บทที่ 1

บทนำ

ปัญหาและความเป็นมาของปัญหา

การวัดผลการศึกษาเป็นองค์ประกอบที่สำคัญประการหนึ่ง ของกระบวนการศึกษาซึ่งใช้ควบคู่ไปกับการเรียนการสอน ทั้งนี้เนื่องจากการวัดผลการศึกษา จะเป็นตัวชี้ให้ครูอาจารย์หรือผู้ที่เกี่ยวข้องกับการศึกษาทราบว่า ผลการจัดการเรียนการสอนที่ผ่านมานั้น สัมฤทธิ์ผลมากน้อยเพียงใด อีกทั้งให้ข้อมูลเพื่อนำมาปรับปรุงและพัฒนาคุณภาพของการศึกษาในระดับต่างๆ การวัดผล จะมีความถูกต้องเมื่ออาศัยเครื่องมือที่ดี และมีเกณฑ์สำหรับการเปรียบเทียบที่เหมาะสม เพื่อใช้ในการแปลความหมายผลการวัด แบบสอบ (Test) นับเป็นเครื่องมือที่ใช้กันแพร่หลายมากที่สุดทางการศึกษา เพราะถือว่าเป็นเครื่องมือที่วัดคุณลักษณะ ความสามารถหรือสมรรถภาพของผู้สอบได้ ซึ่งแบบสอบที่ดีนั้น จะต้องสามารถวัดได้ตรงกับความสามารถที่แท้จริงของแต่ละบุคคล และเกิดความยุติธรรมแก่ทุกคนด้วย (สงบ ลักษณะ, 2525 :47)

การสร้างแบบสอบคู่ขนาน (Parallet Test) เป็นความพยายามในการแก้ปัญหาการใช้แบบสอบฉบับเดิมในการวัดผลการศึกษา ซึ่งมักจะเกิดปัญหาข้อสอบรั่ว หรือการได้เปรียบเสียเปรียบกันระหว่างผู้สอบเมื่อดำเนินการสอบไม่พร้อมกัน ทำนองเดียวกันในการตรวจสอบคุณภาพการศึกษาโดยการนำคะแนนที่ได้จากการสอบในแต่ละปีการศึกษามาเปรียบเทียบกัน เพื่อให้เห็นแนวโน้มที่เพิ่มขึ้นหรือลดลงของผลสัมฤทธิ์ทางการเรียนในแต่ละปี มักจะมีการสร้างแบบสอบชุดใหม่ขึ้นทุกปี เพื่อหลีกเลี่ยงปัญหาการได้ข้อมูลที่ไม่ใช่ข้อมูลที่แท้จริง และความไม่ยุติธรรมซึ่งเกิดจากการใช้แบบสอบชุดเดิมในสถานการณ์การสอบที่แตกต่างกัน โดยพยายามสร้างแบบสอบชุดใหม่ให้มีลักษณะคู่ขนานกับแบบสอบชุดเดิมมากที่สุด เพื่อจะได้นำคะแนนผลการสอบจากแบบสอบชุดเดิมและแบบสอบชุดใหม่ มาเปรียบเทียบกันได้โดยตรงเหมือนกับเป็นแบบสอบชุดเดียวกัน แต่ในทางปฏิบัติการสร้างแบบสอบคู่ขนานให้มีลักษณะข้อสอบ คือ เนื้อหา ค่าสถิติของข้อสอบและแบบสอบทั้งฉบับเท่ากัน มีโอกาสเป็นไปได้น้อยมาก มักได้เพียงแบบสอบที่เทียบเคียงกันเท่านั้น (Gulliksen, 1950 : 174) ดังนั้นจึงมีผู้ประยุกต์วิธีการทางสถิติ มาใช้ในการที่จะทำให้คะแนนจากแบบสอบที่วัดคุณลักษณะเดียวกัน แต่ต่างฉบับกัน สามารถนำมาเปรียบเทียบกันได้อย่างมีความหมายมากขึ้น วิธีการดังกล่าว เรียกว่า การเทียบมาตรา (Equating) ซึ่งสามารถนำไปใช้กับสถานการณ์ของการสอบได้ดังนี้ (สงบ ลักษณะ, 2525 : 23)

1) บุคคล 1 กลุ่ม ได้รับการทดสอบด้วยแบบทดสอบ 2 ฉบับ หรือมากกว่าในเวลาใกล้เคียงกัน แล้วต้องการเทียบระดับคะแนนระหว่างแบบทดสอบฉบับต่างๆ เหล่านั้น

