การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนโดยใช้วิธีการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบร่วมกับวิธีการปรับเทียบเชิงเส้นตรง

นางสาวพัชรี จันทร์เพ็ง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรมหาบัณฑิต
สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2547

ISBN 974-53-1900-7

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

A STUDY OF STUDENTS' MATHEMATICAL ABILITY GROWTH USING VERTICAL EQUATING COMBINING IRT AND LINEAR EQUATING METHODS

Miss Putcharee Junpeng

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Education in Educational Measurement and Evaluation
Department of Educational Research and Psychology
Faculty of Education
Chulalongkorn University
Academic Year 2004

ISBN 974-53-1900-7

หัวข้อวิทยานิพนธ์	การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนโดยใช้วิธีการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบร่วมกับวิธีการปรับเทียบเชิงเส้นตรง
โดย	นางสาวพัชรี จันทร์เพ็ง
สาขาวิชา	การวัดและประเมินผลการศึกษา
อาจารย์ที่ปรึกษา	อาจารย์ ดร.ณัฏฐภรณ์ หลาวทอง

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....คณบดีคณะครุศาสตร์
(รองศาสตราจารย์ ดร.พฤทธิ์ ศิริบรรณพิทักษ์)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(รองศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี)

.....อาจารย์ที่ปรึกษา
(อาจารย์ ดร. ณัฏฐภรณ์ หลาวทอง)

.....กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.อวยพร เรืองตระกูล)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

พัชรี จันทร์เพ็ง : การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนโดยใช้วิธีการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบร่วมกับวิธีการปรับเทียบเชิงเส้นตรง. (A STUDY OF STUDENTS' MATHEMATICAL ABILITY GROWTH USING VERTICAL EQUATING COMBINING IRT AND LINEAR EQUATING METHODS) อ.ที่ปรึกษา : ดร. ณัฏฐภรณ์ หลาวทอง, 194 หน้า. ISBN 974-53-1900-7

การวิจัยครั้งนี้มีจุดมุ่งหมาย เพื่อศึกษาประสิทธิภาพของการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบแบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรงที่ใช้แบบสอบร่วมภายใน และศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ในแต่ละระดับชั้นโดยภาพรวมและจำแนกตามกลุ่มความสามารถ คือ กลุ่มต่ำ กลุ่มปานกลาง และกลุ่มสูง

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ของโรงเรียนในเขตพื้นที่การศึกษาสกลนคร เขต 1 จำนวน 1,941 คน โดยสุ่มกลุ่มตัวอย่างด้วยวิธีการสุ่มแบบแบ่งชั้น กลุ่มตัวอย่างแบ่งเป็น 2 กลุ่ม คือ กลุ่มปรับเทียบคะแนน เป็นนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 จำนวน 543, 534 และ 504 คน ตามลำดับ และกลุ่มสอบทานผล เป็นนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 จำนวน 113, 120 และ 127 คน ตามลำดับ เครื่องมือที่ใช้ในการวิจัยครั้งนี้เป็นแบบวัดความสามารถทางคณิตศาสตร์ เรื่อง "จำนวนและการดำเนินการ" สำหรับระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 รวม 3 ฉบับ ฉบับละ 30 ข้อ ที่มีค่าความเที่ยงของแบบสอบเท่ากับ 0.803, 0.912 และ 0.857 ตามลำดับ โดยมีข้อสอบร่วมระหว่างระดับชั้น ม.1 และ ม.2 จำนวน 6 ข้อ ข้อสอบร่วมระหว่างระดับชั้น ม.2 และ ม.3 จำนวน 6 ข้อ โดยวิเคราะห์ข้อมูลเพื่อประมาณค่าความสามารถของนักเรียน ด้วยโปรแกรม BILOG 3.04 คำนวณหาค่าคงที่เพื่อสร้างสมการปรับเทียบคะแนนด้วยโปรแกรม EQUATE 2.0 ปรับเทียบคะแนนเชิงเส้น ด้วยวิธี "กำลังสองน้อยสุด" ด้วยโปรแกรม SPSS for Window Version 11.0

สรุปผลการวิจัยได้ดังนี้

1) ความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนตามรูปแบบที่ศึกษาสำหรับผู้สอบที่มีความสามารถแตกต่างกัน 3 ระดับ พบว่ามีค่าเท่ากับ 0.063 ซึ่งเป็นความคลาดเคลื่อนมาตรฐานที่ต่ำ และความเพียงพอของการปรับเทียบคะแนน พบว่า ค่าดัชนีความแตกต่างของการปรับเทียบคะแนนอยู่ในระดับน่าพอใจ (C=0.225)

2) ผลการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียน ในกลุ่มรวม พบว่า ค่าเฉลี่ยของคะแนนความสามารถทั้ง 3 ระดับชั้น แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยความสามารถทางคณิตศาสตร์ของนักเรียนเพิ่มขึ้นตามระดับชั้นที่สูงขึ้น นั่นคือ ระดับชั้น ม.1 เมื่อขึ้น ม.2 มีพัฒนาการสูงขึ้น 0.167 คะแนน นักเรียนระดับชั้น ม.2 เมื่อขึ้น ม.3 มีพัฒนาการสูงขึ้น 0.174 คะแนน และนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.3 มีพัฒนาการมากที่สุด เท่ากับ 0.312 คะแนน

ส่วนผลการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนจำแนกตามกลุ่มความสามารถ คือ กลุ่มต่ำ กลุ่มปานกลาง และกลุ่มสูง พบว่า นักเรียนในแต่ละกลุ่มความสามารถ มีคะแนนพัฒนาการเพิ่มขึ้นตามระดับชั้นที่สูงขึ้น โดยพิจารณาจากอัตราพัฒนาการ และนักเรียนในกลุ่มความสามารถสูงมีอัตราพัฒนาการมากที่สุด รองลงมา คือ กลุ่มปานกลาง และกลุ่มต่ำ ตามลำดับ

ภาควิชา วิจัยและจิตวิทยาการศึกษา ลายมือชื่อนิสิต.....

สาขาวิชา การวัดและประเมินผลการศึกษา ลายมือชื่ออาจารย์ที่ปรึกษา.....

ปีการศึกษา 2547 ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

4683721127 : MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION
KEY WORD : VERTICAL EQUATING / IRT EQUATING / LINEAR EQUATING / GROWTH OF MATIMATICAL ABILITY
PUCHAREE JUNPENG : A STUDY OF STUDENTS' MATHEMATICAL ABILITY GROWTH USING VERTICAL EQUATING COMBINING IRT AND LINEAR EQUATING METHODS.
THESIS ADVISOR : NATTAPORN LAWTHONG, Ph.D., 194 pp. ISBN 974-53-1900-7

The purposes of this research were to investigate the efficiency of vertical equating combining IRT by three-parameter logistic model and linear equating methods with internal anchor tests, and to study of Matayomsuksa 1-3 students' mathematical ability growth . The students in each grade level was used to equate the students' mathematical ability scores across grade levels so that they were on the same scale scores. The students in each level were divied into three sub-groups accordingly to their abilities; high, fair, and low ability groups.

The sample was 1,941 students obtained by stratified random sampling from Matayomsuksa 1-3 students of Office of Sakonnakhon Educational Service Area Zone I. The students in each grade level were divided into two groups; the equated and validated groups. There were altogether 543, 534, and 504 students respectively for equated groups, while there were 113, 120, and 127 students for the validated groups. They were used for data collecting with the Mathematical ability test, the numerical and operation, for the Matayomsuksa 1-3. Each test had 30 items. The reliabilities of test were 0.803, 0.912, as well as 0.857. There were 6 internal anchor items between the M.1 test and the M.2 test, 6 internal anchor items between the M.2 test and the M.3 test. The research data were analyzed by employing BILOG 3.04 for estimating students' abilities parameters, employing EQUATE 2.0 for expressing linear interpolation equations, employing SPSS for Windows version 11.0 for expressing linear regression equation by mean of simple least square method.

The results were as follows :

1. The standard error of the equating study model from students of three different ability levels was 0.063 that was low. The discrepancy index of equating in range of satisfy was the acceptable criterion. (C=0.225)

2. The study of the growth rate on students' Mathematical abilities found the significance at .01 level of mean scores of three different levels. It could be concluded that the growth rate on Mathematical ability from M.1 students' to M.2 students' was 0.167, and M.2 student's to M.3 students' was 0.174. Moreover, M.1 students' to M.3 students' was 0.312, that was the highest.

In the growth rates study on students' Mathematical abilities were divided into three groups of abilities namely: low, fair, and high groups. considering the growth rates were found that all groups had gain scores accordingly class levels. The student in the high ability group had the highest growth rate that followed by the fair ability group and the low ability group respectively.

Field of study …… Educational Measurement and Evaluation…….. Student's……..……..……..
Department ……Educational Research and Psychology……..Advisor's.....
Academic year …..2004…….....……Co-advisor's signature……...

กิตติกรรมประกาศ

การทำวิทยานิพนธ์ฉบับนี้ ผู้วิจัยได้รับความรู้และคำแนะนำต่างๆ ด้วยความกรุณาอย่างยิ่งจากอาจารย์ที่ปรึกษาวิทยานิพนธ์ คือ อาจารย์ ดร.ณัฏฐภรณ์ หลาวทอง และได้รับคำแนะนำในการดำเนินการวิจัยเพิ่มเติมจาก รองศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี ซึ่งอาจารย์ทั้งสองท่านได้ให้ความรู้ในหลักวิชามากมายอันทรงคุณค่ายิ่ง ทั้งยังได้ดูแลเอาใจใส่คอยติดตามความก้าวหน้าของผู้วิจัยตลอดมา ผู้วิจัยรู้สึกซาบซึ้งและขอกราบขอบพระคุณเป็นอย่างยิ่งด้วยความเคารพมา ณ ที่นี้ด้วย

ขอกราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.อวยพร เรืองตระกูล ซึ่งเป็นกรรมการสอบวิทยานิพนธ์ที่ได้กรุณาสละเวลาอันมีค่ามาร่วมสอบวิทยานิพนธ์ ตลอดจนให้ข้อเสนอแนะที่เป็นประโยชน์ต่อการทำวิทยานิพนธ์ฉบับนี้ และขอกราบขอบพระคุณคณาจารย์ภาควิชาวิจัยและจิตวิทยาการศึกษา จุฬาลงกรณ์มหาวิทยาลัย ตลอดจนคณาจารย์ภาควิชาประเมินผลและวิจัยการศึกษา มหาวิทยาลัยขอนแก่น ที่คอยให้การสนับสนุนและเป็นกำลังใจด้วยดีตลอดมา

ขอขอบพระคุณ ดร.อดิศร ศรีบุญวงศ์ และ ดร.ภัทราพร เกษสังข์ ที่ให้คำแนะนำในการทำวิทยานิพนธ์จนสำเร็จลุล่วงไปด้วยดี และขอขอบคุณเป็นพิเศษสำหรับ คุณอิทธิฤทธิ์ พงษ์ปิยะรัตน์ คุณสุวิมล กฤชคฤหาสน์ คุณตติยา เชยชุ่ม คุณชยุตม์ ภิรมย์สมบัติ คุณสิทธิพันธ์ ยศยอดยิ่ง และคุณพงษ์ลิขิต เพชรผล ที่กรุณาให้ความช่วยเหลือและข้อเสนอแนะที่เป็นประโยชน์แก่ผู้วิจัยมาโดยตลอด รวมทั้งพี่ๆและเพื่อนๆ ภาควิชาวิจัยและจิตวิทยาการศึกษาที่ให้กำลังใจด้วยดีตลอดมา

ขอขอบพระคุณบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย และ ภาควิชาประเมินผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น ที่ให้ทุนอุดหนุนในการทำวิทยานิพนธ์และงานวิจัยนี้จะไม่ประสบความสำเร็จ ถ้าไม่ได้คำแนะนำจากผู้เชี่ยวชาญในการสร้างเครื่องมือ ตลอดจนความร่วมมือจากโรงเรียนที่เป็นกลุ่มทดลองเครื่องมือและกลุ่มตัวอย่างทั้ง 10 โรงเรียน ซึ่งอำนวยความสะดวกในการเก็บข้อมูล ผู้วิจัยขอขอบพระคุณเป็นอย่างยิ่งและจะไม่ลืมพระคุณที่ให้ความอนุเคราะห์ในครั้งนี้

คุณค่าและประโยชน์ที่อาจมีต่อผู้อื่นของวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแก่ คุณพ่อธีระพล คุณแม่วิลัย คุณปู่ และคุณย่า ตลอดจนญาติพี่น้องทุกคน ที่ได้ให้ความอบอุ่นห่วงใยและสนับสนุน ในทุกๆด้าน จนผู้วิจัยมีความสำเร็จมาถึงจุดนี้

สารบัญ

บทที่		หน้า
บทคัดย่อภาษาไทย		ง
บทคัดย่อภาษาอังกฤษ		จ
กิตติกรรมประกาศ		ฉ
สารบัญ		ช
สารบัญตาราง		ญ
สารบัญแผนภาพ		ฐ
บทที่		
1	บทนำ……..	1
	ความเป็นมาและความสำคัญของปัญหา……	1
	คำถามการวิจัย……	5
	วัตถุประสงค์ของการวิจัย…….	6
	สมมติฐานการวิจัย……...	7
	ขอบเขตของการวิจัย……	8
	คำจำกัดความเชิงปฏิบัติการ……	9
	กรอบแนวคิดของการวิจัย……	13
	ประโยชน์ที่คาดว่าจะได้รับ……...	14
2	เอกสารและงานวิจัยที่เกี่ยวข้อง……....	15
	ตอนที่ 1 การปรับเทียบคะแนนระหว่างแบบสอบ…….……..	15
	ตอนที่ 2 หลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 ในกลุ่มสาระการเรียนรู้ คณิตศาสตร์ ช่วงชั้นที่ 3……	53
	ตอนที่ 3 งานวิจัยที่เกี่ยวข้องกับเทคนิคการปรับเทียบคะแนนและการศึกษาพัฒนาการ……..	57
3	วิธีดำเนินการวิจัย…….	79
	ประชากร……..	79
	กลุ่มตัวอย่าง……..….	80
	แบบแผนการเก็บข้อมูลในการปรับเทียบของการวิจัย…….....	83
	เครื่องมือที่ใช้ในการวิจัย…….....	84

หน้า

วิธีดำเนินการสร้างแบบสอบ..... 86

การเก็บรวบรวมข้อมูล..... 88

การวิเคราะห์ข้อมูล..... 89

4 ผลการวิเคราะห์ข้อมูล..... 97

ตอนที่ 1 ผลการวิเคราะห์ของแบบสอบ..... 100

1.1 ผลการวิเคราะห์ค่าสถิติพื้นฐานของแบบสอบแต่ละฉบับ..... 100

1.2 ค่าพารามิเตอร์ของข้อสอบร่วม..... 101

ตอนที่ 2 ผลการสร้างการดำเนินการปรับเทียบคะแนนแบบสอบ..... 104

2.1 การปรับเทียบระดับความสามารถของผู้สอบ โดยใช้ทฤษฎีตอบสนองข้อสอบแบบโลจิสติก 3 พารามิเตอร์..... 104

2.2 การปรับเทียบระดับความสามารถของผู้สอบโดยใช้การปรับเทียบตามทฤษฎีตอบสนองข้อสอบ (IRT) แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง.. 109

2.3 การปรับเทียบระดับคะแนน..... 114

ตอนที่ 3 ผลการศึกษาประสิทธิภาพของการปรับเทียบคะแนน..... 117

3.1 ค่าสถิติพื้นฐานของคะแนนสอบจากกลุ่มสอบทานผล..... 117

3.2 ผลการวิเคราะห์ความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน..... 119

3.3 ผลการวิเคราะห์ความเพียงพอของความคลาดเคลื่อนของการปรับเทียบคะแนน..... 122

ตอนที่ 4 ผลการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียน..... 130

4.1 การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนในกลุ่มรวม..... 131

4.2 การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนจำแนกตามกลุ่มความสามารถ..... 135

5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ..... 147

สรุปผลการวิจัย..... 149

อภิปรายผล..... 152

ข้อเสนอแนะ..... 159

หน้า

รายการอ้างอิง..... 162
ภาคผนวก ก รายชื่อผู้เชี่ยวชาญ..... 169
ภาคผนวก ข แผนผังข้อสอบ..... 171
ภาคผนวก ค ค่าความยาก และ ค่าอำนาจจำแนก ของการวิเคราะห์ข้อสอบตามทฤษฎี CTT (ค่า p และ r) และตามทฤษฎี IRT (ค่า b และ a) ของแบบสอบ เรื่อง "จำนวนและการดำเนินการ" ของระดับชั้น ม.1, 2 และ 3..... 177
ภาคผนวก ง ตารางปรับเทียบคะแนน..... 184
ภาคผนวก จ การวิเคราะห์ด้วยโปรแกรม BILOG 3.04..... 189
ภาคผนวก ฉ การวิเคราะห์ด้วยโปรแกรม EQUATE version 2.0..... 192
ประวัติผู้เขียนวิทยานิพนธ์..... 195

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

ตารางที่ หน้า

1.1 เกณฑ์การจำแนกนักเรียนตามกลุ่มความสามารถ…….. 10

1.2 รูปแบบของแบบสอบร่วมภายใน…….. 11

2.1 รูปแบบการเก็บรวบรวมข้อมูลสำหรับการปรับเทียบคะแนน…… 27

2.2 มาตรฐานการเรียนรู้ช่วงชั้น (ช่วงชั้นที่ 3) สาระที่ 1 : จำนวนและการดำเนินการ…….. 56

2.3 แสดงการศึกษาพัฒนาการของนักเรียนในวงการศึกษาตั้งแต่อดีตถึงปัจจุบัน…… 74

3.1 จำนวนประชากรจำแนกตามขนาดของโรงเรียนและระดับชั้น……. 79

3.2 จำนวนนักเรียนและจำนวนโรงเรียนที่ใช้เป็นกลุ่มตัวอย่างในการวิจัย จำแนกตามขนาดของโรงเรียน และระดับชั้น…….. 81

3.3 จำนวนกลุ่มตัวอย่างที่ใช้ในการวิจัยจำแนกตาม ขนาดของโรงเรียนและกลุ่มตัวอย่างสำหรับการวิจัย เขตพื้นที่การศึกษาที่ 1 จังหวัดสกลนคร……... 82

3.4 แบบแผนการปรับเทียบคะแนน…… 83

3.5 แบบแผนการเก็บรวบรวมข้อมูลในการปรับเทียบคะแนน…… 84

3.6 แบบแผนการเก็บรวบรวมข้อมูลในการสอบทานผล……. 84

3.7 รูปแบบที่ได้จากการวิเคราะห์การปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ……. 90

3.8 เกณฑ์การแบ่งนักเรียนแต่ละระดับชั้นตามความสามารถ (θ) โดยใช้คะแนนที่ได้จากการทำแบบสอบภายหลังการปรับเทียบคะแนนให้อยู่ในสเกลเดียวกันตามระดับชั้นของตน…… 95

4.1 ค่าสถิติพื้นฐานของแบบสอบแต่ละฉบับ…….. 100

4.2 ค่าพารามิเตอร์ของข้อสอบร่วมและค่าพารามิเตอร์ของข้อสอบร่วมเฉลี่ยระหว่างชั้นม.1 กับ ม.2…….. 102

4.3 ค่าพารามิเตอร์ของข้อสอบร่วมและค่าพารามิเตอร์ของข้อสอบร่วมเฉลี่ยระหว่างชั้น ม.2 กับ ม.3……. 103

4.4 ค่าความสามารถ (θ) และคะแนนสอบ (X) ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1-3 กลุ่มปรับเทียบคะแนน…….. 105

4.5 การปรับเทียบคะแนนความสามารถและคะแนนจริงของนักเรียนกลุ่มปรับเทียบคะแนนระดับชั้น ม. 2 ให้อยู่ในสเกลเดียวกับระดับชั้น ม.1 และ ม.3 ตามทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์…… 107

ตารางที่ หน้า

4.6 ค่าคงที่ A, K และ สมการเชิงเส้นตรงในการแปลงระดับความสามารถ (θ) ของผู้สอบที่ได้จากแบบสอบต่างระดับชั้นให้อยู่ในสเกลเดียวกัน ตามรูปแบบการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ร่วมกับการปรับเทียบเชิงเส้นตรง…….. 110

4.7 การปรับเทียบคะแนนความสามารถและคะแนนจริงของนักเรียนกลุ่มปรับเทียบคะแนนให้อยู่ในสเกลเดียวกันกันตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง…….. 112

4.8 การปรับเทียบสเกลของคะแนนสอบวัดความสามารถทางคณิตศาสตร์ระหว่างระดับชั้น ม.1, ม.2 และ ม.3…….. 115

4.9 ค่าสถิติพื้นฐานของคะแนนสอบในระดับชั้น ม.1, ม.2 และ ม.3 จากกลุ่มสอบทานผล…. 117

4.10 คะแนนความคลาดเคลื่อนเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานของคะแนนความคลาดเคลื่อน (SD_E) และคะแนนความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน (SEE) จากคะแนนแบบสอบของนักเรียนในแต่ละระดับชั้นและในระดับคะแนนทั้งหมดของทุกระดับชั้น…… 120

4.11 การประเมินค่าดัชนีความแตกต่าง (C) เมื่อแปลงคะแนนแบบสอบระดับชั้น ม.1 ไปยังคะแนนแบบสอบระดับ ม.2 จากกลุ่มสอบทานผลในระดับ ม.1……. 123

4.12 การประเมินค่าดัชนีความแตกต่าง (C) เมื่อแปลงคะแนนแบบสอบระดับชั้น ม.1 ไปยังคะแนนแบบสอบระดับ ม.3 จากกลุ่มสอบทานผลในระดับ ม.1…….. 124

4.13 การประเมินค่าดัชนีความแตกต่าง (C) เมื่อแปลงคะแนนแบบสอบระดับชั้น ม.2 ไปยังคะแนนแบบสอบระดับ ม.1 จากกลุ่มสอบทานผลในระดับ ม.2…….. 125

4.14 การประเมินค่าดัชนีความแตกต่าง (C) เมื่อแปลงคะแนนแบบสอบระดับชั้น ม.2 ไปยังคะแนนแบบสอบระดับ ม.3 จากกลุ่มสอบทานผลในระดับ ม.2…….. 126

4.15 การประเมินค่าดัชนีความแตกต่าง (C) เมื่อแปลงคะแนนแบบสอบระดับชั้น ม.3 ไปยังคะแนนแบบสอบระดับ ม.1 จากกลุ่มสอบทานผลในระดับ ม.3…….. 127

4.16 การประเมินค่าดัชนีความแตกต่าง (C) เมื่อแปลงคะแนนแบบสอบระดับชั้น ม.3 ไปยังคะแนนแบบสอบระดับ ม.2 จากกลุ่มสอบทานผลในระดับ ม.3…….. 128

ตารางที่ หน้า

4.17 การประเมินค่าดัชนีความแตกต่าง (C) ซึ่งเป็นความคลาดเคลื่อนรวมจากทุกรูปแบบการแปลงคะแนนตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง..... 129

4.18 ค่าสถิติพื้นฐานของคะแนนความสามารถ (θ) และ คะแนนจริง (ξ) จากผลการสอบของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 ภายหลังการปรับเทียบคะแนนในกลุ่มรวม.... 131

4.19 ผลการวิเคราะห์ความแปรปรวนทางเดียวของคะแนนความสามารถของนักเรียนระดับชั้น ม.1 ม.2 และ ม.3 ในกลุ่มรวม..... 133

4.20 ผลการทดสอบความแตกต่างระหว่างคะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1-3 ในกลุ่มรวม..... 133

4.21 คะแนนความสามารถเฉลี่ยและอัตราความงอกงาม (Growth Rate) ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1-3 โดยภาพรวม..... 134

4.22 ค่าสถิติพื้นฐานของนักเรียนจำแนกตามระดับชั้นจำแนกตามกลุ่มความสามารถ..... 136

4.23 ผลการวิเคราะห์ความแปรปรวนทางเดียวของคะแนนความสามารถของนักเรียนระดับชั้น ม.1 ม.2 และ ม.3 ตามกลุ่มความสามารถ..... 139

4.24 ผลการทดสอบความแตกต่างระหว่างคะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 จำแนกตามกลุ่มความสามารถ..... 141

4.25 คะแนนความสามารถเฉลี่ยและอัตราความงอกงาม (Growth Rate) ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1-3 จำแนกตามกลุ่มความสามารถ..... 142

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญแผนภาพ

แผนภาพที่ **หน้า**

1.1 กรอบแนวคิดของการวิจัย..... 13

2.1 การปรับเทียบคะแนนตามแนวนอน (Horizontal Equating)..... 20

2.2 การปรับเทียบคะแนนตามแนวตั้ง (Vertical Equating)..... 21

2.3 ขั้นตอนแรกของการพิจารณาการปรับเทียบโดยใช้ IRT..... 45

2.4 สรุปขั้นตอนของการปรับเทียบคะแนนโดยใช้ IRT..... 48

2.5 การออกแบบรวบรวมข้อมูลของ Lord และ Wingersky (1984)..... 63

3.1 รูปแบบการจัดฉบับแบบสอบที่มีข้อสอบร่วมภายในระหว่างชั้น..... 85

3.2 สรุปขั้นตอนการวิเคราะห์ข้อมูล..... 96

4.1 พัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนในกลุ่มรวม..... 135

4.2 แสดงพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 จำแนกตามกลุ่มความสามารถของนักเรียน..... 144

4.3 แสดงพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนตามกลุ่มความสามารถแยกตามระดับชั้น 145

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

ด้วยพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 และแก้ไขเพิ่มเติม (ฉบับที่ 2) พ.ศ. 2545 มาตรา 26 ได้ให้ความสำคัญกับการประเมินผลการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญโดยพิจารณาจากพัฒนาการของผู้เรียน ความประพฤติ การสังเกต พฤติกรรมการเรียน การร่วมกิจกรรม การทดสอบควบคู่กันไปในกระบวนการเรียนการสอน ตามความเหมาะสมของแต่ละระดับและรูปแบบการศึกษา ซึ่งเป็นสิ่งสำคัญที่จะช่วยให้เกิดความเข้าใจในตัวเด็ก เข้าใจสภาพการเรียนการสอนที่เป็นอยู่ อันจะนำไปสู่การจัดการเรียนการสอนที่เหมาะสมกับเนื้อหาและตัวผู้เรียน เพื่อปรับปรุงกระบวนการเรียนการสอนให้มีประสิทธิภาพยิ่งขึ้น

การประเมินผลการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ เชื่อว่านักเรียนแต่ละคนมีพัฒนาการที่แตกต่างกัน ดังนั้นการประเมินและทดสอบต้องเป็นแบบเฉพาะบุคคล และมีความเหมาะสมกับพัฒนาการ พร้อมทั้งให้สารสนเทศของวิธีการที่จะนำมาสอนผู้เรียนได้อย่างมีประสิทธิภาพ และสรรค์สร้างผู้เรียนให้มีความสำเร็จมากขึ้น ซึ่ง บุญเชิด ภิญโญอนันตพงษ์ (2544) กล่าวว่า “การประเมินที่เหมาะสมกับพัฒนาการ ได้มาจากการพัฒนาหลักสูตรอย่างเหมาะสมกับพัฒนาการ และในทางกลับกัน หลักสูตรมีความเหมาะสมกับการพัฒนาของผู้เรียน ได้มาจากการประเมินที่เหมาะสมกับพัฒนาการ” และลักษณะสำคัญของการประเมินผลการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญมีลักษณะมุ่งรวบรวมสารสนเทศของการพัฒนาการ และการเรียนรู้มุ่งเน้นพัฒนาการที่เกิดขึ้นอย่างเด่นชัด สอดคล้องกับพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 และแก้ไขเพิ่มเติม (ฉบับที่ 2) พ.ศ. 2545 มาตรา 30 ได้กำหนดให้สถานศึกษาส่งเสริมให้ผู้สอนสามารถทำวิจัย เพื่อพัฒนาการเรียนรู้ที่เหมาะสมกับผู้เรียนในแต่ละระดับการศึกษา แต่จากการศึกษาสภาพการจัดการเรียนรู้ตามหลักสูตรการศึกษาขั้นพื้นฐาน พ.ศ. 2544 ของโรงเรียนนำร่องและโรงเรียนเครือข่าย (สำนักวิชาการและมาตรฐานการศึกษา, 2547) พบว่า ผู้ที่เกี่ยวข้องกับการจัดการศึกษาในสถานศึกษาเกือบร้อยละ 70 ไม่ได้ใช้การวิจัยเป็นส่วนหนึ่งของกระบวนการเพื่อพัฒนาการเรียนรู้ของผู้เรียน นอกจากนี้ผลการวิจัยยังพบว่า ผลการเรียนของนักเรียนยังไม่ได้มาตรฐาน เขียนไม่ได้ คิดไม่เป็น โดยเฉพาะในกลุ่มสาระการเรียนรู้คณิตศาสตร์ ดังนั้นจึงมีความจำเป็นที่จะต้องมีการทำวิจัยเพื่อพัฒนาการเรียนรู้ของผู้เรียนควบคู่ไปในกระบวนการเรียนการสอน เช่น การศึกษาวิจัยเชิงพัฒนาการโดยเฉพาะการทราบถึงพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียน เนื่องจากคณิตศาสตร์

เป็นกลุ่มสาระที่สำคัญในการที่จะพัฒนาเยาวชนของประเทศให้มีความสามารถและมีศักยภาพที่จะแข่งขันกับประเทศอื่นได้ (อวยพร เรืองตระกูล, 2544) และการรู้ถึงพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียน จะเป็นแนวทางในการวางแผนพัฒนาการเรียนการสอนที่เหมาะสม ทั้งยังก่อให้เกิดการเรียนรู้ พัฒนาการความสามารถและแนวโน้มเป็นไปในทิศทางและรูปแบบใด อันจะเป็นการพัฒนาผู้เรียนให้ไปสู่จุดมุ่งหมายของหลักสูตรคณิตศาสตร์ได้อย่างเหมาะสม เพื่อนำผลการวิจัยที่ได้ไปใช้เป็นข้อมูลป้อนกลับ (Feedback) ในการปรับปรุงและวางแผนการเรียนการสอนให้มีประสิทธิภาพมากขึ้น

การวิจัยเชิงพัฒนาการ เป็นการศึกษาที่ดูความก้าวหน้าหรือพัฒนาการของสิ่งใดสิ่งหนึ่ง ไม่เพียงแต่ศึกษาสภาพที่เกิดขึ้น และความสัมพันธ์ของปรากฏการณ์เท่านั้น ยังรวมถึงการวิจัยที่มุ่งศึกษาการเปลี่ยนแปลงตามเงื่อนไขเวลา (Time Dimention) ซึ่ง ผ่องพรรณ ตรัยมงคลกูล (2545) ได้แบ่งงานวิจัยในกลุ่มการวิจัยเชิงพัฒนาการออกเป็น 2 ประเภท คือ การศึกษาระยะยาว (Longitudinal Study) เป็นศึกษาจากกลุ่มตัวอย่างกลุ่มเดียวกันในช่วงเวลาต่างกัน ซึ่งช่วงเวลาดังกล่าว อาจจะไม่ยาวมากนักหรืออาจยาวนานเป็นสิบปี ส่วนการศึกษาภาคตัดขวาง (Cross Sectional Study) เป็นการศึกษาจากกลุ่มตัวอย่างต่าง ๆ กันหลายกลุ่มที่เป็นตัวแทนของแต่ละช่วงพัฒนาการขณะใดขณะหนึ่งพร้อมกัน ซึ่งจากลักษณะของการศึกษาระยะยาว จะเป็นการวิจัยที่น่าตอบคำถามได้ดีในเรื่องการเปลี่ยนแปลงของผู้เรียน แต่มีข้อจำกัดที่เด่นชัดในเรื่องของความสิ้นเปลือง ความไม่ต่อเนื่องในการเก็บข้อมูลกับกลุ่มเดิม โดยเฉพาะข้อจำกัดในเรื่องระยะเวลา ซึ่งการจัดการศึกษาในปัจจุบันจำเป็นที่จะต้องมีการวิจัยเชิงพัฒนาการของนักเรียนในระดับชั้นต่างๆอย่างต่อเนื่อง เพื่อนำไปใช้ในการปรับปรุงแก้ไขได้ทันท่วงที ดังนั้นการศึกษาระยาวอาจจะไม่เหมาะสม ซึ่งในสถานการณ์ดังกล่าวการศึกษาภาคตัดขวางน่าจะเหมาะสมกว่าในด้านความรวดเร็วในการรู้ผลการวิจัย มีปัญหาน้อยในเรื่องการสูญหายของกลุ่มตัวอย่างและความคลาดเคลื่อนในการวัดจากกลุ่มตัวอย่าง

แม้ว่าการศึกษาพัฒนาการภาคตัดขวาง จะให้สารสนเทศเพื่อหาข้อบกพร่องในกระบวนการเรียนการสอนเป็นสำคัญ อันจะนำไปสู่การเร่งแก้ปัญหาได้ทันท่วงทีก่อนที่นักเรียนจะจบการศึกษาออกไปจากโรงเรียน และเป็นประโยชน์ต่อการกำหนดนโยบาย หรือการพัฒนาการเรียนการสอนในระดับหนึ่ง แต่ก็ยังนับว่าไม่เพียงพอ ในการสรุปโยงความสัมพันธ์ระหว่างผลการประเมินต่างระดับชั้นเหล่านั้นอย่างน่าเชื่อถือได้ เพราะการประเมินในช่วงเวลาใดเวลาหนึ่ง โดยใช้กลุ่มตัวอย่างที่ต่างกันและใช้เครื่องมือวัดที่ต่างกันมาเปรียบเทียบกันโดยตรง ทำให้ผลการเปรียบเทียบมีความคลาดเคลื่อนสูง ขาดความแม่นยำของสารสนเทศอันส่งผลสืบเนื่องให้การตัดสินใจสรุปผลผิดพลาดได้ (ศิริชัย กาญจนวาสี, 2545) ดังนั้นการแก้ไขปัญหาการเปรียบเทียบ

ผลการเรียนของบุคคลต่างกลุ่มที่ได้จากเครื่องมือวัดที่ต่างกัน ในปีการศึกษาเดียวกัน เพื่อนำมาเปรียบเทียบและศึกษาพัฒนาการในการเรียนได้ จึงควรใช้เทคนิควิธีการที่เรียกว่า "**การปรับเทียบแนวตั้ง**" (Vertical Equating) ในปีการศึกษาเดียวกันระหว่างแบบสอบต่างฉบับ ได้จากนักเรียนต่างกลุ่มกัน แล้วจึงนำผลที่ได้มาเปรียบเทียบกัน การปรับเทียบคะแนนดังกล่าว เป็นวิธีการเปลี่ยนระบบหน่วยคะแนนของแบบสอบฉบับหนึ่ง ไปสู่ระบบหน่วยคะแนนของแบบสอบอีกฉบับหนึ่ง เพื่อให้ได้คะแนนที่แปลงจากแบบสอบทั้งสองฉบับมีความเท่าเทียมกัน (Angoff, 1971) ผลการเปรียบเทียบจึงให้สารสนเทศที่ถูกต้อง แม่นยำ และเพียงพอต่อการตัดสินใจ

ศิริชัย กาญจนวาสี (2545) ได้ให้ความหมายของการปรับเทียบคะแนนว่า เป็นกระบวนการแปลงคะแนนของแบบสอบต่างฉบับ ที่มุ่งวัดคุณลักษณะเดียวกันสามารถนำมาเปรียบเทียบกันได้ ซึ่งถ้าพิจารณาการนำไปใช้กับผู้สอบและแบบสอบ Hambleton และ Swaminathan (1984) ได้แบ่งการปรับเทียบคะแนนออกเป็นสองสถานการณ์ ได้แก่ การปรับเทียบคะแนนตามแนวนอน (Horizontal Equating) ใช้ในกรณีที่ต้องการเปรียบเทียบแบบสอบที่มีระดับความยากพอ ๆ กัน และการกระจายความสามารถของผู้ตอบแบบสอบคล้ายคลึงกัน และการปรับเทียบแนวตั้ง (Vertical Equating) ใช้ในกรณีเปรียบเทียบแบบสอบที่มีระดับความยากต่างกันและการกระจายความสามารถของผู้สอบต่างกัน

จากการศึกษางานวิจัยทั้งในประเทศและต่างประเทศ นิยมนำวิธีการปรับเทียบแนวตั้ง (Vertical Equating) ตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ในการศึกษาพัฒนาการหรือความก้าวหน้าในการเรียนซึ่งใช้แบบสอบหลายระดับกับผู้สอบที่มีระดับความสามารถแตกต่างกัน เนื่องจากให้ผลการปรับเทียบที่มีความเพียงพอและน่าพอใจกว่าวิธีอื่น (Slinde & Linn ,1977-1979; Kolen, 1980; Kolen & Brennan,1995; พรพิมล นาคเวช, 2537; วรรณ์ดี แสงประทีปทอง, 2538; ภัทรพร เกษสังข์, 2546) แต่ปัญหาที่มักจะเกิดขึ้นเสมอเมื่อปรับเทียบคะแนนกับกลุ่มผู้สอบที่มีความสามารถแตกต่างกันมาก ๆ คือเกิดความคลาดเคลื่อนของการปรับเทียบคะแนนสูงโดยเฉพาะการปรับเทียบคะแนนในระดับชั้นที่ไม่ได้อยู่ติดกัน เช่น การปรับเทียบคะแนนความสามารถระหว่างระดับชั้น ม.1 และ ม.3 เป็นต้น โดยส่งผลต่อความถูกต้อง (Accuracy) และความแม่นยำ (Precision) ที่แสดงถึงประสิทธิภาพของการปรับเทียบคะแนนโดยตรง ซึ่งเป็นปัญหาที่เรียกว่า "Out-of–Level Effect" (O'Brien and Tohn, 1984; Bielinksi, Thurlow & Scottt, 2000; Leung, 2003)

Cook, Dunbar และ Eignor (1981) ได้ให้ข้อเสนอแนะว่าสำหรับการประมาณค่าแตกต่างกันมากๆ โดยเฉพาะการประมาณค่าที่ต่ำกว่าระดับการเดา (Lower Asymtote) ที่มักจะเกิดขึ้นเสมอสำหรับการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ สามารถใช้วิธีการเชิงเส้นตรง (Linear Interpolation) ซึ่งให้ผลการปรับเทียบที่ดีกว่าในสถานการณ์

เช่นนี้ ข้อเสนอแนะดังกล่าวสอดคล้องกับผลการศึกษาของ Leung (2003) ได้เสนอวิธีการปรับเทียบแนวตั้ง (Vertical Equating) โดยใช้การปรับเทียบร่วมกันระหว่างการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ กับวิธีการปรับเทียบเชิงเส้นตรง โดยใช้ข้อสอบร่วม จุดมุ่งหมายของการศึกษาครั้งนี้เพื่อศึกษาพัฒนาการความสามารถทางด้านภาษาจีนของนักเรียนระดับประถมศึกษาในฮ่องกง โดยกลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนระดับประถมศึกษาที่ทำการทดสอบในปี 1999 ซึ่งแบ่งนักเรียนออกเป็นระดับความสามารถ 6 ระดับ โดยใช้การปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ เพื่อประมาณค่าความสามารถของผู้สอบที่อยู่ในระดับชั้นที่ต่อเนื่องกัน สำหรับการปรับเทียบเชิงเส้นตรงจะใช้ในการประมาณค่าโดยวิธี “Simple Least Squared Regression” สำหรับการประมาณค่าในระดับความสามารถที่แตกต่างกันมากๆ ผลการตรวจสอบประสิทธิภาพของการปรับเทียบคะแนน โดยพิจารณาความสัมพันธ์ระหว่างการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ กับการปรับเทียบหลังจากมีการปรับแก้ค่าโดยใช้วิธีการปรับเทียบเชิงเส้นตรง ซึ่งค่าความสัมพันธ์ที่ได้มีค่าเข้าใกล้ 1 เกือบทุกค่า แสดงให้เห็นว่าคุณภาพของการปรับเทียบคะแนนอยู่ในระดับดี

ผลจากการศึกษาของ Leung (2003) จะเห็นได้ว่า มีการเพิ่มประสิทธิภาพของการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ โดยใช้วิธีร่วมกับการปรับเทียบเชิงเส้นตรง ซึ่งยังไม่มีผู้ที่ศึกษามาก่อน และมีการตรวจสอบประสิทธิภาพของการปรับเทียบพิจารณาจากความสัมพันธ์ระหว่างค่าพารามิเตอร์ที่ได้จากการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ กับการปรับเทียบหลังจากมีการปรับแก้ค่าจากการปรับเทียบเชิงเส้นตรง ผลที่ได้จากการปรับเทียบจะนำมาใช้ในการศึกษาพัฒนาการทางด้านภาษา ดังนั้นผู้วิจัยจึงนำวิธีการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรงเพื่อเพิ่มประสิทธิภาพของการปรับเทียบแนวตั้ง แต่จะแตกต่างจากการวิจัยของ Leung (2003) คือศึกษาพัฒนาการความสามารถทางด้านคณิตศาสตร์จากกลุ่มผู้สอบที่แตกต่างกัน 3 ระดับชั้น คือ ม.1, ม.2 และ ม.3 โดยใช้วิธีการตรวจสอบประสิทธิภาพของการปรับเทียบจากการศึกษาสมการปรับเทียบจากกลุ่มตัวอย่างปรับเทียบคะแนน และนำสมการปรับเทียบที่ได้ไปใช้ตรวจสอบคุณภาพกับกลุ่มตัวอย่างสอบทานผล ซึ่งจะพิจารณาค่าความคลาดเคลื่อนมาตรฐานของการปรับเทียบ (Standard Error of Equating) และวิธีการประเมินความเพียงพอของการปรับเทียบคะแนนจากการวิเคราะห์ผลในกลุ่มตัวอย่างสอบทานผล ที่เสนอโดย Kolen (1982) โดยได้ใช้ข้อมูลจากผู้สอบเองเป็นเกณฑ์ในการหาความแตกต่าง ดังนั้นการใช้คะแนนของตนเองเป็นเกณฑ์จึงมีความเป็นอิสระ ไม่ขึ้นกับกระบวนการแปลงคะแนนอื่น ๆ (พรพิมล นาคเวช, 2537) แต่การหาค่าดัชนีได้ดัดแปลงจากสูตรของโคเลน ไปใช้ตามแนวคิดของ Petersen และคณะ (1982) คือ ใช้

ค่าความแปรปรวนเป็นตัวถ่วงน้ำหนักเพื่อให้ค่าที่ได้มีหน่วยเป็นมาตรฐาน ซึ่งวิธีการตรวจสอบประสิทธิภาพดังกล่าวมีความเป็นมาตรฐานและยอมรับกันโดยทั่วไป

ดังนั้นผู้วิจัยจึงสนใจศึกษาการปรับเทียบแนวตั้ง (Vertical Equating) ตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง เพื่อเพิ่มประสิทธิภาพของการปรับเทียบคะแนน และนำมาใช้ในการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ ของนักเรียนในช่วงชั้นที่ 3 คือ ม.1, ม.2 และ ม.3 ซึ่งนักเรียนทั้ง 3 ระดับชั้น เป็นนักเรียนในภาคการศึกษาและปีการศึกษาเดียวกัน อันเป็นการศึกษาพัฒนาการในลักษณะภาคตัดขวาง (Cross-Sectional Study) โดยจัดให้มีแบบสอบร่วมภายใน ข้อมูลที่ได้จากแบบสอบร่วมจะนำมาใช้ในการปรับความแตกต่างระหว่างความสามารถของนักเรียนต่างระดับชั้นให้อยู่ในสเกลเดียวกัน อันจะทำให้เห็นพัฒนาการที่ต่อเนื่องของแต่ละระดับชั้นได้อย่างชัดเจน วิธีการดังกล่าว สามารถนำไปใช้ในการศึกษาพัฒนาการในระดับชั้นอื่น เนื้อหาวิชาอื่น หรือความสามารถในด้านอื่นๆ และสามารถนำข้อมูลที่ได้จากการวิจัยไปใช้ในการวางแผนการเรียนการสอนให้มีประสิทธิภาพมากขึ้นทั้งในระดับสถานศึกษา เขตพื้นที่การศึกษา และระดับจังหวัด เพื่อเป็นการเตรียมความพร้อมสู่การทดสอบระดับชาติ (National Test) ในการยกระดับมาตรฐานการศึกษาไทย

คำถามการวิจัย

งานวิจัยในอดีตที่ผ่านมา นิยมนำวิธีการปรับเทียบแนวตั้ง แบบโลจิสติก 3 พารามิเตอร์ มาใช้ในการศึกษาพัฒนาการหรือความก้าวหน้าในการเรียนกับนักเรียนในระดับชั้นที่แตกต่างกัน ซึ่งมีระดับความสามารถที่แตกต่างกัน แต่ปัญหาที่มักจะเกิดขึ้นเสมอเมื่อปรับเทียบคะแนนกับกลุ่มผู้สอบที่มีความสามารถแตกต่างกันมาก ๆ มักจะเกิดความคลาดเคลื่อนของการปรับเทียบคะแนน ซึ่งปัญหาดังกล่าวเรียกว่า “Out-of–Level Effect” ดังนั้นผู้วิจัยจึงนำวิธีการปรับเทียบเชิงเส้นตรง (Linear Equating) มาช่วยปรับแก้ปัญหาดังกล่าว เพื่อที่จะสามารถประมาณค่าความสามารถของผู้สอบได้ทุกระดับความสามารถ โดยสามารถลดความคลาดเคลื่อนของการปรับเทียบคะแนน และเป็นการเพิ่มประสิทธิภาพของการปรับเทียบคะแนน ซึ่งงานวิจัยในประเทศไทยยังไม่มีใครศึกษาและการศึกษาถึงพัฒนาการความสามารถทางคณิตศาสตร์ จำเป็นที่จะต้องศึกษาพัฒนาการโดยภาพรวมของกลุ่มตัวอย่าง และควรมีการศึกษาตามระดับความสามารถของนักเรียน ซึ่งเป็นสิ่งจำเป็นอย่างยิ่งต่อบุคคลที่เกี่ยวข้องกับผู้เรียน ดังนั้น เพื่อให้ได้สารสนเทศที่นำไปสู่ข้อเสนอแนะเกี่ยวกับการปรับปรุงการเรียนการสอนอันจะนำไปสู่การพัฒนาผู้เรียนตามระดับความสามารถและเต็มตามศักยภาพ ผู้วิจัยได้กำหนดคำถามวิจัยสำคัญดังนี้

1) การปรับเทียบแนวตั้งตามวิธีทฤษฎีการตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง มีประสิทธิภาพของการปรับเทียบคะแนนเป็นเช่นไร

2) ความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 โดยภาพรวมมีลักษณะพัฒนาการเป็นอย่างไร ภายหลังการใช้วิธีการปรับเทียบคะแนนตาม ข้อ (1)

3) ความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 จำแนกตามกลุ่มความสามารถ คือ กลุ่มต่ำ กลุ่มปานกลาง และ กลุ่มสูง มีลักษณะพัฒนาการเป็นอย่างไร กลุ่มใดมีอัตราพัฒนาการมากที่สุดและกลุ่มใดมีอัตราพัฒนาการต่ำที่สุด ภายหลังการใช้วิธีการปรับเทียบคะแนนตาม ข้อ (1)

วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีจุดมุ่งหมายหลัก เพื่อศึกษาประสิทธิภาพของการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง และศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ จากกลุ่มผู้สอบต่างกัน 3 ระดับชั้น ได้แก่ ระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 โดยมีจุดมุ่งหมายเฉพาะดังนี้

1) เพื่อศึกษาประสิทธิภาพของการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบแบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรงที่ใช้แบบสอบร่วมภายใน

2) เพื่อศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 โดยเปรียบเทียบคะแนนความสามารถทางคณิตศาสตร์ของนักเรียน ในแต่ละระดับชั้นโดยภาพรวม และจำแนกตามกลุ่มความสามารถ คือ กลุ่มต่ำ กลุ่มปานกลาง และ กลุ่มสูง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สมมติฐานการวิจัย

จากผลการศึกษาของ Leung (2003) ได้ทำการศึกษาการปรับเทียบแนวตั้งตามทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง มาใช้ในการศึกษาพัฒนาการความสามารถทางภาษา สำหรับแบบสอบหลายระดับสำหรับผู้สอบที่มีความสามารถแตกต่างกัน 6 ระดับความสามารถ ซึ่งใช้วิธีการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบโดยการปรับเทียบตามแนวตั้ง แบบโลจิสติก 3 พารามิเตอร์ เพื่อประมาณค่าความสามารถของผู้สอบที่อยู่ระดับชั้นติดกัน (Adjacent Grades) สำหรับการปรับเทียบเชิงเส้นตรงจะใช้ในการประมาณค่าความสามารถเพื่อป้องกันปัญหาที่เรียกว่า “Out-of-Level Effect” ที่มักจะเกิดขึ้นเสมอภายหลังการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ ผลจากการศึกษาประสิทธิภาพของการปรับเทียบคะแนนโดยพิจารณาความสัมพันธ์ระหว่างการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ กับการปรับเทียบหลังจากมีการปรับแก้ค่าจากการปรับเทียบเชิงเส้นตรง พบว่าค่าความสัมพันธ์ที่ได้มีค่าเข้าใกล้ 1 เกือบทุกค่า แสดงให้เห็นว่าคุณภาพของการปรับเทียบคะแนนอยู่ในระดับดี ผู้วิจัยจึงได้กำหนดสมมติฐานของการวิจัยตามวัตถุประสงค์ข้อที่ 1 ดังนี้

1) การปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรงที่ใช้แบบสอบร่วมภายใน มีความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน (Standard Error of Equating : SEE) น่าจะอยู่ในระดับต่ำ และ ความเพียงพอของความคลาดเคลื่อนในการปรับเทียบคะแนนน่าจะมีค่าดัชนีความแตกต่างของการปรับเทียบคะแนน (C) อยู่ในระดับน่าพอใจขึ้นไป ($C \leq (0.1\ SD_x)^2$)

ในการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียน ภัทราพร เกษสังข์ (2546) ได้ทำการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียน จากกลุ่มตัวอย่างที่มีความสามารถแตกต่างกันสามระดับชั้น คือ มัธยมศึกษาปีที่ 1, 2 และ 3 ผลการวิจัยสรุปว่าความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.1 ไประดับชั้น ม. 2 กับนักเรียนชั้น ม. 2 ไประดับชั้น ม.3 มีพัฒนาการสูงขึ้น และนักเรียนระดับชั้น ม. 1 ไปชั้น ม.3 มีพัฒนาการสูงขึ้นมาก สำหรับงานวิจัยที่ศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนในช่วงชั้นที่ 3 ที่**แบ่งตามกลุ่มความสามารถ คือ กลุ่มต่ำ กลุ่มปานกลาง และกลุ่มสูง** ยังไม่มีผู้ศึกษาโดยตรง แต่เนื่องจากนักเรียนระดับชั้น ม.1 ในกลุ่มต่ำและกลุ่มปานกลาง เป็นนักเรียนชั้นมัธยมศึกษาปีแรกของหลักสูตร ซึ่งต้องมีการปรับตัวในเรื่องต่างๆ เมื่อเรียนในชั้นที่สูงขึ้นนักเรียนมีโอกาสพัฒนาตัวเองให้ก้าวหน้ามากขึ้น และเป็นนักเรียนที่มีพื้นฐานความรู้น้อยกว่านักเรียนชั้นมัธยมศึกษาอื่นๆ

เมื่อเรียนในชั้นที่สูงขึ้นนักเรียนมีโอกาสพัฒนาตนเองให้ก้าวหน้ามากขึ้น ดังนั้นพัฒนาการจึงน่าจะเพิ่มขึ้นตามระดับชั้นที่สูงขึ้น และสำหรับนักเรียนที่มีความสามารถสูง มักจะพัฒนาตัวเองอย่างสม่ำเสมอในทุกระดับชั้น ดังนั้นพัฒนาการจึงน่าจะสูงขึ้นตามระดับความสามารถที่สูงขึ้น ผู้วิจัยจึงได้กำหนดสมมติฐานของการวิจัยตามวัตถุประสงค์ข้อที่ 2 และ 3 ดังนี้

2) ความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 โดยภาพรวมน่าจะมีพัฒนาการสูงขึ้นตามระดับชั้นที่สูงขึ้น

3) ความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ในกลุ่มต่ำ กลุ่มปานกลาง และกลุ่มสูง น่าจะมีพัฒนาการเพิ่มขึ้นตามระดับชั้นที่สูงขึ้น โดยนักเรียนที่อยู่ในกลุ่มสูงน่าจะมีอัตราการพัฒนาการมากที่สุดและนักเรียนที่อยู่ในกลุ่มต่ำน่าจะมีอัตราพัฒนาการต่ำที่สุด

ขอบเขตของการวิจัย

1) ประชากรที่ใช้ในการวิจัยครั้งนี้เป็นนักเรียนช่วงชั้นที่ 3 คือนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ของโรงเรียนในเขตพื้นที่การศึกษาสกลนคร เขต 1 ซึ่งกลุ่มตัวอย่างที่ใช้ได้มาจากการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) มีขนาดของโรงเรียนเป็นตัวแปรสำหรับแบ่งชั้น คือ ขนาดใหญ่พิเศษ ขนาดใหญ่ ขนาดกลาง และขนาดเล็ก

2) ตัวแปรในการวิจัย คือ

2.1 ประสิทธิภาพของการปรับเทียบคะแนน

2.2 พัฒนาการความสามารถทางคณิตศาสตร์

3) ความสามารถทางคณิตศาสตร์ที่นำมาศึกษาในครั้งนี้ เป็นความสามารถในการเรียนรู้เรื่อง “จำนวนและการดำเนินการ” ของกลุ่มสาระการเรียนรู้คณิตศาสตร์ ในช่วงชั้นที่ 3 ที่วัดจากแบบสอบเลือกตอบ 4 ตัวเลือก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

คำจำกัดความเชิงปฏิบัติการ

การปรับเทียบคะแนนของแบบสอบ (Test Equating) หมายถึง กระบวนการแปลงคะแนนจากแบบสอบวัดความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 โดยวิธีการทางสถิติเพื่อปรับเทียบคะแนนระหว่างแบบสอบต่างชุด ให้อยู่ในมาตรวัดหรือสเกล (Scale) เดียวกันเพื่อให้สามารถปรับเทียบคะแนนกันได้

การปรับเทียบแนวตั้ง (Vertical Equating) หมายถึง การปรับเทียบคะแนนที่ได้จากกลุ่มตัวอย่างที่ระดับความสามารถต่างกัน และแบบสอบที่ระดับความยากต่างกัน ในที่นี้ศึกษากลุ่มตัวอย่างและแบบสอบที่มีความยาก 3 ระดับ ได้แก่ ระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 และใช้แบบสอบ 3 ฉบับ ได้แก่ แบบสอบสำหรับชั้น ม.1, ม.2 และ ม.3 ตามลำดับ

ความสามารถทางคณิตศาสตร์ หมายถึง ความสามารถของนักเรียนที่ได้จากการตอบแบบสอบวัดความสามารถทางคณิตศาสตร์ เรื่อง “จำนวนและการดำเนินการ” ระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ที่วัดจากข้อสอบแบบเลือกตอบ 4 ตัวเลือก

พัฒนาการความสามารถทางคณิตศาสตร์ หมายถึง คะแนนความสามารถทางคณิตศาสตร์ของนักเรียนที่เพิ่มขึ้น ภายหลังจากการปรับเทียบคะแนนของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ให้อยู่ในสเกลเดียวกันแล้ว ซึ่งเป็นนักเรียนคนละกลุ่มที่เรียนอยู่ในภาคการศึกษาและปีการศึกษาเดียวกัน ซึ่งพิจารณาจาก “อัตราพัฒนาการ” (Growth Rate) นั่นคือ อัตราพัฒนาการของนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.2, ม.2 เมื่อขึ้น ม.3 และ ม.1 เมื่อขึ้น ม.3 มีค่าเพิ่มขึ้นตามระดับชั้นที่สูงขึ้น

อัตราพัฒนาการ (Growth Rate) หมายถึง วิธีการวัดคะแนนพัฒนาการตามทฤษฎีตอบสนองข้อสอบ ของอวยพร เรืองตระกูล (2544) ซึ่งได้ประยุกต์แนวคิดในการวัดคะแนนพัฒนาการตามทฤษฎีตอบสนองข้อสอบแบบดั้งเดิมของ ศิริชัย กาญจนวาสี (2532) และแนวคิดของ วินิจ เทือกทอง (2537) มาใช้ในการวัดคะแนนพัฒนาการจากความสามารถที่แท้จริงสัมพัทธ์ กรณีการวัด 2 ครั้ง โดยเทียบกับศักยภาพการพัฒนา นำมาประยุกต์ใช้ในการศึกษาแนวโน้มการเปลี่ยนแปลงคะแนนเฉลี่ยของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 โดยในที่นี้เป็นการประมาณค่าจากอัตราส่วนผลต่างของความสามารถที่แท้จริงของนักเรียนที่มีระดับชั้นสูงกว่าและระดับชั้นที่ต่ำกว่ากับศักยภาพการพัฒนา โดยนิยามศักยภาพการพัฒนาว่าเป็นผลต่างของความสามารถสูงสุด (Max = 4) กับความสามารถที่แท้จริงของนักเรียนที่มีระดับชั้นที่ต่ำกว่า

กลุ่มความสามารถ หมายถึง การแบ่งนักเรียนตามระดับความสามารถ (θ) ออกเป็น 3 กลุ่มความสามารถ คือ กลุ่มต่ำ, กลุ่มปานกลาง และกลุ่มสูง จะแบ่งกลุ่มโดยใช้คะแนนของนักเรียนที่ได้จากการทำแบบสอบปรับเทียบคะแนนตามชั้นของตนตามเกณฑ์ของ วรรณ์ดี แสงประทีปทอง (2538) ดังนี้

ตารางที่ 1.1 เกณฑ์การจำแนกนักเรียนตามกลุ่มความสามารถ

กลุ่มความสามารถ	เกณฑ์การจำแนก
ต่ำ	$\theta \leq \theta$ - 0.84 SD.
ปานกลาง	θ - 0.25 SD.$< \theta \leq \theta$ + 0.25 SD.
สูง	$\theta \geq \theta$ + 0.84 SD.

แบบสอบปรับเทียบคะแนน หมายถึง แบบสอบต่างๆ ที่ใช้ในการปรับเทียบคะแนนสำหรับการวิจัยนี้ หมายถึง แบบสอบคณิตศาสตร์ 3 ฉบับ สำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ที่สร้างจากหลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 ช่วงชั้นที่ 3 เรื่อง "จำนวนและการดำเนินการ" แบบสอบแต่ละฉบับจะประกอบด้วยข้อสอบร่วมภายในร้อยละ 20 ของข้อสอบในแต่ละฉบับ

แบบสอบร่วมภายใน (Internal Anchor Test) หมายถึง กลุ่มข้อสอบร่วมที่อยู่ในแบบสอบปรับเทียบคะแนนทั้ง 3 ฉบับ

ข้อสอบร่วมภายใน (Internal Anchor Items) หมายถึง ข้อสอบร่วมระหว่างระดับชั้นที่ต่อเนื่องที่อยู่ในแบบสอบร่วมภายในแต่ละระดับชั้นที่ต่อเนื่องนั้น แบ่งออกเป็นสองชุด ดังนี้

ข้อสอบร่วมระหว่างระดับชั้นที่ต่อเนื่องชั้น ม.1 กับ ม.2 จำนวน 6 ข้อ จัดอยู่ในแบบสอบปรับเทียบคะแนนของระดับชั้น ม.1 และ ระดับชั้น ม.2

ข้อสอบร่วมระหว่างระดับชั้นที่ต่อเนื่องชั้น ม.2 กับ ม.3 จำนวน 6 ข้อ จัดอยู่ในแบบสอบปรับเทียบคะแนนของระดับชั้น ม.2 และ ระดับชั้น ม.3

สามารถเขียนเป็นรูปแบบของแบบสอบร่วมภายใน ได้ดังตารางที่ 1.2

ตารางที่ 1.2 รูปแบบของแบบสอบร่วมภายใน

ระดับชั้น	แบบสอบเฉพาะ	ข้อสอบร่วมระหว่างชั้น ม.1 กับ ม.2	ข้อสอบร่วมระหว่างชั้น ม.2 กับ ม.3
ม.1	24 ข้อ	6 ข้อ	-
ม.2	18 ข้อ	6 ข้อ	6 ข้อ
ม.3	24 ข้อ	-	6 ข้อ

คะแนนสมมูล (Equivalent Score) หมายถึง คะแนนจากแบบสอบฉบับ X ที่แปลงคะแนนจากตารางปรับเทียบคะแนนไปเป็นคะแนนจากแบบสอบฉบับ Y ของนักเรียนคนที่ i โดยเทียบคะแนนกันได้ว่าเป็นคะแนนที่เท่าเทียมกัน ซึ่งใช้กระบวนการปรับเทียบคะแนนจากกลุ่มปรับเทียบคะแนน

คะแนนเกณฑ์ (Criterion Score) หมายถึง คะแนนจากแบบสอบฉบับ Y ที่นักเรียนคนที่ i สอบได้จริงของนักเรียนกลุ่มสอบทานผล

กลุ่มปรับเทียบคะแนน (Equating Samples) หมายถึง กลุ่มที่สอบแบบสอบปรับเทียบคะแนนในระดับชั้นของตนมีแบบสอบร่วมภายใน ผลที่ได้จากกลุ่มตัวอย่างนี้ นำมาใช้ในการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์และการสร้างคะแนนแปลงตามรูปแบบการปรับเทียบแนวตั้ง 3 ระดับชั้น ในแต่ละรูปแบบของตัวแปรที่ศึกษา

กลุ่มสอบทานผล (Cross-Validation Samples) หมายถึง กลุ่มตัวอย่างที่ได้จากการสุ่มประชากรนักเรียนระดับช่วงชั้นที่ 3 ซึ่งเป็นประชากรเดียวกันกับกลุ่มตัวอย่างปรับเทียบคะแนนโดยไม่มีหน่วยตัวอย่างซ้ำกันเลย กลุ่มตัวอย่างสอบทานผลนี้ได้จากการสุ่มประชากรครั้งเดียวกับการสุ่มประชากรของกลุ่มตัวอย่างปรับเทียบคะแนน โดยนักเรียนทุกคนต้องทำแบบสอบปรับเทียบคะแนนที่มีแบบสอบร่วมภายในทั้ง 3 ฉบับ นำมาใช้ในการวิเคราะห์สอบทานผลการปรับเทียบคะแนนในแนวตั้ง 3 ระดับ เพื่อประเมินประสิทธิภาพของการปรับเทียบคะแนน ในรูปแบบของตัวแปรที่ศึกษา

ประสิทธิภาพของการปรับเทียบคะแนน (Efficiency of Equating) หมายถึง ความถูกต้อง (Accuracy) และความแม่นยำ (Precision) ของการปรับเทียบคะแนน ซึ่งพิจารณาจากความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน และความเพียงพอของความคลาดเคลื่อนของการปรับเทียบคะแนน

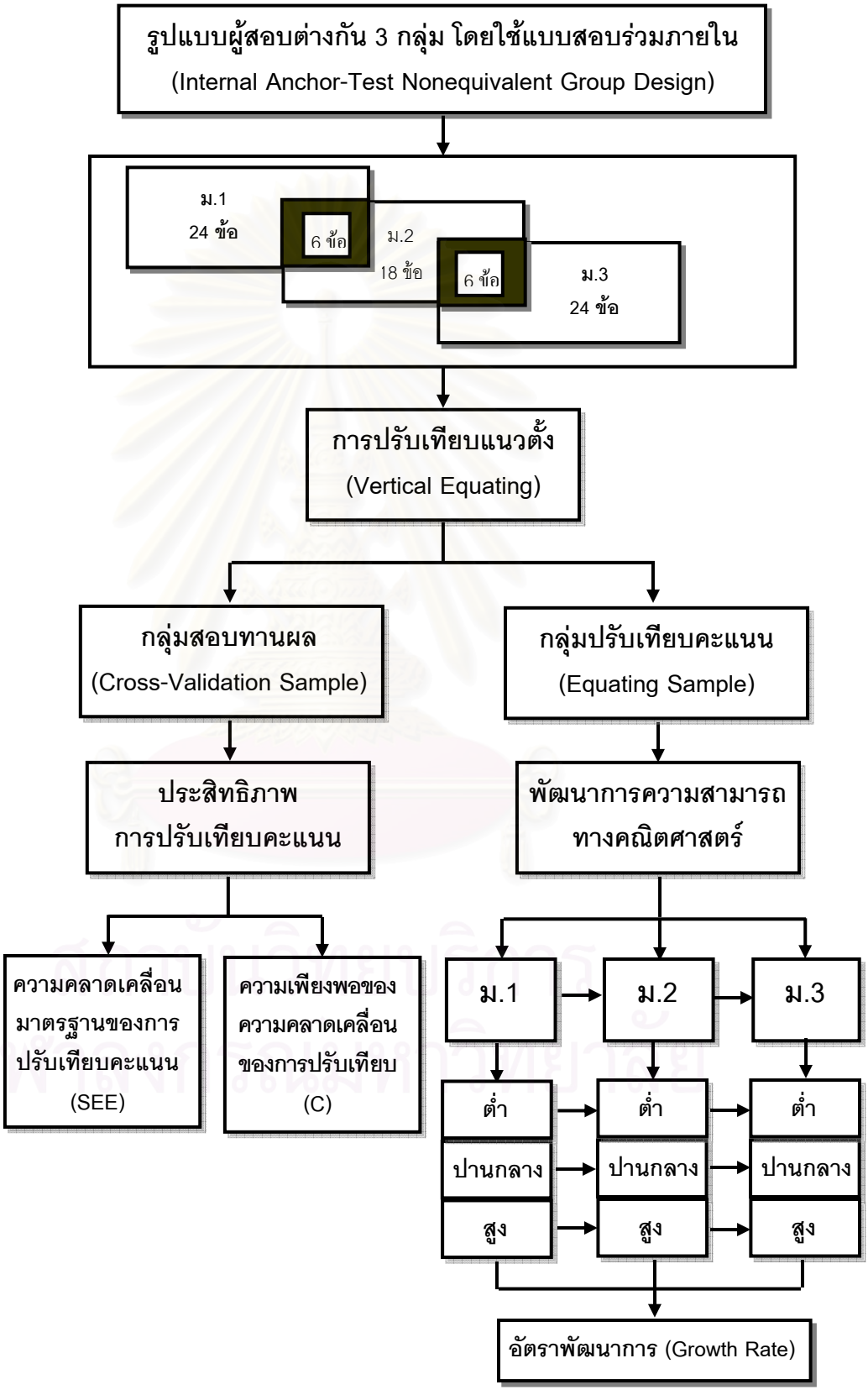
ความคลาดเคลื่อนมาตรฐานในการปรับเทียบคะแนน (Standard Error of Equating : SEE) หมายถึง ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ยของความแตกต่างระหว่างคะแนนสมมูลซึ่งเป็นคะแนนจากแบบสอบฉบับ X ที่แปลงคะแนนจากตารางปรับเทียบคะแนนไปเป็นคะแนนจากแบบสอบฉบับ Y ของนักเรียนคนที่ i กับคะแนนเกณฑ์ซึ่งเป็นคะแนนจากแบบสอบฉบับ Y ที่นักเรียนคนที่ i สอบได้จริงของนักเรียนกลุ่มสอบทานผลโดยไม่คิดทิศทาง

ดัชนีความแตกต่าง (Discrepancy Comparison : C) หมายถึง ค่าความคลาดเคลื่อนรวม (Total Error : Te) ซึ่งคำนวณได้จากค่าเฉลี่ยกำลังสองของความแตกต่างของคะแนนที่สมมูล ซึ่งได้จากการปรับเทียบคะแนนกับคะแนนเกณฑ์ที่ได้ถ่วงน้ำหนัก ด้วยความแปรปรวนของคะแนนเกณฑ์ ดังนั้นค่าที่ได้จึงเป็นค่ามาตรฐานที่นำมาเปรียบเทียบกับเกณฑ์ของความพึงพอใจในการปรับเทียบคะแนนได้ ดัชนีได้จากการวิเคราะห์กลุ่มสอบทานผล ใช้ประเมินประสิทธิภาพหรือความเพียงพอของวิธีการปรับเทียบคะแนน

ความเพียงพอของความคลาดเคลื่อนในการปรับเทียบคะแนน (Adequacy of Equating) หมายถึง ระดับความถูกต้องของการปรับเทียบคะแนน ของแบบสอบชุดหนึ่งไปสู่แบบสอบชุดอื่นตามวิธีการของการปรับเทียบคะแนน ในแต่ละรูปแบบของตัวแปรที่ศึกษา ซึ่งการตัดสินความเพียงพอของความคลาดเคลื่อนของการปรับเทียบคะแนนที่ใช้ในการวิจัยนี้ ประเมินจากค่าดัชนีความแตกต่างกับเกณฑ์การยอมรับของ Petersen และคณะ (1982) ดังนี้

		C	≤	$(0.05\ SD_x)^2$	หมายถึง	ระดับน่าพอใจมาก
$(0.05\ SD_x)^2$	<	C	≤	$(0.10\ SD_x)^2$	หมายถึง	ระดับน่าพอใจ
$(0.10\ SD_x)^2$	<	C	≤	$(0.15\ SD_x)^2$	หมายถึง	ระดับปานกลาง
$(0.15\ SD_x)^2$	<	C	≤	$(0.20\ SD_x)^2$	หมายถึง	ระดับไม่น่าพอใจ
$(0.20\ SD_x)^2$	<	C			หมายถึง	ระดับไม่น่าพอใจมาก

กรอบแนวคิดในการวิจัย



แผนภาพที่ 1.1 กรอบแนวคิดในการวิจัย

ประโยชน์ที่คาดว่าจะได้รับ

1) ได้ตารางปรับเทียบคะแนนความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3

2) การวิจัยครั้งนี้ช่วยให้ได้แนวทางในการตัดสินใจที่จะนำรูปแบบการปรับเทียบแนวตั้งโดยใช้ทฤษฎีตอบสนองข้อสอบร่วมกับวิธีเชิงเส้นตรง ไปใช้ในการวัดและประเมินผลตามหลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 ได้อย่างเหมาะสม

3) ได้ข้อมูลพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนช่วงชั้นที่ 3 ซึ่งสามารถนำไปใช้ในการวางแผนการเรียนการสอนให้มีประสิทธิภาพยิ่งขึ้น

