

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้แบ่งการดำเนินการออกเป็น 2 ระยะคือ ระยะที่ 1 เป็นการสร้างข้อสอบเพื่อนำไปใช้ในการทดสอบแบบ ซี เอ ที ระยะที่ 2 เป็นการเปรียบเทียบคุณภาพของการทดสอบแบบ ซี เอ ที และแบบประเพณีนิยม ดังรายละเอียดต่อไปนี้

การสร้างข้อสอบเพื่อนำไปใช้ในการทดสอบแบบ ซี เอ ที

ลักษณะและปริมาณของประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัยครั้งนี้คือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2530 ของโรงเรียนในสังกัดกรมสามัญศึกษา จังหวัดนนทบุรี จำนวน 6,201 คน จากโรงเรียนทั้งหมด 17 โรงเรียน

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ได้แก่ นักเรียนของโรงเรียนที่เหลือจากการทดลองเครื่องมือ เพื่อการวิจัยครั้งนี้ ซึ่งมีจำนวน 5,632 คน จากโรงเรียน 16 โรงเรียน และเพื่อให้ได้กลุ่มตัวอย่างมากเพียงพอที่จะใช้ในการวิเคราะห์ข้อสอบตามรูปแบบโลจิสติก 3 พารามิเตอร์ จึงสุ่มตัวอย่างโดยมีเป้าหมายว่าให้ได้กลุ่มตัวอย่างจากโรงเรียนแต่ละแห่งประมาณร้อยละ 75 จากนมนำแบบสอที่ เป็นเครื่องมือ เพื่อการวิจัยครั้งนี้ไปสอบ ปรากฏว่าได้กลุ่มตัวอย่างทั้งสิ้นจำนวน 4,800 คน จากโรงเรียน 16 โรงเรียน มีรายละเอียดดังนี้

ตารางที่ 3 จำนวนกลุ่มตัวอย่างจำแนกตามโรงเรียน

ที่	โรงเรียน	จำนวนนักเรียนทั้งหมด	จำนวนกลุ่มตัวอย่าง	%ของนักเรียนจากโรงเรียน	%ของกลุ่มตัวอย่าง
1.	รัตนธิเบศร์	475	398	83.79	8.29
2.	วัดเขมาภิรตาราม	562	500	88.97	10.42
3.	ศรีบุญยานนท์	570	503	88.25	10.48
4.	สตรีนนทบุรี	603	566	93.86	11.79
5.	เบญจมาราชานุสรณ์	772	701	90.80	14.60

ตารางที่ 3 (ต่อ)

ที่	โรงเรียน	จำนวน นักเรียน ทั้งหมด	จำนวน กลุ่มตัวอย่าง	%ของนักเรียน จากโรงเรียน	%ของกลุ่ม ตัวอย่าง
6.	นนทบุรีพิทยาคม	325	265	81.54	5.52
7.	ไทรน้อย	96	87	90.63	1.81
8.	ราชดรุณีนิยม	77	73	94.81	1.52
9.	บางกรวย	218	181	83.03	3.77
10.	วัดนไชยศรีบุญญาคม	140	97	69.29	2.02
11.	บางบัวทอง	268	175	47.55	3.65
12.	บางใหญ่	249	232	93.17	4.83
13.	นนทกิจประชาอุปถัมภ์	36	36	100.00	0.75
14.	คลองพระอุดมวิทยาคม	159	137	86.16	2.86
15.	สวนกุหลาบวิทยาลัยนนทบุรี	614	557	90.72	11.61
16.	โพธิ์นิมิตวิทยาคม	368	292	79.35	6.08
	รวม	5,632	4,800	85.23	100.00

จากตารางที่ 3 แสดงว่า กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ได้มาจากโรงเรียนมัธยมศึกษาในสังกัดกรมสามัญศึกษา จังหวัดนนทบุรีทุกโรงเรียน ซึ่งคิดเป็นร้อยละจากจำนวนนักเรียนในแต่ละโรงเรียนจะมีระหว่างร้อยละ 47 ถึงร้อยละ 100 ซึ่งเฉลี่ยแล้วประมาณร้อยละ 85.23 จากประชากรทั้งหมดหรือ 4,800 คน ซึ่งเป็นจำนวนที่มากเพียงพอสำหรับการวิจัยต่อไป

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้เป็นแบบสอบถามวัดความสามารถด้านคำศัพท์ภาษาอังกฤษที่ผู้วิจัยสร้างขึ้น ประกอบด้วยข้อสอบแบบเลือกตอบ 4 ตัวเลือก จำนวน 500 ข้อ โดยมีขั้นตอนในการสร้างดังนี้

1. ศึกษาหลักสูตร คู่มือครู แบบเรียนและเอกสารที่เกี่ยวข้อง ผู้วิจัยได้ศึกษาหลักสูตรวิชาภาษาอังกฤษระดับมัธยมศึกษาตอนต้น

2. กำหนดขอบเขตและรูปแบบของข้อสอบ โดยสำรวจรายชื่อหนังสือแบบเรียนภาษาอังกฤษชั้นมัธยมศึกษาตอนต้น จากโรงเรียนต่าง ๆ ในสังกัดกรมสามัญศึกษา จังหวัดนนทบุรี พบว่า โรงเรียนต่าง ๆ ใช้หนังสือแบบเรียนภาษาอังกฤษดังรายชื่อต่อไปนี้



1. Junior Active Context English
2. Learning Kit
3. English for a Changing World
4. Odyssey

จากเอกสารที่เกี่ยวข้อง พบว่า ควรใช้ศัพท์ที่มีความหมายในตัวเอง (Content Word) ซึ่งได้แก่ คำนาม คำกริยา คำคุณศัพท์ และคำวิเศษณ์ (Madsen 1983:19-27) ในการทดสอบ คำศัพท์ จึงกำหนดน้ำหนักของคำเป็น คำนาม:คำกริยา:คำคุณศัพท์:คำวิเศษณ์ เป็น 5.5:2.0:1.5:1.0 (กัลยา ไกวัลดิช 2527: 48-119 และทพาวรรณ จันทร์พิบูลย์ 2519:62) โดยใช้รูปแบบในการสร้างข้อสอบเป็น 4 รูปแบบคือ

1. แบบเลือกคำจากข้อความที่ตัดให้ เหมาะกับคำที่กำหนดให้
2. แบบเลือกคำที่มีความหมายเหมือนกับคำที่ขีดเส้นใต้
3. แบบเลือกคำที่มีความหมายตรงข้ามกับคำที่ขีดเส้นใต้
4. แบบเติมคำให้สมบูรณ์ โดยให้เลือกคำที่มีความหมายเหมาะกับประโยคที่กำหนดให้

นำรูปแบบดังกล่าวให้ผู้เชี่ยวชาญกำหนดน้ำหนักของรูปแบบ ผู้เชี่ยวชาญมีความเห็นว่ารูปแบบเติมคำให้สมบูรณ์ โดยให้เลือกคำที่มีความหมายเหมาะกับประโยคที่กำหนดให้ เป็นรูปแบบที่ให้ลารสนะเทศมากกว่ารูปแบบอื่น ๆ และมีความเห็นว่าควรกำหนดน้ำหนักเป็น 1:1:1:7 ผู้วิจัยจึงนำมากำหนดเป็นจำนวนข้อดังตารางที่ 4

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4 จำนวนข้อสอบจำแนกตามรูปแบบของคำถาม