2) แบบทดสอบฉบับเดียวกัน แต่นำไปใช้ทดสอบบุคคลหลายกลุ่มที่อาจมีความสามารถแตกต่างกัน แล้วต้องการเปรียบเทียบบุคคลทั้งหมด ทุกกลุ่ม บนพื้นฐานของคะแนนจริงหรือความสามารถแท้

3) แบบทดสอบ 1 ฉบับ แต่จัดแบ่งกลุ่มของข้อสอบใน 1 ฉบับนี้ เป็นหลายระดับเรียงตามความง่ายไปหายากโดยมีข้อสอบร่วม (Anchor Items) ระหว่างกลุ่มที่อยู่ติดๆกันบ้างเป็นบางข้อ ต้องการนำแบบทดสอบนี้ไปสอบกับบุคคลต่างระดับความสามารถกัน ผลคือ คนต่างกลุ่มกันแต่ทำข้อสอบบางข้อร่วมกัน และมีอีกหลายข้อที่คนต่างกลุ่มต่างก็ทำอิสระแก่กัน แล้วต้องการเทียบระดับความสามารถหรือระดับคะแนนของผู้สอบทุกกลุ่มในคราวเดียวกัน ซึ่งกรณีนี้เป็นกรณีสำคัญของการสร้างข้อสอบมาตรฐานชนิดหลายระดับ (Multi Level) ไว้ในฉบับเดียวกัน เพื่อประเมินความก้าวหน้าของนักเรียนคนเดิม หรือกลุ่มเดิมในแต่ละปีที่ผ่านไป

จากสถานการณ์ของการสอบ ถ้าพิจารณาจากการนำการเทียบมาตราไปใช้กับผู้สอบและแบบสอบ สามารถจำแนกได้ 2 ลักษณะคือ (สงบ ลักษณะ, 2523 : 23, อ้างถึงใน วิรัช วรรณรัตน์, 2530 : 69-70)

1. การเทียบมาตราในแนวนอน (Horizontal Equating) เป็นการเทียบคะแนนระหว่างแบบสอบสองชุดที่มีความยากพอๆ กัน กล่าวคือ ต้องการเทียบคะแนนจากแบบสอบฉบับหนึ่งว่าจะเท่ากับกี่หน่วยคะแนนของแบบสอบอีกฉบับหนึ่ง ดังนั้นในการทดสอบจะต้องสอบกับบุคคลกลุ่มเดียวกัน

2. การเทียบมาตราในแนวตั้ง (Vertical Equating) เป็นการเทียบคะแนนระหว่างแบบสอบสองชุดที่มีความยากต่างกัน เมื่อนำไปสอบกับบุคคลที่มีความสามารถต่างกัน กล่าวคือ ต้องการเทียบคะแนนจากแบบสอบสองชุด เมื่อความยากของข้อสอบต่างกัน และเมื่อสอบกับบุคคลที่มีความสามารถต่างกัน การเทียบมาตราในลักษณะนี้จะกระทำได้โดยมีแบบสอบร่วมฉบับหนึ่ง

ในการเทียบมาตรา เป็นการกำหนดเงื่อนไขและสร้างกฎการแปลงคะแนนในแต่ละครั้งที่ทำการเทียบมาตรา ดังนั้นจึงมีรูปแบบการเทียบมาตรา (Equating Model) ที่แตกต่างกันหลายรูปแบบ ถ้าพิจารณาจากการใช้เกณฑ์ในเชิงทฤษฎีการวัดผล แบ่งได้ 2 รูปแบบ คือ

1. รูปแบบการเทียบมาตราโดยการใช้ทฤษฎีดั้งเดิม ซึ่งมี 2 วิธี คือ วิธีอิควิเปอร์เซนไทล์ (Equipercentile Equating) และวิธีเทียบแบบเส้นตรง (Linear Equating) (Kolen and Whitney,1982, อ้างถึงใน ภาวิณี ศรีสุขวัฒนานนท์ , 2529 : 22)

2. รูปแบบการเทียบมาตราโดยการใช้ทฤษฎีการตอบสนองข้อสอบ ซึ่งภายใต้เงื่อนไขบางประการของทฤษฎีการตอบสนองข้อสอบ หรือที่เรียกว่า Latent Trait Theory จะสามารถคำนวณค่าความสามารถแท้ของผู้สอบได้ชัดเจน และสามารถเชื่อมโยงความสามารถของผู้สอบเข้ากับโอกาสที่เขาจะทำข้อสอบได้ถูกต้องได้อย่างชัดเจนในลักษณะของฟังก์ชัน Logistic ที่แสดงในรูปเส้นโค้งลักษณะเฉพาะของข้อสอบแต่ละข้อ (Item Characteristic Curve)

จากการวิจัยการนำรูปแบบการเทียบมาตราไปใช้ในการเทียบมาตราในแนวนอน โคเลน (Kolen, 1981) กล่าวว่า การเทียบมาตราโดยใช้ทฤษฎีดั้งเดิมไม่มีเหตุผลเพียงพอ ในการที่จะอธิบายถึงความทัดเทียมกัน (Equity) ความสมมาตร (Symmetry) และความไม่ผันแปรตามกลุ่มตัวอย่าง (Invariance) แต่ในการเทียบมาตราโดยใช้ทฤษฎีการตอบสนองข้อสอบสามารถแก้ปัญหานี้ได้จากการศึกษาของปีเตอร์เซนและคณะ (Petersen and others. 1982 : 71-135) ที่เปรียบเทียบผลการใช้รูปแบบเชิงเส้นตรง รูปแบบอิควิเปอร์เซนไทล์ และรูปแบบการใช้ทฤษฎีการตอบสนองข้อสอบ พบว่า รูปแบบการเทียบมาตราโดยใช้ทฤษฎีการตอบสนองข้อสอบสามารถประมาณค่าของการสอบได้แน่นอนและเหมาะสมกว่าทฤษฎีดั้งเดิม เช่นเดียวกับการศึกษาของโคเลนและวิทนีย์ (Kolen and Whitney. 1982 : 279-293) ได้ศึกษาความเพียงพอของการเทียบมาตรา โดยเปรียบเทียบวิธีเดียวกับปีเตอร์เซนและคณะ พบว่ารูปแบบการเทียบมาตราโดยการใช้ทฤษฎีการตอบสนองข้อสอบให้ความเพียงพอมากที่สุด

สำหรับการเทียบมาตราในแนวตั้งนั้น งานวิจัยส่วนใหญ่นำทฤษฎีการตอบสนองข้อสอบแบบจำลองของราส์ช และแบบจำลองสามพารามิเตอร์ เพื่อการศึกษาเปรียบเทียบความคลาดเคลื่อน ประเมินความเพียงพอของวิธีการเทียบมาตรา เมื่อใช้ในสถานการณ์เกี่ยวกับ การออกแบบการเก็บรวบรวมข้อมูล ลักษณะกลุ่มตัวอย่าง ขนาดกลุ่มตัวอย่าง ลักษณะของแบบสอบ วิธีประมาณค่าพารามิเตอร์ ที่แตกต่างกัน และประยุกต์ใช้ในการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนของผู้เรียน ซึ่งยังไม่สามารถสรุปได้แน่นอนว่า การเทียบมาตรารูปแบบใด มีประสิทธิภาพมากที่สุดในแต่ละสถานการณ์

จากที่กล่าวมาผู้วิจัยจึงสนใจที่จะศึกษาเพื่อเปรียบเทียบความคลาดเคลื่อนและตรวจสอบความเพียงพอของการเทียบมาตรา ระหว่างรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์และสามพารามิเตอร์ เมื่อนำมาใช้ในการเทียบมาตราในแนวตั้งกับสถานการณ์การออกแบบเก็บรวบรวมข้อมูลที่ใช้แบบสอบร่วม กับกลุ่มผู้สอบที่ความสามารถแตกต่างกัน 2 ระดับ เพื่อนำสารสนเทศที่ได้จากการศึกษาช่วยในการตัดสินใจเลือกวิธีการเทียบมาตราในแนวตั้งและนำไปประยุกต์ใช้ในการตรวจสอบคุณภาพการศึกษา

วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบคุณภาพของการเทียบมาตราในแนวตั้งรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์กับสามพารามิเตอร์ ดังนี้

1. เพื่อเปรียบเทียบความคลาดเคลื่อนของการเทียบมาตราในแนวตั้งระหว่างรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์กับสามพารามิเตอร์

2. เพื่อตรวจสอบความเพียงพอของการเทียบมาตราในแนวตั้งรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์และสามพารามิเตอร์

สมมติฐานของการวิจัย

การวิจัยครั้งนี้กำหนดสมมติฐานดังนี้

1. ความคลาดเคลื่อนของการเทียบมาตราในแนวตั้งระหว่างรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์กับสามพารามิเตอร์แตกต่างกัน

2. ความเพียงพอของการเทียบมาตราในแนวตั้งรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์อยู่ในระดับที่ไม่น่าพอใจ รูปแบบทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์อยู่ในระดับที่น่าพอใจ

ความสำคัญและประโยชน์

ผลการวิจัยครั้งนี้มีความสำคัญและประโยชน์ดังนี้

1. ทำให้ทราบถึงคุณภาพของการเทียบมาตราในแนวตั้งรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์และสามพารามิเตอร์ เพื่อเป็นแนวทางในการตัดสินใจเลือกรูปแบบทฤษฎีการตอบสนองข้อสอบ ไปใช้ให้เหมาะสมกับวิธีการเทียบมาตราในแนวตั้ง

2. นำวิธีการเทียบมาตราในแนวตั้งรูปแบบทฤษฎีตอบสนองข้อสอบที่มีคุณภาพ ไปใช้สำหรับการศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียน และนำข้อมูลที่ได้เป็นแนวทางในการวางแผนเพื่อปรับปรุงพัฒนาการเรียนการสอนให้มีประสิทธิภาพยิ่งขึ้น

ขอบเขตของการวิจัย

1. การวิจัยครั้งนี้เป็นการเทียบมาตราในแนวตั้ง (Vertical Equating) ของนักเรียนที่มีความสามารถต่างกัน 2 ระดับ คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 และระดับชั้นมัธยมศึกษาปีที่ 2 ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดพัทลุง

2. การเปรียบเทียบคุณภาพของการเทียบมาตราในแนวตั้ง เป็นการเปรียบเทียบคุณภาพของการเทียบมาตราระหว่างรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์กับสามพารามิเตอร์ เมื่อแบบสอบมีความยากต่างกันและผู้สอบมีระดับความสามารถต่างกัน โดยใช้วิธีแปลง (Transformation) ค่าความสามารถของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 2 ให้อยู่ในสเกลเดียวกับความสามารถของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 ตามสมการเชิงเส้นตรง (Linear Equation) ของลอร์ดและฮูเวอร์ (L0yd and Hoover, 1980 : 180; Hamilton and Swanminathan, 1985, อ้างจาก พรพิมล นาคเวช, 2537 : 63) โดยมีแบบสอบร่วมภายนอก

3. เครื่องมือที่ใช้ในการวิจัย เป็นแบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา คณิตศาสตร์ เรื่อง สมการ สร้างจากเนื้อหาในรายวิชา ค 102 และ ค 204 แบ่งเป็นแบบสอบที่ใช้สำหรับกลุ่มเทียบมาตรา จำนวน 2 ฉบับ แบบสอบร่วม 1 ฉบับ และแบบสอบทานผล 1 ฉบับ

นิยามศัพท์เฉพาะ

1. การเทียบมาตรา (Equating) หมายถึง กระบวนการใช้วิธีการทางสถิติ เพื่อปรับคะแนนที่ได้จากแบบสอบต่างฉบับ ซึ่งวัดสิ่งเดียวกัน ให้สามารถเปรียบเทียบกันได้

2. การเทียบมาตราในแนวตั้ง (Vertical Equating) หมายถึง การเทียบมาตราระหว่างแบบสอบที่มีความยากต่างกันเมื่อใช้กับบุคคลที่มีความสามารถต่างกัน ในการวิจัยใช้แบบสอบที่มีความยากต่างกันสองระดับ คือ แบบสอบระดับชั้นมัธยมศึกษาปีที่ 1 และแบบสอบระดับชั้นมัธยมศึกษาปีที่ 2 เมื่อนำไปใช้กับกลุ่มผู้สอบที่มีระดับความสามารถต่างกัน คือ ผู้สอบซึ่งเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 และนักเรียนชั้นมัธยมศึกษาปีที่ 2

3. แบบสอบเทียบมาตรา หมายถึง แบบสอบที่ใช้สำหรับกลุ่มเทียบมาตรา ในการวิจัยครั้งนี้ หมายถึง แบบสอบวิชาคณิตศาสตร์ เรื่อง สมการ สำหรับผู้สอบชั้นมัธยมศึกษาปีที่ 1 และแบบสอบวิชาคณิตศาสตร์ เรื่อง สมการ สำหรับผู้สอบชั้นมัธยมศึกษาปีที่ 2 ที่สร้างขึ้นตามหลักสูตรมัธยมศึกษาตอนต้นของกระทรวงศึกษาธิการ พุทธศักราช 2521 (ฉบับปรับปรุง 2533)

4. แบบสอบทานผล หมายถึง แบบสอบที่ใช้สำหรับกลุ่มสอบทานผล ในการวิจัยครั้งนี้ หมายถึง แบบสอบวิชาคณิตศาสตร์ เรื่อง สมการ จำนวน 48 ข้อ ซึ่งได้จากการนำแบบสอบเทียบมาตราทั้งสองฉบับมารวมกัน

5. แบบสอบร่วม (Anchor Test) หมายถึง แบบสอบที่ใช้สอบทั้งระดับชั้นมัธยมศึกษาปีที่ 1 และระดับชั้นมัธยมศึกษาปีที่ 2 สำหรับนำค่าพารามิเตอร์ไปใช้ในสมการแปลงเชิงเส้นตรง ในการวิจัยครั้งนี้ เลือกข้อสอบที่มีค่าความยากสูง ในระดับชั้นมัธยมศึกษาปีที่ 1 จำนวน 6 ข้อ เป็นข้อสอบร่วม

6. กลุ่มเทียบมาตรา หมายถึง กลุ่มนักเรียนชั้นมัธยมศึกษาปีที่ 1 และมัธยมศึกษาปีที่ 2 ที่ได้จากการสุ่มประชากรนักเรียนชั้นมัธยมศึกษาปีที่ 1 และชั้นมัธยมศึกษาปีที่ 2 ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดพัทลุง ซึ่งต้องสอบแบบสอบเทียบมาตราตามระดับชั้นของตน และแบบสอบร่วมอีก 1 ฉบับ

7. กลุ่มสอบทานผล หมายถึง กลุ่มนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่ได้จากการสุ่มประชากรนักเรียนชั้นมัธยมศึกษาปีที่ 2 ของโรงเรียนในสังกัดกรมสามัญศึกษา จังหวัดพัทลุง ซึ่งไม่มีหน่วยที่ซ้ำกับกลุ่มเทียบมาตรา และแต่ละคนจะทำแบบสอบทานผล

8. คุณภาพของการเทียบมาตรา หมายถึง ความถูกต้องแม่นยำของการเทียบมาตรา ที่พิจารณาจาก ความคลาดเคลื่อนของการเทียบมาตราและความเพียงพอของการเทียบมาตรา

9. ดัชนีความแตกต่าง (Discrepancy Index) หมายถึง ค่าความคลาดเคลื่อนรวม (Total Error) ในการเทียบมาตรา คำนวณจากค่าเฉลี่ยกำลังสองของความแตกต่างของคะแนนที่สมมูลกัน แล้วถ่วงน้ำหนักด้วยความแปรปรวนของคะแนนเกณฑ์

10. คะแนนสมมูล หมายถึง ค่าความสามารถที่ได้จากการทำแบบสอบต่างฉบับที่เทียบกันได้ว่าเป็นความสามารถที่เท่าเทียมกัน

11. คะแนนเกณฑ์ หมายถึง ค่าความสามารถของผู้สอบกลุ่มสอบทานผล ซึ่งได้จากการทำข้อสอบในแบบสอบทานผล ในส่วนที่เป็นเนื้อหาในระดับชั้นมัธยมศึกษาปีที่ 1

12. ความเพียงพอของการเทียบมาตรา หมายถึง ระดับความถูกต้องของการเทียบมาตราของแบบสอบฉบับหนึ่งไปสู่แบบสอบอีกฉบับหนึ่ง การตัดสินความเพียงพอของการเทียบมาตราในการวิจัยครั้งนี้ ใช้การเปรียบเทียบดัชนีความแตกต่าง โดยใช้เกณฑ์ของปีเตอร์เซนและคณะ (Petersen, et al. 1982 : 93-94, อ้างจาก ภาวิณี ศรีสุขวัฒนานันท์, 2529 : 103) ดังนี้

ระดับน่าพอใจอย่างยิ่ง	เมื่อ	$C \leq (0.05\ SD_x)^2$
ระดับน่าพอใจ	เมื่อ	$(0.05\ SD_x)^2 < C \leq (0.10\ SD_x)^2$
ระดับปานกลาง	เมื่อ	$(0.10\ SD_x)^2 < C \leq (0.15\ SD_x)^2$
ระดับไม่น่าพอใจ	เมื่อ	$(0.15\ SD_x)^2 < C \leq (0.20\ SD_x)^2$
ระดับไม่น่าพอใจอย่างยิ่ง	เมื่อ	$(0.20\ SD_x)^2 < C$