

## รายงานวิจัย

รายงานวิจัย ผลของการวิเคราะห์คุณภาพของข้อสอบ วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6  
 กลางภาคเรียนที่ 2 ปีการศึกษา 2559  
 ชื่อผู้วิจัย นางสาวภัชชรา ไ้ม่างาม  
 ปีที่ทำการวิจัย 2559  
 ที่ปรึกษา นางสมร อุทานวรวจน์

## บทคัดย่อ

การวิจัยเรื่อง รายงานการวิเคราะห์ข้อสอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 กลางภาคเรียนที่ 2 ปีการศึกษา 2559 มีวัตถุประสงค์เพื่อต้องการหาคุณภาพของข้อสอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 กลางภาคเรียนที่ 2 ปีการศึกษา 2559 กลุ่มตัวอย่างที่ใช้ในการวิจัยคือนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2558 โรงเรียนปี่คอนเฮาส์แย้มสอาดรังสิต

ขอบเขตในการวิจัยด้านเนื้อหา คือข้อสอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 กลางภาคเรียนที่ 2 ปีการศึกษา 2559 หน่วยที่ 7 เรื่อง ทศนิยมและการบวก การลบ การคูณทศนิยม หน่วยที่ 8 เรื่อง การหารทศนิยม หน่วยที่ 9 เรื่อง รูปเรขาคณิต

เครื่องมือที่ใช้ในการวิจัย ได้แก่ ข้อสอบกลางภาคเรียนที่ 2 ปีการศึกษา 2559 วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 การวิเคราะห์ข้อสอบใช้โปรแกรมวิเคราะห์ข้อสอบ ช่วยในการคำนวณและวิเคราะห์ข้อสอบบ่งชี้คุณภาพของข้อสอบ เรื่องระดับความยากง่ายของข้อสอบ อำนาจจำแนกของข้อสอบและประสิทธิภาพของตัวลอง การวิเคราะห์ข้อมูล หาค่าความยากง่าย ค่าอำนาจจำแนกเป็นรายข้อ และหาค่าความเชื่อถือได้ทั้งฉบับ โดยใช้โปรแกรมวิเคราะห์ข้อสอบของนายธีรวัฒน์ สุพัตกุล จังหวัดบุรีรัมย์

สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ค่าความยากง่าย ค่าอำนาจจำแนกและค่าความเชื่อถือได้

ผลการวิจัยพบว่าข้อสอบที่ดีควรเก็บไว้มีจำนวน 24 ข้อ คิดเป็นร้อยละ 80 ข้อสอบที่ควรปรับปรุงหรือตัดทิ้งมีจำนวน 6 ข้อ คิดเป็นร้อยละ 20 เป็นไปตามวัตถุประสงค์ของผู้ทำการวิจัย ที่ต้องการวัดคุณภาพของข้อสอบให้เป็นไปตามมาตรฐาน และนำผลการวิเคราะห์ไปปรับปรุงคุณภาพของข้อสอบให้ดียิ่งขึ้นต่อไป

## รายงานวิจัย

ผลของการวิเคราะห์คุณภาพของข้อสอบ วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6  
กลางภาคเรียนที่ 2 ปีการศึกษา 2559

### บทที่ 1

#### บทนำ

##### 1.1 ความเป็นมาและความสำคัญของปัญหา

การตรวจสอบคุณภาพของแบบทดสอบ เพื่อหาคุณลักษณะของแบบทดสอบที่ดี โดยปกติ จะตรวจสอบ ตั้งแต่ก่อนนำแบบทดสอบไปใช้ โดยตรวจสอบว่าข้อสอบที่สร้างขึ้นเป็นไปตามหลักการ เขียน ข้อสอบ ตรงตามแผนสร้าง แบบทดสอบ (Test blue print) ข้อสอบที่สร้างมีความเป็นปรนัย ตรงตามพฤติกรรมที่ต้องการจะวัด และตรวจสอบหลังนำไปใช้ โดยการคำนวณค่าสถิติ ของแบบทดสอบเป็น รายข้อและตรวจสอบคุณภาพแบบทดสอบทั้งฉบับโดยการตรวจสอบคุณภาพแบบทดสอบเป็นรายข้อ มีจุดมุ่งหมายเพื่อตรวจสอบคุณภาพ 2 ประการ คือ ความยากของข้อสอบ (difficulty : p) และอำนาจจำแนกของข้อสอบ (discrimination : r, D) ส่วนการตรวจสอบ คุณภาพของแบบทดสอบทั้งฉบับ มีจุดมุ่งหมายเพื่อหาคุณภาพ 2 ประการ คือ ความตรง (validity) และความเที่ยงของแบบทดสอบ (reliability)

ในด้านการเรียนการสอน แบบทดสอบเป็นเครื่องมือที่ใช้วัดผลสัมฤทธิ์ทางการเรียนของนักเรียนหาก แบบทดสอบไม่มีคุณภาพ อาจจะทำให้ยากไปหรืออาจจะง่ายไป จะมีผลต่อคะแนนหรือผลสัมฤทธิ์ของผู้เรียน คือ หากแบบทดสอบยากนักเรียนก็ได้คะแนนน้อยและหากแบบทดสอบง่ายนักเรียนก็ได้คะแนนมาก

เนื่องจากโรงเรียนปีคอนเฮาส์แยมส์อาดริงลิต เป็นโรงเรียนเอกชนขนาดใหญ่ จึงมีการแข่งขันใน มาตรฐาน และคุณภาพของผู้เรียน ดังนั้นครูผู้สอนมีภาระหน้าที่มาก บางครั้งจึงมีเวลาน้อยในการออกข้อสอบ ทำให้ข้อสอบที่ออกมาในแต่ละครั้งบางครั้งที่ปรึกษาตรวจแล้ว ข้อสอบจะถูกแก้ไขมาก จึงทำให้เกิด ความล่าช้าในการออกข้อสอบ

ผู้วิจัยซึ่งทำหน้าที่รับผิดชอบงานด้านวัดผล และเป็นผู้รับผิดชอบด้านการจัดสอบของโรงเรียน ทั้งกลางภาคและปลายภาค ดังนั้นผู้วิจัยจึงสนใจที่จะศึกษาเกี่ยวกับการหาคุณภาพของข้อสอบ และการเก็บ รวบรวมข้อสอบที่มีคุณภาพไว้ในธนาคารข้อสอบ เพื่อที่จะได้สะดวกและรวดเร็วในการออกข้อสอบ ของครูผู้สอนในครั้งต่อไป

##### 1.2 วัตถุประสงค์ของการวิจัย

เพื่อต้องการหาคุณภาพของข้อสอบ วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 กลางภาคเรียนที่ 2 ปีการศึกษา 2559

### 1.3 สมมติฐานของการวิจัย

การศึกษาในครั้งนี้จะทำให้ได้ข้อสอบที่มีคุณภาพและมีความเชื่อถือได้

### 1.4 ขอบเขตของการวิจัย

#### 1.4.1 ขอบเขตด้านเนื้อหา

วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 กลางภาคเรียนที่ 2 ปีการศึกษา 2559

เนื้อหาวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 กลางภาคเรียนที่ 2 ปีการศึกษา 2559

#### หน่วยที่ 7

##### เรื่อง ทศนิยมและการบวก การลบ การคูณทศนิยม

- การอ่านและการเขียนทศนิยมไม่เกินสามตำแหน่ง
- ค่าประจำหลักของทศนิยมและเขียนทศนิยมในรูปการกระจาย
- การเปรียบเทียบทศนิยม
- ความสัมพันธ์ระหว่างทศนิยมกับเศษส่วน
- การประมาณค่าใกล้เคียงทศนิยม
- โจทย์ปัญหาการบวก ลบทศนิยม
- การคูณทศนิยม

#### หน่วยที่ 8

##### เรื่อง การหารทศนิยม

- การหารทศนิยมกับจำนวนนับและการประมาณคำตอบ
- การเขียนเศษส่วนให้อยู่ในรูปของทศนิยม
- การหารทศนิยมเมื่อตัวหารเป็นทศนิยม
- โจทย์ปัญหาการหารทศนิยม
- โจทย์ปัญหาระคน

#### หน่วยที่ 9

##### เรื่อง รูปเรขาคณิต

- เส้นทแยงมุมและสมบัติของเส้นทแยงมุม
- การหาความยาวรอบรูปสี่เหลี่ยม
- การหาพื้นที่รูปสี่เหลี่ยมโดยใช้สมบัติต่างๆ
- การคาดคะเนพื้นที่ของรูปสี่เหลี่ยม
- โจทย์ปัญหาเกี่ยวกับความยาวรอบรูปของรูปสี่เหลี่ยม
- โจทย์ปัญหาเกี่ยวกับพื้นที่รูปสี่เหลี่ยม
- การหาความยาวรอบรูปวงกลม

- การหาพื้นที่รูปวงกลม
- โจทย์ปัญหาเกี่ยวกับวงกลม

#### 1.4.2 ขอบเขตด้านประชากร

กลุ่มตัวอย่างที่ใช้ในการทดลอง

นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2558 โรงเรียนปีคอนเฮาส์แย้มสอาดรังสิต

### 1.5 นิยามศัพท์

วิเคราะห์ หมายถึง การศึกษาค้นคว้าหรือการศึกษาหาคำตอบอย่างละเอียดรอบคอบ ตามกระบวนการวิเคราะห์ต่อประเด็นตามที่เกิดขึ้นเป็นการนำข้อมูลที่จัดทำไว้มารวบรวมวิเคราะห์โดยเลือกใช้สถิติให้เหมาะสมกับวัตถุประสงค์ของการวิเคราะห์ / วิจัย และลักษณะของข้อมูล สถิติ (Statistics) หมายถึง ศาสตร์แขนงหนึ่งซึ่งประกอบด้วยวิธีการเก็บรวบรวมข้อมูล การจัดหมวดหมู่ข้อมูลการวิเคราะห์ข้อมูล การแปลความหมายข้อมูล

คุณภาพ หมายถึง สิ่งที่เป็นไปตามมาตรฐาน ความต้องการ ความเหมาะสมกับวัตถุประสงค์ และสิ่งที่ปราศจากข้อบกพร่อง

ความยากง่าย (Difficulty) หมายถึง ค่าที่แสดงว่าเครื่องมือหรือแบบทดสอบนั้นมีความยากหรือความง่ายเพียงใด แบบทดสอบที่ดีควรมีค่าความยากง่ายปานกลาง

อำนาจจำแนก (Discrimination) หมายถึง ค่าที่แสดงว่า เครื่องมือหรือแบบประเมินนั้นสามารถแยกนักเรียนเก่งหรืออ่อนได้ โดยคนเก่งจะตอบถูก ส่วนคนอ่อนจะตอบผิด

ความเชื่อมั่น (Reliability) หมายถึง เครื่องมือนั้นมีผลการวัดที่สม่ำเสมอคงที่ จะวัดกี่ครั้งก็ได้ผลเหมือนเดิม นักเรียนที่ได้คะแนนสูงวัดกี่ครั้งก็จะได้คะแนนสูง นักเรียนที่ได้คะแนนต่ำวัดกี่ครั้งก็จะได้คะแนนต่ำ

### 1.6 ประโยชน์ของการวิจัย

1. ช่วยให้ทราบว่าข้อสอบแต่ละข้อมีคุณภาพอย่างไร ข้อใดใช้ได้หรือข้อใด ใช้ไม่ได้ ทำให้สามารถคัดเลือกข้อสอบที่มีคุณภาพดี นำไปใช้ประโยชน์ต่อไป
2. ช่วยในการพัฒนาปรับปรุงข้อสอบที่ใช้ไม่ได้ให้ใช้ได้ หรือข้อที่ใช้ได้ อยู่แล้วให้มีคุณภาพดียิ่งขึ้นอีกด้วย ในกรณีที่เป็นข้อสอบแบบเลือกตอบ คุณภาพของข้อสอบแต่ละข้อนอกจากจะขึ้นอยู่กับ การเขียนตัวคำถาม และตัวเลือกที่เป็นตัวถูกแล้ว ยังขึ้นอยู่กับประสิทธิภาพของตัวลวง แต่ละตัวด้วยว่าสามารถ ทำหน้าที่เป็นตัวลวงได้มากน้อยเพียงใด และลวงผู้เรียนประเภทไหนมากกว่ากัน ดังนั้นควรวิเคราะห์ คุณภาพ ของตัวลวงด้วย จะช่วยให้สามารถพัฒนาหรือปรับปรุงแก้ไขข้อสอบ ได้ตรงจุดยิ่งขึ้น
3. ช่วยเป็นแนวทางให้ครูเขียนข้อสอบได้ดีขึ้น เพราะผลการวิเคราะห์จะแสดงให้เห็นว่าข้อใดเขียนกำกวมหรือไม่ชัดเจน และจะช่วยให้ครูระมัด ระวังการใช้สำนวนภาษาที่ใช้ในการเขียนคำถามรวมทั้งตัวเลือกที่เป็นตัวถูกและตัวลวงด้วย

4. ชี้ให้เห็นถึงจุดอ่อนที่ผู้สอนหรือผู้เรียนต้องปรับปรุงแก้ไข เพื่อพัฒนาความสามารถและทักษะของผู้เรียนให้เป็นตามที่คาดหวัง
5. เป็นพื้นฐานสำคัญในการปรับปรุงพัฒนาการเรียนการสอน โดยเฉพาะอย่างยิ่งด้านพฤติกรรมการเรียนรู้ที่เหมาะสมสำหรับเนื้อหาวิชานั้นๆ
6. ทำให้สามารถคัดเลือกข้อสอบที่มีคุณภาพมารวมเป็นฉบับข้อสอบที่ดี สามารถจัดทำข้อสอบคู่ขนานและเป็นรากฐานสำคัญในการพัฒนาแบบสอบมาตรฐาน
7. นำข้อสอบที่ได้จากการวิเคราะห์ข้อสอบแล้วนำมาสะสมเป็นธนาคารข้อสอบ เพื่อความสะดวกต่อการใช้ข้อสอบในครั้งต่อไป

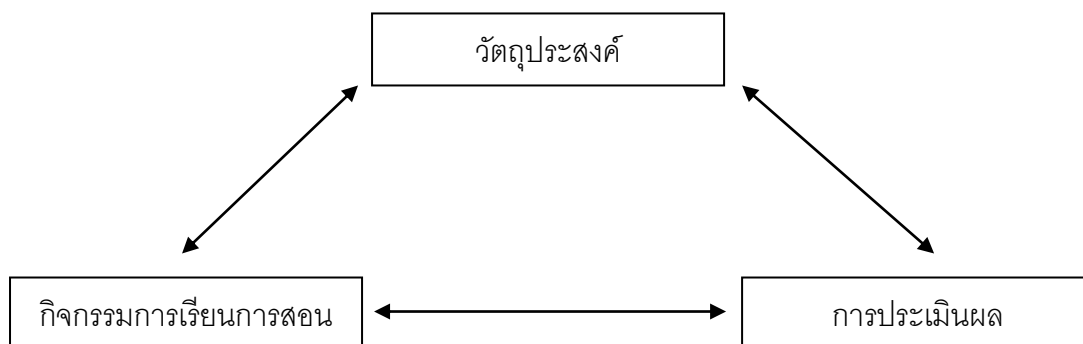
## บทที่ 2

### เอกสารและงานวิจัยที่เกี่ยวข้อง

ผู้ศึกษาได้ค้นคว้าเอกสารและทฤษฎีที่เกี่ยวข้องกับการวิเคราะห์ข้อสอบจากรายงานการวิเคราะห์ข้อสอบและเว็บไซต์ต่าง ๆ พอสรุปได้ดังนี้

#### 2.1 ความสำคัญของการวิเคราะห์ข้อสอบ ของ ดร.วรรณดี แสงประทีปทอง

กล่าวถึงความสำคัญของการวิเคราะห์ข้อสอบว่าเนื่องจากการจัดการศึกษาประกอบด้วยองค์ประกอบที่สำคัญ 3 ส่วน ได้แก่ วัตถุประสงค์ กิจกรรมการเรียน การสอน และการประเมินผล องค์ประกอบทั้งสามมีความสัมพันธ์กันดังนี้



วัตถุประสงค์เป็นเป้าหมายของการจัดการศึกษา และเป็นตัวบ่งชี้กิจกรรมการเรียนการสอน และการประเมินผล กิจกรรมการเรียนการสอนเป็นกระบวนการที่จะทำให้ผู้เรียนได้พัฒนาไปตามวัตถุประสงค์และเป็นแนวทางในการประเมินผล ส่วนการประเมินผลเป็นเครื่องชี้วัดผลที่เกิดขึ้นในตัวผู้เรียน ซึ่งเกิดจากการจัดกิจกรรมการเรียนการสอนตามวัตถุประสงค์ และเป็นเครื่องชี้วัดความเหมาะสมของการจัดกิจกรรม ดังนั้น การประเมินผลจึงมีความสำคัญอย่างมากต่อการจัดการศึกษา

ในการประเมินผล ผู้สอนจะใช้เครื่องมือชนิดต่าง ๆ เพื่อวัดความรู้ความสามารถที่เกิดขึ้นในตัวผู้เรียน เครื่องมือที่สร้างขึ้นนี้ ถ้ามีสมบัติของเครื่องมือวัดที่ดี จะทำให้การวัดความรู้ ความสามารถของผู้เรียนทำได้ อย่างถูกต้อง ตรงกับความสามารถที่แท้จริง เครื่องมือวัดที่นิยมใช้มากที่สุด คือ แบบทดสอบ โดยทั่วไป แบบทดสอบที่ใช้เป็นแบบทดสอบที่ครูผู้สอนสร้างขึ้น ดังนั้น จึงจำเป็นต้องมีการตรวจสอบคุณภาพของ เครื่องมือวัด ซึ่งเรียกว่าการวิเคราะห์ข้อสอบ เพื่อนำผลการวิเคราะห์มาใช้เป็นข้อมูลในการปรับปรุงข้อสอบให้ มีคุณภาพ และสามารถวัดความสามารถที่แท้จริงของผู้เรียนได้

## 2.2 ทฤษฎีเกี่ยวกับการวิเคราะห์ข้อสอบ

การวิเคราะห์ข้อสอบ วิเคราะห์ได้ 2 ลักษณะ คือ วิเคราะห์ข้อสอบเป็นรายข้อ และวิเคราะห์ข้อสอบทั้งฉบับ วิธีการวิเคราะห์สามารถทำได้หลายแนวทาง ขึ้นกับทฤษฎีการวัดผลที่นำมาใช้ทฤษฎีการวัดผลที่นำมาใช้โดยทั่วไปในการวิเคราะห์ข้อสอบ ได้แก่ ทฤษฎีการทดสอบแบบคลาสสิก

หรือเรียกว่า ทฤษฎีการทดสอบแบบดั้งเดิม หรือทฤษฎีการทดสอบแบบประเพณีนิยม (Classical Test Theory ; CTT) และอีกทฤษฎีหนึ่งคือทฤษฎีการตอบสนองข้อสอบ (Item Response Theory ; IRT)

2.2.1 ทฤษฎีการทดสอบแบบคลาสสิก ตามทฤษฎีนี้การวิเคราะห์ข้อสอบเป็นรายข้อจะวิเคราะห์ค่าความยาก (Item difficulty) ค่าอำนาจจำแนก (Discrimination Power) และประสิทธิภาพของตัวลวง (Effectiveness Of distracters) ส่วนการวิเคราะห์ข้อสอบทั้งฉบับจะวิเคราะห์ค่าความตรงหรือความเที่ยงตรง (Validity) ค่าความเที่ยงหรือความเชื่อมั่น (Reliability) ของแบบทดสอบ

2.2.2 ทฤษฎีการตอบสนองข้อสอบ ตามทฤษฎีนี้ การวิเคราะห์ข้อสอบเป็นรายข้อจะวิเคราะห์ค่าความยาก ค่าอำนาจจำแนก ค่าความน่าจะเป็นของการเดาถูก ค่าฟังก์ชันสารสนเทศของข้อสอบ (Item information function) ส่วนการวิเคราะห์ข้อสอบทั้งฉบับจะวิเคราะห์ค่าฟังก์ชันสารสนเทศของแบบทดสอบ (Test information function) และความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ ซึ่งถ้าความคลาดเคลื่อนมาตรฐานของการประมาณค่ามีค่าต่ำ แสดงว่าการประมาณค่าความสามารถของผู้เรียนมีความแม่นยำสูง

สำหรับข้อสอบที่ใช้ในการวัดผล ระหว่างเรียนหรือใช้ในการวัดผลปลายภาคเรียนนั้น ส่วนใหญ่เป็นข้อสอบที่ครูผู้สอนสร้างขึ้น และยังไม่ได้พัฒนาเป็นแบบทดสอบมาตรฐาน ดังนั้นในการวิเคราะห์ข้อสอบจึงควรวิเคราะห์ตามแนวทฤษฎีการทดสอบแบบคลาสสิก ซึ่งจากการวิเคราะห์จะได้ทราบประสิทธิภาพของตัวลวงด้วย ทำให้สามารถปรับปรุงข้อสอบได้ทั้งตัวคำถาม คำตอบถูกและตัวลวง อันนำไปสู่การพัฒนาข้อสอบให้เป็นมาตรฐาน นอกจากนี้การวิเคราะห์ข้อสอบตามแนวทฤษฎีการตอบสนองข้อสอบ เหมาะที่จะใช้วิเคราะห์แบบทดสอบที่ได้พัฒนาแล้ว ดังนั้นในการกล่าวถึงการวิเคราะห์ข้อสอบต่อไปนี้จะกล่าวเฉพาะการวิเคราะห์ข้อสอบตามแนวทฤษฎีการทดสอบแบบคลาสสิก

ตามทฤษฎีการทดสอบแบบคลาสสิกนั้น เชื่อว่าคะแนนที่ได้จากการทดสอบแต่ละครั้ง (X) ประกอบด้วยคะแนนความสามารถที่แท้จริง (T) ของผู้สอบ และความคลาดเคลื่อนในการวัด (E) ซึ่งสามารถเขียนเป็นสมการทางคณิตศาสตร์ได้ดังนี้

$$X = T + E$$

เนื่องจากการทดสอบแต่ละครั้ง จะเกิดความคลาดเคลื่อนขึ้นได้ไม่มากนักน้อย แต่เป้าหมายของการวัดต้องการให้คะแนนที่วัดหรือทดสอบได้มีค่าใกล้เคียงคะแนนความสามารถที่แท้จริงของผู้สอบมากที่สุด ดังนั้น จึงต้องพยายามหาวิธีการต่าง ๆ เพื่อให้การวัดเกิดความคลาดเคลื่อนน้อยที่สุดเท่าที่จะทำได้

## 2.3 การวิเคราะห์ข้อสอบรายข้อ

การวิเคราะห์ข้อสอบรายข้อ เป็นการตรวจสอบคุณภาพของข้อสอบแต่ละข้อ โดยพิจารณาจากสมบัติที่สำคัญ 3 ประการ ได้แก่ ความยาก อำนาจจำแนก และประสิทธิภาพของตัวลง

### 2.3.1 ความยากของข้อสอบ

ความหมายของความยาก และการคำนวณค่าความยาก ความยากของข้อสอบ คือ สัดส่วนหรือร้อยละของผู้สอบที่ตอบข้อสอบข้อนั้นถูก

ค่าความยากคำนวณได้จากสูตร

$$\text{ความยาก (P)} = \frac{\text{จำนวนผู้สอบที่ตอบตัวเลือกนั้น}}{\text{จำนวนผู้สอบ}}$$

$$\text{หรือความยาก (P)} = \frac{\text{จำนวนผู้สอบที่ตอบตัวเลือกนั้น} \times 100}{\text{จำนวนผู้สอบทั้งหมด}}$$

$$\text{หรือ (P)} = \frac{H + L}{N + N}$$

เมื่อ H คือ จำนวนผู้สอบในกลุ่มที่ได้คะแนนรวมสูงที่ตอบตัวเลือกนั้น

เมื่อ L คือ จำนวนผู้สอบในกลุ่มที่ได้คะแนนรวมต่ำที่ตอบตัวเลือกนั้น

เมื่อ  $N_H$  คือ จำนวนคนในกลุ่มที่ได้คะแนนรวมสูง

เมื่อ  $N_L$  คือ จำนวนคนในกลุ่มที่ได้คะแนนรวมต่ำ

ความยากของข้อสอบเปรียบเทียบกับความชันของเนิน ข้อสอบข้อใดยากมาก ก็เสมือนกับเนินนั้นชันมาก เด็กปีนไม่ค่อยไหว ทำผิดมาก จัดเป็นข้อสอบที่ยากมากและในทำนองเดียวกัน ถ้าข้อสอบข้อใดมีผู้ทำถูกมาก แสดงว่าเป็นข้อสอบที่ง่าย หรือมีระดับความยากต่ำ

#### 2.3.1.1 ลักษณะของความยาก ความยากของข้อสอบมีลักษณะดังนี้

1. ค่าความยากของข้อสอบในรูปสัดส่วน (P) มีค่าอยู่ระหว่าง 0 ถึง 1 และค่าความยากของข้อสอบในรูปร้อยละ (P) มีค่าอยู่ระหว่าง 0 ถึง 100
2. ข้อสอบข้อใดที่ผู้สอบทำถูกหมดทุกคน แสดงว่าข้อสอบง่ายมาก มีค่าความยาก (P) เท่ากับ 1.0 หรือ p เท่ากับ 100

2.3.1.2 ข้อสอบข้อใดที่ผู้สอบทำผิดหมดทุกคน แสดงว่าข้อสอบยากมาก มีค่าความยาก (P) เท่ากับ 0 หรือ P เท่ากับ 0

2.3.1.3 ข้อสอบข้อใดที่ผู้สอบครึ่งหนึ่งทำถูกและผู้สอบอีกครึ่งหนึ่งทำผิด แสดงว่าเป็นข้อสอบที่ยากปานกลาง หรือยากพอเหมาะ



2.3.1.4 ข้อสอบข้อใดที่มีค่าความยากสูง แสดงว่าข้อสอบข้อนั้นมีผู้ตอบถูกจำนวนมากจึงเป็นข้อสอบที่ง่าย

2.3.1.5 ข้อสอบข้อใดที่มีค่าความยากต่ำ แสดงว่า ข้อสอบข้อนั้นมีผู้ตอบถูกจำนวนน้อย จึงเป็นข้อสอบที่ยาก

#### เกณฑ์การแปลความหมายของความยากของข้อสอบ

ระดับความยาก (P)	ความหมาย
0.81 – 1.00	ง่ายมาก
0.61 – 0.80	ง่าย
0.51 – 0.60	ค่อนข้างง่าย
0.50	ยากง่ายพอเหมาะ
0.40 – 0.49	ค่อนข้างยาก
0.20 – 0.39	ยาก
0.0– 0.19	ยากมาก

#### 2.3.2 อำนาจจำแนกของข้อสอบ

##### 2.3.2.1 ความหมายของอำนาจจำแนก และการคำนวณค่าอำนาจจำแนก

อำนาจจำแนกของข้อสอบหมายถึง ความสามารถของข้อสอบที่จะจำแนกความแตกต่างของสิ่งที่ต้องการวัด โดยสามารถจำแนกกลุ่มผู้สอบที่ได้คะแนนรวมสูงออกจากกลุ่มผู้สอบที่ได้คะแนนรวมต่ำ การคำนวณค่าอำนาจจำแนกอาจคำนวณโดยใช้สูตรอย่างง่าย ( $r$ ) และสูตรสัมประสิทธิ์สหสัมพันธ์แบบพอยท์ ไบซีเรียล (Point biserial correlation coefficient ;  $r_{pb}$ )

1. การคำนวณค่าอำนาจจำแนกโดยใช้สูตรอย่างง่าย ( $r$ ) เป็นการเปรียบเทียบจำนวนผู้สอบในกลุ่มที่ได้คะแนนรวมสูง กับกลุ่มที่ได้คะแนนรวมต่ำที่ตอบข้อสอบข้อนั้นถูก การคำนวณใช้สูตร

$$r = H - L_{n_H} \quad \text{หรือ} \quad r = H - L_{n_L}$$

โดยที่สัญลักษณ์แต่ละตัวมีความหมายเหมือนในการคำนวณค่าความยาก

2. การคำนวณค่าอำนาจจำแนกโดยใช้สูตรสัมประสิทธิ์สหสัมพันธ์แบบพอยท์ไบซีเรียล การหาค่าอำนาจจำแนกโดยวิธีนี้ใช้ในกรณีที่มีการกระจายของคะแนนรวม หรือการกระจายคะแนนของผู้สอบที่ตอบข้อสอบถูก หรือการกระจายคะแนนของผู้สอบที่ตอบข้อสอบผิดไม่เป็นโค้งปกติ การคำนวณค่า  $r_{pb}$  ใช้สูตร

$$r_{pb} = \frac{X\bar{p} - X\bar{q} \cdot \sqrt{pq}}{S_x}$$

เมื่อ  $X\bar{p}$  = คือค่าเฉลี่ยของคะแนนของผู้สอบ ที่ตอบข้อสอบข้อนั้นถูก

เมื่อ  $X\bar{q}$  = คือค่าเฉลี่ยของคะแนนของผู้สอบที่ตอบข้อสอบข้อนั้นผิด

เมื่อ  $S_x$  = คือค่าเบี่ยงเบนมาตรฐานของคะแนนรวมของผู้สอบทั้งหมด

เมื่อ  $p$  = คือสัดส่วนของผู้สอบที่ตอบข้อสอบข้อนั้นถูก

เมื่อ  $q$  = คือสัดส่วนของผู้สอบที่ตอบข้อสอบข้อนั้นผิด

### 2.3.3 ลักษณะของอำนาจจำแนก

1. อำนาจจำแนกมีค่าตั้งแต่ -1.00 ถึง 100
2. ข้อสอบข้อใดที่ผู้สอบในกลุ่ม ได้คะแนนรวมสูง ทำถูกทุกคนแต่ผู้สอบในกลุ่มที่ได้คะแนนรวมต่ำทำผิดทุกคน อำนาจจำแนกมีค่าเท่ากับ 1 ข้อสอบข้อนั้นเป็นข้อสอบที่มีอำนาจจำแนกดีเลิศ
3. ข้อสอบข้อใดที่ผู้สอบในกลุ่มที่ได้คะแนนรวมสูงทำผิดทุกคน แต่ผู้สอบในกลุ่มที่ได้คะแนนรวมต่ำทำถูกทุกคน อำนาจจำแนกมีค่าเท่ากับ -1 ข้อสอบข้อนั้นเป็นข้อสอบที่ไม่ดี
4. ข้อสอบข้อใดที่ผู้สอบในกลุ่มที่ได้คะแนนรวมสูงตอบถูกเท่า ๆ กับผู้สอบในกลุ่มที่ได้คะแนนรวมต่ำ อำนาจจำแนกจะเท่ากับ 0 หรือมีค่าใกล้ศูนย์ ข้อสอบข้อนั้นมีอำนาจจำแนกต่ำ
5. โดยทั่วไปการสร้างข้อสอบต้องการข้อสอบที่มีค่าอำนาจจำแนกเป็นบวก และถ้าอำนาจจำแนกที่เป็นบวกยิ่งมีค่ามากก็ยิ่งดี

### เกณฑ์การแปลความหมายค่าอำนาจจำแนกที่เป็นบวก

อำนาจจำแนก	ความหมาย
1.00	จำแนกได้ดีเลิศ
0.80 – 0.99	จำแนกได้ดีมาก
0.60 – 0.79	จำแนกได้ดี
0.40 – 0.59	จำแนกได้ปานกลาง
0.20 – 0.49	จำแนกได้เล็กน้อย
ต่ำกว่า .19	จำแนกไม่ได้เลย

## 2.4 ประสิทธิภาพของตัวลวง

ตัวเลือกที่เป็นตัวลวง เป็นส่วนประกอบที่สำคัญของข้อสอบ ตัวลวงที่มีประสิทธิภาพต้องสามารถลวงผู้ตอบที่ไม่มีความสามารถได้ การวิเคราะห์ประสิทธิภาพของตัวลวง ทำได้โดยวิเคราะห์ความยาก

และอำนาจจำแนกของตัวลงแต่ละตัว

การคำนวณความยากของตัวลง ทำโดยวิธีการเดียวกับการคำนวณความยากของตัวถูก โดยตัวลงที่ติดจะต้องเป็นตัวลงที่มีผู้เลือกตอบบ้าง ในทางปฏิบัติใช้ตัวลงที่มีผู้เลือกตอบอย่างน้อยร้อยละ 5

อำนาจจำแนกของตัวลง คำนวณได้จากสูตร

$$r = \frac{L-H}{n_H} \quad \text{หรือ} \quad r = \frac{L-H}{n_L}$$

$$\text{หรือ } r = P_L - P_H$$

เมื่อ  $P_L$  คือ สัดส่วนของผู้สอบในกลุ่มที่ได้คะแนนรวม ต่ำ ที่ตอบตัวเลือกนั้น

$P_H$  คือ สัดส่วนของผู้สอบในกลุ่มที่ได้คะแนนรวม สูง ที่ตอบตัวเลือกนั้น

การวิเคราะห์ข้อสอบแบบอิงกลุ่มมีหลายวิธี นอกจากวิธีคำนวณสัมประสิทธิ์สหสัมพันธ์แบบพอย์ไบซีเรียล และวิธีใช้สูตรอย่างง่ายแล้ว ยังสามารถวิเคราะห์โดยใช้ตารางวิเคราะห์ข้อทดสอบของ จุง เต ฟาน (Chung – The - Fan) ตารางนี้จะช่วยประเมินค่าความยากและอำนาจจำแนกของข้อสอบแต่ละข้อ (ทั้งตัวถูกและตัวลง) ถ้าทราบร้อยละของผู้สอบในกลุ่มที่มีคะแนนรวมสูง และผู้สอบที่มีคะแนนรวมต่ำที่ตอบข้อสอบแต่ละข้อ การใช้ตารางวิเคราะห์ข้อทดสอบของ จุง เต ฟาน นี้มีข้อตกลงว่า การแจกแจงคะแนนของกลุ่ม ผู้สอบที่นำมาวิเคราะห์ต้องมีลักษณะเป็นโค้งปกติ และใช้เทคนิคร้อยละ 27 คือ นำข้อมูลผลการสอบของกลุ่ม ผู้สอบที่มีคะแนนรวมสูง ร้อยละ 27 และกลุ่มผู้สอบที่มีคะแนนรวมต่ำ ร้อยละ 27 มาวิเคราะห์ในทางปฏิบัติ เทคนิค 27% นี้จะใช้สำหรับการสอบที่มีผู้สอบจำนวนมาก เพื่อให้ได้การแจกแจงคะแนนของผู้สอบตามข้อตกลงเบื้องต้นซึ่งมักถือเกณฑ์ว่าผู้สอบในแต่ละกลุ่มที่นำมาวิเคราะห์ควรมีกลุ่มละประมาณ 100 คน นั่นคือจำนวนผู้สอบทั้งหมดจะต้องมีตั้งแต่ 470 คนขึ้นไป และในกรณีที่ผู้สอบมีจำนวนน้อยก็ไม่นิยมใช้เทคนิค 27% เนื่องจากอาจทำให้ผลการวิเคราะห์ที่คลาดเคลื่อนได้

โดยทั่วไปแบบทดสอบวัดผลสัมฤทธิ์ที่ครูผู้สอนสร้างขึ้นใช้ในโรงเรียนมักใช้กับผู้สอบจำนวนไม่มากนัก การวิเคราะห์ข้อสอบอาจวิเคราะห์โดยใช้สูตรอย่างง่ายดังกล่าวไปแล้ว การพิจารณาตัดสินข้อสอบว่ามีคุณภาพ นิยมใช้เกณฑ์ว่า ถ้าข้อสอบมีค่าความยาก ระหว่าง 0.20 ถึง 0.80 และมีค่าอำนาจจำแนกตั้งแต่ .20 ขึ้นไป จึงจะถือว่าเป็นข้อสอบที่มีคุณภาพ ขั้นตอนการวิเคราะห์ข้อสอบโดยใช้สูตรอย่างง่ายจะกล่าวถึงในหัวข้อต่อไป

2.4.1 ขั้นตอนการวิเคราะห์ข้อสอบโดยใช้สูตรอย่างง่าย เมื่อนำแบบทดสอบแบบเลือกตอบที่สร้างขึ้น ไป ทดสอบและนำกระดาษคำตอบมาตรวจให้คะแนนแล้ว การวิเคราะห์ข้อสอบโดยใช้สูตรอย่างง่ายมีขั้นตอน ดังนี้

2.4.1.1 เรียงกระดาษคำตอบของผู้สอบแต่ละคน จากผู้ที่ได้คะแนนรวมมากไปหาน้อย สมมติว่ามีกระดาษคำตอบ 20 คน

2.4.1.2 แบ่งกระดาษคำตอบ เป็นกลุ่มผู้ที่ได้คะแนนรวมสูง (หรือเรียกว่ากลุ่มสูง) และกลุ่มผู้ที่ได้คะแนนรวมต่ำ (เรียกว่ากลุ่มต่ำ) การแบ่งกลุ่มนี้อาจแบ่ง 2 ส่วน เป็นครึ่งต่อครึ่ง หรือแบ่ง 4 ส่วน โดยใช้ 44% เป็นเกณฑ์แบ่งหรือแบ่งด้วยเกณฑ์ 27% โดยกรณี que แบ่งมากกว่า 2 ส่วน จะนำเฉพาะผลสอบของผู้สอบในกลุ่มและกลุ่มต่ำมาวิเคราะห์ สมมติว่าในกรณีตัวอย่างนี้แบ่งกระดาษคำตอบออกเป็น 2 ส่วน ได้จำนวนผู้สอบกลุ่มสูง และกลุ่มต่ำ กลุ่มละ 10 คน

2.4.1.3 นำกระดาษคำตอบมาแจกแจงความถี่ของการตอบในแต่ละตัวเลือก สมมติว่าการแจกแจงความถี่ของการตอบข้อสอบข้อที่ 1 เป็นดังนี้

ข้อที่	ตัวเลือก	กลุ่ม (H)	กลุ่มต่ำ (L)
1	ก	1	0
	ข	1	4
	(ค)	6	4
	ง	2	2
	จ	0	0
	เว้น	-	-

2.4.1.4 คำนวณค่าความยาก (P) ของตัวถูกและตัวลวง จากสูตร

$$p = \frac{H+L}{n_H + n_L}$$

คำนวณค่า อำนาจจำแนกของข้อสอบจากสูตร

$$r = \frac{H-L}{n_H} \quad \text{หรือ} \quad r = \frac{H-L}{n_L}$$

อำนาจจำแนกของตัวลง

$$r = \frac{L-H}{n_H} \quad \text{หรือ} \quad r = \frac{L-H}{n_L}$$

การคำนวณความยากและอำนาจจำแนก แสดงดังตาราง

ข้อที่	กลุ่มสูง (H)	กลุ่มต่ำ (L)	P	r
1	1	0	$\frac{1+0}{20} = 0.05$	$\frac{0-1}{10} = .10$
	6	4	$\frac{1+4}{20} = .25$	$\frac{4-1}{10} = .40$
	6	4	$\frac{6+4}{20} = .50$	$\frac{6-4}{10} = .20$
	2	2	$\frac{2+2}{20} = .20$	$\frac{2-2}{10} = 0$
	0	-	$\frac{0+0}{20} = 0$	$\frac{0-0}{10} = 0$

เมื่อพิจารณาความยากและอำนาจจำแนกของข้อสอบ ข้อที่ 1 สรุปได้ว่า

1. ตัวเลือก ค เป็นตัวเลือกที่ถูก มีค่าความยากพอเหมาะ ( $p = .50$ ) โดยสามารถจำแนกผู้สอบได้เล็กน้อย ( $r = .20$ ) ตัวเลือกนี้ใช้ได้
  2. ตัวลง ก แม้จะลงผู้สอบได้บ้าง ( $p = .05$ ) แต่ลงผู้สอบในกลุ่มสูง มากกว่ากลุ่มต่ำ อำนาจจำแนกมีค่าลบ ตัวลงนี้ใช้ไม่ได้
  3. ตัวลง ข ลงผู้สอบได้บ้าง ( $p = .25$ ) และลงผู้สอบในกลุ่มต่ำ มากกว่ากลุ่มสูง อำนาจจำแนก มีค่าเป็นบวก ตัวลงนี้ใช้ไม่ได้
  4. ตัวลง ง แม้จะลงผู้สอบได้ ( $p = .20$ ) แต่ไม่มีอำนาจจำแนกผู้สอบในกลุ่มสูง และกลุ่มต่ำ ตัวเลือกนี้ใช้ไม่ได้
  5. ตัวลง จ ลงผู้สอบไม่ได้ และไม่มีอำนาจจำแนก ตัวเลือกนี้ใช้ไม่ได้
- ดังนั้น ข้อสอบข้อที่ 1 เป็นข้อสอบที่ใช้ได้ แต่ควรปรับปรุงตัวลง ข ง และ จ

โดยสรุปแล้ว การวิเคราะห์ห้ข้อสอบเป็นการคำนวณค่าทางสถิติ เพื่อยืนยันคุณภาพข้อสอบที่สร้างขึ้น ข้อสอบจะมีคุณภาพดีหรือไม่ ไม่ได้อยู่ที่ค่าสถิติจากการวิเคราะห์ข้อสอบเพียงอย่างเดียว ข้อสอบที่ดีต้องวัดตรงกับวัตถุประสงค์ที่ต้องการ มีลักษณะข้อสอบที่ดี และเมื่อนำไปทดลองสอบแล้วนำผลมาวิเคราะห์ได้ ค่าสถิติอยู่ในเกณฑ์ใช้ได้ จึงจะถือว่าเป็นข้อสอบที่มีคุณภาพ

## 2.5 การวิเคราะห์ข้อสอบทั้งฉบับ

การวิเคราะห์ข้อสอบทั้งฉบับ เป็นการตรวจสอบคุณภาพของเครื่องมือวัด คุณสมบัติที่สำคัญมากของเครื่องมือวัดใด ๆ มี 2 ประการ คือ

1. ความตรงหรือความเที่ยงตรง (Validity)
2. ความเที่ยงหรือความเชื่อมั่น (Reliability)

## 2.6 ความตรงหรือความเที่ยงตรง

ความตรง หมายถึง ความสามารถของเครื่องมือวัด ที่สามารถวัดได้ในสิ่งที่ต้องการวัด เป็นความสอดคล้องระหว่างผลการวัด กับสิ่งที่ต้องการวัด ความตรงที่ใช้ในการทดสอบจำแนกเป็น 3 ชนิด ได้แก่

1. ความตรงตามเนื้อหา
2. ความตรงตามโครงสร้าง
3. ความตรงตามเกณฑ์ที่เกี่ยวข้อง

แบบทดสอบวัดผลสัมฤทธิ์ จะเกี่ยวข้องกับความตรงตามเนื้อหา มากกว่าความตรงชนิดอื่น ๆ

1. ความตรงตามเนื้อหา หมายถึง ความสอดคล้องของเนื้อหาสาระของข้อสอบ กับเนื้อหาวิชาที่สอน หรือข้อสอบสอดคล้องและครอบคลุมเนื้อหา และวัตถุประสงค์ของวิชาที่สอน

การตรวจสอบความตรงตามเนื้อหา ทำได้โดยการวิเคราะห์เนื้อหาของแบบทดสอบทั้งฉบับโดยอาศัยผู้เชี่ยวชาญเป็นผู้พิจารณาว่าเนื้อหาสาระของแบบทดสอบ สอดคล้องกับแบบเรียนรายละเอียดของวิชา และหลักสูตรหรือไม่ ซึ่งในทางปฏิบัติจะใช้ตารางวิเคราะห์หลักสูตรเป็นตัวเทียบ

2. ความตรงตามโครงสร้าง หมายถึง ความสามารถของแบบทดสอบที่จะวัดคุณลักษณะหรือพฤติกรรมตามโครงสร้างทฤษฎีได้ การตรวจสอบความตรงตามโครงสร้าง มีหลายวิธีเช่น

2.1 การเทียบกลุ่มอ้างอิง (Known group) วิธีการนี้จะนำแบบทดสอบที่สร้างขึ้นไปใช้กับกลุ่มที่ทราบคุณลักษณะทางจิตวิทยาตามที่ต้องการวัดโดยใช้ 2 กลุ่ม ที่มีลักษณะตรงข้ามกันแล้วทดสอบความแตกต่างของค่าที่วัดได้จากกลุ่มทั้งสอง ถ้าความแตกต่างมีนัยสำคัญเชิงสถิติ แสดงว่าแบบทดสอบนั้นมีความตรงตามโครงสร้าง

2.2 การวิเคราะห์ตัวประกอบ (Factor Analysis) โดยอาศัยวิธีการทางสถิติสำหรับตรวจหาคุณภาพทางจิตวิทยา ด้วยการวิเคราะห์ความสัมพันธ์ระหว่างข้อสอบแต่ละข้อว่า ข้อสอบทั้งหมดนั้นวัดองค์ประกอบอะไรบ้าง ถ้าตรงตามทฤษฎีหรือสมมุติฐานที่ตั้งไว้ แสดงว่าแบบทดสอบมีความตรงตามโครงสร้าง

2.3 การหาความสัมพันธ์กับเกณฑ์ที่มีโครงสร้างเหมือนกัน วิธีนี้ทำโดยหาสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากแบบทดสอบที่เราสร้างขึ้นกับแบบทดสอบอื่นที่วัดในโครงสร้าง หรือทฤษฎีเดียวกัน ซึ่งพิสูจน์ไว้แล้วว่ามีความตรงตามโครงสร้าง ถ้าแบบทดสอบที่สร้างขึ้นใหม่มีสหสัมพันธ์กับแบบทดสอบที่เป็นเกณฑ์สูง แสดงว่าแบบทดสอบที่สร้างขึ้นมีความตรงตามโครงสร้าง

3. ความตรงตามเกณฑ์ที่เกี่ยวข้อง เป็นการพิจารณาความสัมพันธ์ระหว่างแบบทดสอบที่สร้างขึ้นกับเกณฑ์ภายนอกบางอย่าง ซึ่งเป็นสภาพความเป็นจริงที่ได้จากการปฏิบัติงานความตรงตามเกณฑ์ที่เกี่ยวข้อง แบ่งออกเป็น 2 ประเภทดังนี้

3.1 ความตรงตามสภาพ (Concurrent Validity) เป็นความสามารถของแบบทดสอบที่วัดได้ตรงกับสภาพความเป็นจริงของบุคคลในขณะนั้น เช่น แบบทดสอบวัดผลสัมฤทธิ์วิชาภาษาไทยนำไปให้ผู้เรียนคนหนึ่งสอบ ปรากฏว่าได้คะแนนสูง ซึ่งในสภาพความเป็นจริงผู้เรียนมีความสามารถทางภาษาไทยสูงจริง แสดงว่าแบบทดสอบวัดได้ตรงตามสภาพที่เป็นอยู่

3.2 ความตรงตามการพยากรณ์ (Predictive Validity) เป็นความสามารถของแบบทดสอบที่วัดผลได้ตรงกับสภาพความเป็นจริงที่เกิดขึ้นในอนาคต เช่น แบบทดสอบคัดเลือกบุคคลเข้าศึกษาต่อเมื่อนำไปใช้สอบคัดเลือกบุคคลเข้าศึกษา ปรากฏว่าผู้ที่สอบคัดเลือกได้คะแนนดี เมื่อเข้าศึกษามีผลการเรียนอยู่ในเกณฑ์ดี แสดงว่าแบบทดสอบมีความตรงตามการพยากรณ์

## 2.7 ความเที่ยงหรือความเชื่อมั่น

ความเที่ยง หมายถึง ความคงที่ของคะแนนที่วัดได้แต่ละครั้ง วิธีการหาค่าความเที่ยงของแบบทดสอบทำได้หลายวิธี คือ

1. วิธีสอบซ้ำ
2. วิธีแบบทดสอบคู่ขนาน
3. วิธีหาความสอดคล้องภายในแบ่งเป็น
  - 3.1 วิธีแบ่งครึ่งแบบทดสอบ
  - 3.2 วิธีหาจากสูตรคูเดอร์และริชาร์ดสัน
  - 3.3 วิธีหาจากสูตรสัมประสิทธิ์แอลฟา

### 1. วิธีสอบซ้ำ

การหาความเชื่อมั่นโดยวิธีสอบซ้ำ เป็นการหาความสัมพันธ์ของคะแนนจากการทำแบบทดสอบฉบับเดียวกันสองครั้ง โดยทิ้งช่วงห่างให้เหมาะสม (ประมาณ 2 สัปดาห์) การหาความเชื่อมั่นโดยวิธีนี้เป็น การตรวจสอบความคงที่ของการแสดงออกของผู้สอบสองครั้งว่า จะมีความคงที่หรือไม่ วิธีการนี้มีจุดอ่อนที่ความแปรเปลี่ยนภายในตัวผู้สอบในระหว่างทิ้งช่วงการสอบ ดังนั้น การหาความเชื่อมั่นโดยวิธีนี้ควรนำไปใช้กับแบบทดสอบวัดคุณลักษณะที่ค่อนข้างจะคงที่ไม่แปรเปลี่ยนโดยง่าย

## 2. วิธีใช้แบบทดสอบคู่ขนาน

การหาความเชื่อมั่นโดยใช้วิธีแบบทดสอบคู่ขนาน เป็นการหาความสัมพันธ์ของคะแนนจากการนำแบบทดสอบ 2 ฉบับ ที่เทียบเท่ากันไปสอบกับบุคคลกลุ่มเดียวกัน วิธีการนี้มีจุดอ่อนที่ความเป็นคู่ขนานกันของแบบทดสอบ 2 ฉบับ ซึ่งสร้างได้ยาก

## 3. วิธีหาความสอดคล้องภายใน

### 3.1 วิธีแบ่งครึ่งแบบทดสอบ

การหาความเที่ยงโดยวิธีนี้ เป็นการหาความสัมพันธ์ของคะแนนจากการใช้แบบทดสอบฉบับเดียว และสอบเพียงครั้งเดียว โดยนำผลการสอบมาแบ่งเป็นข้อมูล 2 ชุด โดยอาจแบ่งเป็นข้อคู่ - ข้อคี่ แบ่งเป็นครึ่งฉบับแรก ครึ่งฉบับหลัง จากการหาค่าสัมประสิทธิ์สหสัมพันธ์จะได้สัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบครึ่งฉบับ แล้วจึงนำไปปรับขยายเป็นสัมประสิทธิ์สหสัมพันธ์ของแบบทดสอบทั้งฉบับ

จากสูตรของสเปียร์แมน บราวน์ (Spearman Brown) ดังนี้

$$r_{tt} = \frac{2r_{xy}}{1+r_{xy}}$$

$r_{tt}$  คือ ความเที่ยงตรงของแบบทดสอบทั้งฉบับ

$r_{xy}$  คือ ความเที่ยงของแบบทดสอบครึ่งฉบับ

### 3.2 วิธีหาจากสูตรของคูเดอร์และริชาร์ดสัน

การหาความเที่ยงโดยวิธีนี้ เป็นการหาความสัมพันธ์ของคะแนนจากการใช้แบบทดสอบฉบับเดียว และสอบเพียงครั้งเดียวโดยนำผลการสอบมาคำนวณค่าสัมประสิทธิ์ ใช้สูตรของคูเดอร์และริชาร์ดสัน สูตรที่ใช้มี 2 สูตร คือ สูตร KR-20 กับสูตร KR-21

สูตร KR-20

$$r_u = \frac{k}{K-1} \left( \frac{s_x^2 - \sum Pq}{s_v^2} \right)$$

เมื่อ  $k$  คือ จำนวนข้อของแบบทดสอบ

$s_x^2$  คือ ความแปรปรวนของคะแนนรวม

$p$  คือ สัดส่วนของผู้ตอบข้อสอบถูกในแต่ละข้อ

$q$  คือ สัดส่วนของผู้ตอบข้อสอบผิดในแต่ละข้อ ( $1 - p$ )

สูตร KR-21



$$r_u = \frac{k}{K-1} \left( 1 - \frac{\overline{X(k-x)}}{Ks^2} \right)$$

เมื่อ  $k$  คือ จำนวนข้อของแบบทดสอบ

$\bar{X}$  คือ ค่าเฉลี่ยของคะแนนรวม

$s_x^2$  คือ ความแปรปรวนของคะแนนรวม

สูตร KR-20 และ KR-21 นี้ใช้ได้เฉพาะการหาความเที่ยงของแบบทดสอบที่ให้คะแนนแต่ละข้อเป็นแบบ 0 กับ 1 เท่านั้น สูตร KR-21 ใช้ในกรณีข้อสอบทุกข้อมีค่าความยากเท่ากัน ซึ่งในทางปฏิบัติต้องพิจารณาเงื่อนไขที่เป็นจริงด้วย

### 3.3 วิธีหาสูตรสัมประสิทธิ์ แอลฟา

การหาค่าความเที่ยงโดยใช้สูตรของครอนบาค (Cronbach) นี้ปรับมาจากสูตร KR-20 ใช้หาความเที่ยงของเครื่องมือวัดที่ให้คะแนนแตกต่างกันไปในแต่ละข้อได้โดยไม่จำเป็นต้องเป็นระบบการให้คะแนนแบบ 1 กับ 0 สูตรการคำนวณเป็นดังนี้

$$\alpha = \frac{k}{k-1} \left( \frac{1 - \sum s_x^2}{s_t^2} \right)$$

เมื่อ  $\alpha$  คือ ค่าความเที่ยง

เมื่อ  $k$  คือ จำนวนข้อสอบ

$s_x^2$  คือ ความแปรปรวนของคะแนนแต่ละข้อ

$s_t^2$  คือ ความแปรปรวนของคะแนนทั้งฉบับ

การหาค่าความเที่ยงโดยใช้สูตรสัมประสิทธิ์แอลฟา สามารถหาได้โดยใช้ผลการสอบจากแบบทดสอบฉบับเดียว นำไปสอบกับบุคคลกลุ่มเดียว และนำไปใช้กันได้อย่างกว้างขวาง โดยไม่จำกัดเฉพาะแบบทดสอบที่ให้คะแนนแบบ 1 กับ 0 เนื่องจากแบบทดสอบที่ใช้ส่วนใหญ่เป็นแบบเลือกตอบ ซึ่งให้คะแนนเป็นเวลา 1 กับ 0 ดังนั้นในการเสนอตัวอย่างการคำนวณค่าความเที่ยงของแบบทดสอบ จะเสนอตัวอย่างการหาค่าเที่ยงโดยใช้สูตร KR-20 และ KR-21

**ตัวอย่าง** แบบทดสอบแบบเลือกตอบฉบับหนึ่ง มี 10 ข้อ ให้คะแนนการตอบแต่ละข้อ

โดยตอบถูกให้ 1 คะแนน ตอบผิดให้ 0 คะแนน นำไปสอบกับผู้เรียน 10 คน ได้คะแนนดังตาราง

จงหาความเที่ยงของแบบทดสอบ โดยใช้สูตร KR-20 และ KR-21

ผู้สอบคนที่	ข้อ1	ข้อ2	ข้อ3	ข้อ4	ข้อ5	ข้อ6	ข้อ7	ข้อ8	ข้อ9	ข้อ10	รวม (X)	X <sup>2</sup>
1	1	1	0	1	1	0	0	0	0	0	4	16
2	0	1	1	1	1	0	1	0	0	0	5	25
3	0	1	0	1	0	0	0	1	1	0	4	16
4	1	1	1	1	0	1	0	1	1	1	8	64
5	1	0	1	1	0	1	0	0	1	0	5	25
6	1	1	0	0	1	0	0	1	1	0	5	25
7	1	1	1	0	0	0	0	0	0	1	4	16
8	1	0	1	1	0	1	0	1	1	1	7	49
9	1	1	1	1	1	1	0	1	1	1	9	81
10	1	1	1	1	1	0	1	1	1	1	9	81
P	.8	.8	.7	.8	.5	.4	.2	.6	.7	.5	$\sum x = 60$	$\sum x^2 = 398$
Q	.2	.2	.4	.2	.5	.6	.8	.4	.4	.5		
Pq.	.16	.16	.21	.16	.25	.24	.16	.21	.21	.25		$\sum pq = 2.04$

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

$$\bar{x} = \frac{60}{10} = 6$$

$$S_x^2 = \frac{1}{N-1} [\sum X^2 - N\bar{x}^2]$$

จากสูตร KR-20

$$r_{tt} = \frac{k}{k-1} \left( S_{tt}^2 - \frac{\sum p_2 q}{S_x^2} \right)$$

$$= \frac{10}{10-1} \frac{(4.22 - 2.04)}{4.22}$$

$$= 0.57$$

จากสูตร KR – 21

$$\begin{aligned}
 r_{tt} &= \frac{k}{k-1} \left\{ X \frac{(\bar{k} - \bar{X})}{S_x^2} \right\} \\
 &= \frac{10-1}{9} \left\{ \frac{6(10-6)}{10(4.22)} \right\} \\
 &= 0.48
 \end{aligned}$$

## 2.8 การวิเคราะห์ข้อสอบโดยใช้โปรแกรมคอมพิวเตอร์

ในปัจจุบันมีการนำคอมพิวเตอร์มาช่วยในการวิเคราะห์ข้อมูลอย่างแพร่หลาย สำหรับการวิเคราะห์ ข้อสอบมีโปรแกรมที่พัฒนาขึ้นใช้หลายโปรแกรม ทั้งการวิเคราะห์ข้อสอบตามทฤษฎีการสอบแบบคลาสสิก และการวิเคราะห์ข้อสอบตามทฤษฎีการตอบข้อสอบ เช่น โปรแกรม BILOG โปรแกรม ITEM โปรแกรม IRT โปรแกรม SIRI เป็นต้น โปรแกรมทั้งหลายที่พัฒนาขึ้นส่วนใหญ่เป็นโปรแกรมสำเร็จรูป ซึ่งให้ความสะดวกต่อ ผู้ใช้เป็นอย่างมาก ผู้สนใจใช้โปรแกรมคอมพิวเตอร์ช่วยในการวิเคราะห์ข้อสอบสามารถศึกษาวิธีการใช้ได้จากโปรแกรมดังกล่าว

## บทที่ 3 วิธีดำเนินการศึกษา

ในการดำเนินการจัดทำข้อสอบและวิเคราะห์ข้อสอบของผู้วิจัยครั้งนี้ แบ่งการดำเนินการเป็น 8 ขั้นตอน คือ

1. การรวบรวมจัดทำข้อสอบ
2. การทดสอบ
3. การวิเคราะห์ข้อสอบ
4. ประชากร
5. เครื่องมือที่ใช้ในการวิจัย
6. การเก็บรวบรวมข้อมูล
7. การวิเคราะห์ข้อมูล
8. สถิติที่ใช้ในการวิเคราะห์ข้อมูล

### 3.1 การรวบรวมจัดทำข้อสอบ

3.1.1 ผู้วิจัยได้ทำการศึกษารายละเอียดของหลักสูตร จุดประสงค์ มาตรฐานรายวิชาและ คำอธิบายรายวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6

3.1.2 ศึกษาเนื้อหาสาระที่ใช้ในการสอนจากหนังสือและเอกสารต่าง ๆ ตลอดจนแผนการสอนรายวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6

3.1.3 ทำการออกข้อสอบ โดยคำนึงถึงจุดประสงค์รายวิชาและมาตรฐานรายวิชา ทั้งหมด 30 ข้อ

### การจัดการทดสอบ

กำหนดวันทดสอบ 1 วัน 1 ชั่วโมง โดยแจ้งให้นักเรียนทราบล่วงหน้า และแจ้งเนื้อหาสอบให้ทราบ

### 3.3 การวิเคราะห์ข้อสอบ

ใช้โปรแกรมวิเคราะห์ข้อสอบ ช่วยในการคำนวณและวิเคราะห์ข้อสอบ บ่งชี้คุณภาพของข้อสอบดังนี้

- ระดับความยากง่ายของข้อสอบ
- อำนาจจำแนกของข้อสอบ
- ประสิทธิภาพของตัวลวง

### 3.4 ประชากรและกลุ่มตัวอย่าง

- กลุ่มตัวอย่าง นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2559

### 3.5 เครื่องมือที่ใช้ในการวิจัย

ได้ใช้ข้อสอบกลางภาคเรียนที่ 2 ปีการศึกษา 2559 ของครูประจำวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 มาใช้

### 3.6 การเก็บรวบรวมข้อมูล

นำกระดาษคำตอบจากการทดสอบกลางภาคเรียนที่ 2 ปีการศึกษา 2559  
วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6

### 3.7 การวิเคราะห์ข้อมูล

นำคำตอบที่นักเรียนตอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 นำมาหาค่าความยากง่าย  
หาค่าอำนาจจำแนก เป็นรายข้อ และหาค่าความเชื่อถือได้ทั้งฉบับ โดยใช้โปรแกรมวิเคราะห์ข้อสอบ  
ของนายธีรวัฒน์ สุพัตกุล จังหวัดบุรีรัมย์

### 3.8 สถิติที่ใช้ในการวิเคราะห์ข้อมูล

- หาค่าความยากง่าย
- หาค่าอำนาจจำแนก
- หาค่าความเชื่อถือได้

## บทที่ 4

### ผลการวิเคราะห์ข้อมูล

ผู้วิจัยได้ทำการศึกษาเรื่อง ผลของการวิเคราะห์คุณภาพของข้อสอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 กลางภาคเรียนที่ 2 ประจำปีการศึกษา 2559 มีประชากรในการศึกษา คือ นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2559 จำนวนข้อสอบทั้งหมด 30 ข้อ วิเคราะห์ข้อสอบได้ดังนี้

**ตารางที่ 1 ผลของการวิเคราะห์คุณภาพข้อสอบวิชาประวัติศาสตร์ ชั้นประถมศึกษาปีที่ 1**

ข้อ	ค่าความยากง่าย (P)	ค่าอำนาจจำแนก (B-index)	เฉลย	แปลผล
1. สรุป				ดีมาก
ก	.16	-1	ง	ปรับตัวเลือก ข ค
ข	.00	.00		
ค	.16	.20		
ง	.66	.80		
2. สรุป				ดีมาก
ก	.83	1	ก	ปรับตัวเลือก ข ค
ข	.00	.00		
ค	.00	.00		
ง	.16	-1		
3. สรุป				ดีมาก
ก	.00	.00	ค	ปรับตัวเลือก ก ง
ข	.16	.1		
ค	.83	1		
ง	.00	.00		
4. สรุป				ใช้ไม่ได้
ก	.00	.00	ค	ปรับตัวเลือก ก ข ง
ข	.00	.00		
ค	1	.00		
ง	.00	.00		
5. สรุป				ดีมาก
ก	.16	-1	ง	ปรับตัวเลือก ข ค
ข	.16	.20		
ค	.00	.00		
ง	.66	.80		
ข้อ	ค่าความยากง่าย (P)	ค่าอำนาจจำแนก (B- index)	เฉลย	แปลผล

6. สรุป				ดีมาก
ก	.00	.00		ปรับตัวเลือก ก ข
ข	.00	.00	ง	
ค	.16	-.1		
ง	.83	1		
7. สรุป				ดีมาก
ก	.00	.00		ปรับตัวเลือก ก ง
ข	.16	-1	ค	
ค	.50	.60		
ง	.33	.40		
8. สรุป				ดีมาก
ก	.16	-1		ปรับตัวเลือก ค ง
ข	.83	1	ข	
ค	.00	.00		
ง	.00	.00		
9. สรุป				ใช้ไม่ได้
ก	.16	.20		ปรับตัวเลือก ก ข ค
ข	.00	.00	ง	
ค	.16	.20		
ง	.66	-.40		
10. สรุป				ดีมาก
ก	.16	-1		ปรับตัวเลือก
ข	.66	.80	ข	ค ง
ค	.16	.20		
ง	.00	.00		
11. สรุป				ดีมาก
ก	.00	.00		ปรับตัวเลือก
ข	.33	-.80	ค	ก ง
ค	.66	.80		
ง	.00	.00		
ข้อ	ค่าความยากง่าย (P)	ค่าอำนาจจำแนก (B-index)	เฉลย	แปลผล

12. สรุป				ดีมาก
ก	.00	.00		ปรับตัวเล็กน้อย ก ง
ข	.66	.80	ข	
ค	.33	-.80		
ง	.00	.00		
13. สรุป				ใช้ไม่ได้
ก	.1	.00		ปรับตัวเล็กน้อย ข ค ง
ข	.00	.00	ค	
ค	.00	.00		
ง	.00	.00		
14. สรุป				ดีมาก
ก	.00	.00		ปรับตัวเล็กน้อย ก ข
ข	.00	.00	ง	
ค	.16	-1		
ง	.83	1		
15. สรุป				ใช้ไม่ได้
ก	.00	.00		ปรับตัวเล็กน้อย ก ค ง
ข	.83	-.20	ข	
ค	.16	.20		
ง	.00	.00		
16. สรุป				ดีมาก
ก	.00	.00		ปรับตัวเล็กน้อย ก ข
ข	.00	.00	ง	
ค	.16	.1		
ง	.83	1		
17. สรุป				ดีมาก
ก	.16	.20		ปรับตัวเล็กน้อย ก ข
ข	.16	.20	ค	
ค	.33	.40		
ง	.33	-.80		
ข้อ	ค่าความยากง่าย (P)	ค่าอำนาจจำแนก (B-index)	เฉลย	แปลผล



18. สรุป				ใช้ไม่ได้
ก	.33		-.80	ปรับตัวเลือก ข ค
ข	.00		.00	
ค	.50		.60	
ง	.16		.20	
19. สรุป				ดีมาก
ก	.16		-1	ปรับตัวเลือก ค ง
ข	.83		1	
ค	.00		.00	
ง	.00		.00	
20. สรุป				ดีมาก
ก	.00		.00	ปรับตัวเลือก ก ค
ข	.16		-.1	
ค	.00		.00	
ง	.83		1	
21. สรุป				ดีมาก
ก	.33		.40	ปรับตัวเลือก ข ง
ข	.33		.40	
ค	.33		-.80	
ง	.00		.00	
22. สรุป				ดีมาก
ก	.66		.80	ปรับตัวเลือก ข ง
ข	.00		.00	
ค	.16		-1	
ง	.16		.20	
23. สรุป				ดีมาก
ก	.00		.00	ปรับตัวเลือก ก ค
ข	.16		-1	
ค	.16		.20	
ง	.66		.80	
ข้อ	ค่าความยากง่าย (P)	ค่าอำนาจจำแนก (B-index)	เฉลี่ย	แปลผล

24. สรุป				ใช้ไม่ได้
ก	.16	-1		ปรับตัวเลือก ข ค
ข	.50	.60	ง	
ค	.33	.40		
ง	.00	.00		
25. สรุป				ดีมาก
ก	.16	-1		ปรับตัวเลือก ข ค
ข	.16	.20	ง	
ค	.16	.20		
ง	.50	.60		
26. สรุป				ดีมาก
ก	.00	.00		ปรับตัวเลือก ก ง
ข	.50	.60	ข	
ค	.33	-.80		
ง	.16	.20		
27. สรุป				ดีมาก
ก	.33	.40		ปรับตัวเลือก ข ง
ข	.00	.00	ก	
ค	.16	-1		
ง	.50	.60		
28. สรุป				ดีมาก
ก	.50	.60		ปรับตัวเลือก ค ง
ข	.16	.16	ก	
ค	.00	.00		
ง	.33	.33		
29. สรุป				ดีมาก
ก	.33	.40		ปรับตัวเลือก ก ค
ข	.50	.60	ข	
ค	.00	.00		
ง	.16	-1		
30. สรุป				ดีมาก

ก	.16	-1		ปรับตัวเล็ก ค ง
ข	.50	.60	ข	
ค	.33	.40		
ง	.00	.00		

จากตารางที่ 1 พบว่าผลของการวิเคราะห์คุณภาพของข้อสอบ จากจำนวน 30 ข้อ  
ข้อสอบที่มีระดับคุณภาพดีมาก จำนวน 24 ข้อ ข้อสอบที่มีระดับคุณภาพใช้ไม่ได้ จำนวน 6 ข้อ

## บทที่ 5

### สรุปผลการศึกษา อภิปรายผล และข้อเสนอแนะ

#### 5.1 สรุปผลการศึกษา

จากการวิเคราะห์ข้อสอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 กลางภาคเรียนที่ 2 ปีการศึกษา 2559 โรงเรียนบึงคอนแฮงส์แยมสอาดรังสิต ในประเด็นความยาก – ง่าย ของข้อสอบและอำนาจจำแนกของข้อสอบ จำนวน 30 ข้อ พบว่าข้อสอบจำนวน 30 ข้อ วิเคราะห์คุณภาพของข้อสอบได้ดังนี้

1. ข้อสอบที่ดีควรเก็บไว้ มีจำนวน 24 ข้อ คิดเป็นร้อยละ 80 ได้แก่ ข้อสอบที่มีผลสรุปดีมาก มี 24 ข้อ คือข้อ 1,2,3,5,6,7,8,10,11,12,14,16,17,19,20,21,22,23,25,26,27,28,29,30
2. ข้อสอบที่ควรตัดทิ้ง มี 6 ข้อ คิดเป็นร้อยละ 20 ได้แก่ ข้อสอบที่มีผลสรุปใช้ไม่ได้ มี 6 ข้อ คือข้อ 4,9,13,15,18,24

#### 5.2 อภิปรายผล

จากการวิเคราะห์ข้อสอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 กลางภาคเรียนที่ 2 ปีการศึกษา 2559 ทำการวิเคราะห์ข้อสอบ 30 ข้อ แล้วพบว่าโดยภาพรวมข้อสอบที่ดีควรเก็บไว้มีจำนวน 24 ข้อ คิดเป็นร้อยละ 80 ส่วนข้อสอบที่ควรตัดทิ้งมีจำนวน 6 ข้อ คิดเป็นร้อยละ 20 เป็นไปตามวัตถุประสงค์ของผู้ทำการศึกษาที่ต้องการวัดคุณภาพของข้อสอบให้เป็นไปตามมาตรฐานและนำผลการวิเคราะห์ไปปรับปรุงคุณภาพของข้อสอบให้ดียิ่งขึ้นต่อไป

#### 5.3 ข้อเสนอแนะ

1. ควรนำผลการวิเคราะห์ข้อสอบที่ดีไปใช้ร่วมกับผู้เรียนกลุ่มอื่น ๆ อีกครั้งว่าเป็นข้อสอบที่มีมาตรฐานหรือไม่
2. ควรนำข้อสอบที่วิเคราะห์ว่าควรปรับปรุง นำไปปรับปรุงแล้วทำการวิเคราะห์กับกลุ่มผู้เรียนซ้ำอีกครั้งหนึ่ง

## บรรณานุกรม

ธีรวัฒน์ สุพัตกุล.วิเคราะห์ข้อสอบ ทฤษฎีและคู่มือโปรแกรมฯ.กรุงเทพฯ : อักษราพิพัฒน์,2551.

บุญชม ศรีสะอาด.การวิจัยเบื้องต้น.กรุงเทพมหานคร : สุวีริยาสาส์น,2543

ศิริชัย กาญจนวาสี.ทฤษฎีการทดสอบแบบดั้งเดิม(พิมพ์ครั้งที่ 4) กทม : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย

ศิริชัย กาญจนวาสี. การวิเคราะห์ข้อสอบ เอกสารประกอบการอบรม การวิเคราะห์ข้อสอบเพิ่มมาตรฐานการศึกษา 26 ธันวาคม 2546 คณะสัตวแพทยศาสตร์ มหาวิทยาลัยขอนแก่น

นันทพร หาญวิทย์สกุล. รายงานการวิเคราะห์ข้อสอบ วิชาหลักการประชาสัมพันธ์ รหัส 2201 – 2701

แผนกวิชาการตลาด 2552 วิทยาลัยบริหารธุรกิจและท่องเที่ยวกรุงเทพ