4) ได้แนวทางการศึกษาพัฒนาการความสามารถของนักเรียนต่างระดับชั้น ซึ่งสามารถนำไปประยุกต์ใช้ในการเปรียบเทียบความสามารถของนักเรียนในเนื้อหาวิชาอื่นหรือในระดับอื่น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการนำเสนอเอกสารและงานวิจัยที่เกี่ยวข้องกับการวิจัยนี้ ผู้วิจัยได้นำเสนอโดยจำแนกออกเป็นตอน ๆ ดังนี้

ตอนที่ 1 การปรับเทียบคะแนนระหว่างแบบสอบ

ตอนที่ 2 หลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 ในกลุ่มสาระการเรียนรู้ คณิตศาสตร์ ช่วงชั้นที่ 3

ตอนที่ 3 งานวิจัยที่เกี่ยวข้องกับเทคนิคการปรับเทียบคะแนนและการศึกษาพัฒนาการ

ตอนที่ 1 การปรับเทียบคะแนนระหว่างแบบสอบ

1.1 แนวคิดเชิงทฤษฎีของการปรับเทียบคะแนนระหว่างแบบสอบ

คำศัพท์ที่เกี่ยวข้องกับคำว่า “**การปรับเทียบคะแนนระหว่างแบบสอบ**” (Test Equating) มีอยู่หลายคำ Mislevy (1992 cited by Johnson & Owen, 1998) และ Linn (1993 cited by Johnson & Owen, 1998) ได้อธิบายให้เห็นความแตกต่างของคำศัพท์ 4 คำ คือ Equating, Calibration, Projection และ Moderation ต่างเป็นรูปแบบหนึ่งของการเชื่อมต่อ (linking) กระบวนการจำเป็นที่จะทำให้การเชื่อมต่อได้รับผลสำเร็จ ขึ้นอยู่กับว่าการประเมินทั้งสองมีความเหมือนกันในด้านเป้าหมาย การครอบคลุมเนื้อหาและคุณสมบัติของการวัดเพียงใด Mislevy และ Linn ได้อธิบายคำทั้ง 4 คำข้างต้นว่ามีความแตกต่างกันในข้อตกลงเบื้องต้นที่กำหนด ซึ่งลดน้อยลงตามลำดับ ข้อกำหนดของการปรับเทียบคะแนน (Equating) จะมีข้อตกลงที่เข้มงวดกว่าและมอดเดอเรชั่น (Moderation) มีข้อตกลงที่ผ่อนปรนมากที่สุด

Mislevy และ Linn ได้กล่าวถึง **การปรับเทียบคะแนน** ว่าเป็นการเชื่อมต่อการประเมินสองการประเมินด้วยแบบสอบที่สร้างจากตารางวิเคราะห์รายละเอียดเดียวกัน ข้อกำหนดคือความเท่าเทียมกันในเชิงโครงสร้าง เช่น ความยาก รูปแบบของคำถาม วิธีการจัดการ จำนวนข้อสอบและความแม่นยำในการวัดแต่ละระดับของคะแนน ภายใต้สภาพการณ์ที่ควบคุม ทำให้ผล

การประเมินสามารถสับเปลี่ยนกันได้ จึงสร้างตารางคะแนนฉบับ X ที่สอดคล้องกับคะแนนฉบับ Y ได้ ดังนั้นทุกๆ ข้อคำถามจึงกำหนดคะแนนจากการประเมิน X ซึ่งกำหนดในวิธีเดียวกันกับคะแนนแปลงจากการประเมิน Y เมื่อพิจารณาเทียบกับ **การปรับ** (Calibration) ซึ่งเป็นการเชื่อมต่อที่เงื่อนไขน้อยกว่าการปรับเทียบคะแนน คือ ถ้าการประเมิน Y ถูกสร้างจากตารางวิเคราะห์รายละเอียดเดียวกัน กับการประเมิน X แต่มีความแม่นยำแตกต่างกันหรือระดับของความยากแตกต่างกัน ในกรณีที่ไม่สามารถเทียบคะแนนได้ แต่ผลของการประเมินทั้งสองสามารถถูกปรับให้เหมาะสมได้ และสำหรับ **โปรเจกชัน** (Projection) และ **มอดเดอเรชั่น** (Moderation) เป็นการเชื่อมต่อมาตรวัดเมื่อข้อสอบวัดในโครงสร้างแตกต่างกัน ในขณะเดียวกันได้มีผู้ให้ความหมายของการปรับเทียบคะแนนไว้หลายท่านด้วยกัน ดังนี้

Guliksen (1995) ได้ให้ความหมายของการปรับเทียบระหว่างแบบสอบว่าเป็นวิธีการปรับเทียบคะแนนที่ได้จากแบบสอบสองฉบับที่วัดเนื้อหาเดียวกันให้เป็นคะแนนสมมูลกัน (Equivalent Scores) สามารถเปรียบเทียบกันได้โดยตรง โดยได้เสนอวิธีการอย่างง่าย ๆ โดยให้ผู้สอบกลุ่มเดียวทำแบบสอบทั้งสองฉบับที่พยายามสร้างให้คู่ขนานกัน แล้วคำนวณคะแนนมาตรฐานของแต่ละฉบับ คะแนนแปลงที่ตรงกับคะแนนมาตรฐานเดียวกัน ถือว่าเป็นคะแนนสมมูลกัน สามารถนำมาเทียบกันได้โดยตรง

Flanagan (1951 cited by Angoff, 1984) ได้ให้นิยามเชิงทฤษฎีของการปรับเทียบคะแนนระหว่างแบบสอบว่าเป็นวิธีการที่ทำคะแนนจากแบบสอบต่างชุดให้มีคุณลักษณะที่ปรับเทียบกันได้โดยแบบสอบทั้งสองฉบับ จะต้องมีการแจกแจงของคะแนนจริงในประชากรเท่ากันหรือมีความเที่ยงเท่ากันเสียก่อน จึงสามารถนำคะแนนดิบมาปรับเทียบกันได้ จากความหมายดังกล่าวเป็นนิยามเชิงทฤษฎี ในทางปฏิบัติ Flanagan ได้เสนอให้ตรวจสอบจากคะแนนเฉลี่ยของแบบสอบที่ได้จากกลุ่มตัวอย่างขนาดใหญ่ ถ้าคะแนนเฉลี่ยจากแบบสอบทั้งสองไม่แตกต่างกันอย่างมีนัยสำคัญ จึงสามารถนำคะแนนดิบมาปรับเทียบกันได้ แต่ถ้าคะแนนเฉลี่ยแตกต่างกันอย่างมีนัยสำคัญ ควรสร้างแบบสอบทั้งสองให้มีคุณสมบัติของความเป็นคู่ขนาน แล้วเลือกวิธีการปรับเทียบคะแนนที่เหมาะสม ซึ่งเสนอไว้ 4 วิธี คือ

1. ใช้ค่าเฉลี่ยโดยคำนวณค่าเฉลี่ยของการแจกแจงคะแนนทั้งสองชุด ถ้าความแตกต่างของค่าเฉลี่ยอยู่ภายในขอบเขตของการแปรผันเชิงสุ่มแล้ว ถือว่าคะแนนทั้งสองชุดเปรียบเทียบกันได้ แต่ถ้ามีความแตกต่างกันอย่างมีนัยสำคัญ ให้ใช้วิธีบวกเข้าหรือลบออกเท่ากับจำนวนที่แตกต่างจากคะแนนชุดที่หนึ่ง เพื่อให้เกิดคะแนนสมมูลกับอีกชุดหนึ่ง

2. ใช้เทคนิคของสมการถดถอย โดยหาค่าประมาณที่ดีที่สุดของคะแนนจากแบบสอบชุดที่หนึ่งซึ่งรู้ค่าของอีกชุดหนึ่ง

3. ใช้คะแนนมาตรฐานที่เท่ากันเป็นตัวปรับเทียบคะแนน

4. ใช้วิธีอิควิเปอร์เซ็นไทล์ โดยคะแนนสมมูลกันเมื่อคะแนนจากแบบสอบต่างฉบับมีตำแหน่งเปอร์เซ็นต์ไทล์เท่ากัน

Angoff (1971) ได้ให้ความหมายของการปรับเทียบคะแนนระหว่างแบบสอบว่าเป็นกระบวนการแปลงคะแนนของแบบสอบฉบับหนึ่งไปสู่ระบบคะแนนของแบบสอบฉบับหนึ่ง ซึ่งวัดสิ่งเดียวกัน เพื่อให้คะแนนแปลงจากแบบสอบทั้งสองฉบับสามารถเทียบเท่ากัน (Equipvalent) และใช้แทนกันได้ (Interchangeable)

Lord (1980 อ้างใน อดิศร ศรีบุญวงศ์, 2545) ได้กล่าวถึงวัตถุประสงค์ของการปรับเทียบคะแนนระหว่างแบบสอบ เพื่อกำหนดความเท่าเทียมกันระหว่างคะแนนดิบจากแบบสอบต่างฉบับ และเนื่องจากวิธีการปรับเทียบคะแนนแบบสอบเป็นวิธีการหาหลักฐานเชิงประจักษ์ จึงจำเป็นต้องกำหนดแบบแผนในการเก็บรวบรวมข้อมูลและกฎเกณฑ์ของการปรับเทียบคะแนนจากแบบสอบหนึ่งไปสู่แบบสอบหนึ่ง Lord ได้กำหนดเงื่อนไขของการปรับเทียบคะแนนซึ่งนักวัดผลหลายคนได้เห็นด้วยในกรณีที่ว่า แบบสอบฉบับ X และฉบับ Y จะสามารถนำมาปรับเทียบกันได้ก็ต่อเมื่อแบบสอบทั้งสองมีคุณสมบัติ 4 ประการ คือ

1. แบบสอบทั้งสองฉบับจะต้องวัดความสามารถเดียวกัน (Same ability) คือ แบบสอบทั้งสองวัดในคุณลักษณะนี้อาจเป็นคุณลักษณะแฝง หรือความสามารถหรือทักษะ อย่างใดอย่างหนึ่งก็ได้

2. มีความเสมอภาค (Equity) คือ เมื่อทุกกลุ่มมีความสามารถเดียวกัน การแจกแจงคะแนนของแบบสอบ Y หลังจากที่มีการปรับคะแนนแล้วจะมีการแจกแจงเหมือนกับการแจกแจงของคะแนนจากแบบสอบ X

3. ประชากรไม่แปรเปลี่ยน (Population invariance) คือ การปรับคะแนนต้องเป็นไปในแนวเดียวกัน ไม่ว่าคะแนนจะมาจากกลุ่มตัวอย่างใดก็ตาม

4. มีความสมมาตร (Symmetry) การปรับคะแนนสามารถแปรเปลี่ยนกลับได้เช่นกัน การแปลงคะแนนจากฉบับ X ไปยังฉบับ Y มีผลเช่นเดียวกันกับฉบับ Y ไปสู่ฉบับ X

Petersen และคณะ (1989) ได้ให้ความหมายของการปรับเทียบคะแนนระหว่างแบบสอบว่าเป็นกระบวนการเชิงประจักษ์สำหรับสร้างระบบความสัมพันธ์ระหว่างคะแนนของแบบสอบต่างฉบับให้สามารถแปลงคะแนนของแบบสอบฉบับหนึ่งไปสู่คะแนนของแบบสอบอีกฉบับหนึ่งได้

ชูศักดิ์ ขัมภลิขิต (2534) ได้ให้ความหมายของการปรับเทียบคะแนนระหว่างแบบสอบไว้ว่า การปรับเทียบคะแนนเป็นกระบวนการที่เกี่ยวข้องกับกิจกรรม 2 ประการ ดังนี้

1. กระบวนการที่ทำให้แบบสอบ 2 ฉบับใด ๆ มีความทัดเทียมกัน หรือเท่ากันในเชิงโครงสร้าง

2. การใช้วิธีการทางสถิติเพื่อปรับ (Adjust) คะแนนที่ได้จากแบบสอบแต่ละฉบับให้อยู่ในสเกลเดียวกันและเทียบกันได้

ศิริชัย กาญจนวาสี (2545) ได้ให้ความหมายของการปรับเทียบคะแนนว่า เป็นกระบวนการแปลงคะแนนของแบบสอบต่างฉบับ ที่มุ่งวัดคุณลักษณะเดียวกันสามารถนำมาเปรียบเทียบกันได้

จากแนวคิดและการให้ความหมายดังกล่าวข้างต้น สรุปได้ว่าการปรับเทียบคะแนนเป็นกระบวนการทางสถิติที่ใช้ในการแปลงระบบคะแนนระหว่างแบบสอบต่างฉบับกัน จากแบบสอบหนึ่ง ให้อยู่บนหน่วยหรือสเกลของแบบสอบฉบับอื่นที่มุ่งวัดคุณลักษณะเดียวกัน เพื่อให้คะแนนสามารถเปรียบเทียบกันได้อย่างมีความหมาย

1.2 เงื่อนไขของการปรับเทียบคะแนน

การปรับเทียบคะแนนระหว่างแบบสอบ สามารถกระทำได้ภายใต้เงื่อนไขที่จำเป็นบางประการ ซึ่งมีผู้เสนอเงื่อนไขต่าง ๆ ดังนี้

Gulliksen (1950) ได้เสนอว่า แบบสอบที่นำมาเทียบคะแนนกันได้ต้องทำเป็นแบบสอบคู่ขนานกัน ซึ่งสร้างโดยการจับคู่ข้อสอบที่เหมือนกันแล้วสุ่มแยกเป็นแบบสอบคู่ขนานสองฉบับ

Levine (1955 อ้างใน ศิริชัย กาญจนวาสี, 2541) ได้เสนอแบบสอบสองฉบับ ที่ต้องการนำมาปรับเทียบคะแนนให้อยู่ในสเกลเดียวกันนั้น ต้องเป็นแบบสอบคู่ขนาน (Parallel Forms) กันในด้าน 1) เนื้อเรื่อง (Subject) 2) โครงสร้าง (Structure) 3) รูปแบบ (Formats) 4) ชนิดของข้อสอบ (Item Type) และ 5) เวลาที่ใช้สอบ

Angoff (1984) ได้เสนอว่า แบบสอบที่จะนำมาเทียบคะแนนกันได้นั้นต้องเป็นแบบสอบที่วัดคุณลักษณะเดียวกัน และใช้วัดกลุ่มผู้สอบที่มีระดับความสามารถใกล้เคียงกัน

Hambleton และ Swaminathan (1985) ได้เสนอข้อกำหนดของการปรับเทียบคะแนน ดังนี้

1. ความเท่าเทียมกัน (Equity) ซึ่ง Lord (1980 อ้างใน ภัทราพร เกษสังข์, 2546) ได้ให้ข้อสังเกตของความเท่าเทียมกันในแบบสอบของการปรับเทียบคะแนน ว่าถ้าการปรับเทียบคะแนนของแบบสอบฉบับ X และ ฉบับ Y มีความยุติธรรมต่อผู้สอบแล้ว จะต้องไม่มีความแตกต่างที่ทุก ๆ ระดับความสามารถไม่ว่าจะสอบด้วยแบบสอบฉบับ X หรือฉบับ Y แบบสอบที่วัดคุณลักษณะ (Traits) หรือความสามารถ (Abilities) ต้องไม่ต่างกัน มีความคู่ขนานกัน ถ้าคะแนนของแบบสอบมีค่าความเที่ยง (Reliable) และระดับความยาก (Diffculty Levels) ไม่เท่ากันแล้วไม่สามารถเทียบกันได้ เนื่องจากคะแนนจริง (True Score) ของแบบสอบมีความสัมพันธ์กันแบบไม่เป็นเส้นตรง (Nonlinear) และทำให้แบบสอบมีความเที่ยงไม่เท่ากัน ณ ระดับความสามารถที่ต่างกัน และการแจกแจงความถี่แบบมีเงื่อนไข (The Conditionnal Frequency Distribution) ณ ระดับความสามารถ θ, $f\ [\ x \mid \theta]$ ของคะแนน θ จากแบบสอบฉบับ X จะต้องเป็นการแจกแจงความถี่แบบมีเงื่อนไข เช่นเดียวกับการแปลงคะแนน x (y), $f[\ x\ (y) \mid \theta]$ เมื่อ x(y) เป็นฟังก์ชันแบบหนึ่งต่อหนึ่งของ y

2. ความไม่แปรเปลี่ยนไปตามกลุ่ม (Invariance Across Groups) คือ คะแนนของกลุ่มต่างๆ ที่ผ่านการปรับเทียบคะแนนแล้ว จะมีค่าไม่แตกต่างกันไม่ว่าจะปรับเทียบคะแนนจากกลุ่มตัวอย่างใด

3. ความสมมาตร (Symmetry) คือ ผลการปรับเทียบคะแนนจากแบบสอบฉบับ X ไปสู่ฉบับ Y หรือจากฉบับ Y ไปสู่ฉบับ X จะต้องเหมือนกัน

4. ความเป็นมิติเดียวกันของแบบสอบ (Unidimentional of the tests) คือแบบสอบทั้งสองฉบับจะต้องวัดคุณลักษณะเดียวกัน (The Same Trait) (คุณลักษณะแฝง, ความสามารถหรือทักษะ)

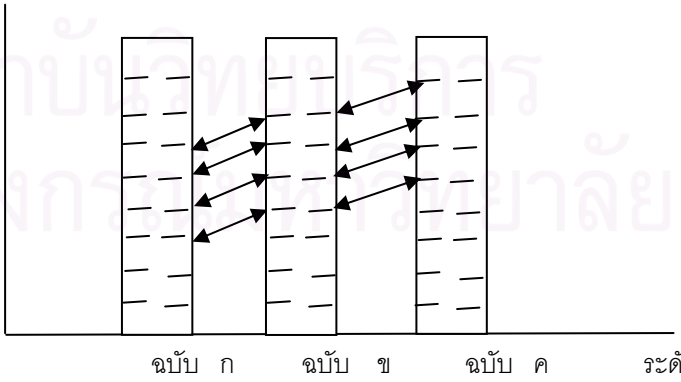
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

1.3 ประเภทของการปรับเทียบคะแนน

ในการทดสอบที่มีการใช้เทคนิคการปรับเทียบคะแนนระหว่างแบบสอบตามเงื่อนไขและจุดมุ่งหมายของการทดสอบ จำแนกได้เป็น 2 สถานการณ์ (Hambleton & Swaminathan, 1985) ซึ่งต้องใช้รูปแบบการปรับเทียบคะแนน (Equating Form) 2 รูปแบบ ดังนี้

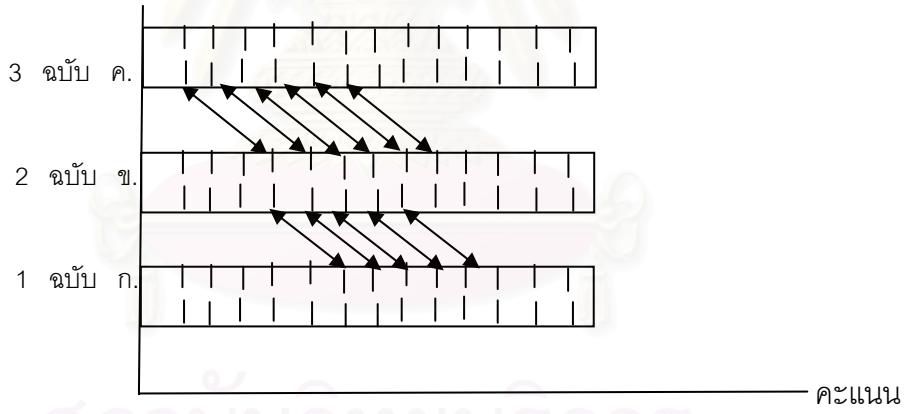
1. **การปรับเทียบคะแนนตามแนวนอน** (Horizontal Equating) เป็นการปรับเทียบคะแนนระหว่างแบบสอบต่างฉบับกัน เมื่อแต่ละฉบับมุ่งวัดคุณลักษณะเดียวกัน มีระดับความยากง่ายใกล้เคียงกันและกลุ่มผู้สอบมีการแจกแจงความสามารถ อยู่ในประชากรเดียวกันหรือมีความสามารถใกล้เคียงกัน เป็นเทคนิคที่เหมาะสมในสถานการณ์ที่มีความจำเป็นต้องสร้างแบบสอบเนื้อหาหลาย ๆ ฉบับเพื่อให้การทดสอบเกิดความยุติธรรมและป้องกันความลับของข้อสอบเมื่อใช้ต่างเวลากันสำหรับผู้สอบขนาดใหญ่ ปรับเทียบว่าคะแนนที่ได้จากแบบสอบฉบับหนึ่งเทียบเป็นเท่าไรของอีกฉบับหนึ่ง ซึ่งวัดในระดับความสามารถเดียวกัน นำไปใช้ในการปรับเทียบคะแนนระหว่างแบบสอบวิชาต่างฉบับกัน เช่น การปรับเทียบคะแนนระหว่างแบบสอบวิชาคณิตศาสตร์ 2 ฉบับ สำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 เป็นต้น

แบบสอบต่างฉบับที่ต้องการนำมาปรับเทียบคะแนนกันตามแนวนอนนั้น ควรเป็นแบบสอบที่พยายามสร้างหรือมุ่งให้คู่ขนานกัน (Alternate Forms) นอกจากนี้ การแจกแจงความสามารถของผู้สอบที่ทำการสอบด้วยแบบสอบแต่ละฉบับควรมีการแจกแจงที่ประมาณได้ว่ามีระดับความสามารถที่ใกล้เคียงกัน



แผนภาพที่ 2.1 การปรับเทียบคะแนนตามแนวนอน (Horizontal Equating)(ศิริชัย กาญจนวาสี, 2545)

2. การปรับเทียบคะแนนตามแนวตั้ง (Vertical Equating) เป็นการปรับเทียบระหว่างแบบสอบต่างฉบับกัน เมื่อแบบสอบต่างฉบับมุ่งวัดลักษณะเดียวกัน แต่มีระดับความยากแตกต่างกันและกลุ่มผู้สอบมีการแจกแจงความสามารถอยู่คนละประชากรหรือมีความสามารถแตกต่างกัน เหมาะกับสถานการณ์ที่มีความจำเป็นต้องสร้างแบบสอบต่างฉบับวัดเนื้อหาเดียวกันมุ่งวัดความสามารถของผู้สอบที่ต่างระดับกัน เพื่อปรับเทียบว่าคะแนนที่สอบได้จากฉบับหนึ่งเทียบเป็นเท่าไรของฉบับอื่นที่วัดต่างระดับกัน นำไปใช้ในการปรับเทียบคะแนนระหว่างแบบสอบวิชาเดียวกันแต่ต่างระดับชั้น เช่น การปรับเทียบคะแนนระหว่างแบบสอบวิชาคณิตศาสตร์ ระดับชั้น ม.1-ม.3 แบบสอบต่างฉบับที่วัดเนื้อเรื่องเดียวกัน แต่ต่างระดับกันที่ต้องการนำมาปรับเทียบคะแนนกันนั้น แบบสอบแต่ละฉบับจึงมีระดับความยากแตกต่างกันไป รวมทั้งกลุ่มผู้สอบแต่ละฉบับมีการแจกแจงความสามารถอยู่ต่างประชากร หรือมีการแจกแจงความสามารถที่อยู่ในระดับที่แตกต่างกัน ดังนั้นการปรับเทียบคะแนนตามแนวตั้ง จึงมีความสลับซับซ้อนกว่าการปรับเทียบคะแนนในแนวนอนทั้งในแง่ทฤษฎีและการปฏิบัติ



แผนภาพที่ 2.2 การปรับเทียบคะแนนตามแนวตั้ง (Vertical Equating) (ศิริชัย กาญจนวาสี, 2545)

1.4 ขั้นตอนการดำเนินการปรับเทียบคะแนน

ขั้นตอนที่สำคัญของการปรับเทียบคะแนนระหว่างแบบสอบ มีดังต่อไปนี้ (Kolen & Brennan, 1995)

1. กำหนดจุดมุ่งหมายของการปรับเทียบคะแนนว่าจะนำผลจากการปรับเทียบเพื่อดูพัฒนาการเปลี่ยนแปลงทางการศึกษา หรือเพื่อเทียบความสามารถผู้สอบเป็นข้อมูลใช้ตัดสินผลร่วมกัน ใช้แทนกันได้

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

3. เลือกวิธีการเก็บรวบรวมข้อมูล มีรูปแบบกลุ่มสุ่ม (Random Group Design) โดยสุ่มกลุ่มตัวอย่างจากประชากรเดียวกัน แต่ละกลุ่มทำแบบสอบคนละชุด รูปแบบผู้สอบกลุ่มเดียว (Single-Group Design) ผู้สอบกลุ่มเดียวทำแบบสอบทั้งสองชุด รูปแบบกลุ่มเดียวที่ได้รับการจัดให้สมดุล (Single-Group Design with Counterbalancing) แบ่งผู้สอบแต่ละกลุ่มแบบสอบแต่ละชุดเป็น 3 ส่วน ให้ผู้สอบกลุ่มย่อยแรก สอบแบบสอบชุดที่ 1 ตอนแรก ตามด้วยแบบสอบชุดที่ 2 ตอนหลัง และผู้สอบกลุ่มย่อยที่ 2 ทำแบบสอบชุดที่ 2 ตอนแรก ตามด้วยแบบสอบชุดที่ 1 ตอนหลัง และรูปแบบผู้สอบกลุ่มไม่เท่าเทียมกันโดยใช้แบบสอบร่วม (Common-Item Nonequivalent Group Design) ผู้สอบต่างกลุ่มประชากรทำแบบสอบคนละชุด และผู้สอบทุกคนทำแบบสอบร่วมที่อาจจะแทรกภายในแบบสอบทั้งสองฉบับ (Internal Common Item) หรือแยกออกจากแบบสอบ (External Common Item)

4. เก็บรวบรวมข้อมูลตามรูปแบบที่กำหนดไว้

5. เลือกนิยามเชิงปฏิบัติการของการปรับเทียบคะแนน เพื่อจะตัดสินใจว่าจะใช้วิธีการปรับเทียบคะแนนเชิงเส้นตรง (Linear Equating Method) หรือวิธีการปรับเทียบที่ไม่เป็นเส้นตรง (Nonlinear Equating Method)

6. เลือกสถิติประมาณค่าที่ใช้วิเคราะห์ เลือกให้สอดคล้องกับนิยามเชิงปฏิบัติการที่กำหนด มีวิธีการปรับจากค่าเฉลี่ย (Mean Equating) โดยพิจารณาคะแนนสมมูลกันเมื่อคะแนนจากแบบสอบต่างฉบับเบี่ยงเบนไปจากคะแนนเฉลี่ยเท่ากัน วิธีปรับเทียบเชิงเส้นตรง (Linear Equating) พิจารณาคะแนนสมมูลกันเมื่อคะแนนจากแบบสอบต่างฉบับมีคะแนนมาตรฐานเท่ากัน วิธีการปรับเทียบอิควิเปอร์เซ็นไทล์ (Equipercentile Equating) ที่คะแนนสมมูลกันเมื่อแบบสอบต่างฉบับมีตำแหน่งเปอร์เซ็นไทล์เท่ากัน และวิธีการปรับเทียบโดยใช้สมการถดถอย (Regression Equating) เป็นการสร้างสมการทำนายคะแนนจากแบบสอบชุดหนึ่งไปยังอีกชุดหนึ่ง หรือได้จากคะแนนสมมูลกันเมื่อคะแนนของแบบสอบแต่ละฉบับทำนายคะแนนเกณฑ์ได้เท่ากัน ทั้ง 4 วิธีนี้เป็นวิธีการปรับเทียบคะแนนตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) ส่วนวิธีการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ (Item Response Theory) เป็นการหาสัมประสิทธิ์การปรับเทียบ หรือค่าความชัน (Slope) และ ค่าคงที่ ของฟังก์ชันเชิงเส้นตรงที่เป็นความสัมพันธ์ของการปรับเทียบคะแนน วิธีการหาสัมประสิทธิ์การปรับเทียบ มีวิธีใช้ค่าเฉลี่ยของ

ค่าอำนาจจำแนกและค่าเฉลี่ยของค่าความยากของข้อสอบ (Mean and σ Method) วิธีการทำให้ความแตกต่างระหว่างคะแนนจริงเดิมกับคะแนนจริงที่ปรับแล้วมีค่าน้อยที่สุด โดยใช้สถิติ F-test (Characteristic Curve Method) และวิธีทำให้ความแตกต่างระหว่างพารามิเตอร์ของข้อสอบเดิมกับที่ปรับแล้วมีค่าน้อยที่สุด โดยใช้สถิติ χ^2 –test (Minimum χ^2)

7. ประเมินผลการปรับเทียบคะแนน มีเกณฑ์ที่ใช้ดังนี้ (Hariss & Crouse, 1993) ความเสมอภาค (Weak Equity) ของ Divgi และ Yen ที่พิจารณาจากความเท่าเทียมกันของการแจกแจงตามเงื่อนไขของคะแนนที่ได้จากแบบสอบต่างฉบับหลังจากการปรับเทียบแล้ว ดัชนีสำหรับการเปลี่ยนแปลงคะแนน (Indices) ของ Angoff ความคลาดเคลื่อนมาตรฐาน (Standard Error) ของ Angoff เป็นการวิเคราะห์เพื่อประมาณความสามารถของการปรับเทียบจากการสุ่มตัวอย่าง ข้อมูลที่จำลองขึ้น (Generated Data) ของ Lord เป็นการปรับเทียบคะแนนกลับสู่แบบสอบเดิมโดยตรง หรือปรับผ่านแบบสอบอื่นก่อนปรับกลับสู่แบบสอบเดิม กลุ่มตัวอย่างขนาดใหญ่ (Large Sample) ของ Angoff เป็นการใช้กลุ่มตัวอย่างขนาดใหญ่ในการปรับเทียบคะแนนซึ่งคล้ายกับการปรับเทียบคะแนนจากประชากร และใช้เปรียบเทียบกับผลการปรับเทียบกับคะแนนที่มีขนาดกลุ่มตัวอย่างน้อยกว่า ความคงเส้นคงวา (Consistency) เป็นการประเมินผลการปรับเทียบคะแนนข้ามวิธี เพื่อหาความคงเส้นคงวา (Stability) ของ Angoff เป็นการปรับเทียบคะแนนซ้ำเพื่อตรวจความคงที่ของผลการปรับเทียบคะแนน

เงื่อนไขที่ส่งเสริมให้การปรับเทียบคะแนนเป็นที่น่าพอใจมีรายละเอียด คือ (Kolen & Brennan, 1995)

1. ลักษณะทั่วไป

1.1 มีการระบุเป้าหมายของการที่นำคะแนนมาปรับเทียบอย่างชัดเจน

1.2 มีการออกแบบเก็บรวบรวมข้อมูล การวางแผนเชื่อมโยงการปรับเทียบคะแนน วิธีการทางสถิติที่ใช้ และการเลือกผลที่ได้จากการปรับเทียบหลายวิธี ให้มีความเหมาะสมกับความสำเร็จของการปรับเทียบคะแนน

1.3 มีการติดตามและควบคุมคุณภาพของกระบวนการปรับเทียบคะแนนอย่างเพียงพอ

2. การพัฒนาแบบสอบสำหรับการเก็บรวบรวมข้อมูลทุกรูปแบบ

2.1 มีการกำหนดเนื้อหาแบบสอบและกำหนดแบบแผนทางสถิติข้อสอบเป็นอย่างดี มีความคงที่

2.2 มีการสร้างแบบสอบเพื่อนำมาปรับเทียบพิจารณาจากสถิติข้อสอบ เช่น ค่าความยาก ค่าอำนาจจำแนก จากการสอบครั้งก่อน

2.3 มีการใช้ข้อสอบที่มีความยาวพอเหมาะ อย่างน้อยมีจำนวน 30 ข้อ หรือมากกว่า ตัวเฉลยสำหรับตรวจให้คะแนนมีความคงที่เมื่อใช้กับแบบสอบแบบเลือกตอบ

3. การพัฒนาแบบสอบสำหรับผู้สอบกลุ่มไม่เท่าเทียมกันใช้แบบสอบร่วม

3.1 ข้อสอบร่วมต้องเป็นตัวแทนของแบบสอบที่ใช้ในการปรับเทียบทั้งคุณลักษณะของเนื้อหาและค่าสถิติของข้อสอบ

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

3.3 ข้อสอบร่วมแต่ละข้อสร้างให้มีลักษณะที่ใกล้เคียงกับแบบสอบที่นำมาปรับเทียบ ทั้งตัวคำถามและตัวเลือก

4. กลุ่มตัวอย่าง

4.1 กลุ่มตัวอย่างมีลักษณะเป็นตัวแทน

4.2 กลุ่มตัวอย่างมีความคงที่

4.3 กลุ่มตัวอย่างมีขนาดใหญ่เพียงพอ

4.4 กลุ่มตัวอย่าง 2 กลุ่ม สำหรับกลุ่มผู้สอบไม่เท่าเทียมกันใช้แบบสอบร่วม มีความสามารถไม่แตกต่างกันมากนัก

5. การบริหารการสอบ

5.1 ดำเนินการทดสอบอย่างปลอดภัย

5.2 ดำเนินการสอบในแต่ละครั้งอย่างระมัดระวัง และจัดสภาพการสอบให้เหมือนกัน

6. หลักสูตร เนื้อหาการปฏิบัติ หรือสาระที่นำมาใช้ในการปรับเทียบคะแนนมีความคงที่

ผลของการปรับเทียบคะแนนมีความเชื่อถือได้ หรือมีความถูกต้องแม่นยำขึ้นอยู่กับการเลือกสถิติที่ใช้ในการปรับเทียบคะแนนให้มีความเหมาะสมกับแบบแผนการออกข้อสอบ การเก็บรวบรวมข้อมูล และการดูแลแบบสอบให้เป็นไปอย่างมีมาตรฐาน

1.5 การออกแบบวิธีการเก็บรวบรวมข้อมูล

ในการปรับเทียบคะแนนระหว่างแบบสอบแต่ละครั้งจำเป็นต้องมีแบบแผนในการเก็บรวบรวมข้อมูล สามารถจำแนกเป็นแบบแผนการเก็บรวบรวมข้อมูล 4 รูปแบบ (Lord, 1975b; Angoff, 1984; Petersen & others, 1989 อ้างใน ศิริชัย กาญจนวาสี, 2541) ดังนี้

1. **รูปแบบผู้สอบกลุ่มเดียว** (Single-Group Design)

1.1 รูปแบบผู้สอบกลุ่มเดียวที่ไม่ได้รับการจัดให้สมดุล (Uncounterbalanced Design)

รูปแบบนี้ใช้ผู้สอบกลุ่มเดียวทำหน้าที่เป็นผู้สอบร่วมซึ่งถือเป็นรูปแบบที่ง่ายที่สุด ซึ่งมีวิธีการเก็บรวบรวมข้อมูลโดยใช้กลุ่มตัวอย่างผู้สอบกลุ่มเดียว แต่ละคนทำแบบสอบทั้งสองฉบับ โดยให้ผู้สอบทำแบบสอบฉบับหนึ่งแล้วตามด้วยแบบสอบอีกฉบับหนึ่ง การให้ผู้สอบทำแบบสอบทั้งสองฉบับ ระดับความสามารถของผู้สอบของแบบสอบทั้งสองฉบับเท่ากัน จึงไม่เป็นปัญหาแทรกซ้อน รวมทั้งไม่มีปัญหาในเรื่องความยากของแบบสอบต่างฉบับ คะแนนของแบบสอบจึงสามารถปรับเทียบกันได้บนพื้นฐานความสามารถที่เท่ากันจากคนกลุ่มเดียวกัน

1.2 รูปแบบผู้สอบกลุ่มเดียวที่ได้รับการจัดให้สมดุล (Counterbalanced Design)

รูปแบบนี้ปรับมาจากรูปแบบแรก เพื่อขจัดผลของลำดับการทดสอบก่อน-หลัง โดยการสุ่มผู้สอบแยกเป็นสองกลุ่มย่อย แต่ละกลุ่มย่อยได้รับการทดสอบด้วยแบบสอบทั้งสองฉบับในลักษณะที่กลุ่มย่อยหนึ่งให้ทำแบบสอบฉบับที่ 1 แล้วตามด้วยแบบสอบฉบับที่ 2 ส่วนอีกกลุ่มย่อยหนึ่งให้ทำแบบสอบฉบับที่ 2 แล้วตามด้วยแบบสอบฉบับที่ 1 สลับกันเพื่อให้เกิดความสมดุล

การใช้รูปแบบนี้ กลุ่มผู้สอบจึงได้รับอิทธิพลจากลำดับการทดสอบ การเรียนรู้ การฝึกฝนและความเมื่อยล้าที่สมดุลกันในทั้งสองกลุ่มย่อย

2. **รูปแบบผู้สอบกลุ่มสมมูล** (Equivalent-Group Design)

รูปแบบนี้เป็นการจัดกลุ่มผู้สอบให้มีความคล้ายคลึงกันมากที่สุดเท่าที่จะทำได้ (Equivalent Group) โดยใช้กลุ่มเท่าเทียมกัน แล้วให้ผู้สอบในแต่ละกลุ่มทำแบบสอบฉบับเดียว ถึงแม้การจัดกลุ่มจะไม่เหมือนกันทุกประการ (Identical Group) แต่ควรมีลักษณะทั่วไปที่คล้ายกันมากที่สุด กลุ่มดังกล่าวอาจได้มาโดยการสุ่ม (Random) ข้อดีของรูปแบบนี้คือ แต่ละกลุ่มทำแบบสอบฉบับเดียว

รูปแบบนี้สามารถหลีกเลี่ยงปัญหาเรื่องการเรียนรู้ การฝึกฝน และความเมื่อยล้าจากการทำแบบสอบฉบับแรก แต่ปัญหาที่ตามมาคือ กลุ่มที่ใช้ไม่เหมือนกันอาจมีการแจกแจงความสามารถที่ต่างกัน และไม่มีข้อมูลที่นำมาใช้ปรับความแตกต่างของกลุ่ม ความแตกต่างที่เกิดขึ้นแม้เพียงเล็กน้อยย่อมส่งผลกระทบต่อความลำเอียงในการปรับเทียบคะแนนได้ วิธีการที่จะช่วยลดความแตกต่างระหว่างกลุ่มได้โดยการใช้กลุ่มตัวอย่างสุ่มขนาดใหญ่

3. รูปแบบผู้สอบต่างกลุ่มโดยใช้แบบสอบร่วม (Anchor-Test Design)

3.1 รูปแบบผู้สอบกลุ่มสุ่มโดยใช้แบบสอบร่วม (Anchor-Test Random Group Design)

รูปแบบนี้เป็นการใช้ผู้สอบที่ได้จากการสุ่ม แล้วให้ผู้สอบทำแบบสอบฉบับเดียว โดยแบบสอบแต่ละฉบับมีแบบสอบร่วมจำนวนหนึ่ง (Common or Anchor Items) ซึ่งเรียกว่าแบบสอบร่วม (Anchor Test) การใช้แบบสอบร่วมนี้สามารถกระทำได้ 2 ลักษณะดังนี้

3.1.1 แบบสอบร่วมภายใน (Internal Anchor Test) เป็นแบบสอบที่มีข้อสอบร่วมที่จัดรวมไว้อยู่ในแบบสอบที่ต้องการนำมาปรับเทียบคะแนน

3.1.2 แบบสอบร่วมภายนอก (External Anchor Test) เป็นแบบสอบที่มีข้อสอบร่วมที่ถูกจัดแยกออกมาต่างหากจากตัวแบบสอบ ที่ต้องการนำมาปรับเทียบคะแนน ซึ่งกลุ่มผู้สอบจะต้องทำและจับเวลาแยกออกจากตัวแบบสอบที่ต้องการนำมาปรับเทียบคะแนน

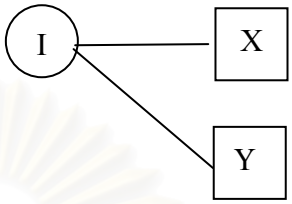
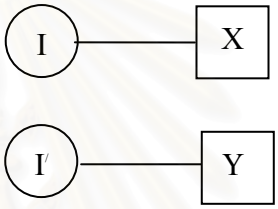
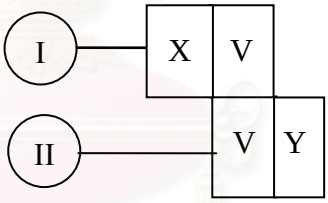
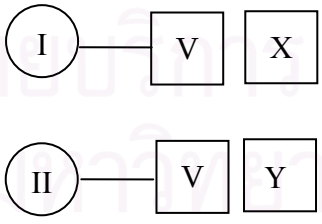
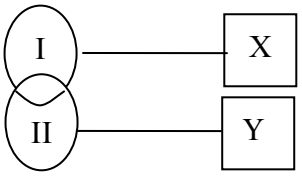
3.2 รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกันโดยใช้แบบสอบร่วม (Anchor-Test Nonequivalent Group Design)

รูปแบบนี้เป็นการใช้ผู้สอบซึ่งลักษณะไม่เท่าเทียมกัน โดยให้ผู้สอบแต่ละกลุ่มทำแบบสอบฉบับเดียว กลุ่มละฉบับ ซึ่งแบบสอบแต่ละฉบับมีแบบสอบร่วมจำนวนหนึ่ง (Common or Anchor Items) ซึ่งเรียกว่า แบบสอบร่วม (Anchor Test) ซึ่งอาจใช้แบบสอบร่วมภายในหรือแบบสอบร่วมภายนอกก็ได้ ในการทดสอบบางสถานการณ์ อาจมีความจำเป็นต้องใช้แบบสอบร่วมกับกลุ่มผู้สอบต่างประชากร เช่น กลุ่มผู้สอบต่างโปรแกรม เวลา หรือระดับ เป็นต้น ซึ่งคะแนนจากแบบสอบร่วมที่ใช้เป็นกลไกสำคัญสำหรับการปรับเทียบคะแนนระหว่างแบบสอบ

4. รูปแบบผู้สอบต่างกลุ่มโดยมีผู้สอบร่วม (Common-Person Design)

รูปแบบนี้เป็นการใช้ผู้สอบต่างกลุ่มกันทำแบบสอบต่างฉบับ แต่มีจำนวนผู้สอบส่วนหนึ่งจากทั้งสองกลุ่มทำแบบสอบทั้งสองฉบับ ซึ่งเป็นรูปแบบของการใช้ผู้สอบร่วม

ตารางที่ 2.1 รูปแบบการเก็บรวบรวมข้อมูลสำหรับการปรับเทียบคะแนน (ศิริชัย กาญจนวาสี, 2545)

การออกแบบ (Designs)	กลุ่มผู้สอบ แบบสอบ	ปัจจัยเกี่ยวข้อง
1. การใช้ผู้สอบกลุ่มเดียว (Single-Group Design)		• ความเมื่อยล้า • ลำดับการทำแบบสอบ • การเรียนรู้จากแบบสอบ
2. รูปแบบผู้สอบกลุ่มสมมูล (Equivalent-Group Design)		• การสุ่ม • ความทัดเทียม
3. การใช้แบบสอบร่วม (Anchor-Test Design) 3.1 แบบสอบร่วมภายใน (Internal Anchor Test)		• ความยาวของ V • ความสัมพันธ์ระหว่าง V กับ X และ Y
3.2 แบบสอบร่วมภายนอก (External Anchor Test)		
4. การใช้ผู้สอบร่วม (Common-Person Design)		• ความเมื่อยล้าของกลุ่มผู้สอบ • ลำดับการทำแบบสอบ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

1.6 รูปแบบของการปรับเทียบคะแนนตามทฤษฎีการวัดแบบดั้งเดิม

การออกแบบวิธีการปรับเทียบคะแนนระหว่างแบบสอบเป็นการกำหนดเงื่อนไข และกฎเกณฑ์การแปลงคะแนน (Transformation) จากแบบสอบต่างชุดที่ต้องการนำมาปรับเทียบกันให้เปรียบเทียบกันได้อย่างมีความหมาย วิธีการที่ใช้ในการปรับเทียบคะแนนตามทฤษฎีการวัดแบบดั้งเดิม จำแนกเป็น 4 วิธี ดังนี้

1. การปรับเทียบอิควิเปอร์เซ็นไทล์ (Equipercentile Equating)

การปรับเทียบคะแนนแบบอิควิเปอร์เซ็นไทล์ใช้นิยามของการปรับเทียบว่าคะแนนจากแบบสอบ X และแบบสอบ Y จะมีความเท่าเทียมกันก็ต่อเมื่อคะแนนทั้งสองอยู่ที่ตำแหน่งเปอร์เซ็นไทล์เดียวกัน และลักษณะการแจกแจงของประชากรผู้สอบต้องมีลักษณะเหมือนกัน วิธีอิควิเปอร์เซ็นไทล์เป็นการสร้างคะแนนสมมูลระหว่างแบบสอบบนพื้นฐานความเท่าเทียมกันของค่าโมเมนต์ทั้ง 3 ทางสถิติ ซึ่งได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และการแจกแจงของคะแนน จึงเป็นกระบวนการแปลงคะแนนแบบไม่ใช่เส้นตรง (Nonlinear Transformation) ยังผลให้คะแนนระหว่างคะแนนดิบและคะแนนแปลงไม่เป็นเส้นตรง

คะแนน X_i และ Y_i ถือเป็นคะแนนสมมูลกัน เมื่อ $X_{PR_k} = Y_{PR_k}$

โดย X_{PR_k} แทน คะแนนจากแบบสอบฉบับ X ตรงตำแหน่งเปอร์เซ็นไทล์ที่ k

Y_{PR_k} แทน คะแนนจากแบบสอบฉบับ Y ตรงตำแหน่งเปอร์เซ็นไทล์ที่ k

และ
$$X_{PR_k} = Lo + \left[\frac{\frac{kn}{100} - F_b}{f}\right](C)$$

เมื่อ Lo แทน คะแนนขีดจำกัดล่างที่แท้จริงของคะแนนในชั้นเปอร์เซ็นไทล์ที่ k

n แทน จำนวนสมาชิกในกลุ่มตัวอย่าง

f แทน ความถี่ของคะแนนในชั้นเปอร์เซ็นไทล์ที่ k

C แทน อันตรภาคชั้น

สำหรับการวิเคราะห์ในการปรับเทียบคะแนนอิควิเปอร์เซ็นไทล์ มีขั้นตอนดังนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

1. สร้างกราฟตำแหน่งเปอร์เซ็นไทล์ของคะแนนที่ได้จากแบบสอบแต่ละฉบับลงบนแกนเดียวกัน

2. หาจุดที่เท่าเทียมกันของกราฟทั้งสอง โดยพิจารณาจากตำแหน่งเปอร์เซ็นไทล์เดียวกันจะตกกับคะแนน x เท่าใด คะแนน y เท่าใด

3. ทำกราฟแสดงความสัมพันธ์ของความเท่าเทียมกันของคะแนน โดยแกนหนึ่งเป็นคะแนนจากแบบสอบชุด X ส่วนอีกแกนหนึ่งเป็นคะแนนจากแบบสอบชุด Y ที่มีตำแหน่งเปอร์เซ็นไทล์เดียวกัน

ข้อเสนอแนะสำหรับการปรับเทียบอิควิเปอร์เซ็นไทล์ พอสรุปได้ดังนี้ (Angoff, 1984)

1. แบบสอบที่นำมาปรับเทียบคะแนนกัน ควรมีความเที่ยงใกล้เคียงกัน ถ้าต่างกันมาก จะทำให้การปรับคะแนนขาดความคงที่

2. การสร้างกราฟเปอร์เซ็นไทล์ เพื่อปรับเทียบคะแนนระหว่างแบบสอบควรสร้างด้วยความระมัดระวังและไม่ลำเอียง

3. การปรับเทียบคะแนนควรใช้กลุ่มตัวอย่างที่มีขนาดใหญ่ ซึ่งเป็นกลุ่มที่มีความสามารถค่อนข้างกระจาย และกระจายพอ ๆ กัน ถ้าใช้กลุ่มตัวอย่างที่มีขนาดเล็กทำให้คะแนนมีความไวต่อความแปรปรวนเชิงสุ่ม ผลการปรับเทียบคะแนนสามารถแปรผันไปตามกลุ่มผู้สอบได้

4. การปรับเทียบคะแนนไม่ควรอยู่นอกพิสัยของคะแนนสังเกตได้ เพราะจะมีความคลาดเคลื่อนสูง

2. วิธีการปรับเทียบเชิงเส้นตรง

วิธีการปรับเทียบเชิงเส้นตรงเป็นวิธีที่ยึดถือนิยามของการปรับเทียบคะแนนว่า คะแนนจากแบบสอบสองฉบับ X และ Y เมื่อแบบสอบ X และ Y วัดสิ่งเดียวกันและมีความเที่ยงใกล้เคียงกัน จะถือว่าคะแนนสมมูลกัน (Equivalent Score) เมื่อคะแนนของแต่ละฉบับมีคะแนนมาตรฐาน (Standard Score) เท่ากัน

คะแนน X_i และ Y_i ถือเป็นคะแนนสมมูลกัน เมื่อ $X_{Zk} = Y_{Zk}$

โดย X_{Zk} แทน คะแนนจากแบบสอบฉบับ X ตรงตำแหน่งมาตรฐาน k

Y_{Zk} แทน คะแนนจากแบบสอบฉบับ Y ตรงตำแหน่งมาตรฐาน k

คะแนน ณ ตำแหน่ง Z_k ของแบบสอบ X และ Y

$$\frac{X_i - \bar{x}}{s_x} = \frac{Y_i - \bar{y}}{s_y}$$

สามารถเขียนเป็นสมการแปลงคะแนนได้ดังนี้

$$\left(X_i - x\right)\frac{S_y}{S_x} = Y_i - y$$

$$\therefore\ T_{(x_i)} = Y_i = y + \left(X_i - x\right)\frac{S_y}{S_x} = Y + S_y Z_x$$

$$\text{หรือ}\ T_{(x_i)} = \left[Y - \left\langle \frac{S_y}{S_x} \right\rangle x \right] + \left\langle \frac{S_y}{S_x} \right\rangle X_i$$

$$\therefore\ T_{(x_i)} = B_{YX} + A_{YX} X_i$$

เมื่อ $T_{(Xi)}$ แทน คะแนนแปลงของ X_i ที่มีความสมมูลกับ Y_i

A_{YX} แทน ความชัน (Slope) = $\dfrac{S_y}{S_x}$

B_{yx} แทน ค่าคงที่ (Constant) = $\bar{Y} - A_{yx}\bar{x}$

จากสูตรจะเห็นได้ว่า วิธีปรับเทียบคะแนนเชิงเส้นตรง เป็นการแปลงคะแนนจากแบบสอบ X ให้สมมูลกับ Y โดยให้คะแนนแปลงของ X มีค่าเฉลี่ยเป็น $\bar{Y}$ และส่วนเบี่ยงเบนมาตรฐานเป็น S_y นั่นเอง

วิธีการปรับเทียบคะแนนเชิงเส้นตรงนั้นมีด้วยกันหลายเทคนิควิธี อาจกล่าวได้ว่าวิธีการปรับเทียบคะแนนเชิงเส้นตรงเป็นกรณีเฉพาะของวิธีปรับเทียบคะแนนอีควิเปอร์เซ็นไทล์ เนื่องจากว่าถ้าการแจกแจงคะแนนจากแบบสอบ X และ Y เหมือนกัน วิธีอิควิเปอร์เซ็นไทล์ให้ผลการปรับเทียบคะแนนระหว่างแบบสอบเป็นเส้นตรง ซึ่งเหมือนกับวิธีการปรับเทียบเชิงเส้นตรง แต่

ถ้าการแจกแจงคะแนนจากแบบสอบ X และ Y แตกต่างกัน และแบบสอบทั้งสองมีความยากแตกต่างกัน ความสัมพันธ์ระหว่างแบบสอบจะมีลักษณะเป็นเส้นโค้ง วิธีอิควิเปอร์เซ็นไทล์น่าจะให้ผลการปรับเทียบที่ดีกว่า วิธีการปรับเทียบเชิงเส้นตรงไม่ค่อยเหมาะสมสำหรับการปรับเทียบคะแนนตามแนวตั้ง (Vertical Equating) เนื่องจากการปรับเทียบดังกล่าวมีกลุ่มตัวอย่างที่ระดับความสามารถไม่เท่ากัน อีกทั้งการแจกแจงคะแนนอาจมีรูปทรงที่ต่างกันได้ทั้งในแง่ความโด่ง (Kurtosis) และความเบ้ (Skewness)

ถึงแม้ว่าวิธีการเชิงเส้นตรงไม่เหมาะสำหรับการปรับเทียบคะแนนตามแนวตั้ง (Vertical Equating) แต่จากการศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องพบว่าวิธีการเชิงเส้นตรงมีหลายเทคนิควิธีที่เหมาะสำหรับการปรับเทียบตามแนวตั้ง (Vertical Equating) เช่น เทคนิควิธีของ Thurstone (1925) โดยการปรับเทียบคะแนนแบบ Thurstone สามารถปรับเทียบคะแนนของกลุ่มประชากรที่มีความสามารถแตกต่างกันได้ดี

นอกจากนี้ ยังสามารถนำวิธีการเชิงเส้นตรงมาช่วยแก้ปัญหาที่เกิดจากการปรับเทียบตามแนวตั้ง (Vertical Equating) ตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ เนื่องจากการปรับเทียบความสามารถของผู้สอบที่แตกต่างกันมาก ๆ มักจะเกิดความคลาดเคลื่อนของการปรับเทียบคะแนน โดยเฉพาะที่คะแนนส่วนปลายของข้อมูลปัญหา ซึ่งเป็นปัญหาที่เรียกว่า “Out-of-Level Effect” (O’Brien and Tohn, 1984; Bielinksi, Thurlow & Scottt, 2000; Leung, 2003) โดยใช้วิธีการเชิงเส้นตรงที่เรียกว่า “วิธีกำลังสองน้อยสุด” (Least Square Method) ซึ่งเป็นวิธีการที่ไม่ยุ่งยาก โดยสถานการณ์ที่นำมาใช้สำหรับการวิจัย สามารถแสดงได้ดังสมการ (Glass & Hopkins, 1996)

$$Y_i = bX_i + c$$

เมื่อ $b_{Y.X} = b = r\left(\frac{s_y}{s_x}\right)$

และ $c = \bar{Y} - b\bar{X}$

เมื่อ Y_i แทน คะแนนแปลงของ X_i ที่สมมูลกับ Y_i

b แทน ความชัน (Slope) = $r\left(\frac{s_y}{s_x}\right)$

C แทน ค่าคงที่ (Constant) = $\bar{Y} - b\bar{X}$

ดังนั้นจะเห็นได้ว่า วิธีการปรับเทียบเชิงเส้นตรง (Linear Equating) จึงมีลักษณะการคำนวณคล้ายกับวิธีการปรับเทียบโดยสร้างสมการทำนายคะแนนด้วยสมการถดถอย แต่มีวิธีการเก็บรวบรวมที่แตกต่างกัน และเป้าประสงค์ของการปรับเทียบ คือการประมาณค่าพารามิเตอร์ มากกว่าการทำนาย

ข้อแนะนำสำหรับการปรับเทียบเชิงเส้นตรง มีดังนี้

1. ถ้าแบบสอบที่นำมาปรับเทียบคะแนนมีความยากของแบบสอบแตกต่างกัน การแจกแจงคะแนนอาจมีความสัมพันธ์เชิงเส้นโค้ง การใช้วิธีการปรับเทียบเชิงเส้นตรงคงไม่เหมาะสม การใช้วิธีการปรับเทียบคะแนนแบบอิควิเปอร์เซ็นไทล์น่าจะเหมาะสมกว่า

2. ข้อตกลงเบื้องต้นของการปรับเทียบคะแนนแบบเส้นตรงถือเป็นสิ่งสำคัญ จึงควรตรวจสอบความสอดคล้องของการแจกแจงก่อนปรับเทียบและหลังการปรับเทียบ ว่าเป็นความสัมพันธ์เชิงเส้นตรงหรือไม่

3. ถ้าต้องการปรับเทียบของกลุ่มประชากรที่มีความสามารถแตกต่างกัน การปรับเทียบแบบเธอร์สโตน จึงเหมาะสมกับการปรับเทียบตามแนวตั้ง แต่มีข้อเสนอแนะเพิ่มเติม คือ

แบบสอบทั้งสองฉบับการกระจายค่าความยากของข้อสอบร่วมต้องมีลักษณะการแจกแจงแบบโค้งปกติ

- กลุ่มตัวอย่างที่นำมาเทียบคะแนนทั้งสองกลุ่ม ควรมีการแจกแจงแบบโค้งปกติ
- คะแนนมาตรฐานของแบบสอบทั้งสองฉบับต้องมีความสัมพันธ์กันเชิงเส้นตรง
- แบบสอบต้องสามารถวัดได้ถึงระดับอันตรภาค (Interval Scale) และต้องประกอบด้วยข้อสอบซึ่งเรียงลำดับจากข้อง่ายไปหาข้อยาก

3. วิธีการปรับเทียบคะแนนโดยใช้สมการถดถอย (Regression Method)

การปรับเทียบคะแนนโดยใช้สมการถดถอย เป็นวิธีที่ใช้เทคนิคการวิเคราะห์สมการถดถอย (Regression Equation) สำหรับปรับเทียบคะแนนระหว่างแบบสอบ การปรับเทียบคะแนนโดยวิธีนี้เป็นการใช้สมการเส้นตรงในการทำนายตัวแปรตาม (Dependent Variable) จากตัวแปรอิสระ (Independent Variable) ซึ่งมีลักษณะที่ไม่สมมาตร (Symmetry) กล่าวคือ สมการทำนาย Y จาก X หรือการทำนาย X จาก Y เป็นสมการที่ไม่สมมาตร หรือให้ผลทำนายในทิศทางเดียวกัน นอกจากนี้คะแนนจากแบบสอบที่ใช้เป็นตัวทำนาย ยังมีข้อตกลงเบื้องต้นว่า จะต้องมีค่าความเที่ยงเป็น 1

ข้อเสนอแนะสำหรับการปรับเทียบคะแนนโดยใช้สมการถดถอย ดังนี้

1. แบบสอบที่ต้องการนำมาปรับเทียบคะแนนกัน ต้องมีความสัมพันธ์กับคะแนนเกณฑ์เท่าเทียมกัน มิเช่นนั้นแล้วจะทำให้คะแนนจากแบบสอบฉบับหนึ่งจะสามารถทำนายคะแนนเกณฑ์ได้แม่นยำกว่าอีกฉบับหนึ่ง ทำให้การปรับเทียบคะแนนมีความผันแปรตามกลุ่มที่ศึกษา

2. แบบสอบที่นำมาปรับเทียบคะแนนกัน ต้องเป็นแบบสอบที่วัดคุณลักษณะเดียวกัน มีความเป็นคู่ขนาน และมีความเที่ยงสูง

4. วิธีการปรับเทียบค่าเฉลี่ย (Mean Equating)

วิธีนี้ใช้สำหรับรูปแบบการเก็บรวบรวมข้อมูลแบบกลุ่มสุ่ม เมื่อมีการปรับเทียบคะแนนสังเกตได้ (Observed Score Equating) เป็นการพิจารณาค่าความยากของแบบสอบชุดที่ 1 ต่างจากชุดที่ 2 คงที่ตลอดตั้งแต่คะแนนน้อยที่สุดจนถึงคะแนนมากที่สุด เช่น แบบสอบชุดที่ 1 ง่ายกว่าแบบสอบชุดที่ 2 อยู่ 5 คะแนนในกลุ่มผู้ได้คะแนนสูง และง่ายกว่ากลุ่มต่ำ 5 คะแนน เช่นเดียวกัน มีกระบวนการปรับเทียบคะแนนดังนี้

กำหนดแบบสอบ X เป็นแบบสอบใหม่ที่ต้องการปรับเทียบ χ เป็นคะแนนที่ได้จากแบบสอบ X ทั้งหมด และ x แทน คะแนนเฉพาะที่ต้องการปรับเทียบ

และกำหนดแบบสอบ Y เป็นแบบสอบชุดเดิม γ แทนคะแนนที่ได้จากแบบสอบ Y ทั้งหมด และ y แทนคะแนนเฉพาะจากแบบสอบ Y

ให้ $\mu(\chi)$ เป็นค่าเฉลี่ยของคะแนนที่ได้จากแบบสอบ X

ให้ $\mu(\gamma)$ เป็นค่าเฉลี่ยของคะแนนที่ได้จากแบบสอบ Y

ความหมายของการปรับเทียบคะแนนคือ คะแนนจาก 2 แบบสอบ มีความแตกต่างจากค่าเฉลี่ยแต่ละกลุ่มเท่ากัน กำหนดโดย

$$x - \mu(\chi) = y - \mu(\gamma)$$

$$\text{จะได้} \quad m_y(x) = y = x - \mu(\chi) + \mu(\gamma)$$

เมื่อ $m_y(x)$ แทน คะแนน x จากแบบสอบ X แปลงไปสู่เสกลของแบบสอบ Y โดยใช้วิธีการปรับเทียบค่าเฉลี่ย

ตัวอย่างเช่น เมื่อคะแนนเฉลี่ยของแบบสอบ X เท่ากับ 68 ค่าเฉลี่ยของแบบสอบ Y เท่ากับ 74 เมื่อแทนค่าในสมการข้างบนได้ดังนี้

$$m_y(x) = x - 68 + 74 = x + 6$$

คะแนนเฉลี่ยของแบบสอบ X เป็น 68 คะแนน เมื่อพิจารณาแล้วจะอยู่ในระดับความสามารถเท่ากับ 74 คะแนนในแบบสอบ Y และคะแนนจากแบบสอบ X 73 คะแนน จะอยู่ในระดับความสามารถเดียวกันกับแบบสอบชุด Y เท่ากับ 79 คะแนน

1.7 การปรับเทียบคะแนนตามวิธีทฤษฎีตอบสนองข้อสอบ

วิธีการปรับเทียบคะแนนที่กล่าวมาเป็นรูปแบบการปรับเทียบโดยใช้ทฤษฎีดั้งเดิมทำการปรับเทียบคะแนนดิบ (Raw Score) ระหว่างแบบสอบ วิธีดังกล่าวมีจุดอ่อนทั้งในด้านความเสมอภาค (Equity) ความสมมาตร (Symmetry) และความไม่แปรผันตามกลุ่ม (Invariance) การปรับเทียบโดยใช้ทฤษฎีการตอบสองข้อสอบ สามารถแก้ปัญหาเหล่านี้ได้ ถ้ามีการเลือกใช้โมเดล IRT ที่มีความสอดคล้องกับข้อมูล (Kolen, 1981)

การปรับเทียบคะแนนตามวิธีการนี้อาศัยรูปแบบความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้สอบข้อสอบใด ๆ ว่าขึ้นอยู่กับความสามารถของผู้สอบ ซึ่งประมาณค่าได้จากโค้งแสดงลักษณะข้อสอบ (Item Characteristic Curve : ICC) ของแต่ละรูปแบบที่ใช้ ซึ่งอาจจะเป็นรูปแบบ 1, 2 หรือ 3 พารามิเตอร์ โดยไม่ได้ขึ้นอยู่กับการแจกแจงคะแนนของกลุ่มตัวอย่าง (Lord & Novick, 1968) ซึ่ง Lord (1980) ได้ตั้งข้อกำหนดที่สำคัญ 3 ประการ ในการปรับเทียบคะแนนแบบสอบฉบับที่วัดคุณลักษณะเดียว ซึ่งเทียบคะแนนที่เทียบเท่ากันหรือสมมูลกันที่ระดับความสามารถเดียวกันไว้ดังนี้

1. ความเสมอภาค (Equity) คือ ถ้าพิจารณาสำหรับทุก ๆ ค่าความสามารถ (θ) การแจกแจงความถี่อย่างมีเงื่อนไขของคะแนนแปลง x(y) ที่ θ กำหนดให้ต้องเหมือนกันกับการแจกแจงความถี่อย่างมีเงื่อนไขของคะแนน x

2. ความไม่แปรเปลี่ยนไปตามกลุ่ม (Invariance) คือ คะแนนแปลง x(y) คะแนนแปลงคงที่โดยไม่ขึ้นกับตัวแปรอื่น ๆ ของประชากรที่นำมาปรับเทียบคะแนน

3. ความสมมาตร (Symmetry) คือ การปรับเทียบคะแนนจะต้องเหมือนกัน ไม่ว่าการปรับเทียบคะแนนนั้นจะปรับเทียบจากแบบสอบ X ไปยังแบบสอบ Y และแบบสอบ Y ไปยังแบบสอบ X

Hambleton และ Swaminathan (1985) ได้เสนอข้อตกลงเบื้องต้นเกี่ยวกับทฤษฎี IRT ไว้ 4 ประการ ดังนี้

1. ความเป็นเอกมิติ (Unidimentional) คือ คุณลักษณะ หรือความสามารถที่เป็นตัวกำหนดพฤติกรรมการตอบข้อสอบแต่ละข้อ (ตอบผิดหรือถูก) มีลักษณะเด่นหรือความสำคัญเพียงลักษณะเดียว

2. ความเป็นอิสระ (Local Independence) การตอบข้อสอบแต่ละข้อของผู้สอบมีความเป็นอิสระจากกัน หรือไม่มีอิทธิพลต่อกัน ทั้งระหว่างข้อสอบและระหว่างผู้สอบ

3. โค้งคุณลักษณะข้อสอบสามารถอธิบายพฤติกรรมการตอบข้อสอบ (Item Characteristic Curves : Item Response Model) ความสัมพันธ์ระหว่างระดับความสามารถของผู้สอบกับโอกาสในการตอบข้อสอบถูกสามารถแสดงได้ด้วยโค้งคุณลักษณะข้อสอบ

4. ข้อสอบที่ใช้ต้องไม่เป็นข้อสอบประเภทความเร็ว (Nonspeededness of the Test) ผู้สอบทุกคนควรมีโอกาสทำข้อสอบทุกข้อ เพื่อให้คะแนนรวมจากการสอบเป็นค่าความสามารถที่แท้จริง ไม่มีข้อจำกัดเกี่ยวกับเวลาที่ใช้ในการสอบ

เนื่องจากวิธีการเก็บรวบรวมข้อมูลทั้ง 4 วิธี วิธีการใช้ข้อสอบร่วม (Anchor-Test Design) เป็นวิธีที่เป็นไปได้ในทางปฏิบัติมากที่สุด ดังนั้นการคำนวณค่าสถิติ สำหรับการปรับเทียบสเกล (Scaling constants) ในที่นี้จะกล่าวถึงโดยอิงวิธีการเก็บข้อมูล โดยใช้ข้อสอบร่วมเป็นหลัก วิธีการคำนวณค่าคงที่สำหรับใช้ปรับเทียบคะแนนระหว่างแบบสอบตามวิธีทฤษฎีตอบข้อสอบที่สำคัญ 4 วิธี (ศิริชัย กาญจนวาสี, 2545) ดังนี้

1. วิธีสมการถดถอย (Regression Method)

การปรับค่าประมาณพารามิเตอร์ความยาก (b) ของข้อสอบร่วมจากแบบสอบ X ให้อยู่บนมาตรหรือสเกลเดียวกันกับแบบสอบ Y ด้วยสมการเส้นตรง ดังนี้

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

เมื่อ b_{y_c}, b_{X_c} แทน ค่าประมาณพารามิเตอร์ความยากของข้อสอบร่วมจากแบบสอบ Y และ X

α, β แทน ค่าคงที่ หรือสัมประสิทธิ์การถดถอย

โดย $$\alpha = r\left(\frac{S_{y_c}}{S_{x_c}}\right)$$

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

เมื่อ r แทน สัมประสิทธิ์สหสัมพันธ์ระหว่าง b_x และ b_y ของข้อสอบร่วม

$\bar{b}_{y_c}, \bar{b}_{x_c}$ แทน ค่าเฉลี่ยของ b สำหรับข้อสอบร่วมจากฉบับ X และ Y

s_{y_c}, s_{x_c} แทน ค่าส่วนเบี่ยงเบนมาตรฐานของ b สำหรับข้อสอบร่วมจากฉบับ X และ Y

ในทำนองเดียวกัน การปรับค่าประมาณความสามารถ (θ) ของผู้สอบร่วมที่สอบฉบับ X บนมาตรหรือสเกลเดียวกับฉบับ Y ด้วยสมการดังนี้

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

วิธีสมการถดถอยมีปัญหาในด้านความสมมาตร (Symmetry) ของผลการปรับเทียบคะแนน การปรับเทียบคะแนนจากแบบสอบ X และ Y จะให้ผลต่างจากการปรับเทียบคะแนนจากแบบสอบ Y ไป X

2. **วิธีค่าเฉลี่ยและซิกมา** (Mean and Sigma Method)

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

$$\therefore \bar{b}_{y_c} = \alpha \bar{b}_{X_c} + \beta$$

และ $s_{y_c} = \alpha s_{x_c}$

ดังนั้น ค่าคงที่จึงคำนวณได้ ดังนี้ $\alpha = \left(\frac{S_{y_c}}{S_{x_c}} \right)$

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

เนื่องจาก $b_{y_c} = \alpha b_{X_c} + \beta$ การแปลง b_{y_x} เป็น b_{x_x} จึงคำนวณได้เช่นเดียวกัน ดังนี้ $b_{x_c} = \frac{b_{y_x} - \beta}{\alpha}$

วิธีค่าเฉลี่ยและซิกมา จึงมีลักษณะสมมาตร การปรับเทียบคะแนนจากแบบสอบ X ไป Y หรือจากแบบสอบ Y ไป X ให้ผลเหมือนกัน

- กรณี 1- พารามิเตอร์

$$\bar{b}^{*}_{y_c} = \alpha \bar{b}_{X_c} + \beta \qquad (\because \alpha = 1)$$

$$\therefore \quad \bar{b}_{y_c} = \alpha \bar{b}_{X_c} + \beta$$

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

- กรณี 2 และ 3-พารามิเตอร์

$$\bar{b}^{*}_{y_c} = \alpha \bar{b}_{X_c} + \beta$$

และ $$a_y{}^{*} = \frac{a_x}{\alpha}$$

3. วิธีปรับแก้ค่าเฉลี่ยและซิกมา (Robust Mean and Sigma Method)

การคำนวณตามวิธีค่าเฉลี่ยและซิกมามิได้พิจารณาถึงระดับความแม่นยำที่เกิดขึ้นจากการประมาณค่าพารามิเตอร์ของข้อสอบแต่ละตัวที่ใช้ในสมการ Linn และคณะ (1981) ได้เสนอวิธีการปรับแก้ค่าเฉลี่ยและซิกมา โดยนำความคลาดเคลื่อนมาตรฐานของการประมาณค่าพารามิเตอร์ของข้อสอบร่วม เป็นรายข้อมาใช้ถ่วงน้ำหนัก ข้อสอบร่วมที่มีค่าแปรปรวนระหว่างฉบับสูงจะมีน้ำหนักน้อย ส่วนข้อที่มีความแปรปรวนระหว่างฉบับต่ำจะมีน้ำหนักมาก ค่าน้ำหนักเป็นส่วนกลับของความแปรปรวนของค่าประมาณตัวที่มีค่ามาก โดยมีขั้นตอนการคำนวณ ดังนี้

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

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

เมื่อ $V\left(b_{y_{ci}}\right)$ และ $v\left(b_{x_{ci}}\right)$ เป็นค่าความแปรปรวนของค่าประมาณพารามิเตอร์ b สำหรับข้อสอบร่วมฉบับ Y และ X

(2) กำหนดค่าประมาณของพารามิเตอร์ที่ถ่วงน้ำหนัก

$$W'_i = \frac{w_i}{\sum_{j=1}^{k} w_j}$$

เมื่อ k เป็นจำนวนข้อสอบร่วมสำหรับแบบสอบฉบับ Y และ X

(3) คำนวณค่าประมาณของพารามิเตอร์ที่ถ่วงน้ำหนัก

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

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

(4) คำนวณค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าประมาณพารามิเตอร์ที่ถ่วงน้ำหนัก

(5) คำนวณค่าคงที่ α และ β โดยใช้วิธีค่าเฉลี่ยและซิกมา ดังกล่าวข้างต้น

4. **วิธีโค้งลักษณะข้อสอบ** (Characteristic Curve Method)

Stocking และ Lord (1983) ได้เสนอวิธีโค้งลักษณะข้อสอบโดยนำสารสนเทศที่ได้จากค่าพารามิเตอร์ของข้อสอบทั้งค่าพารามิเตอร์ความยาก และอำนาจจำแนก มาใช้ในการคำนวณค่าคงที่ สำหรับการปรับเทียบคะแนนระหว่างแบบสอบ ดังนี้

สมมุติให้ T_{Xa} แทนคะแนนจริงของข้อสอบที่มีความสามารถ θ_a จากการทำแบบสอบ X ซึ่งมีข้อสอบร่วม k ข้อ

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

ในทำนองเดียวกันให้ T_{Ya} แทนคะแนนจริงของผู้สอบที่มีความสามารถเท่ากัน (θ_a) จากการทำแบบสอบ Y ซึ่งมีข้อสอบร่วม k ข้อ

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

สำหรับชุดของข้อสอบร่วม k ข้อ

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

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

$$c_{y_{ci}} = c_{x_{ci}}$$

การคำนวณค่าคงที่ α และ β ได้มาจากการทำให้ฟังก์ชัน F มีค่าต่ำสุด

$$F = \frac{1}{N}\sum_{a=1}^{N}\left(T_{Xa} - T_{Ya}\right)^2$$

เมื่อ N เป็นจำนวนข้อสอบ และ F เป็นฟังก์ชันแสดงความแตกต่างระหว่างคะแนนจริงที่ได้จากแบบสอบทั้งสองของผู้สอบที่มีความสามารถเดียวกัน การคำนวณค่าคงที่ α และ β ใช้กระบวนการกระทำซ้ำหลายรอบจนได้ค่าที่ดีที่สุด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รูปแบบการตอบข้อสอบรายข้อ เนื่องจากข้อมูลที่ได้จากการสอบมีหลายลักษณะ ได้แก่ ข้อมูลแบบทวิวิภาค (Dichotomous) และข้อสอบแบบหลายค่า (Polytomous) ดังนั้นจึงมีการพัฒนารูปแบบเพื่อให้สอดคล้องกับลักษณะข้อมูลดังกล่าวขึ้นมากมาย แต่สำหรับข้อมูลที่เป็นแบบทวิ รูปแบบที่นิยมใช้ ได้แก่ รูปแบบโลจิสติก (Logistic Model) ซึ่งแตกต่างไปตามจำนวนพารามิเตอร์ที่ใช้ในแต่ละรูปแบบ ซึ่งเรียงลำดับตามการพัฒนารูปแบบโลจิสติก มีรายละเอียดดังนี้

1. รูปแบบโลจิสติก 2 พารามิเตอร์ (Two-Parameter Logistic Model)

รูปแบบนี้ได้เสนอโดย Birnbaum เมื่อปี ค.ศ. 1957 เป็นโค้งคุณลักษณะข้อสอบและเป็นฟังก์ชันของการแจกแจงที่มีสองพารามิเตอร์

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

เมื่อ $P_i(\theta)$ เป็นความน่าจะเป็นของผู้สอบที่มีความสามารถ θ สามารถตอบข้อสอบ i ได้ถูกต้อง b_i และ a_i เป็นพารามิเตอร์ของข้อสอบ i D เป็นค่าองค์ประกอบสเกล (Scaling Factor) ซึ่งถ้ากำหนดให้มีค่าเท่ากับ 1.7 แล้วค่า $P_i(\theta)$ จากโค้งความถี่สะสมกับโค้งโลจิสติกจะมีค่าต่างกันน้อยกว่า .01 สำหรับทุกค่าของ θ

จากรูปแบบนี้อยู่บนข้อตกลงที่ว่าการเดาคำตอบจะไม่เกิดขึ้น ซึ่งจะเป็นเช่นนี้ได้ก็ต่อเมื่อค่าพารามิเตอร์ $a_i > 0$ (ข้อสอบที่มีความสัมพันธ์ด้านบวกระหว่างคะแนนจากการสอบกับความสามารถของผู้สอบที่วัดโดยแบบสอบนั้น) และค่าความน่าจะเป็นในการตอบข้อสอบได้ถูกจะลดลงถึงศูนย์เมื่อความสามารถลดลง

2. แบบโลจิสติก 3 พารามิเตอร์ (Three-Parameter Logistic Model)

รูปแบบโลจิสติกสามพารามิเตอร์ เป็นการปรับปรุงสองพารามิเตอร์เพียงแต่เพิ่มพารามิเตอร์ตัวที่สาม คือ พารามิเตอร์การเดาคำตอบ หรือ พารามิเตอร์ C_i เข้าไปรูปแบบนี้เขียนในรูปสมการเชิงคณิตศาสตร์ได้ดังนี้

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

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

b_i คือ พารามิเตอร์ความยาก

a_i คือ พารามิเตอร์อำนาจจำแนก

c_i คือ พารามิเตอร์โอกาสการเดาข้อสอบได้ถูก

D คือ ค่าองค์ประกอบสเกล (D=1.7)

พารามิเตอร์ c_i เป็นจุดต่ำสุดที่โค้งคุณลักษณะข้อสอบ ซึ่งเป็นพารามิเตอร์นี้จะใช้เมื่อคิดว่าการเดาเป็นองค์ประกอบในการตอบข้อสอบ บางครั้งเรียกพารามิเตอร์นี้ว่าเป็นโอกาสที่จะตอบได้ถูกต้องสำหรับคนที่มีความสามารถต่ำ

ในการปรับรูปแบบสามพารามิเตอร์ ให้เป็นรูปแบบสองพารามิเตอร์ ต้องอยู่บนข้อตกลงที่ว่าพารามิเตอร์การเดามีค่าเท่ากับศูนย์

3. **รูปแบบโลจิสติก 1 พารามิเตอร์** (One-Parameter Logistic Model)

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

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

ถึงแม้ว่ารูปแบบนี้จะเป็นกรณีเฉพาะของรูปแบบสองพารามิเตอร์ และสามพารามิเตอร์แต่ก็ยังมีคุณสมบัติพิเศษที่ทำให้นิยมใช้กัน คือ ประการแรก เนื่องจากรูปแบบนี้มีจำนวนพารามิเตอร์ไม่มากจึงสะดวกต่อการใช้งาน ประการที่สอง ปัญหาที่เกิดจากการประมาณค่าพารามิเตอร์มีน้อยกว่า การประมาณค่าพารามิเตอร์สำหรับรูปแบบที่มีพารามิเตอร์หลาย ๆ ตัว

การปรับเทียบคะแนนระหว่างแบบสอบโดยใช้ทฤษฎีการตอบข้อสอบ มีหลายวิธีที่สามารถนำมาใช้ในการปรับเทียบคะแนนได้ โดยพิจารณาให้สอดคล้องกับรูปแบบการเก็บรวบรวมข้อมูลและการวิเคราะห์ ในกรณีที่กลุ่มตัวอย่างมีขนาดไม่ใหญ่นักการเลือกรูปแบบ 1 พารามิเตอร์มีความเหมาะสมกับการปรับเทียบคะแนนในนอน แต่ถ้ากลุ่มตัวอย่างขนาดใหญ่การเลือกรูปแบบ 2 พารามิเตอร์ หรือ 3 พารามิเตอร์ จะเหมาะสมมากกว่า สำหรับการปรับเทียบแนวตั้ง ดังนั้นในการวิจัยครั้งนี้ ผู้วิจัยจึงเลือกวิธีวิเคราะห์ รูปแบบ 3 พารามิเตอร์

การกำหนดสเกลสำหรับรายงานผล

การปรับเทียบคะแนนโดยใช้ทฤษฎีตอบสนองข้อสอบ มีการปรับเทียบโดยใช้คะแนนจริง และคะแนนสังเกตได้ ทั้งสองวิธีมีทั้งข้อได้เปรียบและเสียเปรียบ กล่าวคือ การปรับเทียบคะแนนโดยใช้คะแนนจริงจะไม่สามารถอธิบายคะแนนที่อยู่ต่ำกว่าระดับการเดาได้ โดยจะให้ความหมายของคะแนนสมมูลเฉพาะคะแนนที่อยู่เหนือค่าเฉลี่ยของการเดา ซึ่งถึงแม้จะเป็นการปรับเทียบโดยใช้คะแนนจริง แต่ยังเป็นค่าที่ประมาณได้จากสูตรในการคำนวณ ดังนั้นจึงยังต้องมีความคลาดเคลื่อนอยู่ ส่วนการปรับเทียบโดยใช้คะแนนสังเกตได้นั้นเป็นการปรับเทียบคะแนนโดยประมาณ ซึ่งอธิบายคะแนนสมมูลจาก X และ Y ได้ครอบคลุมพิสัยของคะแนนสังเกตได้ Lord ได้กล่าวว่า ทั้งสองวิธีนี้มีความสอดคล้องกันมาก แต่การสรุปผลในการอ้างอิงต้องทำอย่างรอบคอบ (Lord, 1980 อ้างใน พรพิมล นาคเวช, 2537)

การรายงานผลคะแนนที่ปรับเทียบแล้ว สามารถเลือกมาตรสำหรับการรายงานเป็นคะแนนความสามารถ (θ) คะแนนจริง (ξ) และ คะแนนดิบ (Γ)

กรณีที่ต้องการรายงานผลการปรับเทียบคะแนนความสามารถ (θ) จากสมการ การปรับเทียบ θ ดังกล่าว เมื่อคำนวณผลตามสมการ ก็สามารถสร้างตารางรายงานผลคะแนนสมมูลกับตำแหน่ง θ เดียวกัน แต่ในกรณีที่ต้องการรายงานผลการปรับเทียบด้วยคะแนนจริง (ξ) จำเป็นต้องสร้างตารางปรับเทียบ หรือพล็อตกราฟที่จะทำให้ได้คะแนนจริงที่สมมูลกัน ส่วนกรณีของการรายงานผลการปรับเทียบด้วยคะแนนดิบ (Γ) หรือคะแนนที่สังเกตได้ (Observed Score) จำเป็นต้องมีกระบวนการ (ศิริชัย กาญจนวาสี, 2545) ดังต่อไปนี้

1. **การปรับเทียบคะแนนจริง**

เพื่อให้ง่ายต่อการแปลความหมาย จึงเกิดแนวคิดการแปลงคะแนนความสามารถ (θ) ให้อยู่ในรูปของคะแนนดิบ (True Score) เพื่อให้มีความหมายสอดคล้องกับสภาพความเป็นจริงและเข้าใจได้ง่ายขึ้น การปรับเทียบคะแนนจริงระหว่างแบบสอบ มีวิธีดำเนินการดังนี้

ให้ ξ_x = คะแนนจริงของผู้สอบแบบสอบฉบับ X ที่มีระดับความสามารถ θ_x

ซึ่ง $\xi_x = \sum_{i=1}^{n} P_i(\theta_x)$

ในทำนองเดียวกันให้ ξ_y = คะแนนจริงของผู้สอบแบบสอบฉบับ X ที่มีระดับความสามารถ θ_y

ซึ่ง $\xi_y = \sum_{j=1}^{m} P_j(\theta_y)$

แต่เนื่องจาก $\theta_y = T_{(\theta_x)} = \alpha\theta_x + \beta$

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

เนื่องจากแต่ละค่าของ θ_x สามารถแปลงเป็นคะแนนจริงที่สมมูลกับ θ_y ได้คือ (ξ_y , ξ_x) จากการคำนวณค่าสมมูลเป็นรายคู่ จึงสามารถทำเป็นตารางปรับเทียบคะแนนจริงระหว่างฉบับได้ หรือพล็อตกราฟ ซึ่งเส้นกราฟที่แสดงความสัมพันธ์ระหว่างคะแนนความสามารถ θ_x และ θ_y ที่สมมูลกัน สามารถสร้างได้โดยการคำนวณค่าคงที่ในสมการด้วยวิธีใดวิธีหนึ่งดังกล่าวมาแล้ว เมื่อทราบค่า θ_x และ θ_y สมมูลกัน แล้วนำมาแปลงเป็นคะแนนจริง ξ_x และ ξ_y ที่สมมูลกัน ความสัมพันธ์ของ ξ_x และ ξ_y จะไม่เป็นเส้นตรง อันเป็นผลมาจากการประมาณค่า θ_x และ θ_y มีความคลาดเคลื่อนเกิดขึ้นในระดับหนึ่ง เนื่องจากความคลาดเคลื่อนของการสุ่ม ซึ่งจะมีขนาดใหญ่ขึ้น สำหรับค่าที่อยู่สุดขั้วของพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ ดังนั้นการปรับเทียบคะแนนจริงตรงคะแนนปลายขั้ว โดยใช้เส้นกราฟจึงควรใช้วิธีการแก้ของ Angoff (182) หรือหลีกเลี่ยงปัญหาโดยการคำนวณด้วยสูตร ξ_x และ ξ_y โดยตรงดังกล่าวข้างต้น

2. **การปรับเทียบคะแนนสังเกตได้** (Observed-Score Equating)

การปรับเทียบคะแนนตามแนว IRT สามารถทำได้โดยการทราบค่าพารามิเตอร์ความสามารถ (θ) หรือคะแนนจริง คำถามที่ตามมาคือ สามารถปรับเทียบคะแนนที่ได้จากการสอบหรือไม่ สำหรับกรณีที่ต้องการปรับเทียบคะแนนระหว่างคะแนนสอบที่สังเกตได้โดยตรง (Observed score) หรือคะแนนดิบ (Raw score) โดยใช้ IRT ก็สามารถกระทำได้เช่นเดียวกัน

เนื่องจากคะแนนจริง (ξ) จะอยู่บนสเกลเดียวกันกับคะแนนดิบ (Γ)

ก็ต่อเมื่อ $\Gamma = \sum_{i=1}^{n} U_i$

เมื่อ U_i = คะแนนรายข้อ (U_i =1 เมื่อตอบถูก และ U_i =0 เมื่อตอบผิด)

ยิ่งกว่านั้น ถ้า IRT มีความสมเหตุสมผลจะได้ $E(\Gamma) = \xi$ ดังนั้นสิ่งที่ทำได้ในตอนนี้ คือ

1. หาความสัมพันธ์ระหว่างคะแนนจริง ξ_x และ ξ_y จากแบบสอบดังที่ได้กล่าวมาแล้ว

2. ให้การกระทำกับความสัมพันธ์คล้ายกับเป็นความสัมพันธ์ระหว่างคะแนนดิบ Γ_x และ Γ_y แล้วปรับเทียบคะแนนดิบ

Lord (1980) ได้อธิบายว่าความสัมพันธ์ที่มีอยู่ระหว่าง ξ_x และ ξ_y ไม่จำเป็นต้องเหมือนกับที่มีอยู่ระหว่างคะแนนดิบ Γ_x และ Γ_y เรื่องนี้สามารถอธิบายได้จากโมเดล 3 พารามิเตอร์ คือ $\xi_x > \sum_{i=1}^{n} c_i$ และ $\xi_y > \sum_{i=1}^{m} c_i$ ขณะที่คะแนนที่ได้จากการสอบ (คะแนนดิบ) Γ_x และ Γ_y อาจมีค่าเป็น 0 ดังนั้นการปรับเทียบคะแนนโดยใช้คะแนนจริงจะไม่ให้สารสนเทศการปรับเทียบของผู้สอบที่มีคะแนนดิบต่ำกว่าระดับการเดาเพื่อหลีกเลี่ยงปัญหานี้อาจจะใช้สูตรการแปลงคะแนนเข้ามาช่วยได้ อย่างไรก็ตามถ้าจะกล่าวกันตามหลักเกณฑ์โดยตรงแล้วความสัมพันธ์ระหว่างคะแนนจริง ξ_x และ ξ_y อาจจะไม่นำมาใช้ในการปรับเทียบคะแนนดิบ Γ_x และ Γ_y

ทฤษฎีตอบสนองข้อสอบ (IRT) ได้ให้วิธีการสำหรับทำนายการแจกแจงคะแนนดิบของแบบสอบที่กำหนดได้ เมื่อได้ลักษณะการแจกแจงคะแนนดิบของแบบสอบฉบับ X และ ฉบับ Y อาจจะใช้การปรับเทียบอีควิเปอร์เซ็นไทล์ แต่ข้อได้เปรียบอย่างมากของ IRT คือ แบบสอบที่นำมาใช้นั้นไม่จำเป็นต้องมีระดับความยากใกล้เคียงกัน

ตามทฤษฎีแล้วการแจกแจงของคะแนนดิบ $f(\Gamma|\theta)$ ของแบบสอบสามารถกำหนดจากความเป็นเอกลักษณ์ (Identity) (Lord, 1980) ดังนี้

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

การปรับเทียบคะแนนที่สังเกตได้หรือคะแนนดิบ โดยใช้ IRT สามารถสรุปเป็นขั้นตอนดังนี้

1. นำคะแนนจากกลุ่มผู้สอบแบบสอบฉบับ X มาวิเคราะห์เพื่อประมาณค่าพารามิเตอร์ของข้อสอบและผู้สอบ แล้วนำค่า θ_x ของผู้สอบแต่ละคนมาคำนวณความถี่สัมพัทธ์อย่างมีเงื่อนไขทางทฤษฎีของการแจกแจงคะแนนดิบ $f_x(\Gamma \mid \theta_x)$

2. แปลงค่า θ_x ให้สมมูลกับ θ_y ด้วยสูตร $T_{(\theta_x)} = \alpha\theta_x + \beta$ ตามวิธีที่กล่าวมาข้างต้น แล้วนำ $T_{(\theta_x)}$ ของกลุ่มผู้สอบมาคำนวณความถี่ของคะแนนดิบทางทฤษฎี $f_x(\Gamma)$ จึงได้ตารางการแจกแจงความถี่ของคะแนนดิบทางทฤษฎีของกลุ่มผู้สอบ ฉบับ X

3. นำคะแนนจากกลุ่มผู้สอบแบบสอบฉบับ Y มาวิเคราะห์เพื่อประมาณค่าพารามิเตอร์ของข้อสอบและผู้สอบ แล้วนำค่า θ_y ของผู้สอบแต่ละคนมาคำนวณความถี่สัมพัทธ์อย่างมีเงื่อนไขทางทฤษฎีของการแจกแจงคะแนนดิบ $f_y(\Gamma \mid \theta_y)$

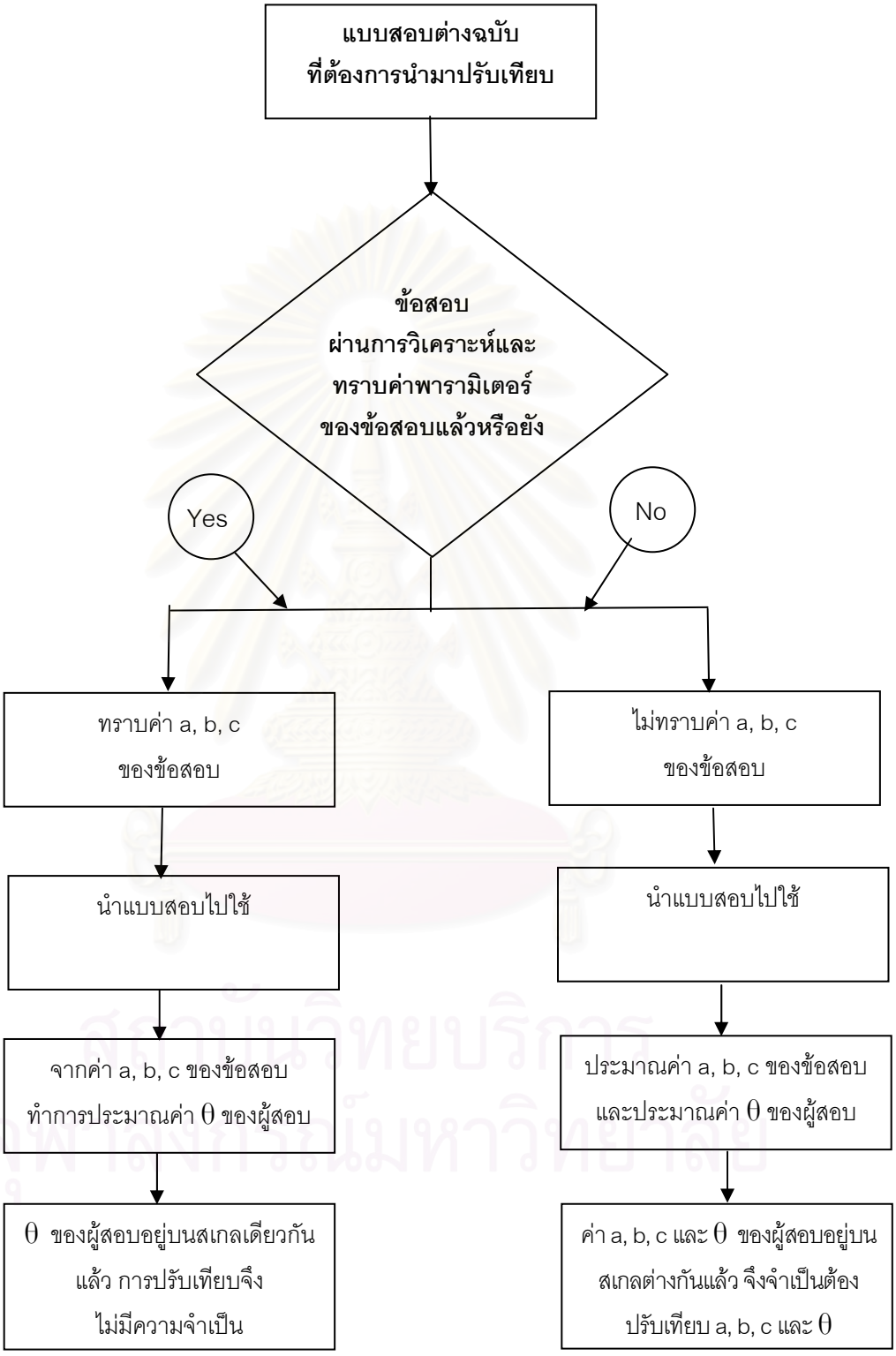
4. นำค่า θ_y ของกลุ่มผู้สอบมาคำนวณความถี่ของคะแนนดิบทางทฤษฎี $f_y(\Gamma)$ จึงได้ตารางการแจกแจงความถี่ของคะแนนดิบทางทฤษฎีของการแจกแจงคะแนนดิบทางทฤษฎีของกลุ่มผู้สอบฉบับ Y

5. ปรับเทียบคะแนนระหว่างคะแนนดิบของแบบสอบ X และ Y โดยใช้วิธีอีควิเปอร์เซ็นไทล์

สรุปขั้นตอนการปรับเทียบคะแนนโดยวิธีทฤษฎีตอบสนองข้อสอบ

ขั้นตอนแรกของการใช้ทฤษฎีตอบสนองข้อสอบสำหรับการปรับเทียบคะแนนระหว่างแบบสอบ จะต้องพิจารณาว่า แบบสอบต่างฉบับที่จะนำมาใช้นั้นได้มีการวิเคราะห์และคำนวณค่าพารามิเตอร์ของข้อสอบแล้วหรือไม่ ถ้าทราบค่าพารามิเตอร์ของข้อสอบทุกข้อแล้ว สามารถนำแบบสอบฉบับเหล่านั้นไปใช้ได้เลย ผลการตอบข้อสอบที่ได้ ประมาณค่าความสามารถของผู้สอบ (θ) ซึ่งตามปกติค่า θ จะลู่เข้าสู่ค่าที่แท้จริง เมื่อมีจำนวนข้อสอบเพิ่มมากขึ้น และค่า θ จะอยู่บนสเกลร่วมกันอยู่แล้ว ดังนั้นในกรณีนี้การปรับเทียบคะแนนจึงไม่มีความจำเป็นแต่ในกรณีที่แบบสอบต่างฉบับกันนั้นไม่ทราบค่าพารามิเตอร์ของข้อสอบมาก่อน จึงต้องใช้วิธีการประมาณค่าร่วมพร้อมกันไปทั้งค่าพารามิเตอร์ของข้อสอบ และค่าความสามารถของผู้สอบ (θ) ของกลุ่มต่างกัน จึงจำเป็นต้องมีการปรับเทียบระหว่างแบบสอบ ดังแสดงในแผนภาพที่ 2.3

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



แผนภาพที่ 2.3 ขั้นตอนแรกของการพิจารณาการปรับเทียบโดยใช้ IRT (ศิริชัย กาญจนวาสี, 2541)

กระบวนการดำเนินงานปรับเทียบคะแนนตามวิธีทฤษฎีตอบสนองข้อสอบ มีขั้นตอนดังนี้

1. เลือกรูปแบบการปรับเทียบ (Equating Design) รูปแบบการปรับเทียบคะแนนอาจเป็นรูปแบบผู้สอบกลุ่มเดียว (Single-Group Design) รูปแบบผู้สอบกลุ่มเท่าเทียมกัน (Equivalent-Group Design) รูปแบบผู้สอบต่างกลุ่มโดยใช้แบบสอบร่วม (Anchor-Test Design) และรูปแบบผู้สอบต่างกลุ่มโดยใช้ผู้สอบร่วม (Common-Person Design) สำหรับการเลือกรูปแบบการปรับเทียบที่เหมาะสมขึ้นอยู่กับธรรมชาติ ของแบบสอบและกลุ่มผู้สอบ

2. กำหนดโมเดลการตอบสนองข้อสอบ (Item Response Model) โมเดลการตอบสนองข้อสอบประเภทให้คะแนนรายข้อแบบ 0, 1 ที่นิยมใช้กันมี 3 โมเดล ได้แก่ โมเดล 1, 2 และ 3 พารามิเตอร์ การตัดสินใจเลือกใช้โมเดลใดขึ้นอยู่กับลักษณะของการปรับเทียบและขนาดกลุ่มตัวอย่าง

ในกรณีที่กลุ่มตัวอย่างมีขนาดไม่ใหญ่นัก การเลือกใช้โมเดล1 พารามิเตอร์ น่าจะมีความเหมาะสมกับการปรับเทียบในแนวนอน (Horizontal Equating) แต่ถ้ากลุ่มตัวอย่างมีขนาดใหญ่ การใช้โมเดล 2 พารามิเตอร์ หรือ 3 พารามิเตอร์ น่าจะเหมาะสมกว่า

สำหรับการปรับเทียบในแนวตั้ง (Vertical Equating) การใช้โมเดล 1 พารามิเตอร์ ไม่ค่อยเหมาะสม เนื่องจากลักษณะข้อสอบอาจจะไม่สอดคล้องกับโมเดล (Slinde & Linn, 1979 อ้างใน ศิริชัย กาญจนวาสี, 2541)

3. สร้างเมตริกซ์สเกลร่วมของความสามารถและพารามิเตอร์ข้อสอบ เนื่องจากพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบมีความสัมพันธ์กันเชิงเส้นตรง จึงทำให้สามารถสร้างเมตริกซ์สเกลร่วมได้ โดยการคำนวณค่าคงที่ในสมการ สำหรับการเปรียบเทียบค่าพารามิเตอร์ของข้อสอบให้อยู่บนสเกลเดียวกัน เพื่อปรับเทียบ θ ให้อยู่บนสเกลที่สามารถเปรียบเทียบกันได้โดยตรง

4. กำหนดสเกลสำหรับรายงานคะแนนสอบ (Test Score) ที่ปรับแล้ว การรายงานผลคะแนนสอบที่ปรับแล้ว สามารถเลือกสเกลสำหรับการรายงานเป็นคะแนนความสามารถ (θ) คะแนนจริง (ξ) และ คะแนนดิบ (Γ)

4.1 ถ้าต้องการรายงานผลการปรับเทียบด้วยคะแนนความสามารถ (θ) จากสมการการปรับเทียบ θ เมื่อคำนวณผลตามสมการ ก็สามารถสร้างตารางรายงานผลเปรียบเทียบกันได้เลย

4.2 ถ้าต้องการรายงานผลการปรับเทียบด้วยคะแนนจริง (ξ) สามารถคำนวณ ξ ที่สมมูลกันตรงตำแหน่ง θ ต่าง ๆ ได้ เพื่อนำมาสร้างตารางปรับเทียบ หรือพล็อตเป็นกราฟ ก็จะได้ค่าคะแนนที่สมมูลกัน

4.3 ถ้าต้องการรายงานผลการปรับเทียบด้วยคะแนนดิบ หรือ คะแนนที่สังเกตได้ (Observe Score) โดยมีขั้นตอนดังนี้

-คำนวณความถี่ของการแจกแจงคะแนนดิบอย่างมีเงื่อนไขตามทฤษฎีจากกลุ่มตัวอย่าง

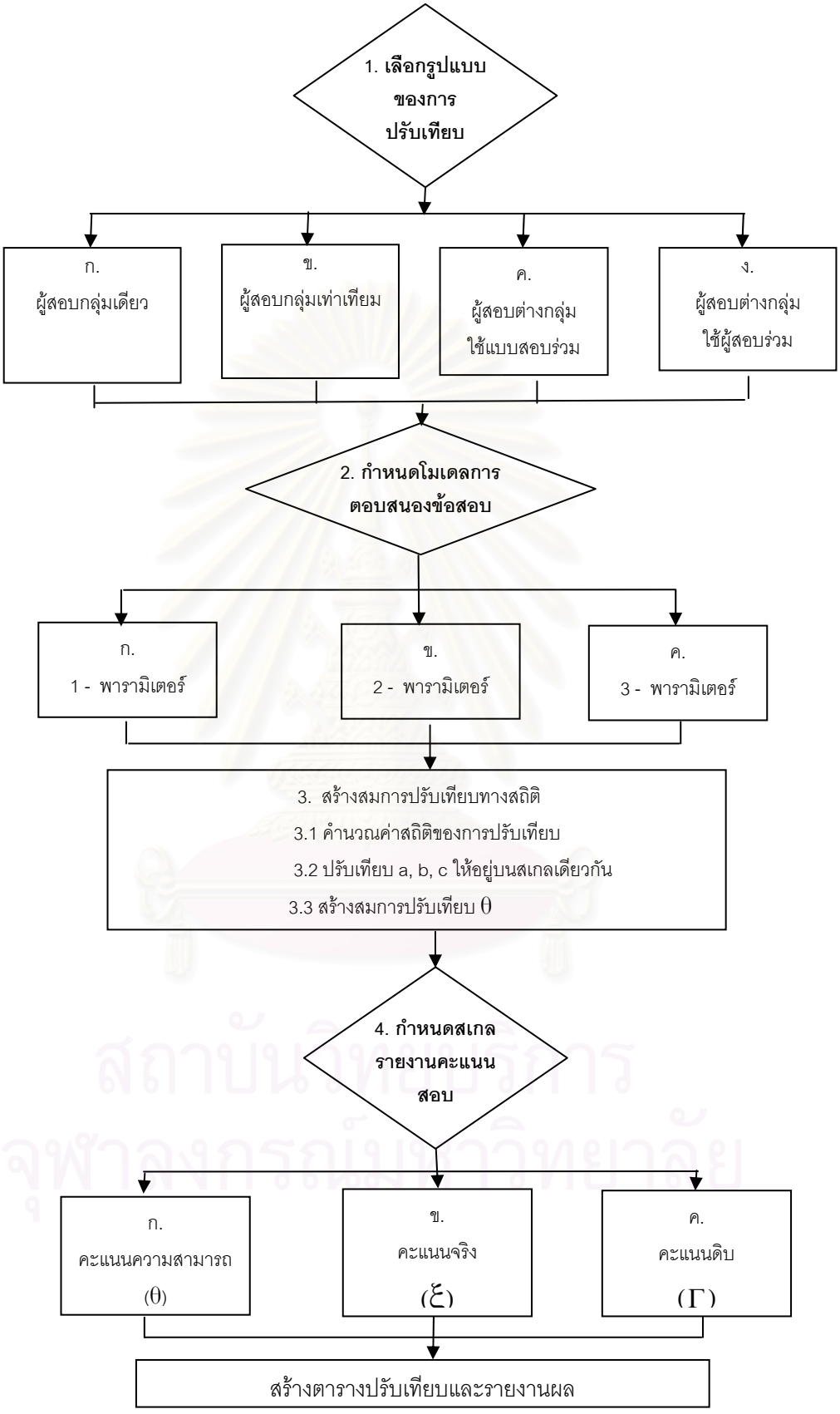
-คำนวณความถี่รวมของการแจกแจงคะแนนดิบตามทฤษฎีของแต่ละแบบสอบ

-คำนวณค่าเปอร์เซ็นไทล์ของคะแนนจากแต่ละแบบสอบ โดยปรับเทียบโดยหลักการของอิควิเปอร์เซ็นไทล์

-สร้างตารางปรับเทียบหรือ พล็อตกราฟที่ปรับเทียบค่าคะแนนดิบที่สมมูลกันจากต่างแบบสอบ

ขั้นตอนดังกล่าวสามารถสรุปได้ ดังแผนภาพที่ 2.4

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



แผนภาพที่ 2.4 สรุปขั้นตอนของการปรับเทียบคะแนนโดยใช้ IRT

1.8 ประสิทธิภาพของการปรับเทียบคะแนน

ในการประเมินประสิทธิภาพของการปรับเทียบคะแนนระหว่างแบบสอบ นิยมศึกษาสมการปรับเทียบจากกลุ่มตัวอย่างปรับเทียบคะแนน และนำสมการปรับเทียบที่ได้ไปใช้ตรวจสอบคุณภาพกับกลุ่มตัวอย่างสอบทานผล มีผู้เสนอผลการปรับเทียบไว้สามารถจัดได้เป็น 2 ประเภท (ศิริชัย กาญจนวาสี, 2545) ดังนี้

1. **ความคลาดเคลื่อนมาตรฐานของการปรับเทียบ** (Standard Error of Equating)

วิธีการปรับเทียบทุกวิธี กลุ่มตัวอย่างที่ได้จากการสุ่มจากประชากรเดียวกันหรือหลายประชากร จะมีความผันแปรเชิงสุ่มเกิดขึ้นกับกลุ่มตัวอย่างที่ทำการศึกษา ดังนั้นจึงสามารถวิเคราะห์ความคลาดเคลื่อนมาตรฐานที่เกิดขึ้นในเชิงทฤษฎี สำหรับการปรับเทียบคะแนนระหว่างแบบสอบจากกลุ่มสอบทานผล ได้ดังนี้

ให้ E เป็นความคลาดเคลื่อนของการปรับเทียบคะแนนของ X บนสเกลของ Y

$$\therefore E_i = \left| y_i - y_i^* \right|$$

เมื่อ y_i แทน คะแนนคนที่ i จากแบบสอบ Y

y_i^* แทน คะแนนคนที่ i จากการปรับเทียบคะแนนแบบสอบ X บนสเกลของแบบสอบ Y

ส่วนเบี่ยงเบนมาตรฐานของคะแนนความคลาดเคลื่อนจากการปรับเทียบเป็นดังนี้

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

เมื่อ n แทน ขนาดกลุ่มตัวอย่างในกลุ่มสอบทานผล

ดังนั้น ความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน (SEE) จึงคำนวณได้จากสูตร

$$SEE = \frac{SD_E}{\sqrt{n}}$$

2. **ความเพียงพอของการปรับเทียบคะแนน** (Adequacy of Equating Models)

วิธีการปรับเทียบคะแนนแต่ละวิธี ประกอบขึ้นด้วยรูปแบบของการปรับเทียบคะแนน (Equating Model) ซึ่งมีข้อตกลงเบื้องต้นของแต่ละรูปแบบและการออกแบบ เพื่อเก็บข้อมูลให้เป็นไปตามข้อตกลงเบื้องต้นอย่างถูกต้อง (Accurate) และแม่นยำ (Precise) ตามทฤษฎี แต่ความเป็นจริงของการสอบมักไม่เป็นตามนั้น เพราะมีหลายสิ่งหลายอย่างอยู่นอกเหนือการควบคุม เช่น โปรแกรมการทดสอบระดับชาติ ซึ่งมีทั้งกฎเกณฑ์และตัวนโยบายเป็นตัวกำหนด จึงไม่สามารถควบคุมการสอบให้เป็นไปตามข้อตกลงเชิงทฤษฎี ดังนั้นข้อมูลที่นำมาใช้ในการศึกษาจึงไม่สามารถระบุได้ว่าเป็นตัวอย่างของประชากรได้อย่างชัดเจน และโดยความเป็นจริงเป็นการจัดกระทำกับประชากรมากกว่า (Population Quantities) ไม่ใช่ค่าประมาณ (Sample Estimate) (Braun & Holland, 1982)

ด้วยสภาพดังกล่าวทำให้การปรับเทียบคะแนนที่จัดกระทำอยู่นั้นอยู่ในภาวะที่มีเงื่อนไขน้อยกว่าข้อตกลงที่มีในแต่ละรูปแบบ ดังนั้นการปรับเทียบคะแนนที่ได้พัฒนาขึ้นจึงจำเป็นต้องมีการตรวจสอบความเพียงพอของรูปแบบ วิธีการประเมินความเพียงพอมีผู้เสนอแนวคิดและวิธีปฏิบัติไว้ ดังนี้

2.1 **ดัชนีเปรียบเทียบเปอร์เซ็นไทล์** (Percentile Comparison Index) เป็นมาตรการวัดความไม่สอดคล้องระหว่างการแจกแจงของคะแนนในแบบสอบ ฉบับ X กับแบบสอบฉบับ Y ที่ได้แปลงไปอยู่ในสเกลคะแนนของแบบสอบฉบับ X แล้ว ตามวิธีที่ได้พัฒนาขึ้น ดัชนีเปรียบเทียบเปอร์เซ็นไทล์ คือ ค่าเฉลี่ยกำลังสองของความแตกต่าง (The Mean-Squared Difference) ที่ได้จากการแจกแจงของคะแนนต่าง ๆ ของเกณฑ์ X กับคะแนนแปลง X^* ที่แปลงมาจาก Y ด้วยวิธีปรับเทียบคะแนนที่ระบุไว้ ณ ที่ตำแหน่งเปอร์เซ็นไทล์เดียวกัน ดัชนีนี้ได้เสนอโดย Kolen (1982) ได้แนะนำให้ใช้ข้อมูลจากกลุ่มตัวอย่างสอบทานผล ซึ่งได้สุ่มมาจากประชากรเดียวกันกับกลุ่มตัวอย่างที่ได้พัฒนาตารางคะแนนแปลง สูตรการคำนวณมีดังนี้ คือ

$$C = \sum \frac{(X_i - X^*)}{nk}$$

n คือ จำนวนของคะแนนดิบของกลุ่มสอบทานผล

k คือ จำนวนข้อสอบในแบบสอบร่วมที่ใช้

ค่า C ที่ได้ถ้ามีค่าน้อยจะให้ความหมายว่า รูปแบบการปรับเทียบคะแนนที่นำมาสร้างตารางปรับเทียบคะแนนนั้นมีความเหมาะสมและเพียงพอที่จะให้ผลการแปลงคะแนนอย่างคงเส้นคงวา

2.2 **สัมประสิทธิ์การทำนาย** (Cofficient of Prediction) สัมประสิทธิ์การทำนายเป็นค่าสหสัมพันธ์กำลังสอง (R^2) สำหรับการทำนายคะแนน y จากคะแนนปรับเทียบ y^* มีสูตรในการคำนวณ ดังนี้

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

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

2.3 **ดัชนีความแตกต่าง** (Discrepancy Index) Petersen และคณะ (1982) เสนอดัชนีความแตกต่างระหว่างคะแนนเกณฑ์กับคะแนนปรับเทียบ โดยถ่วงน้ำหนักด้วยค่าความแปรปรวน เพื่อให้ได้ค่าคงที่เป็นมาตรฐาน ดังนี้

$$C = \sum \frac{(X_{ij} - X_{kij^*})^2}{nS_{Xi}^2}$$

เมื่อ C แทน ค่าคะแนนเป็นดัชนีความแตกต่างในการปรับเทียบคะแนน

X_{ij} แทน คะแนนจากแบบสอบฉบับ i ของนักเรียนคนที่ j

X_{kij}^* แทน คะแนนจากแบบสอบฉบับ k ที่แปลงจากตารางปรับเทียบคะแนนไปเป็นคะแนนจากแบบสอบฉบับที่ i ของนักเรียนคนที่ j

n แทน จำนวนในกลุ่มตัวอย่างสอบทานผล

S_{Xi} แทน ความแปรปรวนคะแนนจากแบบสอบฉบับ X

วิธีการประเมินความเพียงพอที่ได้กล่าวมานี้ เป็นการประเมินผลของการปรับเทียบคะแนน อาศัยคะแนนเกณฑ์ที่เลือกสรรแล้วเป็นหลักในการเทียบหาความแตกต่าง สำหรับดัชนีที่

เสนอโดย Petersen และคณะ (1982) นั้น คะแนนเกณฑ์ที่ใช้ คือ ผลการแปลงคะแนนค่าดัชนีที่คำนวณมีลักษณะเป็นหน่วยมาตรฐานแล้ว สามารถนำค่าเหล่านี้ที่ได้จากการใช้รูปแบบที่ต่างกันตลอดจนสถานการณ์ที่ได้ข้อมูลจากผู้สอบเองเป็นเกณฑ์ในการหาความแตกต่าง ข้อมูลเหล่านี้ได้จากการออกแบบโดยการใช้กลุ่มตัวอย่างสอบทานผล ซึ่งผู้สอบในกลุ่มตัวอย่างนี้ได้รับการสอบด้วยแบบสอบทั้งฉบับ ดังนั้นการใช้คะแนนของตนเองเป็นเกณฑ์จึงมีความเป็นอิสระไม่ขึ้นกับกระบวนการแปลงคะแนนอื่น ๆ เช่น วิธีที่เสนอโดย Petersen และคณะ ดังนี้

		$C \leq (0.05\ SD_x)^2$	หมายถึง	ระดับน่าพอใจมาก
$(0.05\ SD_x)^2$	$<$	$C \leq (0.10\ SD_x)^2$	หมายถึง	ระดับน่าพอใจ
$(0.10\ SD_x)^2$	$<$	$C \leq (0.15\ SD_x)^2$	หมายถึง	ระดับน่าปานกลาง
$(0.15\ SD_x)^2$	$<$	$C \leq (0.20\ SD_x)^2$	หมายถึง	ระดับน่าไม่พอใจ
$(0.20\ SD_x)^2$	$<$	C	หมายถึง	ระดับน่าไม่พอใจมาก

ในการตรวจสอบประสิทธิภาพนั้น สามารถศึกษาได้จากความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนและความเพียงพอของการปรับเทียบคะแนน ซึ่งความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนนั้น พิจารณาจากองค์ประกอบที่สำคัญ 2 องค์ประกอบ คือ 1) ความลำเอียงของการปรับเทียบคะแนน และ 2) ความแปรผันเชิงสุ่มและความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน เป็นการอาศัยวิธีการทางสถิติในการตรวจสอบ ส่วนความเพียงพอของการปรับเทียบคะแนน วิธีการประเมินนั้นมีวิธีการปฏิบัติสองวิธี 1) ดูค่าดัชนีความแตกต่างของการปรับเทียบกับเกณฑ์ และ 2) ดูค่าดัชนีปรียบเทียบเปอร์เซ็นไทล์ ซึ่งค่าที่ได้มีค่าน้อยจะให้ความหมายว่า รูปแบบการปรับเทียบคะแนนที่นำมาสร้างตารางปรับเทียบคะแนนนั้นมีความเหมาะสมและเพียงพอที่จะให้ผลการแปลงคะแนนอย่างคงเส้นคงวา จากการตรวจสอบความเพียงพอของการปรับเทียบคะแนนนี้ ถือว่าเป็นการอาศัยดุลยพินิจในการตรวจสอบ

ดังนั้นในการวิจัยครั้งนี้ ผู้วิจัยได้หาประสิทธิภาพของการปรับเทียบคะแนนครั้งนี้ คือ การประเมินความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนและความเพียงพอของการปรับเทียบคะแนนโดยใช้เกณฑ์ของ Petersen และคณะ (1982) เพื่อช่วยในการตัดสินใจว่า ในการปรับเทียบคะแนนที่นำรูปแบบวิธีการปรับเทียบคะแนนและตัวแปรต่าง ๆ ที่ศึกษาไปใช้นั้น มีความเหมาะสมหรือไม่อย่างไร

ตอนที่ 2 หลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 ในกลุ่มสาระการเรียนรู้คณิตศาสตร์ ช่วงชั้นที่ 3

เนื่องจากหลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 ในกลุ่มสาระการเรียนรู้คณิตศาสตร์ ช่วงชั้นที่ 3 เป็นหลักสูตรแกนกลางของประเทศที่เริ่มใช้ในโรงเรียนทั่วประเทศที่จัดการศึกษาขั้นพื้นฐาน ในปีการศึกษา 2546 โดยเริ่มตั้งแต่ชั้นประถมศึกษาปีที่ 1 ประถมศึกษาปีที่ 4 มัธยมศึกษาปีที่ 1 และมัธยมศึกษาปีที่ 4 จนครบทุกชั้นในปีการศึกษา 2548 โดยได้กำหนดกรอบสาระและมาตรฐานการเรียนรู้เมื่อผู้เรียนจบการศึกษาขั้นพื้นฐานและเมื่อผู้เรียนจบในแต่ละช่วงชั้นไว้ สถานศึกษามีหน้าที่จัดทำสาระการเรียนรู้และกำหนดผลการเรียนรู้ที่คาดหวังรายปี หรือรายภาค โดยสาระการเรียนรู้ที่สถานศึกษากำหนดจากมาตรฐานการเรียนรู้ช่วงชั้นที่กำหนดไว้ในหลักสูตรเป็นสาระการเรียนรู้พื้นฐานสำหรับผู้เรียนทุกคน นอกจากนี้สถานศึกษาสามารถกำหนดสาระและมาตรฐานการเรียนรู้เพิ่มขึ้นเองให้เหมาะสมกับศักยภาพและความต้องการของผู้เรียนได้อีก สำหรับหลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 ในกลุ่มสาระการเรียนรู้ หลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 ในกลุ่มสาระการเรียนรู้คณิตศาสตร์ ช่วงชั้นที่ 3 มีสาระสำคัญ ดังต่อไปนี้

วิสัยทัศน์

การศึกษาคณิตศาสตร์สำหรับหลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 เป็นการศึกษาเพื่อปวงชนที่เปิดโอกาสให้เยาวชนทุกคนได้เรียนรู้คณิตศาสตร์อย่างต่อเนื่องและตลอดชีวิตตามศักยภาพ ทั้งนี้เพื่อให้เยาวชนเป็นผู้ที่มีความรู้ความสามารถทางคณิตศาสตร์ที่พอเพียง สามารถนำความรู้และทักษะกระบวนการทางคณิตศาสตร์ที่จำเป็นไปพัฒนาคุณภาพชีวิตให้ดีขึ้น รวมทั้งนำไปสร้างเครื่องมือในการเรียนรู้สิ่งต่าง ๆ และเป็นพื้นฐานสำหรับการศึกษาต่อ ดังนั้นจึงเป็นความรับผิดชอบของสถานศึกษาที่ต้องจัดสาระการเรียนรู้ที่เหมาะสมกับผู้เรียนแต่ละคน ทั้งนี้เพื่อให้บรรลุตามมาตรฐานการเรียนรู้ที่กำหนดไว้

สาระและมาตรฐานการเรียนรู้

สาระที่เป็นองค์ความรู้และมาตรฐานการเรียนรู้ที่จำเป็นสำหรับผู้เรียน เมื่อจบการศึกษาขั้นพื้นฐานกลุ่มสาระการเรียนรู้คณิตศาสตร์ มีดังนี้

สาระที่ 1 : จำนวนและการดำเนินการ

มาตรฐาน ค 1.1 : เข้าใจถึงความหลากหลายของการแสดงจำนวนและการใช้จำนวนในชีวิตจริง

มาตรฐาน ค 1.2 : เข้าใจจากผลที่เกิดขึ้นจากการดำเนินการของจำนวนและความสัมพันธ์ระหว่างการดำเนินการต่าง ๆ และสามารถใช้การดำเนินการในการแก้ปัญหาได้

มาตรฐาน ค 1.3 : ใช้การประมาณค่าในการคำนวณและแก้ปัญหาได้

มาตรฐาน ค 1.4 : เข้าใจในระบบจำนวนและสามารถนำสมบัติเกี่ยวกับจำนวนไปใช้ได้

สาระที่ 2 : การวัด

มาตรฐาน ค 2.1 : เข้าใจพื้นฐานเกี่ยวกับการวัด

มาตรฐาน ค 2.2 : วัดและคาดคะเนของสิ่งที่ต้องการวัด

มาตรฐาน ค 2.3 : แก้ปัญหาเกี่ยวกับการวัดได้

สาระที่ 3 : เรขาคณิต

มาตรฐาน ค 3.1 : อธิบายและวิเคราะห์รูปเรขาคณิตสองมิติและสามมิติได้

มาตรฐาน ค 3.2 : ใช้การนึกภาพ (Visualization) ใช้เหตุผลเกี่ยวกับปริภูมิ (Spatial Reasoning) และใช้แบบจำลองทางเรขาคณิต (Geometric Model) ในการแก้ปัญหาได้

สาระที่ 4 : พีชคณิต

มาตรฐาน ค 4.1 : อธิบายและวิเคราะห์รูปแบบ (Pattern) ความสัมพันธ์ และฟังก์ชันต่าง ๆ ได้

มาตรฐาน ค 4.2 : ใช้นิพจน์ สมการ อสมการ กราฟ และแบบจำลองทางคณิตศาสตร์อื่น ๆ แทนสถานการณ์ต่าง ๆ ตลอดจนแปลความหมายและนำไปใช้แก้ปัญหาได้

สาระที่ 5 : การวิเคราะห์ข้อมูลและความน่าจะเป็น

มาตรฐาน ค 5.1 : เข้าใจและใช้วิธีการทางสถิติในการวิเคราะห์ข้อมูลได้

มาตรฐาน ค 5.2 : ใช้วิธีการทางสถิติและความรู้เกี่ยวกับความน่าจะเป็นในการคาดการณ์ได้อย่างสมเหตุสมผล

มาตรฐาน ค 5.3 : ใช้ความรู้เกี่ยวกับสถิติและความน่าจะเป็นช่วยในการตัดสินใจและแก้ปัญหาได้

สาระที่ 6 : ทักษะ/กระบวนการทางคณิตศาสตร์

มาตรฐาน ค 6.1 : มีความสามารถในการแก้ปัญหา

มาตรฐาน ค 6.2 : มีความสามารถในการให้เหตุผล

มาตรฐาน ค 6.3 : มีความสามารถในการสื่อสาร การสื่อความหมายทางคณิตศาสตร์ และการนำเสนอ

มาตรฐาน ค 6.4 : มีความสามารถในการเชื่อมโยงความรู้ต่าง ๆ ทางคณิตศาสตร์และเชื่อมโยงคณิตศาสตร์กับศาสตร์อื่น

มาตรฐาน ค 6.5 : มีความคิดริเริ่มสร้างสรรค์

คุณภาพของผู้เรียนเมื่อจบช่วงชั้นที่ 3 (ชั้นมัธยมศึกษาปีที่ 1-3)

เมื่อผู้เรียนจบการเรียน ช่วงชั้นที่ 3 (ม.1-ม.3) ผู้เรียนควรที่จะมีความสามารถ ดังนี้

1. รู้ เข้าใจ มีทักษะการคำนวณจำนวนลบ เลขยกกำลัง จำนวนเต็ม กรณฑ์ที่ 2-3 จำนวนอตรรกยะ และสามารถนำความรู้เกี่ยวกับจำนวนไปใช้ในชีวิตจริงได้

2. รู้ เข้าใจ ทักษะการวัดพื้นที่ ปริมาตร รูปเรขาคณิต การแปลงคะแนน พร้อมทั้งนำความรู้เกี่ยวกับการวัดไปใช้ในชีวิตจริงได้

3. รู้ เข้าใจ มีทักษะในการใช้ทฤษฎีบทเรขาคณิตความเท่ากันของรูปสามเหลี่ยม ทฤษฎีบทปีทาโกรัส พร้อมทั้งสามารถนำความรู้เกี่ยวกับการวัดไปใช้ในชีวิตจริงได้

4. รู้ เข้าใจ ทักษะการแก้สมการ อสมการ กราฟ หรือแบบจำลองทางคณิตศาสตร์อื่นๆ ในการแก้ปัญหาได้

5. กำหนดประเด็นปัญหา และดำเนินการตามหลักวิชาสถิติเพื่อตอบปัญหา การใช้สถิติในการคาดการณ์และตัดสินใจ

6. สามารถนำความรู้และกระบวนการทางคณิตศาสตร์ไปใช้ในการเรียนรู้ในชีวิตจริงได้

7. มีความคิดริเริ่มสร้างสรรค์ในการทำงาน

จากคุณภาพของผู้เรียนเมื่อจบช่วงชั้นที่ 3 พบว่ามีสาระที่สำคัญที่จะนำไปสู่คุณภาพของผู้เรียน คือ สาระที่ 1 เรื่อง “จำนวนและการดำเนินการ” ดังนั้นผู้วิจัยจึงสนใจศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ในสาระที่ 1 เรื่อง “**จำนวนและการดำเนินการ**” ซึ่งมีมาตรฐานการเรียนรู้ ดังนี้

ตารางที่ 2.2 มาตรฐานการเรียนรู้ช่วงชั้น (ช่วงชั้นที่ 3) สาระที่ 1 : จำนวนและการดำเนินการ

มาตรฐานการเรียนรู้	มาตรฐานการเรียนรู้ช่วงชั้น ม.1-3
มาตรฐาน ค 1.1 : เข้าใจถึงความหลากหลายของการแสดงจำนวนและการใช้จำนวนในชีวิตจริง	1. มีความคิดรวบยอดเกี่ยวกับจำนวนเต็มบวก จำนวนเต็มลบ ศูนย์ และจำนวนตรรกยะ 2. รู้จักจำนวนอตรรกยะ และจำนวนจริง 3. เข้าใจเกี่ยวกับอัตราส่วน สัดส่วน ร้อยละ และนำไปใช้ในการแก้ปัญหา 4. เข้าใจเกี่ยวกับเลขยกกำลังที่มีเลขชี้กำลังเป็นจำนวนเต็ม และสามารถเขียนจำนวนเต็มและสามารถเขียนจำนวนให้อยู่ในรูปสัญกรณ์วิทยาศาสตร์ ($Ax10^n$ เมื่อ $1 \leq A<10$ และ เป็นจำนวนเต็ม) ได้ 5. เข้าใจเกี่ยวกับรากที่สองและรากที่สามของจำนวนจริง
มาตรฐาน ค 1.2 : เข้าใจจากผลที่เกิดขึ้นจากการดำเนินการของจำนวนและความสัมพันธ์ระหว่างการดำเนินการต่าง ๆ และสามารถใช้การดำเนินการในการแก้ปัญหาได้	1. บวก ลบ คูณ และหารจำนวนเต็มเศษส่วน ทศนิยม เลขยกกำลัง 2. หารากที่สองและรากที่สามของจำนวนเต็ม โดยการแยกตัวประกอบและนำไปใช้แก้ปัญหาได้ 3. อธิบายผลที่เกิดขึ้นจากการบวก การลบ การคูณ การหาร การยกกำลัง และการหารากของจำนวนเต็มและจำนวนตรรกยะ พร้อมทั้งบอกความสัมพันธ์ของการดำเนินการของจำนวนต่างๆ ได้ 4. ตระหนักถึงความสมเหตุสมผลของคำตอบที่ได้จากการคำนวณและการแก้ปัญหา
มาตรฐาน ค 1.3 : ใช้การประมาณค่าในการคำนวณและแก้ปัญหาได้	1. เข้าใจเกี่ยวกับการประมาณค่า และนำไปใช้แก้ปัญหาได้อย่างเหมาะสม 2. หารากที่สองและรากที่สามของจำนวนจริง โดยการประมาณ การเปิดตาราง หรือการใช้เครื่องคำนวณ และนำไปใช้แก้ปัญหาได้

มาตรฐานการเรียนรู้	มาตรฐานการเรียนรู้ช่วงชั้น ม.1-3
มาตรฐาน ค 1.4 : เข้าใจในระบบจำนวนและสามารถนำสมบัติเกี่ยวกับจำนวนไปใช้ได้	1. เข้าใจสมบัติต่างๆ เกี่ยวกับระบบจำนวนเต็ม และนำไปใช้แก้ปัญหาได้ 2. มีความคิดรวบยอดเกี่ยวกับจำนวนในระบบจำนวนจริง

ตอนที่ 3 งานวิจัยที่เกี่ยวข้องกับเทคนิคการปรับเทียบคะแนนและการศึกษาพัฒนาการ

3.1 งานวิจัยที่เกี่ยวข้องกับเทคนิคการปรับเทียบคะแนน

ปัจจุบันได้มีการศึกษาเทคนิคการปรับเทียบคะแนนอย่างต่อเนื่องและกว้างขวาง Fluour และคณะ (1999 cited by Bastari, 2000) ได้กล่าวว่าเป็นสิ่งท้าทายอย่างมากต่อการอ้างอิงข้อค้นพบเกี่ยวกับการปรับเทียบคะแนน ทั้งนี้เพราะมีความแตกต่างในความคลอบคลุมเนื้อหา ประเภทของข้อสอบ การดำเนินการสอบ และการใช้แบบสอบในสถานการณ์ที่เป็นไปได้ตลอดจนพัฒนาการเชื่อมต่อมาตรวัด เพื่อเปรียบเทียบการประเมินที่แตกต่างกัน วารสาร Applied Psychological Measurement , Applied Measurement in Education และ Journal of Education Measurement ซึ่งเป็นวารสารที่เป็นผู้นำทางการวัดผลการศึกษาและจิตวิทยา ได้เผยแพร่ประเด็นสำคัญของหัวข้อการศึกษาการปรับเทียบคะแนน การศึกษาส่วนใหญ่เป็นการเปรียบเทียบความแตกต่างของวิธีการแปลงคะแนน นอกจากนี้ Skagg และ Lissitz (1986) ได้ทบทวนวรรณคดีที่เกี่ยวข้องกับการปรับเทียบคะแนนภายใต้ทฤษฎี IRT อย่างกว้างขวาง และยังมีหนังสืออีก 3 เล่มที่เกี่ยวข้องเป็นพิเศษ (Holland & Rubin, 1982; Kolen & Brennan, 1995; ศิริชัย กาญจนวาสี, 2541) จึงเห็นได้ว่า การศึกษาการปรับเทียบคะแนนได้มุ่งประเด็นไปยังข้อสอบที่ตรวจให้คะแนนสองค่า (Dichotomous) ถึงแม้ว่าเทคนิคที่ได้จะไม่ใช่วิธีการที่ดีที่สุด เนื่องจากผลที่ได้จากการศึกษาวิธีการปรับเทียบคะแนนในแต่ละบริบทยังคงแตกต่างกัน ขึ้นอยู่กับวัตถุประสงค์และสถานการณ์เฉพาะ โดยทั่วไปวิธีการภายใต้ทฤษฎีตอบสนองข้อสอบ ให้ผลที่ดีกว่าเมื่อเปรียบเทียบกับวิธีการปรับเทียบคะแนนภายใต้ทฤษฎีการวัดแบบดั้งเดิม แต่ก็ไม่ได้เป็นเช่นนั้นทุกสถานการณ์ (Petersen, Cook & Stocking, 1983) ขึ้นอยู่กับบริบทที่ทำการศึกษา ดังรายละเอียดที่จะกล่าวต่อไป

งานวิจัยการปรับเทียบคะแนนต่างประเทศ

งานวิจัยที่เกี่ยวข้องกับการปรับเทียบคะแนนที่ผ่านมานั้น ส่วนใหญ่ทำการเปรียบเทียบผลการปรับเทียบคะแนน โดยใช้วิธีการปรับเทียบคะแนนแบบต่างๆ นอกจากนี้ได้มีการนำวิธีการปรับเทียบคะแนนไปประยุกต์ใช้ในการศึกษาความก้าวหน้าทางการเรียน ซึ่งผู้วิจัยได้เสนอตามลำดับ ดังนี้

Slinde และ Linn (1977–1979) ได้ทำการตรวจสอบปัญหาของการปรับเทียบคะแนนตามแนวตั้ง (Vertical Equating) ของแบบสอบสองชุดที่สร้างขึ้นเพื่อใช้กับประชากรที่มีความสามารถต่างกันด้วยวิธีอ้อม

จากการศึกษาได้ให้ข้อเสนอแนะว่า การใช้รูปแบบการปรับเทียบคะแนน 3 รูปแบบ คือ เชิงเส้นตรง อิควิเปอร์เซ็นไทล์ ทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 1 พารามิเตอร์ อาจมีข้อจำกัดในกระบวนการปรับเทียบคะแนนในแนวตั้ง โดยเฉพาะอย่างยิ่ง เมื่อทำการปรับเทียบคะแนนโดยใช้แบบสอบสองชุดที่มีความแตกต่างกันมาก และกลุ่มตัวอย่างสองกลุ่มที่มีระดับความสามารถต่างกันมาก จากการศึกษานี้ได้ให้ข้อสังเกตว่า ถ้าใช้การปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบชนิด 3 พารามิเตอร์ อาจให้ผลการปรับเทียบคะแนนที่ดีกว่าในสถานการณ์เช่นนั้น

Marco, Petersen และ Stewart (1979) ได้ประเมินและเปรียบเทียบวิธีการปรับเทียบคะแนนเชิงเส้นตรง วิธีการปรับเทียบคะแนนแบบอิควิเปอร์เซ็นไทล์ และวิธีการใช้โค้งลักษณะข้อสอบโดยใช้แบบสอบ SAT

ผลจากการวิจัยพบว่าแบบสอบที่มีความยากแตกต่างกัน วิธีการปรับเทียบคะแนนโดยใช้โค้งลักษณะข้อสอบให้ผลดีที่สุด และวิธีการปรับเทียบคะแนนเชิงเส้นตรงให้ผลน้อยที่สุด และวิธีการปรับเทียบคะแนนให้ผลน้อยที่สุด และเมื่อ Petersen และคณะทำซ้ำอีก ในปี ค.ศ. 1983 โดยวิเคราะห์ซ้ำกับปี ค.ศ. 1979 ผลคงปรากฏเหมือนเดิม แต่เมื่อ Kolen (1981) ได้ศึกษารูปแบบการปรับเทียบคะแนนระหว่างรูปแบบดั้งเดิม 2 วิธี คือรูปแบบเชิงเส้นตรงและรูปแบบอิควิเปอร์เซ็นไทล์ กับรูปแบบการปรับเทียบคะแนนโดยใช้โค้งลักษณะข้อสอบที่ศึกษาทั้ง 1, 2 และ 3 พารามิเตอร์ แต่ละวิธีเทียบด้วยค่าประมาณจากค่าจริง (Estimated True Score Equating) และเทียบด้วยการประมาณจากค่าสังเกต (Estimated Observed Score Equating) ใช้ข้อมูลจากโครงการ IOWA Test of Education Development (ITED) ค.ศ. 1978 ศึกษา 2 รูปแบบ คือ เทียบจากฟอร์มที่มีความยากเท่าเทียมกัน และเทียบกับฟอร์มที่มีความยากแตกต่างกัน กลุ่มตัวอย่างเป็นนักเรียนระดับที่ 1 คือ ระดับเกรด 9 และ 10 ระดับ 2 คือ ระดับเกรด 11 และ 12 จำนวน

ทั้งหมด 10,728 คน จากโรงเรียนในรัฐไอโอวา 34 แห่ง การศึกษานี้ใช้เกณฑ์ Cross-Validation Criterion

ผลการวิจัยพบว่า การปรับเทียบคะแนนทั้งหมดแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และ .05 การปรับเทียบคะแนนโดยใช้วิธีอิควิเปอร์เซ็นไทล์เหมาะสมที่สุดกับแบบสอบที่มีความยากแตกต่างกัน ส่วนการปรับเทียบคะแนนโดยใช้โค้งลักษณะข้อสอบ พบว่า กรณีใช้พารามิเตอร์เดียว ยังเป็นวิธีที่ไม่ถูกต้องสำหรับแบบสอบที่ต่างกัน ซึ่งเนื่องมาจากการเดาเกิดขึ้นมาก ส่วนกรณีใช้ 3 พารามิเตอร์ มีปัญหาที่เกิดขึ้นอย่างหลีกเลี่ยงไม่ได้ คือ การประมาณค่าไม่ถูกต้องของ Lower Asymtote Parameter

Cook, Dunbar และ Eignor (1981) ได้ศึกษาผลการปรับเทียบคะแนนตามทฤษฎี IRT เปรียบเทียบกับการปรับเทียบแบบดั้งเดิม 2 วิธี คือ การปรับเทียบเชิงเส้นตรง และการปรับเทียบแบบอิควิเปอร์เซ็นไทล์ การปรับเทียบตามทฤษฎี IRT โมเดล 3 พารามิเตอร์ ซึ่งประมาณค่าพารามิเตอร์โดยใช้โปรแกรม LOGIST ด้วยวิธีความเป็นไปได้สูงสุด (Maximum Lielihood Procedure) สำหรับการประมาณค่าที่อยู่ต่ำกว่าระดับการเดา ใช้วิธีการเชิงเส้นตรง (Linear Interpolation) ในการปรับเทียบ ศึกษาทั้งสองกรณี คือ มีแบบสอบร่วม และไม่มีแบบสอบร่วม โดยใช้ประชากรที่ไม่ได้สุ่มมาจากประชากรเดียวกัน ทั้งนี้เพื่อศึกษาความเป็นไปได้ (Feasibility) ของวิธีการที่ใช้ทฤษฎีตอบสนองข้อสอบ ในการนำไปแก้ปัญหาทางปฏิบัติ การเปรียบเทียบพิจารณา 2 ประการ คือ ความสอดคล้องสัมพันธ์ (Relative agreement) ระหว่างวิธีที่ใช้ทฤษฎีตอบสนองข้อสอบกับวิธีดั้งเดิมแต่ละวิธี โดยพิจารณาจากการพล็อตกราฟ อีกประการหนึ่งคือ ดัชนีความแตกต่างสำหรับการแจกแจงคะแนนรวม และแต่ละส่วนของการแจกแจง (สูงกว่า 20 % , ตรงกลาง 60 % และต่ำกว่า 20 % ของการแจกแจง) ซึ่งคำนวณจากค่าที่ถ่วงน้ำหนักของค่าเฉลี่ยกำลังสองของความแตกต่างระหว่างค่าประมาณจากการปรับเทียบคะแนนแต่ละวิธีที่ศึกษากับคะแนนเกณฑ์ เกณฑ์ที่ใช้ศึกษาในที่นี้ คือ คะแนนแปลงด้วยการปรับเทียบคะแนนที่ใช้ทฤษฎีตอบสนองข้อสอบชนิด 3 พารามิเตอร์ ซึ่งงานวิจัยที่มีมาก่อน (Kolen, 1981; Slind & Linn, 1977) แนะนำว่ามีความเหมาะสมในการปรับเทียบคะแนนในกรณีที่ผู้สอบมีความยากต่างกัน สอบกับกลุ่มตัวอย่างที่มีความสามารถต่างกัน

ข้อมูลที่ใช้ในการศึกษาครั้งนี้ เป็นผลการสอบ 2 ครั้งที่ดำเนินการแล้ว โดย The Collage Board Admissions Testing Program ซึ่งใช้แบบสอบ 2 ชุด ที่มีความแตกต่างกันในความยาวและความยาก แต่ละชุดมีข้อสอบ 2 ตอน คือ ข้อสอบภาษาและคณิตศาสตร์ ใช้กับกลุ่มตัวอย่างที่ไม่ได้เท่าเทียมกันโดยการสุ่ม

ผลการวิจัยพบว่า วิธีการปรับเทียบคะแนนแบบดั้งเดิม กับวิธีการปรับเทียบตามทฤษฎี IRT มีความสอดคล้องกันมาก เนื่องจากโดยความจริงแล้ว การแจกแจงคะแนนดิบจากแบบสอบทั้งสองชุดมีรูปร่างคล้ายคลึงกันมาก จึงทำให้รูปแบบเชิงเส้นตรงและเชิงเส้นโค้งหรืออิควิเปอร์เซ็นไทล์มีผลใกล้เคียงกัน โดยการเทียบเชิงเส้นโค้งมีผลใกล้เคียงกับตามวิธีทฤษฎีตอบสนองข้อสอบมากกว่าการปรับเทียบเชิงเส้นตรง อย่างไรก็ตาม ความแตกต่างที่เกิดขึ้นที่ส่วนปลายของการปรับเทียบแบบดั้งเดิมทุกวิธีที่ต่างไปจากการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ เพราะไม่มีผลเนื่องจากการขาดแคลนข้อมูลในส่วนปลายของการแจกแจง อย่างไรก็ตาม การปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ ก็ไม่สามารถเตรียมการปรับเทียบคะแนนในส่วนที่อยู่ปลายด้านล่างที่ต่ำกว่าระดับการเดาได้ ซึ่งจำเป็นต้องใช้วิธีอื่นประมาณค่าเพิ่มเติม

นอกจากนี้ผลการวิจัยพบว่า การปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ โดยตรงเมื่อไม่มีแบบสอบร่วม และเมื่อกลุ่มตัวอย่างไม่ได้เป็นกลุ่มตัวอย่างที่เท่าเทียมกันโดยการสุ่มนั้นให้ผลแตกต่างจากการปรับเทียบเมื่อมีแบบสอบร่วมเพียงเล็กน้อย

Kolen (1981 อ้างใน ภัทราพร เกษสังข์, 2546) ได้เปรียบเทียบผลการปรับเทียบคะแนนระหว่างวิธีดั้งเดิม 2 วิธี กับทฤษฎีตอบสนองข้อสอบ 7 วิธีโดยใช้ข้อมูลจากการทดสอบนักเรียนในรัฐไอโอวา ปี ค.ศ. 1978 ในโครงการปรับเทียบคะแนนชื่อ The Iowa Test of Education Development (ITED) โดยทำการศึกษาทั้งสองแบบ คือ การปรับเทียบคะแนนจากแบบฉบับที่มีความยากเท่ากันและการปรับเทียบคะแนนจากฉบับที่มีความยากต่างกัน แบบสอบที่ใช้ คือ แบบสอบที่พิมพ์ใหม่ครั้งที่ 7 เทียบไปสู่แบบสอบฉบับเก่าครั้งที่ 6 ในแบบสอบที่พิมพ์ครั้งที่ 6 มีฉบับหนึ่ง คือ X-6 ใช้ทดสอบกับนักเรียนทั้งสองระดับ คือ ระดับที่ 1 เป็นนักเรียนเกรด 9 และ 10 และระดับที่ 2 เป็นนักเรียนเกรด 11 และ 12 ส่วนแบบสอบที่พิมพ์ครั้งที่ 7 มีสองฉบับที่คู่ขนานกัน คือ X-7 และ Y-7 แต่ละฉบับแบ่งแบบสอบออกเป็นสองระดับ ซึ่งมีความยากแตกต่างกัน เพื่อแยกสอบนักเรียนแต่ละคน โดยมีแบบแผนการรวบรวมข้อมูลดังนี้

นักเรียนระดับที่ 1 สอบแบบสอบ X-6, X-7 ระดับ 1 และ Y-7 ระดับ 1

นักเรียนระดับที่ 2 สอบแบบสอบ X-6, X-7 ระดับ 2 และ Y-7 ระดับ 2

ทั้งนี้นักเรียนแต่ละคนจะได้รับแบบสอบเพียง 1 ฉบับ คือ X-6, X-7 หรือ Y-7 โดยการสุ่ม ดังนั้นจำนวนนักเรียนที่สอบแบบสอบแต่ละฉบับจะมี 1 ใน 3 จากการสอบแบบสอบแต่ละระดับ สำหรับกลุ่มสอบทานผลใช้ข้อมูลจากนักเรียนทุก ๆ คนที่ 3 จากการสอบแต่ละระดับ ซึ่งการปรับเทียบคะแนนแบบดั้งเดิม ใช้วิธีการแบบเชิงเส้นตรงและแบบอิควิเปอร์เซ็นไทล์ ซึ่งจัดกระทำตามวิธีการของ Angoff (1971) ส่วนการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ ใช้รูปแบบ

โลจิสติก ชนิด 1, 2 และ 3 พารามิเตอร์ โดยแต่ละรูปแบบเทียบสองแบบ คือ เทียบด้วยค่าประมาณคะแนนจริงและเทียบด้วยค่าประมาณคะแนนสังเกต ส่วนการเทียบอีกวิธีหนึ่ง คือ รูปแบบราสช์

ดัชนีสำหรับตรวจสอบผล Kolen ใช้วิธีคำนวณจากค่าเฉลี่ยกำลังสองของความแตกต่าง (สำหรับผู้สอบแบบทดสอบ X-6 ในกลุ่มตรวจสอบทานผล) ระหว่างคะแนนจากแบบสอบ X-6 กับแบบสอบ X-7 และ Y-7 ที่แปลงคะแนนแล้ว ณ ตำแหน่งที่เปอร์เซ็นไทล์เดียวกันโดยใช้วิธีการปรับเทียบคะแนนแบบต่างๆ ทั้ง 9 วิธี หากค่าดัชนีจากกลุ่มสอบทานผลที่ได้จากการปรับเทียบคะแนนใดมีค่าน้อยกว่า แสดงว่า วิธีนั้นมีความสอดคล้องระหว่างการแจกแจงคะแนนของแบบสอบฉบับเก่ากับคะแนนของแบบสอบฉบับใหม่ที่แปลงไปในสเกลของแบบสอบฉบับเก่ามากกว่า นั่นคือ วิธีการนั้นมีความคงเส้นคงวามากกว่า ทั้งนี้ทดสอบความแตกต่างของค่าดัชนีที่ได้จากการปรับเทียบคะแนนแบบต่าง ๆ เพื่อเปรียบเทียบผลของแต่ละวิธี โดยใช้ค่าสถิติฟรายด์แมน (Friedman Test)

ผลการวิจัยพบว่า วิธีการปรับเทียบคะแนนแบบต่างๆ ให้ผลแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และ .05 เมื่อใช้แบบสอบฉบับที่ 1 และ 2 ตามลำดับ เมื่อพิจารณาแบบสอบระดับที่ 1 วิธีการปรับเทียบคะแนนตามวิธีทฤษฎีตอบสนองข้อสอบ ชนิด 3 พารามิเตอร์ และวิธีการเชิงเส้นตรงให้ผลแน่นอนลดลงตามลำดับ ส่วนการสอบแบบสอบระดับที่ 2 วิธีที่ให้ผลแน่นอนที่สุด คือ วิธีการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ ชนิด 3 พารามิเตอร์ โดยกะประมาณคะแนนจริงและให้ผลแน่นอนถัดมา คือ รูปแบบวิธีเชิงเส้นตรง และวิธีแบบทฤษฎีตอบสนองข้อสอบ ชนิด 1 พารามิเตอร์ ตามลำดับ และ Kolen ได้อภิปรายเสนอ สรุปว่า วิธีที่ให้ผลการปรับเทียบคะแนนที่มีความเพียงพอน้อยที่สุด คือ วิธีตามวิธีทฤษฎีตอบสนองข้อสอบ ชนิด 1 พารามิเตอร์ ซึ่งอาจจะมาจากข้อสอบที่มีค่าความยากต่างกัน การปรับเทียบคะแนนอาจจะได้รับผลกระทบและให้ประสิทธิภาพต่ำลง

Kolen และ Whitney (1982) ได้เปรียบเทียบความถูกต้องของการปรับเทียบคะแนนโดยใช้แบบสอบร่วม 4 วิธี คือ วิธีอิควิเปอร์เซ็นไทล์ วิธีเชิงเส้นตรง วิธีทฤษฎีตอบสนองข้อสอบแบบ 1 และ 3 พารามิเตอร์ ใช้แบบสอบ General Education Development (GRE) ซึ่งเป็นแบบสอบผลสัมฤทธิ์ เพื่อใช้ในการตัดสินให้ประกาศนียบัตร ที่ต้องการเทียบความรู้ระดับเตรียมอุดมศึกษาทั่วประเทศ แบบสอบมี 12 ชุด โดยใช้ฟอร์มที่ 12 เป็นแบบสอบร่วม ที่เหลืออีก 11 ฟอร์ม เป็น แบบสอบที่ใช้เทียบสเกลผู้เข้าสอบแต่ละคนต้องทำแบบสอบ 2 ฟอร์ม ในจำนวน 11 พอร์ม ส่วนฟอร์มที่ 12 ต้องทำทุกคน กลุ่มตัวอย่างมาจากการสอบปี 1980 จำนวนมากกว่า 800,000 คน ทำการสุ่มแบบแบ่งชั้นหลายชั้นจากประเภทของโรงเรียน เขตภูมิศาสตร์ สถานภาพ

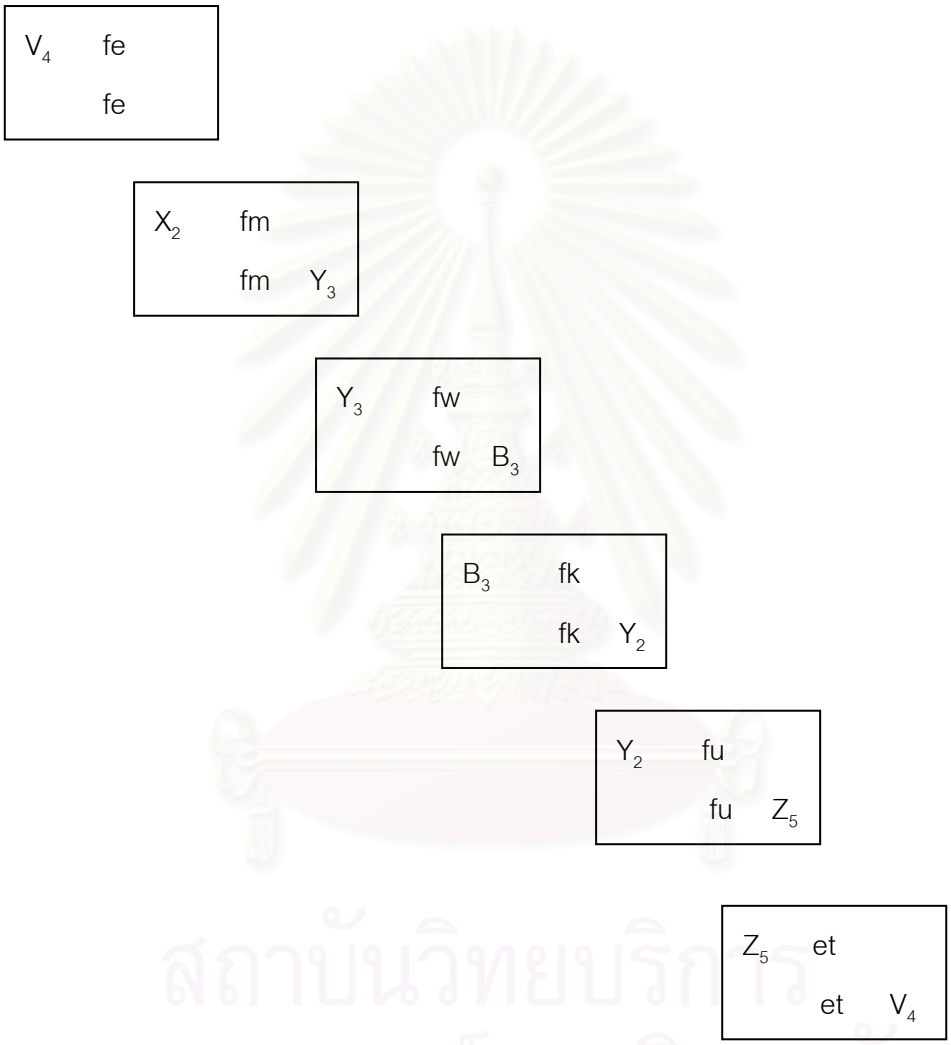
ทางสังคม สุดท้ายได้นักเรียนที่เป็นตัวอย่างโรงเรียนละ 22 คน จำนวนคนในแต่ละชุดของแบบสอบประมาณ 200 คน ประมาณค่าพารามิเตอร์โดยใช้โปรแกรม LOGIST แล้วสร้างตารางปรับเทียบคะแนนระหว่างแบบสอบอีก 11 ชุด การปรับเทียบคะแนนโดยใช้อิควิเปอร์เซ็นไทล์โดยใช้โปรแกรม LOGIST แล้วสร้างตารางปรับเทียบคะแนนระหว่างแบบสอบอีก 11 ชุด การปรับเทียบคะแนนโดยใช้อิควิเปอร์เซ็นไทล์และใช้โค้งลักษณะข้อสอบ ได้ผลไม่เป็นที่ยอมรับ ขณะเดียวกันวิธีการปรับเทียบคะแนนโดยใช้พารามิเตอร์ตัวเดียวให้ผลเพียงพอเทียบเท่ากับรูปแบบสามพารามิเตอร์ ในการประเมินผลการศึกษาครั้งนี้มีการเปรียบเทียบอื่นๆ ซึ่งสรุปได้ว่า ความถูกต้องของการปรับเทียบคะแนนขึ้นอยู่กับองค์ประกอบหลายอย่าง เช่น ลักษณะของแบบสอบ รูปแบบของการปรับเทียบคะแนน ขนาดของกลุ่มตัวอย่าง เป็นต้น

Hutten (1982) ได้ศึกษาความเหมาะสมของข้อมูลจริงกับรูปแบบทฤษฎีตอบสนองข้อสอบ 2 รูปแบบ คือ รูปแบบราส์ช กับรูปแบบ 3 พารามิเตอร์ ในการประมาณค่าพารามิเตอร์ความสามารถ และความยากของข้อสอบ สำหรับกลุ่มตัวอย่างขนาดเล็ก (250 คน) และแบบสอบสั้น (20 ข้อ)

ผลการวิจัยพบว่า รูปแบบของราส์ชและรูปแบบ 3 พารามิเตอร์ มีความเหมาะสมกับข้อมูลประมาณร้อยละ 80 ทั้งสองรูปแบบ เมื่อเปรียบเทียบคะแนนที่ทำนายกับคะแนนที่วัดได้ โดยใช้สถิติ Kolmogorov–Smirnov แล้ว รูปแบบของราส์ชเหมาะสมกับข้อมูลทั้งหมดดีกว่ารูปแบบ 3 พารามิเตอร์ ผู้วิจัยสนับสนุนให้ใช้รูปแบบ 3 พารามิเตอร์ กับกลุ่มตัวอย่างที่มีขนาดตั้งแต่ 1,000 คนขึ้นไป เพื่อจะได้ความแม่นยำในการประมาณค่าพารามิเตอร์ กับกลุ่มตัวอย่างตั้งแต่ 1,000 คนขึ้นไป เพื่อจะได้ความแม่นยำในการประมาณค่าพารามิเตอร์ต่าง ๆ ผลการวิจัยนี้สอดคล้องกับที่ Douuglass (1981) ได้ศึกษาเปรียบเทียบการใช้รูปแบบทฤษฎีตอบสนองข้อสอบกับผลการสอบในระดับห้องเรียน โดยใช้รูปแบบทฤษฎีตอบสนองข้อสอบประมาณค่าพารามิเตอร์ของข้อสอบ 100 ข้อ ที่เป็นข้อสอบปลายปีของ 4 ปีการศึกษา จากกลุ่มตัวอย่างระดับมหาวิทยาลัยจำนวน 594-1,082 คน จัดข้อสอบแบ่งเป็น 4 ชุด จากนั้นทำการสุ่มข้อสอบชุดละ 43-53 ข้อ สุ่มผู้สอบจำนวน 200, 600 และ 800 คน มาวิเคราะห์ ปรากฏว่า การวิเคราะห์ข้อสอบด้วยรูปแบบของราส์ชมีความคงที่ของค่าพารามิเตอร์ดีกว่ารูปแบบ 3 พารามิเตอร์ และเมื่อมีการปรับเทียบคะแนนของแบบสอบข้ามกลุ่มตัวอย่าง รูปแบบของราส์ชก็มีความคงที่ดีกว่าด้วย

Lord และ Wingersky (1984 อ้างใน พรพิมล นาคเวช, 2537) ได้เปรียบเทียบการปรับเทียบคะแนนระหว่างรูปแบบตามวิธีทฤษฎีตอบสนองข้อสอบ ที่เทียบโดยใช้คะแนนจริง และรูปแบบตามวิธีทฤษฎีตอบสนองข้อสอบ ที่ปรับเทียบโดยใช้คะแนนจริง และรูปแบบตามทฤษฎี IRT

ที่ปรับเทียบคะแนนสังเกตโดยใช้อิควิเปอร์เซ็นไทล์ จากแบบสอบ SAT ด้านภาษา จำนวน 6 ชุด เทียบเข้าสู่กันในลักษณะลูกโซ่ โดยใช้แบบสอบร่วม ทั้งนี้ แบบสอบชุดแรกและชุดสุดท้ายเป็นชุดเดียวกัน การออกแบบรวบรวมข้อมูลดำเนินการ ดังแผนภาพที่ 2.5



แผนภาพที่ 2.5 การออกแบบรวบรวมข้อมูลของ Lord และ Wingersky (1984)

จากแผนภาพที่ 2.6 แบบสอบ V_4 เทียบไปสู่แบบสอบ X_2 โดยใช้แบบสอบร่วม fe ส่วนแบบสอบ X_2 เทียบไปสู่แบบสอบ Y_3 โดยใช้แบบสอบร่วม fm ซึ่งจะได้ผลการเทียบจาก V_4 ไป Y_3 กระทำการปรับเทียบคะแนนในลักษณะเดียวกันไปยังแบบสอบชุดอื่น สุดท้ายจะได้ผลการเทียบ V_4 กับ V_4 ซึ่งเป็นการปรับเทียบแบบสอบกับตัวเอง และใช้เกณฑ์ในการเปรียบเทียบความ

คงเส้นคงวาของวิธีการปรับเทียบคะแนนทั้ง 2 วิธี ทั้งนี้เกณฑ์นี้แตกต่างจากที่ Kolen (1981) ได้เคยศึกษาความคงเส้นคงวาของวิธีการปรับเทียบมาตราแบบต่าง ๆ Kolen ใช้เกณฑ์จากกลุ่มสอบทานผล ซึ่งเป็นกลุ่มอิสระอีกกลุ่มหนึ่ง โดยผู้วิจัยกล่าวไว้ว่า เกณฑ์ที่ Kolen ใช้ ไม่ใช่เกณฑ์ที่มีคุณสมบัติสำหรับการเลือกวิธีการปรับเทียบคะแนนที่ดีที่สุด เพราะวิธีการปรับเทียบคะแนนที่ไม่ถูกต้อง อาจให้ผลการเทียบคงเส้นคงวามากกว่าวิธีการที่ถูกต้อง เกณฑ์การปรับเทียบแบบสอบกับตัวเองที่ใช้ในการศึกษาครั้งนี้อาจไม่เที่ยงตรงในเชิงทฤษฎี และไม่น่าพอใจในการปฏิบัติสำหรับการเทียบแบบสอบสองฉบับที่แตกต่างกัน แต่เป็นเกณฑ์ที่มีประโยชน์ต่อการเปลี่ยนวิธีการปรับเทียบคะแนนใหม่ หากพบว่าคะแนนที่แปลงแล้วในขั้นสุดท้ายไม่เป็นอย่างเดียวกันกับคะแนนเดิมของตนเอง

ผลการวิจัยยังไม่มีการตัดสินใจเชิงทฤษฎีที่ชัดเจนในการเลือกวิธีการปรับเทียบคะแนนแบบทฤษฎีตอบสนองข้อสอบที่ใช้คะแนนจริง กับแบบทฤษฎีตอบสนองข้อสอบที่ใช้คะแนนสังเกตโดยอิควิเปอร์เซ็นไทล์ เนื่องจากทั้งสองวิธีมีความสอดคล้องกัน

Dorrans และ Kingston (1985) ได้ศึกษาผลการฝ่าฝืนเกี่ยวกับความเป็นมิติเดียว ในการประมาณค่าพารามิเตอร์ของข้อสอบ และพารามิเตอร์ความสามารถ โดยใช้การปรับเทียบคะแนนแบบสามพารามิเตอร์กับแบบสอบวัดความถนัดชุด GRE ด้านภาษา จำนวน 4 ชุด แต่ละชุดวัด 2 ด้าน คือ ความเข้าใจในการอ่าน และหลักภาษา การออกแบบการปรับเทียบคะแนนมี 2 แบบ คือ เทียบคะแนนรวมทั้งชุดซึ่งมีความเป็นวิวิธพันธ์ (Heterogeneous) และเทียบคะแนนรวมแต่ละด้านซึ่งมีความเป็นเอกพันธ์ (Homogeneous) โดยการออกแบบการรวบรวมข้อมูล เมื่อใช้กลุ่มที่เท่าเทียมกัน และเมื่อใช้แบบสอบร่วม จำนวนกลุ่มตัวอย่างมีตั้งแต่ 2,579 ถึง 4,351 คน

ผลการวิจัยพบว่า การปรับเทียบคะแนนตามวิธีทฤษฎีตอบสนองข้อสอบ เมื่อฝ่าฝืนข้อตกลงเบื้องต้นเกี่ยวกับแบบสอบต้องมิติเดียวนั้น อาจมีผลกระทบ (Impact) ต่อการปรับเทียบคะแนน แต่ไม่อาจยืนยันได้ เนื่องจากมีความคล้ายคลึงกัน ระหว่างผลการปรับเทียบเมื่อข้อสอบเป็นเอกพันธ์และวิวิธพันธ์ ดังนั้นผู้วิจัยจึงแนะนำว่าการปรับเทียบคะแนนโดยใช้ทฤษฎีตอบสนองข้อสอบมีความแกร่งเพียงพอต่อการฝ่าฝืนความเป็นมิติเดียว

Averye (1992) ได้ศึกษาเปรียบเทียบประสิทธิภาพของการปรับเทียบคะแนนตามแนวตั้งด้วยวิธีอิควิเปอร์เซ็นไทล์ และวิธี 3 พารามิเตอร์ โดยใช้การจำลองข้อมูลภายใต้เงื่อนไขของขนาดกลุ่มตัวอย่าง ความยาวของแบบสอบ และความยากของแบบสอบร่วมการปรับเทียบทำ 3 กรณี คือ 1) เปรียบเทียบประสิทธิภาพของวิธีการปรับเทียบคะแนนทั้งสองภายใต้เงื่อนไขทุกเงื่อนไข 2) เปรียบเทียบประสิทธิภาพของวิธีการปรับเทียบคะแนนแต่ละวิธีภายใต้เงื่อนไขทุกเงื่อนไข 3) ตรวจสอบผลของตัวแปรอิสระแต่ละตัวในการปรับเทียบคะแนนแต่ละวิธี ขนาดกลุ่ม

ตัวอย่างที่ใช้มี 3 ขนาด คือ 200, 500 และ 1,000 ความยาวของข้อสอบมี 2 ขนาด คือ 30 ข้อ และ 60 ข้อ ส่วนความยาวของแบบทดสอบร่วมมี 2 ขนาด คือ 5 ข้อ และ 10 ข้อ สถิติใช้เป็นเกณฑ์การประเมินมี 2 อย่าง คือ ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนซึ่งถ่วงน้ำหนักแล้ว (Weighted Mean Square Error : WMSE) และค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนซึ่งไม่ถ่วงน้ำหนัก (UMSE)

ผลการวิจัยพบว่า การปรับเทียบคะแนนตามเงื่อนไขต่าง ๆ โดยใช้วิธีการปรับเทียบคะแนนทั้งสองวิธีไม่แตกต่างกัน สำหรับการปรับเทียบคะแนนโดยใช้วิธีอิควิเปอร์เซ็นไทล์ พบว่า ความยาวของแบบสอบกับความยาวของแบบสอบร่วมเป็นองค์ประกอบสำคัญ ส่วนการปรับเทียบคะแนน รูปแบบ 3 พารามิเตอร์ พบว่า ขนาดกลุ่มตัวอย่างเป็นองค์ประกอบสำคัญ ขนาดตัวอย่างเล็ก (200) มีแนวโน้มที่จะให้ผลการปรับเทียบคะแนนที่ไม่ถูกต้อง ในขณะที่ขนาดกลุ่มตัวอย่างที่ใหญ่ขึ้น (500 และ 1,000) จะให้ผลการปรับเทียบคะแนนที่มีความถูกต้องมากกว่า

Glowacki (1991) ได้ตรวจสอบโมเดลของการปรับเทียบคะแนนที่มีความเหมาะสมกับการสอบของบัณฑิตวิทยาลัยแห่งมหาวิทยาลัยอลาบามา ปัญหาในการวิจัยคือโมเดลของการปรับเทียบที่ตรวจสอบ มีการแจกแจงคะแนนดิบ หรือคะแนนที่ผ่านจากการสอบแบบสอบการอ่าน และคณิตศาสตร์ แตกต่างกันหรือไม่ โมเดลที่ใช้ในการตรวจสอบ คือ โมเดลเชิงส้นตรง อิควิเปอร์เซ็นไทล์ และ IRT ชนิด 1,2 และ 3 พารามิเตอร์

ผลการวิจัยพบว่า วิธีการปรับเทียบคะแนนทั้ง 5 โมเดล ในการสอบการอ่านและคณิตศาสตร์ให้ผลที่คล้ายคลึงกันแสดงว่าโมเดลทั้งหมดสามารถนำมาใช้กับการปรับเทียบคะแนนได้ โดยไม่มีโมเดลใดดีที่สุด

Kim และ Cohen (1995) ได้ศึกษาเปรียบเทียบวิธีการปรับเทียบคะแนนระหว่างแบบสอบที่มีการตรวจให้คะแนนแบบพหุวิภาค ตามโมเดล Graded Response Model วิธีประมาณค่าไคว์สแคว์ต่ำสุด (A Minimum χ^2 Method) 2 วิธี คือ วิธี Diagonal χ^2 Method กับวิธี Full χ^2 Method และ วิธีโค้งคุณลักษณะข้อสอบ ใช้สถานการณ์จำลองข้อมูลศึกษาปรับเทียบแนวนอน ออกแบบการเก็บข้อมูลใช้รูปแบบผู้สอบกลุ่มเท่าเทียมกัน ใช้แบบสอบกลุ่มละ 30 ข้อ กลุ่มตัวอย่าง 300 คน พบว่า สัมประสิทธิ์การปรับเทียบคะแนนหรือค่าคงที่จากทั้งสามวิธีมีค่าใกล้เคียงกัน นอกจากนี้ใช้ในการปรับเทียบในแนวตั้ง ออกแบบการเก็บข้อมูลรูปแบบใช้ข้อสอบร่วมภายใน พบว่า สัมประสิทธิ์การปรับเทียบคะแนนหรือค่าคงที่ที่ได้จากทั้ง 3 วิธี มีค่าใกล้เคียงกันด้วย

Yang (1997) ได้ทำการศึกษาผลของคุณลักษณะของแบบสอบ ในความถูกต้องของการปรับเทียบคะแนน ที่ใช้วิธีการปรับเทียบคะแนนที่แตกต่างกัน ทำการปรับเทียบแบบสอบสองรูปแบบ มีจุดมุ่งหมายเฉพาะ เพื่อต้องการปรับปรุงแบบสอบที่ข้อสอบมีเนื้อหาเป็นเอกพันธ์เพื่อ

ต้องการปรับปรุงข้อสอบเกี่ยวกับแบบสอบร่วม ที่เป็นตัวแทนของเนื้อหาของแบบสอบทั้งหมด และเพื่อดูผลของเนื้อหาที่ขึ้นอยู่กับวิธีการปรับเทียบคะแนน โดยมีเป้าหมายหลักเพื่อการปรับปรุงแบบสอบ ซึ่งนำไปสู่เกณฑ์การตัดสินทางการศึกษา ในการปรับเทียบคะแนนครั้งนี้ใช้ข้อสอบร่วมเนื่องจากผู้สอบสองกลุ่มมีความแตกต่างกัน แบบสอบที่ใช้มี 4 ฉบับ โดยสุ่มข้อสอบที่แตกต่างกันในความเป็นเอกพันธ์ของเนื้อหาของข้อสอบ และความเป็นตัวแทนในเนื้อหาของข้อสอบร่วม โดยแต่ละแบบสอบมี 2 รูปแบบ และมีชุดของข้อสอบร่วมอยู่ในแบบสอบในแต่ละแบบสอบใช้วิธีการปรับเทียบคะแนน แบบเชิงเส้นตรง และแบบอิควิเปอร์เซ็นไทล์ และวิธีการปรับเทียบคะแนนที่มีพื้นฐาน IRT ทำการปรับเทียบคะแนนแบบสอบ 2 รูปแบบ ของแต่ละแบบสอบ ออกแบบการสุ่มข้อสอบโดยใช้ค่าเฉลี่ย เพื่อสร้างเกณฑ์การปรับปรุง 2 เกณฑ์ ที่มีพื้นฐานบนคะแนนจริงสำหรับการประเมินความถูกต้องของผลการปรับเทียบคะแนนจากวิธีต่างๆ เหล่านี้ โดยการตรวจสอบด้านความเที่ยงตรงเชิงโครงสร้างและมิติของแบบสอบที่สัมพันธ์กับการปรับเทียบคะแนนของแบบสอบ

ผลการวิจัยพบว่า ทุกวิธีการปรับเทียบคะแนนผลมีความถูกต้องระดับปานกลาง และทุกการพิจารณา ผลมีความถูกต้องมาก เมื่อข้อสอบร่วมมีความเป็นตัวแทนของแบบสอบทั้งหมดมาก หรือข้อสอบนั้นมีเนื้อหาเป็นเอกพันธ์ ดังนั้นการที่จะปรับปรุงการปรับเทียบคะแนนให้ดีขึ้นข้อสอบร่วมที่ใส่เข้าไปในแบบสอบปรับเทียบคะแนน ควรที่จะสะท้อนถึงเนื้อหาของแบบสอบร่วมทั้งหมด และพบด้วยว่า ผลการปรับเทียบคะแนนที่มีพื้นฐาน IRT มีความถูกต้องมากกว่าวิธีอื่น

Wiley (1999) ได้ทำการศึกษารูปแบบการปรับเทียบคะแนน 2 รูปแบบที่สอบกับรูปแบบข้อสอบที่หลากหลาย โดยใช้แบบแผนกลุ่มไม่เท่าเทียมกันที่มีแบบสอบร่วม ข้อสอบที่ใช้เป็นแบบเลือกตอบ (MC) และแบบเขียนตอบ (แบบเขียนตอบคำตอบที่ได้ยาวและต้องใช้ความจำมากกว่า MC ซึ่งไม่สามารถใช้ได้กับข้อสอบร่วม แบบสอบเหล่านี้ยากที่จะสร้างให้แบบสอบร่วมส่งผลให้แบบสอบรวมถูกต้อง ซึ่งทางเลือกหนึ่งของการปรับเทียบคะแนนในคะแนนรวมของ MC มากกว่าคะแนนรวมคะแนนของแบบเขียนตอบสามารถเป็นมาตรฐานที่มีค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานเหมือนกัน) การศึกษาครั้งนี้เป็นการเปรียบเทียบความถูกต้องของการปรับเทียบคะแนนที่ใช้รูปแบบการปรับเทียบที่หลากหลายถึงความถูกต้องของการปรับเทียบคะแนนของแบบสอบ ไปยังคะแนนของแบบสอบรวม แบบสอบประกอบด้วย แบบ MC 80 ข้อ และแบบเขียนตอบ 2 ข้อ ทดสอบกับกลุ่มผู้สอบที่มีความสามารถแตกต่างกัน วิธีการปรับเทียบคะแนนที่แตกต่างกัน 3 วิธี คือ Tucker, Levine และ Equipercentile ผลปรากฏว่า ความคลาดเคลื่อนของการปรับเทียบคะแนนส่วนใหญ่ได้รับอิทธิพลมาจากความสามารถของผู้สอบ

Leung (2003) ได้เสนอวิธีการปรับเทียบตามแนวตั้ง (Vertical Equating) โดยใช้การปรับเทียบร่วมกันระหว่างวิธีการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ กับวิธีการปรับเทียบเชิงเส้นตรง โดยใช้ข้อสอบร่วม จุดมุ่งหมายของการศึกษาครั้งนี้เพื่อตรวจสอบพัฒนาการความสามารถทางด้านภาษาจีนของนักเรียนระดับประถมศึกษาในฮ่องกง โดยกลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนระดับประถมศึกษาที่ทำการทดสอบในปี 1999 ซึ่งข้อมูลที่ใช้ในการศึกษาเป็นข้อมูลทุติยภูมิ ได้มาจากการศึกษาที่เป็นลักษณะของ Pilot Studies ของกลุ่มตัวอย่างขนาดใหญ่ (A Large Scaled Research) โดยการทดสอบความสามารถทางภาษาจีน ในปี 1999 ซึ่งในแต่ละระดับชั้น จะแบ่งออกเป็นระดับความสามารถ 6 ระดับ (Level 1-6)

จากงานวิจัยดังกล่าวจะใช้วิธีการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ โดยการปรับเทียบตามแนวตั้ง แบบโลจิสติก 3 พารามิเตอร์ เพื่อประมาณค่าความสามารถของผู้สอบที่อยู่ในระดับชั้นติดกัน (Adjacent grades) หรือต่อเนื่องกัน สำหรับการปรับเทียบแบบเชิงเส้นตรงจะใช้สำหรับการประมาณค่า โดยวิธี “Simple Least Squared Regression” เพื่อป้องกันปัญหาที่เรียกว่า “Out-of-Level Effect” ที่มักจะเกิดขึ้นจากการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ ผลการศึกษาพบว่าในแต่ละระดับชั้นนักเรียนมีพัฒนาการความสามารถทางภาษาจีนเพิ่มขึ้นและสำหรับการประเมินคุณภาพของการปรับเทียบคะแนน โดยใช้วิธีดูความสัมพันธ์ระหว่างการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ กับ การปรับเทียบการปรับเทียบหลังจากมีการปรับแก้ค่าจากการปรับเทียบเชิงเส้นตรง ซึ่งค่าความสัมพันธ์ที่ได้มีค่าเข้าใกล้หนึ่งเกือบทุกค่า 1 แสดงให้เห็นว่าคุณภาพของการปรับเทียบคะแนนอยู่ในระดับดี

งานวิจัยการปรับเทียบคะแนนในประเทศ

สำหรับงานวิจัยในประเทศไทย เท่าที่สำรวจได้เป็นงานวิจัยในข้อมูลที่เกี่ยวข้องกับเทคนิคการปรับเทียบแนวตั้ง ดังนี้

อาภรณ์ กาญจนกิจโสภณ (2531) ได้ทำการศึกษาการสร้างแบบสอบและตารางปรับเทียบคะแนนตามแนวตั้ง ในวิชาคณิตศาสตร์ เรื่องสมการและอสมการ โดยวิธีปรับเทียบคะแนนแบบราส์ช สำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น กลุ่มตัวอย่างที่ใช้ปรับเทียบคะแนนเป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 จำนวนระดับชั้นละ 490, 482 และ 415 คน ตามลำดับ ผู้วิจัยได้นำแบบสอบระดับชั้นละ 90 ข้อ ไปทดสอบกับกลุ่มตัวอย่างอื่น แล้วคัดข้อสอบที่มีคุณภาพ เรียงตามความยากที่ต้องการระดับชั้นละ 40, 30 และ 30 ข้อ โดยแต่ละชั้นจะมีข้อสอบร่วมระหว่างชั้นอยู่ 10 ข้อ หลังจากนั้นนำไปทดสอบกับกลุ่มตัวอย่างที่ใช้ปรับเทียบคะแนน

และสร้างตารางปรับเทียบคะแนน เพื่อใช้ศึกษาพัฒนาการของนักเรียนกลุ่มเดิมที่ผ่านขึ้นไปเรียนในชั้นที่สูงขึ้นว่า นักเรียนมีความก้าวหน้าเป็นไปตามปกติหรือเบี่ยงเบนไปจากปกติ

สุจินดา ผ่องอักษร (2533) ได้นำวิธีการปรับเทียบคะแนนรูปแบบราสช์ไปใช้ในการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียน ในกลุ่มทักษะของนักเรียนระดับชั้น ป.6 ที่เรียนจบหลักสูตรประถมศึกษา พุทธศักราช 2521 ในช่วงระยะเวลา 3 ปีการศึกษา (ปีการศึกษา 2529-2531) ในการวิจัยครั้งนี้นอกจากการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนแล้ว ยังมีการเปรียบเทียบความแตกต่างของผลสัมฤทธิ์ทางการเรียนระหว่างปีการศึกษาต่างๆ หลังจากที่ได้ทำการปรับเทียบคะแนนแล้วและเปรียบเทียบเศษส่วนจำนวนนักเรียนที่มีผลสัมฤทธิ์ผ่านเกณฑ์ที่น่าพอใจ ระหว่างการปรับเทียบคะแนนกับไม่ปรับเทียบคะแนน กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนระดับชั้น ป.6 ระดับประเทศเมื่อปีการศึกษา 2529, 2530 และ 2531 ในสังกัดสำนักงานประถมศึกษาจังหวัดปทุมธานี รวมจำนวนนักเรียนกลุ่มตัวอย่างในแต่ละปีการศึกษาเท่ากับ 534, 453 และ 467 คน ตามลำดับ แบบสอบที่ใช้เป็นแบบวัดผลสัมฤทธิ์ทางการเรียนในกลุ่มทักษะวิชาคณิตศาสตร์และภาษาไทย ซึ่งสำนักงานการประถมศึกษาสร้างขึ้น แบบแผนการเก็บรวบรวมข้อมูลใช้แบบผู้สอบร่วมเนื่องจากแบบสอบทั้ง 3 ฉบับไม่มีข้อสอบร่วม

ผลการวิจัยพบว่า จากการปรับเทียบคะแนนระหว่างแบบสอบทั้ง 3 ฉบับ คะแนนที่ปรับเทียบแล้ว ซึ่งตรงกับคะแนนดิบเดียวกัน มีค่าแตกต่างกันทั้งแบบสอบวิชาคณิตศาสตร์ และแบบสอบวิชาภาษาไทย เมื่อศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียน พบว่า ผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ในปีการศึกษา 2530 สูงกว่าปีการศึกษา 2529 แต่ลดลงในปีการศึกษา 2531 สำหรับผลสัมฤทธิ์วิชาภาษาไทย พบว่าอัตราความก้าวหน้าของคะแนนเฉลี่ยเพิ่มขึ้นทุกปีการศึกษา และอัตราความก้าวหน้าของค่าร้อยละนักเรียนที่มีผลน่าพอใจเพิ่มขึ้นทุกปีการศึกษา

พรพิมล นาคเวช (2537) ได้ทำการศึกษา เพื่อตรวจสอบคุณภาพการปรับเทียบคะแนนระหว่างแบบสอบในแนวตั้ง โดยใช้ทฤษฎีตอบสนองข้อสอบ สำหรับผู้ที่มีความสามารถต่างกัน 3 ระดับ คือ ม.1 ม.2 และ ม.3 กลุ่มตัวอย่างเป็นนักเรียนสังกัดกรมสามัญศึกษา จำนวน 6,429 คน เครื่องมือที่ใช้เป็นแบบสอบคณิตศาสตร์ 3 ฉบับ ฉบับละ 25 ข้อ โดยมีข้อสอบร่วมภายใน 5 ข้อ

ผลการวิจัยพบว่า ความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนสำหรับผู้มีความสามารถต่างกัน 3 ระดับ ค่อนข้างต่ำ มีค่าอยู่ระหว่าง 0.031-0.051 และความเพียงพอของ

การปรับเทียบคะแนนระหว่างแบบสอบตามแนวตั้ง สำหรับผู้มีความสามารถต่างกัน 3 ระดับ พบว่าค่าดัชนีความแตกต่างของการปรับเทียบคะแนน (C) อยู่ในระดับน่าพอใจ

วรรณ์ดี แสงประทีปทอง (2538) ได้ทำการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1, ม.2 และ ม.3 โดยใช้วิธีการปรับเทียบคะแนนตามแนวตั้ง ตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติกสามพารามิเตอร์ ใช้ข้อสอบร่วมภายนอกปรับคะแนนให้อยู่บนสเกลเดียวกัน แล้วเปรียบเทียบผลสัมฤทธิ์ทางการเรียนของนักเรียนแต่ละชั้น

ผลการวิจัย พบว่า 1) ผลสัมฤทธิ์ทางการเรียนภาษาไทยของนักเรียนระดับชั้น ม.2 ไป ม.3 ก้าวหน้ามากกว่าปกติ แต่ผลสัมฤทธิ์ทางการเรียนภาษาไทยระดับชั้น ม.1 ไป ม.2 ก้าวหน้าตามปกติ 2) ในกลุ่มนักเรียนที่มีความสามารถสูง ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยจากระดับชั้น ม.2 ไป ม.3 และ จากระดับชั้น ม.1 ไป ม.2 ก้าวหน้าตามปกติ 3) ในกลุ่มนักเรียนที่มีความสามารถปานกลาง ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของระดับชั้น ม.2 ไป ม.3 และระดับชั้น ม.1 ไป ม.2 ก้าวหน้ามากกว่าปกติ 4) ในกลุ่มนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนต่ำ ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนระดับชั้น ม.2 ไป ม.3 ก้าวหน้ามากกว่าปกติ แต่ผลสัมฤทธิ์ทางการเรียนของระดับชั้น ม.1 ไป ม.2 ก้าวหน้าตามปกติ

วันทยา วงศ์ศิลปะภิรมย์ (2543) ได้ทำการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ป.1-ป.6 โดยวิธีการปรับเทียบแนวตั้ง ซึ่งใช้วิธีการปรับเทียบ 2 วิธี คือ การปรับเทียบตามทฤษฎีการทดสอบแบบดั้งเดิมวิธีของเธอร์สโตนโดยเสนอค่าสัมบูรณ์ และการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 2 พารามิเตอร์ โดยเสนอค่าคะแนนจริง กลุ่มตัวอย่างที่ใช้ในการวิจัยเป็นนักเรียนระดับชั้น ป.1-ป.6 ที่สุ่มจากทุกจังหวัดทั่วประเทศ จำนวน 72,000 คน การออกแบบวิธีการเก็บรวบรวมข้อมูลโดยใช้รูปแบบกลุ่มไม่เท่าเทียมกันโดยใช้แบบสอบร่วม (Anchor Test Nonequivalent Group Design) เครื่องมือที่ใช้วัดมีทั้งสิ้น 7 ฉบับ

ผลการวิจัย พบว่า 1) นักเรียนระดับชั้น ป.1-ป.6 ทั้ง 6 ระดับชั้น มีคะแนนเฉลี่ยต่ำกว่าร้อยละ 50 2) การปรับเทียบตามทฤษฎีการทดสอบแบบดั้งเดิมวิธีของเธอร์สโตนโดยเสนอค่าสัมบูรณ์ การตรวจสอบคุณสมบัติความสอดคล้องของข้อคำถามร่วมกับโมเดลของเธอร์สโตน โดยการหาค่าสัมประสิทธิ์ระหว่างค่า Z ของชั้นที่ต่อเนื่อง ปรากฏว่าข้อคำถามร่วมของแบบสอบ 3 ฉบับจากทั้งหมด 15 ฉบับ ไม่สอดคล้องกับโมเดล ฉบับอื่นๆนอกเหนือจากนี้อีก 12 ฉบับ ค่าความสัมพันธ์มีนัยสำคัญทางสถิติที่ระดับต่ำกว่า .005 และผลการเปรียบเทียบคะแนนความสามารถทางคณิตศาสตร์ตามโมเดลของเธอร์สโตน แสดงให้เห็นถึงพัฒนาการในระดับชั้นที่

สูงขึ้นอย่างต่อเนื่อง 3) การปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 2 พารามิเตอร์ ปรากฏว่า การตรวจสอบข้อคำถามกับโมเดลโดยใช้การทดสอบไคสแควร์ด้วยโปรแกรม BILOG พบว่าแบบสอบแต่ละฉบับมีข้อคำถามที่ไม่สอดคล้องกับโมเดล และผลการปรับเทียบคะแนนความสามารถทางคณิตศาสตร์โดยเฉลี่ยตามการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 2 พารามิเตอร์ โดยเสนอค่าคะแนนจริง คิดเป็นค่าร้อยละ พบว่านักเรียนมีพัฒนาการระหว่างชั้นต่อเนื่องในทุกสมรรถนะ เมื่อพิจารณาพัฒนาการระหว่างชั้น พบว่านักเรียนระดับชั้น ป.3 แสดงพัฒนาการค่อนข้างสูงกว่านักเรียนชั้นอื่นๆ รองลงมาคือระดับชั้น ป.4

ทิภาภรณ์ ทุงฤทธิ์ (2544) ทำการปรับเทียบคะแนนโดยเส้นตรง เมื่อใช้วิธีการแปลงคะแนนแตกต่างกันในแบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ เพื่อเปรียบเทียบคะแนนที่ได้ ความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน และความเพียงพอของการปรับเทียบคะแนนในแนวนอน โดยใช้วิธีแบบเส้นตรง เมื่อใช้วิธีการแปลงคะแนนแตกต่างกัน 3 วิธี ดังนี้ 1) จากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบสอบกลุ่มละฉบับ 2) จากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบสอบทั้งสองฉบับที่เรียงลำดับการสอบต่างกัน และ 3) จากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบสอบคนละฉบับ และมีแบบสอบร่วมที่ต้องสอบทั้งสองกลุ่ม เครื่องมือเป็นแบบวัดผลสัมฤทธิ์ทางคณิตศาสตร์ จำนวน 3 ฉบับ ฉบับ X และ Y และฉบับร่วม กลุ่มตัวอย่างเป็นนักเรียนชั้น ป.6 จำนวน 820 คน ผลการวิจัยสรุปได้ว่า คะแนนที่ได้จากการปรับเทียบคะแนน และความเพียงพอของการปรับเทียบคะแนน 3 วิธี ไม่แตกต่างกัน ส่วนความเพียงพอของการปรับเทียบคะแนน 3 วิธี แตกต่างกัน

ภัทราพร เกษสังข์ (2546) ได้ศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียน และศึกษาประสิทธิภาพของการปรับเทียบคะแนนในแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติกสองและสามพารามิเตอร์ ที่ใช้แบบทดสอบร่วมภายในและแบบทดสอบเชื่อมโยงภายนอก ที่มีความยากแตกต่างกัน (ยากและปานกลาง) และ ความยาว (15 10 และ 5 ข้อ) ของแบบทดสอบร่วม จากกลุ่มผู้สอบที่มีความสามารถแตกต่างกันสามระดับชั้น คือ ชั้น ม.1,2 และ 3 โดยการวิเคราะห์ความคลาดเคลื่อนมาตรฐานละความเพียงพอของการปรับเทียบคะแนนแต่ละรูปแบบที่ศึกษา แล้วนำรูปแบบการปรับเทียบคะแนนที่เหมาะสม มาใช้ในการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียน

กลุ่มตัวอย่างของการวิจัย เป็นนักเรียนมัธยมศึกษาปีที่ 1,2 และ 3 สังกัดกองกลางมัธยมศึกษา กรมสามัญศึกษาในจังหวัดชัยนาท จำนวน 4,041 คน ที่ได้จากการสุ่มกลุ่มตัวอย่างแบบแบ่งชั้น แบ่งกลุ่มตัวอย่างเป็นสองกลุ่ม กลุ่มแรก จำนวน 3,081 คน ใช้สำหรับศึกษาการ

ปรับเทียบคะแนน เพื่อนำผลไปเทียบคะแนนระหว่างแบบสอบ และศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ กลุ่มที่สอง จำนวน 960 คน ใช้สำหรับสอบทานผล เพื่อนำผลไปตรวจสอบประสิทธิภาพของการปรับเทียบคะแนน เครื่องมือวัดที่ใช้เป็น แบบสอบคณิตศาสตร์ เรื่อง “สมการและอสมการ” ที่มีค่าความเที่ยงของแบบสอบระดับชั้นมัธยมศึกษาปีที่ 1,2 และ 3 เท่ากับ 0.745, 0.945 และ 0.651 ตามลำดับ และแบบทดสอบวัดความถนัดด้านตัวเลข ที่มีค่าความเที่ยงเท่ากับ 0.792, 0.847 และ 0.708 ตามลำดับ

ผลการวิจัยพบว่า 1) การวิเคราะห์ความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน ตามรูปแบบที่ศึกษาสำหรับผู้ที่มีความสามารถต่างกัน 3 ระดับ พบว่า มีค่าระหว่าง 0.1011 ถึง 0.1830 เป็นค่าระดับต่ำ ความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนต่ำที่สุด เป็นการปรับเทียบคะแนนในแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติกสองพารามิเตอร์ ที่แบบทดสอบร่วมภายใน 15 ข้อ มีความยากของข้อสอบร่วมเฉลี่ยอยู่ในระดับยากมาก 2) การวิเคราะห์ความเพียงพอของการปรับเทียบคะแนน พบว่า ค่าดัชนีความแตกต่างของการปรับเทียบคะแนน อยู่ในระดับน่าพอใจมาก และระดับน่าพอใจ มีความเพียงพอของการปรับเทียบคะแนนดีที่สุด เป็นการปรับเทียบคะแนนในแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบแบบโลจิสติกสองพารามิเตอร์ ที่แบบสอบร่วมภายใน 15 ข้อ ที่ลักษณะความยากของข้อสอบร่วมเฉลี่ยระดับยากมาก 3) การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ จากรูปแบบการปรับเทียบคะแนนที่มีประสิทธิภาพ พบว่าความสามารถทางคณิตศาสตร์ ของนักเรียนชั้น ม.1 ไปชั้น ม.2 ไปชั้น ม.3 มีพัฒนาสูงขึ้นและนักเรียน ม.1 ไป ม.3 มีพัฒนาการสูงขึ้นมาก

ข้อค้นพบที่ได้จากการศึกษาการปรับเทียบคะแนน สามารถสรุปได้ตามองค์ประกอบที่เกี่ยวข้องกับการปรับเทียบคะแนน ดังนี้

1. แบบแผนการเชื่อมต่อ (Linking Design) แบบแผนแบบสอบร่วมเป็นแบบแผนที่มีความเป็นไปได้และเป็นที่นิยมในทางปฏิบัติ การศึกษาส่วนมากกำหนดและอภิปรายในแบบแผนนี้ (Angoff, 1984; Petersen, Kolen & Hoover, 1989; Kolen & Brennan, 1995) ซึ่งไม่มีความซับซ้อนในการจัดการกับแบบสอบหลายๆ ฉบับ

2. ความยาวแบบสอบ (Test Length) แต่ละการศึกษาใช้แบบสอบที่มีความยาวแตกต่างกัน เช่นงานวิจัยของเบเกอร์ ใช้ 40 ข้อ (Baker, 1996) ขณะที่การศึกษาอื่นๆ ประยุกต์ใช้กับข้อสอบจำนวน 50 ข้อ (Kim & Cohen, 1998) จึงยังไม่มีข้อสรุปที่ชัดเจนเกี่ยวกับความยาวของแบบสอบ

3. ความยาวของข้อสอบร่วม (Anchor Length) เช่นเดียวกัน มีความแตกต่างในแต่ละการศึกษา แต่บางการศึกษาใช้ข้อสอบทุกข้อเป็นข้อสอบร่วมจำนวน 2 ข้อ และให้ผลดี (Wingersky & Lord, 1984; Vale, 1986) ซึ่งผลที่ได้อาจจะต้องตีความด้วยความรอบคอบ อย่างไรก็ตามจำนวนข้อสอบร่วมที่ยาวกว่าย่อมให้ผลที่ดีกว่า

4. ขนาดของหน่วยตัวอย่าง (Sample Size) ไม่มีการตกลงเกี่ยวกับประเด็นนี้ แต่ละการศึกษาประยุกต์ใช้หน่วยตัวอย่างขนาดแตกต่างกัน ซึ่ง Brennan และ Kolen (1987) ได้กล่าวไว้ว่ากลุ่มตัวอย่างที่ใช้ในการปรับเทียบคะแนนแต่ละกลุ่มไม่ควรต่ำกว่า 400 คน แต่การศึกษาหนึ่งใช้หน่วยตัวอย่างจริงจำนวน 3,000 คน ขณะที่การศึกษาอื่น ๆใช้ผู้เข้าสอบ 1,000 คน และ 500 คน (Kim & Cohen, 1998) โดยทั่วไปแล้วหน่วยตัวอย่างขนาดใหญ่จะทำให้เกิดความแม่นยำสำหรับการประมาณค่าพารามิเตอร์ข้อสอบ

5. การแจกแจงความสามารถ (Ability Distribution) เมื่อนำแบบสอบหลายฉบับดำเนินการต่อกลุ่มสุ่มหรือกลุ่มที่สมมูลกัน ผลที่ได้จะให้ความแม่นยำมากกว่า เมื่อแบบสอบหลายฉบับดำเนินการในกลุ่มที่ไม่เท่าเทียมกัน แต่ถ้าต้องการปรับเทียบคะแนนในกลุ่มที่ไม่เท่าเทียมกัน ควรใช้แบบแผนแบบสอบร่วมจะดีกว่า (Vale, 1986)

6. วิธีปรับเทียบคะแนนมาตรวัด (Method of Equating the Scale) Stocking และ Lord (1983) ได้เปรียบเทียบวิธีโค้งลักษณะข้อสอบกับวิธี robust mean and sigma และพบว่าวิธีโค้งคุณลักษณะข้อสอบให้ผลที่ดีกว่า และเมื่อเปรียบเทียบวิธีโค้งคุณลักษณะข้อสอบกับวิธี mean and mean เมื่อข้อมูลมีความซับซ้อนต่อการปรับ วิธีโค้งคุณลักษณะข้อสอบให้ผลที่แม่นยำกว่า Petersen, Cook และ Stocking ได้เปรียบเทียบวิธีการปรับค่าพารามิเตอร์พร้อมกัน ด้วยวิธีกำหนดค่า b (fixed b) กับวิธีโค้งลักษณะข้อสอบ ผลที่ได้ปรากฏว่าโดยภาพรวมวิธีการปรับค่าพารามิเตอร์พร้อมกันเป็นวิธีการที่ดีที่สุด นอกเหนือจากนี้ Kim และ Cohen (1998) ได้พบว่า เมื่อจำนวนของข้อสอบร่วมมีขนาดเล็ก วิธีโค้งลักษณะข้อสอบให้ผลที่ดีกว่าวิธีการปรับค่าพารามิเตอร์พร้อมกัน

7. ในวงการศึกษามีการนำวิธีการปรับเทียบคะแนนไปใช้ในการเปรียบเทียบคะแนนแบบสอบต่างฉบับที่วัดคุณลักษณะเดียวกัน และใช้ในการศึกษาพัฒนาการหรือความก้าวหน้าทางการเรียน ซึ่งจากการศึกษาความก้าวหน้าในการเรียนทั้งในประเทศและต่างประเทศจะเป็นลักษณะของการปรับเทียบคะแนนตามแนวตั้ง (Vertical Equating) และผลการศึกษาโดยทั่วไปวิธีการภายใต้ทฤษฎีตอบสนองข้อสอบ ให้ผลที่ดีกว่า เมื่อเปรียบเทียบกับวิธีการปรับเทียบคะแนนภายใต้ทฤษฎีการวัดแบบดั้งเดิม ซึ่งโดยส่วนใหญ่ เป็นการศึกษาการปรับเทียบเชิงเส้นตรง

(Linear Equating) และการปรับเทียบแบบอิควิเปอร์เซ็นไทล์ (Equipercentile Equating) แต่ก็ไม่ได้เป็นเช่นนั้นทุกสถานการณ์ (Petersen, Cook & Stocking, 1983) ขึ้นอยู่กับบริบทที่ต้องการศึกษา

ดังได้กล่าวมาแล้วข้างต้น โดยภาพรวมข้อค้นพบที่ได้จากการศึกษาการปรับเทียบคะแนน อาจจะสามารถอ้างอิงหรือไม่สามารถอ้างอิงไปยังสถานการณ์และเงื่อนไขอื่นได้ เพราะว่าแต่ละการศึกษาจะเกี่ยวข้องกับบริบทเฉพาะ ผลสรุปอาจจะเหมือนหรือไม่เหมือนกับการศึกษาอื่น ๆ นอกเหนือจากนี้ บางการศึกษาใช้ข้อมูลจริง และบางการศึกษาอาจจำลองข้อมูล ทำให้ผลของการศึกษาที่ได้มีความแตกต่างกัน แต่ก็ยังให้ภาพรวมว่าเงื่อนไขใดทำให้การปรับเทียบคะแนนในแต่ละวิธีให้ผลที่แม่นยำที่สุด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

3.2 งานวิจัยที่เกี่ยวข้องกับการศึกษาพัฒนาการ

การศึกษาเกี่ยวกับพัฒนาการที่ผ่านมามีการศึกษาวิจัยแตกต่างกัน ขึ้นอยู่กับจุดมุ่งหมายที่ทำการศึกษา ดังแสดงในตารางที่ 2.3

ตารางที่ 2.3 แสดงการศึกษาพัฒนาการของนักเรียนในวงการศึกษาตั้งแต่อดีตถึงปัจจุบัน

ปีการศึกษา	ผู้ที่ทำการศึกษา	แนวทาง/รูปแบบการศึกษา
2524	สำนักงานคณะกรรมการประถมศึกษาแห่งชาติ (สปช.)	จัดโครงการประเมินคุณภาพนักเรียนตั้งแต่เริ่มก่อตั้งหน่วยงาน คือ ปีการศึกษา 2524 แต่การดำเนินงานมีข้อจำกัดของเวลา บุคลากร และงบประมาณ
2527-2532	สำนักงานคณะกรรมการประถมศึกษาแห่งชาติ (สปช.)	งานประเมินคุณภาพเริ่มจัดอย่างเป็นลักษณะในการประเมินระดับชาติ (National Assessment) อย่างชัดเจนและต่อเนื่อง โดยกำหนดจุดมุ่งหมาย เพื่อศึกษาผลสัมฤทธิ์ทางการเรียนของนักเรียน ป.6 ในช่วงเวลาต่างๆ สถิติที่ใช้ คือ ร้อยละ ประเมินกับเกณฑ์ที่กำหนดไว้ การนำเสนอผลการประเมินได้นำเสนอเป็นภาพรวมของระดับประเทศ เขตการศึกษา และรายจังหวัด และจัดการเผยแพร่เพื่อให้มีผลไปใช้อย่างกว้างขวาง
2529	สำนักงานคณะกรรมการประถมศึกษาแห่งชาติ (สปช.)	ทำการสำรวจการนำเอาผลประเมินระดับชาติ (National Assessment) ไปใช้ประโยชน์ด้านต่าง ๆ กิจกรรมที่มีการนำไปใช้วางแผนพัฒนาคุณภาพนักเรียน การจัดสรรงบประมาณ และการนิเทศทั้งภายในและภายนอก หลังการสำรวจครั้งนั้นไม่ได้นำผลการประเมินไปใช้ ไม่ได้มีการติดตามการนำผลการประเมินไปใช้อย่างเป็นทางการ แต่ในทางปฏิบัติมีการนำไปใช้กันอย่างต่อเนื่อง แต่ด้วยข้อจำกัดด้านประสิทธิภาพในการจัดสารสนเทศ การนำผลไปใช้จึงไม่ได้ผลเท่าที่ควร

ปีการศึกษา	ผู้ที่ทำการศึกษา	แนวทาง/รูปแบบการศึกษา
2533-2541	กรมวิชาการ	ได้จัดให้มีการประเมินในลักษณะของการประเมินระดับชาติ โดยปีการศึกษาที่ลงท้ายด้วยเลขคี่ (2533,2535,...) เป็นการประเมินคุณภาพของนักเรียนชั้น ป.6 ส่วนในปีการศึกษาที่ลงท้ายด้วยเลขคู่เป็นการประเมิน ม.3 และ ม.6 เพื่อให้เห็นแนวโน้มของการพัฒนาคุณภาพนักเรียนในภาพรวมของประเทศที่ผ่านมาตลอด 15 ปี ได้ชัดเจนขึ้น จึงนำข้อมูลมาทำการวิเคราะห์ค่าถดถอย แต่เนื่องจาก สปช. ทำการประเมินในบางปี เพื่อแก้ปัญหาจึงใช้วิธีประมาณค่าเฉลี่ยของปีที่ขาดหายไปจากข้อมูล ผลการประเมินในช่วงแรก คือ ปีการศึกษา 2527-2532, 2534 (เฉพาะภาษาไทย) และ 2536 (เฉพาะคณิตศาสตร์) ซึ่งดำเนินการโดย สปช. ได้แสดงแนวโน้มเพิ่มขึ้นทุกกลุ่มประสบการณ์ (ยกเว้นภาษาไทย) มีค่าเฉลี่ยเริ่มต้นต่ำกว่าร้อยละ 50 และมีอัตราความก้าวหน้าประมาณร้อยละ 1 -4 ผลการประเมินช่วงหลัง คือ ปีการศึกษา 2533-2541 ซึ่งดำเนินการปีเว้นปี โดยกรมวิชาการได้แสดงแนวโน้มลดลงทุกกลุ่มประสบการณ์ ยกเว้นกลุ่มการงานพื้นฐานอาชีพ ซึ่งเป็นกลุ่มเดียวที่มีค่าเฉลี่ยเริ่มต้นต่ำกว่าร้อยละ 50 แต่มีอัตราความก้าวหน้าเฉลี่ยประมาณร้อยละ 1 สำหรับกลุ่มประสบการณ์อื่น ๆ มีค่าเฉลี่ยเริ่มต้นสูงกว่าร้อยละ 50 แต่อัตราความก้าวหน้าลดลงโดยเฉลี่ยประมาณร้อยละ 1-3
2541	กรมวิชาการ ร่วมกับ สปช.	ทำการวิเคราะห์ถดถอยจากผลการประเมินของกรมวิชาการ กับ สปช. รวมกัน ตั้งแต่ปีการศึกษา 2527-2541 สรุปได้ว่า ทุกกลุ่มประสบการณ์มีอัตราความก้าวหน้าเพิ่มขึ้นประมาณร้อยละ 0.5-1.2 สรุปได้ว่า สำหรับค่าเฉลี่ยเริ่มต้นต่ำกว่าร้อยละ 40 ส่วนกลุ่มทักษะภาษาไทยและกลุ่มสร้างเสริมประสบการณ์ชีวิต มีค่าเฉลี่ยเริ่มต้นสูงกว่าร้อยละ 50
2544	สำนักนิเทศและพัฒนามาตรฐานการศึกษา	ศึกษาเปรียบเทียบผลสัมฤทธิ์ของ 4 กลุ่มสาระสรุปได้ว่า กลุ่มทักษะภาษาไทยมีค่าเฉลี่ยสูงที่สุด รองลงมาคือ กลุ่มสร้างเสริมประสบการณ์ชีวิต กลุ่มการงานและพื้นฐานอาชีพ และกลุ่มทักษะคณิตศาสตร์ ตามลำดับ สำหรับอัตราความก้าวหน้า กพอ.มีค่าสูงที่สุด รองลงมา คือ คณิตศาสตร์ ภาษาไทย และ สปช.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ปี การศึกษา	ผู้ที่ทำการศึกษา	แนวทาง/รูปแบบการศึกษา
2544	อวยพร เรืองตระกูล	ทำการพัฒนาวิธีการวัดคะแนนพัฒนาการ ตามทฤษฎีตอบสนองข้อสอบ 2 วิธี ซึ่ง 1 ใน 2 วิธีที่พัฒาขึ้น คือ วิธีการวัดคะแนนพัฒนาการจากความสามารถที่แท้จริงสัมพัทธ์เมื่อเทียบกับศักยภาพการพัฒนา ($S\theta$) เนื่องจากวิธีการวัดคะแนนพัฒนาการโดยทั่วไปมักมีอิทธิพลเพดานติดอยู่ กล่าวคือ ผู้สอบที่มีความสามารถที่แท้จริงก่อนเรียนสูงจะมีพิสัยของศักยภาพการพัฒนาแคบกว่าพิสัยของผู้สอบที่มีความสามารถที่แท้จริงก่อนเรียนต่ำ คะแนนพัฒนาการของผู้สอบที่มีความสามารถที่แท้จริงก่อนเรียนสูงจึงต่ำกว่าคะแนนของผู้สอบที่มีความสามารถที่แท้จริงก่อนเรียนต่ำ วิธีการที่พัฒนาขึ้นสามารถปรับแก้จุดอ่อนข้อนี้ของวิธีการวัดคะแนนพัฒนาการจากความแตกต่างของความสามารถที่แท้จริงผู้วิจัยจึงประยุกต์แนวคิดของ ศิริชัย กาญจนวาสี (2532) มาแก้ปัญหาเรื่องอิทธิพลเพดานโดยพัฒนาเป็นวิธีการวัดคะแนนพัฒนาการจากความสามารถที่แท้จริงสัมพัทธ์ โดยประมาณค่าจากอัตราส่วนผลต่างของความสามารถที่แท้จริงก่อนและหลังการเรียนกับศักยภาพการพัฒนาได้ โดยนิยามศักยภาพการพัฒนาว่าเป็นผลต่างของความสามารถสูงสุดกับความสามารถที่แท้จริงก่อนเรียน โดยใช้สูตร ดังนี้ $$S\theta = \frac{\theta_2 - \theta_1}{MAX - \theta_1}$$ เมื่อ $S\theta$ = วิธีการวัดคะแนนพัฒนาการจากความสามารถที่แท้จริงสัมพัทธ์เมื่อเทียบกับศักยภาพการพัฒนา θ_1 = ความสามารถที่แท้จริงก่อนเรียน θ_2 = ความสามารถที่แท้จริงหลังเรียน MaX = ความสามารถสูงสุด (Max = 4) ผลจากการวัดคะแนนพัฒนาการในครั้งนี้พบว่า ผลสัมฤทธิ์ทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.2 ในโรงเรียนโครงการขยายโอกาสการศึกษาขั้นพื้นฐาน และระดับ ม.3 โรงเรียนในสังกัดสามัญศึกษา เขตกรุงเทพมหานคร มีพัฒนาการค่อนข้างต่ำ ส่วนใหญ่จะมีพัฒนาการต่ำกว่าร้อยละ 50 ของคะแนนเต็ม

จากที่กล่าวมาข้างต้นแสดงให้เห็นการศึกษาพัฒนาการในภาคตามยาวของนักเรียนระดับชั้นใดชั้นหนึ่ง ในช่วงเวลาที่ผ่านมา และการประเมินอีกลักษณะหนึ่งที่แสดงให้เห็นถึงพัฒนาการของนักเรียนในภาคตัดขวาง เช่น

สำนักงานคณะกรรมการศึกษาแห่งชาติ (2534) ได้ประเมินคุณภาพนักเรียนชั้น ป.1-6 ในกลุ่มทักษะภาษาไทยในปีการศึกษา 2534 วิเคราะห์ข้อมูลโดยใช้สถิติร้อยละ เพื่อศึกษาความสามารถในการเขียนของนักเรียนในเชิงปริมาณและเชิงคุณภาพ รวมทั้งศึกษาข้อบกพร่องในการเขียน

สำนักงานคณะกรรมการศึกษาแห่งชาติ (2536) ได้ทำการประเมินความสามารถของนักเรียนในกลุ่มทักษะคณิตศาสตร์ ในลักษณะเดียวกันกับที่เคยดำเนินงานการประเมินทางภาษาไทยเพื่อให้ได้สารสนเทศทั้งในเชิงปริมาณและเชิงคุณภาพ หรือภาพรวมผลสัมฤทธิ์ทางการเรียนทั้งจุดเด่นและข้อบกพร่องของนักเรียน อันเป็นประโยชน์ต่อการพัฒนาการเรียนการสอนตลอดแนว และการดำเนินการครั้งนี้ ได้พัฒนารูปแบบการประเมินเพิ่มเติมเพื่อให้สามารถศึกษาพัฒนาการระหว่างชั้นได้ด้วย โดยจัดให้มีข้อสอบร่วม (Anchor Test) ระหว่างชั้นจำนวนหนึ่งสำหรับใช้ในการปรับเทียบคะแนนของทุกระดับชั้นให้อยู่ในมาตรฐานเดียวกัน อันจะทำให้เห็นพัฒนาการที่ต่อเนื่องของแต่ละชั้นได้อย่างชัดเจน และใช้การปรับเทียบตามทฤษฎีการวัดแบบดั้งเดิมตามโมเดลของเธอร์สโตน ซึ่งเป็นโมเดลที่ใช้กับการปรับเทียบตามแนวตั้งโดยเฉพาะ และ IRT โดยใช้โมเดล 2 พารามิเตอร์ ผลการประเมินสรุปได้ดังนี้

คะแนนเฉลี่ยและจำนวนนักเรียนที่มีผลน่าพอใจทั้งในรายสมรรถภาพและในภาพรวม 3 สมรรถภาพ มีค่าต่ำกว่าร้อยละ 50 และผลรวม 3 สมรรถภาพ ทั้ง 6 ระดับชั้นมีคะแนนเฉลี่ยประมาณร้อยละ 40-47 โดยนักเรียนที่มีผลน่าพอใจมีประมาณร้อยละ 26-40 และไม่มีนักเรียนที่ไม่ได้คะแนนเลย ยกเว้นนักเรียนชั้น ป.6 ที่มีเพียงร้อยละ 0.1

ผลการปรับเทียบคะแนนตามทฤษฎีการวัดแบบดั้งเดิม โดยวิธีการของเธอร์สโตน โดยใช้ ป.1 เป็นฐาน แสดงให้เห็นพัฒนาการระหว่างชั้นได้ดี โดยทักษะการคำนวณนักเรียนชั้น ป.1 มีคะแนนที่ปรับแล้ว เท่ากับ 3.270 เพิ่มขึ้นเป็น 3.270 ในชั้น ป.6 ด้านความเข้าใจคณิตศาสตร์ คะแนนในชั้น ป.1 เท่ากับ 14.333 เพิ่มขึ้นเป็น 34.296 ในชั้น ป.6 ด้านทักษะการแก้โจทย์ปัญหาคะแนนเพิ่มขึ้นจาก 16.345 ในชั้น ป.1 เป็น 41.698 ในชั้น ป.6 เมื่อรวมผล 3 สมรรถภาพพัฒนาการระหว่างชั้นจะยิ่งเห็นได้ชัดเจนขึ้น โดยชั้น ป.1 มีคะแนน 16.345 เพิ่มขึ้นเป็น 41.698 ในชั้น ป.6 อย่างไรก็ดี เมื่อเปรียบเทียบระหว่างสมรรถภาพจะเห็นได้ว่าทักษะการแก้โจทย์ปัญหามี

พัฒนาการสูงกว่าสมรรถภาพอื่นๆ และเมื่อเปรียบเทียบระหว่างชั้น พบว่า ชั้น ป.4 และ ป.3 มีพัฒนาการสูงกว่าชั้นอื่นๆ รองลงมา คือ ชั้น ป.6 และ ป.2 สำหรับชั้น ป.5 มีพัฒนาการน้อยมาก

ผลการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ โดยใช้รูปแบบ 2 พารามิเตอร์ ผลการปรับเทียบคะแนนตามกระบวนการปรับเทียบคะแนนแบบลูกโซ่ โดยนำเสนอผลเป็นคะแนนจริง พัฒนาการความสามารถของนักเรียนชั้น ป.1-ป.6 ในด้านทักษะการคำนวณ นักเรียนชั้น ป.1 มีคะแนนที่ปรับแล้ว เท่ากับ ร้อยละ 70 เพิ่มขึ้นเป็นร้อยละ 96.18 ในชั้น ป.6 ด้านความเข้าใจทางคณิตศาสตร์มีคะแนนร้อยละ 56.76 ในชั้น ป.1 เพิ่มขึ้นเป็นร้อยละ 85.52 ในชั้น ป.6 ด้านทักษะการแก้ปัญหา มีคะแนนร้อยละ 61.89 ในชั้น ป.1 เพิ่มเป็นร้อยละ 87.15 ในชั้น ป.6 เมื่อเปรียบเทียบพัฒนาการระหว่างสมรรถภาพ จะเห็นได้ว่าด้านความเข้าใจมีพัฒนาการระหว่างชั้นโดยเฉลี่ยมากกว่าสมรรถภาพด้านอื่นๆ เมื่อเปรียบเทียบพัฒนาการระหว่างระดับชั้น ปรากฏว่าชั้น ป.3 มีพัฒนาการสูงกว่าชั้นอื่นๆ รองลงมาคือ ชั้น ป.4

จากงานวิจัยที่ศึกษาพัฒนาการที่กล่าวมา มีทั้งการศึกษากับกลุ่มตัวอย่างเดิมแต่โดยเวลาต่างกัน ซึ่งเป็นการศึกษาภาคตามยาว (Longitudinal study) และศึกษากับกลุ่มตัวอย่างต่างกันแต่เวลาที่ศึกษาขณะใดขณะหนึ่งพร้อมกัน ซึ่งเป็นการศึกษาภาคตามขวาง (Cross-Section Study) อีกทั้งเป็นการศึกษาที่ใช้รูปแบบวิธีที่ต่างกัน ที่จะได้มาซึ่งสารสนเทศ (Information) ที่จะนำมาใช้ในการศึกษาพัฒนาการ แต่การศึกษานั้นมีสิ่งที่เหมือนกัน คือ การนำแนวคิดเรื่องการปรับเทียบคะแนนเป็นเครื่องมือมาใช้เพื่อความชัดเจน เพื่อให้ได้มาซึ่งข้อมูลที่ใช้ในการศึกษาพัฒนาการ ประกอบกับการวิจัยนี้เป็นการศึกษากับกลุ่มตัวอย่างต่างกัน ที่ศึกษาในเวลาเดียวกัน ผู้วิจัยจึงสนใจที่นำรูปแบบการปรับเทียบคะแนนที่มีประสิทธิภาพ ที่จะได้มาซึ่งสารสนเทศที่เหมาะสมในการศึกษาพัฒนาการ ดังนั้นผู้วิจัยจึงสนใจในการศึกษาการปรับเทียบแนวตั้ง (Vertical Equating) ตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง เพื่อนำมาใช้ในการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ ของนักเรียนในช่วงชั้นที่ 3 คือ ม.1, 2 และ 3 โดยนักเรียนทั้ง 3 ระดับชั้น เป็นนักเรียนในภาคการศึกษาและปีการศึกษาเดียวกัน อันเป็นการศึกษาพัฒนาการในลักษณะภาคตัดขวาง (Cross-Section Study) โดยจัดให้มีแบบสอบร่วมภายใน นอกจากนี้ยังศึกษาประสิทธิภาพของการปรับเทียบแนวตั้ง ในตัวแปรที่ศึกษา อันเป็นสารสนเทศในการประยุกต์ใช้การปรับเทียบคะแนนที่เหมาะสมในสาขาวิชาและระดับชั้นเรียนต่อไปในอนาคต

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยนี้เป็นการศึกษาเชิงประจักษ์ (Empirical Study) เพื่อศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนช่วงชั้นที่ 3 คือ ชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 และศึกษาประสิทธิภาพของการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง รายละเอียดของวิธีการวิจัย มีดังนี้

ประชากร

ประชากรในการวิจัยครั้งนี้ เป็นนักเรียนช่วงชั้นที่ 3 คือ ชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ภาคเรียนที่ 2 ปีการศึกษา 2547 ของโรงเรียนในเขตพื้นที่การศึกษาที่ 1 จังหวัดสกลนคร จำนวน 15 โรงเรียน รวม 9,975 คน เป็นนักเรียน ม.1 จำนวน 3,579 คน นักเรียน ม.2 จำนวน 3,488 คน และนักเรียน ม.3 จำนวน 2,727คน (ข้อมูลจากสำนักงานเขตพื้นที่การศึกษาสกลนคร เขต 1, 2547) ดังแสดงในตารางที่ 3.1

ตารางที่ 3.1 จำนวนประชากรจำแนกตามขนาดของโรงเรียนและระดับชั้น

ขนาดโรงเรียน	จำนวนโรงเรียน	ชั้น ม.1	ชั้น ม.2	ชั้น ม.3	รวมจำนวนนักเรียน
		จำนวนนักเรียน	จำนวนนักเรียน	จำนวนนักเรียน	
ใหญ่พิเศษ	1	685	697	623	2,005
ใหญ่	1	573	444	398	1,415
กลาง	10	2,238	2,082	1,527	5,848
เล็ก	3	263	264	179	707
รวม	15	3,759	3,488	2,725	9,975

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาครั้งนี้เป็นนักเรียนช่วงชั้นที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2547 ของโรงเรียนในเขตพื้นที่การศึกษาที่ 1 จังหวัดสกลนคร จำนวน 1,941 คน ซึ่งได้โดยการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) มีขนาดของโรงเรียนเป็นตัวแปรสำหรับแบ่งชั้น คือ ขนาดใหญ่พิเศษ ขนาดใหญ่ ขนาดกลาง และขนาดเล็ก ซึ่งมีขั้นตอนการได้มาซึ่งกลุ่มตัวอย่าง ดังนี้

1. คำนวณขนาดกลุ่มตัวอย่างจากนักเรียน 9,975 คน จากสูตร (ศิริชัย กาญจนวาสี, ทวีวัฒน์ ปิตยานนท์ และ ดิเรก ศรีสุโข, 2544) ดังนี้

$$n_{\hat{\mu}} = \frac{NZ^2\sigma^2}{NE^2 + Z^2\sigma^2}$$

เมื่อ $n_{\hat{\mu}}$ แทน ขนาดของกลุ่มตัวอย่างที่ต้องการศึกษาค่าเฉลี่ย (μ) จากประชากร

N แทน ขนาดของประชากร

Z แทน ค่าที่กำหนดจากค่าความเชื่อมั่นที่ผู้วิจัยต้องการจะใช้เพื่อการสรุปผลที่ระดับความเชื่อมั่น 95% ซึ่งมีค่าเท่ากับ 1.96 หรือประมาณ 2

σ^2 แทน ค่าความแปรปรวนของตัวแปรหลักที่ต้องการศึกษา

E แทน ค่าความคลาดเคลื่อนที่สามารถยอมรับได้ในการสรุปผล

แทนค่าในสูตร

$$n = \frac{9,975 \times (2)^2 (1)^2}{\left[9,975 \times (0.05)^2\right] + \left[(2)^2 (1)^2\right]}$$

$$= 1,378$$

เพื่อให้งานวิจัยมีความแข็งแกร่งของข้อมูลและลดความคลาดเคลื่อนที่เกิดขึ้น ผู้วิจัยใช้กลุ่มตัวอย่าง มากกว่า 1,378 คน ประกอบกับจากการศึกษาของ Kolen และ Brennan (1987) เสนอว่า ในการปรับเทียบคะแนนแต่ละกลุ่มด้วยวิธีทฤษฎีตอบสนองข้อสอบ (IRT) ผู้สอบในแต่ละกลุ่มของกลุ่มปรับเทียบคะแนนไม่ควรต่ำกว่า 400 คน นอกจากนี้ Kolen และ Whitney (1982) ได้เสนอว่า กลุ่มสอบทานผลควรมีขนาดไม่น้อยกว่า 10% ของกลุ่มปรับเทียบคะแนน ดังนั้นเพื่อให้งานวิจัยมีความแข็งแกร่งของข้อมูลและลดความ คลาดเคลื่อนที่เกิดขึ้น ผู้วิจัยใช้กลุ่มปรับเทียบคะแนนระดับชั้นละประมาณ 500 คน และกลุ่ม สอบทานผลประมาณ 20 % ของกลุ่มปรับเทียบคะแนน ทำให้ได้กลุ่มสอบทานผลระดับชั้นละประมาณ 100 คน รวมเป็นกลุ่มตัวอย่างในแต่ละระดับชั้นประมาณระดับชั้นละ 600 คน รวมเป็นกลุ่มตัวอย่างทั้งหมดประมาณ 1,800 คน

2. คำนวณหาขนาดของกลุ่มตัวอย่างที่พอดีของประชากรในแต่ละขนาดของโรงเรียนเมื่อกำหนดให้กลุ่มตัวอย่างในแต่ละระดับชั้น มีจำนวน 600 คน โดยใช้สูตรดังนี้ (Freund, 1979)

$$n_i = \frac{N_i}{N} \cdot n$$

เมื่อ n_i แทน ขนาดของกลุ่มตัวอย่างที่พอดีในแต่ละขนาดของโรงเรียน

N_i แทน ขนาดของประชากรในแต่ละโรงเรียน

N แทน ขนาดของประชากรที่ใช้ในการวิจัยทั้งหมด

n แทน ขนาดของกลุ่มตัวอย่างที่ใช้ในการวิจัย

จากการคำนวณได้กลุ่มตัวอย่างในแต่ละขนาดของโรงเรียน ดังแสดงในตารางที่ 3.2

ตารางที่ 3.2 จำนวนนักเรียนและจำนวนโรงเรียนที่ใช้เป็นกลุ่มตัวอย่างในการวิจัย จำแนกตามขนาดของโรงเรียน และระดับชั้น

ขนาดโรงเรียน	จำนวนโรงเรียน	ชั้น ม.1	ชั้น ม.2	ชั้น ม.3	รวมจำนวนนักเรียน
		จำนวนนักเรียน	จำนวนนักเรียน	จำนวนนักเรียน	
ใหญ่พิเศษ	1	109	120	137	366
ใหญ่	1	92	77	88	257
กลาง	5	357	358	336	1,051
เล็ก	2	42	45	39	126
รวม	9	600	600	600	1,800

3. หลังจากได้จำนวนกลุ่มตัวอย่างในแต่ละขนาดของโรงเรียนของแต่ละระดับชั้นแล้วทำการสุ่มตัวอย่างโดยวิธีสุ่มแบบแบ่งชั้น (Stratified Random Sampling) ได้กลุ่มตัวอย่างที่ใช้ในการวิจัยจำแนกตามขนาดของโรงเรียน โรงเรียน และกลุ่มสำหรับการวิจัยจำนวน 1,941 คน ดังแสดงในตารางที่ 3.3

ตารางที่ 3.3 จำนวนกลุ่มตัวอย่างที่ใช้ในการวิจัยจำแนกตาม ขนาดของโรงเรียนและกลุ่มตัวอย่างสำหรับการวิจัย เขตพื้นที่การศึกษาสกลนคร เขต 1

ขนาดโรงเรียน / โรงเรียน	อำเภอ	กลุ่มปรับเทียบคะแนน			กลุ่มสอบทานผล		
		ม.1	ม.2	ม.3	ม.1	ม.2	ม.3
ขนาดใหญ่พิเศษ							
สกลราชวิทยานุกูล	เมือง	92	97	80	20	20	20
ขนาดใหญ่							
ธาตุนารายณ์วิทยา	เมือง	68	54	58	13	14	14
ขนาดกลาง							
สกลนครพัฒนศึกษา	เมือง	58	61	79	15	15	16
คำเพิ่มพิทยา	ภูพาน	70	70	70	15	15	16
เต่างอยพัฒนศึกษา	เต่างอย	71	79	56	15	7	15
กุสุมาลย์วิทยาคม	กุสุมาลย์	69	69	46	15	17	15
ร่มเกล้า	โคกศรีสุพรรณ	59	60	73	12	29	16
ขนาดเล็ก							
โพนพิทยาคม	โพนนาแก้ว	26	23	21	4	4	4
ด่านม่วงคำพิทยา	โคกศรีสุพรรณ	30	21	21	4	6	4
รวม		543	534	504	113	127	120

จากตารางที่ 3.3 กลุ่มตัวอย่างสำหรับการวิจัยนี้ ผู้วิจัยใช้ขั้นตอนการเก็บรวบรวมข้อมูลโดยในแต่ละการสอบของแต่ละห้อง ในแต่ละระดับชั้น จัดแบ่งกลุ่มนักเรียน ออกเป็น 2 กลุ่ม คือ

1) กลุ่มปรับเทียบคะแนน จำนวน 1,581 คน ซึ่งแยกเป็นกลุ่มย่อยตามระดับชั้น และแต่ละระดับชั้นทำแบบสอบปรับเทียบสเกลเฉพาะในระดับของตนเอง คือนักเรียนในระดับชั้น ม.1, ม.2 และ ม.3 จำนวนระดับชั้นละ 543, 534 และ 504 คน ตามลำดับ เพื่อนำมาใช้ในการสร้างตารางปรับเทียบคะแนน และศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ โดยตอบแบบสอบปรับเทียบคะแนนระดับชั้นของตนที่มีแบบสอบร่วมภายใน

2) กลุ่มสอบทานผล จำนวน 360 คน โดยแยกเป็นกลุ่มย่อยตามระดับชั้น ม.1, ม.2 และ ม.3 จำนวนระดับชั้นละ 113, 127 และ 120 คน ตามลำดับ เป็นกลุ่มตัวอย่างที่ใช้ในการสอบทานผล เพื่อนำมาใช้ในการประเมินประสิทธิภาพของการปรับเทียบคะแนน มีจำนวนประมาณร้อยละ 20 ของกลุ่มตัวอย่างปรับเทียบคะแนน โดยตอบแบบสอบปรับเทียบคะแนนที่มีข้อสอบร่วมภายใน (Internal Anchor Test)

แบบแผนการเก็บข้อมูลในการปรับเทียบของการวิจัย

การวิจัยครั้งนี้ใช้แบบแผนผู้สอบกลุ่มไม่เท่าเทียมกันโดยมีแบบสอบร่วม (Anchor Test Nonequivalent Group Design) ซึ่งแบบสอบร่วมมีลักษณะเป็นแบบสอบร่วมภายใน (Internal Anchor Test) แบบแผนการปรับเทียบคะแนน แสดงดังตารางที่ 3.4

ตารางที่ 3.4 แบบแผนการปรับเทียบคะแนน

กลุ่มตัวอย่าง	แบบสอบ		
	ฉบับที่ 1	ฉบับที่ 2	แบบสอบร่วม
กลุ่มที่ 1	✓	-	✓
กลุ่มที่ 2	-	✓	✓

การปรับเทียบคะแนนในครั้งนี้ออกแบบการเก็บรวบรวมข้อมูลโดยใช้แบบสอบร่วม (Anchor Test) กับกลุ่มที่มีความสามารถต่างกัน 3 กลุ่ม ได้แก่ นักเรียนระดับชั้น ม.1, ม.2 และ ม.3 ซึ่งเป็นกลุ่มปรับเทียบคะแนน โดยแต่ละกลุ่มทำแบบสอบในระดับชั้นของตนเองเพียง 1 ชุด ซึ่งในแบบสอบจะมีข้อสอบร่วมระหว่างระดับชั้น ม.1 กับ ม.2 ข้อสอบร่วมระหว่างระดับชั้น ม.2 กับ ม.3 อย่างละ 6 ข้อ สรุปได้ดังตารางที่ 3.5

ตารางที่ 3.5 แบบแผนการเก็บรวบรวมข้อมูลในการปรับเทียบคะแนน

กลุ่มตัวอย่างปรับเทียบคะแนน	แบบสอบระดับชั้น ม.1	แบบสอบระดับชั้น ม.1 - ม.2	แบบสอบระดับชั้น ม.2	แบบสอบระดับชั้น ม.2 - ม.3	แบบสอบระดับชั้น ม.3
ม.1	✓	✓	-	-	-
ม.2	-	✓	✓	✓	-
ม.3	-	-	-	✓	✓

กลุ่มตัวอย่างสอบทานผลในครั้งนี้ เป็นนักเรียนระดับชั้น ม.1 ม.2 และ ม.3 โดยในแต่ละระดับต้องตอบแบบสอบรวมของทั้ง 3 ระดับ คือ ตอบแบบสอบในทุกระดับชั้นที่มีแบบสอบร่วมภายในซึ่งจะตัดข้อสอบร่วมระหว่างชั้นที่ต่อเนื่องซ้ำกันออก รวมแล้วจำนวน 78 ข้อ ดังตารางที่ 3.6

ตารางที่ 3.6 แบบแผนการเก็บรวบรวมข้อมูลในการสอบทานผล

กลุ่มตัวอย่างปรับเทียบคะแนน	แบบสอบระดับชั้น ม.1	แบบสอบระดับชั้น ม.1 - ม.2	แบบสอบระดับชั้น ม.2	แบบสอบระดับชั้น ม.2 - ม.3	แบบสอบระดับชั้น ม.3
ม.1	✓	✓	✓	✓	✓
ม.2	✓	✓	✓	✓	✓
ม.3	✓	✓	✓	✓	✓

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการศึกษาครั้งนี้เป็นแบบสอบที่ผู้วิจัยสร้างขึ้น ในเนื้อหาสาระการเรียนรู้คณิตศาสตร์ เรื่อง “จำนวนและการดำเนินการ” ช่วงชั้นที่ 3 แบ่งเป็น 3 ระดับ ตามหลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 ระดับละ 30 ข้อ ดังนี้

ระดับที่ 1 แบบสอบความสามารถทางคณิตศาสตร์ เรื่อง จำนวนและการดำเนินการ ม.1

ระดับที่ 2 แบบสอบความสามารถทางคณิตศาสตร์ เรื่อง จำนวนและการดำเนินการ ม.2

ระดับที่ 3 แบบสอบความสามารถทางคณิตศาสตร์ เรื่อง จำนวนและการดำเนินการ ม.3

แบบสอบแต่ละฉบับจะมีข้อสอบร่วมภายใน (Internal Anchor Items) จำนวน 6 ข้อ ในระดับชั้นที่ต่อเนื่องกัน ดังแผนภาพที่ 3.1

	ข้อสอบเฉพาะ	ข้อสอบร่วม	ข้อสอบเฉพาะ	ข้อสอบร่วม	ข้อสอบเฉพาะ
ม.1	1,2,3,…,24	25,26,…,30			
ม.2		1,2,3,……,6	7,8,9,…,24	25,26,27,…,30	
ม.3				1,2,3,……..,6	7,8,9,…,30

แผนภาพที่ 3.1 รูปแบบการจัดฉบับแบบสอบที่มีข้อสอบร่วมภายในระหว่างชั้นที่ต่อเนื่อง

จากแผนภาพที่ 3.1 จะเห็นว่า แบบสอบระดับชั้น ม.1 มีจำนวน 30 ข้อ โดยข้อที่ 24-30 ซึ่งเป็นเนื้อหาในระดับชั้น ม.1 ใช้เป็นข้อสอบร่วมระหว่างระดับชั้น ม.1 กับ ม.2 แบบสอบระดับชั้น ม.2 จำนวน 30 ข้อ โดยข้อที่ 1-6 เป็นข้อสอบร่วมระหว่างระดับชั้น ม.1 กับ ม.2 ส่วนข้อที่ 7-24 ซึ่งเป็นเนื้อหาในระดับชั้น ม.2 และใช้ข้อที่ 24-30 เป็นข้อสอบร่วมระหว่างชั้น ม.2 กับ ม.3 แบบสอบชั้น ม.3 มีจำนวน 30 ข้อ โดยข้อที่ 1-6 เป็นข้อสอบร่วมระหว่างชั้น ม.2 กับ ม.3 และข้อที่ 7-30 ซึ่งเป็นเนื้อหาในระดับชั้น ม.3

สำหรับข้อสอบร่วมระหว่างระดับชั้น ม.1 กับ ม.2 ซึ่งเป็นเนื้อหาของระดับชั้น ม.1 ได้มาจากข้อสอบที่มีค่าความยาก (b) สูงสุดและครอบคลุมวัตถุประสงค์ที่ต้องการวัดจากแบบสอบของระดับชั้น ม.1 และข้อสอบร่วมระหว่างระดับชั้น ม.2 กับ ม.3 ซึ่งเป็นเนื้อหาของระดับชั้น ม.2 ได้มาจากข้อสอบที่มีค่าความยาก (b) สูงสุดและครอบคลุมวัตถุประสงค์ที่ต้องการวัดจากแบบสอบของระดับชั้น ม.2 ทั้งนี้เพราะนักเรียนที่อยู่ในระดับชั้น ม.2 ถือได้ว่าผ่านการเรียนรู้จากระดับชั้น ม.1 มาแล้ว ในทำนองเดียวกัน นักเรียนในระดับชั้น ม.3 ได้ผ่านการเรียนรู้จากระดับชั้น ม.1 และ ม.2 มาแล้ว เช่นกัน

ส่วนเครื่องมือที่ใช้กับกลุ่มตัวอย่างสอบทานผลนั้น เนื่องจากนักเรียนในกลุ่มนี้ แต่ละคนต้องทำแบบสอบของทั้ง 3 ระดับ ดังนั้นเพื่อความสะดวกในการรวบรวมข้อมูล ผู้วิจัยจึงรวมแบบสอบของทั้ง 3 ระดับ เป็นข้อสอบฉบับเดียว มีทั้งหมด 78 ข้อ โดยตัดข้อสอบร่วมระหว่างชั้นข้อที่ซ้ำกันออก แล้วเรียงลำดับข้อ จากแบบสอบของระดับชั้น ม.1 เป็นข้อที่ 1-30 ต่อด้วยแบบสอบระดับชั้น ม.2 เป็นข้อที่ 31-53 และสุดท้ายเป็นแบบสอบของระดับชั้น ม.3 เป็นข้อที่ 54-78 ตามลำดับ ในการนำคะแนนมาวิเคราะห์แยกข้อมูลออกเป็นระดับชั้นละ 30 ข้อเช่นเดิม

วิธีดำเนินการสร้างแบบสอบ

แบบสอบที่สร้างขึ้นเป็นแบบสอบเลือกตอบ 4 ตัวเลือก ในสาระการเรียนรู้คณิตศาสตร์ เรื่อง “จำนวนและการดำเนินการ” สำหรับนักเรียนช่วงชั้นที่ 3 โดยมีวิธีดำเนินการสร้าง ดังนี้

1. ศึกษาหลักสูตร สาระการเรียนรู้คณิตศาสตร์ มาตรฐานการเรียนรู้ช่วงชั้นและผลการเรียนรู้ที่คาดหวัง เรื่อง จำนวนและการดำเนินการ สำหรับนักเรียนช่วงชั้นที่ 3 จากแผนการจัดการเรียนรู้ และเอกสารที่เกี่ยวข้อง

2. สร้างเครื่องมือตามมาตรฐานการเรียนรู้และผลการเรียนรู้ที่คาดหวัง โดยแต่ละมาตรฐานการเรียนรู้จะสร้างข้อสอบให้มีจำนวนมากกว่าที่ต้องการ ซึ่งเครื่องมือที่สร้างขึ้นเป็นแบบสอบเลือกตอบ 4 ตัวเลือก ระดับละ 60 ข้อ

3. นำแบบสอบตรวจสอบคุณภาพขั้นต้น โดยให้ผู้เชี่ยวชาญทางการวัดผลและครูผู้สอนทางคณิตศาสตร์ จำนวน 7 ท่าน พิจารณาความสอดคล้องของข้อสอบกับมาตรฐานการเรียนรู้ เพื่อตรวจสอบความตรงเชิงเนื้อหา โดยหาค่าดัชนีความสอดคล้องของข้อสอบ (Item Objective Congruence : IOC) ข้อสอบที่ใช้ได้มีค่า IOC ตั้งแต่ 0.60 ถึง 1.00 พร้อมทั้งปรับปรุงตามคำแนะนำของผู้เชี่ยวชาญ

4. นำแบบสอบที่ได้จากข้อ 3 ไปทดลองกับนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 ที่ไม่ใช่กลุ่มตัวอย่างมาระดับชั้นละ 214 คน (4 ห้องเรียน), 150 คน (3 ห้องเรียน) และ 99 คน (2 ห้องเรียน) ตามลำดับ จากโรงเรียนกัลยาณวัตร จังหวัดขอนแก่น เนื่องจากกลุ่มทดลองเครื่องมือดังกล่าวมีลักษณะคล้ายกลุ่มตัวอย่างสำหรับการวิจัยในครั้งนี้ โดยนักเรียนแต่ละระดับชั้นจะทำข้อสอบในระดับชั้นของตน

5. ตรวจให้คะแนน โดยให้คะแนนข้อที่ตอบถูก 1 คะแนน ข้อที่ตอบผิด 0 คะแนน

6. นำผลจากข้อ 5. มาวิเคราะห์หาคุณภาพข้อสอบ โดยใช้โปรแกรมวิเคราะห์ข้อสอบ CTIA วิเคราะห์ข้อสอบรายข้อ คัดเลือกข้อสอบที่มีค่าความยาก (p) 0.20-0.80 ค่าอำนาจจำแนก (r) ตั้งแต่ 0.20 ขึ้นไป

7. นำข้อสอบที่ได้จากข้อ 5. มาวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบ คือความยากของข้อสอบ (b) และค่าอำนาจจำแนกของข้อสอบ (a) โดยใช้โปรแกรม BILOG 3.04 แล้วคัดเลือกข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบ (a) ที่มีค่าตั้งแต่ 0.05 ขึ้นไป ผลปรากฏว่า

แบบสอบระดับชั้น ม.1 มีค่าสถิติของข้อสอบตามทฤษฎี CTT ดังนี้ ค่าความยาก (p) อยู่ระหว่าง 0.229 ถึง 0.734 และค่าอำนาจจำแนก (r) อยู่ระหว่าง 0.228 ถึง 0.579 และมีค่าสถิติ

ของข้อสอบตามทฤษฎี IRT มีความยาก (b) อยู่ระหว่าง -0.194 ถึง 1.837 และค่าอำนาจจำแนก (a) อยู่ระหว่าง 0.416 ถึง -1.372

แบบสอบระดับชั้น ม.2 มีค่าสถิติของข้อสอบตามทฤษฎี CTT ดังนี้ ค่าความยาก (p) อยู่ระหว่าง 0.139 ถึง 0.715 และค่าอำนาจจำแนก (r) อยู่ระหว่าง 0.200 ถึง 0.600 และมีค่าสถิติของข้อสอบตามทฤษฎี IRT มีความยาก (b) อยู่ระหว่าง -1.090 ถึง 1.968 และค่าอำนาจจำแนก (a) อยู่ระหว่าง 0.528 ถึง 1.420

แบบสอบระดับชั้น ม.3 มีค่าสถิติของข้อสอบตามทฤษฎี CTT ดังนี้ ค่าความยาก (p) อยู่ระหว่าง 0.222 ถึง 0.697 และค่าอำนาจจำแนก (r) อยู่ระหว่าง 0.231 ถึง 0.731 และมีค่าสถิติของข้อสอบตามทฤษฎี IRT มีความยาก (b) อยู่ระหว่าง -1.441 ถึง 1.262 และค่าอำนาจจำแนก (a) อยู่ระหว่าง 0.575 ถึง 1.433

8. ปรับปรุงคุณภาพของข้อสอบ โดยคัดเลือกข้อสอบเฉพาะแต่ละระดับชั้น และจากการกระจายของระดับความยากของข้อสอบและความครอบคลุมตามมาตรฐานการเรียนรู้ให้ได้น้ำหนักตามโครงสร้างเนื้อหา ในกรณีที่สาระการเรียนใดมีข้อสอบ จำนวน 1 ข้อ จะพิจาณาให้เป็นข้อสอบร่วม และกรณีที่มีจำนวนข้อสอบเกินจากจำนวนที่ต้องการในแต่ละสาระการเรียนรู้จะพิจาณาข้อสอบที่มีค่าอำนาจจำแนกอยู่ในเกณฑ์ดี (โดยพิจารณาระดับพฤติกรรมการวัดไปพร้อมกัน) ในการคัดเลือกข้อสอบพิจารณาแยกออกเป็น 2 กรณี คือ

8.1 ข้อสอบร่วมระหว่างระดับชั้นที่ต่อเนื่อง จะพิจารณาให้กระจายไปตามเนื้อหาและจะพิจารณาจากค่าความยากของข้อสอบ (b) โดยแยกออกเป็น 2 ระดับ คือ

- ระดับชั้น ม.1 กับ ม.2 จะพิจารณาจากข้อสอบที่กระจายไปตามเนื้อหาของระดับชั้น ม.1 และมีค่าความยากสูง จำนวน 6 ข้อ แต่ถ้าข้อใดมีค่าความยากในระดับชั้น ม.1 สูง แต่จำนวนข้อสอบร่วมที่ได้เกาะกลุ่มอยู่ในเนื้อหาใดเนื้อหาหนึ่ง จะทำการเลือกข้อสอบที่มีค่าความยากต่ำลงมาแต่ครอบคลุมเนื้อหาแทน จากการพิจารณาดังกล่าวได้ค่าความยากตั้งแต่ 0.791 ถึง 1.407

- ระดับชั้น ม.2 กับ ม.3 จะพิจารณาจากข้อสอบที่กระจายไปตามเนื้อหาของระดับชั้น ม.2 และมีค่าความยากในระดับสูง จำนวน 6 ข้อ ซึ่งมีหลักการพิจารณาเช่นเดียวกับแบบสอบร่วมระดับชั้น ม.1 กับ ม.2 ได้ค่าความยากตั้งแต่ 1.290 ถึง 1.747

8.2 ข้อสอบเฉพาะในแต่ละระดับชั้น พิจารณาจากการกระจายของข้อสอบกับจุดประสงค์การเรียนรู้โดยพิจารณาเลือกข้อสอบสำหรับระดับชั้น ม.1 จำนวน 24 ข้อ ม.2 จำนวน 18 ข้อ และ ม.3 จำนวน 24 ข้อ

9. นำข้อสอบที่คัดเลือกจากข้อ 8 มาจัดฉบับแบบสอบเรียงตามลำดับ ดังแผนภาพที่ 3.1 ได้จำนวนข้อดังนี้

9.1 แบบสอบระดับชั้น ม.1 ข้อสอบเฉพาะระดับชั้น 24 ข้อ คือข้อที่ 1-24 และข้อสอบร่วมกับระดับชั้น ม.2 จำนวน 6 ข้อ คือข้อที่ 25-30 รวมเป็น 30 ข้อ ซึ่งมีค่าความเที่ยงของแบบสอบเท่ากับ 0.807

9.2 แบบสอบระดับชั้น ม.2 มีข้อสอบเฉพาะระดับชั้น 18 ข้อ คือ ข้อที่ 7-24 ข้อสอบร่วมกับระดับชั้น ม.1 จำนวน 6 ข้อ คือข้อที่ 1-6 และข้อสอบร่วมกับระดับชั้น ม.3 จำนวน 6 ข้อ คือข้อที่ 25-30 รวมเป็น 30 ข้อ ซึ่งมีค่าความเที่ยงของแบบสอบ เท่ากับ 0.774

9.3 แบบสอบระดับชั้น ม.3 ข้อสอบเฉพาะระดับชั้น 24 ข้อ คือข้อที่ 7-30 และข้อสอบร่วมกับระดับชั้น ม.2 จำนวน 6 ข้อ คือข้อที่ 1-6 รวมเป็น 30 ข้อ ซึ่งมีค่าความเที่ยงของแบบสอบเท่ากับ 0.719

การเก็บรวบรวมข้อมูล

1. ขอข้อมูลรายชื่อโรงเรียนและจำนวนนักเรียนระดับช่วงชั้นที่ 3 จากสำนักงานเขตพื้นที่การศึกษาสกลนคร เขต 1

2. ขอหนังสือแนะนำตนเองจากบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย เพื่อขอความร่วมมือในการเก็บข้อมูลจากโรงเรียน

3. ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลด้วยตนเองจากโรงเรียนที่เป็นกลุ่มตัวอย่าง

4. นำแบบสอบที่เก็บมาแล้วตรวจให้คะแนนและวิเคราะห์ข้อมูล

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลดำเนินการโดยแบ่งเป็น 4 ตอน ได้แก่

1. การวิเคราะห์แบบสอบ
2. การสร้างตารางปรับเทียบคะแนน
3. การศึกษาประสิทธิภาพของการปรับเทียบคะแนน
4. การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนช่วงชั้นที่ 3

รายละเอียดของการวิเคราะห์ข้อมูล มีดังนี้

ตอนที่ 1 การวิเคราะห์แบบสอบ

วิเคราะห์ค่าสถิติพื้นฐานของแบบสอบฉบับต่างๆ ก่อนปรับเทียบคะแนน ได้แก่ คะแนนเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด ค่าต่ำสุด ค่ามัธยฐาน ฐานนิยม ความเบ้ ความโด่ง และค่าความเที่ยงของแบบสอบ โดยใช้โปรแกรมสำเร็จรูป SPSS for Windows Version 11.0

ตอนที่ 2 การสร้างตารางปรับเทียบคะแนนแบบสอบ

2.1 การปรับเทียบระดับความสามารถของผู้สอบโดยใช้ทฤษฎีตอบสนองข้อสอบแบบโลจิสติก 3 พารามิเตอร์ โดยมีขั้นตอนดังนี้

2.1.1 ประมาณค่าพารามิเตอร์ของข้อสอบในแบบสอบร่วมและประมาณค่าพารามิเตอร์ความสามารถของผู้สอบแต่ละคน ด้วยโปรแกรม BILOG 3.04

2.2.2 หาค่าคงที่แปลงคะแนนความสามารถจากระดับชั้น ม.2 ไปสเกลเดียวกันกับคะแนนความสามารถระดับชั้น ม.1 และ ม.3 โดยใช้สูตรการแปลงคะแนนความสามารถจากโปรแกรม EQUATE 2.0 ของ Bake, Al-kani และ Dosary (1991) ได้สมการเชิงเส้นตรง ดังนี้

$$\theta_{21} = A_1\theta_2 + K_1$$

$$\theta_{23} = A_2\theta_2 + K_2$$

เมื่อ A_1 และ A_2 แทน ความชัน

K_1 และ K_2 แทน ค่าสัมประสิทธิ์ (Intercept)

จากสูตรการแปลงคะแนนความสามารถจะได้คะแนนความสามารถของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ดังตารางที่ 3.7

ตารางที่ 3.7 รูปแบบที่ได้จากการวิเคราะห์การปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ

ระดับชั้น	ผลการวิเคราะห์	
	θ_{21}	θ_{23}
1	✓	
2	✓	✓
3		✓

จากตารางที่ 3.7 พบว่า นักเรียนแต่ละคนในระดับชั้น ม.2 จะมีคะแนนความสามารถ 2 ค่า นั่นคือ จะมีคะแนนความสามารถที่มีสเกลเดียวกันกับทั้งระดับชั้น ม.1 (θ_{12}) และระดับชั้น ม.3 (θ_{23})

2.2.3 นำค่าประมาณพารามิเตอร์ที่ได้ผ่านการแปลงให้อยู่ในสเกลเดียวกันดังกล่าวแล้วนั้น คำนวณหาค่าคะแนนจริงของผู้สอบแบบสอบแต่ละชุดในแต่ละระดับชั้น โดยใช้สูตรในการคำนวณ (อ้างใน ศิริชัย กาญจนวาสี, 2545) ดังนี้

$$\xi_2 = \sum_{i=1}^{30} P_i(\theta_2)$$

$$\xi_{21} = \sum_{j=1}^{30} P_j(A_1\theta_2 + K_1) = \sum_{j=1}^{30} P_j(\theta_{21})$$

$$\xi_{23} = \sum_{k=1}^{30} P_k(A_2\theta_2 + K_2) = \sum_{k=1}^{30} P_k(\theta_{23})$$

โดยที่ ξ_2 แทน คะแนนจริงของนักเรียนระดับชั้น ม.2 ซึ่งได้จากการตอบแบบสอบชั้น ม.2

ξ_{21} แทน คะแนนจริงของนักเรียนระดับชั้น ม.2 ที่ได้จากการแปลงคะแนนความสามารถไปสู่สเกลเดียวกันกับระดับชั้น ม.1

ξ_{23} แทน คะแนนจริงของนักเรียนระดับชั้น ม.2 ที่ได้จากการแปลงคะแนนความสามารถไปสู่สเกลเดียวกันกับระดับชั้น ม.3

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2.2 การปรับเทียบระดับความสามารถของผู้สอบโดยใช้การปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ (IRT) แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง โดยมีขั้นตอนดังนี้

2.2.1 ทำการปรับแก้ค่าที่อาจจะเกิดความคลาดเคลื่อนสูงในการปรับเทียบคะแนนในระดับคะแนนที่มีความสามารถแตกต่างกันมาก ๆ โดยเฉพาะการแปลงคะแนนความสามารถระหว่างระดับชั้น ม.1 และ ม.3 เนื่องจากมีระดับความสามารถที่แตกต่างกันอย่างเห็นได้ชัด ด้วยเหตุนี้อาจเกิดปัญหาที่เรียกว่า “Out-of-Level Effect” (O’Brien & Tohn, 1984; Leung, 2003) ที่อาจส่งผลต่อความถูกต้องแม่นยำของการปรับเทียบคะแนน ผู้วิจัยจึงทำการปรับแก้ค่าที่ได้จากการปรับเทียบคะแนนความสามารถตามทฤษฎีตอบสนองข้อสอบดังกล่าวข้างต้น โดยใช้วิธีการเชิงเส้นตรง ที่เรียกว่า “วิธีกำลังสองน้อยสุด” (Simple Least Square Regression) โดยใช้โปรแกรม SPSS for Window Version 11.0 โดยกำหนดให้ θ_{21} เป็นตัวแปรอิสระ และ θ_{23} เป็นตัวแปรเกณฑ์ ที่เป็นผลมาจากการแปลงคะแนนความสามารถจากระดับชั้น ม.2 ไปสู่สเกลเดียวกันกับคะแนนความสามารถของนักเรียนระดับชั้น ม.1 โดยคะแนนความสามารถดังกล่าวจะส่งผลต่อการแปลงคะแนนความสามารถจากระดับชั้น ม.2 ไปสู่สเกลเดียวกันกับคะแนนความสามารถของนักเรียนระดับชั้น ม.3 ค่าที่ได้จะใช้ในการประมาณค่าพารามิเตอร์เพื่ออธิบายการแปลงค่าความสามารถจากระดับชั้น ม.1 ไปสู่สเกลเดียวกันกับระดับชั้น ม. 3 โดยตรง สามารถแสดงได้ดังสมการต่อไปนี้

$$\theta_{23} = b\theta_{21} + a$$
$$\theta_{13} = b\theta_{1} + a$$

เมื่อ $b = r\left(\frac{S_{\theta_{23}}}{S_{\theta_{21}}}\right)$ และ $a = \bar{\theta}_{23} - b\bar{\theta}_{21}$

2.2.3 คำนวณค่าคะแนนจริง ในแต่ละคะแนนความสามารถ

2.2.2 สร้างตารางปรับเทียบคะแนน และนำค่าคะแนนความสามารถของนักเรียนแต่ละระดับจากตารางไปเทียบหาคะแนนดิบในแต่ละระดับความสามารถ

ตอนที่ 3 การศึกษาประสิทธิภาพของการปรับเทียบคะแนน

3.1 ประสิทธิภาพของการปรับเทียบคะแนน พิจารณาจากค่าคะแนนความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน ดังนี้

3.1.1 หาค่าความคลาดเคลื่อนของการปรับเทียบคะแนน โดยใช้สูตร

$$E = \left| x_j - x_j^* \right|$$

เมื่อ E แทน ค่าความคลาดเคลื่อน

x_j แทน คะแนนของแบบสอบชุดที่ j ที่ได้จากการสอบ

x_j^* แทน คะแนนของแบบสอบชุด j ที่ได้จากตารางปรับเทียบคะแนน

3.1.2 หาค่าคะแนนความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน โดยใช้สูตร

$$SEE = \frac{SD_e}{\sqrt{n}}$$

เมื่อ SEE แทน คะแนนความคลาดเคลื่อนมาตรฐานของคะแนนความคลาดเคลื่อน

SD_e แทน ค่าส่วนเบี่ยงเบนมาตรฐานของคะแนนความคลาดเคลื่อนของการปรับเทียบคะแนน

n แทน จำนวนนักเรียนในกลุ่มตัวอย่างสอบทานผล

3.2 การหาดัชนีความแตกต่าง (Discrepancy Comparison : C) เพื่อพิจารณาความเพียงพอของความคลาดเคลื่อนของการปรับเทียบคะแนน โดยใช้สูตรของ Petersen และคณะ ดังนี้

$$C = \sum \frac{(X_{ij} - X_{kij}^*)^2}{nS_{Xi}^2}$$

เมื่อ C แทน ค่าดัชนีความแตกต่างในการปรับเทียบคะแนน

X_{ij} แทน คะแนนจากแบบสอบฉบับ i ของนักเรียนคนที่ j

X_{kij}^* แทน คะแนนจากแบบสอบฉบับ k ที่แปลงคะแนนเป็นตารางเทียบคะแนนไปเป็นคะแนนจากแบบสอบฉบับที่ i ของนักเรียนคนที่ j

n แทน จำนวนในกลุ่มตัวอย่างสอบทานผล

S^2_{Xi} แทน ความแปรปรวนคะแนนจากแบบสอบฉบับ i

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ซึ่งการตัดสินความเพียงพอของความคลาดเคลื่อนของการปรับเทียบคะแนนที่ใช้ในการวิจัยนี้ ประเมินจากค่าดัชนีความแตกต่างกับเกณฑ์การยอมรับของ Petersen และคณะ (1982) ดังนี้

		C	≤	$(0.05\ SD_x)^2$	หมายถึง	ระดับน่าพอใจมาก
$(0.05\ SD_x)^2$	<	C	≤	$(0.10\ SD_x)^2$	หมายถึง	ระดับน่าพอใจ
$(0.10\ SD_x)^2$	<	C	≤	$(0.15\ SD_x)^2$	หมายถึง	ระดับปานกลาง
$(0.15\ SD_x)^2$	<	C	≤	$(0.20\ SD_x)^2$	หมายถึง	ระดับไม่น่าพอใจ
$(0.20\ SD_x)^2$	<	C			หมายถึง	ระดับไม่น่าพอใจมาก

ตอนที่ 4 การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนช่วงชั้นที่ 3

การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 จะศึกษาจากผลต่างของค่าเฉลี่ยคะแนนความสามารถทางคณิตศาสตร์ของนักเรียน แต่ละระดับชั้น ภายหลังจากการแปลงคะแนนความสามารถของนักเรียนระดับชั้น ม.2 และ ม.3 ให้อยู่ในสเกลเดียวกันกับคะแนนของแบบสอบระดับชั้น ม.1 แล้ว การวิเคราะห์ข้อมูลดำเนินการ 2 ครั้ง ครั้งแรก ศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียน 3 ชั้น โดยภาพรวม โดยใช้ข้อมูลจากนักเรียนในกลุ่มปรับเทียบคะแนนทุกคน และครั้งที่ 2 ศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนจากลุ่มปรับเทียบคะแนน โดยแยกวิเคราะห์ข้อมูลในกลุ่มนักเรียนที่มีความสามารถต่ำ นักเรียนที่มีความสามารถปานกลาง และนักเรียนที่มีความสามารถสูงแต่ละกลุ่ม มีขั้นตอนในการวิเคราะห์ข้อมูลดังนี้

4.1 การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.1 ม.2 และ ม.3 ในกลุ่มรวม

4.1.1 แปลงคะแนนที่นักเรียนระดับชั้น ม.1 และ ม.2 แต่ละคนได้จากการตอบแบบสอบตามชั้นของตน ไปสู่สเกลคะแนนแบบสอบระดับชั้น ม.3 โดยใช้ตารางปรับเทียบคะแนนที่สร้างไว้ในตอนที่ 2

4.1.2 คำนวณค่าสถิติพื้นฐานของคะแนนภายหลังการปรับเทียบคะแนนที่นักเรียนแต่ละชั้นทำได้ โดยใช้ โปรแกรมสำเร็จรูป SPSS for Windows Version 11.0

4.1.3 ทดสอบความแตกต่างระหว่างคะแนนความสามารถเฉลี่ยของนักเรียนทั้งสามระดับชั้นของคะแนนเฉลี่ย ตามข้อ 4.1.2 ด้วยการวิเคราะห์ความแปรปรวนแบบทางเดียว (One-Way ANOVA) โดยใช้โปรแกรมสำเร็จรูป SPSS for Windows Version 11.0

4.1.4 คำนวณหาอัตราพัฒนาการ (Growth Rate) จากวิธีการวัดคะแนนพัฒนาการตามทฤษฎีตอบสนองข้อสอบ ของอวยพร เรืองตระกูล (2544) ซึ่งได้ประยุกต์แนวคิดในการวัดคะแนนพัฒนาการตามทฤษฎีตอบสนองข้อสอบแบบดั้งเดิมของ ศิริชัย กาญจนวาสี (2532) และแนวคิดของ วินิจ เทือกทอง (2537) มาใช้ในการวัดคะแนนพัฒนาการจากความสามารถที่แท้จริง สัมพัทธ์ กรณีการวัด 2 ครั้ง โดยเทียบกับศักยภาพการพัฒนา โดยแทนด้วยสัญลักษณ์ $S\theta$ ในกรณีนี้ผู้วิจัยขอเรียกสัญลักษณ์ดังกล่าวว่า อัตราพัฒนาการ (Growth Rate) โดยได้นำสูตรดังกล่าวมาประยุกต์ใช้ในการศึกษาแนวโน้มการเปลี่ยนแปลงคะแนนเฉลี่ยของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ดังนี้

$$S\theta_{12} = \frac{\theta_2 - \theta_1}{MAX - \theta_1}$$

$$S\theta_{23} = \frac{\theta_3 - \theta_2}{MAX - \theta_2}$$

$$S\theta_{13} = \frac{\theta_3 - \theta_1}{MAX - \theta_1}$$

เมื่อ $S\theta_{12}$ แทน อัตราพัฒนาการของนักเรียน ม.1 เมื่อขึ้น ม.2

$S\theta_{23}$ แทน อัตราพัฒนาการของนักเรียน ม.2 เมื่อขึ้น ม.3

$S\theta_{13}$ แทน อัตราพัฒนาการของนักเรียน ม.1 เมื่อขึ้น ม.3

θ_1 แทน คะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้น ม.1

θ_2 แทน คะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้น ม.2

θ_3 แทน คะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้น ม.3

MAX แทน ความสามารถสูงสุด ($\theta_{MAX} = 4$)

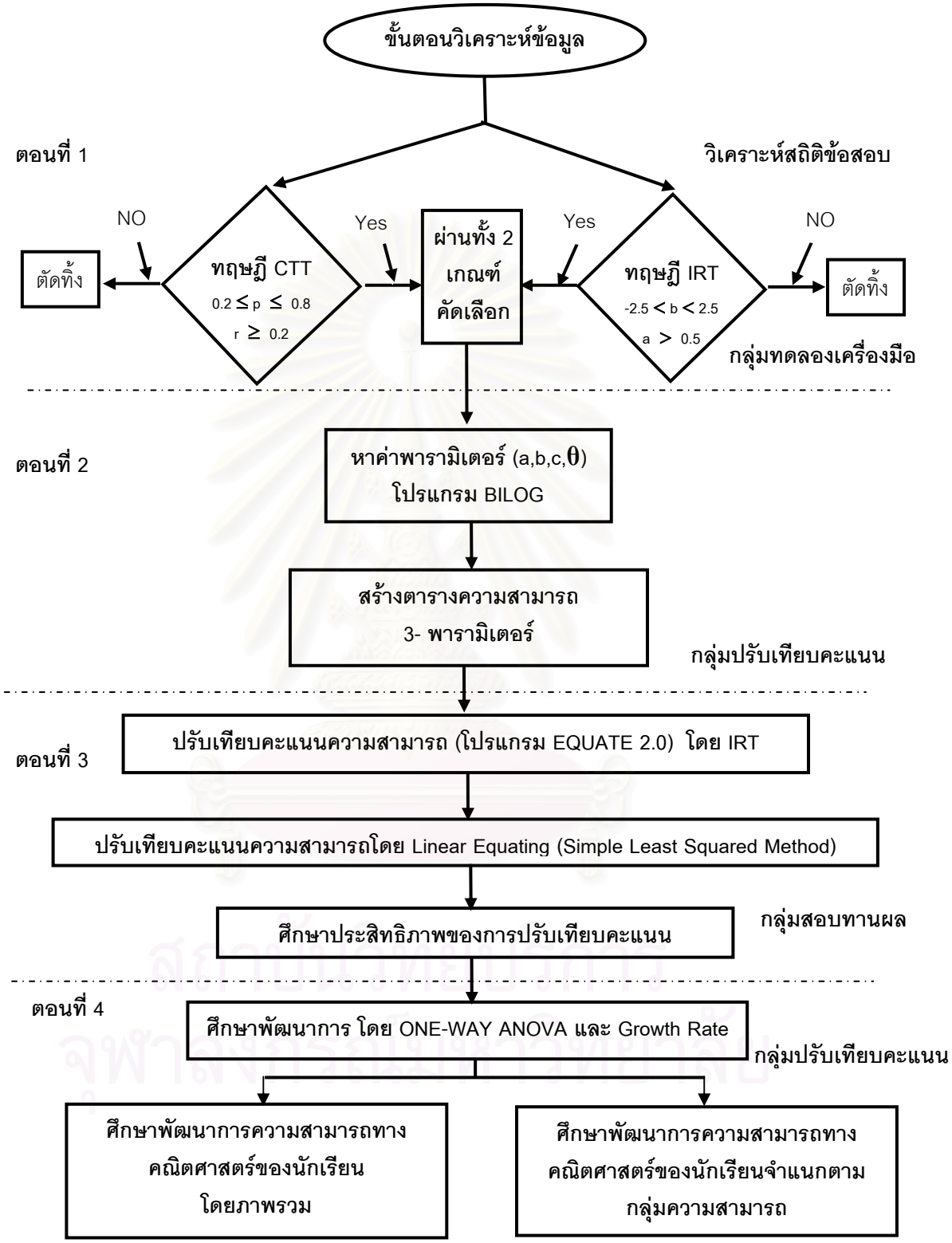
หมายเหตุ : เนื่องจากค่า θ เป็นคะแนนมาตรฐานที่มีค่าจาก $-\alpha$ ถึง $+\alpha$ ซึ่งสามารถกำหนดช่วงได้ตั้งแต่ 4 ถึง -4 เพราะสามารถครอบคลุมการแจกแจงของข้อมูลได้ถึงร้อยละ 99.994 (Kirk, 1995 อ้างใน อวยพร เรืองตระกูล, 2544) ดังนั้น Max ที่เป็นความสามารถที่แท้จริงสูงสุด สำหรับการวิจัยนี้จึงกำหนดให้มีค่าเท่ากับ 4 ตามแนวคิดของ อวยพร เรืองตระกูล (2544)

4.2 การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนจำแนกตามกลุ่มความสามารถ การแบ่งกลุ่มผู้สอบตามระดับความสามารถ (θ) จะแบ่งกลุ่มโดยใช้คะแนนของนักเรียนที่ได้จากการทำแบบสอบปรับเทียบคะแนนตามชั้นของตน ตามเกณฑ์ของวรรณ์ดี แสงประทีปทอง (2538) ซึ่งตามปกติจะมีทั้งหมด 5 เกณฑ์ สำหรับนักเรียนที่มีระดับความสามารต่ำ ปานกลางค่อนข้างต่ำ ปานกลาง ปานกลางค่อนข้างสูง และ สูง เพื่อให้เห็นสารสนเทศที่ชัดเจนยิ่งขึ้น สำหรับการศึกษาในครั้งนี้ผู้วิจัยได้ใช้ข้อมูลจากนักเรียนที่อยู่ในกลุ่มต่ำ, กลุ่มปานกลาง และกลุ่มสูง เท่านั้น ส่วนนักเรียนที่จัดอยู่ในกลุ่มปานกลางค่อนข้างต่ำ และกลุ่มปานกลางค่อนข้างสูง จะไม่นำมาวิเคราะห์ ซึ่งเกณฑ์ที่ใช้ในการแบ่งนักเรียนแต่ละระดับชั้นตามกลุ่มความสามารถ (θ) ปรากฏดังตารางที่ 3.8

ตารางที่ 3.8 เกณฑ์การแบ่งนักเรียนแต่ละระดับชั้นตามกลุ่มความสามารถ (θ) โดยใช้คะแนนที่ได้จากการทำแบบสอบภายหลังการปรับเทียบคะแนนให้อยู่ในสเกลเดียวกันตามระดับชั้นของตน

ระดับชั้น	เกณฑ์การจำแนกกลุ่มความสามารถ		
	ต่ำ ($\theta \leq \theta$ - 0.84 SD.)	ปานกลาง (θ-0.25 SD.$<\theta\leq\theta$+0.25 SD.)	สูง ($\theta \geq \theta$ +0.84 SD.)
ม.1	$\theta \leq$ -2.586	-2.060 $<\theta<$ 1.614	$\theta \geq$ -1.088
ม.2	$\theta \leq$ -1.697	-1.112 $<\theta<$ -0.616	$\theta \geq$ -0.030
ม.3	$\theta \leq$ -0.779	-0.244 $<\theta<$ 0.208	$\theta \geq$ 0.743

การวิเคราะห์ข้อมูลในกลุ่มผู้สอบจำแนกตามระดับความสามารถดำเนินการตามขั้นตอนในทำนองเดียวกับข้อ 4.1.1 ถึง 4.1.4 ซึ่ง จากการวิเคราะห์ข้อมูลที่กล่าวมา ทั้งหมดแบ่งออกเป็น 4 ขั้นตอน ซึ่งสามารถสรุปขั้นตอนการวิเคราะห์ ดังแผนภาพที่ 3.2



แผนภาพที่ 3.2 สรุปขั้นตอนการวิเคราะห์ข้อมูล

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีจุดมุ่งหมาย เพื่อศึกษาประสิทธิภาพของการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง และศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ในแต่ละระดับชั้นโดยภาพรวมและในแต่ละกลุ่มความสามารถ คือ กลุ่มต่ำ สูง กลุ่มปานกลาง และกลุ่มสูง โดยใช้แบบสอบร่วมภายใน

แบบสอบที่ใช้ เป็นแบบวัดความสามารถทางคณิตศาสตร์ เรื่อง “จำนวนและการดำเนินการ” ระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ระดับชั้นละ 30 ข้อ ซึ่งในแบบสอบระดับชั้น ม.1 กับ ม.2 มีข้อสอบร่วมระหว่างระดับชั้นจำนวน 6 ข้อ และในแบบสอบระดับชั้น ม.2 กับ ม.3 มีข้อสอบร่วมระหว่างระดับชั้น 6 ข้อ นำแบบสอบดังกล่าวไปสอบกับกลุ่มปรับเทียบคะแนน และกลุ่มสอบทานผล กลุ่มปรับเทียบคะแนน ประกอบด้วย กลุ่มตัวอย่างนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 ซึ่งทำแบบวัดความสามารถทางคณิตศาสตร์ ระดับชั้น ม.1, ม.2 และ ม.3 ตามลำดับ ส่วนในกลุ่มสอบทานผล ประกอบด้วย กลุ่มตัวอย่างนักเรียนทุกระดับชั้นทำแบบสอบทั้ง 3 ฉบับ

ผู้วิจัยได้นำผลการสอบจากกลุ่มปรับเทียบคะแนนมาวิเคราะห์ และปรับเทียบคะแนนโดยใช้การปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ โดยอาศัยค่าคงที่ ที่ได้จากการวิเคราะห์โดยใช้โปรแกรม EQUATE 2.0 ปรับค่าความสามารถของผู้สอบทุกระดับชั้นให้อยู่ในมาตรวัดหรือสเกลเดียวกัน และทำการปรับแก้ค่าความสามารถที่มีระดับความสามารถแตกต่างกันมากๆ โดยการปรับเทียบเชิงเส้นตรง ด้วยวิธี “Simple Least Squared Method” แล้วสร้างเป็นตารางปรับเทียบคะแนนโดยอาศัยค่าความสามารถที่ปรับแล้วนั้น มาใช้ในการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนทั้ง 3 ระดับชั้น โดยภาพรวมและในแต่ละกลุ่มความสามารถ ส่วนคะแนนจากกลุ่มสอบทานผล นำมาวิเคราะห์หาความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนและความเพียงพอของการปรับเทียบคะแนน อันเป็นสารสนเทศที่สำคัญในการตรวจสอบประสิทธิภาพของการปรับเทียบคะแนน ดังเสนอผลการวิเคราะห์ข้อมูลเป็นดังนี้

ตอนที่ 1 ผลการวิเคราะห์ของแบบสอบ

ตอนที่ 2 ผลการสร้างการดำเนินการปรับเทียบคะแนนแบบสอบ

ตอนที่ 3 ผลการศึกษาประสิทธิภาพของการปรับเทียบคะแนน

ตอนที่ 4 ผลการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียน

ในการเสนอผลการวิเคราะห์ข้อมูล ผู้วิจัยได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ดังนี้

X_1 แทน คะแนนสอบของกลุ่มปรับเทียบคะแนน ระดับชั้น ม.1

X_2 แทน คะแนนสอบของกลุ่มปรับเทียบคะแนน ระดับชั้น ม.2

X_3 แทน คะแนนสอบของกลุ่มปรับเทียบคะแนน ระดับชั้น ม.3

X_{11} แทน คะแนนสอบของกลุ่มสอบทานผลระดับชั้น ม.1 จากการทำแบบสอบระดับชั้น ม.1

X_{12} แทน คะแนนสอบของกลุ่มสอบทานผลระดับชั้น ม.1 จากการทำแบบสอบระดับชั้น ม.2

X_{13} แทน คะแนนสอบของกลุ่มสอบทานผลระดับชั้น ม.1 จากการทำแบบสอบระดับชั้น ม.3

X_{21} แทน คะแนนสอบของกลุ่มสอบทานผลระดับชั้น ม.2 จากการทำแบบสอบระดับชั้น ม.1

X_{22} แทน คะแนนสอบของกลุ่มสอบทานผลระดับชั้น ม.2 จากการทำแบบสอบระดับชั้น ม.2

X_{23} แทน คะแนนสอบของกลุ่มสอบทานผลระดับชั้น ม.2 จากการทำแบบสอบระดับชั้น ม.3

X_{31} แทน คะแนนสอบของกลุ่มสอบทานผลระดับชั้น ม.3 จากการทำแบบสอบระดับชั้น ม.1

X_{32} แทน คะแนนสอบของกลุ่มสอบทานผลระดับชั้น ม.3 จากการทำแบบสอบระดับชั้น ม.2

X_{33} แทน คะแนนสอบของกลุ่มสอบทานผลระดับชั้น ม.3 จากการทำแบบสอบระดับชั้น ม.3

X^*_{12} แทน คะแนนสอบของกลุ่มสอบทานผลระดับชั้น ม.1 ที่แปลงไปสู่ระดับชั้น ม.2

X^*_{13} แทน คะแนนสอบของกลุ่มสอบทานผลระดับชั้น ม.1 ที่แปลงไปสู่ระดับชั้น ม.3

X^*_{21} แทน คะแนนสอบของกลุ่มสอบทานผลระดับชั้น ม.2 ที่แปลงไปสู่ระดับชั้น ม.1

X^*_{23} แทน คะแนนสอบของกลุ่มสอบทานผลระดับชั้น ม.2 ที่แปลงไปสู่ระดับชั้น ม.3

X^*_{31} แทน คะแนนสอบของกลุ่มสอบทานผลระดับชั้น ม.3 ที่แปลงไปสู่ระดับชั้น ม.1

X^*_{32} แทน คะแนนสอบของกลุ่มสอบทานผลระดับชั้น ม.3 ที่แปลงไปสู่ระดับชั้น ม.2

X^*_1 แทน คะแนนสอบของนักเรียนระดับชั้น ม.1 จากตารางปรับเทียบคะแนน

X^*_2 แทน คะแนนสอบของนักเรียนระดับชั้น ม.2 จากตารางปรับเทียบคะแนน

X^*_3 แทน คะแนนสอบของนักเรียนระดับชั้น ม.3 จากตารางปรับเทียบคะแนน

θ_1 แทน ค่าความสามารถของนักเรียนระดับชั้น ม.1

θ_2 แทน ค่าความสามารถของนักเรียนระดับชั้น ม.2

θ_3 แทน ค่าความสามารถของนักเรียนระดับชั้น ม.3

θ_{21} แทน ค่าความสามารถของนักเรียนระดับชั้น ม.2 ที่แปลงให้อยู่ในสเกลเดียวกับค่าความสามารถของนักเรียนระดับชั้น ม.1

θ_{23} แทน ค่าความสามารถของนักเรียนระดับชั้น ม.2 ที่แปลงให้อยู่ในสเกลเดียวกับค่าความสามารถของนักเรียนระดับชั้น ม.3

θ_{13} แทน ค่าความสามารถของนักเรียนระดับชั้น ม.1 ที่แปลงให้อยู่ในสเกลเดียวกับค่าความสามารถของนักเรียนระดับชั้น ม.3

ξ_2 แทน ค่าคะแนนจริงของนักเรียนระดับชั้น ม.2

ξ_{21} แทน ค่าคะแนนจริงของนักเรียนระดับชั้น ม.2 ที่แปลงคะแนนความสามารถให้อยู่ในสเกลเดียวกันกับระดับชั้น ม.1

ξ_{23} แทน ค่าคะแนนจริงของนักเรียนระดับชั้น ม.2 ที่แปลงคะแนนความสามารถให้อยู่ในสเกลเดียวกันกับระดับชั้น ม.3

ξ_{13} แทน ค่าคะแนนจริงของนักเรียนระดับชั้น ม.1 ที่แปลงคะแนนความสามารถให้อยู่ในสเกลเดียวกันกับระดับชั้น ม.3

SD_E แทน ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนของการปรับเทียบคะแนน

SEE แทน ความคลาดเคลื่อนมาตรฐาน (Standard Error) ของการปรับเทียบคะแนน

N แทน จำนวนนักเรียน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 1 ผลการวิเคราะห์ของแบบสอบ

1.1 ผลการวิเคราะห์ค่าสถิติพื้นฐานของแบบสอบแต่ละฉบับ

การวิเคราะห์ตอนนี้ ผู้วิจัยได้นำคะแนนแบบวัดความสามารถทางคณิตศาสตร์เฉพาะเนื้อหาในระดับชั้นของตน จากกลุ่มปรับเทียบคะแนนระดับชั้นมัธยมศึกษาปีที่ 1-3 มาวิเคราะห์หาค่าสถิติพื้นฐานของแบบสอบแต่ละฉบับ ดังแสดงในตารางที่ 4.1

ตารางที่ 4.1 ค่าสถิติพื้นฐานของแบบสอบแต่ละฉบับ

ค่าสถิติ	ระดับชั้นของแบบสอบ		
	ม.1	ม.2	ม.3
จำนวนนักเรียน	543	534	504
คะแนนเต็ม	30	30	30
คะแนนสูงสุด	29	30	30
คะแนนต่ำสุด	3	2	3
ค่าเฉลี่ย ($\bar{X}$)	13.070	12.636	13.654
ส่วนเบี่ยงเบนมาตรฐาน (SD)	5.454	7.739	6.295
ค่าความแปรปรวน (Variance)	29.747	59.901	39.638
มัธยฐาน (Median)	12	10	11
ฐานนิยม (Mode)	9	8	8
ความเบ้ (Sk)	0.722	1.093	1.408
ความโด่ง (Ku)	-0.008	-0.083	1.114
ความเที่ยง (KR-20)	0.803	0.912	0.857

ผลการวิเคราะห์จากตารางที่ 4.1 แสดงว่า จากการทำแบบวัดความสามารถทางคณิตศาสตร์ เรื่อง “จำนวนและการดำเนินการ” ของนักเรียนระดับมัธยมศึกษาปีที่ 1-3 จากกลุ่มปรับเทียบคะแนน ปรากฏว่า ในระดับชั้น ม.1 ตอบแบบสอบในระดับชั้นของตนได้คะแนนสูงสุดและต่ำสุด เท่ากับ 29 และ 3 ตามลำดับ โดยมีคะแนนเฉลี่ย 13.070 คะแนน ส่วนเบี่ยงเบนมาตรฐาน

เท่ากับ 5.454 การแจกแจงของคะแนนมีลักษณะเบ้ขวา แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อย และมีความโด่งมากกว่าโค้งปกติ แบบสอบนี้มีค่าความเที่ยง เท่ากับ 0.803

ในระดับชั้น ม.2 ตอบแบบสอบในระดับชั้นของตนได้คะแนนสูงสุดและต่ำสุด เท่ากับ 30 และ 2 ตามลำดับ โดยมีคะแนนเฉลี่ย 12.636 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 7.739 การแจกแจงของคะแนนมีลักษณะเบ้ขวา แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อย และมีความโด่งแบนกว่าโค้งปกติเล็กน้อย แบบสอบนี้มีค่าความเที่ยง เท่ากับ 0.912

สำหรับในระดับชั้น ม.3 ตอบแบบสอบในระดับชั้นของตนได้คะแนนสูงสุดและต่ำสุด เท่ากับ 30 และ 3 ตามลำดับ โดยมีคะแนนเฉลี่ย 13.654 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 6.295 การแจกแจงของคะแนนมีลักษณะเบ้ขวา แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อย และมีความโด่งมากกว่าโค้งปกติเล็กน้อย แบบสอบนี้มีค่าความเที่ยง เท่ากับ 0.857

เมื่อพิจารณาโดยรวม ค่าเฉลี่ย ค่ามัธยฐาน และฐานนิยมของคะแนนจากแบบสอบทั้ง 3 ฉบับ ต่ำกว่าครึ่งหนึ่งของคะแนนเต็ม แสดงว่าระดับความยากของแบบสอบทั้งสามฉบับค่อนข้างมาทางยาก จากค่าความเบ้และความโด่งของแบบสอบทั้งสามฉบับ จะเห็นว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อย และคะแนนจากแบบสอบทั้ง 3 ฉบับ มีการแจกแจงใกล้โค้งปกติ โดยคะแนนจากแบบสอบระดับชั้นมัธยมศึกษาปีที่ 1 มีความโด่งเกือบเท่าโค้งปกติ ส่วนในระดับชั้น ม.2 มีความโด่งแบนกว่าโค้งปกติเล็กน้อย และระดับชั้น ม. 3 มีความโด่งมากกว่าโค้งปกติเล็กน้อย

1.2 ค่าพารามิเตอร์ของข้อสอบร่วม

การวิเคราะห์ตอนนี้ผู้วิจัย ได้นำคะแนนแต่ละฉบับจากกลุ่มปรับเทียบคะแนนทั้ง 3 ระดับชั้น ไปวิเคราะห์ตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ เพื่อหาค่าพารามิเตอร์ของข้อสอบร่วม ปรากฏค่าพารามิเตอร์ของข้อสอบร่วมและค่าพารามิเตอร์ของข้อสอบร่วมเฉลี่ย ดังแสดงในตารางที่ 4.2 และ 4.3

ตารางที่ 4.2 ค่าพารามิเตอร์ของข้อสอบร่วมและค่าพารามิเตอร์ของข้อสอบร่วมเฉลี่ยระหว่างชั้น ม.1-ม.2

ข้อสอบร่วมจากแบบสอบระดับชั้น ม.1				ข้อสอบร่วมจากแบบสอบระดับชั้น ม.2			
ข้อ	a	b	C	**ข้อ**	a	b	c
25	1.646	1.462	0.160	1	2.432	0.830	0.161
26	4.366	1.458	0.130	2	4.770	1.059	0.133
27	1.126	1.749	0.213	3	2.974	1.129	0.312
28	2.494	0.811	0.176	4	6.326	0.597	0.226
29	1.460	1.000	0.215	5	2.819	0.805	0.229
30	2.261	1.573	0.171	6	3.781	0.973	0.200
$\bar{X}$	2.226	1.342	0.177	$\bar{X}$	3.850	0.898	0.210
SD	1.063	0.328	0.029	SD	1.343	0.178	0.056

จากตารางที่ 4.2 แสดงว่า ข้อสอบร่วมระหว่างระดับชั้น ม.1 กับ ม.2 ที่ได้จากการตอบแบบสอบของนักเรียนระดับชั้น ม.1 มีค่าอำนาจจำแนกตั้งแต่ 1.126 ถึง 4.366 ค่าความยากตั้งแต่ 0.811 ถึง 1.749 ค่าการเดามีค่าตั้งแต่ 0.171 ถึง 0.215 และข้อที่มีค่าความยากสูงสุด คือข้อ 27 ส่วนข้อสอบร่วมระหว่างระดับชั้น ม.1 กับ ม.2 ที่ได้จากการตอบแบบสอบของนักเรียนระดับชั้น ม.2 มีค่าอำนาจจำแนก ตั้งแต่ 2.432 ถึง 6.326 ค่าความยากตั้งแต่ 0.597 ถึง 1.129 ค่าการเดามีค่าตั้งแต่ 0.133 ถึง 0.229 และข้อที่ยากที่สุด คือ ข้อ 3 ซึ่งเป็นข้อเดียวกันกับข้อที่ยากที่สุดในระดับชั้น ม.1

ข้อสอบร่วมเฉลี่ยระหว่างระดับชั้น ม.1 กับ ม.2 ที่ได้จากการตอบแบบสอบของนักเรียนระดับชั้น ม.1 มีค่าอำนาจจำแนกเฉลี่ย 2.217 และมีค่าการเดาเฉลี่ย 0.177 ซึ่งค่าพารามิเตอร์โดยเฉลี่ยทั้งสองค่ามีค่าต่ำกว่าข้อสอบร่วมเฉลี่ยที่ได้จากการตอบแบบสอบของนักเรียนระดับชั้น ม.2 ที่มีค่าอำนาจจำแนกเฉลี่ย 3.850 และค่าการเดาเฉลี่ย 0.210 ส่วนค่าความยากของข้อสอบร่วม พบว่าระดับชั้น ม.1 มีค่าความยากเฉลี่ย 1.342 ซึ่งมีค่ามากกว่า ค่าความยากเฉลี่ยของนักเรียนระดับชั้น ม.2 โดยมีค่าความยากเฉลี่ย 0.898 โดยสรุปพบว่า ข้อสอบร่วมทั้ง 6 ข้อ มีความยากสำหรับระดับชั้น ม.1 มากกว่าระดับชั้น ม.2

ตารางที่ 4.3 ค่าพารามิเตอร์ของข้อสอบร่วมและค่าพารามิเตอร์ของข้อสอบร่วมเฉลี่ยระหว่างชั้น ม.2-ม.3

ข้อสอบร่วมจากแบบสอบระดับชั้น ม.2				ข้อสอบร่วมจากแบบสอบระดับชั้น ม.3			
ข้อ	**a**	**b**	**C**	**ข้อ**	**a**	**b**	**c**
25	2.275	1.807	0.222	1	2.113	1.735	0.142
26	6.497	0.658	0.230	2	2.072	0.396	0.123
27	2.602	1.127	0.202	3	2.878	0.905	0.126
28	4.015	0.915	0.206	6	1.171	0.659	0.231
29	3.871	0.656	0.173	5	2.632	0.424	0.364
30	4.426	1.131	0.169	6	2.520	0.907	0.251
$\bar{X}$	3.948	1.049	0.200	$\bar{X}$	2.231	0.837	0.206
SD	1.375	0.390	0.022	SD	0.551	0.449	0.086

