บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์ เพื่อตรวจสอบคุณภาพของการเทียบมาตราเชิงเส้นตรง เมื่อใช้แบบสอบร่วมภายในที่มีช่วงความยาวต่างกัน 4 ขนาด คือ 10 15 20 และ 25 ข้อ โดยเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบมาตราเชิงเส้นตรง (Standard error of Equating: SEE) และเปรียบเทียบดัชนีความแตกต่าง (Discrepancy Index: C) จากการวิเคราะห์กลุ่มสอบทานผล (Cross Validation Group)

กลุ่มตัวอย่างที่ใช้ในการวิจัย ได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2534 ของโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา กลุ่มโรงเรียนที่สอง จังหวัดขอนแก่น จำนวน 927 คน ซึ่งได้จากการวางแผนสุ่มตัวอย่างเพื่อการจัดสอบเป็นกลุ่มที่ใช้ในการวิจัยเพื่อนำผลการสอบไปวิเคราะห์เทียบมาตราคะแนนของแบบสอบ ได้นักเรียนเป็นกลุ่มตัวอย่างเทียบมาตราจำนวน 810 คน และเป็นกลุ่มที่ใช้ในการวิจัยเพื่อตรวจสอบคุณภาพของวิธีการเทียบมาตราเชิงเส้นตรง ได้นักเรียนเป็นกลุ่มสอบทานผลจำนวน 117 คน

เครื่องมือที่ใช้ในงานวิจัยประกอบด้วยแบบสอบเทียบมาตรา 2 ฉบับเป็นแบบสอบวัดผลสัมฤทธิ์ปลายภาค ในวิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 2 ภาคต้น (ค 203) ของกลุ่มโรงเรียนมัธยมศึกษาตอนต้น จังหวัดขอนแก่น กลุ่มที่ 2 เป็นแบบสอบที่สร้างขึ้นใช้ในสถานการณ์การสอบจริงมาแล้วเมื่อปีการศึกษา 2532 (ฉบับ X) และ 2533 (ฉบับ Y) ฉบับละ 60 ข้อ เป็นแบบสอบเลือกตอบชนิด 4 ตัวเลือก แบบสอบทั้ง 2 ฉบับ วัดความรู้ครอบคลุมเนื้อหา 6 เรื่อง คือ จำนวนและตัวเลข ระบบจำนวนเต็ม พื้นที่ ความเท่ากันทุกประการ เศษส่วนและทศนิยม คุณสมบัติของรูปสามเหลี่ยมมุมฉาก ทดลองใช้แบบสอบทั้งสองฉบับกับนักเรียนชั้นมัธยมศึกษาปีที่สอง โรงเรียนประชารัฐวิทยาเสริม จำนวน 69 คน ซึ่งไม่ซ้ำกับกลุ่มตัวอย่างเทียบมาตราและกลุ่มสอบทานผล เพื่อหาคุณภาพของแบบสอบ พร้อมทั้งปรับปรุงคำถาม และตัวเลือกของข้อสอบบางข้อที่ไม่ผ่านเกณฑ์ให้เหมาะสม แล้วให้ผู้เชี่ยวชาญด้านเนื้อหาวิชาตรวจความตรงและปรับปรุงภาษาของคำถาม ตัวเลือก และตัวลวงก่อนนำไปใช้กับกลุ่มตัวอย่าง

สำหรับแบบสอบร่วมที่ใช้ในการเทียบมาตราเป็นแบบสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 ภาคต้น (ค203) ซึ่งเป็นเนื้อหาเดียวกันกับแบบสอบเทียบมาตรา ซึ่งผู้วิจัยเป็นผู้สร้าง จำนวน 40 ข้อ สร้างจากจุดประสงค์สำคัญ 7 จุดประสงค์ ทดลองใช้แบบสอบเพื่อหาคุณภาพและความเหมาะสมที่จะนำไปใช้กับกลุ่มตัวอย่างจริง โดยนำไปสอบกับนักเรียนชั้นมัธยมศึกษาปีที่ 2 โรงเรียนสวนกุหลาบวิทยาลัยนนทบุรีจำนวน 40 คน และโรงเรียนนนทรีวิทยา จำนวน 37 คน แล้วคัดเลือกข้อสอบจากกลุ่มข้อสอบที่หาคุณภาพแล้ว ตามจำนวนข้อในแต่ละจุดประสงค์ได้แบบสอบร่วมที่นำไปใช้ในการเทียบมาตราจำนวน 25 ข้อ นำไปรวมกับแบบสอบเทียบมาตราในแต่ละฉบับซึ่งเรียงข้อสอบร่วมตามเนื้อหาวิชาในแต่ละบท ได้แบบสอบสำหรับนำไปใช้ในการเทียบมาตรา 2 ฉบับ คือ ฉบับ X และฉบับ Y ซึ่งมีความยาวฉบับละ 85 ข้อ แต่ละฉบับมีแบบสอบร่วม 25 ข้อ

