



บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ต้องการศึกษาความแม่นยำในการวัดแบบความเรียง ซึ่งมีตัวประกอบ
เกี่ยวข้องหลายตัว จึงใช้ทฤษฎีการอ้างอิงสรุปแบบศึกษาหลายตัวประกอบ (Multifacet Gene-
ralizability Studies) ผู้วิจัยดำเนินการตามลำดับขั้นดังนี้

เอกภพของการวิจัย

ทฤษฎีการอ้างอิงสรุปของ Cardinet, et al. (1976, 1981, 1982, 1983) ใช้
คำว่า "เอกภพ" ในความหมายเดียวกันกับ "ประชากร" เพื่อให้สามารถอธิบาย วิธีการศึกษาได้
สอดคล้องกับแนวคิดของทฤษฎี ผู้วิจัยจึงขอใช้คำว่า เอกภพ แทนคำว่า ประชากรของการวิจัย ใน
การวิจัยครั้งนี้กำหนดเอกภพของการวิจัยไว้ดังนี้

1. เอกภพของการสังเกตที่สามารถยอมรับได้ (Universe of admissible
observation)

1.1 เอกภพของนักเรียน (p) ได้แก่ นักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียน
สังกัดสำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ (สพช.) อำเภอโนนสัง จังหวัดอุดรธานี
ในปีการศึกษา 2531 จำนวน 1448 คน

1.2 เอกภพของผู้ตรวจ (r) ได้แก่ ครูที่สอนกลุ่มวิชาสร้างเสริมประสบการณ์
ชีวิตในชั้นประถมศึกษาปีที่ 6 โรงเรียนประถมศึกษาสังกัด สพช.อำเภอโนนสัง จังหวัดอุดรธานี
ปีการศึกษา 2531 จำนวน 154 คน

1.3 เอกภพของข้อสอบ (i) ได้แก่ ข้อสอบความเรียงวัดผลสัมฤทธิ์ทางการ
เรียนตามจุดประสงค์การเรียนรู้ของกลุ่มวิชาสร้างเสริมประสบการณ์ชีวิต ชั้นประถมศึกษาปีที่ 6
ที่ว่า "ให้นักเรียนสามารถแสดงความคิดเห็นเกี่ยวกับข่าวและเหตุการณ์ต่าง ๆ ได้ "

2. เอกภพของการอ้างอิงสรุป (Universe of generalization) เอกภพของ
การอ้างอิงสรุป หมายถึง เอกภพที่ผู้วิจัยต้องการอ้างอิงสรุปไปถึง ประกอบด้วยเอกภพของข้อสอบ

และผู้ตรวจ ซึ่งมีขนาดและประเภทดังนี้

2.1 เอกภพของข้อสอบและผู้ตรวจ โดยที่เอกภพของข้อสอบมีขนาดไม่จำกัด ($N_i \rightarrow \infty$) แต่เอกภพผู้ตรวจมีขนาดจำกัดคือ 154 คน ($N_i = 154$)

2.2 เอกภพของผู้ตรวจซึ่งเป็นเอกภพที่มีขนาดจำกัด ($N_i = 154$) อย่างเดียว โดยที่กำหนดให้ข้อสอบเป็นตัวประกอบคงที่

การสุ่มเงื่อนไขการวัดหรือกลุ่มตัวอย่าง

1. การกำหนดขนาดของกลุ่มตัวอย่าง

เป้าหมายของทฤษฎีการอ้างอิงสรุป คือการประมาณค่าความแปรปรวนของตัวประกอบที่นำมาศึกษา และการประมาณค่าสัมประสิทธิ์การอ้างอิงสรุป ดังนั้นขนาดกลุ่มตัวอย่างที่เหมาะสม จึงได้แก่ขนาดที่ให้ค่าประมาณความแปรปรวนไม่ลำเอียง (bias) สามารถนำไปใช้ประมาณค่าสัมประสิทธิ์การอ้างอิงสรุปได้อย่างมั่นใจ Smith (1978: 319-346) เสนอว่าขนาดของกลุ่มตัวอย่างขั้นต่ำที่ใช้ในการการศึกษา 3 ตัวประกอบ ควรมีขนาด $n_p \times n_i \times n_r$ อย่างน้อย 1000 ในการศึกษาครั้งนี้ จำนวน $n_p \times n_i \times n_r$ มีค่าเป็น 1500 กล่าวคือ จำนวนนักเรียน (n_p) เป็น 30 คน จำนวนข้อสอบ (n_i) เป็น 5 ข้อ และจำนวนผู้ตรวจ (n_r) เป็น 10 คน ซึ่งมีค่ามากกว่าเกณฑ์ขั้นต่ำ จึงเป็นขนาดกลุ่มตัวอย่างที่เหมาะสม

2. การสุ่มเงื่อนไขการวัด

2.1 การสุ่มผู้ตรวจ ใช้วิธีสุ่มอย่างง่ายโดยการจับสลากจากรายชื่อของครูในเอกภพที่กล่าวมาแล้ว ได้จำนวน 20 คน เป็นชาย 13 คนและหญิง 7 คน สุ่มแบ่งเป็น 2 กลุ่ม ๆ ละ 10 คน จากนั้นสุ่มแต่ละกลุ่มเข้ารับวิธีการทดลองคือวิธีตรวจต่างกัน 2 วิธี ดังนั้นจำนวนผู้ตรวจในแต่ละวิธีจึงเป็น 10 คน ($n_p = 10$) ผลการสุ่มปรากฏดังตารางที่ 2

ตารางที่ 2 จำนวนผู้ตรวจจำนวนตามวิธีการตรวจและเพศ

เพศ	วิธีการตรวจ		รวม
	ประเมินรวม	วิเคราะห์	
ชาย	7	6	13
หญิง	3	4	7
รวม	10	10	20

2.2 การสุ่มนักเรียน ทำการสุ่มสองครั้งดังนี้

2.2.1 สุ่มเพื่อทำการทดลองเครื่องมือ สุ่มชื่อโรงเรียนประถมศึกษาสังกัด สบข. อำเภอโนนสัง จังหวัดอุดรธานี จำนวน 1 โรงเรียน โดยวิธีการสุ่มอย่างง่าย ได้โรงเรียนชุมชนบ้านกุดคู้ซึ่งมีนักเรียนชั้นประถมศึกษาปีที่ 6 อยู่ 2 ห้องเรียน สุ่มห้องเรียนมา 1 ห้อง และใช้นักเรียนทั้งห้อง จำนวน 32 คน เป็นกลุ่มตัวอย่างใช้ในขั้นการพัฒนาแบบสอบความเรียง

2.2.2 สุ่มเพื่อเก็บรวบรวมข้อมูล ผู้วิจัยทำการทดสอบนักเรียนชั้นประถมศึกษาปีที่ 6 ของปีการศึกษา 2531 ที่เข้าศึกษาต่อในชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2532 ของโรงเรียนมัธยมศึกษาในเขตอำเภอโนนสัง ซึ่งมีทั้งหมด 3 โรงเรียน โดยทำการสอบในสัปดาห์แรกของภาคเรียนที่ 1 จึงได้ถือว่า ความสามารถที่ใช้ตอบแบบสอบความเรียงในการวิจัยครั้งนี้เป็นผลจากการเรียนในชั้นประถมศึกษาปีที่ 6 ของปีการศึกษา 2531 ใช้วิธีการสุ่มแบบแบ่งชั้นภูมิ (stratified random sampling) โดยให้โรงเรียนมัธยมศึกษาแต่ละโรงเรียนเป็นหนึ่งชั้นภูมิ (strata) แล้วสุ่มห้องเรียน โรงเรียนละ 1 ห้อง โดยวิธีสุ่มอย่างง่าย แต่ละห้องทำการสอบนักเรียนทุกคน แล้วสุ่มกระดาษคำตอบของนักเรียน ห้องละ 10 คน โดยวิธีสุ่มอย่างง่าย ได้กระดาษคำตอบของนักเรียนจำนวน 30 ชุด ซึ่งเป็นจำนวนที่ใกล้เคียงกับสถิติจำนวนนักเรียนชั้น

ประถมศึกษาปีที่ 6 ต่อหนึ่งห้องเรียน ในปีการศึกษา 2531 ของอำเภอโนนสัง ที่มีจำนวนนักเรียนเฉลี่ยต่อห้องเป็น 29.6 คน ใช้เป็นกระดาษคำตอบในการทดลองต่อไป แม้ว่าการสุ่มครั้งนี้เป็นการสุ่มแบบชั้นภูมิ แต่ในการศึกษาครั้งนี้ผู้วิจัยไม่ต้องการศึกษาอิทธิพลของโรงเรียน และห้องเรียน ผลการสุ่มปรากฏดังตารางที่ 3

ตารางที่ 3 จำนวนนักเรียนที่ใช้เป็นกลุ่มตัวอย่างในการเก็บรวบรวมข้อมูล

โรงเรียน	ห้องเรียน	นักเรียน ที่ทดสอบ	กระดาษคำตอบ ของนักเรียน
1. กุดคู้พิทยาคม	1	46	10
2. โนนสังวิทยาคาร	1	45	10
3. โนนเมืองวิทยาคาร	1	32	10
รวม	3	123	30

เนื่องจากนักเรียนชั้นประถมศึกษาปีที่ 6 ในพื้นที่บริการรอบ ๆ โรงเรียนมัธยมศึกษาทั้ง 3 แห่ง มีจำนวนใกล้เคียงกัน ดังนั้นแม้ว่านักเรียนชั้นมัธยมศึกษาปีที่ 1 ของโรงเรียนโนนเมืองวิทยาคารจะมีน้อยกว่าอีก 2 โรงเรียนที่เหลือ แต่เพื่อให้ได้กลุ่มตัวอย่างนักเรียนกระจายและเป็นสัดส่วนที่สอดคล้องกับจำนวนประชากร ผู้วิจัยจึงสุ่มนักเรียนแต่ละโรงเรียนในจำนวนที่เท่ากัน

2.3 การสุ่มข้อสอบ ผู้วิจัยได้สร้างข้อสอบวัดความคิดเห็นเกี่ยวกับข่าวและเหตุการณ์ตามลักษณะเฉพาะของข้อสอบ (item specification) จำนวน 5 ข้อ แม้ว่าในทางปฏิบัติจะไม่ได้สุ่มจากเอกภพของข้อสอบจริง ๆ แต่ตามทฤษฎีการวัดถือว่า ข้อสอบที่สร้างขึ้นเพื่อวัดเรื่องใดเรื่องหนึ่งเป็นตัวอย่างสุ่มจากโดเมนเชิงสมมุติของข้อสอบนั้น ๆ (Nunnally 1976: 193-194) ดังนั้นข้อสอบที่ใช้ในการศึกษาครั้งนี้ ถือเป็นตัวอย่างสุ่ม มีขนาดเป็น 5 ($n_i = 5$) ใช้เวลาสอบ 75 นาที

