บทที่ 4

ผลการวิเคราะห์ข้อมูล

ในการวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาการออกแบบโค้งแสดงสารสนเทศของแบบสอบผลสัมฤทธิ์ด้วยการจัดข้อสอบให้มีความยากง่ายเหมาะสมกับระดับความสามารถของนักเรียน ผลการวิเคราะห์ข้อมูลกระทำเพื่อวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบ อันจะนำมาซึ่งการออกแบบแบบสอบที่ผู้วิจัยต้องการ ผู้วิจัยได้ดำเนินการตามกระบวนการสร้างแบบสอบผลสัมฤทธิ์ที่มีลักษณะเป็นแบบสอบชนิดเลือกตอบ 4 ตัวเลือก จากขอบเขตเนื้อหาหลัก (Domain) 7 ขอบเขตเนื้อหา แยกเป็นเนื้อหาย่อย ๆ (Sub Domain) 19 เนื้อหาย่อย ครอบคลุมจุดประสงค์การเรียนรู้ที่วัด 29 จุดประสงค์ ได้ข้อสอบที่ใช้ในการวิจัยจำนวน 120 ข้อ

ผลการวิเคราะห์ข้อมูลผู้วิจัยเสนอเป็น 3 ตอนดังนี้

ตอนที่ 1 ผลการใช้รูปแบบของราสช์วิเคราะห์แบบสอบผลสัมฤทธิ์

1. ผลการตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว
2. ผลการวิเคราะห์แบบสอบผลสัมฤทธิ์ด้วยโปรแกรมไบคาล

ตอนที่ 2 การออกแบบแบบสอบ

1. การเทียบค่าความสามารถของนักเรียนที่สอบแบบสอบทั้ง 2 ฉบับ
2. ผลการคัดเลือกข้อสอบเพื่อออกแบบแบบสอบตามระดับความสามารถของนักเรียนทั้ง 3 ระดับ คือ ความสามารถสูง ความสามารถปานกลาง และความสามารถต่ำ
3. ผลการวิเคราะห์แบบสอบผลสัมฤทธิ์ตามกลุ่มความสามารถของนักเรียนทั้ง 3 ระดับ

ตอนที่ 3 ผลการตรวจสอบความแม่นยำ (precision) ของแบบสอบผลสัมฤทธิ์

1. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของนักเรียนทั้ง 3 ระดับ
2. เปรียบเทียบค่าความเที่ยงของแบบสอบที่วัดความสามารถของนักเรียนทั้ง 3 ระดับ

รายละเอียดของผลการวิเคราะห์ข้อมูลมีดังต่อไปนี้

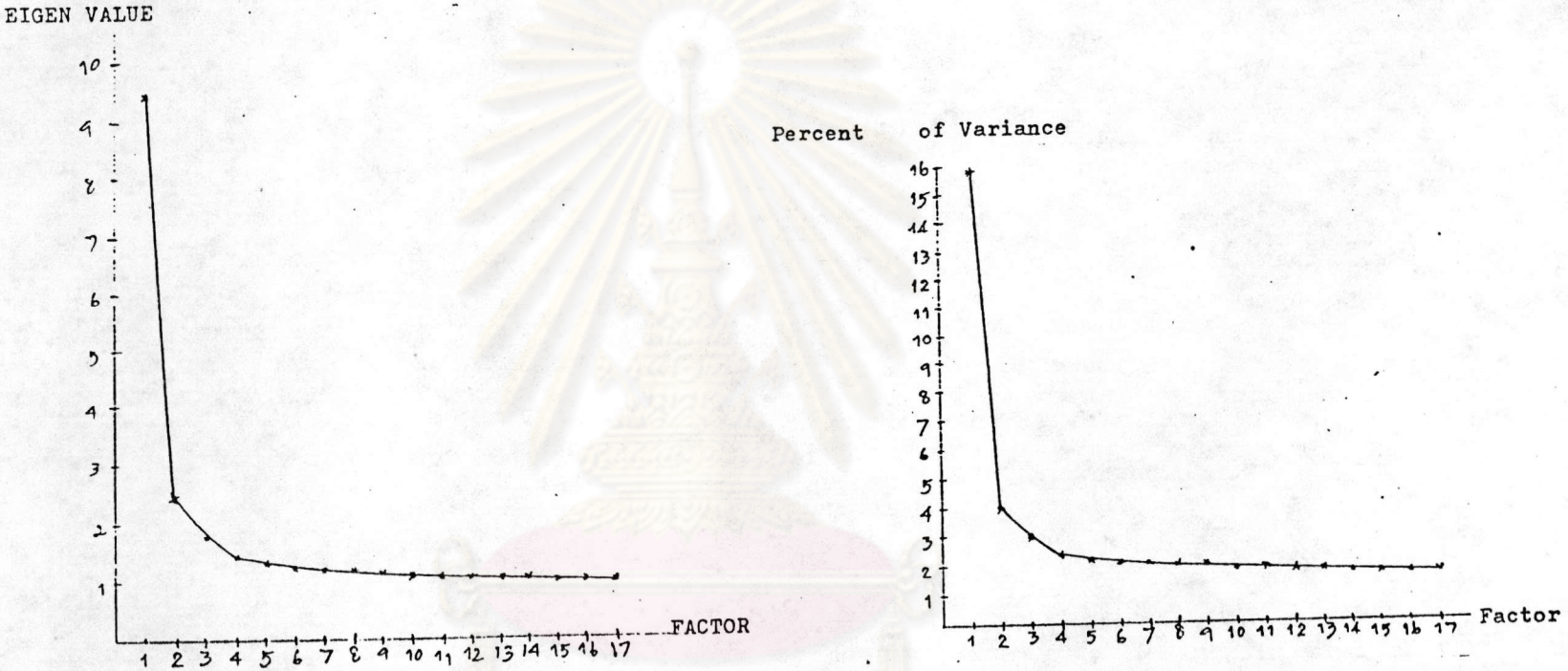
ตอนที่ 1 ผลการใช้รูปแบบของราสช์วิเคราะห์แบบสอบผลสัมฤทธิ์

1. ผลการตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว (Unidimensionality)

เนื่องจากคุณสมบัติการวัดเพียงมิติเดียวเป็นข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ (Item Response Theory)ในการวิเคราะห์ข้อมูลของแบบสอบผลสัมฤทธิ์และจากการวิจัยของนักวัดผลหลายท่าน อาทิ Soriyan(1971) Chen(1981) Hutten(1981) และ Bloomquist(1984) พบว่าความเหมาะสมกับรูปแบบของราสช์มีความสัมพันธ์กับคุณสมบัติการวัดเพียงมิติเดียวอย่างมีนัยสำคัญ ดังนั้นผู้วิจัยจึงนำแบบสอบทั้ง 2 ฉบับจำนวน 120 ข้อ ตรวจสอบด้วยการวิเคราะห์ตัวประกอบ (Factor Analysis) ซึ่งปรากฏผลดังตารางที่ 32 และตารางที่ 33 ในภาคผนวก ซึ่งสามารถนำค่าไอเกนและค่า Percent of variance มาเขียนเป็นภาพแสดงให้เห็นความแตกต่างของค่าไอเกนและค่า Percent of variance ของแต่ละตัวประกอบซึ่งแสดงได้ด้วยแผนภาพที่ 12 และแผนภาพที่ 13

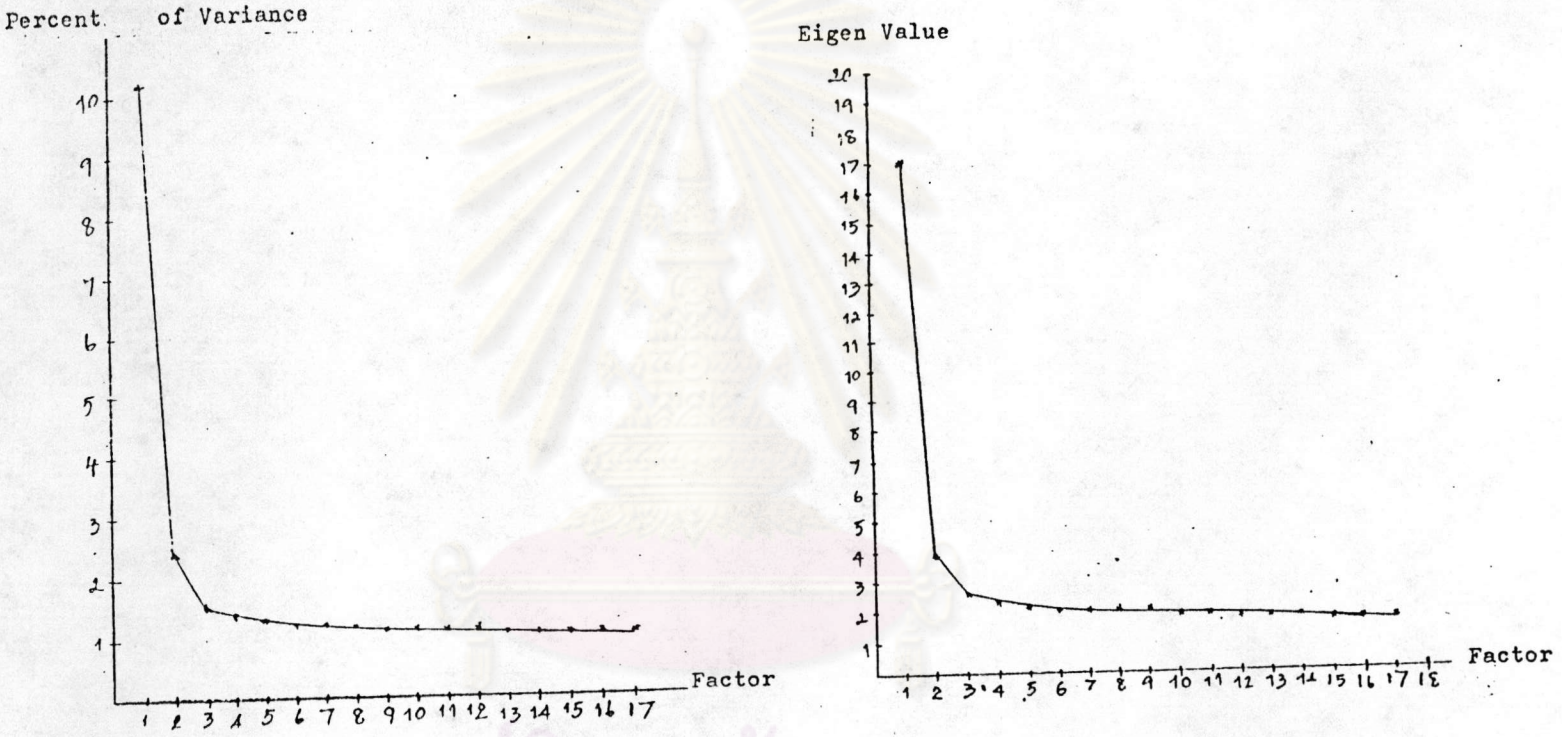
จากตารางที่ 32 ตารางที่ 33 แผนภาพที่ 12 และแผนภาพที่ 13 พบว่าเมื่อนำแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ทั้ง 2 ฉบับ ๆละ 60 ข้อ ทำการวิเคราะห์ตัวประกอบจะได้ตัวประกอบที่มีค่าไอเกนเกิน 1.00 อยู่ 17 ตัวประกอบ ซึ่งถ้าพิจารณาค่าไอเกนและค่า Percent of variance ของตัวประกอบที่ 1 ของแบบสอบทั้ง 2 ฉบับ จะเห็นว่ามีค่าไอเกนและค่า Percent of variance ของตัวประกอบที่ 1 สูงกว่าตัวประกอบที่ 2 อย่างมาก แต่ในตัวประกอบที่ 2 จะมีค่าไอเกนและค่า Percent of variance สูงกว่าตัวประกอบในลำดับต่อไปเพียงเล็กน้อย โดยมีค่าไอเกนของตัวประกอบที่ 1 มายังตัวประกอบที่ 2 มีค่าลดลงอย่างรวดเร็ว ในขณะที่ตัวประกอบที่ 2 มายังตัวประกอบที่ 3 จนกระทั่งถึงตัวประกอบที่ 17 มีค่าไอเกนลดลงทีละน้อย ซึ่งลักษณะความแตกต่างเช่นนี้สอดคล้องกับข้อเสนอของ Lord (Lord 1980: 21) ที่ว่าลักษณะเช่นนี้พอจะเชื่อถือได้ว่าแบบสอบมีการวัดเพียงมิติเดียว ดังนั้นแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ทั้ง 2 ฉบับจึงน่าจะเชื่อถือได้ว่ามีคุณสมบัติการวัดเพียงมิติเดียว

แผนภาพที่ 12 แสดงค่าไอเกนและค่า Percent of variance ของตัวประกอบที่มีค่าไอเกนเกิน 1.00 ของแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ ฉบับที่ 1 จำนวน 60 ข้อ



จากแผนภาพจะเห็นว่าค่าไอเกนและค่า Percent of variance ของตัวประกอบที่ 1 มีค่าสูงกว่าตัวประกอบที่ 2 อย่างมาก สังเกตได้ว่าจากตัวประกอบที่ 1 มายังตัวประกอบที่ 2 ค่าจะลดลงอย่างรวดเร็ว แต่ในตัวประกอบที่ 2 มายังตัวประกอบที่ 3 และมายังตัวประกอบอื่น ๆ จนกระทั่งตัวประกอบที่ 17 ตามลำดับ ค่าไอเกนและค่า Percent of variance จะลดลงเพียงเล็กน้อยเท่านั้น

แผนภาพที่ 13 แสดงค่าไอเกนและค่า Percent of variance ของตัวประกอบที่มีค่าไอเกนเกิน 1.00 ของแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ ฉบับที่ 2 จำนวน 60 ข้อ



จากแผนภาพจะเห็นว่าค่าไอเกนและค่า Percent of variance ของตัวประกอบที่ 1 มีค่าสูงกว่าตัวประกอบที่ 2 อย่างมาก สังเกตได้ว่าจากตัวประกอบที่ 1 มายังตัวประกอบที่ 2 ค่าจะลดลงอย่างรวดเร็ว แต่ในตัวประกอบที่ 2 มายังตัวประกอบที่ 3 และมายังตัวประกอบอื่น ๆ จนกระทั่งตัวประกอบที่ 17 ตามลำดับ ค่าไอเกนและค่า Percent of variance จะลดลงเพียงเล็กน้อยเท่านั้น

ตารางที่ 7 แสดงค่าพารามิเตอร์ของข้อสอบและความเหมาะสมกับรูปแบบของราสช์ของแบบสอบผลสัมฤทธิ์ ฉบับที่ 1 จำนวน 60 ข้อ

ข้อสอบข้อที่	ค่าความยาก	ค่าความคลาดเคลื่อนมาตรฐาน	ค่าอำนาจจำแนก	t-test	ข้อสอบข้อที่	ค่าความยาก	ค่าความคลาดเคลื่อนมาตรฐาน	ค่าอำนาจจำแนก	t-test
1	0.587	0.058	0.95	0.49	31	-1.085	0.068	1.27	-1.97
2	0.998	0.060	1.26	-2.15	32	-0.314	0.060	1.01	-0.01
3	0.378	0.058	1.10	-0.64	33	0.014	0.058	1.06	-0.38
4	0.044	0.058	0.85	1.23	34	1.646	0.066	1.04	-1.09
5	0.054	0.058	0.52	3.25**	35	0.557	0.058	1.15	-1.07
6	-0.124	0.059	1.08	-0.44	36	0.604	0.058	-0.20	8.20**
7	1.195	0.061	0.78	1.05	37	0.024	0.058	0.47	3.21**
8	1.394	0.063	0.89	0.11	38	1.103	0.061	0.21	5.80**
9	-0.828	0.064	1.28	-2.27	39	1.553	0.065	0.69	1.46
10	-1.263	0.070	1.15	-1.48	40	-0.670	0.063	1.29	-2.16
11	-1.085	0.068	1.40	-3.12**	41	-0.772	0.064	1.45	-3.08**
12	-1.389	0.073	1.50	-3.37**	42	0.190	0.058	0.93	0.66
13	-0.736	0.063	1.27	-2.09	43	-0.040	0.058	1.17	-1.42
14	-0.902	0.065	1.31	-2.43	44	1.073	0.061	0.67	2.05
15	-0.915	0.065	1.28	-2.15	45	1.316	0.063	0.62	2.08
16	1.495	0.064	-0.30	8.40**	46	0.325	0.058	1.08	-0.64
17	-0.006	0.058	1.11	-0.84	47	0.183	0.058	0.98	0.23
18	0.732	0.059	1.08	-0.83	48	-1.080	0.068	1.12	-0.50
19	-1.283	0.071	1.17	-1.50	49	0.886	0.059	1.03	-0.34
20	-0.304	0.060	0.93	0.59	50	-1.416	0.073	1.34	-1.97
21	-1.210	0.070	1.29	-2.22	51	-1.139	0.068	1.12	-0.91
22	0.137	0.058	1.04	-0.20	52	-0.114	0.059	1.20	-1.27
23	-0.020	0.058	1.03	-0.22	53	0.322	0.058	1.06	-0.40
24	0.368	0.058	1.12	-0.98	54	0.438	0.058	1.27	-2.18
25	-0.311	0.060	1.03	-0.22	55	0.772	0.059	0.83	1.01
26	-2.059	0.089	1.38	-2.07	56	0.286	0.058	1.03	-0.02
27	-1.559	0.076	1.38	-2.28	57	0.769	0.059	1.35	-2.66**
28	-0.527	0.061	1.40	-2.28	58	0.071	0.058	0.94	0.47
29	1.766	0.068	0.82	0.79	59	-0.200	0.059	1.11	-0.74
30	-0.321	0.060	1.21	-1.60	60	-0.094	0.059	1.20	-1.04

** $P < .01$

ตารางที่ 8 แสดงค่าพารามิเตอร์ของข้อสอบและความเหมาะสมกับรูปแบบของราสช์ของแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ ฉบับที่ 2 จำนวน 60 ข้อ

ข้อสอบข้อที่	ค่าความยาก	ค่าความคลาดเคลื่อนมาตรฐาน	ค่าอำนาจจำแนก	t-test	ข้อสอบข้อที่	ค่าความยาก	ค่าความคลาดเคลื่อนมาตรฐาน	ค่าอำนาจจำแนก	t-test
1	0.688	0.059	1.13	-1.13	31	-0.293	0.060	1.38	-2.83*
2	0.369	0.058	1.02	-1.02	32	-0.562	0.061	1.26	-1.94
3	0.299	0.058	1.00	0.07	33	-0.667	0.062	1.51	-3.82*
4	0.346	0.058	0.93	0.44	34	0.934	0.060	0.93	0.27
5	1.276	0.063	0.88	-0.20	35	0.332	0.058	0.49	3.92*
6	1.179	0.062	0.05	6.20**	36	-0.898	0.064	1.20	-1.77
7	0.275	0.058	0.61	3.08**	37	1.225	0.063	0.69	1.75
8	-0.960	0.065	1.21	-1.11	38	-0.438	0.060	1.40	-2.95*
9	-1.275	0.069	1.35	-2.58	39	-0.304	0.060	1.46	-3.66*
10	-0.882	0.064	1.24	-2.04	40	-0.973	0.065	1.33	-2.50
11	-1.623	0.076	1.54	-3.39**	41	0.771	0.060	1.09	-1.08
12	-0.842	0.064	1.21	-1.80	42	-0.870	0.064	1.28	-2.03
13	-0.763	0.063	1.37	-2.76	43	0.510	0.059	0.75	1.87
14	0.232	0.058	1.08	-0.59	44	1.487	0.065	0.51	2.67*
15	-0.489	0.061	1.27	-2.05	45	0.582	0.059	0.97	0.26
16	0.108	0.058	1.37	-2.81**	46	0.978	0.061	0.43	3.86*
17	-1.481	0.073	1.45	-2.83**	47	-0.183	0.059	1.28	-2.27
18	1.156	0.062	-0.29	9.35**	48	-0.328	0.060	1.21	-1.62
19	0.558	0.059	0.67	2.44	49	0.473	0.059	1.21	-1.89
20	0.372	0.058	0.72	2.27	50	0.490	0.059	1.19	-1.78
21	0.650	0.059	0.97	0.24	51	-1.962	0.084	1.21	-1.46
22	0.678	0.059	0.78	1.57	52	-1.556	0.074	1.41	-2.59*
23	0.456	0.059	1.03	-0.02	53	-1.214	0.068	1.21	-1.52
24	1.291	0.063	0.77	1.22	54	0.490	0.059	0.76	1.59
25	-0.485	0.061	1.02	-0.12	55	-0.674	0.062	1.19	-1.41
26	-0.152	0.059	1.09	-0.71	56	0.007	0.059	1.15	-1.08
27	-1.475	0.073	1.40	-2.73**	57	0.609	0.059	0.66	2.48*
28	-1.423	0.072	1.51	-3.34**	58	0.470	0.059	0.85	1.26
29	2.072	0.074	0.63	1.27	59	1.040	0.061	0.71	2.32*
30	-0.307	0.060	1.45	-3.21**	60	0.674	0.059	0.85	1.02

** $P < .01$

จากตารางที่ 7 และตารางที่ 8 จะเห็นว่าแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 มีความเหมาะสม (fit)กับรูปแบบของราสช์จากค่าสถิติทดสอบที (t-test) ที่มีนัยสำคัญทางสถิติที่ระดับ .01 ดังนี้

แบบสอบฉบับที่ 1 ข้อสอบที่ไม่เหมาะสมกับรูปแบบมีอยู่ 9 ข้อ ได้แก่ ข้อ 5,11,12,16,36,37,38,41 และ 57 ดังนั้นแบบสอบฉบับนี้จึงมีความเหมาะสมกับรูปแบบ 85 %

แบบสอบฉบับที่ 2 ข้อสอบที่ไม่เหมาะสมกับรูปแบบมีอยู่ 17 ข้อ ได้แก่ ข้อ 6,7,11,13,16,17,18,27,30,31,33,35,38,39,44,46 และ 52 ดังนั้นแบบสอบฉบับนี้จึงมีความเหมาะสมกับรูปแบบ 70 %

เมื่อพิจารณาแบบสอบทั้ง 2 ฉบับแล้วปรากฏว่าแบบสอบมีความเหมาะสมกับรูปแบบของราสช์ 77.5 %

<u>หมายเหตุ</u> แบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ทั้ง 2 ฉบับได้ใช้การวิเคราะห์โดยโปรแกรมวิเคราะห์ข้อสอบตามแนวคิดของทฤษฎีมาตรฐานดั้งเดิมเพื่อหาความเที่ยง ปรากฏว่าฉบับที่ 1 ได้ค่าความเที่ยง .8996 และฉบับที่ 2 ได้ค่าความเที่ยง .9074 แสดงให้เห็นถึงคุณภาพของแบบสอบที่นำมาใช้เป็นเครื่องมือในการวิจัยกับกลุ่มตัวอย่างทั้ง 1,429 คนได้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 9 แสดงค่าความสามารถ (θ) และค่าความคลาดเคลื่อนมาตรฐานของค่าความสามารถของนักเรียนที่สอบแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ ฉบับที่ 1

คะแนนดิบ	จำนวนคน	ค่าความสามารถ	ค่าความคลาดเคลื่อนมาตรฐาน	คะแนนดิบ	จำนวนคน	ค่าความสามารถ	ค่าความคลาดเคลื่อนมาตรฐาน
59	0	4.37	1.02	29	49	-0.07	0.28
58	1	3.66	0.73	28	39	-0.15	0.28
57	8	3.23	0.60	27	46	-0.23	0.28
56	8	2.92	0.53	26	27	-0.31	0.28
55	11	2.67	0.48	25	34	-0.39	0.28
54	15	2.46	0.45	24	33	-0.47	0.29
53	13	2.27	0.42	23	26	-0.55	0.29
52	28	2.11	0.40	22	42	-0.63	0.29
51	17	1.96	0.38	21	25	-0.72	0.29
50	18	1.83	0.36	20	33	-0.80	0.30
49	21	1.70	0.35	19	20	-0.89	0.30
48	21	1.58	0.34	18	20	-0.98	0.30
47	26	1.47	0.33	17	35	-1.07	0.31
46	34	1.36	0.33	16	18	-1.16	0.31
45	29	1.26	0.32	15	15	-1.26	0.32
44	34	1.17	0.31	14	9	-1.36	0.33
43	36	1.07	0.31	13	8	-1.47	0.33
42	43	0.98	0.30	12	9	-1.58	0.34
41	51	0.89	0.30	11	4	-1.70	0.35
40	46	0.80	0.30	10	3	-1.83	0.37
39	48	0.72	0.29	9	0	-1.97	0.38
38	50	0.64	0.29	8	0	-2.11	0.40
37	52	0.55	0.29	7	0	-2.28	0.42
36	32	0.47	0.29	6	0	-2.46	0.45
35	48	0.39	0.28	5	0	-2.67	0.48
34	42	0.31	0.28	4	0	-2.92	0.53
33	54	0.24	0.28	3	0	-3.24	0.61
32	55	0.16	0.28	2	0	-3.67	0.73
31	53	0.08	0.28	1	0	-4.38	1.02
30	40	0.00	0.28				

จากตารางจะเห็นว่านักเรียน 1,429 คนที่สอบแบบสอบฉบับนี้มีพิสัยของค่าความสามารถ (θ) อยู่ระหว่าง -1.83 ถึง 3.66 โดยมีค่าความสามารถเฉลี่ยเท่ากับ 0.37 และค่าความเบี่ยงเบนมาตรฐานเท่ากับ 0.90

ตารางที่ 10 แสดงค่าความสามารถ (θ) และค่าความคลาดเคลื่อนมาตรฐานของค่าความสามารถของนักเรียนที่สอบแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ ฉบับที่ 2

คะแนนดิบ	จำนวนคน	ค่าความสามารถ	ค่าความคลาดเคลื่อนมาตรฐาน	คะแนนดิบ	จำนวนคน	ค่าความสามารถ	ค่าความคลาดเคลื่อนมาตรฐาน
59	0	4.36	1.02	29	30	-0.07	0.28
58	3	3.65	0.73	28	40	-0.15	0.28
57	5	3.22	0.60	27	51	-0.22	0.28
56	6	2.91	0.53	26	37	-0.30	0.28
55	13	2.66	0.48	25	55	-0.38	0.29
54	16	2.45	0.45	24	37	-0.46	0.29
53	17	2.27	0.42	23	37	-0.54	0.29
52	18	2.11	0.40	22	34	-0.63	0.29
51	16	1.96	0.38	21	29	-0.71	0.29
50	19	1.83	0.36	20	41	-0.80	0.30
49	22	1.70	0.35	19	29	-0.89	0.30
48	22	1.58	0.34	18	31	-0.93	0.30
47	34	1.47	0.33	17	22	-1.07	0.31
46	23	1.37	0.32	16	21	-1.16	0.31
45	32	1.27	0.32	15	15	-1.26	0.32
44	35	1.17	0.31	14	17	-1.37	0.33
43	42	1.08	0.31	13	13	-1.47	0.34
42	43	0.98	0.30	12	6	-1.59	0.34
41	28	0.90	0.30	11	8	-1.71	0.35
40	40	0.81	0.29	10	5	-1.83	0.37
39	37	0.73	0.29	9	1	-1.97	0.38
38	56	0.64	0.29	8	2	-2.12	0.40
37	40	0.56	0.29	7	0	-2.28	0.42
36	42	0.48	0.29	6	0	-2.47	0.45
35	38	0.40	0.28	5	0	-2.68	0.48
34	46	0.32	0.28	4	0	-2.93	0.53
33	43	0.24	0.28	3	0	-3.25	0.61
32	30	0.17	0.28	2	0	-3.68	0.73
31	49	0.09	0.28	1	0	-4.39	1.02
30	52	0.01	0.28				

จากตารางจะเห็นว่านักเรียน 1,429 คนที่สอบแบบสอบฉบับนี้มีพิสัยของค่าความสามารถ (θ) อยู่ระหว่าง -2.12 ถึง 3.65 โดยมีค่าความสามารถเฉลี่ยเท่ากับ 0.30 และค่าความเบี่ยงเบนมาตรฐานเท่ากับ 0.94

ตารางที่ 11 แสดงค่าสถิติพื้นฐานของ คะแนนดิบ ค่าความสามารถ และข้อสอบแต่ละข้อ ของแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ ฉบับที่ 1

นักเรียน จำนวนคน	คะแนนดิบ	ค่าเฉลี่ยความสามารถ	ค่าความคลาดเคลื่อนมาตรฐาน	จำนวนข้อ	ข้อสอบข้อที่
1	58	3.70	0.73		
8		3.50	0.73		
8	57	3.30	0.60		
+3 S.D.		3.10	0.60		
8	56	2.90	0.53		
11	55	2.70	0.48		
15	54	2.50	0.45		
13	53	2.30	0.42		
+2 S.D. 28	52	2.10	0.40		
35	50	1.90	0.36		
21	49	1.70	0.35	2	29 34
47	47	1.50	0.33	2	16 39
+1 S.D. 63	45	1.30	0.32	2	8 45
70	43	1.10	0.31	3	7 38 44
140	40	0.90	0.30	2	2 49
98	38	0.70	0.29	4	18 36 55 57
84	36	0.50	0.29	4	1 5 35 54
MEAN 144	33	0.30	0.28	5	3 24 46 53 56
148	30	0.10	0.28	7	4 22 33 37 42 47 58
88	28	-0.10	0.28	7	6 17 23 43 52 59 60
107	25	-0.30	0.28	4	20 25 30 32
-1 S.D. 59	23	-0.50	0.29	1	28
67	21	-0.70	0.29	3	13 40 41
73	18	-0.90	0.30	3	9 14 15
53	16	-1.10	0.31	4	11 31 48 51
24	14	-1.30	0.33	4	10 12 19 21
-2 S.D. 17	12	-1.50	0.34	2	27 50
4	11	-1.70	0.35		
3	9	-1.90	0.38		
	8	-2.10	0.40	1	26

ตารางที่ 12 แสดงค่าสถิติพื้นฐานของ คะแนนดิบ ค่าความสามารถ และข้อสอบแต่ละข้อของแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ ฉบับที่ 2

นักเรียน จำนวนคน	คะแนนดิบ	ค่าเฉลี่ยความสามารถ	ค่าความคลาดเคลื่อนมาตรฐาน	จำนวนข้อ	ข้อสอบข้อที่
3	58	3.70	0.73		
		3.50	0.73		
5	57	3.30	0.60		
+3 S.D.		3.10	0.60		
6	56	2.90	0.53		
13	55	2.70	0.48		
16	54	2.50	0.45		
17	53	2.30	0.42		
+2 S.D. 18	52	2.10	0.40	1	29
35	50	1.90	0.36		
22	49	1.70	0.35		
56	47	1.50	0.33	1	44
55	45	1.30	0.32	3	5 24 37
+1 S.D. 77	43	1.10	0.31	3	6 18 59
111	40	0.90	0.29	2	34 46
93	38	0.70	0.29	6	1 21 22 41 57 60
120	35	0.50	0.28	8	19 23 43 45 49 50 54 58
MEAN 89	33	0.30	0.28	7	2 3 4 7 14 20 35
131	30	0.10	0.28	2	16 56
70	28	-0.10	0.28	2	26 47
143	25	-0.30	0.29	4	30 31 39 48
74	23	-0.50	0.29	4	15 25 32 38
-1 S.D. 104	20	-0.70	0.30	3	13 33 55
60	18	-0.90	0.30	6	8 10 12 36 40 42
43	16	-1.10	0.31		
32	14	-1.30	0.33	2	9 53
-2 S.D. 19	12	-1.50	0.34	4	17 27 28 52
8	11	-1.70	0.35	1	11
6	9	-1.90	0.38	1	51
2	8	-2.10	0.40		

ตารางที่ 11 และตารางที่ 12 แสดงให้เห็นถึงความเกี่ยวข้องกันของคะแนนดิบ ค่าความสามารถและข้อสอบแต่ละข้อของแบบสอบผลสัมฤทธิ์ทั้ง 2 ฉบับ ดังนี้

แบบสอบฉบับที่ 1 จะเห็นว่าข้อสอบส่วนใหญ่จะเป็นข้อสอบที่เหมาะสมกับนักเรียนกลุ่มความสามารถปานกลางซึ่งมีอยู่ 31 ข้อ ข้อสอบที่เหมาะสำหรับกลุ่มความสามารถต่ำ 21 ข้อ และข้อสอบที่เหมาะสำหรับกลุ่มความสามารถสูง 8 ข้อ ตัวอย่างเช่น เมื่อพิจารณาข้อสอบกับความสามารถจะเห็นว่า

ข้อที่ 29,34 ตรงกับระดับความสามารถ (θ) 1.70 ซึ่งอยู่ในกลุ่มความสามารถสูงเป็นข้อที่ยากที่สุด

ข้อที่ 4,22,33,37,42,47 และ 58 ตรงกับความสามารถ (θ) เท่ากับ 0.10 ซึ่งอยู่ในกลุ่มความสามารถปานกลาง

ข้อที่ 26 ตรงกับระดับความสามารถ (θ) เท่ากับ -2.10 ซึ่งอยู่ในกลุ่มความสามารถต่ำเป็นข้อสอบที่ง่ายที่สุด

แบบสอบฉบับที่ 2 จะเห็นว่าข้อสอบส่วนใหญ่จะเป็นข้อสอบที่เหมาะสำหรับนักเรียนกลุ่มความสามารถปานกลางซึ่งมีอยู่ 30 ข้อ ข้อสอบที่เหมาะสำหรับกลุ่มความสามารถต่ำมีอยู่ 21 ข้อ และข้อสอบที่เหมาะสำหรับกลุ่มความสามารถสูงมีอยู่ 9 ข้อ ตัวอย่างเช่นเมื่อพิจารณาข้อสอบกับความสามารถจะเห็นว่า

ข้อที่ 29 ตรงกับระดับความสามารถ (θ) เท่ากับ 2.10 ซึ่งอยู่ในกลุ่มความสามารถสูงเป็นข้อที่ยากที่สุด

ข้อที่ 16 และ 56 ตรงกับระดับความสามารถเท่ากับ 0.10 ซึ่งอยู่ในกลุ่มความสามารถปานกลาง

ข้อที่ 51 ตรงกับระดับความสามารถเท่ากับ -2.10 ซึ่งอยู่ในกลุ่มความสามารถต่ำเป็นข้อสอบที่ง่ายที่สุด

ตอนที่ 2 การออกแบบแบบสอบ

1. การเทียบความสามารถของนักเรียนที่สอบแบบทั้ง 2 ฉบับ

เนื่องจากนักเรียนกลุ่มนี้ได้สอบแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ที่ผู้วิจัยได้ทำการเขียนข้อสอบให้มีความคล้ายคลึงกันมากที่สุดโดยใช้ Item form ของข้อสอบที่วัดจุดประสงค์การเรียนรู้เดียวกัน แต่ผลการสอบของนักเรียนคนหนึ่งในแบบสอบทั้ง 2 ฉบับ มีค่าความสามารถ(θ) ที่แตกต่างกันทั้งนี้เพราะได้คะแนนดิบไม่เท่ากัน ดังนั้นผู้วิจัยจึงได้ทำการเทียบคะแนนของแบบสอบฉบับที่ 2 มายังฉบับที่ 1 โดยอาศัยมโนทัศน์ของการเทียบคะแนนของทฤษฎีการตอบสนองข้อสอบ จากรูปแบบผู้สอบร่วมที่วอร์ม(Warm 1978:116) ได้เสนอไว้ ทำการวิเคราะห์คำนวนโดยใช้โปรแกรมภาษาฟอร์แทรน เมื่อได้ผลแล้วนำมาจัดแบ่งกลุ่มตามความสามารถดังตารางต่อไปนี้

ตารางที่ 13 แสดงการแบ่งกลุ่มความสามารถ(θ) ของนักเรียนที่สอบแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ทั้ง 2 ฉบับ ที่ได้เทียบคะแนนเรียบร้อยแล้ว

กลุ่มความสามารถ	ค่าเบี่ยงเบนมาตรฐาน	คะแนนดิบ	พิสัยความสามารถ	ค่าความคลาดเคลื่อนมาตรฐาน	จำนวนคน
สูง	> 1 SD.	45 - 58	1.30 - 3.70	0.32 - 0.73	246
ปานกลาง	± 1 SD.	25 - 44	-0.30 - 1.29	0.28 - 0.31	754
ต่ำ	< 1 SD.	8 - 24	-2.10 - -0.29	0.29 - 0.40	429

จากตาราง 13แสดงให้เห็นว่านักเรียนทั้ง 1,429 คนที่สอบแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ทั้ง 2 ฉบับเมื่อทำการเทียบ คะแนนแล้วสามารถแบ่งกลุ่มตามความสามารถ โดยใช้ค่าเบี่ยงเบนมาตรฐานของความสามารถเป็นเกณฑ์ ซึ่งปรากฏว่ากลุ่มความสามารถปานกลางมีจำนวนนักเรียนมากที่สุด รองลงมาได้แก่กลุ่มความสามารถต่ำและกลุ่มความสามารถสูงตามลำดับ

2. ผลการคัดเลือกข้อสอบเพื่อออกแบบแบบสอบด้วยการจัดข้อสอบที่มีความยากง่ายเหมาะสมกับระดับความสามารถของนักเรียนทั้ง 3 ระดับ คือ ความสามารถสูง ความสามารถปานกลาง และความสามารถต่ำ

ผู้วิจัยได้นำข้อสอบทั้ง 120 ข้อที่สอบโดยนักเรียนจำนวน 1,429 คน มาพิจารณาคัดเลือกตามเกณฑ์ที่ใช้คัดเลือกข้อสอบของทฤษฎีการตอบสนองข้อสอบ ซึ่งมีรายละเอียดดังนี้

เกณฑ์ในการคัดเลือกข้อสอบครั้งนี้ใช้ข้อสอบที่มีความเหมาะสมกับรูปแบบของราสช์ ซึ่งในการวิเคราะห์ได้ข้อสอบที่เหมาะสมกับรูปแบบของราสช์จำนวน 94 ข้อ และพิจารณาค่าความยาก (b) ในช่วง -2.5 ถึง 2.5 เมื่อได้ข้อสอบแล้วผู้วิจัยได้พิจารณาขอบเขตของเนื้อหาทั้งหมด 7 ขอบเขตเนื้อหาเพื่อที่จะนำข้อสอบที่มีความสอดคล้องกับเนื้อหานั้น ๆมาออกแบบแบบสอบสำหรับวัดความสามารถของนักเรียนที่แตกต่างกัน 3 ระดับ ทั้งนี้ข้อสอบที่บรรจุอยู่ในแบบสอบแต่ละฉบับที่สร้างขึ้นใหม่มีข้อสอบที่เหมาะกับความสามารถนั้น ๆเป็นหลักและมีข้อสอบที่เหมาะกับความสามารถอื่น ๆประกอบด้วยเพื่อที่ให้ได้โค้งแสดงสารสนเทศของแบบสอบที่มีช่วงความสามารถใกล้เคียงช่วงความสามารถตั้งแต่ -3.00 ถึง 3.00 ดังผลการคัดเลือกข้อสอบในตารางที่ 14 ต่อไปนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 14 แสดงข้อสอบที่คัดเลือกเพื่อวัดตามระดับความสามารถของนักเรียนทั้ง 3 ระดับ คือ ความสามารถสูง ปานกลาง และต่ำ

กลุ่มความสามารถ	ข้อสอบที่คัดเลือก ฉบับที่ 1	ข้อสอบที่คัดเลือก ฉบับที่ 2	พิสัยของค่าความยาก
สูง	1 3 8 10 17 18 25 29 32 33 34 35 39 42 43 44 45 52 54 55 56 58	2 5 14 22 23 24 25 29 32 34 37 41 43 45 49 57 59 60	-1.263 ถึง 2.072
ปานกลาง	1 3 4 6 17 18 20 22 23 24 25 30 32 33 35 42 43 44 46 47 53 55 56	1 2 3 4 14 15 20 22 25 32 34 41 50 54 56 57 60	-0.562 ถึง 1.073
ต่ำ	1 6 9 10 12 13 14 15 19 21 25 26 27 29 31 32 40 42 48 50 51 52 60	3 5 8 9 10 12 25 27 36 40 41 42 51 52 53 55 56	-2.059 ถึง 1.276

จากตารางพบว่าข้อสอบที่คัดเลือกเพื่อวัดตามระดับความสามารถของนักเรียนทั้ง 3 ระดับนั้นจะมีบางข้อที่อยู่หลายระดับดังจะแยกให้เห็นในแบบสอบแต่ละฉบับต่อไปนี้

<u>แบบสอบสำหรับความสามารถต่ำ</u>

กลุ่มความสามารถ	ฉบับที่ 1	ฉบับที่ 2
สูง	29	5
ปานกลาง	1 6 25 32 42 52 60	3 25 41 56
ต่ำ	9 10 12 13 14 15 19 21 26 27 31 40 48 50 51	8 9 10 12 27 36 40 42 51 52 53 55

แบบสอบสำหรับความสามารถปานกลาง

กลุ่มความสามารถ	ข้อสอบที่คัดเลือก	
	ฉบับที่ 1	ฉบับที่ 2
สูง	-	-
ปานกลาง	1 3 4 6 17 18 20 22 23 24 25 30 32 33 35 42 43 44 46 47 53 55 56	1 2 3 4 14 15 20 22 25 32 34 41 50 54 56 57 60
ต่ำ	-	-

แบบสอบสำหรับความสามารถสูง

กลุ่มความสามารถ	ข้อสอบที่คัดเลือก	
	ฉบับที่ 1	ฉบับที่ 2
สูง	8 29 34 39 45	5 24 29 37
ปานกลาง	1 3 17 18 25 32 33 35 42 43 44 52 54 55 56 58	2 14 22 23 25 32 34 41 43 45 49 57 59 60
ต่ำ	-	-

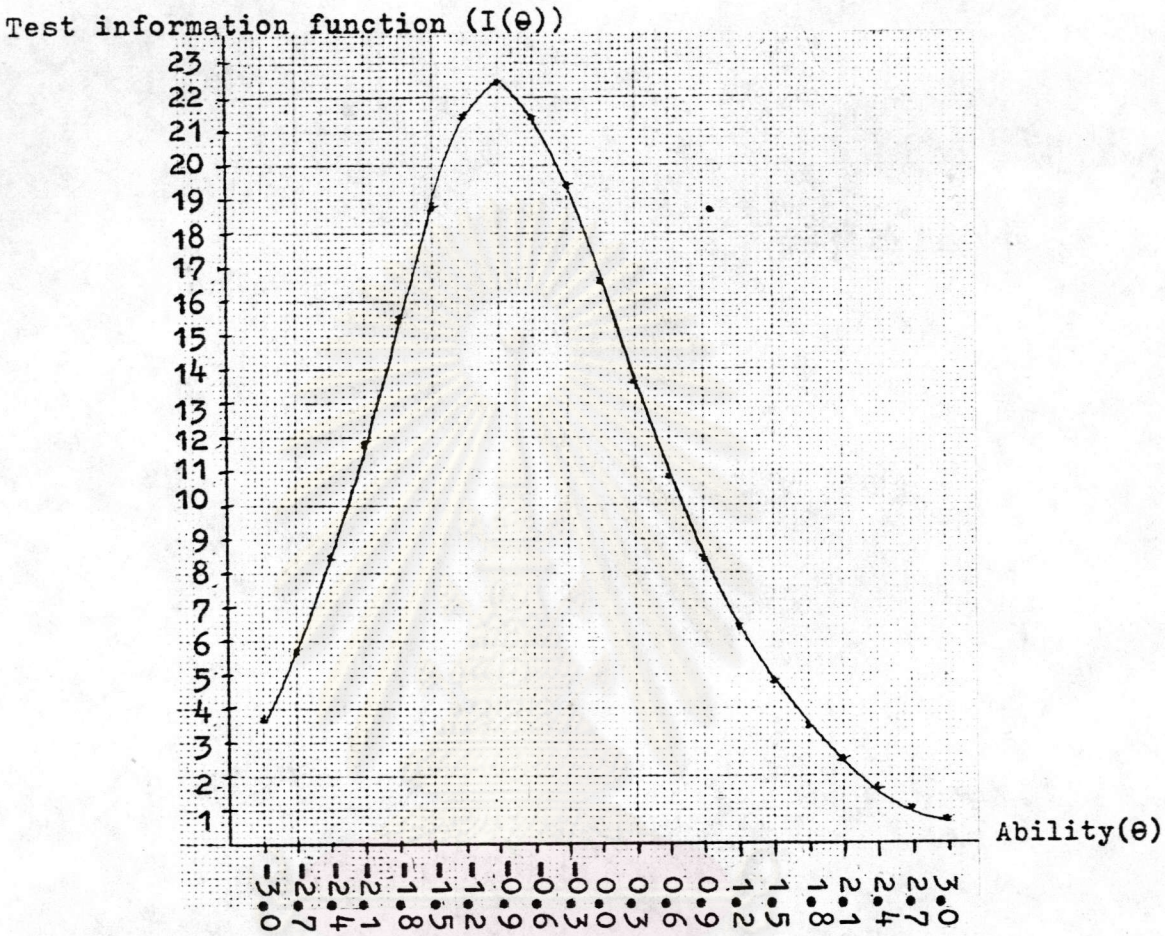
การคัดเลือกข้อสอบต้องพิจารณาให้ครอบคลุมขอบเขตเนื้อหาหลักเพื่อความเท่าเทียมกันในการวัดของแต่ละเนื้อหา เมื่อได้ข้อสอบตามที่ต้องการแล้วผู้วิจัยได้นำค่าความยาก (b) ของข้อสอบจำนวน 40 ข้อในแต่ละฉบับมาคำนวนหาค่าฟังก์ชั่นแสดงสารสนเทศของแบบสอบโดยใช้โปรแกรมภาษาฟอร์แทรนที่ผู้วิจัยเขียนขึ้นคำนวนหาค่าฟังก์ชั่นแสดงสารสนเทศของแบบสอบดังกล่าว ซึ่งช่วงความสามารถที่ผู้วิจัยได้นำมาพิจารณานั้นอยู่ในช่วง -3.00 ถึง 3.00 ผลการคำนวนได้ค่าฟังก์ชั่นแสดงสารสนเทศของแบบสอบทั้ง 3 ฉบับดังตารางต่อไปนี้

ตารางที่ 15 แสดงค่าฟังก์ชั่นแสดงสารสนเทศของแบบสอบ (Test information function) ที่ออกแบบเพื่อใช้สอบตามระดับความสามารถของผู้สอบทั้ง 3 ระดับ

ค่าความสามารถ (θ)	ค่าฟังก์ชั่นแสดงสารสนเทศของแบบสอบที่ออกแบบสำหรับกลุ่มความสามารถ		
	สูง	ปานกลาง	ต่ำ
-3.00	0.5266	0.9795	3.7648
-2.70	0.8571	1.5799	5.7574
-2.40	1.3768	2.5013	8.4606
-2.10	2.1708	3.8552	11.8162
-1.80	3.3405	5.7341	15.5145
-1.50	4.9934	8.1719	18.9589
-1.20	7.2290	11.1204	21.4015
-0.90	10.0638	14.4372	22.2709
-0.60	13.4080	17.8395	21.4759
-0.30	16.1251	20.8519	19.4021
0.00	20.0686	22.8684	16.6300
0.30	**22.2065**	**23.3578**	**13.6724**
0.60	**22.8522**	**22.0981**	**10.8824**
0.90	21.8837	19.3134	8.4551
1.20	19.5767	15.6147	6.4534
1.50	16.4341	11.7470	4.8469
1.80	12.9909	8.3026	3.5650
2.10	9.7001	5.5794	2.5452
2.40	6.8720	3.6096	1.7522
2.70	4.6501	2.2733	1.1626
3.00	3.0316	1.4063	0.7476

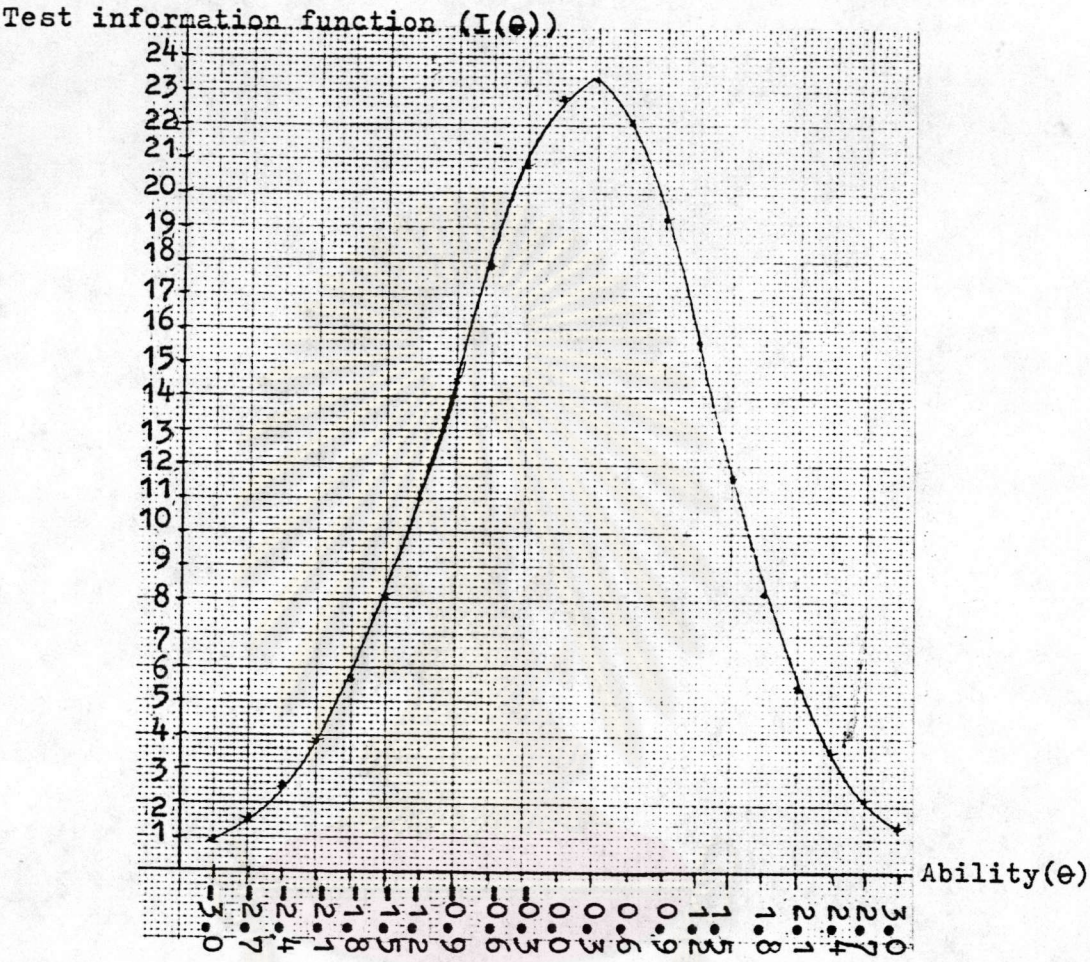
จากตารางพบว่าค่าฟังก์ชั่นแสดงสารสนเทศของแบบสอบที่ออกแบบตามระดับกลุ่มความสามารถทั้ง 3 ระดับนั้นมีค่าสูงที่ระดับความสามารถต่างกันคือแบบสอบที่วัดกลุ่มความสามารถสูงมีค่าสูงสุดที่ระดับความสามารถเท่ากับ 0.60 แบบสอบที่วัดกลุ่มความสามารถปานกลางมีค่าสูงสุดที่ระดับความสามารถเท่ากับ 0.30 และแบบสอบที่วัดกลุ่มความสามารถต่ำมีค่าสูงสุดที่ระดับความสามารถเท่ากับ -0.90 ซึ่งสามารถแสดงให้เห็นอย่างชัดเจนในแผนภาพที่ 14 - 16 ต่อไปนี้

แผนภาพที่ 14 แสดงโค้งแสดงสารสนเทศของแบบสอบ (Test information curve) ที่ใช้สอบวัดผู้สอบที่มีความสามารถต่ำ



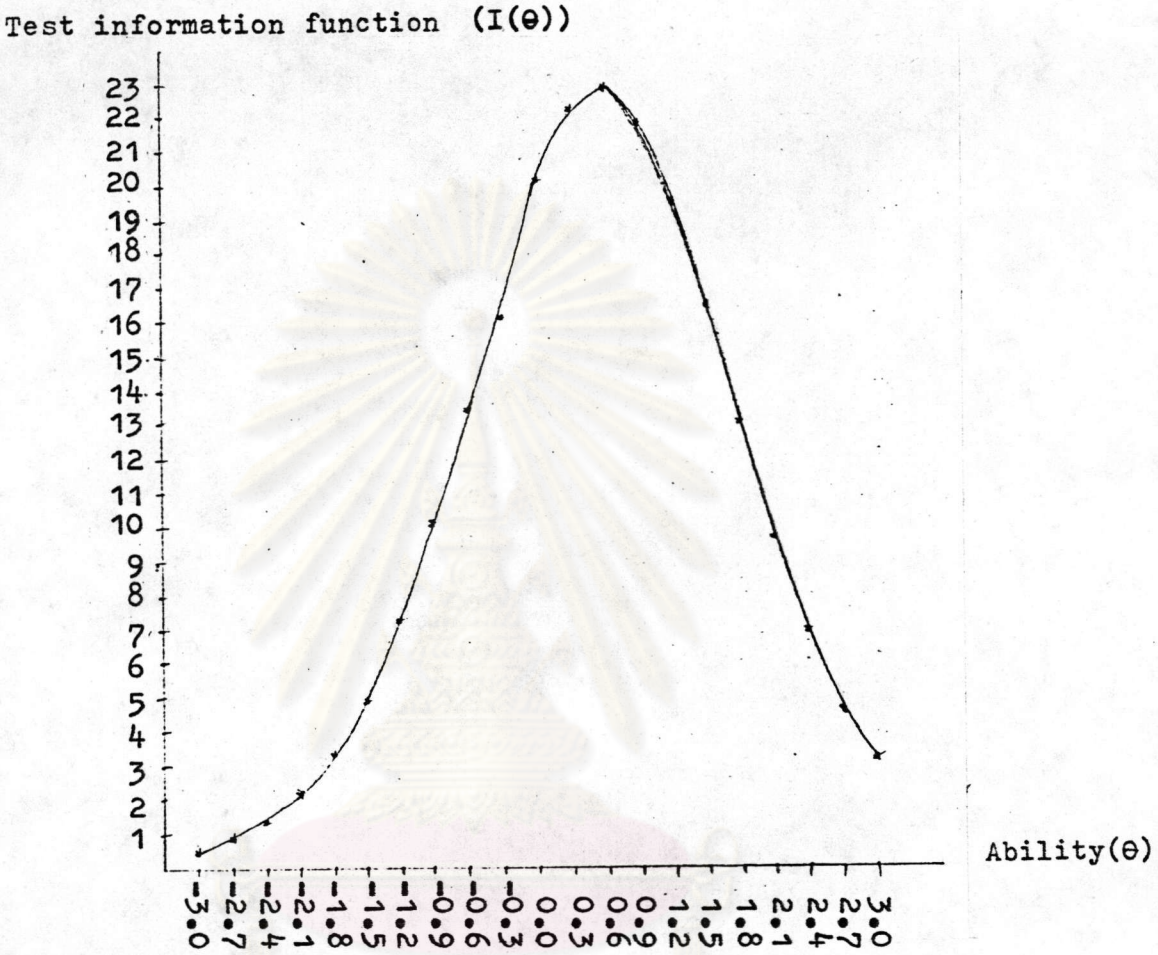
จากแผนภาพที่ 14 จะเห็นได้ว่าโค้งแสดงสารสนเทศของแบบสอบที่ใช้วัดกลุ่มความสามารถต่ำมีจุดยอดอยู่ระดับความสามารถ -0.9 ซึ่งเป็นจุดที่มีค่าฟังก์ชั่นแสดงสารสนเทศของแบบสอบสูงที่สุดคือ 22.2709

แผนภาพที่ 15 แสดงโค้งแสดงสารสนเทศของแบบสอบ (Test information curve) ที่ใช้สอบวัดผู้สอบที่มีความสามารถปานกลาง



จากแผนภาพที่ 15 จะเห็นได้ว่าโค้งแสดงสารสนเทศของแบบสอบที่ใช้วัดกลุ่มความสามารถปานกลางมีจุดยอดอยู่ที่ระดับความสามารถ 0.3 ซึ่งเป็นจุดที่มีค่าฟังก์ชั่นแสดงสารสนเทศของแบบสอบสูงที่สุดคือ 23.3578

แผนภาพที่ 16 แสดงโค้งแสดงสารสนเทศของแบบสอบ (Test information curve) ที่ใช้สอบวัดผู้สอบที่มีความสามารถสูง



จากแผนภาพที่ 16 จะเห็นได้ว่าโค้งแสดงสารสนเทศของแบบสอบที่ใช้วัดกลุ่มความสามารถสูงมีจุดยอดอยู่ที่ระดับความสามารถ 0.60 ซึ่งเป็นจุดที่มีค่าฟังก์ชั่นแสดงสารสนเทศของแบบสอบสูงที่สุดคือ 22.8522

ตอนที่ 3 ผลการตรวจสอบความแม่นยำ (precision) ของแบบสอบผลสัมฤทธิ์

1. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของนักเรียนทั้ง 3 ระดับ

การออกแบบแบบสอบเป็นการที่ผู้ดำเนินการสอบต้องการแบบสอบที่มีความคลาดเคลื่อนน้อยที่สุดการสร้างแบบสอบที่ตรงกับความสามารถของผู้สอบก็เพื่อที่จะให้ได้คะแนนผลการสอบเป็นที่เชื่อถือได้เมื่อกลุ่มความสามารถนั้น ๆสอบ อีกทั้งการออกแบบแบบสอบเป็นวิธีการที่จะทำให้แบบสอบที่ผู้ดำเนินการสอบสร้างขึ้นมีความเชื่อถือได้ในด้านค่าสถิติที่ต้องการ เช่น ความสามารถเฉลี่ยของกลุ่มผู้สอบที่ต้องการวัด จำนวนข้อสอบที่เหมาะสมกับกลุ่มผู้สอบ เป็นต้น

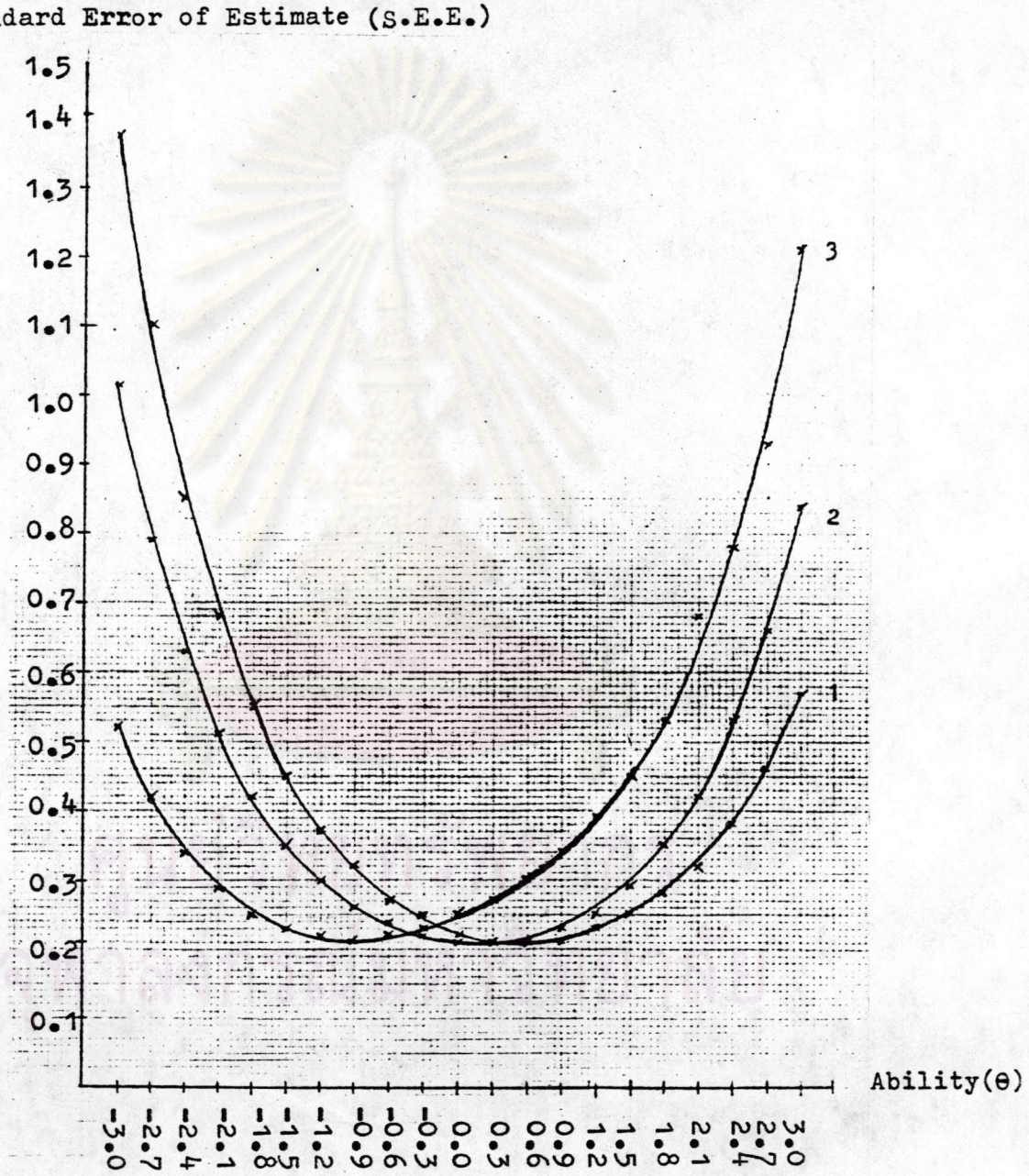
แบบสอบที่ผู้วิจัยออกแบบเพื่อวัดความสามารถของกลุ่มผู้สอบที่มีความสามารถแตกต่างกัน 3 ระดับ คือ ความสามารถสูง ปานกลาง และต่ำ นี้เป็นตัวอย่างที่ทำการทดลองเท่านั้น หากต้องการแบบสอบที่เหมาะสำหรับวัดกลุ่มความสามารถต่าง ๆกันอย่างแท้จริงจะต้องมีจำนวนข้อสอบที่อยู่ในคลังข้อสอบเป็นจำนวนมากกว่านี้เพื่อที่จะได้แยกความสามารถของนักเรียนได้อย่างชัดเจน แบบสอบที่ผู้วิจัยออกแบบไว้ทั้ง 3 ฉบับนี้สามารถคำนวนหาค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถได้จากสูตรที่เสนอไว้ในตอนต้นแล้วซึ่งค่าดังกล่าวผู้วิจัยได้เสนอไว้ในตารางที่ 16 ต่อไปนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 16 แสดงค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ (Standard Error of Estimate : S.E.E.) ของแบบสอบที่ออกแบบเพื่อใช้สอบตามระดับความสามารถของผู้สอบทั้ง 3 ระดับ

ระดับความสามารถ	ค่าความสามารถ (Θ)	ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของแบบสอบที่ออกแบบสำหรับกลุ่มความสามารถ สูง		ปานกลาง		ต่ำ	
	-3.00	1.37		1.01		0.52	
	-2.70	1.10		0.79		0.42	
	-2.40	0.85		0.63		0.34	
	-2.10	0.68		0.51		0.29	
ต่ำ	-1.80	0.55		0.42		0.25	
	-1.50	0.45	$\bar{X}$	0.35	$\bar{X}$	0.23	$\bar{X}$
	-1.20	0.37	.66	0.30	.50	0.22	.30
	-0.90	0.32	SD.	0.26	SD.	0.21	SD.
	-0.60	0.27	.36	0.24	.25	0.22	.10
	-0.30	0.25		0.22		0.23	
	0.00	0.22		0.21		0.25	
	0.30	0.21	$\bar{X}$	0.21	$\bar{X}$	0.27	$\bar{X}$
กลาง	0.60	0.21	.22	0.21	.22	0.30	.30
	0.90	0.21	SD.	0.23	SD.	0.34	SD.
	1.20	0.23	.01	0.25	.01	0.39	.05
	1.50	0.25		0.29		0.45	
	1.80	0.28		0.35		0.53	
	2.10	0.32	$\bar{X}$	0.42	$\bar{X}$	0.68	$\bar{X}$
สูง	2.40	0.38	.38	0.53	.52	0.78	.76
	2.70	0.46	SD.	0.66	SD.	0.93	SD.
	3.00	0.57	.11	0.84	.19	1.16	.24
		$\bar{X}$ = 0.45		$\bar{X}$ = 0.43		$\bar{X}$ = 0.43	
		SD. = 0.31		SD. = 0.23		SD. = 0.25	

แผนภาพที่ 17 กราฟแสดงการเปรียบเทียบการกระจายของค่าความคลาดเคลื่อนมาตรฐาน ในการประมาณค่าความสามารถของแบบสอบที่ออกแบบเพื่อใช้สอบตามระดับความสามารถของผู้สอบทั้ง 3 ระดับ



หมายเหตุ เส้นกราฟเส้นที่ 1 เป็นค่า S.E.E ของแบบสอบที่ใช้กับกลุ่มความสามารถสูง

เส้นกราฟเส้นที่ 2 เป็นค่า S.E.E ของแบบสอบที่ใช้กับกลุ่มความสามารถปานกลาง

เส้นกราฟเส้นที่ 3 เป็นค่า S.E.E ของแบบสอบที่ใช้กับกลุ่มความสามารถต่ำ

จากตารางที่ 16 พบว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ เมื่อพิจารณาโดยความสามารถรวมไม่แบ่งตามกลุ่มความสามารถปรากฏว่าค่าเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าไม่ต่างกันมาก แต่เมื่อพิจารณาแต่ละกลุ่มความสามารถจะเห็นว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของกลุ่มผู้สอบที่มีความสามารถตรงกับแบบสอบระดับนั้น ๆจะมีค่าเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถที่ต่ำกว่ากลุ่มอื่น ๆ

จากแผนภาพที่ 17 แสดงการเปรียบเทียบการกระจายของค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของนักเรียน(Standard Error of Estimate) ในช่วงความสามารถ -3.00 ถึง 3.00 ของแบบสอบทั้ง 3 ฉบับที่ใช้วัดความสามารถแตกต่างกัน จากเส้นกราฟลักษณะการกระจายของค่า S.E.E. ทั้ง 3 เส้นมีลักษณะการกระจายที่คล้าย ๆกัน คือ ในความสามารถสูงและต่ำค่า S.E.E. จะเพิ่มขึ้น ส่วนความสามารถปานกลางเส้นกราฟจะลดต่ำลง เส้นกราฟเส้นที่ 1 เป็นค่า S.E.E. ของแบบสอบที่ใช้กับกลุ่มความสามารถสูง เส้นกราฟจะสูงเมื่ออยู่ในช่วงความสามารถต่ำและจะลดลงในช่วงที่มีความสามารถปานกลางและจะสูงขึ้นแต่ไม่มากนักในช่วงความสามารถสูง แสดงว่าค่าความสามารถกระจายมากในช่วงความสามารถต่ำในขณะที่แบบสอบใช้ทดสอบกับกลุ่มความสามารถสูง เส้นกราฟเส้นที่ 2 เป็นกราฟค่า S.E.E. ของแบบสอบที่ใช้กับกลุ่มความสามารถปานกลาง เส้นกราฟจะสูงเมื่ออยู่ในช่วงความสามารถต่ำและสูง และลดลงเมื่ออยู่ในช่วงความสามารถปานกลาง แสดงว่ามีการกระจายสูงเมื่อใช้ผู้สอบที่มีความสามารถต่ำและสูงสอบ แต่ในช่วงความสามารถปานกลางจะมีการกระจายน้อย เส้นกราฟเส้นที่ 3 เป็นค่า S.E.E. ของแบบสอบที่ใช้กับกลุ่มความสามารถต่ำ เส้นกราฟจะสูงมากในช่วงความสามารถสูงและจะลดต่ำลงในช่วงความสามารถปานกลางจนกระทั่งต่ำในช่วงความสามารถต่ำ แสดงว่าค่าความสามารถกระจายสูงมากในช่วงที่มีความสามารถสูงในขณะที่แบบสอบใช้สอบกับความสามารถต่ำ

จากเส้นกราฟของค่า S.E.E. ทั้ง 3 เส้น แสดงว่าค่า S.E.E. จะมีการกระจายสูงมากเมื่อใช้กลุ่มผู้สอบที่มีความสามารถตรงกันข้ามกับแบบสอบที่ออกแบบไว้สำหรับกลุ่มความสามารถนั้น ๆ

ตารางที่ 17 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถจากแบบสอบทั้ง 3 ฉบับ เมื่อผู้สอบไม่จำแนกระดับความสามารถ (รวม)

แบบสอบที่ใช้สำหรับกลุ่มความสามารถ	ช่วงความสามารถ	ΣR	χ^2	df
สูง	21	676.5	0.0121	2
ปานกลาง	21	664.5		
ต่ำ	21	665		

จากตารางที่17 เปรียบเทียบค่าSEE ของแบบสอบที่ใช้สำหรับผู้สอบที่มีความสามารถแตกต่างกัน โดยใช้Kruskall Wallis Test โดยพิจารณาช่วงความสามารถตั้งแต่ -3.00 ถึง 3.00 ค่าΣR คือ ผลรวมของอันดับที่ของค่าS.E.Eทั้ง 3 แบบสอบ โดยแบบสอบที่ใช้สำหรับกลุ่มความสามารถสูงมีค่าΣR เท่ากับ 676.5 แบบสอบที่ใช้สำหรับกลุ่มความสามารถปานกลางมีค่าΣR เท่ากับ 664.5 และแบบสอบที่ใช้สำหรับกลุ่มความสามารถต่ำมีค่า ΣR เท่ากับ 665 จากผลการทดสอบปรากฏว่าไม่มีหลักฐานเพียงพอที่จะสรุปว่าค่าS.E.Eของทั้ง 3 แบบสอบแตกต่างกัน นั่นคือ การกระจายของค่าS.E.Eในช่วงความสามารถ -3.00 ถึง 3.00 ของทั้ง 3 แบบสอบไม่อาจสรุปได้ว่าแตกต่างกัน

แต่เมื่อพิจารณาผลการเปรียบเทียบค่าSEE ในแต่ละช่วงของความสามารถที่ผู้วิจัยกำหนดขึ้นตามกลุ่มความสามารถทั้ง 3 ระดับคือ

กลุ่มความสามารถสูง อยู่ระหว่าง 1.30 ถึง 3.00
กลุ่มความสามารถปานกลาง อยู่ระหว่าง -0.30 ถึง 1.29
กลุ่มความสามารถต่ำ อยู่ระหว่าง -3.00 ถึง -2.90

จะได้ผลการเปรียบเทียบดังต่อไปนี้

ตารางที่ 18 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถจากแบบสอบทั้ง 3 ฉบับ เมื่อกลุ่มผู้สอบมีความสามารถต่ำ

แบบสอบที่ใช้สำหรับกลุ่มความสามารถ	ช่วงความสามารถ	ΣR	χ^2	df
สูง	9	169	6.3871*	2
ปานกลาง	9	137.5		
ต่ำ	9	61.5		

* $P < .05$ จากตารางที่ 18 เปรียบเทียบค่า SEE ของแบบสอบที่ใช้สำหรับวัดผู้สอบที่มีความสามารถแตกต่างกัน 3 ฉบับ โดยใช้ Kruskall Wallis Test เมื่อผู้สอบมีความสามารถต่ำ ค่าΣR คือ ผลรวมของอันดับของค่า S.E.Eในแบบสอบแต่ละฉบับ โดยแบบสอบที่ใช้วัดกลุ่มความสามารถสูงมีค่า ΣR เท่ากับ 169 แบบสอบที่ใช้วัดกลุ่มความสามารถปานกลางมีค่า ΣR เท่ากับ 137.5 และแบบสอบที่ใช้วัดกลุ่มความสามารถต่ำมีค่า ΣR เท่ากับ 61.5 จากผลการทดสอบปรากฏว่าค่า χ^2 ของการทดลองมีค่า 6.3871 ซึ่งมากกว่าค่า $\chi^2_{2:.95}$ มีค่า 5.99 ที่ใช้เปรียบเทียบ แสดงว่าค่าเฉลี่ยของค่า S.E.E. ของทั้ง 3 แบบสอบเมื่อสอบโดยผู้สอบที่มีความสามารถต่ำแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05

ตารางที่ 19 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถจากแบบสอบทั้ง 3 ฉบับ เมื่อกลุ่มผู้สอบมีความสามารถปานกลาง

แบบสอบที่ใช้สำหรับกลุ่มความสามารถ	ช่วงความสามารถ	ΣR	χ^2	df
สูง	6	89	1.9825	2
ปานกลาง	6	22		
ต่ำ	6	41		

จากตารางที่ 19 เปรียบเทียบค่า SEE ของแบบสอบที่ใช้สำหรับวัดผู้ที่มีความสามารถแตกต่างกัน 3 ฉบับโดยใช้ Kruskall Wallis Test เมื่อผู้สอบมีความสามารถปานกลาง ค่า ΣR คือ ผลรวมของอันดับของค่า SEE ในแบบสอบแต่ละฉบับ โดยแบบสอบที่ใช้วัดกลุ่มความสามารถสูงมีค่า ΣR เท่ากับ 89 แบบสอบที่ใช้วัดกลุ่มความสามารถปานกลางมีค่า ΣR เท่ากับ 22 และแบบสอบที่ใช้วัดกลุ่มความสามารถต่ำมีค่า ΣR เท่ากับ 41 จากผลการทดสอบปรากฏว่า ค่า χ^2 ของการทดลองมีค่า 1.9825 ซึ่งน้อยกว่าค่า $\chi^2_{2:.95}$ มีค่า 5.99 ที่ใช้เปรียบเทียบ แสดงว่าไม่มีหลักฐานที่แสดงว่าค่าค่าเฉลี่ยของค่า SEE ของทั้ง 3 แบบสอบเมื่อผู้สอบที่มีความสามารถปานกลางไม่อาจสรุปได้ว่าแตกต่างกัน

ตารางที่ 20 **เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถจากแบบสอบทั้ง 3 ฉบับ เมื่อกลุ่มผู้สอบมีความสามารถสูง**

แบบสอบที่ใช้สำหรับกลุ่มความสามารถ	ช่วงความสามารถ	ΣR	χ^2	df
สูง	6	82.5	6.9327*	2
ปานกลาง	6	54.5		
ต่ำ	6	34		

* $P < .05$ จากตารางที่ 20 เปรียบเทียบค่า SEE ของแบบสอบที่ใช้สำหรับวัดผู้ที่มีความสามารถแตกต่างกัน 3 ฉบับ โดยใช้ Kruskall Wallis Test เมื่อผู้สอบมีความสามารถสูง ค่า ΣR คือ ผลรวมของอันดับของค่า SEE ในแบบสอบแต่ละฉบับโดยแบบสอบที่ใช้วัดกลุ่มความสามารถสูงมีค่า ΣR เท่ากับ 82.5 แบบสอบที่ใช้วัดกลุ่มความสามารถปานกลางมีค่า ΣR เท่ากับ 54.5 และแบบสอบที่ใช้วัดกลุ่มความสามารถต่ำมีค่า ΣR เท่ากับ 34 จากผลการทดสอบปรากฏว่าค่า ของการทดลองมีค่า 6.9327 ซึ่งมากกว่าค่า $\chi^2_{2:.95}$ มีค่า 5.99 ที่ใช้เปรียบเทียบ แสดงว่าค่าเฉลี่ยของค่า SEE ของทั้ง 3 แบบสอบเมื่อสอบโดยผู้สอบที่มีความสูงแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05

2. เปรียบเทียบค่าความเที่ยงของแบบสอบที่วัดความสามารถของนักเรียนทั้ง 3 ระดับ

แบบสอบที่ผู้วิจัยออกแบบและสุ่มข้อสอบบรรจุลงในโค้งแสดงสารสนเทศของแบบสอบดังกล่าวทั้ง 3 ฉบับ เมื่อสุ่มผู้สอบที่มีความสามารถแตกต่างกัน 3 ระดับมาทดสอบกับแบบสอบทั้ง 3 ฉบับดังกล่าวแล้วนำผลการสอบมาวิเคราะห์โดยใช้โปรแกรมวิเคราะห์ข้อสอบได้ค่าสถิติพื้นฐานดังต่อไปนี้

ตารางที่ 21 แสดงค่าสถิติพื้นฐานของแบบสอบที่เหมาะสำหรับความสามารถต่ำ

กลุ่มความสามารถ	n	$\bar{X}$	S.D.	r_{tt}	S.E.M.
ต่ำ	200	17.430	4.373	0.5344	2.384
ปานกลาง	200	17.275	4.271	0.5006	3.018
สูง	200	36.397	2.832	0.6450	1.687

ตารางที่ 22 แสดงค่าสถิติพื้นฐานของแบบสอบที่เหมาะสำหรับความสามารถปานกลาง

กลุ่มความสามารถ	n	$\bar{X}$	S.D.	r_{tt}	S.E.M.
ต่ำ	200	11.475	2.900	0.0475	2.830
ปานกลาง	200	26.840	4.476	0.6411	2.681
สูง	200	32.070	4.609	0.7571	2.272

ตารางที่ 23 แสดงค่าสถิติพื้นฐานของแบบสอบที่เหมาะสมสำหรับความสามารถสูง

กลุ่มความสามารถ	n	$\bar{X}$	S.D.	r_{tt}	S.E.M.
ต่ำ	200	10.640	2.966	0.1517	2.731
ปานกลาง	200	15.085	3.513	0.3126	2.913
สูง	200	29.355	5.158	0.7633	2.509

จากตารางที่ 21,22และ 23 จะเห็นว่าเมื่อใช้แบบสอบที่เหมาะสมกับกลุ่มความสามารถต่ำสอบวัดผู้สอบที่มีความสามารถแตกต่างกันค่าความเที่ยงของกลุ่มที่มีความสามารถสูงมีค่าสูงที่สุด รองลงมาได้แก่ผู้สอบที่มีความสามารถต่ำและปานกลาง เมื่อใช้แบบสอบที่เหมาะสำหรับกลุ่มความสามารถปานกลางและสูงสอบวัดผู้สอบที่มีความสามารถแตกต่างกันค่าความเที่ยงของกลุ่มที่มีความสามารถสูงมีค่าสูงสุด รองลงมาได้แก่ผู้สอบที่มีความสามารถปานกลางและต่ำทั้ง 2 ฉบับ

เมื่อพิจารณาในแต่ละกลุ่มความสามารถที่สอบแบบสอบที่แตกต่างกัน 3 ฉบับ กลุ่มที่มีความสามารถต่ำเมื่อสอบแบบสอบที่เหมาะสำหรับความสามารถต่ำจะมีค่าความเที่ยงสูงที่สุด รองลงมาได้แก่แบบสอบที่เหมาะสำหรับกลุ่มความสามารถสูงและปานกลางตามลำดับ กลุ่มที่มีความสามารถปานกลางเมื่อสอบแบบสอบที่เหมาะสำหรับความสามารถปานกลางมีค่าความเที่ยงสูงที่สุด รองลงมาได้แก่แบบสอบที่เหมาะสำหรับความสามารถต่ำและสูงตามลำดับ กลุ่มที่มีความสามารถสูงเมื่อสอบแบบสอบที่เหมาะสำหรับความสามารถสูงมีค่าความเที่ยงสูงที่สุด รองลงมาได้แก่แบบสอบที่เหมาะสำหรับความสามารถปานกลางและต่ำ นำค่าความเที่ยงของแบบสอบทั้ง 3 ฉบับที่สอบโดยกลุ่มความสามารถแตกต่างกัน มาแปลงให้อยู่ในรูปของคะแนนมาตรฐานฟิชเชอร์ซี แล้วทำการทดสอบรวมด้วยค่าไคสแควร์ (Chi square χ^2) ซึ่งได้ผลดังนี้

ตารางที่ 24 แสดงค่าความเที่ยงของแบบสอบที่ใช้วัดกลุ่มความสามารถทั้ง 3 ระดับเมื่อแปลงให้อยู่ในรูปของคะแนนมาตรฐานฟิชเชอร์ซีและผลการทดสอบนัยสำคัญของค่าความแตกต่างโดยใช้ค่าไคสแควร์

กลุ่มความสามารถ	แบบสอบที่ออกแบบสำหรับกลุ่มความสามารถ						χ^2
	ต่ำ		ปานกลาง		สูง		
	r_{tt}	Z_r	r_{tt}	Z_r	r_{tt}	Z_r	
ต่ำ	.5344	.596	.0475	.048	.1517	.153	33.3309**
ปานกลาง	.5006	.550	.6411	.760	.3126	.323	18.8199**
สูง	.6450	.767	.7571	.989	.7633	1.004	6.9468*

** $P < .01$

* $P < .05$

จากตารางที่ 24 จะเห็นได้ว่าเมื่อทำการเปรียบเทียบค่าความเที่ยงของแบบสอบที่เหมาะสำหรับกลุ่มความสามารถต่าง ๆกันกับกลุ่มผู้สอบในแต่ละระดับด้วยค่าไคสแควร์(Chi square: χ^2) ซึ่งเป็นการทดสอบรวมพบว่าค่าความเที่ยงของแบบสอบทั้ง 3 ฉบับเมื่อสอบด้วยผู้สอบที่มีความสามารถแตกต่างกันมีค่าความเที่ยงต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เมื่อกลุ่มผู้สอบมีความสามารถต่ำและปานกลาง ส่วนแบบสอบทั้ง 3 ฉบับดังกล่าวเมื่อผู้สอบมีความสามารถสูงสอบ ค่าความเที่ยงแตกต่างกันที่ระดับ .05 ดังนั้นจึงต้องทำการทดสอบรายคู่ด้วยสัมประสิทธิ์ฟิชเชอร์ซี ซึ่งผลการทดสอบสามารถแสดงให้เห็นได้ดังตารางต่อไปนี้

ตารางที่ 25 แสดงความแตกต่างระหว่างค่าความเที่ยงของกลุ่มผู้สอบที่มีความสามารถต่ำเมื่อสอบแบบสอบที่เหมาะสำหรับความสามารถต่าง ๆกัน ในรูปอัตราส่วนซี

แบบสอบที่เหมาะสำหรับความสามารถ	ต่ำ	ปานกลาง	สูง
ต่ำ	-	5.4392**	4.3970**
ปานกลาง	-	-	1.0422
สูง	-	-	-

** $P < .01$

จากตารางจะเห็นว่ากลุ่มผู้สอบที่มีความสามารถต่ำเมื่อสอบแบบสอบที่เหมาะสำหรับความสามารถต่าง ๆกัน มีความแตกต่างกันทางสถิติระหว่างค่าความเที่ยงของแบบสอบที่เหมาะสำหรับความสามารถปานกลางและต่ำ และแบบสอบที่เหมาะสำหรับความสามารถสูงและต่ำ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนค่าความเที่ยงของแบบสอบที่เหมาะสำหรับความสามารถปานกลางและสูงไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แต่อย่างใด

ตารางที่ 26 แสดงความแตกต่างระหว่างค่าความเที่ยงของกลุ่มผู้สอบที่มีความสามารถปานกลาง เมื่อสอบแบบสอบที่เหมาะสำหรับความสามารถต่าง ๆกัน ในรูปอัตราส่วนซี

แบบสอบที่เหมาะสำหรับความสามารถ	ต่ำ	ปานกลาง	สูง
ต่ำ	-	2.0844*	2.2531*
ปานกลาง	-	-	4.3375**
สูง	-	-	-

** P < .01
* P < .05

จากตารางจะเห็นว่ากลุ่มผู้สอบที่มีความสามารถปานกลางเมื่อสอบแบบสอบที่มีความเหมาะสำหรับความสามารถต่าง ๆกัน มีความแตกต่างกันทางสถิติระหว่างค่าความเที่ยงของแบบสอบที่เหมาะสำหรับความสามาถปานกลางและต่ำ และแบบสอบที่เหมาะสำหรับความสามารถสูงและต่ำอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 สำหรับแบบสอบที่เหมาะสำหรับความสามารถปานกลางและสูงมีนัยสำคัญทางสถิติที่ระดับ .01

ตารางที่ 27 แสดงความแตกต่างระหว่างค่าความเที่ยงของกลุ่มผู้สอบที่มีความสามารถสูง เมื่อสอบแบบสอบที่เหมาะสำหรับความสามารถต่าง ๆกันในรูปอัตราส่วนซี

แบบสอบที่เหมาะสำหรับความสามารถ	ต่ำ	ปานกลาง	สูง
ต่ำ	-	2.2035*	2.3524*
ปานกลาง	-	-	0.1489
สูง	-	-	-

* P < .05

จากตารางจะเห็นว่ากลุ่มผู้สอบที่มีความสามารถสูงเมื่อสอบแบบสอบที่เหมาะสำหรับความสามารถต่าง ๆกัน มีความแตกต่างกันทางสถิติระหว่างค่าความเที่ยงของแบบสอบที่เหมาะสำหรับความสามารถปานกลางและต่ำ และแบบสอบที่เหมาะสำหรับความสามารถสูงและต่ำอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนค่าความเที่ยงของแบบสอบที่เหมาะสำหรับความสามารถปานกลางและสูงไม่มีนัยสำคัญทางสถิติแต่อย่างใด