การรวบรวมข้อมูล

ผู้วิจัยได้นำแบบสอบที่ใช้ในการเทียบมาตรา ซึ่งจัดเตรียมไว้ทั้ง 2 ฉบับ คือ ฉบับ X และฉบับ Y ฉบับละ 85 ข้อ ไปทดสอบกับกลุ่มตัวอย่างเทียบมาตรา โดยสุ่มอย่างมีระบบให้นักเรียนในแต่ละห้อง สอบแบบสอบคนละฉบับสลับกัน แล้วนำคะแนนผลการสอบของนักเรียนมาจัดแยกออกตามลักษณะของแบบสอบเทียบมาตรา แล้วตรวจคะแนนของนักเรียนแต่ละคนที่สอบแบบสอบขนาด 85 ข้อในแต่ละฉบับโดยตรวจตามความยาวของแบบสอบร่วม จะได้คะแนนของนักเรียนเมื่อทำข้อสอบ 70 ข้อ 75 ข้อ 80 ข้อ และ 85 ข้อ ทั้งนักเรียนที่สอบแบบสอบฉบับ X และฉบับ Y

ผู้วิจัยยังได้นำแบบสอบเทียบมาตราทั้งสองฉบับ ๆ ละ 85 ข้อ ไปสอบกับกลุ่มสอบทานผลเพื่อหาคุณภาพของวิธีการเทียบมาตราเชิงเส้นตรง โดยให้กลุ่มตัวอย่างแต่ละคนสอบแบบสอบเทียบมาตราทั้งสองฉบับ โดยสุ่มอย่างมีระบบให้แต่ละคนสอบแบบสอบฉบับ X แล้วตามด้วยฉบับ Y หรือสอบฉบับ Y ก่อน แล้วตามด้วยฉบับ X โดยสอบวันละ 1 ฉบับ เว้น 1 วัน จึงสอบอีก 1 ฉบับ นำคะแนนผลการสอบของกลุ่มสอบทานผลมาดำเนินการเช่นเดียวกับกลุ่มตัวอย่างเทียบมาตรา ได้คะแนนผลการสอบแต่ละคู่ในแต่ละความยาวของแบบสอบร่วม

การวิเคราะห์ข้อมูล

1. **การวิเคราะห์ข้อมูลกลุ่มตัวอย่างเทียบมาตรา**

1.1 วิเคราะห์ข้อสอบตามทฤษฎีดั้งเดิม (Classical Theory) เพื่อหาค่าความยาก อำนาจจำแนก และความเที่ยงของแบบสอบแต่ละฉบับ

1.2 หาค่าสถิติพื้นฐานของกลุ่มตัวอย่างโดยใช้โปรแกรมคอมพิวเตอร์ SPSS-X เพื่อหาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน

1.3 คำนวณค่าสถิติต่าง ๆ ที่ใช้ในการเทียบมาตราเชิงเส้นตรง อันได้แก่ ค่าเฉลี่ยและความแปรปรวนของแบบสอบฉบับ X และ Y ค่าเฉลี่ยและความแปรปรวนของแบบสอบร่วมจากกลุ่ม α และ β และกลุ่มตัวอย่างรวม $t(\alpha+\beta)$ ตามวิธีการประมาณค่าด้วยวิธีความเป็นไปได้สูงสุดของ ลอร์ด (Lord 1950 cited by Angoff 1984: 105-106) จากการรวบรวมข้อมูลของแองกอฟรูปแบบที่ 3

1.4 คำนวณค่าความชัน (A) และจุดตัดแกน (B) ที่ใช้ในสมการเส้นตรงและสร้างสมการเส้นตรงที่ใช้ในการเทียบมาตราในแต่ละความยาวของแบบสอบร่วม

1.5 เทียบมาตราเชิงเส้นตรง โดยการแปลงคะแนนจากแบบสอบฉบับ Y (ปีการศึกษา 2533) ซึ่งเป็นฉบับใหม่ ไปยังแบบสอบฉบับ X (ปีการศึกษา 2532) ซึ่งเป็นฉบับเก่า ได้ตารางเทียบมาตรา 4 ตาราง

1.6 คำนวณค่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตรา โดยใช้สูตรการประมาณความคลาดเคลื่อนมาตรฐานที่เสนอโดยลอร์ด (Lord 1950 cited by Angoff 1984: 106) และเปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานจากค่าประสิทธิภาพสัมพัทธ์ (RE)

2. **การวิเคราะห์ข้อมูลกลุ่มสอบทานผล**

2.1 คำนวณค่าดัชนีความแตกต่าง (C) จากกลุ่มสอบทานผลโดยนำผลการสอบจากแบบสอบฉบับ Y มาเทียบคะแนนจากตารางเทียบมาตราได้คะแนนสมมูลกับคะแนนในแบบสอบฉบับ X (X^*) แล้ววัดความแตกต่างระหว่างคะแนนสมมูลกับคะแนนที่ได้จากผลการสอบฉบับ X

2.2 ประเมินค่าดัชนีความแตกต่าง (C) ตามเกณฑ์ของบีเตอร์เซนและคณะ และทดสอบนัยสำคัญของดัชนีความแตกต่าง โดยใช้สถิติทดสอบนันพาราเมตริก Friedman Test

<u>สรุปผลการวิจัย</u>

1. ผลการเทียบมาตราจากแบบสอบฉบับ Y ไปยังฉบับ X มีลักษณะสอดคล้องกัน คะแนนสมมูลส่วนใหญ่ของแบบสอบฉบับ Y ต่ำกว่าคะแนนแบบสอบฉบับ X ในแต่ละความยาวของแบบสอบร่วม คือ

1.1 ในการเทียบมาตราที่ใช้แบบสอบร่วมขนาด 10 ข้อ คะแนนสมมูลจากแบบสอบฉบับ Y ในช่วงคะแนน 1-11 สูงกว่าคะแนนของแบบสอบฉบับ X มีค่าตั้งแต่ .88-.03 ส่วนในช่วงคะแนน 12-คะแนนสูงสุด ต่ำกว่าคะแนนของแบบสอบฉบับ X มีค่าตั้งแต่ .05-4.96

1.2 ในการเทียบมาตราที่ใช้แบบสอบร่วมขนาด 15 ข้อ คะแนนสมมูลจากแบบสอบฉบับ Y ทุกระดับคะแนน ต่ำกว่า คะแนนของแบบสอบฉบับ X มีค่าตั้งแต่ .07-5.88

1.3 ในการเทียบมาตราที่ใช้แบบสอบร่วมขนาด 20 ข้อ คะแนนสมมูลจากแบบสอบฉบับ Y ในช่วงคะแนน 1-7 สูงกว่าคะแนนของแบบสอบฉบับ X มีค่าตั้งแต่ .54-.08 สำหรับที่ระดับคะแนน 8 มีค่าเท่ากันกับคะแนนของแบบสอบฉบับ X ส่วนในช่วงคะแนน 9-คะแนนสูงสุด ต่ำกว่าคะแนนของแบบสอบฉบับ X มีค่าตั้งแต่ .08-5.61

1.4 ในการเทียบมาตราที่ใช้แบบสอบร่วมขนาด 25 ข้อ คะแนนสมมูลจากแบบสอบฉบับ Y ต่ำกว่าคะแนนของแบบสอบฉบับ X มีค่าตั้งแต่ .02-5.56 เฉพาะคะแนนสมมูลจากแบบสอบ Y ที่ระดับคะแนน 1 สูงกว่าคะแนนของแบบสอบฉบับ X อยู่ .05

2. ผลการวิเคราะห์ความคลาดเคลื่อนมาตรฐานในการเทียบมาตราเชิงเส้นตรง เมื่อใช้แบบสอบร่วมขนาด 10 15 20 และ 25 ข้อ ให้ค่าความคลาดเคลื่อนมาตรฐานที่แตกต่างกัน โดยแบบสอบร่วมที่ยาวกว่าให้ค่าความคลาดเคลื่อนมาตรฐานที่ต่ำกว่าในทุกช่วงคะแนนสอดคล้องกับสมมุติฐานข้อที่หนึ่งตามที่กำหนดไว้ คือ

2.1 ความคลาดเคลื่อนมาตรฐานในแต่ละความยาวของแบบสอบร่วม มีลักษณะสอดคล้องกันคือมีค่าต่ำสุดที่ระดับคะแนนเฉลี่ยของแต่ละความยาวของแบบสอบร่วม แล้วเพิ่มไปสู่ระดับคะแนนสูงและต่ำ โดยเมื่อใช้แบบสอบร่วมขนาด 10 ข้อ ให้ค่าความคลาดเคลื่อนมาตรฐานที่ระดับคะแนนเฉลี่ย (31) .4874 แล้วมีค่าเพิ่มขึ้นไปสู่ระดับคะแนนสูงและต่ำเท่ากับ 1.6911 และ 1.3332

เมื่อใช้แบบสอบร่วมขนาด 15 ข้อ ให้ค่าความคลาดเคลื่อนมาตรฐานที่ระดับคะแนนเฉลี่ย (33) .4662 แล้วมีค่าเพิ่มขึ้นไปสู่ระดับคะแนนสูงและต่ำเท่ากับ 1.6418 และ 1.2895

เมื่อใช้แบบสอบร่วมขนาด 20 ข้อ ให้ค่าความคลาดเคลื่อนมาตรฐานที่ระดับคะแนนเฉลี่ย (35) .4412 แล้วมีค่าเพิ่มขึ้นไปสู่ระดับคะแนนสูงและต่ำเท่ากับ 1.5782 และ 1.2274

เมื่อใช้แบบสอบร่วมขนาด 25 ข้อ ให้ค่าความคลาดเคลื่อนมาตรฐานที่ระดับคะแนนเฉลี่ย (37) .4084 แล้วมีค่าเพิ่มขึ้นไปสู่ระดับคะแนนสูงและต่ำเท่ากับ 1.4642 และ 1.1513

นั่นคือ ที่ระดับคะแนนต่ำสุดและสูงสุด ในแต่ละความยาวของแบบสอบร่วมขนาด 10 15 20 และ 25 ข้อ ค่าความคลาดเคลื่อนมาตรฐานของแบบสอบร่วมลดลงตามลำดับ คือที่ระดับคะแนนต่ำสุด มีค่า 1.6911 1.6418 1.5782 และ 1.4642 ตามลำดับ และที่ระดับคะแนนสูงสุดมีค่า 1.3332 1.2895 1.2274 และ 1.1513 ตามลำดับ

2.2 ประเมินค่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตรา ณ คะแนนมาตรฐานของ y (Z_y) จากค่าเฉลี่ยของประสิทธิภาพสัมพัทธ์ เมื่อใช้แบบสอบร่วมขนาด 10 15 และ 20 ข้อ ให้ค่าความแม่นยำของการเทียบมาตรา เมื่อเทียบกับแบบสอบร่วมขนาด 25 ข้อ ร้อยละ 86.26 89.08 และ 93.30 ตามลำดับ นั่นคือ แบบสอบร่วมที่ยาวกว่าให้ค่าประสิทธิภาพสัมพัทธ์ที่สูงกว่า

3. ผลการประเมินคุณภาพของการเทียบมาตราเชิงเส้นตรง โดยใช้ค่าดัชนีความแตกต่าง (ดัชนี C) ซึ่งได้จากการวิเคราะห์ข้อมูลของกลุ่มสอบทานผล เมื่อใช้แบบสอบร่วมขนาดต่างกัน คือ

3.1 ค่าดัชนีความแตกต่างจากการเทียบมาตราเชิงเส้นตรง เมื่อใช้แบบสอบร่วมขนาด 10 15 20 และ 25 ข้อ มีค่าเท่ากับ .4319 .3886 .3630 และ .3354 ตามลำดับ ค่าดัชนีความแตกต่างที่ได้จากการใช้แบบสอบร่วมที่ยาวกว่าจะมีค่าน้อยกว่า เมื่อนำดัชนีแต่ละค่าที่ได้มาประเมินด้วยเกณฑ์ของบีเตอร์เซนและคณะ พบว่าผลการเทียบมาตราจากการใช้แบบสอบร่วม 4 ขนาด เมื่อเทียบกับเกณฑ์ให้ผลอยู่ในระดับที่น่าพอใจ

3.2 เปรียบเทียบความแตกต่างของค่าดัชนีความแตกต่าง ซึ่งเป็นค่าความคลาดเคลื่อนรวมในการเทียบมาตราของแบบสอบร่วม 4 ขนาด พบว่าแตกต่างอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 สอดคล้องกับสมมุติฐานข้อที่สองที่กำหนดไว้ แสดงว่าแบบสอบร่วมขนาด 10 15 20 และ 25 ข้อ ให้ความคลาดเคลื่อนในการเทียบมาตราไม่แตกต่างกัน

การอภิปรายผลการวิจัย

1. การเทียบมาตราเชิงเส้นตรงในแต่ละความยาวของแบบสอบร่วม เพื่อปรับคะแนนที่ได้จากแบบสอบต่างชุด ให้อยู่ในมาตราคะแนนเดียวกัน โดยเทียบมาตราจากแบบสอบฉบับใหม่ (ฉบับ Y) ไปสู่มาตราคะแนนของแบบสอบฉบับเก่า (ฉบับ X) ผลของการเทียบมาตราจะได้คะแนนสมมูลส่วนใหญ่จากแบบสอบฉบับ Y ต่ำกว่าคะแนนของแบบสอบฉบับ X เมื่อวิเคราะห์ในเชิงเหตุผล พบว่าแบบสอบเทียบมาตราทั้งสองฉบับนั้นสร้างจากจุดประสงค์ชุดเดียวกัน ดำเนินการสอบกับกลุ่มตัวอย่างที่สุ่มจากประชากรเดียวกัน แต่ได้ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของแบบสอบฉบับ Y ต่ำกว่าฉบับ X เมื่อนำมาเทียบมาตราเชิงเส้นตรงซึ่งเป็นการแปลงคะแนนของแบบสอบสองชุดให้เท่าเทียมกัน ถ้าต่างก็ตรงกับคะแนนมาตรฐานค่าเดียวกัน (Angoff 1971: 564; Petersen and others 1982: 73) นั่นคือเป็นการแปลงคะแนนของแบบสอบสองชุดที่ตอบสนองต่อจำนวนเดียวกันของส่วนเบี่ยงเบนมาตรฐานที่อยู่เหนือ หรือต่ำกว่าค่าเฉลี่ย (Petersen and others 1989: 251) โดยการวิเคราะห์คะแนนทั้งหมด (Angoff 1971: 564) และเป็นการแปลงคะแนนอย่างคงที่ ตลอดการแจกแจง (Flanagan 1951: 747-748) การแปลงคะแนนเชิงเส้นตรงจึงเป็นการเปลี่ยนจุดเริ่มต้นของมาตราคะแนน และหน่วยการวัด ซึ่งเป็นการปรับค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน คะแนนที่แปลงแล้วจะไม่เปลี่ยนรูปร่างในการแจกแจงคะแนน

คะแนนสมมูลที่ได้จากการเทียบมาตราในแต่ละความยาวของแบบสอบร่วม จึงแปรเปลี่ยนตามค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานที่ปรับแล้วของแบบสอบแต่ละชุด ได้คะแนนสมมูลส่วนใหญ่จากคะแนนแบบสอบฉบับ Y ต่ำกว่าคะแนนจากแบบสอบ X สอดคล้องกับผลการเทียบมาตรา จากงานวิจัยของภาวิณี ศรีสุขวัฒนานันท์ (2529: 155-170) และเรวดี อินทสะระ (2531) ซึ่งเทียบมาตราจากแบบสอบฉบับ Y ที่มีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานสูงกว่าไปยังแบบสอบฉบับ X พบว่าคะแนนสมมูลจากแบบสอบฉบับ Y สูงกว่าคะแนนของแบบสอบฉบับ X ในทางกลับกันหากเทียบมาตราจากแบบสอบฉบับ Y ซึ่งมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานต่ำกว่าไปยังแบบสอบฉบับ X จะได้คะแนนสมมูลจากแบบสอบฉบับ Y ต่ำกว่าคะแนนของแบบสอบฉบับ X เช่นกัน

2. ในการตรวจสอบคุณภาพของเทียบมาตรา โดยการประเมินความแม่นยำของการเทียบมาตรา เมื่อใช้แบบสอบร่วมขนาดต่าง ๆ กัน ด้วยการวิเคราะห์ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรา (SEE) นั้น ค่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตราในแต่ละความยาวของแบบสอบร่วม จะแตกต่างกันตามระดับคะแนน (Brennan & Kolen 1987: 251) และมีลักษณะใกล้เคียงกันคือช่วงคะแนนต่ำและสูง จะได้ค่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตราสูงกว่าช่วงคะแนนตรงกลาง ซึ่งสอดคล้องกับงานวิจัยที่ ชูชีพ พงษ์สมบูรณ์ (2528: 117) ได้ข้อสังเกตไว้ว่า ถ้ากลุ่มตัวอย่างความสามารถต่ำจะได้ค่าความคลาดเคลื่อนมาตรฐานสูง และถ้ากลุ่มตัวอย่างมีความสามารถสูง จะได้ค่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตราสูงเช่นกัน แต่กลุ่มที่มีความสามารถปานกลาง เมื่อนำมาเทียบมาตรา จะได้ค่าความคลาดเคลื่อนมาตรฐานต่ำ เมื่อเปรียบเทียบปริมาณความคลาดเคลื่อนมาตรฐาน ณ จุดค่าคะแนนมาตรฐานเดียวกันของ Y แบบสอบร่วมที่ยาวกว่าให้ค่าความคลาดเคลื่อนมาตรฐานน้อยกว่า ผลการวิจัยเป็นไปตามสมมุติฐานข้อที่หนึ่ง ซึ่งสอดคล้องกับงานวิจัยของภาวิณี ศรีสุขวัฒนานันท์ (2529: 159) ทั้งนี้ด้วยเหตุผลที่ว่า แบบสอบร่วมที่ใช้ในการเทียบมาตราจะทำหน้าที่ลดความคลาดเคลื่อนในการเทียบมาตรา ซึ่งเป็นผลจากความแตกต่างระหว่างกลุ่มผู้สอบและแบบสอบ (Petersen and Cook 1987: 233) แบบสอบร่วมที่ยาวกว่าจึงให้ความแม่นยำในการเทียบมาตรามากกว่า (Klein and Kolen 1985 cited by Petersen and Cook 1987) และเมื่อพิจารณาค่าเฉลี่ยของประสิทธิภาพสัมพัทธ์ของแบบสอบร่วมที่ยาวกว่าให้ค่าประสิทธิภาพสัมพัทธ์ของการเทียบมาตราสูงกว่า และค่าเฉลี่ยของประสิทธิภาพสัมพัทธ์เมื่อใช้แบบสอบร่วมขนาด 10 15 และ 20 ข้อ ให้ความแม่นยำของการเทียบมาตราเมื่อเทียบกับแบบสอบร่วม 25 ข้อ อยู่ในช่วงร้อยละ 86.26-93.30 ถือว่าอยู่ในเกณฑ์ค่อนข้างสูง เมื่อเทียบกับค่าประสิทธิภาพสัมพัทธ์ของการเทียบมาตราเชิงเส้นตรงในแบบสอบวัดผลสัมฤทธิ์ ซึ่งภาวิณี ศรีสุขวัฒนานันท์ (2529) ได้วิจัยไว้พบว่าแบบสอบร่วมขนาด 7 14 ข้อ ให้ความแม่นยำในการเทียบมาตรา เมื่อเทียบกับแบบสอบร่วมขนาด 21 ข้อ ร้อยละ 75.97 และ 87.81 ซึ่งเป็นไปได้ว่าความยาวของข้อสอบร่วมตั้งแต่ 10 ข้อขึ้นไป สามารถให้สารสนเทศอย่างเพียงพอที่จะไปปรับความแตกต่างของกลุ่มผู้สอบและแบบสอบที่ต่างกัน

3. ในการตรวจสอบคุณภาพของการเทียบมาตราเชิงเส้นตรง จากค่าดัชนีความแตกต่างที่วิเคราะห์โดยกลุ่มสอบทานผล เมื่อใช้แบบสอบร่วมขนาดต่างกัน ซึ่งประเมินโดยใช้เกณฑ์ของปีเตอร์เซนและคณะ (Petersen and others 1982: 93-94) พบว่าค่าความคลาดเคลื่อนรวมของการเทียบมาตราเชิงเส้นตรง เมื่อใช้แบบสอบร่วมทั้ง 4 ขนาด มีค่าความคลาดเคลื่อนเมื่อเทียบกับเกณฑ์แล้วน้อยกว่ากำลังสองของร้อยละ 10 ของส่วนเบี่ยงเบนมาตรฐานของแบบสอบฉบับ X ซึ่งอยู่ในระดับที่น่าพอใจทั้งสิ้น เมื่อวิเคราะห์ในเชิงเหตุผล พบว่าในการเทียบมาตราใช้กลุ่มตัวอย่างสมมูลสองกลุ่ม ที่ได้จากการสุ่มจากประชากรเดียวกัน ถึงแม้แบบสอบที่นำมาเทียบมาตราจะเป็นแบบสอบที่สร้างในแต่ละปีการศึกษา แต่ก็มีความเที่ยงอยู่ในเกณฑ์สูง วัดในเนื้อหา และจุดประสงค์ชุดเดียวกัน ทำให้การแจกแจงคะแนนของแบบสอบสองชุดมีลักษณะคล้ายคลึงกัน เป็นไปตามนิยามการเทียบมาตราเชิงเส้นตรงที่กำหนดไว้ ผลการเทียบมาตราเชิงเส้นตรงจึงมีคุณภาพอยู่ในระดับที่น่าพอใจ สอดคล้องกับงานวิจัยของภาวิณี ศรีสุขวัฒนานันท์ (2529: 160) ซึ่งพบว่าการเทียบมาตราเชิงเส้นตรงในแบบสอบวัดผลสัมฤทธิ์เมื่อใช้แบบสอบร่วมทั้ง 3 ขนาด ให้คุณภาพของการเทียบมาตราอยู่ในระดับที่น่าพอใจ และจากงานวิจัยของโคเลนและวิทนีย์ (Kolen and Whitney 1982: 279-293) ซึ่งพบว่าในการเทียบมาตราตามแนวระนาบซึ่งกลุ่มตัวอย่างสุ่มสมมูลแล้ว รูปแบบการเทียบมาตราเชิงเส้นตรงเป็นรูปแบบที่เหมาะสม และเมื่อพิจารณากลุ่มสอบทานผล ซึ่งเป็นกลุ่มตัวอย่างที่ใช้ในการประเมินคุณภาพของการเทียบมาตราเชิงเส้นตรง ซึ่งโคเลนและวิทนีย์ (Kolen and Whitney 1982) ได้แนะนำให้ใช้ข้อมูลคะแนนจากผู้สอบเป็นเกณฑ์ในการหาความแตกต่างนั้น พบว่าสุ่มมาจากประชากรเดียวกันกับกลุ่มตัวอย่างเทียบมาตราในจำนวนไม่น้อยกว่า 10% ของกลุ่มตัวอย่างเทียบมาตรา ซึ่งเป็นจำนวนที่เหมาะสม (Kolen and Whitney 1982: 279-293) และได้รับการทดสอบจากแบบสอบเทียบมาตราทั้งสองฉบับโดยสุ่มอย่างมีระบบ ให้ผู้สอบแต่ละคนถ้าสอบแบบสอบฉบับ X แล้วตามด้วยฉบับ Y หรือสอบด้วยฉบับ Y แล้วตามด้วยฉบับ X เป็นการป้องกันความคลาดเคลื่อนในการสอบ เพื่อให้คะแนนบนแบบสอบทั้งสองฉบับได้รับผลกระทบอย่างเท่าเทียมกันในปัจจัยของการเรียนรู้และความเหนื่อยล้า ในการวิเคราะห์ข้อมูลพบว่าค่าสถิติพื้นฐานของแบบสอบในฉบับ X สูงกว่าฉบับ Y ในทุกความยาวของแบบสอบร่วม ซึ่งเป็นลักษณะที่สอดคล้องกับค่าสถิติพื้นฐานของกลุ่มตัวอย่างเทียบมาตรา จะเห็นได้ว่าการดำเนินการเทียบมาตราและสอบทานผลอยู่ในลักษณะของการออกแบบที่ดี ทำให้ผลการเทียบมาตราเชิงเส้นตรงมีคุณภาพอยู่ในระดับที่น่าพอใจ

4. ดัชนีความแตกต่างของแบบสอบร่วมที่ยาวกว่า มีค่าน้อยกว่าดัชนีความแตกต่างของแบบสอบร่วมที่สั้นกว่า เมื่อนำค่าดัชนีความแตกต่างมาเปรียบเทียบกัน พบว่าแตกต่างอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 ผลการวิจัยเป็นไปตามสมมติฐานข้อที่สอง ที่ว่าค่าดัชนีความแตกต่างในการเทียบมาตราเชิงเส้นตรง ของแบบสอบร่วมตั้งแต่ 10 ข้อขึ้นไป น่าจะแตกต่างอย่างไม่มีนัยสำคัญทางสถิติ อย่างไรก็ตามข้อค้นพบนี้ยังไม่อาจสรุปเป็นการทั่วไปได้ เนื่องจากสถานการณ์ของการวิจัยครั้งนี้ ใช้แบบสอบร่วมที่สร้างจากเนื้อหาเดียวกันกับแบบสอบเทียบมาตราประกอบกับแบบสอบเทียบมาตรามีความยาวเพียง 60 ข้อ ถ้าแบบสอบเทียบมาตรามีความยาวเพิ่มขึ้น แบบสอบร่วมขนาด 10 ข้อ อาจจะไม่เพียงพอ ทั้งนี้ยังขาดข้อมูลเชิงประจักษ์สนับสนุนในส่วนนี้

<u>ข้อเสนอแนะในการนำไปใช้</u>

1. ในการเทียบมาตราเชิงเส้นตรง ถึงแม้ว่าแบบสอบร่วมขนาดยาวกว่าจะให้ความแม่นยำในการเทียบมาตรามากกว่า แต่ในทางปฏิบัติ ต้องคำนึงถึงเวลาที่ใช้สอบ การจดจำข้อสอบร่วมได้ การเป็นความลับของแบบสอบ ดังนั้น แบบสอบร่วมขนาด 10 ข้อ ซึ่งมีความแม่นยำของเทียบมาตราอยู่ในระดับที่น่าพอใจ ไม่แตกต่างจากข้อสอบขนาด 25 ข้อ จึงสมควรใช้แบบสอบร่วมที่มีขนาดสั้นกว่า กับแบบสอบเทียบมาตรา ซึ่งเป็นแบบสอบวัดผลสัมฤทธิ์ ที่มีความยาวไม่เกินฉบับละ 60 ข้อ

2. การประเมินผลสัมฤทธิ์ของหลักสูตร ในการสอบวัดผลสัมฤทธิ์ทางการเรียนของนักเรียน แบบสอบที่ใช้กับนักเรียนแต่ละปีการศึกษา มักมีการเปลี่ยนแปลงในตัวแบบสอบ แต่เนื้อหาหลักสูตรยังคงเดิม การประเมินโดยการใช้คะแนนดิบของแบบสอบแต่ละครั้ง ไม่สามารถนำมาเปรียบเทียบข้ามรุ่นกันได้ เพราะคะแนนดิบจากแบบสอบต่างชุดจะอยู่ในสเกลที่ต่างกัน หากนำเทคนิคการเทียบมาตราโดยใช้แบบสอบร่วมมาใช้กับการสอบในตอนปลายภาคทุกปี ผลการสอบนี้จะนำไปใช้ในการประเมินผลอย่างมีมาตรฐาน และสามารถควบคุมการตัดสินได้-ตก ที่ตรงกับความสามารถของผู้สอบได้อย่างมีประสิทธิภาพกว่า

3. ในการสอบวัดผลสัมฤทธิ์ปลายภาคนั้น กลุ่มโรงเรียนได้ร่วมกันออกข้อสอบเพื่อนำมาใช้กับโรงเรียนภายในกลุ่ม หากมีการสร้างแบบสอบวัดผลสัมฤทธิ์ปลายภาคมากกว่าหนึ่งฉบับ โดยมีแบบสอบร่วมในแต่ละฉบับ โรงเรียนภายในกลุ่ม อาจใช้แบบสอบต่างฉบับกัน การวัดผลภายในกลุ่มสามารถตัดสินผลการสอบของนักเรียนที่สอบแบบสอบต่างฉบับกันได้ ด้วยการเทียบมาตราโดยใช้แบบสอบร่วม

ข้อเสนอแนะในการวิจัย

1. ควรจะมีการศึกษาถึงขนาดความยาวของแบบสอบเทียบมาตรา ว่ามีผลต่อความแม่นยำของการเทียบมาตราอย่างไร รวมทั้งศึกษาถึงจำนวนข้อที่เหมาะสมของแบบสอบร่วมตามขนาดความยาวของแบบสอบเทียบมาตราที่ต่างกันไป

2. ควรศึกษาถึงความเป็นไปได้ในการเทียบมาตราโดยใช้แบบสอบอัตนัย และจำนวนข้อของแบบสอบร่วมที่เหมาะสมในการนำไปใช้กับแบบสอบเทียบมาตราที่เป็นแบบสอบอัตนัย

3. ควรจะได้มีการตรวจสอบความยาวของแบบสอบร่วมภายใน ที่ใช้ในการเทียบมาตราเชิงเส้นตรง เมื่อกลุ่มตัวอย่างไม่ใช่กลุ่มสุ่ม ซึ่งเป็นรูปแบบที่ใช้ในกลุ่มโรงเรียนตามสภาพการเรียนการสอนปัจจุบัน โดยออกแบบให้มีแบบสอบร่วมในแบบสอบวัดผลสัมฤทธิ์ปลายภาคของกลุ่มโรงเรียนในแต่ละปีการศึกษา

4. ควรจะมีการศึกษาวิธี Pre-equating ซึ่งเป็นเทคนิคใหม่ที่มีการนำเสนอจากนักวัดผล เพื่อเป็นข้อมูลสำหรับการตัดสินใจเลือกใช้แบบสอบต่าง ๆ เนื่องจากเป็นวิธีการเทียบมาตราโดยอาศัยทฤษฎีการตอบข้อสอบในการหาค่าประจำข้อสอบ (parameter) ไว้ก่อน แล้วสร้างตารางแปลงคะแนนสมมูลจากแบบสอบชุดหนึ่งไปยังอีกชุดหนึ่ง การใช้วิธีการนี้จึงควรมีคลังข้อสอบที่เคยผ่านการทดลองใช้มาแล้ว