

๖๗๖.12๗

๑๒๔๕๓

๖.3

การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย
ของนักเรียนชั้นมัธยมศึกษาตอนต้น โดยการเก็บคะแนนตามแนวตั้ง
ตามทฤษฎีการตอบสนองข้อสอบ แบบจำลองโลจิสติก สามพารามิเตอร์

12 ส.ค. 2539

ปริญญานิพนธ์
ของ
วรรณดี แสงประทีปทอง

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษาคุษบัณฑิต สาขาการทดสอบและวัดผลการศึกษา
พฤษภาคม 2538
ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย
ของนักเรียนชั้นมัธยมศึกษาตอนต้น โดยการเทียบคะแนนตามแนวตั้ง
ตามทฤษฎีการตอบสนองข้อสอบ แบบจำลองโลจิสติก สามพารามิเตอร์

บทคัดย่อ
ของ
วรรณดี แสงประทีปทอง

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษาคุณวุฒิปดฺพิท สาขาการทดสอบและวัดผลการศึกษา
พฤษภาคม 2538

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 โดยใช้วิธีการเทียบคะแนนตามแนวดิ่ง ตามทฤษฎีการตอบสนองข้อสอบ แบบจำลองโลจิสติก สามพารามิเตอร์ การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนศึกษาทั้งในกลุ่มรวม และในกลุ่มย่อยตามระดับความสามารถของนักเรียน กลุ่มตัวอย่างการวิจัยเป็นนักเรียนชั้น ม.1 ม.2 และ ม.3 ของโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษาในจังหวัดสิงห์บุรี ภาคเรียนที่ 1 ปีการศึกษา 2537 กลุ่มตัวอย่างแบ่งเป็น 2 กลุ่ม คือกลุ่มตัวอย่างเทียบคะแนน เป็นนักเรียนชั้น ม.1 ม.2 และ ม.3 จำนวน 1245 1280 และ 1171 คน ตามลำดับ และกลุ่มตัวอย่างสอบทานผลเป็นนักเรียนชั้น ม.1 ม.2 และ ม.3 จำนวน 373 384 และ 352 คนตามลำดับ เครื่องมือที่ใช้ในการวิจัยประกอบด้วยแบบทดสอบเทียบคะแนน 3 ฉบับซึ่งเป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย ท 101(ม.1) ท 203(ม.2) และ ท 305(ม.3) และแบบทดสอบรวม ซึ่งเป็นแบบทดสอบความสามารถทั่วไปทางภาษาไทย 1 ฉบับ ข้อมูลจากการวิจัยรวบรวมจากสถานการณ์การสอบจริง ของการสอบปลายภาคเรียนที่ 1 ปีการศึกษา 2537 พร้อมกันทุกระดับชั้น การวิเคราะห์ข้อมูลใช้วิธีการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ แบบจำลองโลจิสติก สามพารามิเตอร์ เพื่อเทียบคะแนนจากการสอบของนักเรียนชั้น ม.1 ม.2 และ ม.3 จากแบบทดสอบต่างฉบับให้อยู่ในสเกลเดียวกัน แล้วเปรียบเทียบผลสัมฤทธิ์ทางการเรียนของนักเรียนแต่ละชั้น ภายหลังจากเทียบคะแนนนักเรียนแต่ละชั้นให้อยู่ในสเกลเดียวกันแล้วเพื่อศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียน ผลการวิจัยพบว่า

1. ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.3 สูงกว่า ม.2 แต่ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.2 และ ม.1 ไม่แตกต่างกัน หรือกล่าวได้ว่า ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.2 ไป ม.3 ก้าวหน้ามากกว่าปกติ แต่ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ไป ม.2 ก้าวหน้าตามปกติ

2. ในกลุ่มนักเรียนที่มีความสามารถสูง ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.3 และ ม.2 ไม่แตกต่างกัน และ ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.2 และ ม.1 ไม่แตกต่างกัน หรือกล่าวได้ว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ไป ม.3 ก้าวหน้าตามปกติ และผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ไป ม.2 ก้าวหน้าตามปกติ

3. ในกลุ่มนักเรียนที่มีความสามารถปานกลาง ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.3 สูงกว่า ม.2 และผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.2 สูงกว่า ม.1 หรือกล่าวได้ว่าผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.2 ไป ม.3 ก้าวหน้ามากกว่าปกติ และผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ไป ม.2 ก้าวหน้ามากกว่าปกติ

4. ในกลุ่มนักเรียนที่มีความสามารถต่ำ ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย ของนักเรียนชั้น ม.3 สูงกว่า ม.2 แต่ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.2 และ ม.1 ไม่แตกต่างกัน หรือกล่าวได้ว่าผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.2 ไป ม.3 ก้าวหน้ามากกว่าปกติ แต่ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ไป ม.2 ก้าวหน้าตามปกติ

A PROGRESSIVE STUDY OF THE THAI LANGUAGE ACHIEVEMENT
OF LOWER SECONDARY SCHOOL STUDENTS BY USING
THE THREE - PARAMETER LOGISTIC MODEL TO VERTICAL EQUATING

AN ABSTRACT
BY
VANDEE SANGPRATEEPTONG

Presented in partial fulfillment of the requirements for
the Doctor of Education degree in Testing and Measurement
at Srinakharinwirot University

May 1995

The purpose of this study was to investigate the progress of the students' achievement from Mathayom Suksa 1(M.1) to Mathayom Suksa 3 (M.3) by comparing their equated scores. The samples were consisted of the M.1, M.2 and M.3 students studying in the secondary schools under the jurisdiction of the Department of General Education in Singburi province. The students in each grade level were divided into two groups; the equated and the validated ones, and those in each group were divided into three sub-groups accordingly to their ability, namely, high fair and low. There were altogether 1,245 1,280 and 1,171 students respectively for M.1 M.2 and M.3 equated groups while there were 373 384 and 352 students for the validated ones. The instruments to be used for data collecting were 4 achievement tests, three of which measured the students achievement in Thai language courses, namely, Th.101 (for M.1 students), Th.203 (for M 2 students) and Th.305 (for M.3 students) while the fourth one which was the anchor test, measured general ability in Thai language. All the four tests were administered simultaneously to the students at the end of the first semester of the academic year 1994.

Vertical equating technique, based on item response theory 3 - parameter logistic model, was used to equate the students achievement scores across grade levels so that they were on the same scale. Progress of the students' achievement across grade levels was analysed by comparing their equated scores. The major findings were as follows;

1. The M.3 students' achievement in Thai language was higher than the M.2's while no difference was found between M.1 and M.2 students' achievement. In other words, the students' progress from M.2 to M.3 was higher than normal while that from M.1 to M.2 was normal.

2. Regarding only to the high ability students, no difference were found neither between M.2 and M.3 students' achievement nor between M.1's and M.2's. It can be concluded for the high ability

students that their progress both from M.1 to M.2 and M.2 to M.3 was normal.

3. As for the fair ability students, M.3 students' achievement was higher than M.2's and M.2's higher than M.1's, or it can be said that the progress of the fair ability students from M.1 to M.2 as well as from M.2 to M.3 were higher than normal.

4. Concerning the low ability students, it was found that M.3 students' achievement was higher than M.1's but no difference was found between M.1's and M.2's. The conclusion was that the progress of the low ability students' achievement from M.2 to M.3 was higher than normal but that from M.1 to M.2 was normal.

คณะกรรมการควบคุมและคณะกรรมการสอบ ได้พิจารณาปริญญานิพนธ์ฉบับนี้แล้ว
เห็นสมควรเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาการศึกษาดุขบัณฑิต สาขาการทดสอบ
และวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒได้

คณะกรรมการควบคุม

.....รณ. 31/5/๐๗.....ประธาน
(อ.ดร.วัน สังข์สะอาด)
.....กรรมการ
(อ.ดร.ชูศักดิ์ ขัมภลิจิต)
.....กรรมการ
(รศ.ดร.ศิริชัย กาญจนวาสี)

คณะกรรมการสอบ

.....รณ. 31/5/๐๗.....ประธาน
(อ.ดร.วัน สังข์สะอาด)
.....กรรมการ
(อ.ดร.ชูศักดิ์ ขัมภลิจิต)
.....กรรมการ
(รศ.ดร.ศิริชัย กาญจนวาสี)
.....กรรมการที่แต่งตั้งเพิ่มเติม
(รศ.ดร.ดิเรก ศรีสุโข)

บัณฑิตวิทยาลัยอนุมัติให้รับปริญญฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษาดุขบัณฑิต สาขาการทดสอบและวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒ

.....คณบดีบัณฑิตวิทยาลัย
(ดร.ศิริยุภา พูลสุวรรณ)
วันที่ 11...เดือน พฤษภาคม...พ.ศ. 2538

ผู้วิจัยสำนึกในพระมหากรุณาธิคุณเป็นล้นเกล้าล้นกระหม่อมหาที่สุดมิได้

ที่ได้พระราชทาน "ทุนภูมิพล" เพื่อสนับสนุนการทำปริญญานิพนธ์นี้

ประกาศคุณูปการ

ในการทำปริญญานิพนธ์ฉบับนี้ ผู้วิจัยได้แนวคิดในการกำหนดปัญหาการวิจัยและแนวทางในการดำเนินการวิจัยจาก รศ.ดร.ศิริชัย กาญจนวาสี และ อ.ดร.ชูศักดิ์ ขัมภลิจิต และได้รับความรู้และคำแนะนำต่าง ๆ ด้วยความกรุณาอย่างยิ่งจากคณะกรรมการที่ปรึกษาปริญญานิพนธ์ อ.ดร.วัน สังข์สะอาด อ.ดร.ชูศักดิ์ ขัมภลิจิต และ รศ.ดร.ศิริชัย กาญจนวาสี ทำให้ปริญญานิพนธ์ฉบับนี้สำเร็จด้วยดี ผู้วิจัยขอขอบพระคุณเป็นอย่างสูงไว้ ณ ที่นี้

ขอขอบพระคุณ รศ.ดร.ติเรก ศรีสุโข กรรมการสอบปริญญานิพนธ์ที่ช่วยให้ข้อคิดเห็นต่าง ๆ อันทำให้ปริญญานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณ ผู้อำนวยการสำนักศึกษาจังหวัดสิงห์บุรี ผู้อำนวยการ อาจารย์ใหญ่ ผู้ช่วยผู้บริหารฝ่ายวิชาการ และคณาจารย์ของโรงเรียนมัธยมศึกษาที่เป็นกลุ่มตัวอย่างการวิจัย และขอขอบใจนักเรียนทุกคนที่ให้ความร่วมมือเป็นอย่างดี

ขอขอบคุณ คุณวัฒนา แสงประทีปทอง อ.นพชัย สุวรรณะ อ.วราภรณ์ อรัญपाल และ อ.จุฑาทิพย์ เเคราะห์ดี ที่ช่วยประสานงานการเก็บรวบรวมข้อมูล อ.สมพร บุญอ้อม อ.บังอร กมลวัฒนา ที่ช่วยวิเคราะห์ข้อมูล รศ.กานดา พุนลาภทวี ที่ช่วยเหลืองานทุกเรื่อง คุณนิรมล แสงประทีปทอง และ คุณเอื้องพร อมรหิรัญ ที่ช่วยจัดพิมพ์ปริญญานิพนธ์ อ.อนุวัติ คุณแก้ว ตลอดจนเพื่อน ๆ ที่มีจากกล่าวนามได้หมด ที่คอยช่วยเหลือ และให้กำลังใจในการทำปริญญานิพนธ์นี้

ขอขอบพระคุณ ครู อาจารย์ ที่ได้อบรม สั่งสอน ผู้วิจัย จนมีโอกาสสำเร็จการศึกษาขึ้นนี้

ขอขอบพระคุณ พ่อ แม่ และญาติพี่น้อง ที่สนับสนุน ช่วยเหลือ และเป็นกำลังใจในการศึกษาตลอดมา

ท้ายที่สุด ผู้วิจัยขอกล่าวว่า

"พระคุณท่านทั้งหลายมากมายนัก สุดที่จักแทนทดให้หมดได้
อันประโยชน์ที่เกิดจากการวิจัย ขอมอบให้แทนคำบอกขอบพระคุณ"

วรรณดี แสงประทีปทอง

สารบัญ

บทที่	หน้า
1	บทนำ..... 1 ภูมิหลัง..... 1 จุดมุ่งหมายของการวิจัย..... 7 สมมติฐานของการวิจัย..... 8 ความสำคัญของการวิจัย..... 8 ขอบเขตของการวิจัย..... 9 ข้อตกลงเบื้องต้นของการวิจัย..... 9 ความจำกัดของการวิจัย..... 9 นิยามศัพท์เฉพาะ..... 9
2	เอกสารและงานวิจัยที่เกี่ยวข้อง..... 12 แนวคิดทั่วไปเกี่ยวกับการเทียบคะแนน..... 12 การเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ..... 20 งานวิจัยที่เกี่ยวข้องกับทฤษฎีการตอบสนองข้อสอบ..... 29
3	วิธีดำเนินการวิจัย..... 54 ประชากร..... 54 กลุ่มตัวอย่าง..... 54 เครื่องมือที่ใช้ในการวิจัย..... 57 การเก็บรวบรวมข้อมูล..... 59 การวิเคราะห์ข้อมูล..... 61
4	ผลการวิเคราะห์ข้อมูล..... 68 ตอนที่ 1 ผลการวิเคราะห์แบบทดสอบ..... 68 ตอนที่ 2 ผลการสร้างตารางเทียบคะแนนแบบทดสอบ..... 100 ตอนที่ 3 ผลการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียน วิชาภาษาไทยของนักเรียนชั้นมัธยมศึกษาตอนต้น..... 108

ตอนที่ 4 ผลการศึกษาประสิทธิภาพของวิธีการเทียบคะแนน.....	120
5 สรุป อภิปรายผลการวิจัยและข้อเสนอแนะ.....	123
สรุปผลการวิจัย.....	126
อภิปรายผลการวิจัย.....	130
ข้อเสนอแนะ.....	133
บรรณานุกรม.....	135
ภาคผนวก.....	140
ภาคผนวก ก.....	141
ภาคผนวก ข.....	166
ประวัติของผู้วิจัย.....	178

บัญชีตาราง

ตาราง		หน้า
1	ประเด็นการวิจัยเกี่ยวกับการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ.....	31
2	จำนวนประชากรการวิจัย จำแนกตามโรงเรียน.....	55
3	จำนวนกลุ่มตัวอย่างการวิจัย จำแนกตามโรงเรียน.....	56
4	โครงสร้างของแบบทดสอบความสามารถทั่วไปทางภาษาไทย.....	58
5	จำนวนนักเรียนที่ใช้ในการทดลองสอบแบบทดสอบเทียบคะแนนและแบบทดสอบรวม	59
6	จำนวนข้อสอบและจำนวนกลุ่มตัวอย่างผู้สอบแบบทดสอบแต่ละฉบับ.....	60
7	เกณฑ์การประเมินดัชนีความแตกต่าง.....	67
8	ค่าไอเกนขององค์ประกอบที่มีค่าเกิน 1.0 ของแบบทดสอบชั้น ม.1.....	68
9	ค่าไอเกนขององค์ประกอบที่มีค่าเกิน 1.0 ของแบบทดสอบชั้น ม.2.....	71
10	ค่าไอเกนขององค์ประกอบที่มีค่าเกิน 1.0 ของแบบทดสอบชั้น ม.3.....	73
11	ค่าไอเกนขององค์ประกอบที่มีค่าเกิน 1.0 ของแบบทดสอบความสามารถทั่วไป ทางภาษาไทย.....	76
12	ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน และแบบทดสอบความสามารถทั่วไปทางภาษาไทย.....	78
13	ค่าประมาณพารามิเตอร์ของข้อสอบในแบบทดสอบชั้น ม.1.....	79
14	ค่าประมาณพารามิเตอร์ของข้อสอบในแบบทดสอบชั้น ม.2.....	85
15	ค่าประมาณพารามิเตอร์ของข้อสอบในแบบทดสอบชั้น ม.3.....	90
16	ค่าประมาณพารามิเตอร์ของข้อสอบในแบบทดสอบความสามารถทั่วไปทาง ภาษาไทย.....	95
17	ค่าสถิติพื้นฐานของแบบทดสอบฉบับต่าง ๆ ก่อนการเทียบคะแนน.....	99
18	สมการที่ใช้ในการแปลงค่าพารามิเตอร์ของข้อสอบและพารามิเตอร์ของผู้สอบ...	100
19	คะแนนสมมูลจากการเทียบคะแนนแบบทดสอบชั้น ม.1 ม.2 และ ม.3.....	101
20	คะแนนสมมูลจากการเทียบคะแนนแบบทดสอบชั้น ม.1 กับ แบบทดสอบ ความสามารถทั่วไปทางภาษาไทย.....	104
21	คะแนนสมมูลจากการเทียบคะแนนแบบทดสอบชั้น ม.2 กับแบบทดสอบความ สามารถทั่วไปทางภาษาไทย.....	105

22	คะแนนสมมูลจากการเทียบคะแนนแบบทดสอบชั้น ม.3 กับแบบทดสอบความสามารถทั่วไปทางภาษาไทย.....	107
23	ค่าสถิติพื้นฐานของผลการสอบของนักเรียนชั้น ม.1 ม.2 และ ม.3 ภายหลังการเทียบคะแนน.....	109
24	ผลการวิเคราะห์ความแปรปรวนของคะแนนของนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มรวม.....	110
25	ผลการทดสอบความแตกต่างระหว่างคะแนนเฉลี่ยของนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มรวม.....	110
26	ผลการวิเคราะห์แนวโน้มของการเปลี่ยนแปลงคะแนนเฉลี่ยของนักเรียน ในกลุ่มรวม.....	111
27	ค่าสถิติพื้นฐานของผลการสอบของนักเรียนจำแนกตามชั้นและระดับความสามารถ	113
28	ผลการวิเคราะห์ความแปรปรวนของคะแนนของนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มนักเรียนที่มีความสามารถสูง.....	114
29	ผลการทดสอบความแตกต่างระหว่างคะแนนเฉลี่ยของคะแนนนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มนักเรียนที่มีความสามารถสูง.....	114
30	ผลการวิเคราะห์ความแปรปรวนของคะแนนของนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มนักเรียนที่มีความสามารถปานกลาง.....	115
31	ผลการทดสอบความแตกต่างระหว่างคะแนนเฉลี่ยของคะแนนนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มนักเรียนที่มีความสามารถปานกลาง.....	115
32	ผลการวิเคราะห์ความแปรปรวนของคะแนนของนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มนักเรียนที่มีความสามารถต่ำ.....	116
33	ผลการทดสอบความแตกต่างระหว่างคะแนนเฉลี่ยของคะแนนนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มนักเรียนที่มีความสามารถต่ำ.....	117
34	ผลการวิเคราะห์แนวโน้มของการเปลี่ยนแปลงคะแนนของนักเรียนในกลุ่มย่อย...	118
35	ค่าสถิติพื้นฐานภายหลังการเทียบคะแนนของกลุ่มสอบทานผล.....	121
36	การคำนวณดัชนีความแตกต่าง.....	122
37	ค่าฟังก์ชันสารสนเทศของแบบทดสอบและค่าฟังก์ชันลักษณะข้อสอบของ แบบทดสอบชั้น ม.1.....	142

38	ค่าฟังก์ชันสารสนเทศของแบบทดสอบและค่าฟังก์ชันลักษณะข้อสอบของ แบบทดสอบชั้น ม.1.....	148
39	ค่าฟังก์ชันสารสนเทศของแบบทดสอบและค่าฟังก์ชันลักษณะข้อสอบของ แบบทดสอบชั้น ม.1.....	154
40	ค่าฟังก์ชันสารสนเทศของแบบทดสอบและค่าฟังก์ชันลักษณะข้อสอบของ แบบทดสอบชั้น ม.1.....	160
41	น้ำหนักองค์ประกอบของแบบทดสอบชั้น ม.1.....	167
42	น้ำหนักองค์ประกอบของแบบทดสอบชั้น ม.2.....	170
43	น้ำหนักองค์ประกอบของแบบทดสอบชั้น ม.3.....	175

บัญชีภาพประกอบ

ภาพประกอบ		หน้า
1	สถานการณ์ของการสอบและระบบอย่างง่ายของการเทียบคะแนน..	4
2	สถานการณ์ของการวิจัย.....	7
3	แบบแผนการเก็บรวบรวมข้อมูลเพื่อเทียบคะแนนแบบทดสอบ.....	16
4	ค่าไอเกนขององค์ประกอบ 4 ค่าแรกของแบบทดสอบชั้น ม.1.....	70
5	ค่าไอเกนขององค์ประกอบ 4 ค่าแรกของแบบทดสอบชั้น ม.2.....	72
6	ค่าไอเกนขององค์ประกอบ 4 ค่าแรกของแบบทดสอบชั้น ม.3.....	75
7	ค่าไอเกนขององค์ประกอบ 4 ค่าแรกของแบบทดสอบความสามารถทั่วไป ทางภาษาไทย.....	77
8	ไค้่งสารสนเทศของแบบทดสอบและไค้่งลักษณะข้อสอบของแบบทดสอบชั้น ม.1...	84
9	ไค้่งสารสนเทศของแบบทดสอบและไค้่งลักษณะข้อสอบของแบบทดสอบชั้น ม.2...	89
10	ไค้่งสารสนเทศของแบบทดสอบและไค้่งลักษณะข้อสอบของแบบทดสอบชั้น ม.3...	94
11	ไค้่งสารสนเทศของแบบทดสอบและไค้่งลักษณะข้อสอบของแบบทดสอบ ความสามารถทั่วไปทางภาษาไทย.....	97
12	ไค้่งสารสนเทศของแบบทดสอบและไค้่งลักษณะข้อสอบของแบบทดสอบฉบับต่าง ๆ	98
13	แนวโน้มของการเปลี่ยนแปลงคะแนนเฉลี่ยของนักเรียนในกลุ่มรวม.....	112
14	แนวโน้มของการเปลี่ยนแปลงคะแนนเฉลี่ยของนักเรียนในกลุ่มย่อย.....	119

บทที่ 1

บทนำ

ภูมิหลัง

องค์ประกอบของการศึกษา ประกอบด้วยส่วนที่สำคัญ 3 ส่วน ได้แก่ จุดมุ่งหมาย กิจกรรมการเรียนการสอนและการประเมินผล การจัดหลักสูตรการศึกษาจะจัดกิจกรรมการเรียนการสอนเพื่อให้บรรลุจุดมุ่งหมายที่กำหนดไว้ และมีการประเมินผลการจัดดำเนินการว่าบรรลุจุดมุ่งหมายหรือไม่ เพียงใดเพื่อนำข้อมูลที่ได้มาใช้เป็นแนวทางการพัฒนาหลักสูตรให้มีประสิทธิภาพยิ่งขึ้น

ในการจัดการศึกษา เป้าหมายและจุดมุ่งหมายของหลักสูตรทุก ๆ ระดับมุ่งไปยังผลที่เกิดขึ้นในตัวผู้เรียน การจัดกิจกรรมการเรียนการสอนและการจัดกิจกรรมบริหารหลักสูตรด้านต่าง ๆ อันเป็นขั้นตอนของการใช้หลักสูตรก็ล้วนแต่มุ่งไปยังผลที่เกิดขึ้นแก่ตัวผู้เรียนทั้งสิ้น ผลที่เกิดขึ้นที่ตัวผู้เรียนก็คือความสำเร็จของผู้เรียน (วัลลภ กันทรพย. 2528 : 60-61) ดังนั้นการประเมินผลการจัดการศึกษาจึงประเมินได้จากผลที่เกิดขึ้นกับผู้เรียนดังกล่าว กระทรวงศึกษาธิการเป็นหน่วยงานหลักที่มีหน้าที่รับผิดชอบการจัดการศึกษาระดับประถมศึกษาและมัธยมศึกษา ในปี พ.ศ. 2521 กระทรวงศึกษาธิการประกาศใช้หลักสูตรประถมศึกษา พุทธศักราช 2521 และหลักสูตรมัธยมศึกษาตอนปลาย พุทธศักราช 2521 และในปี พ.ศ. 2524 ประกาศใช้หลักสูตรมัธยมศึกษาตอนปลาย พุทธศักราช 2524 ต่อมหลักสูตรที่ประกาศใช้นี้ กระทรวงศึกษาธิการมอบให้สถานศึกษาเป็นผู้ดำเนินการวัดผลประเมินผลการเรียน ตลอดจนตัดสินผลการเรียน โดยกำหนดให้กรมวิชาการ และหน่วยงานเจ้าสังกัดดำเนินการตรวจสอบเป็นระยะ ๆ เพื่อศึกษาและพัฒนาคุณภาพการเรียนให้มีระดับสูงขึ้น

การประเมินคุณภาพการเรียนของนักเรียนนั้น มีหน่วยงานเจ้าสังกัดและหน่วยงานที่เกี่ยวข้อง ดำเนินการหลายหน่วยงาน ได้แก่ กรมวิชาการ หน่วยงานนิเทศก์ กรมสามัญศึกษา กองวิชาการ สำนักงานคณะกรรมการ การประถมศึกษาแห่งชาติ กองวิจัยการศึกษา สำนักงานคณะกรรมการการศึกษาแห่งชาติ นอกจากนี้ยังมีการประเมินผลสัมฤทธิ์ทางการเรียนของนักเรียนในรูปแบบงานวิจัยของอาจารย์และนิสิตนักศึกษา ในสถาบันการศึกษาต่าง ๆ

กรมวิชาการได้ประเมินผลการใช้หลักสูตรประถมศึกษา และมัธยมศึกษา ตั้งแต่ปี พ.ศ. 2526 โดยดำเนินการตรวจสอบคุณภาพการศึกษาในลักษณะต่อเนื่อง เพื่อศึกษาผลโดยรวมในระดับชาติ เขตการศึกษา ภาคภูมิศาสตร์ และสังกัดต่าง ๆ เพื่อศึกษาผลการพัฒนาการศึกษา ตลอดจนศึกษาการบรรลุเป้าหมายของหลักสูตร โดยดำเนินการตรวจสอบระดับชั้นตัวประโยคในระบะแรกประเมินเฉพาะผลสัมฤทธิ์ทางการเรียนและต่อมาประเมินกระบวนการจัดการศึกษาของ

โรงเรียนด้วย เครื่องมือที่ใช้ในการประเมินเป็นแบบทดสอบและแบบประเมินซึ่งสร้างขึ้นใน
แต่ละครั้งที่ประเมิน งานประเมินการใช้หลักสูตรดังกล่าว กรมวิชาการดำเนินการต่อเนื่องมาจน
ถึงปัจจุบัน

หน่วยศึกษานิเทศ กรมสามัญศึกษาได้ประเมินคุณภาพการศึกษาโรงเรียนมัธยมศึกษา
ซึ่งเป็นการดำเนินงานตามนโยบายของกรมสามัญศึกษาตามโครงการพัฒนาการเรียนการสอนและ
การประเมินผลตามหลักสูตร โดยได้มีการศึกษาเชิงประเมินคุณภาพการศึกษาระดับมัธยมศึกษา
เพื่อนำข้อมูลมาประกอบการตัดสินใจวางแผนพัฒนาเป็นระยะ ๆ อย่างต่อเนื่อง การประเมิน
ดังกล่าวดำเนินการมาตั้งแต่ปีงบประมาณ 2531 โดยประเมินคุณภาพการศึกษาระดับชั้นมัธยมศึกษา
ปีที่ 1 ซึ่งมีการประเมินผลสัมฤทธิ์ทางการเรียน การวัดความถนัดทางการเรียนและการประเมิน
พฤติกรรมการสอนของครู ในปีงบประมาณ 2532 ได้ประเมินคุณภาพในระดับชั้นมัธยมศึกษาปีที่ 2
และ 5 โดยมีการประเมินผลสัมฤทธิ์ทางการเรียนการประเมินบรรยากาศและสิ่งแวดล้อมของ
โรงเรียน และประเมินพฤติกรรมการสอนของครู ในปีงบประมาณ 2533 ได้ประเมินคุณภาพผล
การจัดการเรียนการสอนโดยไม่มีการประเมินผลสัมฤทธิ์ทางการเรียนและในปีงบประมาณ 2534
ได้ประเมินคุณภาพในระดับชั้นมัธยมศึกษาปีที่ 3 และ 4 โดยมีการประเมินผลสัมฤทธิ์ทางการเรียน
รายวิชาสามัญของแต่ละระดับชั้น

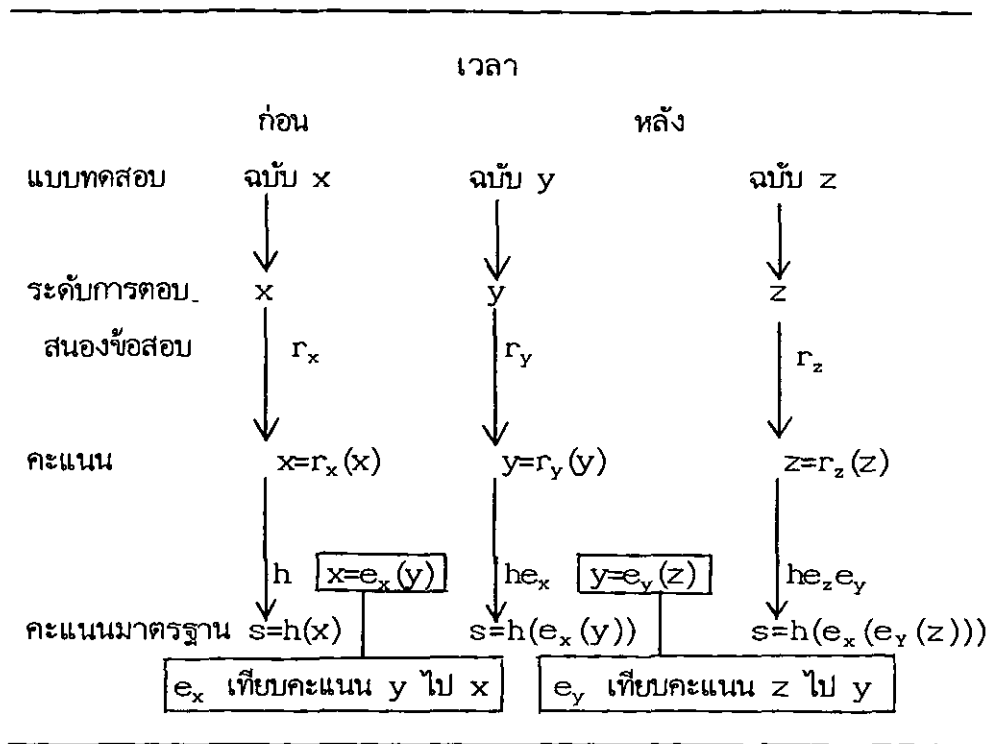
สำนักงานการประถมศึกษาแห่งชาติได้ประเมินคุณภาพนักเรียนชั้นประถมศึกษาโดยมี
จุดมุ่งหมายสำคัญเพื่อให้ได้ข้อมูลเกี่ยวกับคุณภาพของการศึกษาว่าผลผลิตของการจัดการศึกษาตาม
หลักสูตร การประเมินพิจารณาจากผลสัมฤทธิ์ทางการเรียนของนักเรียนที่เรียนจบตามหลักสูตรใน
ปีการศึกษาต่าง ๆ ว่ามีคุณภาพและมีแนวโน้มที่เพิ่มขึ้นหรือลดลงเพียงใด เครื่องมือที่ใช้คือแบบ
ทดสอบซึ่งสร้างขึ้นใหม่ทุกปีการศึกษา ตามขอบเขตสมรรถภาพที่ต้องการวัดซึ่งกำหนดไว้เหมือนกัน
ทุกปี เพื่อให้เป็นแนวทางในการประเมินในปีการศึกษาต่อไปและเพื่อให้สามารถเปรียบเทียบผล
ระหว่างปีการศึกษาต่าง ๆ ได้ และเมื่อวิเคราะห์ผลการสอบแล้ว จะเปรียบเทียบความก้าวหน้า
ของผลสัมฤทธิ์ทางการเรียนระหว่างปีการศึกษาต่าง ๆ ได้แก่ ช่วง 1 ปี 2 ปี และ 3 ปี
การศึกษาที่ผ่านมาเพื่อจัดทำข้อมูลแสดงให้เห็นแนวโน้มของผลการศึกษาของแต่ละจังหวัด และ
ส่วนรวมทั้งประเทศ

นอกจากการประเมินผลสัมฤทธิ์ทางการเรียนของนักเรียน โดยหน่วยงานต้นสังกัดของ
โรงเรียน ได้แก่ กรมวิชาการ หน่วยศึกษานิเทศก์ กรมสามัญศึกษา สำนักงานคณะกรรมการ
การประถมศึกษาแห่งชาติ ยังมีการประเมินจากหน่วยงานนอกสังกัดของโรงเรียนได้แก่สำนักงาน
คณะกรรมการการศึกษาแห่งชาติซึ่งเป็นการศึกษาวิจัยที่พิจารณาประสิทธิภาพของการศึกษาทั้งระบบ
มีการศึกษาสภาพทั่วไปของการมัธยมศึกษา ความเสมอภาคในการจัดสรรทรัพยากรทางการศึกษา
สัมฤทธิ์ผลทางการเรียนซึ่งยึดถือจุดมุ่งหมายของหลักสูตรเป็นเกณฑ์วัด ตลอดจนความสัมพันธ์

ระหว่างด้านต่าง ๆ ทางการศึกษา และด้านอื่น ๆ ที่เกี่ยวข้องกับผลสัมฤทธิ์ทางการเรียนของนักเรียน โครงการนี้ศึกษาวิจัยในระดับประเทศ โดยใช้กลุ่มตัวอย่างเป็นโรงเรียนมัธยมศึกษาทั้งในส่วนกลางและส่วนภูมิภาค ทั้งนี้เพื่อที่จะนำผลการวิจัยดังกล่าวไปใช้ในการกำหนดนโยบาย และจัดทำแผนการศึกษาในระดับมัธยมศึกษาให้มีประสิทธิภาพ กว้างขวางและทั่วถึงยิ่งขึ้น (สำนักงานและกรมการศึกษาแห่งชาติ. 2530)

จากการตรวจสอบงานวิจัยที่ผ่านมาพบว่าการศึกษาลักษณะผลสัมฤทธิ์ทางการเรียนของนักเรียนระดับชั้นต่าง ๆ จะศึกษา ณ ช่วงเวลาใดช่วงเวลาหนึ่งที่ทำวิจัยและในกรณีที่ต้องการศึกษาความเปลี่ยนแปลง หรือความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนจะนำผลสัมฤทธิ์ทางการเรียน ซึ่งได้จากการวิจัยในช่วงเวลาที่ต่างกันของปีการศึกษาต่าง ๆ โดยใช้กลุ่มตัวอย่างการวิจัยที่ต่างกัน และใช้เครื่องมือวัดที่ต่างกันมาเปรียบเทียบกันโดยตรง โดยทั่วไปคะแนนที่ได้จากเครื่องมือวัดที่ต่างกันมิได้แทนระดับความสามารถเดียวกัน ดังนั้นการนำคะแนนที่ได้ไปเปรียบเทียบโดยตรง จึงทำให้ผลการเปรียบเทียบขาดความแม่นยำและเกิดความคลาดเคลื่อนสูง ผู้วิจัยในฐานะที่เป็นบุคลากรด้านการวัดผลการศึกษา จึงมีความสนใจที่จะแก้ไขปัญหการเปรียบเทียบผลการเรียนของบุคคลต่างกลุ่มที่ได้จากเครื่องมือวัดที่ต่างกัน เพื่อให้ผลการเรียนดังกล่าวสามารถนำมาเปรียบเทียบและศึกษาความก้าวหน้าในการเรียนได้ การเปรียบเทียบคะแนนระหว่างแบบทดสอบต่างฉบับซึ่งใช้วัดบุคคลต่างกลุ่มนี้สามารถทำได้ โดยใช้วิธีการเทียบคะแนนระหว่างแบบทดสอบต่างฉบับที่ได้จากนักเรียนต่างกลุ่มก่อน แล้วจึงนำผลที่ได้มาเปรียบเทียบกัน

การเทียบคะแนนเป็นกระบวนการแปลงคะแนนของแบบทดสอบฉบับหนึ่งไปสู่คะแนนของแบบทดสอบอีกฉบับหนึ่งซึ่งวัดคุณลักษณะเดียวกันโดยการสร้างความสัมพันธ์ระหว่างคะแนนของแบบทดสอบ 2 ฉบับ เพื่อให้คะแนนที่แปลงแล้วของแบบทดสอบต่างฉบับนำมาเปรียบเทียบกันได้ การเทียบคะแนนเป็นกระบวนการที่เกี่ยวข้องกับกิจกรรม 2 ประการ (ชูศักดิ์ ชัมภลิจิต. 2527: 2) คือกระบวนการที่ทำให้แบบทดสอบ 2 ฉบับ มีความเท่าเทียมกันในเชิงโครงสร้างและการใช้วิธีการทางสถิติเพื่อปรับคะแนนที่ได้จากการสอบแต่ละครั้งให้อยู่ในมาตราเดียวกันและเปรียบเทียบกันได้ สถานการณ์ของการสอบและระบอบอย่างง่ายของการเทียบคะแนนระหว่างแบบทดสอบต่างฉบับ แสดงดังภาพประกอบ 1 (Braun and Holland. 1982 : 11-13)



ภาพประกอบ 1 สถานการณ์ของการสอบและระบบอย่างง่ายของการเทียบคะแนน

จากภาพประกอบ 1 แบบทดสอบ X Y และ Z นำไปใช้สอบในสถานการณ์ที่มีติของเวลาเปลี่ยนไป และต้องการเปรียบเทียบคะแนนความสามารถของบุคคลซึ่งได้รับการสอบโดยแบบทดสอบต่างฉบับกัน การเทียบคะแนนแบบทดสอบต่างฉบับทำได้โดยการแปลงคะแนนให้เป็นคะแนนมาตรฐาน แล้วนำคะแนนมาตรฐานมาเปรียบเทียบกัน

การเทียบคะแนนแบ่งได้เป็น 2 ประเภท คือ (Hambleton and Swaminathan, 1985:197)

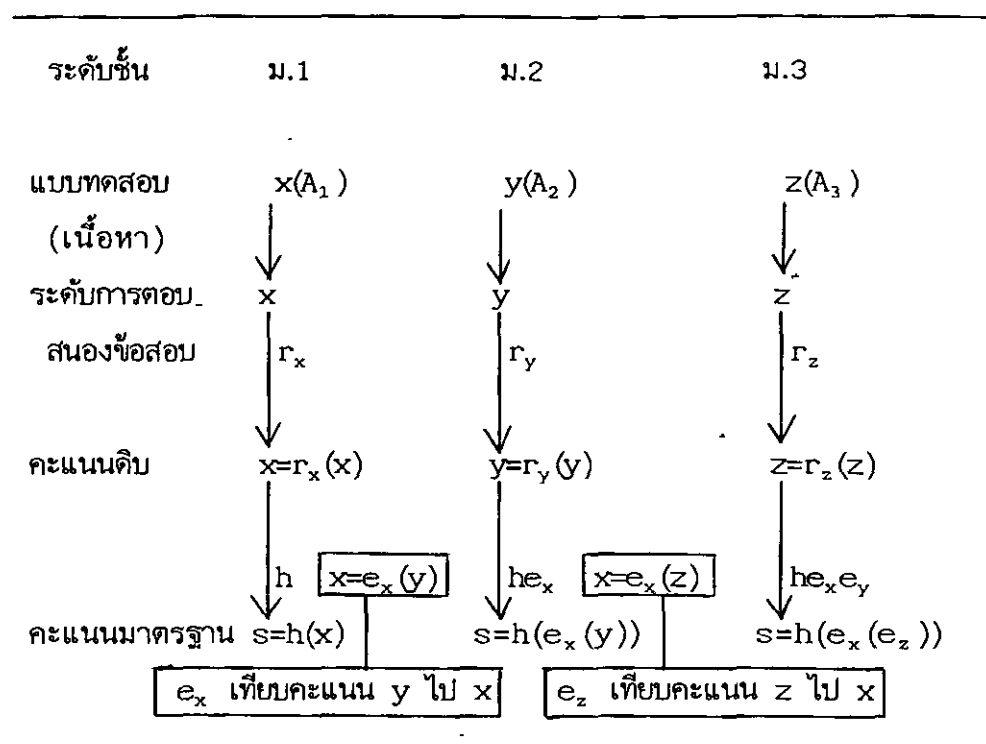
1. การเทียบคะแนนตามแนวนอน (Horizontal Equating) เป็นการเทียบคะแนนในกรณีที่แบบทดสอบต่างฉบับที่นำมาเปรียบเทียบกัน มีระดับความยากใกล้เคียงกัน และกลุ่มผู้สอบมีความสามารถใกล้เคียงกัน

2. การเทียบคะแนนตามแนวตั้ง (Vertical Equating) เป็นการเทียบในกรณีที่แบบทดสอบต่างฉบับที่นำมาเปรียบเทียบกันมีระดับความยากต่างกันและกลุ่มผู้สอบมีความสามารถต่างกัน

วิธีการเทียบคะแนนที่นิยมใช้โดยทั่วไปมี 3 วิธีได้แก่วิธีการเทียบคะแนนอิกวิเปอร์เซ็นไทล์ วิธีการเทียบคะแนนเชิงเส้นตรงและวิธีการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ สำหรับวิธีการ 2 วิธีแรก เป็นวิธีการตามทฤษฎีการสอบแบบดั้งเดิม ส่วนวิธีที่ 3 เป็นวิธีการตามทฤษฎีการตอบสนองข้อสอบ ทฤษฎีการตอบสนองข้อสอบมีข้อดีเหนือกว่าวิธีการตามทฤษฎีการสอบแบบ

ดั้งเดิมในแง่ที่ว่าผลการเทียบคะแนนไม่ขึ้นกับระดับความสามารถของกลุ่มที่นำมาเปรียบเทียบและไม่ได้ได้รับความกระทบกระเทือนจากแบบทดสอบที่มีความยากต่างกัน ดังนั้นผู้วิจัยจึงศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนโดยใช้วิธีการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ สำหรับการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ นิยมใช้แบบจำลองโลจิสติก 1 พารามิเตอร์ และแบบจำลองโลจิสติก 3 พารามิเตอร์ แต่จากการศึกษาวิจัยที่ผ่านมา มีข้อสรุปที่ชัดเจนว่าการเทียบคะแนนตามแนวดังด้วยแบบจำลองโลจิสติก 1 พารามิเตอร์หรือแบบจำลองของราซส์ เป็นวิธีการที่ไม่เพียงพอ (Cook and Eignor. 1989 : 161 173; Gary Skaggs and others. 1986:495-529) ดังนั้นในการวิจัยครั้งนี้ ผู้วิจัยจึงใช้วิธีการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ แบบจำลองโลจิสติก 3 พารามิเตอร์ โดยจะศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนระดับชั้นมัธยมศึกษาตอนต้น โดยมีจุดมุ่งหมายเพื่อนำวิธีการเทียบคะแนนตามแนวดังตามทฤษฎีการตอบสนองข้อสอบมาใช้เพื่อศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียน วิธีการดังกล่าวสามารถนำไปใช้ในการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนในระดับชั้นอื่นหรือเนื้อหาวิชาอื่น และสามารถนำผลการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้นมัธยมศึกษาตอนต้น ที่ได้ไปใช้เป็นข้อมูลย้อนกลับในการวางแผนการเรียนการสอนให้มีประสิทธิภาพมากขึ้น สำหรับเหตุผลที่ศึกษาวิชาภาษาไทยนั้น เนื่องจากภาษาไทยมีความสำคัญในฐานะเป็นภาษาประจำชาติ แผนการศึกษาแห่งชาติ พุทธศักราช 2535 ได้กำหนดแนวนโยบายการศึกษาเกี่ยวกับวิชาภาษาไทยไว้อย่างชัดเจนว่า "ส่งเสริมให้ผู้เรียนทุกระดับรวมทั้งประชาชนโดยทั่วไปมีความรู้ความสามารถในการใช้ภาษาไทยได้อย่างถูกต้องและเหมาะสมเพื่อการสื่อสารเพื่อสร้างสรรค์และพัฒนาความคิดและองค์ความรู้และเพื่อธำรงรักษาวัฒนธรรมชาติ" (สำนักงานคณะกรรมการการศึกษาแห่งชาติ. 2535 : 20) หลักสูตรการศึกษาระดับมัธยมศึกษากำหนดให้วิชาภาษาไทยเป็นวิชาบังคับที่ทุกคนต้องเรียน ภาษาไทยเป็นภาษาประจำชาติ ถือเป็นวัฒนธรรมของชาติ รวมทั้งเป็นเครื่องมือสื่อสารของคนในชาติ ดังนั้นนักเรียนทุกคนจึงควรใช้ภาษาไทยอย่างถูกต้อง เหมาะสมกับวัยและศักยภาพ จุดประสงค์ของหลักสูตรมุ่งเน้นเรื่องการใช้ภาษาเป็นเครื่องมือในการติดต่อสื่อสารได้อย่างมีประสิทธิภาพ มุ่งพัฒนาทักษะทางภาษาตามวัยและตามศักยภาพของผู้เรียนแต่ละคน (ปราชญ์ บุญชุ่ม. 2537 : 260) โดยในชั้นมัธยมศึกษาปีที่ 1 มีจุดมุ่งหมายเพื่อพัฒนาความสามารถในการใช้ภาษาชั้นมัธยมศึกษาปีที่ 2 มีจุดมุ่งหมายเพื่อให้สามารถใช้ภาษาได้อย่างถูกต้องเหมาะสม และชั้นมัธยมศึกษาปีที่ 3 มีจุดมุ่งหมายเพื่อให้สามารถใช้ภาษาสื่อสารได้อย่างมีประสิทธิภาพ (กระทรวงศึกษาธิการ. 2533 : 15-16) ผู้วิจัยจึงสนใจศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยในระดับชั้นมัธยมศึกษาตอนต้นดังกล่าว

สถานการณ์ของการวิจัยครั้งนี้เป็นการศึกษาเปรียบเทียบความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ชั้นมัธยมศึกษาปีที่ 2 และชั้นมัธยมศึกษาปีที่ 3 โดยนักเรียน 3 ชั้นเป็นนักเรียนในภาคการศึกษาและปีการศึกษาเดียวกัน นักเรียนแต่ละชั้นสอบแบบทดสอบตามชั้นของตน คือ ม.1 สอบแบบทดสอบ X ม.2 สอบแบบทดสอบ Y และ ม.3 สอบแบบทดสอบ Z แบบทดสอบทั้ง 3 ฉบับ เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนภาษาไทย (A) ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง 2533) เช่นเดียวกัน แต่มีระดับความยากแตกต่างกันตามระดับชั้น (A_1 , A_2 , A_3) ผลการวัดผลสัมฤทธิ์ทางการเรียนภาษาไทยของนักเรียน 3 ชั้น โดยใช้เครื่องมือวัดที่ต่างกัน 3 ฉบับนี้สามารถนำมาเปรียบเทียบกันได้โดยให้นักเรียนแต่ละคนทำแบบทดสอบรวมซึ่งเป็นแบบทดสอบวัดความสามารถทั่วไปทางภาษาไทยอีกฉบับหนึ่ง ข้อมูลที่ได้จากแบบทดสอบรวมซึ่งมีข้อมูลของนักเรียนทุกคนนี้ จะนำมาใช้ในการปรับความแตกต่างระหว่างความสามารถของนักเรียนต่างระดับชั้น และปรับค่าความยากของข้อสอบจากแบบทดสอบต่างฉบับกัน เพื่อทำให้คะแนนของนักเรียน 3 ชั้นที่ได้จากการตอบแบบทดสอบตามชั้นของตนมีสเกลการวัดเดียวกัน หลังจากนั้นคะแนนของนักเรียนชั้น ม.2 ซึ่งทำแบบทดสอบ Y จะถูกแปลงไปสู่คะแนนตามสเกลของแบบทดสอบ X ซึ่งเป็นแบบทดสอบของนักเรียนชั้น ม.1 และในทำนองเดียวกันคะแนนของนักเรียนชั้น ม.3 ซึ่งทำแบบทดสอบ Z จะถูกแปลงไปสู่คะแนนตามสเกลของแบบทดสอบ X ดังนั้นผลการวัดที่ได้จากนักเรียน 3 ชั้น ซึ่งตอบแบบทดสอบที่ต่างกัน 3 ฉบับ จึงนำมาเปรียบเทียบกันเพื่อศึกษาความก้าวหน้าของผล การเรียนได้ สถานการณ์ของการวิจัยแสดงดังภาพประกอบ 2



ภาพประกอบ 2 สถานการณ์ของการวิจัย

จุดมุ่งหมายของการวิจัย

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 โดยใช้วิธีการเทียบคะแนนตามแนวตั้งตาม ทฤษฎีการตอบสนองข้อสอบ แบบจำลองโลจิสติก สามพารามิเตอร์เป็นแนวทางในการศึกษา โดยมีจุดมุ่งหมายเฉพาะดังนี้

1. เพื่อศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 โดยเปรียบเทียบจากผลสัมฤทธิ์ทางการเรียนของนักเรียนแต่ละชั้น ภายหลังจากเทียบคะแนนนักเรียนแต่ละชั้นให้อยู่ในสเกลเดียวกันแล้ว
2. เพื่อศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 ในแต่ละกลุ่มความสามารถ คือกลุ่มความสามารถสูง กลุ่มความสามารถปานกลาง และ กลุ่มความสามารถต่ำ โดยเปรียบเทียบจากผลสัมฤทธิ์ทางการเรียนของนักเรียนแต่ละชั้นภายหลังจากเทียบคะแนนนักเรียนแต่ละชั้นให้อยู่ในสเกลเดียวกันแล้ว

สมมติฐานของการวิจัย

1. ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 มีความก้าวหน้าขึ้นตามลำดับ โดยผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1

2. ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มนักเรียนที่มีความสามารถสูง มีความก้าวหน้าขึ้นตามลำดับ โดยผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1

3. ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มนักเรียนที่มีความสามารถปานกลางมีความก้าวหน้าขึ้นตามลำดับ โดยผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1

4. ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มนักเรียนที่มีความสามารถต่ำมีความก้าวหน้าขึ้นตามลำดับ โดยผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1

ความสำคัญของการวิจัย

1. ได้ข้อมูลความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้นมัธยมศึกษาตอนต้น ซึ่งสามารถนำไปใช้ในการวางแผนการเรียนการสอนให้มีประสิทธิภาพยิ่งขึ้น

2. ได้ทราบผลการเปรียบเทียบความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้นมัธยมศึกษาตอนต้น จังหวัดสิงห์บุรี

3. ได้แนวทางการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนของนักเรียนต่างระดับชั้น ซึ่งสามารถนำไปประยุกต์ใช้ในการเปรียบเทียบความสามารถของนักเรียนในเนื้อหาวิชาอื่นหรือในระดับชั้นอื่น

4. ได้ตารางการเปรียบเทียบความสามารถทางภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3

5. ได้สารสนเทศเกี่ยวกับแบบทดสอบ ซึ่งจะ เป็นแนวทางในการพัฒนาแบบทดสอบให้มีประสิทธิภาพยิ่งขึ้น

ขอบเขตของการวิจัย

1. การเทียบคะแนนในการวิจัยนี้ เป็นการเทียบคะแนนในแนวตั้งของนักเรียน 3 ชั้น คือ ชั้น ม.1 ม.2 และ ม.3 ที่ได้จากแบบทดสอบที่ต่างกัน 3 ฉบับ วิธีการเทียบคะแนนใช้วิธีการเทียบคะแนนจริงตามทฤษฎีการตอบสนองข้อสอบ แบบจำลองโลจิสติก ที่มีพารามิเตอร์ 3 ตัว

2. การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย ในการวิจัยนี้ ศึกษาจากคะแนนเฉลี่ยของผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ม.2 และ ม.3 จากวิชาบังคับแกนคือ ท101 ท203 และ ท305 ที่ได้จากแบบทดสอบเลือกตอบ ในการสอบปลายภาคเรียนที่ 1 ปีการศึกษา 2537 การศึกษาในลักษณะดังกล่าวเป็นการศึกษาภาคตัดขวาง (Cross Section Study) โดยมีข้อสมมติว่านักเรียน 3 ชั้น เป็นนักเรียนในรุ่น (Cohort) เดียวกัน

3. การศึกษาประสิทธิภาพของวิธีการเทียบคะแนน ศึกษาประสิทธิภาพของวิธีการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ แบบจำลองโลจิสติกสามพารามิเตอร์ เมื่อนำมาใช้ในการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3

ข้อตกลงเบื้องต้นของการวิจัย

หลักสูตรชั้นมัธยมศึกษาตอนต้นพุทธศักราช 2521 (ฉบับปรับปรุง 2533) จัดโครงสร้างเนื้อหาวิชาภาษาไทย โดยจัดเป็นลำดับต่อเนื่องกัน และสอดคล้องกับพัฒนาการของผู้เรียนระดับชั้นต่าง ๆ

ความจำกัดของการวิจัย

การวิจัยนี้ใช้วิธีการเทียบคะแนนจริง ตามทฤษฎีการตอบสนองข้อสอบ แล้วเทียบคะแนนจริงไปเป็นคะแนนสังเกต ซึ่งตรงกับตำแหน่งเปอร์เซ็นต์ไทล์เดียวกัน เพื่อให้ตารางเทียบคะแนนที่สร้างขึ้น นำไปใช้ได้สะดวกยิ่งขึ้น ดังนั้นผลการเทียบคะแนน อาจมีความคลาดเคลื่อนเพิ่มขึ้นในช่วงการเทียบคะแนนจริงไปเป็นคะแนนสังเกต

นิยามศัพท์เฉพาะ

1. การเทียบคะแนน หมายถึง กระบวนการแปลงคะแนนของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย 3 ฉบับและแบบทดสอบความสามารถทางภาษาไทย ของระดับชั้น ม.1 ม.2 และ ม.3 ให้มีสเกลคะแนนเดียวกันและเปรียบเทียบกันได้

2. การเทียบคะแนนตามแนวตั้ง หมายถึงการเทียบคะแนนแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยที่มีระดับความยากต่างกัน 3 ระดับชั้น คือชั้น ม.1 ม.2 และ ม.3 และนำไปใช้สอบกับกลุ่มผู้สอบที่มีความสามารถต่างกัน 3 ระดับ คือผู้สอบชั้น ม.1 ม.2 และ ม.3

3. คะแนนสมมูล หมายถึงคะแนนจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนของชั้น ม.1 ม.2 และ ม.3 และคะแนนจากแบบทดสอบความสามารถทั่วไปทางภาษาไทย ซึ่งตรงกับคะแนนจริงค่าเดียวกันหรือตรงกับคะแนนจริงที่มีค่าใกล้เคียงกัน

4. ตารางเทียบคะแนน หมายถึง ตารางแสดงการเปรียบเทียบคะแนนที่สมมูลกันจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน ชั้น ม.1 ม.2 และ ม.3

5. ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย หมายถึงคะแนนของนักเรียน แต่ละคนที่ได้จากการตอบแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของระดับชั้น ม.1 ม.2 และ ม.3

6. ความก้าวหน้าของผลสัมฤทธิ์ทางการเรียน หมายถึง ผลสัมฤทธิ์ทางการเรียนของนักเรียนที่เพิ่มขึ้นภายหลังจากเทียบคะแนนของนักเรียนชั้น ม.1 ม.2 และ ม.3 ซึ่งเป็นนักเรียนคนละกลุ่มที่เรียนอยู่ในภาคการศึกษาและปีการศึกษาเดียวกันให้อยู่ในสเกลเดียวกันแล้ว โดยค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 จะสูงกว่าค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 และค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 จะสูงกว่าค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1

7. ดัชนีความแตกต่าง (Discrepancy Index) หมายถึงค่าความคลาดเคลื่อนรวม (Total Error) ในการเทียบคะแนน ซึ่งคำนวณได้จากค่าเฉลี่ยกำลังสองของความแตกต่างระหว่างคะแนนแปลงที่ได้จากแบบทดสอบตามระดับชั้นของผู้สอบ กับคะแนนเกณฑ์ที่ได้จากแบบทดสอบรวม แล้วถ่วงน้ำหนักให้เป็นค่ามาตรฐานด้วยความแปรปรวนของคะแนนเกณฑ์จากแบบทดสอบรวม ดัชนีนี้ได้รับการวิเคราะห์กลุ่มสอบทานผล (Validated Group) เป็นค่าที่ใช้ประเมินประสิทธิภาพของวิธีการเทียบคะแนนโดยใช้เกณฑ์การยอมรับของ ปีเตอร์สันและคนอื่น ๆ (Petersen and others. 1982:93-94) ดังนี้

ระดับน่าพอใจอย่างมาก หมายถึง $C \leq (0.05 SD_x)^2$

ระดับน่าพอใจ หมายถึง $(.05 SD_x)^2 < C \leq (.10 SD_x)^2$

ระดับปานกลาง หมายถึง $(.10 SD_x)^2 < C \leq (.15 SD_x)^2$

ระดับไม่น่าพอใจ หมายถึง $(.15 SD_x)^2 < C \leq (.20 SD_x)^2$

ระดับไม่น่าพอใจอย่างยิ่ง หมายถึง $(.20 SD_x)^2 < C$

8. แบบทดสอบเทียบคะแนน หมายถึง แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของระดับชั้น ม.1 ม.2 และ ม.3 ซึ่งใช้ในการสอบปลายภาคเรียนที่ 1 ปีการศึกษา 2537 แบบทดสอบนี้ใช้กับนักเรียนชั้นละ 1 ฉบับ โดยนักเรียนชั้น ม.1 สอบแบบทดสอบวิชา ท 101

นักเรียนชั้น ม.2 สอบแบบทดสอบวิชา ท 203 และนักเรียนชั้น ม.3 สอบแบบทดสอบวิชา ท 305

9. แบบทดสอบรวม หมายถึง แบบทดสอบความสามารถทั่วไปทางภาษาไทย ใช้สอบกับนักเรียนชั้น ม.1 ม.2 และ ม.3 เพื่อนำผลมาใช้ในการปรับค่าพารามิเตอร์ของแบบทดสอบแต่ละชั้นให้อยู่บนสเกลเดียวกัน

10. กลุ่มเทียบคะแนน หมายถึง กลุ่มนักเรียนชั้น ม.1 ม.2 และ ม.3 ที่สอบแบบทดสอบเทียบคะแนนตามชั้นของตน 1 ฉบับและสอบแบบทดสอบรวมอีก 1 ฉบับ เพื่อนำผลมาใช้ในการสร้างตารางเทียบคะแนนและศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ม.2 และ ม.3

11. กลุ่มสอบทานผล หมายถึง กลุ่มนักเรียนชั้น ม.1 ม.2 และ ม.3 ที่สอบแบบทดสอบเทียบคะแนนตามชั้นของตน 1 ฉบับ และสอบแบบทดสอบรวมอีก 1 ฉบับ เพื่อนำผลมาใช้ในการประเมินประสิทธิภาพของวิธีการเทียบคะแนน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้ ผู้วิจัยศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการเทียบคะแนน โดยเน้นการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ ผลการศึกษาเสนอตามลำดับดังนี้

ตอนที่ 1 แนวคิดทั่วไปเกี่ยวกับการเทียบคะแนน

- 1.1 ความหมายของการเทียบคะแนน
- 1.2 เงื่อนไขของการเทียบคะแนน
- 1.3 ประเภทของการเทียบคะแนน
- 1.4 แบบแผนการเก็บรวบรวมข้อมูล
- 1.5 วิธีการที่ใช้ในการเทียบคะแนน
- 1.6 การนำวิธีการเทียบคะแนนไปใช้ทางการศึกษา

ตอนที่ 2 การเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ

- 2.1 แนวคิดของทฤษฎีการตอบสนองข้อสอบ
- 2.2 ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ
- 2.3 แบบจำลองการตอบสนองข้อสอบ
- 2.4 กระบวนการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ

ตอนที่ 3 งานวิจัยที่เกี่ยวข้องกับการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ

ตอนที่ 1 แนวคิดทั่วไปเกี่ยวกับการเทียบคะแนน

ความหมายของการเทียบคะแนน

กัลลิกเสน (Gulliksen. 1950 : 298) ให้ความหมายของการเทียบคะแนนว่าเป็นวิธีการทำคะแนนจากแบบทดสอบ 2 ฉบับของวิชาเดียวกัน ให้เป็นคะแนนสมมูล (Equivalent scores) ที่เปรียบเทียบกันได้โดยตรง

ฟลานาแกน (Flanagan. 1951 : 747-748) ให้ความหมายของการเทียบคะแนนว่าเป็นวิธีการทำคะแนนจากแบบทดสอบต่างฉบับให้สามารถนำมาเปรียบเทียบกันได้

แองกอฟ (Angoff. 1984 : 85) ให้ความหมายของการเทียบคะแนนว่าเป็นการแปลงระบบของหน่วยการวัดของแบบทดสอบฉบับหนึ่งไปสู่ระบบของหน่วยการวัดของแบบทดสอบฉบับอื่น ซึ่งหลังจากการแปลงคะแนนแล้วคะแนนที่ได้จากแบบทดสอบต่างฉบับสามารถนำมาเทียบหาคะแนนที่สมมูลกันได้

ปีเตอร์สันและคนอื่น ๆ (Petersen and others. 1989 : 242) กล่าวว่า การเทียบคะแนนเป็นกระบวนการเชิงประจักษ์เพื่อสร้างความสัมพันธ์ระหว่างคะแนนของแบบทดสอบ 2 ฉบับ ทำให้สามารถแปลงคะแนนของแบบทดสอบฉบับหนึ่ง ไปสู่คะแนนของแบบทดสอบอีกฉบับหนึ่งได้

ชูศักดิ์ ชัมภลิจิต (2527 : 2) กล่าวว่า การเทียบคะแนนเป็นกระบวนการที่เกี่ยวข้องกับกิจกรรม 2 ประการ คือ 1) กระบวนการที่ทำให้แบบทดสอบ 2 ฉบับ มีความเท่าเทียมกันในเชิงโครงสร้าง และ 2) การใช้วิธีการทางสถิติเพื่อปรับคะแนนที่ได้จากการสอบแต่ละฉบับให้อยู่ในมาตราเดียวกัน และเปรียบเทียบกันได้

จากที่กล่าวมาพอจะสรุปได้ว่า การเทียบคะแนนเป็นวิธีการที่ทำให้คะแนนจากแบบทดสอบต่างฉบับซึ่งวัดสิ่งเดียวกัน สามารถนำมาเปรียบเทียบกันได้

เงื่อนไขของการเทียบคะแนน

การเทียบคะแนนมีจุดมุ่งหมายที่จะทำให้คะแนนที่ได้จากแบบทดสอบ 2 ฉบับใด ๆ มีความเท่าเทียมกันมากที่สุดเท่าที่จะเป็นไปได้ แต่เนื่องจากการเทียบคะแนนเป็นกระบวนการเชิงประจักษ์ ซึ่งต้องเกี่ยวข้องกับแบบแผนในการเก็บรวบรวมข้อมูล และวิธีการเทียบคะแนนจากแบบทดสอบฉบับหนึ่งไปยังแบบทดสอบอีกฉบับหนึ่ง ดังนั้นเมื่อพิจารณาจุดมุ่งหมายของการเทียบคะแนนที่ต้องการสร้างคะแนนสมมูลจากแบบทดสอบต่างฉบับกัน จึงมีความจำเป็นที่จะต้องกำหนดเงื่อนไขบางประการในการเทียบคะแนนเลวิน (Holland and Rubin.1982 : 10; citing Levin) กันได้ก็ต่อเมื่อแบบทดสอบทั้งสองนั้นคู่ขนานกันในด้านต่อไปนี้ คือ 1) โครงสร้าง 2) เวลา 3) ชนิดของข้อสอบ 4) รูปแบบ และ 5) เนื้อหาวิชา จะแตกต่างกันก็เพียงความยาก ซึ่งการเทียบคะแนนก็คือการปรับคะแนนที่ได้จากแบบทดสอบ 2 ฉบับใด ๆ ที่มีความยากต่างกัน"

ลอร์ด (Lord. 1980 : 199) ตั้งข้อกำหนดที่สำคัญ 3 ประการของการเทียบคะแนนแบบทดสอบ 2 ฉบับซึ่งวัดคุณลักษณะเดียว (Unidimensionality) ที่ระดับความสามารถเดียวกันไว้ดังนี้

1) ความเสมอภาค (Equity) กล่าวคือ คะแนนที่ผ่านการเทียบแล้วจะมีการแจกแจงเหมือนกับคะแนนที่กลุ่มนั้นทำได้จริง

2) ความไม่แปรผันตามกลุ่ม (Invariance Across Groups) กล่าวคือ คะแนนของกลุ่มต่าง ๆ ที่ผ่านการเทียบแล้วจะมีค่าไม่ต่างกัน ไม่ว่าจะเทียบจากกลุ่มตัวอย่างใด

3) ความสมมาตร (Symmetry) กล่าวคือ ผลของการเทียบคะแนนจะต้องเหมือนกัน ไม่ว่าจะเทียบจากแบบทดสอบฉบับที่ 1 ไปยังแบบทดสอบฉบับที่ 2 หรือเทียบจากแบบทดสอบฉบับที่ 2 ไปยังแบบทดสอบฉบับที่ 1

ประเภทของการเทียบคะแนน

การเทียบคะแนนแบ่งได้ 2 ประเภท คือ

1. การเทียบคะแนนตามแนวนอนเป็นการเทียบคะแนนในกรณีที่แบบทดสอบต่างฉบับที่นำมาเทียบคะแนนมีระดับความยากใกล้เคียงกันและกลุ่มผู้สอบมีความสามารถใกล้เคียงกัน โดยต้องการเปรียบเทียบว่าที่คะแนนของแบบทดสอบฉบับหนึ่งจึงจะเทียบเท่ากับที่คะแนนของแบบทดสอบอีกฉบับหนึ่ง เช่น การเทียบคะแนนแบบทดสอบวิชาเดียวกันในกลุ่มผู้สอบระดับชั้นเดียวกัน เป็นต้น
2. การเทียบคะแนนตามแนวตั้ง เป็นการเทียบคะแนนในกรณีที่แบบทดสอบต่างฉบับที่นำมาเทียบคะแนนมีระดับความยากต่างกันและกลุ่มผู้สอบมีความสามารถต่างกัน โดยต้องการเทียบคะแนนระหว่างแบบทดสอบนั้นตามระดับความสามารถของผู้สอบ เช่น การเทียบคะแนนแบบทดสอบวิชาเดียวกัน ในกลุ่มผู้สอบต่างระดับชั้นกัน เป็นต้น

แบบแผนการเก็บรวบรวมข้อมูล

ข้อมูลจากการสอบโดยใช้แบบทดสอบหลายฉบับและเก็บรวบรวมข้อมูล โดยใช้รูปแบบต่าง ๆ นั้น ข้อมูลที่ได้จากบางรูปแบบไม่สามารถนำมาเทียบคะแนนได้ เช่น ข้อมูลที่ได้จากการสอบโดยใช้แบบทดสอบ 2 ฉบับ กับผู้สอบที่มีความแตกต่างกัน 2 กลุ่ม จะนำคะแนนที่ได้จากแบบทดสอบต่างฉบับมาเปรียบเทียบกันไม่ได้ โดยทั่วไปคะแนนจากแบบทดสอบต่างฉบับที่จะนำมาเปรียบเทียบกันได้ ต้องได้มาจากผู้สอบกลุ่มเดียวกัน หรือผู้สอบกลุ่มเท่าเทียมกันหรือได้จากผู้สอบที่สอบข้อสอบร่วมกันรูปแบบใดรูปแบบหนึ่ง แบบแผนการเก็บรวบรวมข้อมูล ซึ่งสามารถนำคะแนนจากแบบทดสอบต่างฉบับมาเปรียบเทียบกันได้ จำแนกได้ดังนี้ (Petersen and others. 1989:244-246)

1. รูปแบบผู้สอบกลุ่มเดียว (Single Group Design) รูปแบบนี้เป็นรูปแบบที่ง่ายที่สุด การเก็บรวบรวมข้อมูลรูปแบบนี้ใช้กลุ่มตัวอย่างผู้สอบกลุ่มเดียวโดยให้ผู้สอบแต่ละคนทำแบบทดสอบทั้ง 2 ฉบับ (ดูภาพ 3.1) การบริหารการสอบทำโดยให้ผู้สอบทำแบบทดสอบฉบับหนึ่งแล้วตามด้วยแบบทดสอบอีกฉบับหนึ่ง การให้ทำแบบทดสอบทั้ง 2 ฉบับ ทำให้ไม่มีปัญหาในเรื่องความยากของแบบทดสอบต่างฉบับ แต่การสอบครั้งหลัง ผู้สอบอาจได้รับผลกระทบจากการสอบครั้งแรก เนื่องจากการเรียนรู้ ความเมื่อยล้า และการฝึกฝน
2. รูปแบบผู้สอบกลุ่มเดียวที่ได้รับการจัดให้สมดุล (Counterbalanced Random Groups Design) ในทางปฏิบัติ ลำดับของการสอบแบบทดสอบก่อน-หลัง บ่อผลต่อผลการสอบ ดังนั้นรูปแบบนี้จึงปรับมาจากรูปแบบแรกโดยจัดผู้สอบเป็นกลุ่มสุ่ม 2 กลุ่ม แต่ละกลุ่มย่อยได้รับการสอบด้วยแบบทดสอบทั้ง 2 ฉบับ ในลักษณะการจัดลำดับก่อน - หลัง สลับให้เกิด

ความสมดุล (ดูภาพ 3.2) การจัดรูปแบบนี้เพื่อให้การสอบโดยใช้แบบทดสอบทั้ง 2 ฉบับ ได้รับอิทธิพลจากองค์ประกอบในเรื่องการเรียนรู้ ความเมื่อยล้าและการฝึกฝนในลักษณะที่สมดุลกัน

3. รูปแบบผู้สอบกลุ่มเท่าเทียมกัน (Equivalent Groups Design) ในทางปฏิบัติ การให้ผู้สอบทำแบบทดสอบทั้ง 2 ฉบับ อาจจะทำให้ยาก การใช้ผู้สอบกลุ่มเท่าเทียมกันเป็นรูปแบบที่สามารถนำมาใช้ได้ รูปแบบนี้จำเป็นต้องจัดกลุ่มผู้สอบให้มีความคล้ายคลึงกันมากเท่าที่จะทำได้ ผู้สอบในแต่ละกลุ่มจะทำแบบทดสอบฉบับเดียว (ดูภาพ 3.3) ดังนั้นจึงหลีกเลี่ยงปัญหาเรื่องการเรียนรู้ ความเมื่อยล้าและการฝึกฝนจากการสอบครั้งแรกได้ แต่จะไม่มีข้อมูลที่จะนำมาใช้ปรับความแตกต่างของกลุ่มตัวอย่างทั้งสอง วิธีการที่จะลดความแตกต่างระหว่างกลุ่มทำได้โดยใช้กลุ่มตัวอย่างขนาดใหญ่

4. รูปแบบผู้สอบกลุ่มสุ่มสองกลุ่มโดยมีแบบทดสอบร่วม (Anchor Test Random Groups Design) การใช้รูปแบบผู้สอบกลุ่มเท่าเทียมกันอาจมีข้อจำกัดในเรื่องความแตกต่างระหว่างกลุ่มตัวอย่าง การใช้แบบทดสอบร่วมด้วยทำให้สามารถปรับความแตกต่างระหว่างกลุ่มได้ตามรูปแบบนี้กลุ่มสุ่มแต่ละกลุ่มจะได้รับการสอบแบบทดสอบกลุ่มละ 1 ฉบับและทั้ง 2 กลุ่มจะได้รับแบบทดสอบร่วมด้วย (ดูภาพ 3.4) แบบทดสอบร่วมนี้อาจจัดไว้รวมกับแบบทดสอบที่ต้องการเทียบคะแนนซึ่งเรียกว่า แบบทดสอบร่วมภายใน (Internal anchor test) หรือจัดแยกไว้จากแบบทดสอบฉบับที่ต้องการเทียบคะแนนก็ได้ซึ่งเรียกว่าแบบทดสอบร่วมภายนอก (External anchor test)

5. รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกันโดยมีแบบทดสอบร่วม (Anchor Test Nonequivalent Groups Design) ในการสอบบางครั้งจำเป็นต้องใช้แบบทดสอบฉบับใหม่กับผู้สอบกลุ่มใหม่ในโปรแกรมการสอบเดียวกัน การสอบโดยใช้แบบทดสอบต่างฉบับและเวลาสอบต่างกันนี้ ผู้สอบอาจไม่ได้มาจากกลุ่มประชากรเดียวกัน จึงจำเป็นต้องใช้แบบทดสอบร่วมในการปรับความแตกต่างระหว่างกลุ่ม ตามรูปแบบนี้ผู้สอบในแต่ละกลุ่มจะได้รับแบบทดสอบ 1 ฉบับและแบบทดสอบร่วม (ดูภาพ 3.5) ซึ่งต้องมีเนื้อหาและระดับความยากคล้ายกับแบบทดสอบที่ต้องการเทียบคะแนนให้มากที่สุดเท่าที่จะเป็นไปได้ คะแนนจากแบบทดสอบร่วมจะใช้เพื่อลดความลำเอียงในการเทียบคะแนนจากกลุ่มที่ต่างกัน

3.1 รูปแบบผู้สอบกลุ่มเดียว

กลุ่มตัวอย่าง	แบบทดสอบ	
	ฉบับที่ 1	ฉบับที่ 2
P ₁	X	X

3.2 รูปแบบผู้สอบกลุ่มเดียวที่ได้รับการจัดให้สมดุล

กลุ่มตัวอย่าง	แบบทดสอบ			
	ฉบับที่ 1		ฉบับที่ 2	
	สอบครั้งที่ 1	สอบครั้งที่ 2	สอบครั้งที่ 1	สอบครั้งที่ 2
P ₁	X			X
P ₂		X	X	

3.3 รูปแบบผู้สอบกลุ่มเท่าเทียมกัน

กลุ่มตัวอย่าง	แบบทดสอบ	
	ฉบับที่ 1	ฉบับที่ 2
P ₁	X	
P ₂		X

3.4 รูปแบบผู้สอบกลุ่มผสมสองกลุ่มโดยมีแบบทดสอบร่วม

กลุ่มตัวอย่าง	แบบทดสอบ		แบบทดสอบร่วม
	ฉบับที่ 1	ฉบับที่ 2	
P ₁	X		X
P ₂		X	X

3.5 รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกันโดยมีแบบทดสอบร่วม

กลุ่มตัวอย่าง	แบบทดสอบ		แบบทดสอบร่วม
	ฉบับที่ 1	ฉบับที่ 2	
P ₁	X		X
Q ₁		X	X

ภาพประกอบ 3 แบบแผนการเก็บรวบรวมข้อมูลเพื่อเทียบคะแนนแบบทดสอบ

วิธีการเทียบคะแนน

วิธีการที่ใช้ในการเทียบคะแนนมีหลายวิธี วิธีการเทียบคะแนนที่เป็นพื้นฐานมี 3 วิธี ได้แก่

1. วิธีเทียบคะแนนอีควิเปอร์เซ็นไทล์
2. วิธีเทียบคะแนนเชิงเส้นตรง
3. วิธีเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ

วิธีการเทียบคะแนนแต่ละวิธี กล่าวโดยสรุปได้ดังนี้

1. วิธีเทียบคะแนนอีควิเปอร์เซ็นไทล์ การเทียบคะแนนตามวิธีการนี้มีนิยามว่า สำหรับกลุ่มผู้สอบที่กำหนด x คะแนนที่ได้จากแบบทดสอบต่างฉบับกันจะถือว่าเทียบเท่ากันถ้าต่างก็ตรงกับคะแนน ณ ตำแหน่งเปอร์เซ็นต์ไทล์เดียวกัน กระบวนการเทียบคะแนนอีควิเปอร์เซ็นไทล์ประกอบด้วย 2 ขั้นตอน ได้แก่

- 1) การจัดทำตารางแจกแจงความถี่สะสมของแบบทดสอบ 2 ฉบับที่จะนำมาเทียบกัน
- 2) การเทียบคะแนนของแบบทดสอบ 2 ฉบับจากการแจกแจงความถี่สะสมของคะแนน

ที่ได้

ในทางปฏิบัติ การดำเนินการในขั้นตอนที่ 2 - อาจมีปัญหาเนื่องจากอาจไม่มีคะแนนของแบบทดสอบฉบับหนึ่งที่ตรงกับคะแนนจากแบบทดสอบอีกฉบับหนึ่ง ณ ตำแหน่งเปอร์เซ็นต์ไทล์เดียวกัน ดังนั้นเพื่อให้สามารถเทียบคะแนนด้วยวิธีอีควิเปอร์เซ็นไทล์ได้ จึงจำเป็นต้องปรับการแจกแจงคะแนนของแบบทดสอบให้มีลักษณะต่อเนื่อง

2. การเทียบคะแนนเชิงเส้นตรง การเทียบคะแนนตามวิธีการนี้มีนิยามว่าสำหรับกลุ่มผู้สอบที่กำหนด คะแนนที่ได้จากแบบทดสอบต่างฉบับกันจะถือว่าเทียบเท่ากันถ้าต่างก็ตรงกับคะแนนมาตรฐานค่าเดียวกัน

วิธีเทียบคะแนนเชิงเส้นตรง ประกอบด้วย 2 ขั้นตอน ได้แก่

- 1) การคำนวณค่าเฉลี่ย (M) และค่าเบี่ยงเบนมาตรฐาน (S) ของคะแนนที่ได้จากแบบทดสอบทั้ง 2 ฉบับของกลุ่มตัวอย่าง
- 2) การเทียบคะแนนของแบบทดสอบฉบับหนึ่งไปยังอีกฉบับหนึ่งตามสมการแปลงเชิงเส้นตรง ดังนี้

$$\frac{y - M_y}{S_y} = \frac{x - M_x}{S_x}$$

เมื่อ x และ y คือคะแนนที่ได้จากแบบทดสอบฉบับ X และฉบับ Y ตามลำดับ

M_x และ M_y คือค่าเฉลี่ยของการแจกแจงคะแนนที่ได้จากแบบทดสอบฉบับ X และฉบับ Y ตามลำดับ

S_x และ S_y คือค่าเบี่ยงเบนมาตรฐานของการแจกแจงคะแนน จากแบบทดสอบฉบับ X และฉบับ Y ตามลำดับ

การเทียบคะแนนเชิงเส้นตรงสามารถนำไปใช้กับข้อมูลการสอบที่ได้จากแบบแผนการเก็บข้อมูลรูปแบบต่าง ๆ ดังที่กล่าวแล้ว ซึ่งทำให้เกิดรูปแบบเฉพาะของวิธีการได้แก่แบบจำลองของทักเกอร์และแบบจำลองของเลวิน เป็นต้น

3. วิธีเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ การเทียบคะแนนตามวิธีนี้นี้อาศัยรูปแบบความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้ตอบข้อสอบใด ๆ ว่าขึ้นอยู่กับความสามารถของผู้สอบซึ่งประมาณค่าได้จากโค้งลักษณะข้อสอบ (Item Characteristic Curve) ของแต่ละรูปแบบที่ใช้ ซึ่งอาจจะเป็นแบบ หนึ่ง สอง หรือสามพารามิเตอร์ โดยไม่ขึ้นกับการแจกแจงคะแนนของกลุ่มตัวอย่าง (สุจินดา ผ่องอักษร. 2533 : 17 - 28 อ้างจาก Lord and Novick. 1968) การเทียบคะแนนโดยใช้ทฤษฎีการตอบสนองข้อสอบมีหลักการว่าคะแนนจากแบบทดสอบต่างฉบับที่วัดในสิ่งเดียวกันจะถือว่าเทียบเท่ากันได้ถ้าผู้สอบทำแบบทดสอบต่างฉบับแล้วได้คะแนนจริง (True scores) หรือความสามารถแท้ (Ability) เท่ากัน (สงบ ลักษณะ. 2525 : 22) การเทียบคะแนนตามทฤษฎีนี้แบ่งเป็น 2 แบบ ได้แก่ 1) การเทียบคะแนนโดยใช้คะแนนจริง (True Score Equating) และการเทียบคะแนนโดยใช้คะแนนสังเกต (Observed Score Equating) รูปแบบการเทียบคะแนนทั้งสองนี้มีข้อได้เปรียบกัน คือ วิธีเทียบคะแนนโดยใช้คะแนนจริงไม่สามารถอธิบายคะแนนที่อยู่ต่ำกว่าระดับการเดาได้ โดยจะให้ความหมายของคะแนนสมมูลเฉพาะคะแนนที่อยู่เหนือค่าเฉลี่ยของการเดา ซึ่งแม้จะเป็นการเทียบคะแนนโดยใช้คะแนนจริง แต่ก็ยังเป็นค่าประมาณที่ได้จากสูตรการคำนวณ ดังนั้นจึงยังคงมีความคลาดเคลื่อนอยู่ . ส่วนการเทียบคะแนนโดยใช้คะแนนสังเกตเป็นการเทียบคะแนนโดยประมาณแต่สามารถอธิบายคะแนนสมมูลจากแบบทดสอบทั้ง 2 ฉบับ ได้ครอบคลุมพิสัยของคะแนนที่สังเกตได้ วิธีการทั้งสองนี้มีความสอดคล้องกันมากแต่การสรุปผล เพื่อนำไปใช้อ้างอิงต้องทำอย่างพิถีพิถัน อย่างไรก็ตามการเทียบคะแนนโดยใช้คะแนนสังเกตนี้นำไปใช้ในการวิจัยน้อยมาก ซึ่งอาจเป็นเพราะว่าวิธีนี้มีความซับซ้อนและการลงทุนสูงกว่าวิธีการเทียบคะแนนโดยใช้คะแนนจริง (Lord. 1980 : 202 - 203; Lord and Wingersky. 1984 : 453)

การนำวิธีการเทียบคะแนนไปใช้ทางการศึกษา

ในทางการศึกษามีการนำวิธีการเทียบคะแนนไปใช้ในการเปรียบเทียบคะแนนแบบทดสอบต่างฉบับที่วัดคุณลักษณะเดียวกันและใช้ในการศึกษาความก้าวหน้าในการเรียน ดังนี้

1. การเปรียบเทียบคะแนนจากแบบทดสอบต่างฉบับ

เนื่องจากการสอบวิชาเดียวกันอาจมีการสอบหลาย ๆ ครั้ง โดยใช้แบบทดสอบต่างฉบับกัน คะแนนที่ได้จากแบบทดสอบแต่ละฉบับจะมีหน่วยการวัดที่ต่างกัน การรายงานเปรียบเทียบผลการสอบในโปรแกรมการสอบดังกล่าว จึงจำเป็นต้องมีการทำให้ผลการวัดจากแบบทดสอบต่างฉบับอยู่ในสเกลเดียวกันก่อน ซึ่งทำได้ด้วยวิธีการเทียบคะแนน เช่น การเปรียบเทียบความสามารถทางภาษาไทยของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ของนักเรียนในจังหวัดต่าง ๆ โดยใช้แบบทดสอบที่ต่างกันในแต่ละจังหวัด ทำได้โดยการแปลงคะแนนจากแบบทดสอบแต่ละฉบับให้มีหน่วยการวัดเดียวกันแล้วจึงนำคะแนนที่ได้มาเปรียบเทียบกัน ดังเช่นการวิจัยของสุรัตน์ ขวัญบุญจันทร์ (2531) ซึ่งสร้างตารางเทียบคะแนนแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์เรื่องพลังงานและการเปลี่ยนแปลง สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2 หรือต้องการเปรียบเทียบผลการศึกษานักเรียนระดับชั้นเดียวกันที่อยู่ต่างปีการศึกษา ซึ่งในกรณีนี้นักเรียนที่อยู่ต่างปีการศึกษาคือคนต่างกลุ่มกันและแบบทดสอบที่ใช้ก็ต่างฉบับกัน การเปรียบเทียบผลการสอบดังกล่าว ทำได้โดยการเทียบคะแนนจากแบบทดสอบแต่ละฉบับให้อยู่ในสเกลเดียวกันแล้วจึงนำคะแนนที่เทียบแล้วมาเปรียบเทียบกัน

การเปรียบเทียบคะแนนจากแบบทดสอบต่างฉบับในอีกลักษณะหนึ่งเป็นการเทียบคะแนนแบบทดสอบต่างฉบับที่มีความยากต่างกัน และนำไปใช้สอบกับผู้ที่มีระดับความสามารถต่างกัน เช่น การวิจัยของอาภรณ์ กาญจนกิจโสภณ (2531) ซึ่งสร้างตารางเทียบคะแนนแบบทดสอบวิชาคณิตศาสตร์ เรื่องสมการและอสมการ ระดับชั้น ม.1 ม.2 และ ม.3 เป็นต้น

2. การศึกษาความก้าวหน้าในการเรียน

ในการศึกษาความเปลี่ยนแปลง หรือความก้าวหน้าในการเรียนแต่ละวิชาสามารถทำได้โดยนำผลการเรียนรู้แต่ละระดับของผู้สอบคนเดิมมาเปรียบเทียบกันเพื่อดูความเปลี่ยนแปลงที่เกิดขึ้นในตัวผู้เรียน ในการศึกษาดังกล่าวแบบทดสอบที่ใช้วัดแต่ละครั้งจะวัดความรู้ในระดับที่ต่างกัน และผู้สอบก็มีระดับความสามารถต่างกันในการสอบแต่ละครั้ง ดังนั้นการนำคะแนนจากการสอบโดยแบบทดสอบที่มีระดับความยากต่างกัน และผู้สอบมีความสามารถต่างกันมาเปรียบเทียบกันจึงต้องมีการเทียบคะแนนจากการสอบ 2 ครั้งดังกล่าวให้อยู่ในสเกลเดียวกันโดยใช้วิธีการเทียบคะแนนแล้วนำผลการสอบของผู้สอบคนเดียวกันหรือผู้สอบที่มีความสามารถเดียวกันมาเปรียบเทียบกันเพื่อดูความก้าวหน้าในการเรียนจากความเปลี่ยนแปลงที่เกิดขึ้น

ในต่างประเทศมีการศึกษาความก้าวหน้าในการเรียนโดยใช้แบบทดสอบหลายระดับกับผู้สอบระดับชั้นต่าง ๆ และใช้วิธีการเทียบคะแนนเพื่อเปรียบเทียบคะแนนระหว่างระดับชั้นต่าง ๆ สำหรับในประเทศไทยการใช้ผลการเทียบคะแนนในลักษณะนี้จะได้เห็นได้จากการศึกษาของสุจินดา ผ่องอักษร (2532) ซึ่งศึกษาความก้าวหน้าทางผลสัมฤทธิ์ทางการเรียนของนักเรียนระดับชั้นประถมศึกษาปีที่ 6 ระหว่างปีการศึกษา 2529 - 2531 เพื่อศึกษาความเปลี่ยนแปลงของผลการศึกษาในระหว่างปีการศึกษาดังกล่าว เป็นต้น

ตอนที่ 2 การเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ

แนวความคิดของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ หรือบางครั้งเรียกว่าทฤษฎีคุณลักษณะแฝงเป็นทฤษฎีที่อธิบายความสัมพันธ์ระหว่างคุณลักษณะภายในหรือความสามารถที่มีอยู่ในตัวบุคคลกับพฤติกรรมการตอบสนองข้อสอบของบุคคลนั้นว่ามีโอกาสตอบข้อสอบถูกมากน้อยเพียงใด ทฤษฎีนี้มีพื้นฐานความเชื่อว่าการตอบข้อสอบของผู้สอบซึ่งเป็นสิ่งที่สังเกตได้โดยตรงว่าถูกหรือผิด จะถูกกำหนดโดยคุณลักษณะภายใน หรือความสามารถที่มีอยู่ในตัวบุคคลซึ่งเป็นสิ่งที่ไม่สามารถสังเกตได้โดยตรง ทฤษฎีนี้อธิบายความสัมพันธ์ดังกล่าวในรูปของฟังก์ชันทางคณิตศาสตร์ หรือแบบจำลองที่แสดงความสัมพันธ์ระหว่างระดับความสามารถ คุณลักษณะของข้อสอบ และโอกาสการตอบสนองข้อสอบได้ ถูกที่เรียกว่าฟังก์ชันการตอบสนองข้อสอบ ซึ่งมีหลายรูปแบบ เช่น ฟังก์ชันนอร์มอล ออจีไว (Normal Ogive Function) ฟังก์ชันโลจิสติก (Logistic Function) ฟังก์ชันการตอบสนองข้อสอบสามารถนำมาใช้ศึกษาความสัมพันธ์ระหว่างความน่าจะเป็นในการตอบสนองข้อสอบแต่ละข้อได้ถูกต้อง กับระดับความสามารถของผู้สอบที่วัดได้โดยแบบทดสอบฉบับนั้น เมื่อนำมาเขียนกราฟจะได้โค้งลักษณะข้อสอบ (Item Characteristic Curve ; ICC) โค้งลักษณะข้อสอบมีได้หลายลักษณะขึ้นอยู่กับแบบจำลองหรือโมเดลที่ใช้อธิบายความสัมพันธ์ดังกล่าว (ศิริชัย กาญจนวาสี. 2534 : 119 - 120) แบบจำลองต่าง ๆ สามารถสร้างขึ้นได้โดยการกำหนดข้อตกลงเบื้องต้นเกี่ยวกับข้อมูลที่แบบจำลองนั้นนำมาใช้ และยังมีข้อตกลงเบื้องต้นในแบบจำลองที่เป็นไปได้มากเพียงใดก็ยังเป็นประโยชน์มากขึ้นเท่านั้น

ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

แบบจำลองทางคณิตศาสตร์โดยทั่วไปจะประกอบด้วยข้อตกลงเบื้องต้นเกี่ยวกับข้อมูลที่แบบจำลองนั้นนำมาใช้ และความสัมพันธ์ระหว่างตัวแปรที่สังเกตได้ และตัวแปรที่สังเกตไม่ได้

ซึ่งความสัมพันธ์นี้เรียกว่าไอแก้มลักษณะข้อสอบข้อตกลงเบื้องต้นโดยทั่วไปของแบบจำลองการตอบสนองข้อสอบมี 3 ประการดังนี้

1. มิติของแบบจำลอง (Dimensionality of the Latent Space) ในทฤษฎีคุณลักษณะแฝงโดยทั่วไปมีข้อตกลงว่าคุณลักษณะต่าง ๆ หรือความสามารถของผู้สอบเป็นสิ่งที่อยู่ในผลของการทดสอบ คุณลักษณะนี้ถ้ากำหนดให้มี k ลักษณะหมายถึงมี k มิติด้วย การนำทฤษฎีคุณลักษณะแฝงหรือ ทฤษฎีการตอบสนองข้อสอบไปใช้นั้นมีข้อตกลงเบื้องต้นว่า ความสามารถเดียวหรือคุณลักษณะเดียวก็เพียงพอต่อการอธิบายความสามารถในการทำข้อสอบของผู้สอบแล้ว แบบจำลองการตอบสนองข้อสอบที่กำหนดข้อตกลงให้มีคุณลักษณะเพียงอย่างเดียวนี้ หมายถึงความเป็นมิติเดียว (Unidimensionality) นั่นเอง ข้อตกลงนี้เป็นลักษณะโดยทั่วไปในการสร้างแบบทดสอบซึ่งมักจะสร้างให้มีความเป็นมิติเดียว เพื่อให้ง่ายในการแปลความหมายของคะแนน แต่ข้อตกลงเบื้องต้นในเรื่องความเป็นมิติเดียวนี้อาจไม่สามารถเกิดขึ้นได้เสมอไป เนื่องจากมักจะมีคุณลักษณะอื่นเกิดขึ้นเสมอ เช่น คุณลักษณะด้านบุคลิกภาพและองค์ประกอบอื่น ๆ ที่มีอิทธิพลต่อการทำแบบทดสอบองค์ประกอบเหล่านี้ ได้แก่ แรงจูงใจในการทำแบบทดสอบ ความวิตกกังวลในการทำแบบทดสอบ ความเร็วในการทำแบบทดสอบ ความรู้เกี่ยวกับการใช้กระดาษคำตอบ และทักษะในการคิด เป็นต้น ถ้าตั้งข้อตกลงเบื้องต้นให้มีความเป็นมิติเดียว ข้อมูลที่จะนำมาใช้ต้องเป็นข้อมูลที่มีองค์ประกอบหรือส่วนประกอบที่มีลักษณะเด่นเพียงลักษณะเดียวนั้นและลักษณะเด่นนี้ แสดงถึงผลการวัดที่ได้จากแบบทดสอบด้วย

แบบจำลองบางรูปแบบกำหนดข้อตกลงเบื้องต้นของการวัดความสามารถมากกว่า 1 มิติ เรียกว่าข้อตกลงดังกล่าวว่ามีความเป็นหลายมิติ (Multidimensionality) ผู้ที่ศึกษาเกี่ยวกับความเป็นหลายมิติได้แก่ บอคและแอทกิน (Bock and Aitkin, 1981) มูเลค (Mulaik, 1972) เรคคาส (Reckase, 1985) และ ซามิจิมา (Samejima, 1974) ซึ่งการพัฒนาเทคนิคนี้ในปัจจุบันยังจำกัดอยู่

2. ความเป็นอิสระในการตอบข้อสอบ (Local Independence) ข้อตกลงเบื้องต้นข้อนี้ กล่าวว่า การตอบข้อสอบแต่ละข้อของผู้สอบเป็นอิสระต่อกัน ความหมายของข้อตกลงเบื้องต้นข้อนี้ตรงกับข้อตกลงเบื้องต้นในเรื่องความเป็นมิติเดียวของแบบทดสอบ ถ้าข้อตกลงเบื้องต้นข้อนี้เป็นจริง จะพบว่าผลการสอบของผู้สอบในแต่ละข้อจะไม่มีผลต่อการตอบข้อสอบข้ออื่น ๆ ตัวอย่างเช่น เนื้อหาของข้อสอบใด ๆ จะต้องไม่แนะนำคำตอบข้อสอบข้ออื่น ๆ สิ่งที่ข้อตกลงเบื้องต้นในข้อนี้เน้นก็คือจะต้องเป็นความสามารถของผู้สอบและคุณลักษณะของข้อสอบเท่านั้นที่มีผลต่อการทำข้อสอบ เมื่อข้อสอบมีความเป็นอิสระในการตอบ ความน่าจะเป็นของรูปแบบการตอบข้อสอบของผู้สอบเป็นผลจากความน่าจะเป็นในการตอบข้อสอบแต่ละข้อ

ถ้า U_i เป็นตัวแปรสุ่มของการตอบข้อสอบที่สุ่มได้จากกลุ่มผู้สอบที่ทำข้อสอบข้อที่ i ($i = 1, 2, 3, \dots, n$) โดยให้ $U_i = 1$ ถ้าผู้สอบตอบข้อสอบได้ถูก และ $U_i = 0$ ถ้าผู้สอบตอบข้อสอบผิด กำหนดสัญลักษณ์ P_i และ Q_i แทนความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบได้ถูก และผิด ตามลำดับ ความน่าจะเป็น P_i และ Q_i นี้ขึ้นอยู่กับ ความสามารถของกลุ่มผู้สอบซึ่งอาจใช้สัญลักษณ์ $P_i(\theta)$ และ $Q_i(\theta)$ ได้ ข้อตกลงเบื้องต้นในเรื่องความเป็นอิสระในการตอบข้อสอบสามารถเขียนในรูปของคณิตศาสตร์ได้ดังนี้

$$\begin{aligned} \text{Prob} [U_1 = u_1, U_2 = u_2, \dots, U_n = u_n | \theta] \\ = P_1(\theta)^{u_1} Q_1(\theta)^{1-u_1} P_2(\theta)^{u_2} Q_2(\theta)^{1-u_2} \dots P_n(\theta)^{u_n} Q_n(\theta)^{1-u_n} \\ = \prod_{i=1}^n P_i(\theta)^{u_i} Q_i(\theta)^{1-u_i} \end{aligned}$$

3. โค้งลักษณะข้อสอบ (Item Characteristic Curves) โค้งลักษณะข้อสอบหรือฟังก์ชันลักษณะข้อสอบเป็นฟังก์ชันทางคณิตศาสตร์ที่แสดงความสัมพันธ์ของความน่าจะเป็นของความสำเร็จในการตอบข้อสอบที่ได้จากการวัดด้วยแบบทดสอบ โดยทั่วไปเป็นฟังก์ชันที่ไม่ใช่เส้นตรงของการถดถอยของคะแนนรายข้อบนความสามารถ ความแตกต่างส่วนมากที่พบในแบบจำลองการตอบสนองข้อสอบ คือฟังก์ชันทางคณิตศาสตร์ของโค้งลักษณะข้อสอบขึ้นอยู่กับผู้ที่พัฒนาแบบทดสอบหรือขึ้นอยู่กับผู้เลือกฟังก์ชันทางคณิตศาสตร์มาใช้อธิบายโค้งลักษณะข้อสอบ เนื่องจากความน่าจะเป็นของผู้สอบแต่ละคนที่จะตอบข้อสอบถูกขึ้นอยู่กับโค้งลักษณะข้อสอบ ความน่าจะเป็นนี้เป็นการแจกแจงความสามารถของผู้สอบที่เป็นอิสระกัน ดังนั้นความน่าจะเป็นของการตอบข้อสอบแต่ละคนจึงไม่ขึ้นอยู่กับจำนวนของผู้สอบที่มีระดับความสามารถเดียวกัน คุณสมบัติของความไม่แปรผันของโค้งลักษณะข้อสอบนี้เป็นลักษณะเด่นของแบบจำลองการตอบสนองข้อสอบ โค้งลักษณะข้อสอบมีอยู่หลายรูปแบบ ขึ้นอยู่กับลักษณะเฉพาะของแบบจำลองของการตอบสนองข้อสอบ โดยทั่วไปอาจจะมีพารามิเตอร์ หนึ่ง สอง หรือ สาม ตัวก็ได้

แบบจำลองการตอบสนองข้อสอบ (Item Response Model)

แบบจำลองการตอบสนองข้อสอบมีหลายรูปแบบ ข้อแตกต่างที่สำคัญระหว่างแบบจำลองคือ ฟังก์ชันทางคณิตศาสตร์ที่นำมาใช้ แบบจำลองการตอบสนองข้อสอบจำแนกตามลักษณะข้อมูลผลการตอบข้อสอบได้ 3 ประเภทดังนี้ (Hambleton and Swaminathan, 1985 : 35-52)

1. แบบจำลองที่ใช้กับข้อมูลการตอบที่ให้คะแนนแบบสองค่า (Dichotomous)

แบบจำลองประเภทนี้มีหลายรูปแบบ ดังนี้

1.1 แบบจำลองนอร์มอล ออจีไว์ ที่มี 2 พารามิเตอร์ (Two Parameter Normal Ogive Model) ลอร์ดเสนอแบบจำลองนี้ใน ปี ค.ศ. 1952 แบบจำลองที่เสนอมี

พารามิเตอร์ 2 ตัว ซึ่งมีไค่งลักษณะข้อสอบอยู่ในรูปของนอร์มอล ออโจว ดังนี้

$$P_i(\theta) = \int_{-\infty}^{a_i(\theta-b_i)} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \quad i = 1, 2, \dots, n$$

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบที่มีความสามารถ จะตอบข้อสอบข้อที่ i ได้ถูก

b_i คือ ความยากของข้อสอบข้อที่ i

a_i คือ อำนาจจำแนกของข้อสอบข้อที่ i

Z คือ ค่าเบี่ยงเบนปกติจากการแจกแจง ซึ่งมีค่าเฉลี่ยเท่ากับ b_i และค่าเบี่ยงเบนมาตรฐานเท่ากับ $1/a_i$

ค่าความยากจะอยู่ระหว่าง -2.0 ถึง $+2.0$ ค่าความยากที่ใกล้ -2.0 แสดงว่าข้อสอบง่ายและถ้าความยากมีค่าใกล้ $+2.0$ แสดงว่าข้อสอบยาก

ส่วนค่าอำนาจจำแนกโดยทฤษฎีแล้วมีค่าระหว่าง $-\infty$ ถึง $+\infty$ แต่โดยปกติแล้ว ค่าอำนาจจำแนก จะอยู่ระหว่าง 0 ถึง $+2.0$

1.2 แบบจำลองโลจิสติกที่มี 2 พารามิเตอร์ (Two - Parameter Logistic Model) เบิร์นบอม (Birnbaum) เสนอแบบจำลองการตอบสนองข้อสอบที่มีไค่งลักษณะข้อสอบอยู่ในการแจกแจงแบบโลจิสติกที่มีพารามิเตอร์ 2 ตัว ดังนี้

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

$P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบที่มีความสามารถจะตอบข้อสอบข้อที่ i ได้ถูก

b_i คือ ความยากของข้อสอบข้อที่ i

a_i คือ อำนาจจำแนกของข้อสอบข้อที่ i

n คือ จำนวนข้อสอบในแบบทดสอบ

D คือ ค่าปรับสเกล มีค่าเท่ากับ 1.7

เบิร์นบอม ได้นำฟังก์ชันการแจกแจงสะสมแบบโลจิสติก ที่มี 2 พารามิเตอร์ไปใช้แทนฟังก์ชันนอร์มอล ออโจว ที่มี 2 พารามิเตอร์ ในการสร้างไค่งลักษณะข้อสอบ ปรากฏว่าการ

คำนวณทางคณิตศาสตร์ของแบบจำลองโลจิสติกทำได้ง่ายกว่าแบบจำลองนอร์มอล ออโจว ซึ่งได้มีการแสดงให้เห็นว่า เมื่อ D เท่ากับ 1.7 ค่าของ $P_i(\theta)$ ที่ได้จากแบบจำลองทั้งสองมีค่าแตกต่างกันน้อยกว่า .01 สำหรับทุก ๆ ค่าของ θ (Hambleton and Swaminathan, 1985:37 citeng Haley, 1952)

1.3 แบบจำลองโลจิสติกที่มี 3 พารามิเตอร์ (Three - Parameter Logistic Model) แบบจำลองนี้ได้จากแบบจำลองโลจิสติกที่มีพารามิเตอร์ 2 ตัว โดยเพิ่มพารามิเตอร์ตัวที่สามเข้าไป ดังนี้

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

โดยที่ $P_i(\theta)$, b_i , a_i และ D มีความหมายเช่นเดียวกับในแบบจำลองโลจิสติกที่มีพารามิเตอร์ 2 ตัว แต่มีความต่างที่ b_i บนสเกลความสามารถกล่าวคือ b_i มีค่าเท่ากับ $(1+c_i)/2$ แทนที่จะเป็น .50

พารามิเตอร์ c เป็นค่ากำกับต่ำสุด (Lower Asymptote) ของฟังก์ชัน ซึ่งเป็นความน่าจะเป็นที่ผู้สอบที่มีความสามารถต่ำมาก ๆ จะตอบข้อสอบข้อที่ i ได้ถูก การนำพารามิเตอร์ c นี้เข้ามาในแบบจำลองเพื่อใช้เป็นข้อมูลการตอบของผู้สอบที่มีความสามารถต่ำในการทำข้อสอบแบบเลือกตอบเนื่องจากการเดาก็เป็นปัจจัยหนึ่งที่มีผลต่อการตอบข้อสอบ ดังนั้นจึงมักเรียกพารามิเตอร์ c นี้ว่าเป็นพารามิเตอร์การเดา (Guessing Parameter)

1.4 แบบจำลองโลจิสติกที่มี 1 พารามิเตอร์ (One - Parameter Logistic Model or Rasch Model) ตั้งแต่ ค.ศ. 1960 เป็นต้นมา มีนักวิจัยหลายคนหันมาสนใจแนวคิดเกี่ยวกับแบบจำลองการตอบสนองข้อสอบของจอร์จ ราสช์ (George Rasch) ซึ่งเป็นนักคณิตศาสตร์ชาวเดนมาร์ก ผลงานของเขาระตุ้นความสนใจของนักพัฒนาแบบทดสอบต่าง ๆ และมีการศึกษาจุดเด่น และความเป็นไปได้ในการนำแบบจำลองนี้ไปใช้ มีการพัฒนาโปรแกรมคอมพิวเตอร์ที่ใช้ได้ง่ายและรวดเร็วมากขึ้นจึงทำให้แบบจำลองของราสช์นี้แพร่หลายมาก

แม้ว่าแบบจำลองของราสช์ จะได้รับการพัฒนาโดยอิสระ จากแบบจำลองการตอบสนองข้อสอบแบบอื่น ๆ แต่ก็ถือได้ว่าแบบจำลองนี้เป็นกรณีเฉพาะของแบบจำลองโลจิสติกที่มี 3 พารามิเตอร์ โดยมีข้อตกลงเพิ่มขึ้น 2 ประการ คือ 1) อำนาจจำแนก ของข้อสอบแต่ละข้อ (a_i) เท่ากัน และ 2) ค่าการเดามีค่าน้อยมาก ดังนั้นฟังก์ชันจะเป็น

$$P_i(\theta) = \frac{e^{D(\theta-b_i)}}{1 + e^{D(\theta-b_i)}} \quad i = 1, 2, \dots, n$$

เมื่อ a เป็นค่าที่ไม่มีกฏนิยาม ซึ่งถือว่าเป็นอำนาจจำแนกของข้อสอบทุกข้อ
แบบจำลองราล์ซ เป็นแบบจำลองที่มีคุณสมบัติบางประการที่ดึงดูดใจผู้ใช้ เนื่องจาก

1) แบบจำลองนี้เกี่ยวข้องกับพารามิเตอร์น้อยกว่าแบบจำลองอื่น ซึ่งทำให้ง่ายต่อการดำเนินงาน

2) ปัญหาในการประมาณค่าพารามิเตอร์มีน้อยกว่าแบบจำลองอื่น

1.5 แบบจำลองโลจิสติกที่มี 4 พารามิเตอร์ (Four-Parameter Logistic Model)

แมคโดนัลด์ (Mc Donald, 1967) บาร์ตันและลอร์ด (Barton and Lord, 1981) อธิบายแบบจำลองโลจิสติกที่มีพารามิเตอร์ 4 ตัวดังนี้

$$P_i(\theta) = c_i + (\gamma_i - c_i) \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}} \quad i = 1, 2, \dots, n$$

แบบจำลองนี้ต่างจากแบบจำลองโลจิสติกที่มี 3 พารามิเตอร์ ตรงที่มีค่า ซึ่งสมมติให้เป็นค่าที่ต่ำกว่า 1 ซึ่งหมายความว่า ไค่งลักษณะข้อสอบอาจมีค่ากำกับสูงสุด (Upper Asymptote) น้อยกว่า 1 การเสนอแบบจำลองนี้ก็เพื่อจะแก้ปัญหาผู้สอบที่มีความสามารถสูงอาจเกิดความสับสนหรืออาจรู้มากกว่าผู้ออกข้อสอบจึงทำให้เลือกคำตอบที่ไม่ใช่ข้อที่เฉลยไว้ ปัจจุบันแบบจำลองนี้ได้ได้รับความสนใจเฉพาะเชิงทฤษฎีเท่านั้น ยังไม่มีผู้นำมาใช้ปฏิบัติ

1.6 แบบจำลองนอร์มอล ออโวก์ ที่มี 1 3 และ 4 พารามิเตอร์ (One-, Three-, and Four- Parameter Nomal Ogive Models) รูปแบบทางคณิตศาสตร์ของแบบจำลองนอร์มอล ออโวก์ที่มีพารามิเตอร์ต่าง ๆ เป็นดังนี้

1.6.1 แบบจำลองนอร์มอล ออโวก์ที่มี 1 พารามิเตอร์

$$P_i(\theta) = \int_{-\infty}^{\theta-b_i} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \quad i = 1, 2, \dots, n$$

1.6.2 แบบจำลองนอร์มอล ออโจว์ที่มี 3 พารามิเตอร์

$$P_i(\theta) = c_i + (1 - c_i) \int_{-\infty}^{a_i(\theta - b_i)} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

1.6.3 แบบจำลองนอร์มอล ออโจว์ ที่มี 4 พารามิเตอร์

$$P_i(\theta) = c_i + (\gamma_i - c_i) \int_{-\infty}^{a_i(\theta - b_i)} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

แบบจำลองนอร์มอล ออโจว์ มีความยุ่งยากในการคำนวณมากกว่าแบบจำลองโลจิสติกและยังมีจำนวนพารามิเตอร์เพิ่มมากขึ้น การคำนวณยิ่งยุ่งยากมากขึ้น ดังนั้นในปัจจุบันแบบจำลองนอร์มอล ออโจว์ จึงเป็นเพียงความสนใจที่เกิดขึ้นในอดีตเท่านั้น

2. แบบจำลองที่ใช้กับข้อมูลการตอบที่ให้คะแนนแบบหลายค่า (Multicategory Scoring) แบบจำลองประเภทนี้ใช้กับแบบทดสอบที่ให้คะแนนเป็นแบบลำดับขั้นเช่น แบบวัดเจตคติ แบบวัดบุคลิกภาพ เป็นต้น แบบจำลองประเภทนี้ได้แก่แบบจำลองนอมินัล เรสพอน (Nominal Response Model) แบบจำลองเกรด เรสพอน (Grade Response Model) แบบจำลอง ไบโนเมียล ไทรอัล (Binomial Trials Model) แบบจำลอง เรตติ้ง สเกล (Rating Scale Model) และ แบบจำลอง พาร์เชียล เครดิต (Partial Credit Model) ..

2.1 แบบจำลองนอมินัล เรสพอน บอค (Bock) เป็นคนแรกที่พัฒนาแบบจำลองการตอบสนองข้อสอบที่ให้คะแนนแบบหลายค่า แบบจำลองนี้มีจุดมุ่งหมายเพื่อประมาณค่าความสามารถของผู้สอบให้มีความแม่นยำมากที่สุด โดยใช้สารสนเทศจากการตอบสนองข้อสอบแต่ละข้อภายใต้เงื่อนไขที่ว่าผลรวมของความน่าจะเป็นในการตอบสนองที่เป็นไปได้ทั้งหมดที่ระดับความสามารถที่กำหนดต้องเท่ากับ 1 แม้ว่าจะมีการเว้นไม่ตอบสำหรับการตอบสนองที่ถูกต้องนั้น ไค้งของความสามารถจะเป็นแบบเพิ่มขึ้นโดยตลอด ส่วนรูปร่างของไค้งความสามารถของการตอบสนองที่ไม่ถูกต้องอาจมีลักษณะเพิ่มขึ้นโดยตลอดหรือลดลงโดยตลอดหรือเป็นไค้งบอคเดียวที่มีค่าความโค้งต่างกันไป ทั้งนี้ขึ้นอยู่กับผู้ตอบที่มีระดับความสามารถต่าง ๆ จะเลือกตอบอย่างไร

2.2 แบบจำลองเกรด เรสพอน ซามิเจมา (Samejima) เป็นผู้เสนอแบบจำลองนี้เพื่อใช้กับสเกลที่มีลักษณะเป็นลำดับขั้น (Ordered Categories) เช่น ลิเคิร์ตสเกล และสเกลแบบซีแมนติก (Semantic Differential rating Scale) บางครั้งเรียกแบบจำลองนี้ว่า แบบจำลองความแตกต่าง (Difference Model) เพราะความน่าจะเป็นในการ

เลือกตัวเลือกที่ตอบสนองมีความแตกต่างกันตามความน่าจะเป็นของโค้งลักษณะข้อสอบของแต่ละตัวเลือกที่ตอบ

2.3 แบบจำลองไบโนเมียล ไทรอัล แอนดริช (Andrich) ขยายแบบจำลองโลจิสติกที่มี 1 พารามิเตอร์ เพื่อนำมาประยุกต์ใช้กับข้อมูลประเภทการนับจำนวนครั้งของความสำเร็จและจำนวนครั้งที่ไม่สำเร็จในการตอบคำถามแต่ละข้อ

2.4 แบบจำลองเรทติงสเกล แอนดริช เป็นผู้พัฒนาขึ้นจากแบบจำลองโลจิสติกที่มี 1 พารามิเตอร์ เพื่อนำมาประยุกต์ใช้กับข้อคำถามที่มีรูปแบบการตอบมากกว่า 2 ระดับขึ้นไป (Polychotomous) โดยนำมาใช้กับการวัดเจตคติ ซึ่งมีการตรวจให้คะแนนตามระดับความคิดเห็นซึ่งแบ่งเป็นหลายระดับในลักษณะลำดับขั้น

2.5 แบบจำลองพาเซียนลเกรดิต มาสเตอร์ (Masters) ได้พัฒนาแบบจำลองโลจิสติกที่มี 1 พารามิเตอร์เพื่อนำมาใช้กับข้อคำถามที่มีรูปแบบการตรวจ ให้คะแนนเป็นแบบลำดับขั้นเช่นเดียวกับแบบจำลองเรทติงสเกล แต่ให้ผลการวิเคราะห์ที่ละเอียดมากกว่า

3. แบบจำลองที่ใช้กับข้อมูลการตอบที่ให้คะแนนแบบต่อเนื่อง (continuous) ได้แก่ แบบจำลองการตอบสนองแบบต่อเนื่อง (Continuous Response Model) ซามิเจมาเป็นผู้เสนอแบบจำลองนี้ในปี 1973 เพื่อใช้กับสเกลที่เป็นแบบต่อเนื่อง แบบจำลองนี้ใช้ประโยชน์สำหรับการวัดทางจิตวิทยา และการวัดเจตคติ

กระบวนการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ

ในการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ เมื่อเลือกแบบจำลองการตอบสนองข้อสอบที่เหมาะสมได้แล้ว และไม่ว่าการเทียบคะแนนจะเป็นการเทียบคะแนนแนวนอนหรือการเทียบคะแนนตามแนวตั้ง กระบวนการในการเทียบคะแนนประกอบด้วย 3 ขั้นตอน คือ

1. การเลือกแบบแผนการเก็บรวบรวมข้อมูล
2. การจัดค่าพารามิเตอร์ของแบบทดสอบต่างฉบับให้อยู่ในสเกลเดียวกัน
3. การเทียบค่าประมาณความสามารถของผู้สอบ

รายละเอียดของกระบวนการเทียบคะแนนแต่ละขั้นตอนมีดังนี้

1. การเลือกแบบแผนการเก็บรวบรวมข้อมูล แบบแผนการเก็บรวบรวมข้อมูลจากแบบทดสอบต่างฉบับเพื่อนำมาเทียบคะแนนมีหลายรูปแบบ ดังได้กล่าวไว้ในตอนที่ 1 การเลือกรูปแบบใดพิจารณาได้จากธรรมชาติของแบบทดสอบและกลุ่มผู้ตอบ ซึ่งอาจเลือกใช้รูปแบบใดรูปแบบหนึ่งตามความเหมาะสม

2. การจัดทำพารามิเตอร์ของแบบทดสอบต่างฉบับให้อยู่ในสเกลเดียวกัน เนื่องจากการประมาณค่าพารามิเตอร์ของข้อสอบในแบบทดสอบต่างฉบับ อาจกระทำพร้อมกันหรือกระทำแยกกัน ในกรณีที่การประมาณค่าพารามิเตอร์ของข้อสอบต่างฉบับทำพร้อมกัน ค่าประมาณที่ได้จะอยู่ในสเกลเดียวกันจึงสามารถนำมาเปรียบเทียบกันได้เลย แต่ในบางกรณีที่จำเป็นต้องประมาณค่าพารามิเตอร์ของข้อสอบแยกกัน เช่นกรณีที่ผู้สอบแบบทดสอบแต่ละฉบับมีระดับความสามารถต่างกัน (ดังเช่นในการวิจัยครั้งนี้) การประมาณค่าพารามิเตอร์ของข้อสอบต้องกระทำแยกกัน ค่าประมาณที่ได้จึงไม่อยู่ในสเกลเดียวกัน ดังนั้นต้องมีการปรับค่าประมาณพารามิเตอร์ของข้อสอบจากแบบทดสอบต่างฉบับให้อยู่ในสเกลเดียวกันก่อนจึงจะนำค่าที่ได้ไปใช้ในการเปรียบเทียบกัน การทำให้ค่าพารามิเตอร์ของแบบทดสอบ 2 ฉบับใด ๆ อยู่ในสเกลเดียวกันต้องอาศัยข้อสอบร่วมในแบบทดสอบทั้ง 2 ฉบับ หรือต้องมีผู้สอบร่วมส่วนหนึ่ง ซึ่งทำแบบทดสอบทั้ง 2 ฉบับ จึงจะสามารถนำข้อมูลที่ร่วมกัน 2 ฉบับมาใช้ปรับค่าพารามิเตอร์ของข้อสอบจากแบบทดสอบต่างฉบับนั้นให้อยู่ในสเกลเดียวกันได้ วิธีการโดยทั่วไปที่ใช้ในการจัดให้ค่าพารามิเตอร์ของข้อสอบจากแบบทดสอบต่างฉบับมาอยู่ในสเกลเดียวกันมีดังนี้ (Petersen and others. 1989 : 256)

2.1 วิธีประมาณค่าพร้อมกัน (Concurrent Calibration) ตามวิธีการนี้ พารามิเตอร์ของแบบทดสอบ 2 ฉบับ จะถูกประมาณค่าพร้อมกัน โดยใช้ข้อมูลจากผู้สอบร่วมหรือข้อสอบร่วม (แล้วแต่กรณี) จะทำให้ค่าประมาณพารามิเตอร์ของข้อสอบและค่าประมาณความสามารถอยู่ในสเกลเดียวกัน

2.2 วิธีกำหนดค่าความยากให้คงที่ (The Fixed bs Method) ตามวิธีการนี้จะประมาณค่าพารามิเตอร์ของแบบทดสอบทีละฉบับ หลังจากประมาณค่าพารามิเตอร์แบบทดสอบฉบับแรกแล้ว จะใช้ค่าความยากของข้อสอบร่วมไปประมาณค่าพารามิเตอร์ของแบบทดสอบฉบับที่สอง โดยไม่ต้องประมาณค่าพารามิเตอร์จากแบบทดสอบฉบับที่สองซ้ำ ดังนั้น พารามิเตอร์ของความสามารถและพารามิเตอร์ของข้อสอบจากแบบทดสอบฉบับที่สองจึงอยู่บนสเกลเดียวกับแบบทดสอบฉบับแรก

2.3 วิธีเทียบค่าความยาก (The Equated bs Method) ตามวิธีการนี้ พารามิเตอร์ของแบบทดสอบแต่ละฉบับจะถูกประมาณค่าแยกกัน ดังนั้นในกลุ่มของแบบทดสอบร่วมค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของความยากของแบบทดสอบทั้ง 2 ฉบับจะถูกกำหนดให้เท่ากันโดยใช้การแปลงเชิงเส้นตรงเพื่อแปลงค่าพารามิเตอร์ของความสามารถและค่าพารามิเตอร์ความยาก (b) และอำนาจจำแนก (a) ของแบบทดสอบฉบับที่ 2 ไปสู่สเกลของแบบทดสอบฉบับแรก

2.4 วิธีใช้โค้งลักษณะข้อสอบ (The Characteristic Curve Method) ตามวิธีการนี้พารามิเตอร์ของข้อสอบจากแบบทดสอบ 2 ฉบับ จะประมาณค่าแยกกัน พารามิเตอร์

ที่ได้จากการประมาณค่า 2 ครั้งนี้ จะแตกต่างกันเฉพาะจุดเริ่มต้น (Origin) และหน่วยของการวัดเท่านั้น ดังนั้นจึงสามารถแปลงค่าประมาณพารามิเตอร์ที่ได้ให้อยู่ในสเกลเดียวกันโดยใช้การแปลงเชิงเส้นตรง

3. การเทียบค่าประมาณความสามารถของผู้สอบต้องพิจารณาว่าคะแนนจากแบบทดสอบจะรายงานในรูปคะแนนความสามารถหรือคะแนนจริงหรือคะแนนสังเกต การเทียบคะแนนมีวิธีการแตกต่างกัน ดังนี้

3.1 ในกรณีที่การรายงานผลคะแนนจากแบบทดสอบรายงานในรูปคะแนนความสามารถการเทียบคะแนนสามารถทำได้เลยโดยใช้ค่าประมาณความสามารถที่ได้จากแบบทดสอบต่างฉบับ

3.2 ในกรณีที่การรายงานผลทำในรูปคะแนนจริง จะต้องมีการประมาณค่าคะแนนจริงจากแบบทดสอบแต่ละฉบับโดยใช้ค่าประมาณพารามิเตอร์ที่ได้ในขั้นตอนที่ 2 คะแนนจริงจะถูกประมาณค่าหลาย ๆ ค่า เพื่อนำไปใช้ สร้างกราฟหรือตารางเพื่อเทียบคะแนนจริงที่ตรงกันจากแบบทดสอบแต่ละฉบับ

3.3 ในกรณีที่การรายงานคะแนนทำในรูปคะแนนสังเกตจะต้องทำการแจกแจงคะแนนสังเกตอย่างมีเงื่อนไขตามทฤษฎี สำหรับความสามารถที่ตรงกันของกลุ่มผู้สอบเพื่อเทียบคะแนนที่สร้างขึ้นนี้ด้วยวิธีเทียบคะแนนอิกวิเปอร์เซ็นไทล์ แล้วทำการเทียบคะแนนสังเกตที่แท้จริงจากตารางหรือกราฟของการเปรียบเทียบ

ตอนที่ 3 งานวิจัยที่เกี่ยวข้องกับการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ

การศึกษาเกี่ยวกับการเทียบคะแนนเริ่มมีการกล่าวถึงตั้งแต่ราว ค.ศ. 1920 แต่งานที่พิมพ์เผยแพร่มีน้อยมาก ต่อมาความสนใจเรื่องการเทียบคะแนนมีมากขึ้นนับตั้งแต่แองกอฟ เขียนเรื่อง "Scales, Norms, and Equivalent scores" ในหนังสือ Educational Measurement (2nd.) (Brennan. 1987:221) ในช่วงปี ค.ศ. 1978 - 1980 สถาบันบริการการทดสอบทางการศึกษาของสหรัฐอเมริกา (Educational Testing Service: ETS) ได้จัดทำโครงการวิจัยเพื่อทบทวน ตรวจสอบวิธีการเทียบคะแนนที่ใช้อยู่และมีการจัดสัมมนา โดยเน้นการพัฒนาเทคนิควิธีการเทียบคะแนน การผลิตเอกสารและบทความเกี่ยวกับการเทียบคะแนน (Holland and Rubin. 1982 : XIII) ในช่วงต้นทศวรรษ 1980 มีการกล่าวอ้างอิงถึงการเทียบคะแนนมากขึ้น ซึ่งจะเห็นได้จาก ในปี ค.ศ. 1985 สมาคมจิตวิทยาของสหรัฐอเมริกา (American Psychological Association) ได้กล่าวถึงการเทียบคะแนนไว้ในข้อกำหนดสำหรับการทดสอบทางการศึกษาและจิตวิทยา (American Psychological

Association. 1985 : 31-34) ซึ่งในข้อกำหนดฉบับที่ผ่านมาไม่มีการกล่าวถึงการเทียบคะแนนไว้

ตามแนวความคิดของการเทียบคะแนนนั้นเป็นเรื่องที่ไม่ซับซ้อน แต่โดยที่การเทียบคะแนนเป็นกระบวนการในเชิงประจักษ์ ซึ่งต้องเกี่ยวข้องกับการพิจารณาทางเลือกเกี่ยวกับแบบแผนในการเก็บรวบรวมข้อมูลและวิธีการที่ใช้ในการเทียบคะแนน ดังนั้นประเด็นของการศึกษาวิจัยจึงเกี่ยวข้องกับประเด็นที่กล่าวข้างต้น ในระยะแรกมีการศึกษาวิจัยเกี่ยวกับการเทียบคะแนนตามวิธีการดั้งเดิม ต่อมาประมาณปลายทศวรรษ 1960 มีการนำทฤษฎีการตอบสนองข้อสอบมาประยุกต์ใช้กับการเทียบคะแนน จากบทความของลอว์ค (1977) และไรท์ (1977) ที่ชี้ให้เห็นข้อดีในเชิงทฤษฎีของการนำทฤษฎีการตอบสนองข้อสอบมาใช้ในการเทียบคะแนน โดยเฉพาะการเทียบคะแนนตามแนวตั้ง ทำให้นักวิจัยหันมาศึกษาเปรียบเทียบผลการเทียบคะแนนตามวิธีการดั้งเดิมกับการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบในสถานการณ์ของการสอบจริง (Cook and Eignor. 1989:162) โดยศึกษาผลของการเทียบคะแนนในตัวแปรต่าง ๆ เช่น ลักษณะของกลุ่มตัวอย่างลักษณะของแบบทดสอบที่ใช้เทียบคะแนน ลักษณะของข้อสอบรวม วิธีการทางสถิติที่ใช้ในการประมาณค่าคะแนนแปลง เป็นต้น สำหรับงานวิจัยที่เกี่ยวข้องกับการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบมีประเด็นเรื่องการศึกษาสรุป ได้ดังตาราง 1 (Skaggs and Lissitz. 1986: 498 - 501)

ตาราง 1 ประเด็นการวิจัยเกี่ยวกับการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ

ผู้วิจัย ปีที่ศึกษา	แบบทดสอบ	กลุ่ม ตัวอย่าง	แบบจำลอง การเทียบ คะแนน	ตัวแปรอิสระ	วิธีการ ประเมินผล
ไรท์ 1968	LSAT	นักเรียน กฎหมาย 980 คน	ราล์ช	1.แบบทดสอบ ย่อยฉบับง่าย และฉบับยาก 2.กลุ่มตัวอย่าง ความสามารถ สูงและต่ำ	1.ความแตกต่าง มาตรฐาน 2.การเปรียบเทียบ เชิงโค้งลักษณะ ข้อสอบของ แบบทดสอบ 3.ความสัมพันธ์ ของความยาก
แอนเดอร์สัน และ อีเวอเรท 1968	แบบทดสอบ ตรวจสอบ เชาวน์ปัญญา	ทหาร 610 และ 875 คน	ราล์ช	1.การเปรียบเทียบ ค่าประมาณ ความยากของ ข้อสอบจาก กลุ่มตัวอย่าง 2 กลุ่ม	1.ความสัมพันธ์ ของค่าความ ยาก
ไวท์ลี และ เดวิส 1974	แบบทดสอบ อุปมาอุปมัย	นักเรียน ระดับ วิทยาลัย และระดับ มัธยม 950 คน	ราล์ช	1.แบบทดสอบ ย่อยฉบับง่าย และฉบับยาก 2.แบบทดสอบ ย่อย ข้อคู่และ ข้อคี่	1.ความแตกต่าง มาตรฐานของ ค่าประมาณ ความสามารถ

ตาราง 1 (ต่อ)

ผู้วิจัย ปีที่ศึกษา	แบบทดสอบ	กลุ่ม ตัวอย่าง	แบบจำลอง การเทียบ คะแนน	ตัวแปรอิสระ	วิธีการ ประเมินผล
ทินสเลย์ และ เดวิส 1975	แบบทดสอบ อุปมาอุปมัย	นักเรียน ระดับมัธยม ข้าราชการ ลูกจ้าง 120-630คน	ราล์ช	1.การเปรียบเทียบค่าประมาณความยากของข้อสอบซึ่งได้จากกลุ่มตัวอย่างที่ต่างกัน	สหสัมพันธ์ของค่าความยาก
เรนท และ บาชอร์ 1977	แบบทดสอบ การอ่าน CAT,CTBS ITBS,MAT STEP SRA,SAT	นักเรียน ระดับ 4-6 1300 - 2000 คน	ราล์ช	การพัฒนาสเกลอ้างอิงสำหรับการอ่านในระดับชาติและการประเมินผลการเทียบคะแนนด้วยราล์ชโดยการอ่าน	ค่าเบี่ยงเบนมาตรฐานของความยากและค่าเฉลี่ยความสามารถจากกลุ่มตัวอย่างที่ต่างกันและใช้แบบทดสอบที่ต่างกัน
สไลด์ และลิน 1977	แบบทดสอบ วัดผลสัมฤทธิ์ ทาง คณิตศาสตร์ CEED	นักศึกษา ระดับ วิทยาลัย ปีที่ 1 390-510 คน	ราล์ช	1.การเทียบ-คะแนนตามแนวทางในกลุ่มตัวอย่างย่อยที่มีความสามารถต่างกัน	ความแตกต่างมาตรฐานของค่าประมาณความสามารถ

ตาราง 1 (ต่อ)

ผู้วิจัย ปีที่ศึกษา	แบบทดสอบ	กลุ่ม ตัวอย่าง	แบบจำลอง การเทียบ คะแนน	ตัวแปรอิสระ	วิธีการ ประเมินผล
สไลด์ และลิน 1978	แบบทดสอบ ความเข้าใจ ในการอ่าน และคำศัพท์ SRA	นักเรียน ระดับ 5 510-570 คน	ราล์ช	1.การเทียบ คะแนนตาม แนวตั้งในกลุ่ม ตัวอย่างย่อยที่ มีความสามารถ ต่างกัน	ความแตกต่าง มาตรฐาน ของค่าประมาณ ความสามารถ
มารีโค ปีเตอร์สัน และ สตริต 1979	SAI (ภาษา)	นักเรียน ระดับมัธยม 1580 คน	1.ราล์ช 2.สาม พารามิเตอร์ 3.วิธีอีควิ- เปอร์เซ็นไทล์ 4.วิธีเชิง เส้นตรง	1.การเทียบ คะแนนตาม แนวนอนและ แนวตั้ง 2.กลุ่มตัวอย่าง แบบสุ่ม และ กลุ่มคล้ายคลึงกัน 3.แบบสอบร่วม ภายในและ ภายนอก 4.ความยากของ แบบทดสอบร่วม 5.ฟังก์ชันการ เทียบคะแนน แบบ IRT และ อีควิ- เปอร์เซ็นไทล์	ค่าเฉลี่ยกำลัง สองของความ คลาดเคลื่อน และดัชนี ความลำเอียง

ตาราง 1 (ต่อ)

ผู้วิจัย ปีที่ศึกษา	แบบทดสอบ	กลุ่ม ตัวอย่าง	แบบจำลอง การเทียบ คะแนน	ตัวแปรอิสระ	วิธีการ ประเมินผล
ลอยด์ และ ชูเวอร์ 1980	แบบทดสอบ ความเข้าใจ ทาง คณิตศาสตร์ ITBS	นักเรียน ระดับ 6-8 90, 245- 300 คน	ราล์ซ	6. เปรียบเทียบ แบบจำลองของ การเทียบคะแนน การเทียบคะแนน ตามแนวตั้งใน กลุ่มตัวอย่าง ต่างระดับชั้น	การเปรียบเทียบ จากกราฟของ ฟังก์ชันการ เทียบคะแนน
ดิฟกี 1981	MAT: การอ่าน	นักเรียน ระดับ 6-8 5500 คน	ราล์ซ	การเทียบคะแนน ตามแนวดิ่งของ แบบทดสอบย่อย ฉบับต่าง ๆ	1. ความแตกต่าง มาตรฐาน 2. ความลำเอียง ของคะแนนรวม
โคเลน 1981	ITED	นักเรียน ระดับ 9 - 12 1580 - 1925 คน	1. ราล์ซ 2. สอง พารามิเตอร์ 3. สาม พารามิเตอร์ 4. อีควิ- เปอร์เซ็นไทล์ 5. เชิงเส้นตรง	1. การเทียบ- คะแนนตามแนว นอนและแนวตั้ง 2. การเปรียบเทียบ- เทียบแบบจำลอง การเทียบคะแนน 3. ข้อสอบ ประเภทคำศัพท์ และประเภท ตัวเลข	ค่าเฉลี่ยกำลัง สองของความ แตกต่างจาก กลุ่มสอบทานผล

ตาราง 1 (ต่อ)

ผู้วิจัย ปีการศึกษา	แบบทดสอบ	กลุ่ม ตัวอย่าง	แบบจำลอง การเทียบ คะแนน	ตัวแปรอิสระ	วิธีการ ประเมินผล
ปีเตอร์สัน คุก และ สต็อกคิง 1981	SAT ภาษาและ คณิตศาสตร์	นักเรียน ระดับมัธยม 2670 คน	1.สาม พารามิเตอร์ 2.เชิงเส้นตรง 3.อีควิ- เปอร์เซ็นไทล์	1.การเปรียบเทียบ- เทียบแบบจำลอง การเทียบคะแนน 2.การเปรียบเทียบ เทียบการแปลง ค่าเมตริกซ์ของ แบบจำลองสาม พารามิเตอร์ 3.ข้อสอบภาษา กับข้อสอบ คณิตศาสตร์	1.การเทียบ คะแนนกับแบบ ทดสอบฉบับเดิม โดยเทียบผ่าน แบบทดสอบฉบับ อื่น 2.ค่าเฉลี่ยกำลัง สองของความ คลาดเคลื่อน 3.ดัชนีความ ลำเอียง
คุก คันนา และ ไอเนอร์ 1981	SAT/PSAT ภาษาและ คณิตศาสตร์	นักเรียน ระดับมัธยม 2000 คน	1.สาม พารามิเตอร์ 2.เชิงเส้นตรง 3.อีควิเปอร์- เซ็นไทล์	1.การเปรียบเทียบ- เทียบแบบจำลอง การเทียบคะแนน 2.ข้อสอบภาษา และ คณิตศาสตร์	1.ค่าเฉลี่ยกำลัง สองของความ คลาดเคลื่อน 2.ดัชนีความ - ลำเอียงโดยใช้ แบบจำลองสาม- พารามิเตอร์เป็น เกณฑ์
โคเลน และ ไวทนี 1981	แบบทดสอบ GED	ผู้ใหญ่ 200 คน	1.สามพารา มิเตอร์ 2.ราส์ซ 3.เชิงเส้นตรง	การเปรียบเทียบ แบบจำลองของ การเทียบคะแนน กับแบบทดสอบ ฉบับอื่น	ค่าเฉลี่ยกำลังสอง ของความแตกต่าง จากกลุ่มสอบทาน ผล

ตาราง 1 (ต่อ)

ผู้วิจัย ปีการศึกษา	แบบทดสอบ	กลุ่ม ตัวอย่าง	แบบจำลอง การเทียบ คะแนน	ตัวแปรอิสระ	วิธีการ ประเมินผล
เพเทียน 1981	แบบทดสอบ ITED	นักเรียน ระดับ 9-12 650-944 คน	4. อีคิว- เปอร์เซ็นต์ไทล์ 1. ราส์ช 2. สอง พารามิเตอร์ 3. สาม พารามิเตอร์ 4. อีคิว- เปอร์เซ็นต์ไทล์	การเปรียบเทียบ แบบจำลองของ การเทียบคะแนน ตามแนวตั้ง	1. ความสัมพันธ์ ของคะแนนสอบ บนสเกลการ เทียบคะแนนกับ คะแนนเดิมโดย ใช้การเทียบ คะแนนอีคิว- เปอร์เซ็นต์ไทล์ กับกลุ่มตัวอย่าง รวมเป็นเกณฑ์
คุกและ ไอเนอร์ 1983	แบบทดสอบ วัดผลสัมฤทธิ์ CB	นักเรียน มัธยม 2000- 4000 คน	1. สาม พารามิเตอร์ 2. อีคิว- เปอร์เซ็นต์ไทล์ 3. เชิงเส้น ตรง	1. วิธีการสร้าง เมตริกซ์ร่วม 2 วิธี 2. การเปรียบเทียบ แบบจำลอง ของการเทียบ คะแนน	1. การเทียบ คะแนนกับแบบ ทดสอบฉบับเดิม 2. ค่าเฉลี่ยกำลัง สองของความ คลาดเคลื่อน 3. ดัชนีความ ลำเอียง 4. กราฟ
โจแฮนสัน 1988	-	จำลองข้อมูล	สาม พารามิเตอร์	1. ความยาวของ แบบทดสอบรวม 2. พิสัยความยาก ของแบบทดสอบ	

ตาราง 1 (ต่อ)

ผู้วิจัย ปีการศึกษา	แบบทดสอบ	กลุ่ม ตัวอย่าง	แบบจำลอง การเทียบ คะแนน	ตัวแปรอิสระ	วิธีการ ประเมินผล
เพเชียน 1990	ITED	นักเรียน ระดับ 9,10,11	อีควิ- เปอร์เซ็นต์ไทล์ มาตรฐาน หนึ่ง ,สอง และสาม พารามิเตอร์	3.วิธีการ ประมาณค่า พารามิเตอร์ ในการแปลง คะแนน 1.การเปรียบเทียบ- เทียบแบบจำลองของการ เทียบคะแนน ตามแนวตั้ง 2.ระดับความ ยากของ แบบทดสอบ	
แฮริส 1991	EACT PACT	นักเรียน ระดับ 11,12 2050- 2176	1.อีควิ- เปอร์เซ็นต์ไทล์ 2. สาม พารามิเตอร์	1.แบบแผนการ เก็บข้อมูล 2.การเปรียบเทียบ- แบบจำลอง ของการเทียบ- คะแนน	ค่าเฉลี่ยกำลัง สองของความ แตกต่าง
อเยอเว 1992	-	จำลอง ข้อมูล	1.อีควิ- เปอร์เซ็นต์ไทล์ 2. สาม พารามิเตอร์	1.ขนาดตัวอย่าง 2.ความยาว ของแบบทดสอบ 3.ความยาวของ แบบทดสอบรวม	1.ค่าเฉลี่ยกำลัง สองของความ คลาดเคลื่อนซึ่ง ถ่วงน้ำหนักและ ไม่ถ่วงน้ำหนัก

งานวิจัยเกี่ยวกับการนำทฤษฎีการตอบสนองข้อสอบมาใช้ในการเทียบคะแนนในระยะแรก เป็นการศึกษาคุณสมบัติในเรื่องความไม่แปรผันของประชากร โดยใช้แบบจำลองของราล์ซ การศึกษาทำได้ 2 ลักษณะ คือ 1) การเปรียบเทียบค่าประมาณความยากของแบบทดสอบ ฉบับเดียวกันจากกลุ่มตัวอย่างที่ต่างกัน 2 กลุ่ม ซึ่งเป็นการศึกษาความเป็นอิสระของค่าประมาณ พารามิเตอร์ของข้อสอบจากกลุ่มตัวอย่างผู้สอบ (Person - Free Item Calibration) และ 2) การเปรียบเทียบค่าประมาณความสามารถของกลุ่มตัวอย่างกลุ่มเดียวกันจากแบบทดสอบ 2 ฉบับ ซึ่งเป็นการศึกษาความเป็นอิสระของค่าประมาณพารามิเตอร์ของความสามารถจากกลุ่ม ข้อสอบ (Item - Free Person Calibration) จากผลการศึกษาของ ไรท์ (Write. 1968) แอนเดอร์สัน เคอร์เนย์และอีเวอเรท (Anderson, Kearney and Everett. 1968) ทินสเลย์และเดวิส (Tinsley and Davis. 1975) ได้ข้อสรุปว่า ผลการศึกษาโดยใช้ แบบจำลองของราล์ซมีคุณสมบัติในเรื่องความไม่แปรผันของประชากรทั้ง 2 ลักษณะ

เมื่อพิจารณางานวิจัยที่ใช้แบบจำลองของราล์ซในการเทียบคะแนนจากงานวิจัยของเรนท์ และบาชอร์ (Renty and Bashow. 1977) ซึ่งพัฒนาสเกลอ้างอิงระดับชาติสำหรับการอ่าน สเกลนี้เป็นการเปรียบเทียบคะแนนการจัดกลุ่มของแบบทดสอบ 28 ฉบับ/ระดับ โดยใช้ข้อสอบ จากคลังข้อสอบ จำนวน 2644 ข้อ ในการศึกษาครั้งนี้มีการประเมินความเพียงพอของการเทียบ คะแนน โดยพิจารณาจากความแปรปรวนของค่าประมาณความสามารถจากการสอบหลายครั้ง โดยใช้แบบทดสอบฉบับเดิม มีการคำนวณค่าดัชนีจากค่าเฉลี่ยของค่าเบี่ยงเบนมาตรฐานของคะแนน ความสามารถแต่ละระดับของกลุ่มตัวอย่าง พิสัยของดัชนีมีค่าโลจิทระหว่าง .008 ถึง .041 เมื่อเปรียบเทียบสเกลของความสามารถกับความคลาดเคลื่อนมาตรฐานของค่าประมาณความ สามารถพบว่าค่าที่ได้มีค่าน้อย ผู้วิจัยจึงสรุปว่าการเทียบคะแนนโดยใช้แบบจำลองของราล์ซมี คุณสมบัติเรื่องความไม่แปรผันของประชากร จากผลการวิจัยดังกล่าวแสดงให้เห็นว่าในการ เปรียบเทียบโดยใช้แบบทดสอบที่มีความยากคล้ายกันกับกลุ่มตัวอย่างที่มีความสามารถใกล้เคียงกัน สามารถนำแบบจำลองของราล์ซมาใช้ได้อย่างเหมาะสม การเทียบคะแนนโดยใช้แบบจำลองของ ราล์ซมีข้อดีเหนือกว่าการใช้วิธีการดั้งเดิม 2 ประการ คือ 1) สามารถจัดแบบทดสอบฉบับย่อย ให้เหมาะสมกับกลุ่มผู้สอบ และ 2) สามารถพัฒนาแบบทดสอบที่เท่าเทียมกันในเชิงสถิติแม้ว่า ค่าความยากจะแตกต่างกัน นอกจากนี้นงานวิจัยนี้ยังยกประเด็นคำถามเกี่ยวกับข้อจำกัดในเรื่อง ความไม่แปรผันของประชากรโดยใช้แบบจำลองของราล์ซ ประเด็นคำถามนี้ไวท์ลี และเดวิส (White and Dawis. 1974) แนะนำว่า เมื่อแบบทดสอบย่อยมีค่าความยากแตกต่างกัน ค่าประมาณความสามารถมีค่าไม่คงที่นัก ส่วนทินสเลย์ และเดวิส ซึ่งประเด็นปัญหาในกรณีที่กลุ่ม ตัวอย่างมีขนาดเล็ก และกรณีที่กลุ่มตัวอย่างมีระดับความสามารถแตกต่างกันมาก ข้อจำกัดใน การนำแบบจำลองของราล์ซมาใช้นี้เห็นได้จากการเทียบคะแนนตามแนวทางซึ่งกลุ่มตัวอย่างที่ศึกษา

มีระดับความสามารถต่างกัน และแบบทดสอบที่ใช้มีความยากต่างกัน ดังจะพบได้จากงานวิจัยของสไลด์และลินน์ (Slind and Linn, 1977) ซึ่งพบว่าการเทียบคะแนนตามแนวตั้งโดยวิธีอีควิเปอร์เซ็นไทล์ให้ผลแตกต่างอย่างมากในการเทียบระดับชั้น และในสเกลของผู้สอบกลุ่มเดิมที่สอบแบบทดสอบต่างระดับชั้น ผู้วิจัยจึงแนะนำให้ใช้แบบจำลองการตอบสนองข้อสอบในการเทียบคะแนนตามแนวตั้ง ต่อมาสไลด์และลินด์ (1978) ศึกษาการเทียบคะแนนตามแนวตั้งโดยใช้แบบจำลองของราล์ชโดยศึกษาชั้นงานเดิมและขยายประเด็นการศึกษาเพิ่มเติม จากงานวิจัยของไวท์ลีและเดวิส (1974) และงานของไรท์ (1968) ซึ่งไวท์ลีและเดวิสตั้งข้อสังเกตว่าความแปรปรวนของคะแนนความแตกต่างระหว่างแบบทดสอบย่อยฉบับง่ายและฉบับยากสูงกว่าค่าคาดหวังของความคลาดเคลื่อนจากการวัดเล็กน้อย นอกจากนี้สไลด์และลินศึกษาความคงที่ของการเทียบคะแนนเมื่อนำค่าประมาณความยากที่ได้จากกลุ่มตัวอย่างหนึ่งไปใช้กับกลุ่มตัวอย่างที่ต่างไป ทั้งนี้เนื่องจากในทางปฏิบัติกลุ่มตัวอย่างที่นำมาเทียบคะแนนอาจมีความสามารถแตกต่างกันมาก สไลด์และลินศึกษาโดยใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนซึ่งแบ่งเป็นฉบับง่ายและฉบับยาก กลุ่มตัวอย่างเป็นนักศึกษาระดับวิทยาลัยปีที่ 1 จำนวน 1365 คน ซึ่งแบ่งเป็นกลุ่มที่มีความสามารถสูง ปานกลาง และต่ำ โดยใช้คะแนนจากแบบทดสอบฉบับง่ายเป็นเกณฑ์ กลุ่มสูงและกลุ่มต่ำใช้ในการประมาณค่าพารามิเตอร์ของความสามารถและของข้อสอบ การศึกษาจะพิจารณาจากค่าสถิติของความแตกต่าง เมื่อเปรียบเทียบค่าประมาณของความสามารถที่ตรงกันจากกลุ่มผู้สอบที่มีความสามารถสูงและกลุ่มผู้สอบที่มีความสามารถต่ำ ผลการศึกษาสนับสนุนข้อค้นพบของไวท์ลี เดวิสและไรท์ อย่างไรก็ตามมีข้อสรุปที่แตกต่างเมื่อนำค่าสถิติที่ได้ไปใช้กับกลุ่มตัวอย่างอื่น ค่าประมาณความสามารถที่ได้จากแบบทดสอบย่อย 2 ฉบับต่างกันถึง 1.2 โลจิท กล่าวคือคะแนนที่ได้จากการเทียบคะแนน ณ ระดับความสามารถต่าง ๆ แปรผันไปตามระดับความสามารถของกลุ่มตัวอย่างที่นำมาใช้ในการเทียบคะแนน ข้อสรุปนี้แสดงให้เห็นอย่างชัดเจนว่าแบบจำลองของราล์ชละเมิดข้อตกลงเบื้องต้นเรื่องความไม่แปรผันของประชากร

ต่อมาสไลด์และลิน (1978) ศึกษาชั้นงานวิจัยเดิมกับนักเรียนระดับ 5 โดยจัดเป็นผู้สอบเป็น 3 กลุ่ม ตามความสามารถที่ได้จากแบบทดสอบความเข้าใจในการอ่าน (the California Tests of Basic Skill Reading Comprehension subtest) แต่ใช้ผลการวิเคราะห์จากแบบทดสอบการอ่าน SRA การศึกษาใช้วิธีการเดิมยกเว้นการประมาณค่าความสามารถ และพารามิเตอร์ของข้อสอบใช้กลุ่มตัวอย่างผู้สอบที่มีความสามารถระดับกลาง ข้อสรุปโดยทั่วไปสนับสนุนการศึกษาครั้งที่ผ่านมา ยิ่งไปกว่านั้นเมื่อใช้กลุ่มผู้สอบที่มีความสามารถต่ำ ค่าประมาณความสามารถที่ได้มีความแตกต่างกันมาก ผลสรุปใช้ได้เฉพาะกรณีผู้สอบมีความสามารถสูงและปานกลาง ผู้วิจัยจึงสรุปว่าการเดามีผลในกรณีผู้สอบความสามารถต่ำ ซึ่งการนำแบบจำลองของราล์ชไปใช้จะให้ผลดีกว่าในสถานการณ์สอบที่มีการเดาน้อย

นักวิจัยหลายคนได้ศึกษาความเที่ยงตรงของการเทียบคะแนนตามแนวตั้ง โดยใช้แบบจำลองของราล์ซเมื่อความแตกต่างระหว่างกลุ่มของข้อสอบและกลุ่มของผู้สอบมีค่าน้อย (Less Extreme) ลอยด์และฮูเวอร์ (Lloyd and Hoover. 1980) ศึกษาโดยใช้แบบทดสอบทักษะการคำนวณพื้นฐาน (The Iowa Tests of Basic Skills Math Computations Test) กับกลุ่มตัวอย่างนักเรียนระดับ 6 ถึง 8 การเทียบคะแนนเทียบทั้งระดับชั้นที่อยู่ติดกันและข้ามระดับชั้น ข้อสรุปจากกราฟของคะแนนสังเกตที่เทียบแล้วสนับสนุนงานวิจัยของสไลด์และลิน ที่ว่าการเทียบคะแนนระหว่าง 2 ระดับใด ๆ ได้รับอิทธิพลจากกลุ่มผู้สอบที่นำมาใช้ในการประมาณค่าพารามิเตอร์ เมื่อพิจารณาสาเหตุของการขาดความเพียงพอในการเทียบคะแนนโดยใช้แบบจำลองของราล์ซ ลอยด์และฮูเวอร์สนใจว่า เนื้อหาวิชาตามหลักสูตรของระดับชั้นที่ต่างกัน โดยเฉพาะวิชาคณิตศาสตร์นั้นอาจไม่ได้วัดคุณลักษณะเดียวกัน เขาจึงศึกษาเรื่องนี้โดยใช้วิธีวิเคราะห์องค์ประกอบ ผลการวิเคราะห์พบว่าข้อสอบจากคลังข้อสอบ (Item Pool) แสดงว่าวัดเนื้อหามากกว่า 1 มิติ และเนื่องจากข้อตกลงเบื้องต้นของแบบจำลองการตอบสนองข้อสอบกำหนดว่าข้อสอบต้องวัดมิติเดียว ดังนั้นสำหรับเนื้อหาวิชาบางประเภทอาจจะเทียบคะแนนกันไม่ได้ เพราะเนื้อหามีการเปลี่ยนแปลงจากระดับหนึ่งไปอีกระดับหนึ่ง

ดิฟกี (Divgi. 1981) ศึกษาการเทียบคะแนนตามแนวตั้งโดยใช้แบบจำลองของราล์ซ ข้อมูลการวิจัยได้จากแบบทดสอบการอ่าน แบบทดสอบซึ่งประกอบด้วยข้อสอบ 60 ข้อ ถูกแบ่งเป็น 2 ฉบับย่อย คือฉบับง่ายและฉบับยาก การวิจัยเทียบคะแนนแบบทดสอบย่อย 2 ฉบับดังกล่าว ค่าสถิติของความแตกต่างมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานเป็น $-.037$ และ 1.06 โลจิทตามลำดับ ค่าดังกล่าวเป็นดัชนีชี้ความสมเหตุสมผลของการวิจัย อย่างไรก็ตามเมื่อนำค่าความแตกต่างของค่าประมาณความสามารถจากแบบทดสอบย่อยทั้งสองมาพล็อตกราฟเพื่อแสดงฟังก์ชันของการทำนายคะแนนรวมจากแบบทดสอบ ผลการวิจัยแสดงให้เห็นความลำเอียงในการเทียบคะแนน ผู้สอบที่มีความสามารถสูงและต่ำจะได้รับคะแนนเทียบเท่าที่สูงกว่าเมื่อใช้แบบทดสอบฉบับที่ต่างไป สเกลของแบบจำลองของราล์ซเหมาะสมกับผู้สอบที่มีความสามารถปานกลางที่สอบแบบทดสอบย่อยฉบับง่าย ผลการวิจัยครั้งนี้สนับสนุนงานวิจัยที่ผ่านมาในเรื่องความไม่เพียงพอของการเทียบคะแนนโดยใช้แบบจำลองของราล์ซ อย่างไรก็ตาม จากงานวิจัยนี้มีข้อควรระวังเกี่ยวกับการใช้ค่าสถิติของความแตกต่างในการประเมินผลการเทียบคะแนนที่ระดับความสามารถของผู้สอบระดับต่าง ๆ

ฟิลลิป (Phillips. 1983 : 267 - 281) ศึกษาเปรียบเทียบวิธีการเทียบคะแนนอีควิเปอร์เซ็นไทล์ วิธีใช้แบบจำลองของราล์ซ วิธีปรับขยายแบบจำลองโลจิสติก แบบ 1 และ 2 พารามิเตอร์ (กำหนดให้การเดามีค่าคงที่) เพื่อเทียบคะแนนแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนแบบหลายระดับชั้น โดยใช้แบบทดสอบมาตราส่วน (Scaling Test) และใช้กลุ่ม

ตัวอย่างขนาดเล็กประมาณ 300-500 คน แบบทดสอบเทียบคะแนนประกอบด้วยแบบทดสอบย่อย 3 ฉบับ ได้แก่แบบทดสอบด้านการอ่าน ด้านภาษาและด้านคณิตศาสตร์ แบบทดสอบมาตราส่วนประกอบด้วยข้อสอบจำนวนน้อยข้อซึ่งมีเนื้อหา และระดับความยากเหมาะสมกับแต่ละระดับชั้น การเทียบคะแนนจะเทียบจากแบบทดสอบแต่ละระดับชั้นไปยังแบบทดสอบมาตราส่วน ผลการวิจัยพบว่า ในสถานการณ์ที่ใช้กลุ่มตัวอย่างขนาดเล็กและใช้แบบทดสอบมาตราส่วนในการเทียบคะแนนแบบทดสอบหลายระดับชั้น วิธีใช้ทฤษฎีการตอบสนองข้อสอบ โดยปรับขยายแบบจำลองที่มีพารามิเตอร์ 1 และ 2 ตัว เป็นวิธีการที่สามารถนำมาใช้แทนวิธีอีควิเปอร์เซ็นไทล์ นอกจากนี้เขาสรุพบว่าแบบจำลองของราล์ซให้ผลที่ไม่น่าพอใจ เมื่อนำมาใช้ในการเทียบคะแนนตามแนวตั้งในกรณีที่ศึกษาวิจัยดังกล่าว

สเกกส์และลิสสิท (Cook and Eignor, 1989:168 citing Skaggs and Lissitz, 1986) ศึกษาความสามารถของแบบจำลองของราล์ซและแบบจำลองสามพารามิเตอร์ในการเทียบคะแนนตามแนวตั้ง การวิจัยครั้งนี้ใช้การจำลองข้อมูลเพื่อสร้างตัวแปรจำนวนหนึ่ง กลุ่มตัวอย่างที่มีระดับความสามารถต่างกันจะทำแบบทดสอบ 2 ฉบับ เพื่อเปรียบเทียบกันประเด็นที่สนใจศึกษาคือความไม่แปรผันของวิธีการเทียบคะแนน 3 วิธีที่ใช้กับผู้สอบที่มีระดับความสามารถที่เฉพาะ ผู้วิจัยพบว่า ผลการวิจัยสนับสนุนข้อค้นพบที่ผ่านมาที่ว่าแบบจำลองของราล์ซไม่เพียงพอสำหรับการเทียบคะแนนตามแนวตั้งถ้าข้อมูลไม่เหมาะสมกับแบบจำลองอย่างแท้จริง วิธีอีควิเปอร์เซ็นไทล์และวิธีใช้แบบจำลอง 3 พารามิเตอร์ ให้ผลการเทียบคะแนนที่ดีกว่าเมื่อเปรียบเทียบกับการใช้วิธีการเทียบคะแนนจริง

จากผลการวิจัยที่กล่าวมา แสดงให้เห็นว่าการใช้แบบจำลองของราล์ซกับการเทียบคะแนนตามแนวตั้งให้ผลการวิจัยที่ไม่น่าพอใจ สเกกส์และลิสสิท (Skaggs and Lissitz, 1986 : 509) แนะนำว่าการนำแบบจำลองของราล์ซ ไปใช้ในการเทียบคะแนนควรตรวจสอบความเหมาะสมของแบบจำลองกับข้อมูลในการเทียบคะแนนตามแนวนอน แต่ไม่ควรใช้แบบจำลองของราล์ซกับการเทียบคะแนนตามแนวตั้ง

ในช่วงที่มีการศึกษาวิจัยเกี่ยวกับการนำแบบจำลองของราล์ซมาใช้ในการเทียบคะแนนนั้น นักวิจัยส่วนหนึ่งก็ได้ศึกษาวิจัยเกี่ยวกับการนำแบบจำลองการตอบสนองแบบอื่นมาใช้เช่นกัน แบบจำลองที่มีการศึกษากันมาก ได้แก่ แบบจำลองโลจิสติก แบบสามพารามิเตอร์ ประเด็นการศึกษาวิจัยเน้นที่การเปรียบเทียบวิธีการเทียบคะแนนที่ต่างกันโดยใช้ข้อมูลชุดเดียวกัน สิ่งที่ศึกษาเปรียบเทียบกันมากที่สุดคือเปรียบเทียบแบบจำลองแบบสามพารามิเตอร์กับแบบจำลองของราล์ซ ดังจะเห็นได้จากงานวิจัยต่อไปนี้

มาร์โค ปีเตอร์สันและสตีวต (Marco, Petersen and Stewart, 1983) ศึกษาโดยใช้ข้อมูลจากแบบทดสอบความถนัดเชิงวิชาการ ด้านภาษา การศึกษาครั้งนี้มีประเด็นที่ศึกษา

หลายประเด็น วิธีการเทียบคะแนนที่ใช้มีหลายวิธี ได้แก่ วิธีเชิงเส้นตรง 4 วิธี วิธีอีคิว-เปอร์เซ็นต์ไทล์ 2 วิธี วิธีเทียบคะแนนโดยใช้แบบจำลองโลจิสติกแบบหนึ่ง และสามพารามิเตอร์ วิธีเทียบคะแนนดังกล่าวนำมาใช้ภายใต้เงื่อนไขต่าง ๆ ได้แก่ กลุ่มตัวอย่างคล้ายคลึงกันและต่างกัน แบบทดสอบร่วมภายในและภายนอก คะแนนเกณฑ์และค่าสถิติรวมที่ต่างกัน การศึกษาครั้งนี้แบ่งเป็น 2 โครงการ โครงการแรกเป็นการเทียบคะแนนแบบทดสอบกับแบบทดสอบฉบับเดิม ในลักษณะของการเทียบคะแนนตามแนวนอน และโครงการที่สองเป็นการเทียบคะแนนแบบทดสอบที่มีความยากต่างกัน ในลักษณะของการเทียบคะแนนตามแนวตั้ง ในทุกสถานการณ์นี้ใช้แบบทดสอบร่วม โดยกลุ่มตัวอย่างกลุ่มที่หนึ่งสอบแบบทดสอบฉบับที่ 1 และแบบทดสอบร่วม เช่นเดียวกับกลุ่มตัวอย่างที่สองซึ่งสอบแบบทดสอบฉบับที่สองและแบบทดสอบร่วม การประเมินความเพียงพอของการเทียบ-คะแนนนี้ใช้สถิติ 2 อย่าง คือ ความคลาดเคลื่อนรวม หรือค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน ซึ่งเป็นค่าเฉลี่ยของผลรวมกำลังสองของความแตกต่างระหว่างคะแนนเกณฑ์กับคะแนนแปลงที่เท่าเทียมกัน เกณฑ์การประเมินอีกอย่างหนึ่งคือค่าความลำเอียงยกกำลังสอง ซึ่งเป็นค่าเฉลี่ยกำลังสองของความแตกต่างระหว่างคะแนนเกณฑ์กับคะแนนจากการสอบที่เท่าเทียมกัน ซึ่งค่าความลำเอียงยกกำลังสองนี้ก็คือส่วนหนึ่งของค่าความคลาดเคลื่อนรวม ในกรณีการเทียบคะแนนตามแนวนอนเกณฑ์คือคะแนนที่ได้จากการเทียบคะแนนกับแบบทดสอบฉบับเดิม (เทียบกับตัวเอง) ส่วนในกรณีการเทียบคะแนนตามแนวตั้ง เกณฑ์ 2 อย่างได้มาจากกลุ่มตัวอย่างขนาดใหญ่ที่ใช้ในการวิจัย เกณฑ์หนึ่งได้จากการเทียบคะแนนด้วยวิธีอีคิวเปอร์เซ็นต์ไทล์โดยตรง อีกเกณฑ์หนึ่งได้จากการเทียบคะแนนด้วยวิธีอีคิวเปอร์เซ็นต์ไทล์โดยใช้คะแนนจริงจากวิธีการตามทฤษฎีการตอบสนอง-ข้อสอบ

ในการเทียบคะแนนตามแนวนอน มาร์โค และคนอื่น ๆ ได้ข้อสรุปว่าเมื่อใช้แบบทดสอบร่วมภายนอก และแบบทดสอบฉบับรวมมีค่าความยากเท่ากันทั้งวิธีเชิงเส้นตรง และวิธีตามทฤษฎีการตอบสนองข้อสอบให้ผลดี และเมื่อใช้แบบทดสอบร่วมภายใน วิธีอีคิวเปอร์เซ็นต์ไทล์ให้ผลดีเช่นกัน ผลจากวิธีใช้แบบจำลองของราล์สให้ผลดีกว่าวิธีอื่น ๆ เล็กน้อยเมื่อใช้แบบทดสอบร่วมภายใน เมื่อใช้แบบทดสอบร่วมที่ง่ายหรือยากกว่าแบบทดสอบฉบับรวม การเทียบคะแนนโดยใช้กลุ่มตัวอย่างสุ่มแสดงให้เห็นว่ามีความคลาดเคลื่อนน้อยมาก ในทางตรงข้ามแบบจำลองตามทฤษฎีการตอบสนองข้อสอบให้ผลดีกว่าวิธีการตามทฤษฎีการสอบแบบดั้งเดิมมากในกลุ่มตัวอย่างที่ไม่คล้ายคลึงกัน

ในกรณีของการเทียบคะแนนตามแนวตั้ง วิธีเทียบคะแนนเชิงเส้นตรงมีความคลาดเคลื่อนมากที่สุด รองลงมาคือวิธีใช้แบบจำลองของราล์ส วิธีสามพารามิเตอร์มีความคลาดเคลื่อนน้อยที่สุด เมื่อใช้คะแนนเกณฑ์ตามทฤษฎีการตอบสนองข้อสอบ เมื่อใช้คะแนนเกณฑ์ตามวิธีอีคิวเปอร์เซ็นต์ไทล์พบว่าวิธีอีคิวเปอร์เซ็นต์ไทล์ดีที่สุด

ผลจากการศึกษาของมาร์โคและคณะแสดงให้เห็นข้อดีของการเทียบคะแนนตามแนวนอนตามทฤษฎีการตอบสนองข้อสอบในกรณีกลุ่มตัวอย่างไม่ได้มาจากการสุ่มซึ่งกรณีเช่นนี้พบได้มากในทางปฏิบัติ สำหรับการเทียบคะแนนตามแนวดิ่ง วิธีเทียบคะแนนโดยใช้แบบจำลองของราล์สมีความคลาดเคลื่อนรวมสูงมาก ผลการวิจัยนี้สอดคล้องกับการศึกษาของสไลด์และลิน (1977, 1978) และการศึกษาของลอร์ดและชวอร์ (1980) ข้อค้นพบเกี่ยวกับการนำวิธีเทียบคะแนนอีคิวเปอร์เซ็นไทล์ไปใช้ยังคลุมเครือ การนำวิธีเทียบคะแนนอีคิวเปอร์เซ็นไทล์ไปใช้มีข้อสรุปชัดเจน ในกรณีที่ประเมินด้วยเกณฑ์อีคิวเปอร์เซ็นไทล์ ประเด็นนี้ขัดแย้งกับการศึกษาของสไลด์และลิน (1977) อย่างไรก็ตามมีข้อสังเกตว่าการวิจัยทั้งสองเรื่องใช้แบบทดสอบต่างฉบับกัน ซึ่งเป็นองค์ประกอบที่สำคัญอย่างหนึ่งในการเปรียบเทียบผลการเทียบคะแนน

เมื่อพิจารณาความคลาดเคลื่อนรวมของการเทียบคะแนนตามแนวดิ่ง พบว่าไม่ว่าจะใช้เกณฑ์ใดในการพิจารณา แบบจำลองแบบสามพารามิเตอร์ให้ผลดีกว่าแบบจำลองของราล์ส เมื่อเทียบคะแนนแบบทดสอบฉบับที่มีความยากปานกลางกับฉบับง่ายพบว่าแบบจำลองของราล์สมีความลำเอียงน้อยกว่าแบบจำลองสามพารามิเตอร์ และเมื่อเทียบคะแนนระหว่างแบบทดสอบฉบับยากปานกลางกับฉบับยากมาก พบว่าแบบจำลองของราล์สมีความลำเอียงมากกว่าแบบจำลองสามพารามิเตอร์ ส่วนการเทียบคะแนนแบบทดสอบฉบับง่ายกับฉบับยากมาก พบว่าแบบจำลองสามพารามิเตอร์มีค่าความลำเอียงและค่าความคลาดเคลื่อนรวมน้อยกว่า ซึ่งอาจจะเป็นไปได้ที่ว่าในแบบทดสอบฉบับที่ง่ายกว่า การเดาจะเกิดขึ้นน้อยกว่า ดังนั้นแบบจำลองของราล์สอาจจะเหมาะสมกับข้อมูลมากกว่าเมื่อใช้แบบทดสอบฉบับที่ยากขึ้น ข้อสรุปนี้สอดคล้องกับงานวิจัยเกี่ยวกับแบบจำลองของราล์สที่ผ่านมา (Guskey, 1981) เมื่อการเดามีผลน้อยมาก

โคเลน (Kolen, 1981 : 1-11) ได้ศึกษาเพิ่มเติมจากผลการศึกษาของมาร์โคและคณะ โดยใช้วิธีเทียบคะแนนหลายวิธีทั้งวิธีการตามทฤษฎีการตอบสนองข้อสอบวิธีเชิงเส้นตรงและวิธีอีคิวเปอร์เซ็นไทล์ โคเลนเทียบคะแนนแบบทดสอบพัฒนาการทางการศึกษา (The Iowa Tests of Educational Development : ITED) 2 ฉบับย่อยโดยเทียบคะแนนแบบทดสอบฉบับใหม่ไปยังฉบับเก่า มีแบบทดสอบฉบับหนึ่งซึ่งมีความยาก 2 ระดับ การวิจัยครั้งนี้ดำเนินการเทียบคะแนนทั้งแนวนอนและแนวดิ่งเพื่อเปรียบเทียบผลของแบบจำลอง หนึ่ง สอง และ สามพารามิเตอร์ การเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบศึกษาโดยใช้วิธีการเทียบคะแนนจริงและวิธีการเทียบคะแนนสังเกต นอกจากนี้โคเลนยังศึกษาโดยใช้ข้อมูลจากกลุ่มตัวอย่างสอบทานผลและค่าสถิติเฉพาะที่กำหนดขึ้นเพื่อประเมินผล แบบทดสอบที่ใช้ไม่มีข้อสอบร่วมกันเลย แบบทดสอบแต่ละฉบับจัดให้ผู้สอบทำโดยการสุ่ม ดังนั้นในแบบแผนการวิจัยจึงไม่มีการวัดซ้ำเลย แต่เนื่องจากกลุ่มตัวอย่างทั้งหมดได้มาโดยการสุ่ม จึงคาดหวังว่าการแจกแจงความสามารถของกลุ่มผู้สอบจะเหมือนกัน สถิติที่ใช้ในการประเมินคือ ค่าเฉลี่ยกำลังสองของความแตกต่างระหว่างคะแนนที่ได้

จากการสอบกับคะแนนแปลงที่ได้จากการเทียบคะแนนในกลุ่มตัวอย่างสอบทานผล

ผลการวิจัยของโคเลน พบว่าสำหรับการเทียบคะแนนตามแนวตั้ง วิธีอีควิเปอร์เซ็นไทล์ และวิธีสามพารามิเตอร์ให้ผลดี วิธีเชิงเส้นตรง และวิธีใช้แบบจำลองของราล์ซให้ผลดีน้อยกว่าสองวิธีแรก สำหรับวิธีการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ 2 วิธี วิธีประมาณค่าจากคะแนนสังเกตให้ผลดีกว่าวิธีใช้คะแนนจริงเล็กน้อย ส่วนการเทียบคะแนนตามแนวนอนผลการวิจัยไม่คงที่ ผลการวิจัยมีความแตกต่างกันมากระหว่างแบบทดสอบฉบับต่าง ๆ โดยทั่วไปวิธีใช้ค่าประมาณคะแนนจริงสำหรับแบบจำลองสามพารามิเตอร์ให้ผลดีที่สุด แต่วิธีเชิงเส้นตรงก็ให้ผลดีเช่นกัน

การวิจัยนี้สนับสนุนการวิจัยที่ผ่านมาซึ่งพบว่าการใช้แบบจำลองของราล์ซไม่เพียงพอในการเทียบคะแนนตามแนวตั้ง ผลการวิจัยของมาร์โคและโคเลนเสนอแนะว่า แบบจำลองแบบสามพารามิเตอร์ใช้ได้ดีกว่าในสถานการณ์ต่าง ๆ แต่การใช้วิธีเทียบคะแนนจริงอาจมีปัญหาการประมาณค่าคะแนนจริงของค่าพารามิเตอร์ที่อยู่ต่ำกว่าระดับการเดา ซึ่งทำให้วิธีใช้คะแนนสังเกตน่าจะสมเหตุสมผลมากกว่าในการเทียบคะแนนตามแนวตั้ง

เมื่อพิจารณาการเทียบคะแนนตามแนวตั้ง งานวิจัยส่วนใหญ่แสดงให้เห็นว่าแบบจำลองสามพารามิเตอร์มีข้อดีเหนือกว่าแบบจำลองของราล์ซ แต่ก็ยังมีปัญหาในส่วนของคะแนนที่อยู่ต่ำกว่าระดับการเดา งานวิจัยที่ให้ผลขัดแย้งกับข้อสรุปนี้พบในงานวิจัยของเพเทียนซึ่งจะกล่าวถึงต่อไป

เพเทียน (Patience, 1981) ศึกษาการเทียบคะแนนตามแนวตั้งโดยใช้แบบทดสอบซึ่งแบ่งเป็นฉบับย่อยที่ ยาก ยากง่ายปานกลาง และง่าย โดยใช้ผลการตอบของนักเรียนระดับ 12 แบบทดสอบฉบับยากใช้กับนักเรียนระดับ 11 แบบทดสอบฉบับยากง่ายปานกลางใช้กับนักเรียนระดับ 10 และแบบทดสอบฉบับง่ายใช้กับนักเรียนระดับ 9 แบบทดสอบย่อยระดับที่อยู่ติดกันมีข้อสอบรวมภายใน 6 ข้อ ผู้สอบแต่ละระดับมี 1000 คน วิธีเทียบคะแนนที่ใช้คือ วิธีอีควิเปอร์เซ็นไทล์ วิธีราล์ซ วิธีสอง และ สามพารามิเตอร์ การเทียบคะแนนเทียบจากแบบทดสอบย่อยฉบับยากและฉบับง่ายไปยังแบบทดสอบย่อยฉบับยากง่ายปานกลาง ดัชนีที่ใช้ในการประเมินผลการเทียบคะแนนคือค่าสหสัมพันธ์ระหว่างคะแนนสอบกับคะแนนที่ผ่านการเทียบแล้ว ในการวิจัยนี้ไม่ได้คำนวณค่าสหสัมพันธ์ระหว่างแบบจำลองสามพารามิเตอร์กับแบบจำลองอื่นอีก 3 แบบจำลอง แม้ว่าจะใช้คะแนนที่อยู่ต่ำกว่าระดับการเดาด้วยแล้ว สหสัมพันธ์ของแบบจำลองสามพารามิเตอร์ก็ต่ำกว่าแบบจำลองอื่น ๆ เพเทียนให้เหตุผลหลายประการในเรื่องนี้ได้แก่ ความสั้น-ยาวของแบบทดสอบรวม ขนาดตัวอย่างมีขนาดเล็ก และการขาดคุณสมบัติการวัดมิติเดียว ยิ่งกว่านั้นค่าสหสัมพันธ์ยังสามารถบ่งความลำเอียงในการเทียบคะแนน ผลการวิจัยครั้งนี้ให้ข้อเสนอแนะว่ามืองค์ประกอบหลายอย่างที่มีอิทธิพลต่อผลการเทียบคะแนนและวิธีการที่นำมาใช้ประเมินผลการเทียบคะแนนเป็นสิ่งที่สำคัญมาก

มีงานวิจัยอีกหลายเรื่องซึ่งศึกษาเปรียบเทียบวิธีการที่ใช้ในการเทียบคะแนน ซึ่งได้แก่ งานวิจัยของ คูก ดันบาร์และไอเนอร์ (Cook, Dunbar and Ignor. 1981) โคเวล (Cowell. 1981) โคลเลนและไวท์นีย์ (Kolen and Whitney. 1982) ปีเตอร์สัน คูก และสต็อกคิงส์ (Petersen, Cook and Stocking. 1983) การวิจัยแต่ละเรื่องมีประเด็น ปัญหาการศึกษาแตกต่างกัน

จากงานวิจัยของเพเทียน (Patience. 1981) ซึ่งเสนอแนะว่า กลุ่มตัวอย่าง ขนาดเล็กเป็นองค์ประกอบที่มีผลต่อการเทียบคะแนนโดยแบบจำลองสามพารามิเตอร์เนื่องจาก กระบวนการแปลงค่าทำได้ยาก โคเวลจึงนำขนาดตัวอย่างมาเป็นตัวแปรในการศึกษาผลการเทียบ คะแนนแบบทดสอบภาษาอังกฤษ (TOEFL) เขาศึกษาเปรียบเทียบผลโดยใช้กลุ่มตัวอย่างขนาดใหญ่ ประมาณ 2000-3000 คน และกลุ่มตัวอย่างขนาดเล็กประมาณ 300 คน และใช้คะแนนที่เท่า เทียมกัน ซึ่งได้จากการแปลงเชิงเส้นตรงเป็นคะแนนเกณฑ์ จากการเทียบคะแนนแบบทดสอบ TOEFL หลาย ๆ ฉบับ ความแตกต่างของผลการเทียบคะแนนโดยใช้วิธีการเทียบคะแนนต่าง ๆ และโดยใช้กลุ่มตัวอย่างขนาดต่างๆ มีขนาดเล็กมาก แบบทดสอบที่ใช้แต่ละฉบับอาจจะมีความยาก คล้ายคลึงกันและกลุ่มตัวอย่างอาจมีความสามารถใกล้เคียงกัน ในการวิจัยมีการศึกษาความคงที่ ของผลการเทียบคะแนนตามแบบจำลองสามพารามิเตอร์โดยใช้กลุ่มตัวอย่างขนาดเล็ก ผลการ ศึกษาปรากฏว่าความแตกต่างระหว่างผลการเทียบคะแนนโดยใช้กลุ่มตัวอย่างขนาดเล็กและขนาด ใหญ่ มีค่าน้อยกว่าความแตกต่างของผลการเทียบคะแนน โดยใช้แบบจำลองหนึ่งพารามิเตอร์ และแบบจำลองสามพารามิเตอร์ การที่ผลการวิจัยเป็นเช่นนี้ก็อาจเป็นเพราะใช้แบบจำลองเชิง เส้นตรงเป็นเกณฑ์ ดังนั้นความแตกต่างระหว่างวิธีการเทียบคะแนนจึงมากกว่าความคลาดเคลื่อน ของการเทียบโดยส่วนรวม

ในทางตรงกันข้าม โคลเลนและไวท์นีย์ (Kolen and Whitney. 1982:279-292) ศึกษาโดยใช้แบบทดสอบ GED (General Education Development Tests) กับกลุ่ม ตัวอย่างขนาดเล็ก (170-198 คน) และใช้วิธีเทียบคะแนนตามแบบจำลองสามพารามิเตอร์ ผลการวิจัยพบว่า ความคลาดเคลื่อนของการเทียบคะแนนโดยใช้วิธีการประมาณค่าพารามิเตอร์ ของแบบจำลองสามพารามิเตอร์จากแบบทดสอบร่วม (เพื่อนำค่าไปใช้ในการเทียบคะแนนกับ แบบทดสอบฉบับอื่นๆ) มีนัยสำคัญ ผลการวิจัยครั้งนี้สอดคล้องกับผลการวิจัยของเพเทียนดังได้ กล่าวมาแล้ว

จากผลการวิจัยดังกล่าวข้างต้น ไม่สามารถสรุปได้ว่า วิธีการเทียบคะแนนด้วยแบบ-จำลองสามพารามิเตอร์ให้ผลการเทียบคะแนนเหนือกว่าวิธีการตามทฤษฎีการตอบสนองข้อสอบ ด้วยแบบจำลองอื่น ๆ และวิธีการตามทฤษฎีการสอบแบบดั้งเดิม จากการวิจัยพบว่าองค์ประกอบ หลายอย่างให้ผลปะปนกัน (confound) เช่นปัญหาการประมาณค่าพารามิเตอร์ แบบแผนการ

รวบรวมข้อมูลที่แตกต่างกัน วิธีการเชื่อมต่อกำพารามิเตอร์ที่ต่างกันและการวัดที่มีหลายมิติ การวิจัยในระยะต่อมาจึงเน้นที่การศึกษาประเด็นดังกล่าวแต่ละประเด็นหรือศึกษาประเด็นดังกล่าวร่วมกันว่าจะมีผลต่อผลการเทียบคะแนนอย่างไร

ลอร์ด (Loyd. 1983) ศึกษาการเทียบคะแนนตามแนวคิดโดยใช้แบบทดสอบร่วม 2 แบบ และใช้วิธีการเทียบคะแนนโดยแบบจำลองของราล์ช และแบบจำลองสามพารามิเตอร์ การวิจัยใช้แบบทดสอบ ITBS (the Iowa Tests of Basic Skills Mathematics Concept Subtest) ซึ่งประกอบด้วยเนื้อหาครบถ้วน 6 ระดับชั้น การเทียบคะแนนระดับชั้นที่อยู่ติดกันใช้แบบทดสอบร่วมภายใน สำหรับแบบทดสอบร่วมภายนอกสร้างขึ้นโดยประกอบด้วยข้อสอบที่มาจากระดับชั้นต่าง ๆ ทั้ง 6 ระดับชั้น แบบทดสอบภายนอกนี้ใช้ในการเทียบคะแนนระดับชั้นที่อยู่ติดกันด้วย ผลการวิจัยให้ข้อแนะนำว่า แบบทดสอบร่วมภายนอกให้ผลการเทียบคะแนนที่น่าพอใจมากกว่าเมื่อใช้แบบทดสอบร่วมภายใน ยิ่งไปกว่านั้นแบบจำลองของราล์ช และแบบจำลองสามพารามิเตอร์ให้ผลการเทียบคะแนนที่ต่างกันอย่างชัดเจน โดยเฉพาะที่ระดับความสามารถต่ำ ผลจากการวิจัยนี้เป็นการยืนยันผลการวิจัยที่ผ่านมาเกี่ยวกับการเทียบคะแนนตามแนวคิด แต่ก็แสดงให้เห็นด้วยว่าแบบแผนของการเทียบคะแนนมีอิทธิพลต่อการเทียบคะแนนด้วยเช่นกัน

ปีเตอร์สัน คูก และสต็อกคิงส์ (Petersen, Cook and Stocking. 1983) ศึกษาเปรียบเทียบวิธีการเทียบคะแนนด้วยแบบจำลองสามพารามิเตอร์ วิธีซึ่งเส้นตรง 3 แบบ และวิธีอีคิวเปอร์เซ็นไทล์ แบบทดสอบที่ใช้คือ SAT ฉบับภาษา และฉบับคณิตศาสตร์ ในการวิจัยนี้แบบทดสอบฉบับหนึ่งเทียบคะแนนกับตัวเองโดยผ่านแบบทดสอบฉบับอื่น การเทียบคะแนนแต่ละครั้งทำในลักษณะสายโซ่ โดยใช้แบบทดสอบร่วมภายนอก การเทียบคะแนนด้วยแบบจำลองสามพารามิเตอร์ทำ 3 แบบ ตามวิธีประมาณค่าพารามิเตอร์ของข้อสอบเพื่อเทียบคะแนนจากแบบทดสอบ 2 ฉบับ โดยทำให้ค่าพารามิเตอร์อยู่บนสเกลเดียวกัน วิธีการแรกคือวิธีประมาณค่าพร้อมกัน (Concurrent Calibration) ตามวิธีการนี้พารามิเตอร์ของแบบทดสอบทั้ง 2 ฉบับและพารามิเตอร์ของแบบทดสอบร่วมจะถูกประมาณค่าพร้อมกันครั้งเดียว โดยใช้โปรแกรม LOGIST วิธีการที่สองเรียกว่าวิธีกำหนดค่าความยากให้คงที่ (Fixed b's Method) ตามวิธีการนี้โปรแกรม LOGIST จะประมาณค่าพารามิเตอร์ของแบบทดสอบฉบับที่หนึ่ง และแบบทดสอบร่วม แล้วใช้ค่าพารามิเตอร์ของแบบทดสอบร่วมจากการประมาณค่าครั้งแรก เป็นค่าพารามิเตอร์ของแบบทดสอบร่วม ในการประมาณค่าแบบทดสอบฉบับที่สอง วิธีการที่สาม คือวิธีแปลงค่าโดยใช้โค้งลักษณะข้อสอบ (Characteristic Curve Transformation Method) ตามวิธีการนี้โปรแกรม LOGIST จะประมาณค่าพารามิเตอร์ของแบบทดสอบแต่ละฉบับ รวมทั้งแบบทดสอบร่วมแยกกัน 2 ครั้ง แล้วใช้วิธีการแปลงเชิงเส้นตรงเพื่อแปลงค่าพารามิเตอร์ของ

แบบทดสอบฉบับที่สองไปสู่สเกลของแบบทดสอบฉบับที่หนึ่ง การประเมินผลการเทียบคะแนนใช้วิธีพิจารณาจากความคลาดเคลื่อนรวมและดัชนีความลำเอียงยกกำลังสอง เช่นเดียวกับดัชนีที่มาร์โคและคนอื่น ๆ (Macro et al. 1983) ใช้ ผลการวิจัยแสดงให้เห็นว่าแบบทดสอบฉบับภาษาและฉบับคณิตศาสตร์ให้ผลการตอบสนองที่แตกต่างกันอย่างชัดเจนเมื่อใช้วิธีเทียบคะแนนวิธีเดียวกัน ผลการวิจัยโดยใช้วิธีเทียบคะแนนที่ต่างกันมีความแตกต่างกันมากกว่าผลที่ได้จากแต่ละวิธีและผลการวิจัยแตกต่างกันมากในแบบทดสอบฉบับภาษา สำหรับแบบทดสอบฉบับภาษา พบว่าวิธีการตามแบบจำลองสามพารามิเตอร์ทั้ง 3 วิธีการ ให้ผลการเทียบคะแนนมีความคลาดเคลื่อนต่ำกว่าเมื่อใช้วิธีเทียบคะแนนตามวิธีการดั้งเดิม วิธีการอีควิเปอร์เซ็นไทล์ให้ผลที่แย่ที่สุด และสำหรับแบบทดสอบฉบับคณิตศาสตร์วิธีการเทียบคะแนนทุกวิธีให้ผลที่น่าเชื่อถือ วิธีเชิงเส้นตรง 2 วิธี และวิธีใช้แบบจำลองสามพารามิเตอร์โดยประมาณค่าพร้อมกันให้ผลการเทียบคะแนนที่มีความคลาดเคลื่อนต่ำสุด กรณีที่แบบทดสอบฉบับภาษาและฉบับคณิตศาสตร์ให้ผลไม่สอดคล้องกันนี้ ไม่ปรากฏเหตุผลที่ชัดเจน ทั้งนี้อาจเนื่องจากแบบทดสอบแต่ละฉบับมีความเป็นคู่ขนานแตกต่างกัน หรืออาจเนื่องจากความเหมาะสมของแบบจำลองที่ใช้กับข้อมูลจากแบบทดสอบแต่ละฉบับแตกต่างกัน เป็นต้น แต่อย่างไรก็ตาม ผลการศึกษาในแบบทดสอบฉบับภาษาล้ำคลึงกับผลการศึกษาของมาร์โคและคนอื่น ๆ ซึ่งแนะนำว่ากระบวนการที่เป็นระบบบางอย่างมีผลเกี่ยวข้องกับผลการเทียบคะแนนดังกล่าว

การเทียบคะแนนตามแนวทางโดยใช้วิธีการที่คล้ายกันนี้มีการศึกษาโดยโฮมและดอดี้ โบแกน (Homes and Doody-Bogan. 1983) โดยใช้ข้อมูลจากแบบทดสอบ CTBS (The vocabulary and Reading Comprehension Subtests of Comprehensive Tests of Basic Skill) การวิจัยใช้แบบแผนการวิจัยแบบผู้สอบกลุ่มเดียว (Single Group Design) และผู้สอบหลายกลุ่ม (Multiple Group Designs) ใช้กลุ่มตัวอย่างสอบทานผลเพื่อประมาณค่าความสามารถของผู้สอบแต่ละคน ด้วยวิธีการ 2 วิธี ใช้แบบจำลองสามพารามิเตอร์เพื่อประมาณค่าความสามารถในการเทียบคะแนนทั้งหมด ผลการวิจัยจาก 3 วิธีสอดคล้องกันมาก แต่แบบแผนผู้สอบกลุ่มเดียวให้ผลดีกว่าเล็กน้อย อย่างไรก็ตามมีการศึกษาผลของเนื้อหาวิชาที่ใช้ด้วยเช่นเดียวกับงานวิจัยของปีเตอร์สัน สำหรับแบบทดสอบด้านภาษาวิธีการประมาณค่าพร้อมกันมีความคลาดเคลื่อนน้อยกว่าวิธีการแปลงค่าพารามิเตอร์ แต่สำหรับแบบทดสอบความเข้าใจ ผลการศึกษาคือตรงกันข้าม

สำหรับการศึกษาเปรียบเทียบระหว่างข้อสอบด้านภาษาและด้านตัวเลข ในการวิจัยอื่น โคลเลน (Kolen. 1981) ก็พบว่าแบบทดสอบทั้ง 2 ประเภท ให้ผลที่ไม่สอดคล้องกัน เมื่อพิจารณาจากความเพียงพอของแบบจำลองการเทียบคะแนนที่ต่างกัน โคลเลนและไวท์นีย์ (Kolen and Whitney. 1982) พบว่า แบบทดสอบวัดผลสัมฤทธิ์ประเภทต่าง ๆ ให้ผลแตกต่างกันน้อยกว่า แต่ไม่ได้ตั้งข้อสังเกตเกี่ยวกับความแตกต่างด้านโครงสร้างของแบบทดสอบเหล่านั้น

จากการศึกษาของคุกและไอเนอร์ (Cook and Eignor. 1983) ซึ่งเทียบคะแนนโดยใช้แบบทดสอบวัดผลสัมฤทธิ์วิชาชีพ คณิตศาสตร์และสังคมศึกษาหลายฉบับในการสอบคัดเลือกเข้าศึกษา ใช้แบบแผนการวิจัยแบบสายโซ่คล้ายกับการศึกษาของปีเตอร์สันและคนอื่น ๆ วิธีการเทียบคะแนนที่ใช้ได้แก่ วิธีเชิงเส้นตรง 3 วิธี วิธีอีควิเปอร์เซ็นไทล์ และวิธีเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ แบบจำลองสามพารามิเตอร์ โดยใช้วิธีการประมาณค่าตามแบบจำลองนี้ 2 วิธี คือ วิธีประมาณค่าพร้อมกัน และวิธีแปลงโดยใช้ไค่งลักษณะข้อสอบ ผลการเทียบคะแนนประเมินโดยใช้ค่าเฉลี่ยของความคลาดเคลื่อนยกกำลังสอง และค่าสถิติของความลำเอียง ผลการวิจัยยืนยันผลการวิจัยที่ผ่านมาว่า ผลการเทียบคะแนนแปรผันไปตามเนื้อหาของแบบทดสอบ แต่วิธีอีควิเปอร์เซ็นไทล์เป็นวิธีที่ไม่เพียงพอในทุกกรณี

โจแฮนสัน (Johanson. 1988:488) ศึกษาผลของความยากของแบบทดสอบร่วมความสามารถของผู้สอบ และวิธีการประมาณค่าพารามิเตอร์ที่มีต่อการเทียบคะแนนตามแนวตั้งตามทฤษฎีการตอบสนองข้อสอบ ข้อมูลการศึกษาได้จากการจำลองขึ้นตามแบบจำลองโลจิสติกประชากรมี 3 กลุ่มแต่ละกลุ่มประกอบด้วยกลุ่มตัวอย่าง 2 กลุ่มย่อย การประมาณค่าพารามิเตอร์ของประชากรใช้โปรแกรม LOGIST แบบทดสอบร่วมที่ใช้ มีความยาว 4 ขนาด ในการศึกษาจะศึกษาผลของแบบทดสอบร่วมไปพร้อมกับ ลักษณะประชากร และวิธีการประมาณค่าพารามิเตอร์ ความยากของแบบทดสอบร่วม มีการแจกแจงแบบยูนิฟอร์ม วิธีการประมาณค่าพารามิเตอร์ใช้วิธีประมาณค่าพร้อมกัน วิธีใช้ไค่งลักษณะข้อสอบ วิธีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน วิธีกำลังสองน้อยสุดแบบออร์ทอกอนอลและวิธีกำลังสองน้อยสุดแบบปกติ

ผลการวิจัยพบว่า วิธีเทียบคะแนนโดยใช้ไค่งลักษณะข้อสอบเป็นวิธีที่มีความถูกต้องมากที่สุด นอกจากนี้วิธีเทียบคะแนนโดยใช้ไค่งลักษณะข้อสอบยังเป็นวิธีการเดียวที่ให้ผลการเทียบคะแนนในระดับที่ยอมรับได้โดยใช้แบบทดสอบร่วมจำนวนน้อยที่สุด สำหรับวิธีการเทียบคะแนนทุกวิธีเมื่อใช้แบบทดสอบร่วมที่มีจำนวนมากที่สุดและใช้กลุ่มผู้สอบที่มีความสามารถแตกต่างกันน้อยที่สุดให้ผลการเทียบคะแนนที่ยอมรับได้ เมื่อใช้กลุ่มตัวอย่างที่มีความแตกต่างกันมาก การประมาณค่าพารามิเตอร์ของข้อสอบทำได้ไม่ดัดนัก โดยสรุปแล้วความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ของข้อสอบอาจทำให้ยากในการเทียบคะแนนแบบทดสอบที่มีระดับความยากแตกต่างกันมาก ถ้าแบบทดสอบร่วมที่ใช้สั้นกว่าและแบบทดสอบฉบับรวมมีขนาดสั้น

เพเทียน (Patience. 1990:486) ศึกษาเปรียบเทียบผลของการเทียบคะแนนระหว่างระดับชั้นโดยใช้วิธีเทียบคะแนน 5 วิธี คือวิธีการตามทฤษฎีการสอบแบบดั้งเดิม 2 วิธี ได้แก่วิธี อีควิเปอร์เซ็นไทล์ และวิธีการของเธอร์สโตน และวิธีการตามทฤษฎีการตอบสนองข้อสอบ 3 วิธี ได้แก่ วิธีใช้แบบจำลองโลจิสติกแบบหนึ่ง สอง และสามพารามิเตอร์ โดยในการวิจัยนี้ต้องการศึกษาว่าวิธีการที่ใช้วิธีใดให้ผลการเทียบคะแนนที่เหมาะสมที่สุด แบบทดสอบที่ใช้ใน

การวิจัยแบ่งเป็น 3 ระดับ คือง่าย ยากง่าย ปานกลางและยาก การเทียบคะแนนแบบทดสอบแต่ละระดับใช้วิธีการเทียบคะแนนทั้ง 5 วิธี แบบทดสอบฉบับง่ายใช้กับนักเรียนระดับ 9 แบบทดสอบฉบับยากง่ายปานกลาง ใช้กับนักเรียนระดับ 10 และแบบทดสอบฉบับยากใช้กับนักเรียนระดับ 11 คะแนนที่ได้จากแบบทดสอบแต่ละระดับจะนำไปเปรียบเทียบกับคะแนนโดยรวมของนักเรียน วิธีการเทียบคะแนนที่เหมาะสมที่สุดกำหนดว่าเป็นวิธีการที่ให้ผลการเทียบคะแนนใกล้เคียงกับผลที่ได้จากคะแนนของนักเรียนจากแบบทดสอบรวมทุกระดับ

ผลการวิจัยพบว่าวิธีการเทียบคะแนนทั้ง 5 วิธีให้ผลการเทียบคะแนนที่คล้ายคลึงกันเมื่อใช้แบบทดสอบฉบับง่าย และฉบับยากง่ายปานกลาง เมื่อใช้แบบทดสอบฉบับยากพบว่า ผลการเทียบคะแนนโดยวิธีอีควิเปอร์เซ็นไทล์ วิธีการของเธอร์สโตน และวิธีสามพารามิเตอร์ให้ผลคล้ายคลึงกับคะแนนที่ได้จากแบบทดสอบรวมมากกว่าวิธีหนึ่งพารามิเตอร์ และวิธีสองพารามิเตอร์

แฮริส (Harris. 1991:221-235) ศึกษาผลของแบบแผนการรวบรวมข้อมูล 2 วิธี คือแบบแผนที่ 1 ของแองกอฟ (ผู้สอบกลุ่มเดียวจัดให้สมดุล) และแบบแผนที่ 2 ของแองกอฟ (ผู้สอบกลุ่มเท่าเทียมกัน) ที่มีต่อการเทียบคะแนนตามแนวตั้งโดยใช้วิธีเทียบคะแนน 2 วิธี คือวิธีอีควิเปอร์เซ็นไทล์ และวิธีตามทฤษฎีการตอบสนองข้อสอบแบบสามพารามิเตอร์ แบบทดสอบที่ใช้ คือแบบทดสอบวิชาคณิตศาสตร์ 2 ฉบับ คือ EACT มีจำนวน 60 ข้อ และฉบับ PACT มีจำนวน 40 ข้อ แบบทดสอบทั้ง 2 ฉบับมีเนื้อหาคาบเกี่ยวกับกัน แต่มีเนื้อหาเฉพาะและระดับความยากต่างกัน กลุ่มตัวอย่างเป็นนักเรียนระดับ 11 และ 12 ซึ่งแบ่งเป็นกลุ่มย่อยหลายกลุ่ม เกณฑ์ที่ใช้ในการประเมิน คือ ค่าเฉลี่ยกำลังสองของความแตกต่าง (RMSD) ผลการวิจัยสรุปได้ว่าผลการเทียบคะแนนตามแนวตั้งโดยใช้แบบแผนการเก็บข้อมูลทั้ง 2 แบบมีความเพียงพอและให้ผลคล้ายคลึงกัน โดยรูปแบบที่ใช้ผู้สอบกลุ่มเดียวให้ผลคงที่กว่ารูปแบบที่ใช้ผู้สอบกลุ่มเท่าเทียมกันเล็กน้อย

อเยอเว (Ayerve. 1992:1841) ศึกษาเปรียบเทียบประสิทธิภาพของการเทียบคะแนนตามแนวตั้งด้วยวิธีอีควิเปอร์เซ็นไทล์ และวิธีสามพารามิเตอร์ โดยใช้การจำลองข้อมูลภายใต้เงื่อนไขของขนาดตัวอย่าง ความยาวของแบบทดสอบ และความยากของแบบทดสอบร่วมการเปรียบเทียบทำ 3 กรณี คือ 1) เปรียบเทียบประสิทธิภาพของวิธีการเทียบคะแนนทั้งสองภายใต้เงื่อนไขทุกเงื่อนไข 2) เปรียบเทียบประสิทธิภาพของวิธีการเทียบคะแนนแต่ละวิธีภายใต้เงื่อนไขของตัวแปรอิสระแต่ละตัว และ 3) ตรวจสอบผลของตัวแปรอิสระแต่ละตัวในการเทียบคะแนนแต่ละวิธี ขนาดตัวอย่างที่ใช้มี 3 ขนาด คือ 200 500 และ 1000 ความยาวของแบบทดสอบมี 2 ขนาด คือ 30 ข้อและ 60 ข้อ ส่วนความยาวของแบบทดสอบร่วมมี 2 ขนาด คือ 5 ข้อ และ 10 ข้อ สถิติที่ใช้เป็นเกณฑ์การประเมินมี 2 อย่าง คือ ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนซึ่งถ่วงน้ำหนักแล้ว (Weighted Mean Square Error : WMSE) และ ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนซึ่งไม่ถ่วงน้ำหนัก (UMSE)

ผลการวิจัยพบว่า การเทียบคะแนนตามเงื่อนไขต่าง ๆ โดยที่ใช้วิธีการเทียบคะแนนทั้งสองวิธีไม่แตกต่างกัน สำหรับการเทียบคะแนนโดยวิธีอีควิเปอร์เซ็นไทล์พบว่าความยาวของแบบทดสอบกับความยาวของแบบทดสอบรวมเป็นองค์ประกอบที่สำคัญ ส่วนการเทียบคะแนนตามวิธีสามพารามิเตอร์ พบว่าขนาดตัวอย่างเป็นองค์ประกอบที่สำคัญ ขนาดตัวอย่างที่เล็ก (200) มีแนวโน้มที่จะให้ผลการเทียบคะแนนที่ไม่ถูกต้อง ในขณะที่ขนาดตัวอย่างที่ใหญ่ขึ้น (500 และ 1000) จะให้ผลการเทียบคะแนนที่มีความถูกต้องมากกว่า

สำหรับงานวิจัยในประเทศไทยที่เกี่ยวกับการเทียบคะแนนตามทฤษฎีการตอบสนองข้อสอบ นั้น ชูชีพ พงษ์สมบูรณ์. (2528 : 112-118) ศึกษาเปรียบเทียบประสิทธิภาพของการเทียบคะแนนโดยใช้รูปแบบผู้สอบรวม และรูปแบบแบบทดสอบรวม การเทียบคะแนนใช้วิธีการ 3 วิธี คือวิธีอีควิเปอร์เซ็นไทล์ วิธีเชิงเส้นตรง และวิธีสามพารามิเตอร์ กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2526 แบบทดสอบที่ใช้แบ่งเป็น 2 ฉบับ ๆ ละ 38 ข้อ มีข้อสอบรวมภายใน 20 ข้อ แบบทดสอบทั้ง 2 ฉบับมีความยากใกล้เคียงกัน ผลการวิจัยพบว่าประสิทธิภาพของการเทียบคะแนนระหว่างรูปแบบผู้สอบรวมและรูปแบบแบบทดสอบรวมไม่แตกต่างกัน

ภาวณี ศรีสุวรรณานันท์ (2529 : 155-170) ศึกษาเปรียบเทียบผลของการเทียบคะแนนตามแนวนอนโดยใช้วิธีการเทียบคะแนน 3 วิธีเช่นเดียวกับ ชูชีพ พงษ์สมบูรณ์ แต่ในการวิจัยนี้ใช้ แบบทดสอบรวมที่มีขนาดยาวต่างกัน 3 ขนาด คือ ขนาดร้อยละ 20 (7 ข้อ) ขนาดร้อยละ 40 (14 ข้อ) และขนาดร้อยละ 60 (21 ข้อ) แบบทดสอบทั้งหมดมี 100 ข้อ จัดแบ่งเป็น 2 ฉบับ ๆ ละ 35 ข้อ แบบทดสอบทั้ง 2 ฉบับ มีความยากใกล้เคียงกัน กลุ่มตัวอย่างที่ใช้จำแนกเป็น 2 กรณี คือกรณีการสอบคัดเลือกเพื่อเข้าทำงาน และกรณีการสอบวัดผลสัมฤทธิ์ กลุ่มตัวอย่างแต่ละกรณีมี 2 กลุ่ม ๆ ละ 1500 คน ซึ่งเป็นกลุ่มที่ใช้ในการเทียบคะแนน 1 กลุ่ม และใช้ในการสอบทานผลการเทียบคะแนน 1 กลุ่ม การประเมินผลการเทียบคะแนนแต่ละวิธีจะประเมินจากความเพี้ยนพองของการเทียบคะแนนโดยใช้ดัชนีเปรียบเทียบความแตกต่างที่ได้จากกลุ่มสอบทานผล ผลการวิจัยพบว่า เมื่อใช้แบบทดสอบรวมที่มีขนาดยาวกว่าจะให้ผลการเทียบคะแนนที่มีประสิทธิภาพสูงกว่าเมื่อใช้แบบทดสอบรวมที่มีขนาดสั้นกว่าและให้ผลการเทียบคะแนนที่น่าพอใจกว่า ทั้งกรณีการสอบคัดเลือกและกรณีการสอบวัดผลสัมฤทธิ์ ส่วนการศึกษาค้นคว้าผลการเทียบคะแนนโดยใช้วิธีการที่ต่างกันพบว่า ในกรณีการสอบคัดเลือก วิธีการเทียบคะแนนที่ให้ผลการเทียบคะแนนที่มีความเพี้ยนพองมากที่สุดคือ วิธีอีควิเปอร์เซ็นไทล์ วิธีสามพารามิเตอร์และวิธีเชิงเส้นตรงตามลำดับ แต่ในกรณีการสอบวัดผลสัมฤทธิ์ วิธีการที่ให้ผลการเทียบคะแนนที่มีความเพี้ยนพองมากที่สุด คือวิธีเชิงเส้นตรง วิธีอีควิเปอร์เซ็นไทล์ และวิธีสามพารามิเตอร์ตามลำดับ

เรวดี อินทะสระ (2530 : 64-67) ศึกษาเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน และความเพี้ยนพองจากการใช้วิธีการเทียบคะแนนตามทฤษฎีการตอบสนอง

ข้อสอบ แบบสามพารามิเตอร์ กับวิธีเทียบคะแนนโดยใช้เทคนิคการวิเคราะห์หองค์ประกอบ กลุ่มตัวอย่างในการศึกษา เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2529 จำนวน 2823 คน แบบทดสอบที่ใช้เป็นแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ 2 ฉบับ ๆ ละ 45 ข้อ โดยมีแบบทดสอบรวมภายในฉบับละ 15 ข้อ ผลการวิจัยพบว่าความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนโดยวิธีสามพารามิเตอร์มีน้อยกว่าวิธีใช้เทคนิคการวิเคราะห์หองค์ประกอบ ส่วนความเพี่ยงพของความคลาดเคลื่อนในการเทียบคะแนน โดยวิธีสามพารามิเตอร์อยู่ในระดับที่น่าพอใจ และความเพี่ยงพของการเทียบคะแนนโดยวิธีใช้เทคนิคการวิเคราะห์หองค์ประกอบอยู่ในระดับปานกลาง

พรพิมล นาคเวช (2537 :99-101) ศึกษาคุณภาพของการเทียบคะแนนในแนวตั้ง โดยใช้ทฤษฎีการตอบสนองข้อสอบ สำหรับผู้มีความสามารถต่างกัน 3 ระดับ ซึ่งได้แก่ นักเรียนชั้น ม.1 ม.2 และม.3 โดยวิเคราะห์ความคลาดเคลื่อนมาตรฐานและความเพี่ยงพของการเทียบคะแนน กลุ่มตัวอย่างเป็นนักเรียนชั้น ม.1 ม.2 และ ม.3 สังกัดกองการมัธยมศึกษา กรมสามัญศึกษา จำนวน 6429 คน ที่ได้จากการสุ่มกลุ่มตัวอย่างแบบแบ่งชั้นหลายขั้นตอน แบ่งเป็นกลุ่มตัวอย่างเทียบคะแนนเพื่อนำผลไปเทียบคะแนนระหว่างแบบทดสอบ จำนวน 3279 คน และกลุ่มตัวอย่างสอบทานผลเพื่อนำผลไปตรวจสอบคุณภาพของการเทียบคะแนน จำนวน 3105 คน เครื่องมือที่ใช้เป็นแบบทดสอบคณิตศาสตร์ เรื่องสมการและอสมการ สำหรับระดับชั้น ม.1 ม.2 และ ม.3รวม 3 ฉบับ ๆ ละ 25 ข้อ โดยมีข้อสอบรวมระหว่างระดับชั้นที่ติดกันระดับชั้นละ 5 ข้อ ผลการศึกษาพบว่า ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนมีค่าอยู่ระหว่าง .031-.051 ซึ่งเป็นความคลาดเคลื่อนมาตรฐานที่ค่อนข้างต่ำ และผลการวิเคราะห์ความเพี่ยงพของการเทียบคะแนนตามแนวตั้งพบว่า ค่าดัชนีความแตกต่างของการเทียบคะแนนอยู่ในระดับที่น่าพอใจ

สำหรับการนำวิธีการเทียบคะแนนไปประยุกต์ใช้ในการศึกษา มีผู้นำวิธีการเทียบคะแนนตามแนวตั้งและวิธีการเทียบคะแนนตามแนวนอนไปใช้สร้างตารางเทียบคะแนนแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนและใช้ในการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียน การวิจัยดังกล่าวใช้วิธีการเทียบคะแนนโดยใช้แบบจำลองของราล์ซ ดังนี้

สุรัตน์ ขวัญบุญจันทร์ (2531 : 77-81) สร้างแบบทดสอบและตารางเทียบคะแนนตามแนวนอนวิชาวิทยาศาสตร์ที่วิเคราะห์ด้วยแบบจำลองของราล์ซ แบบทดสอบที่สร้างขึ้นมี 2 ฉบับ ฉบับละ 60 ข้อ แต่ละฉบับมีข้อสอบรวมภายในจำนวน 10 ข้อ แบบทดสอบที่สร้างขึ้นใช้สำหรับการเรียนการสอนวิชาวิทยาศาสตร์ เรื่องพลังงานและการเปลี่ยนแปลง ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 การวิจัยครั้งนี้มีการประเมินผลการเทียบคะแนนโดยใช้ดัชนีความเพี่ยงพของความคลาดเคลื่อนในการเทียบคะแนนโดยวิเคราะห์จากกลุ่มสอบทานผล ผลการตรวจสอบพบว่าความเพี่ยงพของการเทียบคะแนนดังกล่าว อยู่ในระดับที่น่าพอใจมาก

อาภรณ์ กาญจนกิจโสภณ (2531 : 66-69) สร้างแบบทดสอบและตารางเทียบคะแนนตามแนวดิ่ง ในวิชาคณิตศาสตร์ เรื่องสมการและอสมการ สำหรับนักเรียนชั้นมัธยมศึกษาตอนต้นที่วิเคราะห์ด้วย แบบจำลองของราล์ช แบบทดสอบที่สร้างขึ้นมี 3 ฉบับ ๆ ละ 40 ข้อ สำหรับชั้นมัธยมศึกษาปีที่ 1 2 และ 3 ชั้นละ 1 ฉบับ แบบทดสอบแต่ละฉบับมีข้อสอบร่วมระหว่างชั้นฉบับละ 10 ข้อ นอกจากนี้ผู้วิจัยสร้างตารางเทียบคะแนนระหว่างชั้นซึ่งเทียบที่ความสามารถที่เท่ากัน ผลการวิจัยสามารถนำไปใช้ในการศึกษาความก้าวหน้าทางการเรียนเรื่องสมการและอสมการของนักเรียนได้

ในการนำวิธีการเทียบคะแนนไปประยุกต์ใช้ในการศึกษาความก้าวหน้าทางการเรียน สุจินดา ผ่องอักษร (2533 : 143-157) นำวิธีการเทียบคะแนนโดยใช้แบบจำลองของราล์ชไปใช้ในการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนในกลุ่มทักษะของนักเรียนชั้นประถมศึกษาปีที่ 6 ที่เรียนจบตามหลักสูตรประถมศึกษา พุทธศักราช 2521 ในช่วงระยะเวลา 3 ปีการศึกษา (ปีการศึกษา 2529 ถึง 2531) ในการวิจัยครั้งนี้นอกจากการศึกษาค้นคว้าความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนแล้วยังมีการเปรียบเทียบความแตกต่างของผลสัมฤทธิ์ทางการเรียนระหว่างปีการศึกษาต่าง ๆ หลังจากที่ได้เทียบคะแนนแล้ว และเปรียบเทียบสัดส่วนจำนวนนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนผ่านเกณฑ์ที่น่าพอใจระหว่างการเทียบคะแนนกับไม่เทียบคะแนน กลุ่มตัวอย่างการวิจัยเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ตามโครงการประเมินคุณภาพนักเรียนชั้นประถมศึกษาปีที่ 6 ระดับประเทศในจังหวัดปทุมธานี แบบทดสอบที่ใช้เป็นแบบทดสอบวัดผลสัมฤทธิ์ในกลุ่มทักษะ (วิชาคณิตศาสตร์และภาษาไทย) ซึ่งสำนักงานคณะกรรมการการศึกษาแห่งชาติสร้างขึ้นและ ใช้ทดสอบนักเรียนเมื่อปีการศึกษา 2529 2530 และ 2531 แบบแผนการเก็บรวบรวมข้อมูลใช้รูปแบบผู้สอบร่วม เนื่องจากแบบทดสอบที่ใช้ทั้ง 3 ฉบับ ไม่มีข้อสอบร่วมกัน

ผลการวิจัยพบว่า จากการเทียบคะแนนระหว่างแบบทดสอบ 3 ฉบับ คะแนนที่เทียบแล้วของแบบทดสอบ 3 ฉบับซึ่งตรงกับคะแนนดิบค่าเดียวกันมีค่าแตกต่างกันทั้งแบบทดสอบวิชาคณิตศาสตร์และแบบทดสอบวิชาภาษาไทย เมื่อศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนพบว่าผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ในปีการศึกษา 2530 สูงกว่าปีการศึกษา 2529 แต่ลดลงในปีการศึกษา 2531 สำหรับผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย พบว่าอัตราความก้าวหน้าของคะแนนเฉลี่ยเพิ่มขึ้นทุกปีการศึกษา และอัตราความก้าวหน้าของค่าร้อยละนักเรียนที่มีผลน่าพอใจเพิ่มขึ้นทุกปีการศึกษา

จากผลงานวิจัยทั้งต่างประเทศและในประเทศที่ผ่านมาดังกล่าว จะเห็นว่าการวิจัยเกี่ยวกับการเทียบคะแนนส่วนใหญ่ศึกษาวิจัยเทคนิคและวิธีการเทียบคะแนนเพื่อให้ได้ผลการเทียบคะแนนที่ถูกต้องแม่นยำ ผลจากการวิจัยสรุปได้ว่า ไม่มีวิธีการใด เพียงวิธีการเดียวที่ให้ผลดีในทุกสถานการณ์ วิธีการเทียบคะแนนที่ใช้ได้ผลดีที่สุดดูเหมือนจะขึ้นกับองค์ประกอบหลายอย่างได้แก่

ความยาวของแบบทดสอบ เนื้อหาของแบบทดสอบ วิธีการประมาณค่าพารามิเตอร์ วิธีปรับค่าพารามิเตอร์ การกระจายของค่าความยากของข้อสอบ และค่าพารามิเตอร์ความสามารถของกลุ่มตัวอย่างที่ใช้ วิธีการประเมินความเพียงพอของการเทียบคะแนน ซึ่งอาจมีอิทธิพลต่อการแปลความหมายของการเทียบคะแนน แต่อย่างไรก็ตาม ในการเทียบคะแนนตามแนวดิ่ง ผลการวิจัยพบว่า วิธีการตามทฤษฎีการตอบสนองข้อสอบ แบบจำลองโลจิสติก สามพารามิเตอร์ เป็นวิธีการที่ให้ผลดีในหลาย ๆ สถานการณ์ ผู้วิจัยจึงนำวิธีการเทียบคะแนนดังกล่าวมาใช้ในการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 ซึ่งจะได้กล่าวถึงรายละเอียดของวิธีการในบทต่อไป

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยนี้ เป็นการศึกษาเชิงประจักษ์ (Empirical study) มีจุดมุ่งหมายเพื่อศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 โดยใช้วิธีการเทียบคะแนนตามแนวดิ่ง ตามทฤษฎีการตอบสนองข้อสอบ แบบจำลองโลจิสติกสามพารามิเตอร์ รายละเอียดของวิธีดำเนินการวิจัยมีดังนี้

ประชากร

ประชากรในการวิจัยครั้งนี้เป็นนักเรียนชั้น ม.1 ม.2 และ ม.3 ภาคเรียนที่ 1 ปีการศึกษา 2537 ของโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา กระทรวงศึกษาธิการ ในจังหวัดสิงห์บุรี จำนวนรวม 7372 คน เป็นนักเรียนชั้น ม. 1 จำนวน 2507 คน นักเรียนชั้น ม.2 จำนวน 2553 คน และนักเรียนชั้น ม.3 จำนวน 2312 คน ดังแสดงในตาราง 2

กลุ่มตัวอย่าง

จากประชากรการวิจัยดังกล่าว ผู้วิจัยสุ่มกลุ่มตัวอย่างด้วยวิธีการสุ่มอย่างเป็นระบบจากนักเรียนทุกชั้นทุกโรงเรียน เพื่อให้ได้กลุ่มตัวอย่างการวิจัย 2 กลุ่ม คือ

1. กลุ่มตัวอย่างที่ใช้ในการเทียบคะแนน จำนวนร้อยละ 50 ของจำนวนประชากรการวิจัยในแต่ละชั้น แต่ละโรงเรียน ผลการสุ่มกลุ่มตัวอย่างได้นักเรียนชั้น ม.1 ม.2 และ ม.3 จำนวน 1245 1280 และ 1171 คน ตามลำดับ
2. กลุ่มตัวอย่างที่ใช้ในการสอบทานผล จำนวนร้อยละ 30 ของกลุ่มตัวอย่างเทียบคะแนน ผลการสุ่มกลุ่มตัวอย่าง ได้นักเรียนชั้น ม.1 ม.2 และ ม.3 จำนวน 373 384 และ 352 คน ตามลำดับ

จำนวนกลุ่มตัวอย่างเทียบคะแนนและกลุ่มตัวอย่างสอบทานผล แสดงดังตาราง 3

ตาราง 2 จำนวนประชากรการวิจัยจำแนกตามโรงเรียน

โรงเรียน	จำนวนนักเรียน		
	ม.1	ม.2	ม.3
1. สิงห์บุรี	692	679	576
2. สิงห์บุรี"ประสานมิตรอุปถัมภ์"	210	285	281
3. หัวไผ่	89	88	97
4. อินทร์บุรี	290	316	252
5. ศรีวิณีวิทยาคม	126	116	110
6. ทองเอนวิทยา	108	91	93
7. พรหมบุรีรัชดาภิเษก	109	113	102
8. บ้านแป่งวิทยา	109	102	118
9. บางระจันวิทยา	251	254	260
10. ศรีศักดิ์สุวรรณวิทยา	126	165	139
11. ท่าช้างวิทยาการ	196	162	140
12. ค่ายบางระจันวิทยาคม	195	181	143
รวม	2507	2553	2312

ตาราง 3 จำนวนกลุ่มตัวอย่างการวิจัยจำแนกตามโรงเรียน

โรงเรียน	จำนวนนักเรียน					
	กลุ่มเก็บคะแนน			กลุ่มสอบทานผล		
	ม.1	ม.2	ม.3	ม.1	ม.2	ม.3
1. สิงห์บุรี	348	342	290	104	102	87
2. สิงห์บุรี"ประสานมิตรอุปถัมภ์"	103	144	143	31	43	43
3. หัวไผ่	44	44	47	13	13	14
4. อินทร์บุรี	145	160	128	44	48	38
5. ศรีวิณีวิทยาคม	64	59	56	19	18	17
6. ทองเอนวิทยา	54	50	45	16	15	14
7. พรหมบุรีรัชดาภิเษก	55	57	53	17	17	16
8. บ้านแป่งวิทยา	53	52	61	16	16	18
9. บางระจันวิทยา	121	122	132	36	37	40
10. ศรีศักดิ์สุวรรณวิทยา	63	82	71	19	25	21
11. ท่าช้างวิทยาคาร	97	78	72	29	23	22
12. ค่ายบางระจันวิทยาคม	98	190	73	29	27	22
รวม	1245	1280	1171	373	384	352

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยมี 2 ประเภท คือ แบบทดสอบเทียบคะแนนและแบบทดสอบรวม

1. แบบทดสอบเทียบคะแนน เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย ของชั้น ม.1(ท 101) ชั้น ม.2(ท 203) และชั้น ม.3(ท 305) ชั้นละ 1 ฉบับ แบบทดสอบ แต่ละฉบับประกอบด้วยข้อสอบแบบเลือกตอบ 4 ตัวเลือก จำนวน 90 ข้อ แบบทดสอบเทียบคะแนน ทั้ง 3 ฉบับนี้ ผู้วิจัยร่วมกับอาจารย์ประจำวิชาภาษาไทยของกลุ่มโรงเรียนมัธยมศึกษาจังหวัด สิงห์บุรีเป็นผู้สร้าง

ขั้นตอนการสร้างแบบทดสอบเทียบคะแนน มีดังนี้

1.1 คณะกรรมการกลุ่มโรงเรียนมัธยมศึกษาแต่งตั้งคณะกรรมการเพื่อจัดทำแบบ ทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย โดยประกอบด้วยคณะกรรมการออกข้อสอบ คณะกรรมการคัดเลือกข้อสอบ และคณะกรรมการจัดพิมพ์แบบทดสอบ

1.2 คณะกรรมการออกข้อสอบ ออกข้อสอบตามจุดประสงค์การเรียนรู้ที่ คณะกรรมการวิชาการกลุ่มโรงเรียนมัธยมศึกษากำหนด

1.3 คณะกรรมการคัดเลือกข้อสอบ คัดเลือกข้อสอบตามจุดประสงค์การเรียนรู้ ที่กำหนดโดยคัดเลือกข้อสอบตามแผนผังการวัด จำนวนชั้นละ 120 ข้อ

1.4 ผู้วิจัยจัดข้อสอบแต่ละชั้นเป็น 2 ฉบับ ๆ ละ 60 ข้อ เพื่อนำไปทดลองสอบ กับนักเรียนชั้นละประมาณ 200 คน โดยให้นักเรียนตอบแบบทดสอบคนละ 1 ฉบับ

เนื่องจากข้อมูลในการวิจัยครั้งนี้ เก็บจากการสอบในสถานการณ์จริงของการสอบปลาย ภาคการศึกษาที่ 1 ปีการศึกษา 2537 ของนักเรียนชั้นมัธยมศึกษาในจังหวัดสิงห์บุรี ซึ่งนักเรียน ทุกคนในประชากรการวิจัยต้องสอบแบบทดสอบเทียบคะแนนและแบบทดสอบรวมฉบับที่ใช้ในการวิจัย ดังนั้นเพื่อรักษาความลับของแบบทดสอบ จึงนำแบบทดสอบที่สร้างขึ้นไปทดลองสอบกับนักเรียนชั้น มัธยมศึกษาตอนต้น สังกัดกรมสามัญศึกษาในจังหวัดอื่น คือ นักเรียนโรงเรียนมัธยมด่านสำโรง จังหวัด สมุทรปราการ จำนวน 651 คน จำนวนนักเรียนที่ทดลองสอบแบบทดสอบแต่ละฉบับแสดง ดังตาราง 5

1.5 นำผลการสอบจากการทดลองสอบมาวิเคราะห์คุณภาพของข้อสอบเป็นรายข้อ ตามทฤษฎีการสอบแบบดั้งเดิม คือ ค่าความยาก (p) และค่าอำนาจจำแนก (r) ของข้อสอบ

1.6 คัดเลือกข้อสอบซึ่งวัดจุดประสงค์ต่าง ๆ ตามโครงสร้างของแบบทดสอบที่ กำหนด จำนวนชั้นละ 90 ข้อ เพื่อใช้เป็นแบบทดสอบเทียบคะแนน ผลการทดลองสอบได้ข้อสอบ ชั้น ม.1 มีค่าความยากระหว่าง .202 ถึง .835 และค่าอำนาจจำแนกระหว่าง .065 ถึง .492 ข้อสอบชั้น ม.2 มีค่าความยากระหว่าง .143 ถึง .911 และค่าอำนาจจำแนกระหว่าง .021 ถึง .534 ข้อสอบชั้น ม.3 มีค่าความยากระหว่าง .128 ถึง .836 และค่าอำนาจจำแนก

ระหว่าง .029 ถึง .471

2. แบบทดสอบรวม เป็นแบบทดสอบวัดความสามารถทั่วไปทางภาษาไทยสำหรับชั้น ม.1 ม.2 และ ม.3 ผู้วิจัยจัดทำแบบทดสอบรวมขึ้นโดยนำข้อสอบของสำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ ซึ่งสร้างขึ้นและวิเคราะห์คุณภาพแล้วมาจัดฉบับโดยใช้โครงสร้างของแบบทดสอบวัดความสามารถทางภาษาไทย ตามแนวของสำนักงานทดสอบทางการศึกษาดังตาราง 4

ตาราง 4 โครงสร้างของแบบทดสอบความสามารถทั่วไปทางภาษาไทย

ความสามารถ	จำนวนข้อ
1. ความสามารถในการรับสารอย่างมีประสิทธิภาพ	15
1.1 จับใจความสำคัญหรือใจความสำคัญรองได้	6
1.2 แยกข้อเท็จจริงจากข้อคิดเห็นได้	3
1.3 บอกเจตนารมณ์ของผู้ส่งสารได้	3
1.4 แสดงความคิดเห็นในเชิงวิจารณ์ได้อย่างมีเหตุผล	3
2. ความสามารถในการส่งสารได้ตามวัตถุประสงค์อย่างมีประสิทธิภาพ	15
2.1 ใช้ถ้อยคำสำนวนได้เหมาะสมและตรงตามจุดประสงค์อย่างมีประสิทธิภาพ	7
2.2 ลำดับข้อความ เรียงราวได้อย่างเหมาะสม	4
2.3 ใช้ภาษาเชิงสร้างสรรค์ได้	4

ขั้นตอนการสร้างแบบทดสอบความสามารถทั่วไปทางภาษาไทย มีดังนี้

2.1 คัดเลือกข้อสอบที่วัดจุดประสงค์ตามโครงสร้างของแบบทดสอบความสามารถทั่วไปทางภาษาไทย และมีค่าความยากง่ายต่าง ๆ กัน ทั้งข้อสอบที่ง่าย ข้อสอบที่ยากและข้อสอบที่ยากง่ายปานกลาง จำนวน 90 ข้อ แล้วจัดเป็นแบบทดสอบ 2 ฉบับ ๆ ละ 45 ข้อ

2.2 นำแบบทดสอบทั้ง 2 ฉบับไปทดลองสอบกับกลุ่มตัวอย่างนักเรียนชั้น ม.1 ม.2 และ ม.3 จำนวน 1723 คน ใน 4 โรงเรียนของจังหวัดลพบุรี ได้แก่ โรงเรียนโคกกระเทียมวิทยาลัย โรงเรียนดงตาลวิทยา โรงเรียนโคกสำโรงวิทยา และโรงเรียนท่าวังวิทยาคาร จำนวนนักเรียนที่ใช้ในการทดลองสอบแต่ละฉบับ แสดงดังตาราง 5

ตาราง 5 จำนวนนักเรียนที่ใช้ในการทดลองสอบแบบทดสอบเทียบคะแนนและแบบทดสอบร่วม

โรงเรียน	ม.1		ม.2		ม.3	
	จ.1	จ.2	จ.1	จ.2	จ.1	จ.2
<u>แบบทดสอบเทียบคะแนน</u> มัธยมด่านสำโรง	108	104	109	113	108	109
<u>แบบทดสอบร่วม</u> โคกกระเทียมวิทยาลัย ดงตาลวิทยา โคกสำโรงวิทยา ท่าวังวิทยาคาร รวม		195		208		205
	138		138		100	
	96	92	71	75	63	69
	89		94		90	
	323	287	295	283	261	274

2.3 นำผลจากการทดลองสอบมาวิเคราะห์คุณภาพของข้อสอบตามทฤษฎีการสอบแบบดั้งเดิม และวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ คือ ค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าสัมประสิทธิ์การเดา (c)

2.4 คัดเลือกข้อสอบที่มีค่าสัมประสิทธิ์การเดาต่ำกว่า .30 และวัดความสามารถตามโครงสร้างของแบบทดสอบความสามารถทั่วไปทางภาษาไทยจำนวน 30 ข้อ เพื่อใช้เป็นแบบทดสอบร่วมโดยเป็นแบบทดสอบที่สามารถจำแนกความสามารถของนักเรียน 3 ชั้นได้ ข้อสอบร่วมที่ใช้มีค่าความยากระหว่าง -1.1651 ถึง 2.819 ค่าอำนาจจำแนกระหว่าง $.296$ ถึง 1.411 และค่าสัมประสิทธิ์การเดาระหว่าง $.112$ ถึง $.271$ และเมื่อวิเคราะห์ข้อสอบร่วมตามทฤษฎีการสอบแบบดั้งเดิม ปรากฏว่าข้อสอบมีความยากระหว่าง $.212$ ถึง $.816$ และมีค่าอำนาจจำแนกระหว่าง $.277$ ถึง $.590$

การเก็บรวบรวมข้อมูล

ข้อมูลในการวิจัยนี้เก็บรวบรวมจากสถานการณ์การสอบจริงของการสอบปลายภาคเรียนที่ 1 ปีการศึกษา 2537 ซึ่งนักเรียนในประชากรการวิจัยต้องสอบแบบทดสอบภาษาไทยทั้ง 2

ฉบับเหมือนกันทุกคน คือแบบทดสอบผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย ชั้นละ 1 ฉบับ และแบบทดสอบร่วมซึ่งเป็นแบบทดสอบความสามารถทั่วไปทางภาษาไทยที่ใช้กับนักเรียนทุกชั้นเหมือนกันอีก 1 ฉบับ กล่าวคือ นักเรียนชั้น ม.1 สอบแบบทดสอบชั้น ม.1 และแบบทดสอบร่วมซึ่งในการวิจัยนี้จัดรวมไว้ในฉบับเดียวกัน ในทำนองเดียวกัน นักเรียนชั้น ม.2 สอบแบบทดสอบชั้น ม.2 และแบบทดสอบร่วมซึ่งจัดรวมไว้ในฉบับเดียวกัน และนักเรียนชั้น ม.3 สอบแบบทดสอบชั้น ม.3 และแบบทดสอบร่วมซึ่งจัดรวมไว้ในฉบับเดียวกัน เมื่อได้กระดาษคำตอบของนักเรียนทุกคนมาแล้วผู้วิจัยคัดเลือกเฉพาะผลการตอบแบบทดสอบของนักเรียนที่เป็นกลุ่มตัวอย่างเทียบคะแนน และกลุ่มตัวอย่างสอบทานผลเพื่อนำไปวิเคราะห์ต่อไป

รายละเอียดเกี่ยวกับจำนวนข้อสอบและจำนวนกลุ่มตัวอย่างผู้สอบแบบทดสอบแต่ละฉบับแสดงดังตาราง 6

ตาราง 6 จำนวนข้อสอบและจำนวนกลุ่มตัวอย่างผู้สอบแบบทดสอบแต่ละฉบับ

ชั้น	แบบทดสอบ		จำนวนผู้สอบ		
	เทียบคะแนน	รวม	กลุ่มเทียบ	กลุ่มสอบ	รวม
	(จำนวนข้อ)	(จำนวนข้อ)	คะแนน	ทานผล	
ม.1	T ₁	V	1245	373	1618
	(90)	(30)			
ม.2	T ₂	V	1280	384	1664
	(90)	(30)			
ม.3	T ₃	V	1171	352	1523
	(90)	(30)			
รวม			3696	1109	4805

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลดำเนินการโดยแบ่งเป็น 4 ตอน ได้แก่

1. การวิเคราะห์แบบทดสอบ
2. การสร้างตารางเทียบคะแนนแบบทดสอบ
3. การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้นมัธยมศึกษาตอนต้น
4. การศึกษาประสิทธิภาพของวิธีการเทียบคะแนน

รายละเอียดของการวิเคราะห์ข้อมูล มีดังนี้

ตอนที่ 1 การวิเคราะห์แบบทดสอบ

1.1 ตรวจสอบคุณสมบัติการวัดเพียงมิติเดียวของแบบทดสอบแต่ละฉบับด้วยการวิเคราะห์องค์ประกอบ โดยสกัดองค์ประกอบด้วยวิธีวิเคราะห์องค์ประกอบหลัก (Principle Component; PC) แล้วหมุนแกนด้วยวิธีแวนิแมกซ์ (Varimax) โดยใช้โปรแกรมสำเร็จรูป SPSS/PC⁺ และตรวจสอบความสัมพันธ์ระหว่างแบบทดสอบเทียบคะแนนแต่ละฉบับกับแบบทดสอบร่วม โดยการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน โดยใช้โปรแกรมสำเร็จรูป SPSS/PC⁺

1.2 วิเคราะห์ข้อสอบในแบบทดสอบเทียบคะแนนและแบบทดสอบร่วมเพื่อประมาณค่าพารามิเตอร์ของข้อสอบได้แก่ ค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าสัมประสิทธิ์การเดา (c) โดยใช้โปรแกรมสำเร็จรูป BILOG เวอร์ชัน 3.04

1.3 คำนวณค่าความน่าจะเป็นที่ผู้สอบแต่ละชั้นจะตอบข้อสอบแต่ละข้อในแบบทดสอบเทียบคะแนนตามชั้นที่สอบได้ถูกต้อง โดยใช้สูตร (Hambleton and Swaminathan, 1985 : 37-38)

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

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นในการตอบข้อสอบที่ i ถูกเมื่อผู้สอบมีความสามารถ θ

- a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i
- b_i คือ ค่าความยากของข้อสอบข้อที่ i
- c_i คือ ค่าการเดาของข้อสอบข้อที่ i
- D คือ ค่าปรับสเกล มีค่าประมาณ 1.702
- e คือ ค่าคงที่ มีค่าประมาณ 2.718

1.4 คำนวณค่าฟังก์ชันสารสนเทศของแบบทดสอบแต่ละฉบับ โดยนำค่าพารามิเตอร์ของข้อสอบจากแบบทดสอบแต่ละฉบับมาคำนวณ ณ ระดับความสามารถ θ จาก -4 ถึง $+4$ โดยในการวิจัยนี้กำหนด θ มีค่าแตกต่างกันช่วงละ $.05$ ซึ่งจะได้ค่า θ รวม 161 ค่า ดังนี้ $-4.00, -3.95, \dots, +4.00$ การคำนวณค่าฟังก์ชันสารสนเทศของแบบทดสอบ คำนวณจากสูตร (Hambleton and Swaminathan. 1985:104)

$$I(\theta) = \sum_{i=1}^n \frac{P'_i(\theta)^2}{P_i(\theta)Q_i(\theta)}$$

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

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

1.5 คำนวณค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบ ณ ระดับความสามารถ (θ) ต่าง ๆ จาก $+4$ ถึง -4 โดยกำหนดค่า θ เหมือนข้อ 1.4 การคำนวณค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบ ใช้สูตร (Lord. 1982:199)

$$\xi = \sum_{i=1}^n P_i(\theta)$$

1.6 วิเคราะห์ค่าสถิติพื้นฐานของแบบทดสอบฉบับต่าง ๆ ก่อนเทียบคะแนน ได้แก่ คะแนนเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน คะแนนสูงสุด คะแนนต่ำสุด มัธยฐาน ฐานนิยม ความเบ้ ความโด่ง และความเชื่อมั่นของแบบทดสอบ โดยใช้โปรแกรมสำเร็จรูป SPSS/PC⁺

ตอนที่ 2 การสร้างตารางเทียบคะแนนแบบทดสอบ

2.1 ประมวลค่าพารามิเตอร์ของข้อสอบในแบบทดสอบรวม และประมวลค่าพารามิเตอร์ความสามารถของผู้สอบแต่ละคน โดยใช้โปรแกรมสำเร็จรูป BILOG เวอร์ชัน 3.04 (Mislevy and Bock. 1990)

2.2 ประมวลค่าพารามิเตอร์ของข้อสอบในแบบทดสอบ ชั้น ม.1 ม.2 และ ม.3 แต่ละฉบับ และประมวลค่าพารามิเตอร์ความสามารถของผู้สอบแต่ละคน โดยใช้โปรแกรมสำเร็จรูป BILOG เวอร์ชัน 3.04

2.3 แปลงค่าพารามิเตอร์ความยาก (b) อำนาจจำแนก (a) และพารามิเตอร์ความสามารถ (θ) ที่ได้จากแบบทดสอบเทียบคะแนนแต่ละระดับชั้น ให้อยู่ในสเกลเดียวกับสเกลของพารามิเตอร์จากแบบทดสอบรวม โดยใช้สูตร (Warm, 1978:116)

$$\theta'_T = \left[\frac{SD_{\theta_v}}{SD_{\theta_T}} \right] \theta_T + \left[\bar{\theta}_v - \frac{SD_{\theta_v}}{SD_{\theta_T}} \bar{\theta}_T \right]$$

$$b'_T = \left[\frac{SD_{\theta_v}}{SD_{\theta_T}} \right] b_T + \left[\bar{\theta}_v - \frac{SD_{\theta_v}}{SD_{\theta_T}} \bar{\theta}_T \right]$$

$$a_T = \left[\frac{SD_{\theta_v}}{SD_{\theta_T}} \right] a_T$$

- โดยที่ θ_T คือ พารามิเตอร์ความสามารถของผู้สอบที่ได้แปลงให้อยู่ในสเกลของแบบทดสอบรวม
- θ_T คือ พารามิเตอร์ความสามารถของผู้สอบที่ได้จากแบบทดสอบเทียบคะแนน
- SD_{θ_v} คือ ส่วนเบี่ยงเบนมาตรฐานของพารามิเตอร์ความสามารถในสเกลของแบบทดสอบรวม
- SD_{θ_T} คือ ส่วนเบี่ยงเบนมาตรฐานของพารามิเตอร์ความสามารถในสเกลของแบบทดสอบเทียบคะแนน
- $\bar{\theta}_v$ คือ ค่าเฉลี่ยของพารามิเตอร์ความสามารถในสเกลของแบบทดสอบรวม
- $\bar{\theta}_T$ คือ ค่าเฉลี่ยของพารามิเตอร์ความสามารถในสเกลของแบบทดสอบเทียบคะแนน

- b_T คือ พารามิเตอร์ความยากของข้อสอบในสเกลของ
แบบทดสอบรวม
- b_T คือ พารามิเตอร์ความยากของข้อสอบในสเกลของแบบทดสอบ
เทียบคะแนน
- a_T คือ พารามิเตอร์อำนาจจำแนกในสเกลของแบบทดสอบรวม
- a_T คือ พารามิเตอร์อำนาจจำแนกในสเกลของแบบทดสอบรวม
ตามระดับชั้น

2.4 ประมาณค่าคะแนนจริงของผู้สอบ แต่ละคนที่ได้จากการตอบแบบทดสอบแต่ละ
ฉบับ โดยใช้สูตร (Lord. 1982 : 199)

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

$$\xi_y = \sum_{i=1}^m P_i(\theta_y)$$

$$\xi_z = \sum_{i=1}^k P_i(\theta_z)$$

$$\xi_v = \sum_{i=1}^l P_i(\theta_v)$$

- โดยที่ ξ_x คือ คะแนนจริงของนักเรียนชั้น ม.1 ซึ่งได้จากการตอบ
แบบทดสอบชั้น ม.1 (X)
- ξ_y คือ คะแนนจริงของนักเรียนชั้น ม.2 ซึ่งได้จากการตอบ
แบบทดสอบชั้น ม.2 (Y)
- ξ_z คือ คะแนนจริงของนักเรียนชั้น ม.3 ซึ่งได้จากการตอบ
แบบทดสอบชั้น ม.3 (Z)
- ξ_v คือ คะแนนจริงของนักเรียนชั้น ม.1 ม.2 และ ม.3
ซึ่งได้จากการตอบแบบทดสอบรวม (V)

2.5 คำนวณหาค่าเฉลี่ยของคะแนนสังเกตโดยคิดจากผู้สอบที่ได้คะแนนจริงเท่ากัน

2.6 นำค่าเฉลี่ยคะแนนสังเกตของคะแนนจริงจากแบบทดสอบแต่ละฉบับ ตามข้อ

2.5 ไปคำนวณหาคะแนน ณ ตำแหน่งเปอร์เซ็นต์ไทล์ต่าง ๆ

2.7 สร้างตารางเทียบคะแนนสังเกตตามค่าเปอร์เซ็นต์ไทล์ที่ได้ในข้อ 2.6 โดยเรียงจากเปอร์เซ็นต์ไทล์น้อยไปมาก เพื่อเทียบคะแนนแบบทดสอบ ม.1 ม.2 และ ม.3 และแบบทดสอบรวม

ตอนที่ 3 การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้นมัธยมศึกษาตอนต้น

การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 จะศึกษาจากผลต่างของค่าเฉลี่ยผลสัมฤทธิ์ทางการเรียนของนักเรียนแต่ละชั้น ภายหลังจากเทียบคะแนนของนักเรียนชั้น ม.2 และ ม.3 ให้อยู่ในสเกลเดียวกับคะแนนของแบบทดสอบชั้น ม.1 แล้ว การวิเคราะห์ข้อมูลดำเนินการ 2 ครั้ง ครั้งแรกศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนของนักเรียน 3 ชั้นในกลุ่มรวม โดยใช้ข้อมูลจากนักเรียนในกลุ่มตัวอย่างเทียบคะแนนทุกคนและครั้งที่ 2 ศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนของนักเรียนจากกลุ่มตัวอย่างเทียบคะแนน โดยแยกวิเคราะห์ข้อมูลในกลุ่มนักเรียนที่มีความสามารถสูง นักเรียนที่มีความสามารถปานกลางและนักเรียนที่มีความสามารถต่ำแต่ละกลุ่ม ขั้นตอนการวิเคราะห์ข้อมูลในแต่ละครั้งดำเนินการดังนี้

3.1 การศึกษาความก้าวหน้าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ม.2 และม.3 ในกลุ่มรวม

3.1.1 แปลงคะแนนที่นักเรียนชั้น ม. 2 และ ม.3 แต่ละคนได้จากการตอบแบบทดสอบตามชั้นของตน ไปสู่สเกลคะแนนแบบทดสอบชั้น ม.1 โดยใช้ตารางเทียบคะแนนที่สร้างไว้แล้วในตอนที 2

3.1.2 คำนวณหาค่าสถิติพื้นฐานของคะแนนภายหลังการเทียบคะแนนแล้วที่นักเรียนแต่ละชั้นทำได้ โดยใช้โปรแกรมสำเร็จรูป SPSS/PC⁺

3.1.3 ทดสอบนัยสำคัญของความแตกต่างของคะแนนเฉลี่ยตามข้อ 3.1.2 ด้วยการวิเคราะห์ความแปรปรวนทางเดียว โดยใช้โปรแกรมสำเร็จรูป SPSS/PC⁺

3.1.4 คำนวณหาค่าผลต่างของคะแนนเฉลี่ยของนักเรียนแต่ละชั้นในข้อ

3.1.2 ค่าที่ได้นี้เป็นค่าที่แสดงความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนของนักเรียนในแต่ละชั้น

3.1.5 วิเคราะห์แนวโน้มการเปลี่ยนแปลงคะแนนเฉลี่ยของนักเรียนชั้น

ม.1 ม.2 และ ม.3

3.2 การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนของนักเรียนในแต่ละกลุ่มความสามารถ

การแบ่งกลุ่มผู้สอบตามความสามารถ (๑) จะแบ่งกลุ่มโดยใช้คะแนนของนักเรียนที่ได้จากการทำแบบทดสอบเทียบคะแนนตามชั้นของตน ตามเกณฑ์ดังนี้

ผู้สอบที่มีความสามารถต่ำ ($\theta \leq \bar{\theta} - .84 \text{ SD.}$)

ผู้สอบที่มีความสามารถปานกลางค่อนข้างต่ำ ($\bar{\theta} - .84 \text{ SD.} < \theta \leq \bar{\theta} - .25 \text{ SD.}$)

ผู้สอบที่มีความสามารถปานกลาง ($\bar{\theta} - .25 \text{ SD.} < \theta \leq \bar{\theta} + .25 \text{ SD.}$)

ผู้สอบที่มีความสามารถปานกลางค่อนข้างสูง ($\bar{\theta} + .25 \text{ SD.} < \theta \leq \bar{\theta} + .84 \text{ SD.}$)

ผู้สอบที่มีความสามารถสูง ($\theta > \bar{\theta} + .84 \text{ SD.}$)

ในการศึกษาความก้าวหน้าและเปรียบเทียบผลสัมฤทธิ์ทางการเรียนในกลุ่มผู้สอบที่มีระดับความสามารถต่างกันนั้นเพื่อให้เห็นผลการเปรียบเทียบอย่างชัดเจน จึงวิเคราะห์โดยใช้ข้อมูลจากผู้สอบที่มีความสามารถต่ำ ผู้สอบที่มีความสามารถปานกลาง และผู้สอบที่มีความสามารถสูงเท่านั้น ส่วนผู้สอบที่จัดอยู่ในกลุ่มที่มีความสามารถปานกลางค่อนข้างต่ำ และกลุ่มที่มีความสามารถปานกลางค่อนข้างสูงจะไม่นำมาวิเคราะห์

การวิเคราะห์ข้อมูลในกลุ่มผู้สอบแต่ละระดับความสามารถดำเนินการตามขั้นตอนในทำนองเดียวกับข้อ 3.1.1 ถึง 3.1.4

ตอนที่ 4 การศึกษาประสิทธิภาพของวิธีการเทียบคะแนน

การศึกษาประสิทธิภาพของวิธีการเทียบคะแนนนี้จะศึกษาโดยใช้ข้อมูลจากกลุ่มสอบทานผล ประสิทธิภาพของวิธีการเทียบคะแนนจะประเมินจากความเพียงพอของวิธีการเทียบคะแนน โดยคำนวณค่าดัชนีความแตกต่างซึ่งเป็นค่าความคลาดเคลื่อนรวมของการเทียบคะแนนแล้วเปรียบเทียบดัชนีดังกล่าวกับระดับการยอมรับความเพียงพอของการเทียบคะแนนตามเกณฑ์ของปีเตอร์สัน และคนอื่น ๆ (Petersen and others. 1982 : 91-93)

ขั้นตอนการศึกษาประสิทธิภาพของการเทียบคะแนน ดำเนินการโดยใช้ข้อมูลจากกลุ่มสอบทานผล ดังนี้

4.1 แปลงคะแนนที่นักเรียนแต่ละคนทำได้จากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนตามชั้นของตนไปสู่สเกลคะแนนของแบบทดสอบรวมโดยใช้ตารางการเทียบคะแนนที่สร้างในตอนที 2

4.2 คำนวณความแตกต่างระหว่างคะแนนแปลงที่นักเรียนแต่ละคนทำได้ตามข้อ

4.1 กับ คะแนนจากแบบทดสอบรวมที่นักเรียนแต่ละคน ทำได้จริง

4.3 นำคะแนนความแตกต่างที่ได้ตามข้อ 4.2 มาคำนวณดัชนีความแตกต่าง (C) จากสูตร (Petersen and others. 1982 : 91)

$$C = \sum_{i=1}^n (x_i - x_i')^2 / n S_x^2$$

เมื่อ C คือ ดัชนีความแตกต่าง

x_i คือ คะแนนของนักเรียนคนที่ 1 ที่ได้จากการทำแบบทดสอบรวม

x_i' คือ คะแนนของนักเรียนคนที่ 1 ที่ได้จากการแปลงคะแนน

แบบทดสอบวัดผลสัมฤทธิ์ตามระดับชั้นโดยใช้ตารางเทียบคะแนน

n คือ จำนวนนักเรียนในกลุ่มสอบทานผล

S_x^2 คือ ความแปรปรวนของคะแนนจากแบบทดสอบรวม

4.4 ประเมินค่าดัชนีความแตกต่างในข้อ 4.3 โดยนำไปเปรียบเทียบกับเกณฑ์ของปีเตอร์สันและคนอื่น ๆ ดังตาราง 7 แล้วสรุปผลการประเมินตามเกณฑ์ที่ตั้งไว้ว่าวิธีการเทียบคะแนนที่ใช้ในการวิจัยครั้งนี้มีความเพียงพอในระดับใด

ตาราง 7 เกณฑ์การประเมินดัชนีความแตกต่าง

ดัชนีความแตกต่าง	ความหมาย
$C \leq (.05 SD_x)^2$	น่าพอใจอย่างมาก
$(.05 SD_x)^2 < C \leq (.10 SD_x)^2$	น่าพอใจ
$(.10 SD_x)^2 < C \leq (.15 SD_x)^2$	ปานกลาง
$(.15 SD_x)^2 < C \leq (.20 SD_x)^2$	ไม่น่าพอใจ
$(.20 SD_x)^2 < C$	ไม่น่าพอใจอย่างยิ่ง

บทที่ 4
ผลการวิเคราะห์ข้อมูล

การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย ของนักเรียนชั้นมัธยมศึกษาตอนต้น โดยใช้วิธีการเทียบคะแนนตามแนวดิ่ง ตามทฤษฎีการตอบสนองข้อสอบ แบบจำลองโลจิสติก สามพารามิเตอร์นี้ ผู้วิจัยแบ่งการเสนอผลการวิเคราะห์ข้อมูลเป็น 4 ตอน ดังนี้

ตอนที่ 1 ผลการวิเคราะห์แบบทดสอบ

ตอนที่ 2 ผลการสร้างตารางเทียบคะแนนแบบทดสอบ

ตอนที่ 3 ผลการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย
ของนักเรียนชั้นมัธยมศึกษาตอนต้น

ตอนที่ 4 ผลการศึกษาประสิทธิภาพของวิธีการเทียบคะแนน

ผลการวิเคราะห์ข้อมูลแต่ละตอน มีดังนี้

ตอนที่ 1 ผลการวิเคราะห์แบบทดสอบ

1.1 ผลการตรวจสอบคุณสมบัติการวัดเพียงมิติเดียวของแบบทดสอบ

จากข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบที่กำหนดว่าแบบทดสอบต้องวัดเพียงมิติเดียวนั้น ผู้วิจัยนำคะแนนที่ได้จากการตอบ แบบทดสอบวิชาภาษาไทย ชั้น ม.1 ม.2 และม.3 และแบบทดสอบความสามารถทั่วไปทางภาษาไทย มาตรวจสอบคุณสมบัติการวัดเพียงมิติเดียวด้วยการวิเคราะห์องค์ประกอบ โดยสกัดองค์ประกอบด้วยวิธีวิเคราะห์องค์ประกอบหลัก แล้วหมุนแกนด้วยวิธีแวนแมกซ์ ผลการวิเคราะห์แสดงดังตาราง 8 - 11

ตาราง 8 ค่าไอเกนขององค์ประกอบที่มีค่าเกิน 1.0 ของแบบทดสอบชั้น ม.1

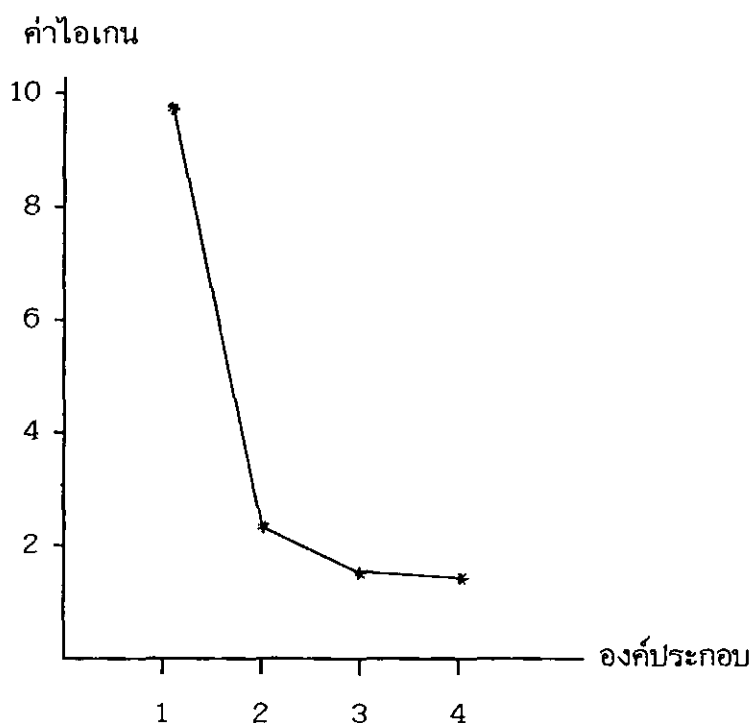
องค์ประกอบ	ค่าไอเกน
1	9.945
2	2.294
3	1.616
4	1.466
5	1.439
6	1.422

ตาราง 8 (ต่อ)

องค์ประกอบ	ค่าไอเกน
7	1.345
8	1.331
9	1.309
10	1.307
11	1.289
12	1.263
13	1.238
14	1.204
15	1.192
16	1.161
17	1.144
18	1.141
19	1.121
20	1.092
21	1.082
22	1.077
23	1.072
24	1.062
25	1.051
26	1.027
27	1.018
28	1.011

จากตาราง 8 แสดงว่าเมื่อวิเคราะห์องค์ประกอบของแบบทดสอบชั้น ม.1 ได้ องค์ประกอบที่มีค่าไอเกนเกิน 1.0 จำนวน 28 องค์ประกอบ ค่าไอเกนที่มีค่าสูงสุด คือ 9.945 ค่าไอเกนที่มีค่ารองลงมาเป็นอันดับ 2 3 และ 4 มีค่า 2.294 1.616 และ 1.466 ตามลำดับ

เมื่อนำค่าไอเกนที่มีค่าสูงสุด 4 ค่าแรก มาเขียนกราฟแสดงความสัมพันธ์ระหว่าง ตำแหน่งขององค์ประกอบกับค่าไอเกน ได้ภาพดังภาพประกอบ 4



ภาพประกอบ 4 ค่าไอเกนขององค์ประกอบ 4 ค่าแรกของแบบทดสอบชั้น ม.1

เมื่อพิจารณาค่าไอเกนขององค์ประกอบของแบบทดสอบชั้น ม.1 จากตารางที่ 8 ประกอบกับ ภาพประกอบ 4 จะเห็นว่าค่าไอเกนขององค์ประกอบที่ 1 สูงกว่าค่าไอเกนของ องค์ประกอบที่ 2 มากกว่า 4 เท่า แต่ค่าไอเกนขององค์ประกอบที่ 2 สูงกว่าค่าไอเกนของ องค์ประกอบที่ 3 และองค์ประกอบที่ 4 เพียงเล็กน้อย แปลความหมายได้ว่า แบบทดสอบชั้น ม.1 จำนวน 90 ข้อนั้น สามารถวัดคุณลักษณะที่เด่น (Dominant) เพียง 1 คุณลักษณะหรืออาจกล่าว ได้ว่า แบบทดสอบชั้น ม.1 มีคุณสมบัติการวัดเพียงมิติเดียว (Hambleton and Swaminathan. 1985 : 17; Lord. 1980 : 21) ซึ่งเป็นไปตามข้อตกลงเบื้องต้นของแบบจำลองโลจิสติก

ตาราง 9 ค่าไอเกนขององค์ประกอบที่มีค่าเกิน 1.0 ของแบบทดสอบชั้น ม.2

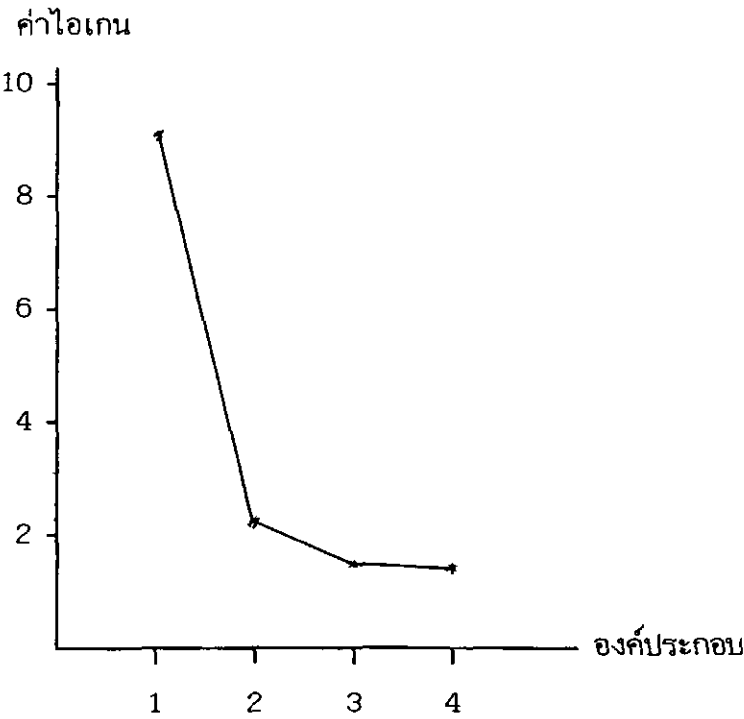
องค์ประกอบ	ค่าไอเกน
1	8.917
2	2.240
3	1.538
4	1.464
5	1.424
6	1.393
7	1.359
8	1.347
9	1.318
10	1.290
11	1.277
12	1.260
13	1.230
14	1.203
15	1.193
16	1.190
17	1.169
18	1.159
19	1.145
20	1.141
21	1.121
22	1.107
23	1.080
24	1.069
25	1.058
26	1.045
27	1.035

ตาราง 9 (ต่อ)

องค์ประกอบ	ค่าไอเกน
28	1.028
29	1.003

จากตาราง 9 แสดงว่าเมื่อวิเคราะห์องค์ประกอบแบบทดสอบชั้น ม.2 ได้องค์ประกอบที่มีค่าไอเกนเกิน 1.0 จำนวน 29 องค์ประกอบ โดยค่าไอเกนที่มีค่าสูงสุด คือ 8.917 ค่าไอเกนที่มีค่ารองลงมาเป็นอันดับ 2 3 และ 4 มีค่า 2.240 1.538 และ 1.464 ตามลำดับ

เมื่อนำค่าไอเกนที่มีค่าสูงสุด 4 ค่าแรก มาเขียนกราฟแสดงความสัมพันธ์ระหว่างตำแหน่งขององค์ประกอบกับค่าไอเกน ได้ภาพดังภาพประกอบ 5



ภาพประกอบ 5 ค่าไอเกนขององค์ประกอบที่มีค่าเกิน 1.0 ของแบบทดสอบชั้น ม.2

เมื่อพิจารณาค่าไอเอนขององค์ประกอบของแบบทดสอบชั้น ม.2 จากตารางที่ 9 ประกอบกับภาพประกอบ 5 จะเห็นว่าค่าไอเอนขององค์ประกอบที่ 1 สูงกว่าค่าไอเอนขององค์ประกอบที่ 2 ประมาณ 4 เท่า แต่ค่าไอเอนขององค์ประกอบที่ 2 สูงกว่าค่าไอเอนขององค์ประกอบที่ 3 และองค์ประกอบที่ 4 เพียงเล็กน้อย แปลความหมายได้ว่า แบบทดสอบชั้น ม.2 จำนวน 90 ข้อนี้ สามารถวัดคุณลักษณะที่เด่นเพียง 1 คุณลักษณะ จึงอาจกล่าวได้ว่าแบบทดสอบชั้น ม.2 มีคุณสมบัติการวัดเพียงมิติเดียว ซึ่งเป็นไปตามข้อตกลงเบื้องต้นของแบบจำลองโลจิสติก

ตาราง 10 ค่าไอเอนขององค์ประกอบที่มีค่าเกิน 1.0 ของแบบทดสอบชั้น ม.3

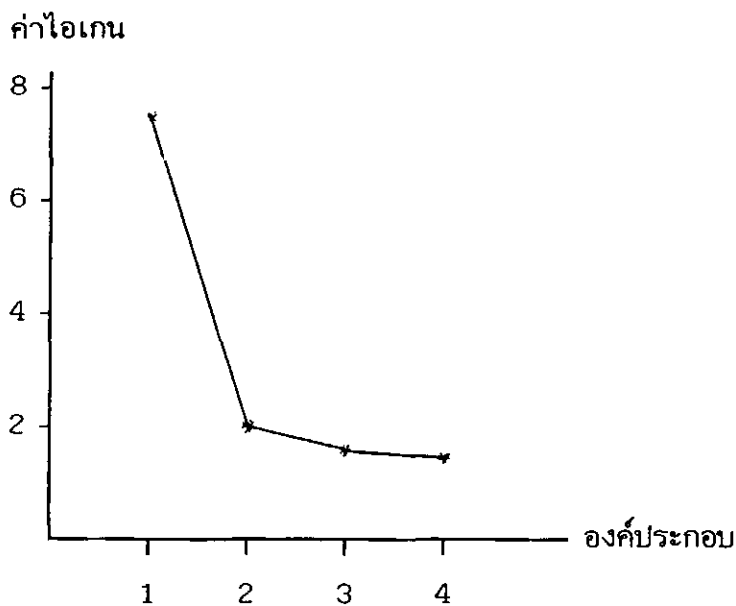
องค์ประกอบ	ค่าไอเอน
1	7.528
2	2.094
3	1.558
4	1.470
5	1.456
6	1.435
7	1.434
8	1.390
9	1.350
10	1.329
11	1.315
12	1.301
13	1.293
14	1.260
15	1.239
16	1.218
17	1.212
18	1.193
19	1.615

ตาราง 10 (ต่อ)

องค์ประกอบ	ค่าไอเกน
20	1.147
21	1.127
22	1.121
23	1.100
24	1.097
25	1.084
26	1.076
27	1.066
28	1.056
29	1.047
30	1.018
31	1.010
32	1.005

จากตาราง 10 แสดงว่า เมื่อวิเคราะห์องค์ประกอบของแบบทดสอบชั้น ม.3 ได้ องค์ประกอบที่มีค่าไอเกน เกิน 1.0 จำนวน 32 องค์ประกอบ โดยค่าไอเกนที่มีค่าสูงสุด คือ 7.528 ค่าไอเกนที่มีค่ารองลงมาเป็นอันดับ 2 3 และ 4 มีค่า 2.094 1.558 และ 1.470 ตามลำดับ

เมื่อนำค่าไอเกนที่มีค่าสูงสุด 4 ค่าแรก มาเขียนกราฟแสดงความสัมพันธ์ระหว่าง ตำแหน่งขององค์ประกอบกับค่าไอเกน ได้ภาพดังภาพประกอบ 6



ภาพประกอบ 6 ค่าไอเกนขององค์ประกอบ 4 ค่าแรก ของแบบทดสอบชั้น ม.3

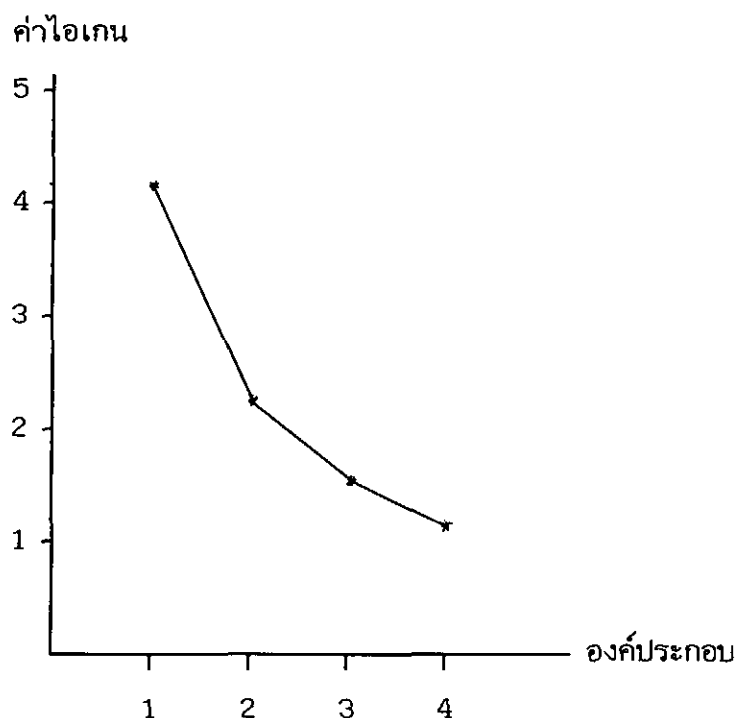
เมื่อพิจารณาค่าไอเกนขององค์ประกอบของแบบทดสอบชั้น ม.3 จากตาราง 9 ประกอบกับภาพประกอบ 6 จะเห็นว่าค่าไอเกนขององค์ประกอบที่ 1 สูงกว่าค่าไอเกนขององค์ประกอบที่ 2 มากกว่า 3 เท่า แต่ค่าไอเกนขององค์ประกอบที่ 2 สูงกว่าค่าไอเกนขององค์ประกอบที่ 3 และองค์ประกอบที่ 4 เพียงเล็กน้อย แปลความหมายได้ว่า แบบทดสอบชั้น ม.3 จำนวน 90 ข้อนี้สามารถวัดคุณลักษณะที่เด่นเพียง 1 คุณลักษณะ ซึ่งเป็นไปตามข้อตกลงเบื้องต้นของแบบจำลองโลจิสติก

ตาราง 11 ค่าไอเกนขององค์ประกอบที่มีค่าเกิน 1.0 ของแบบทดสอบความสามารถทั่วไป
ทางภาษาไทย

องค์ประกอบ	ค่าไอเกน
1	4.147
2	2.216
3	1.468
4	1.135
5	1.030
6	1.004

จากตาราง 11 แสดงว่าเมื่อวิเคราะห์องค์ประกอบแบบทดสอบรวม ได้องค์ประกอบที่มีค่าไอเกนเกิน 1.0 จำนวน 6 องค์ประกอบ โดยค่าไอเกนที่มีค่าสูงสุดมีค่า 4.147 ค่าไอเกนที่มีค่ารองลงมาเป็นอันดับ 2 3 และ 4 มีค่า 2.216 และ 1.468 และ 1.135 ตามลำดับ

เมื่อนำค่าไอเกนที่มีค่าสูงสุด 4 ค่าแรก มาเขียนกราฟแสดงความสัมพันธ์ระหว่างตำแหน่งขององค์ประกอบกับค่าไอเกน ได้ภาพดังภาพประกอบ 7



ภาพประกอบ 7 ค่าไอเกนขององค์ประกอบ 4 ค่าแรกของแบบทดสอบ
ความสามารถทั่วไปทางภาษาไทย

เมื่อพิจารณาค่าไอเกนขององค์ประกอบของแบบทดสอบวัดความสามารถทั่วไปทางภาษาไทย จากตาราง 10 ประกอบกับภาพประกอบ 7 จะเห็นว่า ค่าไอเกนขององค์ประกอบที่ 1 สูงกว่าค่าไอเกนขององค์ประกอบที่ 2 ประมาณ 2 เท่า และค่าไอเกนขององค์ประกอบที่ 2 สูงกว่าค่าไอเกนขององค์ประกอบที่ 3 และ 4 เพียงเล็กน้อย แสดงว่าแบบทดสอบร่วมมีแนวโน้มที่จะวัดคุณลักษณะที่เด่นเพียง 1 คุณลักษณะ แต่ไม่เด่นชัด (Dominant) เท่ากับแบบทดสอบทั้ง 3 ฉบับที่กล่าวมา จึงกล่าวได้ว่าแบบทดสอบความสามารถทั่วไปทางภาษาไทยมีคุณสมบัติการวัดเพียงมิติเดียวซึ่งเป็นไปตามข้อตกลงเบื้องต้นของแบบจำลองโลจิสติก

เมื่อตรวจสอบความสัมพันธ์ของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยแต่ละฉบับ กับแบบทดสอบวัดความสามารถทั่วไปทางภาษาไทยซึ่งเป็นแบบทดสอบที่ใช้ในการปรับค่าพารามิเตอร์ในการเทียบคะแนน โดยคำนวณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน ได้ขนาดความสัมพันธ์ดังแสดงในตาราง 12

ตาราง 12 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน กับแบบทดสอบความสามารถทั่วไปทางภาษาไทย

แบบทดสอบ	ม.1	ม.2	ม.3
แบบทดสอบความสามารถทั่วไป	.548**	.541**	.621**

** มีนัยสำคัญที่ระดับ .001

จากตาราง 12 พบว่าแบบทดสอบวัดความสามารถทั่วไป ซึ่งในการวิจัยนี้ใช้เป็นแบบทดสอบร่วมกับแบบทดสอบเทียบคะแนนแต่ละฉบับมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 ทั้ง 3 ฉบับ โดยขนาดความสัมพันธ์อยู่ในระดับปานกลาง แบบทดสอบชั้น ม.3 มีขนาดความสัมพันธ์กับแบบทดสอบความสามารถทั่วไปทางภาษาไทยสูงที่สุด รองลงไปที่แบบทดสอบชั้น ม.1 และแบบทดสอบชั้น ม.2 ตามลำดับ

1.2 ผลการวิเคราะห์ข้อสอบรายข้อ

ผู้วิจัยนำผลการตอบแบบทดสอบเทียบคะแนนและแบบทดสอบรวมมาวิเคราะห์เพื่อประมาณค่าพารามิเตอร์ของข้อสอบเป็นรายข้อ ปรากฏว่ามีข้อสอบบางข้อที่ไม่สามารถประมาณค่าพารามิเตอร์ความยากได้ คือแบบทดสอบชั้น ม.1 ข้อ 58 66 67 68 และ 84 แบบทดสอบชั้น ม.2 ข้อ 19 28 58 81 และ 83 แบบทดสอบชั้น ม.3 ข้อ 4 และ 20 แบบทดสอบรวม ข้อ 10 ดังนั้นในการวิเคราะห์ข้อมูลจึงต้องตัดข้อสอบข้อที่ไม่สามารถประมาณค่าพารามิเตอร์ความยากดังกล่าวออกไป และเพื่อให้การแปลความหมายผลการเทียบคะแนนชัดเจน เข้าใจได้ง่าย ผู้วิจัยจึงตัดข้อสอบชั้น ม.3 ออกอีก 3 ข้อ เพื่อให้เหลือจำนวนข้อสอบเท่ากับจำนวนข้อสอบในแบบทดสอบ ชั้น ม.1 และชั้น ม.2 โดยตัดข้อ 8 35 และ 84 ซึ่งเป็นข้อสอบที่มีความยากมาก และอ่านจําแนกต่ำกว่าข้อสอบข้ออื่น ๆ ออก ดังนั้นในการวิเคราะห์ข้อมูลขั้นต่อ ๆ ไป ผู้วิจัยจึงวิเคราะห์จากแบบทดสอบชั้น ม.1 ม.2 และ ม.3 ซึ่งประกอบด้วยข้อสอบฉบับละ 85 ข้อและแบบทดสอบรวมซึ่งประกอบด้วยข้อสอบ 29 ข้อ ค่าพารามิเตอร์ของข้อสอบจากแบบทดสอบแต่ละฉบับแสดงดังตาราง 13-16 และไค้สารสนเทศของแบบทดสอบฉบับต่าง ๆ แสดงดังภาพประกอบ

ตาราง 13 ค่าประมาณพารามิเตอร์ของข้อสอบในแบบทดสอบชั้น ม.1

ข้อ	a	S.E. _a	b	S.E. _b	c	S.E. _c
1	0.534	0.103	0.546	0.219	0.308	0.066
2	0.399	0.131	2.853	0.999	0.306	0.046
3	0.725	0.239	2.584	0.979	0.249	0.025
4	0.606	0.123	1.167	0.321	0.263	0.048
5	1.256	0.227	0.421	0.227	0.445**	0.035
6	0.952	0.149	0.220	0.162	0.331	0.047
7	0.514	0.060	-1.337	0.187	0.209	0.083
8	0.926	0.099	-1.118	0.154	0.199	0.071
9	0.772	0.070	-2.099*	0.217	0.170	0.073
10	0.358	0.074	1.168	0.303	0.202	0.068
11	0.456	0.058	-1.170	0.182	0.207	0.083
12	0.352	0.072	0.944	0.264	0.216	0.073
13	0.730	0.159	1.861	0.491	0.196	0.028
14	0.371	0.068	0.457	0.179	0.213	0.075
15	0.193*	0.056	3.168	0.960	0.338	0.080
16	0.259	0.060	0.950	0.302	0.284	0.089
17	0.992	0.086	-1.360	0.148	0.131	0.056
18	0.846	0.075	-1.213	0.138	0.157	0.062
19	0.305	0.050	-0.320	0.144	0.214	0.083
20	0.475	0.060	-0.830	0.150	0.195	0.077
21	0.701	0.073	-1.796	0.213	0.194	0.080
22	0.623	0.070	-1.961	0.245	0.231	0.091
23	0.653	0.070	-1.233	0.165	0.199	0.077
24	0.958	0.120	0.060	0.119	0.219	0.045
25	1.089	0.138	-0.346	0.118	0.277	0.053
26	0.799	0.302	2.479	1.113	0.354	0.025

ตาราง 13 (ต่อ)

ข้อ	a	S.E. _a	b	S.E. _b	c	S.E. _c
27	0.512	0.088	0.370	0.170	0.229	0.067
28	0.612	0.091	0.380	0.150	0.194	0.055
29	0.598	0.081	0.111	0.119	0.167	0.057
30	0.464	0.052	-1.778	0.221	0.198	0.082
31	0.987	0.084	-1.532	0.167	0.170	0.068
32	0.739	0.118	0.793	0.211	0.225	0.041
33	0.361	0.067	0.726	0.211	0.202	0.071
34	0.959	0.242	2.023	0.664	0.206	* 0.021
35	1.138	0.165	0.616	0.199	0.253	0.032
36	0.428	0.166	4.348**	1.769	0.282	0.029
37	0.761	0.197	2.137	0.659	0.215	0.026
38	0.414	0.070	0.308	0.156	0.203	0.072
39	0.764	0.099	0.208	0.125	0.188	0.048
40	0.391	0.096	1.779	0.495	0.223	0.059
41	0.553	0.112	0.979	0.286	0.285	0.055
42	1.170	0.160	0.428	0.162	0.238	0.034
43	0.961	0.103	-0.669	0.118	0.201	0.059
44	0.552	0.113	1.282	0.339	0.255	0.049
45	0.842	0.124	0.337	0.153	0.246	0.046
46	0.341	0.049	-0.394	0.128	0.182	0.074
47	0.638	0.147	1.929	0.516	0.174	* 0.032
48	0.639	0.078	0.124	0.109	0.150*	0.050
49	0.691	0.151	1.705	0.455	0.211	0.033
50	0.579	0.082	0.326	0.137	0.176	0.055
51	0.560	0.098	0.547	0.191	0.236	0.060
52	0.608	0.076	-0.334	0.116	0.183	0.064

ตาราง 13 (ต่อ)

ข้อ	a	S.E. _a	b	S.E. _b	c	S.E. _c
53	0.916	0.171	0.876	0.275	0.303	0.037
54	1.005	0.119	-0.191	0.107	0.207	0.049
55	0.598	0.098	0.784	0.206	0.206	0.049
56	1.435	0.507	2.277	1.262	0.164	0.014
57	0.538	0.090	0.833	0.210	0.182	0.052
58	0.714	0.252	2.882	1.144	0.205	0.022
59	0.908	0.271	2.239	0.835	0.231	0.022
60	1.047	0.143	0.973	0.219	0.153	0.025
61	1.024	0.154	0.633	0.208	0.308	0.036
62	1.005	0.149	0.867	0.230	0.251	0.031
63	1.294	0.253	1.175	0.416	0.354	0.025
64	0.686	0.108	0.662	0.193	0.234	0.047
65	0.813	0.354	3.252	1.621	0.208	0.017
66	1.246	0.182	-0.694	0.151	0.346	0.062
67	0.963	0.133	0.716	0.180	0.160	0.032
68	0.669	0.107	-0.042	0.150	0.291	0.067
69	1.505**	0.240	0.998	0.320	0.215	0.022
70	0.585	0.189	2.661	0.949	0.257	0.031
71	0.919	0.161	0.924	0.272	0.293	0.036
72	0.866	0.091	-0.589	0.110	0.166	0.058
73	1.140	0.136	-0.519	0.122	0.275	0.056
74	0.652	0.103	-0.377	0.165	0.334	0.079
75	0.483	0.067	-0.110	0.118	0.176	0.067
76	1.112	0.188	0.808	0.253	0.258	0.032
77	0.459	0.103	1.300	0.363	0.250	0.059
78	1.366	0.221	0.937	0.295	0.249	0.025

ตาราง 13 (ต่อ)

ข้อ	a	S.E. _a	b	S.E. _b	c	S.E. _c
79	1.062	0.150	0.945	0.232	0.190	0.028
80	0.994	0.170	0.972	0.273	0.253	0.031
81	0.725	0.115	0.414	0.174	0.276	0.052
82	1.112	0.122	0.015	0.111	0.228	0.039
83	0.988	0.113	-0.546	0.115	0.209	0.058
84	0.755	0.116	0.750	0.198	0.201	0.041
85	0.766	0.112	0.241	0.143	0.226	0.052
ค่าเฉลี่ย	.758		.552		.231	
ส่วนเบี่ยงเบนมาตรฐาน.289			1.274		.056	

★★ หมายถึง ค่าพารามิเตอร์ของข้อสอบที่มีค่าสูงสุด

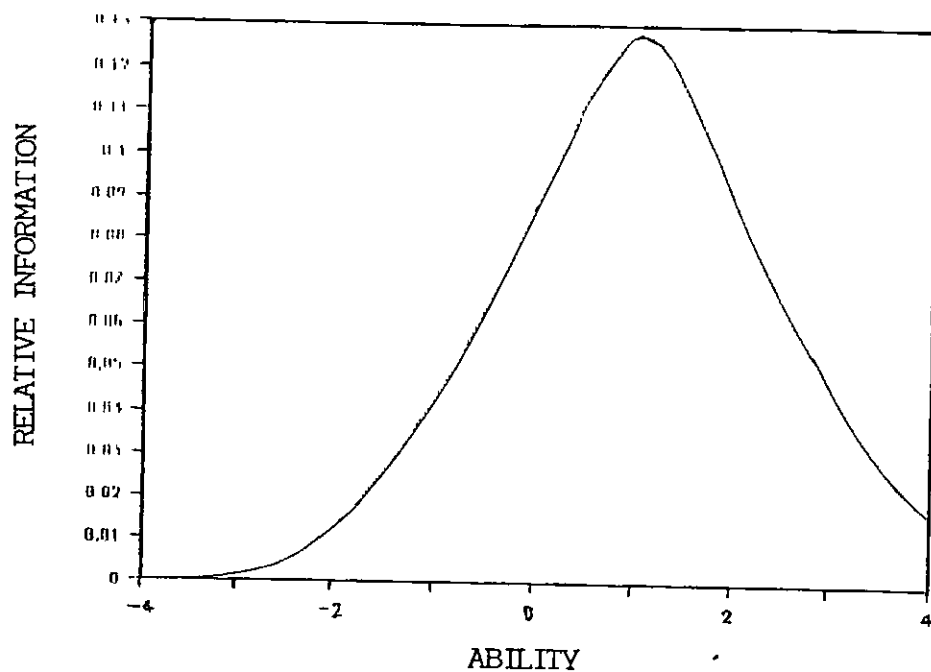
* หมายถึง ค่าพารามิเตอร์ของข้อสอบที่มีค่าต่ำสุด

ผลการวิเคราะห์ข้อสอบในแบบทดสอบชั้น ม.1 จากตาราง 13 ได้ค่าประมาณพารามิเตอร์ความยากของข้อสอบมีค่าสูงสุด 4.348 และต่ำสุด -2.099 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .552 และ 1.274 ตามลำดับ ค่าประมาณพารามิเตอร์อำนาจจำแนกของข้อสอบมีค่าสูงสุด 1.505 และต่ำสุด .193 โดยมีค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน .758 และ .289 ตามลำดับ ค่าประมาณพารามิเตอร์การเดามีค่าสูงสุด .445 และต่ำสุด .150 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .231 และ .056 ตามลำดับ

เมื่อนำค่าพารามิเตอร์ของข้อสอบ มาคำนวณค่าฟังก์ชันสารสนเทศของแบบทดสอบ และค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบ ที่ระดับความสามารถ (0) ระหว่าง -4 ถึง +4 รวม 161 ระดับ ปรากฏว่าแบบทดสอบชั้น ม.1 มีค่าฟังก์ชันสารสนเทศของแบบทดสอบระหว่าง .005 ถึง 10.779 โดยมีค่าสูงสุดที่ระดับความสามารถ 1.050 และแบบทดสอบชั้น ม.1 นี้มีค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบระหว่าง 19.846 ถึง 83.513 (ดังตาราง 37 ในภาคผนวก) เมื่อนำค่าฟังก์ชันสารสนเทศของแบบทดสอบมาเขียนภาพได้โค้งสารสนเทศของแบบทดสอบ ดังภาพประกอบ 8(ก) และนำค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบมาเขียนภาพ

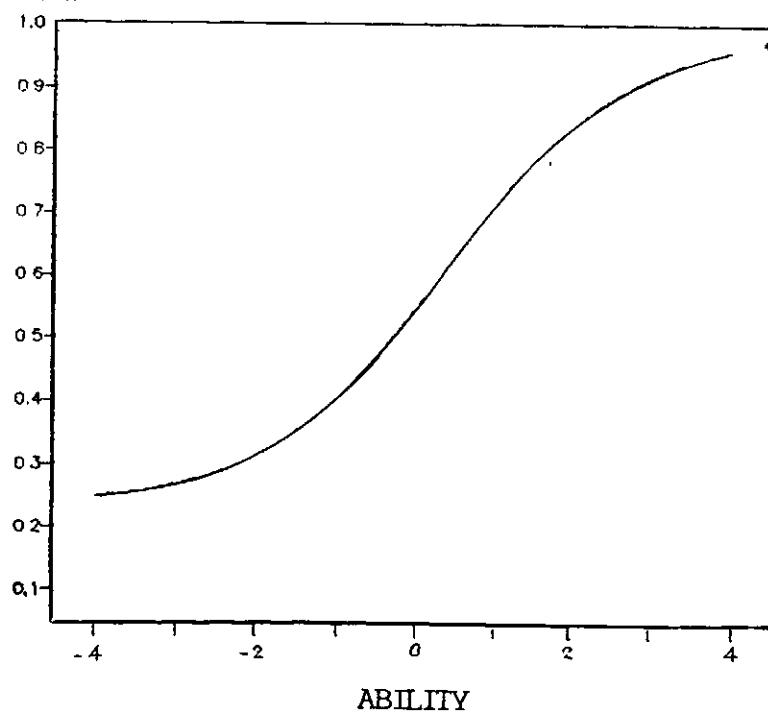
ได้โค้งลักษณะข้อสอบของแบบทดสอบ ดังภาพประกอบ 8(ข)

อนึ่งค่าฟังก์ชันสารสนเทศของแบบทดสอบชั้น ม.1 ซึ่งแสดงในภาพประกอบ 8(ก) และค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบชั้น ม.1 ซึ่งแสดงในภาพประกอบ 8(ข) ใช้ค่าฟังก์ชันซึ่งเฉลี่ยตามจำนวนข้อสอบเพื่อให้สเกลการวัดอยู่ระหว่าง 0 ถึง 1



ภาพประกอบ 8(ก) โค้งสารสนเทศของแบบทดสอบชั้น ม.1

$\Sigma P(\theta)/1$



ภาพประกอบ 8(ข) โค้งลักษณะข้อสอบของแบบทดสอบชั้น ม.1

ภาพประกอบ 8 โค้งสารสนเทศของแบบทดสอบ และโค้งลักษณะข้อสอบของแบบทดสอบชั้น ม.1

ตาราง 14 ค่าประมาณพารามิเตอร์ของข้อสอบในแบบทดสอบชั้น ม.2

ข้อ	a	S.E. _a	b	S.E. _b	c	S.E. _c
1	0.439	0.074	-0.778	0.209	0.338	0.099
2	0.577	0.070	-2.198	0.289	0.256	0.096
3	0.279	0.049	-1.000	0.229	0.282	0.099
4	0.345	0.123	3.108	1.174	0.431	0.048
5	0.637	0.136	0.598	0.247	0.342	0.059
6	0.519	0.082	-0.282	0.161	0.291	0.082
7	0.578	0.099	-0.121	0.168	0.326	0.077
8	1.243	0.204	1.020	0.306	0.254	0.025
9	0.394	0.050	-2.784*	0.366	0.229	0.090
10	0.578	0.239	3.941	1.738	0.152	0.019
11	0.480	0.125	1.675	0.508	0.264	0.052
12	0.438	0.067	-0.098	0.135	0.209	0.074
13	0.577	0.112	0.746	0.240	0.265	0.057
14	1.200	0.267	1.326	0.484	0.332	0.025
15	0.455	0.056	-0.314	0.112	0.168	0.065
16	0.755	0.148	1.023	0.302	0.301	0.040
17	0.460	0.098	0.755	0.266	0.310	0.070
18	0.367	0.045	-2.329	0.302	0.220	0.087
19	0.404	0.131	2.789	0.963	0.263	0.045
20	0.353	0.050	-1.134	0.200	0.229	0.087
21	0.413	0.124	2.647	0.848	0.206	0.044
22	0.503	0.084	0.137	0.156	0.243	0.072
23	0.673	0.140	0.865	0.293	0.378	0.048
24	0.417	0.144	3.155	1.152	0.249	0.040
25	0.342	0.120	3.337	1.232	0.346	0.047
26	0.302	0.064	0.929	0.269	0.230	0.078
27	0.409	0.111	1.583	0.504	0.301	0.063

ตาราง 14 (ต่อ)

ข้อ	a	S.E. _a	b	S.E. _b	c	S.E. _c
28	0.397	0.065	-0.051	0.146	0.229	0.080
29	0.728	0.074	-0.889	0.130	0.182	0.067
30	0.412	0.048	-1.360	0.188	0.202	0.080
31	0.475	0.063	-1.100	0.192	0.258	0.091
32	0.947	0.188	1.251	0.370	0.286	0.030
33	0.643	0.076	-0.720	0.136	0.206	0.072
34	1.026	0.347	2.196	0.980	0.281	0.021
35	0.777	0.118	0.914	0.220	0.196	0.036
36	0.906	0.158	1.160	0.305	0.247	0.030
37	0.983	0.118	-0.533	0.136	0.321	0.061
38	1.058	0.116	-0.807	0.127	0.196	0.060
39	0.413	0.106	1.900	0.546	0.224	0.054
40	1.185	0.173	1.240	0.294	0.144	0.019
41	0.537	0.055	-1.243	0.152	0.165	0.067
42	0.729	0.171	1.590	0.474	0.272	0.034
43	0.700	0.078	-0.809	0.137	0.209	0.072
44	0.661	0.076	-1.311	0.185	0.240	0.085
45	0.644	0.085	-0.041	0.126	0.213	0.060
46	0.942	0.099	-0.421	0.102	0.171	0.051
47	0.735	0.115	0.426	0.172	0.251	0.050
48	0.851	0.089	-0.966	0.142	0.218	0.070
49	0.674	0.065	-0.709	0.116	0.175	0.062
50	0.666	0.086	-0.665	0.148	0.252	0.077
51	0.363	0.102	2.238	0.686	0.240	0.057
52	0.853	0.090	-0.676	0.125	0.225	0.063
53	0.659	0.093	-0.308	0.135	0.244	0.070
54	0.903	0.082	-1.211	0.136	0.135*	0.055

ตาราง 14 (ต่อ)

ข้อ	a	S.E. _a	b	S.E. _b	c	S.E. _c
55	0.281	0.081	2.580	0.790	0.260	0.066
56	0.899	0.084	-1.341	0.157	0.179	0.068
57	0.826	0.085	-1.170	0.155	0.205	0.074
58	0.692	0.089	-0.369	0.128	0.231	0.067
59	0.929	0.127	0.701	0.178	0.175	0.032
60	0.865	0.131	0.254	0.168	0.342	0.048
61	0.825	0.072	-0.892	0.111	0.137	0.053
62	1.375	0.192	0.534	0.196	0.259	0.029
63	1.204	0.183	0.730	0.233	0.272	0.029
64	1.314	0.214	0.610	0.240	0.333	0.030
65	1.262	0.196	1.203	0.319	0.177	0.021
66	0.171*	0.053	4.254**	1.375	0.500*	0.071
67	1.355	0.321	1.321	0.546	0.332	0.023
68	1.113	0.230	1.328	0.424	0.256	0.025
69	1.679	0.232	0.695	0.250	0.249	0.023
70	0.724	0.177	1.509	0.474	0.310	0.037
71	0.237	0.074	3.405	1.121	0.337	0.067
72	0.786	0.144	1.416	0.340	0.179	0.029
73	0.664	0.086	-0.367	0.127	0.215	0.068
74	1.942**	0.322	1.155	0.455	0.200	0.017
75	0.758	0.129	0.719	0.218	0.259	0.044
76	0.831	0.142	0.720	0.232	0.311	0.041
77	0.847	0.115	-0.222	0.135	0.297	0.059
78	1.002	0.154	1.038	0.269	0.255	0.029
79	0.952	0.162	0.552	0.211	0.308	0.041
80	0.453	0.075	-0.182	0.157	0.260	0.084
81	0.539	0.134	1.616	0.480	0.260	0.047

ตาราง 14 (ต่อ)

ข้อ	a	S.E. _a	b	S.E. _b	c	S.E. _c
82	0.218	0.057	2.155	0.615	0.272	0.081
83	0.502	0.122	1.528	0.444	0.251	0.051
84	0.613	0.283	4.435	2.214	0.181	0.017
85	0.463	0.095	0.779	0.248	0.247	0.067
ค่าเฉลี่ย	.710		.569		.252	
ส่วนเบี่ยงเบนมาตรฐาน	.335		1.494		.064	

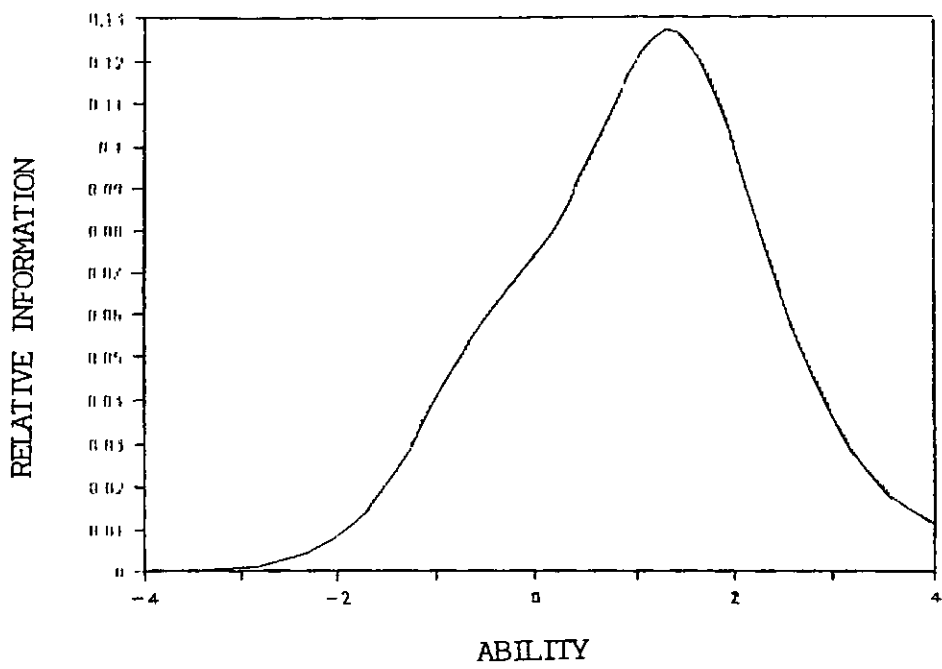
** หมายถึง ค่าพารามิเตอร์ของข้อสอบที่มีค่าสูงสุด

* หมายถึง ค่าพารามิเตอร์ของข้อสอบที่มีค่าต่ำสุด

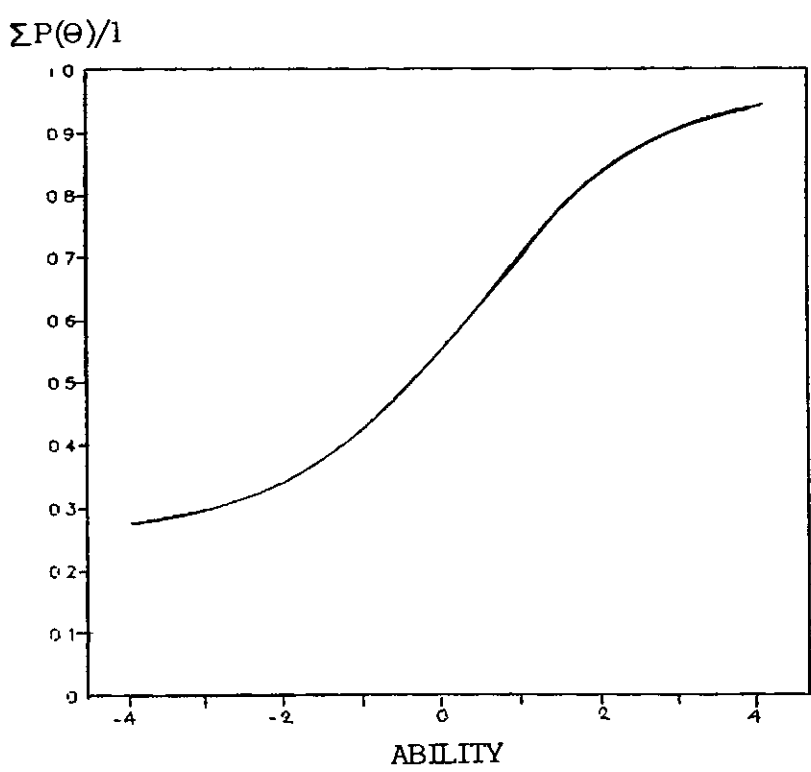
ผลการวิเคราะห์ข้อสอบในแบบทดสอบชั้น ม.2 จากตาราง 14 ได้ค่าประมาณพารามิเตอร์ความยากของข้อสอบมีค่าสูงสุด 4.254 และต่ำสุด -2.784 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .569 และ 1.494 ตามลำดับ ค่าประมาณพารามิเตอร์อำนาจจำแนกของข้อสอบมีค่าสูงสุด 1.942 และต่ำสุด .171 โดยมีค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน .710 และ .335 ตามลำดับ ค่าประมาณพารามิเตอร์การเดามีค่าสูงสุด .500 และต่ำสุด .135 โดยมีค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน .252 และ .064 ตามลำดับ

เมื่อนำค่าพารามิเตอร์ของข้อสอบ มาคำนวณค่าฟังก์ชันสารสนเทศของแบบทดสอบ และค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบ ที่ระดับความสามารถ (0) ระหว่าง -4 ถึง +4 รวม 161 ระดับ ปรากฏว่าแบบทดสอบชั้น ม.2 มีค่าฟังก์ชันสารสนเทศของแบบทดสอบระหว่าง .007 ถึง 10.760 โดยมีค่าสูงสุดที่ระดับความสามารถ 1.350 และแบบทดสอบชั้น ม.2 นี้มีค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบระหว่าง 21.734 ถึง 82.613 (ดังตาราง 38 ในภาคผนวก) เมื่อนำค่าฟังก์ชันสารสนเทศของแบบทดสอบมาเขียนภาพได้โค้งสารสนเทศของแบบทดสอบ ดังภาพประกอบ 9(ก) และนำค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบมาเขียนภาพได้โค้งลักษณะข้อสอบของแบบทดสอบ ดังภาพประกอบ 9(ข)

อนึ่งค่าฟังก์ชันสารสนเทศของแบบทดสอบซึ่งแสดงในภาพประกอบ 9(ก) และ ค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบซึ่งแสดงในภาพประกอบ 9(ข) ใช้ค่าฟังก์ชันมีสเกลการวัดระหว่าง 0 ถึง 1 เช่นเดียวกับค่าฟังก์ชันของแบบทดสอบชั้น ม.1



ภาพประกอบ 9(ก) โค้งสารสนเทศของแบบทดสอบชั้น ม.2



ภาพประกอบ 9(ข) โค้งลักษณะข้อสอบของแบบทดสอบชั้น ม.2

ภาพประกอบ 10 โค้งสารสนเทศของแบบทดสอบ และโค้งลักษณะข้อสอบของแบบทดสอบชั้น ม.2

ตาราง 15 ค่าประมาณพารามิเตอร์ของข้อสอบในแบบทดสอบชั้น ม.3

ข้อ	a	S.E. _a	b	S.E. _b	c	S.E. _c
1	0.704	0.112	-2.693	0.459	0.247	0.100
2	0.423	0.069	-0.065	0.149	0.224	0.080
3	0.227*	0.041	-1.902	0.367	0.218	0.091
4	0.411	0.065	-0.470	0.155	0.221	0.085
5	0.496	0.063	-1.745	0.245	0.210	0.088
6	0.354	0.053	-2.064	0.326	0.209	0.089
7	0.309	0.098	2.749	0.920	0.266	0.062
8	0.466	0.064	-0.258	0.117	0.160	0.067
9	0.728	0.094	-2.913*	0.408	0.209	0.089
10	0.673	0.091	-2.885	0.417	0.201	0.087
11	0.740	0.079	-2.060	0.249	0.194	0.082
12	0.494	0.080	-0.606	0.176	0.265	0.091
13	0.216	0.041	-1.414	0.295	0.217	0.090
14	0.416	0.057	-2.521	0.361	0.200	0.086
15	0.379	0.147	3.720**	1.506	0.311	0.040
16	0.377	0.058	0.040	0.124	0.165	0.069
17	0.381	0.049	-1.445	0.209	0.177	0.077
18	0.741	0.177	2.383	0.650	0.089	0.019
19	0.478*	0.065	-0.222	0.118	0.164	0.067
20	0.412	0.055	-1.339	0.205	0.185	0.080
21	0.684	0.091	-0.502	0.128	0.188	0.069
22	0.783	0.345	3.076	1.549	0.208	0.020
23	0.592	0.075	-0.956	0.168	0.229	0.082
24	0.511	0.059	-1.874	0.234	0.168	0.074
25	0.852	0.093	-1.544	0.196	0.155	0.067
26	0.421	0.054	-1.364	0.199	0.172	0.075
27	0.444	0.063	-0.766	0.167	0.222	0.085

ตาราง 15 (ต่อ)

ข้อ	a	S.E. _a	b	S.E. _b	c	S.E. _c
28	0.430	0.060	-0.046	0.112	0.148	0.063
29	0.496	0.059	-1.910	0.250	0.200	0.085
30	0.514	0.064	-0.634	0.126	0.158	0.067
31	0.540	0.073	-0.067	0.121	0.169	0.064
32	0.482	0.075	0.411	0.147	0.156	0.060
33	0.335	0.060	-0.002	0.158	0.231	0.086
34	0.461	0.059	-1.066	0.179	0.211	0.084
35	0.547	0.104	0.654	0.223	0.246	0.062
36	0.566	0.119	1.346	0.352	0.183	0.047
37	0.428	0.055	-2.037	0.282	0.202	0.086
38	0.604	0.066	-1.528	0.192	0.166	0.072
39	0.577	0.115	1.227	0.315	0.187	0.047
40	0.556	0.123	0.989	0.311	0.270	0.058
41	0.815	0.175	0.816	0.294	0.314	0.046
42	0.698	0.114	0.718	0.201	0.185	0.046
43	0.741	0.127	0.699	0.223	0.275	0.046
44	0.537	0.083	0.574	0.165	0.153	0.054
45	0.372	0.120	3.087	1.042	0.187	0.046
46	0.372	0.121	3.094	1.053	0.189	0.046
47	0.447	0.082	0.017	0.173	0.275	0.084
48	0.351	0.105	1.898	0.635	0.303	0.069
49	0.251	0.045	-0.675	0.182	0.214	0.088
50	0.272	0.086	2.839	0.947	0.290	0.068
51	0.599	0.121	1.284	0.331	0.194	0.044
52	0.363	0.070	0.884	0.236	0.183	0.067
53	0.632	0.074	-0.612	0.123	0.169	0.064
54	0.415	0.068	0.110	0.143	0.196	0.074

ตาราง 15 (ต่อ)

ข้อ	a	S.E. _a	b	S.E. _b	c	S.E. _c
55	0.876	0.113	-1.488	0.225	0.218	0.084
56	0.895	0.158	1.513	0.360	0.141*	0.024
57	0.614	0.086	-0.310	0.137	0.234	0.071
58	0.761	0.100	-0.560	0.136	0.225	0.069
59	0.638	0.103	0.910	0.217	0.154	0.043
60	0.663	0.081	-0.847	0.148	0.199	0.075
61	0.745	0.084	-0.953	0.147	0.189	0.071
62	0.450	0.059	-0.985	0.162	0.171	0.073
63	1.244	0.355	1.661	0.739	0.272	0.023
64	0.450	0.147	2.597	0.915	0.246	0.044
65	1.192	0.214	0.657	0.258	0.314**	0.034
66	1.285*	0.214	0.391	0.204	0.328	0.036
67	0.979	0.222	1.358	0.455	0.294	0.030
68	0.301	0.086	1.835	0.586	0.295	0.077
69	0.403	0.073	0.432	0.176	0.202	0.073
70	0.381	0.096	2.068	0.567	0.176	0.054
71	0.655	0.067	-0.832	0.120	0.134	0.058
72	0.693	0.219	2.090	0.779	0.295	0.033
73	0.462	0.099	1.200	0.324	0.194	0.059
74	0.638	0.151	1.559	0.453	0.226	0.040
75	0.599	0.139	1.412	0.409	0.244	0.046
76	0.710	0.082	-0.466	0.121	0.188	0.064
77	0.659	0.101	0.034	0.145	0.229	0.065
78	0.721	0.073	-1.418	0.173	0.180	0.073
79	0.726	0.261	2.911	1.176	0.150	0.021
80	1.096	0.198	0.877	0.292	0.282	0.032
81	0.475	0.082	0.835	0.208	0.153	0.055

ตาราง 15 (ต่อ)

ข้อ	a	S.E. _a	b	S.E. _b	c	S.E. _c
82	0.332	0.072	0.891	0.269	0.222	0.078
83	0.655	0.139	0.965	0.299	0.263	0.050
84	0.870	0.118	-0.057	0.134	0.266	0.054
85	0.571	0.169	2.118	0.710	0.257	0.039
ค่าเฉลี่ย	.575		.151		.213	
ส่วนเบี่ยงเบนมาตรฐาน	.219		1.567		.049	

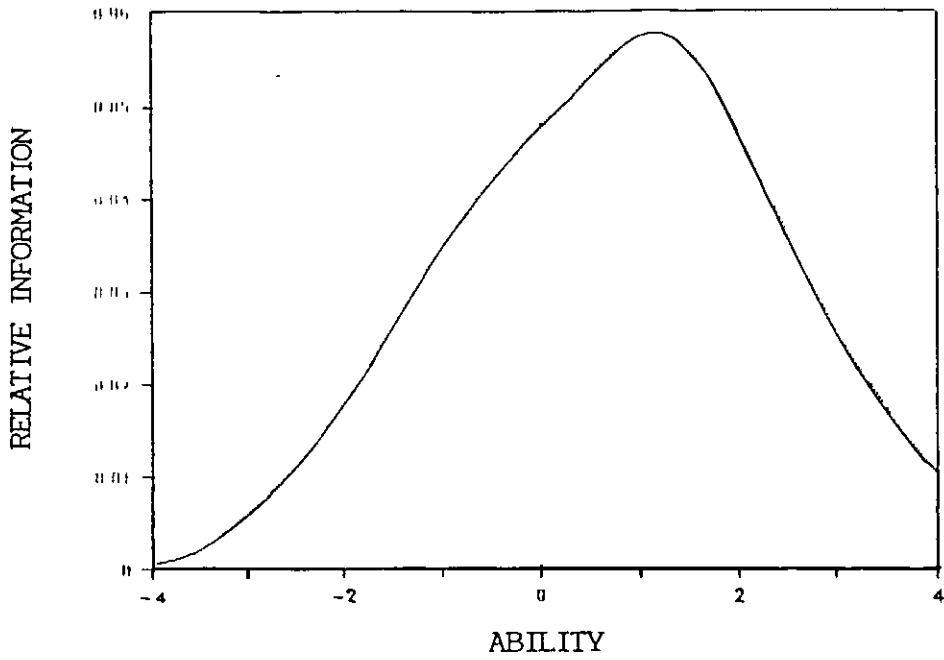
★★ หมายถึง ค่าพารามิเตอร์ของข้อสอบที่มีค่าสูงสุด

* หมายถึง ค่าพารามิเตอร์ของข้อสอบที่มีค่าต่ำสุด

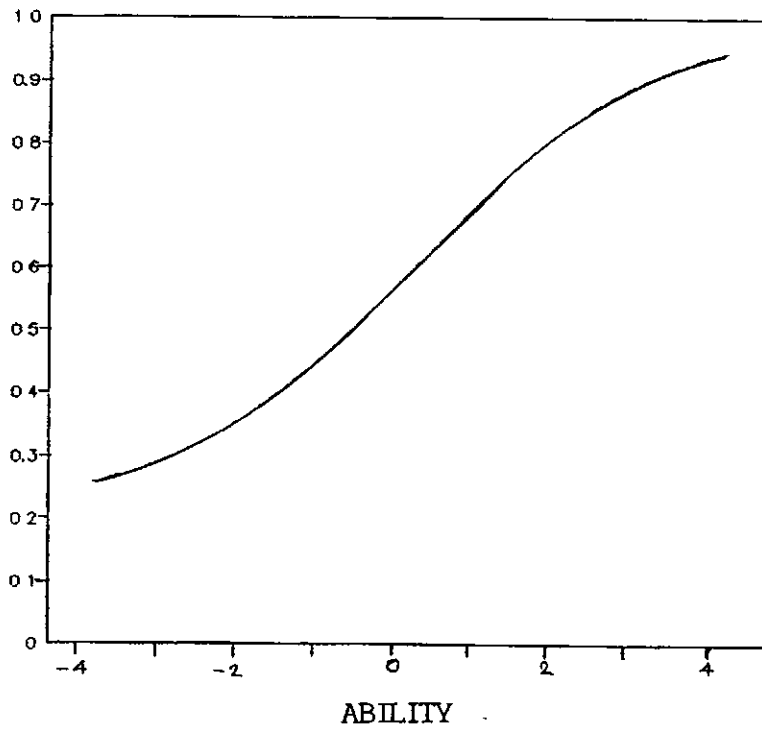
ผลการวิเคราะห์ข้อสอบในแบบทดสอบชั้น ม.3 จากตาราง 15 แสดงว่าค่าประมาณพารามิเตอร์ความยากของข้อสอบมีค่าสูงสุด 3.720 และต่ำสุด -2.913 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .151 และ 1.567 ตามลำดับ ค่าประมาณพารามิเตอร์อำนาจจำแนกของข้อสอบมีค่าสูงสุด 1.285 และต่ำสุด .227 โดยมีค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน .575 และ .219 ตามลำดับ ค่าประมาณพารามิเตอร์การเดามีค่าสูงสุด .314 และต่ำสุด .141 โดยมีค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน .213 และ .049 ตามลำดับ

เมื่อนำค่าพารามิเตอร์ของข้อสอบ มาวิเคราะห์ค่าฟังก์ชันสารสนเทศของแบบทดสอบ และฟังก์ชันลักษณะข้อสอบของแบบทดสอบ ที่ระดับความสามารถ (0) ระหว่าง -4 ถึง +4 รวม 161 ระดับ ปรากฏว่าแบบทดสอบชั้น ม.3 มีค่าสารสนเทศระหว่าง .051 ถึง 4.892 โดยมีค่าสูงสุดที่ระดับความสามารถ 1.150 และแบบทดสอบชั้น ม.3 นี้มีค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบระหว่าง 18.733 ถึง 83.644 (ดังตาราง 39 ในภาคผนวก)เมื่อนำค่าฟังก์ชันสารสนเทศของแบบทดสอบมาเขียนภาพได้โค้งสารสนเทศของแบบทดสอบ ดังภาพประกอบ 10(ก) และนำค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบมาเขียนภาพได้โค้งลักษณะข้อสอบของแบบทดสอบ ดังภาพประกอบ 10(ข)

อนึ่งค่าฟังก์ชันสารสนเทศของแบบทดสอบชั้น ม.3 ซึ่งแสดงในภาพประกอบ 10(ก)และค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบชั้น ม.3 ซึ่งแสดงในภาพประกอบ 10 (ข) ใช้ค่าฟังก์ชันที่มีสเกลการวัดระหว่าง 0 ถึง 1 เช่นเดียวกับค่าฟังก์ชันของแบบทดสอบชั้น ม.1



ภาพประกอบ 10(ก) โค้งสารสนเทศของแบบทดสอบชั้น ม.3
 $\Sigma P(\theta)/1$



ภาพประกอบ 10(ข) โค้งลักษณะข้อสอบของแบบทดสอบชั้น ม.3

ภาพประกอบ 10 โค้งสารสนเทศของแบบทดสอบ และโค้งลักษณะข้อสอบของแบบทดสอบชั้น ม.3

ตาราง 16 ค่าประมาณพารามิเตอร์ของข้อสอบในแบบทดสอบความสามารถทั่วไปทางภาษาไทย

ข้อ	a	S.E. _a	b	S.E. _b	c	S.E. _c
1	2.960	1.403	1.413	2.526	0.263	0.011
2	1.774	0.322	1.613	0.648	0.226	0.010
3	1.283	0.233	1.509	0.501	0.373**	0.016
4	1.716	0.116	0.912	0.169	0.187	0.013
5	1.798	0.367	1.572	0.730	0.277	0.011
6	0.449	0.043	0.274	0.102	0.105	0.048
7	1.299	0.105	0.734	0.148	0.246	0.017
8	1.628	0.264	1.833	0.569	0.133	0.008
9	1.950	0.258	1.335	0.489	0.240	0.012
10	1.830	0.225	1.338	0.426	0.199	0.011
11	3.885**	2.237	1.371	3.946	0.141	0.007
12	2.544	0.224	0.991	0.361	0.277	0.013
13	0.386	0.033	0.423	0.084	0.071	0.037
14	3.376	1.422	1.348	2.526	0.150	0.009
15	0.536	0.037	-0.586	0.086	0.100*	0.049
16	0.451	0.140	3.459**	1.134	0.280	0.024
17	0.582	0.117	2.310	0.528	0.196	0.023
18	0.373	0.050	1.187	0.208	0.112	0.048
19	0.694	0.034	-1.061*	0.072	0.060	0.032
20	0.346	0.082	2.139	0.566	0.266	0.053
21	0.378	0.044	0.741	0.143	0.109	0.049
22	0.458	0.028	-0.282	0.050	0.048	0.026
23	0.475	0.029	0.341	0.049	0.033	0.018
24	0.623	0.036	-0.482	0.063	0.065	0.034
25	0.412	0.034	0.253	0.077	0.073	0.038
26	0.516	0.032	-0.285	0.058	0.060	0.032
27	0.383	0.030	0.846	0.089	0.044*	0.024

ตาราง 16 (ต่อ)

ข้อ	a	S.E. _a	b	S.E. _b	c	S.E. _c
28	0.437	0.070	1.341	0.276	0.175	0.047
29	0.343*	0.039	0.536	0.126	0.111	0.053
ค่าเฉลี่ย	1.169		.936		.160	
ส่วนเบี่ยงเบนมาตรฐาน	1.011		.964		.091	

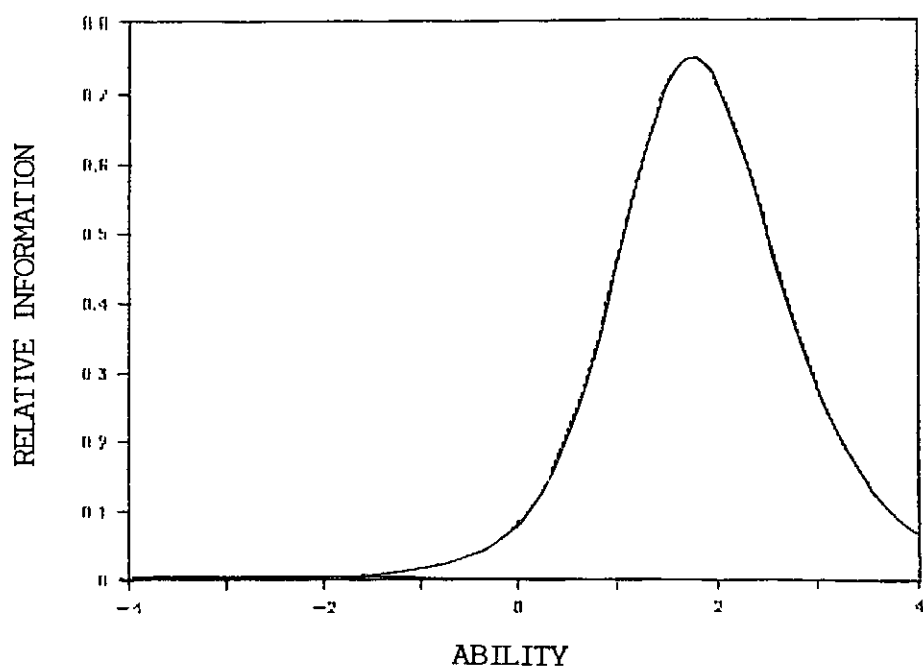
★★ หมายถึง ค่าพารามิเตอร์ของข้อสอบที่มีค่าสูงสุด

* หมายถึง ค่าพารามิเตอร์ของข้อสอบที่มีค่าต่ำสุด

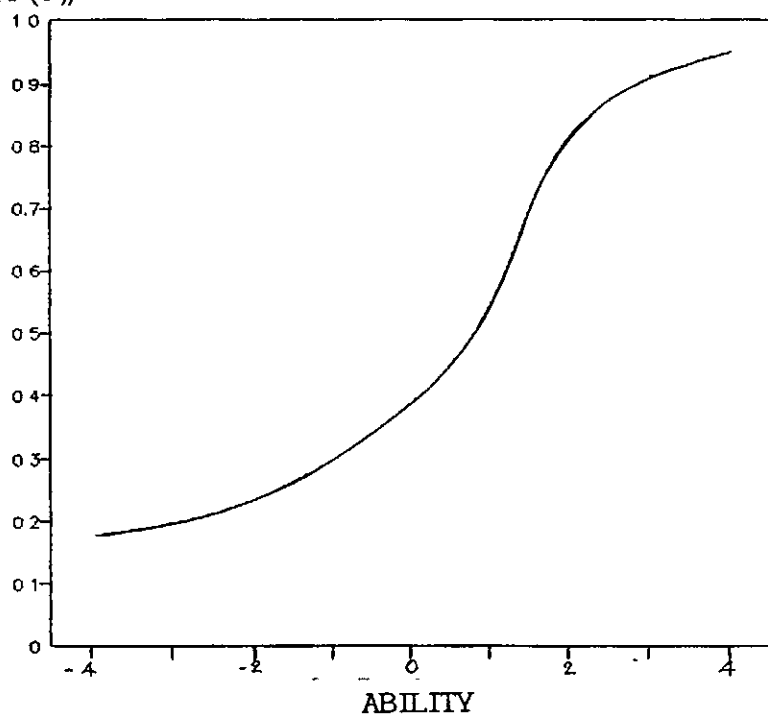
ผลการวิเคราะห์ข้อสอบในแบบทดสอบความสามารถทั่วไปทางภาษาไทย จากตาราง 16 แสดงว่าค่าประมาณพารามิเตอร์ความยากมีค่าสูงสุด 3.474 และต่ำสุด -1.069 โดยมีค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน .936 และ .964 ตามลำดับ ค่าประมาณพารามิเตอร์อำนาจจำแนกของข้อสอบมีค่าสูงสุด 3.886 และต่ำสุด .342 โดยมีค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน 1.169 และ 1.011 ตามลำดับ ค่าประมาณพารามิเตอร์การเดามีค่าสูงสุด .375 และต่ำสุด .100 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .160 และ .091 ตามลำดับ

เมื่อนำค่าพารามิเตอร์ของข้อสอบ มาวิเคราะห์ค่าฟังก์ชันสารสนเทศของแบบทดสอบ และฟังก์ชันลักษณะข้อสอบของแบบทดสอบ ที่ระดับความสามารถ (0) ระหว่าง -4 ถึง +4 รวม 161 ระดับ ปรากฏว่าแบบทดสอบความสามารถทั่วไปทางภาษาไทย มีค่าฟังก์ชันสารสนเทศของแบบทดสอบระหว่าง 0 ถึง 21.532 โดยมีค่าสูงสุดที่ระดับความสามารถ 1.750 และมีค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบระหว่าง 4.640 ถึง 28.575 (ดังตาราง 40 ใน ภาคผนวก) เมื่อนำค่าฟังก์ชันสารสนเทศของแบบทดสอบมาเขียนภาพได้โค้งสารสนเทศของแบบทดสอบ ดังภาพประกอบ 11(ก) และนำค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบมาเขียนภาพได้โค้งลักษณะข้อสอบของแบบทดสอบ ดังภาพประกอบ 11(ข)

อนึ่งค่าฟังก์ชันสารสนเทศของแบบทดสอบซึ่งแสดงในภาพประกอบ 11(ก) และค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบซึ่งแสดงในภาพประกอบ 11(ข) ใช้ค่าฟังก์ชันที่มีสเกลการวัด 0 ถึง 1 เช่นเดียวกับแบบทดสอบฉบับอื่น ๆ



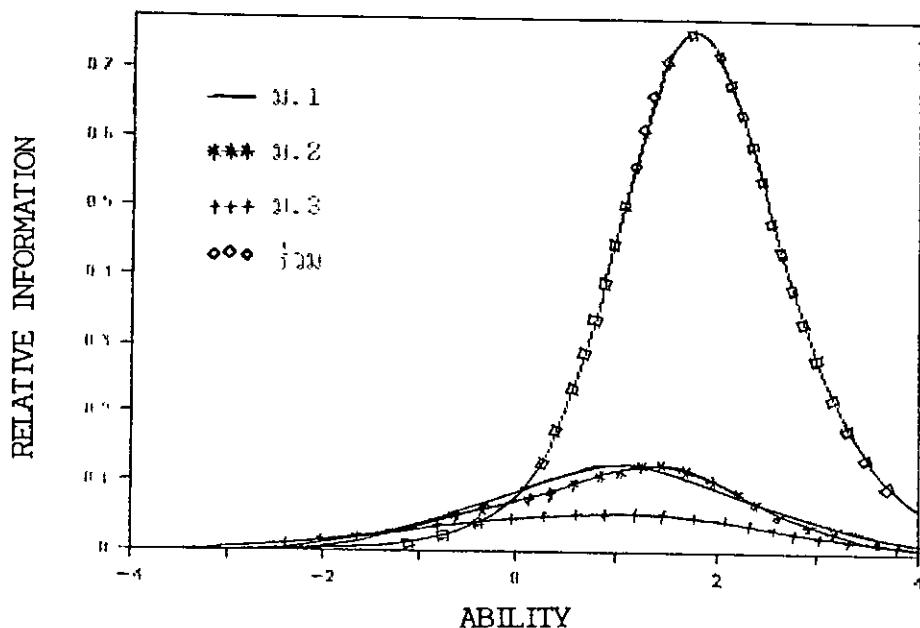
ภาพประกอบ 11(ก) โค้งสารสนเทศของแบบทดสอบความสามารถทั่วไปทางภาษาไทย
 $\Sigma P(\theta)/1$



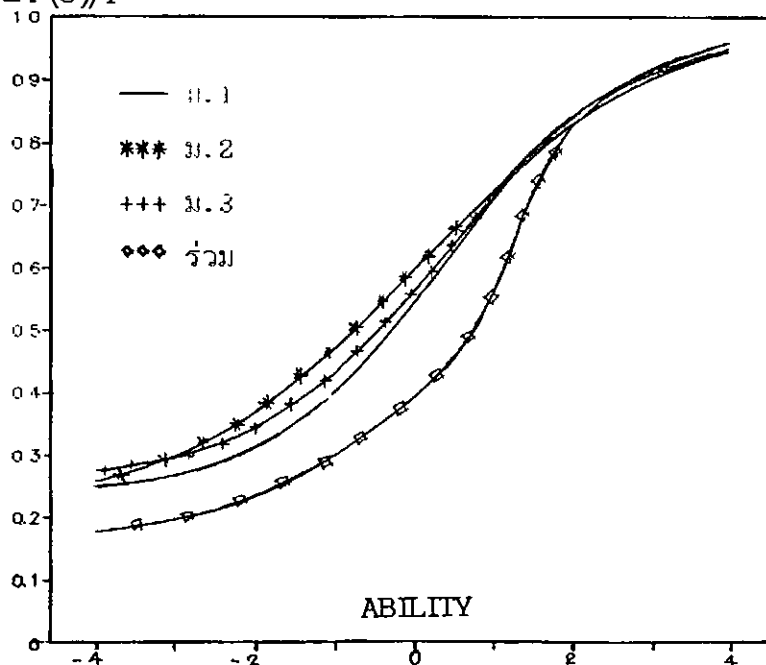
ภาพประกอบ 11(ข) โค้งลักษณะข้อสอบของแบบทดสอบความสามารถทั่วไปทางภาษาไทย

ภาพประกอบ 11 โค้งสารสนเทศของแบบทดสอบ และโค้งลักษณะข้อสอบของ
 แบบทดสอบความสามารถทั่วไปทางภาษาไทย

เพื่อให้เห็นการเปรียบเทียบโค้งสารสนเทศของแบบทดสอบและโค้งลักษณะข้อสอบ
ของแบบทดสอบฉบับต่าง ๆ ได้ชัดเจนขึ้น ผู้วิจัยนำค่าสารสนเทศของแบบทดสอบทั้ง 4 ฉบับ มา
เขียนในภาพเดียวกัน ดังภาพประกอบ 12(ก) และนำภาพโค้งลักษณะข้อสอบของแบบทดสอบทั้ง
4 ฉบับ มาเขียนในภาพเดียวกัน ดังภาพประกอบ 12(ข)



ภาพประกอบ 12(ก) โค้งสารสนเทศของแบบทดสอบฉบับต่าง ๆ
 $\Sigma P(\theta)/1$



ภาพประกอบ 12(ข) โค้งลักษณะข้อสอบของแบบทดสอบฉบับต่าง ๆ

ภาพประกอบ 12 โค้งสารสนเทศและโค้งลักษณะข้อสอบของแบบทดสอบฉบับต่าง ๆ

1.3 ผลการวิเคราะห์ค่าสถิติพื้นฐานของแบบทดสอบฉบับต่าง ๆ

จากการตรวจให้คะแนนผลการตอบแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนและแบบทดสอบความสามารถทั่วไปจากกรสอบของนักเรียนชั้น ม.1 ม.2 และ ม.3 ที่เป็นกลุ่มตัวอย่างการวิจัย สรุปค่าสถิติพื้นฐานของแบบทดสอบ 4 ฉบับ ของนักเรียนชั้นมัธยมศึกษาตอนต้น ได้ดังตาราง 17

ตาราง 17 ค่าสถิติพื้นฐานของแบบทดสอบฉบับต่าง ๆ ก่อนการเทียบคะแนน

ค่าสถิติ	แบบทดสอบ			
	ม.1	ม.2	ม.3	ความสามารถทั่วไป
จำนวนข้อสอบ	85	85	85	29
จำนวนผู้สอบ	1245	1280	1171	3696
คะแนนสูงสุด	78	78	77	28
คะแนนต่ำสุด	16	17	15	1
คะแนนเฉลี่ย	46.601	48.362	50.344	12.581
ส่วนเบี่ยงเบนมาตรฐาน	12.709	12.007	11.138	5.116
มัธยฐาน	47.000	49.000	51.000	12.000
ฐานนิยม	36.000	53.000	50.000	9.000
ความโด่ง	-.655	-.641	-.249	-.324
ความเบ้	-.003	-.121	-.237	.482
ความเชื่อมั่น	.894	.880	.867	.771

ผลการตอบแบบทดสอบเทียบคะแนนของนักเรียน 3 ชั้นจากตาราง 17 แสดงว่าก่อนการเทียบคะแนน นักเรียนชั้น ม.3 มีค่าเฉลี่ยของคะแนนสูงสุด รองลงมาคือ นักเรียนชั้น ม.2 และ นักเรียนชั้น ม.1 ตามลำดับ พิสัยของคะแนนที่ได้จากแบบทดสอบเทียบคะแนน 3 ฉบับ มีค่าใกล้เคียงกัน แต่พิสัยของคะแนนที่ได้จากแบบทดสอบความสามารถทั่วไปทางภาษาไทยกว้างมาก

ตอนที่ 2 ผลการสร้างตารางเทียบคะแนนแบบทดสอบ

การสร้างตารางเทียบคะแนนแบบทดสอบวิชาภาษาไทย ชั้น ม.1 ม.2 ม.3 และแบบทดสอบความสามารถทั่วไปทางภาษาไทย ดำเนินการตามขั้นตอนที่กล่าวไว้ในบทที่ 3 กล่าวคือ ผู้วิจัยได้วิเคราะห์ข้อสอบในแบบทดสอบแต่ละฉบับด้วยโปรแกรม BILOG เวอร์ชัน 3.04 โดยใช้แบบจำลองโลจิสติกที่มีพารามิเตอร์ 3 ตัว เพื่อให้ได้ค่าประมาณพารามิเตอร์ของข้อสอบแต่ละข้อ คือค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าการเดา (c) และประมาณค่าความสามารถของผู้สอบ (θ) แต่ละคน จากนั้นนำค่าพารามิเตอร์ดังกล่าว มาคำนวณหาค่าคงที่เพื่อใช้ในการแปลงค่าประมาณพารามิเตอร์ของข้อสอบและพารามิเตอร์ของผู้สอบให้อยู่ในสเกลเดียวกัน คืออยู่ในสเกลของแบบทดสอบความสามารถทั่วไปทางภาษาไทย จากนั้นนำค่าพารามิเตอร์ที่แปลงให้อยู่ในสเกลเดียวกันแล้วโดยใช้สมการในตาราง 18 ไปคำนวณหาคะแนนจริงของผู้สอบแต่ละคน ซึ่งได้จากแบบทดสอบแต่ละฉบับ แล้วนำคะแนนสังเกตของคะแนนจริงที่ได้จากแบบทดสอบแต่ละฉบับซึ่งตรงกับตำแหน่งเปอร์เซ็นต์ไทล์เดียวกันมาสร้างเป็นตารางเทียบคะแนน ตารางเทียบคะแนนแบบทดสอบ ม.1 ม.2 และ ม.3 ที่สร้างขึ้นแสดงดังตาราง 19

ตาราง 18 สมการที่ใช้ในการแปลงค่าพารามิเตอร์ของข้อสอบและพารามิเตอร์ของผู้สอบ

ชั้น	b	a	θ
ม.1	$b'_{1i} = .910b_{1i} - .004$	$a'_{1i} = 1.099a_{1i}$	$\theta'_{1i} = .910\theta_{1i} - .004$
ม.2	$b'_{2i} = .913b_{2i} - .012$	$a'_{2i} = 1.096a_{2i}$	$\theta'_{2i} = .913\theta_{2i} - .012$
ม.3	$b'_{3i} = .929b_{3i} + .001$	$a'_{3i} = 1.076a_{3i}$	$\theta'_{3i} = .929\theta_{3i} + .001$

ตาราง 19 คะแนนสมมูลจากการเทียบคะแนนแบบทดสอบชั้น ม.1 ม.2 และ ม.3

ม.1	ม.2	ม.3
12-13	13-14	10-11
14-15	15-16	12-13
16-17	17-18	14-15
18	19-20	16-17
19	21	18
20	22	19
21	23-24	20-21
22-24	25-26	22
25-26	27-28	23-25
27	29	26-27
28	30	28-29
29-30	31-32	30
31-32	33	31
33	34	32-33
34-35	35-37	34
36	38	35
37	39	36-37
38-39	40	38
40-42	41-42	39-40
43	43-44	41-42
44	45	43
45	46	44
46-47	47-48	45-46
48	49	47
49	50	48
50-51	51-52	49-50

ตาราง 19 (ต่อ)

ม.1	ม.2	ม.3
52	53	51
53-54	54	52-53
55	55-56	54
56	57	55-56
57-59	58-59	57-58
60-61	60	59-60
62-63	61-63	61-62
64-66	64-65	63-66
67-68	66-67	67
69	68	68
70	69	69
71-72	70-71	70-72
73	72-73	73
74-76	74-75	74
77	76	75-76
78	77-78	77
79	79	78
80	80	79
81	81	80
82-83	82-83	81
84	84	82
85	85	83-85

ตาราง 19 แสดงคะแนนที่สมมูลกันจากแบบทดสอบชั้น ม.1 ม.2 และม.3 เรียงจากคะแนนน้อยไปมาก จากตารางคะแนนที่อยู่ในแถวเดียวกันเป็นคะแนนที่สมมูลกัน และนักเรียนที่ได้คะแนนสมมูลกันดังกล่าวเป็นนักเรียนที่มีความสามารถเท่าเทียมกัน เช่น นักเรียนชั้น ม.1 ที่

สอบแบบทดสอบชั้น ม.1 ได้ 20 คะแนน นักเรียนชั้น ม.2 ที่สอบแบบทดสอบชั้น ม.2 ได้ 22 คะแนน และนักเรียนชั้น ม. 3 ที่สอบแบบทดสอบชั้น ม. 3 ได้ 19 คะแนน เป็นนักเรียนที่มีความสามารถเท่าเทียมกันหรือกล่าวได้ว่า สำหรับนักเรียนแต่ละชั้นที่มีความสามารถเท่าเทียมกัน (มีคะแนนจริงเท่ากัน) ถ้านักเรียนชั้น ม.1 ทำแบบทดสอบชั้น ม.1 ได้ 20 คะแนน นักเรียนชั้น ม.2 ทำแบบทดสอบชั้น ม.2 ได้ 22 คะแนน และนักเรียนชั้น ม.3 ทำแบบทดสอบชั้น ม.3 ได้ 19 คะแนน ถือว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนที่มีระดับความสามารถดังกล่าว มีความก้าวหน้าตามปกติหรือกล่าวได้ว่าสำหรับนักเรียนแต่ละชั้นที่มีความสามารถเท่าเทียมกัน ถ้านักเรียนชั้น ม.1 ทำแบบทดสอบชั้น ม.1 ได้ 20 คะแนน นักเรียนชั้น ม.2 ทำแบบทดสอบ ม.2 ได้ 22 คะแนน นักเรียนชั้น ม. 3 ทำแบบทดสอบชั้น ม. 3 ได้ 19 คะแนน ถือว่านักเรียนที่มีระดับความสามารถดังกล่าวมีผลสัมฤทธิ์ทางการเรียนจากชั้น ม.1 ไป ม.2 ก้าวหน้ามากกว่าปกติและผลสัมฤทธิ์ทางการเรียนจากชั้น ม.2 ไป ม.3 ก้าวหน้ามากกว่าปกติ

ตารางเทียบคะแนนที่สร้างขึ้นสามารถนำไปใช้ในการเปรียบเทียบความสามารถทางภาษาไทยของนักเรียนชั้นมัธยมศึกษาตอนต้นได้ เช่นต้องการเปรียบเทียบความสามารถของนักเรียนชั้น ม.1 และชั้น ม.2 ก็ดำเนินการโดยให้นักเรียนชั้น ม.1 สอบแบบทดสอบชั้น ม.1 (ท 101) และให้นักเรียนชั้น ม.2 สอบแบบทดสอบชั้น ม.2(ท 203) แล้วตรวจให้คะแนนนักเรียนแต่ละคน สมมติว่านักเรียนชั้น ม.1 ได้ 49 คะแนน และนักเรียนชั้น ม.2 ได้ 49 คะแนนเท่ากัน เมื่อนำคะแนนของนักเรียนทั้ง 2 คน ไปเปรียบเทียบกับตารางเทียบคะแนน จะเห็นว่า นักเรียนชั้นม.1 ที่ได้ 49 คะแนน มีความสามารถเท่ากับนักเรียนชั้น ม.2 ที่ได้ 50 คะแนน แสดงว่า นักเรียนชั้น ม.1 คนนี้มีความสามารถสูงกว่า นักเรียนชั้น ม.2 ที่นำมาเปรียบเทียบกัน

นอกจากการสร้างตารางเทียบคะแนนแบบทดสอบชั้น ม.1 ม.2 และ ม.3 แล้ว ผู้วิจัยสร้างตารางเทียบคะแนนแบบทดสอบเทียบคะแนนแต่ละชั้นกับแบบทดสอบรวมโดยใช้วิธีการเดียวกับการสร้างตารางเทียบคะแนนแบบทดสอบชั้นต่าง ๆ เพื่อใช้ในการศึกษาประสิทธิภาพของวิธีการเทียบคะแนน ตารางเทียบคะแนนแบบทดสอบชั้น ม.1 กับ แบบทดสอบความสามารถทั่วไป แสดงดังตาราง 20 ตารางเทียบคะแนนแบบทดสอบชั้น ม.2 กับ แบบทดสอบความสามารถทั่วไปแสดงดังตาราง 21 และ ตารางเทียบคะแนนแบบทดสอบชั้น ม.3 กับแบบทดสอบความสามารถทั่วไปแสดงดังตาราง 22

ตาราง 20 คะแนนสมมูลจากการเทียบคะแนนแบบทดสอบชั้น ม.1 กับแบบทดสอบความสามารถทั่วไปทางภาษาไทย

แบบทดสอบ ม.1	แบบทดสอบความสามารถทั่วไป
12-13	1.0
14-15	2.0
16-17	3.0
18-22	4.0
23	4.5
24-25	5.0
26	5.5
27	6.5
28-29	7.0
30-32	8.0
33	9.5
34-35	10.0
36	10.5
37-38	11.0
39	12.0
40	13.0
41-42	14.0
43	15.0
44	16.0
45	16.5
46-57	17.0
58	18.0
59	19.0
60	19.5
61	20.0

ตาราง 20 (ต่อ)

แบบทดสอบ ม.1	แบบทดสอบความสามารถทั่วไป
62	21.5
63-65	23.0
66	23.5
67-69	24.0
70	24.5
71-72	25.0
73	25.5
74-75	26.0
76-78	27.5

ตาราง 21 คะแนนสมมูลจากการเทียบคะแนนแบบทดสอบชั้น ม.2 กับแบบทดสอบความสามารถทั่วไปทางภาษาไทย

แบบทดสอบ ม.2	แบบทดสอบความสามารถทั่วไป
13-14	1.0
15-16	2.0
17-18	3.0
19-25	4.0
26-27	5.0
28-29	6.0
30	6.5
31	7.0
32-33	8.0
34	9.0

ตาราง 21 (ต่อ)

แบบทดสอบ.ม.2	แบบทดสอบความสามารถทั่วไป
35-37	10.0
38	11.0
39	11.5
40	12.0
41	13.0
42	14.0
43	15.0
44-45	16.0
46	16.5
47-57	17.0
58	17.5
59	19.0
60	20.0
61	21.5
62	22.0
63-64	23.0
65	23.5
66-69	24.0
70-71	25.0
72-74	26.0
75-78	27.5

ตาราง 22 คะแนนสมมูลจากการเทียบคะแนนแบบทดสอบชั้น ม.3 กับ แบบทดสอบความสามารถทั่วไปทางภาษาไทย

แบบทดสอบ ม.3	แบบทดสอบความสามารถทั่วไป
10-11	1.0
12-13	2.0
14-15	3.0
16-22	4.0
23	5.0
24-27	6.0
28-29	7.0
30	7.5
31	8.0
32	9.0
33-34	10.0
35-36	11.0
37	11.5
38	12.0
39	13.0
40	13.5
41	15.0
42	15.5
43	16.0
44	16.5
45-56	17.0
57	17.5
58	18.5
59	19.5
60	20.5

ตาราง 22 (ต่อ)

แบบทดสอบ ม.3	แบบทดสอบความสามารถทั่วไป
61	22.0
62	22.5
63-65	23.0
66-68	24.0
69	24.5
70-72	25.0
73	25.5
74	26.0
75-77	27.5

จากตาราง 20-22 จะเห็นว่า คะแนนจากแบบทดสอบความสามารถทั่วไปทางภาษาไทย 1 คะแนน อาจตรงกับคะแนนจากแบบทดสอบเทียบคะแนนหลายค่าและคะแนนจากแบบทดสอบเทียบคะแนนที่เทียบไปเป็นคะแนนตามสเกลของแบบทดสอบความสามารถทั่วไปทางภาษาไทย อาจมีค่าเป็นเลขทศนิยม จากตาราง 20 นักเรียนชั้น ม.1 ที่ได้คะแนนจากแบบทดสอบ ม.1 25 คะแนน จะถือว่านักเรียนคนนี้มี kemampuan เท่าเทียมกับนักเรียนที่ได้คะแนนจากแบบทดสอบความสามารถทั่วไปทางภาษาไทย 5 คะแนน และในทำนองเดียวกัน โดยใช้ข้อมูลจากตาราง 21 นักเรียนชั้น ม.2 ที่ได้คะแนนจากแบบทดสอบ ชั้น ม.2 26 คะแนน จะถือว่านักเรียนคนนี้มี kemampuan เท่าเทียมกับนักเรียนที่ได้คะแนนจากแบบทดสอบความสามารถทั่วไปทางภาษาไทย 5 คะแนน และในทำนองเดียวกัน โดยใช้ข้อมูลจากตาราง 22 นักเรียนชั้น ม.3 ที่ได้คะแนนจากแบบทดสอบชั้น ม.3 23 คะแนน จะถือว่านักเรียนคนนี้มี kemampuan เท่าเทียมกับนักเรียนที่ได้คะแนนจากแบบทดสอบความสามารถทั่วไปทางภาษาไทย 5 คะแนน

ตอนที่ 3 ผลการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียน ชั้นมัธยมศึกษาตอนต้น

การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ม.2 และ ม.3 ดำเนินการโดยเทียบคะแนนที่นักเรียนชั้น ม.2 ทำได้จากการตอบแบบทดสอบชั้น ม.2

ไปสู่คะแนนที่เท่าเทียมกันตามสเกลคะแนนของแบบทดสอบชั้น ม.1 และในทำนองเดียวกัน เทียบคะแนนนักเรียนชั้น ม.3 ไปสู่คะแนนที่เท่าเทียมกันตามสเกลคะแนนของแบบทดสอบชั้น ม.1 ค่าสถิติพื้นฐานของคะแนนของนักเรียนชั้น ม.2 และ ม.3 ซึ่งเทียบให้อยู่ในสเกลเดียวกันคะแนนของนักเรียนชั้น ม.1 แล้ว แสดงดังตาราง 23

ตาราง 23 ค่าสถิติพื้นฐานของผลการสอบของนักเรียนชั้น ม.1 ม.2 และ ม.3 ภายหลังการเทียบคะแนน

ค่าสถิติ	ม.1	ม.2	ม.3
จำนวนข้อสอบ	85	85	85
จำนวนผู้สอบ	1245	1280	1171
คะแนนสูงสุด	77	78	78
คะแนนต่ำสุด	16	16.5	16.5
คะแนนเฉลี่ย	46.601	47.566	51.192
ส่วนเบี่ยงเบนมาตรฐาน	12.709	12.793	11.072
ความเอนเอียง	-.665	-.652	-.194
ความเบ้	-.033	-.083	-.286

เมื่อเทียบคะแนนผลการสอบของนักเรียนชั้น ม.1 ม.2 และ ม.3 ให้อยู่ในสเกลเดียวกันแล้ว ผู้วิจัยทดสอบความแตกต่างระหว่างคะแนนเฉลี่ยของนักเรียน 3 ระดับชั้น โดยการวิเคราะห์ความแปรปรวนแบบทางเดียว (ONE-WAY ANALYSIS OF VARIANCE) ผลการวิเคราะห์ความแปรปรวนของนักเรียน 3 ระดับชั้น สรุปได้ดังตาราง 24

ตาราง 24 ผลการวิเคราะห์ความแปรปรวนของคะแนนนักเรียนชั้น ม.1 ม.2 และ ม.3
ในกลุ่มรวม

แหล่งความแปรปรวน	SS	df	MS	F
ระหว่างกลุ่ม	14049.283	2	7024.852	46.852*
ภายในกลุ่ม	553706.039	3693	149.934	
รวม	567755.322	3696		

* มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง 24 แสดงว่าเมื่อเทียบคะแนนของนักเรียน 3 ชั้นให้อยู่ในสเกลเดียวกัน แล้วผลการเปรียบเทียบความแตกต่างของค่าเฉลี่ยของคะแนนนักเรียน 3 ชั้น แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .01

เมื่อทดสอบความแตกต่างของคะแนนเฉลี่ยของนักเรียนแต่ละชั้นเป็นรายคู่ ได้ผลดังตาราง 25

ตาราง 25 ผลการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของคะแนนนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มรวม

ระดับชั้น	ม.1	ม.2	ม.3
ม.1	—	.965	4.591*
ม.2		—	3.626*
ม.3			—

* มีนัยสำคัญทางสถิติที่ระดับ .01

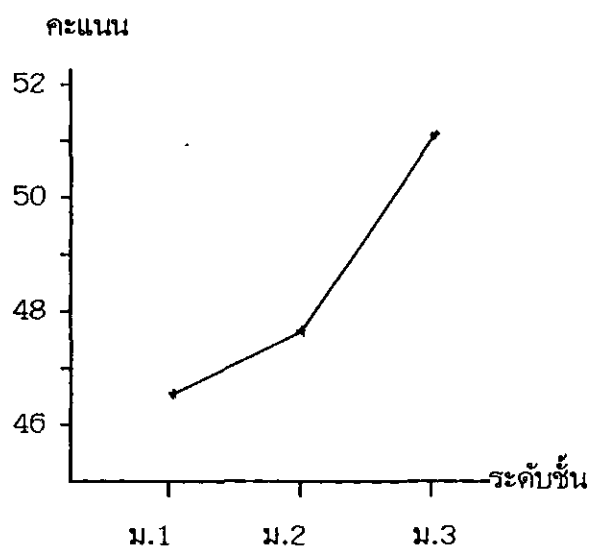
ผลการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของคะแนนนักเรียนชั้นต่าง ๆ ในกลุ่มรวมพบว่า คะแนนเฉลี่ยของนักเรียนชั้น ม.1 กับ ม.3 แตกต่างกัน คะแนนเฉลี่ยของนักเรียนชั้น ม.2

กับ ม.3 แตกต่างกัน แต่คะแนนเฉลี่ยของนักเรียนชั้น ม.1 กับ ม.2 ไม่แตกต่างกัน สรุปได้ว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.2 มีความก้าวหน้าตามปกติ ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.3 มีความก้าวหน้ามากกว่าปกติโดยเฉลี่ย 4.591 คะแนน ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ไป ม.3 มีความก้าวหน้ามากกว่าปกติโดยเฉลี่ย 3.626 คะแนน

เมื่อวิเคราะห์แนวโน้มการเปลี่ยนแปลงค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้นต่าง ๆ ปรากฏว่าการเปลี่ยนแปลงค่าเฉลี่ยของคะแนนนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มรวมมีแนวโน้มไม่ใช่เส้นตรง ผลการวิเคราะห์แนวโน้มแสดงดังตาราง 26 และภาพแสดงแนวโน้มแสดงดังภาพประกอบ 13

ตาราง 26 ผลการวิเคราะห์แนวโน้มของการเปลี่ยนแปลงคะแนนเฉลี่ยของนักเรียนในกลุ่มรวม

แหล่งความแปรปรวน	SS	df	MS	F	p
ระหว่างกลุ่ม	14049.283	2	7024.641	46.852	.000
แนวโน้มเชิงเส้นตรง	12568.689	1	12568.689	83.828	.000
แนวโน้มที่ไม่ใช่เส้นตรง	1480.594	1	1480.594	9.875	.002
ภายในกลุ่ม	553706.039	3693	149.934		
รวม	567755.322	3696			



ภาพประกอบ 13 แนวโน้มการเปลี่ยนแปลงคะแนนเฉลี่ยของนักเรียนในกลุ่มรวม

ข้อมูลจากตาราง 26 และ ภาพประกอบ 13 แสดงให้เห็นว่า การเปลี่ยนแปลงค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ม.2 และ ม.3 มีแนวโน้มไม่ใช่เส้นตรง แต่เนื่องจากผลรวมกำลังสองของแนวโน้มที่มีไม่ใช่เส้นตรงมีชั้นแห่งความเป็นอิสระเพียง 1 จึงไม่สามารถวิเคราะห์ต่อว่าแนวโน้มของการเปลี่ยนแปลงค่าเฉลี่ยเป็นรูปใด

ในการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้นมัธยมศึกษาตอนต้น 3 ชั้นนี้ เพื่อให้ได้ข้อมูลที่ชัดเจนเกี่ยวกับความก้าวหน้าในการเรียนของนักเรียนที่มีระดับความสามารถแตกต่างกัน ผู้วิจัยได้แบ่งนักเรียนแต่ละระดับชั้นเป็น 3 กลุ่มตามระดับความสามารถ ผลจากการจัดแบ่งกลุ่มนักเรียนดังกล่าว แสดงดังตาราง 27

ตาราง 27 ค่าสถิติพื้นฐานของนักเรียนจำแนกตามชั้นและระดับความสามารถ

ค่าสถิติ	ระดับความสามารถ		
	ต่ำ	กลาง	สูง
ม.1			
จำนวนผู้สอบ	246	236	248
คะแนนสูงสุด	41.000	54.000	77.000
คะแนนต่ำสุด	16.000	37.000	53.000
คะแนนเฉลี่ย	29.045	45.920	64.347
ส่วนเบี่ยงเบนมาตรฐาน	5.130	2.918	4.772
ม.2			
จำนวนผู้สอบ	256	264	264
คะแนนสูงสุด	43.000	55.000	78.000
คะแนนต่ำสุด	16.500	38.500	55.000
คะแนนเฉลี่ย	29.787	47.119	64.845
ส่วนเบี่ยงเบนมาตรฐาน	5.190	2.848	4.747
ม.3			
จำนวนผู้สอบ	229	237	237
คะแนนสูงสุด	43.000	56.000	78.000
คะแนนต่ำสุด	16.500	46.500	56.000
คะแนนเฉลี่ย	35.085	51.458	65.768
ส่วนเบี่ยงเบนมาตรฐาน	5.920	2.300	4.748

การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนของนักเรียนในแต่ละระดับความสามารถ ผู้วิจัยศึกษาความก้าวหน้าโดยเปรียบเทียบคะแนนเฉลี่ยของนักเรียน 3 ชั้น จำแนกตามระดับความสามารถ โดยวิธีการในตนเองเกี่ยวกับการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนของนักเรียน 3 ชั้นในกลุ่มรวม ผลการวิเคราะห์ความแปรปรวนของคะแนนของนักเรียน 3 ชั้นในกลุ่มย่อยแต่ละกลุ่ม แสดงดังตาราง 28-33

ตาราง 28 ผลการวิเคราะห์ความแปรปรวนของคะแนนของนักเรียน ชั้น ม.1 ม.2 และ ม.3
ในกลุ่มนักเรียนที่มีความสามารถสูง

แหล่งความแปรปรวน	SS	df	MS	F
ระหว่างกลุ่ม	251.365	2	125.683	6.048*
ภายในกลุ่ม	15502.046	746	20.780	
รวม	15753.411	749		

* มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง 28 แสดงว่าในกลุ่มนักเรียนที่มีความสามารถสูง คะแนนเฉลี่ยของนักเรียน
ชั้น ม.1 ม.2 และ ม.3 แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อทดสอบความแตกต่างของคะแนนเฉลี่ยของนักเรียนแต่ละชั้นเป็นรายคู่ได้ผลดังตาราง
29

ตาราง 29 ผลการทดสอบความแตกต่างระหว่างคะแนนเฉลี่ยของนักเรียนชั้น ม.1 ม.2 และ
ม.3 ในกลุ่มนักเรียนที่มีความสามารถสูง

ระดับชั้น	ม.1	ม.2	ม.3
ม.1	-	.498	1.421*
ม.2		-	.923
ม.3			-

* มีนัยสำคัญทางสถิติที่ระดับ .01

ผลการทดสอบความแตกต่างระหว่างคะแนนเฉลี่ยของนักเรียนชั้นต่าง ๆ ในกลุ่มนักเรียน
ที่มีความสามารถสูงพบว่าคะแนนเฉลี่ยของนักเรียนชั้น ม.1 กับ ม.2 ไม่แตกต่างกัน คะแนนเฉลี่ย
ของนักเรียนชั้น ม.2 กับ ม.3 ไม่แตกต่างกัน แต่คะแนนเฉลี่ยของนักเรียนชั้น ม.1 กับ ม.3

แตกต่างกัน สรุปได้ว่าในกลุ่มนักเรียนที่มีความสามารถสูง ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.2 ก้าวหน้าตามปกติ ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ไป ม.3 ก้าวหน้าตามปกติ แต่ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.3 ก้าวหน้ามากกว่าปกติโดยเฉลี่ย 1.421 คะแนน

ตาราง 30 ผลการวิเคราะห์ความแปรปรวนของคะแนนของนักเรียน ชั้น ม.1 ม.2 และ ม.3 ในกลุ่มนักเรียนที่มีความสามารถปานกลาง

แหล่งความแปรปรวน	SS	df	MS	F
ระหว่างกลุ่ม	4047.455	2	2023.728	275.975*
ภายในกลุ่ม	5382.540	734	7.333	
รวม	9429.995	737		

* มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง 30 แสดงว่าในกลุ่มนักเรียนที่มีความสามารถปานกลางคะแนนเฉลี่ยของนักเรียนชั้น ม.1 ม.2 และ ม.3 แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อทดสอบความแตกต่างของคะแนนเฉลี่ยของนักเรียนแต่ละชั้นเป็นรายคู่ ได้ผลดังตาราง 31

ตาราง 31 ผลการทดสอบความแตกต่างระหว่างคะแนนเฉลี่ยของนักเรียน ชั้น ม.1 ม.2 และ ม.3 ในกลุ่มนักเรียนที่มีความสามารถปานกลาง

ระดับชั้น	ม.1	ม.2	ม.3
ม.1	—	1.199	5.538*
ม.2		—	4.339*
ม.3			—

* มีนัยสำคัญทางสถิติที่ระดับ .01

ผลการทดสอบความแตกต่างระหว่างคะแนนเฉลี่ยของนักเรียนชั้นต่างๆ ในกลุ่มนักเรียนที่มีความสามารถปานกลางพบว่า คะแนนเฉลี่ยของนักเรียน ชั้น ม.1 กับ ม.2 แตกต่างกัน คะแนนเฉลี่ยของนักเรียนชั้น ม.2 กับ ม.3 แตกต่างกันและคะแนนเฉลี่ยของนักเรียนชั้น ม.1 กับ ม.3 แตกต่างกัน สรุปได้ว่าในกลุ่มนักเรียนที่มีความสามารถปานกลาง ผลสัมฤทธิ์ทางการเรียนของนักเรียนทุกชั้นก้าวหน้ามากกว่าปกติ โดยผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.2 ก้าวหน้ามากกว่าปกติโดยเฉลี่ย 1.199 คะแนน ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ไป ม.3 ก้าวหน้ามากกว่าปกติโดยเฉลี่ย 4.339 คะแนน และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.3 ก้าวหน้ามากกว่าปกติโดยเฉลี่ย 5.538 คะแนน

ตาราง 32 ผลการวิเคราะห์ความแปรปรวนของคะแนนของนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มนักเรียนที่มีความสามารถต่ำ

แหล่งความแปรปรวน	SS	df	MS	F
ระหว่างกลุ่ม	5128.295	2	2564.148	87.711*
ภายในกลุ่ม	21545.801	728	29.234	
รวม	26674.096	731		

* มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง 32 แสดงว่าในกลุ่มนักเรียนที่มีความสามารถต่ำ คะแนนเฉลี่ยของนักเรียนชั้น ม.1 ม.2 และ ม.3 แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.1

เมื่อทดสอบความแตกต่างของคะแนนเฉลี่ยของนักเรียนแต่ละชั้นเป็นรายคู่ได้ผลดัง

ตาราง 33

ตาราง 33 ผลการทดสอบความแตกต่างระหว่างคะแนนเฉลี่ยของนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มนักเรียนที่มีความสามารถต่ำ

ระดับชั้น	ม.1	ม.2	ม.3
ม.1	-	.742	6.041*
ม.2		-	5.298*
ม.3			-

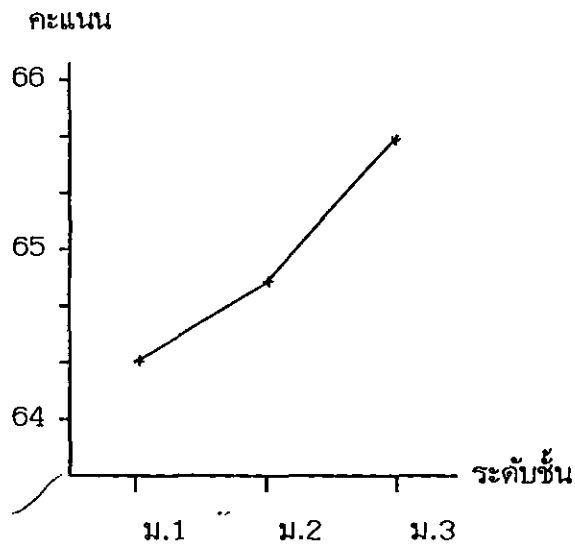
* มีนัยสำคัญทางสถิติที่ระดับ .01

ผลการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของคะแนนนักเรียนชั้นต่าง ๆ ในกลุ่มนักเรียนที่มีความสามารถต่ำ พบว่าคะแนนเฉลี่ยของนักเรียนชั้น ม.1 กับ ม.2 ไม่แตกต่างกัน แต่คะแนนเฉลี่ยของนักเรียนชั้น ม.2 กับ ม.3 แตกต่างกัน และคะแนนเฉลี่ยของนักเรียนชั้น ม.1 กับ ม.3 แตกต่างกัน สรุปได้ว่าในกลุ่มนักเรียนที่มีความสามารถต่ำ ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.2 ก้าวหน้าตามปกติ แต่ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ไป ม.3 ก้าวหน้ามากกว่าปกติโดยเฉลี่ย 5.298 คะแนน และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.3 ก้าวหน้ามากกว่าปกติโดยเฉลี่ย 6.041 คะแนน

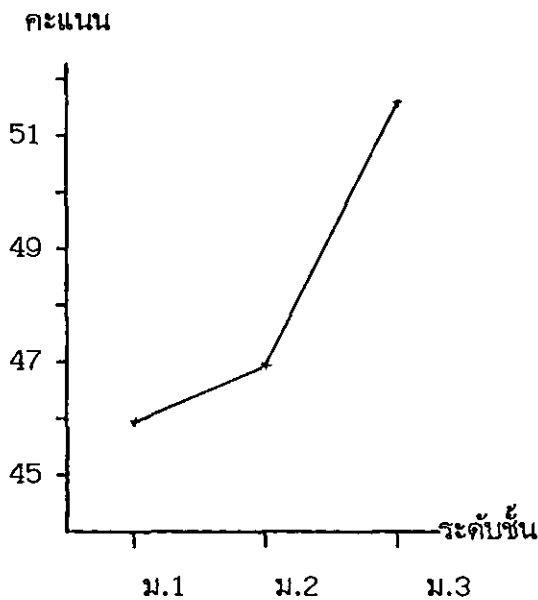
สำหรับการวิเคราะห์แนวโน้มการเปลี่ยนแปลงค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียนของนักเรียนในกลุ่มย่อยตามระดับความสามารถ ผู้วิจัยดำเนินการวิเคราะห์ด้วยวิธีการเดียวกับการวิเคราะห์ในกลุ่มรวม ผลการวิเคราะห์แสดงดังตาราง 34 และแนวโน้มการเปลี่ยนแปลงค่าเฉลี่ยแสดงดังภาพประกอบ 14

ตาราง 34 ผลการวิเคราะห์แนวโน้มของการเปลี่ยนแปลงคะแนนของนักเรียนในกลุ่มย่อย

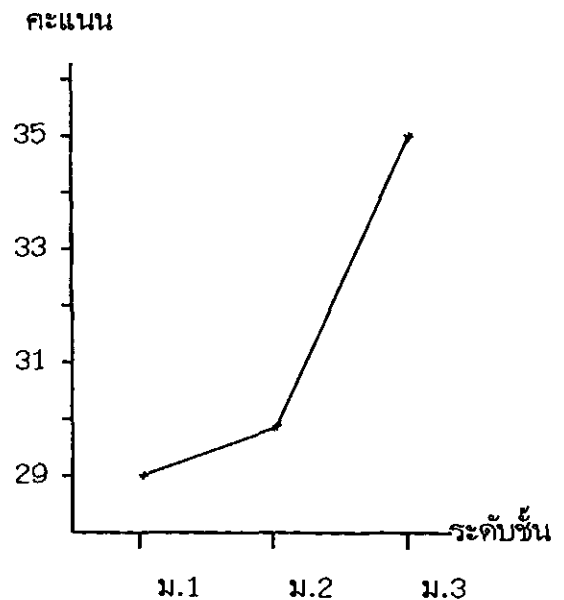
แหล่งความแปรปรวน	SS	df	MS	F	p
<u>กลุ่มความสามารถสูง</u>					
ระหว่างกลุ่ม	251.365	2	125.683	6.048	.003
แนวโน้มเชิงเส้นตรง	243.636	1	243.636	11.724	.001
แนวโน้มที่ไม่ใช่เส้นตรง	7.729	1	7.729	.372	.542
ภายในกลุ่ม	15502.046	746	20.780		
<u>กลุ่มความสามารถปานกลาง</u>					
ระหว่างกลุ่ม	4047.455	2	2023.728	275.975	.000
แนวโน้มเชิงเส้นตรง	3630.177	1	3630.177	495.036	.000
แนวโน้มที่ไม่ใช่เส้นตรง	417.278	1	417.278	56.903	.000
ภายในกลุ่ม	5382.540	734	7.333		
<u>กลุ่มความสามารถต่ำ</u>					
ระหว่างกลุ่ม	5128.295	2	2564.147	87.711	.000
แนวโน้มเชิงเส้นตรง	4245.886	1	4245.886	145.236	.000
แนวโน้มที่ไม่ใช่เส้นตรง	882.409	1	882.409	30.184	.000
ภายในกลุ่ม	21545.801	728	29.234		



ภาพประกอบ 14(ก) กลุ่มความสามารถสูง



ภาพประกอบ 14(ข) กลุ่มความสามารถปานกลาง



ภาพประกอบ 14(ค) กลุ่มความสามารถต่ำ

ภาพประกอบ 14 แนวโน้มของการเปลี่ยนแปลงคะแนนเฉลี่ยของนักเรียนในกลุ่มย่อย

ข้อมูลจากตาราง 34 และภาพประกอบ 14 แสดงให้เห็นว่าในกลุ่มนักเรียนที่มีความสามารถสูง การเปลี่ยนแปลงค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ม.2 และ ม.3 มีแนวโน้มเป็นเส้นตรง ส่วนในกลุ่มนักเรียนที่มีความสามารถปานกลาง และในกลุ่มนักเรียนที่มีความสามารถต่ำ การเปลี่ยนแปลงค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียนของนักเรียน 3 ชั้น

มีแนวโน้มที่ไม่ใช่เส้นตรง แต่เนื่องจากชั้นแห่งความเป็นอิสระของผลรวมกำลังสองของความแปรปรวนของแนวโน้มที่ไม่ใช่เส้นตรง มีเพียง 1 จึงไม่สามารถวิเคราะห์ต่อได้ว่าแนวโน้มเป็นรูปใด

ตอนที่ 4 ผลการศึกษาประสิทธิภาพของวิธีการเทียบคะแนน

การศึกษาประสิทธิภาพของวิธีเทียบคะแนน ศึกษาโดยใช้ข้อมูลจากกลุ่มสอบทานผล ซึ่งเป็นนักเรียนชั้นมัธยมศึกษาตอนต้น จากประชากรการวิจัยเดียวกับกลุ่มตัวอย่างเทียบคะแนน โดยกลุ่มนักเรียนมาร้อยละ 30 ของจำนวนนักเรียนในกลุ่มเทียบคะแนน นักเรียนแต่ละคนจะมีคะแนน 2 ส่วน คือคะแนนจากแบบทดสอบตามระดับชั้นของตน และคะแนนจากแบบทดสอบความสามารถทั่วไปทางภาษาไทย ประสิทธิภาพของวิธีการเทียบคะแนนได้จากการเปรียบเทียบดัชนีความแตกต่าง ซึ่งเป็นค่าความคลาดเคลื่อนรวมของการเทียบคะแนนกับเกณฑ์การยอมรับดัชนีความแตกต่าง

การคำนวณดัชนีความแตกต่าง เริ่มจากการใช้ตารางเทียบคะแนนคือ ตาราง 20-22 เทียบคะแนนที่นักเรียนแต่ละคนได้จากแบบทดสอบภาษาไทยตามระดับชั้นของตนไปเป็นคะแนนที่เท่าเทียมกับคะแนนจากแบบทดสอบความสามารถทั่วไปทางภาษาไทยแล้วคำนวณหาความแตกต่างของผลการเทียบคะแนน ค่าสถิติพื้นฐานของกลุ่มสอบทานผลหลังจากเทียบคะแนนแบบทดสอบแต่ละชั้นไปยังสเกลของแบบทดสอบความสามารถทั่วไปทางภาษาไทย แสดงดังตาราง 35 และผลการคำนวณดัชนีความแตกต่าง แสดงดังตาราง 36

ตาราง 35 ค่าสถิติพื้นฐานภายหลังการเทียบคะแนนของกลุ่มสอบทานผล

ค่าสถิติ	ระดับความสามารถ	
	เทียบคะแนน	ความสามารถทั่วไป
ม.1		
จำนวนผู้สอบ	373	373
คะแนนสูงสุด	24.500	16.000
คะแนนต่ำสุด	3.000	2.000
คะแนนเฉลี่ย	12.618	10.306
ส่วนเบี่ยงเบนมาตรฐาน	4.835	3.454
ม.2		
จำนวนผู้สอบ	384	384
คะแนนสูงสุด	21.500	20.000
คะแนนต่ำสุด	2.000	3.000
คะแนนเฉลี่ย	13.051	10.831
ส่วนเบี่ยงเบนมาตรฐาน	4.473	3.373
ม.3		
จำนวนผู้สอบ	352	352
คะแนนสูงสุด	27.000	27.000
คะแนนต่ำสุด	4.000	5.000
คะแนนเฉลี่ย	17.735	17.690
ส่วนเบี่ยงเบนมาตรฐาน	3.950	4.122

ตาราง 36 การคำนวณดัชนีความแตกต่าง

$\sum_{i=1}^n (x_i - \bar{x})^2$	n	S_x^2	S_x	C
12233.000	1109	24.357	4.935	.464*

* น่าพอใจในระดับปานกลาง (.244 < C ≤ .548)

จากตาราง 36 ผลการคำนวณดัชนีความแตกต่างของคะแนนที่นักเรียนแต่ละคนทำได้จากแบบทดสอบความสามารถทั่วไปทางภาษาไทยกับคะแนนจากแบบทดสอบตามระดับชั้นที่เทียบให้มีสเกลเดียวกับแบบทดสอบความสามารถทั่วไปทางภาษาไทยมีค่าเท่ากับ .464 ซึ่งเมื่อเปรียบเทียบกับเกณฑ์ที่กำหนดไว้ ปรากฏว่า ค่าดัชนีที่ได้อยู่ในระดับปานกลาง แสดงว่าการเทียบคะแนนโดยใช้วิธีการตามทฤษฎีการตอบสนองข้อสอบแบบจำลองโลจิสติก 3 พารามิเตอร์ที่ใช้ในการวิจัยนี้ให้ผลการเทียบคะแนนมีประสิทธิภาพในระดับปานกลาง

บทที่ 5

สรุป อภิปรายผลและข้อเสนอแนะ

การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้นมัธยมศึกษาตอนต้น โดยการเทียบคะแนนตามแนวตั้งตามทฤษฎีการตอบสนองข้อสอบ แบบจำลองโลจิสติกสามพารามิเตอร์นี้ มีประเด็นการศึกษาสรุปได้ดังนี้

จุดมุ่งหมายของการวิจัย

1. เพื่อศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 โดยเปรียบเทียบจากผลสัมฤทธิ์ทางการเรียนซึ่งเป็นคะแนนที่เทียบให้อยู่ในสเกลเดียวกันแล้ว
2. เพื่อศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 ในแต่ละกลุ่มความสามารถ โดยเปรียบเทียบจากผลสัมฤทธิ์ทางการเรียนซึ่งเป็นคะแนนที่เทียบให้อยู่ในสเกลเดียวกันแล้ว

สมมติฐานของการวิจัย

1. ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 มีความก้าวหน้าขึ้นตามลำดับ โดยผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1
2. ในกลุ่มนักเรียนที่มีความสามารถสูง ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 มีความก้าวหน้าขึ้นตามลำดับ โดยผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1
3. ในกลุ่มนักเรียนที่มีความสามารถปานกลาง ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 มีความก้าวหน้าขึ้นตามลำดับ โดยผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1
4. ในกลุ่มนักเรียนที่มีความสามารถต่ำ ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 มีความก้าวหน้าขึ้นตามลำดับ โดยผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 และผลสัมฤทธิ์ทาง

การเรียนของนักเรียนชั้น ม.2 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1

กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างที่ใช้ในการวิจัย ได้แก่ นักเรียนชั้น ม.1 ม.2 และ ม.3 ในภาคเรียนที่ 1 ปีการศึกษา 2537 ของโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา จังหวัดสิงห์บุรี กลุ่มตัวอย่างการวิจัย มี 2 กลุ่ม ได้แก่กลุ่มตัวอย่างเทียบคะแนนเป็นนักเรียนชั้น ม.1 ม.2 และ ม.3 จำนวน 1245 1280 และ 1171 คน ตามลำดับ และกลุ่มตัวอย่างสอบทานผล เป็นนักเรียนชั้น ม.1 ม.2 และ ม.3 จำนวน 373 384 และ 352 คน ตามลำดับ

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยมี 2 ประเภท คือ แบบทดสอบเทียบคะแนนและแบบทดสอบรวม แบบทดสอบเทียบคะแนนเป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย 3 วิชา คือ ท 101 ท 203 และ ท 305 ใช้กับนักเรียนชั้น ม.1 ม.2 และ ม.3 ตามลำดับ แบบทดสอบแต่ละฉบับประกอบด้วยข้อสอบแบบเลือกตอบ จำนวน 90 ข้อ แบบทดสอบรวมเป็นแบบทดสอบวัดความสามารถทั่วไปทางภาษาไทย ประกอบด้วยข้อสอบแบบเลือกตอบ จำนวน 30 ข้อ แบบทดสอบรวมนี้ใช้กับนักเรียนทุกชั้นเหมือนกัน

การเก็บรวบรวมข้อมูล

นักเรียนที่เป็นตัวอย่างการวิจัยแต่ละคน จะต้องตอบแบบทดสอบ 2 ฉบับ คือนักเรียนชั้น ม.1 ตอบแบบทดสอบชั้น ม.1 (ท 101) และแบบทดสอบรวม นักเรียนชั้น ม.2 ตอบแบบทดสอบชั้น ม.2 (ท 203) และแบบทดสอบรวม และนักเรียนชั้น ม.3 ตอบแบบทดสอบชั้น ม.3 (ท 305) และแบบทดสอบรวม ในการวิจัยนี้แบบทดสอบเทียบคะแนนและแบบทดสอบรวม จัดรวมไว้ด้วยกันเป็นฉบับเดียว ข้อมูลจากกลุ่มตัวอย่างได้จากการสอบปลายภาคเรียนของนักเรียนพร้อมกันทุกชั้น เมื่อวันที่ 4 ตุลาคม 2537 ระหว่างเวลา 13:00 – 15:30 น.เมื่อได้กระดาษคำตอบของนักเรียนที่เป็นกลุ่มตัวอย่างการวิจัยแล้ว ผู้วิจัยนำมามันทึกเพื่อวิเคราะห์ข้อมูลต่อไป

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูล วิเคราะห์ในเรื่องต่าง ๆ ต่อไปนี้

1. การวิเคราะห์แบบทดสอบ

วิเคราะห์แบบทดสอบ เพื่อตรวจสอบ และวิเคราะห์ค่าสถิติต่อไปนี้

1.1 ตรวจสอบคุณสมบัติการวัดเพียงมิติเดียวของแบบทดสอบ

1.2 วิเคราะห์ข้อสอบรายข้อ

1.3 วิเคราะห์ค่าสถิติพื้นฐานของแบบทดสอบฉบับต่าง ๆ

2. การสร้างตารางเทียบคะแนนแบบทดสอบ

การสร้างตารางเทียบคะแนนแบบทดสอบ ดำเนินการตามขั้นตอนต่อไปนี้

2.1 วิเคราะห์ข้อสอบ เพื่อประมาณค่าพารามิเตอร์ของข้อสอบแต่ละข้อ ในแบบทดสอบเทียบคะแนน และแบบทดสอบรวม

2.2 แปลงค่าพารามิเตอร์ของข้อสอบในแบบทดสอบเทียบคะแนน และค่าพารามิเตอร์ของความสามารถไปสู่สเกลของข้อสอบจากแบบทดสอบรวม

2.3 คำนวณหาคะแนนจริงของผู้สอบแต่ละคนที่ได้จากแบบทดสอบแต่ละฉบับ

2.4 คำนวณหาค่าเฉลี่ยของคะแนนสังเกตโดยคิดจากผู้สอบที่ได้คะแนนจริงเท่ากัน

2.5 คำนวณหาคะแนน ณ ตำแหน่งเปอร์เซ็นต์ไทล์ต่าง ๆ ของค่าเฉลี่ยคะแนนสังเกตในข้อ 2.4

2.6 สร้างตารางเทียบคะแนนสังเกตของแบบทดสอบฉบับต่าง ๆ ตามค่าเปอร์เซ็นต์ไทล์ในข้อ 2.5

3. การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย

การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย ศึกษาความก้าวหน้าของนักเรียนในกลุ่มรวมทั้งชั้น และศึกษาความก้าวหน้าของนักเรียนในกลุ่มย่อยโดยแบ่งกลุ่มย่อยตามระดับความสามารถของนักเรียนเป็นกลุ่มความสามารถสูง กลุ่มความสามารถปานกลางและกลุ่มความสามารถต่ำ

การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนดำเนินการตามขั้นตอนต่อไปนี้

3.1 เทียบคะแนนที่นักเรียน ชั้น ม.2 และ ม.3 ได้จากการตอบแบบทดสอบชั้น ม.2 และ ม.3 ตามลำดับ ไปยังสเกลของคะแนนแบบทดสอบชั้น ม.1

3.2 คำนวณค่าสถิติพื้นฐานของคะแนนที่เทียบให้อยู่ในสเกลเดียวกันแล้วของนักเรียนแต่ละชั้น

3.3 เปรียบเทียบความแตกต่างของคะแนนเฉลี่ยของนักเรียนชั้น ม.1 ม.2 และ ม.3 ซึ่งเทียบให้อยู่ในสเกลเดียวกันแล้ว

3.4 คำนวณหาผลต่างของคะแนนเฉลี่ยของนักเรียน 3 ระดับชั้น ค่าที่ได้นี้คือคะแนนเฉลี่ยความก้าวหน้าของผลการเรียนของนักเรียน

3.5 วิเคราะห์แนวโน้มการเปลี่ยนแปลงคะแนนเฉลี่ยของนักเรียนชั้น ม.1 ม.2 และ ม.3

สำหรับการศึกษาความก้าวหน้าของผลการเรียนของนักเรียนในแต่ละกลุ่มความสามารถ เมื่อแบ่งกลุ่มนักเรียนแต่ละชั้นตามระดับความสามารถแล้ว ดำเนินการศึกษาค้นคว้าความก้าวหน้าของผลการเรียนในแต่ละกลุ่มย่อยตามขั้นตอนเดียวกับ ข้อ 3.1 ถึง 3.5

4. การศึกษาประสิทธิภาพของวิธีการเทียบคะแนน

ประสิทธิภาพของวิธีการเทียบคะแนน พิจารณาจากผลรวมของความแตกต่างระหว่างคะแนนที่นักเรียนสอบได้จริง กับคะแนนที่เทียบจากตารางเทียบคะแนนที่สร้างขึ้นแล้วเปรียบเทียบดัชนีความแตกต่าง (C) ที่คำนวณได้ กับระดับการยอมรับความเพียงพอของการเทียบคะแนนตามเกณฑ์ของปีเตอร์สันและคนอื่น ๆ ดังนี้

$C \leq (.05 SD_x)^2$	หมายถึง น่าพอใจอย่างมาก
$(.05 SD_x)^2 < C \leq (.10 SD_x)^2$	หมายถึง น่าพอใจ
$(.10 SD_x)^2 < C \leq (.15 SD_x)^2$	หมายถึง ปานกลาง
$(.15 SD_x)^2 < C \leq (.20 SD_x)^2$	หมายถึง ไม่น่าพอใจ
$(.20 SD_x)^2 < C$	หมายถึง ไม่น่าพอใจอย่างยิ่ง

สรุปผลการวิจัย

ผลการวิจัยสรุปได้ดังนี้

1. ผลการวิเคราะห์แบบทดสอบ

1.1 ผลการตรวจสอบคุณสมบัติการวัดเพียงมิติเดียวของแบบทดสอบ

จากการวิเคราะห์องค์ประกอบของแบบทดสอบภาษาไทย 3 ฉบับ และแบบทดสอบความสามารถทั่วไปทางภาษาไทย ปรากฏว่า แบบทดสอบแต่ละฉบับสามารถวัดคุณลักษณะที่เด่นเพียง 1 คุณลักษณะ หรืออาจกล่าวได้ว่าแบบทดสอบแต่ละฉบับ มีคุณสมบัติการวัดเพียงมิติเดียว ซึ่งเป็นไปตามข้อตกลงเบื้องต้นของแบบจำลองโลจิสติก

1.2 ผลการวิเคราะห์ข้อสอบรายข้อ

จากการวิเคราะห์แบบทดสอบชั้น ม.1 ชั้น ม.2 ชั้น ม.3 และแบบทดสอบร่วมซึ่งประกอบด้วยข้อสอบแบบเลือกคำตอบจำนวน 90 90 90 และ 30 ข้อ ตามลำดับ ด้วยโปรแกรม BILOG ปรากฏว่าสามารถประมาณค่าพารามิเตอร์ของข้อสอบจากแบบทดสอบ 4 ฉบับดังกล่าวได้ 85 85 88 และ 29 ข้อ ตามลำดับ ผู้วิจัยได้ตัดข้อสอบข้อที่ประมาณค่าพารามิเตอร์ไม่ได้ออก และได้ตัดข้อสอบชั้น ม.3 ออกอีก 3 ข้อ ดังนั้นแบบทดสอบเทียบคะแนนที่ใช้ในการเทียบคะแนนจึงประกอบด้วยข้อสอบฉบับละ 85 ข้อ และแบบทดสอบร่วมประกอบด้วยข้อสอบ 29 ข้อ

1.2.1 ผลการวิเคราะห์ข้อสอบชั้น ม.1 ปรากฏว่าค่าประมาณพารามิเตอร์

ความยากมีค่าระหว่าง -2.099 ถึง 4.348 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .552 และ 1.274 ตามลำดับ ค่าประมาณพารามิเตอร์อำนาจจำแนกมีค่าระหว่าง .193 ถึง 1.505 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .758 และ .289 ตามลำดับ ค่าประมาณพารามิเตอร์การเดา มีค่าระหว่าง .150 ถึง .445 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .231 และ .056 ตามลำดับ และเมื่อคำนวณค่าฟังก์ชันสารสนเทศของแบบทดสอบ ปรากฏว่า แบบทดสอบชั้น ม.1 มีค่าฟังก์ชันสารสนเทศสูงสุด ที่ระดับความสามารถ 1.050

1.2.2 ผลการวิเคราะห์ข้อสอบชั้น ม.2 ปรากฏว่าค่าประมาณพารามิเตอร์ความยากมีค่าระหว่าง -2.784 ถึง 4.254 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .569 และ 1.494 ตามลำดับ ค่าประมาณพารามิเตอร์อำนาจจำแนก มีค่าระหว่าง .171 ถึง 1.942 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .710 และ .335 ตามลำดับ ค่าประมาณพารามิเตอร์การเดามีค่าระหว่าง .135 และ .500 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .252 และ .064 ตามลำดับ และเมื่อวิเคราะห์ค่าฟังก์ชันสารสนเทศของแบบทดสอบ ปรากฏว่าแบบทดสอบชั้นม.2 มีค่าฟังก์ชันสารสนเทศสูงสุด ที่ระดับความสามารถ 1.350

1.2.3 ผลการวิเคราะห์ข้อสอบชั้น ม.3 ปรากฏว่าค่าประมาณพารามิเตอร์ความยากมีค่าระหว่าง -2.913 ถึง 3.720 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .151 และ 1.567 ตามลำดับ ค่าประมาณพารามิเตอร์อำนาจจำแนกมีค่าระหว่าง .227 ถึง 1.285 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .575 และ .219 ตามลำดับ ค่าประมาณพารามิเตอร์การเดา มีค่าระหว่าง .141 ถึง .314 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .213 และ .049 ตามลำดับ และเมื่อวิเคราะห์ค่าฟังก์ชันสารสนเทศของแบบทดสอบ ปรากฏว่าแบบทดสอบชั้น ม.3 มีค่าฟังก์ชันสารสนเทศสูงสุด ที่ระดับความสามารถ 1.150

1.2.4 ผลการวิเคราะห์ข้อสอบในแบบทดสอบความสามารถทั่วไปทางภาษาไทย ปรากฏว่าค่าประมาณพารามิเตอร์ความยาก มีค่าระหว่าง -1.069 ถึง 3.474 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .936 และ .964 ตามลำดับ ค่าประมาณพารามิเตอร์อำนาจจำแนกมีค่าระหว่าง .342 ถึง 3.886 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน 1.169 และ 1.011 ตามลำดับ ค่าประมาณพารามิเตอร์การเดา มีค่าระหว่าง .102 ถึง .375 โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน .160 และ .091 ตามลำดับ และเมื่อวิเคราะห์ค่าฟังก์ชันสารสนเทศของแบบทดสอบ ปรากฏว่าแบบทดสอบความสามารถทั่วไปมีค่าฟังก์ชันสารสนเทศสูงสุด ที่ระดับความสามารถ 1.750

1.3 ผลการวิเคราะห์ค่าสถิติพื้นฐานของแบบทดสอบฉบับต่าง ๆ ก่อนเทียบคะแนน

1.3.1 จากการวิเคราะห์ผลการสอบแบบทดสอบชั้น ม.1 ปรากฏว่า

นักเรียนชั้น ม.1 ได้คะแนนระหว่าง 16-78 คะแนน จากคะแนนเต็ม 85 คะแนน โดยค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานเป็น 46.808 และ 12.697 คะแนน ตามลำดับ และความเชื่อมั่นของแบบทดสอบ เท่ากับ .895

1.3.2 จากการวิเคราะห์ผลการสอบแบบทดสอบชั้น ม.2 ปรากฏว่า นักเรียนชั้น ม.2 ได้คะแนนระหว่าง 17-78 คะแนน จากคะแนนเต็ม 85 คะแนน โดยค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานเป็น 48.362 และ 12.007 คะแนน ตามลำดับ และความเชื่อมั่นของแบบทดสอบ เท่ากับ .879

1.3.3 จากการวิเคราะห์ผลการสอบแบบทดสอบชั้น ม.3 ปรากฏว่า นักเรียนชั้น ม.3 ได้คะแนนระหว่าง 15 - 77 คะแนน จากคะแนนเต็ม 85 คะแนน โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน 50.344 และ 11.138 คะแนน ตามลำดับ และความเชื่อมั่นของแบบทดสอบ เท่ากับ .862

1.3.4 จากการวิเคราะห์ผลการสอบแบบทดสอบรวมของนักเรียนทั้ง 3 ชั้น ปรากฏว่านักเรียนชั้นมัธยมศึกษาตอนต้น 3 ชั้น ได้คะแนนระหว่าง 1 - 28 คะแนน จากคะแนนเต็ม 29 คะแนน โดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน 12.581 และ 5.116 คะแนน ตามลำดับ และความเชื่อมั่นของแบบทดสอบเท่ากับ .771

2. ผลการสร้างตารางเทียบคะแนนแบบทดสอบ

จากการใช้วิธีการเทียบคะแนนตามแนวตั้งด้วยแบบจำลองโลจิสติก สามพารามิเตอร์ ผู้วิจัยสร้างตารางเทียบคะแนนแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย ของชั้น ม.1 ม.2 และ ม.3 ดังแสดงในตาราง 19 และสร้างตารางเทียบคะแนนแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยแต่ละชั้น กับแบบทดสอบความสามารถทั่วไปทางภาษาไทย ดังตาราง 20-22 คะแนนสัมมูลจากแบบทดสอบต่างฉบับส่วนใหญ่อยู่ในรูปช่วงคะแนน

3. ผลการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย

3.1 ผลการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนในกลุ่มรวม

เมื่อแปลงคะแนนของนักเรียนชั้น ม.2 และ ม.3 ให้อยู่ในสเกลของ ม.1 แล้ว ปรากฏว่าค่าเฉลี่ยของคะแนนของนักเรียนชั้น ม.1 ม.2 และ ม.3 เป็น 46.601 และ 47.566 และ 51.192 คะแนน ตามลำดับ ซึ่งเมื่อวิเคราะห์ความแตกต่างของคะแนนดังกล่าวแล้ว ปรากฏว่าค่าเฉลี่ยของคะแนนของนักเรียน 3 ชั้น แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 และชั้น ม.1 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 แต่ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไม่แตกต่างกัน ซึ่งสรุปได้ว่าผลสัมฤทธิ์ทาง

การเรียนรู้ของนักเรียนชั้น ม.1 ไป ม.2 มีความก้าวหน้าตามปกติ แต่ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ไป ม.3 มีความก้าวหน้ามากกว่าปกติโดยเฉลี่ย 3.626 คะแนน และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.3 มีความก้าวหน้าขึ้นโดยเฉลี่ย 4.591 คะแนน การเปลี่ยนแปลงคะแนนเฉลี่ยของนักเรียน 3 ชั้นนี้ มีแนวโน้มไม่ชัดเจน

3.2 ผลการศึกษาความก้าวหน้า ของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มย่อย

เมื่อแบ่งนักเรียนแต่ละระดับชั้นเป็นกลุ่มย่อยตามระดับความสามารถแล้วศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียน 3 ชั้น ปรากฏว่า

3.2.1 ในกลุ่มนักเรียนที่มีความสามารถสูง คะแนนเฉลี่ยของนักเรียน 3 ชั้นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ไม่แตกต่างกัน ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไม่แตกต่างกัน แต่ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ซึ่งสรุปได้ว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.2 มีความก้าวหน้าตามปกติ และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ไป ม.3 มีความก้าวหน้าตามปกติ แต่ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.3 มีความก้าวหน้ามากกว่าปกติโดยเฉลี่ย 1.421 คะแนน การเปลี่ยนแปลงคะแนนเฉลี่ยของนักเรียน 3 ชั้นนี้ มีแนวโน้มเป็นเส้นตรง

3.2.2 ในกลุ่มนักเรียนที่มีความสามารถปานกลาง คะแนนเฉลี่ยของนักเรียน 3 ชั้น แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.1 โดยผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ซึ่งสรุปได้ว่าผลสัมฤทธิ์ทางการเรียนของนักเรียน 3 ชั้นมีความก้าวหน้าขึ้นตามลำดับ โดยผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.2 มีความก้าวหน้ามากกว่าปกติโดยเฉลี่ย 1.199 คะแนน ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ไป ม.3 มีความก้าวหน้ามากกว่าปกติโดยเฉลี่ย 4.339 คะแนน และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.3 มีความก้าวหน้ามากกว่าปกติโดยเฉลี่ย 5.538 คะแนน การเปลี่ยนแปลงคะแนนเฉลี่ยของนักเรียน 3 ชั้นนี้ มีแนวโน้มไม่ชัดเจน

3.2.3 ในกลุ่มนักเรียนที่มีความสามารถต่ำ คะแนนเฉลี่ยของนักเรียน 3 ชั้นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 และผลสัมฤทธิ์ทางการเรียนของนักเรียน

ชั้น ม.3 สูงกว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 แต่ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 และ ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไม่แตกต่างกัน ซึ่งสรุปได้ว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.2 มีความก้าวหน้าตามปกติ แต่ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ไป ม.3 มีความก้าวหน้ามากกว่าปกติโดยเฉลี่ย 5.298 คะแนน และผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.3 มีความก้าวหน้ามากกว่าปกติโดยเฉลี่ย 6.040 คะแนน การเปลี่ยนแปลงคะแนนเฉลี่ยของนักเรียน 3 ชั้นนี้ มีแนวโน้มไม่เปลี่ยนแปลง

4. ผลการศึกษาประสิทธิภาพของวิธีการเทียบคะแนน

จากการคำนวณดัชนีความแตกต่าง (C) ซึ่งเป็นค่าที่แสดงถึงผลรวมความคลาดเคลื่อนของการเทียบคะแนนได้ค่าดัชนีเท่ากับ .464 เมื่อเปรียบเทียบกับดัชนีความแตกต่างดังกล่าวกับเกณฑ์การยอมรับความเพียงพอตามเกณฑ์ของปีเตอร์สันและคนอื่น ๆ สรุปได้ว่าวิธีการเทียบคะแนนที่ใช้ในการวิจัยนี้ให้ผลการเทียบคะแนนมีประสิทธิภาพในระดับปานกลาง

อภิปรายผลการวิจัย

การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 โดยการเทียบคะแนนตามแนวดิ่ง ตามทฤษฎีการตอบสนองข้อสอบ แบบจำลองโลจิสติกสามพารามิเตอร์ มีประเด็นสำคัญที่นำมาอภิปรายดังนี้

1. ลักษณะของการศึกษาความก้าวหน้าในการวิจัยนี้

การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยในการวิจัยนี้ ศึกษาเฉพาะวิชาภาษาไทยที่เป็นวิชาบังคับแกน คือวิชา ท 101 (ชั้น ม.1) ท 203 (ชั้น ม.2) และ ท 305 (ชั้น ม.3) และผลการสอบที่นำมาศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนเป็นผลการสอบที่ได้จากคะแนนการตอบข้อสอบแบบเลือกตอบซึ่งวัดบางจุดประสงค์เท่านั้น สำหรับผลการวัดในบางจุดประสงค์ เช่น จุดประสงค์ที่วัดทักษะการฟัง หรือจุดประสงค์ที่วัดทักษะการพูด ไม่รวมอยู่ในการสอบปลายภาค ดังนั้น ผลสรุปจากการศึกษาวิจัย จึงเป็นผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยบางส่วน มีข้อทั้งหมดของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3

สำหรับรูปแบบการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 ในการวิจัยนี้ใช้การศึกษาภาคตัดขวาง (Cross Section Study) โดยผลสัมฤทธิ์ทางการเรียนของนักเรียน 3 ระดับชั้น ได้จากนักเรียนคนละกลุ่มที่เรียนอยู่ในภาคการศึกษาและปีการศึกษาเดียวกัน โดยตั้งข้อสมมติว่านักเรียน 3 ระดับชั้นเป็นนักเรียนใน รุ่น (Cohort) เดียวกัน ซึ่งแตกต่างไปจากรูปแบบการศึกษาระยะยาว ซึ่งติดตามนักเรียน

รุ่นใดรุ่นหนึ่งไปทุกชั้น ตั้งแต่ชั้น ม.1 ไปจนจบชั้น ม.3

2. ผลการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียน 3 ชั้นในกลุ่มรวม

ผลการศึกษาพบว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1ไป ม.2 ก้าวหน้าตามปกติแต่ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ไป ม.3 ก้าวหน้ามากกว่าปกติซึ่งไม่สอดคล้องกับสมมติฐานการวิจัยข้อ 1 ที่คาดว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนมีความก้าวหน้าขึ้นตามลำดับที่เป็นเช่นนี้น่าจะเนื่องมาจาก นักเรียนชั้น ม.1 เป็นนักเรียนที่เปลี่ยนจากการเป็นนักเรียนระดับประถมศึกษาเป็นนักเรียนระดับมัธยมศึกษาเป็นปีแรกจึงต้องปรับตัวในเรื่องต่าง ๆ เช่น วิธีการเรียน สถานที่เรียน กลุ่มเพื่อน นอกจากนี้ พัฒนาการทางภาษา มีลักษณะค่อยเป็นค่อยไปและมีการขยายวงความรู้ไปรอบด้าน ดังนั้นพัฒนาการทางภาษาจะเพิ่มขึ้นตามช่วงเวลาที่ผ่านมา ตามอายุที่เพิ่มขึ้น และตามประสบการณ์ที่สูงขึ้นของนักเรียน นอกจากนี้ นักเรียนชั้น ม.3 เป็นนักเรียนปีสุดท้ายของหลักสูตรชั้นมัธยมศึกษาตอนต้น ซึ่งในปีการศึกษาต่อไป ต้องมีการเรียนหลักสูตรที่ต่างไป จึงต้องใช้ความรู้ในการสอบแข่งขัน หรือสอบคัดเลือกเข้าเรียนต่อ จึงเป็นเหตุให้มีการฝึกฝน และเตรียมความรู้ในวิชาต่าง ๆ ให้พร้อม ดังนั้นผลสัมฤทธิ์ทางการเรียนจากชั้น ม.1 ไป ม.2 จึงก้าวหน้าตามปกติ แต่ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ไป ม.3 ก้าวหน้ามากกว่าปกติ

3. ผลการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนในกลุ่มย่อย

ผลการวิจัยพบว่า ความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนกลุ่มที่มีความสามารถปานกลางมีความก้าวหน้าขึ้นตามลำดับ สอดคล้องกับสมมติฐานการวิจัยข้อ 3 ส่วนผลการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนในกลุ่มที่มีความสามารถสูงและในกลุ่มที่มีความสามารถต่ำไม่สอดคล้องกับสมมติฐานการวิจัย ดังนี้

3.1 ผลการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนกลุ่มที่มีความสามารถสูงพบว่า ผลสัมฤทธิ์ทางการเรียนของนักเรียนแต่ละชั้น มีความก้าวหน้าตามปกติ ซึ่งไม่สอดคล้องกับสมมติฐานการวิจัยข้อ 2 ที่คาดว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนมีความก้าวหน้าขึ้นตามลำดับ ที่เป็นเช่นนี้น่าจะเนื่องมาจากนักเรียนที่มีความสามารถสูงมีการพัฒนาความรู้ ความสามารถในระดับสูงแล้วในทุกชั้น ดังนั้นผลสัมฤทธิ์ทางการเรียนจึงก้าวหน้าตามปกติในระดับชั้นที่ต่อเนื่องกัน นอกจากนี้ เมื่อพิจารณาแนวโน้มการเปลี่ยนแปลงค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียนของนักเรียนทั้ง 3 ระดับชั้น พบว่ามีแนวโน้มเป็นเส้นตรง ซึ่งสนับสนุนว่านักเรียนกลุ่มที่มีความสามารถสูง มีการพัฒนาอย่างสม่ำเสมอในทุกระดับชั้น ดังนั้นในระดับชั้นที่ต่อเนื่องกัน เช่น ม.1 ไป ม.2 หรือ ม.2 ไป ม.3 ความก้าวหน้าของผลสัมฤทธิ์ทาง

การเรียนรู้จึงเป็นความก้าวหน้าตามปกติ ส่วนในระดับชั้นที่อยู่ห่างกัน คือ ม.1 ไป ม.3 ผลสัมฤทธิ์ทางการเรียนของนักเรียนก้าวหน้ามากกว่าปกติ

3.2 ผลการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนของนักเรียนกลุ่มที่มีความสามารถต่ำ พบว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.2 ก้าวหน้าตามปกติ แต่ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ไป ม.3 ก้าวหน้ามากกว่าปกติ ซึ่งไม่สอดคล้องกับสมมติฐานการวิจัย ข้อ 4 ที่คาดว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนมีความก้าวหน้าขึ้นตามลำดับ ที่เป็นเช่นนี้น่าจะเนื่องมาจากนักเรียนชั้น ม.1 เป็นนักเรียนชั้นมัธยมศึกษาเป็นปีแรกของหลักสูตร ซึ่งต้องมีการปรับตัวในเรื่องต่าง ๆ ดังอภิปรายในข้อ 2 และเป็นนักเรียนที่มีพื้นฐานความรู้น้อยกว่านักเรียนมัธยมศึกษาชั้นอื่น ๆ เมื่อเรียนในชั้นที่สูงขึ้นนักเรียนมีโอกาสนำตนเองให้ก้าวหน้ามากขึ้น ดังนั้นผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ไป ม.3 จึงก้าวหน้ามากกว่าปกติ นอกจากนี้เมื่อพิจารณาแนวโน้มการเปลี่ยนแปลงค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียนของนักเรียน 3 ระดับชั้น พบว่าแนวโน้มมีลักษณะไม่ใช่เส้นตรง โดยคะแนนเฉลี่ยของผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 และ ม.2 ไม่แตกต่างกัน แต่คะแนนเฉลี่ยของผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่า ม.2 มาก

4. ผลการตรวจสอบคุณสมบัติการวัดเพียงมิติเดียวของแบบทดสอบ

ผลการตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว พบว่า แบบทดสอบทั้ง 4 ฉบับ มีแนวโน้มที่จะวัดคุณลักษณะเดียว การตรวจสอบคุณสมบัติการวัดเพียงมิติเดียวนี้ใช้วิธีวิเคราะห์องค์ประกอบโดยศึกษาคุณลักษณะที่เด่นเพียงด้านเดียวของแบบทดสอบ การดำเนินการนี้ไม่ได้ครอบคลุมไปถึงรายละเอียดของการวัดองค์ประกอบของแบบทดสอบแต่ละฉบับว่าวัดองค์ประกอบใด แต่อย่างไรก็ตาม แบบทดสอบทั้ง 4 ฉบับสร้างจากจุดประสงค์ของการวัดความสามารถทางภาษาไทย และตามโครงสร้างของแบบทดสอบ (Blue Print) เดียวกัน จากหลักสูตรการเรียนการสอนวิชาภาษาไทยชั้นมัธยมศึกษาตอนต้นหลักสูตรเดียวกัน จึงทำให้ผู้วิจัยคาดว่า อย่างน้อยแบบทดสอบทั้ง 4 ฉบับก็วัดความสามารถทางภาษาไทยอย่างเดียวกัน การตรวจสอบคุณสมบัติการวัดเพียงมิติเดียวของแบบทดสอบก็เพื่อให้มั่นใจได้ว่า แบบทดสอบที่นำมาใช้ทั้ง 4 ฉบับ มีคุณสมบัติพื้นฐานของการวัดสอดคล้องกับแบบจำลองโลจิสติกที่นำมาใช้ในการวิจัยนี้

นอกจากนี้ผลการตรวจสอบคุณสมบัติการวัดเพียงมิติเดียวของแบบทดสอบความสามารถทั่วไปทางภาษาไทย พบว่า แบบทดสอบความสามารถทั่วไปทางภาษาไทยมีแนวโน้มที่วัดคุณลักษณะเดียว แต่แนวโน้มดังกล่าวไม่เด่นชัดเหมือนผลการตรวจสอบแบบทดสอบฉบับอื่น ๆ ทั้งนี้อาจเนื่องจากการวิจัยนี้ แบบทดสอบความสามารถทั่วไปทางภาษาไทยซึ่งใช้เป็นแบบทดสอบร่วมภายนอกวัดพฤติกรรมต่าง ๆ ตามโครงสร้างของแบบทดสอบโดยครอบคลุมพฤติกรรมการวัดของทั้ง 3 ชั้น ซึ่งแตกต่างไปจากการวัดคุณลักษณะของแบบทดสอบแต่ละชั้นที่เฉพาะเจาะจงมากกว่า ทำให้คุณ-

ลักษณะการวัดเพียงมิติเดียวของแบบทดสอบความสามารถทั่วไปทางภาษาไทยไม่เด่นชัดเท่าผลการตรวจสอบที่พบในแบบทดสอบฉบับอื่น ๆ

5. ประสิทธิภาพของวิธีการเทียบคะแนน

จากการประเมินความเพียงพอของวิธีการเทียบคะแนน ผลการประเมินสรุปว่าวิธีการเทียบคะแนน ตามทฤษฎีการตอบสนองข้อสอบ โดยใช้แบบจำลองโลจิสติก สามพารามิเตอร์ ที่นำมาใช้กับสถานการณ์การวิจัยนี้มีประสิทธิภาพในระดับปานกลาง แม้ว่าจากผลการวิจัยหลายเรื่องพบว่า สำหรับการเทียบคะแนนตามแนวตั้ง วิธีเทียบคะแนนโดยใช้แบบจำลองโลจิสติก สามพารามิเตอร์ ให้ผลดีกว่าเมื่อเปรียบเทียบกับวิธีอื่นในหลายสถานการณ์ (Skaggs and Lissitz, 1986) ทั้งนี้อาจเนื่องมาจากการวิจัยนี้ใช้การเทียบคะแนนจากแบบทดสอบแต่ละชั้น ไปยังแบบทดสอบความสามารถทั่วไปทางภาษาไทยซึ่งเป็นแบบทดสอบร่วมชนิดภายนอกที่ใช้ในการปรับค่าพารามิเตอร์ของแบบทดสอบต่างระดับชั้น และโดยที่แบบทดสอบร่วมมีขนาดสั้นกว่าแบบทดสอบเทียบคะแนนของชั้น ม.1 ม.2 และ ม.3 มาก ดังนั้นในขั้นตอนการเทียบคะแนน สเกลคะแนนของแบบทดสอบแต่ละชั้นจะถูกรวบรวมให้แคบลง คะแนนจากแบบทดสอบเทียบคะแนนหลายค่า จึงตรงกับคะแนนจากแบบทดสอบร่วมค่าเดียวกัน ทำให้ความคลาดเคลื่อนในการเทียบคะแนนเพิ่มสูงขึ้น

ข้อเสนอแนะ

จากการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย ของนักเรียนชั้น ม.1 ม.2 และ ม.3 โดยใช้การเทียบคะแนนตามแนวตั้ง ตามทฤษฎีการตอบสนองข้อสอบ แบบจำลองโลจิสติก สามพารามิเตอร์ ผู้วิจัยมีข้อเสนอแนะในการนำผลการวิจัยไปใช้ และในการวิจัยเกี่ยวกับการเทียบคะแนนตามแนวตั้ง ดังนี้

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1.1 จากผลการวิจัยพบว่า ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้น ม.1 ม.2 และ ม.3 ในกลุ่มรวมและในกลุ่มที่มีความสามารถต่ำมีความก้าวหน้า โดยผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 ไป ม.2 ก้าวหน้าตามปกติ แต่ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.2 ไปชั้น ม.3 ก้าวหน้ามากกว่าปกติ ดังนั้นในการพัฒนานักเรียนเพื่อให้ผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยก้าวหน้าทุกระดับชั้น จึงควรหาวิธีการจัดการเรียนการสอนที่จะส่งเสริมนักเรียนชั้น ม.2 โดยเฉพาะนักเรียนที่มีความสามารถต่ำ ให้มีผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยก้าวหน้ามากยิ่งขึ้น

1.2 จากผลการวิจัยพบว่า นักเรียนที่มีความสามารถปานกลางมีผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยก้าวหน้ามากกว่าปกติทุกระดับชั้น ดังนั้นจึงควรส่งเสริมให้นักเรียนกลุ่ม

นี้มีผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยก้าวหน้าเช่นนี้ต่อไป

2. ข้อเสนอแนะในการวิจัยเกี่ยวกับการเทียบคะแนนตามแนวดิ่ง

2.1 การนำวิธีการเทียบคะแนนตามแนวดิ่งไปใช้ในการเปรียบเทียบคะแนนจากแบบทดสอบฉบับต่าง ๆ ควรคำนึงถึงสิ่งต่าง ๆ ต่อไปนี้

2.1.1 เนื้อหาวิชาที่ศึกษา ควรเป็นเนื้อหาที่มีโครงสร้างการวัดเดียวกัน หรือมีลักษณะเนื้อหาวิชาที่ต่อเนื่องกัน

2.1.2 คุณภาพของแบบทดสอบ ต้องมีคุณภาพสูงเพื่อให้ผลการวัดมีความถูกต้องแม่นยำ

2.1.3 แบบทดสอบที่นำมาเทียบคะแนนควรมีความยาวไม่แตกต่างกันมากนัก

2.2 ควรมีการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนในลักษณะนี้ โดยศึกษากับนักเรียนระดับชั้นอื่น

2.3 ควรมีการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนในลักษณะนี้ที่จังหวัดอื่น โดยเฉพาะจังหวัดที่มีการใช้ภาษาพื้นบ้าน ว่าจะให้ผลแตกต่างจากการวิจัยนี้หรือไม่อย่างไร

2.4 ควรมีการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนในลักษณะนี้ โดยศึกษากับเนื้อหาวิชาอื่น ทั้งเนื้อหาที่มีลักษณะการวัดความสามารถทั่วไป เช่น ภาษาอังกฤษ และเนื้อหาวิชาที่วัดความสามารถเฉพาะ เช่น คณิตศาสตร์ เป็นต้น

2.5 ควรมีการศึกษาเกี่ยวกับผลของความยาวของแบบทดสอบร่วมที่มีต่อผลการเทียบคะแนน เมื่อนำมาใช้ในการวิจัยในลักษณะนี้

2.6 ควรมีการศึกษาเกี่ยวกับความยาวของแบบทดสอบเทียบคะแนนที่ทำให้ผลการเทียบคะแนนมีประสิทธิภาพสูงสุด

2.7 ควรมีการศึกษาผลของแบบทดสอบร่วมลักษณะต่างๆที่มีต่อผลการเทียบคะแนน เช่น แบบทดสอบร่วมที่ประกอบด้วยข้อสอบที่มีลักษณะเป็นเอกพันธ์ แบบทดสอบร่วมที่มีลักษณะเป็นวิวิธพันธ์ เป็นต้น

บรรณานุกรม

บรรณานุกรม

- คณะกรรมการการประถมศึกษาแห่งชาติ, สำนักงาน. รายงานการประเมินความก้าวหน้า : คุณภาพนักเรียนชั้นประถมปีที่ 6 ระดับประเทศ ปีการศึกษา 2530. กรุงเทพฯ : สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ, 2531.
- คณะกรรมการการประถมศึกษาแห่งชาติ, สำนักงาน. แผนการศึกษาแห่งชาติพุทธศักราช 2535. กรุงเทพฯ : สำนักนายกรัฐมนตรี, 2535.
- ชูชีพ พงษ์สมบูรณ์. การเปรียบเทียบการมาตรฐานระหว่างรูปแบบที่ใช้ผู้สอบร่วมกับรูปแบบที่ใช้แบบสอบร่วม. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2530. ถ่ายเอกสาร.
- ชูศักดิ์ ชัมภลิจิต. "การปรับเทียบคะแนนทดสอบ," ใน บทความสรรทางวิชาการทดสอบ. นนทบุรี : ศูนย์พัฒนศึกษาแห่งชาติของประเทศไทย ทบวงมหาวิทยาลัย, 2527.
- ปราณี บุญชุ่ม. "การศึกษาวิเคราะห์องค์ประกอบและหลักการพัฒนาหลักสูตรภาษาไทยระดับมัธยม," ใน ประมวลสาระชุดวิชาสารัตถะและวิทยวิธีทางวิชาภาษาไทย. นนทบุรี : บัณฑิตวิทยาลัย สาขาวิชาศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช, 2537.
- ผจงจิต อินทสุวรรณ. การวิเคราะห์แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนด้วยโมเดลโลจิสติก. กรุงเทพฯ : สถาบันวิจัยพฤติกรรมศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2534.
- พรพิมล นาคเวช. การศึกษาคุณภาพของการเทียบคะแนนในแนวตั้งโดยใช้ทฤษฎีการตอบสนอง-ข้อสอบ. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2537. ถ่ายเอกสาร.
- ไพศาล หวังพานิช. เอกสารประกอบการสอนวิธีการวิจัยทางพฤติกรรมศาสตร์และสังคมศาสตร์. กรุงเทพฯ : สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2523.
- ภาวิณี ศรีสุขวัฒนานนท์. การเปรียบเทียบผลจากการใช้รูปแบบการเทียบมาตรฐานที่ต่างกันเมื่อแบบสอบร่วมมีความยาวต่างกัน. วิทยานิพนธ์ ค.ค. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2530. ถ่ายเอกสาร.
- เรวดี อินทสระ. การเปรียบเทียบความคลาดเคลื่อนของการเทียบมาตรฐานระหว่างรูปแบบอิงทฤษฎีการตอบข้อสอบกับรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ. ปรินญานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2530. อัดสำเนา.

- วัลลภ กันทรพ์. "การประเมินผลที่ได้รับของหลักสูตร," วารสารการวิจัยการศึกษา. 2(15) : 60 - 72 ; เมษายน - มิถุนายน 2528.
- วิชาการ,กรม. คู่มือหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533). กรุงเทพฯ : กระทรวงศึกษาธิการ, 2534.
- ศิริชัย กาญจนวาสี. "ทฤษฎีการทดสอบ," ใน เอกสารการสอนชุดวิชาการพัฒนาแบบทดสอบ วัดผลสัมฤทธิ์ทางการเรียน. นนทบุรี : สาขาวิชาศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช, 2535.
- สงบ ลักษณะ. "การเทียบระดับคะแนนระหว่างแบบทดสอบ (EQUATING)," วารสารการวัดผลการศึกษา. 4(2) : 21 - 32 ; กันยายน - ธันวาคม 2525.
- สุจินดา ผ่องอักษร. การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนในกลุ่มทักษะของนักเรียนชั้นประถมศึกษาปีที่ 6 ที่เรียนจบตามหลักสูตรประถมศึกษา พุทธศักราช 2521 ในช่วงระยะเวลา 3 ปีการศึกษา(ปีการศึกษา 2529 ถึง 2531) โดยใช้การเทียบคะแนนรูปแบบบราซส์. ปรินซ์ตันนิพนธ์ กศ.ด. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2533. ถ่ายเอกสาร.
- สุรัตน์ ขวัญบุญจันทร์. การสร้างแบบทดสอบและตารางเทียบคะแนนในแนวนอน วิชาวิทยาศาสตร์ ที่วิเคราะห์ด้วยบราซส์โมเดล. ปรินซ์ตันนิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2531. ถ่ายเอกสาร.
- อนันต์ ศรีโสภณ. ทฤษฎีการวัดและการทดสอบ. กรุงเทพฯ : ไทยวัฒนาพานิช, 2525.
- อาภรณ์ กาญจนกิจโสภณ. การสร้างแบบทดสอบและตารางเทียบคะแนนตามตั้งในวิชาคณิตศาสตร์. ปรินซ์ตันนิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2531. อัดสำเนา.
- American Psychological Association. Standards for Educational and Psychological Testing. Washington, DC : American Psychological Association, 1985.
- Angoff, W.H. Scales, norms , and equivalent scores. Princeton, NJ : Educational Testing Service, 1984.
- Ayerve, R.I. "The Effectiveness of the Equipercntile Method and IRT Three - Parameter Model on Vertical Equating Under Varying Conditions of Sample Size, Test Length, and Anchor Test Length : A Simulation Study," Dissertation Abtracts International. 53 : 1841 - A ; December, 1992.

- Braun, H.I. and P.W. Holland. " A Mathematic Analysis of Some ETS Equating Procedure," in P.W. Holland and D.B. Rubin (Ed.), Test Equating. p. 9 - 49. New York : Academic Press, 1982.
- Brennan, R.L. "Introduction to Problems, Perspective, and Practical Issues in Equating," Applied Psychological Measurement. 11(3) : 221 - 224 ; 1987.
- Cook, L.L and D.R. Eignor. "Using Item Response Theory in Test Score Equating," International Journal of Educational Research. 13(2) : 161 - 173 ; 1989.
- Flanagan, T.C. "Unit, Scores and Norms," in E.F. Lindquist (Ed.) Educational Measurement. p.695-763. Washington D.C. : American Council on Education, 1951.
- Gulliksen, H. Theory of Mental Tests. New York:John Wiley and Sons, Inc., 1950.
- Hambleton, R.K. "Principles and Selected Applications of Item Response Theory" in R.L. Linn (Ed.) Educational Measurement. p. 9 - 49. 3 rd. New York : Macmillan, 1989.
- Hambleton, R.K. and H. Swaminathan. Item Response Theory. Kluwer. Nijhoff, Boston, 1985.
- Harris, D.J. and H.D. Hoover. "An Application of the Three - Parameter IRT Model to Vertical Equating," Applied Psychological Measurement. 11(2) : 151 - 159 ; June, 1987.
- Holland, P.W. and Donald B. Rubin. Test Equating. New Jersey: Educational Testing Service, Princeton, Academic Press, 1982.
- Johanson, G.A. " A Study of Item Response Theory Equating with an Anchor Test Design," Dissertation Abstracts International. 49 : 488 - A ; September, 1988.
- Kolen, M.J. "Comparison of Traditional and Item Response Theory Methods for Equating Tests,"Journal of Edncational Measurement. 18(1) : 1 - 11 Spring, 1981.

- Lord, F.M. Application of Item Response Theory to Practical Testing Problems. Hillsdale, N.J. Erbaum, 1980.
- Mislevy, R.J. and Bock, R.K. BILOG 3 Item Analysis and Test Scoring with Binary Logistic Models. Mooresville IN:Scientific Software, Inc. 1990.
- Patience, W.M. "A Comparison of Classical and Item Response Theory Methods of Vertical Equating Tests," Dissertation Abstracts International. 51 : 486 - A ; August, 1990.
- Pertersen, N.S., M.J. Kolen and H.D. Hoover. "Scaling, Norming, and Equating," in R.L. Linn (Ed.) Educational Measurement. p.221 - 262. 3rd. New York : Macmillan, 1989.
- Pertersen, N.S., G.L. Marco. and E.E. Stewart. "A Test of the Adequacy of Linear Score Equating Models," in P.W. Holland and D.B. Rubin (Eds.) Test Equating. p. 71 - 135. New York: Academic Press. 1982.
- Phillips, S.E. "Comparison of Equipercentile and Item Response Theory Equating When the Scaling Test Method Is Applied to a Multilevel Achievement Battery," Applied Psychological Measurement. 7(3) : 267 - 281 ; Summer, 1983.
- Skaggs, G. and R.W. Lissitz. "IRT Test Equating : Relevant Issues and a Review of Recent Research," Review of Educational Research. 56(4) : 495 - 529; Winter, 1986.
- Warm, T.A. A Primer of Item Response Theory. Oklahoma: U.S.A. Coast Guard Institute, 1978.

ภาคผนวก

ภาคผนวก ก

ตาราง 37 ค่าฟังก์ชันสารสนเทศของแบบทดสอบและค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบ
ชั้น ม.1

ลำดับที่	θ	$\sum P(\theta)$	$\sum P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
1	-4.0000	19.8460	0.2335	0.0050	0.0001
2	-3.9500	19.8625	0.2337	0.0059	0.0001
3	-3.9000	19.8804	0.2339	0.0069	0.0001
4	-3.8500	19.8998	0.2341	0.0081	0.0001
5	-3.8000	19.9209	0.2344	0.0095	0.0001
6	-3.7500	19.9437	0.2346	0.0112	0.0001
7	-3.7000	19.9685	0.2349	0.0131	0.0002
8	-3.6500	19.9953	0.2352	0.0154	0.0002
9	-3.6000	20.0243	0.2356	0.0180	0.0002
10	-3.5500	20.0558	0.2360	0.0211	0.0002
11	-3.5000	20.0898	0.2364	0.0246	0.0003
12	-3.4500	20.1266	0.2368	0.0287	0.0003
13	-3.4000	20.1663	0.2373	0.0335	0.0004
14	-3.3500	20.2093	0.2378	0.0390	0.0005
15	-3.3000	20.2557	0.2383	0.0453	0.0005
16	-3.2500	20.3057	0.2389	0.0526	0.0006
17	-3.2000	20.3596	0.2395	0.0609	0.0007
18	-3.1500	20.4178	0.2402	0.0705	0.0008
19	-3.1000	20.4803	0.2409	0.0814	0.0010
20	-3.0500	20.5476	0.2417	0.0939	0.0011
21	-3.0000	20.6200	0.2426	0.1080	0.0013
22	-2.9500	20.6976	0.2435	0.1240	0.0015
23	-2.9000	20.7809	0.2445	0.1421	0.0017
24	-2.8500	20.8702	0.2455	0.1625	0.0019
25	-2.8000	20.9658	0.2467	0.1853	0.0022
26	-2.7500	21.0680	0.2479	0.2108	0.0025

ตาราง 37 (ต่อ)

ลำดับที่	θ	$\sum P(\theta)$	$\sum P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
27	-2.7000	21.1772	0.2491	0.2392	0.0028
28	-2.6500	21.2937	0.2405	0.2708	0.0032
29	-2.6000	21.4179	0.2520	0.3057	0.0036
30	-2.5500	21.5501	0.2535	0.3442	0.0040
31	-2.5000	21.6905	0.2552	0.3864	0.0045
32	-2.4500	21.8397	0.2569	0.4325	0.0051
33	-2.4000	21.9978	0.2588	0.4828	0.0057
34	-2.3500	22.1652	0.2608	0.5373	0.0063
35	-2.3000	22.3422	0.2628	0.5962	0.0070
36	-2.2500	22.5291	0.2650	0.6596	0.0078
37	-2.2000	22.7262	0.2674	0.7276	0.0086
38	-2.1500	22.9337	0.2698	0.8002	0.0094
39	-2.1000	23.1520	0.2724	0.8775	0.0103
40	-2.0500	23.3812	0.2751	0.9595	0.0113
41	-2.0000	23.6215	0.2779	1.0461	0.0123
42	-1.9500	23.8733	0.2809	1.1372	0.0134
43	-1.9000	24.1367	0.2840	1.2329	0.0145
44	-1.8500	24.4118	0.2872	1.3330	0.0157
45	-1.8000	24.6989	0.2906	1.4374	0.0169
46	-1.7500	24.9981	0.2941	1.5460	0.0182
47	-1.7000	25.3096	0.2978	1.6585	0.0195
48	-1.6500	25.6335	0.3016	1.7749	0.0209
49	-1.6000	25.9700	0.3055	1.8950	0.0223
50	-1.5500	26.3192	0.3096	2.0186	0.0237
51	-1.5000	26.6812	0.3139	2.1456	0.0252
52	-1.4500	27.0562	0.3243	2.2757	0.0268
53	-1.4000	27.4443	0.3229	2.4090	0.0283

ตาราง 37 (ต่อ)

ลำดับที่	θ	$\Sigma P(\theta)$	$\Sigma P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
54	-1.3500	27.8456	0.3276	2.5451	0.0299
55	-1.3000	28.2602	0.3325	2.6842	0.0316
56	-1.2500	28.6883	0.3375	2.8259	0.0332
57	-1.2000	29.1301	0.3427	2.9704	0.0349
58	-1.1500	29.5856	0.3481	3.1176	0.0367
59	-1.1000	30.0550	0.3539	3.2675	0.0384
60	-1.0500	30.5384	0.3593	3.4201	0.0402
61	-1.0000	31.0359	0.3651	3.5754	0.0421
62	-0.9500	31.5478	0.3712	3.7334	0.0439
63	-0.9000	32.0741	0.3773	3.8943	0.0458
64	-0.8500	32.6149	0.3837	4.0581	0.0477
65	-0.8000	33.1704	0.3902	4.2249	0.0497
66	-0.7500	33.7406	0.3969	4.3946	0.0517
67	-0.7000	34.3257	0.4038	4.5673	0.0537
68	-0.6500	34.9256	0.4109	4.7431	0.0558
69	-0.6000	35.5405	0.4181	4.9220	0.0579
70	-0.5500	36.1703	0.4255	5.1040	0.0600
71	-0.5000	36.8149	0.4331	5.2892	0.0622
72	-0.4500	37.4744	0.4409	5.4776	0.0644
73	-0.4000	38.1486	0.4488	5.6693	0.0667
74	-0.3500	38.8373	0.4569	5.8642	0.0690
75	-0.3000	39.5403	0.4652	6.0624	0.0713
76	-0.2500	40.2574	0.4736	6.2640	0.0737
77	-0.2000	40.9882	0.4822	6.4689	0.0761
78	-0.1500	41.7323	0.4907	6.6771	0.0786
79	-0.1000	42.4893	0.4999	6.8886	0.0810
80	-0.0500	43.2586	0.5089	7.1033	0.0836

ตาราง 37 (ต่อ)

ลำดับที่	θ	$\sum P(\theta)$	$\sum P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
81	0.0000	44.0396	0.5181	7.3208	0.0861
82	0.0500	44.8318	0.5274	7.5409	0.0887
83	0.1000	45.6342	0.5369	7.7632	0.0913
84	0.1500	46.4461	0.5464	7.9871	0.0940
85	0.2000	47.2665	0.5551	8.2119	0.0966
86	0.2500	48.0946	0.5658	8.4366	0.0993
87	0.3000	48.9293	0.5756	8.6604	0.1019
88	0.3500	49.7694	0.5856	8.8818	0.1045
89	0.4000	50.6137	*0.5955	9.0995	0.1071
90	0.4500	51.4612	0.6054	9.3119	0.1096
91	0.5000	52.3104	0.6154	9.5174	0.1120
92	0.5500	53.1602	0.6254	9.7140	0.1143
93	0.6000	54.0091	0.6354	9.8999	0.1165
94	0.6500	54.8559	0.6422	10.0731	0.1185
95	0.7000	55.6991	0.6553	10.2315	0.1204
96	0.7500	56.5375	0.6651	10.3732	0.1220
97	0.8000	57.3698	0.6749	10.4963	0.1235
98	0.8500	58.1946	0.6846	10.5990	0.1247
99	0.9000	59.0108	0.6942	10.6798	0.1256
100	0.9500	59.8171	0.7037	10.7374	0.1263
101	1.0000	60.6125	0.7131	10.7708	0.1267
102	1.0500	61.3959	0.7223	10.7793	0.1268
103	1.1000	62.1665	0.7314	10.7626	0.1266
104	1.1500	62.9233	0.7402	10.7208	0.1261
105	1.2000	63.6656	0.7490	10.6544	0.1253
106	1.2500	64.3929	0.7576	10.5641	0.1243
107	1.3000	65.1045	0.7659	10.4512	0.1230

ตาราง 37 (ต่อ)

ลำดับที่	θ	$\sum P(\theta)$	$\sum P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
108	1.3500	65.8000	0.7741	10.3173	0.1214
109	1.4000	66.4791	0.7821	10.1641	0.1196
110	1.4500	67.1416	0.7899	9.9936	0.1176
111	1.5000	67.7873	0.7975	9.8081	0.1154
112	1.5500	68.4162	0.8049	9.6098	0.1131
113	1.6000	69.0283	0.8121	9.4011	0.1106
114	1.6500	69.6237	0.8191	9.1842	0.1080
115	1.7000	70.2024	0.8259	8.9615	0.1054
116	1.7500	70.7648	0.8325	8.7348	0.1028
117	1.8000	71.3111	0.8390	8.5062	0.1001
118	1.8500	71.8415	0.5452	8.2774	0.0974
119	1.9000	72.3563	0.8513	8.0496	0.0947
120	1.9500	72.8559	0.8571	7.8242	0.0920
121	2.0000	73.3406	0.8628	7.6021	0.0894
122	2.0500	73.8107	0.8684	7.3839	0.0869
123	2.1000	74.2665	0.8737	7.1702	0.0844
124	2.1500	74.7084	0.8790	6.9610	0.0819
125	2.2000	75.1367	0.8840	6.7566	0.0795
126	2.2500	75.5517	0.8888	6.5569	0.0771
127	2.3000	75.9537	0.8936	6.3616	0.0748
128	2.3500	76.3430	0.8982	6.1704	0.0726
129	2.4000	76.7199	0.9026	5.9831	0.0704
130	2.4500	77.0846	0.9096	5.7991	0.0682
131	2.5000	77.4373	0.9110	5.6183	0.0661
132	2.5500	77.7784	0.9150	5.4401	0.0640
133	2.6000	78.1080	0.9189	5.2644	0.0619
134	2.6500	78.4264	0.9227	5.0909	0.0599

ตาราง 37 (ต่อ)

ลำดับที่	θ	$\Sigma P(\theta)$	$\Sigma P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
135	2.7000	78.7337	0.9263	4.9192	0.0579
136	2.7500	79.0302	0.9298	4.7494	0.0559
137	2.8000	79.3160	0.9331	4.5814	0.0539
138	2.8500	79.5914	0.9364	4.4150	0.0519
139	2.9000	79.8565	0.9395	4.2505	0.0500
140	2.9500	80.1115	0.9425	4.0877	0.0481
141	3.0000	80.3566	0.9454	3.9269	0.0462
142	3.0500	80.5920	0.9481	3.7682	0.0443
143	3.1000	80.8179	0.9508	3.6119	0.0425
144	3.1500	81.0344	0.9533	3.4579	0.0407
145	3.2000	81.2418	0.9558	3.3067	0.0389
146	3.2500	81.4404	0.9581	3.1584	0.0372
147	3.3000	81.6302	0.9604	3.0131	0.0354
148	3.3500	81.8116	0.9625	2.8711	0.0338
149	3.4000	81.9847	0.9645	2.7326	0.0321
150	3.4500	82.1499	0.9665	2.5977	0.0306
151	3.5000	82.3073	0.9683	2.4666	0.0290
152	3.5500	82.4572	0.9701	2.3394	0.0275
153	3.6000	82.5999	0.9718	2.2164	0.0261
154	3.6500	82.7357	0.9734	2.0974	0.0247
155	3.7000	82.8647	0.9749	1.9828	0.0233
156	3.7500	82.9873	0.9763	1.8725	0.0220
157	3.8000	83.1037	0.9777	1.7666	0.0208
158	3.8500	83.2141	0.9790	1.6651	0.0196
159	3.9000	83.3190	0.9802	1.5680	0.0184
160	3.9500	83.4184	0.9814	1.4753	0.0174
161	4.0000	83.5126	0.9825	1.3870	0.0163

ตาราง 38 ค่าฟังก์ชันสารสนเทศของแบบทดสอบและค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบ
ชั้น ม.2

ลำดับที่	θ	$\sum P(\theta)$	$\sum P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
1	-4.0000	21.7338	0.2557	0.0071	0.0001
2	-3.9500	21.7576	0.2560	0.0083	0.0001
3	-3.9000	21.7833	0.2563	0.0096	0.0001
4	-3.8500	21.8110	0.2566	0.0111	0.0001
5	-3.8000	21.8408	0.2570	0.0129	0.0002
6	-3.7500	21.8729	0.2573	0.0149	0.0002
7	-3.7000	21.9075	0.2577	0.0171	0.0002
8	-3.6500	21.9447	0.2582	0.0197	0.0002
9	-3.6000	21.9847	0.2586	0.0226	0.0003
10	-3.5500	22.0276	0.2591	0.0259	0.0003
11	-3.5000	22.0737	0.2597	0.0296	0.0003
12	-3.4500	22.1230	0.2603	0.0337	0.0004
13	-3.4000	22.1759	0.2609	0.0384	0.0005
14	-3.3500	22.2324	0.2616	0.0436	0.0005
15	-3.3000	22.2928	0.2623	0.0494	0.0006
16	-3.2500	22.3574	0.2630	0.0558	0.0007
17	-3.2000	22.4263	0.2638	0.0629	0.0007
18	-3.1500	22.4997	0.2647	0.0708	0.0008
19	-3.1000	22.5779	0.2656	0.0795	0.0009
20	-3.0500	22.6612	0.2666	0.0890	0.0010
21	-3.0000	22.7496	0.2676	0.0995	0.0012
22	-2.9500	22.8436	0.2687	0.1111	0.0013
23	-2.9000	22.9434	0.2699	0.1237	0.0015
24	-2.8500	23.0491	0.2712	0.1375	0.0016
25	-2.8000	23.1612	0.2728	0.1525	0.0018
26	-2.7500	23.2798	0.2739	0.1690	0.0020

ตาราง 38 (ต่อ)

ลำดับที่	θ	$\sum P(\theta)$	$\sum P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
27	-2.7000	23.4053	0.2754	0.1870	0.0022
28	-2.6500	23.5380	0.2769	0.2066	0.0024
29	-2.6000	23.6782	0.2786	0.2279	0.0027
30	-2.5500	23.8262	0.2803	0.2513	0.0030
31	-2.5000	23.9824	0.2821	0.2768	0.0033
32	-2.4500	24.1471	0.2841	0.3047	0.0036
33	-2.4000	24.3207	0.2861	0.3352	0.0039
34	-2.3500	24.5036	0.2883	0.3686	0.0043
35	-2.3000	24.6962	0.2905	0.4051	0.0048
36	-2.2500	24.8988	0.2929	0.4452	0.0052
37	-2.2000	25.1119	0.2954	0.4891	0.0058
38	-2.1500	25.3359	0.2981	0.5372	0.0063
39	-2.1000	25.5713	0.3008	0.5899	0.0069
40	-2.0500	25.8185	0.3037	0.6477	0.0076
41	-2.0000	26.0778	0.3068	0.7108	0.0084
42	-1.9500	26.3497	0.3100	0.7798	0.0092
43	-1.9000	26.6347	0.3133	0.8550	0.0101
44	-1.8500	26.9331	0.3169	0.9368	0.0110
45	-1.8000	27.2452	0.3205	1.0255	0.0121
46	-1.7500	27.5714	0.3244	1.1214	0.0132
47	-1.7000	27.9120	0.3284	1.2248	0.0144
48	-1.6500	28.2673	0.3326	1.3357	0.0157
49	-1.6000	28.6374	0.3369	1.4542	0.0171
50	-1.5500	29.0226	0.3414	1.5802	0.0186
51	-1.5000	29.4229	0.3461	1.7137	0.0202
52	-1.4500	29.8384	0.3510	1.8544	0.0218
53	-1.4000	30.2690	0.3561	2.0018	0.0236

ตาราง 38 (ต่อ)

ลำดับที่	θ	$\Sigma P(\theta)$	$\Sigma P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
54	-1.3500	30.7147	0.3613	2.1555	0.0254
55	-1.3000	31.1754	0.3668	2.3148	0.0272
56	-1.2500	31.6508	0.3724	2.4791	0.0292
57	-1.2000	32.1406	0.3781	2.6476	0.0311
58	-1.1500	32.6445	0.3841	2.8193	0.0332
59	-1.1000	33.1622	0.3901	2.9935	0.0352
60	-1.0500	33.6931	0.3964	3.1690	0.0373
61	-1.0000	34.2369	0.4028	3.3450	0.0394
62	-0.9500	34.7929	0.4093	3.5204	0.0414
63	-0.9000	35.3606	0.4160	3.6944	0.0435
64	-0.8500	35.9394	0.4228	3.8660	0.0455
65	-0.8000	36.5288	0.4298	4.0345	0.0475
66	-0.7500	37.1281	0.4368	4.1991	0.0494
67	-0.7000	37.7366	0.4440	4.3593	0.0513
68	-0.6500	38.3538	0.4512	4.5147	0.0531
69	-0.6000	38.9790	0.4586	4.6650	0.0549
70	-0.5500	39.6117	0.4660	4.8101	0.0566
71	-0.5000	40.2512	0.4735	4.9502	0.0582
72	-0.4500	40.8971	0.4811	5.0855	0.0598
73	-0.4000	41.5487	0.4888	5.2165	0.0614
74	-0.3500	42.2056	0.4965	5.3439	0.0629
75	-0.3000	42.8674	0.5043	5.4685	0.0643
76	-0.2500	43.5337	0.5122	5.5913	0.0658
77	-0.2000	44.2039	0.5200	5.7134	0.0672
78	-0.1500	44.8779	0.5280	5.8360	0.0687
79	-0.1000	45.5553	0.5359	5.9606	0.0701
80	-0.0500	46.2358	0.5440	6.0883	0.0716

ตาราง 38 (ต่อ)

ลำดับที่	θ	$\sum P(\theta)$	$\sum P(\theta)/l$	$I(\theta)$	$I(\theta)/l$
81	0.0000	46.9192	0.5520	6.2205	0.0732
82	0.0500	47.6051	0.5601	6.3584	0.0748
83	0.1000	48.2935	0.5682	6.5034	0.0765
84	0.1500	48.9842	0.5763	6.6563	0.0783
85	0.2000	49.6768	0.5844	6.8180	0.0802
86	0.2500	50.3712	0.5926	6.9891	0.0822
87	0.3000	51.0672	0.6008	7.1699	0.0844
88	0.3500	51.7645	0.6090	7.3605	0.0866
89	0.4000	52.4629	0.6172	7.5604	0.0889
90	0.4500	53.1621	0.6254	7.7692	0.0914
91	0.5000	53.8617	0.6337	7.9857	0.0939
92	0.5500	54.5615	0.6419	8.2086	0.0966
93	0.6000	55.2609	0.6501	8.4364	0.0993
94	0.6500	55.9595	0.6583	8.6670	0.1020
95	0.7000	56.6568	0.6666	8.8981	0.1047
96	0.7500	57.3524	0.6747	9.1275	0.1074
97	0.8000	58.0455	0.6829	9.3523	0.1100
98	0.8500	58.7355	0.6910	9.5699	0.1126
99	0.9000	59.4219	0.6991	9.7774	0.1150
100	0.9500	60.1039	0.7071	9.9719	0.1173
101	1.0000	60.7809	0.7151	10.1507	0.1194
102	1.0500	61.4520	0.7230	10.3110	0.1213
103	1.1000	62.1167	0.7308	10.4503	0.1229
104	1.1500	62.7741	0.7385	10.5662	0.1243
105	1.2000	63.4235	0.7462	10.6567	0.1254
106	1.2500	64.0644	0.7536	10.7201	0.1261
107	1.3000	64.6959	0.7611	10.7549	0.1265

ตาราง 38 (ต่อ)

ลำดับที่	θ	$\sum P(\theta)$	$\sum P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
108	1.3500	65.3176	0.7684	10.7603	0.1266
109	1.4000	65.9287	0.7756	10.7355	0.1263
110	1.4500	66.5287	0.7827	10.6805	0.1257
111	1.5000	67.1172	0.7897	10.5956	0.1247
112	1.5500	67.6937	0.7964	10.4814	0.1233
113	1.6000	68.2577	0.8030	10.3391	0.1216
114	1.6500	68.8091	0.8095	10.1700	0.1196
115	1.7000	69.3475	0.8159	9.9760	0.1174
116	1.7500	69.8726	0.8220	9.7592	0.1148
117	1.8000	70.3845	0.8281	9.5217	0.1120
118	1.8500	70.8829	0.8339	9.2660	0.1090
119	1.9000	71.3679	0.8396	8.9947	0.1058
120	1.9500	71.8395	0.8452	8.7102	0.1025
121	2.0000	72.2978	0.8506	8.4152	0.0990
122	2.0500	72.7428	0.8558	8.1122	0.0954
123	2.1000	73.1748	0.8601	7.8036	0.0918
124	2.1500	73.5939	0.8658	7.4916	0.0881
125	2.2000	74.0004	0.8706	7.1784	0.0845
126	2.2500	74.3946	0.8752	6.8659	0.0808
127	2.3000	74.7767	0.8797	6.5560	0.0771
128	2.3500	75.1470	0.8841	6.2502	0.0735
129	2.4000	75.5059	0.8883	5.9498	0.0700
130	2.4500	75.8536	0.8924	5.6560	0.0665
131	2.5000	76.1906	0.8964	5.3699	0.0632
132	2.5500	76.5170	0.9002	5.0923	0.0599
133	2.6000	76.8333	0.9039	4.8238	0.0568
134	2.6500	77.1398	0.9075	4.5649	0.0537

ตาราง 38 (ต่อ)

ลำดับที่	θ	$\Sigma P(\theta)$	$\Sigma P(\theta)/l$	$I(\theta)$	$I(\theta)/l$
135	2.7000	77.4367	0.9110	4.3161	0.0508
136	2.7500	77.7245	0.9144	4.0775	0.0480
137	2.8000	78.0034	0.9177	3.8493	0.0453
138	2.8500	78.2736	0.9209	3.6316	0.0427
139	2.9000	78.5355	0.9239	3.4243	0.0403
140	2.9500	78.7894	0.9269	3.2273	0.0380
141	3.0000	79.0354	0.9298	3.0405	0.0358
142	3.0500	79.2739	0.9326	2.8637	0.0337
143	3.1000	79.5050	0.9354	2.6966	0.0317
144	3.1500	79.7290	0.9380	2.5389	0.0299
145	3.2000	79.9460	0.9405	2.3904	0.0281
146	3.2500	80.1564	0.9430	2.2506	0.0265
147	3.3000	80.3602	0.9454	2.1193	0.0249
148	3.3500	80.5576	0.9477	1.9961	0.0235
149	3.4000	80.7488	0.9500	1.8807	0.0221
150	3.4500	80.9341	0.9522	1.7726	0.0209
151	3.5000	81.1134	0.9543	1.6716	0.0197
152	3.5500	81.2870	0.9563	1.5773	0.0186
153	3.6000	81.4551	0.9583	1.4893	0.0175
154	3.6500	81.6177	0.9602	1.4073	0.0166
155	3.7000	81.7749	0.9621	1.3310	0.0157
156	3.7500	81.9270	0.9638	1.2599	0.0148
157	3.8000	82.0740	0.9656	1.1938	0.0140
158	3.8500	82.2160	0.9672	1.1324	0.0133
159	3.9000	82.3532	0.9689	1.0753	0.0127
160	3.9500	82.4856	0.9704	1.0221	0.0120
161	4.0000	82.6133	0.9719	0.9727	0.0114

ตาราง 39 ค่าฟังก์ชันสารสนเทศของแบบทดสอบและค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบ
ชั้น ม.3

ลำดับที่	θ	$\sum P(\theta)$	$\sum P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
1	-4.0000	18.7331	0.2204	0.0514	0.0006
2	-3.9500	18.7875	0.2210	0.0591	0.0007
3	-3.9000	18.8459	0.2217	0.0678	0.0008
4	-3.8500	18.9087	0.2225	0.0776	0.0009
5	-3.8000	18.9761	0.2232	0.0886	0.0010
6	-3.7500	19.0484	0.2241	0.1009	0.0012
7	-3.7000	19.1258	0.2250	0.1147	0.0013
8	-3.6500	19.2088	0.2260	0.1299	0.0015
9	-3.6000	19.2974	0.2270	0.1467	0.0017
10	-3.5500	19.3921	0.2281	0.1652	0.0019
11	-3.5000	19.4932	0.2293	0.1854	0.0022
12	-3.4500	19.6009	0.2306	0.2074	0.0024
13	-3.4000	19.7155	0.2319	0.2312	0.0027
14	-3.3500	19.8373	0.2334	0.2569	0.0030
15	-3.3000	19.9667	0.2349	0.2844	0.0033
16	-3.2500	20.1039	0.2365	0.3138	0.0037
17	-3.2000	20.2492	0.2382	0.3449	0.0041
18	-3.1500	20.4029	0.2400	0.3778	0.0044
19	-3.1000	20.5653	0.2419	0.4124	0.0049
20	-3.0500	20.7367	0.2440	0.4485	0.0053
21	-3.0000	20.9173	0.2461	0.4862	0.0057
22	-2.9500	21.1075	0.2483	0.5252	0.0062
23	-2.9000	21.3073	0.2507	0.5655	0.0067
24	-2.8500	21.5172	0.2531	0.6070	0.0071
25	-2.8000	21.7374	0.2557	0.6496	0.0076
26	-2.7500	21.9680	0.2584	0.6933	0.0082

ตาราง 39 (ต่อ)

ลำดับที่	θ	$\sum P(\theta)$	$\sum P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
27	-2.7000	22.2094	0.2613	0.7379	0.0087
28	-2.6500	22.4617	0.2642	0.7835	0.0092
29	-2.6000	22.7251	0.2674	0.8302	0.0098
30	-2.5500	22.9998	0.2706	0.8778	0.0103
31	-2.5000	23.2861	0.2740	0.9265	0.0109
32	-2.4500	23.5840	0.2775	0.9764	0.0115
33	-2.4000	23.8937	0.2811	1.0275	0.0121
34	-2.3500	24.2154	0.2849	1.0799	0.0127
35	-2.3000	24.5491	0.2888	1.1338	0.0133
36	-2.2500	24.8950	0.2929	1.1892	0.0140
37	-2.2000	25.2531	0.2971	1.2463	0.0147
38	-2.1500	25.6235	0.3015	1.3051	0.0154
39	-2.1000	26.0062	0.3060	1.3656	0.0161
40	-2.0500	26.4011	0.3106	1.4280	0.0168
41	-2.0000	26.8083	0.3154	1.4922	0.0176
42	-1.9500	27.2276	0.3203	1.5581	0.0183
43	-1.9000	27.6591	0.3254	1.6257	0.0191
44	-1.8500	28.1026	0.3306	1.6948	0.0199
45	-1.8000	28.5580	0.3360	1.7654	0.0208
46	-1.7500	29.0250	0.3415	1.8372	0.0216
47	-1.7000	29.5036	0.3471	1.9100	0.0225
48	-1.6500	29.9935	0.3529	1.9838	0.0233
49	-1.6000	30.4946	0.3588	2.0581	0.0242
50	-1.5500	31.0065	0.3648	2.1328	0.0251
51	-1.5000	31.5290	0.3709	2.2078	0.0260
52	-1.4500	32.0620	0.3772	2.2827	0.0269
53	-1.4000	32.6050	0.3836	2.3573	0.0277

ตาราง 39 (ต่อ)

ลำดับที่	θ	$\sum P(\theta)$	$\sum P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
54	-1.3500	33.1578	0.3901	2.4316	0.0286
55	-1.3000	33.7202	0.3967	2.5053	0.0295
56	-1.2500	34.2917	0.4031	2.5783	0.0303
57	-1.2000	34.8722	0.4103	2.6505	0.0312
58	-1.1500	35.4612	0.4172	2.7217	0.0320
59	-1.1000	36.0585	0.4242	2.7920	0.0328
60	-1.0500	36.6636	0.4313	2.8612	0.0337
61	-1.0000	37.2763	0.4385	2.9293	0.0345
62	-0.9500	37.8961	0.4458	2.9962	0.0352
63	-0.9000	38.5228	0.4532	3.0619	0.0360
64	-0.8500	39.1559	0.4607	3.1263	0.0368
65	-0.8000	39.7950	0.4682	3.1895	0.0375
66	-0.7500	40.4398	0.4758	3.2513	0.0383
67	-0.7000	41.0898	0.4834	3.3117	0.0390
68	-0.6500	41.7447	0.4911	3.3707	0.0397
69	-0.6000	42.4040	0.4989	3.4283	0.0403
70	-0.5500	43.0673	0.5067	3.4844	0.0410
71	-0.5000	43.7342	0.5145	3.5392	0.0416
72	-0.4500	44.4043	0.5224	3.5925	0.0423
73	-0.4000	45.0771	0.5303	3.6445	0.0429
74	-0.3500	45.7522	0.5383	3.6953	0.0435
75	-0.3000	46.4293	0.5462	3.7449	0.0441
76	-0.2500	47.1079	0.5542	3.7936	0.0446
77	-0.2000	47.7876	0.5622	3.8414	0.0452
78	-0.1500	48.4681	0.5702*	3.8885	0.0457
79	-0.1000	49.1489	0.5782	3.9351	0.0463
80	-0.0500	49.8297	0.5862	3.9814	0.0468

ตาราง 39 (ต่อ)

ลำดับที่	θ	$\sum P(\theta)$	$\sum P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
81	0.0000	50.5101	0.5942	4.0276	0.0474
82	0.0500	51.1899	0.6022	4.0738	0.0479
83	0.1000	51.8685	0.6102	4.1203	0.0485
84	0.1500	52.5459	0.6181	4.1671	0.0490
85	0.2000	53.2215	0.6261	4.2143	0.0496
86	0.2500	53.8952	0.6341	4.2619	0.0501
87	0.3000	54.5665	0.6420	4.3100	0.0507
88	0.3500	55.2352	0.6498	4.3583	0.0513
89	0.4000	55.9011	0.6577	4.4069	0.0518
90	0.4500	56.5636	0.6655	4.4553	0.0524
91	0.5000	57.2227	0.6732	4.5034	0.0530
92	0.5500	57.8779	0.6809	4.5508	0.0535
93	0.6000	58.5289	0.6886	4.5970	0.0541
94	0.6500	59.1754	0.6962	4.6417	0.0546
95	0.7000	59.8171	0.7037	4.6842	0.0551
96	0.7500	60.4535	0.7112	4.7242	0.0556
97	0.8000	61.0844	0.7186	4.7611	0.0560
98	0.8500	61.7094	0.7260	4.7943	0.0564
99	0.9000	62.3282	0.7333	4.8234	0.0567
100	0.9500	62.9403	0.7405	4.8480	0.0570
101	1.0000	63.5454	0.7476	4.8675	0.0573
102	1.0500	64.1433	0.7546	4.8816	0.0574
103	1.1000	64.7334	0.7616	4.8899	0.0575
104	1.1500	65.3156	0.7684	4.8923	0.0576
105	1.2000	65.8895	0.7752	4.8884	0.0575
106	1.2500	66.4549	0.7818	4.8780	0.0574
107	1.3000	67.0113	0.7884	4.8612	0.0572

ตาราง 39 (ต่อ)

ลำดับที่	θ	$\sum P(\theta)$	$\sum P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
108	1.3500	67.5588	0.7948	4.8377	0.0569
109	1.4000	68.0969	0.8011	4.8077	0.0566
110	1.4500	68.6255	0.8074	4.7710	0.0561
111	1.5000	69.1446	0.8135	4.7279	0.0556
112	1.5500	69.6538	0.8195	4.6785	0.0550
113	1.6000	70.1532	0.8253	4.6230	0.0544
114	1.6500	70.6427	0.8311	4.5616	0.0537
115	1.7000	71.1221	0.8367	4.4946	0.0529
116	1.7500	71.5917	0.8423	4.4223	0.0520
117	1.8000	72.0512	0.8477	4.3452	0.0511
118	1.8500	72.5008	0.8530	4.2636	0.0502
119	1.9000	72.9405	0.8581	4.1780	0.0492
120	1.9500	73.3704	0.8632	4.0889	0.0481
121	2.0000	73.7906	0.8681	3.9968	0.0470
122	2.0500	74.2012	0.8730	3.9021	0.0459
123	2.1000	74.6023	0.8777	3.8055	0.0448
124	2.1500	74.9941	0.8823	3.7074	0.0436
125	2.2000	75.3765	0.8868	3.6083	0.0425
126	2.2500	75.7499	0.8912	3.5087	0.0413
127	2.3000	76.1143	0.8955	3.4090	0.0401
128	2.3500	76.4698	0.8997	3.3098	0.0389
129	2.4000	76.8166	0.9037	3.2113	0.0378
130	2.4500	77.1548	0.9077	3.1139	0.0366
131	2.5000	77.4844	0.9116	3.0179	0.0355
132	2.5500	77.8056	0.9154	2.9235	0.0344
133	2.6000	78.1185	0.9190	2.8309	0.0333
134	2.6500	78.4230	0.9226	2.7402	0.0322

ตาราง 39 (ต่อ)

ลำดับที่	θ	$\Sigma P(\theta)$	$\Sigma P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
135	2.7000	78.7193	0.9261	2.6516	0.0312
136	2.7500	79.0074	0.9295	2.5650	0.0302
137	2.8000	79.2874	0.9328	2.4804	0.0292
138	2.8500	79.5592	0.9360	2.3979	0.0282
139	2.9000	79.8229	0.9391	2.3172	0.0273
140	2.9500	80.0785	0.9421	2.2384	0.0263
141	3.0000	80.3260	0.9450	2.1613	0.0254
142	3.0500	80.5653	0.9478	2.0858	0.0245
143	3.1000	80.7966	0.9505	2.0117	0.0237
144	3.1500	81.0198	0.9532	1.9390	0.0228
145	3.2000	81.2349	0.9557	1.8675	0.0220
146	3.2500	81.4420	0.9581	1.7971	0.0211
147	3.3000	81.6411	0.9605	1.7278	0.0203
148	3.3500	81.8323	0.9627	1.6594	0.0195
149	3.4000	82.0156	0.9649	1.5921	0.0187
150	3.4500	82.1911	0.9670	1.5257	0.0179
151	3.5000	82.3588	0.9689	1.4603	0.0172
152	3.5500	82.5190	0.9708	1.3960	0.0164
153	3.6000	82.6717	0.9726	1.3327	0.0157
154	3.6500	82.8171	0.9743	1.2706	0.0149
155	3.7000	82.9553	0.9759	1.2097	0.0142
156	3.7500	83.0866	0.9775	1.1501	0.0135
157	3.8000	83.2110	0.9790	1.0919	0.0128
158	3.8500	83.3287	0.9803	1.0353	0.0122
159	3.9000	83.4401	0.9816	0.9802	0.0115
160	3.9500	83.5453	0.9829	0.9267	0.0109
161	4.0000	83.6444	0.9841	0.8750	0.0103

ตาราง 40 ค่าฟังก์ชันสารสนเทศของแบบทดสอบและค่าฟังก์ชันลักษณะข้อสอบของแบบทดสอบ
ความสามารถทั่วไปทางภาษาไทย

ลำดับที่	θ	$\Sigma P(\theta)$	$\Sigma P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
1	-4.0000	4.6396	0.1600	0.0001	0.0000
2	-3.9500	4.6414	0.1600	0.0001	0.0000
3	-3.9000	4.6433	0.1601	0.0001	0.0000
4	-3.8500	4.6453	0.1602	0.0001	0.0000
5	-3.8000	4.6475	0.1603	0.0002	0.0000
6	-3.7500	4.6500	0.1603	0.0002	0.0000
7	-3.7000	4.6526	0.1604	0.0002	0.0000
8	-3.6500	4.6555	0.1605	0.0003	0.0000
9	-3.6000	4.6586	0.1606	0.0003	0.0000
10	-3.5500	4.6620	0.1608	0.0004	0.0000
11	-3.5000	4.6657	0.1609	0.0004	0.0000
12	-3.4500	4.6697	0.1610	0.0005	0.0000
13	-3.4000	4.6741	0.1611	0.0006	0.0000
14	-3.3500	4.6789	0.1613	0.0007	0.0000
15	-3.3000	4.6840	0.1615	0.0008	0.0000
16	-3.2500	4.6896	0.1617	0.0010	0.0000
17	-3.2000	4.6957	0.1619	0.0012	0.0000
18	-3.1500	4.7024	0.1621	0.0014	0.0000
19	-3.1000	4.7096	0.1624	0.0016	0.0001
20	-3.0500	4.7174	0.1628	0.0019	0.0001
21	-3.0000	4.7259	0.1630	0.0022	0.0001
22	-2.9500	4.7351	0.1633	0.0026	0.0001
23	-2.9000	4.7451	0.1636	0.0031	0.0001
24	-2.8500	4.7559	0.1640	0.0036	0.0001
25	-2.8000	4.7677	0.1644	0.0042	0.0001
26	-2.7500	4.7804	0.1648	0.0050	0.0002

ตาราง 40 (ต่อ)

ลำดับที่	θ	$\sum P(\theta)$	$\sum P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
27	-2.7000	4.7943	0.1653	0.0058	0.0002
28	-2.6500	4.8093	0.1658	0.0068	0.0002
29	-2.6000	4.8255	0.1664	0.0080	0.0003
30	-2.5500	4.8431	0.1670	0.0093	0.0003
31	-2.5000	4.8621	0.1677	0.0109	0.0004
32	-2.4500	4.8826	0.1684	0.0127	0.0004
33	-2.4000	4.9049	0.1691	0.0148	0.0005
34	-2.3500	4.9289	0.1700	0.0172	0.0006
35	-2.3000	4.9548	0.1709	0.0199	0.0007
36	-2.2500	4.9828	0.1718	0.0231	0.0008
37	-2.2000	5.0130	0.1729	0.0267	0.0009
38	-2.1500	5.0456	0.1740	0.0308	0.0011
39	-2.1000	5.0806	0.1752	0.0355	0.0012
40	-2.0500	5.1183	0.1765	0.0408	0.0014
41	-2.0000	5.1589	0.1779	0.0469	0.0016
42	-1.9500	5.2025	0.1794	0.0537	0.0019
43	-1.9000	5.2492	0.1810	0.0613	0.0021
44	-1.8500	5.2994	0.1827	0.0699	0.0024
45	-1.8000	5.3531	0.1846	0.0794	0.0027
46	-1.7500	5.4106	0.1866	0.0901	0.0031
47	-1.7000	5.4720	0.1887	0.1019	0.0035
48	-1.6500	5.5376	0.1910	0.1149	0.0040
49	-1.6000	5.6076	0.1934	0.1293	0.0045
50	-1.5500	5.6822	0.1959	0.1450	0.0050
51	-1.5000	5.7615	0.1987	0.1622	0.0056
52	-1.4500	5.8458	0.2016	0.1808	0.0062
53	-1.4000	5.9352	0.2047	0.2011	0.0069

ตาราง 40 (ต่อ)

ลำดับที่	θ	$\Sigma P(\theta)$	$\Sigma P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
54	-1.3500	6.0300	0.2079	0.2230	0.0077
55	-1.3000	6.1303	0.2114	0.2467	0.0085
56	-1.2500	6.2364	0.2150	0.2720	0.0094
57	-1.2000	6.3484	0.2189	0.2992	0.0103
58	-1.1500	6.4664	0.2230	0.3283	0.0113
59	-1.1000	6.5906	0.2272	0.3594	0.0124
60	-1.0500	6.7212	0.2318	0.3925	0.0135
61	-1.0000	6.8584	0.2365	0.4277	0.0147
62	-0.9500	7.0021	0.2415	0.4653	0.0160
63	-0.9000	7.1527	0.2466	0.5052	0.0174
64	-0.8500	7.3100	0.2521	0.5478	0.0189
65	-0.8000	7.4743	0.2577	0.5933	0.0205
66	-0.7500	7.6457	0.2636	0.6419	0.0221
67	-0.7000	7.8241	0.2698	0.6940	0.0239
68	-0.6500	8.0097	0.2762	0.7500	0.0259
69	-0.6000	8.2025	0.2828	0.8104	0.0279
70	-0.5500	8.4025	0.2897	0.8758	0.0302
71	-0.5000	8.6098	0.2969	0.9469	0.0327
72	-0.4500	8.8243	0.3043	1.0245	0.0353
73	-0.4000	9.0461	0.3119	1.1094	0.0383
74	-0.3500	9.2751	0.3198	1.2028	0.0415
75	-0.3000	9.5114	0.3280	1.3058	0.0450
76	-0.2500	9.7548	0.3364	1.4198	0.0490
77	-0.2000	10.0053	0.3450	1.5463	0.0533
78	-0.1500	10.2629	0.3539	1.6870	0.0582
79	-0.1000	10.5276	0.3630	1.8435	0.0636
80	-0.0500	10.7991	0.3724	2.0179	0.0696

ตาราง 40 (ต่อ)

ลำดับที่	θ	$\Sigma P(\theta)$	$\Sigma P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
81	0.0000	11.0774	0.3820	2.2123	0.0763
82	0.0500	11.3625	0.3918	2.4289	0.0838
83	0.1000	11.6541	0.4019	2.6699	0.0921
84	0.1500	11.9521	0.4121	2.9377	0.1013
85	0.2000	12.2563	0.4226	3.2347	0.1115
86	0.2500	12.5665	0.4333	3.5632	0.1229
87	0.3000	12.8826	0.4442	3.9255	0.1354
88	0.3500	13.2042	0.4553	4.3238	0.1491
89	0.4000	13.5311	0.4666	4.7600	0.1641
90	0.4500	13.8630	0.4780	5.2357	0.1805
91	0.5000	14.1996	0.4896	5.7521	0.1983
92	0.5500	14.5405	0.5014	6.3101	0.2176
93	0.6000	14.8854	0.5133	6.9099	0.2383
94	0.6500	15.2339	0.5253	7.5510	0.2604
95	0.7000	15.5855	0.5374	8.2324	0.2839
96	0.7500	15.9398	0.5496	8.9520	0.3087
97	0.8000	16.2963	0.5619	9.7069	0.3347
98	0.8500	16.6546	0.5743	10.4934	0.3618
99	0.9000	17.0140	0.5867	11.3065	0.3899
100	0.9500	17.3742	0.5991	12.1406	0.4186
101	1.0000	17.7345	0.6115	12.9890	0.4479
102	1.0500	18.0944	0.6239	13.8439	0.4774
103	1.1000	18.4532	0.6363	14.6971	0.5068
104	1.1500	18.8105	0.6486	15.5393	0.5358
105	1.2000	19.1656	0.6609	16.3612	0.5642
106	1.2500	19.5179	0.6730	17.1530	0.5915
107	1.3000	19.8668	0.6851	17.9046	0.6174

ตาราง 40 (ต่อ)

ลำดับที่	θ	$\Sigma P(\theta)$	$\Sigma P(\theta)/1$	$I(\theta)$	$I(\theta)/1$
108	1.3500	20.2118	0.6970	18.6067	0.6416
109	1.4000	20.5523	0.7087	19.2498	0.6638
110	1.4500	20.8877	0.7203	19.8257	0.6836
111	1.5000	21.2175	0.7316	20.3266	0.7009
112	1.5500	21.5412	0.7428	20.7463	0.7154
113	1.6000	21.8584	0.7537	21.0796	0.7269
114	1.6500	22.1685	0.7644	21.3230	0.7353
115	1.7000	22.4713	0.7749	21.4742	0.7405
116	1.7500	22.7663	0.7850	21.5327	0.7425
117	1.8000	23.0532	0.7949	21.4994	0.7414
118	1.8500	23.3318	0.8045	21.3766	0.7371
119	1.9000	23.6019	0.8139	21.1681	0.7299
120	1.9500	23.8632	0.8229	20.8786	0.7200
121	2.0000	24.1156	0.8316	20.5139	0.7074
122	2.0500	24.3591	0.8400	20.0807	0.6924
123	2.1000	24.5936	0.8481	19.5862	0.6754
124	2.1500	24.8190	0.8558	19.0379	0.6565
125	2.2000	25.0355	0.8633	18.4436	0.6360
126	2.2500	25.2431	0.8705	17.8112	0.6142
127	2.3000	25.4419	0.8773	17.1482	0.5913
128	2.3500	25.6320	0.8839	16.4622	0.5677
129	2.4000	25.8137	0.8901	15.7602	0.5435
130	2.4500	25.9870	0.8961	15.0486	0.5189
131	2.5000	26.1522	0.9018	14.3336	0.4943
132	2.5500	26.3096	0.9072	13.6206	0.4697
133	2.6000	26.4593	0.9124	12.9144	0.4453
134	2.6500	26.6017	0.9173	12.2194	0.4214

ตาราง 40 (ต่อ)

ลำดับที่	θ	$\Sigma P(\theta)$	$\Sigma P(\theta)/l$	$I(\theta)$	$I(\theta)/l$
135	2.7000	26.7369	0.9220	11.5391	0.3979
136	2.7500	26.8654	0.9264	10.8766	0.3751
137	2.8000	26.9873	0.9306	10.2345	0.3529
138	2.8500	27.1029	0.9346	9.6148	0.3315
139	2.9000	27.2126	0.9384	9.0191	0.3110
140	2.9500	27.3165	0.9419	8.4485	0.2913
141	3.0000	27.4150	0.9453	7.9036	0.2725
142	3.0500	27.5083	0.9486	7.3851	0.2547
143	3.1000	27.5967	0.9506	6.8928	0.2377
144	3.1500	27.6804	0.9545	6.4268	0.2216
145	3.2000	27.7597	0.9572	5.9866	0.2064
146	3.2500	27.8348	0.9598	5.5718	0.1921
147	3.3000	27.9059	0.9623	5.1815	0.1787
148	3.3500	27.9731	0.9646	4.8151	0.1660
149	3.4000	28.0368	0.9668	4.4716	0.1542
150	3.4500	28.0970	0.9689	4.1501	0.1431
151	3.5000	28.1540	0.9708	3.8495	0.1327
152	3.5500	28.2079	0.9728	3.5689	0.1231
153	3.6000	28.2588	0.9744	3.3071	0.1140
154	3.6500	28.3069	0.9761	3.0633	0.1056
155	3.7000	28.3523	0.9778	2.8363	0.0978
156	3.7500	28.3952	0.9791	2.6251	0.0905
157	3.8000	28.4356	0.9805	2.4288	0.0838
158	3.8500	28.4736	0.9818	2.2465	0.0775
159	3.9000	28.5095	0.9831	2.0772	0.0716
160	3.9500	28.5432	0.9842	1.9200	0.0662
161	4.0000	28.5749	0.9853	1.7743	0.0612

๖. ความหมาย

ตาราง 41 น้ำหนักองค์ประกอบของแบบทดสอบชั้น ม.1

ข้อ	องค์ประกอบที่ 1	องค์ประกอบที่ 2	องค์ประกอบที่ 3
1	.277	.030	.010
2	.116	.064	.088
3	.108	.127	.165
4	.264	.043	.205
5	.360	.033	.463
6	.406	.066	.304
7	.330	.105	.018
8	.467	.286	.060
9	.353	.412	.010
10	.217	.135	.122
11	.300	.160	.093
12	.219	.084	.165
13	.212	.190	.087
14	.232	.116	.033
15	.063	.316	.125
16	.143	.044	.013
17	.481	.409	.004
18	.454	.364	.004
19	.208	.149	.109
20	.320	.158	.002
21	.367	.290	.048
22	.320	.273	.005
23	.378	.268	.017
24	.476	.011	.017
25	.493	.083	.066
26	.100	.002	.119
27	.302	.044	.222

ตาราง 41 (ต่อ)

ข้อ	องค์ประกอบที่ 1	องค์ประกอบที่ 2	องค์ประกอบที่ 3
28	.364	.029	.031
29	.388	.046	.025
30	.288	.207	.157
31	.453	.332	.139
32	.347	.093	.088
33	.236	.064	.034
34	.162	.246	.034
35	.431	.082	.198
36	.072	.045	.100
37	.164	.237	.011
38	.275	.106	.080
39	.430	.021	.020
40	-.011	.059	.106
41	.259	.004	.379
42	.463	.090	.356
43	.502	.106	.142
44	.254	.117	.257
45	.408	.006	.148
46	.248	.143	.073
47	.232	.059	.077
48	.410	.092	.081
49	.238	.133	.043
50	.367	.051	.022
51	.320	.014	.076
52	.398	.044	.022
53	.334	.123	.028
54	.512	.028	.013

ตาราง 41 (ต่อ)

ข้อ	องค์ประกอบที่ 1	องค์ประกอบที่ 2	องค์ประกอบที่ 3
55	.322	.006	.010
56	.043	.211	.225
57	.313	.051	.057
58	.051	.010	.001
59	.097	.083	.070
60	.124	.108	.037
61	.412	.225	.044
62	.378	.205	.118
63	.368	.226	.114
64	.253	.306	.123
65	.342	.173	.018
66	.227	.102	.003
67	.225	.008	.110
68	.100	.110	.002
69	.227	.008	.254
70	.485	.010	.170
71	.442	.169	.001
72	.361	.074	.152
73	.386	.351	.098
74	.120	.143	.124
75	.328	.161	.155
76	.498	.006	.241
77	.510	.001	.173
78	.346	.075	.158
79	.332	.008	.250
80	.390	.208	.103
81	.224	.058	.076

ตาราง 41 (ต่อ)

ข้อ	องค์ประกอบที่ 1	องค์ประกอบที่ 2	องค์ประกอบที่ 3
82	.378	.238	.020
83	.397	.218	.086
84	.100	.010	.001
85	.350	.220	.045
86	.360	.123	.029
87	.508	.106	.213
88	.506	.017	.104
89	.370	.130	.056
90	.407	.077	.041

ตาราง 42 น้ำหนักองค์ประกอบของแบบทดสอบชั้น ม.2

ข้อ	องค์ประกอบที่ 1	องค์ประกอบที่ 2	องค์ประกอบที่ 3
1	.263	.055	.135
2	.301	.147	.132
3	.179	.056	.217
4	.076	.068	.001
5	.292	.154	.145
6	.313	.052	.282
7	.325	.084	.111
8	.348	.254	.023
9	.229	.186	.183
10	.085	.009	.065
11	.201	.053	.167
12	.303	.000	.078

ตาราง 42 (ต่อ)

ข้อ	องค์ประกอบที่ 1	องค์ประกอบที่ 2	องค์ประกอบที่ 3
13	.300	.099	.073
14	.243	.173	.003
15	.336	.120	.099
16	.290	.178	.067
17	.239	.016	.240
18	.238	.229	.020
19	.210	.143	.091
20	.129	.052	.063
21	.244	.113	.059
22	.160	.002	.098
23	.308	.080	.180
24	.253	.055	.226
25	.120	.047	.158
26	.093	.079	.317
27	.193	.002	.007
28	.006	.136	.200
29	.182	.082	.190
30	.269	.095	.126
31	.445	.186	.151
32	.288	.203	.063
33	.299	.157	.150
34	.275	.241	.094
35	.405	.122	.143
36	.101	.189	.024
37	.362	.063	.233
38	.308	.224	.003
39	.459	.130	.149

ตาราง 42 (ต่อ)

ข้อ	องค์ประกอบที่ 1	องค์ประกอบที่ 2	องค์ประกอบที่ 3
40	.533	.180	.019
41	.195	.036	.118
42	.361	.313	.038
43	.370	.178	.156
44	.232	.171	.124
45	.433	.089	.066
46	.381	.159	.146
47	.395	.058	.100
48	.523	.090	.126
49	.376	.081	.054
50	.457	.239	.051
51	.434	.170	.110
52	.397	.100	.111
53	.165	.007	.099
54	.472	.143	.031
55	.393	.064	.114
56	.491	.275	.096
57	.136	.077	.046
58	.010	.002	.001
59	.460	.297	.293
60	.441	.301	.266
61	.414	.065	.051
62	.423	.105	.050
63	.374	.087	.081
64	.493	.262	.253
65	.463	.237	.023
66	.407	.225	.029

ตาราง 42 (ต่อ)

ข้อ	องค์ประกอบที่ 1	องค์ประกอบที่ 2	องค์ประกอบที่ 3
67	.397	.241	.022
68	.349	.318	.043
69	.022	.266	.111
70	.263	.178	.129
71	.287	.174	.087
72	.440	.280	.104
73	.221	.148	.020
74	.080	.134	.104
75	.292	.205	.063
76	.419	.025	.053
77	.324	.400	.099
78	.343	.134	.067
79	.335	.190	.094
80	.434	.028	.115
81	.045	.013	.009
82	.182	.021	.005
83	.064	.168	.169
84	.384	.132	.270
85	.295	.009	.145
86	.224	.137	.129
87	.118	.123	.096
88	.231	.048	.118
89	.039	.032	.170
90	.261	.032	.097

ตาราง 43 น้ำหนักองค์ประกอบของแบบทดสอบชั้น ม.3

ข้อ	องค์ประกอบที่ 1	องค์ประกอบที่ 2	องค์ประกอบที่ 3
1	.292	.163	.151
2	.291	.046	.170
3	.154	.002	.063
4	.134	.033	.013
5	.287	.076	.180
6	.320	.095	.008
7	.244	.134	.072
8	.028	.234	.108
9	.124	.089	.020
10	.354	.055	.048
11	.289	.476	.121
12	.282	.415	.051
13	.376	.417	.035
14	.305	.084	.070
15	.142	.131	.087
16	.262	.250	.090
17	.078	.072	.170
18	.285	.056	.082
20	.016	.035	.281
21	.206	.163	.222
22	.347	.009	.089
23	.298	.159	.065
24	.427	.013	.071
25	.020	.088	.242
26	.373	.114	.000
27	.340	.267	.006
28	.456	.327	.122

ตาราง 43 (ต่อ)

ข้อ	องค์ประกอบที่ 1	องค์ประกอบที่ 2	องค์ประกอบที่ 3
29	.309	.228	.117
30	.301	.167	.028
31	.327	.059	.020
32	.308	.247	.047
33	.369	.126	.029
34	.374	.061	.077
35	.069	.061	.110
36	.335	.020	.026
37	.233	.018	.170
38	.320	.079	.092
39	.306	.127	.216
40	.285	.166	.107
41	.286	.010	.281
42	.393	.160	.364
43	.295	.044	.237
44	.270	.019	.035
45	.323	.175	.104
46	.376	.210	.061
47	.332	.066	.064
48	.351	.077	.059
49	.144	.015	.047
50	.142	.057	.162
51	.266	.031	.155
52	.151	.002	.014
53	.183	.098	.008
54	.111	.077	.168
55	.289	.132	.070

ตาราง 43 (ต่อ)

ข้อ	องค์ประกอบที่ 1	องค์ประกอบที่ 2	องค์ประกอบที่ 3
56	.243	.065	.036
57	.423	.017	.064
58	.224	.087	.021
59	.444	.054	.091
60	.305	.225	.165
61	.381	.135	.198
62	.446	.150	.165
63	.348	.212	.010
64	.417	.164	.171
65	.452	.102	.305
66	.334	.060	.015
67	.203	.120	.116
68	.143	.048	.198
69	.372	.263	.056
70	.420	.286	.108
71	.256	.234	.055
72	.139	.039	.033
73	.266	.104	.122
74	.203	.144	.030
75	.442	.031	.009
76	.143	.116	.061
77	.263	.010	.238
78	.240	.202	.014
79	.245	.151	.128
80	.433	.058	.021
81	.008	.004	.201
82	.153	.031	.004

ตาราง 43 (ต่อ)

ข้อ	องค์ประกอบที่ 1	องค์ประกอบที่ 2	องค์ประกอบที่ 3
83	.103	.047	.095
84	.002	.054	.331
85	.347	.234	.066
86	.303	.007	.236
87	.209	.044	.002
88	.288	.143	.020
89	.435	.047	.236
90	.159	.223	.092

ประวัติของผู้วิจัย

ชื่อ นามสกุล	วรรณดี แสงประทีปทอง
ที่อยู่	223/232 หมู่ 4 ถนนแจ้งวัฒนะ ตำบลปากเกร็ด อำเภอปากเกร็ด จังหวัดนนทบุรี 11120
ตำแหน่งหน้าที่	อาจารย์ ระดับ 7
ที่ทำงาน	สำนักทะเบียนและวัดผล มหาวิทยาลัยสุโขทัยธรรมาธิราช
ประวัติการศึกษา	พ.ศ. 2513 ป.7 จากโรงเรียนสหราษฎร์วิทยา จ.สิงห์บุรี พ.ศ. 2516 ม.ศ.3 จากโรงเรียนอินทร์บุรี จ.สิงห์บุรี พ.ศ. 2518 ม.ศ.5 จากโรงเรียนสิงห์บุรี จ.สิงห์บุรี พ.ศ. 2522 กศ.บ.(คณิตศาสตร์) จากมหาวิทยาลัยศรีนครินทรวิโรฒ บางแสน พ.ศ. 2525 ค.ม.(สถิติการศึกษา) จากจุฬาลงกรณ์มหาวิทยาลัย พ.ศ. 2538 กศ.ด.(การทดสอบและวัดผลการศึกษา) จากมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร