

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบชนิดพหุมิติ ด้วยวิธีแมนเทิล-แฮนส์เชลกับวิธีถดถอยโลจิสติก เมื่อใช้เกณฑ์จับคู่เปรียบเทียบแตกต่างกัน ผู้วิจัยได้ดำเนินการศึกษาตามลำดับดังต่อไปนี้

1. กำหนดประชากรและกลุ่มตัวอย่าง
2. สร้างเครื่องมือที่ใช้ในการวิจัย
3. เก็บรวบรวมข้อมูล
4. กำหนดขั้นตอนการวิเคราะห์ข้อมูล
5. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล-แฮนส์เชลและวิธีซิบเทสท์
6. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีโลจิสติก
7. คำนวณอัตราความถูกต้องของการตรวจสอบและความคลาดเคลื่อน
8. ทดสอบอัตราความถูกต้องของการตรวจสอบและความคลาดเคลื่อนด้วยสถิติ Z

1. กำหนดประชากรและกลุ่มตัวอย่าง

ประชากร

ประชากรที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2540 ภาคเรียนที่ 2 ในโรงเรียนระดับประถมศึกษา สังกัดสำนักงานการประถมศึกษากรุงเทพมหานคร จำนวน 38 โรงเรียน ซึ่งแบ่งออกเป็น 6 กลุ่มโรงเรียน มีจำนวนทั้งสิ้น 6,707 คน (ข้อมูลและสารสนเทศ ปีการศึกษา 2540 สำนักงานการประถมศึกษากรุงเทพมหานคร)

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการวิจัย ได้จากการกำหนดขนาดกลุ่มตัวอย่างในตาราง Yamane ที่ความเชื่อมั่น 97% ยอมให้เกิดความคลาดเคลื่อน $\pm 3\%$ ขนาดประชากร 6,707 คน ได้จำนวนกลุ่มตัวอย่าง 953 คน (Yamane, T., 1973: 1088) แต่ในการเก็บรวบรวมข้อมูลจริง ผู้วิจัยใช้

กลุ่มตัวอย่างทั้งสิ้นจำนวน 1,076 คน เป็นนักเรียนหญิง 483 คน และนักเรียนชาย 593 คน คิดเป็นร้อยละ 44.9 และ 55.1 ตามลำดับ โดยมีรายละเอียดดังนี้

1. กลุ่มตัวอย่างเพื่อการทดลองใช้(Try-out) จำนวน 2 ครั้ง ผู้วิจัยใช้วิธีเลือกโรงเรียนโดยใช้กลุ่มตัวอย่างในการทดลองใช้เครื่องมือวิจัยจำนวน 623 คน ดังรายละเอียดใน ตารางที่ 4

ตารางที่ 4 กลุ่มตัวอย่างในการทดลองใช้เครื่องมือวิจัย

ครั้งที่ทดลองใช้	รายชื่อโรงเรียน	จำนวนนักเรียน(คน)
1	โรงเรียนบ้านหนองจิม (สิงห์จันทร์บำรุง)	60
2	โรงเรียนวัดโบสถ์	118
	โรงเรียนวัดชนะสงคราม	70
	โรงเรียนราชวินิต	208
	โรงเรียนวัดโสมนัส	96
	โรงเรียนวัดใหม่ช่องลม	71
	รวม	623

2. กลุ่มตัวอย่างที่ใช้ในการเก็บข้อมูลจริงได้จากการสุ่มตัวอย่างโรงเรียนจากกลุ่มโรงเรียนระดับประถมศึกษา สังกัดสำนักงานการประถมศึกษากรุงเทพมหานครกลุ่มละ 1 โรงเรียน ได้กลุ่มตัวอย่างจำนวน 1,076 คน ดังรายละเอียดในตารางที่ 5

ตารางที่ 5 กลุ่มตัวอย่างที่ใช้ในการเก็บข้อมูลจริง

รายชื่อโรงเรียน	กลุ่มโรงเรียน	จำนวนนักเรียน (คน)
โรงเรียนราชวินิตประถมบางแค	กรุงธน	189
โรงเรียนดาราคาม	รัตนโกสินทร์	194
โรงเรียนวัดพลับพลารชัย	สุโขทัย	283
โรงเรียนอนุบาลสามเสนฯ	อโยธยา	200
โรงเรียนอนุบาลวัดปรีณายก	ทวารวดี	210
	รวม	1,076

2. สร้างเครื่องมือที่ใช้ในการวิจัย

ลักษณะของแบบทดสอบ

แบบทดสอบความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ชุดนี้ เป็นข้อสอบแบบเลือกตอบชนิด 4 ตัวเลือก ประกอบด้วยแบบสอบย่อยจำนวน 2 ฉบับ ๆ ละ 40 ข้อ ซึ่งแต่ละฉบับวัดความสามารถเกี่ยวกับองค์ประกอบ 4 ด้าน ดังต่อไปนี้

1. ด้านความรู้ความเข้าใจทางคณิตศาสตร์ เป็นองค์ประกอบที่วัดความสามารถเกี่ยวกับความคิดรวบยอดทางคณิตศาสตร์ ซึ่งเป็นคำถามประเภทความเข้าใจในหลักการพื้นฐานในวิชาคณิตศาสตร์

2. ด้านความสามารถด้านการคิดคำนวณ เป็นองค์ประกอบที่ใช้ความเข้าใจและความสามารถในการคิดคำนวณ โดยใช้ความรู้พื้นฐานทางคณิตศาสตร์ ได้แก่ การบวก การลบ การคูณ การหาร และการประยุกต์

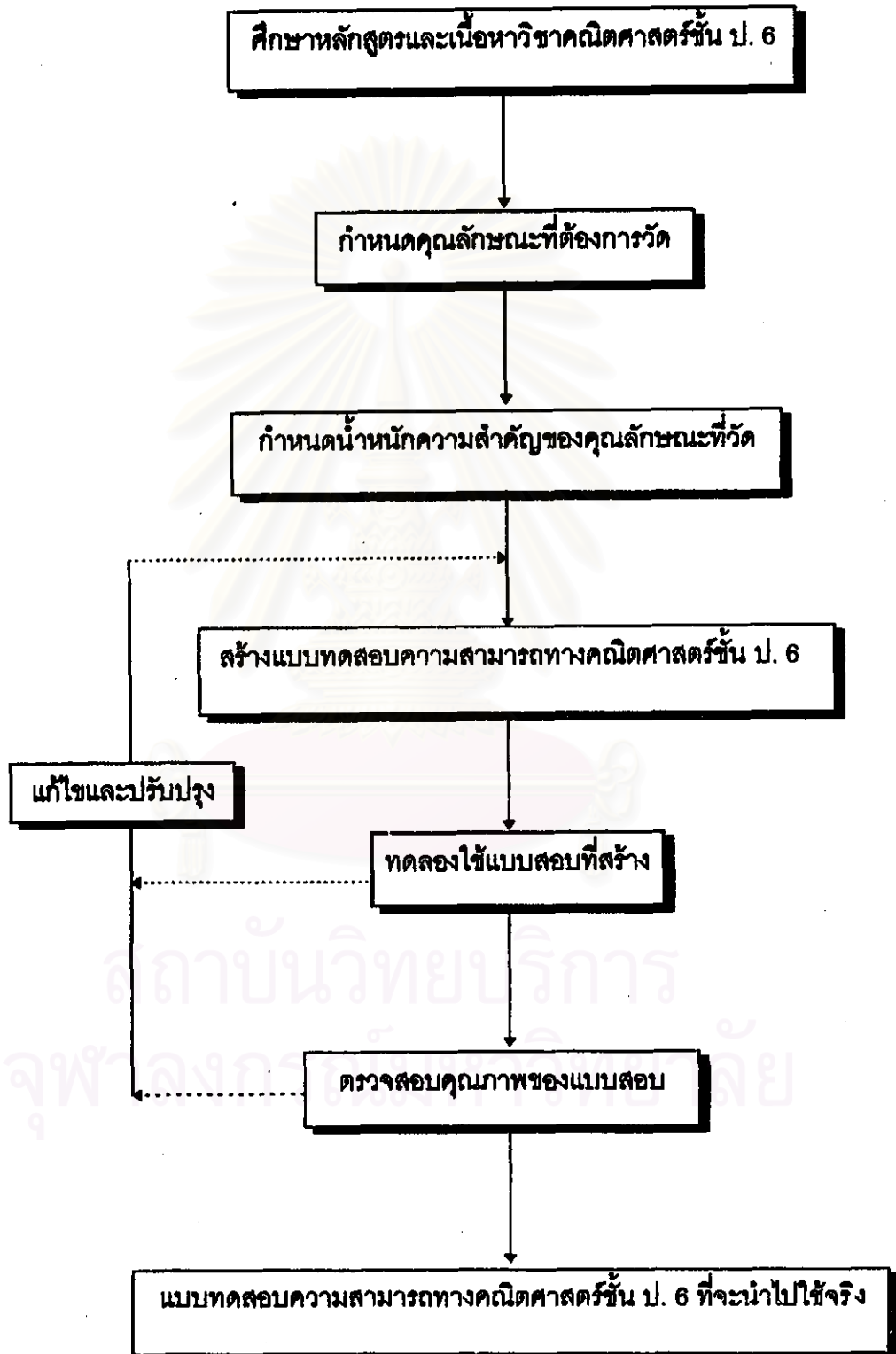
3. ด้านความสามารถเกี่ยวกับการพิจารณาผลลัพธ์อย่างสมเหตุสมผล เป็นองค์ประกอบที่ใช้ความรู้ความเข้าใจเกี่ยวกับทฤษฎีและหลักการต่าง ๆ ทางคณิตศาสตร์เพื่อใช้ในการพิจารณาคำตอบที่ถูกต้องที่สุด

4. ด้านความสามารถด้านการแก้โจทย์ปัญหา เป็นองค์ประกอบที่วัดความสามารถในการนำความรู้ในวิชาคณิตศาสตร์ไปประยุกต์ใช้เพื่อแก้ปัญหาในสถานการณ์ที่กำหนดให้

ในการวิจัยครั้งนี้ ผู้วิจัยได้ดำเนินการสร้างและพัฒนาเครื่องมือที่ใช้ในการวิจัย ซึ่งเป็นแบบสอบที่วัดความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 โดยมีขั้นตอนตามแผนภาพที่ 3

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

แผนภาพที่ 3 ขั้นตอนการสร้างเครื่องมือในการวิจัย



ขั้นตอนการสร้างเครื่องมือวิจัย

1. ทำการศึกษาเนื้อหาเกี่ยวกับวิชาคณิตศาสตร์ในระดับชั้นประถมศึกษาปีที่ 6 จากหลักสูตร คู่มือครู หนังสือแบบเรียนวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 และเอกสารที่เกี่ยวข้อง ซึ่งรายวิชานี้มีเนื้อหาทั้งหมด 8 เรื่อง นำเนื้อหาที่ต้องการวัดไปสร้างเป็นข้อกระทงได้ 104 ข้อ ซึ่งแบ่งเป็นแบบสอบย่อย จำนวน 4 ชุด ๆ ละ 26 ข้อ สรุปเป็นตารางโครงสร้างได้ดังตารางที่ 6

ตารางที่ 6 ตารางโครงสร้างของแบบทดสอบความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6

แบบสอบย่อย ฉบับที่	องค์ประกอบ ที่ต้องการวัด	เนื้อหาที่มุ่งวัด	จำนวน ข้อ	หมายเลขข้อ
1	ความรู้ความเข้าใจ ทางคณิตศาสตร์	1.1 จำนวนและตัวเลข	3	1, 2, 3
		1.2 ตัวประกอบของจำนวนนับ	4	4, 5, 6, 7,
		1.3 สมการและการแทนค่า	3	8, 9, 10
		1.4 ทศนิยม	3	11, 12, 13
		1.5 เศษส่วน	3	14, 15, 16
		1.6 เส้นและมุม	4	17, 18, 19, 20
		1.7 พื้นที่และปริมาตร	3	21, 22, 23
		1.8 ร้อยละ	3	24, 25, 26
2	การคิดคำนวณ	2.1 การบวก ลบ คูณ หาร จำนวนนับ	4	1, 2, 3, 4
		2.2 ตัวประกอบของจำนวนนับ	3	5, 6, 7
		2.3 สมการและการแทนค่า	4	8, 9, 10, 11
		2.4 การบวก ลบ คูณ หาร เศษส่วน	5	12, 13, 14, 15, 16
		2.5 การบวก ลบ คูณ หาร ทศนิยม	6	17, 18, 19, 20, 21, 22
		2.6 พื้นที่และปริมาตร	4	23, 24, 25, 26

ตารางที่ 6 (ต่อ)

แบบสอบ ฉบับที่	องค์ประกอบ ที่ต้องการวัด	เนื้อหาที่มุ่งวัด	จำนวน ข้อ	หมายเลขข้อ
3.	การพิจารณาผลลัพธ์ ที่สมเหตุสมผล	3.1 จำนวนและตัวเลข	9	1, 2, 3, 4, 5, 6, 7, 8, 9,
		3.2 สมการและการแทนค่า	4	10, 11, 12, 13
		3.3 ตัวประกอบของจำนวนนับ	4	14, 15, 16, 17
		3.4 เศษส่วน	4	18, 19, 20, 21
		3.5 เส้นและมุม	4	22, 23, 24, 25
		3.6 พื้นฐานและปริมาตร	1	26
4.	การแก้โจทย์ปัญหา	4.1 การบวก ลบ คูณ หารของ จำนวนนับ	8	1, 2, 3, 4, 5, 6, 7, 8
		4.2 สมการและการแทนค่า	2	9, 10
		4.3 การบวก ลบ คูณ หาร ของเศษส่วน	2	11, 12
		4.4 การบวก ลบ คูณ หาร ของทศนิยม	2	13, 14
		4.5 พื้นที่และปริมาตร	6	15, 16, 17, 18, 19, 20
		4.6 ร้อยละ	6	21, 22, 23, 24, 25, 26

2. นำแบบสอบที่จัดพิมพ์เรียบร้อยแล้วไปทดลองใช้ (Try-out) กับกลุ่มตัวอย่างซึ่งเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนบ้านหนองจิม(สิงห์จันทร์บำรุง) จำนวน 60 คน นำผลมาวิเคราะห์คุณภาพของแบบสอบตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) โดยใช้โปรแกรมสำเร็จรูป IRT (BAY) 1.0 ทำการตรวจสอบหาค่าความยาก (p) และค่าอำนาจจำแนก (r) รายข้อ พิจารณาคัดเลือกข้อที่มีค่าความยากระหว่าง 0.2 - 0.8 และมีค่าอำนาจจำแนกตั้งแต่ 0.2 ขึ้นไป รวมทั้งพิจารณาถึงเวลาที่เหมาะสมในการทำแบบสอบ

3. นำแบบสอบที่ปรับปรุงแล้วไปทดลองใช้ (Try - out) ครั้งที่สอง กับกลุ่มตัวอย่างซึ่งเป็นนักเรียนชั้น ประถมศึกษาปีที่ 6 ในโรงเรียนสังกัดสำนักงานการประถมศึกษากรุงเทพมหานคร ได้แก่ โรงเรียนวัดโบสถ์ โรงเรียนวัดชนะสงคราม โรงเรียนราชวินิต โรงเรียนวัดโสมนัส และโรงเรียนวัดใหม่ช่องลม จำนวน 583 คน นำผลมาวิเคราะห์คุณภาพข้อสอบโดยใช้โปรแกรมสำเร็จรูป IRT (BAY) 1.0 ตรวจสอบหาค่าความยาก(p) และค่าอำนาจจำแนก (r) รายข้อ โดยผลการวิเคราะห์แบบสอบทั้ง 4 ฉบับ ปรากฏผลดังตารางที่ 7

ตารางที่ 7 ค่าความยาก (p) และค่าอำนาจจำแนก (r) รายข้อของแบบสอบในการทดลองใช้

ข้อ	แบบสอบฉบับที่ 1		แบบสอบฉบับที่ 2		แบบสอบฉบับที่ 3		แบบสอบฉบับที่ 4	
	p	r	p	r	p	r	p	r
1.	0.759	0.385	0.748	0.511	0.558	0.500	0.317	0.287
2.	0.139	0.042	0.709	0.472	0.462	0.427	0.968	0.380
3.	0.370	0.403	0.660	0.579	0.490	0.401	0.537	0.679
4.	0.491	0.406	0.252	0.187	0.510	0.501	0.317	0.377
5.	0.760	0.522	0.252	0.459	0.490	0.320	0.512	0.423
6.	0.296	0.638	0.515	0.158	0.365	0.346	0.512	0.571
7.	0.333	0.284	0.350	0.104	0.231	0.186	0.455	0.642
8.	0.667	0.514	0.728	0.291	0.173	0.151	0.756	0.506
9.	0.426	0.512	0.515	0.538	0.490	0.337	0.455	0.505
10.	0.361	0.027	0.544	0.661	0.327	0.142	0.431	0.502
11.	0.315	0.355	0.291	0.402	0.529	0.485	0.439	0.185
12.	0.639	0.466	0.398	0.396	0.644	0.289	0.537	0.523
13.	0.657	0.351	0.262	0.412	0.452	0.327	0.585	0.657
14.	0.296	0.439	0.427	0.465	0.423	0.431	0.431	0.674
15.	0.435	0.302	0.398	0.229	0.356	0.155	0.650	0.574
16.	0.778	0.453	0.408	0.295	0.317	0.402	0.325	0.246
17.	0.509	0.507	0.641	0.359	0.442	0.169	0.179	0.320
18.	0.278	0.417	0.738	0.309	0.125	0.037	0.463	0.643
19.	0.417	0.511	0.427	0.461	0.414	0.262	0.472	0.521
20.	0.454	0.549	0.583	0.309	0.365	0.400	0.520	0.500

ตารางที่ 7 (ต่อ)

ข้อ	แบบสอบฉบับที่ 1		แบบสอบฉบับที่ 2		แบบสอบฉบับที่ 3		แบบสอบฉบับที่ 4	
	p	r	p	r	p	r	p	r
21.	0.167	0.276	0.544	0.334	0.750	0.527	0.455	0.588
22.	0.472	0.373	0.262	0.275	0.654	0.306	0.130	0.234
23.	0.204	0.054	0.330	0.327	0.289	0.023	0.382	0.612
24.	0.630	0.460	0.388	0.426	0.279	0.086	0.447	0.583
25.	0.352	0.234	0.369	0.386	0.308	0.299	0.252	0.446
26.	0.287	0.376	0.146	0.501	0.404	0.120	0.309	0.071
KR-20	0.712		0.703		0.529		0.816	

จากตารางที่ 7 ปรากฏว่าแบบสอบฉบับที่ 1 มีค่าความยาก(p) ตั้งแต่ 0.139 ถึง 0.778 และมีค่าอำนาจจำแนก(r) ตั้งแต่ 0.027 ถึง 0.638 แบบสอบฉบับที่ 2 มีค่าความยาก(p) ตั้งแต่ 0.146 ถึง 0.748 และมีค่าอำนาจจำแนก(r) ตั้งแต่ 0.104 ถึง 0.661 แบบสอบฉบับที่ 3 มีค่าความยาก(p) ตั้งแต่ 0.126 ถึง 0.750 และมีค่าอำนาจจำแนก(r) ตั้งแต่ 0.023 ถึง 0.527 และแบบสอบฉบับที่ 4 มีค่าความยาก(p) ตั้งแต่ 0.130 ถึง 0.968 และมีค่าอำนาจจำแนก(r) ตั้งแต่ 0.071 ถึง 0.679 เมื่อพิจารณาผลการวิเคราะห์คุณภาพแบบสอบทั้งฉบับ พบว่า แบบสอบฉบับที่ 1, 2, 3 และ 4 มีค่าความเที่ยง(KR-20) เท่ากับ 0.712, 0.703, 0.529 และ 0.816 ตามลำดับ ซึ่งจะเห็นว่าคุณภาพของข้อสอบบางข้อยังไม่ดีนักผู้วิจัยจึงได้แก้ไขปรับข้อคำถามและตัวเลือกใหม่ในแบบสอบทั้ง 4 ฉบับ และคัดเลือกข้อสอบที่มีคุณภาพมาสร้างเป็นแบบสอบที่จะนำไปใช้ในการเก็บข้อมูลจริงจำนวน 80 ข้อโดยตัดข้อสอบที่มีคุณภาพต่ำออกไปและใช้หลักในการเลือกข้อสอบผลสัมฤทธิ์โดยทั่วไปที่นิยมใช้ข้อสอบที่มีค่าความยาก(p) ตามสัดส่วน ดังนี้ ค่าความยาก ระหว่าง 0.20 - 0.40 (25%), 0.41 - 0.60 (60 %) และ 0.61 - 0.80 (25 %) ส่วนค่าอำนาจจำแนก(r)รายข้อควรมีค่ามากกว่า 0.2 ขึ้นไป (ศิริชัย กาญจนวาสี, 2535) เพื่อสะดวกแก่การเก็บรวบรวมข้อมูลและหลีกเลี่ยงความคลาดเคลื่อนอันจะเกิดจากความเมื่อยล้าของผู้สอบในการทำข้อสอบ ผู้วิจัยจึงได้แบ่งแบบสอบออกเป็น 2 ฉบับ คือ ฉบับ A และฉบับ B มีจำนวนข้อสอบฉบับละ 40 ข้อ ใช้เวลาสอบฉบับละ 60 นาที(1 คาบ) และแบบสอบทั้งสองฉบับยังแบ่งออกเป็น 4 ด้านย่อย ๆ โดยแบ่งตามองค์ประกอบในการวัด ได้แก่

- องค์ประกอบที่ 1 เป็นเนื้อหาที่วัดด้านความรู้ความเข้าใจทางคณิตศาสตร์
จำนวนข้อสอบฉบับละ 10 ข้อ (ตั้งแต่ข้อ 1 – 10)
- องค์ประกอบที่ 2 เป็นเนื้อหาที่วัดความสามารถด้านการคิดคำนวณ
จำนวนข้อสอบฉบับละ 10 ข้อ (ตั้งแต่ข้อ 11 – 20)
- องค์ประกอบที่ 3 เป็นเนื้อหาที่วัดความสามารถด้านการพิจารณาผลลัพธ์อย่าง
สมเหตุสมผล จำนวนข้อสอบฉบับละ 10 ข้อ (ตั้งแต่ข้อ 21 – 30)
- องค์ประกอบที่ 4 เป็นเนื้อหาที่วัดความสามารถด้านการแก้ปัญหา
จำนวนข้อสอบฉบับละ 10 ข้อ (ตั้งแต่ข้อ 31 – 40)

ซึ่งสามารถสรุปรายละเอียดเกี่ยวกับแบบสอบทั้งสองฉบับดังในตารางที่ 8

ตารางที่ 8 ค่าความยาก และค่าอำนาจจำแนกของแบบทดสอบความสามารถทางคณิตศาสตร์
ชั้นประถมศึกษาปีที่ 6 หลังจากปรับปรุงแก้ไข

องค์ประกอบในการวัด	ข้อ	แบบสอบฉบับ A		แบบสอบฉบับ B	
		p	r	p	r
1. ด้านความรู้ความเข้าใจ ทางคณิตศาสตร์(K1)	1.	0.759	0.385	0.370	0.403
	2.	0.491	0.406	0.759	0.522
	3.	0.296	0.638	0.333	0.284
	4.	0.426	0.512	0.667	0.514
	5.	0.667	0.351	0.639	0.466
	6.	0.315	0.355	0.435	0.302
	7.	0.296	0.439	0.509	0.507
	8.	0.778	0.453	0.278	0.417
	9.	0.417	0.511	0.454	0.549
	10.	0.472	0.373	0.630	0.460
2. ด้านการคิดคำนวณ (K2)	11.	0.748	0.511	0.660	0.579
	12.	0.709	0.472	0.252	0.459
	13.	0.544	0.661	0.728	0.291
	14.	0.291	0.402	0.515	0.538
	15.	0.398	0.396	0.427	0.465

ตารางที่ 8 (ต่อ)

องค์ประกอบในการวัด	ข้อ	แบบสอบฉบับ A		แบบสอบฉบับ B	
		p	r	p	r
2. ด้านการคิดคำนวณ(ต่อ)	16.	0.262	0.412	0.408	0.295
	17.	0.641	0.369	0.738	0.309
	18.	0.427	0.461	0.330	0.327
	19.	0.583	0.309	0.388	0.426
	20.	0.544	0.434	0.369	0.385
3. ด้านการพิจารณาผลลัพธ์ ที่สมเหตุสมผล(K3)	21.	0.490	0.320	0.558	0.500
	22.	0.231	0.286	0.462	0.427
	23.	0.365	0.346	0.490	0.401
	24.	0.490	0.337	0.510	0.504
	25.	0.644	0.289	0.327	0.242
	26.	0.452	0.327	0.529	0.485
	27.	0.317	0.402	0.423	0.431
	28.	0.442	0.209	0.414	0.262
	29.	0.760	0.527	0.365	0.400
	30.	0.654	0.306	0.308	0.298
4. ด้านการแก้โจทย์ปัญหา (K4)	31.	0.512	0.571	0.317	0.287
	32.	0.455	0.643	0.968	0.380
	33.	0.455	0.506	0.317	0.377
	34.	0.431	0.502	0.512	0.423
	35.	0.431	0.674	0.756	0.506
	36.	0.650	0.574	0.637	0.523
	37.	0.325	0.246	0.685	0.657
	38.	0.463	0.643	0.472	0.521
	39.	0.447	0.583	0.520	0.500
	40.	0.252	0.446	0.455	0.586

เมื่อพิจารณาจากตารางที่ 8 พบว่า แบบสอบที่ได้คัดเลือกข้อสอบที่มีคุณภาพมาแล้วทั้งสองฉบับ มีค่าความยากและค่าอำนาจจำแนกอยู่ในช่วงดังต่อไปนี้ คือ แบบสอบฉบับ A มีค่าความยาก(p) ตั้งแต่ 0.231 ถึง 0.788 และมีค่าอำนาจจำแนก(η) ตั้งแต่ 0.209 ถึง 0.674 สำหรับแบบสอบฉบับ B มีค่าความยาก(p) ตั้งแต่ 0.252 ถึง 0.968 และมีค่าอำนาจจำแนก(η) ตั้งแต่ 0.242 ถึง 0.657 แสดงให้เห็นว่าข้อสอบแต่ละข้อมีคุณภาพเพียงพอที่จะนำไปใช้

ตารางที่ 9 ค่าสถิติของแบบสอบในองค์ประกอบที่ 1 (วัดความสามารถด้านความรู้ความเข้าใจทางคณิตศาสตร์)

ค่าสถิติ	MEAN	MIN	MEDIAN	MAX	S.D.	VAR
คะแนนผลการทดสอบ	9.98	3.00	11.50	20.00	4.01	16.11
ค่าความยาก	0.50	0.28	0.58	0.78	0.53	0.28
ดัชนีความยากมาตรฐาน	13.00	9.90	12.65	15.40	1.81	3.27
ค่าอำนาจจำแนก	0.50	0.35	0.52	0.69	0.12	0.01
ค่าความเที่ยงของแบบสอบ (KR-20)			0.761			
ค่าความคลาดเคลื่อนในการวัด			1.964			

จากตารางที่ 9 พบว่าแบบทดสอบความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่มีองค์ประกอบในการวัดด้านความรู้ความเข้าใจทางคณิตศาสตร์ มีค่าความยากและดัชนีความยากมาตรฐานโดยเฉลี่ย เท่ากับ 0.50 และ 13.00 ซึ่งอยู่ในเกณฑ์ที่เหมาะสม ค่าอำนาจจำแนกของแบบสอบโดยเฉลี่ยเท่ากับ 0.50 ซึ่งอยู่ในเกณฑ์ดี ส่วนผลการวิเคราะห์ค่าความเที่ยงของแบบสอบเพื่อหาความสอดคล้องภายใน (internal consistency) ของแบบสอบ ซึ่งคำนวณแบบ KR-20 มีค่าเท่ากับ 0.761 และมีความคลาดเคลื่อนในการวัดเท่ากับ 1.964 แสดงให้เห็นว่า แบบสอบซึ่งวัดองค์ประกอบในด้านนี้มีคุณภาพอยู่ในเกณฑ์เหมาะสม สามารถนำไปใช้ในการวิจัยได้

ตารางที่ 10 ค่าสถิติของแบบสอบในองค์ประกอบที่ 2 (วัดความสามารถด้านการคิดคำนวณ)

ค่าสถิติ	MEAN	MIN	MEDIAN	MAX	S.D.	VAR
คะแนนผลการทดสอบ	9.96	1.00	10.00	19.00	3.87	14.98
ค่าความยาก	0.50	0.25	0.50	0.75	0.52	0.27
ดัชนีความยากมาตรฐาน	13.01	10.29	13.00	0.70	1.78	3.17
ค่าอำนาจจำแนก	0.48	0.26	0.48	0.70	0.16	0.03
ค่าความเที่ยงของแบบสอบ (KR-20)			0.738			
ค่าความคลาดเคลื่อนในการวัด			1.982			

จากตารางที่ 10 พบว่าแบบทดสอบความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่มีองค์ประกอบในการวัดด้านการคิดคำนวณ มีค่าความยากและดัชนีความยากมาตรฐาน โดยเฉลี่ย เท่ากับ 0.50 และ 13.01 ซึ่งอยู่ในเกณฑ์พอเหมาะ ค่าอำนาจจำแนกของแบบสอบโดยเฉลี่ย เท่ากับ 0.48 ซึ่งอยู่ในเกณฑ์ดี ส่วนผลการวิเคราะห์ค่าความเที่ยงของแบบสอบ (KR-20) มีค่า เท่ากับ 0.738 และมีความคลาดเคลื่อนในการวัดเท่ากับ 1.982 แสดงให้เห็นว่าแบบสอบซึ่งวัด องค์ประกอบในด้านนี้มีคุณภาพอยู่ในเกณฑ์เหมาะสม สามารถนำไปใช้ในการวิจัยได้

ตารางที่ 11 ค่าสถิติของแบบสอบในองค์ประกอบที่ 3 (วัดความสามารถด้านการพิจารณาผลลัพธ์ที่สมเหตุสมผล)

ค่าสถิติ	MEAN	MIN	MEDIAN	MAX	S.D.	VAR
คะแนนผลการทดสอบ	9.22	2.00	9.50	17.00	3.38	11.44
ค่าความยาก	0.46	0.23	0.49	0.75	0.52	0.27
ดัชนีความยากมาตรฐาน	13.41	10.26	13.12	15.98	1.38	1.90
ค่าอำนาจจำแนก	0.41	0.14	0.36	0.57	0.15	0.02
ค่าความเที่ยงของแบบสอบ (KR-20)			0.624			
ค่าความคลาดเคลื่อนในการวัด			2.074			

จากตารางที่ 11 พบว่าแบบทดสอบความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่มีองค์ประกอบในการวัดด้านการพิจารณาผลลัพธ์ที่สมเหตุสมผล มีค่าความยากและดัชนีความ ยากมาตรฐาน โดยเฉลี่ยเท่ากับ 0.46 และ 13.41 ซึ่งอยู่ในเกณฑ์พอเหมาะ ค่าอำนาจจำแนกของ

แบบสอบโดยเฉลี่ยเท่ากับ 0.41 ซึ่งอยู่ในเกณฑ์ดี ส่วนผลการวิเคราะห์ค่าความเที่ยงของแบบสอบ (KR-20) มีค่าเท่ากับ 0.624 และมีความคลาดเคลื่อนในการวัดเท่ากับ 2.074 แสดงให้เห็นว่าแบบสอบซึ่งวัดองค์ประกอบในด้านนี้มีคุณภาพอยู่ในเกณฑ์เหมาะสม สามารถนำไปใช้ในการวิจัยได้

ตารางที่ 12 ค่าสถิติของแบบสอบในองค์ประกอบที่ 4 (วัดความสามารถด้านการแก้ปัญหา)

ค่าสถิติ	MEAN	MIN	MEDIAN	MAX	S.D.	VAR
คะแนนผลการทดสอบ	9.86	2.00	10.50	19.00	4.64	21.50
ค่าความยาก	0.49	0.25	0.61	0.97	0.52	0.27
ดัชนีความยากมาตรฐาน	12.95	5.68	10.65	15.71	2.14	4.59
ค่าอำนาจจำแนก	0.58	0.12	0.39	0.67	0.20	0.04
ค่าความเที่ยงของแบบสอบ (KR-20)			0.832			
ค่าความคลาดเคลื่อนในการวัด			1.900			

จากตารางที่ 12 พบว่าแบบทดสอบความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่มีองค์ประกอบในการวัดด้านความรู้ความเข้าใจทางคณิตศาสตร์ มีค่าความยากและดัชนีความยากมาตรฐาน โดยเฉลี่ย เท่ากับ 0.49 และ 12.95 ซึ่งอยู่ในเกณฑ์พอเหมาะ ค่าอำนาจจำแนกของแบบสอบโดยเฉลี่ยเท่ากับ 0.58 ซึ่งอยู่ในเกณฑ์ดี ส่วนผลการวิเคราะห์ค่าความเที่ยงของแบบสอบ (KR-20) มีค่าเท่ากับ 0.832 มีความคลาดเคลื่อนในการวัดเท่ากับ 1.900 แสดงให้เห็นว่าแบบสอบซึ่งวัดองค์ประกอบในด้านนี้มีคุณภาพอยู่ในเกณฑ์ดี สามารถนำไปใช้ในการวิจัยได้

3. เก็บรวบรวมข้อมูล

3.1 ขั้นตอนทดลองใช้ (Try-out) ผู้วิจัยดำเนินการดังนี้

ทดลองใช้ครั้งที่ 1 ใช้กลุ่มตัวอย่างซึ่งเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนบ้านหนองจิม(สิงห์จันทร์บำรุง) จำนวน 60 คน เพื่อหาข้อบกพร่องในด้านคำสั่ง คำชี้แจงความเข้าใจในข้อคำถาม เวลาในการทำแบบสอบ ตลอดจนความชัดเจนของข้อคำถามและตัวเลือกปรับปรุงแก้ไขข้อบกพร่องที่พบ

ทดลองใช้ครั้งที่ 2 ใช้กลุ่มตัวอย่าง ซึ่งเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนในสังกัดสำนักงานการประถมศึกษากรุงเทพมหานคร ได้แก่ โรงเรียนวัดโบสถ์ โรงเรียนราชวินิต

โรงเรียนวัดชนะสงคราม โรงเรียนวัดโสมนัส และโรงเรียนวัดใหม่ช่องลม จำนวน 663 คน เพื่อวิเคราะห์หาคุณภาพข้อสอบทั้งค่าความยาก(p) ค่าอำนาจจำแนก(q) ค่าความเที่ยงของแบบสอบ และคัดเลือกข้อสอบที่มีคุณภาพสำหรับนำไปใช้ในการเก็บข้อมูลจริง

3.2 การเก็บข้อมูลจริง

ผู้วิจัยทำหนังสือขอความร่วมมือในการวิจัยจากบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัยถึงผู้อำนวยการการประถมศึกษากรุงเทพมหานครให้ออกหนังสือถึงผู้บริหารโรงเรียนประถมศึกษาในสังกัดสำนักงานการประถมศึกษากรุงเทพมหานคร(สป.กทม.) เพื่อขอเก็บข้อมูลจากโรงเรียนที่เป็นกลุ่มตัวอย่าง หลังจากนั้นนำหนังสือจากสป.กทม.ไปติดต่อกับผู้บริหารโรงเรียนและผู้ช่วยฝ่ายวิชาการเพื่อชี้แจงวัตถุประสงค์ในการทำวิจัย นัดหมายวัน เวลาและสถานที่สอบ ตลอดจนติดต่ออาจารย์ที่สอนในคาบที่มีการทดสอบ เพื่อขอความร่วมมือพร้อมทั้งชี้แจงการดำเนินการสอบ

การเก็บรวบรวมข้อมูลโดยผู้วิจัยมีการดำเนินการ ดังนี้

3.2.1 จัดเตรียมแบบสอบที่ปรับปรุงแล้วทั้งสองฉบับ โดยในแต่ละชุดแบบสอบประกอบด้วยแบบสอบที่มีค่าชี้แจงในการสอบ กระดาษคำตอบ และกระดาษเปล่าสำหรับใช้ในการทดเลข ซึ่งแต่ละของแบบสอบมีการจัดแบบสอบเรียงตามลำดับเลขที่ 1, 2, 3, ไปจนครบตามจำนวนนักเรียน

3.2.2 นำแบบสอบไปใช้ทดสอบกับกลุ่มตัวอย่าง กลุ่มตัวอย่างแต่ละคนจะได้ทำแบบสอบทั้งสองฉบับ โดยจะแบ่งเวลาทดสอบออกเป็นช่วงเช้า 1 ฉบับและช่วงบ่ายอีก 1 ฉบับ

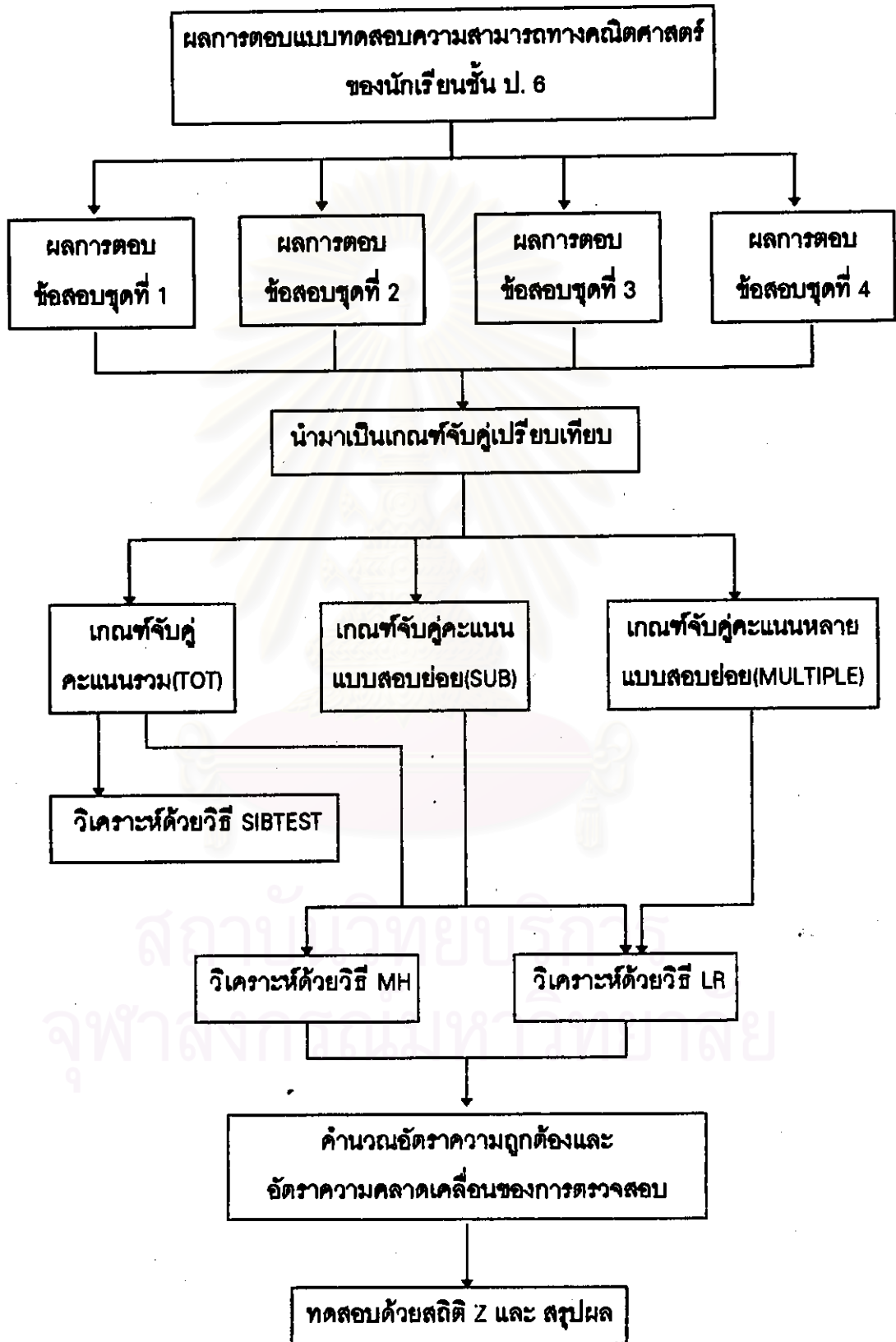
3.2.3 นำกระดาษคำตอบที่ได้จากการดำเนินการสอบมาตรวจสอบความเรียบร้อย จากนั้นเรียงกระดาษคำตอบทั้งหมดให้เป็นระบบ

3.2.4 นำผลการตอบข้อสอบทั้งสองฉบับของกลุ่มตัวอย่างแต่ละคนมารวมกัน ลงรหัส (coding) ผลการตอบข้อสอบของกลุ่มตัวอย่าง จากนั้นตรวจคะแนนและวิเคราะห์คุณภาพข้อสอบด้วยโปรแกรมวิเคราะห์ข้อสอบ

4. กำหนดขั้นตอนการวิเคราะห์ข้อมูล

ในการวิจัยครั้งนี้ ผู้วิจัยได้ดำเนินการวิเคราะห์ข้อมูล โดยมีขั้นตอนในการดำเนินการปรากฏดังแผนภาพที่ 4

แผนภาพที่ 4 ขั้นตอนการวิเคราะห์ข้อมูล



ผู้วิจัยดำเนินการวิเคราะห์ข้อมูลตามลำดับต่อไปนี้

4.1 จัดเตรียมแฟ้มข้อมูลสำหรับใช้ในการวิเคราะห์ โดยแยกผลการตอบข้อสอบตามกลุ่มผู้สอบ ได้แก่ กลุ่มเปรียบเทียบและกลุ่มอ้างอิง แยกตามองค์ประกอบในการวัด โดยแยกผลการตอบข้อสอบของกลุ่มตัวอย่างทั้งหมดออกเป็นผลการตอบข้อสอบย่อย จำนวน 4 ชุด โดยจำแนกตามองค์ประกอบในการวัดที่แตกต่างกัน ซึ่งจะได้ผลการตอบข้อสอบชุดที่ 1 เป็นผลการตอบข้อสอบที่วัดความสามารถเกี่ยวกับความรู้ความเข้าใจทางคณิตศาสตร์ ผลการตอบข้อสอบชุดที่ 2 เป็นผลการตอบข้อสอบที่วัดความสามารถด้านการคิดคำนวณ ผลการตอบข้อสอบชุดที่ 3 เป็นผลการตอบข้อสอบที่วัดความสามารถด้านการพิจารณาผลลัพธ์ที่สมเหตุสมผล และผลการตอบข้อสอบชุดที่ 4 เป็นผลการตอบข้อสอบที่วัดความสามารถด้านการแก้โจทย์ปัญหา

4.2 คำนวณค่าสถิติพื้นฐาน เพื่อบรรยายลักษณะการแจกแจงของคะแนนจากแบบทดสอบความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ซึ่งวิเคราะห์โดยใช้โปรแกรม SPSS/PC⁺

4.3 ใช้โปรแกรมวิเคราะห์ข้อสอบ CTIA (สุพัฒน์ สุกมลสันต์, 2538) ซึ่งเป็นโปรแกรมที่วิเคราะห์ข้อสอบตามทฤษฎีแบบดั้งเดิม (Classical Test Theory) โดยดำเนินการ ดังนี้

4.3.1 วิเคราะห์คุณภาพของข้อสอบ โดยการวิเคราะห์ค่าความยาก(p) และค่าอำนาจจำแนก(ถรวายข้อ จากสูตร

$$p = \frac{P_H + P_L}{2}$$

$$\text{และ } r = P_H - P_L$$

เมื่อ p = ค่าความยากของข้อสอบ

r = ค่าอำนาจจำแนกของข้อสอบ

P_H = สัดส่วนของคนที่ตอบถูกในกลุ่มสูง

P_L = สัดส่วนของคนที่ตอบถูกในกลุ่มต่ำ

4.3.2 วิเคราะห์คุณภาพของแบบทดสอบความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 โดยการคำนวณค่าความเที่ยงเชิงความสอดคล้องภายในของแบบสอบด้วยสูตร KR-20 ของคูเดอร์-ริชาร์ดสัน (Kuder-Richardson) ดังนี้

$$r_u = \left[\frac{n}{n-1} \right] \left[\frac{1 - \sum pq}{S_x^2} \right]$$

- เมื่อ r_u = ค่าความเที่ยง
 n = จำนวนข้อสอบ
 $\sum pq$ = ผลรวมของความแปรปรวนรายข้อ
 S_x^2 = ความแปรปรวนของคะแนนรวม
 p = สัดส่วนของผู้ที่ตอบถูก
 q = สัดส่วนของผู้ที่ตอบผิด

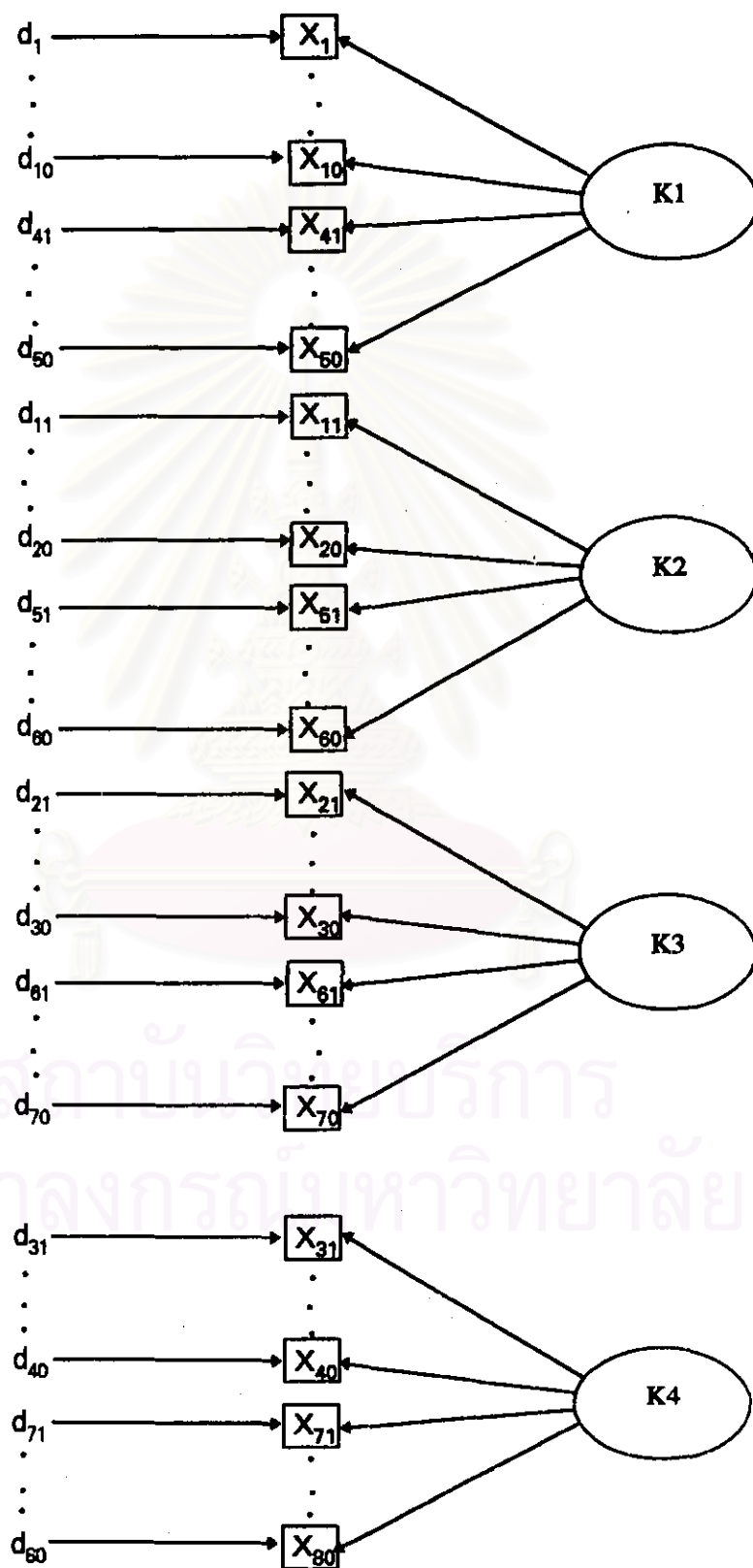
4.4 ทดสอบความตรงเชิงโครงสร้างของแบบทดสอบความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ด้วยวิธีวิเคราะห์องค์ประกอบเชิงยืนยัน(CFA) โดยดำเนินการ ดังนี้

4.4.1 คำนวณค่ามัธยฐานเลขคณิต และส่วนเบี่ยงเบนมาตรฐานของคะแนนแต่ละองค์ประกอบของผู้สอบทั้งหมด และคำนวณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันของข้อสอบแต่ละด้าน โดยใช้โปรแกรมคอมพิวเตอร์ SPSS/PC⁺

4.4.2 วิเคราะห์องค์ประกอบเชิงยืนยัน (CFA) ของแบบทดสอบความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 โดยใช้โปรแกรมคอมพิวเตอร์ LISREL 8.10 ซึ่งมีโมเดลองค์ประกอบของแบบสอบ ดังแผนภาพที่ 6

สถาบันวิทยบริการ
 จุฬาลงกรณ์มหาวิทยาลัย

แผนภาพที่ 5 โมเดลองค์ประกอบแบบทดสอบความสามารถทางคณิตศาสตร์
ชั้นประถมศึกษาปีที่ 6



- เมื่อ
- K1 = ความรู้ความเข้าใจทางคณิตศาสตร์
 - K2 = ความสามารถด้านการคิดคำนวณ
 - K3 = ความสามารถด้านการพิจารณาผลลัพธ์ที่สมเหตุสมผล
 - K4 = ความสามารถด้านการแก้โจทย์ปัญหา
 - X_i = คะแนนข้อสอบข้อที่ i , $i = 1, 2, 3, \dots, 80$
 - d_i = ความคลาดเคลื่อนของคะแนนข้อสอบข้อที่ i , $i = 1, 2, 3, \dots, 80$

4.5 วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล - แชนส์เซล (MH) และวิธีซิบเทสต์ (SIBTEST) แยกตามการใช้เกณฑ์จับคู่เปรียบเทียบ ได้แก่ คะแนนรวมและคะแนนแบบสอบย่อย โดยนำผลการตอบข้อสอบที่ได้จากการเก็บข้อมูลมาคำนวณหาดัชนี α_{MH} และดัชนี β_{SB} และทดสอบนัยสำคัญ ซึ่งค่าเหล่านี้สามารถคำนวณได้พร้อมกันด้วยโปรแกรม SIBTEST ของ Stout และ Roussos(1992)

4.6 วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีถดถอยโลจิสติก (LR) แยกตามการใช้เกณฑ์จับคู่เปรียบเทียบ ซึ่งได้แก่ คะแนนรวม คะแนนแบบสอบย่อย และคะแนนหลายแบบสอบย่อย โดยนำเพิ่มข้อมูลผลการตอบข้อสอบของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบไปวิเคราะห์ด้วยโปรแกรม SPSS/PC* ซึ่งวิธีถดถอยโลจิสติกเป็นการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ทดสอบด้วยโมเดลทางสถิติ Logistic Regression เพื่อทำนายโอกาสในการตอบข้อสอบได้ถูกต้อง ซึ่งพิจารณาจากจุดตัดและค่าความชันของพารามิเตอร์

4.7 คำนวณอัตราความถูกต้องและอัตราความคลาดเคลื่อนในการตรวจสอบ โดยนำผลการวิเคราะห์ด้วยวิธีแมนเทิล - แชนส์เซล (MH) และวิธีถดถอยโลจิสติก (LR) มาเปรียบเทียบกับผลการตรวจสอบด้วยวิธีซิบเทสต์ (SIBTEST) โดยจะใช้เกณฑ์ว่า ข้อสอบที่ถูกตรวจพบว่าทำหน้าที่ต่างกันด้วยวิธีซิบเทสต์ถือว่าเป็นข้อสอบที่ทำหน้าที่ต่างกัน(DIF) หากวิธีแมนเทิล-แชนส์เซลหรือวิธีถดถอยโลจิสติกตรวจพบข้อสอบที่ทำหน้าที่ต่างกันตรงกับที่ตรวจพบด้วยวิธีซิบเทสต์ ถือว่าวิธีนั้นมีความถูกต้องในการตรวจสอบ แต่ถ้าตรวจพบข้อสอบข้อใดทำหน้าที่ต่างกันอย่างอื่นจากที่ตรวจพบด้วยวิธีซิบเทสต์ ถือว่ามีความคลาดเคลื่อนในการตรวจสอบ

4.8 ทดสอบอัตราความถูกต้องและอัตราความคลาดเคลื่อนในการตรวจสอบระหว่างวิธีแมนเทิล - แชนส์เซล (MH) กับวิธีถดถอยโลจิสติก(LR) ซึ่งแยกตามเกณฑ์จับคู่เปรียบเทียบ ด้วยค่าสถิติ Z ที่ระดับนัยสำคัญ .05

5. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล-แฮนส์เชลและวิธีซิบเทสท์

การตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันด้วยวิธีแมนเทิล-แฮนส์เชล (MH) และวิธีซิบเทสท์ (SIBTEST) สามารถวิเคราะห์ได้พร้อมกันโดยใช้โปรแกรม SIBTEST ของ Stout และ Roussos(1992) ผลจากการวิเคราะห์จะได้ค่าดัชนี α_{MH} , ดัชนี β_{SIB} และทดสอบนัยสำคัญทางสถิติ สำหรับเพิ่มข้อมูลที่ใช้ป้อนเข้าไปในโปรแกรม คือ เพิ่มข้อมูลผลการตอบข้อสอบของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ โดยภายในเพิ่มข้อมูลแต่ละเพิ่มจะมีเพียงผลการตอบข้อสอบของผู้สอบตามจำนวนข้อสอบที่จะนำมาใช้เป็นเกณฑ์จับคู่ (matching criteria) เท่านั้น การใช้โปรแกรมวิเคราะห์ในแต่ละครั้งต้องใช้เพิ่มข้อมูลของทั้งกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ ในขั้นแรกของการวิเคราะห์เริ่มด้วยการใช้โปรแกรมย่อย SIBIN (SIB INPUT) เพื่อจัดเตรียมข้อมูลจากเพิ่มข้อมูลผลการตอบของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ ขั้นตอนที่สองนำเพิ่มข้อมูลซึ่งเป็นผล (out put) จากโปรแกรมย่อย SIBIN มาเป็นเพิ่มข้อมูลสำหรับป้อนเข้าโปรแกรมย่อย SIBTEST ซึ่งผลที่ได้ประกอบด้วยค่า α_{MH} , ค่า χ^2_{MH} , ดัชนี β_{SIB} และการทดสอบนัยสำคัญทางสถิติโดยพิจารณาจากค่า p-value ถ้าพบว่ามีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่าข้อสอบนั้นทำหน้าที่ต่างกัน (DIF)

ในการวิจัยครั้งนี้เป็นการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ(DIF) ระหว่างวิธีแมนเทิล-แฮนส์เชลและวิธีถดถอยโลจิสติกเมื่อใช้เกณฑ์จับคู่เปรียบเทียบที่ต่างกัน การวิเคราะห์ DIF ด้วยวิธีซิบเทสท์ซึ่งเป็นวิธีที่นำมาประเมินประสิทธิภาพของการตรวจสอบใช้คะแนนรวม (total test score) เป็นเกณฑ์จับคู่เปรียบเทียบ ส่วนวิธีแมนเทิล-แฮนส์เชลเป็นวิธีหนึ่งนำมาเปรียบเทียบใช้เกณฑ์จับคู่ 2 เกณฑ์ คือ คะแนนรวม (total test score) และคะแนนแบบสอบย่อย (subtest score)

6. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีถดถอยโลจิสติก

การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีถดถอยโลจิสติก(LR) วิธีนี้เป็นการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ทดสอบด้วยโมเดลทางสถิติ Logistic Regression เพื่อทำนายโอกาสหรือความน่าจะเป็นในการตอบข้อสอบได้ถูกหรือผิดของผู้สอบ โดยตัวที่ให้ทำนายได้แก่ กลุ่มผู้สอบ (g) ซึ่งแบ่งเป็นกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ ความสามารถของผู้สอบ(x) และปฏิสัมพันธ์ระหว่างกลุ่มผู้สอบและความสามารถของผู้สอบ (g by x) ในการตรวจ DIF ด้วยวิธีนี้วิเคราะห์โดยใช้โปรแกรม SPSS/PC ส่วนข้อมูลที่ป้อนเข้าเป็นรหัสของกลุ่มผู้สอบและผลการตอบข้อสอบที่ให้คะแนนแบบ 0,1

ในการวิจัยนี้ วิธี LR เป็นอีกวิธีหนึ่งที่น่ามาเปรียบเทียบประสิทธิภาพในการตรวจ DIF กรณีที่ใช้เกณฑ์จับคู่เปรียบเทียบแตกต่างกัน 3 เกณฑ์ คือ คะแนนรวม คะแนนแบบสอบย่อย และคะแนนหลายแบบสอบย่อย

สำหรับการวิเคราะห์เมื่อใช้คะแนนรวมและคะแนนแบบสอบย่อยเป็นเกณฑ์จับคู่ จะใช้กลุ่มผู้สอบ(g), ความสามารถของผู้สอบ(x) และปฏิสัมพันธ์ระหว่างกลุ่มผู้สอบและความสามารถของผู้สอบ(g by x) เป็นตัวทำนาย ส่วนเมื่อใช้คะแนนหลายแบบสอบย่อยเป็นเกณฑ์จับคู่ นอกจากจะใช้ตัวทำนายดังกล่าวแล้วยังนำความสามารถของผู้สอบทุกแบบสอบย่อย(x1, x2, x3, x4) เข้ามาเป็นตัวทำนายร่วมด้วย

การพิจารณาข้อสอบที่ทำหน้าที่ต่างกัน (DIF) จากการวิเคราะห์ด้วยวิธี LR พิจารณาจากการมีนัยสำคัญทางสถิติ (Significant) ของค่าปฏิสัมพันธ์ระหว่างกลุ่มผู้สอบและความสามารถของผู้สอบ(g by x) หรือพิจารณาจากการมีนัยสำคัญทางสถิติ (Significant) ของกลุ่มผู้สอบ(g)

7. คำนวนอัตราความถูกต้องของการตรวจสอบและอัตราความคลาดเคลื่อน

คำนวนหาอัตราความถูกต้องของการตรวจสอบ โดยนำผลการตรวจ DIF ด้วยวิธีแมนเทิล-แฮนส์เซล (MH) และวิธีถดถอยโลจิสติก (LR) เมื่อใช้เกณฑ์จับคู่เปรียบเทียบแตกต่างกันมาเปรียบเทียบกับผลการตรวจสอบด้วยวิธีซิปเทสท์ ซึ่งเป็นวิธีที่ใช้ประเมินประสิทธิภาพของการตรวจสอบ ดังนั้นหากข้อสอบข้อใดถูกตรวจพบว่าทำหน้าที่ไม่ต่างกันด้วยวิธีซิปเทสท์จะถือว่าเป็นข้อสอบที่ทำหน้าที่ไม่ต่างกัน หากวิธีแมนเทิล-แฮนส์เซลหรือวิธีถดถอยโลจิสติกตรวจพบข้อสอบที่ทำหน้าที่ไม่ต่างกันตรงกับที่ตรวจพบด้วยวิธีซิปเทสท์ ก็ถือว่าวิธีนั้นมีความถูกต้องประเภทที่ 1 ในการตรวจสอบและถ้าตรวจพบว่าข้อสอบข้อใดทำหน้าที่ต่างกันตรงกับที่ตรวจพบด้วยวิธีซิปเทสท์ถือว่าวิธีนั้นมีความถูกต้องประเภทที่ 2 ในการตรวจสอบ การคำนวนอัตราความถูกต้องในการตรวจสอบประเภทที่ 1 และ 2 ทำได้ดังนี้

อัตราความถูกต้องประเภทที่ 1

$$= \frac{\text{จำนวนข้อสอบที่ตรวจพบว่าทำหน้าที่ไม่ต่างกันได้ถูกต้อง}}{\text{จำนวนข้อสอบที่ทำหน้าที่ไม่ต่างกันทั้งหมดในแบบสอบ}}$$

อัตราความถูกต้องประเภทที่ 2

$$= \frac{\text{จำนวนข้อสอบที่ตรวจพบว่าทำหน้าที่ต่างกันได้ถูกต้อง}}{\text{จำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้งหมดในแบบสอบ}}$$

สำหรับอัตราความคลาดเคลื่อน ค่าวนได้โดยนำผลการตรวจ DIF ด้วยวิธี MH และวิธี LR เมื่อใช้เกณฑ์จับคู่เปรียบเทียบแตกต่างกันมาเปรียบเทียบกับผลการตรวจสอบด้วยวิธีซิปเทสท์ ถ้าข้อสอบข้อใดที่วิธีแมนเทล-แฮนส์เซลหรือวิธีดอดอยโลจิสติกตรวจพบว่าทำหน้าที่ต่างกันแต่วิธีซิปเทสท์ตรวจพบว่าเป็นข้อสอบที่ทำหน้าที่ไม่ต่างกัน ก็จะถือว่าวิธีนั้นมีการระบุผิดพลาดซึ่งทำให้เกิดความคลาดเคลื่อนประเภทที่ 1 และถ้าตรวจพบว่าข้อสอบข้อใดทำหน้าที่ไม่ต่างกันแต่วิธีซิปเทสท์ตรวจพบว่าเป็นข้อสอบที่ทำหน้าที่ต่างกันก็จะถือว่าวิธีนั้นมีการระบุผิดพลาดซึ่งทำให้เกิดความคลาดเคลื่อนประเภทที่ 2 ซึ่งการคำนวณอัตราความคลาดเคลื่อนในการตรวจสอบประเภทที่ 1 และ 2 ทำได้ดังนี้

อัตราความคลาดเคลื่อนชนิดที่ 1 (Fault Positive : FP)

$$= \frac{\text{จำนวนข้อสอบที่ระบุผิดพลาดว่าทำหน้าที่ต่างกัน}}{\text{จำนวนข้อสอบที่ทำหน้าที่ไม่ต่างกันทั้งหมดในแบบสอบ}}$$

อัตราความคลาดเคลื่อนชนิดที่ 2 (Fault Negative : FN)

$$= \frac{\text{จำนวนข้อสอบที่ระบุผิดพลาดว่าทำหน้าที่ไม่ต่างกัน}}{\text{จำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้งหมดในแบบสอบ}}$$

8. ทดสอบอัตราความถูกต้องของการตรวจสอบและอัตราความคลาดเคลื่อนด้วยสถิติ Z

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในการวิจัยนี้เป็นการเปรียบเทียบซีในการตรวจสอบ 2 วิธีภายใต้เกณฑ์จับคู่เปรียบเทียบ 3 เกณฑ์ ดังนั้นจึงมีระดับการศึกษา 5 ลักษณะ คือ ในลักษณะที่ 1 เป็นการตรวจ DIF ด้วยวิธีแมนเทล-แฮนส์เซล(MH)เมื่อใช้คะแนนรวมเป็นเกณฑ์จับคู่(MH_{TOT}) ลักษณะที่ 2 เป็นการตรวจ DIF ด้วยวิธีแมนเทล-แฮนส์เซล(MH)เมื่อใช้คะแนนแบบสอบย่อยเป็นเกณฑ์จับคู่(MH_{SUB}) ลักษณะที่ 3, 4, และ 5 เป็นการตรวจ DIF ด้วยวิธีดอดอยโลจิสติก (LR) เมื่อใช้คะแนนรวม (LR_{TOT}), คะแนนแบบสอบย่อย(LR_{SUB})และคะแนนหลายแบบลอบย่อย($LR_{MULTIPLE}$)เป็นเกณฑ์จับคู่ จากสมมติฐานการวิจัยสามารถสรุปได้ดังนี้

พิจารณาอัตราความถูกต้องในการตรวจสอบ จะได้สมมติฐานการวิจัย คือ

$$H_0 : LR = MH$$

$$H_1 : LR > MH$$

และพิจารณาอัตราความคลาดเคลื่อน จะได้สมมติฐานการวิจัย คือ

$$H_0 : LR = MH$$

$$H_1 : LR < MH$$

ดังนั้นการทดสอบอัตราความถูกต้องและอัตราความคลาดเคลื่อนในการตรวจสอบด้วยสถิติ Z จึงเป็นการเปรียบเทียบความแตกต่างของอัตราความถูกต้องและอัตราความคลาดเคลื่อนในการตรวจสอบระหว่างวิธีถดถอยโลจิสติกกับวิธีแมนเทิล-แฮนส์เซล โดยการทดสอบด้วยค่าสถิติ Z ที่ระดับนัยสำคัญ .05

$$Z_p = \frac{|P_1 - P_2|}{\sqrt{PQ(1/n_1 + 1/n_2)}}$$

เมื่อ P_1, P_2 เป็นสัดส่วนที่ได้จากการตรวจสอบด้วยวิธี LR และวิธี MH

P เท่ากับ $\frac{(P_1)(n_1) + (P_2)(n_2)}{n_1 + n_2}$

Q เป็นค่าที่ได้จาก $1 - P$

n_1, n_2 จำนวนของข้อสอบที่ตรวจสอบด้วยวิธี LR และวิธี MH

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย