

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันแบบไม่สม่ำเสมอของข้อสอบ ระหว่างวิธีแมนเทิล-แฮนส์เซล แบบปกติ กับ วิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ โดยใช้ข้อมูลเชิงประจักษ์ ซึ่งมีวิธีดำเนินการวิจัย ดังนี้

ประชากร

ประชากรในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชายและหญิง ชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2539 ของโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา อำเภอเมือง จังหวัดชลบุรี จำนวน 7 โรงเรียน 71 ห้องเรียน รวมนักเรียนทั้งสิ้น 3,050 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชายและหญิง ชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2539 ของโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา อำเภอเมือง จังหวัดชลบุรี จำนวน 1,200 คน จำแนกเป็นเพศชาย 600 คน และเพศหญิง 600 คน ซึ่งได้มาโดยวิธีสุ่มตัวอย่างแบบหลายขั้นตอน และมีนักเรียนเป็นหน่วยการสุ่ม ขั้นตอนในการสุ่มตัวอย่าง เป็นดังนี้

1. กำหนดขนาดกลุ่มตัวอย่างทั้งหมดที่ความเชื่อมั่นระดับ 95 % ($\alpha = .05$) เมื่อเทียบกับตารางขนาดกลุ่มตัวอย่างขั้นต่ำ สำหรับศึกษาสัดส่วนของประชากรได้เท่ากับ 359 คน (Yamane, 1973)

สำหรับขนาดกลุ่มตัวอย่างที่ใช้ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ด้วยวิธีแมนเทิล-แฮนส์เซล Hambleton, et al. (1993) ได้แนะนำว่า ควรใช้กลุ่มผู้สอบระหว่าง 200 ถึง 1,000 คนต่อกลุ่ม แต่เนื่องจากวิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ต้องวิเคราะห์ข้อมูลโดยการแบ่งกลุ่มผู้สอบออกเป็น กลุ่มผู้สอบที่มี

ความสามารถสูงกับกลุ่มผู้สอบที่มีความสามารถต่ำ ดังนั้น เพื่อให้กลุ่มผู้สอบย่อยทั้งสองกลุ่มมีโอกาสที่จะมีจำนวนผู้สอบตั้งแต่ 200 คนขึ้นไป ผู้วิจัยจึงเลือกใช้กลุ่มตัวอย่างขนาด 1,200 คน จำแนกเป็นกลุ่มอ้างอิง (เพศหญิง) 600 คน และกลุ่มสนใจ (เพศชาย) 600 คน

2. การได้มาซึ่งกลุ่มตัวอย่างใช้วิธีสุ่มตัวอย่างแบบหลายขั้นตอน มีขั้นตอน ดังนี้

ขั้นตอนที่ 1 ใช้วิธีสุ่มตัวอย่างแบบแบ่งชั้น มีขนาดของโรงเรียนเป็นตัวแปรแบ่งชั้นภูมิ จำแนกโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา อำเภอเมือง จังหวัดชลบุรี ออกเป็น 2 ขนาด ได้แก่ โรงเรียนขนาดใหญ่ (มีระดับมัธยมศึกษาตอนต้นและตอนปลาย) 5 โรงเรียน และโรงเรียนขนาดเล็ก (มีเฉพาะระดับมัธยมศึกษาตอนต้น) 2 โรงเรียน สุ่มจำนวนโรงเรียนในอัตราส่วนขนาดใหญ่ : ขนาดเล็ก เท่ากับ 2 : 1 ได้โรงเรียนขนาดใหญ่ 2 โรงเรียน ได้แก่ โรงเรียนชลบุรี “สุขบท” และโรงเรียนแสนสุข กับ โรงเรียนขนาดเล็ก 1 โรงเรียน ได้แก่ โรงเรียนอ่างศิลาพิทยาคม

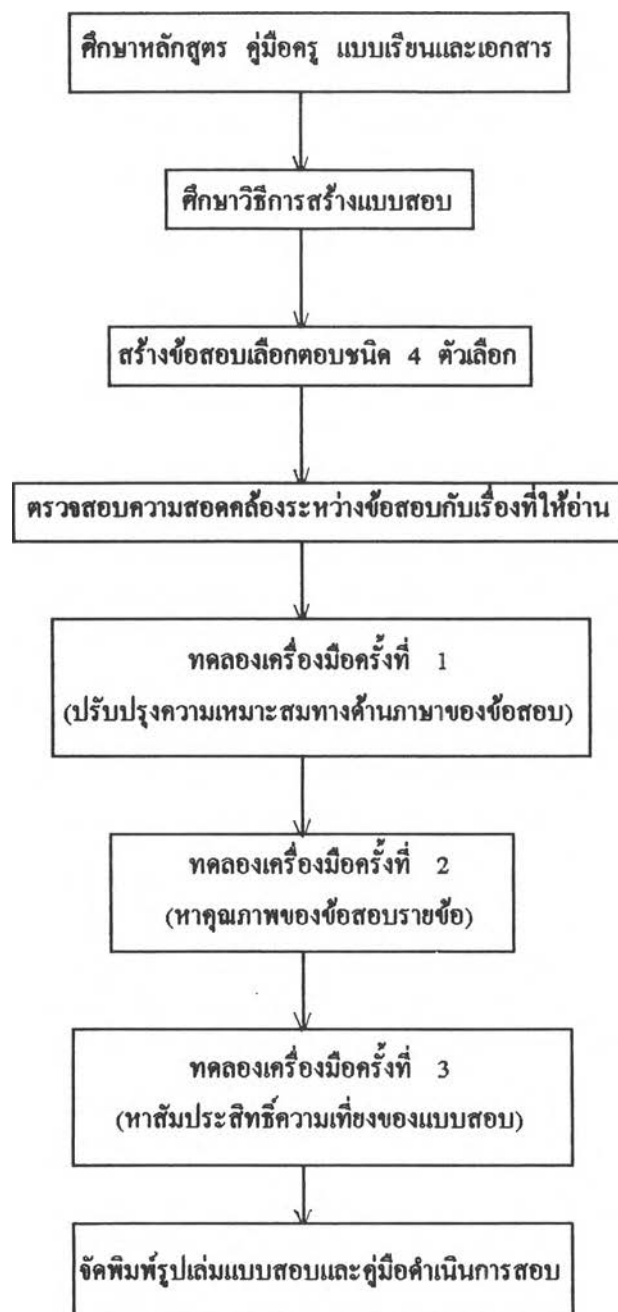
ขั้นตอนที่ 2 ใช้วิธีสุ่มตัวอย่างแบบง่าย สุ่มนักเรียนในโรงเรียนที่ได้จากการสุ่มขั้นตอนที่ 1 ได้จำนวนนักเรียน 1,200 คน จำแนกเป็นเพศชาย 600 คน และเพศหญิง 600 คน รายละเอียดดังตารางที่ 4

ตารางที่ 4 รายชื่อโรงเรียน และ จำนวนนักเรียนที่ใช้เป็นกลุ่มตัวอย่างในการวิจัย

โรงเรียน	จำนวนนักเรียนที่ใช้เป็นกลุ่มตัวอย่าง		
	รวม	ชาย	หญิง
1. ชลบุรี “สุขบท”	642	319	323
2. แสนสุข	383	188	195
3. อ่างศิลาพิทยาคม	175	93	82
รวม	1,200	600	600

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ เป็นแบบสอบวัดความสามารถในการอ่านภาษาไทย ชั้นมัธยมศึกษาปีที่ 1 ชนิดเลือกตอบ 4 ตัวเลือก จำนวน 75 ข้อ ซึ่งผู้วิจัยสร้างขึ้น การสร้างแบบสอบวัดความสามารถในการอ่านภาษาไทย ได้ดำเนินการตามขั้นตอน ดังนี้



รายละเอียดของการสร้างแบบสอบแต่ละขั้นตอน มีดังนี้

1. ศึกษาหลักสูตร คู่มือครู แบบเรียนและเอกสารต่าง ๆ ในระดับชั้นมัธยมศึกษาปีที่ 1 เพื่อกำหนดเรื่องที่จะให้อ่าน

2. ศึกษาวิธีการสร้างแบบสอบวัดความสามารถในการอ่านภาษาไทย ตามหลักการสร้างข้อสอบวัด Cognitive Domain เรื่อง ข้อสอบวัดความเข้าใจ ของ Benjamin S. Bloom (1956, quoted in Linn and Gronlund, 1995)

3. การสร้างข้อสอบ

ผู้วิจัยได้ดำเนินการดังนี้

3.1 เรื่องที่ให้อ่านพิจารณาจากเอกสารสำคัญ ๆ ดังต่อไปนี้

(1) หลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) กระทรวงศึกษาธิการ

(2) หนังสือเรียนภาษาไทย ท 101 ท 102 ชุดทักษะสัมพันธ์ เล่ม 1 ชั้นมัธยมศึกษาปีที่ 1 กรมวิชาการ กระทรวงศึกษาธิการ

(3) หนังสือเรียนภาษาไทย ท 101 ท 102 หลักภาษาไทย เล่ม 1 ชั้นมัธยมศึกษาปีที่ 1 กรมวิชาการ กระทรวงศึกษาธิการ

(4) หนังสือคหกรรมศิลป์ สำหรับชั้นมัธยมศึกษาตอนต้น ของอมรรัตน์ เจริญชัย และ คูเพลิน สิริหงส์

(5) หนังสือเรื่องสารพันในบ้าน ของสมาคมคหกรรมศาสตร์แห่งประเทศไทย

(6) พจนานุกรม ฉบับราชบัณฑิตยสถาน พ.ศ. 2525

(7) หนังสือพิมพ์มติชนสุดสัปดาห์ ฉบับวันที่ 12 ธันวาคม 2538 และฉบับวันที่ 7 พฤษภาคม 2539

(8) หนังสือพิมพ์มติชนรายวัน ฉบับวันที่ 7 กรกฎาคม 2539 และวันที่ 8 กรกฎาคม 2539

3.2 ลักษณะของแบบสอบ

3.2.1 แบบสอบประกอบด้วย ข้อสอบเลือกตอบ 4 ตัวเลือก จำนวน 150 ข้อ

3.2.2 เรื่องที่ให้อ่าน ได้แก่ ข้อความ บทความ บทประพันธ์ บทสนทนา ที่มีเนื้อหาสาระยากง่ายเหมาะสมกับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง 2533) กระทรวงศึกษาธิการ

3.2.3 ข้อสอบแต่ละข้อวัดความสามารถในการอ่านภาษาไทย หมายถึง วัดความสามารถในการจับใจความของเรื่องที่ให้อ่าน การวัดแยกย่อยออกไปอีก 3 ระดับ ดังนี้

1. *การแปลความ* เป็นข้อสอบที่วัดความสามารถในการแปลความหมายของเรื่องที่ให้อ่านจากรูปแบบหนึ่งไปยังอีกรูปแบบหนึ่ง (คำศัพท์ ข้อความ ตัวเลข หรือ สัญลักษณ์)

ลักษณะของคำถาม เป็นดังนี้

(1) แปลความหมายของคำศัพท์เทคนิคหรือวิชาการ ข้อความ ส่วนวน คำพังเพย หรือข้อคิด ให้เป็นภาษาสามัญทั่ว ๆ ไป หรือ ในทางกลับกัน

(2) พิจารณาความสอดคล้องของเรื่องที่ให้อ่านหรือข้อความในเรื่องที่ให้อ่านกับ ข้อความ ส่วนวน ข้อคิด หรือคำพังเพย

2. *การตีความ* เป็นข้อสอบที่วัดความสามารถในการอธิบาย หรือ สรุปเรื่องที่อ่านว่า เป็นอย่างไร

ลักษณะของคำถาม เป็นดังนี้

(1) เรื่องที่ให้อ่านนำเสนอในลักษณะใด มีจุดมุ่งหมายเพื่ออะไร

(2) เรื่องที่ให้อ่านมีใจความสำคัญอย่างไร เน้นในสิ่งใด เกี่ยวข้องหรือคล้ายกับสิ่งใด ให้ข้อคิดหรือข้อควรปฏิบัติอย่างไร หรือควรตั้งชื่อเรื่องที่อ่านว่าอย่างไร

(3) ผู้เขียนหรือผู้พูดเรื่องที่ให้อ่านเป็นบุคคลเช่นไร มีความรู้สึกอย่างไร น่าจะเป็นใครหรือประกอบอาชีพใด กล่าวกับใคร หรือกล่าวถึงบุคคลใด

(4) ผู้ที่ถูกกล่าวถึงในเรื่องที่ให้อ่านเป็นใคร หรือเป็นบุคคลเช่นไร มีความสัมพันธ์กันอย่างไร

(5) คำหรือข้อความในเรื่องที่ให้อ่านหมายถึงใคร สิ่งใด ใช้ในสถานการณ์ใด

(6) ผู้ปฏิบัติตนตามเรื่องที่ให้อ่าน ควรเป็นบุคคลเช่นไร

(7) เรื่องที่ให้อ่านเกิดขึ้นเมื่อใด ที่ใด ในสถานการณ์ใด

(8) ตัวบุคคลในเรื่องที่ให้อ่านเป็นใคร เป็นบุคคลเช่นไร มีความรู้สึกอย่างไร

3. *การขยายความ* เป็นข้อความวัดความสามารถในการประมาณแนวโน้มของเรื่องที่ให้อ่านว่าเป็นอย่างไร โดยทำนายผลหรือสิ่งที่อาจเกิดขึ้นตามมา

ลักษณะของคำถาม เป็นดังนี้

(1) ถ้าปฏิบัติตามเรื่องที่ให้อ่านหรือข้อความในเรื่องที่ให้อ่าน ผลจะเป็นอย่างไร

(2) ถ้าไม่ปฏิบัติตามเรื่องที่ให้อ่านหรือข้อความในเรื่องที่ให้อ่าน ผลจะเป็นอย่างไร

3.3 ลักษณะของเรื่องที่ให้อ่าน

ผู้วิจัยได้แบ่งลักษณะของเรื่องที่ให้อ่านเป็น 2 ลักษณะ ดังนี้

ลักษณะที่ 1 มีเนื้อหาสาระทั่ว ๆ ไป สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 จำนวน 21 เรื่อง เรื่องหนึ่ง ๆ มีจำนวนข้อสอบระหว่าง 3 ถึง 10 ข้อ รวมจำนวนข้อสอบ 120 ข้อ

ลักษณะที่ 2 มีเนื้อหาสาระเกี่ยวข้องกับเพศหญิงมากกว่าเพศชาย จำนวน 3 เรื่อง เรื่องหนึ่ง ๆ มีจำนวนข้อสอบระหว่าง 7 ถึง 12 ข้อ รวมจำนวน 30 ข้อ

ส่วนการกำหนดเรื่องที่ให้อ่านในลักษณะนี้ ผู้วิจัยใช้หลักการเทียบเคียงกับแนวคิดของ Ironson (1982) และ Tittle (1982)

4. นำแบบสอบที่สร้างขึ้นไปให้ผู้เชี่ยวชาญพิจารณา ซึ่งจำแนกผู้เชี่ยวชาญเป็น 2 กลุ่มได้แก่

4.1 ผู้เชี่ยวชาญทางด้านเนื้อหา ได้แก่ อาจารย์ผู้สอนวิชาภาษาไทย ชั้นมัธยมศึกษาปีที่ 1 จำนวน 3 คน โดยมีหลักเกณฑ์การคัดเลือก ดังนี้

(1) จบการศึกษาระดับปริญญาโท สาขาวิชาการสอนภาษาไทยหรือสาขาวิชาที่สัมพันธ์กัน

(2) มีประสบการณ์การสอนภาษาไทยมาแล้ว ตั้งแต่ 10 ปีขึ้นไป

(3) มีประสบการณ์การสอนภาษาไทย ชั้นมัธยมศึกษาปีที่ 1 มาแล้ว ตั้งแต่ 3 ปีขึ้นไป

4.2 ผู้เชี่ยวชาญทางด้านกรอ่าน จำนวน 1 คน ได้แก่ อาจารย์มหาวิทยาลัยที่สอนวิชาหลักสูตรและการสอนภาษาไทย และจบการศึกษาทางด้านกรอ่าน

รายนามผู้เชี่ยวชาญ ปรากฏในภาคผนวก ก

ผู้วิจัยให้ผู้เชี่ยวชาญพิจารณาแบบสอบ ในประเด็นต่อไปนี้

(1) เรื่องที่ให้อ่านมีเนื้อหาสาระยากง่ายเหมาะสม กับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 หรือไม่ เพราะเหตุใด

(2) ข้อสอบสอดคล้องกับเรื่องที่ให้อ่าน หรือไม่

ตัวอย่าง การพิจารณาความเหมาะสมของเรื่องที่ให้อ่านและความสอดคล้องระหว่าง
ข้อสอบกับเรื่องที่ให้อ่าน

เรื่องที่ให้อ่าน

ผู้เชี่ยวชาญพิจารณา

“เริ่มเฮ้ย เริ่มชีวิต
ประจวบคองอาทิตย์เริ่มใจแสง
รุ่งอรุณแจ่มใสไม่ร้อนแรง
สีเหลืองแดงเด่นหล้าภาพธรรม
เปรียบวัยเยาว์ยังไม่มีราศีคละ
ชีวิตจะเจริญได้ดังใฝ่ฝัน
ก็ด้วยการศึกษาคีเสริมชีวิต
ให้เกิดปัญญาสว่างนำทางเฮย”

เรื่องที่ให้อ่านมีเนื้อหาสาระยากง่าย
เหมาะสมกับนักเรียน ม. 1 หรือไม่

เหมาะสม	ไม่เหมาะสม	เหตุผล

ข้อคำถาม

1. คำว่า “เริ่มใจแสง” หมายความว่า
อย่างไร
- ก. กำลังตกคิน
- ข. กำลังเที่ยงวัน
- (ค) กำลังขึ้นท้องฟ้า
- ง. กำลังเกิดสุริยุปราคา

ข้อสอบสอดคล้องกับเรื่องที่ให้อ่านหรือไม่

สอดคล้อง(1)	ไม่แน่ใจ (0)	ไม่สอดคล้อง(-1)

ผู้เชี่ยวชาญได้พิจารณาความเหมาะสมของเรื่องที่ให้อ่านกับนักเรียนในระดับชั้นมัธยมศึกษา
ปีที่ 1 แล้ว ได้ตัดเรื่องที่ให้อ่านในลักษณะที่ 1 ออกจำนวน 3 เรื่อง จึงเหลือเรื่องที่ให้อ่าน
ในลักษณะที่หนึ่ง 19 เรื่อง จำนวนข้อสอบ 100 ข้อ และได้ตัดเรื่องที่ให้อ่านในลักษณะที่ 2
ออกจำนวน 1 เรื่อง จึงเหลือเรื่องที่ให้อ่านในลักษณะที่สอง 2 เรื่อง จำนวนข้อสอบ 19 ข้อ
รายละเอียดผลการพิจารณาของผู้เชี่ยวชาญ ปรากฏดังตารางที่ 5

ตารางที่ 5 สรุปความเห็นของผู้เชี่ยวชาญในประเด็นเรื่องที่ให้อ่านมีเนื้อหาสาระ
เหมาะสมกับนักเรียนชั้นมัธยมศึกษาปีที่ 1

เรื่องที่ให้อ่าน	เนื้อหาสาระยากง่ายเหมาะสมกับนักเรียน ม. 1	
	เหมาะสม	ไม่เหมาะสม
ลักษณะที่ 1 เนื้อหาสาระทั่วไป จำนวน 21 เรื่อง		
1. คนดี (3)	/	
2. การเก็บเล็กผสมน้อย (4)	/	
3. ปัญหา (5)	/	
4. การลงแขก (4)		/
5. ลางเนื้อกลางยา (5)	/	
6. ปรีนิพพาน (6)	/	
7. ความพยายาม (5)	/	
8. การทะเลาะวิวาท (5)	/	
9. อุปสรรค (5)	/	
10. ความห่างไกล (4)	/	
11. คนคุยโว (5)	/	
12. กลองไทย (4)	/	
13. การใช้จ่าย (6)	/	
14. รักของแม่ (10)	/	
15. การชุบชีวิต (4)	/	
16. การแต่งกาย (6)	/	
17. การดำเนินชีวิต (6)		/
18. หน้าที่ (9)	/	
19. เริ่มชีวิต (8)	/	

ตารางที่ 5 (ต่อ) สรุปความเห็นของผู้เชี่ยวชาญในประเด็นเรื่องที่ให้อ่านมีเนื้อหาสาระ
 ง่ายเหมาะสมกับนักเรียนชั้นมัธยมศึกษาปีที่ 1

เรื่องที่ให้อ่าน	เนื้อหาสาระง่ายเหมาะสมกับนักเรียน ม. 1	
	เหมาะสม	ไม่เหมาะสม
20. เส้นพริก (6)	/	
21. สิทธิสตรี (10)		/
ลักษณะที่ 2 เนื้อหาสาระเกี่ยวข้องกับเพศหญิงมากกว่าเพศหญิง จำนวน 3 เรื่อง		
22. ฝรั่งผลไม้ (7)	/	
23. ขนมหิน-น้ำยาไก่ (12)	/	
24. การแต่งตัว (11)		/

ตัวเลขในวงเล็บ คือ จำนวนข้อสอบในแต่ละเรื่อง

หลักเกณฑ์การพิจารณาความเหมาะสมของเรื่องที่ให้อ่าน คือ ผู้เชี่ยวชาญ จำนวน 3 ใน 4 คน มีความเห็นร่วมกันว่า เรื่องที่ให้อ่านมีเนื้อหาสาระง่ายเหมาะสมกับนักเรียนชั้นมัธยมศึกษาปีที่ 1

ผู้เชี่ยวชาญได้พิจารณาความสอดคล้องระหว่างข้อสอบกับเรื่องที่ให้อ่าน โดยพิจารณาตัดข้อสอบในลักษณะที่ 1 ออกจำนวน 9 ข้อ เหลือข้อสอบในลักษณะที่ 1 จำนวน 91 ข้อและรวมกับข้อสอบในลักษณะที่ 2 อีกจำนวน 19 ข้อ รวมเหลือข้อสอบ 110 ข้อ รวมทั้งได้ปรับปรุงแก้ไขเนื้อหาสาระของข้อสอบให้เหมาะสมยิ่งขึ้น รายละเอียดผลการพิจารณาเฉพาะข้อสอบที่นำไปใช้ในการวิจัย จำนวน 75 ข้อ ปรากฏดังตารางที่ 6

ตารางที่ 6 สรุปความเห็นของผู้เชี่ยวชาญ ในประเด็นความสอดคล้องระหว่างข้อสอบกับเรื่อง
ที่ให้อ่าน

เรื่องที่ให้อ่าน	ความสามารถในการอ่าน		ดัชนีความสอดคล้องระหว่าง ข้อสอบกับเรื่องที่ให้อ่าน
	แปลความ	ตีความ ขยายความ	
1. เริ่มชีวิต (4)			
ข้อที่ 1	/		1.00
2		/	1.00
3	/		1.00
4		/	1.00
2. เส้นพวง (4)			
ข้อที่ 5		/	1.00
6		/	.75
7	/		1.00
8		/	1.00
3. ขนมหิน-น้ำยาไก่ (9)			
ข้อที่ 9	/		1.00
10	/		1.00
11	/		1.00
12	/		1.00
13	/		1.00
14	/		1.00
15		/	1.00
16	/		1.00
17	/		1.00

ตารางที่ 6 (ต่อ) สรุปความเห็นของผู้เชี่ยวชาญ ในประเด็นความสอดคล้องระหว่างข้อสอบ
กับ เรื่องที่ให้อ่าน

เรื่องที่ให้อ่าน	ความสามารถในการอ่าน			ดัชนีความสอดคล้องระหว่าง ข้อสอบกับเรื่องที่ให้อ่าน
	แปลความ	ตีความ	ขยายความ	
4. ปรินิพพาน (6)				
ข้อที่ 18	/			1.00
19	/			1.00
20	/			1.00
21	/			1.00
22	/			1.00
23	/			1.00
5. ความพยายาม (4)				
ข้อที่ 24	/			.75
25	/			1.00
26	/			1.00
27	/			1.00
6. การทะเลาะวิวาท (4)				
ข้อที่ 28	/			.75
29	/			1.00
30	/			1.00
31	/			1.00
7. อุปสรรค (2)				
ข้อที่ 32	/			1.00
33	/			1.00

ตารางที่ 6 (ต่อ) สรุปความเห็นของผู้เชี่ยวชาญ ในประเด็นความสอดคล้องระหว่างข้อสอบ
กับ เรื่องที่ให้อ่าน

เรื่องที่ให้อ่าน	ความสามารถในการอ่าน			ดัชนีความสอดคล้องระหว่าง ข้อสอบกับเรื่องที่ให้อ่าน
	แปลความ	ตีความ	ขยายความ	
8. หน้าที่ (5)				
ข้อที่ 34	/			1.00
35	/			1.00
36	/			1.00
37	/			1.00
38	/			1.00
9. ความห่างไกล (3)				
ข้อที่ 39	/			1.00
40	/			1.00
41	/			1.00
10. คนคุยโว (4)				
ข้อที่ 42	/			1.00
43	/			1.00
44	/			1.00
45	/			1.00
11. กลองไทย (3)				
ข้อที่ 46	/			1.00
47	/			1.00
48	/			1.00

ตารางที่ 6 (ต่อ) สรุปความเห็นของผู้เชี่ยวชาญ ในประเด็นความสอดคล้องระหว่างข้อสอบ
กับ เรื่องที่ให้อ่าน

เรื่องที่ให้อ่าน	ความสามารถในการอ่าน			ดัชนีความสอดคล้องระหว่าง ข้อสอบกับเรื่องที่ให้อ่าน
	แปลความ	ตีความ	ขยายความ	
12. การใช้ง่าย (4)				
ข้อที่ 49	/			1.00
50	/			1.00
51	/			.75
52	/			1.00
13. รักของแม่ (7)				
ข้อที่ 53	/			1.00
54	/			1.00
55	/			1.00
56	/			1.00
57	/			1.00
58	/			1.00
59	/			.75
14. การรูปชีวิต (3)				
ข้อที่ 60	/			1.00
61	/			1.00
62	/			1.00
15. การแต่งกาย (4)				
ข้อที่ 63	/			1.00
64	/			1.00
65	/			.75

ตารางที่ 6 (ต่อ) สรุปความเห็นของผู้เชี่ยวชาญ ในประเด็นความสอดคล้องระหว่างข้อสอบ
กับ เรื่องที่ให้อ่าน

เรื่องที่ให้อ่าน	ความสามารถในการอ่าน			ดัชนีความสอดคล้องระหว่าง ข้อสอบกับเรื่องที่ให้อ่าน
	แปลความ	ตีความ	ขยายความ	
ข้อที่ 66	/			1.00
16. ฝรั่งผลไม้ (6)				
ข้อที่ 67	/			1.00
68	/			1.00
69			/	.75
70	/			1.00
71	/			1.00
72			/	1.00
17. การเก็บเล็กผสมน้อย (3)				
ข้อที่ 73	/			1.00
74	/			1.00
75	/			1.00

ตัวเลขในวงเล็บ คือจำนวนข้อสอบในแต่ละเรื่อง

เกณฑ์การพิจารณาความสอดคล้องระหว่างข้อสอบกับเรื่องที่ให้อ่าน หาได้ดังนี้

$$\text{ดัชนีความสอดคล้องระหว่างข้อสอบกับเรื่องที่ให้อ่าน} = \frac{\text{ผลรวมความเห็นของผู้เชี่ยวชาญ}}{\text{จำนวนผู้เชี่ยวชาญ}}$$

การแปลผล ถ้าดัชนีมีค่าตั้งแต่ .75 ขึ้นไป ถือว่าข้อสอบข้อนั้นมีความตรงตามเนื้อหา

5. นำแบบสอบที่ได้จากข้อที่ 4 ไปทดลองสอบครั้งที่ 1 กับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 1 ปีการศึกษา 2539 ของโรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตรศาสตร์ จำนวน 10 คน แบ่งเป็นนักเรียนชาย 5 คน ซึ่งประกอบด้วยนักเรียนเก่ง (1 คน) ปานกลาง (2 คน) อ่อน (2 คน) และนักเรียนหญิง 5 คน ซึ่งประกอบด้วย นักเรียนเก่ง (1 คน) ปานกลาง (2 คน) อ่อน (2 คน) โดยทดลองสอบในวันที่ 12 กันยายน 2539 เพื่อปรับปรุงความเหมาะสมทางด้านภาษาของข้อสอบ

6. นำแบบสอบที่ได้ปรับปรุงแล้วจากข้อที่ 5 ไปทดลองสอบครั้งที่ 2 กับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2539 ของโรงเรียนบ้านค่าย อำเภอบ้านค่าย จังหวัดระยอง จำนวน 100 คน จำแนกเป็นนักเรียนชาย 50 คน และนักเรียนหญิง 50 คน ในวันที่ 4 พฤศจิกายน 2539 เพื่อหาคุณภาพของข้อสอบรายข้อ (ค่าความยากและค่าอำนาจจำแนกของข้อสอบ) ตามทฤษฎีการสอบแบบดั้งเดิม แล้วคัดเลือกข้อสอบไว้จำนวน 75 ข้อ ตามเกณฑ์การคัดเลือกข้อสอบ ดังนี้

6.1 คัดเลือกข้อสอบที่มีค่าความยาก (p) อยู่ระหว่าง .20 - .80

6.2 คัดเลือกข้อสอบที่มีค่าอำนาจจำแนก (r) ตั้งแต่ .20 ขึ้นไป

รายละเอียดของผลการวิเคราะห์ข้อสอบ ปรากฏในภาคผนวก ข

ลักษณะของข้อสอบที่นำไปใช้ในการวิจัย จำนวน 75 ข้อ จำแนกตามเรื่องที่ให้อ่าน 17 เรื่อง ดังตารางที่ 7

ตารางที่ 7 ลักษณะของข้อสอบภาษาไทย ชั้นมัธยมศึกษาปีที่ 1 จำนวนตามเรื่องที่ให้อ่าน

เรื่องที่ให้อ่าน	จำนวนข้อ	ข้อที่
ลักษณะที่ 1 เนื้อหาสาระทั่วไป	จำนวน 60 ข้อ (80 %)	
1 เริ่มชีวิต	4	1 2 3 4
2 เส้นพวงรัก	4	5 6 7 8
3 ปรีนิทพาน	6	18 19 20 21 22 23
4 ความพยายาม	4	24 25 26 27
5 การทะเลาะวิวาท	4	28 29 30 31
6 อุปสรรค	2	32 33
7 หน้าที่	5	34 35 36 37 38
8 ความห่างไกล	3	39 40 41
9 คนคุยโว	4	42 43 44 45
10 กลองไทย	3	46 47 48
11 การใช้ง่าย	4	49 50 51 52
12 รักของแม่	7	53 54 55 56 57 58 59
13 การชุบชีวิต	3	60 61 62
14 การแต่งกาย	4	63 64 65 66
15 การเก็บเล็กผสมน้อย	3	73 74 75
ลักษณะที่ 2 เนื้อหาสาระเกี่ยวข้องกับเพศหญิงมากกว่าชาย	จำนวน 15 ข้อ (20 %)	
16 ขนมหิน-น้ำยาไก่อ	9	9 10 11 12 13 14 15 16 17
17 วันผลไม้	6	67 68 69 70 71 72
รวม	75	

ตัวอย่างข้อสอบ

คำชี้แจง จากบทประพันธ์นี้ จงตอบคำถามข้อ 1 - 4

เริ่มเอ๋ย เริ่มชีวิต	ประจวบดวงอาทิตย์เริ่มไยแสง
รุ่งอรุณแจ่มใสไม่ร้อนแรง	สีเหลืองแดงเด่นหล้าภาพธรรม
เปรียบวัยเยาว์ยังไม่มีราศีคละ	ชีวิตจะเจริญได้ดังไฟฝัน
ก็ด้วยการศึกษาดีเสริมชีวิต	ให้เกิดปัญญาสว่างนำทางเอย

- (1) คำว่า “เริ่มไยแสง” หมายความว่าอย่างไร
 - ก. กำลังตกดิน
 - ข. กำลังเที่ยงวัน
 - ✓ ค. กำลังขึ้นท้องฟ้า
 - ง. กำลังเกิดสุริยุปราคา
- (2) ผู้แต่งบทประพันธ์ต้องการเน้นความสำคัญของสิ่งใด
 - ก. วัยเด็กเป็นวัยที่บริสุทธิ์
 - ข. ชีวิตจะเจริญได้ด้วยความหวัง
 - ✓ ค. ชีวิตจะเจริญก้าวหน้าด้วยการศึกษา
 - ง. ชีวิตเปรียบเหมือนพระอาทิตย์แรกขึ้น
- (3) ข้อความว่า “ให้เกิดปัญญาสว่างนำทางเอย” คำที่ขีดเส้นใต้หมายความว่าอย่างไร
 - ก. แนวทางการแก้ปัญหาสังคม
 - ✓ ข. แนวทางการดำเนินชีวิตที่ถูกต้อง
 - ค. แนวทางการช่วยเหลือเพื่อนมนุษย์
 - ง. แนวทางการเอาชนะอุปสรรคของชีวิต
- (4) บทประพันธ์นี้เน้นให้เห็นคุณค่าของอะไร
 - ✓ ก. การเรียนรู้
 - ข. การทำมาหากิน
 - ค. การเลี้ยงดูเด็ก
 - ง. การดำเนินชีวิต

รายละเอียดของข้อสอบ ปรากฏในภาคผนวก ค

7. นำแบบสอบในข้อ 6 ไปทดลองสอบครั้งที่ 3 กับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2539 ของโรงเรียนบ้านค่าย อำเภอบ้านค่าย จังหวัดระยอง จำนวน 100 คน จำแนกเป็นนักเรียนชาย 50 คน และนักเรียนหญิง 50 คน ในวันที่ 8 พฤศจิกายน 2539 เพื่อหาค่าความเที่ยงของแบบสอบด้วยสูตร KR - 20 ของคูเคอร์ริชาร์คสัน โดยใช้โปรแกรมคอมพิวเตอร์ CTIA/GRADING Ver 6.30 (สุวัฒน์ สุขมลสันต์, 2539) ได้ค่าความเที่ยงของแบบสอบ เท่ากับ 0.90

8. นำแบบสอบตามข้อ 7 ไปจัดพิมพ์เป็นรูปเล่มและจัดทำคู่มือดำเนินการสอบเพื่อนำไปใช้เก็บรวบรวมข้อมูลต่อไป

การเก็บรวบรวมข้อมูล

ผู้วิจัยได้ดำเนินการ ดังนี้

1. ติดต่อประสานงานกับโรงเรียนที่เป็นกลุ่มตัวอย่าง เพื่อกำหนดวัน เวลา ในการนำแบบสอบวัดความสามารถในการอ่านภาษาไทย ไปทดสอบกับนักเรียนที่เป็นกลุ่มตัวอย่าง

2. ชี้แจงทำความเข้าใจกับผู้ประสานงานซึ่งอยู่ที่โรงเรียน อาจารย์ที่คุมห้องสอบ และนิสิตคณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา ซึ่งช่วยเก็บข้อมูล ในเรื่องลักษณะของแบบสอบวัดความสามารถในการอ่านภาษาไทย และวิธีดำเนินการสอบ

3. นำแบบสอบวัดความสามารถในการอ่านภาษาไทย ไปทดสอบกับนักเรียนที่เป็นกลุ่มตัวอย่าง ในช่วงวันที่ 11 ถึง 22 พฤศจิกายน 2539

4. นำผลการตอบข้อสอบในแบบสอบวัดความสามารถในการอ่านภาษาไทย มาตรวจดูความเรียบร้อย แล้วลงรหัสผลการตอบข้อสอบทุกข้อ (ตอบถูกได้ 1 และ ตอบผิดได้ 0) ของนักเรียนที่เป็นกลุ่มตัวอย่างทุกคน ในกระดาษแบบการลงรหัส

5. การจัดทำแฟ้มข้อมูล

5.1 การจัดทำแฟ้มข้อมูล ของกลุ่มตัวอย่าง 1,200 คน จำนวน 1 แฟ้ม ตามรูปแบบการจัดเตรียมข้อมูลสำหรับการวิเคราะห์ข้อมูล ตามทฤษฎีการตอบสนองข้อสอบ แบบ 2 พารามิเตอร์ ด้วยโปรแกรมคอมพิวเตอร์ BILOG Ver 3.04 ทั้งนี้เพื่อนำค่าประมาณความสามารถของผู้สอบและค่าประมาณความยากของข้อสอบ ไปใช้แบ่งผู้สอบและข้อสอบ ตามวิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ

5.2 การจัดทำเพิ่มข้อมูล เพื่อวิเคราะห์การทำหน้าที่ต่างกันแบบไม่สม่ำเสมอของข้อสอบ ตามวิธี IRT area แบบ 2 พารามิเตอร์ ด้วยโปรแกรมคอมพิวเตอร์ BILOG Ver 3.04 , โปรแกรมคอมพิวเตอร์ EQUATE และ โปรแกรมคอมพิวเตอร์ IRTDIF ดังนี้

5.2.1 นำผลการตอบข้อสอบทุกข้อ (ตอบถูกได้ 1 คะแนน และตอบผิด 0 คะแนน) ของผู้สอบกลุ่มสนใจทุกคน มาจัดทำเพิ่มข้อมูลเป็นรายคน จำนวน 1 เพิ่ม

5.2.2 นำผลการตอบข้อสอบทุกข้อ (ตอบถูกได้ 1 คะแนน และผิดได้ 0 คะแนน) ของผู้สอบกลุ่มอ้างอิงทุกคน มาจัดทำเพิ่มข้อมูลเป็นรายคน จำนวน 1 เพิ่ม

5.3 การจัดทำเพิ่มข้อมูล เพื่อวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ ตามวิธี แมนเทิล-แฮนส์เชล แบบปกติ จำนวน 1 เพิ่ม

5.4 การจัดทำเพิ่มข้อมูล เพื่อวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ ตามวิธี แมนเทิล-แฮนส์เชล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ดำเนินการ ดังนี้

5.4.1 นำค่าความสามารถของผู้สอบทุกคนที่คำนวณได้ตามข้อ 5.1 มาหาค่าความสามารถเฉลี่ยของผู้สอบ ได้เท่ากับ 0.0016

5.4.2 แบ่งผู้สอบกลุ่มสนใจ ออกเป็นกลุ่มผู้สอบที่มีความสามารถสูง (ค่าความสามารถของแต่ละคน สูงกว่าหรือเท่ากับ 0.0016) และกลุ่มผู้สอบที่มีความสามารถต่ำ (ค่าความสามารถของแต่ละคน ต่ำกว่า 0.0016)

5.4.3 แบ่งผู้สอบกลุ่มอ้างอิง ออกเป็นกลุ่มผู้สอบที่มีความสามารถสูง กับกลุ่มผู้สอบที่มีความสามารถต่ำ

5.4.4 นำผลการตอบข้อสอบของกลุ่มผู้สอบที่มีความสามารถสูงทั้งของกลุ่มสนใจ และกลุ่มอ้างอิง ในข้อ 5.4.2 และ 5.4.3 มาจัดทำเพิ่มข้อมูล จำนวน 1 เพิ่ม

5.4.5 นำผลการตอบข้อสอบของกลุ่มผู้สอบที่มีความสามารถต่ำ ทั้งของกลุ่มสนใจและกลุ่มอ้างอิง ในข้อที่ 5.4.2 และ 5.4.3 มาจัดทำเพิ่มข้อมูล จำนวน 1 เพิ่ม

5.4.6 นำค่าความยากของข้อสอบที่คำนวณได้ตามข้อ 5.1 มาแบ่งข้อสอบทั้งหมด ออกเป็นกลุ่มข้อสอบยาก ($b \geq 1.00$ ข้อ) ข้อสอบยากง่ายปานกลาง ($-1.00 < b < 1.00$) และ ข้อสอบง่าย ($b \leq -1.00$)

5.4.7 นำเพิ่มข้อมูลกลุ่มผู้สอบที่มีความสามารถสูง ในข้อ 5.4.4 มาแบ่งตามกลุ่มข้อสอบ (ข้อ 5.4.6) ได้เพิ่มข้อมูลอีก 3 เพิ่ม ได้แก่ 1. เพิ่มข้อมูลกลุ่มผู้สอบที่มีความ

สามารถสูงและเป็นข้อสอบยาก 2. เพิ่มข้อมูลกลุ่มผู้สอบที่มีความสามารถสูงและเป็นข้อสอบ
ยากง่ายปานกลาง 3. เพิ่มข้อมูลกลุ่มผู้สอบที่มีความสามารถสูงและเป็นข้อสอบง่าย

5.4.8 นำเพิ่มข้อมูลกลุ่มผู้สอบที่มีความสามารถต่ำ ในข้อ 5.4.5 มาแบ่งตาม
กลุ่มข้อสอบ (ข้อ 5.4.6) ได้เพิ่มข้อมูลอีก 3 เพิ่ม ได้แก่ 1. เพิ่มข้อมูลกลุ่มผู้สอบที่มี
ความสามารถต่ำและเป็นข้อสอบยาก 2. เพิ่มข้อมูลกลุ่มผู้สอบที่มีความสามารถต่ำและเป็น
ข้อสอบยากง่ายปานกลาง 3. เพิ่มข้อมูลกลุ่มผู้สอบที่มีความสามารถต่ำและเป็นข้อสอบง่าย

5.5 การจัดทำเพิ่มข้อมูล เพื่อวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ ตามวิธี
แมนเทิล-แฮนส์เซล แบบปกติ กับ แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของ
ข้อสอบ จากการจำลองกลุ่มตัวอย่างจากข้อมูลจริง จำนวน 10 ครั้ง ๆ ละ 9 เพิ่มข้อมูล
รวมจำนวน 90 เพิ่ม ดำเนินการเช่นเดียวกับข้อ 5.3 และ ข้อ 5.4.2 - 5.4.5, 5.4.7, 5.4.8

6. นำเพิ่มข้อมูลที่จัดเตรียมไว้ ไปดำเนินการวิเคราะห์ต่อไป

โปรแกรมคอมพิวเตอร์ที่ใช้ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

ผู้วิจัยได้ติดตามความก้าวหน้าทางด้านโปรแกรมคอมพิวเตอร์ที่ใช้ตรวจสอบการทำหน้าที่
ต่างกันของข้อสอบ พบว่า ได้มีผู้ทรงคุณวุฒิในสหรัฐอเมริกาได้จัดทำโปรแกรมคอมพิวเตอร์ไว้
แล้ว แต่ในเมืองไทยยังไม่มีโปรแกรมคอมพิวเตอร์เหล่านี้ จึงได้ดำเนินการดังนี้

1. โปรแกรมคอมพิวเตอร์ IRTDIP

จากการติดตามความก้าวหน้าทางด้านโปรแกรมคอมพิวเตอร์ที่ใช้ตรวจสอบการทำหน้าที่
ต่างกันของข้อสอบ ตามวิธี IRT area พบว่า Kim และ Cohen (1992) แห่ง Testing
and Evaluation Services, University of Wisconsin, Madison ได้จัดทำโปรแกรมคอมพิวเตอร์ไว้
แล้ว ผู้วิจัยจึงได้ติดต่อขอความอนุเคราะห์โปรแกรมคอมพิวเตอร์ IRTDIP จาก Dr. Allan
S. Cohen โดยตรง ซึ่งก็ได้รับความอนุเคราะห์โปรแกรมคอมพิวเตอร์ ดังกล่าว

2. โปรแกรมคอมพิวเตอร์ MHDIP

จากการติดตามความก้าวหน้าทางด้านโปรแกรมคอมพิวเตอร์ที่ใช้ตรวจสอบการทำหน้าที่
ต่างกันของข้อสอบ ตามวิธีแมนเทิล-แฮนส์เซล แบบ 2 ขั้นตอน พบว่า Hambleton และ
Rogers (1994) แห่ง Laboratory of Psychometric and Evaluative, University of
Massachusetts, Amherst ได้จัดทำโปรแกรมคอมพิวเตอร์ไว้แล้ว จึงได้ติดต่อขอซื้อโปรแกรม

คอมพิวเตอร์ MHDIF จาก Dr. Ronald K. Hambleton ในราคา 85 เหรียญสหรัฐ (ราคาในสหรัฐอเมริกา) ซึ่งก็ได้รับโปรแกรมคอมพิวเตอร์ดังกล่าว

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลแบ่งออกเป็น 5 ตอน คือ ตอนที่ 1 การวิเคราะห์ข้อมูลเบื้องต้น เพื่อประมาณค่าความสามารถของผู้สอบและความยากของข้อสอบ ตามทฤษฎีการตอบสนองข้อสอบ แบบ 2 พารามิเตอร์ ซึ่งจะนำไปใช้แบ่งกลุ่มผู้สอบและกลุ่มข้อสอบ ในตอนที่ 3.2 ตอนที่ 2 เป็นการวิเคราะห์หาข้อสอบทำหน้าที่ย่างกันแบบไม่สม่ำเสมอ ตามวิธี IRT area แบบ 2 พารามิเตอร์ ตอนที่ 3 เป็นการวิเคราะห์หาข้อสอบทำหน้าที่ย่างกันแบบไม่สม่ำเสมอ ตามวิธีแมนเทิล-แฮนส์เซล แบบปกติ กับ แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ตอนที่ 4 เป็นการเปรียบเทียบผลการตรวจพบข้อสอบทำหน้าที่ย่างกันแบบไม่สม่ำเสมอ ระหว่างวิธีแมนเทิล-แฮนส์เซล แบบปกติ กับ แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ตอนที่ 5 เป็นการวิเคราะห์หาข้อสอบทำหน้าที่ย่างกันแบบไม่สม่ำเสมอ ตามวิธีแมนเทิล-แฮนส์เซล แบบปกติ กับ แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ จากการจำลองกลุ่มตัวอย่างจากข้อมูลจริง จำนวน 10 ครั้ง

ตอนที่ 1 การวิเคราะห์ข้อมูลเบื้องต้น

การวิเคราะห์ข้อสอบเพื่อประมาณค่าความสามารถของผู้สอบรายคน ($\hat{\theta}$) และค่าความยากของข้อสอบรายข้อ (\hat{b}) ของกลุ่มตัวอย่าง จำนวน 1,200 คน ตามทฤษฎีการตอบสนองข้อสอบ แบบ 2 พารามิเตอร์ โดยใช้โปรแกรมคอมพิวเตอร์ BILOG Ver 3.04 (Mislevy and Bock, 1990)

โปรแกรมคอมพิวเตอร์ BILOG Ver 3.04 เป็นโปรแกรมที่ใช้กับเครื่อง PC เขียนด้วยภาษา FORTRAN 77 เพื่อวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ ข้อมูลผลการตอบข้อสอบต้องเป็นแบบ 2 ประเภท (dichotomous) คือ ตอบถูกได้คะแนน 1 คะแนน และตอบผิดได้คะแนน 0 คะแนน การวิเคราะห์ข้อมูลใช้วิธีการประมาณแบบ marginal bayes estimation (MBE) แบ่งการวิเคราะห์เป็น 3 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 เป็นการแปลงข้อมูลดิบให้อยู่ในรูปของข้อมูลสำหรับวิเคราะห์ ในขั้นตอนที่ 2 โดยใช้โปรแกรมย่อยชื่อ INPUT

ขั้นตอนที่ 2 เป็นการประมาณค่าพารามิเตอร์ของข้อสอบ และทดสอบความเหมาะสมของข้อสอบกับโมเดล logistic โดยใช้โปรแกรมย่อยชื่อ CALIBRATE การประมาณค่าพารามิเตอร์ของข้อสอบใช้การคำนวณซ้ำทวน (iteration) ด้วยวิธี EM 5 รอบ และ ตามด้วย Newton-Gauss อีก 2 รอบ ส่วนการทดสอบความเหมาะสมของข้อสอบกับโมเดล logistic ใช้การทดสอบ χ^2 - Goodness of Fit index

ขั้นตอนที่ 3 เป็นการประมาณค่าความสามารถของผู้สอบ โดยใช้โปรแกรมย่อยชื่อ SCORE

ตอนที่ 2 การวิเคราะห์หาข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ตามวิธี IRT area

การวิเคราะห์หาจำนวนข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ในแบบสอบวัดความสามารถในการอ่านภาษาไทย ใช้วิธี IRT area แบบ 2 พารามิเตอร์ โดยมีวิธีการสำคัญ ๆ ดังนี้

2.1 คำนวณเมตริกซ์ค่าความแปรปรวนและความแปรปรวนร่วม (Variance - Covariance matrices) จากข้อมูลผลการตอบในแฟ้มข้อมูลกลุ่มสนใจ (เพศชาย) และกลุ่มอ้างอิง (เพศหญิง) โดยใช้โปรแกรมคอมพิวเตอร์ BILOG Ver 3.04 (Mislevy and Bock, 1990) ซึ่งนำผลการวิเคราะห์ไปใช้กับโปรแกรมคอมพิวเตอร์ EQUATE และ IRTDIF ต่อไป

2.2 เทียบมาตรของค่าพารามิเตอร์ของข้อสอบที่ประมาณได้จากกลุ่มสนใจกับกลุ่มอ้างอิง โดยใช้โปรแกรมคอมพิวเตอร์ EQUATE Ver 2.0 (Baker, Al-Karni, and Al-Dosary, 1991)

โปรแกรมคอมพิวเตอร์ EQUATE เป็นโปรแกรมที่ใช้กับเครื่อง PC เขียนด้วยภาษา FORTRAN 77 เพื่อแปลงค่าพารามิเตอร์ของข้อสอบกลุ่มสนใจให้อยู่บนมาตรฐานเดียวกับค่าพารามิเตอร์ของข้อสอบกลุ่มอ้างอิง ดังสมการ

$$a_2 = a_2^* / A \quad (1)$$

$$\text{Var} (a_2) = \text{Var} (a_2^*) / A^2 \quad (2)$$

$$b_2 = Ab_2^* + K \quad (3)$$

$$\text{Var} (b_2) = A^2 \text{Var} (a_2^*) \quad (4)$$

เมื่อ A แทน ค่าความชันของเส้นถดถอย
K แทน intercept

โปรแกรมจะใช้เมตริกซ์ค่าความแปรปรวนและความแปรปรวนร่วม ของกลุ่มสนใจ และกลุ่มอ้างอิง ที่คำนวณได้จากโปรแกรมคอมพิวเตอร์ BILOG Ver 3.04 เพื่อคำนวณหาค่า equating slope A และ equating intercept K ซึ่งจะนำไปใช้ในโปรแกรมคอมพิวเตอร์ IRTDIF ต่อไป

2.3 การคำนวณหาข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ตามวิธี IRT area แบบ 2 พารามิเตอร์ ของ Raju (1988 ; 1990) โดยใช้โปรแกรมคอมพิวเตอร์ IRTDIF Ver 1.0 (Kim and Cohen, 1992)

โปรแกรมคอมพิวเตอร์ IRTDIF Ver 1.0 เป็นโปรแกรมที่ใช้กับเครื่อง PC เขียนด้วยภาษา FORTRAN 77 เพื่อคำนวณการทำหน้าที่ต่างกันแบบไม่สม่ำเสมอของข้อสอบ โดยโปรแกรมใช้เมตริกซ์ค่าความแปรปรวนและความแปรปรวนร่วมของกลุ่มสนใจกับกลุ่มอ้างอิง ที่คำนวณได้จากข้อ 2.1 และใช้ค่า A กับค่า K ที่คำนวณได้จากข้อ 2.2 ไปคำนวณหาค่าพื้นที่ความแตกต่างระหว่างโค้งลักษณะข้อสอบของกลุ่มอ้างอิงกับกลุ่มสนใจ คือ ค่า Exact Unsigned Area Measure (EUA) ในกรณี $a_1 \neq a_2$ ดังสมการ

$$EUA = |H| \quad (1)$$

$$E(EUA) = \sqrt{(2/\pi) \{ \text{Var}(H) \}} \quad (2)$$

และ

$$\text{Var}(EUA) = (1 - 2/\pi) \{ \text{Var}(H) \} \quad (3)$$

เมื่อ

$$H = \frac{2(a_2 - a_1)}{Da_1a_2} \ln \left[1 + \exp \left\{ \frac{Da_1a_2(b_2 - b_1)}{a_2 - a_1} \right\} \right] - (b_2 - b_1) \quad (4)$$

และ

$$\begin{aligned} \text{Var}(H) = & A_1^2 \text{Var}(a_1) + A_2^2 \text{Var}(a_2) + B_1^2 \text{Var}(b_1) + B_2^2 \text{Var}(b_2) + \\ & 2A_1B_1 \text{Cov}(a_1b_1) + 2A_2B_2 \text{Cov}(a_2b_2) \end{aligned} \quad (5)$$

ค่านำหนักของ A_1 , A_2 , B_1 และ B_2 หาได้ดังนี้

$$A_1 = \frac{2}{Da_1^2} \left[\frac{Y \exp(Y)}{1 + \exp(Y)} - \ln \{1 + \exp(Y)\} \right] \quad (6)$$

$$A_2 = -\frac{a_1^2}{a_2^2} A_1 \quad (7)$$

$$B_1 = \frac{1 - \exp(Y)}{1 + \exp(Y)} \quad (8)$$

และ $B_2 = -B_1 \quad (9)$

เมื่อ $Y = \frac{Da_1 a_2 (b_2 - b_1)}{a_2 - a_1} \quad (10)$

การทดสอบนัยสำคัญทางสถิติ (Ruju, 1990) ให้ใช้สถิติทดสอบ ดังนี้

$$Z(H) = \frac{H}{\sqrt{\text{Var}(H)}} \quad (11)$$

ผู้วิจัยวิเคราะห์หาข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ในแบบสอบวัดความสามารถในการอ่านภาษาไทย โดยวิธี IRT area แบบ 2 พารามิเตอร์ เพื่อนำผลการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ไปใช้เป็นเกณฑ์สำหรับเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันแบบไม่สม่ำเสมอของข้อสอบ ระหว่างวิธีแมนเทิล-แฮนส์เชล แบบปกติ กับ วิธีแมนเทิล-แฮนส์เชล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ

ตอนที่ 8 การวิเคราะห์หาข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ตามวิธีแมนเทิล-แฮนส์เชล แบบปกติ กับ แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ

ผู้วิจัยทำการวิเคราะห์หาข้อสอบทำหน้าที่ต่างกัน ตามวิธีแมนเทิล-แฮนส์เชล โดยใช้โปรแกรมคอมพิวเตอร์ MHDIF (Hambleton and Rogers, 1994) ซึ่งเป็นโปรแกรมที่ใช้กับเครื่อง PC เขียนด้วยภาษา FORTRAN 77 เพื่อคำนวณค่าสถิติ MH- χ^2 และ MH D-DIF

ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่างกลุ่มผู้สอบ 2 กลุ่ม ข้อมูลผลการตอบข้อสอบต้องเป็นแบบ 2 ประเภท คือ ตอบถูกได้ 1 คะแนน ตอบผิดได้ 0 คะแนน ผู้สอบในแต่ละกลุ่มถูกจัดกลุ่มความสามารถ โดยใช้คะแนนรวมของแบบสอบเป็นเกณฑ์การจัดกลุ่มผู้สอบ 2 กลุ่ม โปรแกรมคำนวณค่าสถิติ MH แบบ 2 ขั้นตอน ดังนี้

ขั้นตอนที่หนึ่ง โปรแกรมคำนวณค่าสถิติ MH ของข้อสอบแต่ละข้อ ถ้าไม่พบข้อสอบทำหน้าที่ต่างกัน โปรแกรมจะหยุดทำงาน แต่ถ้าพบข้อสอบทำหน้าที่ต่างกัน โปรแกรมจะคำนวณคะแนนรวมของผู้สอบแต่ละคนอีกครั้งหนึ่ง โดยไม่นำเอาคะแนนจากข้อสอบทำหน้าที่ต่างกันเข้ามารวมด้วย ทั้งนี้เพื่อให้คะแนนรวมที่ใช้แทนความสามารถมีความบริสุทธิ์ขึ้น

ขั้นตอนที่สอง หลังจากรวมคะแนนใหม่แล้ว กลุ่มผู้สอบจะถูกจัดกลุ่มใหม่ตามคะแนนรวมที่ทำให้บริสุทธิ์ แล้วคำนวณค่าสถิติ MH ของข้อสอบแต่ละข้ออีกครั้งหนึ่ง ในกรณีที่ต้องการคำนวณค่าสถิติ MH ของข้อสอบข้อที่ทำหน้าที่ต่างกัน ตามที่ตรวจพบในขั้นตอนที่ 1 ต้องนำคะแนนของข้อสอบข้อนั้น รวมเข้าไว้ในคะแนนรวมของผู้สอบด้วย โดยที่การจัดกลุ่มผู้สอบไม่เปลี่ยนแปลง

3.1 วิเคราะห์หาข้อสอบทำหน้าที่ต่างกัน ตามวิธีแมนเทิล-แฮนส์เชล แบบปกติ (traditional MH procedure) ดังนี้

3.1.1 วิเคราะห์หาข้อสอบทำหน้าที่ต่างกัน ในแบบสอบทั้งฉบับ โดยใช้กลุ่มผู้สอบทั้งหมด

3.1.2 เกณฑ์ที่ใช้ตัดสินข้อสอบทำหน้าที่ต่างกัน คือ ค่าสถิติ $MH-\chi^2$ มีนัยสำคัญที่ระดับ .05

3.2 วิเคราะห์หาข้อสอบทำหน้าที่ต่างกัน ตามวิธีแมนเทิล-แฮนส์เชล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ดังนี้

3.2.1 แบ่งผู้สอบทั้งหมดโดยใช้ค่าความสามารถของผู้สอบจากกลุ่มตัวอย่าง 1,200 คน ซึ่งประมาณได้ตามทฤษฎีการตอบสนองข้อสอบ แบบ 2 พารามิเตอร์ ออกเป็น 2 กลุ่มย่อย ได้แก่ กลุ่มผู้สอบที่มีความสามารถสูง (ค่าความสามารถของผู้สอบมากกว่าหรือเท่ากับความสามารถเฉลี่ยของกลุ่มผู้สอบทั้งหมด) และ กลุ่มผู้สอบที่มีความสามารถต่ำ (ค่าความสามารถของผู้สอบน้อยกว่าความสามารถเฉลี่ยของกลุ่มผู้สอบทั้งหมด)

3.2.2 แบ่งข้อสอบทั้งหมดโดยใช้ค่าความยากของข้อสอบจากกลุ่มตัวอย่าง 1,200 คน ซึ่งประมาณได้ตามทฤษฎีการตอบสนองข้อสอบ แบบ 2 พารามิเตอร์ ออกเป็น 3 กลุ่มย่อย ได้แก่ กลุ่มข้อสอบยาก กลุ่มข้อสอบยากง่ายปานกลาง และ กลุ่มข้อสอบง่าย

3.2.3 วิเคราะห์หาข้อสอบทำหน้าที่ต่างกัน แต่ละส่วนแยกจากกันเป็นอิสระ ดังนี้
ส่วนที่ 1 วิเคราะห์หาข้อสอบทำหน้าที่ต่างกัน ในแบบสอบทั้งฉบับ โดยใช้กลุ่ม
ผู้สอบที่มีความสามารถสูง

ส่วนที่ 2 วิเคราะห์หาข้อสอบทำหน้าที่ต่างกัน ในแบบสอบทั้งฉบับ โดยใช้กลุ่ม
ผู้สอบที่มีความสามารถต่ำ

ส่วนที่ 3 วิเคราะห์หาข้อสอบทำหน้าที่ต่างกัน ในกลุ่มข้อสอบยาก โดยใช้กลุ่ม
ผู้สอบที่มีความสามารถสูง

ส่วนที่ 4 วิเคราะห์หาข้อสอบทำหน้าที่ต่างกัน ในกลุ่มข้อสอบยากง่ายปานกลาง
โดยใช้กลุ่มผู้สอบที่มีความสามารถสูง

ส่วนที่ 5 วิเคราะห์หาข้อสอบทำหน้าที่ต่างกัน ในกลุ่มข้อสอบง่าย โดยใช้กลุ่ม
ผู้สอบที่มีความสามารถสูง

ส่วนที่ 6 วิเคราะห์หาข้อสอบทำหน้าที่ต่างกัน ในกลุ่มข้อสอบยาก โดยใช้กลุ่ม
ผู้สอบที่มีความสามารถต่ำ

ส่วนที่ 7 วิเคราะห์หาข้อสอบทำหน้าที่ต่างกัน ในกลุ่มข้อสอบยากง่ายปานกลาง
โดยใช้กลุ่มผู้สอบที่มีความสามารถต่ำ

ส่วนที่ 8 วิเคราะห์หาข้อสอบทำหน้าที่ต่างกัน ในกลุ่มข้อสอบง่าย โดยใช้กลุ่มผู้สอบ
ที่มีความสามารถต่ำ

3.2.4. เกณฑ์ที่ใช้ตัดสินข้อสอบทำหน้าที่ต่างกัน คือ สถิติ $MH-\chi^2$ มีนัยสำคัญ
ที่ระดับ .05

3.3 เกณฑ์การตัดสินว่า ข้อสอบทำหน้าที่ต่างกัน ที่ตรวจพบตามข้อ 3.1 และ 3.2
เป็นข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ คือ สอดคล้องกับข้อสอบทำหน้าที่ต่างกันแบบ
ไม่สม่ำเสมอที่ตรวจพบ ตามวิธี IRT area แบบ 2 พารามิเตอร์

ตอนที่ 4 การเปรียบเทียบผลการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ระหว่าง
วิธีแมนเทิล-แฮนส์เชล แบบปกติ กับ แบบแบ่งกลุ่มความสามารถของผู้สอบ
และความยากของข้อสอบ

4.1 กำหนดให้ข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ที่ตรวจพบตามวิธี IRT area
แบบ 2 พารามิเตอร์ เป็นเกณฑ์สำหรับใช้เปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกัน

แบบไม่สม่ำเสมอของข้อสอบ ระหว่างวิธีแมนเทิล-แฮนส์เซล แบบปกติ กับ แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ โดยมีหลักเกณฑ์การตัดสินข้อสอบทำหน้าที่ต่างกันว่าตรวจพบตามวิธีแมนเทิล-แฮนส์เซล ว่าเป็น ข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ คือ สอดคล้องกับข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ตามวิธี IRT area แบบ พารามิเตอร์

4.2 เปรียบเทียบสัดส่วนและเปอร์เซ็นต์การตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ระหว่างวิธีแมนเทิล-แฮนส์เซล แบบปกติ กับ วิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ เมื่อใช้กับจำนวนข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอที่ตรวจพบ ตามวิธี IRT area เป็นเกณฑ์

ตอนที่ 5 การวิเคราะห์หาข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ตามวิธีแมนเทิล-แฮนส์เซล แบบปกติ กับ แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ จากการจำลองกลุ่มตัวอย่างจากข้อมูลจริง

ผู้วิจัยได้ทำการวิเคราะห์หาข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ตามวิธีแมนเทิล-แฮนส์เซล แบบปกติ กับ แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ จากการจำลองกลุ่มตัวอย่างจากข้อมูลจริง จำนวน 10 ครั้ง เพื่อตรวจสอบความคงเส้นคงวาของสัดส่วนการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ระหว่างวิธีทั้งสอง เมื่อใช้จำนวนข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอที่ตรวจพบ ตามวิธี IRT area เป็นเกณฑ์โดยดำเนินการดังนี้

5.1 ใช้โปรแกรมคอมพิวเตอร์ SPSS สุ่มผู้สอบจากกลุ่มตัวอย่าง 1,200 คน แบบไม่ใส่คืน โดยสุ่มแยกตามเพศ เป็นนักเรียนชาย 500 คน และนักเรียนหญิง 500 คน ได้กลุ่มตัวอย่าง ขนาด 1,000 คน ทำการสุ่มผู้สอบจากกลุ่มตัวอย่างซ้ำแบบเดิมอีก รวมจำนวนทั้งหมด 10 ครั้ง ได้กลุ่มตัวอย่าง จำนวน 10 กลุ่ม

5.2 ดำเนินการวิเคราะห์ข้อมูลเช่นเดียวกับตอนที่ 3 การวิเคราะห์หาข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ตามวิธีแมนเทิล-แฮนส์เซล แบบปกติ กับ แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ

ตอนที่ 6 สถิติที่ใช้วิเคราะห์ข้อมูล

6.1 สถิติที่ใช้แสดงคุณภาพของข้อสอบรายข้อ ได้แก่

6.1.1 ค่าความยากของข้อสอบ (Crocker and Algina, 1986)

$$P_j = \frac{\text{จำนวนผู้สอบที่ตอบข้อสอบข้อ } j \text{ ถูก}}{N}$$

เมื่อ P แทน ค่าความยากของข้อสอบ
 N แทน จำนวนผู้สอบทั้งหมด
 j แทน ข้อสอบแต่ละข้อ

6.1.2 ค่าอำนาจจำแนกของข้อสอบ (Crocker and Angina, 1986)

$$\text{ตัวเลือกถูก (r)} = P_H - P_L$$

$$\text{ตัวลวง (r)} = P_L - P_H$$

เมื่อ r แทน ค่าอำนาจจำแนกของข้อสอบ
 P_H แทน สัดส่วนของคนที่ตอบข้อสอบถูกในกลุ่มสูง
 P_L แทน สัดส่วนของคนที่ตอบข้อสอบถูกในกลุ่มต่ำ

6.2 สถิติที่ใช้แสดงคุณภาพของแบบสอบ ได้แก่ ความเที่ยงของคะแนนแบบสอบ หาจากการประมาณค่าสัมประสิทธิ์ความคงที่ภายใน โดยวิธี Kuder Richardson 20 (KR 20) (Crocker and Algina, 1986)

$$KR_{20} = \frac{k}{k-1} \left[1 - \frac{\sum pq}{\sigma_1^2} \right]$$

- เมื่อ k แทน จำนวนข้อสอบในแบบสอบ
 p แทน สัดส่วนของผู้สอบที่ตอบข้อสอบถูก (1) ในแต่ละข้อ
 q แทน สัดส่วนของผู้สอบที่ตอบข้อสอบผิด (0) ในแต่ละข้อ
 $(q = 1 - p)$
 σ_x^2 แทน ความแปรปรวนของคะแนนแบบสอบทั้งฉบับ

6.3 สถิติทดสอบการทำหน้าที่ต่างกันของข้อสอบตามวิธีแมนเทิล-เฮนส์เชล (Dorans and Holland, 1993)

$$MH-\chi^2 = \left[\sum_m R_{rm} - \sum_m E(R_{rm}) \right]^2 / \sum_m \text{Var}(R_{rm})$$

$$E(R_{rm}) = E(R_{rm} \mid \alpha = 1) = N_{rm} R_{tm} / N_{tm}$$

$$\begin{aligned} \text{Var}(R_{rm}) &= \text{Var}(R_{rm} \mid \alpha = 1) \\ &= [N_{rm} R_{tm} N_{fm} W_{tm}] / [N_{rm}^2 (N_{tm} - 1)] \end{aligned}$$

เมื่อ R_{rm} แทน จำนวนผู้สอบกลุ่มอ้างอิงที่ตอบข้อสอบถูก (1)
 Ω ระดับชั้นคะแนน m

$E(R_{rm})$ แทน ความถี่ที่คาดว่าจะเป็นของจำนวนผู้สอบกลุ่มอ้างอิง
 ที่ตอบข้อสอบถูก (1) Ω ระดับชั้นคะแนน m

$\text{Var}(R_{rm})$ แทน ความแปรปรวนของจำนวนผู้สอบกลุ่มอ้างอิงที่
 ตอบข้อสอบถูก (1) Ω ระดับชั้นคะแนน m

N_{rm} แทน จำนวนผู้สอบกลุ่มอ้างอิง Ω ระดับชั้นคะแนน m

R_{tm} แทน จำนวนผู้สอบที่ตอบข้อสอบถูก (1) Ω ระดับชั้นคะแนน m

N_{tm} แทน จำนวนผู้สอบทั้งหมดของสองกลุ่ม Ω ระดับชั้นคะแนน m

N_{fm} แทน จำนวนผู้สอบกลุ่มสนใจ Ω ระดับชั้นคะแนน m

W_{tm} แทน จำนวนผู้สอบที่ตอบข้อสอบผิด (0) Ω ระดับชั้นคะแนน m

6.4 คำนวณวัดขนาดการทำหน้าที่ต่างกันของข้อสอบ ตามวิธีแมนเทล-เฮนส์เชล (Dorans and Holland, 1993)

$$\text{MH D-DIF} = -2.35 \ln [\alpha_{\text{MH}}]$$

$$\alpha_{\text{MH}} = \left[\sum_m R_{\text{fm}} W_{\text{fm}} / N_{\text{tm}} \right] / \left[\sum_m R_{\text{fm}} W_{\text{fm}} / N_{\text{tm}} \right]$$

เมื่อ	MH D-DIF	แทน	ค่านีวัดขนาดการทำหน้าที่ต่างกันของข้อสอบ
	-2.35	แทน	ค่าปรับสเกลของ หน่วยงานบริการทดสอบทางการศึกษา แห่งสหรัฐอเมริกา
	$\ln [\alpha_{\text{MH}}]$	แทน	ลอการริทึมธรรมชาติของตัวประมาณค่าอัตราส่วนแถมต่อคงที่
	R_{fm}	แทน	จำนวนผู้สอบกลุ่มอ้างอิงที่ตอบข้อสอบถูก (1) ณ ระดับชั้นคะแนน m
	W_{fm}	แทน	จำนวนผู้สอบกลุ่มสนใจที่ตอบข้อสอบผิด (0) ณ ระดับชั้นคะแนน m
	N_{tm}	แทน	จำนวนผู้สอบทั้งหมดของสองกลุ่ม ณ ระดับชั้นคะแนน m
	R_{fm}	แทน	จำนวนผู้สอบกลุ่มสนใจที่ตอบข้อสอบถูก (1) ณ ระดับชั้นคะแนน m
	W_{fm}	แทน	จำนวนผู้สอบกลุ่มอ้างอิงที่ตอบข้อสอบผิด (0) ณ ระดับชั้นคะแนน m

6.5 สถิติที่ใช้ในการทดสอบสัดส่วนของข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอที่ตรวจพบ ระหว่างวิธีแมนเทล-เฮนส์เชล แบบปกติ กับ วิธีแมนเทล-เฮนส์เชล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ใช้ Z - statistic (Glass and Hopkins, 1984)

$$Z = \frac{(P_1 - P_2)}{\sqrt{P(1-P) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

$$P = \frac{f_1 + f_2}{n_1 + n_2}$$

เมื่อ	Z แทน	ค่าสถิติ Z
P_1	แทน	สัดส่วนของข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอที่ตรวจพบตามวิธีแมนเทิล-แฮนส์เซิล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ
P_2	แทน	สัดส่วนของข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอที่ตรวจพบตามวิธีแมนเทิล-แฮนส์เซิล แบบปกติ
f_1	แทน	จำนวนข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอที่ตรวจพบตามวิธีแมนเทิล-แฮนส์เซิล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ
f_2	แทน	จำนวนข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอที่ตรวจพบตามวิธีแมนเทิล-แฮนส์เซิล แบบปกติ
n_1	แทน	จำนวนข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอทั้งหมด ตามวิธีแมนเทิล-แฮนส์เซิล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ
n_2	แทน	จำนวนข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอทั้งหมด ตามวิธีแมนเทิล-แฮนส์เซิล แบบปกติ
P	แทน	สัดส่วนของข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอทั้งหมดของสองกลุ่ม