

391.211

ก396ก

8.3

การศึกษาความเที่ยงตรงและประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ  
ในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน  
และแบบทดสอบแฟล็กชีลเวิล

10 พ.ย. 2537

ปริญญานิพนธ์  
ของ  
อรินทร์ น่วมถนนอม

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา  
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

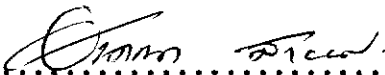
พฤษภาคม 2537

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

191460

คณะกรรมการควบคุมและคณะกรรมการสอบได้พิจารณาปริญญานิพนธ์ฉบับนี้แล้ว  
เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาการศึกษามหาบัณฑิต  
วิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒได้

คณะกรรมการควบคุม

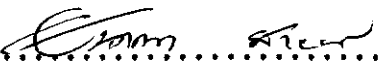
..........ประธาน

(รองศาสตราจารย์อังคณา สายยศ)

..........กรรมการ

(ผู้ช่วยศาสตราจารย์อาวุธ วัฒนสิน)

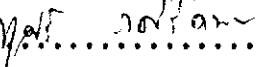
คณะกรรมการสอบ

..........ประธาน

(รองศาสตราจารย์อังคณา สายยศ)

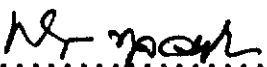
..........กรรมการ

(ผู้ช่วยศาสตราจารย์อาวุธ วัฒนสิน)

..........กรรมการที่แต่งตั้งเพิ่มเติม

(รองศาสตราจารย์ชูศรี วงศ์รัตนะ)

บัณฑิตวิทยาลัยอนุมัติให้รับปริญญานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร  
ปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒ

..........คณบดีบัณฑิตวิทยาลัย

(ดร.ศิริยุภา พูลสุวรรณ)

วันที่ .4. เดือน .พฤษภาคม. พ.ศ. 2537..

## ประกาศคุณูปการ

ในการทำปริญญานิพนธ์ฉบับนี้ ผู้วิจัยได้รับความกรุณาจากรองศาสตราจารย์อังคณา สายยศ และผู้ช่วยศาสตราจารย์อาวุธ วัฒนสิน กรรมการที่ปรึกษาปริญญานิพนธ์ ที่กรุณาให้ คำปรึกษา แนะนำ ตลอดจนข้อคิดเห็นต่าง ๆ ในการทำปริญญานิพนธ์มาโดยตลอดตั้งแต่ต้น จนสำเร็จลุล่วงไปด้วยดี ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง

ขอกราบขอบพระคุณ รองศาสตราจารย์ชูศรี วงศ์รัตนะ กรรมการสอบปริญญานิพนธ์ ที่กรุณาให้ความรู้และคำแนะนำเพิ่มเติม อันเป็นประโยชน์อย่างยิ่งสำหรับการทำปริญญานิพนธ์ ในครั้งนี้

ขอขอบพระคุณ อาจารย์ ดร.ชูศักดิ์ ขัมภลิขิต และอาจารย์รุ่งฤดี ทรัพย์นิธิ สำนักทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ ที่กรุณาให้ความอนุเคราะห์ ในการวิเคราะห์ข้อมูลด้วยเครื่องคอมพิวเตอร์

ขอขอบพระคุณ ผู้อำนวยการ ผู้ช่วยผู้อำนวยการฝ่ายวิชาการ และคณะครูของ โรงเรียนในกลุ่มตัวอย่าง และขอขอบใจนักเรียนทุกคนที่ให้ความร่วมมือในการเก็บข้อมูล เป็นอย่างดี

สุดท้ายนี้ผู้วิจัยขออภิวาถึงพระคุณของบิดา มารดา และครู อาจารย์ ที่ได้อบรม สั่งสอน ตลอดจนให้การสนับสนุนการศึกษาแก่ผู้วิจัยตลอดมา

อรินทร์ น่วมถนนอม

## สารบัญ

| บทที่                                   | หน้า |
|-----------------------------------------|------|
| 1 บทนำ .....                            | 1    |
| ภูมิหลัง .....                          | 1    |
| ความมุ่งหมายของการศึกษาค้นคว้า .....    | 5    |
| ความสำคัญของการศึกษาค้นคว้า .....       | 6    |
| ขอบเขตของการศึกษาค้นคว้า .....          | 6    |
| คำนิยามศัพท์เฉพาะ .....                 | 7    |
| 2 เอกสารและงานวิจัยที่เกี่ยวข้อง .....  | 12   |
| ทฤษฎีการตอบสนองข้อสอบ .....             | 12   |
| การประมาณค่าพารามิเตอร์ .....           | 25   |
| คุณภาพของแบบทดสอบ .....                 | 26   |
| ประสิทธิภาพของแบบทดสอบ .....            | 31   |
| การทดสอบแบบเทเลอร์ .....                | 33   |
| งานวิจัยที่เกี่ยวข้อง .....             | 44   |
| สมมติฐานของการศึกษาค้นคว้า .....        | 49   |
| 3 วิธีดำเนินการศึกษาค้นคว้า .....       | 50   |
| ประชากร .....                           | 50   |
| กลุ่มตัวอย่าง .....                     | 50   |
| เครื่องมือที่ใช้ในการศึกษาค้นคว้า ..... | 51   |
| วิธีดำเนินการสร้างแบบทดสอบ .....        | 52   |
| วิธีดำเนินการเก็บรวบรวมข้อมูล .....     | 64   |
| การวิเคราะห์ข้อมูลและสถิติที่ใช้ .....  | 64   |

| บทที่                                                | หน้า |
|------------------------------------------------------|------|
| 4 ผลการวิเคราะห์ข้อมูล .....                         | 69   |
| สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล ..... | 69   |
| การเสนอผลการวิเคราะห์ข้อมูล .....                    | 70   |
| ผลการวิเคราะห์ข้อมูล .....                           | 70   |
| 5 สรุป อภิปรายผล และข้อเสนอแนะ .....                 | 79   |
| ความมุ่งหมายของการศึกษาค้นคว้า .....                 | 79   |
| กลุ่มตัวอย่าง .....                                  | 79   |
| เครื่องมือที่ใช้ในการศึกษาค้นคว้า .....              | 79   |
| วิธีดำเนินการเก็บรวบรวมข้อมูล .....                  | 80   |
| การเสนอผลการวิเคราะห์ข้อมูล .....                    | 80   |
| สรุปผลการวิเคราะห์ข้อมูล .....                       | 80   |
| อภิปรายผล .....                                      | 81   |
| ข้อเสนอแนะ .....                                     | 87   |
| บรรณานุกรม .....                                     | 88   |
| ภาคผนวก .....                                        | 93   |
| ประวัติย่อของผู้วิจัย .....                          | 140  |

## บัญชีตาราง

| ตาราง |                                                                                                                                                                                   | หน้า |
|-------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| 1     | จำนวนกลุ่มตัวอย่างที่ใช้ในการวิจัย จำนวนตามเขต โรงเรียน<br>จำนวนห้องเรียน และจำนวนนักเรียน .....                                                                                  | 51   |
| 2     | ผลการคัดเลือกข้อสอบที่มีคุณภาพ ของแบบทดสอบ 6 ฉบับ .....                                                                                                                           | 57   |
| 3     | ค่าความเที่ยงตรงและผลการทดสอบความแตกต่างของค่าความเที่ยงตรง<br>ของการประมาณค่าความสามารถของผู้สอบ จากการทำแบบทดสอบที่<br>แตกต่างกัน 3 รูปแบบ .....                                | 71   |
| 4     | ผลการทดสอบความแตกต่างของค่าความเที่ยงตรงของการประมาณค่า<br>ความสามารถของผู้สอบ จากการทำแบบทดสอบที่แตกต่างกัน 3 รูปแบบ<br>โดยทดสอบเป็นรายคู่ .....                                 | 72   |
| 5     | ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ<br>ระหว่างแบบทดสอบที่แตกต่างกัน 3 รูปแบบ ที่ระดับความสามารถ<br>ตั้งแต่ -4.00 ถึง 4.00 จำนวน 81 ระดับความสามารถ .....     | 73   |
| 6     | ค่าไอเกน (มากกว่า 1) เบอร์เช็นต์ของความแปรปรวน และเบอร์เช็นต์<br>ของความแปรปรวนสะสม จากการนำแบบทดสอบ จำนวน 6 ฉบับ<br>มาวิเคราะห์องค์ประกอบ .....                                  | 95   |
| 7     | ผลการประมาณค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าความ<br>คลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ (SEE)<br>และค่าความสอดคล้องกับไคเดล (t-test) ของแบบทดสอบฉบับที่ 1 ..... | 100  |
| 8     | ผลการประมาณค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าความ<br>คลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ (SEE)<br>และค่าความสอดคล้องกับไคเดล (t-test) ของแบบทดสอบฉบับที่ 2 ..... | 101  |

|    |                                                                                                                                                                                                            |     |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 9  | ผลการประมาณค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ (SEE) และค่าความสอดคล้องกับไคสแควร์ (t-test) ของแบบทดสอบฉบับที่ 3 ..                                 | 102 |
| 10 | ผลการประมาณค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ (SEE) และค่าความสอดคล้องกับไคสแควร์ (t-test) ของแบบทดสอบฉบับที่ 4 ..                                 | 103 |
| 11 | ผลการประมาณค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ (SEE) และค่าความสอดคล้องกับไคสแควร์ (t-test) ของแบบทดสอบฉบับที่ 5 ..                                 | 104 |
| 12 | ผลการประมาณค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ (SEE) และค่าความสอดคล้องกับไคสแควร์ (t-test) ของแบบทดสอบฉบับที่ 6 ..                                 | 105 |
| 13 | เกรดเฉลี่ยสะสม (G.P.A.) กับผลการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) ระหว่างแบบทดสอบดั้งเดิม (CONT) แบบทดสอบสองขั้นตอน (TWOT) และแบบทดสอบแฟล็กซีเบิลเวล (FLEXT) .....                               | 107 |
| 14 | ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ (TIF) ระหว่างแบบทดสอบดั้งเดิม (CONT) แบบทดสอบสองขั้นตอน (TWOT) และแบบทดสอบแฟล็กซีเบิลเวล (FLEXT) .....                                                                  | 120 |
| 15 | ค่าประสิทธิภาพสัมพัทธ์ (RE) ของการประมาณค่าความสามารถของผู้สอบ ระหว่างแบบทดสอบดั้งเดิม (CONT) แบบทดสอบสองขั้นตอน (TWOT) และแบบทดสอบแฟล็กซีเบิลเวล (FLEXT) .....                                            | 133 |
| 16 | รายละเอียดจุดประสงค์การเรียนรู้ เนื้อหา และจุดประสงค์เชิงพฤติกรรม วิชาคณิตศาสตร์ (ค 011) ชั้นมัธยมศึกษาปีที่ 3 เรื่องเลขยกกำลัง ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) ..... | 138 |

## บัญชีภาพประกอบ

| ภาพประกอบ                                                                                                                    | หน้า |
|------------------------------------------------------------------------------------------------------------------------------|------|
| 1 ไค้ลักษณะข้อสอบของ นอร์มัล ออใจฟ์ ไมเคิล ที่มีพารามิเตอร์ 3 ตัว ...                                                        | 19   |
| 2 ไค้ลักษณะข้อสอบของ ราล์ฟ ไมเคิล .....                                                                                      | 23   |
| 3 ไค้อินฟอร์เมชันของข้อสอบ 5 ข้อ และ ไค้อินฟอร์เมชันของแบบทดสอบ ..                                                           | 29   |
| 4 ไค้ประสิทธิภาพสัมพัทธ์ของแบบทดสอบค้ำศัพท์ STEP เทียบกับแบบทดสอบ<br>ค้ำศัพท์ MAT .....                                      | 32   |
| 5 โครงสร้างของแบบทดสอบสองขั้นตอน ที่มีแบบทดสอบแยกกลุ่มมีค่าความยาก<br>ใกล้เคียงกัน .....                                     | 38   |
| 6 โครงสร้างของแบบทดสอบสองขั้นตอน ที่มีแบบทดสอบแยกกลุ่มมีค่าความยาก<br>กระจายในช่วงกว้าง .....                                | 39   |
| 7 โครงสร้างของแบบทดสอบสองขั้นตอน ที่มีแบบทดสอบแยกกลุ่มสองครั้ง<br>หรือแบบทดสอบสามขั้นตอน .....                               | 40   |
| 8 แสดงตัวอย่างข้อสอบ ในแบบทดสอบเพิลิกซ์เลเวล .....                                                                           | 41   |
| 9 แผนผังแสดงขั้นตอนการดำเนินการสร้างแบบทดสอบ .....                                                                           | 52   |
| 10 โครงสร้างของแบบทดสอบสองขั้นตอน .....                                                                                      | 60   |
| 11 โครงสร้างของแบบทดสอบเพิลิกซ์เลเวล .....                                                                                   | 61   |
| 12 กระดาษคำตอบแผ่นบนของแบบทดสอบเพิลิกซ์เลเวล .....                                                                           | 62   |
| 13 กระดาษคำตอบแผ่นล่างของแบบทดสอบเพิลิกซ์เลเวล .....                                                                         | 63   |
| 14 ไค้ประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ<br>ระหว่างแบบทดสอบดั้งเดิม เทียบกับแบบทดสอบสองขั้นตอน .....      | 76   |
| 15 ไค้ประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ<br>ระหว่างแบบทดสอบดั้งเดิม เทียบกับแบบทดสอบเพิลิกซ์เลเวล .....   | 77   |
| 16 ไค้ประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ<br>ระหว่างแบบทดสอบสองขั้นตอน เทียบกับแบบทดสอบเพิลิกซ์เลเวล ..... | 78   |

|    |                                                                                          |     |
|----|------------------------------------------------------------------------------------------|-----|
| 17 | โค้งอินฟอร์เมชันของแบบทดสอบดั้งเดิม .....                                                | 124 |
| 18 | โค้งอินฟอร์เมชันของแบบทดสอบสองขั้นตอน (แบบทดสอบวัดผล ฉบับที่ 1) ..                       | 125 |
| 19 | โค้งอินฟอร์เมชันของแบบทดสอบสองขั้นตอน (แบบทดสอบวัดผล ฉบับที่ 2) ..                       | 126 |
| 20 | โค้งอินฟอร์เมชันของแบบทดสอบสองขั้นตอน (แบบทดสอบวัดผล ฉบับที่ 3) ..                       | 127 |
| 21 | โค้งอินฟอร์เมชันของแบบทดสอบสองขั้นตอน (แบบทดสอบวัดผล ฉบับที่ 4) ..                       | 128 |
| 22 | โค้งอินฟอร์เมชันของแบบทดสอบสองขั้นตอน (เฉลี่ยแบบทดสอบวัดผลทั้ง<br>4 ฉบับ) .....          | 129 |
| 23 | โค้งอินฟอร์เมชันของแบบทดสอบแฟล็กชิเลเวล .....                                            | 130 |
| 24 | โค้งอินฟอร์เมชันของแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และ<br>แบบทดสอบแฟล็กชิเลเวล ..... | 131 |

## บทที่ 1

### บทนำ

#### บทนำ

การวัดผลการศึกษาและจิตวิทยาเป็นการวัดความสามารถหรือคุณลักษณะแฝง (Ability or Latent Trait) ของบุคคล ซึ่งไม่สามารถวัดได้โดยตรงเหมือนกับการวัดทางกายภาพ แต่ต้องอาศัยการวัดทางอ้อมโดยใช้แบบทดสอบเป็นเครื่องมือในการวัด แล้วนำผลการสอบหรือคะแนน (Performance or Score) ไปอธิบายหรือทำนายความสามารถที่แท้จริงของบุคคลที่ถูกวัด (Lord, 1980 : 3) การวัดหรือการประมาณค่าความสามารถโดยอาศัยการวัดทางอ้อมนี้มักจะมีผลคลาดเคลื่อนค่อนข้างสูง ซึ่งจะทำให้ผลสรุปที่ได้ขาดความถูกต้อง ดังนั้นนักวัดผลจึงได้สร้างทฤษฎีการวัดขึ้นมา เพื่อที่จะได้สรุปอ้างอิงผลของการวัด ซึ่งเป็นตัวแทนของความสามารถหรือคุณลักษณะแฝงของบุคคล ให้มีความถูกต้องแม่นยำและเชื่อถือได้มากที่สุด ทฤษฎีการวัดที่สำคัญได้แก่ ทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory : CTT) และทฤษฎีการตอบสนองข้อสอบ (Item Response Theory : IRT) เป็นต้น

ทฤษฎีการทดสอบแบบดั้งเดิมเป็นทฤษฎีการวัดที่ใช้กันอย่างแพร่หลายมาตั้งแต่ในอดีตจนถึงปัจจุบัน ซึ่งเน้นการประมาณค่าคะแนนจริง หรือความสามารถที่แท้จริง โดยใช้คะแนนของกลุ่มผู้สอบมาประมาณค่า ทฤษฎีการวัดแบบนี้ได้รับการวิจารณ์ว่าเป็นทฤษฎีที่ตั้งอยู่บนข้อตกลงเบื้องต้นที่อ่อน (Lord, 1980 : 9) โดยเฉพาะในเรื่องค่าสถิติของข้อสอบ ซึ่งได้แก่ค่าความยากของข้อสอบ และค่าอำนาจจำแนกของข้อสอบ จะแปรเปลี่ยนไปตามกลุ่มตัวอย่างที่มีระดับความสามารถแตกต่างกัน และการเปรียบเทียบความสามารถของผู้สอบจะต้องใช้ผลการสอบจากแบบทดสอบฉบับเดียวกัน หรือแบบทดสอบคู่ขนานกันเท่านั้น ถ้าใช้แบบทดสอบหลายฉบับที่วัดความสามารถเดียวกันแต่ความยากง่ายแตกต่างกัน แล้วจะนำคะแนนจากแบบทดสอบคนละฉบับมาเปรียบเทียบกันไม่ได้ (ผจงจิต อินทสุวรรณ, 2525 : 53 ; สำเร็จ บุญเรืองรัตน์, 2525 : 1 - 2 ; Hambleton and Swaminathan, 1985 : 1 - 2) นอกจากนี้แล้วทฤษฎีการทดสอบแบบดั้งเดิมยังไม่สามารถพัฒนาแบบทดสอบให้มีความเหมาะสมกับระดับความสามารถของผู้สอบ ดังนั้นนักวัดผลจึงได้พัฒนาทฤษฎีการวัดแบบอื่น ๆ ที่เหมาะสมกว่า

เพื่อที่จะได้นำไปใช้ในการวัด และแก้ปัญหาข้อบกพร่องดังกล่าว และทฤษฎีใหม่ที่น่าสนใจคือ ทฤษฎีการตอบสนองข้อสอบ ซึ่งทฤษฎีนี้มีชื่อเรียกต่าง ๆ กัน เช่น ทฤษฎีคุณลักษณะแฝง (Latent Trait Theory) ทฤษฎีโค้งลักษณะข้อสอบ (Item Characteristic Curve Theory) หรือทฤษฎีการทดสอบแบบใหม่ (Modern Test Theory) (Hambleton and others. 1978 : 468) เป็นต้น

ทฤษฎีการตอบสนองข้อสอบเป็นทฤษฎีการวัดที่เชื่อกันว่าสามารถแก้จุดอ่อน หรือ ข้อบกพร่องต่าง ๆ ของทฤษฎีการทดสอบแบบดั้งเดิมได้ เนื่องจากทฤษฎีนี้ได้ใช้โมเดลทาง คณิตศาสตร์แสดงความสัมพันธ์ระหว่างผลการสอบ กับความสามารถของผู้สอบ ซึ่งจะทำให้ค่า พารามิเตอร์ของข้อสอบ ได้แก่ ค่าความยาก ค่าอำนาจจำแนก และค่าการเดา ไม่แปรเปลี่ยน ไปตามกลุ่มตัวอย่างของผู้สอบ ถึงแม้ว่าผู้สอบจะใช้แบบทดสอบคนละฉบับกันก็สามารถเปรียบเทียบ ความสามารถกันได้ ทั้งยังสามารถหาความแม่นยำในการประมาณค่าความสามารถของผู้สอบได้ อีกด้วย (Hambleton and Swaminathan. 1985 : 11)

จากคุณสมบัติของทฤษฎีการตอบสนองข้อสอบ จึงทำให้นักวิจัยนำไปประยุกต์ใช้ในการ วิเคราะห์ผลการศึกษาและจิตวิทยาหลาย ๆ อย่าง เช่น การศึกษาประสิทธิภาพสัมพัทธ์ (Relative Efficiency) การสร้างคลังข้อสอบ (Item Bank) การศึกษาความลำเอียง ของข้อสอบ (Item Bias) การเปรียบเทียบคะแนนการสอบ (Test Score Equating) การทดสอบแบบอิงเกณฑ์ (Criterion - Reference Testing) และการทดสอบแบบ เทเลอร์ (Tailored Testing) (Hambleton and others. 1978 : 492 - 499 ; Lord. 1980 : 83 - 231) เป็นต้น

การทดสอบแบบเทเลอร์ หรือมีชื่อที่เรียกแตกต่างกันไป เช่น การทดสอบที่ปรับตาม ระดับความสามารถ (Adaptive Testing) การทดสอบโดยใช้คอมพิวเตอร์ (Computerized Testing) การทดสอบเป็นรายบุคคล (Individualized Testing) การทดสอบแบบโปรแกรม (Programmed Testing) และการทดสอบที่มีการจัดเรียงลำดับ ข้อสอบ (Sequential Item Testing) (Lord. 1980 : 150 - 151) เป็นต้น การทดสอบในรูปแบบนี้จะจัดข้อสอบให้มีค่าความยากเหมาะสมกับระดับความสามารถของผู้สอบ เป็นรายบุคคล (Lord. 1971b : 3 ; Hambleton and Swaminathan. 1985 : 296 - 297) ซึ่งเป็นการทดสอบที่แตกต่างไปจากการทดสอบแบบดั้งเดิม (Conventional

Paper and Pencil Testing) เนื่องจากการทดสอบแบบดั้งเดิมจะประกอบด้วยกลุ่มข้อสอบที่มีจำนวนข้อ และการจัดเรียงลำดับไว้แน่นอน โดยผู้สอบจะต้องทำข้อสอบทุก ๆ ข้อเหมือนกันทุกคน สำหรับข้อสอบที่จัดไว้เช่นจะเป็นข้อสอบที่มีค่าความยากเหมาะสมกับกลุ่มผู้สอบที่มีความสามารถในระดับปานกลางเท่านั้น แต่จะ ไม่เหมาะสมกับกลุ่มผู้สอบที่มีความสามารถในระดับสูงหรือต่ำ (ฉัตรนภา พรหมมา. 2525 : 45 ; พงษ์จิต อินทสุวรรณ. 2525 : 53) ดังนั้นผู้สอบที่มีความสามารถในระดับสูงหรือต่ำกว่าระดับปานกลางของกลุ่มมาก ๆ จะให้รายละเอียดเกี่ยวกับตัวผู้สอบต่ำ (Vale and Weiss. 1975 : 6) กล่าวคือ ถ้าผู้สอบได้ตอบข้อสอบที่มีค่าความยากสูงกว่าระดับความสามารถของผู้สอบ แล้วผู้สอบจะมีโอกาสตอบข้อสอบด้วยการเดา หรือถ้าผู้สอบได้ตอบข้อสอบที่มีค่าความยากต่ำกว่าระดับความสามารถของผู้สอบ แล้วข้อสอบข้อนั้นจะ ไม่ช่วยให้ผู้สอบได้ใช้ความสามารถในระดับสูงสุด ซึ่งมีผลทำให้ค่าความเชื่อมั่นของคะแนนจากการสอบต่ำกว่าที่ควรจะเป็น (Weiss. 1974 : 1) ปัญหาเหล่านี้สามารถแก้ไขได้โดยใช้การทดสอบแบบเทเลอร์ เนื่องจากผู้สอบจะได้ตอบข้อสอบที่มีค่าความยากเหมาะสมกับระดับความสามารถ ซึ่งจะทำให้สามารถวัดสมรรถภาพที่แท้จริงของผู้สอบได้ (สุธรรม จันทร์หอม. 2529 : 10)

การทดสอบแบบเทเลอร์ มีหลักการสำคัญคือผู้สอบไม่จำเป็นต้องทำข้อสอบเหมือน ๆ กันทุกข้อ จำนวนข้อก็ไม่จำเป็นต้องเท่ากัน ทั้งนี้ขึ้นอยู่กับรูปแบบและผลการตอบข้อสอบในแบบทดสอบของแต่ละคน (Weiss and Kingsbury. 1984 : 361) โดยทั่วไปแล้วจะให้ผู้สอบเริ่มทำข้อสอบข้อที่มีค่าความยากในระดับปานกลางก่อน ถ้าผู้สอบตอบข้อสอบข้อแรกถูกข้อต่อไปจะ ได้ทำข้อที่มีค่าความยากมากขึ้น แต่ถ้าผู้สอบตอบข้อสอบข้อนั้นผิดข้อต่อไปก็จะ ได้ทำข้อที่ง่ายขึ้น และการทำข้อสอบข้อต่อไปก็จะดำเนินการในลักษณะนี้ไปจนสิ้นสุดการสอบ (Green and others. 1984 : 347) สำหรับการคัดเลือกข้อสอบที่จะให้ผู้สอบทำในแต่ละข้อนั้น จะเป็นไปตามรูปแบบและเงื่อนไขของการทดสอบแบบเทเลอร์นั้น ๆ ซึ่งแบ่งออกเป็น 2 ยุทธวิธีใหญ่ ๆ (Weiss. 1974 : 3) คือ ยุทธวิธีสองขั้นตอน (Two - Stage Strategies) และยุทธวิธีหลายขั้นตอน (Multi-Stage Strategies) โดยที่ยุทธวิธีหลายขั้นตอนแบ่งออกเป็น 2 รูปแบบ คือ รูปแบบแยกทางคงที่ (Fixed - Branching Model) และรูปแบบแยกทางแปรผัน (Variable - Branching Model) สำหรับรูปแบบแยกทางคงที่ยังแบ่งออกเป็น 3 รูปแบบย่อย ๆ คือ รูปแบบปิรามิด (Pyramidal Model)

แบบทดสอบแฟล็กซีเลเวล (Flexilevel Test) และแบบทดสอบปรับระดับขึ้น (Stradaptive Test) ส่วนรูปแบบแยกทางแปรผันแบ่งออกเป็น 2 ยุทธวิธีย่อย ๆ คือ ยุทธวิธีของเบส์ (Bayesian Strategies) และยุทธวิธีความเป็นไปได้สูงสุด (Maximum Likelihood Strategies)

○ นักวัดผลการศึกษาและจิตวิทยาได้นำการทดสอบแบบเทเลอร์มาศึกษาเปรียบเทียบประสิทธิภาพกับการทดสอบแบบดั้งเดิม ผลการศึกษาพบว่า การทดสอบแบบเทเลอร์จะมีความแม่นยำในการประมาณค่าความสามารถสูงกว่าการทดสอบแบบดั้งเดิม เมื่อผู้สอบมีความสามารถระดับสูงหรือระดับต่ำ ส่วนผู้สอบที่มีความสามารถระดับปานกลาง การทดสอบแบบดั้งเดิมจะมีความแม่นยำในการประมาณค่าความสามารถสูงกว่าการทดสอบแบบเทเลอร์ (Lord. 1971a : 805 - 813 ; Vale. 1975 : 5 - 57 ; Sequin. 1978 : 1196 - A) สำหรับข้อได้เปรียบของการทดสอบแบบเทเลอร์ก็คือ สามารถลดค่าใช้จ่ายในการดำเนินการสอบได้ เนื่องจากใช้จำนวนข้อสอบในแบบทดสอบน้อยกว่าการทดสอบแบบดั้งเดิม ซึ่งจะทำให้ลดเวลาในการดำเนินการสอบ (Linn, Rock and Cleary. 1969 : 145 ; Waters and Bayroff. 1971 : 125 ; Urry. 1977 : 194) โดยสามารถลดเวลาได้มากกว่าร้อยละ 50 ของเวลาในการทดสอบแบบดั้งเดิม (Spinetti and Hambleton. 1977 : 156) ในขณะที่ความเชื่อมั่นและความเที่ยงตรงคงเดิม (Larkin and Weiss. 1975 : 1 ; citing Betz and Weiss. 1973) และจากการศึกษาของไวส์ (Weiss. 1982 : 473 - 492) โมรีโนและคนอื่น ๆ (Moreno and others. 1984 : 155 - 163) พบว่า การทดสอบแบบเทเลอร์จะมีความเชื่อมั่นและความเที่ยงตรงคงเดิม หรือมากกว่า ทั้ง ๆ ที่ใช้จำนวนข้อสอบน้อยกว่าประมาณครึ่งหนึ่งของการทดสอบแบบดั้งเดิม นอกจากนี้แล้วการทดสอบแบบเทเลอร์จะทำให้ผู้สอบไม่เกิดความเบื่อหน่ายในการทำข้อสอบเพราะมีการปรับความยากของข้อสอบให้มีความเหมาะสมกับระดับความสามารถของผู้สอบอยู่ตลอดเวลา ซึ่งจะส่งผลให้ผู้สอบมีระดับการสนใจในการสอบสูงขึ้น นั่นคือ ผู้สอบมีทัศนคติที่ดีต่อการทดสอบแบบเทเลอร์ (งามนิตย์ ธาตุทอง และปรีชา เครือวรรณ. 2527 : 92 ; Schmidt, Urry and Gugel. 1978 : 272)

ในปัจจุบันได้มีการนำเอาเครื่องคอมพิวเตอร์มาประยุกต์ใช้กับการทดสอบแบบเทเลอร์ แทนการทดสอบที่ใช้การเขียนตอบแบบเดิม (Paper and Pencil Test) ซึ่ง

เครื่องคอมพิวเตอร์จะประมาณค่าความสามารถของผู้สอบ แล้วจัดข้อสอบให้เหมาะสมกับระดับความสามารถของผู้สอบ (Weiss and Kingsbury, 1984 : 361 - 375) พร้อมทั้งนำทฤษฎีการตอบสนองข้อสอบมาประยุกต์ใช้ในการประมาณค่าความสามารถของผู้สอบ จึงทำให้การทดสอบแบบเทเลอร์มีประสิทธิภาพมากยิ่งขึ้น (Rudner, 1983 : 951)

อย่างไรก็ตาม เนื่องจากการทดสอบในประเทศไทยยังคงใช้การทดสอบแบบดั้งเดิม ส่วนการทดสอบแบบเทเลอร์ ที่ใช้เครื่องพีซีปรับค่าความยากของข้อสอบให้เหมาะสมกับระดับความสามารถของผู้สอบ ซึ่งเป็นที่ยอมรับและใช้กันอย่างกว้างขวางในต่างประเทศ เช่น ประเทศสหรัฐอเมริกา โอกาสที่จะนำมาใช้ในประเทศไทยในปัจจุบันคงเป็นเรื่องที่ยาก แต่แบบทดสอบสองขั้นตอน (Two-Stage Test) และแบบทดสอบเฟลิกซ์เลเวล (Flexilevel Test) เป็นรูปแบบอย่างง่ายของการทดสอบแบบเทเลอร์ ซึ่งสามารถนำมาทดสอบโดยนำมาประยุกต์ใช้กับการทดสอบเขียนตอบแบบเดิมได้ ดังนั้นผู้วิจัยจึงมีความสนใจที่จะนำแบบทดสอบทั้งสองรูปแบบนี้มาทำการศึกษาร่วมกับแบบทดสอบดั้งเดิม (Conventional Test) ที่ใช้กันอยู่ทั่ว ๆ ไปในปัจจุบัน เพื่อที่จะได้ทราบว่าแบบทดสอบสองขั้นตอนและแบบทดสอบเฟลิกซ์เลเวลสามารถนำมาใช้แทนแบบทดสอบดั้งเดิมได้หรือไม่ โดยศึกษามุ่งเน้นในเรื่องความเที่ยงตรงและประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบที่แตกต่างกัน 3 รูปแบบ คือ แบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบเฟลิกซ์เลเวล สำหรับการศึกษาครั้งนี้ผู้วิจัยได้ศึกษาในสถานการณ์การทดสอบจริง โดยนำผลการทดสอบไปประมาณค่าพารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบ แบบราส์ช โมเดล (Rasch Model) ด้วยวิธีการประมาณค่าความเป็นไปได้สูงสุดแบบไม่มีเงื่อนไข (Unconditional Maximum Likelihood Procedure) โดยใช้โปรแกรมไบคาล (BICAL)

#### ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อศึกษาค่าความเที่ยงตรงของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมกับแบบทดสอบสองขั้นตอน แบบทดสอบดั้งเดิมกับแบบทดสอบเฟลิกซ์เลเวล และแบบทดสอบสองขั้นตอนกับแบบทดสอบเฟลิกซ์เลเวล

2. เพื่อศึกษาค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบสองขั้นตอน แบบทดสอบดั้งเดิมเทียบกับแบบทดสอบเพลิกซ์เลเวล และแบบทดสอบสองขั้นตอนเทียบกับแบบทดสอบเพลิกซ์เลเวล

#### ความสำคัญของการศึกษาค้นคว้า

ผลการศึกษาค้นคว้าครั้งนี้ ทำให้ทราบถึงความเที่ยงตรงและประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบที่แตกต่างกัน 3 รูปแบบ คือ แบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบเพลิกซ์เลเวล ซึ่งจะ เป็นประโยชน์ในการพิจารณาเลือกรูปแบบของแบบทดสอบเลือกตอบให้มีความเหมาะสมกับระดับความสามารถของผู้สอบ ทั้งยังช่วยให้สามารถตัดสินใจเลือกให้แบบทดสอบเลือกตอบที่มีคุณภาพและประสิทธิภาพสูงในการวัดผลสัมฤทธิ์ทางการเรียน เพื่อปรับปรุงการเรียนการสอน หรือใช้ในการสอนซ่อมเสริมนักเรียนที่มีระดับความสามารถแตกต่างกัน นอกจากนี้ยังเป็นแนวทางในการนำทฤษฎีการตอบสนองข้อสอบมาประยุกต์ใช้กับการทดสอบแบบเทเลอร์ เพื่อพัฒนาแบบทดสอบ และวิธีการดำเนินการสอบให้มีประสิทธิภาพต่อไป

#### ขอบเขตของการศึกษาค้นคว้า

1. ประชากรที่ใช้ในการศึกษาค้นคว้านี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2536 ของโรงเรียนสังกัดกรมสามัญศึกษา ส่วนกลาง กลุ่มที่ 5 กรุงเทพมหานคร จำนวนโรงเรียน 15 โรงเรียน จำนวนห้องเรียน 204 ห้องเรียน จำนวนนักเรียน 10,390 คน
2. กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้านี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2536 ของโรงเรียนสังกัดกรมสามัญศึกษา ส่วนกลาง กลุ่มที่ 5 กรุงเทพมหานคร จำนวนโรงเรียน 9 โรงเรียน จำนวนห้องเรียน 9 ห้องเรียน จำนวนนักเรียน 441 คน ซึ่งได้มาจากวิธีการสุ่มแบบแบ่งชั้น (Stratified Random Sampling)
3. ตัวแปรที่ใช้ในการศึกษา
  - 3.1 ตัวแปรอิสระ ได้แก่ รูปแบบของแบบทดสอบ 3 รูปแบบ ดังนี้

3.1.1 แบบทดสอบดั้งเดิม

3.1.2 แบบทดสอบสองขั้นตอน

3.1.3 แบบทดสอบแฟล็กซีเลเวล

3.2 ตัวแปรตาม ได้แก่

3.2.1 ค่าความเที่ยงตรงของการประมาณค่าความสามารถ

3.2.2 ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ

4. เนื้อหาที่ใช้ในการสร้างแบบทดสอบครั้งนี้ เป็นเนื้อหาวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3 ( ค 011) เรื่องเลขยกกำลัง ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533)

5. การศึกษาครั้งนี้ผู้วิจัยได้ศึกษาในสถานการณ์การทดสอบจริง ซึ่งผู้สอบจะทำแบบทดสอบโดยใช้การเขียนตอบแบบเดิม แล้วนำผลการสอบไปประมาณค่าความสามารถของผู้สอบตามทฤษฎีการตอบสนองข้อสอบ แบบราสช์ โมเดล ด้วยวิธีการประมาณค่าความเป็นไปได้สูงสุดแบบไม่มีเงื่อนไข โดยใช้โปรแกรมไบคาล

#### คานิยามศัพท์เฉพาะ

1. แบบทดสอบดั้งเดิม (Conventional Test) หมายถึงแบบทดสอบที่ประกอบด้วยข้อสอบชนิดเลือกตอบ ที่มี 5 ตัวเลือก ซึ่งมีตัวเลือกที่เป็นคำตอบถูกเพียงคำตอบเดียว และมีการจัดเรียงข้อสอบตามลำดับเนื้อหาวิชา โดยในแต่ละเนื้อหาจะเรียงข้อสอบจากง่ายไปหายาก สำหรับจำนวนข้อสอบในฉบับนี้ทั้งหมด 30 ข้อ

2. แบบทดสอบสองขั้นตอน (Two - Stage Test) หมายถึงแบบทดสอบที่ประยุกต์มาจากแบบทดสอบดั้งเดิม แต่มีการจัดเรียงข้อสอบโดยยึดค่าความยากเป็นหลัก แบบทดสอบในรูปแบบนี้จะประกอบด้วยแบบทดสอบแยกกลุ่ม และแบบทดสอบวัดผล ซึ่งมีรายละเอียดดังนี้

2.1 แบบทดสอบแยกกลุ่ม (Routing Test) เป็นแบบทดสอบขั้นแรกให้ผู้สอบทุกคนจะต้องสอบเหมือนกัน โดยสอบเพื่อที่จะจัดแยกระดับความสามารถของผู้สอบว่าผู้สอบคนไหนจะได้ไปสอบแบบทดสอบวัดผลฉบับใดในขั้นที่ 2 ซึ่งจะยึดจำนวนข้อที่ตอบถูกเป็นเกณฑ์ในการแยก สำหรับแบบทดสอบฉบับนี้ประกอบด้วยข้อสอบที่มีค่าความยากในช่วงระหว่าง  $-0.50$  ถึง  $0.50$  ซึ่งจัดเป็นข้อสอบที่มีค่าความยากใกล้เคียงกัน (Peaked) จำนวนข้อสอบในฉบับนี้ทั้งหมด 10 ข้อ

2.2 แบบทดสอบวัดผล (Measurement Test) เป็นแบบทดสอบขั้นที่ 2 ที่ผู้สอบจะต้องเลือกทำตามระดับความสามารถของผู้สอบ โดยมีแบบทดสอบฉบับย่อย ๆ ให้เลือกจำนวน 4 ฉบับ ๆ ละ 30 ข้อ สำหรับแบบทดสอบทั้ง 4 ฉบับนี้จะจัดแยกตามค่าความยากของข้อสอบตั้งแต่ง่ายที่สุดจนถึงยากที่สุด ซึ่งมีรายละเอียดดังนี้

2.2.1 ฉบับที่ 1 มีค่าความยากอยู่ในช่วงระหว่าง -3.00 ถึง -1.50 สำหรับผู้ที่ทำแบบทดสอบแยกกลุ่มถูก จำนวน 0 - 3 ข้อ

2.2.2 ฉบับที่ 2 มีค่าความยากอยู่ในช่วงระหว่าง -1.50 ถึง 0.00 สำหรับผู้ที่ทำแบบทดสอบแยกกลุ่มถูก จำนวน 4 - 5 ข้อ

2.2.3 ฉบับที่ 3 มีค่าความยากอยู่ในช่วงระหว่าง 0.00 ถึง 1.50 สำหรับผู้ที่ทำแบบทดสอบแยกกลุ่มถูก จำนวน 6 - 7 ข้อ

2.2.4 ฉบับที่ 4 มีค่าความยากอยู่ในช่วงระหว่าง 1.50 ถึง 3.00 สำหรับผู้ที่ทำแบบทดสอบแยกกลุ่มถูก จำนวน 8 - 10 ข้อ

3. แบบทดสอบเฟล็กซิเลเวล (Flexilevel Test) หมายถึงแบบทดสอบที่ประยุกต์มาจากแบบทดสอบดั้งเดิม โดยแบ่งข้อสอบออกเป็น 2 กลุ่ม ๆ ละ 29 ข้อ กลุ่มแรกเป็นข้อสอบกลุ่มง่าย มีค่าความยากอยู่ในช่วงระหว่าง -3.00 ถึง 0.00 ซึ่งจะเรียงลำดับข้อสอบจากข้อที่ยากที่สุดไปหาข้อที่ง่ายที่สุด กลุ่มที่สองเป็นข้อสอบกลุ่มยากมีค่าความยากอยู่ในช่วงระหว่าง 0.00 ถึง 3.00 ซึ่งจะเรียงลำดับข้อสอบจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด แล้วคัดเลือกข้อสอบที่มีค่าความยากตรงกึ่งกลางระหว่างข้อสอบกลุ่มง่าย และข้อสอบกลุ่มยากไว้อีก 1 ข้อ ซึ่งมีค่าความยากเท่ากับ 0.00 โดยข้อสอบข้อนี้จะเป็นข้อเริ่มต้นสำหรับผู้สอบทุกคน เมื่อผู้สอบตอบข้อสอบข้อนี้ได้ถูกต้อง ข้อต่อไปจะได้ทำข้อสอบที่มีความยากมากขึ้น แต่ถ้าผู้สอบตอบข้อสอบข้อนั้นผิด ข้อต่อไปก็จะได้ทำข้อสอบที่ง่ายขึ้น สำหรับข้อต่อ ๆ ไปผู้สอบจะได้ทำแบบทดสอบในลักษณะนี้จนสิ้นสุดการทดสอบซึ่งผู้สอบจะทำแบบทดสอบเพียง 30 ข้อเท่านั้น จากจำนวนข้อสอบในแบบทดสอบทั้งหมด 59 ข้อ

4. โค้งลักษณะข้อสอบ (Item Characteristic Curve : ICC) หมายถึงกราฟของฟังก์ชันการตอบสนองข้อสอบที่แสดงความสัมพันธ์ระหว่างระดับความสามารถของผู้สอบกับความน่าจะเป็นที่ผู้สอบซึ่งมีความสามารถระดับนั้น จะตอบข้อสอบข้อนั้นได้ถูกต้อง

5. ค่าความยากของข้อสอบ (Item Difficulty : b) หมายถึงค่าความสามารถตรงจุดเปลี่ยนโค้งลักษณะข้อสอบแต่ละข้อ ซึ่งเป็นระดับความสามารถที่จะตอบข้อสอบข้อนั้นได้ถูกต้องด้วยความน่าจะเป็น 0.50 เมื่อไม่มีการเดา ค่าความยากของข้อสอบมีค่าเป็นไปได้อยู่ระหว่าง  $-\infty$  ถึง  $+\infty$  แต่ในการศึกษาครั้งนี้จะคัดเลือกข้อสอบที่มีค่าความยากอยู่ในช่วง  $-3.00$  ถึง  $3.00$  (Warm. 1978 : 52)

6. ค่าความสามารถของผู้สอบ (Examinee's Ability :  $\theta$ ) หมายถึงระดับความสามารถของผู้สอบแต่ละคน ซึ่งวัดในหน่วย  $\theta$  ที่มีค่าเฉลี่ยเท่ากับ 0 และค่าเบี่ยงเบนมาตรฐาน เท่ากับ 1 สำหรับค่าความสามารถมีค่าเป็นไปได้ตั้งแต่  $-\infty$  ถึง  $+\infty$  แต่ในการศึกษาครั้งนี้ผู้วิจัยใช้ค่าความสามารถในช่วง  $-4.00$  ถึง  $4.00$  โดยแบ่งระดับความสามารถออกเป็น 5 ระดับ ดังนี้

6.1 ระดับต่ำมาก  $\theta < -3.00$

6.2 ระดับต่ำ  $-3.00 \leq \theta < -1.00$

6.3 ระดับปานกลาง  $-1.00 \leq \theta \leq 1.00$

6.4 ระดับสูง  $1.00 < \theta \leq 3.00$

6.5 ระดับสูงมาก  $\theta > 3.00$

7. การประมาณค่าความยากของข้อสอบและค่าความสามารถของผู้สอบ หมายถึงการนำผลการทดสอบจากแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กซีเลเวล ไปประมาณค่าตามทฤษฎีการตอบสนองข้อสอบ แบบราสซัส โมเดล ด้วยวิธีการประมาณค่าความเป็นไปได้สูงสุดแบบไม่มีเงื่อนไข โดยใช้โปรแกรมสำเร็จรูปใบคาล

8. ค่าฟังก์ชันอินฟอร์เมชันของข้อสอบ (Item Information Function) หมายถึงค่าความแม่นยำในการประมาณค่าความสามารถจริงจากผลการตอบข้อสอบข้อนั้น ๆ ซึ่งมีค่าเท่ากับอัตราส่วนของกำลังสองของความชันโค้งลักษณะข้อสอบต่อค่าความแปรปรวนของข้อสอบข้อนั้น โดยเขียนเป็นสมการทางคณิตศาสตร์ได้ดังนี้ (Lord and Novick. 1968 : 454)

$$I(\theta, u_i) = [P'_i(\theta)]^2 / P_i(\theta)Q_i(\theta)$$

- เมื่อ  $I(\theta, U_1)$  แทน ค่าฟังก์ชันอินฟอร์เมชันของข้อสอบข้อที่  $i$
- $P_i(\theta)$  แทน ค่าความน่าจะเป็นที่ผู้สอบซึ่งมีระดับความสามารถ  $\theta$  จะตอบข้อสอบข้อที่  $i$  ได้ถูกต้อง
- $Q_i(\theta)$  แทน  $1 - P_i(\theta)$
- $P'_i(\theta)$  แทน ความชันโค้งลักษณะข้อสอบที่ระดับความสามารถ  $\theta$

9. ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ (Test Information Function : TIF) หมายถึงผลรวมของค่าฟังก์ชันอินฟอร์เมชันของข้อสอบทุก ๆ ข้อ ในแบบทดสอบนั้น ซึ่งเขียนเป็นสมการทางคณิตศาสตร์ได้ดังนี้

$$I(\theta) = \sum_{i=1}^n I(\theta, U_i)$$

10. ค่าความเที่ยงตรง (Validity) ของการประมาณค่าความสามารถ ซึ่งในที่นี้ เป็นการหาค่าความเที่ยงตรงเชิงสภาพ หมายถึงดัชนีที่บ่งชี้ถึงคุณภาพของแบบทดสอบในการประมาณค่าความสามารถของผู้สอบได้ถูกต้องตรงกับสภาพที่เป็นจริงในปัจจุบัน ซึ่งคำนวณได้จากความสัมพันธ์ระหว่างความสามารถที่ได้จากการประมาณค่า กับเกรดเฉลี่ยสะสมในวิชาคณิตศาสตร์ที่เรียนผ่านมาในชั้นมัธยมศึกษาปีที่ 1 และปีที่ 2 โดยใช้สูตรสัมประสิทธิ์สหสัมพันธ์เพียร์สัน

11. ค่าประสิทธิภาพสัมพัทธ์ (Relative Efficiency : RE) ของการประมาณค่าความสามารถ หมายถึงดัชนีที่บ่งชี้ถึงประสิทธิภาพของแบบทดสอบ ซึ่งคำนวณได้จากการหาอัตราส่วนระหว่างค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบแต่ละฉบับที่นำมาเปรียบเทียบกัน ณ ระดับความสามารถเดียวกัน โดยใช้สูตรดังนี้ (Lord and Novick. 1968 : 472)

$$RE(\theta, X_1, X_2) = I(\theta, X_1) / I(\theta, X_2)$$

เมื่อ  $RE(\theta, X_1, X_2)$  แทน ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่า

ความสามารถระหว่างแบบทดสอบรูปแบบที่ 1  
( $X_1$ ) เทียบกับแบบทดสอบรูปแบบที่ 2 ( $X_2$ )

$I(\theta, X_1)$  แทน ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบรูปแบบที่ 1  
( $X_1$ ) ณ ระดับความสามารถ  $\theta$

$I(\theta, X_2)$  แทน ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบรูปแบบที่ 2  
( $X_2$ ) ณ ระดับความสามารถ  $\theta$

สำหรับค่าประสิทธิภาพสัมพัทธ์ที่คำนวณได้ นำมาแปลผลดังนี้

(Warm. 1978 : 76)

เมื่อ  $RE(\theta, X_1, X_2) > 1$  แสดงว่า แบบทดสอบรูปแบบที่ 1 ( $X_1$ )  
มีประสิทธิภาพสูงกว่าแบบทดสอบรูปแบบ  
ที่ 2 ( $X_2$ ) ณ ระดับความสามารถ  $\theta$

$RE(\theta, X_1, X_2) < 1$  แสดงว่า แบบทดสอบรูปแบบที่ 1 ( $X_1$ )  
มีประสิทธิภาพต่ำกว่าแบบทดสอบรูปแบบ  
ที่ 2 ( $X_2$ ) ณ ระดับความสามารถ  $\theta$

$RE(\theta, X_1, X_2) = 1$  แสดงว่า แบบทดสอบรูปแบบที่ 1 ( $X_1$ )  
มีประสิทธิภาพพอ ๆ กับแบบทดสอบรูปแบบ  
ที่ 2 ( $X_2$ ) ณ ระดับความสามารถ  $\theta$

## บทที่ 2

### เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยขอแนะนำเสนอตามลำดับดังนี้

1. ทฤษฎีการตอบสนองข้อสอบ
  - 1.1 หลักการของทฤษฎีการตอบสนองข้อสอบ
  - 1.2 ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ
  - 1.3 โมเดลของทฤษฎีการตอบสนองข้อสอบ
2. การประมาณค่าพารามิเตอร์
3. คุณภาพของแบบทดสอบ
4. ประสิทธิภาพของแบบทดสอบ
5. การทดสอบแบบเทเลอร์
  - 5.1 หลักการของการทดสอบแบบเทเลอร์
  - 5.2 ประเภทของการทดสอบแบบเทเลอร์
  - 5.3 การให้คะแนนของการทดสอบแบบเทเลอร์
6. งานวิจัยที่เกี่ยวข้อง
  - 6.1 งานวิจัยภายในประเทศ
  - 6.2 งานวิจัยในต่างประเทศ

#### ทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบมีชื่อเรียกอื่น ๆ อีก เช่น ทฤษฎีคุณลักษณะแฝง ทฤษฎีโครงสร้างข้อสอบ หรือทฤษฎีการทดสอบแบบใหม่ (ผจงจิต อินทสุวรรณ, 2525 : 54 ; Hambleton and others. 1978 : 468) เป็นต้น ทฤษฎีนี้ได้เข้ามามีบทบาทในวงการวัดผลการศึกษาและจิตวิทยาตั้งแต่ปี ค.ศ. 1950 - 1960 โดยในระยะแรกการพัฒนาเป็นไปอย่างช้า ๆ เพราะขาดแคลนคอมพิวเตอร์ที่ใช้ในการวิเคราะห์ข้อมูล แต่ในปัจจุบันได้มีการพัฒนาโปรแกรมคอมพิวเตอร์ จึงทำให้ได้รับความสะดวก และเป็นที่สนใจของนักวัดผลมากยิ่งขึ้น

สำหรับข้อดีของทฤษฎีนี้ก็คือ เชื่อกันว่าสามารถแก้ข้อบกพร่องของทฤษฎีการทดสอบแบบดั้งเดิมที่เราใช้กันมาเป็นเวลานานได้ เนื่องจากทฤษฎีการทดสอบแบบดั้งเดิมมีจุดอ่อนหลายประการดังนี้ (Hambleton and Swaminathan. 1985 : 1 - 3)

1. ค่าสถิติของข้อสอบ ซึ่งได้แก่ค่าความยากของข้อสอบ และค่าอำนาจจำแนกของข้อสอบ มีค่าไม่คงที่ จะแปรเปลี่ยนไปตามกลุ่มผู้สอบ ดังนั้นการใช้ค่าสถิติดังกล่าวมีข้อจำกัด โดยจะใช้ได้กับกลุ่มผู้สอบที่มีลักษณะเหมือน ๆ กันเท่านั้น
2. การเปรียบเทียบความสามารถของผู้สอบ จะต้องใช้แบบทดสอบฉบับเดียวกัน หรือแบบทดสอบคู่ขนานกันเท่านั้น เพราะถ้าใช้แบบทดสอบคนละฉบับที่วัดความสามารถเดียวกัน แต่ความยากง่ายแตกต่างกัน แล้วจะนำคะแนนจากแบบทดสอบคนละฉบับมาเปรียบเทียบกันไม่ได้
3. นิยามความเชื่อมั่นของแบบทดสอบในรูปของการวัดคู่ขนาน ซึ่งเป็นไปได้ยากที่จะสัมฤทธิ์ผลในทางปฏิบัติ เนื่องจากผู้ตอบอาจจะตอบ ไม่เหมือนเดิมในครั้งที่สอง
4. ไม่มีพื้นฐานในการที่จะบอกได้ว่าผู้สอบคนหนึ่งจะมีความน่าจะเป็นในการตอบข้อสอบข้อหนึ่ง ได้ถูกต้องเพียงใด ซึ่งทำให้ไม่สามารถพัฒนาแบบทดสอบให้เหมาะสมกับระดับความสามารถของผู้สอบแต่ละคนได้
5. การกำหนดข้อตกลงเบื้องต้นให้ความแปรปรวนของความคลาดเคลื่อนในการวัดของผู้สอบทุกคนมีค่าเท่ากัน ซึ่งขัดแย้งกับผลการวิจัยของ ลอร์ดและโนวิก (Lord and Novick. 1968 : 183) ที่พบว่า ความคลาดเคลื่อนในการวัดของผู้สอบแต่ละคนมีค่าไม่เท่ากัน โดยค่าความคลาดเคลื่อนในการวัดสำหรับคะแนนปานกลาง จะมีค่าน้อยกว่าค่าความคลาดเคลื่อนในการวัดสำหรับคะแนนสูงหรือคะแนนต่ำ

จากปัญหาข้อบกพร่องของทฤษฎีการทดสอบแบบดั้งเดิมดังที่กล่าวมาแล้ว จึงทำให้นักวิชาการได้ให้ความสนใจกับทฤษฎีการตอบสนองข้อสอบ ซึ่งทฤษฎีนี้มีจุดเด่นดังนี้ (Hambleton and Swaminathan. 1985 : 11)

1. ค่าพารามิเตอร์ของข้อสอบ เป็นอิสระจากกลุ่มผู้สอบที่ใช้ในการประมาณค่า
2. ค่าความสามารถของผู้สอบ เป็นอิสระจากข้อสอบที่ใช้ในการประมาณค่า
3. สามารถหาความแม่นยำ ในการประมาณค่าความสามารถของผู้สอบได้

### 1. หลักการของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ เป็นทฤษฎีที่ใช้โมเดลทางคณิตศาสตร์อธิบายถึงความสัมพันธ์ระหว่างความสามารถหรือคุณลักษณะภายในของผู้สอบ กับผลการสอบหรือคะแนนของผู้สอบนั้น โดยทฤษฎีนี้เชื่อว่าคะแนนของผู้สอบ ซึ่งเป็นสิ่งที่วัดหรือสังเกตได้โดยตรง สามารถอธิบายหรือพยากรณ์ได้จากความสามารถของผู้สอบคนนั้น ซึ่งเป็นสิ่งที่ไม่สามารถวัดหรือสังเกตได้ (Lord and Novick. 1968 : 358 ; Hambleton and Swaminathan. 1985 : 9) และโมเดลทางคณิตศาสตร์ที่ใช้อธิบายถูกกำหนดด้วยฟังก์ชันเพิ่มทางเดียว (Monotonically Increasing Function) ซึ่งเรียกว่า ฟังก์ชันลักษณะข้อสอบ หรือ โค้งลักษณะข้อสอบ (Hambleton and Swaminathan. 1985 : 13)

สำหรับความสัมพันธ์ระหว่างความสามารถของผู้สอบกับผลการสอบ สามารถเขียนให้อยู่ในรูปความสัมพันธ์ทั่ว ๆ ไป ได้ดังนี้ (สำเร็จ บุญเรืองรัตน์. 2525 : 3)

$$p = f(\theta)$$

เมื่อ  $p$  แทน ผลการสอบ  
 $\theta$  แทน ความสามารถ  
 $f$  แทน ฟังก์ชัน

### 2. ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

การที่จะนำทฤษฎีการตอบสนองข้อสอบไปใช้ จะต้องเป็นไปตามข้อตกลงเบื้องต้นดังนี้

#### 2.1 มิติลาเท็นท์สเปส (Dimensionality of the Latent Space)

พฤติกรรมการตอบแบบทดสอบสามารถอธิบายได้ด้วยความสามารถหรือคุณลักษณะหลาย ๆ อย่าง ความสามารถทั้งหมดรวมเรียกว่า ลาเท็นท์สเปส (Latent Space) ความสามารถแต่ละอย่างจะกำหนดมิติของลาเท็นท์สเปสขึ้นมาหนึ่งมิติ โดยทั่วไปถ้าผลการสอบหรือคะแนนของผู้สอบ สามารถอธิบายได้ด้วยความสามารถหรือคุณลักษณะเดียวก็เพียงพอแล้ว ซึ่งจะเรียกเงื่อนไขว่า ความเป็นมิติเดียว (Unidimensional) (Hambleton and Swaminathan. 1985 : 16 - 17) และสามารถทำการตรวจสอบความเป็นมิติเดียวได้โดยการวิเคราะห์

องค์ประกอบ (Factor Analysis) เพื่อสังเกตค่าไอเกน (Eigen Value) ขององค์ประกอบที่มีค่าสูงสุดว่าแตกต่างจากค่าอื่น ๆ อย่างชัดเจนหรือไม่ ในขณะที่ค่าไอเกนขององค์ประกอบที่เหลือทั้งหมดมีค่าใกล้เคียงกัน (Warm. 1978 : 104 and Lord. 1980 : 21) นอกจากนี้ยังสามารถเลือกข้อสอบที่มีความเป็นมิติเดียวอย่างง่าย ๆ โดยการคัดเลือกข้อสอบที่มีค่าอำนาจจำแนกสูง ๆ ซึ่งไม่ควรต่ำกว่า 0.8 (Green and others. 1984 : 349 ; citing Urry. 1981)

2.2 ความเป็นอิสระในการตอบข้อสอบ (Local Independence) หมายถึงความน่าจะเป็นของผู้สอบในการตอบข้อสอบแต่ละข้อในแบบทดสอบใดถูกต้องจะเป็นอิสระจากกัน นั่นคือ การตอบข้อสอบข้อใดข้อหนึ่ง ได้ถูกหรือผิดจะ ไม่มีผลต่อการตอบข้ออื่น ๆ ซึ่งสามารถเขียนเป็นสมการทางคณิตศาสตร์ได้ดังนี้ (Hambleton and Swaminathan. 1985 : 23)

$$\text{Prob} [U_1 = u_1, U_2 = u_2, \dots, U_n = u_n / \theta] = \prod_{i=1}^n P_i(\theta)^{u_i} Q_i(\theta)^{1-u_i}$$

เมื่อ  $U_i$  แทน ผลการตอบข้อสอบข้อที่  $i$  ของผู้สอบคนหนึ่ง ( $i = 1, 2, 3, \dots, n$ )

ถ้าตอบถูก  $U_i = 1$  และตอบผิด  $U_i = 0$ )

$P_i(\theta)$  แทน ความน่าจะเป็นที่ผู้สอบ ซึ่งมีระดับความสามารถ  $\theta$  จะตอบข้อสอบข้อที่  $i$  ได้ถูกต้อง

$Q_i(\theta)$  แทน  $1 - P_i(\theta)$

สำหรับการตรวจสอบความเป็นอิสระในการตอบข้อสอบนั้น ให้สังเกตความเป็นมิติเดียวของแบบทดสอบ โดยถ้าแบบทดสอบมีความเป็นมิติเดียว แล้วแบบทดสอบนั้นจะมีความเป็นอิสระในการตอบข้อสอบด้วย (Warm. 1978 : 107 ; Lord. 1980 : 19 ; Hambleton and Swaminathan. 1985 : 24)

2.3 โค้งลักษณะข้อสอบ (Item Characteristic Curve : ICC) เป็นฟังก์ชันทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างความน่าจะเป็นในการตอบข้อสอบข้อนั้นได้ถูกต้อง กับระดับความสามารถที่วัดได้โดยใช้ชุดข้อสอบข้อนั้น และความน่าจะเป็นที่ผู้สอบจะ

ตอบข้อสอบ ได้ถูกต้องจะขึ้นอยู่กับ ใ้คงลักษณะข้อสอบในแต่ละ ไม่เคลที่ใช้ แต่จะ ไม่ขึ้นอยู่กับ การ  
 แจกแจงความสามารถของกลุ่มตัวอย่าง (Hambleton and Swaminathan. 1985 :  
 25 - 26) โดยผู้สอบที่มีความสามารถในระดับสูง จะมีความน่าจะเป็นในการตอบข้อสอบ  
 ได้ถูกต้องมากกว่าผู้สอบที่มีความสามารถในระดับต่ำ

### 3. โมเดลของทฤษฎีการตอบสนองข้อสอบ

โมเดลของทฤษฎีการตอบสนองข้อสอบมีลักษณะต่าง ๆ กัน ตามฟังก์ชันที่ใช้  
 แสดงความสัมพันธ์ระหว่างผลการสอบกับความสามารถของผู้สอบ และจำนวนพารามิเตอร์ที่ใช้  
 แสดงลักษณะข้อสอบ ซึ่งถ้าจะจัดแบ่ง โมเดลตามลักษณะของข้อมูล หรือการตรวจให้คะแนน  
 สามารถแบ่งได้ 3 ประเภท ดังนี้ (Hambleton and Swaminathan. 1985 : 35)

3.1 โมเดลที่มีการตรวจให้คะแนนเป็นสองค่า (Dichotomous) โดย  
 ถ้าตอบถูกให้ 1 คะแนน และถ้าตอบผิดให้ 0 คะแนน ซึ่งแบ่งออกเป็นโมเดลต่าง ๆ ดังนี้

3.1.1 กัตต์แมน เพอร์เฟกต์ โมเดล (Guttman Perfect Model)

3.1.2 ลาเทนท์ ดิสเทนซ์ โมเดล (Latent Distance Model)

3.1.3 ลาเทนท์ ลินีเยร์ โมเดล (Latent Linear Model)

3.1.4 นอร์มัล ออจีฟ โมเดล ที่มีพารามิเตอร์ 1,2,3 ตัว (One-,  
 Two -, Three - Parameter Normal Ogive Model)

3.1.5 โลจิสติก โมเดล ที่มีพารามิเตอร์ 1,2,3,4 ตัว (One-,  
 Two -, Three -, Four - Parameter Logistic Model)

3.1.6 ราส์ช โมเดล (Rasch Model)

3.2 โมเดลที่มีการตรวจให้คะแนนเป็นหลายค่า (Multi - Dichotomous)  
 โดยใช้กับข้อสอบที่ให้คะแนนแบบลำดับขั้น เช่น แบบทดสอบวัดเจตคติ แบบทดสอบวัดภาคปฏิบัติ  
 และแบบทดสอบวัดบุคลิกภาพ เป็นต้น ซึ่งแบ่งออกเป็นโมเดลต่าง ๆ ดังนี้

3.2.1 นอมินอล เรสพอน โมเดล (Nominal Response Model)

3.2.2 เกรด เรสพอน โมเดล (Graded Response Model)

3.2.3 พาเชียล เครดิต โมเดล (Partial Credit Model)

3.3 โมเดลที่มีการตรวจให้คะแนนแบบต่อเนื่อง (Continuous) ได้แก่  
 คันทันนิวอัล เรสพอน โมเดล (Continuous Response Model) โมเดลนี้พัฒนาขึ้นโดย

ซามิไจมา (Hambleton and Swaminathan. 1985 : 52 ; citing Samejima. 1973) ซึ่งสามารถนำไปใช้เป็นเครื่องมือของการทดสอบทางจิตวิทยา หรือการศึกษาทางด้านเจตคติ โดยผู้สอบต้องตอบคำถามลงในมาตราวัดแบบต่อเนื่อง (Continuous Scale)

สำหรับในการศึกษาครั้งนี้ ผู้วิจัยจะนำเสนอเฉพาะโมเดลที่มีการตรวจให้คะแนนเป็นสองค่าของ นอร์มัล ออจีฟ โมเดล, โลจิสติก โมเดล และราสช์ โมเดลเท่านั้น ซึ่งมีรายละเอียด ดังนี้

#### นอร์มัล ออจีฟ โมเดล

โมเดลนี้ ลอว์เลย์ เป็นผู้พัฒนาขึ้นมาตั้งแต่ปี ค.ศ. 1943 - 1944 (Lord and Novick. 1968 : 369; citing Lawley. 1943,1944) ต่อมาในปี ค.ศ. 1952 ลอร์ด ได้ศึกษาและพัฒนาวิธีการประมาณค่าพารามิเตอร์ของโมเดล จนสามารถนำไปใช้ได้กับข้อมูลที่เป็นแบบทดสอบวัดผลสัมฤทธิ์และแบบทดสอบวัดความถนัดได้สำเร็จ (Hambleton and Swaminathan. 1985 : 4 - 7; citing Lord. 1952) โดยโมเดลนี้มีความเชื่อว่าความสัมพันธ์ระหว่างความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องกับระดับความสามารถ จะอยู่ในรูปของฟังก์ชันการแจกแจงสะสมปกติ (Normal Cumulative Distribution Function) ซึ่งในครั้งแรกได้นำไปประยุกต์ใช้กับ นอร์มัล ออจีฟ โมเดล ที่มีพารามิเตอร์ 2 ตัว คือ พารามิเตอร์ค่าความยาก ( $b_i$ ) และพารามิเตอร์ค่าอำนาจจำแนก ( $a_i$ ) โดยมีลักษณะของโมเดลดังนี้ (Hambleton and Swaminathan. 1985 : 35)

$$P_i(\theta) = \int_{-\infty}^{a_i(\theta - b_i)} (1/\sqrt{2\pi}) \exp^{-Z^2/2} dZ$$

เมื่อ  $P_i(\theta)$  แทน ความน่าจะเป็นที่ผู้สอบซึ่งมีระดับความสามารถ  $\theta$  จะตอบข้อสอบข้อที่  $i$  ได้ถูกต้อง

$a_i$  แทน ค่าอำนาจจำแนกของข้อสอบข้อที่  $i$

$b_i$  แทน ค่าความยากของข้อสอบข้อที่  $i$

$\theta$  แทน ระดับความสามารถของผู้สอบ

Z แทน ความเบี่ยงเบนปกติ (Normal Deviation) จากการแจกแจง  
ซึ่งมีค่าเฉลี่ยเท่ากับ  $b_i$  และค่าความเบี่ยงเบนมาตรฐาน  
เท่ากับ  $1/a_i$

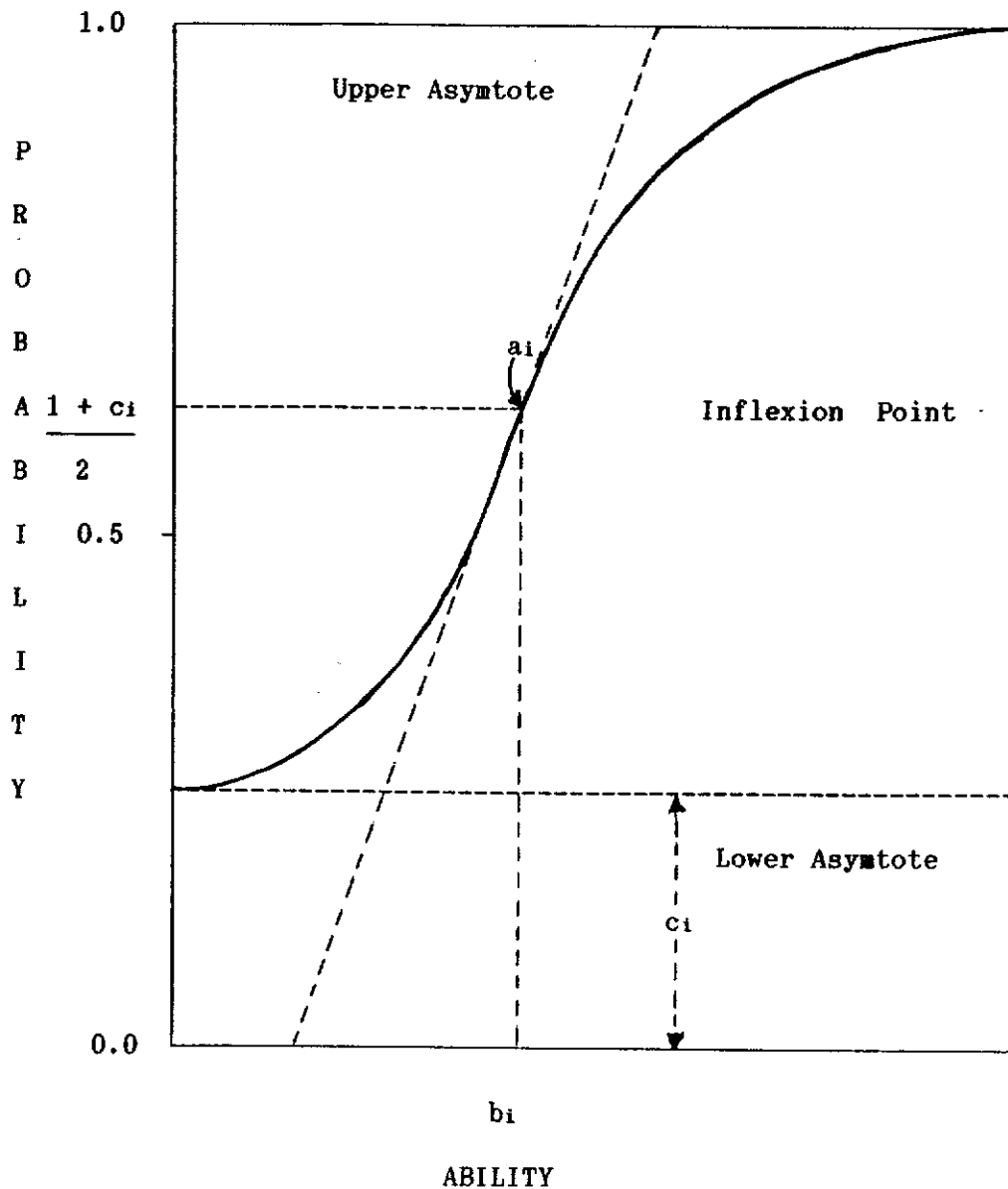
สำหรับการสอบที่ใช้แบบทดสอบเลือกตอบกับผู้สอบที่มีความสามารถในระดับ  
ต่ำ ๆ ผู้สอบอาจจะตอบข้อสอบถูกต้องด้วยการเดา โมเดลที่มีพารามิเตอร์เพียง 2 ตัว  
ไม่เหมาะสมที่จะอธิบายผลการสอบในลักษณะนี้ ดังนั้นจึงได้มีการเพิ่มพารามิเตอร์ของข้อสอบ  
เข้าไปอีก 1 ตัว คือ พารามิเตอร์ค่าการเดา ( $c_i$ ) เป็นนอร์มัล ออโจฟ โมเดล ที่มี  
พารามิเตอร์ 3 ตัว ซึ่งมีลักษณะของโมเดล ดังนี้ (Hambleton and Swaminathan,  
1985 : 50)

$$P_i(\theta) = c_i + (1 - c_i) \int_{-\infty}^{a_i(\theta - b_i)} (1/\sqrt{2\pi}) \exp^{-Z^2/2} dZ$$

ในกรณีที่กำหนดค่าให้ข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากันหมด และมีค่าการเดา  
เป็นศูนย์ โมเดลก็จะเหลือพารามิเตอร์ค่าความยาก ( $b_i$ ) เพียงตัวเดียว ซึ่งจะเรียกว่า  
นอร์มัล ออโจฟ โมเดล ที่มีพารามิเตอร์ 1 ตัว โดยมีลักษณะของโมเดล ดังนี้ (Hambleton  
and Swaminathan, 1985 : 50)

$$P_i(\theta) = \int_{-\infty}^{(\theta - b_i)} (1/\sqrt{2\pi}) \exp^{-Z^2/2} dZ$$

เส้นโค้งลักษณะข้อสอบของ นอร์มัล ออโจฟ โมเดล จะมีลักษณะเป็นรูปตัว  
เอส (S) ซึ่งแต่ละข้อจะมีลักษณะแตกต่างกันไป ทั้งนี้จะขึ้นอยู่กับค่าพารามิเตอร์ของข้อสอบ  
ดังภาพประกอบ 1 (Lord, 1980 : 14)



ภาพประกอบ 1 โค้งลักษณะข้อสอบของ นอร์มัล ออกไซด์ โมเดล ที่มีพารามิเตอร์ 3 ตัว

จากภาพประกอบ 1 พารามิเตอร์  $a_i$ ,  $b_i$ ,  $c_i$  และ  $\theta$  มีความหมายดังนี้

1. ค่าอำนาจจำแนกของข้อสอบ (Item Discrimination :  $a_i$ ) หมายถึง ค่าที่เป็นสัดส่วนกับความชันของโค้งลักษณะข้อสอบ หรือความชันของความน่าจะเป็นที่ผู้สอบ ซึ่งมีระดับความสามารถ  $\theta$  จะตอบข้อสอบข้อนั้นได้ถูกต้อง ณ จุดเปลี่ยนโค้ง ค่าอำนาจจำแนกของข้อสอบมีค่าเป็นไปได้อยู่ระหว่าง 0 ถึง  $+\infty$  แต่ในทางปฏิบัติจะคัดเลือกข้อสอบที่มีค่าอำนาจจำแนกอยู่ในช่วง 0.50 ถึง 2.50 (Warm. 1978 : 52)

2. ค่าความยากของข้อสอบ (Item Difficulty :  $b_1$ ) หมายถึงค่าความสามารถตรงจุดเปลี่ยนโค้งลักษณะข้อสอบแต่ละข้อ ซึ่งเป็นระดับความสามารถที่จะตอบข้อสอบข้อนั้นได้ถูกต้องด้วยความน่าจะเป็น 0.50 เมื่อไม่มีการเดา ค่าความยากของข้อสอบมีค่าเป็นไปได้อยู่ระหว่าง  $-\infty$  ถึง  $+\infty$  แต่ในทางปฏิบัติจะคัดเลือกข้อสอบที่มีค่าความยากอยู่ในช่วง  $-3.00$  ถึง  $+3.00$  (Warm. 1978 : 52)

3. ค่าการเดาของข้อสอบ (Guessing Parameter or Pseudo-Chance Score Level :  $c_1$ ) หมายถึงค่าความน่าจะเป็นที่บุคคลหนึ่งซึ่งปราศจากความสามารถจะตอบข้อสอบข้อนั้นได้ถูกต้อง ค่าการเดาของข้อสอบมีค่าเป็นไปได้อยู่ระหว่าง 0 ถึง 1 แต่ในทางปฏิบัติจะคัดเลือกข้อสอบที่มีค่าการเดาไม่เกิน 0.3 (Urry. 1977 : 184)

4. ค่าความสามารถของผู้สอบ (Examinee's Ability :  $\theta$ ) หมายถึงระดับความสามารถของผู้สอบแต่ละคน ซึ่งวัดในหน่วย  $\theta$  ที่มีค่าเฉลี่ยเท่ากับ 0 และค่าความเบี่ยงเบนมาตรฐานเท่ากับ 1 ค่าความสามารถของผู้สอบมีค่าเป็นไปได้อยู่ระหว่าง  $-\infty$  ถึง  $+\infty$  แต่ในทางปฏิบัติจะใช้ค่าความสามารถในช่วง  $-4.00$  ถึง  $+4.00$

### โลจิสติก ไมเคิล

ไมเคิลนี้ เบิร์นบอม (Birnbau. 1968 : 399 - 400) ได้เป็นผู้เสนอขึ้นโดยในครั้งแรกได้พัฒนา โลจิสติก ไมเคิล ที่มีพารามิเตอร์ 2 ตัว คือ พารามิเตอร์ค่าอำนาจจำแนก ( $a_1$ ) และพารามิเตอร์ค่าความยาก ( $b_1$ ) ซึ่งเป็นไมเคิลที่มีโค้งลักษณะข้อสอบเป็นรูปตัวเอส (s) ที่มีลักษณะใกล้เคียงกับโค้งลักษณะข้อสอบแบบ นอร์มัล ออโจฟ ไมเคิล ที่มีพารามิเตอร์ 2 ตัว โดยมีค่าแตกต่างกันน้อยกว่า 0.01 สำหรับทุกค่าของความสามารถ (Birnbau. 1968 : 399 ; Lord. 1980 : 14) แต่ โลจิสติก ไมเคิล สามารถนำไปประยุกต์ใช้ได้ง่ายและสะดวกกว่า จึงทำให้ไมเคิลนี้เป็นที่นิยมใช้กันมาก สำหรับความสัมพันธ์ระหว่างความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องกับระดับความสามารถ จะอยู่ในรูปของฟังก์ชันการแจกแจงสะสมโลจิสติก (Logistic Cumulative Distribution Function) โดยมีลักษณะของ ไมเคิล ดังนี้ (Hambleton and Swaminathan. 1985 : 36)

$$P_i(\theta) = \frac{\exp(Da_i(\theta - b_i))}{1 + \exp(Da_i(\theta - b_i))}$$

โดย  $a_i$  และ  $b_i$  เป็นค่าพารามิเตอร์ของข้อสอบซึ่งมีความหมายเช่นเดียวกับ นอร์มัล ออไฟท์ โมเดล ส่วน  $\exp$  (Exponential) เป็นค่าคงที่ มีค่าประมาณ 2.71828... และ  $D$  (Scaling Factor) เป็นค่าคงที่ ที่ใช้ปรับโค้งลักษณะข้อสอบ มีค่าเท่ากับ 1.7

ต่อมา เบิร์นบอม ได้เพิ่มพารามิเตอร์ค่าการเดา ( $c_i$ ) อีก 1 ตัว จึงเป็น โลจิสติก โมเดล ที่มีพารามิเตอร์ 3 ตัว ซึ่งแสดงได้ด้วยโมเดลคณิตศาสตร์ ดังนี้ (Hambleton and Swaminathan. 1985 : 37)

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

สำหรับในกรณีที่กำหนดค่าการเดาของข้อสอบมีค่าเท่ากับศูนย์ ( $c_i = 0$ ) และ ถือว่าข้อสอบมีค่าอำนาจจำแนกเท่ากัน ( $a_i = a$ ) ซึ่งจะเรียกว่า โลจิสติก โมเดล ที่มี พารามิเตอร์ 1 ตัว โดยมีลักษณะของโมเดลดังนี้ (Hambleton and Swaminathan. 1985 : 47)

$$P_i(\theta) = \frac{\exp(D(\theta - b_i))}{1 + \exp(D(\theta - b_i))}$$

นอกจากนี้ผู้สอบที่มีความสามารถในระดับสูง อาจตอบข้อสอบผิดด้วยความสะเพร่า ทั้งที่ผู้สอบมีความสามารถเหนือกว่าแบบทดสอบ จึงทำให้ไม่สามารถเลือกคำตอบที่ถูกต้องได้ ซึ่งปัญหานี้ แม็คโดแนลด์, บาร์ตัน และลอร์ด (McDonald, Barton and Lord) ได้เสนอ เพิ่มพารามิเตอร์  $r$  อีก 1 ตัว โดยเรียกว่า โลจิสติก โมเดล ที่มีพารามิเตอร์ 4 ตัว ซึ่งมี ลักษณะของโมเดล ดังนี้ (Hambleton and Swaminathan. 1985 : 48 ; citing McDonald, Barton and Lord. 1967, 1981)

$$P_i(0) = \frac{c_i + (r_i - c_i) \exp \left( \frac{D_{ai}(\theta - b_i)}{[1 + \exp(D_{ai}(\theta - b_i))] \right)}{1}$$

เมื่อ  $r_i$  แทน ค่าที่ผู้สอบซึ่งมีความสามารถในระดับสูง แต่มีความสะเพร่าตอบ  
ข้อสอบไม่ถูกต้อง โดยมีค่าน้อยกว่า 1 เล็กน้อย

### ราล์ฟ ไบเคิล

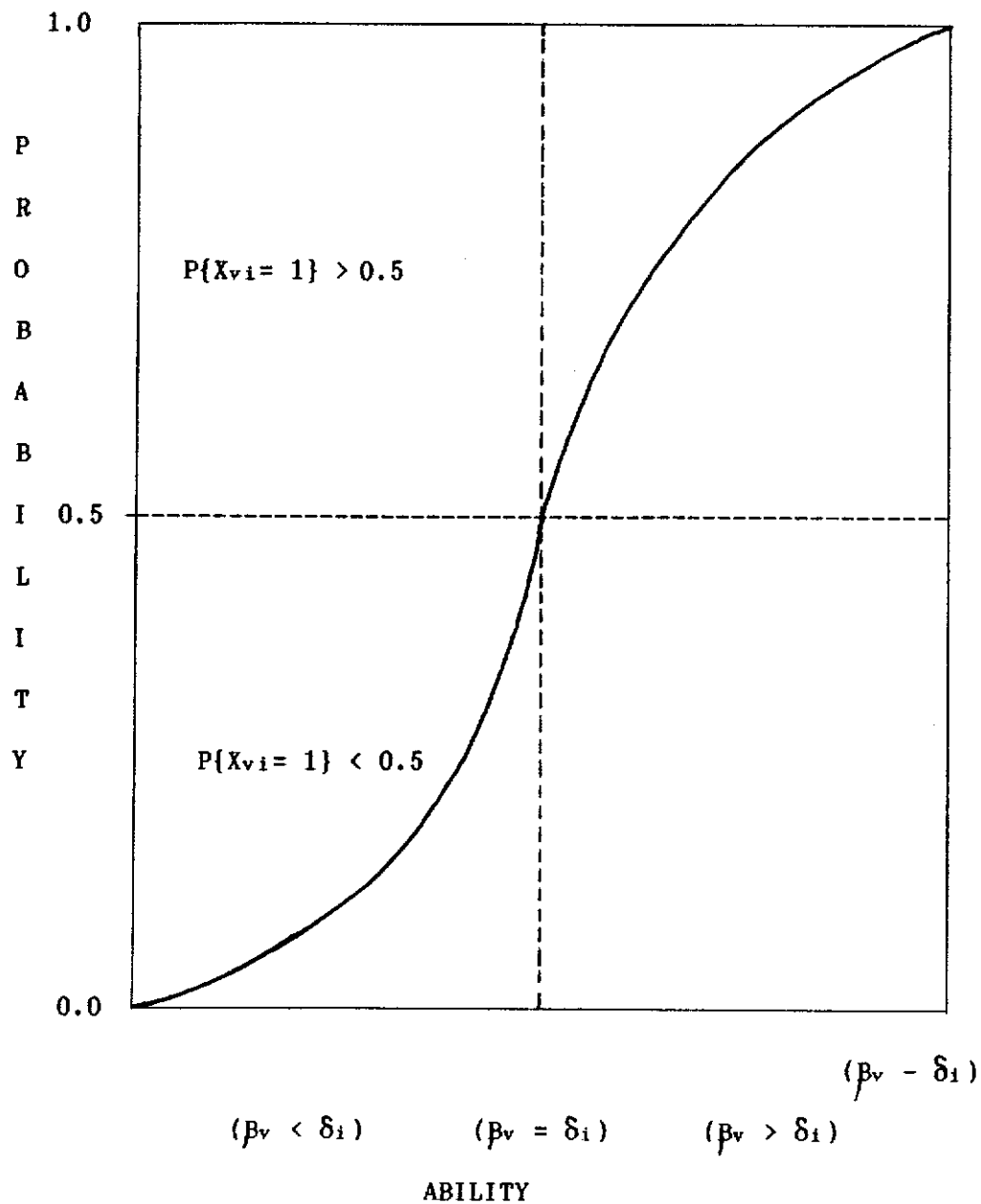
ไบเคิลนี้ได้รับการพัฒนาขึ้นโดย จอร์จ ราล์ฟ (Georg Rasch) ในปี ค.ศ. 1951-1960 โดยมีลักษณะคล้ายกับ โลจิสติก ไบเคิล ที่มีพารามิเตอร์ 1 ตัว ซึ่งพัฒนาขึ้นโดย เบิร์นบอม (Birnbau) ทั้งที่ไบเคิลทั้งสองถูกพัฒนาขึ้นมาอย่างอิสระจากกัน (Hambleton and Cook. 1977 : 82) สำหรับ ราล์ฟ ไบเคิล มีความเชื่อว่า ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบได้ถูกต้องขึ้นอยู่กับผลต่างระหว่างความสามารถของผู้สอบ ( $\beta_v$ ) กับค่าความยากของข้อสอบ ( $\delta_i$ ) ซึ่งสามารถอธิบายได้ด้วยการแจกแจงบัวซอง (Poisson Distribution) แต่ต่อมาได้พัฒนาให้อยู่ในรูป โลจิสติก เนื่องจากสามารถนำไปประยุกต์ใช้ได้ดีกว่า โดยมีลักษณะของไบเคิล ดังนี้ (Wright and Stone. 1979 : 15)

$$P\{X_{vi} = 1/\beta_v, \delta_i\} = \exp \left( \frac{(\beta_v - \delta_i)}{[1 + \exp(\beta_v - \delta_i)]} \right)$$

เมื่อ  $P\{X_{vi} = 1/\beta_v, \delta_i\}$  แทน ความน่าจะเป็นที่ผู้สอบซึ่งมีระดับความสามารถ  $\beta$   
จะตอบข้อสอบข้อที่  $i$  ได้ถูกต้อง

$\beta_v$  แทน ระดับความสามารถของผู้สอบคนที่  $v$   
 $\delta_i$  แทน ค่าความยากของข้อสอบข้อที่  $i$

สำหรับโลจิสติกส์ข้อสอบของ ราล์ฟ ไบเคิล มีลักษณะดังนี้ (Wright and Stone. 1979 : 14)



ภาพประกอบ 2 โค้งลักษณะข้อสอบของ ราสช์ โมเดล

จากภาพประกอบ 2 จะเห็นว่าค่าความยากของข้อสอบ ( $\delta_i$ ) จะอยู่บนแกนระดับความสามารถของผู้สอบ ( $\beta_v$ ) ตรงจุดของเส้นที่ลากมาจากจุดเปลี่ยนโค้งมาตั้งฉากกับแกนระดับความสามารถ หรือความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องมีค่าเท่ากับ 0.5 ดังนั้นพฤติกรรมการตอบข้อสอบจึงขึ้นอยู่กับความแตกต่างระหว่างความสามารถของผู้สอบ ( $\beta_v$ ) กับ

ค่าความยากของข้อสอบ ( $\delta_i$ ) ถ้าความสามารถของผู้สอบมีมากกว่าค่าความยากของข้อสอบ ( $\beta_v - \delta_i > 0$ ) ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบข้อนั้นได้ถูกต้องจะมีค่ามากกว่า 0.5 แต่ถ้าความสามารถของผู้สอบมีน้อยกว่าค่าความยากของข้อสอบ ( $\beta_v - \delta_i < 0$ ) ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบข้อนั้นได้ถูกต้องจะมีค่าน้อยกว่า 0.5 หรือถ้าผู้สอบมีความสามารถพอ ๆ กับค่าความยากของข้อสอบ ( $\beta_v - \delta_i = 0$ ) ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบข้อนั้นได้ถูกต้องจะมีค่าครึ่งต่อครึ่ง (Wright and Stone. 1979 : 12)

ราล์ช ไมเคิล เป็นโมเดลที่มีผู้นิยมมาประยุกต์ใช้กันมาก เนื่องจากเป็นโมเดลที่มีความแกร่ง (Robustness) ซึ่งมีรายละเอียดดังนี้

1. เป็นโมเดลที่นำมาใช้ได้อย่างง่าย ๆ ไม่ซับซ้อนเหมือนกับโมเดลอื่น ๆ ของทฤษฎีการตอบสนองข้อสอบ ถ้าข้อมูลมีไม่มากเกินไปสามารถคำนวณด้วยเครื่องคิดเลขธรรมดาได้ (Wright and Stone. 1979 : 28 - 45)

2. ความเป็นอิสระจากกลุ่มตัวอย่าง (Sample - Free) และความเป็นอิสระจากข้อสอบ (Item - Free) กล่าวคือ ค่าความยากของข้อสอบจะไม่แปรเปลี่ยนไปตามกลุ่มตัวอย่าง และค่าความสามารถของผู้สอบจะไม่แปรเปลี่ยนไปตามค่าความยากของข้อสอบ (Wright and Panchapakesan. 1969 : 23)

3. การกระจายค่าความสามารถของกลุ่มตัวอย่าง หรือค่าความยากของข้อสอบ ไม่จำเป็นต้องเป็นโค้งปกติ เพียงแต่ว่าแนวโน้มว่าจะเป็นโค้งปกติเท่านั้นก็ได้ และกลุ่มตัวอย่างไม่จำเป็นต้องได้มาด้วยการสุ่มก็ได้ (Wright and Stone. 1979 : 20)

4. ข้อสอบไม่จำเป็นต้องมีค่าอำนาจจำแนกเท่ากัน และไม่จำเป็นต้องคำนึงถึงค่าการเดา (Wright and Panchapakesan. 1969 : 25)

5. ขนาดของกลุ่มตัวอย่างที่ใช้กับราล์ช ไมเคิล ไม่จำเป็นต้องใช้มากเหมือนกับโมเดลอื่น ๆ ที่มีจำนวนพารามิเตอร์หลายตัว ดังเช่น โลจิสติก โมเดล ที่มีพารามิเตอร์ 3 ตัว จะต้องใช้กลุ่มตัวอย่างขนาดใหญ่ประมาณ 1,000 คนขึ้นไป (Warm. 1978 : 109) แต่ราล์ช ไมเคิลจะใช้กลุ่มตัวอย่างขนาดเล็ก ซึ่งจากการศึกษาของ แวน เดอ ไวจ์เวอร์ (Van de Vijver. 1986 : 48) พบว่า กลุ่มตัวอย่างขนาด 100 - 500 คน เมื่อใช้กับข้อสอบตั้งแต่ 25 ข้อขึ้นไป สามารถประมาณค่าพารามิเตอร์ได้อย่างมีประสิทธิภาพ

สำหรับในการศึกษาครั้งนี้ ผู้วิจัยได้ใช้พารามิเตอร์ค่าความยากในการจัดข้อสอบตามค่าความยาก โดยไม่ได้ศึกษาพารามิเตอร์ที่เกี่ยวกับค่าอำนาจจำแนก และค่าการเดาของข้อสอบ ดังนั้นผู้วิจัยจึงเลือกใช้ ราล์ช ไมเคิล ในการวิเคราะห์ข้อมูลตามทฤษฎีการตอบสนองข้อสอบ

### การประมาณค่าพารามิเตอร์

การนำทฤษฎีการตอบสนองข้อสอบมาประยุกต์ใช้ นอกจากจะต้องเลือกโมเดลให้มีความเหมาะสมแล้ว วิธีการประมาณค่าพารามิเตอร์ก็มีความสำคัญอีกประการหนึ่ง ซึ่งจะต้องเลือกใช้ให้เหมาะสมกับสภาพของการวัดแต่ละครั้ง เนื่องจากวิธีการประมาณค่าพารามิเตอร์ที่มีประสิทธิภาพจะส่งผลทำให้การสรุปอ้างอิงเกี่ยวกับข้อมูลทางสถิติได้ถูกต้อง แต่ถ้าวิธีการประมาณค่าพารามิเตอร์ขาดประสิทธิภาพแล้ว ย่อมจะทำให้การสรุปอ้างอิงมีความคลาดเคลื่อนจากความเป็นจริงได้ สำหรับวิธีการประมาณค่าพารามิเตอร์ของราล์ช ไมเคิล มีด้วยกันหลายวิธีแต่จะขอนำเสนอวิธีที่สำคัญเพียง 3 วิธี ดังนี้

1. วิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข (Conditional Maximum Likelihood Procedure : CON) เป็นวิธีการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ และค่าความยากของข้อสอบได้อย่างอิสระซึ่งกันและกัน โดยการกำหนดเงื่อนไขให้การประมาณค่าพารามิเตอร์ความสามารถของผู้สอบแยกออกจากค่าความยากของข้อสอบ ซึ่งทำให้การประมาณค่าความสามารถไม่จำเป็นต้องอาศัยค่าความยาก และในทางเองเดียวกันการประมาณค่าความยากไม่จำเป็นต้องรู้ความสามารถก็ได้ สำหรับการประมาณค่าโดยวิธีนี้ไม่มีการกำหนดข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงค่าความสามารถ และค่าความยากของข้อสอบ อย่างไรก็ตามการประมาณค่าโดยวิธีนี้จะไม่มีความยุ่งยากในการคำนวณ และมีข้อจำกัดในกรณีที่แบบทดสอบมีจำนวนข้อสอบมาก ๆ โดยไม่สามารถวิเคราะห์ข้อสอบที่มีจำนวนข้อมากกว่า 25 ข้อขึ้นไปได้

2. วิธีความเป็นไปได้สูงสุดแบบไม่มีเงื่อนไข (Unconditional Likelihood Procedure : UCON) เป็นวิธีการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ และค่าความยากของข้อสอบไปพร้อม ๆ กัน โดยมีการกำหนดข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงความสามารถของผู้สอบ และค่าความยากของข้อสอบมีการแจกแจงแบบปกติ สำหรับการ

ประมาณค่า โดยวิธีนี้มีข้อดีก็คือ สามารถใช้วิเคราะห์ข้อสอบที่มีความยาวทุกขนาด ถึงแม้ว่าจะมีความลำเอียง ในการคำนวณที่เกิดจากความยาวของข้อสอบ แต่ก็สามารถปรับแก้ได้

3. วิธีพรีออกซ์ (Cohen's Approximation Procedure : PROX) เป็นวิธีการประมาณค่าพารามิเตอร์ที่สามารถคำนวณได้อย่างง่าย ๆ ถ้าข้อมูลที่น่าสนใจวิเคราะห์ไม่มากนัก สามารถคำนวณได้โดยไม่ต้องใช้เครื่องคอมพิวเตอร์ก็ได้ การประมาณค่าพารามิเตอร์โดยวิธีนี้มีข้อดกหลงเบื้องต้นเกี่ยวกับค่าความสามารถของผู้สอบ และค่าความยากของข้อสอบ มีการแจกแจงในลักษณะเบ้หรือมีแนวโน้มเป็นแบบปกติ สำหรับจุดเด่นของการประมาณค่าด้วยวิธีพรีออกซ์ก็คือ ค่าที่ประมาณได้จะเป็นอิสระจากกลุ่มตัวอย่าง และความยาวของแบบทดสอบ โดยแปลงคะแนนให้อยู่ในรูปของคะแนนโลจิต (Logit Score)

ในปัจจุบันนี้ ได้มีการนำเอาทฤษฎีการตอบสนองข้อสอบมาใช้ในการวัดผลการศึกษา และจิตวิทยามากยิ่งขึ้น เนื่องจากการพัฒนาโปรแกรมคอมพิวเตอร์เพื่อใช้ในการวิเคราะห์ข้อมูล จึงทำให้ได้รับความสะดวกและรวดเร็วเป็นอย่างมาก ในปี ค.ศ. 1976 ลอร์ดและคนอื่น ๆ ได้พัฒนาโปรแกรมโลจิสต์ (LOGIST) (Hambleton and Swaminathan. 1985 : 5 ; citing Lord and others. 1976) ซึ่งต่อมาได้มีการปรับปรุงและพัฒนาเพิ่มขึ้นเรื่อย ๆ จนสามารถนำมาใช้ในการวิเคราะห์ข้อมูลได้เป็นอย่างดี โดยสามารถประมาณค่าพารามิเตอร์แบบโลจิสติก โมเดล ที่มีพารามิเตอร์ 1, 2 หรือ 3 ตัว และในปี ค.ศ. 1976 ไรท์ และเมด ได้พัฒนาโปรแกรมไบคาล (BICAL) (Hambleton and Cook. 1977 : 90 ; citing Wright and Mead. 1976) เพื่อใช้ประมาณค่าพารามิเตอร์ แบบราสส์ โมเดล นอกจากนี้ยังมีการพัฒนาโปรแกรมอื่น ๆ อีก เช่น โปรแกรมแอลซิลเลส (ANCILLES) พีเอ็มแอล (PML) และไบล็อก (BILOG) (Hambleton and Swaminathan. 1985 : 148) เป็นต้น

สำหรับในการศึกษาครั้งนี้ผู้วิจัยได้เลือกใช้การประมาณค่าพารามิเตอร์ ด้วยวิธีความเป็นไปไคสูงสุดแบบไม่มีเงื่อนไข โดยใช้โปรแกรมสำเร็จรูปไบคาล

#### คุณภาพของแบบทดสอบ

การตรวจสอบคุณภาพของแบบทดสอบตามทฤษฎีการตอบสนองข้อสอบจะพิจารณาจากความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบ แบบทดสอบที่ดีหรือมีคุณภาพสูงนั้น

จะต้องเป็นแบบทดสอบที่ให้ผลการสอบ ไปประมาณค่าความสามารถของผู้สอบ ได้ถูกต้องแม่นยำมากที่สุด และดัชนีที่ใช้วัดความแม่นยำในการประมาณค่าความสามารถของผู้สอบก็คือ ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ (Test Information Function : TIF) (Birnbaum, 1968 : 418) ซึ่งได้จากผลรวมของค่าฟังก์ชันอินฟอร์เมชันของข้อสอบ (Item Information Function) โดยที่ค่าฟังก์ชันอินฟอร์เมชันของข้อสอบหรือแบบทดสอบ เป็นดัชนีผสมที่สร้างจากดัชนีบอกคุณลักษณะของข้อสอบหลาย ๆ ลักษณะด้วยกัน เช่น ค่าความยาก ค่าอำนาจจำแนก และค่าการเดาของข้อสอบ เป็นต้น โดยรวมกันเป็นดัชนีเพียงตัวเดียว ซึ่งทำให้สามารถบ่งชี้ถึงคุณภาพของข้อสอบหรือแบบทดสอบได้ และด้วยคุณสมบัติความไม่แปรเปลี่ยนของค่าพารามิเตอร์ของข้อสอบ จึงทำให้ค่าฟังก์ชันอินฟอร์เมชันมีความเหมาะสมที่จะใช้เป็นดัชนีบอกถึงคุณภาพของข้อสอบหรือแบบทดสอบ ตามแนวทฤษฎีการตอบสนองข้อสอบ ซึ่งสามารถใช้แทนการหาค่าความเชื่อมั่น และค่าความคลาดเคลื่อนมาตรฐานในการวัด ตามแนวทฤษฎีการทดสอบแบบดั้งเดิมได้ (Hambleton and Cook, 1977 : 84 ; Hambleton and Swaminathan, 1985 : 236) เนื่องจากการหาค่าความเชื่อมั่นเป็นวิธีการที่หยาบ และค่าความคลาดเคลื่อนมาตรฐานในการวัดตั้งอยู่บนสมมติฐานที่ผิด ที่กำหนดให้ความแปรปรวนของความคลาดเคลื่อนมีค่าเท่ากันสำหรับทุก ๆ คะแนน (Green and others, 1984 : 352) แต่ความเป็นจริงแล้ว ความแปรปรวนของความคลาดเคลื่อนของผู้สอบแต่ละคนมีค่าไม่เท่ากัน โดยค่าความคลาดเคลื่อนสำหรับคะแนนปานกลาง จะมีค่าน้อยกว่าค่าความคลาดเคลื่อนสำหรับคะแนนต่ำหรือคะแนนสูง (Lord and Novick, 1968 : 183) ดังนั้นจึงเป็นจุดอ่อนในการบ่งชี้ถึงคุณภาพของข้อสอบหรือแบบทดสอบ ส่วนค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบสามารถเขียนเป็นสมการได้ดังนี้ (Lord and Novick, 1968 : 454)

$$I(\theta) = \sum_{i=1}^n I(\theta, U_i) = \sum_{i=1}^n \{ [P'_i(\theta)]^2 / P_i(\theta) Q_i(\theta) \}$$

เมื่อ  $I(\theta)$  แทน ฟังก์ชันอินฟอร์เมชันของแบบทดสอบ

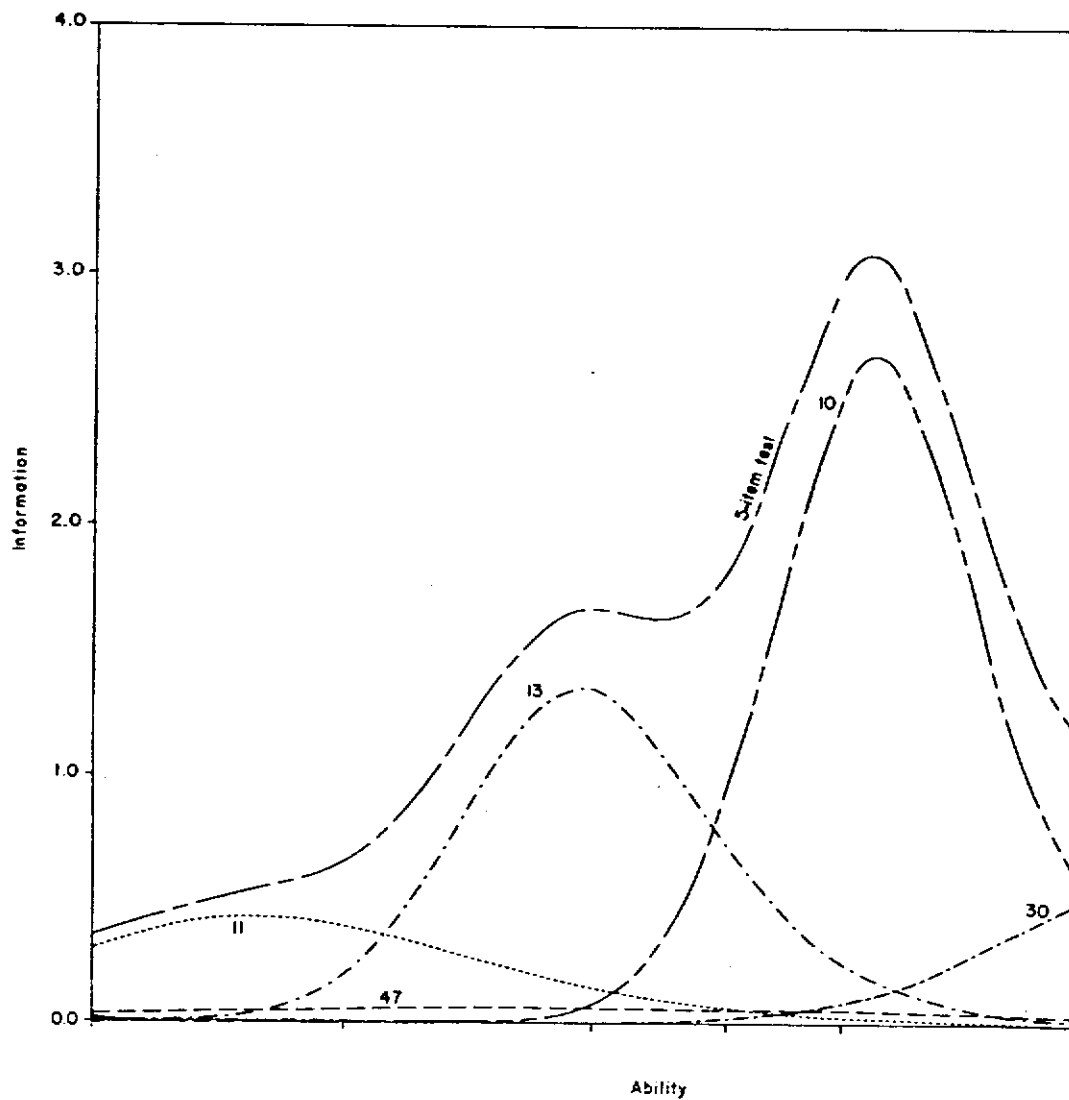
$I(\theta, U_i)$  แทน ฟังก์ชันอินฟอร์เมชันของข้อสอบแต่ละข้อ

$P_i(\theta)$  แทน ความน่าจะเป็นที่ผู้สอบซึ่งมีระดับความสามารถ  $\theta$   
จะตอบข้อสอบข้อที่  $i$  ได้ถูกต้อง

$Q_i(\theta)$  แทน  $1 - P_i(\theta)$

$P'_i(\theta)$  แทน ความชันของโค้งลักษณะข้อสอบที่ระดับความสามารถ  $\theta$

ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบฉบับหนึ่ง ๆ จะมีได้หลายค่าแตกต่างกันไปตามระดับความสามารถ ซึ่ง ๓ ระดับความสามารถใดที่ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบมีค่าสูง แสดงว่าแบบทดสอบฉบับนั้นมีความแม่นยำในการประมาณค่าความสามารถที่ระดับนั้นจะมีค่ามาก แต่ถ้า ๓ ระดับความสามารถใดที่ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบมีค่าต่ำ แสดงว่าแบบทดสอบฉบับนั้นจะมีความแม่นยำในการประมาณค่าความสามารถที่ระดับนั้นจะมีค่าน้อย (Hambleton and Cook. 1977 : 84) และเมื่อแบบทดสอบมีจำนวนข้อสอบมากขึ้นค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบจะมีค่าเพิ่มมากขึ้นด้วย (Hambleton and Swaminathan. 1985 : 107) สำหรับประโยชน์ที่ได้รับจากการทราบค่าฟังก์ชันอินฟอร์เมชันของข้อสอบแต่ละข้อนั้น ทำให้สามารถสร้างแบบทดสอบให้มีค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ ๓ ระดับหนึ่งของความสามารถตามที่ต้องการได้ เช่น ถ้าต้องการสร้างแบบทดสอบเพื่อใช้ในการคัดเลือกนักเรียนให้ได้รับทุน ก็ต้องใช้แบบทดสอบที่มีค่าฟังก์ชันอินฟอร์เมชันสูงสุด ๓ ระดับความสามารถสูง ๆ เพื่อจะทำได้แบบทดสอบที่มีประสิทธิภาพสูงสุด (ผจจจิต อินทสุวรรณ. 2525 : 66) ดังภาพประกอบ 3 (Hambleton and Cook. 1977 : 86 ; citing Lord. 1968)



ภาพประกอบ 3 โค้งอินฟอร์เมชันของข้อสอบ 5 ข้อ และ โค้งอินฟอร์เมชันของแบบทดสอบ

จากภาพประกอบ 3 แสดงว่าข้อสอบทั้ง 5 ข้อนี้ มีความเหมาะสมกับกลุ่มผู้สอบที่มีความสามารถในระดับสูง นั่นคือ ถ้านำไปทดสอบกับกลุ่มผู้สอบที่มีความสามารถในระดับสูงแล้วจะทำให้การประมาณค่าความสามารถของผู้สอบมีความถูกต้องแม่นยำสูง (Hambleton and Cook, 1977 : 87)

ในการวัดหรือการประมาณค่าใด ๆ มักจะมีความคลาดเคลื่อนเกิดขึ้นเสมอ โดยเฉพาะอย่างยิ่งในการประมาณค่าความสามารถของผู้สอบทางการศึกษาและจิตวิทยา จะมีความคลาดเคลื่อนค่อนข้างสูง ซึ่งจะส่งผลต่อความถูกต้องแม่นยำในการประมาณค่าความสามารถ โดยค่าความคลาดเคลื่อนในที่นี้จะอยู่ในรูปของความเบี่ยงเบนมาตรฐาน หรือเรียกว่า ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ (Standard Error of Estimate : SEE ( $\theta$ )) ซึ่งแตกต่างจากความคลาดเคลื่อนมาตรฐานในการวัด (Standard Error of Measurement : SE<sub>meas</sub>) ตรงที่ความคลาดเคลื่อนมาตรฐานในการประมาณค่า จะใช้ความคลาดเคลื่อนที่เกิดจากการประมาณค่าความสามารถของผู้สอบ ที่คลาดเคลื่อนไปจากความสามารถจริง ส่วนความคลาดเคลื่อนมาตรฐานในการวัดจะใช้ความคลาดเคลื่อนที่เกิดจากคะแนนดิบหรือคะแนนที่ได้จากการวัด ที่คลาดเคลื่อนไปจากคะแนนจริง (Lord and Novick. 1968 : 66 - 67) สำหรับค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า มีความสัมพันธ์กับค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ ดังนี้ (Warm. 1978 : 77 ; Hambleton and Swaminathan. 1985 : 104)

$$SEE(\theta) = 1/\sqrt{I(\theta)}$$

จากความสัมพันธ์ดังกล่าวแสดงว่า ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบแปรผกผันกับค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ กล่าวคือ ถ้าแบบทดสอบฉบับใดมีค่าฟังก์ชันอินฟอร์เมชันสูง ณ ระดับความสามารถ  $\theta$  แล้วความคลาดเคลื่อนมาตรฐานในการประมาณค่าจะมีค่าต่ำ แสดงว่าแบบทดสอบฉบับนั้นมีความแม่นยำในการประมาณค่าสูง ณ ระดับความสามารถ  $\theta$  นั้น แต่ถ้าแบบทดสอบฉบับใดมีค่าฟังก์ชันอินฟอร์เมชันต่ำ ณ ระดับความสามารถ  $\theta$  แล้วความคลาดเคลื่อนมาตรฐานในการประมาณค่าจะมีค่าสูง แสดงว่าแบบทดสอบฉบับนั้นมีความแม่นยำในการประมาณค่าต่ำ ณ ระดับความสามารถ  $\theta$  นั้น

สำหรับค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ( $\overline{SEE}(\theta)$ ) มีความสัมพันธ์กับค่าความเชื่อมั่น (Reliability :  $r_{tt}$ ) ดังนี้ (Warm. 1978 : 77)

$$r_{tt} = 1 - [\overline{SEE}(\theta)]^2$$

### ประสิทธิภาพของแบบทดสอบ

ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบนอกจากจะใช้พิจารณาถึงคุณภาพของแบบทดสอบแล้ว ยังสามารถนำมาใช้เปรียบเทียบประสิทธิภาพของแบบทดสอบได้อีกด้วย โดยการคำนวณหาอัตราส่วนของค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ 2 ฉบับ ที่นำมาเปรียบเทียบกัน ค่าที่คำนวณได้เรียกว่า ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่า (Relative Efficiency of Estimators : RE) ซึ่งคำนวณได้จากสูตรดังนี้ (Lord and Novick. 1968 : 471 - 472)

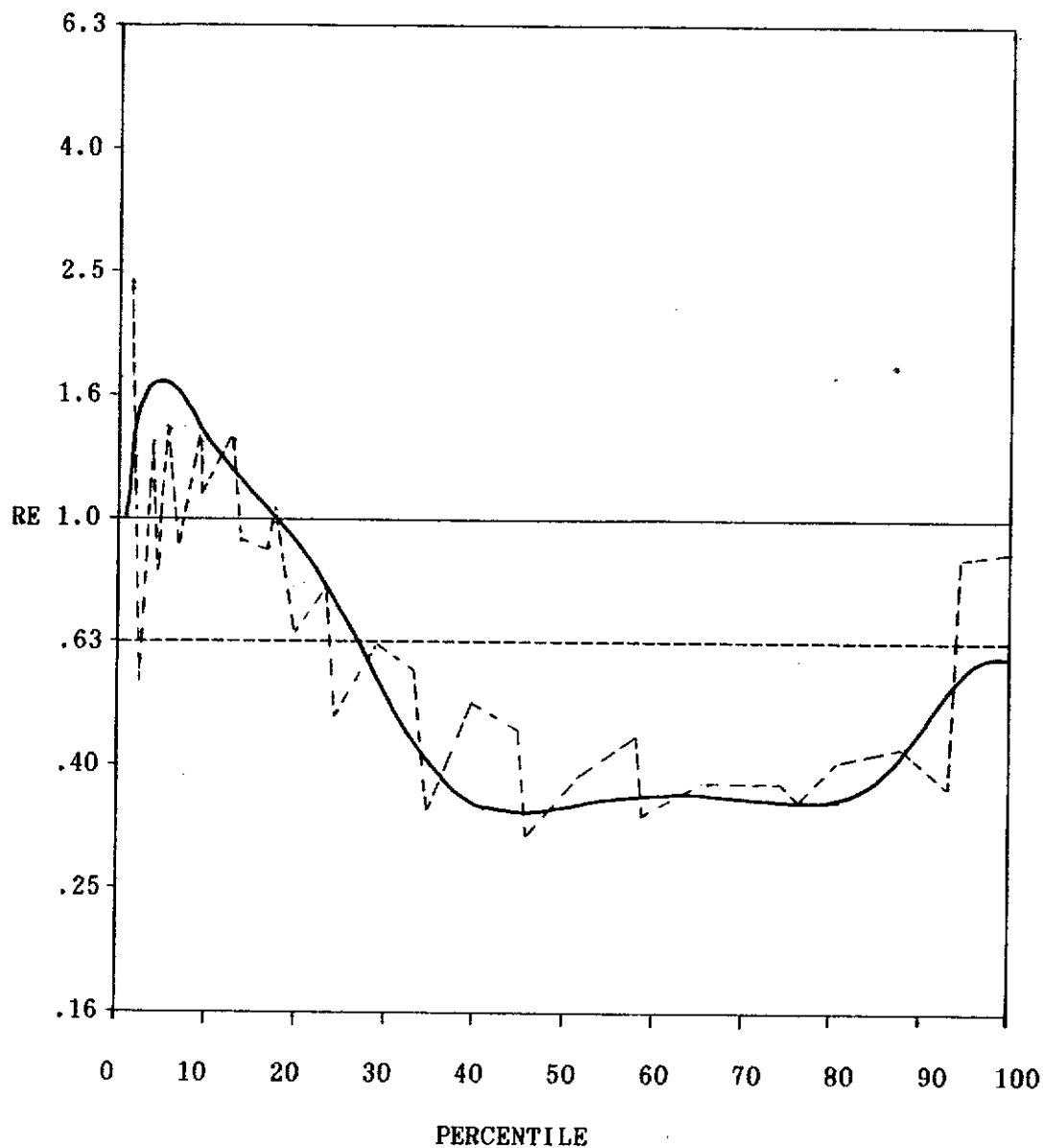
$$RE(\theta, X_1, X_2) = I(\theta, X_1) / I(\theta, X_2)$$

เมื่อ  $RE(\theta, X_1, X_2)$  แทน ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถระหว่างแบบทดสอบรูปแบบที่ 1 ( $X_1$ ) เทียบกับแบบทดสอบรูปแบบที่ 2 ( $X_2$ )

$I(\theta, X_1)$  แทน ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบรูปแบบที่ 1 ( $X_1$ ) ณ ระดับความสามารถ  $\theta$

$I(\theta, X_2)$  แทน ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบรูปแบบที่ 2 ( $X_2$ ) ณ ระดับความสามารถ  $\theta$

สำหรับการแปลความหมายของค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ ระหว่างแบบทดสอบที่นำมาเปรียบเทียบกันพิจารณาดังนี้ ถ้าได้  $RE(\theta, X_1, X_2)$  มากกว่า 1 แสดงว่าแบบทดสอบรูปแบบที่ 1 ( $X_1$ ) มีประสิทธิภาพสูงกว่าแบบทดสอบรูปแบบที่ 2 ( $X_2$ ) ณ ระดับความสามารถ  $\theta$  แต่ถ้า  $RE(\theta, X_1, X_2)$  น้อยกว่า 1 แสดงว่าแบบทดสอบรูปแบบที่ 1 ( $X_1$ ) มีประสิทธิภาพต่ำกว่าแบบทดสอบรูปแบบที่ 2 ( $X_2$ ) ณ ระดับความสามารถ  $\theta$  หรือถ้า  $RE(\theta, X_1, X_2)$  เท่ากับ 1 แสดงว่าแบบทดสอบทั้ง 2 รูปแบบมีประสิทธิภาพพอ ๆ กัน ณ ระดับความสามารถ  $\theta$  (Warm. 1978 : 75 - 76) ซึ่งสามารถแสดงได้ด้วยภาพประกอบ 4 ดังนี้ (Lord. 1980 : 97 ; citing Lord. 1974)



ภาพประกอบ 4 โค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบคำศัพท์ STEP เทียบกับแบบทดสอบคำศัพท์ MAT

จากภาพประกอบ 4 แบบทดสอบคำศัพท์ STEP มีประสิทธิภาพสูงกว่าแบบทดสอบคำศัพท์ MAT ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำ ๆ แต่ต่ำกว่าในกลุ่มผู้สอบที่มีความสามารถระดับสูง ๆ สำหรับเส้นประตามแนวนอนเป็นอัตราส่วนของจำนวนข้อของแบบทดสอบทั้ง 2 ฉบับ

### การทดสอบแบบเทเลอร์

การทดสอบแบบดั้งเดิม ที่เราใช้กันอยู่ทั่ว ๆ ไปนั้น เป็นการนำแบบทดสอบซึ่งประกอบด้วยข้อสอบที่มีค่าความยากในระดับปานกลาง ไปทดสอบกับกลุ่มผู้สอบ โดยผู้สอบแต่ละคนจะต้องทำข้อสอบในแบบทดสอบทุกข้อ ซึ่งเป็นข้อสอบที่มีการจัดเรียงลำดับไว้แน่นอน ดังนั้นแบบทดสอบในลักษณะนี้จึงมีความเหมาะสมกับกลุ่มผู้สอบที่มีความสามารถในระดับเฉลี่ยของกลุ่มเท่านั้น แต่จะไม่เหมาะสมกับกลุ่มผู้สอบที่มีความสามารถในระดับสูงหรือต่ำ (จักรณา พรหมมา, 2525 : 45 ; ผจงจิต อินทสุวรรณ, 2525 : 53) แบบทดสอบที่มีประสิทธิภาพสูงจะต้องประกอบด้วย ข้อสอบที่มีค่าความยากเหมาะสมกับระดับความสามารถของผู้สอบ นั่นคือ ข้อสอบต้องไม่ยากหรือง่ายเกินไปสำหรับผู้สอบแต่ละคน (Hambleton and others, 1978 : 496 - 497 ; citing Wood, 1973) ด้วยเหตุนี้ จึงทำให้นักวิจัยได้คิดค้นวิธีการสร้างข้อสอบและวิธีดำเนินการสอบแบบใหม่ เพื่อที่จะทำให้วัดความสามารถของผู้สอบได้ครอบคลุมทุกระดับความสามารถ โดยการปรับข้อสอบให้เหมาะสมกับระดับความสามารถของผู้สอบแต่ละคน ซึ่งวิธีการทดสอบดังกล่าวนี้มีชื่อเรียกแตกต่างกันออกไป เช่น การทดสอบที่ปรับตามระดับความสามารถ การทดสอบโดยใช้คอมพิวเตอร์ การทดสอบเป็นรายบุคคล การทดสอบแบบโปรแกรม การทดสอบที่มีการจัดเรียงลำดับข้อสอบ หรือข้อที่ เทอร์นบูล เรียกว่า การทดสอบแบบเทเลอร์ (Lord, 1980 : 150 - 151; citing Turnbull, 1951) โดยในระยะแรกยังไม่ได้นำมาใช้อย่างจริงจังเนื่องจากมีปัญหาเกี่ยวกับความยุ่งยากในการดำเนินการสอบที่ใช้การเขียนตอบแบบเดิม (Urry, 1977 : 181) ต่อมาในระยะหลัง ได้มีการพัฒนาโดยใช้เครื่องคอมพิวเตอร์มาช่วยในการดำเนินการสอบ ซึ่งสามารถจัดข้อสอบให้มีความเหมาะสมกับระดับความสามารถของผู้สอบ พร้อมทั้งนำเอาทฤษฎีการตอบสนองข้อสอบมาประยุกต์ใช้ในการประมาณค่าความสามารถของผู้สอบ จึงทำให้การทดสอบแบบเทเลอร์มีประสิทธิภาพมากยิ่งขึ้น (Rudner, 1983 : 951)

#### 1. หลักการทดสอบแบบเทเลอร์

การทดสอบแบบเทเลอร์ เป็นการทดสอบที่มีการจัดข้อสอบให้มีความเหมาะสมกับระดับความสามารถของผู้สอบแต่ละคน (Lord, 1971b : 3 ; Hambleton and others, 1978 : 497) ซึ่งผู้สอบแต่ละคนไม่จำเป็นต้องทำข้อสอบเหมือนกันทุกข้อ หรือจำนวนของข้อสอบก็อาจไม่เท่ากัน ทั้งนี้ขึ้นอยู่กับรูปแบบและผลการตอบข้อสอบใน

แบบทดสอบนั้น ๆ (Weiss and Kingsbury. 1984 : 361) โดยทั่วไปแล้วจะให้ผู้สอบ  
เริ่มทำข้อสอบที่มีค่าความยากปานกลางก่อน ถ้าตอบถูกข้อต่อไปจะได้ทำข้อสอบที่มีค่าความยาก  
เพิ่มมากขึ้น แต่ถ้าตอบข้อสอบนั้นผิดข้อต่อไปจะได้ทำข้อสอบที่ง่ายขึ้น และการทำข้อสอบข้อต่อ ๆ  
ไปก็จะดำเนินการในลักษณะนี้ไปจนสิ้นสุดการทดสอบ (Green and others. 1984 : 347)  
ตลอดช่วงระยะเวลาที่ดำเนินการสอบจะมีการประมาณค่าระดับความสามารถของผู้สอบโดย  
การวินิจฉัยจากผลการตอบของผู้สอบในข้อที่ตอบมาแล้วนั้น ดังนั้นผู้สอบแต่ละคนจะได้รับจำนวน  
ข้อสอบและการจัดเรียงลำดับข้อสอบที่แตกต่างกัน (Water and Bayroff. 1971 : 125 ;  
Spinetti and Hambleton. 1977 : 140)

## 2. ประเภทของการทดสอบแบบเทเลอร์

การทดสอบแบบเทเลอร์แบ่งออกเป็น 2 ประเภทใหญ่ ๆ ดังนี้ (Weiss.  
1974 : 78 - 110 ; Hambleton and Swaminathan. 1985 : 297)

2.1 ยุทธวิธีสองขั้นตอน (Two - Stage Strategies) เป็นยุทธวิธีของ  
การทดสอบแบบเทเลอร์อย่างง่าย ๆ โดยผู้สอบจะต้องทำแบบทดสอบถึงสองครั้งหรือสองขั้นตอน  
(Two - Stage Test) ขั้นตอนแรกเป็นแบบทดสอบแยกกลุ่ม (Routing Test) เพื่อแยก  
กลุ่มผู้สอบตามระดับความสามารถ ส่วนขั้นตอนที่สองเป็นแบบทดสอบวัดผล (Measurement  
Test) ซึ่งแบ่งออกเป็นแบบทดสอบฉบับย่อย ๆ หลายฉบับตามระดับความยาก ถ้าผู้สอบได้รับ  
การประเมินจากแบบทดสอบแยกกลุ่มว่ามีความสามารถในระดับต่ำ ระดับปานกลาง หรือ  
ระดับสูง แล้วจะได้ทำแบบทดสอบวัดผลที่มีค่าความยากในระดับต่ำ ระดับปานกลาง หรือ  
ระดับสูง ตามลำดับ

2.2 ยุทธวิธีหลายขั้นตอน (Multi - Stage Strategies) เป็นยุทธวิธี  
ของการทดสอบแบบเทเลอร์ที่ผู้สอบจะต้องตัดสินใจแยกทางหลาย ๆ ครั้ง จากการตอบข้อสอบ  
แต่ละข้อ ซึ่งแตกต่างกับยุทธวิธีสองขั้นตอนที่ผู้สอบจะต้องตัดสินใจแยกทางเพียงครั้งเดียว  
จากการตอบแบบทดสอบแยกกลุ่มไปยังแบบทดสอบวัดผล สำหรับรูปแบบของยุทธวิธีหลายขั้นตอน  
แบ่งออกเป็น 2 รูปแบบ ดังนี้

2.2.1 รูปแบบแยกทางคงที่ (Fixed - Branching Model) เป็น  
รูปแบบที่มีการกำหนดข้อสอบ และเส้นทางในการตอบข้อสอบไว้ล่วงหน้า เพื่อกำหนดให้ผู้สอบ  
แต่ละคนทำข้อสอบจากขั้นหนึ่ง ไปยังอีกขั้นหนึ่ง ซึ่งแบ่งออกเป็น 3 รูปแบบ ดังนี้

### 2.2.1.1 รูปแบบปิรามิด (Pyramidal Model)

หรือรูปแบบโครงสร้างต้นไม้ โดยแบ่งออกเป็นรูปแบบย่อย ๆ 5 รูปแบบ คือ รูปแบบขนาดขั้นคงที่ (Constant Step Size Pyramid) รูปแบบขนาดขั้นแปรผัน (Variable Step Size Pyramid) รูปแบบขั้นขาด (Truncated Pyramid) รูปแบบมีหลายข้อในแต่ละขั้น (Multiple - Item Pyramid) และรูปแบบให้ นักทักแก่ตัวเลือกของข้อสอบเพื่อแยกทาง (Pyramid Using Differential Response Option Branching)

### 2.2.1.2 แบบทดสอบเฟล็กซีเลเวล (Flexilevel Test)

เป็นแบบทดสอบที่ดัดแปลงมาจากรูปแบบปิรามิด ซึ่งแตกต่างตรงที่รูปแบบปิรามิดมีข้อสอบมากกว่าหนึ่งข้อในแต่ละระดับความยาก แต่แบบทดสอบเฟล็กซีเลเวลจะมีข้อสอบเพียงข้อเดียวในแต่ละระดับความยาก ดังนั้น จึงทำให้การแยกทางของผู้ตอบมีความแตกต่างกัน

### 2.2.1.3 แบบทดสอบปรับระดับขั้น (Stradaptive Test)

เป็นแบบทดสอบที่นำข้อสอบมาจัดกลุ่ม (Stratified) ตามค่าความยากซึ่งแบ่งออกเป็นระดับขั้น (Stratum) ในแต่ละระดับขั้นจะประกอบด้วยข้อสอบที่มีค่าความยากใกล้เคียงกัน และจะเรียงลำดับขั้นตามค่าความยากของข้อสอบที่สูงขึ้น สำหรับการดำเนินการสอบจะใช้ เครื่องคอมพิวเตอร์

### 2.2.2 รูปแบบแยกทางแปรผัน (Variable - Branching Model)

เป็นรูปแบบที่ไม่มีการกำหนดข้อสอบและเส้นทางในการตอบข้อสอบไว้ล่วงหน้า และจะดำเนินการสอบโดยใช้เครื่องคอมพิวเตอร์ ซึ่งเครื่องคอมพิวเตอร์จะคัดเลือกข้อสอบ จากกลุ่มข้อสอบที่มีค่าความยากของข้อสอบใกล้เคียงกับระดับความสามารถของผู้สอบ โดยใช้ข้อมูลจากผลการตอบในแต่ละข้อ การทดสอบจะยุติเมื่อการประมาณค่าความสามารถของผู้สอบมีค่าคงที่ และให้ค่าความคลาดเคลื่อนมาตรฐานเท่ากับเกณฑ์ที่ตั้งไว้ สำหรับรูปแบบแยกทางแปรผันที่ใช้กัน ได้แก่ ยุทธวิธีความเป็นไปได้สูงสุด (Maximum Likelihood Strategies) และยุทธวิธีของเบย์ (Bayesian Strategies)

ในการศึกษาครั้งนี้ ผู้วิจัยได้นำแบบทดสอบดั้งเดิม มาทำการศึกษาร่วมกับแบบทดสอบสองขั้นตอนและแบบทดสอบเฟล็กซีเลเวล ซึ่งเป็นรูปแบบหนึ่งของการทดสอบแบบเทเลอร์ ดังนั้นผู้วิจัยจึงขอเสนอรายละเอียดเฉพาะแบบทดสอบสองขั้นตอนและแบบทดสอบเฟล็กซีเลเวลเท่านั้น

### แบบทดสอบสองขั้นตอน

แบบทดสอบสองขั้นตอน จัดเป็นยุทธวิธีสองขั้นตอนของการทดสอบแบบเทเลอร์อย่างง่าย ๆ โดยสามารถนำมาประยุกต์ใช้กับการทดสอบเขียนตอบแบบเดิมได้เป็นอย่างดี ซึ่งผู้สอบจะต้องสอบถึงสองครั้ง คือ ครั้งแรกสอบแบบทดสอบแยกกลุ่ม และครั้งที่สองสอบแบบทดสอบวัดผล (Lord, 1971 d : 227 - 228 ; Linn , Rock and Cleary. 1969 : 132) โดยมีรายละเอียดดังนี้

1. แบบทดสอบแยกกลุ่ม เป็นแบบทดสอบที่ผู้สอบทุกคนจะต้องสอบเหมือนกัน โดยสอบเพื่อที่จะจัดแยกระดับความสามารถของผู้สอบว่าผู้สอบคนไหนจะได้ไปทำแบบทดสอบวัดผลฉบับใด ซึ่งจะยึดจำนวนข้อของข้อสอบที่ผู้สอบตอบถูกเป็นเกณฑ์ในการแยก ตามปกติแล้วแบบทดสอบแยกกลุ่มนี้จะมีจำนวนข้อสอบ ไม่มากนัก อาจจะมีประมาณ 10 ข้อ ซึ่งจะทำให้ตรวจข้อสอบได้รวดเร็ว เพื่อจะได้จัดแยกผู้สอบไปทำแบบทดสอบวัดผลต่อไป สำหรับค่าความยากของข้อสอบในแบบทดสอบแยกกลุ่ม อาจจะมีประกอบด้วยข้อสอบที่มีค่าความยากใกล้เคียงกัน (Peaked) หรือมีค่าความยากกระจายในช่วงกว้าง (Rectangular) หรือเป็นแบบทดสอบแยกกลุ่มสองครั้ง (Double) เป็นต้น

2. แบบทดสอบวัดผล เป็นแบบทดสอบที่ประกอบด้วยแบบทดสอบฉบับย่อย ๆ ประมาณ 3 - 5 ฉบับ โดยปกติแล้วจะมีฉบับละประมาณ 20 - 30 ข้อ จัดแยกตามระดับค่าความยากของข้อสอบ ซึ่งในแต่ละฉบับจะประกอบด้วยข้อสอบที่มีค่าความยากใกล้เคียงกัน โดยแบบทดสอบฉบับแรกจะเป็นแบบทดสอบที่ง่าย และฉบับต่อ ๆ ไปก็จะเริ่มยากขึ้นเรื่อย ๆ ผู้สอบแต่ละคนไม่จำเป็นต้องทำแบบทดสอบฉบับเดียวกัน ทั้งนี้ขึ้นอยู่กับความสามารถที่สามารถในแบบทดสอบแยกกลุ่ม โดยผู้ที่มีความสามารถในระดับต่ำจะได้ทำแบบทดสอบฉบับที่ง่าย ส่วนผู้ที่มีความสามารถในระดับปานกลางจะได้ทำแบบทดสอบฉบับที่มีความยากปานกลาง สำหรับผู้ที่มีความสามารถในระดับสูงจะได้ทำแบบทดสอบฉบับที่ยาก

แบบทดสอบฉบับย่อย ๆ ของแบบทดสอบวัดผล ควรมีข้อสอบบางส่วนที่มีค่าความยากของข้อสอบคาบเกี่ยวกันเป็นรายข้อ เพื่อป้องกันความคลาดเคลื่อนในการแยกทางของแบบทดสอบแยกกลุ่ม กล่าวคือ ถ้าผู้สอบได้คะแนนใกล้เคียง ๆ กับคะแนนจุดตัดจะทำให้แยกทาง ไปทำแบบทดสอบวัดผลที่ไม่เหมาะสมกับระดับความสามารถของผู้สอบ ดังนั้น ถ้ามีข้อสอบที่มีค่าความยากคาบเกี่ยวกันในแต่ละฉบับ จะสามารถทำให้ปรับความคลาดเคลื่อน

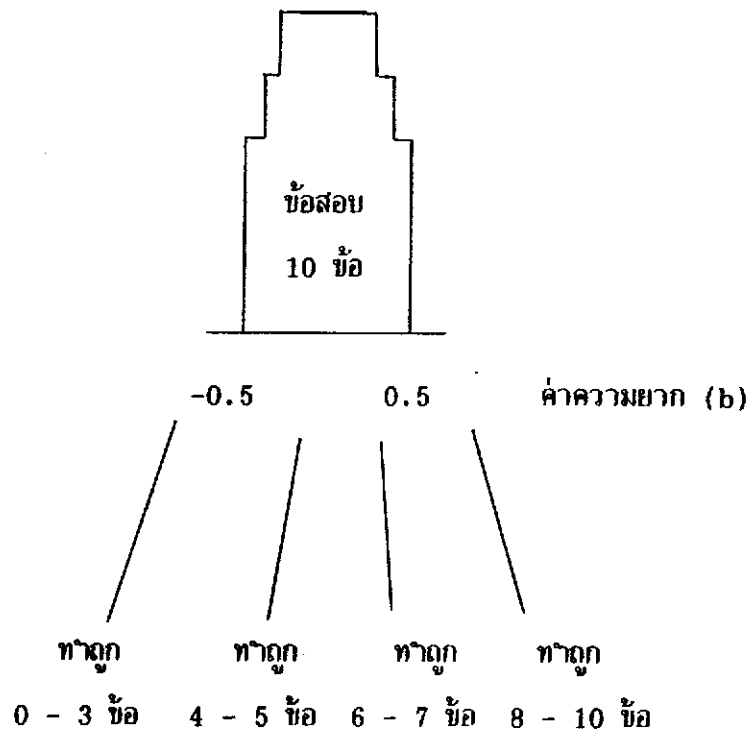
ตรงคะแนนจุดตัดได้ สำหรับการคำนวณจำนวนข้อสอบที่คาบเกี่ยวกันใช้สูตร ดังนี้ (Lord.  
1980 : 140 - 141)

$$m = (n_1 L - n) / (L - 1)$$

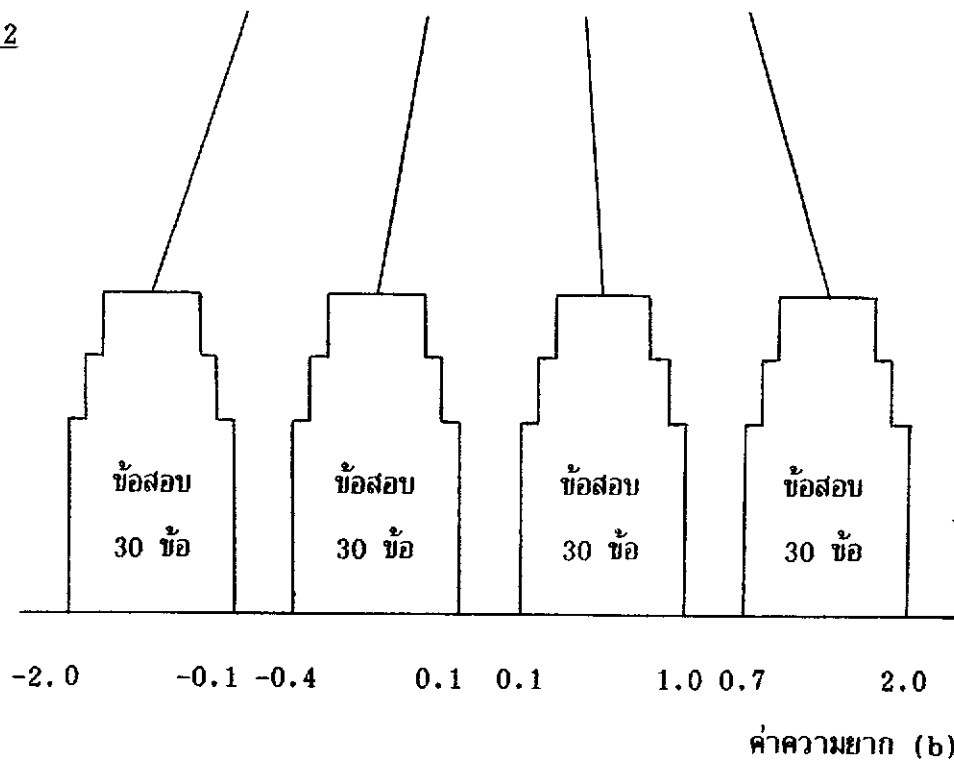
เมื่อ  $m$  แทน จำนวนข้อสอบที่คาบเกี่ยวกัน  
 $L$  แทน จำนวนฉบับของแบบทดสอบวัดผล  
 $n$  แทน จำนวนข้อสอบทั้งหมด  
 $n_1$  แทน จำนวนข้อสอบแต่ละฉบับย่อย ๆ

แบบทดสอบสองขั้นตอนมีหลายลักษณะตามรูปแบบต่าง ๆ ของแบบทดสอบแยกกลุ่ม  
 ดังตัวอย่างภาพประกอบ 5, 6 และ 7 (Weiss. 1974 : 4, 6, 9)

ขั้นที่ 1



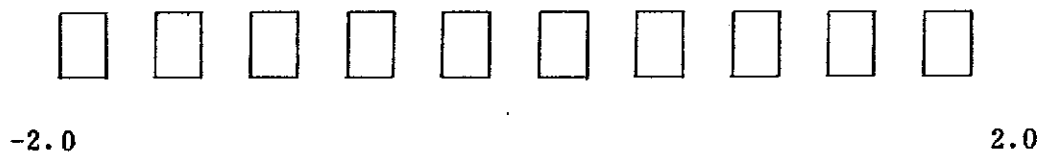
ขั้นที่ 2



ภาพประกอบ 5 โครงสร้างของแบบทดสอบสองขั้นตอน ที่มีแบบทดสอบแยกกลุ่มมีค่าความยากใกล้เคียงกัน

ชั้นที่ 1

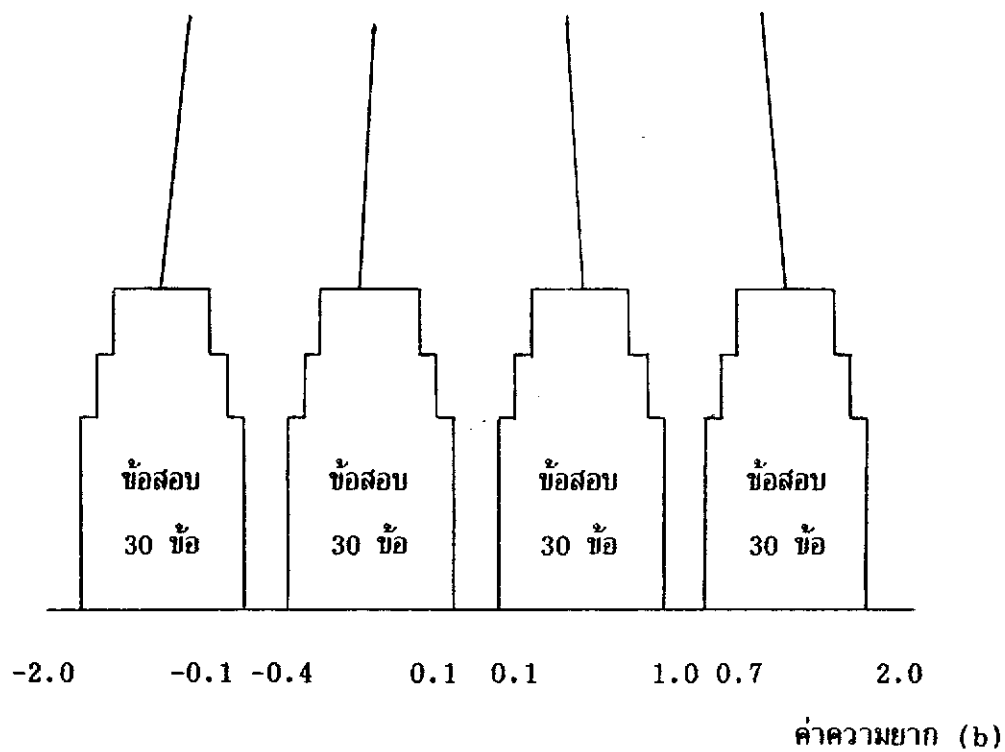
ข้อสอบ 10 ข้อ



ค่าความยาก (b)

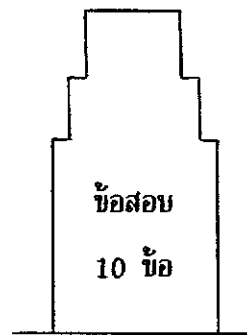
ชั้นที่ 2

ทำถูก  
0 - 3 ข้อ    ทำถูก  
4 - 5 ข้อ    ทำถูก  
6 - 7 ข้อ    ทำถูก  
8 - 10 ข้อ



ภาพประกอบ 6 โครงสร้างของแบบทดสอบสองชั้นคอน ที่มีแบบทดสอบแยกกลุ่มมีค่าความยากกระจายในช่วงกว้าง

ชั้นที่ 1



-0.5      0.5      ค่าความยาก (b)  
ทำถูก      ทำถูก

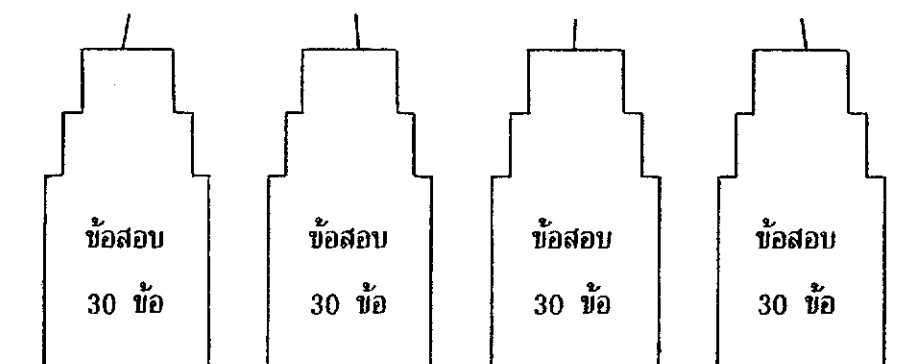
0 - 4 ข้อ

5 - 10 ข้อ



ชั้นที่ 2

-1.0      0.0      0.0      1.0      ค่าความยาก (b)  
ทำถูก      ทำถูก      ทำถูก      ทำถูก  
0 - 4 ข้อ      5 - 10 ข้อ      0 - 4 ข้อ      5 - 10 ข้อ



-2.0      -0.1      -0.4      0.1      0.1      1.0      0.7      2.0

ค่าความยาก (b)

ภาพประกอบ 7 โครงสร้างของแบบทดสอบสองขั้นตอน ที่มีแบบทดสอบแยกกลุ่มสองครั้ง  
หรือแบบทดสอบสามขั้นตอน

### แบบทดสอบเฟลิกซ์เลเวล

แบบทดสอบเฟลิกซ์เลเวล จัดเป็นยุทธวิธีหลายขั้นตอนของการทดสอบแบบ เทเลอร์ มีลักษณะเป็นรูปแบบแยกทางคงที่ โดย ลอร์ด (Weiss. 1974 : 36; citing Lord) ได้เสนอขึ้นเพื่อใช้กับการทดสอบโดยการเขียนตอบแบบเดิม ซึ่งปรับมาจากแบบทดสอบ รูปปิรามิดเพื่อให้มีจำนวนข้อสอบในแบบทดสอบน้อยลง จึงทำให้ในแต่ละระดับความยากจะมีข้อสอบเพียงข้อเดียวเท่านั้น (Hambleton and Swaminathan. 1985 : 300 ; citing Betz and Weiss. 1975) สำหรับตัวข้อสอบของแบบทดสอบเฟลิกซ์เลเวลจะแบ่ง ข้อสอบออกเป็น 2 กลุ่ม คือ ข้อสอบกลุ่มง่ายและข้อสอบกลุ่มยาก แล้วจะคัดเลือกข้อสอบที่มีค่า ความยากตรงกึ่งกลางของข้อสอบทั้งหมดในแบบทดสอบ เป็นข้อเริ่มต้นที่ผู้สอบทุกคนต้องทำ ภายในข้อสอบกลุ่มง่ายจะเรียงลำดับความยากของข้อสอบจากข้อที่ยากที่สุดภายในกลุ่ม ไปหาข้อ ที่ง่ายที่สุด ซึ่งจะพิมพ์ด้วยหมึกสีแดง ส่วนข้อสอบในกลุ่มยากจะเรียงลำดับความยากของข้อสอบ จากข้อที่ง่ายที่สุดภายในกลุ่ม ไปหาข้อที่ยากที่สุด ซึ่งจะพิมพ์ด้วยหมึกสีน้ำเงิน ดังภาพประกอบ 8 (Lord. 1980 : 114)

| ข้อ 0. (ข้อสอบที่มีระดับความยากตรงมัธยฐาน) |                                     |
|--------------------------------------------|-------------------------------------|
| ข้อสอบกลุ่มง่าย<br>(พิมพ์สีแดง)            | ข้อสอบกลุ่มยาก<br>(พิมพ์สีน้ำเงิน)  |
| ข้อ 1. (ข้อที่ยากที่สุดในกลุ่มง่าย)        | ข้อ 1. (ข้อที่ง่ายที่สุดในกลุ่มยาก) |
| ข้อ 2.                                     | ข้อ 2.                              |
| ข้อ 3.                                     | ข้อ 3.                              |
| .                                          | .                                   |
| .                                          | .                                   |
| ข้อ 37. (ข้อที่ง่ายที่สุดในกลุ่มง่าย)      | ข้อ 37. (ข้อที่ยากที่สุดในกลุ่มยาก) |

ภาพประกอบ 8 แสดงตัวอย่างข้อสอบในแบบทดสอบเฟลิกซ์เลเวล

การดำเนินการสอบจะเริ่มต้นด้วยการให้ผู้สอบทำข้อสอบข้อแรกที่ยากที่สุด ความยากตรงมัธยมฐานของข้อสอบทั้งหมดในแบบทดสอบ เมื่อผู้สอบทำข้อสอบข้อแรกถูกข้อต่อไป จะได้ทำข้อสอบที่มีความยากเพิ่มมากขึ้น แต่เมื่อทำข้อสอบข้อแรกผิดข้อต่อไปจะได้ทำข้อสอบที่มีระดับความยากน้อยกว่าข้อแรก สำหรับข้อต่อไป ผู้สอบจะต้องทำข้อสอบในลักษณะนี้ จนถึงที่สุดการทดสอบ ซึ่งถ้าให้  $N$  เป็นจำนวนข้อสอบของแบบทดสอบดั้งเดิม ผู้ที่ผ่านแบบทดสอบเพื่อกีฬาสีเลเวล จะทำข้อสอบเพียง  $(N + 1)/2$  ข้อ (Lord, 1971a : 805)

กระดาคำตอบที่ใช้กับแบบทดสอบเพื่อกีฬาสีเลเวลจะต้องออกแบบมาเป็นพิเศษ เพื่อให้ให้ผู้สอบทราบผลการตอบข้อสอบในแต่ละข้อโดยทันทีว่าถูกหรือผิด กล่าวคือ ถ้าผู้สอบตอบถูกจุดนี้เงินจะปรากฏขึ้น แล้วผู้สอบจะต้องไปตอบข้อสอบที่มีเลขน้อยที่สุด ซึ่งยังไม่ได้ตอบในกลุ่มยาก (พิมพ์สีน้ำเงิน) และถ้าผู้สอบตอบผิดจุดสีแดงจะปรากฏขึ้น แล้วผู้สอบจะต้องไปตอบข้อสอบที่มีเลขน้อยที่สุด ซึ่งยังไม่ได้ตอบในกลุ่มง่าย (พิมพ์สีแดง) เป็นข้อต่อไป (Lord, 1980 : 116)

### 3. วิธีการให้คะแนนของการทดสอบแบบเทเลอร์

การทดสอบแบบเทเลอร์ ผู้สอบแต่ละคนจะได้ทำข้อสอบที่มีความแตกต่างกัน ทั้งจำนวนข้อสอบ ค่าพารามิเตอร์ของข้อสอบ และเวลาในการดำเนินการสอบ การให้คะแนนเพื่อประมาณค่าความสามารถของผู้สอบโดยการนับจำนวนข้อสอบที่ตอบถูก ซึ่งเป็นวิธีการที่เราใช้กันทั่วไปนั้น เป็นวิธีการที่ไม่เหมาะสม ดังนั้นวิธีการให้คะแนนของการทดสอบแบบเทเลอร์ จึงต้องใช้วิธีใหม่ โดยการนำทฤษฎีการตอบสนองข้อสอบมาประยุกต์ใช้ในการประมาณค่าความสามารถของผู้สอบ เนื่องจากการให้คะแนนของการทดสอบแบบเทเลอร์มีพื้นฐานมาจากค่าความยากของข้อสอบที่ผู้สอบแต่ละคนตอบ ซึ่งสามารถจัดให้อยู่บนมาตราวัดเดียวกันได้ (Hambleton and Swaminathan, 1985 : 296) ถึงแม้ว่าผู้สอบจะตอบข้อสอบที่แตกต่างกันก็ตาม สำหรับวิธีการให้คะแนนที่สำคัญมีดังนี้ (Weiss, 1974 : 7 ; Weiss and Kingsbury, 1984 : 363)

3.1 ให้คะแนนตามค่าความยากของข้อที่ยากที่สุดที่ตอบถูก (The most difficulty of item answered correctly) เป็นการให้คะแนนโดยดูจากข้อสอบในแบบทดสอบที่ผู้สอบทำมาทั้งหมด ข้อใดมีค่าความยากสูงสุดที่ทำถูก แสดงว่าเป็นความสามารถของผู้สอบ

3.2 ให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ตอบถูก (Average Difficulty of All Items Answered Correctly) เป็นการให้คะแนนโดยนำเอาค่าความยากของข้อสอบในแบบทดสอบที่ผู้สอบตอบถูกมาหาค่าเฉลี่ย ค่าเฉลี่ยความยากที่คำนวณได้ แสดงว่าเป็นค่าความสามารถของผู้สอบ

3.3 ให้คะแนนตามค่าเฉลี่ยความยากของทุกข้อที่ตอบ (Average Difficulty) เป็นการให้คะแนนที่นำเอาค่าความยากของข้อสอบทุกข้อที่ผู้สอบได้ทำมาหาค่าเฉลี่ย โดยไม่คำนึงว่าผู้สอบจะตอบถูกหรือผิด ค่าเฉลี่ยความยากที่คำนวณได้ แสดงว่าเป็นค่าความสามารถของผู้สอบ

3.4 ให้คะแนนตามค่าความยากของข้อสุดท้ายที่ตอบ (Difficulty of the final item) ซึ่งจะไม่คำนึงว่าข้อสุดท้ายจะตอบถูกหรือผิด ค่าความยากของข้อสอบสุดท้ายที่ผู้สอบได้ทำ แสดงว่าเป็นค่าความสามารถของผู้สอบ

3.5 ให้คะแนนด้วยวิธีความเป็นไปได้สูงสุด (Maximum Likelihood Method) เป็นการประมาณค่าความสามารถโดยการนำเอาคะแนนที่ผู้สอบทำได้ทั้งหมดจากแบบทดสอบ ไปประมาณค่าโดยใช้เครื่องคอมพิวเตอร์ ซึ่งมีหลายวิธีย่อย ๆ เช่น วิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข วิธีความเป็นไปได้สูงสุดแบบไม่มีเงื่อนไข เป็นต้น แต่การประมาณค่าด้วยวิธีนี้มีข้อจำกัดในการตีค่าผู้สอบทำข้อสอบถูกหมดหรือผิดหมด จะไม่สามารถประมาณค่าได้

3.6 ให้คะแนนด้วยวิธีของเบย์ส์ (Bayesian Method) เป็นการประมาณค่าความสามารถที่พัฒนาขึ้นมาเพื่อแก้ปัญหาเกี่ยวกับข้อจำกัดของการประมาณค่าด้วยวิธีความเป็นไปได้สูงสุด ซึ่งจะทำให้มีความถูกต้องแม่นยำมากกว่า แต่วิธีนี้มีการใช้สถิติค่อนข้างซับซ้อน จึงทำให้ไม่มีการนำมาใช้กันอย่างแพร่หลายมากนัก

ในการศึกษาครั้งนี้ ผู้วิจัยนำผลการสอบจากแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน (แบบทดสอบวัดผล) และแบบทดสอบแฟล็กชิเลเวล ไปประมาณค่าความสามารถของผู้สอบ และพารามิเตอร์ค่าความยากของข้อสอบ ตามทฤษฎีการตอบสนองข้อสอบ แบบราส์ซไคเดล ด้วยวิธีการประมาณค่าสูงสุดแบบไม่มีเงื่อนไข โดยใช้โปรแกรมสำเร็จรูปไบคาล

## งานวิจัยที่เกี่ยวข้อง

### 1. งานวิจัยภายในประเทศ

งามนิตย์ ธาตุทอง และปรีชา เครือวรรณ (2527) ได้สร้างและประเมินผลการใช้แบบทดสอบแฟล็กชีเลเวล ในวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 โดยใช้แบบทดสอบแฟล็กชีเลเวล 1 ฉบับ ที่มี 18 ชิ้น มีข้อสอบทั้งหมด 35 ข้อ แต่นักเรียนจะทำเพียง 18 ข้อ และแบบทดสอบดั้งเดิม 1 ฉบับ มีข้อสอบทั้งหมด 18 ข้อ จากผลการศึกษาพบว่าความสัมพันธ์ระหว่างคะแนนแบบทดสอบแฟล็กชีเลเวลกับคะแนนผลสัมฤทธิ์ทางการเรียน มีค่าเท่ากับ 0.7238 และความสัมพันธ์ระหว่างคะแนนจากแบบทดสอบดั้งเดิมกับคะแนนผลสัมฤทธิ์ทางการเรียน มีค่าเท่ากับ 0.6109 เมื่อทดสอบความแตกต่างระหว่างค่าสัมประสิทธิ์สหสัมพันธ์ดังกล่าวปรากฏว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และผู้สอบส่วนมากมีทัศนคติที่ดีต่อการทดสอบด้วยแบบทดสอบแฟล็กชีเลเวล

เสรี พรหมสวัสดิ์โรจน์ (2531) ได้ศึกษาความเที่ยงตรงในการจำแนกความรอบรู้ และการประมาณค่าความสามารถ ในวิชาคณิตศาสตร์ของแบบทดสอบแฟล็กชีเลเวลและแบบทดสอบรูปปริมาตรที่ใช้ขนาดขั้นแปรรผัน โดยใช้แบบทดสอบแฟล็กชีเลเวล จำนวน 2 ฉบับ ขนาด 6 ชิ้น และ 12 ชิ้น และแบบทดสอบรูปปริมาตรขนาดขั้นแปรรผัน จำนวน 2 ฉบับ ขนาด 6 ชิ้น และ 12 ชิ้น กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 5,987 คน ผลการศึกษาพบว่า ความเที่ยงตรงในการจำแนกความรอบรู้เมื่อใช้ข้อสอบที่สร้างขึ้นเป็นเกณฑ์ ปรากฏว่า แบบทดสอบแฟล็กชีเลเวลและแบบทดสอบรูปปริมาตรขนาดขั้นแปรรผันทุกฉบับ มีความเที่ยงตรงในการจำแนกความรอบรู้ได้อย่างมีนัยสำคัญทางสถิติ และเมื่อเปรียบเทียบค่าความเที่ยงตรงแต่ละฉบับพบว่า แบบทดสอบแฟล็กชีเลเวลทั้ง 2 ฉบับ มีความเที่ยงตรงในการจำแนกความรอบรู้ได้สูงกว่าแบบทดสอบรูปปริมาตรอยู่ 1 ฉบับ อย่างมีนัยสำคัญทางสถิติ

เกรียงศักดิ์ สุวรรณภาค (2532) ได้ศึกษาความสัมพันธ์ระหว่างรูปแบบบางรูปแบบของแบบทดสอบแยกกลุ่มกับความสามารถทางการเรียน ในวิชาคณิตศาสตร์ โดยทำการศึกษา 2 ตัวแปร คือ ตัวแปรแรกเป็นรูปแบบของแบบทดสอบกลุ่มซึ่งมี 3 รูปแบบ ได้แก่ แบบทดสอบแยกกลุ่มที่มีค่าความยากของข้อสอบใกล้เคียงกัน แบบทดสอบแยกกลุ่มที่มีค่าความยากของข้อสอบกระจายจากน้อยที่สุดถึงมากที่สุด และแบบทดสอบแยกกลุ่มที่มีการแยกกลุ่ม 2 ครั้ง และตัวแปรที่สอง คือ วิธีการให้คะแนนซึ่งมี 2 วิธี ได้แก่ วิธีการให้คะแนนตามค่าความยาก

ของข้อที่ยากที่สุดที่ตอบถูก และวิธีการให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ตอบถูก โดยทำการศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 2,455 คน ผลการศึกษาพบว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผลการสอบที่ใช้แบบทดสอบแยกกลุ่มแต่ละฉบับ ซึ่งตรวจให้คะแนนในแต่ละวิธี กับความสามารถทางการเรียน มีค่าเป็นบวกอย่างมีนัยสำคัญทางสถิติทุกค่า และผลการทดสอบความแตกต่างของค่าสัมประสิทธิ์สหสัมพันธ์ดังกล่าวเป็นรายคู่พบว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติทุกคู่ สำหรับค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจให้คะแนนทั้ง 2 วิธี ดังกล่าว โดยแยกวิเคราะห์ตามแบบทดสอบแยกกลุ่มแต่ละรูปแบบ มีค่าเป็นบวกอย่างมีนัยสำคัญทางสถิติทุกค่า และผลการทดสอบความแตกต่างของค่าสัมประสิทธิ์สหสัมพันธ์ดังกล่าวเป็นรายคู่พบว่า ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการให้คะแนนทั้ง 2 วิธี เมื่อพิจารณาในแบบทดสอบแยกกลุ่มแต่ละรูปแบบ แตกต่างกันอย่างมีนัยสำคัญทางสถิติทุกคู่

นริศ สวัสดิ์ (2533) ได้ศึกษาความสัมพันธ์ระหว่างคะแนนของแบบทดสอบสองขั้นตอนและคะแนนของแบบทดสอบรูปปริมาตร โดยใช้แบบทดสอบสองขั้นตอนซึ่งเป็นแบบทดสอบแยกกลุ่ม จำนวน 15 ข้อ และแบบทดสอบวัดผล 4 ฉบับ ๆ ละ 20 ข้อ ส่วนแบบทดสอบรูปปริมาตรที่ใช้เป็นแบบทดสอบที่มีขนาดขึ้นแปรผัน ขนาด 4 ชั้น 8 ชั้น และ 12 ชั้น อย่างละ 1 ฉบับ และใช้วิธีการให้คะแนนตามค่าความยากเฉลี่ยของข้อสอบที่ตอบถูก โดยทำการศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน 815 คน ผลการศึกษาพบว่า ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนของแบบทดสอบสองขั้นตอนกับคะแนนของแบบทดสอบรูปปริมาตรทั้ง 3 ขนาด มีความสัมพันธ์ทางบวกอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกค่า

## 2. งานวิจัยในต่างประเทศ

ลินน์ ร็อค และเคลียร์ (Linn, Rock and Cleary, 1969) ได้ศึกษาวิธีการสร้างแบบทดสอบโปรแกรม (Program Test) เพื่อประเมินความแตกต่างของยุทธวิธีแยกทาง 7 วิธี คือ แบบทดสอบแยกกลุ่ม 2 ครั้ง (Two - stage routing test) แบบทดสอบแยกกลุ่มพิสัยกว้าง (Broad range routing test) แบบทดสอบแยกกลุ่มที่ใช้กลุ่มจำแนก (Group discrimination routing test) แบบสุ่มตามลำดับข้อสอบ 4 กลุ่ม (Four - group sequential item sampling) แบบทดสอบรูปปริมาตร ขนาด 10 ชั้น (Ten item branching test) และแบบทดสอบรูปปริมาตร ขนาด 5 ชั้น โดยแต่ละชั้นมีข้อสอบอยู่ 5 ข้อ (Twenty five item blocks of five branching test) ในการ

ศึกษาครั้งนี้ใช้กลุ่มตัวอย่าง จำนวน 4,885 คน โดยสร้างกลุ่มข้อสอบ จำนวน 190 ข้อ ผลการศึกษาพบว่า การใช้แบบทดสอบโปรแกรมโดยจัดให้ผู้สอบได้ตอบข้อสอบที่มีค่าความยากเหมาะสมกับระดับความสามารถของผู้สอบ คะแนนที่ได้มีความสัมพันธ์กับเกณฑ์การวัดสูง นอกจากนี้ยังใช้ข้อสอบและเวลาในการทดสอบน้อยกว่าการใช้แบบทดสอบดั้งเดิม ในขณะที่ค่าความเที่ยงตรงยังมีค่าเท่าเดิมหรือสูงกว่า

ลอร์ด (Lord. 1971a) ได้ทำการประเมินคุณภาพของแบบทดสอบแฟล็กชีเลเวล โดยใช้เครื่องคอมพิวเตอร์จำลองสถานการณ์ การประมาณค่าพารามิเตอร์ใช้ทฤษฎีการตอบสนองข้อสอบ แบบนอร์มัล ออโจฟ โมเดล ที่มีพารามิเตอร์ 3 ตัว ในการศึกษาครั้งนี้ ได้ตั้งเกณฑ์ในการคัดเลือกข้อสอบที่มีค่าอำนาจจำแนกเท่ากับ 0.50 โดยนัยแบบทดสอบดั้งเดิม จำนวน 60 ข้อ มาเปรียบเทียบกับคุณภาพกับแบบทดสอบแฟล็กชีเลเวล จำนวน 119 ข้อ ซึ่งผู้สอบจะทำข้อสอบเพียง 60 ข้อ เท่ากับแบบทดสอบดั้งเดิม ผลการศึกษาพบว่าเมื่อแบบทดสอบแฟล็กชีเลเวลมีขนาดของช่วงห่างระหว่างค่าความยากของข้อสอบ ( $d$ ) ใหญ่ขึ้นจะทำให้ค่าฟังก์ชันอินฟอร์เมชันของผู้สอบลดลงทุกระดับความสามารถ และแบบทดสอบจะมีประสิทธิภาพสูงสุดเมื่อ  $d$  มีค่าอยู่ระหว่าง 0.033 ถึง 0.067 แต่จะ ไม่มีประสิทธิภาพเมื่อ  $d$  มีค่าใกล้เคียงกับ 0.1 สำหรับการเปรียบเทียบค่าฟังก์ชันอินฟอร์เมชันระหว่างแบบทดสอบแฟล็กชีเลเวลกับแบบทดสอบดั้งเดิมพบว่า แบบทดสอบแฟล็กชีเลเวลจะให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่าแบบทดสอบดั้งเดิม เมื่อผู้สอบมีความสามารถในระดับสูงหรือต่ำ ส่วนผู้สอบที่มีความสามารถในระดับปานกลาง แบบทดสอบดั้งเดิมจะให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่าแบบทดสอบแฟล็กชีเลเวล

เวลล์ (Vale. 1975) ได้รายงานผลการประเมินประสิทธิภาพของแบบทดสอบแฟล็กชีเลเวล แบบทดสอบสองขั้นตอน แบบทดสอบสามขั้นตอน แบบทดสอบรูปปริมาตร แบบทดสอบปรับระดับขึ้น ยุทธวิธีของเบส และแบบทดสอบดั้งเดิม โดยทุกฉบับให้ผู้สอบตอบจำนวน 8 ข้อ สำหรับกลุ่มข้อสอบที่นำมาสร้างมี จำนวน 68 ข้อ แล้วแบ่งออกเป็น 17 กลุ่มซึ่งแต่ละกลุ่มประกอบด้วยข้อสอบ จำนวน 4 ข้อ ในการศึกษาครั้งนี้ได้ใช้เครื่องคอมพิวเตอร์จำลองสถานการณ์ โดยมีข้อตกลงเบื้องต้น 4 ประการ คือ ข้อสอบทั้งหมดอยู่ในเกณฑ์ที่มีค่าอำนาจจำแนกเท่ากัน ไม่มีการเดา ให้คะแนนด้วยวิธีของเบส และค่าประมาณความสามารถเริ่มต้นของผู้สอบมีค่าเท่ากับ 0.5 สำหรับเกณฑ์ที่ใช้ประเมินแบบทดสอบคือ ค่าฟังก์ชัน

อินฟอร์เมชัน ผลการศึกษาพบว่า แบบทดสอบแฟล็กซีเลเวลที่ประกอบด้วยข้อสอบที่มีค่าความยากกระจายตั้งแต่ง่ายที่สุด ไปจนถึงยากที่สุด จะให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่าแบบทดสอบแบบดั้งเดิม ทุกระดับความสามารถ และแบบทดสอบแฟล็กซีเลเวลที่ประกอบด้วยข้อสอบที่มีค่าความยากปานกลางจะให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่าแบบทดสอบแบบดั้งเดิม เมื่อผู้สอบมีความสามารถในระดับสูงหรือต่ำมาก ๆ เท่านั้น แต่จะให้ค่าฟังก์ชันอินฟอร์เมชันต่ำกว่าแบบทดสอบดั้งเดิม เมื่อผู้สอบมีความสามารถในระดับเฉลี่ยใกล้ ๆ ศูนย์ สำหรับการเปรียบเทียบกับการทดสอบแบบเทเลอร์อีก 5 รูปแบบนั้น พบว่า แบบทดสอบแฟล็กซีเลเวลจะให้ค่าฟังก์ชันอินฟอร์เมชันต่ำกว่าทั้ง 5 รูปแบบ ทุกระดับความสามารถ

ซีควิน (Sequin, 1978) ได้ศึกษาเปรียบเทียบแบบทดสอบแฟล็กซีเลเวลกับแบบทดสอบดั้งเดิม โดยใช้กลุ่มตัวอย่าง จำนวน 1,317 คน เพื่อตอบข้อสอบ 2 กลุ่มวิชาคือ วิชาภาษาฝรั่งเศส จำนวน 35 ข้อ และวิชาคณิตศาสตร์ จำนวน 27 ข้อ แล้วนำไปวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบโดยใช้ทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบแบบนอร์มัล ออโจฟ โมเดล ที่มีพารามิเตอร์ 3 ตัว แล้วนำข้อสอบทั้ง 2 กลุ่มมาสร้างเป็นแบบทดสอบแฟล็กซีเลเวล จำนวน 6 ฉบับ และแบบทดสอบดั้งเดิม จำนวน 6 ฉบับ โดยการจัดโครงสร้างของแบบทดสอบทั้ง 12 ฉบับ ได้ยึดกระบวนการคัดเลือกข้อสอบ 3 ประการ คือ การเลือกข้อสอบที่มีค่าอำนาจจำแนกสูงสุด การคัดเลือกข้อสอบที่มีค่าความยากใกล้เคียงกับระดับความสามารถเฉลี่ยของกลุ่มตัวอย่าง และการคัดเลือกข้อสอบที่มีค่าความยากกระจายในช่วงกว้าง นอกจากนี้ยังได้นำข้อสอบในกลุ่มมาสร้างเป็นแบบทดสอบต้นฉบับ  $N$  ข้อ กับแบบทดสอบที่ได้จากการลดจำนวนข้อสอบในแบบทดสอบต้นฉบับให้เหลือ  $n$  ข้อ สำหรับแบบทดสอบทุกฉบับที่สร้างขึ้นมานี้ใช้วิธีคำนวณการสอบ โดยใช้สถานการณ์จำลอง ผลการศึกษาพบว่า แบบทดสอบแฟล็กซีเลเวลที่เรียงข้อสอบตามค่าความยากของข้อสอบ แบบนอร์มัล ออโจฟ โมเดล ที่มีพารามิเตอร์ 3 ตัว ดีกว่าแบบทดสอบแฟล็กซีเลเวลที่เรียงข้อสอบตามค่าสัดส่วนการตอบถูกต้องอย่างมีนัยสำคัญทางสถิติ แต่ไม่แตกต่างกับแบบทดสอบชนิดเดียวกันที่เรียงข้อสอบตามความสามารถของผู้สอบ ที่ข้อสอบแต่ละข้อให้ค่าฟังก์ชันอินฟอร์เมชันสูงสุด สำหรับแบบทดสอบดั้งเดิมที่เรียงข้อสอบตามค่าอำนาจจำแนกสูงสุด จะให้ผลการวัดที่น่าพอใจมากที่สุด นอกจากนี้ซีควิน ยังได้เสนอแนะว่า แบบทดสอบแฟล็กซีเลเวลไม่เหมาะสมที่จะใช้ประมาณค่าความสามารถของผู้สอบที่มีความสามารถในระดับปานกลาง แต่จะใช้ได้ดีกับผู้สอบที่มีความสามารถในระดับสูงหรือต่ำมาก ๆ

แองกอฟฟ์ และฮัดเดิลสตัน (Angoff and Huddleston. 1958) ได้ทำการตรวจสอบแบบทดสอบสองขั้นตอนพบว่า แบบทดสอบสองขั้นตอนมีเทคนิคเหนือกว่าแบบทดสอบดั้งเดิมทั่ว ๆ ไป แต่การนำแบบทดสอบสองขั้นตอนไปใช้ยังไม่เพียงพอ ที่จะยอมรับเนื่องจากยังมีปัญหาเกี่ยวกับการดำเนินการสอบ สำหรับการศึกษาแบบทดสอบสองขั้นตอนของทั้งสองคนนี้ มุ่งเน้นในเรื่องความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบ มากกว่าที่จะมุ่งเน้นในเรื่องของการลดเวลาในการสอบ

โอลิเวียร์ (Olivier. 1974) ได้ศึกษาเปรียบเทียบขนาดของความคลาดเคลื่อนในการวัดที่เกิดจากการใช้แบบทดสอบแฟล็กซีเลเวลและแบบทดสอบดั้งเดิม โดยทำการศึกษาในสถานการณ์การทดสอบจริง ผู้วิจัยได้นำแบบทดสอบทั้ง 2 รูปแบบ ซึ่งมีความยาวเท่ากัน ไปทดสอบกับกลุ่มนักเรียน 2 กลุ่ม โดยแต่ละกลุ่มทำแบบทดสอบเพียงรูปแบบเดียว แล้วนำผลการสอบไปวิเคราะห์หาค่าคุณภาพของข้อสอบ ผลการศึกษาพบว่า แบบทดสอบแฟล็กซีเลเวลให้ค่าความคลาดเคลื่อนมาตรฐานของการวัดสูงกว่าแบบทดสอบดั้งเดิม นอกจากนี้ยังพบว่า แบบทดสอบแฟล็กซีเลเวลมีปัญหามากมาย ๆ ประการที่เกี่ยวกับการสร้างข้อสอบและการดำเนินการสอบ

จากเอกสารและผลงานวิจัยที่เกี่ยวข้องดังที่กล่าวมาแล้วนั้น ส่วนมากจะพบว่า แบบทดสอบสองขั้นตอนและแบบทดสอบแฟล็กซีเลเวลมีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบได้ดีกว่าแบบทดสอบดั้งเดิม เมื่อผู้สอบมีความสามารถในระดับสูงหรือต่ำมาก ๆ แต่ถ้าผู้สอบมีความสามารถในระดับปานกลาง แล้วแบบทดสอบดั้งเดิมจะมีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบได้ดีกว่า นอกจากนี้แล้วยังพบว่าแบบทดสอบสองขั้นตอนและแบบทดสอบแฟล็กซีเลเวล ใช้จำนวนข้อสอบและเวลาในการดำเนินการสอบน้อยกว่าแบบทดสอบดั้งเดิม ในขณะที่ความเชื่อมั่นและความเที่ยงตรงยังมีค่าเท่าเดิม และผู้สอบส่วนมากมีความพึงพอใจในการทำข้อสอบมากกว่า ด้วยเหตุผลเหล่านี้ ผู้วิจัยจึงมีความสนใจที่จะนำแบบทดสอบทั้ง 2 รูปแบบนี้ มาทำการศึกษาร่วมกับแบบทดสอบดั้งเดิม เพื่อจะได้ทราบว่าสามารถนำมาใช้แทนแบบทดสอบดั้งเดิมได้หรือไม่ โดยศึกษามุ่งเน้นในเรื่องความเที่ยงตรงและประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์

### สมมติฐานของการศึกษาค้นคว้า

จากหลักการของการทดสอบแบบเทเลอร์ และผลงานของวิจัยที่เกี่ยวข้องจะเห็นได้ว่าแบบทดสอบสองขั้นตอนและแบบทดสอบเพิลิกซีเลเวลมีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่าแบบทดสอบดั้งเดิม เมื่อผู้สอบมีความสามารถในระดับสูงหรือต่ำมาก ๆ แต่เมื่อผู้สอบมีความสามารถในระดับปานกลางแล้ว แบบทดสอบดั้งเดิมจะมีประสิทธิภาพในการประมาณค่าความสามารถได้สูงกว่าแบบทดสอบสองขั้นตอนและแบบทดสอบเพิลิกซีเลเวล (Lord, 1971a : 805 - 813 ; Vale, 1975 : 5 - 75 ; Sequin, 1978 : 1196 - A) ดังนั้น ผู้วิจัยจึงตั้งสมมติฐานดังนี้

1. ค่าความเที่ยงตรงของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมกับแบบทดสอบสองขั้นตอน มีค่าแตกต่างกัน
2. ค่าความเที่ยงตรงของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมกับแบบทดสอบเพิลิกซีเลเวล มีค่าแตกต่างกัน
3. ค่าความเที่ยงตรงของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบสองขั้นตอนกับแบบทดสอบเพิลิกซีเลเวล มีค่าแตกต่างกัน
4. ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบสองขั้นตอน มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง ( $-1.00 < \theta < 1.00$ ) แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมา ( $\theta < -1.00$ ) และระดับสูงขึ้นไป ( $\theta > 1.00$ )
5. ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบเพิลิกซีเลเวล มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง ( $-1.00 < \theta < 1.00$ ) แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมา ( $\theta < -1.00$ ) และระดับสูงขึ้นไป ( $\theta > 1.00$ )
6. ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบสองขั้นตอนเทียบกับแบบทดสอบเพิลิกซีเลเวล มีค่าสูงกว่า 1 ในกลุ่มผู้สอบทุกระดับความสามารถ

### บทที่ 3

#### วิธีดำเนินการศึกษาค้นคว้า

##### ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2536 ของโรงเรียนสังกัดกรมสามัญศึกษา ส่วนกลาง กลุ่มที่ 5 (เขตบางกะปิ ลาดพร้าว บางกะปิ ลาดกระบัง หนองจอก และมีนบุรี) กรุงเทพมหานคร จำนวนโรงเรียน 15 โรงเรียน จำนวนห้องเรียน 204 ห้องเรียน จำนวนนักเรียน 10,390 คน

##### กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2536 ของโรงเรียนสังกัดกรมสามัญศึกษา ส่วนกลาง กลุ่มที่ 5 (เขตบางกะปิ ลาดพร้าว บางกะปิ ลาดกระบัง หนองจอก และมีนบุรี) กรุงเทพมหานคร จำนวนโรงเรียน 9 โรงเรียน จำนวนห้องเรียน 9 ห้องเรียน จำนวนนักเรียน 441 คน ซึ่งมีรายละเอียดและขั้นตอนการสุ่มดังนี้

ขั้นที่ 1 กำหนดขนาดกลุ่มตัวอย่างด้วยความเชื่อมั่น 95 เปอร์เซนต์ ( $\alpha = .05$ ) เมื่อเทียบจากตารางขนาดตัวอย่างปรากฏว่าต้องใช้กลุ่มตัวอย่าง 385 คน ประมาณ 9 ห้องเรียน (ล้วน สายยศ และอังกศ สายยศ. 2524 : 287 ; อ้างอิงมาจาก Mercado. 1977)

ขั้นที่ 2 สุ่มโรงเรียนจากทุกเขตในประชากร ด้วยวิธีการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) โดยมีเขตเป็นชั้น (Strata) และโรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit) ได้จำนวนโรงเรียน 9 โรงเรียน จำนวนห้องเรียน 120 ห้องเรียน จำนวนนักเรียน 6,007 คน

ขั้นที่ 3 สุ่มห้องเรียนจากทุกโรงเรียนที่สุ่มได้ในขั้นที่ 2 ด้วยวิธีการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) โดยมีโรงเรียนเป็นชั้น (Strata) และห้องเรียนเป็นหน่วยการสุ่ม (Sampling Unit) ได้จำนวนห้องเรียน 9 ห้องเรียน จำนวนนักเรียน 441 คน ซึ่งมีรายละเอียดดังตาราง 1

ตาราง 1 จำนวนกลุ่มตัวอย่างที่ใช้ในการวิจัย จำแนกตามเขต โรงเรียน จำนวนห้องเรียน และจำนวนนักเรียน

| เขต       | โรงเรียน                         | จำนวนห้องเรียน | จำนวนนักเรียน |
|-----------|----------------------------------|----------------|---------------|
| บางกะปิ   | 1. บางกะปิ                       | 1              | 53            |
|           | 2. มัธยมวัดบึงทองหลาง            | 1              | 55            |
|           | 3. เทพลีลา                       | 1              | 45            |
| ลาดพร้าว  | 4. ลาตปลาเค้าพิทยาคม             | 1              | 48            |
| ปทุม      | 5. นวมินทร์ราชูทิศ กรุงเทพมหานคร | 1              | 46            |
| ลาดกระบัง | 6. เทพศิรินทร์ร่มเกล้า           | 1              | 50            |
|           | 7. รัตนโกสินทร์สุม ไทลลาดกระบัง  | 1              | 49            |
| หนองจอก   | 8. วัดหนองจอก                    | 1              | 45            |
| มีนบุรี   | 9. เศรษฐบุรพาพิทย                | 1              | 50            |
|           | รวม                              | 9              | 441           |

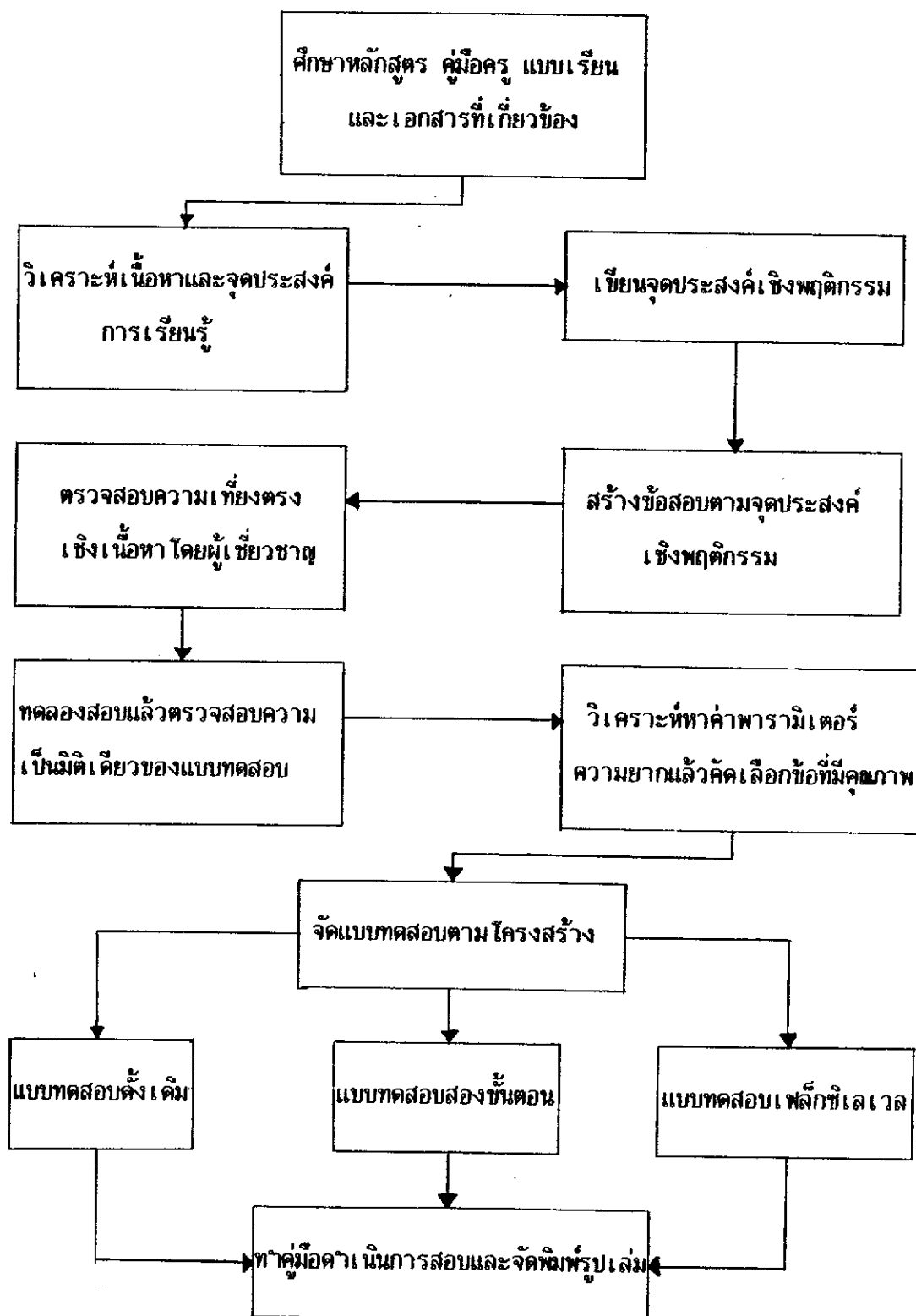
#### เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบเลือกตอบวัดผลสัมฤทธิ์ทางการเรียน ชนิด 5 ตัวเลือก ซึ่งผู้วิจัยสร้างขึ้นเองจากเนื้อหาในวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3 (ค 011) เรื่องเลขยกกำลัง ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) โดยมี 3 รูปแบบ ดังนี้

1. แบบทดสอบดั้งเดิม มีจำนวน 30 ข้อ
2. แบบทดสอบสองขั้นตอน ได้แก่
  - 2.1 แบบทดสอบแยกกลุ่ม มีจำนวน 10 ข้อ
  - 2.2 แบบทดสอบวัดผล มีจำนวน 30 ข้อ
3. แบบทดสอบแฟล็กซีเลเวล มีจำนวน 59 ข้อ

### วิธีดำเนินการสร้างแบบทดสอบ

การสร้างแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กซีเลเวล  
ผู้วิจัยสร้างตามแผนผัง ดังนี้



ภาพประกอบ 9 แผนผังแสดงขั้นตอนการดำเนินการสร้างแบบทดสอบ

รายละเอียดในการสร้างแบบทดสอบ มีดังนี้

# 1. การสร้างกลุ่มข้อสอบ

1.1 ศึกษาเนื้อหาจากหลักสูตร คู่มือครู แบบเรียน และเอกสารที่เกี่ยวข้อง  
ในวิชาคณิตศาสตร์ (ค 011) ชั้นมัธยมศึกษาปีที่ 3 เรื่องเลขยกกำลัง

1.2 วิเคราะห์จุดประสงค์การเรียนรู้และเนื้อหา เพื่อนำมาเขียนเป็น  
จุดประสงค์เชิงพฤติกรรม โดยสามารถเขียนได้ 9 จุดประสงค์ ซึ่งมีรายละเอียดดังตาราง 16  
ในภาคผนวก จ

1.3 สร้างข้อสอบเลือกตอบ ชนิด 5 ตัวเลือก โดยสร้างตามจุดประสงค์  
เชิงพฤติกรรม ซึ่งจะสร้างข้อสอบได้ทั้งหมด จำนวน 270 ข้อ โดยมีรายละเอียดดังนี้

จุดประสงค์ที่ 1 สร้างข้อสอบ จำนวน 36 ข้อ

จุดประสงค์ที่ 2 สร้างข้อสอบ จำนวน 36 ข้อ

จุดประสงค์ที่ 3 สร้างข้อสอบ จำนวน 30 ข้อ

จุดประสงค์ที่ 4 สร้างข้อสอบ จำนวน 30 ข้อ

จุดประสงค์ที่ 5 สร้างข้อสอบ จำนวน 30 ข้อ

จุดประสงค์ที่ 6 สร้างข้อสอบ จำนวน 24 ข้อ

จุดประสงค์ที่ 7 สร้างข้อสอบ จำนวน 30 ข้อ

จุดประสงค์ที่ 8 สร้างข้อสอบ จำนวน 18 ข้อ

จุดประสงค์ที่ 9 สร้างข้อสอบ จำนวน 36 ข้อ

1.4 นำข้อสอบที่สร้างขึ้น 270 ข้อ ไปให้ผู้เชี่ยวชาญซึ่งมีความรู้ในเนื้อหา  
และการเขียนข้อสอบ จำนวน 5 คน เป็นผู้ตรวจสอบความเที่ยงตรงเชิงเนื้อหา (Content  
Validity) โดยการพิจารณาความสอดคล้องระหว่างข้อสอบแต่ละข้อกับจุดประสงค์  
เชิงพฤติกรรม ข้อสอบที่ถือว่าวัด ได้ตรงตามจุดประสงค์เชิงพฤติกรรมจะต้อง ได้รับความเห็นชอบ  
จากผู้เชี่ยวชาญไม่น้อยกว่า 3 คน สำหรับแบบประเมินความสอดคล้องมีลักษณะดังตัวอย่าง  
ต่อไปนี้



จากผลของการตรวจสอบความเที่ยงตรงเชิงเนื้อหาของผู้เชี่ยวชาญ ผลปรากฏว่าในข้อสอบ จำนวน 270 ข้อ ได้รับการพิจารณาว่ามีความสอดคล้องกับจุดประสงค์เชิงพฤติกรรม จำนวน 247 ข้อ ซึ่งแสดงว่ามีข้อสอบที่จะต้องปรับปรุงแก้ไข จำนวน 23 ข้อ เพื่อให้ได้ข้อสอบ จำนวน 270 ข้อ

1.5 ปรับปรุงข้อสอบตามที่ผู้เชี่ยวชาญได้พิจารณาไว้ แล้วนำไปให้ผู้เชี่ยวชาญพิจารณาอีกครั้งหนึ่ง ผลปรากฏว่าได้รับความเห็นชอบทั้งหมด จึงได้ข้อสอบครบตามจำนวน 270 ข้อ แล้วจัดเป็นแบบทดสอบ 6 ฉบับ ๆ ละ 45 ข้อ ซึ่งจัดโดยการสุ่มข้อสอบในแต่ละจุดประสงค์เชิงพฤติกรรม

1.6 นำแบบทดสอบทั้ง 6 ฉบับ ไปทดลองสอบ ในระหว่าง วันที่ 1 เดือน พฤศจิกายน พ.ศ. 2536 ถึงวันที่ 27 เดือนพฤศจิกายน พ.ศ. 2536 กับกลุ่มตัวอย่างที่ใช้ในการวิเคราะห์ค่าพารามิเตอร์ความยากของข้อสอบ ซึ่งเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนสังกัดกรมสามัญศึกษา ส่วนกลาง กลุ่มที่ 5 โดยเป็นนักเรียนคนละกลุ่มกับกลุ่มที่ใช้ในการวิจัย มีจำนวนโรงเรียน 10 โรงเรียน จำนวนห้องเรียน 50 ห้องเรียน และจำนวนนักเรียน 2,403 คน ซึ่งได้มาจากการสุ่มอย่างง่าย (Simple Random Sampling) ดังรายละเอียดต่อไปนี้

|                                     |                                        |
|-------------------------------------|----------------------------------------|
| โรงเรียนบางกะปิ                     | จำนวน 5 ห้องเรียน จำนวนนักเรียน 265 คน |
| โรงเรียนมัธยมวัดบึงทองหลาง          | จำนวน 5 ห้องเรียน จำนวนนักเรียน 230 คน |
| โรงเรียนลาดปลาเค้าพิทยาคม           | จำนวน 5 ห้องเรียน จำนวนนักเรียน 220 คน |
| โรงเรียนนวมินทราชูทิศ กรุงเทพมหานคร | จำนวน 5 ห้องเรียน จำนวนนักเรียน 221 คน |
| โรงเรียนรัตนโกสินทร์สมโภชลาดกระบัง  | จำนวน 5 ห้องเรียน จำนวนนักเรียน 225 คน |
| โรงเรียนเทพศิรินทร์ร่มเกล้า         | จำนวน 5 ห้องเรียน จำนวนนักเรียน 254 คน |
| โรงเรียนวัดหนองจอก                  | จำนวน 5 ห้องเรียน จำนวนนักเรียน 238 คน |
| โรงเรียนเศรษฐบุรพาพิทย              | จำนวน 5 ห้องเรียน จำนวนนักเรียน 242 คน |
| โรงเรียนสตรีเศรษฐบุรพาพิทย          | จำนวน 5 ห้องเรียน จำนวนนักเรียน 245 คน |
| โรงเรียนสตรีวิทยา 2                 | จำนวน 5 ห้องเรียน จำนวนนักเรียน 263 คน |

1.7 ผู้วิจัยได้สุ่มนักเรียนในห้องเรียนแต่ละคนเพื่อทำแบบทดสอบแต่ละฉบับ โดยการจัดการดำเนินการสอบให้เป็นระบบดังนี้ คือ ให้นักเรียนคนที่ 1, 2, 3, 4, 5 และ 6 ทำแบบทดสอบฉบับที่ 1, 2, 3, 4, 5 และ 6 ตามลำดับ และนักเรียนคนที่ 7, 8, 9, 10, 11 และ 12 จะทำแบบทดสอบฉบับที่ 1, 2, 3, 4, 5 และ 6 ตามลำดับ เหมือนเดิม และคนต่อ ๆ ไป ก็จะได้ทำแบบทดสอบในลักษณะนี้ ซึ่งจะทำให้ได้นักเรียนทำแบบทดสอบแต่ละฉบับมีจำนวนใกล้เคียงกัน คือ

ฉบับที่ 1 มีนักเรียนทำแบบทดสอบ จำนวน 400 คน

ฉบับที่ 2 มีนักเรียนทำแบบทดสอบ จำนวน 401 คน

ฉบับที่ 3 มีนักเรียนทำแบบทดสอบ จำนวน 399 คน

ฉบับที่ 4 มีนักเรียนทำแบบทดสอบ จำนวน 403 คน

ฉบับที่ 5 มีนักเรียนทำแบบทดสอบ จำนวน 400 คน

ฉบับที่ 6 มีนักเรียนทำแบบทดสอบ จำนวน 400 คน

1.8 นำผลการสอบมาตรวจให้คะแนน โดยข้อที่ตอบถูกให้ 1 คะแนน และ ข้อที่ตอบผิดให้ 0 คะแนน

1.9 นำผลจากข้อ 1.8 มาวิเคราะห์องค์ประกอบ (Factor Analysis) เพื่อตรวจสอบความเป็นมิติเดียวของแบบทดสอบก่อนที่จะนำไปวิเคราะห์ตาม ทฤษฎีการตอบสนองข้อสอบ โดยใช้โปรแกรม SPSS-X ซึ่งเลือกสกัดองค์ประกอบด้วยวิธี วิเคราะห์องค์ประกอบหลัก (Principal Component Analysis) แล้วหมุนแกนด้วยวิธี แวริแมกซ์ (Varimax) ได้ผลดังตาราง 6 ในภาคผนวก ก ซึ่งปรากฏว่า ค่าไอเกน (Eigen Value) ขององค์ประกอบที่ 1 ของแบบทดสอบทั้ง 6 ฉบับ มีค่าสูงกว่าค่าไอเกน ขององค์ประกอบที่ 2 เป็นอย่างมาก และค่าไอเกนขององค์ประกอบที่ 2 กับองค์ประกอบ ที่เหลือมีค่าไม่แตกต่างกันมากนัก ดังนั้นจึงกล่าวได้ว่าแบบทดสอบทั้ง 6 ฉบับ มีความเป็นมิติ เดียวพอประมาณ (Warm. 1978 : 104 ; Lord. 1980 : 19) นั่นคือ ทำให้สามารถ นำแบบทดสอบทั้ง 6 ฉบับ ไปวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบได้

1.10 นำผลจากข้อ 1.8 ไปวิเคราะห์หาค่าพหุคูณความยากของข้อสอบ ตามทฤษฎีการตอบสนองข้อสอบ แบบราสส์ โมเดล ด้วยวิธีการประมาณค่าความเป็นไปได้สูงสุด แบบไม่มีเงื่อนไข โดยใช้โปรแกรมไบคาล ได้ผลการวิเคราะห์ที่มีค่าความยากของข้อสอบ (b)

กระจายอยู่ในช่วงตั้งแต่ -2.81 ถึง 2.77 และค่าความสอดคล้องกับโมเดล (Fit Model : t-test) กระจายอยู่ในช่วงตั้งแต่ -3.29 ถึง 5.05 ซึ่งมีรายละเอียดดังตาราง 7, 8, 9, 10, 11 และ 12 ในภาคผนวก ข

1.11 คัดเลือกข้อสอบที่มีคุณภาพ โดยมีค่าความยากอยู่ในช่วง -3.00 ถึง 3.00 และมีค่าความสอดคล้องกับโมเดลอยู่ในช่วง -2.00 ถึง 2.00 ได้ผลดังตาราง 2

ตาราง 2 ผลการคัดเลือกข้อสอบที่มีคุณภาพ ของแบบทดสอบ 6 ฉบับ

| แบบทดสอบ<br>ฉบับที่ | จำนวนข้อ<br>(n) | ค่าความยาก<br>(b) | ค่าความสอดคล้อง<br>(t-test) |
|---------------------|-----------------|-------------------|-----------------------------|
| 1                   | 38              | -2.75 ถึง 2.62    | -1.99 ถึง 1.74              |
| 2                   | 35              | -2.12 ถึง 2.63    | -1.96 ถึง 1.56              |
| 3                   | 43              | -2.45 ถึง 2.08    | -1.87 ถึง 1.73              |
| 4                   | 39              | -2.81 ถึง 2.77    | -1.94 ถึง 1.75              |
| 5                   | 39              | -2.01 ถึง 2.21    | -1.88 ถึง 1.81              |
| 6                   | 43              | -1.40 ถึง 2.27    | -1.99 ถึง 1.80              |

จากตาราง 2 ผลปรากฏว่าได้ข้อสอบที่มีคุณภาพ จำนวน 237 ข้อ โดยมีค่าความยากของข้อสอบกระจายอยู่ในช่วงตั้งแต่ -2.81 ถึง 2.77 และมีค่าความสอดคล้องกับโมเดลกระจายอยู่ในช่วงตั้งแต่ -1.99 ถึง 1.81 แล้วจึงเรียงข้อสอบทั้งหมด จำนวน 237 ข้อ จากข้อที่มีค่าความยากต่ำสุดไปหาข้อที่มีค่าความยากสูงสุด

## 2. การสร้างแบบทดสอบตามโครงสร้าง

### 2.1 การสร้างแบบทดสอบสองขั้นตอน แบ่งออกเป็นสองขั้นดังนี้

ขั้นที่ 1 การสร้างแบบทดสอบแยกกลุ่ม ทำโดยสุ่มข้อสอบในข้อ 1.11 ที่มีค่าความยากใกล้เคียงกันในช่วงตั้งแต่ -1.00 ถึง 1.00 มาจำนวน 10 ข้อ ซึ่งทำได้

ข้อสอบที่มีค่าความยากในช่วงตั้งแต่  $-0.579$  ถึง  $0.693$  แล้วจัดเรียงข้อสอบจากข้อที่มีค่าความยากต่ำสุด ไปหาข้อที่มีค่าความยากสูงสุด

ขั้นที่ 2 การสร้างแบบทดสอบวัดผล ทำโดยสุ่มข้อสอบในข้อ 1.11 มาจำนวน 120 ข้อ แล้วจัดเป็นแบบทดสอบฉบับย่อย ๆ จำนวน 4 ฉบับ โดยยึดค่าความยากของข้อสอบเป็นเกณฑ์ในการจัดแบบทดสอบ ซึ่งจะทำได้แบบทดสอบ จำนวน 4 ฉบับ ๆ ละ 30 ข้อ แล้วจัดเรียงข้อสอบจากข้อที่มีค่าความยากต่ำสุด ไปหาข้อที่มีค่าความยากสูงสุด ดังนี้

ฉบับที่ 1 มีค่าความยากอยู่ในช่วงตั้งแต่  $-2.813$  ถึง  $-1.027$

ฉบับที่ 2 มีค่าความยากอยู่ในช่วงตั้งแต่  $-0.999$  ถึง  $-0.006$

ฉบับที่ 3 มีค่าความยากอยู่ในช่วงตั้งแต่  $0.058$  ถึง  $0.962$

ฉบับที่ 4 มีค่าความยากอยู่ในช่วงตั้งแต่  $1.011$  ถึง  $2.630$

## 2.2 การสร้างแบบทดสอบแฟล็กซีเลเวล ทำโดยสุ่มข้อสอบในข้อ 1.11

ที่เหลือจากการสร้างแบบทดสอบสองขั้นตอน มาจำนวน 59 ข้อ ปรากฏว่าได้ข้อสอบที่มีค่าความยากอยู่ในช่วงตั้งแต่  $-2.752$  ถึง  $2.768$  แล้วแบ่งข้อสอบออกเป็นสองกลุ่ม ๆ ละ 29 ข้อ โดยยึดค่าความยากเป็นเกณฑ์ในการแบ่งกลุ่มข้อสอบ จะได้ข้อสอบกลุ่มง่าย (กลุ่ม A) มีค่าความยากอยู่ในช่วงตั้งแต่  $-2.752$  ถึง  $-0.063$  และข้อสอบกลุ่มยาก (กลุ่ม B) มีค่าความยากอยู่ในช่วงตั้งแต่  $0.082$  ถึง  $2.768$  สำหรับข้อสอบที่เหลืออีก 1 ข้อนั้น จะเป็นข้อสอบที่มีค่าความยากตรงกึ่งกลางระหว่างข้อสอบกลุ่มง่ายกับข้อสอบกลุ่มยาก ซึ่งมีค่าความยากเท่ากับ  $0.017$  ข้อสอบข้อนี้จะเป็นข้อเริ่มต้นของผู้สอบทุกคน ข้อสอบกลุ่มง่ายจะเรียงลำดับข้อสอบจากข้อที่ยากที่สุด ไปหาข้อที่ง่ายที่สุด ส่วนข้อสอบกลุ่มยากจะเรียงลำดับข้อสอบจากข้อที่ง่ายที่สุด ไปหาข้อที่ยากที่สุด

สำหรับกระดาษคำตอบที่ใช้กับแบบทดสอบแฟล็กซีเลเวลจะต้องออกแบบมาเป็นพิเศษ เพื่อที่จะทำให้ผู้สอบทราบผลการตอบข้อสอบในแต่ละข้อโดยทันทีว่าถูกหรือผิด ซึ่งจะเป็นตัวกำหนดแนวทางในการตอบข้อสอบด้วย ดังนั้นผู้วิจัยจึงได้สร้างกระดาษคำตอบขึ้นมาใหม่ โดยใช้กระดาษคำตอบซ้อนกัน 2 แผ่น กระดาษแผ่นบนจะพิมพ์หมายเลขข้อและตัวเลือกของแต่ละข้อเช่นเดียวกับกระดาษคำตอบทั่ว ๆ ไป แต่จะปรุรอบตัวเลือก เพื่อให้ผู้สอบสามารถแกะกระดาษคำตอบตามรอยปรุรอบตัวเลือกได้สะดวก ส่วนกระดาษแผ่นล่างจะพิมพ์หมายเลขข้อและตัวเลือกให้ตรงกับกระดาษแผ่นบน แต่ในตัวเลือกของกระดาษคำตอบ

จะพิมพ์ตัวอักษรที่แสดงกลุ่มของข้อสอบ (กลุ่ม A หรือกลุ่ม B) ซึ่งจะเป็นตัวกำหนดให้ผู้สอบทำข้อสอบในกลุ่มใดกลุ่มหนึ่ง กล่าวคือ ถ้าผู้สอบตอบถูกเมื่อแกะรอยปรุรอบตัวเลือกจะพบตัวอักษร B แสดงว่าจะต้อง ไปตอบข้อสอบข้อที่มีเลขน้อยที่สุด ซึ่งยัง ไม่ได้ตอบในกลุ่มยาก (กลุ่ม B) และถ้าผู้สอบตอบผิดเมื่อแกะรอยปรุรอบตัวเลือกจะพบตัวอักษร A แสดงว่าจะต้อง ไปตอบข้อสอบข้อที่มีเลขน้อยที่สุด ซึ่งยังไม่ได้ตอบในกลุ่มง่าย (กลุ่ม A)

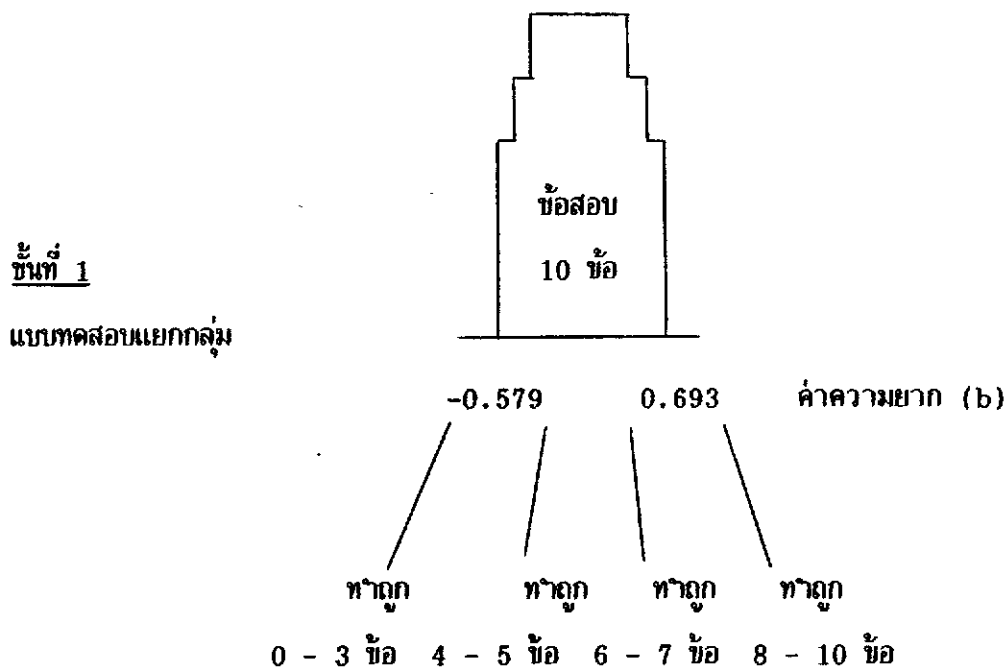
2.3 การสร้างแบบทดสอบดั้งเดิม ทำโดยสุ่มข้อสอบในข้อ 1.11 ที่เหลือจากการสร้างแบบทดสอบสองขั้นตอนและแบบทดสอบเพื่อกี้อีกสิบเลเวล มาจำนวน 30 ข้อ แล้วนำข้อสอบที่สุ่มมานี้มาจัดเรียงข้อสอบตามเนื้อหา และ ในแต่ละเนื้อหาจะเรียงข้อสอบที่มีค่าง่ายที่สุดไปยังข้อสอบที่มีค่ายากที่สุด

3. นำแบบทดสอบทั้ง 3 รูปแบบ มาจัดทำคู่มือคำแนะนำการสอบ และจัดพิมพ์เป็นรูปเล่ม

โครงสร้างของแบบทดสอบสองขั้นตอน แบบทดสอบเพื่อกี้อีกสิบเลเวล และกระดาษคำตอบของแบบทดสอบเพื่อกี้อีกสิบเลเวล แสดงได้ดังภาพประกอบ 10, 11, 12 และ 13

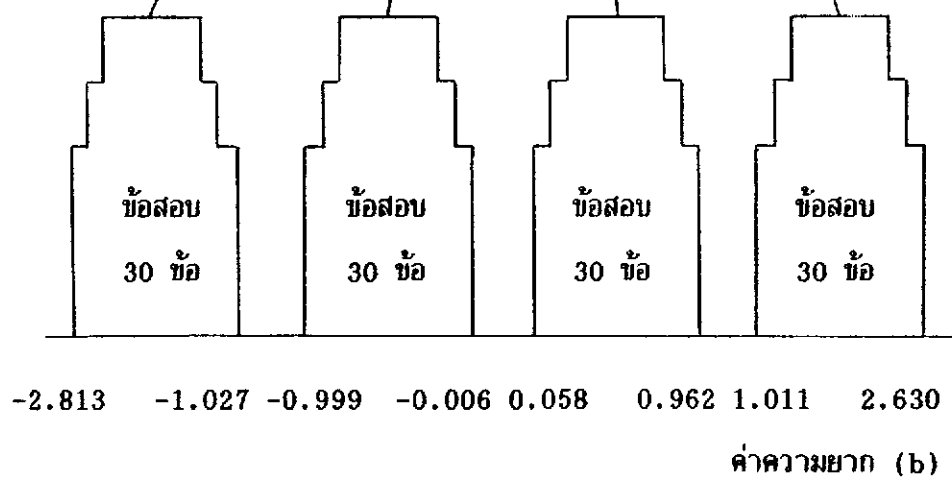
### ขั้นที่ 1

แบบทดสอบแยกกลุ่ม



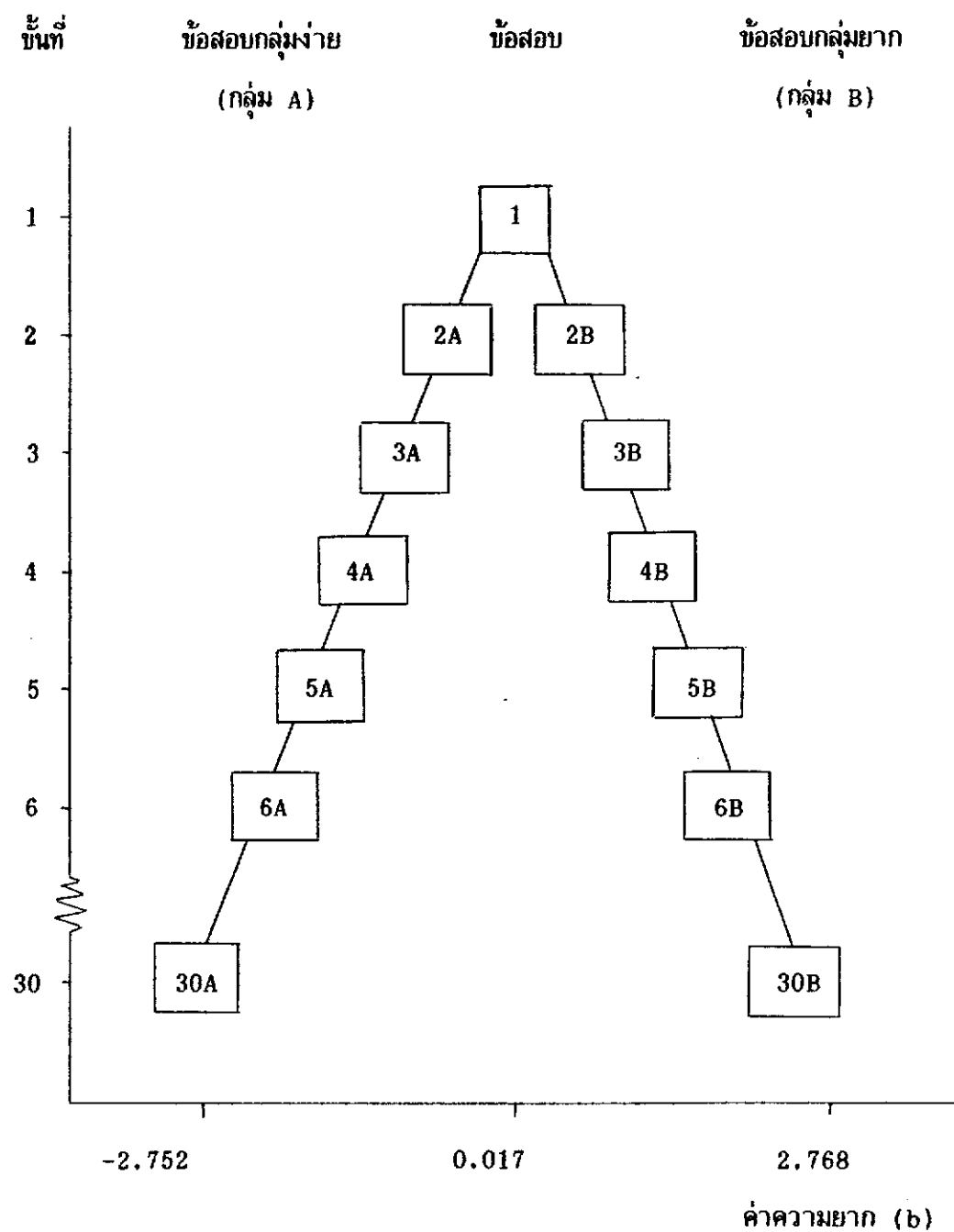
### ขั้นที่ 2

แบบทดสอบวัดผล



ภาพประกอบ 10 โครงสร้างของแบบทดสอบสองขั้นตอน

จากภาพประกอบ 10 ผู้สอบทุกคนจะต้องทำข้อสอบในขั้นที่ 1 ซึ่งมีอยู่จำนวน 10 ข้อ เมื่อผู้สอบทำข้อสอบถูก 0 - 3 ข้อ จะต้องไปทำข้อสอบในขั้นที่ 2 ที่ง่ายที่สุด ส่วนผู้สอบที่ทำข้อสอบในขั้นที่ 1 ถูก 4-5, 6-7 หรือ 8-10 ข้อ จะแยกไปทำข้อสอบในขั้นที่ 2 ที่มีค่าความยากเพิ่มมากขึ้นตามลำดับ



ภาพประกอบ 11 โครงสร้างของแบบทดสอบเฟลิกซ์ชีเลเวล

จากภาพประกอบ 11 ผู้สอบทุกคนต้องทำข้อสอบข้อแรก เมื่อผู้สอบตอบผิดจะต้องไปตอบข้อสอบที่มีเลขน้อยที่สุด ซึ่งยังไม่ได้ตอบในกลุ่มง่าย (กลุ่ม A) และเมื่อผู้สอบตอบถูกจะต้องไปตอบข้อสอบที่มีเลขน้อยที่สุด ซึ่งยังไม่ได้ตอบในกลุ่มยาก (กลุ่ม B) และจะยุติการสอบเมื่อผู้สอบทำข้อสอบไปทั้งหมด จำนวน 30 ข้อ

โรงเรียน.....  
 ชื่อ.....เลขที่.....ชั้น.....  
 วิชา.....สอบวันที่.....เดือน.....พ.ศ.....

| ข้อ     |   |   |   |   |    | ก   |   |   |   |   |    | ข       |   |   |   |   |    | ค   |   |   |   |   |   | ง |  |  |  |  |  | จ |  |  |  |  |  |
|---------|---|---|---|---|----|-----|---|---|---|---|----|---------|---|---|---|---|----|-----|---|---|---|---|---|---|--|--|--|--|--|---|--|--|--|--|--|
| 1       |   |   |   |   |    |     |   |   |   |   |    |         |   |   |   |   |    |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| กลุ่ม A |   |   |   |   |    |     |   |   |   |   |    | กลุ่ม B |   |   |   |   |    |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| ข้อ     | ก | ข | ค | ง | จ  | ข้อ | ก | ข | ค | ง | จ  | ข้อ     | ก | ข | ค | ง | จ  | ข้อ | ก | ข | ค | ง | จ |   |  |  |  |  |  |   |  |  |  |  |  |
| 2       |   |   |   |   | 17 |     |   |   |   |   | 2  |         |   |   |   |   | 17 |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| 3       |   |   |   |   | 18 |     |   |   |   |   | 3  |         |   |   |   |   | 18 |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| 4       |   |   |   |   | 19 |     |   |   |   |   | 4  |         |   |   |   |   | 19 |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| 5       |   |   |   |   | 20 |     |   |   |   |   | 5  |         |   |   |   |   | 20 |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| 6       |   |   |   |   | 21 |     |   |   |   |   | 6  |         |   |   |   |   | 21 |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| 7       |   |   |   |   | 22 |     |   |   |   |   | 7  |         |   |   |   |   | 22 |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| 8       |   |   |   |   | 23 |     |   |   |   |   | 8  |         |   |   |   |   | 23 |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| 9       |   |   |   |   | 24 |     |   |   |   |   | 9  |         |   |   |   |   | 24 |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| 10      |   |   |   |   | 25 |     |   |   |   |   | 10 |         |   |   |   |   | 25 |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| 11      |   |   |   |   | 26 |     |   |   |   |   | 11 |         |   |   |   |   | 26 |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| 12      |   |   |   |   | 27 |     |   |   |   |   | 12 |         |   |   |   |   | 27 |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| 13      |   |   |   |   | 28 |     |   |   |   |   | 13 |         |   |   |   |   | 28 |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| 14      |   |   |   |   | 29 |     |   |   |   |   | 14 |         |   |   |   |   | 29 |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| 15      |   |   |   |   | 30 |     |   |   |   |   | 15 |         |   |   |   |   | 30 |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |
| 16      |   |   |   |   |    |     |   |   |   |   | 16 |         |   |   |   |   |    |     |   |   |   |   |   |   |  |  |  |  |  |   |  |  |  |  |  |

ภาพประกอบ 12 กระดาษคำตอบแผ่นบนของแบบทดสอบเพื่อกิจิเลเวล

| ข้อ ก ข ค ง จ |   |   |   |   |   |
|---------------|---|---|---|---|---|
| 1             | A | B | A | A | A |

| กลุ่ม A |   |   |   |   |   | กลุ่ม B |   |   |   |   |   |     |   |   |   |   |   |    |   |   |   |   |   |
|---------|---|---|---|---|---|---------|---|---|---|---|---|-----|---|---|---|---|---|----|---|---|---|---|---|
| ข้อ     | ก | ข | ค | ง | จ | ข้อ     | ก | ข | ค | ง | จ | ข้อ | ก | ข | ค | ง | จ |    |   |   |   |   |   |
| 2       | A | A | A | B | A | 17      | A | B | A | A | A | 2   | A | A | A | B | A | 17 | A | A | A | B | A |
| 3       | B | A | A | A | A | 18      | A | A | B | A | A | 3   | B | A | A | A | A | 18 | B | A | A | A | A |
| 4       | A | A | B | A | A | 19      | B | A | A | A | A | 4   | A | B | A | A | A | 19 | A | B | A | A | A |
| 5       | B | A | A | A | A | 20      | A | A | A | B | A | 5   | A | A | B | A | A | 20 | A | A | A | A | B |
| 6       | A | A | A | B | A | 21      | A | B | A | A | A | 6   | A | A | A | A | B | 21 | A | A | B | A | A |
| 7       | A | B | A | A | A | 22      | A | A | A | A | B | 7   | A | A | B | A | A | 22 | A | B | A | A | A |
| 8       | A | A | B | A | A | 23      | A | A | B | A | A | 8   | B | A | A | A | A | 23 | A | A | A | B | A |
| 9       | A | A | A | A | B | 24      | B | A | A | A | A | 9   | B | A | A | A | A | 24 | A | A | A | B | A |
| 10      | B | A | A | A | A | 25      | A | A | A | B | A | 10  | A | A | A | B | A | 25 | A | B | A | A | A |
| 11      | A | A | A | B | A | 26      | A | B | A | A | A | 11  | A | A | B | A | A | 26 | A | A | B | A | A |
| 12      | A | B | A | A | A | 27      | A | A | B | A | A | 12  | A | A | A | A | B | 27 | A | A | A | A | B |
| 13      | A | A | A | B | A | 28      | A | A | A | A | B | 13  | B | A | A | A | A | 28 | A | A | A | B | A |
| 14      | B | A | A | A | A | 29      | B | A | A | A | A | 14  | A | B | A | A | A | 29 | A | B | A | A | A |
| 15      | A | A | A | A | B | 30      | A | B | A | A | A | 15  | A | A | B | A | A | 30 | A | A | A | B | A |
| 16      | A | A | B | A | A |         |   |   |   |   |   | 16  | A | B | A | A | A |    |   |   |   |   |   |

ภาพประกอบ 13 กระดาษคำตอบด้านล่างของแบบทดสอบเพ็ญิกิเลเวล

### วิธีดำเนินการเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลในครั้งนี้ ผู้วิจัยดำเนินการตามขั้นตอน ดังนี้

1. วางแผนในการดำเนินการสอบ โดยการติดต่อขอความร่วมมือในการนำแบบทดสอบไปทดสอบกับนักเรียนที่ใช้เป็นกลุ่มตัวอย่าง เพื่อกำหนดวัน เวลา ในการดำเนินการสอบ
2. นำแบบทดสอบทั้ง 3 รูปแบบ คือ แบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กชีลเวล ไปทดสอบกับกลุ่มตัวอย่าง จำนวน 441 คน โดยนักเรียนแต่ละคนจะต้องทำแบบทดสอบทั้ง 3 รูปแบบ สำหรับแบบทดสอบทั้ง 3 รูปแบบนั้น ผู้วิจัยจะสุ่มแต่ละรูปแบบไปทดสอบกับกลุ่มตัวอย่าง จำนวน 3 ครั้ง โดยทั้งช่วงการทดสอบแต่ละครั้งเป็นเวลา 1 สัปดาห์
3. นำผลการสอบในข้อ 2 มาจัดเตรียมข้อมูลสำหรับการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ แล้วนำผลการประมาณค่าไปวิเคราะห์ทางสถิติ และทดสอบสมมติฐานของการวิจัยต่อไป

### การวิเคราะห์ข้อมูลและสถิติที่ใช้

1. หาค่าสถิติพื้นฐาน คือ ค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐาน
2. การตรวจสอบความเป็นมิติเดียว (Unidimensionality) ของแบบทดสอบตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) โดยนำแบบทดสอบจำนวน 6 ฉบับ ๆ ละ 45 ข้อ ไปทดสอบกับนักเรียนที่ใช้ในการทดลองสอบ จำนวน 2,403 คน แล้วนำผลมาวิเคราะห์องค์ประกอบ (Factor Analysis) โดยใช้โปรแกรม SPSS-X ซึ่งเลือกสกัดองค์ประกอบด้วยวิธีวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis : PC) แล้วหมุนแกนด้วยวิธีแวนริแมกซ์ (Varimax)
3. ประมาณค่าพารามิเตอร์ความยากของข้อสอบและค่าความสามารถของผู้สอบ โดยนำผลการทดลองสอบจากแบบทดสอบ จำนวน 6 ฉบับ และผลการสอบในการวิจัยจากแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กชีลเวล มาวิเคราะห์หาค่าพารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบ แบบราสช์ ไมเคิล (Rasch Model) ด้วยวิธีการประมาณค่าความเป็นไปได้สูงสุดแบบไม่มีเงื่อนไข (Unconditional Maximum Likelihood Procedure) โดยใช้โปรแกรมไบคาล (BICAL)

4. คำนวณค่าความเที่ยงตรงเชิงสภาพ (Concurrent Validity) ของการประมาณค่าความสามารถจากแบบทดสอบที่แตกต่างกัน 3 รูปแบบ โดยทำตามลำดับขั้น ดังนี้

4.1 คำนวณค่าความสัมพันธ์ระหว่างความสามารถที่ได้จากการประมาณค่าด้วยการทดสอบแบบทดสอบแต่ละรูปแบบ กับเกรดเฉลี่ยสะสมในวิชาคณิตศาสตร์ที่เรียนผ่านมาในชั้นมัธยมศึกษาปีที่ 1 และปีที่ 2 ซึ่งจะคำนวณแยกตามโรงเรียนของกลุ่มตัวอย่าง โดยใช้สูตรสัมประสิทธิ์สหสัมพันธ์เพียร์สัน (Pearson Product Moment Correlation Coefficient) (ล้วน สายยศ และอังคณา สายยศ. 2522 : 168)

$$r_{\theta y} = \frac{N\sum\theta Y - \sum\theta\sum Y}{\sqrt{[N\sum\theta^2 - (\sum\theta)^2][N\sum Y^2 - (\sum Y)^2]}}$$

เมื่อ  $r_{\theta y}$  แทน ค่าความเที่ยงตรงเชิงสภาพของการประมาณค่าความสามารถ  
 $\theta$  แทน ค่าความสามารถที่ได้จากการประมาณค่า  
 $Y$  แทน เกรดเฉลี่ยสะสม  
 $N$  แทน จำนวนนักเรียนในกลุ่มตัวอย่าง

4.2 เปลี่ยนค่าความเที่ยงตรงเชิงสภาพของกลุ่มตัวอย่างแต่ละโรงเรียนเป็นคะแนนมาตรฐาน (Z) ตารางของฟิชเชอร์ (Fisher's Transformation)

4.3 คำนวณ Fisher's Z เฉลี่ย โดยใช้สูตรดังนี้ (ล้วน สายยศ และอังคณา สายยศ. 2522 : 228)

$$\bar{Z} = \sum(N_i - 3)Z_i / \sum(N_i - 3)$$

เมื่อ  $\bar{Z}$  แทน คะแนนมาตรฐาน (Fisher's Z) เฉลี่ย  
 $Z_i$  แทน คะแนนมาตรฐาน (Fisher's Z) ของกลุ่มที่ i  
 $N_i$  แทน จำนวนนักเรียนของกลุ่มที่ i

4.4 เปลี่ยนค่าคะแนนมาตรฐานเฉลี่ย ( $\bar{Z}$ ) เป็นค่าความเที่ยงตรงเชิงสภาพเฉลี่ย ตามตารางของพิชเชอร์

5. ทดสอบความแตกต่างของค่าความเที่ยงตรงเชิงสภาพ โดยนำค่าความเที่ยงตรงมาแปลงเป็นคะแนนมาตรฐาน ( $Z$ ) ตามตารางของพิชเชอร์ แล้วนำมาทดสอบ โดยใช้สูตร Chi Square :  $\chi^2$  ดังนี้ (Steel and Torrie. 1981 : 281)

$$\chi^2 = \frac{\sum (N - 3) Z^2 - [\sum (N - 3) Z]^2}{\sum (N - 3)}$$

- เมื่อ  $\chi^2$  แทน ค่าจาก  $\chi^2$ -distribution โดยที่  $df = K - 1$   
 $Z$  แทน คะแนนมาตรฐานที่แปลงมาจากค่าความเที่ยงตรงเชิงสภาพตามตารางของพิชเชอร์  
 $N$  แทน จำนวนนักเรียนในกลุ่มตัวอย่าง  
 $K$  แทน จำนวนค่าของความเที่ยงตรงเชิงสภาพ

6. ถ้าพบค่าความเที่ยงตรงเชิงสภาพที่นำมาทดสอบในข้อ 5 มีความแตกต่างกันอย่างมีนัยสำคัญ ให้นำค่าความเที่ยงตรงเชิงสภาพมาทดสอบความแตกต่างเป็นรายคู่ โดยใช้สูตรดังนี้ (ล้วน สายยศ และอังคณา สายยศ. 2522 : 227)

$$Z = \frac{Z_{r1} - Z_{r2}}{\sqrt{1/(N_1 - 3) + 1/(N_2 - 3)}}$$

- เมื่อ  $Z$  แทน คะแนนมาตรฐานของไค้งปกติ  
 $Z_{r1}, Z_{r2}$  แทน คะแนนมาตรฐานที่แปลงมาจากค่าสัมประสิทธิ์สหสัมพันธ์ตามตารางของพิชเชอร์  
 $N_1, N_2$  แทน จำนวนนักเรียนในกลุ่มตัวอย่างที่ทาแบบทดสอบแต่ละรูปแบบ

7. การคำนวณค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบเฟลิกซ์เลเวล โดยนำค่าพารามิเตอร์ของข้อสอบแต่ละฉบับมาคำนวณค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ ณ ระดับความสามารถต่าง ๆ 81 ระดับ โดยใช้สูตรในการคำนวณดังนี้ (Lord and Novick. 1968 : 454)

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

เมื่อ  $I(\theta)$  แทน ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ ณ ระดับความสามารถต่าง ๆ ซึ่งผู้วิจัยแบ่งออกเป็น 81 ระดับ คือ -4.00, -3.90, -3.80, ..., 3.80, 3.90, 4.00

$i$  แทน ข้อสอบข้อที่ 1, 2, 3, ..., 30

$P_i(\theta)$  แทน ความน่าจะเป็นที่ผู้สอบที่มีระดับความสามารถ  $\theta$  จะตอบข้อสอบข้อที่  $i$  ได้ถูกต้อง

$Q_i(\theta)$  แทน  $1 - P_i(\theta)$

$P'_i(\theta)$  แทน ความชันโค้งลักษณะข้อสอบที่ระดับความสามารถ  $\theta$

8. การคำนวณค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ ระหว่างแบบทดสอบที่แตกต่างกัน 3 รูปแบบ โดยนำค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ ณ ระดับความสามารถต่าง ๆ 81 ระดับ มาเปรียบเทียบกับกันทีละคู่ โดยใช้สูตรในการคำนวณ ดังนี้ (Lord and Novick. 1968 : 472)

$$I(\theta, X_1, X_2) = I(\theta, X_1) / I(\theta, X_2)$$

เมื่อ  $RE(\theta, X_1, X_2)$  แทน ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ  
ระหว่างแบบทดสอบรูปแบบที่ 1 ( $X_1$ ) เทียบกับแบบทดสอบ  
รูปแบบที่ 2 ( $X_2$ )

$I(\theta, X_1)$  แทน ค่าอินฟอร์เมชันของแบบทดสอบรูปแบบที่ 1 ( $X_1$ ) ณ ระดับ  
ความสามารถ  $\theta$

$I(\theta, X_2)$  แทน ค่าอินฟอร์เมชันของแบบทดสอบรูปแบบที่ 2 ( $X_2$ ) ณ ระดับ  
ความสามารถ  $\theta$

๑. ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ ระหว่างแบบทดสอบ  
ที่นำมาเปรียบเทียบกับในข้อ 8 นำมาแปลผลดังนี้ (Warm. 1978 : 76)

เมื่อ  $RE(\theta, X_1, X_2) > 1$  แสดงว่า แบบทดสอบรูปแบบที่ 1 ( $X_1$ )  
มีประสิทธิภาพสูงกว่าแบบทดสอบรูปแบบ  
ที่ 2 ( $X_2$ ) ณ ระดับความสามารถ  $\theta$

$RE(\theta, X_1, X_2) < 1$  แสดงว่า แบบทดสอบรูปแบบที่ 1 ( $X_1$ )  
มีประสิทธิภาพต่ำกว่าแบบทดสอบรูปแบบ  
ที่ 2 ( $X_2$ ) ณ ระดับความสามารถ  $\theta$

$RE(\theta, X_1, X_2) = 1$  แสดงว่า แบบทดสอบรูปแบบที่ 1 ( $X_1$ )  
มีประสิทธิภาพพอ ๆ กับแบบทดสอบรูปแบบ  
ที่ 2 ( $X_2$ ) ณ ระดับความสามารถ  $\theta$

#### บทที่ 4

#### ผลการวิเคราะห์ข้อมูล

ในการศึกษาครั้งนี้เพื่อให้เกิดความเข้าใจตรงกันในการนำเสนอและแปลผลการวิเคราะห์ข้อมูล ผู้วิจัยจึง ได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล ดังนี้

##### สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

|          |     |                                                                        |
|----------|-----|------------------------------------------------------------------------|
| X        | แทน | ค่าเฉลี่ย                                                              |
| S        | แทน | ค่าความเบี่ยงเบนมาตรฐาน                                                |
| N        | แทน | จำนวนนักเรียนในกลุ่มตัวอย่าง                                           |
| n        | แทน | จำนวนข้อสอบในแบบทดสอบ                                                  |
| CONT     | แทน | แบบทดสอบดั้งเดิม                                                       |
| TWOT     | แทน | แบบทดสอบสองขั้นตอน                                                     |
| FLEXT    | แทน | แบบทดสอบเฟลิกซ์เลเวล                                                   |
| b        | แทน | ค่าความยากของข้อสอบ                                                    |
| $\theta$ | แทน | ค่าความสามารถของผู้สอบ                                                 |
| Y        | แทน | เกรดเฉลี่ยสะสม                                                         |
| rey      | แทน | ค่าสัมประสิทธิ์สหสัมพันธ์ (ค่าความเที่ยงตรงของการประมาณค่าความสามารถ)  |
| z        | แทน | คะแนนมาตรฐาน ซึ่งแปลงมาจากค่าสัมประสิทธิ์สหสัมพันธ์ตามตารางของพิชเชอร์ |
| k        | แทน | จำนวนค่าของสัมประสิทธิ์สหสัมพันธ์                                      |
| df       | แทน | ชั้นของความเป็นอิสระ                                                   |
| $\chi^2$ | แทน | ค่าไคสแควร์                                                            |
| TIF      | แทน | ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ                                     |
| RE       | แทน | ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ                        |
| SEE      | แทน | ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า                               |

### การเสนอผลการวิเคราะห์ข้อมูล

ในการศึกษาครั้งนี้ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับ ดังนี้

1. ผลการหาค่าความเที่ยงตรงและการทดสอบความแตกต่างของค่าความเที่ยงตรงของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กซีเลเวล
2. ผลการหาค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กซีเลเวล

### ผลการวิเคราะห์ข้อมูล

1. ผลการหาค่าความเที่ยงตรงและการทดสอบความแตกต่างของค่าความเที่ยงตรงของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กซีเลเวล

ผู้วิจัยได้นำค่าประมาณความสามารถของผู้สอบ ที่ได้จากการทดสอบด้วยแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กซีเลเวล จากจำนวนนักเรียน 441 คน ไปคำนวณหาค่าความเที่ยงตรง โดยการคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน ระหว่างค่าความสามารถของผู้สอบจากการทดสอบด้วยแบบทดสอบแต่ละรูปแบบ กับเกรดเฉลี่ยสะสมในวิชาคณิตศาสตร์ แล้วทดสอบความแตกต่างของค่าความเที่ยงตรงระหว่างแบบทดสอบทั้ง 3 รูปแบบ ด้วยวิธีการแปลงค่าสัมประสิทธิ์สหสัมพันธ์ ให้อยู่ในรูปคะแนนมาตรฐานตามตารางของฟิชเชอร์ โดยใช้ค่าไคสแควร์ ได้ผลดังตาราง 3

ตาราง 3 ค่าความเที่ยงตรงและผลการทดสอบความแตกต่างของค่าความเที่ยงตรงของ  
การประมาณค่าความสามารถของผู้สอบ จากการทบทวนแบบทดสอบที่แตกต่างกัน 3 รูปแบบ

| แบบทดสอบ | rey    | Z      | df | $\chi^2$ |
|----------|--------|--------|----|----------|
| CONT     | 0.5694 | 0.6436 | 2  | 7.9006*  |
| TWOT     | 0.6825 | 0.8321 | 2  |          |
| FLEXT    | 0.6175 | 0.7177 | 2  |          |

\*p < .05

จากตาราง 3 แสดงให้เห็นว่าค่าความเที่ยงตรงของการประมาณค่าความสามารถ  
ของผู้สอบ จากการทดสอบด้วยแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบ  
แฟล็กซีเลเวลมีค่าเท่ากับ 0.5694, 0.6825 และ 0.6175 ตามลำดับ ซึ่งค่าความเที่ยงตรง  
ของการประมาณค่าความสามารถจากแบบทดสอบสองขั้นตอนมีค่าสูงสุด ต่ำลงมาเป็นแบบทดสอบ  
แฟล็กซีเลเวล และแบบทดสอบดั้งเดิมตามลำดับ และเมื่อนำมาทดสอบความแตกต่างปรากฏว่าค่า  
ความเที่ยงตรงของการประมาณค่าความสามารถของผู้สอบ ระหว่างแบบทดสอบทั้ง 3 รูปแบบ  
มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่ามีค่าความเที่ยงตรงแตกต่างกัน  
อย่างน้อย 1 คู่ ดังนั้น เพื่อให้ทราบค่าความเที่ยงตรงของการประมาณค่าความสามารถของ  
ผู้สอบจากแบบทดสอบคู่ใดมีความแตกต่างกัน ผู้วิจัยจึงนำค่าความเที่ยงตรงจากแบบทดสอบดั้งเดิม  
แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กซีเลเวล ไปทดสอบความแตกต่างทีละคู่ได้ผลดัง  
ตาราง 4

ตาราง 4 ผลของการทดสอบความแตกต่างของค่าความเที่ยงตรงของการประมาณค่าความสามารถของผู้สอบ จากการทําแบบทดสอบที่แตกต่างกัน 3 รูปแบบ โดยทดสอบเป็นรายคู่

| แบบทดสอบ | Z      | CONT   | TWOT     | FLEXT  |
|----------|--------|--------|----------|--------|
|          |        | 0.6436 | 0.8321   | 0.7177 |
| CONT     | 0.6436 | -      | 2.7885** | 1.0962 |
| TWOT     | 0.8321 |        | -        | 1.6923 |
| FLEXT    | 0.7177 |        |          | -      |

\*\*p < .01

จากตาราง 4 แสดงให้เห็นว่าค่าความเที่ยงตรงของการประมาณค่าความสามารถของผู้สอบ ระหว่างแบบทดสอบดั้งเดิมกับแบบทดสอบสองขั้นตอน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนค่าความเที่ยงตรงของการประมาณค่าความสามารถของผู้สอบระหว่างแบบทดสอบดั้งเดิมกับแบบทดสอบแฟล็กซีเลเวล และแบบทดสอบสองขั้นตอนกับแบบทดสอบแฟล็กซีเลเวล มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. ผลการหาค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กซีเลเวล

ผู้วิจัย ได้คำนวณหาค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ ระหว่างแบบทดสอบดั้งเดิม เทียบกับแบบทดสอบสองขั้นตอน แบบทดสอบดั้งเดิม เทียบกับแบบทดสอบแฟล็กซีเลเวล และแบบทดสอบสองขั้นตอน เทียบกับแบบทดสอบแฟล็กซีเลเวล โดยการนำค่าฟังก์ชันเอ็นฟอร์เมชันของแบบทดสอบแต่ละรูปแบบ ซึ่งมีรายละเอียดดังตาราง 14 และภาพประกอบ 17, 18, 19, 20, 21, 22 และ 23 ในภาคผนวก ง มาคำนวณหาค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ ปรากฏว่าได้ผลดังตาราง 5 (มีรายละเอียดดังตาราง 15 ในภาคผนวก จ)

ตาราง 5 ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ ระหว่าง  
แบบทดสอบที่แตกต่างกัน 3 รูปแบบ ที่ระดับความสามารถตั้งแต่ -4.00 ถึง 4.00 จำนวน  
81 ระดับความสามารถ

| ระดับความสามารถ          | RE              |                 |                 |
|--------------------------|-----------------|-----------------|-----------------|
|                          | CONT/TWOT       | CONT/FLEXT      | TWOT/FLEXT      |
| $\theta < -3.00$         | 0.4568 - 0.6235 | 0.7256 - 0.7291 | 1.1694 - 1.5884 |
| $-3.00 < \theta < -1.00$ | 0.6435 - 0.9681 | 0.7300 - 0.8919 | 0.8690 - 1.1316 |
| $-1.00 < \theta < 1.00$  | 0.9865 - 1.1713 | 0.9218 - 1.3360 | 0.9343 - 1.1463 |
| $1.00 < \theta < 3.00$   | 0.7299 - 1.0186 | 0.6763 - 0.9773 | 0.7583 - 0.9640 |
| $\theta > 3.00$          | 0.3419 - 0.6924 | 0.7086 - 0.7380 | 1.0234 - 2.1583 |

จากตาราง 5 แสดงให้เห็นช่วงการกระจายของค่าประสิทธิภาพสัมพัทธ์ของการ  
ประมาณค่าความสามารถ ระหว่างแบบทดสอบที่แตกต่างกัน 3 รูปแบบ ซึ่งนำมาเปรียบเทียบกับ  
เกณฑ์ 1 (Warm, 1978 : 76) ได้ผลดังนี้

1. ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์  
ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบสองขั้นตอน มีค่าต่ำกว่า 1 ที่ความสามารถระดับ  
ต่ำมาก ( $\theta < -3.00$ ) ระดับต่ำ ( $-3.00 < \theta < -1.00$ ) และระดับสูงมาก ( $\theta > 3.00$ )  
ส่วนที่ความสามารถระดับปานกลาง ( $-1.00 < \theta < 1.00$ ) มีค่าประสิทธิภาพสัมพัทธ์ต่ำกว่า 1  
และสูงกว่า 1 โดยมีค่าต่ำกว่า 1 ที่ระดับความสามารถเท่ากับ -1.10 ( $\theta = -1.10$ ) เพียง  
ระดับเดียวเท่านั้น นอกนั้นที่ความสามารถตั้งแต่ -0.90 ถึง 1.00 ( $-0.90 < \theta < 1.00$ )  
จะมีค่าสูงกว่า 1 แสดงว่าที่ความสามารถระดับปานกลางนี้ส่วนมากมีค่าสูงกว่า 1 สำหรับที่ความ  
สามารถระดับสูง ( $1.00 < \theta < 3.00$ ) มีค่าประสิทธิภาพสัมพัทธ์ต่ำกว่า 1 และสูงกว่า 1 โดย  
มีค่าสูงกว่า 1 ที่ความสามารถเท่ากับ 1.10 ( $\theta = 1.10$ ) เพียงระดับเดียวเท่านั้น นอกนั้น

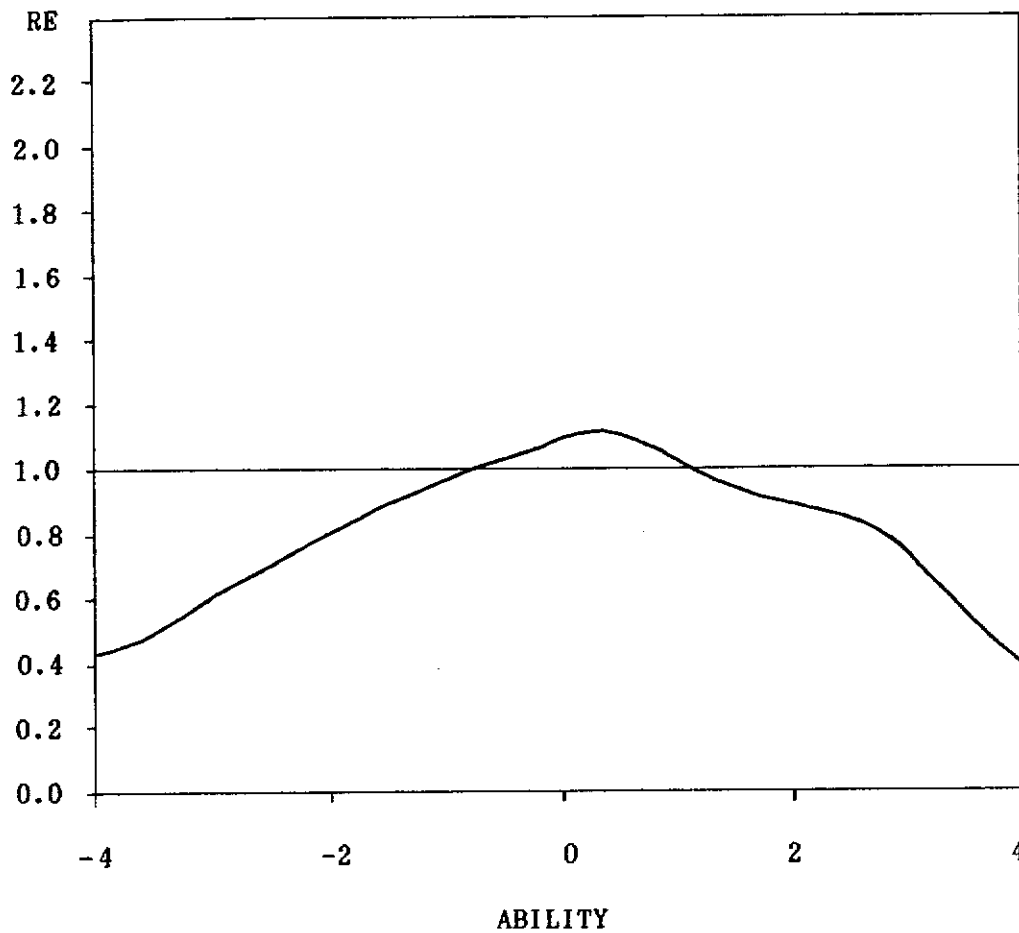
ที่ความสามารถตั้งแต่ 1.20 ถึง 3.00 ( $1.20 < \theta < 3.00$ ) จะมีค่าต่ำกว่า 1 แสดงว่า  
 ที่ความสามารถระดับสูงนี้ส่วนมากมีค่าประสิทธิภาพสัมพัทธ์ต่ำกว่า 1 ดังนั้นจึงสรุปรวมได้ว่า  
 ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบ  
 ดั้งเดิมเทียบกับแบบทดสอบสองขั้นตอน มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับ  
 ปานกลาง ( $-1.00 < \theta < 1.00$ ) แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำ  
 ลงมา ( $\theta < -1.00$ ) และระดับสูงขึ้นไป ( $\theta > 1.00$ ) นั่นคือ แบบทดสอบดั้งเดิมมีประสิทธิภาพ  
 ในการประมาณค่าความสามารถสูงกว่าแบบทดสอบสองขั้นตอน ในกลุ่มผู้สอบที่มีความสามารถ  
 ระดับปานกลาง แต่จะมีประสิทธิภาพในการประมาณค่าความสามารถต่ำกว่าแบบทดสอบสอง  
 ขั้นตอน ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมาและระดับสูงขึ้นไป

2. ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์  
 ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบเพลิกซีเลเวล มีค่าต่ำกว่า 1 ที่ความสามารถระดับ  
 ต่ำมาก ( $\theta < -3.00$ ) ระดับต่ำ ( $-3.00 < \theta < -1.00$ ) ระดับสูง ( $1.00 < \theta < 3.00$ )  
 และระดับสูงมาก ( $\theta > 3.00$ ) ส่วนที่ความสามารถระดับปานกลาง ( $-1.00 < \theta < 1.00$ )  
 มีค่าประสิทธิภาพสัมพัทธ์ต่ำกว่า 1 และสูงกว่า 1 โดยมีค่าต่ำกว่า 1 ที่ระดับความสามารถ  
 เท่ากับ -1.00, -0.90 และ -0.80 ( $-1.00 < \theta < -0.80$ ) เพียง 3 ระดับความสามารถ  
 เท่านั้น นอกนั้นที่ความสามารถตั้งแต่ -0.70 ถึง 1.00 ( $-0.70 < \theta < 1.00$ ) จะมีค่าสูง  
 กว่า 1 แสดงว่าที่ความสามารถระดับปานกลางนี้ส่วนมากมีค่าประสิทธิภาพสัมพัทธ์สูงกว่า 1  
 ดังนั้นจึงสรุปรวมได้ว่า ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์  
 ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบเพลิกซีเลเวล มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มี  
 ความสามารถระดับปานกลาง ( $-1.00 < \theta < 1.00$ ) แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มี  
 ความสามารถระดับต่ำลงมา ( $\theta < -1.00$ ) และระดับสูงขึ้นไป ( $\theta > 1.00$ ) นั่นคือ  
 แบบทดสอบดั้งเดิมมีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่าแบบทดสอบเพลิกซีเลเวล  
 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง แต่จะมีประสิทธิภาพในการประมาณค่าความ  
 สามารถต่ำกว่าแบบทดสอบเพลิกซีเลเวล ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมาและระดับ  
 สูงขึ้นไป

3. ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์  
 ระหว่างแบบทดสอบสองขั้นตอนเทียบกับแบบทดสอบเพลิกซีเลเวล มีค่าต่ำกว่า 1 ที่ความ  
 สามารถระดับสูง ( $1.00 < \theta < 3.00$ ) แต่มีค่าสูงกว่า 1 ที่ระดับความสามารถระดับ

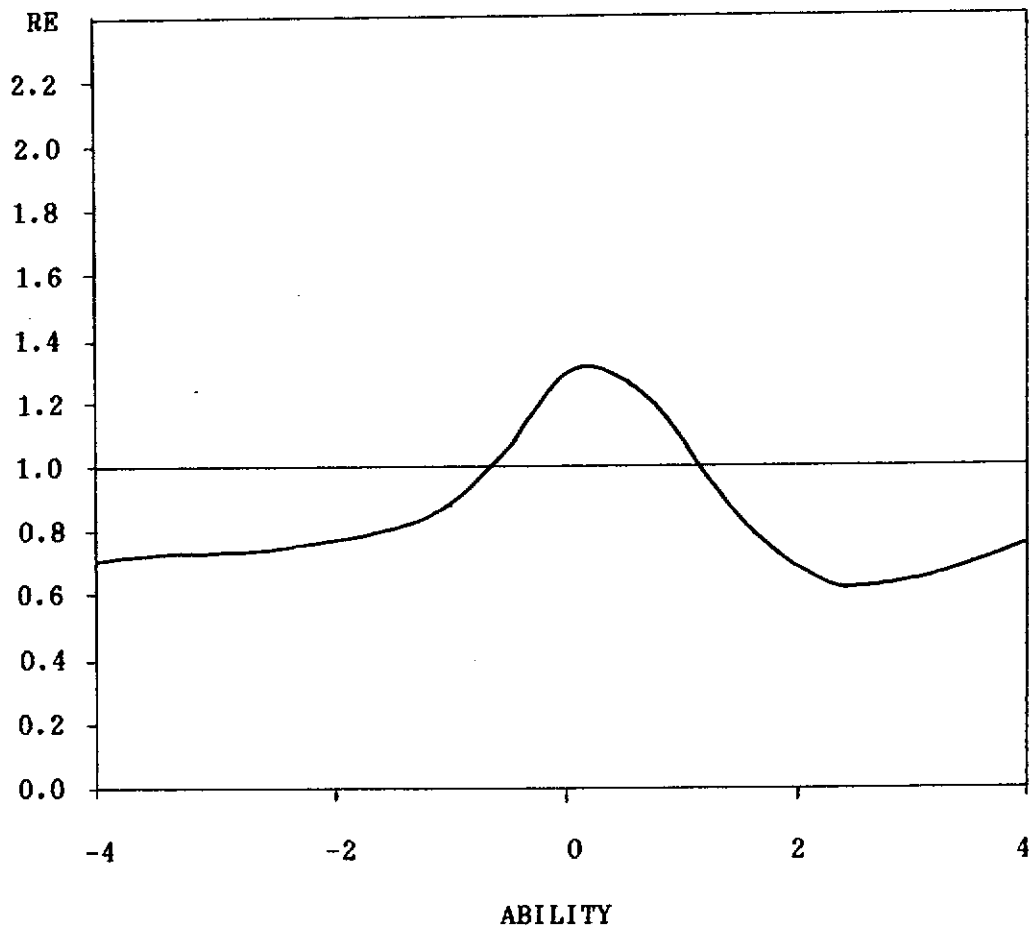
ต่ำมาก ( $\theta < -3.00$ ) และระดับสูงมาก ( $\theta > 3.00$ ) ส่วนที่ความสามารถระดับต่ำ ( $-3.00 < \theta < -1.00$ ) จะมีค่าประสิทธิภาพสัมพัทธ์ต่ำกว่า 1 และสูงกว่า 1 โดยมีค่าสูงกว่า 1 ที่ความสามารถเท่ากับ  $-3.00$ ,  $-2.90$ ,  $-2.80$  และ  $-2.60$  ( $-3.00 < \theta < -2.60$ ) เพียง 5 ระดับความสามารถเท่านั้น นอกนั้นที่ความสามารถตั้งแต่  $-2.50$  ถึง  $-1.10$  ( $-2.50 < \theta < -1.10$ ) มีค่าต่ำกว่า 1 แสดงว่าที่ความสามารถระดับต่ำนี้ส่วนมากมีค่าประสิทธิภาพสัมพัทธ์ต่ำกว่า 1 สำหรับที่ความสามารถระดับปานกลาง ( $-1.00 < \theta < 1.00$ ) จะมีค่าประสิทธิภาพสัมพัทธ์ต่ำกว่า 1 และสูงกว่า 1 โดยมีค่าต่ำกว่า 1 ที่ความสามารถเท่ากับ  $-1.00$ ,  $-0.90$ ,  $-0.80$ ,  $-0.70$  และ  $1.00$  ( $-1.00 < \theta < -0.70$  และ  $\theta = 1.00$ ) เพียง 5 ระดับความสามารถเท่านั้น นอกนั้นที่ความสามารถตั้งแต่  $-0.60$  ถึง  $0.90$  ( $-0.60 < \theta < 0.90$ ) มีค่าสูงกว่า 1 แสดงว่าที่ความสามารถระดับปานกลางนี้ส่วนมากมีค่าประสิทธิภาพสัมพัทธ์สูงกว่า 1 ดังนั้นจึงสรุปได้ว่า ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบสองขั้นตอนเทียบกับแบบทดสอบแฟล็กชีเลเวล มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำมาก ( $\theta < -3.00$ ) ระดับปานกลาง ( $-1.00 < \theta < 1.00$ ) และระดับสูงมาก ( $\theta > 3.00$ ) แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำ ( $-3.00 < \theta < -1.00$ ) และระดับสูง ( $1.00 < \theta < 3.00$ ) นั่นคือ แบบทดสอบสองขั้นตอนมีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่าแบบทดสอบแฟล็กชีเลเวล ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำมาก ระดับปานกลาง และระดับสูงมาก แต่จะมีประสิทธิภาพในการประมาณค่าความสามารถต่ำกว่าแบบทดสอบแฟล็กชีเลเวล ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำและระดับสูง

เพื่อความชัดเจนมากยิ่งขึ้น ผู้วิจัยจึงนำค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ ระหว่างแบบทดสอบที่แตกต่างกัน 3 รูปแบบ มาเสนอในรูปของกราฟ ดังภาพประกอบ 14, 15 และ 16



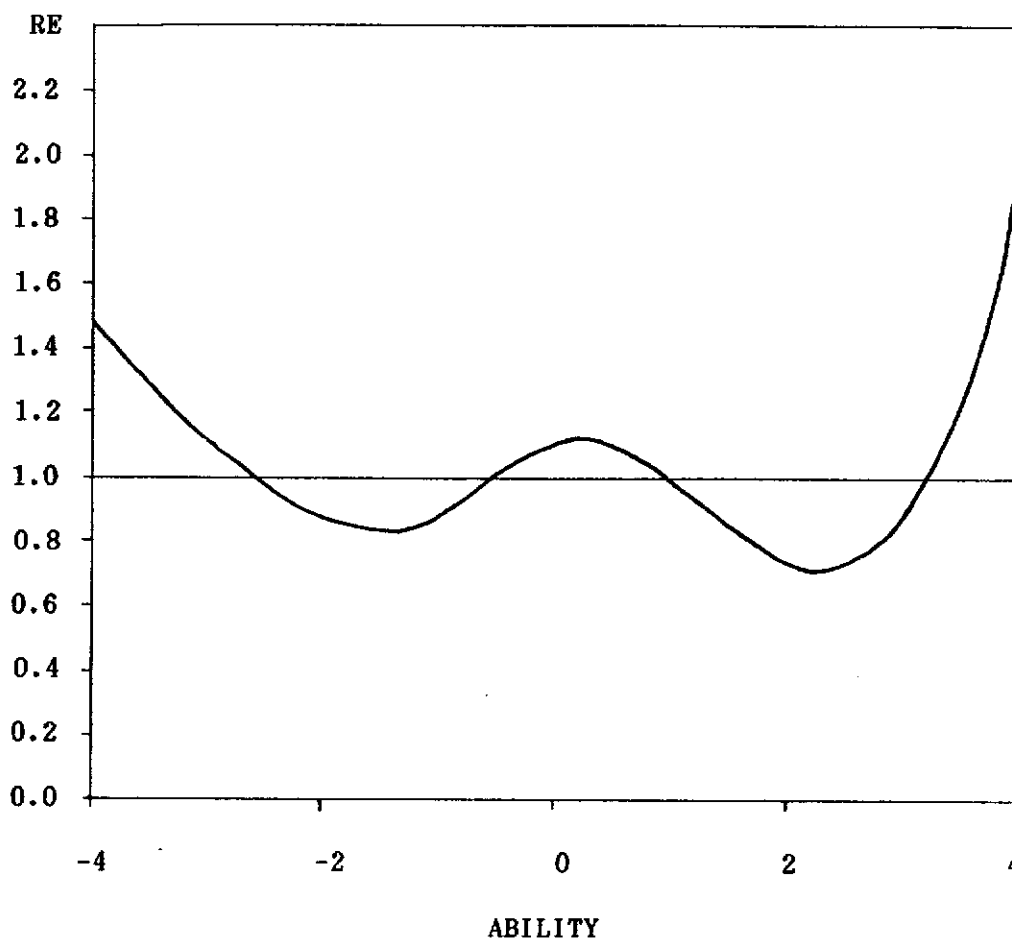
ภาพประกอบ 14 ได้ังประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ  
ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบสองขั้นตอน

จากภาพประกอบ 14 แสดงให้เห็นว่าค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่า  
ความสามารถของผู้สอบ ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบสองขั้นตอน มีค่าสูง  
กว่า 1 ในช่วงระดับความสามารถตั้งแต่  $-0.90$  ถึง  $1.10$  แต่มีค่าต่ำกว่า 1 ในช่วง  
ระดับความสามารถตั้งแต่  $-4.00$  ถึง  $-1.00$  และ  $1.20$  ถึง  $4.00$  นั่นคือ แบบทดสอบ  
ดั้งเดิมจะมีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่าแบบทดสอบสองขั้นตอน  
ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง ( $-0.90 < \theta < 1.10$ ) แต่จะมีประสิทธิภาพ  
ในการประมาณค่าความสามารถต่ำกว่าแบบทดสอบสองขั้นตอน ในกลุ่มผู้สอบที่มีความสามารถ  
ระดับต่ำลงมา ( $\theta < -0.90$ ) และระดับสูงขึ้นไป ( $\theta > 1.10$ )



ภาพประกอบ 15 ได้ังประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบระหว่าง  
แบบทดสอบดั้งเดิมเทียบกับแบบทดสอบแฟล็กชีเลเวล

จากภาพประกอบ 15 แสดงให้เห็นว่าค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่า  
ความสามารถของผู้สอบ ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบแฟล็กชีเลเวล มีค่า  
สูงกว่า 1 ในช่วงระดับความสามารถตั้งแต่  $-0.70$  ถึง  $1.00$  แต่มีค่าต่ำกว่า 1 ในช่วง  
ระดับความสามารถตั้งแต่  $-4.00$  ถึง  $-0.80$  และ  $1.10$  ถึง  $4.00$  นั่นคือ แบบทดสอบ  
ดั้งเดิมจะมีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่าแบบทดสอบแฟล็กชีเลเวล  
ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง ( $-0.70 < \theta < 1.00$ ) แต่จะมีประสิทธิภาพ  
ในการประมาณค่าความสามารถต่ำกว่าแบบทดสอบแฟล็กชีเลเวล ในกลุ่มผู้สอบที่มีความ  
สามารถระดับต่ำลงมา ( $\theta < -0.70$ ) และระดับสูงขึ้นไป ( $\theta > 1.00$ )



ภาพประกอบ 16 ไค้งประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ  
ระหว่างแบบทดสอบสองขั้นตอนเทียบกับแบบทดสอบแฟล็กซีเลเวล

จากภาพประกอบ 16 แสดงให้เห็นว่าค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่า  
ความสามารถของผู้สอบ ระหว่างแบบทดสอบสองขั้นตอนเทียบกับแบบทดสอบแฟล็กซีเลเวล  
มีค่าสูงกว่า 1 ในช่วงระดับความสามารถตั้งแต่ -4.00 ถึง -2.60, -0.60 ถึง 0.90  
และ 3.10 ถึง 4.00 แต่มีค่าต่ำกว่า 1 ในช่วงระดับความสามารถตั้งแต่ -2.50 ถึง -0.70  
และ 1.00 ถึง 3.00 นั่นคือ แบบทดสอบสองขั้นตอนจะมีประสิทธิภาพในการประมาณค่าความ  
สามารถสูงกว่าแบบทดสอบแฟล็กซีเลเวล ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำมาก ( $\theta <$   
 $-2.50$ ) ระดับปานกลาง ( $-0.60 \leq \theta \leq 0.90$ ) และระดับสูงมาก ( $\theta > 3.00$ ) แต่จะมี  
ประสิทธิภาพในการประมาณค่าความสามารถต่ำกว่าแบบทดสอบแฟล็กซีเลเวล ในกลุ่มผู้สอบที่มี  
ความสามารถระดับต่ำ ( $-2.50 \leq \theta < -0.60$ ) และระดับสูง ( $0.90 < \theta \leq 3.00$ )

## บทที่ 5

### สรุป อภิปรายผล และข้อเสนอแนะ

#### ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อศึกษาค่าความเที่ยงตรงของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมกับแบบทดสอบสองขั้นตอน แบบทดสอบดั้งเดิมกับแบบทดสอบเพิลิกซีเลเวล และแบบทดสอบสองขั้นตอนกับแบบทดสอบเพิลิกซีเลเวล
2. เพื่อศึกษาค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบสองขั้นตอน แบบทดสอบดั้งเดิมเทียบกับแบบทดสอบเพิลิกซีเลเวล และแบบทดสอบสองขั้นตอนเทียบกับแบบทดสอบเพิลิกซีเลเวล

#### กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2536 ของโรงเรียนสังกัดกรมสามัญศึกษา ส่วนกลาง กลุ่มที่ 5 กรุงเทพมหานคร จำนวนโรงเรียน 9 โรงเรียน จำนวนห้องเรียน 9 ห้องเรียน จำนวนนักเรียน 441 คน ซึ่งได้มาจากการสุ่มแบบแบ่งชั้น (Stratified Random Sampling)

#### เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบเลือกตอบวัดผลสัมฤทธิ์ทางการเรียน ชนิด 5 ตัวเลือก ซึ่งผู้วิจัยสร้างขึ้นเองจากเนื้อหาในวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3 (ค 011) เรื่องเลขยกกำลัง ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) โดยมี 3 รูปแบบ ดังนี้

1. แบบทดสอบดั้งเดิม มีจำนวน 30 ข้อ
2. แบบทดสอบสองขั้นตอน ได้แก่
  - 2.1 แบบทดสอบแยกกลุ่ม มีจำนวน 10 ข้อ
  - 2.2 แบบทดสอบวัดผล มีจำนวน 4 ฉบับ ๆ ละ 30 ข้อ
3. แบบทดสอบเพิลิกซีเลเวล มีจำนวน 59 ข้อ

### วิธีดำเนินการเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลในครั้งนี้ ผู้วิจัยดำเนินการตามขั้นตอน ดังนี้

1. วางแผนในการดำเนินการสอบ โดยการติดต่อขอความร่วมมือในการนำแบบทดสอบไปทดสอบกับนักเรียนที่ใช้เป็นกลุ่มตัวอย่าง เพื่อกำหนดวัน เวลา ในการดำเนินการสอบ
2. นำแบบทดสอบทั้ง 3 รูปแบบ คือแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กชีลเเวล ไปทดสอบกับกลุ่มตัวอย่าง จำนวน 441 คน โดยนักเรียนแต่ละคน จะต้องทำแบบทดสอบทั้ง 3 รูปแบบ สำหรับแบบทดสอบทั้ง 3 รูปแบบนั้น ผู้วิจัยจะสุ่มแต่ละรูปแบบไปทดสอบกับกลุ่มตัวอย่าง จำนวน 3 ครั้ง โดยทั้งช่วงการทดสอบแต่ละครั้งเป็นเวลา 1 สัปดาห์
3. นำผลการสอบในข้อ 2 มาจัดเตรียมข้อมูลสำหรับการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ แล้วนำผลการประมาณค่า ไปวิเคราะห์ทางสถิติ และทดสอบสมมติฐานของการวิจัยต่อไป

### การเสนอผลการวิเคราะห์ข้อมูล

ในการศึกษาครั้งนี้ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับ ดังนี้

1. ผลการหาค่าความเที่ยงตรงและการทดสอบความแตกต่างของค่าความเที่ยงตรงของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กชีลเเวล
2. ผลการหาค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ระหว่างแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กชีลเเวล

### สรุปผลการวิเคราะห์ข้อมูล

1. ค่าความเที่ยงตรงของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมกับแบบทดสอบสองขั้นตอน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
2. ค่าความเที่ยงตรงของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมกับแบบทดสอบแฟล็กชีลเเวล มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ
3. ค่าความเที่ยงตรงของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบสองขั้นตอนกับแบบทดสอบแฟล็กชีลเเวล มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4. ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบสองขั้นตอน มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง ( $-0.90 < \theta < 1.10$ ) และมีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมา ( $\theta < -0.90$ ) และระดับสูงขึ้นไป ( $\theta > 1.10$ )

5. ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบเพิลิกิเลเวล มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง ( $-0.70 < \theta < 1.00$ ) และมีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมา ( $\theta < -0.70$ ) และระดับสูงขึ้นไป ( $\theta > 1.00$ )

6. ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบสองขั้นตอนเทียบกับแบบทดสอบเพิลิกิเลเวล มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำมาก ( $\theta < -2.50$ ) ระดับปานกลาง ( $-0.60 < \theta < 0.90$ ) และระดับสูงมาก ( $\theta > 3.00$ ) และมีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำ ( $-2.50 < \theta < -0.60$ ) และระดับสูง ( $0.90 < \theta < 3.00$ )

#### อภิปรายผล

1. ค่าความเที่ยงตรงของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมกับแบบทดสอบสองขั้นตอน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ซึ่งเป็นไปตามสมมติฐานที่ตั้งไว้ในข้อ 1 โดยแบบทดสอบสองขั้นตอนมีค่าความเที่ยงตรงของการประมาณค่าความสามารถสูงกว่าแบบทดสอบดั้งเดิมอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และยังสอดคล้องกับผลการศึกษาของ ลินน์ ร็อค และเคลียร์ (Linn, Rock and Cleary. 1969 : 145) ที่พบว่า แบบทดสอบสองขั้นตอนมีค่าความเที่ยงตรงเท่ากับหรือสูงกว่าแบบทดสอบดั้งเดิม ทั้งนี้อาจเนื่องมาจาก แบบทดสอบสองขั้นตอนซึ่งเป็นรูปแบบหนึ่งของการทดสอบแบบเทเลอร์ ที่มีการจัดข้อสอบให้มีค่าความยากเหมาะสมกับระดับความสามารถของผู้สอบเป็นรายบุคคล (Hambleton and Swaminathan. 1985 : 296 - 297) โดยผู้สอบที่มีความสามารถระดับต่ำจะได้ข้อสอบฉบับที่ง่าย ส่วนผู้ที่มีความสามารถระดับสูงขึ้นไปก็จะได้ทำข้อสอบฉบับที่ยากขึ้นไปตามลำดับ ดังนั้นผู้สอบแต่ละคนจะได้ข้อสอบที่มีค่าความยากเหมาะสมกับระดับความสามารถของเขา ซึ่งจะส่งผลให้การประมาณค่าความสามารถของผู้สอบมีความถูกต้อง

แม่นยำตรงกับสภาพที่เป็นจริงมากยิ่งขึ้น สำหรับแบบทดสอบดั้งเดิมที่ใช้กันอยู่ทั่ว ๆ ไปนั้น จะประกอบด้วยข้อสอบที่มีค่าความยากระดับปานกลางของกลุ่มผู้สอบ ซึ่ง ไม่คำนึงถึงความแตกต่างของระดับความสามารถในกลุ่มผู้สอบ โดยกลุ่มผู้สอบที่มีความสามารถระดับต่ำหรือระดับสูง จะตอบข้อสอบที่มีค่าความยากไม่ตรงกับระดับความสามารถของเขา จึงมีผลทำให้การประมาณค่าความสามารถของผู้สอบมีความคลาดเคลื่อนจากสภาพที่เป็นจริง นั่นคือ แบบทดสอบสองขั้นตอนจะมีค่าความเที่ยงตรงของการประมาณค่าความสามารถของผู้สอบสูงกว่าแบบทดสอบดั้งเดิม

2. ค่าความเที่ยงตรงของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมกับแบบทดสอบแฟล็กซีเลเวล มีค่าแตกต่างกันอย่าง ไม่มีนัยสำคัญทางสถิติ ซึ่ง ไม่เป็นไปตามสมมติฐานที่ตั้งไว้ในข้อ 2 แสดงว่าแบบทดสอบแฟล็กซีเลเวลถึงแม้ว่าจะมีการจัดข้อสอบให้มีค่าความยากเหมาะสมกับระดับความสามารถของผู้สอบแล้วก็ตาม แต่ค่าความเที่ยงตรงของการประมาณค่าความสามารถ ไม่แตกต่าง ไปจากแบบทดสอบดั้งเดิม ทั้งนี้อาจเป็นเพราะว่า ในการทำแบบทดสอบแฟล็กซีเลเวล ผู้สอบแต่ละคนจะต้องเริ่มต้นด้วยการตอบข้อสอบข้อแรกที่มีค่าความยากตรงกึ่งกลางของข้อสอบทั้งหมด (Lord, 1971a : 805) จึงทำให้ผู้สอบที่มีความสามารถระดับปานกลางจะตอบข้อสอบข้อแรกเท่านั้นที่ตรงกับระดับความสามารถของเขา ส่วนข้อสอบข้อต่อ ๆ ไป จะมีค่าความยากเบนห่างออกจากระดับความสามารถ โดย ไม่ยากเกินไป ก็ง่ายเกินไป ส่วนผู้สอบที่มีความสามารถระดับต่ำหรือระดับสูงนั้น จะตอบข้อสอบข้อแรก ๆ ในการกำหนดทิศทางเพื่อที่จะนำไปสู่ข้อสอบที่มีค่าความยากตรงกับระดับความสามารถของเขา และเมื่อได้ตอบข้อสอบที่ตรงกับระดับความสามารถของเขาแล้ว ข้อสอบข้อต่อ ๆ ไป จะเป็นข้อสอบที่มีระดับความยากเบนห่างออกไปจากระดับความสามารถของผู้สอบอีกครั้งหนึ่ง ผลกระทบที่เกิดขึ้นในลักษณะนี้อาจเป็นสาเหตุที่ทำให้ผู้สอบตอบข้อสอบด้วยการเดา ซึ่ง ไม่แตกต่าง ไปจากแบบทดสอบดั้งเดิม นั่นคือ ค่าความเที่ยงตรงของการประมาณค่าความสามารถของผู้สอบ ระหว่างแบบทดสอบดั้งเดิมกับแบบทดสอบแฟล็กซีเลเวลมีค่า ใกล้ เคียงกัน

3. ค่าความเที่ยงตรงของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบสองขั้นตอนกับแบบทดสอบแฟล็กซีเลเวล มีค่าแตกต่างกันอย่าง ไม่มีนัยสำคัญทางสถิติ ซึ่ง ไม่เป็นไปตามสมมติฐานที่ตั้งไว้ในข้อ 3 ทั้งนี้อาจเนื่องมาจาก แบบทดสอบสองขั้นตอนกับแบบทดสอบแฟล็กซีเลเวล ซึ่งเป็นรูปแบบหนึ่งของการทดสอบแบบเทเลอร์ ที่มีการจัดข้อสอบให้มีค่าความยากเหมาะสมกับระดับความสามารถของผู้สอบ ถึงแม้ว่าจะมีรูปแบบของแบบทดสอบ

วิธีการสร้าง และการดำเนินการตอบข้อสอบที่มีความแตกต่างกันก็ตาม แต่จะให้ค่าความเที่ยงตรงของการประมาณค่าความสามารถของผู้สอบได้ใกล้เคียงกัน นั่นคือ ถ้าค่าหนึ่งถึงคุณภาพของแบบทดสอบด้านความเที่ยงตรงของการประมาณค่าความสามารถ แล้วจะเลือกใช้แบบทดสอบสองขั้นตอนหรือแบบทดสอบเพิลิกซ์เลเวลก็ได้

4. ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบสองขั้นตอน มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง ( $-0.90 < \theta < 1.10$ ) แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมา ( $\theta < -0.90$ ) และระดับสูงขึ้นไป ( $\theta > 1.10$ ) ซึ่งเป็นไปตามสมมติฐานที่ตั้งไว้ในข้อ 4 แสดงว่าแบบทดสอบดั้งเดิมมีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงกว่าแบบทดสอบสองขั้นตอน ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง แต่มีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบต่ำกว่าแบบทดสอบสองขั้นตอน ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมาและระดับสูงขึ้นไป ทั้งนี้อาจเป็นเพราะว่า การหาค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถระหว่างแบบทดสอบจะขึ้นอยู่กับค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบที่นำมาเปรียบเทียบกัน โดยค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบจะเป็นดัชนีที่บ่งชี้ถึงความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบ กล่าวคือ ๕ ระดับความสามารถใดที่ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบมีค่าสูง แสดงว่าแบบทดสอบฉบับนั้นมีความแม่นยำในการประมาณค่าความสามารถที่ระดับนั้นจะมีค่ามาก แต่ถ้า ๕ ระดับความสามารถใดที่ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบมีค่าต่ำ แสดงว่าแบบทดสอบฉบับนั้นจะมีความแม่นยำในการประมาณค่าความสามารถที่ระดับนั้นจะมีค่าน้อย (Hambleton and Cook, 1977 : 84) และผลจากการศึกษาในครั้งนี้พบว่า ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบดั้งเดิมมีค่าสูงกว่าแบบทดสอบสองขั้นตอน ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง แต่มีค่าต่ำกว่าแบบทดสอบสองขั้นตอน ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมาและระดับสูงขึ้นไป ซึ่งมีรายละเอียดดังตาราง 14 ในภาคผนวก ง ส่วนสาเหตุที่เป็นเช่นนี้อาจเนื่องมาจาก แบบทดสอบสองขั้นตอนมีการจัดข้อสอบให้มีความเหมาะสมกับระดับความสามารถของผู้สอบ ถึงแม้ว่าผู้สอบจะมีความสามารถต่ำหรือสูงก็จะมีผลกระทบในการตอบข้อสอบ จึงทำให้มีความแม่นยำในการประมาณค่าความสามารถสูงกว่าแบบทดสอบดั้งเดิม ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำหรือ

ระดับสูง ส่วนแบบทดสอบดั้งเดิมจะประกอบด้วยข้อสอบที่มีค่าความยากระดับปานกลาง จึงทำให้มีความแม่นยำในการประมาณค่าความสามารถสูงกว่าแบบทดสอบสองขั้นตอน ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง จากเหตุผลดังที่กล่าวมาแล้วนั้น ย่อมส่งผลทำให้การคำนวณค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบสองขั้นตอน มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมาและระดับสูงขึ้นไป

5. ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบเพ็ลิกซ์เลเวล มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง ( $-0.70 < \theta < 1.00$ ) แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมา ( $\theta < -0.70$ ) และระดับสูงขึ้นไป ( $\theta > 1.00$ ) ซึ่งเป็นไปตามสมมติฐานที่ตั้งไว้ในข้อ 5 แสดงว่าแบบทดสอบดั้งเดิมมีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงกว่าแบบทดสอบเพ็ลิกซ์เลเวล ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง แต่มีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบต่ำกว่าแบบทดสอบเพ็ลิกซ์เลเวล ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมาและระดับสูงขึ้นไป ทั้งนี้สามารถอธิบายได้โดยใช้หลักการทางเองเดียวกับกรณีแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบสองขั้นตอน กล่าวคือ จากผลการศึกษาพบว่า ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบดั้งเดิมมีค่าสูงกว่าแบบทดสอบเพ็ลิกซ์เลเวล ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง แต่มีค่าต่ำกว่าแบบทดสอบเพ็ลิกซ์เลเวล ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมาและระดับสูงขึ้นไป ซึ่งมีรายละเอียดดังตาราง 14 ในภาคผนวก ง โดยอาจมีสาเหตุมาจาก แบบทดสอบเพ็ลิกซ์เลเวลมีการปรับระดับความยากของข้อสอบให้มีความเหมาะสมกับระดับความสามารถของผู้สอบอยู่ตลอดเวลา ไม่ว่าผู้สอบจะมีความสามารถต่ำหรือสูงเกินไป จึงทำให้มีความแม่นยำในการประมาณค่าความสามารถสูงกว่าแบบทดสอบดั้งเดิม ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำหรือระดับสูง แต่ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง แบบทดสอบดั้งเดิมจะมีความแม่นยำในการประมาณค่าความสามารถสูงกว่าแบบทดสอบเพ็ลิกซ์เลเวล เนื่องจากแบบทดสอบดั้งเดิมจะประกอบด้วยข้อสอบที่มีค่าความยากระดับปานกลางจึงมีความเหมาะสมกับกลุ่มผู้สอบที่มีความสามารถระดับปานกลางเท่านั้น ด้วยเหตุผลดังที่กล่าว ไปแล้วนั้น จึงมีผลทำให้การคำนวณค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบเพ็ลิกซ์เลเวล มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถ

ระดับปานกลาง แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมาและระดับสูงขึ้นไป ผลการศึกษาดังกล่าวนี้มีความสอดคล้องกับผลการศึกษาของ ลอร์ด (Lord. 1971a : 805 - 813) และเวลล์ (Vale. 1975 : 5 - 75) ที่พบว่า แบบทดสอบดั้งเดิมจะให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่าแบบทดสอบแฟล็กซีเลเวล เมื่อผู้สอบมีความสามารถระดับปานกลาง ส่วนผู้สอบที่มีความสามารถระดับต่ำหรือระดับสูง แบบทดสอบดั้งเดิมจะให้ค่าฟังก์ชันอินฟอร์เมชันต่ำกว่าแบบทดสอบแฟล็กซีเลเวล นอกจากนี้ยังสอดคล้องกับข้อเสนอแนะของซีควิน (Sequin. 1978 : 1196-A) ที่กล่าวไว้ว่า แบบทดสอบแฟล็กซีเลเวล ไม่เหมาะสมที่จะใช้ประมาณค่าความสามารถของผู้สอบที่มีความสามารถระดับปานกลาง แต่จะใช้ได้ดีกับผู้สอบที่มีความสามารถระดับต่ำหรือระดับสูงมาก ๆ

6. ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบสองขั้นตอนเทียบกับแบบทดสอบแฟล็กซีเลเวล มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำมาก ( $\theta < -2.50$ ) ระดับปานกลาง ( $-0.60 \leq \theta \leq 0.90$ ) และระดับสูงมาก ( $\theta > 3.00$ ) แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำ ( $-2.50 \leq \theta < -0.60$ ) และระดับสูง ( $0.90 < \theta \leq 3.00$ ) ซึ่งไม่เกินไปตามสมมติฐานที่ตั้งไว้ในข้อ 6 แสดงว่าแบบทดสอบสองขั้นตอนมีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงกว่าแบบทดสอบแฟล็กซีเลเวล ไม่ครบทุกช่วงระดับความสามารถของกลุ่มผู้สอบ โดยมีประสิทธิภาพสูงกว่าแบบทดสอบแฟล็กซีเลเวล ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำมาก ระดับปานกลาง และระดับสูงมาก แต่มีประสิทธิภาพต่ำกว่าแบบทดสอบแฟล็กซีเลเวล ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำและระดับสูง ทั้งนี้เป็นผลเนื่องมาจาก ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบสองขั้นตอนส่วนมากมีค่าสูงกว่าแบบทดสอบแฟล็กซีเลเวล มีบางช่วงของระดับความสามารถเท่านั้นที่มีค่าต่ำกว่าแบบทดสอบแฟล็กซีเลเวล ซึ่งมีรายละเอียดดังตาราง 14 ในภาคผนวก ง ผลการศึกษาดังกล่าวนี้นี้ขัดแย้งกับผลการศึกษาของเวลล์ (Vale. 1975 : 5 - 75) ที่พบว่า ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบสองขั้นตอนมีค่าสูงกว่าแบบทดสอบแฟล็กซีเลเวล ทุกระดับความสามารถของกลุ่มผู้สอบ สำหรับค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบจะเป็นดัชนีที่บ่งชี้ถึงความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบ โดยถ้าค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบมีค่ามากแล้วค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า

ความสามารถจะมีค่าน้อยแสดงว่าแบบทดสอบฉบับนั้นจะมีความแม่นยำในการประมาณค่าความสามารถสูงสุดที่ระดับความสามารถนั้น ในทางหนึ่งเดียวกันถ้าค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบมีค่าน้อยแล้วค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถก็จะมีค่ามาก แสดงว่าแบบทดสอบฉบับนั้นจะมีความแม่นยำในการประมาณค่าความสามารถน้อยลงที่ระดับความสามารถนั้น (Hambleton and Cook. 1977 : 84) ด้วยเหตุผลดังกล่าวประกอบกับ ผลการศึกษาในครั้งนี้พบว่า ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบสองชิ้นตอนมีค่าสูงกว่า แบบทดสอบเฟลิกซ์เลเวล ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำมาก ระดับปานกลาง และระดับ สูงมาก แต่มีค่าต่ำกว่าแบบทดสอบเฟลิกซ์เลเวล ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำและระดับ สูง ทั้งนี้อาจมีสาเหตุมาจาก ความคลาดเคลื่อนในการประมาณค่าจากแบบทดสอบเฟลิกซ์เลเวล โดยผู้สอบที่มีความสามารถระดับปานกลางจะตอบข้อสอบข้อแรกเท่านั้นที่ตรงกับระดับความสามารถของเขา ส่วนข้อสอบข้อต่อ ๆ ไปจะมีค่าความยากเบนห่างออกจากระดับความสามารถของเขา โดยไม่ยากเกินไปก็ง่ายเกินไป สำหรับผู้สอบที่มีความสามารถระดับต่ำมากและระดับสูงมาก จะตอบ ข้อสอบข้อแรก ๆ ในการกำหนดทิศทางเพื่อที่จะนำไปสู่ข้อสอบที่ตรงกับระดับความสามารถของเขา และเมื่อได้ตอบข้อสอบที่ตรงกับระดับความสามารถของเขาแล้ว ข้อสอบข้อต่อ ๆ ไป ก็จะมีค่าความ ยากเบนห่างออกจากระดับความสามารถของเขาอีกครั้งหนึ่ง ผลกระทบที่เกิดขึ้นในลักษณะนี้อาจ เป็นสาเหตุที่ทำให้ผู้สอบต้องตอบข้อสอบด้วยการเดา ซึ่งจะส่งผลให้ความแม่นยำในการประมาณ ค่าความสามารถลดต่ำลง นั่นคือแบบทดสอบเฟลิกซ์เลเวลจะมีค่าฟังก์ชันอินฟอร์เมชันต่ำกว่าแบบ ทดสอบสองชิ้นตอนที่ระดับความสามารถนั้น ส่วนความคลาดเคลื่อนในการประมาณค่าความสามารถ จากแบบทดสอบสองชิ้นตอน อาจมีสาเหตุมาจากการที่ผู้สอบที่มีความสามารถระดับต่ำและระดับสูง มี ความผิดพลาดในการตอบแบบทดสอบแยกกลุ่มเพื่อที่จะ ไปตอบในแบบทดสอบวัดผล จึงทำให้ผู้สอบได้ ตอบข้อสอบที่มีค่าความยากไม่เหมาะสมกับระดับความสามารถของเขา ซึ่งจะมีผลทำให้ความ แม่นยำในการประมาณค่าความสามารถลดต่ำลง นั่นคือ แบบทดสอบสองชิ้นตอนมีค่าฟังก์ชันอินฟอร์ เมชันต่ำกว่าแบบทดสอบเฟลิกซ์เลเวลที่ระดับความสามารถนั้น จากเหตุผลดังที่กล่าวไปแล้วนั้น จึงมีผลทำให้แบบทดสอบสองชิ้นตอนมีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงกว่า แบบทดสอบเฟลิกซ์เลเวล ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำมาก ระดับปานกลาง และระดับ สูงมาก แต่จะมีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบต่ำกว่าแบบทดสอบ เฟลิกซ์เลเวล ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำและระดับสูง

### ข้อเสนอแนะ

1. ในการเลือกใช้แบบทดสอบเพื่อวัดผลสัมฤทธิ์ทางการเรียนให้มีความถูกต้องแม่นยำสูง ควรเลือกใช้แบบทดสอบที่มีความเหมาะสมกับระดับความสามารถของผู้สอบ โดยในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง ควรเลือกใช้แบบทดสอบดั้งเดิม แต่ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำหรือระดับสูง ควรเลือกใช้แบบทดสอบสองขั้นตอนหรือแบบทดสอบเพลิกซ์เลเวล
2. ควรมีการศึกษาในลักษณะเดียวกัน โดยใช้รูปแบบของการทดสอบแบบเทเลอร์ชนิดอื่น ๆ ที่แตกต่างกันออกไป เช่น แบบทดสอบรูปปริมาตร แบบทดสอบสามขั้นตอน แบบทดสอบปรับระดับขั้น ยุทธวิธีความเป็นไปได้สูงสุด และยุทธวิธีของเบส์ เป็นต้น และควรเปลี่ยนวิธีดำเนินการสอบโดยนำเครื่องคอมพิวเตอร์มาประยุกต์ใช้กับการทดสอบแบบเทเลอร์ แทนการทดสอบที่ใช้การเขียนตอบแบบเดิม
3. ควรมีการศึกษาในสถานการณ์จำลอง ซึ่งอาจใช้เทคนิคมอนติ คาร์โล (Monte Carlo Method) โดยใช้เครื่องคอมพิวเตอร์ช่วยในการสร้างสถานการณ์จำลองข้อมูลแทนการทดสอบในสถานการณ์จริง
4. ควรมีการศึกษาเพิ่มเติมในแง่ผลกระทบต่อประสิทธิภาพของการประมาณค่าความสามารถของผู้สอบ อันเนื่องมาจากจำนวนข้อสอบ เวลาในการทดสอบ วิธีการให้คะแนน และค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ เป็นต้น
5. ในการนำทฤษฎีการตอบสนองข้อสอบมาใช้ประมาณค่าพารามิเตอร์ น่าจะลองใช้โมเดลแบบอื่น ๆ ที่แตกต่างไปจาก ราส์ช โมเดล เช่น โลจิสติก โมเดล ที่มีพารามิเตอร์ 2 หรือ 3 ตัว ส่วนวิธีการประมาณค่าอาจนำวิธีการประมาณค่าความเป็นไปได้สูงสุดแบบมีเงื่อนไข หรือมีวิธีการประมาณค่าแบบเบส์ แทนวิธีการประมาณค่าความเป็นไปได้สูงสุดแบบไม่มีเงื่อนไข เพื่อจะได้ทราบว่า ให้ผลแตกต่าง ไปจากการศึกษาในครั้งนี้อหรือไม่ อย่างไร

บรรณานุกรม

## บรรณานุกรม

- เกรียงศักดิ์ สุวรรณภาค. ความสัมพันธ์ระหว่างรูปแบบทางรูปแบบของแบบทดสอบแยกกลุ่มกับความสามารถทางการเรียน ในวิชาคณิตศาสตร์. ปรินทิฟอนท์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2532. อัดสำเนา.
- งามนิตย์ ธาตุทอง และปรีชา เครือวรรณ. "การให้คะแนนตัวเองในการสอบแบบทดสอบแฟล็กชีลเเวล," รายงานการวิจัย. ขอนแก่น : มหาวิทยาลัยขอนแก่น, 2527.
- จักรนภา พรหมมา. "Tailored Testing : ผลสำเร็จของการประยุกต์ใช้ Latent Trait Theory," วารสารการวัดผลการศึกษา. 4(2) : 45 - 53 ; กันยายน - ธันวาคม 2525.
- นริศ สวัสดิ์. การศึกษาความสัมพันธ์ระหว่างคะแนนของแบบทดสอบสองขั้นตอนและคะแนนของแบบทดสอบรูปปริมาตร. ปรินทิฟอนท์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2533. อัดสำเนา.
- ผจงจิต อินทสุวรรณ. "Latent Trait Theory," วารสารการวัดผลการศึกษา. 3(3) : 51 - 69 ; มกราคม - เมษายน 2525.
- ล้วน สายยศ และอังคณา สายยศ. สถิติวิทยาทางการศึกษา. กรุงเทพฯ : วัฒนาพานิช, 2522.
- \_\_\_\_\_. หลักการวิจัยทางการศึกษา. กรุงเทพฯ : ทวีกิจการพิมพ์, 2524.
- สุธรรม์ จันทรหอม. "แบบทดสอบต่อเนื่อง : Tailored Test," วารสารการวัดผลการศึกษา. 7 (21) : 10 - 22 ; มกราคม - เมษายน 2529.
- สำเร็จ บุญเรืองรัตน์. "การพัฒนาทฤษฎีเลแทนท์เพื่อการวิเคราะห์ข้อสอบ," วารสารการวัดผลการศึกษา. 4 (2) : 1 - 13 ; กันยายน - ธันวาคม 2525.
- เสรี เปรมขวลิตโรจน์. การศึกษาความเที่ยงตรงในการจำแนกความรู้ และการประมาณค่าความสามารถในวิชาคณิตศาสตร์ของแบบทดสอบแฟล็กชีลเเวลและแบบทดสอบรูปปริมาตรที่ใช้ขนาดขั้นแปรผัน. ปรินทิฟอนท์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2531. อัดสำเนา.

- Angoff, W.H. and E.M. Huddleston. "The Multiple - Level Experiment : A Study of a Two - Stage Test System for the College Board Scholastic Aptitude Test," Statistical Report 58 - 21. Princeton N.J. : Educational Testing Service, 1958.
- Betz, N.E. and D.J. Weiss. "An Empirical Study of Computer - Administered Two - Stage Ability Testing," Research Report 73 - 4. Minneapolis : University of Minnesota, Psychometric Methods Program, Department of Psychology, 1973.
- Birnbaum, Allan. "Some Latent Trait Model and Their Use in Inferring in Examinee's Ability," In F.M. Novick. Statistical Theories of Mental Test Scores. Reading, Massachusetts : Addison - Wesley, 1968.
- Green, Bert F. and others. "Technical Guidelines for Assessing Computerized Adaptive Tests," Journal of Educational Measurement. 21(4) : 347 - 360 ; Winter, 1984.
- Hambleton, R.K. and others. "Developments in Latent Trait Theory : Model, Technical Issues, and Applications," Review of Educational Research. 48 (4) : 467 - 510, 1987.
- Hambleton, R.K. and Hariharan Swaminathan. Item Response Theory. Boston : Kluwer - Nijhoff Publishing, 1985.
- Hambleton, R.K. and Linda L. Cook. "Latent Trait Models and Their Use in the Analysis of Educational Test Data," Journal of Education Measurement. 14(2) : 75 - 96 ; Summer. 1977.
- Larkin, K.C. and Weiss, D.J. "An Empirical Comparison of Two - Stage and Pyramidal Adaptive Ability Testing," Research Report 75 - 1. Minneapolis : University of Minnesota, Department of Psychology, Psychometric Methods Program, 1975.
- Linn, R.L., D.A. Rock. and T.A. Cleary. "The Development and Evaluation of Several Programmed Testing Methods," Educational and Psychological Measurement. 29 (1) : 129 - 146, 1969.
- ✧ Lord, F.M. "A Theoretical Study of the Measurement Effectiveness of Flexilevel Tests," Educational and Psychological Measurement. 31(4) : 805 - 813, 1971 a.
- Lord, F.M. "Robbins - Monro Procedures for Tailored Testing," Educational and Psychological Measurement. 31 (1) : 3 - 31, 1971b.
- ✧ ———. "The Self - Scoring Flexilevel Test," Journal of Educational Measurement. 8 (3) : 147 - 151, 1971 c.

- Lord, F.M. "A Theoretical Study of Two - Stage Testing," Psychometrika. 36 (3) : 227 - 242, 1971 d.
- \_\_\_\_\_. "Practical Applications of Item Characteristic Curve Theory," Journal of Educational Measurement. 14 (2) : 117 - 138, 1977.
- \_\_\_\_\_. Applications of Item Response Theory to Practical Testing Problem. Hillsdale, New Jersey : Lawrence Erlbaum Associates, 1980.
- Lord, F.M. and M.R. Novick. Statistical Theories of Mental Test Scores. Massachusetts : Addison - Wesley, 1968.
- Moreno, K.E. and others. "Relationship Between Corresponding Armed Services Vocational Aptitude Battery and Computerized Adaptive Testing Subtests," Applied Psychological Measurement. 8(2) : 155 - 163, 1984.
- \* Olivier, P.A. "An Evaluation of the Self - Scoring Flexilevel Tailored Testing Model," Dissertation Abstracts International. 35 (7) : 4257 - A, 1974.
- \* Rudner, Lawrence, "A Closer Look at Latent Trait Parameter Invariance," Educational and Psychological Measurement. 43 (4) : 951 - 955, 1983.
- Schmidt, F.L., V.W. Urry and J.F. Gugel. "Computer Assisted Tailored Testing : Examinee Reactions and Evaluations," Educational and Psychological Measurement. 38 (2) : 265 - 273, 1978.
- Sequin, S.P. "An Exploratory Study of the Efficiency of the Flexilevel Testing Procedure," Dissertation Abstracts International. 39 (3) : 1196 - A, 1978.
- Spinetti, J.P. and R.K. Hambleton. "A Computer Simulation Study of Tailored Testing Strategies for Objective - Based Instructional Program," Educational and Psychological Measurement. 37 (1) : 139 - 158, 1977.
- Steel, Robert G.D. and James H. Torrie. Principles and Procedures of Statistics. 2nd ed. Singapore : McGraw - Hill, Inc., 1981.
- Urry, V.W. "Tailored Testing : A Successful Application of Latent Trait Theory," Journal of Educational Measurement. 14 (2) : 181 - 196, 1977.

- Vale, C.D. "Problem : Strategies of Branching Through an Item Pool," In th "Computerized Adaptive Trait Measurement : Problems and Prospects," Research Report 75 - 5. D.J. Weiss, Ed. Minneapolis : University of Minnesota, Department of Psychology, Psychometric Methods Program, 1975.
- Vale, C.D. and David J. Weiss. "A Simulation Study of Stradative Ability Testing," Research Report 75 - 6. Psychometric Method Program. Department of Psychology, University of Minnesota Minneapolis, 1975.
- Vijver, Fons J.R. Van De. "The Robustness of Rasch Estimates," Applied Psychological Measurement. 10 (1) : 45 - 57, 1986.
- Warm, Thomas A. "A Primer of Item Response Theory," Technical Report 941278. Oklahoma : U.S. Coast Guard Institute, 1978.
- Waters, C.W. and A.G. Bayroff. "A Comparison of Computer Simulation Conventional and Branching Test," Educational and Psychological Measurement. 31 (1) 125 - 136, 1971.
- Weiss, D.J. "Strategies of Adaptive Ability Measurement," Research Report 74 - 5. Minneapolis : University of Minnesota, Psychometric Methods Program, Department of Psychology, 1974.
- . "Improving Measurement Quality and Efficiency with Adaptive Testing," Applied Psychological Measurement. 6 : 473 - 492, 1982.
- \*Weiss, D.J. and G. Gage Kingsbury. "Application of Computerized Adaptive Testing to Educational Problems," Journal of Educational Measurement. 21 (4) : 361 - 375, 1984.
- Wright, Benjamin D. and N. Panchapakesan. "A Procedure for Sample Free Item Analysis," Educational and Psychological Measurement. 29 : 23 - 48, 1969.
- Wright, Benjamin D. and Mark H. Stone. Best Test Design. Chicago : MESA Press, 1979.

ภาคผนวก

ภาคผนวก ก

ผลการวิเคราะห์องค์ประกอบที่มีค่าไอเกนมากกว่า 1

เพื่อตรวจสอบความเป็นมิติเดียวของแบบทดสอบที่นำไปทดลองสอบ จำนวน 6 ฉบับ

ตาราง 6 ค่าไอเกน (มากกว่า 1) เพอร์เซ็นต์ของความแปรปรวน และเพอร์เซ็นต์ของความแปรปรวนสะสม จากการนำแบบทดสอบ จำนวน 6 ฉบับ มาวิเคราะห์องค์ประกอบ

| แบบทดสอบ<br>ฉบับที่ | องค์ประกอบที่ | ค่าไอเกน | เพอร์เซ็นต์ของความ<br>แปรปรวน | เพอร์เซ็นต์ของความ<br>แปรปรวนสะสม |
|---------------------|---------------|----------|-------------------------------|-----------------------------------|
| 1                   | 1             | 6.74523  | 16.9                          | 16.9                              |
|                     | 2             | 2.42328  | 6.1                           | 22.9                              |
|                     | 3             | 1.87064  | 4.7                           | 27.6                              |
|                     | 4             | 1.71494  | 4.3                           | 31.9                              |
|                     | 5             | 1.56792  | 3.9                           | 35.8                              |
|                     | 6             | 1.41718  | 3.5                           | 39.3                              |
|                     | 7             | 1.34878  | 3.4                           | 42.7                              |
|                     | 8             | 1.29220  | 3.2                           | 46.0                              |
|                     | 9             | 1.20759  | 3.0                           | 49.0                              |
|                     | 10            | 1.15088  | 2.9                           | 51.8                              |
|                     | 11            | 1.12422  | 2.8                           | 54.7                              |
|                     | 12            | 1.06339  | 2.7                           | 57.3                              |
|                     | 13            | 1.00871  | 2.5                           | 59.8                              |
| 2                   | 1             | 4.26465  | 14.2                          | 14.2                              |
|                     | 2             | 1.62193  | 5.4                           | 19.6                              |
|                     | 3             | 1.50751  | 5.0                           | 24.6                              |
|                     | 4             | 1.43894  | 4.8                           | 29.4                              |
|                     | 5             | 1.33098  | 4.4                           | 33.9                              |
|                     | 6             | 1.23350  | 4.1                           | 38.0                              |
|                     | 7             | 1.16377  | 3.9                           | 41.9                              |
|                     | 8             | 1.12804  | 3.8                           | 45.6                              |

ตาราง 6 (ต่อ)

| แบบทดสอบ<br>ฉบับที่ | องค์ประกอบที่ | ค่าไอเกน | เปอร์เซ็นต์ของความ<br>แปรปรวน | เปอร์เซ็นต์ของความ<br>แปรปรวนสะสม |
|---------------------|---------------|----------|-------------------------------|-----------------------------------|
|                     | 9             | 1.11974  | 3.7                           | 49.4                              |
|                     | 10            | 1.09264  | 3.6                           | 53.0                              |
|                     | 11            | 1.04468  | 3.5                           | 56.5                              |
|                     | 12            | 1.01761  | 3.4                           | 59.9                              |
| 3                   | 1             | 6.30774  | 15.8                          | 15.8                              |
|                     | 2             | 2.59525  | 6.5                           | 22.3                              |
|                     | 3             | 2.02232  | 5.1                           | 27.3                              |
|                     | 4             | 1.93505  | 4.8                           | 32.2                              |
|                     | 5             | 1.58039  | 4.0                           | 36.1                              |
|                     | 6             | 1.51022  | 3.8                           | 39.9                              |
|                     | 7             | 1.34438  | 3.4                           | 43.2                              |
|                     | 8             | 1.26612  | 3.2                           | 46.4                              |
|                     | 9             | 1.19516  | 3.0                           | 49.4                              |
|                     | 10            | 1.11191  | 2.8                           | 52.2                              |
|                     | 11            | 1.08054  | 2.7                           | 54.9                              |
|                     | 12            | 1.03880  | 2.6                           | 57.5                              |
| 4                   | 1             | 7.74520  | 19.4                          | 19.4                              |
|                     | 2             | 2.27295  | 5.7                           | 25.0                              |
|                     | 3             | 1.74513  | 4.4                           | 29.4                              |
|                     | 4             | 1.57758  | 3.9                           | 33.4                              |
|                     | 5             | 1.48211  | 3.7                           | 37.1                              |

ตาราง 6 (ต่อ)

| แบบทดสอบ<br>ฉบับที่ | องค์ประกอบที่ | ค่าไอเกน | เปอร์เซ็นต์ของความ<br>แปรปรวน | เปอร์เซ็นต์ของความ<br>แปรปรวนสะสม |
|---------------------|---------------|----------|-------------------------------|-----------------------------------|
|                     | 6             | 1.28956  | 3.2                           | 40.3                              |
|                     | 7             | 1.23765  | 3.1                           | 43.4                              |
|                     | 8             | 1.17722  | 2.9                           | 46.3                              |
|                     | 9             | 1.14106  | 2.9                           | 49.2                              |
|                     | 10            | 1.08365  | 2.7                           | 51.9                              |
|                     | 11            | 1.05409  | 2.6                           | 54.5                              |
|                     | 12            | 1.00501  | 2.5                           | 57.0                              |
| 5                   | 1             | 6.51930  | 16.3                          | 16.3                              |
|                     | 2             | 2.06702  | 5.2                           | 21.5                              |
|                     | 3             | 1.78240  | 4.5                           | 25.9                              |
|                     | 4             | 1.67491  | 4.2                           | 30.1                              |
|                     | 5             | 1.50534  | 3.8                           | 33.9                              |
|                     | 6             | 1.41386  | 3.5                           | 37.4                              |
|                     | 7             | 1.33609  | 3.3                           | 40.7                              |
|                     | 8             | 1.28326  | 3.2                           | 44.0                              |
|                     | 9             | 1.18983  | 3.0                           | 46.9                              |
|                     | 10            | 1.14986  | 2.9                           | 49.8                              |
|                     | 11            | 1.06617  | 2.7                           | 52.5                              |
|                     | 12            | 1.03412  | 2.6                           | 55.1                              |

ตาราง 6 (ต่อ)

| แบบทดสอบ<br>ฉบับที่ | องค์ประกอบที่ | ค่าไอเกน | เปอร์เซ็นต์ของความ<br>แปรปรวน | เปอร์เซ็นต์ของความ<br>แปรปรวนสะสม |
|---------------------|---------------|----------|-------------------------------|-----------------------------------|
| 6                   | 1             | 5.51345  | 13.8                          | 13.8                              |
|                     | 2             | 1.93146  | 4.8                           | 18.6                              |
|                     | 3             | 1.72470  | 4.3                           | 22.9                              |
|                     | 4             | 1.63206  | 4.1                           | 27.0                              |
|                     | 5             | 1.55094  | 3.9                           | 30.9                              |
|                     | 6             | 1.38421  | 3.5                           | 34.3                              |
|                     | 7             | 1.33372  | 3.3                           | 37.7                              |
|                     | 8             | 1.27707  | 3.2                           | 40.9                              |
|                     | 9             | 1.22521  | 3.1                           | 43.9                              |
|                     | 10            | 1.18648  | 3.0                           | 46.9                              |
|                     | 11            | 1.15201  | 2.9                           | 49.8                              |
|                     | 12            | 1.10239  | 2.8                           | 52.5                              |
|                     | 13            | 1.04291  | 2.6                           | 55.1                              |
|                     | 14            | 1.01442  | 2.5                           | 57.7                              |

## ภาคผนวก ข

ค่าประมาณพารามิเตอร์ความยากของข้อสอบ

ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ

และค่าความสอดคล้องกับโมเดล จากแบบทดสอบจำนวน 6 ฉบับนำไปทดลองสอบ

ตาราง 7 ผลการประมาณค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าความคลาดเคลื่อน  
มาตรฐานในการประมาณค่าความยากของข้อสอบ (SEE) และค่าความสอดคล้องกับโมเดล  
(t - test) ของแบบทดสอบฉบับที่ 1

| ข้อ | b     | SEE  | t     | ข้อ | b     | SEE  | t     |
|-----|-------|------|-------|-----|-------|------|-------|
| 1   | -2.75 | 0.20 | -0.75 | 24  | -0.68 | 0.11 | -0.75 |
| 2   | -1.53 | 0.13 | -0.85 | 25  | 0.77  | 0.11 | 1.17  |
| 3   | 0.68  | 0.11 | 0.83  | 26  | 0.72  | 0.11 | -0.09 |
| 4   | -1.37 | 0.13 | 0.70  | 27  | 0.51  | 0.11 | 0.92  |
| 5   | 0.57  | 0.11 | -0.30 | 28  | -0.78 | 0.11 | -1.37 |
| 6   | -0.70 | 0.11 | 0.72  | 29  | -0.42 | 0.11 | -2.69 |
| 7   | -0.62 | 0.11 | 1.16  | 30  | -0.15 | 0.11 | -2.80 |
| 8   | -0.62 | 0.11 | 0.74  | 31  | 0.12  | 0.11 | -3.29 |
| 9   | -1.22 | 0.12 | -1.25 | 32  | -0.43 | 0.11 | 0.15  |
| 10  | -0.29 | 0.11 | -0.24 | 33  | -0.55 | 0.11 | -1.36 |
| 11  | -0.36 | 0.11 | -2.08 | 34  | -0.22 | 0.11 | -1.70 |
| 12  | -1.48 | 0.13 | -2.20 | 35  | -0.08 | 0.11 | -1.51 |
| 13  | -0.70 | 0.11 | -1.96 | 36  | 0.10  | 0.11 | 0.94  |
| 14  | -0.72 | 0.11 | -2.35 | 37  | -1.60 | 0.13 | -0.28 |
| 15  | -0.18 | 0.11 | -0.63 | 38  | -0.40 | 0.11 | -1.99 |
| 16  | -0.42 | 0.11 | -0.45 | 39  | 0.12  | 0.11 | 1.61  |
| 17  | -0.04 | 0.11 | 0.35  | 40  | 0.24  | 0.11 | 1.74  |
| 18  | 0.96  | 0.12 | 0.50  | 41  | 2.62  | 0.18 | 1.33  |
| 19  | 1.17  | 0.12 | 0.04  | 42  | 0.75  | 0.11 | 0.95  |
| 20  | 1.96  | 0.15 | 1.32  | 43  | 0.39  | 0.11 | 2.43  |
| 21  | 1.58  | 0.13 | 0.49  | 44  | 1.60  | 0.13 | 0.53  |
| 22  | 0.37  | 0.11 | 1.24  | 45  | 1.71  | 0.14 | 0.87  |
| 23  | 0.79  | 0.11 | 1.57  |     |       |      |       |

ตาราง 8 ผลการประมาณค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าความคลาดเคลื่อน  
มาตรฐานในการประมาณค่าความยากของข้อสอบ (SEE) และค่าความสอดคล้องกับโมเดล  
(t - test) ของแบบทดสอบฉบับที่ 2

| ข้อ | b     | SEE  | t     | ข้อ | b     | SEE  | t     |
|-----|-------|------|-------|-----|-------|------|-------|
| 1   | -1.39 | 0.13 | 0.60  | 24  | 0.17  | 0.11 | -2.34 |
| 2   | -0.60 | 0.11 | -0.08 | 25  | 0.09  | 0.11 | -0.87 |
| 3   | -1.39 | 0.13 | -0.05 | 26  | 1.67  | 0.13 | 0.36  |
| 4   | -0.13 | 0.11 | -1.96 | 27  | 0.45  | 0.11 | 0.69  |
| 5   | -1.20 | 0.13 | -1.04 | 28  | -0.74 | 0.12 | -1.36 |
| 6   | -0.78 | 0.12 | -2.08 | 29  | -1.25 | 0.13 | -1.17 |
| 7   | -1.29 | 0.13 | -0.25 | 30  | -0.10 | 0.11 | -1.03 |
| 8   | -1.61 | 0.14 | -1.09 | 31  | -0.15 | 0.11 | -1.58 |
| 9   | -2.12 | 0.17 | -0.89 | 32  | 0.21  | 0.11 | 3.53  |
| 10  | -0.75 | 0.12 | -0.97 | 33  | 0.21  | 0.11 | 4.07  |
| 11  | 2.63  | 0.17 | 0.01  | 34  | 1.09  | 0.12 | -2.44 |
| 12  | -0.29 | 0.11 | -1.16 | 35  | 0.49  | 0.11 | -1.19 |
| 13  | 0.13  | 0.11 | -2.95 | 36  | 0.94  | 0.11 | 1.56  |
| 14  | -0.66 | 0.12 | -2.75 | 37  | 0.07  | 0.11 | -1.36 |
| 15  | 0.09  | 0.11 | -1.71 | 38  | -0.42 | 0.11 | 1.00  |
| 16  | -0.27 | 0.11 | -0.32 | 39  | -0.95 | 0.12 | 0.52  |
| 17  | 0.17  | 0.11 | -1.55 | 40  | 1.31  | 0.12 | 0.85  |
| 18  | 1.21  | 0.12 | 0.60  | 41  | 1.44  | 0.12 | 0.21  |
| 19  | -0.27 | 0.11 | 0.25  | 42  | 0.41  | 0.11 | 3.32  |
| 20  | -1.44 | 0.14 | -0.17 | 43  | 1.61  | 0.13 | -0.43 |
| 21  | 1.54  | 0.13 | 2.18  | 44  | 0.60  | 0.11 | 1.04  |
| 22  | 0.68  | 0.11 | -0.70 | 45  | 0.00  | 0.11 | 2.12  |
| 23  | 0.61  | 0.11 | 0.22  |     |       |      |       |

ตาราง 9 ผลการประมาณค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าความคลาดเคลื่อน  
มาตรฐานในการประมาณค่าความยากของข้อสอบ (SEE) และค่าความสอดคล้องกับโมเดล  
(t - test) ของแบบทดสอบฉบับที่ 3

| ข้อ | b     | SEE  | t     | ข้อ | b     | SEE  | t     |
|-----|-------|------|-------|-----|-------|------|-------|
| 1   | -2.45 | 0.25 | 0.02  | 24  | -0.09 | 0.12 | 0.25  |
| 2   | -0.44 | 0.13 | 0.40  | 25  | 0.79  | 0.11 | 0.59  |
| 3   | 0.41  | 0.11 | 1.08  | 26  | 1.09  | 0.11 | -1.34 |
| 4   | 2.38  | 0.13 | 3.46  | 27  | 0.81  | 0.11 | 0.63  |
| 5   | -0.58 | 0.13 | 0.59  | 28  | -0.55 | 0.13 | -0.45 |
| 6   | -0.44 | 0.13 | 1.32  | 29  | -1.29 | 0.16 | -0.48 |
| 7   | 0.77  | 0.11 | 0.08  | 30  | -0.48 | 0.13 | -0.36 |
| 8   | -1.22 | 0.15 | -0.62 | 31  | -0.65 | 0.13 | -1.12 |
| 9   | -0.98 | 0.14 | -0.32 | 32  | -0.14 | 0.12 | -1.17 |
| 10  | -0.90 | 0.14 | -0.50 | 33  | 0.57  | 0.11 | -1.21 |
| 11  | -0.14 | 0.12 | -0.12 | 34  | 0.14  | 0.11 | -0.60 |
| 12  | 0.35  | 0.11 | -0.65 | 35  | 0.51  | 0.11 | -1.26 |
| 13  | 0.39  | 0.11 | -1.55 | 36  | 0.15  | 0.11 | -0.34 |
| 14  | -0.33 | 0.12 | -1.87 | 37  | 0.49  | 0.11 | 0.17  |
| 15  | -0.58 | 0.13 | -1.14 | 38  | -0.56 | 0.13 | -0.05 |
| 16  | -0.23 | 0.12 | 0.36  | 39  | 0.52  | 0.11 | -0.67 |
| 17  | -0.24 | 0.12 | -0.73 | 40  | 1.81  | 0.11 | 0.80  |
| 18  | -0.84 | 0.14 | -0.84 | 41  | 0.39  | 0.11 | -1.36 |
| 19  | -1.00 | 0.14 | -0.69 | 42  | 2.62  | 0.14 | 2.17  |
| 20  | -1.42 | 0.16 | 0.07  | 43  | 0.12  | 0.11 | 0.67  |
| 21  | -0.06 | 0.12 | -0.64 | 44  | 2.08  | 0.12 | 1.73  |
| 22  | -1.42 | 0.16 | -0.54 | 45  | 0.82  | 0.11 | 0.60  |
| 23  | 0.41  | 0.11 | -1.29 |     |       |      |       |

ตาราง 10 ผลการประมาณค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าความคลาดเคลื่อน  
มาตรฐานในการประมาณค่าความยากของข้อสอบ (SEE) และค่าความสอดคล้องกับไมเคิล  
(t - test) ของแบบทดสอบฉบับที่ 4

| ข้อ | b     | SEE  | t     | ข้อ | b     | SEE  | t     |
|-----|-------|------|-------|-----|-------|------|-------|
| 1   | -1.70 | 0.16 | -0.65 | 24  | 0.61  | 0.11 | -0.70 |
| 2   | -1.72 | 0.16 | -2.13 | 25  | 0.19  | 0.11 | -1.00 |
| 3   | 0.83  | 0.11 | -0.28 | 26  | 0.76  | 0.11 | -0.54 |
| 4   | -2.81 | 0.24 | -0.78 | 27  | 0.62  | 0.11 | 0.60  |
| 5   | -1.17 | 0.14 | -0.78 | 28  | -0.78 | 0.13 | -1.58 |
| 6   | 2.77  | 0.16 | 1.60  | 29  | -0.63 | 0.12 | -2.02 |
| 7   | -0.78 | 0.13 | -0.08 | 30  | -0.75 | 0.12 | -1.85 |
| 8   | -0.96 | 0.13 | -2.20 | 31  | -1.30 | 0.14 | -1.41 |
| 9   | 2.01  | 0.13 | 0.35  | 32  | 0.19  | 0.11 | -0.82 |
| 10  | -1.49 | 0.15 | -1.88 | 33  | -0.54 | 0.12 | -1.10 |
| 11  | 1.25  | 0.11 | 5.05  | 34  | 0.08  | 0.11 | -1.10 |
| 12  | -1.70 | 0.16 | -1.62 | 35  | -0.46 | 0.12 | -1.50 |
| 13  | -0.88 | 0.13 | 0.40  | 36  | 0.25  | 0.11 | 2.40  |
| 14  | -0.66 | 0.12 | -0.58 | 37  | -0.09 | 0.11 | -1.48 |
| 15  | -0.28 | 0.12 | 0.71  | 38  | 0.26  | 0.11 | 0.68  |
| 16  | 1.84  | 0.12 | 1.75  | 39  | 0.12  | 0.11 | 0.27  |
| 17  | -0.27 | 0.12 | -1.21 | 40  | 1.23  | 0.11 | 0.16  |
| 18  | -0.81 | 0.13 | -1.12 | 41  | 1.05  | 0.11 | 0.52  |
| 19  | -0.91 | 0.13 | -1.94 | 42  | 0.85  | 0.11 | 2.71  |
| 20  | -0.34 | 0.12 | -1.11 | 43  | 0.49  | 0.11 | 0.80  |
| 21  | 0.60  | 0.11 | 0.31  | 44  | 0.17  | 0.11 | 0.09  |
| 22  | 1.07  | 0.11 | -0.96 | 45  | 2.34  | 0.14 | 0.97  |
| 23  | 1.45  | 0.12 | 0.04  |     |       |      |       |

ตาราง 11 ผลการประมาณค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าความคลาดเคลื่อน  
มาตรฐานในการประมาณค่าความยากของข้อสอบ (SEE) และค่าความสอดคล้องกับโมเดล  
(t - test) ของแบบทดสอบฉบับที่ 5

| ข้อ | b     | SEE  | t     | ข้อ | b     | SEE  | t     |
|-----|-------|------|-------|-----|-------|------|-------|
| 1   | 0.82  | 0.12 | 4.20  | 24  | -0.06 | 0.11 | 0.81  |
| 2   | -0.53 | 0.11 | -2.35 | 25  | 0.19  | 0.11 | -0.56 |
| 3   | 0.63  | 0.12 | -1.82 | 26  | 2.21  | 0.17 | 0.80  |
| 4   | 0.82  | 0.12 | 0.89  | 27  | 1.60  | 0.14 | 0.62  |
| 5   | -1.29 | 0.12 | -0.37 | 28  | -2.01 | 0.14 | -0.56 |
| 6   | -0.48 | 0.11 | 0.86  | 29  | -1.60 | 0.12 | -0.51 |
| 7   | 0.71  | 0.12 | -1.67 | 30  | -0.28 | 0.11 | -1.31 |
| 8   | -1.97 | 0.13 | -0.56 | 31  | 0.26  | 0.11 | 0.82  |
| 9   | -0.80 | 0.11 | 1.64  | 32  | -0.89 | 0.11 | 0.12  |
| 10  | 1.20  | 0.13 | -1.48 | 33  | -1.15 | 0.11 | -0.64 |
| 11  | -1.43 | 0.12 | -1.78 | 34  | 0.28  | 0.11 | -0.29 |
| 12  | -0.99 | 0.11 | -1.82 | 35  | -0.46 | 0.11 | -1.67 |
| 13  | 1.86  | 0.16 | 1.65  | 36  | -0.01 | 0.11 | -0.30 |
| 14  | 1.42  | 0.14 | 1.81  | 37  | -0.66 | 0.11 | 0.29  |
| 15  | -0.27 | 0.11 | -1.65 | 38  | -0.69 | 0.11 | -0.20 |
| 16  | -0.64 | 0.11 | -1.10 | 39  | 0.02  | 0.11 | -1.62 |
| 17  | 0.54  | 0.11 | 0.88  | 40  | 1.88  | 0.16 | 1.07  |
| 18  | 0.11  | 0.11 | -1.35 | 41  | 0.40  | 0.11 | 2.81  |
| 19  | -0.50 | 0.11 | -1.62 | 42  | 0.31  | 0.11 | -0.91 |
| 20  | -0.35 | 0.11 | 0.45  | 43  | 0.25  | 0.11 | -2.09 |
| 21  | 1.25  | 0.13 | -0.08 | 44  | 0.36  | 0.11 | 2.34  |
| 22  | -1.18 | 0.11 | -1.80 | 45  | 0.57  | 0.12 | 2.13  |
| 23  | 0.55  | 0.11 | -0.18 |     |       |      |       |

ตาราง 12 ผลการประมาณค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าความคลาดเคลื่อน  
มาตรฐานในการประมาณค่าความยากของข้อสอบ (SEE) และค่าความสอดคล้องกับโมเดล  
(t - test) ของแบบทดสอบฉบับที่ 6

| ข้อ | b     | SEE  | t     | ข้อ | b     | SEE  | t     |
|-----|-------|------|-------|-----|-------|------|-------|
| 1   | 0.66  | 0.11 | -0.88 | 24  | -0.38 | 0.11 | 0.07  |
| 2   | -0.43 | 0.11 | -1.60 | 25  | 0.06  | 0.11 | -0.87 |
| 3   | -1.40 | 0.12 | -0.16 | 26  | 2.27  | 0.17 | 1.13  |
| 4   | -0.89 | 0.11 | 0.39  | 27  | -0.96 | 0.11 | -1.72 |
| 5   | -0.94 | 0.11 | -0.13 | 28  | -0.31 | 0.11 | -0.43 |
| 6   | 1.20  | 0.12 | -1.17 | 29  | -0.25 | 0.11 | 0.86  |
| 7   | 1.08  | 0.12 | -1.31 | 30  | 0.66  | 0.11 | 0.28  |
| 8   | -1.03 | 0.12 | -0.92 | 31  | -0.56 | 0.11 | -0.82 |
| 9   | -1.22 | 0.12 | -0.83 | 32  | -0.83 | 0.11 | -0.90 |
| 10  | 0.50  | 0.11 | 0.37  | 33  | 0.48  | 0.11 | -0.25 |
| 11  | 1.42  | 0.13 | -0.25 | 34  | 0.40  | 0.11 | -0.63 |
| 12  | 0.62  | 0.11 | 0.91  | 35  | -0.07 | 0.11 | 1.23  |
| 13  | 1.01  | 0.12 | 1.35  | 36  | -0.09 | 0.11 | 1.50  |
| 14  | -0.80 | 0.11 | -1.99 | 37  | -0.47 | 0.11 | -0.60 |
| 15  | -0.66 | 0.11 | 0.96  | 38  | -1.29 | 0.12 | -0.57 |
| 16  | -0.30 | 0.11 | 0.68  | 39  | -0.84 | 0.11 | -1.57 |
| 17  | 0.27  | 0.11 | -1.56 | 40  | 0.05  | 0.11 | 0.83  |
| 18  | -0.99 | 0.11 | -0.04 | 41  | -0.11 | 0.11 | -0.72 |
| 19  | -0.07 | 0.11 | -0.15 | 42  | 0.27  | 0.11 | 0.27  |
| 20  | -1.36 | 0.12 | -0.86 | 43  | 0.93  | 0.12 | 0.21  |
| 21  | 0.40  | 0.11 | 1.80  | 44  | 1.37  | 0.13 | 0.28  |
| 22  | 0.61  | 0.11 | -1.28 | 45  | 0.86  | 0.12 | -0.27 |
| 23  | 0.69  | 0.11 | -1.76 |     |       |      |       |

**ภาคผนวก ค**

**เกรตเจลี่ยสะสมและค่าประมาณพารามิเตอร์ความสามารถของผู้สอบ  
ระหว่างแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบเพิลิกซ์เลเวล**

ตาราง 13 เกณฑ์เฉลี่ยสะสม (G.P.A.) กับผลการประมาณค่าพารามิเตอร์ความสามารถของ  
ผู้สอบ (θ) ระหว่างแบบทดสอบดั้งเดิม (CONT) แบบทดสอบสองขั้นตอน (TWOT) และ  
แบบทดสอบเพื่อกำหนดเวลา (FLEXT)

| คนที่ | G.P.A. | θ     |       |       | คนที่ | G.P.A. | θ     |       |       |
|-------|--------|-------|-------|-------|-------|--------|-------|-------|-------|
|       |        | CONT  | TWOT  | FLEXT |       |        | CONT  | TWOT  | FLEXT |
| 1     | 3.50   | 1.47  | 2.09  | 1.19  | 20    | 3.25   | 1.98  | 1.34  | 1.76  |
| 2     | 1.25   | -0.91 | -0.43 | 0.16  | 21    | 2.50   | 0.28  | 0.27  | -0.67 |
| 3     | 2.75   | 1.08  | 1.23  | 0.54  | 22    | 3.00   | 0.14  | 1.23  | 2.10  |
| 4     | 2.00   | -0.01 | 0.95  | 1.19  | 23    | 3.25   | 1.71  | 1.74  | 1.19  |
| 5     | 3.00   | -0.75 | 1.34  | 0.96  | 24    | 2.25   | -1.08 | 0.41  | 0.16  |
| 6     | 3.50   | 1.47  | 2.77  | 1.46  | 25    | 3.50   | 3.47  | 2.09  | 2.10  |
| 7     | 3.75   | 3.47  | 2.09  | 2.10  | 26    | 3.75   | 1.27  | 2.33  | 0.96  |
| 8     | 3.00   | 2.75  | 1.56  | 2.10  | 27    | 2.75   | 1.47  | 1.14  | 0.96  |
| 9     | 3.25   | 1.98  | 1.80  | 1.46  | 28    | 2.50   | 0.91  | 1.05  | 0.16  |
| 10    | 2.25   | -0.59 | 0.95  | 0.16  | 29    | 1.25   | -0.29 | -0.57 | 1.19  |
| 11    | 2.75   | 1.71  | 1.34  | 1.76  | 30    | 2.50   | -0.15 | 0.95  | 0.54  |
| 12    | 2.00   | 0.28  | 1.14  | 1.76  | 31    | 2.75   | -0.15 | 1.74  | 0.96  |
| 13    | 3.25   | 1.27  | 2.33  | 0.96  | 32    | 3.00   | -1.08 | 1.74  | 1.46  |
| 14    | 1.50   | -0.15 | -0.43 | 0.96  | 33    | 2.50   | 1.27  | 1.31  | 0.54  |
| 15    | 2.75   | 0.58  | 1.74  | 0.96  | 34    | 3.50   | 1.71  | 2.33  | 1.46  |
| 16    | 2.25   | -0.01 | 0.41  | -0.84 | 35    | 1.75   | 1.27  | -0.27 | 0.96  |
| 17    | 2.75   | 1.08  | 0.95  | 2.10  | 36    | 1.50   | -0.59 | 1.56  | 2.10  |
| 18    | 1.50   | -0.01 | -0.57 | 1.46  | 37    | 3.00   | 2.75  | 1.14  | 0.16  |
| 19    | 2.00   | -0.75 | 0.42  | 1.46  | 38    | 2.75   | -1.08 | 0.72  | 1.76  |

ตาราง 13 (ต่อ)

| คนที | G.P.A. | θ     |       |       | คนที | G.P.A. | θ     |       |       |
|------|--------|-------|-------|-------|------|--------|-------|-------|-------|
|      |        | CONT  | TWOT  | FLEXT |      |        | CONT  | TWOT  | FLEXT |
| 39   | 3.00   | 1.08  | 1.31  | 0.16  | 60   | 3.50   | 1.27  | 2.33  | 2.10  |
| 40   | 2.50   | -0.91 | 0.71  | 0.16  | 61   | 1.75   | -1.08 | -0.88 | -0.67 |
| 41   | 1.25   | 0.58  | -0.15 | 2.10  | 62   | 2.00   | -0.75 | 0.95  | 1.19  |
| 42   | 2.00   | 0.74  | 0.44  | 2.10  | 63   | 2.25   | -0.01 | 0.27  | -0.67 |
| 43   | 1.00   | -1.08 | -0.88 | -0.18 | 64   | 2.75   | -1.08 | 0.79  | 1.76  |
| 44   | 3.25   | 2.75  | 2.09  | 2.10  | 65   | 3.00   | 2.75  | 1.56  | 2.10  |
| 45   | 2.25   | 0.28  | 0.27  | 2.10  | 66   | 3.25   | 1.08  | 1.66  | 1.76  |
| 46   | 1.50   | -0.01 | -0.57 | 0.96  | 67   | 2.25   | 0.91  | 0.61  | 0.74  |
| 47   | 2.75   | -0.59 | 1.13  | 0.16  | 68   | 3.25   | 1.27  | 1.23  | 0.74  |
| 48   | 1.75   | -0.15 | -0.58 | 0.54  | 69   | 3.75   | 1.98  | 2.33  | 1.76  |
| 49   | 2.25   | -1.08 | 1.31  | 2.10  | 70   | 2.50   | -0.15 | 0.72  | 1.19  |
| 50   | 1.25   | 0.28  | 1.04  | -0.67 | 71   | 2.75   | 1.98  | 0.95  | 1.76  |
| 51   | 2.75   | 1.98  | 1.74  | 1.46  | 72   | 3.25   | 1.71  | 1.34  | 1.19  |
| 52   | 3.50   | 3.47  | 2.33  | 2.10  | 73   | 3.25   | 0.28  | 1.05  | 0.96  |
| 53   | 2.00   | 1.27  | 0.72  | 0.16  | 74   | 3.00   | -0.59 | 1.31  | -1.01 |
| 54   | 2.50   | 1.71  | 1.23  | 0.96  | 75   | 2.75   | 0.91  | 1.51  | 0.54  |
| 55   | 3.75   | 2.75  | 2.33  | 2.50  | 76   | 1.50   | -1.27 | -0.43 | -1.81 |
| 56   | 2.50   | 1.98  | 0.61  | 1.19  | 77   | 2.75   | -0.15 | 1.56  | 2.10  |
| 57   | 3.00   | 2.31  | 1.74  | 1.76  | 78   | 3.00   | -0.91 | 1.51  | 0.54  |
| 58   | 2.50   | -0.91 | 0.88  | 1.19  | 79   | 2.50   | -0.59 | 0.71  | 1.76  |
| 59   | 2.75   | -1.08 | 1.34  | 1.19  | 80   | 2.75   | 1.08  | 1.51  | 0.54  |

ตาราง 13 (ต่อ)

| คนที่ | G.P.A. | ๑     |       |       | คนที่ | G.P.A. | ๑     |       |       |
|-------|--------|-------|-------|-------|-------|--------|-------|-------|-------|
|       |        | CONT  | TWOT  | FLEXT |       |        | CONT  | TWOT  | FLEXT |
| 81    | 1.75   | 1.47  | -0.57 | 0.54  | 102   | 2.25   | 1.71  | 0.44  | 1.76  |
| 82    | 2.50   | 2.31  | 2.33  | 1.76  | 103   | 2.75   | -0.01 | 0.78  | 0.74  |
| 83    | 2.25   | 0.74  | 0.78  | 0.54  | 104   | 3.75   | 0.91  | 2.33  | 2.10  |
| 84    | 2.00   | 1.71  | 0.61  | 1.46  | 105   | 2.25   | 2.31  | 0.44  | 2.10  |
| 85    | 3.00   | 1.08  | 1.74  | 0.74  | 106   | 2.75   | 1.27  | 1.74  | 0.96  |
| 86    | 2.50   | 1.98  | 0.78  | 1.76  | 107   | 2.25   | -0.59 | 0.44  | -0.67 |
| 87    | 2.50   | 1.47  | 0.95  | 0.35  | 108   | 3.00   | 0.74  | 1.56  | 1.46  |
| 88    | 2.25   | 0.91  | 0.78  | 1.46  | 109   | 2.00   | 1.98  | 0.44  | 1.76  |
| 89    | 3.00   | 1.08  | 1.34  | 0.96  | 110   | 3.25   | -0.01 | 1.56  | 1.46  |
| 90    | 3.50   | 1.98  | 2.33  | 1.76  | 111   | 2.75   | 0.74  | 0.44  | 1.46  |
| 91    | 3.25   | 1.27  | 1.34  | 0.96  | 112   | 3.50   | 3.47  | 2.33  | 2.10  |
| 92    | 2.75   | 0.28  | 1.31  | 0.74  | 113   | 2.75   | 1.27  | 0.78  | -0.51 |
| 93    | 2.50   | 1.47  | 0.61  | -1.01 | 114   | 3.25   | 1.71  | 1.74  | 0.35  |
| 94    | 2.75   | -1.27 | 1.56  | 1.76  | 115   | 2.25   | 0.53  | 0.78  | 2.10  |
| 95    | 3.50   | 1.71  | 2.09  | 1.76  | 116   | 3.25   | 1.08  | 2.25  | 2.10  |
| 96    | 2.25   | -1.08 | 0.78  | 2.10  | 117   | 2.50   | -0.01 | 0.44  | -0.67 |
| 97    | 3.00   | 0.58  | 1.51  | 0.96  | 118   | 2.75   | -0.59 | 0.88  | -1.01 |
| 98    | 2.50   | 1.98  | 0.61  | 1.46  | 119   | 3.00   | 0.91  | 1.93  | 1.76  |
| 99    | 2.75   | 2.75  | 1.56  | 2.99  | 120   | 3.50   | 2.31  | 2.33  | 2.50  |
| 100   | 2.50   | -0.91 | 0.44  | 1.46  | 121   | 3.25   | -0.59 | -0.43 | -0.01 |
| 101   | 2.75   | 3.47  | 1.14  | 2.10  | 122   | 2.75   | -0.01 | 1.93  | 2.10  |

ตาราง 13 (ต่อ)

| คนที่ | G.P.A. | ๑     |       |       | คนที่ | G.P.A. | ๑     |       |       |
|-------|--------|-------|-------|-------|-------|--------|-------|-------|-------|
|       |        | CONT  | TWOT  | FLEXT |       |        | CONT  | TWOT  | FLEXT |
| 123   | 2.50   | 0.43  | 1.14  | 1.76  | 144   | 2.25   | 0.58  | 0.78  | 0.96  |
| 124   | 2.75   | -0.75 | 1.66  | 0.96  | 145   | 2.00   | -0.59 | 0.44  | 0.96  |
| 125   | 1.50   | 1.27  | 0.41  | 1.19  | 146   | 1.25   | -0.75 | -0.14 | 0.74  |
| 126   | 1.75   | 1.08  | 0.44  | 1.76  | 147   | 2.50   | 0.28  | 1.13  | -0.18 |
| 127   | 2.75   | -0.01 | 0.28  | 1.19  | 148   | 2.75   | -1.27 | 0.72  | 0.35  |
| 128   | 2.00   | 0.74  | 0.78  | 2.10  | 149   | 2.00   | -0.01 | 0.44  | -1.76 |
| 129   | 3.00   | 0.43  | 1.43  | 1.76  | 150   | 2.50   | 0.74  | 1.43  | 1.46  |
| 130   | 3.75   | 0.91  | 1.80  | 2.10  | 151   | 1.75   | -1.27 | 0.29  | 0.96  |
| 131   | 3.50   | -0.59 | 1.56  | 1.19  | 152   | 2.50   | 1.47  | 1.31  | 0.54  |
| 132   | 3.00   | -0.01 | 2.33  | 2.50  | 153   | 2.00   | 0.14  | 0.29  | 0.74  |
| 133   | 2.75   | 0.58  | 1.93  | 0.84  | 154   | 1.75   | 0.14  | 0.27  | -0.34 |
| 134   | 3.25   | 2.75  | 2.25  | 2.10  | 155   | 2.50   | 0.28  | 0.42  | 1.46  |
| 135   | 3.75   | -0.44 | 2.33  | 2.50  | 156   | 2.75   | 1.08  | 1.05  | 0.16  |
| 136   | 2.50   | 0.74  | 1.43  | 1.46  | 157   | 2.75   | 0.91  | 1.51  | 0.16  |
| 137   | 3.50   | 1.08  | 2.01  | 2.10  | 158   | 2.00   | 0.43  | 0.28  | -0.18 |
| 138   | 2.00   | 0.43  | 0.14  | 1.19  | 159   | 3.00   | 0.58  | 1.74  | 0.74  |
| 139   | 2.75   | -0.15 | 1.66  | 1.19  | 160   | 2.25   | 1.27  | 0.44  | 1.76  |
| 140   | 2.50   | -0.15 | 1.51  | 0.16  | 161   | 1.00   | -0.15 | -1.05 | 0.54  |
| 141   | 1.75   | 1.08  | 1.05  | 0.16  | 162   | 2.75   | 1.71  | 1.14  | 1.76  |
| 142   | 1.25   | -1.08 | -0.43 | -0.51 | 163   | 2.50   | 0.14  | 1.13  | -0.18 |
| 143   | 1.00   | -0.91 | -0.39 | -1.01 | 164   | 3.00   | 0.91  | 1.31  | 0.35  |

ตาราง 13 (ต่อ)

| คนที่ | G.P.A. | ๑     |       |       | คนที่ | G.P.A. | ๑     |       |       |
|-------|--------|-------|-------|-------|-------|--------|-------|-------|-------|
|       |        | CONT  | TWOT  | FLEXT |       |        | CONT  | TWOT  | FLEXT |
| 165   | 1.50   | -1.08 | -0.29 | -0.67 | 186   | 3.25   | 0.43  | 1.56  | 2.10  |
| 166   | 1.75   | 0.43  | -0.58 | -0.18 | 187   | 1.75   | 0.23  | 0.28  | 1.19  |
| 167   | 1.25   | -1.27 | -0.57 | 0.54  | 188   | 2.25   | 1.27  | 1.04  | 0.74  |
| 168   | 2.25   | 1.47  | 1.23  | 0.74  | 189   | 1.50   | 1.08  | 0.29  | 1.76  |
| 169   | 3.00   | -0.15 | 0.72  | -0.67 | 190   | 2.75   | 2.75  | 1.05  | 2.10  |
| 170   | 1.00   | 0.74  | -0.14 | 1.19  | 191   | 3.50   | 0.74  | 1.66  | 1.76  |
| 171   | 1.25   | 0.58  | 0.29  | 1.19  | 192   | 2.50   | 1.71  | 1.41  | 1.46  |
| 172   | 2.00   | -0.75 | 0.29  | 1.76  | 193   | 3.50   | 0.28  | 1.43  | 0.96  |
| 173   | 2.50   | 0.91  | 0.95  | 0.35  | 194   | 3.00   | 2.31  | 2.01  | 2.10  |
| 174   | 1.00   | -0.01 | -0.14 | 0.74  | 195   | 1.75   | 0.74  | 0.27  | 0.35  |
| 175   | 1.25   | 0.14  | -0.15 | -0.01 | 196   | 2.25   | -1.08 | 0.72  | 0.96  |
| 176   | 1.00   | 0.58  | -0.43 | -0.35 | 197   | 3.75   | 1.08  | 2.69  | 2.50  |
| 177   | 1.50   | 1.08  | 0.88  | 0.54  | 198   | 3.00   | -0.91 | -0.72 | 2.10  |
| 178   | 1.75   | -0.91 | -0.43 | 1.76  | 199   | 1.75   | 0.43  | 0.29  | 0.96  |
| 179   | 1.00   | -1.08 | -1.74 | 1.76  | 200   | 2.50   | 0.91  | 1.43  | 0.35  |
| 180   | 2.25   | -0.15 | 0.27  | -1.01 | 201   | 3.50   | -0.59 | 1.80  | 1.76  |
| 181   | 1.50   | 0.91  | 0.58  | -0.01 | 202   | 1.75   | 0.91  | 0.13  | 2.10  |
| 182   | 2.50   | 0.28  | 0.29  | 0.74  | 203   | 1.00   | 1.08  | -0.15 | 0.54  |
| 183   | 2.75   | 1.71  | 1.51  | 0.35  | 204   | 2.25   | -0.59 | 0.42  | 0.54  |
| 184   | 3.00   | 0.28  | 1.56  | 2.10  | 205   | 2.50   | -0.75 | 2.01  | 1.76  |
| 185   | 2.50   | 2.75  | 0.44  | 2.10  | 206   | 2.75   | 1.47  | 0.95  | 0.54  |

ตาราง 13 (ต่อ)

| คนที่ | G.P.A. | ๑     |       |       | คนที่ | G.P.A. | ๑     |       |       |
|-------|--------|-------|-------|-------|-------|--------|-------|-------|-------|
|       |        | CONT  | TWOT  | FLEXT |       |        | CONT  | TWOT  | FLEXT |
| 207   | 1.50   | 0.43  | 0.13  | 1.19  | 228   | 1.25   | 1.27  | -0.64 | 0.96  |
| 208   | 1.75   | 1.08  | 0.88  | 0.74  | 229   | 2.25   | 0.91  | 0.27  | 0.54  |
| 209   | 1.00   | -1.07 | -0.02 | 1.76  | 230   | 1.75   | -0.15 | -0.24 | -0.01 |
| 210   | 2.75   | 1.27  | 1.51  | 0.54  | 231   | 3.50   | -0.01 | 2.69  | 2.10  |
| 211   | 1.00   | 0.14  | -0.02 | 0.74  | 232   | 1.50   | -0.29 | -0.43 | 0.96  |
| 212   | 1.50   | 0.58  | 0.41  | 0.16  | 233   | 2.25   | 0.14  | 0.56  | -0.18 |
| 213   | 1.25   | -0.75 | -0.57 | 0.96  | 234   | 1.75   | -0.44 | -0.29 | -0.18 |
| 214   | 1.75   | 0.14  | 0.13  | 1.19  | 235   | 1.50   | 0.43  | 0.28  | 1.19  |
| 215   | 1.25   | -0.15 | -0.43 | 0.54  | 236   | 2.25   | -0.59 | 1.13  | 0.74  |
| 216   | 2.00   | 0.43  | 1.04  | 0.54  | 237   | 2.50   | 0.28  | 0.42  | -0.01 |
| 217   | 1.00   | -0.01 | -0.02 | 1.19  | 238   | 2.75   | 0.58  | 0.79  | -0.18 |
| 218   | 2.50   | -0.91 | 1.13  | -0.01 | 239   | 2.00   | -0.15 | 0.00  | -0.01 |
| 219   | 1.25   | 0.74  | 0.88  | -0.01 | 240   | 3.50   | -0.59 | 2.01  | 1.76  |
| 220   | 1.00   | -1.28 | -0.02 | 0.96  | 241   | 1.75   | 1.08  | 0.42  | 1.76  |
| 221   | 1.75   | -0.15 | 0.71  | -0.67 | 242   | 2.25   | -0.44 | 0.72  | 0.96  |
| 222   | 2.50   | 0.58  | 1.04  | -0.18 | 243   | 2.75   | 0.43  | 0.64  | -0.34 |
| 223   | 2.00   | -0.91 | 0.42  | -0.01 | 244   | 3.25   | -0.01 | 1.23  | 0.54  |
| 224   | 1.00   | 1.08  | -0.02 | 0.54  | 245   | 3.50   | 0.74  | 1.56  | 1.46  |
| 225   | 2.25   | -0.91 | 0.27  | -0.34 | 246   | 2.25   | 0.28  | 0.56  | -0.34 |
| 226   | 1.50   | 0.74  | -0.02 | 0.96  | 247   | 3.50   | 1.08  | 2.43  | 2.50  |
| 227   | 2.75   | 0.91  | 1.51  | 0.74  | 248   | 2.00   | -0.59 | 0.41  | 0.54  |

ตาราง 13 (ต่อ)

| คนที่ | G.P.A. | ๑     |       |       | คนที่ | G.P.A. | ๑     |       |       |
|-------|--------|-------|-------|-------|-------|--------|-------|-------|-------|
|       |        | CONT  | TWOT  | FLEXT |       |        | CONT  | TWOT  | FLEXT |
| 249   | 1.75   | -0.29 | 0.05  | -1.01 | 270   | 2.75   | -0.75 | 0.42  | 0.74  |
| 250   | 2.50   | 0.28  | 0.27  | -0.51 | 271   | 2.00   | 1.08  | 0.29  | 1.76  |
| 251   | 2.75   | 0.58  | 0.79  | -0.01 | 272   | 1.75   | -0.29 | 0.13  | -0.18 |
| 252   | 2.50   | -1.27 | 0.64  | -0.67 | 273   | 2.25   | -0.01 | 0.56  | 0.35  |
| 253   | 2.75   | 0.91  | 0.87  | 0.54  | 274   | 3.75   | 0.74  | 2.01  | 2.10  |
| 254   | 1.50   | -0.44 | -1.66 | 0.54  | 275   | 2.25   | -0.59 | 0.13  | 0.35  |
| 255   | 1.25   | 0.43  | -0.24 | -0.84 | 276   | 1.25   | 0.43  | -0.81 | 1.46  |
| 256   | 2.50   | -0.15 | 0.42  | 0.16  | 277   | 2.25   | -0.75 | 0.42  | 0.96  |
| 257   | 1.25   | 0.14  | 0.27  | -0.51 | 278   | 3.50   | 2.31  | 2.01  | 2.10  |
| 258   | 1.50   | -0.15 | -0.24 | -0.67 | 279   | 3.00   | -0.75 | -0.43 | 1.46  |
| 259   | 1.75   | -0.59 | 0.64  | -0.01 | 280   | 2.75   | -0.15 | 0.41  | 0.96  |
| 260   | 2.00   | 0.58  | 0.13  | -0.51 | 281   | 2.00   | 0.74  | -0.02 | 2.10  |
| 261   | 2.50   | -0.44 | 0.28  | 0.35  | 282   | 1.25   | -0.91 | -0.33 | 1.46  |
| 262   | 2.75   | -1.08 | 0.56  | -0.34 | 283   | 1.50   | 0.28  | -0.14 | 1.46  |
| 263   | 2.00   | -0.01 | 0.28  | -0.01 | 284   | 1.50   | -1.97 | -0.33 | -0.84 |
| 264   | 1.75   | -0.15 | 0.41  | 0.16  | 285   | 1.75   | 0.58  | -0.28 | -1.01 |
| 265   | 2.25   | -0.59 | -0.01 | 0.35  | 286   | 2.75   | 0.43  | 1.04  | 0.54  |
| 266   | 1.00   | 0.28  | -0.01 | -0.51 | 287   | 2.25   | 0.74  | 0.56  | 0.35  |
| 267   | 1.50   | -0.44 | -0.71 | -1.18 | 288   | 1.25   | -0.75 | -0.57 | 1.19  |
| 268   | 2.00   | 0.14  | 0.64  | -0.84 | 289   | 4.00   | 2.75  | 2.25  | 2.50  |
| 269   | 2.50   | 0.58  | 0.13  | -0.18 | 290   | 1.50   | -1.08 | -0.81 | -0.34 |

ตาราง 13 (ต่อ)

| คนที่ | G.P.A | θ     |       |       | คนที่ | G.P.A | θ     |       |       |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|       |       | CONT  | TWOT  | FLEXT |       |       | CONT  | TWOT  | FLEXT |
| 291   | 3.75  | -0.91 | -0.15 | 0.16  | 312   | 1.75  | 0.43  | -0.81 | 1.46  |
| 292   | 2.75  | -0.15 | 0.44  | -0.51 | 313   | 1.50  | -0.44 | -0.33 | 0.74  |
| 293   | 3.50  | 1.08  | 1.23  | 0.74  | 314   | 1.50  | -0.44 | -0.14 | 0.74  |
| 294   | 3.75  | 0.58  | 2.01  | 2.10  | 315   | 1.75  | -0.75 | -0.43 | 0.96  |
| 295   | 1.75  | 0.91  | -0.88 | -0.18 | 316   | 1.00  | -0.15 | -0.18 | 1.46  |
| 296   | 2.50  | -0.75 | -0.24 | 0.54  | 317   | 1.25  | -0.44 | -0.64 | 1.46  |
| 297   | 1.50  | -0.44 | -0.28 | 1.46  | 318   | 2.75  | 0.58  | 1.13  | 0.35  |
| 298   | 2.75  | 0.14  | 1.80  | -0.18 | 319   | 2.50  | 0.14  | 0.28  | 0.96  |
| 299   | 2.50  | 0.43  | 0.56  | 1.76  | 320   | 1.75  | -0.91 | -0.64 | 1.19  |
| 300   | 1.50  | -0.29 | 0.55  | -1.01 | 321   | 1.50  | -0.29 | -1.34 | -1.01 |
| 301   | 2.00  | -0.59 | 0.05  | 0.16  | 322   | 3.50  | -1.08 | 0.79  | -0.18 |
| 302   | 2.25  | 0.74  | 1.56  | 1.19  | 323   | 1.50  | 0.74  | -0.14 | 0.96  |
| 303   | 1.50  | -1.70 | -0.24 | -1.01 | 324   | 1.75  | 1.47  | -0.72 | 0.96  |
| 304   | 1.75  | -0.29 | -0.01 | 0.35  | 325   | 3.25  | 0.58  | 1.13  | 0.74  |
| 305   | 3.00  | -1.27 | -0.73 | 0.74  | 326   | 1.50  | -0.75 | -0.15 | -0.67 |
| 306   | 3.25  | 0.58  | 0.71  | 1.46  | 327   | 1.75  | -0.44 | -0.18 | 0.96  |
| 307   | 2.25  | -0.44 | 0.43  | 0.74  | 328   | 2.00  | -0.91 | -0.43 | 1.19  |
| 308   | 2.00  | 0.74  | 0.44  | 1.19  | 329   | 3.25  | 0.91  | 1.23  | -0.01 |
| 309   | 1.50  | -0.15 | -1.34 | 1.19  | 330   | 3.00  | -0.15 | 0.87  | 0.74  |
| 310   | 3.25  | 0.14  | 0.87  | 1.54  | 331   | 1.50  | 0.43  | -0.49 | 1.46  |
| 311   | 1.25  | -0.44 | -1.15 | 0.34  | 332   | 1.50  | -0.01 | -0.43 | 0.96  |

ตาราง 13 (ต่อ)

| คนที่ | G.P.A | θ     |       |       | คนที่ | G.P.A | θ     |       |       |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|       |       | CONT  | TWOT  | FLEXT |       |       | CONT  | TWOT  | FLEXT |
| 333   | 1.25  | -0.44 | -0.18 | 1.19  | 354   | 1.75  | -0.15 | 0.64  | 0.16  |
| 334   | 2.75  | -0.15 | 1.04  | 0.74  | 355   | 1.50  | 0.28  | -0.01 | 1.19  |
| 335   | 2.00  | 0.58  | -0.64 | 1.46  | 356   | 1.15  | -0.29 | -0.29 | 1.19  |
| 336   | 2.25  | -0.29 | -0.14 | 0.74  | 357   | 2.50  | -0.01 | 0.34  | 1.19  |
| 337   | 3.25  | 0.14  | 0.27  | 0.54  | 358   | 1.25  | -0.29 | -0.73 | 0.35  |
| 338   | 1.75  | 0.43  | -1.05 | -0.51 | 359   | 2.50  | 0.43  | 0.49  | -0.01 |
| 339   | 2.00  | -0.75 | -0.43 | 0.14  | 360   | 2.75  | -0.15 | -0.02 | -0.01 |
| 340   | 1.50  | 0.74  | -0.18 | 1.46  | 361   | 1.50  | -0.44 | -0.55 | 0.96  |
| 341   | 2.50  | -0.15 | 0.42  | -0.01 | 362   | 2.75  | -0.01 | 0.13  | 0.16  |
| 342   | 2.25  | -0.29 | 0.28  | -0.18 | 363   | 2.00  | -0.29 | 0.19  | 1.76  |
| 343   | 2.75  | -0.44 | 0.29  | -0.67 | 364   | 2.00  | 0.28  | 0.64  | -0.01 |
| 344   | 2.00  | 0.14  | 0.44  | 0.74  | 365   | 1.75  | -0.44 | 0.34  | -0.34 |
| 345   | 2.50  | -0.29 | -0.02 | 0.74  | 366   | 1.50  | 0.74  | -0.73 | 0.35  |
| 346   | 2.25  | -0.01 | 0.19  | -0.34 | 367   | 2.50  | -0.15 | 0.57  | -0.10 |
| 347   | 2.00  | -0.29 | 0.13  | 0.16  | 368   | 2.25  | 0.43  | 0.41  | -0.01 |
| 348   | 1.75  | 0.14  | -0.49 | 0.96  | 369   | 2.50  | 0.28  | 0.72  | 1.76  |
| 349   | 2.25  | -0.29 | 0.49  | 0.96  | 370   | 2.75  | -0.15 | 0.14  | 0.35  |
| 350   | 1.50  | -0.15 | -0.88 | 0.96  | 371   | 2.75  | -0.01 | 0.28  | 0.16  |
| 351   | 1.75  | -0.15 | 0.41  | -0.18 | 372   | 2.25  | -1.47 | -0.72 | 0.16  |
| 352   | 3.50  | -0.75 | 0.13  | 1.19  | 373   | 3.00  | -0.91 | -0.43 | -0.34 |
| 353   | 1.75  | 0.14  | 0.05  | 0.54  | 374   | 1.25  | -0.23 | -0.73 | -0.18 |

ตาราง 13 (ต่อ)

| คนท. | G.P.A | ๑     |       |       | คนท. | G.P.A | ๑     |       |       |
|------|-------|-------|-------|-------|------|-------|-------|-------|-------|
|      |       | CONT  | TWOT  | FLEXT |      |       | CONT  | TWOT  | FLEXT |
| 375  | 2.00  | 0.14  | 0.41  | -0.34 | 396  | 1.25  | -0.15 | -0.02 | -0.67 |
| 376  | 1.75  | -1.70 | -0.29 | 1.19  | 397  | 1.00  | -0.75 | -1.15 | -0.84 |
| 377  | 2.25  | -0.44 | 0.49  | -0.01 | 398  | 2.00  | 0.14  | 0.14  | 0.74  |
| 378  | 2.25  | -0.01 | 0.56  | -0.01 | 399  | 2.00  | -0.15 | 0.13  | 0.96  |
| 379  | 2.00  | -1.70 | 0.19  | -0.84 | 400  | 3.25  | -1.47 | -0.88 | 0.35  |
| 380  | 1.50  | -0.23 | -1.05 | -0.18 | 401  | 3.25  | -0.75 | 0.87  | 0.35  |
| 381  | 2.50  | -0.15 | 0.13  | 0.35  | 402  | 3.50  | -0.01 | 0.14  | 0.96  |
| 382  | 2.00  | -1.47 | 0.27  | 0.54  | 403  | 3.25  | 0.74  | 0.88  | -0.01 |
| 383  | 2.25  | 0.28  | 0.57  | 0.74  | 404  | 2.25  | -0.44 | 0.41  | -0.01 |
| 384  | 1.25  | -1.28 | -0.01 | 1.76  | 405  | 3.25  | 0.28  | 2.01  | 1.46  |
| 385  | 2.50  | 0.43  | 0.34  | -0.01 | 406  | 2.75  | 0.91  | 0.88  | 0.74  |
| 386  | 2.25  | 0.74  | 0.41  | -0.18 | 407  | 3.25  | 0.14  | 0.05  | 0.35  |
| 387  | 1.35  | -0.44 | -0.39 | -1.18 | 408  | 2.50  | 0.43  | 0.64  | 0.16  |
| 388  | 2.00  | -0.15 | 0.28  | 1.19  | 409  | 1.25  | -1.47 | -0.28 | 1.19  |
| 389  | 1.00  | -0.44 | -1.05 | -0.67 | 410  | 2.50  | -0.44 | 0.19  | -0.18 |
| 390  | 3.00  | -0.23 | -0.02 | 0.54  | 411  | 2.50  | -0.15 | 0.57  | 0.35  |
| 391  | 1.25  | 0.14  | -0.73 | 0.96  | 412  | 2.50  | 0.14  | 0.41  | -0.01 |
| 392  | 1.25  | -1.70 | -0.49 | 0.96  | 413  | 1.25  | -0.29 | -0.28 | -0.01 |
| 393  | 2.75  | -0.15 | 1.13  | 0.74  | 414  | 2.75  | -0.01 | 0.34  | 1.19  |
| 394  | 1.50  | -0.75 | -0.29 | 0.16  | 415  | 1.75  | 0.14  | -0.01 | -0.51 |
| 395  | 2.50  | -0.23 | 0.41  | 0.16  | 416  | 2.25  | -0.44 | 0.41  | 1.06  |

ตาราง 13 (ต่อ)

| คนที่ | G.P.A | θ     |       |       | คนที่ | G.P.A | θ     |       |       |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|       |       | CONT  | TWOT  | FLEXT |       |       | CONT  | TWOT  | FLEXT |
| 417   | 1.75  | -0.01 | -0.28 | -0.51 | 430   | 2.25  | -0.01 | 0.49  | 1.19  |
| 418   | 2.25  | -0.29 | 0.19  | -0.51 | 431   | 2.25  | -0.15 | 0.28  | -0.18 |
| 419   | 2.50  | -0.29 | -0.14 | -0.01 | 432   | 2.75  | -0.75 | -0.43 | 0.54  |
| 420   | 2.75  | -0.15 | -0.28 | 0.54  | 433   | 2.25  | 0.43  | 0.67  | 0.74  |
| 421   | 2.00  | -0.44 | -0.01 | -0.34 | 434   | 1.75  | -0.29 | -0.15 | 0.54  |
| 422   | 1.75  | -0.29 | -0.57 | -0.01 | 435   | 3.00  | 0.14  | 0.05  | -0.34 |
| 423   | 2.00  | 0.14  | 0.05  | -0.67 | 436   | 2.25  | 0.43  | 0.14  | 1.19  |
| 424   | 2.25  | -0.15 | 0.13  | 1.46  | 437   | 2.25  | 0.58  | 0.13  | -0.51 |
| 425   | 3.50  | 0.43  | 1.80  | 1.19  | 438   | 2.75  | 0.43  | 0.87  | 0.35  |
| 426   | 3.25  | -0.29 | -0.14 | 0.74  | 439   | 2.50  | -0.29 | 0.41  | -0.18 |
| 427   | 1.50  | -0.01 | -0.73 | 0.35  | 440   | 1.75  | 0.43  | 0.05  | 0.54  |
| 428   | 2.00  | 0.14  | 0.05  | 1.19  | 441   | 1.50  | 0.14  | -0.71 | -0.51 |
| 429   | 3.00  | -0.44 | -0.01 | 0.54  |       |       |       |       |       |

## ภาคผนวก ง

ค่าฟังก์ชันอินฟอร์เมชันและ ไลน์อินฟอร์เมชันของแบบทดสอบ  
ระหว่างแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบเพิลิกซีเลเวล

ตาราง 14 ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ (TIF) ระหว่างแบบทดสอบดั้งเดิม

(CONT) แบบทดสอบสองขั้นตอน (TWOT) และแบบทดสอบแฟล็กซีเลเวล (FLEXT)

ที่ระดับความสามารถตั้งแต่ -4.00 ถึง 4.00 จำนวน 81 ระดับความสามารถ

(แบบทดสอบสองขั้นตอน ผู้วิจัยได้นำค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบวัดผล ฉบับที่ 1 :

TWOT<sub>1</sub> แบบทดสอบวัดผล ฉบับที่ 2 : TWOT<sub>2</sub> แบบทดสอบวัดผล ฉบับที่ 3 : TWOT<sub>3</sub> และแบบทดสอบวัดผล ฉบับที่ 4 : TWOT<sub>4</sub> มาคำนวณหาฟังก์ชันอินฟอร์เมชันเฉลี่ย : TWOT<sub>mean</sub>)

| ระดับ<br>ความสามารถ<br>(๑) | TIF    |                   |                   |                   |                   |                      |        |
|----------------------------|--------|-------------------|-------------------|-------------------|-------------------|----------------------|--------|
|                            | CONT   | TWOT <sub>1</sub> | TWOT <sub>2</sub> | TWOT <sub>3</sub> | TWOT <sub>4</sub> | TWOT <sub>mean</sub> | FLEXT  |
| -4.00                      | 0.1523 | 0.3628            | 0.1726            | 0.1419            | 0.6490            | 0.3334               | 0.2099 |
| -3.90                      | 0.1803 | 0.4281            | 0.2043            | 0.1765            | 0.7248            | 0.3834               | 0.2484 |
| -3.80                      | 0.2133 | 0.5084            | 0.2417            | 0.2089            | 0.8038            | 0.4398               | 0.2939 |
| -3.70                      | 0.2523 | 0.5946            | 0.2859            | 0.2471            | 0.8853            | 0.5032               | 0.3475 |
| -3.60                      | 0.2984 | 0.6997            | 0.3380            | 0.2924            | 0.9683            | 0.5746               | 0.4108 |
| -3.50                      | 0.3527 | 0.8224            | 0.3995            | 0.3457            | 1.0523            | 0.6550               | 0.4853 |
| -3.40                      | 0.4167 | 0.9652            | 0.4719            | 0.4086            | 1.1369            | 0.7457               | 0.5730 |
| -3.30                      | 0.4919 | 1.1308            | 0.5571            | 0.4828            | 1.2224            | 0.8483               | 0.6760 |
| -3.20                      | 0.5804 | 1.3224            | 0.6572            | 0.5700            | 1.3097            | 0.9648               | 0.7969 |
| -3.10                      | 0.6943 | 1.5431            | 0.7746            | 0.6725            | 1.4003            | 1.0976               | 0.9386 |
| -3.00                      | 0.8059 | 1.7962            | 0.9120            | 0.7928            | 1.4966            | 1.2494               | 1.1041 |
| -2.90                      | 0.9481 | 2.0849            | 1.0726            | 0.9338            | 1.6016            | 1.4232               | 1.2971 |
| -2.80                      | 1.1139 | 2.4125            | 1.2596            | 1.0985            | 1.7197            | 1.6223               | 1.5214 |
| -2.70                      | 1.3068 | 2.7819            | 1.4770            | 1.2907            | 1.8518            | 1.8504               | 1.7813 |
| -2.60                      | 1.5304 | 3.1957            | 1.7286            | 1.5141            | 1.0049            | 2.1108               | 2.0811 |
| -2.50                      | 1.7888 | 3.6557            | 2.0187            | 1.7731            | 2.1820            | 2.4074               | 2.4253 |

ตาราง 14 (ต่อ)

| ระดับ<br>ความสามารถ<br>( $\theta$ ) | TIF     |                   |                   |                   |                   |                      |         |
|-------------------------------------|---------|-------------------|-------------------|-------------------|-------------------|----------------------|---------|
|                                     | CONT    | TWOT <sub>1</sub> | TWOT <sub>2</sub> | TWOT <sub>3</sub> | TWOT <sub>4</sub> | TWOT <sub>mean</sub> | FLEXT   |
| -2.40                               | 2.0861  | 4.1634            | 2.3518            | 2.0722            | 2.3871            | 0.6037               | 2.8185  |
| -2.30                               | 2.4265  | 4.7186            | 2.7319            | 2.4161            | 2.6240            | 3.1227               | 3.2646  |
| -2.20                               | 2.8143  | 5.3207            | 3.1633            | 2.8096            | 2.8965            | 3.5475               | 3.7673  |
| -2.10                               | 3.2535  | 5.9675            | 3.6494            | 3.2573            | 3.2081            | 4.0206               | 4.3287  |
| -2.00                               | 3.7477  | 6.6552            | 4.1928            | 3.7634            | 3.5624            | 4.5435               | 4.9499  |
| -1.90                               | 4.3000  | 7.3787            | 4.7949            | 4.3313            | 3.9631            | 5.1170               | 5.6297  |
| -1.80                               | 4.9125  | 8.1309            | 5.4557            | 4.9637            | 4.4135            | 5.7410               | 6.3643  |
| -1.70                               | 5.5862  | 8.0931            | 6.1728            | 5.6613            | 4.9170            | 6.2111               | 7.1472  |
| -1.60                               | 6.3209  | 9.6843            | 6.9419            | 6.4234            | 5.4763            | 7.1315               | 7.9684  |
| -1.50                               | 7.1146  | 10.4614           | 7.7559            | 7.2468            | 6.0931            | 7.8893               | 8.8145  |
| -1.40                               | 7.9641  | 11.2196           | 8.6054            | 8.1257            | 6.7678            | 8.6766               | 9.6692  |
| -1.30                               | 8.8637  | 11.9422           | 9.4784            | 9.0514            | 7.4985            | 9.4926               | 10.5136 |
| -1.20                               | 9.8063  | 12.6117           | 10.3612           | 10.0123           | 8.2810            | 10.3166              | 11.3271 |
| -1.10                               | 10.7821 | 13.2101           | 11.2381           | 10.9938           | 9.1081            | 11.1375              | 12.0888 |
| -1.00                               | 11.7791 | 13.7203           | 12.0929           | 11.9789           | 9.9694            | 11.9404              | 12.7789 |
| -0.90                               | 12.7823 | 14.1272           | 12.9092           | 12.9481           | 10.8517           | 12.6341              | 13.3799 |
| -0.80                               | 13.7741 | 14.4186           | 13.6711           | 13.8805           | 11.7392           | 13.4274              | 13.8780 |
| -0.70                               | 14.7338 | 14.5865           | 14.3638           | 14.7546           | 12.6143           | 14.0798              | 14.2641 |
| -0.60                               | 15.6378 | 14.6200           | 14.9746           | 15.5491           | 13.4582           | 14.6506              | 14.5341 |
| -0.50                               | 16.4606 | 14.5419           | 15.4926           | 16.2441           | 14.2522           | 15.1327              | 14.6894 |
| -0.40                               | 17.1756 | 14.3366           | 15.9096           | 16.8218           | 14.9780           | 15.4365              | 14.7363 |

ตาราง 14 (ต่อ)

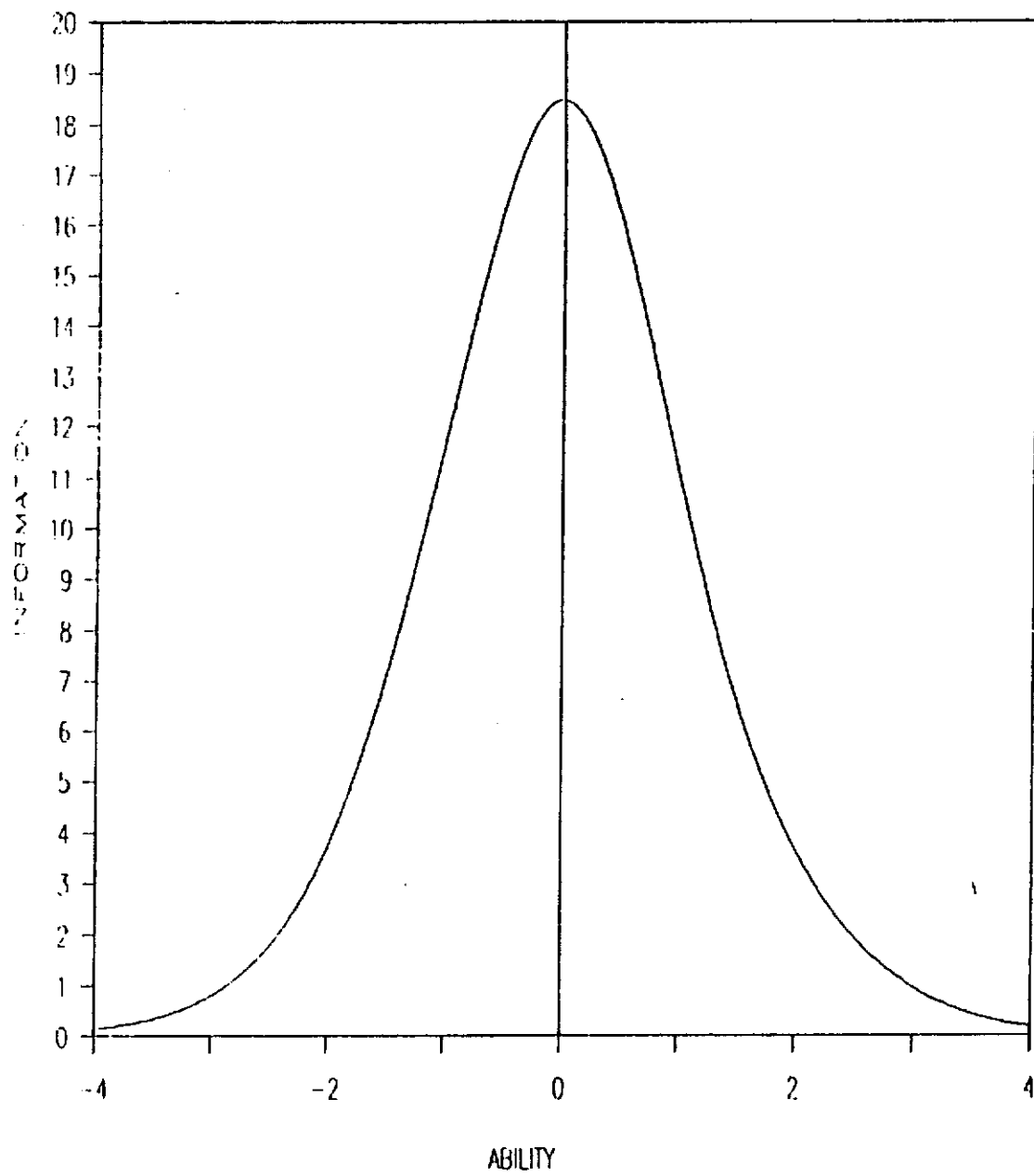
| ระดับ<br>ความสามารถ<br>(๑) | TIF     |                   |                   |                   |                   |                      |         |
|----------------------------|---------|-------------------|-------------------|-------------------|-------------------|----------------------|---------|
|                            | CONT    | TWOT <sub>1</sub> | TWOT <sub>2</sub> | TWOT <sub>3</sub> | TWOT <sub>4</sub> | TWOT <sub>mean</sub> | FLEXT   |
| -0.30                      | 17.7568 | 14.0214           | 16.2196           | 17.2680           | 15.6184           | 15.7819              | 14.6855 |
| -0.20                      | 18.1860 | 13.6096           | 16.4195           | 17.5724           | 16.1584           | 15.9400              | 14.5507 |
| -0.10                      | 18.4280 | 13.1169           | 16.5083           | 17.7291           | 16.5434           | 15.9847              | 14.3479 |
| 0.00                       | 18.4862 | 12.5601           | 16.4874           | 17.7370           | 16.8852           | 15.9174              | 14.0934 |
| 0.10                       | 18.3501 | 11.9567           | 16.3602           | 17.5993           | 17.0521           | 15.6670              | 13.8030 |
| 0.20                       | 18.0234 | 11.3232           | 16.1318           | 17.3231           | 17.0791           | 15.4643              | 13.4902 |
| 0.30                       | 17.5180 | 10.6752           | 15.8088           | 16.9188           | 16.9188           | 15.0915              | 13.1657 |
| 0.40                       | 16.8531 | 10.0265           | 15.3986           | 16.3989           | 16.3989           | 14.6322              | 12.8366 |
| 0.50                       | 16.0540 | 9.3892            | 14.9094           | 15.7776           | 15.7776           | 14.0962              | 12.5060 |
| 0.60                       | 15.1496 | 8.7729            | 14.3499           | 15.0696           | 15.0696           | 13.4939              | 12.1734 |
| 0.70                       | 14.1707 | 8.1855            | 13.7289           | 14.2902           | 15.1413           | 12.8364              | 11.8353 |
| 0.80                       | 13.1467 | 7.6325            | 13.0554           | 13.4546           | 14.3989           | 12.1354              | 11.4860 |
| 0.90                       | 12.1064 | 7.1175            | 12.3384           | 12.5778           | 13.5750           | 11.4022              | 11.1186 |
| 1.00                       | 11.0741 | 6.6422            | 11.5873           | 11.6747           | 12.6901           | 10.6486              | 10.7264 |
| 1.10                       | 10.0702 | 6.2070            | 10.8116           | 10.7598           | 11.7656           | 9.8860               | 10.3038 |
| 1.20                       | 9.1108  | 5.8107            | 10.0212           | 9.8469            | 10.8221           | 9.1252               | 10.3038 |
| 1.30                       | 8.2073  | 5.4516            | 9.2264            | 8.9491            | 9.8788            | 8.3765               | 9.8471  |
| 1.40                       | 7.3672  | 5.1273            | 8.4373            | 8.0781            | 8.9527            | 7.6489               | 8.8300  |
| 1.50                       | 6.5945  | 4.8352            | 7.6638            | 7.2442            | 8.0581            | 6.9503               | 8.2756  |
| 1.60                       | 5.8890  | 4.5727            | 6.9151            | 6.4558            | 7.2063            | 6.2875               | 7.6990  |
| 1.70                       | 5.2497  | 4.3373            | 6.1993            | 5.7191            | 6.4059            | 5.6654               | 7.1083  |

ตาราง 14 (ต่อ)

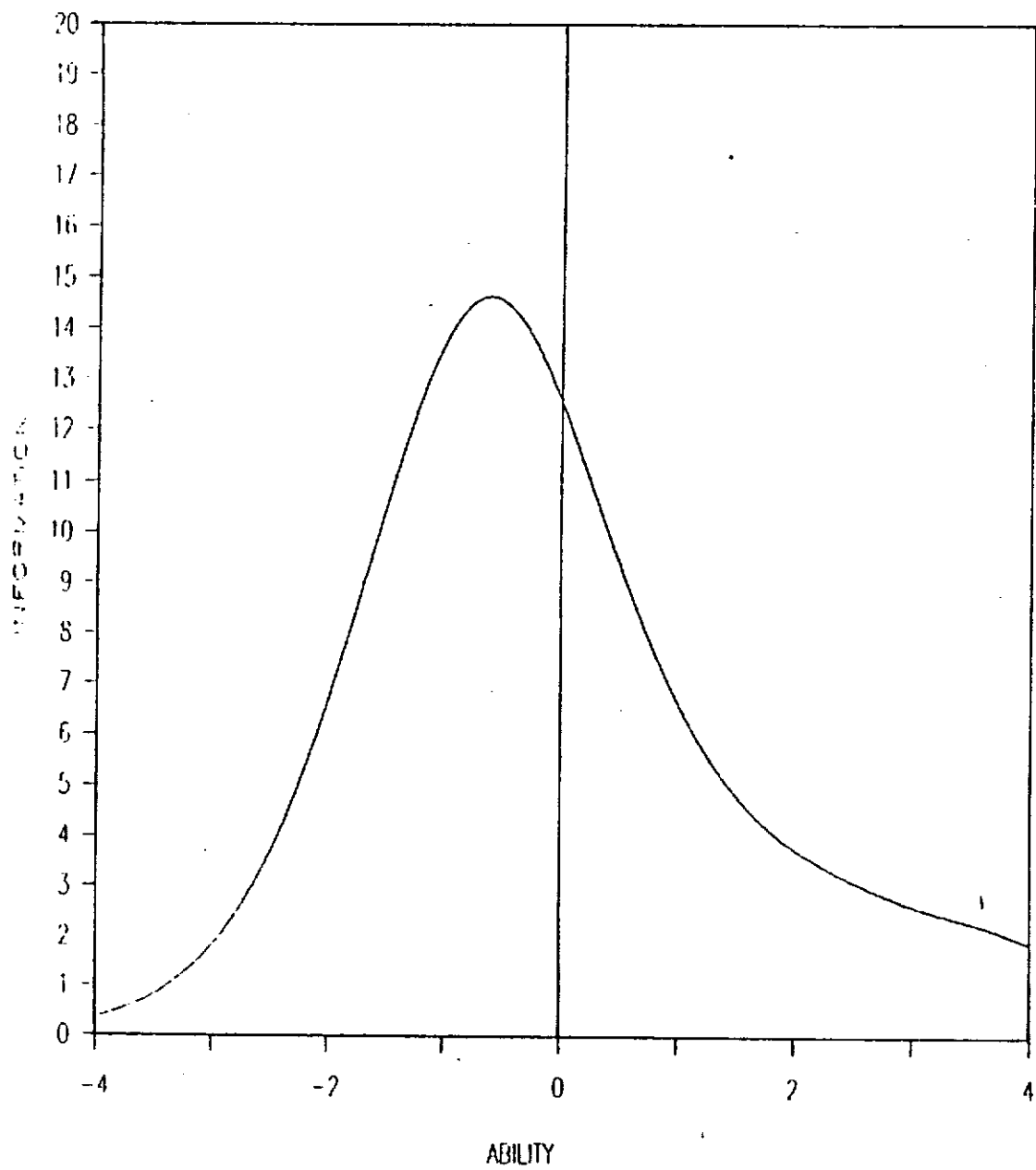
| ระดับ<br>ความสามารถ<br>(๑) | TIF    |                   |                   |                   |                   |                      |        |
|----------------------------|--------|-------------------|-------------------|-------------------|-------------------|----------------------|--------|
|                            | CONT   | TWOT <sub>1</sub> | TWOT <sub>2</sub> | TWOT <sub>3</sub> | TWOT <sub>4</sub> | TWOT <sub>mean</sub> | FLEXT  |
| 1.80                       | 4.6729 | 4.1263            | 5.5233            | 5.0385            | 5.6626            | 5.0877               | 6.5130 |
| 1.90                       | 4.1539 | 3.9373            | 4.8921            | 4.8921            | 4.9795            | 4.5563               | 5.9227 |
| 2.00                       | 3.6875 | 3.7675            | 4.3093            | 4.3093            | 4.3580            | 4.0719               | 5.3464 |
| 2.10                       | 3.2686 | 3.6143            | 3.7765            | 3.7765            | 3.7975            | 3.6338               | 4.7923 |
| 2.20                       | 2.8922 | 3.4748            | 3.2942            | 3.8527            | 3.2959            | 3.2403               | 4.2672 |
| 2.30                       | 2.5538 | 3.3465            | 2.8614            | 2.4977            | 2.8503            | 2.8890               | 3.7760 |
| 2.40                       | 2.2496 | 3.2273            | 2.4759            | 2.1476            | 2.4570            | 2.5770               | 3.3223 |
| 2.50                       | 1.9763 | 3.1154            | 2.1351            | 1.8416            | 2.1190            | 2.3010               | 2.9078 |
| 2.60                       | 1.7311 | 3.0097            | 1.8357            | 1.5756            | 1.8107            | 2.0579               | 2.5331 |
| 2.70                       | 1.5117 | 2.9099            | 1.5747            | 1.3453            | 1.5490            | 1.8446               | 2.1973 |
| 2.80                       | 1.3160 | 2.8160            | 1.3467            | 1.1467            | 1.3225            | 1.6580               | 1.8988 |
| 2.90                       | 1.1421 | 2.7282            | 1.1498            | 0.9759            | 1.1272            | 1.4953               | 1.6355 |
| 3.00                       | 0.9882 | 2.6467            | 0.9800            | 0.8295            | 0.9594            | 1.3539               | 1.4045 |
| 3.10                       | 0.8525 | 2.5712            | 0.8340            | 0.7042            | 0.8154            | 1.2312               | 1.2031 |
| 3.20                       | 0.7336 | 2.5010            | 0.7088            | 0.5973            | 0.6923            | 1.1249               | 1.2031 |
| 3.30                       | 0.6297 | 2.4346            | 0.6018            | 0.5.62            | 0.5873            | 1.0325               | 0.8773 |
| 3.40                       | 0.5393 | 2.3698            | 0.5104            | 0.4287            | 0.4977            | 0.9517               | 0.7427 |
| 3.50                       | 0.4609 | 2.3041            | 0.4326            | 0.3629            | 0.4216            | 0.8803               | 0.6355 |
| 3.60                       | 0.3933 | 2.2348            | 0.3664            | 0.3070            | 0.3568            | 0.8155               | 0.5398 |
| 3.70                       | 0.3350 | 2.1592            | 0.3101            | 0.2596            | 0.3019            | 0.7577               | 0.4581 |
| 3.80                       | 0.2850 | 2.0753            | 0.2623            | 0.2195            | 0.2553            | 0.7031               | 0.3883 |

ตาราง 14 (ต่อ)

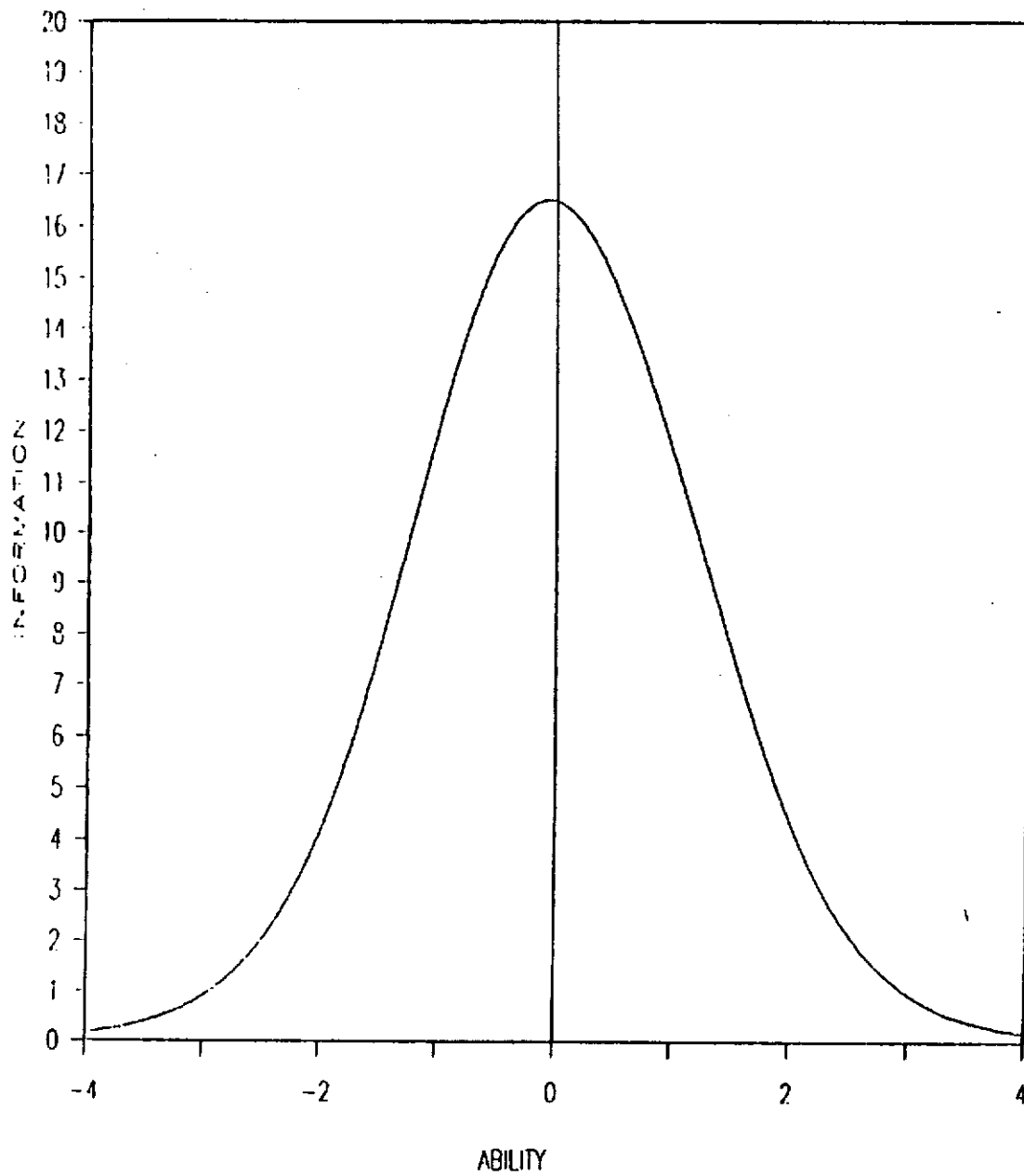
| ระดับ<br>ความสามารถ<br>(๐) | TIF    |                   |                   |                   |                   |                      |        |
|----------------------------|--------|-------------------|-------------------|-------------------|-------------------|----------------------|--------|
|                            | CONT   | TWOT <sub>1</sub> | TWOT <sub>2</sub> | TWOT <sub>3</sub> | TWOT <sub>4</sub> | TWOT <sub>mean</sub> | FLEXT  |
| 3.90                       | 0.2422 | 1.9818            | 0.2218            | 0.1855            | 0.2158            | 0.6512               | 0.3290 |
| 4.00                       | 0.2056 | 1.8785            | 0.1875            | 0.1567            | 0.1824            | 0.6013               | 0.2786 |



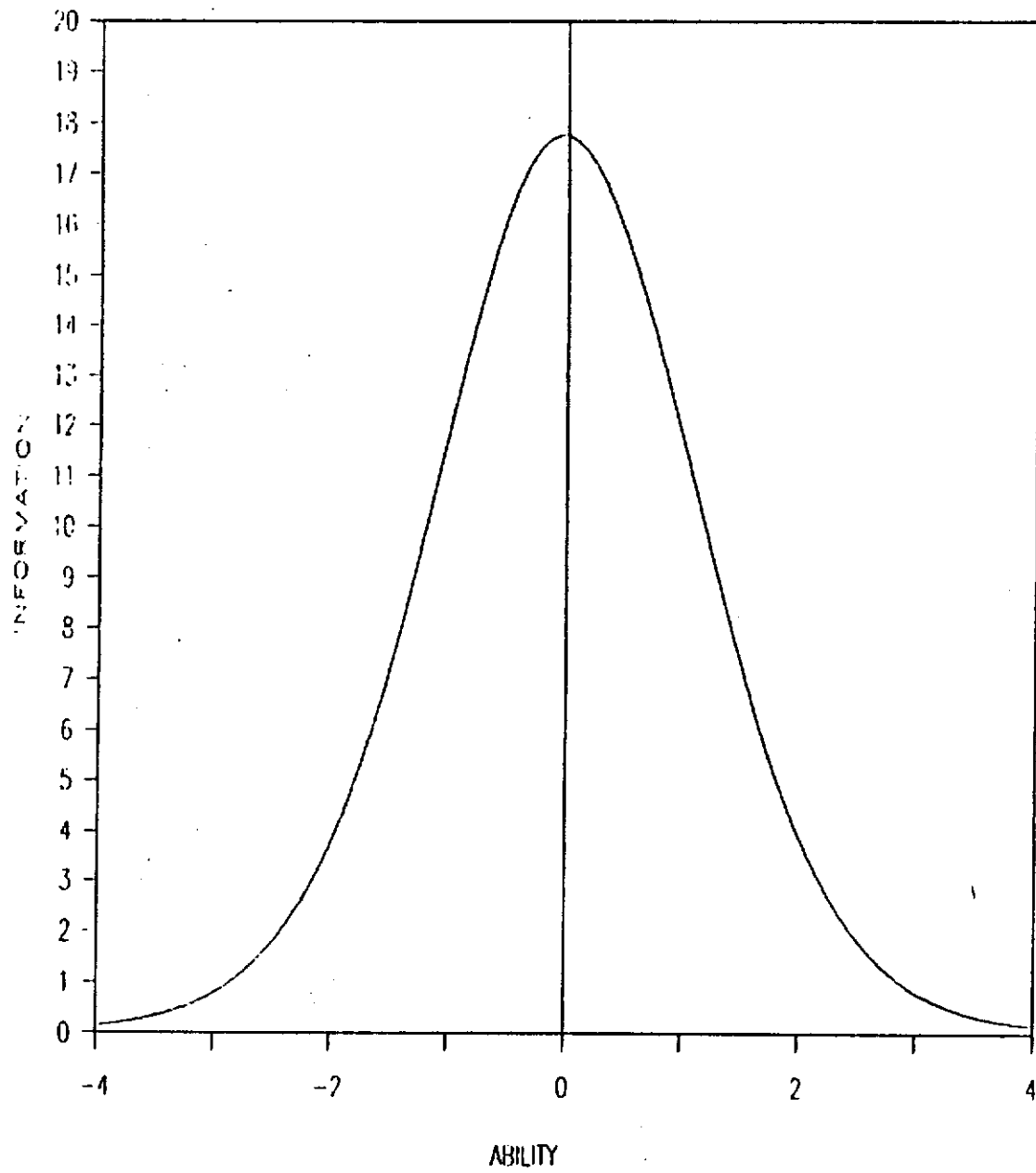
ภาพประกอบ 17 ได้ข้อมูลเมชันของแบบทดสอบดั้งเดิม



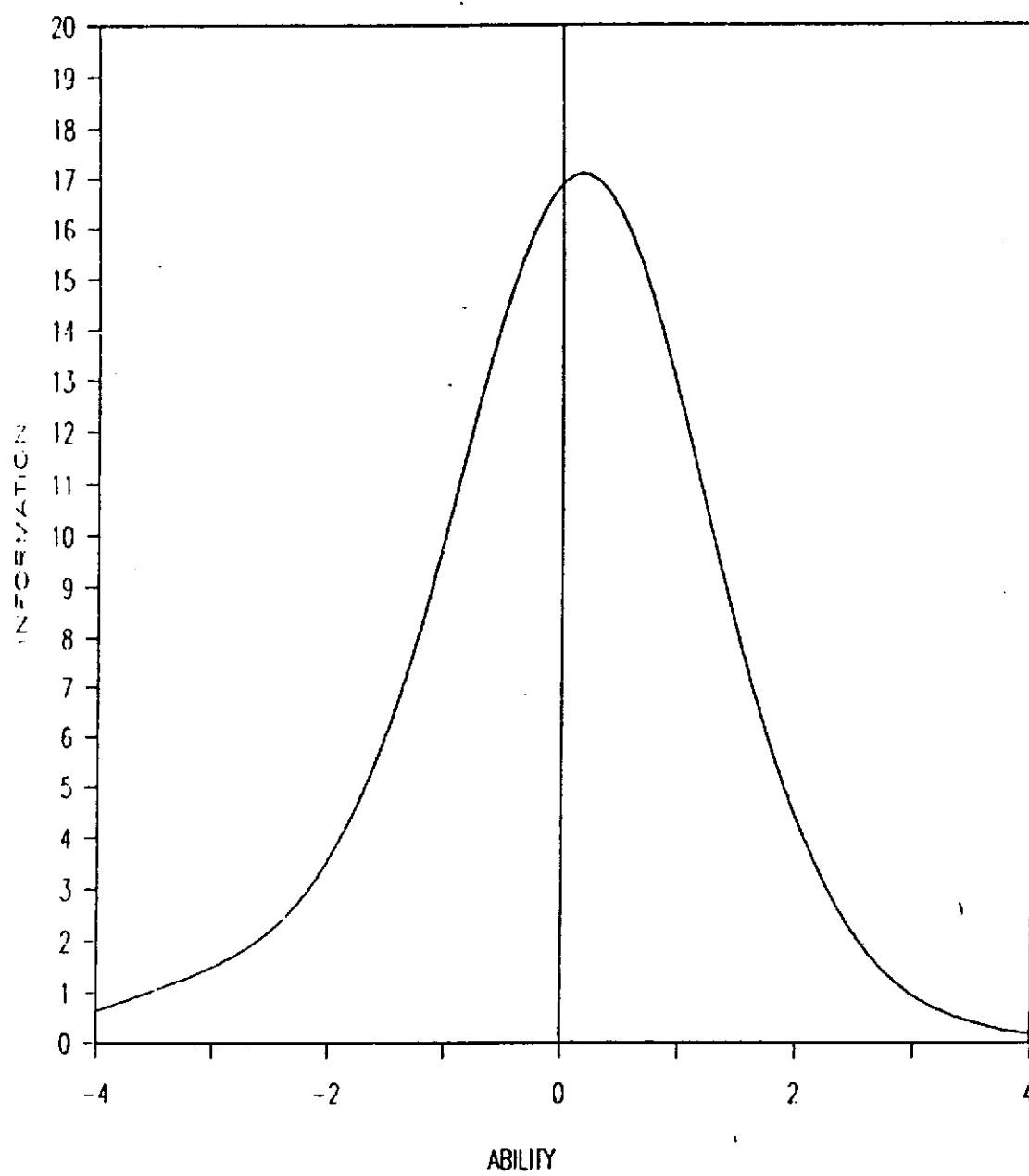
ภาพประกอบ 18 โค้งอินฟอร์เมชันของแบบทดสอบสองขั้นตอน (แบบทดสอบวัดผลฉบับที่ 1)



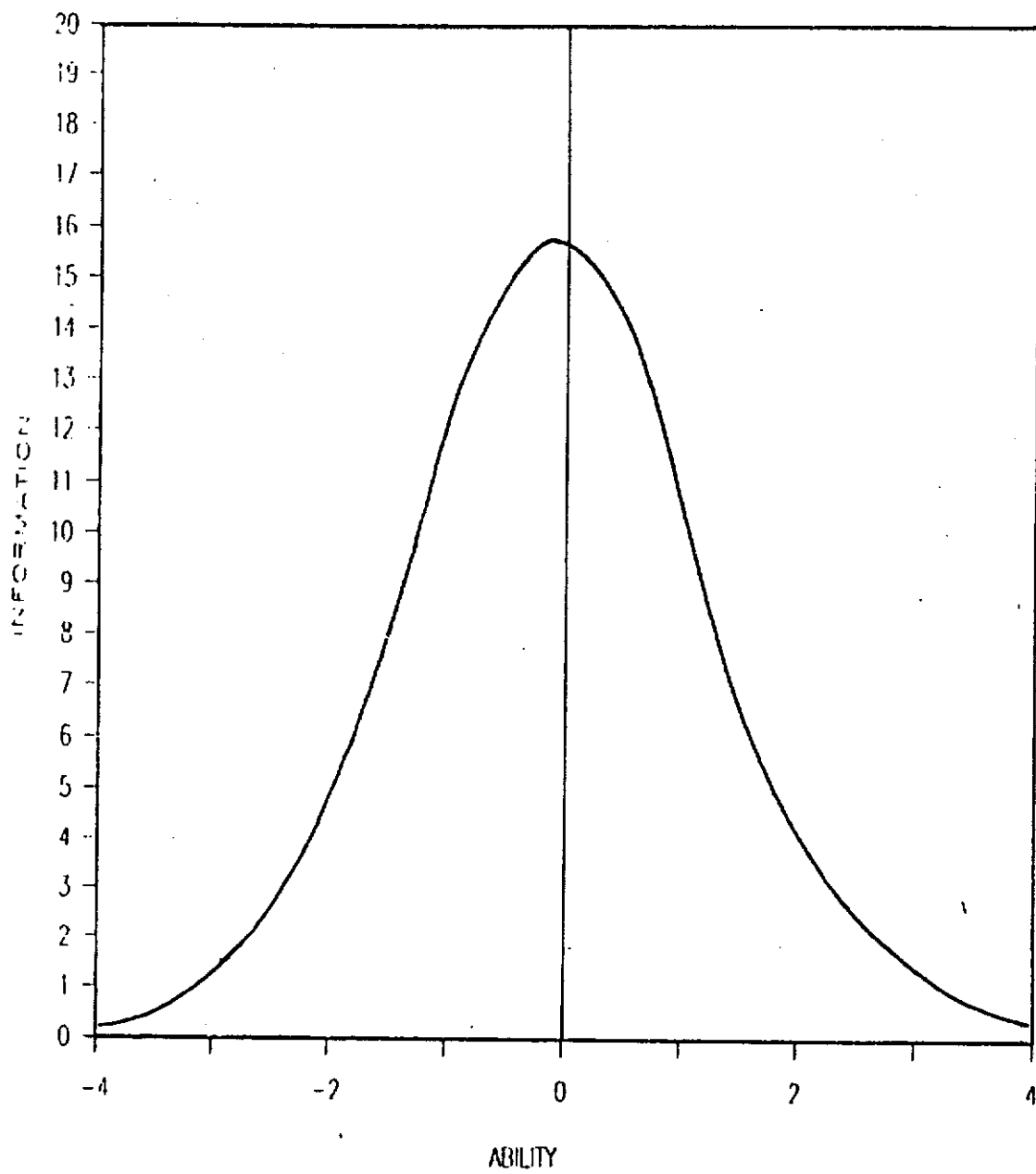
ภาพประกอบ 19 โค้งอินฟอร์เมชันของแบบทดสอบสองขั้นตอน (แบบทดสอบวัดผลสัมฤทธิ์ที่ 2)



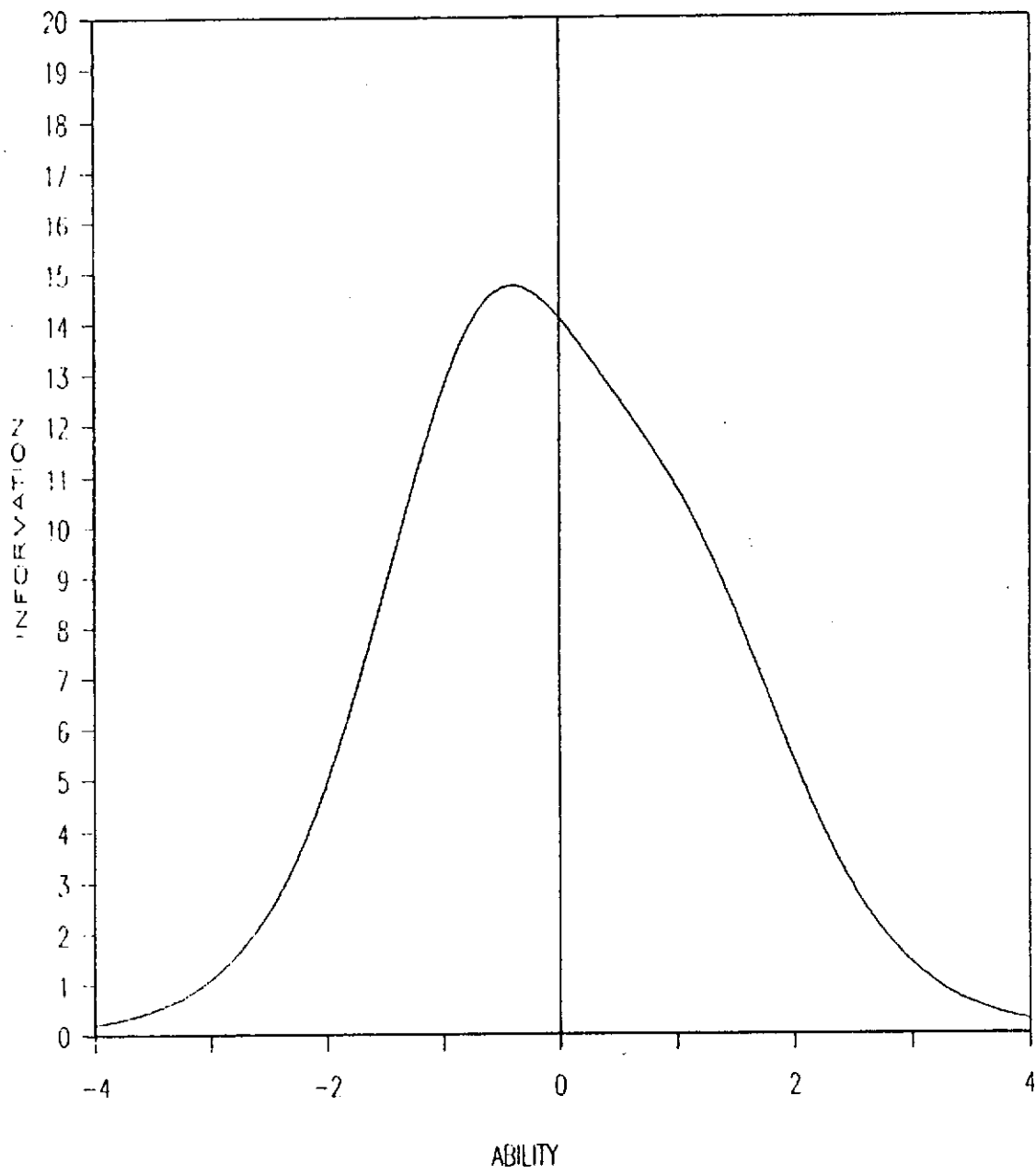
ภาพประกอบ 20 โค้งอินฟอร์เมชันของแบบทดสอบสองขั้นตอน (แบบทดสอบวัดผลฉบับที่ 3)



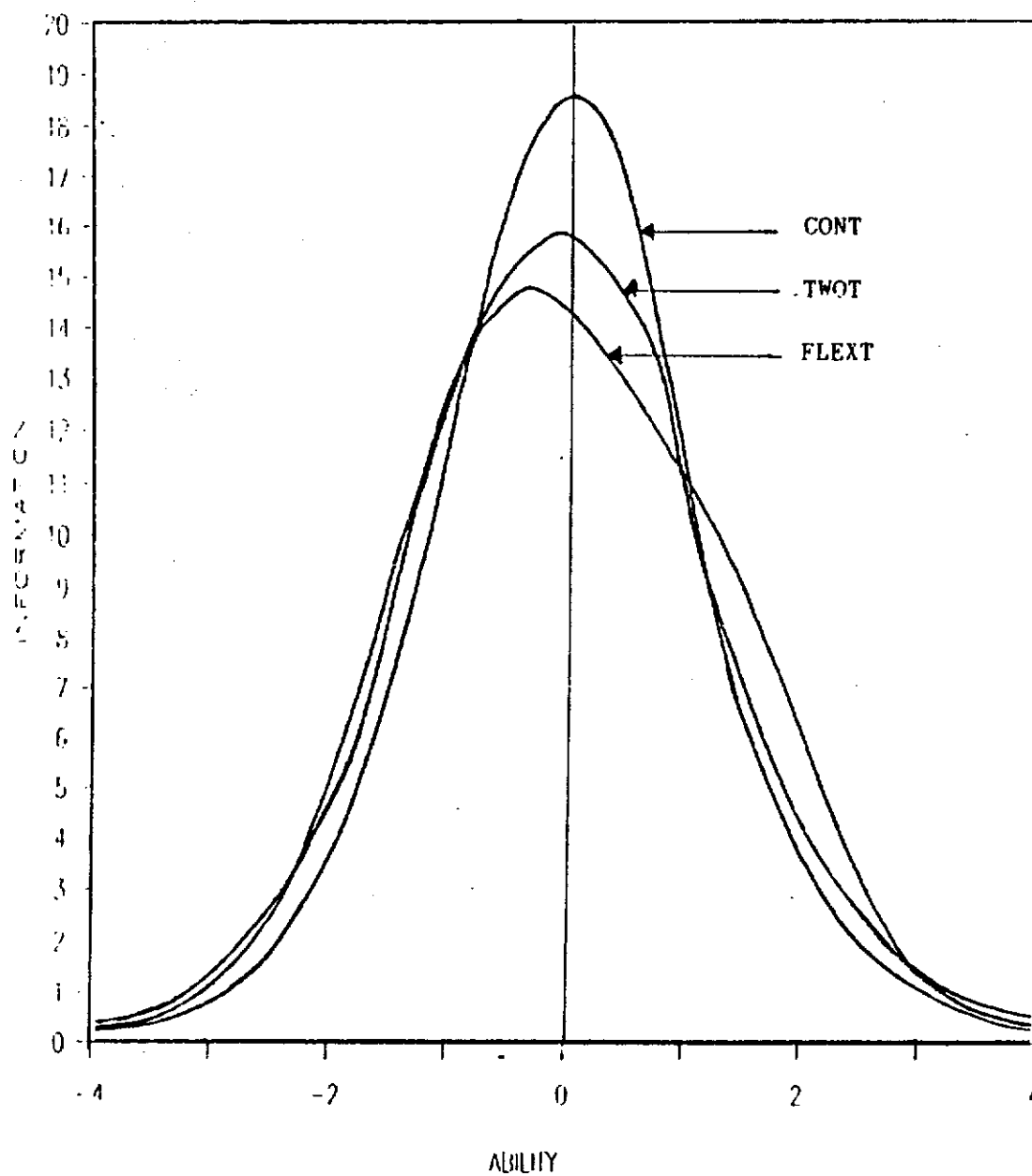
ภาพประกอบ 21 ได้แก่อินฟอร์เมชันของแบบทดสอบสองขั้นตอน (แบบทดสอบวัดผลสัมฤทธิ์ที่ 4)



ภาพประกอบ 22 โค้งอินฟอร์เมชันของแบบทดสอบสองขั้นตอน (เจดีย์แบบทดสอบวัดผล  
ทั้ง 4 ฉบับ)



ภาพประกอบ 23 ได้ังอินฟอร์เมชันของแบบทดสอบเฟลิกซ์เลเวล



ภาพประกอบ 24 ไค้งอินฟอร์เมชันของแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และ  
แบบทดสอบเฟล็กซีเบิลเวล

ภาคผนวก จ

ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถระหว่าง  
แบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กชีเลเวล

ตาราง 15 ค่าประสิทธิภาพสัมพัทธ์ (RE) ของการประมาณค่าความสามารถของผู้สอบ  
ระหว่างแบบทดสอบดั้งเดิม (CONT) แบบทดสอบสองขั้นตอน (TWOT) และแบบทดสอบ  
แฟล็กซีเบิลเวล (FLEXT) ที่ระดับความสามารถตั้งแต่ -4.00 ถึง 4.00 จำนวน 81  
ระดับความสามารถ

| ระดับ<br>ความสามารถ<br>( $\theta$ ) | RE        |            |            |
|-------------------------------------|-----------|------------|------------|
|                                     | CONT/TWOT | CONT/FLEXT | TWOT/FLEXT |
| -4.00                               | 0.4568    | 0.7256     | 1.5884     |
| -3.90                               | 0.4703    | 0.7258     | 1.5435     |
| -3.80                               | 0.4850    | 0.7258     | 1.4964     |
| -3.70                               | 0.5014    | 0.7260     | 1.4481     |
| -3.60                               | 0.5193    | 0.7264     | 1.3987     |
| -3.50                               | 0.5385    | 0.7268     | 1.3497     |
| -3.40                               | 0.5588    | 0.7272     | 1.3014     |
| -3.30                               | 0.5799    | 0.7277     | 1.2549     |
| -3.20                               | 0.6016    | 0.7283     | 1.2107     |
| -3.10                               | 0.6235    | 0.7291     | 1.1694     |
| -3.00                               | 0.6435    | 0.7300     | 1.1316     |
| -2.90                               | 0.6662    | 0.7310     | 1.0972     |
| -2.80                               | 0.6866    | 0.7322     | 1.0663     |
| -2.70                               | 0.7062    | 0.7336     | 1.0388     |
| -2.60                               | 0.7250    | 0.7354     | 1.0143     |
| -2.50                               | 0.7430    | 0.7376     | 0.9926     |
| -2.40                               | 0.7604    | 0.7401     | 0.9734     |
| -2.30                               | 0.7771    | 0.7433     | 0.9565     |
| -2.20                               | 0.7933    | 0.7470     | 0.9417     |

ตาราง 15 (ต่อ)

| ระดับ<br>ความสามารถ<br>(๑) | RE        |            |            |
|----------------------------|-----------|------------|------------|
|                            | CONT/TWOT | CONT/FLEXT | TWOT/FLEXT |
| -2.10                      | 0.8092    | 0.7516     | 0.9288     |
| -2.00                      | 0.8248    | 0.7571     | 0.9162     |
| -1.90                      | 0.8403    | 0.7638     | 0.9089     |
| -1.80                      | 0.8557    | 0.7719     | 0.9021     |
| -1.70                      | 0.8994    | 0.7816     | 0.8690     |
| -1.60                      | 0.8863    | 0.7932     | 0.8950     |
| -1.50                      | 0.9018    | 0.8071     | 0.8959     |
| -1.40                      | 0.9176    | 0.8237     | 0.8977     |
| -1.30                      | 0.9337    | 0.8431     | 0.9029     |
| -1.20                      | 0.9505    | 0.8657     | 0.9108     |
| -1.10                      | 0.9681    | 0.8919     | 0.9213     |
| -1.00                      | 0.9865    | 0.9218     | 0.9343     |
| -0.90                      | 1.0117    | 0.9553     | 0.9443     |
| -0.80                      | 1.0258    | 0.9925     | 0.9675     |
| -0.70                      | 1.0464    | 1.0330     | 0.9871     |
| -0.60                      | 1.0673    | 1.0760     | 1.0080     |
| -0.50                      | 1.0878    | 1.1206     | 1.0302     |
| -0.40                      | 1.1127    | 1.1655     | 1.0475     |
| -0.30                      | 1.1251    | 1.2091     | 1.0747     |
| -0.20                      | 1.1460    | 1.2495     | 1.0955     |
| -0.10                      | 1.1529    | 1.2844     | 1.1141     |

ตาราง 15 (ต่อ)

| ระดับ<br>ความสามารถ<br>(๑) | RE        |            |            |
|----------------------------|-----------|------------|------------|
|                            | CONT/TWOT | CONT/FLEXT | TWOT/FLEXT |
| 0.00                       | 1.1614    | 1.3117     | 1.1294     |
| 0.10                       | 1.1713    | 1.3294     | 1.1350     |
| 0.20                       | 1.1655    | 1.3360     | 1.1463     |
| 0.30                       | 1.1608    | 1.3306     | 1.1463     |
| 0.40                       | 1.1518    | 1.3129     | 1.1399     |
| 0.50                       | 1.1389    | 1.2837     | 1.1272     |
| 0.60                       | 1.1277    | 1.2445     | 1.1085     |
| 0.70                       | 1.1039    | 1.1973     | 1.0846     |
| 0.80                       | 1.0833    | 1.1446     | 1.0565     |
| 0.90                       | 1.0618    | 1.0889     | 1.0225     |
| 1.00                       | 1.0400    | 1.0324     | 0.9927     |
| 1.10                       | 1.0186    | 0.9773     | 0.9595     |
| 1.20                       | 0.9984    | 0.9252     | 0.8954     |
| 1.30                       | 0.9799    | 0.8773     | 0.8954     |
| 1.40                       | 0.9632    | 0.8343     | 0.8662     |
| 1.50                       | 0.9488    | 0.7968     | 0.8399     |
| 1.60                       | 0.9366    | 0.7650     | 0.8167     |
| 1.70                       | 0.9266    | 0.7385     | 0.7970     |
| 1.80                       | 0.9185    | 0.7175     | 0.7812     |
| 1.90                       | 0.9117    | 0.7014     | 0.7693     |
| 2.00                       | 0.9056    | 0.6897     | 0.7616     |

ตาราง 15 (ต่อ)

| ระดับ<br>ความสามารถ<br>(๑) | RE        |            |            |
|----------------------------|-----------|------------|------------|
|                            | CONT/TWOT | CONT/FLEXT | TWOT/FLEXT |
| 2.10                       | 0.8995    | 0.6821     | 0.7583     |
| 2.20                       | 0.8926    | 0.6778     | 0.7594     |
| 2.30                       | 0.8840    | 0.6763     | 0.7651     |
| 2.40                       | 0.8730    | 0.6771     | 0.7757     |
| 2.50                       | 0.8589    | 0.6797     | 0.7913     |
| 2.60                       | 0.8412    | 0.6834     | 0.8124     |
| 2.70                       | 0.8195    | 0.6880     | 0.8395     |
| 2.80                       | 0.7937    | 0.6731     | 0.8732     |
| 2.90                       | 0.7638    | 0.6983     | 0.9143     |
| 3.00                       | 0.7299    | 0.7036     | 0.9640     |
| 3.10                       | 0.6924    | 0.7086     | 1.0234     |
| 3.20                       | 0.6521    | 0.7133     | 1.0938     |
| 3.30                       | 0.6099    | 0.7178     | 1.1769     |
| 3.40                       | 0.5667    | 0.7218     | 1.2737     |
| 3.50                       | 0.5236    | 0.7253     | 1.3852     |
| 3.60                       | 0.4823    | 0.7286     | 1.5107     |
| 3.70                       | 0.4421    | 0.7313     | 1.6540     |
| 3.80                       | 0.4053    | 0.7340     | 1.8107     |
| 3.90                       | 0.3719    | 0.7362     | 1.9723     |
| 4.00                       | 0.3419    | 0.7380     | 2.1583     |

ภาคผนวก จ

จุดประสงค์การเรียนรู้ เนื้อหา และจุดประสงค์เชิงพฤติกรรม  
วิชาคณิตศาสตร์ (ค 011) ชั้นมัธยมศึกษาปีที่ 3 เรื่อง เลขยกกำลัง

ตาราง 16 รายละเอียดจุดประสงค์การเรียนรู้ เนื้อหา และจุดประสงค์เชิงพฤติกรรม  
 วิชาคณิตศาสตร์ (ค 011) ชั้นมัธยมศึกษาปีที่ 3 เรื่องเลขยกกำลัง ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533)

### จุดประสงค์การเรียนรู้

1. นักเรียนสามารถหาผลคูณและผลหารของจำนวนที่เขียนอยู่ในรูปเลขยกกำลังที่มีเลขชี้กำลังเป็นจำนวนเต็ม โดยใช้นิยามและคุณสมบัติต่าง ๆ ของเลขยกกำลัง
2. นักเรียนสามารถใช้เลขยกกำลังในการเขียนแสดงจำนวนที่มีค่าน้อย ๆ หรือมาก ๆ ในรูป  $A \times 10^n$  เมื่อ  $1 < A < 10$  และ  $n$  เป็นจำนวนเต็ม

| เนื้อหา                                              | จุดประสงค์เชิงพฤติกรรม                                                                                                                 |
|------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|
| 1. ความหมายของเลขยกกำลัง                             | 1. เมื่อกำหนดเลขยกกำลังมาให้ นักเรียนสามารถบอกความหมายของเลขยกกำลังได้ และสามารถเขียนจำนวนที่กำหนดมาให้ในรูปของเลขยกกำลังได้           |
| 2. การคูณเลขยกกำลัง เมื่อเลขชี้กำลังเป็นจำนวนเต็มบวก | 2. เมื่อกำหนดโจทย์ปัญหาการคูณของเลขยกกำลังที่มีฐานเท่ากันและมีเลขชี้กำลังเป็นจำนวนเต็มบวก นักเรียนสามารถหาผลคูณได้                     |
| 3. การหารเลขยกกำลัง เมื่อเลขยกกำลังเป็นจำนวนเต็มบวก  | 3. เมื่อกำหนดโจทย์ปัญหาการหารของเลขยกกำลังที่มีฐานเท่ากันและ ไม่เท่ากับศูนย์ โดยมีเลขชี้กำลังเป็นจำนวนเต็มบวก นักเรียนสามารถหาผลหารได้ |
| 4. การคูณเลขยกกำลัง เมื่อเลขชี้กำลังเป็นจำนวนเต็ม    | 4. เมื่อกำหนดโจทย์ปัญหาการคูณของเลขยกกำลังที่มีฐานเท่ากันและมีเลขชี้กำลังเป็นจำนวนเต็ม นักเรียนสามารถหาผลคูณได้                        |

| เนื้อหา                                           | จุดประสงค์เชิงพฤติกรรม                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|---------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 5. การหารเลขยกกำลัง เมื่อเลขชี้กำลังเป็นจำนวนเต็ม | 5. เมื่อกำหนดโจทย์ปัญหาการหารของเลขยกกำลังที่มีฐานเท่ากันและมีเลขชี้กำลังเป็นจำนวนเต็ม นักเรียนสามารถหาผลหารได้                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 6. การนำความรู้เรื่องเลขยกกำลังไปแก้โจทย์ปัญหา    | <p>6. เมื่อกำหนดจำนวนที่มีค่ามาก ๆ มาให้นักเรียนสามารถเขียนแทนด้วยเลขยกกำลังในรูป <math>A \times 10^n</math> เมื่อ <math>1 \leq A &lt; 10</math> และ <math>n</math> เป็นจำนวนเต็มได้</p> <p>7. เมื่อกำหนดจำนวนที่มีค่าน้อย ๆ มาให้นักเรียนสามารถเขียนแทนด้วยเลขยกกำลังในรูป <math>A \times 10^n</math> เมื่อ <math>1 \leq A &lt; 10</math> และ <math>n</math> เป็นจำนวนเต็มได้</p> <p>8. เมื่อกำหนดเลขยกกำลังในรูป <math>A \times 10^n</math> เมื่อ <math>1 \leq A &lt; 10</math> และ <math>n</math> เป็นจำนวนเต็ม นักเรียนสามารถเปลี่ยนให้อยู่ในรูปจำนวนเต็มหรือทศนิยมได้</p> <p>9. เมื่อกำหนดโจทย์ปัญหาที่เกี่ยวกับเลขยกกำลัง มาให้นักเรียนสามารถใช้ความรู้เรื่องเลขยกกำลัง ไปแก้โจทย์ปัญหาได้</p> |

## ประวัติย่อของผู้วิจัย

ชื่อ นายอรินทร์ ชื่อสกุล น่วมถนอม

เกิดวันที่ 4 เดือนมิถุนายน พุทธศักราช 2507

สถานที่เกิด อำเภอวิเศษชัยชาญ จังหวัดอ่างทอง

สถานที่อยู่ปัจจุบัน บ้านเลขที่ 13/1 หมู่ 8 ตำบลศาลเจ้าโรงทอง

อำเภอวิเศษชัยชาญ จังหวัดอ่างทอง โทร. 631613, 631617

ตำแหน่งหน้าที่การงาน อาจารย์ 1 ระดับ 4

สถานที่ทำงานปัจจุบัน โรงเรียนบางกะปิ เขตบางกะปิ กรุงเทพฯ โทร. 3775870

## ประวัติการศึกษา

พ.ศ.2525 มัธยมศึกษาตอนปลาย (สายวิทย์ - คณิต)

จากโรงเรียนวิเศษชัยชาญ "ตันติวิทยานุกูล" จังหวัดอ่างทอง

พ.ศ.2529 ค.บ. (คณิตศาสตร์)

จากวิทยาลัยครูพระนครศรีอยุธยา

พ.ศ.2537 กศ.ม. (การวัดผลการศึกษา)

จากมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

การศึกษาความเที่ยงตรงและประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ  
ในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน  
และแบบทดสอบแฟล็กชีลเวล

บทคัดย่อ  
ของ  
อรินทร์ น่วมกนอม

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา

ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

พฤษภาคม 2537

371.271

๑396 ก

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อศึกษาความเที่ยงตรงและประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ ระหว่างแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กซีเลเวล โดยศึกษากับกลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2536 ของโรงเรียนสังกัดกรมสามัญศึกษา กรุงเทพมหานคร จำนวน 441 คน ซึ่งได้มาจากการสุ่มแบบแบ่งชั้น เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ที่ผู้วิจัยสร้างขึ้นเอง ซึ่งมี 3 รูปแบบ คือ แบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน และแบบทดสอบแฟล็กซีเลเวล การประมาณค่าความยากของข้อสอบ และค่าความสามารถของผู้สอบใช้โปรแกรมไบคาล ตามทฤษฎีการตอบสนองข้อสอบ แบบราส์ไนเคล ด้วยวิธีการประมาณค่าความเป็นไปได้สูงสุดแบบไม่มีเงื่อนไข

ผลการวิจัยสรุปได้ดังนี้

1. ค่าความเที่ยงตรงของการประมาณค่าความสามารถ ระหว่างแบบทดสอบดั้งเดิมกับแบบทดสอบสองขั้นตอน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนค่าความเที่ยงตรงของการประมาณค่าความสามารถ ระหว่างแบบทดสอบดั้งเดิมกับแบบทดสอบแฟล็กซีเลเวล และแบบทดสอบสองขั้นตอนกับแบบทดสอบแฟล็กซีเลเวล มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ ระหว่างแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบสองขั้นตอน และแบบทดสอบดั้งเดิมเทียบกับแบบทดสอบแฟล็กซีเลเวล มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมาและระดับสูงขึ้นไป ส่วนค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ ระหว่างแบบทดสอบสองขั้นตอนเทียบกับแบบทดสอบแฟล็กซีเลเวลมีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำมาก ระดับปานกลาง และระดับสูงมาก แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำและระดับสูง

A STUDY OF VALIDITY AND RELATIVE EFFICIENCY OF ABILITY  
ESTIMATION IN MATHEMATICS BETWEEN CONVENTIONAL,  
TWO - STAGE AND FLEXILEVEL TESTS

AN ABSTRACT  
BY  
ARIN NAUMTANOM

Presented in partial fulfillment of the requirements for the  
Master of Education degree in Educational Measurement  
at Srinakharinwirot University

May 1994

The purpose of this research was to study the validity and relative efficiency of ability estimation in Mathematics between conventional, two - stage and flexilevel tests. A sample of 441 Mathayomsuksa 3 students of the academic year 1993 in Bangkok was selected by stratified random sampling technique. The research instruments used were Mathematics achievement tests constructed by the researcher. Three test formats were conventional, two - stage and flexilevel tests. Unconditional Maximum Likelihood Procedure was used to estimate items difficulty and examinees' ability for the Rasch Model in Item Response Theory by the BICAL computer program.

The results of the research were Summarized as follows.

1. The validity of ability estimation between conventional and two - stage tests had statistically significant different at .01 level. For the validity of ability estimation between conventional and flexilevel, two - stage and flexilevel tests had not statistically significant different.

2. The relative efficiency of ability estimation between conventional and two - stage, conventional and flexilevel tests had more than 1 at the moderate ability levels but less than 1 at the very low, low, high and very high ability levels. For the relative efficiency of ability estimation between two - stage and flexilevel tests had more than 1 at the very low, moderate and very high ability levels but less than 1 at the low and high ability levels.