



บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบคุณภาพของมาตรประมาณค่าเชิงพฤติกรรม (Behaviorally-Anchored Rating Scales) กับมาตรประมาณค่าแบบกราฟิก (Graphic Rating Scales) ในด้านความตรง ความเที่ยง และลักษณะความคลาดเคลื่อนในการประมาณค่า ในสถานการณ์การประเมินพฤติกรรมด้านการประเมินผลของครู ในการดำเนินการวิจัย ได้กำหนดเป้าหมายของการประเมิน ดำเนินการพัฒนาเครื่องมือ เก็บรวบรวมข้อมูล วิเคราะห์ข้อมูล และนำเสนอข้อมูลตามแนวทางดังนี้

เป้าหมายของการประเมิน

1. พฤติกรรมเป้าหมายในการศึกษาครั้งนี้ คือ พฤติกรรมด้านการประเมินผลของครู ซึ่งประกอบด้วยพฤติกรรมย่อย 4 ด้าน คือ (1) การวางแผนการวัดผล-ประเมินผล (2) การวัดผล-ประเมินผลในชั้นสอน หรือระหว่างดำเนินการสอน (3) การวัดผล-ประเมินผลในชั้นสรุปผลการเรียน และ (4) การตรวจสอบและพัฒนาระบบการวัดผล-ประเมินผลของตนเอง

2. กลุ่มตัวอย่างที่เป็นเป้าหมายของการประเมินเพื่อการตรวจสอบคุณภาพของมาตรวัด (Subjects for Scale Evaluation/ Target Ratee) ประกอบด้วย 2 กลุ่ม คือ

2.1 กลุ่มครุคณิตศาสตร์ ในโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ในเขตกรุงเทพมหานคร ที่สอนระดับชั้นมัธยมศึกษาตอนต้น ในภาคต้น ปีการศึกษา 2533 คัดเลือกโดยการสุ่มอย่างง่ายจากครุคณิตศาสตร์ที่สอนในระดับชั้นมัธยมศึกษาปีที่ 3 โรงเรียนละ 1 คน รวมกลุ่มเป้าหมายที่คาดหวัง 103 คน ถือเป็นกลุ่มเป้าหมายที่เป็นตัวแทนของครุคณิตศาสตร์ที่สอนระดับชั้นมัธยมศึกษาตอนต้น ในการศึกษาครั้งนี้ โดยได้ใช้ข้อมูลผลการประเมินพฤติกรรมของครูกลุ่มนี้ในการวิเคราะห์ความตรงชนิด Convergent and Discriminant validity ความเที่ยงระหว่างผู้ประเมิน (Rater Reliability) และลักษณะความคลาดเคลื่อนในการประมาณค่า (Error)

2.2 กลุ่มครุคณิตศาสตร์ที่สอนในระดับชั้นมัธยมศึกษาตอนต้น ในภาคต้น ปีการศึกษา 2533 ที่ได้รับการระบุจากหัวหน้าหมวดวิชาและศึกษานิเทศก์ภาควิชาคณิตศาสตร์ หน่วยงานนิเทศก์ กรมสามัญศึกษา ว่า เป็นครูที่จัดกิจกรรมการเรียนการสอนอย่างเป็นระบบมีประสิทธิภาพ จำนวน 30 คน และ กลุ่มครุคณิตศาสตร์ที่ได้รับการระบุจากหัวหน้าหมวดวิชาว่า ปฏิบัติงานยังไม่เป็นระบบ-ขาดประสิทธิภาพ 30 คน ถือเป็นตัวแทนของกลุ่มครูที่ปฏิบัติงานอย่างมีประสิทธิภาพ และด้อยประสิทธิภาพ ตามลำดับ รวมกลุ่มเป้าหมายที่คาดหวัง 60 คน กลุ่มเป้าหมายกลุ่มนี้ได้ใช้เป็นกลุ่มตัวอย่างสำหรับการตรวจสอบความตรงของมาตรวัดในการจำแนกความแตกต่างระหว่างกลุ่มครูที่ปฏิบัติงานมีประสิทธิภาพต่างกัน

กลุ่มตัวอย่างผู้ประเมิน (Raters)

ผู้ประเมินในการศึกษานี้มี 2 กลุ่ม คือ (1) กลุ่มครุคณิตศาสตร์ที่เป็นกลุ่มเป้าหมายของการประเมินทำหน้าที่ประเมินพฤติกรรมของตนเอง เป็นสถานการณ์ของการประเมินตนเองและ (2) หัวหน้าหมวดวิชา หรือผู้บังคับบัญชาชั้นต้นของกลุ่มเป้าหมาย ทำหน้าที่ประเมินพฤติกรรมของครูกลุ่มเป้าหมายอีกชั้นหนึ่ง ถือเป็นสถานการณ์ของการประเมินโดยผู้บังคับบัญชา

การคัดเลือกตัวอย่างผู้ประเมิน จำนวนผู้ประเมินที่คาดหวัง และจำนวนผู้ประเมินที่ให้ข้อมูลจริง

ในการคัดเลือกกลุ่มผู้ประเมิน ได้ดำเนินการดังนี้

1. กลุ่มผู้ประเมินประเภทครูผู้สอนทั่วไป ใช้วิธีสุ่มอย่างง่ายจากครุคณิตศาสตร์ที่สอนระดับชั้นมัธยมศึกษาปีที่ 3 โรงเรียนละ 1 คน รวมกลุ่มตัวอย่างที่คาดหวัง 103 คน
2. กลุ่มผู้ประเมินประเภทครูผู้สอนที่ปฏิบัติงานเป็นระบบ คัดเลือกโดยให้ศึกษานิเทศก์ภาควิชาคณิตศาสตร์ หน่วยงานนิเทศก์ กรมสามัญศึกษา และหัวหน้าหมวดวิชาคณิตศาสตร์ระบุรายชื่อครุคณิตศาสตร์ที่ปฏิบัติงานอย่างเป็นระบบคือ (1) มีการวางแผนการวัดผล-ประเมินผลอย่างดี เช่น จัดทำแผนการสอน แผนการประเมินผล ปฏิทินการวัดผล เป็นต้น (2) มีกลยุทธ์ที่ดีในการวัดผล-ประเมินผลการเรียน ในชั้นสอน เช่น มีการค้นหาแบบทดสอบย่อย ๆ เพื่อตรวจสอบความสามารถของผู้เรียนในแต่ละเนื้อหา จัดสอนซ่อมเสริมอย่างจริงจัง เป็นต้น

และ (3) มีความกระตือรือร้นที่จะพัฒนาระบบวัดผล เช่น ศึกษาหาความรู้เพิ่มเติมอยู่เสมอ มีการวิเคราะห์คุณภาพของข้อสอบ เป็นต้น แล้วคัดเลือกครูที่ได้รับการระบอบอย่างสอดคล้องกัน ถือเป็นกลุ่มครูที่ปฏิบัติงานเป็นระบบ ส่วนกลุ่มครูที่ปฏิบัติงานยังไม่เป็นระบบ-ขาดประสิทธิภาพและควรได้รับการปรับปรุงนั้น ให้อยู่ในดุลยพินิจของหัวหน้าหมวดวิชาในโรงเรียนที่มีครูประเภทปฏิบัติงานเป็นระบบเป็นผู้คัดเลือก กล่าวโดยสรุปก็คือ โรงเรียนใดที่ได้รับการคัดเลือกครูประเภทปฏิบัติงานเป็นระบบ 1 คน หัวหน้าหมวดวิชาคณิตศาสตร์ในโรงเรียนนั้นจะต้องทำการคัดเลือกครูคณิตศาสตร์ที่ปฏิบัติงานไม่เป็นระบบหรือด้อยประสิทธิภาพกว่าเพิ่มอีก 1 คน รวมกลุ่มตัวอย่างผู้ประเมินตนเองที่คาดหวังในกลุ่มนี้จำนวน 60 คน เป็นผู้ปฏิบัติงานเป็นระบบ 30 คน และปฏิบัติงานยังไม่เป็นระบบ 30 คน

3. กลุ่มผู้ประเมินประเภทหัวหน้าหมวดวิชา คือเอาหัวหน้าหมวดวิชาคณิตศาสตร์ในโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ในเขตกรุงเทพมหานครทุกโรงเรียน เป็นผู้ประเมินพฤติกรรมของครูที่อยู่ในความดูแลของตนเอง รวมกลุ่มตัวอย่างที่คาดหวัง 103 คน

จำนวนตัวอย่างผู้ประเมิน (Raters) ที่ตอบและส่งแบบประเมินกลับคืน และถือเป็นกลุ่มผู้ประเมินจริงในการศึกษาครั้งนี้ ปรากฏดังตารางที่ 1

ตารางที่ 1 จำนวนผู้ประเมินประเภทต่าง ๆ ที่ตอบแบบประเมินกลับคืน และสัดส่วนร้อยละ เมื่อเทียบกับจำนวนที่คาดหวัง

ประเภทของผู้ประเมิน	จำนวน		
	ที่คาดหวัง	ตอบจริง	ร้อยละ
1. ครูผู้สอนทั่วไป	103	97	94.17
2. หัวหน้าหมวดวิชา	103	94	91.26
3. ครูที่ปฏิบัติงานเป็นระบบ	30	30	100.00
4. ครูที่ปฏิบัติงานยังไม่เป็นระบบ	30	30	100.00
รวม	266	249	93.60

จากตารางที่ 1 พบว่า สัดส่วนผู้ส่งแบบสอบถามกลับคืนเมื่อเทียบกับจำนวนที่คาดหวัง อยู่ในระดับเกินร้อยละ 90 ทุกรายการ ซึ่งแสดงถึงความร่วมมือของกลุ่มผู้ประเมินว่าได้ให้ความร่วมมือในระดับที่ดีมาก อย่างไรก็ตาม จำนวนที่นำเสนอนี้เป็นจำนวนสูงสุดของผู้ตอบแบบประเมินโดยภาพรวมเท่านั้น ในระดับรายละเอียดจะพบว่า ผู้ตอบแบบประเมินบางท่านได้ตอบแบบประเมินเพียง 1 ชุด หรือ 1 รูปแบบเท่านั้น ดังนั้นจำนวนแบบประเมินที่ได้รับกลับคืนเมื่อจำแนกตามประเภทของแบบประเมินแล้ว จะน้อยกว่าจำนวนที่นำเสนอในตารางนี้ (รายละเอียดจำนวนแบบประเมินแต่ละรูปแบบที่ได้รับกลับคืนปรากฏดังตารางที่ 3)

สถานภาพของผู้ตอบแบบประเมินกลุ่มต่าง ๆ จำแนกตามตัวแปร เพศ อายุราชการ และ ประสบการณ์ในการวิจัย ปรากฏดังตารางที่ 2

ตารางที่ 2 สถานภาพของผู้ตอบแบบประเมินกลุ่มต่าง ๆ จำแนกตามตัวแปร เพศ อายุราชการ และประสบการณ์ในการวิจัย

รายการตัวแปรสถานภาพ	ครูผู้สอนทั่วไป		ครูที่ปฏิบัติงาน เป็นระบบ		ครูที่ปฏิบัติงาน ยังไม่เป็นระบบ		หัวหน้าหมวดวิชา		รวม	
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ
1. เพศ.....ชาย.....	24	24.74	7	23.33	13	43.33	35	37.23	79	31.47
หญิง.....	73	75.26	23	76.67	17	56.67	59	62.77	172	68.53
2.อายุราชการ...1 - 5 ปี...	3	3.09	-	-	1	3.33	-	-	4	1.59
6 - 10 ปี...	14	14.43	6	20.00	3	10.00	2	2.13	25	9.96
11 - 15 ปี...	46	47.44	14	46.67	19	63.34	92	34.04	111	44.23
16 - 20 ปี...	24	24.74	6	20.00	6	20.00	27	28.72	63	25.10
21 - 25 ปี...	4	4.12	1	3.33	1	3.33	10	10.64	16	6.37
26 - 30 ปี...	3	3.09	2	6.67	-	-	9	9.58	14	5.58
เกิน 30 ปี...	3	3.09	1	3.33	-	-	14	14.89	18	7.17
3. ประสบการณ์ในการวิจัย										
เคยมีผลงานวิจัย	23	23.72	8	26.67	2	6.67	34	36.17	67	26.69
ไม่เคยทำการวิจัย	74	76.28	22	73.33	28	93.33	60	63.83	184	73.31

จากตารางที่ 2 พบว่า กลุ่มผู้ตอบแบบประเมินทุกประเภท ส่วนใหญ่เป็นครูผู้หญิง ในด้านอายุราชการของกลุ่มตัวอย่างโดยภาพรวม พบว่า ส่วนใหญ่มีอายุราชการระหว่าง 11-20 ปี (ร้อยละ 69.33) เมื่อพิจารณาเป็นรายกลุ่ม พบว่า มากกว่าร้อยละ 50 ของครูผู้สอนทั่วไปมีอายุราชการไม่เกิน 15 ปี ในขณะที่หัวหน้าหมวดวิชาร้อยละ 63.83 มีอายุราชการมากกว่า 15 ปี ในจำนวนนี้ ร้อยละ 14.89 มีอายุราชการมากกว่า 30 ปี ในส่วนของประสบการณ์ในการวิจัยนั้น พบว่า โดยภาพรวมแล้วมีจำนวนไม่มากนักที่มีประสบการณ์ในการวิจัย กลุ่มหัวหน้าหมวดวิชาเป็นกลุ่มที่มีประสบการณ์ในการวิจัยมากที่สุดเมื่อเทียบกับกลุ่มอื่น ๆ (ร้อยละ 36.17) เมื่อเปรียบเทียบระหว่างกลุ่มครูที่ปฏิบัติงานเป็นระบบกับกลุ่มที่ปฏิบัติงานยังไม่เป็นระบบ พบว่า กลุ่มที่ปฏิบัติงานเป็นระบบมีส่วนผู้ที่มีประสบการณ์ในการวิจัยสูงกว่า

การพัฒนามาตรวัด

ในการพัฒนามาตรวัด ได้ดำเนินการพัฒนามาตรวัด 2 รูปแบบคือ (1) มาตรฐานค่าเชิงพฤติกรรม และ (2) มาตรฐานค่าแบบกราฟิก โดยได้ดำเนินการควบคู่กันไปทั้งสองรูปแบบ แนวทางการดำเนินการและผลการดำเนินการในแต่ละขั้นตอน เป็นดังนี้

1. กำหนดกรอบพฤติกรรมที่มุ่งประเมิน

การดำเนินการในขั้นตอนนี้ ผู้วิจัยได้ศึกษาบทบาทหน้าที่ที่ควรจะเป็นในด้านการวัดผล-ประเมินผลการเรียนของครูผู้สอนจากงานวิจัย ข้อเสนอแนะ คู่มือครู ระเบียบในการวัดผล-ประเมินผลการเรียน ฯลฯ แล้วกำหนดกรอบรายการพฤติกรรมที่มุ่งประเมิน ผลการดำเนินการปรากฏว่า ได้กรอบรายการพฤติกรรม 4 ด้านย่อย คือ (1) พฤติกรรมด้านการวางแผนการวัดผล-ประเมินผล (2) พฤติกรรมการวัดผล-ประเมินผลในชั้นสอน (3) พฤติกรรมการวัดผล-ประเมินผลในชั้นสรุปผลการเรียน และ (4) พฤติกรรมการ ตรวจสอบและพัฒนาระบบการวัดผล-ประเมินผลของตนเอง รายละเอียดผลการศึกษาในขั้นตอนนี้ ได้สรุปไว้ในบทที่ 2 ตอนที่ 6

2. สำรวจรายการพฤติกรรมการปฏิบัติงานที่ควรจะเป็นตามการรับรู้ของครู

ผู้วิจัยได้สร้างแบบสอบถามปลายเปิดโดยยึดกรอบรายการพฤติกรรมในข้อ 1. เป็นกรอบเนื้อหาของแบบสอบถาม (ดูตัวอย่างแบบสอบถามในภาคผนวก ก.) ส่งแบบสอบถามให้หัวหน้าหมวดวิชา ครูผู้สอน และครูวัดผลจำนวน 56 คน ให้ความเห็น หรือเสนอตัวอย่างพฤติกรรมด้านการประเมินผลของครูซึ่งที่เห็นว่าเป็นพฤติกรรมที่มีประสิทธิภาพ และด้วยประสิทธิภาพที่ส่งผลต่อคุณภาพการเรียนการสอน การดำเนินการในชั้นตอนนั้นนอกจากจะเป็นการดำเนินการตามแนวทางการพัฒนามาตรฐานค่าเชิงพฤติกรรมตามข้อเสนอแนะของ Landy และ Schwab (1974) แล้ว ยังจะเกิดผลดีในชั้นของการประเมินซึ่ง Pierson (1975 : 2718-A) กล่าวว่า การดำเนินการในลักษณะนี้เป็นการกำหนดกรอบโดเมนของถ้อยคำหรือภาษาที่ใช้ให้เป็นที่สื่อความหมายในระหว่างกลุ่มผู้ประเมินและผู้ถูกประเมินได้เป็นอย่างดี ทั้งนี้เพราะว่ารายการพฤติกรรมที่นำไปบรรจุไว้ในมาตรวัดจะเป็นรายการพฤติกรรมที่เป็นที่รับรู้ร่วมกันของกลุ่มบุคคลในสังคมที่กำลังทำการประเมิน

การเก็บรวบรวมข้อมูลในชั้นตอนนี้ ผู้วิจัยได้ดำเนินการในเดือนมิถุนายน 2533 โดยเก็บรวบรวมข้อมูลจากคณะครู-อาจารย์ที่เข้าร่วมประชุมสัมมนาเพื่อพัฒนาวัตรกรรมการเรียนการสอน ณ หอประชุมคุรุสภา กรุงเทพมหานคร ผลการดำเนินการ ปรากฏว่าได้รับแบบสอบถามกลับคืน 42 ฉบับ หรือร้อยละ 75 ของจำนวนแบบสอบถามที่แจกในที่ประชุม

3. วิเคราะห์ จำแนกกลุ่มพฤติกรรม และตั้งชื่อรายการพฤติกรรมหลัก

ในชั้นตอนนี้ ผู้วิจัยได้ทำการวิเคราะห์เนื้อหาของข้อมูลที่ได้จากแบบสอบถามในข้อ 2. แล้วจัดกลุ่มตัวอย่างพฤติกรรมออกเป็นกลุ่ม ๆ ตั้งชื่อกลุ่มพฤติกรรมโดยยึดกรอบรายการพฤติกรรมหลักที่ได้จากการศึกษาแนวคิดที่ว่าด้วยพฤติกรรมการสอนของครูดังที่ปรากฏในบทที่ 2 ตอนที่ 6 เป็นข้อมูลประกอบการพิจารณาด้วย การดำเนินการในชั้นตอนนี้ นอกจากจะได้ตัวอย่างรายการพฤติกรรมที่จะนำไปจัดทำเป็นมาตรฐานค่าแล้ว ยังได้รายการคำคุณศัพท์เชิงคุณภาพที่แฝงมาในตัวอย่างพฤติกรรมด้วย เช่น ข้อเสนอของครูที่ว่า "ควรกำหนดแนวทางการวัดผล-ประเมินผลไว้อย่างชัดเจน" "ควรกำหนดรายการที่จะวัดรวมทั้งสัดส่วนคะแนนไว้อย่างชัดเจน" ข้อเสนอแนะเหล่านี้ทำให้สามารถสรุปได้ว่า ในเรื่องของการกำหนดแนวทางการวัดผล-ประเมินผลการเรียน ซึ่งเป็นพฤติกรรมหลักรายการหนึ่งในด้านการวางแผนนั้น ควรพิจารณาคุณภาพของพฤติกรรม

หลักนี้ในแง่ที่ว่า "ได้วางแผนไว้อย่าง ชัดเจน-ไม่ชัดเจน เพียงใด" ซึ่งคำคุณศัพท์ "ชัดเจน-ไม่ชัดเจน" ในที่นี้ จะถือเป็นคำคุณศัพท์ที่ใช้ในมาตรฐานค่าแบบกราฟิกต่อไป

ผลการดำเนินการในขั้นตอนนี้ ผู้วิจัยได้คัดเลือกกลุ่มพฤติกรรมหลักไว้ 20 รายการ เป็นรายการพฤติกรรมหลักที่จัดอยู่ในด้านต่าง ๆ คือ ด้านการวางแผนการวัดผล-ประเมินผล การวัดผล-ประเมินผลในชั้นสอน การวัดผล-ประเมินผลในชั้นสรุปผลการเรียน และด้านการตรวจสอบและพัฒนาระบบการวัดผล-ประเมินผลของตนเอง ด้านละ 5 รายการ

4. จัดทำมาตรฐานค่าเชิงพฤติกรรม และจัดทำแบบประเมินพฤติกรรมฉบับร่าง

ผู้วิจัยได้นำรายการพฤติกรรมหลัก พร้อมทั้งพฤติกรรมย่อยในกลุ่มมาจัดทำมาตรฐานค่าเชิงพฤติกรรมในลักษณะของมาตรฐานค่าในแนวตั้งที่ใช้ตัวอย่างพฤติกรรมย่อย ๆ ผูกติดไว้บนมาตรวัดอย่างเรียงลำดับ 5 ระดับ จากมีคุณภาพมากไปน้อย หรือน้อยไปมาก การจัดเรียงลำดับในแต่ละชุดของมาตรวัดอาศัยวิธีการเชิงเหตุผลในการพิจารณาตัดสิน ผลการดำเนินการในขั้นตอนนี้ ได้มาตรฐานค่าเชิงพฤติกรรมจำนวน 20 รายการ แล้วนำมาตรฐานค่าเชิงพฤติกรรมทั้ง 20 รายการนี้ไปบรรจุเป็นเนื้อหาของแบบประเมินพฤติกรรมการปฏิบัติงานด้านการวัดผล-ประเมินผลของครูผู้สอน ถือเป็นแบบประเมินชนิดมาตรฐานค่าเชิงพฤติกรรมฉบับร่าง ขณะเดียวกันก็นำเฉพาะชื่อรายการพฤติกรรมหลักไปบรรจุในแบบประเมินชนิดมาตรฐานค่าแบบกราฟิกฉบับร่าง โดยได้เพิ่มในส่วนของสเกลคำตอบพร้อมทั้งคำคุณศัพท์ที่เห็นว่าเหมาะสมที่ได้จากการดำเนินการในขั้นตอนที่ 3. เข้าไปในมาตรฐานค่ารูปแบบนี้

5. ให้ผู้ทรงคุณวุฒิ/กลุ่มครูที่มีประสบการณ์ตรวจสอบความเหมาะสม-ยืนยันคุณภาพของมาตรฐานค่า และทดลองใช้แบบประเมิน (รอบที่ 1)

ในขั้นตอนนี้ได้ดำเนินการกิจกรรม 2 รายการควบคู่กันไปคือ

5.1 นำรายการชุดมาตรฐานค่าทั้ง 20 รายการที่ได้จากขั้นตอนที่ 4. บรรจุในแบบสอบถามสำหรับผู้ทรงคุณวุฒิ หรือกลุ่มครูที่มีประสบการณ์ในการสอนคณิตศาสตร์-มีประสบการณ์ในการวัดผลประเมินผลการเรียนในระดับมัธยมศึกษาเป็นอย่างดี จำนวน 25 ท่าน พิจารณาความเหมาะสมของชุดมาตรวัดทั้ง 20 รายการ โดยในแต่ละรายการให้พิจารณาตัดสิน 3 ลักษณะ คือ (1) พิจารณาว่าการจัดเรียงลำดับพฤติกรรมในแต่ละรายการหรือแต่ละข้อได้เรียงอย่างเป็น

ลำดับจากมีคุณภาพมากไปน้อย หรือน้อยไปมากอย่างสมเหตุสมผลแล้วหรือไม่ (2) วิจารณ์ความเหมาะสมของคำคุณศัพท์ที่เป็นมิตินคุณภาพของแต่ละข้อ และ (3) วิจารณ์ความเหมาะสมของข้อความถ้อยคำ หรือภาษาที่ใช้ (ดูตัวอย่างแบบสอบถามสำหรับผู้ทรงคุณวุฒิ ในภาคผนวก ข.) โดยในการให้ผู้ทรงคุณวุฒิวิจารณ์ความเหมาะสมของเครื่องมือนี้ ผู้วิจัยได้ไปติดต่อและขอพบผู้ทรงคุณวุฒิทุกท่านด้วยตนเอง ได้ชี้แจงจุดประสงค์ของการวิจัยอธิบายที่มาของชุดมาตราประมาณค่าทั้ง 20 รายการอธิบายแนวทางการดำเนินการในขั้นต่อ ๆ ไป พร้อมทั้งมอบแบบประเมินพฤติกรรมการปฏิบัติงานฉบับร่างทั้ง 2 รูปแบบไว้เพื่อประกอบการพิจารณาด้วย และ

5.2 นำแบบประเมินพฤติกรรมทั้ง 2 รูปแบบ ไปให้หัวหน้าหมวดวิชาคณิตศาสตร์ 5 คน ทดลองใช้ประเมินพฤติกรรมของครู และให้ครู 5 คน ทดลองใช้ประเมินพฤติกรรมของตนเอง โดยผู้วิจัยได้ไปพบครูทั้งสองกลุ่มนี้ด้วยตนเอง ทั้งในวันส่งแบบสอบถามและวันรับคืนทั้งนี้เพื่อประโยชน์ในการซักถามความคิดเห็นเกี่ยวกับแบบประเมินในบางลักษณะ เช่น โดยสรุปอ่านแล้วเข้าใจแจ่มแจ้งหรือไม่ มีประเด็นที่สงสัยหรือไม่ เป็นต้น

6. ปรับปรุงมาตราประมาณค่าและแบบประเมินฉบับร่าง พร้อมทั้งทดลองใช้รอบที่ 2

ผู้วิจัยได้ทำการวิเคราะห์ความคิดเห็นของผู้ทรงคุณวุฒิโดยยึดถือเกณฑ์ร้อยละ 80 หรือ 20 ใน 25 เป็นเกณฑ์ยืนยันความเหมาะสมของมาตรวัด โดยเฉพาะในเรื่องของการเรียงลำดับคุณภาพของตัวอย่างพฤติกรรมที่ผูกติดไว้บนมาตรวัด ผลการวิเคราะห์ข้อมูลพบว่า รายการพฤติกรรมหลัก 2 รายการ คือ ข้อ 7 การตรวจสอบความเข้าใจของนักเรียนขณะทำการสอน และ ข้อ 12 การเลือกจุดประสงค์การเรียนรู้เพื่อตรวจสอบความรู้ในปลายภาคเรียน ทั้ง 2 รายการ มีผู้ทรงคุณวุฒิเพียง 18 และ 17 ท่าน ตามลำดับ เห็นว่าเหมาะสมซึ่งถือว่าไม่ถึงเกณฑ์ร้อยละ 80 เฉพาะข้อ 7 มีผู้เสนอแนะว่าเป็นพฤติกรรมที่สังเกตได้ยาก โดยเฉพาะหัวหน้าหมวดวิชาจะไม่ค่อยมีโอกาสได้เห็นพฤติกรรมของครูในลักษณะนั้น ส่วนในข้อ 12 นั้น มีผู้ทรงคุณวุฒิที่เป็นผู้อำนวยการพิเศษด้านการวัดผลของกรมวิชาการ กระทรวงศึกษาธิการ ได้ให้ข้อคิดเห็นว่า เป็นหน้าที่ของกลุ่มโรงเรียนที่จะคัดเลือกจุดประสงค์การเรียนรู้เพื่อตรวจสอบในปลายภาคเรียน ไม่ใช่หน้าที่ของครูผู้สอนคนใดคนหนึ่งโดยเฉพาะ ด้วยอาศัยเกณฑ์การตัดสินร้อยละ 80 และความคิดเห็นเพิ่มเติมของผู้ทรงคุณวุฒิดังกล่าว ผู้วิจัยจึงได้ตัดรายการพฤติกรรมหลักที่ 7 และ 12 ออกจากชุดมาตราวัด

เหลือรายการพฤติกรรมหลักที่มุ่งประเมินเพียง 18 รายการ ทำให้แบบประเมินสั้นขึ้น ซึ่งเป็นการสนองความคิดเห็นของกลุ่มครูที่เป็นตัวอย่างในการทดลองใช้แบบประเมินด้วย ทั้งนี้เพราะกลุ่มตัวอย่างกลุ่มนี้ได้แสดงความคิดเห็นว่าแบบประเมินชนิดมาตราประมาณค่าเชิงพฤติกรรมค่อนข้างยาว ทำให้ต้องใช้ความพยายามหรือความตั้งใจอย่างมาก จึงจะสามารถตอบแบบประเมินได้แล้วเสร็จ

ในส่วนของรายการพฤติกรรม 18 รายการที่ได้รับการยืนยันในเรื่องของความเหมาะสมจากผู้ทรงคุณวุฒินั้น หลายข้อได้รับการปรับปรุงในบางลักษณะ เช่น ข้อ 1 รายการพฤติกรรมหลัก (Stem) เดิมคือ "การวิเคราะห์หลักสูตรในรายวิชาที่รับผิดชอบ" ซึ่งได้รับข้อท้วงติงจากผู้ทรงคุณวุฒิบางท่านว่า กว้างเกินไป หรือเป็นพฤติกรรมที่ซับซ้อนเกินไป อาจมีผลทำให้เกิดความเข้าใจที่สับสนได้ ดังนั้นจึงได้ปรับปรุงใหม่เป็น "การวิเคราะห์จุดประสงค์การเรียนรู้ในรายวิชาที่สอน" พร้อมทั้งได้เปลี่ยนคำคุณศัพท์จาก "เป็นระบบ-ไม่เป็นระบบ" เป็น "เจาะลึก-ผิวเผิน" แทน หรือในข้ออื่น ๆ ก็ได้มีการปรับปรุงถ้อยคำ ภาษาที่ใช้ อาทิ "การตรวจสอบความสามารถของนักเรียนก่อนสอน" เปลี่ยนเป็น "การประเมินความสามารถขั้นพื้นฐานของนักเรียนก่อนเริ่มการสอน" "การประสานแผน หรือแจ้งเงื่อนไขต่าง ๆ ให้นักเรียนทราบ" เปลี่ยนเป็น "การแจ้งเงื่อนไขต่าง ๆ ให้นักเรียนทราบ" "การให้ข้อมูลป้อนกลับเพื่อสะท้อนความเข้าใจ หรือเสริมความเข้าใจแก่นักเรียนหลังจากที่ได้มอบหมายงานให้ทำ" เปลี่ยนเป็น "การให้ข้อมูลป้อนกลับ (Feedback) ในการตรวจสอบผลงานหรือการบ้าน" เป็นต้น

หลังจากที่ได้ปรับปรุงรายการพฤติกรรม ภาษา หรือถ้อยคำที่ใช้แล้ว ผู้วิจัยได้ปรับปรุงแบบประเมินฉบับร่างอีกครั้งหนึ่ง แล้วนำไปให้ครูผู้สอน 5 คน ทดลองใช้ประเมินพฤติกรรมการปฏิบัติงานของตนเอง ซึ่งในการทดลองใช้รอบที่ 2 นี้ ปรากฏว่าไม่พบประเด็นที่เป็นอุปสรรคที่เด่นชัดแต่ประการใด ยกเว้นในด้านความยาวของแบบประเมินชนิดมาตราประมาณค่าเชิงพฤติกรรมเท่านั้นที่กลุ่มตัวอย่างยังคงแสดงความคิดเห็นว่าค่อนข้างยาว ต้องใช้สมาธิในการอ่านมาก ซึ่งจากการสอบถาม พบว่า ต้องใช้เวลาประมาณ 25-30 นาที จึงจะตอบได้แล้วเสร็จ

7. จัดพิมพ์แบบประเมินพฤติกรรมกรรมการปฏิบัติงาน

ผู้วิจัยได้ดำเนินการจัดพิมพ์แบบประเมินพฤติกรรมกรรมการปฏิบัติงานด้านการวัดผล-ประเมินผลของครูผู้สอน ทั้งรูปแบบที่ใช้เทคนิคมาตรฐานค่าเชิงพฤติกรรม และรูปแบบที่ใช้มาตรฐานค่าแบบกราฟิก ตามต้นฉบับที่ได้ปรับปรุงเป็นที่เรียบร้อยแล้ว โดยในการจัดพิมพ์ได้พิมพ์ส่วนปกหน้าแยกสำหรับกลุ่มผู้ประเมินแต่ละกลุ่มโดยเฉพาะ คือ "สำหรับครูผู้สอนใช้ประเมินตนเอง" และ "สำหรับหัวหน้าหมวดวิชาใช้ประเมินครูผู้สอน" พร้อมทั้งมีคำชี้แจงพิเศษสำหรับแต่ละกลุ่มไว้ที่ปกหน้าด้วย

สรุปลักษณะของแบบประเมินที่ใช้ เป็นเครื่องมือเก็บรวบรวมข้อมูลในการวิจัยครั้งนี้
นอกเหนือจากปกหน้าซึ่งจัดพิมพ์แยกสำหรับกลุ่มผู้ประเมินแต่ละกลุ่มโดยเฉพาะแล้ว
ในส่วนของเนื้อหาในแบบประเมิน ประกอบด้วยสาระดังนี้

คำชี้แจงเกี่ยวกับวัตถุประสงค์ของการประเมิน และขอบเขตเนื้อหา

ตอนที่ 1 ข้อมูลส่วนตัวของผู้ตอบแบบประเมิน

ตอนที่ 2 รายการพฤติกรรมกรรมการปฏิบัติงานด้านการวัดผล-ประเมินผลที่ต้องการให้ประเมิน ซึ่งประกอบด้วยรายการพฤติกรรมหลัก 18 ข้อ แยกเป็น

- | | |
|---|-------|
| ก. พฤติกรรมด้านการวางแผนการวัดผล ๙ | 5 ข้อ |
| ข. พฤติกรรมการวัดผล-ประเมินผลในชั้นสอน | 4 ข้อ |
| ค. พฤติกรรมการวัดผล-ประเมินผลในชั้นสรุปผลการเรียน | 4 ข้อ |
| ง. พฤติกรรมการตรวจสอบและพัฒนา | |
| ระบบการวัดผล-ประเมินผลของตนเอง | 5 ข้อ |

ส่วนของคำชี้แจง และเนื้อหาสาระในตอนที่ 1 จะมีใจความเหมือนกันทุกประการ ทั้งฉบับมาตรฐานค่าเชิงพฤติกรรม และฉบับมาตรฐานค่าแบบกราฟิก เฉพาะในตอนที่ 2 จะแตกต่างกัน กล่าวคือ นอกเหนือไปจากความแตกต่างในส่วนของสเกลคำตอบแล้ว แต่ละรูปแบบ จะมีคำชี้แจงและกรณีตัวอย่างการประเมินที่เป็นของตนเองโดยเฉพาะ

โดยสรุป แบบประเมินชนิดมาตรฐานประมาณค่าทั้งสองรูปแบบได้บรรจุไว้ด้วยเนื้อหาสาระในส่วนต่าง ๆ เทียบเท่ากัน ประกอบด้วยรายการพฤติกรรมหลัก (Stem) เหมือนกัน แต่แตกต่างกันในส่วนของสเกลคำตอบที่เป็นทางเลือก (Alternatives) ทั้งนี้โดยถือว่าช่วงคะแนนมาตรฐานวัดทั้งสองรูปแบบอยู่ในระบบช่วงคะแนนที่เทียบเท่ากัน พฤติกรรมที่นำมาผูกติดกับมาตรฐานค่าเชิงพฤติกรรมถือเป็นเพียงตัวอย่างพฤติกรรมการทำงานเท่านั้น (รายละเอียดของแบบประเมินทั้งสองชนิดดูที่ภาคผนวก ง.)

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลในการวิจัยครั้งนี้ ได้ดำเนินการในลักษณะของการวิจัยเชิงสำรวจ แนวทางการดำเนินการและผลการดำเนินการเป็นดังนี้

1. การประเมินพฤติกรรมการทำงานของกลุ่มเป้าหมายประเภทครูผู้สอนทั่วไป

ในรอบแรก ดำเนินการโดยส่งรายชื่อโรงเรียนมาครึ่งหนึ่ง (จำนวน 52 โรงเรียน) แล้วส่งแบบประเมินชนิดมาตรฐานประมาณค่าเชิงพฤติกรรมสำหรับหัวหน้าหมวดวิชาคณิตศาสตร์ใช้ประเมินพฤติกรรมของครูผู้สอน และสำหรับครูผู้สอนใช้ประเมินตนเอง พร้อมทั้งแนบซองปิดผนึกไปด้วย 2 ชุด เพื่อครูผู้สอนและหัวหน้าหมวดวิชาสามารถส่งแบบประเมินกลับคืนได้อย่างอิสระแยกจากกัน ส่วนโรงเรียนอีกครึ่งหนึ่ง (51 โรงเรียน) ได้ส่งแบบประเมินชนิดมาตรฐานประมาณค่าแบบกราฟิกให้ก่อน โดยดำเนินการในลักษณะเดียวกัน

การดำเนินการในขั้นตอนนี้ ได้ดำเนินการในปลายภาคเรียนที่ 1 คือ สัปดาห์สุดท้ายของเดือนกันยายน 2533 ซึ่งปรากฏว่าหลังจากส่งแบบประเมินและรอการส่งกลับคืนเป็นระยะเวลา 2 สัปดาห์ ได้รับแบบประเมินกลับคืน ประมาณร้อยละ 60 ของจำนวนโรงเรียนที่ส่งไปทั้งหมด ในจำนวนนี้มีโรงเรียนที่ส่งแบบประเมินกลับคืนครบถ้วนทั้งของหัวหน้าหมวดวิชาและครูผู้สอน จำนวน 24 โรงเรียนหรือร้อยละ 23.30 ของจำนวนที่ส่งไป ซึ่งโรงเรียนกลุ่มนี้ ผู้วิจัยได้ส่งแบบประเมินรอบที่ 2 ไปให้ในทันที เป็นแบบประเมินอีกรูปแบบหนึ่งที่แตกต่างไปจากรอบแรก ขณะเดียวกันก็ดำเนินการทวงถามแบบประเมินจากโรงเรียนที่ส่งแบบประเมินกลับคืนไม่ครบถ้วน โดย (1) ส่งไปรษณียบัตรถึงหัวหน้าหมวดวิชาโดยตรง และ (2) ส่งไปรษณียบัตรถึงครู-อาจารย์ที่ผู้วิจัยรู้จักเป็นการส่วนตัว ให้ช่วยติดตามให้อีกชั้นหนึ่ง โดยดำเนินการควบคู่ไปกับการทวงถามจากหัวหน้าหมวดวิชา ซึ่งปรากฏว่านับจนถึงสิ้นเดือนตุลาคม 2533 ได้รับแบบประเมินรอบแรกกลับคืนจากโรงเรียนต่าง ๆ ที่สมบูรณ์

ครบถ้วน กล่าวคือ ได้รับคืนทั้งแบบประเมินโดยหัวหน้าหมวดวิชา และครูผู้สอน จำนวน 78 โรง (ซึ่งทุกโรงจะได้รับแบบประเมินรอบที่ 2 ในทันทีที่ได้รับแบบประเมินรอบแรกกลับคืนครบถ้วน) ในส่วนของโรงเรียนส่วนที่เหลือ ผู้วิจัยได้ทวงถามรอบที่ 2 โดยการโทรศัพท์ถึงหัวหน้าหมวดวิชา ผลปรากฏว่า ได้รับแบบประเมินกลับคืนครบถ้วนเพิ่มอีก 6 โรง รวมโรงเรียนที่ได้รับแบบประเมินรอบที่ 1 กลับคืนครบถ้วน จำนวน 94 โรง สำหรับโรงเรียนส่วนที่เหลือผู้วิจัยได้ตัดสินใจส่งแบบประเมินรอบที่ 2 ถึงหัวหน้าหมวดวิชาและครูผู้สอนเช่นเดียวกับโรงเรียนอื่น ๆ โดยดำเนินการในสัปดาห์ที่ 2 ของเดือนพฤศจิกายน 2533 แล้วรอแบบสอบถามกลับคืนจนถึงสิ้นเดือนพฤศจิกายน ซึ่งปรากฏว่าโรงเรียนที่ตอบแบบประเมินในรอบแรกทุกโรงจะตอบแบบประเมินในรอบที่ 2 ด้วย ในขณะเดียวกันมีโรงเรียนที่ไม่ตอบในรอบแรก แต่ตอบแบบประเมินในรอบที่ 2 จำนวน 10 โรง โดยสรุป ถ้านับเป็นรายโรงแล้ว มีโรงเรียนที่ตอบแบบประเมินกลับคืนถึง 97 โรง สรุปจำนวนแบบประเมินรูปแบบต่าง ๆ ที่ได้รับกลับคืน ปรากฏดังตารางที่ 3

ตารางที่ 3 สรุปจำนวนแบบประเมินรูปแบบต่าง ๆ ที่ได้รับกลับคืนจากกลุ่มตัวอย่างผู้ประเมิน ประเภทครูผู้สอนโดยทั่วไป และหัวหน้าหมวดวิชา

กลุ่มผู้ประเมิน	จำนวนที่ได้รับกลับคืน	
	มาตรฐานค่า เชิงสถิติกรรม	มาตรฐานค่า แบบกราฟิก
ครูผู้สอนทั่วไป	89 (86.40 %)	92 (89.32 %)
หัวหน้าหมวดวิชา	85 (82.52 %)	93 (90.29 %)

หมายเหตุ ตัวเลขในวงเล็บ คือ ร้อยละของแบบสอบถามเมื่อเทียบกับจำนวนที่คาดหวัง

จากตาราง พบว่า แบบประเมินที่ได้รับกลับคืนมีจำนวนเกินร้อยละ 80 ของจำนวนที่คาดหวังทุกประเภทของแบบประเมิน และแบบกราฟิกปรากฏสัดส่วนที่ได้รับกลับคืนสูงกว่าแบบประเมินเชิงพฤติกรรม

2. การประเมินพฤติกรรมของครูกลุ่มที่ปฏิบัติงานเป็นระบบ และปฏิบัติงานยังไม่เป็นระบบ

การเก็บรวบรวมข้อมูลส่วนนี้ ผู้วิจัยได้เดินทางไปพบหัวหน้าหมวดวิชาคณิตศาสตร์ของโรงเรียนที่ปรากฏรายชื่อครูที่ปฏิบัติงานเป็นระบบตามการเสนอของศึกษานิเทศก์ภาควิชาคณิตศาสตร์ หน่วยศึกษานิเทศก์ กรมสามัญศึกษา ได้ชี้แจงจุดประสงค์ของการวิจัยให้ทราบ พร้อมทั้งขอความร่วมมือให้ช่วยระบุรายชื่อครูที่ปฏิบัติงานเป็นระบบตามการรับรู้ของหัวหน้าหมวดวิชาจำนวน 3 คน ทั้งนี้โดยยึดเกณฑ์การพิจารณาตามที่ได้ระบุไว้แล้วในตอนที่ว่าด้วยวิธีการคัดเลือกกลุ่มเป้าหมาย หน้า 67 แล้วคัดเลือกรายชื่อที่ปรากฏตรงกัน ถือเป็นครูที่ปฏิบัติงานเป็นระบบภายใต้การดำเนินการโดยวิธีการนี้ ปรากฏว่าได้รายชื่อครูที่ปฏิบัติงานเป็นระบบจำนวน 23 คน ผู้วิจัยจึงได้ขอความร่วมมือจากหัวหน้าหมวดวิชาคณิตศาสตร์โรงเรียนอื่น ๆ อีก 7 โรงเรียนช่วยระบุรายชื่อครูคณิตศาสตร์ที่ปฏิบัติงานอย่างเป็นระบบเพิ่มอีก 7 คน ทำให้ได้จำนวนครูที่ปฏิบัติงานเป็นระบบรวม 30 คน ใช้เป็นกลุ่มตัวอย่างของครูที่ปฏิบัติงานเป็นระบบในการวิจัยครั้งนี้ โดยในการนี้ ได้ขอความร่วมมือจากหัวหน้าหมวดวิชาจากโรงเรียนทั้ง 30 โรงเรียน ช่วยคัดเลือกครูที่ปฏิบัติงานยังไม่เป็นระบบ-ขาดประสิทธิภาพ หรือควรได้รับการพัฒนาโดยเร่งด่วนอีกโรงเรียนละ 1 คน รวมครูกลุ่มนี้จำนวน 30 คน ถือเป็นกลุ่มตัวอย่างครูที่ปฏิบัติงานยังไม่เป็นระบบ ผู้วิจัยได้มอบแบบประเมินรอบที่ 1 ให้แก่หัวหน้าหมวดวิชาในลักษณะเดียวกับการเก็บรวบรวมข้อมูลกรณีของครูผู้สอนทั่วไป โดยนัดหมายรับคืนด้วยตนเอง ในวันรับข้อมูลรอบที่ 1 กลับคืนได้มอบแบบประเมินรอบที่ 2 ไว้อีกหนึ่งชุดหนึ่ง และขอความร่วมมือจากหัวหน้าหมวดวิชาช่วยเก็บรวบรวมข้อมูลและจัดส่งกลับคืนทางไปรษณีย์

การเก็บรวบรวมข้อมูลจากกลุ่มเป้าหมายกลุ่มนี้ เริ่มต้นเก็บรวบรวมข้อมูลในสัปดาห์แรกของเดือนตุลาคม จนถึงกลางเดือน พฤศจิกายน ปรากฏว่าได้แบบประเมินกลับคืนดังตารางที่ 4

ตารางที่ 4 สรุปจำนวนแบบประเมินรูปแบบต่าง ๆ ที่ได้รับคืนจากการประเมินพฤติกรรมตนเองของครูกลุ่มปฏิบัติงานเป็นระบบ และกลุ่มปฏิบัติงานยังไม่เป็นระบบ

กลุ่มผู้ประเมิน	จำนวนที่ได้รับกลับคืน	
	มาตรฐานค่า	มาตรฐานค่า
	เชิงพฤติกรรม	แบบกราฟิก
ครูที่ปฏิบัติงานเป็นระบบ	30	30
ครูที่ปฏิบัติงานไม่เป็นระบบ	30	29

จากตาราง พบว่า ในการประเมินพฤติกรรมการปฏิบัติงานของครูที่ปฏิบัติงานเป็นระบบ และปฏิบัติงานยังไม่เป็นระบบนั้น ปรากฏว่าได้รับแบบประเมินตนเองของครูกลับคืนครบถ้วนเกือบทุกฉบับ มีเพียงฉบับเดียวเท่านั้นที่ไม่สามารถเก็บรวบรวมได้

การวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลเพื่อตรวจสอบสมมุติฐานการวิจัย ในขั้นต้นได้ทำการวิเคราะห์ค่าสถิติพื้นฐานของคะแนนผลการประเมินคือ วิเคราะห์ค่าเฉลี่ยเลขคณิต (Arithmetic Mean) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) ของคะแนนผลการประเมินทั้งระดับรายพฤติกรรม และระดับรายบุคคล และวิเคราะห์ค่าสหสัมพันธ์ของคะแนนผลการประเมินระหว่างด้านพฤติกรรมต่าง ๆ (Dimension Intercorrelatin) แล้วนำค่าสถิติพื้นฐานไปวิเคราะห์เปรียบเทียบเพื่อตรวจสอบสมมุติฐานการวิจัยแต่ละรายการ โดยอาศัยแนวคิดทางทฤษฎีและแนวทางการวิเคราะห์ดังนี้

1. การตรวจสอบความตรงในการประมาณค่า

การตรวจสอบความตรงในการประมาณค่าในที่นี้ เป็นการตรวจสอบความตรงตามภาวะสัมพันธ์ฐาน (Construct Validity) ซึ่งได้ดำเนินการวิเคราะห์ข้อมูลเพื่อตรวจสอบความตรงใน 2 ลักษณะ คือ วิเคราะห์เมตริกสหสัมพันธ์แบบ Multitrait-Multimethod เพื่อตรวจสอบ Convergent & Discriminant Validity และวิเคราะห์เปรียบเทียบผลการประเมินพฤติกรรมตนเองของครูกลุ่มที่ปฏิบัติงานเป็นระบบกับกลุ่มที่ปฏิบัติงานยังไม่เป็นระบบ เพื่อตรวจสอบความตรงของมาตรประมาณค่าในการจำแนกความแตกต่างระหว่างกลุ่มครูที่ปฏิบัติงานมีประสิทธิภาพต่างกัน รายละเอียดแนวทางการวิเคราะห์ข้อมูลเพื่อตรวจสอบความตรงแต่ละลักษณะ เป็นดังนี้

1.1 การวิเคราะห์เมตริกสหสัมพันธ์แบบ Multitrait-Multimethod เพื่อตรวจสอบ Convergent & Discriminant Validity ของมาตรประมาณค่าทั้งสองรูปแบบ การวิเคราะห์ข้อมูลในส่วนนี้ เป็นการประยุกต์แนวทางการอธิบายเมตริกสหสัมพันธ์แบบ Multitrait-Multimethod ตามหลักของ Campbell และ Fiske (1959) โดยถือว่า พฤติกรรมด้านการประเมินผลของครูผู้สอน 4 ด้านย่อย (Multidimension) คือ การวางแผนการวัดผล-ประเมินผล การวัดผล-ประเมินผลในชั้นสอน การวัดผล-ประเมินผลในชั้นสรุปผลการเรียน และการพัฒนาและปรับปรุงระบบการวัดผล-ประเมินผลของตนเอง เป็นเสมือนคุณลักษณะ 4 รายการ (Multitrait) (Campbell, 1973 : 21, Keaveny and McGinn, 1975 : 695-703) และวิธีการวัดคือ มาตรประมาณค่าเชิงพฤติกรรม กับมาตรประมาณค่าแบบกราฟิก เป็นวิธีการ 2 วิธี (Multimethod) ที่ใช้วัดคุณลักษณะดังกล่าว โดยมีชื่อตรงว่า มาตรวัดจะมีความตรงชนิดที่เรียกว่า Convergent & Discriminant Validity ถ้าสะท้อนผลการวัดเป็นไปตามหลักการ 4 ประการ คือ (1) ค่าสหสัมพันธ์ระหว่างคะแนนผลการวัดโดยมาตรวัดต่างรูปแบบที่มุ่งวัดพฤติกรรมด้านเดียวกัน (Different methods measured the same dimension) มีค่าสูงอย่างมีนัยสำคัญทางสถิติ (2) ค่าสหสัมพันธ์ระหว่างคะแนนผลการวัดโดยมาตรวัดต่างรูปแบบที่มุ่งวัดพฤติกรรมด้านเดียวกัน มีค่ามากกว่าค่าสหสัมพันธ์ระหว่างคะแนนจากการวัดพฤติกรรมที่แตกต่างกันด้วยมาตรวัดต่างรูปแบบ (Heterodimension-Heteromethod) (3) ค่าสหสัมพันธ์ระหว่างคะแนนผลการวัดโดยมาตรวัดต่างรูปแบบที่มุ่งวัดพฤติกรรมด้านเดียวกัน มีค่ามากกว่าค่าสหสัมพันธ์ระหว่างคะแนนผลการวัดพฤติกรรมต่างด้านด้วยมาตรวัดรูปแบบเดียวกัน (Heterodimension-Monomethod) และ (4) ค่าสหสัมพันธ์ระหว่างคะแนนผลการวัดพฤติกรรม

ที่แตกต่างกันด้วยมาตรวัดรูปแบบเดียวกันมีค่าใกล้เคียงกันหรือเป็นไปได้ในลักษณะเดียวกันในทุกรูปแบบของมาตรวัด

แนวทางการวิเคราะห์ข้อมูลเพื่อตรวจสอบความตรงภายในได้เงื่อนไขนี้ ได้ใช้วิธีการอธิบายเมตริกซ์สหสัมพันธ์ตามหลักการ 4 ประการ ดังกล่าวข้างต้น แล้วเปรียบเทียบค่าสหสัมพันธ์ของคะแนนผลการประเมินพฤติกรรม 2 ด้านใด ๆ ระหว่างกรณีใช้มาตรประมาณค่าเชิงพฤติกรรมเป็นเครื่องมือประเมิน กับกรณีใช้มาตรประมาณค่าแบบกราฟิกเป็นเครื่องมือประเมินเพื่อตรวจสอบคุณภาพของมาตรประมาณค่าในการจำแนกความแตกต่างระหว่างด้านพฤติกรรมโดยทำการทดสอบความแตกต่างในภาพรวมระหว่างกรณีใช้มาตรวัดต่างรูปแบบกัน ด้วย The Wilcoxon Signed Ranks Tests กรณีกลุ่มตัวอย่างขนาดเล็ก (Siegel, S., 1988 : 88-89)

1.2 การวิเคราะห์เปรียบเทียบผลการประเมินพฤติกรรมตนเองของครูกลุ่มที่ปฏิบัติงานเป็นระบบ กับ กลุ่มที่ปฏิบัติงานยังไม่เป็นระบบ เพื่อตรวจสอบความตรงของมาตรประมาณค่าในการจำแนกความแตกต่างระหว่างกลุ่มครูที่ปฏิบัติงานมีประสิทธิภาพต่างกัน

การตรวจสอบความตรงภายในได้เงื่อนไขนี้ เป็นการดำเนินการโดยอาศัยแนวความคิดที่ว่า มาตรประมาณค่าที่มีคุณภาพควรจะให้ผลการประเมินที่สามารถจำแนกความแตกต่างระหว่างกลุ่มครู 2 กลุ่ม ที่ได้รับการยอมรับว่าเป็นกลุ่มที่ปฏิบัติงานมีประสิทธิภาพแตกต่างกันอย่างคมชัด หรือจำแนกความแตกต่างได้อย่างชัดเจน

แนวทางการวิเคราะห์ข้อมูลในส่วนนี้ ได้ใช้วิธีเปรียบเทียบค่าเฉลี่ยเลขคณิตของคะแนนผลการประเมินพฤติกรรมตนเองของครูกลุ่มที่ปฏิบัติงานเป็นระบบ กับกลุ่มครูที่ปฏิบัติงานยังไม่เป็นระบบ โดยใช้สัมประสิทธิ์สหสัมพันธ์แบบไบซีเรียล (The Biserial Coefficient of Correlation) เป็นดัชนีบ่งชี้น้ำหนักเชิงจำแนกความแตกต่างระหว่างกลุ่ม เป็นรายพฤติกรรมซึ่งคำนวณโดยสูตร (Guilford, J.P., 1978 : 305)

$$r_{b1a} = \frac{\bar{X}_p - \bar{X}_q}{S_x} \cdot \frac{p \cdot q}{Y}$$

เมื่อ

- $r_{i,j}$ = สัมประสิทธิ์สหสัมพันธ์แบบไบซีเรียล
 X_p = ค่าเฉลี่ยเลขคณิตของผลการประเมินพฤติกรรมกลุ่มครูที่ปฏิบัติงานเป็นระบบ
 X_q = ค่าเฉลี่ยเลขคณิตของผลการประเมินพฤติกรรมกลุ่มครูที่ปฏิบัติงาน
 ยังไม่เป็นระบบ
 S_c = ส่วนเบี่ยงเบนมาตรฐานของคะแนนของกลุ่มตัวอย่างทั้งหมด
 p = สัดส่วนจำนวนคนในกลุ่มที่ปฏิบัติงานเป็นระบบ
 q = สัดส่วนจำนวนคนในกลุ่มที่ปฏิบัติงานยังไม่เป็นระบบ
 Y = ค่าความสูงของโค้งปกติ ณ จุดแบ่งสัดส่วน p กับ q

แล้วเปรียบเทียบน้ำหนักเชิงจำแนกความแตกต่างในภาพรวมระหว่างกรณีใช้มาตรวัด
 ต่างรูปแบบกัน ด้วย The Wilcoxon Signed Ranks Tests (Siegel, S., 1988 : 88-89)
 โดยสูตร

$$z = \frac{T^+ - \mu_{T^+}}{\sigma_{T^+}}$$

เมื่อ

- Z = อัตราส่วนซี
 T^+ = ผลรวมของลำดับที่ของคะแนนความแตกต่าง เฉพาะที่มีเครื่องหมายบวก
 ซึ่งเกิดขึ้นในกรณีที่ค่าสังเกตภายใต้เงื่อนไขหลัก มีค่ามากกว่าค่าสังเกต
 ภายใต้อีกเงื่อนไขหนึ่ง เฉพาะในที่นี้ก็คือ กรณีที่ดัชนีหนึ่งซึ่งน้ำหนักเชิง
 จำแนกความแตกต่าง กรณีใช้มาตรประมาณค่าเชิงพฤติกรรม
 เป็นเครื่องมือประเมิน มีค่ามากกว่ากรณีใช้มาตรประมาณค่าแบบกราฟิก
 เป็นเครื่องมือประเมิน
 μ_{T^+} = ค่าเฉลี่ยที่คาดหวังของการเกิดลำดับที่มีเครื่องหมายบวก คำนวณโดยสูตร

$$\mu_{T^+} = \frac{N(N+1)}{4}$$

σ_T = ความคลาดเคลื่อนมาตรฐานของการเกิดลำดับที่ ที่มีเครื่องหมายบวก
คำนวณโดยสูตร

$$\sigma_T^2 = \frac{N(N+1)(2N+1)}{24}$$

N = จำนวนคู่ของรายการเปรียบเทียบ

2. การตรวจสอบความเที่ยงในการประมาณค่า

เนื่องจากการประเมินพฤติกรรมของครูในการวิจัยนี้ กลุ่มเป้าหมาย (Target Ratee) 1 คน ได้ถูกประเมินโดยผู้ประเมิน (Raters) 2 คน คือ ตนเอง และหัวหน้าหมวดวิชาซึ่งถ้าหากมาตรฐานประมาณค่าหรือเครื่องมือที่ใช้ในการประเมินครั้งนี้ มีความคมชัดและมีลักษณะที่สื่อความหมายเป็นอย่างดีแล้ว คะแนนจากผู้ประเมิน 2 คน ควรจะสอดคล้องกันหรือมีค่าใกล้เคียงกัน ดังนั้นด้วยแนวความคิดนี้ ในการตรวจสอบความเที่ยงของการประมาณค่า จึงได้ใช้วิธีการวิเคราะห์ค่าสหสัมพันธ์ระหว่างคะแนนของผู้ประเมิน 2 กลุ่ม ถือเป็นดัชนีบ่งชี้ความเที่ยงระหว่างผู้ประเมิน (Rater Reliability)

แนวทางการวิเคราะห์ความเที่ยงระหว่างผู้ประเมินในกรณีนี้ ได้ใช้วิธีวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ระหว่างชั้นคะแนน (Intraclass Correlation Coefficient) ถือเป็นดัชนีบ่งชี้ความเที่ยงระหว่างผู้ประเมิน ซึ่งคำนวณโดยสูตรดังนี้ (Shrout, 1979 : 420-428)

$$ICC = \frac{BMS - WMS}{BMS + (K - 1)WMS}$$

เมื่อ

BMS = ความแปรปรวนอันเนื่องมาจากความแตกต่างของกลุ่มเป้าหมาย
(Between-targets Mean Square)

WMS = ความแปรปรวนของคะแนนในกลุ่มเป้าหมายแต่ละคน
(Within-target Mean Square)

- K = จำนวนผู้ประเมินต่อเป้าหมาย 1 คน
- ICC = ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างชั้นคะแนน
(Intraclass Correlation Coefficient) หรือดัชนีบ่งชี้
ความเที่ยงระหว่างผู้ประเมิน

การเปรียบเทียบค่าความเที่ยงระหว่างผู้ประเมิน ระหว่างกรณีใช้มาตรฐานค่าต่าง
รูปแบบกัน ได้ใช้วิธีวิเคราะห์เชิงบรรยายและทดสอบความแตกต่างในภาพรวมด้วย The Wilcoxon
Signed Ranks Test

3. การตรวจสอบลักษณะความคลาดเคลื่อนในการวัด

การตรวจสอบลักษณะความคลาดเคลื่อนในการวัดในการวิจัยนี้ เป็นการตรวจสอบ
ลักษณะความคลาดเคลื่อนอย่างเป็นระบบ (Systematic Error) ที่สำคัญ 2 ประการ คือ
Leniency Error หรือ ความคลาดเคลื่อนเชิงปล่อยคะแนน หรือ เลือกประเมินในระดับคะแนน
สูง ๆ อย่างสม่ำเสมอ และ Halo Error หรือ ความคลาดเคลื่อนอันเนื่องมาจากการเคยมี
ปฏิสัมพันธ์มาก่อนระหว่างผู้ประเมินกับผู้ถูกประเมิน แนวทางการวิเคราะห์ข้อมูลเพื่อตรวจสอบลักษณะ
ความคลาดเคลื่อนแต่ละรายการเป็นดังนี้

3.1 การตรวจสอบ Leniency Error

การตรวจสอบ Leniency Error ในที่นี้ ได้ใช้วิธีเปรียบเทียบค่าเฉลี่ย
เลขคณิตของผลการประเมินพฤติกรรมเป็นรายพฤติกรรม ระหว่าง กรณีใช้มาตรฐานค่าเชิง
พฤติกรรมเป็นเครื่องมือประเมิน กับ กรณีใช้มาตรฐานค่าแบบกราฟิกเป็นเครื่องมือประเมิน
เพื่อพิจารณาแนวโน้มของผลการประเมินว่ามาตรฐานค่ารูปแบบใดให้ค่าเฉลี่ยเลขคณิตของคะแนน
ผลการประเมินสูงกว่าอย่างสม่ำเสมอ แล้วทดสอบความแตกต่างระหว่างกรณีใช้มาตรฐานค่า
ต่างรูปแบบกันด้วย The Wilcoxon Signed Ranks Test

3.2 การตรวจสอบ Halo Error

ด้วยแนวความคิดที่ว่า ในการประเมินเป้าหมายใด ๆ นั้น การเคยมีปฏิสัมพันธ์มาก่อน ระหว่างผู้ประเมินกับผู้ถูกประเมินจะมีผลทำให้ผลการประเมินเกิดความคลาดเคลื่อนได้ กล่าวคือ ถ้าผู้ประเมินมีทัศนคติเชิงบวกหรือมีความรู้สึกที่ดีต่อเป้าหมายที่กำลังประเมินก็มีแนวโน้มที่จะประเมินพฤติกรรมทุกรายการในทางมาก หรือทางบวก ในทางกลับกัน ถ้าผู้ประเมินมีทัศนคติเชิงลบ หรือมีความรู้สึกที่ไม่ดีต่อเป้าหมายที่กำลังประเมิน ก็มีแนวโน้มที่จะประเมินหรือให้คะแนนในระดับต่ำ ๆ อย่างสม่ำเสมอ ซึ่งตามข้อสันนิษฐานนี้ จะมีผลทำให้คะแนนระหว่างรายการพฤติกรรมของผู้ประเมินแต่ละคนแตกต่างกันน้อยมาก หรือมีการกระจายน้อย หรือผลการประเมินระหว่างตัวแปรไม่เป็นอิสระจากกัน ดังนั้น ในการตรวจสอบความคลาดเคลื่อนอันเนื่องมาจากการมีปฏิสัมพันธ์มาก่อนในที่นี้จึงได้ทำการวิเคราะห์การกระจายของคะแนนผลการประเมินรายการพฤติกรรมของผู้ประเมินแต่ละคน ถือเป็นคะแนนบ่งชี้การเกิดความคลาดเคลื่อน (Halo Score) โดยมีเงื่อนไขว่า ถ้าคะแนนระหว่างรายการพฤติกรรมแตกต่างกันน้อย หรือมีค่าใกล้เคียงกันอย่างสม่ำเสมอ จะมีผลทำให้ค่าดัชนีบ่งชี้การกระจายของคะแนนหรือ Halo Score มีค่าต่ำ ถือเป็นสภาพที่เกิดความคลาดเคลื่อนในการประเมินสูงกว่า กรณีที่ดัชนีบ่งชี้การกระจายของคะแนนมีค่ามาก (Cook, 1989 : 40) ภายใต้เงื่อนไขนี้ ในการตรวจสอบความคลาดเคลื่อนจึงได้ใช้วิธีการเปรียบเทียบค่าเฉลี่ยเลขคณิตของดัชนีบ่งชี้การกระจายของคะแนน หรือคะแนนความคลาดเคลื่อน (Halo Score) ระหว่างกรณีใช้มาตราประมาณค่าต่างรูปแบบกัน และทดสอบความแตกต่างในประชากรด้วย t Test กรณีกลุ่มตัวอย่างไม่เป็นอิสระจากกัน (Dependent Samples) ซึ่งคำนวณโดยสูตร (Guilford, J.P., 1978 : 158)

$$t = \frac{\bar{X}_d}{\sqrt{\frac{\sum x_d^2}{N(N-1)}}}$$

เมื่อ	t	=	อัตราส่วนที่
	\bar{X}_d	=	ค่าเฉลี่ยเลขคณิตของคะแนนความแตกต่าง
	X_d	=	คะแนนเบี่ยงเบนของคะแนนความแตกต่างใด ๆ จากค่าเฉลี่ยเลขคณิตของคะแนนความแตกต่าง = $(d - \bar{X}_d)$
	N	=	จำนวนข้อมูล

การแปลความหมายค่าเฉลี่ยเลขคณิตของคะแนนผลการประเมินพฤติกรรม

เนื่องจากการประเมินพฤติกรรมการปฏิบัติงานของครูในการวิจัยนี้ ได้ใช้มาตราประมาณค่า ชนิด 5 ระดับทางเลือก ซึ่งได้ตรวจให้คะแนนตามเกณฑ์ คือ

- 5 คะแนน สำหรับการประเมินคุณภาพการปฏิบัติงานในระดับมากที่สุด
- 4 คะแนน สำหรับการประเมินคุณภาพการปฏิบัติงานในระดับมาก
- 3 คะแนน สำหรับการประเมินคุณภาพการปฏิบัติงานในระดับปานกลาง
- 2 คะแนน สำหรับการประเมินคุณภาพการปฏิบัติงานในระดับน้อย
- 1 คะแนน สำหรับการประเมินคุณภาพการปฏิบัติงานในระดับน้อยที่สุด

ดังนั้น ในส่วนของค่าเฉลี่ยเลขคณิตของคะแนนผลการประเมินพฤติกรรม ได้กำหนดช่วงคะแนนไว้เพื่อประโยชน์ในการแปลความหมายดังนี้

- | | | |
|-------------|-------|---|
| 4.50 - 5.00 | | แสดงถึงคุณภาพของพฤติกรรมในระดับมากที่สุด |
| 3.50 - 4.49 | | แสดงถึงคุณภาพของพฤติกรรมในระดับมาก |
| 2.50 - 3.49 | | แสดงถึงคุณภาพของพฤติกรรมในระดับปานกลาง |
| 1.50 - 2.49 | | แสดงถึงคุณภาพของพฤติกรรมในระดับน้อย |
| 1.00 - 1.49 | | แสดงถึงคุณภาพของพฤติกรรมในระดับน้อยที่สุด |

การนำเสนอข้อมูล

ในการนำเสนอผลการวิเคราะห์ข้อมูลหรือข้อค้นพบ ได้นำเสนอเรียงตามลำดับสมมุติฐานการวิจัย คือ เสนอผลการวิเคราะห์ข้อมูลเพื่อตรวจสอบความตรงในการประมาณค่า ผลการวิเคราะห์ข้อมูลเพื่อตรวจสอบความเที่ยงในการประมาณค่า และผลการวิเคราะห์ข้อมูลเพื่อตรวจสอบลักษณะความคลาดเคลื่อนในการประมาณค่าตามลำดับ โดยนำเสนอข้อมูลทั้งสถานการณ์การประเมินพฤติกรรมตนเองของครูผู้สอน และสถานการณ์การประเมินโดยหัวหน้าหมวดวิชาควบคู่กันไป รายละเอียดผลการวิเคราะห์ข้อมูลเพื่อตรวจสอบสมมุติฐานการวิจัยแต่ละรายการ ปรากฏดังบทที่ 4 ส่วนสรุปผลการวิจัย หรือข้อค้นพบปรากฏดังบทที่ 5