จากตารางที่ 4.3 แสดงว่า ข้อสอบร่วมระหว่างระดับชั้น ม.2 กับ ม.3 ที่ได้จากการตอบแบบสอบของนักเรียนชั้นระดับชั้น ม.2 มีค่าอำนาจจำแนกตั้งแต่ 2.275 ถึง 6.497 ค่าความยากตั้งแต่ 0.656 ถึง 1.807 ค่าการเดามีค่าตั้งแต่ 0.169 ถึง 0.230 และข้อที่มีค่าความยากสูงสุด คือข้อ 25 ส่วนข้อสอบร่วมระหว่างระดับชั้น ม.2 กับ ม.3 ที่ได้จากการตอบแบบสอบของนักเรียนระดับชั้น ม.3 มีค่าอำนาจจำแนกตั้งแต่ 1.171 ถึง 2.878 ค่าความยากตั้งแต่ 0.396 ถึง 1.735 ค่าการเดามีค่าตั้งแต่ 0.123 ถึง 0.364 และข้อที่ยากที่สุด คือ ข้อ 1 ซึ่งเป็นข้อเดียวกันกับข้อที่ยากที่สุดในระดับชั้น ม.2

ข้อสอบร่วมเฉลี่ยระหว่างชั้น ม.2 กับ ม.3 ที่ได้จากการตอบแบบสอบของนักเรียนระดับชั้น ม. 2 มีค่าอำนาจจำแนกเฉลี่ย 3.948 และมีค่าการเดาเฉลี่ย 0.200 ส่วนค่าพารามิเตอร์ของข้อสอบร่วมเฉลี่ยที่ได้จากการตอบแบบสอบของนักเรียนระดับชั้น ม.2 ที่มีค่าอำนาจจำแนกเฉลี่ย 2.231 และค่าการเดาเฉลี่ย 0.206 ส่วนค่าความยากของข้อสอบร่วมพบว่า ระดับชั้น ม.2 มีค่าความยากเฉลี่ย 1.049 ซึ่งมีค่ามากกว่าค่าความยากเฉลี่ยของนักเรียนระดับชั้น ม.3 โดยมีความยากเฉลี่ย 0.837 โดยสรุปพบว่า ข้อสอบร่วมทั้ง 6 ข้อ มีความยากสำหรับระดับชั้น ม.2 มากกว่าระดับชั้น ม.3

ตอนที่ 2 ผลการสร้างการดำเนินการปรับเทียบคะแนนแบบสอบ

2.1 การปรับเทียบระดับความสามารถของผู้สอบโดยใช้ทฤษฎีตอบสนองข้อสอบแบบโลจิสติก 3 พารามิเตอร์

ผู้วิจัยได้นำแบบสอบระดับชั้น ม.1 ไปสอบกับนักเรียนระดับชั้น ม.1 จำนวน 543 คน แบบสอบระดับชั้น ม.2 ไปสอบกับนักเรียนระดับชั้น ม.2 จำนวน 534 คน และแบบสอบระดับชั้น ม.3 ไปสอบกับนักเรียนระดับชั้น ม.3 จำนวน 504 คน แล้วนำไปวิเคราะห์ด้วยโปรแกรม BILOG 3.04 ที่ใช้รูปแบบโลจิสติก 3 พารามิเตอร์ เพื่อให้ได้ค่าความสามารถของผู้สอบแต่ละคน แล้วหาค่าความสามารถเฉลี่ยของผู้สอบในแต่ละคะแนน ได้ผลดังตารางที่ 4.4

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.4 ค่าความสามารถ (θ) และคะแนนสอบ (X) ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 กลุ่มปรับเทียบคะแนน

ชั้นมัธยมศึกษาปีที่ 1		ชั้นมัธยมศึกษาปีที่ 2		ชั้นมัธยมศึกษาปีที่ 3	
X_1	θ_1	X_2	θ_2	X_3	θ_3
		2	-1.080		
3	-1.506	3	-0.986	3	-1.181
4	-1.396	4	-0.896	4	-1.421
5	-1.284	5	-0.789	5	-1.356
6	1.178	6	-0.689	6	-1.109
7	-0.083	7	-0.575	7	-1.093
8	0.007	8	-0.448	8	-0.954
9	0.120	9	-0.275	9	-0.818
10	0.251	10	-0.117	10	-0.649
11	0.364	11	-0.017	11	-0.531
12	0.403	12	0.093	12	-0.529
13	0.593	13	0.269	13	-0.441
14	0.775	14	0.390	14	-0.296
15	0.914	15	0.417	15	-0.159
16	1.068	16	0.512	16	-0.166
17	1.186	17	0.651	17	0.166
18	1.293	18	0.743	18	0.343
19	1.351	19	0.801	19	0.402
20	1.432	20	1.037	20	0.565
21	1.595	21	1.248	21	0.618
22	1.671	22	1.329	22	0.719
23	1.846	23	1.435	23	0.890
24	1.931	24	1.556	24	0.959
25	2.078	25	1.656	25	1.099
26	2.139	26	1.764	26	1.312
27	2.335	27	1.866	27	1.377
28	2.499	28	1.985	28	2.229
29	2.560	29	2.151	29	2.328
		30	2.377	30	2.581

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การวิเคราะห์ตอนนี้ ผู้วิจัยได้นำค่าพารามิเตอร์จากการวิเคราะห์ข้อสอบแต่ละฉบับด้วยโปรแกรม BILOG 3.04 โดยใช้รูปแบบโลจิสติก 3 พารามิเตอร์ เพื่อให้ได้ค่าประมาณพารามิเตอร์ของข้อสอบแต่ละข้อ คือ ค่าความยาก (b) ค่าอำนาจจำแนก (a) ค่าการเดา (c) และประมาณค่าความสามารถของผู้สอบ (θ) จากนั้นนำค่าพารามิเตอร์ดังกล่าวมาคำนวณหาค่าคงที่ ด้วยโปรแกรม EQUATE 2.0 เพื่อแปลงค่าพารามิเตอร์ของข้อสอบและประมาณค่าความสามารถของผู้สอบในแต่ละระดับชั้นให้อยู่ในสเกลดียวกัน โดยใช้สมการในการแปลงค่าความสามารถ ดังเสนอต่อไปนี้

$$\theta_{21} = 1.027\ \theta_2 + 0.927$$
$$\theta_{23} = 1.018\ \theta_2 - 0.938$$

จากสมการทั้ง 2 สมการข้างต้นใช้ในการปรับ (Transformation) ค่าความสามารถของนักเรียนทั้ง 3 ระดับชั้น คือ ม.1, ม.2 และ ม.3 ให้อยู่ในสเกลเดียวกัน โดยแปลงคะแนนที่ได้จากระดับชั้น ม. 2 ไปสู่ระดับชั้น ม.1 และระดับชั้น ม.3 นำค่าประมาณพารามิเตอร์ที่ได้ผ่านการแปลงให้อยู่ในสเกลเดียวกันดังกล่าวแล้วนั้น มาทำการคำนวณหาค่าคะแนนจริงของผู้สอบแบบสอบแต่ละชุดในแต่ละระดับชั้น ปรากฏผลดังตารางที่ 4.5 โดยใช้สูตรในการคำนวณ (อ้างใน ศิริชัย กาญจนวาสี, 2545) ดังนี้

$$\xi_2 = \sum_{i=1}^{30} P_i(\theta_2)$$
$$\xi_{21} = \sum_{j=1}^{30} P_j(\theta_{21})$$
$$\xi_{23} = \sum_{k=1}^{30} P_k(\theta_{23})$$

โดยที่ ξ_2 แทน คะแนนจริงของนักเรียนระดับชั้น ม.2 ซึ่งได้จากการตอบแบบสอบชั้น ม.2

ξ_{21} แทน คะแนนจริงของนักเรียนระดับชั้น ม.2 ที่แปลงไปสู่สเกลเดียวกันกับระดับชั้น ม.1

ξ_{23} แทน คะแนนจริงของนักเรียนระดับชั้น ม.2 ที่แปลงไปสู่สเกลเดียวกันกับระดับชั้น ม.3

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.5 การปรับเทียบคะแนนความสามารถและคะแนนจริงของนักเรียนกลุ่มปรับเทียบคะแนนระดับชั้น ม. 2 ให้อยู่ในสเกลเดียวกันกับระดับชั้น ม.1 และ ม.3 ตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์

θ_{21}	ξ_{21}	θ_{2}	ξ_{2}	θ_{23}	ξ_{23}
-0.188	12.62	-1.080	6.54	-2.038	6.48
-0.091	13.04	-0.986	6.55	-1.943	6.48
0.001	13.47	-0.896	6.56	-1.851	6.48
0.111	14.04	-0.789	6.58	-1.742	6.49
0.214	14.62	-0.689	6.61	-1.640	6.49
0.331	15.37	-0.575	6.65	-1.524	6.49
0.462	16.32	-0.448	6.73	-1.395	6.50
0.640	17.78	-0.275	6.89	-1.218	6.51
0.801	19.21	-0.117	7.15	-1.057	6.53
0.904	20.13	-0.017	7.40	-0.956	6.55
1.017	21.12	0.093	7.78	-0.843	6.58
1.198	22.63	0.269	8.73	-0.665	6.65
1.322	23.61	0.390	9.81	-0.541	6.73
1.350	23.82	0.417	10.11	-0.514	6.75
1.448	24.55	0.512	11.41	-0.416	6.84
1.590	25.56	0.651	13.92	-0.275	7.04
1.685	26.19	0.743	15.8	-0.182	7.23
1.745	26.55	0.801	17.02	-0.122	7.37
1.987	27.71	1.037	22.02	0.118	8.24
2.204	28.4	1.248	25.56	0.333	9.45
2.287	28.59	1.329	26.45	0.416	10.02
2.396	28.81	1.435	27.29	0.523	10.85
2.520	29.01	1.556	27.93	0.646	11.96
2.623	29.15	1.656	28.32	0.748	13.11
2.734	29.28	1.764	28.65	0.859	14.71
2.839	29.38	1.866	28.90	0.962	16.47
2.961	29.49	1.985	29.13	1.084	18.39
3.131	29.60	2.151	29.39	1.252	20.68
3.363	29.72	2.377	29.63	1.483	23.46

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

จากตารางที่ 4.5 ผลการวิเคราะห์ความสามารถของนักเรียนระดับชั้น ม.2 ที่ปรับ (Transformation) ไปหาความสามารถของนักเรียนระดับชั้น ม.1 และ ม.3 ซึ่งทำให้ได้ค่าความสามารถอยู่ในสเกลเดียวกันทั้ง 3 ระดับ พบว่า คะแนนความสามารถสมมูลของนักเรียนระดับชั้น ม.2 ที่นำมาปรับ (Transformation) ให้เทียบได้กับความสามารถของนักเรียนระดับชั้น ม.1 มีค่าสูงขึ้น ส่วนคะแนนความสามารถสมมูลของนักเรียนระดับชั้น ม.2 ที่นำมาปรับ (Transformation) ให้เทียบได้กับความสามารถของนักเรียนระดับชั้น ม.3 มีค่าลดลง ซึ่งสอดคล้องกับผลการวิเคราะห์คะแนนจริงของนักเรียนทั้ง 3 ระดับ ภายหลังการแปลงค่าความสามารถให้อยู่ในสเกลเดียวกัน พบว่า คะแนนจริงของนักเรียนระดับชั้น ม.2 ที่มีคะแนนความสามารถสมมูลกับนักเรียนระดับชั้น ม.1 มีค่าเพิ่มขึ้น ส่วนคะแนนจริงระดับชั้น ม.2 ที่มีคะแนนความสามารถสมมูลกับนักเรียนระดับชั้น ม.3 มีค่าลดลง

ตัวอย่างเช่น นักเรียนระดับชั้น ม.2 ที่ทำแบบวัดความสามารถทางคณิตศาสตร์ระดับชั้น ม.2 แล้วมีความสามารถที่ระดับคะแนนมาตรฐาน 0.093 โดยมีคะแนนจริง 7.78 คะแนน ถ้านำมาปรับเทียบกับความสามารถทางคณิตศาสตร์ของนักเรียนในระดับชั้น ม.1 จะมีระดับความสามารถที่ระดับคะแนนมาตรฐาน 1.017 เทียบได้กับคะแนนจริง 21.12 คะแนน และเมื่อปรับเทียบกับความสามารถทางคณิตศาสตร์ของนักเรียนในระดับชั้น ม.3 จะมีระดับความสามารถที่ระดับคะแนนมาตรฐาน -0.843 ซึ่งเทียบได้กับคะแนนจริง 6.58 คะแนน เป็นต้น

ทั้งนี้คะแนนการปรับเทียบคะแนนระหว่าง 3 ระดับชั้นมีค่าใกล้เคียงกันมากในช่วงตรงกับคะแนนความสามารถทางปลายด้านคะแนนสูงมากและต่ำมาก ส่วนในช่วงที่ตรงกับคะแนนความสามารถที่อยู่ตรงกลางจะมีค่าแตกต่างกันมากขึ้น และในการปรับคะแนนความสามารถของนักเรียนระดับชั้น ม.1 ให้อยู่ในสเกลเดียวกันกับระดับชั้น ม.3 ซึ่งเป็นระดับที่มีคะแนนความสามารถแตกต่างกันอย่างชัดเจน ในการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบดังกล่าวข้างต้น อาจเกิดความคลาดเคลื่อนสูงในการปรับเทียบคะแนน อาจเกิดปัญหาที่เรียกว่า “Out-of-Level Effect” ซึ่งส่งผลต่อความถูกต้องแม่นยำของการปรับเทียบคะแนน ผู้วิจัยจึงทำการปรับแก้ค่าที่ได้จากการปรับเทียบคะแนนความสามารถตามวิธีทฤษฎีตอบสนองข้อสอบดังกล่าวข้างต้น โดยใช้วิธีการปรับเทียบเชิงเส้นตรง ด้วยวิธี “Simple Least Square Regression” โดยจะกล่าวถึงในรายละเอียดต่อไป

2.2 การปรับเทียบระดับความสามารถของผู้สอบโดยใช้การปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ (IRT) แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง

จากผลการวิเคราะห์การปรับเทียบระดับความสามารถของผู้สอบตามวิธีทฤษฎีตอบสนองข้อสอบ อาจเกิดความคลาดเคลื่อนสูงในการปรับเทียบคะแนนในระดับคะแนนที่มีความสามารถแตกต่างกันมาก ๆ โดยเฉพาะการปรับเทียบคะแนนความสามารถระหว่างระดับชั้น ม.1 และ ม.3 เนื่องจากมีระดับความสามารถที่แตกต่างกันอย่างเห็นได้ชัด ด้วยเหตุนี้อาจเกิดปัญหาที่เรียกว่า **"Out-of-Level Effect"** (O'Brien and Tohn, 1984; Bielinksi, Thurlow & Scottt, 2000; Leung, 2003) ที่อาจส่งผลต่อความถูกต้องแม่นยำของการปรับเทียบคะแนน ผู้วิจัยจึงทำการปรับแก้ค่าที่ได้จากการปรับเทียบคะแนนความสามารถตามวิธีทฤษฎีตอบสนองข้อสอบดังกล่าวข้างต้น โดยใช้วิธีการเชิงเส้นตรง ที่เรียกว่า **"วิธีกำลังสองน้อยสุด" (Simple Least Square Regression)** ซึ่งเป็นวิธีการที่ไม่ยุ่งยาก โดยใช้โปรแกรม SPSS for Window Version 11.0 โดยกำหนดให้ θ_{21} เป็นตัวแปรอิสระ และ θ_{23} เป็นตัวแปรเกณฑ์ ที่เป็นผลมาจากการแปลงคะแนนความสามารถจากระดับชั้น ม.2 ไปสู่สเกลเดียวกันกับคะแนนความสามารถของนักเรียนระดับชั้น ม.1 โดยคะแนนความสามารถดังกล่าวจะส่งผลต่อการแปลงคะแนนความสามารถจากระดับชั้น ม. 2 ไปสู่สเกลเดียวกันกับคะแนนความสามารถของนักเรียนระดับชั้น ม.3 ค่าที่ได้จะใช้ในการประมาณค่าพารามิเตอร์เพื่ออธิบายการแปลงค่าความสามารถจากระดับชั้น ม.1 ไปสู่สเกลเดียวกันกับระดับชั้น ม. 3 โดยตรง ซึ่งผลจากการวิเคราะห์สามารถแสดงได้ดังสมการต่อไปนี้

$$\theta_{23} = 0.991\,\theta_{21} - 1.852$$

จากสมการดังกล่าวข้างต้นสามารถสรุปรวมเป็นสมการเชิงเส้นตรง 4 สมการ สำหรับแปลงค่าคะแนนความสามารถจากแบบสอบทั้ง 3 ระดับชั้นให้อยู่ในสเกลเดียวกัน ซึ่งปรากฏผลดังตารางที่ 4.6

ตารางที่ 4.6 ค่าคงที่ A, K และ สมการเชิงเส้นตรงในการแปลงระดับความสามารถ (θ) ของผู้สอบที่ได้จากแบบสอบต่างระดับชั้นให้อยู่ในสเกลเดียวกัน ตามรูปแบบการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ร่วมกับการปรับเทียบเชิงเส้นตรง

รูปแบบการแปลงระดับความสามารถ	รูปแบบการปรับเทียบ	ค่าคงที่		สมการเชิงเส้นตรง
		A	K	
1. แปลงจากแบบสอบระดับชั้น ม.2 ไปยังแบบสอบระดับชั้น ม.1	IRT	1.027	0.921	$\theta_{21} = 1.027\theta_2 + 0.921$
2. แปลงจากแบบสอบระดับชั้น ม.2 ไปยังแบบสอบระดับชั้น ม.3	IRT	1.018	-0.938	$\theta_{23} = 1.018\theta_2 - 0.938$
3. แปลงจากแบบสอบระดับชั้น ม.2 ที่มีคะแนนความสามารถสมมูลกับ ม.1 ไปยัง แบบสอบระดับชั้น ม.2 ที่มีคะแนนความสามารถสมมูลกับระดับชั้น ม.3	IRT + การปรับเทียบเชิงเส้นตรง	0.991	-1.852	$\theta_{23} = 0.991\theta_{21} - 1.852$
4. แปลงจากแบบสอบระดับชั้น ม.1 ไปยังแบบสอบระดับชั้น ม.3	IRT + การปรับเทียบเชิงเส้นตรง	0.991	-1.852	$\theta_{13} = 0.991\theta_1 - 1.852$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

จากตารางที่ 4.6 ค่าคงที่ A และ K ที่ได้รับเหล่านี้นำไปแทนในสมการเชิงเส้นตรง เพื่อใช้ในการแปลงค่าประมาณพารามิเตอร์ความสามารถที่ได้จากแบบสอบแต่ละคู่ ดังกล่าว โดยมีขั้นตอนการแปลง ดังนี้

2.3.1 แปลงค่าประมาณพารามิเตอร์ที่ได้จากการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ จากแบบสอบระดับชั้น ม.2 ที่มีคะแนนความสามารถสมมูลกับ ม.1 (θ_{21}) ไปยังแบบสอบระดับชั้น ม.2 ที่มีคะแนนความสามารถสมมูลกับระดับชั้น ม.3 (θ_{23}) โดยใช้สมการที่ 3 จากตารางที่ 4.6

2.3.2 แปลงค่าประมาณพารามิเตอร์จากแบบสอบระดับชั้น ม.1 ไปยังแบบสอบระดับชั้น ม.3 (θ_{13}) โดยใช้สมการที่ 4 จากตารางที่ 4.6 โดยสมการนี้จะมีประโยชน์ในการปรับเทียบคะแนนพัฒนาการความสามารถทางคณิตศาสตร์ จากระดับชั้น ม.1 ไปยังสเกลเดียวกันกับคะแนนความสามารถในระดับชั้น ม.3

ดังนั้นจะได้ค่าประมาณของแบบสอบทั้ง 3 ระดับชั้นที่แปลงไปยังสเกลเดียวกันทั้งหมด พร้อมทั้งสามารถคำนวณหาค่าคะแนนจริงของผู้สอบแบบสอบแต่ละระดับชั้น ซึ่งปรากฏผลดังตารางที่ 4.7

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.7 การปรับเทียบคะแนนความสามารถและคะแนนจริงของนักเรียนกลุ่มปรับเทียบคะแนนทั้ง 3 ระดับชั้นให้อยู่ในสเกลเดียวกัน ตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิตอร์ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง

θ_{21}	ξ_{21}	θ_{2}	ξ_{2}	θ_{23}	ξ_{23}
-0.188	12.62	-1.080	6.54	-2.038	6.48
-0.091	13.04	-0.986	6.55	-1.943	6.48
0.001	13.47	-0.896	6.56	-1.851	6.48
0.111	14.04	-0.789	6.58	-1.742	6.49
0.214	14.62	-0.689	6.61	-1.640	6.49
0.331	15.37	-0.575	6.65	-1.524	6.49
0.462	16.32	-0.448	6.73	-1.395	6.50
0.640	17.78	-0.275	6.89	-1.218	6.51
0.801	19.21	-0.117	7.15	-1.058	6.53
0.904	20.13	-0.017	7.40	-0.956	6.55
1.017	21.12	0.093	7.78	-0.844	6.58
1.198	22.63	0.269	8.73	-0.665	6.65
1.322	23.61	0.390	9.81	-0.542	6.73
1.350	23.82	0.417	10.11	-0.514	6.75
1.448	24.55	0.512	11.41	-0.417	6.84
1.590	25.56	0.651	13.92	-0.276	7.04
1.685	26.19	0.743	15.80	-0.182	7.22
1.745	26.55	0.801	17.02	-0.123	7.37
1.987	27.71	1.037	22.02	0.117	8.24
2.204	28.4	1.248	25.56	0.332	9.45
2.287	28.59	1.329	26.45	0.415	10.02
2.396	28.81	1.435	27.29	0.522	10.85
2.520	29.01	1.556	27.93	0.645	11.95
2.623	29.15	1.656	28.32	0.747	13.10
2.734	29.28	1.764	28.65	0.858	14.69
2.839	29.38	1.866	28.90	0.961	16.45
2.961	29.49	1.985	29.13	1.082	18.37
3.131	29.60	2.151	29.39	1.251	20.66
3.363	29.72	2.377	29.63	1.481	23.44

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

จากตารางที่ 4.7 ผลการแปลงคะแนนความสามารถของนักเรียนกลุ่มปรับเทียบภายหลังการปรับแก้ค่าที่ได้จากการปรับเทียบคะแนนความสามารถตามวิธีทฤษฎีตอบสนองข้อสอบ โดยใช้วิธีการปรับเทียบเชิงเส้นตรง ตามวิธีการ “กำลังสองน้อยสุด“ (Simple Least Square Regression) ซึ่งทำให้ได้ค่าความสามารถอยู่ในสเกลเดียวกันทั้ง 3 ระดับ พบว่า การปรับเทียบคะแนนในแต่ละคู่ มีแนวโน้มเช่นเดียวกับการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ ที่ปรากฏในตาราง 4.5 เช่น คะแนนความสามารถสมมูลของนักเรียนระดับชั้น ม.2 ที่นำมาปรับ (Transformation) ให้เทียบได้กับความสามารถของนักเรียนระดับชั้น ม.1 มีค่าสูงขึ้น ส่วนคะแนนความสามารถสมมูลของนักเรียนระดับชั้น ม.2 ที่นำมาปรับ (Transformation) ให้เทียบได้กับความสามารถของนักเรียนระดับชั้น ม.3 มีค่าลดลง ซึ่งสอดคล้องกับผลการวิเคราะห์คะแนนจริงของนักเรียนทั้ง 3 ระดับ ภายหลังการแปลงค่าความสามารถให้อยู่ในสเกลเดียวกัน พบว่า คะแนนจริงระดับชั้น ม.2 ที่มีคะแนนความสามารถสมมูลกับนักเรียนระดับชั้น ม.1 มีค่าเพิ่มขึ้น ส่วนคะแนนจริงระดับชั้น ม.2 ที่มีคะแนนความสามารถสมมูลกับนักเรียนระดับชั้น ม.3 มีค่าลดลง แต่ในการปรับระดับความสามารถจากระดับชั้น ม.2 ไปยังระดับชั้น ม.3 (θ_{23}) มีค่าเปลี่ยนแปลงไปจากเดิมเล็กน้อย ในทุกค่าคะแนนความสามารถเนื่องจากมีการเปลี่ยนแปลงสมการเชิงเส้นตรงดังกล่าวข้างต้น นอกจากนี้ยังสามารถปรับคะแนนความสามารถระดับชั้น ม.1 ไปหาคะแนนความสามารถของนักเรียนระดับชั้น ม.3 โดยใช้สมการเส้นตรงที่ได้จากการปรับเทียบเชิงเส้นตรงโดยตรง

โดยการแปลผลจากตารางที่ 4.7 พบว่า คะแนนความสามารถสมมูลของนักเรียนระดับชั้น ม.2 ที่นำมาแปลงค่าความสามารถให้เทียบได้กับความสามารถของนักเรียนระดับชั้น ม.1 และ ม.3 มีค่าสูงขึ้น ดังนั้นคะแนนจริงจึงมีแนวโน้มเช่นเดียวกับการปรับเทียบคะแนนความสามารถ ตัวอย่างเช่น

นักเรียนระดับชั้น ม.2 ที่ทำแบบสอบวัดความสามารถทางคณิตศาสตร์ ระดับชั้น ม.2 แล้วมีความสามารถที่ระดับคะแนนมาตรฐาน 0.093 โดยมีคะแนนจริง 7.78 คะแนน จากคะแนนเต็ม 30 คะแนน ถ้านำมาปรับเทียบกับความสามารถทางคณิตศาสตร์ของนักเรียนในระดับชั้น ม.1 จะมีความสามารถที่คะแนนมาตรฐาน 1.017 โดยมีคะแนนจริง 21.12 คะแนน จากคะแนนเต็ม 30 คะแนน และเมื่อปรับเทียบกับความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.3 มีค่าความสามารถที่ระดับคะแนนมาตรฐาน -0.844 โดยมีคะแนนจริง 6.58 คะแนน จากคะแนนเต็ม 30 คะแนน เป็นต้น

2.3 การปรับเทียบระดับคะแนน

ผลการวิเคราะห์ระดับคะแนนความสามารถของผู้สอบที่ทำแบบสอบทั้ง 3 ระดับชั้นให้อยู่ในสเกลเดียวกันแล้วนั้น จะนำมาใช้สำหรับการปรับเทียบหาคะแนนที่สมมูลกันระหว่างแบบสอบในแต่ละระดับชั้น เพื่อนำไปใช้ในการตรวจสอบประสิทธิภาพของการปรับเทียบคะแนน ซึ่งในการปรับเทียบคะแนนที่สมมูลกัน นิยมทำการปรับเทียบหาคะแนนดิบกันระหว่างแบบสอบที่ตรงกับระดับความสามารถที่สมมูลกัน หรือ การปรับเทียบคะแนนจริงระหว่างแบบสอบ ซึ่งปัญหาของการปรับเทียบสเกลโดยใช้คะแนนจริงคือไม่สามารถอธิบายคะแนนที่อยู่ต่ำกว่าระดับการเดา (Lower Asymtote) ได้ โดยจะให้ความหมายของคะแนนสมมูลเฉพาะคะแนนที่อยู่เหนือค่าเฉลี่ยของการเดา ซึ่งถึงแม้จะเป็นการปรับเทียบโดยใช้คะแนนจริง แต่ยังเป็นค่าที่ประมาณได้จากสูตรในการคำนวณ ดังนั้นจึงยังต้องมีความคลาดเคลื่อนอยู่ ส่วนการปรับเทียบโดยใช้คะแนนสังเกตได้นั้นเป็นการปรับเทียบคะแนนโดยประมาณ ซึ่งอธิบายคะแนนสมมูลจาก X และ Y ได้ครอบคลุมพิสัยของคะแนนสังเกตได้ Lord ได้กล่าวว่า ทั้งสองวิธีนี้มีความสอดคล้องกันมาก แต่การสรุปผลในการอ้างอิงต้องทำอย่างรอบคอบ (Lord, 1980; พรพิมล นาคเวช, 2537; พิชัย ละแมนชัย, 2538) ดังนั้นในการวิจัยครั้งนี้ จึงทำการปรับเทียบหาคะแนนสังเกตได้ ในที่นี้หมายถึงคะแนนดิบ ระหว่างแบบสอบที่ตรงกับระดับความสามารถที่สมมูลกัน

การปรับเทียบคะแนนของผู้สอบโดยนำค่าระดับความสามารถของนักเรียนที่แปลงจากระดับชั้น ม.2 ไปยังสเกลเดียวกันกับระดับชั้น ม.1 และระดับชั้น ม.3 จากตารางที่ 4.7 ไปเทียบหาคะแนนดิบ จากตารางที่ 4.4 ปรากฏผลดังตารางที่ 4.8

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.8 การปรับเทียบสเกลของคะแนนสอบวัดความสามารถทางคณิตศาสตร์ระหว่างระดับชั้น ม.1, ม.2 และ ม.3

X_1^*	X_2^*	X_3^*
6	2	
7	3	
8	4	
9	5	3
10	6	3
11	7	4
12	8	5
13	9	6
14	10	7
15	11	8
16	12	9
17	13	10
18	14	11
19	15	12
20	16	13
21	17	14
22	18	15
23	19	16
24	20	17
25	21	18
26	22	19
27	23	20
28	24	21
29	25	22
	26	23
	27	24
	28	25
	29	26
	30	27

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

จากตารางที่ 4.8 การปรับเทียบสเกลของคะแนนสอบวัดความสามารถทางคณิตศาสตร์ระหว่างระดับชั้น ม.1, ม.2 และ ม.3 โดยใช้วิธีการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง พบว่า คะแนนสอบระหว่างระดับชั้น ม.1, ม.2 และ ม.3 เรียงจากคะแนนต่ำสุดไปหาคะแนนสูงสุดที่ระดับชั้นเดียวกัน ซึ่งเห็นได้ว่า ในระดับชั้น ม.1 เริ่มเรียงคะแนนตั้งแต่ 6 ถึง 29 คะแนน ระดับชั้น ม.2 เริ่มเรียงคะแนนตั้งแต่ 2 ถึง 30 คะแนน และระดับชั้น ม.3 เริ่มเรียงคะแนนตั้งแต่ 3 ถึง 27 คะแนน โดยจากตารางปรับเทียบคะแนน พบว่าที่ระดับความสามารถที่สมมูลกัน นักเรียนที่อยู่ระดับชั้น ม.1 ถ้านำมาเทียบคะแนนคณิตศาสตร์ของนักเรียนระดับชั้น ม.3 จะมีคะแนนที่แตกต่างกันอย่างชัดเจน เช่น

ตัวอย่างที่ 1 นักเรียนที่อยู่ระดับชั้น ม.1 ทำคะแนนสอบวัดความสามารถทางคณิตศาสตร์ระดับชั้น ม.1 ได้ 12 คะแนน ถ้านำมาเทียบกับคะแนนคณิตศาสตร์ของนักเรียนระดับชั้น ม.2 มีคะแนนเท่ากับ 8 คะแนน และเมื่อเทียบกับคะแนนสอบคณิตศาสตร์ของนักเรียนระดับชั้น ม.3 มีคะแนน เท่ากับ 5 คะแนน

ตัวอย่างที่ 2 นักเรียนที่อยู่ระดับชั้น ม.3 ทำคะแนนแบบวัดความสามารถทางคณิตศาสตร์ ระดับชั้น ม.3 ได้ 20 คะแนน ถ้านำมาเทียบกับคะแนนแบบวัดความความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.2 มีคะแนนเท่ากับ 23 คะแนน และเมื่อเทียบกับคะแนนสอบวัดความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.1 มีคะแนน เท่ากับ 27 คะแนน

ตัวอย่างที่ 3 นักเรียนที่อยู่ระดับชั้น ม.2 ทำคะแนนความสามารถทางคณิตศาสตร์ระดับชั้น ม.2 ได้ 20 คะแนน ถ้านำมาเทียบกับคะแนนความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.1 มีคะแนนเท่ากับ 24 คะแนน และเมื่อเทียบกับคะแนนความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.3 จะมีคะแนนเท่ากับ 17 คะแนน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 3 ผลการศึกษาประสิทธิภาพของการปรับเทียบคะแนน

การวิเคราะห์ประสิทธิภาพของการปรับเทียบคะแนนตามวิธีทฤษฎีตอบสนองข้อสอบร่วมกับวิธีการปรับเทียบเชิงเส้นตรง ศึกษาโดยใช้ข้อมูลจากกลุ่มสอบทานผล ซึ่งเป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 จากกลุ่มประชากรการวิจัยเดียวกับกลุ่มปรับเทียบคะแนน โดยสุ่มนักเรียนมาประมาณร้อยละ 20 ของจำนวนนักเรียนในกลุ่มปรับเทียบคะแนน ซึ่งมีผลการวิเคราะห์ดังนี้

3.1 ค่าสถิติพื้นฐานของคะแนนสอบจากกลุ่มสอบทานผล

การวิเคราะห์ตอนนี้ ผู้วิจัยได้นำแบบวัดความสามารถทางคณิตศาสตร์ระดับชั้น ม.1, ม.2 และ ม.3 สอบกับกลุ่มสอบทานผล ที่เป็นนักเรียนระดับชั้น ม.1 จำนวน 113 คน นักเรียนระดับชั้น ม.2 จำนวน 127 คน และนักเรียนระดับชั้น ม.3 จำนวน 120 คน โดยทุกคนของแต่ละระดับชั้นตอบแบบวัดความสามารถทางคณิตศาสตร์ทั้งสามฉบับ ดังนั้นในแต่ละระดับจะมีคะแนนสอบของแบบสอบ 3 ฉบับ เมื่อรวมทั้ง 3 ระดับ ได้คะแนนของแบบวัดความสามารถทางคณิตศาสตร์ทั้งหมด 9 ฉบับ ซึ่งค่าสถิติพื้นฐานของคะแนนสอบทั้ง 9 ฉบับ แสดงดังตารางที่ 4.9

ตารางที่ 4.9 ค่าสถิติพื้นฐานของคะแนนสอบในระดับชั้น ม.1, ม.2 และ ม.3 จากกลุ่มสอบทานผล

คะแนน	N	$\bar{X}$	SD	Med	Mode	Max	Min
X_{11}	113	16.443	5.593	16	9	29	9
X_{12}	113	11.664	5.601	11	8	28	0
X_{13}	113	8.823	5.131	8	6	25	0
X_{21}	127	16.898	6.584	17	17	30	5
X_{22}	127	13.205	6.028	12	12	25	5
X_{23}	127	9.480	5.191	9	5	25	1
X_{31}	120	17.708	5.396	17	16	30	7
X_{32}	120	13.358	5.340	13	15	28	4
X_{33}	120	10.158	5.460	9.5	10	22	3

ผลการวิเคราะห์ตามตารางที่ 4.9 พบว่า กลุ่มสอบทานผลในระดับชั้น ม.1 ซึ่งเป็นนักเรียนในระดับชั้น ม.1 ทำแบบสอบระดับชั้น ม.1 ได้คะแนนสูงสุดและคะแนนต่ำสุด เท่ากับ 29 และ 9 คะแนน ตามลำดับ จากคะแนนเต็ม 30 คะแนน ได้คะแนนเฉลี่ย 16.443 คะแนน ส่วนเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 5.593 และนักเรียนส่วนใหญ่ได้คะแนน 9 คะแนน เมื่อนักเรียนระดับชั้น ม.1 กลุ่มนี้ทำแบบสอบระดับชั้น ม.2 ปรากฏว่าได้คะแนนสูงสุดและต่ำสุด เท่ากับ 28 และ 0 คะแนน ตามลำดับ ได้คะแนนเฉลี่ย 11.664 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 5.601 นักเรียนส่วนใหญ่ตอบแบบสอบชุดนี้ได้คะแนน 8 คะแนน และเมื่อตอบแบบสอบระดับชั้น ม.3 ปรากฏว่า ได้คะแนนสูงสุดและต่ำสุด เท่ากับ 25 และ 0 คะแนน ตามลำดับ ได้คะแนนเฉลี่ย 8.823 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 5.131 โดยนักเรียนส่วนใหญ่ตอบแบบสอบชุดนี้ได้ 6 คะแนน

เมื่อกลุ่มสอบทานผลในระดับชั้น ม.2 ซึ่งเป็นนักเรียนในระดับชั้น ม.2 ทำแบบสอบระดับชั้น ม.1 ปรากฏว่าได้คะแนนสูงสุดและต่ำสุด เท่ากับ 30 และ 5 คะแนน ตามลำดับ จากคะแนนเต็ม 30 คะแนน ได้คะแนนเฉลี่ย 16.898 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 6.584 นักเรียนส่วนใหญ่ตอบ แบบสอบชุดนี้ได้ 17 คะแนน เมื่อนักเรียนระดับชั้น ม.2 กลุ่มนี้ทำแบบสอบระดับชั้น ม.2 ได้คะแนนสูงสุดและคะแนนต่ำสุด เท่ากับ 25 และ 5 คะแนน ตามลำดับ ได้คะแนนเฉลี่ย 13.205 คะแนน ส่วนเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 6.028 นักเรียนส่วนใหญ่ได้คะแนน 12 คะแนน และเมื่อตอบแบบสอบระดับชั้น ม.3 ปรากฏว่า ได้คะแนนสูงสุดและต่ำสุด เท่ากับ 25 และ 1 คะแนน ตามลำดับ ได้คะแนนเฉลี่ย 9.480 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 5.191 โดยนักเรียนส่วนใหญ่ตอบแบบสอบชุดนี้ได้ 5 คะแนน

และกลุ่มสอบทานผลในระดับชั้น ม.3 ซึ่งเป็นนักเรียนในระดับชั้น ม.3 ทำแบบสอบระดับชั้น ม.1 ปรากฏว่าได้คะแนนสูงสุดและต่ำสุด เท่ากับ 30 และ 7 คะแนน ตามลำดับ จากคะแนนเต็ม 30 คะแนน ได้คะแนนเฉลี่ย 17.708 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 5.396 นักเรียนส่วนใหญ่ตอบแบบสอบชุดนี้ได้คะแนน 16 คะแนน เมื่อตอบแบบสอบระดับชั้น ม.2 ปรากฏว่า ได้คะแนนสูงสุดและต่ำสุด เท่ากับ 28 และ 4 คะแนน ตามลำดับ ได้คะแนนเฉลี่ย 13.354 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 5.340 นักเรียนส่วนใหญ่ตอบแบบสอบชุดนี้ได้ 4 คะแนน และเมื่อตอบแบบสอบระดับชั้น ม.3 ได้คะแนนสูงสุดและคะแนนต่ำสุด เท่ากับ 22 และ 3 คะแนน ตามลำดับ ได้คะแนนเฉลี่ย 10.154 คะแนน ส่วนเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 5.460 โดยนักเรียนส่วนใหญ่ได้คะแนน 10 คะแนน

3.2 ผลการวิเคราะห์ความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน

การวิเคราะห์ตอนนี้ ผู้วิจัยได้นำคะแนนของกลุ่มสอบทานผลของทุกระดับ แยกวิเคราะห์ในแต่ละระดับ ดังนี้

นำคะแนนของนักเรียนระดับชั้น ม.1 ที่ตอบแบบสอบระดับชั้น ม.1 ไปปรับเทียบคะแนนจากตาราง 4.8 เพื่อปรับเทียบให้เป็นคะแนนในระดับชั้น ม.2 (X^*_{12}) และระดับชั้น ม.3 (X^*_{13}) แล้วนำมาเปรียบเทียบกับคะแนนที่ตอบจากการสอบของแบบสอบระดับชั้น ม.2 (X_{12}) และแบบสอบระดับชั้น ม.3 (X_{13}) ตามลำดับ (คะแนนเกณฑ์) และนำคะแนนของนักเรียนระดับชั้น ม.2 ที่ตอบแบบสอบชั้น ม.2 ไปเทียบคะแนนจากตาราง เพื่อปรับเทียบให้เป็นคะแนนในระดับชั้น ม.1 (X^*_{21}) และระดับชั้น ม.3 (X^*_{23}) แล้วนำมาเปรียบเทียบกับคะแนนที่ตอบจากการสอบของแบบสอบระดับชั้น ม.1 (X_{21}) และแบบสอบระดับชั้น ม.3 (X_{23}) ตามลำดับ (คะแนนเกณฑ์) ในทำนองเดียวกัน นำคะแนนของนักเรียนระดับชั้น ม.3 ที่ตอบแบบสอบระดับชั้น ม.3 ไปปรับเทียบคะแนนจากตาราง เพื่อปรับให้เป็นคะแนนในระดับชั้น ม.1 (X^*_{31}) และระดับชั้น ม.2 (X^*_{32}) แล้วนำมาเปรียบเทียบกับคะแนนที่ตอบจากการสอบของแบบสอบระดับชั้น ม.1 (X_{31}) และแบบสอบระดับชั้น ม.2 (X_{32}) ตามลำดับ (คะแนนเกณฑ์) แล้วหาคะแนนความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนของทุกระดับชั้น ดังแสดงในตารางที่ 4.10

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.10 คะแนนความคลาดเคลื่อนเฉลี่ย ($\bar{X}_E$) ส่วนเบี่ยงเบนมาตรฐานของคะแนนความคลาดเคลื่อน (SD_E) และคะแนนความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน (SEE) จากคะแนนแบบสอบของนักเรียนในแต่ละระดับชั้นและในระดับคะแนนทั้งหมดของทุกระดับชั้น

รูปแบบการแปลงคะแนน	n	$\bar{X}_E$	SD_E	SEE
X^*_{12}	113	2.300	1.841	0.173
X^*_{13}	113	2.380	1.910	0.179
X^*_{21}	127	2.126	1.745	0.154
X^*_{23}	127	2.629	1.794	0.159
X^*_{31}	120	2.566	1.952	0.178
X^*_{32}	120	2.433	1.747	0.159
รวมรูปแบบการแปลงคะแนนทั้งหมดของทุกระดับชั้น	720	2.406	1.833	0.068

ผลการวิเคราะห์ตามตารางที่ 4.10 พบว่า คะแนนความคลาดเคลื่อนเฉลี่ยในระดับคะแนนทั้งหมดเมื่อแปลงคะแนนจากคะแนนแบบสอบระดับชั้น ม.1 ไปยังคะแนนแบบสอบระดับชั้น ม.2 (X^*_{12}) จากการสอบของกลุ่มสอบทานผลในระดับชั้น ม.1 มีค่าความคลาดเคลื่อนเฉลี่ย เท่ากับ 2.300 ส่วนเบี่ยงเบนมาตรฐานของคะแนนความคลาดเคลื่อนเท่ากับ 1.841 และ ค่าความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน เท่ากับ 0.173 ส่วนคะแนนความคลาดเคลื่อนเฉลี่ยในระดับคะแนนทั้งหมดเมื่อแปลงคะแนนจากคะแนน แบบสอบระดับชั้น ม.1 ไปยังคะแนนแบบสอบระดับชั้น ม.3 (X^*_{13}) จากการสอบของกลุ่มสอบทานผลในระดับชั้น ม.1 มีค่าความคลาดเคลื่อนเฉลี่ย เท่ากับ 2.380 ส่วนเบี่ยงเบนมาตรฐานของคะแนนความคลาดเคลื่อน เท่ากับ 1.910 และ ค่าความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนเท่ากับ 0.179 โดยรวมพบว่า คะแนนความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนแบบสอบระดับชั้น ม.1 ไปยังคะแนนแบบสอบระดับชั้น ม.2 มีค่าน้อยกว่าคะแนนความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนแบบสอบระดับชั้น ม.1 ไปยังคะแนนแบบสอบระดับชั้น ม.3 จากการสอบของกลุ่มสอบทานผลระดับชั้น ม.1 แต่มีค่าแตกต่างกันเพียงเล็กน้อย

ส่วนคะแนนความคลาดเคลื่อนเฉลี่ยในระดับคะแนนทั้งหมดเมื่อแปลงคะแนนจากคะแนนแบบสอบระดับชั้น ม.2 ไปยังคะแนนแบบสอบระดับชั้น ม.1 (X^*_{21}) จากการสอบของกลุ่มสอบทานผลในระดับชั้น ม.2 มีค่าความคลาดเคลื่อนเฉลี่ย เท่ากับ 2.126 ส่วนเบี่ยงเบนมาตรฐานของคะแนนความคลาดเคลื่อน เท่ากับ 1.745 และ ค่าความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน เท่ากับ 0.154 ส่วนคะแนนความคลาดเคลื่อนเฉลี่ยในระดับคะแนนทั้งหมดเมื่อแปลงคะแนนจากคะแนนแบบสอบระดับชั้น ม.2 ไปยังคะแนนแบบสอบระดับชั้น ม.3 (X^*_{23}) จากการสอบของกลุ่มสอบทานผลในระดับชั้น ม.2 มีค่าความคลาดเคลื่อนเฉลี่ย เท่ากับ 2.629 ส่วนเบี่ยงเบนมาตรฐานของคะแนนความคลาดเคลื่อน เท่ากับ 1.794 และ ค่าความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน เท่ากับ 0.179 โดยรวมพบว่า คะแนนความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน แบบสอบระดับชั้น ม.2 ไปยังคะแนนแบบสอบระดับชั้น ม.1 มีค่าน้อยกว่าคะแนนความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนแบบสอบระดับชั้น ม.2 ไปยังคะแนนแบบสอบระดับชั้น ม.3 จากการสอบของกลุ่มสอบทานผลระดับชั้น ม.2 แต่มีค่าแตกต่างกันน้อยมาก

และคะแนนความคลาดเคลื่อนเฉลี่ยในระดับคะแนนทั้งหมดเมื่อแปลงคะแนนจากคะแนนแบบสอบระดับชั้น ม.3 ไปยังคะแนนแบบสอบระดับชั้น ม.1 (X^*_{31}) จากการสอบของกลุ่มสอบทานผลในระดับชั้น ม.3 มีค่าความคลาดเคลื่อนเฉลี่ย เท่ากับ 2.566 ส่วนเบี่ยงเบนมาตรฐานของคะแนนความคลาดเคลื่อน เท่ากับ 1.952 และ ค่าความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน เท่ากับ 0.178 ส่วนคะแนนความคลาดเคลื่อนเฉลี่ยในระดับคะแนนทั้งหมดเมื่อแปลงคะแนนจากคะแนนแบบสอบระดับชั้น ม.3 ไปยังคะแนนแบบสอบระดับชั้น ม.2 (X^*_{32}) จากการสอบของกลุ่มสอบทานผลในระดับชั้น ม.3 มีคะแนนความคลาดเคลื่อนเฉลี่ย เท่ากับ 2.433 ส่วนเบี่ยงเบนมาตรฐานของคะแนนความคลาดเคลื่อน เท่ากับ 1.747 และ ค่าความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน เท่ากับ 0.159 โดยรวมพบว่า คะแนนความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนแบบสอบระดับชั้น ม.3 ไปยังคะแนนแบบสอบระดับชั้น ม.1 มีค่ามากกว่าคะแนนความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนแบบสอบระดับชั้น ม.3 ไปยังคะแนนแบบสอบระดับชั้น ม.2 จากการสอบของกลุ่มสอบทานผลระดับชั้น ม.3

เมื่อพิจารณาคะแนนความคลาดเคลื่อนเฉลี่ยในรูปแบบการแปลงคะแนนทั้งหมดของทุกระดับชั้น มีค่าเท่ากับ 2.406 ส่วนเบี่ยงเบนมาตรฐานของคะแนนความคลาดเคลื่อนโดยเฉลี่ย มีค่าเท่ากับ 1.833 และค่าความคลาดเคลื่อมาตรฐานของการปรับเทียบคะแนน มีค่าเท่ากับ 0.068

3.3 ผลการวิเคราะห์ความเพียงพอของความคลาดเคลื่อนของการปรับเทียบคะแนน

ความคลาดเคลื่อนของการปรับเทียบคะแนน ซึ่งเป็นค่าความคลาดเคลื่อนรวมของการปรับเทียบคะแนนทุกระดับชั้นกับเกณฑ์การยอมรับดัชนีความแตกต่าง

การพิจารณาดัชนีความแตกต่าง เริ่มคำนวณหาค่าคะแนนแปลงจากการปรับเทียบคะแนนของนักเรียนในกลุ่มสอบทานผลแต่ละคนตามระดับชั้นของตนไปเป็นคะแนนที่เท่าเทียมกับคะแนนแปลงเป็นระดับชั้นอื่นดังนี้ ระดับชั้น ม.1 มีคะแนนแบบสอบระดับชั้น ม.1 ไปเทียบคะแนนจากตาราง 4.8 เพื่อปรับให้เป็นคะแนนในระดับชั้น ม.2 (X^*_{12}) และระดับชั้น ม.3 (X^*_{13}) ที่อยู่ในสเกลเดียวกัน ส่วนระดับชั้น ม.2 มีคะแนนจากแบบสอบระดับชั้น ม.2 เทียบไปยังคะแนนแบบสอบระดับชั้น ม.1 (X^*_{21}) และคะแนนแบบสอบระดับชั้น ม.2 เทียบไปยังคะแนนแบบสอบระดับชั้น ม.3 (X^*_{23}) ที่อยู่ในสเกลเดียวกัน ในทำนองเดียวกัน นำคะแนนของนักเรียนระดับชั้น ม.3 เทียบคะแนนให้เป็นคะแนนในระดับชั้น ม.1 (X^*_{31}) และระดับชั้น ม.2 (X^*_{32}) ที่อยู่ในสเกลเดียวกัน โดยในแต่ละระดับชั้นใช้ตารางปรับเทียบคะแนนความสามารถตามตาราง 4.8 แปลงคะแนนแล้วหาความแตกต่างระหว่างคะแนนเกณฑ์ X_{12}, X_{13}, X_{21}, X_{23}, X_{31} และ X_{32} กับคะแนนแปลง X^*_{12}, X^*_{13}, X^*_{21}, X^*_{23}, X^*_{31} และ X^*_{32} ตามลำดับ ในระดับชั้น ม.1 จะมีคะแนนความแตกต่างของ X_{12} กับ X^*_{12} และ X_{13} กับ X^*_{13} ส่วนในระดับชั้น ม.2 มีคะแนนความแตกต่างของ X_{21} กับ X^*_{21} และ X_{23} กับ X^*_{23} และในระดับชั้น ม.3 จะมีคะแนนความแตกต่างของ X_{31} กับ X^*_{31} และ X_{32} กับ X^*_{32} ของแต่ละราย ใช้ผลดังกล่าวมาคำนวณหาค่าดัชนีความแตกต่างของความคลาดเคลื่อน (C) ดังสูตรหน้า 50 แล้วประเมินค่าดัชนีความแตกต่าง โดยใช้หลักของ Petersen และคณะ (1982) ซึ่งกำหนดคุณภาพของการปรับเทียบคะแนนดังนี้

		C	≤	$(0.05\ SD_x)^2$	หมายถึง	ระดับน่าพอใจมาก
$(0.05\ SD_x)^2$	<	C	≤	$(0.10\ SD_x)^2$	หมายถึง	ระดับน่าพอใจ
$(0.10\ SD_x)^2$	<	C	≤	$(0.15\ SD_x)^2$	หมายถึง	ระดับปานกลาง
$(0.15\ SD_x)^2$	<	C	≤	$(0.20\ SD_x)^2$	หมายถึง	ระดับไม่น่าพอใจ
$(0.20\ SD_x)^2$	<	C			หมายถึง	ระดับไม่น่าพอใจมาก

ผลการประเมินค่าดัชนีความแตกต่าง ดังแสดงในตารางที่ 4.11

ตารางที่ 4.11 การประเมินค่าดัชนีความแตกต่าง (C) เมื่อแปลงคะแนนแบบสอบระดับชั้น ม.1 ไปยังคะแนนแบบสอบระดับ ม.2 จากกลุ่มสอบทานผลในระดับ ม.1

$\sum_{i=1}^{n}(X_i - X_i^*)^2$	n	S_x^2	S_x	C
978	113	31.368	5.600	0.275

	C	≤	0.078	หมายถึง	ระดับน่าพอใจมาก
0.078	< C**	≤	0.313	หมายถึง	ระดับน่าพอใจ
0.313	< C	≤	0.705	หมายถึง	ระดับปานกลาง
0.705	< C	≤	1.254	หมายถึง	ระดับไม่น่าพอใจ
1.254	< C			หมายถึง	ระดับไม่น่าพอใจมาก

ผลการวิเคราะห์ตามตารางที่ 4.11 แสดงว่า ดัชนีความแตกต่าง (C) ของการปรับเทียบคะแนนเมื่อแปลงคะแนนแบบสอบระดับบชั้น ม.1 ไปยังแบบสอบระดับชั้น ม.2 จากกลุ่มสอบทานผลในระดับชั้น ม.1 มีค่า 0.275 ซึ่งน้อยกว่ากำลังสองของร้อยละ 10 ของค่าเบี่ยงเบนมาตรฐานของแบบสอบระดับ ม.2 ซึ่งเท่ากับ 0.313 นั่นคือ ดัชนีความแตกต่าง (C) ของการแปลงคะแนนจากคะแนนแบบสอบระดับ ม.1 ไปยังคะแนนแบบสอบระดับ ม.2 ให้ความเพียงพอของการปรับเทียบคะแนนอยู่ในระดับน่าพอใจ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.12 การประเมินค่าดัชนีความแตกต่าง (C) เมื่อแปลงคะแนนแบบสอบระดับชั้น ม.1 ไปยังคะแนนแบบสอบระดับ ม.3 จากกลุ่มสอบทานผลในระดับ ม.1

$\sum_{i=1}^{n}(X_i - X_i^*)^2$	n	S_x^2	S_x	C
1,049	113	26.325	5.130	0.352

	C	≤	0.065	หมายถึง	ระดับน่าพอใจมาก
0.065	< C	≤	0.263	หมายถึง	ระดับน่าพอใจ
0.263	< C**	≤	0.592	หมายถึง	ระดับปานกลาง
0.592	< C	≤	1.053	หมายถึง	ระดับไม่น่าพอใจ
1.053	< C			หมายถึง	ระดับไม่น่าพอใจมาก

ผลการวิเคราะห์ตามตารางที่ 4.12 แสดงว่า ดัชนีความแตกต่าง (C) ของการปรับเทียบคะแนนเมื่อแปลงคะแนนแบบสอบระดับบชั้น ม.1 ไปยังแบบสอบระดับชั้น ม.3 จากกลุ่มสอบทานผลในระดับชั้น ม.1 มีค่า 0.352 ซึ่งน้อยกว่ากำลังสองของร้อยละ 15 ของค่าเบี่ยงเบนมาตรฐานของแบบสอบระดับ ม.3 ซึ่งเท่ากับ 0.592นั่นคือ ดัชนีความแตกต่าง (C) ของการแปลงคะแนนจากคะแนนแบบสอบระดับ ม.1 ไปยังคะแนนแบบสอบระดับ ม.3 ให้ความเพียงพอของการปรับเทียบคะแนนอยู่ในระดับปานกลาง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.13 การประเมินค่าดัชนีความแตกต่าง (C) เมื่อแปลงคะแนนแบบสอบระดับชั้น ม.2 ไปยังคะแนนแบบสอบระดับ ม.1 จากกลุ่มสอบทานผลในระดับ ม.2

$\sum_{i=1}^{n}(X_i - X_i^*)^2$	n	S_x^2	S_x	C
958	127	43.346	6.583	0.160

	C	≤	0.108	หมายถึง	ระดับน่าพอใจมาก
0.108	< C**	≤	0.433	หมายถึง	ระดับน่าพอใจ
0.433	< C	≤	0.975	หมายถึง	ระดับปานกลาง
0.975	< C	≤	1.733	หมายถึง	ระดับไม่น่าพอใจ
1.733	< C			หมายถึง	ระดับไม่น่าพอใจมาก

ผลการวิเคราะห์ตามตารางที่ 4.13 แสดงว่า ดัชนีความแตกต่าง (C) ของการปรับเทียบคะแนนเมื่อแปลงคะแนนแบบสอบระดับบชั้น ม.2 ไปยังแบบสอบระดับชั้น ม.1 จากกลุ่มสอบทานผลในระดับชั้น ม.2 มีค่า 0.160 ซึ่งน้อยกว่ากำลังสองของร้อยละ 10 ของค่าเบี่ยงเบนมาตรฐานของแบบสอบระดับ ม.1 ซึ่งเท่ากับ 0.433 นั่นคือ ดัชนีความแตกต่าง (C) ของการแปลงคะแนนจากคะแนนแบบสอบระดับ ม.2 ไปยังคะแนนแบบสอบระดับ ม.1 ให้ความเพียงพอของการปรับเทียบคะแนนอยู่ในระดับน่าพอใจ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.14 การประเมินค่าดัชนีความแตกต่าง (C) เมื่อแปลงคะแนนแบบสอบระดับชั้น ม.2 ไปยังคะแนนแบบสอบระดับ ม.3 จากกลุ่มสอบทานผลในระดับ ม.2

$\sum_{i=1}^{n}(X_i - X_i^*)^2$	n	S_x^2	S_x	C
1,284	127	38.332	6.191	0.263

		C ≤	0.095	หมายถึง	ระดับน่าพอใจมาก
0.095	<	C** ≤	0.383	หมายถึง	ระดับน่าพอใจ
0.383	<	C ≤	0.862	หมายถึง	ระดับปานกลาง
0.862	<	C ≤	1.533	หมายถึง	ระดับไม่น่าพอใจ
1.533	<	C		หมายถึง	ระดับไม่น่าพอใจมาก

ผลการวิเคราะห์ตามตารางที่ 4.14 แสดงว่า ดัชนีความแตกต่าง (C) ของการปรับเทียบคะแนนเมื่อแปลงคะแนนแบบสอบระดับบชั้น ม.2 ไปยังแบบสอบระดับชั้น ม.3 จากกลุ่มสอบทานผลในระดับชั้น ม.2 มีค่า 0.263 ซึ่งน้อยกว่ากำลังสองของร้อยละ 10 ของค่าเบี่ยงเบนมาตรฐานของแบบสอบระดับ ม.3 ซึ่งเท่ากับ 0.383 นั่นคือ ดัชนีความแตกต่าง (C) ของการแปลงคะแนนจากคะแนนแบบสอบระดับ ม.2 ไปยังคะแนนแบบสอบระดับ ม.3 ให้ความเพียงพอของการปรับเทียบคะแนนอยู่ในระดับน่าพอใจ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.15 การประเมินค่าดัชนีความแตกต่าง (C) เมื่อแปลงคะแนนแบบสอบระดับชั้น ม.3 ไปยังคะแนนแบบสอบระดับ ม.1 จากกลุ่มสอบทานผลในระดับ ม.3

$\sum_{i=1}^{n}(X_i - X_i^*)^2$	n	S_x^2	S_x	C
1,244	120	29.115	5.395	0.356

		C	≤	0.072	หมายถึง	ระดับน่าพอใจมาก
0.072	<	C	≤	0.291	หมายถึง	ระดับน่าพอใจ
0.291	<	C**	≤	0.655	หมายถึง	ระดับปานกลาง
0.655	<	C	≤	1.164	หมายถึง	ระดับไม่น่าพอใจ
1.164	<	C			หมายถึง	ระดับไม่น่าพอใจมาก

ผลการวิเคราะห์ตามตารางที่ 4.15 แสดงว่า ดัชนีความแตกต่าง (C) ของการปรับเทียบคะแนนเมื่อแปลงคะแนนแบบสอบระดับบชั้น ม.3 ไปยังแบบสอบระดับชั้น ม.1 จากกลุ่มสอบทานผลในระดับชั้น ม.3 มีค่า 0.356 ซึ่งน้อยกว่ากำลังสองของร้อยละ 15 ของค่าเบี่ยงเบนมาตรฐานของแบบสอบระดับ ม.1 ซึ่งเท่ากับ 0.655 นั่นคือ ดัชนีความแตกต่าง (C) ของการแปลงคะแนนจากคะแนนแบบสอบระดับ ม.3 ไปยังคะแนนแบบสอบระดับ ม.1 ให้ความเพียงพอของการปรับเทียบคะแนนอยู่ในระดับปานกลาง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.16 การประเมินค่าดัชนีความแตกต่าง (C) เมื่อแปลงคะแนนแบบสอบระดับชั้น ม.3 ไปยังคะแนนแบบสอบระดับ ม.2 จากกลุ่มสอบทานผลในระดับ ม.3

$\sum_{i=1}^{n}(X_i - X_i^*)^2$	n	S_x^2	S_x	C
1,074	120	28.517	5.340	0.313

		C	≤	0.071	หมายถึง	ระดับน่าพอใจมาก
0.071	<	C	≤	0.285	หมายถึง	ระดับน่าพอใจ
0.285	<	C**	≤	0.641	หมายถึง	ระดับปานกลาง
0.641	<	C	≤	1.140	หมายถึง	ระดับไม่น่าพอใจ
1.140	<	C			หมายถึง	ระดับไม่น่าพอใจมาก

ผลการวิเคราะห์ตามตารางที่ 4.16 แสดงว่า ดัชนีความแตกต่าง (C) ของการปรับเทียบคะแนนเมื่อแปลงคะแนนแบบสอบระดับบชั้น ม.3 ไปยังแบบสอบระดับชั้น ม.2 จากกลุ่มสอบทานผลในระดับชั้น ม.3 มีค่า 0.313 ซึ่งน้อยกว่ากำลังสองของร้อยละ 15 ของค่าเบี่ยงเบนมาตรฐานของแบบสอบระดับ ม.2 ซึ่งเท่ากับ 0.641 นั่นคือ ดัชนีความแตกต่าง (C) ของการแปลงคะแนนจากคะแนนแบบสอบระดับ ม.3 ไปยังคะแนนแบบสอบระดับ ม.2 ให้ความเพียงพอของการปรับเทียบคะแนนอยู่ในระดับปานกลาง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.17 การประเมินค่าดัชนีความแตกต่าง (C) ซึ่งเป็นความคลาดเคลื่อนรวมจากทุกรูปแบบการแปลงคะแนน ที่ปรับเทียบคะแนนตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง

$\sum_{i=1}^{n}(X_i - X_i^*)^2$	n	S_x^2	S_x	C
6587	720	42.458	6.516	0.215

	C	≤	0.106	หมายถึง	ระดับน่าพอใจมาก
0.106	< C**	≤	0.424	หมายถึง	ระดับน่าพอใจ
0.424	< C	≤	0.955	หมายถึง	ระดับปานกลาง
0.953	< C	≤	1.698	หมายถึง	ระดับไม่น่าพอใจ
1.698	< C			หมายถึง	ระดับไม่น่าพอใจมาก

ผลการวิเคราะห์ตามตารางที่ 4.17 การประเมินค่าดัชนีความแตกต่าง (C) ของคะแนนที่นักเรียนแต่ละคนทำได้จากกลุ่มสอบทานผล ซึ่งเป็น**ความคลาดเคลื่อนรวมทั้งหมดจากรูปแบบการแปลงคะแนน** มีค่าดัชนีความแตกต่าง เท่ากับ 0.215 เมื่อเปรียบเทียบกับเกณฑ์ที่กำหนดไว้ ปรากฏว่า ความเพียงพอในการปรับเทียบคะแนนอยู่ในระดับที่น่าพอใจ เนื่องจากมีค่าน้อยกว่ากำลังสองของร้อยละ 10 ของค่าเบี่ยงเบนมาตรฐานของแบบสอบที่เป็นคะแนนเกณฑ์ แสดงว่าการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ร่วมกับวิธีการปรับเทียบเชิงเส้นตรงที่ใช้ในการวิจัยนี้ให้ผลการปรับเทียบคะแนนมีประสิทธิภาพในระดับน่าพอใจ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 4 ผลการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียน

ในการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 นั้น ผู้วิจัยได้ศึกษาประสิทธิภาพของการปรับเทียบคะแนน เพื่อให้ได้วิธีการที่เหมาะสมมาใช้ในการศึกษาพัฒนาการ ผลปรากฏว่า การปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบร่วมกับวิธีการปรับเทียบเชิงเส้นตรง มีความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนอยู่ในระดับต่ำ และวิธีการดังกล่าวเมื่อเปรียบเทียบกับเกณฑ์ที่กำหนดมีความเพียงพอในการปรับเทียบคะแนนที่อยู่ในระดับน่าพอใจ ผู้วิจัยจึงนำรูปแบบวิธีการดังกล่าวมาใช้ในการศึกษาพัฒนาการต่อ โดยปรับเทียบคะแนนของนักเรียนระดับชั้น ม.1 ที่ได้จากการตอบแบบสอบชั้นระดับชั้น ม.1 และปรับเทียบคะแนนของนักเรียนระดับชั้น ม.2 จากการตอบแบบสอบระดับชั้น ม.2 ทั้ง 2 ระดับไปสู่คะแนนที่สมมูลกันในสเกลคะแนนของนักเรียนระดับชั้น ม.3

สำหรับการวิเคราะห์ในตอนนี้ ภายหลังการปรับเทียบคะแนนความสามารถให้อยู่ในสเกลเดียวกันของทั้ง 3 ระดับชั้น ผู้วิจัยนำคะแนนความสามารถภายหลังการปรับเทียบดังกล่าวมาคำนวณหาค่าคะแนนจริงของแต่ละคน โดยไม่ได้แสดงผลการปรับเทียบที่เป็นคะแนนดิบเนื่องจากจะมีบางค่าคะแนนความสามารถที่ไม่สามารถประมาณค่าคะแนนดิบได้ เนื่องจากระดับความสามารถมีความแตกต่างกัน โดยเฉพาะการปรับเทียบจากระดับชั้น ม.1 ไปสู่ระดับชั้น ม.3 แนวทางแก้ปัญหา คือใช้สมการเชิงเส้นตรงที่ได้จากการปรับเทียบเชิงเส้นตรงมาทำการปรับเทียบระดับความสามารถแล้วคำนวณหาค่าคะแนนจริงของผู้สอบแต่ละคน ซึ่งวิธีการดังกล่าวจะสามารถประมาณค่าความสามารถให้อยู่ในสเกลเดียวกันได้ทุกค่า และสามารถศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนได้ทุกคนแล้วหาค่าสถิติพื้นฐานของคะแนนสอบและผลการศึกษาพัฒนาการแสดงดังต่อไปนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

4.1 การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนโดยภาพรวม

ภายหลังการปรับเทียบคะแนนของนักเรียนระดับชั้น ม.1 และ ม.2 ไปสู่คะแนนที่สมมูลกันในสเกลคะแนนของนักเรียนระดับชั้น ม.3 ปรากฏค่าสถิติพื้นฐานของคะแนนความสามารถ (θ) และ คะแนนจริง (ξ) จากผลการสอบของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 ภายหลังการปรับเทียบคะแนนในกลุ่มรวม ดังตารางที่ 4.18

ตารางที่ 4.18 ค่าสถิติพื้นฐานของคะแนนความสามารถ (θ) และ คะแนนจริง (ξ) จากผลการสอบของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 ภายหลังการปรับเทียบคะแนนโดยภาพรวม

ค่าสถิติ	ม.1		ม.2		ม.3	
	θ	ξ	θ	ξ	θ	ξ
จำนวนข้อสอบ	30	30	30	30	30	30
จำนวนผู้สอบ	543	543	534	534	504	504
คะแนนสูงสุด	0.685	12.370	1.487	23.500	2.581	29.290
คะแนนต่ำสุด	-3.490	6.480	-2.032	6.480	-1.333	6.500
คะแนนเฉลี่ย	-1.837	6.629	-0.864	8.348	-0.018	10.255
ส่วนเบี่ยงเบนมาตรฐาน	0.892	0.582	0.992	4.290	0.906	6.432
มัธยฐาน	-1.845	6.480	-1.148	6.520	-0.205	7.175
ฐานนิยม	-1.475	6.480	1.294	6.490	1.333	6.550
ความเบ้	0.440	6.246	0.913	2.401	1.032	1.974
ความโด่ง	-0.311	46.084	-0.247	4.362	0.392	2.465

ผลการวิเคราะห์ตามตาราง 4.12 ค่าสถิติพื้นฐานจากผลการสอบของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 ภายหลังการแปลงคะแนนความสามารถของนักเรียนในระดับชั้น ม.1 กับ ม.2 ไปสู่คะแนนที่สมมูลกันในสเกลคะแนนของนักเรียนระดับชั้น ม.3 แล้ว ปรากฏว่า ในระดับชั้น ม.1 ได้คะแนนความสามารถสูงสุดและต่ำสุดเท่ากับ 0.685 และ -3.490 ตามลำดับ โดยมีคะแนนความสามารถเฉลี่ย -1.837 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.892 การแจกแจงของคะแนนมีลักษณะเบ้ขวา แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อยและมีความโด่งแบน

กว่าโค้งปกติ ส่วนในระดับชั้น ม. 2 ได้คะแนนความสามารถสูงสุดและต่ำสุด เท่ากับ 1.487 และ -2.032 ตามลำดับ โดยมีคะแนนความสามารถเฉลี่ย -0.864 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.992 การแจกแจงของคะแนนมีลักษณะเบ้ขวา แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อยและมีความโด่งแบนกว่าโค้งปกติเล็กน้อย และในระดับชั้น ม.3 ได้คะแนนความสามารถสูงสุดและต่ำสุด เท่ากับ 2.581 และ -1.333 ตามลำดับ โดยมีคะแนนความสามารถเฉลี่ย -0.018 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.906 การแจกแจงของคะแนนมีลักษณะเบ้ขวา แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อยและมีความโด่งสูงกว่าโค้งปกติเล็กน้อย

เมื่อพิจารณาคะแนนจริงของนักเรียนในระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 พบว่า มีแนวโน้มเช่นเดียวกับคะแนนความสามารถ โดย คะแนนจริงของนักเรียนในระดับชั้น ม.1 ได้คะแนนสูงสุดและต่ำสุด เท่ากับ 12.37 และ 6.48 ตามลำดับ โดยมีคะแนนเฉลี่ย เท่ากับ 6.629 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.582 การแจกแจงของคะแนนมีลักษณะเบ้ขวา แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อยและมีความโด่งมาก ส่วนในระดับชั้น ม. 2 ได้คะแนนสูงสุดและต่ำสุด เท่ากับ 23.50 และ 6.48 คะแนน ตามลำดับ โดยมีคะแนนเฉลี่ย 8.348 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 4.290 การแจกแจงของคะแนนมีลักษณะเบ้ขวา แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อยและมีความโด่งมาก และในระดับชั้น ม.3 ได้คะแนนสูงสุดและต่ำสุด เท่ากับ 29.29 และ 6.50 คะแนน ตามลำดับ โดยมีคะแนนเฉลี่ย 10.255 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 6.432 การแจกแจงของคะแนนมีลักษณะเบ้ขวา แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อยและมีความโด่งมาก

เมื่อพิจารณาโดยรวม พบว่าคะแนนความสามารถและคะแนนจริงโดยเฉลี่ยของนักเรียนทั้ง 3 ระดับชั้น เพิ่มขึ้นตามระดับชั้นที่สูงขึ้น นั่นคือ นักเรียนในระดับชั้น ม.3 มีคะแนนความสามารถและคะแนนจริงโดยเฉลี่ยสูงกว่านักเรียนในระดับชั้น ม.2 และนักเรียนในระดับชั้น ม.2 มีคะแนนความสามารถและคะแนนจริงโดยเฉลี่ยสูงกว่านักเรียนในระดับชั้น ม.1 ตามลำดับ อีกทั้งเป็นที่น่าสังเกตว่าลักษณะการแจกแจงคะแนนของนักเรียนมีลักษณะเบ้ขวาทั้ง 3 ระดับชั้น แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อย

ผู้วิจัยทดสอบความแตกต่างระหว่างคะแนนความสามารถเฉลี่ยของนักเรียนทั้ง 3 ระดับชั้น โดยการวิเคราะห์ความแปรปรวนแบบทางเดียว (One-Way ANOVA) ภายหลังเมื่อปรับเทียบคะแนนผลการสอบของนักเรียนระดับชั้น มัธยมศึกษาปีที่ 1, 2 และ 3 ให้อยู่ในสเกลเดียวกัน ผลการวิเคราะห์สรุปได้ดังนี้

ตารางที่ 4.19 ผลการวิเคราะห์ความแปรปรวนทางเดียวของคะแนนความสามารถของนักเรียนระดับชั้น ม.1 ม.2 และ ม.3 โดยภาพรวม

แหล่งความแปรปรวน	SS	df	MS	F	P
ระหว่างกลุ่ม	868.138	2	434.069	500.203	.000
ภายในกลุ่ม	1369.367	1578	0.868		
รวม	2237.506	1580			

จากตารางที่ 4.19 แสดงว่าเมื่อปรับเทียบคะแนนความสามารถของนักเรียนทั้ง 3 ระดับชั้นให้อยู่ในสเกลเดียวกันแล้ว ผลการเปรียบเทียบความแตกต่างของค่าเฉลี่ยของคะแนนความสามารถ 3 ระดับชั้น แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .01 ($F = 500.203$, $p = .000$) และความแปรปรวนของคะแนนความสามารถของนักเรียน 3 ระดับชั้น มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .01 (Levene Statistic = 4.497, $p = 0.011$) จึงทำการทดสอบความแตกต่างของคะแนนความสามารถเฉลี่ยของนักเรียนแต่ละระดับชั้นเป็นรายคู่ ด้วยวิธีการของ Scheffe ปรากฏผลดังตารางที่ 4.20

ตารางที่ 4.20ผลการทดสอบความแตกต่างระหว่างคะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1-3 ในกลุ่มรวม

ระดับชั้น	ม.1	ม.2	ม.3
ม.1	-		
ม.2	.973**	-	
ม.3	1.819**	.846**	-

** $p < .01$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ผลการทดสอบความแตกต่างเป็นรายคู่ ระหว่างค่าเฉลี่ยของคะแนนความสามารถของนักเรียนระดับชั้นต่างๆ พบว่า คะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้น ม.1 กับ ม.2, คะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้น ม.2 กับ ม.3 และคะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้น ม.1 กับ ม.3 ทั้ง 3 คู่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ผู้วิจัยจึงทำการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ โดยคำนวณหา "อัตราพัฒนาการ" (Growth Rate) จากสูตรของ อวยพร เรืองตระกูล (2544) ซึ่งได้ประยุกต์แนวคิดในการวัดคะแนนพัฒนาการตามทฤษฎีการทดสอบแบบดั้งเดิมของ ศิริชัย กาญจนวาสี (2532) และแนวคิดของวินิจ เทือกทอง (2537) มาใช้ในการวัดคะแนนพัฒนาการตามทฤษฎีตอบสนองข้อสอบ โดยทำการวัดคะแนนพัฒนาการจากความสามารถที่แท้จริงสัมพัทธ์เมื่อเทียบกับศักยภาพการพัฒนา ($S\theta$) ผลการวิเคราะห์สรุปได้ดังนี้

ตารางที่ 4.21 คะแนนความสามารถเฉลี่ยและอัตราพัฒนาการ (Growth Rate) ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1-3 โดยภาพรวม

กลุ่มความสามารถ	คะแนนความสามารถเฉลี่ย			อัตราพัฒนาการ (Growth Rate)		
	ม.1	ม.2	ม.3	ม.1 เมื่อขึ้น ม.2	ม.2 เมื่อขึ้น ม.3	ม.1 เมื่อขึ้น ม.3
กลุ่มรวม	-1.837	-0.864	-0.018	0.167	0.174	0.312

จากตารางที่ 4.21 เมื่อพิจารณาอัตราพัฒนาการ (Growth Rate) สรุปได้ว่า นักเรียนระดับช่วงชั้นที่ 3 (ม.1-ม.3) มีพัฒนาการความสามารถทางคณิตศาสตร์สูงขึ้นตามระดับชั้นที่สูงขึ้น โดยความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.2 พัฒนาการความสามารถสูงขึ้น 0.167 คะแนน นักเรียนระดับชั้น ม.2 เมื่อขึ้น ม.3 พัฒนาการความสามารถสูงขึ้น 0.174 คะแนน และนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.3 พัฒนาการความสามารถสูงขึ้น 0.312 คะแนน ซึ่งจากตารางที่ 4.1 และ 4.4 ก่อนทำการปรับเทียบคะแนน ค่าคะแนนเฉลี่ย และค่าคะแนนความสามารถเฉลี่ย ทั้ง 3 ระดับชั้นไม่ได้ให้สารสนเทศในการเปรียบเทียบคะแนนกันได้ เพราะแบบสอบต่างชุด ต่างระดับชั้น เมื่อหลังการปรับเทียบแล้ว สามารถแสดงแนวโน้มบางประการได้ เพื่อให้เห็นการพัฒนาการความสามารถทางคณิตศาสตร์ที่ชัดเจน ดังแสดงใน แผนภาพที่ 4.1

แผนภาพที่ 4.1 พัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนโดยภาพรวม

4.2 การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนจำแนกตามกลุ่มความสามารถ

ในการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนทั้ง 3 ระดับชั้นนั้น เพื่อให้ได้ข้อมูลที่ชัดเจนเกี่ยวกับพัฒนาการความสามารถทางคณิตศาสตร์ในกลุ่มความสามารถที่แตกต่างกัน ผู้วิจัยได้แบ่งนักเรียนแต่ละระดับชั้นตามความสามารถ (θ) โดยใช้คะแนนที่ได้จากการทำแบบสอบภายหลังการปรับเทียบคะแนนให้อยู่ในสเกลเดียวกันตามระดับชั้นของตน เป็น 3 กลุ่ม ตามระดับความสามารถ ผลจากการจัดแบ่งกลุ่มนักเรียนดังกล่าว ปรากฏค่าสถิติพื้นฐานของนักเรียนจำแนกตามระดับชั้นและตามกลุ่มความสามารถ ปรากฏผลดังตารางที่ 4.22

ตารางที่ 4.22 ค่าสถิติพื้นฐานของนักเรียนจำแนกตามระดับชั้นและกลุ่มความสามารถ

ค่าสถิติ	กลุ่มความสามารถ					
	ต่ำ		ปานกลาง		สูง	
	θ	ξ	θ	ξ	θ	ξ
มัธยมศึกษาปีที่ 1						
จำนวนผู้สอบ	123	123	102	102	102	102
คะแนนสูงสุด	-2.601	6.480	-1.616	6.490	0.685	12.370
คะแนนต่ำสุด	-3.490	6.480	-2.059	6.480	-1.046	6.540
ค่าเฉลี่ย	-2.956	6.480	-1.818	6.485	-0.416	7.246
ส่วนเบี่ยงเบนมาตรฐาน	0.232	0	0.131	0.005	0.406	1.159
มัธยฐาน	-2.935	6.480	-1.811	6.490	-0.537	6.730
ฐานนิยม	-2.811	6.480	-1.928	6.490	0.685	6.740
ความเบ้	-0.299	-	-0.172	-0.040	0.868	2.535
ความโด่ง	-0.819	-	-1.142	-2.039	0.127	6.975
มัธยมศึกษาปีที่ 2						
จำนวนผู้สอบ	124	124	87	87	188	188
คะแนนสูงสุด	-1.710	6.490	-0.618	6.680	1.487	23.500
คะแนนต่ำสุด	-2.030	6.480	-1.108	6.530	0.272	9.070
ค่าเฉลี่ย	-1.858	6.483	-0.857	6.586	0.923	16.350
ส่วนเบี่ยงเบนมาตรฐาน	0.080	0.005	0.139	0.044	0.395	5.050
มัธยฐาน	-1.854	6.480	-0.896	6.570	1.070	18.180
ฐานนิยม	-1.800	6.480	-0.933	6.560	1.294	10.090
ความเบ้	-0.094	0.690	0.208	0.644	-0.240	-0.193
ความโด่ง	-0.972	-1.549	-1.220	-0.925	-1.556	-1.612
มัธยมศึกษาปีที่ 3						
จำนวนผู้สอบ	120	120	71	71	77	77
คะแนนสูงสุด	-0.784	6.600	0.205	8.680	2.581	29.290
คะแนนต่ำสุด	-1.333	6.500	-0.230	7.120	0.898	15.370
ค่าเฉลี่ย	-0.934	6.562	-0.002	7.810	1.730	24.709
ส่วนเบี่ยงเบนมาตรฐาน	0.106	0.022	0.125	0.454	0.462	3.656
มัธยฐาน	-0.922	6.560	0.000	7.750	1.432	22.900
ฐานนิยม	-1.105	6.550	-0.111	7.610	1.333	21.720
ความเบ้	-0.686	-0.090	-0.057	0.322	0.208	-0.193
ความโด่ง	0.365	-0.918	-1.085	-1.066	-1.525	-1.216

ผลการวิเคราะห์ตามตารางที่ 4.22 ค่าสถิติพื้นฐานของผลการสอบของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 ภายหลังการปรับเทียบคะแนนตามกลุ่มความสามารถ ปรากฏผลดังนี้

กลุ่มนักเรียนที่มีความสามารถต่ำ พบว่า ในระดับชั้น ม.1 ได้คะแนนความสามารถสูงสุดและต่ำสุดเท่ากับ -2.603 และ -3.490 ตามลำดับ โดยมีคะแนนความสามารถเฉลี่ย -2.956 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.232 การแจกแจงของคะแนนมีลักษณะเบ้ขวา แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อยและมีความโด่งแบนกว่าโค้งปกติ ส่วนในระดับชั้น ม. 2 ได้คะแนนความสามารถสูงสุดและต่ำสุด เท่ากับ -1.710 และ -2.030 ตามลำดับ โดยมีคะแนนความสามารถเฉลี่ย -1.858 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.080 การแจกแจงของคะแนนมีลักษณะเกือบเท่าโค้งปกติ และมีความโด่งแบนกว่าโค้งปกติ และในระดับชั้น ม.3 ได้คะแนนความสามารถสูงสุดและต่ำสุด เท่ากับ -0.784 และ -1.333 ตามลำดับ โดยมีคะแนนความสามารถเฉลี่ย -0.934 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.106 การแจกแจงของคะแนนมีลักษณะเบ้ขวา แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อยและมีความโด่งสูงกว่าโค้งปกติเล็กน้อย

เมื่อพิจารณาโดยรวมในกลุ่มนักเรียนที่มีความสามารถต่ำ พบว่าคะแนนความสามารถและคะแนนจริงโดยเฉลี่ยของนักเรียนทั้ง 3 ระดับชั้นเพิ่มขึ้นตามระดับชั้นที่สูงขึ้น นั่นคือ นักเรียนในระดับชั้น ม.3 มีคะแนนความสามารถและคะแนนจริงโดยเฉลี่ยสูงกว่านักเรียนในระดับชั้น ม.2 และนักเรียนในระดับชั้น ม.2 มีคะแนนความสามารถและคะแนนจริงโดยเฉลี่ยสูงกว่านักเรียนในระดับชั้น ม.1 ตามลำดับ อีกทั้งเป็นที่น่าสังเกตว่าส่วนเบี่ยงเบนมาตรฐานของคะแนนจริงทั้ง 3 ระดับชั้นมีค่าน้อยมาก ทั้งนี้เป็นผลเนื่องจากเป็นคะแนนที่อยู่ส่วนปลายซึ่งค่าที่ได้จะแตกต่างกันน้อยมาก โดยเฉพาะในระดับชั้น ม.1 คะแนนจริงที่ได้มีเพียงค่าเดียว คือ 6.48 คะแนน ดังนั้นส่วนเบี่ยงเบนมาตรฐานจึงมีค่า เท่ากับ 0

กลุ่มนักเรียนที่มีความสามารถปานกลาง พบว่า ในระดับชั้น ม.1 ได้คะแนนความสามารถสูงสุดและต่ำสุดเท่ากับ -1.616 และ -2.059 ตามลำดับ โดยมีคะแนนความสามารถเฉลี่ย -1.818 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.131 การแจกแจงของคะแนนมีลักษณะเบ้ขวาเล็กน้อย แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อยและมีความโด่งแบนกว่าโค้งปกติ ส่วนในระดับชั้น ม. 2 ได้คะแนนความสามารถสูงสุดและต่ำสุด เท่ากับ -0.618 และ -1.108 ตามลำดับ โดยมีคะแนนความสามารถเฉลี่ย -0.857 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.139 การแจกแจงของคะแนนมีลักษณะเบ้ซ้ายเล็กน้อย แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนมาก และมีความโด่งแบนกว่าโค้งปกติ และในระดับชั้น ม.3 ได้คะแนนความสามารถสูงสุด

และต่ำสุด เท่ากับ 0.205 และ -0.230 ตามลำดับ โดยมีคะแนนความสามารถเฉลี่ย -0.002 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.125 การแจกแจงของคะแนนมีลักษณะเกือบเท่าโค้งปกติ และมีความโด่งแบนกว่าโค้งปกติ

เมื่อพิจารณาโดยรวมในกลุ่มนักเรียนที่มีความสามารถปานกลาง พบว่าคะแนนความสามารถและคะแนนจริงโดยเฉลี่ยของนักเรียนทั้ง 3 ระดับชั้นเพิ่มขึ้นตามระดับชั้นที่สูงขึ้น นั่นคือ นักเรียนในระดับชั้น ม.3 มีคะแนนความสามารถและคะแนนจริงโดยเฉลี่ยสูงกว่านักเรียนในระดับชั้น ม.2 และนักเรียนในระดับชั้น ม.2 มีคะแนนความสามารถและคะแนนจริงโดยเฉลี่ยสูงกว่านักเรียนในระดับชั้น ม.1 ตามลำดับ อีกทั้งเป็นที่น่าสังเกตว่าส่วนเบี่ยงเบนมาตรฐานของคะแนนจริงทั้ง 3 ระดับชั้นมีค่าเพิ่มขึ้นกว่านักเรียนที่มีระดับความสามารถต่ำ ทั้งนี้เป็นผลเนื่องจากเป็นคะแนนที่อยู่ช่วงตรงกลางของข้อมูล ซึ่งค่าที่ได้เริ่มมีการกระจายข้อมูลเพิ่มมากขึ้นโดยเฉพาะในระดับชั้น ม.3 มีส่วนเบี่ยงเบนมาตรฐานของคะแนนจริงสูงกว่าในระดับชั้นอื่นอย่างชัดเจน

กลุ่มนักเรียนที่มีความสามารถสูง พบว่า ในระดับชั้น ม.1 ได้คะแนนความสามารถสูงสุดและต่ำสุดเท่ากับ 0.685 และ -1.046 ตามลำดับ โดยมีคะแนนความสามารถเฉลี่ย -0.416 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.406 การแจกแจงของคะแนนมีลักษณะเบ้ขวา แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อยและมีความโด่งเกือบเท่าโค้งปกติ ส่วนในระดับชั้น ม. 2 ได้คะแนนความสามารถสูงสุดและต่ำสุด เท่ากับ 1.487 และ 0.272 ตามลำดับ โดยมีคะแนนความสามารถเฉลี่ย -0.923 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.139 การแจกแจงของคะแนนมีลักษณะเบ้ซ้ายเล็กน้อย แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนมาก และมีความโด่งแบนกว่าโค้งปกติ และในระดับชั้น ม.3 ได้คะแนนความสามารถสูงสุดและต่ำสุด เท่ากับ 2.581 และ 0.898 ตามลำดับ โดยมีคะแนนความสามารถเฉลี่ย 1.730 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.463 การแจกแจงของคะแนนมีลักษณะเบ้ขวาเล็กน้อย แสดงว่า คะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อย และมีความโด่งแบนกว่าโค้งปกติ

เมื่อพิจารณาโดยรวมในกลุ่มนักเรียนที่มีความสามารถสูง พบว่าคะแนนความสามารถและคะแนนจริงโดยเฉลี่ยของนักเรียนทั้ง 3 ระดับชั้นเพิ่มขึ้นตามระดับชั้นที่สูงขึ้น นั่นคือ นักเรียนในระดับชั้น ม.3 มีคะแนนความสามารถและคะแนนจริงโดยเฉลี่ยสูงกว่านักเรียนในระดับชั้น ม.2 และนักเรียนในระดับชั้น ม.2 มีคะแนนความสามารถและคะแนนจริงโดยเฉลี่ยสูงกว่านักเรียนในระดับชั้น ม.1 ตามลำดับ อีกทั้งเป็นที่น่าสังเกตว่าส่วนเบี่ยงเบนมาตรฐานของคะแนนจริงทั้ง 3 ระดับชั้นมีค่าเพิ่มขึ้นกว่านักเรียนที่มีระดับความสามารถต่ำและปานกลาง แม้ว่า

จะเป็นคะแนนที่อยู่ส่วนปลายข้อมูล โดยเฉพาะนักเรียนในระดับชั้น ม.2 มีส่วนเบี่ยงเบนมาตรฐานของคะแนนจริงสูงกว่าในระดับชั้นอื่นอย่างชัดเจน

การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนในแต่ละระดับความสามารถ ผู้วิจัยศึกษาโดยเปรียบเทียบคะแนนเฉลี่ยของนักเรียน 3 ระดับชั้น จำแนกตามกลุ่มความสามารถ โดยวิธีการในทำนองเดียวกับการศึกษาพัฒนาการความสามารถของนักเรียน 3 ระดับชั้นโดยภาพรวม ผลการวิเคราะห์ความแปรปรวนของคะแนนของนักเรียน 3 ระดับชั้น จำแนกตามกลุ่มย่อยแต่ละกลุ่ม ปรากฏดังตารางที่ 4.23

ตารางที่ 4.23 ผลการวิเคราะห์ความแปรปรวนทางเดียวของคะแนนความสามารถของนักเรียนระดับชั้น ม.1 ม.2 และ ม.3 จำแนกตามกลุ่มความสามารถ

แหล่งความแปรปรวน	SS	df	MS	F	P
<u>กลุ่มต่ำ</u>					
ระหว่างกลุ่ม	249.172	2	124.586	5229.300	.000
ภายในกลุ่ม	8.672	364	.024		
รวม	257.844	366			
<u>กลุ่มปานกลาง</u>					
ระหว่างกลุ่ม	140.848	2	70.424	4037.300	.000
ภายในกลุ่ม	4.483	257	.017		
รวม	145.331	259			
<u>กลุ่มสูง</u>					
ระหว่างกลุ่ม	217.858	2	108.929	638.212	.000
ภายในกลุ่ม	62.127	364	.171		
รวม	279.986	366			

จากตารางที่ 4.23 ผลการวิเคราะห์ความแปรปรวนทางเดียวของคะแนนความสามารถของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 ตามกลุ่มความสามารถปรากฏผลดังนี้

ในกลุ่มนักเรียนที่มีความสามารถต่ำ พบว่า คะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ($F = 5229.300$, $p = .000$) ความแปรปรวนของคะแนนความสามารถของนักเรียน 3 ระดับชั้น มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (Levene Statistic = 83.582, $p = .000$) จึงทำการทดสอบความแตกต่างของคะแนนความสามารถเฉลี่ยของนักเรียนแต่ละระดับชั้นเป็นรายคู่ด้วยวิธีการของ Dunnett T3

ในกลุ่มนักเรียนที่มีความสามารถปานกลาง พบว่า คะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ($F = 4037.3$, $p = .000$) ความแปรปรวนของคะแนนความสามารถของนักเรียน 3 ระดับชั้น มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .01 (Levene Statistic = 1.049, $p = .352$) จึงทำการทดสอบความแตกต่างของคะแนนความสามารถเฉลี่ยของนักเรียนแต่ละระดับชั้นเป็นรายคู่ด้วยวิธีการของ Scheffe

ในกลุ่มนักเรียนที่มีความสามารถสูง พบว่า คะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ($F = 638.212$, $p = .000$) ความแปรปรวนของคะแนนความสามารถของนักเรียน 3 ระดับชั้น มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (Levene Statistic = 9.469, $p = .000$) จึงทำการทดสอบความแตกต่างของคะแนนความสามารถเฉลี่ยของนักเรียนแต่ละระดับชั้นเป็นรายคู่ ด้วยวิธีการของ Dunnett T3

การทดสอบความแตกต่างของคะแนนความสามารถเฉลี่ยของนักเรียนแต่ละระดับชั้นเป็นรายคู่ทั้ง 3 กลุ่มความสามารถ ปรากฏผลดังตารางที่ 4.24

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.24 ผลการทดสอบความแตกต่างระหว่างคะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1-3 จำแนกตามกลุ่มความสามารถ

ระดับชั้น	ม.1	ม.2	ม.3
กลุ่มต่ำ			
ม.1	-		
ม.2	1.098**	-	
ม.3	2.022**	.924**	-
กลุ่มปานกลาง			
ม.1	-		
ม.2	.961**	-	
ม.3	1.816**	.855**	-
กลุ่มสูง			
ม.1	-		
ม.2	1.339**.	-	
ม.3	2.146**	.807**	-

** $p < .01$

ผลการทดสอบความแตกต่างเป็นรายคู่ ระหว่างคะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 จำแนกตามกลุ่มความสามารถ จากตารางที่ 4.24 ปรากฏว่า **นักเรียนทั้ง 3 กลุ่มความสามารถ คือ กลุ่มต่ำ กลุ่มปานกลาง** และ **กลุ่มสูง** มีคะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้น ม.1 กับ ม.2, คะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้น ม.2 กับ ม.3 และคะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้น ม.1 กับ ม.3 ทั้ง 3 คู่ แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ผู้วิจัยจึงทำการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ตามกลุ่มความสามารถ โดยคำนวณหาอัตราพัฒนาการ (Growth Rate) จากสูตรของ อวยพร เรืองตระกูล (2544) ซึ่งได้ประยุกต์แนวคิดในการวัดคะแนนพัฒนาการตามทฤษฎีการทดสอบแบบดั้งเดิมของ ศิริชัย กาญจนวาสี (2532) และแนวคิดของวินิจ เทือกทอง (2537) มาใช้ในการวัดคะแนนพัฒนาการตามทฤษฎีตอบสนองข้อสอบ โดยทำการวัดคะแนนพัฒนาการจากความสามารถที่แท้จริงสัมพัทธ์เมื่อเทียบกับศักยภาพการพัฒนา ($s\theta$) ผลการวิเคราะห์สรุปได้ดังนี้

ตารางที่ 4.25 คะแนนความสามารถเฉลี่ยและอัตราพัฒนาการ (Growth Rate) ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 จำแนกตามกลุ่มความสามารถ

กลุ่มความสามารถ	คะแนนความสามารถเฉลี่ย			อัตราพัฒนาการ (Growth Rate)		
	ม.1	ม.2	ม.3	ม.1 เมื่อขึ้น ม.2	ม.2 เมื่อขึ้น ม.3	ม.1 เมื่อขึ้น ม.3
กลุ่มต่ำ	-2.956	-1.818	-0.416	0.164	0.241	0.365
กลุ่มปานกลาง	-1.858	-0.857	0.923	0.171	0.366	0.475
กลุ่มสูง	-0.934	-0.002	1.730	0.189	0.433	0.540

จากตารางที่ 4.25 เมื่อพิจารณาอัตราพัฒนาการ (Growth Rate) ตามกลุ่มความสามารถของนักเรียน สรุปได้ว่า นักเรียนระดับช่วงชั้นที่ 3 (ม.1-ม.3) มีพัฒนาการความสามารถทางคณิตศาสตร์สูงขึ้นตามระดับชั้นที่สูงขึ้นทุกกลุ่มความสามารถ ซึ่งปรากฏผลดังนี้

ในกลุ่มนักเรียนที่มีระดับความสามารถต่ำ พบว่าความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.2 พัฒนาการความสามารถสูงขึ้น 0.164 คะแนน นักเรียนระดับชั้น ม.2 เมื่อขึ้น ม.3 พัฒนาการความสามารถสูงขึ้น 0.241 คะแนน และนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.3 พัฒนาการความสามารถสูงขึ้น 0.365 คะแนน

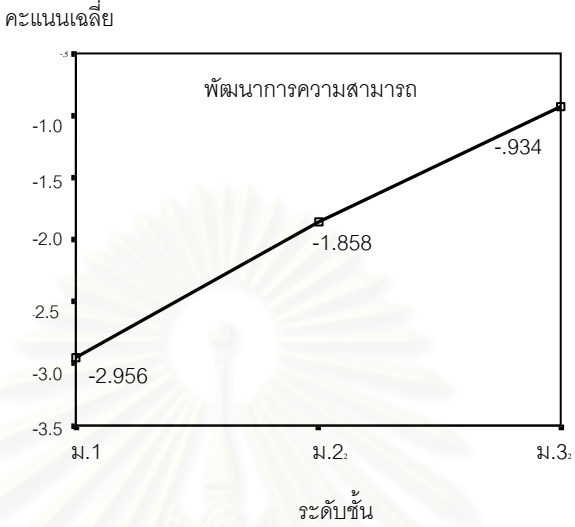
ในกลุ่มนักเรียนที่มีระดับความสามารถปานกลาง พบว่า พบว่าความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.2 พัฒนาการความสามารถสูงขึ้น 0.171 คะแนน นักเรียนระดับชั้น ม.2 เมื่อขึ้น ม.3 พัฒนาการความสามารถสูงขึ้น 0.366 คะแนน และนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.3 พัฒนาการความสามารถสูงขึ้น 0.475 คะแนน

ในกลุ่มนักเรียนที่มีระดับความสามารถสูง พบว่า พบว่าความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.2 พัฒนาการความสามารถสูงขึ้น 0.189 คะแนน นักเรียนระดับชั้น ม.2 เมื่อขึ้น ม.3 พัฒนาการความสามารถสูงขึ้น 0.433 คะแนน และนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.3 พัฒนาการความสามารถสูงขึ้น 0.540 คะแนน

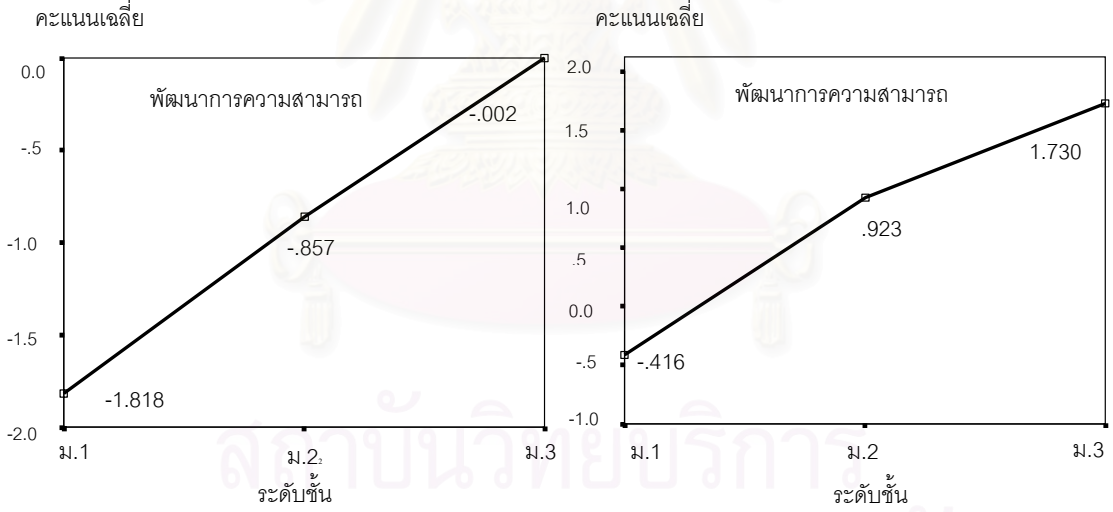
โดยสรุป จากการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนทั้ง 3 กลุ่มความสามารถมีพัฒนาการความสามารถทางคณิตศาสตร์เพิ่มขึ้นตามระดับชั้นที่สูงขึ้น โดยเฉพาะพัฒนาการความสามารถของนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.3 มีพัฒนาการสูงมากทุกกลุ่มความสามารถ เมื่อพิจารณาอัตราพัฒนาการเปรียบเทียบระหว่างกลุ่มต่ำ กลุ่มปานกลาง และกลุ่มสูง พบว่า นักเรียนที่อยู่ในกลุ่มสูงมีอัตราพัฒนาการความสามารถทางคณิตศาสตร์สูงกว่านักเรียนที่อยู่ในกลุ่มต่ำ และกลุ่มปานกลาง โดยนักเรียนที่อยู่ในกลุ่มต่ำมีอัตราพัฒนาการความสามารถทางคณิตศาสตร์ต่ำกว่านักเรียนในกลุ่มความสามารถอื่น

เพื่อให้เห็นพัฒนาการความสามารถทางคณิตศาสตร์ที่ชัดเจนตามกลุ่มความสามารถ คือ กลุ่มความต่ำ กลุ่มปานกลาง และกลุ่มสูง ดังแสดงในแผนภาพที่ 4.2

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



แผนภาพที่ 4.2 (ก) กลุ่มความสามารถต่ำ

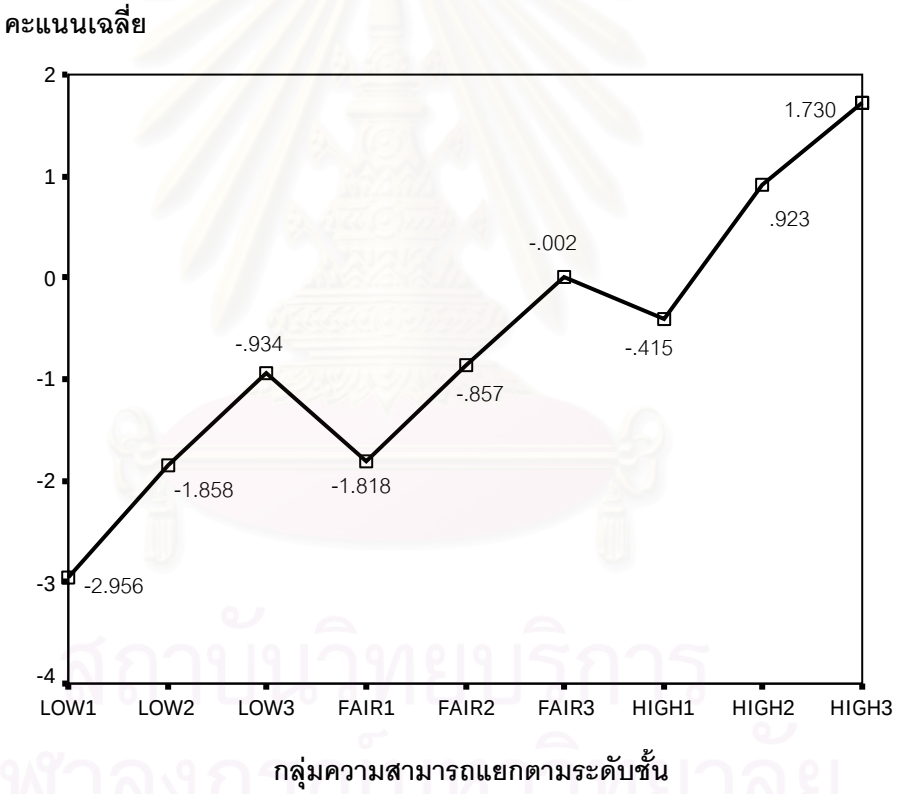


แผนภาพ 4.2 (ข) กลุ่มความสามารถปานกลาง แผนภาพ 4.2(ค) กลุ่มความสามารถสูง

แผนภาพที่ 4.2 แสดงพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนระดับมัธยมศึกษาปีที่ 1-3 แยกตามกลุ่มความสามารถของนักเรียน

จากแผนภาพที่ 4.2 แสดงให้เห็นว่านักเรียนทั้ง 3 กลุ่มความสามารถมีการเปลี่ยนแปลงค่าเฉลี่ยของคะแนนพัฒนาการทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 มีแนวโน้มเพิ่มสูงขึ้น ตามระดับชั้นที่เพิ่มสูงขึ้น โดยเฉพาะในกลุ่มสูง

นอกจากนี้เพื่อให้เห็นแนวโน้มพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนในแต่ละระดับชั้นในแต่ละกลุ่มความสามารถ ดังนี้ กลุ่มต่ำของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 (L0W1, L0W2, L0W3) กลุ่มปานกลางของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 (FAIR1, FAIR 2, FAIR3) และกลุ่มสูงของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 (HIGH1, HIGH2, HIGH3) รวมทั้งหมด 9 กลุ่มความสามารถแยกตามะดับชั้น สามารถแสดงแนวโน้มบางประการได้ดังแผนภาพที่ 4.3



แผนภาพที่ 4.3 แสดงพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนตามกลุ่มความสามารถแยกตามระดับชั้น

จากแผนภาพที่ 4.3 เมื่อศึกษาแนวโน้มพัฒนาการข้ามระดับชั้นและข้ามกลุ่มความสามารถทั้ง 9 กลุ่มความสามารถ สามารถเห็นแนวโน้มของพัฒนาการความสามารถทางคณิตศาสตร์ได้ชัดเจนยิ่งขึ้น ซึ่งสามารถลำดับความสามารถเฉลี่ยของนักเรียนจากน้อยที่สุดไปหามากที่สุด ได้ดังนี้ กลุ่มต่ำของระดับชั้น ม.1, กลุ่มต่ำของระดับชั้น ม.2, กลุ่มปานกลางของระดับชั้น ม.1, กลุ่มปานกลางของระดับชั้น ม.2, กลุ่มต่ำของระดับชั้น ม.3, กลุ่มสูงของระดับชั้น ม.1, กลุ่มปานกลางของระดับชั้น ม.3, กลุ่มสูงของระดับชั้น ม.2 และ กลุ่มสูงของระดับชั้น ม.3 ซึ่งเป็นกลุ่มที่มีคะแนนความสามารถเฉลี่ยสูงสุด

นอกจากนี้เมื่อพิจารณาแผนภาพที่ 4.3 จะเห็นได้ว่ากลุ่มต่ำของนักเรียนระดับชั้น ม.2 มีคะแนนความสามารถเฉลี่ยใกล้เคียงกับกลุ่มความสามารถปานกลางของนักเรียนระดับชั้น ม.1 ส่วนกลุ่มต่ำของนักเรียนระดับชั้น ม.3 มีคะแนนความสามารถเฉลี่ยใกล้เคียงกับกลุ่มปานกลางของนักเรียนระดับชั้น ม.2 โดยกลุ่มปานกลางของนักเรียนระดับชั้น ม.2 มีคะแนนความสามารถเฉลี่ยสูงกว่านักเรียนระดับชั้น ม.1 และ ม.3 ที่อยู่ในกลุ่มต่ำและกลุ่มปานกลาง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้มีจุดมุ่งหมาย เพื่อศึกษาประสิทธิภาพของการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง ที่ใช้แบบสอบร่วมภายใน และศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ในแต่ละระดับชั้นโดยภาพรวมและจำแนกตามกลุ่มความสามารถ คือ กลุ่มต่ำ กลุ่มปานกลาง และ กลุ่มสูง โดยเปรียบเทียบคะแนนความสามารถทางคณิตศาสตร์ของนักเรียนในแต่ละระดับชั้น ภายหลังการปรับเทียบคะแนนนักเรียนแต่ละระดับชั้นให้อยู่ในสเกลเดียวกัน อันเป็นการศึกษาพัฒนาการภาคตัดขวาง (Cross-Sectional Study)

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ของโรงเรียนในเขตพื้นที่การศึกษาสกลนคร เขต 1 ภาคเรียนที่ 2 ปีการศึกษา 2547 จำนวน 9 โรงเรียน โดยสุ่มกลุ่มตัวอย่างด้วยวิธีการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) มีขนาดของโรงเรียนเป็นตัวแปรสำหรับแบ่งชั้น คือ ขนาดใหญ่พิเศษ ขนาดใหญ่ ขนาดกลาง และขนาดเล็ก ผลการสุ่มกลุ่มปรับเทียบคะแนนได้เป็นนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 จำนวน 543, 534 และ 504 คน ตามลำดับ และกลุ่มสอบทานผล เป็นนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 จำนวน 113, 120 และ 127 คน ตามลำดับ

เครื่องมือที่ใช้ในการวิจัยครั้งนี้เป็นแบบวัดความสามารถทางคณิตศาสตร์ เรื่อง “จำนวนและการดำเนินการ” สำหรับระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 รวม 3 ฉบับ ฉบับละ 30 ข้อ โดยเป็นแบบสอบเลือกตอบ 4 ตัวเลือก โดยมีข้อสอบร่วมระหว่างระดับชั้น ม.1 กับ ม.2 จำนวน 6 ข้อ ซึ่งเป็นเนื้อหาของระดับชั้น ม.1 ข้อสอบร่วมระหว่างระดับชั้น ม.2 กับ ม.3 จำนวน 6 ข้อ ซึ่งเป็นเนื้อหาของระดับชั้น ม.2

การเก็บรวบรวมข้อมูล โดยนำแบบสอบไปสอบกับกลุ่มตัวอย่างสองกลุ่ม คือ กลุ่มปรับเทียบคะแนนโดยทำแบบสอบปรับเทียบคะแนนในระดับชั้นของตนที่มีแบบสอบร่วมภายในส่วนกลุ่มสอบทานผล จะต้องทำแบบสอบปรับเทียบคะแนนของทั้ง 3 ระดับชั้น

ในการวิเคราะห์ข้อมูลใช้วิธีการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง โดยได้แบ่งการวิเคราะห์ออกเป็น 4 ตอน ตอนแรกเป็นการวิเคราะห์ค่าสถิติพื้นฐานของแบบสอบ ตอนที่สองเป็นการดำเนินการสร้าง

ตารางค่าความสามารถและคะแนนของกลุ่มปรับเทียบคะแนน โดยการนำคะแนนที่ตอบแบบสอบในระดับชั้นของตนมาวิเคราะห์เพื่อประมาณค่าความสามารถของผู้สอบแต่ละคนของกลุ่มตัวอย่างปรับเทียบคะแนนด้วยโปรแกรม BILOG 3.04 นำค่าที่ได้มาวิเคราะห์ปรับเทียบคะแนนโดยใช้การปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ โดยอาศัยค่าคงที่ที่ได้จากการวิเคราะห์โดยใช้โปรแกรม EQUATE 2.0 ปรับค่าความสามารถของผู้สอบทุกระดับชั้น ให้อยู่ในมาตรวัดหรือสเกลเดียวกัน และทำการปรับแก้ค่าความสามารถที่มีระดับความสามารถแตกต่างกันมากๆ โดยการปรับเทียบเชิงเส้นตรง ด้วยวิธี “Simple Least Squared Method” ซึ่งใช้โปรแกรม SPSS for Window Version 11.0 ตอนที่สามเป็นการศึกษาประสิทธิภาพของการปรับเทียบคะแนนโดยดูผลการวิเคราะห์ความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน และผลการวิเคราะห์ความเพียงพอของความคลาดเคลื่อนของการปรับเทียบคะแนน ที่เปรียบเทียบกับเกณฑ์ของ Petersen และคณะ (1982) และตอนสุดท้ายเป็นการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนทั้ง 3 ระดับชั้น โดยภาพรวมและจำแนกตามกลุ่มความสามารถ คือ กลุ่มต่ำ กลุ่มปานกลาง และกลุ่มสูง ภายหลังการปรับเทียบคะแนนของนักเรียนระดับชั้น ม.1 และคะแนนของนักเรียนระดับชั้น ม.2 ไปยังสเกลเดียวกันกับคะแนนแบบสอบระดับชั้น ม.3 แล้วเปรียบเทียบความแตกต่างค่าเฉลี่ยของคะแนนความสามารถของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 ซึ่งปรับเทียบให้อยู่ในสเกลเดียวกันแล้ว ด้วยการวิเคราะห์ความแปรปรวนทางเดียว (One-Way ANOVA) และคำนวณหาอัตราพัฒนาการ (Growth Rate) จากคะแนนความสามารถเฉลี่ยเพื่อศึกษาพัฒนาความสามารถทางคณิตศาสตร์ของนักเรียน ดังเสนอสรุปผลการวิจัย ดังนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สรุปผลการวิจัย

1. ผลการวิเคราะห์แบบสอบ

ผลการวิเคราะห์ค่าสถิติพื้นฐานของแบบสอบปรับเทียบคะแนน จากการวิเคราะห์แบบสอบระดับชั้น ม.1 ได้คะแนนระหว่าง 3-29 คะแนน จากคะแนนเต็ม 30 คะแนน โดยมีคะแนนเฉลี่ย 13.070 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 5.454 และค่าความเที่ยงของแบบสอบเท่ากับ 0.803 ส่วนแบบสอบระดับชั้น ม.2 ได้คะแนนระหว่าง 2-30 คะแนน โดยมีคะแนนเฉลี่ย 12.636 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 7.739 และมีค่าความเที่ยงของแบบสอบเท่ากับ 0.912 และแบบสอบในระดับชั้น ม. 3 ได้คะแนนระหว่าง 3-30 คะแนน โดยมีคะแนนเฉลี่ย 13.654 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 6.295 ค่าความเที่ยงของแบบสอบ เท่ากับ 0.857

เมื่อพิจารณาแบบสอบโดยรวม ค่าเฉลี่ย ค่ามัธยฐาน และฐานนิยมของคะแนนจากแบบสอบทั้ง 3 ฉบับ ต่ำกว่าครึ่งหนึ่งของคะแนนเต็ม แสดงว่าระดับความยากของแบบสอบทั้งสามฉบับค่อนข้างมาทางยาก จากค่าความเบ้และความโด่งของแบบสอบทั้งสามฉบับ จะเห็นว่าคะแนนส่วนใหญ่ค่อนไปทางคะแนนน้อย และคะแนนจากแบบสอบทั้ง 3 ฉบับ มีการแจกแจงใกล้เคียงโค้งปกติ

2. ผลการสร้างการดำเนินการปรับเทียบคะแนนแบบสอบ

จากการใช้วิธีการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง ผู้วิจัยสร้างตารางปรับเทียบคะแนนความสามารถ คะแนนดิบ และคะแนนจริงของนักเรียนกลุ่มปรับเทียบคะแนนระดับชั้น ม.1, ม.2 และ ม.3 ดังแสดงในตารางที่ 4.7 และ 4.8

3. ผลการศึกษาประสิทธิภาพของการปรับเทียบคะแนน

3.1 ผลการวิเคราะห์ความคลาดเคลื่อนมาตรฐานของการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง จากกลุ่มสอบทานผล พบว่า รูปแบบการแปลงคะแนนทั้งหมดของทุกระดับชั้น มีค่าความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน เท่ากับ 0.068 นั่นคือ ความคลาดเคลื่อนมาตรฐานของการ

ปรับเทียบคะแนนอยู่ในระดับต่ำ และเมื่อพิจารณาในแต่ละรูปแบบของการแปลงคะแนน ทั้ง 6 รูปแบบ (X^*_{12}, X^*_{13}, X^*_{21}, X^*_{23}, X^*_{31}, X^*_{32}) มีค่าความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน อยู่ระหว่าง 0.154 ถึง 0.178 โดยรูปแบบการแปลงคะแนนในระดับชั้นที่แตกต่างกัน คือ การแปลงคะแนนในระดับชั้น ม.1 กับ ระดับชั้น ม.3 (X^*_{13}, X^*_{31}) มีค่าความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนสูงกว่าการแปลงคะแนนในรูปแบบอื่น แต่แตกต่างกันเพียงเล็กน้อย

3.2 ผลการวิเคราะห์ความเพียงพอของการปรับเทียบคะแนน ซึ่งพิจารณาจากค่าดัชนีความแตกต่างของการปรับเทียบคะแนน (C) พบว่า ดัชนีความแตกต่างของการปรับเทียบคะแนน (C) มีค่าเท่ากับ 0.215 เมื่อนำค่าที่ได้มาประเมินด้วยเกณฑ์ของ Petersen และคณะ (1982) พบว่า ผลของการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง มีความเพียงพอของการปรับเทียบคะแนน อยู่ในระดับที่น่าพอใจ และเมื่อพิจารณาในแต่ละรูปแบบของการแปลงคะแนน ทั้ง 6 รูปแบบ คือ X^*_{12}, X^*_{13}, X^*_{21}, X^*_{23}, X^*_{31}, X^*_{32} พบว่า ดัชนีความแตกต่างของการปรับเทียบคะแนน (C) มีค่าเท่ากับ 0.275, 0.352, 0.160, 0.263, 0.356, 0.313 ตามลำดับ เมื่อนำดัชนีเหล่านี้มาประเมินด้วยเกณฑ์ของ Petersen และคณะ (1982) พบว่ารูปแบบการแปลงคะแนนส่วนใหญ่มีความคลาดเคลื่อนรวมเมื่อเทียบกับเกณฑ์แล้วน้อยกว่ากำลังสองของร้อยละ 10 ของค่าส่วนเบี่ยงเบนมาตรฐานของแบบสอบที่เป็นคะแนนเกณฑ์ ซึ่งอยู่ในระดับน่าพอใจ แต่จะมีบางรูปแบบที่มีความคลาดเคลื่อนรวมเมื่อเทียบกับเกณฑ์แล้วมากกว่ากำลังสองของร้อยละ 10 ของค่าส่วนเบี่ยงเบนมาตรฐานของแบบสอบที่เป็นคะแนนเกณฑ์ คือ รูปแบบการแปลงคะแนนในระดับชั้น ม.1 กับ ระดับชั้น ม.3 (X^*_{13}, X^*_{31}) ซึ่งความเพียงพอของการปรับเทียบคะแนนอยู่ในระดับปานกลาง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

4. ผลการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์

4.1 ผลการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์โดยภาพรวม

เมื่อแปลงคะแนนของนักเรียนระดับชั้น ม.1 และ ม.2 ให้อยู่ในสเกลเดียวกับระดับชั้น ม.3 แล้ว ปรากฏว่า คะแนนความสามารถเฉลี่ยของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 มีค่าเท่ากับ -1.837, -0.864 และ -0.018 ตามลำดับ เมื่อวิเคราะห์ความแตกต่างของคะแนนดังกล่าวแล้ว ปรากฏว่า ค่าเฉลี่ยของคะแนนความสามารถทั้ง 3 ระดับชั้น แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เมื่อศึกษาพัฒนาการความสามารถทางคณิตศาสตร์จากอัตราพัฒนาการ (Growth Rate) พบว่า ความสามารถทางการเรียนของนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.2 ความสามารถทางคณิตศาสตร์สูงขึ้น 0.973 คะแนน ส่วนความสามารถทางการเรียนของนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.3 ความสามารถทางคณิตศาสตร์สูงขึ้น 0.312 คะแนน และความสามารถทางการเรียนของนักเรียนระดับชั้น ม.2 เมื่อขึ้น ม.3 ความสามารถทางคณิตศาสตร์สูงขึ้น 0.174 คะแนน ดังนั้นจึงสรุปได้ว่า ความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.1, ม.2 และ ม.3 มีพัฒนาการเพิ่มขึ้นตามระดับชั้นที่สูงขึ้น

4.2 ผลการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์จำแนกตามกลุ่มความสามารถ

เมื่อแบ่งนักเรียนแต่ละระดับชั้นเป็นกลุ่มย่อยตามระดับความสามารถ แล้วศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ พบว่า นักเรียนทั้ง 3 กลุ่มความสามารถ **คือ กลุ่มต่ำ กลุ่มปานกลาง** และ **กลุ่มสูง** มีค่าเฉลี่ยของคะแนนความสามารถทั้ง 3 ระดับชั้น แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เมื่อศึกษาพัฒนาการความสามารถทางคณิตศาสตร์จากการพิจารณาอัตราพัฒนาการ (Growth Rate) ตามกลุ่มความสามารถของนักเรียน สรุปได้ว่า นักเรียนระดับชั้น ม.1, ม.2 และ ม.3 มีพัฒนาการความสามารถทางคณิตศาสตร์เพิ่มขึ้นตามระดับชั้นที่สูงขึ้นทุกกลุ่มความสามารถ ซึ่งปรากฏผลดังนี้

ในกลุ่มต่ำ พบว่าความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.2 พัฒนาการความสามารถสูงขึ้น 0.164 คะแนน นักเรียนระดับชั้น ม.2 เมื่อขึ้น ม.3 พัฒนาการความสามารถสูงขึ้น 0.241 คะแนน และนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.3 พัฒนาการความสามารถสูงขึ้น 0.365 คะแนน

ในกลุ่มปานกลาง พบว่าความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.2 พัฒนาการความสามารถสูงขึ้น 0.171 คะแนน นักเรียนระดับชั้น ม.2 เมื่อขึ้น ม.3 พัฒนาการความสามารถสูงขึ้น 0.366 คะแนน และนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.3 พัฒนาการความสามารถสูงขึ้น 0.475 คะแนน

ในกลุ่มสูง พบว่าความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.2 พัฒนาการความสามารถสูงขึ้น 0.189 คะแนน นักเรียนระดับชั้น ม.2 เมื่อขึ้น ม.3 พัฒนาการความสามารถสูงขึ้น 0.433 คะแนน และนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.3 พัฒนาการความสามารถสูงขึ้น 0.540 คะแนน

โดยสรุป จากการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนทั้ง 3 กลุ่มความสามารถ พบว่านักเรียนมีพัฒนาการความสามารถทางคณิตศาสตร์เพิ่มขึ้นตามระดับชั้นที่สูงขึ้น โดยเฉพาะพัฒนาการความสามารถของนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.3 มีพัฒนาการสูงมากทุกกลุ่มความสามารถ และเมื่อพิจารณาอัตราพัฒนาการเปรียบเทียบระหว่างกลุ่มความสามารถคือ กลุ่มต่ำ กลุ่มปานกลาง และกลุ่มสูง พบว่า นักเรียนที่อยู่ในกลุ่มสูงมีอัตราพัฒนาการความสามารถทางคณิตศาสตร์สูงกว่านักเรียนที่อยู่ในกลุ่มต่ำ และกลุ่มปานกลาง โดยนักเรียนที่อยู่ในกลุ่มต่ำมีอัตราพัฒนาการความสามารถทางคณิตศาสตร์ต่ำกว่านักเรียนในกลุ่มความสามารถอื่น

อภิปรายผล

การอภิปรายผลในงานวิจัยนี้นำเสนอ 2 ประเด็นหลัก คือ การอภิปรายผลที่ได้จากการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบร่วมซึ่งเป็นผลที่ได้จากการวิเคราะห์ค่าสถิติเบื้องต้น และการอภิปรายตามสมมติฐานการวิจัย มีรายละเอียด ดังนี้

1. การอภิปรายผลที่ได้จากการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบร่วม

ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบร่วมระหว่างระดับชั้น ม.1 กับ ม.2 และ ม.2 กับ ม.3 จากตารางที่ 4.3 และ 4.4 มีประเด็นในการอภิปรายดังนี้

ค่าอำนาจจำแนก (a) ของข้อสอบร่วมจากการตอบของนักเรียนระดับชั้น ม.2 มากกว่า ม.1 อาจเนื่องมาจากข้อสอบมีความยากเกินไปสำหรับนักเรียนระดับชั้น ม.1 โดยพิจารณาจากมีค่าความยาก (b) ที่สูงกว่า ม.2 ซึ่งการที่ข้อสอบมีค่ายากเกินไปทำให้จำแนกกลุ่มต่ำออกจากกลุ่มสูงได้น้อย ส่งผลให้ค่าอำนาจจำแนกของ ม.1 มากกว่า ม.2 ในทำนองเดียวกัน ค่าอำนาจจำแนก (a) ของ ม.3 น้อยกว่า ม.2 อาจเนื่องมาจากข้อสอบง่ายเกินไปสำหรับ ม.3 โดยพิจารณาจากมีค่าความยาก (b) ที่ต่ำกว่า ม.2 ทำให้มีค่าอำนาจจำแนกที่ต่ำ ส่งผลให้ค่าอำนาจจำแนกของ ม.2 มากกว่า ม.3

ค่าการเดา (C) ของข้อสอบร่วมที่ได้จากการตอบของนักเรียนระดับชั้น ม.1 น้อยกว่า ม.2 อาจเนื่องมาจาก นักเรียน ม.2 มีความรู้พื้นฐานมากกว่า และเรียนเนื้อหานั้นมาก่อน ทำให้เกิดความรู้บางส่วนที่จะช่วยให้มีโอกาสในการตอบข้อสอบได้สูงกว่า ม.1 ส่วนค่าการเดาของข้อสอบร่วมที่ได้จากการตอบของนักเรียนระดับชั้น ม.2 น้อยกว่า ม.3 จะจำแนกได้ 2 กลุ่ม คือ ข้อสอบร่วม 3 ข้อแรกของ ม.3 มีค่าการเดาต่ำกว่า ม.2 ในขณะที่ 3 ข้อหลัง ม.3 มีค่าการเดาสูงกว่า ม.2 อาจเนื่องมาจากข้อสอบ 3 ข้อแรก เป็นเรื่อง อัตราส่วน สัดส่วนและร้อยละ ซึ่งนักเรียน ม.3 อาจจะเรียนเนื้อหา (Content) ในเรื่องนั้นมานานทำให้เกิดการลืมในเนื้อหา ส่วน ม.2 เพิ่งเรียนเนื้อหา ดังกล่าวอาจจะจำเนื้อหาได้บางส่วนที่อาจช่วยให้มีโอกาสการเดาสูงกว่า ม.3

โดยสรุปจะพบว่าสอดคล้องกับ ศิริชัย กาญจนวาสี (2544) ที่กล่าวไว้ว่าข้อสอบที่มีความยากเกินไปหรือง่ายเกินไป มีโอกาสให้ค่าอำนาจจำแนกต่ำ อาจไม่สามารถจำแนกกลุ่มต่ำออกจากกลุ่มสูงได้ ซึ่งข้อสอบมีค่าความยากง่ายปานกลางจะมีแนวโน้มทำให้จำแนกได้ดีกว่า นอกจากนี้ค่าการเดา (C) ของข้อสอบจะเกี่ยวข้องกับเนื้อหา (Content) ซึ่งถ้านักเรียนมีความรู้บางส่วนในเนื้อหานั้นอาจส่งผลให้มีค่าการเดาที่สูง หรือถ้าเกิดพฤติกรรมการลืมในเนื้อหาอาจจะส่งผลให้มีค่าการเดาที่ต่ำ

2. การอภิปรายตามสมมติฐานการวิจัย

2.1 การอภิปรายตามสมมติฐานการวิจัยข้อที่ 1 ประสิทธิภาพของการปรับเทียบคะแนน

จากการศึกษาประสิทธิภาพของการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง สำหรับนักเรียนที่มีความสามารถต่างกัน 3 ระดับ โดยพิจารณาจากความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนและความเพียงพอของความคลาดเคลื่อนของการปรับเทียบคะแนน พบว่า สอดคล้องกับ

สมมติฐานการวิจัยข้อที่ 1 นั่นคือ ความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนอยู่ในระดับต่ำ ส่วนผลการวิเคราะห์ความเพียงพอของความคลาดเคลื่อนของการปรับเทียบคะแนน โดยพิจารณาจากดัชนีความแตกต่าง ประเมินโดยใช้เกณฑ์ของ Petersen และคณะ (1982) พบว่าความคลาดเคลื่อนของการปรับเทียบคะแนนจากรูปแบบที่ศึกษาเมื่อเปรียบเทียบกับเกณฑ์แล้วน้อยกว่ากำลังสองของร้อยละ 10 ของค่าส่วนเบี่ยงเบนมาตรฐานของแบบสอบ X (คะแนนเกณฑ์) โดยอยู่ในระดับที่น่าพอใจ ซึ่งสอดคล้องกับผลการวิจัยของ Leung (2003) ที่ได้เสนอวิธีการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ร่วมกับวิธีการปรับเทียบเชิงเส้นตรง ซึ่งผลการตรวจสอบประสิทธิภาพของการปรับเทียบด้วยวิธีดังกล่าว โดยพิจารณาความสัมพันธ์ระหว่างการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบกับการปรับเทียบหลังจากมีการปรับแก้ค่าจากการปรับเทียบเชิงเส้นตรงมีค่าเข้าใกล้ 1 แสดงให้เห็นว่าคุณภาพของการปรับเทียบคะแนนอยู่ในระดับดี เนื่องจากเกิดความคลาดเคลื่อนของการปรับเทียบคะแนนอยู่ในระดับต่ำ

ข้อสังเกตที่ได้จากการวิจัยพบว่า ในการแปลงคะแนนระหว่างระดับชั้นที่อยู่ติดกัน (Adjacent Grade) เช่น การแปลงคะแนนจากแบบสอบระดับชั้น ม.2 ไปยังระดับชั้น ม.1 และการแปลงคะแนนจากแบบสอบระดับชั้น ม.2 ไปยังระดับชั้น ม.3 เป็นต้น พบว่ารูปแบบการแปลงคะแนนดังกล่าว มีความคลาดเคลื่อนของการปรับเทียบคะแนน (SEE) อยู่ในระดับที่ต่ำกว่า และมีความเพียงพอของการปรับเทียบคะแนนอยู่ในระดับที่น่าพอใจ ซึ่งบ่งบอกถึงความมีประสิทธิภาพของการปรับเทียบคะแนนมากกว่ารูปแบบการแปลงคะแนนในระดับชั้นที่ไม่ได้อยู่ติดกัน เช่น การแปลงคะแนนจากแบบสอบระดับชั้น ม.1 ไปยังแบบสอบระดับชั้น ม.3 และการแปลงคะแนนจากแบบสอบระดับชั้น ม.3 ไปยังระดับชั้น ม.1 เป็นต้น สาเหตุดังกล่าวอาจเนื่องมาจาก ในแต่ละรูปแบบของการแปลงคะแนนจะต้องเกิดความคลาดเคลื่อนของการปรับเทียบคะแนนในระดับหนึ่งเสมออาจจะมากหรือน้อยแตกต่างกันไป สำหรับในกรณีนี้ อาจเกิดปัญหาที่หลีกเลี่ยงไม่ได้ในกรณีที่มีการปรับเทียบกับกลุ่มผู้สอบ ที่มีระดับความสามารถหรือระดับชั้นที่แตกต่างกันมากๆ ที่เรียกว่า “Out-of-Level Effect” (O’Brien & Tohn, 1984; Bielinksi, Thurlow & Scottt, 2000; Leung, 2003) สำหรับการปรับเทียบคะแนนในระดับชั้นที่แตกต่างกันมากๆ โดยเฉพาะการปรับเทียบระหว่างระดับชั้น ม.1 กับ ม.3 โดยเป็นระดับชั้นที่กลุ่มผู้สอบมีความสามารถแตกต่างกันอย่างชัดเจน โดยเฉพาะคะแนนที่อยู่ในส่วนปลายของข้อมูล ซึ่งสอดคล้องกับการศึกษาของ พรพิมล นาคเวช (2537) ที่พบว่าความแตกต่างของคะแนนสมมูลจากการปรับเทียบคะแนนจากตารางปรับเทียบคะแนนกับคะแนนที่สอบจริง จะมีความคลาดเคลื่อนเฉลี่ยของการปรับเทียบคะแนนสูงตรงระดับคะแนนส่วนปลาย

ทั้งนี้เนื่องมาจากที่คะแนนส่วนปลายมีผู้ได้คะแนนในส่วนนี้น้อย จึงเกิดปัญหาข้อมูลไม่เพียงพอจึงทำให้ความคลาดเคลื่อนส่วนนี้มากกว่าส่วนอื่น

อย่างไรก็ตามค่าความคลาดเคลื่อนที่ได้จากการปรับเทียบตามทฤษฎีตอบสนองข้อสอบแบบโลจิสติก 3 พารามิเตอร์ ร่วมกับการปรับเทียบเชิงเส้นตรง สามารถลดความคลาดเคลื่อนของการปรับเทียบคะแนนได้ เนื่องจากเมื่อพิจารณาจากค่าความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน พบว่ามีค่าแตกต่างกันเพียงเล็กน้อยจากรูปแบบการแปลงคะแนนในระดับชั้นที่ติดกัน และความเพียงพอของการปรับเทียบคะแนนอยู่ในระดับที่ยอมรับได้ และเมื่อพิจารณารูปแบบการปรับเทียบโดยรวม ที่บ่งบอกถึงความถูกต้อง (Accuracy) และความแม่นยำ (Precision) แต่ข้อค้นพบที่ได้จากการวิจัยครั้งนี้ พบว่า การปรับแก้ค่าในรูปแบบของการนำวิธีการปรับเทียบเชิงเส้นตรงมาใช้ในการปรับแก้ค่าที่ได้จากการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบนั้น ค่าที่ได้จากการปรับเทียบตามทฤษฎีตอบสนองข้อสอบ กับค่าที่ได้จากการปรับเทียบเชิงเส้นตรงด้วยวิธี “Simple Least Square Regression” แตกต่างกันเพียงเล็กน้อยเท่านั้น ซึ่งแตกต่างกับงานวิจัยในต่างประเทศ เช่น จากการศึกษาของ Leung (2003) ที่นำวิธีการปรับเทียบเชิงเส้นตรงมาช่วยแก้ปัญหา “Out-of-Level Effect” แล้วค่าที่ได้จะแตกต่างจากการปรับเทียบเชิงเส้นตรงอย่างชัดเจน และมีประสิทธิภาพมากกว่าการปรับเทียบตามทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์เพียงอย่างเดียว ซึ่งไม่สอดคล้องกับการศึกษาในครั้งนี้ สาเหตุอาจเนื่องมาจากการศึกษาของ Leung (2003) เป็นการศึกษาถึง 6 ระดับความสามารถ คือนักเรียนในระดับ 1-6 ซึ่งต้องใช้สมการเชิงเส้นตรงที่ได้จากการปรับเทียบเชิงเส้นตรง 4 สมการ จึงจะสามารถปรับแก้ค่าที่ได้จากการปรับเทียบตามทฤษฎีตอบสนองข้อสอบแบบโลจิสติก 3 พารามิเตอร์ได้ครบทั้ง 6 ระดับความสามารถ แต่ในบริบทที่ทำการศึกษาในครั้งนี้จะศึกษาเพียง 3 ระดับความสามารถ ซึ่งใช้สมการเชิงเส้นตรงที่ได้จากการปรับเทียบเชิงเส้นตรงเพียง 1 สมการ ดังนั้นค่าที่ได้จากการปรับเทียบเชิงเส้นตรง จึงแตกต่างกันเพียงเล็กน้อยจากการปรับเทียบตามทฤษฎีตอบสนองข้อสอบ

ประเด็นที่น่าสังเกตที่ได้จากการวิจัยครั้งนี้อีกประการหนึ่ง คือ ประสิทธิภาพของการปรับเทียบคะแนนเมื่อพิจารณาจาก ความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน (SEE) และ ดัชนีความเพียงพอของการปรับเทียบคะแนน (C) ซึ่งถ้าทั้ง 2 ค่านี้ มีค่าน้อย จะบ่งบอกถึงความมีประสิทธิภาพของการปรับเทียบคะแนนเพิ่มมากขึ้น ซึ่งค่าที่ส่งผลต่อความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน คือ ค่าส่วนเบี่ยงเบนมาตรฐานของคะแนนความคลาดเคลื่อนของการปรับเทียบคะแนนระหว่างคะแนนที่ได้จากตารางปรับเทียบคะแนนกับคะแนนที่นักเรียน

สอบได้จริง (SD_E) ซึ่งถ้าค่าคะแนนความคลาดเคลื่อนของการปรับเทียบคะแนนระหว่างคะแนนที่ได้จากตารางปรับเทียบคะแนนกับคะแนนที่นักเรียนสอบได้จริงแตกต่างกันไม่มากและมีการกระจาย (Diviation) ของข้อมูลที่ได้น้อย และค่าที่ได้ค่อนข้างมีความคงที่ (Stable) จะทำให้ความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนน (SEE) ต่ำ ขณะเดียวกันก็ใช้ดัชนีความแตกต่างของการปรับเทียบคะแนน (C) จากกลุ่มสอบทานผลซึ่งเป็นดัชนีที่แสดงให้เห็นว่าคะแนนแปลงจากตารางปรับเทียบคะแนนนั้นแตกต่างจากคะแนนที่ผู้สอบทำได้จริงเท่าใด ความแตกต่างนี้ถือเป็นความคลาดเคลื่อนรวมที่เกิดจากการปรับเทียบคะแนนในแต่ละวิธี ซึ่งจะถ่วงน้ำหนักด้วยค่าความแปรปรวนของคะแนนที่นักเรียนสอบได้จริงซึ่งถือเป็นคะแนนเกณฑ์ เพื่อให้ค่าที่ได้มีความเป็นมาตรฐาน ซึ่งค่าคะแนนความคลาดเคลื่อนของการปรับเทียบคะแนนระหว่างคะแนนที่ได้จากตารางปรับเทียบคะแนนกับคะแนนที่นักเรียนสอบได้จริงควรมีค่าน้อย และกลุ่มตัวอย่างสอบทานผลควรมีค่าที่เพียงพอต่อการตรวจสอบประสิทธิภาพของการปรับเทียบคะแนน ซึ่งสอดคล้องกับผลการศึกษาของ Kolen และ Whitney (1982) ที่ได้แนะนำให้ใช้ข้อมูลคะแนนจากผู้สอบเป็นเกณฑ์ในการหาความแตกต่าง ซึ่งสุ่มตัวอย่างมาจากประชากรชุดเดียวกันกับกลุ่มปรับเทียบคะแนนในจำนวนไม่น้อยกว่าร้อยละ 10 ของกลุ่มปรับเทียบคะแนน ซึ่งเป็นจำนวนที่เหมาะสม

2.2 การอภิปรายตามสมมติฐานการวิจัยข้อที่ 2 การศึกษาพัฒนาการความสามารถของนักเรียนโดยภาพรวม

ผลการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 โดยภาพรวม จากการพิจารณาอัตราพัฒนาการ (Growth Rate) พบว่าความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.1 เมื่อขึ้น ม.2 และระดับชั้น ม.2 เมื่อขึ้น ม.3 มีพัฒนาการสูงขึ้น ส่วนในระดับชั้นที่อยู่ห่างกัน คือ ระดับชั้น ม.1 เมื่อขึ้น ม.3 ความสามารถทางคณิตศาสตร์ของนักเรียนมีพัฒนาการสูงมาก ซึ่งสอดคล้องกับสมมติฐานการวิจัย ข้อ 2 ที่ตั้งไว้ว่าความสามารถทางคณิตศาสตร์ของนักเรียนมีพัฒนาการสูงขึ้นตามระดับชั้นที่สูงขึ้น โดยผลการวิจัยที่ได้สอดคล้องกับผลการศึกษาของ ภัทรพร เกษสังข์ (2546) ที่เป็นเช่นนี้น่าจะเนื่องมาจาก พัฒนาการความสามารถทางคณิตศาสตร์เพิ่มขึ้นตามประสบการณ์ที่สูงขึ้นของนักเรียนเพราะมีโอกาสพัฒนาตัวเองได้มากขึ้น และเนื้อหาที่เรียนเป็นเนื้อหาที่ต่อเนื่องกันและนักเรียนมีความรู้พื้นฐานในเนื้อหานั้น ทำให้การเรียนในเนื้อหาที่มีระดับสูงขึ้นจึงเป็นการต่อยอดจากเนื้อหาเดิม ซึ่งสามารถอธิบายได้ด้วยทฤษฎีด้วยทฤษฎีสติปัญญาของ Cattell (Cattell's Theory of Fluid and Crystallized Intelligence) ที่กล่าวว่าความสามารถทางสติปัญญาทางสมองของมนุษย์แบ่งออกได้เป็น 2 ส่วน คือ ความสามารถทางสติปัญญาที่ตกผลึก (Crystallized

Ability) ซึ่งเป็นความสามารถที่ขึ้นอยู่กับการเรียนรู้ และได้เข้ามาจากสิ่งที่เข้ามาในชีวิตประจำวัน เช่น ความเข้าใจภาษา ความเข้าใจตัวเลข ความสามารถในการประเมินผล ความสามารถในการประเมินค่า เป็นต้น ความสามารถดังกล่าวเป็นความสามารถที่ได้รับการสะสมมาเป็นเวลานานจนกล่าวได้ว่าเป็นความสามารถที่ตกผลึก ซึ่งเป็นความสามารถที่เป็นผลมาจากการเรียนรู้ หรือประสบการณ์ในเนื้อหาวิชา ดังนั้นการเรียนรู้ที่เกิดขึ้นในชั้นเรียน หรือหลักการที่เป็นความรู้ในเนื้อหาวิชา จึงถือได้ว่าเป็นความสามารถที่ตกผลึก และความสามารถที่เลื่อนไหลได้ (Fluid Ability) เป็นความสามารถที่เป็นผลมาจากการเรียนรู้ และประสบการณ์ ผู้ที่มีความสามารถด้านนี้สูงจะสามารถทำงานชนิดต่างๆ ได้ดี ความสามารถด้านนี้จะแทรกอยู่ในทุกอริยบทของกิจกรรม เป็นความสามารถในด้านความพร้อมในการเรียนรู้ จะเห็นว่าความสามารถทางสติปัญญาก็คือความสามารถที่ตกผลึก และความสามารถทางสติปัญญาที่เลื่อนไหล ซึ่งเป็นความสามารถที่สำคัญในการเผชิญโลกทางวิชาการของแต่ละบุคคล ดังนั้น เมื่อนักเรียนเริ่มมีวุฒิภาวะเพิ่มขึ้น จึงมีประสบการณ์เพิ่มมากขึ้น จึงส่งผลให้มีพัฒนาการเพิ่มมากขึ้นตามระดับชั้นที่สูงขึ้นตามไปด้วย

ข้อสังเกตที่ได้จากผลการศึกษาในครั้งนี้คือ ถึงแม้พัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนทั้ง 3 ระดับชั้น จะเพิ่มขึ้นตามประสบการณ์ที่มีมากขึ้น แต่คะแนนความสามารถเฉลี่ยมีค่าติดลบทั้ง 3 ระดับ ซึ่งแสดงว่านักเรียนส่วนใหญ่มีความสามารถอยู่ในระดับปานกลางค่อนข้างต่ำ นอกจากนี้คะแนนที่ได้ทั้งก่อนและหลังการปรับเทียบมีคะแนนต่ำกว่าร้อยละ 50 ของคะแนนเต็ม ซึ่งสอดคล้องกับผลการศึกษาของ อวยพร เรืองตระกูล (2544) ที่พบว่าผลสัมฤทธิ์ทางคณิตศาสตร์ของนักเรียนระดับชั้น ม.2 และระดับชั้น ม.3 มีพัฒนาการค่อนข้างต่ำ ส่วนใหญ่จะมีพัฒนาการทางคณิตศาสตร์ต่ำกว่าร้อยละ 50 ของคะแนนเต็ม นอกจากนี้ถ้าพิจารณาพัฒนาการความสามารถทางคณิตศาสตร์ในระดับประถมศึกษาซึ่งเป็นระดับชั้นที่ต่อเนื่องก่อนถึงระดับมัธยมศึกษา พบว่ามีแนวโน้มคล้ายคลึงกัน เช่น จากผลการประเมินของกรมวิชาการร่วมกับ สปช. ตั้งแต่ปี 2527-2541 พบว่า คณิตศาสตร์มีอัตราความก้าวหน้าเพิ่มขึ้นประมาณร้อยละ 0.5-1.2 แต่มีค่าเฉลี่ยต่ำกว่าร้อยละ 40 ของคะแนนเต็ม ซึ่งสอดคล้องกับผลการศึกษาของ วันทยา วงศ์ศิลปะภิรมย์ (2543) ที่ศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้น ป.1-6 ในปีการศึกษา 2536 เป็นการศึกษาภาคตัดขวางที่ใช้นักเรียนทั้ง 6 ระดับชั้นในปีการศึกษาเดียวกันโดยใช้วิธีการปรับเทียบแนวตั้งทั้งวิธีการปรับเทียบคะแนนตามทฤษฎีการสอบแบบดั้งเดิม (เฉพาะวิธีการของเธอร์สโตน) และวิธีการตามทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 2 พารามิเตอร์ โดยทำการประเมิน 3 สมรรถภาพหลัก คือ ทักษะการคำนวณ ความเข้าใจทางคณิตศาสตร์ และ ทักษะการแก้โจทย์ปัญหา ผลการวิจัยพบว่า นักเรียนมีพัฒนาการระหว่างระดับชั้นอย่างต่อเนื่องในทุกสมรรถภาพ แต่ไม่มีสมรรถภาพใดใน

ชั้นใดที่มีค่าเฉลี่ยร้อยละสูงกว่าร้อยละ 50 หรือมี นักเรียนที่มีผลน่าพอใจมากกว่าร้อยละ 50 และความสามารถด้านการแก้โจทย์ปัญหาและการคำนวณซึ่งเป็นสมรรถภาพที่มีผลการประเมินต่ำกว่าสมรรถภาพอื่น

จากผลการวิจัยดังกล่าวข้างต้น ตั้งแต่อดีตจนถึงปัจจุบัน โดยทำการศึกษาตั้งแต่ระดับประถมศึกษาจนถึงระดับมัธยมศึกษา พบว่าพัฒนาการทางด้านคณิตศาสตร์มีแนวโน้มเพิ่มขึ้นตามระดับชั้นที่สูงขึ้น แต่มีอัตราที่เพิ่มขึ้นน้อย และคะแนนที่ได้ค่อนข้างไปทางคะแนนน้อย สาเหตุอาจเนื่องมาจากคณิตศาสตร์จำเป็นที่จะต้องอาศัยความสามารถหลายด้านประกอบกัน เช่น ความสามารถด้านการคิดคำนวณ ความสามารถด้านการแก้โจทย์ปัญหา เป็นต้น ซึ่งความสามารถด้านการแก้โจทย์ปัญหาเป็นสมรรถภาพที่มีผลการประเมินต่ำกว่าสมรรถภาพอื่นๆ มาโดยตลอด

นอกจากนี้ เมื่อพิจารณาถึงแบบวัดที่นำมาใช้ในการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์สำหรับงานวิจัยในครั้งนี้ ได้ทำการทดสอบความสามารถทางคณิตศาสตร์ เรื่อง “จำนวนและการดำเนินการ” ซึ่งจากงานวิจัยที่ผ่านมา ได้ศึกษาความสัมพันธ์ระหว่างสมรรถภาพทางสมองบางประการกับผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียนระดับชั้นต่างๆ ที่พบว่าสมรรถภาพสมองด้านจำนวนเป็นตัวพยากรณ์ที่ดีตัวหนึ่งในการพยากรณ์ นอกจากนี้มีการศึกษาโดยทำการวิเคราะห์องค์ประกอบความสามารถทางคณิตศาสตร์ ของนักเรียนระดับชั้นต่างๆ ผลปรากฏว่าองค์ประกอบด้านจำนวน เป็นองค์ประกอบหนึ่งของความสามารถทางคณิตศาสตร์ และสรุปว่าความสามารถด้านจำนวน เป็นความสามารถหนึ่งที่จำเป็นอย่างยิ่งต่อการเรียนคณิตศาสตร์เพราะการเรียนคณิตศาสตร์จะเกี่ยวข้องกับตัวเลขและการแก้โจทย์ปัญหาต่างๆ ที่เป็นสถานการณ์ใหม่อยู่เสมอ ดังนั้นผู้เรียนคณิตศาสตร์ได้ดีจะต้องเป็นผู้ที่มีความสามารถด้านจำนวนด้วย ซึ่งสอดคล้องกับกิติพงษ์ ลิขิตบุญฤทธิ์ (2537) ที่ได้ทำการวิเคราะห์องค์ประกอบแบบสอบความถนัดที่สัมพันธ์กับผลสัมฤทธิ์ทางการเรียนาคณิตศาสตร์ของนักเรียนระดับชั้น ม.1 พบว่า ความถนัดด้านจำนวนมีความสัมพันธ์กับผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ และเป็นองค์ประกอบที่พยากรณ์ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ได้ดีด้วย ดังนั้นสาเหตุที่สำคัญอีกประการหนึ่งที่ทำให้การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์มีค่าเฉลี่ยอยู่ในระดับไม่เป็นที่น่าพอใจ อาจเนื่องมาจากการขาดทักษะในด้านจำนวน ดังนั้นควรหาวิธีการแก้ปัญหานี้อย่างเร่งด่วนและจริงจัง

2.3 การอภิปรายตามสมมติฐานการวิจัยข้อที่ 3 การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนจำแนกตามกลุ่มความสามารถ

ผลการศึกษาพัฒนาการความสามารถทางคณิตศาสตร์จำแนกตามกลุ่มความสามารถ พบว่า พัฒนาการความสามารถทางคณิตศาสตร์จำแนกตามกลุ่มความสามารถ

คือ กลุ่มต่ำ กลุ่มปานกลาง และกลุ่มสูง สอดคล้องกับสมมติฐานการวิจัยข้อ 3 คือพัฒนาการทุกกลุ่มความสามารถเพิ่มขึ้นตามระดับชั้นที่สูงขึ้น นอกจากนี้เมื่อพิจารณาอัตราพัฒนาการ (Growth Rate) ระหว่างกลุ่ม พบว่า นักเรียนที่อยู่ในกลุ่มความสามารถสูงมีพัฒนาการความสามารถทางคณิตศาสตร์สูงกว่านักเรียนที่อยู่ในกลุ่มปานกลางและกลุ่มต่ำ ตามลำดับ ที่เป็นเช่นนี้อาจจะเนื่องมาจากนักเรียนที่มีความสามารถสูงมีการพัฒนาความรู้ ความสามารถในระดับสูงอยู่แล้วและก่อนที่จะเรียนเนื้อหาใหม่จะมีความรู้บางส่วนที่ดีกว่า และเป็นการเรียนเพื่อการต่อยอด ทำให้เกิดการเรียนรู้ได้รวดเร็วกว่านักเรียนที่มีความสามารถ ปานกลาง และต่ำ ที่ต้องใช้เวลาในการเรียนรู้ ซึ่งสอดคล้องกับหลักการเรียนรู้ทางการศึกษาที่สามารถอธิบายได้ด้วยทฤษฎีการเรียนรู้ที่มีความเชื่อตามหลักการศึกษาที่เท่ากันว่า “นักเรียนเรียนรู้ได้เท่ากันแต่ต้องใช้เวลาในการเรียนรู้หรือการศึกษาที่ต่างกัน” (อ้างใน ศิริชัย กาญจนวาสี, 2544)

นอกจากนี้ผู้วิจัยพบว่านักเรียนที่อยู่ในกลุ่มความสามารถสูง ส่วนใหญ่เป็นนักเรียนในโรงเรียนประจำจังหวัด ซึ่งเป็นโรงเรียนที่มีชื่อเสียงทางวิชาการ มีความพร้อมในบุคลากรและสื่อการศึกษา อีกทั้งนักเรียนจะมีความกระตือรือร้นในการเรียนอยู่ตลอดเวลา เพื่อเตรียมพร้อมกับการศึกษาในระดับที่สูงขึ้นไป โดยเฉพาะการเตรียมความพร้อมเพื่อศึกษาในระดับช่วงชั้นที่ 4 (ม.4-6) เพื่อการศึกษาต่อในระดับอุดมศึกษาที่ต้องใช้คะแนน GPA เพื่อเป็นส่วนหนึ่งในการสอบคัดเลือกเข้าศึกษาต่อ ดังนั้นนักเรียนที่มีความสามารถสูงจึงมีแรงจูงใจใฝ่สัมฤทธิ์สูงในการศึกษา จึงส่งผลให้พัฒนาการความสามารถทางคณิตศาสตร์สูงกว่าทุกกลุ่ม

ข้อเสนอแนะ

1. ข้อเสนอแนะเชิงนโยบาย

ถึงแม้ผลการวิจัยครั้งนี้พบว่า พัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนนักเรียนระดับชั้น ม.1, 2 และ 3 มีพัฒนาการเพิ่มขึ้นตามระดับชั้นที่สูงขึ้น แต่ผลการวิจัยพบว่านักเรียนส่วนใหญ่มีความสามารถอยู่ในระดับปานกลางค่อนข้างต่ำ นอกจากนี้คะแนนที่ได้ต่ำกว่าร้อยละ 50 ของคะแนนเต็มอย่างเห็นได้ชัด ดังนั้นหน่วยงานที่เกี่ยวข้องน่าจะมีมาตรการส่งเสริมสนับสนุนเพื่อหาแนวทางแก้ไขอย่างเร่งด่วน

2. ข้อเสนอแนะในการนำไปปฏิบัติ

2.1 การแปลความหมายของคะแนนที่ปรับเทียบแล้ว ควรระวังการสื่อความหมายหรือการนำไปใช้อ้างอิง ด้วยความหมายของคะแนนเหล่านี้มิได้แปลความหมายเหมือนคะแนนจากการสอบ เพราะคะแนนที่ได้ไม่ได้สื่อความถึงตัวแทนความรู้จากประชากรความรู้ในระดับชั้นนั้นๆ แต่เป็นค่าความสัมพันธ์เมื่อเปรียบเทียบกับระดับชั้นอื่น

2.2 เนื่องจากผลการวิจัยในครั้งนี้ พบว่าการแปลงคะแนนในระดับชั้น ม.2 ไปยังระดับชั้น ม.1 และ ม.3 มีความคลาดเคลื่อนมาตรฐานของการปรับเทียบคะแนนต่ำกว่ารูปแบบการแปลงคะแนนรูปแบบอื่น และมีความเพียงพอของการปรับเทียบคะแนนอยู่ในระดับน่าพอใจ ซึ่งบ่งบอกถึงความคงที่ของการปรับเทียบคะแนน ดังนั้นในการศึกษาเพื่อปรับเทียบคะแนนแนวตั้งจากกลุ่มผู้สอบที่มีความสามารถแตกต่างกัน 3 ระดับชั้น คือ ระดับชั้น ม.1 ม.2 และ ม.3 ควรใช้รูปแบบการแปลงคะแนนในระดับชั้นที่อยู่ติดกัน (Adjacent Grade) เพื่อลดความคลาดเคลื่อนของการปรับเทียบคะแนน และเป็นการเพิ่มประสิทธิภาพของการปรับเทียบคะแนน เช่น การแปลงคะแนนจากระดับชั้น ม.2 ให้อยู่ในสเกลเดียวกับ ม.1 และ ม.3 เป็นต้น

2.3 จากข้อค้นพบที่ได้จากการวิจัยในครั้งนี้ พบว่าการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ หรือการปรับเทียบตามทฤษฎีตอบสนองข้อสอบแบบโลจิสติก 3 พารามิเตอร์ ร่วมกับการปรับเทียบเชิงเส้นตรง ค่าที่ได้แตกต่างกันเพียงเล็กน้อย ดังนั้นในกรณีที่ต้องการศึกษาการปรับเทียบสำหรับกลุ่มผู้สอบที่มีความสามารถต่างกัน 3 ระดับชั้น ครูหรือผู้ที่เกี่ยวข้องกับการศึกษา สามารถใช้เฉพาะวิธีการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบ ก็เพียงพอสำหรับการศึกษาพัฒนาการความสามารถหรือความก้าวหน้าของผู้เรียนเนื่องจากเป็นวิธีการที่ไม่ยุ่งยากซับซ้อนและประหยัดเวลามากกว่า

2.4 เนื่องจากผลการวิจัยพบว่า นักเรียนที่อยู่ในกลุ่มต่ำมีพัฒนาการความสามารถทางคณิตศาสตร์ต่ำกว่ากลุ่มความสามารถอื่น ดังนั้นครูควรเอาใจใส่กลุ่มนี้เป็นพิเศษ โดยเฉพาะความสามารถทางด้าน “จำนวนและการดำเนินการ” ซึ่งนักเรียนยังขาดทักษะทางด้านการแก้โจทย์ปัญหาและการคำนวณ

3. ข้อเสนอแนะเพื่อทำวิจัยต่อ

3.1 ควรมีการพัฒนาวิธีการปรับเทียบคะแนนในกรณีที่แบบสอบที่มีลักษณะหลายมิติ (Multidimentional) เนื่องจากสถานการณ์สอบในปัจจุบันเป็นไปได้น้อยที่ข้อสอบหลายฉบับที่นำมาปรับเทียบกันจะมีลักษณะที่มีเนื้อหาเพียงมิติเดียว (Unidimentional) ซึ่งวัดคุณลักษณะเด่นเพียงคุณลักษณะเดียว

3.2 ควรมีการศึกษาเปรียบเทียบประสิทธิภาพของวิธีการปรับเทียบแนวตั้ง ในหลายๆ วิธี เพื่อนำไปศึกษากับผู้สอบที่มีความสามารถมากกว่า 3 ระดับความสามารถเพื่อนำวิธีการที่ดีที่สุดมาใช้ในการศึกษาความก้าวหน้าหรือพัฒนาการในด้านต่างๆ เนื่องจากในการวิจัยครั้งนี้พบว่าวิธีการปรับเทียบตามวิธีทฤษฎีตอบสนองข้อสอบแบบโลจิสติก 3 พารามิเตอร์ กับวิธีการปรับเทียบตามทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับการปรับเทียบเชิงเส้นตรง ที่ศึกษาสำหรับผู้สอบที่มีความสามารถแตกต่างกัน 3 ระดับความสามารถ มีค่าที่แตกต่างกันเพียงเล็กน้อย ดังนั้นในการศึกษาครั้งต่อไปควรที่จะศึกษาในประเด็นดังกล่าว

3.3 ควรมีการสร้างแบบสอบแบบเทเลอร์ (Tailored Testing) คือ การสอบที่มีการจัดข้อสอบให้มีความยากที่เหมาะสมกับระดับความสามารถของผู้สอบ โดยผู้สอบแต่ละคนจะได้รับชุดของข้อสอบที่แตกต่างกันตามระดับความสามารถ แต่ยังสามารถนำค่าความสามารถที่ประมาณค่าได้จากการสอบนี้มาเปรียบเทียบกันได้ ทั้งนี้เพราะความสามารถที่ประมาณค่าได้อยู่ในสเกลเดียวกันแล้ว เนื่องจากการศึกษาในครั้งนี้ได้ศึกษาพัฒนาการตามกลุ่มความสามารถ คือ กลุ่มต่ำ กลุ่มปานกลาง และกลุ่มสูง แต่แบบสอบที่ใช้ในการวัดยังคงเป็นแบบสอบชุดเดียวกันทั้ง 3 ระดับความสามารถโดยไม่ได้คำนึงถึงความแตกต่างระหว่างบุคคล

3.4 ควรมีการสร้างแบบสอบเฉพาะสำหรับผู้สอบที่มีความสามารถต่ำ (Disability or Low Ability) โดยเฉพาะ ที่เรียกว่า “Out-of-Level Testing” ซึ่งเป็นการสอบที่มีการจัดข้อสอบให้มีความยากที่เหมาะสมกับระดับความสามารถของผู้สอบที่มีระดับความสามารถต่ำ เพื่อลดโอกาสการเดาของข้อสอบ และสามารถลดปัญหาที่เรียกว่า “Floor Effect” ได้ แล้วจึงนำคะแนนที่ได้มาทำการปรับเทียบแนวตั้งกับแบบสอบของนักเรียนที่อยู่ในระดับความสามารถตามปกติที่เรียกว่า “In-Level Testing” ซึ่งสามารถนำค่าความสามารถที่ประมาณค่าได้จากการสอบนี้มาเปรียบเทียบกันได้ ทั้งนี้เพราะความสามารถที่ประมาณค่าได้อยู่ในสเกลเดียวกันแล้ว

รายการอ้างอิง

ภาษาไทย

กิติพงษ์ ลิขิตบุญฤทธิ์. (2537). *การวิเคราะห์องค์ประกอบแบบทดสอบความถนัดที่สัมพันธ์กับผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 ในจังหวัดอุดรธานี*. วิทยานิพนธ์ปริญญามหาบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร.

คณะกรรมการการศึกษาแห่งชาติ, สำนักงาน. (2545). *พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 (ฉบับแก้ไข 2545)*. กรุงเทพมหานคร: พริกหวานกราฟฟิก.

ชูศักดิ์ ขัมภลิขิต. (2528). *การเปรียบเทียบการเทียบมาตราระหว่างรูปแบบที่ใช้ผู้สอบร่วมกับรูปแบบที่ใช้แบบสอบร่วม*. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

ชูศักดิ์ ขัมภลิขิต. (2534). การปรับเทียบคะแนนสอบ. *บทคัดสรรทางวิชาการทดสอบ*. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร: มหาวิทยาลัยสุโขทัยธรรมาธิราช ศูนย์พัฒนาศึกษาแห่งชาติของประเทศไทย ทบวงมหาวิทยาลัย, 143-450.

ทิภาภรณ์ ทุงฤทธิ์. (2544). *การเปรียบเทียบผลของการปรับเทียบคะแนนโดยวิธีเชิงเส้นตรง เมื่อใช้วิธีการแปลงคะแนนที่แตกต่างกันในแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์*. วิทยานิพนธ์ปริญญามหาบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒประสานมิตร.

บุญเชิด ภิญโญอนันตพงษ์. (2544). *การประเมินการเรียนรู้ ที่เน้นผู้เรียนเป็นสำคัญ แนวคิดและหลักการ*. กรุงเทพมหานคร: โรงพิมพ์อมรินทร์พริ้นติ้งแอนด์พับลิชชิ่ง.

ผ่องพรรณ ตรัยมงคลกูล และ สุภาพ ฉัตราภรณ์. (2545). *การออกแบบการวิจัย*. พิมพ์ครั้งที่ 4. กรุงเทพมหานคร: สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์.

พรพิมล นาคเวช. (2537). *การศึกษาคุณภาพการเทียบมาตราในแนวตั้งโดยใช้ทฤษฎีตอบสนองข้อสอบ*. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

วันทยา วงศ์ศิลปะภิรมย์. (2543). การปรับเทียบคะแนนความสามารถทางคณิตศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 1-6. *วารสารวิชาการ* 3(พฤษภาคม 2543) : 47-54.

วรรณ์ดี แสงประทีปทอง. (2538). *การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้นมัธยมศึกษาตอนต้น โดยการปรับเทียบคะแนนตามแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบแบบจำลองโลจิสติกสามพารามิเตอร์*. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

วินิจ เทือกทอง. (2537). การศึกษาเปรียบเทียบประสิทธิภาพของการคำนวณคะแนนเพิ่มวิธีต่าง ๆ ด้วยวิธีการมอนติคาร์โล. วิทยานิพนธ์ปริญญามหาบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

ภัทราพร เกษสังข์. (2546). *การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาตอนต้นโดยการปรับเทียบคะแนนในแนวตั้งที่ใช้วิธีการที่เหมาะสม*. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

ภาวินี ศรีสุขวัฒนานันท์. (2529). *การเปรียบเทียบผลจากการใช้รูปแบบการเทียบมาตราที่ต่างกันเมื่อแบบสอบร่วมมีความยาวต่างกัน*. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต ภาควิชาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

เรวดี อินทสะระ. (2530). *การเปรียบเทียบความคลาดเคลื่อนของการเทียบมาตราระหว่างรูปแบบอิงทฤษฎีตอบสนองข้อสอบกับรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ*. วิทยานิพนธ์ปริญญามหาบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

ศิริชัย กาญจนวาสี. (2541). *การปรับเทียบคะแนนระหว่างแบบสอบ*. กรุงเทพมหานคร : คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2544). *ทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory)*. พิมพ์ครั้งที่ 4. กรุงเทพมหานคร : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2545). *ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theories)*. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี, ทวีวัฒน์ ปิตยานนท์ และ ดิเรก ศรีสุโข. (2544). *การเลือกใช้สถิติที่เหมาะสมสำหรับการวิจัย*. พิมพ์ครั้งที่ 3. กรุงเทพมหานคร: บริษัทบุญศิริการพิมพ์ จำกัด.

สุจินดา ผ่องอักษร. (2533). *การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนในกลุ่มทักษะของนักเรียนชั้นประถมศึกษาปีที่ 6 ที่เรียนจบตามหลักสูตรประถมศึกษา พุทธศักราช 2526 ในช่วงระยะเวลา 3 ปีการศึกษา (ปีการศึกษา 2529-2531) โดยใช้การปรับเทียบคะแนนรูปแบบราชส์*.

วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

สุนิสา จุ้ยม่วงศรี. (2536). *ผลของความยาวของแบบสอบร่วมที่มีต่อคุณภาพของวิธีการเทียบมาตราเชิงเส้นตรง.* วิทยานิพนธ์ปริญญามหาบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

สุนิสา จุ้ยม่วงศรี. (2546). *การศึกษาผลของการปรับเทียบคะแนนข้อสอบที่ตรวจให้คะแนนแบบหลายค่า.* วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

สำนักวิชาการและมาตรฐานการศึกษา. (2547). "การศึกษาสภาพการจัดการเรียนรู้ตามหลักสูตรการศึกษาขั้นพื้นฐานของโรงเรียนนำร่องและโรงเรียนเครือข่าย". *เอกสารในการสัมมนาทางวิชาการเรื่อง การวัดและประเมินผลมาตรฐานการศึกษาขั้นพื้นฐาน* เสนอที่ โรงแรมแอมบาสเดอร์ 5-6 สิงหาคม 2547.

ศึกษาธิการ, กระทรวง, กรมวิชาการ. (2544). *หลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544.* กรุงเทพมหานคร : โรงพิมพ์องค์การรับส่งสินค้าและครุภัณฑ์ (ร.ส.พ.).

ศึกษาธิการ, กระทรวง, สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ, สำนักนิเทศและพัฒนามาตรฐานการศึกษาแห่งชาติ. (2544). *พัฒนาการของคุณภาพนักเรียนประถมศึกษาและแนวทางการประเมิน.* กรุงเทพมหานคร : โรงพิมพ์การศาสนา.

อดิศร ศรีบุญวงศ์. (2545). *การพัฒนาเกณฑ์ตัดสินคุณภาพการปรับเทียบคะแนนตามวิธีทฤษฎีตอบสนองข้อสอบ.* วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต ภาควิชาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

อาภรณ์ กาญจนกิจโสภณ. (2635). *การสร้างแบบทดสอบและตารางเทียบคะแนนในแนวตั้งในวิชาคณิตศาสตร์.* วิทยานิพนธ์ปริญญามหาบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

อวยพร เรืองตระกูล. (2544). การพัฒนาและวิเคราะห์คุณภาพขององวิธีการวัดคะแนนพัฒนาการตามทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต ภาควิชาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

ภาษาอังกฤษ

Angoff, W.H. (1971). Scales, norm and equivalent score. In R.L. Throndike(Ed.), *Education Measurement*. (2nd ed.), 508-600. Washing, D.C. : American Council on Education.

Angoff, W.H. (1984). Scale, norms and equipvalent score, Princeton, New Jersey : *Educationnal measurement.*

Ayerve, R.I. (1992). The effectiveness of the equipercentile method and IRT three parameter model on vertical equating under verying condition of sample size, test length and anchor test length : Asimulation study. Doctoral Dissertation. Columbia university*. Dissertation Abstracts International* 53: 1841

Baker, F.B. (1996). Aninvestigation of the Sampling Distributions of Equating Coefficients. *Applied Phychological Measurement* 20(1) : 45-57

Bastari, B. (2000). *Linking Multiple-choice and Construces-Response Item to a Common Proficiency Scale*. (Online) Dissertation.Ec. D. America: Graduate School, Massachusetts Amherst University. Available : UMI/ Dissertration (2001) p. 1-78

Bielinksi, J., Thurlow, M. and Scott., J. (2000). *How out-of-level testing affects the psychometric quality of test scores (Out-of-Level Testing Project Report 2*). Minneapolis, MN : University of Minnesota, National Center on Education Outcomes.Retrieved (today's date),From http://education.umn.edu/NCEO/OnlinePubs/OOLT2.html

Braun, H.L. and Holland, P.W. (1982). Observed-Score Test Equating : A Mathematical Analysis of Some ETS Equating Procedures, In Holland, P.W. and Rubin, D.B (Eds.) *Test Equating*. New York : Academic Press.

Cook, L. and Eignor, D.R. (1989). Using item response theory in test score equating. *International Journal of Education Research*.13: 161-173

Flanagan, T.C. (1951). Units, Score and Norms. In EF . Linguist(Ed.) *Education Measurement*. Washing, D.C. : American Council on Education.

Glass, G. V. and Hopkins, K.D. (1996). *Statistical Methods in Education and Psychology*.(3th ed.). New Jersey : Prentice-Hall Inc.

Glowacki, M.L. (1991). The analysis of test equating models for the Alabama high school graduation examination. Doctoral Dissertation, university of Alabama. *Dissertation Abstracts International* 52(1991): 1722

Guliksen, H. (1950). *Theory of mental test*. New York : Wiley

Jodoin, M.G. and others. (2003). Acomparison of Linear, Fixed Common, and Concurrent Paramiter Estimation Equating Procedures in Capturimg Academic Growth. *The Journal of Experimental Education*. 71(3): 229-50

Hambleton, R.K. and Swaminathan, H. (1985). *Item Response Theory : Principle and Applications*. (2nd ed). Boston : Kluwer Nijhoff Publishing.

Johnson, E.G. and Owen, E. (1998). Linking the Nation Assessment of Educaation Progress (NAEP) and the third International Mathemation and Science Study (Timss) : A technical Report. (Online). U.S.A. : National Center For Education Statistics, *Research and Development Report*. Available : http://www.ETS.ORG.

Kolen, M.L. and Brennan, R.L. (1995). *Test Equating Method and Practices*. New York : Spring

Leung , S.O. (2003). A Practical Use of Vertical Equating by Combining IRT Equating and Linearing Equating. *Practical Assessment, Research & Evaluation.* Retrieved June 8, 2004 From http://PARE online.net/getven.asp? v=8&n=23

Lissitz. R.W., and Huynh. H. (2003). Vertical Equating for State Assessment: Issues and Solution in Determination of Adequating Yearly Progress and School Accountability. *Practical Assessment, Research & Evaluation*. Retrieved June 8, (2004) From http://PARE online.net/getven.asp? v=8&n=10

Lord, F.M. (1980). *Application of Item Response Theory to Practical Testing Problem.* New Jersey: Hillssdale Erlbaun.

O'Brien, M.l. and Tohn, D. (1984). Applying and Evaluating Rasch vertical Equating Proceduce for Out-of-Level Testing. *Paper presented at the Annual Meeting of the Eastern Educational Research Association* (West Palm Beach, FL, February 10, 1984).

Petersen, N.S., Marco, G.L. and Stewart, E.E. (1982). A test of the Adequarcy of linear score equating Methods, In Holland, P.W. and Rubin, D.B (Eds.) *Test Equating*. New York : Academic Press.

Petersen and others. (1982). A test of the Adequacy of linear score equating Methods, In R.L. Linn (ED.) *Educational Measurement*. (3rd ed.) New York : Macmillan

Slinde, J.A., and Linn, R.L. (1978). Vertical equating test: Fact or Phantom. *Journal of Educational Measurement* 14: 23-31.

Vale, D. (1986). Linkink Item Parameters onto a Common Scale. *Applied Phychological Measurement* 10(4) : 333-344.

Wingersky, M.S, and Lord, F.M. (1984). an investigation of methods for reducing sampling error in curtain IRT procures. *Applied Phychological Measurement* 8: 347-364.

Wiley, A. (1999). An Investigation into Two Model For Equating Examinations with Multiple Item Formats (Anchor Testing), Doctoral dissertation, Fordham University. *Dissertation Abstracts International*. 59(7-A).

Yang, Wen-ling. (1997). The effects of Common-Item Test Equating, Doctoral dissertation, Michigan State University. *Dissertation Abstracts International*. 59(7-A).

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ก
รายชื่อผู้เชี่ยวชาญ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รายชื่อผู้เชี่ยวชาญ

1. รองศาสตราจารย์ ดร. สุวัฒนา เอี่ยมอรพรรณ	อาจารย์ประจำภาควิชาหลักสูตร การสอน และเทคโนโลยีการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
2. รองศาสตราจารย์ ดร. ฉวีวรรณ เศวตมาลย์	อาจารย์ประจำภาควิชาหลักสูตรและการสอน คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร
3. อาจารย์ ไตรรงค์ เจนการ	นักวัดและประเมินผลการศึกษา กลุ่มส่งเสริมการเรียนการสอนและประเมินผล สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ
4. ดร. อรนุช มั่งมีสุขศิริ	นักวิชาการสอบ สำนักพัฒนานวัตกรรมการจัดการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ
5. อาจารย์ อัญฑิกา ทั่งทอง	อาจารย์กลุ่มสาระการเรียนรู้คณิตศาสตร์ โรงเรียนท่าบ้านหลวง จังหวัดชัยนาท
6. อาจารย์ นิรันดร์ ตัณฑัยย์	อาจารย์กลุ่มสาระการเรียนรู้คณิตศาสตร์ โรงเรียนวัดหงส์รัตนาราม กรุงเทพมหานคร
7. อาจารย์ เจตนา จี้กังวาฬ	อาจารย์กลุ่มสาระการเรียนรู้คณิตศาสตร์ โรงเรียนยานนาเวศวิทยา กรุงเทพมหานคร

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ข
ตารางแสดงแผนผังข้อสอบ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางภาคผนวกที่ 1 สรุปแผนผังข้อสอบวัดความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 เรื่อง “จำนวนและการดำเนินการ”

สาระการเรียนรู้	ผลการเรียนรู้	ระดับชั้น		
		ม.1	ม.2	ม.3
1. ระบบจำนวน	1. หา ห.ร.ม. หรือ ค.ร.น. ของจำนวนนับที่กำหนดให้ได้	✓	-	-
	2. ใช้ความรู้เกี่ยวกับ ห.ร.ม. หรือ ค.ร.น. แก้ปัญหาได้	✓	-	✓
	3. ระบุหรือยกตัวอย่างจำนวนเต็มบวก จำนวนเต็มลบ และจำนวนเต็มศูนย์ได้	✓	-	-
	4. เปรียบเทียบจำนวนเต็มได้	✓	-	-
	5. บวก ลบ คูณ หรือ หาร จำนวนเต็ม พร้อมทั้งนำความรู้และสมบัติเกี่ยวกับจำนวนเต็มไปใช้ได้	✓	✓	✓
	6. ระบุหรือยกตัวอย่างจำนวนจริง จำนวนตรรกยะ หรือจำนวนอตรรกยะได้	-	✓	-
	7. หารากที่สองหรือรากที่สามของจำนวนเต็มโดยการหารแยกตัวประกอบได้	-	✓	✓
	8. บวก ลบ คูณ หรือหาร จำนวนที่อยู่ในรูปเครื่องหมายกรณฑ์ได้	-	-	✓
	9. แก้โจทย์ปัญหาเกี่ยวกับรากที่สองหรือรากที่สามได้	-	✓	✓
	10. หารากที่สองหรือรากที่สามของจำนวนจริงที่กำหนดให้ โดยการประมาณค่าได้	-	-	✓
3. เลขยกกำลัง	11. เขียนเลขยกกำลังเป็นจำนวนเต็มแทนจำนวนที่กำหนดให้ได้	✓	-	-
	12. ใช้เลขยกกำลังในการเขียนแสดงจำนวนในรูปสัญกรณ์วิทยาศาสตร์ได้	✓	-	-
	13. คูณหรือหารเลขยกกำลังที่มีฐานเดียวกันและเลขยกกำลังเป็นจำนวนเต็ม พร้อมทั้งนำไปแก้โจทย์ปัญหาได้	✓	✓	-
	14. คูณหรือหารเลขยกกำลัง พร้อมทั้งนำไปแก้โจทย์ปัญหาได้	-	-	✓
4. เศษส่วนและทศนิยม	15. เขียนเศษส่วนในรูปทศนิยมซ้ำศูนย์ หรือเขียนทศนิยมในรูปเศษส่วนได้	✓	✓	-
	16. เขียนเศษส่วนในรูปทศนิยมซ้ำหรือเขียนทศนิยมซ้ำในรูปเศษส่วนได้	-	-	✓
	17. เปรียบเทียบเศษส่วนหรือทศนิยมได้	✓	✓	-
	18. บวก ลบ คูณ หรือหาร เศษส่วนหรือทศนิยม พร้อมทั้งนำไปแก้โจทย์ปัญหาได้	✓	✓	✓
5.อัตราส่วนและร้อยละ	19. หาค่าอัตราส่วน สัดส่วน หรือร้อยละ และนำไปแก้โจทย์ปัญหาได้	-	✓	✓

ตารางภาคผนวกที่ 2 น้ำหนักคะแนนของแบบสอบวัดความสามารถทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 เรื่อง “จำนวนและการดำเนินการ”

สาระการเรียนรู้	ผลการเรียนรู้	ระดับชั้น / น้ำหนัก (%)		
		ม.1	ม.2	ม.3
1. ระบบจำนวน	1. หา ห.ร.ม. หรือ ค.ร.น. ของจำนวนนับที่กำหนดให้ได้	13	-	-
	2. ใช้ความรู้เกี่ยวกับ ห.ร.ม. หรือ ค.ร.น. แก้ปัญหาได้	10	3	3
	3. ระบุหรือยกตัวอย่างจำนวนเต็มบวก จำนวนเต็มลบ และจำนวนเต็มศูนย์ได้	6	-	-
	4. เปรียบเทียบจำนวนเต็มได้	10	-	-
	5. บวก ลบ คูณ หรือ หาร จำนวนเต็ม พร้อมทั้งนำความรู้และสมบัติเกี่ยวกับจำนวนเต็มไปใช้ได้	26	6	10
	6. ระบุหรือยกตัวอย่างจำนวนจริง จำนวนตรรกยะ หรือจำนวนอตรรกยะได้	3	10	-
	7. หารากที่สองหรือรากที่สามของจำนวนเต็มโดยการหารแยกตัวประกอบได้	10	10	-
	8. บวก ลบ คูณ หรือหาร จำนวนที่อยู่ในรูปเครื่องหมายกรณฑ์ได้	-	6	38
	9. แก้โจทย์ปัญหาเกี่ยวกับรากที่สองหรือรากที่สามได้	-	3	3
	10. หารากที่สองหรือรากที่สามของจำนวนจริงที่กำหนดให้ โดยการประมาณค่า ได้	-	-	10
3. เลขยกกำลัง	11. เขียนเลขยกกำลังเป็นจำนวนเต็มแทนจำนวนที่กำหนดให้ได้	-	3	-
	12. ใช้เลขยกกำลังในการเขียนแสดงจำนวนในรูปสัญกรณ์วิทยาศาสตร์ได้	-	-	-
	13. คูณหรือหารเลขยกกำลังที่มีฐานเดียวกันและเลขยกกำลังเป็นจำนวนเต็ม พร้อมทั้งนำไปแก้โจทย์ปัญหาได้	3	3	-
	14. คูณหรือหารเลขยกกำลัง พร้อมทั้งนำไปแก้โจทย์ปัญหาได้	-	-	3
4. เศษส่วนและทศนิยม	15. เขียนเศษส่วนในรูปทศนิยมซ้ำศูนย์ หรือเขียนทศนิยมในรูปเศษส่วนได้	6	-	-
	16. เขียนเศษส่วนในรูปทศนิยมซ้ำหรือเขียนทศนิยมซ้ำในรูปเศษส่วนได้	-	-	10
	17. เปรียบเทียบเศษส่วนหรือทศนิยมได้	3	3	-
	18. บวก ลบ คูณ หรือหาร เศษส่วนหรือทศนิยม พร้อมทั้งนำไปแก้โจทย์ปัญหาได้	10	10	3
5.อัตราส่วนและร้อยละ	19. หาค่าอัตราส่วน สัดส่วน หรือร้อยละ และนำไปแก้โจทย์ปัญหาได้	-	43	20
	รวม	100	100	100

ตารางภาคผนวกที่ 3 แผนผังข้อสอบ (Testing Blueprint) **ระดับ ม.1**

แบบวัดความสามารถทางคณิตศาสตร์ เรื่อง "จำนวนและการดำเนินการ"
หลักสูตรการศึกษาขั้นพื้นฐาน พ.ศ. 2544 การสอบวัดด้านความสามารถ
ระดับชั้น มัธยมศึกษาปีที่ 1 จุดมุ่งหมายการทดสอบ : ประเมินพัฒนาการของผู้เรียน
เวลาที่ใช้ในการทดสอบ 60 นาที ผู้วางแผน พัชรี จันทร์เพ็ง

สาระการเรียนรู้	ผลการเรียนรู้	น้ำหนักรายข้อ	
		คิดเป็น %	จำนวนข้อ
1. ระบบจำนวนเต็ม	1. หา ห.ร.ม.หรือ ค.ร.น.ของจำนวนนับที่กำหนดให้ได้	13	4
	2. ใช้ความรู้เกี่ยวกับ ห.ร.ม. หรือ ค.ร.น. แก้โจทย์ปัญหาได้	10	3
	3. ระบุหรือยกตัวอย่างจำนวนเต็มบวก จำนวนเต็มลบ และจำนวนเต็มศูนย์ได้	6	2
	4. เปรียบเทียบจำนวนเต็มได้	10	3
	5. บวก ลบ คูณ หรือหาร จำนวนเต็ม พร้อมทั้งนำความรู้และสมบัติเกี่ยวกับจำนวนเต็มไปใช้ได้	26	7
2. เลขยกกำลัง	6. เขียนเลขยกกำลังเป็นจำนวนเต็มแทนจำนวนที่กำหนดให้ได้	3	1
	7. ใช้เลขยกกำลังในการเขียนแสดงจำนวนในรูปสัญกรณ์วิทยาศาสตร์ได้	10	3
	8. คูณหรือหารเลขยกกำลังที่มีฐานเดียวกัน และเลขชี้กำลังเป็นจำนวนเต็ม พร้อมทั้งนำไปแก้โจทย์ปัญหาในสถานการณ์ต่างๆ ได้	3	1
3. เศษส่วนและทศนิยม	9. เขียนเศษส่วนในรูปทศนิยมหรือเขียนทศนิยมในรูปเศษส่วนได้	6	2
	10. เปรียบเทียบเศษส่วนหรือทศนิยมได้	3	1
	11. บวก ลบ คูณ และหาร เศษส่วนหรือทศนิยม พร้อมทั้งนำไปแก้โจทย์ปัญหาในสถานการณ์ต่างๆ ได้	10	3
	รวม	100	30

ตารางภาคผนวกที่ 4 แผนผังข้อสอบ (Testing Blueprint) **ระดับ ม.2**

แบบวัดความสามารถทางคณิตศาสตร์ เรื่อง "จำนวนและการดำเนินการ"
หลักสูตรการศึกษาขั้นพื้นฐาน พ.ศ. 2544 การสอบวัดด้านความสามารถ
ระดับชั้น มัธยมศึกษาปีที่ 2 จุดมุ่งหมายการทดสอบ : ประเมินพัฒนาการของผู้เรียน
เวลาที่ใช้ในการทดสอบ 60 นาที ผู้วางแผน พัชรี จันทร์เพ็ง

สาระการเรียนรู้	ผลการเรียนรู้	น้ำหนักรายข้อ	
		คิดเป็น %	จำนวนข้อ
1. ระบบจำนวน	1. ใช้ความรู้เกี่ยวกับ ห.ร.ม. หรือ ค.ร.น. แก้โจทย์ปัญหาได้	3	1
	2. บวก ลบ คูณ หรือ หาร จำนวนเต็ม พร้อมทั้งนำความรู้และสมบัติเกี่ยวกับจำนวนเต็มไปใช้ได้	6	2
	3. ระบุหรือยกตัวอย่างจำนวนจริง จำนวนตรรกยะ หรือจำนวนอตรรกยะได้	10	3
	4. หารากที่สองหรือรากที่สามของจำนวนเต็ม โดยการหารแยกตัวประกอบพร้อมทั้งนำไปใช้แก้โจทย์ปัญหาได้	10	3
	5. บวก ลบ คูณ หรือหาร จำนวนที่อยู่ในรูปเครื่องหมายกรณฑ์ได้	6	2
	6. แก้โจทย์ปัญหาเกี่ยวกับรากที่สองหรือรากที่สามได้	3	1
	7. เขียนเลขยกกำลังเป็นจำนวนเต็มแทนจำนวนที่กำหนดให้ได้	3	1
2. เลขยกกำลัง	8. คูณหรือหารเลขยกกำลังที่มีฐานเดียวกันและเลขชี้กำลังเป็นจำนวนเต็ม พร้อมทั้งนำไปแก้โจทย์ปัญหาได้	3	1
	9. เปรียบเทียบเศษส่วนหรือทศนิยมได้	3	1
3. เศษส่วนและทศนิยม	10. บวก ลบ คูณ หรือหาร เศษส่วนหรือทศนิยม พร้อมทั้งนำไปแก้โจทย์ปัญหาได้	10	3
4.อัตราส่วนและร้อยละ	11. หาค่าอัตราส่วน สัดส่วน หรือร้อยละ พร้อมทั้งแก้โจทย์ปัญหาในสถานการณ์ต่างๆ ได้	43	12
	รวม	100	30

ตารางภาคผนวกที่ 5 แผนผังข้อสอบ (Testing Blueprint) ระดับ ม.3

แบบวัดความสามารถทางคณิตศาสตร์ เรื่อง "จำนวนและการดำเนินการ"
หลักสูตรการศึกษาขั้นพื้นฐาน พ.ศ. 2544 การสอบวัดด้านความสามารถ
ระดับชั้น มัธยมศึกษาปีที่ 3 จุดมุ่งหมายการทดสอบ : ประเมินพัฒนาการของผู้เรียน
เวลาที่ใช้ในการทดสอบ 60 นาที ผู้วางแผน พัชรี จันทร์เพ็ง

สาระการเรียนรู้	ผลการเรียนรู้	น้ำหนักรายข้อ	
		คิดเป็น %	จำนวนข้อ
1. ระบบจำนวน	1.ใช้ความรู้เกี่ยวกับ ห.ร.ม. และ ค.ร.น. แก้โจทย์ปัญหาได้	3	1
	2. บวก ลบ คูณ และหาร จำนวนเต็ม พร้อมทั้งนำความรู้และสมบัติเกี่ยวกับจำนวนเต็มไปใช้ได้	10	3
	3. บวก ลบ คูณ หรือหาร จำนวนที่อยู่ในรูปเครื่องหมายกรณฑ์ได้	38	11
	4. แก้โจทย์ปัญหาเกี่ยวกับรากที่สองหรือรากที่สามได้	3	1
	5. หารากที่สองหรือรากที่สามของจำนวนจริงที่กำหนดให้โดยการประมาณค่า การเปิดตาราง ได้	10	3
2. เลขยกกำลัง	6. คูณหรือหารเลขยกกำลัง พร้อมทั้งนำไปแก้โจทย์ปัญหาได้	3	1
3. เศษส่วนและทศนิยม	7. เขียนเศษส่วนในรูปทศนิยมซ้ำหรือเขียนทศนิยมซ้ำในรูปเศษส่วนได้	10	3
	8. บวก ลบ คูณ หรือหาร เศษส่วนและทศนิยม พร้อมทั้งนำไปแก้โจทย์ปัญหาได้	3	1
4.อัตราส่วนและร้อยละ	9. หาค่าอัตราส่วน สัดส่วน หรือร้อยละ พร้อมทั้งแก้โจทย์ปัญหาในสถานการณ์ต่างๆ ได้	20	6
	รวม	100	30

ภาคผนวก ค
ค่าความยาก และ ค่าอำนาจจำแนก ของการวิเคราะห์ข้อสอบ ตามทฤษฎี CTT (ค่า p และ r) และตามทฤษฎี IRT (ค่า b และ a) ของแบบสอบเรื่อง “จำนวนและการดำเนินการ” ของระดับชั้น ม.1, 2 และ 3

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางภาคผนวกที่ 6 ค่าพารามิเตอร์ของการวิเคราะห์ข้อสอบตามทฤษฎี CTT คือ ความยาก (p) และค่าอำนาจจำแนก (r) และตามทฤษฎี IRT คือ ความยาก (b) และ อำนาจจำแนก (a) ของแบบสอบเรื่อง “จำนวนและการดำเนินการ” ในระดับชั้น ม.1

ข้อที่	p	r	b	a	หมายเหตุ
1	0.575	0.333	-0.454	0.752	คัดเลือก
2	0.514	0.404	-0.790	0.835	คัดเลือก
3	0.734	0.228	-0.635	0.974	คัดเลือก
4	0.579	0.491	-0.349	1.102	คัดเลือก
5	0.734	0.228	-1.930	0.551	คัดเลือก
6	0.400	0.614	0.379	1.372	คัดเลือก
7	0.356	0.333	1.136	0.593	คัดเลือก
8	0.495	0.386	0.024	0.745	คัดเลือก
9	0.561	0.281	-0.595	0.424	คัดเลือก
10	0.453	0.509	0.242	0.850	คัดเลือก
11	0.593	0.281	-0.618	0.659	คัดเลือก
12	0.453	0.123	0.550	0.383	ตัดทิ้ง
13	0.533	0.281	-0.194	0.752	คัดเลือก
14	0.425	0.474	0.331	1.049	คัดเลือก
15	0.636	0.579	-0.556	1.233	คัดเลือก
16	0.425	0.368	0.791	0.416	คัดเลือก
17	0.491	0.404	0.053	0.728	คัดเลือก
18	0.509	0.316	-0.066	0.626	คัดเลือก
19	0.332	0.193	1.538	0.472	ตัดทิ้ง
20	0.285	0.158	1.902	0.503	ตัดทิ้ง
21	0.509	0.439	-0.047	1.041	คัดเลือก
22	0.659	0.368	-0.936	0.772	คัดเลือก
23	0.322	0.123	1.807	0.423	ตัดทิ้ง
24	0.280	0.316	1.392	0.733	คัดเลือก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ข้อที่	p	r	b	a	หมายเหตุ
25	0.229	0.298	1.837	0.711	คัดเลือก
26	0.678	0.421	-0.910	0.927	คัดเลือก
27	0.421	0.228	0.703	0.473	คัดเลือก
28	0.215	0.105	3.246	0.409	ตัดทิ้ง
29	0.31	0.246	1.315	0.523	คัดเลือก
30	0.145	0.158	2.494	0.774	ตัดทิ้ง
31	0.112	0.000	4.880	0.435	ตัดทิ้ง
32	0.336	0.316	1.051	0.695	คัดเลือก
33	0.224	-0.018	4.054	0.311	ตัดทิ้ง
34	0.453	0.386	0.279	0.721	คัดเลือก
35	0.327	0.474	0.790	1.061	คัดเลือก
36	0.379	0.386	0.707	0.764	คัดเลือก
37	0.220	0.193	0.672	2.014	ตัดทิ้ง
38	0.350	0.281	0.985	0.670	คัดเลือก
39	0.668	0.316	-1.111	0.678	คัดเลือก
40	0.304	0.281	1.407	0.626	คัดเลือก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางภาคผนวกที่ 7 ค่าพารามิเตอร์ของการวิเคราะห์ข้อสอบตามทฤษฎี CTT คือ ความยาก (p) และค่าอำนาจจำแนก (r) และตามทฤษฎี IRT คือ ความยาก (b) และ อำนาจจำแนก (a) ของแบบสอบเรื่อง “จำนวนและการดำเนินการ” ในระดับชั้น ม.2

ข้อที่	p	r	b	a	หมายเหตุ
1	0.450	0.530	-0.198	0.735	คัดเลือก
2	0.400	0.252	1.290	0.930	คัดเลือก
3	0.133	-0.075	5.017	0.420	ตัดทิ้ง
4	0.536	0.450	-0.194	1.008	คัดเลือก
5	0.536	0.575	-1.720	1.305	คัดเลือก
6	0.172	0.125	2.821	0.580	ตัดทิ้ง
7	0.325	0.250	1.278	0.596	คัดเลือก
8	0.079	0.075	4.010	0.639	ตัดทิ้ง
9	0.384	0.475	0.472	1.097	คัดเลือก
10	0.344	0.050	0.423	1.549	ตัดทิ้ง
11	0.318	0.575	0.607	1.530	คัดเลือก
12	0.139	0.100	2.612	0.745	ตัดทิ้ง
13	0.430	0.600	0.210	1.420	คัดเลือก
14	0.219	0.250	1.699	0.810	คัดเลือก
15	0.351	0.150	1.199	0.527	ตัดทิ้ง
16	0.285	0.200	1.610	0.592	คัดเลือก
17	0.377	0.175	0.971	0.528	ตัดทิ้ง
18	0.200	0.404	0.730	0.544	คัดเลือก
19	0.364	0.525	0.511	1.232	คัดเลือก
20	0.391	0.575	0.362	1.394	คัดเลือก
21	0.497	0.500	-0.011	0.915	คัดเลือก
22	0.510	0.125	-0.098	0.519	ตัดทิ้ง
23	0.715	0.350	0.603	0.829	คัดเลือก
24	0.384	0.350	0.603	0.829	คัดเลือก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ข้อที่	p	r	b	a	หมายเหตุ
25	0.371	0.275	0.863	0.636	คัดเลือก
26	0.411	0.200	0.624	0.592	คัดเลือก
27	0.185	0.200	1.879	0.858	คัดเลือก
28	0.450	0.575	0.158	1.249	คัดเลือก
29	0.338	0.050	1.483	0.464	คัดเลือก
30	0.636	0.225	-1.090	0.543	คัดเลือก
31	0.205	0.250	1.747	0.841	คัดเลือก
32	0.404	0.500	0.417	0.997	คัดเลือก
33	0.238	0.200	1.968	0.618	คัดเลือก
34	0.371	0.375	0.771	0.719	คัดเลือก
35	0.384	0.275	0.933	0.517	คัดเลือก
36	0.291	0.275	1.450	0.642	คัดเลือก
37	0.338	0.050	1.551	0.443	ตัดทิ้ง
38	0.146	0.100	3.122	0.591	ตัดทิ้ง
39	0.166	-0.050	3.940	0.419	ตัดทิ้ง
40	0.172	0.075	3.152	0.514	ตัดทิ้ง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางภาคผนวกที่ 8 ค่าพารามิเตอร์ของการวิเคราะห์ข้อสอบตามทฤษฎี CTT คือ ความยาก (p) และค่าอำนาจจำแนก (r) และตามทฤษฎี IRT คือ ความยาก (b) และ อำนาจจำแนก (a) ของแบบสอบเรื่อง “จำนวนและการดำเนินการ” ในระดับชั้น ม.3

ข้อที่	p	r	b	a	หมายเหตุ
1	0.596	0.615	-0.430	1.146	คัดเลือก
2	0.162	0.077	2.651	0.652	ตัดทิ้ง
3	0.051	-0.038	4.643	0.658	ตัดทิ้ง
4	0.465	0.615	0.097	1.531	คัดเลือก
5	0.566	0.615	-0.268	1.368	คัดเลือก
6	0.586	0.385	-0.516	0.787	คัดเลือก
7	0.667	0.500	0.960	-0.852	คัดเลือก
8	0.538	0.364	0.565	1.132	คัดเลือก
9	0.458	0.654	-0.192	1.219	คัดเลือก
10	0.343	0.154	0.660	1.023	ตัดทิ้ง
11	0.566	0.308	-0.394	0.801	คัดเลือก
12	0.505	0.538	-0.044	1.248	คัดเลือก
13	0.424	0.462	0.385	0.823	คัดเลือก
14	0.111	-0.154	4.541	0.467	ตัดทิ้ง
15	0.495	0.654	-0.004	1.249	คัดเลือก
16	0.404	0.538	1.262	0.355	คัดเลือก
17	0.424	0.731	0.250	1.433	คัดเลือก
18	0.692	0.577	0.464	1.319	คัดเลือก
19	0.465	0.538	0.147	0.935	คัดเลือก
20	0.414	0.538	0.277	1.504	คัดเลือก
21	0.222	-0.077	3.051	0.417	ตัดทิ้ง
22	0.364	0.538	0.609	1.024	คัดเลือก
23	0.077	0.202	0.600	2.385	ตัดทิ้ง
24	0.354	0.077	1.359	0.447	ตัดทิ้ง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ข้อที่	p	r	b	a	หมายเหตุ
25	0.364	0.385	0.655	0.935	คัดเลือก
26	0.263	-0.154	2.617	0.399	ตัดทิ้ง
27	0.596	0.308	-0.744	0.585	คัดเลือก
28	0.182	-0.115	3.538	0.433	ตัดทิ้ง
29	0.505	0.538	-0.053	0.854	คัดเลือก
30	0.293	0.423	1.131	0.846	คัดเลือก
31	0.596	0.538	-0.424	1.165	คัดเลือก
32	0.475	0.538	0.089	1.641	คัดเลือก
33	0.152	0.038	2.927	0.614	ตัดทิ้ง
34	0.303	0.038	1.707	0.498	คัดเลือก
35	0.141	0.000	3.487	0.533	ตัดทิ้ง
36	0.525	0.154	-0.223	0.575	คัดเลือก
37	0.192	-0.154	3.556	0.411	คัดเลือก
38	0.081	0.000	4.179	0.604	ตัดทิ้ง
39	0.687	0.231	0.675	-1.292	คัดเลือก
40	0.697	0.308	-1.441	0.637	คัดเลือก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ง
ตารางปรับเทียบคะแนน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางภาคผนวกที่ 9 การปรับเทียบค่าความสามารถ (θ) และค่าคะแนนจริง (ξ) ของนักเรียนกลุ่มปรับเทียบคะแนนระดับ ม.2 ที่มีค่าความสามารถตั้งแต่ 3 ถึง -3 ให้อยู่ในสเกลเดียวกับระดับ ม.1 และ ม.3 โดยใช้วิธีการปรับเทียบแนวตั้งตามวิธีทฤษฎีตอบสนองข้อสอบ แบบโลจิสติก 3 พารามิเตอร์ ร่วมกับการปรับเทียบเชิงเส้นตรง

θ_{12}	ξ_{21}	θ_{2}	ξ_{2}	θ_{23}	ξ_{23}
-2.154	6.71	-3.000	6.51	-3.993	6.48
-2.000	6.52	-2.850	6.51	-3.840	6.48
-1.949	6.56	-2.800	6.51	-3.789	6.48
-1.897	6.61	-2.750	6.51	-3.738	6.48
-1.846	6.66	-2.700	6.51	-3.687	6.48
-1.795	6.71	-2.650	6.51	-3.637	6.48
-1.743	6.77	-2.600	6.51	-3.586	6.48
-1.692	6.82	-2.550	6.51	-3.535	6.48
-1.641	6.88	-2.500	6.51	-3.484	6.48
-1.589	6.95	-2.450	6.51	-3.433	6.48
-1.538	7.01	-2.400	6.51	-3.382	6.48
-1.487	7.08	-2.350	6.51	-3.331	6.48
-1.435	7.15	-2.300	6.51	-3.280	6.48
-1.384	7.22	-2.250	6.51	-3.229	6.48
-1.332	7.30	-2.200	6.51	-3.178	6.48
-1.281	7.37	-2.150	6.51	-3.128	6.48
-1.230	7.45	-2.100	6.51	-3.077	6.48
-1.127	7.62	-2.000	6.51	-2.975	6.48
-1.076	7.71	-1.950	6.51	-2.924	6.48
-1.024	7.81	-1.900	6.51	-2.873	6.48
-0.973	7.91	-1.850	6.51	-2.822	6.48
-0.921	8.02	-1.800	6.51	-2.771	6.48
-0.870	8.13	-1.750	6.51	-2.720	6.48
-0.819	8.24	-1.700	6.51	-2.669	6.48
-0.767	8.37	-1.650	6.51	-2.618	6.48
-0.716	8.50	-1.600	6.51	-2.568	6.48
-0.665	8.63	-1.550	6.51	-2.517	6.48
-0.613	8.77	-1.500	6.51	-2.466	6.48

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

θ_{12}	ξ_{21}	θ_{2}	ξ_{2}	θ_{23}	ξ_{23}
-0.562	8.92	-1.450	6.52	-2.415	6.48
-0.511	9.08	-1.400	6.52	-2.364	6.48
-0.459	9.24	-1.350	6.52	-2.313	6.48
-0.408	9.41	-1.300	6.52	-2.262	6.48
-0.356	9.59	-1.250	6.53	-2.211	6.48
-0.305	9.78	-1.200	6.53	-2.160	6.48
-0.254	9.97	-1.150	6.53	-2.109	6.48
-0.202	10.17	-1.100	6.54	-2.059	6.48
-0.151	10.39	-1.050	6.54	-2.008	6.48
-0.100	10.61	-1.000	6.55	-1.957	6.48
-0.048	10.84	-0.950	6.55	-1.906	6.48
0.003	11.08	-0.900	6.56	-1.855	6.48
0.055	11.34	-0.850	6.57	-1.804	6.49
0.106	11.61	-0.800	6.58	-1.753	6.49
0.157	11.90	-0.750	6.59	-1.702	6.49
0.209	12.20	-0.700	6.61	-1.651	6.49
0.260	12.52	-0.650	6.62	-1.600	6.49
0.311	12.87	-0.600	6.64	-1.550	6.49
0.363	13.23	-0.550	6.67	-1.499	6.49
0.414	13.61	-0.500	6.69	-1.448	6.50
0.465	14.01	-0.450	6.73	-1.397	6.50
0.517	14.44	-0.400	6.77	-1.346	6.50
0.568	14.88	-0.350	6.81	-1.295	6.51
0.620	15.33	-0.300	6.86	-1.244	6.51
0.671	15.81	-0.250	6.93	-1.193	6.52
0.722	16.29	-0.200	7.00	-1.142	6.52
0.774	16.78	-0.150	7.09	-1.091	6.53
0.825	17.28	-0.100	7.19	-1.041	6.54
0.876	17.79	-0.050	7.31	-0.990	6.55
0.928	18.30	0.000	7.45	-0.939	6.56
0.979	18.80	0.050	7.61	-0.888	6.57
1.030	19.31	0.100	7.80	-0.837	6.58

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

θ_{12}	ξ_{21}	θ_{2}	ξ_{2}	θ_{23}	ξ_{23}
1.082	19.82	0.150	8.03	-0.786	6.60
1.133	20.32	0.200	8.29	-0.735	6.62
1.185	20.82	0.250	8.60	-0.684	6.64
1.236	21.32	0.300	8.97	-0.633	6.67
1.287	21.81	0.350	9.40	-0.582	6.70
1.339	22.29	0.400	9.92	-0.531	6.74
1.390	22.77	0.450	10.52	-0.481	6.78
1.441	23.25	0.500	11.22	-0.430	6.83
1.493	23.72	0.550	12.03	-0.379	6.89
1.544	24.17	0.600	12.93	-0.328	6.96
1.595	24.62	0.650	13.90	-0.277	7.04
1.647	25.05	0.700	14.91	-0.226	7.13
1.698	25.46	0.750	15.95	-0.175	7.24
1.750	25.84	0.800	17.00	-0.124	7.37
1.801	26.20	0.850	18.05	-0.073	7.51
1.852	26.54	0.900	19.12	-0.022	7.67
1.904	26.84	0.950	20.20	0.029	7.86
1.955	27.12	1.000	21.26	0.079	8.07
2.006	27.37	1.050	22.28	0.130	8.30
2.058	27.60	1.100	23.25	0.181	8.55
2.109	27.80	1.150	24.13	0.232	8.83
2.161	27.99	1.200	24.91	0.283	9.14
2.212	28.16	1.250	25.58	0.334	9.46
2.263	28.31	1.300	26.16	0.385	9.80
2.315	28.45	1.350	26.64	0.436	10.17
2.366	28.57	1.400	27.04	0.487	10.56
2.417	28.69	1.450	27.38	0.538	10.97
2.469	28.79	1.500	27.67	0.588	11.42
2.520	28.89	1.550	27.91	0.639	11.90
2.571	28.98	1.600	28.12	0.690	12.43
2.623	29.06	1.650	28.30	0.741	13.02
2.674	29.13	1.700	28.46	0.792	13.70

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

θ_{12}	ξ_{21}	θ_{2}	ξ_{2}	θ_{23}	ξ_{23}
2.726	29.20	1.750	28.61	0.843	14.46
2.777	29.26	1.800	28.74	0.894	15.29
2.828	29.32	1.850	28.86	0.945	16.17
2.880	29.37	1.900	28.97	0.996	17.02
2.931	29.42	1.950	29.07	1.047	17.83
2.982	29.46	2.000	29.16	1.098	18.59
3.034	29.50	2.050	29.25	1.148	19.30
3.085	29.54	2.100	29.32	1.199	19.98
3.136	29.58	2.150	29.39	1.250	20.65
3.188	29.61	2.200	29.45	1.301	21.31
3.239	29.64	2.250	29.51	1.352	21.95
3.291	29.67	2.300	29.56	1.403	22.56
3.342	29.69	2.350	29.61	1.454	23.14
3.393	29.72	2.400	29.65	1.505	23.69
3.455	29.74	2.460	29.69	1.566	24.30
3.496	29.76	2.500	29.72	1.607	24.69
3.547	29.78	2.550	29.75	1.657	25.17
3.599	29.79	2.600	29.77	1.708	25.64
3.650	29.81	2.650	29.80	1.759	26.10
3.701	29.82	2.700	29.82	1.810	26.52
3.753	29.84	2.750	29.84	1.861	26.90
3.804	29.85	2.800	29.85	1.912	27.24
4.010	29.89	3.000	29.91	2.116	28.19

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก จ
การวิเคราะห์ด้วยโปรแกรม BILOG 3.04

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การวิเคราะห์ด้วยโปรแกรม BILOG 3.04

ตัวอย่าง File คำสั่ง

```
FILE: TEST1
RUN 3PLM
>GLO DFN='C:\TEST1.DAT', NPA=3, LOG, SAV;
>SAV SCO=' C:\TEST1.SCO', PAR=' C:\TEST1.PAR', COV= 'C:\FOCALT1.COV';
>LEN NIT=30;
>INP NTO=30, SAM=543, NAL=4, NID=4;
      (4A1,30A1)
>TES ITE=(1(1)30);
>CAL FLOAT;
>SCO MET=2,INF=2,POP,FIT;
```

เมื่อ	DFN='C:\TEST1.DAT'	หมายถึง	File ข้อมูลดิบ
	NPA=3	หมายถึง	จำนวนพารามิเตอร์
	COV= 'C:\FOCALT1.COV'	หมายถึง	File ข้อมูล variance-covarience Matrices
	NIT=30;	หมายถึง	จำนวนข้อสอบ
	NID=4	หมายถึง	จำนวน ID
	4A1	หมายถึง	จำนวน ID
	30A1	หมายถึง	จำนวนข้อสอบที่วิเคราะห์
	FLOAT	หมายถึง	การประมาณค่าพารามิเตอร์
	INF=2	หมายถึง	ให้ระบุรูปแบบของสารสนเทศของแบบสอบโดยรายงานเป็นโค้งสารสนเทศของแบบสอบและตารางค่าสารสนเทศของแบบสอบ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เมตริกซ์ Covariances ของพารามิเตอร์ข้อสอบ จากโปรแกรม BILOG

FILE: TEST1

RUN 3PLM

0001TEST0001	.589852	.210532	.280439	.018909	.030929
	.308742	.006323	.051805	.010087	
0002TEST0001	.877861	1.055319	.133564	.027746	.004625
	.059117	.005318	.009945	.003121	
0003TEST0001	1.350296	.702348	.306797	.100490	.035919
	.055415	.014597	.013298	.004483	
0004TEST0001	1.105837	.542611	.286404	.068122	.040075
	.086145	.014331	.021060	.006529	
0005TEST0001	.903171	-.327426	.186804	.022114	.023372
	.079671	.005175	.019800	.006502	
0006TEST0001	2.240762	1.046755	.177950	.252835	.004806
	.012523	.010566	.002041	.001170	
0007TEST0001	1.395432	1.212232	.211815	.113331	.010627
	.036548	.011942	.006235	.002566	
0008TEST0001	.937932	.092983	.228577	.035644	.031738
	.098333	.009306	.024921	.007649	
0009TEST0001	1.289226	1.079693	.100141	.054047	-.000818
	.023098	.005453	.003452	.001570	

.

.

.

0027TEST0001	1.126430	1.748906	.212995	.104968	-.013724
	.072803	.011501	.005745	.002512	
0028TEST0001	2.493582	.811204	.175978	.214351	.009135
	.010094	.008664	.002148	.001179	
0029TEST0001	1.460496	.999720	.214993	.111989	.016321
	.031475	.012091	.006598	.002734	
0030TEST0001	2.260920	1.573165	.171055	.307811	-.008352
	.015330	.007364	.000998	.000615	

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ฉ
การวิเคราะห์ด้วยโปรแกรม EQUATE version 2.0

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การวิเคราะห์ด้วยโปรแกรม EQUATE version 2.0

คุณสมบัติของโปรแกรม

1. ใช้วิเคราะห์ปรับเทียบพารามิเตอร์ของข้อสอบและผู้สอบภายใต้ ทฤษฎี IRT โมเดล 1,2 และ 3 พารามิเตอร์

2. ใช้ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบและผู้สอบ (a,b,c และ θ) จากโปรแกรม BILOG

ตัวอย่าง File คำสั่ง และ Printout

EQUATE TEST2 TO TEST1
NUMBER OF ABILITY SCALE POINTS= 21
ICC MODEL HAS 3 PARAMETERS

FROM METRIC ITEM PARAMETER FILE NAME IS
C:\BILOG\FOCALT2.COV
FILE FORMAT IS (12X,3F12.6,24X/)
NUMBER OF ITEM IN “FROM” TEST IS 30
FROM TEST IS LOGISTIC OGIVE METRIC

TO METRIC ITEM PARAMETER FILE NAME IS
C:\BILOG\FOCALT1.COV
FILE FORMAT IS (12X,3F12.6,24X/)
NUMBER OF ITEM IN “TO” TEST IS 30
TO TEST IS LOGISTIC OGIVE METRIC

TRANSFORMED ITEM PARAMETER FILE NAME IS
EQUATE21

TRANSFORMED PARAMETERS WILL BE IN
LOGISTIC OGIVE METRIC

ANCHOR ITEM IDs FOR “FORM” INSTRUMENT ARE

1-6:

ANCHOR ITEM IDs FOR “TO” INSTRUMENT ARE

25-30:

INITIAL VALUE FOR A= 1.1304 INITIAL VALUE FOR K= 1.4930

FUNCTION AT INITIAL VALUES = 0.056937

METRIC TRANSFORMATION COEFFICIENTS ARE

A= 1.027397 K= 0.921714

FUNCTION VALUE = 0.024238

SUMMARY STATISTIC FOR TRANSFORMED ITEMS

MEAN B = 1.183 VARIANCE B = 10.304 STD DEV B = 3.210

MEAN A = 1.533 VARIANCE A = 0.858 STD DEV A = 0.925

MEAN C = 0.592 VARIANCE C = 0.176 STD DEV C = 0.419

LABORATORY OF EXPERIMENTAL DESIGN

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

UNIVERSITY OF WISCONSIN

PROGRAM DATED 17 JUNE 1991

เมื่อ 12X หมายถึง อ่านข้อมูลข้ามไป 12 คอลัมภ์

3F12.6 หมายถึง ให้อ่านข้อมูล 12 คอลัมภ์ 3 ครั้ง เป็นตัวเลขทศนิยม 6 ตำแหน่ง

24X/ หมายถึง ให้อ่านข้อมูลข้ามไป 24 คอลัมภ์และ 1 บรรทัด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวพัชรี จันทร์เพ็ง เกิดเมื่อวันที่ 9 พฤษภาคม พ.ศ. 2523 ที่อำเภอเมือง จังหวัดสกลนคร สำเร็จการศึกษาปริญญาศึกษาศาสตรบัณฑิต (เกียรตินิยมอันดับ 2) สาขาการมัธยมศึกษา วิชาเอกคณิตศาสตร์ศึกษา จากมหาวิทยาลัยขอนแก่น ในปีการศึกษา 2545 และเข้าศึกษาต่อในหลักสูตรครุศาสตรมหาบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2546

ปัจจุบันได้รับทุนพัฒนาอาจารย์ (UDC) จากภาควิชาประเมินผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น เพื่อเข้าศึกษาต่อในหลักสูตรครุศาสตรดุษฎีบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในภาคการศึกษาต้น ปีการศึกษา 2548

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย