

การสร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 2

- 3 ก.พ. 2526

สำนักหอสมุดกลาง มหาวิทยาลัยศรีนครินทรวิโรฒ
 ถนนวิภาวดี 23 พระโขนง กรุงเทพฯ 11 โทร. 3921 575, 391 5058
 ปรินญาณินพนธ์

ของ

สมถวิล วิจิตรวรรณา

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประธานมิต
 เพื่อเป็นส่วนหนึ่งของการศึกษาค้นคว้าหลักสูตร
 ปรินญาณการศึกษามหาบัณฑิต

กรกฎาคม 2524

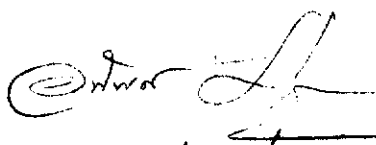
ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

150544

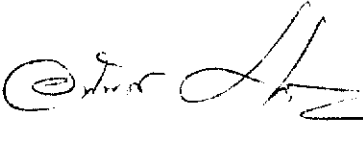
คณะกรรมการที่ปรึกษาประจำตัวนิสิตและกรรมการสอบ ได้พิจารณาปัญหานี้นพนธ์ฉบับนี้แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาการศึกษามหาบัณฑิตของมหาวิทยาลัยศรีนครินทรวิโรฒ

คณะกรรมการที่ปรึกษา

คณะกรรมการสอบ



ประธาน



ประธาน

ดร. สungsan

กรรมการ

ดร. สungsan

กรรมการ

(ดร. ประสงค์) กรรมการ

กรรมการ

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้สำเร็จได้ด้วยดี เนื่องจากผู้วิจัยได้รับคำแนะนำและช่วยเหลือ
ตลอดจนแก้ไขข้อบกพร่องต่างๆจาก รองศาสตราจารย์ ดร. อนันต์ ศรีโสภา และ
อาจารย์ ดร. ผจงจิต อินทสุวรรณ ผู้วิจัยรู้สึกซาบซึ้งและขอกราบขอบพระคุณเป็น
อย่างสูงไว้ ณ ที่นี้

ขอกราบขอบพระคุณ อาจารย์ ดร. สงบ ลักษณะ อาจารย์ ดร. โกวิท
ประมวลพฤษ์ อาจารย์สมศักดิ์ สินธุระเวชญ์ และคณะศึกษานิเทศก์จังหวัด
สมุทรสงคราม ที่กรุณาให้คำแนะนำและคำปรึกษาในการสร้างแบบทดสอบ

ขอขอบพระคุณ ผู้อำนวยความสะดวก อาจารย์ใหญ่ ครูใหญ่ และคณะครูของ
โรงเรียนในกลุ่มตัวอย่าง และขอขอบใจนักเรียนทุกคนที่ให้ความร่วมมือในการเก็บ
รวบรวมข้อมูล

ขอขอบคุณ คุณสำราญ ชุมพะวัฑ์ คุณหริน สงบธรรม คุณพิศเพลิน
หอมสุวรรณ คุณประวิง รอดเข็ม คุณศศิธร เล็กสุขศรี คุณสุภาพร เลหาสดีตย์
ตลอดจนเพื่อนอีกหลายท่านที่ให้ความช่วยเหลือในการทำปริญญานิพนธ์จนสำเร็จลุล่วง
มากด้วยดี

ท้ายที่สุดนี้ ผู้วิจัยขอน้อมรำลึกถึงพระคุณของบิดามารดา ที่ให้กำลังใจ
และสนับสนุนการศึกษาของผู้วิจัยตลอดมา พระคุณนี้หาที่เปรียบเทียบมิได้

สมณวิไล วิจิตรวรรณ

สารบัญ

บทที่	หน้า
1 บทนำ	1
ภูมิหลัง	1
ความมุ่งหมายของการศึกษาค้นคว้า	3
ความสำคัญของการศึกษาค้นคว้า	4
ขอบเขตของการศึกษาค้นคว้า	4
คำจำกัดความเฉพาะ	5
2 เอกสารและงานวิจัยที่เกี่ยวข้อง	6
ประวัติที่มาของการวัดผลแบบอิงเกณฑ์	6
ลักษณะการวัดผลและแบบทดสอบอิงเกณฑ์	8
การเปรียบเทียบความแตกต่างระหว่างการวัดผลแบบอิงกลุ่ม-อิงเกณฑ์	12
การสร้างแบบทดสอบอิงเกณฑ์	13
การกำหนดคะแนนเกณฑ์ในแบบทดสอบอิงเกณฑ์	15
การวิเคราะห์รายข้อ	26
ความเชื่อมั่นของแบบทดสอบอิงเกณฑ์	33
ความเที่ยงตรงของแบบทดสอบอิงเกณฑ์	38
งานวิจัยที่เกี่ยวข้องกับแบบทดสอบอิงเกณฑ์	41
3 วิธีดำเนินการศึกษาค้นคว้า	43
แหล่งข้อมูลและกลุ่มตัวอย่าง	43
วิธีดำเนินการสร้างแบบทดสอบอิงเกณฑ์	44
วิธีดำเนินการเก็บข้อมูล	60
วิธีจัดการกระทำกับข้อมูล	61

บทที่	หน้า
4 ผลการวิเคราะห์ข้อมูล	62
สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล	62
การเสนอผลการวิเคราะห์ข้อมูล	62
ผลการวิเคราะห์ข้อมูล	63
5 สรุป อภิปรายผล และข้อเสนอแนะ	112
ความมุ่งหมายในการศึกษาค้นคว้า	112
กลุ่มตัวอย่าง	112
เครื่องมือที่ใช้ในการรวบรวมข้อมูล	112
วิธีดำเนินการรวบรวมข้อมูล	113
การวิเคราะห์ข้อมูล	113
สรุปผลการวิเคราะห์ข้อมูล	114
อภิปรายผล	120
ข้อเสนอแนะ	124
บรรณานุกรม	126
ภาคผนวก	133

บัญชีตาราง

ตาราง

หน้า

1	การเปรียบเทียบความแตกต่างระหว่างการวัดผลแบบอิงกลุ่ม และอิงเกณฑ์	12
2	ความน่าจะเป็นของความสัมพันธ์ระหว่างคะแนนจากแบบทดสอบ กับเกณฑ์ภายนอก	20
3	ความสัมพันธ์ระหว่างคะแนนจากแบบทดสอบกับสถานการณ์ การเรียนรู้	22
4	การจำแนกเกณฑ์และคะแนนพยากรณ์ของนักเรียน 2 กลุ่ม	23
5	ความเที่ยงตรงตามเนื้อหาที่ได้จากคะแนนนักเรียน 2 กลุ่ม	40
6	รายชื่อโรงเรียน จำนวนนักเรียนในกลุ่มตัวอย่าง และกริที่ใช้ ทดสอบ	44
7	เนื้อหาและจุดประสงค์เชิงพฤติกรรมเรื่องการหาร ชั้นประถมศึกษาปีที่ 2	46
8	ประชากรเนื้อหาที่ได้จากขอบเขตเนื้อหาเรื่องการนับจุด	50
9	ประชากรเนื้อหาที่ได้จากขอบเขตเนื้อหาเรื่องการหาร	51
10	รายชื่อแบบทดสอบย่อย จุดประสงค์เชิงพฤติกรรม และจำนวนข้อสอบ 52	
11	ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน ค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 1 จากการ ทดสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม	64
12	ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน ค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 2 จากการ ทดสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม	70

13	ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน ค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 3 จากการ ทดสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม	75
14	ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน ค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 4 จากการ ทดสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม	79
15	ความน่าจะเป็นในการตัดสินใจที่ถูกต้อง ความคลาดเคลื่อนของการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์ เชิงพฤติกรรม 1.1.1 ในแบบทดสอบฉบับที่ 1	83
16	ความน่าจะเป็นในการตัดสินใจที่ถูกต้อง ความคลาดเคลื่อนของการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์ เชิงพฤติกรรม 1.1.2 ในแบบทดสอบฉบับที่ 1	86
17	ความน่าจะเป็นในการตัดสินใจที่ถูกต้อง ความคลาดเคลื่อนของการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์ เชิงพฤติกรรม 1.1.3 ในแบบทดสอบฉบับที่ 1	88
18	ความน่าจะเป็นในการตัดสินใจที่ถูกต้อง ความคลาดเคลื่อนของการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์ เชิงพฤติกรรม 1.2.1 ในแบบทดสอบฉบับที่ 1	90
19	ความน่าจะเป็นในการตัดสินใจที่ถูกต้อง ความคลาดเคลื่อนของการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์ เชิงพฤติกรรม 1.2.2 ในแบบทดสอบฉบับที่ 1	92
20	ความน่าจะเป็นในการตัดสินใจที่ถูกต้อง ความคลาดเคลื่อนของการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์ เชิงพฤติกรรม 2.1 ในแบบทดสอบฉบับที่ 2	95

21	ความน่าจะเป็นในการตัดสินใจที่ถูกต้อง ความคลาดเคลื่อนของการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์ เชิงพฤติกรรม 2.2 ในแบบทดสอบฉบับที่ 2	97
22	ความน่าจะเป็นในการตัดสินใจที่ถูกต้อง ความคลาดเคลื่อนของการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์ เชิงพฤติกรรม 3.1 ในแบบทดสอบฉบับที่ 3	100
23	ความน่าจะเป็นในการตัดสินใจที่ถูกต้อง ความคลาดเคลื่อนของการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์ เชิงพฤติกรรม 3.2 ในแบบทดสอบฉบับที่ 3	102
24	ความน่าจะเป็นในการตัดสินใจที่ถูกต้อง ความคลาดเคลื่อนของการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์ เชิงพฤติกรรม 4.1 ในแบบทดสอบฉบับที่ 4	105
25	ความน่าจะเป็นในการตัดสินใจที่ถูกต้อง ความคลาดเคลื่อนของการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์ เชิงพฤติกรรม 4.2 ในแบบทดสอบฉบับที่ 4	107
26	ค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ทั้งสี่ฉบับ จากการทดสอบ ครั้งที่สาม	109
27	ค่าความเที่ยงตรงตามเนื้อหาของแบบทดสอบอิงเกณฑ์ทั้งสี่ฉบับ จากการทดสอบครั้งที่สาม	110

บัญชีประกอบภาพ

ภาพประกอบ	หน้า
1 แสดงความคลาดเคลื่อนของการประเมิน	16
2 การจำแนกนักเรียนตามคะแนนพยากรณ์ที่ได้จากจุดตัดของ การกระจายคะแนนของนักเรียน 2 กลุ่ม	24
3 ลำดับชั้นในการดำเนินการสร้างแบบทดสอบอิงเกณฑ์ วิชาคณิตศาสตร์	45
4 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ ก่อนเรียนและคะแนนในการสอบหลังเรียน ของจุดประสงค์ 1.1.1 ในแบบทดสอบฉบับที่ 1	82
5 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ ก่อนเรียนและคะแนนในการสอบหลังเรียน ของจุดประสงค์ 1.1.2 ในแบบทดสอบฉบับที่ 1	85
6 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ ก่อนเรียนและคะแนนในการสอบหลังเรียน ของจุดประสงค์ 1.1.3 ในแบบทดสอบฉบับที่ 1	87
7 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ ก่อนเรียนและคะแนนในการสอบหลังเรียน ของจุดประสงค์ 1.2.1 ในแบบทดสอบฉบับที่ 1	89
8 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ ก่อนเรียนและคะแนนในการสอบหลังเรียน ของจุดประสงค์ 1.2.2 ในแบบทดสอบฉบับที่ 1	91

9	คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ ก่อนเรียนและสอบหลังเรียน ของจุดประสงค์เชิงพฤติกรรม 2.1 ในแบบทดสอบฉบับที่ 2	94
10	คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ ก่อนเรียนและสอบหลังเรียน ของจุดประสงค์เชิงพฤติกรรม 2.2 ในแบบทดสอบฉบับที่ 2	96
11	คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ ก่อนเรียนและสอบหลังเรียน ของจุดประสงค์เชิงพฤติกรรม 3.1 ในแบบทดสอบฉบับที่ 3	99
12	คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ ก่อนเรียนและสอบหลังเรียน ของจุดประสงค์เชิงพฤติกรรม 3.2 ในแบบทดสอบฉบับที่ 3	101
13	คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ ก่อนเรียนและสอบหลังเรียน ของจุดประสงค์เชิงพฤติกรรม 4.1 ในแบบทดสอบฉบับที่ 4	104
14	คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ ก่อนเรียนและสอบหลังเรียน ของจุดประสงค์เชิงพฤติกรรม 4.2 ในแบบทดสอบฉบับที่ 4	106

ภูมิหลัง

การวัดผลและประเมินผลทางการศึกษา เป็นสิ่งจำเป็นควบคู่กับการเรียน การสอน ในขณะที่สังคมเปลี่ยนแปลงไป การศึกษาก็ต้องปรับตัวให้เข้ากับความ เปลี่ยนแปลงนั้นๆ เพื่อยังประโยชน์ให้เกิดแก่สังคม ขณะเดียวกันการวัดผลและการประเมิน ผลการศึกษาก็ต้องปรับปรุงให้สอดคล้องกับวัตถุประสงค์ของการศึกษา เพื่อเป็นเครื่อง ชี้แนะว่าการเรียนการสอนได้ดำเนินไปตามวัตถุประสงค์ของการศึกษาหรือไม่ เพียงไร. จากการประกาศใช้หลักสูตรประถมศึกษาและมัธยมศึกษา พุทธศักราช 2521 การศึกษาได้เปลี่ยนแปลงไปในลักษณะที่มุ่งพัฒนาชีวิตและสังคมของผู้เรียน ซึ่งจะช่วย พัฒนาประเทศชาติในที่สุด การวัดผลและประเมินผลตามระเบียบกระทรวงศึกษาธิการ พ.ศ. 2520, พ.ศ. 2521 และ พ.ศ. 2523 กำหนดวิธีการวัดและประเมินผลเพื่อให้สอดคล้องกับหลักสูตรดังกล่าวเป็น 2 ประเภทคือ การประเมินผลเพื่อปรับปรุง การเรียนการสอนหรือการประเมินระหว่างภาคเรียน และการประเมินเพื่อตัดสิน การเรียนหรือการประเมินผลปลายภาค

การประเมินผลเพื่อตัดสินผลการเรียนหรือการประเมินผลปลายภาค (Summative Evaluation) มีจุดมุ่งหมายเพื่อทราบผลการพัฒนาการเรียนทั้งหมด ในรายวิชาที่เรียนว่าถึงปลายภาคเรียน นักเรียนมีความสามารถระดับใด (กระทรวงศึกษาธิการ 2523 : 10) การตัดสินต้องอาศัยการเปรียบเทียบกับกลุ่ม ซึ่งเป็นลักษณะการวัดผลแบบอิงกลุ่ม (ส.ว.สนา ประมวลพญักษ์ 2522 : 3, สมบูรณ์ ชิตพงศ์ 2520 : 27, และบุญเชิด ภิญโญนนตพงศ์ 2521 : 27 อ้างอิงมาจาก Bloom. 1971) การวัดผลแบบอิงกลุ่มแปลความหมายของผลการวัด ได้ว่า นักเรียนมีความสามารถมากหรือน้อยกว่าคนอื่น มี การเปรียบเทียบกับกลุ่ม นักเรียนที่ใช้เครื่องมือวัดอย่างเดียวกัน จะต้องเป็นเครื่องมือที่สามารถจำแนกนักเรียน

ออกลดหลั่นกันตามความสามารถ (กมล ภูประเสริฐ 2520 : 12)

ส่วนการประเมินผลเพื่อปรับปรุงการเรียนการสอนหรือประเมินผลระหว่างภาค (Formative Evaluation) X มีจุดมุ่งหมายเพื่อทราบว่านักเรียนมีความรู้ความสามารถตามจุดประสงค์ที่กำหนดไว้หรือไม่ (กระทรวงศึกษาธิการ 2523 : 9)

ในการประเมินผลเพื่อปรับปรุงการเรียนการสอน การวัดผลแบบอิงเกณฑ์มีความเหมาะสมกับการประเมินดังกล่าว (ส.วาสนา ประवालพุกษ์ 2522 : 3)

กมล สุกประเสริฐ 2520 : 21 - 23, สมบูรณ์ ชิตพงศ์ 2520 : 92 และ บุญเจิด ภิญโญนนทพงศ์ 2521 : 27 อ้างอิงมาจาก Bloom. 1971)

การวัดผลแบบอิงเกณฑ์แปลความหมายของผลการวัดได้ว่า นักเรียนมีความสามารถหรือไม่ จะต้องมีการเปรียบเทียบผลการวัดของนักเรียนคนนั้นกับเกณฑ์ที่กำหนดให้ โดยไม่คำนึงถึงผลการวัดของนักเรียนคนอื่นๆ เครื่องมือที่ใช้วัดต้องสอดคล้องกับเกณฑ์ที่กำหนดไว้ให้มากที่สุด คำว่า "เกณฑ์" ในที่นี้หมายถึงกลุ่มของพฤติกรรมที่กำหนดไว้ อย่างดี ซึ่งมีความหมายกว้างกว่าเกณฑ์การผ่านในจุดประสงค์เชิงพฤติกรรม เพราะรวมทั้งจุดประสงค์เชิงพฤติกรรมและรายละเอียดเพิ่มเติมอย่างอื่น ที่จะทำให้เกณฑ์มีความหมายชัดเจนมากที่สุด (กมล ภูประเสริฐ 2520 : 12 - 13)

การวัดผลแบบอิงเกณฑ์เป็นระบบการวัดผลแบบใหม่ซึ่งแพร่หลายในช่วงเวลา กว่า 10 ปีที่ผ่านมา มีระเบียบวิธีการต่างไปจากการวัดแบบเดิมหรือแบบอิงกลุ่ม การวัดผลแบบอิงเกณฑ์มาจากรากฐานความเชื่อว่า ทุกคนสามารถจะเรียนรู้ทุกสิ่งทุกอย่างได้เหมือนกันหมดถ้าให้เวลาไม่จำกัด จากความเชื่อนี้ได้พัฒนาเป็นทฤษฎี การเรียนรู้ที่เรียกว่า Mastery Learning (ส.วาสนา ประवालพุกษ์ 2522 : 2)

กมล สุกประเสริฐ ได้กล่าวถึงการเขียนแบบทดสอบอิงเกณฑ์ว่า ต้องให้ครอบคลุมจุดประสงค์เชิงพฤติกรรมของการสอน และตัดสินตามจุดประสงค์เชิงพฤติกรรมในแต่ละหน่วยเรียนเป็นสิ่งที่ครูทำได้ง่าย สิ่งใดที่นักเรียนตอบไม่ได้ก็ให้ไปดูสิ่งนั้นเพิ่มเติม หรือซ่อมเสริมให้ และสอบแก้ตัวในสิ่งที่ตกเสียใหม่ (กมล สุกประเสริฐ 2520 : 7)

ถ้าแบบทดสอบวัดหลายจุดประสงค์ก็ต้องเป็นแบบทดสอบที่มีชุดย่อยๆ หลายชุด การแปล

ความหมายของคะแนนจะแยกเป็นแต่ละจุดประสงค์ โดยใช้คะแนนจุดคิดของแต่ละจุดประสงค์เชิงพฤติกรรมหรือความรู้ขั้นต่ำของผู้สอบ (สงบ ลักษณะ 2523 : 5)

ระเบียบการประเมินผลที่มีส่วนของการวัดผลแบบอิงเกณฑ์อยู่มาก ได้แก่ ระเบียบการประเมินผลในระดับประถมศึกษา ซึ่งกำหนดให้มีการประเมินผลเป็นรายตัวในสมุดประจำชั้นอีกด้วย (โกวิท ประวาลพุกษ์ 2523 : 16) โดยครูผู้สอนจะต้องระบุว่า นักเรียนของเรามีความรู้ความสามารถ"ผ่าน"จุดประสงค์เชิงพฤติกรรมที่กำหนดไว้ในสมุดประจำชั้นในแต่ละวิชา(กลุ่มวิชา)หรือไม่ (วิชัย ราษฎร์ศิริ 2522 : 56) เนื่องจากการวัดผลแบบอิงเกณฑ์เป็นการวัดผลในรูปแบบใหม่คงกล่าวแล้ว ความรู้ในเรื่องนี้ยังไม่แพร่หลายทำให้ครูผู้สอนและบุคคลทั่วไปยังไม่เข้าใจเพียงพอ เป็นปัญหาในการสร้างแบบทดสอบและหาคุณภาพของแบบทดสอบ ผู้วิจัยมีความสนใจที่จะสร้างแบบทดสอบอิงเกณฑ์และหาคุณภาพของแบบทดสอบที่สร้างขึ้น โดยจะหาคะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรมด้วย ซึ่งใช้วิชาคณิตศาสตร์อันเป็นวิชาในกลุ่มทักษะและมีเนื้อหาที่มีลำดับของพฤติกรรม โดยเฉพาะเรื่องการหาร ในระดับชั้นประถมศึกษาปีที่ 2 นักเรียนส่วนใหญ่ยังขาดทักษะและความคิดรวบยอดในเรื่องนี้ แบบทดสอบอิงเกณฑ์ที่สร้างขึ้นนอกจากจะเป็นแนวทางในการสร้างแบบทดสอบอิงเกณฑ์ต่อไปแล้ว ยังเป็นเครื่องมือที่วัดผลสัมฤทธิ์ การศึกษาตามจุดประสงค์การเรียน ซึ่งสอดคล้องกับระเบียบการประเมินผลในระดับประถมศึกษาอีกด้วย

☆ ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อสร้างแบบทดสอบอิงเกณฑ์วัดผลสัมฤทธิ์ในการเรียนวิชาคณิตศาสตร์เรื่องการหาร สำหรับชั้นประถมศึกษาปีที่ 2 และหาคุณภาพของแบบทดสอบที่สร้างขึ้น
2. เพื่อหาคะแนนเกณฑ์ที่เหมาะสมในแต่ละจุดประสงค์เชิงพฤติกรรมของแบบทดสอบอิงเกณฑ์ที่สร้างขึ้น

★ ความสำคัญของการศึกษาค้นคว้า

ผลการศึกษาค้นคว้าครั้งนี้ทำได้

1. แบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่องการหาร ชั้นประถมศึกษาปีที่ 2 ที่มีคุณภาพที่สามารถนำไปใช้วัดผลเมื่อสิ้นสุดการเรียนในแต่ละตอน เพื่อนำผลมาปรับปรุงการเรียนการสอน และเป็นแนวทางในการสอนซ่อมเสริมนักเรียนที่ยังไม่ผ่านเกณฑ์ในจุดประสงค์นั้นๆ

คะแนนเกณฑ์ที่ใช้ตัดสินการผ่านจุดประสงค์เชิงพฤติกรรมของแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่องการหาร ชั้นประถมศึกษาปีที่ 2 ที่สร้างขึ้น

3. แนวทางการวิจัยในส่วนของการวัดผลแบบอิงเกณฑ์ และพัฒนาวิธีการสร้างแบบทดสอบอิงเกณฑ์ในเรื่องอื่น ระดับอื่น และวิชาอื่นๆต่อไป

→ ขอบเขตของการศึกษาค้นคว้า

1. ประชากรที่ใช้ในการศึกษาค้นคว้านี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 2 ปีการศึกษา 2523 ของโรงเรียนในสังกัดสำนักงานการประถมศึกษา จังหวัดสมุทรสงคราม

2. กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้านี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 2 ปีการศึกษา 2523 ของโรงเรียนในสังกัดสำนักงานการประถมศึกษา จังหวัดสมุทรสงคราม จำนวนทั้งสิ้น 370 คน ซึ่งเลือกมาโดยการสุ่มแบบแบ่งชั้น

(Stratified Random Sampling)

3. เนื้อหาที่นำมาศึกษาค้นคว้านี้ เป็นเนื้อหาวิชาคณิตศาสตร์ เรื่องการหาร ชั้นประถมศึกษาปีที่ 2 ตามหลักสูตรประถมศึกษา พุทธศักราช 2521

คำจำกัดความเฉพาะ

1. แบบทดสอบเชิงเกณฑ์ หมายถึง แบบทดสอบวัดผลสัมฤทธิ์ที่ประกอบด้วยข้อคำถามที่สร้างจากจุดประสงค์เชิงพฤติกรรม ซึ่งอาจสุ่มมาจากประชากรของพฤติกรรม มีคะแนนเกณฑ์ที่ตัดสินการสอบผ่านจุดประสงค์เชิงพฤติกรรมของนักเรียน ความหมายของคะแนนได้จากการเปรียบเทียบคะแนนของนักเรียนกับคะแนนเกณฑ์ที่กำหนด โดยไม่คำนึงถึงผลการสอบของคนอื่นๆ
2. จุดประสงค์เชิงพฤติกรรม หมายถึง จุดประสงค์ที่กำหนดในการสอน ซึ่งเขียนในลักษณะที่บอกการกระทำหรือพฤติกรรมของนักเรียนว่า เมื่อสิ้นสุดในแต่ละเนื้อหาเรียบร้อยแล้ว นักเรียนสามารถทำอะไรได้บ้าง
3. คะแนนเกณฑ์ หมายถึง คะแนนที่ใช้ตัดสินการสอบผ่านในจุดประสงค์เชิงพฤติกรรมนั้นๆ
4. การวิเคราะห์รายข้อ หมายถึง การตรวจสอบคุณภาพของข้อสอบเป็นข้อๆ ซึ่งประกอบด้วย
 - 4.1 ค่าดัชนีความยากของข้อสอบ หมายถึง เปอร์เซนต์ของจำนวนนักเรียนที่ตอบข้อสอบข้อนั้นได้ถูกต้อง
 - 4.2 ค่าดัชนีอำนาจจำแนกของข้อสอบ หมายถึง ข้อสอบที่สามารถจำแนกนักเรียนที่ตอบผิดในการสอบก่อนเรียน แต่ตอบถูกในการสอบหลังเรียนออกจากกลุ่มผู้สอบทั้งหมด
5. คุณภาพของแบบทดสอบ หมายถึง คุณภาพทางด้านความเชื่อมั่นและความเที่ยงตรง
 - 5.1 ความเชื่อมั่นของแบบทดสอบ หมายถึง ค่าสัมประสิทธิ์ของความสอดคล้องของการตัดสินที่คงเส้นคงวาในกลุ่มผู้สอบ
 - 5.2 ความเที่ยงตรงตามเนื้อหาของแบบทดสอบ หมายถึง แบบทดสอบสามารถวัดได้ตรงตามเนื้อหา ข้อสอบเป็นตัวแทนของพฤติกรรมในแต่ละจุดประสงค์เชิงพฤติกรรมที่ต้องการทดสอบ

เอกสารและงานวิจัยที่เกี่ยวข้อง

การสร้างแบบทดสอบอิงเกณฑ์ครั้งนี้ ผู้วิจัยได้ศึกษาเนื้อหาต่าง ๆ เกี่ยวกับแบบทดสอบ
ซึ่งเสนอเนื้อหาตามหัวข้อ ดังนี้

1. ประวัติที่มาของการวัดผลแบบอิงเกณฑ์
2. ลักษณะของการวัดผล และแบบทดสอบอิงเกณฑ์
3. การเปรียบเทียบความแตกต่างระหว่างการวัดผลแบบอิงกลุ่มและอิงเกณฑ์
4. การสร้างแบบทดสอบอิงเกณฑ์
5. การกำหนดคะแนนเกณฑ์ของแบบทดสอบอิงเกณฑ์
6. การวิเคราะห์รายข้อของแบบทดสอบอิงเกณฑ์
7. ความเชื่อมั่นของแบบทดสอบอิงเกณฑ์
8. ความเที่ยงตรงของแบบทดสอบอิงเกณฑ์
9. งานวิจัยที่เกี่ยวข้องกับแบบทดสอบอิงเกณฑ์

ประวัติที่มาของการวัดผลแบบอิงเกณฑ์

เกลเซอร์ (Glaser, 1963 : 519 - 520) เป็นบุคคลแรกที่ใช้คำว่า "การวัดผล
แบบอิงกลุ่ม" และ "การวัดผลแบบอิงเกณฑ์" โดยให้สิ่งกีดขวางของการวัดผลแบบอิงกลุ่ม (Norm-
Referenced Measurement) เป็นการวัดผลสัมฤทธิ์ที่ยึดหลักมาตรฐานสัมพัทธ์ระหว่างผู้สอบ
ส่วนการวัดผลแบบอิงเกณฑ์ (Criterion-Referenced Measurement) เป็นการวัดผล
สัมฤทธิ์ที่ยึดหลักมาตรฐานของผู้สอบอย่างสมบูรณ์ เป็นการทดสอบที่พิสูจน์ว่าผู้สอบมีความรู้หรือไม่
ในงานที่ปฏิบัติ ซึ่งไม่ขึ้นอยู่กับผลการปฏิบัติงานปฏิบัติของคนอื่น ๆ อย่างไรก็ตามแนวคิดเกี่ยวกับ
เรื่องนี้ได้เริ่มมีมาตั้งแต่ ค.ศ. 1913 เมื่อธรอนไคค์ (Thron-dike, 1913) เขียนบทความ
ตีพิมพ์ในวารสารจิตวิทยาการศึกษา ได้กล่าวถึงกระบวนการทั้งสองนี้ แต่ไม่ได้ใช้คำว่า "อิงกลุ่ม"
และ "อิงเกณฑ์" อย่างชัดเจนเท่านั้น แม้วาธรอนไคค์ กล่าวถึงการวางกรอบปัญหาในการอธิบาย
อธิบายเรื่องเกณฑ์ แต่ก็ได้นำมาสู่การวัดผลแบบอิงเกณฑ์ (Popham, 1978 : 9 - 10)

นักวิจัยคนอื่น ๆ ที่กล่าวถึงความแตกต่างของการวัดผลเรื่องนั้นบ้าง ได้ยก ฟลานาแกน (Flanagan. 1951) โดยอธิบายถึงการทดสอบแบบอิงเกณฑ์ในลักษณะของจำนวนคำที่แปลถูกของผู้เรียนว่า "ผู้เรียนจะต้องแปลหนังสือซีเชไว์ จำนวน 1 หน้า โดยผิดได้ไม่เกิน 3 แห่ง" และบทความของอีเบล (Ebel. 1962) กล่าวว่า การวัดผลใด ๆ จะมีความหมายอย่างสมบูรณ์ เมื่อมีการอ้างอิงถึงเนื้อหาอย่างชัดเจน คะแนนจะต้องแปลว่าอยู่ในระดับใดของเนื้อหาทั้งหมด ที่ทำการวัด เช่น การที่ผิดได้ 25 คะแนน นอกจากจะแปลความในลักษณะว่าเขาได้เปอร์เซ็นต์ที่เท่าไรแล้ว คะแนนจะมีความหมายยิ่งขึ้น ถ้าสามารถบอกได้ว่าเขาเรียนรู้แล้วเป็นสัดส่วนเท่าใดของเนื้อหาทั้งหมด (โกวิท ประวาลพุกุล 2523 : 18 อ้างอิงมาจาก Flanagan. 1951 และ Ebel. 1962) ความสนใจการวัดผลแบบอิงเกณฑ์นั้นกล่าวได้ว่า เกิดจากองค์ประกอบหลายประการด้วยกัน คือ

1. การวัดผลแบบอิงกลุ่มอย่างเคียวไม่เพียงพอ ในการให้ข้อมูลที่จะปรับปรุงการเรียนการสอน เช่น เนื้อหาวิชา วิธีการสอน และวิธีสอน
2. แนวคิดใหม่และความเจริญก้าวหน้าเทคโนโลยีการสอน เช่น การสอนรายบุคคล การเรียนเดี่ยว อุปกรณ์แบบโปรแกรม ซึ่งมีหลักสำคัญอยู่ที่มุ่งหมายของการสอนที่ชัดเจน และสามารถวัดได้ว่านักเรียนมีพฤติกรรมตามจุดมุ่งหมายนั้นหรือไม่ ยิ่งในการสอนที่เน้นลำดับขั้นการเรียนรู้ ข้อมูลจากการวัดผลใช้ตัดสินว่านักเรียนควรจะเรียนในหน่วยต่อไปหรือไม่ โดยไม่เกี่ยวข้องกับความสามารถของนักเรียนคนอื่น
3. การวัดผลแบบอิงกลุ่มก่อให้เกิดสถานการณ์การเรียนแบบแข่งขัน ไม่ร่วมมือกัน และลดแรงจูงใจในการเรียนสำหรับนักเรียนที่เรียนระดับปานกลางและต่ำ (กมล ภูประเสริฐ 2518 : 78 - 80 อ้างอิงมาจาก Cronbach. 1969, Tyler. 1970, Glaser and Nitko. 1971 and Gentile. 1971)

✓ ลักษณะการวัดผลและแบบทดสอบอิง เกณฑ์

ได้มีผู้กล่าวถึงลักษณะของการวัดผลและแบบทดสอบอิง เกณฑ์ไว้หลายคน เช่น

เกลเซอร์ (Glaser. 1963 : 519 - 521) กล่าวว่า การวัดผลแบบอิง เกณฑ์ ยึดหลักมาตรฐานของคุณภาพอย่างสมบูรณ์ ขณะที่การวัดผลแบบอิงกลุ่มยึดหลักมาตรฐานสัมพัทธ์ และ เป็นการวัดผลสัมฤทธิ์ระหว่างตั้งแต่ยังมิได้รับการฝึกปฏิบัติจนถึงการกระทำได้อย่างมีรูปแบบ ระบุกับ ผลสัมฤทธิ์ของนักเรียนชั้นต่ำสามารถประเมินได้โดยการวัดผลแบบอิง เกณฑ์ ซึ่งไม่เกี่ยวข้องกับ การอ้างอิงงานปฏิบัติของคนอื่น ซึ่งนิตโก (Nitko. 1970 : 38) ก็ได้ให้ความหมายของ การวัดผลแบบอิง เกณฑ์ในทำนองเดียวกันว่า ระบบอิง เกณฑ์สนใจแต่ความรู้ขั้นต่ำของผู้เรียนโดยไม่สนใจการปฏิบัติของคนอื่น ๆ และคะแนนผลของการกระทำตามลำดับขั้นจนกระทั่งสามารถรอบรู้

ปอปแฮม และฮูเซก (Popham and Husek. 1969 : 1 - 9) กล่าวถึงการวัดผล แบบอิง เกณฑ์ว่า เป็นการตรวจสอบสัมฤทธิ์ผลของบุคคลในการกระทำพฤติกรรมต่าง ๆ โดยเปรียบเทียบ กับเกณฑ์ที่กำหนดให้ การวัดผลชนิดนี้จะใช้ในสถานการณ์ที่ต้องการทราบถึงสมรรถภาพของมนุษย์ เช่น ผู้เรียนมีทักษะถึง เกณฑ์หรือไม่ โดยไม่มีเงื่อนไขเกี่ยวกับจำนวนคน เกณฑ์ในที่นี้หมายถึง กลุ่มของ พฤติกรรมที่ได้รับการกำหนดไว้ล่วงหน้าอย่างชัดเจน พฤติกรรมต่าง ๆ ภายใต้โคเมนหนึ่ง ๆ จะมี ลักษณะเป็นเอกพจน์กัน ในการให้ความหมายของการวัดผลแบบอิง เกณฑ์ในลักษณะนี้ เกลเซอร์ และ นิตโก (Glaser and Nitko. 1971 : 653) กล่าวว่า แบบทดสอบอิง เกณฑ์สร้างขึ้นเพื่อให้ได้ผลจากการวัดที่แปลความหมายว่า ผู้สอบสามารถปฏิบัติอะไรได้บ้างหรือทำข้อสอบได้ถึงระดับใด เมื่อเปรียบเทียบกับมาตรฐานของการปฏิบัติที่กำหนดเกณฑ์ไว้ล่วงหน้า มาตรฐานการปฏิบัติก็คือการ ระบุนิยามของบทเรียนหรือกลุ่ม กลุ่มของงาน (Set of Tasks) ที่นักเรียนสามารถปฏิบัติงานได้ การวัดผลเป็นการนำเอากลุ่มของงานอ้างอิงไปยังขอบเขตเหล่านั้น ส่วน แฮริส และสจวร์ท

(Harris and Stewart. 1971 : 3) ให้ความหมายของแบบทดสอบอิง เกณฑ์ในทำนองคล้ายกัน โดยกล่าวว่า แบบทดสอบอิง เกณฑ์แท้จริงประกอบด้วย รายชื่อที่มุ่งให้เป็นตัวแทนของโคเมนหรือ ขอบเขตเนื้อหาทั้งหมดที่ได้รับการกำหนดไว้อย่างชัดเจน ในการทดสอบต้องใช้จำนวนรายชื่อที่ผู้สอบ

ตอบถูกในกลุ่มตัวอย่างพฤติกรรมนี้ นำมาประมาณค่าของอัตราส่วนที่แท้จริงของการทำถูก
เมื่อเทียบกับพฤติกรรมทั้งหมดในขอบเขตนั้น

ฮานโลเวนส์ (Ivens, 1970) กล่าวถึงแบบทดสอบอิงเกณฑ์ว่า เป็นแบบทดสอบ
ที่ประกอบด้วยรายชื่อย่อย ๆ ที่สร้างขึ้นเพื่อวัดจุดประสงค์เชิงพฤติกรรมต่าง ๆ ที่กำหนดไว้ และ
เป็นอิสระจากการทดสอบความแปรปรวน (สงวนลิขสิทธิ์ 2523 : 2) ซึ่งแฮมเบตัน และ
โนวิก (Hambleton and Novick, 1973 : 161) ก็กล่าวถึงลักษณะของแบบทดสอบ
อิงเกณฑ์ในทำนองเดียวกัน และกล่าวถึงปัญหาเบื้องต้นในการวัดผลแบบอิงเกณฑ์คือ การจำแนก
เด็กเป็นหนึ่งในการแสดงอาการของการรอบรู้ ซึ่งแทนด้วยระดับการปฏิบัติที่แตกต่างกันใน
ขอบเขตของพฤติกรรมแต่ละจุดประสงค์ในแบบทดสอบ กระนั้นผู้คิด หรือคะแนนที่ระดับความรู้
ที่จะนำมาแบ่งนักเรียนเป็น 2 กลุ่มคือ รอบรู้และไม่รอบรู้ บนพื้นฐานของปฏิบัติของแบบทดสอบ
แต่ละชุด

อเล็กซ์ ซีโรสก (อเล็กซ์ ซีโรสก 2522 : 256) กล่าวถึงแบบทดสอบอิงเกณฑ์
ว่าใช้สำหรับวัดความรู้ความสามารถของนักเรียนแต่ละคนว่าถึงเกณฑ์ขั้นต่ำที่กำหนดไว้หรือไม่
บางทีก็เรียกแบบทดสอบนี้ว่า Mastery Test แบบทดสอบนี้จึงเน้นการวัดความรู้และทักษะ
ต่าง ๆ ในตัวนักเรียนว่ามีถึงเกณฑ์ที่กำหนดไว้หรือไม่ โดยไม่คำนึงถึงความแตกต่างระหว่าง
นักเรียน ซึ่งสอดคล้องกับความหมายของแบบทดสอบอิงเกณฑ์ กมล ฐประเสริฐ (กมล
ฐประเสริฐ 2520 : 13) กล่าวว่า การประเมินแบบอิงเกณฑ์เป็นผลสืบเนื่องมาจาก
พฤติกรรมการเรียนเพื่อรอบรู้ของนักเรียน เป็นการประเมินโดยอาศัยการวัดเอาหลักเกณฑ์ภายนอก
มาเทียบ นักเรียนจะมีความสามารถอย่างใดอย่างหนึ่งหรือไม่นั้น จะต้องเปรียบเทียบผลการวัด
ของนักเรียนแต่ละคนกับเกณฑ์ โดยไม่คำนึงถึงผลการวัดของนักเรียนคนอื่น ๆ เครื่องมือที่ใช้
จะต้องสอดคล้องกับเกณฑ์ที่กำหนด และสอดคล้องกับความหมายที่สมบูรณ์ อิตทงส์
(สมบูรณ์ อิตทงส์ 2520 : 28) กล่าวถึงแบบทดสอบอิงเกณฑ์ว่า เป็นแบบทดสอบที่ต้องการ
วัดความรู้ความสามารถของบุคคลโดยอาศัยเกณฑ์ที่กำหนดไว้ในจุดมุ่งหมายของรายวิชานั้น ๆ ความหมาย
ของคะแนนที่ได้จากการสอบขึ้นอยู่กับความยากหรือไม่ง่ายของเกณฑ์ที่กำหนดไว้เป็นสำคัญ การสอบแบบนี้

เพื่อต้องการทราบว่าสิ่งใดบ้างที่เด็กยังทำไม่ได้และสิ่งใดบ้างที่เด็กทำได้แล้ว ไม่ต้องการทราบว่าเขาทำได้ดีกว่าคนอื่นมากนักน้อยเพียงใด เพราะการทำได้ก็หรือเร็วกว่าคนอื่นไม่ได้หมายความว่าเด็กทำสำเร็จหรือไม่สำเร็จในสิ่งนั้น และสมศักดิ์ สิ้นธุระเวช (สมศักดิ์ /! สิ้นธุระเวช 2522 : 1) ก็กล่าวถึงแบบทดสอบอิงเกณฑ์ในทำนองเดียวกัน แต่การแปลผลได้ 2 แบบคือ ผู้สอบปฏิบัติได้ตามเกณฑ์ที่กำหนดไว้หรือไม่ และการแปลผลอีกแบบหนึ่งโดยที่เราได้จำกัดขอบเขตของโคเมนไว้ เช่น นักเรียนสะกดคำได้ 8 คำ จาก 10 คำ ซึ่งสุ่มมาจากคำให้สะกดในชั้นเรียนมี 6 ประมาว่าเขาสะกดได้ 80 % ของคำทั้งหมดซึ่งเป็นการแปลผลของโคเมน

✳ สำหรับ สบม ลักษณะ (สบม ลักษณะ 2523 : 1 - 2) ได้กล่าวถึงลักษณะของการทดสอบแบบอิงเกณฑ์ว่า คือการทดสอบที่สามารถนำคะแนนมาแปลความหมายได้ว่า ผู้สอบมีความสัมฤทธิ์ผลในการกระทำกิจกรรมอะไรได้บ้าง ลักษณะสำคัญของแบบทดสอบอิงเกณฑ์คือ

1. เป็นแบบทดสอบที่สร้างพื้นฐานของจุดประสงค์หรือสิ่งที่ต้องการให้ผู้สอบปฏิบัติที่ได้รับการนิยามไว้อย่างชัดเจน
2. รายชื่อของแบบทดสอบจะต้องวัดจุดประสงค์หรือสิ่งที่ต้องการให้ผู้สอบนั้นปฏิบัติได้
3. คะแนนที่ได้ควรแปลความหมายว่า ผู้สอบสามารถบรรลุตามเกณฑ์ที่กำหนดไว้หรือไม่

ส่วน โกวิท ประวาลพฤษ และ ส.วาสนา ประวาลพฤษ (โกวิท ประวาลพฤษ และ ส.วาสนา ประวาลพฤษ 2518 : 22 - 23) ให้คำจำกัดความของแบบทดสอบอิงเกณฑ์ไว้ 2 ประเด็น คือ

1. แบบทดสอบอิงเกณฑ์ คือ ข้อสอบที่เป็นกลุ่มของงานที่เลือกสรรมาเพื่ออธิบายกลุ่มของจุดมุ่งหมายในการสอบหนึ่ง ๆ แบบทดสอบอิงเกณฑ์จะประกอบด้วยการวัดพฤติกรรมหลาย ๆ อย่างที่แสดงออกมาถึงการบรรลุจุดมุ่งหมายข้อหนึ่ง

2. แบบทดสอบอิงเกณฑ์ คือ ข้อสอบที่ใช้ให้คะแนนว่านักเรียนนั้นได้บรรลุถึงขั้นใดในจุดมุ่งหมายที่ตั้งไว้ คำว่า "รอบรู้" หมายความว่าในระดับจุดมุ่งหมายหนึ่ง ๆ นั้นนักเรียนนั้นได้รู้ถึงขั้นไหน เป็นสัดส่วนหรือเป็นกี่เปอร์เซ็นต์ของจุดมุ่งหมายนั้น ซึ่งจะต้องอิงเกณฑ์ว่า นักเรียนนั้นรู้หรือแสดงพฤติกรรมใดขนาดไหนจึงจะถือว่า "รอบรู้"

โกวิท ประวาทพุก (โกวิท ประวาทพุก 2523 : 18 - 19) ได้สรุปลักษณะของแบบทดสอบอิงเกณฑ์ว่ามี 2 ประเภทคือ

1. การปัดจุดประสงค์เป็นหลัก หรือจะเรียกว่าอิงจุดประสงค์ก็ได้ กล่าวคือในการเรียนการสอนมีการกำหนดจุดประสงค์เชิงพฤติกรรม และการสอบวัดก็ยึดเอาจุดประสงค์เหล่านั้นเป็นเกณฑ์ เพื่อจะชี้แจงว่าใครผ่านเกณฑ์ที่กำหนดไว้ในจุดประสงค์แล้วหรือไม่

นักวิจัยที่สนใจในเรื่องนี้คือ ลินวอลล์ (Linvall) และนิตโก (Nitko) (สงวนลิขสิทธิ์ 2522 : 49)

2. การขยายจุดประสงค์ให้ครอบคลุมประชากรของพฤติกรรม ที่เรียกว่าโดเมน ข้อสอบต่าง ๆ ได้จากการสุ่มตัวอย่างของพฤติกรรมที่กำหนดจากโดเมน แบบทดสอบในลักษณะนี้เรียกว่า Domain - Referenced Test ซึ่งเมื่อสอบวัดแล้วสามารถอ้างถึงได้ว่า บุคคลนั้นมีความสามารถในระดับใดตามคะแนนมาตรฐานเนื้อหา หรือมีระดับความรู้ขนาดใด

นักวิจัยในขณะนี้ ได้แก่ แฮร์ริส (Harris) และสจวร์ต (Stewart) (สงวนลิขสิทธิ์ 2522 : 49)

ดังนี้จากความหมายของลักษณะแบบทดสอบอิงเกณฑ์ที่กล่าวมาแล้วนี้ พอสรุปได้ว่าแบบทดสอบอิงเกณฑ์คือ

1. แบบทดสอบวัดผลสัมฤทธิ์หรือประกอบด้วยการวัดค่าตามที่สร้างจากจุดประสงค์เชิงพฤติกรรม ซึ่งอาจเป็นมาจากการพิจารณาของพฤติกรรมหรือโดเมนที่เฉพาะเจาะจง
2. มีเกณฑ์ขั้นต่ำที่ตัดสินการผ่านจุดประสงค์เชิงพฤติกรรมของข้อสอบ
3. ความหมายของคะแนนได้จาก การเปรียบเทียบคะแนนของผู้สอบกับเกณฑ์ที่กำหนดโดยไม่ว่าถึงผลการสอบของนักเรียน ๆ

★ การเปรียบเทียบความแตกต่างระหว่างการวัดผลแบบอิงกลุ่มและอิงเกณฑ์

ได้มีผู้เปรียบเทียบความแตกต่างระหว่างการวัดผลทั้งสองแบบไว้ ดังนี้

ตาราง 1 เปรียบเทียบความแตกต่างระหว่างการวัดผลแบบอิงกลุ่มและอิงเกณฑ์

การวัดผลแบบอิงเกณฑ์	การวัดผลแบบอิงกลุ่ม
1. การประเมินผลเป็นการเปรียบเทียบ ความสามารถของนักเรียนกับพฤติกรรม เป็นหลักไม่ว่าถึงคะแนนของผู้อื่น ๆ	1. การประเมินผลเป็นการเปรียบเทียบความสามารถ ของนักเรียนกับคะแนนของผู้อื่น ๆ
2. ใช้เป็นแบบอิงกลุ่มได้	2. ใช้เป็นแบบอิงเกณฑ์ไม่ได้
3. เหมาะสมกับการเริ่มการสอน และ การประเมินความก้าวหน้า	3. เหมาะสมกับการคัดเลือกและการแยกประเภท
4. ประเมินทั้งตัวนักเรียนและวิธีสอน	4. ประเมินแต่เพียงตัวนักเรียน
5. คะแนนมีการกระจายค่อนข้างต่ำ	5. คะแนนมีการกระจายสูง
6. เขียนคำถามถามเนื้อหาและจุดประสงค์ ที่กำหนดไว้	6. มีการสุ่มตัวอย่างเนื้อหา ข้อสอบต้องมีความยาก พอเหมาะ
7. เน้นความเที่ยงตรงตามเนื้อหามากที่สุด	7. ความเที่ยงตรงทุกแบบมีความสำคัญ รวมทั้ง คุณลักษณะของคน ๆ ของข้อสอบ
8. คะแนนแปลความหมายออกมาในรูปของ ผ่าน - ไม่ผ่านเกณฑ์	8. คะแนนแปลความหมายออกมาในรูปของเปอร์เซ็นต์ หรือคะแนนมาตรฐานอื่น ๆ ด้วย
9. ใช้วัดระดับความรู้หรือทักษะ ความสามารถ	9. ใช้จำแนกนักเรียนเป็นกลุ่มเก่ง และกลุ่มอ่อน
10. ประเมินผลออกมาเป็นระบบทวีคูณ ผ่าน - ไม่ผ่าน แต่สามารถแปลเป็น ระบบเกรก	10. ประเมินผลออกมาเป็นระบบเกรก

(บุญเชิด วิทยุโชนันตพงศ์ 2521 : 29 - 31, ส.วาสนา ประมวลพจนานุกรม 2517 : 84)

การสร้างแบบทดสอบอิงเกณฑ์

สำหรับการสร้างแบบทดสอบอิงเกณฑ์นั้น มีผู้กล่าวถึงหลายคนด้วยกันซึ่งจะนำมาเป็นแนวทางในการสร้างได้ ดังนี้

แอร์ราเซียน และเมคัส (Airasian and Maduas. 1972 : 8) แบ่งขั้นตอนของการสร้างแบบทดสอบอิงเกณฑ์ออกเป็น 4 ขั้นตอน ดังนี้

1. เขียนสมรรถภาพที่จะวัดให้ชัดเจน โดยเขียนในรูปของพฤติกรรมที่สามารถสังเกตและสอบวัดได้
2. กำหนดเกณฑ์การรวมไว้ในสมรรถภาพนั้น
3. ศึกษาสถานการณ์เพื่อให้นักเรียนได้แสดงถึงความสามารถที่ต้องการวัด แล้วเขียนข้อสอบตามสถานการณ์นั้น
4. ตัดสินความรอบรู้ของนักเรียน โดยเทียบกับมาตรฐานการปฏิบัติขั้นต่ำสุดตามจุดประสงค์

เมโย (Mayo. 1970) กล่าวถึงขั้นตอนต่าง ๆ ของแบบทดสอบวัดผลสัมฤทธิ์โดยทั่วไป ดังนี้

1. ตั้งจุดมุ่งหมายที่จะใช้ข้อสอบวัด
2. เขียนข้อสอบตามเนื้อหาและวัดพฤติกรรมของจุดมุ่งหมาย
3. วิเคราะห์ข้อสอบเพื่อหาข้อคำถามที่มีความยากประมาณ 50 - 60 % และมีค่าอำนาจจำแนกให้สูงที่สุด
4. ตีความหมายของคะแนนโดยใช้กลุ่มเป็นหลักอ้างอิง

สำหรับการสร้างข้อสอบแบบการเรียนรู้เพื่อรอบรู้ขึ้น ขั้นแรกและขั้นที่สองเหมือนเดิม ส่วนขั้นที่สามข้อคำถามไม่ต้องการให้มีค่าอำนาจจำแนก ส่วนค่าความยากควรประมาณ 85 % (มีผู้ทำถูก 85 %) สำหรับขั้นที่สี่ในการตีความหมายของคะแนนก็ใช้เกณฑ์จากจุดมุ่งหมายไม่เกี่ยวข้องกับความสามารถของนักเรียนคนอื่น ๆ ในกลุ่ม (มฤตชาติ สุวรรณวงศ์ 2520 : 36 อ้างอิงมาจาก Mayo. 1970)

๒ สมศักดิ์ สินธุระเวชญ์ (สมศักดิ์ สินธุระเวชญ์ 2521 : 79 - 80) กล่าวถึงลำดับขั้นในการสร้างแบบทดสอบอิงเกณฑ์ ดังนี้

1. เขียนรูบร่างหมายให้ชัดเจนในรูปของรูปประสงค้เชิงพฤติกรรม
2. เตรียม Content Outline ที่จะออกข้อสอบ โดยมีวิธีการดังนี้

ในกรณีข้อสอบของงานที่ให้ทำหรือชิ้นของงานที่เรียนอยู่ในวงจำกัด เช่น "การบวกเศษส่วนที่มีส่วนเท่ากัน" การทำการร่างวิเคราะห์เพื่อสร้างข้อสอบก็ไม่จำเป็น อาจใช้วิธีการง่าย ๆ โดยเพียงแต่เขียนรายละเอียดของงานแต่ละงาน ดังนี้

- 1) บวกเศษส่วน 2 จำนวน ที่มีส่วนเท่ากัน ค่าตอบน้อยกว่าหนึ่ง ($\frac{1}{3} + \frac{1}{3}$)
- 2) บวกเศษส่วน 2 จำนวน ที่มีส่วนเท่ากัน ค่าตอบเท่ากับหนึ่ง ($\frac{1}{3} + \frac{2}{3}$)
- 3) บวกเศษส่วน 2 จำนวน ที่มีส่วนเท่ากัน ค่าตอบมากกว่าหนึ่ง ($\frac{4}{6} + \frac{4}{6}$)
- 4) บวกเศษส่วนมากกว่า 2 จำนวน แต่ละจำนวนมีส่วนเท่ากัน เวลาตอบตอบเป็นเศษส่วนอย่างต่ำ

จากการที่เขียนรายละเอียดของงานแต่ละขั้นนี้ ก็จะทำให้เขียนข้อสอบได้ ซึ่งแต่ละงานจะใช้ข้อสอบประมาณ 4 ถึง 6 ข้อ

แต่ถ้าในกรณีชิ้นส่วนของงานต้องครอบคลุมพฤติกรรมใหญ่ เช่น "การบวกเศษส่วน" ก็จำเป็นต้องสร้างตารางวิเคราะห์

3. สร้างข้อคำถามให้สอดคล้องกับ Content Outline โดยทดสอบเฉพาะส่วนที่สำคัญ ๆ

4. ตรวจสอบแบบทดสอบที่สร้างว่าสอดคล้องตามรูบร่างหมายที่กำหนดไว้หรือไม่

นอกจากนี้สมศักดิ์ สินธุระเวชญ์ ได้กล่าวถึงแบบทดสอบอิงเกณฑ์อีกว่า ตัวคำถามต้องชัดเจนไม่คลุมเครือ เพราะต้องการให้เห็นว่าผู้สอบทำได้หรือไม่แล้ว แผนการต่าง ๆ ต้องเขียนให้ชัดเจนเช่นกัน ไม่ว่าจะเป็นตารางวิเคราะห์รูบร่างหมายหรือรูปแบบของข้อสอบ เพราะสิ่งต่าง ๆ เหล่านี้จะทำให้การสร้างข้อสอบเป็นไปอย่างมีประสิทธิภาพ (สมศักดิ์ สินธุระเวชญ์ 2522 : 1)

โกวิท ประวาลพุกษ์ (โกวิท ประวาลพุกษ์ 2522 : 124 - 149) ได้เสนอแนะ
วิธีการสร้างแบบทดสอบอิงเกณฑ์ ดังนี้

1. กำหนดจุดมุ่งหมายทั่วไปของการเรียนรู้ในเนื้อหาตอนหนึ่ง ๆ ถ้ามีตารางวิเคราะห์
หลักสูตรอยู่แล้วก็นำมาใช้ได้เลย พฤติกรรมที่พึงประสงค์จะครอบคลุมไปถึงพฤติกรรมขั้นสูง ๆ ด้วย
2. นำจุดมุ่งหมายทั่วไปมาเขียนเป็นจุดมุ่งหมายเชิงพฤติกรรม โดยกำหนดและจำแนก
เป็นพฤติกรรมปลายทาง พฤติกรรมเบื้องต้น และพฤติกรรมดำเนินการ
3. จัดสร้างข้อสอบตามจุดมุ่งหมายเชิงพฤติกรรมไว้มาก ๆ เพื่อจะสามารถทดสอบได้
ตลอดเวลา

4. จัดเนื้อหา เอกสาร และวัสดุประกอบการสอนที่เหมาะสมที่จะทำให้เกิดพฤติกรรม
5. ทำการทดสอบผู้เรียนเพื่อตรวจสอบดูว่า นักเรียนยังต้องเรียนเสริมก่อนเริ่มสอน
บทเรียนที่กำหนดไว้ ถ้าผู้เรียนยังไม่พร้อมก็ควรมีการฝึกฝนเพิ่มเติมพิเศษ หรือสอนให้คำดังมา
จากเดิม

6. ทำการสอนโดยใช้แบบทดสอบระหว่างการสอน เข้ามาช่วยในการตัดสินใจ
ดำเนินการให้มากที่สุด เช่นจะสอนซ้ำหรือไม่

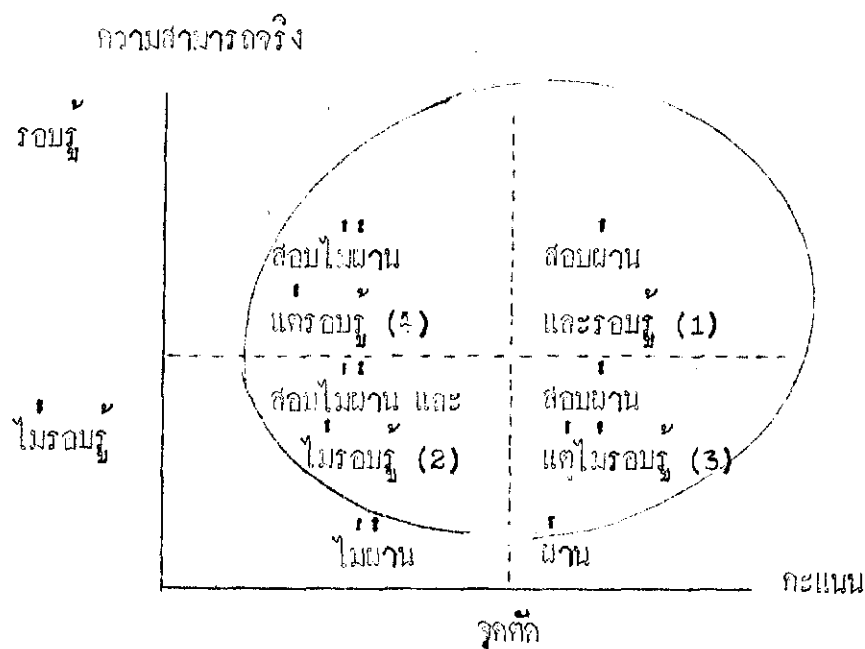
7. ทดสอบเมื่อสอนสิ้นสุดลง การทดสอบระยะนี้การจักกระทำเมื่อการทดสอบแต่ละตอน
สิ้นสุด ผู้เรียนมีโอกาสดูจะแสดงความสามารถตามขั้นสุดท้ายได้มาก

8. เมื่อผู้เรียนไม่สามารถผ่านแบบทดสอบสุดท้ายนี้ได้ ถือว่าเป็นจะต้องมีการซ่อมเสริม
บทวน กระบวนการเรียนการสอนสิ้นสุด เมื่อผู้เรียนทุก ๆ คนได้แสดงความสามารถตาม
จุดมุ่งหมายได้ตามเกณฑ์ที่กำหนดไว้

การกำหนดคะแนนเกณฑ์ในแบบทดสอบอิงเกณฑ์

แกสส์ (Glass, 1978 : 237 - 261) และแฮมเบิลตัน (Hambleton, 1978 :
277 - 288) กล่าวถึงคะแนนเกณฑ์หรือคะแนนจุดตัดว่า เป็นมาตรฐานการปฏิบัติหรือความสามารถ
ขั้นต่ำของผลการสอนแบบนาน-ไม่นาน นักวิจัยที่สนใจจะหาคะแนนจุดตัด (Cut-off Score)

เพื่อจำแนกผู้สอบออกเป็น 2 ประเภทคือ ผู้ที่ผ่านเกณฑ์กับผู้ที่ไม่ผ่านเกณฑ์ ได้แก่ เกลเซอร์ (Glaser, 1963 : 520) เมเจอร์ (Mager, 1962 : 44) และ ป็อปแฮม และ ฮุซเซก (Popham and Husek, 1969 : 2) ซึ่งส่วนใหญ่คะแนนเกณฑ์เหล่านี้อาจอยู่ในรูปเปอร์เซ็นต์ของจำนวนข้อที่ถูกต้องหรือถูกต้องในแต่ละจุดประสงค์ การกำหนดคะแนนเกณฑ์ทำหรือสูงไปจะทำให้เกิดความคลาดเคลื่อนได้ทั้งภาพประกอบ 1



ภาพประกอบ 1 แสดงความคลาดเคลื่อนของการประเมิน (สงวนลิขสิทธิ์ 2522 : 50)

พื้นที่ในวงรีสมมติแทนการกระจายของผู้สอบ แบ่งเป็น 4 ส่วนคือ

- (1) คนที่สอบผ่านและมีความสามารถจริง การเรียนต่อระดับสูงขึ้น
- (2) คนที่สอบไม่ผ่านและไม่มีความสามารถ การซ่อมเสริม
- (3) คนที่สอบผ่านแต่ไม่มีความสามารถตามจุดประสงค์
- (4) คนที่สอบไม่ผ่านแต่มีความสามารถจริงตามจุดประสงค์

ในส่วนที่ (3) และ (4) เป็นความคลาดเคลื่อนในการประเมิน ซึ่งส่วนที่ (3) เกิดจากการกำหนดคะแนนเกณฑ์ต่ำไป ทำให้นักเรียนที่ไม่มีคุณภาพสอบผ่านไปได้ และในส่วนที่ (4)

เนื่องจากการกำหนดคะแนนเกณฑ์สูงไป นักเรียนบางคนทำให้เสียเวลาและกำลังใจ นักวัดผลบางท่านได้ถือเอาการลดความคลาดเคลื่อนนี้เป็นคุณสมบัติที่เรียกว่า ความเชื่อมั่นของแบบทดสอบยิ่งเกณฑ์ มีการใช้สถิติเพื่อประมาณค่าคะแนนจริง และถูกต้องเพื่อให้เกิดความคลาดเคลื่อนน้อยที่สุด นอกเหนือไปจากการกำหนดคะแนนเกณฑ์โดยวิธีวิธีการปฏิบัติอื่น ๆ

ในการกำหนดคะแนนเกณฑ์หรือคะแนนจุดตัด เพื่อใช้ในการตัดสินการสอบผ่าน โดยมีผู้กล่าวไว้หลายวิธี ดังนี้

1. กำหนดไว้อย่างหนา เช่น ทองทำได้ 85 % แต่เป็นวิธีการที่ไม่เหมาะสม เพราะขาดพื้นฐานที่มาของตัวเลขเหล่านั้น (สงม จักรณะ 2522 : 51)

2. กำหนดโดยผู้เชี่ยวชาญร่วมกันพิจารณา ได้แก่ ความเห็นของครูผู้สอน นักพัฒนาหลักสูตร ผู้มีความรู้ทางเนื้อหาวิชาอื่น ๆ แล้วดูคะแนนส่วนใหญ่ วิธีทำให้ได้ข้อมูลเชิงประจักษ์การใช้บุคคลหลายคนทำให้เกิดความเชื่อถือสูง แต่ของระวังการใช้ตัวชี้วัด (Buton. 1978 : 267 - 268 และสงม จักรณะ 2522 : 53) นอกจากนี้ แฮมเบิลตัน (Hambleton. 1978 : 286) เห็นว่าการให้คะแนน ผู้บริหารโรงเรียน และนักเรียนระดับมัธยมศึกษาตอนปลายขึ้นไปร่วมพิจารณาค้น โดยคะแนนที่ได้จากแต่ละกลุ่มนำมาพิจารณาความสำคัญของแบบทดสอบและจำนวนรายละเอียดของโคเมนด้วย

3. กำหนดโดยใช้ทฤษฎีการเรียนรู้ของห้อง วิธีนี้คือการเรียนรู้จะมีกลุ่มของงานที่ต้องกระทำต่อเนื่องกัน การสอบเพื่อค้นหาจุดแบ่งกลุ่มของงานที่ทำได้กับทำไม่ได้ แต่วิชาไมอาจแบ่งเป็นขั้นตอนต่อเนื่องของการเรียนรู้ได้ง่าย ๆ ยังต้องมีการวิจัยตรวจสอบว่า คะแนนจุดตัดนั้นลดความคลาดเคลื่อนของการจำแนกผิดได้หรือไม่ (Buton. 1978 : 266 - 267 และ สงม จักรณะ 2522 : 52)

4. กำหนดโดยการแบ่งงานปฏิบัติของนักเรียน ๆ มาเป็นเกณฑ์ อ้างอิงจากประชากรของผู้สอบ เช่น คะแนนมัธยมฐานจากคะแนนการสอบและเกณฑ์ปกติ (Norm) แล้ววิธีนี้ไม่สอดคล้องกับหลักการวัดผลแบบอิงเกณฑ์ ซึ่งจะเปรียบเทียบความสามารถของผู้สอบกับนักเรียนคนอื่น นักทฤษฎีอิงเกณฑ์พิจารณาว่าวิธีนี้ไม่เหมาะสม (Glass. 1978 : 243 - 244 และ Millman. 1973 : 205 - 216)

5. การหาคะแนนเกณฑ์โดยการนับถอยหลังจาก 100% ผู้เขียนข้อสอบตั้งระดับการปฏิบัติเป็น 100% แต่เนื่องจากอาจมีความคลาดเคลื่อนทางการเรียนรู้ทางสติปัญญา ความไม่สนใจ ฯลฯ ซึ่งยากที่เกณฑ์จะสมบูรณ์ได้จึงอนุญาตให้ลดเกณฑ์ลง 5 ถึง 20% หรือมากกว่า (Glass . 1978 : 244)

6. คะแนนเกณฑ์ที่ได้จากการเพิ่มเติมเกณฑ์อื่นาร่วมด้วย เป็นการตัดสินการผ่านเกณฑ์จากชุดของแบบทดสอบกับการระบุ " ความสำเร็จหรือความรอบรู้ " ภายนอก เช่น ใบรับรองอาชีพว่าเป็นผู้สามารถ (นักศิลปะ ทันตแพทย์ ช่างตัดผม ฯลฯ) คะแนนเฉลี่ยและการกระจายที่บุคคลเหล่านี้ทำได้จากแบบทดสอบนำมาศึกษา บางครั้งอาจทั้งเป็นคะแนนเกณฑ์ของแบบทดสอบได้ ปัญหาในวิธีนี้มี 2 ประการคือ ผลการสอบอาจไม่ให้ความจริงตามสภาพเกณฑ์ภายนอก ซึ่งผู้ที่มิใบรับรองอาชีพทำคะแนนน้อยกว่าเกณฑ์และผู้ผ่านเกณฑ์ในแบบทดสอบแต่ไม่มีคุณวุฒิ ตำแหน่งของคะแนนเกณฑ์ที่ไม่สอดคล้องกับคุณวุฒิภายนอกนี้ อาจแก้ไขโดยใช้ทฤษฎีการตัดสินใจของ " การจำแนกผิดทางบวก " และ " การจำแนกผิดทางลบ " (Glass . 1978 : 244 - 245)

7. คะแนนเกณฑ์ที่ได้โดยวิธีวิจัยปฏิบัติการ (Glass . 1978 : 253 - 257) หรือวิธีศึกษาผลที่ตามมา (Millman . 1973 : 205 - 216) โดยอาศัยความสัมพันธ์ระหว่างคะแนนในแบบทดสอบกับผลการปฏิบัติงาน ช่วงคะแนนใดที่ปรากฏผลงานให้ค่าสูงสุกก็ถือเป็นคะแนนเกณฑ์โดยดูจากกราฟ วิธีวิจัยปฏิบัติการนี้บล็อก (Glass . 1978 : 253-- 257 citing Block . 1972) นำมาทดสอบกับนักเรียนระดับชั้น 8 จำนวน 91 คนในวิชาเมตริกซ์ แบ่งนักเรียนเป็น 5 กลุ่มคือ กลุ่มควบคุม (5%), 65%, 75%, 85% และ 95% พบว่าสำหรับแบบทดสอบอิงเกณฑ์ที่วัดความสามารถทางสมองควรมีคะแนนเกณฑ์ 95% และที่วัดทางจิตใจควรมีเกณฑ์ 85% นอกจากนี้ มิลแมนยังกล่าวว่า การกำหนดคะแนนเกณฑ์การผ่านต้องคำนึงถึงประโยชน์ในด้านการศึกษามากที่สุด ถ้าทักษะใดเป็นพื้นฐานที่สำคัญควรตั้งเกณฑ์ให้สูง และสงบ ลักษณะ

กล่าวถึงผลที่ตามมาของก็คือ มุ่งเรียนสอนทางใด วิชาเดียวกันอาจมีจุดติดค้างกัน

(สงม. ลักษณะ 2522 : 54)

8. คะแนนเกณฑ์ได้จากความสามารถขั้นต่ำที่นักเรียนควรจะได้ตามรูปประสงค์การเรียนรู้ ซึ่งเนลสัน และอิลเบอ เสนอวิธีการไว้ เนลสัน (Glass, 1978 : 247 - 251 citing Nedelsky, 1954) กำหนดคะแนนการสอบตามขั้นต่ำของแบบทดสอบปรนัยชนิดเลือกตอบ ที่ว่าเลือกคำตอบซึ่งนักเรียนที่คะแนนต่ำสุดไม่เลือกเรียกว่า F-responses คะแนนการบ้านขั้นต่ำสุดในแบบทดสอบ (D) ได้จากนักเรียนที่ไม่เลือก F-responses แต่เลือกคำตอบที่เหลือ (F-D students) ดังนี้

$$D = M_{FD} + k \sigma_{FD}$$

เมื่อ M_{FD} คือ คะแนนเฉลี่ย F-D students ได้คะแนนจากแบบทดสอบ หรือคะแนนเฉลี่ย

σ_{FD} คือ ความเบี่ยงเบนมาตรฐานของนักเรียนกลุ่มนี้

k คือ ค่าคงที่ใดจากการพิจารณาหลาย ๆ ครั้ง

k ที่กำหนดเป็น -1, 0, 1, 2 นั้นได้แก่ 16 %, 50 %, 84 % และ 98 % ของ F-D students ที่ถูกต้องมีค่า k เพราะ F-responses ในการสอบแปรเปลี่ยนได้ 2 แบบคือ เป็นตัวจริงที่แทนที่กับตัวลวงธรรมดา ค่า k ที่สูงจะสอดคล้องกับประเภทแรก ปกติค่า k มีช่วงระหว่าง .5 ถึง 1.5 ถ้าเท่ากับ 1 จัดเป็นปกติและมีค่าต่ำมากเมื่อเป็น 0

สำหรับวิธีการของอิลเบอ (Glass, 1978 : 247 - 248 citing Ebel, 1972) กล่าวถึงจุดบกพร่องของคะแนนการบ้าน โอกาสที่ผู้สอบจะผ่านหรือไม่ขึ้นอยู่กับข้อสอบในแบบทดสอบ มากกว่าความสามารถจริงของผู้สอบ ข้อสอบอาจยากหรือมีค่ากำหนดทางจากพื้นฐานแบบทดสอบทั้งใจไว้ อิลเบอจึงใช้วิธีพิจารณาจากความสำคัญ และความยากของข้อสอบ โดยแบ่งความสำคัญเป็น 4 ประเภทคือ ประเด็น ความหมาย การยอมรับและลักษณะคำถาม แบ่งระดับความยากเป็น 3 ระดับคือ ง่าย ปานกลาง และยาก ผู้สร้างข้อสอบให้เปอร์เซ็นต์

การคาดหวังระดับความยากของแต่ละประเภทไว้ จะแนะนำการผ่านไต่จาก ผลรวมที่ได้จาก จำนวนข้อสอบแต่ละประเภทดูกับความสำเร็จที่คาดหวัง แล้วนำมาหารด้วยจำนวนข้อ

แอมบรูว์ และเฮชท์ (Glass. 1978 : 249 citing Andrew & Hecht. 1976) ได้ศึกษาเปรียบเทียบวิธีการของเนลสันและอเบิลจากข้อสอบปรนัย 180 ข้อ พบว่า คะแนนเกณฑ์ของเนลสันคือ 49 % ส่วนของอเบิลคือ 68 % ซึ่งประมาณ 95 % ของผู้สอบ ทำข้อสอบผ่านโดยใช้เกณฑ์ของเนลสัน มีเพียง 50 % ของผู้สอบที่ผ่าน โดยใช้เกณฑ์ของอเบิล นอกจากนี้แฮมเบิลตัน (Hambleton. 1978 : 288 - 289) แนะนำว่าวิธีการของเนลสัน และอเบิลควรทดลองกับหลาย ๆ กลุ่มตัวอย่าง ความแตกต่างระหว่างกลุ่มนำมาแก้ปัญหา

ส่วนแองกอฟฟ์ เสนอวิธีการหาเกณฑ์คล้ายคลึงกับอเบิล แต่ไม่นำความยากของข้อสอบ มาเกี่ยวข้อง โดยใช้ความน่าจะเป็นที่นักเรียนสอบผ่านในชั้นต่ำสุดในแบบทดสอบ โดยให้กลุ่ม พิจารณารายข้อถึงความน่าจะเป็นในการทำถูกมากน้อยเพียงใด (มีค่าระหว่าง 0 ถึง 1)

นำค่าเฉลี่ยมาเป็นจุดตัดของแบบทดสอบ (Millman. 1973 : 210 citing Angoff.

1971)

9. กำหนดคะแนนเกณฑ์โดยใช้หลักการตัดสินใจ โดยใช้เกณฑ์ภายนอกจัดแบ่งผู้สอบ เป็น 2 กลุ่ม เช่น ผู้จบปริญญา-ไม่จบปริญญา ลูกจ้าง-นายจ้าง ฯลฯ บุคคลทั้งสองกลุ่มนี้ จะจัดเป็นประเภท "ผ่าน" หรือ "ไม่ผ่าน" คะแนนจากการทำแบบทดสอบอิงเกณฑ์หาจากกำหนด จากผู้เชี่ยวชาญ (วิธีการกำหนดเกณฑ์ในข้อ 2) จากการสอบได้บุคคลเป็น 4 พวก ดังตาราง 2

ตาราง 2 แสดงความน่าจะเป็นของความสัมพันธ์ระหว่างคะแนนจากแบบทดสอบกับเกณฑ์ภายนอก

		เกณฑ์ภายนอก	
		ผ่าน	ไม่ผ่าน
แบบทดสอบ	ไม่ผ่าน	P_A	P_B
	ผ่าน	P_C	P_D

- P_A เป็นสัดส่วนของการจำแนกผิดทางลบ คือผู้ผ่านเกณฑ์ภายนอกแต่ไม่ผ่านเกณฑ์
 ในแบบทดสอบ
- P_B เป็นสัดส่วนของผู้ที่ไม่ผ่านเกณฑ์ภายนอกและเกณฑ์ในแบบทดสอบ
- P_C เป็นสัดส่วนของผู้ผ่านเกณฑ์ภายนอกและเกณฑ์ในแบบทดสอบ
- P_D เป็นสัดส่วนของการจำแนกผิดทางบวก คือผู้ที่ไม่ผ่านเกณฑ์ภายนอกแต่ผ่านเกณฑ์
 ในแบบทดสอบ

คะแนนเกณฑ์นี้อาจเปลี่ยนแปลงได้ตามความเหมาะสม ทำให้สัดส่วนทั้งสอง
 เปลี่ยนไปเรื่อย ๆ พิจารณาการตั้งคะแนนเกณฑ์ $(P_B + P_C)$ และ $(P_A + P_D)$

ดังนั้นฟังก์ชันของคะแนนเกณฑ์ต่ำสุด เมื่อ C_X เป็นคะแนนเกณฑ์คือ

$$f(C_X) = (P_A + P_D) / (P_B + P_C)$$

ถ้ามีค่าต่ำสุดจะได้อะไรคะแนนที่สมเหตุสมผล โดยที่สมมติให้การจำแนกผิดทางบวก
 และการจำแนกผิดทางลบเท่ากัน ในทางตรงกันข้าม ให้บุคคลที่ไม่ผ่านแบบทดสอบภายนอก
 ภายนอกเป็น A และบุคคลที่ผ่านแบบทดสอบแต่ไม่ผ่านเกณฑ์ภายนอกเป็น B ได้ฟังก์ชันต่ำสุด
 ของคะแนนเกณฑ์คือ

$$f(C_X) = (A P_A + B P_D) / (P_B + P_C)$$

ฟังก์ชันในข้อความ A, B ซึ่งพิจารณาจากสิ่งที่เกี่ยวข้อง เช่น การสอบผ่าน
 แบบทดสอบของนักเรียนออกจากโรงเรียนกลางคืน การสอบตกของนักเรียนที่จบการศึกษา
 แวน เดอร์ ลินเดน และเมลเลนเบิร์ก (Van Der Linden & Mollenbergh, 1977)
 ได้พัฒนาทฤษฎีการตัดสินใจโดยนักจิตศาสตร์มาใช้ ทั้งสองเห็นด้วยกับเกณฑ์มาตรฐานแบบทดสอบอิงเกณฑ์
 ถึง 80 % ของการทำข้อสอบได้ถูกต้อง (Glass, 1978 : 251 - 253)

งานวิจัยที่เฝ้าดูการตัดสินใจในการหาคะแนนเกณฑ์ได้แก่ แฮมเบิลตัน และอัลไจนา
 (Hambleton, Swaminatan and Algina, 1975 : 87 - 98) โดยใช้วิธีของเบย์

ทดสอบกับนักเรียน 30 คน ทำข้อสอบ 10 ข้อ โดยคำนึงถึงความคลาดเคลื่อนในการจำแนกผิดและประเภท กำหนด $\alpha = 1$, $\beta = 2$ ตามลำดับ พบว่าเมื่อคะแนนถูกตัดเป็น .90 ความคลาดเคลื่อนน้อยที่สุดเมื่อใหญ่ที่ใด 9 คะแนนผ่าน เมื่อถูกตัดเป็น .60 และ .80 ความคลาดเคลื่อนน้อยที่สุดเมื่อใหญ่ที่โดยน้อยกว่า 6 ข้อของเรียนใหม่ ส่วน 7 ถึง 9 ข้อ ต้องทบทวนและปฏิบัติให้ถูกต้อง

10. คะแนนเกณฑ์ได้จากผลการเปรียบเทียบนักเรียน 2 กลุ่ม คือ กลุ่มที่เรียนรู้อย่างดีกับกลุ่มที่เรียนรู้อย่างน้อย หรือกลุ่มที่รอบรู้เกี่ยวกับกลุ่มที่ยังไม่รอบรู้ ซึ่งนักเรียนที่เรียนแล้วควรมีนักเรียน "ผ่าน" มาก และกลุ่มนักเรียนที่ยังไม่เรียนควรจะมีนักเรียน "ไม่ผ่าน" มาก จึงจะทำให้เกณฑ์มีความเที่ยงตรงมากขึ้น (Hambleton. 1978 : 289) วิธีการนี้คล้ายกับข้อ 9 แต่ใช้สถานการณ์การเรียนแบ่งกลุ่มนักเรียน เมื่อทำแบบทดสอบอิงเกณฑ์แล้วจะได้นักเรียน 4 กลุ่ม ดังตาราง 3

ตาราง 3 แสดงความสัมพันธ์ระหว่างคะแนนจากแบบทดสอบกับสถานการณ์การเรียนรู้อย่างดี

แบบทดสอบ			
		ตอบถูก	ตอบผิด
สถานการณ์การเรียนรู้อย่างดี	รอบรู้	1 - B	B
	ไม่รอบรู้	α	1 - α

α คือ ความน่าจะเป็นที่นักเรียนไม่รอบรู้แต่ตอบถูก

β คือ ความน่าจะเป็นที่นักเรียนรอบรู้แต่ตอบผิด

เอมริค (Emrich. 1971 : 321 - 326) ได้ศึกษาคะแนนเกณฑ์โดยวิธีหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างสถานการณ์รอบรู้กับการตอบเป็น

$$\phi = \frac{1 - \alpha - B}{\sqrt{1 - (\alpha - B)^2}}$$

คะแนนเกณฑ์จากการคำนวณวิธีอยู่ในรูปเปอร์เซ็นต์ (k) เมื่อ RR ได้จากการ
ถอดกราฟที่สองของความเชื่อมั่นของแบบทดสอบ โดยสูตรของสเปียร์แมน-บราวน์

$$k = \frac{\log B / (1 - \alpha) + n^{-1} \log(RR)}{\log \alpha (1 - \alpha) \beta (1 - \beta)}$$

จากการคำนวณของอเมริกา ได้คะแนนเกณฑ์เหมาะสมคือ 53.5 % ซึ่งนักเรียนจะต้อง
ทำได้อีก 9 ข้อ จากข้อสอบที่ใช้ในการศึกษารั้งนี้ 16 ข้อ ก่อนาวิลคอกซ์ และแฮริส

(Wilcox and Harris. 1977 : 215 - 217) ได้วิเคราะห์วิธีการของอเมริกา ได้สูตร
สัมพันธ์ซึ่งสัมพันธ์กันไปจากสูตรเดิม และโดยเท่ากันเมื่อ $k = 1/2$ และ $\alpha = \beta = 0$
นั่นคือต้องมีคะแนนเกณฑ์เป็นครึ่งหนึ่งและความคลาดเคลื่อนไม่มีการวัด ความหมายของการเรียนรู้
ได้ทั้งหมดหรือไม่โดยของอเมริกา ต้องนิยามคะแนนวิธีด้วยวิธีของลอว์ และโนวิก มากกว่า
จะใช้การนิยามแบบธรรมดา

ส่วนการหาคะแนนเกณฑ์โดยใช้กลุ่มนักเรียนที่ไม่ได้รับการสอนและนักเรียนที่ได้รับการสอนแล้ว
คือวิธีการของเบอร์ก (Berk. 1976 : 4 - 9) การกระจายคะแนนของ 2 กลุ่ม จะถาม
เกี่ยวกับ จุดที่ฟังก์ชันทั้งสองตัดกันคือคะแนนพยากรณ์จะแบ่งการเรียนรู้เป็น 4 พวก ดังตาราง 4

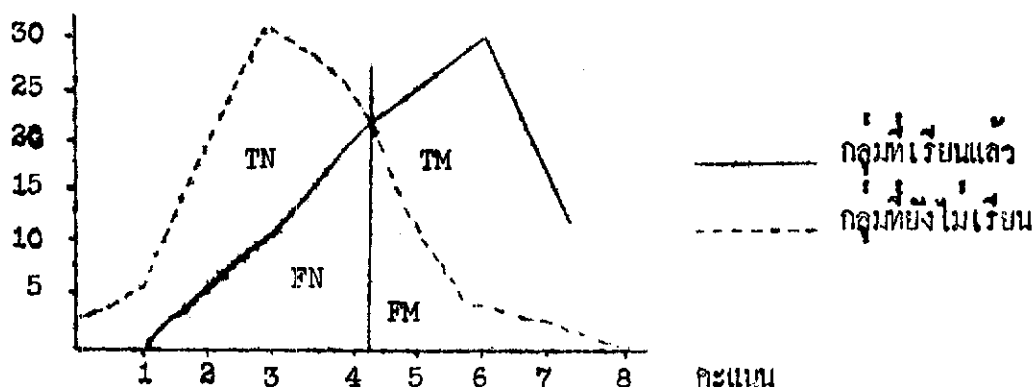
ตาราง 4 การจำแนกเกณฑ์และคะแนนพยากรณ์ของนักเรียน 2 กลุ่ม

		การจำแนกเกณฑ์	
		ได้รับการสอน	ไม่ได้รับการสอน
คะแนนพยากรณ์	รอบรู้	รอบรู้จริง (TN)	รอบรู้เท็จ (FM)
	ไม่รอบรู้	ไม่รอบรู้เท็จ (FN)	ไม่รอบรู้จริง (TN)

เบอร์ก ได้ศึกษาเกี่ยวกับนักเรียนเกรด 6 และนักเรียนเกรด 5 จำนวนกลุ่มละ 100 คน
ทำข้อสอบ 8 ข้อ วัดจุดประสงค์การเรียนรู้แล้ว การกระจายของคะแนนแสดงในภาพประกอบ 2
จุดตัดที่ได้เป็นคะแนนพยากรณ์แบ่งนักเรียนเป็น 4 พวก

จำนวนนักเรียน

คะแนนพยากรณ์



ภาพประกอบที่ 2 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายในกลุ่มที่เรียนแล้ว กับกลุ่มที่ยังไม่เรียน (Berk. 1978 : 6)

คะแนนจุดตัดแต่ละคะแนนนำมาหาค่าความน่าจะเป็นของการตัดสินใจถูกต้องจากสูตร $P(TM) + P(TN)$ และความคลาดเคลื่อนในการจำแนกประเภทที่ 1 (ผู้ที่เรียนแล้ว แต่สอบไม่ผ่านเกณฑ์) คือ $P(FN)$ และความคลาดเคลื่อนในการจำแนกประเภทที่ 2 (ผู้ที่ยังไม่เรียน แต่สอบผ่านเกณฑ์) คือ $P(FM)$

$$\text{เมื่อ } P(TM) = TM/(M+N)$$

$$P(TN) = TN/(M+N)$$

$$P(FM) = FM/(M+N)$$

$$P(FN) = FN/(M+N)$$

$$N = \text{จำนวนนักเรียนทั้งหมดที่ยังไม่เรียน}$$

$$M = \text{จำนวนนักเรียนทั้งหมดที่เรียนแล้ว}$$

คะแนนจุดตัดแต่ละคะแนนสามารถตรวจสอบความเที่ยงตรง โดยใช้สูตรสัมประสิทธิ์ความเที่ยงตรงของเกตต์พัฒนาจากรูปแบบของ Menemer เพื่อเลือกค่าสัมประสิทธิ์สูงสุดของความน่าจะเป็นในการตัดสินใจของคะแนนจุดตัดมาเป็นคะแนนเกณฑ์ สูตรการหาความเที่ยงตรงของเกณฑ์ได้จาก

$$\phi_{vc} = \frac{P(TM) - BR(SR)}{\sqrt{BR(1 - BR) SR(1 - SR)}}$$

เมื่อ ϕ_{VC}	คือ	ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์
BR	คือ	ความน่าจะเป็นของผู้น้อมน้อมในประชากร = $P(FN) + P(TM)$
SR	คือ	ความน่าจะเป็นของการพบการน้อมน้อมในประชากร = $P(PM) + P(TM)$

คะแนนถูกจัดที่เฉพาะสำหรับคะแนนเกณฑ์จากแบบทดสอบของเบอร์กลี 5 คะแนน ซึ่งมีความน่าจะเป็นของการจัดในเกณฑ์ถูกต้องเท่ากับ .74 ความคลาดเคลื่อนในการจำแนกนักเรียนประเภทที่ 1 และประเภทที่ 2 เท่ากับ .14 และ .12 ตามลำดับ ค่าสัมประสิทธิ์ความเที่ยงตรงของเกณฑ์คือ .48

นอกจากวิธีการ หาคะแนนเกณฑ์ทั้ง 10 วิธีแล้ว มิลแมน แนะนำว่าการพิจารณาจากหลาย ๆ ด้านประกอบกัน เช่น คะแนนของนักเรียน อาจเป็นคะแนนจากกลุ่มที่เรียนมาก่อนหรือกำหนดจำนวนนักเรียนที่จะสอบผ่านเนื้อหาวิชา โดยการนำสัดส่วนผู้สอบทำได้นักเรียนมารวมเป็นคะแนนเกณฑ์ ตลอดจนผลการเรียนที่ผ่านมา เช่น เป็นพื้นฐานของวิชาอื่น การมีส่วนร่วมและความคลาดเคลื่อนในการวัด (Millman, 1973 : 205 - 216) และแฮมเบิลตัน แนะนำถึงเนื้อหาวิชาเกี่ยวกับ การนำรายละเอียดของโคเมนร่วมอภิปรายตัดสินเกณฑ์ ถ้าถูกมุ่งหมายที่สำคัญและจำเป็นควรตั้งเกณฑ์ที่สูง ผู้พิจารณาเกณฑ์ควรมีความรู้ในการใช้แบบทดสอบและลักษณะของกลุ่มนักเรียน หากหลักสูตรและวิธีสอนมีการเปลี่ยนแปลง ต้องเปลี่ยนแปลงเกณฑ์การตัดสินด้วย (Hambleton, 1978 : 288 - 289)

จากหลักฐานเหล่านี้เห็นว่า วิธีหาคะแนนเกณฑ์หรือคะแนนถูกจัดทั้งวิธีที่เกี่ยวกับการปฏิบัติ โดยดูจากสิ่งประกอบต่าง ๆ และการพิจารณาของผู้เชี่ยวชาญ กับวิธีการหาคะแนนเกณฑ์โดยใช้ทฤษฎีและความรู้ทางคณิตศาสตร์ สงบ อภิณะ กล่าวว่า การปรับปรุงคะแนนเกณฑ์อยู่เสมอโดยนำผลต่อเนื่องของการตัดสินการบ้าน-โฮมวอร์ มาเป็นข้อมูลในการปรับปรุง เพื่อไม่ให้คะแนนเกณฑ์ค่าเรือสูงไปจนชาวภาควิชาและเด็กตกหนัก ๆ (สงบ อภิณะ 2522 : 53) ซึ่งสงบ อภิณะ ได้กล่าวถึงประโยชน์ของคะแนนเกณฑ์ในบทความเดียวกันว่า

1. เป็นการควบคุมมาตรฐานของการเรียนการสอนตามหลักสูตรอย่างแท้จริง นั่นก็ช่วยให้นักเรียนสามารถแยกการเรียนรู้ของเล็กได้เป็น 2 ประเภทใหญ่ ๆ คือ ผู้ที่เรียนรู้แล้วได้ปฏิบัติตามจุดประสงค์และผู้เรียนได้ไม่ครบ (Mastery - Non-Mastery)

2. ช่วยให้เราตรวจสอบได้ว่าเด็กคนใดมีความสามารถครบถ้วน และสามารถเรียนต่อไปตามลำดับขั้นการเรียนรู้

3. ใช้แยกประเภทนักเรียนออกเป็นประเภทย่อย ๆ ตามระดับปริมาณและคุณภาพของการเรียนรู้ เช่น ใช้เพื่อให้อะกับคะแนน A,B,C เป็นต้น

ในการศึกษาค้นคว้าวิจัยเกี่ยวกับวิธีการหาคะแนนเกณฑ์ของเบอร์ก (Berk. 1976 :

4 - 9) ร่วมกับการพิจารณาคะแนนเกณฑ์ของครูผู้สอน มาเป็นคะแนนเกณฑ์ของรายชื่อบุคคลประสงค์แจ้งพฤติกรรมในแบบทดสอบอิงเกณฑ์ของเบอร์ก เพราะเป็นวิธีการที่สอดคล้องกับการสอนนักเรียน 2 ครั้ง (ซึ่งใช้ในการวิเคราะห์รายชื่อบุคคล) โดยที่การสอบก่อนเรียนถือเป็นกลุ่มนักเรียนที่ยังไม่ได้รับการสอน (Uninstructed Group) การสอบหลังเรียนเป็นกลุ่มนักเรียนที่ได้รับการสอนแล้ว (Instructed Group) ผู้ทรงคุณวุฒิสามารถคำนวณหาความคลาดเคลื่อนในการจำแนกและสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์ที่หาได้ ในขั้นนี้จะอาศัยเกณฑ์ภายนอกเพื่อพิจารณาความเหมาะสมของคะแนนเกณฑ์ของการสอบ 2 กลุ่ม ที่ตามเกี่ยวกับที่แสดงในกราฟ นำมาหาคะแนนพยากรณ์และคำนวณหาค่าในสูตร คะแนนจุดตัดหรือคะแนนเกณฑ์ได้พิจารณาพร้อมกับความเห็นของครูผู้สอน เพราะมีประสบการณ์ในด้านการสอนและรู้จักนักเรียนของตนดี วิธีนี้แม้จะมีจุดอ่อนคือต้องระวังอคติด้วย แต่มิใช่จุดเด่นคือได้มาซึ่งข้อมูลเชิงประจักษ์ (Empirical Data) เราช่วยในการตั้งคะแนนเกณฑ์ และการใช้บุคคลหลายท่านทำให้เกิดความเชื่อถือสูง

๕ การวิเคราะห์รายชื่อบุคคล

ในแบบทดสอบอิงเกณฑ์ที่ครูผู้สอนหมายความว่า นักเรียนแต่ละคนมีความสามารถถึงเกณฑ์ขั้นต่ำที่ครูผู้สอนคาดหวังไว้หรือไม่ ในการคัดเลือกข้อสอบแบบอิงเกณฑ์ โดยการเปรียบเทียบระหว่างค่าคะแนนที่นักเรียนได้เรียนในเนื้อหา กับที่ไม่ได้เรียนในเนื้อหาที่เขาเคย หรืออาจใช้วิธีเปรียบเทียบระหว่างการสอบก่อนสอนกับการสอบหลังสอน เพื่อดูคะแนนความก้าวหน้าที่ได้ ซึ่งช่วยให้ทราบว่าวิธีสอนของครูได้ผลเพียงใด และแบบทดสอบที่ใช้วัดมีคุณภาพเพียงใด การทดสอบ

แบบอิงเกณฑ์ซึ่งไม่คำนึงถึงควา้จำแนกให้ใด แต่คำนึงว่าข้อสอบสามารถวัดจุดมุ่งหมาย ได้ชัดเจนเพียงใด และมีความเที่ยงตรงตามเนื้อหาเพียงใดมากกว่า (อนันต์ ศรีโสภณ 2522 : 256) ซึ่งสอดคล้องกับความเห็นของ แฮร์ริส และสจวร์ต (Harris and Stewart. 1971 : 13) ว่า เทคนิคการจำแนกและความยากของข้อสอบไม่สามารถนำมาไว้ใน แบบทดสอบที่มีโครงสร้างประเมินความรู้ในแบบทักษะ ซึ่งลักษณะการปฏิบัติได้รับการ นิยามที่แน่ชัด และความเห็นของป็อพแฮม และฮูเซก (Popham and Husek. 1969 : 1 - 9) ที่ให้ความสนใจต่อความจำแนกและความยากของข้อสอบแบบอิงเกณฑ์น้อยที่สุด อย่างไรก็ตามการวัดความจำแนกก็ยังคงให้ความสนใจ เพราะอาจเนื่องมาจากความบกพร่อง ของข้อสอบ ผลการสอบ หรือนักเรียนไม่ได้รับความรู้ที่ถูกต้อง

การความยากและความจำแนกเหล่านี้ เป็นค่าความยากและความจำแนก แบบเดิมที่ใช้ในแบบทดสอบเดิมอยู่ อย่างไรก็ตามได้มีการเสนอวิธีการวิเคราะห์ข้อสอบในรูปแบบ ของการวิเคราะห์แบบอิงเกณฑ์ แต่ก็ยังไม่ชัดเจนรูปแบบในเรื่องนี้ ส่วนใหญ่ต้องใช้การสอบ กับนักเรียน 2 ครั้ง ดังนี้

แนวทางการวิเคราะห์ความยากของข้อสอบ เบอร์ก (Berk. 1978 : 4) เสนอ วิธีการหาความยากโดยพิจารณาเปอร์เซ็นต์ของผู้ตอบถูกในรอบนั้น คำนวณค่าเฉลี่ยของข้อสอบ มีค่าตั้งแต่ 0 ถึง 100 เมื่อนักเรียนทั้งหมดตอบถูกค่านี้น่าจะเป็น 100 ถ้าทุกคนตอบผิดมีค่า 0 หากค่านี้น้อยกว่าค่าที่กลุ่มหนึ่งไม่ได้รับการสอน แสดงว่านักเรียนกลุ่มนั้นไม่มีความรู้เพิ่มขึ้น ข้อสอบทั้งภายในกลุ่มหลังเรียนอาจจะเป็นข้อสอบทั้งภายในกลุ่มก่อนเรียนก็ได้เช่นกัน ข้อสอบที่คิด การเป็นข้อสอบที่ยากในกลุ่มก่อนเรียนแต่ภายในกลุ่มหลังเรียน กอทซ์ และวาร์กาส (ส่งบ ลัษณะ 2523 : 41 อ้างอิงมาจาก Cox and Vargas. 1966) มีความเห็นว่า ถ้า ไม่มีการปรับเปลี่ยนความยากของข้อสอบในการสอบทั้ง 2 ครั้ง ถือว่าขาดความเที่ยงตรง ตามโครงสร้าง โดยทั่วไปข้อสอบชนิดนี้จะมีค่าความยากสูง (อนันต์ ศรีโสภณ 2522 : 256, กบด ภูประเสริฐ 2518 : 84 อ้างอิงมาจาก Mayo. 1970)

สำหรับแนวทางการวิเคราะห์ค่าอำนาจจำแนกของข้อสอบระบบอิง เกณฑ์ มีผู้เสนอ ดังนี้

การหาค่าอำนาจจำแนกจากสัดส่วนของการสอบก่อนเรียนและหลังเรียน Pretest - Posttest ไกลแก ลอกซ์ และวาร์กัส โดยคำนวณจากผลต่างของสัดส่วนผู้ตอบถูกในการสอบหลังเรียน กับสัดส่วนของผู้ตอบถูกในข้อนี้ในการสอบก่อนเรียน ซึ่งจะมีค่าพิสัยระหว่าง -1.00 ถึง 1.00 ข้อจำกัดของวิธีนี้คือ ถ้าชี้แนะไม่ให้เกิดความไวกับการแปรเปลี่ยนของการสอบรายบุคคล แต่เป็นผลจากความก้าวหน้าของกลุ่มผู้สอบทั้งหมด (Berk. 1978 : 5 citing Cox and Vargas. 1966) ส่วนโกรสทิน และเฟรดคูสัน ได้เสนอดัชนีเอส (S-index) ซึ่งวัดการคล้ายคลึงกับวิธีแรกมาก โดยคำนวณจากผลต่างของจำนวนผู้ตอบถูกในการสอบหลังเรียน กับจำนวนผู้ตอบถูกในกลุ่มก่อนเรียนในข้อนั้น แล้วหารด้วยจำนวนผู้เข้าสอบทั้งหมด ดัชนีเอสมีค่าพิสัยระหว่าง -1.00 ถึง 1.00 เช่นกัน ค่าที่เป็นลบแสดงว่ามีปัญหาเกี่ยวกับ 3 ประการคือ ข้อสอบง่ายมากเกินไปถูกในก่อนเรียนหรือข้อสอบยากเกินไปทุกคนในวิธีการเคาท์ในกลุ่มก่อนเรียนและหลังเรียน ข้อสอบค่าความแปรได้หลายนัย และประการสุดท้ายการสอนของครูอาจสร้างสิ่งกีดขวางให้แก่ผู้เรียน (สมศักดิ์ สันธุระเวช 2522 : 11 - 13 อ้างอิงมาจาก Krispin and Feldhuson. 1974) นอกจากนี้ รอนบูช (Rounbush) และโคสคอฟฟ์ และเคลน (Kosecoff and Klein) ได้เสนอค่าอำนาจจำแนกโดยยึดหลักคะแนนความก้าวหน้ารายบุคคล โดยที่รวมผลต่างจากสัดส่วนของผู้ตอบผิดในการเรียนก่อนเรียนแต่ตอบถูกในการสอบหลังเรียนกับจำนวนผู้เข้าสอบทั้งหมด แต่โคสคอฟฟ์ และเคลน (คำนวณจากสัดส่วนของผู้ตอบผิดในการสอบก่อนเรียนกับจำนวนผู้เข้าสอบทั้งหมด) ลบด้วยสัดส่วนของผู้ตอบผิดในการสอบทั้ง 2 ครั้ง กับจำนวนผู้เข้าสอบทั้งหมด ดัชนีทั้งสองนี้มีพิสัยตั้งแต่ -1.00 ถึง 1.00 ซึ่งดัชนีที่คำนวณได้โดยวิธีของโคสคอฟฟ์ และเคลน จะมีค่าต่ำกว่าสูตรของรอนบูช (Berk. 1978 : 5 citing Rounbush, 1973 and Kosecoff and Klein. 1974)

การหาค่าอำนาจจำแนกของการสอบกลุ่มหนึ่งไม่เรียนกับอีกกลุ่มที่เรียนแล้ว (Uninstructed-Instructor Group) ซึ่งเคลนและโคสคอฟฟ์ แนะนำโดยการคำนวณจากสัดส่วนของผู้ตอบถูก

ในกลุ่มที่เรียนเร็วกับจำนวนผู้เข้าสอบทั้งหมด ความสัมพันธ์ของผู้ตอบถูกในกลุ่มที่ยังไม่เรียน กับจำนวนผู้เข้าสอบทั้งหมด วิธีการนี้อาจคล่องกับวิธีของมาร์ตัน (Marton) ซึ่งมีข้อจำกัด เกี่ยวเกี่ยวกับวิธีของคอกซ์ และวาร์กาส (Berk. 1978 : 6 citing Klein and Kosecoff. 1976 and Marton. 1973)

✓ การหาค่าอำนาจจำแนกจากการผสมเพียงครั้งเดียว โดยหาความแตกต่างของกลุ่ม ผู้ตอบถูกและไม่ตอบถูก Mastery - Nonmastery) ซึ่งเบรนนัน (Brannan. 1972 : 289 - 303) ได้เสนอดัชนีบี (B-index) ซึ่งประยุกต์จากการวิเคราะห์รายข้อแบบอิงกลุ่ม โดยคำนวณจากผลต่างของสัดส่วนผู้ตอบถูกในกลุ่มสูงกับสัดส่วนของผู้ตอบถูกในกลุ่มต่ำ ผู้สร้าง ข้อสอบเลือกจุดกึ่งกลางระหว่างกลุ่มสูงและกลุ่มต่ำหรือกลุ่มที่ตอบถูกและกลุ่มไม่ตอบถูกได้อย่างอิสระ เพราะเชื่อว่าผู้สร้างแบบทดสอบอิงเกณฑ์หวังว่าผู้สอบส่วนใหญ่จะทำข้อสอบถูก การกระจายจะเบ้ ไม่หางอม วิธีการนี้อาจคล่องกับวิธีการหาค่าอำนาจจำแนก D % ของฮู (Hsu. 1971:34) ซึ่งดัชนีทั้งสองวัดความเที่ยงตรงของเกณฑ์จะกลายเป็นเงื่อนไขสำคัญของความเที่ยงตรงสถิติรายข้อ ดัชนีทั้งสองจะคำนวณได้โดยถ้านักเรียนทั้งหมดหรือไม่มีก็เรียนจนได้คะแนนถึงเกณฑ์

การหาค่าอำนาจจำแนกโดยใช้สูตรสัมประสิทธิ์สหสัมพันธ์และผลการถดถอยของกลุ่ม อ้างอิง (Criterion Group) 2 กลุ่ม อาจจะเป็นกลุ่มตอบถูก - ไม่ตอบถูก กลุ่มเรียนแล้ว - ยังไม่เรียนหรือกลุ่มก่อน - หลังเรียน ซึ่งได้แก่ วิธีของฮาธาไดนา (Haladyna. 1974 : 93 - 99) ใช้วิธีสัมประสิทธิ์สหสัมพันธ์หาค่าอำนาจจำแนกของคะแนนทั้งหมด ยังกลุ่มตัวอย่าง ที่ประกอบด้วยกลุ่มอ้างอิง 2 กลุ่ม (Combined group item-total r) ดัชนีี้สามารถ คำนวณกับโปรแกรมการวิเคราะห์รายข้อของแบบทดสอบอิงกลุ่ม เช่น ITEMAN และวิธีนี้ค่า สัมประสิทธิ์สหสัมพันธ์กับวิธีของคอกซ์ และวาร์กาส .64 ถึง .84 (Berk. 1978 : 6 citing Berk and Crisemer. 1976) ข้อจำกัดของวิธีนี้คือ ไม่มีความไวเมื่อคะแนน ของนักเรียนใกล้เคียงและต่ำมาก ซึ่งอาจจะไม่ได้วัดอำนาจจำแนกระหว่างกลุ่มอ้างอิงที่ระบุ ส่วนการเว้นต้น และบัพ (Millman. 1974 : 369 - 370 citing Dalrinton and Bishop. 1966) หาค่าอำนาจจำแนกโดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์รายข้อกับการจำแนก

กลุ่มอ้างอิง ขอบจำกัดของวิธีนี้คือ การใช้ข้อสอบกับการคำนวณที่ซับซ้อน ซึ่งต้องใช้โปรแกรม การวิเคราะห์รายข้อเพิ่มเติมหรือโปรแกรมที่มีประสิทธิภาพสูงพอสมควร ส่วนการวิเคราะห์ข้อ โดยไร้ซึ่งการทดลองในการหาคำอ่านจำแนก มิลแมนเป็นผู้เสนอโดยให้ข้อสอบเป็นตัวแปร พยายามและการจำแนกกลุ่มอ้างอิงเป็นตัวแปรแทน การวิเคราะห์รวมทั้งการหาสัมประสิทธิ์ สหสัมพันธ์ระหว่างข้อกับ ข้อจำกัดของวิธีนี้คือต้องใช้กลุ่มตัวอย่างขนาดใหญ่ ถ้าไม่ได้อ่านจำแนกผลการเลือกก็ย่อมสรุปผลการพยากรณ์ สัมประสิทธิ์ถดถอยรายข้อนำมาใช้ได้ น้ำหนักข้อสอบโดยไร้พิจารณาการวัดจุดประสงค์ของข้อสอบ (Millman. 1974 : 367 - 368)

ในการวิเคราะห์รายข้อได้มีผู้เสนอวิธีการหาความเป็นเอกพันธ์ของข้อสอบ โดยใช้ การสถิติ โปปแฮม ปอปแฮม (Popham. 1971 : 79 - 98) เสนอวิธีสถิติการกระจาย โดยนำเอาความถี่ของการสอบก่อนเรียน-หลังเรียนมาเปรียบเทียบกับความถี่มาตรฐาน (ความ คาดหวัง) ของข้อสอบ ความถี่มาตรฐานของข้อสอบที่ได้จากค่าการวัดรายข้อในจุดประสงค์ การสอนเกี่ยวกับเท่านั้น ข้อจำกัดของวิธีนี้คือ ไม่มีความถูกต้องนำมาใช้สำหรับข้อสอบที่มีปกติ ก่อมาเพื่อที่ และข้อทดสอบสหสัมพันธ์ (Prttie and Oosterhof. 1976) ได้พัฒนาวิธีการ ของปอปแฮม ด้วยการสอบเพียงครั้งเดียว (สอบก่อนเรียนเท่านั้น) โดยนำดัชนีความยากของ แต่ละข้อที่ได้จากจุดประสงค์ข้อสอบแล้ว มาเปรียบเทียบกับความยากมาตรฐานของชุดข้อสอบโดยใช้ การสถิติการกระจาย วิธีนี้มีข้อจำกัดเช่นเดียวกับวิธีการหาคำอ่านจำแนกของปอปแฮม สำหรับ แองเกิล และลาดาลา (Engel and LaDala. 1977) ได้พัฒนาวิธีการของลอง (Long. 1974) เพื่อหาความเป็นเอกพันธ์ของข้อสอบ โดยใช้การวัดประสิทธิภาพสัมพัทธ์ระหว่างข้อกับ คะแนนทั้งหมด ซึ่งคล้ายกับดัชนีอ่านจำแนกของข้อสอบ ข้อจำกัดของวิธีการแปลงเมื่อดัชนีนี้ มีค่า ๐ ยังไม่ทราบ ไม่ทราบสถิติการกระจายกลุ่มตัวอย่าง ดังนั้นไม่มีนัยสำคัญทางสถิติได้ (Berk. 1978 : 8 citing White and Saltz)

งานวิจัยที่เกี่ยวข้องกับการวิเคราะห์รายข้อของแบบทดสอบอิงเกณฑ์ ได้แก่
 ฮู (Hsu . 1971 : 85 - 91) ได้ศึกษาค้นคว้าจำแนกโดยใช้ค่าสถิติ 3 ชนิดคือ
 $D\%$, phi-coefficient (ϕ) และสัมประสิทธิ์สหสัมพันธ์พอยท์ไบซีเรียล r_{pbis}
 กับสถานการณ์ 3 แบบคือ การสอบก่อน-หลังเรียน การสอบกับนักเรียนต่างระดับชั้น
 (เกรด 3 - 4 และเกรด 2 - 3) และการจัดข้อสอบเข้ากลุ่มตามดัชนีความยาก 3
 ระดับ พบว่าดัชนีอำนาจจำแนกทั้ง 3 แบบมีค่าสัมประสิทธิ์สหสัมพันธ์สูงในเกือบทุกกรณี
 แต่ไม่มีความสอดคล้องของดัชนีทั้งสามชนิดในกลุ่มนักเรียนที่มีความสามารถแปรเปลี่ยนใน
 ช่วงกว้างเมื่อเปรียบเทียบกับกลุ่มสูงที่คัดเลือกไว้ ฮูไม่แนะนำให้ดัชนีทั้งสามใช้ในการ
 คัดเลือกข้อสอบ เพราะการคัดเลือกจากกลุ่มหนึ่งจะไม่สามารถวัดการปฏิบัติอย่างเดียวกัน
 ในกลุ่มอื่นๆได้ ในการศึกษาครั้งนี้ผู้ทดลองกับนักเรียนในโครงการคณิตศาสตร์ 5 จุดประสงค์

ต่อมาฮาลาดินาก็ได้ศึกษาการวิเคราะห์รายข้อแบบ $D\%$ และ r_{pbis}
 โดยให้กลุ่มตัวอย่างจำนวน 189 คนทำแบบทดสอบอิงเกณฑ์จำนวน 3 แบบ เกณฑ์ตัดสิน
 ร้อยละ 70 พบว่าดัชนีทั้งสองให้ค่าใกล้เคียงกัน ซึ่งสอดคล้องกับการศึกษาของฮู
 ฮาลาดินามีความเห็นว่ $D\%$ ตรงกับความหมายของแบบทดสอบอิงเกณฑ์มากกว่าแค่ใช้ในการ
 การสอบ 2 ครั้ง ส่วน r_{pbis} ใช้ในการทดสอบเพียงครั้งเดียวแต่มีประสิทธิภาพในการ
 ให้ความถูกต้องของข้อมูลรายข้อดีกว่า (Haladyna . 1974 : 93 - 97)

ในปีค.ศ. 1977 แทรช (Thrash . 1977 : 247 - A) เปรียบเทียบ
 ดัชนีอำนาจจำแนก 3 ชนิดคือ ดัชนีของคอกซ์และวาร์กาส (C-V) รอนนุส (R)
 และแบรนแนน-สจวร์ต (B-S) โดยศึกษาทั้งจุดประสงค์ 16 ข้อในวิชาคณิตศาสตร์
 และการอ่าน ทดลองกับนักเรียน 2 ระดับชั้นคือมัธยมและระดับที่สูงขึ้นไป และการ
 ทดลอง 2 แบบคือ แบบกำหนดจุดประสงค์ (A) กับแบบที่ให้นักเรียนเลือกจุดประสงค์
 (B) พบว่าดัชนีอำนาจจำแนกทั้งสามชนิดนำมาใช้ในวิชาคณิตศาสตร์ดีกว่าเมื่อเปรียบเทียบกับ
 การอ่าน การทดลอง B ได้ดีกว่าการทดลอง A ส่วนระดับชั้นให้ผลคล้ายกัน

สัมประสิทธิ์สหสัมพันธ์ระหว่างดัชนี C-V กับดัชนี R เท่ากับ .84 ดัชนี C-V กับดัชนี B-S
 เท่ากับ .70 และดัชนี B-S กับดัชนี R เท่ากับ .36 สำหรับกลุ่มตัวอย่างขนาดใหญ่

ดัชนี C-V ให้ค่าประมาณดีกว่าดัชนี R และในการวิเคราะห์รายข้อของการสอบก่อนเรียนและหลังเรียน แรชแมนแนะนำให้ใช้วิธีของคอกซ์และวาร์กาส

นอกจากนี้เวคแมน (Wedman. 1974 : 63) ศึกษาดัชนีอำนาจจำแนก 5 แบบกับนักเรียนมัธยมปลาย 285 คนและนักศึกษาครู 20 คนในวิชาทฤษฎีความน่าจะเป็น คือ ความแตกต่างของกลุ่มสอบก่อน-หลังเรียน (D:PP) ความแตกต่างของกลุ่มสูง-กลุ่มต่ำ 27% (D:27%) ความแตกต่างของกลุ่มผู้ชำนาญ-ไม่ชำนาญ (D:ENE) ความแตกต่างของกลุ่มสูงและกลุ่มต่ำ (O:13) และค่าสถิติโค-สแควของบอยแฮม พบว่าส่วนใหญ่มีค่าสัมประสิทธิ์สหสัมพันธ์ค่อนข้างต่ำ ที่นับว่าสูงได้แก่สัมประสิทธิ์สหสัมพันธ์ระหว่างสถิติโค-สแคว กับ D:27% , กับ D:ENE และกับ O:13 คือ .58, .52 และ .75 ตามลำดับ สัมประสิทธิ์สหสัมพันธ์ระหว่าง D:27% กับ D:PP คือ .62 ส่วนค่าสัมประสิทธิ์สหสัมพันธ์ของ D:27% กับ O:13 มีค่าสูงสุดคือ .91

สำหรับผู้ศึกษาเปรียบเทียบการวิเคราะห์รายข้อของแบบทดสอบอิงเกณฑ์ที่มีผลต่อค่าความเชื่อมั่นและค่าความเที่ยงตรงของแบบทดสอบคือ เครเฮน (Crehen. 1974 : 255 - 269) ศึกษาโดยใช้แบบทดสอบคู่ขนานและใช้วิธีการคัดเลือกข้อสอบ 6 วิธีคือ ดัชนีของคอกซ์และวาร์กาส ดัชนีของแบรนแนน วิธีการเรียงลำดับตามสัดส่วนการตอบของผู้สอบหลังเรียน เรียงตามการเลือกของครูผู้สร้างแบบทดสอบ เรียงตามค่าสัมประสิทธิ์พอยท์ไบซีเรียล และเรียงตามการสุ่ม พบว่าวิธีการทั้ง 6 วิธีให้ผลต่อค่าความเชื่อมั่นของแบบทดสอบไม่แตกต่างกัน ส่วนดัชนีอำนาจจำแนกของคอกซ์และวาร์กาส และดัชนีอำนาจจำแนกของแบรนแนนให้ค่าความเที่ยงตรงของแบบทดสอบสูงสุด

ในการศึกษาการสร้างแบบทดสอบอิงเกณฑ์ครั้งนี้ ผู้วิจัยได้เลือกดัชนีความยากของเบอร์ก (Berk. 1978 : 4) และดัชนีอำนาจจำแนกของโคสคอฟฟ์และเคลน (Kosecoff and Klein. 1976 : 276 - 293) ในการวิเคราะห์รายข้อของแบบทดสอบที่สร้างขึ้น เพราะแม้ว่าแบบทดสอบจะมีความเที่ยงตรงตามเนื้อหาตามความเห็นของผู้ออกข้อสอบ แต่ลักษณะข้อสอบอาจยากหรือง่ายเกินไปไม่เหมาะสมกับสภาพผู้สอบ จึงใช้ดัชนีความยากและดัชนีอำนาจจำแนกในการปรับปรุงและคัดเลือกข้อสอบ

ลักษณะอำนาจจำแนกให้ใช้กรณี เป็นสูตรที่เสนอผลความก้าวหน้ารายบุคคล โดยคำนวณจาก สัดส่วนของผู้ตอบผิดในการสอบก่อนเรียนและตอบถูกในการสอบหลังเรียน ควบคุมสัดส่วนของผู้ตอบผิดในการสอบทั้ง 2 ครั้ง ซึ่งสัดส่วนตัวแรกเป็นผลมาจากการเรียนรู้ที่เพิ่มขึ้น สัดส่วน ตัวหลังนำไปลบทิ้งเนื่องจากความคลาดเคลื่อนที่อาจเกิดจากการสร้างแบบทดสอบและ การเรียนรู้ ดังนั้นอัตราที่อำนาจจำแนกในสูตรนี้จึงมีค่าต่ำกว่าสูตรอื่น สูตรการหาค่าอำนาจ จำแนกในแบบทดสอบอิงระดับส่วนใหญ่เป็นการศึกษาคะแนนความก้าวหน้า แต่เป็นรูปแบบ ความก้าวหน้าเป็นกลุ่ม ซึ่งไม่มีความไวต่อการเปลี่ยนแปลงความก้าวหน้าของบุคคลเดียวกัน ในการสอบทั้ง 2 ครั้ง

ความเชื่อมั่นของแบบทดสอบอิงเกณฑ์

การหาความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ ได้มีผู้เสนอวิธีการประมาณค่าหลายวิธี แต่พอแบ่งกว้าง ๆ ได้ 2 ประเภทคือ ประมาณค่าความเชื่อมั่นจากการสอบครั้งเดียวและ การสอบ 2 ครั้ง ส่วนวิธีการของวิพังกัน อาจจัดอยู่ในประเภทใดก็ได้

วิพังกัน (Livingston, 1972 : 17 - 19) เสนอวิธีการหาความเชื่อมั่น โดยอาศัยแนวคิดบนถึงกลุ่มซึ่งคะแนนควรจะมีการกระจายกว้าง แต่แบบอิงเกณฑ์เมื่อนำไป เปรียบเทียบกับเกณฑ์คะแนนควรจะเกาะกลุ่มกัน วิพังกันได้พิจารณาความเข้มแข็งแบบคะแนน เป็นความเบี่ยงเบนจากเกณฑ์ไปจึงจะแตกต่าง โดย สูตรการหาความเชื่อมั่นของแบบทดสอบ อิงเกณฑ์ คือ

$$K_{tx}^2 = \frac{P_{tx}^2 \sigma_x^2 + (\mu_x - c_x)^2}{\sigma_x^2 + (\mu_x - c_x)^2}$$

เมื่อ K_{tx}^2

คือ ค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์

c_x

คือ คะแนนเกณฑ์

P_{tx}^2 , μ_x และ σ_x^2 เป็นตัวแปรของความเชื่อมั่น, ค่าเฉลี่ย และ

การกระจายของแบบทดสอบอิงกลุ่ม

เมื่อคะแนนเฉลี่ยเท่ากับคะแนนเกณฑ์ ค่าความเชื่อมั่นที่ได้จะเท่ากับค่าความเชื่อมั่น
ของแบบทดสอบซึ่งกลุ่มนั้นมีค่าต่ำสุด ถ้าคะแนนที่ได้จากแบบทดสอบเท่ากันผู้ทดสอบจะได้
การนิยาม และระยะห่างระหว่างคะแนนตัวหนึ่งห่างจากเกณฑ์มากเท่าใด คะแนนจริงจะ
ตกอยู่ในเกณฑ์เดียวกับเกณฑ์ ดังนั้นค่าความเชื่อมั่นที่ได้จึงมีความไวต่อช่วงห่างของคะแนนเฉลี่ย
และคะแนนเกณฑ์ วิธีที่ใช้ได้กับการกระจายของคะแนนฐานนิยมเดียว แฮร์ริส (Harris.
1972) ซึ่งเห็นว่าวิธีการของดีไวงั้นนี้ให้ความเชื่อมั่นที่สูงกว่าที่คำนวณจากวิธีการแบบเดิม
แต่ทั้งนี้ก็ให้ความคลาดเคลื่อนในการวัดเท่ากัน แฮร์ริสสรุปว่าสัมประสิทธิ์ที่มากกว่า
ไม่ได้แสดงถึงความแม่นยำของการกำหนดว่าคะแนนจริงจะตกอยู่ใต้หรือเหนือกว่าเกณฑ์หรือไม่
(ผจญจิต อันตพรพรณ 2523 : 48 - 49 อ้างอิงมาจาก Harris. 1972)

การหาความเชื่อมั่นของแบบทดสอบจากการสอบ 2 ครั้ง คือ วิธีการของสวามินาธาน
และคณะ (Swaminatan, Hambleton and Algina. 1974 : 263 - 267) โดยดู
สัดส่วนของนักเรียนที่ถูกกำหนดไว้อย่างคงเส้นคงวาในแต่ละประเทศ (บ้าน-ในมา) ของ
การเรียนรู้ ผู้เรียนเอนมาจากตนเองและตนเองแบบเปิดต้นและโนวิก (Hambleton &
Novick. 1973) ทว่าค่าของความเชื่อมั่นการสะท้อนถึงระดับที่นักเรียนถูกกำหนดในสภาวะ
เรียนรู้ของการเรียนรู้ จากผลการสอบ 2 ครั้งของชุดสอบยาวนาน ระดับความคงเส้นคงวา
วัดด้วยสัมประสิทธิ์ของความสอดคล้องระหว่างการสอบทั้งสอง

$$k = (P_0 - P_C) / (1 - P_C)$$

เมื่อ P_0 เป็นสัดส่วนของความสอดคล้องที่สังเกตได้จากการสอบซ้ำ
 P_C เป็นสัดส่วนของความสอดคล้องที่คาดหวัง

ค่าของ k ขึ้นอยู่กับสัดส่วนที่สังเกตและที่คาดหวังบนแนวคะแนนของเมตริกซ์ ซึ่งมีค่า
สูงสุดคือ 1.00 อาจเกิดขึ้นได้เมื่อสัดส่วนตรงปลายจากผลของการสอบแต่ละครั้งเท่ากัน อย่างไรก็ตาม
ถ้าผู้สอบคนใดถูกกำหนดในวิธีที่แตกต่างกันของการสอบทั้ง 2 ครั้ง ก็จะน้อยกว่า 1.00 ค่าที่ลดลง
จะไม่สนใจเพราะแสดงถึงความไม่คงเส้นคงวาในการตัดสินและไม่ได้ความเชื่อมั่นของแบบทดสอบ

ส่วนการหาความเชื่อมั่นของแบบทดสอบเชิงความถี่อาศัยการสอบครั้งเดียว ได้แก่

ฮวิน (Huynh. 1976 : 253 - 264) กำหนดหาความเชื่อมั่นของแบบทดสอบ โดยไร้ออกจากการสอบเพียงครั้งเดียว ซึ่งวิธีการเช่นเดียวกับแฮมเบิลตัน และโนวิก (Hambleton and Novick) โดยเป็นวิธีที่แสดงถึงความคงเส้นคงวาของการตัดสินใจจากการสอบซ้ำ และผลของผู้สอบหนึ่งที่จะมีความสามารถจริงถูกเปิดเผยด้วยการกระจายแบบไบนารี ซึ่งผู้สอบคงวางข้อสอบแต่ละชุดที่มีความยากพอ ๆ กัน ภายใต้ข้อสันนิษฐานเป็นอิสระของลักษณะ ฮวินเสนอสมการ 2 สมการสำหรับการคำนวณสัมประสิทธิ์ ดังนี้

$$K = \frac{P_{11} - P_1^2}{P_1 - P_1^2} \quad \text{เมื่อคะแนนเกณฑ์ค่าสูงใกล้เคียงคะแนนเต็ม}$$

และ

$$K = \frac{P_{00} - P_0^2}{P_0 - P_0^2} \quad \text{เมื่อคะแนนเกณฑ์ค่าต่ำใกล้ 0}$$

เมื่อ	P_{11}	เป็นสัดส่วนของผู้สอบที่ผ่านการสอบทั้ง 2 ครั้ง
	P_1	เป็นสัดส่วนของผู้สอบที่ผ่านการสอบเพียงครั้งเดียว
	P_{00}	เป็นสัดส่วนของผู้สอบที่ไม่ผ่านการสอบทั้ง 2 ครั้ง
	P_0	เป็นสัดส่วนของผู้สอบที่ไม่ผ่านการสอบเพียงครั้งเดียว

วิธีการของฮวินเป็นวิธีที่ขึ้นกับความรู้ทางสถิติในด้านการแจกแจงมาก มีความยุ่งยากในการคำนวณและซับซ้อนกว่า ถ้าจำนวนข้อสอบน้อยค่าที่ได้จะมีค่าความเชื่อมั่นไม่ดีเท่าที่ควร

ซึบโคเวียค (Subkoviak. 1976 : 265 - 276) เสนอวิธีการหาความเชื่อมั่นในการสอบครั้งเดียวที่ให้คะแนนเป็น 2 แบบ โดยมีข้อตกลงว่าการกระจายของคะแนนดิบจากการสอบ 2 ครั้ง ของบุคคลหนึ่งเป็นแบบไบนารีเหมือนกันตลอด และคะแนนทั้งสองเป็นอิสระต่อกัน นิยามของสัมประสิทธิ์ความสอดคล้องสำหรับบุคคล S ว่าเป็นความน่าจะเป็นที่บุคคลนั้น

ถูกกำหนดลงในชั้นเกี่ยวกับการเรียนรู้ อันเนื่องมาจากการสอบจากแบบทดสอบดูขนาด เมื่อใช้คะแนนเกณฑ์ 0 สำหรับชั้นของการเรียนรู้เป็น 2 ระดับคือ แบบผ่าน-ไม่ผ่าน เขียนได้ดังนี้

$$\text{เมื่อ } P_c(s) = P(X_s \geq 0, X_s = 0) + P(X_s \geq 0, X_s \leq 0)$$

เทอมสองเทอมทางขวามือของสมการ แทนความน่าจะเป็นของการตัดสินใจในการสอบได้-ได้และตก-ตก ตามลำดับ

X_s เป็นจำนวนของการตอบถูก

n เป็นจำนวนข้อสอบในแบบทดสอบ

ถ้าความเชื่อมั่นของแบบทดสอบได้จาก การประมาณค่าสัมประสิทธิ์ของความสอดคล้องของการตัดสินใจสำหรับบุคคล N คน ทั้งสมการ

$$P_c = \sum_{s=1}^N P_c(s) / N$$

$$\text{เมื่อ } P_c = [P(X_s \geq 0)]^2 + [1 - (X_s \geq 0)]^2$$

$$P(X_s \geq 0) = \sum_{X_s=0}^n \binom{n}{X_s} P_c^{X_s} (1 - P_c)^{n-X_s}$$

ค่าประมาณที่ได้ขึ้นอยู่กับรูปของความน่าจะเป็นของการตัดสินใจทั้งสองเส้นคงวาในกลุ่มผู้สอบ วิธีการของซิมโคเวียคมีความยุ่งยากในวิธีปฏิบัติบ้าง แต่ก็ได้ให้ความคลาดเคลื่อนค่อนข้างเล็ก ซึ่งแสดงถึงการประมาณค่าที่ค่อนข้างเที่ยง (แมงจิต อินทสุวรรณ 2523 : 60)

แฮร์ริส (Harris. 1972 : 41 - 45) ได้เสนอค่าสัมประสิทธิ์ในกรณีที่กำหนดสภาพการเรียนรู้เป็น 2 ระดับ ใช้กับข้อสอบที่เป็นแบบทดสอบวัดความรู้ ข้อสอบแต่ละข้อมีระบบการให้คะแนนเป็นถูกหรือผิดเท่านั้น ความเชื่อมั่นที่ได้เป็นค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างระดับความรู้กับคะแนนรวม

$$r_c = \frac{SS_B}{SS_B + SS_W}$$

เมื่อ MS_B คือ กำลังสองของสัมประสิทธิ์สหสัมพันธ์
 SS_B และ SS_W เป็นผลรวมของกำลังสองของการวิเคราะห์ระหว่างกลุ่ม
 และภายในกลุ่ม

โลเวทท์ (สมถิติ สนธิประเวศ 2522 : 4 - 6 อ้างอิงมาจาก Lovett. 1977) เป็นวิธีการหาความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ โดยวิธีการวิเคราะห์ความแปรปรวน คล้ายคลึงกับวิธีการของเกรซี อาศัยหลักการว่าแบบทดสอบนั้นสามารถจำแนกระดับการปฏิบัติของ แต่ละคนกับระดับเกณฑ์

$$\text{ความเชื่อมั่นของแบบทดสอบ} = \frac{MS_F - MS_e}{MS_F}$$

เมื่อ MS_F คือ ค่าเฉลี่ยกำลังสองของคะแนนระหว่างบุคคล
 MS_e คือ ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน

สำหรับงานวิจัยที่เกี่ยวข้องกับความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ ได้แก่ ชัมโคเวียค (Subkoviak. 1978 : 111 - 116) โดยศึกษาวิธีการ 4 แบบคือ วิธีการของสวามินาธาน และคณะ วิธีการของฮวิน วิธีการของมาร์แชลและแฮร์เทด และวิธีการของชัมโคเวียคเอง ซึ่ง 3 วิธีหลังใช้การสอบเพียงครั้งเดียว ทำการสอบกับนักเรียน 1,500 คน โดยสุ่มตัวอย่าง กลุ่มละ 30 คน รวม 50 กลุ่ม ทำแบบทดสอบประมาณวัดความถนัดเชิงวิชาการใช้เกณฑ์ตัดสิน รอยละ 50, 60, 70 และ 80 พบว่า วิธีการของสวามินาธานและคณะ ให้ค่าประมาณไม่ เอนเอียง แต่ความคลาดเคลื่อนมาตรฐานแนวจะเล็กน้อยแนวโน้มใหญ่ อาจลดลงด้วยการเพิ่ม ขนาดกลุ่มตัวอย่าง วิธีการของฮวินให้ค่าความเชื่อมั่นต่ำกว่าวิธีอื่น ๆ มาก แต่ให้ค่าความคลาดเคลื่อนต่ำกว่าวิธีอื่น ส่วนวิธีการของชัมโคเวียคเองให้ค่าความคลาดเคลื่อนค่อนข้างเล็ก แสดงถึง การประมาณค่าค่อนข้างเที่ยง

ในการสร้างแบบทดสอบอิงเกณฑ์ครั้งนี้ ผู้วิจัยเลือกใช้วิธีการหาความเชื่อมั่นของชัมโคเวียค ในการพิจารณาได้ยึดหลักการเลือกวิธีการที่มีขอบเขตของนัยที่สุด และถูกต้องตามทฤษฎีการวัดผล แบบอิงเกณฑ์ พิจารณาแล้วพบว่า สูตรของฮวินดีกว่า เป็นสูตรสะดวกในการปฏิบัติและเข้าใจง่าย แต่ได้รับการวิพากษ์วิจารณ์ว่าเป็นสูตรที่ไม่เหมาะสม เพราะไม่ได้คำนึงถึงการจำแนกผู้สอบว่าถูก หรือผิด (สงขม อักษร 2523 : 6) และไม่มีเจตนาที่จะเน้นทั้งหมัดเท่ากันหรือออกที่คะแนนเกณฑ์

วิธีการของสวามินาธานและคณะ ไม่สะดวกในแง่การปฏิบัติเพราะอาศัยผลการทดสอบครั้ง
ของแบบทดสอบคูณานัน วิธีการของโอเวห์ และวิธีการของแฮร์รี ทั้งสองวิธีนี้คำนวณจาก
การวิเคราะห์ความแปรปรวนของคะแนน ซึ่งไม่สอดคล้องกับการวัดผลแบบอิงเกณฑ์นัก ส่วน
วิธีการของฮวินมีความยุ่งยากซับซ้อนในการหาฟังก์ชัน โดยเฉพาะเมื่อคะแนนเกณฑ์ห่างจาก
คะแนนเต็มมาก และการในค่าประมาณค่าความน่าจะเป็นจริงมาก (นพจจิต อินทสุวรรณ
2523 : 61) วิธีการของซัคโคเวียคแม้จะมีความยุ่งยากในวิธีปฏิบัติบ้าง แต่ก็ได้หาความ
คลาดเคลื่อนก่อนรางวัลเล็ก ซึ่งแสดงถึงการประมาณค่าที่ค่อนข้างเที่ยง (นพจจิต อินทสุวรรณ
2523 : 60) ดังนั้นจึงได้เลือกเอาวิธีการหาค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์สร้างขึ้น
แต่ละฉบับโดยวิธีของซัคโคเวียค

ความเที่ยงตรงของแบบทดสอบอิงเกณฑ์

แบบทดสอบอิงเกณฑ์จำเป็นต้องมีความเที่ยงตรง 2 ลักษณะคือ ความเที่ยงตรงตาม
เนื้อหาและความเที่ยงตรงตามโครงสร้าง ในการตรวจหาความเที่ยงตรงตามเนื้อหา นั้น เป็น
การตรวจสอบรายข้อต่าง ๆ ของแบบทดสอบนั้นว่า สามารถเป็นตัวแทนของพฤติกรรมหรือรายข้อ
ต่าง ๆ ในขอบเขตของการวัดได้หรือไม่ ผู้สอบที่มีข้อนี้จะช่วยให้การแปลคะแนนของข้อสอบ
มีความหมายยิ่งขึ้น และในแง่ความเที่ยงตรงตามโครงสร้าง ถ้าว่า "โครงสร้าง" คือ คุณลักษณะ
ภายในตัวผู้สอบที่แสดงการมีความสามารถใดครอบคลุมในจุดประสงค์ที่วัดโดยข้อสอบนั้น ซึ่งเรียกว่า
ผู้รอบรู้ หรือคุณลักษณะที่แสดงถึงการไม่ครอบคลุมใดครอบคลุมในจุดประสงค์ที่กำลังวัดซึ่งเรียกว่าผู้ไม่รอบรู้
(สงขล ลักษณะ 2523 : 37 - 38)

สำหรับความเที่ยงตรงตามเนื้อหาเป็นสิ่งที่สำคัญที่สุดในแบบทดสอบอิงเกณฑ์ (Glaser and
Nitko. 1973 : 159 - 170) การตั้งจุดประสงค์ไว้ชัดเจนจะช่วยให้การตรวจสอบได้มาก
(กมล ภูประเสริฐ 2520 : 86 อ้างอิงมาจาก Merhren and Lehman. 1973)
จุดประสงค์ในการเรียนจึงมีความสำคัญมากในการพิจารณาความเที่ยงตรงประเภทนี้ ซึ่งตัดสินโดย
ผู้เชี่ยวชาญทางเนื้อหา การจะใช้วิธีการของระบบอิงเกณฑ์โดยตรง (Millman. 1974 : 63)
มีผู้เสนอไว้หลายวิธีการ ดังนี้

โรวินแลนดี และแฮมเบิลตัน (Hambleton. 1978 : 34 citing Rovinalli Hambleton. 1977) เสนอวิธีกำหนดจุดประสงค์และรายชื่อกาง ๆ ที่ใช้วัดจุดประสงค์นั้น เสนอต่อผู้เชี่ยวชาญหลายคน โดยผู้เชี่ยวชาญจะต้องให้คะแนนในแต่ละข้อ โดยมีหลักการว่า

- 1 คือ แม้ใจอาจสงสัยนั้นวัดจุดประสงค์ได้จริง
- 0 คือ ไม่แม้ใจอาจสงสัยนั้นวัดจุดประสงค์นั้น
- 1 คือ แม้ใจอาจสงสัยนั้นไม่วัดจุดประสงค์นั้น

แล้วรวมคะแนนจากผู้เชี่ยวชาญในแต่ละข้อ ตัวเลขคะแนนนี้จะเรียกว่าดัชนีของความสอดคล้องกัน ระหว่างข้อสอบแต่ละข้อกับจุดประสงค์ แล้วหาจุดตัดของคะแนนเพื่อจำแนกรายชื่อที่วัดจุดประสงค์นั้น กับที่ไม่ได้วัดจุดประสงค์นี้

วิธีที่สองในการหาความเที่ยงตรงตามเนื้อหา ซึ่งเสนอโดยบุคคลผู้เกี่ยวข้องกับวิธีแรก โดยใช้มาตราส่วนประมาณค่าเป็นเครื่องมือ ผู้เชี่ยวชาญทั้งหลายจะได้รับรายการจุดประสงค์กับข้อสอบหลายข้อเหมือนวิธีแรก ผู้เชี่ยวชาญจะให้คะแนนความเหมาะสมที่คิดว่าแต่ละข้อสามารถวัดจุดประสงค์นั้นได้ เช่น เหมาะสมมากที่สุดให้ 5 เหมาะสมน้อยที่สุดให้ 1 เป็นต้น แล้วหาค่าเฉลี่ยของข้อสอบแต่ละข้อ นำมาพิจารณาพร้อมกับค่าความเบี่ยงเบนมาตรฐาน ก็จะได้อัตราความเที่ยงตรงของเนื้อหา

และอีกวิธีหนึ่งที่โรวินแลนดี และแฮมเบิลตัน เสนอวิธีการหาความเที่ยงตรงตามเนื้อหา โดยกำหนดให้ผู้เชี่ยวชาญทางเนื้อหาหลายคนพิจารณารายการของข้อสอบ และรายการ องค์จุดประสงค์ กาง ๆ แล้วให้จับคู่ว่าข้อใดวัดจุดประสงค์ใด แล้วอาจใช้สถิติไค-สแควร์ทดสอบหาความเป็นอิสระของข้อมูล หรืออาจใช้การพิจารณาด้วยสายตาถ้าก็พอจะเห็นรายข้อแสดงความสอดคล้องหรือไม่สอดคล้องของผู้เชี่ยวชาญ

นอกจากนี้ครอนบาร์ช เสนอแนะความเที่ยงตรงตามเนื้อหาของแบบทดสอบถึง เกิดขึ้น โดยใช้การทดลองของผู้เขียนข้อสอบ 2 กลุ่ม ที่มีความสามารถใกล้เคียงกัน เขียนข้อสอบกลุ่มละ 1 ชุด วัดจุดประสงค์เดียวกัน แล้วนำไปสอบกับนักเรียนกลุ่มหนึ่งที่เรียนจุดประสงค์นั้นแล้ว ซึ่งถ้ากำหนดจุดประสงค์อย่างชัดเจนและการสุ่มข้อสอบจากโคเนนท์ทำได้เหมาะสม ผลการทำข้อสอบ 2 ชุด ของนักเรียนกลุ่มนี้จะค่อนข้างสอดคล้องกัน

ข้อที่ไม่สอดคล้องจะถูกกำจัดออกในฐานะที่ค้อยคุณสมบัติของความเที่ยงตรงตามเนื้อหา (สงบ. ลักษณะ 2523: 40 อ้างอิงมาจาก Gronbach. 1971)

คาร์เวอร์ (Carver. 1970: 256) เสนอวิธีหาความเที่ยงตรงตามเนื้อหา ของแบบทดสอบอิงเกณฑ์โดยใช้ค่าสถิติ ข้อมูลได้จากการสอบนักเรียนที่เรียนแล้วกับ นักเรียนที่ยังไม่เรียน คำนวณหาค่าความเที่ยงตรงตามตาราง 5

ตาราง 5 ค่าความเที่ยงตรงตามเนื้อหาที่ได้จากคะแนนนักเรียน 2 กลุ่ม

	กลุ่มที่ยังไม่เรียน	กลุ่มที่เรียนแล้ว
ผ่าน	b	a
ไม่ผ่าน	c	d

$$\text{ความเที่ยงตรง} = (a + c) / N$$

$$\text{เมื่อ } N = a + b + c + d$$

ส่วนการหาความเที่ยงตรงตามโครงสร้างนั้น เป็นการตรวจสอบการแปล ความหมายของคะแนน ซึ่งอาจหาสถิติบางตัวมาตรวจสอบ กระบวนการนี้อาจจะรวม อยู่กับการวิเคราะห์รายข้อ ได้แก่ ความคล้ายของข้อสอบและคำอ่านอาจจำแนก โดยใช้การแปรเปลี่ยนของค่าความยากง่ายของรายข้อ เมื่อใช้กับนักเรียนก่อนเรียน และหลังเรียนเป็นทวีคูณนะ ถ้าไม่มีการแปรเปลี่ยนของการสอบทั้งสองครั้งนี้ถือว่า ขาดความเที่ยงตรงตามโครงสร้าง ข้อสอบต่างๆที่วัดจุดประสงค์เกี่ยวกับควรมีค่า ความยากง่ายใกล้เคียงกัน(สงบ. ลักษณะ 2523: 41 อ้างอิงมาจาก Cox and Vargas. 1966) นอกจากนี้ยังมีผู้เสนอวิธีการทางสถิติอื่นๆเพื่อตรวจสอบความเที่ยง ตรงตามโครงสร้าง เช่น มาตรฐานของกัทท์แมน (GUTTMAN) และการวิเคราะห์ องค์ประกอบ เป็นต้น (Hambleton. 1978: 39)

ในการสร้างแบบทดสอบอิงเกณฑ์ครั้งนี้ ผู้วิจัยได้หาความเที่ยงตรง
 ความเนื้อหาโดยใช้สูตรของคาร์เวอร์ ค่าที่คำนวณได้จากสูตรจะมีค่าสูงเมื่อจำนวน
 นักเรียนที่เรียนแล้วสอบผ่านและจำนวนนักเรียนที่ยังไม่เรียนซึ่งสอบไม่ผ่าน มีจำนวน
 มากกว่านักเรียนที่เรียนแล้วสอบไม่ผ่านและนักเรียนที่ยังไม่เรียนซึ่งสอบผ่าน
 สอดคล้องกับลักษณะการเรียนการสอนที่นักเรียนจะได้รับความรู้เพิ่มขึ้นเมื่อเรียนเรื่อง
 นั้นมาแล้ว ซึ่งค่าความเที่ยงตรงนี้จะขึ้นกับประสิทธิภาพในการสอนด้วย นอกจากจะ
 หาค่าความเที่ยงตรงความเนื้อหาด้วยวิธีดังกล่าวแล้ว ยังใช้วิธีการของโรวินแนลลี
 และแสมเปิดค้นร่วมด้วย โดยใช้วิธีให้ผู้เชี่ยวชาญทางเนื้อหาวิชาพิจารณาความชัดเจน
 ของจุดประสงค์เชิงพฤติกรรม และจับคู่รายการของจุดประสงค์เชิงพฤติกรรมกับข้อสอบ
 รายข้อ เพื่อให้ข้อสอบที่สร้างขึ้นเป็นตัวแทนของพฤติกรรมที่ห้องการวัด ข้อสอบที่ไม่ได้
 การเลือกหรือจับคู่จะนำมาปรับปรุงใหม่

งานวิจัยที่เกี่ยวข้องกับการสร้างแบบทดสอบอิงเกณฑ์

แกนโนโพล (Ganoplo. 1978 : 246 - A) สร้างแบบทดสอบอิงเกณฑ์
 วัดความสามารถในวิชาพื้นฐานการอ่าน เพื่อพัฒนาและหาข้อสอบที่เหมาะสมในการ
 วัดความสามารถในการอ่านของนักเรียนระดับมัธยมศึกษา โดยตัดสินความรอบรู้
 ที่ประเมินจากแบบทดสอบ และนิยามชี้แจงจำกัดของการวัดโดยแบบทดสอบ ความเชื่อมั่น
 ของแบบทดสอบได้จากการสอบซ้ำ แบบทดสอบคู่ขนานและรวมทั้ง 2 วิธีในการ
 ประเมินค่า ซึ่งเป็นการหาในแบบของความคงตัวและความเสมอเหมือนโดยคำนึงถึง
 1) ระเบียบวิธีวัดและสหสัมพันธ์ที่ใช้คำนวณค่าประมาณในการวัดแบบอิงกลุ่ม และ
 2) ความสอดคล้องของการตัดสินและสถานการณ์ที่ใช้วัดแบบอิงเกณฑ์ ส่วนค่า
 ความเที่ยงตรงได้จากการเลือกข้อสอบจากโคเมนและการแปลผลของลักษณะพรรณนา
 แบบทดสอบ

โฮลเมส์ (Holmes. 1978 : 4197 - A) ศึกษาผลสัมฤทธิ์และทัศนคติของนักเรียนระดับมัธยมศึกษา เมื่อใช้การประเมินผลแบบอิงเกณฑ์ในวิชาคณิตศาสตร์ ตัวแปรอิสระคือ การใช้ชุดการสอน นักเรียน และวิธีการวัดผลแบบอิงเกณฑ์ ตัวแปรตามคือ ผลสัมฤทธิ์และทัศนคติที่เกี่ยวข้องกับวิชาคณิตศาสตร์ พบว่า ค่าเฉลี่ยของคะแนนผลสัมฤทธิ์ในการสอบหลังเรียนของกลุ่มทดลองกับกลุ่มควบคุมแตกต่างกันอย่างไม่มีนัยสำคัญ ค่าเฉลี่ยของทัศนคติของ 2 กลุ่มแตกต่างกันอย่างไม่มีนัยสำคัญ และสรุปได้ว่า เครื่องมือที่ใช้ในการทดสอบที่สัมพันธ์กับวิธีการ ไม่ทำให้ผลสัมฤทธิ์หรือทัศนคติทางบวกเพิ่มขึ้น ชุดการสอนและวิธีการวัดผลมีผลกระทบถึงความสามารถในการอ่านและเพศของนักเรียนในเชิงเปรียบเทียบ

ฮอร์เนอร์ (Horner. 1978 : 4198 - A) ศึกษาการใช้แบบทดสอบอิงเกณฑ์และแบบทดสอบอิงกลุ่มในวิชาคณิตศาสตร์กับนักเรียนระดับสี่จำนวน 133 คน ในการทดสอบสองทางพบว่า สัมประสิทธิ์ระหว่างตัวแปรของแบบทดสอบทั้งสองชนิดคือ .87 ซึ่งไม่มีนัยสำคัญของความแตกต่างระหว่างตัวแปรของแบบทดสอบทั้งสองชนิด

กาญจนา วันสุนทร (กาญจนา วันสุนทร 2521 : 23 - 198) สร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์เรื่องสมการ ทดสอบกับนักเรียนชั้นมัธยมศึกษาปีที่ 1 จำนวน 60 คน โดยกำหนดเกณฑ์ต่างกัน 3 จุดคือ .6 , .7 , และ .8 การคัดเลือกข้อสอบใช้ดัชนีเอนท์ในการหาค่าอ่านจำแนก กำหนดค่าความยากและค่าอ่านจำแนกที่สูงกว่า .20 แบบทดสอบอิงเกณฑ์ประกอบด้วยแบบทดสอบย่อย 4 ฉบับๆละ 10 ข้อ หาค่าความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบโดยใช้วิธีการของลิฟวิงตันและคาร์เวอร์ตามลำดับ และพบว่าคะแนนเกณฑ์มีผลต่อค่าความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบ เมื่อคะแนนเกณฑ์สูงค่าความเชื่อมั่นและความเที่ยงตรงจะลดลง

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

แหล่งข้อมูลและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการศึกษาค้นคว้า เป็นนักเรียนชั้นประถมศึกษาปีที่ 2 ปีการศึกษา 2523 ของโรงเรียนในสังกัดสำนักงานการประถมศึกษา ในจังหวัดสมุทรสงครามจำนวน 110 โรงเรียน จำนวนนักเรียน 3,442 คน

กลุ่มตัวอย่างในการศึกษาค้นคว้า เป็นนักเรียนชั้นประถมศึกษาปีที่ 2 ปีการศึกษา 2523 ของโรงเรียนในสังกัดสำนักงานการประถมศึกษา ในจังหวัดสมุทรสงครามจำนวน 9 โรงเรียน จำนวนนักเรียน 370 คน ซึ่งเลือกมาโดยการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) ดังนี้

จำแนกโรงเรียนเป็น 3 ประเภทตามที่ตั้งในอำเภอทั้งหมด 3 อำเภอ ในจังหวัดสมุทรสงคราม คือ อำเภอเมือง อำเภออัมพวา และอำเภอบางคนที โดยการสุ่มโรงเรียนมาอำเภอละ 3 โรงเรียน รวมทั้งหมด 9 โรงเรียน จำนวนนักเรียนที่ได้จากการสุ่ม 370 คน ซึ่งเป็นกลุ่มตัวอย่างในการทดสอบเครื่องมือทั้ง 3 ครั้ง

ครั้งที่ 1 สุ่มโรงเรียนมาอำเภอละโรงเรียน ได้โรงเรียนจำนวน 3 โรงเรียน จากโรงเรียน 9 โรงเรียนที่สุ่มในตอนแรก จำนวนนักเรียนที่เป็นกลุ่มตัวอย่างจากการสุ่มครั้งนี้ 100 คน เพื่อใช้ในการทดสอบครั้งที่ 1 นำผลมาวิเคราะห์และปรับปรุงข้อสอบ

ครั้งที่ 2 สุ่มโรงเรียนมาอำเภอละโรงเรียน ได้โรงเรียนจำนวน 3 โรงเรียนที่เหลือจากการสุ่มครั้งที่ 1 จำนวนนักเรียนที่เป็นกลุ่มตัวอย่าง 146 คน เพื่อใช้ในการทดสอบครั้งที่ 2 นำผลมาวิเคราะห์ คัดเลือก และปรับปรุงข้อสอบ

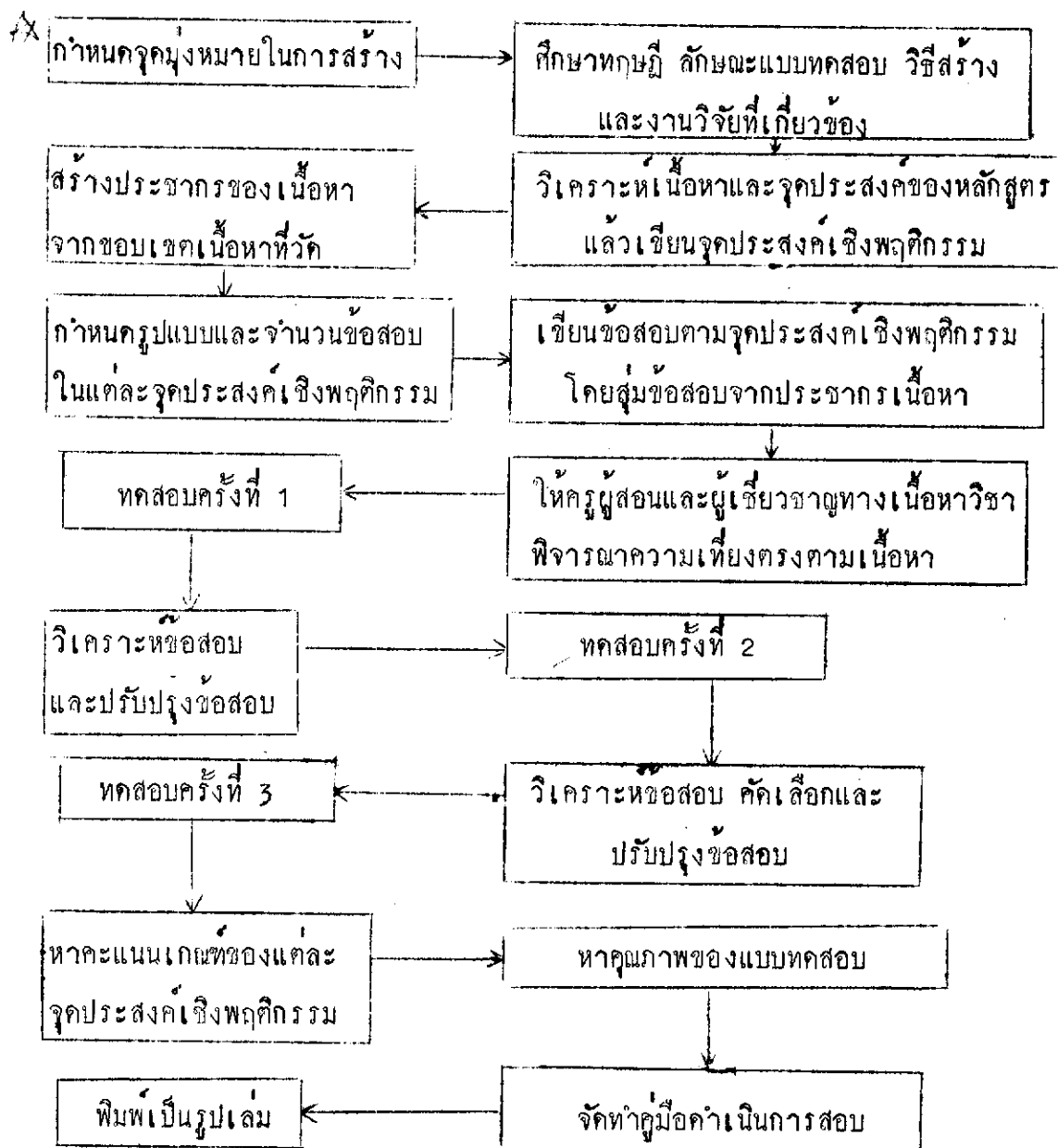
ครั้งที่ 3 นักเรียนที่เหลือจำนวน 124 คนในโรงเรียน 3 โรงเรียน ใช้เป็นกลุ่มตัวอย่างในการทดสอบครั้งที่ 3 ซึ่งนำแบบทดสอบที่ปรับปรุงแก้ไขแล้วมาทดสอบเพื่อนำมาคะแนนเกณฑ์และคุณภาพของแบบทดสอบ

ตาราง 6 รายชื่อโรงเรียน จำนวนนักเรียนในกลุ่มตัวอย่าง และครั้งที่ใช้ทดสอบ

โรงเรียน	จำนวนนักเรียน	ทดสอบครั้งที่
<u>อำเภอเมือง</u>		
โรงเรียนวัดช่องลม(ธรรมโชติ)	40	1
โรงเรียนวัดศรีรัตนาราม	34	2
โรงเรียนวัดธรรมาราม	47	3
<u>อำเภออัมพวา</u>		
โรงเรียนวัดสี่แยกราษฎร์บำรุง	39	1
โรงเรียนวัดประชาโสภาการาม	63	2
โรงเรียนวัดช้างเผือก	21	3
<u>อำเภอบางคนที</u>		
โรงเรียนวัดเกตุการาม	21	1
โรงเรียนวัดเจริญสุขาราม	49	2
โรงเรียนวัดบางน้อย	56	3

★ วิธีดำเนินการสร้างแบบทดสอบ

ในการดำเนินการสร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์เรื่องการหาร
 ชั้นประถมศึกษาปีที่ 2 ผู้วิจัยได้ดำเนินการสร้างตามลำดับขั้น ดังแสดงในภาพประกอบ 3



ภาพประกอบ 3 ลำดับขั้นในการดำเนินการสร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์

การสร้างแบบทดสอบอิงเกณฑ์ครั้งนี้ ดำเนินการสร้างตามลำดับขั้น (ภาพประกอบ 3) ดังนี้

★ 1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบ

จุดมุ่งหมายเพื่อสร้างแบบทดสอบอิงเกณฑ์วัดผลสัมฤทธิ์ในการเรียนวิชา ศนิตศาสตร์ เรื่องการหาร สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 เพื่อนำไปประเมินผลระหว่างเรียนว่า นักเรียนมีความรู้ความสามารถเรื่องการหาร ตามจุดประสงค์ที่กำหนดไว้ในแต่ละหน่วยการเรียนรู้หรือไม่ จุดประสงค์เชิงพฤติกรรมใดที่นักเรียนไม่ผ่าน ควรทบทวนเพิ่มเติมหรือซ่อมเสริมก่อนที่จะเรียนหน่วยการเรียนรู้ต่อไป

2. ศึกษาทฤษฎี ลักษณะแบบทดสอบ วิธีการสร้างแบบทดสอบและงานวิจัยที่เกี่ยวข้อง เพื่อเป็นแนวทางในการสร้างแบบทดสอบอิงเกณฑ์

3. วิเคราะห์เนื้อหาและจุดประสงค์เรื่องการหาร ชั้นประถมศึกษาปีที่ 2 ซึ่งมีขั้นตอนการวิเคราะห์ ดังนี้

3.1 ศึกษาเนื้อหา หลักการ และจุดประสงค์เรื่องการหาร ชั้นประถมศึกษาปีที่ 2 จากหลักสูตรประถมศึกษา พุทธศักราช 2521

3.2 ศึกษารายละเอียดเนื้อหาเรื่องการหารทั้งหมด และจุดประสงค์เชิงพฤติกรรมในเรื่องนี้สำหรับชั้นประถมศึกษาปีที่ 2 ตลอดจนเนื้อหาที่เกี่ยวข้องกับการหารในชั้นประถมศึกษาปีที่ 1 จากคู่มือครู ประมวลการเรียนรู้ แบบเรียน และสมุดประจำชั้น ป. 02/2

3.3 วิเคราะห์เนื้อหาเรื่องการหาร เพื่อจัดเป็นหน่วยการเรียนรู้ย่อยจากหลักสูตรประถมศึกษา พุทธศักราช 2521 ซึ่งแบ่งเนื้อหา 5 หัวข้อคือ ความหมายของการหารและสัญลักษณ์การหาร ประโยคสัญลักษณ์การหาร การหารจำนวนที่ตัวหารมีหลักเดียวจนหารไม่เกิน 10 แบบคงตัวและเหลือเศษ การตรวจสอบคำตอบโดยใช้ความสัมพันธ์ระหว่างการคูณและการหาร และโจทย์ปัญหาการหาร เนื่องจากผู้ชำนาญในสาขาศนิตศาสตร์และอาจารย์ผู้สอนมีความเห็นว่า ควรให้รวม "การนับลด" เข้ากับเรื่องนี้และการวิเคราะห์ลำดับขั้นการเรียนรู้เรื่องการหารของ กรม ภูประเสริฐ

(กมล ภูประเสริฐ 2520 : 35) ได้จัดเรื่องการนับลคเป็นพฤติกรรมพื้นฐานในการเรียนเรื่องอาหาร ตลอดจนในแบบเรียนก็ได้จัดหัวข้อการนับลคไว้ก่อนที่นักเรียนจะเรียนเรื่องอาหารนี้ ดังนั้นเพื่อความเหมาะสมและสะดวกในการประเมินจุดประสงค์และการออกข้อสอบ จึงจัดแบ่งเป็นเนื้อหาย่อย 4 หัวข้อ ดังนี้

- 1) การนับลคและความหมายของการหาร
- 2) ประโยชน์สัญลักษณ์การหาร และความสัมพันธ์ระหว่างการคูณและการหาร
- 3) โจทย์การหาร
- 4) โจทย์ปัญหาการหาร

3.4 วิเคราะห์จุดประสงค์ในหลักสูตรว่า วัดพฤติกรรมในหน่วยเนื้อหาย่อยตอนใด จุดประสงค์ที่วัดเนื้อหาเดียวกัน ซึ่งมีวิธีการหรือพฤติกรรมอย่างเดียวกัน แต่ผลการกระทำต่างกัน เช่น " เมื่อกำหนดโจทย์การหารที่มีตัวหารหลักเดียว ผลหารไม่เกิน 10 นักเรียนสามารถทำโจทย์เกี่ยวกับการหารแบบลงตัวและมีเศษได้ " เมื่อคำนึงถึงผลการประเมินความสามารถของนักเรียนให้ถูกจุด และการกำหนดขอบเขตเนื้อหาที่ทดสอบ จึงแยกเป็นจุดประสงค์เชิงพฤติกรรม 2 จุดประสงค์ คือ "เมื่อกำหนดโจทย์การหารที่มีตัวหารหลักเดียว ผลหารไม่เกิน 10 นักเรียนสามารถทำโจทย์เกี่ยวกับการหารแบบลงตัวได้" และ "เมื่อกำหนดโจทย์การหารที่มีตัวหารหลักเดียว ผลหารไม่เกิน 10 นักเรียนสามารถทำโจทย์เกี่ยวกับการหารแบบมีเศษได้"

เนื่องจากในการสร้างแบบทดสอบครั้งนี้ ผู้วิจัยมีจุดมุ่งหมายในการหาคะแนนเกณฑ์ เพื่อตัดสินการสอบผ่านของนักเรียนในแต่ละจุดประสงค์เชิงพฤติกรรม ดังนั้นในจุดประสงค์เชิงพฤติกรรมที่เขียนขึ้น จึงมิได้ระบุส่วนที่เป็น"เกณฑ์"ไว้

3.5 จากการวิเคราะห์เนื้อหาและจุดประสงค์ ในการเรียนเรื่องอาหาร ชั้นประถมศึกษาปีที่ 2 ได้เขียนจุดประสงค์เชิงพฤติกรรมในหน่วยเนื้อหาย่อยดังแสดงในตาราง 7 เพื่อเป็นแบบแผนในการสร้างแบบทดสอบ และเป็นแนวทางในการพิจารณาความเที่ยงตรงตามเนื้อหา ดังนี้

ตาราง 7 เนื้อหาและจุดประสงค์เชิงพฤติกรรมเรื่องการหาร ชั้นประถมศึกษาปีที่ 2

เนื้อหา	จุดประสงค์เชิงพฤติกรรม
1. ความหมายของการหาร 1.1 การนับลค 1.2 ความรู้เบื้องต้น- เรื่องการหาร	1.1.1 เมื่อกำหนดจำนวนนับลค 2 - 3 จำนวน นักเรียนสามารถบอกจำนวนต่อไปหรือจำนวนที่ ขาดหายไปได้ 1.1.2 เมื่อกำหนดจำนวนเลขหรือสิ่งของ นักเรียนสามารถบอกจำนวนครั้งที่เอาออกจนหมด หรือนับลคได้ 1.1.3 เมื่อกำหนดจำนวนเลขหรือสิ่งของ นักเรียนสามารถบอกจำนวนครั้งที่เอาออกจนหมด และเศษที่เหลือได้ 1.2.1 เมื่อกำหนดสิ่งของให้จำนวนหนึ่งและบอก จำนวนกลุ่ม นักเรียนสามารถหาจำนวนในแต่ละ กลุ่มได้ 1.2.2 เมื่อกำหนดจำนวนสิ่งของที่เท่ากันของแต่ละ กลุ่มและของทั้งหมด นักเรียนสามารถหาจำนวนกลุ่มได้
2. สัญลักษณ์การหาร 2.1 ประโยคสัญลักษณ์การหาร 2.2 ความสัมพันธ์ระหว่าง- การคูณและการหาร	2.1 เมื่อกำหนดภาพหรือโจทย์ให้ นักเรียน สามารถเขียนอยู่ในรูปประโยคสัญลักษณ์ได้ 2.2 เมื่อกำหนดโจทย์การหาร นักเรียนสามารถ ใช้ความสัมพันธ์ระหว่างการคูณและการหารในการหา ตัวตั้ง ตัวหาร และผลหารได้

ตาราง 7 (ต่อ)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม
3. โจทย์การหาร	
3.1 การหารลงตัว	3.1 เมื่อกำหนดโจทย์การหารที่มีตัวหารหลักเดียว ผลหารไม่เกิน 10 นักเรียนสามารถทำโจทย์เกี่ยวกับการหารลงตัวได้
3.2 การหารมีเศษ	3.2 เมื่อกำหนดโจทย์การหารที่มีตัวหารหลักเดียว ผลหารไม่เกิน 10 นักเรียนสามารถทำโจทย์เกี่ยวกับการหารมีเศษได้
4. โจทย์ปัญหาการหาร	
4.1 การหารลงตัว	4.1 เมื่อกำหนดโจทย์ปัญหาการหารที่มีตัวหารหลักเดียว ผลหารไม่เกิน 10 นักเรียนสามารถหาคำตอบแบบลงตัวได้
4.2 การหารมีเศษ	4.2 เมื่อกำหนดโจทย์ปัญหาการหารที่มีตัวหารหลักเดียว ผลหารไม่เกิน 10 นักเรียนสามารถหาคำตอบแบบมีเศษได้

4. จากตารางเนื้อหาและจุดประสงค์เชิงพฤติกรรม ได้ขอบเขตเนื้อหาจากการศึกษาและการวิเคราะห์เนื้อหาได้ข้อ 3 พบว่ามีขอบเขตเนื้อหา ดังนี้

ขอบเขตที่ 1 การนับลค โดยนับลคที่ละ 4, 6, 7, 8 และ 9 ส่วนการนับลคที่ละ 2, 3, 5 และ 10 นักเรียนเรียนมาแล้วในชั้นประถมศึกษาปีที่ 1

ขอบเขตที่ 2 การนับลคแบบมีเศษ โดยนับลคที่ละ 2, 3, 4, ..., 9, 10

ขอบเขตที่ 3 การหารลงตัว ที่มีตัวหารหลักเดียวและผลหารไม่เกิน 10

ขอบเขตที่ 4 การหารมีเศษ ที่มีตัวหารหลักเดียวและผลหารไม่เกิน 10

จากขอบเขตเนื้อหาดังกล่าว นำมาสร้างประชากรเนื้อหาที่มุ่งจะวัดได้ ดังนี้

ตาราง 8 ประชากรเนื้อหาที่ได้จากขอบเขตเนื้อหาเรื่องการนับลค

การนับลคแบบลงตัว	การนับลคแบบมีเลข
จำนวนครั้งที่นับลคทีละ 2 จนหมด 20,18,16,...,4,2 (10, 9, 8,...,2,1 ครั้ง) จำนวนครั้งที่นับลคทีละ 3 จนหมด 30,27,24,...,6,3 (10, 9, 8,...,2,1 ครั้ง) . . . จำนวนครั้งที่นับลคทีละ 9 จนหมด 90,81,72,...,18,9 (10, 9, 8,..., 2,1 ครั้ง) จำนวนครั้งที่นับลคทีละ 10 จนหมด 100,90,80,...,20,10 (10, 9, 8,..., 2, 1 ครั้ง)	จำนวนครั้งที่นับลคทีละ 2 จนหมด แล้วเหลือ 1 จำนวนครั้งที่นับลคทีละ 3 จนหมด แล้วเหลือ 1,2 . . . จำนวนครั้งที่นับลคทีละ 9 จนหมด แล้วเหลือเศษ 1,2,3,4,5,6,7 และ 8 จำนวนครั้งที่นับลคทีละ 10 จนหมด แล้วเหลือเศษ 1,2,3,4,5,6,7,8 และ9

ตาราง 9 ประชากรเนื้อหาที่ได้จากขอบเขตเนื้อหาการหาร

การหารลงตัว	การหารมีเศษ
จำนวนที่หารด้วย 2 ลงตัว 20, 18, 16, ..., 4, 2	จำนวนที่หารด้วย 2 แล้วเหลือเศษ 1
จำนวนที่หารด้วย 3 ลงตัว 30, 27, 24, ..., 6, 3	จำนวนที่หารด้วย 3 แล้วเหลือเศษ 1, 2 .
.	.
.	.
จำนวนที่หารด้วย 9 ลงตัว 90, 81, 72, ..., 18, 9	จำนวนที่หารด้วย 9 แล้วเหลือเศษ 1, 2 3, 4, 5, 6, 7, 8

5. กำหนดรูปแบบของแบบทดสอบ และจำนวนข้อสอบในแต่ละจุดประสงค์เชิงพฤติกรรม รูปแบบของแบบทดสอบอิงเกณฑ์ที่นำมาทดสอบนักเรียนชั้นประถมศึกษาปีที่ 2 นี้ ได้เลือกใช้แบบทดสอบปรนัยชนิดเลือกตอบ 3 ตัวเลือก เนื่องจากแบบทดสอบปรนัยสามารถออกข้อสอบได้ครอบคลุมเนื้อหาและพฤติกรรม เหมาะสมกับการวิเคราะห์เพื่อหาคุณภาพรายข้อ สามารถควบคุมระดับความยากและค่าอำนาจจำแนก มีค่าความเชื่อมั่นและความเที่ยงตรงสูงกว่าแบบอื่นๆ ส่วนจำนวนตัวเลือกที่เหมาะสมกับนักเรียนระดับประถมศึกษาตอนต้นคือ 3 ตัวเลือก และแบบทดสอบจัดแบ่งเป็นแบบทดสอบย่อย 4 ฉบับ โดยมีจุดประสงค์นำไปใช้วัดเนื้อหาย่อยทั้งสี่หัวข้อ เพื่อจะสามารถประเมินผลการเรียนของนักเรียนในจุดประสงค์เชิงพฤติกรรมในแต่ละตอน แบบทดสอบย่อยทั้ง 4 ฉบับ ได้แก่

- ฉบับที่ 1 ความหมายของการหาร
 ตอนที่ 1 การนับลดครั้งละเท่าๆกัน
 ตอนที่ 2 ความรู้เบื้องต้นในเรื่องการหาร
- ฉบับที่ 2 สัญลักษณ์การหาร
 ตอนที่ 1 ประโยคสัญลักษณ์การหาร
 ตอนที่ 2 ความสัมพันธ์ระหว่างการคูณและการหาร
- ฉบับที่ 3 โจทย์การหาร
 ตอนที่ 1 การหารลงตัว
 ตอนที่ 2 การหารมีเศษ
- ฉบับที่ 4 โจทย์ปัญหาการหาร
 ตอนที่ 1 การหารลงตัว
 ตอนที่ 2 การหารมีเศษ

แบบทดสอบย่อยฉบับที่ 1 ประกอบด้วยจุดประสงค์เชิงพฤติกรรมจำนวน 5 จุดประสงค์ ฉบับที่ 2 ประกอบด้วยจุดประสงค์เชิงพฤติกรรมจำนวน 2 จุดประสงค์ ฉบับที่ 3 ประกอบด้วยจุดประสงค์เชิงพฤติกรรม 2 จุดประสงค์ และฉบับที่ 4 ประกอบด้วยจุดประสงค์เชิงพฤติกรรม 2 จุดประสงค์ จากจุดประสงค์ดังกล่าวได้ กำหนดจำนวนข้อสอบที่เหมาะสมในแต่ละจุดประสงค์ โดยคำนึงถึงความยากง่ายของ เนื้อหาและจำนวนคาบการสอน และความสำคัญของจุดประสงค์ โดยพิจารณาจำนวน ข้อร่วมกับครูผู้สอน ดังแสดงในตาราง 10

ตาราง 1ก รายชื่อแบบทดสอบย่อย จุดประสงค์เชิงพฤติกรรม และจำนวนข้อสอบ

แบบทดสอบย่อยฉบับที่	จุดประสงค์เชิงพฤติกรรม	จำนวนข้อสอบ	ข้อสอบทั้งหมด
1. ความหมายของการहार			
* 1.1 การนับเลข	1.1.1	5	29
	1.1.2	5	
	1.1.3	5	
1.2 ความรู้เบื้องต้นในการहार	1.2.1	7	
	1.2.2	7	
2. สัญลักษณ์การहार			
2.1 ประโยคสัญลักษณ์การहार	2.1	8	16
* 2.2 ความสัมพันธ์ระหว่างการคูณและการहार	2.2	8	
3. โจทย์การहार			
* 3.1 การहारลงตัว	3.1	8	16
3.2 การहारมีเศษ	3.2	8	
4. โจทย์ปัญหาการहार			
4.1 การहारลงตัว	4.1	8	16
4.2 การहारมีเศษ	4.2	8	

6. สร้างข้อสอบที่ใช้วัดความสามารถของนักเรียน ตามจุดประสงค์เชิงพฤติกรรมแต่ละจุดประสงค์ โดยสุ่มข้อสอบจากประชากรเนื้อหาที่กำหนด ซึ่งแบบทดสอบย่อยฉบับที่ 1 อยู่ในขอบเขตเนื้อหาที่ 1 และขอบเขตเนื้อหาที่ 2 ตัวเลขจึงได้จากการสุ่มประชากรเนื้อหาในตาราง 8 ส่วนแบบทดสอบย่อยฉบับที่ 2, 3 และ 4 อยู่ในขอบเขตเนื้อหาที่ 4 ตัวเลขที่ใช้ออกข้อสอบจึงได้จากการสุ่มประชากรเนื้อหาในตาราง 9 สำหรับข้อสอบที่เกี่ยวข้องกับโจทย์ปัญหาการหาร ซึ่งเป็นข้อสอบในแบบทดสอบย่อยฉบับที่ 4 และในแบบทดสอบย่อยฉบับที่ 2 ตอนที่ 1 นำลักษณะโจทย์คำถามที่นักเรียนเคยพบเห็นหรือใช้เรียนในหนังสือแบบเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 2 ส่วนตัวเลขในแต่ละข้อของแบบทดสอบทุกฉบับ นำมาจากข้อบกพร่องในการทำเรื่องการหารที่สอบถามจากครูผู้สอน เช่น การใช้วิธีการลบแทนการหาร การสลับที่ตำแหน่งการหาร การคิดผลหารผิดของนักเรียน เป็นต้น และจำนวนข้อสอบมีมากกว่าที่กำหนดไว้ในจุดประสงค์เชิงพฤติกรรมละ 2 - 5 ข้อ แล้วจัดเรียงข้อสอบตามลำดับจุดประสงค์ในแบบทดสอบย่อยแต่ละฉบับ

7. นำแบบทดสอบที่ประกอบด้วยรายข้อที่สร้างขึ้นทั้งสี่ฉบับชุดหนึ่ง และรายการของจุดประสงค์เชิงพฤติกรรมชุดหนึ่ง ให้ครูผู้สอนจำนวน 5 ท่านพิจารณาจับคู่ข้อสอบกับจุดประสงค์เชิงพฤติกรรม เป็นการหาความเที่ยงตรงตามเนื้อหา (Hambleton . 1978 : 34) ว่าจุดประสงค์เชิงพฤติกรรมแต่ละอันสามารถวัดได้ด้วยข้อสอบใดบ้าง ข้อสอบใดที่ไม่ได้รับการจับคู่หรือจับคู่ผิดจุดประสงค์เชิงพฤติกรรมที่ต้องการมากกว่า 2 เสียงก็ต้องปรับปรุงข้อสอบนั้นใหม่ แล้วนำข้อสอบที่ปรับปรุงแล้วไปให้ครูผู้สอนพิจารณาว่าวัดจุดประสงค์ที่ต้องการวัดหรือยัง ทำการปรับปรุงจนได้ผลที่พอใจว่า ข้อสอบแต่ละข้อวัดจุดประสงค์ที่ต้องการวัด

8. นำแบบทดสอบอิงเกณฑ์ที่สุ่มจากประชากรเนื้อหาและมีความเที่ยงตรงตามเนื้อหาดีแล้ว คือสามารถวัดจุดประสงค์เชิงพฤติกรรมที่ต้องการวัด ไปทดสอบกับนักเรียนชั้นประถมศึกษาปีที่ 2 ในจังหวัดสมุทรสงครามจำนวนกลุ่มตัวอย่าง 100 คน โดยทำการสอบก่อนเรียนและเว้นช่วงเมื่อนักเรียนเรียนจบเนื้อหาย่อยแต่ละตอน

จึงสอบหลังเรียน กำหนดเวลาสอบครั้งนี้ แบบทดสอบฉบับที่ 1 ใช้เวลา 45 นาที ส่วนแบบทดสอบฉบับที่ 2,3 และ 4 ใช้เวลา 30 นาที โดยพิจารณาพร้อมกับครูผู้สอน

9. นำข้อสอบมาวิเคราะห์รายข้อ เพื่อปรับปรุงคุณภาพข้อสอบ

9.1 คำนึงความยากของข้อสอบ โดยหาเปอร์เซ็นต์ของผู้ตอบถูกในข้อสอบนั้น (Berk. 1978 : 4) จากการหาค่านี้นี้ความยากในการสอบก่อนเรียนและหลังเรียน ข้อสอบที่ดีควรเป็นข้อสอบยากในกลุ่มก่อนเรียนแต่ง่ายในกลุ่มหลังเรียน ซึ่งเป็นการแปรเปลี่ยนค่าความยากในการสอบทั้งสองครั้ง ดังนั้นในการพิจารณาคัดเลือกข้อสอบจึงพิจารณากำหนดค่าความยากในการสอบหลังเรียนตั้งแต่ 50% ขึ้นไป และมีค่าความยากในการสอบก่อนเรียนต่ำกว่า 50 %

9.2 คำนึงอำนาจจำแนกของข้อสอบ คำนวณจากสัดส่วนของผู้ตอบผิดในการสอบก่อนเรียนแก่ตอบถูกในการสอบหลังเรียน ลบด้วยสัดส่วนของผู้ตอบผิดในการสอบทั้งสองครั้ง (Kosecoff and Klein. 1976 : 276 - 293) ซึ่งเป็นสูตรเสนอความก้าวหน้ารายบุคคล ข้อสอบที่ดีควรมีค่าตั้งแต่ .00 ขึ้นไป แต่ในการคัดเลือกต้องการให้ข้อสอบมีค่าอำนาจจำแนกสูงพอสมควร จึงกำหนดค่านี้อำนาจจำแนกของข้อสอบ สูงกว่าหรือเท่ากับ .20

9.3 ปรับปรุงแก้ไขข้อสอบที่ไม่ได้เกณฑ์ในข้อ 9.1 และ 9.2

10. ทดลองใช้แบบทดสอบครั้งที่ 2 กับกลุ่มตัวอย่างจำนวน 146 คน โดยทำการสอบก่อนเรียนและหลังเรียนเมื่อจบเนื้อหาย่อยแต่ละตอน

11. วิเคราะห์ข้อสอบรายข้อ คัดเลือกและปรับปรุงข้อสอบโดยใช้หลักเกณฑ์ เช่นเดียวกับข้อ 9.1 - 9.3 คัดเลือกข้อสอบตามจำนวนข้อที่ระบุในตาราง 10 กรณีข้อสอบที่มีคุณภาพรายข้อเกินกว่าจำนวนที่กำหนด ก็จะใช้ข้อสอบที่มีอำนาจจำแนกต่ำออกไป

12. ทดสอบครั้งสุดท้ายกับกลุ่มตัวอย่าง 124 คน โดยสอบก่อนเรียนและหลังเรียน เมื่อจบเนื้อหาย่อยแต่ละตอน

13. กำหนดหาคะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรม

ในแบบทดสอบอิงเกณฑ์ทั้งสี่ฉบับด้วยวิธีการของเบอร์ก (Berk. 1976 : 4 - 9) ร่วมกับการพิจารณาคะแนนเกณฑ์ของครูผู้สอน โดยนำคะแนนที่ได้จากทั้ง 2 วิธีมาหาค่าเฉลี่ยเป็นคะแนนเกณฑ์ เพื่อตัดสินการสอบผ่าน - ไม่ผ่านในแต่ละจุดประสงค์เชิงพฤติกรรม

14. ตรวจสอบหาคุณภาพของแบบทดสอบ

14.1 หาค่าความเชื่อมั่นของแบบทดสอบแต่ละฉบับ โดยใช้สูตรการหาความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ของซับโคเวียค (Subkoviak. 1976 : 265 - 276)

14.2 หาความเที่ยงตรงตามเนื้อหาของแบบทดสอบแต่ละฉบับ โดยใช้วิธีการ 2 วิธีคือ ตรวจสอบโดยนำแบบทดสอบที่สร้างขึ้นและรายการของจุดประสงค์เชิงพฤติกรรม ให้ครูผู้สอนและผู้เชี่ยวชาญทางเนื้อหาวิชาได้แก่ศึกษานิเทศก์จำนวน 2 ท่านพิจารณาจับคู่ข้อสอบกับรายจุดประสงค์เชิงพฤติกรรม ซึ่งพิจารณาว่าข้อสอบแต่ละข้อเป็นตัวแทนของพฤติกรรมที่ต้องการวัดหรือไม่ และอีกวิธีหนึ่งโดยใช้สูตรหาความเที่ยงตรงตามเนื้อหาของคาร์เวอร์ (Carver. 1970 : 256)

15. จัดทำคู่มือดำเนินการสอบ

16. จัดพิมพ์รูปเล่ม

ตัวอย่างข้อสอบ

แบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่องการหาร ชั้นประถมศึกษาปีที่ 2 ที่สร้างขึ้นเป็นแบบทดสอบปรนัยชนิดเลือกตอบ 3 ตัวเลือก แบ่งเป็นฉบับย่อย 4 ฉบับ ดังนี้

ฉบับที่ 1 ความหมายของการหาร

ตอนที่ 1 การนับลด

จุดประสงค์เชิงพฤติกรรม 1.1.1 "เมื่อกำหนดจำนวนนับลด 2 - 3 จำนวน นักเรียนสามารถบอกจำนวนต่อไปหรือจำนวนที่ขาดหายไปได้" ดังตัวอย่างข้อ ๐)

ข้อ ๐) 21, 18, 15, ... , 9

ก. 14

ข. 12

ค. 11

จุดประสงค์เชิงพฤติกรรม 1.1.2 "เมื่อกำหนดจำนวนเลขหรือสิ่งของ นักเรียนสามารถบอกจำนวนครั้งที่เอาออกจนหมดหรือนับลดได้" ดังตัวอย่างข้อ ๐)

ข้อ ๐) เริ่มจาก 16 ลบออกทีละ 4 ต้องลบออกกี่ครั้งจึงหมด

ก. 3 ครั้ง

ข. 4 ครั้ง

ค. 5 ครั้ง

จุดประสงค์เชิงพฤติกรรม 1.1.3 "เมื่อกำหนดจำนวนเลขหรือสิ่งของ นักเรียนสามารถบอกจำนวนครั้งที่เอาออกจนหมดและเศษที่เหลือได้" ดังตัวอย่างข้อ ๐)

ข้อ ๐) กระดุม 16 เม็ด หยิบมาใช้ครั้งละ 5 เม็ด เอาออกมาใช้กี่ครั้ง จะเหลือกระดุมกี่เม็ด



ก. 4 ครั้ง เหลือ 2 เม็ด

ข. 3 ครั้ง เหลือ 2 เม็ด

ค. 3 ครั้ง เหลือ 1 เม็ด

ตอนที่ 2 ความรู้เบื้องต้นเรื่องการหาร

จุดประสงค์เชิงพฤติกรรม 1.2.1 "เมื่อกำหนดสิ่งของให้จำนวนหนึ่งและบอกจำนวนกลุ่ม นักเรียนสามารถหาจำนวนในแต่ละกลุ่มได้" ดังตัวอย่างข้อ ๐)

ข้อ ๐) ลูกหิน 24 ลูก แบ่งเป็น 6 กองเท่าๆกัน จะมีลูกหินกองละกี่ลูก



ก. 4 ลูก

ข. 5 ลูก

ค. 6 ลูก

จุดประสงค์เชิงพฤติกรรม 1.2.2 " เมื่อกำหนดจำนวนสิ่งของที่เท่ากันของแต่ละกลุ่ม และของทั้งหมด นักเรียนสามารถหาจำนวนกลุ่มได้ " ดังตัวอย่างข้อ ๐)

ข้อ ๐) ลูกบอลทั้งหมด 9 ลูก เก็บใส่กล่องละ 3 ลูก ต้องใช้กล่องกี่ใบ



ก. 6 ใบ

ข. 4 ใบ

ค. 3 ใบ

ฉบับที่ 2 สัญลักษณ์การหาร

ตอนที่ 1 ประโยกสัญลักษณ์การหาร

จุดประสงค์เชิงพฤติกรรม 2.1 " เมื่อกำหนดภาพหรือโจทย์ให้ นักเรียนสามารถเขียน อยู่ในรูปประโยคสัญลักษณ์ได้ " ดังตัวอย่างข้อ ๐) และข้อ ๐๐)

ข้อ ๐)



จากภาพจะแทนด้วยข้อใดจึงจะถูกต้อง

ก. $8 \div 2 = 4$ ข. $8 \div 4 = 4$ ค. $4 \div 2 = 2$

ข้อ ๐๐) ข้อใดมีความหมายเช่นเดียวกับ $10 \div 2 = 5$

ก. แดงมีเงิน 10 บาท ซื้อขนม 2 บาท แดงจะมีเงินกี่บาท

ข. แดงมีเงิน 10 บาท ได้ทำงาน 2 บาท ได้เงินกี่บาท

ค. แดงมีเงิน 10 บาท ซื้อสมุดที่ละ 2 บาท ใ้สมุดกี่เล่ม

ตอนที่ 2 ความสัมพันธ์ระหว่างการคูณและการหาร

จุดประสงค์เชิงพฤติกรรม 2.2 " เมื่อกล่าวถึงโจทย์การหาร นักเรียนสามารถใช้ ความสัมพันธ์ระหว่างการคูณและการหารในการหาตัวตั้ง ตัวหาร และผลหารได้ " ดังตัวอย่างข้อ ๐) , ข้อ ๐๐) และ ข้อ ๐๐๐)

ข้อ ก) $8 \times 5 = 40$ ดังนั้น $40 \div 8 = \square$

ก. 32

ข. 8

ค. 5

ข้อ ง) $4 \times \square = 16$ ดังนั้น $16 \div \square = 4$

ก. 6

ข. 4

ค. 3

ข้อ ๐๐) $\square \times 7 = 42$ ดังนั้น $42 \div 7 = \square$

ก. 35

ข. 7

ค. 6

ฉบับที่ 3 โจทย์การหาร

ตอนที่ 1 การหารลงตัว

จุดประสงค์เชิงพฤติกรรม 3.1 " เมื่อกำหนดโจทย์การหารที่มีตัวหารหลักเดียว ผลหารไม่เกิน 10 นักเรียนสามารถทำโจทย์เกี่ยวกับการหารลงตัวได้ "

ดังตัวอย่าง ข้อ ๐) , ข้อ ๐๐) และ ข้อ ๐๐๐)

ข้อ ๐) $40 \div 4 = \square$

ก. 36

ข. 10

ค. 1

ข้อ ๐๐) $28 \div \square = 7$

ก. 21

ข. 4

ค. 3

ข้อ ๐๐๐) $\square \div 5 = 5$

ก. 25

ข. 15

ค. 1๐

ตอนที่ 2 การหารมีเศษ

จุดประสงค์เชิงพฤติกรรม 3.2 " เมื่อกำหนดโจทย์การหารที่มีตัวหารหลักเดียว ผลหารไม่เกิน 10 นักเรียนสามารถทำโจทย์เกี่ยวกับการหารมีเศษได้ "

ดังตัวอย่างข้อ ๐)

ข้อ ๐) $7 \div 2 = \square$ เศษ $\square$

ก. 1 เศษ 2

ข. 2 เศษ 1

ค. 3 เศษ 2

ฉบับที่ 4 โจทย์ปัญหาการหาร

ตอนที่ 1 การหารลงตัว

จุดประสงค์เชิงพฤติกรรม 4.1 " เมื่อกำหนดโจทย์ปัญหาการหารที่มีตัวหารหลักเดียว ผลหารไม่เกิน 10 นักเรียนสามารถหาคำตอบแบบลงตัวได้ " ดังตัวอย่างข้อ 0)

ข้อ 0) ลำไย 27 ผล แบ่งให้นักอง 3 คนเท่าๆกัน จะได้คนละกี่ผล

ก. 9 ผล

ข. 8 ผล

ค. 7 ผล

ตอนที่ 2 การหารมีเศษ

จุดประสงค์เชิงพฤติกรรม 4.2 " เมื่อกำหนดโจทย์ปัญหาการหารที่มีตัวหารหลักเดียว ผลหารไม่เกิน 10 นักเรียนสามารถหาคำตอบแบบมีเศษได้ " ดังตัวอย่างข้อ ๑)

ข้อ ๑) ข้าวโพด 46 ผล จัดใส่ถาดละ 5 ผล จะต้องใช้ถาดกี่ใบเหลือกี่ผล

ก. 9 ใบ เหลือ 1 ผล

ข. 8 ใบ เหลือ 1 ผล

ค. 7 ใบ เหลือ 1 ผล

วิธีดำเนินการรวบรวมข้อมูล

1. ติดต่อโรงเรียนที่ใช้เป็นกลุ่มตัวอย่าง เพื่อกำหนดวันและเวลาในการสอบ
2. เตรียมข้อสอบให้พอเพียงกับจำนวนนักเรียนที่สอบในแต่ละครั้ง และวางแผนในการดำเนินการสอบล่วงหน้า เช่น จัดกรรมการคุมสอบ การชี้แจงขั้นตอน และวิธีดำเนินการสอบต่อกรรมการคุมสอบ เป็นต้น

3. อธิบายให้นักเรียนในกลุ่มตัวอย่างเข้าใจวัตถุประสงค์ของการทำข้อสอบ

4. นำแบบทดสอบไปทดสอบกับกลุ่มตัวอย่าง โดยแต่ละฉบับกำหนดเวลาในการสอบดังนี้ ฉบับที่ 1 เวลา 45 นาที ส่วนในฉบับที่ 2, 3 และ 4 เวลา 30 นาที การทดสอบมีจำนวน 3 ครั้ง ซึ่งแต่ละครั้งมีจุดมุ่งหมาย ดังนี้

4.1 ทดสอบครั้งที่ 1 เพื่อวิเคราะห์รายข้อ และปรับปรุงข้อสอบ

4.2 ทดสอบครั้งที่ 2 เพื่อวิเคราะห์รายข้อ คัดเลือกและปรับปรุงข้อสอบ

ผลการวิเคราะห์ข้อมูล

สัญลักษณ์ และอักษรรย่อที่ใช้ในการวิเคราะห์ข้อมูล

เพื่อให้เกิดความเข้าใจตรงกันในการแปลความหมายของการวิเคราะห์ข้อมูล ผู้วิจัยกำหนดสัญลักษณ์และอักษรรย่อที่ใช้ในการวิเคราะห์ข้อมูล ดังนี้

P_1	แทน	ค่าความยากของข้อสอบในการสอบก่อนเรียน
P_2	แทน	ค่าความยากของข้อสอบในการสอบหลังเรียน
G	แทน	ค่าอำนาจจำแนกของข้อสอบในการสอบทั้งสองครั้ง

การเสนอผลการวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลและแปลผลการวิเคราะห์ข้อมูลในการศึกษาค้างนี้ ผู้วิจัยเสนอผลการวิเคราะห์ข้อมูล ดังนี้

1. ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 1 จากการทดสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม
2. ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 2 จากการทดสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม
3. ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 3 จากการทดสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม

4. ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และ
ค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 4 จากการทดสอบครั้งที่หนึ่ง ครั้งที่สอง
และครั้งที่สาม
5. คะแนนเกณฑ์ของจุดประสงค์เชิงพฤติกรรมในแบบทดสอบฉบับที่ 1 จากการทดสอบ
ครั้งที่สาม
6. คะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรมในแบบทดสอบฉบับที่ 2 จาก
การทดสอบครั้งที่สาม
7. คะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรมในแบบทดสอบฉบับที่ 3 จาก
การทดสอบครั้งที่สาม
8. คะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรมในแบบทดสอบฉบับที่ 4 จาก
การทดสอบครั้งที่สาม
9. ค่าความเชื่อมั่นของแบบทดสอบแต่ละฉบับ จากการทดสอบครั้งที่สาม
10. ค่าความเที่ยงตรงของแบบทดสอบแต่ละฉบับ จากการทดสอบครั้งที่สาม

ผลการวิเคราะห์ข้อมูล

1. ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และ
ค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 1 จากการทดสอบครั้งที่หนึ่ง ครั้งที่สอง
และครั้งที่สาม

ผู้วิจัยนำคะแนนจากการทดสอบทั้งสามครั้ง มาคำนวณหาความยาก
โดยใช้สูตรของ เบอร์ก์ ในการสอบก่อนเรียนและสอบหลังเรียน และหาค่าอำนาจจำแนก
โดยใช้วิธีของโคสคอฟฟ์ และเคลน ดังแสดงในตาราง 11

ตาราง 11 ค่าความยากในการสอบก่อนเรียนและค่าความยากในการสอบหลังเรียน

ค่าอ่านาจำแนกของข้อสอบในแบบทดสอบฉบับที่ 1 จากการสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม

จุดประสงค์ เชิงพฤติกรรม	ทดสอบครั้งที่หนึ่ง				ทดสอบครั้งที่สอง				ทดสอบครั้งที่สาม			
	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G
1.1.1	1	29	79	.24	1	31	73	.34	1	32	75	.28
	2	31	79	.39	2	25	76	.35	2	34	72	.29
	3	28	75	.37	3	30	72	.38	3	33	68	.32
	4	17	45	.24	4*	25	59	.28				
	5	25	66	.35	5	15	63	.37	4	30	67	.37
	6	17	70	.38	6	18	54	.37	5	28	71	.20
	7	19	53	.29	7*	22	58	.28				
1.1.2	8	32	68	.28	8	41	76	.34	6	40	75	.30
	9	46	83	.30	9*	52	87	.23	7	35	73	.25
	10	46	80	.23	10*	48	77	.15	8	30	61	.25
	11	35	71	.32	11	25	80	.36	9	45	.67	.25
	12	30	65	.31	12	42	79	.35	10	39	71	.32
	13	30	65	.31	13	34	58	.32				
	14	20	61	.22	14	35	61	.37				

ตาราง 11 (ต่อ)

จุดประสงค์ เชิงพฤติกรรม	ทดสอบครั้งที่หนึ่ง				ทดสอบครั้งที่สอง				ทดสอบครั้งที่สาม			
	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G
1.1.3	15	27	65	.28	15	30	62	.34	11	36	61	.25
	16	25	80	.30	16	25	76	.35	12	32	72	.35
	17	27	65	.22	17	24	66	.31	13	37	58	.25
	18	25	59	.19	18	24	66	.35	14	30	58	.28
	19	28	62	.21	19	25	69	.30	15	33	61	.32
	20	34	59	.18	20*	34	62	.22				
	21	25	55	.21	21*	30	61	.24				
1.2.1	22	42	79	.33	22	30	80	.37	16	46	83	.30
	23	41	78	.29	23	28	72	.34	17	44	84	.26
	24	31	80	.42	24	23	77	.32	18	50	85	.23
	25	34	83	.37	25	21	77	.28	19	46	80	.23
	26	38	75	.24	26	30	70	.25	20	35	73	.20
	27	34	81	.31	27	32	72	.29	21	37	78	.21
	28	31	73	.26	28	38	82	.39	22	41	80	.27
	29	27	72	.21	29*	31	65	.23				
	30	29	68	.18	30*	28	61	.21				

ตาราง 11 (ต่อ)

จุดประสงค์ เชิงพฤติกรรม	ทดสอบครั้งที่หนึ่ง				ทดสอบครั้งที่สอง				ทดสอบครั้งที่สาม			
	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G
1.2.2	31	28	72	.27	31	35	79	.31	23	28	80	.34
	32	36	81	.32	32	31	73	.34	24	32	75	.28
	33	40	84	.36	33	38	72	.26	25	37	75	.23
	34	32	76	.20	34*	39	66	.18				
	35	27	71	.22	35	24	65	.22	26	27	75	.32
	36	27	76	.26	36	21	73	.34	27	32	69	.22
	37	24	70	.25	37	28	72	.24	28	28	74	.27
	38	34	65	.17	38*	31	65	.19				
	39	36	58	.14	39*	34	55	.10				
	40	23	59	.24	40	27	66	.24	29	28	74	.27

* ข้อที่คัดออกจากแบบทดสอบ

ผลการวิเคราะห์ข้อมูลตามตาราง 11 ปรากฏว่า แบบทดสอบฉบับที่ 1 จุดประสงค์
เชิงพฤติกรรม 1.1.1 จำนวน 7 ข้อ เมื่อนำไปทดสอบครั้งที่หนึ่งแล้ว ข้อสอบมีค่าความยาก
ในการสอบก่อนเรียนตั้งแต่ 17 % - 31 % ค่าความยากในการสอบหลังเรียน ตั้งแต่ 45 % -
79 % ข้อสอบที่มีค่าความยากในการสอบหลังเรียนไม่ถึง 50 % มีเพียงข้อเดียว ส่วนค่าอำนาจ
จำแนกของการสอบทั้งสองครั้งมีค่าตั้งแต่ .24 - 39 ข้อสอบทุกข้อมีค่าอำนาจจำแนกสูงกว่า .20

ตามต้องการ แสดงว่าข้อสอบจำนวน 6 ข้อมีคุณภาพตามเกณฑ์ที่ต้องการปรับปรุงข้อสอบที่มีค่าความยากไม่ถึงเกณฑ์ 1 ข้อให้สูงขึ้น เมื่อนำไปทดสอบครั้งที่สอง ปรากฏว่า ข้อสอบมีค่าความยากในการสอบก่อนเรียน ตั้งแต่ 15 % - 31 % ค่าความยากในการสอบหลังเรียน มีค่าตั้งแต่ 54 % - 76 % ข้อสอบทุกข้อมีค่าความยากก่อนเรียนต่ำกว่า 50% และค่าความยากหลังเรียนสูงกว่า 50 % ตามต้องการ ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .28 - .38 ซึ่งมีค่าอำนาจจำแนกสูงกว่า .20 แสดงว่าข้อสอบทุกข้อมีคุณภาพรายข้อตามต้องการ แต่จุดประสงค์เชิงพฤติกรรมข้อนี้กำหนดข้อสอบไว้จำนวน 5 ข้อ ฉะนั้นจึงต้องตัดออก 2 ข้อ โดยคัดเลือกข้อสอบที่มีค่าอำนาจจำแนกต่ำสุดออกไป เมื่อนำไปทดสอบครั้งที่สาม ปรากฏว่า ข้อสอบมีค่าความยากในการสอบก่อนเรียนตั้งแต่ 28 % - 34 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 67 % - 75 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .20 - .37 แสดงว่าข้อสอบทุกข้อมีคุณภาพรายข้อตามต้องการ

จุดประสงค์เชิงพฤติกรรม 1.1.2 จำนวน 7 ข้อ เมื่อนำไปทดสอบครั้งที่หนึ่งแล้ว ปรากฏว่าข้อสอบมีค่าความยากตั้งแต่ 20 % - 46 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 61 % - 83 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .22 - .32 แสดงว่าข้อสอบทุกข้อมีค่าความยากและค่าอำนาจจำแนกตามต้องการ นำข้อสอบไปทดสอบครั้งที่สอง ปรากฏว่า ข้อสอบมีค่าความยากในการสอบก่อนเรียนตั้งแต่ 25 % - 52 % ค่าความยากในการสอบหลังเรียน มีค่าตั้งแต่ 58 % - 87 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .15 - .37 ข้อสอบที่มีค่าอำนาจจำแนกต่ำกว่า .20 มีเพียงข้อเดียว แสดงว่ามีข้อสอบจำนวน 6 ข้อ มีค่าความยากและค่าอำนาจจำแนกตามต้องการ จุดประสงค์เชิงพฤติกรรมข้อนี้กำหนดข้อสอบจำนวน 5 ข้อ ฉะนั้นจึงคัดเลือกข้อสอบที่ไม่มีคุณภาพและข้อสอบที่มีค่าอำนาจจำแนกต่ำสุดออกไป เมื่อทดสอบครั้งที่สามแล้วปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 30 % - 45 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 61 % - 75 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .25 - .32 แสดงว่าข้อสอบทุกข้อมีคุณภาพรายข้อตามต้องการ

จุดประสงค์เชิงพฤติกรรม 1.1.3 จำนวน 7 ข้อ เมื่อนำไปทดสอบครั้งที่หนึ่งแล้ว
ปรากฏว่าข้อสอบมีค่าความยากในการสอบก่อนเรียนตั้งแต่ 25% - 34 % ค่าความยากในการสอบ
หลังเรียนมีค่าตั้งแต่ 55 % - 80 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .18 - .30 ข้อสอบที่มีค่า
อำนาจจำแนกไม่ถึง .20 มีจำนวน 2 ข้อ ปรับปรุงข้อสอบที่มีคุณภาพไม่ถึงเกณฑ์ที่ต้องการใหม่
แล้วนำไปทดสอบครั้งที่สอง ปรากฏค่าความยากในการสอบก่อนเรียนตั้งแต่ 24 % - 34 %
ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 61 % - 76 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่
.22 - .35 แสดงว่าข้อสอบทุกข้อมีคุณภาพรายข้อตามต้องการ จุดประสงค์ข้อนี้กำหนดข้อสอบ
จำนวน 5 ข้อ ฉะนั้นจึงคัดข้อออก 2 ข้อ โดยคัดข้อสอบที่มี ค่าอำนาจจำแนกต่ำสุดออกไป
เมื่อนำไปทดสอบครั้งที่สาม ปรากฏว่าข้อสอบมีค่าความยากในการสอบก่อนเรียนตั้งแต่ 30 % -
37 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 58 % - 72 % ส่วนค่าอำนาจจำแนกมีค่า
ตั้งแต่ .25 - .35 แสดงว่าข้อสอบทุกข้อมีคุณภาพรายข้อตามต้องการ

จุดประสงค์เชิงพฤติกรรม 1.2.1 จำนวน 9 ข้อ เมื่อนำไปทดสอบครั้งที่หนึ่งแล้ว
ปรากฏว่าข้อสอบมีค่าความยากในการสอบก่อนเรียนตั้งแต่ 27 % - 42 % ค่าความยากใน
การสอบหลังเรียนมีค่าตั้งแต่ 68 % - 83 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .18 - .42
แสดงว่าข้อสอบมีคุณภาพรายข้อตามต้องการ เมื่อนำไปทดสอบครั้งที่สอง ปรากฏว่าข้อสอบ
มีค่าความยากในการสอบก่อนเรียนตั้งแต่ 21 % - 38 % ค่าความยากในการสอบหลังเรียนมีค่า
ตั้งแต่ 61 % - 82 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .21 - .39 แสดงว่าข้อสอบทุกข้อมีค่า
ความยากและค่าอำนาจจำแนกตามต้องการ จุดประสงค์เชิงพฤติกรรมข้อนี้กำหนดข้อสอบไว้จำนวน
7 ข้อ ฉะนั้นจึงคัดข้อสอบที่มีค่าอำนาจจำแนกต่ำสุดออกไป 2 ข้อ เมื่อนำไปทดสอบครั้งที่สาม
ปรากฏว่าข้อสอบมีค่าความยากในการสอบก่อนเรียนตั้งแต่ 35 % - 50 % ค่าความยากในการสอบ
หลังเรียนมีค่าตั้งแต่ 73-85% ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .20 - .30 แสดงว่าข้อสอบทุกข้อ
มีคุณภาพรายข้อตามต้องการ

จุดประสงค์เชิงพฤติกรรม 1.2.2 จำนวน 10 ข้อ เมื่อนำไปทดสอบครั้งที่หนึ่งแล้ว ปรากฏว่าข้อสอบมีค่าความยากในการสอบก่อนเรียนตั้งแต่ 23 % - 40 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 58 % - 84 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .14 - .36 ข้อสอบที่มีค่าอำนาจจำแนกต่ำกว่า .20 มีจำนวน 2 ข้อ แสดงว่าต้องปรับปรุงข้อสอบทั้งสองข้อนี้ เมื่อนำไปทดสอบครั้งที่สองปรากฏว่า ข้อสอบมีค่าความยากในการสอบก่อนเรียนตั้งแต่ 21 % - 39 % ค่าความยากในการสอบหลังเรียนตั้งแต่ 55 % - 79 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .10 - .34 ข้อสอบที่มีค่าอำนาจจำแนกต่ำกว่า .20 มีจำนวน 3 ข้อ จุดประสงค์เชิงพฤติกรรมข้อนี้ กำหนดข้อสอบไว้จำนวน 7 ข้อ จึงตัดข้อสอบที่มีคุณภาพไม่ถึงเกณฑ์ที่ต้องการออกไป 3 ข้อ เมื่อนำไปทดสอบครั้งที่สาม ปรากฏว่าข้อสอบมีค่าความยากในการสอบก่อนเรียนตั้งแต่ 27 % - 37 % ค่าความยากในการสอบหลังเรียนตั้งแต่ 69 % - 80 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .22 - .34 แสดงว่าข้อสอบทุกข้อมีคุณภาพรายข้อตามต้องการ

2. ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และ
ค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 2 จากการทดสอบครั้งที่หนึ่ง ครั้งที่สอง
และครั้งที่สาม

ผู้วิจัยได้นำคะแนนจากการสอบทั้งสามครั้ง มาคำนวณหาค่าความยาก โดยใช้สูตรของ เบอร์ก และหาค่าอำนาจจำแนกโดยใช้วิธีของโคลสคอฟฟ์ และเคลน ก็แสดงไว้ในตาราง 12

ตาราง 12 ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และ
ค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 2 จากการทดสอบครั้งที่หนึ่ง ครั้งที่สอง
และครั้งที่สาม

จุดประสงค์ เชิงพฤติกรรม	ทดสอบครั้งที่หนึ่ง				ทดสอบครั้งที่สอง				ทดสอบครั้งที่สาม			
	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G
2.1	1	18	68	.32	1	22	61	.23	1	27	65	.22
	2	22	65	.24	2	24	66	.25	2	24	76	.31
	3	30	62	.22	3	21	63	.25	3	21	61	.26
	4	24	69	.28	4	20	66	.32	4	31	76	.30
	5	22	73	.35	5	28	70	.27	5	33	61	.21
	6	27	75	.29	6	32	72	.37	6	27	64	.35
	7	21	59	.18	7	26	63	.25	7	26	58	.22
	8	24	49	.14	8*	17	49	.18				
	9	25	79	.38	9	14	63	.34	8	30	66	.26
	10	29	56	.24	10*	18	46	.17				
2.2	11	33	52	.19	11*	26	41	.23				
	12	27	68	.26	12	14	54	.27	9	37	64	.21
	13	27	76	.37	13	20	61	.21	10	30	66	.28
	14	28	75	.32	14	23	68	.28	11	22	69	.30
	15	28	86	.41	15	24	70	.28	12	32	62	.21

ตาราง 12 (ต่อ)

จุดประสงค์ เชิงพฤติกรรม	ทดสอบครั้งที่หนึ่ง				ทดสอบครั้งที่สอง				ทดสอบครั้งที่สาม			
	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G
2.2 (ต่อ)	16	25	66	.27	16	18	69	.34	13	24	67	.24
	17	27	69	.24	17	13	72	.39	14	30	67	.22
	18	29	58	.17	18	21	65	.24	15	26	61	.23
	19	31	66	.20	19*	18	49	.19				
	20	34	48	.15	20	21	58	.26	16	29	53	.22

* ข้อที่ตัดออกจากแบบทดสอบ

ผลการวิเคราะห์ข้อมูลตามตาราง 12 ปรากฏว่า แบบทดสอบฉบับที่ 2 จำนวน 20 ข้อ เมื่อนำไปทดสอบครั้งที่หนึ่งแล้ว จุดประสงค์เชิงพฤติกรรม 2.1 จำนวน 10 ข้อ มีค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 18 % - 30 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 49 % - 79 % ข้อที่มีค่าความยาก 49 % มีเพียงข้อเดียว นอกนั้นมีค่าสูงกว่า 50 % ทั้งสิ้น ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .14 - .38 ข้อที่มีค่าอำนาจจำแนกต่ำกว่า .20 มี 2 ข้อ นอกนั้นมีค่าอำนาจจำแนกสูงกว่า .20 ทั้งสิ้น แสดงว่าข้อสอบส่วนใหญ่เป็นไปตามเกณฑ์ที่ต้องการ ส่วนข้อสอบที่ไม่ถึง เกณฑ์จำนวน 2 ข้อ นำมาปรับปรุงแก้ไข แล้วนำข้อสอบทั้งหมดไปทดสอบครั้งที่สอง ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 14 % - 32 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 46 % - 72 % ข้อสอบที่มีค่าความยากในการสอบหลังเรียนต่ำกว่า 50 %

มีจำนวน 2 ข้อ ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .17 - .34 ข้อสอบที่มีค่าอำนาจจำแนกต่ำกว่า .20 มีจำนวน 2 ข้อ ซึ่งเป็นข้อเดียวกับสองข้อแรกที่มีค่าความยากต่ำกว่า 50 % แสดงว่ามีข้อสอบจำนวน 8 ข้อ ที่มีค่าความยากและค่าอำนาจจำแนกตามเกณฑ์ที่ต้องการ ในจุดประสงค์เชิงพฤติกรรมนี้กำหนดข้อสอบจำนวน 8 ข้อ จึงตัดข้อสอบที่ไม่ได้ตามเกณฑ์ออกไป แล้วนำไปทดสอบครั้งที่สาม ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 21 % - 33 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 58 % - 76 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .21 - .35 แสดงว่าข้อสอบทุกข้อมีคุณภาพรายข้อตามเกณฑ์ที่ต้องการ

จุดประสงค์เชิงพฤติกรรม 2.2 จำนวน 10 ข้อ เมื่อนำไปทดสอบครั้งที่หนึ่งแล้ว ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 25 % - 34 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 48 % - 86 % ข้อสอบที่มีค่าความยากในการสอบหลังเรียนต่ำกว่า 50 % มีจำนวน 1 ข้อ ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .15 - .41 ข้อสอบที่มีค่าอำนาจจำแนกต่ำกว่า .20 มีจำนวน 3 ข้อ มีข้อสอบ 1 ข้อที่ซ้ำกับข้อที่มีค่าความยากในการสอบหลังเรียนไม่ถึง 50 % ปรับปรุงข้อสอบที่ไม่ได้ตามเกณฑ์ แล้วนำข้อสอบทั้งหมดไปทดสอบครั้งที่สอง ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 13 % - 26 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 41 % - 72 % ข้อสอบที่มีค่าความยากในการสอบหลังเรียนไม่ถึง 50 % มีจำนวน 2 ข้อ ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .19 - .39 ข้อสอบที่มีค่าอำนาจจำแนกไม่ถึง .20 มีจำนวน 11 ข้อ ซึ่งเป็นข้อเดียวกับหนึ่งในสองข้อแรกที่มีค่าความยากในการสอบหลังเรียนต่ำกว่า 50 % แสดงว่ามีข้อสอบจำนวน 8 ข้อ ที่มีค่าความยากและค่าอำนาจจำแนกตามเกณฑ์ที่ต้องการ จุดประสงค์เชิงพฤติกรรมนี้กำหนดจำนวนข้อสอบ 8 ข้อ จึงตัดข้อสอบที่ไม่ได้ตามเกณฑ์ออกไป แล้วนำไปทดสอบครั้งที่สาม ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 22 % - 37 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 53 % - 69 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .21 - .30 แสดงว่าข้อสอบทุกข้อมีคุณภาพรายข้อตามเกณฑ์ที่ต้องการ

เมื่อพิจารณาแบบทดสอบทั้งฉบับ จำนวน 20 ข้อ หลังจากนำไปทดสอบครั้งที่หนึ่ง ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 18 % - 34 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 48 % - 86 % ข้อสอบที่มีค่าความยากในการสอบหลังเรียนต่ำกว่า 50 % จำนวน 2 ข้อ ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .14 - .41 ข้อสอบที่มีค่าอำนาจจำแนกต่ำกว่า .20 มีจำนวน 5 ข้อ ซึ่งสองในห้าข้อเป็นข้อเดียวกับข้อสอบที่มีค่าความยากในการสอบหลังเรียนต่ำกว่า 50 % แสดงว่ามีข้อสอบจำนวน 15 ข้อ ที่มีค่าความยากและค่าอำนาจจำแนกตามเกณฑ์ที่ต้องการ ส่วนข้อสอบที่ไม่ได้ตามเกณฑ์ปรับปรุงใหม่ แล้วนำข้อสอบทั้งหมดไปทดสอบครั้งที่สอง ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 13 % - 32 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 41 % - 72 % ข้อสอบที่มีค่าความยากต่ำกว่า 50 % จำนวน 4 ข้อ ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .17 - .39 ข้อสอบที่มีค่าอำนาจจำแนกต่ำกว่า .20 จำนวน 3 ข้อ ซึ่งเป็นข้อสอบเดียวกับข้อสอบที่มีค่าความยากในการสอบหลังเรียนต่ำกว่า 50 % แสดงว่าข้อสอบจำนวน 16 ข้อ ที่มีค่าความยากในการสอบหลังเรียนและค่าอำนาจจำแนกตามเกณฑ์ที่ต้องการ จึงได้คัดเลือกข้อสอบที่ไม่ได้ตามเกณฑ์ในแต่ละตอนออกไป แล้วนำไปทดสอบครั้งที่สาม ปรากฏว่า ค่าความยากในการสอบก่อนเรียน มีค่าตั้งแต่ 21 % - 37 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 53 % - 76 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .21 - .35 แสดงว่าข้อสอบทุกข้อ มีคุณภาพรายข้อตามเกณฑ์ที่ต้องการ

3. ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และ
ค่าอำนาจจำแนกในการสอบทั้งสองครั้งของข้อสอบในแบบทดสอบฉบับที่ 3 จากการทดสอบ
ครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม

ผู้วิจัยได้นำคะแนนจากการสอบทั้งสามครั้งมาคำนวณหาค่าความยาก โดยใช้
สูตรของเบอร์ก และหาค่าอำนาจจำแนกโดยใช้วิธีของโคสคอฟฟ์ และเกสเนน ดังแสดงไว้
ในตาราง 13

ตาราง 13 ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และ
ค่าอำนาจจำแนกในการสอบทั้งสองครั้งของข้อสอบในแบบทดสอบฉบับที่ 3 จากการทดสอบ
ครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม

จุดประสงค์ เชิงพฤติกรรม	ทดสอบครั้งที่หนึ่ง				ทดสอบครั้งที่สอง				ทดสอบครั้งที่สาม			
	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G
3.1	1	34	81	.35	1	20	82	.38	1	30	80	.32
	2	30	82	.30	2	31	76	.25	2	34	78	.27
	3	21	78	.28	3	28	82	.36	3	25	76	.25
	4	26	71	.26	4	18	87	.38	4	33	72	.22
	5	23	79	.28	5	25	79	.27	5	27	74	.27
	6	33	75	.21	6	31	76	.23	6	34	63	.20
	7	32	64	.22	7*	34	68	.22				
	8	23	59	.15	8	13	66	.31	7	25	64	.24
	9	21	54	.20	9*	18	52	.25				

ตาราง 13 (ต่อ)

จุดประสงค์ เชิงพฤติกรรม	ทดสอบครั้งที่หนึ่ง				ทดสอบครั้งที่สอง				ทดสอบครั้งที่สาม			
	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G
3.1 (ทอ)	10	23	53	.20	10	20	72	.34	8	19	62	.31
	11	25	64	.30	* 11	15	53	.24				
	12	35	51	.13	* 12	25	56	.21				
3.2	13	22	55	.23	* 13	23	44	.08				
	14	28	61	.21	* 14	17	48	.15				
	15	28	43	.14	15	14	53	.24	9	19	58	.27
	16	25	59	.21	* 16	17	51	.19				
	17	18	64	.26	17	14	77	.32	10	20	63	.28
	18	22	57	.22	18	17	63	.24	11	18	68	.31
	19	13	61	.32	19	23	70	.28	12	18	66	.29
	20	13	54	.26	* 20	27	63	.22				
	21	20	58	.25	21	25	61	.23	13	23	54	.21
	22	21	77	.27	22	27	79	.24	14	22	68	.28
	23	17	71	.34	23	25	61	.23	15	23	65	.25
	24	21	71	.31	24	18	68	.28	16	23	66	.22

* ข้อที่คัดออกจากแบบทดสอบ

ผลการวิเคราะห์ตามตาราง 13 ปรากฏว่าแบบทดสอบฉบับที่ 3 จำนวน 24 ข้อ เมื่อนำไปทดสอบครั้งที่หนึ่งแล้ว จุดประสงค์เชิงพฤติกรรม 3.1 จำนวน 12 ข้อ ข้อสอบมีค่าความยากในการสอบก่อนเรียนตั้งแต่ 21 % - 35 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 51 % - 82 % ค่าอำนาจจำแนกในการสอบทั้งสองครั้งมีค่าตั้งแต่ .13 - .35 ข้อสอบที่มีค่าอำนาจจำแนกต่ำกว่า .20 มีจำนวน 2 ข้อ แสดงว่าข้อสอบที่มีค่าความยากในการสอบหลังเรียนและค่าอำนาจจำแนกตามเกณฑ์ที่ต้องการมีจำนวน 10 ข้อ ปรับปรุงข้อสอบที่ไม่ได้ตามเกณฑ์ทั้งสองข้อ แล้วนำไปสอบครั้งที่สอง ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 13 % - 34 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 52 % - 87 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .21 - .38 แสดงว่าข้อสอบทุกข้อมีค่าความยากในการสอบหลังเรียนและค่าอำนาจจำแนกตามเกณฑ์ที่ต้องการ ในจุดประสงค์เชิงพฤติกรรมนี้กำหนดข้อสอบจำนวน 8 ข้อ จึงตัดข้อสอบที่มีค่าอำนาจจำแนกต่ำสุดออกไปจำนวน 4 ข้อ แล้วนำไปทดสอบครั้งที่สาม ปรากฏว่าค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 19 % - 34 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 62 % - 80 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .20 - .32 แสดงว่าข้อสอบทุกข้อมีคุณภาพรายข้อตามเกณฑ์ที่ต้องการ

จุดประสงค์เชิงพฤติกรรม 3.2 จำนวน 12 ข้อ เมื่อนำไปทดสอบครั้งที่หนึ่ง ปรากฏว่าค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 13 % - 28 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 43 % - 77 % ข้อสอบที่มีค่าความยากในการสอบหลังเรียนต่ำกว่า 50 % มีเพียงข้อเดียว ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .14 - .34 ข้อสอบที่มีค่าอำนาจจำแนกต่ำกว่า .20 มีเพียงข้อเดียว ซึ่งเป็นข้อเดียวกับข้อสอบที่มีค่าความยากในการสอบหลังเรียนต่ำกว่า 50 % แสดงว่าข้อสอบจำนวน 11 ข้อ มีค่าความยากและค่าอำนาจจำแนกตามเกณฑ์ที่ต้องการ ปรับปรุงข้อสอบที่ไม่ได้คุณภาพตามเกณฑ์ แล้วนำไปทดสอบครั้งที่สอง ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 14 % - 27 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 44 % - 79 %

ข้อสอบที่มีค่าความยากต่ำกว่า 50 % มีจำนวน 2 ข้อ ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .08 - .32 ข้อสอบที่มีค่าอำนาจจำแนกต่ำกว่า .20 มีจำนวน 3 ข้อ ซึ่งสองในสามข้อเป็นข้อเกี่ยวกับข้อสอบที่มีค่าความยากในการสอบหลังเรียนต่ำกว่า 50 % แสดงว่าข้อสอบจำนวน 9 ข้อ มีคุณภาพตามเกณฑ์ที่กำหนด จุดประสงค์เชิงพฤติกรรมนี้กำหนดข้อสอบจำนวน 8 ข้อ จึงต้องคัดข้อสอบที่ไม่มีคุณภาพตามเกณฑ์และข้อที่มีค่าอำนาจจำแนกต่ำสุดออกไป แล้วนำข้อสอบจำนวน 8 ข้อ ไปทดสอบครั้งที่สาม ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 18 % - 23 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 54 % - 68 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .21 - .31 แสดงว่าข้อสอบทุกข้อมีคุณภาพรายข้อตามที่ต้องการ

เมื่อพิจารณาแบบทดสอบทั้งฉบับ จำนวน 24 ข้อ หลังจากนำไปทดสอบครั้งที่หนึ่ง ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 13 % - 35 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 43 % - 82 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .13 - .35 ข้อสอบที่มีค่าความยากในการสอบหลังเรียนต่ำกว่า 50 % และมีค่าอำนาจจำแนกต่ำกว่า .20 มีจำนวน 3 ข้อ นำมาปรับปรุงแก้ไขใหม่ แล้วนำไปทดสอบครั้งที่สอง ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 13 % - 34 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 44 % - 87 % ส่วนค่าอำนาจจำแนกของข้อสอบมีค่าตั้งแต่ .08 - .38 ข้อสอบที่มีค่าความยากในการสอบหลังเรียนต่ำกว่า 50 % และมีค่าอำนาจจำแนกต่ำกว่า .20 มีจำนวน 3 ข้อ เมื่อคัดเลือกข้อสอบในแต่ละตอนตามจำนวนข้อสอบที่กำหนด ได้ข้อสอบจำนวน 16 ข้อ โดยคัดข้อที่มีคุณภาพไม่ถึง เกณฑ์ที่ต้องการและข้อสอบที่มีค่าอำนาจจำแนกต่ำสุดออกไป แล้วนำไปทดสอบครั้งที่สาม ปรากฏว่าค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 18 % - 34 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 54 % - 80 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .20 - .32 แสดงว่าข้อสอบทุกข้อมีคุณภาพรายข้อตามที่ต้องการ

4. ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และ
ค่าอำนาจจำแนกในการสอบทั้งสองครั้งของข้อสอบในแบบทดสอบฉบับที่ 4 จากการทดสอบ
ครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม

ผู้วิจัยได้นำคะแนนจากการทดสอบทั้งสามครั้ง มาคำนวณหาความยากโดยใช้สูตรของเบอร์ก และหาค่าอำนาจจำแนกโดยใช้วิธีของโคลสคอฟฟ์ และเคนน ดังแสดงไว้ในตาราง 14

ตาราง 14 ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และ
ค่าอำนาจจำแนกในการสอบทั้งสองครั้งของข้อสอบในแบบทดสอบฉบับที่ 4 จากการทดสอบ
ครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม

จุดประสงค์ เชิงพฤติกรรม	ทดสอบครั้งที่หนึ่ง				ทดสอบครั้งที่สอง				ทดสอบครั้งที่สาม			
	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G
4.1	1	25	60	.26	1	34	73	.27	1	28	78	.32
	2	23	74	.28	2	17	82	.35	2	22	73	.30
	3	16	75	.33	3	21	73	.27	3	26	64	.23
	4	14	69	.31	4*	31	79	.22				
	5	19	77	.36	5	20	75	.30	4	26	72	.28
	6	23	60	.19	6	28	75	.24	5	22	70	.23
	7	33	65	.13	7	17	70	.32	6	22	66	.21
	8	24	54	.09	8	25	72	.27	7	24	71	.26
	9	22	60	.23	9*	34	65	.20				

ตาราง 14 (ต่อ)

จุดประสงค์ เชิงพฤติกรรม	ทดสอบครั้งที่หนึ่ง				ทดสอบครั้งที่สอง				ทดสอบครั้งที่สาม			
	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G	ข้อที่	P ₁	P ₂	G
4.1 (ต่อ)	10	21	62	.23	10	25	70	.27	8	24	64	.23
	11	17	48	.21	11*	31	55	.16				
	12	14	56	.24	12*	22	55	.18				
4.2	13	25	62	.24	13	25	66	.25	9	21	67	.24
	14	24	65	.26	14	18	63	.31	10	25	58	.21
	15	23	54	.18	15*	21	61	.22				
	16	24	58	.20	16	28	68	.23	11	18	61	.26
	17	18	48	.25	17	17	54	.25	12	26	58	.20
	18	15	64	.32	18	19	65	.23	13	23	62	.23
	19	18	73	.36	19	21	68	.22	14	24	69	.24
	20	22	48	.15	20*	38	59	.16				
	21	24	54	.22	21	30	68	.31	15	19	64	.25
	22	24	54	.20	22	17	54	.23	16	20	54	.22
	23	24	60	.21	23*	25	55	.21				
	24	22	64	.26	24*	24	59	.21				

* ข้อที่คัดออกจากแบบทดสอบ

ผลการวิเคราะห์ตามตาราง 14 ปรากฏว่า แบบทดสอบฉบับที่ 4 จำนวน 24 ข้อ เมื่อนำไปทดสอบครั้งที่หนึ่งแล้ว จุดประสงค์เชิงพฤติกรรม 4.1 จำนวน 12 ข้อ ข้อสอบมีค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 14 % - 33 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 48 % - 77 % ข้อสอบที่มีค่าความยากในการสอบหลังเรียนต่ำกว่า 50 % มีเพียงข้อเดียว ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .09 - .36 ข้อสอบที่มีค่าอำนาจจำแนกต่ำกว่า .20 มีจำนวน 3 ข้อ แสดงว่าข้อสอบจำนวน 8 ข้อ ที่มีคุณภาพตามเกณฑ์ที่ต้องการ ปรับปรุงข้อสอบที่มีคุณภาพไม่ถึงเกณฑ์ใหม่ แล้วนำไปทดสอบครั้งที่สอง ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 17 % - 34 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 55 % - 82 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .16 - .35 ข้อสอบที่มีค่าอำนาจจำแนกต่ำกว่า .20 มีจำนวน 2 ข้อ แสดงว่าข้อสอบจำนวน 10 ข้อ มีคุณภาพตามเกณฑ์ที่ต้องการ จุดประสงค์เชิงพฤติกรรมนี้กำหนดข้อสอบจำนวน 8 ข้อ จึงตัดข้อสอบที่ไม่ได้ตามเกณฑ์และข้อสอบที่มีค่าอำนาจจำแนกต่ำสุดออกไป 4 ข้อ แล้วนำไปทดสอบครั้งที่สาม ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 22 % - 28 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 64 % - 78 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .21 - .32 แสดงว่าข้อสอบทุกข้อมีคุณภาพรายข้อตามเกณฑ์ที่ต้องการ

จุดประสงค์เชิงพฤติกรรม 4.2 จำนวน 12 ข้อ เมื่อนำไปทดสอบครั้งที่หนึ่ง ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 15 % - 25 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 48 % - 73 % ข้อสอบที่มีค่าความยากในการสอบหลังเรียนต่ำกว่า 50 % มีจำนวน 2 ข้อ ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .15 - .36 ข้อสอบที่มีค่าอำนาจจำแนกต่ำกว่า .20 มีจำนวน 2 ข้อ ซึ่งหนึ่งในสองข้อนี้เป็นข้อเดียวกับข้อสอบที่มีค่าความยากในการสอบหลังเรียนต่ำกว่า 50 % แสดงว่ามีข้อสอบจำนวน 9 ข้อ ที่มีคุณภาพตามเกณฑ์ที่ต้องการ ปรับปรุงข้อสอบที่ไม่ได้คุณภาพใหม่ แล้วนำข้อสอบทั้งหมดไปทดสอบครั้งที่สอง ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 17 % - 38 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 54 % - 68 %

ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .16 - .31 ข้อสอบที่มีค่าอำนาจจำแนกต่ำกว่า .20 มีจำนวน 1 ข้อ แสดงว่าข้อสอบจำนวน 11 ข้อ มีคุณภาพตามเกณฑ์ที่ต้องการ จุดประสงค์เชิงพฤติกรรมนี้กำหนดข้อสอบจำนวน 8 ข้อ จึงคัดข้อสอบที่ไม่ได้คุณภาพและข้อสอบที่มีค่าอำนาจจำแนกต่ำสุดออกไป กรณีที่มีค่าอำนาจจำแนกเท่ากันเลือกข้อสอบที่มีค่าความยากสูงไว้ แล้วนำข้อสอบทั้ง 8 ข้อนี้ ไปทดสอบครั้งที่สาม ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 18 % - 26 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 54 % - 69 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .20 - .26 แสดงว่าข้อสอบทุกข้อมีคุณภาพรายข้อตามต้องการ

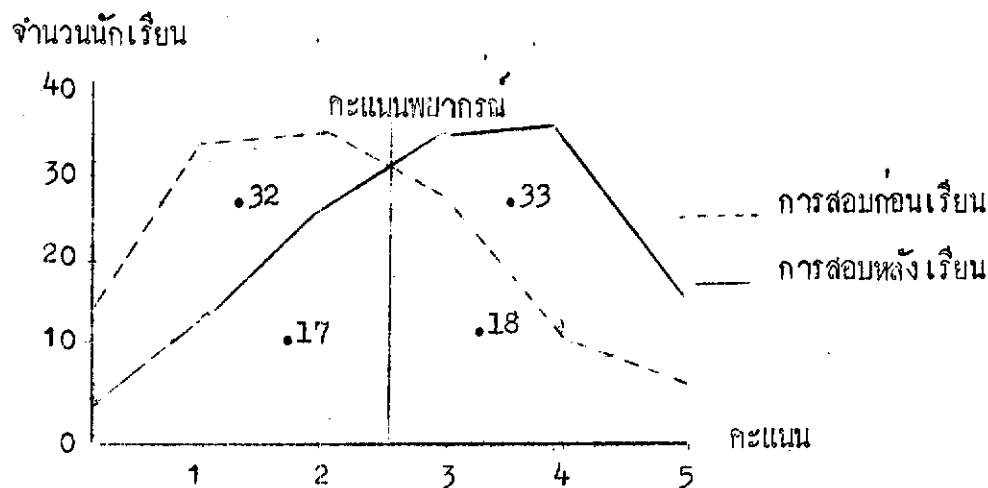
เมื่อพิจารณาแบบทดสอบทั้งฉบับ จำนวน 24 ข้อ หลังจากนำไปทดสอบครั้งที่หนึ่ง ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 14 % - 33 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 48 % - 77 % ข้อสอบที่มีค่าความยากหลังเรียนต่ำกว่า 50 % มีจำนวน 3 ข้อ ส่วนค่าอำนาจมีค่าตั้งแต่ .09 - .36 ข้อสอบมีค่าอำนาจจำแนกต่ำกว่า .20 มีจำนวน 5 ข้อ ซึ่งมีข้อสอบหนึ่งในห้าข้อนี้มีค่าความยากในการสอบหลังเรียนต่ำกว่า 50 % ด้วย แสดงว่ามีข้อสอบจำนวน 17 ข้อ ที่มีคุณภาพตามต้องการ เมื่อปรับปรุงข้อสอบที่ไม่ได้คุณภาพตามเกณฑ์ใหม่แล้ว นำไปทดสอบครั้งที่สอง ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 17 % - 38 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 54 % - 82 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .16 - .35 ข้อสอบที่มีค่าอำนาจจำแนกไม่ถึง .20 มีจำนวน 3 ข้อ แสดงว่ามีข้อสอบจำนวน 21 ข้อ ที่มีคุณภาพตามเกณฑ์ที่ต้องการ เมื่อคัดเลือกข้อสอบตามจำนวนในแต่ละตอนแล้ว โดยคัดข้อสอบที่มีคุณภาพไม่ถึง เกณฑ์และข้อที่มีอำนาจจำแนกต่ำสุดออกไป กรณีที่ข้อสอบมีค่าอำนาจจำแนกเท่ากันเลือกข้อที่มีค่าความยากในการสอบหลังเรียนสูง แล้วนำข้อสอบจำนวน 16 ข้อ ไปทดสอบครั้งที่สาม ปรากฏว่า ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 18 % - 28 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 54 % - 78 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .20 - .32 แสดงว่าข้อสอบทุกข้อมีคุณภาพรายข้อตามต้องการ

5. คะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรมในแบบทดสอบฉบับที่ 1

จากการทดสอบครั้งที่สาม

5.1 คะแนนเกณฑ์ของจุดประสงค์เชิงพฤติกรรม 1.1.1 ในแบบทดสอบฉบับที่ 1

ผู้วิจัยได้นำคะแนนผลการสอบของจุดประสงค์เชิงพฤติกรรม 1.1.1 จำนวน 5 ข้อ จากการสอบก่อนเรียนซึ่งถือเป็นกลุ่มนักเรียนที่ยังไม่ได้รับการสอน และจากการสอบหลังเรียนซึ่งถือเป็นกลุ่มนักเรียนที่ได้รับการสอนแล้ว ในการทดสอบครั้งที่สามกับกลุ่มตัวอย่างจำนวน 124 คน มาหาคะแนนเกณฑ์โดยใช้วิธีของเบอร์ก ได้คะแนนพยากรณ์จากจุดตัดของการกระจายคะแนนในการสอบทั้งสองครั้ง ดังภาพประกอบ 4



ภาพประกอบ 4 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบก่อนเรียน และคะแนนในการสอบหลังเรียนของจุดประสงค์เชิงพฤติกรรม 1.1.1 ในแบบทดสอบฉบับที่ 1

จากภาพประกอบ 4 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ
ทั้งสองครั้ง คือที่ 3 คะแนน

คะแนนแต่ละคะแนนนำมาหาความน่าจะเป็นในการตัดสินที่ถูกต้อง ความคลาดเคลื่อน
ในการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด เพื่อหาคะแนนเกณฑ์ใช้
ตัดสินการผ่านจุดประสงค์นี้ ดังในตาราง 15

ตาราง 15 ความน่าจะเป็นของการตัดสินที่ถูกต้อง ความคลาดเคลื่อนของการจำแนกผิด
และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัดของจุดประสงค์เชิงพฤติกรรม 1.1.1
ในแบบทดสอบฉบับที่ 1

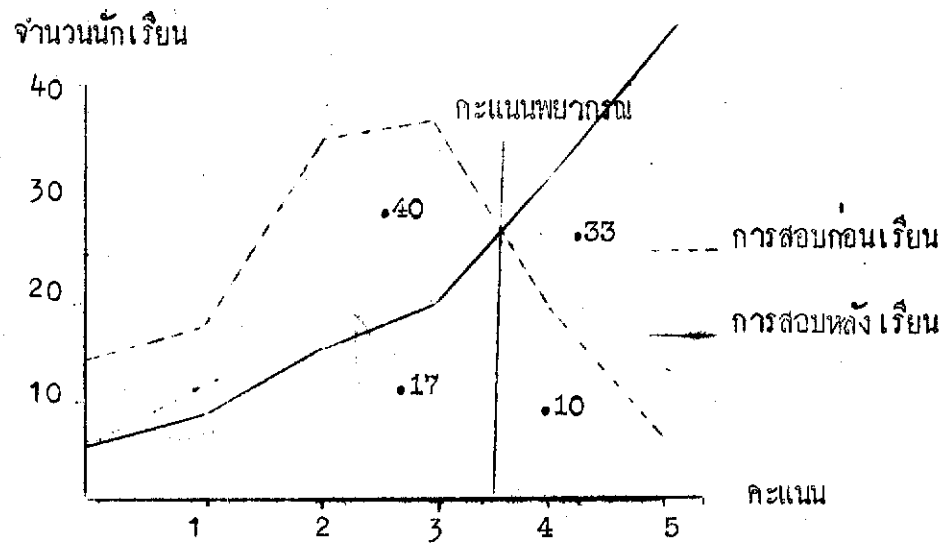
คะแนนจุดตัด	ความน่าจะเป็นของ การตัดสินที่ถูกต้อง	ความคลาดเคลื่อนของการจำแนกผิด ประเภทที่ 1 / ประเภทที่ 2	สัมประสิทธิ์ ความเที่ยงตรง
5	.53	.34 / .03	.08
4	.63	.30 / .07	.29
3*	.66	.17 / .18	.34
2	.62	.06 / .32	.29
1	.54	.01 / .45	.16

* คะแนนจุดตัดที่เหมาะสม

จากตาราง 15 ที่คะแนนจุดตัด 3 คะแนน ได้ค่าความน่าจะเป็นในการตัดสิน ที่ถูกต้องเท่ากับ .66 ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัดเท่ากับ .34 ซึ่งเป็นค่าสูงสุด แสดงว่าคะแนนจุดตัดที่เหมาะสมหรือคะแนนเกณฑ์โดยวิธีของ เบอร์กคือที่ 3 คะแนน ส่วนการพิจารณาของครูผู้สอนกำหนดคะแนนเกณฑ์เป็น 4 คะแนน เมื่อนำมาเฉลี่ยจะได้คะแนนเกณฑ์ที่ 4 คะแนน ได้ค่าความน่าจะเป็นในการตัดสินที่ถูกต้องจากตาราง 15 เท่ากับ .63 ความคลาดเคลื่อนประเภทที่ 1 (ผู้ที่ยืนยันแล้วแต่สอบไม่ผ่านเกณฑ์) และประเภทที่ 2 (ผู้ที่ยังไม่เรียนแต่สอบผ่านเกณฑ์) เท่ากับ .30 และ .07 ตามลำดับ ความเที่ยงตรงของคะแนนเกณฑ์ที่ 4 คะแนนเท่ากับ .26 ดังนั้นนักเรียนที่จะสอบผ่านจุดประสงค์เชิงพฤติกรรม 1.1.1 ที่ว่า "เมื่อกำหนดจำนวนนับค 2 - 3 จำนวน นักเรียนสามารถบอกจำนวนต่อไปหรือจำนวนที่ขาดหายไปได้" จะต้องทำข้อสอบได้ถูกต้องอย่างน้อย 4 ข้อ ในจำนวนข้อสอบทั้งหมด 5 ข้อ

5.2 คะแนนเกณฑ์ของจุดประสงค์เชิงพฤติกรรม 1.1.2 ในแบบทดสอบฉบับที่ 1

ผู้วิจัยได้นำคะแนนผลการสอบของจุดประสงค์เชิงพฤติกรรม 1.1.2 จำนวน 5 ข้อ จากการสอบก่อนเรียนและสอบหลังเรียน ในการทดสอบครั้งที่สามกับกลุ่มตัวอย่างจำนวน 124 คน มาหาคะแนนเกณฑ์โดยวิธีของ เบอร์ก ได้คะแนนพยากรณ์จากจุดตัดของการกระจายคะแนนในการสอบทั้งสองครั้ง ถึงภาพประกอบ 5



ภาพประกอบ 5 คะแนนพหุการณที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ
ก่อนเรียนและหลังเรียน ของจุดประสงค์เชิงพฤติกรรม 1.1.2
ในแบบทดสอบฉบับที่ 1

จากภาพประกอบ 5 คะแนนพหุการณที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ
ทั้งสองครั้ง คือที่ 4 คะแนน

คะแนนแต่ละคะแนนนำมาหาค่าความน่าจะเป็นในการตัดสินใจถูกต้อง ความคลาดเคลื่อน
ในการจำแนกชนิดและสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด เพื่อหาคะแนนเกณฑ์ที่ใช้ตัดสิน
การสอบในจุดประสงค์นี้ จึงลงในตาราง 16

ตาราง 16 ค่าความน่าจะเป็นในการตัดสินใจ ความคลาดเคลื่อนของการจำแนกผิด
และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์เชิงพฤติกรรม 1.1.2
ในแบบทดสอบฉบับที่ 1

คะแนนจุดตัด	ความน่าจะเป็นของ การตัดสินใจถูกต้อง	ความคลาดเคลื่อนของการจำแนกผิด ประเภทที่ 1 / ประเภทที่ 2	สัมประสิทธิ์ ความเที่ยงตรง
5	.69	.29 / .02	.45
4*	.74	.17 / .10	.48
3	.66	.10 / .25	.33
2	.57	.04 / .39	.19
1	.53	.02 / .45	.12

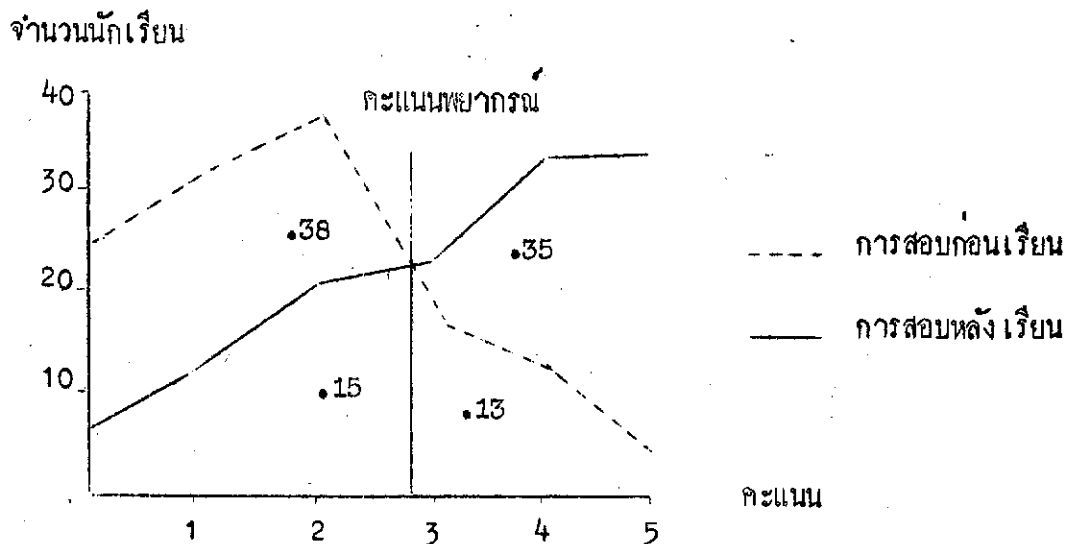
* คะแนนจุดตัดที่เหมาะสม

จากตาราง 16 ที่คะแนนจุดตัด 4 คะแนน ได้ค่าความน่าจะเป็นในการตัดสินใจ
เท่ากับ .74 ความคลาดเคลื่อนในการจำแนกผิดประเภทที่ 1 (ผู้ที่เรียนแล้วแต่สอบไม่ผ่านเกณฑ์)
และประเภทที่ 2 (ผู้ที่ยังไม่เรียนแต่สอบผ่านเกณฑ์) เท่ากับ .17 และ .10 ตามลำดับ
ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัดที่ 4 คะแนนเท่ากับ .48 ซึ่งมีค่าสูงสุด แสดงว่า
คะแนนจุดตัดที่เหมาะสมหรือคะแนนเกณฑ์โดยวิธีของ เบอร์กคือที่ 4 คะแนน และจากการพิจารณา
ของครูผู้สอนกำหนดคะแนนเกณฑ์เป็น 4 คะแนนเช่นกัน ดังนั้นนักเรียนที่จะสอบผ่านจุดประสงค์
เชิงพฤติกรรม 1.1.2 ที่ว่า "เมื่อกำหนดจำนวนเลข นักเรียนสามารถบอกจำนวนครั้งที่เอาออก
จนหมดหรือจำนวนที่นับลัดได้" นักเรียนจะต้องทำข้อสอบได้ถูกต้องอย่างน้อย 4 ข้อ ในจำนวนข้อสอบ
ทั้งหมด 5 ข้อ

5.3 คะแนนเกณฑ์ของจุดประสงค์เชิงพฤติกรรม 1.1.3 ในแบบทดสอบ

ฉบับที่ 1

ผู้วิจัยนำคะแนนจากการทดสอบครั้งที่สาม ของจุดประสงค์เชิงพฤติกรรม 1.1.3 จำนวน 5 ข้อ จากการสอบก่อนเรียนซึ่งเป็นกลุ่มนักเรียนที่ยังไม่ได้รับการสอน และการสอบหลังเรียนซึ่งเป็นกลุ่มนักเรียนที่ได้รับการสอนแล้ว มาหาคะแนนเกณฑ์โดยใช้วิธีของ เบอร์ก ได้คะแนนพยากรณ์ที่ได้จากจุดตัดในการสอบทั้งสองครั้ง ดังภาพประกอบ 6



ภาพประกอบ 6 คะแนนพยากรณ์ที่ได้จากจุดตัดการกระจายคะแนนในการสอบทั้งสองครั้ง
3 คะแนน

จากภาพประกอบ 6 คะแนนแต่ละคะแนนนำมาหาความน่าจะเป็นในการตัดสินใจถูกต้อง ความคลาดเคลื่อนในการจำแนกชนิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด เพื่อหาคะแนนเกณฑ์ที่ใช้ตัดสินการผ่านจุดประสงค์นี้ ดังตาราง 17

ตาราง 17 ค่าความน่าจะเป็นในการตัดสินใจ ความคลาดเคลื่อนของการจำแนกชนิด
และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์เชิงพฤติกรรม 1.1.3
ในแบบทดสอบฉบับที่ 1

คะแนนจุดตัด	ความน่าจะเป็นของ การตัดสินใจถูกต้อง	ความคลาดเคลื่อนของการจำแนกชนิด ประเภทที่ 1 / ประเภทที่ 2	สัมประสิทธิ์ ความเที่ยงตรง
5	.62	.37 / .01	.34
4	.71	.23 / .06	.44
3*	.73	.15 / .13	.46
2	.65	.07 / .28	.33
1	.58	.02 / .40	.22

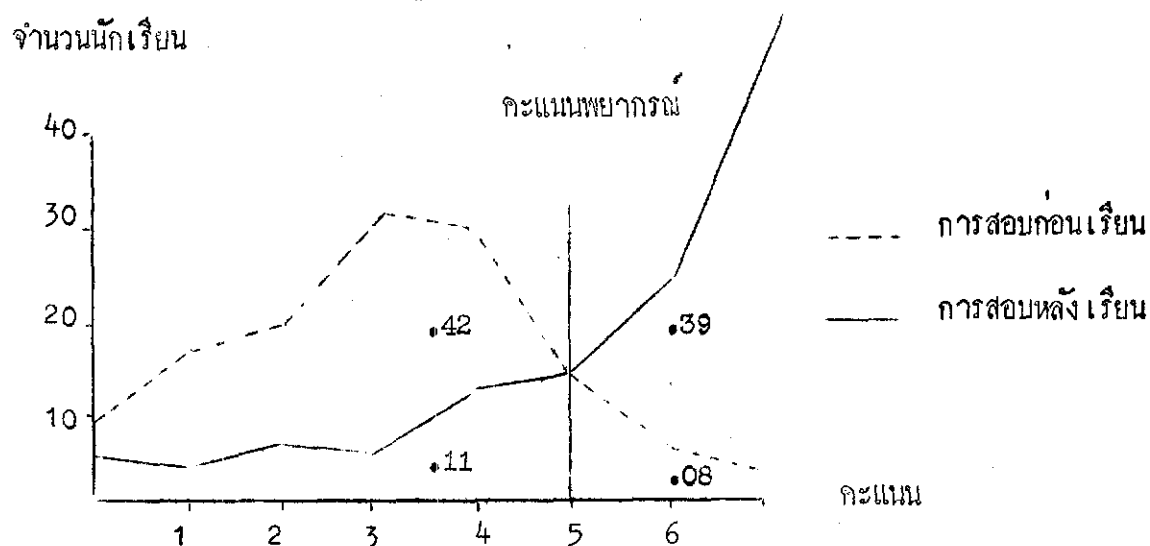
* คะแนนจุดตัดที่เหมาะสม

จากตาราง 17 ที่คะแนนจุดตัด 3 คะแนน ได้ค่าความน่าจะเป็นของการตัดสินใจถูกต้องเท่ากับ .73 ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัดเท่ากับ .46 ซึ่งมีค่าสูงสุด แสดงว่าคะแนนจุดตัดที่เหมาะสมหรือคะแนนเกณฑ์โดยวิธีของ เบอร์กคือที่ 3 คะแนน ส่วนการพิจารณาของครูผู้สอนกำหนดคะแนนเกณฑ์เป็น 4 คะแนน เมื่อนำมาเฉลี่ยจะได้คะแนนเกณฑ์ 4 คะแนน ซึ่งมีค่าความน่าจะเป็นของการตัดสินใจถูกต้องเท่ากับ .71 ความคลาดเคลื่อนในการจำแนกชนิดประเภทที่ 1 (ผู้ที่เรียนแล้วแต่สอบไม่ผ่าน) และประเภทที่ 2 (ผู้ที่ยังไม่เรียนแต่สอบผ่าน) เท่ากับ .23 และ .06 ตามลำดับ ค่าความเที่ยงตรงของคะแนนเกณฑ์ 4 คะแนน เท่ากับ .44 ดังนั้นนักเรียน

ที่จะสอบผ่านจุดประสงค์เชิงพฤติกรรม 1.1.3 ที่ว่า "เมื่อกำหนดจำนวนเลข นักเรียนสามารถบอกจำนวนครั้งที่เอาออกจนหมดและจำนวนเศษที่เหลือได้" จะต้องทำข้อสอบได้ถูกต้องอย่างน้อย 4 ข้อ ในจำนวนข้อสอบทั้งหมด 5 ข้อ

5.4 คะแนนเกณฑ์ของจุดประสงค์เชิงพฤติกรรม 1.2.1 ในแบบทดสอบฉบับที่ 1

ผู้วิจัยนำคะแนนผลการสอบของจุดประสงค์เชิงพฤติกรรม 1.2.1 จำนวน 7 ข้อ จากการสอบก่อนเรียนและการสอบหลังเรียน ในการทดสอบครั้งที่สาม กับกลุ่มตัวอย่างจำนวน 124 คน มาหาคะแนนเกณฑ์โดยวิธีของเบอร์ก ได้คะแนนพยากรณ์จากจุดตัดของการกระจายคะแนนในการสอบทั้งสองครั้ง ดังภาพประกอบ 7



ภาพประกอบ 7 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบก่อนเรียนและการสอบหลังเรียน ของจุดประสงค์เชิงพฤติกรรม 1.2.1 ในแบบทดสอบฉบับที่ 1

จากภาพประกอบ 7 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบทั้งสองครั้ง คือ 5 คะแนน แล้วหาค่าความน่าจะเป็นในการตัดสินที่ถูกต้อง ความคลาดเคลื่อนในการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด เพื่อหาคะแนนที่ใช้ตัดสินการสอบผ่านจุดประสงค์นี้ ดังตาราง 18

คะแนนจุดตัด	ความน่าจะเป็นของการตัดสินที่ถูกต้อง	ความคลาดเคลื่อนในการจำแนกผิดประเภทที่ 1 / ประเภทที่ 2	สัมประสิทธิ์ความเที่ยงตรง
7	.73	.25 / .01	.53
6	.81	.16 / .03	.64
5*	.82	.01 / .08	.77
4	.73	.08 / .20	.46
3	.63	.05 / .32	.39
2	.57	.03 / .40	.20
1	.52	.01 / .47	.10

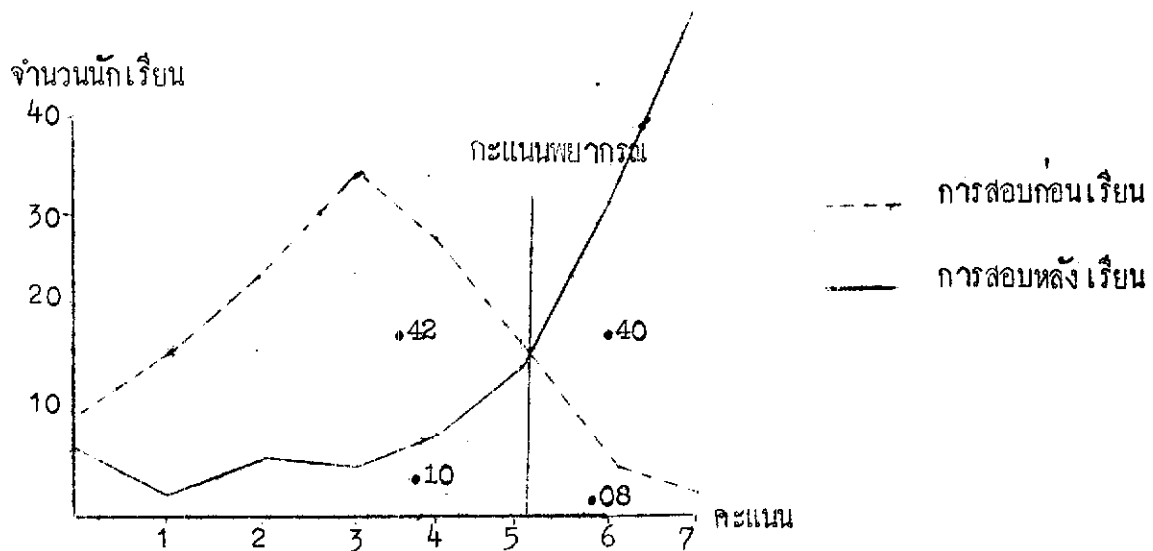
* คะแนนจุดตัดที่เหมาะสม

จากตาราง 18 ที่คะแนนจุดตัด 5 คะแนน ได้ค่าความน่าจะเป็นในการตัดสินที่ถูกต้องเป็น .82 ความคลาดเคลื่อนในการจำแนกผิดประเภทที่ 1 (ผู้ที่เรียนแล้วแต่สอบไม่ผ่านเกณฑ์) เท่ากับ .11 และ .08 ตามลำดับ ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัดเท่ากับ .77 ซึ่งมีค่าสูงสุด แสดงว่าคะแนนจุดตัดที่เหมาะสมหรือคะแนนเกณฑ์โดยวิธีของ เบอร์กคือที่ 5 คะแนน

และจากการพิจารณาของครูผู้สอนกำหนดคะแนนเกณฑ์ 5 คะแนนเช่นกัน ดังนั้นนักเรียนจะสอบผ่านจุดประสงค์เชิงพฤติกรรม 1.2.1 ที่ว่า "เมื่อกำหนดโจทย์ปัญหาหรือข้อให้จำนวนหนึ่ง และบอกจำนวนกลุ่ม นักเรียนสามารถหารจำนวนในแต่ละกลุ่มได้" นักเรียนจะต้องทำข้อสอบถูกต้องอย่างน้อย 5 ข้อ ในจำนวนข้อสอบทั้งหมด 7 ข้อ

5.5 คะแนนเกณฑ์ของจุดประสงค์เชิงพฤติกรรม 1.2.2 ในแบบทดสอบฉบับที่ 1

ผู้วิจัยได้นำคะแนนผลการสอบในจุดประสงค์เชิงพฤติกรรม 1.2.2 จำนวน 7 ข้อ จากการสอบก่อนเรียนซึ่งถือเป็นกลุ่มนักเรียนที่ยังไม่ได้รับการสอน และการสอบหลังเรียนซึ่งถือเป็นกลุ่มนักเรียนที่ได้รับการสอนแล้ว ในการทดสอบครั้งที่สามกับกลุ่มตัวอย่าง : 124 คน มาหาคะแนนเกณฑ์โดยใช้วิธีของ เบอร์ก ได้คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบทั้งสองครั้ง ดังภาพประกอบ 8



ภาพประกอบ 8 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายในการสอบก่อนเรียน และการสอบหลังเรียน ของจุดประสงค์เชิงพฤติกรรม 1.2.2 ในแบบทดสอบฉบับที่ 1

จากภาพประกอบ 8 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนน
ในการสอบทั้งสองครั้ง คือที่ 6 คะแนน

คะแนนแต่ละคะแนนนำมาหาค่าความน่าจะเป็นในการตัดสินที่ถูกต้อง ความคลาดเคลื่อน
ในการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด เพื่อหาคะแนนเกณฑ์ใน
จุดประสงค์นี้ ดังตาราง 19

ตาราง 19 ค่าความน่าจะเป็นในการตัดสินที่ถูกต้อง ความคลาดเคลื่อนในการจำแนกผิด
และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์เชิงพฤติกรรม 1.2.2
ในแบบทดสอบฉบับที่ 1

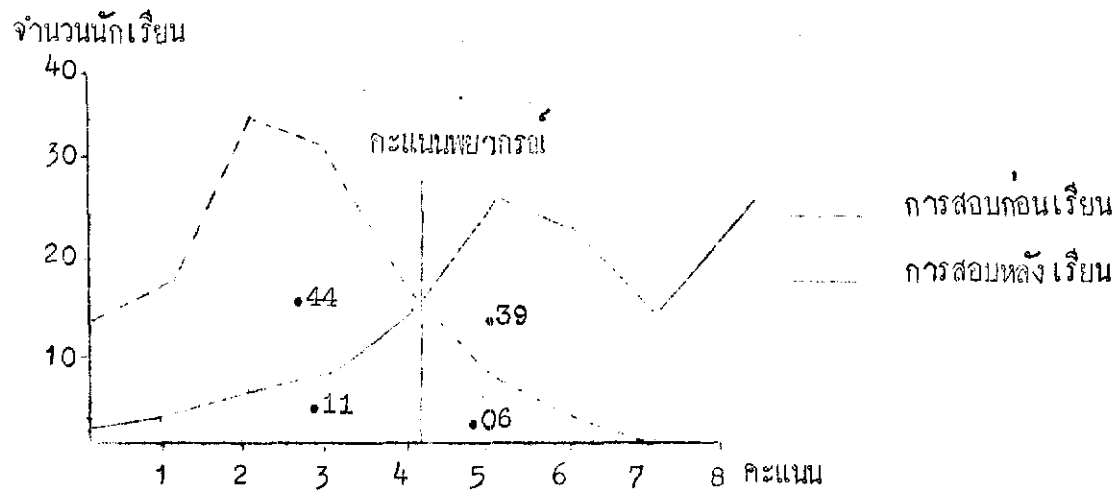
คะแนนจุดตัด	ความน่าจะเป็นของ การตัดสินที่ถูกต้อง	ความคลาดเคลื่อนในการจำแนกผิด ประเภทที่ 1 / ประเภทที่ 2	สัมประสิทธิ์ ความเที่ยงตรง
7	.72	.27 / .01	.44
6	.83	.15 / .02	.68
5*	.82	.10 / .03	.80
4	.74	.07 / .19	.50
3	.63	.05 / .32	.30
2	.56	.03 / .40	.24
1	.51	.02 / .46	.08

* คะแนนจุดตัดที่เหมาะสม

จากตาราง 19 ที่คะแนนจุดตัด 6 คะแนน ค่าความน่าจะเป็นของการตัดสินใจถูกต้องเท่ากับ .83 แต่ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัดที่ 6 คะแนนเท่ากับ .68 ในขณะที่คะแนนจุดตัด 5 คะแนน ค่าความน่าจะเป็นในการตัดสินใจถูกต้องเท่ากับ .82 ซึ่งมีค่าใกล้เคียงกับค่าความน่าจะเป็นในการตัดสินใจถูกต้องที่คะแนนจุดตัด 6 คะแนนมาก แต่ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัดที่ 5 คะแนนเท่ากับ .80 ซึ่งแสดงว่าที่คะแนนจุดตัด 5 คะแนนเป็นคะแนนที่สมเหตุสมผลมากกว่า และในภาพประกอบ 8 จุดตัดของกราฟทั้งสองก็อยู่ใกล้เคียงกับที่คะแนน 5 คะแนน ดังนั้นจึงเลือกเอาคะแนนจุดตัดที่เหมาะสมที่คะแนน 5 คะแนน และจากการพิจารณาของครูผู้สอนกำหนดคะแนนเกณฑ์เป็น 5 คะแนน ฉะนั้นนักเรียนจะสอบผ่านจุดประสงค์เชิงพฤติกรรม 1.2.2 ที่ว่า "เมื่อกำหนดจำนวนสิ่งของที่เท่ากันของแต่ละกลุ่มและของทั้งหมด นักเรียนสามารถหาจำนวนกลุ่มได้" ซึ่งนักเรียนจะต้องทำข้อสอบได้ถูกต้องอย่างน้อย 5 ข้อ ในจำนวนข้อสอบทั้งหมด 7 ข้อ

6. คะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรมในแบบทดสอบฉบับที่ 2 จาก การทดสอบครั้งที่สาม

6.1 คะแนนเกณฑ์ของจุดประสงค์เชิงพฤติกรรม 2.1 ในแบบทดสอบฉบับที่ 2
ผู้วิจัยได้นำคะแนนผลการสอบในจุดประสงค์เชิงพฤติกรรม 2.1 จำนวน 8 ข้อ จากการสอบก่อนเรียนซึ่งถือเป็นกลุ่มนักเรียนที่ยังไม่ได้รับการสอน และการสอบหลังเรียน ซึ่งเป็นกลุ่มนักเรียนที่ได้รับการสอนแล้ว ในการทดสอบครั้งที่สามกับกลุ่มตัวอย่าง จำนวน 124 คน มาหาคะแนนเกณฑ์โดยใช้วิธีของ เบอร์ก ได้คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายในการสอบทั้งสองครั้ง ดังภาพประกอบ 9



ภาพประกอบ 9 คะแนนพวกรวมที่ได้จากจุดตัดของการกระจายคะแนนในการสอบก่อนเรียน และการสอบหลังเรียน ของจุดประสงค์เชิงพฤติกรรม 2.1 ในแบบทดสอบ ฉบับที่ 2

จากภาพประกอบ 9 คะแนนพวกรวมที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ ทั้งสองครั้ง คือที่ 5 คะแนน

คะแนนแต่ละคะแนนนำมาหาค่าความน่าจะเป็นในการตัดสินที่ถูกต้อง ความคลาดเคลื่อน ในการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด เพื่อหาคะแนนเกณฑ์ใน จุดประสงค์นี้ ถึงการฯ 20

ตาราง 20 ค่าความน่าจะเป็นของการตัดสินใจที่ถูกต้อง ความคลาดเคลื่อนในการจำแนกชนิด
และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์เชิงพฤติกรรม 2.1
ในแบบทดสอบฉบับที่ 2

คะแนนจุดตัด	ความน่าจะเป็นของ การตัดสินใจที่ถูกต้อง	ความคลาดเคลื่อนในการจำแนกชนิด ประเภทที่ 1 / ประเภทที่ 2	สัมประสิทธิ์ ความเที่ยงตรง
8	.61	.40 / .00	.34
7	.67	.33 / .00	.46
6	.76	.23 / .01	.60
5 *	.83	.11 / .06	.67
4	.82	.07 / .11	.65
3	.73	.04 / .24	.38
2	.61	.01 / .38	.33
1	.55	.00 / .44	.21

* คะแนนจุดตัดที่เหมาะสม

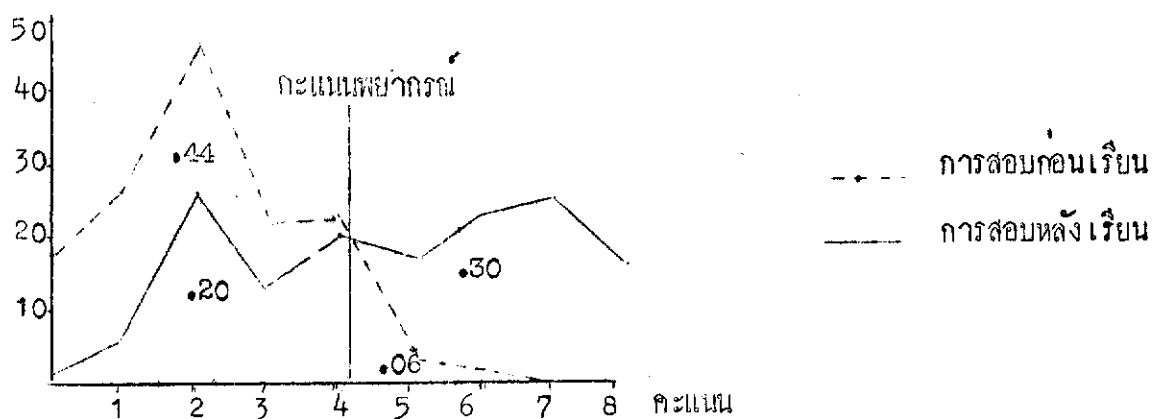
จากตาราง 20 ที่คะแนนจุดตัด 5 คะแนน ค่าความน่าจะเป็นของการตัดสินใจที่ถูกต้อง
เท่ากับ .83 ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัดเท่ากับ .67 แสดงว่าคะแนนจุดตัด
ที่เหมาะสมหรือคะแนนเกณฑ์โดยวิธีของ เบอร์กคือที่ 5 คะแนน ส่วนการพิจารณาของครูผู้สอน
กำหนดคะแนนเกณฑ์เท่ากับ 6 คะแนน เมื่อนำเฉลี่ยจะได้คะแนนเกณฑ์ที่ 6 คะแนน ซึ่งจาก

ตาราง 20 ค่าความน่าจะเป็นในการตัดสินใจถูกต้องเท่ากับ .76 ความคลาดเคลื่อนในการจำแนกนิพประเภทที่ 1 (ผู้ที่เรียนแล้วแต่สอบไม่ผ่านเกณฑ์) และประเภทที่ 2 (ผู้ที่ยังไม่เรียนแต่สอบผ่านเกณฑ์) เท่ากับ .23 และ .01 ตามลำดับ ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัดที่ 6 คะแนนเท่ากับ .60 ดังนั้นนักเรียนที่จะสอบผ่านจุดประสงค์เชิงพฤติกรรม 2.1 ที่ว่า "เมื่อกำหนดโจทย์หรือภาพให้ นักเรียนสามารถเขียนอยู่ในรูปประโยคสัญลักษณ์ได้" จะต้องทำคะแนนได้ถูกต้องอย่างน้อย 6 ข้อ ในจำนวนข้อสอบทั้งหมด 8 ข้อ

6.2 คะแนนเกณฑ์ของจุดประสงค์เชิงพฤติกรรม 2.2 ในแบบทดสอบฉบับที่ 2

ผู้วิจัยได้นำคะแนนผลการสอบในจุดประสงค์เชิงพฤติกรรม 2.2 จำนวน 8 ข้อ จากการสอบก่อนเรียนซึ่งเป็นกลุ่มนักเรียนที่ยังไม่ได้รับการสอน และการสอบหลังเรียนซึ่งเป็นกลุ่มนักเรียนที่ได้รับการสอนแล้ว ในการทดสอบครั้งที่สามกับกลุ่มตัวอย่าง จำนวน 124 คน มาหาคะแนนเกณฑ์โดยใช้วิธีของ เบอร์ก ได้คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายในการสอบทั้งสองครั้ง ทั้งภาพประกอบที่ 10

จำนวนนักเรียน



ภาพประกอบ 10 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบก่อนเรียน และการสอบหลังเรียน ของจุดประสงค์เชิงพฤติกรรม 2.2 ในแบบทดสอบฉบับที่ 2

จากภาพประกอบ 10 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ
ทั้งสองครั้ง คือที่ 5 คะแนน

คะแนนแต่ละคะแนนนำมาหาค่าความน่าจะเป็นในการตัดสินที่ถูกต้อง ความคลาดเคลื่อน
ในการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด เพื่อหาคะแนนเกณฑ์ใน
จุดประสงค์นี้ พิจารณาราย 21

ตาราง 21 ค่าความน่าจะเป็นของการตัดสินที่ถูกต้อง ความคลาดเคลื่อนในการจำแนกผิด
และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์เชิงพฤติกรรม 2.2
ในแบบทดสอบฉบับที่ 2

คะแนนจุดตัด	ความน่าจะเป็นของ การตัดสินที่ถูกต้อง	ความคลาดเคลื่อนในการจำแนกผิด ประเภทที่ 1 / ประเภทที่ 2	สัมประสิทธิ์ ความเที่ยงตรง
8	.54	.46 / .00	.20
7	.59	.41 / .00	.30
6	.69	.31 / .04	.47
5	.78	.20 / .02	.62
4*	.78	.14 / .08	.72
3	.75	.07 / .17	.52
2	.61	.03 / .36	.19
1	.53	.01 / .47	.15

* คะแนนจุดตัดที่เหมาะสม

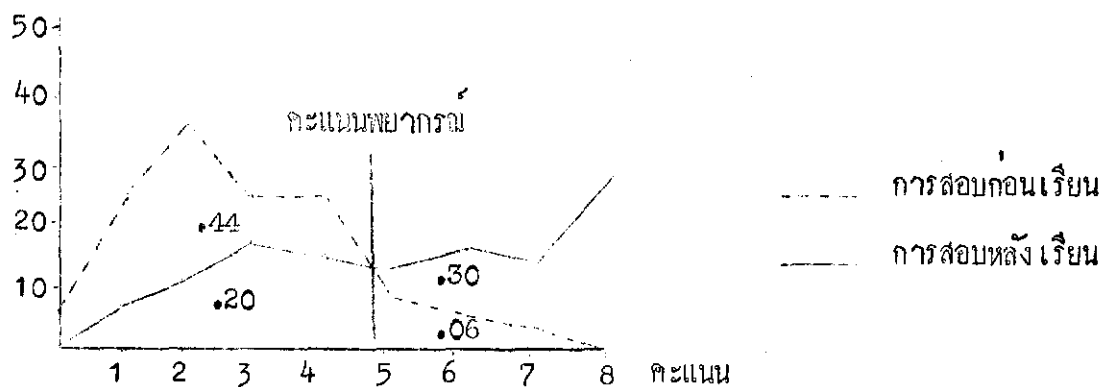
จากตาราง 21 คะแนนจุดตัด 4 คะแนน และ 5 คะแนน มีค่าความน่าจะเป็นในการตัดสินใจถูกต้อง .78 เท่ากัน แต่ที่คะแนนจุดตัด 4 คะแนน มีสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัดเท่ากับ .72 ซึ่งมีความมากกว่าค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัดที่ 5 คะแนน ที่มีค่าเท่ากับ .62 จึงเลือกใช้คะแนนจุดตัดที่เหมาะสมที่ 4 คะแนน ส่วนการพิจารณาของครูผู้สอนกำหนดคะแนนเกณฑ์เป็น 6 คะแนน เมื่อนำมาเฉลี่ยได้คะแนนเกณฑ์เป็น 5 คะแนน จากตาราง 21 ที่คะแนนจุดตัด 5 คะแนน ค่าความน่าจะเป็นในการตัดสินใจถูกต้อง เป็น .78 ความคลาดเคลื่อนในการจำแนกชนิดประเภทที่ 1 (ผู้เรียนแล้วแต่สอบไม่ผ่านเกณฑ์) และประเภทที่ 2 (ผู้ยังไม่เรียนแต่สอบผ่านเกณฑ์) เท่ากับ .20 และ .02 ตามลำดับ ความเที่ยงตรงที่คะแนนเกณฑ์ 5 คะแนน เท่ากับ .72 ดังนั้นนักเรียนจะสอบผ่านจุดประสงค์เชิงพฤติกรรม 2.2 ที่ว่า "เมื่อกำหนดโจทย์การหาร นักเรียนสามารถใช้สัมพันธ์ระหว่างการคูณและการหารในการหาตัวตั้ง ตัวหาร และผลหารได้" นักเรียนจะต้องทำข้อสอบได้ถูกต้องอย่างน้อย 5 ข้อ ในจำนวนข้อสอบทั้งหมด 8 ข้อ

7. คะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรม ในแบบทดสอบฉบับที่ 3

ในการทดสอบครั้งที่สาม

7.1 คะแนนเกณฑ์ในจุดประสงค์เชิงพฤติกรรม 3.1 ในแบบทดสอบฉบับที่ 3 ผู้วิจัยได้นำคะแนนผลการสอบในจุดประสงค์เชิงพฤติกรรม 3.1 จำนวน 8 ข้อ จากการสอบก่อนเรียนซึ่งเป็นกลุ่มที่ยังไม่ได้รับการสอน และการสอบหลังเรียนซึ่งเป็นกลุ่มที่ได้รับการสอนแล้ว ในการทดสอบครั้งที่สามกับกลุ่มตัวอย่างจำนวน 124 คน มาหาคะแนนเกณฑ์ โดยใช้วิธีของ เบอร์ก ได้คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบทั้งสองครั้ง ดังภาพประกอบ 11

จำนวนนักเรียน



ภาพประกอบ 11 คะแนนพยากรณ์ที่ได้จากจุดตัดการกระจายคะแนนในการสอบก่อนเรียน และการสอบหลังเรียน ของจุดประสงค์เชิงพฤติกรรม 3.1 ในแบบทดสอบฉบับที่ 3

จากภาพประกอบ 11 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบทั้งสองครั้ง คือที่ 5 คะแนน

คะแนนแต่ละคะแนนนำมาหาค่าความน่าจะเป็นในการตัดสินที่ถูกต้อง ความคลาดเคลื่อนในการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด เพื่อหาคะแนนเกณฑ์ในจุดประสงค์นี้ ดังตาราง 22

ตาราง 22 ค่าความน่าจะเป็นในการตัดสินใจถูกต้อง ความคลาดเคลื่อนในการจำแนกชนิด
และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์เชิงพฤติกรรม 3.1
ในแบบทดสอบฉบับที่ 3

คะแนนจุดตัด	ความน่าจะเป็นของ การตัดสินใจถูกต้อง	ความคลาดเคลื่อนในการจำแนกชนิด ประเภทที่ 1 / ประเภทที่ 2	สัมประสิทธิ์ ความเที่ยงตรง
8	.61	.39 / .00	.36
7	.66	.32 / .01	.42
6	.71	.26 / .03	.48
5 *	.74	.20 / .06	.71
4	.71	.14 / .15	.43
3	.68	.07 / .24	.38
2	.58	.02 / .39	.27
1	.52	.00 / .48	.13

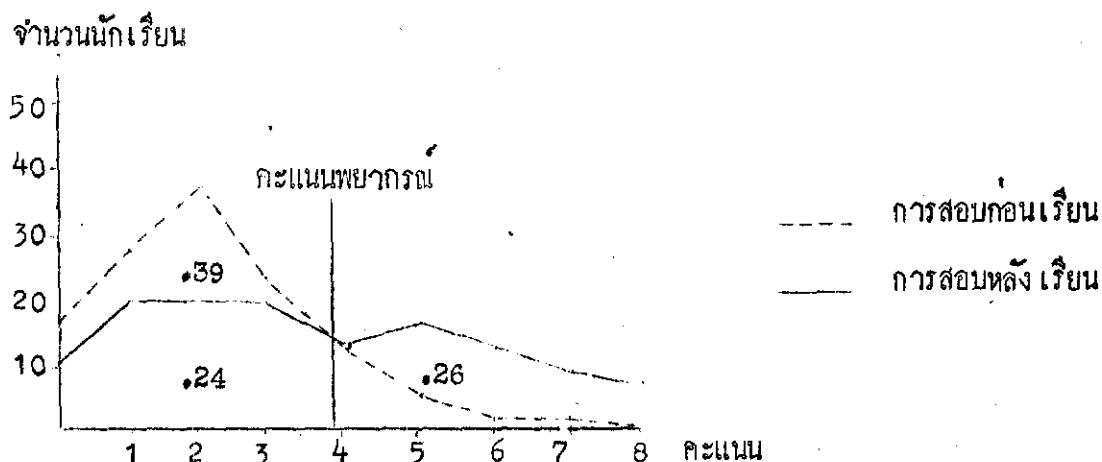
* คะแนนจุดตัดที่เหมาะสม

จากตาราง 22 ที่คะแนนจุดตัด 5 คะแนน ค่าความน่าจะเป็นในการตัดสินใจถูกต้อง
เป็น .74 ความคลาดเคลื่อนในการจำแนกชนิดประเภทที่ 1 (ผู้ที่เรียนแล้วแต่สอบไม่ผ่านเกณฑ์)
และประเภทที่ 2 (ผู้ที่ยังไม่เรียนแต่สอบผ่านเกณฑ์) เท่ากับ .20 และ .06 ตามลำดับ
ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัดที่ 5 คะแนนเท่ากับ .71 แสดงว่าคะแนนจุดตัด

ที่เหมาะสมหรือคะแนนเกณฑ์โดยวิธีของ เบอร์กคือที่ 5 คะแนน แต่จากการพิจารณาของครู ผู้สอนกำหนดคะแนนเกณฑ์ที่ 5 คะแนน เมื่อนำมาเฉลี่ยจะได้คะแนนเกณฑ์ที่ 5 คะแนน ดังนั้น นักเรียนจะสอบผ่านจุดประสงค์เชิงพฤติกรรม 3.1 ที่ว่า "เมื่อกำหนดโภชนาการที่มีตัวหารหลักเดียว ผลหารไม่เกิน 10 นักเรียนสามารถทำโจทย์เกี่ยวกับการหารลงตัวได้" นักเรียนจะต้องทำข้อสอบได้ถูกต้องอย่างน้อย 5 ข้อ ในจำนวนข้อสอบทั้งหมด 8 ข้อ

7.2 คะแนนเกณฑ์ของจุดประสงค์เชิงพฤติกรรม 3.2 ในแบบทดสอบฉบับที่ 3

ผู้วิจัยได้นำคะแนนผลการสอบในจุดประสงค์เชิงพฤติกรรม 3.2 จำนวน 8 ข้อ จากการสอบก่อนเรียนซึ่งเป็นกลุ่มนักเรียนที่ยังไม่ได้รับการสอน และการสอบหลังเรียนซึ่งเป็นกลุ่มนักเรียนที่ได้รับการสอนแล้ว ในการทดสอบครั้งที่สามกับกลุ่มตัวอย่าง 124 คน มาหาคะแนนเกณฑ์โดยวิธีของ เบอร์ก ได้คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบทั้งสองครั้ง ดังภาพประกอบ 12



ภาพประกอบ 12 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบก่อนเรียน และสอบหลังเรียน ของจุดประสงค์เชิงพฤติกรรม 3.2 ในแบบทดสอบฉบับที่ 3

จากภาพประกอบ 12 คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนน
ในการสอบทั้งสองครั้ง คือ 4 คะแนน

คะแนนแต่ละคะแนนนำมาหาค่าความน่าจะเป็นในการตัดสินที่ถูกต้อง ความคลาดเคลื่อน
ในการจำแนก และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัดเพื่อหาคะแนนเกณฑ์ใน
จุดประสงค์นี้ ดังตาราง 23

ตาราง 23 ค่าความน่าจะเป็นในการตัดสินที่ถูกต้อง ความคลาดเคลื่อนในการจำแนก
และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด

คะแนนจุดตัด	ความน่าจะเป็นของ การตัดสินที่ถูกต้อง	ความคลาดเคลื่อนในการจำแนก ประเภทที่ 1 / ประเภทที่ 2	สัมประสิทธิ์ ความเที่ยงตรง
8	.52	.49 / .00	.13
7	.55	.45 / .04	.21
6	.59	.40 / .01	.29
5	.64	.33 / .03	.35
4*	.65	.24 / .11	.39
3	.59	.20 / .21	.26
2	.52	.12 / .36	.14
1	.49	.04 / .47	.10

* คะแนนจุดตัดที่เหมาะสม

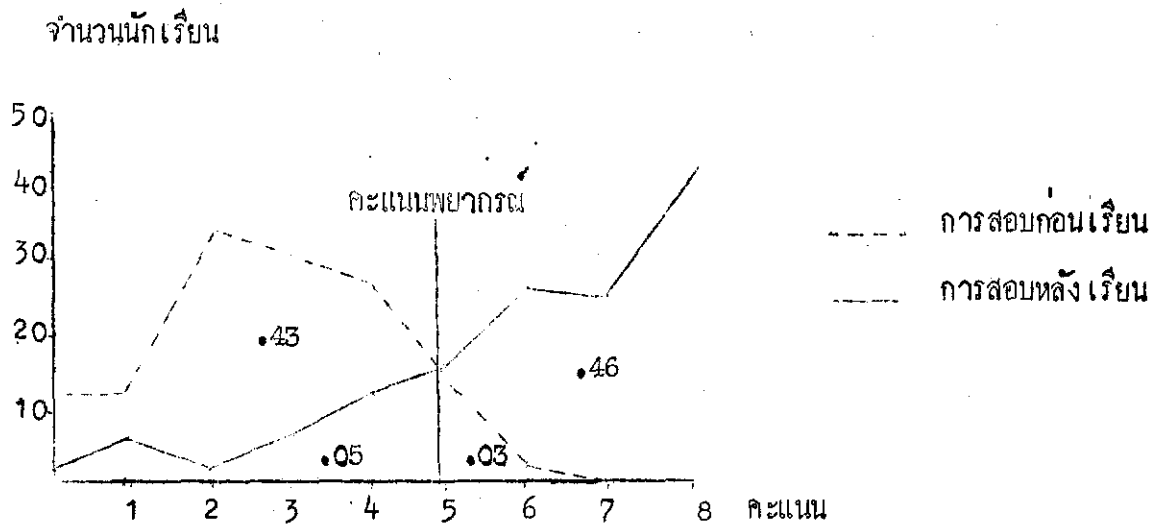
จากตาราง 23 ที่คะแนนจุดตัดที่ 4 คะแนน ค่าความน่าจะเป็นของการตัดสินใจถูกต้องเท่ากับ .65 ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัดที่เท่ากับ .39 ซึ่งมีค่าสูงสุด แสดงว่าคะแนนจุดตัดที่เหมาะสมหรือคะแนนเกณฑ์โดยวิธีของ เบอร์กคือที่ 4 คะแนน ส่วนการพิจารณาของผู้สอนกำหนดคะแนนเกณฑ์เป็น 6 คะแนน เมื่อนำมาเฉลี่ยจะได้คะแนนเกณฑ์ที่ 5 คะแนน จากตาราง 23 ค่าความน่าจะเป็นของการตัดสินใจถูกต้องเท่ากับ .64 ความคลาดเคลื่อนในการจำแนกนิพประเภที่ 1 (ผู้ที่เรียนแล้วแต่สอบไม่ผ่านเกณฑ์) และประเภที่ 2 (ผู้ที่ยังไม่เรียนแต่สอบผ่านเกณฑ์) มีค่าเท่ากับ .33 และ .03 ตามลำดับ ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์ที่ 5 คะแนนเท่ากับ .35 ดังนั้นนักเรียนจะสอบผ่านจุดประสงค์เชิงพฤติกรรม 3.2 ที่ว่า "เมื่อกำหนดโจทย์การหารให้ นักเรียนสามารถทำโจทย์เกี่ยวกับการหารมีเศษได้" จะต้องทำข้อสอบได้ถูกต้องอย่างน้อย 5 ข้อ ในจำนวนข้อสอบทั้งหมด 8 ข้อ

8. คะแนนเกณฑ์ในแต่ละจุดประสงค์เชิงพฤติกรรม ในแบบทดสอบฉบับที่ 4

ในการทดสอบครั้งที่สาม

8.1 คะแนนเกณฑ์ในจุดประสงค์เชิงพฤติกรรม 4.1 ในแบบทดสอบฉบับที่ 4

ผู้วิจัยได้นำคะแนนผลการสอบในจุดประสงค์เชิงพฤติกรรม 4.1 จำนวน 8 ข้อ จากการสอบก่อนเรียนซึ่งเป็นกลุ่มนักเรียนที่ยังไม่ได้รับการสอน และการสอบหลังเรียนซึ่งเป็นกลุ่มนักเรียนที่ได้รับการสอนแล้ว ในการทดสอบครั้งที่สามกับกลุ่มตัวอย่าง 124 คน มาหาคะแนนเกณฑ์โดยวิธีของ เบอร์ก ได้คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนนในการสอบทั้งสองครั้ง ดังภาพประกอบ 13



ภาพประกอบ 13 คะแนนพหุการณที่ได้จากจุดตัดของการกระจายของการสอบก่อนเรียน และการสอบหลังเรียน ของจุดประสงค์เชิงพฤติกรรม 4.1 ใน แบบทดสอบฉบับที่ 4

จากภาพประกอบ 13 คะแนนพหุการณที่ได้จากจุดตัดของการกระจายคะแนนในการสอบ ทั้งสองครั้ง คือ 5 คะแนน

คะแนนแต่ละคะแนนนำมาหาค่าความน่าจะเป็นในการตัดสินที่ถูกต้อง ความคลาดเคลื่อน ในการจำแนก และสัมพันธ์ความเที่ยงตรงของคะแนนจุดตัด เพื่อหาคะแนนเกณฑ์ใน จุดประสงค์นี้ ดังตาราง 24

ตาราง 24 ค่าความน่าจะเป็นในการตัดสินใจที่ถูกต้อง ความคลาดเคลื่อนในการจำแนกชนิด
และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์เชิงพฤติกรรม 4.1
ในแบบทดสอบฉบับที่ 4

คะแนนจุดตัด	ความน่าจะเป็นของ การตัดสินใจที่ถูกต้อง	ความคลาดเคลื่อนในการจำแนกชนิด ประเภทที่ 1 / ประเภทที่ 2	สัมประสิทธิ์ ความเที่ยงตรง
8	.71	.27 / .00	.52
7	.80	.20 / .00	.54
6	.88	.11 / .01	.78
5*	.89	.05 / .07	.79
4	.79	.04 / .17	.60
3	.69	.03 / .29	.44
2	.57	.02 / .41	.22
1	.54	.01 / .49	.19

* คะแนนจุดตัดที่เหมาะสม

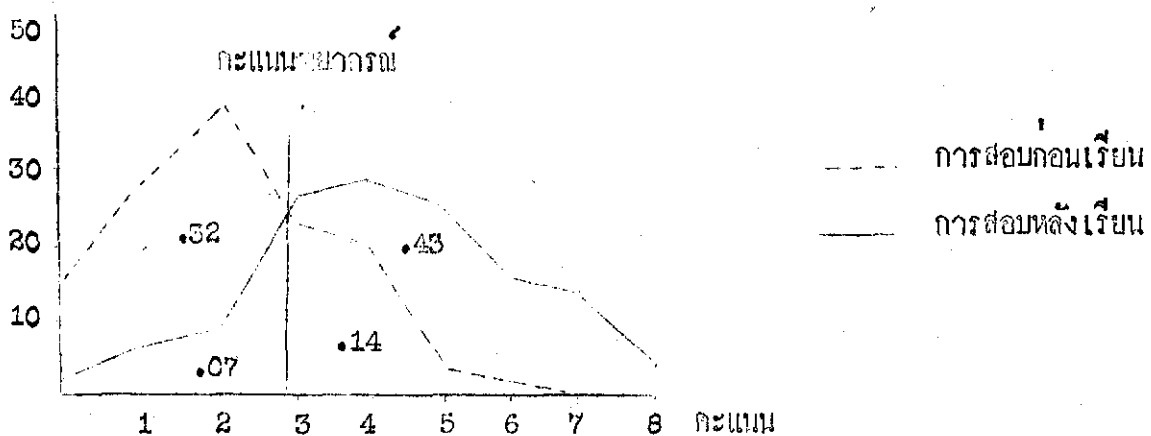
จากตาราง 24 ที่คะแนนจุดตัด 5 คะแนน ค่าความน่าจะเป็นในการตัดสินใจที่ถูกต้อง
เป็น .89 ความคลาดเคลื่อนในการจำแนกชนิดประเภทที่ 1 (ผู้ที่เรียนแล้วแต่สอบไม่ผ่านเกณฑ์)
และประเภทที่ 2 (ผู้ที่ยังไม่เรียนแต่สอบผ่านเกณฑ์) เท่ากับ .05 และ .07 ตามลำดับ
ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัดที่ 5 คะแนนมีค่าเท่ากับ .79 ซึ่งมีค่าสูงสุด
แสดงว่าคะแนนจุดตัดที่เหมาะสมหรือคะแนนเกณฑ์โดยวิธีของเบอร์ก์ คือที่คะแนน 5 คะแนน

แต่จากการพิจารณาของครูผู้สอนกำหนดคะแนนเกณฑ์เป็น 6 คะแนน เมื่อเข้ามาเฉลี่ยจะได้
คะแนนเกณฑ์เป็น 5 คะแนน ดังนั้นนักเรียนจะสอบผ่านจุดประสงค์เชิงพฤติกรรม 4.1 ทว่า
"เมื่อกำหนดโจทย์ปัญหาการหารที่มีตัวหารหลักเดียว และผลหารไม่เกิน 10 นักเรียนสามารถทำโจทย์
หารลงตัวได้" นั้น นักเรียนจะต้องทำข้อสอบที่ถูกต้องอย่างน้อย 6 ข้อ ในจำนวนข้อสอบ
ทั้งหมด 8 ข้อ

8.2 คะแนนในจุดประสงค์เชิงพฤติกรรม 4.2 ในแบบทดสอบฉบับที่ 4

ผู้วิจัยได้นำคะแนนผลการสอบในจุดประสงค์เชิงพฤติกรรม 4.2 จำนวน
8 ข้อ จากการสอบก่อนเรียนซึ่งเป็นกลุ่มนักเรียนที่ยังไม่ได้รับการสอน และการสอบหลังเรียน
ซึ่งเป็นกลุ่มนักเรียนที่ได้รับการสอนแล้ว ในการทดสอบครั้งที่สามกับกลุ่มตัวอย่างจำนวน 124 คน
มาหาคะแนนเกณฑ์โดยวิธีของเบอร์ก ได้คะแนนพยากรณ์ที่ได้จากจุดตัดของการกระจายคะแนน
ในการสอบทั้งสองครั้ง ดังภาพประกอบ 14

จำนวนนักเรียน



ภาพประกอบ 14 คะแนนพยากรณ์ได้จากจุดตัดของการกระจายคะแนนในการสอบก่อนเรียน
และการสอบหลังเรียน ของจุดประสงค์เชิงพฤติกรรม 4.2 ในแบบทดสอบ
ฉบับที่ 4

จากภาพประกอบ 14 คะแนนพบว่ากรณีใดจากจุดตัดของการกระจายคะแนน
ในการสอบทั้งสองครั้ง คือที่ 3 คะแนน

คะแนนแต่ละคะแนนนำมาหาค่าความน่าจะเป็นในการตัดสินใจถูกต้อง ความคลาดเคลื่อน
ในการจำแนกผิด และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด เพื่อหาคะแนนเกณฑ์ใน
จุดประสงค์นี้ ดังตาราง 25

ตาราง 25 ค่าความน่าจะเป็นในการตัดสินใจถูกต้อง ความคลาดเคลื่อนในการจำแนกผิด
และสัมประสิทธิ์ความเที่ยงตรงของคะแนนจุดตัด ของจุดประสงค์เชิงพฤติกรรม 4.2
ในแบบทดสอบฉบับที่ 4

คะแนนจุดตัด	ความน่าจะเป็นของ การตัดสินใจถูกต้อง	ความคลาดเคลื่อนในการจำแนกผิด ประเภทที่ 1 , ประเภทที่ 2	สัมประสิทธิ์ ความเที่ยงตรง
8	.51	.49 / .00	.12
7	.56	.44 / .00	.25
6	.61	.39 / .01	.33
5	.69	.29 / .02	.45
4	.73	.17 / .10	.47
3 *	.75	.07 / .19	.64
2	.63	.03 / .34	.32
1	.54	.01 / .45	.19

* คะแนนจุดตัดที่เหมาะสม

จากตาราง 25 ที่คะแนนถูกต้อง 3 คะแนน ถ้าความน่าจะเป็นในการตัดสินที่ถูกต้องเท่ากับ .75 ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนถูกต้องเท่ากับ .64 ซึ่งเป็นค่าสูงสุด แสดงว่าคะแนนถูกต้องที่เฉพาะสมหรือคะแนนเกณฑ์โดยวิธีของเบอร์กัสคือ 3 คะแนน จากการพิจารณาของครูผู้สอนกำหนดคะแนนเกณฑ์เป็น 5 คะแนน เมื่อนำมาเฉลี่ยจะได้คะแนนเกณฑ์ 4 คะแนน และจากตาราง 25 ถ้าความน่าจะเป็นในการตัดสินที่ถูกต้องเท่ากับ .73 ความคลาดเคลื่อนของการจำแนกประเภทที่ 1 (ผู้ที่เรียนแล้วแต่สอบไม่ผ่านเกณฑ์) และประเภทที่ 2 (ผู้ที่ยังไม่เรียนแต่สอบผ่านเกณฑ์) เท่ากับ .17 และ .10 ตามลำดับ ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนถูกต้อง 4 คะแนน เท่ากับ .47 ดังนั้นนักเรียนจะสอบผ่านจุดประสงค์เชิงพฤติกรรม 4.2 ที่ว่า "เมื่อกำหนดโจทย์ปัญหาการหารที่มีตัวหารหลักเดียวและผลหารไม่เกิน 10 นักเรียนสามารถหาคำตอบอย่างมีเงื่อนไข" จะคงทำข้อสอบใดถูกต้องอย่างน้อย 4 ข้อ ในจำนวนข้อสอบทั้งหมด 8 ข้อ

9. ค่าความเชื่อมั่นของแบบทดสอบแต่ละฉบับ จากการทดสอบครั้งที่สาม

ผู้วิจัยนำคะแนนจากการทดสอบครั้งที่สาม มาคำนวณหาค่าความเชื่อมั่นของแบบทดสอบอิง เกณฑ์ตัดสิน โดยใช้สูตรของซิปโคเวียค ดังแสดงผลในตาราง 26

ตาราง 26 ค่าความเชื่อมั่นของแบบทดสอบอิง เกณฑ์ตัดสิน จากการทดสอบครั้งที่สาม

แบบทดสอบฉบับที่	จำนวน ข้อสอบ	คะแนน เกณฑ์	ค่าเฉลี่ย	ค่าเบี่ยงเบน มาตรฐาน	ความเชื่อมั่น
1. ความหมายการหาร	29	22	22.2743	7.3593	.6354
2. สัญลักษณ์การหาร	16	11	10.0968	2.2874	.6147
3. โจทย์การหาร	16	10	10.1774	4.3069	.7835
4. โจทย์ปัญหาการหาร	16	10	10.5726	2.8783	.7553

จากตาราง 26 แสดงให้เห็นว่าค่าความเชื่อมั่นเชิงแบบทดสอบอิง เกณฑ์ตัดสิน ที่คำนวณโดยใช้สูตรของซิปโคเวียคมีค่าตั้งแต่ .6147 ถึง .7835 แบบทดสอบทั้งค่าความเชื่อมั่นสูงสุดคือ แบบทดสอบฉบับที่ 3 (โจทย์การหาร) แบบทดสอบทั้งค่าความเชื่อมั่นต่ำสุดคือ แบบทดสอบฉบับที่ 2 (สัญลักษณ์การหาร) ซึ่งทุกฉบับมีค่าสูงกว่า .60 ทั้งสิ้น

10. ความเที่ยงตรงของแบบทดสอบแต่ละฉบับ จากการทดสอบครั้งที่สาม

ผู้วิจัยนำคะแนนจากการทดสอบครั้งที่สาม มาหาความเที่ยงตรงของแบบทดสอบ
ดังนี้

10.1 ความเที่ยงตรงตามเนื้อหา ซึ่งหาโดยวิธีการสองวิธีคือ

10.1.1 ความเที่ยงตรงตามเนื้อหา ซึ่งตรวจสอบโดยนำแบบทดสอบ
ถึงเกณฑ์ทั้งสี่ฉบับ และสุกประสงค์เชิงพฤติกรรมรายข้อไปให้ผู้เชี่ยวชาญทางเนื้อหา เป็นผู้
พิจารณาว่าการกำหนดสุกประสงค์เชิงพฤติกรรมได้แน่ชัดชัดเจนหรือไม่ ข้อสอบแต่ละข้อวัดการ
ตามสุกประสงค์เชิงพฤติกรรมที่ต้องการสอบหรือไม่ โดยให้พิจารณาจับคู่ข้อสอบรายข้อกับ
สุกประสงค์เชิงพฤติกรรม ปรากฏว่าสุกประสงค์เชิงพฤติกรรมทุกข้อเขียนไว้ชัดเจน ข้อสอบ
แต่ละข้อวัดได้ตามสุกประสงค์เชิงพฤติกรรมที่ต้องการทดสอบ แสดงว่าแบบทดสอบอิงเกณฑ์
ทั้งสี่ฉบับมีความเที่ยงตรงตามเนื้อหา

10.1.2 ความเที่ยงตรงตามเนื้อหา โดยคำนวณจากสูตรของ
การวัด รัง แสดงผลในตาราง 27

ตาราง 27 ค่าความเที่ยงตรงตามเนื้อหาของแบบทดสอบอิงเกณฑ์ทั้งสี่ฉบับ จากการทดสอบ
ครั้งที่สาม

แบบทดสอบฉบับที่	จำนวนข้อสอบ	คะแนนเกณฑ์	กลุ่มตัวอย่าง	ความเที่ยงตรง
1. ความหมายการหาร	29	22	124	.8226
2. สัญลักษณ์การหาร	16	11	124	.7177
3. โจทย์การหาร	16	10	124	.8427
4. โจทย์ปัญหาการหาร	16	10	124	.7460

จากตาราง 27 แสดงให้เห็นว่าค่าความเที่ยงตรงตามเนื้อหาของแบบทดสอบ
 อิงเกณฑ์สี่อันดับ ที่คำนวณโดยใช้สูตรของคาร์เวอรมีค่าตั้งแต่ .7177 ถึง .8427
 แบบทดสอบที่มีค่าความเที่ยงตรงตามเนื้อหาสูงสุดคือ แบบทดสอบฉบับที่ 3 (โจทย์การหาร)
 แบบทดสอบที่มีค่าความเที่ยงตรงตามเนื้อหาต่ำสุดคือ แบบทดสอบฉบับที่ 2 (สัญลักษณ์การหาร)
 ซึ่งแบบทดสอบทุกฉบับมีค่าความเที่ยงตรงตามเนื้อหาสูงกว่า .70 ทั้งสิ้น แสดงว่าแบบทดสอบ
 อิงเกณฑ์สี่อันดับมีค่าความเที่ยงตรงตามเนื้อหาอยู่ในเกณฑ์สูง

สรุป อภิปรายผล และข้อเสนอแนะ

ความมุ่งหมายในการศึกษาค้นคว้า

1. เพื่อสร้างแบบทดสอบอิงเกณฑ์ในวิชาคณิตศาสตร์ เรื่องการหาร สำหรับนักเรียนชั้นประถมศึกษาปีที่ 3

2. เพื่อหาคะแนนเกณฑ์ตัดสินการสอบผ่านในจุดประสงค์เชิงพฤติกรรมของแบบทดสอบที่สร้างขึ้น

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้า เป็นนักเรียนชั้นประถมศึกษาปีที่ 2 ปีการศึกษา 2523 ของโรงเรียนในสังกัดสำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ ในจังหวัดสมุทรสงคราม ซึ่งเลือกมาโดยการสุ่มแบบแบ่งชั้น

เครื่องมือที่ใช้ในการรวบรวมข้อมูล

เครื่องมือที่ใช้ในการศึกษาค้นคว้าคือ แบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่องการหาร ซึ่งประกอบด้วยแบบทดสอบย่อยสี่ฉบับ ดังนี้

แบบทดสอบฉบับที่ 1 ความหมายการหาร

แบบทดสอบฉบับที่ 2 สัญลักษณ์การหาร

แบบทดสอบฉบับที่ 3 โจทย์การหาร

แบบทดสอบฉบับที่ 4 โจทย์ปัญหาการหาร

★ วิธีดำเนินการรวมข้อสอบ

1. จัดข้อโรงเรียนที่ใช้เป็นกลุ่มตัวอย่าง เพื่อกำหนดวันและเวลาในการสอบ
2. เตรียมข้อสอบให้เพียงพอกับจำนวนนักเรียนที่สอบในแต่ละครั้ง และวางแผนในการดำเนินการสอบล่วงหน้า เช่น จัดกรรมการคุมสอบ การชี้แจงขั้นตอนและวิธีดำเนินการสอบต่อกรรมการคุมสอบ
3. อธิบายให้นักเรียนในกลุ่มตัวอย่างเข้าใจวัตถุประสงค์ของการทำแบบทดสอบ
4. นำแบบทดสอบไปทดสอบกับกลุ่มตัวอย่าง การทดสอบเป็นสามครั้ง ดังนี้
 - 4.1 ทดสอบครั้งที่ 1 เพื่อหาคุณภาพรายข้อ และปรับปรุงข้อสอบ
 - 4.2 ทดสอบครั้งที่ 2 เพื่อหาคุณภาพรายข้อ คัดเลือกและปรับปรุงข้อสอบ
 - 4.3 ทดสอบครั้งที่ 3 เพื่อหาคุณภาพรายข้อ คัดแนบเกณฑ์ในแต่ละจุดประสงค์เชิงพฤติกรรม และหาคุณภาพของแบบทดสอบ

✕ การวิเคราะห์ข้อสอบ

1. หากค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 1 จากการทดสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม
2. หากค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 2 จากการทดสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม
3. หากค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 3 จากการทดสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม

4. ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 4 จากการทดสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม

5. คะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรมในแบบทดสอบฉบับที่ 1 จากการทดสอบครั้งที่สาม

6. คะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรมในแบบทดสอบฉบับที่ 2 จากการทดสอบครั้งที่สาม

7. คะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรมในแบบทดสอบฉบับที่ 3 จากการทดสอบครั้งที่สาม

8. คะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรมในแบบทดสอบฉบับที่ 4 จากการทดสอบครั้งที่สาม

9. ค่าความเชื่อมั่นของแบบทดสอบแต่ละฉบับ จากการทดสอบครั้งที่สาม

10. ค่าความเที่ยงตรงของแบบทดสอบแต่ละฉบับ จากการทดสอบครั้งที่สาม

สรุปผลการวิเคราะห์ข้อมูล

1. ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 1 จากการทดสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม

1.1 จากการทดสอบครั้งที่หนึ่ง ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 17 % - 46 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 45 % - 83 % ค่าอำนาจจำแนกมีค่าตั้งแต่ .14 - .42

1.2 จากการทดสอบครั้งที่สอง ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 15 % - 52 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 54 % - 87 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .10 - .39

1.3' จากการทดสอบครั้งที่สาม ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 27 % - 50 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 58 % - 85 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .20 - .37

2. ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 2 จากการทดสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม

2.1 ทดสอบครั้งที่หนึ่ง ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 18 % - 34 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 48 % - 86 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .14 - .41

2.2 ทดสอบครั้งที่สอง ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 13 % - 32 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 41 % - 72 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .17 - .39

2.3 ทดสอบครั้งที่สาม ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 21 % - 37 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 53 % - 76 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .21 - .35

3. ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 3 จากการทดสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม

3.1 ทดสอบครั้งที่หนึ่ง ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 13 % - 35 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 43 % - 82 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .13 - .35

3.2 ทดสอบครั้งที่สอง ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 13 % - 34 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 44 % - 87 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .08 - .38

3.3 ทดสอบครั้งที่สาม ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 18 % - 34 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 54 % - 80 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .20 - .32

4. ค่าความยากในการสอบก่อนเรียน ค่าความยากในการสอบหลังเรียน และ ค่าอำนาจจำแนกของข้อสอบในแบบทดสอบฉบับที่ 4 จากการทดสอบครั้งที่หนึ่ง ครั้งที่สอง และครั้งที่สาม

4.1 ทดสอบครั้งที่หนึ่ง ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 14 % - 33 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 48 % - 77 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .09 - .36

4.2 ทดสอบครั้งที่สอง ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 17 % - 38 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 54 % - 82 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .16 - .35

4.3 ทดสอบครั้งที่สาม ค่าความยากในการสอบก่อนเรียนมีค่าตั้งแต่ 18 % - 28 % ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 54 % - 78 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .20 - .32

5. คะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรมในแบบทดสอบฉบับที่ 1 จากการทดสอบครั้งที่สาม

จุดประสงค์เชิงพฤติกรรม 1.1.1 คะแนนเกณฑ์คือ 4 ใน 5 ข้อ ค่าความน่าจะเป็นในการตัดสินเลือกเกณฑ์ถูกต้องเท่ากับ .63 ค่าความคลาดเคลื่อนในการจำแนกนิคประเภทที่ 1 และประเภทที่ 2 เท่ากับ .30 และ .07 ตามลำดับ ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์เท่ากับ .29

จุดประสงค์เชิงพฤติกรรม 1.1.2 คะแนนเกณฑ์คือ 4 ใน 5 ข้อ ค่าความน่าจะเป็นในการตัดสินใจเลือกเกณฑ์ถูกต้องเท่ากับ .74 ค่าความคลาดเคลื่อนในการจำแนกประเภทที่ 1 และประเภทที่ 2 เท่ากับ .17 และ .10 ตามลำดับ ส่วนค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์เท่ากับ .48

จุดประสงค์เชิงพฤติกรรม 1.1.3 คะแนนเกณฑ์คือ 4 ใน 5 ข้อ ค่าความน่าจะเป็นในการตัดสินใจเลือกเกณฑ์ถูกต้องเท่ากับ .71 ค่าความคลาดเคลื่อนในการจำแนกประเภทที่ 1 และประเภทที่ 2 เท่ากับ .23 และ .06 ตามลำดับ ส่วนค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์เท่ากับ .44

จุดประสงค์เชิงพฤติกรรม 1.2.1 คะแนนเกณฑ์คือ 5 ใน 7 ข้อ ค่าความน่าจะเป็นในการตัดสินใจเลือกเกณฑ์ถูกต้องเท่ากับ .82 ค่าความคลาดเคลื่อนในการจำแนกประเภทที่ 1 และประเภทที่ 2 เท่ากับ .11 และ .08 ตามลำดับ ส่วนค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์เท่ากับ .77

จุดประสงค์เชิงพฤติกรรม 1.2.2 คะแนนเกณฑ์คือ 5 ใน 7 ข้อ ค่าความน่าจะเป็นในการตัดสินใจเลือกเกณฑ์ถูกต้องเท่ากับ .82 ค่าความคลาดเคลื่อนในการจำแนกประเภทที่ 1 และประเภทที่ 2 เท่ากับ .10 และ .08 ตามลำดับ ส่วนค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์เท่ากับ .80

6. คะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรมในแบบทดสอบฉบับที่ 2 จากการทดสอบครั้งที่สาม

จุดประสงค์เชิงพฤติกรรม 2.1 คะแนนเกณฑ์คือ 6 ใน 8 ข้อ ค่าความน่าจะเป็นในการตัดสินใจเลือกเกณฑ์ถูกต้องเท่ากับ .76 ค่าความคลาดเคลื่อนในการจำแนกประเภทที่ 1 และประเภทที่ 2 เท่ากับ .23 และ .01 ตามลำดับ ส่วนค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์เท่ากับ .60

รูปประสงค์เชิงพฤติกรรม 2.2 คะแนนเกณฑ์คือ 5 ใน 8 ข้อ ค่าความน่าจะเป็นในการตัดสินใจเลือกเกณฑ์ถูกต้องเท่ากับ .78 ค่าความคลาดเคลื่อนในการจำแนกนิคมประเภทที่ 1 และประเภทที่ 2 เท่ากับ .20 และ .02 ส่วนค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์เท่ากับ .62

7. คะแนนเกณฑ์ในแต่ละรูปประสงค์เชิงพฤติกรรมในแบบทดสอบฉบับที่ 3 จากการทดสอบครั้งที่สาม

รูปประสงค์เชิงพฤติกรรม 3.1 คะแนนเกณฑ์คือ 5 ใน 8 ข้อ ค่าความน่าจะเป็นของการตัดสินใจเลือกเกณฑ์ถูกต้องเท่ากับ .74 ค่าความคลาดเคลื่อนในการจำแนกนิคมประเภทที่ 1 และประเภทที่ 2 เท่ากับ .20 และ .06 ตามลำดับ ส่วนค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์เท่ากับ .71

รูปประสงค์เชิงพฤติกรรม 3.2 คะแนนเกณฑ์คือ 5 ใน 8 ข้อ ค่าความน่าจะเป็นในการตัดสินใจเลือกเกณฑ์ถูกต้องเท่ากับ .64 ค่าความคลาดเคลื่อนในการจำแนกนิคมประเภทที่ 1 และประเภทที่ 2 เท่ากับ .33 และ .03 ตามลำดับ ส่วนค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์เท่ากับ .35

8. คะแนนเกณฑ์ของแต่ละรูปประสงค์เชิงพฤติกรรมในแบบทดสอบฉบับที่ 4 จากการทดสอบครั้งที่สาม

รูปประสงค์เชิงพฤติกรรม 4.1 คะแนนเกณฑ์คือ 6 ใน 8 ข้อ ค่าความน่าจะเป็นในการตัดสินใจเลือกเกณฑ์ถูกต้องเท่ากับ .88 ค่าความคลาดเคลื่อนในการจำแนกนิคมประเภทที่ 1 และประเภทที่ 2 เท่ากับ .11 และ .01 ตามลำดับ ส่วนค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์เท่ากับ .79

รูปประสงค์เชิงพฤติกรรม 4.2 คะแนนเกณฑ์คือ 4 ใน 8 ข้อ ค่าความน่าจะเป็นในการตัดสินใจเลือกเกณฑ์ถูกต้องเท่ากับ .73 ค่าความคลาดเคลื่อนในการจำแนกนิคมประเภทที่ 1 และประเภทที่ 2 เท่ากับ .17 และ .10 ตามลำดับ ส่วนค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์เท่ากับ .47

9. ค่าความเชื่อมั่นของแบบทดสอบแต่ละฉบับ จากการทดสอบครั้งที่สาม ซึ่งคำนวณโดยใช้สูตรของซัทโคเวียค ปรากฏว่า แบบทดสอบแต่ละฉบับมีค่าความเชื่อมั่นเรียงตามลำดับดังนี้ .6354, .6147, .7835, และ .7553

10. ค่าความเที่ยงตรงของแบบทดสอบแต่ละฉบับ จากการทดสอบครั้งที่สาม โดยหาความเที่ยงตรงตามเนื้อหา 2 แบบคือ

10.1 ความเที่ยงตรงตามเนื้อหาโดยให้ผู้เชี่ยวชาญตรวจสอบ ซึ่งได้รับการยืนยันจากผู้เชี่ยวชาญทางเนื้อหาวิชาและครูผู้สอนว่า การกำหนดจุดประสงค์เชิงพฤติกรรมได้ชัดเจน ข้อสอบแต่ละข้อวัดตรงตามจุดประสงค์เชิงพฤติกรรมที่ต้องการ แสดงว่าแบบทดสอบทั้งสี่ฉบับมีความเที่ยงตรงตามเนื้อหา

10.2 ความเที่ยงตรงตามเนื้อหา ซึ่งคำนวณโดยใช้สูตรของคาร์เวอร์ ปรากฏว่า แบบทดสอบฉบับที่หนึ่ง ฉบับที่สอง ฉบับที่สาม และฉบับที่สี่ มีค่าความเที่ยงตรงตามเนื้อหาเรียงตามลำดับดังนี้ .8226, .7177, .8427, และ .7460 ตามลำดับ

อภิปรายผล

ผลที่ได้จากการวิเคราะห์ข้อมูล อภิปรายผลได้ดังนี้

1. ค่าความยากและค่าอำนาจจำแนกของข้อสอบในแบบทดสอบทั้งสี่ฉบับ จากการทดสอบครั้งที่หนึ่ง ข้อสอบมีค่าความยากในการสอบก่อนเรียนตั้งแต่ 14% - 46% ค่าความยากในการสอบหลังเรียนมีค่าตั้งแต่ 43% - 86% ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .09 - .42 ข้อสอบส่วนใหญ่มีค่าความยากและค่าอำนาจจำแนกตามที่ต้องการ มีข้อสอบบางข้อมีค่าความยากในการสอบหลังเรียนต่ำกว่า 50%

ต่ำกว่า 50 % และค่าอำนาจจำแนกต่ำกว่า .20 ตามที่กำหนด เพราะการทดสอบครั้งนี้ เป็นการทดสอบครั้งแรกที่นำแบบทดสอบอิง เกณฑ์ทั้งสี่ฉบับไปทดสอบ ข้อสอบยังไม่ได้วิเคราะห์ มาก่อนย่อมมีข้อบกพร่องอยู่บาง เช่น ความกำกวมในการใช้ภาษา ความชัดเจนในคำสั่งชี้แจง และโจทย์ เมื่อปรับปรุงข้อสอบแล้ว นำแบบทดสอบทั้งสี่ฉบับไปทดสอบครั้งที่สอง ปรากฏว่า ข้อสอบมีค่าความยากในการสอบก่อนเรียนตั้งแต่ 13 % - 52 % ค่าความยากในการสอบ หลังเรียนมีค่าตั้งแต่ 44 % - 81 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .08 - .39 ข้อสอบ ส่วนใหญ่มีความยากและค่าอำนาจจำแนกที่ต้องการ และมีค่าความยากในการสอบหลัง เพิ่มขึ้น แสดงว่าข้อสอบง่ายขึ้น ทำให้ค่าอำนาจจำแนกลดต่ำลง แต่มีข้อสอบบางข้อมีความ ยากในการสอบหลังเรียนต่ำกว่า 50 % และค่าอำนาจจำแนกต่ำกว่า .20 แต่ก็มี จำนวนน้อยข้อ ข้อสอบส่วนใหญ่อยู่ในเกณฑ์นี้ เพราะได้ผ่านการปรับปรุงมาแล้ว เมื่อคัดเลือก ข้อสอบที่มีค่าความยากในการสอบก่อนเรียนต่ำกว่า 50 % ค่าความยากในการสอบหลัง เรียน สูงกว่า 50 % และมีค่าอำนาจจำแนกไม่ต่ำกว่า .20 ซึ่งคัดเลือกได้ครบจำนวนที่ต้องการ นำแบบทดสอบทั้งสี่ฉบับไปทดสอบครั้งที่สาม ซึ่งเป็นการทดสอบครั้งสุดท้ายของการสร้าง แบบทดสอบครั้งนี้ ปรากฏว่าข้อสอบมีค่าความยากในการสอบก่อนเรียนตั้งแต่ 18 % - 34 % ค่าความยากในการสอบหลัง เรียนมีค่าตั้งแต่ 53 % - 85 % ส่วนค่าอำนาจจำแนกมีค่าตั้งแต่ .20 - .37 ซึ่งมีค่าความยากและค่าอำนาจจำแนกอยู่ในเกณฑ์ที่ต้องการทุกข้อ เพราะข้อสอบ ได้วิเคราะห์และปรับปรุงจนมีคุณภาพแล้ว ดังนั้นข้อสอบทุกข้อจึงมีลักษณะยากในการสอบก่อนเรียน และง่ายเมื่อนำไปสอบหลัง เรียนตามลักษณะของแบบอิง เกณฑ์ (Berk. 1978 : 7) ตามต้องการ

2. คะแนนเกณฑ์ของจุดประสงค์เชิงพฤติกรรม ในแบบทดสอบอิง เกณฑ์ทั้งสี่ฉบับ

จากการทดสอบครั้งที่สาม ได้หาคะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรมในรูปของจำนวนข้อ ซึ่งเทียบในรูปเปอร์เซ็นต์ได้ค่าตั้งแต่ 50 % (4 ใน 8 ข้อ) - .80 % (4 ใน 5 ข้อ) ค่า ความน่าจะเป็นของการตัดสินเลือกเกณฑ์ถูกต้องมีค่าตั้งแต่ .63 - .88 ค่าความคลาดเคลื่อน

ในการจำแนกนิคทั้งสองประเภทเมื่อรวมกันมีค่าตั้งแต่ .12 - .37 ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์มีค่าตั้งแต่ .29 - .80 ซึ่งค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์แบ่งได้เป็น 2 ระดับใหญ่คือ ระดับที่มีค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์สูงกว่า .50 มีอยู่จำนวน 6 จุดประสงค์ ซึ่งมีค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์ตั้งแต่ .60 - .80 ค่าความน่าจะเป็นในการตัดสินใจเลือกเกณฑ์ที่ถูกต้องมีค่าตั้งแต่ .74 - .88 และค่าความคลาดเคลื่อนในการจำแนกนิคทั้งสองประเภทเมื่อรวมกันมีค่าตั้งแต่ .11 - .26 และที่ระดับค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์ต่ำกว่า .50 มีจำนวน 5 จุดประสงค์เชิงพฤติกรรม ซึ่งมีค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์ตั้งแต่ .29 - .48 ค่าความน่าจะเป็นของการตัดสินใจเลือกเกณฑ์ที่ถูกต้องมีค่าตั้งแต่ .63 - .74 และค่าความคลาดเคลื่อนในการจำแนกนิคทั้งสองประเภทเมื่อรวมกันมีค่าตั้งแต่ .27 - .37 แสดงว่าจุดประสงค์เชิงพฤติกรรมที่มีค่าความน่าจะเป็นในการตัดสินใจเลือกเกณฑ์ที่ถูกต้องสูง และค่าความคลาดเคลื่อนในการจำแนกนิคทั้งสองประเภทเมื่อรวมกันมีค่าต่ำ ย่อมมีค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์สูงด้วย เพราะค่าความน่าจะเป็นในการตัดสินใจเลือกเกณฑ์ที่ถูกต้อง ได้จากผลรวมของความน่าจะเป็นของกลุ่มที่เรียนแล้วและสอบผ่านเกณฑ์ กับความน่าจะเป็นของกลุ่มที่ยังไม่เรียน และสอบไม่ผ่านเกณฑ์ ส่วนค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์ ขึ้นอยู่กับค่าความน่าจะเป็นของกลุ่มที่เรียนแล้วและสอบผ่านเกณฑ์ กับค่าความคลาดเคลื่อนในการจำแนกนิคทั้งสองประเภท เมื่อค่าความน่าจะเป็นของกลุ่มที่เรียนแล้วมีค่าสูง จะทำให้ค่าความน่าจะเป็นในการตัดสินใจเลือกเกณฑ์ที่ถูกต้องและค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์สูงขึ้นด้วย ขณะเดียวกันค่าความคลาดเคลื่อนในการจำแนกนิคก็ลดลงด้วย เห็นได้จากเกณฑ์ที่มีค่าความเที่ยงตรงสูงสุดคือ จุดประสงค์เชิงพฤติกรรม 1.2.2 ซึ่งมีค่าสัมประสิทธิ์ความเที่ยงตรงเท่ากับ .80 มีค่าความน่าจะเป็นในการตัดสินใจเลือกเกณฑ์ที่ถูกต้องเท่ากับ .82 ซึ่งมีค่าในระดับสูงมาก และความคลาดเคลื่อนในการจำแนกนิค

ทั้งสองประเภทเมื่อรวมกันเท่ากับ .18 ซึ่งเป็นค่าในระดับต่ำ เกณฑ์ที่มีค่าสัมประสิทธิ์ความเที่ยงตรงต่ำสุดคือ จุฬประสงค์เชิงพฤติกรรม 1.1.1 มีค่าเท่ากับ .29 มีค่าความน่าจะเป็นในการตัดสินใจเลือกเกณฑ์ที่ถูกต้องเท่ากับ .63 ซึ่งมีค่าระดับต่ำสุด และค่าความคลาดเคลื่อนในการจำแนกทั้งสองประเภทเมื่อรวมกันเท่ากับ .37 ซึ่งมีค่าในระดับสูงสุด

3. ค่าความเชื่อมั่นของแบบทดสอบ ผลปรากฏว่าค่าความเชื่อมั่นของแบบทดสอบทั้งสี่ฉบับเรียงลำดับดังนี้ .6354, .6147, .7835 และ .7553 จะเห็นได้ว่าค่าความเชื่อมั่นของแบบทดสอบสองฉบับหลังมีค่าสูงกว่าสองฉบับแรก ทั้งนี้เพราะเกณฑ์ในแบบทดสอบฉบับที่สามและฉบับที่สี่ต่ำกว่าสองฉบับแรก จำนวนคนที่ผ่านเกณฑ์มีจำนวนมาก และสูตรหาค่ารวมของความน่าจะเป็นที่บุคคลตกในชั้นคะแนนจากคะแนนเกณฑ์จนถึงคะแนนเต็ม ซึ่งการประมาณค่าสัมประสิทธิ์ของความสอดคล้องสำหรับกลุ่มบุคคล อยู่ในรูปของค่าเฉลี่ยของความน่าจะเป็นในการตัดสินใจที่คงเส้นคงวาในกลุ่มจึงมีค่าออกมาสูง อย่างไรก็ตามแบบทดสอบทั้งสี่ฉบับมีค่าความเชื่อมั่นสูงกว่า .60 ซึ่งเป็นค่าที่สูงพอสมควร จึงแสดงว่าแบบทดสอบอิงเกณฑ์ทั้งสี่ฉบับมีความคงที่ในการตัดสินใจทดสอบ

4. ค่าความเที่ยงตรงของแบบทดสอบ ผลจากการตรวจสอบความเที่ยงตรงตามเนื้อหา ปรากฏว่าการกำหนดจุฬประสงค์เชิงพฤติกรรมเขียนไว้ชัดเจน ข้อสอบแต่ละข้อวัดได้ตรงตามจุฬประสงค์เชิงพฤติกรรมที่ต้องการทดสอบ ซึ่งสอดคล้องกับที่ สงบ ลักษณะ (สงบ ลักษณะ 2523 : 37 - 39) กล่าวไว้ว่า วิธีตรวจสอบความเที่ยงตรงตามเนื้อหาของแบบทดสอบอิงเกณฑ์นำมาใช้มากที่สุดคือ การให้ผู้เชี่ยวชาญทางเนื้อหาเป็นผู้พิจารณา ถ้าการกำหนดจุฬประสงค์เชิงพฤติกรรมเขียนไว้ชัดเจน และข้อสอบแต่ละข้อวัดได้ตรงตามจุฬประสงค์เชิงพฤติกรรมที่ต้องการทดสอบจริง แสดงว่าแบบทดสอบมีความเที่ยงตรงตามเนื้อหา

นอกจากนี้ผู้วิจัยได้คำนวณหาความเที่ยงตรงตามเนื้อหาโดยใช้สูตรของครวเวอร์ ได้วัดความเที่ยงตรงตามเนื้อหาของแบบทดสอบอิงเกณฑ์ทั้งสี่ฉบับ ดังนี้ .8226, .7177,

.8427 และ .7460 ตามลำดับ ซึ่งแบบทดสอบชุดฉบับนี้มีค่าความเที่ยงตรงตามเนื้อหา อยู่ในเกณฑ์สูงและมีค่าใกล้เคียงกัน แสดงว่าแบบทดสอบชุดฉบับนี้สามารถวัดความรู้ ในเนื้อหาของผู้สอบได้จริง และวัดได้ตามเนื้อหาที่ต้องการสอบ

★ ข้อเสนอแนะ

ข้อเสนอแนะในด้านการวิจัย ผู้วิจัยขอเสนอแนะ ดังนี้

1. แบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์เรื่องการหาร สำหรับนักเรียน ชั้นประถมศึกษาปีที่ 2 ในจังหวัดสมุทรสงครามเท่านั้น ถ้านำไปใช้ในท้องถิ่นอื่น ควรจะหาคะแนนเกณฑ์ในแต่ละจุดประสงค์เชิงพฤติกรรมใหม่ เพราะคะแนนเกณฑ์ ในที่นี้อาจจะไม่เหมาะสมกับนักเรียนในท้องถิ่นนั้น
2. ควรสร้างแบบทดสอบอิงเกณฑ์ในระดับชั้นต่างๆ ในเนื้อหาอื่นๆ เพื่อให้สอดคล้องกับระเบียบการประเมินผลตามหลักสูตรใหม่ โดยเฉพาะการประเมิน ผลระหว่างภาคที่กำหนดให้ตัดสินการสอบผ่านของนักเรียนตามรายข้อจุดประสงค์เชิง พฤติกรรม
3. ควรได้มีการศึกษาเปรียบเทียบวิธีการวัดผลระบบอิงเกณฑ์ในแบบต่างๆ เช่น การเปรียบเทียบวิธีการวิเคราะห์รายข้อหลายๆวิธี วิธีการหาความเชื่อมั่น ของแบบทดสอบอิงเกณฑ์วิธีอื่นๆ ตลอดจนเปรียบเทียบวิธีการกำหนดคะแนนเกณฑ์ในแต่ละ จุดประสงค์เชิงพฤติกรรมในแบบทดสอบอิง เกณฑ์ต่อไป

અનુભવ

บรรณานุกรม

กมล สุประเสริฐ "การเปรียบเทียบการวัดผลแบบอิงกลุ่มและอิงเกณฑ์" พัฒนาบทคัด 11 :

77 - 87 กรกฎาคม 2518

_____ การเปรียบเทียบวิธีการสองวิธีในการทดสอบความเที่ยงตรงของลำดับขั้นการเรียนรู้

ปริญญานิพนธ์ กศ.ภ. มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร 2520, 130 หน้า

อัครสำเนา

_____ "แนวความคิดที่เกี่ยวกับการวัดผลการเรียนการสอน" พัฒนาบทคัด 13 :

1 - 17 กรกฎาคม 2520

กมล สุประเสริฐ "ข้อคิดเกี่ยวกับการวัดผลการศึกษา" พัฒนาบทคัด 13 : 18 - 24

กรกฎาคม 2520

กาญจนา วัฒนสุนทร การสร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ วิทยานิพนธ์ ค.ม.

จุฬาลงกรณ์มหาวิทยาลัย 2521, 123 หน้า อัครสำเนา

โกวิท ประวาลพุกษ์ การประเมินผลแนวใหม่ : คู่มือสร้างสำหรับครูประถมศึกษา (ป.1)

กรุงเทพ มหานครนิเทศมาเยาชน 2522, 169 หน้า

_____ "การทดสอบแบบอิงเกณฑ์" วารสารการวัดผลการศึกษา 1(3) : 16 - 25

มกราคม - เมษายน 2523

โกวิท ประวาลพุกษ์ และ ส.วาสนา ประวาลพุกษ์ "การวัดผลการเรียนรู้แบบ Mastery

Learning Mastery Test" ศึกษาศาสตร์ 1(1) : 22 - 27 มกราคม - เมษายน

2518

จินนากา สัตบุทร การศึกษายลสัมฤทธิ์ทวงกรเรียนซ่อมเสริมคณิตศาสตร์วิธีการของนักเรียน

ชั้นประถมศึกษาปีที่ 2 ปริญญานิพนธ์ กศ.ม. มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

2521, 134 หน้า อัครสำเนา

บันลือ พุกกะวัน แนวการสอนสำเร็จรูป คู่มือครูสอน ป.2 ไทยวัฒนาพานิช 2519,
275 หน้า

บุญเชิด วิญญูอนันตพงศ์ การวัดและประเมินผลการศึกษาทฤษฎีและการประยุกต์
คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร 2521, 416 หน้า

นางจิต อิ่มสุวรรณ "ความเชื่อถือได้ของข้อสอบที่ใช้เกณฑ์กำหนด" วารสารการวัดผลการศึกษา
1(3) : 46 - 62 มกราคม - เมษายน 2523

นางชชาติ สุวรรณวงศ์ "การเขียนเพื่อรู้และแบบทดสอบอิงเกณฑ์" ศรีนครินทรสาร
2(2) : 32 - 38 ตุลาคม 2518 - มกราคม 2519

ไพศาล หวังพานิช "การสอบเพื่อปรับปรุงการเรียนการสอน" พัฒนากิจฉล 14 :
39 - 44 กรกฎาคม 2521

วิชัย ราษฎร์ศิริ หลักสูตรและแบบเรียนประถมศึกษา (หลักการและแนวปฏิบัติให้หลักสูตรใหม่
สัมฤทธิ์ผล) ไทยวัฒนาพานิช 2522, 170 หน้า

* วิชาการ, กรม แนวการใช้หลักสูตรประถมศึกษา พุทธศักราช 2521 เข็มตรัดเอ็กซ์
และสหศึกษาการพิมพ์ 2522, 211 หน้า

* _____ ระเบียบกระทรวงศึกษาธิการว่าด้วยการประเมินผลการเรียนระดับประถมศึกษา
เอกสารโรเนียว หมายเลข 9, 20 หน้า

* _____ รายละเอียดจุดประสงค์การเรียนรู้ในสมุทพริทัศน์ประถมศึกษาปีที่ 2
โรงพิมพ์การศาสนา 2522, 59 หน้า

ศึกษาธิการ, กระทรวง แบบเรียนวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 2 เล่ม 1, 2
โรงพิมพ์คุรุสภา 2521, 177 หน้า

_____ หลักสูตรประถมศึกษา พุทธศักราช 2521 โรงพิมพ์คุรุสภา 2522, 426 หน้า

สงม ลักษณะ "การตรวจสอบความเที่ยงตรงของข้อสอบอิงเกณฑ์" วารสารการวัดผลการศึกษา
1(3) : 36 - 44 มกราคม - เมษายน 2523

สงบ ลักษณะ "ปัญหาพหุฐานและเกณฑ์" วารสารการวัดและการศึกษา 2(1) : 49-55

กันยายน - ธันวาคม 2522

_____ . ทฤษฎีการสร้างแบบทดสอบอิงเกณฑ์ เอกสารอักษรา 8 หน้า, 2523

ส.วาสนา ประวาลพฤษ N.R. Versus C.R. Measurement พัฒนาวิทย์ 10 :

80 - 85 กรกฎาคม 2517

_____ . "การประเมินผลอิงอะไรก็เอียบ" วารสารวัดและการศึกษา 1(1) : 1 - 17

พฤษภาคม - สิงหาคม 2522

✓ สมบูรณ์ ชิตพงศ์ "การวัดผลกับหลักสูตร" พัฒนาวิทย์ 13 : 84 - 94 กรกฎาคม 2520

_____ . "จุดมุ่งหมายเชิงพฤติกรรม" พัฒนาวิทย์ 14 : 8 - 20 กรกฎาคม 2521

สมศักดิ์ สินธุระเวชญ์ "การประเมินผลแบบอิงกลุ่มและอิงเกณฑ์" วารสารวิจัยทางการศึกษา

1(8) : 74 - 86 มิถุนายน 2521

_____ . แบบทดสอบอิงเกณฑ์ เอกสารโรเนียว 2522, 10 หน้า

สำเริง บุญเรืองรัตน์ "การเรียนรู้เพื่อรู้" วารสารการวัดและการศึกษา 1(3) : 1 - 9

มกราคม - เมษายน 2523

อนันต์ ศรีโสภณ ทฤษฎีการวัดและการทดสอบ คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ

ประสานมิตร 2522, 281 หน้า

_____ . การวัดผลแบบอิงกลุ่มและอิงเกณฑ์ เอกสารโรเนียว 2524, 9 หน้า

Airasian, P.W. & Madaus, G.F. "Criterion - Referenced Testing in the Classroom," Measurement in Education. 15 : 1 - 8, Summer, 1972.

Berk, R.A. "Determination of Optimal (Sic.) Cutting Scores in Criterion - Referenced Measurement," Journal of Experimental Education. 45 : 4 = 9, July, 1976.

✓ _____ . "A Consumers Guide to Criterion - Referenced Test Item Statistics," A Series of Special Reports of N C M E, 9(1) : Winter 1978. 8 p.

- Brannan, R.C. "A Generalized Upper - Lower Discrimination Index," Educational and Psychological Measurement. 32(2) : 289 - 303, Summer, 1972.
- Brannan, R.C. and Stolurow, L.M. "An Empirical Decision Process for Formative Evaluation," Research Memorandum No. 4. Harvard CA I Laboratory, Cambridge, Mass, 8 : 5 - 10, June, 1971.
- ✓ Buter, N.W. "Societal Standards," Journal of Educational Measurement. 15 : 264 - 271, Winter, 1978.
- Carver, R.P. "Special Problems in Measuring Change with Psychometric Devices," Evaluation Research : Strategies and Methods. 10 : 250 - 258, Fall, 1970.
- Crehen, K.D. "Item Analysis for Teacher - made Mastery Test," Journal of Educational Measurement. 11 : 255 - 264, April, 1974.
- ✓ Enrich, J.A. "An Evaluation Model for Mastering Testing," Journal of Educational Measurement. 8 : 321 - 326, Winter, 1971.
- Ganopole, S.J. "The Construction and Validation of a Criterion Referenced Test of Fundamental Reading Competences," Dissertation Abstracts. 39(1) : 246 - A, July, 1978.
- ✓ Glaser, R. "Instructional Technology and the measurement of learning outcomes," American Psychologist. 18 : 519 - 521, May, 1963.
- Glaser, R. and Nitko, A.J. "Measurement in learning and Instruction in R.L. Thorndike (ed.)," Journal of Educational Measurement. 8 : 653, Winter, 1971.
- Glass, Gene V. "Standards and Criteria," Journal of Educational Measurement. 15 : 237 - 261, Winter, 1978.
- Haladyna, T.M. "Effects of Different Samples on Item and Test Characteristics of Criterion - Referenced Test," Journal of Educational Measurement. 11 : 93 - 97, April, 1974.
- ✓ Hambleton, R.K. "Criterion - Referenced Testing and Measurement : A Review of Technical Issues and Developments," Review of Educational Research. 48 (1) : 1 - 49, Winter - Fall, 1978.
- ✓ Hambleton, R.W. and Novick, M.R. "Toward an integration of theory and method for criterion - referenced tests," Journal of Educational Measurement. 10 (3) : 159 - 170, Fall, 1973.

- ✓ Holmes, D.H. "The Effects on Pupil Achievement and Attitude of Using a Criterion - Referenced Evaluation and Pupil Monitoring System in General Mathematics," Dissertation Abstracts. 39(7) : 4197 - A., January, 1979.
- Harris, C.W. "An Index of efficiency for fixed - length mastery tests," Research Association.
- ✓ Harris, C.W. and Stewart, D.M. "Application of classical strategies to criterion - referenced test construction," American Educational Research Association. 33 : 3 - 13, 1971.
- Hornor, D.M. "A Canonical Analysis of a Criterion - Referenced Test and Norm - Referenced Test in Mathematics," Dissertation Abstracts. 39(7) : 4193 - A, January, 1977.
- Hsu, Tse - Chi. "Empirical data on criterion - referenced tests," Paper Presented to The American Educational Research Association. 1971.
- Huynh, H. "On the Reliability of Decisions in Domain - Referenced Testing," Journal of Educational Measurement. 13(4) : 253 - 264, Spring, 1976.
- Kosecoff, J.B. and Klein, S.P. "Instructional sensitivity statistics appropriate for objectives - based test item," Center of the Study of Evaluation University of California, 1974.
- ✓ Livingston, S.A. "Criterion - referenced applications of classical test theory," Journal of Educational Measurement. 9(1) : 11 - 21, Winter - Fall, 1972.
- ✓ Mager, R.F. Preparing Instructional Objectives. Palo Alto, CA : Fearon Publishers 1962 : 280 pp.
- Millman, J. "Passing Scores and Test Lengths for Domain - Referenced Measurement," Review of Educational Research. 43 : 205 - 216, June, 1974.
- ✓ Millman, J. "Criterion - Referenced Measurement," Evaluation in Education. 307 - 397, 1974.
- Nikto, A.J. "Problems in the Development of Criterion - Referenced Tests," The IPI Pittsburg Experience. 1970 : 90 pp.
- Perla T. Garsia. "An Analysis of a Criterion - Referenced Test in Elementary Mathematics," Education Quarterly. 14(3), Fall, 1978.

✓ Popham, W.J. Criterion - Referenced Measurement. Prentice - Hall Inc., Englewood Cliffs, New Jersey, 1978, 260 pp.

Popham, W.J. and Husek, T.R. "Implication of Criterion - Referenced Measurement," Journal of Educational Measurement. 6(1) : 1 - 9, Spring, 1969.

Subkoviak, M.J. "Estimating Reliability a Single Administration of a Cristerion - Referenced Test," Journal of Educational Measurement. 13(4) : 265 - 276, Winter, 1976.

Swaminathan, H., Hambleton, R.K. and Algina, J. "Reliability of Criterion Reference Test," Journal of Educational Measurement. 11(4) : 263-267, Spring, 1974.

✓ Thrash, Susan Kaye, "A Model for Criterion - Referenced Measurement and a Comparison of Item Analysis Procedures," Dissertations Abstracts. 39(1) : 247 - A, July, 1978.

Wedman, I. "On the Evaluation of Criterion - Referenced Tests," In Cronbag, H.F. and De Gruijater, D.N. (Ed'S) Contemporary Issures in Educational Testing Mouton and Co., The Hague, 1974 : 80 pp.

Wilcox, R.R. and Harries, C.W. "On Emrick's "An Evaluation model for Mastering Testing," Journal of Educational Measurement. 14(3) : 215 - 217, Fall, 1977.

การผนวก

ภาคผนวก ก
สถิติที่ใช้วิเคราะห์ข้อมูล

สถิติที่ใช้วิเคราะห์ข้อมูล

1. สถิติพื้นฐาน คือ ค่าเฉลี่ย และค่าความเบี่ยงเบนมาตรฐาน

$$\bar{X} = \frac{\sum x}{N}$$

และ

$$S.D. = \sqrt{\frac{N \sum x^2 - (\sum x)^2}{N(N-1)}}$$

2. ค่าความยากของข้อสอบโดยใช้สูตรของ เบอร์ก (Berk. 1978 : 4)

$$DIFF = \frac{C}{N} \times 100$$

เมื่อ C คือ จำนวนนักเรียนที่ตอบข้อนั้นได้ถูกต้อง

N คือ จำนวนนักเรียนทั้งหมด

3. ค่าอำนาจจำแนก โดยใช้สูตรของโคสคอฟฟ์ และเคลน (Kosecoff and Klein. 1974 : 276 - 293)

$$DISMG = \frac{I_c}{N} - \frac{I_i}{N}$$

เมื่อ I_c คือ จำนวนนักเรียนที่ตอบผิดในการสอบก่อนเรียน
แต่ตอบถูกในการสอบหลังเรียน

I_i คือ จำนวนนักเรียนที่ตอบผิดในการสอบทั้งสองครั้ง

N คือ จำนวนนักเรียนที่เข้าสอบ

4. คะแนนเกณฑ์โดยใช้วิธีการของเบอร์ก (Berk. 1978 : 4 - 9)

ความน่าจะเป็นในการตัดสินใจถูกต้อง ได้จาก $P(TM) + P(FN)$

ความคลาดเคลื่อนของการจำแนกผิดประเภทที่ 1 คือ $P(FN)$

ความคลาดเคลื่อนของการจำแนกผิดประเภทที่ 2 คือ $P(FM)$

สัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์ ได้จาก

$$\phi_{vc} = \frac{P(TM) - BR(SR)}{\sqrt{BR(1 - BR) SR(1 - SR)}}$$

$$BR = P(FN) + P(TM)$$

$$SR = P(FM) + P(TM)$$

$$\text{เมื่อ } P(TM) = TM / (M + N)$$

$$P(TN) = TN / (M + N)$$

$$P(FM) = FM / (M + N)$$

$$P(FN) = FN / (M + N)$$

$$M = \text{จำนวนนักเรียนทั้งหมดที่ยังไม่เรียน}$$

$$N = \text{จำนวนนักเรียนทั้งหมดที่เรียนแล้ว}$$

5. ความเชื่อมั่นของแบบทดสอบ โดยใช้สูตรของซึบโกเวียค (Subkovisk.

1976 : 265 - 276)

$$P_c = \sum_{s=1}^N P_c^{(s)} / N$$

$$\text{เมื่อ } P_c = \left[P(X_s \geq 0) \right]^2 + \left[1 - P(X_s \geq 0) \right]^2$$

$$P(X_s \geq 0) = \sum_{X_s=0}^s \binom{n}{X_s} P_s^{X_s} (1 - P_s)^{n - X_s}$$

$$P_s = X_s / n$$

$$X_s \text{ เป็นจำนวนของการตอบถูก}$$

$$n \text{ เป็นจำนวนคำถามในแบบทดสอบ}$$

6. ความเที่ยงตรงตามเนื้อหาของแบบทดสอบ โดยใช้สูตรของคาร์เวอร์
(Carver. 1970 : 256)

	กลุ่มที่ยังไม่เรียน	กลุ่มที่เรียนแล้ว
ผ่าน	b	a
ไม่ผ่าน	c	d

$$\text{ความเที่ยงตรง} = (a + c) / N$$

$$N = a + b + c + d$$

ภาคผนวก ข

คู่มือการใช้แบบทดสอบอิงเกณฑ์

คู่มือการใช้แบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่องการหาร

แบบทดสอบชุดนี้สร้างขึ้นเพื่อนำไปใช้ในการทดสอบผลสัมฤทธิ์ในการเรียนวิชาคณิตศาสตร์เรื่องการหาร สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 ว่านักเรียนมีความรู้ความสามารถตามเกณฑ์ที่กำหนดไว้ในจุดประสงค์เชิงพฤติกรรมหรือไม่ จุดประสงค์ใดที่นักเรียนไม่ผ่าน การทบทวนเพิ่มเติมหรือซ่อมเสริมก่อนที่จะเรียนในจุดประสงค์ถัดไป แบบทดสอบชุดนี้ประกอบด้วยแบบทดสอบย่อย 4 ฉบับ คือ

แบบทดสอบฉบับที่ 1 ความหมายของการหาร

แบบทดสอบฉบับที่ 2 สัญลักษณ์การหาร

แบบทดสอบฉบับที่ 3 โจทย์การหาร

แบบทดสอบฉบับที่ 4 โจทย์ปัญหาการหาร

โครงสร้างของแบบทดสอบ

แบบทดสอบฉบับที่ 1 เป็นแบบทดสอบปรนัยชนิดเลือกตอบ 3 ตัวเลือก
จำนวน 29 ข้อ แบ่งเป็น 2 ตอน คือ

ตอนที่ 1 การนับลคครั้งละเท่ากัน ได้แก่ การหาจำนวนนับลคที่ถัดไป
จำนวนครั้งที่นับลคจนหมด และจำนวนครั้งที่นับลคจนหมดและเศษที่เหลือ

ตอนที่ 2 ความรู้เบื้องต้นในเรื่องการหาร ได้แก่ การหาจำนวนกลุ่ม
และการหาจำนวนในแต่ละกลุ่ม

ตัวอย่างข้อสอบ

ข้อ (0) จาก 13 นับลคทีละ 4 จะนับลคได้กี่ครั้งและเหลือเศษเท่าไร

ก. 4 ครั้ง เหลือ 1

ข. 3 ครั้ง เหลือ 1

ค. 3 ครั้ง เหลือ 2

คำตอบที่ถูกต้อง คือ ข.

ข้อ (00) ส้มทั้งหมด 12 ผล แบ่งเป็น 3 กอง จะมีกองละกี่ผล



ก. 5 ผล

ข. 4 ผล

ค. 3 ผล

คำตอบที่ถูกต้อง คือ ข.

แบบทดสอบฉบับที่ 2 เป็นแบบทดสอบปรนัยชนิดเลือกตอบ 3 ตัวเลือก

จำนวน 16 ข้อ แบ่งเป็น 2 ตอน ดังนี้

ตอนที่ 1 ประโยคสัญลักษณ์การหาร

ตอนที่ 2 ความสัมพันธ์ระหว่างการคูณและการหาร

ตัวอย่างข้อสอบ

ข้อ (๑)



จากภาพจะแทนด้วยข้อใดจึงจะถูกต้อง

ก. $8 \div 2 = 4$

ข. $8 \div 2 = 2$

ค. $4 \div 2 = 2$

คำตอบที่ถูกต้อง คือ ก.

ข้อ (00) $4 \times \square = 16$ ดังนั้น $16 \div \square = 4$

ก. 6

ข. 4

ค. 3

คำตอบที่ถูกต้อง คือ ข.

แบบทดสอบฉบับที่ 3 เป็นแบบทดสอบปรนัยชนิดเลือกตอบ 3 ตัวเลือก

จำนวน 16 ข้อ แบ่งเป็น 2 ตอน ดังนี้

ตอนที่ 1 การหารลงตัว

ตอนที่ 2 การหารมีเศษ

ตัวอย่างข้อสอบ

ข้อ (๐) $20 \div 5 = \square$

ก. 6

ข. 5

ค. 4

คำตอบที่ถูกต้อง คือ ค.

ข้อ (๐๐) $56 \div 6 = ?$

ก. 9 เศษ 2

ข. 9

ค. 8 เศษ 2

คำตอบที่ถูกต้อง คือ ก.

แบบทดสอบฉบับที่ 4 เป็นแบบทดสอบปรนัยชนิดเลือกตอบ 3 ตัวเลือก

จำนวน 16 ข้อ แบ่งเป็น 2 ตอน คือ

ตอนที่ 1 การหารลงตัว

ตอนที่ 2 การหารมีเศษ

ตัวอย่างข้อสอบ

ข้อ (๐) ลำไย 32 ผล แบ่งให้นัก 4 คน จะได้คนละกี่ผล

ก. 9 ผล

ข. 6 ผล

ค. 7 ผล

คำตอบที่ถูกต้อง คือ ข.

ข้อ (๐๐) ข้าวโพด 46 ผล จัดใส่ถาดละ 5 ผล จะต้องใช้ถาดกี่ใบ
และเหลือกี่ผล

ก. 9 ใบ เหลือ 1 ผล

ข. 8 ใบ เหลือ 1 ผล

ค. 7 ใบ เหลือ 1 ผล

คำตอบที่ถูกต้อง คือ ก.

ลักษณะของแบบทดสอบ

แบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่องการหาร มีจุดมุ่งหมายในการสร้าง คือ ใช้ทดสอบนักเรียนหลังจากการสอนเนื้อหาแต่ละเนื้อหาสิ้นสุดลง ว่านักเรียนมีความรู้ความสามารถถึงเกณฑ์ที่กำหนดไว้ในจุดประสงค์เชิงพฤติกรรมหรือไม่ นักเรียนที่สอบไม่ผ่านเกณฑ์ในจุดประสงค์ใดถือว่านักเรียนยังบกพร่องในเรื่องนั้น ครูผู้สอนควรทบทวนเพิ่มเติมหรือซ่อมเสริมก่อนที่จะเรียนในจุดประสงค์ถัดไป คะแนนรวมของนักเรียนจึงไม่มีความสำคัญ เพราะไม่มีการเปรียบเทียบคะแนนของนักเรียนในกลุ่ม แบบทดสอบชุดนี้อาจนำไปใช้ทดสอบนักเรียนเป็นกลุ่มหรือรายบุคคลก็ได้

การพัฒนาแบบทดสอบ

การสร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่องการหารชุดนี้ เริ่มดำเนินการสร้างเมื่อพ.ศ. 2523 และสร้างสำเร็จเมื่อพ.ศ. 2524 โดยใช้ นักเรียนชั้นประถมศึกษาปีที่ 2 ในจังหวัดสมุทรสงคราม จำนวน 370 คนเป็นกลุ่มตัวอย่าง โดยทำการทดสอบ 3 ครั้ง คือ การทดสอบครั้งที่ 1 และครั้งที่ 2 เพื่อหาคุณภาพและคัดเลือกข้อสอบ ครั้งที่ 3 เพื่อหาคุณภาพรายข้อ คะแนนเกณฑ์ในแต่ละจุดประสงค์เชิงพฤติกรรม และหาคุณภาพของแบบทดสอบที่สร้างขึ้น

คุณภาพของแบบทดสอบ

1. คุณภาพรายข้อ หมายถึง ค่าความยากของ ข้อสอบ คำนวณโดยใช้ สูตรของเบอร์ก และค่าอำนาจจำแนกของข้อสอบ โดยใช้สูตรของโคสคอฟฟ์และเคลน ได้ค่าความยากและค่าอำนาจจำแนกของข้อสอบในแบบทดสอบทั้ง 4 ฉบับ ดังแสดงใน ตาราง 1

ตาราง 1 ค่าความยากและค่าอำนาจจำแนกของข้อสอบในแบบทดสอบทั้ง 4 ฉบับ

แบบทดสอบฉบับที่	ค่าความยาก		ค่าอำนาจจำแนก
	ก่อนเรียน	หลังเรียน	
1. ความหมายการหาร	27%-50%	58%-85%	.20-.37
2. สัญลักษณ์การหาร	21%-37%	53%-76%	.21-.35
3. โจทย์การหาร	18%-34%	54%-80%	.20-.32
4. โจทย์ปัญหาการหาร	18%-28%	54%-78%	.20-.32

2. ค่าความเชื่อมั่นของแบบทดสอบ หมายถึง ค่าสัมประสิทธิ์ของความสอดคล้องของการตัดสินที่คงเส้นคงวาในกลุ่มผู้สอบ ความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ทั้ง 4 ฉบับนี้ได้จากการใช้สูตรของซิปโคเวียค ดังแสดงในตาราง 2

ตาราง 2 ค่าสถิติพื้นฐานและค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ทั้ง 4 ฉบับ

แบบทดสอบฉบับที่	จำนวนข้อสอบ	คะแนนเกณฑ์	ค่าเฉลี่ย	ค่าเบี่ยงเบนมาตรฐาน	ค่าความเชื่อมั่น
1. ความหมายการหาร	29	22	22.2743	7.3593	.6354
2. สัญลักษณ์การหาร	16	11	10.0968	2.2874	.6147
3. โจทย์การหาร	16	10	10.1774	4.3069	.7835
4. โจทย์ปัญหาการหาร	16	10	10.5726	2.8783	.7553

3. ค่าความเที่ยงตรงของแบบทดสอบ มี 2 ชนิดคือ

3.1 ความเที่ยงตรงตามเนื้อหา หมายถึง แบบทดสอบสามารถวัดได้ตรงตามเนื้อหา ข้อสอบเป็นตัวแทนของพฤติกรรมในแต่ละจุดประสงค์เชิงพฤติกรรมที่ต้องการทดสอบ โดยให้ผู้เชี่ยวชาญทางเนื้อหาเป็นผู้พิจารณาและลงความเห็น

แบบทดสอบอิงเกณฑ์ทั้งสี่ฉบับสามารถวัดเนื้อหาในเรื่องการหารได้จริง นอกจากนี้ได้หาค่าความเที่ยงตรงตามเนื้อหาโดยใช้สูตรของคาร์เวอร์ ดังแสดงในตาราง 3

ตาราง 3 ค่าความเที่ยงตรงตามเนื้อหาของแบบทดสอบอิงเกณฑ์ทั้ง 4 ฉบับ

แบบทดสอบฉบับที่	จำนวนข้อสอบ	คะแนนเกณฑ์	กลุ่มตัวอย่าง	ความเที่ยงตรง
1. ความหมายการหาร	29	22	124	.8226
2. สัญลักษณ์การหาร	16	11	124	.7177
3. โจทย์การหาร	16	10	124	.8427
4. โจทย์ปัญหาการหาร	16	10	124	.7460

3.2 ความเที่ยงตรงตามโครงสร้างของแบบทดสอบ หมายถึง การตรวจสอบคุณลักษณะภายในของผู้สอบ ว่ามีความสามารถหรือมีพฤติกรรมตามจุดประสงค์เชิงพฤติกรรมทั้งมวลที่สอบวัด ค่าความเที่ยงตรงตามโครงสร้างหาโดยอาศัยค่าสัมประสิทธิ์ความสอดคล้องภายใน ดังแสดงในตาราง 4

ตาราง 4 ค่าความเที่ยงตรงตามโครงสร้างของแบบทดสอบอิงเกณฑ์ทั้ง 4 ฉบับ

แบบทดสอบฉบับที่	จำนวนข้อสอบ	คะแนนเกณฑ์	ค่าความเที่ยงตรง
1. ความหมายของการหาร	29	22	.5276
2. สัญลักษณ์การหาร	16	11	.4670
3. โจทย์การหาร	16	10	.5412
4. โจทย์ปัญหาการหาร	16	10	.5820

เวลาที่ใช้ในการทดสอบ

แบบทดสอบอิงเกณฑ์เป็นแบบทดสอบที่สร้างขึ้นเพื่อใช้ในการประเมินผลย่อยหรือประเมินผลเพื่อปรับปรุงการเรียนการสอน เวลาที่ใช้ในการทดสอบจึงควรเป็นเวลาที่ใช้ในระหว่างการเรียนการสอน เช่น เมื่อจบเนื้อหาย่อยในแต่ละจุดประสงค์เชิงพฤติกรรม หรือในคอนันที่จะขึ้นเนื้อหาใหม่ กำหนดเวลาที่นักเรียนควรจะทำแบบทดสอบแต่ละฉบับเสร็จครบทุกข้อ ดังแสดงในตาราง 4

ตาราง 4 เวลาที่ใช้ในการทดสอบคิดเป็นนาที

แบบทดสอบฉบับที่	ชี้แจง	ทำข้อสอบ	รวม
1. ความหมายการหาร	5	40	45
2. สัญลักษณ์การหาร	5	25	30
3. โจทย์การหาร	5	25	30
4. โจทย์ปัญหาการหาร	5	25	30

วิธีดำเนินการทดสอบ

1. การเตรียมตัวก่อนทดสอบ

1.1 เตรียมแบบทดสอบ กระจายคำตอบและกระจายหาคให้มีมากกว่าจำนวนผู้สอบ สำหรับนักเรียนบางคนที่ไม่ชอบแบบทดสอบไม่ชัด กระจายคำตอบหาก เป็นกัน

1.2 ผู้ดำเนินการสอบต้องอ่านและศึกษาคำชี้แจงวิธีการทำแบบทดสอบอย่างน้อยหนึ่งครั้ง เพื่อให้สามารถดำเนินการสอบได้ชัดเจน

2. วิธีดำเนินการขณะสอบ

2.1 แจกแบบทดสอบ กระจายคำตอบ และกำชับนักเรียนไม่ให้ลงมือทำจนกว่าจะได้รับคำชี้แจง

2.2 ชี้แจงวิธีการทำแบบทดสอบ ให้นักเรียนเขียนรายละเอียดต่างๆ ที่เกี่ยวกับตัวนักเรียนเอง เสร็จแล้วครูอ่านคำชี้แจงให้นักเรียนฟังโดยใช้เวลา 5 นาที

2.3 เมื่อเวลาผ่านไปครึ่งหนึ่งของเวลาทั้งหมด ควรบอกเวลาที่เหลือแก่นักเรียน และเมื่อเหลือเวลาอีก 5 นาที ควรบอกเวลาแก่นักเรียนอีกครั้งหนึ่ง เพื่อให้นักเรียนเกิดแรงจูงใจ และรีบทำข้อสอบอย่างตั้งใจ

2.4 เมื่อหมดเวลา ให้นักเรียนนำแบบทดสอบและกระดาษคำตอบส่งผู้ดำเนินการสอบ

วิธีการตรวจให้คะแนนและการรายงานผลการสอบ

1. นำกระดาษคำตอบของนักเรียนแต่ละคนมาตรวจให้คะแนน ตอบถูกให้ 1 ตอบผิด ไม่ตอบ หรือตอบเกินกว่าหนึ่งคำตอบให้ 0

2. รวมคะแนนของนักเรียนแต่ละคนที่ทำข้อสอบถูกในแต่ละจุดประสงค์เชิงพฤติกรรม ถ้าคะแนนที่ได้น้อยกว่าคะแนนเกณฑ์ในแต่ละจุดประสงค์เชิงพฤติกรรม ก็แสดงว่านักเรียนบกพร่องในจุดประสงค์เชิงพฤติกรรมนั้น ถ้าคะแนนที่ได้มากกว่าหรือเท่ากับคะแนนเกณฑ์ในแต่ละจุดประสงค์เชิงพฤติกรรม ก็แสดงว่านักเรียนมีความรู้ในจุดประสงค์นั้น

3. นำผลการสอบบันทึกลงใบแจ้งผลการเรียน ดังตัวอย่าง

แบบรายงานผลการเรียนวิชาคณิตศาสตร์ เรื่องการหาร

โรงเรียน..... ชั้นประถมศึกษาปีที่ 2

ชื่อ ปีการศึกษา

แบบทดสอบ ฉบับที่	จุดประสงค์ เชิงพฤติกรรม	คะแนน เกณฑ์	ข้อที่								คะแนน รวม	ผลการ สอบ
			1	2	3	4	5	6	7	8		
1	1.1.1	4										
	1.1.2	4										
	1.1.3	4										
	1.2.1	5										
	1.2.2	5										
2	2.1	6										
	2.2	5										
3	3.1	5										
	3.2	5										
4	4.1	6										
	4.2	4										

นักเรียนทำข้อสอบถูกข้อใดก็กา "X" ลงในช่องข้อสอบนั้น ถ้าผิดก็ทำเครื่องหมาย "-" ลงในข้อนั้น คะแนนรวมในจุดประสงค์เชิงพฤติกรรมใดสูงกว่าหรือเท่ากับคะแนนเกณฑ์ ใส่ "ผ" ในช่องผลการสอบ แต่ถ้าคะแนนรวมต่ำกว่าคะแนนเกณฑ์ ใส่ "มผ" ลงในช่องผลการสอบ

ข้อเสนอแนะ

1. แบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์เรื่องการหาร ที่ผู้วิจัยสร้างขึ้น ใช้สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 ในจังหวัดสมุทรสงครามเท่านั้น ถ้าจะนำไปใช้ในท้องถิ่นอื่น ควรจะหาคะแนนเกณฑ์ในแต่ละจุดประสงค์เชิงพฤติกรรมใหม่ เพราะคะแนนเกณฑ์ในที่นี้อาจไม่เหมาะสมกับนักเรียนในท้องถิ่นนั้น
2. ควรใช้แบบทดสอบนี้ทดสอบนักเรียนหลังจากที่ทำการสอนเนื้อหาย่อยแต่ละเนื้อหาสิ้นสุดลง นักเรียนที่ไม่ผ่านในจุดประสงค์ใด ครูผู้สอนควรทบทวนเพิ่มเติมหรือซ่อมเสริมให้ก่อนที่นักเรียนจะเรียนในจุดประสงค์ถัดไป
3. เมื่อพบว่านักเรียนส่วนใหญ่ได้คะแนนรวมต่ำกว่าคะแนนเกณฑ์ ควรปรับปรุงวิธีการสอนใหม่ หรือเพิ่มเวลาการสอนในจุดประสงค์เชิงพฤติกรรมนั้นให้มากขึ้น

การสร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 2

บทคัดย่อ

ของ

สมถวิล วิจิตรวรรณ

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาการศึกษามหาบัณฑิต

กรกฎาคม 2524

การศึกษาค้นคว้าครั้งนี้มีจุดมุ่งหมายเพื่อสร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่องการหาร ชั้นประถมศึกษาปีที่ 2 และคำนวณหาคะแนนเกณฑ์ในแต่ละจุดประสงค์เชิงพฤติกรรมทั้ง 11 จุดประสงค์ กลุ่มตัวอย่างในการศึกษาค้นคว้าครั้งนี้ประกอบด้วย นักเรียนชั้นประถมศึกษาปีที่ 2 ปีการศึกษา 2523 ในจังหวัดสมุทรสงคราม จำนวน 370 คน ซึ่งเลือกมาโดยการสุ่มแบบแบ่งชั้น แบบทดสอบอิงเกณฑ์ที่สร้างขึ้น ประกอบด้วยแบบทดสอบย่อย 4 ฉบับที่สร้างตามจุดประสงค์เชิงพฤติกรรม มีวิธีดำเนินการดังนี้ ทดสอบครั้งที่หนึ่งและครั้งที่สอง เพื่อวิเคราะห์รายข้อ และทดสอบครั้งที่สามเพื่อหาคุณภาพของแบบทดสอบย่อยแต่ละฉบับ และคำนวณหาคะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรม

ผลจากการศึกษาพบว่า แบบทดสอบอิงเกณฑ์ที่สร้างขึ้นมีค่าความยากของ ข้อสอบสูงกว่า 50% โดยใช้วิธีการของเบอร์ก ค่าอำนาจจำแนกของข้อสอบหา ได้จากวิธีถ่วงของโคสคอฟฟ์และเคลนมีค่าตั้งแต่ .20 ขึ้นไป ค่าความเชื่อมั่นของ แบบทดสอบย่อยแต่ละฉบับคำนวณหาโดยใช้สูตรของซัทโคเวียคมีค่า .6354, .6147, .7835 และ .7553 ตามลำดับ ความเที่ยงตรงตามเนื้อหาของแต่ละแบบ ทดสอบย่อยทั้งสี่ฉบับใช้วิธีการของโรวินเนลลีและแฮมเปิลตัน โดยให้ผู้เชี่ยวชาญ ทางเนื้อหาวิชาเป็นผู้พิจารณาตรวจสอบ และคำนวณหาค่าความเที่ยงตรงตามเนื้อหา โดยใช้สูตรของคาร์เวอร์ ปรากฏค่า .8226, .7177, .8427 และ .7460 ตามลำดับ สำหรับคะแนนเกณฑ์ของแต่ละจุดประสงค์เชิงพฤติกรรม มีค่าตั้งแต่ 50% (4 ใน 8 ข้อ) ถึง 80% (4 ใน 5 ข้อ)

A CONSTRUCTION OF THE CRITERION - REFERENCED TEST
IN PRATHOM SUKSA II MATHEMATICS

AN ABSTRACT

BY

SOMTAPIN WIJITANNA

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

July 1981

The purposes of this study were to construct a high qualified criterion test in Mathematics on "Division" for Prathom Suksa II and to find the appropriate criteria of all 11 behavioral objectives. The sample of the study consisted of 370 Prathom Suksa II students of the 1980 academic year in Samut Songkhram province was selected by using stratified random sampling. According to the behavioral objectives, 4 subtests were developed, then the tests were tried out twice for the item analysis purpose. The final form of the test was administered to the selected sample and the responses were analyzed

The results of study indicated that the difficulty index of each item was more than 50 %, and the discrimination index of each item was equal to or more than .20. The indices were calculated by Berk's formula and Kosecoff and Klein's formula, respectively. The reliability coefficient of the 4 subtests, calculated by using Subkoviak's method, were .6354, .6147, .7835, and .7553. The content validity of each subtest was studied by using two methods proposed by Rovinalli and Rovinalli and Hambleton and Caver. The Carver's formula gave the coefficient of .8226, .7177, .8427, and .7460. The criterion of each behavioral objective, calculated by Berk's formula ranged from 50 % (4 out of 8 item) to 80 % (4 out of 5 items)