การกำหนดความสัมพันธ์ของพหุคูณที่ศึกษา

ในการศึกษาครั้งนี้ ผู้วิจัยกำหนดความสัมพันธ์ของตัวประกอบที่ศึกษา ในชั้นการวิเคราะห์ ความแปรปรวน (G study) และในชั้นการอ้างอิงสรุป (D study) ในแต่ละสถานการณ์เป็นแบบ crossed design กล่าวคือ $p \times i \times r$ สำหรับชั้นการวิเคราะห์ความแปรปรวนและ $p \times i \times r$ สำหรับ ชั้นการอ้างอิงสรุป ความสัมพันธ์ของพหุคูณนี้หมายความว่า ในการสอบวัดแบบความเรียงครั้งนี้ นักเรียนทุกคนสอบด้วยข้อสอบชุดเดียวกัน แล้วผู้ตรวจทุกคนตรวจคำตอบของนักเรียนทุกคนและตรวจ ทุก ๆ ข้อ รูปแบบความสัมพันธ์ของพหุคูณแบบ crossed 3 พหุคูณนี้ องค์ประกอบของคะแนน (score components) และองค์ประกอบความแปรปรวน (variance components) 7 องค์ประกอบดังนี้

	<u>องค์ประกอบของคะแนน</u>	<u>องค์ประกอบความแปรปรวน</u>
	$X_{pir} = u$	$\sigma^2(X_{pir}) =$
นักเรียน	$p \quad + u_{p-u}$	$+ \sigma^2(p)$
ข้อสอบ	$i \quad + u_{i-u}$	$+ \sigma^2(i)$
ผู้ตรวจ	$r \quad + u_{r-u}$	$+ \sigma^2(r)$
ผลรวม		
	$p_i \quad + u_{pi} - u_p - u_i + u$	$+ \sigma^2(p_i)$
	$p_r \quad + u_{pr} - u_p - u_r + u$	$+ \sigma^2(p_r)$
	$i_r \quad + u_{ir} - u_i - u_r + u$	$+ \sigma^2(i_r)$
เศษเหลือ	$pir \quad + X_{pir} - u_{pi} - u_{pr} - u_{ir} + u_p + u_i + u_r - u$	$+ \sigma^2(pir)$
เมื่อ		

X_{pir} แทน คะแนนสังเกตของนักเรียน p ที่ตอบข้อสอบ i และได้รับการตรวจ จากผู้ตรวจ r

u แทน ค่าเฉลี่ยรวม (grand mean)

u_{p-u} แทน ผลของนักเรียน p

u_{i-u} แทน ผลของข้อสอบ i

$ur-u$ แทน ผลของผู้ตรวจ r

$upi-up-ui+u$ แทน ผลร่วมระหว่างนักเรียน p กับข้อสอบ i

$upr-up-ur+u$ แทน ผลร่วมระหว่างนักเรียน p กับผู้ตรวจ r

$uir-ui-ur+u$ แทน ผลร่วมระหว่างข้อสอบ i กับผู้ตรวจ r

$Xpir-upi-upr-uir+up+ui+ur-u$ แทน ผลร่วมระหว่างนักเรียน ข้อสอบ

และผู้ตรวจ ซึ่งปะปน (confound)

กับความคลาดเคลื่อนเชิงสุ่ม

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ได้แก่แบบสอบถามเรียงวัดจุดประสงค์การเรียนรู้ เรื่อง ข่าวก้าวและเหตุการณ์ของกลุ่มสร้างเสริมประสบการณ์ชีวิตชั้นประถมศึกษาปีที่ 6 และคู่มือการตรวจ ผู้วิจัยได้ดำเนินการพัฒนาเครื่องมือตามลำดับขั้นดังนี้

1. เลือกและวิเคราะห์จุดประสงค์ การสอบความเรียงเป็นการสอบที่ต้องลงทุนลงแรงมากทั้งผู้ตอบและผู้ตรวจ ดังนั้นการตัดสินใจใช้แบบสอบถามความเรียง จำเป็นต้องพิจารณาให้รอบคอบ ตรวจสอบจุดประสงค์ที่ต้องการวัดให้แน่ชัด Thorndike and Hagen (1977: 265) กล่าวว่า ถ้าต้องการวัดความจำควรใช้แบบสอบถามจะดีกว่า ถ้าวัดความเข้าใจ การประยุกต์ และการวิเคราะห์ใช้ได้ทั้งแบบปรนัยและแบบความเรียง แต่แบบสอบถามค่อนข้างดีกว่าเล็กน้อย ถ้าต้องการให้นักเรียนแสดงเหตุผล อธิบายความสัมพันธ์ทางความคิด ข้อสอบแบบความเรียง ประเภทจำกัดคำตอบจะใช้ได้ดีกว่า แต่ถ้าเป็นจุดประสงค์ขั้นสูงวิเคราะห์ และประเมินค่า ควรใช้แบบสอบถามความเรียงประเภทไม่จำกัดคำตอบ ในการเลือกและวิเคราะห์จุดประสงค์ ผู้วิจัยดำเนินการดังนี้

1.1 ตรวจสอบจุดประสงค์การเรียนรู้ ที่กองวิชาการ สบข. ใช้ประเมินคุณภาพทางการศึกษาสำหรับชั้นประถมศึกษาปีที่ 6 พบว่าจุดประสงค์ข้อที่ 41 ของกลุ่มวิชาสร้างเสริมประสบการณ์ชีวิต (กระทรวงศึกษาธิการ, สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ 2525: 5) ระบุว่า "ให้นักเรียนสามารถแสดงความคิดเห็นเกี่ยวกับข่าวและเหตุการณ์ต่าง ๆ

ได้" เป็นจุดประสงค์ที่เน้นความเป็นอิสระในการแสดงความคิดเห็น สะท้อนถึงพฤติกรรมความรู้ ความคิดระดับสูง สมควรได้รับการสอบวัดด้วยแบบสอบถามเรียง จึงเลือกจุดประสงค์ข้อนี้เป็น โดเมนของ เนื้อหาของผลการเรียนที่มุ่งวัดในการวิจัยครั้งนี้

1.2 ศึกษาพฤติกรรมเฉพาะของ "การแสดงความคิดเห็นเกี่ยวกับข่าวและ เหตุการณ์ต่าง ๆ" โดยอาศัยเอกสารหลักสูตร ได้แก่ แผนการสอนกลุ่มสร้าง เสริมประสบการณ์ ชีวิต ชั้นประถมศึกษาปีที่ 6 พบว่าการแสดงความคิดเห็นเกี่ยวกับข่าวและ เหตุการณ์ 2 ลักษณะ ได้แก่ (1) แสดงความคิดเห็น วิพากษ์วิจารณ์ สนับสนุน โต้แย้ง หรือ แสดงทัศนะของตนต่อ เหตุการณ์ หรือบุคคลในข่าว เช่น หลังจากที่ให้ฟังหรือให้อ่านข่าว แล้วจะถามว่า "ถ้านักเรียน อยู่ในเหตุการณ์ตามข่าวนั้น นักเรียนจะทำอย่างไร เพราะเหตุใด?" "จากข่าวหรือเหตุการณ์ที่ เสนอมา นั้น บุคคลในข่าวเป็นใคร" เป็นต้น (2) แสดงความคิดเห็น โดยวิเคราะห์ตาม รูปแบบมาตรฐาน เช่น ระบุแหล่งข่าว สรุปเนื้อหาข่าว แยกประเภทของข่าว เป็นต้น

1.3 ตรวจสอบเนื้อหาของข่าวและ เหตุการณ์ ตามหลักสูตรชั้นประถมศึกษาปีที่ 6 พบว่า แนวการจัดเนื้อหาเปิดกว้างมาก กล่าวคือ นักเรียนในระดับการศึกษาและวัยนี้สามารถ แสดงความคิดเห็นเกี่ยวกับข่าวและ เหตุการณ์ได้ทุกประเภท ดังตารางที่ 4 (กระทรวงศึกษาธิการ, กรมวิชาการ 2525: 504)

ตารางที่ 4 แนวการจัดเนื้อหาในการเรียนการสอนข่าวและ เหตุการณ์ ของชั้นประถมศึกษา

ป.1	ป.2	ป.3	ป.4	ป.5	ป.6
-ข่าวบ้านและ โรงเรียน	-ข่าวบ้านและ โรงเรียน	-ข่าวบ้านและ โรงเรียน	-ข่าวบ้านและ โรงเรียน	-ข่าวบ้านและ โรงเรียน	-ข่าวบ้านและ โรงเรียน
-ข่าวในราช สำนัก	-ข่าวในราช สำนัก	-ข่าวในราช สำนัก	-ข่าวในราช สำนัก	-ข่าวในราช สำนัก	-ข่าวในราช สำนัก
	-ข่าวราชการ	-ข่าวราชการ	-ข่าวราชการ	-ข่าวราชการ	-ข่าวราชการ
		-ข่าวในชุมชน	-ข่าวในชุมชน	-ข่าวในชุมชน	-ข่าวในชุมชน

ตารางที่ 4 แนวการจัดเนื้อหาในการเรียนการสอนข่าวและเหตุการณ์ ของชั้นประถมศึกษา

ป.1	ป.2	ป.3	ป.4	ป.5	ป.6
		-ข่าวเศรษฐกิจ (ในชุมชน)	-ข่าวเศรษฐกิจ (ในชุมชน)	-ข่าวเศรษฐกิจ (ในชุมชน)	-ข่าวเศรษฐกิจ (ในชุมชน)
			-ข่าวต่างประเทศ	-ข่าวต่างประเทศ	-ข่าวต่างประเทศ
			เทศ	เทศ	เทศ
			(เน้นประเทศเพื่อนบ้าน)	(เน้นประเทศเพื่อนบ้าน)	(เน้นประเทศเพื่อนบ้าน)
		-ข่าวสำคัญ ทั่วไป	-ข่าวสำคัญ ทั่วไป	-ข่าวสำคัญ ทั่วไป	
		(ข่าวทั่วโลก)	(ข่าวทั่วโลก)	(ข่าวทั่วโลก)	
			-ข่าวต่างประเทศที่มี ความสัมพันธ์ กับประเทศไทย	-ข่าวต่างประเทศที่มี ความสัมพันธ์ กับประเทศไทย	
					-ข่าวเศรษฐกิจ และสังคม
					-ข่าวการเมือง

1.4 เลือกพฤติกรรมการแสดงความคิดเห็นและเนื้อหา จากการพิจารณาทั้งพฤติกรรม "การแสดงความคิดเห็น" และ "เนื้อหาข่าวและเหตุการณ์" ที่กล่าวมา ผู้วิจัยตัดสินใจเลือกการแสดงความคิดเห็นเชิงวิพากษ์วิจารณ์ สนับสนุนโต้แย้ง แสดงทัศนะต่อข่าว และเลือกข่าวสำคัญทั่วไปเฉพาะข่าวภายในประเทศเป็นเนื้อหาของเครื่องมือในการวิจัย เนื่องจากพิจารณาเห็น

ว่า พฤติกรรมดังกล่าวแสดงถึงการบูรณาการความรู้ ความคิด การสังเคราะห์ และประเมินค่า ซึ่งตรงกับลักษณะพฤติกรรมที่สมควรวัดโดยแบบสอบความเรียง และตรงกับจุดประสงค์การเรียนรู้ที่ต้องการมากที่สุด การเลือกข้อสำคัญในประเทศเป็นเพราะพิจารณาเห็นว่า เป็นเรื่องที่ใกล้ตัว น่าจะทำให้ นักเรียนมีความรู้พื้นฐาน สามารถอภิปรายได้กว้างขวาง

2. การพัฒนาแบบสอบ ดำเนินการสร้างตามลำดับขั้นดังนี้

ขั้นที่ 1. กำหนดลักษณะเฉพาะของข้อสอบ เพื่อจะได้ทราบขอบเขตของพฤติกรรม การแสดงความคิดเห็น ผู้วิจัยได้สัมภาษณ์ผู้เชี่ยวชาญซึ่งมีประสบการณ์ในการประเมินผลโครงการ การศึกษาของกระทรวงศึกษาธิการ มีวุฒิปริญญาเอกทางการศึกษา จำนวน 5 คน (รายชื่อ แสดงในภาคผนวก ข) ส่วนมากให้ความเห็นว่า ความคิดเห็นวิพากษ์วิจารณ์ หรือโต้แย้งก็ตาม จะต้องเป็นไปในทิศทางที่พึงประสงค์ ครอบคลุมพฤติกรรมย่อย 2 พฤติกรรม ดังนี้ คือ (1) ความคิดเห็นมีคุณค่าเชิงจริยธรรม (2) ความคิดเห็นถูกต้องตามหลักวิชา ดังนั้นในการกำหนดลักษณะเฉพาะของข้อสอบครั้งนี้จึงกำหนดพฤติกรรมย่อย 2 พฤติกรรมดังนี้

ลักษณะเฉพาะของข้อสอบความเรียงที่ใช้ในการวิจัย (Item Specifications)

ก. พฤติกรรมที่ต้องการวัด

แสดงความคิดเห็นเกี่ยวกับข่าวหรือเหตุการณ์ต่าง ๆ ได้

ข. พฤติกรรมย่อยที่ 1 สามารถแสดงความคิดเห็นเกี่ยวกับพฤติกรรมของบุคคล ในข่าวหรือเหตุการณ์ ด้วยเหตุผลเชิงจริยธรรม

ค. คำอธิบายทั่วไป

กำหนดเนื้อหาข่าวจากหนังสือพิมพ์หรือเหตุการณ์อย่างใดอย่างหนึ่งให้ภายในกรอบสี่เหลี่ยม ให้นักเรียนอ่านข่าวหรือเหตุการณ์นั้น แล้วให้แสดงความคิดเห็นว่า ถ้าตนเป็นบุคคล ในข่าวจะปฏิบัติอย่างไร พร้อมทั้งแสดงเหตุผลสนับสนุนการกระทำของตนเอง

ง. ลักษณะคำถาม

(1) เป็นข้อสอบแบบสถานการณ์ ใช้ข่าวหรือเหตุการณ์เป็นสถานการณ์ ความยาวไม่เกิน 10 บรรทัดตัวพิมพ์ดีด

(2) ใช้ภาษาที่นักเรียนชั้นประถมศึกษาปีที่ 6 สามารถอ่านเข้าใจง่าย

(3) คำชี้แจงในการตอบจะต้องระบุแนวการตอบให้นักเรียนเข้าใจชัดเจนว่า จะต้องใช้ความรู้ความสามารถด้านใด ในที่นี้ ให้นักเรียนแสดงเหตุผลประกอบการปฏิบัติของตน ที่สอดคล้องกับข้อพร้อมทั้งชี้แนะให้ยกตัวอย่างหรือข้อมูล ประกอบเหตุผลเพื่อให้คำตอบมีความสมบูรณ์ยิ่งขึ้น

(4) แจ้งเกณฑ์การตรวจให้คะแนนไว้ในคำชี้แจง เพื่อให้นักเรียนรู้แต่แรกว่า คำตอบที่ดีจะต้องประกอบด้วย ความคิดเห็นและเหตุผลตรงประเด็น สามารถยกเหตุผลประกอบ เหตุผลมีความเป็นไปได้ และมีคุณค่า แต่จะไม่พิจารณาความถูกต้องในการสะกตการันต์และความสะอาดเรียบร้อย

จ. ลักษณะคำตอบ

นักเรียนจะต้องตอบโดยการเขียนบรรยายในกระดาษคำตอบที่จัดให้โดยไม่มีเว้นบรรทัด

ฉ. ตัวอย่างข้อสอบ

ข้อ 0. ข่า

หนังสือพิมพ์ไทยรัฐ ฉบับวันที่ 23 พฤศจิกายน พ.ศ.2532 ลงข่าวว่า นายวัชรินทร์ เจริญวัชร ได้ขับรถเก๋งส่วนตัวพาลูกเมียมุ่งหน้าเข้ากรุงเทพ พอขึ้นสะพานลอยบางเขน ได้ชนกับรถบรรทุกสินค้าขนาดใหญ่อย่างจังจนเสียหาย ลูกเมียได้รับบาดเจ็บ โดยเฉพาะลูกถึงขนาดแขนขาหัก แต่รถบรรทุกสินค้าได้ชะตะบึงหนีไป ขณะนั้น นายอนุสรณ์ พูนพิริยะ ซึ่งนั่งรถแท็กซี่ตามหลังมา เห็นเหตุการณ์จึงหยุดลงมากดู แล้วบอกให้นายวัชรินทร์พาลูกเมียไปโรงพยาบาล ส่วนตัวเองได้เข้าแจ้งตำรวจให้ตามจับรถบรรทุกสินค้าคันนั้น ในที่สุดตำรวจก็สามารถตามจับคนขับรถบรรทุกสินค้าได้

คำถาม

ถ้านักเรียนเป็นคนขับรถบรรทุกสินค้าขนาดใหญ่คันนั้น นักเรียนจะทอย่างไรให้แสดงความคิดเห็นอย่างเสรี กล่าวคือนักเรียนจะเลือกทำตามคนขับในข่า หรือจะทอย่างอื่นก็ได้ แต่ต้องสามารถยกเหตุผลสนับสนุนความคิดของตน เพื่อโน้มน้าวให้ผู้อ่านมีความเห็นคล้อยตาม หากสามารถยกตัวอย่าง หรือข้อมูลประกอบเหตุผลได้ด้วย จะทำให้คำตอบมีความสมบูรณ์มากยิ่งขึ้น

ข. พดติกรรมย่อที่ 2 ให้นักเรียนสามารถแสดงความคิดเห็นเกี่ยวกับเหตุการณ์ในข่าว โดยใช้เหตุผลโดยอ้างอิงหลักวิชา

ช. คำอธิบายทั่วไป

กำหนดเนื้อหาข่าวจากหนังสือพิมพ์ให้ภายในกรอบสี่เหลี่ยม ให้นักเรียนอ่านข่าวนั้น แล้วให้แสดงความคิดเห็นเกี่ยวกับ การกระทำหรือปรากฏการณ์ในข่าว โดยใช้เหตุผลประกอบความเชื่ออย่างใดอย่างหนึ่งซึ่งสอดคล้องกับเหตุการณ์ในข่าว

ฉ. ลักษณะคำถาม

(1) เป็นข้อสอบแบบสถานการณ์ โดยใช้ข่าวจากหนังสือพิมพ์รายวัน เป็นตัวสถานการณ์ ความยาวไม่เกิน 10 บรรทัดตัวอักษรพิมพ์ดีด

(2) ใช้ภาษาที่นักเรียนประถมศึกษาปีที่ 6 สามารถเข้าใจได้ง่าย

(3) คำชี้แจงในการตอบจะต้องระบุแนวการตอบ ให้นักเรียนเข้าใจชัดเจนว่าจะต้องใช้ความรู้ความสามารถด้านใดในการตอบ ในที่นี้ให้นักเรียนแสดงเหตุผลประกอบความเชื่อการกระทำ หรือปรากฏการณ์อย่างใดอย่างหนึ่งในข่าว พร้อมทั้งชี้แนะให้ยกตัวอย่างหรือข้อมูลประกอบเหตุผลเพื่อให้คำตอบมีความสมบูรณ์มากยิ่งขึ้น

(4) แจ้งเกณฑ์การตรวจให้คะแนน ในคำชี้แจง เพื่อให้ นักเรียนรู้แต่แรกว่า คำตอบที่ดีประกอบไปด้วยคุณสมบัติใดบ้าง ในที่นี้จะประกอบด้วยความคิดเห็นและ เหตุผลตรงประเด็น สามารถยกเหตุผลประกอบได้ เหตุผลมีความเป็นไปได้ และมีคุณค่า

ญ. ลักษณะคำตอบ

นักเรียนจะต้องตอบโดยการเขียนบรรยายในกระดาษคำตอบที่จัดให้โดยไม่มีเว้นบรรทัด

ด. ตัวอย่างข้อสอบ

ข้อ 00. ข่าว

ในเหตุการณ์นี้ท่ามกลางภาคใต้เมื่อเดือนพฤศจิกายน พ.ศ. 2531 ซึ่งท่าความเสียหายให้กับพี่น้องภาคใต้อย่างหนักนั้น มีเหตุการณ์แปลกประหลาดอย่างหนึ่งเกิดขึ้นคือ มีโคลนไหลทะลักจากยอดเขาลงมาทับถมบ้านเรือนและ ไร่กสิกรรมชีวิตราษฎร จมหายไปไหนเลยโคลนอย่างน่าอนาถ เมื่อน้ำลดลงทะเลโคลนนี้นั้นก็ยังเป็นอุปสรรคอย่างหนักในการค้นหาศพผู้เสียชีวิตอีกด้วย

คำถาม

นักเรียนเชื่อหรือไม่ว่า การตัดไม้ทำลายป่าบนภูเขาเป็นเหตุการณ์สำคัญทำให้เกิดเหตุการณ์ประหลาดนั้น ให้แสดงความคิดเห็นอย่างเสรี กล่าวคือ นักเรียนจะ เชื่อหรือไม่เชื่อก็ได้ แต่ต้องยกเหตุผลสนับสนุนความเชื่อของตนเอง เพื่อโน้มน้าวให้ผู้อ่านมีความคิดเห็นคล้ายตาม หากสามารถยกตัวอย่างหรือข้อมูลประกอบเหตุผลได้ด้วย จะทำให้คำตอบมีความสมบูรณ์มากยิ่งขึ้น

ขั้นที่ 2. เขียนข้อความให้สอดคล้องกับลักษณะเฉพาะของข้อสอบจำนวน 8 ข้อ

ขั้นที่ 3. ตรวจสอบความตรงตามวัตถุประสงค์ โครงสร้างและความเหมาะสมกับระดับขั้นที่สอบ ตามหลักการสร้างแบบสอบความเรียง เมื่อเขียนข้อความเสร็จควรให้เพื่อนครู หรือผู้เชี่ยวชาญอย่างน้อย 1 คน ตรวจสอบในด้านต่อไปนี้ (1) ข้อความตรงกับจุดประสงค์หรือไม่ (2) โครงสร้างข้อความชัดเจนหรือไม่ (3) ความยากและความเหมาะสม (Hopkins and Stanley 1981:221-222; Gronlund 1977:80;83; Ebel and Frisbie 1986: 132; Coffman 1971: 288) ผู้วิจัยจึงได้ดำเนินการดังนี้

1. ให้ผู้เชี่ยวชาญซึ่ง เป็นนักวิชาการฝ่ายประเมินผลหลักสูตร กองวิชาการ สบช. จำนวน 10 คน พิจารณาความสอดคล้องระหว่างคำถามแต่ละข้อกับจุดประสงค์ ความเหมาะสมของ โครงสร้าง และความยากของข้อสอบ
2. วิเคราะห์ความสอดคล้องระหว่างคำถามแต่ละข้อกับจุดประสงค์ ความเหมาะสมของ โครงสร้าง และความยากของข้อสอบ โดยรวมคะแนนความคิดเห็นในข้อ เห็นด้วย (+1) ไม่แน่ใจ (0) ไม่เห็นด้วย (-1) แล้วหาค่าเฉลี่ย รวมทั้งแจกน้ำหนักของข้อ เสนอแนะ
3. คัดลึนความตรง ความเหมาะสมของโครงสร้างและความยาก โดยใช้เกณฑ์ความสอดคล้องระหว่างความคิดเห็นของผู้เชี่ยวชาญเฉลี่ยอย่างต่ำ .50 (Hamphill and Westic (1950 อ้างถึงใน โกวิท ประวาลพตษ์ และ สมศักดิ์ ลินธุ์เวชย์ 2527: 226) ผลการวิเคราะห์ปรากฏดังตารางที่ 5 ดังนี้

ตารางที่ 5 ค่าเฉลี่ยความคิดเห็นของผู้เชี่ยวชาญ 10 คน ในการประเมินความเหมาะสมของ
ข้อสอบ

รายการประเมิน	ข้อสอบข้อที่							
	1	2	3	4	5	6	7	8
1 ความวัดตรงจุดประสงค์	.8	.8	.8	.8	.8	.8	.8	.8
2 ความระบุความสามารถที่ต้องใช้ตอบได้ชัดเจน	.8	.8	.8	.7	.8	.8	.7	.6
3 คำชี้แจงในการให้คะแนนชัดเจน	.6	.7	.8	.6	.8	.7	.8	.7
4 เกณฑ์การให้คะแนนเหมาะสม	.3	.4	.4	.4	.4	.4	.4	.4
5 ใช้ภาษาเหมาะสมกับระดับชั้นที่วัด	.5	.7	.7	.8	.6	.6	.7	.6
6 เนื้อหาของข้อเหมาะสมกับระดับชั้นที่วัด	.4	.8	.6	.7	.4	.8	.7	.7

ผลการวิเคราะห์พบว่าข้อสอบทั้ง 8 ข้อ วัดตรงตามจุดประสงค์ที่ต้องการ ส่วนความเหมาะสมของโครงสร้างคำถามและความยาก ยังต้องปรับปรุง กล่าวคือ โครงสร้างคำถามเกี่ยวกับเกณฑ์การให้คะแนนที่ระบุไว้ในข้อคำถามที่ว่า "ในการให้คะแนนนั้น จะไม่พิจารณาความถูกต้องในการสะกดการันต์ และความสะอาดเรียบร้อย" ผู้เชี่ยวชาญเสนอแนะว่าควรตัดออก โดยให้เหตุผลสอดคล้องกันว่า จะทำให้นักเรียนไม่ตั้งใจตอบและยังส่งเสริมให้เด็กเป็นคนขาดระเบียบ ส่วนความยาก และความเหมาะสมพบว่า ข้อสอบข้อที่ 1 และข้อที่ 5 ผู้เชี่ยวชาญส่วนใหญ่เห็นว่า ไม่เหมาะสมกับระดับชั้นที่วัดเนื่องจากข้อหนึ่งเป็นข้ออาชญากรรมที่อาจเป็นตัวอย่างย่ำแย่ในทางไม่ดี ส่วนอีกข้อหนึ่ง เป็นข้อทางการแพทย์ ซึ่งใช้ความรู้สูงเกินความสามารถของนักเรียนชั้นประถมศึกษาปีที่ 6 เมื่อพิจารณาความคิดเห็นของผู้เชี่ยวชาญ ผู้วิจัยจึงปรับปรุงโครงสร้างคำถาม โดยตัดเกณฑ์การตรวจที่กล่าวถึงความถูกต้องในการสะกดการันต์และความสะอาดเรียบร้อยออก พร้อมทั้งปรับปรุงคำชี้แจงการให้คะแนนรายชื่อ โดยยกบางส่วนของข้อที่ชี้กันไปไว้ในคำชี้แจงรวม ตัดข้อสอบออกไป 2 ข้อ เหลือข้อสอบที่จะนำไปทดลองใช้ 6 ข้อ

ขั้นที่ 4. การทดลองแบบสอบ หน้าแบบสอบความเรียงที่ปรับปรุงโครงสร้างแล้ว จำนวน 6 ข้อ ใบทดลองสอบกับนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนชุมชนบ้านกุดคู อำเภอนิคมสังข์จังหวัดอุดรธานี จำนวน 32 คน เพื่อให้นักเรียนตอบอย่างตั้งใจ และทำเต็มความสามารถ ผู้วิจัยได้มอบให้ครูประจำวิชาสร้างเสริมประสบการณ์ชีวิต เป็นผู้ดำเนินการสอบ ให้ดำเนินการดังนี้

1. ชี้แจงกับนักเรียนว่าเป็นการสอบเก็บคะแนนที่ค้างอยู่ ขอให้ นักเรียนตั้งใจทำเต็มความสามารถ
2. แจกข้อสอบทีละข้อพร้อมกระดาษสำหรับเขียนคำตอบ ให้นักเรียนเขียนชื่อในกระดาษคำตอบ เสร็จแล้วจึงให้ลงมือตอบพร้อมกัน
3. เมื่อนักเรียนคนใดตอบเสร็จ เก็บกระดาษคำตอบเสร็จให้ครูบันทึกเวลาไว้ที่หัวกระดาษ
4. ผู้วิจัยวิเคราะห์ผลการทดลองแบบสอบ เพื่อตรวจสอบความเหมาะสมของภาษา ความยาก และเวลาที่ใช้สอบ

ผลการวิเคราะห์พบว่านักเรียนเกือบทุกคนตอบคำถามได้ตรงประเด็น ตามที่ข้อคำถามต้องการ แสดงให้เห็นว่าภาษาที่ใช้ในคำถาม มีความเหมาะสมกับระดับชั้นของนักเรียน ส่วนความยากของข้อสอบเมื่อพิจารณาจากความยาวของคำตอบ พบว่า มีข้อสอบอยู่ 1 ข้อ ซึ่งให้แสดงความคิดเห็นเกี่ยวกับความก้าวหน้าทางวิทยาศาสตร์ นักเรียนส่วนมากตอบได้ไม่เกิน 3 บรรทัด ในขณะที่ข้ออื่น ๆ ตอบได้เกิน 8 บรรทัด แสดงว่าข้อสอบข้อนั้นยากเกินไป สำหรับนักเรียนในประชากรที่ศึกษา หรืออาจเป็นเพราะ ไม่มีประสบการณ์เพียงพอ ส่วนเวลาที่นักเรียนประมาณ 90 % ใช้ในการตอบ อยู่ระหว่าง 12-18 นาที จากผลการทดลองใช้ข้อสอบ ผู้วิจัยจึงตัดข้อสอบออกไป 1 ข้อ เหลือข้อสอบที่นำไปใช้เก็บรวบรวมข้อมูลจริง ๆ 5 ข้อ ให้เวลาทำข้อละ 15 นาทีรวมเวลาที่ใช้สอบทั้งฉบับ 75 นาที

3. การพัฒนาคู่มือการตรวจ ดำเนินการตามลำดับขั้นดังนี้

1. ศึกษาเอกสารและผลงานวิจัยเกี่ยวกับวิธีการตรวจความเรียง พบว่า การตรวจมีอยู่ 2 วิธี ดังที่ได้กล่าวไว้ในบทที่ 2 ในแต่ละวิธีต้องกำหนดเกณฑ์ไว้ล่วงหน้าอย่างชัดเจน จึงจะให้ผลการวัดมีความเชื่อถือได้ ผู้วิจัยได้พัฒนาเกณฑ์แต่ละวิธีดังนี้

1.1 วิธีตรวจแบบประเมินรวม ใช้คำตอบแบบ (Essay model)

โดยการคัดเลือกจากคำตอบของนักเรียนจากการทดสอบจริง Mehrens and Lenmann (1972: 232) กล่าวว่า ผู้ตรวจสามารถคัดเลือกคำตอบแบบจากคำตอบของนักเรียนที่สอบในแต่ละครั้งได้โดยไม่จำเป็นต้องเขียนคำตอบขึ้นเอง ผู้วิจัยจึงได้ดำเนินการดังนี้

1.1.1 ตั้งกรรมการคัดเลือกคำตอบ ประกอบด้วยศึกษา
ศึกษานิเทศก์และครูวิชาการของอำเภอโนนสัง จังหวัดอุดรธานี จำนวน 3 คน

1.1.2 ผู้วิจัยชี้แจงวัตถุประสงค์ของการคัดเลือกคำตอบ
จุดประสงค์การเรียนรู้วัด ลักษณะแบบสอบที่ใช้วัด ระดับชั้นของนักเรียนที่ทดสอบ เพื่อให้คณะกรรมการมีความรู้พื้นฐานตรงกัน

1.1.3 ให้คณะกรรมการคัดเลือกคำตอบของนักเรียนที่ทำการสอบ
จริงแต่ไม่ถูกส่งเป็นคำตอบสำหรับให้ครูตรวจในชั้นการทดลอง โดยให้พิจารณาคัดเลือกที่มีคุณภาพ
ต่างกัน 3 ระดับ คือ คุณภาพต่ำสุด คุณภาพปานกลาง และคุณภาพสูงสุด อย่างละ 1 คำตอบ แต่
ละระดับคุณภาพ สามารถให้คะแนนระบบเต็ม 10 ได้ดังนี้ คุณภาพต่ำสุด ให้ 1 คะแนน คุณภาพ
ปานกลางให้ 5 คะแนน และ คุณภาพสูงสุด ให้ 10 คะแนน

1.1.4 เมื่อได้คำตอบแบบครบทุกข้อ แต่ละข้อมีครบ 3 ระดับ
แล้ว เขียนคะแนนบนหัวกระดาษให้ชัดเจน ถ่ายเอกสารให้ได้ 10 ชุด ตามจำนวนครูที่ตรวจวิธี
ประเมินรวม

1.2 วิธีตรวจแบบวิเคราะห์ ผู้วิจัยดำเนินการตามลำดับขั้นดังนี้

1.2.1 กำหนดประเด็นการตรวจ เป็น 4 ด้าน คือ (1) ด้าน
เนื้อหา (2) ด้านการเรียบเรียง (3) ด้านลีลาและความงดงามทางภาษา และ (4) ด้านกล
ไกการเขียน พร้อมทั้งกำหนดประเด็นย่อย ๆ ในแต่ละด้านเพื่อให้มีความชัดเจนยิ่งขึ้น (Baker
1982: 291-297; Wesdorp, et al. 1982: 299-315)

1.2.2 ให้ผู้เชี่ยวชาญ 3 ท่านที่เคยช่วยวิเคราะห์พฤติกรรมการ
แสดงความคิดเห็น พิจารณาความเหมาะสมของประเด็นการให้คะแนนที่ผู้วิจัยกำหนด

ผลการวิเคราะห์พบว่า ผู้เชี่ยวชาญให้ความเห็นสอดคล้อง
กันว่า ประเด็นที่ผู้วิจัยกำหนด ไม่ตรงกับผลการเรียนที่มุ่งวัด โดยเฉพาะ 2 ประเด็นหลัก เป็น

ความสามารถทางภาษาพิเศษ ไม่ใช่ความสามารถด้านการแสดงความคิดเห็น นอกจากนั้น การกำหนดประเด็นย่อย ๆ มากเกินไป ยังทำให้ผู้ตรวจตัดสินใจลำบาก ผู้เชี่ยวชาญให้ข้อเสนอแนะในตนเองเดี๋ยวกันว่า ควรรวมประเด็นย่อย ๆ ของด้านเนื้อหาเข้ากันให้เหลือเพียง แสดงความคิดเห็นและยกเหตุผลได้ตรงประเด็นไม่ออกนอกเรื่อง ยกตัวอย่างหรือมีข้อมูลอื่นสนับสนุนความคิดเห็นและมีคุณค่าเชิงคุณธรรมหรือถูกต้องตามหลักวิชา ผู้วิจัยจึงกำหนดประเด็นการตรวจใหม่ให้เหลือเพียง 3 ด้าน คือ (1) แสดงความคิดเห็นหรือเหตุผลได้ตรงประเด็น (2) ยกตัวอย่างหรือข้อมูลสนับสนุน (3) ความคิดเห็นมีคุณค่าเชิงคุณธรรมหรือถูกต้องตามหลักวิชา

1.2.3 การกำหนดน้ำหนักคะแนน ผู้วิจัยได้ขอให้นักวิชาการฝ่ายประเมินผลหลักสูตร กองวิชาการ สบช. ซึ่งเคยตัดสินความตรงเชิงเนื้อหาและความเหมาะสมของข้อสอบจำนวน 10 คน ให้นักคะแนน พบว่าค่าเฉลี่ยของคะแนนในแต่ละประเด็น เมื่อกำหนดคะแนนเต็มเป็น 10 ได้ผลดังตารางที่ 6 ดังนี้

ตารางที่ 6 ค่าเฉลี่ยน้ำหนักคะแนนในแต่ละประเด็นการตรวจของผู้เชี่ยวชาญจำนวน 10 คน

ประเด็นที่ต้อง	ค่าเฉลี่ย	
	ค่าความ	ค่าประมาณ
การให้คะแนน		
1. แสดงความคิดเห็นหรือผลได้ตรงประเด็น	3.2	3
2. ยกตัวอย่างหรือข้อมูลสนับสนุน	2.9	3
3. ความคิดเห็นมีคุณค่าเชิงจริยธรรมหรือถูกต้องตามหลักวิชา	4.3	4

ดังนั้น ในการตรวจคำตอบความเรียง โดยวิธีวิเคราะห์ในการวิจัยครั้งนี้ จึงให้คะแนนประเด็นที่ต้องการเป็น 3, 3 และ 4 รวม 10 คะแนน

1.2.4 เขียนคำอธิบายวิธีการตรวจ เพื่อให้ผู้ตรวจวิธีวิเคราะห์ ให้คะแนนได้ตรงกัน ผู้วิจัยดำเนินการดังนี้

(1) เขียนคำอธิบายวิธีให้คะแนนประเด็นที่ 1 และที่ 2 ในแต่ละประเด็นชี้ให้เห็นระดับคุณภาพของคำตอบแตกต่างกันเป็น 3 ระดับ สอดคล้องกับน้ำหนักคะแนน ในหัวข้อ 1.2.3

(2) ให้ครูผู้ที่มีประสบการณ์เคยสอนในชั้นประถมศึกษาปีที่ 6 จำนวน 2 คน พิจารณาว่าข้อความที่อธิบายไว้สามารถอ่านเข้าใจและให้คะแนนได้จริงหรือไม่ ผลปรากฏว่าครูทั้งสองให้ความเห็นตรงกันว่าข้อความที่อธิบายวิธีการให้คะแนนสามารถอ่านเข้าใจและนำไปใช้ในการให้คะแนนได้

(3) สำหรับคำอธิบายวิธีการตรวจประเด็นที่ 3 ด้านคุณค่าเชิงจริยธรรม ผู้วิจัยใช้เกณฑ์การให้คะแนนเหตุผลเชิงคุณธรรม ตามเกณฑ์ของกรมวิชาการ (2526 ข: 104)

(4) สำหรับคำอธิบายประเด็นการตรวจด้านความถูกต้องตามหลักวิชา ผู้วิจัยศึกษาจากเนื้อหาในแผนการสอน กลุ่มวิชาสร้างเสริมประสบการณ์ชีวิต ชั้นประถมศึกษาปีที่ 6 ของกรมวิชาการ (2526 ก: 214-217) และ สัมภาษณ์ศึกษานิเทศก์ และครูวิชาการ ของสำนักงานการประถมศึกษาอำเภอโนนสัง จังหวัดอุดรธานี (คู่มือการตรวจที่ใช้ในการเก็บรวบรวมข้อมูลครั้งนี้แสดงในภาคผนวก ก)

กระบวนการทดลองในการวิจัย

1. การเตรียมการก่อนการทดลอง ผู้วิจัยเตรียมเอกสารต่าง ๆ ที่ใช้ในการทดลอง พร้อมทั้งดำเนินการดังนี้

1.1 คำตอบความเรียง

1.1.1 ผู้วิจัยติดต่อประสานงานกับผู้บริหารโรงเรียนที่เป็นกลุ่มตัวอย่าง ได้แก่ โรงเรียนกุดคูพิทยานุสรณ์ โรงเรียนโนนสังวิทยาคาร และโรงเรียนโนนเมืองวิทยาคาร เพื่อขอสอบนักเรียนชั้นมัธยมศึกษาปีที่ 1 โรงเรียนละ 1 ห้องเรียน

1.1.2 หลังจากได้คำตอบแล้ว สุ่มกระดาษคำตอบของนักเรียนโรงเรียน

ละ 10 คน รวมเป็น 30 คน

1.1.3 ถ่ายสำเนาคำตอบของนักเรียน 30 คนเป็น 20 ชุด เตรียมไว้ให้ครูจำนวน 20 คนตรวจ

1.2 คำตอบตัวแบบ

1.2.1 คัดเลือกคำตอบของนักเรียนที่เหลือ จากข้อ 1.1.2 โดยขอความร่วมมือจากศึกษานิเทศก์และครูวิชาการของอำเภอ พิจารณาคัดเลือกคำตอบที่มีคุณภาพต่างกัน ข้อละ 3 ระดับ ดังที่กล่าวไว้ในหัวข้อการพัฒนาคู่มือการตรวจ

1.2.2 ถ่ายสำเนาคำตอบตัวแบบข้อละ 10 ชุด เพื่อเตรียมไว้ให้ครู 10 คนที่ตรวจวิธีประเมินรวม

1.3 เอกสารผลการเรียนของนักเรียน

1.3.1 ผู้วิจัยสมมุติผลการเรียนของนักเรียนแต่ละคน เป็น 2 ลักษณะ คือ มีผลการเรียนดี กับ ผลการเรียนอ่อน โดยเขียนข้อมูลลงในสมุดรายงานของนักเรียนส่วนที่เกี่ยวข้องกับรายงานผลการเรียนตามแบบที่โรงเรียนประถมศึกษาใช้ กรอกข้อมูลแสดงผลการเรียนแต่ละคนเป็น 2 ลักษณะ ดังนี้

(1) ข้อมูลผลการเรียนดี ระบุผลการสอบปลายภาคใน 5 กลุ่มวิชา ดังนี้ ให้เกรด 4 อย่างน้อย 3 กลุ่มวิชา ที่เหลือ 1 หรือ 2 กลุ่มวิชา ให้เกรด 3 เช่น เด็กชายสมศักดิ์ ได้เกรด กลุ่มวิชาทักษะ 3 ที่เหลืออีก 4 กลุ่ม ได้เกรด 4 จัดว่าเป็นคนมีผลการเรียนดี

(2) ข้อมูลผลการเรียนอ่อน ระบุผลการสอบปลายภาคใน 5 กลุ่มวิชาดังนี้ ให้เกรด 1 อย่างน้อย 3 กลุ่มวิชา ที่เหลือ 1 หรือ 2 กลุ่มวิชา ให้เกรด 2 จากตัวอย่าง เด็กชายสมศักดิ์ ได้เกรดกลุ่มวิชาภาษาอังกฤษเป็น 2 แต่อีก 4 กลุ่มวิชาที่เหลือได้เกรด 1 ทั้งหมด จัดว่าเป็นคนมีผลการเรียนอ่อน (แสดงในภาคผนวก ก)

1.3.2 ถ่ายสำเนาเอกสารผลการเรียนของนักเรียนแต่ละคน 20 ชุด ตามจำนวนผู้ตรวจ

1.3.3 เย็บเอกสารผลการเรียนรวมกัน ผู้ตรวจแต่ละคนจะรับรู้ผลการเรียนของนักเรียนคนหนึ่ง ๆ เพียงด้านเดียว คือ เรียนดีหรืออ่อนเท่านั้น ดังนั้น แม้ว่าจะจัดทำผลการเรียนของเด็กนักเรียนแต่ละคนเป็น 2 ด้าน ในการเย็บรวมเพื่อเสนอ

ต่อครูจะ ไม่มีผลของนักเรียนคนเดียวกันขึ้นอยู่กับตัวนักเรียน

2. การดำเนินการทดลอง

2.1 สถานการณ์ที่ 1 ครูตรวจโดยใช้ประสบการณ์วิชาชีพของตนเอง

2.1.1 ผู้วิจัยติดต่อขอใช้อาคารเรียน โรงเรียนกุดชุมพยูวาม เป็นสถานที่ตรวจ

ที่ตรวจ

2.1.2 จัดผู้ตรวจแยกเป็น 2 ห้องโดยวิธีสุ่มตามที่กล่าวมาแล้ว แต่ละห้องอยู่ห่างกันเว้นระยะ 3 ห้องเรียนเพื่อป้องกันการได้ยินคำชี้แจงวิธีการตรวจที่แตกต่างกัน

2.1.3 แจกเอกสารที่ใช้ในการตรวจ ได้แก่ ข้อสอบ กระดาษคำตอบของนักเรียนซึ่งเก็บรวมกันทั้ง 30 คำตอบ และกระดาษรอกคะแนนซึ่งเรียง เลขที่นักเรียนจาก 1 ถึง 30 มีช่องทางขวามือให้กรอกคะแนน

2.1.4 ผู้วิจัยชี้แจงต่อผู้ตรวจทั้งสองกลุ่มว่า "กระดาษคำตอบนี้เป็นของนักเรียนชั้น ป.6 โรงเรียนแห่งหนึ่ง ครูประจำวิชากลุ่มสร้างเสริมประสบการณ์ชีวิตที่สอนนักเรียนกลุ่มนี้ ต้องการประเมินผลการเรียนรู้ของนักเรียน ตามจุดประสงค์การเรียนรู้ที่หลักสูตรระบุไว้ว่า ให้นักเรียนสามารถแสดงความคิดเห็นเกี่ยวกับข่าวหรือเหตุการณ์ต่าง ๆ ได้ เขาจึงสร้างข้อสอบแบบบรรยาย วัดความสามารถในการแสดงความคิดเห็นเกี่ยวกับข่าวและ เหตุการณ์ต่าง ๆ จำนวน 5 ข้อตั้งที่ท่านมีอยู่ในมือขณะนี้ สมมุติว่าท่านเป็นครูคนนั้น ขอให้ท่านให้คะแนนคำตอบของนักเรียนกลุ่มนี้ โดยกรอกคะแนนลงในใบกรอกคะแนนที่แจกให้ กรุณาอย่าขีดเขียนบนกระดาษคำตอบของนักเรียน"

2.1.5 กลุ่มที่ตรวจโดยวิธีประเมินรวมได้รับคำชี้แจงเพิ่มเติมว่า "ในการตรวจ ให้ท่านอ่านคำตอบของนักเรียนอย่างรวดเร็ว แล้วใช้ดุลยพินิจของท่านพิจารณาจากคุณภาพรวมของคำตอบแต่ละข้อ ให้คะแนนตามที่เห็นสมควร คำตอบที่ดีที่สุดจะได้คะแนนเต็ม 10 คะแนน ถ้ามีคุณภาพด้อยกว่า จะได้คะแนนลดหลั่นลง ไปขึ้นอยู่กับพิจารณาของท่าน ขอให้ถือว่าท่านเป็นผู้มีสิทธิ์ โดยสมบูรณ์ในการตัดสินผลการเรียนตามจุดประสงค์นี้"

2.1.6 กลุ่มที่ตรวจโดยวิธีวิเคราะห์จะได้รับคำชี้แจงเพิ่มเติมว่า "ในการตรวจ ให้ท่านอ่านคำตอบของนักเรียนอย่างละเอียด แล้วให้คะแนนแยกเป็นด้าน ๆ ตามที่เห็นสมควร แต่ละด้านอาจจะให้คะแนนแตกต่างกันได้ แต่เมื่อรวมแล้วจะต้องได้ไม่เกินร้อยละ 10"

คะแนน ส่วนจะให้คะแนนในประเด็นใดบ้างนั้น ขึ้นอยู่กับดุลยพินิจของท่านเอง อาจตรงหรือไม่ตรงกับคนอื่นก็ได้ ขอให้ถือว่า ท่านเป็นผู้มีสิทธิ์โดยสมบูรณ์ในการตัดสินผลการเรียนตามจุดประสงค์นี้"

2.1.7 เมื่อครูตรวจกระดาษคำตอบเสร็จทุกคนซึ่งใช้เวลาประมาณ 1 ชั่วโมงสำหรับการตรวจวิธีประเมินรวม และ 2 ชั่วโมงสำหรับการตรวจวิธีวิเคราะห์ ผู้วิจัยเก็บใบกรอกคะแนนและเขียนชื่อผู้ตรวจไว้หัวกระดาษ แล้วให้ครูหยุดพักรับประทานอาหารกลางวันเป็นเวลา 1 ชั่วโมง

2.2 สถานการณ์ที่ 2 ครูตรวจตามวิธีการที่ได้รับการชี้แจงจากผู้วิจัย

2.2.1 หลังจากที่ครูได้พักรับประทานอาหาร ผู้วิจัยได้จัดการชี้แจงให้ความรู้เกี่ยวกับหลักการและวิธีการตรวจคำตอบแบบบรรยาย เสนอแนะวิธีใช้คู่มือการตรวจที่ผู้วิจัยสร้างขึ้น โดยอบรมรวมกันทั้งสองกลุ่ม เนื้อหาที่ชี้แจงเป็นข้อความรู้เกี่ยวกับการวัดผลแบบความเรียงที่ค่อนข้างรวบรัดดังนี้

"ข้อสอบแบบความเรียง หมายถึงข้อสอบที่ต้องการคำตอบชนิดที่นักเรียนเป็นคนจัดหาหรือเขียนขึ้นเอง บางคนอาจเรียกว่าข้อสอบอัตนัย ข้อสอบประเภทนี้มีคุณค่าสูงยิ่งต่อการเรียนการสอน เป็นเครื่องมือการประเมินผลที่สร้างง่าย สามารถวัดความคิดขั้นสูงได้แก่ การวิเคราะห์ การสังเคราะห์และการประเมินค่า จินตนาการและความคิดสร้างสรรค์ได้ดี แต่แบบสอบชนิดนี้มีจุดอ่อนอยู่ที่ คะแนนที่ครูให้มักขึ้นอยู่กับตัวแปรภายนอกที่มีใช้ความสามารถที่แท้จริงของนักเรียน ครูที่ตรวจคำตอบของนักเรียนคนเดิม 2 ครั้ง จะให้คะแนนต่างกัน ถ้าให้ครู 2 คน ขึ้นไปตรวจก็จะให้คะแนนแตกต่างกันไป นั่นแสดงว่า คะแนนที่นักเรียนได้รับไม่ได้สะท้อนความสามารถที่แท้จริงของเขา แต่สะท้อนถึงความแตกต่างของสิ่งอื่น เช่น โอกาสที่ตรวจ ผู้ตรวจอารมณ์ของครู เป็นต้น ที่เป็นเช่นนี้เพราะผู้ตรวจแต่ละคนหรือในการตรวจแต่ละครั้ง ผู้ตรวจเปลี่ยนแปลงเกณฑ์หรือมาตรฐาน หรือใช้มาตรฐานที่แตกต่างกัน ครูบางคนชอบให้คะแนนสูง บางคนชอบให้ต่ำ บางโอกาสครูอารมณ์ดีนักเรียนได้คะแนนสูง แต่บางโอกาสครูอารมณ์ไม่ดี นักเรียนได้คะแนนต่ำ ผลการให้คะแนนลักษณะนี้ทางการวัดผลเรียกว่า คะแนนขาดความเชื่อมั่นหรือขาดความเที่ยง คะแนนที่ขาดความเที่ยงนี้ หากนำไปใช้ตัดสินใจทางการศึกษา เช่น การตัดสินได้-ตก หรือการประเมินโครงการทางการศึกษาอื่น ๆ จะมีผลเสียอย่างใหญ่หลวงได้

วิธีแก้ไข การที่จะให้คะแนนมีความเที่ยง ครูจะต้องควบคุมความแตกต่างที่ไม่ได้สะท้อนความสามารถที่แท้จริงของนักเรียนไว้ให้ได้ โดยมุ่งสนใจอยู่แต่เฉพาะผลการ

เรียนที่ต้องการวัด ทามาตรฐานที่มีอยู่ในใจให้ปรากฏออกมาเป็นกฎเกณฑ์ที่เป็นรูปธรรม เพื่อยึดถือ เป็นคู่มือในการตรวจอย่างคง เส้นคงวาไม่โอนเอน

2.2.2 หลังจากชี้แจงรวมกลุ่มแล้ว เพื่อให้การตรวจแต่ละวิธีมีอิสระต่อกันมากขึ้น ผู้วิจัยได้แยกดำเนินการชี้แจงตามวิธีการตรวจดังนี้

ก. กลุ่มที่ตรวจแบบประเมินรวม ผู้วิจัยแจกตัวแบบคำตอบให้ครูพร้อมชี้แจงว่า "ที่ท่านถืออยู่ในมือนั้นคือ มาตรฐานที่เป็นรูปธรรมอย่างหนึ่งตามที่ผู้วิจัยกล่าวถึง ขอให้ครูช่วยตรวจกระดาษคำตอบของนักเรียนอีกครั้ง คราวนี้ให้ปฏิบัติดังนี้

(1) อ่านคำตอบแต่ละข้อของนักเรียนทุก ๆ คนคร่าว ๆ 1 เที้ยว โดยไม่ต้องให้คะแนน

(2) อ่านคำตอบตัวแบบ ซึ่งถือว่าเป็นมาตรฐานที่เป็นรูปธรรมเพื่อให้เห็นภาพชัดว่า คำตอบลักษณะใดถือว่ามีคุณภาพต่ำที่ให้คะแนน 1 คะแนน คำตอบลักษณะใดถือว่ามีคุณภาพปานกลางที่ให้คะแนน 5 คะแนน และคำตอบลักษณะใดถือว่ามีคุณภาพสูงที่ให้คะแนน 10 คะแนน

(3) อ่านคำตอบของนักเรียนแต่ละข้อแล้วให้คะแนน โดยเทียบคุณภาพกับคำตอบตัวแบบในข้อ (2) ถ้ามีคุณภาพเท่าเทียมกับคำตอบตัวแบบใด ก็ให้คะแนนเท่ากับคะแนนในตัวแบบนั้น แต่ถ้ามีคุณภาพแตกต่างออกไป ก็ให้คะแนนแตกต่างไปตามที่เห็นสมควร

(4) เมื่อท่านตรวจเสร็จให้ตรวจซ้ำ โดยอ่านผ่าน ๆ อย่างรวดเร็ว ถ้าคะแนนตรงกันก็แล้วไป แต่ถ้าต่างกันให้ใช้คะแนนเฉลี่ย 2 ครั้ง นั้นแทนคะแนนความสามารถของนักเรียน

(5) ให้ตรวจทีละข้อจนครบทุกคนแล้วจึงเริ่มตรวจข้อใหม่"

ข. กลุ่มที่ตรวจแบบวิเคราะห์ ผู้วิจัยแจกคู่มือการตรวจให้ครู (คู่มือการตรวจนี้ผู้วิจัยได้กล่าวถึงอย่างละเอียดในหัวข้อเครื่องมือที่ใช้ในการวิจัยแล้ว) พร้อมชี้แจงว่า "ที่ท่านถืออยู่ขณะนี้คือ มาตรฐานที่เป็นรูปธรรมอย่างหนึ่งตามที่ผู้วิจัยกล่าวถึง ขอให้ครูช่วยตรวจกระดาษคำตอบของนักเรียนอีกครั้ง คราวนี้ให้ปฏิบัติดังนี้

(1) อ่านและทำความเข้าใจกับคู่มือการตรวจเป็นรายข้อ หากมีปัญหาให้ซักถามทันที (ในช่วงนี้ผู้วิจัยคอยกระตุ้นให้ครูได้ซักถามและชี้แจงจนเห็นว่าครูเข้าใจทุกคน)

- (2) ให้ตรวจทีละข้อ โดยให้คะแนนตามคู่มือการตรวจจนครบทุกคนจึง เริ่มตรวจข้อใหม่
- (3) ให้คะแนนแยกทีละประเด็น ตามที่แยกไว้ในกระดาษกรอกคะแนน
- (4) ในการตรวจแต่ละประเด็นหรือแต่ละข้อไม่ควรกลับไปพิจารณาคะแนน ของประเด็นหรือข้อที่ตรวจผ่านมาแล้ว พึงระลึกไว้ว่า คนที่เก่งประเด็นหนึ่งหรือข้อหนึ่ง อาจจะ ไม่เก่งประเด็นอื่นหรือข้ออื่นก็ได้"

2.2.3 เมื่อดำเนินการเสร็จตามสถานการณ์ที่ 2 ซึ่งใช้เวลาประมาณ 1 ชั่วโมง 30 นาทีสำหรับการตรวจวิธีประเมินรวม และประมาณ 3 ชั่วโมง สำหรับการตรวจวิธีวิเคราะห์ ผู้วิจัยเก็บอุปกรณ์ต่าง ๆ ที่แจกให้กับครูคืน กล่าวขอบคุณที่ให้ความร่วมมือ

2.3 สถานการณ์ที่ 3 ครูตรวจตามวิธีการที่ได้ชี้แจงจากผู้วิจัยแต่รับรู้ผลการเรียนของนักเรียน

2.3.1 หลังจากจบการตรวจตามสถานการณ์ที่ 2 ผ่านไป 1 สัปดาห์ ผู้วิจัยติดต่อกับผู้ตรวจอีกครั้ง เพื่อขอความร่วมมือในการตรวจครั้งที่สามและพบว่าครูให้ความร่วมมือเป็นอย่างดี

2.3.2 แจกกระดาษคำตอบ ใบกรอกคะแนน เกณฑ์การตรวจ และเอกสารแสดงผลการเรียนของนักเรียนให้ครูทุกคน

2.3.3 ผู้วิจัยชี้แจงว่า ในการตรวจครั้งนี้ให้ปฏิบัติเหมือนการตรวจที่ผู้วิจัยเคยเสนอแนะ ไว้เมื่อสัปดาห์ที่ผ่านมา พร้อมกับทบทวนความรู้เดิมจนเห็นว่าผู้ตรวจทุกคนเข้าใจ

2.3.4 ก่อนตรวจขอให้ผู้ตรวจหยิบเอกสารผลการเรียนของนักเรียนแต่ละคน ซึ่งผู้วิจัยกล่าวว่าได้ถ่ายสำเนามาจากต้นฉบับตัวจริง และตรวจดูผลการเรียนของนักเรียนกลุ่มนี้ทีละคน แล้วประเมินผลการเรียนจากผลนี้ โดยการเขียนข้อความลงในช่องหน้าชื่อนักเรียนในกระดาษคำตอบว่า เรียนดี หรือ เรียนอ่อน เพียงอย่างเดียวหนึ่งเท่านั้น

2.3.5 หลังจากประเมินนักเรียนแต่ละคนเสร็จแล้ว ขอให้ผู้ตรวจลงมือตรวจกระดาษคำตอบได้

2.3.6 หลังจากที่ครูตรวจคำตอบเสร็จซึ่งใช้เวลาน้อยกว่าการตรวจครั้งที่ 2 เล็กน้อย ผู้วิจัยกล่าวขอบคุณผู้ตรวจที่ให้ความร่วมมืออย่างดียิ่ง

การวิเคราะห์ข้อมูล

1. ประมาณค่าความแปรปรวนของตัวประกอบที่ศึกษาในชั้นทฤษฎีการวิเคราะห์ความแปรปรวนของ Cardinet, et al. ตามแบบแผนการทดลอง ที่เรียกว่า 3 Factorial Design แบบ crossed Design โดยวิเคราะห์ตามตัวแบบสุ่ม (random model)

2. ประมาณค่าความแปรปรวนตามตัวแบบผสม (mixed model) โดยวิเคราะห์ให้สอดคล้องกับรูปแบบการวัด (measurement design) ที่กำหนดพาเซ็คคองที่ 2 แบบดังนี้

2.1 รูปแบบการวัดที่วัดความสามารถของนักเรียน ในการแสดงความคิดเห็นเกี่ยวกับข่าวหรือเหตุการณ์ อ้างอิงไปยังเอกภพของข้อสอบ และผู้ตรวจ เมื่อข้อสอบเป็นเอกภพที่มีขนาดไม่จำกัด แต่เอกภพของผู้ตรวจเป็นเอกภพที่มีขนาดจำกัด แทนด้วยสัญลักษณ์ $M(p/-/-/I, R^*)$

2.2 รูปแบบการวัดที่วัดความสามารถของนักเรียน ในการแสดงความคิดเห็นเกี่ยวกับข่าวหรือเหตุการณ์ อ้างอิงไปยังเอกภพผู้ตรวจเพียงอย่างเดียว โดยที่เอกภพผู้ตรวจมีขนาดจำกัด และข้อสอบเป็นพาเซ็คคองที่ เป็นพาเซ็คคองที่ แทนด้วยสัญลักษณ์ $M(p/-/I/R^*)$

สัญลักษณ์ของรูปแบบข้างต้นนี้ใช้ตามแบบของ Cardinet, et al. (1983) ซึ่งใช้สัญลักษณ์ในรูปทั่วไปว่า $M(D^R/D^F/I^F/I^R)$

เมื่อ M คือ รูปแบบการวัด (Measurement Design)

D^R คือ ตัวประกอบของสิ่งที่ถูกวัดที่เป็นตัวประกอบสุ่ม

D^F คือ ตัวประกอบของสิ่งที่ถูกวัดที่เป็นตัวประกอบคงที่

I^F คือ ตัวประกอบของเครื่องมือที่เป็นตัวประกอบคงที่ หรือเรียกอีกอย่าง

ว่าเป็นตัวประกอบควบคุม (Control Facet)

I^R คือ ตัวประกอบของเครื่องมือที่เป็นตัวประกอบสุ่ม เป็นตัวประกอบที่

ต้องการอ้างอิงสรุปผล

3. ประมาณค่าความแปรปรวนในชั้นการสรุปอ้างอิง (D study variance)

4. ประมาณค่าสัมประสิทธิ์การอ้างอิงสรุปของรูปแบบการวัดแต่ละแบบภายใต้การทดลองแต่ละสถานการณ์

5. เปรียบเทียบสัมประสิทธิ์การอ้างอิงสรุปของรูปแบบการวัดแต่ละแบบภายใต้การทดลองแต่ละสถานการณ์ ตามสมมุติฐานของการวิจัย

6. เพิ่มจำนวนข้อสอบและผู้ตรวจ ในการศึกษาชั้นการอ้างอิงสรุป แล้วประมาณค่าความแม่นยำในการวัด เลือกจำนวนข้อสอบและผู้ตรวจที่ให้ค่าสัมประสิทธิ์การอ้างอิงสรุป 0.50 เป็นอย่างต่ำ

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. การประมาณค่าความแปรปรวนของตัวประกอบที่ศึกษาแบบ 3 Factorial Design ใช้สูตรในการคำนวณตามลำดับดังนี้

1.1 คำนวณผลบวกของกำลังสอง (SS) โดยใช้สูตรจากการวิเคราะห์ความแปรปรวนแบบแฟคตอเรียลสามตัวแปร CRF-pqr (อุทุมพร ทองอุไทย 2523:311-314) เมื่อกำหนดให้จำนวนชั้นหรือจำนวนข้อมูลในแต่ละ Cell เป็น 1 ($n=1$)

หลักการและสูตรในการคำนวณ เป็นดังนี้

ขั้นที่ 1 เขียนตารางสรุป ABCS, ABC, AB, AC, BC เพื่อแจกแจงหรือรวมคะแนนตาม Combination ของตัวแปร

ขั้นที่ 2 คำนวณค่าต่าง ๆ ดังนี้

$$\begin{aligned}
 (1) \quad & N \sum_{1} ABCS \\
 (2) \quad & N \sum_{1} (ABCS)^2 - [ABCS] \\
 (3) \quad & \frac{\left(\sum_{1}^{N} ABCS \right)^2}{npqr} - [X] \\
 (4) \quad & \sum_{1} \frac{p \left(\sum_{1}^q A \right)^2}{npr} - [A]
 \end{aligned}$$

$$(5) \quad \sum_1^q \frac{\binom{p}{\sum B}^2}{npr} = [B]$$

$$(6) \quad \sum_1^r \frac{\binom{p}{\sum C}^2}{npq} = [C]$$

$$(7) \quad \sum_1^p \sum_1^q \frac{(AB)^2}{nr} = [AB]$$

$$(8) \quad \sum_1^p \sum_1^r \frac{(AC)^2}{nq} = [AC]$$

$$(9) \quad \sum_1^q \sum_1^r \frac{(BC)^2}{np} = [BC]$$

$$(10) \quad \sum_1^p \sum_1^q \sum_1^r \frac{(ABC)^2}{n} = [ABC]$$

ขั้นที่ 3 แทนค่าในสูตรผลบวกของกำลังสอง (SS) แต่ละค่าดังนี้

$$SS_{total} = [ABCS] - [X]$$

$$SS_A = [A] - [X]$$

$$SS_B = [B] - [X]$$

$$SS_C = [C] - [X]$$

$$SS_{AB} = [AB] - [A] - [B] + [X]$$

$$SS_{AC} = [AC] - [A] - [C] + [X]$$

$$SS_{BC} = [BC] - [B] - [C] + [X]$$

$$SS_{ABC} = [ABC] - [AB] - [AC] - [BC] + [A] \\ + [B] + [C] - [X]$$

$$SS_{w.cell} = [ABCS] - [ABC]$$

1.2 คำนวณค่าเฉลี่ยกำลังสอง (MS) โดยสูตร (อุทุมพร ทองอุไทย 2523:

120)

$$MS = \frac{SS(\alpha)}{df(\alpha)}$$

df คือ ชั้นแห่งความเป็นอิสระที่สอดคล้องกับผลของตัวประกอบ

1.3 การประมาณค่าความแปรปรวนสำหรับแบบจำลองลุ่ม

คำนวณความแปรปรวน เมื่อตัวประกอบทุกตัวเป็นตัวประกอบลุ่มโดยใช้สูตร

(Tourneur cited by Cardinet, et al. 1983: 25) ดังนี้

$$\hat{\sigma}^2(\alpha) = \frac{1}{f(\alpha)} \left[MS(\alpha) + \sum_{i=1}^{i=j} (-1)^i \left\{ \begin{array}{l} \text{ผลรวมของ MS ที่ตัวกำกับรวมมี } \alpha \text{ ร่วมอยู่} \\ \text{บวกกับพหุเชิงเส้น } i \text{ ที่เพิ่มเข้ามาโดยที่พหุเชิงเส้น } i \\ \text{จะต้องปรากฏอยู่ในชั้น } (i-1) \text{ สำหรับ } i > 1 \end{array} \right\} \right]$$

เมื่อ α คือ ผลขององค์ประกอบใด ๆ

$f(\alpha)$ คือ ผลคูณของขนาดกลุ่มตัวอย่างที่ไม่ปรากฏอยู่ใน α

j คือ จำนวนนิพจน์ (Expression) ที่นำมารวมกันภายใน

เครื่องหมายแสดงผลรวม

i คือ ลำดับที่ (rank) ของนิพจน์จาก 1 ถึง j ภายในวงเล็บ

ปีกกา

จากความสัมพันธ์ของตัวประกอบ $pxixr$ จะหาค่าคอบแปรปรวนได้จาก

สูตรต่อไปนี้

$$\hat{\sigma}^2(pir) = MS(pir)$$

$$\hat{\sigma}^2(pi) = \frac{1}{nr} [MS(pi) - MS(pir)]$$

$$\hat{\sigma}^2(pr) = \frac{1}{ni} [MS(pr) - MS(pir)]$$

$$\hat{\sigma}^2(ir) = \frac{1}{np} [MS(ir) - MS(pir)]$$

$$\hat{\sigma}^2(r) = \frac{1}{npni} [MS(ir) - MS(pr) + MS(pir)]$$

$$\hat{\sigma}^2(i) = \frac{1}{n_p n_r} [MS(i) - MS(ir) - MS(pi) + MS(pir)]$$

$$\hat{\sigma}^2(p) = \frac{1}{n_i n_r} [MS(p) - MS(pi) - MS(pr) + MS(pir)]$$

1.4 การประมาณค่าความแปรปรวนสำหรับแบบจำลองผสม เมื่อกำหนดให้ตัวประกอบบางตัวเป็นตัวประกอบคงที่ แต่ตัวประกอบที่เหลือเป็นตัวประกอบสุ่มใช้สูตร คำนวณโดยอาศัยความแปรปรวนในแบบจำลองสุ่มในหัวข้อ 1.3 โดยใช้สูตร (Brennan 1983: 51) ดังนี้

$$\hat{\sigma}^2(\alpha|M) = [\hat{\sigma}^2(\beta)/F(\beta|\alpha)]$$

เมื่อ $\hat{\sigma}^2(\alpha|M)$ คือ ค่าประมาณความแปรปรวนของ α ในแบบ

จำลองผสม

β คือ ค่าองค์ประกอบใด ๆ ที่อย่างน้อยจะต้องมีตัวกำกับ

ซึ่งแทนผลของ α อยู่ด้วย

และ $F(\beta|\alpha) =$ ผลคูณของขนาดของเอกภพ (N) ของตัวกำกับทุกตัว
 ใน β แต่ไม่อยู่ใน
 แต่ถ้า $\alpha = \beta, F(\beta|\alpha) = 1$

2. ประมาณค่าความแปรปรวนในชั้นการสุ่มอ้างอิง

2.1 หาค่าคาดหวังของความแปรปรวน (Expectancies of variance)

เป็นค่าปรับแก้ค่าความแปรปรวนในชั้น G study ที่มีพาเซตสุ่มจำกัด หรือ พาเซตคงที่อยู่ในตัวกำกับแรก วิธีปรับใช้สูตร (Cardinet, et al. 1983: 34-35)

$$E^2(\alpha) = \frac{(N_f - 1)\hat{\sigma}^2(\alpha)}{N_f}$$

เมื่อ $E^2(\alpha)$ คือ ค่าคาดหวังของความแปรปรวนของ α เมื่อ α มีตัวกำกับแรกของพาเซตคงที่ หรือพาเซตสุ่มจำกัดอยู่ด้วย

N_f คือ ขนาดของเอกภพของพาเซตคงที่หรือสุ่มจำกัด

2.2 ประมาณค่าความแปรปรวนของคะแนนเอกภพ $\hat{\sigma}^2(\mathcal{T})$ ใช้หลักการดังนี้
(Cardinet and Allal 1983: 35-36)

2.2.1 หาค่าองค์ประกอบของ active variance ซึ่งมีตัวกำกับของ
สิ่งที่ถูกวัดอยู่ในตัวกำกับแรกๆ ทุก ๆ ค่า

2.2.2 หาค่าผลบวกขององค์ประกอบตามข้อ 1 เมื่อ Active variance
คือ ความแปรปรวนที่มีผลต่อการสรุปอ้างอิง ได้แก่ ความแปรปรวนของตัวประกอบทุกตัวยกเว้น
ความแปรปรวนที่มีพาเซ็คเครื่องมือที่เป็นพาเซ็คคงที่ (I^F) บนอยู่ในตัวกำกับ ค่าความแปรปรวน
ที่ไม่มีผลต่อการอ้างอิงสรุปนี้เรียกว่า passive variance

2.3 คำนวณความแปรปรวนของความคลาดเคลื่อนสัมบูรณ์ $\hat{\sigma}^2(\Delta)$ โดยดำเนินการ
การดังนี้

2.3.1 หาค่าประมาณความแปรปรวนของพาเซ็คต่าง ๆ โดยสูตร ดังนี้
(Brennan 1983: 57)

$$\hat{\sigma}^2(\bar{\alpha}) = \hat{\sigma}^2(\alpha) / d(\bar{\alpha})$$

เมื่อ $d(\bar{\alpha}) = 1$ ถ้า $\alpha = p$ ซึ่งเป็นพาเซ็คของสิ่งที่ถูก
วัดแต่มีค่าเท่ากับผลคูณของขนาดตัวอย่างใน D study ที่ อยู่ใน α ยกเว้น p

2.3.2 ปรับแก้ความแปรปรวนของพาเซ็คกลุ่มที่มีขนาดจำกัดโดยการคูณ
 $(N-n)/(N-1)$ ซึ่งเรียกว่า finite population correction (Cardinet and Allal
1983: 37-38)

2.3.3 จากค่าประมาณความแปรปรวนของพาเซ็คใน 2.3.1 หักค่า
ความแปรปรวนของสิ่งที่ถูกวัดออกไป แล้วหาผลบวกของความแปรปรวนที่เหลือ โดยการคูณด้วย
 $(N-n)/(N-1)$ เข้ากับความแปรปรวนของพาเซ็คเครื่องมือที่มีขนาดจำกัด

3. การประมาณค่าสัมประสิทธิ์การสรบอ้างอิง ใช้สูตร (Cardinet and Allal
1983: 40)

$$E\hat{\rho}^2(\Delta) = \frac{\hat{\sigma}^2(\mathcal{T})}{\hat{\sigma}^2(\mathcal{T}) + \hat{\sigma}^2(\Delta)}$$

เมื่อ $E\hat{\rho}^2(\Delta)$ คือค่าสัมประสิทธิ์การอ้างอิงสรบ เมื่อตัดสินใจแบบสัมบูรณ์

$\sigma^2(T)$ คือความแปรปรวนของคะแนนเอกภพ

$\sigma^2(\Delta)$ คือความแปรปรวนของความคลาดเคลื่อนแบบสัมบูรณ์

4. เปรียบเทียบสัมประสิทธิ์การอ้างอิงสรุป

4.1 ทดสอบความแตกต่างของค่าสัมประสิทธิ์การอ้างอิงสรุปมากกว่า 2 ค่าที่ได้จากการทดสอบกลุ่มตัวอย่างกลุ่มเดียว ใช้สูตร (Woodruff and Feldt 1986: 393-413)

$$UX1 = \frac{\sum_{i=1}^n (u_i - \bar{u})^2}{S_u^2 - c_u}$$

เมื่อ $u_i = \left[\frac{1}{1 - a_i} \right]^{1/3}$

$$\bar{u} = \frac{\sum_{i=1}^n u_i}{n}$$

$$S_u^2 = \frac{2}{9n(N_c - 1)} \sum_{i=1}^n u_i^2$$

$$c_u = \frac{4}{9n(n-1)(N_c - 1)} \sum_{i=2}^n \sum_{j=1}^{i-1} [r(X_i, X_j)]^2 u_i u_j$$

$$N_c = N \left(\frac{n-1}{n+1} \right)$$

เมื่อ UX 1 คือ สถิติทดสอบเพื่อการแจกแจงคล้ายกับ χ^2 ที่ $df = n-1$

N คือ จำนวนนักเรียนที่ทดสอบ

n คือ จำนวนค่าสังเกต หรือจำนวนข้อสอบ

a_i คือ ค่าสัมประสิทธิ์การอ้างอิงสรุป หรือที่ Cronbach, et

al. ใช้ว่า $E\rho^2$

S_u^2 คือ ความแปรปรวนของคะแนน u_i

\bar{u} คือ ค่าเฉลี่ยของคะแนน u_i

c_u คือ ค่าปรับแก้ความแปรปรวนของ u_i ในกรณีที่ค่า

สัมประสิทธิ์การอ้างอิงสรุปคำนวณมาจากกลุ่มตัวอย่างกลุ่มเดียว

$r(X_i, X_j)$ คือ ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบสอบ i

และ j เมื่อใช้คะแนนรวมเป็นหน่วยการวิเคราะห์

4.2 การทดสอบภายหลังเมื่อผลการทดสอบ UX1 มีนัยสำคัญทางสถิติ ใช้สูตรวิธีการ Large Sample Multiple Comparison (Marascuilo 1966: 280-290) ซึ่งเป็นวิธีการตรงกับวิธีทำ pos hoc ของ Scheffe มีหลักการทั่วไปว่า เมื่อผลการทดสอบ UX1 มีนัยสำคัญทางสถิติ แสดงว่า มี contrast อยู่อย่างน้อย 1 contrast ที่ช่วงความเชื่อมั่น คลุม 0 ในการทดสอบเท่าดังนี้

ด้วยความน่าจะเป็น $1-\alpha$ ที่ค่าของ จะตกอยู่ในช่วง

$$\psi \pm \sqrt{X_{k-1}^2 (1-\alpha) \text{Var}(\hat{\psi})}$$

$$\text{เมื่อ } \psi = a_1\phi_1 + a_2\phi_2 + \dots + a_k\phi_k$$

ϕ คือค่า parameter ที่ต้องการทดสอบ ในที่นี้คือค่าสัมประสิทธิ์

การอ้างอิงสรุปที่แปลงอยู่ในรูปคะแนนมาตรฐาน Woodruff and Feldt ใช้สัญลักษณ์ว่า ϕ และค่าประมาณใช้ u_i

$$\text{Var}(\hat{\psi}) = a^2 \text{Var}(\phi) \text{ ซึ่งก็คือ}$$

$$\text{Var}(\hat{\psi}) = a^2 \text{Var}(u_i)$$

$$\text{Var}(u_i) = \frac{[2]}{9(Nc-1)} u_i^2$$

a คือค่าสัมประสิทธิ์ของการเปรียบเทียบที่ $a_1 + a_2 + \dots + a_k = 0$