รูปแบบของคำถาม	จำนวนข้อ				น้ำหนัก คิดเป็น ร้อยละ	รวม
	คำถาม	คำกริยา	คำคุณศัพท์	คำวิเศษณ์		
1. เลือกคำจำกัดความ ที่ดีที่สุด ให้เหมาะกับ คำที่กำหนดให้	28	10	7	5	10	50
2. เลือกคำที่มีความหมาย เหมือนกับคำที่ขีดเส้นใต้	27	10	8	5	10	50
3. เลือกคำที่มีความหมาย ตรงข้ามกับคำที่ขีดเส้นใต้	27	10	8	5	10	50
4. เติมคำให้สมบูรณ์โดย ให้เลือกคำที่มีความ หมายเหมาะกับประโยค ที่กำหนดให้	193	70	52	35	70	350
รวม	275	100	75	50	100	500

3. สร้างแบบสอบตามรูปแบบและขอบเขตที่กำหนดขึ้น

3.1 คัดเลือกคำศัพท์ที่ปรากฏในแบบเรียนทั้ง 4 เล่ม นำมาสร้างเป็นข้อสอบแบบ
เลือกตอบ 4 ตัวเลือก ตามรูปแบบดังตารางข้างต้น

3.2 นำแบบสอบต้นฉบับที่สร้างขึ้นไปให้ผู้เชี่ยวชาญ (ดังภาคผนวก) ตรวจสอบแก้ไข
ความถูกต้องทางภาษาของข้อสอบทั้งตัวคำถามและตัวเลือก

3.3 นำแบบสอบต้นฉบับที่ผ่านการพิจารณาจากผู้เชี่ยวชาญมาปรับปรุงแก้ไขตาม
ข้อเสนอแนะ

3.4 จัดข้อสอบเป็นชุด ๆ ละ 100 ข้อ โดยให้แต่ละชุดประกอบด้วยข้อสอบที่มีรูปแบบของคำถามและจำนวนข้อให้มีน้ำหนักดังตารางข้างต้น

4. ทดลองใช้ นำข้อสอบที่สร้างและปรับปรุงแล้วจำนวน 500 ข้อ มาจัดเป็นชุด ๆ ละ 100 ข้อตามน้ำหนักของรูปแบบที่กำหนดไว้ นำไปทดสอบกับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โรงเรียนปากเกร็ด จังหวัดนนทบุรี จำนวน 569 คน แจกข้อสอบแบบสุ่ม นำผลการทดสอบมาวิเคราะห์โดยใช้โปรแกรมวิเคราะห์ข้อสอบ ITEMX (สัพพจน์ สกมลสันต์ 2530:67) ทั้งนี้เพื่อนำผลการวิเคราะห์มาปรับปรุงคุณภาพข้อสอบ ได้ผลการวิเคราะห์ข้อสอบดังนี้

ตารางที่ 5 ผลการวิเคราะห์ข้อสอบรายข้อ

แบบสอบ ฉบับที่	จำนวน ข้อ	คะแนน เฉลี่ย	ส่วนเบี่ยงเบน มาตรฐาน	ความยาก เฉลี่ย	ค่าอำนาจ จำแนก	ค่าความเที่ยง	ความคลาด เคลื่อน มาตรฐาน
1	100	40.32	15.85	0.41	0.43	.94	3.81
2	100	40.92	13.04	0.41	0.35	.86	4.89
3	100	40.73	14.80	0.40	0.38	.90	4.61
4	100	38.66	15.23	0.39	0.40	.93	4.15
5	100	41.38	12.05	0.41	0.32	.84	4.79

จากตารางที่ 5 แสดงให้เห็นว่าโดยเฉลี่ยแล้วแบบสอบมีระดับความยากง่ายปานกลาง ข้อสอบมีค่าอำนาจจำแนกเฉลี่ยแล้วเป็นข้อสอบที่มีอำนาจจำแนกดี แต่เมื่อพิจารณาเป็นรายข้อแล้วพบว่า มีข้อสอบบางข้อที่มีค่าอำนาจจำแนกดีกลับ แสดงว่า นักเรียนอ่อนทำได้ถูกแต่นักเรียนเก่งทำผิด ซึ่งเป็นข้อสอบที่ไม่ดี แบบสอบทั้ง 5 ฉบับมีค่าความเที่ยงอยู่ระหว่าง .84 และ .94 และค่าความคลาดเคลื่อนมาตรฐานในการวัดอยู่ในช่วง 3.81 และ 4.89

เพื่อให้ได้ข้อสอบที่มีคุณภาพดีเพื่อใช้ในการทดสอบแบบ ซี เอ ที ผู้วิจัยจึงพิจารณาโดยใช้เกณฑ์ในการคัดเลือกข้อสอบคือ ข้อสอบจะต้องมีค่าอำนาจจำแนกตั้งแต่ .20 ขึ้นไป ค่าอำนาจจำแนกของตัวลวงมีค่าดีดลบและมีค่ามากเท่า ๆ กัน แต่มีค่าน้อยกว่าค่าอำนาจจำแนกของตัวเลือกทุก คะแนนรวมเฉลี่ยของผู้สอบที่เลือกตอบข้อสอบ (mean criterion score) ของตัวเลือกทุก ข้อมีค่าสูงกว่าตัวลวงและค่าคะแนนมาตรฐานแบบ T-score ของตัวเลือกทุก เป็นค่าที่มากที่สุดและค่าความยากง่ายของตัวลวงควรมีค่าเท่า ๆ กัน (สัพพจน์ สกมลสันต์ 2530:84) หลังจากคัดเลือกโดยใช้เกณฑ์ดังกล่าวแล้วได้ข้อสอบที่มีคุณภาพจำนวน 440 ข้อ ผู้วิจัยจึงนำข้อสอบมาละกันแล้วสุ่มข้อสอบขึ้นมาจัดเป็นชุด ๆ ละ 88 ข้อ จำนวน 5 ชุด โดยมีรูปแบบและการกระจายตามน้ำหนักดังตารางที่ 4

การเก็บรวบรวมข้อมูล

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูล โดยนำหนังสือขอความร่วมมือจากบัณฑิตวิทยาลัย ไปยังสำนักงานศึกษาธิการเขต เขตการศึกษา 1 เพื่อขอความอนุเคราะห์จากโรงเรียนต่าง ๆ ที่เป็นกลุ่มตัวอย่างที่ใช้ในการเก็บรวบรวมข้อมูล จากนั้นผู้วิจัยได้นำหนังสือจากสำนักงานศึกษาธิการเขต เขตการศึกษา 1 ไปติดต่อกับทางโรงเรียนเหล่านั้น โดยมีวิธีการดำเนินการเก็บรวบรวมข้อมูล ดังนี้

1. ในการสอบผู้วิจัยได้ติดต่อกับผู้บริหารโรงเรียน ผู้ช่วยฝ่ายวิชาการ เพื่อชี้แจงวัตถุประสงค์ของการวิจัย และนัดหมายเรื่องวัน เวลา และสถานที่สอบ ตลอดจนขอความร่วมมือจากอาจารย์ที่สอนในคาบที่มีการทดสอบให้ช่วยควบคุมและดำเนินการสอบ
2. จัดเตรียมแบบสอบทั้ง 5 ฉบับ โดยในแต่ละฉบับประกอบด้วยข้อสอบ กระดาษคำตอบ คำแนะนำในการตอบแบบสอบและคำแนะนำในการบริหารการสอบ จัดเรียงแบบสอบทั้ง 5 ฉบับให้เป็นระบบเดียวกันคือ เรียงแบบสอบฉบับที่ 1, 2, 3, 4 และ 5 ต่อเนื่องกันไป ทั้งนี้เพื่อให้การสอบมีลักษณะเหมือนเล่ม หนึ่งและบรรจุของแบบสอบตามจำนวนของนักเรียนในแต่ละห้องเรียน
3. ผู้วิจัยนำแบบสอบไปหาการสอบกับนักเรียนตามโรงเรียนต่าง ๆ ดังแสดงไว้ในตารางที่ 3 ตามวัน เวลา ที่นัดหมายไว้ ในด้านการบริหารการสอบผู้วิจัยได้รับความอนุเคราะห์จากอาจารย์ที่สอนในคาบที่มีการทดสอบ ในการสอบแต่ละห้องเรียนให้แจกข้อสอบให้เป็นระบบเดียวกันและให้นักเรียนตรวจดูความถูกต้องของกระดาษคำตอบให้ตรงกับชุดข้อสอบที่ได้รับ
4. หลังจากดำเนินการสอบแล้วนำกระดาษคำตอบที่ได้มาตรวจสอบความเรียบร้อย คัดกระดาษคำตอบที่ไม่สมบูรณ์ออก เช่น กากบาทคำตอบข้อใดข้อหนึ่งทั้งฉบับ แล้วนำกระดาษคำตอบทั้งหมดจัดเรียงลำดับให้เป็นระบบในแต่ละโรงเรียน ได้จำนวนนักเรียนที่ตอบแบบสอบแต่ละชุด ดังตารางที่ 6

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 6 จำนวนกลุ่มตัวอย่างที่ใช้ในการวิจัย จำแนกตามชุดของแบบสอบ

โรงเรียน	แบบสอบฉบับที่					รวม
	1	2	3	4	5	
วัดนาถิเบศร์	107	101	88	49	53	398
วัดเขมาภิรตาราม	115	108	105	102	70	500
ศรีบุญยานนท์	111	101	99	97	95	503
สตรีนนทบุรี	113	116	114	114	109	566
เบญจมาราชานุสรณ์	138	144	145	151	123	701
นนทบุรีพิทยาคม	-	-	88	88	89	265
ไทรน้อย	19	18	19	17	14	87
ราชบุรีนิยม	15	15	15	14	14	73
บางกรวย	39	36	36	35	35	181
วัดนโชนศรีบุญฤศคาม	21	20	17	19	20	97
บางบัวทอง	40	41	38	31	25	175
บางใหญ่	48	47	46	45	46	232
นนทบุรีประชาอุปถัมภ์	8	7	7	7	7	36
คลองพระอุดมวิทยาคม	32	25	26	26	28	137
สวนกุหลาบวิทยาลัยนนทบุรี	112	111	112	111	111	557
โพชนิมิตวิทยาคม	59	61	60	61	51	292
รวม	977	951	1,015	967	890	4,800

จากตารางที่ 6 แสดงว่ากลุ่มตัวอย่างที่ใช้ในการศึกษาค้างนี้ได้มาจากโรงเรียนมัธยมศึกษาสังกัดกรมสามัญศึกษา จำนวน 16 โรงเรียน มีกลุ่มตัวอย่างที่ทดสอบแบบสอบฉบับที่ 1, 2, 3, 4 และ 5 จำนวน 977, 951, 1,015, 967 และ 890 คน ตามลำดับ รวมจำนวนทั้งสิ้น 4,800 คน

5. นักกระดาษคำตอบมาตรวจให้คะแนนและบันทึกผลลงในกระดาษลงรหัส โดยให้ข้อที่ตอบถูกเป็น 1 ข้อที่ตอบผิดเป็น 0 ข้อที่ทำข้ามเป็น 2 ข้อที่ทำไม่ได้ถึงข้อสุดท้ายเป็น 3 (สัพพัญญู สุกมลสันต์ 2530:110)

6. บันทึกข้อมูลลงแผ่นบันทึกข้อมูลแล้วนำไปให้เจ้าหน้าที่ของสถาบันคอมพิวเตอร์ถ่ายข้อมูลลงเทป เพื่อใช้ในการวิเคราะห์ต่อไป

การวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลเพื่อการวิจัยครั้งนี้ใช้เครื่องคอมพิวเตอร์ของสถาบันบริการคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย โดยใช้โปรแกรมสำเร็จรูปชื่อ SPSS* โปรแกรมสำหรับการวิเคราะห์ข้อสอบแบบประเพณีนิยมชื่อ ITEMX ซึ่ง รศ.ดร.สัมพันธ์ สกมลสันต์ ได้พัฒนาขึ้นโปรแกรม Logist 5 Version 2.5 สำหรับการวิเคราะห์ข้อสอบโดยรูปแบบโลจิสติกที่มี 3 พารามิเตอร์และโปรแกรมภาษาฟอร์แทรนคำนวณค่าสารสนเทศแบบสอบ ซึ่งนายธีรศักดิ์ อัจฉรานนท์ เขียนขึ้น

ขั้นตอนต่าง ๆ ในการวิเคราะห์ข้อมูลมีดังนี้

1. ใช้โปรแกรมคอมพิวเตอร์ชื่อ ITEMX วิเคราะห์ข้อสอบรายข้อของแบบสอบแต่ละฉบับ เพื่อหาค่าความยากง่าย (\bar{p} และ $\bar{\Delta}$) ค่าอำนาจจำแนก (r_{b1b}) และค่าความเที่ยงของแบบสอบ
2. การตรวจสอบคุณสมบัติการวัดเที่ยงมิติเดียวและความเป็นอิสระในการตอบข้อสอบ ตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ (IRT) ได้กำหนดว่าแบบสอบที่จะใช้วิเคราะห์ด้วยโมเดลต่าง ๆ ในกรอบทฤษฎีจะต้องมีคุณสมบัติการวัดเที่ยงมิติเดียวและความเป็นอิสระในการตอบข้อสอบ ผู้วิจัยจึงทำการตรวจสอบคุณสมบัติดังกล่าว โดยใช้ข้อมูลที่รวบรวมได้นั้นมาวิเคราะห์ด้วยประกอบของแบบสอบแต่ละชุด โดยใช้โปรแกรม SPSS* การวิเคราะห์ด้วยประกอบนี้ใช้การวิเคราะห์ด้วยประกอบสำคัญ (Principal Component) แล้วหมุนแกนด้วยวิธีแวนซ์ (Varimax) (สัมพันธ์ สกมลสันต์ 2526:90) จากการวิเคราะห์ด้วยประกอบ ปรากฏว่าแบบสอบทั้ง 5 ฉบับมีค่าไอเกนมากกว่า 1.00 อยู่ 27-30 จำนวน ซึ่งถือว่าเป็นตัวประกอบร่วม (Lindeman and Others 1980:25 อ้างถึงในสัมพันธ์ สกมลสันต์ 2527:91) ค่าไอเกนของตัวประกอบที่ 1 ของแบบสอบแต่ละฉบับมีขนาดแตกต่างจากค่าไอเกนของตัวประกอบที่ 2 มากและค่าไอเกนของตัวประกอบที่ 2 ของแบบสอบแต่ละฉบับมีขนาดแตกต่างจากค่าไอเกน ลำดับต่อไปที่เหลือไม่มากนักจึงอาจกล่าวได้ว่าแบบสอบแต่ละฉบับวัดเที่ยงมิติเดียว (Lord 1980:21; Warm 1978:104) (รายละเอียดการวิเคราะห์ด้วยประกอบแสดงไว้ในแผนภาพที่ 9-13 ในภาคผนวก) และเนื่องจากแบบสอบทั้ง 5 ฉบับมีคุณสมบัติการวัดเที่ยงมิติเดียว ดังนั้นจึงมีคุณสมบัติความเป็นอิสระในการตอบข้อสอบด้วย (Hambleton and Swaminathan 1985:24; Warm 1978:107) เนื่องจากแบบสอบทั้ง 5 ฉบับมีลักษณะสอดคล้องกับข้อตกลงเบื้องต้นของวิธีการวิเคราะห์ข้อสอบรายข้อโดยใช้รูปแบบโลจิสติก 3 พารามิเตอร์ จึงสามารถนำผลการสอบไปวิเคราะห์โดยใช้โปรแกรม Logist 5 Version 2.5 ได้
3. การประมาณค่าพารามิเตอร์ของข้อสอบ โดยนำคะแนนผลการสอบของแบบสอบแต่ละชุดมาวิเคราะห์ด้วยโปรแกรม Logist 5 Version 2.5 เพื่อให้ได้ค่าพารามิเตอร์ของข้อสอบ คือ ค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าการเดาของข้อสอบ ซึ่งเป็นการวิเคราะห์ข้อสอบรายข้อตามโมเดลโลจิสติกที่มี 3 พารามิเตอร์ โดยสามารถเขียนเป็นรูปแบบความสัมพันธ์เชิงคณิตศาสตร์ของค่าพารามิเตอร์ได้ดังนี้

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{D a_i(\theta - b_i)}}{1 + e^{D a_i(\theta - b_i)}} \quad (\text{Hambleton 1985:49})$$

เมื่อ $P_i(\theta)$	คือ	ความน่าจะเป็นที่ผู้สอบที่มีระดับความสามารถตอบข้อสอบข้อ i ถูก
θ	คือ	ระดับความสามารถของผู้สอบ
a_i	คือ	ค่าอำนาจจำแนกของข้อสอบข้อ i
b_i	คือ	ค่าความยากของข้อสอบข้อ i
c_i	คือ	ค่าการเดาของข้อสอบข้อ i
D	คือ	Scaling Factor มีค่าเท่ากับ 1.7
e	คือ	ค่าคงที่มีค่าเท่ากับ 2.71822

4. การคำนวณค่าฟังก์ชันสารสนเทศของแบบสอบแต่ละฉบับ n ระดับความสามารถต่าง ๆ โดยการนำค่าพารามิเตอร์ของข้อสอบในแบบสอบแต่ละฉบับมาคำนวณหาค่าฟังก์ชันสารสนเทศของแบบสอบแต่ละฉบับ n ระดับความสามารถต่าง ๆ ซึ่งได้แบ่งระดับความสามารถออกเป็น 31 ระดับคือ -3.0 -2.8 -2.6 -2.4 -2.2 -2.0 -1.8 -1.6 -1.4 -1.2 -1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 โดยใช้สูตรในการคำนวณดังนี้

ค่าฟังก์ชันสารสนเทศแบบสอบ n ระดับความสามารถ $\theta_j(I(\theta_j))$

$$I(\theta_j) = \sum_{i=1}^n \frac{P_i(\theta_j)^2}{P_i(\theta_j)Q_i(\theta_j)}$$

เมื่อ i แทนข้อสอบข้อ $1, 2, \dots, n$

$$P_i'(\theta_j) = \frac{1.7 a_i (1 - c_i)}{e^{D a_i(\theta_j - b_i)} + 1 - e^{D a_i(\theta_j - b_i)}}$$

$$P_i(\theta_j) = \frac{c_i + e^{D a_i(\theta_j - b_i)}}{1 + e^{D a_i(\theta_j - b_i)}}$$

$$Q_i(\theta_j) = \frac{1 - c_i}{1 + e^{D a_i(\theta_j - b_i)}}$$

การศึกษาเปรียบเทียบคุณภาพของการทดสอบแบบซี เอ ที และการทดสอบแบบประเพณีนิยม

ลักษณะและปริมาณของประชากรและกลุ่มตัวอย่าง

ในการดำเนินการวิจัยระยะที่สองนี้ประชากรในการวิจัยได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 3 โรงเรียนสวนกุหลาบวิทยาลัยนนทบุรี และนักเรียนโรงเรียนปากเกร็ด จังหวัดนนทบุรี ที่กำลังศึกษาในภาคเรียนที่ 1 ปีการศึกษา 2531 จำนวน 24 ห้องเรียน จำนวน 1,194 คน

กลุ่มตัวอย่างได้มาโดยการสุ่มกลุ่มโดยใช้ห้องเรียนเป็นหน่วยการสุ่ม ได้กลุ่มตัวอย่างจำนวน 4 ห้องเรียน สุ่มนักเรียนได้ 120 คน ดังตารางที่ 7

ตารางที่ 7 จำนวนนักเรียนที่เป็นกลุ่มตัวอย่าง

โรงเรียน	จำนวน	
	ห้องเรียน	นักเรียน
ปากเกร็ด	2	60
สวนกุหลาบวิทยาลัยนนทบุรี	2	60
รวม	4	120

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยประกอบด้วย โปรแกรมการทดสอบแบบ ซี เอ ที และแบบสอบที่ใช้ในการทดสอบแบบประเพณีนิยม โดยมีลำดับขั้นตอนในการดำเนินการ ดังนี้

1. การสร้างโปรแกรมการทดสอบแบบ ซี เอ ที

1.1 การจัดเตรียมข้อสอบที่จะใช้เป็นกลุ่มข้อสอบเพื่อการทดสอบแบบ ซี เอ ที

1.1.1 คัดเลือกข้อสอบจากแบบสอบทั้ง 5 ฉบับ เพื่อสร้างเป็นกลุ่มข้อสอบ โดยพิจารณาจากผลการวิเคราะห์ข้อสอบ ตามที่วอร์ม (Warm 1978:52-53) เสนอไว้ว่าข้อสอบที่ดีควรมีค่าอำนาจจำแนก (a_1) มากกว่า .8 ค่าความยากของข้อสอบ (b_1) อยู่ในช่วง -2.5 ถึง 2.5 และค่าการเดาของข้อสอบ (c_1) น้อยกว่า .3 และยูริ (Urry 1977:196) กล่าวว่า ข้อสอบที่จัดไว้ในคลังข้อสอบตามความสามารถ (Ability Bank) ควรจะมีค่าอำนาจจำแนก (a_1) มากกว่า

.8 ค่าความยากของข้อสอบ (b_1) อยู่ในช่วง -2.0 ถึง 2.0 และค่าการเดาของข้อสอบ (c_1) น้อยกว่า .3 ผู้วิจัยจึงใช้เกณฑ์ดังกล่าวในการคัดเลือกข้อสอบแต่อย่างไรก็ตาม ผู้วิจัยต้องการให้ได้ข้อสอบที่มีความยากครอบคลุมผู้สอบที่มีความสามารถสูง จึงคัดเลือกข้อสอบที่มีความยากมากกว่า 2.5 ไปด้วย

ข้อสอบที่คัดเลือกไว้ครั้งหนึ่งมีค่าอำนาจจำแนกอยู่ในช่วง 0.80334 ถึง 2.00000 ค่าความยากของข้อสอบอยู่ในช่วง -0.90522 ถึง 3.61860 ค่าการเดาของข้อสอบอยู่ในช่วง 0.04421 ถึง 0.29886 ได้ข้อสอบจำนวน 361 ข้อ

1.1.2 ข้อสอบที่คัดเลือกไว้จำนวน 361 ข้อนี้ผู้วิจัยนำมาจัดเป็นชุดโดยกำหนดช่วงค่าความยากของข้อสอบ (b_1) ช่วงละ 0.5 โดยเริ่มจากค่าความยากปานกลาง ($b=0$) เป็นตัวกำหนด จัดชุดแบบสอบได้ 8 ช่วง แต่ละช่วงประกอบด้วยข้อสอบดังแสดงในตารางที่ 8

ตารางที่ 8 จำนวนข้อสอบในแต่ละช่วงความยากของข้อสอบ

ช่วงที่	ความยากของข้อสอบ (b_1)	จำนวนข้อสอบ
1.	-1.00 ถึง -0.74	8
2.	-0.75 ถึง -0.24	18
3.	-0.25 ถึง 0.25	27
4.	0.26 ถึง 0.75	43
5.	0.76 ถึง 1.25	102
6.	1.26 ถึง 1.75	107
7.	1.76 ถึง 2.25	41
8.	2.26 ขึ้นไป	15
	รวม	361

1.2 การสร้างและการออกแบบโปรแกรมการทดสอบแบบซี เอ ที วิธีการทดสอบแบบ ซี เอ ที เป็นวิธีการที่ให้นักเรียนนั่งหน้าจอคอมพิวเตอร์ คอมพิวเตอร์จะส่งข้อสอบและแสดงข้อสอบบนจอภาพ นักเรียนทำข้อสอบโดยการเลือกคำตอบที่ถูกต้องแล้วกดแป้นตัวเลขที่ตรงกับตัวเลือกที่ต้องการคอมพิวเตอร์จะตรวจข้อสอบแล้วส่งข้อสอบข้อต่อไปตามกฎการเลือกข้อสอบที่ผู้วิจัยได้มาแสดงบนจอภาพอีก คอมพิวเตอร์จะหยุดทำงานก็ต่อเมื่อพบกับเกณฑ์การยุติการสอบ สำหรับโปรแกรมคอมพิวเตอร์ที่ใช้ในการดำเนินการสอบในครั้งนี้ผู้วิจัยดำเนินการตามขั้นตอนดังต่อไปนี้

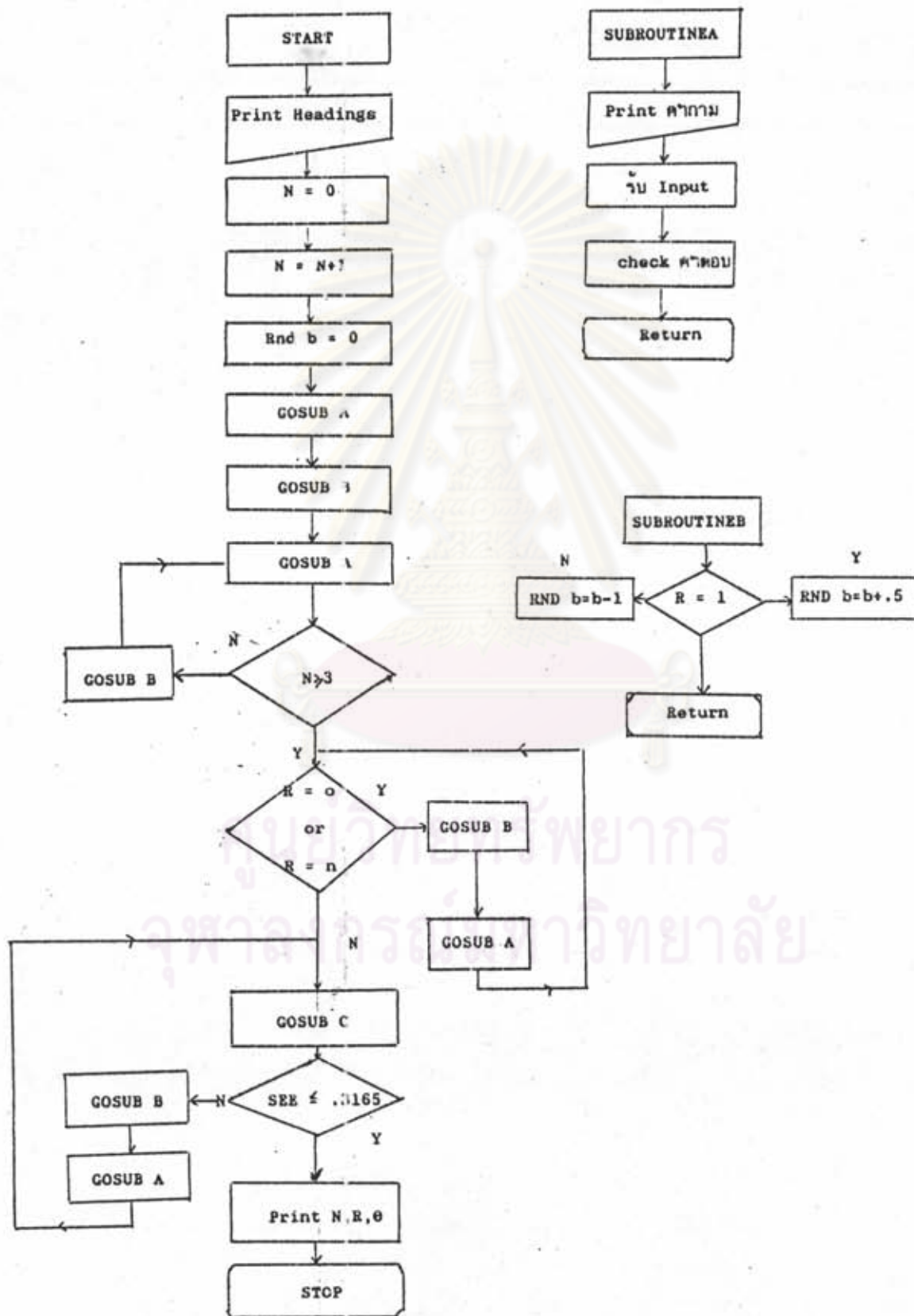
1.2.1 ศึกษาทำความเข้าใจกระบวนการทดสอบแบบ ซี เอ ที โดยละเอียด และดำเนินงานตามขั้นตอนดังต่อไปนี้

- 1) กลุ่มข้อสอบประกอบด้วยข้อสอบที่มีค่าพารามิเตอร์ 3 พารามิเตอร์
- 2) ระดับในการเริ่มต้นทำข้อสอบ ผู้วิจัยกำหนดให้ผู้สอบเริ่มทำข้อสอบที่มีความยากปานกลาง ซึ่งเป็นความยากอยู่ในช่วง -0.25 ถึง 0.25
- 3) กฎในการคัดเลือกข้อสอบ ผู้วิจัยใช้วิธีการจับคู่ค่าความยาก (b_i) กับค่าความสามารถ (θ) โดยใช้หลักว่า ถ้าผู้สอบตอบข้อสอบถูกต้องคอมพิวเตอร์ก็จะสุ่มข้อสอบที่มีค่าความยาก (b_i) เพิ่มขึ้น 0.5 ถ้าตอบผิดคอมพิวเตอร์ก็จะสุ่มข้อสอบที่มีค่าความยากลดลง 1.0
- 4) วิธีการให้คะแนนหรือการประมาณค่าความสามารถ ใช้วิธีการประมาณค่าแบบความเป็นไปได้สูงสุดอย่างมีเงื่อนไข (Conditional Maximum Likelihood Estimation) (Hambleton and Swaminathan 1985:81-88)
- 5) เกณฑ์การยุติการสอบ ผู้วิจัยใช้เกณฑ์ดังต่อไปนี้ กล่าวคือ ตามหลักการของการประมาณค่าความสามารถแบบความเป็นไปได้สูงสุด (Maximum Likelihood Estimation) จะไม่สามารถประมาณค่าความสามารถได้ตามรูปแบบการตอบของผู้สอบไม่เป็น 0, 1 ดังนั้น ผู้วิจัยจึงใช้เกณฑ์ว่าหากผู้สอบตอบข้อสอบครบ 8 ข้อแล้วถูกข้อหรือผิดทุกข้อ ให้ยุติการสอบ โดยถือว่าข้อสอบชุดนี้ไม่เหมาะสมกับความสามารถของผู้สอบคนนั้น ๆ เกณฑ์อีกข้อหนึ่งคือ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ $.3165$ และหากค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบไม่เท่ากับหรือไม่น้อยกว่า $.3165$ แต่ผู้สอบทำข้อสอบครบ 40 ข้อให้ยุติการทดสอบ

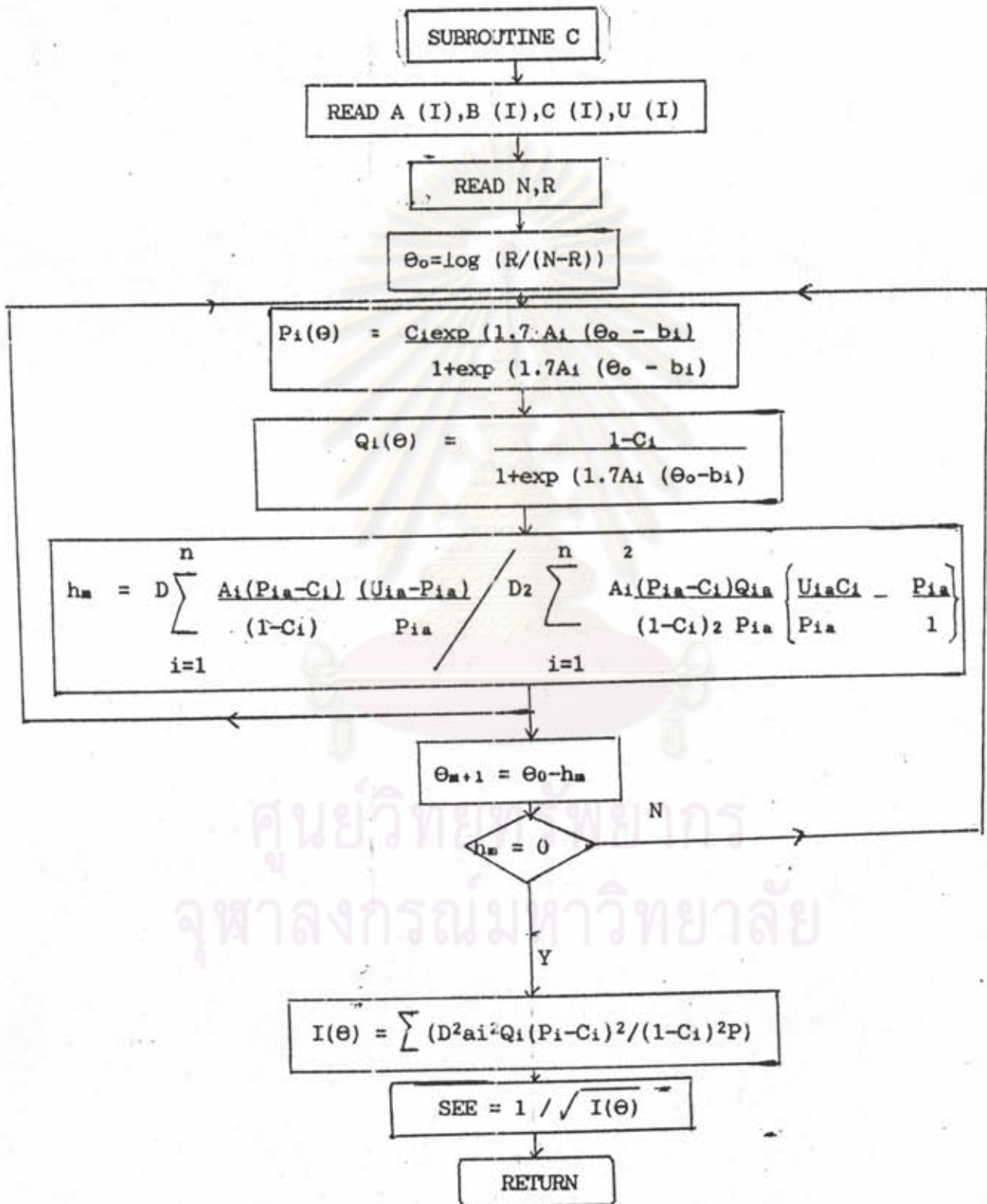
1.2.2 นำขั้นตอนต่าง ๆ ในการดำเนินการสอบแบบ ซี เอ ที มาเขียนเป็นผังงาน (Flowchart) ดังต่อไปนี้

ศูนย์วิจัยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ผังงาน แสดงลำดับขั้นตอนการดำเนินการสร้างและออกแบบโปรแกรมการทดสอบโดยใช้ไมโครคอมพิวเตอร์



ผังงาน (ต่อ)





1.2.3 ในการสร้างโปรแกรมได้รับความอนุเคราะห์จากนายนิรันดร์ สาโรธา นิสิตปริญญาโทบัณฑิต ภาควิชาสถิติศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ช่วยในการเขียนโปรแกรม โดยที่ผู้วิจัยนำฟังก์ชันที่เขียนขึ้นไปให้ผู้เขียนโปรแกรมศึกษาและทำความเข้าใจขั้นตอน และกระบวนการในการทดสอบแบบ ซี เอ ที

1.2.4 บันทึกข้อสอบทั้งตัวคำถาม ตัวเลือก เฉลย และค่าพารามิเตอร์ของ ข้อสอบลงในแผ่นบันทึกข้อมูลโดยเก็บไว้ในไฟล์ชื่อ 1, 2, 3, 4, 5, 6, 7 และ 8 ตามช่วง ความยากที่กำหนดไว้ ตรวจสอบความถูกต้องของข้อมูลบันทึกไว้

1.2.5 นำโปรแกรมที่สร้างเสร็จมาทดลอง ตรวจสอบความถูกต้อง เพื่อหา ข้อบกพร่องต่าง ๆ หากพบข้อผิดพลาดและข้อบกพร่องต่าง ๆ ที่เกิดขึ้น จนกระทั่งได้โปรแกรมที่สมบูรณ์ จึงนำโปรแกรมที่สร้างเสร็จไปใช้ทดสอบกับกลุ่มตัวอย่าง

2. การสร้างแบบสอบเพื่อทดสอบแบบประเพณีนิยม ผู้วิจัยสุ่มข้อสอบจากกลุ่มข้อสอบที่ คัดเลือกไว้จำนวน 361 ข้อ โดยวิธีการสุ่มอย่างง่ายเพื่อให้ได้ข้อสอบจำนวน 40 ข้อ ข้อสอบชุดนี้ มีค่าอำนาจจำแนกกระจายอยู่ในช่วง 0.94069 ถึง 2.00000 ค่าความยากกระจายอยู่ในช่วง -0.76488 ถึง 2.97884 ค่าการเลกากระจายอยู่ในช่วง 0.03462 ถึง 0.28620

การเก็บรวบรวมข้อมูล

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลระยะที่สอง โดยนำหนังสือขอความร่วมมือจากทางบัณฑิต วิทยาลัยไปยังสำนักงานศึกษาธิการเขต เขตการศึกษา 1 เพื่อขอความอนุเคราะห์จากโรงเรียนที่เป็น กลุ่มตัวอย่าง เพื่อขออนุญาตใช้เครื่องไมโครคอมพิวเตอร์ขนาด 16 บิต ของโรงเรียนที่เป็น กลุ่มตัวอย่าง และขอความอนุเคราะห์จากทางโรงเรียนให้จัดนักเรียนที่เป็นกลุ่มตัวอย่าง โดยให้ในวัน เวลา และสถานที่ทำการทดสอบ หลังจากนั้นผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลตามขั้นตอน ดังต่อไปนี้

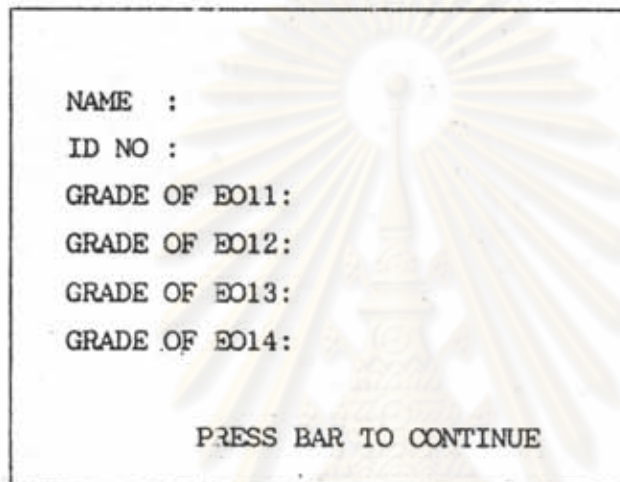
1. เนื่องจากนักเรียนที่เป็นกลุ่มตัวอย่างส่วนมากไม่มีความรู้ในการใช้คอมพิวเตอร์มาก่อน ดังนั้น เมื่อนักเรียนเข้ามาในห้องคอมพิวเตอร์ซึ่งเป็นห้องที่ใช้ดำเนินการสอบ ผู้วิจัยจึงอธิบายวิธีการใน การดำเนินการสอบดังนี้

1.1 นักเรียนทุกคนจะต้องทำข้อสอบให้ครบ 2 ชุด ชุดหนึ่งเป็นการทดสอบแบบ ประเพณีนิยม คือ ให้นักเรียนอ่านและเขียนตอบในกระดาษคำตอบ อีกชุดหนึ่งเป็นการทดสอบโดยการ ให้ผู้สอบนั่งทดสอบหน้าจอคอมพิวเตอร์

1.2 อธิบายขั้นตอนในการทดสอบกับเครื่องไมโครคอมพิวเตอร์ดังนี้

1.2.1 บนจอภาพจะปรากฏข้อความที่สอบถามเกี่ยวกับข้อมูลส่วนตัวของนักเรียน
ให้นักเรียนบันทึกข้อมูลส่วนตัวของนักเรียนแล้วกดแป้น ENTER ทุกครั้ง เมื่อบันทึกครบแล้วให้กดแคร่ยาว
1 ครั้ง ดังภาพที่ 4

ภาพที่ 4 การบันทึกประวัติ ข้อมูลส่วนตัวของนักเรียน



NAME :

ID NO :

GRADE OF EO11:

GRADE OF EO12:

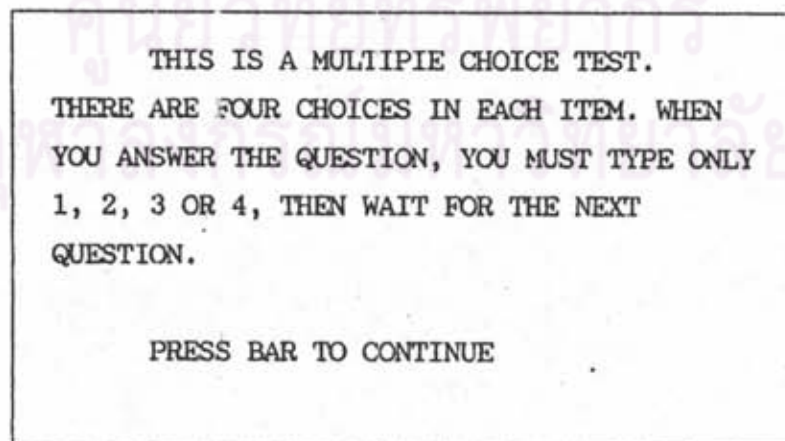
GRADE OF EO13:

GRADE OF EO14:

PRESS BAR TO CONTINUE

1.2.2 หลังจากกดแคร่ยาว 1 ครั้ง จะปรากฏคำอธิบายการตอบข้อสอบ
เช่นนี้ ดังต่อไปนี้

ภาพที่ 5 คำอธิบายในการตอบข้อสอบ



THIS IS A MULTIPLE CHOICE TEST.
THERE ARE FOUR CHOICES IN EACH ITEM. WHEN
YOU ANSWER THE QUESTION, YOU MUST TYPE ONLY
1, 2, 3 OR 4, THEN WAIT FOR THE NEXT
QUESTION.

PRESS BAR TO CONTINUE

1.2.3 จากนั้นผู้วิจัยอธิบายเพิ่มเติมอีกครั้งหนึ่งว่า การตอบข้อสอบชนิดนี้ให้ ผู้สอบกดเฉพาะตัวเลข 1, 2, 3 หรือ 4 ซึ่งตรงกับคำตอบที่นักเรียนคิดว่าถูกต้องที่สุดเพียง 1 ครั้ง เท่านั้น แล้วคอมพิวเตอร์จะสุ่มข้อสอบข้อต่อไปให้นักเรียนทำโดยที่นักเรียนไม่ต้องกดตัวเลขซ้ำ หากนักเรียนกดตัวเลขซ้ำ ตัวเลขนั้นจะกลายเป็นคำตอบของข้อต่อไป นักเรียนแต่ละคนจะได้รับข้อสอบ ไม่เหมือนกัน เนื่องจากมีข้อสอบเก็บไว้ในแผ่นบันทึกข้อมูลจำนวนมาก แต่ทุกคนจะทำข้อสอบประมาณ 20 ข้อเท่านั้น ก็จะสิ้นสุดการทดสอบ โดยให้นักเรียนสังเกตข้อความบนจอภาพจะมีคำว่า SEE ถ้า SEE มีค่าน้อยกว่า หรือเท่ากับ .3165 ซึ่งเป็นเกณฑ์ผู้วิจัยตั้งไว้ก็จะสิ้นสุดการทดสอบ แต่อย่างไรก็ตาม ถ้านักเรียนที่ทำข้อสอบโดยใช้การเดามากกว่าความสามารถจริงของนักเรียนแล้วนักเรียนก็อาจจะต้อง ทำข้อสอบถึง 40 ข้อ เมื่อสิ้นสุดการทดสอบแล้วนักเรียนก็จะทราบว่าตนทำข้อสอบทั้งหมดกี่ข้อ ถกข้อ ความสามารถของนักเรียนเท่ากับเท่าไร ความสามารถของนักเรียนสามารถเปรียบเทียบกันได้ ถึงแม้ว่า นักเรียนจะได้รับข้อสอบต่างกัน หลังจากให้นักเรียนได้รับคำอธิบายเพิ่มเติมแล้ว นักเรียนที่เข้าสอบเข้าใจ วิธีการในการทดสอบเป็นอย่างดี และมีความสนใจเข้ารับการทดสอบมากและตั้งใจทำข้อสอบอย่างเต็ม ความสามารถ

2. ลำดับที่ในการทดสอบแบบประเพณีนิยมและแบบ ซี เอ ที เป็นไปโดยสุ่ม

3. หลังจากให้นักเรียนทดสอบกับเครื่องไมโครคอมพิวเตอร์แล้ว เครื่องจะบันทึกข้อมูล การตอบของนักเรียนไว้ใน Drive A ในไฟล์ข้อมูลชื่อ DATA เมื่อสิ้นสุดการดำเนินการสอบแล้ว ผู้วิจัยรวบรวมข้อมูลต่าง ๆ เพื่อนำมาวิเคราะห์ต่อไป

4. ขอเก็บตัวซ้ำ ๑011 - ๑014 ของนักเรียนที่เข้ารับการทดสอบจากฝ่ายทะเบียนวัดผล ของโรงเรียนที่เป็นกลุ่มตัวอย่างเพื่อตรวจสอบความถูกต้องของเกรดที่นักเรียนบันทึกไว้ตอนเริ่มเข้ารับการทดสอบ

5. จากการทดสอบกับเครื่องไมโครคอมพิวเตอร์ ผู้วิจัยตั้งเกณฑ์การยุติการทดสอบไว้ 3 ข้อ กล่าวคือ หากผู้สอบบรรจุเกณฑ์ข้อใดข้อหนึ่งในสามข้อต่อไปนี้จะถือว่า การทดสอบสิ้นสุดลง

5.1 นักเรียนทำข้อสอบครบ 40 ข้อ

5.2 นักเรียนทำข้อสอบ 8 ข้อแรกถูกต้องทั้งหมดหรือผิดทั้งหมด

5.3 ความคลาดเคลื่อนมาตรฐานในการทำข้อสอบของนักเรียนถึงเกณฑ์ที่กำหนดไว้คือ มีค่าน้อยกว่าหรือเท่ากับ .3165

ผู้วิจัยจึงคัดเลือกเรียนที่ทำข้อสอบครบ 40 ข้อ แล้วยังไม่สามารถหาความสามารถและความคลาดเคลื่อนมาตรฐานในการประมาณค่าได้และนักเรียนที่ทำข้อสอบครบ 8 ข้อ แล้วยปรากฏว่าทำข้อสอบถูกหมดหรือผิดหมดซึ่งแสดงว่าข้อสอบชุดนี้ไม่เหมาะสมกับความสามารถของเขาจึงคัดออกไม่นำมาวิเคราะห์

6. นำข้อมูลที่ได้จากการทดสอบแบบประเพณีนิยมมาตรวจให้คะแนนแบบผิดได้ 0 ถูกได้ 1 เพื่อนำไปวิเคราะห์โดยใช้โปรแกรม Logist 5 Version 2.5 โดยใช้คำสั่งเพิ่มเติมสำหรับข้อมูลที่มิตค่าพารามิเตอร์ของข้อสอบแล้ว ให้ประมาณเฉพาะค่าความสามารถของผู้สอบเท่านั้น

การวิเคราะห์ข้อมูล

1. การคำนวณค่าความสามารถของผู้สอบ

1.1 การคำนวณค่าความสามารถของผู้สอบจากการทดสอบแบบประเพณีนิยมจำนวน 40 ข้อ ผู้วิจัยใช้โปรแกรม Logist 5 Version 2.5 ใช้คำสั่งเพิ่มเติมสำหรับข้อสอบที่มีค่าพารามิเตอร์ของข้อสอบอยู่แล้ว โดยให้ประมาณค่าเฉพาะค่าความสามารถของผู้สอบ (Wingersky, M.S., et al. 1982:24-36)

1.2 การคำนวณค่าความสามารถของผู้สอบจากการทดสอบแบบ ซี เอ ที ผู้วิจัยสร้างโปรแกรมการทดสอบตามสูตรการคำนวณค่าความสามารถของผู้สอบอย่างมีเงื่อนไขโดยวิธีการประมาณค่าแบบความเป็นไปได้สูงสุด (Hambleton and Swaminathan 1985:83) โดยใช้สูตรในการคำนวณ ดังนี้

ขั้นที่ 1 ประมาณค่าความสามารถเริ่มต้นของผู้สอบ โดยใช้สูตร

$$\theta_0 = \ln (r_a / (n - r_a))$$

เมื่อ r_a = จำนวนข้อของการตอบถูก

n = จำนวนข้อที่ทำ

ขั้นที่ 2 คำนวณค่า P_1 , Q_1 , h_m ดังนี้

1) คำนวณค่า P_1 , Q_1 โดยใช้สูตร

$$P_i = \frac{D a_i (\theta_0 - b_i)}{c_i + e} \cdot \frac{D a_i (\theta_0 - b_i)}{1 + e}$$

$$Q_i = \frac{1 - c_i}{D a_i (\theta_0 - b_i) \cdot (1 + e)}$$

2) คำนวณค่า correction factor

$$h_m = D \sum_{i=1}^n \frac{a_i (P_{ia} - c_i) (U_{ia} - P_{ia})}{(1 - c_i) P_{ia}} \bigg/ D^2 \sum_{i=1}^n \frac{a^2 (P_{ia} - c_i) Q_{ia} (U_{ia} c_i - P_{ia})}{(1 - c_i)^2 P_{ia} P_{ia}}$$

3) คำนวณค่า θ ใหม่ โดยใช้สูตร

$$\theta_m + 1 = \theta_0 - h_m$$

4) คำนวณค่าในข้อ 1, 2 และ 3 ซ้ำ ๆ จนกระทั่ง $h_m < .001$

2. การคำนวณค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถโดยใช้สูตร

$$SEE = 1 / \sqrt{I(\theta)}$$

เมื่อ $I(\theta)$ สารสนเทศของแบบสอ
สำหรับข้อสอบที่มี 3 พารามิเตอร์แล้ว

$$I(\theta) = \sum_{i=1}^n \frac{D^2 a^2 Q_i (P_i - c_i)^2}{(1 - c_i)^2 P_i}$$

(Hambleton and Swaminathan 1985:91)

3. คำนวณค่าสหสัมพันธ์ระหว่างการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบประเพณีนิยม การทดสอบแบบ ซี เอ ที และเกรดเฉลี่ยโดยใช้โปรแกรมสำเร็จรูป SPSS* โดยใช้คำสั่ง PEARSON CORR GPA THETA1 THETA2

4. ทดสอบสมมติฐานความแตกต่างระหว่างค่าความคลาดเคลื่อนมาตรฐานโดยการประมาณค่าความสามารถ โดยใช้สูตร t - test (Ferguson,G.A 1981:180)

$$t = \frac{\sum D}{\sqrt{[N \sum D^2 - (\sum D)^2]/(N - 1)}}, df = n-1$$

D = ความแตกต่างระหว่างค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถจากการทดสอบทั้งสองวิธี

N = จำนวนผู้เข้าสอบ

5. ทดสอบความแตกต่างของค่าความตรงของการทดสอบในการประเมินคุณภาพของการทดสอบแบบประเพณีนิยมและแบบ ซี เอ ที ในการวิจัยครั้งนี้ผู้วิจัยใช้เกณฑ์ความตรงซึ่งเกณฑ์สัมพัทธ์ซึ่งค่าความตรงนี้เป็นค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนความสามารถจากวิธีการทดสอบแต่ละวิธีกับคะแนนเกณฑ์คือ เกรดเฉลี่ยวิชาภาษาอังกฤษ ดังนั้น ผู้วิจัยจึงใช้สูตร t - test เพื่อหาความแตกต่างระหว่างค่าสัมประสิทธิ์ความตรงทั้งสองค่านี้ ดังนี้ (Hotelling อ้างถึงใน Guilford 1985:164)

$$t_{ar} = (r_{12} - r_{13}) \sqrt{\frac{(N-3) (1+r_{23})}{2(1-r_{23} - r_{12}^2 - r_{13}^2 + 2r_{23}r_{12}r_{13})}}, df=N-3$$

เมื่อ r_{12} = ค่าสหสัมพันธ์ระหว่างเกรดเฉลี่ยกับคะแนนความสามารถ (θ) ที่ได้จากการทดสอบแบบ ซี เอ ที

r_{13} = ค่าสหสัมพันธ์ระหว่างเกรดเฉลี่ยกับคะแนนความสามารถ (๑)
ที่ได้จากการทดสอบแบบประเพณีนิยม

r_{23} = ค่าสหสัมพันธ์ระหว่างคะแนนความสามารถที่ได้จากการทดสอบแบบ
ซี เอ ที และแบบประเพณีนิยม

N = จำนวนนักเรียนที่เป็นกลุ่มตัวอย่าง



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย