรูปแบบการตอบข้อสอบที่จำลองได้จากโปรแกรม WINGEN มีลักษณะเป็นผลการตอบข้อสอบรายข้อของผู้สอบแต่ละคน ผู้วิจัยจึงนำเข้าข้อมูลแบบผลการตอบข้อสอบรายข้อของผู้สอบแต่ละคน และทำการประมาณค่าพารามิเตอร์ข้อสอบโมเดล 3PL และ GPCM พร้อมกัน (simultaneous calibration) เนื่องจากลักษณะของแบบสอบรูปแบบผสมจะมีส่วนประกอบของข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนหลายค่าในแบบสอบ หากผู้วิจัยประมาณค่าพารามิเตอร์ข้อสอบแยกโมเดล จะทำให้ไม่สามารถประมาณค่าความเที่ยงของแบบสอบทั้งฉบับได้

ส่วนการปรับเทียบด้วยการปรับค่าพารามิเตอร์พร้อมกันนั้น โปรแกรม MULTILOG สามารถดำเนินการได้ และการวิจัยครั้งนี้ได้จำลองข้อมูลเพื่อให้ได้รูปแบบการตอบข้อสอบสำหรับกลุ่มผู้สอบที่ต้องทำแบบสอบฐาน (Test A 1) แบบสอบปรับเทียบสำหรับกลุ่มเท่าเทียม (Test B1) และ แบบสอบปรับเทียบสำหรับกลุ่มไม่เท่าเทียม (Test B2) การปรับเทียบโดยการปรับค่าพารามิเตอร์พร้อมกันทำโดยการวางชุดของพารามิเตอร์ข้อสอบทั้ง 2 ชุด บนสเกลร่วมกัน ขั้นตอนแรกในการดำเนินการคือ รวมชุดของรูปแบบการตอบจากกลุ่มผู้สอบทั้ง 2 กลุ่มเข้าด้วยกันแล้วกำหนดให้ข้อสอบที่อีกกลุ่มหนึ่งไม่ได้ทำเป็น not reach หรือ missing ในขณะที่การระบุสมาชิกของกลุ่มเป็น 1 และ 2 ไว้ในคอลัมภ์แรก ต่อมาเป็นขั้นการเลือกสเกล (pre –select a scale) ก่อนทำการรันโปรแกรมเพื่อทำการประมวลผล (สเกลความสามารถของกลุ่มผู้สอบกลุ่ม 1) โดยกำหนดข้อสอบร่วมเป็นข้อสอบหลัก หลังจากนั้นทำการประมวลผลการวิเคราะห์โมเดลเอกมิติแบบ 3 พารามิเตอร์ และโมเดล GPCM เพื่อประมาณค่าพารามิเตอร์โมเดล 3PL และ GPCM ไปพร้อมๆ กัน ในการวิเคราะห์ (run) โปรแกรมเพียงครั้งเดียว ซึ่งเป็นส่วนที่การันตีได้ว่าพารามิเตอร์ที่ประมาณค่าทั้งหมดอยู่บนสเกลเดียวกัน งประเด็นที่สำคัญในการปรับค่าพารามิเตอร์พร้อมกันด้วยโปรแกรม MULTILOG คือ โดยทั่วไปจะไม่ทราบค่าส่วนเบี่ยงเบนมาตรฐาน ยกเว้นในกรณีพิเศษ เช่น การศึกษาจากการจำลองข้อมูล เพื่อแก้ปัญหาดังกล่าว การศึกษานี้จึงใช้พารามิเตอร์ของการ

แจกแจงปกติที่เลือกสรรแล้วสำหรับประมาณค่ากลุ่มผู้สอบกลุ่มใหม่ (generating new groups' examinees) เพื่อระบุค่าส่วนเบี่ยงเบนมาตรฐาน นั่นคือ การศึกษาครั้งนี้ได้เตรียมค่าความชันที่แท้จริงของการแปลงคะแนนเชิงเส้น สำหรับกระบวนการปรับค่าพารามิเตอร์พร้อมกัน (concurrent calibration) แทนที่การประมาณค่าพารามิเตอร์เหล่านั้นโดยใช้กระบวนการเชื่อมโยง ดังนั้น กระบวนการปรับค่าพารามิเตอร์พร้อมกันจึงอาจจะน่าพอใจกว่าวิธีอื่น และด้วยเหตุนี้อาจจะเป็นผลให้เกิดข้อค้นพบหลักของการวิจัยครั้งนี้ได้ เมื่อโปรแกรม MULTILOG ใช้สำหรับการประมาณค่า (calibration) แบบสอบ ลักษณะของการปรับค่าพารามิเตอร์พร้อมกันที่เด่นกว่าการปรับค่าพารามิเตอร์แยกกัน ดังที่ได้รายงานจากการวิจัยก่อนหน้านี้ เช่น Kim และ Cohen (2002) และ Hanson และ Beguin (2002) อย่างไรก็ตาม สิ่งที่เกี่ยวข้องกับลักษณะเด่นของการปรับค่าพารามิเตอร์พร้อมกันนั้น การศึกษาที่ผ่านมายังไม่มีประเด็นที่ชัดเจนเกี่ยวกับการระบุส่วนเบี่ยงเบนมาตรฐานในการอภิปราย จึงพอจะสันนิษฐานได้ว่าที่ผ่านมามีการกำหนดค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 ซึ่งกำหนดโดยค่าตั้งต้นของโปรแกรม MULTILOG การแก้ปัญหาของสเกลที่ไม่ชัดเจนในการกำหนดโมเดล IRT ทำโดยการกำหนดค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของการแจกแจงความสามารถสำหรับกลุ่มอ้างอิง (กลุ่มเก่า) โดยกำหนดให้ค่าเฉลี่ยเป็น 0 และส่วนเบี่ยงเบนมาตรฐานเป็น 1 ถ้ามีกลุ่มผู้สอบเพียงกลุ่มเดียวที่นำมาใช้ในการประมาณค่า การแก้สมการก็จะยอมรับค่าปกติที่โปรแกรมใช้ในการประมาณค่า แต่ถ้ามีกลุ่มผู้สอบสองกลุ่ม หรือมากกว่านั้น โปรแกรม MULTILOG ก็จะใช้ค่าเฉลี่ยของการแจกแจงความสามารถสำหรับกลุ่มที่ไม่ใช่กลุ่มอ้างอิง (กลุ่มใหม่) แต่ไม่ใช่ค่าส่วนเบี่ยงเบนมาตรฐาน ซึ่งค่าส่วนเบี่ยงเบนมาตรฐานต้องระบุ ในทางตรงกันข้าม พารามิเตอร์ข้อสอบและพารามิเตอร์ความสามารถที่เกี่ยวข้องกับกลุ่มที่ไม่ใช่กลุ่มอ้างอิงอาจจะไม่ถูกวางบนสเกลร่วมของกลุ่มอ้างอิงโดยปราศจากการระบุค่าส่วนเบี่ยงเบนมาตรฐานที่เหมาะสมสำหรับแต่ละกลุ่ม

2.3 เกณฑ์ที่ใช้ในการประเมินผลการปรับเทียบคะแนน

เกณฑ์ที่ใช้ในการประเมินคุณภาพการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมครั้งนี้ พิจารณาจากดัชนีค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (Mean square error : MSE) ซึ่งได้มาจากเกณฑ์โค้งลำดับขั้นการตอบ (category response curve : CRC) เป็นค่าที่แสดงถึงความแตกต่างระหว่างโค้งคุณลักษณะลำดับขั้นการตอบที่แท้จริงกับโค้งคุณลักษณะลำดับขั้นการตอบที่ประมาณค่าได้

ในการประเมินประสิทธิภาพของกระบวนการเชื่อมโยงหลาย ๆ วิธี จำเป็นต้องมีการกำหนดเกณฑ์การประเมินที่เหมาะสมกับองค์ประกอบที่ศึกษาในการจำลองข้อมูล ส่วนใหญ่จะมีการสร้างเกณฑ์การประเมินสำหรับแต่ละการศึกษาโดยการที่รู้พารามิเตอร์ หรือการอนุมานผลลัพธ์จากพารามิเตอร์ ในการศึกษาจำลองข้อมูลครั้งนี้ ได้นิยามเกณฑ์การประเมินโดยพิจารณาจากปริมาณความแตกต่างระหว่างค่าพารามิเตอร์ที่แท้จริงและค่าพารามิเตอร์ที่ประมาณค่าได้ ซึ่งเป็นเกณฑ์ที่มีการใช้ในการศึกษาจำลองข้อมูลเพื่อเปรียบเทียบผลการปรับเทียบคะแนนในการศึกษาวิจัยก่อนหน้านี้ (Hanson, Beguin, and Glas, 2000; Hanson and Beguin, 2002; Kim and Lee, 2004) เกณฑ์การประเมินครั้งนี้สามารถใช้เพื่อเป็นการยืนยันผลจากการศึกษาก่อนหน้านี้หรือได้สารสนเทศที่เป็นประโยชน์มากขึ้น ซึ่งการใช้เกณฑ์ที่แตกต่างกันออกไปยังมีความจำเป็นสำหรับการประเมินความตรงภายในและความตรงภายนอกของการศึกษาในการจำลองข้อมูล สำหรับเกณฑ์ที่ใช้ในการวิจัยครั้งนี้ ผู้วิจัยพิจารณาจากเกณฑ์ที่ใช้ในการศึกษาวิจัยในทำนองเดียวกันกับการวิจัยที่ผ่านมา นั่นคือ ในการศึกษาครั้งนี้เป็นการจำลองข้อมูลและมีการทำซ้ำ 100 ครั้ง สำหรับการเปรียบเทียบคุณภาพของการปรับเทียบคะแนน พิจารณาจากเกณฑ์การประเมิน 2 ชนิด ซึ่งเป็นการสรุปทางสถิติที่ครอบคลุมการทำซ้ำ 100 ครั้ง นั่นคือ เกณฑ์ค่าคงที่การแปลงสเกล (Scale Transformation constants : STC) และเกณฑ์โค้งคุณลักษณะลำดับขั้นการตอบ (Category response curve : CRC) โดยที่เกณฑ์ STC มีพื้นฐานมาจากความแตกต่างระหว่างค่าคงที่ที่ประมาณค่าได้ กับค่าคงที่ที่เป็นค่าที่แท้จริง ซึ่งประยุกต์ใช้กับการปรับเทียบตามแนวทฤษฎีการตอบสนองข้อสอบ 4 วิธีที่ใช้การปรับค่าพารามิเตอร์แยกกันเท่านั้นคือ วิธี moment 2 วิธี และวิธีโค้งคุณลักษณะ 2 วิธี ส่วนเกณฑ์ CRC มีพื้นฐานมาจากความแตกต่างระหว่างโค้งลำดับขั้นการตอบที่ประมาณค่าได้ กับโค้งลำดับขั้นการตอบที่เป็นค่าที่แท้จริง ทั้งเกณฑ์ STC และ CRC ที่นำมาใช้ในการเปรียบเทียบคุณภาพของการปรับเทียบนี้ กล่าวได้ว่าให้สารสนเทศที่เป็นประโยชน์และก่อให้เกิดความเข้าใจเป็นอย่างมากเกี่ยวกับประสิทธิภาพของกระบวนการเชื่อมโยงคะแนน แต่เกณฑ์ CRC มีลักษณะที่เด่นกว่าเกณฑ์ STC เนื่องจากสามารถพิจารณาทั้งค่าความชันและจุดตัดไปพร้อมๆ กันได้ในกระบวนการประมาณค่า (recovering) โค้งลำดับขั้นการตอบ เกณฑ์ทั้ง 2 ชนิดที่ใช้นี้เป็นค่าของ MSE สำหรับประมาณค่าจากการทำซ้ำ 100 ครั้ง ผู้วิจัยศึกษาผลการปรับเทียบคุณภาพของการปรับเทียบโดยพิจารณาจากเกณฑ์ CRC อันเนื่องมาจากลักษณะที่เด่นกว่าเกณฑ์ STC ดังกล่าวข้างต้น

ในแต่ละเงื่อนไขของการจำลองข้อมูล จะมีค่าประมาณพารามิเตอร์ข้อสอบ 100 ชุดสำหรับแบบสอบแต่ละฟอร์ม ทั้งเก่าและใหม่ เนื่องจากมีการทำซ้ำ 100 ครั้งของการจัดกระทำ

ข้อมูล จึงคาดหวังว่าแต่ละเงื่อนไขการจำลองข้อมูลพารามิเตอร์ข้อสอบสำหรับกลุ่มใหม่ควรจะอยู่บนสเกลเดียวกันเหมือนกับว่าเป็นประชากรของค่าพารามิเตอร์จากกลุ่มเดียวกัน หลังจากการแปลงค่าประมาณพารามิเตอร์ของฟอร์มใหม่ ในอดีตที่ผ่านมามีการใช้เกณฑ์การประเมิน 2 กลุ่ม เพื่อประเมินความแม่นยำของกระบวนการปรับเทียบโดยการปรับค่าพารามิเตอร์แยกกันทั้ง 4 วิธี ในการเปรียบเทียบกับวิธีการปรับค่าพารามิเตอร์พร้อมกัน (concurrent calibration) ซึ่งอยู่ภายใต้สถานการณ์ในอุดมติที่โมเดล IRT เป็นไปตามข้อตกลงเบื้องต้นทุกประการ ซึ่งเกณฑ์การประเมินที่ใช้คือ เกณฑ์ค่าคงที่ของการแปลงสเกล (scale transformation constants : STC) และเกณฑ์โค้งคุณลักษณะการตอบ (category response curve : CRC) ซึ่งเกณฑ์ STC มีพื้นฐานมาจากความแตกต่างระหว่างค่าคงที่ของการแปลงคะแนนที่ประมาณค่าได้กับค่าคงที่ของการแปลงคะแนนจริง (estimated and true transformation constant) ดังนั้น เกณฑ์ STC จึงไม่ประยุกต์ไปสู่วิธีการปรับค่าพารามิเตอร์พร้อมกัน ส่วนเกณฑ์ CRC มีพื้นฐานมาจากความแตกต่างระหว่างโค้งคุณลักษณะที่ประมาณค่าได้ กับโค้งคุณลักษณะจริง (estimated and true category response curve) สำหรับเกณฑ์ทั้ง STC และ CRC มีเพียงข้อสอบของฟอร์มใหม่เท่านั้นที่ถูกนำมาคิด โดยประเมินประสิทธิภาพของการปรับเทียบโดยการปรับค่าพารามิเตอร์แยกกันทั้ง 4 วิธี สามารถประเมินโดยการตรวจสอบค่าประมาณพารามิเตอร์ข้อสอบของฟอร์มใหม่วางอย่างเหมาะสมบนสเกลเก่าหรือไม่ อย่างไร ถ้าค่าประมาณพารามิเตอร์ข้อสอบของฟอร์มใหม่วางอย่างเหมาะสมบนสเกลเก่าแล้วดังนั้นก็จะเป็นฟังก์ชันการตอบข้อสอบของฟอร์มใหม่ด้วย

## ข้อเสนอแนะ

### 1. ข้อเสนอแนะในการนำไปใช้

การจำลองข้อมูลเพื่อการศึกษาครั้งนี้ เป็นการกำหนดข้อมูลให้สอดคล้องกับโมเดลการตอบข้อสอบแบบโลจีสติก 3 พารามิเตอร์ และโมเดล GPC ที่มีพารามิเตอร์ข้อสอบที่มีคุณภาพปรับเทียบคะแนนโดยใช้กลุ่มผู้สอบกลุ่มละ 1000 คน ผลการศึกษาเป็นองค์ความรู้ที่ยืนยันคุณภาพของการปรับเทียบคะแนนวิธีโค้งคุณลักษณะ และการปรับค่าพารามิเตอร์พร้อมกันเมื่อมีเงื่อนไขที่ต้องการปรับเทียบแตกต่างกันตามสถานการณ์การทดสอบ ซึ่งในสถานการณ์การทดสอบที่เกิดขึ้นจริง ข้อมูลสำหรับการปรับเทียบอาจไม่สอดคล้องกับโมเดลการตอบข้อสอบที่ต้องการและ/หรือ อาจมีการฝ่าฝืนข้อตกลงเบื้องต้นในการปรับเทียบ เช่น ความเป็นเอกมิติของข้อมูล ความคู่ขนานของ

แบบสอบ 2 ฉบับ ความไม่เป็นตัวแทนของข้อสอบร่วม เป็นต้น ดังนั้น ผลจากการวิจัยจึงเป็นแนวทางในการพิจารณาเลือกใช้เงื่อนไขการศึกษาที่สอดคล้องกับสถานการณ์การทดสอบมากที่สุด

ข้อค้นพบหลักสำหรับการวิจัยครั้งนี้คือ กระบวนการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม ในภาพรวมวิธีการปรับค่าพารามิเตอร์พร้อมกันมีประสิทธิภาพในการปรับเทียบสูงกว่าการปรับเทียบด้วยวิธีโค้งคุณลักษณะ ที่ใช้กับการออกแบบการเก็บรวบรวมข้อมูลทั้งกับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน และกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกัน โดยมีประเด็นในการพิจารณาเพื่อนำไปใช้ ดังนี้

1.1 เนื่องจากการศึกษาครั้งนี้เป็นการจำลองข้อมูล ดังนั้นพารามิเตอร์ข้อสอบจึงเป็นไปตามทฤษฎี สอดคล้องกับโมเดลการตอบข้อสอบและมีความเป็นเอกมิติ นั่นคือ เป็นไปตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบทุกประการ ดังนั้น ผลการปรับเทียบที่เกิดขึ้น จึงเป็นผลในเชิงทฤษฎี หากมีการประยุกต์ใช้ในการทดสอบสถานการณ์จริง ต้องมีการพิจารณาค่าพารามิเตอร์ของข้อสอบ และมีการตรวจสอบข้อมูลว่าเป็นไปตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบหรือไม่ แล้วจึงพิจารณาเลือกใช้วิธีการปรับเทียบ และการเก็บรวบรวมข้อมูล ที่เหมาะสมกับสถานการณ์ การทดสอบนั้นๆ

1.2 การนำวิธีการปรับเทียบไปใช้ ให้พิจารณาจากลักษณะการออกแบบการเก็บรวบรวมข้อมูล นั่นคือ ถ้าการบริหารจัดการทดสอบนั้นๆ กลุ่มผู้สอบสองกลุ่มมีความเท่าเทียมกัน สามารถใช้วิธีการปรับเทียบได้ทั้งวิธีการโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน แต่ถ้าการทดสอบนั้นกลุ่มผู้สอบสองกลุ่มมีความสามารถไม่เท่าเทียมกัน ควรใช้วิธีการปรับเทียบโดยการปรับค่าพารามิเตอร์พร้อมกันเนื่องจากเป็นการนำข้อมูลผลการตอบจากทั้งสองกลุ่มวิเคราะห์ข้อมูลด้วยการ run โปรแกรมคอมพิวเตอร์เพียงครั้งเดียว ซึ่งจะทำให้มั่นใจได้ว่าผลการวิเคราะห์อยู่บนสเกลเดียวกัน แต่มีข้อควรพิจารณาว่า หากการทดสอบมีการใช้แบบสอบ 2 ฉบับที่แตกต่างกัน จำเป็นต้องมีการกำหนดข้อสอบร่วมเพื่อเป็นชุดของข้อสอบที่ใช้ในการเชื่อมโยงสเกลระหว่างการปรับเทียบคะแนนจากแบบสอบทั้งสองฉบับ ในสถานการณ์การทดสอบที่เกิดขึ้นจริงบางครั้งในการทดสอบด้วยแบบสอบ 2 ฉบับที่แตกต่างกัน กับกลุ่มผู้สอบที่มีความสามารถใกล้เคียงกัน Cao (2008) เสนอแนะให้พิจารณาว่ากลุ่มผู้สอบเหล่านั้นเป็นกลุ่มไม่เท่าเทียม และควรออกแบบการปรับเทียบคะแนนให้มีข้อสอบร่วมด้วย

1.3 การปรับเทียบคะแนนจากแบบสอบ 2 ฉบับที่แตกต่างกัน ควรพิจารณากำหนดข้อสอบร่วมให้ประกอบไปด้วยข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนหลายค่า โดยที่จำนวนข้อสอบทั้ง 2 ชนิดควรรวมกันได้มากกว่าร้อยละ 20 ของข้อสอบในแบบสอบทั้ง

ฉบับ ทั้งนี้ จำนวนข้อสอบที่ตรวจให้คะแนนสองค่าในแบบสอบร่วมไม่ควรมีจำนวนมากเกินไป และสัดส่วนของคะแนนระหว่างข้อสอบที่ตรวจให้คะแนนสองค่ากับข้อสอบที่ตรวจให้คะแนนหลายค่าควรมีสัดส่วนใกล้เคียงกัน นอกจากนี้ในการกำหนดข้อสอบร่วมให้พิจารณาถึงความเป็นตัวแทนทั้งในด้านสถิติ และเนื้อหาของข้อสอบด้วย เนื่องจากมีงานวิจัยสนับสนุนว่าเนื้อหาหาของข้อสอบและค่าสถิติของข้อสอบมีผลต่อคุณภาพการปรับเทียบคะแนน โดยเฉพาะอย่างยิ่งถ้ากลุ่มผู้สอบสองกลุ่มมีความสามารถไม่เท่ากัน ความเป็นตัวแทนด้านสถิติจะมีอิทธิพลอย่างมากต่อคุณภาพการปรับเทียบคะแนน (Gao, Hanson and Harris, 1999)

1.4 การพิจารณาสัดส่วนของข้อสอบที่ตรวจให้คะแนนแบบสองค่ากับข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่าในแบบสอบ ในปัจจุบันมีการทดสอบจำนวนมากที่แบบสอบมีรูปแบบผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า และข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า สัดส่วนของข้อสอบทั้งสองชนิดในแบบสอบที่จะทำให้ผลการปรับเทียบคะแนนเกิดความคลาดเคลื่อนต่ำคือต้องเพิ่มจำนวนข้อสอบที่ตรวจให้คะแนนสองค่าในแบบสอบให้มากขึ้น

1.5 การปรับเทียบคะแนนที่เสนอในการวิจัยนี้สามารถนำไปประยุกต์ใช้กับแบบสอบที่ประกอบด้วยแบบสอบย่อย (Testlet) ได้ โดยให้พิจารณาแบบสอบย่อยนั้นในลักษณะของข้อสอบที่ตรวจให้คะแนนหลายค่า ซึ่งจะทำให้ได้ผลการปรับเทียบที่มีความประสิทธิภาพมากขึ้น (Lee, 2001) โดยจัดการกับข้อสอบที่ตรวจให้คะแนนสองค่าเป็นกลุ่มของแบบสอบย่อยหลาย ๆ ชุด และพิจารณาชุดของแบบสอบย่อยแต่ละชุดเป็นข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า

จากข้อเสนอแนะการนำไปใช้ที่กล่าวมาข้างต้น อยู่ภายใต้ข้อตกลงเบื้องต้นของข้อมูลที่สอดคล้องกับโมเดลการตอบข้อสอบแบบโลจีสติก 3 พารามิเตอร์ และโมเดล GPC ที่มีการตรวจให้คะแนน 6 ลำดับขั้น (0 – 5 คะแนน) โดยการปรับเทียบกับกลุ่มผู้สอบกลุ่มละ 1000 คน ดังนั้นในการพิจารณาเพื่อนำไปใช้ ควรคำนึงถึงข้อมูลที่สอดคล้องกับโมเดลการตอบข้อสอบที่ผสมระหว่างข้อสอบที่ตรวจให้คะแนนสองค่า (Dichotomous) และโมเดลที่ตรวจให้คะแนนมากกว่าสองค่า (Polytomous) โดยแบบสอบ 2 ฉบับที่ใช้ในการปรับเทียบมีความคู่ขนานกันและมีจำนวนข้อสอบเท่าๆ กันทั้ง 2 ฉบับ ข้อมูลมีความเป็นเอกมิติ ในทำนองเดียวกัน จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับคุณภาพของวิธีการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่มีการศึกษาในลักษณะอื่นๆ เช่น ข้อมูลเป็นพหุมิติ ข้อสอบร่วมขาดความเป็นตัวแทนทางสถิติ เป็นต้น ผลการวิจัยแสดงให้เห็นถึงความแกร่งของวิธีการปรับเทียบทั้งวิธีโค้งคุณลักษณะ และวิธีการปรับค่าพารามิเตอร์พร้อมกัน ซึ่งถือได้ว่าเป็นวิธีการปรับเทียบที่มีคุณภาพ เกิดความคลาดเคลื่อนในการปรับเทียบต่ำ แต่สิ่งที่ควรพิจารณานอกเหนือจากนี้คือ การกำหนดข้อสอบร่วมที่มีความเป็นตัวแทน

ของแบบสอบทั้งฉบับ สำหรับแบบสอบรูปแบบผสมควรกำหนดให้ข้อสอบทั้งชนิดตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่าเป็นข้อสอบร่วม เพื่อให้มีความเป็นตัวแทนของแบบสอบทั้งฉบับตาม สำหรับกลุ่มผู้สอบสองกลุ่มหากมีการทำแบบสอบต่างฉบับกัน ให้พิจารณาว่าผู้สอบสองกลุ่มมีการออกแบบการเก็บรวบรวมข้อมูลแบบกลุ่มไม่เท่าเทียมกัน (Cao, 2008)

## 2. ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

การศึกษาจำลองข้อมูลเพื่อศึกษาคุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมครั้งนี้ เป็นการผสมระหว่างโมเดลการตอบข้อสอบที่ตรวจให้คะแนนสองค่า (Dichotomous) และโมเดลการตอบข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า (Polytomous) ซึ่งเป็นการพิจารณาในมิติของการให้คะแนนในแต่ละโมเดลที่นำมาผสมในแบบสอบภายใต้รูปแบบการตอบข้อสอบที่ต่างกัน โดยมีความเชื่อว่าโมเดลการตอบข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่าเป็นลักษณะของข้อสอบเขียนตอบที่ต้องใช้ทักษะขั้นสูงในการสร้างคำตอบ (constructed response) มากกว่าข้อสอบที่ตรวจให้คะแนนแบบสอบค่าที่มักจะมีลักษณะเป็นข้อสอบแบบหลายตัวเลือก โดยธรรมชาติของการวัดทางการศึกษาแบบสอบฉบับเดียวกัน ควรมุ่งวัดในโครงสร้างเนื้อหาและ/หรือคุณลักษณะเดียวกันซึ่งแสดงถึงความเป็นเอกมิติของแบบสอบ แต่ในสถานการณ์การทดสอบในปัจจุบัน อาจไม่เป็นไปตามข้อตกลงเบื้องต้นตามทฤษฎีการตอบสนองข้อสอบ เช่น แบบสอบไม่เป็นเอกมิติ ข้อมูลไม่สอดคล้องกับโมเดล ข้อสอบร่วมไม่เป็นตัวแทน เป็นต้น จากลักษณะดังกล่าวข้างต้น แสดงให้เห็นว่าควรมีการศึกษาวิจัยเกี่ยวกับการปรับเทียบคะแนนในประเด็นอื่น ๆ ดังนี้

2.1 การศึกษาปรับเทียบแบบสอบรูปแบบผสมโมเดลการตอบข้อสอบที่ตรวจให้คะแนนสองค่า (Dichotomous) กับโมเดลการตอบข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า (Polytomous) เป็นการศึกษาในมิติของโมเดลการให้คะแนนที่ต่างกัน โดยไม่ได้พิจารณาถึงอิทธิพลของรูปแบบการตอบ (format effect) ดังนั้น จึงควรขยายองค์ความรู้ในการศึกษาปรับเทียบคะแนนสำหรับแบบสอบที่มีการให้คะแนนเหมือนกัน แต่รูปแบบการตอบต่างกัน เช่น ข้อสอบแบบหลายตัวเลือกที่มีการให้คะแนนแบบตอบถูกได้ 1 และตอบผิดได้ 0 กับแบบสอบชนิดตอบสั้นที่มีการให้คะแนนแบบตอบถูกได้ 1 และตอบผิดได้ 0 เช่นเดียวกัน ซึ่งข้อสอบชนิดเลือกตอบนั้นถึงแม้ผู้สอบจะไม่มีความรู้ในเรื่องนั้นๆ แต่ผู้สอบก็มีโอกาสในการเดาคำตอบจากตัวเลือกที่กำหนดให้ได้ ส่วนข้อสอบแบบตอบสั้น ผู้สอบจะไม่มีโอกาสในการเดาคำตอบ กระบวนการได้มาซึ่งคำตอบต้องผ่านกระบวนการคิดจนมั่นใจว่าคำตอบที่ต้องการนั้นถูกต้องที่สุด

ดังนั้นข้อสอบชนิดตอบสั้นจึงต้องใช้กระบวนการคิดในการหาคำตอบมากกว่าข้อสอบแบบเลือกตอบเนื่องจากไม่มีโอกาสในการเดา หรือศึกษากับรูปแบบของข้อสอบ 2 ชนิดที่ต่างกัน แต่มีการตรวจให้คะแนนแบบหลายค่าเหมือนกัน เช่น ข้อสอบแบบหลายตัวเลือกที่มีการให้คะแนนแบบหลายค่า กับแบบสอบชนิดเขียนตอบที่มีการให้คะแนนแบบหลายค่าเช่นเดียวกัน เป็นต้น เพื่อศึกษาอิทธิพลของรูปแบบการตอบในแบบสอบที่ส่งผลต่อคุณภาพการปรับเทียบคะแนน เนื่องจากแบบสอบที่ต่างชนิดกัน ย่อมมีอิทธิพลต่อกระบวนการคิดในการตอบข้อสอบต่างกันด้วย

2.2 การทดสอบในปัจจุบันมีการทดสอบต่างเวลาต่างแบบสอบ ดังเช่น การทดสอบวัดผลสัมฤทธิ์ของผู้เรียน ที่มีการจัดทดสอบในแต่ละปีการศึกษา ซึ่งในการทดสอบแต่ละปีการศึกษาอาจจะมีจำนวนข้อสอบและจำนวนคะแนนแตกต่างกันไป และในการศึกษาเกี่ยวกับการปรับเทียบส่วนใหญ่เป็นการปรับเทียบจากแบบสอบ 2 ฉบับที่มีจำนวนข้อสอบเท่ากัน ปรับค่าพารามิเตอร์ข้อสอบโดยใช้ข้อสอบร่วมเป็นตัวเชื่อมโยงระหว่างแบบสอบ 2 ฉบับ ดังนั้น ควรมีการศึกษาเกี่ยวกับวิธีการปรับเทียบคะแนนจากแบบสอบ 2 ฉบับ ที่มีจำนวนข้อแตกต่างกัน โดยศึกษาวิธีการที่มีคุณภาพในการปรับเทียบสูงนั่นคือเกิดความคลาดเคลื่อนในการปรับเทียบต่ำ ที่สามารถนำมาใช้กับการปรับเทียบคะแนนจากแบบสอบ 2 ฉบับที่มีจำนวนข้อสอบแตกต่างกัน เนื่องจากจำนวนข้อสอบที่ต่างกันจะมีความสัมพันธ์กับปริมาณโครงสร้างเนื้อหาที่ต่างกันด้วย รวมทั้งศึกษามิติข้อสอบร่วมสำหรับแบบสอบ 2 ฉบับที่ใช้การทดสอบแต่ละครั้งด้วย

2.3 การศึกษาปรับเทียบคะแนนที่ผ่านมาเป็นการศึกษาคุณภาพการปรับเทียบเมื่อการออกแบบการเก็บรวบรวมข้อมูลเป็นกลุ่มเท่าเทียม และกลุ่มไม่เท่าเทียม ดังนั้นจึงควรศึกษาผลการปรับเทียบคะแนนโดยมีการออกแบบการเก็บรวบรวมข้อมูลกับผู้สอบสองกลุ่มที่มีความสามารถไม่เท่าเทียมกันที่มีระดับความสามารถแตกต่างกันกันหลายๆ ระดับ เนื่องจากการศึกษาที่ผ่านมาในอดีต เป็นเพียงการศึกษากับกลุ่มผู้สอบที่มีความสามารถเท่าเทียมกัน และกลุ่มผู้สอบที่มีความสามารถไม่เท่าเทียมกันเพียง 1 ระดับเท่านั้น เพื่อให้ผลการศึกษาสามารถสรุปอ้างอิงไปสู่กลุ่มผู้สอบได้หลายระดับแตกต่างกันไป

2.4 ธรรมชาติของแบบสอบรูปแบบผสมระหว่างข้อสอบต่างชนิดกัน ย่อมแสดงถึงความสามารถที่ต่างกันที่ต้องใช้ในการตอบข้อสอบเหล่านั้น ดังนั้น ควรมีการศึกษาการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสมที่ถูกพิจารณาในลักษณะที่แบบสอบมีความเป็นพหุมิติ ซึ่งลักษณะของความเป็นพหุมิติแสดงให้เห็นถึงการฝ่าฝืนข้อตกลงเบื้องต้นตามทฤษฎีการตอบสนองข้อสอบ โดยกำหนดระดับความสัมพันธ์ของความเป็นพหุมิติหลายระดับแตกต่างกันไป เพื่อศึกษาผลของวิธีการปรับเทียบคะแนนแต่ละวิธีที่เหมาะสมกับสถานการณ์การทดสอบดังกล่าว

2.5 ศึกษาการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม กับสถานการณ์การทดสอบจริง เพื่อยืนยันองค์ความรู้ที่เกิดขึ้น เนื่องจากการวิจัยเกี่ยวกับการปรับเทียบที่ผ่านมาส่วนใหญ่เป็นการจำลองข้อมูลเพื่อให้สอดคล้องกับโมเดลที่ต้องการ และเป็นไปตามข้อตกลงเบื้องต้น หากสถานการณ์การทดสอบจริงไม่สอดคล้องกับโมเดลและไม่เป็นไปตามข้อตกลงเบื้องต้น วิธีการปรับเทียบทั้ง 2 วิธีข้างต้นจะยังคงมีความแกร่งหรือไม่



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

# รายการอ้างอิง

**ภาษาไทย**


บุญเชิด ภิญโญอนันตพงษ์. (2544). *การประเมินการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ: แนวคิดและวิธีการ*. กรุงเทพฯ: โอเดียนสโตร์.

พรพิมล นาคเวช. (2537). *การศึกษาคุณภาพของการเทียบคะแนนในแนวตั้งโดยใช้ทฤษฎีการตอบสนองข้อสอบ*. วิทยานิพนธ์ปริญญามหาบัณฑิต, สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา จุฬาลงกรณ์มหาวิทยาลัย.

พวงแก้ว ปุณยกนก. (2546). การประเมินพุทธิพิสัย. ใน สุวิมล ว่องวาณิช (บรรณาธิการ), *การประเมินผลการเรียนรู้แนวใหม่,* หน้า 167 – 194. พิมพ์ครั้งที่ 1. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ไพรัตน์ วงษ์นาม. (2533). *สัมประสิทธิ์การอ้างอิงสรุปของแบบสอบความเรียง.* วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต, สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

เยาวดี วิบูลย์ศรี. (2548). *การวัดผลและการสร้างแบบสอบผลสัมฤทธิ์.* พิมพ์ครั้งที่ 4. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2545). *ทฤษฎีการทดสอบแนวใหม่.* พิมพ์ครั้งที่ 2: กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2546). นโยบายการประเมินผลการเรียนรู้ตามพระราชบัญญัติการศึกษาแห่งชาติ. ใน สุวิมล ว่องวาณิช (บรรณาธิการ), *การประเมินผลการเรียนรู้แนวใหม่,* หน้า 3 – 24. พิมพ์ครั้งที่ 1. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2548). *ทฤษฎีการทดสอบแบบดั้งเดิม.* พิมพ์ครั้งที่ 5. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

สุนิสา จุ้ยม่วงศรี. (2546). *การศึกษาผลของการเทียบคะแนนข้อสอบที่ตรวจให้คะแนนหลายค่า.* วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต, สาขาวิชาการทดสอบและวัดผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

อดิศร ศรีบุญวงศ์. (2545). *การพัฒนาเกณฑ์ตัดสินคุณภาพการปรับเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ.* วิทยานิพนธ์ปริญญาคดุษฎีบัณฑิต, สาขาวิชาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

อรุณี เร้าอรุณ. (2536). *การเปรียบเทียบคะแนนของแบบสอบความเรียงที่ได้จากวิธีการตรวจต่างแบบ.* วิทยานิพนธ์ปริญญามหาบัณฑิต, สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

**ภาษาอังกฤษ**

Andrich,D. (1978). A Rating Formulation for Ordered Response Categories. *Psychometrika,* 43,4: 561-573.

Angoff,W.H. (1971). Scale,norm and equivalent score. InR.L.Throndike (Ed.), *Educational Measurement*. (2nd ed.), 508-600. Washing, D.C.: American Council on Education.

Angoff,W.H. (1984). *Scales, Norms, and Equivalent Scores*. 2nd ed. New Jersey: Princeton, Educational testing Service.

Baker, F.B. (1992). Equating tests under the grade response model. *Applied Psychological Measurement,* 16(1): 87 -96.

Baker, F.B. (1992). *Item Response Theory : Parameter Estimation Techniques*. New York: Marcel Dekker.

Baker, F.B. (1993). Equating tests under the nominal response model. *Applied Psychological Measurement,* 17(3): 239 - 251.

Baker, F.B. (1997). Empirical sampling distributions of Equating coefficients for grade and nominal response instruments . *Applied Psychological Measurement,* 21(2): 157 - 172.

Baker, F.B. and Al–Karni,A. (1991). A comparison of two Procedures for Computing IRT Equating Coefficients. *Journal of Education Measuremen,.* 28, 2: 147 – 162.

Baker,F.B. & Kim,S.H. (2004). *Item Response Theory: Parameter estimation Techniques* (2nd ed.) New York: Marcel Dekker.

Bastari,B. (2000). Linking Multiple – Choice and Constructed – Response Items to a common Proficiency Scale. Doctoral Dissertation, Educational Evaluation, Graduate School, Massachusetts Amherst University.

Beguin,A.A., Hanson, B.A., & Glas, C.A.W. (2000). Effect of multidimensionality on separate and concurrent estimation in IRT equating. *Paper presented at the annual meeting of the National Council on Measurement and Education*, New Orleans, LA. [Online]. 2000. Available from http://www.b-a-h.com/papers/paper0002.html [2005, October]

Brennan, R.L. (1987). Introduction to problems, perspectives, and practical issues in Equating. *Applied psychological measurement* 11, 3: 221-224.

Cao, Y. (2008). *Mixed-Format test equating: Effects of Test dimensionality and common item sets*. Doctoral Dissertation, Department of Measurement Statistics and Evaluation, Faculty of the Graduate School, University of Maryland.

Cohen, Allan S. & Kim, S.H. (1998). An investigation of linking methods under the grade response model. *Applied psychological measurement,* 22, 2: 116-130.

DeMars,C.E. (1998). Gender Differences in Mathematics and Science on a High School Proficiency Exam: The Role of Response Format. *Applied Measurement in Education,* 11,3: 279 – 299.

Donoghue, J.R. (1994). An Empirical Examination of The IRT Information of Polytomously Scored Reading Items Under the Generalized Partial Credit Model. *Journal of Educational Measurement,* 31,4: 295 – 311.

Dorans, N. J. (2000). Scaling and Equating. In H. Wainer (Ed.), *Computerized Adaptive Testing: A Primer (2nd)*. New Jersey: Lawrence.

Embretson, S.E. and Reise,S.P. (2000). *Item Response Theory for Psychologists.* New Jersey: Lawrence Erlbaum Associates.

Feuer, M. J., Holland, P. W., Green, B. F., Bertenthal, M. W., & Hemphill, F. C. (1999). *Uncommon Measures: Equivalence and Linkage among Educational Tests*. Washington, DC: National Academy Press.

Gulliksen,H. (1950). *Theory of mental test*. New York: Wiley.

Hambleton, R.K. (1996). Advances in Assessment Models, Methods, and Practices. *Handbook of Educational Psychology*, Edited by Berliner D.C. and Calfee R.C.P. 899 – 925. New York; Simon & Schuster Maemillan.

Hambleton,R.K, Swaminatan,H & Roger,H.J. (1991). *Fundamentals of Item Response Theory*. Newbury Park: C.A. SAGE Publications.

Hambleton,R.K, Swaminatan,H. (1985). *Item Response Theory: Principles and Application*. 2nd ed. Boston: Kluwer Nijhoff Publishing.

Hambleton,R.K. and Linda,L.Cook. (1977). Latent Trait Models and Their Use in the Analysis of Educational Test Data. *Journal of Educational Measurement,* 14, 2: 75-96.

Han, K. T. (2007). WinGen: Windows software that generates IRT parameters and item responses. *Applied Psychological Measurement,* 31(5), 457-459.

Han, K. T., & Hambleton, R. K. (2007). User's Manual: WinGen (*Center for Educational Assessment Report No. 642)*. Amherst, MA: University of Massachusetts, School of Education.

Han, K. T. (2009). IRTEQ: Windows application that implements IRT scaling and equating [computer program]. *Applied Psychological Measurement,* 33(6), 491-493.

Hanick, P.L., Huang, C-Y. (2002). Effects of decreasing the number of common items in equating link item sets*. Paper presented at the Annual Meeting of the American Educational Research Association*, New Orleans, LA. [Online]. 2002. Available from http://www.aera.net/papers.html [2005, October]

Hanson,B.A. & , Beguin,A.A. (2002). Ontaining a common scale for item response theory item parameters using sperate versus concurrent estimation in the common-item equating design. *Applied psychological measurement* 26: 3-24

Harris, D.J. (1991). Equating with non-representative common item sets and non-equivalent groups*. Paper presented at the Annual Meeting of the American Educational Research Association*, Chicago, IL.

Harris, D.J., & Course, J.D. (1993). A study of criteria used in equating. *Applied Measurement in Education,* 6: 195 -240.

Harwell, M., Stone, C.A., Hsu, T-C, & Kirisci, L. (1996). Monte Carlo studies in Item Response Theory. *Applied psychological measurement,* 20(2)*:* 101-125.

Hennings, Sara, S., Hirsch, T.M. & Zhang, L. (1996). A comparison of equating methods applied to performance-Based assessment. *Paper presented at the Annual Meeting of National Council on Measurement in Education*. (CD-ROM). New York City. P. 1 – 13. Available : ERIC (1992 – 2009/9) Acc. No.ED400295.

Holland,P.W. and Rubin,D.B.(Eds). (1982). *Test Equating*. New Jersey: Educational Testing service, Princeton Academic Press.

Hung, P., Wu, Y., & Chen, Y. (1991). IRT item parameter linking : Relevant issues for the purpose of item banking. *Paper presented at the International Academic Symposium on Psychological Measurement, Taiwan.*

Hyunh, H. & Ferrara S. (1994). A comparison of equal percentile and partial credit equatings for Performance-Based Assessments compose of Free-Response item. *Journal of Educational Measurement.* 31,2: 125 – 141.

Johnson, E.G. and Owen,E. (1998). *Linking the National Assessment of educational progress (NAEP) and the third international mathematics and science study (Timss):* A technical report. National center for education statistics, Research and Development Report. [Online]. 1998. Available from http://ETS.ORG. [2005, September]

Kennedy,P. and Walstad,W.B. (1997). Combining Multiple – Choice and Constructed Response Test Scores: An Economist's View. *Applied Measurement in Education* 10,4: 359 – 375.

Kim. (2004). *Unidimentional IRT scaling procedures for mixed format test and their robustness to multidimentionality*. Doctoral of Philosophy degree in education. Educational Measurement and Statistic. Graduate collage of the university of IOWA.

Kim. (2006). Robustness to Format Effects of IRT linking methods for mixed-format tests. *Applied Measurement in Education* 19,4: 357 – 3.81.

Kim,S.-H, Lee, W-C. (2004). IRT scale linking methods for mixed-format tests. *Paper presented at the annual meeting of the National Council on Measurement and Education*, San Diego, CA.

Kim,S.-H, Lee, W-C. (2006). An Extension of Four IRT Linking Methods for Mixed-Format Tests. *Journal of Educational Measurement*. 43(1). 53 – 70.

Kim,S.-H. & Kolen, M.J. (2006). Robussness to format effects of IRT linking methods for Mixed-Format Tests. *Journal of Educational Measurement* 19: 357 - 381.

Kim,S.-H. & Kolen, M.J. (2004). Optimally defining criterion functions for the characteristic curve procedures in the IRT scale linking. *Paper presented at the annual meeting of the American Educational Research Association*, San Diego, CA.

Kim, S., Walker, M.E., & McHale, F. (2007). Equating of Mixed-format tests in large scale assessments. *Paper Presented at the Annual Meeting of the American Educational Research Association (AERA) and The National Council on Measurement in Education (NCME).*

Kim,S.-H. & Cohen,A.S. (1992). Effects of linking methods on detection of DIF. *Journal of Educational Measurement* 29: 51 - 66.

Kim,S.-H. & Cohen,A.S. (1995). A minimal chi-square method for equating test under grade response model. *Applied Psychology Measurement* 19, 2: 167 -176.

Kim,S.-H. & Cohen,A.S. (1998). A comparison of linking and concurrent calibration under item response theory. *Applied psychological measurement,* 22(2)*:* 131-143.

Kim,S.-H. & Cohen,A.S. (2002). A comparison of linking and concurrent calibration under the graded response model. *Applied psychological measurement,* 26*:* 25-41.

Kim,S.-H. & Cohen,A.S. (2002). Robusness to format effects of IRT linking methods for mixed-format tests. *Applied Measurement in Education,* 19*,* 357-381.

Kolen, M.J., & Whitney, D.R. (1982). Comparison of four procedures for equating the tests of general educational development. *Journal of Educational Measurement* 19, 4*:* 279 – 293.

Kolen,M.J. and Brennan,R.L. (1995). *Test Equating: Method and Practices.* New York: Springer-Verlag.

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

Kubiszyn, T., and Borich, G. (2003). *Educational testing and measurement classroom application and practice*. 7th ed. New York: Wiley & Sons.

Lee, G. and others. (2001). Comparison of Dichotomous and Polytomous Item Response Models in Equating Scores From Test Composed of Testlets. *Applied Psychological Measurement* 25, 4: 357-372.

Li,Y.H., Lissitz, R.W. and Yang, Y.N. (1999), April. Estimatimg IRT Equating Coefficients for Tests with Polytomously and Dichotomously Scored Items. Paper Presented at the Annual Meeting of the National Council on Measurement in Education (NCME). (CD-ROM). p.1-32. *Available: ERIC (1992-2000/09)*; Acc.No. ED431800.

Linn, R.L. (1993). Linking results of distinct assessments. *Applied Measurement in Education* 16,1: 83 – 102.

Linn, R.L., Levine, M.V., Hastings, C.N., & Wardrop, J.L. (1981). Item bias in a test of reading comprehension. . *Applied Psychological Measurement* 5,24: 159-173.

Lord, F.M. (1980). *Application of Item Response Theory to Practical Testing Problems*. N.J.: Hillsdate, Erblum.

Lord, F.M. and Novick,M.R. (1968). *Statistical Theory of Mental Test Scores*. Massachusetts: Addison-Wesley Publishing Company Inc.

Lukhele,R., Thissen,D. and Wainner,H. (1994). On the Relative Value of Multiple – choice, Constructed-Response and Examinee-Selected Items on Two Achievement Tests. *Journal of Educational Measurement* 31,3: 234 – 250.

Macro, G.L. (1977). Item Characteristic Curve Solutions to Three Intractable Testing Problems. *Journal of Educational Measurement* 14,2: 139 – 160.

Master, G.N. (1982). A Rasch Model for Partial Credit Scoring. *Psychometrika* 47,2: 149 – 174.

Milliman, J., & Greene, J. (1989). *The specification and development of tests of achievement and ability.* In R.L. Linn(Ed), Educational Measurement (3rd ed.) Ney York: Macmillan.

Mislevy, R.J. (1992). *Linking educational assessment: Concepts issues, methods and prospects*. Princeton, NJ: ETS Policy Information Center.

Muraki, E. (1992). A Generalized Partial Credit Model: Application of an EM Algolithm. *Applied Psychological Measurement,* 16,2: 159 – 176.

Muraki, E. (1993). Information Functions of the Generalized Partial Credit Model. *Applied Psychological Measurement,* 17, 4: 351 – 363.

Muraki, E. (1998). *RESGEN: Item response generator*. Program manual published by Educational Testing Service, Princeton, NJ.

Muraki, E., Hombo,C.M. and Lee, Yong – Won. (2000). Equating and Linking of Performance Assessments. *Applied Psychological Measurement,* 24, 4: 325 – 337.

Muraki,E. (1990). Fitting a Polytomous Item Response Model to Likert-type Data. *Applied Psychological Measurement,* 14,1: 59-71.

Ogasawara,H. (2001). Least squares estimation of item response theory linking coefficients. *Applied Psychological Measuremen,t* 25: 373-383.

Ogasawara,H. (2002). Stable response functions with unstable item parameter estimates. *Applied Psychological Measurement,* 26,3: 239 – 254.

Pertersen, N.S., Cook,L.L. & Stocking,M.L. (1983). IRT versus conventional equating methods: a comparative study of scale stability. *Journal of educational statistic,* 8, 2: 137-156.

Petersen, N.S., Macro, G.L. and Stewart,E.E. (1982). A Test of the Adequacy of Linear Score Equting Models. *Test Equating*. Edited by Hollan, P.W. and Rubin,D.B. P.71-135. New Jersey: Academic Press.

Simon,A.B., Budescu,D.V., and Nevo, B. (1997). A Comparative Study of Measurement of Partial Knowledge in Multiple – Choice Tests. *Applied Psychological Measurement,* 21, 1: 65 – 88.

Stocking, M.L., & Lord, F.M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, 7: 201-210.

Tate, R. (1999). A Cautionary Note on IRT – Based Linking of Tests with Polytomous Item. *Journal of Educational Measurement,* 36, 4: 336 – 346.

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 Measuremen,* 37, 4: 329 – 346.

Thorndike, R.L., and Hagen, E.P. (1977). *Measurement and evaluation in psychology and education*. 4th ed. New York: John wiley & sons.

Vale,D. (1986). Linking Item Parameters onto a Common Scale. *Applied Psychological Measurement,* 10, 4*:* 333 – 344.

Walberge, H.J., and Haertel, G.D. (1992). *The International Encyclopedia of Education Evaluation (advanced in education).* Oxford: Pergamon.

Wiersma,W. and Jurs,S.G. (1990). *Educational Measurement and Testing*. 2nd ed. Needham Heights. Massachusetts: Simon and Schuster.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

# ภาคผนวก

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

# ภาคผนวก ก

## พารามิเตอร์ข้อสอบสำหรับการจำลองข้อมูลรูปแบบการตอบข้อสอบ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 41 ค่าพารามิเตอร์ของข้อสอบที่ใช้ในการจำลองข้อมูลผลการตอบข้อสอบสำหรับ **แบบสอบฉบับ** A

| Item | Model | พารามิเตอร์ข้อสอบ | | | | | |
|---|---|---|---|---|---|---|---|
| | | a/$\alpha$ | b/$\delta_{i1}$ | c/$\delta_{i2}$ | $\delta_{i3}$ | $\delta_{i4}$ | $\delta_{i5}$ |
| 1 | 3PLM | 1.275 | -2.014 | 0.149 | | | |
| 2 | 3PLM | 1.257 | -0.424 | 0.092 | | | |
| 3 | 3PLM | 1.023 | -0.471 | 0.182 | | | |
| 4 | 3PLM | 1.057 | 0.928 | 0.235 | | | |
| 5 | 3PLM | 0.836 | -0.616 | 0.095 | | | |
| 6 | 3PLM | 0.927 | -0.375 | 0.185 | | | |
| 7 | 3PLM | 1.367 | 1.374 | 0.137 | | | |
| 8 | 3PLM | 1.343 | -0.2 | 0.058 | | | |
| 9 | 3PLM | 0.984 | -0.023 | 0.044 | | | |
| 10 | 3PLM | 0.759 | -0.416 | 0.077 | | | |
| 11 | 3PLM | 1.627 | -0.07 | 0.219 | | | |
| 12 | 3PLM | 0.737 | -1.458 | 0.045 | | | |
| 13 | 3PLM | 1.296 | 1.363 | 0.089 | | | |
| 14 | 3PLM | 1.154 | 1.723 | 0.112 | | | |
| 15 | 3PLM | 0.993 | 0.411 | 0.146 | | | |
| 16 | 3PLM | 1.194 | -1.121 | 0.03 | | | |
| 17 | 3PLM | 1.017 | -1.273 | 0.237 | | | |
| 18 | 3PLM | 0.997 | 1.169 | 0.27 | | | |
| 19 | 3PLM | 1.156 | 1.023 | 0.147 | | | |
| 20 | 3PLM | 0.806 | 0.488 | 0.115 | | | |
| 21 | 3PLM | 1.186 | 0.297 | 0.148 | | | |
| 22 | 3PLM | 0.984 | -0.237 | 0.059 | | | |
| 23 | 3PLM | 0.882 | -1.213 | 0.079 | | | |
| 24 | 3PLM | 1.14 | -0.640 | 0.114 | | | |
| 25 | 3PLM | 0.784 | -0.296 | 0.091 | | | |
| 26 | 3PLM | 0.922 | -0.451 | 0.187 | | | |
| 27 | 3PLM | 0.705 | 0.014 | 0.082 | | | |
| 28 | 3PLM | 1.099 | 0.368 | 0.094 | | | |
| 29 | 3PLM | 1.376 | 0.519 | 0.108 | | | |
| 30 | 3PLM | 1.163 | -2.062 | 0.224 | | | |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 41 ค่าพารามิเตอร์ของข้อสอบที่ใช้ในการจำลองข้อมูลผลการตอบข้อสอบสำหรับ **แบบสอบฉบับ A** (ต่อ)

| Item | Model | พารามิเตอร์ข้อสอบ | | | | | |
|---|---|---|---|---|---|---|---|
| | | a | b/b1 | c/b2 | b3 | b4 | b5 |
| 31 | 3PLM | 1.045 | -0.083 | 0.080 | | | |
| 32 | 3PLM | 0.895 | -0.487 | 0.106 | | | |
| 33 | 3PLM | 0.904 | -0.343 | 0.196 | | | |
| 34 | 3PLM | 0.875 | 1.13 | 0.234 | | | |
| 35 | 3PLM | 1.052 | 0.75 | 0.058 | | | |
| 36 | 3PLM | 1.247 | -0.486 | 0.094 | | | |
| 37 | 3PLM | 0.927 | 0.082 | 0.239 | | | |
| 38 | 3PLM | 1.067 | -0.385 | 0.123 | | | |
| 39 | 3PLM | 0.840 | -0.700 | 0.170 | | | |
| 40 | 3PLM | 1.313 | 0.183 | 0.221 | | | |
| 41 | 3PLM | 0.991 | -0.369 | 0.095 | | | |
| 42 | 3PLM | 1.316 | -0.897 | 0.116 | | | |
| 43 | 3PLM | 1.132 | -0.296 | 0.219 | | | |
| 44 | 3PLM | 0.955 | 0.635 | 0.110 | | | |
| 45 | 3PLM | 1.178 | -0.108 | 0.079 | | | |
| 46 | 3PLM | 1.234 | 0.176 | 0.128 | | | |
| 47 | 3PLM | 0.908 | 1.007 | 0.067 | | | |
| 48 | 3PLM | 0.967 | 0.781 | 0.181 | | | |
| 49 | 3PLM | 1.162 | -0.999 | 0.06 | | | |
| 50 | 3PLM | 1.163 | -2.555 | 0.054 | | | |
| 51 | 3PLM | 0.833 | -0.227 | 0.079 | | | |
| 52 | 3PLM | 0.958 | 1.673 | 0.183 | | | |
| 53 | 3PLM | 1.152 | -1.216 | 0.081 | | | |
| 54 | 3PLM | 0.73 | 0.019 | 0.091 | | | |
| 55 | 3PLM | 0.953 | -0.826 | 0.243 | | | |
| 56 | 3PLM | 1.360 | 0.580 | 0.069 | | | |
| 57 | 3PLM | 1.123 | 0.670 | 0.161 | | | |
| 58 | 3PLM | 0.920 | 1.099 | 0.036 | | | |
| 59 | 3PLM | 1.110 | 1.868 | 0.06 | | | |
| 60 | 3PLM | 1.052 | -1.106 | 0.053 | | | |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 41 ค่าพารามิเตอร์ของข้อสอบที่ใช้ในการจำลองข้อมูลผลการตอบข้อสอบสำหรับ **แบบสอบฉบับ** A (ต่อ)

| | | พารามิเตอร์ข้อสอบ | | | | | |
|---|---|---|---|---|---|---|---|
| Item | Model | a | b/b1 | c/b2 | b3 | b4 | b5 |
| 61 | 3PLM | 1.035 | -1.541 | 0.108 | | | |
| 62 | 3PLM | 0.748 | -0.732 | 0.053 | | | |
| 63 | 3PLM | 0.811 | 0.141 | 0.127 | | | |
| 64 | 3PLM | 0.738 | 0.087 | 0.149 | | | |
| 65 | 3PLM | 1.012 | 0.138 | 0.094 | | | |
| 66 | 3PLM | 0.858 | -0.774 | 0.095 | | | |
| 67 | 3PLM | 1.275 | 0.723 | 0.196 | | | |
| 68 | 3PLM | 1.344 | -0.011 | 0.071 | | | |
| 69 | 3PLM | 1.281 | -1.55 | 0.19 | | | |
| 70 | 3PLM | 1.098 | 1.345 | 0.131 | | | |
| 71 | 3PLM | 1.010 | 1.044 | 0.061 | | | |
| 72 | 3PLM | 1.106 | 0.875 | 0.117 | | | |
| 73 | 3PLM | 1.202 | -0.177 | 0.052 | | | |
| 74 | 3PLM | 1.071 | 1.734 | 0.068 | | | |
| 75 | 3PLM | 1.512 | 0.722 | 0.16 | | | |
| 76 | 3PLM | 0.88 | -0.299 | 0.145 | | | |
| 77 | 3PLM | 0.738 | 1.483 | 0.226 | | | |
| 78 | 3PLM | 1.176 | -1.164 | 0.145 | | | |
| 79 | 3PLM | 0.684 | -1.087 | 0.214 | | | |
| 80 | 3PLM | 0.939 | 0.385 | 0.173 | | | |
| 1 | GPCM | 0.873 | -1.342 | -0.033 | 0.233 | 0.844 | 1.76 |
| 2 | GPCM | 1.192 | -1.105 | 0.340 | 0.766 | 0.930 | 1.314 |
| 3 | GPCM | 0.763 | -0.379 | -0.061 | -0.058 | 0.141 | 0.351 |
| 4 | GPCM | 0.987 | -0.731 | -0.671 | 0.032 | 0.871 | 3.168 |
| 5 | GPCM | 1.208 | -0.408 | -0.166 | -0.079 | 0.349 | 2.763 |
| 6 | GPCM | 1.041 | -1.681 | -0.915 | -0.007 | 0.206 | 0.995 |
| 7 | GPCM | 0.974 | 0.377 | 0.453 | 1.358 | 1.599 | 2.715 |
| 8 | GPCM | 1.237 | -0.126 | 0.212 | 0.591 | 0.701 | 1.141 |
| 9 | GPCM | 0.943 | 0.172 | 0.192 | 0.278 | 0.446 | 1.624 |
| 10 | GPCM | 0.783 | -0.185 | 0.149 | 0.272 | 0.351 | 1.045 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 42 ค่าพารามิเตอร์ของข้อสอบที่ใช้ในการจำลองข้อมูลผลการตอบข้อสอบ สำหรับ **แบบสอบฉบับ** B1

| Item | Model | พารามิเตอร์ข้อสอบ | | | | | |
|---|---|---|---|---|---|---|---|
| | | a | b/b1 | c/b2 | b3 | b4 | b5 |
| 1 | 3PLM | 0.787 | -2.273 | 0.088 | | | |
| 2 | 3PLM | 0.885 | -0.538 | 0.094 | | | |
| 3 | 3PLM | 0.661 | 1.724 | 0.158 | | | |
| 4 | 3PLM | 1.083 | 0.385 | 0.178 | | | |
| 5 | 3PLM | 0.796 | -0.697 | 0.144 | | | |
| 6 | 3PLM | 0.853 | 1.669 | 0.060 | | | |
| 7 | 3PLM | 1.002 | -0.671 | 0.053 | | | |
| 8 | 3PLM | 1.020 | 0.870 | 0.091 | | | |
| 9 | 3PLM | 1.430 | -0.585 | 0.089 | | | |
| 10 | 3PLM | 1.276 | 0.211 | 0.155 | | | |
| 11 | 3PLM | 1.391 | -1.653 | 0.171 | | | |
| 12 | 3PLM | 0.950 | -0.367 | 0.128 | | | |
| 13 | 3PLM | 0.892 | -0.326 | 0.035 | | | |
| 14 | 3PLM | 0.804 | -0.985 | 0.156 | | | |
| 15 | 3PLM | 1.362 | -1.175 | 0.185 | | | |
| 16 | 3PLM | 1.091 | 0.964 | 0.090 | | | |
| 17 | 3PLM | 0.954 | 0.673 | 0.104 | | | |
| 18 | 3PLM | 0.958 | 0.025 | 0.104 | | | |
| 19 | 3PLM | 1.019 | 1.911 | 0.194 | | | |
| 20 | 3PLM | 0.785 | -0.261 | 0.113 | | | |
| 21 | 3PLM | 1.617 | 2.018 | 0.298 | | | |
| 22 | 3PLM | 0.978 | -0.580 | 0.057 | | | |
| 23 | 3PLM | 1.554 | 3.094 | 0.044 | | | |
| 24 | 3PLM | 1.194 | -1.975 | 0.111 | | | |
| 25 | 3PLM | 1.581 | 1.278 | 0.119 | | | |
| 26 | 3PLM | 1.067 | 0.033 | 0.125 | | | |
| 27 | 3PLM | 1.092 | -1.118 | 0.059 | | | |
| 28 | 3PLM | 0.846 | 0.877 | 0.076 | | | |
| 29 | 3PLM | 0.838 | 1.029 | 0.216 | | | |
| 30 | 3PLM | 0.890 | 1.339 | 0.146 | | | |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 42 ค่าพารามิเตอร์ของข้อสอบที่ใช้ในการจำลองข้อมูลผลการตอบข้อสอบ สำหรับ **แบบสอบฉบับ** B1 (ต่อ)

| Item | Model | พารามิเตอร์ข้อสอบ | | | | | |
|---|---|---|---|---|---|---|---|
| | | a | b/b1 | c/b2 | b3 | b4 | b5 |
| 31 | 3PLM | 0.680 | -0.607 | 0.172 | | | |
| 32 | 3PLM | 0.894 | -0.785 | 0.044 | | | |
| 33 | 3PLM | 0.980 | 0.528 | 0.240 | | | |
| 34 | 3PLM | 0.798 | 0.137 | 0.101 | | | |
| 35 | 3PLM | 1.320 | 1.049 | 0.125 | | | |
| 36 | 3PLM | 0.980 | -0.340 | 0.124 | | | |
| 37 | 3PLM | 0.794 | -1.011 | 0.096 | | | |
| 38 | 3PLM | 0.986 | 0.725 | 0.078 | | | |
| 39 | 3PLM | 0.683 | 0.799 | 0.124 | | | |
| 40 | 3PLM | 1.088 | -0.061 | 0.162 | | | |
| 41 | 3PLM | 1.166 | 0.392 | 0.127 | | | |
| 42 | 3PLM | 0.896 | -0.761 | 0.098 | | | |
| 43 | 3PLM | 0.741 | 0.060 | 0.084 | | | |
| 44 | 3PLM | 0.975 | -0.375 | 0.057 | | | |
| 45 | 3PLM | 0.639 | 1.289 | 0.172 | | | |
| 46 | 3PLM | 1.074 | -0.344 | 0.032 | | | |
| 47 | 3PLM | 0.906 | -0.561 | 0.163 | | | |
| 48 | 3PLM | 1.170 | 0.093 | 0.078 | | | |
| 49 | 3PLM | 1.002 | -0.458 | 0.089 | | | |
| 50 | 3PLM | 0.856 | -0.409 | 0.208 | | | |
| 51 | 3PLM | 0.907 | -1.433 | 0.078 | | | |
| 52 | 3PLM | 1.201 | 0.957 | 0.157 | | | |
| 53 | 3PLM | 1.430 | 0.673 | 0.247 | | | |
| 54 | 3PLM | 1.187 | 0.589 | 0.060 | | | |
| 55 | 3PLM | 0.859 | 0.420 | 0.086 | | | |
| 56 | 3PLM | 1.234 | 0.209 | 0.191 | | | |
| 57 | 3PLM | 1.082 | -0.197 | 0.200 | | | |
| 58 | 3PLM | 1.156 | -0.602 | 0.157 | | | |
| 59 | 3PLM | 1.100 | -0.322 | 0.103 | | | |
| 60 | 3PLM | 1.095 | -0.806 | 0.114 | | | |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 42 ค่าพารามิเตอร์ของข้อสอบที่ใช้ในการจำลองข้อมูลผลการตอบข้อสอบ สำหรับ **แบบสอบฉบับ** B1 (ต่อ)

| Item | Model | พารามิเตอร์ข้อสอบ | | | | | |
|---|---|---|---|---|---|---|---|
| | | a | b/b1 | c/b2 | b3 | b4 | b5 |
| 61 | 3PLM | 1.005 | 1.233 | 0.067 | | | |
| 62 | 3PLM | 1.520 | 1.275 | 0.107 | | | |
| 63 | 3PLM | 1.305 | -0.102 | 0.141 | | | |
| 64 | 3PLM | 1.204 | -1.484 | 0.176 | | | |
| 65 | 3PLM | 1.008 | -1.520 | 0.129 | | | |
| 66 | 3PLM | 1.023 | -0.623 | 0.112 | | | |
| 67 | 3PLM | 1.030 | -0.855 | 0.159 | | | |
| 68 | 3PLM | 1.032 | 0.588 | 0.061 | | | |
| 69 | 3PLM | 1.005 | -0.212 | 0.095 | | | |
| 70 | 3PLM | 1.150 | -0.308 | 0.134 | | | |
| 71 | 3PLM | 0.953 | 1.846 | 0.119 | | | |
| 72 | 3PLM | 0.898 | 0.252 | 0.068 | | | |
| 73 | 3PLM | 0.912 | -0.166 | 0.259 | | | |
| 74 | 3PLM | 0.958 | 0.460 | 0.187 | | | |
| 75 | 3PLM | 1.238 | -0.985 | 0.172 | | | |
| 76 | 3PLM | 0.661 | 0.282 | 0.119 | | | |
| 77 | 3PLM | 0.957 | -0.505 | 0.059 | | | |
| 78 | 3PLM | 0.984 | 0.667 | 0.086 | | | |
| 79 | 3PLM | 1.352 | 1.658 | 0.139 | | | |
| 80 | 3PLM | 1.101 | 0.573 | 0.082 | | | |
| 1 | GPCM | 1.088 | -0.662 | -0.517 | -0.476 | 0.713 | 0.927 |
| 2 | GPCM | 1.090 | -0.227 | -0.215 | 0.032 | 0.730 | 0.843 |
| 3 | GPCM | 1.039 | -0.881 | -0.209 | -0.025 | 1.593 | 2.873 |
| 4 | GPCM | 1.255 | -0.629 | -0.628 | -0.050 | 0.017 | 0.497 |
| 5 | GPCM | 1.173 | -0.948 | 0.142 | 0.256 | 0.589 | 0.655 |
| 6 | GPCM | 0.675 | -0.372 | 0.066 | 0.303 | 1.103 | 1.205 |
| 7 | GPCM | 1.050 | -0.879 | -0.280 | 0.235 | 0.344 | 2.394 |
| 8 | GPCM | 1.063 | -1.334 | -1.038 | -0.325 | -0.200 | 2.522 |
| 9 | GPCM | 0.911 | -1.383 | -0.804 | -0.616 | -0.541 | -0.242 |
| 10 | GPCM | 0.896 | 0.170 | 0.771 | 1.322 | 1.499 | 2.046 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 43 ค่าพารามิเตอร์ของข้อสอบที่ใช้ในการจำลองข้อมูลผลการตอบข้อสอบ สำหรับ **แบบสอบฉบับ** B2

| Item | Model | พารามิเตอร์ข้อสอบ | | | | | |
|---|---|---|---|---|---|---|---|
| | | a | b/b1 | c/b2 | b3 | b4 | b5 |
| 1 | 3PLM | 0.877 | -0.598 | 0.256 | | | |
| 2 | 3PLM | 1.077 | 0.518 | 0.048 | | | |
| 3 | 3PLM | 1.150 | -1.044 | 0.127 | | | |
| 4 | 3PLM | 0.766 | 0.733 | 0.093 | | | |
| 5 | 3PLM | 1.344 | -0.283 | 0.030 | | | |
| 6 | 3PLM | 1.147 | 0.153 | 0.147 | | | |
| 7 | 3PLM | 0.932 | 0.405 | 0.139 | | | |
| 8 | 3PLM | 1.010 | 0.244 | 0.092 | | | |
| 9 | 3PLM | 1.382 | 1.616 | 0.076 | | | |
| 10 | 3PLM | 0.968 | -0.574 | 0.098 | | | |
| 11 | 3PLM | 1.033 | -0.571 | 0.225 | | | |
| 12 | 3PLM | 0.950 | 0.680 | 0.210 | | | |
| 13 | 3PLM | 0.784 | 2.500 | 0.083 | | | |
| 14 | 3PLM | 1.298 | 0.667 | 0.080 | | | |
| 15 | 3PLM | 0.877 | 1.193 | 0.201 | | | |
| 16 | 3PLM | 0.863 | 0.472 | 0.159 | | | |
| 17 | 3PLM | 0.666 | 0.628 | 0.101 | | | |
| 18 | 3PLM | 0.750 | 1.199 | 0.016 | | | |
| 19 | 3PLM | 0.952 | -0.991 | 0.065 | | | |
| 20 | 3PLM | 1.001 | -0.239 | 0.114 | | | |
| 21 | 3PLM | 1.021 | -0.626 | 0.201 | | | |
| 22 | 3PLM | 0.836 | -1.365 | 0.114 | | | |
| 23 | 3PLM | 1.154 | 1.146 | 0.069 | | | |
| 24 | 3PLM | 0.746 | -0.888 | 0.038 | | | |
| 25 | 3PLM | 0.824 | -0.570 | 0.139 | | | |
| 26 | 3PLM | 0.969 | -0.544 | 0.061 | | | |
| 27 | 3PLM | 1.045 | 1.133 | 0.198 | | | |
| 28 | 3PLM | 1.208 | -0.978 | 0.154 | | | |
| 29 | 3PLM | 0.745 | 1.068 | 0.195 | | | |
| 30 | 3PLM | 0.967 | 0.674 | 0.118 | | | |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 43 ค่าพารามิเตอร์ของข้อสอบที่ใช้ในการจำลองข้อมูลผลการตอบข้อสอบ สำหรับ **แบบสอบฉบับ** B2 (ต่อ)

| Item | Model | พารามิเตอร์ข้อสอบ | | | | | |
|---|---|---|---|---|---|---|---|
| | | a | b/b1 | c/b2 | b3 | b4 | b5 |
| 31 | 3PLM | 1.222 | 0.126 | 0.153 | | | |
| 32 | 3PLM | 0.995 | -0.592 | 0.108 | | | |
| 33 | 3PLM | 1.065 | 0.481 | 0.151 | | | |
| 34 | 3PLM | 0.783 | -1.647 | 0.156 | | | |
| 35 | 3PLM | 1.156 | 2.118 | 0.069 | | | |
| 36 | 3PLM | 0.542 | -1.127 | 0.182 | | | |
| 37 | 3PLM | 0.942 | 0.588 | 0.114 | | | |
| 38 | 3PLM | 0.892 | 1.937 | 0.057 | | | |
| 39 | 3PLM | 0.905 | -2.164 | 0.059 | | | |
| 40 | 3PLM | 1.105 | -1.843 | 0.121 | | | |
| 41 | 3PLM | 1.110 | 1.985 | 0.263 | | | |
| 42 | 3PLM | 1.062 | 0.421 | 0.075 | | | |
| 43 | 3PLM | 0.836 | -0.642 | 0.260 | | | |
| 44 | 3PLM | 1.061 | -0.882 | 0.052 | | | |
| 45 | 3PLM | 1.105 | -0.770 | 0.037 | | | |
| 46 | 3PLM | 1.044 | 0.912 | 0.223 | | | |
| 47 | 3PLM | 0.763 | -1.339 | 0.210 | | | |
| 48 | 3PLM | 1.099 | -2.028 | 0.090 | | | |
| 49 | 3PLM | 0.977 | 0.045 | 0.045 | | | |
| 50 | 3PLM | 1.057 | 0.717 | 0.145 | | | |
| 51 | 3PLM | 0.764 | 0.673 | 0.099 | | | |
| 52 | 3PLM | 1.023 | -0.184 | 0.057 | | | |
| 53 | 3PLM | 0.821 | -0.310 | 0.335 | | | |
| 54 | 3PLM | 1.418 | 0.335 | 0.028 | | | |
| 55 | 3PLM | 1.193 | 1.012 | 0.068 | | | |
| 56 | 3PLM | 0.935 | 0.199 | 0.089 | | | |
| 57 | 3PLM | 0.985 | 0.333 | 0.116 | | | |
| 58 | 3PLM | 1.505 | -0.470 | 0.183 | | | |
| 59 | 3PLM | 0.926 | 1.238 | 0.062 | | | |
| 60 | 3PLM | 1.321 | 0.049 | 0.144 | | | |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 43 ค่าพารามิเตอร์ของข้อสอบที่ใช้ในการจำลองข้อมูลผลการตอบข้อสอบ สำหรับ **แบบสอบฉบับ** B2 (ต่อ)

| Item | Model | พารามิเตอร์ข้อสอบ | | | | | |
|---|---|---|---|---|---|---|---|
| | | a | b/b1 | c/b2 | b3 | b4 | b5 |
| 61 | 3PLM | 0.911 | 1.847 | 0.202 | | | |
| 62 | 3PLM | 0.813 | -0.270 | 0.069 | | | |
| 63 | 3PLM | 0.704 | -1.944 | 0.050 | | | |
| 64 | 3PLM | 0.806 | -1.125 | 0.115 | | | |
| 65 | 3PLM | 1.830 | 0.435 | 0.184 | | | |
| 66 | 3PLM | 0.982 | 2.093 | 0.068 | | | |
| 67 | 3PLM | 1.045 | 0.745 | 0.045 | | | |
| 68 | 3PLM | 0.827 | 0.908 | 0.184 | | | |
| 69 | 3PLM | 0.933 | -1.635 | 0.178 | | | |
| 70 | 3PLM | 1.079 | -1.556 | 0.134 | | | |
| 71 | 3PLM | 0.890 | 0.499 | 0.122 | | | |
| 72 | 3PLM | 1.030 | -0.266 | 0.107 | | | |
| 73 | 3PLM | 0.868 | -1.813 | 0.090 | | | |
| 74 | 3PLM | 1.024 | 1.390 | 0.197 | | | |
| 75 | 3PLM | 1.059 | -0.779 | 0.153 | | | |
| 76 | 3PLM | 0.773 | -0.813 | 0.167 | | | |
| 77 | 3PLM | 0.839 | -0.261 | 0.051 | | | |
| 78 | 3PLM | 1.235 | -0.378 | 0.062 | | | |
| 79 | 3PLM | 0.775 | 1.003 | 0.191 | | | |
| 80 | 3PLM | 0.871 | -0.258 | 0.122 | | | |
| 1 | GPCM | 0.987 | -1.093 | -0.604 | -0.549 | -0.468 | 0.029 |
| 2 | GPCM | 0.916 | -1.422 | -0.264 | -0.085 | 0.130 | 0.235 |
| 3 | GPCM | 1.287 | -0.708 | 0.148 | 0.317 | 0.769 | 1.505 |
| 4 | GPCM | 1.106 | -0.416 | -0.221 | 0.094 | 0.458 | 2.080 |
| 5 | GPCM | 0.912 | -1.964 | -1.808 | 0.184 | 0.328 | 0.735 |
| 6 | GPCM | 0.873 | -0.343 | 0.274 | 1.504 | 1.506 | 2.956 |
| 7 | GPCM | 1.199 | -1.316 | -0.535 | -0.402 | 0.168 | 0.760 |
| 8 | GPCM | 1.130 | -0.461 | 0.268 | 0.303 | 0.727 | 1.744 |
| 9 | GPCM | 0.865 | -0.406 | -0.009 | 0.379 | 1.276 | 2.413 |
| 10 | GPCM | 1.259 | 0.735 | 1.189 | 1.262 | 1.337 | 2.184 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

# ภาคผนวก ข

## การตรวจสอบคุณภาพของข้อมูลที่ได้จากการจำลองข้อมูล

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

**ผลการเปรียบเทียบค่าเฉลี่ยพารามิเตอร์ข้อสอบ**
**ที่ใช้ในการจำลองข้อมูลที่ได้จากโปรแกรม** Wingen
**กับผลจากการวิเคราะห์พารามิเตอร์ข้อสอบ**
**ที่ได้จากโปรแกรม** MULTILOG

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 44 ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test A จากการประมาณค่า 100 ครั้ง

| ครั้งที่ | a | | | b | | | c | | | alpha | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD | t | sig | MD | t | sig | MD | t | sig | MD | t | sig |
| 1 | 0.53 | 0.53 | 0.60 | 0.03 | 0.23 | 0.82 | 0.00 | 0.30 | 0.76 | -0.05 | -0.99 | 0.35 |
| 2 | -0.19 | -0.19 | 0.85 | -0.01 | -0.06 | 0.95 | 0.00 | 0.43 | 0.67 | 0.10 | 1.18 | 0.27 |
| 3 | -0.14 | -0.14 | 0.89 | -0.01 | -0.03 | 0.97 | 0.00 | 0.22 | 0.82 | 0.06 | 0.56 | 0.59 |
| 4 | 0.91 | 0.91 | 0.37 | 0.03 | 0.19 | 0.85 | 0.00 | 0.04 | 0.97 | -0.02 | -0.29 | 0.78 |
| 5 | 0.91 | 0.91 | 0.37 | 0.03 | 0.19 | 0.85 | 0.00 | 0.04 | 0.97 | 0.11 | 1.31 | 0.22 |
| 6 | -0.58 | -0.58 | 0.56 | -0.02 | -0.13 | 0.90 | 0.00 | -0.28 | 0.78 | -0.06 | -0.84 | 0.42 |
| 7 | -0.58 | -0.58 | 0.56 | -0.02 | -0.13 | 0.90 | 0.00 | -0.28 | 0.78 | 0.00 | 0.01 | 0.99 |
| 8 | -0.98 | -0.98 | 0.33 | 0.04 | 0.30 | 0.77 | -0.01 | -1.02 | 0.31 | -0.02 | -0.44 | 0.67 |
| 9 | 1.04 | 1.04 | 0.30 | 0.06 | 0.38 | 0.70 | 0.00 | 0.33 | 0.74 | -0.04 | -0.47 | 0.65 |
| 10 | -1.68 | -1.68 | 0.10 | 0.20 | 1.34 | 0.18 | 0.02 | 1.42 | 0.16 | 0.08 | 0.88 | 0.40 |
| 11 | -1.08 | -1.08 | 0.28 | -0.01 | -0.09 | 0.93 | 0.00 | -0.07 | 0.95 | 0.03 | 0.38 | 0.72 |
| 12 | -1.44 | -1.44 | 0.15 | 0.06 | 0.38 | 0.71 | -0.01 | -0.81 | 0.42 | 0.06 | 1.11 | 0.30 |
| 13 | -0.74 | -0.74 | 0.46 | -0.02 | -0.11 | 0.91 | -0.01 | -1.24 | 0.22 | 0.00 | 0.04 | 0.97 |
| 14 | -0.61 | -0.61 | 0.54 | 0.07 | 0.46 | 0.65 | 0.00 | 0.08 | 0.93 | -0.04 | -0.50 | 0.63 |
| 15 | -0.61 | -0.61 | 0.54 | 0.07 | 0.46 | 0.65 | 0.00 | 0.08 | 0.93 | -0.06 | -0.47 | 0.65 |
| 16 | -1.55 | -1.55 | 0.13 | -0.11 | -0.68 | 0.50 | 0.00 | 0.21 | 0.83 | 0.02 | 0.20 | 0.85 |
| 17 | -1.55 | -1.55 | 0.13 | -0.11 | -0.68 | 0.50 | 0.00 | 0.21 | 0.83 | 0.00 | 0.10 | 0.93 |
| 18 | -1.60 | -1.60 | 0.11 | -0.02 | -0.14 | 0.89 | 0.00 | 0.23 | 0.82 | 0.04 | 0.39 | 0.70 |
| 19 | 0.36 | 0.36 | 0.72 | 0.26 | 1.65 | 0.10 | 0.00 | -0.26 | 0.80 | 0.00 | -0.02 | 0.98 |
| 20 | -1.62 | -1.62 | 0.11 | -0.16 | -1.10 | 0.27 | 0.01 | 0.78 | 0.44 | 0.07 | 0.74 | 0.48 |
| 21 | -1.07 | -1.07 | 0.29 | 0.03 | 0.19 | 0.85 | 0.00 | -0.24 | 0.81 | 0.06 | 0.81 | 0.44 |
| 22 | -0.55 | -0.55 | 0.58 | -0.02 | -0.12 | 0.91 | -0.01 | -1.08 | 0.29 | 0.12 | 1.33 | 0.22 |
| 23 | -0.38 | -0.38 | 0.70 | 0.02 | 0.11 | 0.92 | 0.01 | 0.89 | 0.38 | 0.01 | 0.13 | 0.90 |
| 24 | -0.69 | -0.69 | 0.49 | -0.18 | -1.13 | 0.26 | 0.00 | 0.39 | 0.69 | -0.05 | -0.98 | 0.35 |
| 25 | -0.69 | -0.69 | 0.49 | -0.18 | -1.13 | 0.26 | 0.00 | 0.39 | 0.69 | -0.02 | -0.30 | 0.77 |
| 26 | -1.70 | -1.70 | 0.09 | 0.01 | 0.06 | 0.95 | -0.01 | -0.94 | 0.35 | 0.17 | 2.10 | 0.06 |
| 27 | -1.70 | -1.70 | 0.09 | 0.01 | 0.06 | 0.95 | -0.01 | -0.94 | 0.35 | 0.15 | 2.40 | 0.04 |
| 28 | -1.52 | -1.52 | 0.13 | 0.03 | 0.22 | 0.83 | -0.01 | -0.01 | 0.13 | 0.06 | 0.71 | 0.50 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 44 ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test A จากการประมาณค่า 100 ครั้ง (ต่อ)

| ครั้งที่ | a | | | b | | | c | | | alpha | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD | t | sig | MD | t | sig | MD | t | sig | MD | t | sig |
| 29 | -0.95 | -0.95 | 0.35 | -0.02 | -0.09 | 0.93 | 0.00 | -0.33 | 0.74 | -0.03 | -0.40 | 0.70 |
| 30 | -0.68 | -0.68 | 0.50 | 0.06 | 0.36 | 0.72 | 0.00 | -0.04 | 0.96 | 0.12 | 1.05 | 0.32 |
| 31 | -0.69 | -0.69 | 0.49 | 0.07 | 0.47 | 0.64 | 0.00 | -0.46 | 0.64 | 0.13 | 1.33 | 0.22 |
| 32 | -0.36 | -0.36 | 0.72 | 0.09 | 0.55 | 0.59 | 0.01 | 1.18 | 0.24 | 0.06 | 0.53 | 0.61 |
| 33 | -1.25 | -1.25 | 0.22 | 0.12 | 0.79 | 0.43 | 0.01 | 1.36 | 0.18 | -0.07 | -0.80 | 0.45 |
| 34 | -1.15 | -1.15 | 0.25 | 0.18 | 1.17 | 0.24 | 0.00 | 0.29 | 0.77 | 0.07 | 0.83 | 0.43 |
| 35 | -1.15 | -1.15 | 0.25 | 0.18 | 1.17 | 0.24 | 0.00 | 0.29 | 0.77 | 0.00 | 0.04 | 0.97 |
| 36 | -1.32 | -1.32 | 0.19 | -0.06 | -0.43 | 0.67 | 0.00 | -0.17 | 0.86 | -0.01 | -0.15 | 0.89 |
| 37 | -1.32 | -1.32 | 0.19 | -0.06 | -0.43 | 0.67 | 0.00 | -0.17 | 0.86 | 0.04 | 0.45 | 0.66 |
| 38 | -1.14 | -1.14 | 0.26 | 0.11 | 0.76 | 0.45 | -0.01 | -1.25 | 0.21 | 0.00 | -0.03 | 0.97 |
| 39 | -1.21 | -1.21 | 0.23 | 0.12 | 0.73 | 0.47 | -0.01 | -0.86 | 0.39 | -0.02 | -0.35 | 0.73 |
| 40 | -0.77 | -0.77 | 0.44 | -0.05 | -0.34 | 0.73 | 0.00 | 0.26 | 0.80 | -0.01 | -0.11 | 0.92 |
| 41 | -0.82 | -0.82 | 0.42 | -0.13 | -0.77 | 0.45 | -0.01 | -0.88 | 0.38 | 0.10 | 1.05 | 0.32 |
| 42 | -1.16 | -1.16 | 0.25 | 0.21 | 1.26 | 0.21 | 0.00 | -0.48 | 0.63 | -0.02 | -0.26 | 0.80 |
| 43 | -1.38 | -1.38 | 0.17 | -0.08 | -0.54 | 0.59 | 0.00 | 0.41 | 0.68 | 0.07 | 0.67 | 0.52 |
| 44 | -0.60 | -0.60 | 0.55 | 0.09 | 0.53 | 0.60 | 0.01 | 0.97 | 0.34 | 0.07 | 0.83 | 0.43 |
| 45 | -0.60 | -0.60 | 0.55 | 0.09 | 0.53 | 0.60 | 0.01 | 0.97 | 0.34 | 0.01 | 0.09 | 0.93 |
| 46 | -1.33 | -1.33 | 0.19 | -0.15 | -1.02 | 0.31 | 0.01 | 0.73 | 0.47 | 0.04 | 0.72 | 0.49 |
| 47 | -1.33 | -1.33 | 0.19 | -0.15 | -1.02 | 0.31 | 0.01 | 0.73 | 0.47 | 0.13 | 1.49 | 0.17 |
| 48 | -1.25 | -1.25 | 0.21 | 0.02 | 0.11 | 0.91 | -0.01 | -0.81 | 0.42 | -0.02 | -0.19 | 0.86 |
| 49 | 0.36 | 0.36 | 0.72 | -0.05 | -0.29 | 0.77 | 0.00 | -0.01 | 0.99 | 0.11 | 1.15 | 0.28 |
| 50 | -2.27 | -2.27 | 0.21 | -0.05 | -0.28 | 0.78 | 0.01 | 1.35 | 0.18 | -0.06 | -0.97 | 0.36 |
| 51 | -1.15 | -1.15 | 0.25 | 0.19 | 1.25 | 0.22 | 0.00 | -0.05 | 0.96 | 0.07 | 0.96 | 0.36 |
| 52 | -1.89 | -1.89 | 0.06 | 0.01 | 0.09 | 0.93 | 0.00 | -0.04 | 0.97 | 0.02 | 0.17 | 0.87 |
| 53 | -0.85 | -0.85 | 0.40 | -0.05 | -0.32 | 0.75 | 0.00 | -0.08 | 0.94 | 0.08 | 0.76 | 0.47 |
| 54 | -1.66 | -1.66 | 0.10 | -0.18 | -1.26 | 0.21 | -0.01 | -0.68 | 0.50 | -0.07 | -0.83 | 0.43 |
| 55 | -1.66 | -1.66 | 0.10 | -0.18 | -1.26 | 0.21 | -0.01 | -0.68 | 0.50 | -0.06 | -0.70 | 0.50 |
| 56 | -0.19 | -0.19 | 0.85 | -0.02 | -0.12 | 0.91 | 0.00 | -0.13 | 0.90 | 0.17 | 1.88 | 0.09 |
| 57 | -0.19 | -0.19 | 0.85 | -0.02 | -0.12 | 0.91 | 0.00 | -0.13 | 0.90 | 0.02 | 0.37 | 0.72 |
| 58 | -2.04 | -2.04 | 0.05 | 0.04 | 0.28 | 0.78 | 0.01 | 0.57 | 0.57 | -0.04 | -0.65 | 0.53 |

ตารางที่ 44 ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test A จากการประมาณค่า 100 ครั้ง (ต่อ)

| ครั้งที่ | a | | | b | | | c | | | alpha | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD | t | sig | MD | t | sig | MD | t | sig | MD | t | sig |
| 59 | -0.81 | -0.81 | 0.42 | 0.03 | 0.19 | 0.85 | 0.00 | 0.51 | 0.61 | -0.07 | -1.50 | 0.17 |
| 60 | -2.16 | -2.16 | 0.03 | 0.23 | 1.56 | 0.12 | 0.00 | -0.02 | 0.99 | 0.08 | 1.24 | 0.25 |
| 61 | -0.74 | -0.74 | 0.46 | -0.03 | -0.18 | 0.85 | -0.01 | -1.09 | 0.28 | -0.11 | -1.71 | 0.12 |
| 62 | 0.05 | 0.05 | 0.96 | 0.03 | 0.18 | 0.86 | 0.00 | 0.44 | 0.66 | 0.09 | 0.84 | 0.42 |
| 63 | -1.26 | -1.26 | 0.21 | 0.17 | 1.00 | 0.32 | 0.00 | -0.05 | 0.96 | -0.05 | -0.59 | 0.57 |
| 64 | 0.83 | 0.83 | 0.41 | -0.06 | -0.42 | 0.67 | 0.01 | 0.77 | 0.44 | -0.05 | -0.64 | 0.54 |
| 65 | 0.83 | 0.83 | 0.41 | -0.06 | -0.42 | 0.67 | 0.01 | 0.77 | 0.44 | 0.00 | 0.05 | 0.96 |
| 66 | -0.09 | -0.09 | 0.93 | -0.06 | -0.45 | 0.65 | 0.00 | -0.36 | 0.72 | -0.05 | -0.54 | 0.60 |
| 67 | -0.09 | -0.09 | 0.93 | -0.06 | -0.45 | 0.65 | 0.00 | -0.36 | 0.72 | -0.02 | -0.31 | 0.77 |
| 68 | -0.88 | -0.88 | 0.38 | 0.15 | 0.89 | 0.38 | 0.00 | -0.24 | 0.81 | 0.02 | 0.21 | 0.84 |
| 69 | -1.04 | -1.04 | 0.30 | -0.09 | -0.61 | 0.54 | 0.01 | 0.51 | 0.61 | 0.05 | 0.53 | 0.61 |
| 70 | -0.17 | -0.17 | 0.86 | 0.04 | 0.25 | 0.80 | 0.00 | -0.02 | 0.98 | -0.05 | -0.49 | 0.64 |
| 71 | -0.64 | -0.64 | 0.52 | -0.05 | -0.35 | 0.73 | 0.00 | 0.06 | 0.95 | 0.02 | 0.29 | 0.78 |
| 72 | -0.76 | -0.76 | 0.45 | 0.07 | 0.44 | 0.66 | 0.00 | 0.28 | 0.78 | -0.05 | -0.74 | 0.48 |
| 73 | -0.67 | -0.67 | 0.51 | 0.09 | 0.63 | 0.53 | -0.01 | -1.42 | 0.16 | 0.01 | 0.18 | 0.86 |
| 74 | -0.35 | -0.35 | 0.73 | 0.10 | 0.60 | 0.55 | -0.01 | -0.69 | 0.49 | 0.09 | 1.04 | 0.32 |
| 75 | -0.35 | -0.35 | 0.73 | 0.10 | 0.60 | 0.55 | -0.01 | -0.69 | 0.49 | 0.01 | 0.12 | 0.91 |
| 76 | -1.29 | -1.29 | 0.20 | -0.08 | -0.57 | 0.57 | 0.00 | -0.44 | 0.66 | -0.03 | -0.44 | 0.67 |
| 77 | -1.29 | -1.29 | 0.20 | -0.08 | -0.57 | 0.57 | 0.00 | -0.44 | 0.66 | 0.00 | 0.03 | 0.98 |
| 78 | -0.04 | -0.04 | 0.97 | 0.04 | 0.29 | 0.77 | 0.00 | -0.49 | 0.63 | -0.01 | -0.12 | 0.90 |
| 79 | -0.70 | -0.70 | 0.48 | -0.03 | -0.18 | 0.86 | 0.00 | -0.10 | 0.92 | 0.14 | 2.06 | 0.07 |
| 80 | 0.47 | 0.47 | 0.64 | -0.07 | -0.44 | 0.66 | 0.01 | 0.92 | 0.36 | -0.06 | -0.84 | 0.42 |
| 81 | -1.18 | -1.18 | 0.24 | -0.12 | -0.77 | 0.44 | 0.00 | -0.40 | 0.69 | -0.02 | -0.28 | 0.79 |
| 82 | -0.78 | -0.78 | 0.44 | 0.06 | 0.39 | 0.70 | 0.00 | 0.29 | 0.77 | -0.02 | -0.35 | 0.73 |
| 83 | 0.04 | 0.04 | 0.97 | 0.00 | -0.02 | 0.98 | 0.00 | 0.32 | 0.75 | 0.05 | 0.86 | 0.41 |
| 84 | -0.11 | -0.11 | 0.91 | -0.18 | -1.18 | 0.24 | 0.00 | -0.03 | 0.98 | -0.06 | -0.67 | 0.52 |
| 85 | -0.11 | -0.11 | 0.91 | -0.18 | -1.18 | 0.24 | 0.00 | -0.03 | 0.98 | 0.03 | 0.46 | 0.65 |
| 86 | 0.43 | 0.43 | 0.67 | 0.10 | 0.64 | 0.53 | 0.00 | 0.38 | 0.70 | 0.11 | 1.14 | 0.28 |
| 87 | 0.43 | 0.43 | 0.67 | 0.10 | 0.64 | 0.53 | 0.00 | 0.38 | 0.70 | -0.01 | -0.06 | 0.96 |
| 88 | -1.33 | -1.33 | 0.19 | 0.20 | 1.21 | 0.23 | 0.00 | 0.41 | 0.68 | 0.00 | 0.04 | 0.97 |
| 89 | -0.31 | -0.31 | 0.76 | 0.12 | 0.81 | 0.42 | 0.00 | 0.47 | 0.64 | -0.01 | -0.15 | 0.89 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 44 ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test A จากการประมาณค่า 100 ครั้ง (ต่อ)

| ครั้งที่ | a | | | b | | | c | | | alpha | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD | t | sig | MD | t | sig | MD | t | sig | MD | t | sig |
| 90 | -1.04 | -1.04 | 0.30 | 0.17 | 1.19 | 0.24 | 0.01 | 0.73 | 0.47 | 0.15 | 2.18 | 0.06 |
| 91 | -0.95 | -0.95 | 0.35 | -0.04 | -0.29 | 0.77 | -0.01 | -0.86 | 0.39 | -0.01 | -0.13 | 0.90 |
| 92 | -1.49 | -1.49 | 0.14 | 0.12 | 0.12 | 0.45 | 0.00 | -0.46 | 0.64 | 0.13 | 1.19 | 0.26 |
| 93 | -0.72 | -0.72 | 0.47 | 0.01 | 0.09 | 0.93 | 0.00 | 0.13 | 0.90 | 0.16 | 1.48 | 0.17 |
| 94 | -1.69 | -1.69 | 0.09 | 0.11 | 0.71 | 0.48 | 0.00 | -0.51 | 0.61 | 0.11 | 0.99 | 0.35 |
| 95 | -1.69 | -1.69 | 0.09 | 0.11 | 0.71 | 0.48 | 0.00 | -0.51 | 0.61 | 0.09 | 0.76 | 0.46 |
| 96 | -1.91 | -1.91 | 0.06 | -0.05 | -0.33 | 0.75 | 0.01 | 1.00 | 0.32 | 0.00 | -0.03 | 0.97 |
| 97 | -1.91 | -1.91 | 0.06 | -0.05 | -0.33 | 0.75 | 0.01 | 1.00 | 0.32 | 0.01 | 0.11 | 0.92 |
| 98 | -0.26 | -0.26 | 0.80 | -0.03 | -0.18 | 0.86 | 0.00 | 0.28 | 0.78 | 0.01 | 0.15 | 0.88 |
| 99 | -1.14 | -1.14 | 0.26 | 0.08 | 0.50 | 0.62 | 0.01 | 0.49 | 0.63 | 0.00 | 0.00 | 1.00 |
| 100 | -1.02 | -1.02 | 0.31 | 0.10 | 0.72 | 0.47 | 0.00 | 0.34 | 0.74 | 0.08 | 1.00 | 0.34 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 45 ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test B1 จากการประมาณค่า 100 ครั้ง

| ครั้งที่ | a | | | b | | | c | | | alpha | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD | t | sig | MD | t | sig | MD | t | sig | MD | t | sig |
| 1 | -0.01 | -0.18 | 0.86 | -0.12 | -0.82 | 0.42 | 0.00 | -0.46 | 0.65 | 0.00 | 0.02 | 0.98 |
| 2 | -0.01 | -0.44 | 0.66 | -0.06 | -0.37 | 0.71 | 0.00 | 0.24 | 0.81 | 0.11 | 0.12 | 0.11 |
| 3 | 0.01 | 0.43 | 0.67 | -0.05 | -0.32 | 0.75 | 0.01 | 1.13 | 0.26 | 0.03 | 0.33 | 0.75 |
| 4 | -0.05 | -1.46 | 0.15 | 0.04 | 0.25 | 0.80 | -0.01 | -1.28 | 0.20 | 0.02 | 0.22 | 0.83 |
| 5 | -0.04 | -1.26 | 0.21 | -0.10 | -0.64 | 0.52 | 0.00 | -0.01 | 0.99 | 0.09 | 1.05 | 0.32 |
| 6 | 0.00 | 0.02 | 0.99 | -0.05 | -0.34 | 0.74 | -0.01 | -1.40 | 0.16 | 0.00 | -0.03 | 0.98 |
| 7 | -0.05 | -1.65 | 0.10 | -0.25 | -1.52 | 0.13 | 0.01 | 1.31 | 0.20 | -0.01 | -0.21 | 0.84 |
| 8 | -0.04 | -1.24 | 0.22 | -0.13 | -0.80 | 0.43 | 0.00 | 0.49 | 0.63 | 0.04 | 0.61 | 0.56 |
| 9 | -0.04 | -1.15 | 0.25 | -0.17 | -1.01 | 0.31 | 0.00 | 0.23 | 0.82 | -0.06 | -0.69 | 0.51 |
| 10 | -0.02 | -0.64 | 0.53 | -0.30 | -2.04 | 0.04 | 0.01 | 0.64 | 0.52 | 0.00 | -0.05 | 0.96 |
| 11 | 0.03 | 0.77 | 0.44 | -0.05 | -0.29 | 0.77 | 0.00 | 0.44 | 0.66 | 0.16 | 2.32 | 0.05 |
| 12 | -0.02 | -0.58 | 0.57 | -0.14 | -0.79 | 0.43 | -0.01 | -0.85 | 0.40 | -0.11 | -1.45 | 0.18 |
| 13 | -0.02 | -0.53 | 0.60 | -0.07 | -0.43 | 0.67 | 0.00 | 0.00 | 1.00 | 0.04 | 0.55 | 0.60 |
| 14 | -0.03 | -0.90 | 0.37 | -0.02 | -0.13 | 0.90 | 0.00 | -0.29 | 0.77 | -0.04 | -0.36 | 0.73 |
| 15 | -0.03 | -0.83 | 0.41 | -0.09 | -0.59 | 0.55 | 0.00 | 0.29 | 0.78 | -0.06 | -0.84 | 0.42 |
| 16 | -0.01 | -0.29 | 0.77 | -0.12 | -0.80 | 0.42 | 0.00 | -0.43 | 0.67 | 0.00 | 0.03 | 0.98 |
| 17 | -0.04 | -1.36 | 0.18 | 0.07 | 0.51 | 0.61 | 0.00 | 0.12 | 0.91 | 0.06 | 0.59 | 0.57 |
| 18 | -0.04 | -1.34 | 0.19 | -0.11 | -0.73 | 0.47 | 0.00 | -0.08 | 0.94 | -0.01 | -0.08 | 0.94 |
| 19 | -0.06 | -1.66 | 0.10 | -0.12 | -0.74 | 0.46 | 0.00 | 0.11 | 0.91 | -0.03 | -0.41 | 0.69 |
| 20 | 0.03 | 1.00 | 0.32 | 0.00 | 0.03 | 0.98 | -0.01 | -0.97 | 0.34 | -0.10 | -1.42 | 0.19 |
| 21 | 0.00 | -0.10 | 0.92 | 0.07 | 0.40 | 0.69 | 0.00 | -0.18 | 0.86 | -0.01 | -0.10 | 0.92 |
| 22 | 0.00 | 0.11 | 0.91 | -0.08 | -0.49 | 0.63 | 0.01 | 1.37 | 0.17 | 0.01 | 0.21 | 0.84 |
| 23 | -0.04 | -1.11 | 0.27 | -0.11 | -0.71 | 0.48 | -0.01 | -0.68 | 0.50 | 0.05 | 0.89 | 0.39 |
| 24 | -0.01 | -0.29 | 0.77 | -0.15 | -0.90 | 0.37 | 0.00 | 0.16 | 0.88 | -0.05 | -0.60 | 0.56 |
| 25 | -0.01 | -0.42 | 0.68 | 0.05 | 0.29 | 0.77 | 0.01 | 1.20 | 0.23 | 0.04 | 0.53 | 0.61 |
| 26 | 0.02 | 0.63 | 0.53 | -0.11 | -0.73 | 0.47 | 0.00 | 0.36 | 0.72 | 0.05 | 0.47 | 0.65 |
| 27 | 0.00 | 0.07 | 0.95 | -0.25 | -1.67 | 0.10 | 0.00 | 0.13 | 0.90 | -0.10 | -1.31 | 0.22 |
| 28 | 0.01 | 0.29 | 0.78 | -0.24 | -1.60 | 0.11 | 0.00 | 0.45 | 0.65 | 0.06 | 0.75 | 0.47 |
| 29 | -0.04 | -1.15 | 0.25 | -0.08 | -0.52 | 0.60 | 0.00 | -0.41 | 0.68 | -0.06 | -0.67 | 0.52 |
| 30 | -0.02 | -0.55 | 0.58 | -0.19 | -1.11 | 0.27 | 0.00 | 0.21 | 0.83 | -0.07 | -0.76 | 0.47 |

ตารางที่ 45 ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test B1 จากการประมาณค่า 100 ครั้ง (ต่อ)

| ครั้งที่ | a | | | b | | | c | | | alpha | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD | t | sig | MD | t | sig | MD | t | sig | MD | t | sig |
| 31 | -0.03 | -0.99 | 0.32 | 0.06 | 0.41 | 0.68 | 0.00 | -0.22 | 0.82 | 0.01 | 0.08 | 0.94 |
| 32 | -0.05 | -1.50 | 0.14 | -0.01 | -0.07 | 0.94 | 0.00 | -0.27 | 0.79 | 0.06 | 0.86 | 0.41 |
| 33 | -0.01 | -0.29 | 0.77 | -0.28 | -1.75 | 0.08 | 0.00 | 0.55 | 0.58 | 0.07 | 0.64 | 0.54 |
| 34 | -0.04 | -1.15 | 0.25 | 0.02 | 0.13 | 0.90 | 0.00 | 0.20 | 0.84 | -0.06 | -0.63 | 0.54 |
| 35 | -0.01 | -0.34 | 0.74 | -0.25 | -1.50 | 0.14 | 0.01 | 1.13 | 0.26 | -0.17 | -2.12 | 0.06 |
| 36 | 0.05 | 1.36 | 0.18 | -0.10 | -0.70 | 0.48 | 0.00 | 0.39 | 0.69 | -0.03 | -0.28 | 0.79 |
| 37 | 0.00 | 0.11 | 0.92 | 0.05 | 0.31 | 0.76 | 0.00 | 0.34 | 0.73 | 0.05 | 0.70 | 0.50 |
| 38 | -0.02 | -0.73 | 0.46 | 0.14 | 0.75 | 0.46 | 0.00 | 0.31 | 0.75 | 0.00 | -0.02 | 0.98 |
| 39 | 0.01 | 0.25 | 0.80 | -0.05 | -0.33 | 0.74 | 0.00 | -0.43 | 0.67 | -0.18 | -1.75 | 0.11 |
| 40 | -0.03 | -0.98 | 0.33 | -0.23 | -1.39 | 0.17 | 0.00 | -0.30 | 0.77 | -0.01 | -0.11 | 0.92 |
| 41 | -0.01 | -0.40 | 0.69 | -0.24 | -1.43 | 0.16 | 0.00 | 0.43 | 0.67 | -0.05 | -1.00 | 0.34 |
| 42 | -0.02 | -0.61 | 0.54 | 0.01 | 0.05 | 0.96 | -0.01 | -0.85 | 0.40 | 0.00 | 0.04 | 0.97 |
| 43 | -0.04 | -1.11 | 0.27 | -0.07 | -0.39 | 0.70 | -0.01 | -0.59 | 0.55 | 0.01 | 0.08 | 0.94 |
| 44 | -0.02 | -0.50 | 0.62 | 0.00 | 0.00 | 1.00 | 0.00 | -0.36 | 0.72 | 0.12 | 1.52 | 0.16 |
| 45 | 0.02 | 0.51 | 0.61 | -0.16 | -0.94 | 0.35 | 0.00 | 0.12 | 0.90 | -0.08 | -1.17 | 0.27 |
| 46 | -0.01 | -0.16 | 0.88 | -0.07 | -0.38 | 0.70 | 0.00 | -0.19 | 0.85 | -0.07 | -0.84 | 0.42 |
| 47 | 0.00 | 0.01 | 0.99 | 0.07 | 0.37 | 0.71 | 0.00 | -0.59 | 0.56 | 0.00 | 0.02 | 0.98 |
| 48 | -0.02 | -0.74 | 0.46 | -0.19 | -1.15 | 0.25 | -0.01 | -0.68 | 0.50 | 0.01 | 0.08 | 0.94 |
| 49 | -0.02 | -0.63 | 0.53 | -0.02 | -0.12 | 0.90 | 0.01 | 1.01 | 0.32 | 0.03 | 0.28 | 0.78 |
| 50 | -0.02 | -0.61 | 0.54 | 0.10 | 0.61 | 0.54 | 0.00 | 0.05 | 0.96 | 0.11 | 1.58 | 0.15 |
| 51 | -0.02 | -0.52 | 0.60 | 0.05 | 0.27 | 0.79 | 0.00 | -0.23 | 0.82 | -0.09 | -1.68 | 0.13 |
| 52 | -0.03 | -0.75 | 0.46 | -0.08 | -0.47 | 0.64 | 0.01 | 1.20 | 0.24 | 0.05 | 0.60 | 0.56 |
| 53 | 0.00 | 0.07 | 0.95 | -0.20 | -1.34 | 0.19 | 0.00 | 0.19 | 0.85 | -0.01 | -0.17 | 0.87 |
| 54 | 0.00 | -0.11 | 0.92 | -0.14 | -0.88 | 0.38 | 0.00 | 0.46 | 0.65 | -0.02 | -0.23 | 0.82 |
| 55 | -0.02 | -0.54 | 0.59 | -0.01 | -0.04 | 0.97 | 0.01 | 0.66 | 0.51 | -0.02 | -0.26 | 0.80 |
| 56 | -0.01 | -0.27 | 0.79 | -0.24 | -1.43 | 0.16 | 0.00 | 0.22 | 0.83 | 0.08 | 1.04 | 0.33 |
| 57 | -0.01 | -0.31 | 0.76 | -0.05 | -0.30 | 0.77 | 0.02 | 1.97 | 0.05 | -0.03 | -0.47 | 0.65 |
| 58 | 0.02 | 0.44 | 0.66 | -0.06 | -0.39 | 0.70 | 0.01 | 0.84 | 0.40 | -0.05 | -0.77 | 0.46 |
| 59 | 0.01 | 0.41 | 0.68 | 0.02 | 0.11 | 0.92 | 0.00 | -0.41 | 0.68 | -0.03 | -0.48 | 0.65 |
| 60 | 0.02 | 0.50 | 0.62 | 0.00 | -0.02 | 0.98 | -0.01 | -1.08 | 0.28 | 0.01 | 0.14 | 0.89 |

ตารางที่ 45 ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test B1 จากการประมาณค่า 100 ครั้ง (ต่อ)

| ครั้งที่ | a | | | b | | | c | | | alpha | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD | t | sig | MD | t | sig | MD | t | sig | MD | t | sig |
| 61 | -0.04 | -1.47 | 0.15 | 0.04 | 0.27 | 0.79 | 0.00 | -0.54 | 0.59 | 0.14 | 2.08 | 0.07 |
| 62 | -0.02 | -0.51 | 0.61 | -0.17 | -1.03 | 0.31 | 0.00 | -0.11 | 0.91 | 0.01 | 0.17 | 0.87 |
| 63 | 0.00 | 0.06 | 0.96 | -0.15 | -1.13 | 0.26 | -0.01 | -1.24 | 0.22 | -0.04 | -0.55 | 0.60 |
| 64 | 0.00 | -0.12 | 0.91 | 0.00 | 0.02 | 0.99 | 0.01 | 0.61 | 0.54 | 0.03 | 0.48 | 0.65 |
| 65 | -0.03 | -0.72 | 0.47 | -0.20 | -1.39 | 0.17 | 0.01 | 0.71 | 0.48 | 0.09 | 1.00 | 0.35 |
| 66 | -0.02 | -0.50 | 0.62 | -0.05 | -0.28 | 0.78 | 0.00 | -0.16 | 0.87 | -0.03 | -0.36 | 0.73 |
| 67 | -0.02 | -0.54 | 0.59 | -0.10 | -0.71 | 0.48 | 0.01 | 1.15 | 0.25 | -0.10 | -1.63 | 0.14 |
| 68 | -0.05 | -1.45 | 0.15 | -0.11 | -0.73 | 0.47 | 0.01 | 1.18 | 0.24 | -0.07 | -1.02 | 0.33 |
| 69 | 0.00 | 0.02 | 0.99 | 0.08 | 0.52 | 0.61 | 0.01 | 0.73 | 0.47 | 0.03 | 0.52 | 0.61 |
| 70 | -0.02 | -0.70 | 0.49 | -0.12 | -0.78 | 0.44 | 0.00 | -0.42 | 0.67 | -0.01 | -0.18 | 0.86 |
| 71 | 0.01 | 0.30 | 0.76 | 0.01 | 0.05 | 0.96 | 0.00 | -0.01 | 0.99 | 0.06 | 0.54 | 0.61 |
| 72 | -0.02 | -0.69 | 0.49 | 0.07 | 0.44 | 0.66 | 0.00 | -0.06 | 0.96 | 0.02 | 0.26 | 0.80 |
| 73 | -0.01 | -0.18 | 0.86 | -0.08 | -0.52 | 0.60 | 0.00 | 0.27 | 0.79 | 0.05 | 0.74 | 0.48 |
| 74 | -0.01 | -0.22 | 0.82 | -0.02 | -0.12 | 0.90 | 0.00 | -0.05 | 0.96 | 0.00 | -0.01 | 0.99 |
| 75 | -0.04 | -1.20 | 0.23 | 0.02 | 0.16 | 0.88 | -0.01 | -0.83 | 0.41 | -0.06 | -0.66 | 0.52 |
| 76 | -0.04 | -1.28 | 0.21 | -0.17 | -1.14 | 0.26 | 0.02 | 2.09 | 0.04 | 0.11 | 1.84 | 0.10 |
| 77 | 0.06 | 1.64 | 0.11 | -0.08 | -0.52 | 0.60 | 0.01 | 1.46 | 0.15 | 0.04 | 0.32 | 0.76 |
| 78 | 0.01 | 0.23 | 0.82 | -0.17 | -1.25 | 0.21 | -0.01 | 0.01 | 0.15 | 0.03 | 0.43 | 0.68 |
| 79 | -0.02 | -0.42 | 0.68 | -0.06 | -0.38 | 0.71 | -0.01 | -1.10 | 0.28 | -0.08 | -0.78 | 0.45 |
| 80 | 0.05 | 1.38 | 0.17 | -0.03 | -0.03 | 0.88 | 0.01 | 0.60 | 0.55 | -0.02 | -0.19 | 0.86 |
| 81 | 0.00 | -0.11 | 0.91 | -0.08 | -0.44 | 0.66 | 0.00 | 0.20 | 0.84 | 0.03 | 0.56 | 0.59 |
| 82 | -0.03 | -0.71 | 0.48 | 0.14 | 0.80 | 0.42 | 0.00 | -0.01 | 0.99 | 0.05 | 0.68 | 0.51 |
| 83 | -0.03 | -0.81 | 0.42 | -0.12 | -0.80 | 0.43 | 0.00 | 0.03 | 0.97 | 0.01 | 0.08 | 0.94 |
| 84 | -0.02 | -0.70 | 0.49 | -0.07 | -0.46 | 0.65 | 0.01 | 0.63 | 0.53 | -0.05 | -0.55 | 0.60 |
| 85 | 0.00 | -0.06 | 0.95 | 0.23 | 1.48 | 0.14 | 0.00 | -0.13 | 0.90 | -0.16 | -2.08 | 0.07 |
| 86 | -0.03 | -0.81 | 0.42 | -0.13 | -0.83 | 0.41 | 0.00 | 0.37 | 0.71 | 0.03 | 0.48 | 0.64 |
| 87 | 0.00 | 0.08 | 0.94 | -0.14 | -0.93 | 0.36 | 0.00 | 0.53 | 0.60 | -0.07 | -1.02 | 0.33 |
| 88 | 0.00 | -0.04 | 0.97 | 0.04 | 0.22 | 0.82 | 0.01 | 1.40 | 0.17 | 0.07 | 0.82 | 0.43 |
| 89 | -0.04 | -1.15 | 0.25 | -0.20 | -1.53 | 0.13 | 0.00 | 0.04 | 0.97 | -0.15 | -2.08 | 0.07 |
| 90 | -0.03 | -0.76 | 0.45 | -0.25 | -1.50 | 0.14 | -0.01 | -0.67 | 0.51 | -0.01 | -0.13 | 0.90 |

ตารางที่ 45 ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test B1 จากการประมาณค่า 100 ครั้ง (ต่อ)

| ครั้งที่ | a | | | b | | | c | | | alpha | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD | t | sig | MD | t | sig | MD | t | sig | MD | t | sig |
| 91 | 0.01 | 0.32 | 0.75 | 0.06 | 0.34 | 0.74 | 0.00 | 0.22 | 0.83 | 0.07 | 0.91 | 0.39 |
| 92 | -0.03 | -0.95 | 0.34 | -0.22 | -1.29 | 0.20 | 0.01 | 0.68 | 0.50 | -0.02 | -0.26 | 0.80 |
| 93 | -0.04 | -1.16 | 0.25 | -0.16 | -1.01 | 0.32 | 0.00 | 0.01 | 0.99 | -0.05 | -0.63 | 0.55 |
| 94 | -0.06 | -1.53 | 0.13 | -0.02 | -0.12 | 0.91 | 0.00 | 0.45 | 0.65 | -0.02 | -0.23 | 0.83 |
| 95 | 0.04 | 1.13 | 0.26 | -0.10 | -0.65 | 0.52 | 0.00 | 0.44 | 0.66 | -0.18 | -2.02 | 0.07 |
| 96 | -0.01 | -0.16 | 0.87 | -0.08 | -0.52 | 0.60 | 0.01 | 0.62 | 0.54 | -0.03 | -0.38 | 0.71 |
| 97 | 0.00 | -0.12 | 0.90 | -0.23 | -1.58 | 0.12 | 0.00 | 0.17 | 0.87 | 0.06 | 0.52 | 0.62 |
| 98 | 0.02 | 0.66 | 0.51 | -0.03 | -0.18 | 0.86 | -0.01 | -0.84 | 0.40 | 0.03 | 0.35 | 0.73 |
| 99 | -0.02 | -0.55 | 0.58 | -0.10 | -0.67 | 0.51 | 0.01 | 1.03 | 0.31 | 0.02 | 0.02 | 0.79 |
| 100 | -0.05 | -1.31 | 0.20 | 0.05 | 0.31 | 0.76 | 0.01 | 0.82 | 0.41 | 0.03 | 0.24 | 0.81 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 46 ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test B2 จากการประมาณค่า 100 ครั้ง

| ครั้งที่ | a | | | b | | | c | | | alpha | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD | t | Sig | MD | t | sig | MD | t | sig | MD | t | sig |
| 1 | 0.02 | 0.50 | 0.62 | -0.02 | -0.13 | 0.90 | 0.01 | 0.53 | 0.60 | -0.09 | -1.38 | 0.20 |
| 2 | 0.04 | 1.29 | 0.20 | -0.08 | -0.57 | 0.57 | 0.01 | 0.62 | 0.54 | -0.04 | -0.57 | 0.58 |
| 3 | 0.03 | 0.88 | 0.38 | -0.02 | -0.12 | 0.90 | 0.00 | -0.14 | 0.89 | -0.05 | -0.48 | 0.65 |
| 4 | 0.02 | 0.55 | 0.59 | -0.08 | -0.46 | 0.65 | 0.00 | -0.16 | 0.87 | -0.13 | -2.04 | 0.07 |
| 5 | 0.03 | 0.80 | 0.42 | 0.09 | 0.53 | 0.60 | 0.00 | -0.13 | 0.90 | -0.04 | -0.47 | 0.65 |
| 6 | 0.04 | 1.20 | 0.23 | -0.06 | -0.32 | 0.75 | 0.00 | -0.21 | 0.83 | -0.03 | -0.32 | 0.76 |
| 7 | 0.02 | 0.55 | 0.59 | 0.10 | 0.65 | 0.52 | -0.01 | -0.53 | 0.60 | -0.07 | -0.87 | 0.41 |
| 8 | 0.02 | 0.50 | 0.62 | 0.03 | 0.20 | 0.84 | -0.01 | -0.66 | 0.51 | 0.02 | 0.19 | 0.85 |
| 9 | 0.03 | 1.07 | 0.29 | -0.06 | -0.35 | 0.73 | -0.01 | -0.52 | 0.61 | -0.04 | -0.36 | 0.73 |
| 10 | 0.03 | 0.99 | 0.33 | 0.06 | 0.35 | 0.73 | 0.00 | 0.10 | 0.92 | -0.06 | -1.34 | 0.21 |
| 11 | 0.02 | 0.71 | 0.48 | 0.10 | 0.66 | 0.51 | -0.01 | -0.70 | 0.49 | -0.09 | -0.98 | 0.35 |
| 12 | 0.05 | 1.72 | 0.09 | 0.01 | 0.06 | 0.95 | 0.00 | -0.21 | 0.83 | -0.02 | -0.22 | 0.83 |
| 13 | 0.01 | 0.47 | 0.64 | -0.07 | -0.46 | 0.65 | -0.01 | -0.86 | 0.40 | -0.02 | -0.26 | 0.80 |
| 14 | 0.04 | 1.28 | 0.21 | -0.14 | -0.82 | 0.41 | 0.00 | -0.21 | 0.84 | -0.02 | -0.29 | 0.77 |
| 15 | -0.89 | 0.06 | 0.12 | -0.12 | -0.81 | 0.42 | -0.01 | -0.62 | 0.53 | 0.03 | 0.31 | 0.76 |
| 16 | 0.04 | 1.09 | 0.28 | 0.03 | 0.18 | 0.86 | 0.00 | 0.09 | 0.93 | -0.15 | -1.59 | 0.15 |
| 17 | 0.06 | 1.75 | 0.08 | -0.13 | -0.76 | 0.45 | 0.00 | -0.03 | 0.98 | 0.01 | 0.13 | 0.90 |
| 18 | 0.03 | 0.88 | 0.38 | 0.01 | 0.05 | 0.96 | -0.01 | -1.20 | 0.23 | 0.04 | 0.69 | 0.51 |
| 19 | 0.02 | 0.58 | 0.56 | 0.03 | 0.15 | 0.88 | 0.00 | 0.09 | 0.93 | -0.01 | -0.05 | 0.96 |
| 20 | 0.03 | 0.96 | 0.34 | 0.07 | 0.47 | 0.64 | 0.00 | -0.13 | 0.89 | -0.03 | -0.45 | 0.66 |
| 21 | 0.01 | 0.35 | 0.73 | 0.02 | 0.10 | 0.92 | 0.00 | -0.24 | 0.81 | -0.06 | -0.65 | 0.53 |
| 22 | 0.06 | 1.91 | 0.06 | -0.07 | -0.48 | 0.63 | 0.00 | -0.36 | 0.72 | 0.00 | 0.03 | 0.98 |
| 23 | 0.04 | 1.30 | 0.20 | 0.07 | 0.41 | 0.68 | 0.00 | 0.18 | 0.86 | -0.08 | -1.09 | 0.30 |
| 24 | 0.02 | 0.76 | 0.45 | -0.06 | -0.39 | 0.70 | 0.00 | 0.35 | 0.73 | -0.05 | -0.49 | 0.64 |
| 25 | -0.01 | -0.26 | 0.80 | 0.00 | -0.03 | 0.97 | -0.01 | -1.10 | 0.28 | -0.05 | -0.58 | 0.58 |
| 26 | 0.01 | 0.20 | 0.84 | 0.18 | 1.12 | 0.27 | 0.02 | 1.90 | 0.06 | 0.01 | 0.11 | 0.91 |
| 27 | 0.04 | 1.31 | 0.19 | -0.24 | -1.44 | 0.15 | 0.00 | 0.32 | 0.75 | 0.08 | 0.87 | 0.41 |
| 28 | 0.03 | 0.87 | 0.39 | -0.15 | -0.98 | 0.33 | 0.00 | -0.38 | 0.70 | -0.09 | -1.11 | 0.30 |
| 29 | 0.04 | 1.13 | 0.26 | -0.17 | -1.08 | 0.28 | 0.00 | -0.18 | 0.86 | -0.03 | -0.44 | 0.67 |
| 30 | -0.01 | -0.23 | 0.82 | 0.08 | 0.49 | 0.63 | -0.01 | -0.86 | 0.39 | -0.03 | -0.59 | 0.57 |

ตารางที่ 46 ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test B2 จากการประมาณค่า 100 ครั้ง (ต่อ)

| ครั้งที่ | a | | | b | | | c | | | alpha | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD | t | Sig | MD | t | sig | MD | t | sig | MD | t | sig |
| 31 | 0.03 | 0.96 | 0.34 | 0.13 | 0.76 | 0.45 | 0.01 | 0.94 | 0.35 | 0.00 | -0.01 | 0.99 |
| 32 | 0.02 | 0.61 | 0.55 | 0.02 | 0.16 | 0.87 | 0.00 | -0.24 | 0.81 | -0.07 | -1.12 | 0.29 |
| 33 | 0.03 | 0.91 | 0.36 | -0.07 | -0.44 | 0.66 | 0.00 | 0.27 | 0.79 | -0.04 | -0.47 | 0.65 |
| 34 | 0.05 | 1.32 | 0.19 | 0.04 | 0.21 | 0.83 | 0.01 | 0.66 | 0.51 | 0.02 | 0.29 | 0.78 |
| 35 | 0.06 | 1.71 | 0.09 | -0.07 | -0.42 | 0.67 | 0.00 | 0.22 | 0.83 | 0.05 | 0.45 | 0.66 |
| 36 | 0.06 | 1.61 | 0.11 | -0.11 | -0.66 | 0.51 | 0.00 | 0.20 | 0.84 | 0.02 | 0.21 | 0.84 |
| 37 | 0.03 | 1.16 | 0.25 | -0.03 | -0.19 | 0.85 | -0.01 | -1.13 | 0.26 | -0.11 | -1.71 | 0.12 |
| 38 | 0.06 | 1.74 | 0.09 | -0.09 | -0.52 | 0.61 | -0.01 | -0.76 | 0.45 | -0.02 | -0.17 | 0.87 |
| 39 | 0.03 | 0.74 | 0.46 | 0.17 | 1.00 | 0.32 | -0.01 | -0.86 | 0.39 | -0.09 | -1.19 | 0.26 |
| 40 | 0.03 | 1.03 | 0.30 | -0.06 | -0.33 | 0.74 | 0.00 | -0.36 | 0.72 | -0.08 | -0.69 | 0.51 |
| 41 | 0.01 | 0.36 | 0.72 | 0.00 | 0.00 | 1.00 | 0.00 | 0.45 | 0.65 | -0.04 | -0.51 | 0.62 |
| 42 | 0.07 | 2.21 | 0.03 | 0.12 | 0.77 | 0.44 | -0.01 | -0.70 | 0.48 | 0.06 | 0.65 | 0.53 |
| 43 | 0.05 | 1.58 | 0.12 | 0.14 | 0.91 | 0.37 | 0.02 | 1.68 | 0.10 | 0.00 | -0.07 | 0.95 |
| 44 | 0.02 | 0.69 | 0.49 | 0.06 | 0.36 | 0.72 | 0.01 | 0.75 | 0.45 | 0.02 | 0.21 | 0.84 |
| 45 | 0.06 | 1.61 | 0.11 | -0.24 | -1.37 | 0.17 | 0.00 | 0.20 | 0.84 | -0.04 | -0.56 | 0.59 |
| 46 | 0.05 | 1.68 | 0.10 | 0.11 | 0.63 | 0.53 | 0.01 | 0.73 | 0.47 | -0.09 | -1.99 | 0.08 |
| 47 | 0.00 | 0.09 | 0.93 | 0.13 | 0.80 | 0.42 | 0.01 | 1.02 | 0.31 | 0.02 | 0.22 | 0.83 |
| 48 | -0.01 | -0.17 | 0.87 | -0.05 | -0.26 | 0.80 | -0.01 | -1.15 | 0.25 | 0.03 | 0.29 | 0.78 |
| 49 | 0.01 | 0.01 | 0.89 | 0.00 | -0.01 | 0.99 | 0.00 | 0.33 | 0.74 | 0.04 | 0.63 | 0.55 |
| 50 | 0.00 | 0.09 | 0.93 | -0.04 | -0.21 | 0.83 | -0.01 | -0.55 | 0.59 | -0.03 | -0.46 | 0.66 |
| 51 | 0.04 | 1.35 | 0.18 | 0.18 | 1.11 | 0.27 | 0.00 | -0.23 | 0.82 | 0.05 | 0.55 | 0.59 |
| 52 | 0.00 | 0.11 | 0.91 | -0.08 | -0.49 | 0.63 | 0.00 | 0.50 | 0.62 | -0.03 | -0.30 | 0.77 |
| 53 | 0.04 | 1.25 | 0.22 | 0.07 | 0.44 | 0.66 | 0.00 | 0.10 | 0.92 | -0.10 | -1.24 | 0.25 |
| 54 | 0.06 | 1.54 | 0.13 | -0.09 | -0.47 | 0.64 | 0.00 | -0.06 | 0.95 | -0.07 | -0.76 | 0.46 |
| 55 | 0.01 | 0.18 | 0.85 | -0.06 | -0.41 | 0.69 | 0.00 | 0.21 | 0.83 | -0.04 | -0.67 | 0.52 |
| 56 | 0.01 | 0.44 | 0.66 | -0.16 | -1.01 | 0.31 | 0.00 | 0.25 | 0.80 | -0.05 | -0.76 | 0.47 |
| 57 | -0.01 | -0.16 | 0.88 | 0.01 | 0.08 | 0.93 | 0.00 | -0.10 | 0.92 | -0.10 | -1.66 | 0.13 |
| 58 | 0.04 | 1.06 | 0.29 | -0.05 | -0.29 | 0.77 | 0.00 | 0.05 | 0.96 | -0.07 | -0.90 | 0.39 |
| 59 | 0.02 | 0.54 | 0.59 | -0.03 | -0.20 | 0.85 | 0.02 | 2.06 | 0.04 | 0.03 | 0.54 | 0.60 |
| 60 | 0.00 | -0.02 | 0.99 | -0.18 | -1.01 | 0.32 | 0.00 | 0.06 | 0.95 | -0.03 | -0.28 | 0.78 |

ตารางที่ 46 ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test B2 จากการประมาณค่า 100 ครั้ง (ต่อ)

| ครั้งที่ | a | | | b | | | c | | | alpha | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD | t | Sig | MD | t | sig | MD | t | sig | MD | t | sig |
| 61 | 0.06 | 1.98 | 0.05 | -0.13 | -0.90 | 0.37 | 0.00 | 0.38 | 0.70 | 0.00 | -0.03 | 0.98 |
| 62 | 0.05 | 1.60 | 0.11 | 0.02 | 0.14 | 0.89 | -0.01 | -0.59 | 0.56 | -0.13 | -2.51 | 0.03 |
| 63 | 0.02 | 0.67 | 0.51 | -0.08 | -0.47 | 0.64 | 0.00 | 0.48 | 0.63 | 0.01 | 0.12 | 0.91 |
| 64 | 0.03 | 0.91 | 0.36 | -0.17 | -1.18 | 0.24 | 0.00 | -0.01 | 0.99 | 0.01 | 0.08 | 0.94 |
| 65 | 0.02 | 0.63 | 0.53 | 0.03 | 0.17 | 0.87 | 0.00 | 0.19 | 0.85 | -0.12 | -1.70 | 0.12 |
| 66 | 0.02 | 0.46 | 0.64 | -0.05 | -0.31 | 0.76 | 0.00 | 0.11 | 0.91 | 0.03 | 0.37 | 0.72 |
| 67 | 0.04 | 1.30 | 0.20 | -0.07 | -0.49 | 0.62 | 0.01 | 1.26 | 0.21 | 0.02 | 0.32 | 0.76 |
| 68 | 0.05 | 1.62 | 0.11 | -0.23 | -1.36 | 0.18 | 0.00 | 0.20 | 0.84 | 0.09 | 0.91 | 0.39 |
| 69 | 0.00 | 0.10 | 0.92 | -0.14 | -0.79 | 0.43 | 0.01 | 0.98 | 0.33 | -0.02 | -0.28 | 0.78 |
| 70 | 0.06 | 1.98 | 0.05 | 0.02 | 0.10 | 0.92 | 0.00 | -0.31 | 0.76 | 0.01 | 0.07 | 0.95 |
| 71 | -0.02 | -0.59 | 0.56 | -0.17 | -0.98 | 0.33 | -0.01 | -0.52 | 0.60 | -0.02 | -0.24 | 0.82 |
| 72 | 0.01 | 0.34 | 0.73 | 0.01 | 0.05 | 0.96 | 0.01 | 0.58 | 0.56 | -0.05 | -0.38 | 0.71 |
| 73 | 0.06 | 1.95 | 0.06 | -0.07 | -0.42 | 0.68 | 0.00 | -0.06 | 0.95 | 0.03 | 0.22 | 0.83 |
| 74 | 0.01 | 0.26 | 0.79 | 0.10 | 0.67 | 0.51 | 0.00 | 0.14 | 0.89 | -0.09 | -1.01 | 0.34 |
| 75 | 0.03 | 0.84 | 0.40 | -0.13 | -0.96 | 0.34 | 0.00 | 0.20 | 0.84 | -0.08 | -1.58 | 0.15 |
| 76 | 0.05 | 1.54 | 0.13 | 0.12 | 0.76 | 0.45 | 0.01 | 0.81 | 0.42 | -0.12 | -1.67 | 0.13 |
| 77 | 0.02 | 0.64 | 0.52 | -0.16 | -0.93 | 0.36 | 0.01 | 0.66 | 0.51 | -0.06 | -0.58 | 0.58 |
| 78 | 0.04 | 1.08 | 0.29 | -0.06 | -0.41 | 0.68 | 0.00 | 0.19 | 0.85 | -0.07 | -0.78 | 0.46 |
| 79 | 0.06 | 1.59 | 0.12 | 0.18 | 1.20 | 0.23 | 0.01 | 0.80 | 0.43 | -0.06 | -0.86 | 0.41 |
| 80 | 0.00 | -0.13 | 0.90 | -0.13 | -0.85 | 0.40 | 0.00 | -0.28 | 0.78 | -0.07 | -1.13 | 0.29 |
| 81 | 0.04 | 1.22 | 0.23 | 0.00 | -0.02 | 0.98 | 0.00 | 0.34 | 0.74 | 0.08 | 1.00 | 0.34 |
| 82 | 0.04 | 1.15 | 0.25 | -0.02 | -0.12 | 0.91 | 0.00 | -0.27 | 0.79 | 0.06 | 0.67 | 0.52 |
| 83 | 0.04 | 1.34 | 0.18 | -0.01 | -0.09 | 0.93 | -0.01 | -0.69 | 0.49 | -0.03 | -0.34 | 0.74 |
| 84 | 0.03 | 0.97 | 0.33 | -0.06 | -0.36 | 0.72 | 0.01 | 1.53 | 0.13 | -0.03 | -0.43 | 0.67 |
| 85 | 0.03 | 0.90 | 0.37 | -0.08 | -0.52 | 0.60 | 0.01 | 0.86 | 0.39 | -0.04 | -0.60 | 0.56 |
| 86 | 0.02 | 0.69 | 0.49 | -0.04 | -0.23 | 0.82 | 0.00 | 0.01 | 0.99 | -0.04 | -0.48 | 0.64 |
| 87 | -0.01 | -0.43 | 0.67 | 0.11 | 0.73 | 0.47 | 0.00 | -0.22 | 0.82 | 0.01 | 0.12 | 0.91 |
| 88 | 0.05 | 1.74 | 0.09 | 0.14 | 0.92 | 0.36 | 0.00 | 0.32 | 0.75 | -0.05 | -1.02 | 0.33 |
| 89 | -0.02 | -0.50 | 0.62 | 0.02 | 0.15 | 0.88 | -0.01 | -1.05 | 0.30 | -0.03 | -0.36 | 0.73 |
| 90 | -0.02 | -0.46 | 0.64 | -0.21 | -1.24 | 0.22 | 0.00 | -0.15 | 0.88 | -0.06 | -0.79 | 0.45 |

ตารางที่ 46 ผลการเปรียบเทียบพารามิเตอร์จากโปรแกรม WINGEN และ MULTILOG สำหรับ Test B2 จากการประมาณค่า 100 ครั้ง (ต่อ)

| ครั้งที่ | a | | | b | | | c | | | alpha | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD | t | Sig | MD | t | sig | MD | t | sig | MD | t | sig |
| 91 | 0.01 | 0.23 | 0.82 | -0.09 | -0.52 | 0.60 | 0.00 | 0.25 | 0.80 | -0.01 | -0.08 | 0.94 |
| 92 | -0.01 | -0.42 | 0.68 | 0.01 | 0.08 | 0.94 | 0.01 | 0.77 | 0.44 | -0.15 | -2.29 | 0.05 |
| 93 | 0.05 | 1.69 | 0.09 | 0.03 | 0.17 | 0.87 | 0.00 | 0.31 | 0.76 | -0.03 | -0.24 | 0.81 |
| 94 | 0.03 | 0.97 | 0.33 | -0.09 | -0.54 | 0.59 | 0.00 | 0.22 | 0.83 | 0.02 | 0.23 | 0.82 |
| 95 | 0.04 | 1.29 | 0.20 | 0.10 | 0.57 | 0.57 | 0.00 | 0.46 | 0.65 | 0.03 | 0.29 | 0.78 |
| 96 | 0.03 | 1.03 | 0.31 | -0.07 | -0.41 | 0.68 | 0.00 | -0.17 | 0.87 | -0.02 | -0.35 | 0.73 |
| 97 | 0.05 | 1.53 | 0.13 | -0.08 | -0.50 | 0.62 | 0.00 | 0.45 | 0.65 | -0.03 | -0.41 | 0.69 |
| 98 | 0.06 | 1.87 | 0.07 | 0.05 | 0.34 | 0.74 | 0.01 | 0.62 | 0.54 | -0.05 | -0.49 | 0.63 |
| 99 | 0.03 | 0.84 | 0.41 | -0.02 | -0.11 | 0.91 | 0.00 | -0.24 | 0.81 | 0.10 | 1.07 | 0.31 |
| 100 | 0.02 | 0.76 | 0.45 | -0.22 | -1.31 | 0.19 | 0.00 | 0.21 | 0.84 | -0.03 | -0.45 | 0.66 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

**ผลการตรวจสอบความเป็นเอกมิติของข้อมูลโดยการวิเคราะห์องค์ประกอบ**
**จำแนกตามเงื่อนไขสัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่า**
**และข้อสอบที่ตรวจให้คะแนนหลายค่าในแบบสอบ**
**จำนวน 4 เงื่อนไข**

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 47 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 50 : 10 ด้วยการทำซ้ำ 100 ครั้ง

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 1 | 1 | 19.585 | 29.309 | 20.010 | 29.267 | 19.921 | 29.348 |
| | 2 | 4.178 | 5.473 | 4.431 | 5.748 | 3.812 | 4.991 |
| 2 | 1 | 19.476 | 30.033 | 19.610 | 29.798 | 19.969 | 29.298 |
| | 2 | 5.371 | 6.888 | 5.088 | 6.544 | 3.236 | 4.299 |
| 3 | 1 | 19.252 | 28.862 | 18.913 | 29.448 | 19.201 | 28.897 |
| | 2 | 4.289 | 5.597 | 4.352 | 5.652 | 3.606 | 4.694 |
| 4 | 1 | 19.517 | 30.082 | 19.359 | 29.083 | 20.003 | 29.238 |
| | 2 | 3.754 | 4.886 | 3.985 | 5.231 | 5.124 | 6.568 |
| 5 | 1 | 19.536 | 29.577 | 19.909 | 29.086 | 19.057 | 28.689 |
| | 2 | 2.982 | 4.005 | 4.252 | 5.566 | 4.784 | 6.214 |
| 6 | 1 | 19.591 | 29.574 | 19.103 | 29.029 | 20.069 | 28.815 |
| | 2 | 4.517 | 5.849 | 3.213 | 4.268 | 3.325 | 4.370 |
| 7 | 1 | 19.333 | 29.604 | 19.236 | 29.644 | 20.119 | 30.079 |
| | 2 | 5.323 | 6.824 | 4.950 | 6.370 | 5.299 | 6.786 |
| 8 | 1 | 19.646 | 30.151 | 19.673 | 28.785 | 19.196 | 30.113 |
| | 2 | 3.297 | 4.337 | 5.010 | 6.427 | 5.131 | 6.589 |
| 9 | 1 | 19.678 | 29.319 | 19.401 | 29.508 | 19.918 | 29.295 |
| | 2 | 4.808 | 6.252 | 3.700 | 4.854 | 3.059 | 4.077 |
| 10 | 1 | 19.223 | 29.290 | 19.659 | 29.581 | 19.858 | 28.971 |
| | 2 | 3.486 | 4.570 | 3.113 | 4.135 | 3.929 | 5.159 |
| 11 | 1 | 19.282 | 29.203 | 19.488 | 29.381 | 19.235 | 28.922 |
| | 2 | 3.471 | 4.552 | 3.418 | 4.472 | 4.936 | 6.357 |
| 12 | 1 | 19.641 | 29.080 | 20.084 | 29.018 | 20.137 | 29.612 |
| | 2 | 3.684 | 4.829 | 5.414 | 6.965 | 5.102 | 6.560 |
| 13 | 1 | 19.181 | 30.098 | 18.955 | 29.277 | 19.928 | 30.240 |
| | 2 | 4.315 | 5.623 | 5.384 | 6.917 | 4.559 | 5.904 |
| 14 | 1 | 19.951 | 29.828 | 18.960 | 29.280 | 19.571 | 29.827 |
| | 2 | 3.611 | 4.699 | 2.929 | 3.969 | 3.581 | 4.655 |
| 15 | 1 | 19.716 | 29.035 | 20.110 | 28.841 | 19.885 | 30.177 |
| | 2 | 5.219 | 6.686 | 2.918 | 3.960 | 4.492 | 5.811 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 47 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 50 : 10 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 16 | 1 | 20.039 | 28.877 | 19.018 | 28.927 | 19.156 | 29.278 |
| | 2 | 3.041 | 4.058 | 2.973 | 3.998 | 3.795 | 4.971 |
| 17 | 1 | 19.802 | 30.017 | 20.084 | 30.247 | 18.972 | 29.405 |
| | 2 | 3.408 | 4.452 | 4.516 | 5.849 | 4.134 | 5.414 |
| 18 | 1 | 19.551 | 29.327 | 19.474 | 29.506 | 19.960 | 30.191 |
| | 2 | 5.150 | 6.614 | 4.926 | 6.351 | 3.049 | 4.061 |
| 19 | 1 | 19.481 | 29.949 | 19.054 | 28.815 | 19.889 | 29.464 |
| | 2 | 4.793 | 6.237 | 4.556 | 5.891 | 4.410 | 5.722 |
| 20 | 1 | 20.102 | 29.908 | 19.735 | 29.065 | 19.913 | 29.067 |
| | 2 | 5.301 | 6.787 | 4.214 | 5.517 | 4.560 | 5.905 |
| 21 | 1 | 19.235 | 28.881 | 19.018 | 30.061 | 19.997 | 29.440 |
| | 2 | 4.660 | 6.071 | 3.630 | 4.733 | 3.514 | 4.608 |
| 22 | 1 | 20.090 | 28.879 | 19.742 | 29.786 | 19.192 | 29.059 |
| | 2 | 3.774 | 4.923 | 4.890 | 6.318 | 3.112 | 4.132 |
| 23 | 1 | 19.053 | 28.774 | 19.464 | 29.212 | 19.559 | 29.902 |
| | 2 | 5.014 | 6.428 | 3.068 | 4.093 | 5.094 | 6.550 |
| 24 | 1 | 19.474 | 30.120 | 19.151 | 29.398 | 19.905 | 28.980 |
| | 2 | 4.096 | 5.355 | 3.790 | 4.952 | 3.115 | 4.137 |
| 25 | 1 | 19.278 | 28.922 | 19.166 | 29.721 | 19.389 | 29.639 |
| | 2 | 5.380 | 6.898 | 4.719 | 6.126 | 4.089 | 5.332 |
| 26 | 1 | 19.212 | 28.728 | 19.175 | 29.226 | 19.291 | 29.054 |
| | 2 | 3.194 | 4.221 | 5.394 | 6.942 | 2.988 | 4.014 |
| 27 | 1 | 18.988 | 29.709 | 19.232 | 29.591 | 19.498 | 30.160 |
| | 2 | 5.236 | 6.700 | 3.800 | 4.980 | 3.888 | 5.107 |
| 28 | 1 | 20.190 | 29.495 | 19.385 | 29.338 | 19.714 | 29.491 |
| | 2 | 3.463 | 4.534 | 3.251 | 4.299 | 4.367 | 5.674 |
| 29 | 1 | 19.852 | 29.496 | 20.178 | 28.683 | 20.096 | 29.483 |
| | 2 | 5.243 | 6.701 | 4.593 | 5.958 | 2.916 | 3.956 |
| 30 | 1 | 20.210 | 29.704 | 19.230 | 29.962 | 19.732 | 29.009 |
| | 2 | 5.198 | 6.660 | 3.902 | 5.124 | 4.571 | 5.922 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 47 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 50 : 10 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 31 | 1 | 20.033 | 28.752 | 19.341 | 30.093 | 20.041 | 29.008 |
| | 2 | 2.839 | 3.894 | 4.196 | 5.493 | 3.017 | 4.035 |
| 32 | 1 | 19.765 | 30.071 | 19.877 | 29.772 | 19.404 | 29.158 |
| | 2 | 3.032 | 4.043 | 4.618 | 5.977 | 3.736 | 4.859 |
| 33 | 1 | 19.916 | 30.221 | 19.207 | 30.015 | 20.033 | 29.377 |
| | 2 | 4.225 | 5.529 | 3.799 | 4.975 | 4.857 | 6.289 |
| 34 | 1 | 20.021 | 30.231 | 19.642 | 29.731 | 19.277 | 29.830 |
| | 2 | 2.914 | 3.952 | 4.750 | 6.167 | 4.403 | 5.711 |
| 35 | 1 | 19.597 | 29.126 | 19.355 | 29.549 | 19.713 | 28.663 |
| | 2 | 3.112 | 4.132 | 4.323 | 5.629 | 5.274 | 6.734 |
| 36 | 1 | 19.133 | 29.932 | 19.054 | 29.070 | 19.199 | 29.074 |
| | 2 | 4.525 | 5.856 | 4.624 | 6.007 | 5.270 | 6.730 |
| 37 | 1 | 20.051 | 29.702 | 19.077 | 29.857 | 19.266 | 30.165 |
| | 2 | 3.781 | 4.936 | 3.268 | 4.318 | 4.393 | 5.707 |
| 38 | 1 | 19.955 | 30.172 | 18.935 | 29.350 | 19.455 | 29.484 |
| | 2 | 3.116 | 4.139 | 3.367 | 4.402 | 4.360 | 5.662 |
| 39 | 1 | 19.108 | 29.727 | 19.827 | 28.834 | 19.285 | 28.881 |
| | 2 | 3.965 | 5.205 | 5.138 | 6.597 | 5.184 | 6.651 |
| 40 | 1 | 19.038 | 28.869 | 18.963 | 29.144 | 19.178 | 29.673 |
| | 2 | 3.402 | 4.441 | 5.075 | 6.536 | 4.250 | 5.566 |
| 41 | 1 | 20.085 | 29.472 | 19.660 | 29.819 | 19.625 | 29.838 |
| | 2 | 3.489 | 4.576 | 4.195 | 5.490 | 4.215 | 5.524 |
| 42 | 1 | 19.618 | 30.257 | 19.223 | 29.316 | 19.143 | 29.671 |
| | 2 | 5.289 | 6.761 | 4.719 | 6.136 | 4.924 | 6.347 |
| 43 | 1 | 19.829 | 30.080 | 19.767 | 29.776 | 19.570 | 29.493 |
| | 2 | 4.520 | 5.852 | 4.819 | 6.259 | 4.054 | 5.277 |
| 44 | 1 | 19.394 | 29.609 | 19.690 | 29.031 | 19.287 | 29.551 |
| | 2 | 4.395 | 5.707 | 3.212 | 4.266 | 4.023 | 5.246 |
| 45 | 1 | 19.696 | 28.691 | 20.119 | 28.863 | 19.368 | 28.691 |
| | 2 | 3.518 | 4.612 | 3.402 | 4.450 | 3.486 | 4.570 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 47 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 50 : 10 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 46 | 1 | 19.076 | 28.680 | 19.479 | 30.033 | 20.027 | 30.011 |
| | 2 | 4.577 | 5.943 | 3.657 | 4.773 | 4.245 | 5.550 |
| 47 | 1 | 19.009 | 29.262 | 19.710 | 29.653 | 19.550 | 30.130 |
| | 2 | 4.021 | 5.244 | 3.481 | 4.560 | 5.007 | 6.422 |
| 48 | 1 | 18.982 | 29.796 | 19.384 | 30.155 | 19.701 | 29.508 |
| | 2 | 4.118 | 5.390 | 3.367 | 4.406 | 5.060 | 6.527 |
| 49 | 1 | 19.333 | 28.898 | 20.162 | 30.062 | 20.039 | 30.051 |
| | 2 | 5.146 | 6.611 | 4.832 | 6.276 | 4.821 | 6.262 |
| 50 | 1 | 20.026 | 29.379 | 19.665 | 29.222 | 19.203 | 29.075 |
| | 2 | 3.110 | 4.131 | 4.648 | 6.046 | 3.030 | 4.042 |
| 51 | 1 | 19.542 | 30.046 | 19.843 | 28.966 | 20.101 | 28.741 |
| | 2 | 3.977 | 5.216 | 3.379 | 4.431 | 3.504 | 4.600 |
| 52 | 1 | 19.418 | 28.853 | 19.874 | 29.107 | 20.174 | 28.961 |
| | 2 | 5.312 | 6.809 | 3.452 | 4.523 | 5.004 | 6.415 |
| 53 | 1 | 20.156 | 29.505 | 18.998 | 30.064 | 19.155 | 30.140 |
| | 2 | 4.188 | 5.485 | 3.296 | 4.333 | 3.840 | 5.037 |
| 54 | 1 | 19.398 | 29.787 | 19.941 | 29.390 | 19.258 | 29.038 |
| | 2 | 3.924 | 5.146 | 3.414 | 4.459 | 4.902 | 6.334 |
| 55 | 1 | 20.183 | 28.971 | 19.776 | 29.951 | 20.147 | 30.168 |
| | 2 | 3.195 | 4.229 | 4.595 | 5.959 | 4.327 | 5.635 |
| 56 | 1 | 18.918 | 28.817 | 19.443 | 30.141 | 19.738 | 29.180 |
| | 2 | 3.325 | 4.367 | 5.325 | 6.834 | 4.514 | 5.844 |
| 57 | 1 | 19.035 | 30.144 | 19.878 | 29.885 | 19.047 | 29.270 |
| | 2 | 3.139 | 4.169 | 5.002 | 6.411 | 5.180 | 6.650 |
| 58 | 1 | 19.089 | 29.908 | 20.180 | 29.606 | 20.205 | 29.470 |
| | 2 | 2.924 | 3.966 | 5.375 | 6.894 | 4.502 | 5.835 |
| 59 | 1 | 19.276 | 29.343 | 19.966 | 29.060 | 19.724 | 29.368 |
| | 2 | 3.792 | 4.953 | 3.943 | 5.189 | 3.064 | 4.081 |
| 60 | 1 | 19.565 | 29.126 | 19.110 | 29.110 | 19.357 | 28.890 |
| | 2 | 4.231 | 5.536 | 5.145 | 6.608 | 3.435 | 4.498 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 47 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 50 : 10 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 61 | 1 | 19.688 | 30.180 | 19.172 | 29.999 | 20.201 | 28.950 |
| | 2 | 3.306 | 4.354 | 4.845 | 6.281 | 3.012 | 4.031 |
| 62 | 1 | 18.901 | 28.923 | 19.833 | 28.948 | 19.677 | 29.567 |
| | 2 | 4.741 | 6.157 | 3.415 | 4.464 | 4.629 | 6.019 |
| 63 | 1 | 20.130 | 28.884 | 19.512 | 29.164 | 19.424 | 28.847 |
| | 2 | 4.942 | 6.363 | 4.857 | 6.286 | 4.381 | 5.696 |
| 64 | 1 | 19.094 | 29.869 | 20.015 | 29.983 | 19.640 | 29.933 |
| | 2 | 4.857 | 6.284 | 5.039 | 6.474 | 5.356 | 6.874 |
| 65 | 1 | 19.109 | 30.253 | 19.476 | 29.060 | 20.211 | 29.394 |
| | 2 | 5.332 | 6.848 | 3.067 | 4.091 | 3.512 | 4.606 |
| 66 | 1 | 19.036 | 30.251 | 19.426 | 28.739 | 20.032 | 29.190 |
| | 2 | 3.087 | 4.100 | 5.358 | 6.875 | 3.056 | 4.067 |
| 67 | 1 | 19.590 | 28.715 | 19.554 | 29.611 | 20.172 | 29.125 |
| | 2 | 5.018 | 6.432 | 4.682 | 6.099 | 4.799 | 6.241 |
| 68 | 1 | 20.136 | 28.992 | 19.811 | 29.582 | 20.205 | 28.836 |
| | 2 | 4.442 | 5.761 | 4.742 | 6.158 | 4.742 | 6.160 |
| 69 | 1 | 19.371 | 29.380 | 19.342 | 29.457 | 19.650 | 29.044 |
| | 2 | 5.205 | 6.664 | 4.493 | 5.821 | 4.084 | 5.327 |
| 70 | 1 | 18.960 | 29.290 | 19.759 | 28.752 | 19.963 | 28.686 |
| | 2 | 3.064 | 4.080 | 5.284 | 6.751 | 3.215 | 4.275 |
| 71 | 1 | 19.338 | 29.859 | 19.088 | 29.574 | 19.371 | 29.265 |
| | 2 | 4.368 | 5.681 | 3.265 | 4.314 | 4.510 | 5.843 |
| 72 | 1 | 20.027 | 29.569 | 19.937 | 30.083 | 19.367 | 29.719 |
| | 2 | 4.035 | 5.267 | 4.226 | 5.530 | 3.110 | 4.131 |
| 73 | 1 | 19.984 | 28.908 | 19.147 | 28.792 | 19.039 | 29.325 |
| | 2 | 3.815 | 4.993 | 2.916 | 3.956 | 4.212 | 5.512 |
| 74 | 1 | 19.235 | 28.945 | 20.115 | 30.252 | 19.637 | 29.111 |
| | 2 | 5.387 | 6.922 | 4.548 | 5.886 | 4.878 | 6.302 |
| 75 | 1 | 19.086 | 30.143 | 19.031 | 30.193 | 19.244 | 29.413 |
| | 2 | 3.587 | 4.665 | 4.451 | 5.778 | 4.357 | 5.658 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 47 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 50 : 10 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 76 | 1 | 19.391 | 29.762 | 20.039 | 29.379 | 19.433 | 30.233 |
| | 2 | 3.950 | 5.193 | 4.382 | 5.697 | 5.196 | 6.657 |
| 77 | 1 | 19.811 | 28.813 | 19.382 | 29.456 | 20.107 | 30.142 |
| | 2 | 4.254 | 5.566 | 4.859 | 6.293 | 4.047 | 5.275 |
| 78 | 1 | 19.733 | 29.788 | 19.435 | 29.780 | 19.555 | 29.647 |
| | 2 | 4.169 | 5.469 | 5.040 | 6.482 | 3.233 | 4.294 |
| 79 | 1 | 19.304 | 29.562 | 20.009 | 29.060 | 19.691 | 28.912 |
| | 2 | 5.008 | 6.422 | 5.310 | 6.807 | 3.449 | 4.509 |
| 80 | 1 | 19.202 | 29.290 | 19.304 | 29.562 | 19.145 | 28.829 |
| | 2 | 4.452 | 5.783 | 3.189 | 4.213 | 2.859 | 3.920 |
| 81 | 1 | 19.311 | 29.238 | 19.273 | 29.305 | 19.463 | 28.999 |
| | 2 | 4.594 | 5.958 | 4.928 | 6.352 | 3.643 | 4.748 |
| 82 | 1 | 19.486 | 28.676 | 19.135 | 29.446 | 19.749 | 30.233 |
| | 2 | 3.419 | 4.476 | 3.097 | 4.106 | 4.088 | 5.329 |
| 83 | 1 | 19.024 | 29.593 | 19.172 | 29.604 | 19.374 | 29.645 |
| | 2 | 4.509 | 5.841 | 4.084 | 5.322 | 5.175 | 6.649 |
| 84 | 1 | 19.482 | 29.852 | 19.447 | 29.698 | 19.628 | 29.209 |
| | 2 | 4.509 | 5.842 | 2.856 | 3.918 | 3.526 | 4.617 |
| 85 | 1 | 19.162 | 29.206 | 19.682 | 28.822 | 19.746 | 30.180 |
| | 2 | 4.220 | 5.528 | 3.115 | 4.136 | 5.418 | 6.972 |
| 86 | 1 | 19.509 | 29.040 | 19.670 | 29.910 | 19.312 | 29.790 |
| | 2 | 3.206 | 4.260 | 4.710 | 6.125 | 3.585 | 4.663 |
| 87 | 1 | 19.236 | 29.706 | 19.190 | 29.951 | 19.781 | 30.129 |
| | 2 | 4.629 | 6.017 | 3.617 | 4.717 | 3.202 | 4.246 |
| 88 | 1 | 19.650 | 29.521 | 20.122 | 28.673 | 20.129 | 29.085 |
| | 2 | 3.768 | 4.911 | 4.694 | 6.114 | 5.050 | 6.510 |
| 89 | 1 | 19.209 | 28.964 | 19.757 | 29.037 | 19.998 | 30.186 |
| | 2 | 3.066 | 4.084 | 4.682 | 6.099 | 4.345 | 5.651 |
| 90 | 1 | 19.362 | 29.036 | 20.063 | 30.204 | 20.206 | 29.658 |
| | 2 | 4.538 | 5.877 | 5.233 | 6.696 | 4.692 | 6.102 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 47 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 50 : 10 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 91 | 1 | 20.025 | 29.093 | 19.128 | 29.885 | 19.062 | 29.661 |
| | 2 | 5.086 | 6.541 | 5.101 | 6.556 | 4.273 | 5.581 |
| 92 | 1 | 19.205 | 28.808 | 19.929 | 30.047 | 19.357 | 30.266 |
| | 2 | 5.174 | 6.648 | 4.643 | 6.041 | 4.796 | 6.237 |
| 93 | 1 | 19.061 | 29.501 | 19.169 | 29.831 | 19.565 | 29.330 |
| | 2 | 4.072 | 5.301 | 5.188 | 6.652 | 4.623 | 6.000 |
| 94 | 1 | 19.311 | 29.418 | 19.071 | 29.567 | 19.540 | 29.400 |
| | 2 | 4.564 | 5.916 | 3.298 | 4.339 | 3.374 | 4.418 |
| 95 | 1 | 19.355 | 28.811 | 20.155 | 30.176 | 19.387 | 29.652 |
| | 2 | 4.267 | 5.574 | 5.051 | 6.513 | 5.166 | 6.635 |
| 96 | 1 | 20.052 | 29.071 | 19.068 | 29.604 | 19.786 | 29.824 |
| | 2 | 5.333 | 6.848 | 4.928 | 6.354 | 4.196 | 5.493 |
| 97 | 1 | 19.018 | 29.254 | 19.374 | 28.664 | 19.674 | 29.248 |
| | 2 | 3.384 | 4.433 | 4.574 | 5.929 | 3.409 | 4.458 |
| 98 | 1 | 20.187 | 28.937 | 19.332 | 29.689 | 19.307 | 29.443 |
| | 2 | 5.326 | 6.836 | 5.056 | 6.524 | 3.796 | 4.973 |
| 99 | 1 | 20.095 | 29.490 | 19.714 | 29.078 | 20.124 | 28.924 |
| | 2 | 3.953 | 5.198 | 3.251 | 4.300 | 3.876 | 5.087 |
| 100 | 1 | 19.931 | 29.727 | 19.694 | 29.058 | 19.106 | 29.380 |
| | 2 | 3.586 | 4.663 | 3.128 | 4.152 | 3.623 | 4.725 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 48 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 60 : 8 ด้วยการทำซ้ำ 100 ครั้ง

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 1 | 1 | 19.442 | 28.615 | 18.902 | 27.368 | 19.611 | 28.947 |
| | 2 | 3.848 | 5.038 | 3.831 | 5.022 | 4.061 | 5.280 |
| 2 | 1 | 19.307 | 29.069 | 19.746 | 28.750 | 19.489 | 28.705 |
| | 2 | 3.887 | 5.106 | 3.024 | 4.036 | 4.658 | 6.062 |
| 3 | 1 | 19.249 | 28.002 | 19.172 | 28.019 | 19.934 | 28.162 |
| | 2 | 4.540 | 5.882 | 5.353 | 6.873 | 4.455 | 5.787 |
| 4 | 1 | 20.074 | 27.276 | 19.546 | 28.591 | 18.939 | 28.624 |
| | 2 | 3.989 | 5.232 | 4.936 | 6.363 | 4.006 | 5.238 |
| 5 | 1 | 19.187 | 28.970 | 19.420 | 27.243 | 19.174 | 28.190 |
| | 2 | 3.192 | 4.220 | 3.744 | 4.874 | 5.136 | 6.597 |
| 6 | 1 | 19.635 | 29.111 | 19.364 | 28.114 | 20.042 | 28.457 |
| | 2 | 5.028 | 6.462 | 4.162 | 5.461 | 4.819 | 6.261 |
| 7 | 1 | 19.909 | 28.930 | 19.767 | 27.618 | 19.560 | 27.240 |
| | 2 | 4.247 | 5.552 | 5.122 | 6.567 | 4.786 | 6.225 |
| 8 | 1 | 20.076 | 27.654 | 19.752 | 27.498 | 19.172 | 29.100 |
| | 2 | 3.861 | 5.064 | 4.471 | 5.797 | 3.806 | 4.987 |
| 9 | 1 | 20.149 | 28.557 | 19.757 | 28.187 | 20.154 | 28.289 |
| | 2 | 4.653 | 6.052 | 4.089 | 5.337 | 4.014 | 5.242 |
| 10 | 1 | 19.446 | 28.671 | 19.778 | 28.894 | 18.986 | 27.952 |
| | 2 | 3.215 | 4.277 | 4.333 | 5.647 | 3.674 | 4.818 |
| 11 | 1 | 19.307 | 28.000 | 19.849 | 28.529 | 20.043 | 27.918 |
| | 2 | 5.205 | 6.672 | 4.212 | 5.516 | 4.323 | 5.632 |
| 12 | 1 | 19.137 | 27.964 | 19.810 | 29.093 | 19.515 | 28.124 |
| | 2 | 4.753 | 6.170 | 3.878 | 5.088 | 4.621 | 5.984 |
| 13 | 1 | 19.679 | 27.779 | 18.994 | 28.387 | 19.347 | 28.530 |
| | 2 | 4.239 | 5.544 | 3.581 | 4.656 | 3.929 | 5.162 |
| 14 | 1 | 19.192 | 29.091 | 20.204 | 27.418 | 19.752 | 28.003 |
| | 2 | 3.487 | 4.571 | 3.043 | 4.059 | 5.240 | 6.700 |
| 15 | 1 | 20.118 | 29.081 | 19.224 | 28.292 | 19.080 | 27.874 |
| | 2 | 5.363 | 6.875 | 4.265 | 5.572 | 4.089 | 5.338 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 48 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 60 : 8 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 16 | 1 | 19.521 | 27.894 | 20.060 | 28.734 | 20.043 | 28.410 |
| | 2 | 4.584 | 5.954 | 2.905 | 3.942 | 2.906 | 3.943 |
| 17 | 1 | 19.628 | 28.664 | 19.497 | 27.326 | 19.905 | 27.695 |
| | 2 | 4.474 | 5.799 | 4.422 | 5.746 | 4.061 | 5.288 |
| 18 | 1 | 19.003 | 28.777 | 19.222 | 27.794 | 19.890 | 28.980 |
| | 2 | 4.240 | 5.547 | 4.813 | 6.258 | 5.034 | 6.471 |
| 19 | 1 | 19.399 | 27.362 | 19.593 | 28.307 | 20.154 | 28.068 |
| | 2 | 4.530 | 5.874 | 4.789 | 6.227 | 4.783 | 6.205 |
| 20 | 1 | 19.270 | 27.916 | 19.102 | 27.690 | 19.283 | 28.657 |
| | 2 | 3.446 | 4.504 | 5.127 | 6.578 | 3.822 | 5.017 |
| 21 | 1 | 20.026 | 28.949 | 19.611 | 28.112 | 19.080 | 28.838 |
| | 2 | 3.953 | 5.199 | 5.089 | 6.545 | 3.291 | 4.331 |
| 22 | 1 | 19.202 | 27.360 | 19.580 | 29.070 | 20.144 | 28.162 |
| | 2 | 5.337 | 6.862 | 4.697 | 6.115 | 4.637 | 6.035 |
| 23 | 1 | 18.903 | 27.959 | 19.385 | 27.665 | 19.790 | 29.007 |
| | 2 | 3.281 | 4.328 | 5.296 | 6.773 | 4.406 | 5.711 |
| 24 | 1 | 19.468 | 27.858 | 20.088 | 28.320 | 19.507 | 28.924 |
| | 2 | 4.290 | 5.602 | 3.783 | 4.943 | 5.138 | 6.598 |
| 25 | 1 | 19.680 | 28.297 | 19.851 | 28.652 | 20.026 | 28.984 |
| | 2 | 3.027 | 4.036 | 4.079 | 5.320 | 4.599 | 5.968 |
| 26 | 1 | 19.829 | 28.396 | 20.142 | 28.419 | 19.374 | 27.304 |
| | 2 | 3.949 | 5.192 | 2.842 | 3.912 | 3.666 | 4.789 |
| 27 | 1 | 19.910 | 28.098 | 19.460 | 29.135 | 20.177 | 27.494 |
| | 2 | 5.102 | 6.560 | 2.916 | 3.956 | 4.268 | 5.579 |
| 28 | 1 | 19.980 | 27.906 | 20.192 | 27.784 | 19.171 | 28.720 |
| | 2 | 3.923 | 5.140 | 4.409 | 5.722 | 5.020 | 6.442 |
| 29 | 1 | 19.733 | 27.241 | 20.210 | 28.842 | 20.205 | 28.761 |
| | 2 | 4.974 | 6.398 | 4.311 | 5.616 | 4.612 | 5.974 |
| 30 | 1 | 19.178 | 28.952 | 20.022 | 28.705 | 19.019 | 29.019 |
| | 2 | 3.956 | 5.200 | 4.213 | 5.516 | 5.301 | 6.793 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 48 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 60 : 8 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 31 | 1 | 19.706 | 28.923 | 19.915 | 28.140 | 19.716 | 27.471 |
| | 2 | 3.676 | 4.822 | 3.581 | 4.658 | 5.320 | 6.817 |
| 32 | 1 | 19.475 | 28.152 | 20.187 | 27.790 | 20.190 | 27.517 |
| | 2 | 5.321 | 6.823 | 4.222 | 5.528 | 5.297 | 6.780 |
| 33 | 1 | 19.654 | 28.546 | 19.106 | 28.535 | 19.888 | 28.904 |
| | 2 | 4.906 | 6.335 | 5.256 | 6.719 | 4.457 | 5.795 |
| 34 | 1 | 19.261 | 29.113 | 19.105 | 29.032 | 19.989 | 27.822 |
| | 2 | 5.338 | 6.864 | 3.455 | 4.525 | 4.383 | 5.703 |
| 35 | 1 | 18.937 | 28.361 | 19.862 | 27.481 | 19.714 | 28.023 |
| | 2 | 4.388 | 5.706 | 2.855 | 3.917 | 5.125 | 6.570 |
| 36 | 1 | 19.724 | 28.619 | 19.857 | 27.340 | 19.715 | 27.694 |
| | 2 | 4.753 | 6.178 | 3.135 | 4.156 | 5.048 | 6.507 |
| 37 | 1 | 20.136 | 27.403 | 19.253 | 27.935 | 19.135 | 27.250 |
| | 2 | 5.139 | 6.603 | 3.462 | 4.532 | 3.187 | 4.211 |
| 38 | 1 | 19.723 | 28.364 | 20.206 | 27.455 | 19.884 | 28.865 |
| | 2 | 4.537 | 5.877 | 3.983 | 5.227 | 4.560 | 5.906 |
| 39 | 1 | 20.060 | 28.075 | 20.166 | 27.291 | 19.541 | 27.416 |
| | 2 | 5.161 | 6.630 | 3.566 | 4.639 | 3.226 | 4.282 |
| 40 | 1 | 19.697 | 27.838 | 18.999 | 27.615 | 19.385 | 28.256 |
| | 2 | 5.380 | 6.902 | 5.090 | 6.545 | 5.160 | 6.627 |
| 41 | 1 | 19.744 | 27.372 | 19.361 | 28.476 | 19.045 | 27.878 |
| | 2 | 5.388 | 6.937 | 4.054 | 5.279 | 4.758 | 6.187 |
| 42 | 1 | 20.013 | 28.297 | 19.701 | 28.047 | 19.773 | 28.294 |
| | 2 | 3.213 | 4.274 | 4.140 | 5.430 | 5.097 | 6.554 |
| 43 | 1 | 19.543 | 27.304 | 19.919 | 29.071 | 20.034 | 27.624 |
| | 2 | 3.763 | 4.908 | 3.100 | 4.120 | 4.576 | 5.935 |
| 44 | 1 | 19.203 | 27.733 | 20.085 | 27.636 | 19.083 | 28.750 |
| | 2 | 3.092 | 4.105 | 4.536 | 5.875 | 3.202 | 4.257 |
| 45 | 1 | 19.792 | 27.656 | 19.346 | 28.362 | 19.390 | 27.414 |
| | 2 | 3.940 | 5.184 | 4.474 | 5.800 | 5.003 | 6.414 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 48 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 60 : 8 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 46 | 1 | 19.235 | 28.020 | 19.572 | 29.032 | 19.430 | 29.062 |
| | 2 | 4.732 | 6.153 | 4.664 | 6.079 | 3.818 | 5.015 |
| 47 | 1 | 20.154 | 28.837 | 19.844 | 28.555 | 19.309 | 27.677 |
| | 2 | 3.165 | 4.190 | 4.571 | 5.929 | 2.952 | 3.983 |
| 48 | 1 | 18.910 | 27.340 | 19.386 | 28.177 | 18.936 | 27.835 |
| | 2 | 3.982 | 5.226 | 4.553 | 5.886 | 4.445 | 5.767 |
| 49 | 1 | 19.011 | 28.357 | 19.277 | 28.303 | 19.214 | 29.030 |
| | 2 | 4.630 | 6.031 | 4.687 | 6.100 | 5.057 | 6.525 |
| 50 | 1 | 18.972 | 27.837 | 20.019 | 28.888 | 19.503 | 28.092 |
| | 2 | 3.557 | 4.632 | 3.881 | 5.099 | 3.280 | 4.320 |
| 51 | 1 | 19.866 | 27.999 | 19.428 | 28.492 | 19.462 | 27.214 |
| | 2 | 4.648 | 6.052 | 5.411 | 6.962 | 3.756 | 4.899 |
| 52 | 1 | 19.384 | 28.664 | 19.509 | 28.398 | 19.005 | 28.171 |
| | 2 | 3.723 | 4.855 | 4.619 | 5.978 | 5.112 | 6.561 |
| 53 | 1 | 19.089 | 28.438 | 19.980 | 28.026 | 19.075 | 27.777 |
| | 2 | 3.813 | 4.992 | 5.126 | 6.571 | 4.792 | 6.229 |
| 54 | 1 | 19.843 | 27.577 | 18.981 | 29.044 | 20.050 | 27.234 |
| | 2 | 3.685 | 4.831 | 3.886 | 5.105 | 4.680 | 6.097 |
| 55 | 1 | 19.537 | 27.373 | 19.663 | 28.597 | 19.086 | 28.578 |
| | 2 | 4.128 | 5.405 | 4.665 | 6.083 | 3.556 | 4.630 |
| 56 | 1 | 19.123 | 27.481 | 19.866 | 28.792 | 19.306 | 28.196 |
| | 2 | 3.752 | 4.881 | 4.529 | 5.872 | 4.688 | 6.101 |
| 57 | 1 | 19.996 | 28.066 | 20.004 | 29.093 | 19.664 | 27.499 |
| | 2 | 3.798 | 4.974 | 5.275 | 6.740 | 4.623 | 6.005 |
| 58 | 1 | 19.716 | 27.349 | 19.902 | 27.359 | 19.403 | 27.224 |
| | 2 | 4.163 | 5.461 | 4.444 | 5.766 | 4.993 | 6.408 |
| 59 | 1 | 19.056 | 28.691 | 19.524 | 27.646 | 19.309 | 27.738 |
| | 2 | 4.627 | 6.009 | 3.199 | 4.238 | 3.843 | 5.038 |
| 60 | 1 | 19.527 | 27.938 | 19.332 | 28.779 | 20.161 | 27.978 |
| | 2 | 4.852 | 6.283 | 3.788 | 4.946 | 3.376 | 4.427 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 48 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 60 : 8 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 61 | 1 | 19.592 | 27.544 | 19.517 | 28.502 | 19.442 | 29.080 |
| | 2 | 3.687 | 4.844 | 2.911 | 3.948 | 5.394 | 6.944 |
| 62 | 1 | 20.151 | 28.791 | 20.054 | 27.619 | 19.960 | 27.959 |
| | 2 | 3.324 | 4.366 | 5.293 | 6.765 | 5.079 | 6.538 |
| 63 | 1 | 19.098 | 29.106 | 19.368 | 27.999 | 19.615 | 29.015 |
| | 2 | 3.558 | 4.632 | 4.541 | 5.882 | 4.782 | 6.202 |
| 64 | 1 | 19.645 | 28.503 | 19.226 | 28.687 | 18.921 | 27.793 |
| | 2 | 3.965 | 5.206 | 4.232 | 5.536 | 5.365 | 6.882 |
| 65 | 1 | 20.155 | 28.682 | 19.254 | 28.307 | 19.187 | 28.899 |
| | 2 | 3.776 | 4.934 | 4.301 | 5.615 | 3.734 | 4.856 |
| 66 | 1 | 19.996 | 29.026 | 18.962 | 28.159 | 19.798 | 27.309 |
| | 2 | 3.670 | 4.797 | 3.924 | 5.148 | 2.946 | 3.977 |
| 67 | 1 | 19.246 | 27.653 | 19.387 | 28.951 | 19.260 | 27.258 |
| | 2 | 3.908 | 5.137 | 3.469 | 4.549 | 4.762 | 6.193 |
| 68 | 1 | 19.160 | 27.817 | 20.040 | 28.323 | 20.204 | 29.126 |
| | 2 | 4.563 | 5.907 | 2.972 | 3.995 | 4.670 | 6.093 |
| 69 | 1 | 19.424 | 28.873 | 19.797 | 28.240 | 18.967 | 27.480 |
| | 2 | 2.851 | 3.915 | 4.079 | 5.320 | 3.557 | 4.632 |
| 70 | 1 | 19.856 | 28.710 | 19.246 | 28.441 | 19.123 | 28.194 |
| | 2 | 4.487 | 5.808 | 4.337 | 5.650 | 3.150 | 4.185 |
| 71 | 1 | 19.087 | 28.318 | 19.714 | 28.343 | 19.026 | 27.430 |
| | 2 | 5.396 | 6.944 | 3.640 | 4.745 | 3.186 | 4.210 |
| 72 | 1 | 19.067 | 29.101 | 19.813 | 27.971 | 20.099 | 28.387 |
| | 2 | 4.420 | 5.741 | 4.047 | 5.276 | 5.185 | 6.651 |
| 73 | 1 | 19.103 | 27.319 | 19.867 | 28.537 | 19.398 | 28.917 |
| | 2 | 3.875 | 5.079 | 4.484 | 5.801 | 4.969 | 6.392 |
| 74 | 1 | 18.983 | 28.900 | 19.885 | 28.366 | 19.137 | 28.657 |
| | 2 | 2.856 | 3.918 | 3.200 | 4.244 | 3.604 | 4.688 |
| 75 | 1 | 19.628 | 28.047 | 19.535 | 28.541 | 19.347 | 27.753 |
| | 2 | 4.031 | 5.250 | 4.878 | 6.307 | 3.896 | 5.118 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 48 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 60 : 8 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 76 | 1 | 20.053 | 27.557 | 19.419 | 27.941 | 19.692 | 28.597 |
| | 2 | 4.693 | 6.112 | 4.916 | 6.342 | 5.226 | 6.695 |
| 77 | 1 | 19.200 | 27.853 | 20.213 | 27.418 | 19.274 | 29.122 |
| | 2 | 4.831 | 6.274 | 4.269 | 5.580 | 4.141 | 5.432 |
| 78 | 1 | 19.088 | 28.420 | 19.043 | 28.056 | 19.914 | 27.964 |
| | 2 | 3.051 | 4.067 | 5.062 | 6.530 | 3.686 | 4.833 |
| 79 | 1 | 19.776 | 28.552 | 19.170 | 27.358 | 20.082 | 28.608 |
| | 2 | 5.229 | 6.695 | 4.026 | 5.247 | 3.758 | 4.903 |
| 80 | 1 | 19.872 | 27.396 | 18.971 | 28.601 | 19.163 | 28.249 |
| | 2 | 4.150 | 5.449 | 4.238 | 5.543 | 3.385 | 4.434 |
| 81 | 1 | 19.280 | 28.649 | 19.927 | 28.765 | 20.120 | 27.883 |
| | 2 | 3.122 | 4.145 | 3.743 | 4.871 | 4.168 | 5.467 |
| 82 | 1 | 19.248 | 27.526 | 19.853 | 28.127 | 19.930 | 27.362 |
| | 2 | 3.346 | 4.391 | 4.642 | 6.040 | 5.246 | 6.712 |
| 83 | 1 | 19.739 | 27.237 | 19.847 | 28.894 | 19.298 | 28.315 |
| | 2 | 3.994 | 5.237 | 4.182 | 5.480 | 5.126 | 6.578 |
| 84 | 1 | 18.974 | 28.269 | 19.808 | 28.942 | 20.107 | 28.136 |
| | 2 | 4.377 | 5.686 | 5.143 | 6.606 | 5.024 | 6.449 |
| 85 | 1 | 20.053 | 27.957 | 19.680 | 28.919 | 19.208 | 27.795 |
| | 2 | 3.190 | 4.219 | 5.294 | 6.768 | 3.327 | 4.383 |
| 86 | 1 | 19.067 | 28.187 | 19.400 | 28.508 | 19.868 | 27.443 |
| | 2 | 4.129 | 5.409 | 3.326 | 4.376 | 3.934 | 5.164 |
| 87 | 1 | 19.479 | 28.035 | 19.850 | 27.659 | 18.976 | 29.110 |
| | 2 | 4.496 | 5.828 | 3.930 | 5.163 | 3.135 | 4.160 |
| 88 | 1 | 19.866 | 27.729 | 18.925 | 27.588 | 18.968 | 28.958 |
| | 2 | 5.295 | 6.769 | 3.655 | 4.761 | 3.596 | 4.676 |
| 89 | 1 | 20.015 | 28.388 | 20.008 | 27.272 | 20.043 | 27.859 |
| | 2 | 4.136 | 5.419 | 4.233 | 5.539 | 4.637 | 6.036 |
| 90 | 1 | 19.510 | 29.011 | 19.120 | 27.738 | 20.173 | 27.925 |
| | 2 | 3.740 | 4.867 | 3.079 | 4.098 | 4.412 | 5.725 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 48 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 60 : 8 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 91 | 1 | 20.130 | 27.474 | 18.906 | 28.150 | 19.185 | 28.070 |
| | 2 | 4.126 | 5.403 | 5.282 | 6.750 | 4.624 | 6.008 |
| 92 | 1 | 20.208 | 28.056 | 19.529 | 28.167 | 19.247 | 27.458 |
| | 2 | 3.917 | 5.139 | 3.574 | 4.647 | 3.923 | 5.144 |
| 93 | 1 | 19.013 | 28.075 | 19.936 | 27.638 | 19.447 | 27.275 |
| | 2 | 3.469 | 4.543 | 3.518 | 4.612 | 5.409 | 6.957 |
| 94 | 1 | 20.129 | 28.431 | 19.715 | 28.590 | 18.928 | 27.930 |
| | 2 | 5.293 | 6.762 | 3.578 | 4.653 | 5.218 | 6.682 |
| 95 | 1 | 19.167 | 28.428 | 18.947 | 28.006 | 19.171 | 27.952 |
| | 2 | 4.031 | 5.258 | 4.205 | 5.506 | 3.805 | 4.981 |
| 96 | 1 | 19.431 | 28.036 | 19.532 | 27.682 | 19.008 | 28.603 |
| | 2 | 4.064 | 5.290 | 3.611 | 4.707 | 4.810 | 6.256 |
| 97 | 1 | 19.502 | 27.870 | 19.881 | 28.443 | 20.021 | 27.779 |
| | 2 | 4.089 | 5.334 | 3.306 | 4.357 | 2.941 | 3.969 |
| 98 | 1 | 19.247 | 28.384 | 19.030 | 28.620 | 19.287 | 28.684 |
| | 2 | 3.483 | 4.563 | 5.158 | 6.625 | 4.946 | 6.367 |
| 99 | 1 | 19.064 | 29.097 | 19.928 | 28.847 | 20.096 | 29.086 |
| | 2 | 4.796 | 6.238 | 5.017 | 6.432 | 5.060 | 6.529 |
| 100 | 1 | 19.080 | 28.815 | 20.137 | 28.685 | 19.789 | 29.096 |
| | 2 | 3.498 | 4.588 | 3.235 | 4.298 | 2.955 | 3.990 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 49 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 70 : 6 ด้วยการทำซ้ำ 100 ครั้ง

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 1 | 1 | 18.268 | 25.422 | 19.683 | 25.596 | 18.182 | 27.887 |
| | 2 | 3.614 | 4.712 | 3.257 | 4.306 | 5.006 | 6.419 |
| 2 | 1 | 19.883 | 26.912 | 17.59 | 28.861 | 18.113 | 28.26 |
| | 2 | 3.197 | 4.233 | 3.648 | 4.757 | 5.051 | 6.516 |
| 3 | 1 | 17.824 | 26.411 | 19.502 | 26.298 | 17.325 | 26.461 |
| | 2 | 3.681 | 4.826 | 4.840 | 6.279 | 4.567 | 5.919 |
| 4 | 1 | 19.437 | 25.862 | 17.453 | 26.181 | 18.546 | 27.586 |
| | 2 | 3.082 | 4.098 | 3.040 | 4.053 | 3.032 | 4.047 |
| 5 | 1 | 19.539 | 28.363 | 19.637 | 28.474 | 17.488 | 27.68 |
| | 2 | 4.098 | 5.355 | 4.885 | 6.312 | 2.878 | 3.935 |
| 6 | 1 | 17.45 | 26.123 | 19.532 | 27.644 | 17.966 | 28.863 |
| | 2 | 3.227 | 4.290 | 4.692 | 6.104 | 4.545 | 5.882 |
| 7 | 1 | 19 | 25.899 | 19.492 | 28.914 | 19.837 | 25.718 |
| | 2 | 4.959 | 6.381 | 5.224 | 6.691 | 4.164 | 5.463 |
| 8 | 1 | 19.09 | 28.964 | 19.189 | 26.485 | 18.668 | 27.623 |
| | 2 | 3.538 | 4.622 | 4.899 | 6.322 | 2.904 | 3.940 |
| 9 | 1 | 19.622 | 27.49 | 17.897 | 28.37 | 18.579 | 27.063 |
| | 2 | 3.418 | 4.474 | 3.839 | 5.023 | 5.341 | 6.870 |
| 10 | 1 | 18.095 | 28.708 | 17.651 | 26.548 | 18.91 | 29.194 |
| | 2 | 3.853 | 5.052 | 5.216 | 6.676 | 4.474 | 5.799 |
| 11 | 1 | 17.678 | 27.343 | 18.043 | 25.312 | 18.216 | 26.25 |
| | 2 | 5.387 | 6.928 | 4.393 | 5.707 | 3.717 | 4.855 |
| 12 | 1 | 19.231 | 26.344 | 19.063 | 26.64 | 17.914 | 26.506 |
| | 2 | 3.728 | 4.856 | 3.229 | 4.292 | 4.288 | 5.591 |
| 13 | 1 | 18.127 | 27.19 | 19.151 | 28.537 | 19.872 | 25.948 |
| | 2 | 4.231 | 5.536 | 3.906 | 5.137 | 4.287 | 5.585 |
| 14 | 1 | 18.536 | 26.812 | 19.84 | 29.2 | 18.598 | 26.742 |
| | 2 | 3.077 | 4.095 | 4.992 | 6.405 | 4.157 | 5.455 |
| 15 | 1 | 18.591 | 26.264 | 18.052 | 29.112 | 19.673 | 27.756 |
| | 2 | 3.372 | 4.418 | 3.197 | 4.233 | 2.848 | 3.914 |
| 16 | 1 | 18.374 | 26.632 | 18.45 | 25.444 | 17.918 | 28.42 |
| | 2 | 5.368 | 6.884 | 5.327 | 6.836 | 5.254 | 6.717 |
| 17 | 1 | 18.357 | 28.094 | 18.122 | 26.771 | 17.951 | 26.174 |
| | 2 | | | | | | |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 49 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 70 : 6 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 18 | 1 | 17.606 | 27.885 | 18.984 | 27.311 | 17.54 | 25.333 |
| | 2 | 5.049 | 6.508 | 4.164 | 5.463 | 3.329 | 4.383 |
| 19 | 1 | 17.477 | 28.864 | 17.263 | 29.056 | 17.561 | 27.096 |
| | 2 | 4.705 | 6.122 | 3.940 | 5.181 | 5.093 | 6.548 |
| 20 | 1 | 18.243 | 26.638 | 18.345 | 27.622 | 19.647 | 25.856 |
| | 2 | 3.511 | 4.606 | 3.797 | 4.973 | 5.035 | 6.474 |
| 21 | 1 | 19.267 | 25.882 | 18.595 | 27.344 | 19.326 | 28.293 |
| | 2 | 5.215 | 6.675 | 3.283 | 4.330 | 4.193 | 5.489 |
| 22 | 1 | 19.27 | 25.943 | 17.76 | 25.496 | 18.158 | 28.623 |
| | 2 | 2.995 | 4.015 | 4.910 | 6.337 | 4.443 | 5.763 |
| 23 | 1 | 18.454 | 29.094 | 18.373 | 27.721 | 17.983 | 25.684 |
| | 2 | 4.045 | 5.272 | 2.837 | 3.893 | 3.262 | 4.307 |
| 24 | 1 | 19.192 | 27.58 | 18.082 | 28.727 | 17.497 | 25.946 |
| | 2 | 3.367 | 4.412 | 4.345 | 5.652 | 4.066 | 5.290 |
| 25 | 1 | 18.636 | 27.259 | 18.765 | 25.657 | 18.523 | 25.309 |
| | 2 | 4.922 | 6.344 | 5.401 | 6.948 | 3.893 | 5.116 |
| 26 | 1 | 17.807 | 26.39 | 19.765 | 26.684 | 19.214 | 26.518 |
| | 2 | 4.114 | 5.390 | 3.407 | 4.451 | 5.280 | 6.744 |
| 27 | 1 | 17.721 | 27.162 | 18.676 | 28.698 | 19.535 | 25.756 |
| | 2 | 3.194 | 4.227 | 3.771 | 4.912 | 4.165 | 5.465 |
| 28 | 1 | 18.929 | 28.073 | 17.362 | 25.31 | 17.826 | 25.261 |
| | 2 | 4.208 | 5.507 | 5.169 | 6.640 | 4.671 | 6.095 |
| 29 | 1 | 19.55 | 26.52 | 18.377 | 26.245 | 19.086 | 25.74 |
| | 2 | 3.013 | 4.033 | 3.032 | 4.045 | 3.119 | 4.143 |
| 30 | 1 | 17.992 | 27.387 | 18.666 | 26.623 | 18.823 | 27.07 |
| | 2 | 3.475 | 4.553 | 3.108 | 4.121 | 5.131 | 6.589 |
| 31 | 1 | 19.607 | 28.84 | 18.853 | 28.825 | 17.475 | 27.659 |
| | 2 | 4.199 | 5.503 | 4.292 | 5.610 | 4.930 | 6.357 |
| 32 | 1 | 19.814 | 26.219 | 18.93 | 26.102 | 18.521 | 26.045 |
| | 2 | 4.575 | 5.930 | 5.383 | 6.910 | 4.647 | 6.044 |
| 33 | 1 | 19.052 | 27.6 | 18.604 | 25.401 | 18.269 | 29.167 |
| | 2 | 5.324 | 6.830 | 2.863 | 3.922 | 4.765 | 6.200 |
| 34 | 1 | 18.957 | 26.481 | 19.696 | 25.568 | 18.125 | 28.754 |
| | 2 | 5.034 | 6.465 | 2.837 | 3.893 | 3.862 | 5.064 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 49 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 70 : 6 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 35 | 1 | 18.954 | 25.564 | 19.214 | 27.596 | 19.111 | 28.145 |
| | 2 | 3.099 | 4.117 | 2.923 | 3.960 | 3.539 | 4.623 |
| 36 | 1 | 18.341 | 26.736 | 18.716 | 27.821 | 18.29 | 29.205 |
| | 2 | 2.957 | 3.990 | 3.775 | 4.934 | 4.407 | 5.718 |
| 37 | 1 | 18.582 | 26.698 | 17.298 | 27.161 | 18.797 | 27.062 |
| | 2 | 4.923 | 6.346 | 4.824 | 6.264 | 3.883 | 5.100 |
| 38 | 1 | 19.851 | 28.254 | 19.559 | 28.842 | 17.695 | 26.081 |
| | 2 | 4.570 | 5.920 | 3.610 | 4.695 | 4.216 | 5.526 |
| 39 | 1 | 18.488 | 28.102 | 19.249 | 25.456 | 17.6 | 27.756 |
| | 2 | 4.971 | 6.396 | 4.522 | 5.853 | 4.831 | 6.264 |
| 40 | 1 | 19.043 | 28.9 | 17.757 | 28.373 | 18.68 | 28.308 |
| | 2 | 4.293 | 5.610 | 3.359 | 4.401 | 4.436 | 5.758 |
| 41 | 1 | 19.231 | 28.259 | 18.111 | 27.954 | 17.516 | 28.868 |
| | 2 | 3.940 | 5.181 | 3.904 | 5.132 | 3.051 | 4.065 |
| 42 | 1 | 18.222 | 28.714 | 18.686 | 28.212 | 19.897 | 28.364 |
| | 2 | 4.249 | 5.561 | 3.189 | 4.214 | 4.432 | 5.749 |
| 43 | 1 | 17.799 | 28.468 | 17.831 | 27.149 | 19.224 | 27.103 |
| | 2 | 3.478 | 4.560 | 3.057 | 4.072 | 3.928 | 5.155 |
| 44 | 1 | 17.367 | 27.617 | 18.942 | 28.936 | 18.835 | 28.848 |
| | 2 | 3.873 | 5.075 | 3.672 | 4.804 | 3.396 | 4.439 |
| 45 | 1 | 17.657 | 25.587 | 19.149 | 26.519 | 18.476 | 28.465 |
| | 2 | 5.051 | 6.513 | 2.945 | 3.972 | 5.151 | 6.616 |
| 46 | 1 | 19.639 | 29.006 | 18.305 | 28.727 | 18.686 | 26.936 |
| | 2 | 4.596 | 5.961 | 4.189 | 5.486 | 4.413 | 5.731 |
| 47 | 1 | 18.836 | 27.522 | 19.363 | 25.865 | 18.508 | 28.193 |
| | 2 | 4.954 | 6.377 | 3.755 | 4.896 | 3.860 | 5.058 |
| 48 | 1 | 17.541 | 27.635 | 19.125 | 28.972 | 18.841 | 28.096 |
| | 2 | 2.867 | 3.927 | 4.043 | 5.271 | 4.623 | 6.005 |
| 49 | 1 | 18.507 | 26.925 | 18.127 | 26.743 | 17.621 | 28.229 |
| | 2 | 3.612 | 4.710 | 5.414 | 6.971 | 4.214 | 5.522 |
| 50 | 1 | 18.913 | 25.874 | 18.17 | 28.067 | 19.135 | 28.052 |
| | 2 | 4.857 | 6.292 | 4.759 | 6.190 | 4.802 | 6.248 |
| 51 | 1 | 18.842 | 27.201 | 19.464 | 28.819 | 18.891 | 28.674 |
| | 2 | 4.076 | 5.312 | 4.582 | 5.947 | 2.858 | 3.919 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 49 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 70 : 6 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 52 | 1 | 19.413 | 26.425 | 17.752 | 26.289 | 18.449 | 28.299 |
| | 2 | 3.851 | 5.042 | 4.922 | 6.345 | 5.191 | 6.655 |
| 53 | 1 | 18.444 | 25.962 | 18.19 | 25.353 | 19.831 | 27.731 |
| | 2 | 3.675 | 4.819 | 3.666 | 4.788 | 5.305 | 6.800 |
| 54 | 1 | 18.956 | 27.387 | 17.596 | 25.192 | 17.99 | 26.179 |
| | 2 | 4.008 | 5.239 | 5.401 | 6.954 | 4.118 | 5.395 |
| 55 | 1 | 19.812 | 27.804 | 17.715 | 27.251 | 19.73 | 25.191 |
| | 2 | 4.075 | 5.303 | 4.195 | 5.491 | 2.950 | 3.980 |
| 56 | 1 | 19.147 | 26.373 | 18.226 | 26.197 | 19.089 | 26.641 |
| | 2 | 4.391 | 5.707 | 4.693 | 6.110 | 3.699 | 4.853 |
| 57 | 1 | 19.89 | 25.897 | 18.546 | 28.898 | 19.653 | 25.568 |
| | 2 | 3.669 | 4.792 | 3.999 | 5.237 | 3.751 | 4.878 |
| 58 | 1 | 18.608 | 25.211 | 19.789 | 25.97 | 19.235 | 26.792 |
| | 2 | 4.993 | 6.408 | 3.522 | 4.616 | 4.279 | 5.581 |
| 59 | 1 | 17.468 | 26.287 | 17.868 | 28.069 | 19.809 | 25.824 |
| | 2 | 4.031 | 5.248 | 5.367 | 6.882 | 3.703 | 4.855 |
| 60 | 1 | 18.516 | 27.77 | 19.357 | 27.428 | 18.64 | 26.421 |
| | 2 | 4.184 | 5.480 | 4.882 | 6.312 | 3.539 | 4.625 |
| 61 | 1 | 19.02 | 28.132 | 18.958 | 27.29 | 19.661 | 26.244 |
| | 2 | 3.375 | 4.420 | 3.292 | 4.332 | 3.736 | 4.859 |
| 62 | 1 | 19.609 | 28.529 | 18.352 | 25.549 | 18.611 | 25.723 |
| | 2 | 4.110 | 5.363 | 4.130 | 5.413 | 3.868 | 5.071 |
| 63 | 1 | 17.317 | 27.694 | 18.149 | 25.867 | 18.861 | 27.161 |
| | 2 | 4.881 | 6.310 | 3.571 | 4.641 | 4.367 | 5.674 |
| 64 | 1 | 18.993 | 26.853 | 18.053 | 25.227 | 17.403 | 29.031 |
| | 2 | 5.096 | 6.553 | 3.092 | 4.101 | 3.902 | 5.125 |
| 65 | 1 | 18.177 | 26.021 | 19.631 | 28.981 | 17.79 | 27.859 |
| | 2 | 3.806 | 4.985 | 5.022 | 6.448 | 3.980 | 5.223 |
| 66 | 1 | 18.391 | 26.205 | 18.788 | 28.129 | 17.856 | 25.897 |
| | 2 | 4.286 | 5.582 | 5.341 | 6.869 | 4.951 | 6.377 |
| 67 | 1 | 19.668 | 25.263 | 18.065 | 28.508 | 19.757 | 28.764 |
| | 2 | 3.168 | 4.193 | 4.612 | 5.972 | 4.070 | 5.300 |
| 68 | 1 | 19.157 | 26.983 | 17.939 | 27.587 | 19.472 | 26.944 |
| | 2 | 4.105 | 5.363 | 4.664 | 6.077 | 4.332 | 5.643 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 49 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 70 : 6 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 69 | 1 | 17.832 | 27.922 | 19.096 | 28.439 | 19.096 | 26.967 |
| | 2 | 5.154 | 6.619 | 4.945 | 6.365 | 4.718 | 6.126 |
| 70 | 1 | 19.65 | 26.767 | 18.399 | 27.321 | 18.475 | 26.786 |
| | 2 | 3.177 | 4.202 | 5.209 | 6.674 | 5.046 | 6.499 |
| 71 | 1 | 18.518 | 26.004 | 17.949 | 27.122 | 18.154 | 28.428 |
| | 2 | 3.506 | 4.602 | 4.168 | 5.466 | 5.253 | 6.715 |
| 72 | 1 | 18.03 | 26.583 | 17.516 | 28.024 | 18.902 | 26.014 |
| | 2 | 4.961 | 6.381 | 4.001 | 5.238 | 4.770 | 6.201 |
| 73 | 1 | 19.758 | 28.713 | 18.057 | 27.177 | 19.126 | 28.496 |
| | 2 | 5.370 | 6.887 | 3.234 | 4.297 | 4.997 | 6.409 |
| 74 | 1 | 18.893 | 26.633 | 18.544 | 27.931 | 18.807 | 26.995 |
| | 2 | 5.147 | 6.612 | 4.014 | 5.240 | 5.235 | 6.696 |
| 75 | 1 | 19.587 | 28.782 | 19.172 | 27.747 | 19.196 | 26.545 |
| | 2 | 3.289 | 4.331 | 4.137 | 5.424 | 4.831 | 6.274 |
| 76 | 1 | 19.086 | 28.296 | 17.462 | 27.875 | 18.242 | 28.354 |
| | 2 | 4.667 | 6.085 | 5.258 | 6.721 | 3.657 | 4.780 |
| 77 | 1 | 17.664 | 26.781 | 18.6 | 27.501 | 19.639 | 28.377 |
| | 2 | 4.571 | 5.925 | 4.365 | 5.664 | 3.202 | 4.246 |
| 78 | 1 | 17.336 | 27.865 | 19.068 | 28.346 | 19.68 | 27.869 |
| | 2 | 2.915 | 3.955 | 4.635 | 6.031 | 3.426 | 4.498 |
| 79 | 1 | 19.396 | 27.668 | 19.72 | 27.395 | 17.925 | 27.423 |
| | 2 | 3.869 | 5.071 | 3.793 | 4.957 | 4.081 | 5.322 |
| 80 | 1 | 19.064 | 28.943 | 18.038 | 27.263 | 18.674 | 28.361 |
| | 2 | 3.848 | 5.038 | 4.961 | 6.385 | 4.429 | 5.748 |
| 81 | 1 | 17.534 | 25.852 | 18.793 | 27.269 | 18.581 | 25.484 |
| | 2 | 4.909 | 6.337 | 4.724 | 6.147 | 3.389 | 4.435 |
| 82 | 1 | 18.97 | 28.628 | 18.674 | 28.172 | 17.519 | 25.968 |
| | 2 | 3.735 | 4.858 | 3.223 | 4.281 | 3.752 | 4.879 |
| 83 | 1 | 19.337 | 28.42 | 18.369 | 27.501 | 17.306 | 25.461 |
| | 2 | 4.299 | 5.613 | 4.514 | 5.846 | 3.674 | 4.817 |
| 84 | 1 | 18.574 | 26.658 | 19.783 | 29.207 | 18.729 | 25.915 |
| | 2 | 4.860 | 6.297 | 3.659 | 4.783 | 4.583 | 5.950 |
| 85 | 1 | 19.026 | 25.711 | 18.499 | 27.784 | 19.202 | 26.491 |
| | 2 | 3.357 | 4.400 | 4.110 | 5.366 | 4.188 | 5.485 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 49 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 70 : 6 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 86 | 1 | 17.733 | 26.79 | 18.596 | 28.801 | 19.244 | 26.78 |
| | 2 | 4.586 | 5.955 | 3.771 | 4.917 | 3.378 | 4.431 |
| 87 | 1 | 19.278 | 25.56 | 19.508 | 28.18 | 17.873 | 28.703 |
| | 2 | 3.475 | 4.554 | 4.813 | 6.256 | 4.529 | 5.868 |
| 88 | 1 | 17.517 | 29.04 | 19.6 | 28.566 | 18.952 | 28.463 |
| | 2 | 5.368 | 6.887 | 5.327 | 6.846 | 3.465 | 4.537 |
| 89 | 1 | 18.639 | 26.051 | 17.935 | 26.534 | 18.71 | 27.501 |
| | 2 | 3.204 | 4.259 | 4.968 | 6.392 | 3.852 | 5.046 |
| 90 | 1 | 19.665 | 28.576 | 19.005 | 26.911 | 18.211 | 27.095 |
| | 2 | 4.156 | 5.454 | 4.916 | 6.342 | 3.195 | 4.230 |
| 91 | 1 | 19.749 | 28.957 | 17.5 | 27.527 | 18.593 | 26.817 |
| | 2 | 2.917 | 3.958 | 3.772 | 4.922 | 4.135 | 5.414 |
| 92 | 1 | 19.087 | 25.925 | 18.054 | 26.25 | 19.303 | 25.819 |
| | 2 | 2.911 | 3.945 | 4.488 | 5.810 | 2.955 | 3.984 |
| 93 | 1 | 18.84 | 27.297 | 19.53 | 27.104 | 18.709 | 29.007 |
| | 2 | 2.960 | 3.992 | 4.520 | 5.853 | 3.356 | 4.392 |
| 94 | 1 | 18.192 | 25.895 | 18.728 | 29.177 | 18.355 | 26.694 |
| | 2 | 5.136 | 6.596 | 2.890 | 3.939 | 3.662 | 4.784 |
| 95 | 1 | 19.51 | 26.763 | 18.68 | 25.981 | 18.395 | 28.804 |
| | 2 | 5.034 | 6.468 | 5.116 | 6.562 | 3.859 | 5.056 |
| 96 | 1 | 18.589 | 27.856 | 19.418 | 28.273 | 19.858 | 26.606 |
| | 2 | 5.267 | 6.722 | 4.979 | 6.399 | 3.634 | 4.739 |
| 97 | 1 | 19.355 | 28.39 | 19.874 | 25.852 | 17.816 | 27.625 |
| | 2 | 3.984 | 5.229 | 3.036 | 4.048 | 4.991 | 6.404 |
| 98 | 1 | 17.606 | 27.221 | 18.25 | 27.578 | 19.37 | 28.911 |
| | 2 | 4.263 | 5.568 | 2.852 | 3.916 | 3.879 | 5.089 |
| 99 | 1 | 19.056 | 26.14 | 19.108 | 27.645 | 18.357 | 28.95 |
| | 2 | 4.292 | 5.604 | 3.855 | 5.053 | 3.981 | 5.225 |
| 100 | 1 | 19.326 | 28.751 | 19.167 | 25.606 | 19.568 | 26.994 |
| | 2 | 5.318 | 6.811 | 3.207 | 4.266 | 4.180 | 5.474 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 50 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 80 : 4 ด้วยการทำซ้ำ 100 ครั้ง

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 1 | 1 | 17.542 | 24.151 | 18.094 | 25.552 | 17.682 | 25.011 |
| | 2 | 3.425 | 4.492 | 3.375 | 4.426 | 3.612 | 4.711 |
| 2 | 1 | 18.11 | 25.047 | 18.335 | 24.282 | 17.78 | 24.53 |
| | 2 | 3.896 | 5.121 | 3.644 | 4.750 | 4.320 | 5.627 |
| 3 | 1 | 17.887 | 25.312 | 18.434 | 25.095 | 17.955 | 25.262 |
| | 2 | 5.067 | 6.532 | 5.087 | 6.542 | 4.715 | 6.125 |
| 4 | 1 | 17.296 | 25.644 | 17.548 | 25.101 | 17.902 | 24.162 |
| | 2 | 5.255 | 6.717 | 2.884 | 3.937 | 4.846 | 6.282 |
| 5 | 1 | 18.524 | 24.363 | 17.645 | 24.519 | 17.545 | 25.44 |
| | 2 | 3.112 | 4.133 | 4.895 | 6.321 | 4.862 | 6.301 |
| 6 | 1 | 17.699 | 25.27 | 18.425 | 24.173 | 17.47 | 23.887 |
| | 2 | 3.794 | 4.959 | 4.751 | 6.167 | 3.422 | 4.476 |
| 7 | 1 | 18.373 | 24.357 | 18.339 | 24.489 | 17.454 | 23.884 |
| | 2 | 3.979 | 5.222 | 3.487 | 4.575 | 4.662 | 6.074 |
| 8 | 1 | 18.407 | 24.306 | 17.631 | 24.679 | 17.495 | 23.996 |
| | 2 | 3.027 | 4.039 | 3.170 | 4.195 | 3.565 | 4.636 |
| 9 | 1 | 18.263 | 24.914 | 18.33 | 25.629 | 18.297 | 24.61 |
| | 2 | 3.683 | 4.827 | 5.319 | 6.811 | 3.341 | 4.388 |
| 10 | 1 | 17.429 | 25.287 | 17.814 | 25.446 | 18.255 | 25.306 |
| | 2 | 2.987 | 4.007 | 2.863 | 3.924 | 4.068 | 5.290 |
| 11 | 1 | 17.442 | 25.265 | 17.477 | 25.679 | 18.06 | 25.13 |
| | 2 | 4.327 | 5.637 | 3.902 | 5.129 | 4.785 | 6.225 |
| 12 | 1 | 17.609 | 24.678 | 18.183 | 24.895 | 17.563 | 24.399 |
| | 2 | 3.679 | 4.823 | 5.408 | 6.956 | 5.398 | 6.945 |
| 13 | 1 | 17.387 | 24.158 | 17.878 | 25.086 | 17.974 | 24.293 |
| | 2 | 5.025 | 6.455 | 4.294 | 5.612 | 4.792 | 6.231 |
| 14 | 1 | 18.439 | 25.099 | 18.494 | 25.373 | 17.288 | 24.047 |
| | 2 | 5.203 | 6.663 | 5.008 | 6.424 | 5.333 | 6.855 |
| 15 | 1 | 18.286 | 24.47 | 17.785 | 24.117 | 17.949 | 24.982 |
| | 2 | 2.839 | 3.899 | 3.485 | 4.570 | 2.864 | 3.927 |
| 16 | 1 | 18.145 | 24.251 | 18.372 | 24.372 | 18.118 | 24.227 |
| | 2 | 3.637 | 4.742 | 4.745 | 6.164 | 3.610 | 4.696 |
| 17 | 1 | 17.35 | 25.144 | 18.215 | 24.391 | 17.28 | 25.273 |
| | 2 | 3.820 | 5.016 | 4.920 | 6.343 | 5.054 | 6.519 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 50 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 80 : 4 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 18 | 1 | 18.082 | 25.319 | 17.712 | 25.537 | 18.096 | 25.674 |
| | 2 | 3.126 | 4.146 | 5.275 | 6.740 | 5.008 | 6.425 |
| 19 | 1 | 18.484 | 23.921 | 18.468 | 24.345 | 18.376 | 25.219 |
| | 2 | 3.547 | 4.626 | 4.636 | 6.032 | 3.097 | 4.107 |
| 20 | 1 | 17.888 | 24.187 | 17.728 | 25.654 | 17.758 | 25.148 |
| | 2 | 3.233 | 4.295 | 3.826 | 5.019 | 3.196 | 4.231 |
| 21 | 1 | 17.433 | 25.424 | 17.563 | 24.283 | 18.538 | 25.424 |
| | 2 | 3.108 | 4.130 | 3.301 | 4.349 | 3.501 | 4.595 |
| 22 | 1 | 17.611 | 24.677 | 17.289 | 24.289 | 18.028 | 24.287 |
| | 2 | 3.939 | 5.179 | 3.253 | 4.306 | 3.201 | 4.244 |
| 23 | 1 | 18.272 | 23.874 | 18.056 | 24.084 | 18.23 | 25.296 |
| | 2 | 3.141 | 4.170 | 3.585 | 4.663 | 5.327 | 6.846 |
| 24 | 1 | 18.501 | 25.207 | 17.801 | 24.223 | 18.262 | 24.774 |
| | 2 | 4.162 | 5.461 | 3.149 | 4.183 | 4.144 | 5.445 |
| 25 | 1 | 17.299 | 23.862 | 18.544 | 24.659 | 18.455 | 25.592 |
| | 2 | 4.417 | 5.733 | 4.557 | 5.891 | 4.208 | 5.511 |
| 26 | 1 | 18.452 | 25.323 | 17.605 | 24.349 | 18.236 | 25.268 |
| | 2 | 3.862 | 5.071 | 5.041 | 6.486 | 3.983 | 5.228 |
| 27 | 1 | 17.375 | 25.331 | 18.104 | 24.468 | 18.057 | 23.869 |
| | 2 | 5.375 | 6.894 | 2.859 | 3.921 | 5.245 | 6.707 |
| 28 | 1 | 17.951 | 23.957 | 17.604 | 25.512 | 17.752 | 24.962 |
| | 2 | 5.005 | 6.416 | 4.979 | 6.404 | 3.407 | 4.452 |
| 29 | 1 | 18.558 | 24.213 | 18.321 | 24.945 | 18.174 | 24.915 |
| | 2 | 3.127 | 4.148 | 3.344 | 4.389 | 4.858 | 6.292 |
| 30 | 1 | 17.485 | 23.911 | 18.436 | 24.343 | 18.308 | 24.428 |
| | 2 | 4.655 | 6.060 | 4.240 | 5.548 | 4.784 | 6.220 |
| 31 | 1 | 17.396 | 25.598 | 18.063 | 24.223 | 17.872 | 24.443 |
| | 2 | 5.131 | 6.591 | 3.734 | 4.858 | 2.980 | 4.003 |
| 32 | 1 | 17.709 | 24.506 | 18.533 | 25.653 | 18.55 | 24.031 |
| | 2 | 5.145 | 6.609 | 4.998 | 6.410 | 5.271 | 6.731 |
| 33 | 1 | 17.527 | 24.481 | 17.356 | 24.926 | 17.995 | 23.956 |
| | 2 | 5.128 | 6.580 | 3.740 | 4.868 | 4.368 | 5.681 |
| 34 | 1 | 18.331 | 25.163 | 17.5 | 25.49 | 17.754 | 24.327 |
| | 2 | 4.943 | 6.365 | 4.915 | 6.338 | 4.495 | 5.827 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 50 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 80 : 4 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 35 | 1 | 18.003 | 25.31 | 18.327 | 25.321 | 17.915 | 24.959 |
| | 2 | 4.110 | 5.388 | 4.507 | 5.840 | 4.597 | 5.964 |
| 36 | 1 | 18.531 | 24.607 | 17.86 | 23.898 | 18.383 | 24.142 |
| | 2 | 4.157 | 5.456 | 5.256 | 6.720 | 3.973 | 5.216 |
| 37 | 1 | 17.902 | 25.098 | 18.038 | 24.202 | 17.635 | 24.613 |
| | 2 | 5.401 | 6.955 | 3.633 | 4.733 | 3.413 | 4.459 |
| 38 | 1 | 18.148 | 24.689 | 18.483 | 24.052 | 17.827 | 25.353 |
| | 2 | 5.101 | 6.557 | 3.883 | 5.103 | 4.485 | 5.806 |
| 39 | 1 | 17.355 | 24.901 | 18.371 | 24.396 | 18.428 | 24.364 |
| | 2 | 3.936 | 5.169 | 4.563 | 5.910 | 3.575 | 4.650 |
| 40 | 1 | 17.862 | 24.444 | 18.219 | 24.487 | 17.585 | 24.437 |
| | 2 | 4.432 | 5.750 | 4.526 | 5.858 | 3.098 | 4.111 |
| 41 | 1 | 17.635 | 24.703 | 18.423 | 23.94 | 18.131 | 24.983 |
| | 2 | 3.387 | 4.434 | 2.991 | 4.015 | 3.012 | 4.033 |
| 42 | 1 | 18.265 | 25.307 | 18.484 | 24.449 | 18.439 | 24.607 |
| | 2 | 5.029 | 6.463 | 4.628 | 6.010 | 5.083 | 6.539 |
| 43 | 1 | 18.078 | 25.21 | 17.835 | 24.828 | 18.45 | 24.132 |
| | 2 | 5.063 | 6.532 | 3.938 | 5.171 | 4.248 | 5.557 |
| 44 | 1 | 18.019 | 24.831 | 17.835 | 24.283 | 17.487 | 25.66 |
| | 2 | 4.118 | 5.399 | 4.800 | 6.242 | 4.654 | 6.054 |
| 45 | 1 | 18.342 | 25.566 | 18.048 | 25.408 | 17.531 | 23.87 |
| | 2 | 3.512 | 4.606 | 5.031 | 6.463 | 3.762 | 4.904 |
| 46 | 1 | 18.159 | 24.451 | 17.997 | 25.287 | 18.019 | 24.418 |
| | 2 | 5.168 | 6.638 | 3.466 | 4.540 | 5.301 | 6.793 |
| 47 | 1 | 17.906 | 24.869 | 17.932 | 25.409 | 17.554 | 24.734 |
| | 2 | 3.756 | 4.899 | 3.516 | 4.610 | 4.578 | 5.945 |
| 48 | 1 | 17.726 | 24.019 | 18.046 | 24.73 | 17.826 | 24.216 |
| | 2 | 3.616 | 4.717 | 2.942 | 3.971 | 3.941 | 5.188 |
| 49 | 1 | 18.059 | 24.782 | 17.541 | 25.3 | 17.363 | 23.907 |
| | 2 | 3.446 | 4.504 | 5.129 | 6.584 | 3.138 | 4.165 |
| 50 | 1 | 17.598 | 24.069 | 18.155 | 24.205 | 17.615 | 24.799 |
| | 2 | 2.982 | 4.007 | 4.755 | 6.179 | 3.195 | 4.230 |
| 51 | 1 | 17.612 | 25.496 | 17.957 | 24.156 | 17.903 | 24.405 |
| | 2 | 5.195 | 6.655 | 3.297 | 4.337 | 5.393 | 6.937 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 50 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 80 : 4 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 52 | 1 | 18.333 | 24.158 | 18.11 | 24.086 | 18.393 | 25.054 |
| | 2 | 4.971 | 6.397 | 3.633 | 4.737 | 3.151 | 4.187 |
| 53 | 1 | 17.386 | 24.281 | 17.258 | 24.371 | 17.677 | 24.131 |
| | 2 | 3.668 | 4.790 | 3.602 | 4.677 | 3.762 | 4.908 |
| 54 | 1 | 17.807 | 24.453 | 18.3 | 25.017 | 17.484 | 25.545 |
| | 2 | 3.011 | 4.024 | 3.181 | 4.207 | 4.621 | 5.989 |
| 55 | 1 | 17.523 | 24.491 | 17.35 | 24.427 | 17.499 | 23.878 |
| | 2 | 3.671 | 4.802 | 3.888 | 5.110 | 4.783 | 6.208 |
| 56 | 1 | 17.312 | 25.354 | 18.156 | 24.79 | 18.407 | 25.632 |
| | 2 | 4.171 | 5.470 | 4.567 | 5.920 | 4.955 | 6.379 |
| 57 | 1 | 18.115 | 24.774 | 17.544 | 25.516 | 17.636 | 24.025 |
| | 2 | 4.757 | 6.186 | 3.656 | 4.763 | 4.241 | 5.550 |
| 58 | 1 | 17.732 | 24.04 | 17.803 | 25.123 | 17.93 | 25.11 |
| | 2 | 4.320 | 5.627 | 2.846 | 3.913 | 4.230 | 5.535 |
| 59 | 1 | 18.272 | 24.823 | 18.084 | 24.96 | 18.122 | 24.233 |
| | 2 | 4.733 | 6.155 | 3.477 | 4.556 | 3.460 | 4.532 |
| 60 | 1 | 17.739 | 23.864 | 18.554 | 24.772 | 17.509 | 24.205 |
| | 2 | 4.740 | 6.156 | 2.946 | 3.978 | 4.147 | 5.448 |
| 61 | 1 | 18.276 | 25.671 | 18.022 | 24.962 | 17.744 | 24.348 |
| | 2 | 5.341 | 6.871 | 4.808 | 6.254 | 3.571 | 4.644 |
| 62 | 1 | 18.466 | 25.177 | 18.204 | 25.401 | 17.857 | 25.458 |
| | 2 | 4.089 | 5.351 | 3.423 | 4.480 | 3.264 | 4.312 |
| 63 | 1 | 17.964 | 25.074 | 18.232 | 25.07 | 17.92 | 25.201 |
| | 2 | 4.089 | 5.355 | 3.203 | 4.257 | 3.027 | 4.041 |
| 64 | 1 | 18.528 | 25.437 | 18.305 | 23.907 | 18.399 | 24.902 |
| | 2 | 3.577 | 4.651 | 5.204 | 6.663 | 3.178 | 4.205 |
| 65 | 1 | 17.951 | 25.451 | 18.511 | 24.114 | 18.47 | 24.76 |
| | 2 | 3.262 | 4.308 | 4.700 | 6.116 | 3.593 | 4.672 |
| 66 | 1 | 17.561 | 25.659 | 17.844 | 24.042 | 17.563 | 24.516 |
| | 2 | 4.435 | 5.757 | 5.045 | 6.493 | 4.749 | 6.165 |
| 67 | 1 | 17.5 | 24.63 | 17.627 | 24.79 | 17.387 | 23.936 |
| | 2 | 3.300 | 4.344 | 5.414 | 6.971 | 4.889 | 6.318 |
| 68 | 1 | 18.034 | 24.224 | 17.668 | 25.422 | 18.039 | 24.907 |
| | 2 | 3.264 | 4.310 | 3.809 | 4.988 | 5.099 | 6.556 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 50 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 80 : 4 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 69 | 1 | 17.582 | 24.328 | 17.656 | 25.341 | 17.913 | 25.565 |
| | 2 | 5.139 | 6.605 | 4.146 | 5.448 | 4.355 | 5.653 |
| 70 | 1 | 18.02 | 25.411 | 17.275 | 25.529 | 17.942 | 25.614 |
| | 2 | 4.379 | 5.686 | 3.497 | 4.583 | 4.185 | 5.483 |
| 71 | 1 | 18.395 | 25.663 | 17.305 | 25.476 | 17.838 | 25.25 |
| | 2 | 2.880 | 3.936 | 4.838 | 6.276 | 4.643 | 6.041 |
| 72 | 1 | 17.998 | 24.384 | 18.514 | 24.908 | 17.467 | 24.121 |
| | 2 | 4.357 | 5.660 | 3.767 | 4.910 | 4.561 | 5.906 |
| 73 | 1 | 17.905 | 25.547 | 18.42 | 24.546 | 17.608 | 24.417 |
| | 2 | 3.011 | 4.029 | 5.071 | 6.534 | 5.285 | 6.757 |
| 74 | 1 | 18.04 | 24.457 | 18.465 | 24.37 | 17.408 | 25.349 |
| | 2 | 5.375 | 6.897 | 4.142 | 5.438 | 4.098 | 5.357 |
| 75 | 1 | 18.309 | 24.55 | 17.977 | 25.358 | 17.588 | 24.781 |
| | 2 | 3.788 | 4.951 | 3.626 | 4.727 | 5.189 | 6.655 |
| 76 | 1 | 17.36 | 25.424 | 18.462 | 25.494 | 17.577 | 24.277 |
| | 2 | 3.491 | 4.576 | 3.299 | 4.343 | 5.284 | 6.757 |
| 77 | 1 | 17.609 | 24.278 | 18.228 | 24.388 | 18.225 | 24.493 |
| | 2 | 3.260 | 4.307 | 3.934 | 5.169 | 4.417 | 5.739 |
| 78 | 1 | 18.045 | 25.006 | 17.795 | 24.175 | 18.275 | 23.941 |
| | 2 | 4.576 | 5.939 | 2.911 | 3.948 | 3.442 | 4.500 |
| 79 | 1 | 17.351 | 23.953 | 17.52 | 24.733 | 17.606 | 25.651 |
| | 2 | 4.406 | 5.716 | 3.815 | 4.994 | 4.673 | 6.095 |
| 80 | 1 | 18.16 | 24.521 | 18.416 | 24.108 | 17.826 | 24.837 |
| | 2 | 5.040 | 6.484 | 3.564 | 4.634 | 4.608 | 5.970 |
| 81 | 1 | 17.841 | 24.515 | 17.719 | 24.1 | 18.201 | 24.772 |
| | 2 | 4.545 | 5.886 | 3.876 | 5.087 | 3.442 | 4.500 |
| 82 | 1 | 17.613 | 24.833 | 17.674 | 24.148 | 17.807 | 25.034 |
| | 2 | 4.228 | 5.531 | 4.102 | 5.362 | 3.116 | 4.140 |
| 83 | 1 | 18.182 | 25.159 | 17.934 | 24.273 | 18.351 | 24.391 |
| | 2 | 3.318 | 4.361 | 3.002 | 4.019 | 3.136 | 4.162 |
| 84 | 1 | 17.393 | 24.812 | 17.356 | 23.897 | 17.897 | 24.855 |
| | 2 | 5.018 | 6.442 | 3.534 | 4.621 | 4.501 | 5.829 |
| 85 | 1 | 17.905 | 25.218 | 18.075 | 24.452 | 17.922 | 25.449 |
| | 2 | 4.783 | 6.207 | 4.267 | 5.576 | 3.066 | 4.087 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 50 ค่าไอเกนและร้อยละของความแปรปรวนของตัวประกอบ ของข้อมูลรูปแบบการตอบสำหรับแบบสอบ 80 : 4 ด้วยการทำซ้ำ 100 ครั้ง (ต่อ)

| ครั้งที่ | Factor | Test A | | Test B1 | | Test B2 | |
|---|---|---|---|---|---|---|---|
| | | eigen | % of variance | eigen | % of variance | eigen | % of variance |
| 86 | 1 | 17.936 | 23.956 | 17.576 | 24.389 | 17.861 | 24.937 |
| | 2 | 3.645 | 4.752 | 5.245 | 6.703 | 4.396 | 5.709 |
| 87 | 1 | 17.303 | 24.552 | 18.26 | 25.628 | 18.246 | 24.845 |
| | 2 | 5.325 | 6.835 | 3.590 | 4.670 | 4.658 | 6.063 |
| 88 | 1 | 17.476 | 23.953 | 17.91 | 25.079 | 18.364 | 25.302 |
| | 2 | 4.032 | 5.261 | 3.265 | 4.318 | 5.154 | 6.622 |
| 89 | 1 | 17.808 | 25.278 | 17.365 | 23.993 | 17.706 | 24.646 |
| | 2 | 3.605 | 4.690 | 3.815 | 4.996 | 4.175 | 5.471 |
| 90 | 1 | 18.092 | 24.441 | 18.475 | 24.318 | 17.898 | 24.355 |
| | 2 | 4.736 | 6.155 | 3.282 | 4.329 | 2.841 | 3.912 |
| 91 | 1 | 17.561 | 25.23 | 18.275 | 25.14 | 17.69 | 25.009 |
| | 2 | 4.862 | 6.299 | 4.596 | 5.962 | 4.646 | 6.044 |
| 92 | 1 | 17.458 | 25.158 | 17.89 | 25.424 | 17.777 | 24.505 |
| | 2 | 5.385 | 6.918 | 5.223 | 6.690 | 4.365 | 5.665 |
| 93 | 1 | 18.53 | 25.265 | 17.672 | 24.929 | 17.41 | 24.743 |
| | 2 | 2.886 | 3.938 | 4.316 | 5.626 | 4.494 | 5.823 |
| 94 | 1 | 18.166 | 25.151 | 18.45 | 25.231 | 17.442 | 24.062 |
| | 2 | 5.267 | 6.729 | 3.956 | 5.202 | 3.507 | 4.604 |
| 95 | 1 | 17.479 | 25.214 | 17.291 | 24.55 | 18.19 | 24.514 |
| | 2 | 3.627 | 4.728 | 4.433 | 5.755 | 3.595 | 4.675 |
| 96 | 1 | 17.608 | 24.835 | 17.813 | 24.881 | 17.908 | 25.059 |
| | 2 | 3.495 | 4.577 | 3.356 | 4.392 | 5.334 | 6.861 |
| 97 | 1 | 17.386 | 24.133 | 18.171 | 24.779 | 17.682 | 24.495 |
| | 2 | 5.330 | 6.847 | 4.723 | 6.143 | 3.625 | 4.726 |
| 98 | 1 | 17.473 | 24.295 | 17.773 | 24.128 | 17.575 | 24.697 |
| | 2 | 3.098 | 4.111 | 3.468 | 4.542 | 3.226 | 4.283 |
| 99 | 1 | 17.887 | 24.479 | 18.495 | 25.626 | 17.883 | 25.233 |
| | 2 | 3.040 | 4.057 | 4.341 | 5.650 | 3.412 | 4.459 |
| 100 | 1 | 17.803 | 24.553 | 18.32 | 25.142 | 17.564 | 24.677 |
| | 2 | 3.948 | 5.190 | 3.787 | 4.946 | 4.434 | 5.756 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

## คุณภาพด้านความเที่ยงของแบบสอบ

ตารางที่ 51 ความเที่ยงของแบสอบแต่ละฉบับ

| ครั้งที่ | 50_10 | | | 60_8 | | | 70_3 | | | 80_4 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Test A | Test B1 | Test B2 | Test A | Test B1 | Test B2 | Test A | Test B1 | Test B2 | Test A | Test B1 | Test B2 |
| 1 | 0.968 | 0.968 | 0.982 | 0.983 | 0.972 | 0.983 | 0.973 | 0.971 | 0.970 | 0.981 | 0.971 | 0.971 |
| 2 | 0.972 | 0.975 | 0.980 | 0.977 | 0.983 | 0.981 | 0.982 | 0.972 | 0.981 | 0.984 | 0.979 | 0.982 |
| 3 | 0.971 | 0.969 | 0.969 | 0.967 | 0.970 | 0.978 | 0.983 | 0.977 | 0.969 | 0.979 | 0.982 | 0.978 |
| 4 | 0.974 | 0.967 | 0.980 | 0.975 | 0.980 | 0.968 | 0.983 | 0.969 | 0.976 | 0.976 | 0.980 | 0.970 |
| 5 | 0.976 | 0.983 | 0.970 | 0.969 | 0.974 | 0.969 | 0.985 | 0.966 | 0.976 | 0.972 | 0.980 | 0.970 |
| 6 | 0.973 | 0.980 | 0.968 | 0.972 | 0.980 | 0.969 | 0.971 | 0.970 | 0.985 | 0.978 | 0.979 | 0.968 |
| 7 | 0.975 | 0.977 | 0.977 | 0.971 | 0.982 | 0.977 | 0.985 | 0.977 | 0.969 | 0.983 | 0.969 | 0.981 |
| 8 | 0.981 | 0.976 | 0.969 | 0.981 | 0.967 | 0.981 | 0.981 | 0.973 | 0.973 | 0.970 | 0.970 | 0.970 |
| 9 | 0.979 | 0.967 | 0.970 | 0.982 | 0.973 | 0.979 | 0.968 | 0.984 | 0.972 | 0.976 | 0.985 | 0.968 |
| 10 | 0.981 | 0.971 | 0.978 | 0.980 | 0.968 | 0.968 | 0.983 | 0.972 | 0.982 | 0.970 | 0.980 | 0.983 |
| 11 | 0.973 | 0.970 | 0.984 | 0.971 | 0.982 | 0.976 | 0.974 | 0.969 | 0.970 | 0.976 | 0.981 | 0.971 |
| 12 | 0.968 | 0.985 | 0.979 | 0.974 | 0.972 | 0.977 | 0.984 | 0.985 | 0.983 | 0.985 | 0.976 | 0.977 |
| 13 | 0.983 | 0.976 | 0.972 | 0.972 | 0.981 | 0.973 | 0.980 | 0.971 | 0.974 | 0.970 | 0.974 | 0.980 |
| 14 | 0.976 | 0.976 | 0.985 | 0.979 | 0.984 | 0.984 | 0.968 | 0.984 | 0.970 | 0.981 | 0.967 | 0.971 |
| 15 | 0.984 | 0.983 | 0.969 | 0.973 | 0.967 | 0.979 | 0.984 | 0.979 | 0.975 | 0.977 | 0.985 | 0.981 |
| 16 | 0.970 | 0.970 | 0.973 | 0.973 | 0.968 | 0.984 | 0.980 | 0.978 | 0.985 | 0.984 | 0.971 | 0.969 |
| 17 | 0.976 | 0.981 | 0.977 | 0.982 | 0.969 | 0.979 | 0.972 | 0.981 | 0.971 | 0.975 | 0.972 | 0.983 |
| 18 | 0.970 | 0.973 | 0.973 | 0.983 | 0.975 | 0.982 | 0.976 | 0.970 | 0.984 | 0.967 | 0.969 | 0.985 |
| 19 | 0.974 | 0.983 | 0.984 | 0.977 | 0.974 | 0.970 | 0.966 | 0.978 | 0.979 | 0.981 | 0.979 | 0.973 |
| 20 | 0.967 | 0.979 | 0.982 | 0.973 | 0.975 | 0.978 | 0.981 | 0.982 | 0.978 | 0.976 | 0.979 | 0.976 |
| 21 | 0.985 | 0.980 | 0.982 | 0.968 | 0.968 | 0.973 | 0.979 | 0.975 | 0.967 | 0.974 | 0.979 | 0.983 |
| 22 | 0.980 | 0.970 | 0.976 | 0.978 | 0.978 | 0.974 | 0.975 | 0.976 | 0.982 | 0.969 | 0.973 | 0.976 |
| 23 | 0.975 | 0.967 | 0.980 | 0.969 | 0.972 | 0.978 | 0.979 | 0.984 | 0.975 | 0.982 | 0.976 | 0.976 |
| 24 | 0.981 | 0.968 | 0.979 | 0.977 | 0.983 | 0.977 | 0.981 | 0.972 | 0.976 | 0.975 | 0.976 | 0.985 |
| 25 | 0.981 | 0.968 | 0.967 | 0.969 | 0.982 | 0.975 | 0.969 | 0.981 | 0.983 | 0.979 | 0.985 | 0.980 |
| 26 | 0.984 | 0.977 | 0.983 | 0.984 | 0.983 | 0.970 | 0.984 | 0.979 | 0.985 | 0.975 | 0.974 | 0.985 |
| 27 | 0.971 | 0.969 | 0.979 | 0.966 | 0.973 | 0.968 | 0.983 | 0.970 | 0.971 | 0.980 | 0.981 | 0.967 |
| 28 | 0.973 | 0.980 | 0.966 | 0.982 | 0.974 | 0.975 | 0.976 | 0.973 | 0.969 | 0.975 | 0.968 | 0.985 |
| 29 | 0.971 | 0.977 | 0.970 | 0.980 | 0.978 | 0.977 | 0.984 | 0.969 | 0.967 | 0.976 | 0.978 | 0.969 |
| 30 | 0.966 | 0.972 | 0.984 | 0.985 | 0.981 | 0.978 | 0.981 | 0.981 | 0.980 | 0.979 | 0.975 | 0.982 |
| 31 | 0.973 | 0.972 | 0.969 | 0.980 | 0.967 | 0.978 | 0.970 | 0.973 | 0.976 | 0.974 | 0.980 | 0.976 |
| 32 | 0.969 | 0.969 | 0.983 | 0.976 | 0.972 | 0.980 | 0.970 | 0.984 | 0.969 | 0.981 | 0.967 | 0.979 |
| 33 | 0.985 | 0.982 | 0.967 | 0.978 | 0.969 | 0.966 | 0.967 | 0.980 | 0.985 | 0.981 | 0.971 | 0.969 |
| 34 | 0.983 | 0.967 | 0.980 | 0.980 | 0.976 | 0.969 | 0.966 | 0.982 | 0.970 | 0.972 | 0.985 | 0.969 |
| 35 | 0.971 | 0.967 | 0.969 | 0.974 | 0.984 | 0.982 | 0.978 | 0.971 | 0.980 | 0.973 | 0.969 | 0.982 |
| 36 | 0.973 | 0.976 | 0.967 | 0.984 | 0.982 | 0.981 | 0.982 | 0.970 | 0.980 | 0.977 | 0.969 | 0.971 |

ตารางที่ 51 ความเที่ยงของแบสอบแต่ละฉบับ (ต่อ)

| ครั้งที่ | 50_10 | | | 60_8 | | | 70_3 | | | 80_4 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Test A | Test B1 | Test B2 | Test A | Test B1 | Test B2 | Test A | Test B1 | Test B2 | Test A | Test B1 | Test B2 |
| 37 | 0.984 | 0.970 | 0.973 | 0.974 | 0.973 | 0.984 | 0.967 | 0.968 | 0.985 | 0.979 | 0.975 | 0.975 |
| 38 | 0.977 | 0.982 | 0.969 | 0.982 | 0.981 | 0.978 | 0.978 | 0.968 | 0.979 | 0.971 | 0.977 | 0.966 |
| 39 | 0.979 | 0.968 | 0.967 | 0.970 | 0.982 | 0.974 | 0.974 | 0.969 | 0.973 | 0.985 | 0.973 | 0.982 |
| 40 | 0.980 | 0.975 | 0.968 | 0.975 | 0.970 | 0.979 | 0.967 | 0.971 | 0.974 | 0.979 | 0.985 | 0.979 |
| 41 | 0.975 | 0.969 | 0.975 | 0.972 | 0.969 | 0.981 | 0.975 | 0.982 | 0.978 | 0.969 | 0.979 | 0.967 |
| 42 | 0.975 | 0.981 | 0.977 | 0.981 | 0.986 | 0.982 | 0.968 | 0.976 | 0.980 | 0.981 | 0.985 | 0.966 |
| 43 | 0.980 | 0.968 | 0.984 | 0.981 | 0.977 | 0.981 | 0.982 | 0.978 | 0.972 | 0.977 | 0.967 | 0.981 |
| 44 | 0.976 | 0.985 | 0.977 | 0.984 | 0.982 | 0.980 | 0.971 | 0.973 | 0.967 | 0.977 | 0.976 | 0.969 |
| 45 | 0.969 | 0.975 | 0.984 | 0.977 | 0.982 | 0.967 | 0.984 | 0.978 | 0.967 | 0.969 | 0.977 | 0.978 |
| 46 | 0.969 | 0.981 | 0.978 | 0.969 | 0.980 | 0.977 | 0.967 | 0.968 | 0.971 | 0.976 | 0.979 | 0.971 |
| 47 | 0.979 | 0.980 | 0.980 | 0.969 | 0.980 | 0.980 | 0.973 | 0.973 | 0.973 | 0.972 | 0.982 | 0.973 |
| 48 | 0.966 | 0.982 | 0.976 | 0.966 | 0.983 | 0.979 | 0.978 | 0.982 | 0.973 | 0.984 | 0.978 | 0.975 |
| 49 | 0.985 | 0.969 | 0.974 | 0.981 | 0.973 | 0.972 | 0.970 | 0.969 | 0.983 | 0.979 | 0.973 | 0.980 |
| 50 | 0.984 | 0.968 | 0.976 | 0.978 | 0.968 | 0.980 | 0.971 | 0.966 | 0.985 | 0.984 | 0.975 | 0.976 |
| 51 | 0.977 | 0.983 | 0.983 | 0.984 | 0.979 | 0.979 | 0.980 | 0.966 | 0.971 | 0.982 | 0.975 | 0.975 |
| 52 | 0.980 | 0.984 | 0.973 | 0.984 | 0.968 | 0.982 | 0.982 | 0.984 | 0.984 | 0.979 | 0.978 | 0.983 |
| 53 | 0.983 | 0.971 | 0.967 | 0.978 | 0.971 | 0.970 | 0.983 | 0.969 | 0.981 | 0.967 | 0.974 | 0.981 |
| 54 | 0.976 | 0.983 | 0.975 | 0.972 | 0.967 | 0.977 | 0.974 | 0.980 | 0.981 | 0.985 | 0.971 | 0.977 |
| 55 | 0.975 | 0.977 | 0.969 | 0.972 | 0.982 | 0.982 | 0.983 | 0.971 | 0.973 | 0.969 | 0.973 | 0.981 |
| 56 | 0.973 | 0.972 | 0.980 | 0.979 | 0.978 | 0.976 | 0.981 | 0.971 | 0.975 | 0.969 | 0.976 | 0.984 |
| 57 | 0.971 | 0.970 | 0.979 | 0.983 | 0.976 | 0.966 | 0.984 | 0.971 | 0.969 | 0.966 | 0.976 | 0.982 |
| 58 | 0.982 | 0.969 | 0.985 | 0.973 | 0.969 | 0.977 | 0.985 | 0.972 | 0.968 | 0.966 | 0.972 | 0.982 |
| 59 | 0.979 | 0.974 | 0.980 | 0.985 | 0.978 | 0.984 | 0.975 | 0.984 | 0.974 | 0.976 | 0.982 | 0.967 |
| 60 | 0.981 | 0.980 | 0.984 | 0.969 | 0.976 | 0.969 | 0.976 | 0.971 | 0.974 | 0.985 | 0.967 | 0.969 |
| 61 | 0.968 | 0.969 | 0.980 | 0.974 | 0.976 | 0.973 | 0.977 | 0.982 | 0.969 | 0.973 | 0.980 | 0.985 |
| 62 | 0.968 | 0.982 | 0.969 | 0.971 | 0.968 | 0.967 | 0.974 | 0.970 | 0.976 | 0.983 | 0.975 | 0.971 |
| 63 | 0.973 | 0.985 | 0.974 | 0.981 | 0.981 | 0.984 | 0.969 | 0.983 | 0.980 | 0.974 | 0.985 | 0.976 |
| 64 | 0.971 | 0.976 | 0.970 | 0.977 | 0.968 | 0.983 | 0.979 | 0.983 | 0.979 | 0.970 | 0.974 | 0.985 |
| 65 | 0.970 | 0.980 | 0.967 | 0.985 | 0.985 | 0.970 | 0.980 | 0.974 | 0.981 | 0.972 | 0.967 | 0.974 |
| 66 | 0.975 | 0.974 | 0.980 | 0.972 | 0.976 | 0.978 | 0.976 | 0.983 | 0.980 | 0.971 | 0.969 | 0.976 |
| 67 | 0.975 | 0.982 | 0.973 | 0.981 | 0.981 | 0.971 | 0.980 | 0.977 | 0.976 | 0.979 | 0.967 | 0.968 |
| 68 | 0.985 | 0.977 | 0.984 | 0.975 | 0.981 | 0.968 | 0.982 | 0.981 | 0.978 | 0.967 | 0.986 | 0.967 |
| 69 | 0.977 | 0.984 | 0.971 | 0.978 | 0.969 | 0.985 | 0.969 | 0.983 | 0.974 | 0.983 | 0.968 | 0.979 |
| 70 | 0.972 | 0.967 | 0.968 | 0.977 | 0.976 | 0.971 | 0.985 | 0.970 | 0.975 | 0.976 | 0.983 | 0.974 |
| 71 | 0.976 | 0.974 | 0.984 | 0.980 | 0.967 | 0.983 | 0.985 | 0.975 | 0.971 | 0.978 | 0.975 | 0.979 |
| 72 | 0.982 | 0.976 | 0.985 | 0.977 | 0.966 | 0.984 | 0.970 | 0.973 | 0.973 | 0.977 | 0.971 | 0.983 |
| 73 | 0.984 | 0.974 | 0.968 | 0.980 | 0.970 | 0.985 | 0.977 | 0.975 | 0.974 | 0.977 | 0.966 | 0.966 |

ตารางที่ 51 ความเที่ยงของแบสอบแต่ละฉบับ (ต่อ)

| ครั้งที่ | 50_10 | | | 60_8 | | | 70_3 | | | 80_4 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Test A | Test B1 | Test B2 | Test A | Test B1 | Test B2 | Test A | Test B1 | Test B2 | Test A | Test B1 | Test B2 |
| 74 | 0.976 | 0.969 | 0.979 | 0.969 | 0.970 | 0.970 | 0.975 | 0.982 | 0.971 | 0.970 | 0.976 | 0.974 |
| 75 | 0.972 | 0.979 | 0.976 | 0.975 | 0.979 | 0.970 | 0.970 | 0.984 | 0.983 | 0.971 | 0.977 | 0.982 |
| 76 | 0.976 | 0.970 | 0.981 | 0.968 | 0.979 | 0.981 | 0.966 | 0.975 | 0.969 | 0.976 | 0.966 | 0.972 |
| 77 | 0.966 | 0.983 | 0.967 | 0.971 | 0.970 | 0.972 | 0.967 | 0.971 | 0.973 | 0.974 | 0.979 | 0.970 |
| 78 | 0.971 | 0.983 | 0.985 | 0.975 | 0.981 | 0.981 | 0.976 | 0.978 | 0.978 | 0.980 | 0.982 | 0.977 |
| 79 | 0.971 | 0.970 | 0.970 | 0.983 | 0.983 | 0.975 | 0.977 | 0.970 | 0.975 | 0.983 | 0.972 | 0.973 |
| 80 | 0.976 | 0.982 | 0.967 | 0.970 | 0.975 | 0.980 | 0.973 | 0.985 | 0.974 | 0.983 | 0.970 | 0.983 |
| 81 | 0.978 | 0.983 | 0.976 | 0.971 | 0.975 | 0.976 | 0.981 | 0.966 | 0.980 | 0.969 | 0.968 | 0.974 |
| 82 | 0.977 | 0.967 | 0.973 | 0.983 | 0.971 | 0.972 | 0.970 | 0.981 | 0.973 | 0.969 | 0.982 | 0.981 |
| 83 | 0.970 | 0.977 | 0.969 | 0.984 | 0.978 | 0.974 | 0.969 | 0.983 | 0.983 | 0.981 | 0.976 | 0.981 |
| 84 | 0.967 | 0.972 | 0.969 | 0.983 | 0.980 | 0.974 | 0.985 | 0.966 | 0.976 | 0.971 | 0.972 | 0.981 |
| 85 | 0.985 | 0.977 | 0.981 | 0.974 | 0.977 | 0.968 | 0.969 | 0.985 | 0.985 | 0.977 | 0.972 | 0.973 |
| 86 | 0.966 | 0.972 | 0.969 | 0.977 | 0.977 | 0.983 | 0.977 | 0.981 | 0.976 | 0.975 | 0.985 | 0.985 |
| 87 | 0.980 | 0.984 | 0.973 | 0.984 | 0.974 | 0.985 | 0.982 | 0.973 | 0.968 | 0.980 | 0.984 | 0.969 |
| 88 | 0.974 | 0.971 | 0.980 | 0.981 | 0.976 | 0.982 | 0.984 | 0.968 | 0.967 | 0.978 | 0.977 | 0.979 |
| 89 | 0.976 | 0.969 | 0.977 | 0.985 | 0.985 | 0.979 | 0.978 | 0.969 | 0.979 | 0.973 | 0.968 | 0.983 |
| 90 | 0.975 | 0.976 | 0.974 | 0.982 | 0.973 | 0.982 | 0.977 | 0.966 | 0.979 | 0.982 | 0.969 | 0.967 |
| 91 | 0.968 | 0.979 | 0.985 | 0.971 | 0.971 | 0.975 | 0.981 | 0.981 | 0.972 | 0.981 | 0.969 | 0.971 |
| 92 | 0.971 | 0.985 | 0.972 | 0.981 | 0.980 | 0.984 | 0.976 | 0.981 | 0.977 | 0.972 | 0.970 | 0.969 |
| 93 | 0.985 | 0.969 | 0.983 | 0.972 | 0.975 | 0.984 | 0.978 | 0.980 | 0.983 | 0.985 | 0.973 | 0.979 |
| 94 | 0.972 | 0.973 | 0.977 | 0.971 | 0.977 | 0.974 | 0.970 | 0.969 | 0.982 | 0.970 | 0.984 | 0.973 |
| 95 | 0.972 | 0.983 | 0.974 | 0.967 | 0.971 | 0.981 | 0.984 | 0.967 | 0.981 | 0.970 | 0.970 | 0.979 |
| 96 | 0.982 | 0.975 | 0.974 | 0.969 | 0.980 | 0.981 | 0.982 | 0.967 | 0.968 | 0.973 | 0.985 | 0.984 |
| 97 | 0.972 | 0.969 | 0.979 | 0.983 | 0.978 | 0.974 | 0.981 | 0.981 | 0.967 | 0.982 | 0.976 | 0.966 |
| 98 | 0.982 | 0.973 | 0.985 | 0.977 | 0.969 | 0.973 | 0.967 | 0.971 | 0.966 | 0.977 | 0.976 | 0.970 |
| 99 | 0.975 | 0.968 | 0.966 | 0.974 | 0.984 | 0.967 | 0.983 | 0.968 | 0.968 | 0.971 | 0.980 | 0.981 |
| 100 | 0.971 | 0.972 | 0.981 | 0.978 | 0.967 | 0.978 | 0.983 | 0.984 | 0.973 | 0.967 | 0.982 | 0.970 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

# ภาคผนวก ค

## ค่าสัมประสิทธิ์การปรับเทียบคะแนนด้วยวิธีโค้งคุณลักษณะ (Heabara) จำแนกตามสัดส่วนของแบบสอบ และรูปแบบการเก็บรวบรวมข้อมูล

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 52 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 50 **และ** PS 10 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มเท่าเทียมโดย**ใช้ข้อสอบร่วม

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 1 | 0.893 | 0.121 | 0.897 | 0.114 | 0.967 | 0.104 | 0.831 | 0.164 | 0.844 | 0.118 | 0.860 | 0.116 |
| 2 | 0.935 | 0.119 | 0.832 | 0.138 | 0.847 | 0.108 | 0.869 | 0.145 | 0.973 | 0.121 | 0.967 | 0.117 |
| 3 | 0.970 | 0.127 | 0.859 | 0.115 | 0.911 | 0.113 | 0.970 | 0.159 | 0.960 | 0.121 | 0.854 | 0.132 |
| 4 | 0.972 | 0.132 | 0.837 | 0.125 | 0.828 | 0.177 | 0.961 | 0.165 | 0.987 | 0.130 | 0.818 | 0.122 |
| 5 | 0.930 | 0.127 | 0.849 | 0.129 | 0.865 | 0.162 | 0.907 | 0.15 | 0.892 | 0.127 | 0.964 | 0.122 |
| 6 | 0.979 | 0.114 | 0.968 | 0.125 | 0.862 | 0.118 | 0.960 | 0.178 | 0.864 | 0.121 | 0.912 | 0.129 |
| 7 | 0.928 | 0.139 | 0.864 | 0.112 | 0.922 | 0.176 | 0.817 | 0.15 | 0.884 | 0.135 | 0.979 | 0.121 |
| 8 | 0.866 | 0.118 | 0.833 | 0.128 | 0.834 | 0.171 | 0.932 | 0.142 | 0.919 | 0.116 | 0.952 | 0.117 |
| 9 | 0.956 | 0.118 | 0.864 | 0.116 | 0.939 | 0.188 | 0.955 | 0.159 | 0.817 | 0.124 | 0.972 | 0.122 |
| 10 | 0.960 | 0.137 | 0.947 | 0.124 | 0.919 | 0.13 | 0.974 | 0.184 | 0.930 | 0.133 | 0.911 | 0.130 |
| 11 | 0.850 | 0.137 | 0.896 | 0.136 | 0.818 | 0.165 | 0.860 | 0.158 | 0.859 | 0.117 | 0.960 | 0.120 |
| 12 | 0.864 | 0.113 | 0.875 | 0.138 | 0.816 | 0.165 | 0.934 | 0.167 | 0.855 | 0.122 | 0.941 | 0.124 |
| 13 | 0.858 | 0.139 | 0.890 | 0.113 | 0.821 | 0.177 | 0.947 | 0.165 | 0.952 | 0.117 | 0.900 | 0.125 |
| 14 | 0.928 | 0.121 | 0.981 | 0.134 | 0.971 | 0.185 | 0.876 | 0.148 | 0.916 | 0.120 | 0.965 | 0.129 |
| 15 | 0.936 | 0.112 | 0.846 | 0.134 | 0.919 | 0.13 | 0.873 | 0.171 | 0.979 | 0.126 | 0.980 | 0.129 |
| 16 | 0.845 | 0.121 | 0.910 | 0.122 | 0.985 | 0.164 | 0.923 | 0.156 | 0.827 | 0.129 | 0.899 | 0.132 |
| 17 | 0.918 | 0.138 | 0.829 | 0.130 | 0.835 | 0.162 | 0.936 | 0.183 | 0.948 | 0.129 | 0.849 | 0.132 |
| 18 | 0.949 | 0.115 | 0.899 | 0.138 | 0.961 | 0.125 | 0.907 | 0.151 | 0.888 | 0.130 | 0.828 | 0.134 |
| 19 | 0.830 | 0.132 | 0.895 | 0.132 | 0.896 | 0.112 | 0.829 | 0.168 | 0.881 | 0.134 | 0.924 | 0.130 |
| 20 | 0.820 | 0.118 | 0.878 | 0.129 | 0.921 | 0.119 | 0.949 | 0.16 | 0.855 | 0.113 | 0.959 | 0.117 |
| 21 | 0.861 | 0.122 | 0.963 | 0.135 | 0.958 | 0.166 | 0.873 | 0.145 | 0.844 | 0.138 | 0.864 | 0.118 |
| 22 | 0.869 | 0.114 | 0.944 | 0.126 | 0.840 | 0.176 | 0.923 | 0.144 | 0.958 | 0.126 | 0.935 | 0.115 |
| 23 | 0.832 | 0.133 | 0.842 | 0.138 | 0.899 | 0.15 | 0.828 | 0.138 | 0.924 | 0.138 | 0.875 | 0.132 |
| 24 | 0.886 | 0.128 | 0.838 | 0.135 | 0.827 | 0.169 | 0.853 | 0.15 | 0.862 | 0.127 | 0.842 | 0.121 |
| 25 | 0.837 | 0.112 | 0.867 | 0.115 | 0.853 | 0.175 | 0.815 | 0.167 | 0.968 | 0.130 | 0.841 | 0.138 |
| 26 | 0.983 | 0.136 | 0.894 | 0.123 | 0.986 | 0.142 | 0.977 | 0.172 | 0.881 | 0.132 | 0.965 | 0.136 |
| 27 | 0.926 | 0.116 | 0.844 | 0.126 | 0.843 | 0.125 | 0.894 | 0.163 | 0.827 | 0.126 | 0.950 | 0.135 |
| 28 | 0.887 | 0.116 | 0.971 | 0.124 | 0.925 | 0.112 | 0.872 | 0.14 | 0.977 | 0.124 | 0.855 | 0.113 |
| 29 | 0.892 | 0.125 | 0.951 | 0.128 | 0.980 | 0.132 | 0.833 | 0.182 | 0.858 | 0.131 | 0.819 | 0.131 |
| 30 | 0.857 | 0.119 | 0.950 | 0.124 | 0.931 | 0.126 | 0.912 | 0.182 | 0.894 | 0.131 | 0.851 | 0.114 |
| 31 | 0.878 | 0.136 | 0.831 | 0.120 | 0.936 | 0.178 | 0.819 | 0.138 | 0.923 | 0.124 | 0.841 | 0.115 |
| 32 | 0.975 | 0.128 | 0.942 | 0.134 | 0.904 | 0.14 | 0.914 | 0.161 | 0.849 | 0.127 | 0.857 | 0.138 |
| 33 | 0.839 | 0.118 | 0.947 | 0.130 | 0.825 | 0.139 | 0.958 | 0.184 | 0.974 | 0.114 | 0.974 | 0.118 |
| 34 | 0.942 | 0.118 | 0.936 | 0.132 | 0.894 | 0.123 | 0.862 | 0.161 | 0.877 | 0.137 | 0.981 | 0.133 |
| 35 | 0.895 | 0.136 | 0.818 | 0.134 | 0.877 | 0.151 | 0.904 | 0.157 | 0.890 | 0.118 | 0.947 | 0.124 |
| 36 | 0.841 | 0.118 | 0.842 | 0.114 | 0.949 | 0.17 | 0.902 | 0.14 | 0.900 | 0.117 | 0.821 | 0.136 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 52 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 50 **และ** PS 10 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มเท่าเทียมโดย**ใช้ข้อสอบร่วม (ต่อ)

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 37 | 0.843 | 0.136 | 0.847 | 0.130 | 0.957 | 0.136 | 0.829 | 0.15 | 0.836 | 0.120 | 0.965 | 0.132 |
| 38 | 0.963 | 0.136 | 0.933 | 0.134 | 0.944 | 0.165 | 0.927 | 0.188 | 0.941 | 0.137 | 0.883 | 0.129 |
| 39 | 0.951 | 0.135 | 0.920 | 0.133 | 0.984 | 0.134 | 0.856 | 0.15 | 0.920 | 0.118 | 0.969 | 0.138 |
| 40 | 0.876 | 0.128 | 0.981 | 0.112 | 0.933 | 0.183 | 0.820 | 0.171 | 0.862 | 0.117 | 0.925 | 0.127 |
| 41 | 0.860 | 0.121 | 0.849 | 0.130 | 0.823 | 0.141 | 0.935 | 0.154 | 0.830 | 0.128 | 0.872 | 0.129 |
| 42 | 0.934 | 0.121 | 0.912 | 0.127 | 0.896 | 0.131 | 0.900 | 0.17 | 0.851 | 0.127 | 0.890 | 0.135 |
| 43 | 0.842 | 0.132 | 0.895 | 0.126 | 0.962 | 0.138 | 0.964 | 0.182 | 0.878 | 0.118 | 0.954 | 0.138 |
| 44 | 0.970 | 0.129 | 0.857 | 0.123 | 0.846 | 0.134 | 0.973 | 0.151 | 0.943 | 0.113 | 0.824 | 0.128 |
| 45 | 0.971 | 0.118 | 0.840 | 0.137 | 0.897 | 0.161 | 0.972 | 0.156 | 0.944 | 0.137 | 0.970 | 0.120 |
| 46 | 0.838 | 0.129 | 0.929 | 0.130 | 0.938 | 0.119 | 0.891 | 0.187 | 0.860 | 0.115 | 0.912 | 0.122 |
| 47 | 0.976 | 0.120 | 0.923 | 0.135 | 0.951 | 0.141 | 0.908 | 0.18 | 0.867 | 0.117 | 0.838 | 0.134 |
| 48 | 0.904 | 0.138 | 0.920 | 0.128 | 0.926 | 0.154 | 0.919 | 0.186 | 0.967 | 0.126 | 0.851 | 0.113 |
| 49 | 0.901 | 0.125 | 0.937 | 0.115 | 0.902 | 0.118 | 0.818 | 0.139 | 0.940 | 0.127 | 0.867 | 0.128 |
| 50 | 0.866 | 0.121 | 0.852 | 0.122 | 0.920 | 0.123 | 0.967 | 0.149 | 0.882 | 0.117 | 0.979 | 0.118 |
| 51 | 0.962 | 0.137 | 0.898 | 0.139 | 0.891 | 0.169 | 0.843 | 0.147 | 0.818 | 0.124 | 0.884 | 0.125 |
| 52 | 0.950 | 0.118 | 0.944 | 0.114 | 0.977 | 0.153 | 0.841 | 0.147 | 0.940 | 0.124 | 0.867 | 0.125 |
| 53 | 0.958 | 0.130 | 0.885 | 0.115 | 0.836 | 0.159 | 0.880 | 0.142 | 0.899 | 0.130 | 0.911 | 0.138 |
| 54 | 0.837 | 0.114 | 0.984 | 0.120 | 0.899 | 0.177 | 0.828 | 0.167 | 0.895 | 0.124 | 0.964 | 0.116 |
| 55 | 0.836 | 0.130 | 0.952 | 0.138 | 0.967 | 0.121 | 0.818 | 0.183 | 0.973 | 0.114 | 0.973 | 0.136 |
| 56 | 0.977 | 0.138 | 0.845 | 0.130 | 0.928 | 0.152 | 0.908 | 0.157 | 0.963 | 0.128 | 0.859 | 0.116 |
| 57 | 0.947 | 0.131 | 0.864 | 0.112 | 0.837 | 0.18 | 0.915 | 0.154 | 0.959 | 0.123 | 0.890 | 0.124 |
| 58 | 0.872 | 0.128 | 0.843 | 0.125 | 0.860 | 0.173 | 0.910 | 0.186 | 0.888 | 0.135 | 0.862 | 0.123 |
| 59 | 0.935 | 0.131 | 0.826 | 0.136 | 0.890 | 0.153 | 0.912 | 0.137 | 0.867 | 0.119 | 0.954 | 0.114 |
| 60 | 0.965 | 0.129 | 0.853 | 0.134 | 0.923 | 0.107 | 0.852 | 0.169 | 0.849 | 0.115 | 0.902 | 0.133 |
| 61 | 0.854 | 0.132 | 0.963 | 0.120 | 0.915 | 0.183 | 0.973 | 0.177 | 0.917 | 0.139 | 0.846 | 0.124 |
| 62 | 0.898 | 0.128 | 0.903 | 0.137 | 0.826 | 0.135 | 0.946 | 0.162 | 0.833 | 0.139 | 0.948 | 0.117 |
| 63 | 0.976 | 0.131 | 0.825 | 0.114 | 0.946 | 0.11 | 0.906 | 0.14 | 0.934 | 0.120 | 0.875 | 0.133 |
| 64 | 0.879 | 0.136 | 0.877 | 0.136 | 0.934 | 0.113 | 0.949 | 0.165 | 0.986 | 0.136 | 0.820 | 0.135 |
| 65 | 0.877 | 0.124 | 0.872 | 0.117 | 0.902 | 0.114 | 0.828 | 0.142 | 0.910 | 0.119 | 0.892 | 0.138 |
| 66 | 0.888 | 0.122 | 0.943 | 0.115 | 0.841 | 0.165 | 0.979 | 0.157 | 0.887 | 0.132 | 0.938 | 0.131 |
| 67 | 0.875 | 0.135 | 0.936 | 0.115 | 0.844 | 0.131 | 0.855 | 0.172 | 0.884 | 0.130 | 0.910 | 0.121 |
| 68 | 0.933 | 0.112 | 0.927 | 0.117 | 0.942 | 0.18 | 0.831 | 0.187 | 0.835 | 0.123 | 0.916 | 0.117 |
| 69 | 0.819 | 0.124 | 0.819 | 0.125 | 0.959 | 0.167 | 0.832 | 0.164 | 0.934 | 0.120 | 0.903 | 0.114 |
| 70 | 0.971 | 0.124 | 0.846 | 0.112 | 0.818 | 0.15 | 0.867 | 0.168 | 0.944 | 0.115 | 0.916 | 0.138 |
| 71 | 0.941 | 0.135 | 0.833 | 0.117 | 0.850 | 0.122 | 0.847 | 0.141 | 0.880 | 0.118 | 0.816 | 0.132 |
| 72 | 0.857 | 0.134 | 0.914 | 0.132 | 0.974 | 0.126 | 0.894 | 0.182 | 0.907 | 0.132 | 0.858 | 0.131 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 52 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 50 **และ** PS 10 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มเท่าเทียมโดย**ใช้ข้อสอบร่วม (ต่อ)

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 73 | 0.894 | 0.139 | 0.834 | 0.125 | 0.924 | 0.141 | 0.914 | 0.16 | 0.924 | 0.189 | 0.937 | 0.131 |
| 74 | 0.901 | 0.119 | 0.921 | 0.116 | 0.957 | 0.188 | 0.874 | 0.188 | 0.912 | 0.126 | 0.931 | 0.131 |
| 75 | 0.897 | 0.118 | 0.973 | 0.136 | 0.898 | 0.145 | 0.965 | 0.178 | 0.861 | 0.137 | 0.974 | 0.136 |
| 76 | 0.935 | 0.125 | 0.984 | 0.135 | 0.856 | 0.11 | 0.851 | 0.179 | 0.837 | 0.135 | 0.844 | 0.123 |
| 77 | 0.863 | 0.138 | 0.911 | 0.114 | 0.902 | 0.144 | 0.836 | 0.187 | 0.871 | 0.121 | 0.843 | 0.124 |
| 78 | 0.936 | 0.127 | 0.900 | 0.137 | 0.816 | 0.165 | 0.913 | 0.173 | 0.871 | 0.175 | 0.976 | 0.124 |
| 79 | 0.838 | 0.136 | 0.971 | 0.134 | 0.944 | 0.108 | 0.966 | 0.143 | 0.944 | 0.130 | 0.902 | 0.120 |
| 80 | 0.848 | 0.112 | 0.888 | 0.135 | 0.854 | 0.122 | 0.888 | 0.151 | 0.930 | 0.136 | 0.914 | 0.128 |
| 81 | 0.926 | 0.113 | 0.868 | 0.123 | 0.879 | 0.164 | 0.823 | 0.139 | 0.816 | 0.132 | 0.939 | 0.133 |
| 82 | 0.874 | 0.129 | 0.861 | 0.125 | 0.845 | 0.137 | 0.872 | 0.189 | 0.921 | 0.195 | 0.840 | 0.136 |
| 83 | 0.862 | 0.114 | 0.945 | 0.135 | 0.856 | 0.109 | 0.847 | 0.174 | 0.834 | 0.156 | 0.899 | 0.134 |
| 84 | 0.961 | 0.128 | 0.921 | 0.139 | 0.822 | 0.118 | 0.980 | 0.167 | 0.833 | 0.134 | 0.828 | 0.126 |
| 85 | 0.938 | 0.116 | 0.946 | 0.121 | 0.874 | 0.116 | 0.874 | 0.165 | 0.886 | 0.133 | 0.843 | 0.129 |
| 86 | 0.855 | 0.138 | 0.824 | 0.121 | 0.860 | 0.129 | 0.816 | 0.182 | 0.911 | 0.145 | 0.864 | 0.130 |
| 87 | 0.905 | 0.120 | 0.846 | 0.124 | 0.852 | 0.174 | 0.946 | 0.157 | 0.829 | 0.128 | 0.984 | 0.134 |
| 88 | 0.913 | 0.124 | 0.855 | 0.122 | 0.932 | 0.179 | 0.939 | 0.174 | 0.824 | 0.131 | 0.838 | 0.113 |
| 89 | 0.880 | 0.117 | 0.897 | 0.129 | 0.985 | 0.163 | 0.962 | 0.178 | 0.852 | 0.155 | 0.925 | 0.139 |
| 90 | 0.823 | 0.117 | 0.904 | 0.130 | 0.970 | 0.117 | 0.874 | 0.158 | 0.887 | 0.133 | 0.846 | 0.113 |
| 91 | 0.977 | 0.289 | 0.858 | 0.136 | 0.957 | 0.111 | 0.882 | 0.142 | 0.971 | 0.133 | 0.911 | 0.121 |
| 92 | 0.818 | 0.135 | 0.970 | 0.137 | 0.904 | 0.155 | 0.844 | 0.178 | 0.918 | 0.132 | 0.961 | 0.129 |
| 93 | 0.846 | 0.022 | 0.830 | 0.138 | 0.943 | 0.186 | 0.945 | 0.142 | 0.945 | 0.130 | 0.884 | 0.137 |
| 94 | 0.893 | 0.118 | 0.819 | 0.131 | 0.830 | 0.134 | 0.839 | 0.138 | 0.905 | 0.210 | 0.885 | 0.119 |
| 95 | 0.910 | 0.135 | 0.976 | 0.134 | 0.844 | 0.182 | 0.943 | 0.167 | 0.932 | 0.122 | 0.913 | 0.135 |
| 96 | 0.973 | 0.213 | 0.902 | 0.124 | 0.826 | 0.131 | 0.832 | 0.159 | 0.864 | 0.136 | 0.984 | 0.133 |
| 97 | 0.726 | 0.124 | 0.912 | 0.125 | 0.925 | 0.152 | 0.856 | 0.143 | 0.980 | 0.131 | 0.942 | 0.131 |
| 98 | 0.907 | 0.126 | 0.817 | 0.120 | 0.822 | 0.126 | 0.836 | 0.15 | 0.882 | 0.225 | 0.887 | 0.133 |
| 99 | 0.850 | 0.021 | 0.982 | 0.116 | 0.922 | 0.111 | 0.939 | 0.159 | 0.971 | 0.132 | 0.940 | 0.113 |
| 100 | 0.908 | 0.118 | 0.926 | 0.126 | 0.984 | 0.122 | 0.924 | 0.172 | 0.881 | 0.134 | 0.852 | 0.119 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 53 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง **DS 60 และ PS 8** ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มเท่าเทียม**โดยใช้ข้อสอบร่วม

| ข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 1 | 0.859 | 0.124 | 0.935 | 0.134 | 0.842 | 0.13 | 0.922 | 0.12 | 0.858 | 0.131 | 0.899 | 0.091 |
| 2 | 0.894 | 0.134 | 0.842 | 0.129 | 0.882 | 0.111 | 0.926 | 0.139 | 0.878 | 0.137 | 0.829 | 0.039 |
| 3 | 0.870 | 0.116 | 0.843 | 0.129 | 0.925 | 0.107 | 0.974 | 0.133 | 0.891 | 0.136 | 0.817 | 0.103 |
| 4 | 0.948 | 0.118 | 0.891 | 0.13 | 0.899 | 0.093 | 0.895 | 0.118 | 0.877 | 0.129 | 0.903 | 0.033 |
| 5 | 0.818 | 0.135 | 0.884 | 0.118 | 0.958 | 0.105 | 0.888 | 0.114 | 0.965 | 0.119 | 0.960 | 0.036 |
| 6 | 0.883 | 0.121 | 0.825 | 0.123 | 0.825 | 0.109 | 0.938 | 0.113 | 0.893 | 0.131 | 0.860 | 0.05 |
| 7 | 0.832 | 0.133 | 0.927 | 0.125 | 0.953 | 0.12 | 0.902 | 0.129 | 0.909 | 0.125 | 0.975 | 0.062 |
| 8 | 0.975 | 0.135 | 0.926 | 0.128 | 0.936 | 0.131 | 0.920 | 0.136 | 0.852 | 0.123 | 0.917 | 0.108 |
| 9 | 0.866 | 0.117 | 0.821 | 0.138 | 0.864 | 0.091 | 0.914 | 0.133 | 0.820 | 0.127 | 0.868 | 0.083 |
| 10 | 0.907 | 0.118 | 0.933 | 0.131 | 0.926 | 0.094 | 0.944 | 0.115 | 0.828 | 0.117 | 0.818 | 0.107 |
| 11 | 0.987 | 0.133 | 0.892 | 0.117 | 0.894 | 0.123 | 0.844 | 0.119 | 0.926 | 0.121 | 0.903 | 0.109 |
| 12 | 0.903 | 0.134 | 0.893 | 0.117 | 0.942 | 0.127 | 0.907 | 0.113 | 0.860 | 0.12 | 0.921 | 0.026 |
| 13 | 0.979 | 0.133 | 0.858 | 0.127 | 0.889 | 0.1 | 0.864 | 0.131 | 0.836 | 0.134 | 0.849 | 0.077 |
| 14 | 0.919 | 0.13 | 0.881 | 0.124 | 0.937 | 0.131 | 0.890 | 0.128 | 0.872 | 0.137 | 0.931 | 0.115 |
| 15 | 0.971 | 0.135 | 0.915 | 0.134 | 0.985 | 0.102 | 0.976 | 0.12 | 0.905 | 0.134 | 0.952 | 0.043 |
| 16 | 0.929 | 0.115 | 0.918 | 0.136 | 0.859 | 0.113 | 0.987 | 0.123 | 0.977 | 0.113 | 0.857 | 0.078 |
| 17 | 0.834 | 0.122 | 0.883 | 0.137 | 0.957 | 0.108 | 0.958 | 0.115 | 0.886 | 0.121 | 0.838 | 0.093 |
| 18 | 0.911 | 0.132 | 0.886 | 0.137 | 0.971 | 0.127 | 0.856 | 0.138 | 0.944 | 0.122 | 0.935 | 0.077 |
| 19 | 0.823 | 0.121 | 0.874 | 0.132 | 0.896 | 0.112 | 0.896 | 0.128 | 0.926 | 0.129 | 0.893 | 0.06 |
| 20 | 0.908 | 0.134 | 0.820 | 0.123 | 0.956 | 0.101 | 0.959 | 0.126 | 0.929 | 0.12 | 0.863 | 0.049 |
| 21 | 0.961 | 0.127 | 0.854 | 0.131 | 0.877 | 0.106 | 0.861 | 0.113 | 0.852 | 0.13 | 0.854 | 0.036 |
| 22 | 0.892 | 0.133 | 0.957 | 0.131 | 0.969 | 0.122 | 0.868 | 0.126 | 0.936 | 0.119 | 0.830 | 0.106 |
| 23 | 0.864 | 0.118 | 0.922 | 0.117 | 0.819 | 0.129 | 0.882 | 0.119 | 0.984 | 0.13 | 0.917 | 0.122 |
| 24 | 0.873 | 0.154 | 0.968 | 0.134 | 0.968 | 0.105 | 0.983 | 0.131 | 0.873 | 0.118 | 0.855 | 0.039 |
| 25 | 0.845 | 0.132 | 0.961 | 0.118 | 0.851 | 0.124 | 0.881 | 0.115 | 0.876 | 0.139 | 0.905 | 0.061 |
| 26 | 0.918 | 0.122 | 0.924 | 0.117 | 0.891 | 0.132 | 0.917 | 0.127 | 0.818 | 0.118 | 0.983 | 0.07 |
| 27 | 0.941 | 0.125 | 0.896 | 0.116 | 0.880 | 0.103 | 0.966 | 0.125 | 0.829 | 0.115 | 0.957 | 0.034 |
| 28 | 0.907 | 0.117 | 0.954 | 0.116 | 0.832 | 0.13 | 0.936 | 0.12 | 0.841 | 0.115 | 0.824 | 0.116 |
| 29 | 0.848 | 0.13 | 0.972 | 0.137 | 0.822 | 0.122 | 0.968 | 0.137 | 0.964 | 0.138 | 0.940 | 0.117 |
| 30 | 0.982 | 0.126 | 0.842 | 0.115 | 0.961 | 0.104 | 0.863 | 0.133 | 0.971 | 0.131 | 0.914 | 0.054 |
| 31 | 0.913 | 0.125 | 0.979 | 0.119 | 0.935 | 0.108 | 0.898 | 0.118 | 0.883 | 0.131 | 0.922 | 0.049 |
| 32 | 0.887 | 0.133 | 0.889 | 0.116 | 0.945 | 0.131 | 0.844 | 0.117 | 0.909 | 0.125 | 0.843 | 0.097 |
| 33 | 0.918 | 0.126 | 0.832 | 0.119 | 0.896 | 0.132 | 0.967 | 0.12 | 0.979 | 0.127 | 0.842 | 0.119 |
| 34 | 0.955 | 0.116 | 0.938 | 0.129 | 0.915 | 0.102 | 0.879 | 0.137 | 0.828 | 0.114 | 0.861 | 0.045 |
| 35 | 0.834 | 0.119 | 0.851 | 0.122 | 0.929 | 0.094 | 0.852 | 0.121 | 0.962 | 0.123 | 0.964 | 0.109 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 53 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง **DS 60 และ PS 8** ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มเท่าเทียม**โดยใช้ข้อสอบร่วม (ต่อ)

| ข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 36 | 0.963 | 0.136 | 0.840 | 0.116 | 0.923 | 0.109 | 0.844 | 0.127 | 0.905 | 0.121 | 0.938 | 0.04 |
| 37 | 0.876 | 0.125 | 0.949 | 0.125 | 0.948 | 0.121 | 0.909 | 0.115 | 0.847 | 0.121 | 0.973 | 0.042 |
| 38 | 0.910 | 0.126 | 0.896 | 0.119 | 0.840 | 0.095 | 0.925 | 0.129 | 0.837 | 0.139 | 0.820 | 0.051 |
| 39 | 0.838 | 0.134 | 0.949 | 0.136 | 0.953 | 0.109 | 0.964 | 0.136 | 0.938 | 0.116 | 0.936 | 0.097 |
| 40 | 0.870 | 0.129 | 0.836 | 0.13 | 0.979 | 0.099 | 0.840 | 0.114 | 0.840 | 0.131 | 0.831 | 0.079 |
| 41 | 0.926 | 0.126 | 0.815 | 0.137 | 0.955 | 0.111 | 1.236 | 0.119 | 0.979 | 0.134 | 0.880 | 0.039 |
| 42 | 0.975 | 0.125 | 0.920 | 0.121 | 0.857 | 0.126 | 0.937 | 0.129 | 0.849 | 0.118 | 0.826 | 0.04 |
| 43 | 0.897 | 0.118 | 0.851 | 0.134 | 0.827 | 0.117 | 0.914 | 0.134 | 0.923 | 0.137 | 0.916 | 0.12 |
| 44 | 0.833 | 0.134 | 0.914 | 0.128 | 0.933 | 0.102 | 0.894 | 0.134 | 0.864 | 0.121 | 0.939 | 0.074 |
| 45 | 0.832 | 0.12 | 0.932 | 0.122 | 0.861 | 0.115 | 0.823 | 0.133 | 0.844 | 0.118 | 0.966 | 0.092 |
| 46 | 0.884 | 0.123 | 0.935 | 0.132 | 0.845 | 0.123 | 1.056 | 0.115 | 0.963 | 0.119 | 0.932 | 0.097 |
| 47 | 0.960 | 0.126 | 0.958 | 0.131 | 0.929 | 0.109 | 0.934 | 0.128 | 0.977 | 0.117 | 0.928 | 0.118 |
| 48 | 0.971 | 0.129 | 0.864 | 0.114 | 0.821 | 0.098 | 0.973 | 0.115 | 0.931 | 0.133 | 0.820 | 0.083 |
| 49 | 0.852 | 0.117 | 0.823 | 0.136 | 0.828 | 0.091 | 0.903 | 0.121 | 0.855 | 0.126 | 0.929 | 0.078 |
| 50 | 0.946 | 0.117 | 0.875 | 0.129 | 0.890 | 0.135 | 0.957 | 0.122 | 0.857 | 0.13 | 0.891 | 0.027 |
| 51 | 0.971 | 0.125 | 0.968 | 0.138 | 0.834 | 0.125 | 0.954 | 0.123 | 0.872 | 0.12 | 0.861 | 0.099 |
| 52 | 0.909 | 0.129 | 0.876 | 0.127 | 0.855 | 0.092 | 0.905 | 0.116 | 0.839 | 0.129 | 0.945 | 0.069 |
| 53 | 0.829 | 0.135 | 0.877 | 0.129 | 0.836 | 0.135 | 0.827 | 0.137 | 0.851 | 0.125 | 0.832 | 0.065 |
| 54 | 0.824 | 0.127 | 0.830 | 0.132 | 0.928 | 0.119 | 0.868 | 0.13 | 0.891 | 0.122 | 0.913 | 0.051 |
| 55 | 0.919 | 0.121 | 0.960 | 0.128 | 0.919 | 0.11 | 0.838 | 0.126 | 0.914 | 0.126 | 0.959 | 0.09 |
| 56 | 0.815 | 0.13 | 0.853 | 0.114 | 0.904 | 0.101 | 0.898 | 0.112 | 0.837 | 0.113 | 0.890 | 0.124 |
| 57 | 0.872 | 0.127 | 0.963 | 0.119 | 0.853 | 0.133 | 0.925 | 0.135 | 0.869 | 0.118 | 0.836 | 0.035 |
| 58 | 0.934 | 0.129 | 0.906 | 0.131 | 0.939 | 0.11 | 0.903 | 0.123 | 0.840 | 0.123 | 0.911 | 0.061 |
| 59 | 0.942 | 0.125 | 0.976 | 0.133 | 0.912 | 0.104 | 0.822 | 0.134 | 0.930 | 0.138 | 0.976 | 0.03 |
| 60 | 0.930 | 0.117 | 0.976 | 0.131 | 0.891 | 0.092 | 0.900 | 0.138 | 0.899 | 0.126 | 0.847 | 0.121 |
| 61 | 0.874 | 0.131 | 0.981 | 0.114 | 0.867 | 0.115 | 0.935 | 0.138 | 0.916 | 0.136 | 0.978 | 0.123 |
| 62 | 0.964 | 0.153 | 0.940 | 0.128 | 0.841 | 0.122 | 0.872 | 0.126 | 0.907 | 0.129 | 0.829 | 0.076 |
| 63 | 0.955 | 0.123 | 0.825 | 0.118 | 0.896 | 0.114 | 0.968 | 0.126 | 0.837 | 0.137 | 0.919 | 0.066 |
| 64 | 0.966 | 0.129 | 0.930 | 0.119 | 0.876 | 0.124 | 0.859 | 0.137 | 0.893 | 0.134 | 0.938 | 0.092 |
| 65 | 0.933 | 0.126 | 0.940 | 0.113 | 0.965 | 0.118 | 0.938 | 0.137 | 0.963 | 0.127 | 0.650 | 0.096 |
| 66 | 1.003 | 0.131 | 0.960 | 0.112 | 0.946 | 0.106 | 0.924 | 0.114 | 0.888 | 0.114 | 0.838 | 0.032 |
| 67 | 0.829 | 0.134 | 0.951 | 0.119 | 0.981 | 0.1 | 0.840 | 0.125 | 0.918 | 0.127 | 0.973 | 0.098 |
| 68 | 0.928 | 0.124 | 0.897 | 0.125 | 0.848 | 0.127 | 1.213 | 0.121 | 0.978 | 0.123 | 0.621 | 0.123 |
| 69 | 0.958 | 0.125 | 0.905 | 0.126 | 0.935 | 0.099 | 0.968 | 0.118 | 0.981 | 0.127 | 0.849 | 0.042 |
| 70 | 0.909 | 0.127 | 0.819 | 0.114 | 0.907 | 0.114 | 1.056 | 0.118 | 0.923 | 0.115 | 0.888 | 0.096 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 53 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง **DS 60 และ PS 8** ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มเท่าเทียม**โดยใช้ข้อสอบร่วม (ต่อ)

| ข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 71 | 0.970 | 0.149 | 0.945 | 0.127 | 0.837 | 0.101 | 0.929 | 0.121 | 0.925 | 0.129 | 0.698 | 0.028 |
| 72 | 0.946 | 0.165 | 0.870 | 0.119 | 0.895 | 0.112 | 1.256 | 0.138 | 0.824 | 0.134 | 0.641 | 0.093 |
| 73 | 0.958 | 0.126 | 0.823 | 0.129 | 0.950 | 0.097 | 0.828 | 0.133 | 0.846 | 0.133 | 0.856 | 0.09 |
| 74 | 0.865 | 0.148 | 0.910 | 0.137 | 0.812 | 0.134 | 1.056 | 0.12 | 0.928 | 0.121 | 0.986 | 0.108 |
| 75 | 0.866 | 0.126 | 0.845 | 0.136 | 0.836 | 0.103 | 1.321 | 0.117 | 0.961 | 0.127 | 0.862 | 0.043 |
| 76 | 0.972 | 0.133 | 0.872 | 0.138 | 0.911 | 0.114 | 1.321 | 0.125 | 0.957 | 0.125 | 0.621 | 0.111 |
| 77 | 0.963 | 0.135 | 0.894 | 0.127 | 0.912 | 0.126 | 1.125 | 0.123 | 0.972 | 0.125 | 0.899 | 0.06 |
| 78 | 0.918 | 0.154 | 0.853 | 0.118 | 0.800 | 0.094 | 0.856 | 0.118 | 0.908 | 0.118 | 0.838 | 0.043 |
| 79 | 0.953 | 0.127 | 0.938 | 0.119 | 0.803 | 0.099 | 0.930 | 0.121 | 0.842 | 0.136 | 0.632 | 0.031 |
| 80 | 0.989 | 0.146 | 0.871 | 0.125 | 0.852 | 0.1 | 0.887 | 0.13 | 0.862 | 0.132 | 0.876 | 0.106 |
| 81 | 0.975 | 0.125 | 0.966 | 0.12 | 0.876 | 0.126 | 0.939 | 0.116 | 0.937 | 0.12 | 0.958 | 0.031 |
| 82 | 0.946 | 0.134 | 0.985 | 0.115 | 0.831 | 0.093 | 0.103 | 0.114 | 0.818 | 0.124 | 0.903 | 0.096 |
| 83 | 0.906 | 0.127 | 0.911 | 0.117 | 0.801 | 0.114 | 1.006 | 0.124 | 0.888 | 0.131 | 0.874 | 0.103 |
| 84 | 0.904 | 0.156 | 0.913 | 0.114 | 0.871 | 0.119 | 1.089 | 0.123 | 0.925 | 0.114 | 0.931 | 0.117 |
| 85 | 0.927 | 0.131 | 0.827 | 0.133 | 0.826 | 0.113 | 0.866 | 0.114 | 0.881 | 0.116 | 0.702 | 0.033 |
| 86 | 0.967 | 0.121 | 0.966 | 0.124 | 0.836 | 0.099 | 1.026 | 0.117 | 0.971 | 0.117 | 0.913 | 0.103 |
| 87 | 0.898 | 0.133 | 0.941 | 0.12 | 0.873 | 0.11 | 1.036 | 0.134 | 0.858 | 0.126 | 0.948 | 0.06 |
| 88 | 0.854 | 0.121 | 0.952 | 0.128 | 0.975 | 0.098 | 1.002 | 0.127 | 0.973 | 0.132 | 0.701 | 0.03 |
| 89 | 0.995 | 0.164 | 0.886 | 0.132 | 0.836 | 0.106 | 0.949 | 0.115 | 0.907 | 0.116 | 0.690 | 0.073 |
| 90 | 0.882 | 0.131 | 0.829 | 0.113 | 0.801 | 0.117 | 1.360 | 0.112 | 0.945 | 0.133 | 0.630 | 0.079 |
| 91 | 0.844 | 0.124 | 0.933 | 0.117 | 0.861 | 0.123 | 0.914 | 0.139 | 0.827 | 0.124 | 0.855 | 0.07 |
| 92 | 0.949 | 0.127 | 0.894 | 0.117 | 0.898 | 0.102 | 1.009 | 0.12 | 0.920 | 0.12 | 0.680 | 0.118 |
| 93 | 0.955 | 0.128 | 0.827 | 0.126 | 0.941 | 0.113 | 0.884 | 0.119 | 0.851 | 0.135 | 0.868 | 0.078 |
| 94 | 0.987 | 0.134 | 0.883 | 0.122 | 0.854 | 0.136 | 0.978 | 0.131 | 0.937 | 0.128 | 0.888 | 0.076 |
| 95 | 0.885 | 0.132 | 0.837 | 0.123 | 0.848 | 0.095 | 0.962 | 0.129 | 0.899 | 0.12 | 0.869 | 0.076 |
| 96 | 1.023 | 0.148 | 0.862 | 0.136 | 0.880 | 0.113 | 0.974 | 0.128 | 0.944 | 0.136 | 0.870 | 0.033 |
| 97 | 0.888 | 0.133 | 0.899 | 0.126 | 0.861 | 0.13 | 0.819 | 0.117 | 0.966 | 0.127 | 0.982 | 0.061 |
| 98 | 0.964 | 0.132 | 0.956 | 0.126 | 0.852 | 0.09 | 0.939 | 0.131 | 0.934 | 0.118 | 0.983 | 0.078 |
| 99 | 0.882 | 0.132 | 0.828 | 0.137 | 0.934 | 0.132 | 0.964 | 0.123 | 0.957 | 0.125 | 0.868 | 0.035 |
| 100 | 0.934 | 0.165 | 0.980 | 0.13 | 0.937 | 0.124 | 0.840 | 0.112 | 0.842 | 0.122 | 0.933 | 0.088 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 54 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 70 **และ** PS 6 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มเท่าเทียม**โดยใช้ข้อสอบร่วม

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 1 | 0.912 | 0.113 | 0.846 | 0.12 | 0.960 | 0.137 | 0.824 | 0.113 | 0.817 | 0.113 | 0.961 | 0.122 |
| 2 | 0.835 | 0.114 | 0.974 | 0.12 | 0.943 | 0.124 | 0.900 | 0.118 | 0.818 | 0.128 | 0.828 | 0.128 |
| 3 | 0.876 | 0.131 | 0.975 | 0.126 | 0.925 | 0.129 | 0.854 | 0.116 | 0.890 | 0.135 | 0.910 | 0.133 |
| 4 | 0.887 | 0.138 | 0.927 | 0.133 | 0.964 | 0.122 | 0.862 | 0.117 | 0.864 | 0.136 | 0.827 | 0.117 |
| 5 | 0.913 | 0.124 | 0.980 | 0.136 | 0.908 | 0.137 | 0.828 | 0.113 | 0.832 | 0.138 | 0.878 | 0.128 |
| 6 | 0.819 | 0.114 | 0.913 | 0.13 | 0.970 | 0.103 | 0.950 | 0.112 | 0.907 | 0.113 | 0.963 | 0.123 |
| 7 | 0.835 | 0.121 | 0.875 | 0.139 | 0.922 | 0.122 | 0.901 | 0.114 | 0.947 | 0.115 | 0.891 | 0.138 |
| 8 | 0.974 | 0.124 | 0.920 | 0.122 | 0.994 | 0.13 | 0.859 | 0.119 | 0.969 | 0.126 | 0.971 | 0.136 |
| 9 | 0.975 | 0.118 | 0.934 | 0.136 | 0.925 | 0.133 | 0.985 | 0.115 | 0.940 | 0.115 | 0.911 | 0.116 |
| 10 | 0.910 | 0.123 | 0.949 | 0.138 | 0.986 | 0.109 | 0.940 | 0.114 | 0.903 | 0.138 | 0.889 | 0.115 |
| 11 | 0.977 | 0.125 | 0.942 | 0.133 | 0.991 | 0.121 | 0.966 | 0.11 | 0.904 | 0.128 | 0.906 | 0.114 |
| 12 | 0.895 | 0.136 | 0.922 | 0.137 | 0.983 | 0.104 | 0.945 | 0.114 | 0.959 | 0.129 | 0.906 | 0.114 |
| 13 | 0.905 | 0.125 | 0.877 | 0.124 | 0.986 | 0.132 | 0.951 | 0.119 | 0.916 | 0.118 | 0.876 | 0.115 |
| 14 | 0.883 | 0.115 | 0.975 | 0.138 | 0.976 | 0.136 | 0.922 | 0.112 | 0.843 | 0.125 | 0.869 | 0.131 |
| 15 | 0.955 | 0.138 | 0.822 | 0.127 | 0.926 | 0.108 | 0.901 | 0.118 | 0.835 | 0.137 | 0.865 | 0.116 |
| 16 | 0.969 | 0.129 | 0.933 | 0.136 | 0.985 | 0.124 | 0.851 | 0.114 | 0.926 | 0.12 | 0.859 | 0.126 |
| 17 | 0.873 | 0.132 | 0.902 | 0.128 | 0.981 | 0.139 | 0.946 | 0.11 | 0.917 | 0.134 | 0.874 | 0.139 |
| 18 | 0.968 | 0.139 | 0.942 | 0.122 | 0.940 | 0.129 | 0.982 | 0.119 | 0.826 | 0.118 | 0.828 | 0.138 |
| 19 | 0.949 | 0.129 | 0.847 | 0.124 | 0.915 | 0.136 | 0.978 | 0.112 | 0.967 | 0.12 | 0.829 | 0.136 |
| 20 | 0.967 | 0.112 | 0.956 | 0.129 | 0.925 | 0.134 | 0.925 | 0.113 | 0.842 | 0.124 | 0.968 | 0.112 |
| 21 | 0.975 | 0.13 | 0.928 | 0.125 | 0.950 | 0.13 | 0.883 | 0.115 | 0.957 | 0.127 | 0.935 | 0.129 |
| 22 | 0.928 | 0.12 | 0.966 | 0.127 | 0.984 | 0.12 | 0.913 | 0.114 | 0.867 | 0.129 | 0.900 | 0.13 |
| 23 | 0.821 | 0.13 | 0.848 | 0.132 | 0.946 | 0.105 | 0.966 | 0.115 | 0.926 | 0.127 | 0.927 | 0.113 |
| 24 | 0.930 | 0.119 | 0.949 | 0.137 | 0.983 | 0.139 | 0.972 | 0.115 | 0.944 | 0.128 | 0.979 | 0.112 |
| 25 | 0.841 | 0.114 | 0.830 | 0.125 | 0.916 | 0.106 | 0.924 | 0.118 | 0.940 | 0.127 | 0.884 | 0.125 |
| 26 | 0.902 | 0.113 | 0.969 | 0.133 | 0.975 | 0.105 | 0.976 | 0.115 | 0.961 | 0.127 | 0.972 | 0.136 |
| 27 | 0.856 | 0.117 | 0.897 | 0.132 | 0.915 | 0.105 | 0.954 | 0.118 | 0.903 | 0.114 | 0.819 | 0.124 |
| 28 | 0.830 | 0.136 | 0.837 | 0.131 | 0.987 | 0.134 | 0.864 | 0.111 | 0.961 | 0.124 | 0.853 | 0.136 |
| 29 | 0.825 | 0.138 | 0.936 | 0.131 | 0.938 | 0.123 | 0.864 | 0.114 | 0.911 | 0.117 | 0.938 | 0.137 |
| 30 | 0.909 | 0.119 | 0.876 | 0.135 | 0.937 | 0.124 | 0.975 | 0.118 | 0.854 | 0.12 | 0.970 | 0.131 |
| 31 | 0.832 | 0.113 | 0.893 | 0.133 | 0.949 | 0.138 | 0.852 | 0.115 | 0.818 | 0.114 | 0.907 | 0.138 |
| 32 | 0.958 | 0.115 | 0.891 | 0.134 | 0.999 | 0.102 | 0.839 | 0.117 | 0.937 | 0.112 | 0.897 | 0.133 |
| 33 | 0.840 | 0.118 | 0.866 | 0.138 | 0.955 | 0.134 | 0.945 | 0.113 | 0.835 | 0.139 | 0.958 | 0.114 |
| 34 | 0.919 | 0.132 | 0.847 | 0.12 | 0.945 | 0.129 | 0.928 | 0.112 | 0.878 | 0.129 | 0.842 | 0.139 |
| 35 | 0.848 | 0.136 | 0.872 | 0.127 | 0.959 | 0.131 | 0.962 | 0.113 | 0.881 | 0.121 | 0.847 | 0.137 |
| 36 | 0.979 | 0.132 | 0.966 | 0.125 | 0.974 | 0.108 | 0.817 | 0.112 | 0.890 | 0.136 | 0.906 | 0.128 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 54 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 70 **และ** PS 6 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มเท่าเทียม**โดยใช้ข้อสอบร่วม (ต่อ)

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 37 | 0.951 | 0.116 | 0.983 | 0.124 | 0.968 | 0.137 | 0.986 | 0.114 | 0.929 | 0.134 | 0.968 | 0.133 |
| 38 | 0.963 | 0.138 | 0.882 | 0.121 | 0.922 | 0.104 | 0.918 | 0.113 | 0.951 | 0.117 | 0.921 | 0.131 |
| 39 | 0.827 | 0.128 | 0.855 | 0.125 | 0.920 | 0.103 | 0.874 | 0.112 | 0.843 | 0.136 | 0.753 | 0.128 |
| 40 | 0.928 | 0.139 | 0.973 | 0.133 | 0.963 | 0.12 | 0.965 | 0.113 | 0.854 | 0.13 | 0.799 | 0.12 |
| 41 | 0.982 | 0.125 | 0.967 | 0.137 | 0.960 | 0.136 | 0.932 | 0.115 | 0.842 | 0.121 | 0.735 | 0.13 |
| 42 | 0.940 | 0.139 | 0.963 | 0.138 | 0.976 | 0.135 | 0.815 | 0.112 | 0.979 | 0.117 | 0.963 | 0.117 |
| 43 | 0.907 | 0.132 | 0.899 | 0.13 | 0.986 | 0.102 | 0.985 | 0.119 | 0.817 | 0.118 | 0.974 | 0.114 |
| 44 | 0.865 | 0.138 | 0.879 | 0.137 | 0.964 | 0.132 | 0.919 | 0.114 | 0.973 | 0.134 | 0.761 | 0.124 |
| 45 | 0.846 | 0.126 | 0.836 | 0.137 | 0.970 | 0.132 | 0.936 | 0.116 | 0.930 | 0.134 | 0.772 | 0.132 |
| 46 | 0.987 | 0.117 | 0.916 | 0.121 | 0.996 | 0.104 | 0.907 | 0.116 | 0.975 | 0.12 | 0.949 | 0.132 |
| 47 | 0.864 | 0.125 | 0.854 | 0.139 | 0.931 | 0.109 | 0.965 | 0.11 | 0.951 | 0.125 | 0.943 | 0.122 |
| 48 | 0.924 | 0.114 | 0.930 | 0.13 | 0.900 | 0.125 | 0.954 | 0.11 | 0.938 | 0.128 | 0.778 | 0.136 |
| 49 | 0.930 | 0.13 | 0.985 | 0.127 | 0.990 | 0.107 | 0.959 | 0.114 | 0.836 | 0.138 | 0.922 | 0.138 |
| 50 | 0.856 | 0.129 | 0.816 | 0.128 | 0.977 | 0.137 | 0.842 | 0.118 | 0.862 | 0.13 | 0.785 | 0.131 |
| 51 | 0.931 | 0.127 | 0.849 | 0.124 | 0.910 | 0.12 | 0.840 | 0.116 | 0.970 | 0.117 | 0.916 | 0.129 |
| 52 | 0.881 | 0.132 | 0.874 | 0.126 | 0.903 | 0.138 | 0.950 | 0.11 | 0.856 | 0.122 | 0.963 | 0.133 |
| 53 | 0.968 | 0.116 | 0.885 | 0.133 | 0.939 | 0.127 | 0.940 | 0.116 | 0.855 | 0.117 | 0.767 | 0.122 |
| 54 | 0.887 | 0.115 | 0.925 | 0.139 | 0.943 | 0.134 | 0.863 | 0.117 | 0.843 | 0.124 | 0.737 | 0.124 |
| 55 | 0.961 | 0.133 | 0.964 | 0.126 | 0.974 | 0.137 | 0.929 | 0.113 | 0.961 | 0.136 | 0.786 | 0.137 |
| 56 | 0.978 | 0.126 | 0.960 | 0.122 | 0.925 | 0.12 | 0.957 | 0.114 | 0.949 | 0.118 | 0.934 | 0.138 |
| 57 | 0.847 | 0.134 | 0.925 | 0.126 | 0.971 | 0.137 | 0.825 | 0.116 | 0.934 | 0.137 | 0.735 | 0.115 |
| 58 | 0.957 | 0.117 | 0.819 | 0.133 | 0.900 | 0.132 | 0.839 | 0.119 | 0.953 | 0.124 | 0.716 | 0.131 |
| 59 | 0.821 | 0.123 | 0.940 | 0.139 | 0.960 | 0.121 | 0.946 | 0.117 | 0.830 | 0.123 | 0.725 | 0.134 |
| 60 | 0.892 | 0.121 | 0.900 | 0.131 | 0.913 | 0.136 | 0.940 | 0.118 | 0.953 | 0.127 | 0.745 | 0.12 |
| 61 | 0.944 | 0.127 | 0.857 | 0.139 | 0.995 | 0.131 | 0.820 | 0.112 | 0.928 | 0.129 | 0.935 | 0.112 |
| 62 | 0.844 | 0.13 | 0.953 | 0.135 | 0.959 | 0.135 | 0.894 | 0.114 | 0.827 | 0.134 | 0.796 | 0.124 |
| 63 | 0.838 | 0.115 | 0.966 | 0.126 | 0.927 | 0.102 | 0.877 | 0.117 | 0.829 | 0.114 | 0.768 | 0.131 |
| 64 | 0.985 | 0.138 | 0.953 | 0.133 | 0.916 | 0.107 | 0.907 | 0.113 | 0.940 | 0.115 | 0.944 | 0.13 |
| 65 | 0.959 | 0.128 | 0.859 | 0.135 | 0.938 | 0.107 | 0.825 | 0.116 | 0.887 | 0.122 | 0.759 | 0.117 |
| 66 | 0.816 | 0.117 | 0.842 | 0.137 | 0.980 | 0.125 | 0.868 | 0.119 | 0.879 | 0.126 | 0.747 | 0.126 |
| 67 | 0.911 | 0.114 | 0.964 | 0.138 | 0.943 | 0.103 | 0.865 | 0.118 | 0.923 | 0.122 | 0.755 | 0.137 |
| 68 | 0.845 | 0.134 | 0.981 | 0.12 | 0.937 | 0.129 | 0.914 | 0.111 | 0.979 | 0.123 | 0.719 | 0.131 |
| 69 | 0.852 | 0.117 | 0.932 | 0.132 | 0.924 | 0.13 | 0.937 | 0.119 | 0.881 | 0.121 | 0.910 | 0.113 |
| 70 | 0.878 | 0.116 | 0.887 | 0.136 | 0.986 | 0.107 | 0.861 | 0.116 | 0.957 | 0.123 | 0.748 | 0.112 |
| 71 | 0.863 | 0.116 | 0.942 | 0.12 | 0.985 | 0.134 | 0.963 | 0.114 | 0.896 | 0.129 | 0.946 | 0.125 |
| 72 | 0.963 | 0.138 | 0.921 | 0.135 | 0.972 | 0.136 | 0.939 | 0.115 | 0.830 | 0.114 | 0.718 | 0.122 |

ตารางที่ 54 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 70 **และ** PS 6 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มเท่าเทียม**โดยใช้ข้อสอบร่วม (ต่อ)

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 73 | 0.961 | 0.132 | 0.841 | 0.135 | 0.966 | 0.125 | 0.977 | 0.118 | 0.860 | 0.115 | 0.729 | 0.126 |
| 74 | 0.924 | 0.133 | 0.983 | 0.13 | 0.997 | 0.107 | 0.957 | 0.115 | 0.921 | 0.113 | 0.769 | 0.124 |
| 75 | 0.848 | 0.13 | 0.885 | 0.124 | 0.980 | 0.131 | 0.936 | 0.111 | 0.830 | 0.12 | 0.903 | 0.132 |
| 76 | 0.847 | 0.122 | 0.964 | 0.137 | 0.921 | 0.127 | 0.971 | 0.117 | 0.901 | 0.132 | 0.933 | 0.117 |
| 77 | 0.967 | 0.134 | 0.973 | 0.127 | 0.925 | 0.127 | 0.903 | 0.118 | 0.902 | 0.123 | 0.909 | 0.126 |
| 78 | 0.937 | 0.129 | 0.884 | 0.133 | 0.918 | 0.102 | 0.986 | 0.115 | 0.931 | 0.126 | 0.947 | 0.126 |
| 79 | 0.932 | 0.113 | 0.848 | 0.139 | 0.904 | 0.121 | 0.958 | 0.113 | 0.950 | 0.115 | 0.965 | 0.13 |
| 80 | 0.961 | 0.126 | 0.909 | 0.136 | 0.912 | 0.131 | 0.967 | 0.113 | 0.826 | 0.138 | 0.932 | 0.116 |
| 81 | 0.956 | 0.131 | 0.820 | 0.124 | 0.994 | 0.105 | 0.937 | 0.113 | 0.824 | 0.113 | 0.907 | 0.126 |
| 82 | 0.881 | 0.117 | 0.965 | 0.121 | 0.913 | 0.12 | 0.851 | 0.114 | 0.904 | 0.129 | 0.729 | 0.139 |
| 83 | 0.887 | 0.113 | 0.930 | 0.136 | 0.924 | 0.132 | 0.985 | 0.117 | 0.821 | 0.131 | 0.771 | 0.113 |
| 84 | 0.915 | 0.113 | 0.943 | 0.128 | 0.924 | 0.105 | 0.860 | 0.115 | 0.849 | 0.128 | 0.781 | 0.133 |
| 85 | 0.986 | 0.115 | 0.951 | 0.121 | 0.902 | 0.129 | 0.868 | 0.117 | 0.857 | 0.118 | 0.901 | 0.129 |
| 86 | 0.984 | 0.129 | 0.945 | 0.138 | 0.980 | 0.103 | 0.898 | 0.118 | 0.891 | 0.119 | 0.754 | 0.135 |
| 87 | 0.901 | 0.138 | 0.943 | 0.124 | 0.913 | 0.137 | 0.850 | 0.116 | 0.823 | 0.126 | 0.719 | 0.114 |
| 88 | 0.980 | 0.131 | 0.911 | 0.139 | 0.964 | 0.109 | 0.942 | 0.114 | 0.831 | 0.116 | 0.740 | 0.119 |
| 89 | 0.860 | 0.113 | 0.896 | 0.132 | 0.946 | 0.137 | 0.944 | 0.117 | 0.821 | 0.134 | 0.977 | 0.134 |
| 90 | 0.860 | 0.117 | 0.912 | 0.128 | 0.963 | 0.138 | 0.950 | 0.11 | 0.965 | 0.113 | 0.716 | 0.128 |
| 91 | 0.939 | 0.12 | 0.983 | 0.125 | 0.960 | 0.138 | 0.926 | 0.11 | 0.914 | 0.134 | 0.921 | 0.117 |
| 92 | 0.882 | 0.135 | 0.835 | 0.132 | 0.918 | 0.106 | 0.925 | 0.118 | 0.861 | 0.13 | 0.954 | 0.114 |
| 93 | 0.977 | 0.116 | 0.861 | 0.137 | 0.915 | 0.133 | 0.923 | 0.115 | 0.895 | 0.117 | 0.965 | 0.118 |
| 94 | 0.908 | 0.122 | 0.824 | 0.127 | 0.907 | 0.105 | 0.931 | 0.113 | 0.820 | 0.116 | 0.935 | 0.113 |
| 95 | 0.845 | 0.128 | 0.844 | 0.12 | 0.972 | 0.13 | 0.931 | 0.112 | 0.823 | 0.138 | 0.773 | 0.116 |
| 96 | 0.883 | 0.134 | 0.958 | 0.13 | 0.903 | 0.134 | 0.932 | 0.114 | 0.866 | 0.134 | 0.799 | 0.131 |
| 97 | 0.824 | 0.112 | 0.896 | 0.12 | 0.953 | 0.126 | 0.875 | 0.145 | 0.872 | 0.115 | 0.904 | 0.113 |
| 98 | 0.832 | 0.13 | 0.921 | 0.139 | 0.908 | 0.133 | 0.852 | 0.116 | 0.954 | 0.128 | 0.977 | 0.139 |
| 99 | 0.832 | 0.122 | 0.847 | 0.123 | 0.988 | 0.12 | 0.848 | 0.113 | 0.957 | 0.116 | 0.772 | 0.124 |
| 100 | 0.834 | 0.131 | 0.976 | 0.121 | 0.930 | 0.126 | 0.984 | 0.116 | 0.963 | 0.137 | 0.915 | 0.128 |

ตารางที่ 55 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 80 **และ** PS 4 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มเท่าเทียม**โดยใช้ข้อสอบร่วม

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 1 | 0.923 | 0.112 | 0.830 | 0.139 | 0.941 | 0.135 | 0.982 | 0.128 | 0.976 | 0.114 | 0.838 | 0.134 |
| 2 | 0.944 | 0.125 | 0.854 | 0.129 | 0.955 | 0.135 | 0.977 | 0.137 | 0.822 | 0.135 | 0.829 | 0.137 |
| 3 | 0.983 | 0.114 | 0.985 | 0.131 | 0.971 | 0.132 | 0.818 | 0.133 | 0.845 | 0.134 | 0.851 | 0.131 |
| 4 | 0.983 | 0.131 | 0.917 | 0.128 | 0.926 | 0.132 | 0.856 | 0.139 | 0.966 | 0.128 | 0.871 | 0.119 |
| 5 | 0.921 | 0.121 | 0.952 | 0.136 | 0.982 | 0.11 | 0.970 | 0.128 | 0.842 | 0.115 | 0.845 | 0.117 |
| 6 | 0.940 | 0.132 | 0.830 | 0.131 | 0.924 | 0.139 | 0.969 | 0.12 | 0.970 | 0.134 | 0.984 | 0.115 |
| 7 | 0.897 | 0.113 | 0.938 | 0.13 | 0.913 | 0.116 | 0.841 | 0.139 | 0.985 | 0.137 | 0.930 | 0.113 |
| 8 | 0.886 | 0.138 | 0.948 | 0.12 | 0.947 | 0.111 | 0.837 | 0.133 | 0.904 | 0.121 | 0.933 | 0.132 |
| 9 | 0.947 | 0.121 | 0.891 | 0.123 | 0.908 | 0.117 | 0.927 | 0.127 | 0.940 | 0.13 | 0.836 | 0.138 |
| 10 | 0.841 | 0.136 | 0.955 | 0.139 | 0.955 | 0.113 | 0.928 | 0.132 | 0.961 | 0.133 | 0.865 | 0.134 |
| 11 | 0.957 | 0.123 | 0.943 | 0.139 | 0.941 | 0.115 | 0.831 | 0.126 | 0.942 | 0.124 | 0.907 | 0.119 |
| 12 | 0.946 | 0.126 | 0.858 | 0.135 | 0.927 | 0.133 | 0.962 | 0.134 | 0.941 | 0.138 | 0.984 | 0.133 |
| 13 | 0.943 | 0.137 | 0.909 | 0.133 | 0.949 | 0.135 | 0.819 | 0.135 | 0.893 | 0.122 | 0.949 | 0.115 |
| 14 | 0.875 | 0.123 | 0.847 | 0.128 | 0.935 | 0.135 | 0.834 | 0.131 | 0.980 | 0.118 | 0.973 | 0.115 |
| 15 | 0.828 | 0.113 | 0.960 | 0.133 | 0.930 | 0.116 | 0.858 | 0.129 | 0.942 | 0.132 | 0.836 | 0.115 |
| 16 | 0.935 | 0.128 | 0.914 | 0.134 | 0.903 | 0.134 | 0.823 | 0.138 | 0.891 | 0.117 | 0.833 | 0.136 |
| 17 | 0.838 | 0.115 | 0.984 | 0.121 | 0.942 | 0.117 | 0.937 | 0.138 | 0.895 | 0.128 | 0.823 | 0.139 |
| 18 | 0.965 | 0.134 | 0.965 | 0.138 | 0.924 | 0.111 | 0.902 | 0.135 | 0.971 | 0.12 | 0.896 | 0.138 |
| 19 | 0.929 | 0.116 | 0.826 | 0.13 | 0.940 | 0.119 | 0.982 | 0.139 | 0.859 | 0.132 | 0.927 | 0.134 |
| 20 | 0.941 | 0.121 | 0.879 | 0.126 | 0.956 | 0.117 | 0.895 | 0.128 | 0.922 | 0.115 | 0.845 | 0.13 |
| 21 | 0.972 | 0.115 | 0.894 | 0.138 | 0.950 | 0.115 | 0.985 | 0.127 | 0.983 | 0.127 | 0.836 | 0.135 |
| 22 | 0.835 | 0.13 | 0.826 | 0.123 | 0.919 | 0.113 | 0.868 | 0.127 | 0.899 | 0.128 | 0.927 | 0.131 |
| 23 | 0.893 | 0.124 | 0.963 | 0.135 | 0.978 | 0.112 | 0.966 | 0.138 | 0.930 | 0.123 | 0.834 | 0.137 |
| 24 | 0.823 | 0.122 | 0.831 | 0.139 | 0.978 | 0.118 | 0.819 | 0.134 | 0.951 | 0.125 | 0.823 | 0.134 |
| 25 | 0.839 | 0.121 | 0.850 | 0.125 | 0.908 | 0.134 | 0.932 | 0.122 | 0.983 | 0.124 | 0.984 | 0.133 |
| 26 | 0.850 | 0.114 | 0.883 | 0.132 | 0.977 | 0.119 | 0.880 | 0.129 | 0.857 | 0.136 | 0.965 | 0.133 |
| 27 | 0.837 | 0.122 | 0.887 | 0.124 | 0.980 | 0.133 | 0.898 | 0.126 | 0.990 | 0.131 | 0.868 | 0.118 |
| 28 | 0.892 | 0.117 | 0.817 | 0.124 | 0.913 | 0.115 | 0.825 | 0.121 | 0.952 | 0.121 | 0.834 | 0.13 |
| 29 | 0.972 | 0.116 | 0.974 | 0.134 | 0.947 | 0.115 | 0.875 | 0.138 | 0.896 | 0.138 | 0.823 | 0.131 |
| 30 | 0.843 | 0.116 | 0.826 | 0.136 | 0.929 | 0.115 | 0.873 | 0.129 | 0.892 | 0.126 | 0.843 | 0.115 |
| 31 | 0.935 | 0.135 | 0.850 | 0.134 | 0.996 | 0.116 | 0.978 | 0.121 | 0.935 | 0.128 | 0.853 | 0.118 |
| 32 | 0.934 | 0.113 | 0.905 | 0.139 | 0.951 | 0.119 | 0.958 | 0.133 | 0.959 | 0.127 | 0.861 | 0.13 |
| 33 | 0.914 | 0.118 | 0.987 | 0.123 | 0.960 | 0.138 | 0.944 | 0.12 | 0.892 | 0.135 | 0.843 | 0.134 |
| 34 | 0.860 | 0.118 | 0.853 | 0.133 | 0.999 | 0.134 | 0.855 | 0.137 | 0.895 | 0.123 | 0.869 | 0.138 |
| 35 | 0.881 | 0.134 | 0.956 | 0.124 | 0.933 | 0.13 | 0.855 | 0.124 | 0.828 | 0.123 | 0.909 | 0.134 |
| 36 | 0.977 | 0.13 | 0.927 | 0.126 | 0.952 | 0.115 | 0.869 | 0.138 | 0.937 | 0.112 | 0.853 | 0.132 |

ตารางที่ 55 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง **DS 80 และ PS 4** ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มเท่าเทียม**โดยใช้ข้อสอบร่วม (ต่อ)

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 37 | 0.844 | 0.137 | 0.979 | 0.137 | 0.944 | 0.131 | 0.908 | 0.136 | 0.828 | 0.114 | 0.950 | 0.134 |
| 38 | 0.935 | 0.128 | 0.922 | 0.131 | 0.946 | 0.137 | 0.884 | 0.134 | 0.910 | 0.116 | 0.859 | 0.137 |
| 39 | 0.928 | 0.128 | 0.827 | 0.127 | 0.969 | 0.134 | 0.880 | 0.126 | 0.969 | 0.121 | 0.909 | 0.113 |
| 40 | 0.914 | 0.113 | 0.950 | 0.134 | 0.968 | 0.133 | 0.848 | 0.125 | 0.873 | 0.136 | 0.983 | 0.134 |
| 41 | 0.837 | 0.112 | 0.839 | 0.139 | 0.924 | 0.133 | 0.854 | 0.134 | 0.841 | 0.133 | 0.892 | 0.137 |
| 42 | 0.897 | 0.119 | 0.903 | 0.12 | 0.916 | 0.118 | 0.922 | 0.129 | 0.858 | 0.125 | 0.936 | 0.138 |
| 43 | 0.918 | 0.114 | 0.865 | 0.12 | 0.964 | 0.13 | 0.832 | 0.136 | 0.902 | 0.122 | 0.839 | 0.134 |
| 44 | 0.847 | 0.118 | 0.942 | 0.125 | 0.959 | 0.111 | 0.837 | 0.125 | 0.905 | 0.138 | 0.923 | 0.138 |
| 45 | 0.858 | 0.127 | 0.938 | 0.138 | 0.949 | 0.115 | 0.881 | 0.129 | 0.956 | 0.137 | 0.893 | 0.135 |
| 46 | 0.906 | 0.119 | 0.945 | 0.139 | 0.951 | 0.118 | 0.904 | 0.129 | 0.958 | 0.133 | 0.824 | 0.136 |
| 47 | 0.893 | 0.121 | 0.974 | 0.126 | 0.929 | 0.13 | 0.919 | 0.126 | 0.900 | 0.122 | 0.850 | 0.118 |
| 48 | 0.959 | 0.138 | 0.984 | 0.129 | 0.950 | 0.114 | 0.864 | 0.126 | 0.983 | 0.131 | 0.874 | 0.114 |
| 49 | 0.966 | 0.119 | 0.954 | 0.122 | 0.963 | 0.118 | 0.828 | 0.139 | 0.901 | 0.139 | 0.960 | 0.113 |
| 50 | 0.837 | 0.137 | 0.858 | 0.133 | 0.927 | 0.134 | 0.843 | 0.138 | 0.965 | 0.114 | 0.905 | 0.115 |
| 51 | 0.896 | 0.123 | 0.876 | 0.135 | 0.977 | 0.112 | 0.936 | 0.138 | 0.841 | 0.118 | 0.843 | 0.114 |
| 52 | 0.920 | 0.134 | 0.845 | 0.139 | 0.969 | 0.114 | 0.831 | 0.131 | 0.916 | 0.138 | 0.929 | 0.139 |
| 53 | 0.878 | 0.132 | 0.817 | 0.138 | 0.943 | 0.117 | 0.918 | 0.133 | 0.960 | 0.133 | 0.914 | 0.117 |
| 54 | 0.904 | 0.127 | 0.910 | 0.135 | 0.923 | 0.113 | 0.820 | 0.129 | 0.971 | 0.117 | 0.905 | 0.135 |
| 55 | 0.983 | 0.124 | 0.985 | 0.121 | 0.980 | 0.114 | 0.831 | 0.138 | 0.915 | 0.131 | 0.955 | 0.138 |
| 56 | 0.871 | 0.117 | 0.975 | 0.136 | 0.941 | 0.137 | 0.858 | 0.135 | 0.972 | 0.125 | 0.984 | 0.135 |
| 57 | 0.950 | 0.123 | 0.824 | 0.131 | 0.992 | 0.138 | 0.883 | 0.138 | 0.909 | 0.133 | 0.815 | 0.134 |
| 58 | 0.954 | 0.121 | 0.946 | 0.136 | 0.928 | 0.114 | 0.825 | 0.123 | 0.928 | 0.112 | 0.871 | 0.138 |
| 59 | 0.936 | 0.139 | 0.912 | 0.132 | 0.968 | 0.118 | 0.981 | 0.121 | 0.898 | 0.113 | 0.888 | 0.137 |
| 60 | 0.917 | 0.119 | 0.955 | 0.122 | 0.925 | 0.135 | 0.816 | 0.124 | 0.983 | 0.118 | 0.876 | 0.131 |
| 61 | 0.965 | 0.121 | 0.950 | 0.131 | 0.928 | 0.116 | 0.865 | 0.138 | 0.960 | 0.126 | 0.910 | 0.134 |
| 62 | 0.835 | 0.123 | 0.932 | 0.133 | 0.997 | 0.118 | 0.835 | 0.137 | 0.967 | 0.119 | 0.976 | 0.133 |
| 63 | 0.878 | 0.127 | 0.946 | 0.127 | 0.973 | 0.114 | 0.954 | 0.131 | 0.926 | 0.117 | 0.820 | 0.114 |
| 64 | 0.828 | 0.119 | 0.925 | 0.133 | 0.959 | 0.113 | 0.930 | 0.134 | 0.858 | 0.119 | 0.895 | 0.13 |
| 65 | 0.852 | 0.122 | 0.873 | 0.121 | 0.936 | 0.115 | 0.975 | 0.133 | 0.899 | 0.119 | 0.937 | 0.132 |
| 66 | 0.978 | 0.122 | 0.847 | 0.135 | 0.930 | 0.114 | 0.915 | 0.134 | 0.908 | 0.123 | 0.858 | 0.133 |
| 67 | 0.903 | 0.126 | 0.985 | 0.135 | 0.921 | 0.139 | 0.907 | 0.13 | 0.859 | 0.138 | 0.819 | 0.132 |
| 68 | 0.984 | 0.138 | 0.979 | 0.13 | 0.918 | 0.117 | 0.879 | 0.132 | 0.978 | 0.135 | 0.976 | 0.134 |
| 69 | 0.944 | 0.133 | 0.921 | 0.139 | 0.916 | 0.115 | 0.895 | 0.133 | 0.823 | 0.114 | 0.846 | 0.119 |
| 70 | 0.961 | 0.135 | 0.967 | 0.139 | 0.909 | 0.138 | 0.863 | 0.132 | 0.842 | 0.127 | 0.887 | 0.138 |
| 71 | 0.910 | 0.115 | 0.901 | 0.133 | 0.931 | 0.115 | 0.943 | 0.134 | 0.947 | 0.134 | 0.986 | 0.135 |
| 72 | 0.962 | 0.124 | 0.958 | 0.133 | 0.917 | 0.116 | 0.830 | 0.139 | 0.906 | 0.124 | 0.956 | 0.132 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 55 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 80 **และ** PS 4 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มเท่าเทียม**โดยใช้ข้อสอบร่วม (ต่อ)

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 73 | 0.886 | 0.135 | 0.986 | 0.133 | 0.973 | 0.11 | 0.910 | 0.128 | 0.842 | 0.137 | 0.882 | 0.138 |
| 74 | 0.866 | 0.122 | 0.987 | 0.135 | 0.969 | 0.11 | 0.917 | 0.125 | 0.902 | 0.117 | 0.904 | 0.115 |
| 75 | 0.839 | 0.133 | 0.936 | 0.136 | 0.982 | 0.113 | 0.875 | 0.122 | 0.982 | 0.127 | 0.953 | 0.136 |
| 76 | 0.912 | 0.116 | 0.963 | 0.129 | 0.931 | 0.111 | 0.968 | 0.128 | 0.896 | 0.126 | 0.876 | 0.13 |
| 77 | 0.838 | 0.133 | 0.857 | 0.13 | 0.956 | 0.119 | 0.903 | 0.135 | 0.980 | 0.117 | 0.840 | 0.132 |
| 78 | 0.958 | 0.125 | 0.900 | 0.134 | 0.991 | 0.11 | 0.963 | 0.136 | 0.893 | 0.127 | 0.885 | 0.133 |
| 79 | 0.937 | 0.133 | 0.942 | 0.136 | 0.969 | 0.115 | 0.882 | 0.13 | 0.950 | 0.124 | 0.969 | 0.135 |
| 80 | 0.918 | 0.117 | 0.889 | 0.127 | 0.934 | 0.115 | 0.899 | 0.122 | 0.910 | 0.113 | 0.817 | 0.115 |
| 81 | 0.818 | 0.112 | 0.946 | 0.13 | 0.962 | 0.117 | 0.952 | 0.133 | 0.905 | 0.118 | 0.920 | 0.132 |
| 82 | 0.891 | 0.121 | 0.900 | 0.135 | 0.969 | 0.13 | 0.819 | 0.135 | 0.970 | 0.131 | 0.972 | 0.115 |
| 83 | 0.873 | 0.139 | 0.868 | 0.133 | 0.922 | 0.115 | 0.834 | 0.135 | 0.950 | 0.126 | 0.969 | 0.118 |
| 84 | 0.914 | 0.129 | 0.857 | 0.138 | 0.939 | 0.138 | 0.982 | 0.122 | 0.939 | 0.133 | 0.942 | 0.116 |
| 85 | 0.952 | 0.128 | 0.818 | 0.139 | 0.954 | 0.117 | 0.935 | 0.135 | 0.819 | 0.127 | 0.904 | 0.112 |
| 86 | 0.856 | 0.113 | 0.971 | 0.132 | 0.935 | 0.116 | 0.903 | 0.138 | 0.990 | 0.124 | 0.845 | 0.119 |
| 87 | 0.950 | 0.139 | 0.946 | 0.126 | 0.930 | 0.112 | 0.862 | 0.136 | 0.824 | 0.116 | 0.887 | 0.133 |
| 88 | 0.983 | 0.128 | 0.826 | 0.139 | 0.920 | 0.114 | 0.957 | 0.132 | 0.990 | 0.12 | 0.898 | 0.139 |
| 89 | 0.911 | 0.136 | 0.860 | 0.132 | 0.915 | 0.136 | 0.869 | 0.139 | 0.851 | 0.136 | 0.917 | 0.131 |
| 90 | 0.866 | 0.131 | 0.869 | 0.132 | 0.981 | 0.114 | 0.881 | 0.133 | 0.827 | 0.112 | 0.869 | 0.130 |
| 91 | 0.964 | 0.123 | 0.886 | 0.137 | 0.965 | 0.114 | 0.948 | 0.129 | 0.930 | 0.138 | 0.845 | 0.130 |
| 92 | 0.852 | 0.114 | 0.946 | 0.134 | 0.980 | 0.133 | 0.853 | 0.126 | 0.914 | 0.129 | 0.977 | 0.133 |
| 93 | 0.928 | 0.128 | 0.893 | 0.121 | 0.964 | 0.116 | 0.937 | 0.136 | 0.930 | 0.13 | 0.862 | 0.119 |
| 94 | 0.958 | 0.131 | 0.834 | 0.131 | 0.972 | 0.136 | 0.939 | 0.137 | 0.923 | 0.124 | 0.869 | 0.119 |
| 95 | 0.986 | 0.125 | 0.837 | 0.125 | 0.975 | 0.119 | 0.889 | 0.122 | 0.972 | 0.134 | 0.861 | 0.135 |
| 96 | 0.961 | 0.115 | 0.911 | 0.135 | 0.906 | 0.138 | 0.900 | 0.137 | 0.974 | 0.115 | 0.920 | 0.113 |
| 97 | 0.979 | 0.135 | 0.902 | 0.138 | 0.982 | 0.116 | 0.886 | 0.124 | 0.981 | 0.13 | 0.931 | 0.138 |
| 98 | 0.916 | 0.132 | 0.971 | 0.132 | 0.924 | 0.136 | 0.861 | 0.13 | 0.910 | 0.128 | 0.844 | 0.133 |
| 99 | 0.819 | 0.136 | 0.840 | 0.135 | 0.961 | 0.137 | 0.828 | 0.135 | 0.965 | 0.115 | 0.952 | 0.114 |
| 100 | 0.838 | 0.128 | 0.908 | 0.128 | 0.986 | 0.111 | 0.979 | 0.135 | 0.940 | 0.115 | 0.816 | 0.131 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 56 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 50 **และ** PS 10 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มไม่เท่าเทียม**โดยใช้ข้อสอบร่วม

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 1 | 1.002 | 0.723 | 1.122 | 0.828 | 1.211 | 0.713 | 1.257 | 0.820 | 1.121 | 0.736 | 1.139 | 0.882 |
| 2 | 1.093 | 0.653 | 1.138 | 0.725 | 1.119 | 0.741 | 1.290 | 0.731 | 1.370 | 0.770 | 1.171 | 0.859 |
| 3 | 1.100 | 0.740 | 1.129 | 0.797 | 1.244 | 0.553 | 1.288 | 0.756 | 1.272 | 0.728 | 1.154 | 0.918 |
| 4 | 1.045 | 0.660 | 1.123 | 0.607 | 1.273 | 0.552 | 1.207 | 0.742 | 1.111 | 0.743 | 1.145 | 0.836 |
| 5 | 1.021 | 0.743 | 1.105 | 0.835 | 1.236 | 0.638 | 1.298 | 0.846 | 1.296 | 0.837 | 1.188 | 0.903 |
| 6 | 1.026 | 0.659 | 1.133 | 0.836 | 1.116 | 0.724 | 1.298 | 0.672 | 1.274 | 0.857 | 1.126 | 0.937 |
| 7 | 1.107 | 0.653 | 1.103 | 0.827 | 1.118 | 0.754 | 1.276 | 0.748 | 1.115 | 0.794 | 1.137 | 0.884 |
| 8 | 1.007 | 0.654 | 1.109 | 0.794 | 1.263 | 0.715 | 1.212 | 0.749 | 1.120 | 0.763 | 1.183 | 0.818 |
| 9 | 1.092 | 0.654 | 1.174 | 0.796 | 1.286 | 0.552 | 1.259 | 0.672 | 1.350 | 0.737 | 1.139 | 0.915 |
| 10 | 1.093 | 0.657 | 1.122 | 0.607 | 1.263 | 0.748 | 1.298 | 0.660 | 1.310 | 0.725 | 1.112 | 0.915 |
| 11 | 1.096 | 0.656 | 1.125 | 0.652 | 1.221 | 0.799 | 1.289 | 0.699 | 1.310 | 0.769 | 1.198 | 0.944 |
| 12 | 1.002 | 0.658 | 1.198 | 0.758 | 1.234 | 0.639 | 1.296 | 0.737 | 1.370 | 0.828 | 1.105 | 0.833 |
| 13 | 0.999 | 0.655 | 1.122 | 0.729 | 1.248 | 0.744 | 1.293 | 0.855 | 1.296 | 0.764 | 1.121 | 0.904 |
| 14 | 1.106 | 0.658 | 1.167 | 0.752 | 1.287 | 0.733 | 1.242 | 0.840 | 1.108 | 0.752 | 1.102 | 0.881 |
| 15 | 1.009 | 0.655 | 1.174 | 0.600 | 1.204 | 0.794 | 1.222 | 0.672 | 1.299 | 0.838 | 1.171 | 0.848 |
| 16 | 1.061 | 0.724 | 1.180 | 0.857 | 1.278 | 0.798 | 1.227 | 0.802 | 1.296 | 0.785 | 1.127 | 0.872 |
| 17 | 1.090 | 0.653 | 1.169 | 0.650 | 1.290 | 0.639 | 1.218 | 0.747 | 1.233 | 0.721 | 1.107 | 0.834 |
| 18 | 1.060 | 0.657 | 1.176 | 0.853 | 1.297 | 0.715 | 1.217 | 0.692 | 1.330 | 0.844 | 1.119 | 0.839 |
| 19 | 0.998 | 0.735 | 1.134 | 0.815 | 1.295 | 0.759 | 1.205 | 0.751 | 1.380 | 0.713 | 1.128 | 0.885 |
| 20 | 1.071 | 0.655 | 1.130 | 0.822 | 1.292 | 0.731 | 1.213 | 0.806 | 1.248 | 0.748 | 1.143 | 0.922 |
| 21 | 1.024 | 0.655 | 1.116 | 0.600 | 1.291 | 0.752 | 1.231 | 0.698 | 1.115 | 0.794 | 1.187 | 0.826 |
| 22 | 1.017 | 0.658 | 1.107 | 0.816 | 1.119 | 0.721 | 1.295 | 0.681 | 1.330 | 0.780 | 1.155 | 0.915 |
| 23 | 1.008 | 0.654 | 1.180 | 0.716 | 1.104 | 0.742 | 1.273 | 0.839 | 1.122 | 0.770 | 1.123 | 0.830 |
| 24 | 1.059 | 0.656 | 1.121 | 0.603 | 1.122 | 0.796 | 1.255 | 0.855 | 1.232 | 0.726 | 1.161 | 0.897 |
| 25 | 1.045 | 0.727 | 1.158 | 0.751 | 1.108 | 0.740 | 1.275 | 0.672 | 1.380 | 0.741 | 1.127 | 0.927 |
| 26 | 1.036 | 0.703 | 1.113 | 0.831 | 1.227 | 0.638 | 1.211 | 0.826 | 1.231 | 0.747 | 1.127 | 0.867 |
| 27 | 1.119 | 0.749 | 1.196 | 0.608 | 1.241 | 0.639 | 1.209 | 0.836 | 1.103 | 0.750 | 1.107 | 0.835 |
| 28 | 1.003 | 0.752 | 1.107 | 0.728 | 1.118 | 0.797 | 1.203 | 0.689 | 1.112 | 0.741 | 1.156 | 0.895 |
| 29 | 1.036 | 0.710 | 1.132 | 0.795 | 1.113 | 0.737 | 1.218 | 0.691 | 1.121 | 0.725 | 1.104 | 0.922 |
| 30 | 0.999 | 0.654 | 1.150 | 0.843 | 1.262 | 0.798 | 1.279 | 0.755 | 1.106 | 0.740 | 1.119 | 0.943 |
| 31 | 1.066 | 0.738 | 1.118 | 0.799 | 1.260 | 0.799 | 1.275 | 0.837 | 1.310 | 0.825 | 1.120 | 0.895 |
| 32 | 1.054 | 0.660 | 1.103 | 0.654 | 1.289 | 0.715 | 1.000 | 0.734 | 1.350 | 0.829 | 1.189 | 0.911 |
| 33 | 1.115 | 0.737 | 1.138 | 0.792 | 1.226 | 0.750 | 1.278 | 0.854 | 1.380 | 0.725 | 1.178 | 0.914 |
| 34 | 1.102 | 0.707 | 1.154 | 0.753 | 1.210 | 0.734 | 1.228 | 0.663 | 1.204 | 0.725 | 1.141 | 0.883 |
| 35 | 1.023 | 0.707 | 1.167 | 0.758 | 1.268 | 0.715 | 1.220 | 0.842 | 1.293 | 0.813 | 1.107 | 0.926 |
| 36 | 1.039 | 0.753 | 1.125 | 0.657 | 1.242 | 0.795 | 1.272 | 0.811 | 1.105 | 0.783 | 1.194 | 0.871 |

ตารางที่ 56 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 50 **และ** PS 10 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มไม่เท่าเทียม**โดยใช้ข้อสอบร่วม (ต่อ)

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 37 | 1.118 | 0.722 | 0.998 | 0.724 | 1.215 | 0.752 | 1.244 | 0.827 | 1.265 | 0.852 | 1.188 | 0.914 |
| 38 | 1.118 | 0.653 | 1.126 | 0.806 | 1.226 | 0.700 | 1.209 | 0.713 | 1.261 | 0.845 | 1.185 | 0.846 |
| 39 | 1.054 | 0.651 | 1.105 | 0.840 | 1.105 | 0.730 | 1.223 | 0.839 | 1.291 | 0.730 | 1.155 | 0.892 |
| 40 | 1.120 | 0.724 | 1.162 | 0.840 | 1.283 | 0.797 | 1.211 | 0.734 | 1.204 | 0.818 | 1.169 | 0.855 |
| 41 | 1.010 | 0.743 | 1.118 | 0.600 | 1.288 | 0.744 | 1.215 | 0.668 | 1.291 | 0.840 | 1.191 | 0.906 |
| 42 | 1.071 | 0.655 | 1.109 | 0.842 | 1.123 | 0.740 | 1.214 | 0.669 | 1.294 | 0.784 | 1.176 | 0.926 |
| 43 | 1.090 | 0.652 | 1.154 | 0.650 | 1.214 | 0.726 | 1.213 | 0.679 | 1.267 | 0.805 | 1.120 | 0.815 |
| 44 | 1.105 | 0.703 | 1.174 | 0.854 | 1.237 | 0.551 | 1.229 | 0.669 | 1.206 | 0.804 | 1.144 | 0.921 |
| 45 | 1.058 | 0.751 | 1.109 | 0.754 | 1.289 | 0.770 | 1.250 | 0.844 | 1.276 | 0.723 | 1.166 | 0.949 |
| 46 | 1.071 | 0.658 | 1.189 | 0.600 | 1.265 | 0.715 | 1.226 | 0.665 | 1.270 | 0.772 | 1.109 | 0.842 |
| 47 | 1.063 | 0.706 | 1.192 | 0.606 | 1.210 | 0.552 | 1.202 | 0.735 | 1.204 | 0.798 | 1.148 | 0.821 |
| 48 | 1.047 | 0.657 | 1.131 | 0.830 | 1.111 | 0.552 | 1.246 | 0.814 | 1.340 | 0.851 | 1.182 | 0.954 |
| 49 | 1.050 | 0.739 | 1.182 | 0.755 | 1.291 | 0.553 | 1.280 | 0.855 | 1.300 | 0.836 | 1.116 | 0.883 |
| 50 | 1.045 | 0.653 | 1.165 | 0.714 | 1.242 | 0.797 | 1.297 | 0.668 | 1.260 | 0.728 | 1.194 | 0.831 |
| 51 | 1.076 | 0.702 | 1.105 | 0.604 | 1.233 | 0.756 | 1.288 | 0.804 | 1.266 | 0.812 | 1.120 | 0.932 |
| 52 | 1.120 | 0.721 | 1.173 | 0.797 | 1.281 | 0.737 | 1.243 | 0.848 | 1.208 | 0.766 | 1.167 | 0.857 |
| 53 | 1.036 | 0.656 | 1.113 | 0.606 | 1.285 | 0.736 | 1.213 | 0.803 | 1.360 | 0.819 | 1.113 | 0.937 |
| 54 | 1.006 | 0.658 | 1.129 | 0.604 | 1.122 | 0.552 | 1.288 | 0.812 | 1.106 | 0.756 | 1.151 | 0.834 |
| 55 | 1.093 | 0.659 | 1.176 | 0.855 | 1.213 | 0.740 | 1.000 | 0.750 | 1.360 | 0.832 | 1.132 | 0.900 |
| 56 | 1.026 | 0.721 | 1.109 | 0.652 | 1.216 | 0.746 | 1.254 | 0.817 | 1.320 | 0.858 | 1.113 | 0.914 |
| 57 | 1.105 | 0.700 | 1.174 | 0.794 | 1.265 | 0.706 | 1.210 | 0.684 | 1.100 | 0.850 | 1.109 | 0.923 |
| 58 | 1.103 | 0.654 | 1.102 | 0.602 | 1.114 | 0.742 | 1.232 | 0.804 | 1.106 | 0.805 | 1.168 | 0.879 |
| 59 | 1.050 | 0.652 | 1.119 | 0.805 | 1.230 | 0.552 | 1.223 | 0.858 | 1.238 | 0.738 | 1.128 | 0.941 |
| 60 | 1.060 | 0.652 | 1.164 | 0.755 | 1.225 | 0.700 | 1.285 | 0.713 | 1.276 | 0.733 | 1.139 | 0.883 |
| 61 | 1.063 | 0.654 | 1.104 | 0.605 | 1.220 | 0.639 | 1.207 | 0.757 | 1.247 | 0.729 | 1.143 | 0.920 |
| 62 | 1.077 | 0.750 | 1.180 | 0.854 | 1.284 | 0.790 | 1.234 | 0.667 | 1.120 | 0.790 | 1.136 | 0.919 |
| 63 | 1.112 | 0.703 | 1.161 | 0.838 | 1.213 | 0.640 | 1.294 | 0.678 | 1.370 | 0.786 | 1.168 | 0.913 |
| 64 | 1.082 | 0.654 | 1.143 | 0.608 | 1.243 | 0.791 | 1.242 | 0.833 | 1.261 | 0.762 | 1.118 | 0.873 |
| 65 | 1.028 | 0.656 | 1.136 | 0.799 | 1.290 | 0.740 | 1.222 | 0.694 | 1.104 | 0.745 | 1.117 | 0.830 |
| 66 | 1.118 | 0.656 | 1.104 | 0.608 | 1.234 | 0.711 | 1.204 | 0.822 | 1.271 | 0.766 | 1.142 | 0.853 |
| 67 | 1.087 | 0.656 | 1.135 | 0.795 | 1.297 | 0.731 | 1.270 | 0.739 | 1.380 | 0.821 | 1.124 | 0.836 |
| 68 | 1.018 | 0.735 | 1.118 | 0.827 | 1.222 | 0.730 | 1.288 | 0.754 | 1.350 | 0.799 | 1.111 | 0.944 |
| 69 | 1.042 | 0.656 | 1.178 | 0.608 | 1.115 | 0.780 | 1.219 | 0.680 | 1.268 | 0.829 | 1.191 | 0.848 |
| 70 | 1.012 | 0.730 | 1.115 | 0.726 | 1.100 | 0.799 | 1.215 | 0.672 | 1.106 | 0.838 | 1.143 | 0.876 |
| 71 | 1.001 | 0.729 | 0.998 | 0.656 | 1.203 | 0.792 | 1.213 | 0.757 | 1.249 | 0.833 | 1.132 | 0.865 |
| 72 | 1.019 | 0.739 | 1.146 | 0.808 | 1.224 | 0.796 | 1.212 | 0.838 | 1.299 | 0.846 | 1.182 | 0.890 |

ตารางที่ 56 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 50 **และ** PS 10 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มไม่เท่าเทียม**โดยใช้ข้อสอบร่วม (ต่อ)

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 73 | 1.100 | 0.654 | 1.124 | 0.801 | 1.106 | 0.725 | 1.000 | 0.739 | 1.390 | 0.770 | 1.193 | 0.941 |
| 74 | 1.064 | 0.659 | 1.194 | 0.858 | 1.243 | 0.638 | 1.273 | 0.743 | 1.117 | 0.813 | 1.171 | 0.840 |
| 75 | 1.073 | 0.708 | 1.118 | 0.657 | 1.282 | 0.638 | 1.214 | 0.757 | 1.112 | 0.825 | 1.106 | 0.869 |
| 76 | 1.023 | 0.654 | 1.125 | 0.814 | 1.298 | 0.640 | 1.213 | 0.841 | 1.231 | 0.722 | 1.120 | 0.938 |
| 77 | 1.081 | 0.721 | 1.101 | 0.606 | 1.109 | 0.639 | 1.209 | 0.748 | 1.340 | 0.750 | 1.104 | 0.951 |
| 78 | 1.051 | 0.659 | 1.175 | 0.604 | 1.215 | 0.552 | 1.239 | 0.838 | 1.236 | 0.770 | 1.193 | 0.825 |
| 79 | 1.106 | 0.652 | 1.121 | 0.602 | 1.210 | 0.708 | 1.250 | 0.732 | 1.320 | 0.857 | 1.151 | 0.909 |
| 80 | 1.103 | 0.653 | 1.102 | 0.727 | 1.266 | 0.640 | 1.220 | 0.663 | 1.232 | 0.799 | 1.187 | 0.870 |
| 81 | 1.086 | 0.712 | 1.116 | 0.653 | 1.105 | 0.553 | 1.259 | 0.751 | 1.290 | 0.758 | 1.182 | 0.882 |
| 82 | 1.084 | 0.744 | 1.174 | 0.839 | 1.107 | 0.790 | 1.283 | 0.829 | 1.241 | 0.791 | 1.130 | 0.864 |
| 83 | 1.022 | 0.704 | 1.123 | 0.791 | 1.219 | 0.799 | 1.211 | 0.688 | 1.290 | 0.844 | 1.134 | 0.893 |
| 84 | 1.093 | 0.724 | 1.109 | 0.728 | 1.284 | 0.741 | 1.231 | 0.840 | 1.242 | 0.727 | 1.130 | 0.941 |
| 85 | 1.056 | 0.653 | 1.116 | 0.794 | 1.271 | 0.552 | 1.299 | 0.847 | 1.101 | 0.763 | 1.105 | 0.928 |
| 86 | 1.002 | 0.712 | 1.139 | 0.850 | 1.258 | 0.715 | 1.226 | 0.663 | 1.245 | 0.722 | 1.158 | 0.859 |
| 87 | 1.061 | 0.656 | 1.109 | 0.825 | 1.231 | 0.640 | 0.999 | 0.691 | 1.121 | 0.818 | 1.166 | 0.836 |
| 88 | 0.999 | 0.751 | 1.137 | 0.828 | 1.115 | 0.703 | 1.219 | 0.716 | 1.297 | 0.818 | 1.120 | 0.931 |
| 89 | 1.050 | 0.655 | 1.101 | 0.844 | 1.263 | 0.740 | 1.205 | 0.729 | 1.370 | 0.780 | 1.160 | 0.957 |
| 90 | 1.046 | 0.706 | 1.189 | 0.855 | 1.110 | 0.553 | 1.201 | 0.729 | 1.350 | 0.851 | 1.157 | 0.815 |
| 91 | 1.012 | 0.741 | 1.107 | 0.816 | 1.272 | 0.792 | 1.222 | 0.681 | 1.242 | 0.795 | 1.110 | 0.875 |
| 92 | 1.058 | 0.653 | 1.136 | 0.721 | 1.255 | 0.700 | 1.289 | 0.826 | 1.115 | 0.851 | 1.119 | 0.839 |
| 93 | 1.048 | 0.655 | 1.113 | 0.722 | 1.288 | 0.553 | 1.209 | 0.677 | 1.390 | 0.830 | 1.100 | 0.903 |
| 94 | 1.061 | 0.657 | 1.131 | 0.829 | 1.232 | 0.740 | 1.205 | 0.800 | 1.000 | 0.729 | 1.104 | 0.828 |
| 95 | 1.094 | 0.659 | 1.125 | 0.792 | 1.296 | 0.553 | 1.297 | 0.837 | 1.297 | 0.822 | 1.122 | 0.876 |
| 96 | 1.087 | 0.703 | 1.174 | 0.849 | 1.217 | 0.759 | 1.242 | 0.660 | 1.340 | 0.799 | 1.137 | 0.891 |
| 97 | 1.043 | 0.660 | 1.143 | 0.796 | 1.204 | 0.552 | 1.202 | 0.845 | 1.121 | 0.783 | 1.119 | 0.908 |
| 98 | 1.111 | 0.740 | 1.135 | 0.798 | 1.239 | 0.553 | 1.202 | 0.729 | 1.205 | 0.789 | 1.110 | 0.959 |
| 99 | 0.999 | 0.660 | 1.116 | 0.858 | 1.291 | 0.553 | 1.208 | 0.834 | 1.300 | 0.775 | 1.122 | 0.817 |
| 100 | 1.112 | 0.655 | 1.104 | 0.720 | 1.218 | 0.639 | 1.297 | 0.723 | 1.233 | 0.723 | 1.137 | 0.840 |

ตารางที่ 57 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 60 **และ** PS 8 ด้วยการเก็บรวบรวมข้อมูลแบบก**ลุ่มไม่เท่าเทียม**โดยใช้ข้อสอบร่วม

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 1 | 1.125 | 0.760 | 1.222 | 0.838 | 1.222 | 0.752 | 1.128 | 0.720 | 1.100 | 0.785 | 1.030 | 0.700 |
| 2 | 1.137 | 0.736 | 1.123 | 0.845 | 1.040 | 0.604 | 1.158 | 0.785 | 1.250 | 0.842 | 1.119 | 0.845 |
| 3 | 1.133 | 0.789 | 1.251 | 0.817 | 1.079 | 0.617 | 1.171 | 0.858 | 1.240 | 0.806 | 1.090 | 0.840 |
| 4 | 1.104 | 0.755 | 1.201 | 0.844 | 1.017 | 0.772 | 1.131 | 0.776 | 1.136 | 0.720 | 1.021 | 0.769 |
| 5 | 1.152 | 0.726 | 1.233 | 0.745 | 1.003 | 0.734 | 1.108 | 0.858 | 1.103 | 0.785 | 1.038 | 0.754 |
| 6 | 1.118 | 0.765 | 1.231 | 0.795 | 1.011 | 0.777 | 0.999 | 0.803 | 1.280 | 0.721 | 1.044 | 0.762 |
| 7 | 1.160 | 0.515 | 1.227 | 0.830 | 1.211 | 0.746 | 1.110 | 0.859 | 1.178 | 0.844 | 1.002 | 0.880 |
| 8 | 1.149 | 0.507 | 1.000 | 0.813 | 1.216 | 0.753 | 1.102 | 0.788 | 1.134 | 0.742 | 1.008 | 0.777 |
| 9 | 1.177 | 0.501 | 1.237 | 0.844 | 1.018 | 0.601 | 1.115 | 0.610 | 1.193 | 0.786 | 1.056 | 0.850 |
| 10 | 1.101 | 0.758 | 1.250 | 0.806 | 1.054 | 0.730 | 1.103 | 0.720 | 1.107 | 0.851 | 1.080 | 0.781 |
| 11 | 1.103 | 0.503 | 1.238 | 0.786 | 1.221 | 0.751 | 1.176 | 0.799 | 1.250 | 0.821 | 1.031 | 0.830 |
| 12 | 1.101 | 0.531 | 1.246 | 0.725 | 1.078 | 0.744 | 1.131 | 0.771 | 1.141 | 0.771 | 1.043 | 0.762 |
| 13 | 1.126 | 0.797 | 1.268 | 0.789 | 1.068 | 0.640 | 1.110 | 0.763 | 1.125 | 0.857 | 0.998 | 0.824 |
| 14 | 1.115 | 0.741 | 1.113 | 0.794 | 1.010 | 0.761 | 1.130 | 0.756 | 1.105 | 0.779 | 1.024 | 0.764 |
| 15 | 1.132 | 0.529 | 1.241 | 0.799 | 1.094 | 0.639 | 1.118 | 0.751 | 1.141 | 0.790 | 1.044 | 0.820 |
| 16 | 1.134 | 0.720 | 0.998 | 0.787 | 1.045 | 0.649 | 1.191 | 0.766 | 1.158 | 0.763 | 1.120 | 0.750 |
| 17 | 1.101 | 0.529 | 1.113 | 0.770 | 1.039 | 0.767 | 1.144 | 0.600 | 1.195 | 0.778 | 1.072 | 0.853 |
| 18 | 1.148 | 0.555 | 1.247 | 0.822 | 1.007 | 0.723 | 1.163 | 0.792 | 1.191 | 0.778 | 1.041 | 0.850 |
| 19 | 1.141 | 0.770 | 1.259 | 0.839 | 1.027 | 0.730 | 1.119 | 0.793 | 1.154 | 0.818 | 1.057 | 0.840 |
| 20 | 1.112 | 0.515 | 1.209 | 0.776 | 1.003 | 0.659 | 1.144 | 0.807 | 1.121 | 0.785 | 1.046 | 0.750 |
| 21 | 1.181 | 0.531 | 1.243 | 0.811 | 1.050 | 0.657 | 1.105 | 0.630 | 1.230 | 0.721 | 1.056 | 0.810 |
| 22 | 0.999 | 0.757 | 1.229 | 0.717 | 1.204 | 0.720 | 1.104 | 0.789 | 1.148 | 0.809 | 1.086 | 0.860 |
| 23 | 1.141 | 0.791 | 1.274 | 0.846 | 1.003 | 0.655 | 1.108 | 0.760 | 1.181 | 0.738 | 1.085 | 0.825 |
| 24 | 1.183 | 0.739 | 1.110 | 0.840 | 1.057 | 0.763 | 1.146 | 0.778 | 1.280 | 0.725 | 1.001 | 0.754 |
| 25 | 1.139 | 0.546 | 1.277 | 0.836 | 1.201 | 0.782 | 1.117 | 0.739 | 1.159 | 0.812 | 1.104 | 0.850 |
| 26 | 1.142 | 0.505 | 1.269 | 0.751 | 1.089 | 0.757 | 1.193 | 0.808 | 1.173 | 0.835 | 1.112 | 0.870 |
| 27 | 1.169 | 0.544 | 1.264 | 0.745 | 1.039 | 0.763 | 0.998 | 0.768 | 1.147 | 0.820 | 1.029 | 0.810 |
| 28 | 1.160 | 0.753 | 1.299 | 0.740 | 1.002 | 0.719 | 1.190 | 0.740 | 1.200 | 0.771 | 1.010 | 0.784 |
| 29 | 1.182 | 0.758 | 1.223 | 0.774 | 1.006 | 0.736 | 1.199 | 0.782 | 1.150 | 0.720 | 1.084 | 0.859 |
| 30 | 1.117 | 0.506 | 1.229 | 0.746 | 1.059 | 0.732 | 1.189 | 0.777 | 1.107 | 0.729 | 1.055 | 0.799 |
| 31 | 1.114 | 0.791 | 1.273 | 0.785 | 1.204 | 0.776 | 1.190 | 0.751 | 1.156 | 0.721 | 1.090 | 0.801 |
| 32 | 1.153 | 0.506 | 1.264 | 0.723 | 1.039 | 0.760 | 1.100 | 0.721 | 1.180 | 0.751 | 1.073 | 0.810 |
| 33 | 1.106 | 0.501 | 1.257 | 0.842 | 1.206 | 0.650 | 0.998 | 0.799 | 1.179 | 0.742 | 1.071 | 0.838 |
| 34 | 1.177 | 0.535 | 1.247 | 0.827 | 1.031 | 0.741 | 1.112 | 0.730 | 1.107 | 0.808 | 1.100 | 0.800 |
| 35 | 1.195 | 0.553 | 1.269 | 0.749 | 1.027 | 0.755 | 1.115 | 0.762 | 1.195 | 0.814 | 1.091 | 0.769 |
| 36 | 1.128 | 0.796 | 1.280 | 0.831 | 1.030 | 0.634 | 1.157 | 0.620 | 1.194 | 0.837 | 1.067 | 0.870 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 57 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 60 **และ** PS 8 ด้วยการเก็บรวบรวมข้อมูลแบบก**ลุ่มไม่เท่าเทียม**โดยใช้ข้อสอบร่วม (ต่อ)

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 37 | 1.198 | 0.782 | 1.226 | 0.845 | 1.012 | 0.775 | 1.166 | 0.718 | 1.145 | 0.744 | 1.001 | 0.784 |
| 38 | 1.122 | 0.784 | 1.118 | 0.747 | 1.087 | 0.650 | 1.126 | 0.790 | 1.195 | 0.793 | 1.108 | 0.804 |
| 39 | 1.122 | 0.779 | 1.276 | 0.802 | 1.074 | 0.609 | 1.182 | 0.733 | 1.220 | 0.773 | 1.084 | 0.785 |
| 40 | 1.113 | 0.714 | 1.251 | 0.835 | 1.029 | 0.734 | 1.135 | 0.781 | 1.240 | 0.809 | 1.066 | 0.787 |
| 41 | 1.104 | 0.508 | 1.239 | 0.827 | 1.005 | 0.778 | 1.125 | 0.650 | 1.180 | 0.733 | 1.097 | 0.752 |
| 42 | 1.134 | 0.518 | 1.279 | 0.858 | 1.044 | 0.744 | 1.121 | 0.753 | 1.108 | 0.857 | 1.069 | 0.788 |
| 43 | 1.134 | 0.510 | 1.292 | 0.772 | 1.049 | 0.721 | 1.132 | 0.850 | 1.121 | 0.827 | 1.107 | 0.769 |
| 44 | 1.190 | 0.532 | 1.287 | 0.779 | 1.098 | 0.793 | 1.113 | 0.748 | 1.195 | 0.830 | 1.059 | 0.766 |
| 45 | 1.191 | 0.728 | 1.287 | 0.735 | 1.096 | 0.619 | 1.110 | 0.730 | 1.270 | 0.741 | 1.087 | 0.766 |
| 46 | 1.168 | 0.553 | 1.202 | 0.780 | 1.036 | 0.738 | 1.163 | 0.726 | 1.166 | 0.796 | 1.089 | 0.788 |
| 47 | 1.197 | 0.777 | 1.234 | 0.787 | 1.057 | 0.724 | 1.147 | 0.795 | 0.999 | 0.857 | 1.099 | 0.824 |
| 48 | 1.160 | 0.757 | 1.238 | 0.840 | 1.026 | 0.718 | 1.198 | 0.769 | 1.210 | 0.788 | 1.079 | 0.766 |
| 49 | 1.107 | 0.785 | 1.263 | 0.851 | 1.210 | 0.633 | 1.122 | 0.840 | 1.104 | 0.750 | 1.073 | 0.850 |
| 50 | 1.127 | 0.788 | 1.118 | 0.737 | 1.002 | 0.773 | 1.109 | 0.788 | 1.160 | 0.717 | 1.033 | 0.805 |
| 51 | 1.111 | 0.540 | 1.291 | 0.771 | 1.042 | 0.714 | 1.141 | 0.784 | 1.154 | 0.850 | 1.100 | 0.808 |
| 52 | 1.104 | 0.553 | 1.259 | 0.783 | 1.084 | 0.799 | 1.123 | 0.757 | 1.220 | 0.847 | 1.107 | 0.880 |
| 53 | 1.103 | 0.542 | 1.273 | 0.798 | 1.058 | 0.768 | 1.154 | 0.762 | 1.134 | 0.788 | 1.049 | 0.850 |
| 54 | 1.192 | 0.787 | 1.220 | 0.735 | 1.075 | 0.764 | 1.122 | 0.620 | 1.129 | 0.724 | 1.073 | 0.857 |
| 55 | 1.140 | 0.511 | 1.114 | 0.857 | 1.208 | 0.604 | 1.183 | 0.734 | 1.106 | 0.795 | 1.103 | 0.821 |
| 56 | 1.103 | 0.787 | 1.280 | 0.761 | 1.056 | 0.613 | 1.160 | 0.782 | 1.260 | 0.718 | 1.017 | 0.821 |
| 57 | 1.127 | 0.513 | 1.224 | 0.776 | 1.033 | 0.790 | 1.112 | 0.801 | 1.123 | 0.782 | 1.095 | 0.812 |
| 58 | 1.183 | 0.559 | 1.230 | 0.805 | 1.009 | 0.643 | 1.121 | 0.833 | 1.186 | 0.792 | 1.016 | 0.854 |
| 59 | 1.156 | 0.515 | 1.225 | 0.816 | 1.004 | 0.755 | 0.998 | 0.791 | 1.120 | 0.771 | 1.064 | 0.851 |
| 60 | 1.119 | 0.733 | 1.258 | 0.851 | 1.052 | 0.721 | 1.182 | 0.620 | 1.154 | 0.722 | 1.043 | 0.845 |
| 61 | 1.167 | 0.542 | 1.117 | 0.788 | 1.024 | 0.602 | 1.124 | 0.660 | 1.125 | 0.728 | 1.078 | 0.880 |
| 62 | 1.102 | 0.770 | 1.284 | 0.839 | 1.008 | 0.758 | 1.111 | 0.853 | 1.107 | 0.830 | 1.047 | 0.800 |
| 63 | 1.110 | 0.535 | 1.219 | 0.834 | 1.002 | 0.751 | 1.107 | 0.715 | 1.123 | 0.773 | 1.067 | 0.752 |
| 64 | 1.150 | 0.720 | 1.208 | 0.758 | 1.207 | 0.617 | 1.132 | 0.837 | 1.121 | 0.839 | 1.030 | 0.800 |
| 65 | 1.150 | 0.779 | 1.121 | 0.777 | 1.221 | 0.735 | 1.129 | 0.720 | 1.102 | 0.762 | 1.028 | 0.753 |
| 66 | 1.183 | 0.720 | 1.237 | 0.758 | 1.087 | 0.767 | 1.167 | 0.750 | 1.101 | 0.767 | 1.113 | 0.800 |
| 67 | 1.137 | 0.717 | 1.242 | 0.723 | 1.065 | 0.755 | 1.111 | 0.772 | 1.127 | 0.730 | 1.024 | 0.833 |
| 68 | 1.170 | 0.741 | 1.272 | 0.731 | 1.056 | 0.746 | 1.153 | 0.731 | 1.185 | 0.827 | 1.098 | 0.880 |
| 69 | 0.999 | 0.725 | 1.258 | 0.806 | 1.015 | 0.717 | 1.116 | 0.733 | 1.195 | 0.824 | 1.037 | 0.812 |
| 70 | 1.119 | 0.744 | 1.116 | 0.859 | 1.087 | 0.621 | 1.107 | 0.788 | 1.104 | 0.833 | 1.103 | 0.850 |
| 71 | 1.163 | 0.506 | 1.249 | 0.767 | 1.046 | 0.778 | 1.145 | 0.758 | 1.105 | 0.836 | 1.026 | 0.840 |
| 72 | 0.999 | 0.758 | 1.113 | 0.765 | 1.068 | 0.744 | 1.122 | 0.798 | 0.999 | 0.802 | 1.074 | 0.837 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 57 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 60 **และ** PS 8 ด้วยการเก็บรวบรวมข้อมูลแบบก**ลุ่มไม่เท่าเทียม**โดยใช้ข้อสอบร่วม (ต่อ)

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 73 | 1.197 | 0.732 | 1.238 | 0.829 | 1.020 | 0.650 | 1.155 | 0.786 | 1.161 | 0.793 | 1.066 | 0.823 |
| 74 | 1.175 | 0.730 | 1.227 | 0.742 | 1.207 | 0.604 | 1.160 | 0.779 | 1.200 | 0.849 | 1.053 | 0.787 |
| 75 | 1.199 | 0.715 | 1.244 | 0.714 | 1.040 | 0.740 | 1.102 | 0.630 | 1.129 | 0.764 | 1.104 | 0.840 |
| 76 | 1.182 | 0.739 | 1.290 | 0.856 | 1.084 | 0.730 | 1.109 | 0.736 | 1.152 | 0.838 | 1.094 | 0.759 |
| 77 | 1.155 | 0.533 | 1.203 | 0.794 | 1.006 | 0.731 | 1.177 | 0.847 | 1.182 | 0.815 | 1.030 | 0.775 |
| 78 | 1.108 | 0.771 | 1.108 | 0.754 | 1.074 | 0.606 | 1.150 | 0.717 | 1.162 | 0.718 | 1.001 | 0.795 |
| 79 | 1.161 | 0.758 | 1.233 | 0.848 | 1.020 | 0.612 | 1.120 | 0.630 | 1.159 | 0.809 | 1.019 | 0.851 |
| 80 | 1.147 | 0.732 | 1.291 | 0.858 | 1.040 | 0.746 | 1.103 | 0.742 | 1.166 | 0.759 | 1.099 | 0.840 |
| 81 | 1.114 | 0.517 | 1.205 | 0.783 | 1.213 | 0.754 | 1.186 | 0.841 | 1.230 | 0.778 | 1.046 | 0.788 |
| 82 | 1.185 | 0.520 | 1.116 | 0.841 | 1.210 | 0.658 | 1.105 | 0.732 | 1.230 | 0.817 | 1.014 | 0.880 |
| 83 | 1.101 | 0.796 | 1.120 | 0.756 | 0.999 | 0.772 | 1.120 | 0.690 | 1.250 | 0.729 | 1.085 | 0.755 |
| 84 | 1.140 | 0.535 | 1.226 | 0.813 | 1.218 | 0.649 | 1.138 | 0.754 | 1.156 | 0.823 | 1.000 | 0.847 |
| 85 | 1.198 | 0.767 | 1.212 | 0.749 | 1.216 | 0.629 | 1.127 | 0.748 | 1.101 | 0.713 | 1.080 | 0.775 |
| 86 | 1.145 | 0.516 | 1.265 | 0.751 | 1.050 | 0.636 | 1.150 | 0.777 | 1.179 | 0.808 | 1.102 | 0.809 |
| 87 | 1.115 | 0.723 | 1.224 | 0.763 | 1.209 | 0.653 | 1.134 | 0.833 | 1.106 | 0.815 | 1.022 | 0.839 |
| 88 | 1.103 | 0.756 | 1.296 | 0.738 | 1.026 | 0.748 | 1.199 | 0.739 | 1.145 | 0.858 | 1.112 | 0.856 |
| 89 | 1.107 | 0.763 | 1.117 | 0.729 | 1.081 | 0.740 | 1.164 | 0.749 | 1.149 | 0.785 | 1.009 | 0.880 |
| 90 | 1.155 | 0.558 | 1.284 | 0.858 | 1.214 | 0.649 | 1.159 | 0.798 | 1.143 | 0.829 | 1.058 | 0.754 |
| 91 | 1.131 | 0.541 | 1.113 | 0.781 | 1.074 | 0.734 | 1.154 | 0.736 | 1.167 | 0.810 | 1.091 | 0.756 |
| 92 | 1.145 | 0.731 | 1.243 | 0.731 | 1.070 | 0.795 | 1.163 | 0.735 | 1.134 | 0.733 | 1.001 | 0.870 |
| 93 | 1.107 | 0.765 | 1.239 | 0.799 | 1.042 | 0.631 | 1.112 | 0.620 | 1.100 | 0.736 | 1.075 | 0.860 |
| 94 | 1.104 | 0.517 | 1.262 | 0.831 | 1.211 | 0.787 | 1.104 | 0.788 | 1.270 | 0.820 | 1.060 | 0.767 |
| 95 | 1.115 | 0.513 | 1.107 | 0.785 | 1.097 | 0.757 | 1.140 | 0.847 | 1.170 | 0.758 | 1.013 | 0.841 |
| 96 | 1.104 | 0.720 | 1.260 | 0.832 | 1.216 | 0.616 | 1.121 | 0.778 | 1.104 | 0.729 | 1.114 | 0.851 |
| 97 | 1.182 | 0.526 | 1.217 | 0.776 | 1.092 | 0.633 | 1.145 | 0.761 | 1.131 | 0.831 | 1.014 | 0.813 |
| 98 | 1.130 | 0.726 | 1.268 | 0.826 | 1.053 | 0.777 | 1.171 | 0.739 | 1.183 | 0.762 | 1.069 | 0.799 |
| 99 | 1.165 | 0.543 | 1.254 | 0.761 | 1.025 | 0.640 | 1.193 | 0.765 | 1.152 | 0.723 | 1.034 | 0.814 |
| 100 | 1.108 | 0.515 | 1.108 | 0.766 | 1.098 | 0.733 | 1.125 | 0.727 | 1.174 | 0.771 | 1.056 | 0.762 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 58 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 70 **และ** PS 6 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มไม่เท่าเทียม**โดยใช้ข้อสอบร่วม

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 1 | 1.183 | 0.816 | 1.076 | 0.792 | 1.119 | 0.757 | 1.000 | 0.750 | 1.371 | 0.811 | 1.254 | 0.827 |
| 2 | 1.186 | 0.765 | 1.000 | 0.751 | 1.259 | 0.761 | 1.346 | 0.845 | 1.328 | 0.750 | 1.293 | 0.804 |
| 3 | 1.103 | 0.736 | 1.082 | 0.754 | 1.289 | 0.847 | 1.348 | 0.846 | 1.321 | 0.798 | 1.264 | 0.762 |
| 4 | 1.128 | 0.843 | 1.079 | 0.724 | 1.204 | 0.799 | 1.349 | 0.786 | 1.304 | 0.827 | 1.267 | 0.850 |
| 5 | 1.191 | 0.854 | 1.065 | 0.751 | 1.282 | 0.849 | 1.113 | 0.798 | 1.305 | 0.761 | 1.223 | 0.806 |
| 6 | 1.100 | 0.839 | 1.047 | 0.722 | 1.118 | 0.517 | 1.390 | 0.817 | 1.320 | 0.744 | 1.000 | 0.852 |
| 7 | 0.999 | 0.856 | 1.056 | 0.750 | 1.101 | 0.522 | 1.352 | 0.766 | 1.369 | 0.734 | 1.258 | 0.851 |
| 8 | 1.181 | 0.811 | 1.098 | 0.727 | 1.258 | 0.774 | 1.341 | 0.520 | 1.371 | 0.716 | 1.103 | 0.803 |
| 9 | 1.186 | 0.737 | 1.096 | 0.755 | 1.253 | 0.734 | 1.336 | 0.764 | 1.375 | 0.722 | 1.223 | 0.758 |
| 10 | 1.162 | 0.843 | 1.095 | 0.750 | 1.290 | 0.756 | 1.332 | 0.751 | 1.326 | 0.844 | 1.290 | 0.716 |
| 11 | 1.157 | 0.754 | 1.024 | 0.741 | 1.258 | 0.721 | 1.301 | 0.771 | 1.356 | 0.818 | 1.252 | 0.821 |
| 12 | 1.110 | 0.836 | 1.011 | 0.752 | 1.297 | 0.795 | 1.329 | 0.832 | 0.998 | 0.817 | 1.102 | 0.780 |
| 13 | 1.145 | 0.809 | 1.057 | 0.755 | 1.116 | 0.767 | 1.397 | 0.524 | 1.336 | 0.717 | 1.109 | 0.764 |
| 14 | 1.176 | 0.812 | 1.014 | 0.753 | 1.215 | 0.845 | 1.353 | 0.752 | 1.373 | 0.733 | 1.107 | 0.784 |
| 15 | 1.119 | 0.777 | 1.010 | 0.769 | 1.103 | 0.806 | 1.106 | 0.850 | 1.338 | 0.837 | 1.279 | 0.834 |
| 16 | 1.115 | 0.732 | 1.105 | 0.790 | 1.237 | 0.527 | 1.107 | 0.831 | 1.393 | 0.715 | 1.270 | 0.853 |
| 17 | 1.188 | 0.833 | 1.036 | 0.755 | 1.298 | 0.813 | 1.388 | 0.746 | 1.337 | 0.758 | 1.289 | 0.830 |
| 18 | 1.174 | 0.768 | 1.086 | 0.739 | 1.273 | 0.813 | 1.365 | 0.830 | 1.323 | 0.856 | 1.244 | 0.790 |
| 19 | 1.127 | 0.716 | 1.051 | 0.748 | 1.274 | 0.770 | 1.324 | 0.830 | 1.308 | 0.781 | 1.238 | 0.823 |
| 20 | 1.121 | 0.800 | 1.111 | 0.756 | 1.297 | 0.797 | 1.335 | 0.719 | 1.304 | 0.852 | 1.234 | 0.758 |
| 21 | 1.165 | 0.716 | 1.027 | 0.751 | 1.232 | 0.516 | 1.345 | 0.759 | 1.325 | 0.769 | 1.276 | 0.810 |
| 22 | 1.146 | 0.744 | 1.114 | 0.754 | 1.279 | 0.843 | 1.000 | 0.779 | 1.302 | 0.761 | 1.104 | 0.783 |
| 23 | 1.165 | 0.773 | 1.123 | 0.752 | 1.235 | 0.800 | 1.379 | 0.758 | 1.322 | 0.851 | 1.282 | 0.795 |
| 24 | 1.151 | 0.847 | 1.088 | 0.737 | 1.258 | 0.514 | 1.360 | 0.767 | 1.339 | 0.741 | 1.216 | 0.734 |
| 25 | 1.147 | 0.830 | 1.053 | 0.750 | 1.227 | 0.731 | 1.349 | 0.857 | 1.349 | 0.832 | 1.275 | 0.796 |
| 26 | 1.162 | 0.842 | 1.066 | 0.751 | 1.121 | 0.768 | 1.327 | 0.815 | 1.303 | 0.830 | 1.258 | 0.818 |
| 27 | 1.131 | 0.804 | 1.078 | 0.771 | 1.105 | 0.512 | 1.332 | 0.726 | 1.385 | 0.781 | 1.229 | 0.737 |
| 28 | 1.113 | 0.843 | 1.075 | 0.716 | 1.238 | 0.812 | 1.308 | 0.522 | 1.308 | 0.722 | 1.232 | 0.798 |
| 29 | 1.108 | 0.762 | 1.069 | 0.790 | 1.243 | 0.783 | 1.370 | 0.858 | 1.397 | 0.742 | 1.105 | 0.826 |
| 30 | 1.170 | 0.737 | 1.122 | 0.753 | 1.110 | 0.746 | 1.329 | 0.794 | 1.321 | 0.848 | 1.212 | 0.835 |
| 31 | 1.191 | 0.765 | 1.005 | 0.752 | 1.260 | 0.815 | 1.113 | 0.736 | 1.368 | 0.791 | 1.266 | 0.843 |
| 32 | 1.113 | 0.742 | 1.033 | 0.744 | 1.264 | 0.802 | 1.335 | 0.846 | 1.301 | 0.829 | 1.205 | 0.778 |
| 33 | 1.101 | 0.734 | 1.011 | 0.755 | 1.271 | 0.794 | 1.381 | 0.805 | 1.395 | 0.769 | 1.297 | 0.760 |
| 34 | 1.100 | 0.835 | 1.009 | 0.753 | 1.109 | 0.802 | 1.000 | 0.780 | 1.374 | 0.781 | 1.285 | 0.858 |
| 35 | 1.137 | 0.765 | 1.112 | 0.726 | 1.291 | 0.851 | 1.100 | 0.749 | 1.328 | 0.841 | 1.296 | 0.757 |
| 36 | 1.175 | 0.775 | 1.064 | 0.728 | 1.240 | 0.841 | 1.340 | 0.783 | 1.348 | 0.724 | 1.272 | 0.835 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 58 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 70 **และ** PS 6 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มไม่เท่าเทียม**โดยใช้ข้อสอบร่วม (ต่อ)

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 37 | 1.121 | 0.715 | 1.016 | 0.765 | 1.299 | 0.855 | 1.386 | 0.810 | 1.352 | 0.844 | 1.280 | 0.789 |
| 38 | 1.000 | 0.829 | 1.106 | 0.751 | 1.273 | 0.857 | 1.391 | 0.774 | 1.397 | 0.733 | 1.111 | 0.796 |
| 39 | 1.180 | 0.856 | 1.017 | 0.752 | 1.296 | 0.746 | 1.341 | 0.772 | 1.365 | 0.782 | 1.201 | 0.759 |
| 40 | 1.119 | 0.760 | 1.040 | 0.720 | 1.121 | 0.725 | 1.397 | 0.733 | 1.329 | 0.825 | 1.263 | 0.835 |
| 41 | 1.104 | 0.751 | 1.038 | 0.747 | 1.102 | 0.781 | 1.349 | 0.793 | 1.317 | 0.814 | 1.213 | 0.807 |
| 42 | 1.183 | 0.781 | 1.091 | 0.732 | 1.110 | 0.776 | 1.371 | 0.772 | 1.327 | 0.775 | 0.999 | 0.722 |
| 43 | 1.144 | 0.714 | 1.072 | 0.750 | 1.203 | 0.763 | 1.320 | 0.527 | 1.351 | 0.842 | 1.260 | 0.784 |
| 44 | 1.121 | 0.763 | 1.012 | 0.754 | 1.000 | 0.806 | 1.360 | 0.746 | 1.325 | 0.853 | 0.998 | 0.798 |
| 45 | 1.197 | 0.730 | 0.998 | 0.750 | 1.271 | 0.765 | 1.114 | 0.838 | 1.348 | 0.720 | 1.232 | 0.840 |
| 46 | 1.143 | 0.776 | 1.111 | 0.787 | 1.237 | 0.760 | 1.355 | 0.857 | 1.311 | 0.825 | 1.204 | 0.747 |
| 47 | 1.100 | 0.824 | 1.087 | 0.755 | 1.122 | 0.746 | 1.371 | 0.522 | 1.352 | 0.839 | 1.216 | 0.804 |
| 48 | 1.146 | 0.766 | 1.021 | 0.753 | 1.296 | 0.792 | 1.356 | 0.780 | 1.366 | 0.747 | 1.120 | 0.784 |
| 49 | 1.168 | 0.782 | 1.094 | 0.751 | 1.270 | 0.746 | 1.388 | 0.716 | 1.372 | 0.832 | 1.100 | 0.788 |
| 50 | 1.143 | 0.762 | 1.117 | 0.777 | 1.280 | 0.745 | 1.328 | 0.744 | 1.326 | 0.850 | 1.225 | 0.784 |
| 51 | 1.122 | 0.769 | 1.095 | 0.754 | 1.233 | 0.756 | 1.122 | 0.743 | 1.323 | 0.737 | 1.259 | 0.796 |
| 52 | 1.153 | 0.823 | 1.070 | 0.720 | 1.227 | 0.767 | 1.324 | 0.851 | 1.316 | 0.735 | 1.287 | 0.827 |
| 53 | 1.117 | 0.833 | 1.008 | 0.754 | 1.270 | 0.843 | 1.101 | 0.748 | 1.345 | 0.838 | 1.267 | 0.779 |
| 54 | 1.170 | 0.714 | 1.098 | 0.723 | 1.112 | 0.855 | 1.316 | 0.525 | 1.350 | 0.812 | 1.255 | 0.786 |
| 55 | 1.101 | 0.778 | 1.016 | 0.751 | 1.123 | 0.858 | 1.103 | 0.736 | 1.334 | 0.802 | 1.284 | 0.763 |
| 56 | 1.114 | 0.798 | 1.003 | 0.751 | 1.106 | 0.797 | 1.327 | 0.783 | 1.312 | 0.720 | 1.266 | 0.752 |
| 57 | 1.195 | 0.782 | 1.031 | 0.755 | 1.271 | 0.801 | 1.335 | 0.817 | 0.999 | 0.819 | 1.000 | 0.833 |
| 58 | 1.183 | 0.837 | 1.095 | 0.776 | 1.247 | 0.739 | 1.342 | 0.765 | 1.320 | 0.801 | 1.252 | 0.733 |
| 59 | 1.198 | 0.828 | 1.103 | 0.762 | 1.233 | 0.811 | 1.378 | 0.794 | 1.335 | 0.776 | 1.107 | 0.773 |
| 60 | 1.150 | 0.743 | 1.035 | 0.751 | 1.207 | 0.852 | 1.399 | 0.717 | 1.359 | 0.718 | 1.110 | 0.763 |
| 61 | 1.134 | 0.855 | 1.086 | 0.770 | 1.202 | 0.783 | 1.107 | 0.728 | 1.312 | 0.811 | 1.261 | 0.714 |
| 62 | 1.156 | 0.787 | 1.007 | 0.730 | 1.212 | 0.512 | 1.112 | 0.764 | 1.366 | 0.761 | 1.274 | 0.799 |
| 63 | 1.194 | 0.793 | 1.047 | 0.738 | 1.219 | 0.802 | 1.337 | 0.766 | 1.317 | 0.837 | 1.288 | 0.756 |
| 64 | 1.198 | 0.852 | 1.070 | 0.752 | 1.271 | 0.780 | 1.122 | 0.788 | 1.335 | 0.751 | 1.267 | 0.751 |
| 65 | 1.170 | 0.838 | 1.032 | 0.773 | 1.229 | 0.749 | 1.121 | 0.742 | 1.376 | 0.828 | 1.226 | 0.832 |
| 66 | 1.165 | 0.819 | 1.040 | 0.730 | 1.251 | 0.526 | 1.107 | 0.810 | 1.330 | 0.838 | 1.278 | 0.820 |
| 67 | 1.123 | 0.793 | 1.112 | 0.745 | 1.275 | 0.769 | 1.394 | 0.842 | 1.387 | 0.844 | 1.233 | 0.770 |
| 68 | 1.102 | 0.743 | 1.045 | 0.752 | 1.236 | 0.742 | 1.105 | 0.833 | 1.395 | 0.778 | 1.218 | 0.806 |
| 69 | 1.139 | 0.793 | 1.009 | 0.794 | 1.213 | 0.735 | 1.113 | 0.844 | 1.320 | 0.785 | 1.295 | 0.850 |
| 70 | 1.117 | 0.788 | 1.070 | 0.714 | 1.115 | 0.716 | 1.383 | 0.845 | 1.314 | 0.748 | 1.204 | 0.832 |
| 71 | 1.192 | 0.835 | 1.049 | 0.754 | 1.110 | 0.773 | 1.389 | 0.764 | 1.371 | 0.787 | 1.205 | 0.719 |
| 72 | 1.120 | 0.807 | 1.053 | 0.748 | 1.232 | 0.813 | 1.307 | 0.724 | 1.312 | 0.803 | 1.201 | 0.749 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 58 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 70 **และ** PS 6 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มไม่เท่าเทียม**โดยใช้ข้อสอบร่วม (ต่อ)

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 73 | 1.117 | 0.802 | 1.048 | 0.725 | 1.211 | 0.742 | 1.335 | 0.774 | 1.366 | 0.748 | 1.105 | 0.809 |
| 74 | 1.150 | 0.789 | 1.038 | 0.760 | 1.277 | 0.798 | 1.101 | 0.526 | 1.387 | 0.778 | 1.211 | 0.757 |
| 75 | 1.102 | 0.790 | 1.047 | 0.784 | 1.212 | 0.523 | 1.398 | 0.522 | 1.323 | 0.832 | 1.241 | 0.843 |
| 76 | 1.176 | 0.819 | 1.011 | 0.752 | 1.114 | 0.768 | 1.373 | 0.733 | 1.325 | 0.737 | 1.209 | 0.857 |
| 77 | 1.197 | 0.736 | 1.103 | 0.769 | 1.246 | 0.799 | 1.389 | 0.817 | 1.347 | 0.778 | 1.219 | 0.828 |
| 78 | 1.173 | 0.812 | 1.120 | 0.740 | 1.231 | 0.752 | 1.387 | 0.812 | 1.317 | 0.789 | 1.115 | 0.722 |
| 79 | 1.150 | 0.839 | 1.083 | 0.754 | 1.209 | 0.786 | 1.100 | 0.842 | 1.319 | 0.824 | 1.296 | 0.776 |
| 80 | 1.181 | 0.790 | 1.087 | 0.714 | 1.283 | 0.776 | 1.393 | 0.770 | 1.387 | 0.724 | 1.224 | 0.743 |
| 81 | 1.168 | 0.801 | 1.107 | 0.751 | 1.241 | 0.773 | 1.103 | 0.744 | 1.322 | 0.818 | 1.284 | 0.843 |
| 82 | 1.107 | 0.729 | 1.062 | 0.743 | 1.121 | 0.726 | 1.100 | 0.809 | 1.344 | 0.853 | 1.116 | 0.817 |
| 83 | 1.114 | 0.796 | 1.056 | 0.737 | 1.247 | 0.775 | 1.349 | 0.520 | 1.398 | 0.847 | 1.226 | 0.835 |
| 84 | 1.125 | 0.812 | 1.058 | 0.743 | 1.276 | 0.819 | 1.358 | 0.520 | 1.327 | 0.779 | 1.284 | 0.738 |
| 85 | 1.107 | 0.795 | 1.057 | 0.718 | 1.230 | 0.802 | 1.351 | 0.525 | 1.349 | 0.757 | 1.209 | 0.819 |
| 86 | 1.110 | 0.747 | 1.034 | 0.751 | 1.232 | 0.790 | 1.308 | 0.786 | 1.349 | 0.765 | 1.256 | 0.756 |
| 87 | 1.147 | 0.775 | 1.082 | 0.752 | 1.283 | 0.749 | 1.102 | 0.858 | 1.323 | 0.795 | 1.294 | 0.721 |
| 88 | 1.103 | 0.813 | 1.052 | 0.740 | 1.268 | 0.728 | 1.399 | 0.840 | 1.311 | 0.755 | 1.202 | 0.836 |
| 89 | 1.121 | 0.727 | 1.009 | 0.774 | 1.232 | 0.741 | 1.111 | 0.834 | 1.300 | 0.824 | 1.252 | 0.853 |
| 90 | 1.104 | 0.742 | 1.031 | 0.753 | 1.201 | 0.842 | 1.365 | 0.731 | 1.313 | 0.756 | 1.274 | 0.735 |
| 91 | 1.120 | 0.722 | 1.065 | 0.753 | 1.236 | 0.731 | 1.304 | 0.849 | 1.335 | 0.857 | 1.211 | 0.733 |
| 92 | 1.103 | 0.774 | 1.088 | 0.756 | 1.233 | 0.847 | 1.381 | 0.833 | 1.317 | 0.784 | 1.116 | 0.845 |
| 93 | 1.120 | 0.735 | 1.105 | 0.758 | 1.269 | 0.811 | 1.303 | 0.848 | 1.316 | 0.855 | 1.213 | 0.725 |
| 94 | 1.192 | 0.784 | 1.040 | 0.785 | 1.121 | 0.848 | 1.324 | 0.751 | 1.302 | 0.786 | 1.253 | 0.721 |
| 95 | 1.104 | 0.727 | 1.048 | 0.751 | 1.211 | 0.517 | 1.304 | 0.781 | 1.322 | 0.730 | 1.213 | 0.829 |
| 96 | 1.189 | 0.844 | 1.010 | 0.798 | 1.208 | 0.517 | 1.111 | 0.809 | 1.381 | 0.729 | 1.277 | 0.717 |
| 97 | 1.129 | 0.750 | 1.004 | 0.760 | 1.104 | 0.515 | 1.109 | 0.767 | 1.303 | 0.755 | 1.216 | 0.811 |
| 98 | 1.117 | 0.833 | 1.000 | 0.758 | 1.244 | 0.520 | 1.328 | 0.757 | 1.356 | 0.743 | 1.243 | 0.806 |
| 99 | 1.127 | 0.771 | 1.101 | 0.754 | 1.260 | 0.841 | 1.352 | 0.781 | 1.391 | 0.757 | 1.282 | 0.749 |
| 100 | 1.103 | 0.754 | 1.088 | 0.755 | 1.261 | 0.798 | 1.363 | 0.795 | 1.386 | 0.848 | 1.235 | 0.768 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 59 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 80 **และ** PS 4 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มไม่เท่าเทียม**โดยใช้ข้อสอบร่วม

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 1 | 1.275 | 0.835 | 1.121 | 0.831 | 1.095 | 0.715 | 1.014 | 0.642 | 1.117 | 0.768 | 0.999 | 0.764 |
| 2 | 1.204 | 0.807 | 0.998 | 0.855 | 1.034 | 0.730 | 1.014 | 0.679 | 1.259 | 0.736 | 1.014 | 0.804 |
| 3 | 1.236 | 0.822 | 1.117 | 0.825 | 1.108 | 0.773 | 1.121 | 0.666 | 1.253 | 0.788 | 1.095 | 0.748 |
| 4 | 1.237 | 0.754 | 1.136 | 0.827 | 1.016 | 0.730 | 1.015 | 0.522 | 1.257 | 0.825 | 1.029 | 0.788 |
| 5 | 1.277 | 0.838 | 1.139 | 0.825 | 1.115 | 0.729 | 1.016 | 0.667 | 1.107 | 0.823 | 1.036 | 0.857 |
| 6 | 1.267 | 0.791 | 1.121 | 0.829 | 1.096 | 0.722 | 1.110 | 0.634 | 1.260 | 0.791 | 1.077 | 0.802 |
| 7 | 1.243 | 0.781 | 1.000 | 0.802 | 1.096 | 0.737 | 1.012 | 0.522 | 1.251 | 0.735 | 1.045 | 0.831 |
| 8 | 1.287 | 0.741 | 1.142 | 0.818 | 1.072 | 0.766 | 1.016 | 0.697 | 1.258 | 0.833 | 1.046 | 0.779 |
| 9 | 1.201 | 0.735 | 1.195 | 0.813 | 1.049 | 0.731 | 1.114 | 0.647 | 1.253 | 0.854 | 1.041 | 0.782 |
| 10 | 1.108 | 0.830 | 1.114 | 0.825 | 1.015 | 0.712 | 1.015 | 0.805 | 1.252 | 0.524 | 1.055 | 0.787 |
| 11 | 1.211 | 0.785 | 1.148 | 0.827 | 1.079 | 0.713 | 1.019 | 0.686 | 1.257 | 0.802 | 1.035 | 0.761 |
| 12 | 1.259 | 0.844 | 1.123 | 0.809 | 1.058 | 0.722 | 1.011 | 0.632 | 1.256 | 0.532 | 1.080 | 0.790 |
| 13 | 1.105 | 0.846 | 1.177 | 0.826 | 1.065 | 0.775 | 1.017 | 0.645 | 1.122 | 0.805 | 1.081 | 0.806 |
| 14 | 1.218 | 0.737 | 1.150 | 0.857 | 1.079 | 0.733 | 1.011 | 0.849 | 1.257 | 0.535 | 1.038 | 0.730 |
| 15 | 1.231 | 0.816 | 1.123 | 0.824 | 1.078 | 0.769 | 1.100 | 0.619 | 1.259 | 0.778 | 1.058 | 0.780 |
| 16 | 1.241 | 0.800 | 1.101 | 0.829 | 1.017 | 0.739 | 1.016 | 0.676 | 1.260 | 0.809 | 1.004 | 0.792 |
| 17 | 1.275 | 0.724 | 1.187 | 0.827 | 1.099 | 0.805 | 1.016 | 0.832 | 1.256 | 0.833 | 1.321 | 0.805 |
| 18 | 1.223 | 0.733 | 1.188 | 0.823 | 1.020 | 0.805 | 1.014 | 0.849 | 1.252 | 0.834 | 1.081 | 0.758 |
| 19 | 1.249 | 0.808 | 1.110 | 0.838 | 1.029 | 0.853 | 1.016 | 0.525 | 1.253 | 0.785 | 1.080 | 0.804 |
| 20 | 1.238 | 0.777 | 1.123 | 0.827 | 1.117 | 0.775 | 1.016 | 0.526 | 1.257 | 0.840 | 1.023 | 0.720 |
| 21 | 1.283 | 0.826 | 1.161 | 0.827 | 1.099 | 0.760 | 1.020 | 0.819 | 1.253 | 0.757 | 1.083 | 0.764 |
| 22 | 1.222 | 0.804 | 1.110 | 0.829 | 1.033 | 0.808 | 1.116 | 0.828 | 1.256 | 0.753 | 1.099 | 0.814 |
| 23 | 1.207 | 0.809 | 1.117 | 0.814 | 1.110 | 0.729 | 1.014 | 0.842 | 1.258 | 0.822 | 1.321 | 0.724 |
| 24 | 1.277 | 0.751 | 1.162 | 0.826 | 1.050 | 0.777 | 1.118 | 0.636 | 1.252 | 0.800 | 1.080 | 0.729 |
| 25 | 1.202 | 0.720 | 1.143 | 0.846 | 1.052 | 0.714 | 1.013 | 0.667 | 1.252 | 0.854 | 1.322 | 0.737 |
| 26 | 1.266 | 0.837 | 1.124 | 0.829 | 1.070 | 0.718 | 1.102 | 0.697 | 1.256 | 0.531 | 1.042 | 0.785 |
| 27 | 1.209 | 0.733 | 1.180 | 0.829 | 1.092 | 0.785 | 1.019 | 0.805 | 1.260 | 0.767 | 1.061 | 0.745 |
| 28 | 1.277 | 0.804 | 1.189 | 0.826 | 1.106 | 0.841 | 1.019 | 0.842 | 1.253 | 0.538 | 1.049 | 0.761 |
| 29 | 1.218 | 0.826 | 1.116 | 0.826 | 1.040 | 0.756 | 1.011 | 0.695 | 1.253 | 0.792 | 1.034 | 0.837 |
| 30 | 1.204 | 0.780 | 1.117 | 0.828 | 1.052 | 0.755 | 1.110 | 0.636 | 1.251 | 0.753 | 1.098 | 0.812 |
| 31 | 1.215 | 0.740 | 1.121 | 0.829 | 1.099 | 0.721 | 1.011 | 0.809 | 1.251 | 0.535 | 1.322 | 0.741 |
| 32 | 1.208 | 0.723 | 1.117 | 0.823 | 1.011 | 0.791 | 1.017 | 0.819 | 1.260 | 0.534 | 1.078 | 0.725 |
| 33 | 1.239 | 0.843 | 1.118 | 0.817 | 1.076 | 0.768 | 1.019 | 0.681 | 1.257 | 0.794 | 1.086 | 0.830 |
| 34 | 1.232 | 0.852 | 1.141 | 0.826 | 1.098 | 0.729 | 1.012 | 0.664 | 1.106 | 0.814 | 1.056 | 0.767 |
| 35 | 1.270 | 0.771 | 1.113 | 0.824 | 1.027 | 0.763 | 1.013 | 0.800 | 1.253 | 0.816 | 1.063 | 0.815 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 59 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 80 **และ** PS 4 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มไม่เท่าเทียม**โดยใช้ข้อสอบร่วม (ต่อ)

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 36 | 1.284 | 0.806 | 1.147 | 0.822 | 1.058 | 0.728 | 1.013 | 0.807 | 1.255 | 0.838 | 1.093 | 0.725 |
| 37 | 1.119 | 0.838 | 1.108 | 0.827 | 1.102 | 0.717 | 1.010 | 0.831 | 1.256 | 0.789 | 1.066 | 0.786 |
| 38 | 1.222 | 0.755 | 1.184 | 0.829 | 1.050 | 0.729 | 1.012 | 0.656 | 1.257 | 0.533 | 1.073 | 0.781 |
| 39 | 1.266 | 0.849 | 1.107 | 0.826 | 1.111 | 0.799 | 1.019 | 0.521 | 1.259 | 0.857 | 1.045 | 0.726 |
| 40 | 1.106 | 0.805 | 1.186 | 0.829 | 1.014 | 0.798 | 1.016 | 0.696 | 1.259 | 0.858 | 1.079 | 0.726 |
| 41 | 1.252 | 0.761 | 1.117 | 0.827 | 1.011 | 0.858 | 1.107 | 0.660 | 1.251 | 0.772 | 1.089 | 0.739 |
| 42 | 1.208 | 0.829 | 1.188 | 0.822 | 1.017 | 0.790 | 1.018 | 0.631 | 1.259 | 0.836 | 1.096 | 0.795 |
| 43 | 1.233 | 0.841 | 1.154 | 0.819 | 1.034 | 0.731 | 1.015 | 0.822 | 1.251 | 0.818 | 1.321 | 0.812 |
| 44 | 1.289 | 0.803 | 1.151 | 0.827 | 1.061 | 0.720 | 1.012 | 0.645 | 1.254 | 0.853 | 1.021 | 0.856 |
| 45 | 1.113 | 0.733 | 1.183 | 0.825 | 1.081 | 0.762 | 1.020 | 0.678 | 1.108 | 0.758 | 1.019 | 0.831 |
| 46 | 1.280 | 0.759 | 1.134 | 0.823 | 1.079 | 0.760 | 1.015 | 0.831 | 1.258 | 0.856 | 1.070 | 0.719 |
| 47 | 1.282 | 0.829 | 1.169 | 0.826 | 1.057 | 0.744 | 1.012 | 0.827 | 1.255 | 0.524 | 1.053 | 0.761 |
| 48 | 1.245 | 0.820 | 1.100 | 0.835 | 1.014 | 0.777 | 1.020 | 0.524 | 1.256 | 0.743 | 0.999 | 0.853 |
| 49 | 1.283 | 0.781 | 1.157 | 0.827 | 1.062 | 0.722 | 1.112 | 0.654 | 1.259 | 0.778 | 1.059 | 0.757 |
| 50 | 1.258 | 0.795 | 1.127 | 0.826 | 1.030 | 0.727 | 1.013 | 0.664 | 1.255 | 0.737 | 1.048 | 0.790 |
| 51 | 1.216 | 0.795 | 1.122 | 0.845 | 1.064 | 0.853 | 1.010 | 0.649 | 1.257 | 0.845 | 1.320 | 0.754 |
| 52 | 1.250 | 0.735 | 1.101 | 0.828 | 1.074 | 0.714 | 1.017 | 0.809 | 1.108 | 0.774 | 1.068 | 0.732 |
| 53 | 1.258 | 0.763 | 1.180 | 0.829 | 1.042 | 0.733 | 1.012 | 0.831 | 1.252 | 0.763 | 0.998 | 0.849 |
| 54 | 1.269 | 0.832 | 1.162 | 0.822 | 1.012 | 0.780 | 1.108 | 0.632 | 1.256 | 0.734 | 1.082 | 0.781 |
| 55 | 1.281 | 0.777 | 1.109 | 0.810 | 1.007 | 0.786 | 1.017 | 0.521 | 1.253 | 0.743 | 1.320 | 0.748 |
| 56 | 1.287 | 0.841 | 1.144 | 0.830 | 1.113 | 0.848 | 1.010 | 0.523 | 1.256 | 0.526 | 1.034 | 0.745 |
| 57 | 1.233 | 0.757 | 1.103 | 0.827 | 1.104 | 0.713 | 1.018 | 0.662 | 1.258 | 0.752 | 1.321 | 0.723 |
| 58 | 1.239 | 0.823 | 1.176 | 0.825 | 1.069 | 0.752 | 1.118 | 0.695 | 1.121 | 0.772 | 1.027 | 0.721 |
| 59 | 1.213 | 0.775 | 1.104 | 0.817 | 1.009 | 0.734 | 1.017 | 0.614 | 1.250 | 0.791 | 1.003 | 0.810 |
| 60 | 1.248 | 0.825 | 1.103 | 0.830 | 1.111 | 0.726 | 1.017 | 0.817 | 1.252 | 0.852 | 1.026 | 0.802 |
| 61 | 1.228 | 0.766 | 1.169 | 0.823 | 1.002 | 0.744 | 1.018 | 0.637 | 1.257 | 0.854 | 1.065 | 0.734 |
| 62 | 1.278 | 0.807 | 1.124 | 0.822 | 1.104 | 0.843 | 1.015 | 0.613 | 1.251 | 0.529 | 1.321 | 0.767 |
| 63 | 1.111 | 0.829 | 0.998 | 0.844 | 1.001 | 0.840 | 1.016 | 0.630 | 1.259 | 0.811 | 1.322 | 0.771 |
| 64 | 1.219 | 0.825 | 1.111 | 0.827 | 1.004 | 0.760 | 1.019 | 0.664 | 1.259 | 0.814 | 1.089 | 0.801 |
| 65 | 1.105 | 0.745 | 1.105 | 0.829 | 1.001 | 0.783 | 1.011 | 0.802 | 1.259 | 0.800 | 1.068 | 0.829 |
| 66 | 1.235 | 0.815 | 1.119 | 0.824 | 1.109 | 0.777 | 1.014 | 0.839 | 1.257 | 0.844 | 1.066 | 0.823 |
| 67 | 1.284 | 0.855 | 1.137 | 0.846 | 1.011 | 0.779 | 1.108 | 0.643 | 1.260 | 0.731 | 1.033 | 0.782 |
| 68 | 1.293 | 0.781 | 1.157 | 0.846 | 1.057 | 0.736 | 1.014 | 0.662 | 1.250 | 0.785 | 1.057 | 0.850 |
| 69 | 1.254 | 0.817 | 1.111 | 0.841 | 1.122 | 0.779 | 1.014 | 0.635 | 1.251 | 0.848 | 1.050 | 0.768 |
| 70 | 1.277 | 0.820 | 1.197 | 0.817 | 1.062 | 0.758 | 1.016 | 0.615 | 1.259 | 0.854 | 1.322 | 0.782 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 59 ค่าสัมประสิทธิ์การปรับเทียบของแบบสอบรูปแบบผสมระหว่าง DS 80 **และ** PS 4 ด้วยการเก็บรวบรวมข้อมูลแบบ**กลุ่มไม่เท่าเทียม**โดยใช้ข้อสอบร่วม (ต่อ)

| สัดส่วนของข้อสอบร่วม | 15/2 | | 20/2 | | 25/2 | | 15/1 | | 10/2 | | 5/3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| สัมประสิทธิ์การปรับเทียบ | A | B | A | B | A | B | A | B | A | B | A | B |
| 71 | 1.269 | 0.796 | 1.119 | 0.830 | 1.113 | 0.722 | 1.017 | 0.817 | 1.255 | 0.857 | 1.062 | 0.738 |
| 72 | 1.292 | 0.832 | 1.123 | 0.825 | 1.027 | 0.737 | 1.106 | 0.636 | 1.113 | 0.849 | 1.095 | 0.794 |
| 73 | 1.108 | 0.787 | 1.140 | 0.851 | 1.040 | 0.852 | 1.112 | 0.843 | 1.260 | 0.830 | 1.093 | 0.807 |
| 74 | 1.247 | 0.721 | 1.117 | 0.831 | 1.100 | 0.763 | 1.118 | 0.676 | 1.259 | 0.528 | 1.051 | 0.797 |
| 75 | 1.240 | 0.819 | 1.123 | 0.829 | 1.028 | 0.752 | 1.015 | 0.828 | 1.254 | 0.855 | 1.096 | 0.769 |
| 76 | 1.266 | 0.741 | 1.120 | 0.840 | 1.101 | 0.803 | 1.018 | 0.640 | 1.256 | 0.808 | 1.004 | 0.801 |
| 77 | 1.259 | 0.722 | 1.174 | 0.822 | 1.018 | 0.738 | 1.120 | 0.655 | 1.252 | 0.835 | 1.321 | 0.815 |
| 78 | 1.111 | 0.801 | 1.193 | 0.831 | 1.099 | 0.746 | 0.999 | 0.524 | 1.258 | 0.794 | 1.048 | 0.723 |
| 79 | 1.209 | 0.817 | 1.146 | 0.837 | 1.012 | 0.722 | 1.020 | 0.803 | 1.254 | 0.826 | 1.322 | 0.738 |
| 80 | 1.280 | 0.804 | 1.188 | 0.824 | 1.114 | 0.741 | 1.014 | 0.814 | 1.255 | 0.772 | 1.073 | 0.817 |
| 81 | 1.231 | 0.834 | 1.193 | 0.830 | 1.119 | 0.795 | 1.019 | 0.524 | 1.256 | 0.768 | 1.056 | 0.743 |
| 82 | 1.254 | 0.760 | 1.149 | 0.823 | 1.035 | 0.728 | 1.012 | 0.618 | 1.254 | 0.831 | 1.032 | 0.754 |
| 83 | 1.205 | 0.818 | 1.140 | 0.826 | 1.018 | 0.719 | 1.010 | 0.526 | 1.111 | 0.805 | 1.320 | 0.799 |
| 84 | 1.227 | 0.763 | 1.143 | 0.801 | 1.054 | 0.794 | 1.018 | 0.630 | 1.252 | 0.752 | 1.078 | 0.789 |
| 85 | 1.245 | 0.817 | 1.166 | 0.822 | 1.074 | 0.809 | 1.116 | 0.686 | 1.253 | 0.816 | 1.048 | 0.834 |
| 86 | 1.110 | 0.835 | 1.120 | 0.829 | 1.002 | 0.785 | 1.019 | 0.848 | 1.253 | 0.752 | 1.068 | 0.789 |
| 87 | 1.281 | 0.854 | 1.114 | 0.826 | 1.025 | 0.726 | 1.016 | 0.529 | 1.258 | 0.736 | 1.322 | 0.827 |
| 88 | 1.297 | 0.788 | 1.115 | 0.818 | 1.093 | 0.856 | 1.016 | 0.699 | 1.257 | 0.746 | 1.067 | 0.775 |
| 89 | 1.296 | 0.823 | 1.100 | 0.828 | 1.060 | 0.845 | 1.011 | 0.819 | 1.257 | 0.784 | 1.012 | 0.851 |
| 90 | 1.275 | 0.758 | 1.195 | 0.830 | 1.010 | 0.774 | 1.013 | 0.688 | 1.258 | 0.752 | 1.038 | 0.854 |
| 91 | 1.107 | 0.713 | 1.108 | 0.828 | 1.007 | 0.744 | 1.014 | 0.823 | 1.250 | 0.775 | 1.024 | 0.756 |
| 92 | 1.232 | 0.753 | 1.109 | 0.828 | 1.031 | 0.753 | 1.011 | 0.814 | 1.251 | 0.840 | 1.048 | 0.798 |
| 93 | 1.232 | 0.728 | 1.102 | 0.829 | 1.004 | 0.778 | 1.014 | 0.800 | 1.121 | 0.779 | 1.322 | 0.802 |
| 94 | 1.275 | 0.801 | 1.124 | 0.823 | 1.067 | 0.729 | 1.017 | 0.805 | 1.109 | 0.841 | 1.094 | 0.829 |
| 95 | 1.292 | 0.759 | 1.104 | 0.814 | 1.076 | 0.765 | 1.010 | 0.685 | 1.115 | 0.754 | 1.002 | 0.809 |
| 96 | 1.232 | 0.858 | 1.102 | 0.829 | 1.009 | 0.758 | 1.018 | 0.681 | 1.112 | 0.782 | 1.085 | 0.746 |
| 97 | 1.273 | 0.746 | 1.109 | 0.823 | 1.047 | 0.747 | 1.015 | 0.691 | 1.254 | 0.772 | 1.322 | 0.818 |
| 98 | 1.110 | 0.810 | 1.179 | 0.823 | 1.039 | 0.723 | 1.000 | 0.654 | 1.251 | 0.531 | 1.017 | 0.796 |
| 99 | 1.238 | 0.733 | 1.195 | 0.803 | 1.019 | 0.845 | 1.010 | 0.664 | 1.123 | 0.850 | 1.086 | 0.745 |
| 100 | 1.269 | 0.828 | 1.170 | 0.827 | 1.018 | 0.725 | 0.998 | 0.813 | 1.113 | 0.741 | 1.052 | 0.731 |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

# ประวัติผู้เขียนวิทยานิพนธ์

| | |
|---|---|
| ชื่อ – สกุล | นางสาวอัญชลี ศรีกลชาญ |
| วัน เดือน ปี เกิด | 11 เดือนมิถุนายน พุทธศักราช 2521 |
| สถานที่เกิด | โรงพยาบาลศูนย์แม่และเด็ก อำเภอเมือง จังหวัดขอนแก่น |
| สถานที่อยู่ปัจจุบัน | 169 วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา ถ.ลงหาดบางแสน ต.แสนสุข อ.เมือง จ.ชลบุรี 20131 |
| ประวัติการศึกษา | |
| พ.ศ. 2542 | การศึกษาบัณฑิต สาขาวิชา วิทยาศาสตร์ – ฟิสิกส์ มหาวิทยาลัยบูรพา |
| พ.ศ. 2546 | วิทยาศาสตรมหาบัณฑิต สาขาวิชา เทคโนโลยีการวัดทางการศึกษา |
| พ.ศ. 2548 | เข้าศึกษาใน สาขาวิชาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย |
| ปัจจุบัน | อาจารย์ประจำวิทยาลัยวิทยาการวิจัย และวิทยาการปัญญา มหาวิทยาลัยบูรพา |

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย