

370.1524

4436 7

8.3

การพัฒนาการวัดความสามารถด้านเหตุผลเชิงวิเคราะห์
ของนักเรียนระดับชั้นมัธยมศึกษา

ปริญญาณพนธ์

ของ

บุญชู ชลัชเรีอ์

14 ส.ค. 2540

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรการศึกษาคณะศึกษาศาสตร์ สาขาการทดสอบและวัดผลการศึกษา

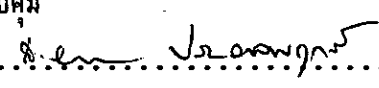
สิงหาคม 2539

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

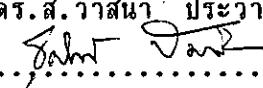
53412

คณะกรรมการควบคุมและคณะกรรมการสอบได้พิจารณาปัญหานี้นี้แล้ว
เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาตรีบัณฑิต
สาขาการทดสอบและวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒได้

คณะกรรมการควบคุม

.....  ประธาน

(รศ.ดร.ส.วาสนา ประवालพฤษ)

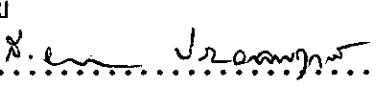
.....  กรรมการ

(ดร.ชุติต ทัมกลีขิต)

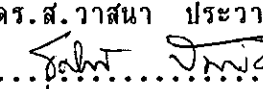
.....  กรรมการ

(รศ.ดร.เชิดศักดิ์ โฆวาสินธุ์)

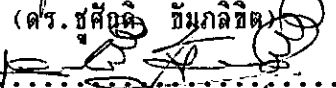
คณะกรรมการสอบ

.....  ประธาน

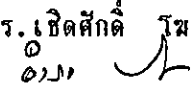
(รศ.ดร.ส.วาสนา ประवालพฤษ)

.....  กรรมการ

(ดร.ชุติต ทัมกลีขิต)

.....  กรรมการ

(รศ.ดร.เชิดศักดิ์ โฆวาสินธุ์)

.....  กรรมการที่แต่งตั้งเพิ่มเติม

(ดร.อาร์ง จันทวานิช)

บัณฑิตวิทยาลัยอนุมัติให้รับปัญหานี้นี้เป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษาบัณฑิต สาขาการทดสอบและวัดผลการศึกษา
ของมหาวิทยาลัยศรีนครินทรวิโรฒ

.....  คณบดีบัณฑิตวิทยาลัย

(ดร.ศิริยา พลสุวรรณ)

วันที่ ๒๔ เดือน สิงหาคม พ.ศ. 2539.

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้สำเร็จได้ด้วยดี เนื่องจากผู้วิจัยได้ความรู้ แนวคิด และคำแนะนำอย่างดียิ่งจาก รองศาสตราจารย์ ดร.ส.วาสนา ประवालพฤกษ์ อาจารย์ ดร.ชูศักดิ์ ชัมภลจิต และรองศาสตราจารย์ ดร.เชิดศักดิ์ ไชวาสินธุ์ อาจารย์ที่ปรึกษาปริญญานิพนธ์ ผู้วิจัยรู้สึกซาบซึ้ง และขอกราบขอบพระคุณในความกรุณาของอาจารย์ทุกท่าน ด้วยความเคารพเป็นอย่างสูง

ขอกราบขอบพระคุณ ดร.อำรุง จันทวานิช อธิบดีกรมวิชาการ กระทรวงศึกษาธิการ ซึ่งร่วมเป็นกรรมการสอบภายนอก

ขอขอบพระคุณผู้เชี่ยวชาญทุกท่าน ที่ได้กรุณาพิจารณาให้ความเห็นอันเป็นประโยชน์ยิ่งต่อการพัฒนาเครื่องมือเพื่อการศึกษาวิจัยครั้งนี้ และขอบส่วนที่เป็นคุณความดีจากผลงานวิจัยครั้งนี้ให้แก่ นักคิด นักเขียน นักวิจัย และนักวัดผลทั้งหลายที่ได้สร้างผลงานวิจัย ตำรา เพื่อการศึกษาอันเป็นแหล่งสำหรับค้นคว้าและอ้างอิงของผู้วิจัย เพื่อให้เกิดพัฒนาการทางความคิดในการสร้างผลงานวิจัยครั้งนี้ ด้วยความสำนึกในบุญคุณเป็นอย่างสูง

ขอขอบพระคุณผู้บริหารโรงเรียน อาจารย์ผู้ให้ความอนุเคราะห์ช่วยเก็บรวบรวมข้อมูล โดยเฉพาะอาจารย์ฝ่ายแนะแนว และนักเรียนที่เป็นกลุ่มตัวอย่างที่ได้ให้ความร่วมมือเป็นอย่างดียิ่งในการเก็บรวบรวมข้อมูลครั้งนี้

ขอขอบคุณอาจารย์บุญเรือง และ อาจารย์รัตนา ศรีเทรัญ คุณรุ่งฤดี ทวีพันธ์ ที่ได้ให้ความช่วยเหลือในการวิเคราะห์และประมวลผลข้อมูลจนแล้วเสร็จ ด้วยคุณภาพของงานเป็นที่น่าภูมิใจ รวมทั้งขอขอบคุณ คุณวราวุธ รุ่งเรือง เจริญกุล และคุณปานัน สุขมาก ซึ่งเป็นผู้พิมพ์ต้นฉบับงานวิจัยฉบับนี้ด้วยความอุตสาหะไว้มิ ๗ ๖๕

ท้ายที่สุดนี้ ขอขอบคุณ เพื่อน พี่ น้อง และสมาชิกทุกคนในครอบครัวของผู้วิจัยที่เป็นกำลังใจและให้ความสนับสนุนผู้วิจัยด้วยดีตลอดมา

บุญชู ชลัษเฐียร

สารบัญ

บทที่	หน้า
1 บทนำ	1
ความเป็นมาและความสำคัญของปัญหา	1
จุดประสงค์ของการศึกษา	7
ความสำคัญของการศึกษาค้นคว้า	7
ขอบเขตของการศึกษา	8
ตัวแปรที่ศึกษา	8
นิยามศัพท์เฉพาะ	9
2 เอกสารและงานวิจัยที่เกี่ยวข้อง	12
แนวคิดและทฤษฎีเกี่ยวกับการคิด ความสามารถด้านเหตุผล	12
โครงสร้างของความคิด	14
ประเภทแบบของการคิด	18
แนวทางในการพัฒนาความสามารถด้านการคิด	21
ความสามารถด้านเหตุผล	29
แนวคิดเกี่ยวกับการวัดและการประเมินผล	35
บทบาทของการวัดและการประเมินผล	35
เครื่องมือวัด	38
คุณภาพและการตรวจสอบคุณภาพของเครื่องมือ	62
ทฤษฎีการตอบสนองข้อสอบ	82
ดัชนีกำหนดคุณภาพและการเปรียบเทียบคุณภาพของแบบทดสอบ	87
3 วิธีดำเนินการ	90
ประชากร	90
กลุ่มตัวอย่าง	90
เครื่องมือที่พัฒนาขึ้นเพื่อการศึกษา	93

การเก็บรวบรวมข้อมูล	93
หลักการและการดำเนินการสร้างเครื่องมือวัด	93
ตัวอย่างเครื่องมือ	100
การศึกษาคูณภาพของข้อคำถามและแบบทดสอบ	119
สถิติที่ใช้ในการวิเคราะห์ข้อมูล	120
 4 ผลการวิเคราะห์ข้อมูล	122
สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล	122
การเสนอผลการวิเคราะห์ข้อมูล	123
ผลการดำเนินการพัฒนาเครื่องมือวัด	124
ผลการวิเคราะห์ข้อมูล	124
การศึกษาคูณภาพของข้อคำถามและแบบทดสอบ	137
ผลการวิเคราะห์ข้อมูล	137
 5 สรุป อภิปรายผล และข้อเสนอแนะ	287
วัตถุประสงค์ของการวิจัย	287
ขอบเขตของการวิจัย	287
วิธีดำเนินการวิจัย	288
สรุปผลการวิจัย	290
อภิปรายผล	295
ข้อเสนอแนะ	303
 บรรณานุกรม	307
 ภาคผนวก	318
 ประวัติย่อของผู้วิจัย	396

บัญชีตาราง

ตาราง

หน้า

1	รายชื่อโรงเรียน และจำนวนนักเรียนที่เป็นกลุ่มตัวอย่าง ในการศึกษาในกระบวนการพัฒนาข้อคำถาม (ทดลองครั้งที่ 1)	91
2	รายชื่อโรงเรียน และจำนวนนักเรียนที่เป็นกลุ่มตัวอย่าง ในกระบวนการศึกษาคุณภาพของเครื่องมือ (ทดลองครั้งที่ 2)	92
3	จำนวนข้อคำถามของแบบทดสอบแต่ละฉบับ จำนวนผู้ตอบทั้งหมด ค่าเฉลี่ย ค่าความเบี่ยงเบนมาตรฐานของคะแนน ค่ามัธยฐาน และช่วงคะแนนที่ได้จากแบบทดสอบแต่ละฉบับ โดยกลุ่มตัวอย่าง ที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3	125
4	จำนวนข้อคำถามของแบบทดสอบแต่ละฉบับ จำนวนผู้ตอบทั้งหมด ค่าเฉลี่ย ค่าความเบี่ยงเบนมาตรฐานของคะแนน ค่ามัธยฐาน และช่วงคะแนนที่ได้จากแบบทดสอบแต่ละฉบับ โดยกลุ่มตัวอย่าง ที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5	126
5	คุณลักษณะของข้อคำถามของแบบทดสอบวัดความสามารถด้านเหตุผล เชิงวิเคราะห์ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3	128
6	คุณลักษณะของข้อคำถามของแบบทดสอบวัดความสามารถด้านเหตุผล เชิงวิเคราะห์ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5	131
7	คุณภาพด้านความเชื่อมั่นของแบบทดสอบจากการทดลองครั้งที่ 1	134
8	ข้อคำถามที่ได้รับการพิจารณาปรับแก้ และตัดออกจากการทดลอง ครั้งที่ 1	136
9	คุณลักษณะของข้อคำถามของแบบทดสอบจากการทดลองครั้งที่ 2 ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3	138
10	คุณลักษณะของข้อคำถามของแบบทดสอบจากการทดลองครั้งที่ 2 ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5	140
11	ค่าความเชื่อมั่นของแบบทดสอบจากการทดลองครั้งที่ 2	141
12	ผลการวิเคราะห์องค์ประกอบของแบบทดสอบแบบเลือกตอบ	143

13	ผลการวิเคราะห์องค์ประกอบของแบบทดสอบแบบเติมคำตอบ	145
14	ค่าสถิติพื้นฐานของค่าอำนาจจำแนก (a) ของแบบทดสอบ ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3	148
15	ค่าสถิติพื้นฐานของค่าอำนาจจำแนก (a) ที่ได้จากกลุ่มตัวอย่าง ที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5	149
16	ค่าสถิติพื้นฐานของค่าความยาก (b) ของแบบทดสอบที่ได้จาก กลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3	151
17	ค่าสถิติพื้นฐานของค่าความยาก (b) ของแบบทดสอบที่ได้จาก กลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5	152
18	ค่าสถิติพื้นฐานของค่าสัมประสิทธิ์การเดา (c) ของแบบทดสอบ ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3	154
19	ค่าสถิติพื้นฐานของค่าสัมประสิทธิ์การเดา (c) ของแบบทดสอบ ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5	155
20	ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ AR แบบเลือกตอบ จากการทดลองครั้งที่ 1	320
21	ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ AX แบบเลือกตอบ จากการทดลองครั้งที่ 1	321
22	ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ CV แบบเลือกตอบ จากการทดลองครั้งที่ 1	322
23	ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ LR แบบเลือกตอบ จากการทดลองครั้งที่ 1	323
24	ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ NLR แบบเลือกตอบ จากการทดลองครั้งที่ 1	324
25	ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ PI แบบเลือกตอบ จากการทดลองครั้งที่ 1	325
26	ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ AR แบบเติมคำตอบ จากการทดลองครั้งที่ 1	326

27	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ AX แบบเติมคำตอบ จากการทดลองครั้งที่ 1	327
28	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ CV แบบเติมคำตอบ จากการทดลองครั้งที่ 1	328
29	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ LR แบบเติมคำตอบ จากการทดลองครั้งที่ 1	329
30	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ NLR แบบเติมคำตอบ จากการทดลองครั้งที่ 1	330
31	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ PI แบบเติมคำตอบ จากการทดลองครั้งที่ 1	331
32	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ AR แบบเลือกตอบ จากการทดลองครั้งที่ 2	334
33	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ AX แบบเลือกตอบ จากการทดลองครั้งที่ 2	335
34	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ CV แบบเลือกตอบ จากการทดลองครั้งที่ 2	336
35	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ LR แบบเลือกตอบ จากการทดลองครั้งที่ 2	337
36	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ NLR แบบเลือกตอบ จากการทดลองครั้งที่ 2	338
37	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ PI แบบเลือกตอบ จากการทดลองครั้งที่ 2	339
38	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ AR แบบเติมคำตอบ จากการทดลองครั้งที่ 2	340
39	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ AX แบบเติมคำตอบ จากการทดลองครั้งที่ 2	341

40	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ CV แบบเติมคำตอบ จากการทดลองครั้งที่ 2	342
41	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ LR แบบเติมคำตอบ จากการทดลองครั้งที่ 2	343
42	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ NLR แบบเติมคำตอบ จากการทดลองครั้งที่ 2	344
43	ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ PI แบบเติมคำตอบ จากการทดลองครั้งที่ 2	345
44	คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษาปีที่ 3 ฉบับ AR	347
45	คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษาปีที่ 3 ฉบับ AX	348
46	คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษาปีที่ 3 ฉบับ CV	349
47	คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษาปีที่ 3 ฉบับ LR	350
48	คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษาปีที่ 3 ฉบับ NLR	351
49	คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษาปีที่ 3 ฉบับ PI	352
50	คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษาปีที่ 5 ฉบับ AR	353
51	คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษาปีที่ 5 ฉบับ AX	354
52	คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษาปีที่ 5 ฉบับ CV	355

53	คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษาปีที่ 5 ฉบับ LR	356
54	คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษาปีที่ 5 ฉบับ NLR	357
55	คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษาปีที่ 5 ฉบับ PI	358
56	จำนวนข้อสอบของแบบทดสอบแต่ละฉบับ (ชั้นมัธยมศึกษาปีที่ 3)	360
57	จำนวนข้อสอบของแบบทดสอบแต่ละฉบับ (ชั้นมัธยมศึกษาปีที่ 5)	361
58	คุณภาพด้านความเชื่อมั่นของแบบทดสอบ ภายหลังจากตัดข้อคำถาม บางข้อที่ไม่เหมาะสมกับโมเดลออกไป	362
59	ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความ คลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ ของแบบทดสอบ ชนิด AR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	364
60	ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความ คลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ ของแบบทดสอบ ชนิด AX แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	366
61	ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความ คลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ ของแบบทดสอบ ชนิด CV แบบเลือกตอบและแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	368
62	ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความ คลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ ของแบบทดสอบ ชนิด LR แบบเลือกตอบและแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	370

- 63 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความ
คลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ
ของแบบทดสอบ ชนิด NLR แบบเลือกตอบและแบบเติมคำตอบ
ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 372
- 64 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความ
คลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ
ของแบบทดสอบ ชนิด PI แบบเลือกตอบและแบบเติมคำตอบ
ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 374
- 65 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความ
คลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ
ของแบบทดสอบ ชนิด AR แบบเลือกตอบและแบบเติมคำตอบ
ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 376
- 66 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความ
คลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ
ของแบบทดสอบ ชนิด AX แบบเลือกตอบและแบบเติมคำตอบ
ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 378
- 67 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความ
คลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของ
แบบทดสอบ ชนิด CV แบบเลือกตอบและแบบเติมคำตอบ
ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 380
- 68 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความ
คลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ
ของแบบทดสอบ ชนิด LR แบบเลือกตอบและแบบเติมคำตอบ
ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 382

69	ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ชนิด NLR แบบเลือกตอบและแบบเติมคำตอบของนักเรียนชั้นมัธยมศึกษาปีที่ 5	384
70	ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ชนิด PI แบบเลือกตอบและแบบเติมคำตอบของนักเรียนชั้นมัธยมศึกษาปีที่ 5	386

บัญชีภาพประกอบ

ภาพประกอบ

หน้า

1 ลำดับขั้นของการคิด (Hierarchy of Thinking)	4
2 ยุทธศาสตร์ของการให้เหตุผล (The Heuristics of Reasoning)	24
3 ทักษะและยุทธศาสตร์การคิดที่สำคัญ	26
4 คุณสมบัติและข้อดีของข้อถามแบบปรนัย	41
5 Taxonomy of Thinking Skills (เฉพาะในส่วนของ Reasoning Skills)	56
6 วิธีการตรวจสอบความถูกต้องของแบบทดสอบ	64
7 แนวทางการวัด	66
8 ชนิดของ Criterion-Related Validation (ยึดความต่างของเวลา)	68
9 ข้อมูลหลักฐานบางประการที่ใช้ในการตรวจสอบความเที่ยงตรง เชิงโครงสร้าง	71
10 วิธีการประมาณค่าความเชื่อมั่น ที่นำเสนอโดย American Psychological Association (Standard for Educational and Psychological Testing, Washington D.C. APA. 1985)	77
11 เกณฑ์การพิจารณาระดับความเชื่อมั่น	81
12 ขั้นตอนการพัฒนาแบบทดสอบ	95
13 โครงสร้างและทักษะย่อยที่ต้องการวัด	97
14 การตรวจสอบความเที่ยงตรงของเครื่องมือวัดที่พัฒนาขึ้น ตามขั้นตอนต่าง ๆ	99
15 ค่าเฉลี่ยสัดส่วนการตอบถูกต้องสำหรับแบบทดสอบ 6 ชนิด ทั้ง 2 รูปแบบ ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3	129

- 16 ค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบ 6 ชนิด ทั้ง 2 รูปแบบ
 ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 130
- 17 ค่าเฉลี่ยสัดส่วนการตอบถูกสำหรับแบบทดสอบ 6 ชนิด ทั้ง 2 รูปแบบ
 ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 132
- 18 ค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบ 6 ชนิด ทั้ง 2 รูปแบบ
 ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 133
- 19 ค่าความเชื่อมั่นของแบบทดสอบ 6 ชนิด ทั้ง 2 รูปแบบ
 ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 135
- 20 ค่าความเชื่อมั่นของแบบทดสอบ 6 ชนิด ทั้ง 2 รูปแบบ
 ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 135
- 21 โคลงลักษณะของแบบทดสอบ ชนิด AR แบบเลือกตอบของ
 นักเรียนชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ 157
- 22 โคลงสารสนเทศของแบบทดสอบ ชนิด AR แบบเลือกตอบ
 ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ 157
- 23 โคลงลักษณะของแบบทดสอบ ชนิด AX แบบเลือกตอบของนักเรียน
 ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ 158
- 24 โคลงสารสนเทศของแบบทดสอบ ชนิด AX แบบเลือกตอบของ
 นักเรียนชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ 158
- 25 โคลงลักษณะของแบบทดสอบ ชนิด CV แบบเลือกตอบของนักเรียน
 ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ 159
- 26 โคลงสารสนเทศของแบบทดสอบ ชนิด CV แบบเลือกตอบของนักเรียน
 ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ 159
- 27 โคลงลักษณะของแบบทดสอบ ชนิด LR แบบเลือกตอบของนักเรียน
 ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ 160
- 28 โคลงสารสนเทศของแบบทดสอบ ชนิด LR แบบเลือกตอบของนักเรียน
 ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ 160

[illegible]

[illegible]

[illegible]

68	คำสั่งสาธนาเทศของแบบทดสอบ ชนิด PI แบบเติมคำตอบของนักเรียน ชั้นมัธยมศึกษาปีที่ 5 ๗ ระดับความสามารถต่าง ๆ	180
69	คำสั่งลักษณะข้อสอบของข้อสอบ ชนิด AR แบบเลือกตอบ ของนักเรียน ชั้นมัธยมศึกษาปีที่ 3	182
70	คำสั่งลักษณะข้อสอบของข้อสอบ ชนิด AX แบบเลือกตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	186
71	คำสั่งลักษณะข้อสอบของข้อสอบ ชนิด CV แบบเลือกตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	189
72	คำสั่งลักษณะข้อสอบของข้อสอบ ชนิด LR แบบเลือกตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	192
73	คำสั่งลักษณะข้อสอบของข้อสอบ ชนิด NLR แบบเลือกตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	196
74	คำสั่งลักษณะข้อสอบของข้อสอบ ชนิด PI แบบเลือกตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	199
75	คำสั่งลักษณะข้อสอบของข้อสอบ ชนิด AR แบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	203
76	คำสั่งลักษณะข้อสอบของข้อสอบ ชนิด AX แบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	207
77	คำสั่งลักษณะข้อสอบของข้อสอบ ชนิด CV แบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	210
78	คำสั่งลักษณะข้อสอบของข้อสอบ ชนิด LR แบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	213
79	คำสั่งลักษณะข้อสอบของข้อสอบ ชนิด NLR แบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	216
80	คำสั่งลักษณะข้อสอบของข้อสอบ ชนิด PI แบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	218

81	โค้งลักษณะข้อสอบของข้อสอบ ชนิด AR แบบเลือกตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	222
82	โค้งลักษณะข้อสอบของข้อสอบ ชนิด AX แบบเลือกตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	226
83	โค้งลักษณะข้อสอบของข้อสอบ ชนิด CV แบบเลือกตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	229
84	โค้งลักษณะข้อสอบของข้อสอบ ชนิด LR แบบเลือกตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	231
85	โค้งลักษณะข้อสอบของข้อสอบ ชนิด NLR แบบเลือกตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	235
86	โค้งลักษณะข้อสอบของข้อสอบ ชนิด PI แบบเลือกตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	238
87	โค้งลักษณะข้อสอบของข้อสอบ ชนิด AR แบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	242
88	โค้งลักษณะข้อสอบของข้อสอบ ชนิด AX แบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	246
89	โค้งลักษณะข้อสอบของข้อสอบ ชนิด CV แบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	249
90	โค้งลักษณะข้อสอบของข้อสอบ ชนิด LR แบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	252
91	โค้งลักษณะข้อสอบของข้อสอบ ชนิด NLR แบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	256
92	โค้งลักษณะข้อสอบของข้อสอบ ชนิด PI แบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	259
93	เปรียบเทียบโค้งสารสนเทศของแบบทดสอบชนิด AR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3	263

[illegible]

107	เปรียบเทียบโครงสร้างสนเทศของแบบทดสอบชนิด AX แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	277
108	ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด AX แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	278
109	เปรียบเทียบโครงสร้างสนเทศของแบบทดสอบชนิด CV แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	279
110	ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด CV แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	280
111	เปรียบเทียบโครงสร้างสนเทศของแบบทดสอบชนิด LR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	281
112	ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด LR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	282
113	เปรียบเทียบโครงสร้างสนเทศของแบบทดสอบชนิด NLR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	283
114	ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด NLR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	284
115	เปรียบเทียบโครงสร้างสนเทศของแบบทดสอบชนิด PI แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	285
116	ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด PI แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5	286

ความเป็นมาและความสำคัญของปัญหา

การเน้น "คน" เป็นศูนย์กลางหรือจุดมุ่งหมายหลักของการพัฒนา ถือเป็นวิสัยทัศน์ (Vision) ใหม่สำหรับการวางแผนพัฒนาประเทศของสังคมไทย ทั้งนี้เพราะต้องการให้ทุกคนในสังคมได้รับการพัฒนาให้เต็มตามศักยภาพ มีภูมิปัญญาความคิดที่จะเป็นผู้ริเริ่มและสร้างการพัฒนา ซึ่งถือเป็นการพัฒนาที่ยั่งยืน แนวความคิดที่เน้นคนเป็นศูนย์กลางของการพัฒนา ดังกล่าว ปรากฏอย่างชัดเจนในการสัมมนาในระดับชาติเกี่ยวกับแนวความคิดและทิศทางการพัฒนาประเทศในช่วงแผนพัฒนาเศรษฐกิจและสังคมแห่งชาติ ฉบับที่ 8 (พ.ศ. 2540-2544) และเพื่อให้บรรลุจุดมุ่งหมายดังกล่าว ได้มีการนำเสนอยุทธศาสตร์การพัฒนาโดยการปรับปรุงกระบวนการเรียนรู้ และการฝึกอบรมให้คิดเป็น ทำเป็น มีการเรียนรู้จากประสบการณ์จริง (คณะกรรมการพัฒนาการเศรษฐกิจและสังคมแห่งชาติ. 2538)

อย่างไรก็ตามแผนพัฒนาเศรษฐกิจและสังคมของประเทศ ฉบับที่ 7 (2535-2539) ที่กำลังใช้อยู่ในปัจจุบัน วัตถุประสงค์ของการพัฒนาก็อยู่ที่การพัฒนาทรัพยากรมนุษย์เช่นกัน วัตถุประสงค์ดังกล่าวก็นำมาซึ่งนโยบายการจัดการศึกษาตามแผนการศึกษาแห่งชาติ พุทธศักราช 2535 อันเป็นแม่บทนำไปสู่มาตรการการจัดการศึกษาของชาติขณะนี้

จากการวิเคราะห์ ปรัชญา จุดมุ่งหมายการพัฒนาซึ่งมุ่งไปที่ "คน" ในระดับต่าง ๆ นับแต่ระดับของแผนพัฒนาฯ แผนการศึกษาแห่งชาติ หลักสูตรการศึกษา ในแต่ละระดับแล้ว จะเห็นว่ามุ่งให้ "คน" มีพัฒนาการทั้งในด้านสติปัญญา อารมณ์ สังคม และสุขภาพอนามัย ทั้งในด้านสุขภาพอนามัย อารมณ์ และสังคม นั้น มุ่งสร้างคุณภาพชีวิตที่ดี และสังคมที่สงบสุขยิ่งขึ้น ส่วนด้านสติปัญญานั้น เป้าหมายอยู่ที่การพัฒนาคนให้มีความสามารถในการคิด มีเหตุผล อันจะนำไปสู่ความสามารถในการแก้ปัญหาที่ต้องเผชิญในชีวิตจริงได้อย่างถูกต้องเหมาะสม

✱ ท่านพุทธทาสภิกขุ (2517 : 41 - 42) ได้ให้ทัศนะไว้เช่นกันว่า การมีเหตุผลที่ดีและมีจริยธรรมคุณธรรมที่สูงส่ง เป็นคุณค่าแห่งความเป็นมนุษย์ที่สมบูรณ์ เหตุผลช่วยให้บุคคลค้นพบความจริงอันเป็นประโยชน์แก่การดำรงชีวิต และจริยธรรมช่วยให้บุคคลเป็นมนุษย์ที่สมบูรณ์ เพราะมนุษย์หมายถึง ผู้มีจิตใจสูง

ครูลิค (Krulik. 1993 : 8 - 9) กล่าวถึง เหตุผลของการสอนซึ่งเน้นที่การแก้ปัญหา (Problem Solving) และทักษะในการใช้เหตุผล (Reasoning Skill) ใน

สหรัฐอเมริกาว่า ผลงานวิจัยได้แสดงให้เห็นว่า เมื่อทักษะการคิด (Thinking Skills) ถูกสอนโดยตรงผลสัมฤทธิ์ทางการเรียนจะเพิ่มขึ้น การแก้ปัญหา และความมีเหตุผลเป็นพื้นฐานสำคัญต่อชีวิตประจำวัน นักเรียนทุกคนจะต้องเผชิญกับปัญหาในชีวิตประจำวันของเขา การแก้ปัญหาเป็นสิ่งที่น่าตื่นเต้นท้าทาย ซึ่งจะนำไปสู่การทำงานอย่างมีผลละ นอกจากนี้ การแก้ปัญหายังเปิดโอกาสให้ผู้เรียนได้ฝึกการคิดเพื่อค้นหาคำตอบ ซึ่งถือเป็นยุทธศาสตร์การเรียนรู้ การแก้ปัญหาเป็นกิจกรรมที่จะต้องเกิดขึ้นตลอดชีวิต เด็กจะเข้าไปเกี่ยวข้องกับปัญหานั้นตั้งแต่เกิด ประสบการณ์การแก้ปัญหามีตลอดเวลา ดังนั้น การสอนการแก้ปัญหาจะต้องกระทำอย่างต่อเนื่อง

Krulik. 1993: 89 หากจะมีผู้ตั้งคำถามว่า เมื่อไรจึงจะสอนความมีเหตุผล คงจะต้องตอบคำถามก่อนว่าความมีเหตุผลเป็นรากฐานของทุกสิ่งที่เรากระทำ กิจกรรมทุกกิจกรรมที่เกิดจากความคิดจะประกอบด้วยเหตุผลบางประการ ทุกสิ่งที่บุคคลกระทำเป็นผลจากการสรุปจากกระบวนการของเหตุผล (Reasoning-Process) ฉะนั้น การพัฒนาทักษะการใช้เหตุผลจะต้องเป็นเป้าหมายแรกของครูทุกคนในทุกห้องเรียน ครูต้องใช้โอกาสทุกสถานการณ์ในชั้นเรียนในการสนับสนุนการใช้เหตุผลของนักเรียน การแก้ปัญหาในชั้นเรียนถือเป็นกิจกรรมหรือเป็นเครื่องมือที่ดีเลิศ ซึ่งจะช่วยให้ผู้เรียนได้พัฒนาความสามารถด้านเหตุผล ทั้งการแก้ปัญหาและความสามารถด้านเหตุผลเป็นทักษะเชิงกระบวนการ (Process Skills) และการคิดไม่อาจเกิดขึ้นได้โดยปราศจากสิ่งที่ต้องคิด การฝึกการใช้เหตุผลจะต้องถูกนำไปใช้กับสถานการณ์เฉพาะที่ซึ่งผู้เรียนสามารถเชื่อมโยงได้ (Krulik. 1993 : 8 - 9)

สมิท (Smith. 1992 : 125) กล่าวว่า ทุกคนเกิดมาประกอบด้วย ความพร้อมที่จะคิดและเรียนรู้ในลักษณะที่เป็นการปฏิบัติอย่างต่อเนื่องของจินตนาการ ที่ผลักดันให้เราทำให้โลกนี้มีความหมาย และจากการวิเคราะห์สถานการณ์การศึกษาตั้งแต่ระดับอนุบาลจนถึงมหาวิทยาลัย การคิดควรเกิดขึ้นเกี่ยวข้องกับทุกกิจกรรมในแต่ละวัน ไม่ใช่ในลักษณะของแบบฝึกหัดหรือเนื้อหาวิชา แต่ในลักษณะการปฏิบัติการทุกอย่าง

นอกจากนี้ การคิดมีหลายประเภทแตกต่างกัน บุคคลใช้วิธีการคิดและการแก้ปัญหาที่ต่างกันผ่านประสบการณ์ การอบรมปลูกฝัง และการศึกษา คุณภาพและวิธีการคิดจะขึ้นอยู่กับพื้นฐานทางพันธุกรรมและชีววิทยาของมนุษย์ (Valette. 1978 : 31)

ด้วยความสำคัญของการคิด ความสามารถด้านเหตุผล และการแก้ปัญหาที่มีต่อการดำเนินชีวิตประจำวันของคนดังกล่าว นักจิตวิทยา นักการศึกษา จึงให้ความสนใจศึกษาค้นคว้าเกี่ยวกับการคิดมาเป็นเวลานาน ทั้งนี้เพื่อความเข้าใจถึงธรรมชาติในแง่มุมต่าง ๆ

อาทิ แบบการคิด องค์ประกอบเพื่อประโยชน์ต่อการพัฒนาต่อไป และจากแนวความคิดและความเชื่อที่ว่า ความคิดเหตุผล และการแก้ปัญหา เป็นความสามารถที่จะพัฒนาให้เกิดแก่ตัวของ "คน" ได้ โดยผ่านกระบวนการเรียนรู้ที่เกิดจากการอบรมปลูกฝัง และการศึกษาการจัดหลักสูตร กระบวนการเรียนการสอนในช่วงแผนพัฒนาฯ ระยะที่ 7 ตั้งแต่ ปี พ.ศ. 2535 ที่ผ่านมา กระทรวงศึกษาธิการในฐานะผู้รับผิดชอบการจัดการศึกษาของชาติได้พยายามหามาตรการและดำเนินการเพื่อตอบสนองแนวความคิดนี้ ดังจะเห็นได้จากการกำหนดจุดมุ่งหมายของหลักสูตรแต่ละระดับชั้นหรือในรายวิชาต่าง ๆ ได้แสดงจุดเน้นเกี่ยวกับความสามารถในการคิดและเหตุผลอย่างชัดเจน เช่น ในวิชาคณิตศาสตร์ระดับมัธยมศึกษาตอนต้น มุ่งเน้นให้ผู้เรียนมีความรู้ความเข้าใจในวิชาคณิตศาสตร์ ข้อมูลที่ปรากฏในสิ่งแวดล้อมสามารถคิดอย่างมีเหตุผล ใช้เหตุผลในการแสดงความคิดเห็นอย่างมีระเบียบชัดเจน และรัดกุม (กรมวิชาการ. 2533 : 40)

ตลอดระยะเวลา 3 ปีที่ผ่านมา นับจากการเริ่มใช้แผนพัฒนาฯ ฉบับที่ 7 ผู้เรียนซึ่งเป็นผลผลิตอันเนื่องมาจากกระบวนการเรียนการสอนที่มุ่งเน้นให้เป็นผู้มีปัญญา รู้จักเหตุและผล รู้จักแยกแยะนิยตชอบชั่วดี มีคุณลักษณะดังมุ่งหวังเพียงใด เป็นคำถามสำคัญที่ต้องการคำตอบ ฉะนั้น หน่วยงานผู้รับผิดชอบในการจัดการศึกษาระดับต่าง ๆ จึงได้ดำเนินการประเมินความก้าวหน้าทางการเรียน และประเมินคุณภาพของผู้เรียนแต่ละระดับชั้นในช่วงระยะเวลาดัง ๆ กัน เพื่อตอบคำถามเกี่ยวกับคุณภาพของผู้เรียนและแนวโน้มของผลผลิตโดยรวม เพื่อประโยชน์ต่อการปรับปรุงแก้ไข และวางแผนการจัดการศึกษาในอนาคต ทั้งนี้การประเมินมุ่งไปที่ผลสัมฤทธิ์ทางการเรียนทั่วไป (General Achievement) ผลสัมฤทธิ์ทางการเรียนทั่วไปแต่เพียงประการเดียวเพียงพอต่อการรองรับ หรือตอบคำถามเกี่ยวกับคุณภาพทางการศึกษาหรือไม่ ผู้วิจัยมีความเห็นว่า แม้เครื่องมือที่ใช้ในการวัดจะยึดเป้าหมายของหลักสูตร จุดประสงค์รายวิชาหรือกลุ่มประสบการณ์ เป็นกรอบสร้างเครื่องมือ โดยมีได้ยึดติดเนื้อหาที่ตาม ผลการสอบวัดดังกล่าวยังไม่เพียงพอและไม่ครอบคลุม เพราะเพียงให้คำตอบว่า ผู้เรียนมีผลสัมฤทธิ์ทางการเรียนเฉพาะส่วนที่เกี่ยวกับความรู้เชิงวิชาการเป็นอย่างไร แต่คุณลักษณะที่พึงประสงค์อื่นของผู้เรียน ได้แก่ ความสามารถในการคิด การให้เหตุผล และการแก้ปัญหา ซึ่งเป็นทักษะที่จำเป็นยังได้รับคำตอบน้อยมากหรือแทบไม่ปรากฏ

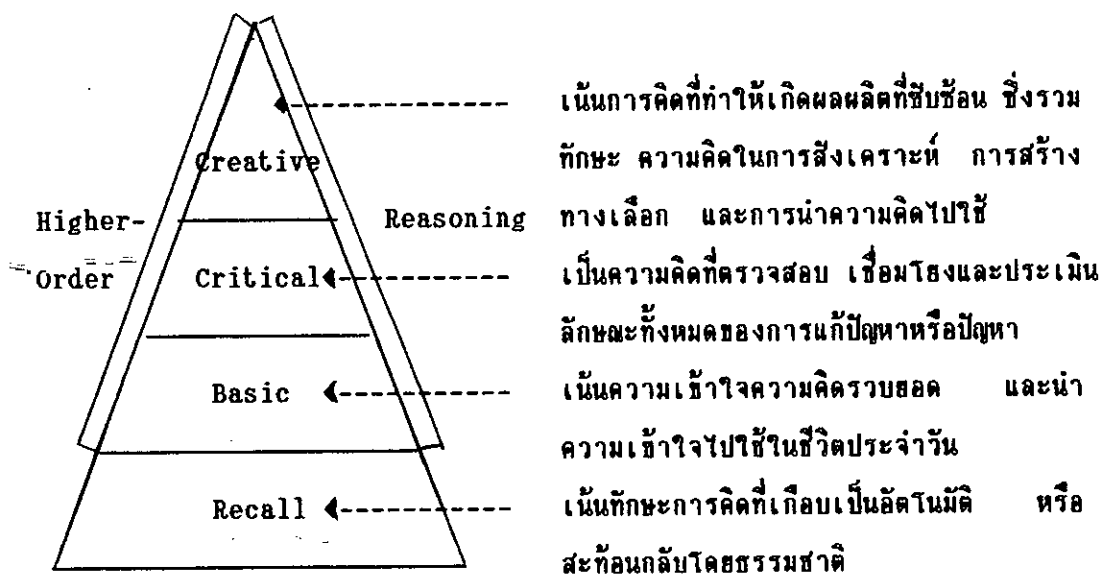
ไรดัง ได้ให้ความเห็นเกี่ยวกับการประเมิน (Assessment) ในอนาคตว่า การพัฒนาการวัดจะเปลี่ยนแปลงได้ ต้องมีความสอดคล้องกับการวิจัยที่เกี่ยวกับการเรียนรู้ การประเมินและการพัฒนาวิธีการแนวทางการวัด จำเป็นต้องเปลี่ยนจากการเน้นที่วิชาการที่แคบ

ไปสู่ประโยชน์ใช้สอยที่กว้างกว่านั้น สิ่งที่ต้องการต่อไปจะอยู่ในรูปของทักษะการเรียนรู้ เน้นความสามารถที่จะเรียนรู้ (Ability to Learn) มิใช่เน้นความจำ ในเรื่องคุณภาพการศึกษจะต้องมีการติดตามความเคลื่อนไหวของระดับผลสัมฤทธิ์ของผู้เรียนที่พัฒนาอย่างต่อเนื่อง (Riding and Butterfield. 1990 : 227)

ครูลิค (Krulik. 1993 : 3 - 5) กล่าวว่า ปัจจุบันคณิตศาสตร์ในโรงเรียน เน้นไปที่การพัฒนาความสามารถในการใช้เหตุผล (Ability to Reason) ของผู้เรียน นักจิตวิทยาและนักวิจัยจำนวนมากพยายามที่จะจำแนกความแตกต่างระหว่าง "ความมีเหตุผล" (Reasoning) และ "การคิด" (Thinking)

"การคิด" (Thinking) หมายถึง ความสามารถของเด็กในการเข้าถึงข้อสรุปที่ถูกต้องจากข้อมูลที่ได้ เด็กจะต้องสร้างความนึกคิดสรุปคุณสมบัติจากความสัมพันธ์ในสถานการณ์ปัญหา จากนั้นตรวจสอบอธิบายข้อสรุปและข้ออ้างนั้น ข้อสรุปเหล่านี้จะถูกนำมารวมกันเพื่อสร้างความคิดใหม่ ๆ

ครูลิค มองความมีเหตุผลว่าเป็นส่วนหนึ่งของความคิดที่อยู่เหนือกว่าระดับความรู้ (Knowledge) หรือการระลึกได้ (Recall) ดังภาพประกอบ 1



ภาพประกอบ 1 ลำดับขั้นของการคิด (Hierarchy of Thinking)

เหตุผลในระดับของการคิดอย่างมีวิจารณญาณ (Critical Thinking) เป็นระดับที่ผู้วิจัยสนใจ เนื่องจากพิจารณาเห็นว่า การคิดในระดับนี้ประกอบด้วยทักษะย่อยหลายประการ คือ การมุ่งไปยังส่วนต่าง ๆ ของปัญหาหรือสถานการณ์ที่เผชิญอยู่ การตรวจสอบความถูกต้อง การวิเคราะห์ข้อมูล รวมทั้งการจดจำและเชื่อมโยงข้อมูลที่เพิ่งได้รับการเรียนรู้ ซึ่งจำเป็นต่อการแก้ปัญหาหรือเผชิญปัญหาในชีวิตจริง มีลักษณะเป็นการวิเคราะห์ (Analytical) และไตร่ตรอง (Reflexive) โดยธรรมชาติ จึงน่าจะสอดคล้องกับจุดหมายของการจัดการศึกษาในระดับสามัญมากที่สุด

การพัฒนาความสามารถด้านเหตุผลให้เกิดขึ้นในตัวผู้เรียน ถือเป็นความมุ่งหวังของนักการศึกษาทั้งไทยและต่างประเทศ ขณะเดียวกันก็ได้ให้ความสนใจต่อการพัฒนากระบวนการหรือกระบวนการสอบวัดความสามารถด้านนี้กันอย่างกว้างขวาง ทั้งนี้เพื่อเป็นยุทธศาสตร์ที่ช่วยให้กระบวนการพัฒนาเป็นไปอย่างสมบูรณ์ ดังที่ ไรดิงค์ และบัทเทอร์ฟิลด์ (Riding- and Butterfield, 1990 : 44 - 45) กล่าวถึงนิยามที่สำคัญของการประเมิน (Evaluation) ว่า การประเมินจะช่วยท้าวความกระจ่างให้กับเป้าหมายและจุดมุ่งหมายของกระบวนการ สำหรับกำหนดขอบเขตการพัฒนาผู้เรียนไปในทิศทางที่ต้องการ ทำหน้าที่เป็นระบบควบคุมคุณภาพ (Quality Control) ที่จะทำให้รู้ว่าแต่ละขั้นของกระบวนการเรียนการสอนมีประสิทธิภาพหรือไม่ และถ้าไม่มีประสิทธิภาพจะต้องเปลี่ยนแปลงในส่วนใดก่อนที่จะสายเกินไป นอกจากนี้ การประเมินเปรียบเสมือนเครื่องมือของการปฏิบัติงานทางการศึกษา สำหรับยืนยันว่าวิธีการที่เลือกมีประสิทธิภาพหรือไม่ต่อการบรรลุเป้าหมายทางการศึกษา

จากการทบทวนวรรณกรรมเกี่ยวกับเครื่องมือ หรือการวัดความสามารถในการใช้เหตุผล (Reasoning Ability) ทั้งที่เป็นแบบทดสอบมาตรฐานซึ่งเป็นที่ยอมรับและใช้กันอย่างแพร่หลายในประเทศ ได้แก่ แบบทดสอบวัดสติปัญญาของสแตนฟอร์ด-บิเนต แบบทดสอบของเวสเลอร์ แบบทดสอบวัดสติปัญญาของโอติส แบบทดสอบ GRE (Graduate Record Examination Board) ของศูนย์บริการการทดสอบทางการศึกษา (Educational Testing Service : ETS) เป็นต้น แบบทดสอบในประเทศ ได้แก่ แบบทดสอบความถนัดด้านเหตุผล ซึ่งสร้างโดยสถาบันหรือบุคคลผู้สนใจ สามารถสรุปได้ว่ารูปแบบของงานที่กำหนดให้ผู้สอบทำในแบบทดสอบ มีทั้งที่เป็นแบบที่ใช้ภาษา (Verbal Material) คือ ใช้คำ ข้อความ ประโยค และแบบไม่ใช้ภาษา (Non-verbal Material) ซึ่งใช้รูปภาพ แผนภาพ หรือให้ตอบโดยการปฏิบัติงาน สำหรับองค์ประกอบ (Component) ของความสามารถที่สอบวัด ได้แก่ ภาษา (Verbal) ใช้ข้อความถามชนิดคำศัพท์ อุปมาอุปไมย สรุปความ และ

การเติมประโยคให้สมบูรณ์ ด้านจำนวนหรือปริมาณ (Quantitative) ที่วัดความคิดรวบยอดเกี่ยวกับจำนวน ปริมาณ ตัวเลข อนุกรม และการจำแนก (Classification)

✓ ในปี ค.ศ. 1977 ได้เพิ่มการวัดความสามารถเชิงวิเคราะห์ (Analytical Ability) ในแบบทดสอบวัดความสามารถด้านเหตุผลของแบบทดสอบ-GRE ทั้งนี้เพื่อมุ่งขยายขอบเขตของการวัดทักษะการใช้เหตุผล (Reasoning Skills) ให้นอกเหนือไปจากการประเมินโดยการวัดภาษาและปริมาณที่ใช้อยู่ โดยระยะเริ่มแรกสร้างข้อคำถาม 4 ชนิด หลังจากได้ทำการศึกษาทดลองพบว่า ชนิดของข้อคำถาม (Item Type) 2 ชนิด จะได้รับผลของการเตรียมตัวเฉพาะในการสอบ (Powers and Swinton. 1983 : 104 - 115 ; Swinton and Powers. 1984 : 266 - 278) และในปี ค.ศ. 1982 จากผลการศึกษาของสวินตัน (Swinton) ไวด์ (Wild) และวอลล์มาร์ค (Wallmark) พบผลของการฝึกหัด (Practice Effect) ข้อคำถาม 2 ชนิด ดังกล่าวจึงต้องถูกตัดทิ้งไปเนื่องจากวัดได้ค่อนข้างแคบ โดยเน้นหนักเฉพาะทักษะการใช้เหตุผลแบบนิรนัย (Deductive Reasoning Skills) และใช้ข้อคำถาม 2 ชนิด คือ เหตุผลเชิงตรรก (Logical Reasoning : LR) และ เหตุผลเชิงวิเคราะห์ (Analytical Reasoning : AR) เนื่องจากการวัดความสามารถด้านการวิเคราะห์ด้วยข้อคำถาม 4 ชนิด ให้ค่าองค์ประกอบด้านการมีเหตุผลที่ต่างไปจากองค์ประกอบด้านภาษาและปริมาณ (Power and Swinton. 1981 : 141 - 158) จึงมีการศึกษาเพื่อพัฒนาชนิดของข้อคำถามเพื่อการวิเคราะห์เพิ่มขึ้น ด้วยหวังว่าจะทำให้ค่าความเที่ยงตรงเชิงจำแนก (Discriminant Validity) และความเที่ยงตรงเชิงเหมือน (Convergent Validity) ของการวัดเพิ่มขึ้น ต่อมา ในปี ค.ศ. 1991 เอ็มเมอร์ริช และคนอื่น ๆ (Emmerich and others. 1991) ได้พัฒนาข้อคำถามเพิ่มขึ้นอีก 4 ชนิด

✗ สำหรับวัดการวิเคราะห์และทำการประเมินข้อคำถามทั้ง 4 ชนิด ดังกล่าว ข้อคำถาม 4 ชนิด ได้แก่ การวิเคราะห์คำอธิบาย (Analysis of Explanations : AX) ความเห็นตรงกันข้าม (Contrasting Views : CV) เหตุผลเชิงตรรกเกี่ยวกับจำนวน (Numerical Logical Reasoning : NLR) และ การระบุรูปแบบ (Pattern Identification : PI) มีการปรับตัวเลือกจากเดิม 5 ตัวเลือก เป็น 3 ตัวเลือก และ 2 ตัวเลือก (ใช่-ไม่ใช่) ผลการศึกษาพบว่า ชนิดของข้อคำถามที่มี 3 ตัวเลือก และตัวเลือกใช่-ไม่ใช่ ให้ค่าความเชื่อมั่นแบบความคงที่ภายในอยู่ในระดับที่น่าพอใจ ข้อคำถามชนิดวิเคราะห์คำอธิบายแสดงความแตกต่างระหว่างเพศเพียงเล็กน้อย ผลของความเห็นออรัลและเพื่อเกิดขึ้นในข้อคำถามชนิดเหตุผลเชิงตรรกเกี่ยวกับจำนวน ผลจากการฝึกฝน พบว่า

เกิดในข้อคำถามชนิดระบุรูปแบบ ส่วนข้อคำถามชนิดความเห็นตรงกันข้าม พบว่า เหมาะกับการวัดทางภาษามากกว่าการวิเคราะห์ หากพิจารณาเกี่ยวกับรูปแบบข้อคำถาม (Item Format) ข้อคำถามชนิด 3 ตัวเลือก จะเหมาะกับชนิดของข้อคำถามที่ศึกษาทดลองทุกชนิด ส่วนแบบ 2 ตัวเลือก ใช่-ไม่ใช่ พบว่า เหมาะกับข้อคำถามชนิดวิเคราะห์ค่าอธิบาย แต่ไม่เหมาะกับชนิดเหตุผลเชิงวิเคราะห์ (Enright, Tucker and Katz. 1995 : 1)

ผู้วิจัยสนใจที่จะพัฒนากระบวนการวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ด้วยการพัฒนารูปแบบของข้อคำถาม และตรวจสอบรูปแบบข้อคำถามที่เหมาะสม เพื่อใช้ในการศึกษาความสามารถด้านเหตุผลเชิงวิเคราะห์ของเด็กไทยขึ้น โดยทำการศึกษากับนักเรียนระดับชั้นมัธยมศึกษาซึ่งถือเป็นวัยที่กำลังพัฒนาการทางสติปัญญาที่สมบูรณ์ ทั้งนี้เพราะตามทฤษฎีพัฒนาทางสติปัญญาของเพียเจต์ (Piaget. 1972) เชื่อว่า ความเข้าใจในการใช้เหตุผลจะเกิดขึ้นอย่างชัดเจนตั้งแต่อายุ 11 ปี และจะพัฒนาอย่างสมบูรณ์เมื่ออายุประมาณ 15 - 16 ปี คือสามารถคิดกับปัญหาที่เป็นรูปธรรมและนามธรรมได้ สร้างสมมติฐานอย่างสมเหตุสมผล คิดย้อนกลับได้อย่างสมบูรณ์ตามแบบแผนทางตรรกวิทยาของเพียเจต์

จุดประสงค์ของการศึกษา

การศึกษานี้มีจุดมุ่งหมายเพื่อพัฒนาการวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ของนักเรียนชั้นมัธยมศึกษา โดยมีจุดมุ่งหมายดังต่อไปนี้

1. สร้างข้อคำถามเพื่อพัฒนารูปแบบการวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ซึ่งประกอบด้วยชนิดของข้อคำถาม (Item Type) 6 ชนิด และรูปแบบข้อคำถาม (Item Format) 2 รูปแบบ

2. ศึกษาคุณภาพของข้อคำถามและแบบทดสอบด้วยวิธีการต่อไปนี้

- 2.1 การวิเคราะห์ด้วยวิธีการแบบดั้งเดิม (Classical Test Theories) เพื่อศึกษาคุณภาพของข้อคำถามและแบบทดสอบในขั้นตอนของการพัฒนาข้อคำถาม

- 2.2 วิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบ (Item Response Theories) เพื่อกำหนดคุณภาพประจำข้อและแบบทดสอบ

ความสำคัญของการศึกษาค้นคว้า

การศึกษานี้ถือเป็นการเสนอแนวทางของการพัฒนาการวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ในส่วนที่เกี่ยวกับรูปแบบชนิดของข้อคำถาม ทั้งนี้เพื่อเพิ่มคุณภาพด้าน

ความเที่ยงตรงของการวัดให้สูงขึ้น โดยการพัฒนาข้อคำถามวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ที่มีรูปแบบแตกต่างไปจากรูปแบบอุปมาอุปไมย อนุกรม การสรุปความ และเติมความให้สมบูรณ์ ซึ่งใช้กันอยู่ ข้อค้นพบจากการศึกษาจะเป็นเครื่องชี้แนวทางในการตัดสินใจเลือกสร้างและใช้แบบทดสอบได้ตามคุณสมบัติ (Properties) ของแบบทดสอบ อนึ่ง ชนิดของข้อคำถามที่ได้พัฒนาขึ้นเป็นต้นแบบ จะเป็นสิ่งที่ท้าทายให้เกิดความสนใจต่อการพัฒนาโปรแกรมการทดสอบ สำหรับหน่วยงานที่รับผิดชอบและต่อการศึกษาวิจัย

ขอบเขตของการศึกษา

- 1) ประชากรที่ใช้ในการศึกษาวิจัยครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 และ 5 ปีการศึกษา 2538 ที่เรียนตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) และหลักสูตรมัธยมศึกษาตอนปลาย พุทธศักราช 2524 (ฉบับปรับปรุง พ.ศ. 2533)
- 2) ความสามารถด้านเหตุผลเชิงวิเคราะห์ที่ผู้วิจัยทำการศึกษาค้นคว้าครั้งนี้ ถือเป็นความคิดระดับสูงซึ่งเทียบได้กับการคิดอย่างมีวิจารณญาณตามแนวคิดของครูลิค (Krulik. 1993)
- 3) ชนิดของข้อคำถามที่พัฒนาขึ้นเพื่อศึกษา ได้แก่ ชนิดของข้อคำถาม 6 ชนิด ดังนี้
 - 3.1 เหตุผลด้านวิเคราะห์
 - 3.2 การวิเคราะห์ค่าอธิบาย
 - 3.3 ความเห็นตรงกันข้าม
 - 3.4 เหตุผลด้านทางตรรก
 - 3.5 เหตุผลทางตรรกเกี่ยวกับจำนวน
 - 3.6 การระบุรูปแบบ
- 4) รูปแบบของข้อคำถามที่ศึกษา ได้แก่ แบบเลือกตอบ (Multiple Choice)
5 ตัวเลือก และ แบบเติมคำตอบ (Supply Type)

ตัวแปรที่ศึกษา

การศึกษาคคุณภาพของข้อคำถามและแบบทดสอบ ได้กำหนดตัวแปรในการศึกษาไว้ดังนี้

1. ตัวแปรอิสระ ได้แก่
 - 1.1 ชนิดของข้อคำถาม

- 1.2 รูปแบบของข้อคำถาม
2. ตัวแปรตาม ได้แก่ คุณภาพของข้อคำถามและแบบทดสอบในด้าน
 - 2.1 ความเชื่อมั่น
 - 2.2 ความเที่ยงตรง
 - 2.3 ค่าฟังก์ชันสารสนเทศของแบบทดสอบ
 - 2.4 ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ (SE e)
 - 2.5 ค่าประสิทธิภาพสัมพัทธ์

นิยามศัพท์เฉพาะ

①. ความสามารถด้านเหตุผลเชิงวิเคราะห์ หมายถึง ความสามารถในการคิด
พิจารณาถึงความเห็นเกี่ยวกับเหตุการณ์ เรื่องราว หรือสิ่งของต่าง ๆ โดยการค้นหา ทำ
ความเข้าใจ จากข้อมูลหลักฐานที่มีอยู่ เพื่อนำไปสู่การประเมิน วินิจฉัย และสร้างข้อสรุปที่
ถูกต้องสอดคล้องกับหลักการของเหตุผล

2. ชนิดของข้อคำถาม (Item Type) หมายถึง ลักษณะงานที่กำหนดให้ผู้ตอบ
 เพื่อพิจารณาตัดสินจากข้อสนเทศ หลักฐานที่ถูกนำเสนอในลักษณะและรูปแบบที่แตกต่างกัน อาทิ
 สถานการณ์ เรื่องราว แผนภูมิ ภาพ

3. รูปแบบข้อคำถาม (Item Format) หมายถึง รูปแบบงานที่กำหนดให้ผู้สอบทำ
 แบ่งเป็น 2 แบบ คือ แบบเลือกตอบ และแบบเติมคำตอบ

④. ผู้เชี่ยวชาญ ในการศึกษครั้งนี้จำเป็นต้องให้ผู้เชี่ยวชาญพิจารณาให้ความเห็น
ในหลายประเด็น อาทิ การพิจารณาโครงสร้างของคุณลักษณะที่ต้องการวัด ตัวบ่งชี้ ความ
สอดคล้องระหว่างชนิดของข้อคำถามและสิ่งที่ต้องการวัด ซึ่งสามารถจะสะท้อนความสามารถ
ที่ต้องการได้ ผู้เชี่ยวชาญจึงประกอบด้วยกลุ่มบุคคลต่อไปนี้ ผู้บริหารการศึกษา ซึ่งเป็น
 บุคคลที่ได้รับการยอมรับว่ามีความรู้ความสามารถ มีความรู้ในศาสตร์หลายด้านทางการศึกษา
 นักวิชาการด้านคณิตศาสตร์ วิทยาศาสตร์ จิตวิทยา วัฒนธรรม หลักสูตร ซึ่งเป็นผู้สำเร็จการ
 ศึกษาตั้งแต่ระดับปริญญาโทขึ้นไป และหรือปฏิบัติงานที่เกี่ยวข้องมีประสบการณ์จากการศึกษา
 วิจัยมาเป็นเวลาไม่น้อยกว่า 3 ปี รวมทั้งมีผลงานด้านการสร้าง การพัฒนา หรือการเขียน
 เอกสาร แผนแม่บทเป็นที่ประจักษ์

5. ค่าพารามิเตอร์ หมายถึง ค่าที่ได้จากการวิเคราะห์แบบทดสอบวัดความ
 สามารถด้านเหตุผลเชิงวิเคราะห์ทั้งแบบย่อยต่าง ๆ เป็นรายชื่อ ได้แก่ ค่าอำนาจจำแนก (a)

ค่าความยาก (b) และค่าการเดาของข้อสอบ (c)

6. ค่าความยากของข้อสอบ (b) หมายถึง ค่าที่แสดงถึงความสามารถที่จุดใดลักษณะข้อสอบมีความชันมากที่สุด (จุดเปลี่ยนโค้ง) ซึ่งเป็นระดับความสามารถที่จะตอบข้อคำถามนั้นถูกต้องด้วยความน่าจะเป็น 0.5 เมื่อไม่มีการเดา ค่าความยากของข้อสอบนี้ปกติจะมีค่าแปรเปลี่ยนอยู่ระหว่าง -2.0 ถึง +2.0 ค่า b ใกล้เคียง 2.0 แสดงว่า ข้อสอบข้อนั้นยากมาก ค่า b ใกล้เคียง -2.0 แสดงว่า ข้อสอบข้อนั้นง่ายมาก (Hambleton, Swaminathan and Rogers. 1991 : 13)

7. ค่าอำนาจจำแนกของข้อสอบ (a) หมายถึง ค่าที่เป็นสัดส่วนกับความชันของโค้ง หรือความชันของความน่าจะเป็นที่ผู้สอบซึ่งมีระดับความสามารถ θ จะตอบข้อสอบข้อนั้นถูก π จุดเปลี่ยนโค้ง ค่าอำนาจจำแนกนี้มีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติจะมีค่าอยู่ระหว่าง -2.0 ถึง +2.0 ค่า 0 แสดงว่า ข้อสอบไม่มีอำนาจจำแนก ค่า 2.0 แสดงว่า ข้อสอบมีค่าอำนาจจำแนกสูง (Hambleton, Swaminathan and Rogers. 1991 : 15)

8. ค่าการเดาของข้อสอบ (c) หมายถึง ค่าที่แสดงถึงความน่าจะเป็นที่ผู้สอบสามารถตอบข้อสอบข้อนั้นถูกโดยไม่มีความรู้ ($\theta = -\infty$) ค่าการเดาของข้อสอบมีค่าที่เป็นไปได้ตั้งแต่ 0 ถึง 1 ถ้าข้อสอบข้อนั้นไม่สามารถที่จะตอบถูกต้องด้วยการเดาแล้ว ค่าการเดาจะเท่ากับ 0 (Lord. 1980 : 12 - 14) ในทางปฏิบัติจะพิจารณาคัดเลือกข้อสอบที่มีค่าการเดาน้อยกว่า 0.30

9. ค่าฟังก์ชันสารสนเทศของแบบทดสอบ หมายถึง ความแม่นยำในการประมาณค่าความสามารถจริงด้วยคะแนนจากการตอบแบบทดสอบ ซึ่งมีค่าเท่ากับผลรวมอัตราส่วนกำลังสองของความชันของโค้งลักษณะข้อสอบต่อค่าความแปรปรวนของข้อสอบข้อนั้น π ระดับความสามารถ θ ของข้อสอบทุก ๆ ข้อ ในแบบทดสอบ

10. ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ หมายถึง อัตราส่วนระหว่างค่าฟังก์ชันสารสนเทศของแบบทดสอบแต่ละชนิด 2 รูปแบบ คือ แบบเลือกตอบ และแบบเติมคำตอบ ซึ่งค่าที่ได้จะช่วยให้สามารถประเมินและเลือกใช้แบบทดสอบได้เหมาะสมยิ่งขึ้น

11. ค่าความสามารถของผู้สอบ (θ) หมายถึง คะแนนจริงของผู้สอบที่ประมาณได้จากคะแนนรวมของการตอบแบบทดสอบโดยแบบจำลองโลจิสติก 3 พารามิเตอร์ ในการวิจัยครั้งนี้ได้จัดช่วงระดับความสามารถ ($\theta = -4.000$ ถึง $+4.000$) เป็น 3/4 ระดับดังนี้

$$\begin{aligned}
 -1.500 \leq \sigma \leq -4.000 &= \text{ระดับความสามารถต่ำ} \\
 -1.250 \leq \sigma \leq +1.250 &= \text{ระดับความสามารถปานกลาง} \\
 1.500 \leq \sigma \leq +4.000 &= \text{ระดับความสามารถสูง}
 \end{aligned}$$

12. แบบจำลองโลจิสติก 3 พารามิเตอร์ หมายถึง แบบจำลองที่แสดงความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบถูกต้องกับระดับความสามารถในรูปของฟังก์ชันการแจกแจงสะสมแบบโลจิสติก โดยแบบจำลองนี้ประกอบด้วยค่าพารามิเตอร์ของข้อสอบ 3 ค่า คือ ค่าความยาก ค่าอำนาจจำแนก และค่าการเดา

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยครั้งนี้ ผู้วิจัยได้ศึกษาและวิเคราะห์เอกสารที่เกี่ยวข้อง โดยนำเสนอตามลำดับ ดังนี้

1. แนวคิดและทฤษฎีเกี่ยวกับการคิด ความสามารถด้านเหตุผล
 - 1.1 โครงสร้างของความคิด
 - 1.2 ประเภทของแบบการคิด
 - 1.3 แนวทางในการพัฒนาความสามารถด้านการคิด
 - 1.4 ความสามารถด้านเหตุผล
2. แนวคิดเกี่ยวกับการวัดและการประเมินผล
 - 2.1 บทบาทของการวัดและการประเมินผล
 - 2.2 เครื่องมือวัด
 - 2.3 คุณภาพและการตรวจสอบคุณภาพของเครื่องมือ
 - 2.3.1 ความเที่ยงตรง
 - 2.3.2 ความเชื่อมั่น
 - 2.3.3 ความสามารถนำไปใช้ประโยชน์
 - 2.4 ทฤษฎีตอบสนองข้อสอบ
 - 2.5 ดัชนีกำหนดคุณภาพ และการเปรียบเทียบคุณภาพของแบบทดสอบ

1. แนวคิดและทฤษฎีเกี่ยวกับการคิด ความสามารถด้านเหตุผล

ไชท์เวียร์ (Schiever. 1991 : 3) ให้ความเห็นว่า ทว่าความคิดของมนุษย์มีความมหัศจรรย์และซับซ้อน คำอธิบาย และแผนภูมิของกระบวนการคิด หรือทักษะการคิดก็ไม่สามารถจะเข้าไปจับความซับซ้อนหรือสาระของการคิดได้ ดังนั้น ตลอดระยะเวลาที่ผ่านมา นักจิตวิทยา และนักวิจัยได้พยายามที่จะอธิบาย ให้ความหมาย และวิเคราะห์โครงสร้างของความสามารถในการคิดในลักษณะต่าง ๆ กัน ตามความเชื่อหรือทฤษฎีที่บุคคลนั้น ๆ ศึกษา ดังจะเห็นได้จากบางกลุ่มเรียกความสามารถในการคิด การแก้ปัญหาว่า สมรรถภาพสมอง บางกลุ่มเรียกหาปัญญา หรือความถนัด ซึ่งชื่อเรียกดังกล่าวนี้ นำไปสู่นิยามของความคิดในลักษณะที่แตกต่างกัน ดังนั้น บรุนเนอร์ และคนอื่น ๆ (Bruner and others)

และทาบ (Taba) ให้ความหมายสอดคล้องกันว่า การคิดเป็นกระบวนการที่ใช้ในการสร้างความคิดรวบยอด (Concept Formation) ด้วยการจำแนกความแตกต่าง การจัดกลุ่ม และการกำหนดชื่อเรียกเกี่ยวกับข้อความจริงที่ได้รับ และเป็นกระบวนการที่ใช้ในการแปลความหมายข้อมูล รวมถึงการสรุปอ้างอิงด้วยการจำแนกรายละเอียด การเชื่อมโยงความสัมพันธ์ของข้อมูลที่ได้รับ ตลอดจนเป็นกระบวนการเกี่ยวกับการนำกฎเกณฑ์ต่าง ๆ ไปประยุกต์ใช้ได้ อย่างมีเหตุผลและเหมาะสม ขณะที่ฮิลการ์ด (Hilgard. 1962) ให้ความหมายว่า การคิดเป็นพฤติกรรมที่เกิดขึ้นในสมองเนื่องจากกระบวนการใช้สัญลักษณ์แทนสิ่งของ หรือสถานการณ์ต่าง ๆ หรือเป็นกระบวนการที่ภาพหรือสัญลักษณ์ของสิ่งของ หรือสถานการณ์ต่าง ๆ ปรากฏในแนวความคิดหรือจิตใจ แต่สำหรับกิลฟอร์ด (Guilford. 1967) ให้ความเห็นว่า การคิดเป็นการค้นหาหลักการโดยการแยกแยะคุณสมบัติของสิ่งต่าง ๆ หรือข้อความจริงที่ได้รับ แล้วทำการวิเคราะห์เพื่อหาข้อสรุปอันเป็นหลักการของข้อความจริงนั้น ๆ รวมถึงการนำหลักการดังกล่าวไปใช้ในสถานการณ์ที่ต่างไปจากเดิม (Generalization) นอกจากนี้ เพียเจต์ (Piaget. 1975) ให้ทัศนะเกี่ยวกับการคิดของบุคคลว่า เป็นกระบวนการที่มี 2 ลักษณะ คือ กระบวนการปรับเปลี่ยนโครงสร้าง (Accommodation) ซึ่งเป็นการดัดแปลงหรือปรับปรุงโครงสร้างที่เกิดจากประสบการณ์เดิมของตนให้เข้ากับสถานการณ์ใหม่ และการปรับเข้ากับโครงสร้าง (Assimilation) โดยการจัดสิ่งเร้าหรือข้อความจริงที่ได้รับเข้าเป็นพวกเดียวกับประสบการณ์เดิมของตน จะเห็นว่าจากทัศนะดังกล่าวข้างต้น มองความคิดในลักษณะที่เป็นกระบวนการอันนำไปสู่ทางแก้ปัญหา (สมเจตน์ ไวยากรณ. 2530; อ้างอิงมาจาก Hilgard. 1962 : 336)

อย่างไรก็ตาม มีมุมมองความคิดในลักษณะที่ต่างออกไป เช่น ครูลิค และรุดนิค (Krulik and Rudnick. 1993 : 1) ที่เห็นว่า "ความคิด" เป็นความสามารถ (Ability) ที่จะเข้าถึง หรือนำไปสู่ข้อสรุปที่ถูกต้องจากข้อมูลที่กำหนดให้ เด็กต้องสร้างความคิดเกี่ยวกับคุณสมบัติเชิงนามธรรมจากความสัมพันธ์ในสถานการณ์ของปัญหา จากนั้นจึงตรวจสอบความถูกต้อง และอธิบายยืนยันข้อสรุปของเขา ข้อสรุปนี้จะถูกรวมไว้ในรูปของความคิดใหม่ (New Idea)

ดังนั้น อาจกล่าวสรุปได้ว่า ความคิดในทัศนะของนักจิตวิทยาและนักวิจัยที่กล่าวมามีลักษณะที่เป็นทั้งกระบวนการ และความสามารถหรือทักษะ

1.1 โครงสร้างของความคิด

เมื่อการวัดทางจิตวิทยาได้พัฒนาขึ้นในครั้งแรกของศตวรรษ มีกลุ่มความคิดเกี่ยวกับความสามารถของมนุษย์ 2 กลุ่ม คือ สเปียร์แมน (Spearman) กล่าวถึงทฤษฎีที่ประกอบด้วย 2 องค์ประกอบ คือ องค์ประกอบทั่วไป (General Factor) และองค์ประกอบเฉพาะ (Specific Factor) องค์ประกอบ "g" ได้รับความสนใจจากนักจิตวิทยามาก ข้อเสนอประการหนึ่ง คือ ความแตกต่างของบุคคลใน "g" ก็คือ ความแตกต่างของบุคคลในระดับของพลังงานทางความคิดที่จะนำมาใช้ในการปฏิบัติงานทางสติปัญญา ในทางตรงกันข้ามและด้วยวิธีการทางสถิติที่แตกต่างกัน เทอร์สโตน (Thurstone) ระบุว่า สมรรถภาพสมองที่เป็นพื้นฐาน (Primary Mental Abilities : PMA) มีอยู่ประมาณ 7 สมรรถภาพ (Sternberg. 1985 : 1 - 5)

ทั้งนี้ องค์ประกอบทั่วไป "g" ของสเปียร์แมนได้ถูกแบ่งเป็นความสามารถย่อย คือ ความถนัดทางภาษา และความถนัดเชิงปริมาณ ซึ่งมีความสอดคล้องกันอย่างมากในเชิงสถิติกับ PMA ของเทอร์สโตน ที่มีความสัมพันธ์กับความสามารถทางวิชาการ ยิ่งไปกว่านั้น สำหรับแบบทดสอบหลักที่รวมการวัดความถนัด และ SAT เข้าด้วยกัน ซึ่งค่าสหสัมพันธ์ระหว่างส่วนของภาษาและปริมาณ มีค่าประมาณ 0.7 ซึ่งชี้ให้เห็นความสามารถพื้นฐานที่อาจถูกเรียกว่า ความถนัดเชิงวิชาการทั่วไป (General Academic Aptitude) หรือความสามารถเชิงเหตุผล (Reasoning Ability) ที่แสดงให้เห็นจากคะแนนรวมของ SAT สมรรถภาพสมองพื้นฐานของเทอร์สโตนซึ่งยังคงอยู่ คือ การวัดความรู้เกี่ยวกับมิติ ได้แก่ ความสามารถที่ต้องการในกลุ่มทันตแพทย์ในการมองเห็นฟันจากมุมต่าง ๆ ซึ่งรวมอยู่ใน Dental Admission Test และการวัดการคิดแบบไดเวอร์เจนซ์ (Divergent Thinking) หรือความคล่องแคล่ว เช่นที่ใช้โดยกิลฟอร์ด (Guilford. 1959) และทอร์เรน (Torrance. 1962) เพื่อทดสอบความคิดสร้างสรรค์มากกว่าทักษะในการรอบรู้ และการนำไปใช้โดยตรงกับเนื้อหาทางวิชาการ (Whimbey. 1984 : 67)

ในความเห็นของสเตอร์นเบิร์ก และคนอื่น ๆ สติปัญญาเป็นส่วนหนึ่งของความสามารถในการจะได้มาซึ่งความสำเร็จในชีวิต จากการศึกษาทฤษฎีของสติปัญญาในเชิงของศักยภาพในการแก้ปัญหา (Implicit Theory) ชี้ให้เห็นว่า สติปัญญาในสภาพแวดล้อมวัฒนธรรมทางสังคมของสหรัฐอเมริกาจะรวมทักษะ 3 กลุ่ม หรือ 3 องค์ประกอบหลัก คือ ความสามารถในการแก้ปัญหา (Problem-Solving Ability) ความสามารถทางภาษา (Verbal Ability) และความสามารถทางสังคม (Social Competence) ในเรื่องความสามารถ

ในการแก้ปัญหา นั้น เป็นความสามารถในการให้เหตุผลตามหลักการของเหตุผลได้อย่างดี การหาความเกี่ยวข้องระหว่างความคิดมองเห็นลักษณะทั้งหมดของปัญหา เปิดใจกว้าง ตอบสนองความคิดของผู้อื่นด้วยความคิด คิดตามสถานการณ์ได้ดี เข้าถึงหัวใจของปัญหา มีความรู้ข้อมูลข่าวสารอย่างแม่นยำ ตัดสินใจดี พึงข้อโต้แย้งทุกด้าน และเข้าจัดการกับปัญหาด้วยการคิดหาหนทาง ความสามารถทางภาษานั้นแสดงออกในทางพูดจาชัดเจน คล่อง ปราศรัย สันทนาได้ดี มีความรู้เฉพาะสาขา เรียนหนัก อ่านด้วยความเข้าใจสูง-กว้าง เข้ากับบุคคลได้อย่างดี เขียนได้โดยปราศจากความสงสัยยาก ใช้คำศัพท์ได้ดี ยอมรับมาตรฐานของสังคม และพยายามในสิ่งใหม่ สำหรับความสามารถทางสังคม ได้แก่ การยอมรับผู้อื่นตามที่เขาเป็น แสดงความสนใจโลกในวงกว้าง ตรงต่อเวลา มีความสำนึกทางสังคม คิดก่อนพูด-ทำ แสดงความอสุยยากเห็น ตัดสินอย่างยุติธรรม รับผิดชอบต่อความต้องการของบุคคลอื่น ใจกว้าง ขี้สัจย์ต่อตนเอง และผู้อื่น แสดงความสนใจในสภาพแวดล้อมใกล้ตัว

เมื่อนำไปให้บุคคลตอบแล้วนำมาแจกแจง พฤติกรรมที่ถูกจัดว่าเป็นสติปัญญา ได้แก่ ความสามารถในการให้เหตุผลตามหลักเหตุผลได้อย่างดี อ่านอย่างกว้างขวาง และอ่านด้วยความเข้าใจสูง จากผลดังกล่าวจะเห็นว่า ความสามารถด้านเหตุผลถือเป็นดัชนีบ่งชี้ตัวหนึ่งที่แสดงความเป็นผู้มีสติปัญญา (Sternberg. 1985 : 58 - 60)

เนื่องจากองค์ประกอบการแก้ปัญหาจะเหมือนกับสิ่งที่แคเทิล (Cattell. 1971) และฮอร์น (Horn. 1968) เรียกว่า Fluid Ability และ Fluid Ability จะวัดได้ดีโดยแบบทดสอบ Reasoning Item เช่น อุปมาอุปไมย (Analogies) และอนุกรม (Series Completions) ความคิดในปัจจุบันแบบทดสอบวัด Fluid Ability มีแนวโน้มเน้นความสามารถที่จะเข้าไปจัดการกับสิ่งใหม่ ๆ สำหรับ Crystallized Abilities วัดได้โดยแบบทดสอบความเข้าใจในการอ่านและคำศัพท์ และมีแนวโน้มเน้นความเป็นอัตโนมัติของกระบวนการระดับสูง (Automatization of High-Level Processes) (Sternberg. 1985 : 77 - 78)

Fluid Abilities เทียบได้กับพฤติกรรมการแก้ปัญหาทั้ง Inductive Reasoning และ Deductive Reasoning เป็นความสามารถด้านนี้ ส่วน Crystallized Abilities มีลักษณะที่มีความรู้เป็นหลัก (Knowledge Based)

ตามความเห็นของกัตตแมน (Guttman. 1965) ซึ่งเป็นผู้เสนอโครงสร้างแบบวงกลม (Radex Structure) สำหรับสติปัญญา เห็นว่าแบบทดสอบแต่ละฉบับที่วัดสติปัญญาสามารถจับมาวาง ณ ที่ใดของวงกลมก็ได้ แบบทดสอบที่อยู่ใกล้จุดศูนย์กลางของวงกลมจะวัด

ความสามารถที่เป็นจุดหลักของสติปัญญา เพราะฉะนั้นแบบทดสอบที่วัดสติปัญญาโดยแท้ จะต้องเป็นจุดศูนย์กลางของวงกลม ความสามารถในการแก้ปัญหาจำเป็นต้องอาศัยองค์ประกอบหลายประการเพื่อช่วยให้การแก้ปัญหาประสบผลสำเร็จ (Sternberg. 1985 : 7) ส่วนจอห์นสันและไรซิง (Johnson and Rising. 1969 : 107 - 110) มีความเห็นว่า กระบวนการแก้ปัญหาเป็นกระบวนการทางสมองที่ซับซ้อน ซึ่งประกอบด้วย

1. การมองเห็นภาพ (Visualizing)
2. การจินตนาการ (Imagining)
3. การจัดท่าอย่างมีทักษะ (Manipulation)
4. การวิเคราะห์ (Analyzing)
5. การสรุปในเชิงนามธรรม (Abstracting)
6. การเชื่อมโยงความคิด (Association Ideas)

ออสซูเบล (Ausubel. 1968 : 581) กล่าวว่า องค์ประกอบที่ทำให้บุคคลแตกต่างกันในการแก้ปัญหา แบ่งออกได้ 3 ประการ คือ

1. ความรู้ในเนื้อหาวิชาและความคุ้นเคยในการคิดเกี่ยวกับเรื่องนั้น
2. การใช้ "แบบการคิด" ที่เกี่ยวข้องการแก้ปัญหา และความรู้ทั่วไปเกี่ยวกับวิธีแก้

ปัญหาที่มีประสิทธิภาพ

3. คุณลักษณะทางบุคลิกภาพ เช่น แรงจูงใจ ความมั่นคงในอารมณ์ ความวิตกกังวล เป็นต้น

ในการวัดทางจิตวิทยาปัญหาของการสร้างโครงสร้างถือเป็นสิ่งสำคัญ เพราะจะทำให้การวัด Fluid Intelligence ในส่วนเฉพาะ และสติปัญญาทั่วไป เช่น

1. ทฤษฎีสปีร์แมน (1923) เกี่ยวกับเหตุผลเชิงอุปมา ให้หลักสำคัญสำหรับทฤษฎีที่เกี่ยวกับสติปัญญาทั่วไป

2. เทอร์สตัน (1938) ใช้การอุปมาด้วยภาษาและรูปภาพเป็นพื้นฐาน สำหรับตรวจสอบความถูกต้องของทฤษฎี Primary Mental Abilities

3. กิลฟอร์ด (1962) ใช้อุปมาอุปไมยหลายชนิด วัดจำนวนของโครงสร้างของความสามารถทางสติปัญญาที่เป็นองค์ประกอบของโมเดลของเขา

ดังนั้น จะเห็นว่า ทฤษฎีการวัดทางจิตวิทยาส่วนใหญ่จะมุ่งเน้นอุปมาอุปไมย (Analogies) ในฐานะเป็นพื้นฐานของการประเมิน

ครูลิค และรุดนิก (Krulik and Rudnick. 1993 : 3 - 5) ถือว่า การคิดประกอบด้วยองค์ประกอบ 4 องค์ประกอบ ที่มีลักษณะเป็นความสามารถหรือทักษะตามลำดับขั้น จากต่ำไปสูง คือ การคิดในระดับการระลึก (Recall Thinking) จะรวมทักษะการคิดที่มีธรรมชาติเกือบเป็นอัตโนมัติ เป็นความสามารถในการระลึกข้อเท็จจริง การคิดขั้นพื้นฐาน (Basic Thinking) เป็นความเข้าใจความคิดรวบยอดอันเป็นประโยชน์ต่อการนำไปใช้ในการแก้ปัญหาในชีวิตประจำวัน และในโรงเรียน การคิดอย่างมีวิจารณญาณ (Critical Thinking) เป็นความคิดที่ใช้ในการตรวจเชื่อมโยงและประเมินลักษณะทั้งหมดของทางแก้ปัญหาหรือปัญหา ประกอบด้วยทักษะย่อย ได้แก่ การมุ่งเน้นไปในส่วนหนึ่งของข้อมูลในปัญหาหรือสถานการณ์ที่เผชิญอยู่ การตรวจสอบความถูกต้องและวิเคราะห์ข้อมูล การจำ และเชื่อมโยงข้อมูลที่เพิ่งได้รับการเรียนรู้ เพื่อกำหนดคำตอบที่มีเหตุผลและความคิดสร้างสรรค์ (Creative Thinking) เป็นความคิดที่เป็นต้นฉบับที่ทำให้เกิดผลผลิตที่ซับซ้อน ความคิดในระดับนี้ เป็นสิ่งประดิษฐ์ที่คิดหรือจินตนาการขึ้นเอง ประกอบด้วยทักษะย่อย ได้แก่ การสังเคราะห์ความคิด การสร้างความคิด และการนำความคิดไปใช้ เพื่อหาประสิทธิภาพของความคิดใหม่ที่สร้างขึ้น

การคิดอย่างมีวิจารณญาณ นั้น ในความเห็นของซีเกล (Siegel. 1988 : 38 - 39) ถือเป็นองค์ประกอบในการประเมินเชิงเหตุผล ผู้มีความคิดวิจารณ์จะต้องสามารถประเมินเหตุผลและความสามารถของเขา เพื่อประกันความเชื่อ การกล่าวอ้าง และการปฏิบัติที่เหมาะสม ดังนั้น ผู้มีความคิดวิจารณ์ต้องมีความเข้าใจที่ดี และมีความสามารถในการเลือกชนิดของหลักการที่เหมาะสมต่อการประเมิน มีการปฏิบัติ และตัดสินใจโดยยึดเหตุผล นอกจากจะมีความสามารถในการประเมินเหตุผลแล้ว ผู้มีความคิดวิจารณ์จะต้องมีเจตคติเชิงวิจารณ์ (Critical Attitude) หรือ Critical Spirit ด้วย โดยต้องมีความเต็มใจที่จะทำตามการตัดสินใจและหลักการ ต้องยอมรับและนับถือเหตุผล ในลักษณะที่เรียกว่า ความรักเหตุผล (A Love of Reason) ซึ่งในประเด็นนี้ เราไม่เพียงต้องการให้นักเรียนของเราสามารถใช้เหตุผลได้อย่างดีเท่านั้น แต่ต้องการให้เขาปฏิบัติได้อย่างจริงจังด้วย

แมคเพค (Mc Peck) เห็นว่า การคิดอย่างมีวิจารณญาณมีลักษณะเป็น Generalized Skill หรือ Abilities หรือชุดของทักษะ หรือความสามารถ ซึ่งใช้หรือประยุกต์กับสถานการณ์หรือสภาพแวดล้อมต่าง ๆ การคิดจำเป็นจะต้องมีสิ่งที่คิด ผู้สอนจะต้องช่วยให้ผู้เรียนพัฒนาทักษะการใช้เหตุผล (Reasoning Skills) ซึ่งจะเป็นลักษณะพื้นฐานทั่วไป และ

สามารถนำไปประยุกต์กับสถานการณ์และเนื้อหาต่าง ๆ ได้ ทั้งนี้เราได้เสนอว่าองค์ประกอบของความคิดรวบยอดเกี่ยวกับการคิดอย่างมีวิจารณญาณ คือ ความสามารถในการประเมินเหตุผลได้อย่างเหมาะสม ความพอใจในการใช้เหตุผล นอกจากนี้ แมคนพได้เน้นบทบาทของเหตุผลที่ดีที่มีต่อความคิดอย่างมีวิจารณญาณ ผู้มีความคิดวิจารณ์จะก้าวไปโดยใช้เหตุผลจะมองหาเหตุผลเพื่อเป็นพื้นฐานในการตัดสินใจและกระทำ ฉะนั้น จึงถือว่าการคิดอย่างมีวิจารณญาณจึงเป็นคุณลักษณะที่เป็นเลิศอันเป็นที่ต้องการ สำหรับกิจกรรมทางการศึกษาเป็นอุดมคติทางการศึกษา (Educational Ideal) ที่จำเป็นต้องมุ่งเข้าไปหา (Siegel. 1988 : 38)

ในการพิจารณาตัดสินหรือลงความเห็นเกี่ยวกับเรื่องใด ๆ บุคคลจำเป็นต้องตกลงใจว่าหลักฐานชนิดใดที่จะนำมาพิจารณา จะให้น้ำหนักหลักฐานแต่ละชนิดอย่างไร จะรวบรวมหลักฐานชนิดต่าง ๆ อย่างไร และจะแปลข้อสรุปของเขาไปสู่ความเป็นไปได้ตามสมมุติฐานเป้าหมายที่ตั้งไว้ เพื่อความสอดคล้องกับผลที่ได้จริงตามเป้าหมายอย่างไร (Sternberg. 1985 : 176)

1.2 ประเภทของแบบการคิด

โดยที่แบบการคิด เป็นกระบวนการหรือวิธีการคิดที่บุคคลใช้ในการรับรู้สภาพที่เป็นสิ่งเร้าต่าง ๆ เพื่อนำไปสู่การเรียนรู้ในสถานการณ์ของสิ่งเร้า นั้น บุคคลจะเลือกใช้แบบการคิดต่างกันขึ้นอยู่กับประสบการณ์การเรียนรู้ของแต่ละบุคคล

มีผู้ศึกษาเกี่ยวกับ "แบบการคิด" ไว้ต่าง ๆ กัน ตามลักษณะของวิธีการที่ใช้ศึกษาทั้งนี้วิธีการศึกษาและการแบ่งแบบการคิดที่นิยมใช้ในการศึกษามี 2 รูปแบบ คือ ตามแนวทางของวิทคิน และคนอื่น ๆ กับการแบ่งแบบการศึกษาตามแนวทางของ เดนกัน มอสส์ และซีเกิล

แบบการคิดตามแนวของวิทคิน และคนอื่น ๆ (Witkin and others. 1977 : 39) ซึ่งได้ศึกษาแบบการคิดจากการทดลองที่ใช้แบบทดสอบ Body Adjustment Test (B.A.T.) จากการทดลอง วิทคินแบ่งแบบการคิดของบุคคลเป็น 2 แบบ คือ

1. การคิดแบบไม่ขึ้นกับสภาพรอบข้าง (Field Independence) เป็นการคิดที่ยึดถือเอาตนเองเป็นศูนย์กลาง การคิดแบบนี้ไม่อาศัยข้อมูลจากภายนอกหรือจากสภาพรอบข้างแต่จะตัดสินปัญหาด้วยเหตุผล โดยยึดสิ่งที่ปรากฏอยู่จริงในสิ่งเร้าเป็นเกณฑ์

2. การคิดแบบขึ้นอยู่กับสภาพรอบข้าง (Field Dependence) เป็นการศึกษาที่บุคคลอาศัยข้อมูลจากภายนอกหรือจากสภาพรอบข้าง โดยอาศัยความรู้และประสบการณ์ของตนเป็นส่วนประกอบช่วยในการตัดสินใจปัญหาโดยไม่ยึดเอาตนเองเป็นศูนย์กลาง

การคิด 2 แบบ นี้ เกิดจากความเชื่อที่ว่า กิจกรรมทางสมองประกอบด้วยกระบวนการ 2 กระบวนการ คือ การอาศัยข้อมูลจากภายนอก และอาศัยข้อมูลจากภายในที่ได้สัมผัสไว้

จากแนวความคิดนี้ โคทส์ (Coates. 1972 : 1 - 8) ได้นำมาสร้างแบบทดสอบชนิดซ้อนภาพ เพื่อศึกษาแบบการคิดของเด็กปฐมวัย เรียกว่า The Preschool Embedded 20 Figures Test (PEFT) ใช้วิธีการทดสอบโดยการให้เด็กหารูปสามเหลี่ยมด้านเท่า ซึ่งมีลักษณะและขนาดเหมือนกับรูปสามเหลี่ยมในภาพตัวอย่างที่กำหนดให้ แบบทดสอบนี้ไม่ต้องอาศัยความเข้าใจทางภาษา

การแบ่งแบบการคิดตามแนวทางของ เคนแกน มอสส์ และซีเกล (Wallach and Kagan. 1956 : 105) ซึ่งใช้แบบทดสอบที่เรียกว่า แบบทดสอบการคิด (Conceptual Style Test) มีลักษณะเป็นภาพลายเส้นขาวดำ แต่ละข้อประกอบด้วยภาพ 3 ภาพ ให้ผู้สอบเลือกภาพ 2 ภาพ ที่ไปด้วยกันได้หรือเข้าคู่กันได้ พร้อมกับให้เหตุผลว่า เพราะเหตุใดจึงเลือกเช่นนั้น เหตุผลที่นำมาจะแสดงถึง "แบบการคิด" ของผู้สอบ ซึ่งเคนแกนได้แบ่งลักษณะของการคิดออกเป็น 3 ประเภท คือ

1. แบบวิเคราะห์เชิงบรรยาย (Descriptive-Analytic Style) เป็นการศึกษาที่จะรวมวัตถุเข้าเป็นพวกเดียวกัน โดยพิจารณาความคล้ายคลึงของลักษณะทางกายภาพที่วัตถุต่าง ๆ มีร่วมกัน ได้แก่ ลักษณะในเรื่อง สี รูปร่าง ขนาด

2. แบบโยงความสัมพันธ์ (Relation Style) เป็นการศึกษาที่จะรวมวัตถุเข้าเป็นพวกเดียวกัน โดยอาศัยความสัมพันธ์ในด้านเวลาหรือสถานที่

3. แบบอ้างอิงจำแนกประเภท (Categorical Inferential Style) การคิดประเภทนี้จะรวมวัตถุเข้าเป็นพวกเดียวกัน โดยพิจารณาสิ่งอื่นที่ร่วมกันเป็นส่วนรวม เช่น หน้าที่และอื่น ๆ ที่นอกเหนือจากลักษณะทางกายภาพ วัตถุในกลุ่มถือว่าเป็นอิสระต่อกัน

การแบ่งแบบการคิดในลักษณะนี้ เกิดจากความเชื่อที่ว่า กิจกรรมทางสมองประกอบด้วยกระบวนการต่าง ๆ 3 ประการ คือ (Kagan and others. 1963 : 433)

1. การอาศัยข้อมูลจากภายนอก
2. การอาศัยข้อมูลจากภายในที่ได้สัมผัสไว้

3. การผสมผสานเชื่อมโยงข้อมูลที่ได้สะสมไว้

นอกจากการแบ่งแบบการคิดเป็น 3 ประเภท ดังกล่าวแล้ว ได้มีการนำหลักการนี้ไปศึกษาโดยแบ่งแบบการคิดเป็นแบบย่อย ๆ ได้ 5 แบบ ดังนี้

1. แบบการคิดแบบวิเคราะห์ (Analytic Style) เป็นการจัดประเภทสิ่งเร้าตามความเหมือนของลักษณะทางกายภาพของสิ่งเร้า นั้น ส่วนประกอบทางกายภาพในที่นี้หมายถึง ส่วนประกอบที่ประกอบขึ้นเป็นโครงสร้างของสิ่งเร้า เช่น ถ้าเสนอสิ่งเร้าเป็นรูปเก๋อี้ โต๊ะ แจกันดอกไม้ พวกที่มีแบบการคิดประเภทนี้จะเลือกจับคู่รูปเก๋อี้กับโต๊ะ โดยให้เหตุผลว่าต่างก็ทำด้วยไม้เหมือนกัน

2. แบบการคิดแบบบรรยาย (Descriptive Style) เป็นการจัดประเภทของสิ่งเร้าตามลักษณะรวมทางกายภาพของสิ่งเร้า นั้น เช่น การบรรยายภาพของวัตถุ ที่ทางของสิ่งเร้า หรือสัตว์ หรือสิ่งของ ที่ปรากฏในสิ่งเร้า ตัวอย่างเช่น ถ้าเสนอภาพ 3 ภาพ ดังเสนอในข้อ 1 ผู้มีแบบการคิดนี้จะเลือกจับคู่ภาพเก๋อี้กับโต๊ะ โดยให้เหตุผลดังกล่าวว่า ต่างก็มีขา 4 ขา เหมือนกัน

3. แบบการคิดแบบจำแนกประเภท (Categorical Style) เป็นการคิดที่จัดประเภทของสิ่งเร้าเข้าเป็นหมวดหมู่ โดยอาศัยประสบการณ์หรือความรู้ที่ได้รับมาเป็นเครื่องตัดสิน เป็นการพิจารณาโดยไม่คำนึงถึงความคล้ายคลึงทางด้านรูปร่าง แต่จะดูที่คุณสมบัติบางประการที่มีร่วมกันอยู่ ตัวอย่างจากภาพทั้ง 3 ดังเสนอ พวกที่มีแบบการคิดแบบนี้จะเลือกจับคู่ภาพเก๋อี้กับโต๊ะ โดยให้เหตุผลว่าต่างก็เป็นเครื่องใช้เหมือนกัน

4. แบบการคิดแบบอ้างสรุป (Inferential Style) เป็นการคิดที่จัดประเภทของสิ่งเร้าเข้าเป็นหมวดหมู่ตามลักษณะหน้าที่ของสิ่งเร้า นั้น หรือจัดตามลักษณะทางอารมณ์ของสิ่งเร้า ตัวอย่างจากสิ่งเร้าที่เป็นรูป เก๋อี้ โต๊ะ และแจกันดอกไม้ พวกที่คิดแบบนี้จะเลือกจับคู่เก๋อี้กับโต๊ะ โดยให้เหตุผลว่าใช้วางสิ่งของได้เหมือนกัน

5. แบบการคิดแบบการหาความสัมพันธ์ (Relational Style) เป็นแบบการคิดที่จัดประเภทของสิ่งเร้า โดยพยายามหาความสัมพันธ์เชื่อมโยงสิ่งเร้าต่าง ๆ ให้สัมพันธ์กัน โดยคำนึงถึงหน้าที่หรือความสัมพันธ์ระหว่างสิ่งเร้า นั้น ๆ เช่น ถ้าเสนอสิ่งเร้าเป็นรูปเก๋อี้ โต๊ะ และแจกันดอกไม้ พวกที่คิดแบบนี้จะเลือกจับคู่โต๊ะกับแจกันดอกไม้ โดยให้เหตุผลว่าแจกันดอกไม้ต้องตั้งไว้บนโต๊ะ (ชาลี อุปกัย 2523 : 13 - 15; อ้างอิงมาจาก Nuanpen Kosolserth. 1964 : 3 - 8)

แบบการคิดทั้ง 5 แบบนี้ อาศัยหลักการแบ่งเช่นเดียวกับของ เคนน และคนอื่น ๆ เพียงแต่แบ่งแยกให้แคบลงอีกเท่านั้น ดังจะเห็นได้ว่า แบบการคิด 2 แบบแรก คือ แบบวิเคราะห์ และแบบบรรยาย เป็นการแยกย่อยมาจากการคิดแบบวิเคราะห์เชิงบรรยาย ส่วนแบบจำแนกประเภท และแบบอ้างสรุป เป็นการแยกย่อยมาจากการคิดแบบจำแนกประเภทเชิงอ้างอิง ส่วนการคิดแบบหาความสัมพันธ์มีลักษณะเหมือนกับการคิดแบบโยงความสัมพันธ์ของ เคนน และคนอื่น ๆ ดังกล่าวแล้ว

โรสแมน (Rossman. 1966) ศึกษาการคิดแบบวิเคราะห์ของนักเรียนชั้นประถมศึกษาปีที่ 1 จำนวน 12 คน ประถมศึกษาปีที่ 2 จำนวน 20 คน โดยใช้แบบทดสอบซ่อนรูป (Hidden Figures Test) แบบทดสอบของวิทกิน (Witkin Embedded Figure Test) แบบทดสอบการก่อรูปความคิด (Conceptual Formation Test) และแบบทดสอบสัมพันธ์คำ (Word Association Test) พบว่า นักเรียนชั้นประถมศึกษาปีที่ 2 คิดแบบวิเคราะห์มากกว่าชั้นประถมศึกษาปีที่ 1 และการคิดแบบวิเคราะห์มีแนวโน้มที่จะเพิ่มขึ้นตามอายุเช่นเดียวกับผลงานของเคนน และคนอื่น ๆ

เมื่อพิจารณาแบบการคิดกับความสำเร็จในการทำงานและการเรียน เคนน และคนอื่น ๆ (Ausuble and others. 1968) พบว่า เด็กที่คิดแบบวิเคราะห์มีแนวโน้มที่จะคิดอย่างรอบคอบมากกว่าเด็กที่คิดแบบจำแนกประเภท และสามารถจะวิเคราะห์วัตถุที่เห็นได้ดีกว่าการคิดแบบอื่น

1.3 แนวทางในการพัฒนาความสามารถด้านการคิด

จากนิยามความหมายของการคิด ซึ่งมีลักษณะทั้งที่เป็นกระบวนการและความสามารถ อาจกล่าวได้ว่า บุคคลสามารถพัฒนาทักษะหรือความสามารถในการคิดของตนได้ ดังที่ ทาบว (Taba. 1965) ได้กล่าวไว้ในข้อตกลงเบื้องต้นของยุทธศาสตร์การสอนของเขาว่า ความคิดเป็นสิ่งที่เรียนรู้ได้ และสามารถพัฒนาการเรียนรู้ไปได้ ดังนั้น การคิดจึงสอนได้ (Schiever. 1991 : 138) ซึ่งสอดคล้องกับแวลท์เลท (Valett. 1978 : 27) ที่กล่าวว่าสรุปว่า บุคคลสามารถเรียนรู้ที่จะพัฒนาทักษะการคิดและความสามารถของเขา นอกจากนี้ทักษะการคิด (Cognitive Skills) และกระบวนการคิด (Thinking Processes) มี แบบแผนการพัฒนาที่ต่างกัน และต้องการเวลาที่เหมาะสมและเป็นลำดับ สำหรับการพัฒนาไปสู่จุดสูงสุด ทั้งนี้ ไชล์เวอร์ ได้สรุปว่า บุคคลเรียนรู้ที่จะคิดในวิธีการที่แตกต่างกัน และการพัฒนาหรือแก้ไขทักษะการคิดที่บกพร่องจะดีที่สุด เมื่อกระทำผ่านลักษณะเฉพาะของบุคคล

สอนในช่วงเวลาที่ยาวนาน และเด็กเรียนรู้ที่จะคิดได้ดีเพียงใดนั้น เป็นสัดส่วนโดยตรงกับ
โครงการสอนเฉพาะ เวลา และความพยายามที่ใช้

เนื่องจากบุคคลมีแบบ (Style) การเรียนรู้และการแก้ปัญหาต่างกัน อันเนื่องมา
จากความแตกต่างทางวัฒนธรรมและการศึกษา คุณภาพและรูปแบบของการคิดจะขึ้นอยู่กับ
พันธุกรรมและชีววิทยาพื้นฐานของร่างกายมนุษย์ จะเห็นได้จากที่เพียเจต์ได้อธิบายการ
เจริญเติบโต และพัฒนาการของความสามารถทางการคิด ในขั้นตอนที่สำคัญหลายขั้นตอน
(Varett. 1978 : 33; citing Inhelder and Piaget. 1958) ด้วยวุฒิภาวะและ
ประสบการณ์ โครงสร้างทางความคิดเหล่านี้จะปรากฏอย่างเป็นธรรมชาติในชีวิตของเด็กและ
แต่ละขั้นของการพัฒนา และจะส่งผลต่อความสามารถในการคิดและการแก้ปัญหาที่ซับซ้อนขึ้น

ความเข้าใจเกี่ยวกับประเภท รูปแบบ พัฒนาการ และโครงสร้างของการคิดความ
สามารถด้านเหตุผล จะนำไปสู่กระบวนการพัฒนาการคิดที่มีประสิทธิภาพ

ในเรื่องของประเภท หรือลักษณะการคิด กาลเอ่ (สมเจตน์ ไวยากรณ์. 2530 :
13; อ้างอิงมาจาก Gagne. 1970 : 283) ได้จำแนกไว้ 2 ประเภท คือ

1. การคิดอย่างไม่มีทิศทาง เป็นการคิดจากสิ่งที่พบเห็นจากประสบการณ์ตรงจาก
สิ่งที่ได้ยินหรือได้ฟังมา เป็นการคิดต่อเนื่อง ได้แก่ การคิดต่อเนื่องอย่างอิสระถึงเหตุการณ์
ที่ล่วงมาแล้ว การคิดที่ถูกควบคุมโดยอาศัยคำสั่งเป็นแนว การคิดผ่นในขณะที่ยังตั้งอยู่ที่เรียกว่า
ผ่นกลางวัน การคิดผ่น การคิดหาเหตุผลเข้าข้างตนเอง โดยอาศัยความเชื่อหรืออารมณ์
ของผู้คิด

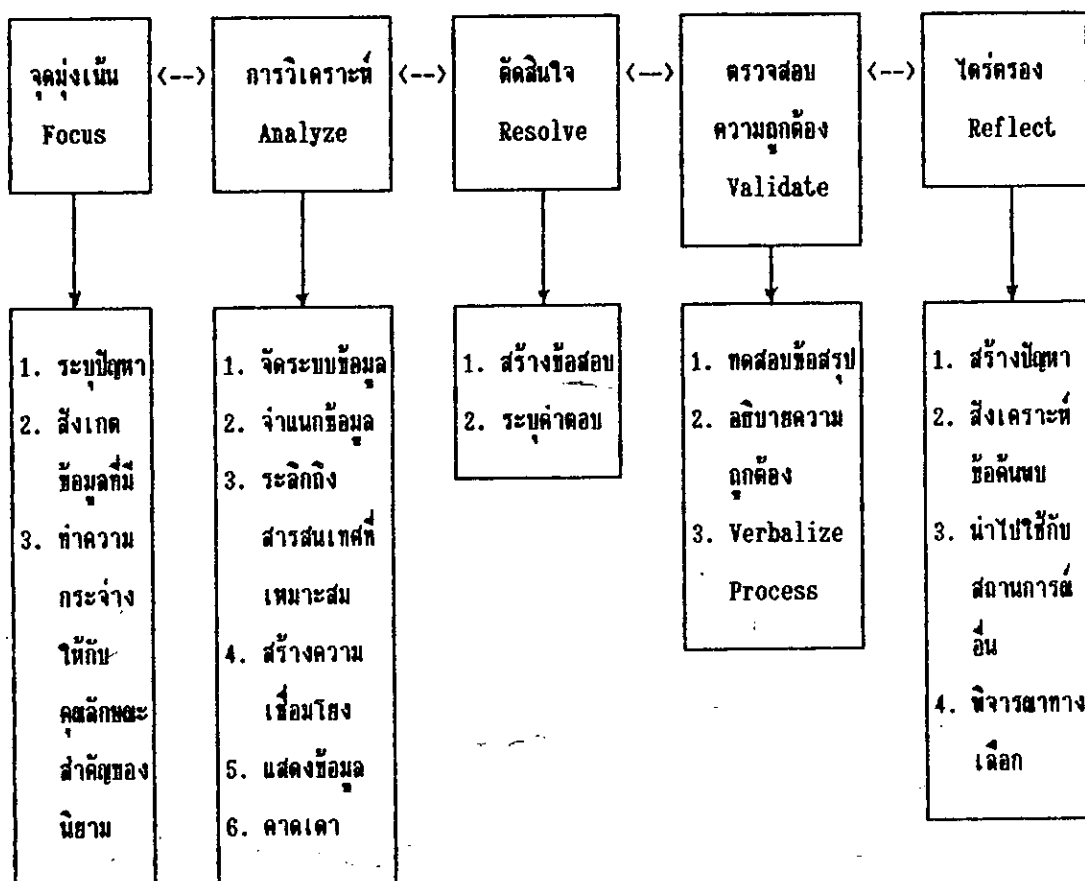
2. การคิดอย่างมีทิศทางหรือจุดมุ่งหมาย คือ การคิดที่บุคคลได้เริ่มใช้ความรู้พื้น-
ฐาน เพื่อทำการกลั่นกรองการคิดที่เพือผ่นเพื่อนำไปสู่จุดหมายบทสรุป ได้แก่ การคิดริเริ่ม
สร้างสรรค์ ซึ่งเป็นการคิดในลักษณะที่คิดได้หลายทางหรือคิดในลักษณะเชื่อมโยง และการ
คิดอย่างมีวิจารณญาณ ซึ่งเป็นการคิดอย่างมีเหตุผลเพื่อนำไปสู่ทางแก้ปัญหา

การคิดอย่างมีทิศทางและมีจุดมุ่งหมายนี้ คลอสมีเซอร์ และริบเบิล (Klausmeir
and Ripple. 1971) ให้ความเห็นว่า เป็นรากฐานสำคัญของการเรียนรู้และการแก้ปัญหา
ของบุคคล ทั้งนี้เพราะในการเรียนรู้และการแก้ปัญหานั้น บุคคลจะต้องรู้จักใช้การคิดอย่างมี
วิจารณญาณ ทำความเข้าใจข้อความจริงหรือปัญหาในลักษณะต่าง ๆ และใช้การคิดเชิงสร้าง
สรรค์ในการค้นหาแนวทางใหม่ ๆ ในการเรียนรู้และการแก้ปัญหา

ความเห็นเกี่ยวกับความสามารถที่จำเป็นสำหรับการคิดอย่างมีวิจารณญาณและการคิดเชิงสร้างสรรค์ ดังกล่าวข้างต้น ชี้ให้เห็นว่าเหตุผลพื้นฐานของทุกอย่างหรือทุกสิ่งที่คุณกระทำ เป็นผลสรุปจากกระบวนการของเหตุผล

ผลจากการศึกษาวิจัยแสดงให้เห็นว่า เมื่อมีการสอนทักษะการคิดโดยตรง ผลสัมฤทธิ์ทางการเรียนจะสูงขึ้น การเน้นการแก้ปัญหา และการใช้เหตุผลในห้องเรียน สามารถทำให้ช่องว่างระหว่างโลกแห่งความเป็นจริงและโลกในห้องเรียนลดลง ซึ่งจะช่วยให้ผู้เรียนเกิดความรู้สึกที่ดีในห้องเรียน การแก้ปัญหาเป็นกิจกรรมที่เกิดขึ้นตลอดชีวิต (Lifetime Activity) เด็กจะเข้าไปเกี่ยวข้องกับปัญหานั้นตั้งแต่เกิด ประสบการณ์การแก้ปัญหาจะมีอยู่ตลอดเวลา ดังนั้น การสอนการแก้ปัญหาจะต้องกระทำอย่างต่อเนื่อง การแสดงความคิดเห็นเกี่ยวกับปัญหา การเสนอแนวทางการแก้ไขและวิธีการเผชิญปัญหา จะต้องได้รับการพิจารณาตลอดเวลา การแก้ปัญหามีความน่าตื่นเต้นท้าทายและน่าสนใจ และยังเปิดโอกาสให้ผู้เรียนได้ฝึกการใช้ประสบการณ์ในการหาคำตอบด้วยตนเอง (Heuristic Thinking) การพิจารณาเลือกปัญหาเป็นสิ่งสำคัญที่จะนำไปสู่การพัฒนาทักษะ ในการแก้ปัญหาและการใช้เหตุผลที่จำเป็นในชีวิตประจำวัน (Krulik and Rudnick. 1993 : 8)

ทั้งนี้ ครูลิด ได้นำเสนอแผนภาพแสดงความต่อเนื่องของการคิดที่ทุกคนจะต้องใช้เมื่อเผชิญกับสถานการณ์ที่ต้องการให้หาทางแก้ไข ดังนี้



ภาพประกอบ 2 สุกฤษศาสตร์ของการให้เหตุผล (The Heuristics of Reasoning)

(Krutlik and Rudnick. 1993 : 28)

กระบวนการจะดำเนินไปมา ผู้มีประสบการณ์มากกว่าจะมีการเคลื่อนกลับไปมาอย่างเป็นอัตโนมัติ ผู้สอนจะต้องพัฒนากระบวนการค้นหาคำตอบด้วยตนเองให้แก่ผู้เรียน (Heuristic Process) ซึ่งโครงสร้างสถานการณ์เหล่านี้ ผู้เรียนจะต้องใช้ทักษะย่อย (Subskills) ต่าง ๆ

อนึ่ง เบเยอร์ (Beyer. 1988 : 26 - 30) ได้เสนอหลักการพื้นฐานที่จะใช้สอนทักษะการคิด โดยเริ่มจากการกำหนดปฏิบัติการการคิด (Thinking Operation) ที่เป็นองค์ประกอบของหลักสูตร ทั้งนี้เบเยอร์ได้เลือกทักษะและปฏิบัติการโดยใช้เกณฑ์ 3 ประการ คือ

1. เป็นทักษะที่จะใช้วิชาที่เรียนเป็นส่วนใหญ่
2. ใช้ทั่วไปนอกโรงเรียน
3. ผู้เชี่ยวชาญระบุว่า ทักษะเหล่านี้เป็นปฏิบัติการสำคัญของการคิด
ปฏิบัติการการคิดซึ่งมีความซับซ้อน 3 ระดับ ที่เบเกอร์ได้เสนอ คือ

I. <u>ยุทธวิธีทางความคิด (Thinking Strategies) จะรวมปฏิบัติการที่ซับซ้อน</u>		
การแก้ปัญหา (Problem Solving)	การตัดสินใจ (Decision Making)	การสร้างความคิดรวบยอด (Conceptualization)
1. ระบุปัญหา	1. นิยาม เป้าหมาย	1. ระบุตัวอย่าง
2. อภิบาล ทำความกระจ่างกับปัญหา	2. ระบบทางเลือก	2. ระบุลักษณะสำคัญ
3. สร้าง/เลือกแผนการแก้ปัญหา	3. วิเคราะห์ทางเลือก	3. จำแนกลักษณะสำคัญ
4. ดำเนินการตามแผน	4. จัดอันดับทางเลือก	4. สร้างความสัมพันธ์
5. ประเมินทางแก้ปัญหา	5. พิจารณาจัดตำแหน่ง ทางเลือกลำดับสูงสุด	ระหว่างคุณลักษณะ
	6. เลือกทางเลือกที่ดีที่สุด	5. ระบุตัวอย่างและไม่ใช่ ตัวอย่างเพิ่มเติม
		6. ปรับความคิดรวบยอด และโครงสร้างของ คุณลักษณะ
II. ทักษะการคิดอย่างมีวิจารณญาณ (Critical Thinking Skills) ไม่ใช่กระบวนการในความหมายเกี่ยวกับยุทธวิธี แต่เป็นปฏิบัติการทางสมองที่ใช้ในการกำหนดคุณค่าหรือความถูกต้องของบางสิ่ง คล้ายกับสิ่งที่อยู่ภายในที่นำมาใช้หรือการปฏิบัติ ได้แก่ 1. จำแนกความแตกต่างระหว่างความจริงที่สามารถตรวจสอบได้ และคุณค่าการกล่าวอ้าง 2. จำแนกความเกี่ยวเนื่องจากข้อมูลที่ไม่มีสอดคล้อง ข้ออ้างหรือเหตุผล 3. กำหนดความแม่นยำของข้อกล่าวอ้าง 4. กำหนดความน่าเชื่อถือของแหล่งข้อมูล 5. ระบุความกำกวมของข้ออ้างหรือการอ้างเหตุผล 6. ระบุข้อดกลองเบี่ยงเบนที่ไม่ได้กำหนดไว้ 7. ตรวจสอบความลำเอียง 8. ระบุการอ้างเหตุผลที่ผิดหลัก 9. Recognizing Logical Inconsistencies in Line of Reasoning 10. กำหนดความแกร่งของการอ้างเหตุผลหรือข้ออ้าง		III. ทักษะการประมวลข้อมูล (Information Processing Skills) เป็นปฏิบัติการขั้นพื้นฐานที่สุดของการคิด แต่ละทักษะจะค่อนข้างง่ายไม่ซับซ้อนในรูปของวิธีการ ทักษะเหล่านี้จะถูกใช้ซ้ำในรูปแบบที่ต่างกัน เพื่อความสำเร็จของยุทธศาสตร์ที่ซับซ้อนกว่าศาสตร์ 1 ระดับ คือ 1. การระลึกได้ 2. การแปลความหมาย 3. การตีความ 4. การขยายความ 5. การนำไปใช้ 6. การวิเคราะห์ (เปรียบเทียบ ความตรงกันข้าม จำแนก ลำดับ) 7. การสังเคราะห์ 8. การประเมิน 9. การใช้เหตุผล (การอ้างเหตุผล) 9.1 อุปมาน (Inductive) 9.2 อนุมาน (Deductive) 9.3 การอุปมา (Analogies)

ภาพประกอบ 3 ทักษะและยุทธศาสตร์การคิดที่สำคัญ

ทักษะการคิดและยุทธศาสตร์การคิดที่เลือกสำหรับการสอน จำเป็นต้องจัดโดยคำนึงถึงระดับชั้นเรียนและเนื้อหาวิชา และทักษะเหล่านั้นจะต้องถูกนำไปใช้ไปในเนื้อหาที่กำหนดให้ จากนั้นจะถูกถ่ายทอดไปยังวิชาอื่น ณ ระดับชั้นถัดไป และถือว่าคณิตศาสตร์และวิทยาศาสตร์เป็นวิชาที่สามารถนำยุทธศาสตร์การคิด และแผนการแก้ปัญหาที่กว้างขวาง ทั้งในรูปแบบและลำดับที่เบเซอร์ได้นำเสนอนั้น เขาได้นำทักษะการประมวลข้อมูล (การจำแนก จัดลำดับ) เข้าสู่หลักสูตรการสอนในระดับอนุบาลถึงเกรด 1 ยุทธวิธีการแก้ปัญหานั้นเริ่มในเกรด 3 ในวิชาคณิตศาสตร์และวิทยาศาสตร์ และปฏิบัติการการคิดอย่างมีวิจารณญาณเริ่มนำเข้าไปในเกรด 3 และ 4 โดยเลือกทักษะที่สามารถนำไปผ่านวิทยาศาสตร์ ภาษา ศิลปะ และสังคมศึกษา ทักษะเหล่านั้นสามารถนำไปใช้กับวิชาในปีต่อไป ซึ่งจะทำให้เกิดทักษะการคิดอย่างมีวิจารณญาณ การฝึกการสร้างรายละเอียด และถ่ายทอดไปยังวิชาอื่น ทักษะการวิเคราะห์เหตุผลอาจถูกสอนในลักษณะเป็น "Umbrella Operation" เพื่อช่วยให้ผู้เรียนเข้าใจทักษะการคิดอย่างมีวิจารณญาณเฉพาะ และการวิเคราะห์ สังเคราะห์ และการประเมิน เริ่มในเกรด 3, 4 หรือ 5 สามารถนำมาใช้กับวิชาที่เหมาะสม ศิลปะการใช้ภาษาเป็นสื่อหน้าที่ดีสำหรับการฝึกทักษะเหล่านี้ เช่น การวิเคราะห์ข้อความ เรื่องราว หรือละคร การสร้างหรือแต่งเรื่องความ ร่ายงาน และการประเมินคุณภาพของผลผลิตด้านการเขียนต่าง ๆ และสุดท้ายก็คือ ยุทธวิธีการตัดสินใจ สังคมศึกษา ศิลปะการใช้ภาษา และพลานามัย จะให้บริบทที่เป็นประโยชน์ต่อการสร้างความสามารถในการตัดสินใจ และทักษะนี้ควรเริ่มในระดับมัธยมศึกษา

เบเซอร์สรุปว่า การรอบรู้ในปฏิบัติการการคิดที่ซับซ้อนใด ๆ ต้องการกระทำด้วยความเอาใจใส่และต้องใช้เวลา

สมิท (Smith, 1988 : 36 - 39) ได้แสดงความเห็นเกี่ยวกับการสอนทักษะการคิดว่า ครูจำเป็นต้องเข้าใจธรรมชาติของนิยามของการคิด เพื่อความสามารถในการอธิบายสร้างความเข้าใจให้แก่ผู้เรียน ครูที่มีประสิทธิภาพมากจะออกแบบการสอนเอง ทั้งนี้ เราสามารถสอนครูในทักษะเกี่ยวกับการสอน การทำงาน การดำเนินงานตามเนื้อหา และจากนั้นปล่อยให้ครูใช้ทักษะเหล่านั้นด้วยตนเอง สมิทไม่เห็นด้วยกับการสอนตามสูตร กฎ มีรูปแบบ เพราะเห็นว่าไม่ควรมีรูปแบบที่คงที่ของพฤติกรรมการสอนในโรงเรียนที่ดีที่สุด หรือแม้แต่ในพฤติกรรมการสอนของครูที่ประสบความสำเร็จมากที่สุด

การคิดเป็นชุดของทักษะซึ่งอาจมีหลายชุด เป็นปฏิบัติการทางความคิดที่แสดงออกทางภาษาในลักษณะของการจัดการกับข้อมูล ข่าวสาร นักคิดที่มีทักษะจะเข้าไปเกี่ยวข้องกับจัด

การกับข้อมูล ข่าวสาร ในวิถีทางที่ผู้ไม่มีทักษะไม่สามารถกระทำได้ และการคิดสามารถปรับปรุงให้ดีขึ้นได้ด้วยการฝึก ทักษะการคิดเป็นวิธีการเก็บรวบรวมและจัดระบบข้อมูล เพื่อที่จะจำแนก อธิบาย พยากรณ์ ทดสอบ และอื่น ๆ เมื่อบุคคลกระทำทักษะเหล่านี้ บุคคลจะมีความยืดหยุ่นและมีประสิทธิภาพในฐานะเป็นนักคิด

ความยืดหยุ่นในความหมายของสมิท คือ สร้างทางเลือกที่มากกว่าวิธีเดียวนั่นเอง ในชีวิตจริงบุคคลอาจพบสถานการณ์ที่ทำให้เกิดความไม่แน่ใจหรืองง ปัญหาเหล่านี้ต้องการทักษะที่แตกต่างไปจากที่จำเป็นต้องใช้ในการแก้ปัญหาที่มีโครงสร้างชัดเจน มีเนื้อหาเฉพาะ ผู้เรียนจำเป็นต้องมีทักษะพื้นฐานของการคิด เพื่อนำไปสู่การพัฒนาระดับความคิดที่สูงขึ้นตามที่คาดหวัง ชนิดของการปฏิบัติของครูผู้สอนมีความสำคัญมากกว่าสิ่งอื่นใด ในกระบวนการสอน ผู้เรียน ครูผู้สอน มิได้ปฏิบัติงานโดยตรงกับกระบวนการคิด (Mental Processes) แต่ทำงานกับความรู้และสิ่งของ คือ ครูจะทำงานกับเนื้อหาของการสอนที่ส่งผลต่อวิถีทางที่ผู้เรียนจะเรียนรู้ที่จะคิดอย่างมีประสิทธิภาพ

สำหรับในประเทศไทย เชิดศักดิ์ โฆวสินธุ์ (2530) ได้ทำการวิจัยเพื่อพัฒนาคุณภาพการคิดของนักเรียนชั้นประถมศึกษาปีที่ 4 ด้วยวิธีการฝึกสมรรถภาพพื้นฐาน 4 ด้าน คือ การสังเกต การประจักษ์ การวิเคราะห์ และการสังเคราะห์ ซึ่งถือเป็นองค์ประกอบการคิดสำคัญที่นักเรียนจะนำมาแก้ปัญหาได้อย่างมีประสิทธิภาพ ไม่ว่าปัญหานั้นจะเป็นปัญหาเชิงวิชาการหรือปัญหาในชีวิตประจำวัน ทั้งนี้ถือเป็นการประเมินคุณภาพการคิดจาก สารการคิด (Productive Thinking) ของบุคคลว่ามีคุณภาพระดับใด จากการให้บุคคลทำงานที่คาดว่าจะต้องใช้วิธีการคิดระดับนั้นมาแก้ปัญหาหรือทำงานนั้นได้สำเร็จ โดยแบบฝึกสมรรถภาพสมองที่สร้างขึ้น ได้ดัดแปลงและปรับปรุงมาจากแบบทดสอบวัดเข้าวัดปัญหาหรือความถนัดที่ใช้รูปภาพ หรือ สัญลักษณ์ เป็นเนื้อหาเพื่อใช้ฝึกสมรรถภาพสมองตามองค์ประกอบทั้ง 4 ด้าน ทั้งนี้กลุ่มทดลองจะได้รับการฝึกสมรรถภาพสมองทุกวันที่มีการเรียนการสอน นักเรียนต้องแก้ปัญหาในแบบฝึกด้วยตนเอง เครื่องมือวัดคุณภาพการคิด คือ แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย และคณิตศาสตร์ ที่สร้างขึ้นตามเนื้อหาของหลักสูตร วัดพฤติกรรมตามทัศนคติของบลูมและคณะ แบบทดสอบมีลักษณะเป็นแบบเลือกตอบ โดยให้เลือกตอบคำตอบที่ถูกต้องที่สุด ดีที่สุด หรือ เหมาะสมที่สุด เพียงคำตอบเดียว

ผลการศึกษา พบว่า การฝึกสมรรถภาพสมองช่วยให้การเรียนรู้หรือคุณภาพการคิดของนักเรียนที่มีผลสัมฤทธิ์ต่ำและอยู่ในสภาพแวดล้อมที่ด้อยกว่า สามารถพัฒนาคุณภาพการคิดให้มีระดับเดียวกับผู้มีสภาพแวดล้อมที่ดีกว่าได้อย่างมีประสิทธิภาพ

สมเจตน์ ไวยากรณ์ (2530) เป็นผู้หนึ่งที่ได้พัฒนารูปแบบการสอน เพื่อพัฒนาความสามารถด้านการใช้เหตุผล ทั้งนี้ได้พัฒนารูปแบบการสอนโดยยึดพฤติกรรมการคิดระดับสูง คือ การคิดแบบวิเคราะห์ สังเคราะห์ และการประเมินค่า ตามแนวคิดของบลูม (Bloom's) เป็นพื้นฐาน โดยเน้นกระบวนการสอน 4 ขั้นตอน คือ การวางแผน การสร้างแนวคิดรวบยอด การนำความรู้ไปประยุกต์ใช้ และประเมินผล และจัดกระทำกับนักเรียนชั้นมัธยมศึกษาปีที่ 6 ทั้งนี้เลือกนักเรียนที่มีความสามารถด้านคณิตศาสตร์ไม่ต่างกัน 2 กลุ่ม จัดเป็นกลุ่มทดลองและกลุ่มควบคุม เครื่องมือวัดความสามารถด้านการใช้เหตุผลประกอบด้วยแบบทดสอบวัดความถนัดการเรียนรู้ด้านการจำแนกประเภท อุปมา-อุปไมย อนุกรมสัมพันธ์ และการสรุปความ ของสำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ การวิเคราะห์ข้อมูลเพื่อตรวจสอบผลการทดลองตามรูปแบบการสอน ใช้วิธีการวิเคราะห์ความแปรปรวนพหุนามแบบการวัดซ้ำ

ผลการทดลอง พบว่า รูปแบบการสอนที่สร้างขึ้นสามารถสอนให้ผู้เรียนเกิดพฤติกรรม การคิดตามที่เสนอของบลูมและคณะได้ทุกระดับพฤติกรรม แต่พฤติกรรมการคิดดังกล่าว ต้องการ เวลาสอนแตกต่างกัน โดยเฉพาะพฤติกรรมการคิดด้านการสังเคราะห์และการประเมินค่า ต้องการเวลาในการสอนมากกว่าพฤติกรรมการคิดด้านการวิเคราะห์ เมื่อเปรียบเทียบกับ การสอนตามปกติแล้ว พบว่า รูปแบบการสอนที่สร้างขึ้นช่วยให้ผู้เรียนทุกระดับการเรียนรู้ทั้งที่มีผล การเรียนดี ปานกลาง และผลการเรียนต่ำ มีความสามารถด้านการใช้เหตุผลในทุก ๆ ด้าน สูงขึ้นอย่างมีนัยสำคัญ ส่วนการสอนปกติช่วยส่งเสริมความสามารถด้านการใช้เหตุผลของผู้-เรียนเฉพาะผู้ที่มีผลการเรียนระดับปานกลางเท่านั้น

โดยสรุปแล้ว จะเห็นว่าความสามารถในการคิดสามารถพัฒนาได้ โดยการสอนการ ฝึกในลักษณะรูปแบบหรือยุทธวิธีที่เหมาะสม ความชัดเจนในโครงสร้างหรือองค์ประกอบของ ความสามารถในการคิด จะนำไปสู่กระบวนการพัฒนาความสามารถดังกล่าวได้อย่างมี ประสิทธิภาพ

1.4 ความสามารถด้านเหตุผล

การให้เหตุผลหรือการรู้จักเหตุผล (Reasoning) เน้นความสามารถของมนุษย์ใน การสร้างข้อสรุปหรืออ้างเหตุผล โดยยึดความรู้หรือความจริงที่คาดล่วงหน้าเป็นหลักในเชิงที่ สอดคล้องกับกฎหรือหลักการของเหตุผล

ความหมายตามนิยามที่กำหนดไว้ในสารานุกรมทางการศึกษา (Encyclopedia of Education) ปี ค.ศ. 1994 ถือเป็นกิจกรรมทางความคิดทั้งหมด ซึ่งการอ้างสรุป (Inference) จะมีบทบาทสำคัญ ทั้งนี้เพราะเป็นกระบวนการที่บุคคลลงความเห็นหรือวินิจฉัย เพื่อสร้างข้อสรุปจากข้อมูลที่มีอยู่ (Bullock. 1994 : 4951) อริสโตเติลได้จำแนกการให้เหตุผลเป็น 2 ลักษณะ คือ การให้เหตุผลที่เป็นไปตามแบบแผน (Formal Reasoning) เป็นการให้เหตุผลที่ใช้โครงสร้างทางตรรกวิทยาเป็นหลักในการอ้างอิง เช่น การจัดประเภท การอ้างเหตุผลของตรรกวิทยาที่เป็นเงื่อนไข (If-Then) และการให้เหตุผลที่ไม่มีแบบแผน (Informal Reasoning) การให้เหตุผลลักษณะนี้จะใช้โครงสร้างของการพูดที่อาจมีเนื้อหา มีการพาดพิงกับความคิดเห็น เนื่องจากเกี่ยวข้องกับสิ่งที่อาจเป็นความจริง โครงสร้างหลักของการให้เหตุผลที่ไม่มีแบบแผนนี้ เป็นการให้เหตุผลเพื่อสนับสนุนข้อสรุปในแวดวงที่ใช้การพูด การเขียน เช่น นักเขียนจะกล่าวว่าโครงสร้างทางตรรกวิทยาที่เป็นระเบียบวิธีการ (Formal Logical) จะมีใช้น้อยมากในโลกแห่งความเป็นจริง และจุดมุ่งเน้นจะเปลี่ยนจากวิธีการเขียนเป็นการวิเคราะห์การอ้างเหตุผลที่ถูกต้อง (Voss. 1994 : 4948)

แม้การนิยามเกี่ยวกับการให้เหตุผลจะค่อนข้างกว้าง อย่างไรก็ตามสามารถจำแนกออกเป็น 2 แนวทาง คือ ความสามารถในการสรุปความเกี่ยวกับวัตถุ เหตุการณ์ หรือกระบวนการในขอบเขตของเนื้อหาต่าง ๆ กัน เช่น เหตุผลทางสังคม ทางจิตศาสตร์ ทางชีววิทยา ทางจักรกล แนวทางหนึ่ง และการให้เหตุผลถูกนิยามในลักษณะที่เป็นความสามารถในการอ้างสรุปจากลักษณะเฉพาะของระบบการลงความเห็น เช่น การให้เหตุผลเชิงตรรก (Logical Reasoning) และการให้เหตุผลเชิงเปรียบเทียบ (Analogical Reasoning)

การให้เหตุผลเชิงตรรก (Logical Reasoning) หรืออาจเรียกอีกชื่อว่า การให้เหตุผลแบบอนุมาน (Deductive Reasoning) การให้เหตุผลลักษณะนี้มีความพิเศษเนื่องจากกระบวนการลงความเห็นหรือวินิจฉัยที่ถูกต้องจะเป็นไปตามกฎของการอนุมาน ซึ่งจำเป็นต้องเป็นจริง ยิ่งไปกว่านั้นการตัดสินใจความต้องการของการให้เหตุผลนี้จะเกี่ยวข้องกับรูปแบบของข้อความมิใช่เนื้อหา การให้เหตุผลเชิงตรรกนี้เป็นพื้นฐานของกิจกรรม เช่น การทดสอบสมมุติฐาน การพิสูจน์ความขัดแย้งกัน หรือการประเมินความต้องการตามตรรกวิทยาของคำอธิบาย

ประเด็นที่น่าสนใจสำหรับความสามารถด้านเหตุผลลักษณะนี้ ก็คือ

1. เด็กสามารถสรุปความอย่างถูกต้องตามเหตุผลได้หรือไม่ และถ้าได้เมื่อไหร่
2. เงื่อนไขใดที่สนับสนุนหรือเป็นอุปสรรคต่อการสรุปความที่ถูกต้อง

3. เด็กเข้าใจการสรุปความแบบอนุมานถูกต้องเป็นจริงหรือไม่ และถ้าเข้าใจเมื่อไหร่

จากหลักฐานการศึกษาในต้นปี ค.ศ. 1990 เกี่ยวกับความสามารถในการสรุปความในระดับอายุต่าง ๆ กัน ให้ภาพที่ไม่ชัดเจนนัก บางสถานการณ์เด็กก่อนวัยเรียนสามารถลงความเห็นที่ถูกต้องตามเหตุผลได้ ขณะที่ในหลาย ๆ สถานการณ์ผู้ใหญ่ลงความเห็นผิดพลาด ข้อค้นพบเหล่านี้เป็นสื่อชี้ให้เห็นว่า ในทางตรงข้ามกับพัฒนาการความสำเร็จหลายประการ การให้เหตุผลเชิงอนุมานมีรูปแบบที่ที่มีความซับซ้อน และไม่สามารถกระทำได้ง่ายนัก อย่างไรก็ตาม ผลการรวบรวมงานวิจัยสรุปได้ว่า เด็กก่อนวัยเรียนจะมีความสามารถเบื้องต้นในการลงความเห็นแบบอนุมานเกี่ยวกับวัตถุที่มีรูปร่างแสดงชัดเจน เช่น ในรูปของ "บ้านคดออก ฉันจะไปเดินเล่น" "ตอนนั้นคดออก ฉันจะไปเดินเล่น" สำหรับผู้ใหญ่สามารถลงความเห็น (สรุปความ) เกี่ยวกับสิ่งที่ป็นนามธรรมได้ (Bullock. 1994 : 4953; citing Braine and Romain. 1983)

อินเฮลเดอร์ และเพียเจต์ (Inhelder and Piaget. 1958) กล่าวว่า การให้เหตุผลเชิงตรรกะก่อนโครงสร้างของความคิด นั้นคือ การจัดระบบความคิดสามารถจัดเป็นเครือข่ายตามเหตุผล ด้วยเครือข่ายนี้เองที่จะนำไปสู่ความสามารถด้านเหตุผลที่ไม่ขึ้นกับเนื้อหา โครงสร้างที่เป็นนามธรรมนี้จะได้อายุอย่างสมบูรณ์ในวัยหนุ่มสาว (Adolescence) ภายหลังจากการสังเกตมาเป็นระยะเวลานาน แต่แนวความคิดของเพียเจต์ถูกวิพากษ์วิจารณ์ในเรื่องที่ว่า การให้เหตุผลเชิงตรรกะมีมากเกินไปที่รูปแบบความสามารถของเขาจะทำได้ และตรรกวิทยาบางอย่างที่เขาใช้อธิบายโครงสร้างทางความคิดยังมีความไม่สมบูรณ์ อย่างไรก็ตาม ถือว่าแนวความคิดของเพียเจต์ถือเป็นรากฐานสำคัญต่อรูปแบบความคิดในปัจจุบัน

มอสแมน (Bullock. 1994 : 4955; citing Moshman. 1990) ได้เสนอระดับขั้นพัฒนาการความสามารถด้านเหตุผลเชิงตรรก 3 ระดับ ดังนี้

ระดับแรก อายุ 6 - 7 ขวบ เด็กยังไม่มีความคิดรวบยอดเกี่ยวกับความสำคัญ ความถูกต้องตามเหตุผล

ระดับที่ 2 อยู่ในช่วงของวัยเด็กตอนกลาง เด็กไม่สามารถให้เหตุผลเกี่ยวกับความถูกต้องของการอ้างเหตุผลโดยอิสระจากความเป็นจริงได้ เช่น ในข้อความว่า "ช้างเป็นพืชหรือสัตว์ อย่างใดอย่างหนึ่ง" "ช้างไม่ใช่สัตว์" เด็กวัยนี้จะไม่ยอมรับข้อสรุปที่ว่า "ช้างเป็น

นี้" ว่าถูกต้อง ทั้งนี้เพราะเด็กยังไม่เข้าใจการสรุปความเชิงตรรกที่ถูกใช้ในลักษณะที่มีความสัมพันธ์อย่างอิสระจากเนื้อหาตนเอง

ระดับที่ 3 ความคิดรวบยอดเกี่ยวกับความถูกต้องของการลงสรุป จะเกิดเมื่อเข้าสู่วัยหนุ่มสาวและเกิดกับบางคนเท่านั้น โดยความรู้ลึกทั่วไป ผู้มีเหตุผลดี (Good Reasoner) จะเป็นผู้ที่ถูกมองว่าเป็นผู้คิดตามหลักตรรกวิทยา อย่างไรก็ตาม ในการให้เหตุผลอย่างไม่เป็นแบบแผนซึ่งมีความสัมพันธ์กับเรื่องราวทางสังคม การอ้างเหตุผลอาจถูกประเมินต่างกันไปในแต่ละบุคคลตามความเชื่อและเจตคติของเขา ซึ่งชี้ให้เห็นว่าการให้เหตุผลดังกล่าวช่วยให้บุคคลตัดสินใจได้ดีที่สุด หรือสร้างข้อสรุป ข้อสนับสนุน อันสมควรที่สุด ความพยายามหรือเป้าหมายของการเรียนในโรงเรียนที่เหมาะสมและต้องการไปสู่ ก็ควรจะเป็นการพัฒนาเด็กให้มีความสามารถในการให้เหตุผล เพื่อสร้างข้อสรุปที่สมเหตุสมผลมากกว่าการหาคำตอบที่ถูกต้องที่สุด

การให้เหตุผลเชิงเปรียบเทียบ (Analogical Reasoning)

สำหรับการให้เหตุผลเชิงเปรียบเทียบนั้น ถือว่าเป็นส่วนประกอบสำคัญของการคิดของมนุษย์ เพราะเป็นหนทางที่นำไปสู่การค้นพบทางวิทยาศาสตร์และความคิดสร้างสรรค์ ทั้งการเปรียบเทียบมีความสำคัญในการเรียนรู้ การแก้ปัญหา และการค้นพบ นักวิจัยที่ให้ความสนใจศึกษาพัฒนาการความสามารถด้านนี้ ได้แก่ เพียเจต์ ตามทฤษฎีของเพียเจต์ (Piagetion Theory) นั้น เชื่อว่าความสามารถในการให้เหตุผลเชิงเปรียบเทียบเป็นทักษะที่พัฒนาช้า จะปรากฏในราวอายุ 11 - 12 ปี เด็กที่อายุน้อยกว่านี้จะไม่สามารถให้เหตุผลเชิงเปรียบเทียบได้ จากการวิจัยในทางจิตวิทยาการวัด พบว่า เด็กหลายคนแม้เมื่ออายุ 11- 12 ปี แล้ว ก็ไม่สามารถเติมอุปมา - อุปไมยง่าย ๆ เช่น หมู : หมูป่า :: สุนัข : ? (สุนัขป่า) ได้ แต่เด็กจะตอบเป็น หมู : หมูป่า :: สุนัข : แมว ซึ่งผิด รูปแบบการวัดพัฒนาการด้านทักษะเหตุผลของเด็กลักษณะนี้ ถือเป็นวิธีการดั้งเดิมที่ได้แนวคิดมาจาก อริสโตเติลที่นิยามการเปรียบเทียบ (Analogy) ในลักษณะของสัดส่วนที่เท่ากัน ซึ่งประกอบด้วยพจน์อย่างน้อย 4 พจน์ เมื่อ 2 สัมพันธ์กับ 1 เท่ากับ 4 สัมพันธ์กับ 3 (Aristotle, Metaphysics) ประสบการณ์ในโรงเรียนจะทำให้ทักษะการให้เหตุผลอย่างไม่มีแบบแผนเพิ่มขึ้นเล็กน้อย ข้อค้นพบนี้เป็นตัวกระตุ้นสำคัญต่อการพัฒนาการปฏิบัติการสอนที่จะทำให้การได้มาซึ่งทักษะการให้เหตุผลง่ายยิ่งขึ้น (Goswami, 1992 : 3)

กล่าวโดยสรุปว่าอายุ และการให้เหตุผลมีความสัมพันธ์กัน ความรู้ที่เพิ่มขึ้นและการเพิ่มการใช้ยุทธวิธีการให้เหตุผล จะสามารถทำให้การให้เหตุผลคล่องหรือง่ายขึ้น

นิคเคอร์สัน (Nickerson. 1984 : 26 - 36) ได้สรุปรูปแบบการสอน เพื่อส่งเสริมทักษะการคิดอย่างมีเหตุผลตามที่ปรากฏในปัจจุบันออกเป็น 5 กลุ่ม ทั้งนี้แต่ละกลุ่มมีแนวทางในการสอนแตกต่างกันออกไป โดยทักษะการคิดอย่างมีเหตุผลที่เป็นเป้าหมายของแต่ละรูปแบบการสอน อาจมีความแตกต่างกันบ้าง แต่ส่วนใหญ่จะเป็นทักษะพื้นฐานของความสามารด้านการใช้เหตุผล อันได้แก่ ความสามารถในการจัดประเภท การจัดเรียงลำดับ การเปรียบเทียบ การเทียบเคียง และการสรุปอ้างอิง ดังรายละเอียดของการจัดการสอนแต่ละกลุ่ม ดังต่อไปนี้

1. กลุ่มที่เน้นแนวทางกระบวนการคิด (Cognitive-Process Approaches) กลุ่มนี้กำหนดข้อตกลงไว้ว่า ความสามารถในการคิดนั้นเป็นสิ่งที่ขึ้นอยู่กับกระบวนการคิดพื้นฐานบางประการ เช่น การเปรียบเทียบ การจัดลำดับ การจำแนกประเภท และการทำนาย กระบวนการดังกล่าวนี้ เป็นกระบวนการคิดอย่างมีระบบเหตุผล ซึ่งนำไปประยุกต์ใช้ในชีวิตประจำวัน

2. กลุ่มที่เน้นแนวทางยุทธศาสตร์การคิด ซึ่งมุ่งเน้นเกี่ยวกับกลวิธีที่นำมาใช้ในการแก้ปัญหา ซึ่งเป็นแนวทางที่เชื่อว่ามีโอกาสจะประสบผลสำเร็จสูงในการนำไปสู่เป้าหมาย แนวทางนี้มักจะพบในงานวิจัยด้านจิตวิทยาที่เกี่ยวกับการคิด การแก้ปัญหา หรือในงานวิจัยเกี่ยวกับเชาว์ปัญญาประดิษฐ์ (Artificial Intelligence) ทั้งนี้งานวิจัยทั้ง 2 แนวทางดังกล่าว ต่างมุ่งทำความเข้าใจเกี่ยวกับวิธีการซึ่งผู้เชี่ยวชาญนำมาใช้ในการแก้ปัญหานั้นต่าง ๆ ว่ามีความแตกต่างไปจากวิธีการที่ผู้ยังขาดประสบการณ์ใช้หรือไม่ โดยมุ่งหวังว่าหากพบข้อแตกต่างก็จักนำวิธีการที่ผู้เชี่ยวชาญได้ใช้ มาเป็นแนวทางช่วยเหลือผู้ที่ยังขาดประสบการณ์ต่อไป และจากการศึกษาพบว่า ผู้ที่มีความเชี่ยวชาญมักใช้เวลาสำหรับกิจกรรมการใช้ความคิดรวบยอดเพื่อพิจารณาปัญหา การกำหนดสิ่งที่จะใช้เป็นตัวแทนปัญหาหลาย ๆ ทาง ตลอดจนการวางแผนเพื่อดำเนินการแก้ปัญหา ก่อนจะลงมือแก้ปัญหา มากกว่าพวกที่ยังขาดประสบการณ์ใช้ จากข้อค้นพบดังกล่าว กลุ่มนี้จึงมุ่งเน้นให้ผู้เรียนได้ฝึกทำกิจกรรมที่ค้นพบดังกล่าวก่อนลงมือแก้ปัญหา ซึ่งคือ การฝึกคิดอย่างวิเคราะห์วิจารณ์ นั่นเอง

3. กลุ่มที่เน้นแนวทางเกี่ยวกับพัฒนาการของการคิดตามทฤษฎีของเพียเจต์ (Formal Thinking or Stage Development) โปรแกรมของกลุ่มนี้สร้างขึ้นตามทฤษฎีเกี่ยวกับพัฒนาการของการคิดของเพียเจต์ โดยมุ่งให้ผู้เรียนสามารถพัฒนาการคิดของตนเอง

จากการคิดเฉพาะด้าน และลักษณะที่เป็นรูปธรรมให้สามารถคิดในแนวทาง และคิดในสิ่งที่ เป็นนามธรรมได้ ซึ่งเป็นพัฒนาการในระดับการใช้เหตุผลเชิงตรรกวิทยาได้ แนวคิดนี้ได้มี การนำไปศึกษากับนักศึกษามหาวิทยาลัย ในขณะที่เรียนเนื้อหาวิชาตามปกติ เช่น คาร์พลิส และผู้ร่วมงานได้ทำการจำแนกกระบวนการเรียนรู้ออกเป็น 3 ขั้นตอน คือ ขั้นการสำรวจ (Exploration) ขั้นการคิดค้น (Invention) และขั้นการนำไปประยุกต์ใช้ (Applica- tion) โดยในขั้นแรกนักศึกษาจะทำการรวบรวมข้อมูลที่ได้จากการสำรวจสิ่งต่าง ๆ รอบตัว โดยไม่กำหนดทิศทาง หรือความสัมพันธ์ของสิ่งต่าง ๆ ซึ่งเป็นรายละเอียดของเนื้อหาวิชา สำหรับขั้นคิดค้น นักศึกษาจะได้รับการกระตุ้นให้ทำการสรุปหลักการที่สอดคล้องกับสถานการณ์ ต่าง ๆ ที่ได้สำรวจมาในขั้นต้น และขั้นสุดท้ายซึ่งเป็นขั้นการนำไปประยุกต์ใช้นั้น นักศึกษา จะนำหลักการที่ได้ไปปรับใช้ในสถานการณ์อื่น ๆ เพื่อเป็นการขยายผลของความรู้ต่อไป

4. กลุ่มที่เน้นในแนวทางของการใช้ภาษาและสัญลักษณ์ (Language and Symbol Manipulation) แนวทางนี้มีความเชื่อว่า การเขียนที่มีประสิทธิภาพนั้น เป็นกิจกรรมที่มี แบบแผนที่ต้องใช้ความสามารถในการแสดงความคิดออกมาให้แจ่มชัด และมีความ ต่อเนื่อง ซึ่งลักษณะดังกล่าวนี้จำเป็นต้องมีการวางแผน ตลอดจนกำหนดแนวทางปฏิบัติเพื่อนำ ไปสู่เป้าหมาย โดยมีการแบ่งงานออกเป็น ส่วน ๆ หรือเป็นขั้นตอนที่ต่อเนื่อง ซึ่งเป็นการฝึก ทักษะการคิดอย่างมีเหตุผลด้านการวิเคราะห์และสังเคราะห์ข้อความ โดยให้เขียนเป็น วิธีการแสดงความคิดออกมาเป็นเครื่องมือในการพัฒนา

5. กลุ่มนี้ยึดการคิดเป็นเนื้อหาสาระของการฝึก หรือการเน้นแนวทางของการคิด เกี่ยวกับการคิด (Thinking About Thinking) กลุ่มนี้เชื่อว่าการเรียนรู้เกี่ยวกับการคิด จะช่วยให้ผู้เรียนสามารถพัฒนากระบวนการคิดของตนเองให้ดีขึ้น เพราะผู้เรียนจะรู้จักสิ่งที่ เป็นความคิดของตนเอง รู้ว่าตนกำลังคิดอะไร และต้องการรู้อะไร อันเป็นแนวทางที่ช่วย ให้ผู้เรียนสามารถควบคุมและตรวจสอบความคิดของตนเองได้ในขณะที่ทำการคิด กลุ่มที่ใช้ แนวทางนี้มีความเห็นว่างานวิจัยเกี่ยวกับการคิดที่ผ่านมา ยังขาดการเน้นให้ผู้เรียนได้เข้าใจ ถึงสิ่งที่ตนเป็นจุดเด่นและจุดด้อยของการคิดของตนเอง หรือขาดการค้นหาข้อผิดพลาดที่มักจะมี เกิดขึ้นในขณะที่ทำการคิดค้นนั้น ฉะนั้น กลุ่มนี้จึงมุ่งที่จะพัฒนาการคิดของผู้เรียนให้ถึงขีดสูงสุดตาม ศักยภาพที่ผู้เรียนมีอยู่ โดยให้ผู้เรียนได้ทำการวางแผนการคิดเป็นขั้นตอน เพื่อใช้เป็นกรอบ ในการตรวจสอบว่าตนเองมักมีข้อผิดพลาดในขั้นตอนใด เพื่อทำการแก้ไขสิ่งที่ตนเป็นข้อผิดพลาด ดังกล่าว ทักษะที่กลุ่มผู้วิจัยตามแนวคิดนี้ใช้ ได้แก่ กระบวนการสืบสวน กระบวนการคิดแก้ปัญหา การสรุปหลักการจากข้อมูลและเงื่อนไขที่มีอยู่ กระบวนการโยงเหตุและผลเข้าด้วยกัน

ทั้งนี้โดยอาศัยความอยากหรืออยากเห็นของผู้เรียนเป็นเครื่องกระตุ้นในการฝึก บางกลุ่มใช้ทักษะการคิดเกี่ยวกับการสรุปอ้างอิง การเปรียบเทียบ การสร้างข้อสันนิษฐาน และการจำแนกประเภท

แนวคิดเพื่อฝึกทักษะการคิดทั้ง 5 กลุ่มนี้ เป็นแนวทางที่กระทำโดยมุ่งเน้นให้ผู้เรียนเกิดทักษะการคิดอย่างมีเหตุผลทุกกลุ่ม เพียงแต่ใช้วิธีการและทักษะการคิดบางทักษะแตกต่างกันเท่านั้น

2. แนวคิดเกี่ยวกับการวัดและการประเมินผล

2.1 บทบาทของการวัดและการประเมินผล

การวัดและการประเมิน มีบทบาทที่สำคัญยิ่งในโปรแกรมการเรียนการสอนของโรงเรียนโดยพื้นฐานที่สุด คือ ให้ข้อมูลสารสนเทศที่สามารถนำไปใช้ในการตัดสินใจทางการศึกษาหลายประการ จุดสำคัญ คือ เน้นการเรียนการสอนในห้องเรียน การตัดสินใจเกี่ยวกับการเรียนรู้และพัฒนาผู้เรียน

จากจุดของการสอน อาจนิยามการประเมินได้ว่าเป็นกระบวนการที่เป็นระบบของการกำหนดว่าผู้เรียนสัมฤทธิ์ผลเพียงใดในจุดประสงค์ของการสอน ซึ่งก็คือ ผลที่ได้จากการเรียนรู้ กระบวนการประเมินจะรวมทั้งวิธีการวัด อันได้แก่ การทดสอบ และการสังเกตอย่างไม่เป็นทางการ สำหรับอธิบายการเปลี่ยนแปลงการปฏิบัติของผู้เรียนเท่า ๆ กับการพิจารณาถึงความเห็นเกี่ยวกับคุณค่า ความต้องการการเปลี่ยนแปลง วิธีการวัดและการประเมิน จะให้หลักฐานทุกอย่างที่เป็นระบบและจุดประสงค์ ซึ่งเป็นพื้นฐานสำหรับการพิจารณาตัดสินใจเกี่ยวกับการสอน กรอนลันด์ และลินน์ (Gronlund and Linn. 1990 : 3 - 15) กล่าวว่า การประเมินการเรียนรู้ของผู้เรียนต้องใช้วิธีการวัดผลสัมฤทธิ์ของนักเรียนมากมาย อย่างไรก็ตาม การประเมินไม่ใช่เพียงวิธีการเก็บรวบรวม แต่เป็นกระบวนการที่เป็นระบบที่มีบทบาทสำคัญต่อการสอนที่มีประสิทธิภาพ การประเมินเริ่มจากการระบุผลที่ได้ ซึ่งต้องการจากการเรียนรู้ และจบด้วยการพิจารณาตัดสินเกี่ยวกับขอบข่ายที่ได้ผลการเรียนรู้มา

ความแตกต่างระหว่างการทดสอบหรือแบบทดสอบ (Test) การวัด (Measurement) และการประเมิน (Evaluation) ซึ่งทั้ง 3 คำ มักทำให้เกิดความสับสน

การทดสอบ หมายถึง เครื่องมือหรือวิธีการมาตรฐานสำหรับวัดพฤติกรรมที่เป็นตัวอย่าง (ตอบคำถามว่า บุคคลกระทำไ้เพียงใด เมื่อเปรียบเทียบกับคนอื่นหรือเมื่อเปรียบเทียบกับประสิทธิภาพของงานที่ให้อำนาจ)

การวัด หมายถึง กระบวนการที่ได้มาซึ่งตัวเลขที่ใช้อธิบายระดับคุณลักษณะเฉพาะ (ตอบคำถาม How much ?)

การประเมิน หมายถึง กระบวนการที่เป็นระบบเกี่ยวกับการเก็บรวบรวมวิเคราะห์ และตีความข้อมูล เพื่อกำหนดการตัดสินใจว่านักเรียนสัมฤทธิ์ไปประสงค์ของการสอน (ตอบคำถาม How good?)

การประเมินมีความหมายที่ซับซ้อนกว่าการวัดและการทดสอบ ทั้งนี้เพราะการวัดจะถูกจำกัดที่คำอธิบายคุณภาพของผู้เรียน ที่ผลของการวัดแสดงในลักษณะของจำนวน ซึ่งไม่รวมคำอธิบายเชิงคุณภาพ (เช่น งานของสหุทยา สละอาด เรียบร้อย) หรือไม่ได้แสดงนัยของการลงความเห็นเกี่ยวกับคุณค่าของผลที่ได้มา

การประเมินมีลักษณะในทางตรงกันข้าม จะรวมทั้งคำอธิบายเชิงปริมาณ (การวัด) และคำอธิบายเชิงคุณภาพของผู้เรียน รวมทั้งมักจะรวมการลงความเห็นเกี่ยวกับคุณค่า (Value Judgement) เกี่ยวกับผลที่ต้องการ การประเมินจะถูกมองเป็นกระบวนการบูรณาการ (Integrated Process) สำหรับการกำหนดธรรมชาติและขยายขอบข่ายการเรียนรู้และพัฒนาการของผู้เรียน

หลักการทั่วไปของการประเมิน

1. ระบุสิ่งที่จะถูกวัดอย่างชัดเจน
2. การเลือกวิธีการประเมิน จะต้องพิจารณาความสอดคล้องกับคุณลักษณะ หรือการปฏิบัติที่ต้องการวัด วิธีการประเมินควรเลือกใช้เกณฑ์ ต่อไปนี้
 - 2.1 ความเป็นปรนัย (Objectivity)
 - 2.2 ความถูกต้องแม่นยำ (Accuracy)
 - 2.3 ความสะดวก (Convenience)
3. การประเมินสิ่งที่ซับซ้อนต้องการวิธีการประเมินที่หลากหลาย
4. การใช้วิธีการประเมินที่เหมาะสมต้องตระหนักถึงข้อจำกัดของมันด้วย

5. การประเมินเป็นวิธีที่จะนำไปสู่จุดหมายปลายทาง (End) แต่ไม่ใช่จุดหมายปลายทาง การประเมินถูกมองดีที่สุดในฐานะเป็นกระบวนการได้มา ซึ่งข้อมูลที่เป็นพื้นฐานในการตัดสินใจ

ชนิดของการทดสอบและวิธีการประเมิน

ลักษณะเด่นของกระบวนการประเมิน คือ การใช้วิธีการที่หลากหลาย ทั้งนี้อาจจำแนกหรืออธิบายในวิธีการที่แตกต่างกัน ขึ้นอยู่กับกรอบการอ้างอิงที่ใช้

การทดสอบและวิธีการประเมินอาจแบ่งเป็น 2 กลุ่ม ตามธรรมชาติของการวัดดังที่ครอนบาค (Cronbach) เรียกว่า การวัด Maximum Performance และ Typical Performances

กลุ่มแรก Maximum Performance จะเป็นตัวชี้ความสามารถ (Abilities) ของบุคคล วิธีการของการวัดชนิดนี้จะเกี่ยวกับการดูว่าบุคคลจะทำได้สูงสุดเท่าที่จะเป็นไปได้ อย่างไรก็ตาม ผลการประเมินจะชี้ให้เห็นว่า บุคคลสามารถทำอะไรได้เมื่อเขาถูกผลักดันไปข้างหน้า แบบทดสอบความถนัดและผลสัมฤทธิ์ทางการเรียนจะถูกรวมอยู่ในกลุ่มนี้ ข้อแตกต่างระหว่างแบบทดสอบ 2 ชนิด คือ การใช้ผลการสอบมากกว่าคุณภาพของแบบทดสอบ

แบบทดสอบความถนัด (Aptitude Test) ถูกออกแบบเพื่อพยากรณ์ความสำเร็จในกิจกรรมการเรียนรู้ในอนาคต ขณะที่แบบทดสอบผลสัมฤทธิ์ทางการเรียนถูกออกแบบเพื่อชี้บ่งระดับความสำเร็จในกิจกรรมการเรียนรู้ที่ผ่านมา และเนื่องจากแบบทดสอบบางฉบับถูกนำไปใช้เพื่อวัตถุประสงค์ทั้ง 2 ประการ เช่น แบบทดสอบพีชคณิตออกแบบเพื่อวัดผลสัมฤทธิ์เมื่อจบรายวิชา และอาจใช้พยากรณ์ความสำเร็จในรายวิชาคณิตศาสตร์ในอนาคต หน้าที่ซึ่งซ้อนทับกันอยู่นี้ เป็นตัวกันการแยกประเภท ฉะนั้น จึงเรียกแบบทดสอบความถนัดและผลสัมฤทธิ์ทางการเรียนว่า การวัดความสามารถ

กลุ่มที่ 2 ถูกออกแบบเพื่อสะท้อนพฤติกรรมปกติ (Typical Behavior) ของบุคคล กระบวนการของการวัดชนิดนี้ จะเกี่ยวกับสิ่งที่บุคคลจะปฏิบัติมากกว่าสามารถปฏิบัติ (Will Do มากกว่า Can Do) วิธีการถูกออกแบบเพื่อประเมินความสนใจ เจตคติ การปรับตัว และลักษณะทางบุคลิกลักษณะต่าง ๆ กลุ่มนี้เน้นการตอบสนองที่เป็นตัวแทนมากกว่าคะแนนที่สูง

อนึ่ง การทดสอบและวิธีการประเมินอาจถูกจำแนกโดยบทบาทหน้าที่ในกิจกรรมการเรียนรู้การสอน ดังนี้

การประเมินเพื่อจัดตำแหน่ง (Placement Evaluation) เพื่อกำหนดผลการปฏิบัติของผู้เรียนว่าผู้เรียนมีความรู้หรือไม่ มีทักษะที่จำเป็นก่อนเริ่มวางแผนการสอนใหม่

การประเมินผลย่อย (Formative Evaluation) เพื่อติดตามความก้าวหน้าในการเรียนระหว่างสอน เพื่อให้ข้อมูลย้อนกลับที่ต่อเนื่องกันทั้งครูและนักเรียนเกี่ยวกับความสำเร็จและความล้มเหลว

การประเมินเพื่อวินิจฉัย (Diagnostic Evaluation) เพื่อวินิจฉัยความยุ่งยากระหว่างการเรียนการสอน

การประเมินผลรวม (Summative Evaluation) เป็นการประเมินผลสัมฤทธิ์เมื่อเสร็จสิ้นการเรียนการสอน

ด้วยบทบาทหน้าที่ของการทดสอบและวิธีการประเมินที่แตกต่างกัน จำเป็นต้องพิจารณาออกแบบการทดสอบที่จะอธิบายผลการปฏิบัติของผู้เรียนได้อย่างพอเพียง ไม่ขัดขวางต่อการปรับหรือการวัดผู้เรียน ซึ่งครูได้ใช้เวลาเป็นเดือนในการสอน ตรวจสอบ สังเกต และมีปฏิสัมพันธ์ระหว่างกันของครูและนักเรียน นักฟิสิกส์ใช้การสังเกต สัมภาษณ์ การสัมผัสความรู้สึก ในการวัดเท่า ๆ กับการใช้แบบทดสอบ และพัฒนาการประเมินและวางแผนโดยการนำข้อมูลจากแหล่งต่าง ๆ มารวมกันมากกว่าเน้นไปที่แหล่งใดแหล่งหนึ่ง ด้วยวิธีนี้การประเมินไม่จำเป็นต้องขึ้นอยู่กับแบบทดสอบหรือการทดสอบแต่เพียงอย่างเดียว อย่างไรก็ตาม ชูลแมน (Shulman, 1980 : 39; citing Baker, 1980) ได้กล่าวสรุปว่า หลักการของการออกแบบแบบทดสอบ จะนำไปสู่การเลือกข้อคำถามสำหรับการทดสอบในลักษณะที่จะทั้งข้อคำถามหรือเขียนข้อคำถามใหม่ ถ้าปรากฏว่าข้อคำถามไม่มั่นคง ถ้าหากมีการขึ้นลงในประเด็นสำคัญ จากวัน--->วัน หรือสัปดาห์--->สัปดาห์ ก็จะต้องใช้ความพยายามที่จะลดการขึ้นลงไม่สม่ำเสมอของแบบทดสอบลง นอกจากนี้ เขายังกล่าวว่าการเรียนรู้เป็นกระบวนการภายในของผู้เรียนที่สังเกตไม่ได้ แต่เราจะอ้างอิงจากพฤติกรรมของผู้เรียน และถ้าการเรียนรู้เป็นกระบวนการอ้างอิงสรุปโดยให้เหตุผลแล้ว การสอนจึงเป็นความสามารถในการใช้สิ่งชี้แนะ (Cues) ชนิดต่าง ๆ เพื่อสร้างข้อสรุปเกี่ยวกับกระบวนการ เพื่อชี้นำการเรียนรู้

2.2 เครื่องมือวัด

เป้าหมายหลักของการทดสอบในชั้นเรียนก็เพื่อให้ได้ข้อมูลที่เที่ยงตรง เชื่อมั่น และเป็นประโยชน์เกี่ยวกับผลสัมฤทธิ์ของผู้เรียน ซึ่งหมายความว่า จะต้องมีการกำหนดว่า

อะไรเป็นสิ่งที่จะถูกวัด การนิยามสิ่งที่วัดให้ชัดเจนจะทำให้สามารถสร้างข้อคำถามได้ตรง ซึ่งจะก่อให้เกิดผลการปฏิบัติตามที่ต้องการ องค์ประกอบที่เกี่ยวข้องกับการวางแผนการสอบซึ่งจะนำไปสู่ข้อมูลที่เป็นประโยชน์และถูกต้อง ได้แก่

1. การกำหนดเป้าประสงค์ของการทดสอบ
2. การพัฒนาคุณลักษณะเฉพาะของแบบทดสอบ
3. การคัดเลือกชนิดของข้อคำถามที่เหมาะสม
4. การเตรียมข้อคำถาม

การคัดเลือกชนิดข้อคำถามที่เหมาะสม ถือเป็นปัจจัยหนึ่งที่จะช่วยสร้างความมั่นใจให้กับผู้สร้างแบบทดสอบ

ทซสุโอกะ (Tatsuoka. 1991 : 1; citing Graser. 1985) เสนอแนะว่าแบบทดสอบวัดผลสัมฤทธิ์สุคใหม่จะต้องสะท้อนให้เห็นลักษณะสำคัญ 4 ประการ ของผลการปฏิบัติหรือผลการสอบ คือ

1. ประเมินผลการปฏิบัติในแบบทดสอบที่ถูกต้องแบบเพื่อวัด
2. ประเมินโครงสร้างหรือตัวแทนความรู้และทักษะทางความคิด
3. วัดการเปลี่ยนแปลงที่เปลี่ยนไปของนักเรียน
4. ประเมินความเป็นอัตโนมัติของทักษะการปฏิบัติ

หลักการพื้นฐานในการคัดเลือกชนิดของข้อคำถามเพื่อใช้ ก็คือ คัดเลือกข้อคำถามชนิดที่ให้งานการปฏิบัติ ที่สามารถวัดผลการเรียนรู้ที่ตั้งใจได้ตรงมากที่สุด

การคัดเลือกชนิดของข้อคำถามที่เหมาะสม ถือเป็นขั้นตอนสำคัญขั้นตอนหนึ่งของกระบวนการพัฒนาแบบทดสอบ ปกติข้อคำถามมักถูกจัดเป็น 2 ประเภท คือ

1. ข้อคำถามแบบปรนัย (Objective Test Item) ซึ่งจะประกอบด้วยข้อคำถามหลายลักษณะที่ต่างกัน แต่อย่างไรก็ตาม เมื่อพิจารณาตามรูปแบบของการตอบสนองแล้วจะแบ่งเป็นรูปแบบที่ต้องการให้ผู้ตอบเติมคำตอบ (Supply The Answer) หรือที่ต้องการให้ผู้ตอบเลือกคำตอบจากตัวเลือกที่ให้ (Select The Answer) รูปแบบที่ให้เติมคำตอบ ได้แก่ แบบเติมคำตอบสั้น (Short Answer) และแบบเติมความให้สมบูรณ์ (Completion) ส่วนรูปแบบเลือกตอบนั้น ได้แก่ การเลือกตอบ แบบถูก-ผิด (True-False) แบบจับคู่ (Matching) และแบบเลือกตอบ (Multiple-Choice)

2. คำถามแบบอัตนัย (Essay Question) ข้อคำถามชนิดนี้วัดปริมาณของความ เป็นอิสระในการตอบสนองที่ให้ผู้เรียน ทั้งนี้จะรวมการตอบสนองที่ผู้เรียนมีอิสระเต็มที่ที่จะตอบ

(Extended Response) และการจำกัดการตอบ (Restricted Response) ทั้งนี้ เพื่อทำให้เกิดประสิทธิภาพในการวัดความรู้เกี่ยวกับความจริงมากขึ้น และเป็นการลดความยุ่งยากในการตรวจให้คะแนน

หากจะกล่าวในรายละเอียดจะพบว่า ข้อคำถามแบบปรนัยหลายชนิดมีลักษณะสำคัญประการหนึ่งที่ทำให้มีความแตกต่างจากแบบทดสอบอัตนัย คือ งานที่กำหนดให้ผู้ตอบตอบ มีลักษณะที่จำกัดชนิดของการตอบสนองค่อนข้างสูง เพื่อจะได้คำตอบที่ถูกต้อง ผู้ตอบจะต้องแสดงความรู้ความสามารถ ความเข้าใจ และทักษะเฉพาะ ซึ่งข้อคำถามต้องการ ผู้ตอบไม่มีอิสระที่จะอธิบายความหมายปัญหาใหม่ หรือจัดคำตอบใหม่ โดยใช้คำพูดของตนเองได้ เขาจะต้องเลือกคำตอบจากตัวเลือกทั้งหลาย หรือเติมคำตอบ จำนวน หรือเครื่องหมายที่ถูกต้องเท่านั้น ด้วยลักษณะของปัญหาที่เป็นแบบแผนและความจำกัดในวิธีการตอบสนอง เป็นผลให้สามารถให้คะแนนได้อย่างเป็นปรนัย รวดเร็ว และแม่นยำ แต่ในด้านลบ รูปแบบของข้อคำถาม ทำให้ข้อคำถามแบบปรนัยไม่มีความเหมาะสมที่จะนำไปใช้วัดความสามารถในการเลือก จัดระบบ หรือรวบรวมความคิด

ฉะนั้น หากจะกล่าวสรุปเกี่ยวกับข้อคำถามแบบปรนัย (Objective Test Item) ตามแนวคิดของกรอนลินน์ และลินน์ (Grönlund and Linn. 1990 : 121 - 140) ได้ว่า แบบทดสอบประเภทนี้จะนำเสนองานที่เป็นแบบแผนมาก ซึ่งต้องการให้ผู้ตอบจำกัดคำตอบของเขาโดยการเติมคำ ประโยค จำนวน หรือ สัญลักษณ์ หรือการเลือกคำตอบตัวเลือกที่กำหนดให้ ลักษณะของข้อคำถามดังกล่าวจะสามารถสนองความต้องการ และเป็นประโยชน์ต่อการวัดผลสัมฤทธิ์ของผู้เรียนในสถานการณ์เฉพาะที่ต้องการ และข้อดีของข้อคำถามรูปแบบนี้ ปรากฏดังภาพประกอบ 4

ด้าน	ลักษณะของข้อคำถามแบบปรนัย
1. ผลการเรียนรู้ที่วัด	มีประสิทธิภาพสำหรับการวัดความรู้ความจริง ความเข้าใจทักษะการคิด และผลที่ได้ซึ่งซับซ้อนบางประการ แต่ไม่เหมาะกับการวัดความสามารถในการเลือกจัดระบบ ความสามารถในการเขียน และทักษะการแก้ปัญหาบางประการ
2. การเตรียมข้อคำถาม	ต้องการข้อคำถามจำนวนมากสำหรับแบบทดสอบ การเตรียมยังอาจต้องใช้เวลามาก
3. การสุ่มตัวอย่างเนื้อหาวิชา	สามารถสุ่มตัวอย่างเนื้อหาวิชาได้กว้างขวาง เนื่องจากแบบทดสอบจะมีข้อคำถามจำนวนมาก
4. การควบคุมคำตอบการตอบสนองของผู้ตอบ	รูปแบบของงานที่มีแบบแผนมากจะจำกัดชนิดของการตอบสนองของผู้ตอบ ป้องกันการตอบ และหลีกเลี่ยงอิทธิพลของทักษะการเขียน แม้ว่าข้อคำถามแบบเลือกตอบจะทำได้
5. การตรวจให้คะแนน	ตรวจให้คะแนนได้อย่างรวดเร็ว ง่าย และถูกต้องตรงกัน
6. อิทธิพลต่อการเรียน	มักสนับสนุนผู้เรียนให้พัฒนาความเข้าใจความรู้ ความจริงเฉพาะ และความสามารถในการจำแนกแหว่งสิ่งเหล่านั้นสามารถสนับสนุนพัฒนาการความเข้าใจทักษะการคิด และผลผลิตที่ซับซ้อนได้ ถ้ามีการสร้างอย่างดี
7. ความเชื่อมั่น	มีความเชื่อมั่นสูง ถ้าได้มาจากแบบทดสอบที่สร้างมาอย่างดี

ภาพประกอบ 4 คุณสมบัติและข้อดีของข้อถามแบบปรนัย

สำหรับมุมมองของฮอปกินส์ และแอนส์ (Hopkins and Antes. 1979 : 26 - 29) แล้ว ข้อคำถามมี 3 ชนิด ตามลักษณะของงานที่ให้ผู้สอบตอบ คือ

1. ข้อคำถามชนิดปัญหา (Problem-Type Item) ซึ่งนำเสนอสถานการณ์ที่เป็นปัญหานำข้อใจให้ผู้ตอบ เพื่อให้ผู้ตอบสนองต่องานที่กำหนดค่าให้ ได้แก่ ปัญหาทางคณิตศาสตร์ (Mathematical Problem) ซึ่งรูปแบบของงานต้องการให้ผู้ตอบตัดสินใจเลือกวิธีที่ใช้ในการแก้ปัญหาและปัญหาวิชาการ (Technical Problem) เป็นปัญหาที่เขียนขึ้นใช้เฉพาะ เช่น ในงานทางวิทยาศาสตร์ เทคโนโลยี วิศวกร และธุรกิจ ทั้งนี้จะมีลักษณะที่ต้องการให้ผู้ตอบคิดแก้ปัญหาจากข้อมูลที่ให้

ข้อคำถามชนิดนี้ใช้วัดการตีความ การวิเคราะห์ความสัมพันธ์ การใช้หลักการ ความสามารถในการสร้างข้อสรุปและประเมินผล ซึ่งเป็นลักษณะของความรู้ความคิด

2. ข้อคำถามชนิดเติมคำตอบ (Supply Type Item) รูปแบบที่คำถามชนิดนี้ต้องการ คือ ให้ผู้ตอบสร้างการตอบสนองหรือคำตอบภายในกรอบโครงสร้างที่กำหนดค่าให้ ข้อคำถามชนิดเติมคำตอบเฉพาะถูกใช้ในวิธีการเหล่านี้ คือ

2.1 ข้อคำถามชนิดเติมคำตอบให้สมบูรณ์ (Completion Items) ต้องการให้ผู้ตอบทำให้ข้อความสมบูรณ์ โดยการเติมคำตอบในช่องว่างในรูปของคำ หรือชุดของคำที่มีความเชื่อมโยงกันอย่างปกติ ข้อคำถามแบบนี้สามารถใช้ในการวัดความรู้ความจริง นิยาม กิยคำ คำศัพท์ และความคิดรวบยอดพื้นฐาน ข้อคำถามแบบนี้จะเป็นประโยชน์มากที่สุดในการวัดพฤติกรรมที่มีความสัมพันธ์กับการระลึกเกี่ยวกับข้อเท็จจริง

2.2 ข้อคำถามชนิดคำตอบสั้น (Short-Answer Items) ต้องการให้ผู้ตอบตอบสนองต่องานที่กำหนดค่าให้ในรูปของวลีหรือประโยคสั้น ๆ ข้อคำถามนี้เหมาะสำหรับวัดความรู้ความจริงเช่นเดียวกับชนิดเติมคำตอบ แต่จะมีอิสระในการตอบสนองมากกว่า

2.3 ข้อคำถามแบบคำตอบขยาย (Extended-Response Essay Items) ข้อคำถามชนิดนี้จะนำเสนองานในลักษณะการเล่าเรื่องหรือคำถาม และต้องการให้ผู้ตอบเตรียมการพูดตอบที่พัฒนาอย่างดีแล้ว ขบวนการของงานที่กำหนดจะไม่จำกัด ผู้ตอบต้องตัดสินใจเกี่ยวกับความครอบคลุม ความลึก ของคำตอบเอง โดยคำนึงถึงเวลาที่กำหนดค่าให้ ข้อคำถามแบบนี้เหมาะกับการวัดการแสดงออกในการเขียน และทักษะในการจัดระบบการนำเสนออย่างมีเหตุผล เช่น ข้อคำถามอาจให้ผู้ตอบอธิบาย แสดงความสัมพันธ์ เปรียบเทียบ ตีความหมาย หรือสรุป

2.4 ข้อคำถามแบบจำกัดคำตอบ (Limited-Response Essay Items) ข้อคำถามนี้จะนำเสนองานซึ่งเหมาะกับการทดสอบผลสัมฤทธิ์ทางการเรียนในเนื้อหาวิชา แม้จะให้ผู้ตอบเป็นผู้สร้างคำตอบเอง แต่จะถูกกำหนดขอบเขตการตอบไว้อย่างชัดเจน

3. ข้อคำถามแบบเลือกตอบ (Selection-Type Item) ข้อคำถามแบบเลือกตอบต้องการให้ผู้ตอบเลือกคำตอบจากทางเลือกที่ให้ ตัวเลือกเหล่านี้อาจถูกนำเสนอเป็นส่วนหนึ่งของข้อคำถาม หรืออาจถูกให้ความหมายจากคำแนะนำของแบบทดสอบ ข้อคำถามแบบนี้จะถูกใช้ในวิธีต่อไปนี้

3.1 ข้อคำถามแบบถูก-ผิด (True-False Items) ผู้ตอบจะต้องระบุว่า ข้อความที่แสดงนั้นถูกหรือผิด ข้อคำถามแบบนี้เป็นประโยชน์ต่อการวัดเกี่ยวกับความรู้ความจริงเฉพาะ ความรู้หรือความเข้าใจหลักการ และการอ้างสรุป ของผู้เรียน

3.2 ข้อคำถามแบบเลือกตอบ (Multiple-Choice Items) ถูกนำเสนอผ่านคำถามหรือข้อคำถามที่ไม่สมบูรณ์ซึ่งทำหน้าที่เหมือนข้อเสนอ และผู้ตอบเลือกคำตอบที่ถูกต้องหรือคำตอบที่ดีที่สุดจากชุดของตัวเลือกที่กำหนดให้ การวัดในระดับความเข้าใจและการนำไปใช้ทำได้ง่าย โดยการใช้ข้อคำถามแบบเลือกตอบ อาจใช้สำหรับวัดพฤติกรรมในทุกระดับ และนำไปประยุกต์ใช้ในการวัดความรู้เกี่ยวกับข้อเท็จจริง ความสัมพันธ์เชิงเหตุผล การวิเคราะห์อย่างมีวิจารณญาณ การตัดสินใจ และความสามารถในการจำแนก

3.3 ข้อคำถามแบบจับคู่ (Matching-Items) ต้องการให้ผู้ตอบระบุความสัมพันธ์ระหว่างชุดของการอ้างหลักฐานและชุดของการตอบสนอง เช่น ชุดของคำถามสามารถวัดความสัมพันธ์ของข้อและเหตุการณ์ วันที่กับเหตุการณ์ ฯลฯ

3.4 ข้อคำถามจำแนกประเภท (Classification Items) ต้องการให้ผู้ตอบจัดประเภทโดยยึดรายละเอียดของการแบ่งพวก ข้อคำถามแบบนี้สามารถนำเสนอในรูปแบบของชุดของคำ ประโยคที่ไม่สมบูรณ์หรือสมบูรณ์ ผู้ตอบจะต้องตอบโดยถูกจำกัดอยู่ในชุดของประเภทเฉพาะ ข้อคำถามแบบนี้จึงใช้ในการวัดความสามารถในการเชื่อมโยงประเภทและชนิดของความรู้เกี่ยวกับข้อเท็จจริง

ดังได้นำเสนอความเห็นของทซึสุโอกะ (Tatsuoka, 1991) เกี่ยวกับลักษณะของแบบทดสอบในยุคใหม่แล้วว่า การทดสอบแนวใหม่จะต้องมีธรรมชาติที่เป็นทั้งปริมาณและคุณภาพ การวัดผลสัมฤทธิ์ที่เกิดขึ้นจะต้องเป็นทั้งการบรรยายและตีความในรูปของกระบวนการที่กำหนดให้ปฏิบัติ

คะแนนจากแบบทดสอบมาตรฐานจะเป็นประโยชน์สำหรับการคัดเลือก การรับรองคะแนนไม่สามารถให้ข้อมูลที่เป็นประโยชน์ต่อครูสำหรับการวางแผนแก้ไข นัก แบบทดสอบชนิดให้ผู้สอบสร้างการตอบสนองเอง (Constructed-Response Formats) เป็นที่ต้องการสำหรับวัดกระบวนการคิดที่ซับซ้อน และเคลื่อนที่เปลี่ยนแปลงตลอดเวลา (Dynamic) (Bennett and others. 1990) ขณะที่ข้อคำถามแบบเลือกตอบเหมาะกับการวัดความรู้ที่คงที่ไม่เคลื่อนไหว (Static Knowledge)

แบบทดสอบที่สร้างขึ้นจะแสดงทักษะการคิด รูปแบบการตอบสนองที่ต้องการจะต่างไปจากเดิม คือ ต้องการข้อคำถามที่ให้ผู้เรียนเขียนแผนภูมิที่แสดงความสัมพันธ์ที่ซับซ้อนระหว่างงานทักษะและทางแก้ปัญหาจะถูกแสดงเป็นกราฟ หรือถามให้ผู้สอบอธิบายความสัมพันธ์ดังกล่าวด้วยคำพูด คำถามอาจเป็นรูปภาพ หรืออาจให้ผู้ตอบลำดับความสัมพันธ์เชิงเหตุระหว่างความคิดรวบยอดต่าง ๆ และเชื่อมโยงมันโดยกราฟ นอกจากนี้ การวัดจำเป็นจะต้องถามด้วยรูปแบบใหม่ที่ทำให้คะแนนที่ซับซ้อนขึ้นมากกว่าถูก-ผิด การตอบสนองที่ถูกต้องสำหรับคำถามข้อหนึ่งอาจกำหนดโดยดูว่า งานการคิดทั้งหมดที่รวมอยู่ในข้อคำถามสามารถตอบคำถามได้ถูกต้องทั้งหมดหรือไม่

ข้อคำถามอีกรูปแบบหนึ่ง ซึ่งกรอนแลนด์ และลินน์ (Gronlund and Linn. 1990 : 192) ได้นำเสนอ และมีความเชื่อว่าสามารถวัดผลการเรียนรู้ที่ซับซ้อน หรือทักษะการคิดระดับสูง อาทิ ความเข้าใจ เหตุผล ความคิดอย่างมีวิจารณญาณ ความคิดเชิงวิทยาศาสตร์ ความคิดริเริ่มสร้างสรรค์ และการแก้ปัญหาได้ เป็นที่ยอมรับและใช้กันอย่างกว้างขวาง รูปแบบนี้ คือ แบบฝึกการตีความหมาย (Interpretive Exercise)

รูปแบบของการวัดจะประกอบด้วยข้อคำถาม ที่จำเป็นต้องใช้ชุดของข้อมูลที่กำหนดให้ ซึ่งอาจอยู่ในรูปของงานเขียน ตาราง แผนภูมิ เส้นภาพ แผนที่ หรือรูปภาพ ในการตอบคำถาม ทั้งนี้ผู้ตอบจะถูกถามให้ระบุความสัมพันธ์ของข้อมูล ให้สร้างข้อสรุปที่ถูกต้อง ประเมินข้อตกลงเบื้องต้นและข้อวินิจฉัย และนำข้อมูลไปใช้อย่างเหมาะสม โดยตอบคำถามในลักษณะดังนี้

1. ใช่-ไม่ใช่
2. สันนิษฐาน-ขัดแย้ง-ไม่ทั้ง 2 อย่าง
3. จำเป็น-ไม่จำเป็น
4. ถูก-ผิด
5. แบบเลือกตอบ

ไช้เวอร์ (Schiever. 1991 : 79) เป็นอีกผู้หนึ่งที่ได้เสนอความเห็นเกี่ยวกับหลักการทั่วไปของการกำหนดงาน หรือการตั้งคำถามที่จะทำให้ผู้ตอบต้องคิด คือ คำถามแบบปลายเปิด (Open-Ended) เพื่อทำให้เกิดความคิดหลากหลาย คำถามที่ต้องใช้ข้อมูลประกอบการตอบปัญหา คำถามที่ให้พิสูจน์หรือหาเหตุผล ซึ่งหลักการนี้สามารถนำไปประยุกต์ใช้ได้กับทุกเนื้อหาวิชา ชั้นเรียน สถานการณ์ และปัญหา

แบบทดสอบนอกจากจะแบ่งตามลักษณะงานที่กำหนดให้ และรูปแบบการตอบสนองแล้ว ยังมีการจำแนกตามลักษณะของการตรวจให้คะแนน คือ แบบทดสอบที่ให้คะแนนเป็นชุดเดียว (Single Score) และแบบที่ให้คะแนนจำแนกตามความสามารถที่สอบวัดอีกด้วย

แบบทดสอบวัดความสามารถด้านการเรียนรู้ ส่วนใหญ่ที่นำไปใช้ในโรงเรียนเป็นแบบทดสอบที่ใช้ดำเนินการทดสอบเป็นกลุ่ม บางฉบับจะให้คะแนนชุดเดียว ขณะที่บางฉบับให้คะแนนที่วัดลักษณะความสามารถ 2 ประการ หรือมากกว่านั้น

แบบทดสอบที่ให้คะแนนชุดเดียวถูกออกแบบเพื่อวัดความสามารถทางการเรียนรู้ทั่วไปของนักเรียน โดยจะมีข้อคำถามหลายชนิดรวมในแบบทดสอบ แต่ไม่เน้นความสามารถ หรือทักษะเฉพาะใดในคะแนนรวม ฉะนั้น ความสามารถเฉพาะ เช่น ภาษา ตัวเลข และเหตุผล จะถูกผสมผสานกันเป็นการวัดความสามารถทั้งหมดโดยรวม ตัวอย่างแบบทดสอบประเภทนี้ที่ใช้กันอย่างกว้างขวาง คือ โอตีส-เลนเนน (Otis-Lennen School Ability Test) ซึ่งใช้กับนักเรียนเกรด 1 - 2

แบบทดสอบที่จำแนกส่วนของคะแนนออกไปมากกว่า 2 ส่วน ได้แก่ School and College Ability Tests (SCAT เกรด 3.5 ถึง 12) จะให้คะแนนความสามารถทางภาษาและปริมาณ ส่วนแบบทดสอบวัดความสามารถทางการคิด (Cognitive Abilities Test) ใช้กับนักเรียนอนุบาล ถึงเกรด 12 ให้การวัดความสามารถด้านเหตุผลที่สำคัญสำหรับการเรียนรู้ในโรงเรียน 3 ประการ และให้คะแนนเป็น 3 ส่วน คือ ภาษา ปริมาณ และไม่ใช่ภาษา (Non-Verbal) ส่วนของภาษา ได้แก่ การจำแนกทางภาษา การเติมคำ การอุปมาอุปไมยทางภาษา ซึ่งเน้นความสามารถด้านเหตุผล ส่วนที่เกี่ยวกับปริมาณ ผู้เรียนต้องการแก้ปัญหาทางเชิงปริมาณที่มีได้สอนในโรงเรียนโดยตรง สำหรับชุดไม่ใช่ภาษาประกอบด้วยแบบทดสอบ 3 ชุด ที่ใช้รูปทรงทางเรขาคณิตและภาพ ซึ่งต้องการให้นักเรียนที่เป็นประสบการณ์นอกโรงเรียน เนื่องจากแบบทดสอบไม่ได้ใช้ทั้งคำพูดและตัวเลข ชุดที่ไม่ใช่ภาษานี้จะสามารถพยากรณ์ความสำเร็จในโรงเรียนได้น้อยกว่าส่วนที่เป็นภาษาและปริมาณ อย่างไรก็ตาม

ก็ตาม ก็ให้ความแม่นยำในการวัดความสามารถด้านเหตุผลของผู้มีความสามารถในการอ่านต่ำ และผู้มีความบกพร่องทางทักษะด้านภาษา

แบบทดสอบที่ออกแบบเพื่อวัดความสามารถในการเรียนเดิมนั้น ถูกเรียกว่า แบบทดสอบวัดสติปัญญา (Intelligence Tests) ซึ่งคำนี้ก็ยังคงใช้อยู่กับแบบทดสอบบางฉบับ แต่ลดลงด้วยเหตุผลหลายประการ คือ

1. ผู้คนส่วนใหญ่หันมาเชื่อมโยงความคิดรวบยอดของสติปัญญากับความสามารถ ที่มีได้ติดตัวมา
2. มีการโต้เถียงกันเกี่ยวกับความเหมาะสมของสติปัญญาและองค์ประกอบที่ควรรวมอยู่ในความคิดรวบยอดของสติปัญญาเพิ่มขึ้น
3. แบบทดสอบในเรื่องนี้ถูกนำไปใช้พยากรณ์ผลสัมฤทธิ์ และอธิบายความสามารถในการเรียนเพิ่มขึ้น จึงได้มีการเปลี่ยนคำว่า แบบทดสอบวัดสติปัญญา (Intelligence Test) ไปเป็นแบบทดสอบวัดความสามารถทางการเรียน (Learning Ability Tests) แบบทดสอบความสามารถในการคิด (Cognitive Ability Tests) และแบบทดสอบวัดความถนัดทางวิชาการ (Scholastic Aptitude Test) ซึ่งคำเหล่านี้เน้นความจริงที่ว่าแบบทดสอบเหล่านี้วัดความสามารถที่พัฒนาได้ (Developed Abilities) ซึ่งเป็นประโยชน์ในการเรียน และไม่ใช้ความสามารถที่ติดตัวมาแต่กำเนิด หรือศักยภาพที่ไม่สามารถพัฒนา ซึ่งเป็นประโยชน์ในการเรียน

คะแนนของผู้เรียนจากแบบทดสอบความถนัดทางวิชาการ หรือแบบทดสอบวัดความสามารถจะตีความหมายได้ดีที่สุด เมื่อวัดความสามารถในการเรียนรู้ปัจจุบัน ผลการสอบได้รับอิทธิพลหรือผลักดันโดยองค์ประกอบต่าง ๆ เช่น ลักษณะที่สืบทอดมา ประสบการณ์เบื้องหลัง แรงจูงใจ ทักษะเฉพาะ ความสนใจ ความมั่นใจในตนเอง และการปรับตัวทางด้านอารมณ์ องค์ประกอบเหล่านี้ เป็นส่วนหนึ่งของความสามารถในปัจจุบันในการปฏิบัติที่จะมีผลต่อคะแนนการทดสอบและผลสัมฤทธิ์ในโรงเรียน องค์ประกอบเหล่านี้สามารถถูกปรับโดยประสบการณ์การศึกษา ดังนั้น ทั้งความสามารถในการเรียนรู้และผลสัมฤทธิ์ในโรงเรียนจะเพิ่มขึ้น

ความสามารถในการเรียนรู้ที่วัดเป็นรายบุคคล มักเรียกกันว่าแบบทดสอบวัดสติปัญญา (Intelligence Tests) แบบทดสอบนี้จะถูกนำไปใช้กับผู้สอบรายบุคคล ณ เวลาใดเวลาหนึ่งในลักษณะตัวต่อตัว ปัญหาจะถูกนำเสนอโดยปากเปล่า โดยผู้ดำเนินการสอบ และผู้สอบจะตอบสนองโดยชี้หรือตอบปากเปล่า แบบทดสอบวัดความสามารถรายบุคคลที่ใช้กันกว้างขวาง ได้แก่ สแตนฟอร์ด-บินเน็ต (Stanford-Binet Scale) ซึ่งการพิมพ์ครั้งที่ 4 ประกอบด้วย

แบบทดสอบย่อย 15 ฉบับ ใน 4 กลุ่มความสามารถ (Gronlund and Linn. 1990 : 302 - 303)

1. เหตุผลทางภาษา (Verbal Reasoning)

- 1.1 คำศัพท์ : ให้นิยามคำ
- 1.2 ความเข้าใจ : ระบุและให้เหตุผลเกี่ยวกับเหตุการณ์
- 1.3 Absurdities : อธิบายว่าอะไรในภาพผิดหรือไม่น่าเป็นไปได้
- 1.4 ความสัมพันธ์ทางภาษา : บอกความแตกต่าง

2. เหตุผลเชิงปริมาณ (Quantitative Reasoning)

- 2.1 ปริมาณ : นับ คำนวณ และแก้ปัญหา
- 2.2 อนุกรม : เติมจำนวนตามลำดับ
- 2.3 สร้างสมการ : จัดลำดับจำนวนและสัญลักษณ์เพื่อสร้างสมการ

3. เหตุผลเชิงนามธรรม (Abstract or Visual Reasoning)

- 3.1 การวิเคราะห์รูปแบบ : สร้างรูปแบบ โดยใช้สี่เหลี่ยมลูกบาศก์และรูปสี่เหลี่ยม
- 3.2 การลอกแบบ : การลอกแบบ การลอกภาพ โดยใช้บล็อก และการวาดภาพ
- 3.3 แมทริกซ์ : เติมแมทริกซ์ที่ขาดหายไปให้สมบูรณ์
- 3.4 การพับและตัดกระดาษ : ระบุภาพของกระดาษที่ไม่ได้มัน หลังจากดูภาพที่มีมันและตัด

4. ความจำระยะสั้น (Short-Term Memory)

- 4.1 ความจำเกี่ยวกับลูกบิด : บอกรูปแบบของลูกบิด
- 4.2 การจำประโยค : กล่าวประโยคซ้ำคำต่อคำ
- 4.3 การจำตัวเลข : กล่าวตัวเลขซ้ำไปข้างหน้า-ถอยหลัง
- 4.4 การจำสิ่งของ : ระบุลำดับที่เหมาะสมของภาพ

แบบทดสอบถูกจัดกลุ่มโดยยึดกลุ่มความสามารถ 4 กลุ่ม สำหรับการตรวจให้คะแนน และสร้างเส้นภาพ แต่ไม่ได้ดำเนินการสอบตามลำดับดังกล่าว

แบบทดสอบวัดความสามารถอีกฉบับที่มีชื่อเสียง คือ เวคสเลอร์ (Wechsler Scales) มี 3 ฉบับ ครอบคลุมทุกระดับอายุ ตั้งแต่ 4 ขวบ จนถึงผู้ใหญ่ คือ Wechsler Preschool and Primary Scale of Intelligence (WPPSI-R) (อายุ 4-6.5 ปี)

Wechsler Intelligence Scale for Children (WISC-R) (อายุ 6.5-16.5 ปี) และ Wechsler Adult Intelligence Scale (WAIS-R) (อายุ 16 ปี ถึงผู้ใหญ่)

แบบทดสอบเวคส์เลอร์จัดเป็นแบบทดสอบชุดย่อย ๆ (Subtests) ประกอบด้วยมาตรวัดที่เป็นภาษา (Verbal Scale) ประกอบด้วยแบบทดสอบย่อย 5 ฉบับ คือ ความรู้ ความคล้ายคลึง เลขคณิต คำศัพท์ และความเข้าใจในการอ่าน มาตรวัดการปฏิบัติ (Performance Scale) ประกอบด้วย 5 ฉบับย่อย เช่นกัน คือ การเติมภาพให้สมบูรณ์ การจัดภาพ การออกแบบบล็อก การวัดสิ่งของ และการลงรหัส

แต่ละฉบับย่อยใน WISC-R จะดำเนินการสอบและให้คะแนนแยกจากกัน โดยจะดำเนินการสอบสลับกันระหว่างแบบทดสอบวัดที่เป็นภาษาและการปฏิบัติ เพื่อให้เกิดความสนใจ คะแนนดิบจะถูกแปลงไปเป็น Scaled Scores ที่มีค่าเฉลี่ย = 10 , SD = 3 แล้วจะถูกนำมารวมกันเป็น IQ ด้านภาษา (Verbal IQ) และ IQ ด้านการปฏิบัติ (Performance IQ) และ IQ รวม ซึ่งเป็นความเบี่ยงเบนของ IQ ที่มีค่าเฉลี่ยเท่ากับ 100 และมีความเบี่ยงเบนมาตรฐานเท่ากับ 15

แบบทดสอบวัดความสามารถของบีเนต์และเวคส์เลอร์ ถูกนำไปใช้อย่างกว้างขวางสำหรับทดสอบนักเรียนรายบุคคล โดยทั้ง 2 ฉบับ จะให้ผลที่เชื่อถือได้สูงประมาณ 0.90 สำหรับคะแนนทั้งฉบับ แบบทดสอบทั้ง 2 ฉบับ สามารถพยากรณ์ความสำเร็จในโรงเรียนได้ดีพอ ๆ กัน

หากพิจารณาในรายละเอียดของแบบทดสอบย่อยแต่ละฉบับ จะเห็นว่าแบบทดสอบของบีเนต์แสดงให้เห็นรูปแบบการวัดความสามารถด้านเหตุผลได้ค่อนข้างชัดเจนกว่าของเวคส์เลอร์

Culture-Fair Intelligence Tests ของ แคทเทล (Cattell) เป็นแบบทดสอบอีกฉบับหนึ่งที่วัดความสามารถ แต่แบบทดสอบฉบับนี้จะมีลักษณะพิเศษ คือ พยายามขจัดปัญหาเกี่ยวกับวัฒนธรรมอันอาจจะมีผลต่อผลการทดสอบออกไป โดยเฉพาะภาษาซึ่งถือว่ามีส่วนค่อนข้างเด่นชัดต่อพฤติกรรมทดสอบ ฉะนั้น แบบทดสอบฉบับนี้จึงเป็นแบบทดสอบที่ไม่ใช้ภาษา (Nonverbal) แต่จะใช้รูปภาพ แผนภูมิ แทนแบบทดสอบ ประกอบด้วยแบบทดสอบย่อย 8 ฉบับ มี 4 งาน คือ อนุกรม (Series) จำแนกประเภท (Classification) แมทริกซ์ (Matrices) และเงื่อนไข (Condition) ใช้ดำเนินการสอบได้กับกลุ่มและรายบุคคล

ผลงานของกิลฟอร์ดเกี่ยวกับโครงสร้างของสติปัญญา (Guilford's structure of Intellect Model) มีอิทธิพลมากต่อการเคลื่อนไหวด้านการทดสอบ จากที่เป็นการวัด

แบบจำกัดจำนวนของความสามารถทั่วไปของสมอง ไปสู่การวัดความสามารถเฉพาะที่หลากหลาย โดยมีพื้นฐานจากการวิจัยนานนับปี และใช้วิธีการวิเคราะห์องค์ประกอบ ทั้งนี้ได้เสนอโครงสร้างทางสถิติปัญารูปแบบ 3 มิติ ซึ่งนำไปสู่การพัฒนาแบบทดสอบวัดความถนัด Differential Aptitude Tests : DAT ชุดของ DAT ประกอบด้วยแบบทดสอบ 8 ฉบับ คือ

1. เหตุผลเชิงภาษา (Verbal Reasoning : VR)
2. ความสามารถด้านตัวเลข (Numerical Ability : NA)
3. เหตุผลเชิงนามธรรม (Abstract Reasoning : AR)
4. ความเร็วและแม่นยำเชิงเสมียน (Clerical Speed and Accuracy : CSA)
5. เหตุผลเชิงจักรกล (Mechanical Reasoning : MR)
6. ความสัมพันธ์เกี่ยวกับระยะทาง (Space Relations : SR)
7. การสะกดคำ (Spelling : S)
8. การใช้ภาษา (Language Usage : LU)

แต่ละฉบับจะให้คะแนนแยกจากกัน แต่มีการรวมคะแนนภาษาและตัวเลข ซึ่งให้ดัชนีของความถนัดเชิงวิชาการเหมือนกับที่ได้จากแบบทดสอบความสามารถทางการเรียนรู้

แบบทดสอบ VR รูปแบบข้อคำถามเป็นปรนัยชนิดเลือกตอบ ทั้งนี้โดยข้อคำถามเป็นการเติมคำในช่องว่าง เพื่อให้ประโยคมีความหมายและเป็นจริง ส่วนแบบทดสอบ AR ซึ่งวัดความสามารถด้านเหตุผล รูปแบบข้อคำถามเป็นปรนัยชนิดเลือกตอบเช่นกัน แต่ข้อคำถามเป็นปัญหาเกี่ยวกับภาพในลักษณะของอนุกรม ดังนั้น ข้อคำถามเป็นแบบปรนัย ชนิดเลือกตอบทั้งหมด ส่วนชนิดของข้อคำถามก็มีลักษณะที่คล้ายกับแบบทดสอบผลสัมฤทธิ์ทั่วไป จะเห็นว่าแบบทดสอบ DAT ที่วัดความสามารถด้านเหตุผลมีรูปแบบข้อคำถามเป็นแบบปรนัยชนิดเลือกตอบทั้งหมด ส่วนชนิดของข้อคำถามมีลักษณะที่คล้ายกับแบบทดสอบวัดผลสัมฤทธิ์ทั่วไป

แบบทดสอบถือเป็นศูนย์กลางของการเรียนการสอน แบบทดสอบใด ๆ และการทดสอบปลายภาคหลักไม่พ้นการฉายภาพงานที่มีมาก่อน ดังนั้น แบบทดสอบไม่เพียงแต่เป็นมาตรฐานการติดตาม แต่จะสร้างมาตรฐานด้วย (Wiggins. 1989 : 703 - 713)

ปัญหารูปแบบหนึ่งที่คนเผชิญในชีวิตจริง เป็นปัญหาที่สำคัญจริง ๆ ในสังคม การเมือง และวิทยาศาสตร์ ในโลกปัจจุบัน ปัญหานี้เรียก "Ill-Structured Problems" ในสถานการณ์ของปัญหาประเภทนี้ กระบวนการแก้ปัญหาที่มีความสำคัญในกรณีที่ไม่มีคำตอบถูกชัดเจน

ทางแก๊จะโน้มเอียงไปกับการให้ความหมายโดยใช้เหตุผลหรือการแสดงหลักฐาน เพื่อพิจารณาตัดสินใจ (Frederikson. 1984 : 193 - 202)

ผลของรูปแบบข้อคำถาม (Item Format) ที่มีต่อการตอบสนองของผู้สอบได้รับความสนใจศึกษาจากนักจิตวิทยาและนักจิตวิทยาอย่างกว้างขวาง มาเป็นเวลานานเช่นกัน โดยเฉพาะข้อคำถามปลายเปิด (Open-Ended) หรือที่รู้จักกันในลักษณะของการตอบสนองอย่างอิสระ (Free-Response) หรือข้อคำถามแบบระลึกได้ (Recall Items) กับข้อคำถามชนิดเลือกตอบ (Multiple-Choice) หรือที่เรียกว่า ข้อคำถามชนิดจำได้ (Recognition Items) จากมุมมองของกลุ่มประมวลผลข้อความจริง เห็นว่ารูปแบบการตอบสนองทั้งสองมีความแตกต่างกัน ขณะที่ข้อคำถามชนิดจำได้ต้องการเพียงความสามารถในการจำแนกระหว่างข้อมูลที่นำเสนอ แต่ข้อคำถามชนิดระลึกได้ต้องการความสามารถในการค้นหาและดึงข้อความจริงมาใช้ตามสถานการณ์และเงื่อนไขได้อย่างเหมาะสม

การเปรียบเทียบระหว่างรูปแบบทั้งสองใช้เกณฑ์หลายประการ เช่น อัตราของความจำเริญ (Estes and DaPolito. 1967 ; Heim and Watts. 1967 ; Loftus and Loftus. 1976) ความยากของข้อคำถาม (Cook. 1955 ; Merwin and Womer. 1969) ลักษณะที่ถูกต้อง (Traub and Fisher. 1977 ; Ward, Frederiken and Carlson. 1980) อัตราการคงไว้ (Retention Rate) (Duchastel and Nungester. 1982) และสถิติที่เตรียมตัวสำหรับการสอบของผู้สอบ (Freund, Brelsford and Atkinson. 1969 ; Loftus. 1971 ; Tversky. 1973) ผลการศึกษาเป็นไปอย่างกว้างขวาง มีเพียงบางส่วนที่ชี้ให้เห็นประโยชน์ของข้อคำถามแบบปลายเปิด ขณะที่ส่วนหนึ่งแสดงให้เห็นว่า ไม่มีความแตกต่างกันระหว่างข้อคำถาม 2 รูปแบบ

การศึกษาส่วนใหญ่ใช้แบบทดสอบความเข้าใจในการอ่านและคำศัพท์ ซึ่งถือเป็นมวลความรู้ที่สามารถนำการระลึกและการจำไปใช้ได้ (Birenbaum and Tatsuoka. 1987 : 385)

ไบเรนบอม และทซุสโอะกะ (Birenbaum and Tatsuoka. 1987 : 385 - 395) ได้ทำการวิจัยเพื่อตรวจสอบผลของการใช้รูปแบบการตอบสนองในข้อคำถามที่เป็นแบบปลายเปิด (Open-Ended) กับแบบเลือกตอบ (Multiple Choice) ว่า เมื่อนำไปใช้ในการวินิจฉัยความเข้าใจผิดในงานที่มีลักษณะเป็นวิธีการแล้ว ข้อมูล 2 ชุด ที่ได้จากข้อคำถาม 2 รูปแบบ สามารถสะท้อนให้เห็นจำนวนชนิดของความคลาดเคลื่อนและวินิจฉัยแหล่งของความเข้าใจผิดได้เพียงใด

กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียน เกรด 8 จำนวน 285 คน จัดกลุ่มให้ทำข้อสอบแบบปลายเปิด 148 คน และทำข้อสอบแบบเลือกตอบ 137 คน ทั้ง 2 กลุ่มประกอบด้วยนักเรียนที่เรียนปฏิบัติการทดลองทางคณิตศาสตร์เป็นเวลา 2 ปี ติดต่อกันเช่นกัน แบบทดสอบรูปแบบปลายเปิดที่ใช้ ประกอบด้วยข้อความ 38 ข้อ ที่จัดเป็นแบบทดสอบคู่ขนาน ชุดละ 19 ข้อ สำหรับแบบทดสอบรูปแบบเลือกตอบประกอบด้วยข้อความเหมือนกัน มี 5 ตัวเลือก ตัวลวงสร้างโดยยึดคำตอบที่นักเรียนตอบผิดจากแบบทดสอบรูปแบบปลายเปิดที่ศึกษาในลักษณะนำร่องไปแล้ว จัดสอบโดยการเขียนตอบทั้ง 2 รูปแบบ ผลการศึกษาแสดงให้เห็นว่าแม้กลุ่มผู้สอบข้อความแบบปลายเปิดและแบบเลือกตอบ จะไม่แตกต่างกันในความสามารถในการแก้ปัญหาการบวกเศษส่วน แต่ข้อความรูปแบบปลายเปิดก็ให้ผลเกี่ยวกับการพิจารณาจำนวนชนิดของความคลาดเคลื่อน และวินิจฉัยแหล่งของความเข้าใจผิดได้เป็นที่น่าสนใจมากกว่า

เบอร์ก (Bergh. 1990 : 1 - 12) เป็นผู้หนึ่งที่มีความสนใจที่จะหาคำตอบว่าข้อความที่มีรูปแบบต่างกัน จะดึงการทำหน้าที่ของสมองที่เหมือนกันออกมาได้หรือไม่เพียงไร ผลการศึกษาวิจัยที่ผ่านมามีค่อนข้างคลุมเครือ ทั้งนี้เพราะนักวิจัยบางคนอ้างว่าไม่มีความแตกต่างระหว่างแบบทดสอบที่มีรูปแบบข้อความที่ต่างกัน ขณะที่บางคนพบข้อสรุปที่ตรงกันข้าม อาทิ ไบเรนบอร์ม และทักส์วอเกะ (1987) ดังได้กล่าวมาแล้ว

การศึกษาเกือบทั้งหมดขาดพื้นฐานเชิงทฤษฎีเกือบทุกกรณี ความแตกต่างของคะแนนที่ได้ซึ่งเกี่ยวข้องกับรูปแบบของข้อความ จะถูกแปลความหมายและอธิบายในเชิงความสัมพันธ์ของความแตกต่างระหว่างการระลึกได้ (Recall) กับการจำได้ (Recognition) ในกรณีนี้สาระของการระลึกได้เป็นการก่อให้เกิดหรือสร้างคำตอบต่อข้อความ ผู้ตอบต้องสร้างคำตอบในรูปของการตอบปากเปล่า หรือการเขียน หรืออธิบายความคิด ถ้าเป็นคำตอบที่ต้องการส่วนการจำได้ตัวเลือกจะถูกนำเสนอให้ผู้ตอบ จึงไม่จำเป็นต้องมีการสร้างคำตอบออกมา และจะใช้ทฤษฎีของการเปรียบเทียบ และตัดออกไปเป็นพื้นฐานในการเลือกคำตอบ ข้อคำถามแบบปลายเปิดและแบบเลือกตอบจะถูกนำมาใช้ในการวัดความสามารถทางภาษาในรูปแบบของการระลึกได้ และจำได้พอ ๆ กับการประเมินความรู้เพียงบางส่วนที่วัดโดยข้อความแบบเลือกตอบ

เหตุผลที่ทำให้ได้ข้อสรุปแตกต่างกัน เกี่ยวกับความสามารถของรูปแบบของข้อความในการวัดความสามารถทางสมอง คือ

ประการแรก การศึกษาขาดกรอบโครงสร้างทางทฤษฎีรองรับ เช่น นักวิจัยบางคน แปลความหมายผลความแตกต่างระหว่างค่าเฉลี่ยของแบบทดสอบ 2 ฉบับ ขณะที่บางคนแปลความหมายจากความแปรปรวน เป็นต้น ซึ่งความแตกต่างในความสามารถด้านสถิติปัญหาที่ถกเถียงอาจเป็นข้อสรุปที่เร็วเกินไป เพราะว่ากระบวนการเดียวกันอาจถูกรวมอยู่ทั้งในการจำได้และระลึกได้

ประการที่สอง ความไม่เหมาะสมของการออกแบบงานวิจัย งานวิจัยส่วนใหญ่มีแบบแผนการวิจัยแบบวัดซ้ำ และเกือบทั้งหมดก็จะให้ทำการทดสอบโดยใช้ข้อคำถามแบบปลายเปิดเป็นลำดับแรก ตามด้วยแบบทดสอบชนิดเลือกตอบอื่น ๆ จากนั้นคำนวณค่าสหสัมพันธ์ระหว่างคะแนนที่ได้จากแบบทดสอบที่ต่างกัน ถ้าผู้ตอบสนองต้องตอบคำถามที่เหมือนกัน แม้จะมีรูปแบบข้อคำถามต่างกันถึง 6 ครั้ง ในช่วงเวลาไม่กี่สัปดาห์ สิ่งที่น่าจะต้องถูกถาม ก็คือ จากครั้งแรกถึงครั้งสุดท้ายวัดคุณลักษณะเดียวกันหรือไม่ (Traub and Fisher. 1977 : 355 - 369) วิธีดำเนินการสอบสามารถพัฒนาให้ดีขึ้นได้โดยการสุ่มดำเนินการสอบในแต่ละโอกาส (Hogan. 1981) การศึกษาลักษณะนี้ ความสนใจหลักจะอยู่ที่การละความแปรปรวนของคะแนนในการทดสอบโดยใช้แบบทดสอบที่มีรูปแบบต่างกัน ดังนั้น สำหรับแบบแผนการวิจัยแบบวัดซ้ำที่แบบทดสอบทั้งหมดถูกนำไปดำเนินการในทุกโอกาส น่าจะเป็นทางแก้ปัญหาที่เหมาะสมที่สุด

อนึ่ง มีผู้ให้ความสนใจศึกษาผลของการสอบที่มีต่อการเรียนการสอน โดยเปรียบเทียบแบบทดสอบเลือกตอบ กับ แบบสร้างคำตอบเอง (Constructed Response) ทั้งนี้ เชปอาร์ด (Shepard. 1991 : 24) แสดงความเห็นว่ แบบทดสอบเลือกตอบส่งเสริมให้เกิดเจตคติที่ไม่ดีต่อการเรียนรู้ และการอ้างสรุปที่ไม่ถูกต้องเกี่ยวกับเป้าหมายของมัน (เช่น มีคำตอบเพียงคำตอบเดียว และคำตอบถูกนั้นผู้สร้างแบบทดสอบและครูรู้แล้วว่าข้อใดถูก ซึ่งเป็นงานของนักเรียนที่จะต้องค้นหาคำตอบ อาจจะโดยการเดาถ้าจำเป็น) ผลงานวิจัยสนับสนุนความคิดของเชปอาร์ดที่พบว่า ผู้เรียนพอใจที่จะสอบโดยใช้ข้อสอบแบบเลือกตอบ แม้ว่ามันจะเห็นได้ว่า แบบทดสอบแบบสร้างคำตอบเองเป็นดัชนีบ่งชี้ที่แม่นยำกว่าความพอใจของนักเรียนต่อแบบทดสอบเลือกตอบ เกิดมาจากการรับรู้ว่าการเตรียมตัวที่น้อยกว่า และสร้างความมั่นใจในตนเองได้มากกว่า (Bennett. 1991 : 24; citing Traub and Mac Rury. 1990)

เจตคติเหล่านี้ชี้ให้เห็นว่า การวัดชนิดให้สร้างคำตอบเองเป็นสาเหตุให้ผู้เรียนทำงานหนักขึ้น เตรียมตัวกว้างกว่า และเรียนหนักกว่า

มีผลงานวิจัยจำนวนมากที่เกิดขึ้นเพื่อศึกษาผลนี้ ในการทำวิจัยสังเคราะห์ของ ลันเดย์เบอร์ก และฟอกซ์ (Lundeborg and Fox. 1991 : 94 - 106) สรุปว่า ผลที่ได้ต่างกันออกไปแล้วแต่สภาวะ ในสภาวะการทดลองที่ใช้ความจำหรือการระลึกได้อย่างง่าย ๆ นักเรียนที่คาดหวังว่าจะได้รับการทดสอบแบบระลึกได้ จะทำได้ดีในงานทั้ง 2 ชนิด มากกว่าผู้ที่คาดว่าเป็นแบบทดสอบที่ใช้การจำได้ ซึ่งเป็นไปตามสมมุติฐานที่ตั้งไว้ อย่างไรก็ตาม ในชั้นเรียน ซึ่งใช้แบบทดสอบเลือกตอบ เต็มคำ และอัตนัย นักเรียนจะทำได้ดีกว่าในรูปแบบที่เขาเหล่านั้นคาดหวัง นอกจากนี้ ยังพบว่าความวิตกกังวลของผู้สอบจะเกิดขึ้นถ้ามีการเผชิญกับรูปแบบข้อสอบที่ไม่คาดคิดไว้ ซึ่งจะส่งผลต่อการตอบสนองของผู้สอบ

เวย์เนอร์ และทิสเซน (Bennett. 1991 : 15; citing Wainer and Thissen. 1991) ศึกษาและได้ตัวบ่งชี้บางประการเกี่ยวกับรูปแบบของข้อคำถาม โดยศึกษาเปรียบเทียบในแง่ต้นทุนการสร้าง การตรวจให้คะแนน เวลาที่ใช้ในการสอบ พบว่า ข้อคำถามแบบอัตนัยจะเสียค่าใช้จ่ายในการพัฒนาสูงกว่า

ชนิดของข้อคำถาม (Item Type)

ชนิดของข้อคำถามที่ใช้ในการสอบวัด จัดว่าเป็นองค์ประกอบสำคัญที่จะทำให้เครื่องมือวัดมีความเที่ยงตรง ให้ผลเป็นที่น่าเชื่อถือ ทั้งนี้เพราะชนิดของข้อคำถามถือเป็นงาน ซึ่งทำหน้าที่เป็นสิ่งเร้าให้ผู้ตอบใช้ความคิดเพื่อแก้ปัญหาแล้วตอบสนองออกมา สิ่งเร้าที่มีคุณภาพควรมีลักษณะที่ท้าทาย ทำให้ผู้ตอบหรือผู้เผชิญต้องใช้ความคิดที่หลายแง่มุมและลึกซึ้งไปกว่าการระลึกถึงข้อเท็จจริงเท่านั้น สามารถสะท้อนให้เห็นความสามารถในการคิดพิจารณาของบุคคลได้ ฉะนั้น จะเห็นว่าแบบทดสอบทั้งหลายที่ได้พัฒนาขึ้นมาโดยเฉพาะแบบทดสอบวัดด้านสติปัญญา แบบทดสอบความถนัด ซึ่งเป้าหมายของการวัดลึกซึ้งเกินกว่าความรู้ความเข้าใจเนื้อหาปกติ ได้มีการพัฒนารูปแบบชนิดของข้อคำถามขึ้นเพื่อใช้ในการวัดองค์ประกอบต่าง ๆ ที่ถือเป็นโครงสร้างสำคัญของคุณลักษณะที่ต้องการวัด

จากการศึกษาวิเคราะห์แบบทดสอบต่างประเทศ ซึ่งเป็นที่นิยมเลือกใช้กันอย่างกว้างขวาง อาทิ Stanford-Binet Scale, Wechsler Scales, Otis-Lennon School Ability Test, Cognitive Abilities Test โดยเฉพาะในส่วนของ การวัดความสามารถด้านเหตุผล พบว่า รูปแบบของการทดสอบมีทั้งในส่วนที่เป็นภาษา และรูปภาพ

ส่วนงานหรือชนิดของข้อคำถามที่ใช้ส่วนใหญ่จะเป็นการให้อุปมาอุปไมย สรุปความ จำแนกประเภท การเติมความให้สมบูรณ์ จัดลำดับ การพิจารณาหาสิ่งที่เกี่ยวข้องกัน หรือ สิ่งที่เขาเดาต่อไป

นูนนัลลี (Nunnally, 1972 : 315 - 317) เป็นผู้หนึ่งที่ศึกษาเกี่ยวกับองค์ประกอบของความสามารถ โดยทำการศึกษาจากการวิเคราะห์องค์ประกอบและกล่าวถึงองค์ประกอบด้านเหตุผล (Reasoning Factors) ว่าเป็นสิ่งที่มีความซับซ้อนที่ขาดต่อการจำแนกองค์ประกอบด้วยการวัดเพียงอย่างเดียวอย่างใดอย่างหนึ่ง หรือขาดต่อการค้นหาวิธีการวัดที่ดีมาใช้ และสิ่งนี้นูนนัลลีถือว่าเป็นตัวกำหนดองค์ประกอบด้านเหตุผล ได้แก่

1. เหตุผลทั่วไป (General Reasoning) เป็นองค์ประกอบด้านเหตุผลที่ธรรมดาที่สุด ซึ่งเกี่ยวข้องกับการค้นหาทางแก้ปัญหา ข้อคำถามที่ใช้วัดองค์ประกอบนี้มักเกี่ยวข้องกับการเติมอนุกรม (Serial Completion)

2. การอนุมาน (Deduction) องค์ประกอบด้านอนุมานจะเกี่ยวข้องกับการสร้างข้อสรุปในลักษณะการอ้างเหตุผลตามหลักตรรกวิทยา เน้นความสามารถที่เกี่ยวข้องกับการประเมินการนำมาซึ่งการอ้างเหตุผล ข้อคำถามที่ใช้มีลักษณะเป็นข้อคำถามเกี่ยวกับการอ้างเหตุผลตามหลักตรรกวิทยา (Logical Syllogism)

3. การมองเห็นความสัมพันธ์ (Seeing Relationships) เป็นความสามารถในการมองเห็นความสัมพันธ์ระหว่างของ 2 สิ่งหรือความคิด 2 ประการ และใช้ความสัมพันธ์ดังกล่าวในการค้นหาสิ่งอื่นหรือความคิดอื่น องค์ประกอบนี้สามารถแสดงได้ดีที่สุด โดยการอุปมาเชิงภาษา หรืออุปมาเชิงเค้าโครง (Design Analogies)

ข้อคำถามบางชุดที่วัดความสามารถด้านเหตุผลจะรวมองค์ประกอบ 1 และ 3 เข้าด้วยกันโดยใช้การเติมอนุกรม

ดังนั้น จะเห็นว่าแนวความคิดของนูนนัลลีเกี่ยวกับโครงสร้างของความสามารถด้านเหตุผลและรูปแบบการสอบวัดก็มีความคล้ายคลึงกัน

สำหรับประเทศไทยนั้น สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ซึ่งถือเป็นสถาบันที่พัฒนาแบบทดสอบมาตรฐานหลายประเภทขึ้น เพื่อใช้ในกิจกรรมการสอบต่าง ๆ จนเป็นที่ยอมรับทั่วไป ก็ใช้ข้อคำถามชนิดอุปมาอุปไมย สรุปความ และจำแนกประเภท สำหรับวัดความสามารถด้านเหตุผล หรือความสามารถในการคิดอย่างมีวิจารณญาณ เช่นกัน

ฉะนั้น อาจจะสามารถได้ว่าชนิดของข้อคำถามที่ใช้วัดความสามารถด้านเหตุผลอย่างแพร่หลายในปัจจุบัน ก็คือ ชนิดอุปมาอุปไมย จำนวนประเภท แมสสเตอร์นเบิร์ก (Sternberg. 1985 : 78) ก็เชื่อว่าความสามารถที่เป็น Fluid Type จะวัดได้ดีโดยข้อคำถามด้านเหตุผล เช่น อุปมาอุปไมย (Analogies) และอนุกรม (Series Completion) นอกจากนี้ ทฤษฎีการวัดการจิตวิทยาส่วนใหญ่ จะมุ่งเน้นไปที่อุปมาอุปไมยในฐานะเป็นพื้นฐานของการประเมินตามทฤษฎีของเขา

ด้วยความเชื่อพื้นฐานดังกล่าว ทำให้แบบทดสอบมาตรฐานส่วนใหญ่จะใช้ข้อคำถามชนิดอุปมาอุปไมยวัดความสามารถด้านเหตุผล

อนึ่ง มีผลงานการศึกษาเกี่ยวกับความสามารถด้านเหตุผล ซึ่งดำเนินการโดยคณะกรรมการการศึกษาาระดับอุดมศึกษาของนิวเจอร์ซีย์ ได้จัดทำโครงการประเมินทักษะพื้นฐานของนักเรียนที่จะเข้าสู่วิทยาระดับวิทยาลัยขึ้นในปี ค.ศ. 1977 (Morante and Ulesky. 1984 : 71 - 74) โครงการดังกล่าวมีจุดประสงค์ 2 ประการ คือ เพื่อการประเมินทักษะพื้นฐานที่จำเป็นของผู้เข้าสู่วิทยาลัยของรัฐขึ้นปีที่ 1 ทั้งหมด และเพื่อประเมินประสิทธิภาพและลักษณะของโครงการสอนซ่อมเสริมของวิทยาลัยเหล่านั้น ในระยะแรกมีความพยายาามที่จะวัดความคิดอย่างมีวิจารณญาณโดยใช้ภาษาและคณิตศาสตร์ ซึ่งต้องให้ความเข้าใจและการคิด แต่พบว่ามึจุดอ่อนบางประการจึงล้มเลิกการวัดบางส่วนไป

ในปี ค.ศ. 1982 ได้มีการดำเนินการใหม่ ทั้งนี้โดยคณะกรรมการได้กำหนดการทำงานใหม่เป็น 3 ระยะ คือ ระยะแรก นิยามสมรรถภาพการคิดของนักศึกษาที่เข้าวิทยาลัยในปีแรกว่า ควรจะมีอะไรบ้าง ระยะที่ 2 สํารวจเครื่องมือวัดสมรรถภาพด้านการคิด ระยะที่ 3 ให้ข้อเสนอแนะเพื่อการพัฒนาทักษะการคิดและเหตุผล

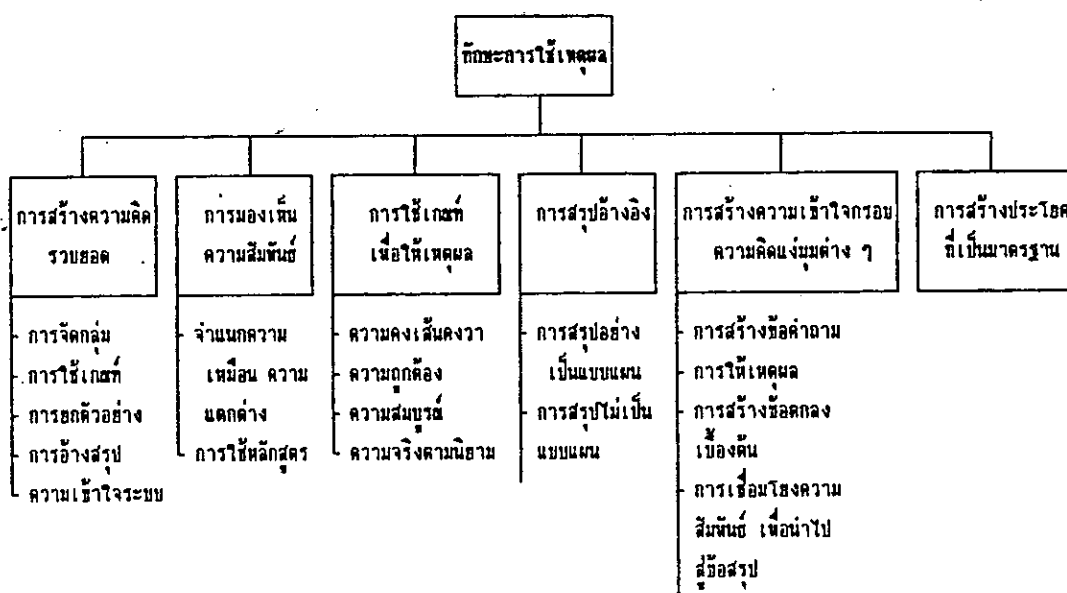
สำหรับการนิยามและการอธิบายสมรรถภาพการคิดนั้น เสนอว่าผู้มีการคิดอย่างมีวิจารณญาณ (Critical Thinking) จะต้องมีความสมรรถภาพต่อไปนี้

1. ความสามารถในการระบุและกำหนดปัญหา ความสามารถในการเสนอและประเมินทางแก้ปัญหา
2. ความสามารถในการระบุ และใช้เหตุผลเชิงอุปมานและอนุมาน และระบุการอ้างเหตุผลที่ไม่ถูกต้องได้
3. ความสามารถในการสร้างข้อสรุปที่สมเหตุสมผลจากข้อมูลข้อสนเทศ ที่พบในแหล่งต่าง ๆ อาจเป็นการเขียน การพูด ตาราง แผนภาพ เพื่อปกป้องข้อสรุปของตนอย่างมีเหตุผล

4. ความสามารถที่จะทำความเข้าใจ พัฒนา และใช้ความคิดรวบยอดและอ้างสรุป

5. ความสามารถในการจำแนกความแตกต่างระหว่างความจริง และความคิดเห็น
อย่างไรก็ตาม ยังถือว่าสมรรถภาพเหล่านี้ยังไม่สามารถอธิบายธรรมชาติของการคิด
อย่างมีเหตุผลได้อย่างสมบูรณ์ ต่อมา ลิปแมน จาคอบ และโคลแมน ได้จัดทำ Taxonomy
of Thinking Skills ขึ้น ซึ่งประกอบด้วยลักษณะของการคิด คือ กิจกรรมทางสมอง
(Mental Acts) สภาวะการคิด (Cognitive States) การรวมกันของกิจกรรมทาง
สมองและสภาวะ

การคิด ทักษะการใช้เหตุผล (Reasoning Skills) และทักษะการค้นคว้า
(Inquiry Skills) ลักษณะการคิดเหล่านี้ได้ถูกจำแนกเป็นทักษะย่อย ๆ สำหรับทักษะการ
ใช้เหตุผลนั้น จำแนกเป็นทักษะย่อยดังภาพประกอบ 5



ภาพประกอบ 5 Taxonomy of Thinking Skills (เฉพาะในส่วนของ Reasoning Skills)

ระยะที่ 2 ซึ่งเป็นการสำรวจแบบทดสอบหรือเครื่องมือวัดสมรรถภาพการคิดนั้น ได้มีการรวบรวมรายชื่อแบบทดสอบ ซึ่งถูกพัฒนาขึ้นเพื่อวัดการคิด การใช้ หรือเหตุผล เพื่อนำมาพิจารณาศึกษาเปรียบเทียบ และเลือกใช้เพื่อการศึกษา พัฒนาการคิด และการใช้เหตุผลต่อไป แบบทดสอบซึ่งใช้กันอย่างกว้างขวางเป็นที่รู้จัก ได้แก่

1. New Jersey Test of Reasoning
2. Whimbey Analytical Skills Inventory
3. Cornell Critical Thinking Test Level X
4. Cognitive Abilities Test Form 3
5. Watson-Glaser Forms A และ B
6. Ross Test of Higher Cognitive Processes

ทำการศึกษาเปรียบเทียบแบบทดสอบเหล่านี้ โดยพิจารณาตรวจสอบข้อคำถามในแง่ประสิทธิภาพและการจำแนกข้อคำถามตามประเภท และความสัมพันธ์ระหว่างข้อคำถามวัดการคิดและทักษะพื้นฐาน

ผลจากการศึกษาพบว่า นักศึกษาชั้นปีที่ 1 ดังกล่าว มีระดับการคิดอย่างมีเหตุผลต่ำกว่าระดับมาตรฐาน ส่วนการศึกษาเครื่องมือพบว่า ข้อคำถามยังไม่ส่งเสริมการสร้างสรรค์ ข้อคำถามบางส่วนคลุมเครือ บางส่วนไม่สัมพันธ์กับคะแนนรวมทั้งหมด และบางส่วนไม่สามารถจำแนกระหว่างผู้มีเหตุผลที่ดีและไม่ดีได้ บางส่วนพบว่ามีความสัมพันธ์กับคณิตศาสตร์และคำสั่งที่มากกว่าเหตุผล และพบว่ามีความสัมพันธ์เชิงบวกระหว่างแบบทดสอบการคิดแต่ละฉบับ (ฉบับที่ 1, 2 และ 3) กับทักษะพื้นฐาน 5 ประการ ดังกล่าวมา

ศูนย์บริการการทดสอบทางการศึกษา (Educational Testing Service : ETS) ของสหรัฐอเมริกาได้ทำการพัฒนาชนิดของข้อคำถาม โดยยึดองค์ประกอบที่ได้จากการวิเคราะห์ องค์ประกอบเป็นพื้นฐานในการพิจารณา สำหรับความสามารถด้านเหตุผลนั้น พบว่ามีองค์ประกอบ 2 องค์ประกอบ ที่แสดงมิติแยกจากกัน คือ Informal Reasoning หรือ Everyday Reasoning และ Formal Deductive Reasoning ทั้งนี้ แกลอตต์ (Enright, Tucker and Katz. 1995 : 3; citing Galotti. 1989) ได้นิยามลักษณะสำคัญของ Formal Reasoning ว่า สารสนเทศทั้งหมดที่จะถูกพิจารณาจะปรากฏอย่างชัดเจนในปัญหา ในทางตรงกันข้าม Informal Reasoning หรือ Everyday Reasoning ต้องการการค้นหาข้อมูลที่เกี่ยวข้อง หรือ ตัดสินใจว่าข้อมูลอะไรที่เกี่ยวข้องกับปัญหาที่กำลังพิจารณา แต่มีกลุ่มหนึ่งที่มีความเห็นต่างไป (Voss and others. 1989) คือ

เชื่อว่าทั้ง Formal และ Informal Reasoning มุ่งเน้นที่การประเมินการอ้างเหตุผล ใน Formal Reasoning กระบวนการมักรวมการแปลงข้อความหรือปัญหา (Proposition) ไปเป็นสัญลักษณ์ การรวมข้อความหรือปัญหาเหล่านี้ก็เพื่อพิจารณาสารสนเทศใหม่ จากหลักการทั่วไป เรื่องเฉพาะ และตัดสินใจว่าความสัมพันธ์ของสัญลักษณ์สอดคล้องไปตามกฎของระบบ เหตุผลชนิดนี้จะถูกนำไปประยุกต์ใช้ใน Formal Deductive System เช่น ตรรกวิทยา และคณิตศาสตร์

ส่วน Informal Arguments ประกอบด้วยข้อสรุปหรือสมมติฐานที่ได้รับการสนับสนุนโดยเหตุผล และถูกประเมินในรูปของความมั่นคงของมัน "Informal" ในที่นี้มีได้หมายความว่าความถึงความไม่รอบคอบในการให้เหตุผล แต่ Informal Reasoning จะมีเกณฑ์ที่ต่างไป ซึ่งจะเกี่ยวข้องกับความเที่ยงตรงและคงเส้นคงวามากกว่า ความสอดคล้องและเข้ากันได้กับองค์ความรู้พื้นฐาน

นักจิตวิทยาหลายคนที่ได้ศึกษาเกี่ยวกับเหตุผล มีจุดเน้นไปที่ Formal Deductive Problems เช่น การอ้างเหตุผลในเชิงตรรก (Syllogisms) ก่อน สำระสำคัญอยู่ที่บุคคลให้เหตุผลโดยมีพื้นฐานอยู่ที่ Formal Logical System หรือไม่ ในมุมมองหนึ่งเห็นว่า การให้เหตุผลสามารถเข้าใจได้ในรูปของชนิดของปฏิบัติการประมวลข้อมูลที่ใช้อธิบายกิจกรรมทางสมองรูปแบบอื่น

ในงานวิจัยเกี่ยวกับการแก้ปัญหาเริ่มแรกจะมุ่งไปที่ Well-Structured Problems หรือปัญหาที่อิงความรู้เป็นหลัก (Knowledge Lean Problems) ใน Well Structured Problems ทางแก้ปัญหาถูกรอบงำโดยระบบของเหตุผลหรือตรรกวิทยา ดังนั้น ความถูกต้องของคำตอบสามารถแสดงได้อย่างชัดเจน บทบาทขององค์ประกอบความหมายทางภาษาจะลดน้อยลง ผู้แก้ปัญหาจะจัดกระทำกับวัตถุ สัญลักษณ์หรือเครื่องหมายที่เป็นนามธรรมไปยังระดับใดระดับหนึ่ง จากบริบทของความหมายทางภาษา ข้อคำถามชนิดเหตุผลเชิงวิเคราะห์ (Analytical Reasoning Item Type) ซึ่งถูกนำมาใช้วัดการวิเคราะห์ใน GRE เป็นตัวอย่างของปัญหาที่มีลักษณะเป็น Well Structured หรือ Knowledge-Lean Problems งานวิจัยเกี่ยวกับปัญหารูปแบบที่มักเกี่ยวกับการอธิบายว่า ผู้เชี่ยวชาญและผู้ไม่เคยมีประสบการณ์มีความแตกต่างกันอย่างไรในวิธีการแก้ปัญหา (Enright, Tucker and Katz. 1995 : 5)

ดังได้กล่าวแล้วว่า ETS ได้มีการพัฒนาชนิดของข้อคำถามวัดความสามารถด้านเหตุผลรูปแบบใหม่ขึ้น ทดลองใช้ในแบบทดสอบ GRE และจากการศึกษาพบว่า ข้อคำถามชนิด

ที่พัฒนาขึ้นใหม่นี้ ทำให้ความเที่ยงตรงเชิงเหมือนของการวัดมีความแข็งแกร่งขึ้น (Emmerich and others : 1991) ชนิดของข้อคำถามที่พัฒนาเหล่านี้ ได้แก่

1. การวิเคราะห์ค่าอธิบาย
2. ความเห็นตรงกันข้าม
3. เหตุผลเชิงตรรกเกี่ยวกับจำนวน
4. การระบุรูปแบบ
5. เหตุผลทางตรรก
6. เหตุผลด้านวิเคราะห์

ข้อคำถามแต่ละชนิด มีคุณลักษณะเฉพาะที่ผู้ใช้สามารถนำไปพิจารณาเลือกใช้ได้อย่างเหมาะสม ทั้งนี้โดยพิจารณาเป้าหมายที่ต้องการวัด สถานการณ์ในการวัด ความเที่ยงตรง และเชื่อมั่นที่ต้องการ และปัจจัยอื่น ๆ อีกหลายประการ

จำนวนตัวเลือกที่เหมาะสมสำหรับข้อคำถามรูปแบบเลือกตอบ เป็นอีกประเด็นที่นักวัดผลหลายท่านพยายามหาคำตอบ การเขียนตัวเลือกสำหรับข้อคำถาม มักถูกชี้แนะโดยประสบการณ์สามัญสำนึก และจากคำบอกเล่าจากที่ปรึกษาหรือตำรา จากการวิเคราะห์ตำราประมาณ 46 เล่ม ที่กล่าวครอบคลุมเกี่ยวกับการเขียนข้อคำถามสำหรับแบบทดสอบชนิดเลือกตอบ ฮาลาดายนา และดาวน์นิ่ง (Haladyna and Downing, 1989 : 37 - 50) รายงานว่า ผู้เขียนมักมีความเห็นไม่สอดคล้องกันเกี่ยวกับจำนวนตัวเลือกที่เป็นอุดมคติ บางคนให้ความเห็นว่าควรเป็น 4 หรือ 5 ตัวเลือก ขณะที่บางคนเห็นว่าการเขียนตัวเลือกที่มากสะดวกหรือเป็นไปได้ ปกติมักจำกัดอยู่ที่ 5 ตัวเลือก แต่ชนิด 3 ตัวเลือก มักไม่ค่อยมีใครกล่าวถึง สำหรับโครงการทดสอบแม้จะมีหลักฐานแสดงให้เห็นว่ารูปแบบ 3 ตัวเลือก เป็นสิ่งที่ต้องการ (Owen and Forman, 1987 : 513 - 522) ข้อดีในทางปฏิบัติของการใช้ตัวลอง 2 ตัว มีหลายประการ คือ

1. ใช้เวลาในการสร้างข้อคำถามน้อยลง
2. ความยาวฉบับข้อสอบจะลดลง
3. ค่าจัดพิมพ์ลดลง
4. เวลาที่ใช้ในการดำเนินการสอบลดลง
5. ข้อคำถามที่จะใช้ดำเนินการสอบจะได้มากขึ้น ณ ช่วงเวลาที่กำหนดให้ ฉะนั้น

ทำให้เป็นการเพิ่มตัวอย่างของเนื้อหาและความเชื่อมั่นของคะแนนได้

ในปี ค.ศ. 1993 ฮาลาดายนา และดาวน์นิง (Haladyna and Downing. 1993 : 999 - 1009) ได้ทำการศึกษาเกี่ยวกับตัวเลือกที่มีความพอดีสำหรับแบบทดสอบเลือกตอบ ขึ้น โดยมีจุดประสงค์หลัก 2 ประการ คือ ประการแรก เพื่อสรุปหลักฐานข้อมูลเชิงทฤษฎีและเชิงประจักษ์เกี่ยวกับความต้องการใช้แบบทดสอบ 3 ตัวเลือก ประการที่สอง เพื่อรายงานผลของการศึกษาเชิงพรรณนาเกี่ยวกับผลการใช้ตัวเลือกในโครงการสอบที่ต่างกัน 3 โครงการ

จากมุมมองเชิงทฤษฎี (Theoretical Perspective) ลอร์ด (Lord. 1977) ได้เป็นผู้ศึกษาเกี่ยวกับจำนวนตัวเลือกที่เหมาะสมที่สุด โดยได้พัฒนาสูตรในการพยากรณ์การเปลี่ยนแปลงค่าความเชื่อมั่น ซึ่งเป็นไปตามจำนวนของตัวเลือกที่เพิ่มเข้าไปในข้อคำถามแบบเลือกตอบ ผลการศึกษาชี้ให้เห็นว่า ข้อคำถามชนิด 3 ตัวเลือก จะดีที่สุดสำหรับผู้สอบทเวอร์สกี (Tversky. 1964 : 386 - 391) ก็พบข้อสรุปที่คล้ายกัน ทั้งนี้เขาวิเคราะห์โดยใช้เกณฑ์ 3 เกณฑ์ คือ ความสามารถในการจำแนก (Discriminability) ความมีอำนาจ (Power) และสารสนเทศของแบบทดสอบ การศึกษาโดยอีเบล (Ebel. 1969 : 565 - 570) ไกรเลอร์ (Grier. 1975 : 109 - 112) และจากลอร์ด (Lord. 1977 : 32 - 38) ซึ่งใช้ทฤษฎีการตอบสนองข้อสอบ (IRT) ก็สนับสนุนข้อสรุปนี้ การศึกษาของลอร์ดให้ข้อมูลมากที่สุดเกี่ยวกับผลสัมฤทธิ์ของมาตรวัดว่าจะทำงานได้ดีที่สุด ณ จุดใด ด้วยการใช้โด่งสารสนเทศของข้อคำถาม ลอร์ดได้แสดงให้เห็นว่า ข้อคำถามชนิด 3 ตัวเลือกให้สารสนเทศมากที่สุด ณ ช่วงกลางของ Score Scale ขณะที่ 2 ตัวเลือก ทำงานได้ดีที่สุด ณ ช่วงที่สูงกว่า Score Scale ส่วน 4 หรือ 5 ตัวเลือก จะทำงานได้ดีที่สุดที่ช่วงต่ำลงมาจากการศึกษาของดาวน์นิง (Downing. 1993) ชี้ให้เห็นว่า ข้อคำถามชนิด 2 ตัวเลือก มีข้อดีหลายประการเช่นเดียวกับที่พบในข้อคำถามชนิด 3 ตัวเลือก

สำหรับ บูดেসคู และนีโว (Budesu and Nevo. 1985 : 183 - 196) ศึกษาจำนวนตัวเลือกที่เหมาะสมเช่นกัน โดยมีมุ่งเพื่อหาข้อสรุปเกี่ยวกับสัดส่วนของเวลาในการทำข้อสอบกับจำนวนตัวเลือกของข้อสอบ ผลการศึกษพบว่า แบบทดสอบที่มีจำนวนตัวเลือกมากจะให้ค่าความแปรปรวน ค่าความเชื่อมั่น และค่าอำนาจจำแนก สูงกว่าแบบทดสอบที่มีจำนวนตัวเลือกน้อย เมื่อหาเวลาเฉลี่ยที่ใช้ตอบต่อตัวเลือกพบว่า ส่วนใหญ่เป็นไปตามข้อสันนิษฐานของลอร์ด (Lord. 1977 : 33 - 38) ที่ว่า เวลาทั้งหมดในการทำแบบทดสอบจะเท่ากับจำนวนข้อสอบคูณด้วยจำนวนตัวเลือกในแต่ละข้อและเวลาที่ใช้ตอบต่อตัวเลือก ฉะนั้น เมื่อใช้ตัวเลือกต่อข้อคำถามมากกว่าเวลาที่ใช้ในการดำเนินการสอบจะมากกว่า และก็ไม่มีใครยัง

อีกเช่นกันว่า การเขียนตัวลงต้องใช้ใช้เวลา รวมทั้งการเขียนตัวลงเป็นสิ่งที่ยากที่สุดในการเขียนข้อความแบบเลือกตอบ

ในแง่การศึกษาทดลอง (Empirical Perspective)

โอเวน และโฟรมาน (Owen and Froman. 1987) ศึกษาผลของจำนวนของตัวเลือกในลักษณะง่าย ๆ คือ การลดข้อความที่มี 5 ตัวเลือก เป็น 3 ตัวเลือก โดยใช้วิธีการวิเคราะห์ข้อความ และการจำแนกตัวเลือกเป็นเกณฑ์ ซึ่งพบว่าไม่มีความแตกต่างที่ชัดเจนในเรื่องของความยากและอำนาจจำแนก ระหว่างข้อความแบบ 3 ตัวเลือก และ 5 ตัวเลือก แต่ที่ปรากฏความแตกต่างชัดเจน คือ เวลาในการดำเนินการสอบ ซึ่งสำหรับข้อความชนิด 3 ตัวเลือก เป็นที่น่าพอใจมากกว่า และการศึกษาชี้ให้เห็นคุณค่าของการประเมินผลการใช้ตัวลง และการนำตัวลงที่ใช้ประโยชน์ไม่ได้ออกไป ว่าเป็นการลดเวลาในการทดสอบที่ปราศจากการสูญเสียความเชื่อมั่น

เทรวิซัน แซค และมิเชล (Trevisan, Sax and Michael. 1991 : 829 - 837) ได้ศึกษาเรื่องนี้เช่นกันโดยสร้างข้อความชนิด 3 และ 4 ตัวเลือก จากข้อความชนิด 5 ตัวเลือก และได้สร้างข้อความชนิด 2 ตัวเลือก และเพิ่มตัวเลือกไปอย่างเป็นระบบ ผลการศึกษาพบว่า แตกต่างกันอย่างไม่มีนัยสำคัญ และการลดจำนวนตัวเลือกไม่มีต่อผลความเชื่อมั่น (Internal Consistency Reliability) ของแบบทดสอบ

สำหรับการศึกษาที่ ฮาลาไดนา และคาวนิง ซึ่งทำการศึกษาโดยใช้แบบทดสอบมาตรฐาน ชนิดเลือกตอบ 3 ฉบับ ซึ่งมีความต่างกันในการใช้และการพัฒนา ฉบับแรกเป็นแบบทดสอบที่ใช้สำหรับโครงการศึกษาแพทย์ โดยจะใช้ทดสอบแพทย์ทางกายภาพประจำปี แบบทดสอบจะให้ข้อมูลเกี่ยวกับการผ่าตัดเฉพาะทาง แบบทดสอบมี 5 ตัวเลือก ฉบับที่สองเป็นแบบทดสอบที่ใช้รับนักศึกษาเข้าเรียน ใช้เพื่อพยากรณ์ระดับผลการเรียนของนักศึกษา ฉะนั้นแบบทดสอบจะเกี่ยวข้องกับความเสี่ยงตรงเชิงพยากรณ์มาก เป็นแบบทดสอบชนิด 4 ตัวเลือก ฉบับที่ 3 เป็นแบบทดสอบเพื่อการรับรองในเรื่องของวิทยาศาสตร์สุขภาพ แต่การดำเนินการสร้างไม่มีมาตรฐานสูงมากเหมือน 2 ฉบับแรก ทำให้ความเสี่ยงตรงค่อนข้างต่ำ หลังจากนำผลการทดลองมาศึกษาในเชิงสถิติเกี่ยวกับการทำหน้าที่ของตัวลง ตัวลงที่มีความถี่ในการถูกเลือกต่ำจะถูกตัดออก สำหรับตัวลงที่เหลืออยู่จะใช้เพื่อทดสอบสมมติฐานเกี่ยวกับการเพิ่มขึ้นหรือลดลงอย่าง Monotonic หรือ Non-Monotonic ผลสรุปการศึกษาครั้งนี้แสดงให้เห็นว่า จำนวนตัวเลือกที่มีประสิทธิภาพของแต่ละข้อความมักใกล้เคียงกับ 1 แต่มีบางข้อที่

มีตัวลวงที่มีประสิทธิภาพมากกว่า 1 ซึ่งจะให้ค่าอำนาจจำแนกสูง ข้อคำถามที่มีตัวลวงซึ่งมีประสิทธิภาพมากกว่า 2 หรือ 3 มีน้อยมาก พบเพียงประมาณ 1.1% ถึง 8.4% จากการสอบแต่ละครั้ง แม้คำถามแต่ละข้อจะต้องการตัวลวงที่มีประสิทธิภาพมากกว่านี้

ดังนั้น อาจกล่าวได้ว่าแม้จะมีการศึกษาวิจัยกันอย่างกว้างขวางเพื่อให้ได้ข้อสรุปว่าจำนวนตัวเลือกเท่าไรจึงจะดีที่สุด แต่เนื่องจากผลการศึกษายังมีความไม่สอดคล้องกันอยู่ ฉะนั้น จึงยังไม่อาจสรุปได้ อย่างไรก็ตาม ครั้งหนึ่งก็ได้มีความพยายามที่จะทำให้ค่าความเชื่อมั่นของคะแนนผลการสอบสูงขึ้น ค่าอำนาจจำแนกสูงขึ้น โดยเพิ่มจำนวนตัวลวงในข้อคำถามแต่ละข้อให้มากขึ้น

2.3 คุณภาพและการตรวจสอบคุณภาพของเครื่องมือ

2.3.1 ความเที่ยงตรง (Validity)

แบบทดสอบและเครื่องมือประเมินอื่น ๆ ถูกนำมาใช้ในด้านต่าง ๆ มากมาย ในโรงเรียน เช่น แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนอาจถูกใช้ในการคัดเลือก การจัดตำแหน่ง การวินิจฉัย หรือรับรองการเรียนรู้ แบบทดสอบความถนัดอาจถูกนำมาใช้สำหรับพยากรณ์ความสำเร็จทางการเรียนในอนาคตหรืออาชีพในอนาคต และการประเมินเรื่องส่วนบุคคล เช่น พัฒนาการทางสังคม อาจใช้เพื่อความเข้าใจปัญหาทางการเรียนให้ดีขึ้น หรือเพื่อประเมินผลของโครงการเฉพาะของโรงเรียน เครื่องมือที่ใช้ในการประเมินเพื่อกิจกรรมต่าง ๆ ของโรงเรียนเหล่านี้ จำเป็นต้องมีคุณลักษณะเฉพาะสำคัญ ซึ่งได้แก่ ความเชื่อมั่น (Reliability) ความเที่ยงตรง (Validity) และความมีประโยชน์ (Usability) (Gronlund and Linn. 1990 : 47) ความเที่ยงตรงถือเป็นคุณสมบัติสำคัญที่สุดที่ควรได้รับการพิจารณาในการสร้าง หรือเลือกใช้แบบทดสอบหรือเครื่องมืออื่น ๆ

ความเที่ยงตรง หมายถึง ความเหมาะสมของการตีความหมายอธิบายคะแนน ผลการสอบและผลการประเมินอื่น ๆ ตามเป้าหมายเฉพาะของการใช้ เช่น ต้องการให้ผลเพื่อการพยากรณ์ความสำเร็จในกิจกรรมบางอย่างของผู้เรียนในอนาคต จำเป็นต้องตีความอธิบายผล โดยยึดความถูกต้องแม่นยำในการประมาณความสำเร็จในอนาคตให้มากที่สุด หรือหากต้องการให้ผลเป็นสิ่งที่วัดความเข้าใจในการอ่าน จำเป็นต้องตีความโดยยึดหลักฐาน ซึ่งคือคะแนนที่สะท้อนความเข้าใจในการอ่านได้อย่างแท้จริงและไม่ถูกบิดเบือน โดยองค์ประกอบ

ที่ไม่เกี่ยวข้อง ในการใช้ข้อมูลเกี่ยวกับความเที่ยงตรงจำเป็นต้องระลึกเสมอว่า ความเที่ยงตรงเกี่ยวข้องกับการแปลความหมายคะแนนผลการสอบ แต่มิใช่ตัวแบบทดสอบ

ธรรมชาติของความเที่ยงตรง (Nature of Validity)

เมื่อใช้คำว่า "ความเที่ยงตรง" (Validity) ในเชิงความสัมพันธ์กับการทดสอบและการประเมิน มีข้อควรระวังที่ต้องคำนึงถึงหลายประการ คือ

1. ความเที่ยงตรง หมายถึง ความเหมาะสมของการแปลความหมายของผลการสอบหรือเครื่องมือประเมินสำหรับกลุ่มบุคคลที่ให้อาจจะกล่าวได้ว่า ความเที่ยงตรงของการตีความกระทำได้จากผลการวัด

2. ความเที่ยงตรงเป็นเรื่องของระดับ ซึ่งไม่ได้อยู่ในลักษณะของการมีหรือไม่มีหรือเที่ยงตรงหรือไม่เที่ยงตรง ความเที่ยงตรงจะได้รับการพิจารณาอย่างดีที่สุดในรูปแบบของการจัดกลุ่มที่ระดับ เช่น ความเที่ยงตรงสูง ความเที่ยงตรงปานกลาง ความเที่ยงตรงต่ำ

3. ความเที่ยงตรงมักจะเกิดขึ้นกับการใช้หรือการแปลความหมายเฉพาะ ไม่มีแบบทดสอบใดที่จะเที่ยงตรงสำหรับทุกเป้าหมายประสงค์ เช่น ผลจากการทดสอบคณิตศาสตร์อาจมีระดับความเที่ยงตรงสูงสำหรับบ่งชี้ทักษะการคำนวณ แต่มีระดับความเที่ยงตรงต่ำสำหรับบ่งชี้เหตุผลทางคณิตศาสตร์ (Arithmetical Reasoning) มีความเที่ยงตรงระดับปานกลาง สำหรับการพยากรณ์ความสำเร็จในรายวิชาคณิตศาสตร์ในอนาคต

4. ความเที่ยงตรงมีแนวคิดเดี่ยว (Unitary Concept) ธรรมชาติของความคิดรวบยอดของความเที่ยงตรง ปกติจะถูกอธิบายเพื่อการยอมรับในรูปแบบของมาตรฐาน (Standards) ที่ถูกเตรียมหรือสร้างขึ้น โดยคณะกรรมการร่วมจากองค์กรวิชาชีพที่เกี่ยวข้องกับการทดสอบทางการศึกษาและจิตวิทยา ความเที่ยงตรงจึงประกอบด้วย ความคิดที่ยึดชนิดของหลักฐานต่าง ๆ มากมาย และมีวิธีการหลายวิธีที่จะรวบรวมหลักฐานเพื่อสนับสนุนความเที่ยงตรงของการแปลความหมายคะแนนผลการสอบ

วิธีการตรวจสอบความถูกต้องของแบบทดสอบ

วิธีการเบื้องต้นในการตรวจสอบความเที่ยงตรงของแบบทดสอบ การตีความหมายของคะแนนผลการสอบจะมีความเที่ยงตรงมากขึ้น เมื่อความเข้าใจเกี่ยวกับสิ่งต่อไปนี้ คือ ที่มาของเนื้อหาของแบบทดสอบและลักษณะเฉพาะ ความสัมพันธ์ของคะแนนผลการสอบกับการวัดที่สำคัญอื่น ๆ และธรรมชาติของสิ่งที่กำลังถูกวัด กรอนลันด์ และลินน์ (Gronlund and Linn.

1990 : 51) ได้นำเสนอวิธีการตรวจสอบความถูกต้องของแบบทดสอบ (Approaches to Test Validation) ดังปรากฏในภาพประกอบ 6

วิธีการตรวจสอบ	วิธีการ	ความหมาย
Content-Related Evidence	เปรียบเทียบงานที่กำหนดในแบบทดสอบกับลักษณะเฉพาะของแบบทดสอบที่อธิบายขอบเขตของงานที่ได้รับการพิจารณา	ตัวอย่างของงานในแบบทดสอบเป็นตัวแทน หรือแสดงขอบเขตของงานที่ถูกวัดได้ดีเพียงใด
Criterion-Related Evidence	เปรียบเทียบคะแนนผลการสอบกับผลการวัดอื่นที่ได้ในเวลาต่อมา (สำหรับพยากรณ์) หรือกับผลการวัดอื่นที่ได้ในปัจจุบัน (เพื่อคาดคะเนสถานการณ์ในปัจจุบัน)	ผลที่ได้จากการสอบพยากรณ์ผลในอนาคตหรือคาดคะเน การปฏิบัติในปัจจุบันจากการวัดอื่นได้ดีเพียงใด
Construct-Related Evidence	สร้างความหมายของคะแนนจากแบบทดสอบโดยการควบคุม (หรือตรวจสอบ) พัฒนาการของแบบทดสอบ การประเมินความสัมพันธ์ของคะแนนกับการวัดอื่นที่เกี่ยวข้อง และทดลองเพื่อกำหนดว่าองค์ประกอบใดมีอิทธิพลต่อผลการสอบ	ผลการสอบสามารถอธิบายการวัดลักษณะ หรือคุณภาพบางประการได้ดีเพียงไร

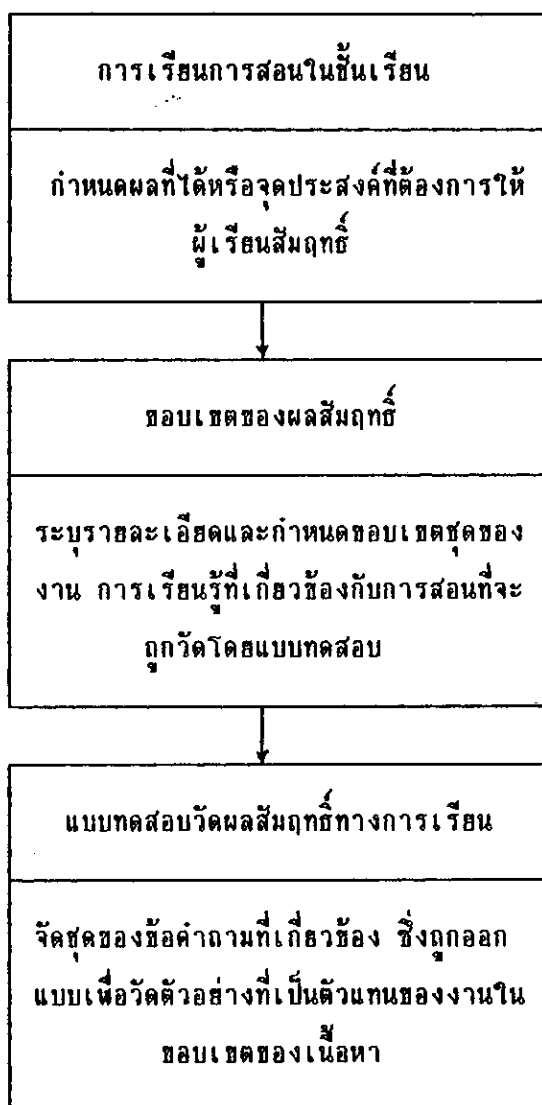
ภาพประกอบ 6 วิธีการตรวจสอบความถูกต้องของแบบทดสอบ

แม้ว่าการตรวจสอบความถูกต้องของแบบทดสอบ จะใช้หลักฐานข้อมูลที่เฉพาะตามเป้าหมายที่ต้องการตรวจสอบก็ตาม แต่ในทางปฏิบัติอาจจำเป็นต้องใช้หลักฐานข้อมูลประเภทอื่น หรือวิธีการอื่นประกอบเพื่อการอธิบายได้อย่างสมบูรณ์ ดังนั้น จะเห็นว่าในการใช้แบบทดสอบความถนัดทางการเรียน (Scholastic Aptitude Test) เพื่อพยากรณ์ความสำเร็จในโรงเรียนในอนาคต Criterion-Related Evidence จะเป็นวิธีการที่เป็นจุดสนใจหลัก แต่อาจจะต้องคำนึงถึงความเหมาะสมของเนื้อหาและองค์ประกอบที่ไม่สอดคล้อง ซึ่งอาจมีอิทธิพลต่อผลการสอบด้วย ฉะนั้น อาจจำเป็นต้องใช้ทั้ง Content และ Construct Related Evidences เพื่อความเข้าใจที่สมบูรณ์ของความหมายของคะแนนการทดสอบ อันจะนำไปสู่การตรวจสอบความถูกต้องในการอธิบายต่อไป

Content-Related Evidence

Content-Related Evidence มีความสำคัญในการสนับสนุนการแปลความหมายการอธิบายคะแนนผลการสอบหรือผลการวัดต่าง ๆ ฉะนั้น สิ่งที่จะนำมาเป็นข้อมูลหลักฐานจะต้องเป็นตัวแทนของกลุ่มตัวอย่าง (Representative Sample) ของงานที่กำหนด

การตรวจสอบความถูกต้องของเนื้อหา (Content Validation) คือ กระบวนการกำหนดความเหมาะสมเพียงพอของการสุ่มตัวอย่างของเนื้อหาที่แสดงความเป็นตัวแทนหรือความสอดคล้องกับขอบเขตของงาน ซึ่งคะแนนที่ได้ถูกนำมาใช้ในการแปลความหมายในสภาพการสอบในชั้นเรียน ขอบเขตของงานผลสัมฤทธิ์ถูกกำหนดโดยการเรียนการสอน การพัฒนาแบบทดสอบ จะรวมถึงการกำหนดขอบเขตของงานการสอนที่สอดคล้องซึ่งจะถูกวัด ดังนั้น การได้มาซึ่งการวัดผลที่ได้จากการเรียนรู้ที่ถูกต้องจะต้องได้มาจากการสอนที่มีมาก่อนแล้ว สิ่งที่จะถูกวัดและตัวอย่างที่เป็นตัวแทนของงานที่เกี่ยวข้อง ดังภาพประกอบ 7



ภาพประกอบ 7 แนวทางการวัด

หลักฐานข้อมูลที่ได้จากการตรวจสอบความเที่ยงตรงของเนื้อหา (Content Validation) กับความเที่ยงตรงเชิงพินิจ (Face Validity)

ความเที่ยงตรงเชิงพินิจเป็นเพียงลักษณะที่ปรากฏของแบบทดสอบเท่านั้น และมีพื้นฐานอยู่ที่การตรวจสอบรูปลักษณะของข้อความว่า แบบทดสอบมีลักษณะที่จะเป็นเครื่องมือวัดที่เหมาะสมด้วยเหตุผลหรือไม่ สรุปได้ว่าลักษณะที่แตกต่างกันอย่างชัดเจนระหว่างวิธีการทั้ง 2

คือ การตรวจสอบความเที่ยงตรงของเนื้อหาที่มีพื้นฐานอยู่กับการสุ่มตัวอย่าง แต่ความเที่ยงตรงเชิงนิยามมีพื้นฐานอยู่ที่ลักษณะที่ปรากฏภายนอก

การตรวจสอบความเที่ยงตรงตามเนื้อหาและการพัฒนาแบบทดสอบ

การตรวจสอบความเที่ยงตรงของเนื้อหา มักเกิดขึ้นระหว่างการพัฒนาแบบทดสอบ ทั้งนี้เพราะเป็นสิ่งที่แรกของการเตรียมรายละเอียดของคคุณลักษณะเฉพาะของแบบทดสอบ และจากนั้นสร้างแบบทดสอบที่ตรงกับคคุณลักษณะเฉพาะดังกล่าว ซึ่งคคุณลักษณะเฉพาะที่อธิบายขอบเขตของผลสัมฤทธิ์ที่จะถูกประเมินจะต้องกลมกลืนกับสิ่งที่ได้สอนไปแล้ว ดังนั้น น้ำหนัก (Weights) ที่กำหนดเพื่อพิจารณาสัดส่วนของเนื้อหา จะสะท้อนส่วนที่เน้นในระหว่างการเรียนการสอน

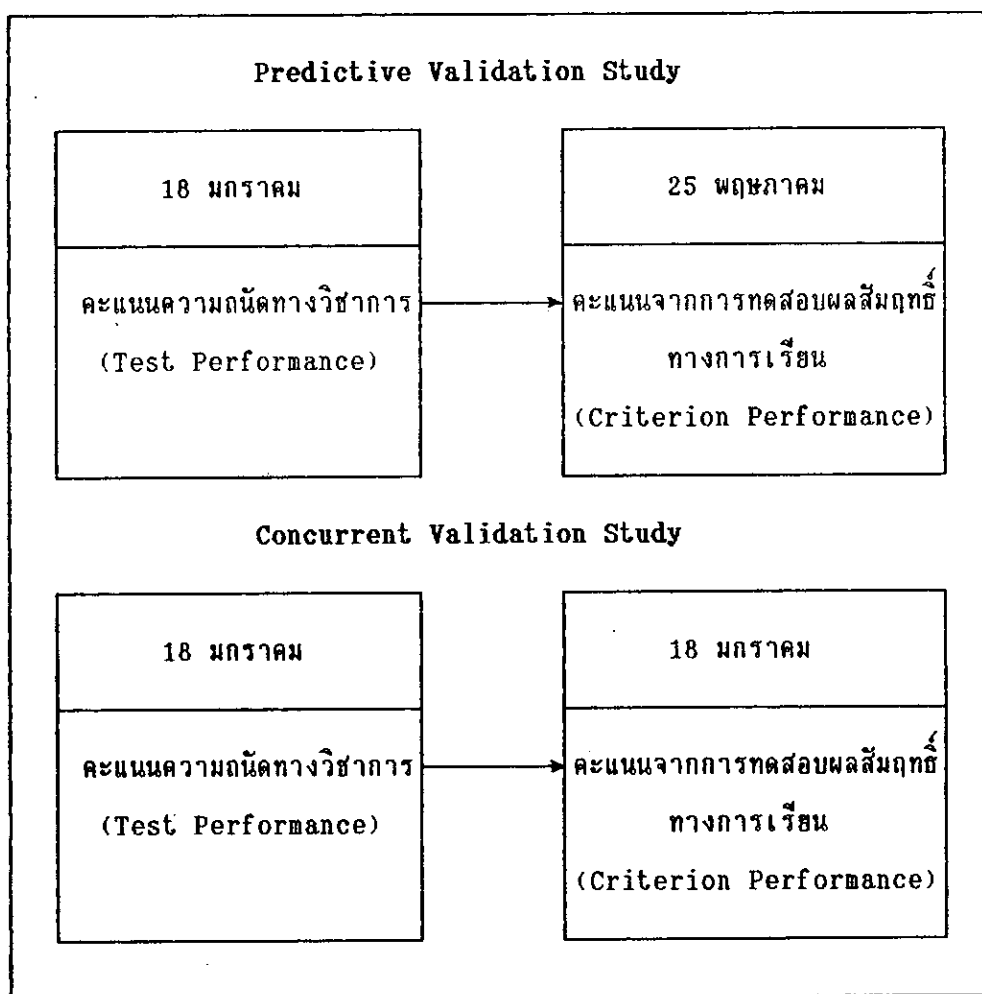
การตรวจสอบความเที่ยงตรงตามเนื้อหาและการเลือกใช้แบบทดสอบ

การเลือกใช้แบบทดสอบ ควรเลือกแบบทดสอบที่ให้การวัดขอบเขตของผลสัมฤทธิ์เฉพาะที่สอดคล้องกันเพราะจะให้ผลที่เที่ยงตรงที่สุด

Criterion-Related Evidence

เมื่อใดก็ตามที่คะแนนผลการสอบต้องถูกใช้เพื่อพยากรณ์ผลการปฏิบัติในอนาคต หรือเพื่อคาดคะเนผลการปฏิบัติในปัจจุบันที่เครื่องมือวัดซึ่งเป็นที่ยอมรับมากกว่าตัวแบบทดสอบเอง จะเป็นเรื่องที่เกี่ยวข้องกับ Criterion-Related Evidence เช่น คะแนนการทดสอบความพร้อมในการอ่าน อาจถูกใช้เพื่อพยากรณ์สัมฤทธิ์ผลในการอ่านของนักเรียนในอนาคต หรือการทดสอบทักษะการใช้พจนานุกรม อาจถูกใช้เพื่อคาดคะเนทักษะในการใช้พจนานุกรมจริงในปัจจุบัน ตัวอย่างแรกสนใจการพยากรณ์ และด้วยเหตุนี้ความสัมพันธ์ระหว่างการวัด หรือเครื่องมือทั้งสองจะขยายออกในช่วงเวลาต่อไป วิธีการนี้จะทำให้ได้มาซึ่งหลักฐานข้อมูลของความเที่ยงตรงที่เรียกว่า การศึกษาการตรวจสอบความเที่ยงตรงเชิงพยากรณ์ (Predictive Validation Study) กรณีการทดสอบทักษะการใช้พจนานุกรมนั้น จะให้ความสนใจไปที่การคาดคะเนสถานการณ์ปัจจุบัน ด้วยเหตุนี้ความสัมพันธ์ระหว่างการวัดหรือเครื่องมือทั้งสองที่ได้มาจะพ้องกัน ความสัมพันธ์ที่สูงในกรณีนี้จะแสดงว่า การทดสอบทักษะการใช้พจนานุกรมเป็นดัชนีบ่งชี้ทักษะที่แท้จริงในการใช้พจนานุกรม วิธีการสำหรับการได้มาของหลักฐานข้อมูลของความเที่ยงตรงนี้จำเป็นต้องศึกษา Concurrent Validation

การตรวจสอบความเที่ยงตรงทั้งสอง คือ ความเที่ยงตรงเชิงพยากรณ์ และความเที่ยงตรงเชิงสภาพ อยู่ในกลุ่มที่เรียกว่า Criterion-Related Validation ซึ่งมีความกว้างกว่า โดยถูกนิยามว่าเป็นกระบวนการของการกำหนดการแผ่ขยายของผลการสอบที่สัมพันธ์กับการปฏิบัติจากการวัดอื่นซึ่งเป็นที่ยอมรับ สิ่งวัดการปฏิบัติอื่นที่ 2 นี้ เรียกว่า เกณฑ์ที่อาจได้มาในอนาคตเมื่อสนใจในการพยากรณ์การปฏิบัติในอนาคต หรือในปัจจุบันเมื่อสนใจคาดคะเนการปฏิบัติขณะนั้น



ภาพประกอบ 8 ชนิดของ Criterion-Related Validation (วัดความต่างของเวลา)

ได้มีความพยายามที่จะทำให้ผลที่ได้สามารถสื่อความหมายได้อย่างดี โดยใช้วิธีการทางสถิติเข้ามาช่วยอธิบายความสัมพันธ์ของคะแนนทั้ง 2 ชุด สัมประสิทธิ์สหสัมพันธ์จะให้ตัวเลขสรุปแสดงระดับของความสัมพันธ์ระหว่างคะแนนทั้ง 2 ชุด ทั้งนี้ได้มีการนำตัวเลขแสดงระดับความสัมพันธ์ดังกล่าวไปทำให้เกิดความเข้าใจที่ง่ายขึ้น โดยนำเสนอผลเป็นกราฟในรูปของ Scatterplot หรือ Expectancy Table สัมประสิทธิ์สหสัมพันธ์ระหว่างชุดของคะแนนที่เป็นตัวพยากรณ์ (Predictor) และตัวเกณฑ์ (Criterion) เรียกว่า สัมประสิทธิ์ความเที่ยงตรง (Validity Coefficient) ถ้าค่าสัมประสิทธิ์ความเที่ยงตรงสูง ชี้ให้เห็นว่าจะมีความแม่นยำในการพยากรณ์ตัวแปรหนึ่งจากตัวแปรอีกตัวหนึ่งมากขึ้น

ปัญหาสำคัญสำหรับการศึกษา Criterion Validation ก็คือ การหาเกณฑ์ของความสำเร็จอันเป็นที่พอใจ ซึ่งกรอนลันด์ และลินน์ เห็นว่า สำหรับเป้าหมายของการศึกษาเกือบทั้งหมดไม่มีเกณฑ์ของความสำเร็จใดซึ่งเป็นที่พอใจอย่างแท้จริง

Construct-Related Evidence

Construct-Related Evidence จะช่วยตัดสินว่าคะแนนผลการสอบแสดงความเป็นตัวแทนขอบเขตของงานที่กำหนดให้ได้ดีเพียงใด และยังเป็นประโยชน์อย่างยิ่งทั้งสำหรับการเตรียมการสร้างและประเมินแบบทดสอบผลสัมฤทธิ์ เมื่อใดก็ตามที่เราต้องการตีความผลการทดสอบในรูปของคุณลักษณะหรือคุณภาพทางจิตวิทยาบางประการ จำเป็นจะต้องเกี่ยวข้องกับ Construct-Related Evidence

โครงสร้าง (Construct) เป็นคุณภาพทางจิตวิทยาที่เราสมมุติว่ามีอยู่ เพื่ออธิบายลักษณะบางประการของพฤติกรรม เช่น เหตุผลทางคณิตศาสตร์ สถิติปัญหา ความคิดสร้างสรรค์ ความวิตกกังวล และความซื่อสัตย์ ถือเป็นโครงสร้าง เพราะว่ามันเป็นโครงสร้างเชิงทฤษฎี (Theoretical Construction) ที่ถูกใช้อธิบายพฤติกรรม เมื่อเราตีความคะแนนผลการสอบในฐานะเป็นการวัดโครงสร้างเฉพาะ เราจะหมายความว่ามีความโครงสร้างดังกล่าวอยู่ ซึ่งมันจะต่างกับโครงสร้างอื่น

Construct Validation อาจถูกนิยามว่าเป็นกระบวนการของการกำหนดขอบเขต ที่ซึ่งผลการสอบสามารถถูกตีความไปถึงในรูปของโครงสร้างทางจิตวิทยานิ่งหรือมากกว่า สำหรับเป้าหมายที่เป็นวิทยาศาสตร์แล้ว ลักษณะที่สำคัญที่สุดของแบบทดสอบ คือ ความเที่ยงตรงเชิงโครงสร้าง ซึ่งหมายถึงระดับที่ซึ่งแบบทดสอบวัดโครงสร้างที่มันออกแบบเพื่อจะวัด ความยากของการสร้างความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบ คือ เกณฑ์

(Criterion) โครงสร้างเป็นสิ่งที่ไม่สามารถวัดได้โดยตรง แต่อาจศึกษาได้ทางอ้อม โดยการวิเคราะห์สหสัมพันธ์ของคะแนนที่ได้จากการสังเกต (Lord and Novick. 1968 : 278)

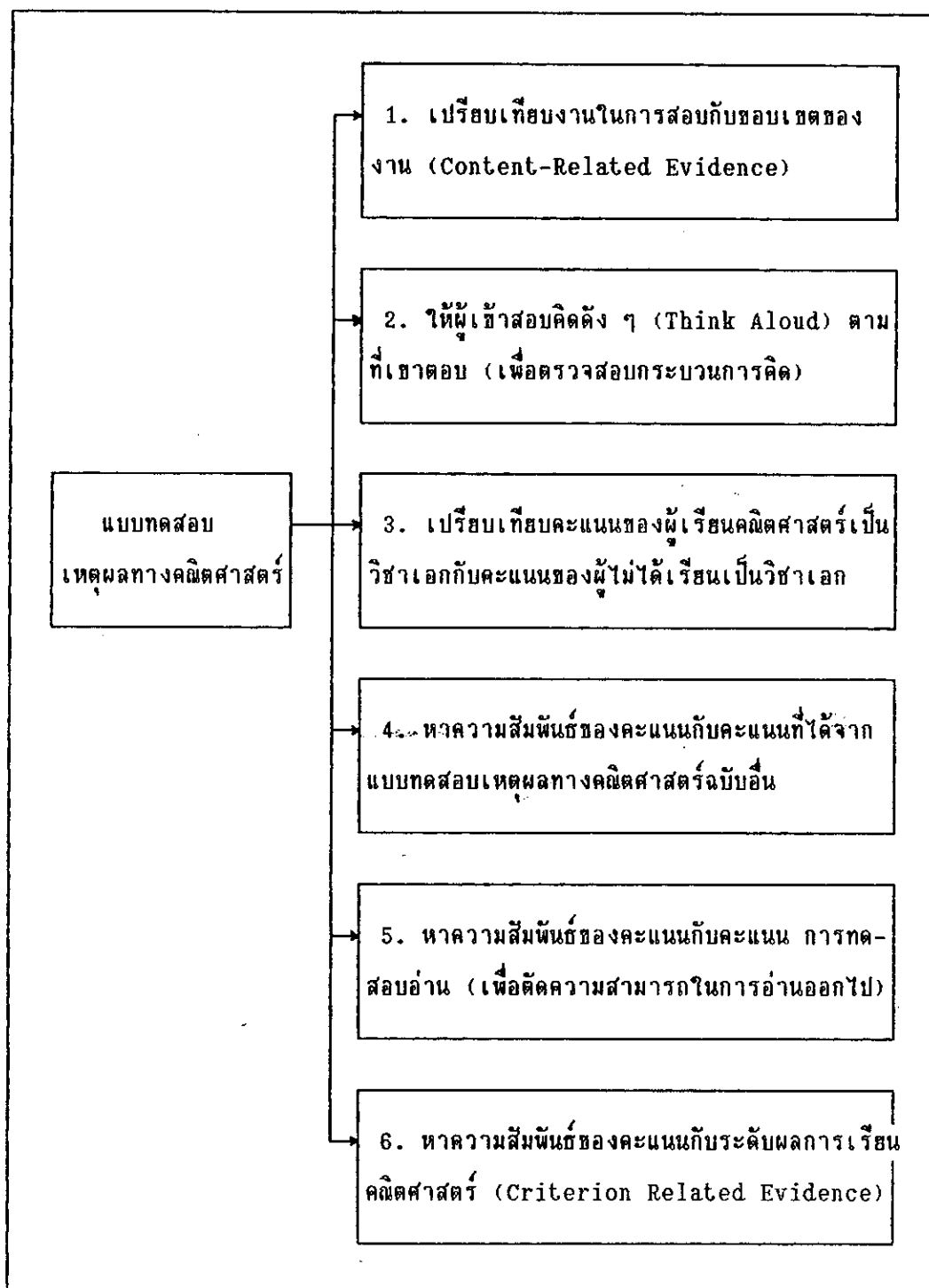
กระบวนการทั่วไปในการตรวจสอบความเที่ยงตรงเชิงโครงสร้าง ได้แก่

1. ระบุและอธิบายความหมายของโครงสร้างที่ถูกวัด โดยวิธีการใช้โครงสร้างเชิงทฤษฎี (Theoretical Framework)

2. สร้างสมมติฐานในเรื่องผลการสอบจากทฤษฎีที่เป็นพื้นฐานของโครงสร้าง

3. พิสูจน์สมมติฐานโดยวิธีการเชิงเหตุผลและเชิงประจักษ์

การตรวจสอบความเที่ยงตรงเชิงโครงสร้างจะเกิดขึ้นตั้งแต่เริ่มแรกระหว่างการพัฒนา และทดลองแบบทดสอบ และมีพื้นฐานอยู่ที่การสะสมข้อมูลหลักฐานจากแหล่งต่าง ๆ ขอบแสดงตัวอย่างชนิดของข้อมูลหลักฐาน ที่อาจนำมาใช้ในการตรวจสอบความเที่ยงตรงเชิงโครงสร้างของเหตุผลทางคณิตศาสตร์ (Mathematical Reasoning) ซึ่งกรอนลันด์ และลินน์ (Gronlund and Linn. 1990 : 68) ได้นำเสนอไว้ ดังภาพประกอบ 9



ภาพประกอบ 9 ข้อมูลหลักฐานบางประการที่ใช้ในการตรวจสอบความเที่ยงตรง
เชิงโครงสร้าง

จะเห็นว่าแหล่งข้อมูลหลักฐานมากมาย ซึ่งผู้สร้างหรือผู้เลือกจะต้องพิจารณาอย่างมีเหตุผลกับข้อมูลที่เกี่ยวข้องมากที่สุด

การตรวจสอบความเที่ยงตรงเชิงโครงสร้างขึ้นอยู่กับ การอ้างอิงอย่างมีเหตุผล ซึ่งถูกนำมาจากข้อมูลชนิดต่าง ๆ และทั้ง Content-Related และ Criterion-Related Evidence ก็ช่วยสนับสนุนการตีความได้ส่วนหนึ่ง แต่ต้องได้รับการเสริมโดยการศึกษาอื่น ๆ อีก เพื่อช่วยทำความเข้าใจความกระจ่างให้กับความหมายของคะแนนการทดสอบ

วิธีการที่มักถูกนำมาใช้ในการตรวจสอบความเที่ยงตรงเชิงโครงสร้าง มีดังนี้

1. นิยามขอบเขตของงาน (Tasks) ที่จะถูกวัด คุณลักษณะเฉพาะของข้อสอบ (Test Specifications) จะถูกนิยามอย่างดี เพื่อความหมายของโครงสร้างที่ชัดเจน และเกิดความเป็นไปได้ที่จะพิจารณาตัดสินว่าแบบทดสอบจะมีความสอดคล้อง และให้ตัวแทน การวัดขอบเขตของงานเพียงใด (Content Validation) ถ้าโครงสร้างได้ถูกวัด ข้อคำถามจะก่อให้เกิดการตอบสนองที่เหมือนกัน และมีความสัมพันธ์ระหว่างกันสูง

2. การวิเคราะห์กระบวนการคิดที่ต้องการโดยให้ข้อคำถาม กระบวนการคิดที่ถูกดึงออกมาโดยข้อคำถาม สามารถถูกกำหนดจากทั้งการตรวจสอบข้อคำถามและจากการดำเนินการใช้แบบทดสอบกับนักเรียนรายบุคคล และให้แต่ละคนคิดดัง ๆ อาจพิจารณาตัดสินโดยตรวจสอบจากการใช้แบบทดสอบกับนักเรียนรายคน และให้ผู้สอบอธิบายว่าเขาได้คำตอบมาอย่างไร เช่นเดียวกับการคิดอย่างดัง ๆ จะทำให้สามารถตรวจสอบได้ว่า ข้อคำถามวัดโครงสร้างใด หรือใช้วิธีการลองผิดลองถูกในการแก้ปัญหา

3. เปรียบเทียบคะแนนของกลุ่มที่รู้แน่ชัด (Known Groups) ในบางกรณีเป็นไปได้ที่จะพยากรณ์ว่า คะแนนจะต่างกันจากกลุ่มหนึ่งไปกลุ่มหนึ่ง อาจเป็นกลุ่มอายุที่ได้รับการฝึกอบรมและไม่ได้รับการฝึกอบรม ปรับตัวได้ดีและปรับตัวไม่ได้ ความสามารถส่วนใหญ่เพิ่มขึ้นตามอายุ อย่างน้อยก็เพิ่มขึ้นระหว่างเด็กผู้ใหญ่ ฉะนั้น จึงพอจะคาดได้ว่า คะแนนการทดสอบผลสัมฤทธิ์ทางการเรียนจะจำแนกความแตกต่างระหว่างกลุ่มที่ได้รับการฝึกอบรมต่างกัน ฉะนั้น การพยากรณ์ความแตกต่างสำหรับแบบทดสอบเฉพาะ สามารถตรวจสอบได้โดยเทียบกับกลุ่มที่รู้ว่ามี ความต่าง และผลจะใช้เป็นส่วนหนึ่งที่สนับสนุนการตรวจสอบความเที่ยงตรงเชิงโครงสร้าง

4. เปรียบเทียบคะแนนก่อนและหลังการทดลองหรือการจัดกระทำบางอย่าง คะแนนการทดสอบบางอย่างถูกคาดหวังว่าค่อนข้างจะทนทานต่อการฝึกฝนเฉพาะ เช่น สถิติปัญญาชน ที่บางอย่างจะเพิ่มขึ้นเช่นกัน คะแนนผลการสอบบางอย่างถูกคาดหวังว่าจะเปลี่ยนแปลงไปกับ

การทดลองหรือการจัดกระทำบางอย่างที่นำไป เช่น คะแนนจากแบบทดสอบความวิตกกังวลถูกคาดหวังว่าจะเปลี่ยนไป เมื่อบุคคลเป็นผู้เคยได้รับประสบการณ์ความวิตกกังวล ดังนั้นจากทฤษฎีรองรับลักษณะที่ถุกวัด เราสามารถพยากรณ์ได้ว่า คะแนนของการทดสอบเฉพาะจะเปลี่ยนแปลงหรืออาจคงที่ภายใต้เงื่อนไขหลายประการ ถ้าการพยากรณ์ของเราถูกตรวจสอบผลที่ได้จะสนับสนุนการตรวจสอบความเที่ยงตรงเชิงโครงสร้าง

5. หากความสัมพันธ์ของคะแนนกับการวัดอื่น ๆ คะแนนจากการสอบวัดหนึ่งอาจมีความสัมพันธ์กับคะแนนจากแบบทดสอบอื่นที่สันนิษฐานว่าวัดสิ่งเดียวกันลักษณะเดียวกัน คะแนนการทดสอบจะมีสหสัมพันธ์ต่ำ ถ้าแบบทดสอบวัดความสามารถที่ต่างกันและแบบทดสอบบางลักษณะอาจได้รับการพยากรณ์ว่ามีความสัมพันธ์กับเกณฑ์ที่ใช้ได้จริงในสภาพการปฏิบัติบางอย่าง

ไอกิน (Aiken. 1991 : 110 - 111) ให้ความเห็นเกี่ยวกับสิ่งแสดงความเที่ยงตรงเชิงโครงสร้าง ซึ่งมีลักษณะใกล้เคียงกับของกรอนลันด์ และลินน์ คือ ความเห็นของผู้รู้หรือผู้เชี่ยวชาญ เกี่ยวกับความสอดคล้องกันของเนื้อหาที่แบบทดสอบวัดกับโครงสร้างที่ต้องการวัด การวิเคราะห์ความคงที่ภายใน การหาความสัมพันธ์ระหว่างแบบทดสอบที่สร้างขึ้นกับแบบทดสอบอื่นหรือตัวแปรอื่นที่คาดว่าจะมีความสัมพันธ์กัน การวิเคราะห์องค์ประกอบ การตั้งคำถามผู้ตอบในรายละเอียดเกี่ยวกับการตอบสนองของเขาต่อแบบทดสอบ เพื่อแสดงกระบวนการคิดเฉพาะที่เกิดขึ้นในการตัดสินใจตอบสนองของเขา

จะเห็นว่า การตรวจสอบความเที่ยงตรงเชิงโครงสร้างจะนำข้อมูลหลายประเภทรวมทั้งใช้กระบวนการตรวจสอบที่เป็นเหตุผลและสถิติเข้ามาผสมผสานกัน

ข้อพิจารณาบางประการที่นำมาใช้ในการศึกษาความเที่ยงตรงเชิงโครงสร้าง ซึ่งแคมป์เบลล์ และฟิสค์ (Lord and Novick. 1968 : 352; citing Campbell and Fiske. 1959) ได้เสนอความคิดเกี่ยวกับความเที่ยงตรงเหมือน (Convergent Validity) และความเที่ยงตรงจำแนก (Discriminant Validity หรือ Divergent Validity) ทั้งนี้ ทั้ง 2 คน เชื่อว่า มาตรวัดที่มีความเหมาะสมสามารถได้มาจากการศึกษาความเที่ยงตรงและความเชื่อมั่นที่แสดงให้เห็นว่า

1. มาตรวัดมีความสัมพันธ์กับมาตรวัดอื่น ซึ่งโดยทางทฤษฎีระบุว่าจะต้องมีความสัมพันธ์กัน หรือกล่าวอย่างง่าย ๆ ว่า ถ้าแบบทดสอบที่สร้างขึ้นมาวัดลักษณะเดียวกันแล้ว ผลจากการสอบด้วยแบบทดสอบเหล่านั้นมีสหสัมพันธ์กันสูง แบบทดสอบนั้นถือว่า มีความเที่ยงตรงเหมือน

2. มาตรฐานไม่มีความสัมพันธ์กับมาตรฐานอื่น ที่โดยทางทฤษฎีนี้ระบุว่าไม่ควรมีความสัมพันธ์กัน หรือกล่าวได้ว่าถ้าแบบทดสอบที่สร้างขึ้นมาวัดลักษณะต่าง ๆ กันแล้ว ผลจากการสอบด้วยแบบทดสอบเหล่านั้นมีสัมประสิทธิ์สหสัมพันธ์ต่ำหรือไม่มีความสัมพันธ์กันเลย ถือว่าแบบทดสอบนั้นมีความเที่ยงตรงจำแนก

การตรวจสอบความเที่ยงตรงเชิงโครงสร้าง มีความสำคัญสำหรับการทดสอบทุกชนิด ผลสัมฤทธิ์ ความถนัด การพัฒนาการทางสังคม ไม่ว่าจะเป็นการคัดเลือก หรือสร้างแบบทดสอบ ความหมายของคะแนนการทดสอบจะขึ้นอยู่กับความระมัดระวังต่อสิ่งซึ่งแบบทดสอบถูกสร้าง และข้อมูลหลักฐานที่สนับสนุนชนิดของการตีความที่จะกระทำ Construct-Related Evidence จะรวม Content และ Criterion Related Evidence ด้วย และมีความสำคัญต่อการตีความหมายของคะแนน ทั้งนี้เมสซิค (Messick) ได้กล่าวเน้นย้ำว่า "ความหมายของการวัดอันเนื่องมาจากความเที่ยงตรงตามโครงสร้าง (Construct Validity) ไม่เพียงแต่มุ่งไปสู่การแปลความหมายของคะแนนเท่านั้น แต่จะทำให้เหตุผลสนับสนุนการใช้แบบทดสอบด้วย" (Gronlund and Linn. 1990 : 70)

องค์ประกอบที่มีอิทธิพลต่อความเที่ยงตรง

มีองค์ประกอบมากมายที่มักทำให้ผลการสอบไม่ถูกต้องกับความตั้งใจที่จะใช้ ได้แก่

1. องค์ประกอบที่มาจากตัวแบบทดสอบเอง องค์ประกอบที่ขัดขวางข้อคำถามจากการทำหน้าที่มุ่งหวัง และทำให้ความเที่ยงตรงในการแปลความหมายคะแนนต่ำลง คือ

- 1.1 ความไม่ชัดเจนของคำชี้แจง
- 1.2 คำศัพท์ที่ให้อ่านและโครงสร้างประโยคยากเกินไป
- 1.3 ระดับความยากของข้อคำถามไม่เหมาะสม
- 1.4 ข้อคำถามถูกสร้างอย่างไม่สมบูรณ์
- 1.5 ความกำกวมของข้อความในข้อคำถาม
- 1.6 ข้อคำถามของแบบทดสอบไม่เหมาะสมกับสิ่งที่ถูกวัด
- 1.7 กำหนดเวลาไม่เหมาะสม
- 1.8 แบบทดสอบสั้นเกินไป
- 1.9 การจัดเรียงลำดับข้อคำถามไม่เหมาะสม
- 1.10 แบบแผนของคำตอบถูกปรากฏชัดเจน

ด้วยจุดบกพร่องของการสร้างแบบทดสอบที่ทำให้ข้อคำถามมิได้ทำหน้าที่ที่ตั้งใจ จะทำให้การแปลความหมายของคะแนนถูกต้องน้อยลง

2. วิธีการสอน แม้แบบทดสอบจะถามคำถามที่ผู้เรียนต้องใช้กระบวนการคิดที่ซับซ้อนระดับสูง แต่หากครูได้สอนวิธีการแก้ปัญหาให้แก่ผู้เรียนไปแล้ว ข้อคำถามของแบบทดสอบดังกล่าวจะวัดได้ไม่มากไปกว่าความรู้ความจำ

3. องค์ประกอบในการดำเนินการสอบและการตรวจให้คะแนน การดำเนินการสอบและการตรวจให้คะแนน อาจนำไปสู่องค์ประกอบที่ทำให้ความเที่ยงตรงในการแปลความหมายจากผลที่ได้ลดลงได้ ในกรณี เช่น แบบทดสอบที่ครูสร้างเอง อาจได้แก่ องค์ประกอบเกี่ยวกับความไม่พอเพียงของเวลาที่ใช้ทำข้อสอบ การล่าเอียงกับผู้สอบบางคน การโกงระหว่างการสอบ และความไม่คงเส้นคงวาของการตรวจให้คะแนนคำตอบประเภทความเรียง (Essay Answers) จะทำให้ความเที่ยงตรงต่ำลง

4. การตอบสนองของผู้สอบ ในบางกรณีความไม่เที่ยงตรงในการแปลความหมายอาจมีสาเหตุเนื่องมาจากองค์ประกอบส่วนบุคคล ที่มีอิทธิพลผลักดันการตอบสนองในสถานการณ์การสอบมากกว่าข้อบกพร่องของเครื่องมือหรือการดำเนินการ ผู้สอบบางคนอาจตกใจในสถานการณ์การสอบ บางคนอาจถูกรบกวนโดยอารมณ์ที่รบกวน การทำแบบทดสอบชุดของการตอบสนอง (Response Set) ถือเป็นองค์ประกอบหนึ่งที่มีอิทธิพลต่อผลที่ได้

5. ธรรมชาติของกลุ่มและเกณฑ์ ความเที่ยงตรงมักใช้เฉพาะกับกลุ่มใดกลุ่มหนึ่ง แบบทดสอบคณิตศาสตร์ที่วัดปัญหาที่เป็นเรื่องราว อาจวัดความสามารถด้านเหตุผลในกลุ่มผู้เรียนช้า และรวมถึงการระลึกข้อมูลอย่างง่าย และทักษะการคำนวณในกลุ่มที่เรียนดีกว่า

สัมประสิทธิ์ความเที่ยงตรงของ Criterion-Related จะต่ำลงเมื่อคุณลักษณะที่วัดโดยตัวพหุศาสตร์และเกณฑ์มีความเหมือนกันน้อย การกระจายของคะแนนต่ำ ความคงที่ของคะแนนต่ำ และช่วงเวลาระหว่างการวัดนานกว่า สิ่งเหล่านี้แสดงให้เห็นว่า ธรรมชาติของเกณฑ์จะช่วยในการประเมินสัมประสิทธิ์ความเที่ยงตรง

กล่าวโดยสรุปจะเห็นว่า ความเที่ยงตรงถือเป็นคุณสมบัติที่สำคัญยิ่งที่ควรได้รับการพิจารณาเมื่อสร้างและเลือกใช้แบบทดสอบ ม้องค์ประกอบมากมายที่มีอิทธิพลต่อความเที่ยงตรงของการแปลความหมายการทดสอบ ฉะนั้น จำเป็นจะต้องควบคุมองค์ประกอบที่อาจส่งผลต่อความเที่ยงตรงและการแปลความผล

2.3.1 ความเชื่อมั่น (Reliability)

คุณลักษณะที่มีความสำคัญมากของผลการประเมินในลำดับต่อจากความเที่ยงตรง คือ ความเชื่อมั่น (Reliability) ความเชื่อมั่นจะให้ความคงเส้นคงวาที่ทำให้ความเที่ยงตรงมีความเป็นไปได้ และเป็นตัวชี้วัดว่าความมั่นใจกับผลที่ได้เพียงใด

ความเชื่อมั่น หมายถึง ความคงเส้นคงวาของคะแนนการทดสอบและผลการประเมินอื่น ๆ จากการวัดหนึ่งไปยังการวัดอื่น ในการแปลความหมายและใช้ข้อมูลเกี่ยวกับความเชื่อมั่น ต้องระลึกไว้ว่าความเชื่อมั่นเป็นผลที่ได้จากเครื่องมือประเมิน แต่ไม่ใช่ตัวเครื่องมือ เครื่องมือประเมินเฉพาะอันหนึ่งอาจมีความเชื่อมั่นต่าง ๆ กัน ขึ้นอยู่กับกลุ่มที่เกี่ยวข้องและสถานการณ์ที่มันถูกใช้ ดังนั้น จึงเป็นการเหมาะสมกว่าหากจะเรียกว่า ความเชื่อมั่นของคะแนนผลการสอบหรือของการวัดมากกว่าของแบบทดสอบหรือเครื่องมือ

โดยทั่วไปคะแนนผลการสอบจะไม่คงที่ มันจะเชื่อถือได้ในช่วงเวลาต่างกัน หรือตัวอย่างข้อคำถามที่ต่างกัน ผู้ประเมินที่ต่างกัน และเป็นไปได้สำหรับคะแนนผลการสอบจะคงที่ในส่วนหนึ่งและไม่คงที่ในอีกส่วนหนึ่ง ความเหมาะสมของชนิดของความคงที่ในกรณีหนึ่งจะถูกกำหนดโดยการใช้นั้น ๆ เช่น ถ้าประสงค์จะรู้ว่าบุคคลจะเป็นเช่นไรในเวลาข้างหน้า ความคงที่ของคะแนนตลอดช่วงเวลาที่นอกเหนือไปจะมีความสำคัญ ดังนั้น สำหรับการแปลความหมายที่ต่างกัน เราจำเป็นต้องวิเคราะห์ความคงที่ต่างกัน แต่ถ้าหากถือว่าความเชื่อมั่นมีลักษณะปกติธรรมดาจะนำไปสู่การแปลความหมายที่คลาดเคลื่อน ความเชื่อมั่นเป็นสิ่งจำเป็นแต่ไม่เป็นสภาพที่พอเพียงสำหรับความเที่ยงตรง แบบทดสอบที่ให้ผลซึ่งไม่คงที่จะไม่มีทางให้ข้อมูลผลการปฏิบัติที่ถูกวัดซึ่งมีความเที่ยงตรง แต่ในทางตรงกันข้ามผลการสอบวัดที่คงที่ อาจวัดสิ่งผิดหรืออาจถูกใช้ในทางที่ไม่เหมาะสม ดังนั้น ความเชื่อมั่นที่ต่ำสามารถคาดคะเนได้ว่าจะไปจำกัดระดับของค่าความเที่ยงตรงที่ได้ แต่ความเชื่อมั่นที่สูงไม่ทำให้แน่ใจว่าระดับความเที่ยงตรงที่น่าพอใจจะเกิดขึ้น กล่าวได้อย่างสั้น ๆ ว่า ความเชื่อมั่นเพียงแต่ให้ความคงที่ ซึ่งทำให้ความเที่ยงตรงมีความเป็นไปได้ และนอกจากนี้ค่าความเชื่อมั่นเป็นเรื่องของสถิติ ความคงที่จะถูกรายงานโดยค่าสัมประสิทธิ์สหสัมพันธ์ที่เรียกว่า สัมประสิทธิ์ความเชื่อมั่น (Reliability Coefficient) และความคลาดเคลื่อนมาตรฐานในการวัด (Standard Error of Measurement)

สัมประสิทธิ์สหสัมพันธ์ หมายถึง ค่าสถิติที่วัดระดับของความสัมพันธ์ระหว่างคะแนน 2 ชุดใด ๆ ที่ได้จากบุคคลกลุ่มเดียวกัน เช่น ความสัมพันธ์ระหว่างความสูงกับน้ำหนัก

สัมประสิทธิ์ความเชื่อมั่น หมายถึง สัมประสิทธิ์สหสัมพันธ์ที่ระดับความสัมพันธ์ระหว่างการวัด 2 ชุด ที่ได้จากเครื่องมือหรือวิธีการอย่างเดียวกัน เช่น ความสัมพันธ์ระหว่างคะแนนที่ได้จากการใช้แบบทดสอบ 2 ฉบับ

การกำหนดค่าความเชื่อมั่นโดยวิธีสหสัมพันธ์

วิธีการประมาณค่าความเชื่อมั่นมีหลายวิธีแตกต่างกัน สัมประสิทธิ์ความเชื่อมั่นที่ได้จากแต่ละวิธีจะถูกแปลความหมายตามชนิดของความคงที่ที่ถูกลึกศึกษา

สมาคมจิตวิทยาแห่งอเมริกา (American Psychological Association) ได้นำเสนอวิธีการในการประมาณค่าความเชื่อมั่นไว้หลายวิธี ดังต่อไปนี้ (Gronlund and Linn. 1990 : 80) ดังภาพประกอบ 10

วิธีการประมาณค่า	ชนิดของการวัดความเชื่อมั่น	วิธีการ
วิธีการสอบซ้ำ (Test-Retest Method)	การวัดความคงที่	ให้คนกลุ่มเดียวกันสอบแบบทดสอบฉบับเดียวกัน 2 ครั้ง โดยมีช่วงเวลาระหว่างการสอบจากหลายนาที่ ---> หลายปี
วิธีการใช้แบบทดสอบที่เท่าเทียมกัน (Equivalent-Forms Method หรือ Forms)	วัดความเท่ากัน	ใช้แบบทดสอบ 2 รูปแบบ กับกลุ่มเดียวกัน ในระยะเวลาติดต่อกัน

ภาพประกอบ 10 วิธีการประมาณค่าความเชื่อมั่น ที่นำเสนอโดย American Psychological Association (Standard for Educational and Psychological Testing, Washington D.C. APA. 1985)

วิธีการประมาณค่า	ชนิดของการวัดความเชื่อมั่น	วิธีการ
ทดสอบซ้ำด้วยแบบทดสอบที่เท่าเทียมกันหรือเหมือนกัน (Test-Retest With Equivalent Forms)	การวัดความคงที่และความเหมือนกัน	ใช้แบบทดสอบ 2 ฟอรม์กับกลุ่มเดียวกัน ด้วยการเพิ่มช่วงเวลาระหว่างฟอรม์
วิธีแบ่งครึ่ง (Split-Half Method)	วัดความคงที่ภายใน	ให้ทดสอบครึ่งเดียว ตรวจสอบให้คะแนนแบบทดสอบเป็น 2 ส่วนเท่า ๆ กัน (เช่น ข้อ คู่-คี่); ตรวจสอบถึงความถูกต้องของสหสัมพันธ์ระหว่าง 2 ส่วน เพื่อให้พอดีกับแบบทดสอบทั้งฉบับ โดยใช้สูตรของสเปียร์แมน-บราวน์
วิธีการคูเดอร์-ริชาร์ด (Kuder-Richardson Method)	วัดความคงที่ภายใน	ให้ทำการทดสอบครึ่งเดียว ตรวจสอบให้คะแนนทั้งฉบับ และใช้สูตรของคูเดอร์-ริชาร์ดสัน

ภาพประกอบ 10 (ต่อ)

จะเห็นว่า การประมาณค่าความเชื่อมั่นสามารถกระทำได้หลายวิธี แต่ละวิธีมีจุดดีและข้อจำกัดที่ต่างกัน ทำให้ค่าประมาณของความเชื่อมั่นที่ได้ต่างกัน ฉะนั้น เมื่อตรวจสอบสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบมาตรฐาน จำเป็นจะต้องพิจารณาวิธีการที่ใช้ในการประมาณค่าความเชื่อมั่น ขนาดของสัมประสิทธิ์ความเชื่อมั่นจะสัมพันธ์กับวิธีการประมาณค่าความเชื่อมั่น คือ

1. ความยาวของแบบทดสอบ (Length of Test)
2. การกระจายของคะแนน (Spread of Scores)
3. ความยากของแบบทดสอบ (Difficulty of Test)
4. ความเป็นปรนัยของข้อคำถามในแบบทดสอบ (Objectivity)

แม้จะพบว่า มีองค์ประกอบบางประการที่มีอิทธิพลต่อค่าความเที่ยงตรงและค่าความเชื่อมั่นของแบบทดสอบ แต่ไม่ได้หมายความว่า ผู้สร้างและผู้เลือกใช้จะพยายามหลีกเลี่ยงแบบทดสอบที่อาจมีองค์ประกอบดังกล่าวเกี่ยวข้องกับโดยไม่ใช้เลย เช่น จะเลือกใช้เฉพาะแบบทดสอบเลือกตอบ เพื่อความเป็นปรนัยในการตรวจให้คะแนน โดยละเลยวิธีการประเมินแบบ Subjective ไป ซึ่งกรณีนี้อาจส่งผลในทางตรงกันข้ามกับความเที่ยงตรง ซึ่งถือว่าเป็นคุณภาพที่สำคัญที่สุดของการประเมินผล ฉะนั้น ทางแก้ปัญหาคือ การคัดเลือกวิธีการประเมินที่เหมาะสมที่สุดสำหรับพฤติกรรมที่กำลังถูกประเมิน และจากนั้นพยายามทำให้วิธีการประเมินมีความเป็นปรนัยให้มากที่สุด เป็นต้น

ความเชื่อมั่นที่ต้องการควรอยู่ในระดับใด เป็นคำถามหนึ่งที่ผู้สร้างแบบทดสอบและผู้เลือกใช้แบบทดสอบต้องการคำตอบ ซึ่งกรอนลินน์ และลินน์ (Gronlund and Linn. 1990 : 100 - 101) ได้ให้ความเห็นว่า ระดับของความเชื่อมั่นที่ต้องการใช้ในทางการศึกษาขึ้นอยู่กับระดับของการตัดสินใจระดับความเชื่อมั่นที่ต้องการ และสิ่งนั้นสามารถยืนยันได้ในเวลาต่อมา ยิ่งต้องการความมั่นใจมากก็จำเป็นต้องใช้ความเชื่อมั่นที่สูงขึ้นดังปรากฏดังภาพประกอบ 11

ความเชื่อมั่นที่ต้องการและธรรมชาติของการตัดสินใจ

ความเชื่อมั่นสูง เป็นที่ต้องการเมื่อ

1. การตัดสินใจมีความสำคัญ
2. เป็นการตัดสินใจครั้งสุดท้าย
3. การตัดสินใจที่ไม่อาจเปลี่ยนแปลงได้
4. การตัดสินใจที่เกี่ยวข้องกับบุคคล
5. การตัดสินใจที่มีผลที่ตามมาจะมีความต่อเนื่องเป็นระยะเวลานาน

ตัวอย่างของกรณีข้างต้น ได้แก่ การคัดเลือกผู้สมัครไว้หรือคัดออก

ความเชื่อมั่นต่ำที่พอยอมรับได้ เมื่อเป็นการตัดสินใจที่

1. มีความสำคัญน้อย
2. กระทำในขั้นต้น
3. เปลี่ยนแปลงได้
4. ถูกยืนยันโดยข้อมูลอื่น
5. เกี่ยวข้องกลุ่ม
6. มีผลชั่วคราว

กรณีตัวอย่าง ได้แก่ การทบทวนบทเรียนในชั้น

ภาพประกอบ 11 เกณฑ์การพิจารณาระดับความเชื่อมั่น

2.3.3 ความสามารถนำไปใช้ประโยชน์ (Usability)

ในการคัดเลือกแบบทดสอบและเครื่องมือประเมินอื่น ไม่อาจละเลยการพิจารณาในเชิงปฏิบัติได้ แบบทดสอบมักถูกนำไปดำเนินการและแปลความโดยครูที่ได้รับการฝึกอบรมด้านการวัดมาเพียงเล็กน้อย เวลาที่ต้องใช้ในการทดสอบมักถูกจำกัด ค่าใช้จ่ายในการทดสอบจะถูกพิจารณาโดยผู้บริหารกองทุนของโรงเรียนเหล่านี้ ถือเป็นสิ่งที่มีความเกี่ยวข้องกับความสามารถนำไปใช้ประโยชน์ของแบบทดสอบและวิธีการประเมินผล ที่ต้องได้รับการพิจารณาเมื่อเลือกใช้เครื่องมือในการประเมิน สิ่งที่ต้องพิจารณาในทางปฏิบัติ ได้แก่

1. ความสะดวกในการดำเนินการสอบ
 2. เวลาที่ต้องใช้ในการดำเนินการสอบ
 3. สะดวกในการแปลความหมายและการนำไปใช้
 4. ความสามารถในการหาแบบทดสอบที่เทียบเท่ากันได้มาใช้
 5. ค่าใช้จ่ายในการทดสอบ อย่างไรก็ตาม ในการวัดทางจิตวิทยาเน้นคุณลักษณะสำคัญของแบบทดสอบที่ดี 3 ประการ คือ ความเป็นปรนัย (Objectivity) ความเที่ยงตรง (Validity) และความเชื่อมั่น (Reliability) (Mitchell. 1992 : 23)
- ดังได้กล่าวมาแล้วจะเห็นว่า การตรวจสอบคุณภาพของเครื่องมือวัดสามารถกระทำได้ 2 วิธี คือ การตรวจสอบคุณภาพก่อนนำเครื่องมือไปใช้ ซึ่งเป็นวิธีการตรวจสอบโดยให้เหตุผลหรือตรรกะ และ การตรวจสอบคุณภาพหลังจากนำเครื่องมือไปใช้แล้ว เป็นการตรวจสอบผลที่ได้จากการวัดโดยการคำนวณค่าทางสถิติ

2.4 ทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory : IRT) ได้ถูกนำมาประยุกต์ใช้ในกระบวนการวัดผลประเมินผลอย่างแพร่หลายในหลาย ๆ ด้าน ทั้งในลักษณะที่นำมาใช้เพื่อสนับสนุน หรือเสริมในจุดอ่อนบางประการของทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory : CTT) และในลักษณะของการนำความก้าวหน้าใหม่ ๆ ของทฤษฎีมาใช้ในการพัฒนากระบวนการวัดผลให้ก้าวหน้ามากขึ้น ขอบข่ายของการนำทฤษฎีการตอบสนองข้อสอบไปใช้ มีดังนี้

1. การวิเคราะห์ข้อสอบ (Item Analysis)
2. การสร้างธนาคารข้อสอบ (Item Bank)
3. ตรวจสอบความลำเอียงของข้อสอบ (Item Bias)
4. ใช้ศึกษาลักษณะการตอบแบบทดสอบของผู้สอบ (Response Pattern)
5. การเทียบคะแนน (Equating)
6. การออกแบบแบบทดสอบที่ดีที่สุด (Best Test Design) และ
7. การสร้างแบบทดสอบที่มีความเหมาะสมกับผู้สอบ (Tailor Test หรือ Adaptive Test)

เนื่องจากข้อมูลรายละเอียดเกี่ยวกับทฤษฎีการตอบสนองข้อสอบมีมากมาย และค่อนข้างทำความเข้าใจได้ยากสำหรับบุคคลทั่วไปซึ่งมิได้ศึกษาโดยตรง ผู้วิจัยจึงขอเสนอแนะเฉพาะ

หลักการที่สำคัญเกี่ยวกับทฤษฎี และประเด็นที่เกี่ยวข้องโดยตรงต่อการศึกษาครั้งนี้เท่านั้น
ดังต่อไปนี้

2.4.1 แนวคิดและลักษณะทั่วไปของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบเป็นทฤษฎีการวัดที่มีวัตถุประสงค์เพื่อเสนอหลักการพื้นฐาน
ที่สำคัญสำหรับการทำนาย การประมาณค่าการอ้างอิงเกี่ยวกับความสามารถหรือลักษณะ
(Ability หรือ Trait) ของบุคคล ซึ่งสามารถวัดได้ด้วยแบบทดสอบ (Hambleton and
Swaminathan. 1985 : 53)

2.4.2 ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบมีข้อตกลงเบื้องต้น ดังนี้

1. ข้อตกลงเกี่ยวกับมิติ (Dimensionality) ทฤษฎีการตอบสนองข้อสอบมีความ
เชื่อว่า ผู้สอบมีความสามารถ k อย่าง ซึ่งความสามารถแต่ละอย่างต่างก็ตอบสนองต่อสิ่งเร้า
ต่าง ๆ กันไป แบบทดสอบที่นำมาวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบนั้น มีข้อตกลงว่า
ต้องมีความเป็นมิติเดียว (Unidimension) นั่นคือ เป็นแบบทดสอบที่ประกอบด้วยข้อสอบ
หลาย ๆ ข้อ ซึ่งใช้สำหรับวัดความสามารถด้านใดด้านหนึ่งของผู้สอบ (สำเร็จ บุญเรืองรัตน์.
2529 : 43) สิ่งที่ยอมรับว่าเป็นมิติเดียว ก็คือ องค์ประกอบที่เด่นชัดหรือมีความสำคัญ
ต่อการอธิบายผลการตอบแบบทดสอบ ซึ่งในการกำหนดข้อตกลงเบื้องต้นเกี่ยวกับความเป็น
มิติดังกล่าวนี้ ก็เพื่อประโยชน์ในการตีความหมายของคะแนนที่ได้จากการสอบนั่นเอง (Lord.
1980 : 19)

การที่จะได้ข้อสอบที่มีความเป็นมิติเดียวนั้น ต้องอาศัยวิธีการคัดเลือกข้อสอบและวิธี
การตรวจสอบที่ดี (Lumsden. 1961, 1976) วิธีที่ได้รับความนิยมมาก ได้แก่ วิธีวิเคราะห์
องค์ประกอบ (Factor Analysis) ซึ่งจะใช้อัตราส่วนของความแปรปรวนขององค์ประกอบ
แรกกับความแปรปรวนขององค์ประกอบที่สอง เป็นดัชนีความเป็นมิติเดียว (Index of Uni-
dimensionality) แต่ปัญหาที่พบในการวิเคราะห์องค์ประกอบ ก็คือ การเลือกใช้ค่าสห-
สัมพันธ์ ระหว่างค่าสหสัมพันธ์ฟี (Phi Correlation) และค่าสหสัมพันธ์เตตราคลอริก
(Tetrachoric Correlation) สำหรับการวิเคราะห์โดยใช้สหสัมพันธ์ฟีนั้น ทำให้ได้
องค์ประกอบหลายองค์ประกอบ อันอาจจะทำให้ได้องค์ประกอบที่ไม่ใช่องค์ประกอบที่แท้จริง
ทั้งนี้ แมคโดนัล และอลาวัต (McDonald and Ahlawat. 1974) ได้ชี้ให้เห็นว่า ที่เกิด

ผลเช่นนี้ก็เพราะข้อมูลมีความสัมพันธ์กันไม่เป็นเส้นตรง (Nonlinearity) ส่วนการใช้ค่าสหสัมพันธ์เตตราคลอริกในการวิเคราะห์นั้น ค่อนข้างยุ่งยากในการคำนวณ ไม่เหมาะสมที่จะนำมาใช้เมื่อความสามารถของผู้สอบไม่แจ่มแจ้งเป็นปกติ ฟังก์ชันของการตอบสนองข้อสอบไม่เป็นแบบปกติสะสมและเมื่อมีการเดา (Lord, 1980 : 20) แต่ข้อดีของการใช้สหสัมพันธ์เตตราคลอริก ก็คือ ผลของการวิเคราะห์จะต้องประกอบที่เด่นชัดเพียงองค์ประกอบเดียว และสิ่งนี้ก็เป็นเงื่อนไขที่พอเพียง สำหรับการตรวจสอบความเป็นมิติเดียวของกลุ่มข้อสอบที่ถูกวิเคราะห์ (Lord and Novick, 1968)

อย่างไรก็ตาม บางครั้งการประเมินความเป็นมิติเดียวด้วยวิธีการวิเคราะห์องค์ประกอบก็ได้อผลไม่เป็นที่น่าพอใจ เพราะการวิเคราะห์องค์ประกอบเป็นวิธีการที่ดำเนินการกับข้อมูลที่มีความสัมพันธ์กันเชิงเส้นตรง ดังนั้น ผลการวิเคราะห์ก็อาจจะไม่ได้มิติเดียว ถึงแม้ว่ากลุ่มข้อสอบจะมีความเป็นมิติเดียวก็ตาม หากข้อมูลไม่สัมพันธ์กันเชิงเส้นตรง และส่วนใหญ่แล้วข้อมูลที่เหมาะสม (Fit) กับโมเดลของทฤษฎีการตอบสนองข้อสอบก็มักมีลักษณะสัมพันธ์กันไม่เป็นเส้นตรง

แฮมเบิลตัน และสวามินาธาน (Hambleton and Swaminathan, 1985 : 307) ได้ให้ข้อเสนอแนะว่า ในการประเมินความเป็นมิติเดียวของกลุ่มข้อสอบนั้น ถ้าอิงความเป็นอิสระของข้อสอบ (Local Independence) ก็อาจมีความเหมาะสมมากกว่า เนื่องจากความเป็นอิสระของข้อสอบนั้น ได้มาในแต่ละระดับความสามารถและไม่ขึ้นอยู่กับกลุ่มผู้สอบ

แบบทดสอบที่น่าจะมีความเป็นมิติเดียว เช่น แบบทดสอบการสะกดคำ คำศัพท์อ่านจับใจความ เหตุผลเชิงคณิตศาสตร์ อุปมาอุปไมย อนุกรมตัวเลข และมิติสัมพันธ์ ส่วนแบบทดสอบที่ไม่เป็นมิติเดียว เช่น แบบทดสอบวัดผลสัมฤทธิ์วิชาเคมี ซึ่งต้องอาศัยความสามารถหลายอย่าง ทั้งที่เกี่ยวกับคณิตศาสตร์และไม่เกี่ยวกับคณิตศาสตร์ในการตอบข้อสอบ (Lord, 1980 : 20)

2. ข้อตกลงเกี่ยวกับความเป็นอิสระของข้อสอบ (Local Independence) ข้อตกลงนี้กล่าวว่า โอกาสในการตอบข้อสอบข้อหนึ่งได้ถูกต้องนั้นจะไม่เกี่ยวข้องกับหรือไม่ผลต่อการตอบข้อสอบข้ออื่น ๆ ในแบบทดสอบ ทั้งนี้โดยถือว่าพฤติกรรมในการตอบข้อสอบต่าง ๆ ในแบบทดสอบของบุคคลใดบุคคลหนึ่งมีความเป็นอิสระเชิงสถิติ และถ้าการตอบข้อสอบมีความเป็นอิสระแล้ว ความน่าจะเป็นร่วมของคะแนนรายชื่อของผู้เข้าสอบแต่ละคน คือ

$$\text{Prob } [U_1 = u_1, U_2 = u_2, \dots, U_n = u_n / \mathbf{e}] = \prod_{i=1}^n P_i(\mathbf{e})^{u_i} Q_i(\mathbf{e})^{1-u_i}$$

เมื่อ $U_i, i = 1, 2, \dots, n$ เป็นผลการตอบข้อสอบ n ข้อ (เมื่อตอบถูก

$u_i = 1$ เมื่อตอบผิด $u_i = 0$)

P_i = ความน่าจะเป็นของผู้สอบคนหนึ่งที่ทำข้อสอบข้อที่ i ถูก

$Q_i = 1 - P_i$

จากสมการกล่าวได้ว่า ความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องทุกข้อ เท่ากับ ผลคูณของความน่าจะเป็นในการตอบข้อสอบในแต่ละข้อได้ถูกต้องและไม่ถูกต้อง การตรวจสอบคุณลักษณะของข้อสอบแต่ละข้อตามข้อตกลงนี้ อาจทำได้โดยการตรวจสอบแบบแผนของการตอบข้อสอบของผู้สอบที่มีความสามารถเท่า ๆ กันว่า ข้อสอบแต่ละคู่ เช่น ข้อ 1 กับข้อ 2 มีลักษณะการตอบแตกต่างกันหรือไม่ ซึ่งสามารถทดสอบได้โดยใช้สถิติการทดสอบไคสแควร์ (Chi-Square) (สำเริง บุญเรืองรัตน์. 2529 : 46) อย่างไรก็ตาม แฮมเบิลตัน และ สวามินาธาน (Hambleton and Swaminathan. 1985 : 24) ได้ให้ความเห็นว่า ถ้าแบบทดสอบมีความเป็นมิติเดียวแล้ว แบบทดสอบจะมีความเป็นอิสระในการตอบข้อสอบด้วย

3. โค้งลักษณะข้อสอบ (Item Characteristic Curve : ICC) เป็นข้อตกลงเกี่ยวกับโค้งลักษณะข้อสอบ กล่าวคือ โค้งลักษณะข้อสอบเป็นกราฟของฟังก์ชันทางคณิตศาสตร์ที่แสดงถึงความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบข้อนั้นถูก กับความสามารถหรือลักษณะที่วัดโดยข้อสอบข้อนั้น โค้งลักษณะข้อสอบมีหลายรูปแบบทั้งนี้ขึ้นอยู่กับแบบจำลองที่ใช้อธิบายความสัมพันธ์ดังกล่าว

2.4.3 แบบจำลองของทฤษฎีการตอบสนองข้อสอบ

เนื่องจากข้อมูลที่ได้จากผู้สอบมีหลายลักษณะ ฉะนั้น จึงได้มีการพัฒนาโมเดลของความสัมพันธ์ในลักษณะต่าง ๆ กัน เพื่อให้แต่ละโมเดลมีความเหมาะสมกับลักษณะของข้อมูลที่ได้ (สำเริง บุญเรืองรัตน์. 2529 : 3) ทั้งนี้แบบจำลองแต่ละแบบจะต่างกันในรูปแบบของฟังก์ชันทางคณิตศาสตร์และจำนวนพารามิเตอร์ที่กล่าวถึง โดยสามารถจำแนกออกได้เป็นกลุ่มใหญ่ 3 กลุ่ม คือ ค่าตอบที่เป็นข้อมูลแบบแบ่งสอง (Dichotomous) ข้อมูลแบบพหุ (Multi-

chotomous) และข้อมูลแบบต่อเนื่อง (Continuous) (Hambleton and Swaminathan. 1985 : 35 - 52)

สำหรับแบบจำลองซึ่งเป็นที่ยอมรับกันมากในระะหลัง ซึ่งใช้กับการให้คะแนนแบบ 2 คำ คือ แบบจำลองปกติอวโจว และแบบจำลองโลจิสติก

แบบจำลองปกติอวโจว (Normal Ogive Model)

แบบจำลองนี้ได้รับการพัฒนาขึ้นโดย ลอเลย์ (Lawley) ในปี ค.ศ. 1944 และ ลอร์ดได้ทำการศึกษาและปรับปรุงแบบจำลองนี้ พร้อมทั้งพัฒนาวิธีการประมาณค่าพารามิเตอร์ของแบบจำลอง โดยนำไปใช้กับข้อมูลที่ได้จากแบบทดสอบวัดผลสัมฤทธิ์และแบบทดสอบความถนัด แบบจำลองนี้เชื่อว่า ความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบถูกต้องกับระดับความสามารถจะอยู่ในรูปฟังก์ชันการแจกแจงปกติสะสม (Cumulative Normal Ogive)

แบบจำลองโลจิสติก (Logistic Model)

แบบจำลองนี้ถูกพัฒนาขึ้นเพื่อให้เหมาะสมกับสถานการณ์การทดสอบยิ่งขึ้น เช่น กรณีมีการเดาเกิดขึ้น หรือ กรณีผู้สอบทำข้อสอบด้วยความสะเพร่า แบบจำลองนี้แบ่งเป็น 4 แบบย่อย ตามพารามิเตอร์ ดังนี้ (Hambleton and Swaminathan. 1985 : 34 - 49)

1. แบบจำลองที่ใช้พารามิเตอร์ตัวเดียว (One-Parameter Logistic Model) หรือราสช์โมเดล (Rasch Model) แบบจำลองนี้ถือว่าข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากัน ($a_1 = a$) และไม่มีการเดา ($c_1 = 0$) ก็บการจะมีค่าพารามิเตอร์ความยาก (b_1) เพียงตัวเดียว

2. แบบจำลองพารามิเตอร์ 2 ตัว (Two-Parameter Logistic Model) เบิร์นบอม (Birnbau) ได้เสนอแบบจำลองนี้ในปี ค.ศ. 1957 และพัฒนาโดยใช้พารามิเตอร์ 2 ตัว คือ ค่าความยาก (b_1) และค่าอำนาจจำแนก (a_1) โดยถือว่า ผู้สอบมีโอกาสตอบถูกโดยการเดาน้อยมาก โดยกำหนดค่าพารามิเตอร์การเดา (c_1) = 0

3. แบบจำลองพารามิเตอร์ 3 ตัว (Three-Parameter Logistic Model) เป็นแบบจำลองที่พัฒนามาจากแบบจำลองที่ใช้พารามิเตอร์ 2 ตัว เพื่อให้เหมาะกับแบบทดสอบที่มีอิทธิพลของการเดาเข้ามาเกี่ยวข้อง โดยเพิ่มพารามิเตอร์ คือ ค่าการเดา เข้าไป

4. แบบจำลองพารามิเตอร์ 4 ตัว (Four-Parameter Logistic Model) แบบจำลองนี้ผู้นำเสนอ คือ แม็คโดนัลด์ (McDonal) ในปี ค.ศ. 1967 ลอร์ด (Lord)

และบาร์ตัน (Barton) ได้พัฒนาต่อในปี ค.ศ. 1981 โดยเพิ่มพารามิเตอร์ 3 ตัว ซึ่งเป็นค่าที่แสดงให้เห็นถึงกรณีที่ผู้สอบที่มีความสามารถสูงตอบข้อสอบผิดอันอาจเนื่องมาจากความไม่รอบคอบ ทำให้ผู้นั้นไม่เลือกคำตอบที่ถูกต้อง

2.5 ดัชนีกำหนดคุณภาพและการเปรียบเทียบคุณภาพของแบบทดสอบ

ทฤษฎีการตอบสนองข้อสอบพิจารณาคุณภาพของแบบทดสอบจากดัชนีที่ใช้กำหนดคุณภาพ คือ ค่าฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Function) ซึ่งเป็นฟังก์ชันของค่าพารามิเตอร์ของข้อสอบที่เป็นดัชนีผสม ที่สร้างจากดัชนีที่ใช้กำหนดคุณภาพของข้อสอบหลาย ๆ ดัชนีของทุก ๆ ข้อในแบบทดสอบ โดยกำหนดว่า ณ ระดับความสามารถใดที่ค่าฟังก์ชันสารสนเทศของแบบทดสอบมีค่าสูง แสดงว่า ณ ความสามารถระดับนั้นแบบทดสอบสามารถใช้ได้อย่างมีประสิทธิภาพ หรืออาจกล่าวได้ว่าคุณภาพของแบบทดสอบสามารถพิจารณาได้จากความเชื่อถือได้ของการประมาณค่าความสามารถที่แท้จริง จากฟังก์ชันสารสนเทศของแบบทดสอบ ซึ่งเป็นผลรวมของฟังก์ชันสารสนเทศของข้อสอบ (Item Information Function) แต่ละข้อ (Lord, 1980 : 72)

2.5.1 ฟังก์ชันสารสนเทศของข้อสอบ (Item Information Function)

ฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อกำหนดได้ดังนี้

$$I\{\theta, U_i\} = (P_i^2) / P_i Q_i$$

เมื่อ $I\{\theta, U_i\}$ คือ ฟังก์ชันสารสนเทศของข้อสอบข้อที่ i
 P_i คือ ความน่าจะเป็นที่ผู้สอบคนหนึ่ง ณ ระดับความสามารถ θ จะตอบข้อสอบ ข้อที่ i ได้ถูกต้อง
 Q_i คือ $1 - P_i$
 P_i คือ ค่าอนุพันธ์ (Derivative) ของโอกาสในการตอบข้อสอบข้อที่ i ถูก หรือความชันของโค้งลักษณะข้อสอบ (ICC) ณ ระดับความสามารถ

ฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อจะรวมกันเป็นฟังก์ชันสารสนเทศของแบบทดสอบ ซึ่งค่าฟังก์ชันสารสนเทศของข้อสอบจะขึ้นอยู่กับความลาดชันของโค้งลักษณะข้อสอบแต่ละข้อ หรือค่าความชันของฟังก์ชันการตอบข้อคำถาม (Item Response Function) และความแปรปรวนของคะแนนจากข้อสอบแต่ละระดับความสามารถ (๑) ยิ่งค่าความชันมากหรือโค้งลักษณะข้อสอบชันมาก ๆ ในขณะที่ยังค่าความแปรปรวนต่ำ จะทำให้ค่าสารสนเทศของข้อสอบมีค่าสูง ซึ่งจะทำให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดมีค่าต่ำ

2.5.2 ฟังก์ชันสารสนเทศของแบบทดสอบ

ฟังก์ชันสารสนเทศของแบบทดสอบ เป็นสัดส่วนกลับของความแปรปรวนอย่างสุ่มของการประมาณค่าความสามารถของผู้สอบแต่ละคน แบบความเป็นไปได้สูงสุดเมื่อเทียบกับความสามารถรวมของผู้เข้าสอบทุกคน (Hulin and others. 1983 : 55) นั้นหมายความว่า เมื่อจำนวนข้อสอบเพิ่มขึ้น $1/\sqrt{I(\theta)}$ ก็จะเป็นความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ

ฟังก์ชันสารสนเทศของแบบทดสอบสามารถเขียนเป็นสมการได้ ดังนี้ (Hambleton and Swaminathan. 1985 : 104)

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

ลักษณะฟังก์ชันสารสนเทศของแบบทดสอบสามารถสรุปได้ ดังต่อไปนี้

1. ฟังก์ชันสารสนเทศของแบบทดสอบ เป็นสิ่งที่กำหนดขึ้นสำหรับชุดของข้อสอบที่แต่ละจุดของสเกลความสามารถ
2. ผลรวมของสารสนเทศเป็นผลมาจากคุณภาพและจำนวนของข้อสอบ
3. ที่ความสามารถเดียวกันเส้นถดถอยที่มีความชันมากกว่าจะให้ค่าสารสนเทศของแบบทดสอบสูงกว่าเส้นถดถอยที่มีความชันน้อยกว่า
4. ค่าความแปรปรวนของข้อสอบที่ต่ำจะให้ค่าสารสนเทศของแบบทดสอบสูง
5. ค่าสารสนเทศของแบบทดสอบไม่ขึ้นอยู่กับการจัดหมู่เฉพาะของข้อสอบ การสนับสนุนของข้อสอบแต่ละข้อก็เป็นอิสระจากกัน
6. ค่าสารสนเทศของแบบทดสอบ มีความสัมพันธ์แบบผกผันกับค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ ณ จุดเดียวกัน นั่นคือ

$$SE(e) = 1/\sqrt{I(e)}$$

เมื่อ $SE(e)$ คือ ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความ
สามารถ ณ ระดับนั้น

2.5.3 ประสิทธิภาพสัมพัทธ์ (Relative Efficiency)

ค่าประสิทธิภาพสัมพัทธ์เป็นดัชนีหนึ่งที่ใช้ในการศึกษาคุณภาพของแบบทดสอบ ในการ
ที่ต้องการเปรียบเทียบแบบทดสอบ 2 ฉบับ ที่วัดคุณลักษณะเดียวกันว่า ฉบับใดจะมีประสิทธิภาพ
สูงกว่ากัน ค่าประสิทธิภาพสัมพัทธ์ได้จากอัตราส่วนของฟังก์ชันสารสนเทศของคะแนนจากแบบ
ทดสอบ 2 ฉบับ ที่ต้องการเปรียบเทียบ เช่น ต้องการเปรียบเทียบคะแนนจากแบบทดสอบ
y และ x สามารถเขียนอยู่ในรูปของสมการดังนี้ (Lord, 1980 : 83)

$$RE(y, x) = I(e, y) / I(e, x)$$

คะแนน x และ y อาจเป็นคะแนนจากแบบทดสอบ 2 ฉบับ ที่ต่างกัน แต่วัด
ความสามารถเดียวกันแต่ให้คะแนนด้วยวิธีการที่ต่างกัน ค่าประสิทธิภาพสัมพัทธ์ของคะแนน
2 ชุดนี้ จะแปรเปลี่ยนไปในแต่ละระดับความสามารถ

การแปลความหมายของค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ณ ระดับความสามารถ
ใด สามารถแปลได้จากการพิจารณาค่าประสิทธิภาพสัมพัทธ์ได้ดังนี้ (Warm, 1978 : 76)

กรณีค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบมีค่ามากกว่า 1 แสดงว่า แบบทดสอบ y
มีคุณภาพสูงกว่าแบบทดสอบ x ณ ระดับความสามารถนั้น

ถ้าค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบมีค่าน้อยกว่า 1 แสดงว่า แบบทดสอบ y
มีคุณภาพต่ำกว่าแบบทดสอบ x ณ ระดับความสามารถนั้น และ

ถ้าค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบมีค่าเท่ากับ 1 แสดงว่า แบบทดสอบ y
และแบบทดสอบ x มีคุณภาพเท่ากัน ณ ระดับความสามารถนั้น

บทที่ 3

วิธีดำเนินการ

การศึกษาเกี่ยวกับการพัฒนาการวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ของนักเรียน ระดับชั้นมัธยมศึกษา ครึ่งนี้ มีจุดมุ่งหมายที่จะพัฒนาข้อคำถามที่สามารถนำไปใช้ประโยชน์ในด้านการศึกษา และพัฒนาความสามารถด้านเหตุผลของผู้เรียนได้อย่างมีประสิทธิภาพกว่ารูปแบบที่ใช้กันอยู่ โดยมีรายละเอียดของวิธีดำเนินการในแต่ละขั้นตอน ดังนี้

ประชากร

ประชากรในการวิจัยครั้งนี้ เป็นนักเรียนระดับชั้นมัธยมศึกษาตอนต้น (ม.3) และมัธยมศึกษาตอนปลาย (ม.5) ปีการศึกษา 2538 ของโรงเรียนมัธยมศึกษาในจังหวัดนครราชสีมา นครปฐม และกรุงเทพมหานคร

กลุ่มตัวอย่าง

กลุ่มตัวอย่างในการวิจัยครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 และ 5 ปีการศึกษา 2538 ของโรงเรียนมัธยมศึกษาในจังหวัดนครราชสีมา นครปฐม และกรุงเทพมหานคร ซึ่งเลือกมาจากกลุ่มประชากรโดยการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) ตามขนาดของโรงเรียน (ขนาดกลาง และใหญ่) มีโรงเรียนเป็นหน่วยการสุ่ม จากนั้นทำการสุ่มอีกโดยการสุ่มอย่างง่าย (Simple Random Sampling) โดยมีห้องเรียนเป็นหน่วยการสุ่ม ทั้งนี้จะใช้นักเรียนทุกคนในห้องเรียนเป็นกลุ่มตัวอย่างในการทดลองเครื่องมือ กลุ่มตัวอย่างแบ่งเป็น 2 กลุ่ม ตามกระบวนการศึกษา คือ

1. กลุ่มตัวอย่างที่ใช้ในการทดลองเครื่องมือวัดในกระบวนการพัฒนาข้อคำถาม
2. กลุ่มตัวอย่างที่ใช้ในการศึกษาคุณภาพของข้อคำถามและแบบทดสอบ

สำหรับขนาดของกลุ่มตัวอย่างที่เพียงพอที่จะเป็นตัวแทนของประชากรได้นั้น กำหนดโดยตารางกำหนดขนาดกลุ่มตัวอย่างของ ฮินเคิล และคนอื่น ๆ (Hinkle and others. 1985 : 275) เมื่อให้ α เท่ากับ 0.05 อำนาจการทดสอบทางสถิติ (Power of Statistical Test) เท่ากับ 0.95 และขนาดอิทธิพล (Effect Size) เท่ากับ 0.10 ได้ขนาดกลุ่มตัวอย่างเท่ากับ 1,300 คน ดังนั้น ในแต่ละกระบวนการของการศึกษาจะใช้กลุ่มตัวอย่างนั้นละประมาณ 1,300 คน รายละเอียดดังแสดงในตาราง 1 และ 2

ตาราง 1 รายชื่อโรงเรียน และจำนวนนักเรียนที่เป็นกลุ่มตัวอย่างในการศึกษาในกระบวนการพัฒนาข้อคำถาม (ทดลองครั้งที่ 1)

ชื่อโรงเรียน	จำนวนนักเรียน	
	มัธยมศึกษาปีที่ 3	มัธยมศึกษาปีที่ 5
1. วัดบวรนิเวศ	33	55
2. วัดสุทัศนวราราม	211	210
3. ราชวินิตมัธยม	217	218
4. สตรีศรีสุวิทย์	273	255
5. สตรีวิทยา	324	197
6. โปธิสารพิทยากร	192	95
7. กรุงเทพมหานครคริสเตียนวิทยาลัย	336	214
8. เขมะสิริอนุสสรณ์	93	96
รวม	1,679	1,275
รวมทั้งสิ้น	2,954	

ตาราง 2 รายชื่อโรงเรียน และจำนวนนักเรียนที่เป็นกลุ่มตัวอย่างในกระบวนการศึกษา
คุณภาพของเครื่องมือ (ทดลองครั้งที่ 2)

ชื่อโรงเรียน	จำนวนนักเรียน	
	มัธยมศึกษาปีที่ 3	มัธยมศึกษาปีที่ 5
1. เสสะเวชวิทยา	66	44
2. อัสสัมชัญคอนแวนต์	156	112
3. สาธิต มศว ประสานมิตร	-	48
4. สวนกุหลาบวิทยาลัย	276	111
5. เบญจมาธาลัย	236	-
6. มัธยมวัดมกุฎกษัตริย์	170	141
7. สันติราษฎร์วิทยาลัย	148	124
8. โยธินบูรณะ	87	117
9. สายน้ำผึ้ง	167	116
10. มหรรณพาราม	143	81
11. สุวรรณาราม	26	-
12. สীগัน (วัดนันทน์)	130	119
13. บดินทร์เคชา 2	162	142
14. วัดนาสโรง	69	60
15. ศิลาจารนิพัฒน์	42	36
16. ศรีอยุธยา	129	74
17. เตรียมอุดมศึกษา	-	233
18. ราชนันบูรณะ (นครปฐม)	304	297
19. โคราชพิทยาคม (นครราชสีมา)	287	237
รวม	2,590	2,092
รวมทั้งสิ้น	4,682	

เครื่องมือที่พัฒนาขึ้นเพื่อการศึกษา

จุดมุ่งหมายของการศึกษาครั้งนี้ เพื่อพัฒนาการวัดในส่วนของการพัฒนาเครื่องมือวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ของนักเรียน ระดับชั้นมัธยมศึกษา ทั้งนี้ข้อคำถามที่สร้างขึ้น 2 รูปแบบ คือ แบบเลือกตอบ 5 ตัวเลือก และแบบเติมคำตอบ แต่ละรูปแบบข้อคำถามจะสร้างเป็น 6 ชนิด ได้แก่

1. เหตุผลด้านวิเคราะห์ (Analytical Reasoning : AR)
2. การวิเคราะห์คำอธิบาย (Analysis of Explanation : AX)
3. ความเห็นตรงกันข้าม (Contrasting Views : CV) ✓
4. เหตุผลทางตรรก (Logical Reasoning : LR.)
5. เหตุผลทางตรรกเกี่ยวกับจำนวน (Numerical Logical Reasoning :

NLR)

6. การระบุรูปแบบ (Pattern Identification : PI)

การตรวจให้คะแนนใช้ระบบ 0, 1

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลทั้ง 2 ขั้นตอน คือ ขั้นการดำเนินการพัฒนาข้อคำถาม และขั้นการศึกษาคุณภาพของเครื่องมือการศึกษาครั้งนี้ ผู้วิจัยจะทำการทดสอบตามระบบห้องเรียนปกติของโรงเรียน ซึ่งเป็นกลุ่มตัวอย่างที่ได้จากการสุ่ม และเนื่องจากแบบทดสอบมีหลายฉบับ การกำหนดแบบทดสอบให้แก่นักเรียนที่เป็นกลุ่มตัวอย่าง จึงใช้วิธีการ Random Assignment

หลักการและการดำเนินการสร้างเครื่องมือวัด

การดำเนินการสร้างเครื่องมือวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ครั้งนี้ ผู้วิจัยมีหลักการและความเชื่อพื้นฐาน ดังนี้

1. ความสามารถด้านเหตุผลเชิงวิเคราะห์ ถือเป็นความสามารถของบุคคลที่สามารถพัฒนาได้ (Developed Ability) อันเป็นผลเนื่องมาจากกระบวนการเรียนรู้ การอบรม และปลูกฝัง

2. ความสามารถด้านเหตุผลเชิงวิเคราะห์ ถือเป็นความสามารถที่เกิดขึ้นกับผู้เรียนอันเป็นผลเนื่องมาจากกระบวนการจัดการเรียนการสอนในรายวิชาต่าง ๆ ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) และหลักสูตร

มัธยมศึกษาตอนปลาย พุทธศักราช 2524 (ฉบับปรับปรุง พ.ศ. 2533) ซึ่งมุ่งเน้นให้ผู้เรียนเป็นผู้มีปัญญา รู้จักเหตุและผล รู้จักแยกแยะผิดชอบชั่วดี

3. เครื่องมือวัดที่พัฒนาขึ้นเพื่อใช้วัดความสามารถด้านเหตุผลเชิงวิเคราะห์ มีขอบข่ายการวัดโครงสร้าง (Construct) ที่สอดคล้องกับกรอบความคิดเกี่ยวกับความสามารถด้านเหตุผลเชิงวิเคราะห์ที่ผู้วิจัยคิดเป็นแนวทางหลัก ทั้งนี้ผู้วิจัยมีความเชื่อว่าชนิดของข้อคำถามที่พัฒนาขึ้นนี้ จะสามารถวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ได้อย่างครอบคลุมและมั่นใจมากยิ่งขึ้น

4. กำหนดเวลาที่ใช้ในการสอบแบบทดสอบแต่ละฉบับ ถือเป็นระยะเวลาเฉลี่ยที่ผู้เรียนในช่วงอายุ และระดับการศึกษานี้ ใช้ในการทำแบบทดสอบแต่ละฉบับได้เสร็จสิ้นอย่างสมบูรณ์

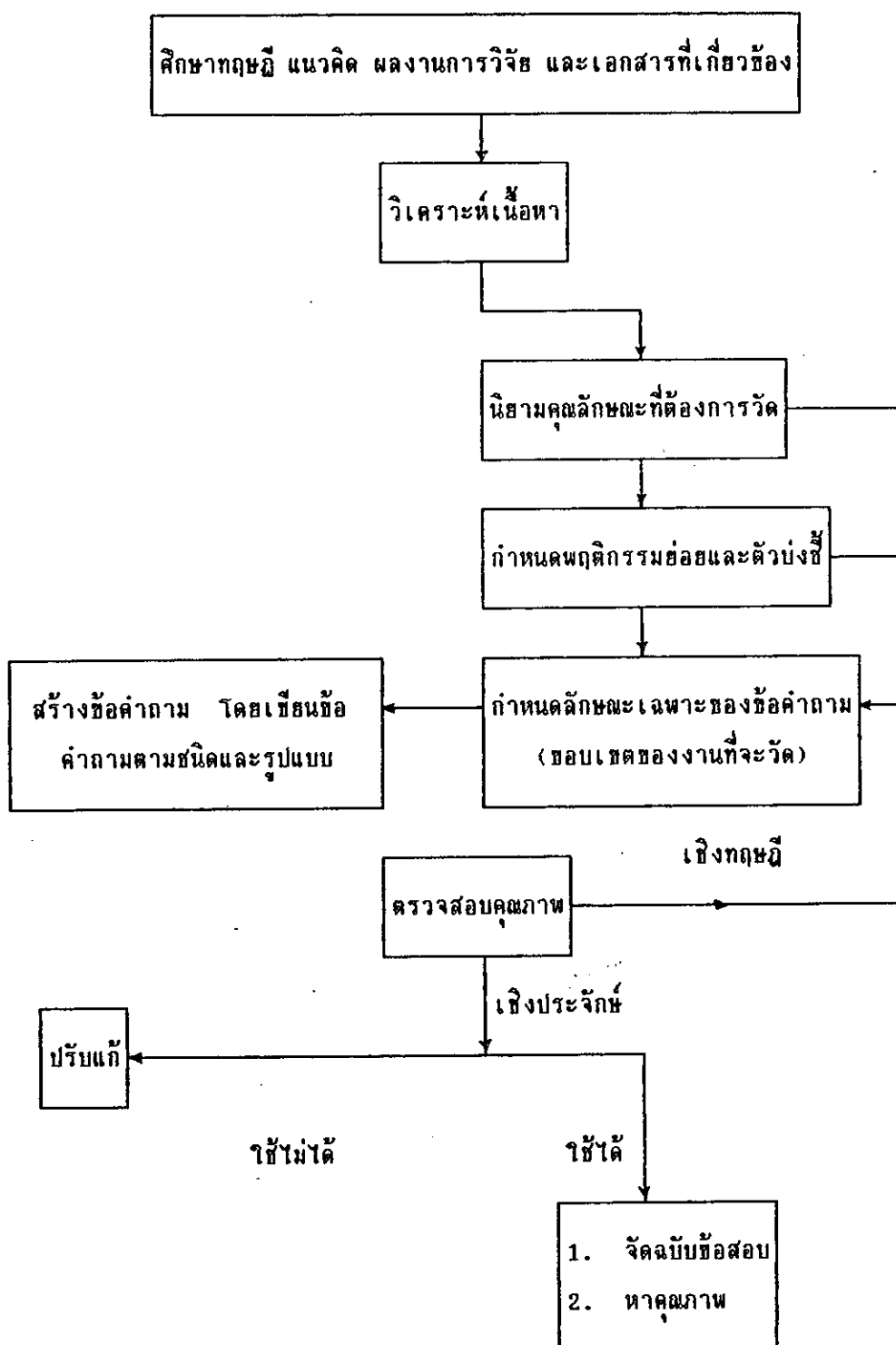
5. กลุ่มประชากรเป้าหมาย ซึ่งเป็นนักเรียนระดับมัธยมศึกษาตอนต้น และตอนปลายนี้ ถือเป็นกลุ่มช่วงอายุที่มีพัฒนาการด้านการคิดอย่างมีเหตุผลสมบูรณ์ตามแนวคิดของเพียเจต์

เนื่องจากการวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อพัฒนาเครื่องมือวัด และศึกษาคุณภาพของเครื่องมือที่พัฒนาขึ้น ฉะนั้น กระบวนการดำเนินการศึกษาจึงแบ่งเป็น 2 ระยะ คือ

1. การดำเนินการพัฒนาเครื่องมือวัด
2. การดำเนินการเพื่อศึกษาคุณภาพของข้อคำถามและแบบทดสอบที่พัฒนาขึ้น
3. กระบวนการดำเนินงานแต่ละระยะ มีดังนี้

3.1 การดำเนินการพัฒนาเครื่องมือวัด จากหลักการและความเชื่อพื้นฐานดังกล่าวข้างต้น ผู้วิจัยได้ดำเนินการสร้างเครื่องมือตามกรอบความคิด และรายละเอียดดังภาพประกอบ 12 ต่อไปนี้

ขั้นตอนการพัฒนาแบบทดสอบ



ภาพประกอบ 12 ขั้นตอนการพัฒนาแบบทดสอบ

1. ศึกษาทฤษฎี แนวคิด งานวิจัย และเอกสารที่เกี่ยวข้องกับการคิดความสามารถ ด้านเหตุผลและการวัดเพื่อนิยามโครงสร้าง (Construct) และกำหนดทักษะซึ่งเป็นคุณสมบัติ (Attribute) ที่สำคัญของโครงสร้างที่ต้องการวัด

2. นิยามโครงสร้างที่ต้องการวัด ในที่นี้ คือ ความสามารถด้านเหตุผลเชิงวิเคราะห์ ซึ่งนิยามได้ว่า หมายถึง ความสามารถในการคิดพิจารณาองความเห็นเกี่ยวกับ เหตุการณ์ เรื่องราว หรือสิ่งของต่าง ๆ โดยการค้นหาทำความเข้าใจ จากข้อมูลหลักฐานที่มี อยู่ เพื่อนำไปสู่การประเมิน วินิจฉัย และสร้างข้อสรุปที่ถูกต้องสอดคล้องกับหลักการและ เหตุผล

3. วิเคราะห์โครงสร้างที่ต้องการวัดเพื่อกำหนดทักษะที่เป็นคุณสมบัติสำคัญ และตัว ชี้บ่ง (Indicator) ทั้งนี้ประมวลสรุปได้ดังภาพประกอบ 13

ทนาย

ทักษะย่อย	พฤติกรรมชี้บ่ง
1. ความสามารถในการสร้าง และใช้ความคิดรวบยอด	1. ระบุตัวอย่าง หลักฐาน 2. ระบุลักษณะสำคัญ 2.1 การจัดกลุ่ม 2.2 การหาหลักการ 2.3 การหาลักษณะที่คล้ายคลึง
2. ความสามารถในการมองเห็นความสัมพันธ์ และโยงความสัมพันธ์	1. มีความรู้ ความเข้าใจระบบและความหมาย 2. จำแนกความเหมือน - ต่าง 3. ใช้หลักการของเหตุ - ผล 4. การระบุความสำคัญหลักฐาน 5. คาดเดาเกี่ยวกับสิ่งที่เกิดขึ้น

ทักษะย่อย	พฤติกรรมที่พึง
3. ความสามารถใช้เกณฑ์ในการพิจารณา	1. ความถูกต้องแม่นยำ 2. ความจริง 3. ความคงเส้นคงวา 4. ความสมบูรณ์ 5. ความสอดคล้องเกะวข้องอย่าง สมเหตุสมผล 6. ความน่าเชื่อถือ/เชื่อถือได้ 7. ความเป็นไปได้ 8. ☒ ความมีเหตุผล 9. ความมีประสิทธิภาพที่ดีที่สุด
4. ความสามารถในการสร้างข้อสรุป	1. การคาดคะเน หรือสร้าง สมมุติฐาน 2. การสร้างข้อตกลงเบื้องต้น 3. การเชื่อมโยงความคิด 4. ☒ การลงความเห็น 5. การสร้างข้อเสนอ
5. ความสามารถในการตัดสินใจ และประเมินปัญหา	1. การเปรียบเทียบคุณค่า 2. การวินิจฉัยคุณค่า 3. เลือกตัดสินใจ 4. ☒ เลือกทางเลือกในการแก้ปัญหา ที่ดีที่สุด

4. ตรวจสอบความเที่ยงตรงของโครงสร้างที่ต้องการวัด (Validity of Construct) โดยผู้เชี่ยวชาญ ทั้งนี้เพราะความเที่ยงตรงเชิงโครงสร้างนี้ โอเกน (1991) กรอนลินด์ และลินน์ (1990) มีความเห็นสอดคล้องกันว่าสามารถตรวจสอบได้จากความเห็นของผู้รู้หรือผู้เชี่ยวชาญ เนื่องจากถือว่าเป็นกระบวนการเชิงเหตุผล

วิธีการตรวจสอบกระทำโดยให้ผู้เชี่ยวชาญจำนวน 5 ท่าน พิจารณานิยามโครงสร้างความสามารถด้านเหตุผลเชิงวิเคราะห์ที่มีความชัดเจน ถูกต้อง เพียงใด จากนั้นตรวจสอบความสอดคล้องความครอบคลุมของทักษะย่อยกับนิยามโครงสร้าง เพื่อให้เกิดความมั่นใจว่าทักษะย่อยที่กำหนดแสดงโครงสร้างที่ต้องการวัดได้ตรงและครอบคลุม รวมทั้งตรวจสอบความสอดคล้องถูกต้องของพฤติกรรมย่อยของแต่ละทักษะย่อย

5. กำหนดลักษณะงาน ขอบเขตงาน ซึ่งคือ ชนิดของข้อคำถาม (Item Type) ที่จะสามารถกระตุ้น หรือสะท้อนพฤติกรรมย่อยที่ต้องการได้ จากนั้นทำการตรวจสอบเชิงเหตุผลโดยผู้เชี่ยวชาญจำนวน 3 ท่าน พิจารณาปรับแก้

6. สร้างข้อคำถามตามชนิด (Item Type) และรูปแบบ (Item Format) ที่กำหนด คือ การสร้างข้อคำถาม 2 รูปแบบ คือ แบบเลือกตอบ 5 ตัวเลือก และแบบเติมคำตอบ ซึ่งทั้ง 2 รูปแบบ ประกอบด้วยข้อคำถาม 6 ชนิด คือ

6.1 เหตุผลด้านวิเคราะห์ (Analytical Reasoning : AR)

6.2 การวิเคราะห์คำอธิบาย (Analysis of Explanation : AX)

6.3 ความเห็นตรงกันข้าม (Contrasting Views : CV)

6.4 เหตุผลเชิงตรรก (Logical Reasoning : LR)

6.5 เหตุผลเชิงตรรกเกี่ยวกับจำนวน (Numerical Logical Reasoning

: NLR)

6.6 การระบุรูปแบบ (Pattern Identification : PI)

7. ตรวจสอบความเที่ยงตรงของข้อคำถามที่สร้างในเชิงตรรก (Logical Analysis) โดยพิจารณาความสอดคล้องของข้อคำถามที่สร้างขึ้นกับขอบเขตของงานที่กำหนดจากผู้เชี่ยวชาญ 8 ท่าน ดังภาพประกอบ 14

การตรวจสอบความเที่ยงตรง	ประเด็นพิจารณา	วิธีการ
1. โครงสร้างที่ต้องการวัด	พิจารณาความชัดเจน ถูกต้อง ของ นิยามโครงสร้างความสามารถด้าน เหตุผลเชิงวิเคราะห์	ผู้เชี่ยวชาญพิจารณา
2. นิยามโครงสร้าง : ทักษะย่อย	สามารถแสดงโครงสร้างได้ตรงและ ครอบคลุม	ผู้เชี่ยวชาญพิจารณา
3. ทักษะย่อย : พฤติกรรมชี้บ่ง	ความเป็นตัวแทนสำคัญของพฤติกรรม การแสดงออก	ผู้เชี่ยวชาญพิจารณา
4. พฤติกรรมชี้บ่ง : งาน (Task) (ชนิดข้อ คำถาม)	ความเป็นตัวแทนของงานที่จะ สามารถกระตุ้น หรือสะท้อนพฤติกรรม ชี้บ่งที่ต้องการได้	ผู้เชี่ยวชาญพิจารณา
5. งาน : ข้อคำถาม ที่ต้องการ	ความครอบคลุมสอดคล้องของตัวแทน ของงานที่ต้องการ	ผู้เชี่ยวชาญพิจารณา
6. โครงสร้าง : ข้อคำถาม	ความเป็นมิติเดียวของข้อคำถาม ความเที่ยงตรงของข้อคำถาม	วิเคราะห์องค์- ประกอบ หาความ สัมพันธ์ระหว่างข้อ กับคะแนนรวมทั้งฉบับ

ภาพประกอบ 14 การตรวจสอบความเที่ยงตรงของเครื่องมือวัดที่พัฒนาขึ้นตามขั้นตอนต่าง ๆ

8. ทำการวิเคราะห์เชิงประจักษ์ (Empirical Analysis) โดยจัดรูปแบบทดสอบ นำไปทดลองในลักษณะนำร่อง (Pilot Study) กับกลุ่มตัวอย่างซึ่งเป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 โรงเรียนมัธยมสาธิต มศว ประสานมิตร จำนวนประมาณ 50 คน ทั้งนี้เพื่อศึกษาความเป็นไปได้ในการตอบสนองของผู้สอบ ความชัดเจนของข้อสอบ สถานการณ์ที่กำหนด คำชี้แจง และเวลาที่ต้องใช้ในการตอบข้อคำถาม รวมทั้งพิจารณาความเที่ยงตรงของข้อคำถาม

9. นำข้อคำถามที่ปรับแก้แล้ว จัดฉบับไปทดสอบกับกลุ่มตัวอย่างของประชากรกลุ่มเป้าหมาย

10. ตรวจให้คะแนนโดยใช้ระบบ 0, 1 ทั้ง 2 รูปแบบ กรณีที่ข้อสอบเป็นแบบเลือกตอบ ตรวจให้คะแนน 1 สำหรับคำตอบซึ่งเป็นตัวเลือกที่ถูกต้องเหมาะสม หรือสอดคล้องกับเงื่อนไขที่กำหนดมากที่สุด ส่วนแบบเติมคำตอบ (Supply Type) นั้น

เกณฑ์การพิจารณาให้คะแนนสำหรับข้อคำถามรูปแบบเติมคำตอบ คำตอบที่จะได้คะแนน 1 ต้องเป็นคำตอบที่แสดงความสอดคล้องถูกต้องตามเงื่อนไข หรือลงความเห็นอย่างสมเหตุสมผล

11. วิเคราะห์ข้อมูลเพื่อหาคุณภาพของข้อคำถามและแบบทดสอบด้วยทฤษฎีวิเคราะห์แบบดั้งเดิม (Classical Test Theory) หาค่าความยาก (p) ค่าอำนาจจำแนก (r) และค่าความเชื่อมั่นของแบบทดสอบ (Reliability)

12. พิจารณารับปรุงแก้ไขโดยให้ข้อมูลจากการดำเนินการตามข้อ 11 เป็นพื้นฐานในการพิจารณา

ตัวอย่างเครื่องมือ

ข้อคำถามแต่ละชนิด

ผู้วิจัยขอเสนอลักษณะของข้อคำถามและตัวอย่างของข้อคำถาม ทั้ง 6 ชนิด ที่ได้พัฒนาขึ้นเป็น 2 รูปแบบ คือ แบบเติมคำตอบ และแบบเลือกตอบ

1. เหตุผลด้านวิเคราะห์ (AR)

ข้อคำถามชนิดนี้มุ่งประเมินทักษะในการจับหลัก และกฎ เพื่อนำไปสู่ข้อสรุป โดยพิจารณาจากเงื่อนไขที่ให้

ทักษะย่อยที่ถือว่าข้อคำถามชนิดนี้จะสามารถวัดได้ ได้แก่ ทักษะย่อยที่ 1, 3, 4 และ 5

AR : แบบเติมคำตอบ

คำชี้แจง : ข้อคำถามในแบบทดสอบฉบับนี้จะนำเสนอสถานการณ์สั้น ๆ พร้อมกับกลุ่มของเงื่อนไข หรือกฎเกณฑ์ ซึ่งนักเรียนจะต้องนำมาพิจารณาวิเคราะห์ประกอบกัน เพื่อตอบคำถามจากสถานการณ์และเงื่อนไขที่กำหนดให้ ทั้งนี้โดยเขียนคำตอบที่นักเรียนพิจารณาเห็นว่าถูกต้องเหมาะสมที่สุด ลงในส่วนที่เป็นคำตอบซึ่งเว้นว่างไว้ของแต่ละข้อ (ตอบเพียงคำตอบเดียว)

ตัวอย่างข้อคำถาม

(0) โตเป็นเจ้าภาพจัดงานเลี้ยงอาหารเย็นให้เพื่อนสนิท 5 คน คือ มุก แยก แก้ว เป้า และป๊ม โตและเพื่อน ๆ จะต้องนั่งรับประทานอาหารที่โต๊ะกลม โตรอบโต๊ะมีเก้าอี้ 6 ที่

เงื่อนไขที่กำหนดคือ :-

- เป้าจะต้องไม่นั่งข้างป๊ม
- แยกต้องมีมุกนั่งติดด้านขวาของเขา

ใช้สถานการณ์ที่กำหนดให้ตอบคำถามข้อ (0) - (00)

คำถาม :

(0) การจัดที่นั่งตามลำดับอย่างไรที่เป็นไปไม่ได้

คำตอบ : โต, แก้ว, ป๊ม, มุก, เป้า, แยก

เมื่อพิจารณาจากข้อกำหนดเงื่อนไข และสถานการณ์ที่กำหนดแล้ว การจัดที่นั่งที่เป็นไปไม่ได้ คือ การจัดเป้าให้นั่งติดจากป้อม ประการหนึ่ง และจัดให้มุกนั่งที่อื่นที่มีใช้ด้านขวาของเอก ซึ่งเป็นการจัดที่จะเมิดเงื่อนไขที่กำหนด ฉะนั้น การจัดที่เป็นไปไม่ได้ ได้แก่

1. โต, แก้ว, ป้อม, มุก, เป้า, เอก
2. โต, แก้ว, ป้อม, เป้า, มุก, เอก

(ตอบเพียงคำตอบเดียว)

คำถาม :

(00) ถ้าแก้วนั่งทางด้านซ้ายของเอก การนั่งอย่างใดจะต้องเป็นจริง

คำตอบ : เป้านั่งข้างโต

การพิจารณาคำตอบของข้อ (00) นี้ เนื่องจากเงื่อนไขกำหนดให้มุกนั่งอยู่ด้านขวาของเอก เป้า และป้อม ต้องนั่งแยกกัน การนั่งที่จะต้องเป็นจริงถ้าแก้วนั่งอยู่ทางด้านซ้ายของเอกก็คือ การนั่งโดยมีเป้านั่งข้างโต หรือป้อมนั่งข้างโต

(ตอบเพียงคำตอบเดียว)

AR : แบบเลือกตอบ /

คำชี้แจง : ข้อคำถามในแบบทดสอบฉบับนี้ จะนำเสนอสถานการณ์ หรือเรื่องราวสั้น ๆ พร้อมกับกลุ่มของเงื่อนไข หรือกฎเกณฑ์ ซึ่งนักเรียนจะต้องพิจารณาวิเคราะห์เพื่อตอบคำถามจากเรื่องราวและเงื่อนไขที่กำหนดให้ ทั้งนี้โดยเลือกตอบจากตัวเลือก 1 - 5 ที่กำหนดให้ จากนั้นขีดเส้นหนาในช่อง ☐ ได้ตัวเลขประจำข้อคำถามที่ต้องการในกระดาษคำตอบ

ตัวอย่างข้อคำถาม

(0) โตเป็นเจ้าภาพจัดงานเลี้ยงอาหารเย็นให้เพื่อนสนิท 5 คน คือ มุก เอก แก้ว เป้า และป๊ม โต และเพื่อน ๆ จะต้องนั่งรับประทานอาหารที่โต๊ะกลม โดยรอบโต๊ะมีเก้าอี้ 6 ที่

เงื่อนไขที่กำหนดคือ :-

- เป้าจะต้องไม่นั่งข้างป๊ม
- เอกต้องมีที่นั่งติดด้านขวาของเขา

ใช้สถานการณ์ที่กำหนดให้ตอบคำถามข้อ (0) - (00)

(0) การจัดที่นั่งตามลำดับทั้งหมดต่อไปนี้ มีความเป็นไปได้ (เริ่มจากโตแล้ววนรอบโต๊ะไปจากด้านซ้าย ---> ขวาของเขา) ยกเว้น การจัดที่นั่งตามข้อใด

1. โต, เป้า, มุก, เอก, แก้ว, ป๊ม
2. โต, ป๊ม, มุก, เอก, แก้ว, เป้า
3. โต, ป๊ม, มุก, เอก, เป้า, แก้ว
4. โต, มุก, เอก, ป๊ม, แก้ว, เป้า
5. โต, แก้ว, ป๊ม, มุก, เป้า, เอก

สำหรับการพิจารณาคำตอบที่ถูกต้องของข้อนี้ จะเห็นว่าการจัดทั้งหมด (1 - 4) มีความเป็นไปได้ ยกเว้น (5) ซึ่งเป็นไปไม่ได้เนื่องจากละเมิดเงื่อนไขที่กำหนดว่า มุกจะต้องนั่งติดด้านขวาของเอก

(00) ถ้าแก้วนั่งทางด้านซ้ายของเอก ข้อใดที่กำหนดไว้ในหัวข้อ I ถึง III ต่อไปนี้ต้องเป็นจริง

- I. เอกนั่งข้างโต
- II. เป้านั่งข้างโต
- III. ป๊มนั่งข้างแก้ว

1. I
2. II
3. I และ II
4. II และ III
5. I, II และ III

การพิจารณาคำตอบของข้อ (00) นี้ เนื่องจากเงื่อนไขที่กำหนดให้มุกนั่งอยู่ด้านซ้ายของเอกแล้ว ทั้ง 3 คน จะต้องนั่งดังนี้ แก้ว, เอก, มุก เนื่องจากปุ้ม และเป้า ต้องนั่งแยกกัน การนั่งที่เป็นไปได้ คือ เป้าและปุ้มนั่งข้างโต แต่คนละด้าน ดังนั้น I ไม่สามารถเป็นจริง III อาจเป็นจริงได้ (แต่ไม่จริงเสมอไป) มี II เท่านั้น ที่จำเป็นต้องเป็นจริง ✓

2. การวิเคราะห์ค่าอธิบาย (AX)

ข้อความถามชนิดนี้มุ่งประเมินความสามารถในการสร้างข้อสรุปที่ถูกต้อง ความสามารถในการสร้างทางเลือก หรือค่าอธิบายที่เป็นไปได้ หรือที่ถูกต้อง

ทักษะย่อยที่ถือว่าข้อความชนิดนี้สามารถวัดได้ ได้แก่ ทักษะย่อยที่ 2, 3, 4 และ

5

ตัวอย่างข้อความ

AX : แบบเติมคำตอบ

คำชี้แจง : แบบทดสอบฉบับนี้ประกอบด้วยข้อความถาม ซึ่งกำหนดข้อความที่แสดงเรื่องราว หรือสถานการณ์ให้ จากนั้นจะมีผลที่เกิดขึ้น (หรือสิ่งที่เกิดตามมา) ซึ่งผลที่เกิดขึ้นนี้ต้องการค่าอธิบายที่สมเหตุสมผล อาจอยู่ในลักษณะของการสนับสนุนหรือคัดค้านก็ได้ ฉะนั้นให้นักเรียนพิจารณาว่าผลที่เกิดขึ้นดังกล่าว ควรได้รับคำอธิบายอย่างไร จึงจะสอดคล้องหรือเพียงพอหรือสมเหตุสมผล เมื่อพิจารณาคำตอบที่เหมาะสมได้แล้ว ให้เขียนคำตอบลงในส่วนที่เป็นคำตอบที่เว้นว่างไว้ให้ในแต่ละข้อความถาม

โปรดอย่าพิจารณาคำอธิบายที่ห่างไกล และไม่จำเป็นไปได้

ใช้สถานการณ์ที่กำหนดให้ต่อไปนี้ ตอบคำถามข้อ (0) - (00)

สถานการณ์ : ฝนตกลงมาอย่างหนักตลอดเวลา 5 ชั่วโมง ฉะนั้น พันดินจึงอึมไปด้วยน้ำ ด้วยเหตุนี้ น้ำจะต้องไหลลงสู่แม่น้ำ และลำคลอง สิ่งตามมา คือ ระดับน้ำในแม่น้ำ และลำคลอง จะสูงขึ้นอย่างรวดเร็ว จากผลนี้ทำให้คาดกันว่า ตอนเช้า น้ำจะท่วมในบริเวณใกล้เคียงริมแม่น้ำ

ผล : ทางเราได้ประกาศให้ประชาชนที่อยู่อาศัยในบริเวณดังกล่าวรีบเคลื่อนย้ายไปอยู่ในพื้นที่ที่สูงกว่า แต่ปรากฏว่าน้ำไม่ได้ท่วมอย่างที่คาดกัน

(0) คำถาม : คำอธิบายที่เกิดขึ้น และมีความเป็นไปได้ควรเป็นอย่างไร

คำตอบ :

(00) คำถาม : คำอธิบายใดเกี่ยวกับ การเตรียมการ ถ้าเป็นจริงจะสอดคล้อง
มีความเป็นไปได้ในการอธิบายผลที่เกิดขึ้น

คำตอบ :

AX : แบบเลือกตอบ

คำชี้แจง : แบบทดสอบฉบับนี้ประกอบด้วยข้อคำถาม ซึ่งกำหนดข้อความที่แสดงเรื่องราว หรือสถานการณ์ให้ จากนั้นจะมีผลที่เกิดขึ้น (หรือสิ่งที่เกิดตามมา) ซึ่งผลที่เกิดขึ้นนี้ต้องการคำอธิบายที่สมเหตุสมผล อาจอยู่ในลักษณะของการสนับสนุนหรือคัดค้านก็ได้ ฉะนั้นให้นักเรียนพิจารณาตัดสินว่าคำอธิบายผล (ตัวเลือก) ใด สอดคล้องหรือเฟียงพลหรือไม่ ต่อการอธิบายผลที่เกิดขึ้น จากสถานการณ์ที่กำหนดให้ในข้อคำถามแต่ละข้อ โดยทำเส้นทแยงในช่อง ☐ ใต้ตัวเลขประจำข้อคำตอบที่ต้องการ ซึ่งมี 5 ตัวเลือก ในกระดาษคำตอบ

โปรดอย่าพิจารณาคำอธิบายที่ห่างไกล และไม่น่าจะเป็นไปได้

ใช้สถานการณ์ที่กำหนดให้ต่อไปนี้ ตอบคำถามข้อ (0) - (00)

สถานการณ์ : ฝนตกลงมาอย่างหนักตลอดเวลา 5 ชั่วโมง ฉะนั้น พันดินจึงอุ่มไปด้วยน้ำ ด้วยเหตุนี้ น้ำจะต้องไหลลงสู่แม่น้ำ และลำคลอง สิ่งที่ตามมา คือ ระดับน้ำในแม่น้ำ และลำคลอง จะสูงขึ้นอย่างรวดเร็ว จากผลนี้ทำให้คาดกันว่า ตอนเช้า น้ำจะท่วมในบริเวณใกล้เคียงริมแม่น้ำ

ผล : ทางการได้ประกาศให้ประชาชนที่อยู่อาศัยในบริเวณดังกล่าวรีบเคลื่อนย้ายไปอยู่ในพื้นที่ที่สูงกว่า แต่ปรากฏว่าน้ำไม่ได้ท่วมอย่างที่คาดกัน

(0) ข้อความใดต่อไปนี้ ถ้าเป็นจริงจะสามารถอธิบายผลที่เกิดขึ้นได้อย่างเพียงพอ

1. ทางการระดมเครื่องมือขุดคลองเพื่อรองรับปริมาณน้ำที่เพิ่มขึ้นทันที
2. ชาวบ้านร่วมกันขนทรายทำเขื่อนขนานไปกับแม่น้ำลำคลองเพื่อกั้นน้ำ
3. หลังจากฝนตกหนัก แดดออก ทำให้อากาศร้อนจัดมากช่วยให้ น้ำระเหย

กลายเป็นไอไปมาก

4. เนื่องจากบริเวณตลอดริมแม่น้ำ ลำคลอง มีต้นไม้ขึ้นหนาแน่น รากของต้นไม้ช่วยดูดซับน้ำไว้

5. ระยะเวลาหลังฝนหยุด นานพอที่จะทำให้ น้ำในแม่น้ำ และคลองระบายน้ำลงสู่แหล่งรับน้ำ

สำหรับการพิจารณาคำตอบที่ถูกต้องของข้อนี้ จะเห็นว่าตัวเลือกที่ 1 - 4 มีความเป็นไปได้น้อย ตัวเลือกที่ 5 จึงเป็นตัวเลือกที่มีความเป็นไปได้ หรือมีความเป็นไปได้เป็นผลพอที่จะอธิบายผลที่เกิดขึ้นได้

(00) ข้อความต่อไปนี ถ้าเป็นจริงจะสอดคล้องมีความเป็นไปได้ในการอธิบาย ผลที่เกิดขึ้น

1. น้ำส่วนหนึ่งถูกโรงงานอุตสาหกรรมนำไปใช้
2. ชาวบ้านร่วมกันระดมเครื่องสูบน้ำออกจากแม่น้ำ ลำคลอง
3. ก่อนหน้านี้อาชาวบ้านได้เตรียมขุดลอกคูคลอง เพื่อขยายทางระบายน้ำไว้

แล้ว

4. หน่วยงานที่รับผิดชอบร่วมกับผู้นำชุมชนได้จัดเตรียมสถานที่ เพื่อแจกจ่าย อาหาร เครื่องนุ่งห่ม

5. ชาวบ้านเตรียมขนย้ายสิ่งของที่อาจได้รับความเสียหายไว้ก่อนฝนตกหนัก

แล้ว

สำหรับการพิจารณาคำตอบที่ถูกต้องของข้อนี้ จะเห็นว่าตัวเลือกที่ 1 และ 2 มีความเป็นไปได้น้อยมาก ส่วนตัวเลือกที่ 4 และ 5 เป็นการกระทำที่ได้มีผลต่อการเกิดน้ำท่วม เป็นเพียงการเตรียมการและการช่วยเหลือประชาชนเท่านั้น ฉะนั้น ตัวเลือกที่มีความเป็นไปได้มากที่สุด คือ ตัวเลือกที่ 3

3. ความเห็นตรงกันข้าม (CV)

ข้อความชนิดนี้ มุ่งวัดความสามารถในการจับความคิดสำคัญของความเห็น หรือ แสดงความเห็นวิจารณ์ ความเห็นที่น่าเสนอ

ทักษะข้อที่ถือว่า ข้อคำถามชนิดนี้สามารถวัดได้ ได้แก่ ทักษะข้อที่ 1, 2 และ 5

ตัวอย่างข้อคำถาม

CV : แบบเดิมคำตอบ

คำชี้แจง : ข้อคำถามของแบบทดสอบฉบับนี้ จะนำเสนอข้อความที่แสดงความเห็น 2 - 3 ความเห็น ให้นักเรียนพิจารณาตอบคำถามตามประเด็นที่กำหนด โดยใช้สาระที่มีอยู่ในแต่ละความเห็นที่น่าเสนอเป็นพื้นฐานในการตอบคำถาม เมื่อพิจารณาประเมิน และลงความเห็นว่าคำตอบใดถูก ควรเป็นเช่นไรแล้ว ให้เขียนตอบลงในส่วนที่เป็นคำตอบส่วนที่เว้นว่างไว้ให้ในแต่ละข้อคำถาม

ตัวอย่าง :

ความเห็นที่ 1 : พุทธสิ่งศักดิ์สิทธิ์ที่นับถือ ค่อนข้างจะผูกพันกับพระพุทธรูปในราช วัดเบญจมบพิตร เข้าวัดนี้มาตั้งแต่เด็ก เข้าไปแล้วเย็นมีความสุข เวลามีปัญหาอะไรจะนึกถึงท่าน สอมนรับว่าเชื่อสิ่งศักดิ์สิทธิ์ เชื่อไสยศาสตร์บ้าง จะไม่ลบหลู่ แต่ไม่มั่งงาย แต่ไม่ถึงขนาดว่า จะต้องทำตามที่คนทักคนบอก เพราะคิดว่าทุกอย่างขึ้นอยู่กับตัวเราและสิ่งแวดล้อม เรื่องราหุ มีคนบอก ทุกพุทธก็จะจุดธูปอธิษฐานถือว่าทำแล้วไม่เสียหาย

ความเห็นที่ 2 : สำหรับตนเองไม่ได้นับถือพระองค์ไหนเป็นพิเศษ คือ ถ้ามีคนบอกว่าองค์ไหนดี แล้วท่านจริงจัง ๆ ใจเราก็จะสรรเสริญนิยมน แต่ไม่ได้จะคิดเอามาบูชาจะไหว ราม ๆ ทำตามคำสอนของพระพุทธเจ้า จะเชื่อในเรื่องการทำดีและเชื่อตัวเองมากกว่า อย่างเรื่องราหุ คิดว่าไม่เกี่ยวหรอก อยู่ที่คนมากกว่าจะทำตัวอย่างไร แต่พอมีคนมาทักก็ อหากทำ

คำถาม : (0) ความเห็นทั้ง 2 สอดคล้องกันมากที่สุดในเรื่องใด

คำตอบ : ความเชื่อเกี่ยวกับการกระทำของตนเอง

เมื่อพิจารณาจากสาระที่ปรากฏในความเห็นทั้ง 2 ในแง่การปฏิบัติตนต่อสิ่งศักดิ์สิทธิ์ การปฏิบัติจริงในชีวิตประจำวันแล้ว ค่อนข้างแตกต่างกัน ประเด็นที่เหมือนกัน หรือตรงกันที่ ชัดเจน คือ เชื่อตนเอง หรือเชื่อการกระทำของตนเอง

โปรดอย่าพิจารณาคำอธิบายที่ห่างไกล และไม่น่าจะเป็นไปได้

CV : แบบเลือกตอบ

คำชี้แจง : ข้อคำถามของแบบทดสอบฉบับนี้ จะนำเสนอข้อความที่แสดงความเห็น 2 - 3 ความเห็น ให้นักเรียนพิจารณาตอบคำถามตามประเด็นที่กำหนด โดยใช้สาระที่มีอยู่ในแต่ละความเห็นที่นำเสนอเป็นพื้นฐานในการตอบคำถาม เมื่อพิจารณาประเมิน และลงความเห็นว่าคุณค่าตอบใดถูกต้องแล้ว ให้ทำเส้นทึบในช่อง ☐ ซึ่งอยู่ใต้ตัวเลขประจำข้อคำถามที่ต้องการในกระดาษคำตอบที่กำหนดให้

ตัวอย่าง :

ความเห็นที่ 1 : พุทธสิ่งศักดิ์สิทธิ์ที่นับถือ ก่อนข้างจะผูกพันกับพระพุทธชินราช วัดเบญจมบพิตร เข้าวัดนี้มาตั้งแต่เด็ก เข้าไปแล้วเย็นมีความสุข เวลามมีปัญหาอะไรจะนึกถึงท่าน ยอมรับว่าเชื่อสิ่งศักดิ์สิทธิ์ เชื่อไสยศาสตร์บ้าง จะไม่ลบหลู่ แต่ไม่ม่งาย แต่ไม่ถึงขนาดว่าจะต้องทำตามที่คุณชัก คนบอก เพราะคิดว่าทุกอย่างขึ้นอยู่กับตัวเรา และสิ่งแวดล้อม เรื่องราหมีคนบอก ทุกๆคนก็จะจุดธูปอธิษฐานถือว่าทำแล้วไม่เสียหาย

ความเห็นที่ 2 : สำหรับตนเองไม่ได้นับถือพระองค์ไหนเป็นพิเศษ คือ ถ้ามีคนบอกว่าองค์ไหนดี แล้วท่านจริงจัง ๆ ใจเราก็จะสรรเสริญนิยม แต่ไม่ได้จะคิดเอามาบูชา จะไหว้วรรม ๆ ทำตามคำสอนของพระพุทธเจ้า จะเชื่อในเรื่องการทำดีและเชื่อตัวเองมากกว่าอย่างเรื่องราว คิดว่าไม่เกี่ยวหรอก อยู่ที่คนมากกว่าจะทำตัวอย่างไร แต่พอมมีคนมาชักก็อยากทำ

คำถาม :

(0) ความเห็นทั้ง 2 สอดคล้องกันมากที่สุดในเรื่องใด

1. การดำเนินชีวิตประจำวัน
2. การเห็นคุณค่าของการทำดี
3. การปฏิบัติตนต่อสิ่งศักดิ์สิทธิ์
4. ความเชื่อเกี่ยวกับอำนาจภายนอก
5. ความเชื่อเกี่ยวกับการกระทำของตนเอง

เมื่อพิจารณาจากสาระที่ปรากฏในความเห็นทั้ง 2 จะเห็นว่า ประเด็น 1 - 4 นั้น ความเห็นทั้ง 2 ค่อนข้างแตกต่างกันในแง่ความเชื่อ และการปฏิบัติ ยกเว้น ความเชื่อเกี่ยวกับการกระทำของตนเองที่แสดงให้เห็นว่า ผู้แสดงความเห็นค่อนข้างเห็นพ้องกันว่า เชื่อตนเองมากกว่า ฉะนั้น ข้อที่เป็นคำตอบก็คือ (5)

4. เหตุผลทางตรรก (LR)

ข้อคำถามชนิดนี้ มุ่งวัดทักษะการใช้เหตุผลอย่างมีวิจารณญาณ เช่น การค้นหา ข้อตกลงเบื้องต้น การประเมินการอ้างเหตุผล การวิเคราะห์ข้อมูลหลักฐาน ที่กำหนดให้ ทักษะย่อยที่ถือว่า ข้อคำถามชนิดนี้สามารถวัดได้ ได้แก่ ทักษะย่อยที่ 1, 2, 3 และ 5

LR : แบบเติมคำตอบ ✓

คำชี้แจง : ข้อคำถามในแบบทดสอบฉบับนี้ จะประกอบด้วยข้อความตอนสั้น ๆ ในลักษณะของการอ้างเหตุผล และจะมีคำถามซึ่งนักเรียนจะต้องพิจารณาการอ้างเหตุผลที่กำหนดให้ตามประเด็นคำถามที่กำหนดไว้ในคำถามแต่ละข้อจากข้อมูล หลักฐานที่กำหนดให้ในข้อความ ด้วยการคิดวิเคราะห์ และใช้เหตุผลในการพิจารณาเพื่อลงสรุป หรือตัดสินใจได้อย่างถูกต้อง มีความหมาย จากนั้นให้เขียนคำตอบที่นักเรียนคิดว่าถูกต้อง สอดคล้อง หรือเป็นไปได้ลงในส่วนที่เป็นคำตอบซึ่งเว้นว่างไว้ให้ของแต่ละข้อ

ตัวอย่างข้อคำถาม

อ่านข้อความในกรอบแล้วตอบข้อคำถาม ข้อ 0

- (0) ฉันไม่เคยรู้สึกลำบากกับการใช้เข็มขัดนิรภัย เมื่อออกไปซื้อของข้างนอกเลย เพราะฉันไม่เคยขับรถเร็วเกินกว่า 60 กิโลเมตร ต่อชั่วโมง

คำถาม :

(0) ผู้กล่าวข้อความข้างต้นมีความคิดเกี่ยวกับการเกิดอุบัติเหตุทางรถยนต์อย่างไร

คำตอบที่น่าจะมีความหมายที่แสดงถึงความคิด ความเชื่อของผู้พูดก็คือ ความเชื่อที่ว่าอุบัติเหตุจะไม่เกิดขึ้น ถ้าขับรถด้วยความเร็วไม่เกิน 60 กิโลเมตร ต่อชั่วโมง ฉะนั้น จะได้คำตอบดังนี้

คำตอบ : อุบัติเหตุทางรถยนต์จะเกิดขึ้นเมื่อร่ว่งด้วยความเร็วเกินกว่า 60 กิโลเมตร ต่อชั่วโมง เท่านั้น

LR : แบบเลือกตอบ

คำชี้แจง : ข้อคำถามในแบบทดสอบฉบับนี้ จะประกอบด้วยข้อความตอนสั้น ๆ ในลักษณะของการอ้างเหตุผล และจะมีคำถามพร้อมคำตอบ 5 ตัวเลือก ซึ่งนักเรียนจะต้องพิจารณาข้อความการอ้างเหตุผลที่กำหนดให้ ตามประเด็นคำถามที่กำหนดไว้ในคำถามแต่ละข้อ จากข้อมูล หลักฐานที่กำหนดให้ในข้อความ ด้วยการคิดวิเคราะห์และใช้เหตุผลในการพิจารณาเพื่อลงสรุปหรือตัดสินใจได้อย่างถูกต้อง จากนั้นให้ขีดเส้นหนาในช่อง ☐ ใต้ตัวเลขประจำข้อคำตอบ ซึ่งเป็นตัวเลือกที่ต้องการในกระดาษคำตอบ

ตัวอย่างข้อคำถาม

อ่านข้อความในกรอบแล้วตอบข้อคำถาม ข้อ 0

(0) ฉันไม่เคยรู้สึกลำบากกับการใช้เข็มขัดนิรภัย เมื่อออกไปซื้อของข้างนอกเลย เพราะฉันไม่เคยขับรถเร็วเกินกว่า 60 กิโลเมตร ต่อชั่วโมง

คำถาม :

(0) ผู้กล่าวข้อความข้างต้นมีความคิดเกี่ยวกับการเกิดอุบัติเหตุทางรถยนต์อย่างไร

1. อุบัติเหตุจะไม่เกิดขึ้นเมื่อผู้ขับรถคาดเข็มขัดนิรภัย
2. อุบัติเหตุทางรถยนต์มักเกิดขึ้น เมื่อรถยนต์วิ่งด้วยความเร็ว 60 กิโลเมตร ต่อชั่วโมงพอดี
3. อุบัติเหตุทางรถยนต์จะเกิดขึ้น เมื่อรถวิ่งด้วยความเร็วเท่ากับ 60 กิโลเมตร ต่อชั่วโมง หรือน้อยกว่านั้น
4. อุบัติเหตุทางรถยนต์จะเกิดขึ้น เมื่อรถวิ่งด้วยความเร็วเกินกว่า 60 กิโลเมตร ต่อชั่วโมง เท่านั้น
5. อุบัติเหตุทางรถยนต์จะเกิดขึ้น เมื่อผู้ขับขี่ไม่ใช่คนที่อาศัยในบริเวณนั้น

จะเห็นว่าคำถาม ถามเพื่อให้นักเรียนมองให้ออกว่า ผู้พูดคิดอย่างไร ขณะที่เขา กล่าวประโยคนั้น จะเห็นว่าตัวเลือกที่ 4 เป็นข้อความที่น่าจะมีความหมายและแสดงให้เห็น ถึงความเชื่อของผู้พูดที่ว่า อุบัติเหตุจะไม่เกิดขึ้น ถ้าขับรถด้วยความเร็วไม่เกิน 60 กิโลเมตร ต่อชั่วโมง หรือความเร็วที่ต่ำกว่านั้น ฉะนั้น ถ้าเป็นจริงก็ถือว่าคำถามข้างต้นมีความหมาย

5. เหตุผลทางตรรกเกี่ยวกับจำนวน (NLR)

ข้อคำถามชนิดนี้ มุ่งวัดความสามารถในการวิเคราะห์ หรือประเมินข้อมูล หรือ ข้อค้นพบ ที่นำเสนอในรูปแบบตาราง แผนภูมิ หรือกราฟ ที่ผู้สอบต้องใช้ความสามารถในการ ผสมผสานเหตุผลทางตัวเลขและภาษา เพื่อแก้ปัญหา

ทักษะย่อยที่ถือว่า ข้อคำถามชนิดนี้สามารถวัดได้ ได้แก่ ทักษะย่อยที่ 2, 3, 4 และ 5

NLR : แบบเติมคำตอบ

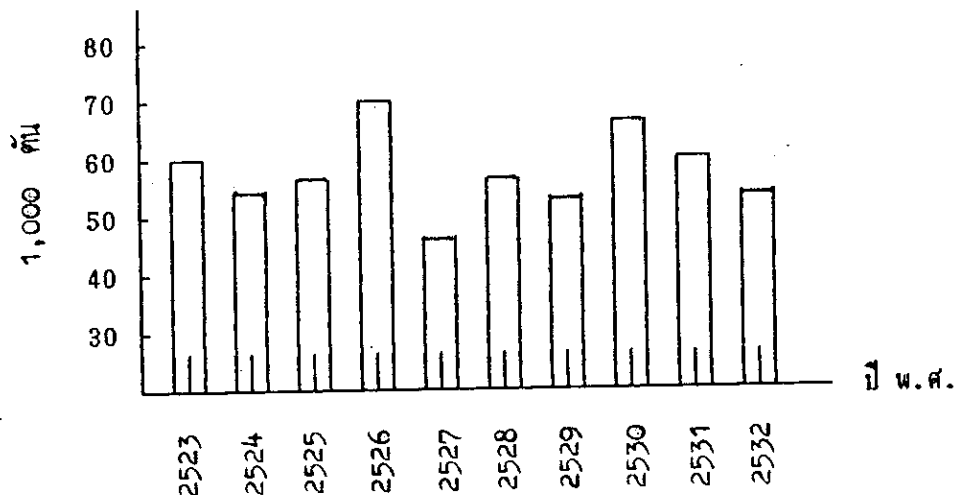
คำชี้แจง : แบบทดสอบฉบับนี้ประกอบด้วยข้อคำถาม ซึ่งจะกำหนดสถานการณ์ในรูปแบบของแผนภูมิ กราฟ ตาราง หรือชุดของข้อมูลต่าง ๆ ให้ นักเรียนจะต้องพิจารณาตอบ คำถามที่อาจกำหนดให้วิเคราะห์ หรือประเมินข้อสรุป สถานการณ์ หรือข้อสันนิษฐานต่าง ๆ จาก

สถานการณ์ข้อมูลหลักฐานที่กำหนดให้ ทั้งนี้โดยการเขียนคำตอบลงในส่วนที่เป็นคำตอบที่เว้นว่างไว้ให้ในแต่ละข้อคำถาม

ตัวอย่างข้อคำถาม

ใช้กราฟต่อไปนี้ ตอบคำถามข้อ (0) - (00)

ปริมาณปลาที่จับได้โดยกลุ่มเรือประมงที่ออกจากท่าเรือสงขลา



(0) คำถาม : โปรดเสนอเหตุผลที่ไม่สามารถให้ค่าอธิบายที่เป็นไปได้เกี่ยวกับปริมาณการจับปลาที่ต่ำกว่าปกติในปี พ.ศ. 2527

คำตอบ :

(00) คำถาม : ค่าอธิบายที่มีความเป็นไปได้ประการหนึ่งสำหรับความผิดปกติของปริมาณปลาที่จับได้ในปี พ.ศ. 2527 ใช้อ้างเป็นหลักฐาน คือ ระหว่างฤดูจับปลาปี พ.ศ. 2526 และ 2527 ราคาน้ำมันดีเซลที่ใช้เป็นเชื้อเพลิงของเรือประมงเพิ่มขึ้น 3 เท่า ราคาของปลาในปี 2527 ก็ยังคงอยู่ในระดับราคาของปี พ.ศ. 2526 แต่ในปลายปี พ.ศ. 2527

ราคาปลาที่ได้ปรับตัวสูงขึ้นอย่างมาก ค่าอธิบายเช่นไร ถ้าเป็นจริงจะสนับสนุนการอ้างหลักฐานข้างต้นเกี่ยวกับปริมาณปลาที่จับได้น้อยผิดปกติในปี 2527 ได้ดีที่สุด

คำตอบ :

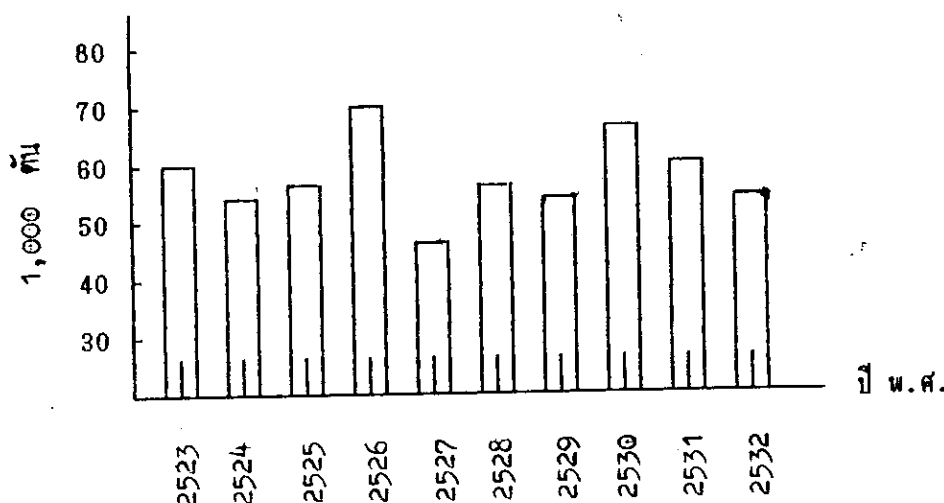
NLR : แบบเลือกตอบ

คำชี้แจง : แบบทดสอบฉบับนี้ประกอบด้วยข้อคำถาม ซึ่งจะกำหนดสถานการณ์ในรูปแบบของแผนภูมิ กราฟ ตาราง หรือชุดของข้อมูลต่าง ๆ ให้ นักเรียนจะต้องพิจารณาตอบคำถามที่อาจกำหนดให้วิเคราะห์ หรือประเมินข้อสรุป สถานการณ์ หรือข้อสังเกตต่าง ๆ จากสถานการณ์ข้อมูลหลักฐานที่กำหนดให้ ทั้งนี้จะต้องเลือกตัวเลือกที่สนับสนุนการตีความดีที่สุด หรือให้เลือกคำอธิบายข้อมูลที่เป็นไปได้มากที่สุด โดยขีดเส้นหนาในช่อง ☐ ให้ตัวเลขประจำข้อคำตอบที่ต้องการในกระดาษคำตอบ

ตัวอย่างข้อคำถาม

ใช้กราฟต่อไปนี้ ตอบคำถามข้อ (0) - (00)

ปริมาณปลาที่จับได้โดยกลุ่มเรือประมงที่ออกจากท่าเรือสงขลา



(0) ข้อใดต่อไปนี้ ถ้าเป็นจริงจะไม่สามารถให้คำอธิบายที่เป็นไปได้เกี่ยวกับ ปริมาณการจับปลาที่ผิดปกติในปี พ.ศ. 2527

1. ในฤดูจับปลาในปี 2527 น้ำมันจำนวนมากได้ไหลลงไปในทะเล ทำให้ ปริมาณปลาลดลงชั่วคราว เนื่องจากมันอพยพไปหาอาหารที่แหล่งอื่น
2. ระหว่างปี พ.ศ. 2527 มีการแบ่งอาณาบริเวณการหาปลาของไทยกับ ประเทศเพื่อนบ้าน แต่ต่อมาได้มีการเจรจาเพื่อทำสนธิสัญญาแบ่งเขตพื้นที่กัน
3. วิธีการหาปลาที่ล้ำสมัย ทำให้ประสิทธิภาพในการจับปลาลดลง แต่หลังจากฤดูจับปลาในปี 2527 ได้มีการนำเครื่องมือและวิธีการที่ทันสมัยเข้ามาใช้
4. ชาวประมงกลุ่มหนึ่งงดการออกไปจับปลาชั่วคราว เนื่องจากได้รับการ ร้องขอจากกลุ่ม Green Peace ซึ่งเดินทางแวะมาที่จังหวัดสงขลา
5. ปี พ.ศ. 2527 เป็นปีที่ชาวสงขลาถือว่ามีความสำคัญต่อความเป็นมา ของจังหวัด จึงร่วมใจกันที่จะงดการทำลายชีวิตสัตว์ทุกชนิด

การพิจารณาคำตอบของข้อ (0) นี้ จะเห็นว่าตัวเลือกที่ 1, 2, 4 และ 5 เป็น คำอธิบายที่มีความสมเหตุสมผลพอสำหรับจะอธิบายว่า เฉพาะในปี พ.ศ. 2527 มีเหตุการณ์ ดังกล่าวเกิดขึ้น แต่ตัวเลือก 3 นั้น หากพิจารณาปริมาณการจับปลาในปี พ.ศ. 2526 จะ เห็นว่าปริมาณการจับปลาได้สูงกว่าปีก่อน ๆ ฉะนั้น ข้ออ้างที่ว่าใช้วิธีการจับปลาที่ล้ำสมัย จึง เป็นไปไม่ได้และไม่สมเหตุสมผลพอ เพราะปีก่อน ๆ ก่อนหน้านั้น ก็ไม่มีภาวะเช่นนี้เกิดขึ้น ดังนั้น ตัวเลือกที่ 3 จึงเป็นตัวเลือกที่ถูก

(00) คำถาม : คำอธิบายที่มีความเป็นไปได้ประการหนึ่งสำหรับความผิดปกติของ ปริมาณปลาที่จับได้ในปี พ.ศ. 2527 ใช้อ้างเป็นหลักฐาน คือ ระหว่างฤดูจับปลาปี พ.ศ. 2526 และ 2527 ราคาน้ำมันดีเซลที่ใช้เป็นเชื้อเพลิงของเรือประมงเพิ่มเป็น 3 เท่า ราคา ของปลาในปี 2527 ก็ยังคงอยู่ในระดับราคาของปี พ.ศ. 2526 แต่ในปลายปี พ.ศ. 2527 ราคาปลาได้ปรับตัวสูงขึ้นอย่างมาก

ถ้าข้อความต่อไปนี้เป็นจริง ข้อใดจะสนับสนุนคำอธิบายที่อ้างเป็นหลักฐานข้างต้นได้
ดีที่สุด

1. ประชาชนงดบริโภคปลาในช่วงที่ปลาทูราคาสูง
2. พายุที่รุนแรงมากอย่างไม่เคยมีมาก่อนทำให้จำนวนวันที่จะออกไปหาปลาน้อยลง
3. เจ้าของเรือหาปลาจำเป็นต้องปลดลูกเรือประมงออกจำนวนหนึ่งเพื่อลดต้นทุน
4. มีการจำกัดขอบเขตการหาปลาของเรือประมงจากท่าสงขลาให้อยู่เฉพาะในพื้นที่ใกล้เคียงกับที่เรือมากที่สุด เพื่อลดการใช้น้ำมัน
5. การหยุดงานประท้วงเป็นระยะเวลานานของคนงานในโรงงานอาหารกระป๋องใกล้ท่าเรือสงขลา ทำให้หาปลาไม่ได้ ชาวประมงจึงหยุดจับปลากลางคืน

การพิจารณาคำตอบสำหรับข้อ (00) นี้ สำหรับตัวเลือกที่ 1, 2, และ 5 เป็นคำอธิบายที่ไม่สอดคล้องกับสถานการณ์ที่กำหนดให้ ตัวเลือก 3 นั้น มีโอกาสเป็นไปได้บ้างแต่ยังไม่ชัดเจน ตัวเลือกที่ 4 ให้คำอธิบายที่ถือว่ามีความสอดคล้องกับสถานการณ์ที่กำหนดมากที่สุด จึงถือเป็นตัวเลือกที่ถูกต้อง

6. การระบุรูปแบบ (PI)

ข้อความชนิดนี้ มุ่งวัดความสามารถในการพิจารณาสร้างกฎที่จะนำไปใช้กับชุดของจำนวนที่กำหนดให้

ทักษะข้อที่ถือว่า ข้อคำถามชนิดนี้สามารถวัดได้ ได้แก่ ทักษะข้อที่ 1, 2, 3 และ 4

PI : แบบเดิมค่าตอบ

คำชี้แจง : ข้อคำถามในแบบทดสอบฉบับนี้ ประกอบด้วยชุดของจำนวนชุดหนึ่งเป็นคำถาม นักเรียนจะต้องพิจารณาว่า จำนวนที่เป็นตัวคำถาม (ตัวปัญหา) ที่กำหนดให้ นั้น มีรูปแบบลำดับของจำนวนเป็นแบบใด ทั้งนี้โดยพิจารณาว่าเป็นไปตามกฎเช่นไร (เครื่องหมายที่กำหนดให้ใช้ในกฎ คือ บวก ลบ และคูณ) จากนั้นให้นักเรียนพิจารณาหาชุดของจำนวนที่มีรูปแบบลำดับของจำนวนที่ใช้กฎเช่นเดียวกับตัวคำถาม (ทั้งนี้จะกำหนดจำนวนที่เป็นตัวตั้งในแต่ละข้อให้) โดยเขียนตอบในส่วนที่เว้นว่างไว้ให้ในแต่ละข้อ

ตัวอย่างข้อคำถาม : (0)

ตัวปัญหา : 3, 4, 12, 13, 39, 40, 120

วิธีตอบ : จากชุดของจำนวนที่เป็นตัวปัญหา นักเรียนจะต้องพิจารณาดังนี้

1. สังเกตว่าชุดของจำนวนที่เป็นปัญหาใช้กฎเช่นไร (คำถามแต่ละข้อจะใช้กฎไม่เหมือนกัน) จากการพิจารณาจะพบว่า

3 ---> 4 (+1)	4 ---> 12 (x 3)
12 ---> 13 (+1)	13 ---> 39 (x 3)
39 ---> 40 (+1)	40 ---> 120 (x 3)

∴ กฎ ที่หาได้จากจำนวนในชุดที่เป็นตัวปัญหาที่กำหนดในตัวอย่างนี้ คือ

(+1), (x3), (+1), (x3), (+1), (x3)

สรุปให้เห็นได้ดังนี้

3,	4,	12,	13,	39,	40,	120
(+1),	(x3),	(+1),	(x3),	(+1),	(x3)	

2. นำกฎที่หาได้ในชุดของจำนวนที่กำหนดไปใช้กับจำนวนชุดใหม่ ซึ่งแต่ละข้อคำถามจะกำหนดจำนวนที่เป็นตัวตั้งไว้ให้ เช่น

จำนวนที่กำหนดให้เป็นตัวตั้ง

คำตอบ : 2,	3,	9,	10,	30,	31,	93
(+1),	(x3),	(+1),	(x3),	(+1),	(x3)	

.. จะเห็นว่าเมื่อนำกฎที่หาได้จาก (1) มาใช้กับจำนวนที่เป็นตัวตั้งที่กำหนดให้จะได้ชุดของจำนวนชุดหนึ่งที่ใช้กฎเช่นเดียวกับชุดที่เป็นตัวปัญหา คือ 2, 3, 9, 10, 30, 31,

93

PI : แบบเลือกตอบ

คำชี้แจง : ข้อคำถามในแบบทดสอบฉบับนี้ ประกอบด้วยชุดของจำนวนชุดหนึ่งเป็นตัวคำถาม และอีกชุดหนึ่งจะเป็นตัวเลือก นักเรียนจะต้องพิจารณาเลือกคำตอบที่มีรูปแบบลำดับของจำนวนที่เทียบได้ หรือเข้ากันได้กับชุดที่เป็นตัวคำถาม ทั้งนี้โดยพิจารณาว่า ลักษณะของลำดับของจำนวนซึ่งกำหนดเป็นไปตามกฎเช่นไร (เครื่องหมายที่กำหนดให้ใช้ในกฎ คือ บวก ลบ และคูณ) เพื่อนำไปพิจารณาใช้เทียบกับตัวเลือกในการหาคำตอบจากตัวเลือก 1 - 5 จากนั้นให้ทำเส้นทึบในช่อง ☐ ได้ตัวเลขประจำข้อคำตอบที่ต้องการในกระดาษคำตอบ

ตัวอย่างข้อคำถาม

(0) 2, 4, 3, 6, 5, 10, 9

รูปแบบ : ชุดของจำนวนชุดนี้ จำนวนลำดับที่ 2, 4 และ 6 จะใช้วิธีการคำนวณวิธีหนึ่ง และลำดับที่ 3, 5 และ 7 จะใช้การคำนวณอีกวิธีหนึ่ง โดยลำดับที่ 2, 4 และ 6 จะใช้กฎลำดับ โดยการนำ 2 คูณจำนวนที่อยู่ข้างหน้า ส่วนลำดับที่ 3, 5 และ 7 จะคำนวณโดยใช้ 1 ลบออกจากจำนวนที่อยู่ข้างหน้า ดังนี้

2, 4, 3, 6, 5, 10, 9

(x2) (-1) (x2) (-1) (x2) (-1)

วิธีตอบ : ในการตอบนักเรียนจะต้องพิจารณาว่า จากกฎที่ใช้ในการคำนวณจำนวนในชุดที่เป็นคำถาม จะเทียบได้กับตัวเลือกใด หรือตัวเลือกใดที่ชุดของจำนวนใช้กฎเช่นเดียวกับจำนวนในชุดที่เป็นคำถาม

- (0) 2, 4, 3, 6, 5, 10, 9
1. 3, 6, 4, 8, 6, 12, 10
2. 5, 7, 6, 8, 7, 9, 8
3. 7, 14, 13, 26, 25, 50, 49

จะเห็นว่าตัวเลือกที่ 3 ใช้กฎเดียวกับชุดของจำนวนที่เป็นคำถาม ดังนั้น ให้ทำเส้นทึบลงใน ☐ ที่ตรงกับ 3 ในกระดาษคำตอบ

การศึกษาคุณภาพของข้อคำถามและแบบทดสอบ

หลังจากสร้างข้อคำถามและตรวจสอบคุณภาพขั้นต้น และทำการปรับปรุงแก้ไขแล้ว นำข้อคำถามที่ปรับปรุงแก้ไขแล้ว จัดฉบับตามชนิดและรูปแบบ เพื่อนำไปทดสอบกับกลุ่มตัวอย่างที่กำหนดไว้ จากนั้นตรวจให้คะแนนผลการสอบที่ได้ นำไปวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ เพื่อศึกษาคุณภาพของข้อคำถามแต่ละชนิดและรูปแบบ ดังนี้

1. ตรวจสอบความเป็นมิติเดียวของข้อสอบ โดยการวิเคราะห์องค์ประกอบ (Factor Analysis) พิจารณาค่าไอเกน (Eigen Value) ขององค์ประกอบที่มีค่าสูงสุดมีค่าสูงกว่าค่าไอเกนขององค์ประกอบแรกมีค่ามากกว่าองค์ประกอบที่ 2 มาก และองค์ประกอบที่ 2 มีค่าไม่มากกว่าองค์ประกอบลำดับถัดไปมากนัก ก็ถือว่าข้อคำถามมีความเป็นมิติเดียว (Lord. 1980 : 21) นอกจากนี้ ได้พิจารณาความสามารถในการอธิบายความแปรปรวนขององค์ประกอบที่อธิบายแบบทดสอบแต่ละฉบับร่วมด้วย

2. ประมวลค่าคุณลักษณะข้อสอบ (Parameter) เพื่อหาค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าสัมประสิทธิ์การเดา (c) ตามทฤษฎีการตอบสนองข้อสอบ โดยใช้โมเดลโลจิสติก 3 พารามิเตอร์ ด้วยโปรแกรม BILOG3 (Mislevy and Bock. 1990 : 6 - 8)

3. คำนวณค่าฟังก์ชันสารสนเทศของข้อสอบ (Item Information Function) และฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Function) เพื่อพิจารณาความแน่นอนในการประมาณค่าความสามารถที่แท้จริง ซึ่งแทนค่าความเชื่อมั่นของแบบทดสอบ (Reliability) ตามทฤษฎีแบบดั้งเดิม และคำนวณหาค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ (Standard-Error of Estimate)

4. คำนวณค่าประสิทธิภาพสัมพัทธ์ (Relative Efficiency) เพื่อศึกษาคุณภาพของแบบทดสอบแต่ละชนิดและรูปแบบ

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. วิเคราะห์หาคคุณภาพของเครื่องมือตามทฤษฎีแบบดั้งเดิม โดยใช้โปรแกรมสำเร็จรูป (ITEM ANALYSIS 01) เพื่อหาค่าความยาก (p) และค่าอำนาจจำแนก (r) และค่าความเชื่อมั่นของแบบทดสอบ

1.1 ค่าอำนาจจำแนก (r) ใช้ดัชนีค่าอำนาจจำแนกที่หาจากสัมประสิทธิ์สหสัมพันธ์แบบไบซีเรียล (Biserial Correlation Coefficient) เนื่องจากพิจารณาเห็นว่าช่วงความยากของข้อคำถามมีมาก รวมทั้งไม่แน่ใจว่ากลุ่มตัวอย่างที่จะใช้ศึกษาต่อไปจะมีความสามารถที่แตกต่างกับกลุ่มที่ใช้ในการวิเคราะห์ปัจจุบันหรือไม่ การใช้สัมประสิทธิ์สหสัมพันธ์แบบไบซีเรียลจะค่อนข้างคงที่ เมื่อมีการเปลี่ยนแปลงกลุ่มตัวอย่าง (Crocker and Algina. 1986 : 315 - 320) โดยใช้สูตรที่พัฒนาโดยเพียร์สัน (Pearson. 1909)

$$r_{b.1.0} = \frac{(\mu_o - \mu_x)(p/y)}{\sigma_x}$$

เมื่อ μ_o = ค่าเฉลี่ยคะแนนเกณฑ์ของผู้ตอบข้อคำถามนั้นถูก
 μ_x = ค่าเฉลี่ยคะแนนเกณฑ์ของผู้สอบทั้งหมด
 σ_x = ความเบี่ยงเบนมาตรฐาน
 p = สัดส่วนของผู้สอบที่ตอบข้อคำถามนั้นถูก
 y = Ordinate ของ Standard Normal Curve

1.2 ค่าความเชื่อมั่น (Reliability) หาสัมประสิทธิ์ความเชื่อมั่นด้วยวิธีการหาความคงที่ภายในของคะแนนผลการสอบ โดยใช้คูเดอร์ และริชาร์ดสัน 20 (Kuder-Richardson : KR-20) เนื่องจากข้อคำถามแต่ละข้อมีความยากที่แตกต่างกันมาก สูตรที่ใช้ในการคำนวณพัฒนาขึ้นโดยคูเดอร์ และริชาร์ดสัน (Kuder and Richardson) (Crocker and Algina. 1986 : 138 - 139)

$$KR_{eo} = \frac{k}{k-1} (1 - \frac{\sum pq}{\hat{\sigma}_x^2})$$

- เมื่อ k = จำนวนข้อในแบบทดสอบ
 $\hat{\sigma}_x^2$ = ความแปรปรวนทั้งหมดของแบบทดสอบ
 pq = ความแปรปรวนของข้อคำถามรายข้อ

2. ประมาณค่าคุณลักษณะของข้อสอบ (Parameter) เพื่อหาค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าสัมประสิทธิ์การเดา (c) ด้วย โปรแกรมสำเร็จรูป BILOG3

3. คำนวณค่าฟังก์ชันสารสนเทศของข้อสอบและแบบทดสอบ โดยใช้สูตรของลอร์ด (Lord. 1980 : 70 - 72)

สูตรฟังก์ชันสารสนเทศของข้อสอบ

$$I\{e, u_i\} = P_i'^2 / P_i Q_i$$

สูตรฟังก์ชันสารสนเทศของแบบทดสอบ

$$I\{e\} = I\{e, e\} = \sum_{i=1}^n P_i'^2 / P_i Q_i$$

เมื่อ $I\{e, U_i\}$ คือ ฟังก์ชันสารสนเทศของข้อสอบข้อที่ i

$I\{e\}$ คือ ฟังก์ชันสารสนเทศของแบบทดสอบ

P_i คือ ความน่าจะเป็นที่ผู้มีความสามารถ e จะตอบข้อสอบข้อที่ i ถูก

Q_i คือ $1 - P_i$

P_i' คือ ความชันของโค้งลักษณะข้อสอบที่ระดับความสามารถ e

4. ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ (Hambleton and Swaminathan. 1985 : 104)

$$SE(e)* = 1/\sqrt{I(e)}$$

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การเสนอผลการวิเคราะห์ข้อมูลการวิจัยครั้งนี้ ผู้วิจัยได้กำหนดความหมายของ
สัญลักษณ์และคำย่อต่าง ๆ เพื่อสื่อความหมายในการเสนอผลการวิจัย ดังนี้

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

p	แทน	ค่าความยากของข้อสอบ โดยวิธีการวิเคราะห์ข้อสอบแบบดั้งเดิม
r	แทน	ค่าอำนาจจำแนกของข้อสอบ โดยวิธีการวิเคราะห์ข้อสอบแบบ ดั้งเดิม
a	แทน	ค่าอำนาจจำแนกของข้อสอบ โดยวิธีการวิเคราะห์ข้อสอบด้วย ทฤษฎีการตอบสนองข้อสอบ
b	แทน	ค่าความยากของข้อสอบ โดยวิธีการวิเคราะห์ข้อสอบด้วยทฤษฎี การตอบสนองข้อสอบ
c	แทน	ค่าการเดาของข้อสอบ โดยวิธีการวิเคราะห์ข้อสอบด้วยทฤษฎี การตอบสนองข้อสอบ
e	แทน	ค่าความสามารถของผู้สอบ
$\bar{a}$	แทน	ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ
$\bar{b}$	แทน	ค่าความยากเฉลี่ยของข้อสอบ
$\bar{c}$	แทน	ค่าการเดาเฉลี่ยของข้อสอบ
a_{SD}	แทน	ความเบี่ยงเบนมาตรฐานของค่าอำนาจจำแนกของข้อสอบ
$P(e)$	แทน	โอกาสในการตอบข้อคำถามได้ถูกต้อง ณ ระดับความสามารถ
b_{SD}	แทน	ความเบี่ยงเบนมาตรฐานของค่าความยากของข้อสอบ
c_{SD}	แทน	ความเบี่ยงเบนมาตรฐานของค่าการเดาของข้อสอบ
RE	แทน	ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ
TCC	แทน	วัดลักษณะของแบบทดสอบ
ICC	แทน	วัดลักษณะของข้อสอบ
TIF	แทน	ค่าฟังก์ชันสารสนเทศของแบบทดสอบ $\{I(e)\}$
$SE(e)$	แทน	ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ

AR	แทน	แบบทดสอบชนิดเหตุผลด้านวิเคราะห์
AX	แทน	แบบทดสอบชนิดการวิเคราะห์คำอธิบาย
CV	แทน	แบบทดสอบชนิดความเห็นตรงกันข้าม
LR	แทน	แบบทดสอบชนิดเหตุผลทางตรรก
NLR	แทน	แบบทดสอบชนิดเหตุผลทางตรรกเกี่ยวกับจำนวน
PI	แทน	แบบทดสอบชนิดการระบุรูปแบบ

กรณีผลการวิเคราะห์นำเสนอข้อแบบทดสอบที่ใช้ Subscript 31, 32, 51 และ 52 แสดงความหมายดังนี้

Subscript ตัวแรก หมายถึง ระดับชั้นของนักเรียน 3 หมายถึง ระดับชั้นมัธยมศึกษาปีที่ 3, 5 หมายถึง ระดับชั้นมัธยมศึกษาปีที่ 5

Subscript ตัวที่สอง หมายถึง รูปแบบของข้อคำถามหรือแบบทดสอบ คือ 1 หมายถึง รูปแบบเลือกตอบ, 2 หมายถึง รูปแบบเติมคำตอบ

การเสนอผลการวิเคราะห์ข้อมูล

ผู้วิจัยแบ่งการนำเสนอผลการวิเคราะห์ข้อมูลออกเป็น 2 ตอน ดังนี้

1. การดำเนินการพัฒนาเครื่องมือวัด

- 1.1 ค่าสถิติพื้นฐานของคะแนนที่ได้จากการสอบแบบทดสอบแต่ละฉบับ
- 1.2 คุณภาพของข้อคำถาม และแบบทดสอบแต่ละฉบับ

2. การศึกษาคุณภาพของข้อคำถาม และแบบทดสอบ

2.1 คุณภาพข้อคำถามและแบบทดสอบจากการวิเคราะห์ด้วยทฤษฎีการวิเคราะห์แบบดั้งเดิม

2.2 คุณภาพของข้อคำถามและแบบทดสอบจากการวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบ

2.2.1 ผลการวิเคราะห์องค์ประกอบเพื่อตรวจสอบความเป็นมิติเดียว

2.2.2 ค่าสถิติพื้นฐานของค่าพารามิเตอร์ของแบบทดสอบชนิดต่าง ๆ

2.2.3 ค่าฟังก์ชันสารสนเทศแบบทดสอบและความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ

2.2.4 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ

1. ผลการดำเนินการพัฒนาเครื่องมือวัด

หลังจากนำข้อคำถามจัดฉบับ เพื่อไปศึกษาในลักษณะนำร่อง (Pilot Study) กับ นักเรียนที่เป็นกลุ่มตัวอย่างแล้ว ผู้วิจัยได้นำข้อมูลทั้งที่เป็นผลการสอบ ความเห็นเกี่ยวกับข้อคำถาม วิธีการตอบมาประกอบการพิจารณาปรับปรุงข้อคำถาม คำชี้แจงการดำเนินการสอบ จากนั้นจัดฉบับนำไปทดสอบกับกลุ่มตัวอย่างเป้าหมายที่กำหนดไว้ ได้แก่ นักเรียนชั้นมัธยมศึกษา ปีที่ 3 และ 5

- 1.1 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบทดสอบแต่ละฉบับ
- 1.2 คุณภาพของข้อคำถาม และแบบทดสอบแต่ละฉบับ

ผลการวิเคราะห์ข้อมูล

- 1.1 ค่าสถิติพื้นฐานของคะแนนที่ได้จากการสอบแบบทดสอบแต่ละฉบับ

การวิเคราะห์คะแนนผลการสอบที่ได้จากแบบทดสอบฉบับย่อยแต่ละฉบับครั้งนี้ ผู้วิจัย ทำการวิเคราะห์ด้วยโปรแกรม Item Analysis 01 เพื่อนำเสนอข้อมูลเบื้องต้นเกี่ยวกับ แบบทดสอบประกอบด้วย ค่าเฉลี่ย (Mean) ค่าความเบี่ยงเบนมาตรฐาน (S.D.) ค่ามัธยฐาน (Median) และช่วงคะแนน เพื่อให้เห็นภาพโดยรวมของแบบทดสอบแต่ละฉบับ ดัง ตาราง 3 และ 4

ตาราง 3 จำนวนข้อคำถามของแบบทดสอบแต่ละฉบับ จำนวนผู้ตอบทั้งหมด ค่าเฉลี่ย ค่าความเบี่ยงเบนมาตรฐานของคะแนน ค่ามัธยฐาน และช่วงคะแนนที่ได้จากแบบทดสอบแต่ละฉบับ โดยกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3

รูปแบบ/ชนิดของข้อคำถาม	จำนวนข้อ	จำนวนผู้สอบ	ค่าเฉลี่ย	S.D.	มัธยฐาน	ช่วงคะแนนที่ได้	
						ต่ำสุด	สูงสุด
แบบเลือกตอบ 5 ตัวเลือก							
AR	21	138	9.39	3.81	10.00	1	19
AX	16	149	5.68	2.28	6.00	0	14
CV	15	175	6.61	2.10	7.00	1	11
LR	20	170	7.26	2.34	7.00	2	13
NLR	15	136	4.75	1.80	5.00	1	10
PI	19	143	11.62	5.35	12.50	0	19
แบบเติมคำตอบ							
AR	21	141	10.00	4.01	10.00	0	18
AX	16	127	4.76	3.09	4.50	0	12
CV	15	136	6.73	2.89	7.00	0	13
LR	20	121	3.78	2.81	3.00	0	14
NLR	15	139	4.62	2.44	4.00	0	11
PI	19	55	16.42	3.39	18.00	6	19

จากตาราง 3 แสดงให้เห็นว่า ค่าเฉลี่ยของคะแนนที่ได้จากแบบทดสอบทั้ง 2 รูปแบบ ทุกฉบับค่อนข้างต่ำมาก ยกเว้นฉบับ PI เมื่อเปรียบเทียบระหว่างรูปแบบพบว่า ฉบับ AR, CV และ PI แบบเติมคำตอบมีค่าเฉลี่ยคะแนนสูงกว่าแบบเลือกตอบ ในขณะที่ฉบับ AX, LR และ NLR แบบเลือกตอบมีค่าสูงกว่าแบบเติมคำตอบ สำหรับค่ามัธยฐานมีค่าใกล้เคียงกับ

ค่าเฉลี่ยทุกฉบับ เมื่อพิจารณาค่าช่วงคะแนนที่ได้ พบว่า ช่วงคะแนนค่อนข้างกว้างทุกฉบับ เช่นกัน โดยมีคะแนนต่ำสุดเป็น 0 แต่คะแนนสูงสุดเต็มหรือเกือบเต็ม

ตาราง 4 จำนวนข้อคำถามของแบบทดสอบแต่ละฉบับ จำนวนผู้ตอบทั้งหมด ค่าเฉลี่ย ค่าความเบี่ยงเบนมาตรฐานของคะแนน ค่ามัชฌิมฐาน และช่วงคะแนนที่ได้จากแบบทดสอบแต่ละฉบับ โดยกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5

รูปแบบ/ชนิดของข้อคำถาม	จำนวน	จำนวน	ค่า	S.D.	มัชฐาน	ช่วงคะแนนที่ได้	
	ข้อ	ผู้สอบ	เฉลี่ย			ต่ำสุด	สูงสุด
แบบเลือกตอบ 5 ตัวเลือก							
AR	21	108	10.29	4.37	12.00	0	19
AX	16	101	6.25	2.17	6.00	1	11
CV	15	106	6.85	2.26	7.00	2	11
LR	20	109	8.02	2.73	8.00	2	15
NLR	15	98	4.52	1.93	4.00	0	9
PI	19	104	14.12	4.73	16.00	1	19
แบบเติมคำตอบ							
AR	21	87	11.68	3.69	12.00	3	19
AX	16	78	4.88	3.39	5.00	0	14
CV	15	101	6.80	2.59	7.00	0	13
LR	20	76	3.84	2.32	4.00	0	10
NLR	15	74	5.91	2.44	6.00	1	13
PI	19	51	15.09	4.20	17.00	5	19

จากตาราง 4 แสดงให้เห็นว่า ค่าเฉลี่ยของคะแนนที่ได้จากแบบทดสอบทั้ง 2 รูปแบบ ทุกฉบับค่อนข้างต่ำมาก ยกเว้น ฉบับ PI เมื่อเปรียบเทียบระหว่างรูปแบบ พบว่า

ฉบับ AR, NLR และ PI แบบเติมคำตอบมีค่าเฉลี่ยสูงกว่าแบบเลือกตอบ ในขณะที่ฉบับ AX, CV และ LR แบบเลือกตอบมีค่าเฉลี่ยสูงกว่าแบบเติมคำตอบ สำหรับค่ามัธยฐานมีค่าใกล้เคียงกับค่าเฉลี่ยทุกฉบับ เมื่อพิจารณาช่วงคะแนนที่ได้ พบว่า ช่วงคะแนนค่อนข้างกว้างทุกฉบับเช่นกัน

1.2 คุณภาพของข้อคำถามและแบบทดสอบแต่ละฉบับ

จากการวิเคราะห์คุณภาพของแบบทดสอบแต่ละฉบับย่อย ปรากฏผลเกี่ยวกับคุณภาพของข้อคำถามรายชื่อ คือ ค่าความยาก (p) ค่าอำนาจจำแนก (r) ค่าเฉลี่ยสัดส่วนของผู้ตอบข้อคำถามถูก ค่าเฉลี่ยของอำนาจจำแนก และค่าความเชื่อมั่น (Reliability) ดังตาราง 5, 6, 7 และ 8 ส่วนคุณภาพข้อคำถามรายชื่อ คือ ค่าความยาก และค่าอำนาจจำแนกได้นำเสนอไว้ในภาคผนวก ก.

ผลการวิเคราะห์คุณภาพของข้อคำถามและแบบทดสอบ มีดังนี้

1. คุณลักษณะของข้อคำถามของแบบทดสอบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียน ชั้นมัธยมศึกษาปีที่ 3
2. คุณลักษณะของข้อคำถามของแบบทดสอบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียน ชั้นมัธยมศึกษาปีที่ 5
3. คุณภาพด้านความเชื่อมั่นของแบบทดสอบ
4. ข้อคำถามที่ได้รับการพิจารณาปรับแก้และตัดออกไปของแบบทดสอบแต่ละฉบับ

ตาราง 5 คุณลักษณะของข้อคำถามของแบบทดสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์
ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3

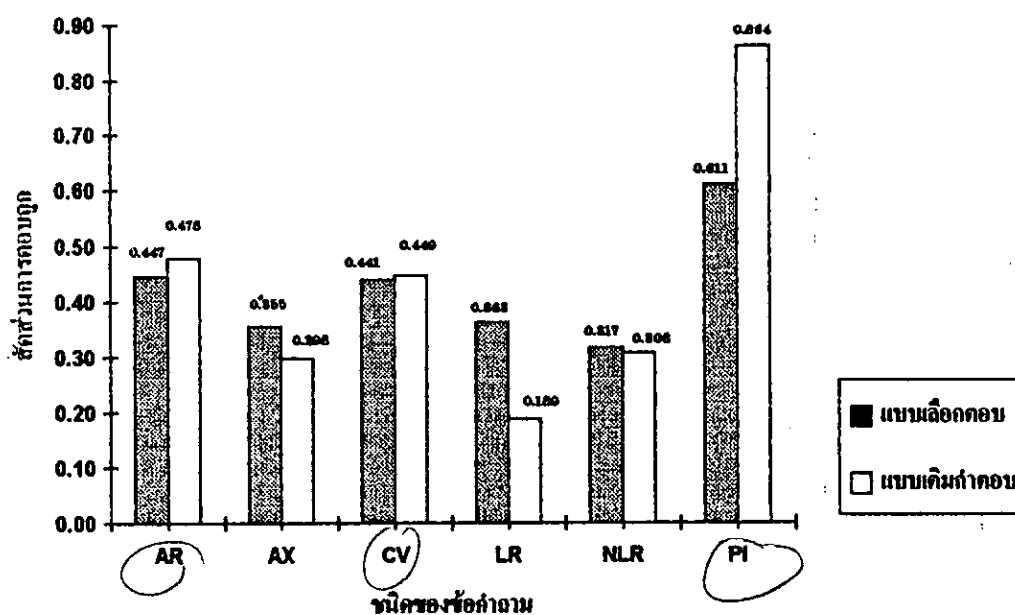
รูปแบบ/ชนิดของข้อคำถาม	ข้อ	ค่าเฉลี่ย สัดส่วน การตอบถูก	ค่าช่วง ความยาก (p)	ค่าเฉลี่ย อำนาจจำแนก (R-bis)	ช่วงค่าอำนาจ จำแนก (r)
แบบเลือกตอบ 5 ตัวเลือก					
AR	21	0.447	0.130-0.819	5.505	0.028-0.751
AX	16	0.355	0.040-0.651	0.394	0.028-0.603
CV	15	0.441	0.057-0.834	0.414	0.038-0.653
LR	20	0.363	0.112-0.812	0.344	0.130-0.653
NLR	15	0.317	0.118-0.809	0.368	0.131-0.567
PI	19	0.611	0.336-0.776	0.749	0.110-0.968
แบบเติมคำตอบ					
AR	21	0.476	0.106-0.823	0.544	0.332-0.811
AX	16	0.298	0.031-0.528	0.588	0.306-0.785
CV	15	0.449	0.081-0.706	0.525	0.307-0.683
LR	20	0.189	0.025-0.438	0.553	0.039-0.447
NLR	15	0.308	0.036-0.561	0.486	0.277-0.619
PI	19	0.864	0.727-0.982	0.784	0.079-0.957

จากตาราง 5 แสดงให้เห็นว่า ค่าเฉลี่ยสัดส่วนการตอบถูกของแบบทดสอบรูปแบบเลือกตอบ มีค่าจาก 0.317 ถึง 0.611 ขณะที่แบบเติมคำตอบมีค่าระหว่าง 0.189 ถึง 0.864 ส่วนค่าเฉลี่ยของค่าอำนาจจำแนกของแบบเลือกตอบ มีค่าจาก 0.344 ถึง 0.749 ขณะที่แบบเติมคำตอบมีค่าจาก 0.486 ถึง 0.784

ในส่วนช่วงของค่าความยากนั้น แบบทดสอบรูปแบบเลือกตอบทุกชนิดของข้อคำถามมีช่วงค่าความยากกว้างกว่าแบบเติมคำตอบ ยกเว้น ชนิด AR ที่รูปแบบเติมคำตอบมีช่วงค่าความยากกว้างกว่าแบบเลือกตอบ

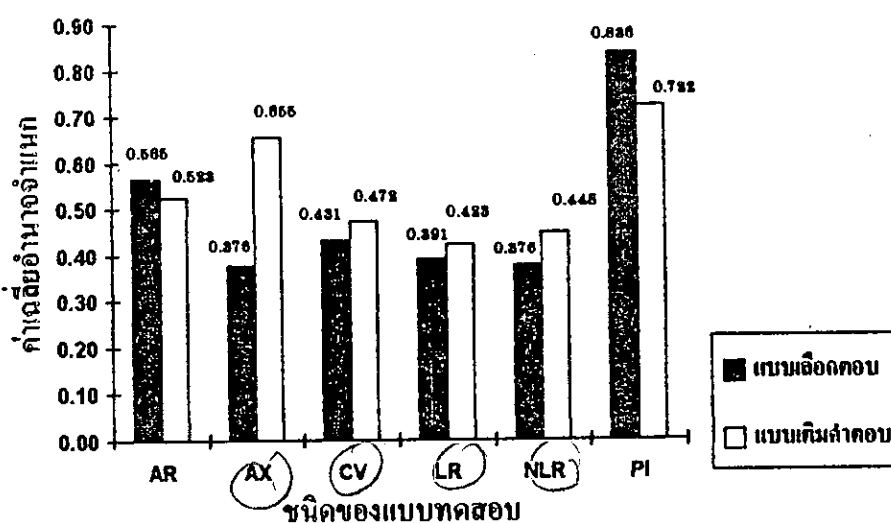
สำหรับช่วงค่าอำนาจจำแนกแบบทดสอบรูปแบบเลือกตอบ ชนิด PI มีช่วงค่าอำนาจจำแนกแคบกว่าแบบเติมคำตอบ ส่วนชนิด AR, AX, CV, LR และ NLR นั้น รูปแบบเลือกตอบมีช่วงค่าอำนาจจำแนกกว้างกว่าแบบเติมคำตอบ

เพื่อให้เห็นระดับความแตกต่างของค่าเฉลี่ยสัดส่วนของการตอบถูก ค่าเฉลี่ยอำนาจจำแนกของแบบทดสอบ 2 รูปแบบ คือ แบบเลือกตอบ และแบบเติมคำตอบ ทั้ง 6 ชนิด ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียน ชั้นมัธยมศึกษาปีที่ 3 ผู้วิจัยจึงนำค่าดังกล่าวมาสร้างกราฟปรากฏผลดังภาพประกอบ 15 - 16



ภาพประกอบ 15 ค่าเฉลี่ยสัดส่วนการตอบถูกสำหรับแบบทดสอบ 6 ชนิด ทั้ง 2 รูปแบบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3

จากภาพประกอบ 15 แสดงให้เห็นว่า แบบทดสอบแบบเติมคำตอบชนิด AR, CV และ PI มีสัดส่วนการตอบถูกสูงกว่าแบบเลือกตอบ ส่วนชนิด AX, AR และ NLR มีสัดส่วนการตอบถูกของแบบเลือกตอบสูงกว่าแบบเติมคำตอบ โดยชนิด NLR แสดงความแตกต่างระหว่างรูปแบบน้อยที่สุด ส่วนชนิด PI แสดงความแตกต่างระหว่างรูปแบบมากที่สุด



ภาพประกอบ 16 ค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบ 6 ชนิด ทั้ง 2 รูปแบบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3

จากภาพประกอบ 16 แสดงให้เห็นว่าแบบทดสอบรูปแบบเติมคำตอบ มีค่าอำนาจจำแนกเฉลี่ยสูงกว่าแบบเลือกตอบทุกชนิด โดยชนิด LR มีค่าความต่างระหว่างรูปแบบมากที่สุด ตามด้วยชนิด AX และ NLR ส่วนชนิด PI แสดงความต่างระหว่างรูปแบบน้อยที่สุด

ตาราง 6 คุณลักษณะของข้อคำถามของแบบทดสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์
ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5

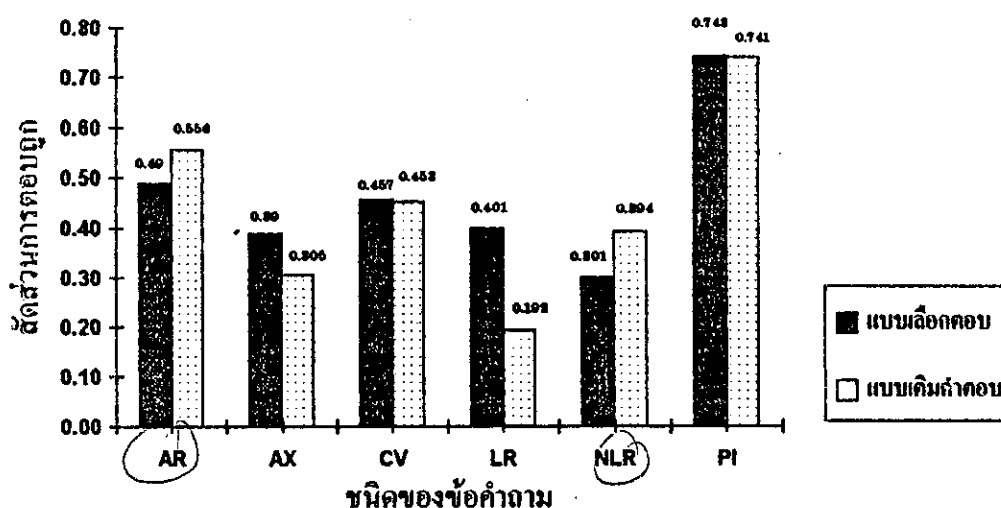
รูปแบบ/ชนิดของข้อคำถาม	ข้อ	ค่าเฉลี่ย สัดส่วน การตอบถูก	ช่วง ความยาก (p)	ค่าอำนาจ จำแนกเฉลี่ย (R-bis)	ช่วงค่าอำนาจ จำแนก (r)
แบบเลือกตอบ 5 ตัวเลือก					
AR	21	0.490	0.157-0.769	0.565	0.078-0.975
AX	16	0.390	0.129-0.624	0.376	0.084-0.654
CV	15	0.457	0.047-0.811	0.431	0.074-0.801
LR	20	0.401	0.119-0.899	0.391	0.111-0.682
NLR	15	0.301	0.051-0.765	0.376	0.112-0.642
PI	19	0.743	0.317-0.904	0.836	0.408-1.070
แบบเติมคำตอบ					
AR	21	0.556	0.069-0.839	0.523	0.097-0.850
AX	16	0.305	0.064-0.608	0.655	0.409-0.839
CV	15	0.453	0.030-0.723	0.472	0.121-0.725
LR	20	0.192	0.039-0.447	0.423	0.090-0.700
NLR	15	0.394	0.277-0.619	0.448	0.162-0.716
PI	19	0.741	0.522-0.957	0.722	0.163-1.050

จากตาราง 6 แสดงให้เห็นว่า ค่าเฉลี่ยสัดส่วนการตอบถูกของแบบทดสอบรูปแบบเลือกตอบมีค่าระหว่าง 0.301 ถึง 0.743 ขณะที่แบบเติมคำตอบมีค่าระหว่าง 0.192 ถึง 0.741 ส่วนค่าอำนาจจำแนกเฉลี่ยของแบบเลือกตอบมีค่าระหว่าง 0.376 ถึง 0.836 ขณะที่แบบเติมคำตอบมีค่าระหว่าง 0.423 ถึง 0.722

ในส่วนช่วงของค่าความชากนั้น แบบทดสอบรูปแบบเลือกตอบชนิด AR และ AX มีช่วงค่าความชากน้อยกว่าแบบเติมคำตอบ ส่วนชนิด CV, LR, NLR และ PI แบบเลือกตอบ มีช่วงค่าความชากกว้างกว่าแบบเติมคำตอบ

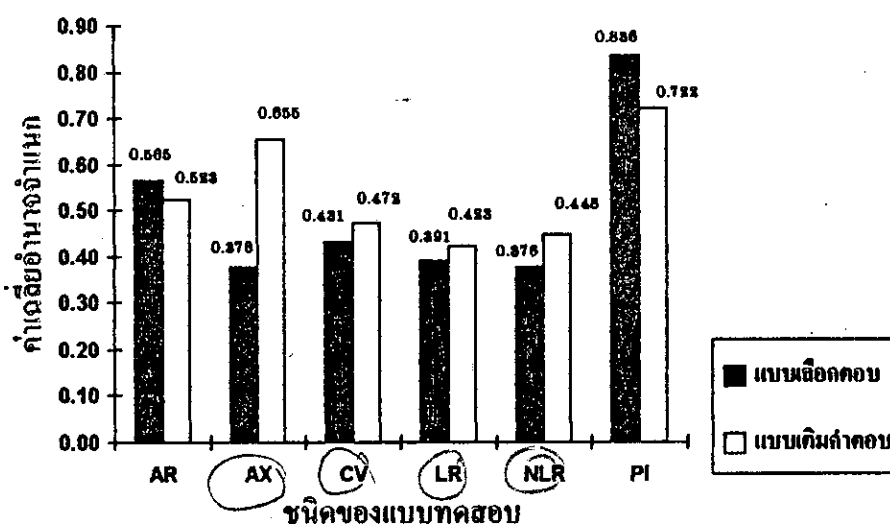
สำหรับช่วงค่าอำนาจจำแนกแบบเลือกตอบชนิด AR, AX และ CV มีช่วงค่าอำนาจจำแนกกว้างกว่าแบบเติมคำตอบ แต่ชนิด LR, NLR และ PI แบบเติมคำตอบมีช่วงค่าอำนาจจำแนกกว้างกว่าแบบเลือกตอบ

เพื่อให้เห็นระดับความแตกต่างของค่าเฉลี่ยสัดส่วนของการตอบถูก ค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบ 2 รูปแบบ คือ แบบเลือกตอบ และแบบเติมคำตอบ ทั้ง 6 ชนิด ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียน ชั้นมัธยมศึกษาปีที่ 5 ผู้วิจัยจึงนำค่าดังกล่าวมาสร้างกราฟปรากฏผลดังภาพประกอบ 17 - 18



ภาพประกอบ 17 ค่าเฉลี่ยสัดส่วนการตอบถูกสำหรับแบบทดสอบ 6 ชนิด ทั้ง 2 รูปแบบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 17 แสดงให้เห็นว่า แบบทดสอบชนิด PI ทั้ง 2 รูปแบบ มีสัดส่วนการตอบถูกสูงกว่าชนิดอื่น ๆ ส่วนชนิด AX และ NLR มีสัดส่วนการตอบถูกต่ำกว่าชนิดอื่น ๆ เมื่อพิจารณาเปรียบเทียบระหว่างรูปแบบ จะเห็นว่าชนิด AX, CV, LR และ PI รูปแบบเลือกตอบมีสัดส่วนการตอบถูกสูงกว่าแบบเติมคำตอบ ส่วนชนิด AR และ NLR รูปแบบเติมคำตอบมีสัดส่วนการตอบถูกสูงกว่าแบบเลือกตอบ ทั้งนี้ แบบทดสอบชนิด LR แสดงความแตกต่างของสัดส่วนการตอบถูกระหว่างรูปแบบมากที่สุด ในขณะที่ชนิด PI แสดงความแตกต่างระหว่างรูปแบบน้อยที่สุด



ภาพประกอบ 18 ค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบ 6 ชนิด ทั้ง 2 รูปแบบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 18 แสดงให้เห็นว่า แบบทดสอบรูปแบบเติมคำตอบชนิด AX, CV, LR และ NLR มีค่าเฉลี่ยอำนาจจำแนกสูงกว่าแบบเลือกตอบ ส่วนชนิด AR และ PI รูปแบบเลือกตอบมีค่าเฉลี่ยอำนาจจำแนกสูงกว่าแบบเติมคำตอบ โดยชนิด AX แสดงความแตกต่างระหว่างรูปแบบมากที่สุด ส่วนชนิด AR และ CV แสดงความแตกต่างระหว่างรูปแบบน้อยที่สุด

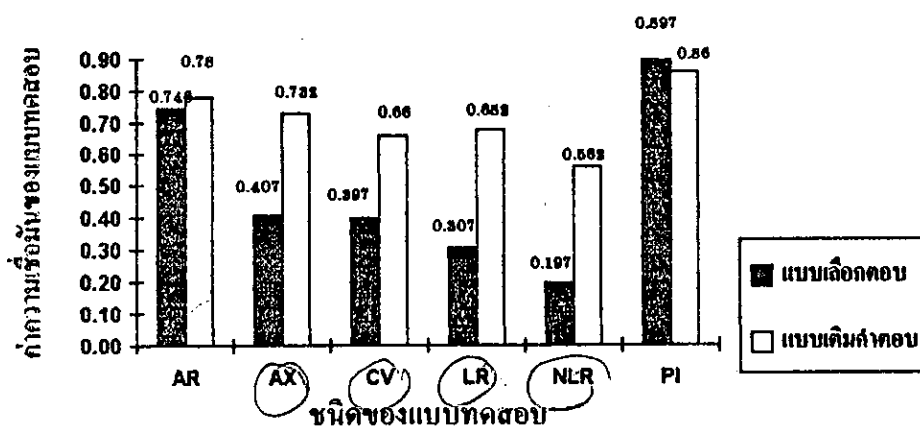
ตาราง 7 คุณภาพด้านความเชื่อมั่นของแบบทดสอบจากการทดลองครั้งที่ 1

ชนิดของข้อคำถาม	ชั้น ม.3				ชั้น ม.5			
	แบบเลือกตอบ		แบบเติมคำตอบ		แบบเลือกตอบ		แบบเติมคำตอบ	
	r_{tt}	SE_{meas}	r_{tt}	SE_{meas}	r_{tt}	SE_{meas}	r_{tt}	SE_{meas}
AR	0.746	1.920	0.780	1.883	0.807	1.917	0.738	1.887
AX	0.407	1.757	0.732	1.600	0.295	1.822	0.788	1.561
CV	0.397	1.629	0.660	1.689	0.466	1.654	0.572	1.695
LR	0.307	1.945	0.682	1.585	0.497	1.937	0.516	1.611
NLR	0.197	1.613	0.562	1.616	0.286	1.629	0.504	1.721
PI	0.897	1.712	0.860	1.266	0.905	1.457	0.880	1.446

จากตาราง 7 แสดงให้เห็นว่าสำหรับชั้นมัธยมศึกษาปีที่ 3 ค่าความเชื่อมั่นของแบบทดสอบแบบเลือกตอบมีค่าอยู่ระหว่าง 0.197 ถึง 0.897 ขณะที่แบบเติมคำตอบมีค่าระหว่าง 0.562 ถึง 0.860 โดยแบบเติมคำตอบทุกชนิดมีค่าความเชื่อมั่นสูงกว่าแบบเลือกตอบ ยกเว้น ชนิด PI

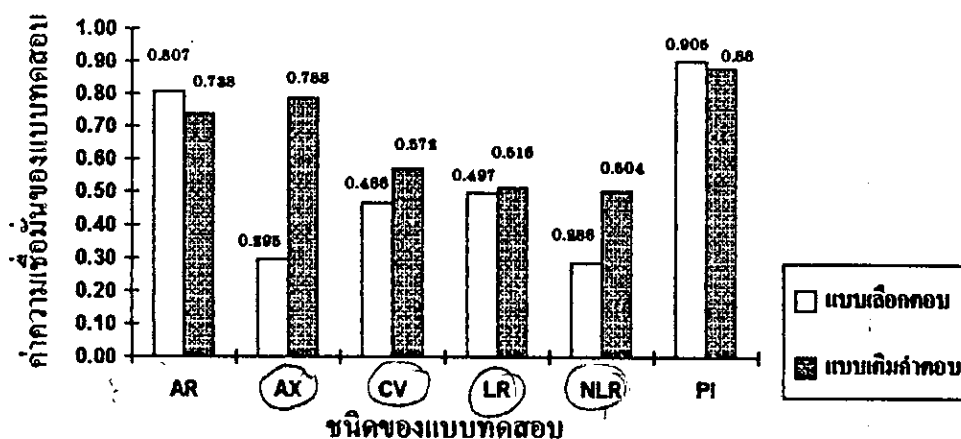
สำหรับชั้นมัธยมศึกษาปีที่ 5 ค่าความเชื่อมั่นของแบบทดสอบแบบเลือกตอบ มีค่าอยู่ระหว่าง 0.286 ถึง 0.905 ขณะที่แบบเติมคำตอบมีค่าระหว่าง 0.504 ถึง 0.880 โดยรูปแบบเลือกตอบชนิด AR และ PI มีค่าความเชื่อมั่นสูงกว่าแบบเติมคำตอบ ส่วนชนิด AX, CV, LR และ NLR แบบเติมคำตอบมีค่าความเชื่อมั่นสูงกว่าแบบเลือกตอบ

เพื่อให้เห็นภาพระดับความเชื่อมั่นของแบบทดสอบ 2 รูปแบบ ได้ชัดเจนขึ้น ผู้วิจัยขอนำเสนอค่าดังกล่าวเป็นกราฟ ปรากฏดังภาพประกอบ 19 - 20



ภาพประกอบ 19 ค่าความเชื่อมั่นของแบบทดสอบ 6 ชนิด ทั้ง 2 รูปแบบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3

จากภาพประกอบ 19 แสดงให้เห็นว่า แบบทดสอบรูปแบบเติมคำตอบให้ค่าความเชื่อมั่นสูงกว่าแบบเลือกตอบทุกชนิด ยกเว้นชนิด PI ซึ่งแบบเลือกตอบให้ค่าความเชื่อมั่นสูงกว่าแบบเติมคำตอบ โดยชนิด NLR มีความแตกต่างระหว่างรูปแบบสูงสุด ส่วน PI มีความแตกต่างระหว่างรูปแบบต่ำสุด เมื่อเปรียบเทียบระหว่างชนิดของแบบทดสอบแล้ว ชนิด NLR ให้ค่าต่ำสุดทั้ง 2 รูปแบบ ส่วนชนิด PI ให้ค่าความเชื่อมั่นสูงสุดทั้ง 2 รูปแบบ เช่นกัน



ภาพประกอบ 20 ค่าความเชื่อมั่นของแบบทดสอบ 6 ชนิด ทั้ง 2 รูปแบบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 20 แสดงให้เห็นว่า แบบทดสอบรูปแบบเดิมคำตอบชนิด AX, CV, LR และ NLR มีค่าความเชื่อมั่นสูงกว่าแบบเลือกตอบ โดยชนิด AX แสดงความแตกต่างสูงสุด ส่วนชนิด AR และ PI นั้น แบบเลือกตอบให้ค่าความเชื่อมั่นสูงกว่าแบบเดิมคำตอบ โดยชนิด PI แสดงความแตกต่างของความเชื่อมั่นระหว่างรูปแบบน้อยที่สุด

เมื่อพิจารณาเปรียบเทียบระหว่างชนิดของแบบทดสอบแล้ว ชนิด NLR มีค่าความเชื่อมั่นต่ำสุดทั้ง 2 รูปแบบ ส่วน ชนิด PI ให้ค่าความเชื่อมั่นสูงสุดทั้ง 2 รูปแบบ

หลังจากผู้วิจัยได้พิจารณาผลการวิเคราะห์คุณภาพของข้อคำถามรายข้อแล้ว และพบว่า มีข้อคำถามของแบบทดสอบบางฉบับมีข้อบกพร่องที่สามารถปรับปรุงแก้ไขให้ดีขึ้นและเหมาะสมขึ้นได้ ผู้วิจัยก็ได้พิจารณาปรับปรุง ส่วนข้อที่ไม่สามารถปรับปรุงได้นั้นเนื่องจากสถานการณ์ที่กำหนดให้เพื่อการตอบคำถามไม่เหมาะสม ผู้วิจัยได้ตัดออกไป รายละเอียดปรากฏดังตาราง 8

ตาราง 8 ข้อคำถามที่ได้รับการพิจารณาปรับแก้ และตัดออกจากการทดลองครั้งที่ 1

ชนิดของข้อคำถาม	จำนวนข้อ	ข้อที่พิจารณาปรับ	ข้อที่ตัดออก	หมายเหตุ
AR	21	2, 7, 17	-	ปรับปรุงประเด็นคำถาม ตัวเลือก
AX	16	2, 3, 4, 6	16	ปรับปรุงประเด็นคำถาม ตัวเลือก
CV	15	9, 13	1, 2, 3	สถานการณ์ที่ใช้ในข้อที่ตัด อาจ ไกลตัวของนักเรียน
LR	20	5, 6, 7, 9, 10, 11, 19, 20	13	ปรับปรุงประเด็นคำถาม ตัวเลือก
NLR	15	2, 3, 6, 11, 12, 13, 14	-	ปรับปรุงประเด็นคำถาม ตัวเลือก
PI	19	2, 10	-	ปรับปรุงคำชี้แจงการตอบ, ตัวเลือก

จากตาราง 8 แสดงให้เห็นว่า แบบทดสอบทุกฉบับมีข้อคำถามที่ต้องการได้รับการพิจารณาปรับแก้ประเด็นคำถามและตัวเลือก (กรณีเป็นแบบเลือกตอบ) ส่วนฉบับที่เป็นแบบเติมคำตอบจะพิจารณาปรับแก้ประเด็นคำถาม ฉบับที่ตัดข้อความออก ได้แก่ AX และ LR ถูกตัดออกไปฉบับละ 1 ข้อ ส่วนฉบับ CV ถูกตัดออก 3 ข้อ เนื่องจากสถานการณ์ที่กำหนดให้ข้อคำถามไม่เหมาะสมกับกลุ่มตัวอย่างในระดับนี้

2. ผลการศึกษาคุณภาพของข้อคำถามและแบบทดสอบ

จากผลการวิเคราะห์ข้อมูล ครั้งที่ 1 ผู้วิจัยได้นำมาเป็นข้อมูลประกอบการพิจารณาปรับแก้ข้อชี้แจงการตอบ ประเด็นคำถาม และตัวเลือก เพื่อให้เกิดความชัดเจน ถูกต้อง มีความเป็นไปได้ต่อการทำให้ผู้ตอบซึ่งมีความสามารถด้านเหตุผลต่ำเลือกตอบ สำหรับฉบับที่เป็นแบบเติมคำตอบจะเป็นการปรับประเด็นคำถาม ภาษาที่ใช้ให้ชัดเจนมากยิ่งขึ้น ตัดข้อคำถามบางข้อที่ไม่เหมาะต่อการนำไปใช้ได้ สำหรับข้อคำถาม 2 รูปแบบ ออกไป จากนั้นนำมาจัดฉบับ เพื่อนำไปใช้กับกลุ่มตัวอย่างเป้าหมายที่กำหนดไว้

ผลการวิเคราะห์

2.1 คุณภาพของข้อคำถามและแบบทดสอบ จากการวิเคราะห์ด้วยทฤษฎีการวิเคราะห์แบบดั้งเดิม

การวิเคราะห์คะแนนผลการสอบที่ได้จากการวิเคราะห์ด้วยโปรแกรม Item Analysis 0.1 ในการศึกษาคุณภาพข้อคำถามและแบบทดสอบ (ทดลองครั้งที่ 2) ปรากฏผลดังตาราง 9, 10 และ 11 ซึ่งได้นำเสนอตามลำดับ ดังนี้

1. คุณลักษณะของข้อคำถามของแบบทดสอบฉบับต่าง ๆ จากการทดลองครั้งที่ 2 ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3
2. คุณลักษณะของข้อคำถามของแบบทดสอบฉบับต่าง ๆ จากการทดลองครั้งที่ 2 ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5
3. ค่าความเชื่อมั่นของแบบทดสอบจากการทดลองครั้งที่ 2

ตาราง 9 คุณสมบัติของข้อคำถามของแบบทดสอบจากการทดลองครั้งที่ 2 ที่ได้จากกลุ่ม
ตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3

รูปแบบ/ชนิดของข้อคำถาม	จำนวนข้อ	ค่าเฉลี่ย สัดส่วน การตอบถูก	ช่วงค่าความยาก	ค่าเฉลี่ย R-bis	ช่วงค่าอำนาจ จำแนก
แบบเลือกตอบ 5 ตัวเลือก					
AR	20	0.415	0.145-0.701	0.486	0.124-0.768
AX	15	0.370	0.108-0.513	0.411	0.114-0.615
CV	12	0.468	0.133-0.751	0.558	0.047-0.347
LR	19	0.343	0.160-0.776	0.358	0.041-0.696
NLR	15	0.352	0.140-0.671	0.431	0.202-0.614
PI	19	0.582	0.410-0.752	0.818	0.670-0.901
แบบเติมคำตอบ					
AR	20	0.421	0.065-0.742	0.581	0.166-0.817
AX	15	0.277	0.083-0.594	0.570	0.282-0.798
CV	12	0.473	0.327-0.641	0.600	0.436-0.745
LR	19	0.245	0.109-0.547	0.462	0.225-0.591
NLR	15	0.294	0.082-0.549	0.419	0.203-0.659
PI	19	0.645	0.480-0.859	0.917	0.603-0.923

จากตาราง 9 แสดงให้เห็นว่า ค่าเฉลี่ยสัดส่วนการตอบถูกของแบบทดสอบรูปแบบ
เลือกตอบ มีค่าจาก 0.343 ถึง 0.582 ขณะที่แบบเติมคำตอบมีค่าระหว่าง 0.245 ถึง
0.645 ทั้งนี้โดยชนิด PI มีค่าเฉลี่ยสัดส่วนการตอบถูกสูงกว่าชนิดอื่นทั้ง 2 รูปแบบ และชนิด
LR ทั้ง 2 รูปแบบ มีค่าเฉลี่ยสัดส่วนการตอบถูกต่ำสุด

ส่วนค่าเฉลี่ยอำนาจจำแนกของแบบเลือกตอบ มีค่าจาก 0.358 ถึง 0.818 ขณะที่แบบเติมคำตอบมีค่าระหว่าง 0.419 ถึง 0.917 ชนิด PI ทั้งรูปแบบเติมคำตอบ และแบบเลือกตอบมีค่าอำนาจจำแนกสูงสุด สำหรับชนิดที่มีค่าอำนาจจำแนกเฉลี่ยต่ำสุดของแบบเลือกตอบ ได้แก่ ชนิด LR และแบบเติมคำตอบ ได้แก่ ชนิด NLR

ในส่วนของช่วงค่าความยากนั้น แบบทดสอบชนิด AR, AX และ PI แบบเติมคำตอบ มีช่วงค่าความยากกว้างกว่าแบบเลือกตอบ สำหรับช่วงค่าอำนาจจำแนกพบว่า แบบทดสอบชนิด AR, AX, CV, NLR และ PI รูปแบบเติมคำตอบมีช่วงค่าอำนาจจำแนกกว้างกว่าแบบเลือกตอบ

ตาราง 10 คุณสมบัติของข้อคำถามของแบบทดสอบจากการทดลองครั้งที่ 2 ที่ได้จาก
กลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5

รูปแบบ/ชนิดของข้อคำถาม	จำนวนข้อ	ค่าเฉลี่ย สัดส่วน การตอบถูก	ช่วงค่าความยาก	ค่าเฉลี่ย R-bis	ช่วงค่าอำนาจ จำแนก
แบบเลือกตอบ 5 ตัวเลือก					
AR	20	0.434	0.155-0.704	0.551	0.234-0.847
AX	15	0.419	0.091-0.561	0.436	0.110-0.623
CV	12	0.543	0.162-0.822	0.534	0.200-0.719
LR	19	0.411	0.161-0.862	0.356	0.081-0.520
NLR	15	0.387	0.157-0.779	0.443	0.182-0.605
PI	19	0.687	0.580-0.792	0.815	0.603-0.923
แบบเติมคำตอบ					
AR	20	0.542	0.098-0.803	0.590	0.261-0.883
AX	15	0.391	0.104-0.627	0.550	0.254-0.868
CV	12	0.600	0.390-0.789	0.620	0.412-0.786
LR	19	0.337	0.127-0.633	0.458	0.064-0.684
NLR	15	0.384	0.101-0.659	0.430	0.149-0.761
PI	19	0.737	0.541-0.926	0.901	0.430-0.901

จากตาราง 10 พบว่า ค่าเฉลี่ยสัดส่วนการตอบถูกของแบบทดสอบ รูปแบบเลือกตอบ มีค่าจาก 0.387 ถึง 0.687 ขณะที่แบบเติมคำตอบมีค่าระหว่าง 0.337 ถึง 0.737 ทั้งนี้ โดยชนิด PI มีค่าเฉลี่ยสัดส่วนการตอบถูกสูงสุด ทั้ง 2 รูปแบบ และ NLR รูปแบบเลือกตอบ และ LR รูปแบบเติมคำตอบ มีค่าเฉลี่ยสัดส่วนการตอบถูกต่ำสุด ส่วนค่าอำนาจเฉลี่ยของแบบเลือกตอบมีค่าระหว่าง 0.356 ถึง 0.815 ขณะที่แบบเติมคำตอบมีค่าระหว่าง 0.430 ถึง 0.901 โดยชนิด PI มีค่าอำนาจจำแนกเฉลี่ยสูงสุด

ในส่วนของช่วงค่าความยากนั้น แบบทดสอบชนิด AR, AX และ PI แบบเติมคำตอบมีค่าความยากกว้างกว่าแบบเลือกตอบ ขณะที่ชนิด CV, LR และ NLR รูปแบบเลือกตอบมีช่วงค่าความยากกว้างกว่าแบบเติมคำตอบ ทั้งนี้สำหรับแบบเลือกตอบชนิด PI มีช่วงค่าความยากน้อยที่สุด

สำหรับช่วงค่าอำนาจจำแนก ข้อคำถามชนิด AR, AX, NLR และ PI รูปแบบเติมคำตอบมีช่วงค่าอำนาจจำแนกกว้างกว่าแบบเลือกตอบ ขณะที่ชนิด CV และ LR รูปแบบเลือกตอบมีช่วงค่าความยากกว้างกว่าแบบเติมคำตอบ

ตาราง 11 ค่าความเชื่อมั่นของแบบทดสอบจากการทดลองครั้งที่ 2

ชนิดของข้อคำถาม	ชั้น ม.3				ชั้น ม.5			
	แบบเลือกตอบ		แบบเติมคำตอบ		แบบเลือกตอบ		แบบเติมคำตอบ	
	r_{L}	SE_{max}	r_{L}	SE_{max}	r_{L}	SE_{max}	r_{L}	SE_{max}
AR	0.721	1.987	0.814	1.899	0.796	1.954	0.824	1.931
AX	0.418	1.780	0.683	1.525	0.498	1.793	0.711	1.653
CV	0.653	1.528	0.699	1.541	0.579	1.533	0.718	1.502
LR	0.368	1.899	0.605	1.721	0.326	1.952	0.619	1.880
NLR	0.461	1.709	0.416	1.656	0.483	1.721	0.461	1.709
PI	0.926	1.650	0.948	1.451	0.919	1.584	0.937	1.401

จากตาราง 11 พบว่า สำหรับชั้นมัธยมศึกษาปีที่ 3 ค่าความเชื่อมั่นของแบบทดสอบชนิดเลือกตอบมีค่าอยู่ระหว่าง 0.368 ถึง 0.926 ขณะที่แบบเติมคำตอบมีค่าระหว่าง 0.416 ถึง 0.948 โดยรูปแบบเติมคำตอบมีค่าความเชื่อมั่นสูงกว่าแบบเลือกตอบทุกชนิด ยกเว้นชนิด NLR และชนิด PI มีความเชื่อมั่นสูงสุดทั้ง 2 รูปแบบ

ส่วนชั้นมัธยมศึกษาปีที่ 5 ค่าความเชื่อมั่นของแบบทดสอบแบบเลือกตอบมีค่าอยู่ระหว่าง 0.326 ถึง 0.919 ขณะที่แบบเติมคำตอบมีค่าระหว่าง 0.461 ถึง 0.937 ทั้งนี้ รูปแบบเติมคำตอบมีค่าความเชื่อมั่นสูงกว่าแบบเลือกตอบทุกชนิดของข้อคำถาม ยกเว้นชนิด NLR เช่นกัน โดยชนิด PI มีความเชื่อมั่นสูงสุดทั้ง 2 รูปแบบ

2.2 คุณภาพของข้อคำถามและแบบทดสอบ จากการวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory)

การวิเคราะห์คะแนนผลการสอบด้วยทฤษฎีการตอบสนองข้อสอบนี้ ถือเป็นการวิเคราะห์หลักของการศึกษาคุณภาพเครื่องมือในการทดลองครั้งที่ 2 นี้ ซึ่งผู้วิจัยได้นำเสนอผลการวิเคราะห์ข้อมูลตามลำดับ ดังนี้

- 2.1 ผลการวิเคราะห์องค์ประกอบเพื่อตรวจสอบความเป็นมิติเดียว
- 2.2 ค่าสถิติพื้นฐานของค่าพารามิเตอร์ของแบบทดสอบชนิดต่าง ๆ
- 2.3 ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบทดสอบ
- 2.4 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ

ผลการวิเคราะห์

- 2.1 ผลการวิเคราะห์องค์ประกอบเพื่อตรวจสอบความเป็นมิติเดียว

ขั้นตอนแรกของการวิเคราะห์ข้อมูลตามหลักการของทฤษฎีการตอบสนองข้อสอบ คือ การตรวจสอบว่าเครื่องมือวัดแต่ละฉบับมีคุณสมบัติด้านความเป็นมิติเดียวหรือไม่ ซึ่งในการวิจัยครั้งนี้ ผู้วิจัยได้ทำการวิเคราะห์องค์ประกอบ (Factor Analysis) โดยการสกัดองค์ประกอบ แบบวิธีวิเคราะห์องค์ประกอบหลัก (Principal-Components Analysis) และหมุนแกน (Rotation) ด้วยวิธีนาวาร์แมกซ์ (Varimax) โดยใช้โปรแกรม SPSS ผลการวิเคราะห์ ปรากฏดังตาราง 12 - 13

ตาราง 12 ผลการวิเคราะห์องค์ประกอบของแบบทดสอบแบบเลือกตอบ

องค์ประกอบที่	ค่าไอเกน (Eigen Value) และเปอร์เซ็นต์ของความแปรปรวน											
	AR		AX		CV		LR		NLR		PI	
	Eigen	% of V	Eigen	% of V	Eigen	% of V	Eigen	% of V	Eigen	% of V	Eigen	% of V
1	3.677	17.5	2.097	14.0	2.567	21.4	2.011	10.6	1.943	13.0	8.040	42.3
2	1.528	7.3	1.231	8.2	1.196	10.0	1.149	7.6	1.343	9.0	1.390	7.3
3	1.381	6.6	1.171	7.8	1.109	9.2	1.362	7.2	1.218	8.1		
4	1.211	5.8	1.099	7.3	1.006	8.4	1.244	6.5	1.144	7.6		
5	1.199	5.7	1.076	7.2			1.155	6.1	1.085	7.2		
6	1.062	5.1	1.003	6.7			1.123	5.9	1.005	6.7		
7	1.008	4.8					1.026	5.4				

จากตาราง 12 การวิเคราะห์โดยวิธีวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis) เมื่อใช้เกณฑ์ค่าไอเกน (Eigen Value) มากกว่าหรือเท่ากับ 1 ในการสกัดองค์ประกอบ พบว่า แบบทดสอบรูปแบบเลือกตอบทั้ง 6 ฉบับ ให้จำนวนองค์ประกอบมากกว่า 1 องค์ประกอบ อย่างไรก็ตาม เมื่อพิจารณาความสามารถในการอธิบายความแปรปรวนขององค์ประกอบแต่ละองค์ประกอบ เทียบกับความสามารถในการอธิบายความแปรปรวนรวมทุกองค์ประกอบของแบบทดสอบแต่ละฉบับ ปรากฏผลดังนี้

ฉบับ AR มีความแปรปรวนรวมทุกองค์ประกอบเท่ากับร้อยละ 52.8 ขณะที่องค์ประกอบแรกมีความสามารถในการอธิบายความแปรปรวนคิดเป็นประมาณร้อยละ 30 ของความแปรปรวนทั้งหมด โดยที่องค์ประกอบลำดับถัดไปมีความสามารถในการอธิบายความแปรปรวนคิดเป็นร้อยละไม่ถึงครึ่งหนึ่งขององค์ประกอบแรก จึงถือว่าองค์ประกอบแรกสามารถอธิบายความแปรปรวนของแบบทดสอบได้มากพอที่จะกล่าวได้ว่า มีความเป็นมิติเดียว

ฉบับ AX มีความแปรปรวนรวมทุกองค์ประกอบเท่ากับร้อยละ 51.2 ขณะที่องค์ประกอบแรกมีความสามารถในการอธิบายความแปรปรวนของแบบทดสอบคิดเป็นประมาณร้อยละ 30 ของความแปรปรวนทั้งหมด โดยองค์ประกอบลำดับถัดไป มีความสามารถในการอธิบายความแปรปรวนคิดเป็นร้อยละไม่ถึงครึ่งหนึ่งขององค์ประกอบแรก จึงถือว่าองค์ประกอบแรกสามารถอธิบายความแปรปรวนของแบบทดสอบได้มากพอที่จะกล่าวได้ว่ามีความเป็นมิติเดียว

ฉบับ CV มีความแปรปรวนรวมทุกองค์ประกอบเท่ากับร้อยละ 49.0 ขณะที่องค์ประกอบแรกมีความสามารถในการอธิบายความแปรปรวนคิดเป็นประมาณร้อยละ 42.9 ของความแปรปรวนทั้งหมด ซึ่งถือว่าสามารถอธิบายความแปรปรวนของแบบทดสอบได้เกือบทั้งหมด จึงมีความเป็นมิติเดียว

ฉบับ LR มีความแปรปรวนรวมทุกองค์ประกอบเท่ากับร้อยละ 49.3 ขณะที่องค์ประกอบแรกมีความสามารถในการอธิบายความแปรปรวนของแบบทดสอบคิดเป็นประมาณร้อยละ 30 ของความแปรปรวนทั้งหมด โดยองค์ประกอบลำดับถัดไปมีความสามารถในการอธิบายความแปรปรวนของแบบทดสอบคิดเป็นร้อยละไม่ถึงครึ่งหนึ่งขององค์ประกอบแรก จึงถือว่าองค์ประกอบแรกสามารถอธิบายความแปรปรวนของแบบทดสอบได้มากพอที่จะกล่าวได้ว่า มีความเป็นมิติเดียว

ฉบับ NLR มีความแปรปรวนรวมทุกองค์ประกอบเท่ากับร้อยละ 51.6 ขณะที่องค์ประกอบแรกมีความสามารถในการอธิบายความแปรปรวนของแบบทดสอบคิดเป็นประมาณร้อยละ 30 ของความแปรปรวนทั้งหมด โดยองค์ประกอบลำดับถัดไปมีความสามารถในการอธิบายความแปรปรวนคิดเป็นร้อยละไม่ถึงครึ่งหนึ่งขององค์ประกอบแรก จึงถือว่าองค์ประกอบแรกสามารถอธิบายความแปรปรวนของแบบทดสอบได้มากพอที่จะกล่าวได้ว่ามีความเป็นมิติเดียว

ฉบับ PI มีความแปรปรวนรวมทุกองค์ประกอบเท่ากับร้อยละ 49.6 ขณะที่องค์ประกอบแรกมีความสามารถอธิบายความแปรปรวนของแบบทดสอบคิดเป็นร้อยละ 85.2 ของความแปรปรวนทั้งหมด ซึ่งถือว่าสามารถอธิบายความแปรปรวนของแบบทดสอบได้เกือบทั้งหมด จึงมีความเป็นมิติเดียว

ตาราง 13 ผลการวิเคราะห์องค์ประกอบของแบบทดสอบแบบเดิมคำตอบ

องค์ประกอบ	ค่าไอเกน (Eigen Value) และเปอร์เซ็นต์ของความแปรปรวน											
	AR		AX		CV		LR		NLR		PI	
	Eigen	% of V	Eigen	% of V	Eigen	% of V	Eigen	% of V	Eigen	% of V	Eigen	% of V
1	5.068	24.1	3.227	21.5	3.048	25.4	2.704	14.2	2.005	13.4	9.558	50.3
2	1.676	8.0	1.327	8.8	1.170	9.8	1.388	7.3	1.345	9.0	1.935	10.2
3	1.457	6.9	1.201	8.0	1.020	8.5	1.353	7.1	1.212	8.1	1.079	5.7
4	1.211	5.8	1.137	7.6			1.169	6.2	1.127	7.5		
5	1.135	5.4	1.036	6.9			1.133	6.0	1.107	7.4		
6	1.068	5.1					1.055	5.6	1.029	6.9		
7							1.004	5.3				

จากตาราง 13 การวิเคราะห์โดยวิธีวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis) เมื่อใช้เกณฑ์ค่าไอเกน (Eigen Value) มากกว่าหรือเท่ากับ 1 ในการสกัดองค์ประกอบ พบว่า แบบทดสอบรูปแบบเดิมคำตอบทั้ง 6 ฉบับ ให้จำนวนองค์ประกอบมากกว่า 1 องค์ประกอบ อย่างไรก็ตาม เมื่อพิจารณาความสามารถในการอธิบายความแปรปรวนขององค์ประกอบแต่ละองค์ประกอบ เทียบกับความสามารถในการอธิบายความแปรปรวนรวมทุกองค์ประกอบของแบบทดสอบแต่ละฉบับ ปรากฏผลดังนี้

ฉบับ AR มีความแปรปรวนรวมทุกองค์ประกอบเท่ากับร้อยละ 55.3 ขณะที่องค์ประกอบแรกมีความสามารถในการอธิบายความแปรปรวนคิดเป็นประมาณร้อยละ 44.0 ของความแปรปรวนทั้งหมด โดยที่องค์ประกอบลำดับถัดไปมีความสามารถในการอธิบายความแปรปรวนคิดเป็นร้อยละไม่ถึงครึ่งหนึ่งขององค์ประกอบแรก จึงถือว่ามีความเป็นมิติเดียว

ฉบับ AX มีความแปรปรวนรวมทุกองค์ประกอบเท่ากับร้อยละ 52.8 ขณะที่องค์ประกอบแรกมีความสามารถในการอธิบายความแปรปรวนคิดเป็นประมาณร้อยละ 40.0 ของความแปรปรวนทั้งหมด โดยที่องค์ประกอบลำดับถัดไปมีความสามารถในการอธิบายความแปรปรวนคิดเป็นร้อยละไม่ถึงครึ่งหนึ่งขององค์ประกอบแรก จึงถือว่ามีความเป็นมิติเดียว

ฉบับ CV มีความแปรปรวนรวมทุกองค์ประกอบเท่ากับร้อยละ 43.7 ขณะที่องค์ประกอบแรกมีความสามารถในการอธิบายความแปรปรวนคิดเป็นร้อยละ 58.0 ของความแปรปรวนทั้งหมด โดยที่องค์ประกอบลำดับถัดไปมีความสามารถในการอธิบายความแปรปรวนคิดเป็นร้อยละไม่ถึงครึ่งหนึ่งขององค์ประกอบแรก จึงถือว่ามีความเป็นมิติเดียว

ฉบับ LR มีความแปรปรวนรวมทุกองค์ประกอบเท่ากับร้อยละ 51.7 ขณะที่องค์ประกอบแรกมีความสามารถในการอธิบายความแปรปรวนของแบบทดสอบคิดเป็นประมาณร้อยละ 30 ของความแปรปรวนทั้งหมด โดยองค์ประกอบลำดับถัดไปมีความสามารถในการอธิบายความแปรปรวนไม่ถึงครึ่งหนึ่งขององค์ประกอบแรก จึงถือว่าองค์ประกอบแรกสามารถอธิบายความแปรปรวนของแบบทดสอบได้มากพอที่จะกล่าวได้ว่า มีความเป็นมิติเดียว

ฉบับ NLR มีความแปรปรวนรวมทุกองค์ประกอบเท่ากับร้อยละ 52.3 ขณะที่องค์ประกอบแรกมีความสามารถอธิบายความแปรปรวนของแบบทดสอบคิดเป็นประมาณร้อยละ 30 ของความแปรปรวนทั้งหมด โดยที่องค์ประกอบลำดับถัดไปมีความสามารถในการอธิบายความแปรปรวนไม่ถึงครึ่งหนึ่งขององค์ประกอบแรก จึงถือว่าองค์ประกอบแรกสามารถอธิบายความแปรปรวนของแบบทดสอบได้มากพอที่จะกล่าวได้ว่า มีความเป็นมิติเดียว

ฉบับ PI มีความแปรปรวนรวมทุกองค์ประกอบเท่ากับร้อยละ 66.2 ขณะที่องค์ประกอบแรกมีความสามารถอธิบายความแปรปรวนของแบบทดสอบคิดเป็นร้อยละ 76.0 ของความแปรปรวนทั้งหมด ซึ่งถือว่าสามารถอธิบายความแปรปรวนของแบบทดสอบได้เกือบทั้งหมด จึงมีความเป็นมิติเดียว

ข้อตกลงเบื้องต้นเกี่ยวกับความเป็นมิติเดียวที่ว่า มีเพียงความสามารถเดียวเท่านั้นที่ถูกวัดโดยชุดของข้อคำถามในแบบทดสอบฉบับหนึ่งนั้น เป็นเงื่อนไขที่ไม่สามารถเข้มงวดได้มากนัก เพราะว่าองค์ประกอบด้านการคิด และบุคลิกภาพ อันได้แก่ แรงจูงใจ ความวิตกกังวล ความสามารถในการทำงาน แนวโน้มการเดา มักมีอิทธิพลต่อผลการสอบ ดังนั้นสิ่งที่ต้องการสำหรับความเป็นมิติเดียว ก็คือ ข้อมูลผลการสอบแสดงให้เห็นองค์ประกอบที่เด่นที่จะมีอิทธิพลต่อผลการสอบ ซึ่งถือว่าความสามารถดังกล่าวถูกวัดโดยแบบทดสอบฉบับดังกล่าว ฉะนั้น ในกรณีนี้จึงอนุมานได้ว่า คะแนนผลการสอบแบบทดสอบเหล่านี้มีคุณสมบัติความเป็นมิติ

เดียว และสามารถนำไปวิเคราะห์ตามหลักการของทฤษฎีการตอบสนองข้อสอบได้ และเมื่อแบบทดสอบมีคุณสมบัติด้านความเป็นมิติเดียวแล้ว แสดงว่าแบบทดสอบมีคุณสมบัติความเป็นอิสระในการตอบข้อสอบด้วย (Hambleton and Swaminathan. 1985 : 24)

2.2 ค่าสถิติพื้นฐานของค่าพารามิเตอร์แบบทดสอบชนิดต่าง ๆ

เพื่อให้ทราบข้อมูลเกี่ยวกับข้อคำถามซึ่งจะเป็นประโยชน์ต่อการทำความเข้าใจ และปรับปรุงข้อคำถาม ผู้วิจัยได้ทำการวิเคราะห์คะแนนผลการสอบเพื่อศึกษาคุณภาพของข้อคำถาม รายข้อและของแบบทดสอบ ด้วยโปรแกรม BILOG3 ได้ค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าการเดา (c) ของข้อสอบรายข้อในแต่ละฉบับ และตัดข้อที่ไม่เหมาะสมกับโมเดลออกไป ซึ่งได้นำเสนอในภาคผนวก ข ตาราง 56 และ 57 และวิเคราะห์ค่าสถิติพื้นฐานของค่าพารามิเตอร์ของแบบทดสอบฉบับต่าง ๆ ปรากฏผลเกี่ยวกับค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานของค่าความยาก ค่าอำนาจจำแนก และค่าการเดา ดังตาราง 14, 15, 16, 17, 18 และ 19

ตาราง 14 ค่าสถิติพื้นฐานของค่าอำนาจจำแนก (a) ของแบบทดสอบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3

ชนิดของข้อคำถาม	รูปแบบ	N	n	$\bar{a}$	a_{sd}
AR	เลือกตอบ	241	20	0.868	0.363
	เติมคำตอบ	155		0.888	0.309
AX	เลือกตอบ	277	12	0.706	0.174
	เติมคำตอบ	192		1.020	0.337
CV	เลือกตอบ	293	11	0.766	0.186
	เติมคำตอบ	153		0.873	0.201
LR	เลือกตอบ	237	17	0.673	0.211
	เติมคำตอบ	201		1.876	0.225
NLR	เลือกตอบ	258	13	0.754	0.187
	เติมคำตอบ	195		0.900	0.213
PI	เลือกตอบ	222	19	2.105	0.821
	เติมคำตอบ	177		1.790	0.782

จากตาราง 14 พบว่า สำหรับกลุ่มตัวอย่างที่เป็นนักเรียน ชั้น ม.3 ค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบรูปแบบเลือกตอบมีค่าระหว่าง 0.673 ถึง 2.105 ขณะที่แบบเติมคำตอบมีค่าระหว่าง 0.873 ถึง 1.790 ทั้งนี้โดยค่าเฉลี่ยอำนาจจำแนกของแบบทดสอบรูปแบบเติมคำตอบทุกชนิดมีค่าสูงกว่าแบบเลือกตอบทุกชนิด ยกเว้น ชนิด PI ที่แบบเลือกตอบมีค่าสูงกว่า

ฉบับที่มีค่าอำนาจจำแนกเฉลี่ยสูงสุด คือ ฉบับ PI ทั้ง 2 รูปแบบ ส่วนฉบับที่มีค่าอำนาจจำแนกเฉลี่ยต่ำสุดสำหรับแบบเลือกตอบ คือ ฉบับ LR และฉบับที่ต่ำที่สุดของแบบเติมคำตอบ คือ ฉบับ CV

ตาราง 15 ค่าสถิติพื้นฐานของค่าอำนาจจำแนก (a) ที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5

ชนิดของข้อคำถาม	รูปแบบ	N	n	$\bar{a}$	a_{SD}
AR	เลือกตอบ	230	20	0.838	0.327
	เติมคำตอบ	155		0.838	0.259
AX	เลือกตอบ	277	12	0.725	0.159
	เติมคำตอบ	192		0.975	0.459
CV	เลือกตอบ	293	11	0.692	0.204
	เติมคำตอบ	153		0.958	0.279
LR	เลือกตอบ	237	17	0.568	0.108
	เติมคำตอบ	201		0.937	0.309
NLR	เลือกตอบ	258	13	0.595	0.173
	เติมคำตอบ	195		1.104	0.308
PI	เลือกตอบ	222	19	0.591	0.412
	เติมคำตอบ	177		2.113	1.040

จากตาราง 15 พบว่า สำหรับกลุ่มตัวอย่างที่เป็นนักเรียน ชั้น ม.5 ค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบรูปแบบเลือกตอบมีค่าระหว่าง 0.568 ถึง 1.591 ขณะที่แบบเติมคำตอบมีค่าระหว่าง 0.838 ถึง 2.113 ทั้งนี้โดยแบบเติมคำตอบชนิด AX, CV, LR, NLR และ PI มีค่าเฉลี่ยอำนาจจำแนกสูงกว่าแบบเลือกตอบ ส่วนฉบับ AR ทั้ง 2 รูปแบบ มีค่าอำนาจจำแนกเฉลี่ยเท่ากัน

ฉบับที่มีค่าอำนาจจำแนกเฉลี่ยสูงสุด คือ ฉบับ PI ทั้ง 2 รูปแบบ ส่วนฉบับที่มีค่าอำนาจจำแนกเฉลี่ยต่ำสุดสำหรับแบบเลือกตอบ คือ ฉบับ LR และฉบับที่มีค่าต่ำที่สุดของแบบเติมคำตอบ คือ ฉบับ AR

ตาราง 16 ค่าสถิติพื้นฐานของค่าความยาก (b) ของแบบทดสอบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3

ชนิดของข้อคำถาม	รูปแบบ	N	n	$\bar{b}$	b_{SD}
AR	เลือกตอบ	230	20	0.915	1.155
	เติมคำตอบ	155		0.432	0.986
AX	เลือกตอบ	277	12	1.533	0.987
	เติมคำตอบ	192		1.215	0.685
CV	เลือกตอบ	293	11	0.371	0.743
	เติมคำตอบ	153		0.610	0.576
LR	เลือกตอบ	237	17	2.030	1.651
	เติมคำตอบ	201		1.619	0.709
NLR	เลือกตอบ	258	13	1.414	0.924
	เติมคำตอบ	195		1.397	0.722
PI	เลือกตอบ	222	19	0.014	0.313
	เติมคำตอบ	177		-0.436	0.488

จากตาราง 16 พบว่า สำหรับกลุ่มตัวอย่างที่เป็นนักเรียน ชั้น ม.3 ค่าความยากเฉลี่ยของแบบทดสอบรูปแบบเลือกตอบมีค่าระหว่าง 0.014 ถึง 2.030 ขณะที่แบบเติมคำตอบมีค่าความยากเฉลี่ยระหว่าง -0.436 ถึง 1.619 ฉบับที่มีค่าความยากสูงสุด คือ ฉบับ LR ทั้ง 2 รูปแบบ และฉบับที่มีค่าความยากต่ำสุด ได้แก่ ฉบับ PI ทั้ง 2 รูปแบบอีกเช่นกัน

แบบทดสอบฉบับ AR, AX, LR, NLR และ PI แบบเลือกตอบมีค่าความยากเฉลี่ยสูงกว่าแบบเติมคำตอบ

ตาราง 17 ค่าสถิติพื้นฐานของค่าความยาก (b) ของแบบทดสอบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5

ชนิดของข้อคำถาม	รูปแบบ	N	n	$\bar{b}$	b_{sd}
AR	เลือกตอบ	230	20	0.778	1.182
	เติมคำตอบ	155		0.039	0.756
AX	เลือกตอบ	277	12	1.225	0.964
	เติมคำตอบ	192		0.633	0.614
CV	เลือกตอบ	293	11	-0.004	0.828
	เติมคำตอบ	153		0.215	0.513
LR	เลือกตอบ	237	17	1.847	1.647
	เติมคำตอบ	201		1.423	0.747
NLR	เลือกตอบ	258	13	1.483	1.391
	เติมคำตอบ	195		1.644	1.329
PI	เลือกตอบ	222	19	-0.436	0.320
	เติมคำตอบ	177		-0.657	0.693

จากตาราง 17 พบว่า สำหรับกลุ่มตัวอย่างที่เป็นนักเรียน ชั้น ม.5 ค่าความยากเฉลี่ยของแบบทดสอบรูปแบบเลือกตอบมีค่าระหว่าง -0.436 ถึง 1.847 ขณะที่แบบเติม-คำตอบมีค่าความยากเฉลี่ยระหว่าง -0.657 ถึง 1.644 ทั้งนี้โดยแบบทดสอบฉบับ AR, AX, LR และ PI มีค่าความยากเฉลี่ยสูงกว่าแบบเติมคำตอบ ส่วนฉบับ CV และ NLR แบบเติม-คำตอบมีค่าความยากเฉลี่ยสูงกว่าแบบเลือกตอบ ฉบับที่มีค่าความยากสูงสุดของแบบเลือกตอบคือ ฉบับ LR ค่าต่ำสุด คือ ฉบับ CV สำหรับแบบเติมคำตอบ ฉบับที่มีค่าความยากเฉลี่ยสูงสุดคือ ฉบับ NLR ค่าต่ำสุด คือ ฉบับ PI

ตาราง 18 ค่าสถิติพื้นฐานของค่าสัมประสิทธิ์การเดา (c) ของแบบทดสอบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3

ชนิดของข้อคำถาม	รูปแบบ	N	n	$\bar{c}$	c_{SD}
AR	เลือกตอบ	230	20	0.144	0.034
	เติมคำตอบ	155		0.081	0.031
AX	เลือกตอบ	277	12	0.202	0.706
	เติมคำตอบ	192		0.110	0.051
CV	เลือกตอบ	293	11	0.112	0.034
	เติมคำตอบ	153		0.199	0.050
LR	เลือกตอบ	237	17	0.184	0.055
	เติมคำตอบ	201		0.121	0.028
NLR	เลือกตอบ	258	13	0.203	0.039
	เติมคำตอบ	195		0.184	0.053
PI	เลือกตอบ	222	19	0.157	0.048
	เติมคำตอบ	177		0.029	0.008

จากตาราง 18 พบว่า สำหรับกลุ่มตัวอย่างที่เป็นนักเรียน ชั้น ม.3 ค่าการเดาเฉลี่ยของแบบทดสอบรูปแบบเลือกตอบมีค่าระหว่าง 0.112 ถึง 0.203 ขณะที่แบบเติมคำตอบมีค่าการเดาเฉลี่ยระหว่าง 0.029 ถึง 0.199 ทั้งนี้โดยแบบทดสอบเลือกตอบชนิด AR, AX, LR NLR และ PI มีค่าการเดาเฉลี่ยสูงกว่าแบบเติมคำตอบ สำหรับแบบเลือกตอบฉบับที่มีค่าการเดาเฉลี่ยสูงสุด คือ ฉบับ NLR ฉบับที่มีค่าต่ำสุด คือ CV

ส่วนแบบเติมคำตอบมีค่าการเดาสูงสุด คือ ฉบับ CV และฉบับที่มีค่าการเดาเฉลี่ยต่ำสุด คือ ฉบับ PI

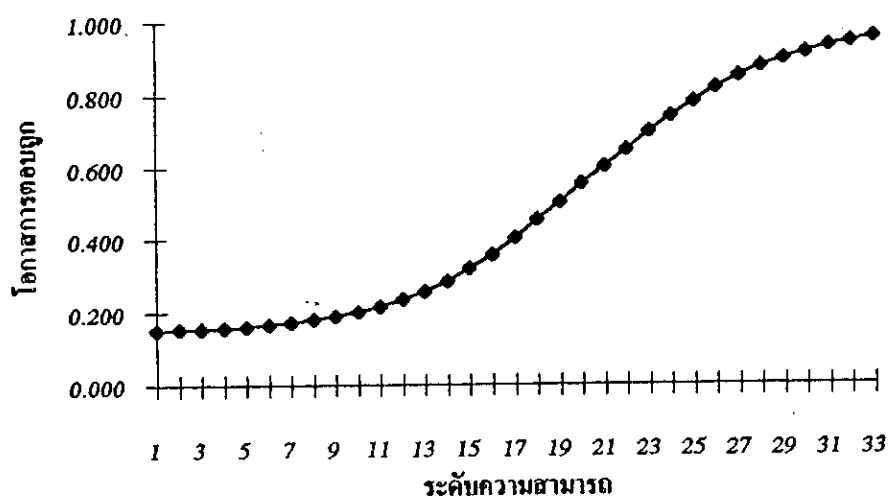
ตาราง 19 ค่าสถิติพื้นฐานของค่าสัมประสิทธิ์การเดา (c) ของแบบทดสอบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5

ชนิดของข้อคำถาม	รูปแบบ	N	n	$\bar{c}$	c_{sd}
AR	เลือกตอบ	230	20	0.093	0.030
	เติมคำตอบ	155		0.133	0.026
AX	เลือกตอบ	277	12	0.227	0.067
	เติมคำตอบ	192		0.152	0.030
CV	เลือกตอบ	293	11	0.157	0.023
	เติมคำตอบ	153		0.292	0.058
LR	เลือกตอบ	237	17	0.244	0.066
	เติมคำตอบ	201		0.200	0.080
NLR	เลือกตอบ	258	13	0.167	0.043
	เติมคำตอบ	195		0.281	0.087
PI	เลือกตอบ	222	19	0.144	0.044
	เติมคำตอบ	177		0.143	0.100

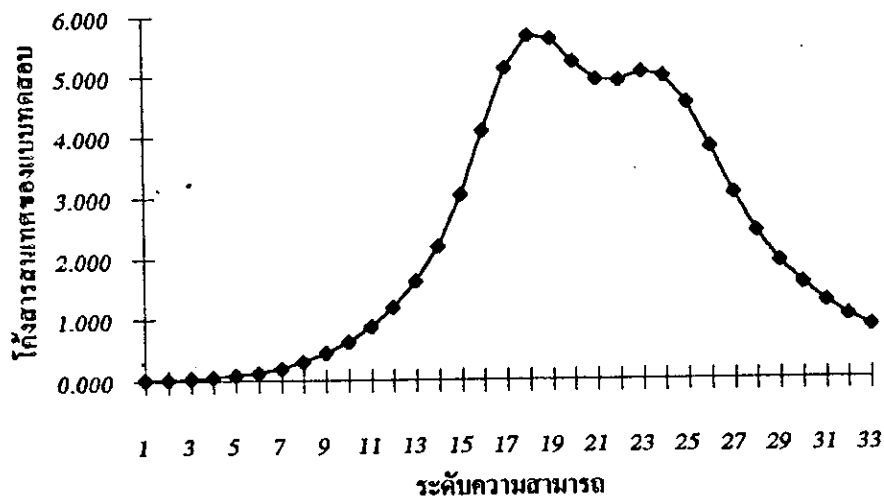
จากตาราง 19 พบว่า สำหรับกลุ่มตัวอย่างที่เป็นนักเรียน ชั้น ม.5 ค่าการเดาเฉลี่ยของแบบทดสอบรูปแบบเลือกตอบมีค่าระหว่าง 0.093 ถึง 0.244 ขณะที่แบบเติมคำตอบมีค่าการเดาเฉลี่ยระหว่าง 0.133 ถึง 0.292 ทั้งนี้โดยแบบทดสอบเลือกตอบชนิด AX, CV, LR และ PI มีค่าการเดาเฉลี่ยสูงกว่าแบบเติมคำตอบ ฉบับเติมคำตอบที่มีค่าการเดาเฉลี่ยสูงกว่าแบบเลือกตอบ คือ AR และ NLR สำหรับแบบเลือกตอบฉบับที่มีค่าการเดาเฉลี่ยสูงสุด คือ ฉบับ LR ฉบับที่มีค่าต่ำสุด คือ AR ส่วนแบบเติมคำตอบที่มีค่าการเดาสูงสุด คือ ฉบับ CV และฉบับที่มีค่าการเดาเฉลี่ยต่ำสุด คือ ฉบับ AR

2.3 ค่าฟังก์ชันสารสนเทศของแบบทดสอบ (TIF) และความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ (SE_o)

ค่าฟังก์ชันสารสนเทศของแบบทดสอบ ถือได้ว่าเป็นดัชนีที่ใช้วัดความแม่นยำในการประมาณค่าความสามารถของผู้สอบที่ได้จากผลรวมของค่าฟังก์ชันสารสนเทศของข้อสอบ (IIF) และเพื่อให้เกิดความชัดเจนในการศึกษา และพิจารณาแนวโน้มของระดับความถูกต้องแม่นยำในการประมาณค่าความสามารถ ผู้วิจัยจึงนำเสนอค่าฟังก์ชันสารสนเทศของแบบทดสอบแต่ละฉบับในรูปกราฟ รวมทั้งเสนอโค้งลักษณะของแบบทดสอบ (TCC) และโค้งลักษณะข้อสอบ (ICC) ของข้อสอบรายข้อ ปรากฏผลดังภาพประกอบ 21 - 92 และตาราง 20 - 21 แสดงค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของแบบทดสอบแต่ละฉบับ สำหรับรายละเอียดดังปรากฏในภาคผนวก ข

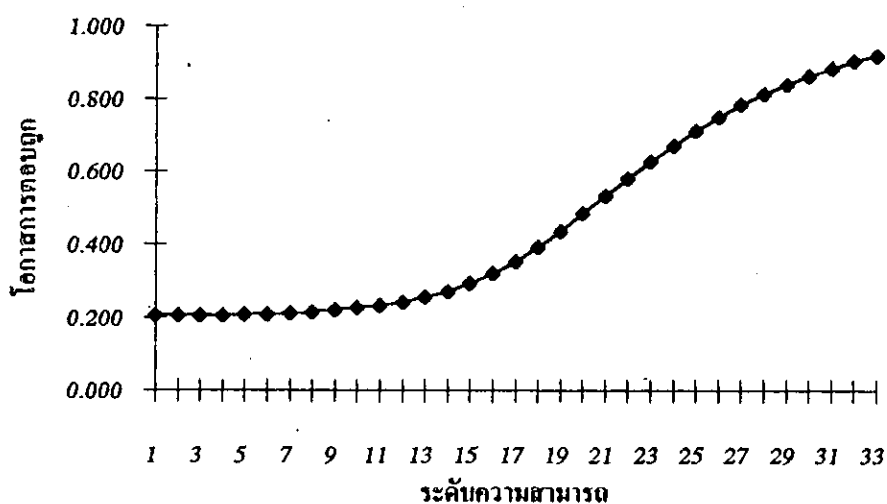


ภาพประกอบ 21 โค้งลักษณะของแบบทดสอบ ชนิด AR แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

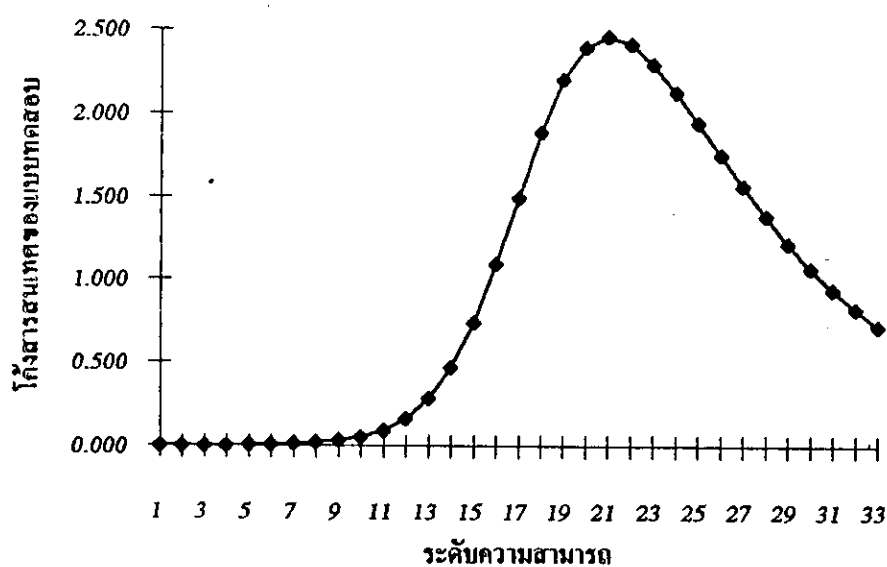


$$AR_{9.1} \quad I(\max) = 5.678 \quad \text{ที่} \quad \phi(\max) = 0.250$$

ภาพประกอบ 22 กำลังสารสนเทศของแบบทดสอบ ชนิด AR แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

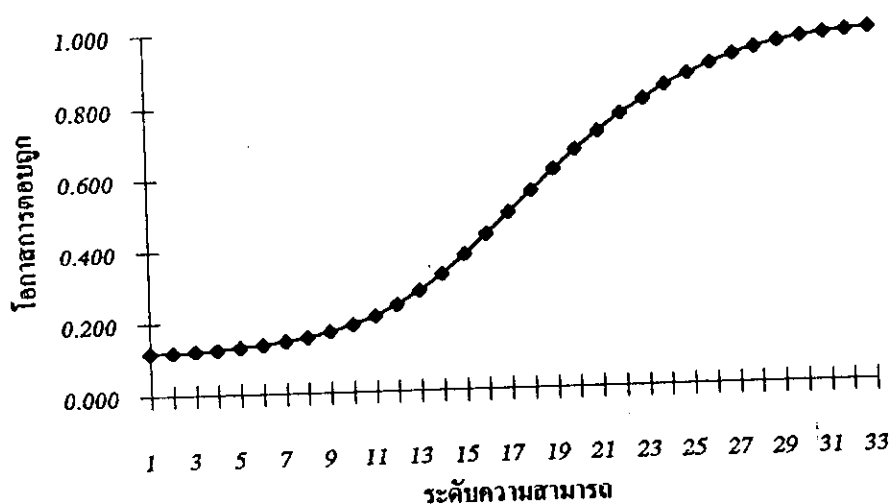


ภาพประกอบ 23 โค้งลักษณะของแบบทดสอบ ชนิด AX แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

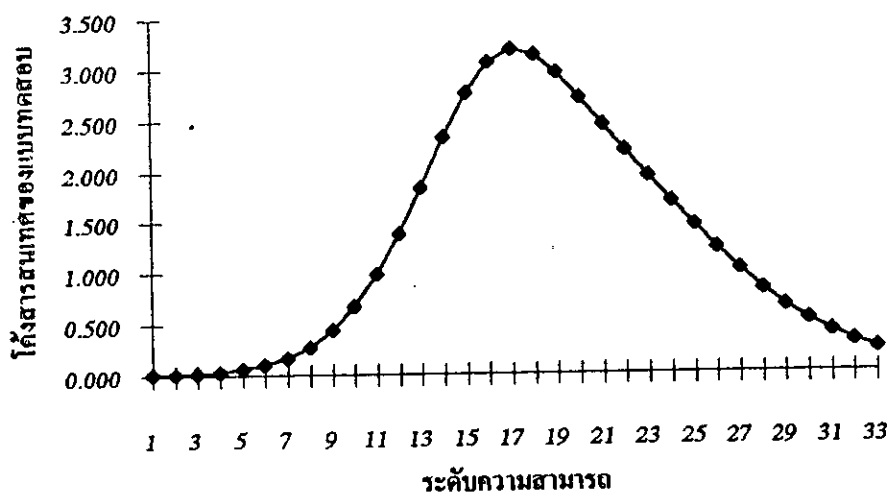


$$AX_{51} \quad I(\max) = 2.453 \quad \text{ที่} \quad \theta(\max) = 1.000$$

ภาพประกอบ 24 โค้งสารสนเทศของแบบทดสอบ ชนิด AX แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

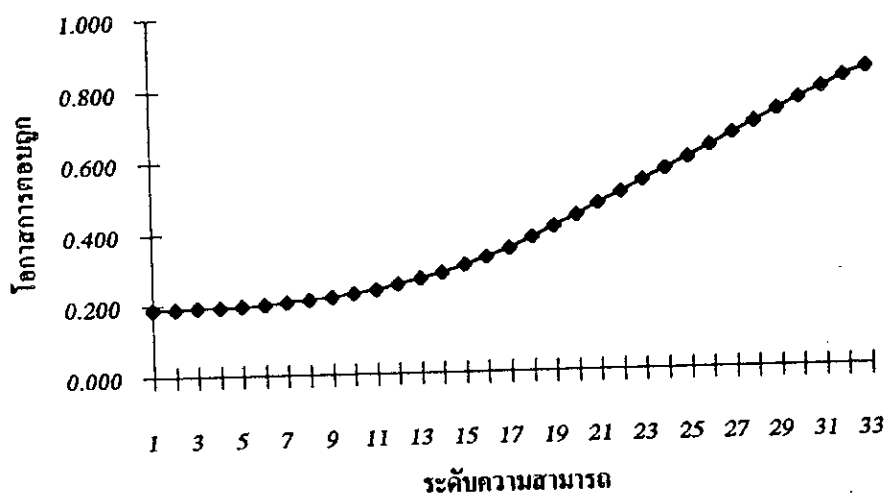


ภาพประกอบ 25 โค้งลักษณะของแบบทดสอบ ชนิด CV แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

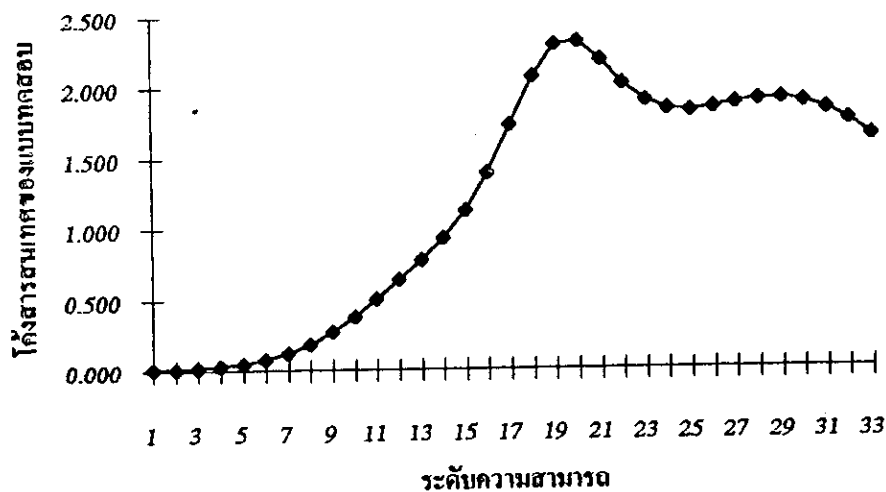


$$CV_{31} I(\max) = 3.184 \text{ ที่ } o(\max) = 0.000$$

ภาพประกอบ 26 โค้งสารสนเทศของแบบทดสอบ ชนิด CV แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

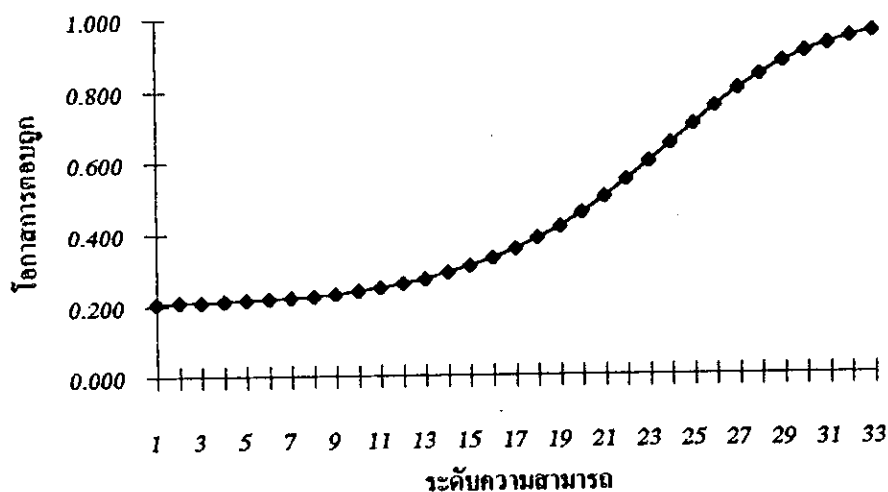


ภาพประกอบ 27 โค้งลักษณะของแบบทดสอบ ชนิด LR แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

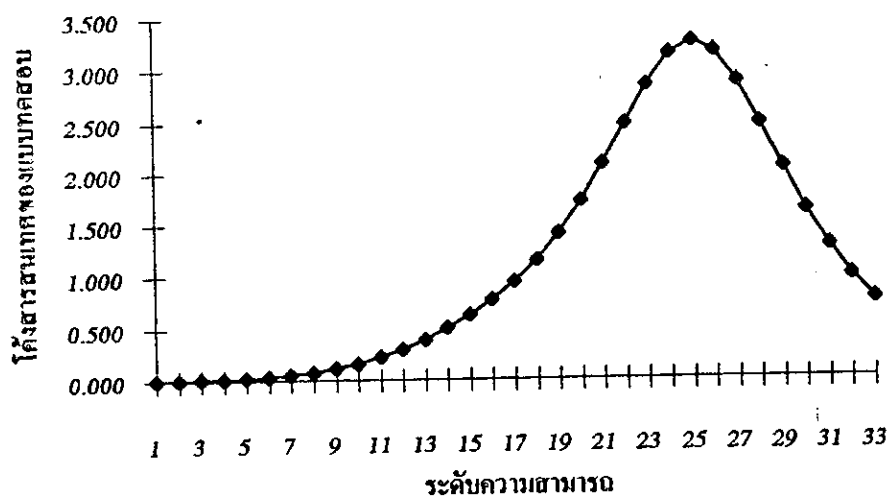


$$LR_{31}, I(\max) = 2.310 \text{ ที่ } \phi(\max) = 0.750$$

ภาพประกอบ 28 โค้งสารสนเทศของแบบทดสอบ ชนิด LR แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

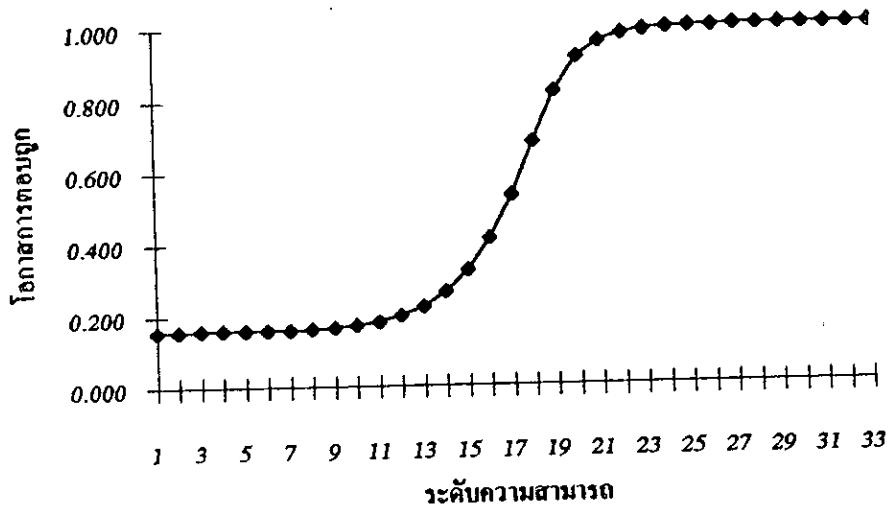


ภาพประกอบ 29 โค้งลักษณะของแบบทดสอบ ชนิด NLR แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

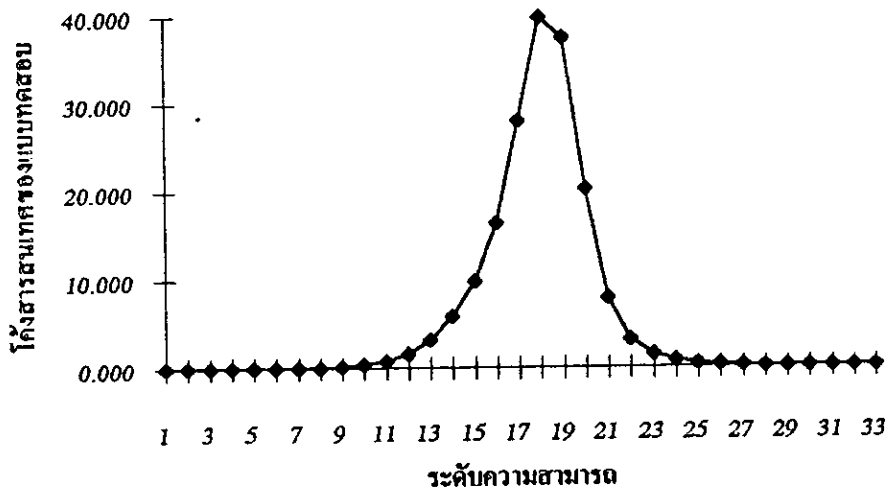


$$NLR_{3.1} I(\max) = 3.266 \text{ ที่ } \theta(\max) = 2.000$$

ภาพประกอบ 30 โค้งสารสนเทศของแบบทดสอบ ชนิด NLR แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

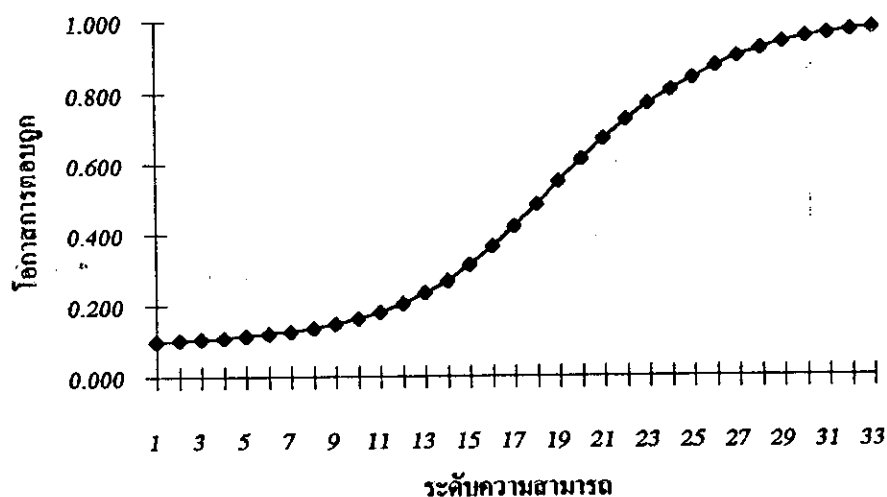


ภาพประกอบ 31 โค้งลักษณะของแบบทดสอบ ชนิด PI แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

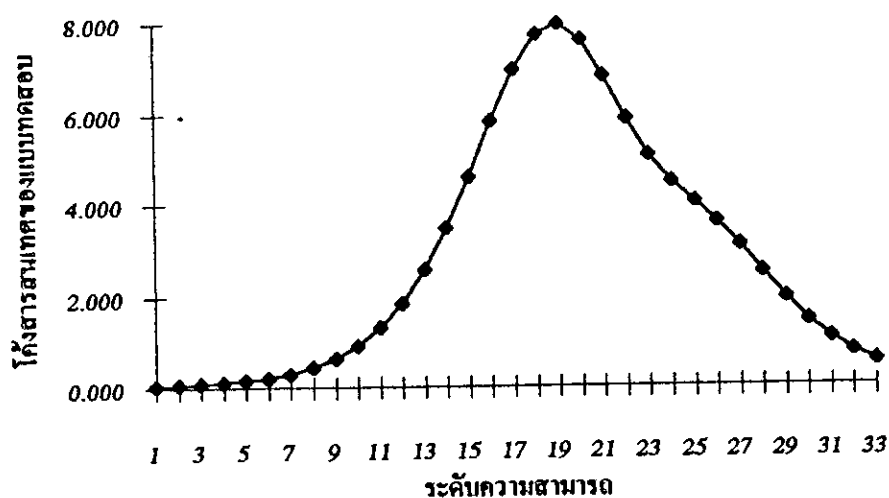


$$PI_{3,1} I(\max) = 39.803 \text{ ที่ } \theta(\max) = 0.250$$

ภาพประกอบ 32 โค้งสารสนเทศของแบบทดสอบ ชนิด PI แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

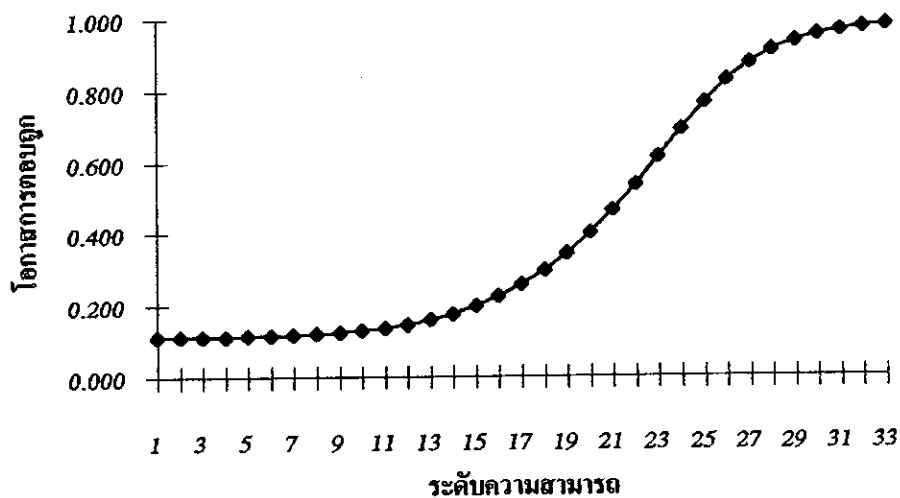


ภาพประกอบ 33 ได้ังลักษณะของแบบทดสอบ ชนิด AR แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

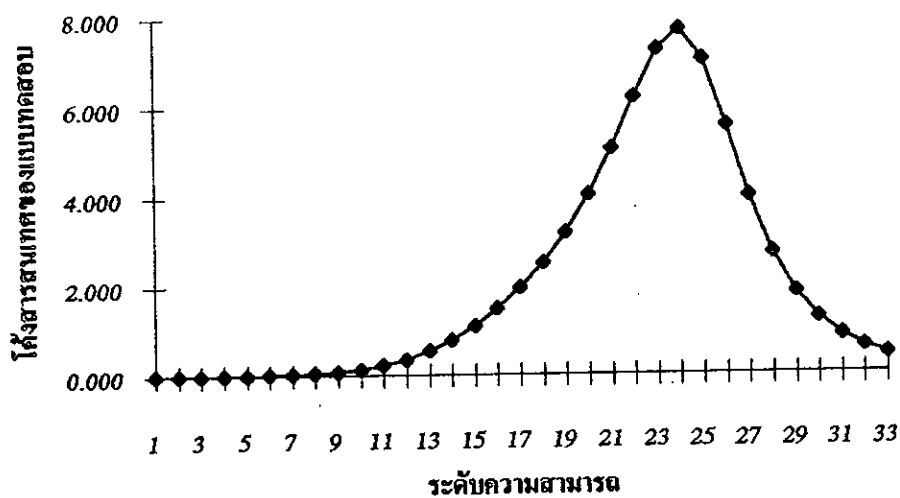


$$AR_{32} I(\max) = 7.954 \text{ ที่ } o(\max) = 0.500$$

ภาพประกอบ 34 ได้ังสาร์สนเทศของแบบทดสอบ ชนิด AR แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

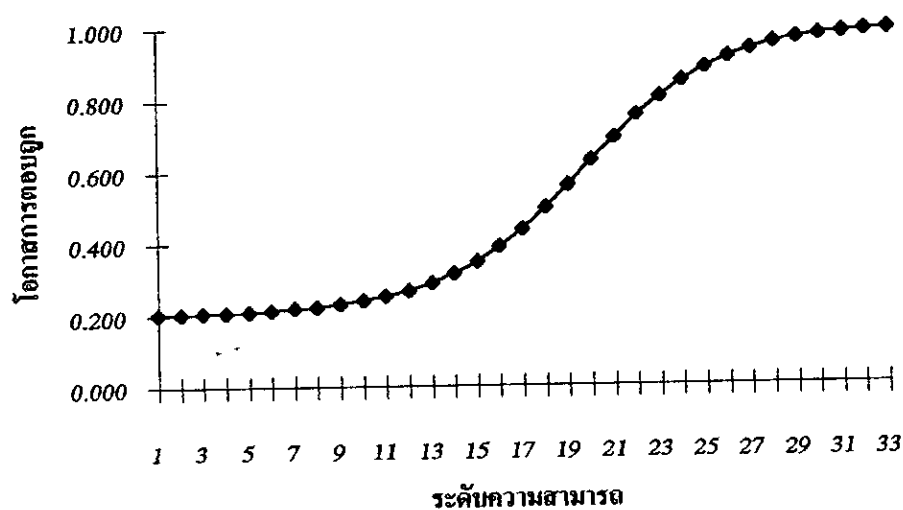


ภาพประกอบ 35 โค้งลักษณะของแบบทดสอบ ชนิด AX แบบเดิมค่าตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

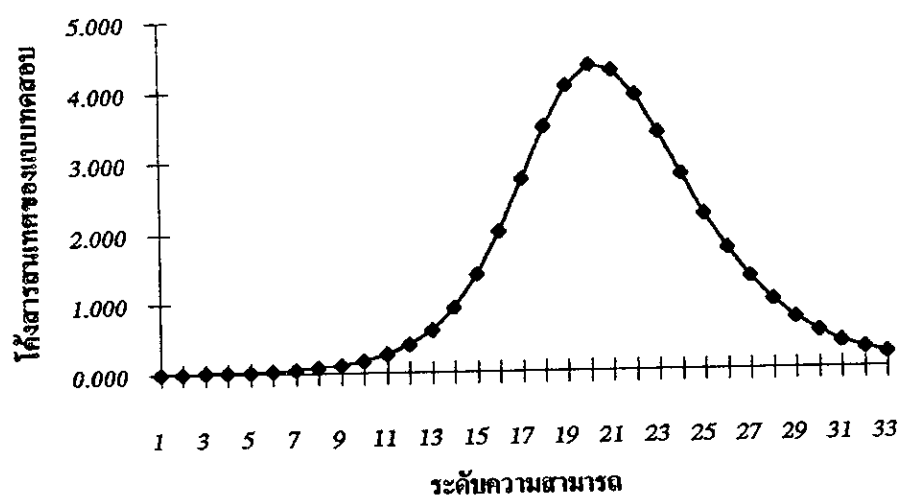


$$AX_{32} \quad I(\max) = 7.702 \quad \text{ที่} \quad \theta(\max) = 1.750$$

ภาพประกอบ 36 ได้สารสนเทศของแบบทดสอบ ชนิด AX แบบเดิมค่าตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

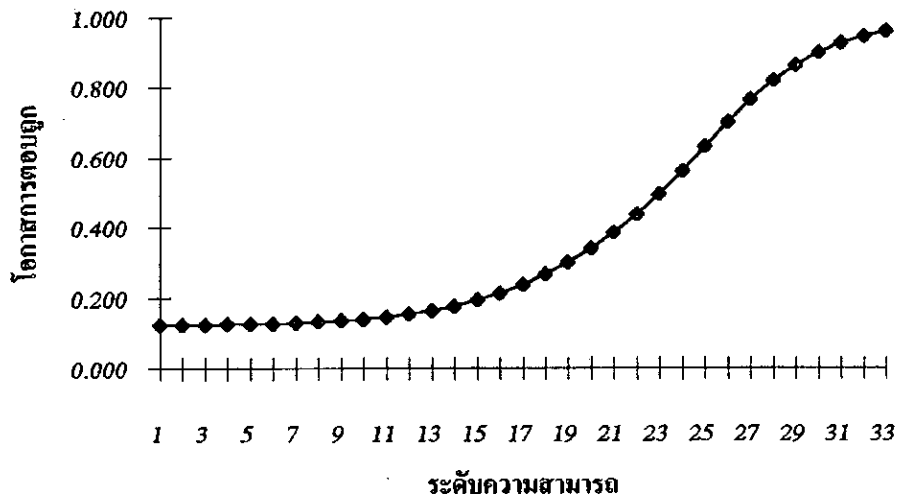


ภาพประกอบ 37 โค้งลักษณะของแบบทดสอบ ชนิด CV แบบเดิมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

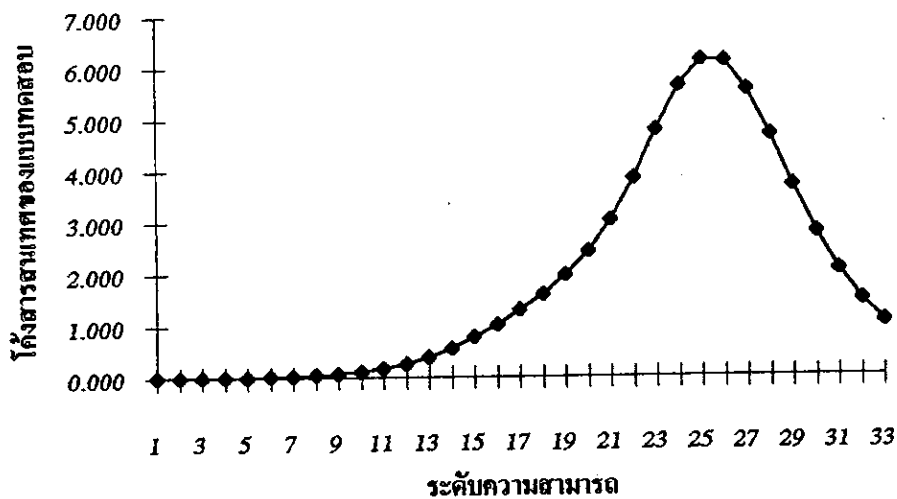


$$CV_{32} I(\max) = 4.322 \text{ ที่ } \theta(\max) = 0.750$$

ภาพประกอบ 38 โค้งสารสนเทศของแบบทดสอบ ชนิด CV แบบเดิมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

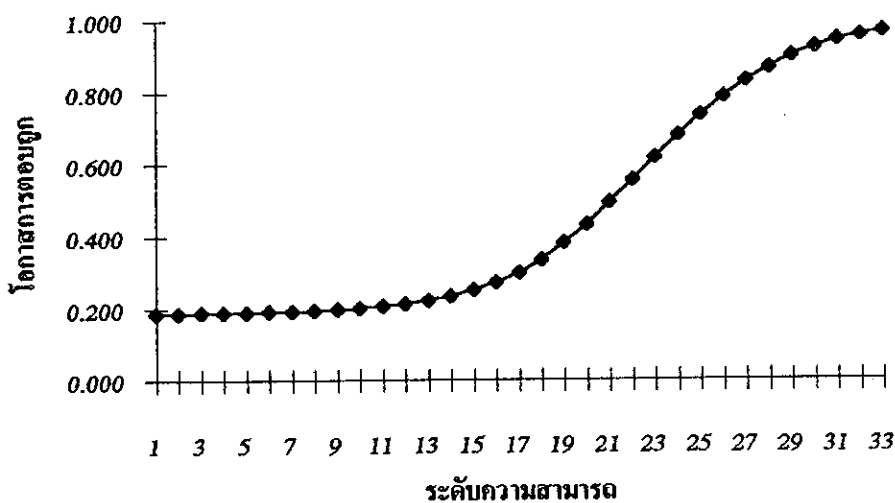


ภาพประกอบ 39 โด่งลักษณะของแบบทดสอบ ชนิด LR แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

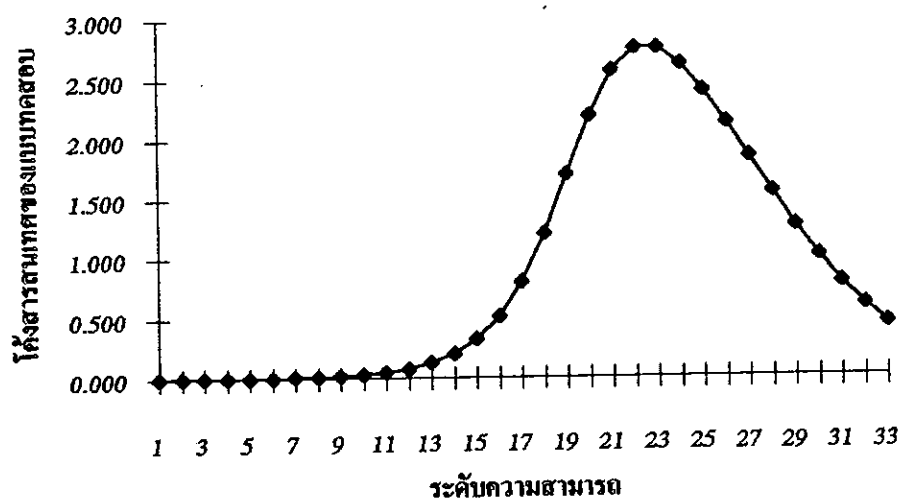


$$LR_{0.2} \quad I(\max) = 6.135 \text{ ที่ } \theta(\max) = 2.000$$

ภาพประกอบ 40 โด่งสารสนเทศของแบบทดสอบ ชนิด LR แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

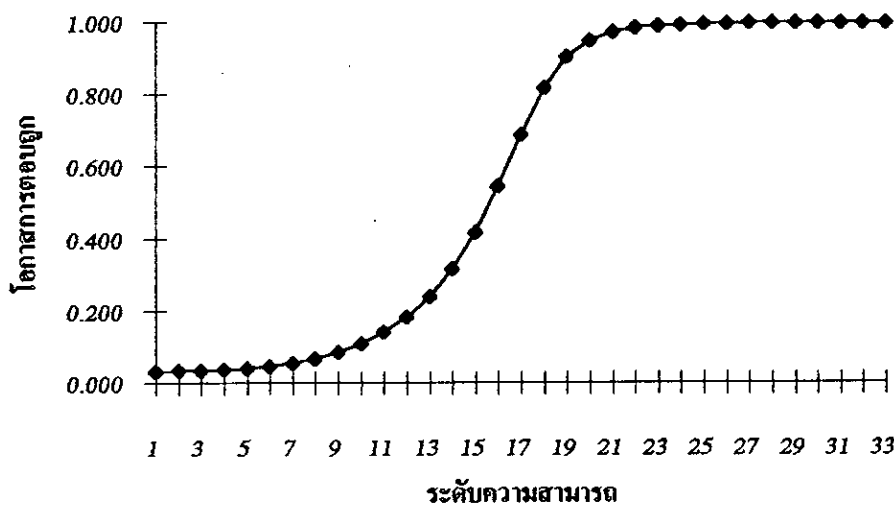


ภาพประกอบ 41 โด่งลักษณะของแบบทดสอบ ชนิด NLR แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

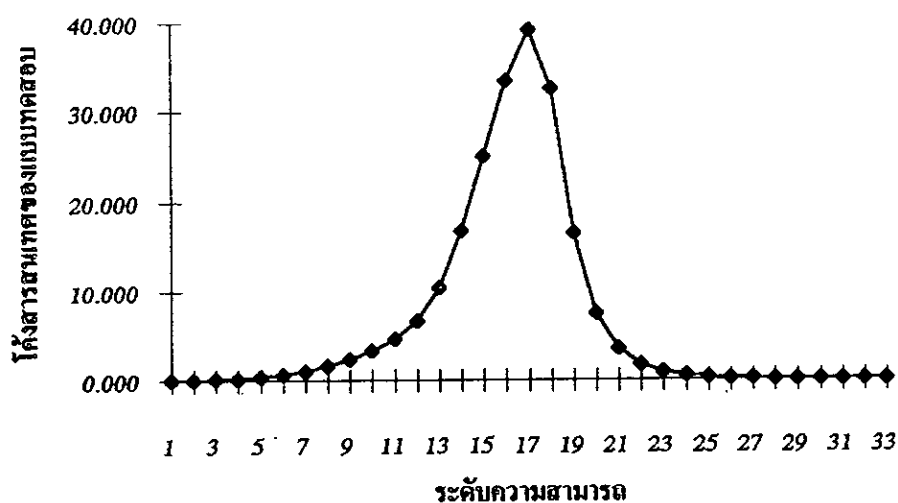


$$NLR_{32} I(\max) = 2.747 \text{ ที่ } \theta(\max) = 1.500$$

ภาพประกอบ 42 ได้สารสนเทศของแบบทดสอบ ชนิด NLR แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

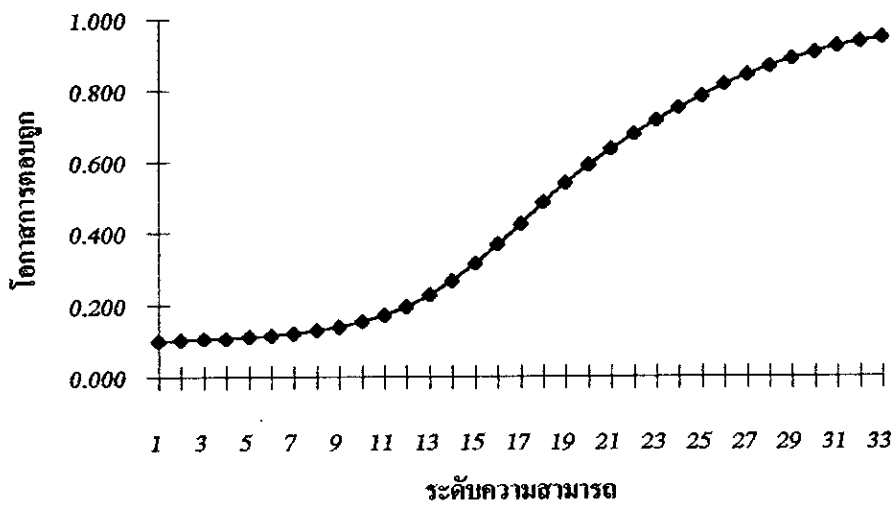


ภาพประกอบ 43 โค้งลักษณะของแบบทดสอบ ชนิด PI แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

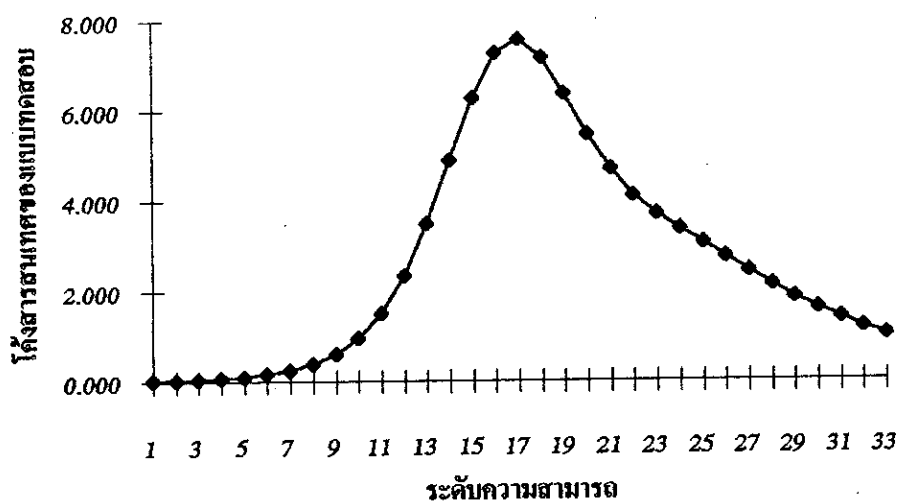


$$PI_{32} I(\max) = 39.2143 \text{ ที่ } o(\max) = 0.000$$

ภาพประกอบ 44 โค้งสารสนเทศของแบบทดสอบ ชนิด PI แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ณ ระดับความสามารถต่าง ๆ

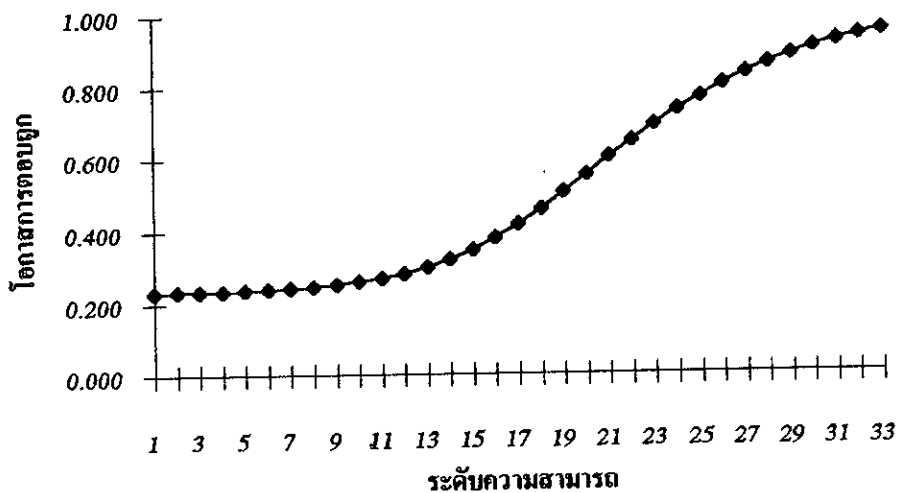


ภาพประกอบ 45 โค้งลักษณะของแบบทดสอบ ชนิด AR แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

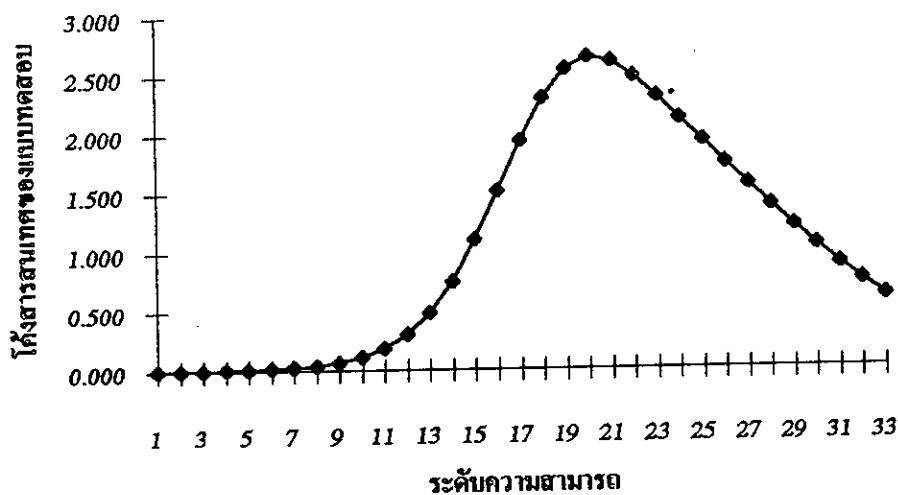


$$AR_{51} \quad I(\max) = 7.563 \text{ ที่ } \theta(\max) = 0.000$$

ภาพประกอบ 46 โค้งสารสนเทศของแบบทดสอบ ชนิด AR แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

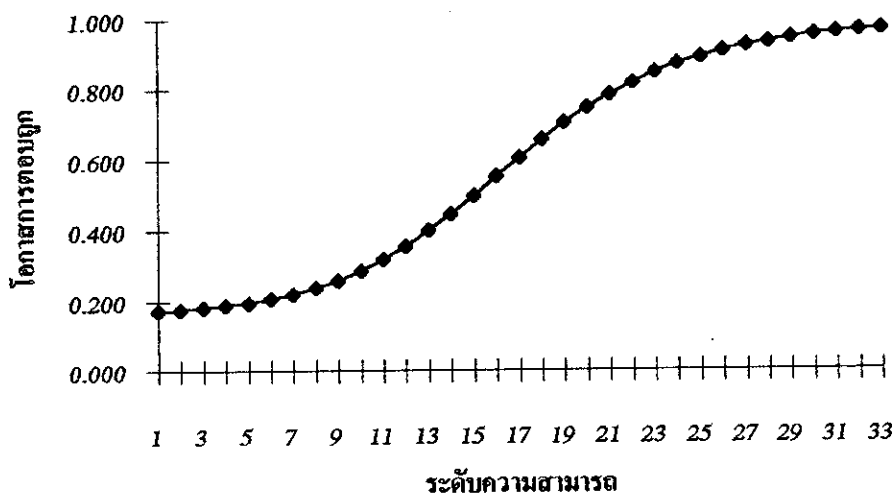


ภาพประกอบ 47 โค้งลักษณะของแบบทดสอบ ชนิด AX แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

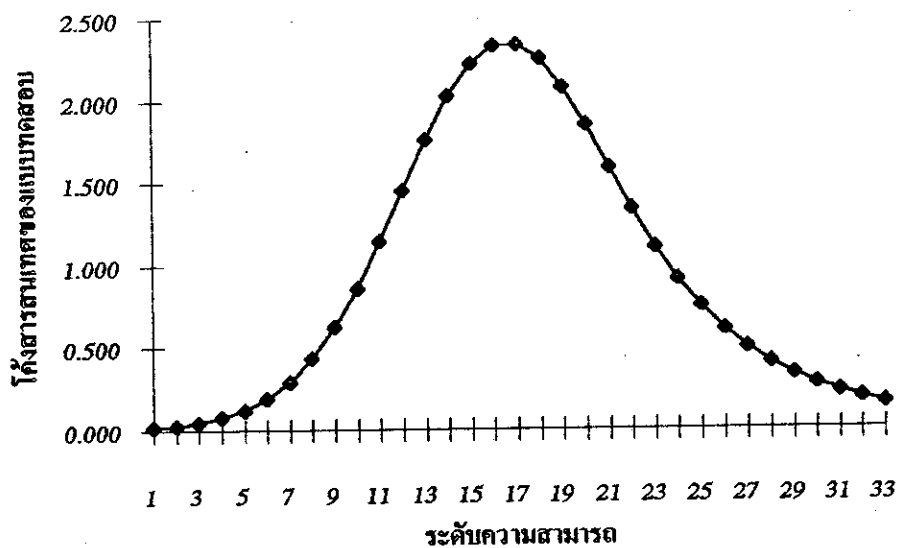


$$AX_{51} \quad I(\max) = 2.684 \quad \text{ที่} \quad \theta(\max) = 0.750$$

ภาพประกอบ 48 โค้งสารสนเทศของแบบทดสอบ ชนิด AX แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

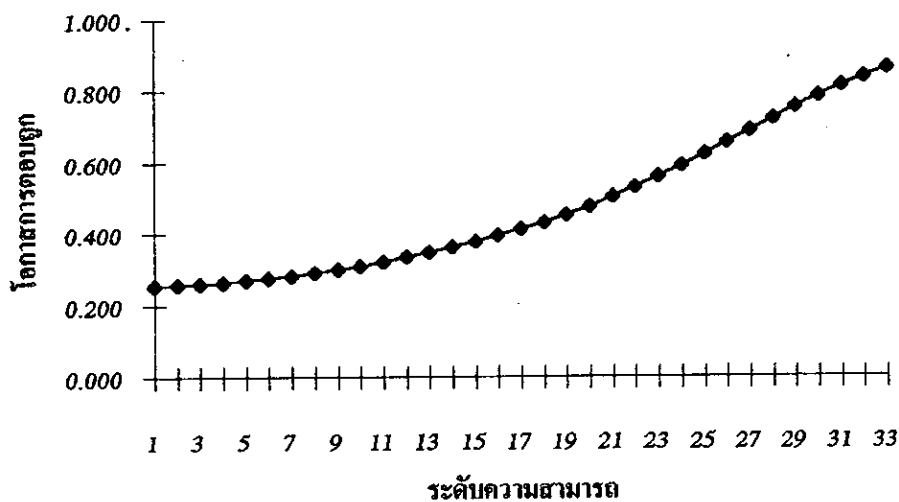


ภาพประกอบ 49 โค้งลักษณะของแบบทดสอบ ชนิด CV แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

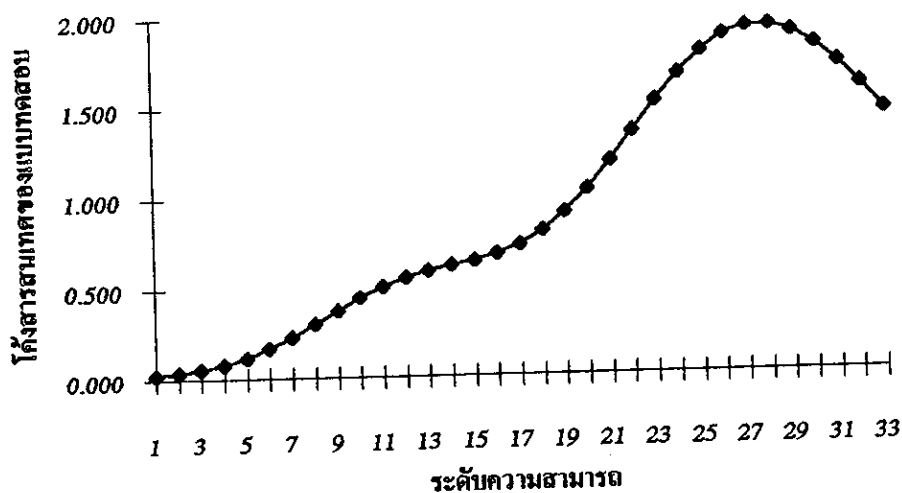


$$CV_{51} \quad I(\max) = 2.338 \quad \text{ที่} \quad o(\max) = 0.000$$

ภาพประกอบ 50 โค้งสารสนเทศของแบบทดสอบ ชนิด CV แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

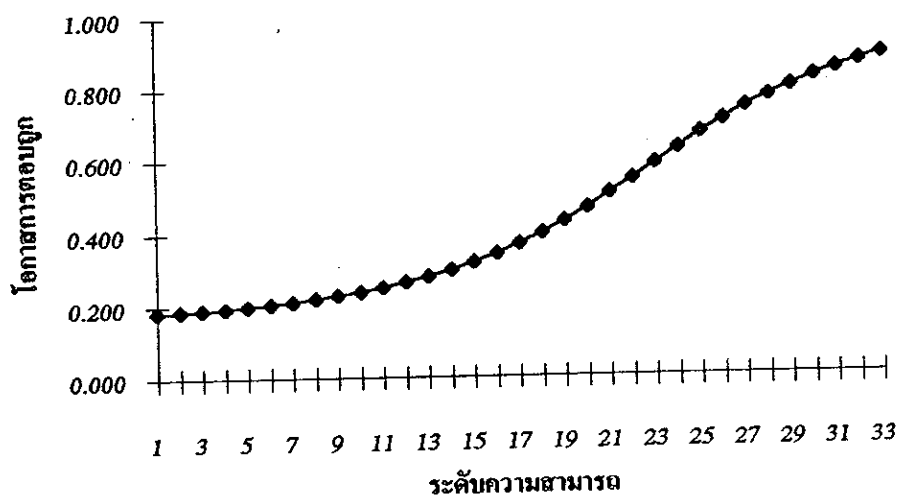


ภาพประกอบ 51 โค้งลักษณะของแบบทดสอบ ชนิด LR แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

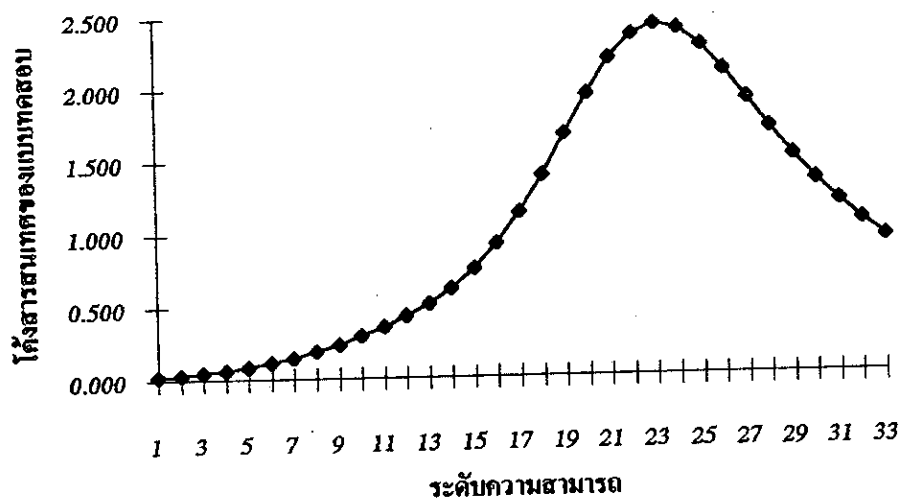


$$LR_{S_1} I(\max) = 1.916 \text{ ที่ } \theta(\max) = 2.750$$

ภาพประกอบ 52 โค้งสารสนเทศของแบบทดสอบ ชนิด LR แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

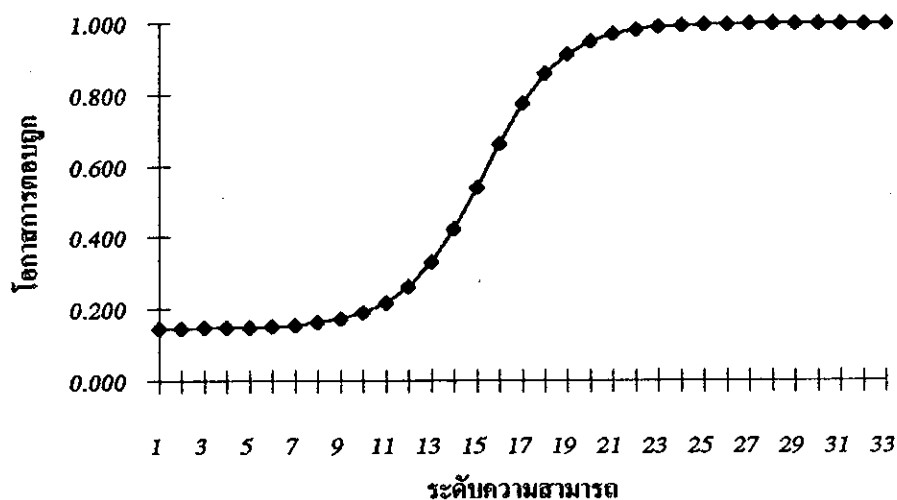


ภาพประกอบ 53 โค้งลักษณะของแบบทดสอบ ชนิด NLR แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

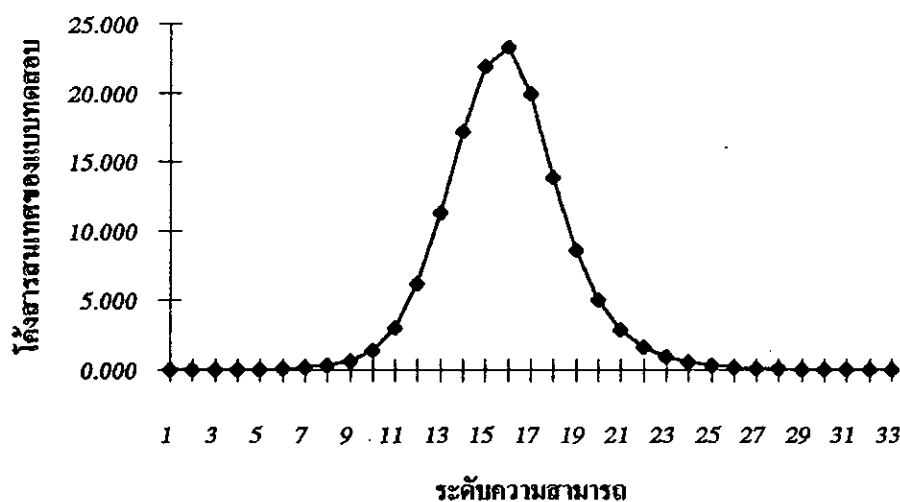


$$NLR_{51} \quad I(\max) = 2.419 \quad \text{ที่} \quad \theta(\max) = 1.500$$

ภาพประกอบ 54 โค้งสารสนเทศของแบบทดสอบ ชนิด NLR แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

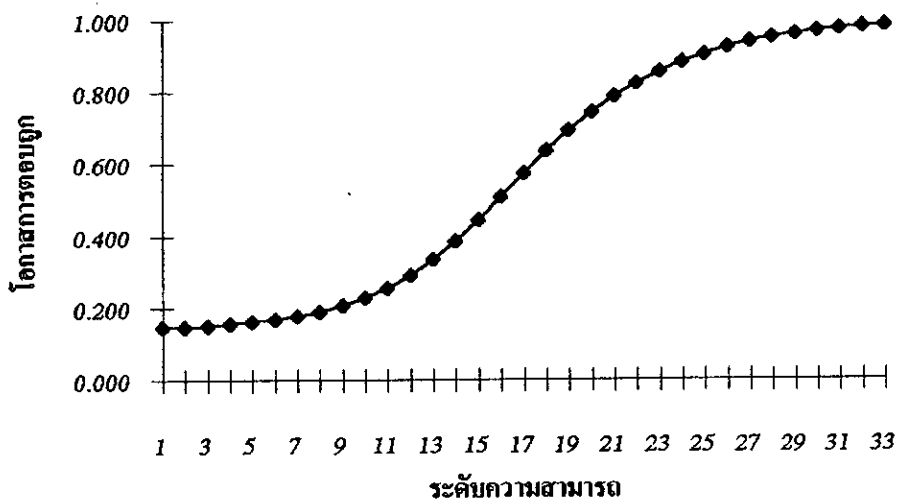


ภาพประกอบ 55 โค้งลักษณะของแบบทดสอบ ชนิด PI แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ



$$PI_{51} I(\max) = 23.316 \text{ ที่ } \theta(\max) = -0.250$$

ภาพประกอบ 56 โค้งสารสนเทศของแบบทดสอบ ชนิด PI แบบเลือกตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

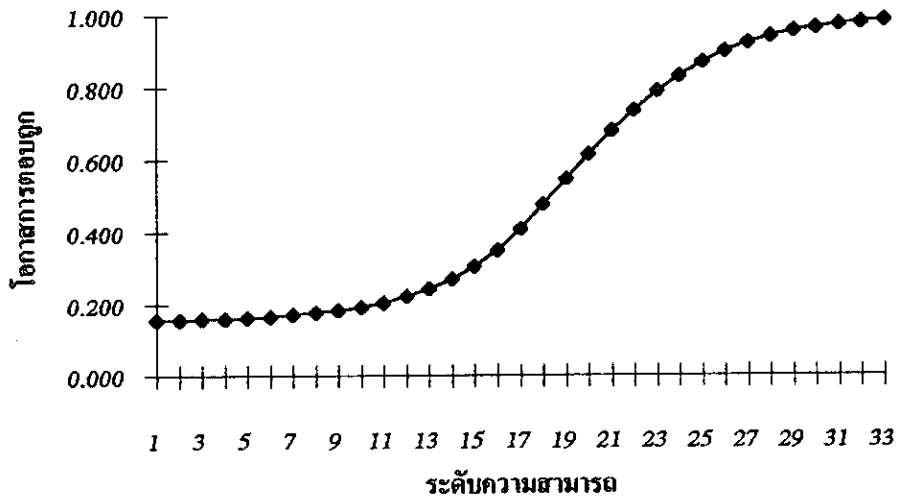


ภาพประกอบ 57 โค้งลักษณะของแบบทดสอบ ชนิด AR แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

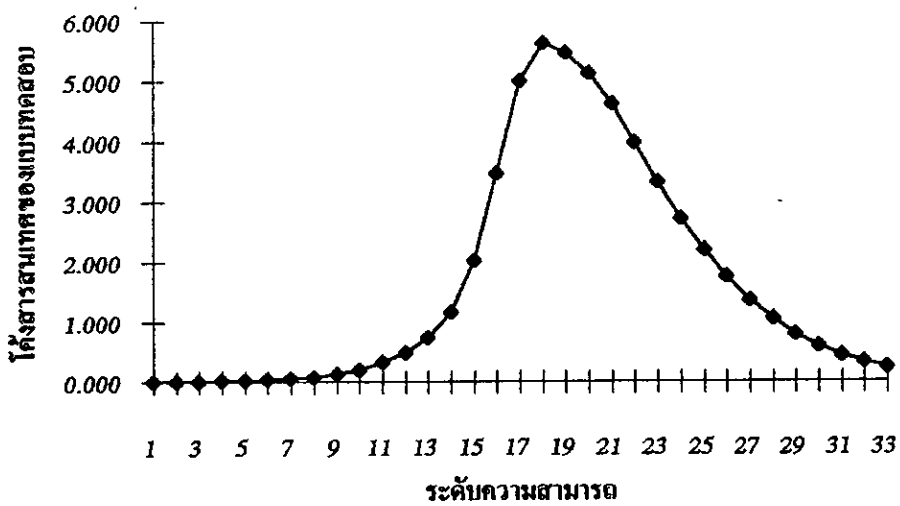


$$AR_{52} \quad I(\max) = 6.637 \text{ ที่ } \theta(\max) = 0.000$$

ภาพประกอบ 58 โค้งสารสนเทศของแบบทดสอบ ชนิด AR แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

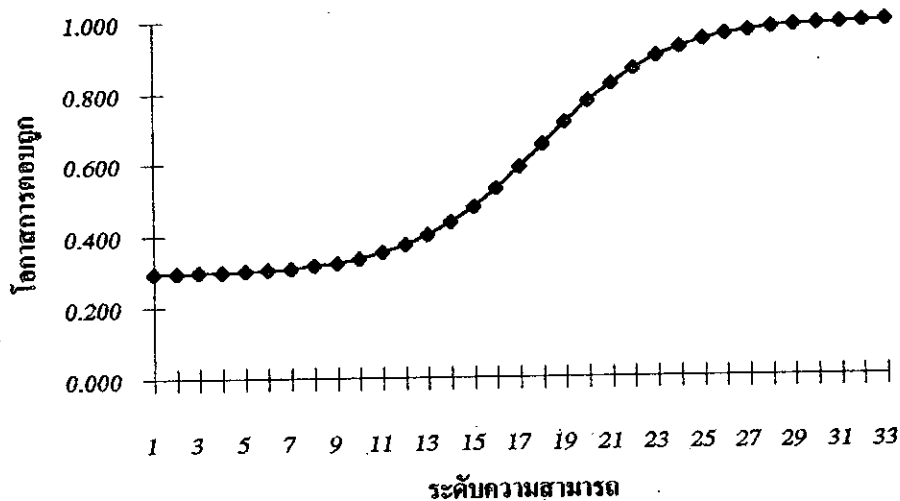


ภาพประกอบ 59 โค้งลักษณะของแบบทดสอบ ชนิด AX แบบเดิมค่าตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

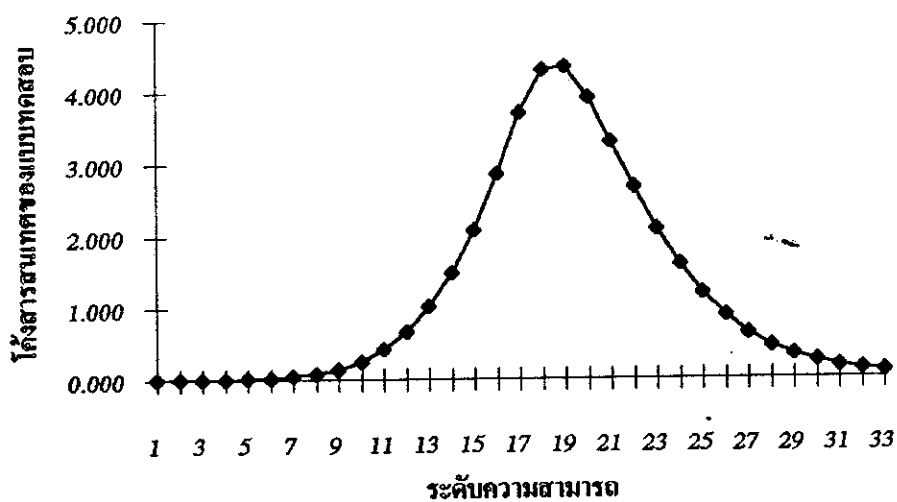


$$AX_{52} I(\max) = 5.617 \text{ ที่ } \theta(\max) = 0.250$$

ภาพประกอบ 60 ได้จำนวนทดสอบของแบบทดสอบ ชนิด AX แบบเดิมค่าตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

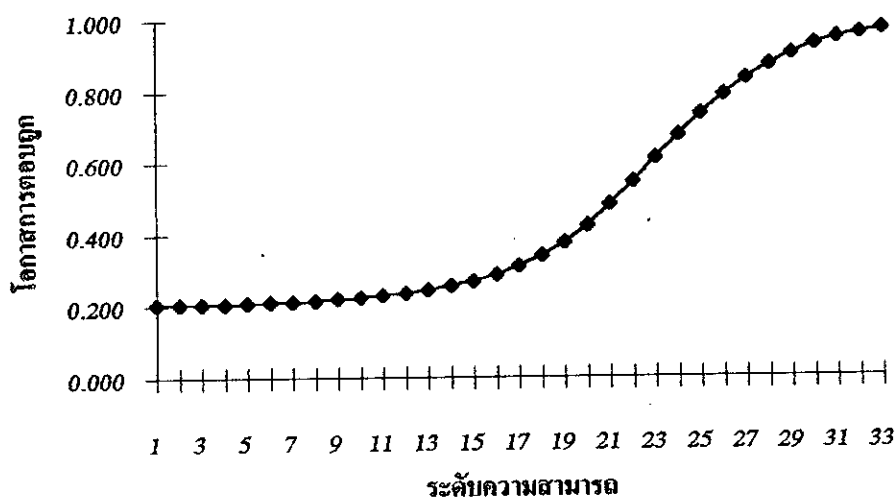


ภาพประกอบ 61 โด่งลักษณะของแบบทดสอบ ชนิด CV แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

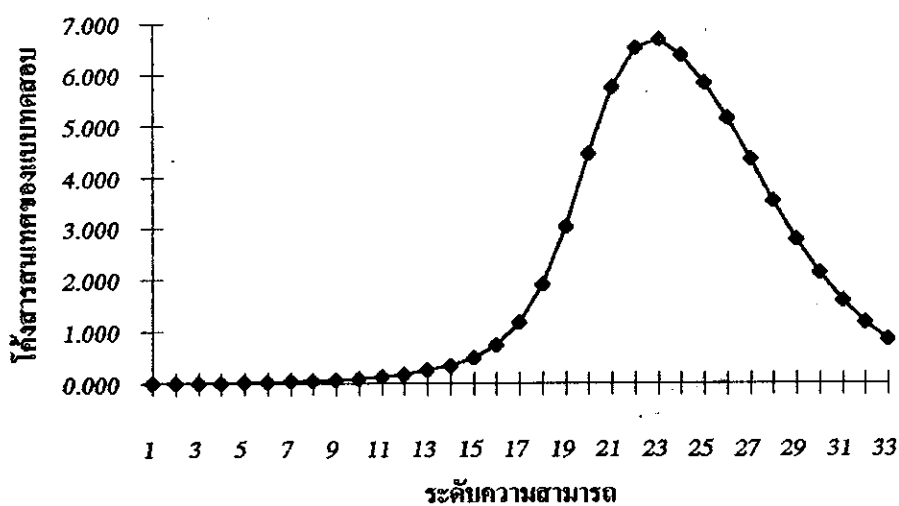


$$CV_{52} I(\max) = 4.336 \text{ ที่ } \sigma(\max) = 0.500$$

ภาพประกอบ 62 โด่งสารสนเทศของแบบทดสอบ ชนิด CV แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

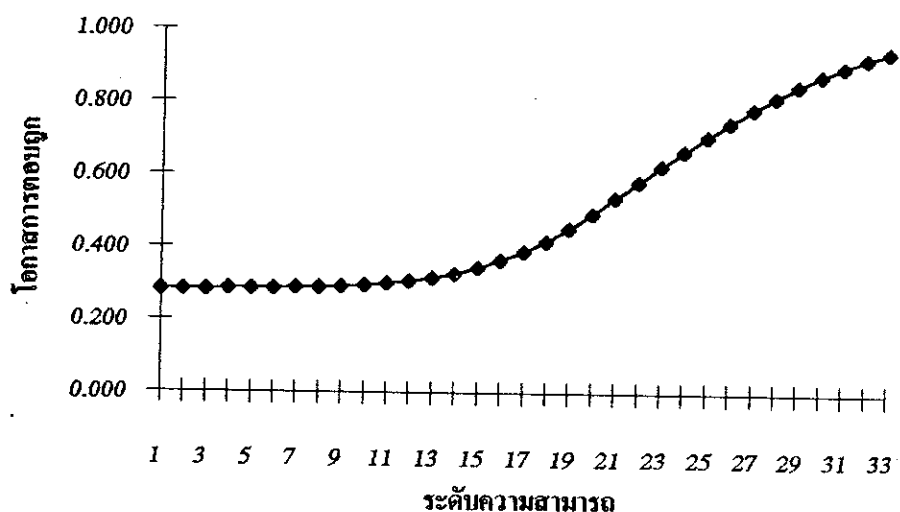


ภาพประกอบ 63 โค้งลักษณะของแบบทดสอบ ชนิด LR แบบเดิมค่าตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

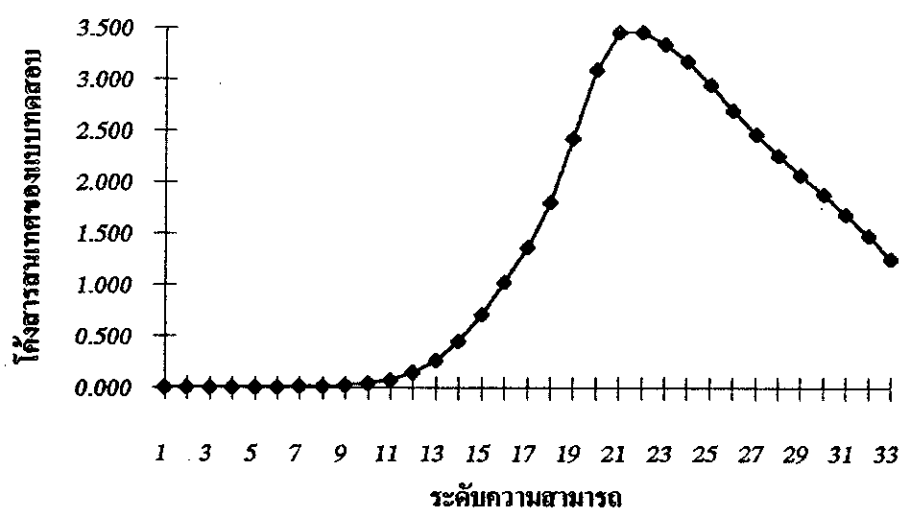


$$LR_{52} \quad I(\max) = 6.680 \quad \text{ที่} \quad \theta(\max) = 1.500$$

ภาพประกอบ 64 โค้งสารสนเทศของแบบทดสอบ ชนิด LR แบบเดิมค่าตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

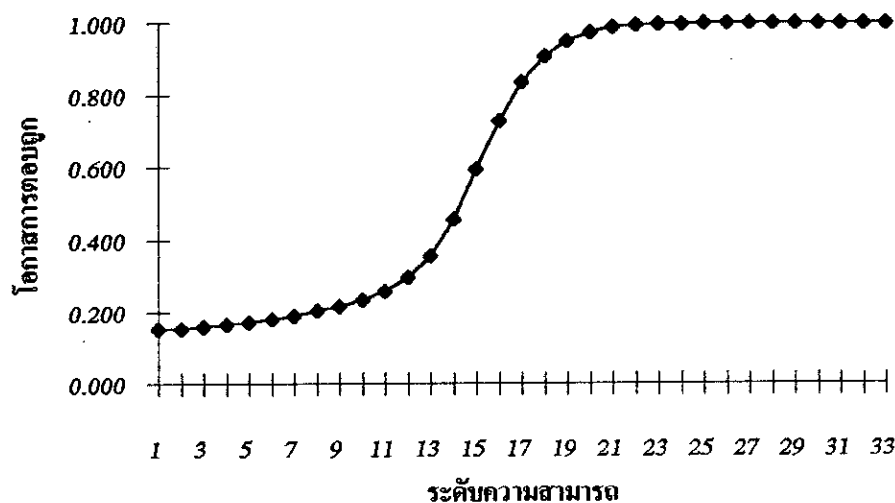


ภาพประกอบ 65 โค้งลักษณะของแบบทดสอบ ชนิด NLR แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

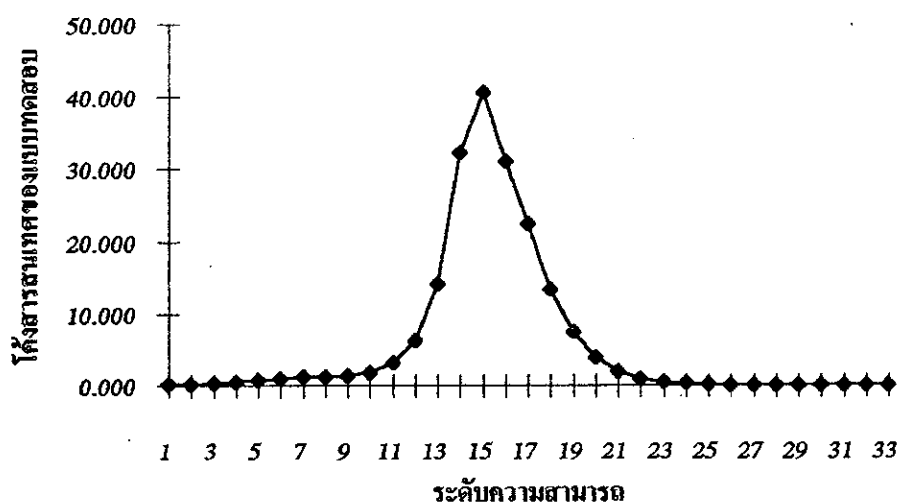


$$NLR_{52} \quad I(\max) = 3.460 \quad \text{ที่} \quad \theta(\max) = 1.250$$

ภาพประกอบ 66 โค้งสารสนเทศของแบบทดสอบ ชนิด NLR แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ



ภาพประกอบ 67 โค้งลักษณะของแบบทดสอบ ชนิด PI แบบเดิมค่าตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ



$$PI_{52} I(\max) = 40.561 \quad \text{ที่ } \theta(\max) = -0.500$$

ภาพประกอบ 68 โค้งสารสนเทศของแบบทดสอบ ชนิด PI แบบเดิมค่าตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ณ ระดับความสามารถต่าง ๆ

จากภาพประกอบ 21 ถึง 68 พบว่า แบบทดสอบแบบเลือกตอบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีจุดสูงสุดของค่าฟังก์ชันสารสนเทศของแบบทดสอบ $\{I(\max)\}$ เรียงตามลำดับจากค่าสูงไปต่ำ คือ ชนิด PI, AR, NLR, CV, LR และ AX โดยมีค่า 39.803, 5.678, 3.266, 3.184, 2.310 และ 1.453 ตามลำดับ ส่วนแบบทดสอบแบบเติมคำตอบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีจุดสูงสุดของค่าฟังก์ชันสารสนเทศของแบบทดสอบ $\{I(\max)\}$ เรียงลำดับจากค่าสูงไปต่ำ คือ ชนิด PI, AR, AX, LR, CV และ NLR โดยมีค่า 39.214, 7.954, 7.702, 6.135, 4.322 และ 2.747 ตามลำดับ

สำหรับแบบทดสอบแบบเลือกตอบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 มีจุดสูงสุดของค่าฟังก์ชันสารสนเทศของแบบทดสอบ $\{I(\max)\}$ เรียงลำดับจากค่าสูงไปต่ำ คือ ชนิด PI, AR, AX, NLR, CV และ LR โดยมีค่า 23.316, 7.563, 2.648, 2.419, 2.338 และ 1.916 ตามลำดับ ส่วนแบบทดสอบแบบเติมคำตอบมีจุดสูงสุดของค่าฟังก์ชันสารสนเทศของแบบทดสอบ $\{I(\max)\}$ เรียงลำดับจากค่าสูงไปต่ำ คือ ชนิด PI, AR, LR, AX, CV และ NLR โดยมีค่า 40.561, 6.937, 6.680, 5.617, 4.336 และ 3.460 ตามลำดับ

กล่าวโดยสรุปได้ว่า แบบทดสอบชนิด PI ทั้งรูปแบบเลือกตอบ และแบบเติมคำตอบจะมีจุดสูงสุดของค่าฟังก์ชันสารสนเทศสูงสุด และชนิด AR มีจุดสูงสุดของค่าฟังก์ชันสารสนเทศสูงสุด ในลำดับรองลงมา

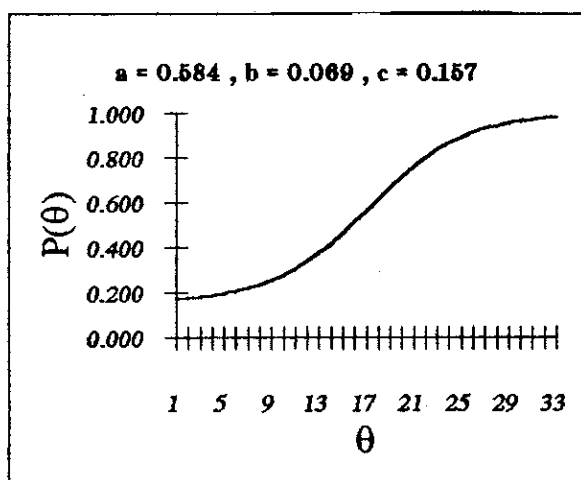


Fig 1

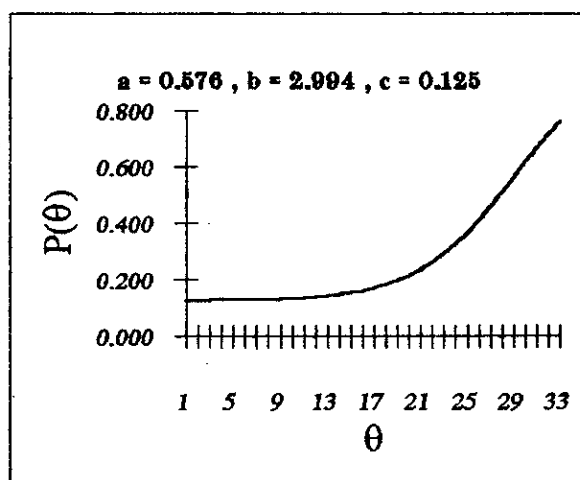


Fig 2

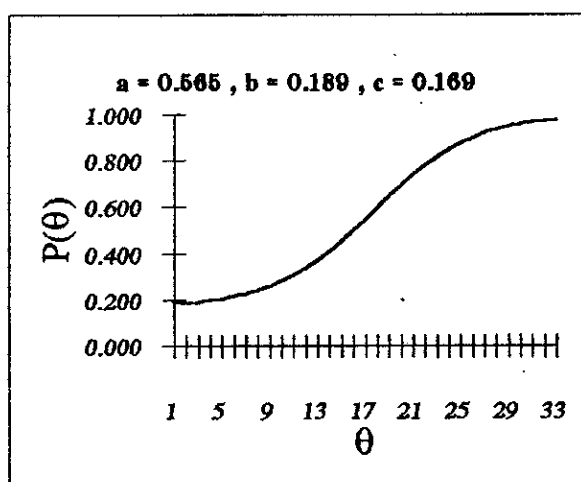


Fig 3

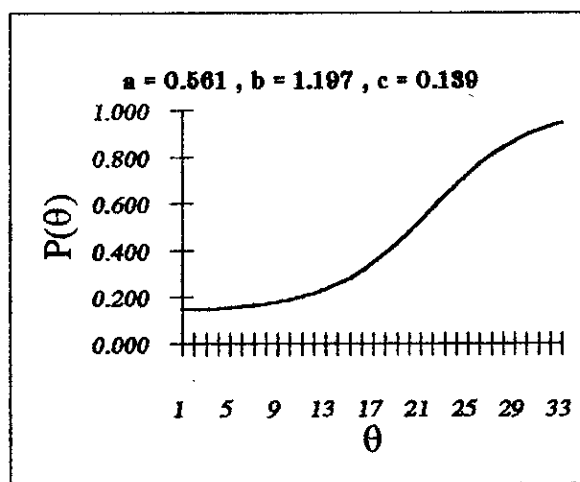


Fig 4

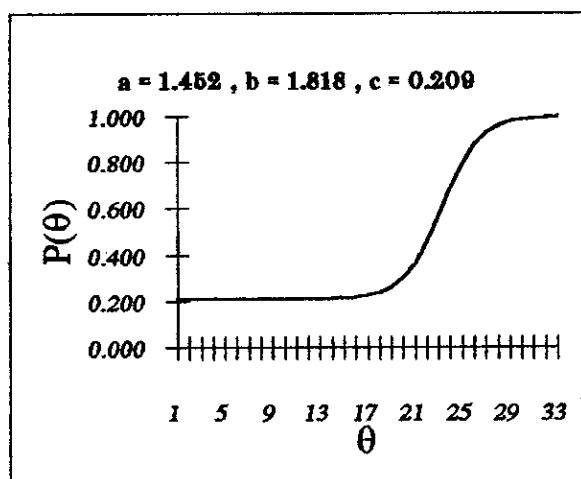


Fig 5

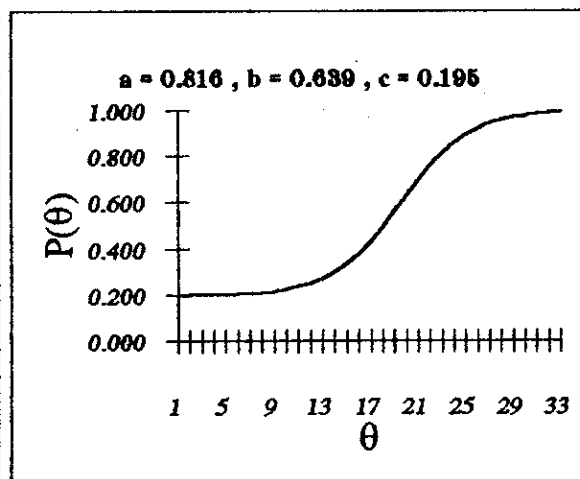


Fig 6

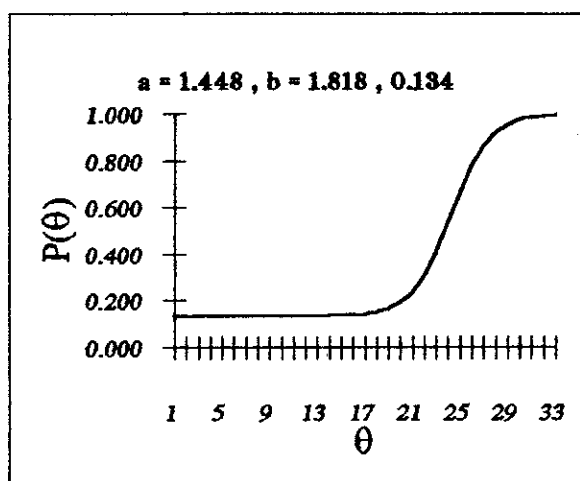


Fig 7

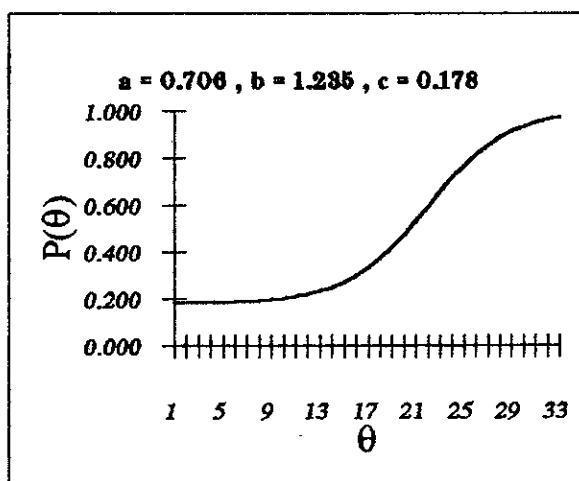


Fig 8

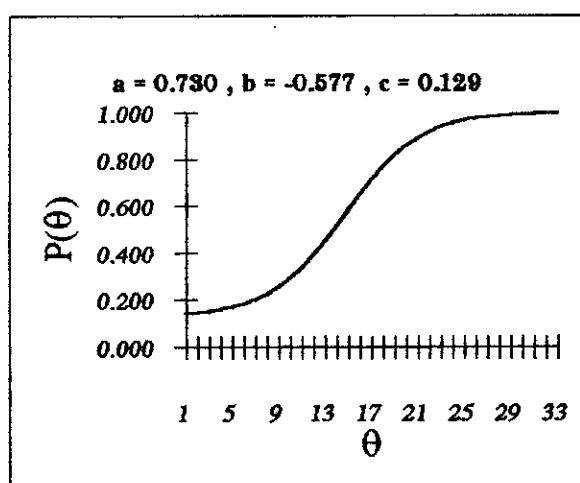


Fig 9

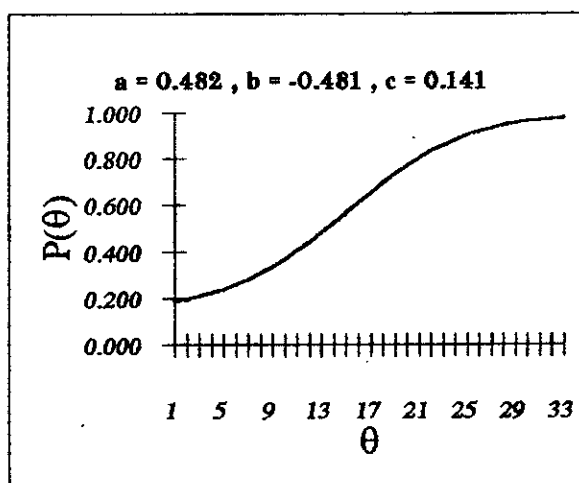


Fig 10

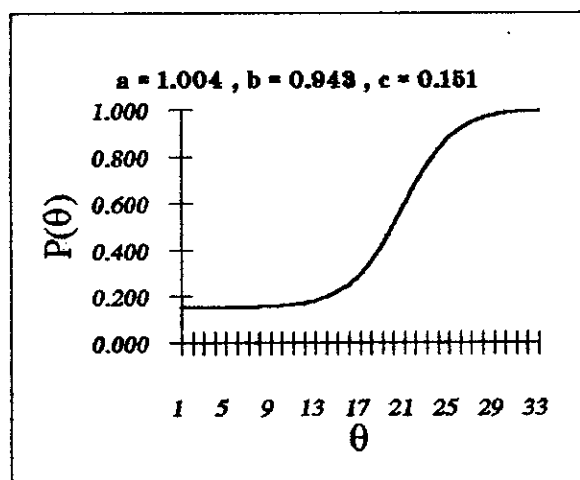


Fig 11

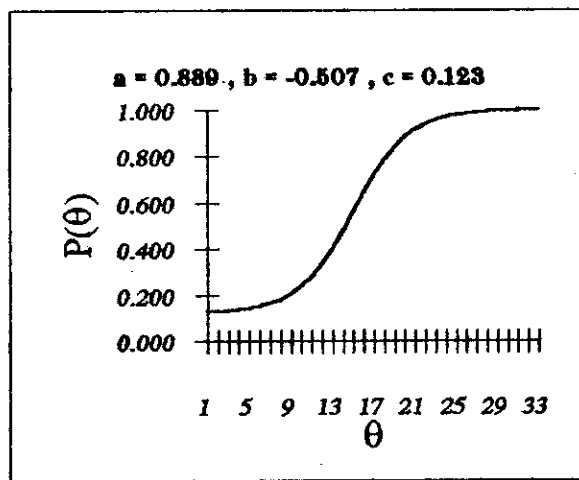


Fig 12

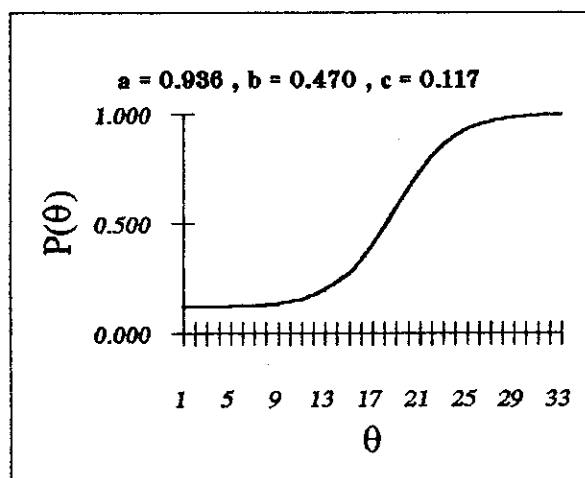


Figure 13

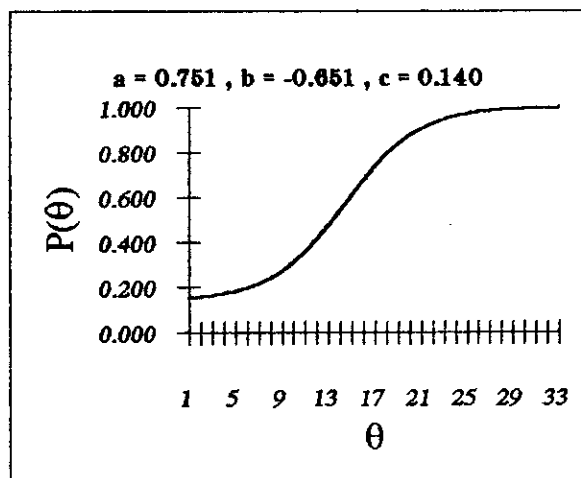


Figure 14

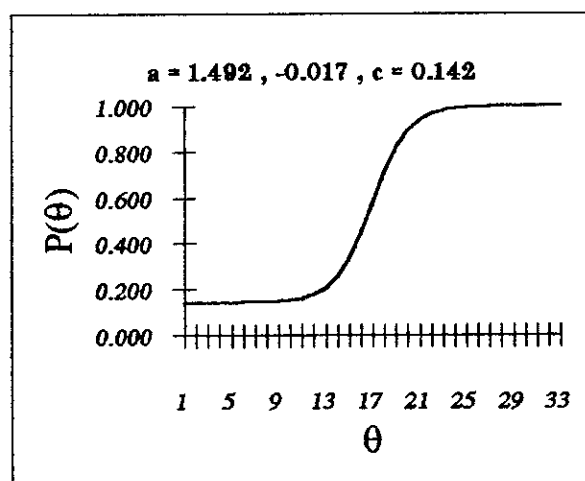


Figure 15

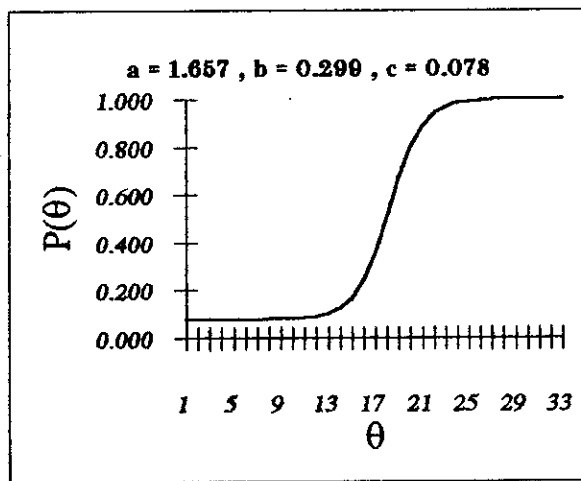


Figure 16

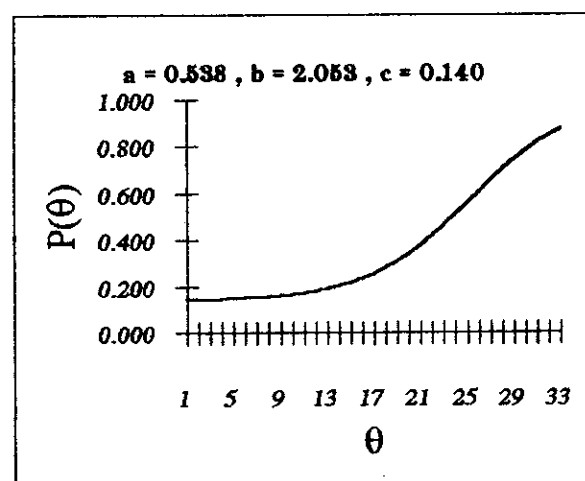


Figure 17

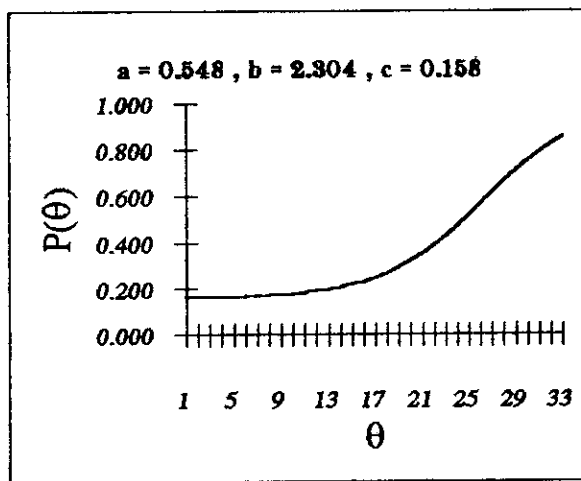
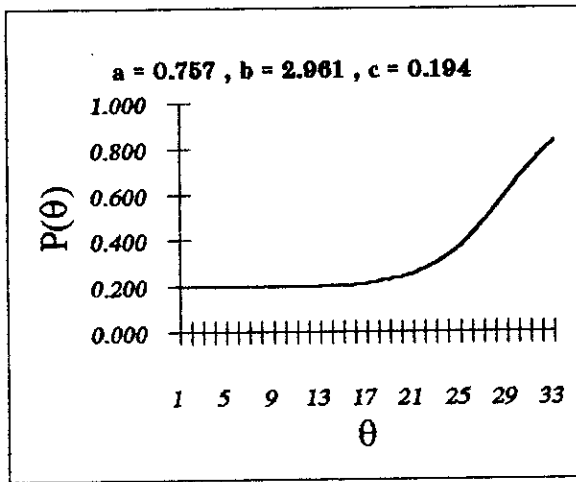
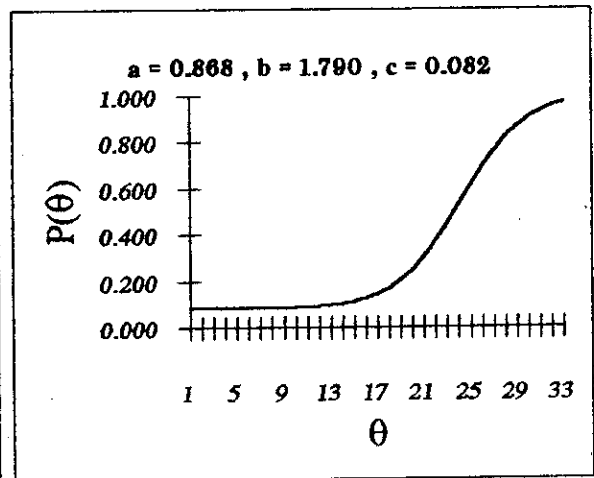


Figure 18



ข้อ 19



ข้อ 20

ภาพประกอบ 69 โค้งลักษณะข้อสอบของข้อสอบ ชนิด AR แบบเลือกตอบ ของนักเรียนชั้น
มัธยมศึกษาปีที่ 3

จากภาพประกอบ 69 พบว่า แบบทดสอบชนิด AR แบบเลือกตอบของนักเรียน ชั้น
มัธยมศึกษาปีที่ 3 ข้อสอบข้อ 2 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 14 มีค่าความ
ยากต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 16 และข้อ 10 มีค่าอำนาจจำแนกต่ำ
สุด สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 5 และข้อ 16 มีค่าการเดาต่ำสุด

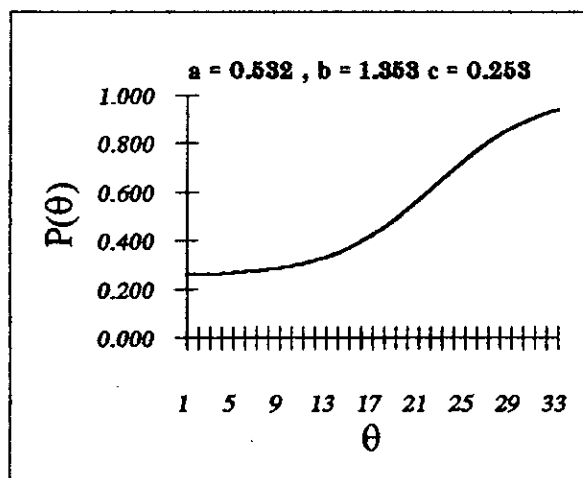


Fig 1

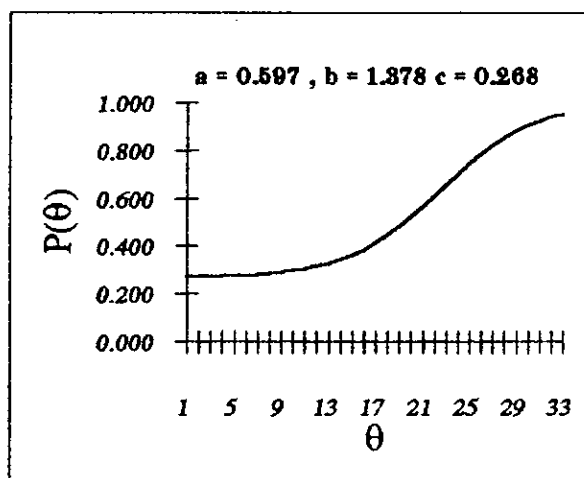


Fig 2

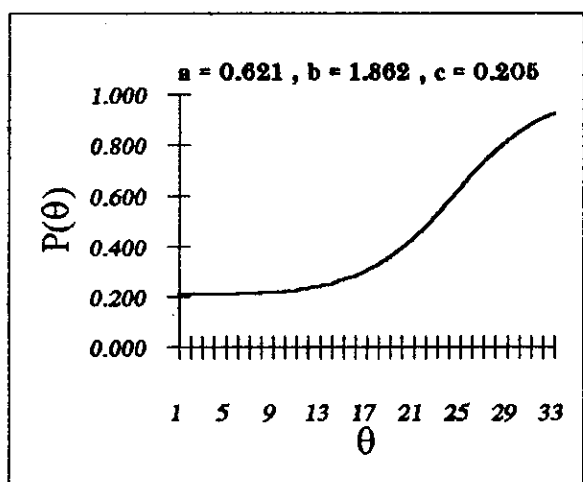


Fig 3

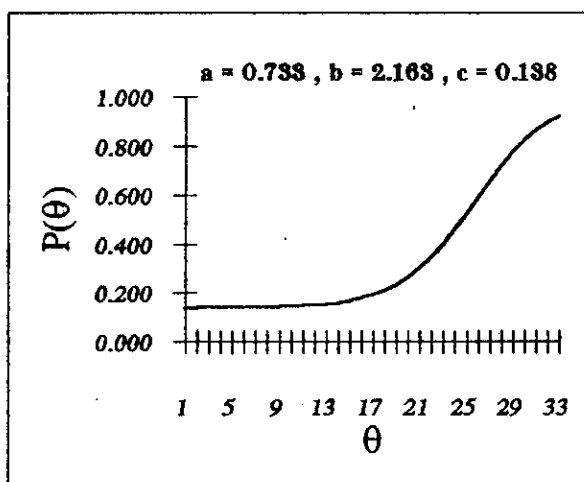


Fig 4

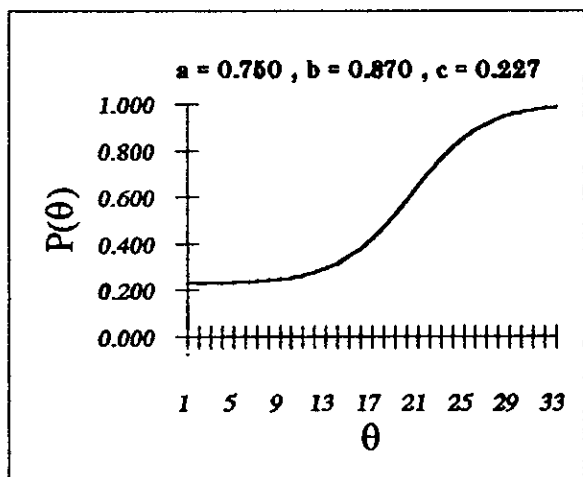


Fig 5

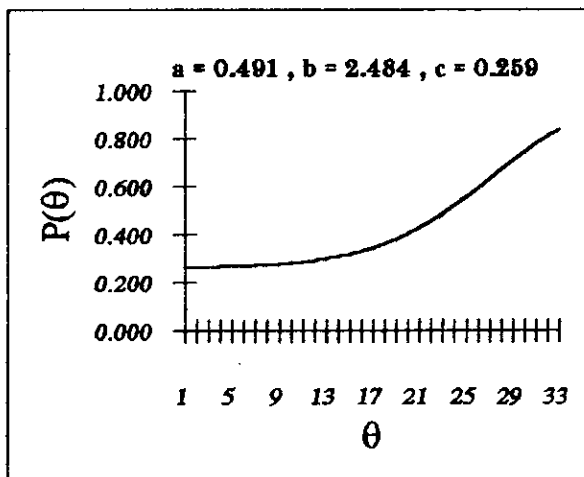
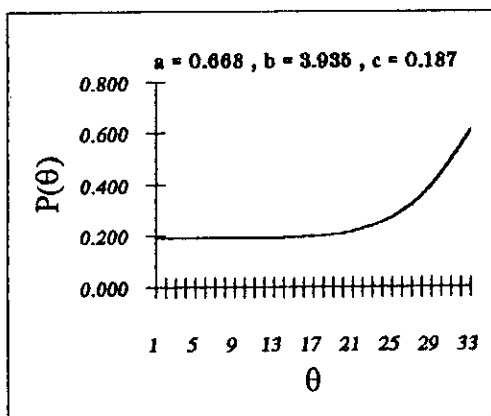
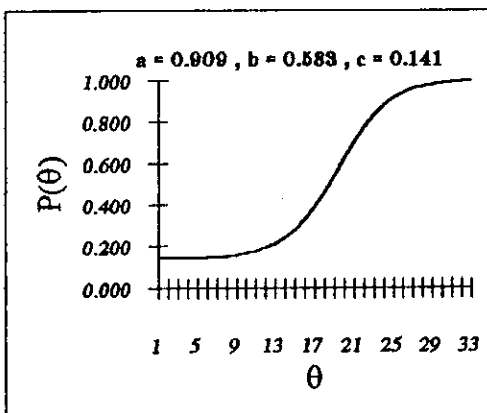


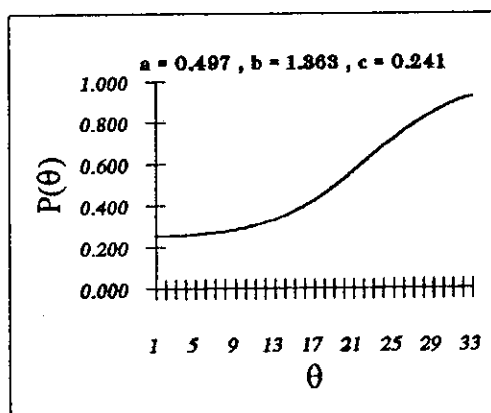
Fig 6



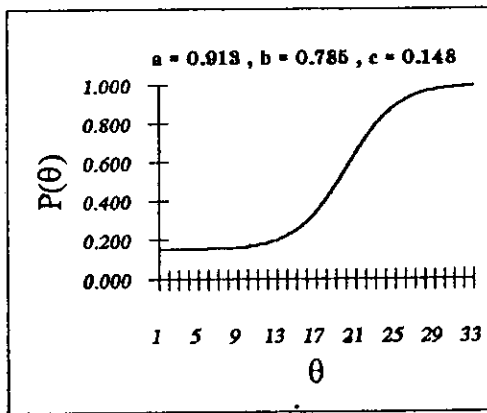
ข้อ 7



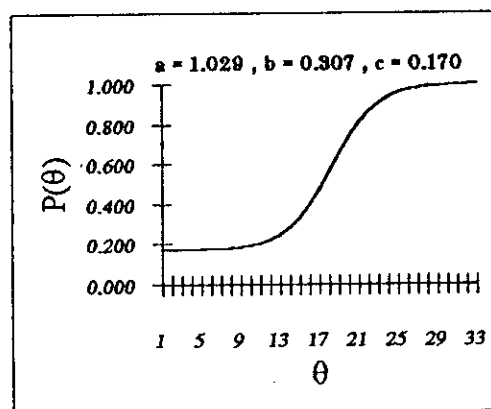
ข้อ 8



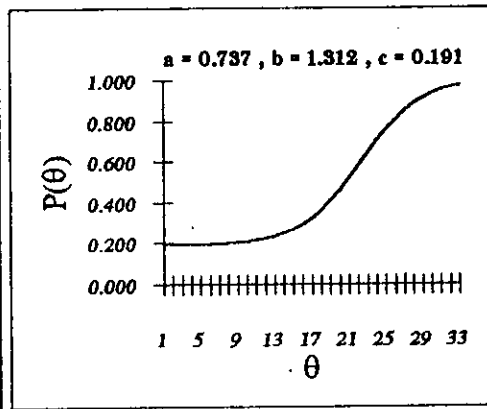
ข้อ 9



ข้อ 10



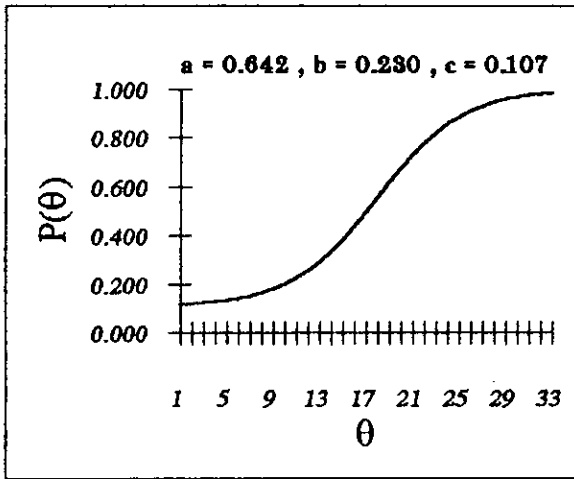
ข้อ 11



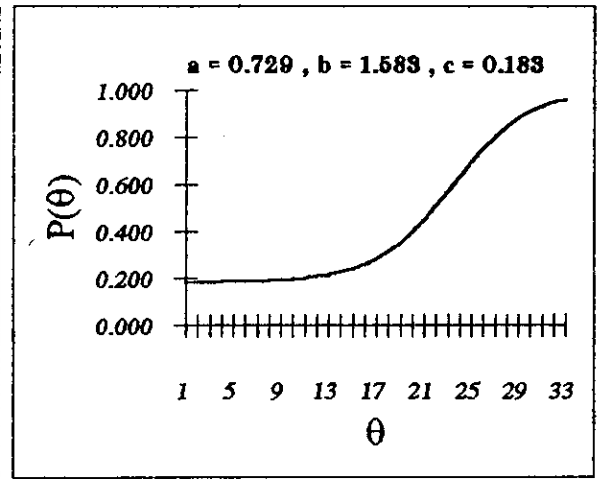
ข้อ 12

ภาพประกอบ 70 ได้ังลักษณะข้อสอบของข้อสอบ ชนิด AX แบบเลือกตอบ ของนักเรียนชั้น
มัธยมศึกษาปีที่ 3

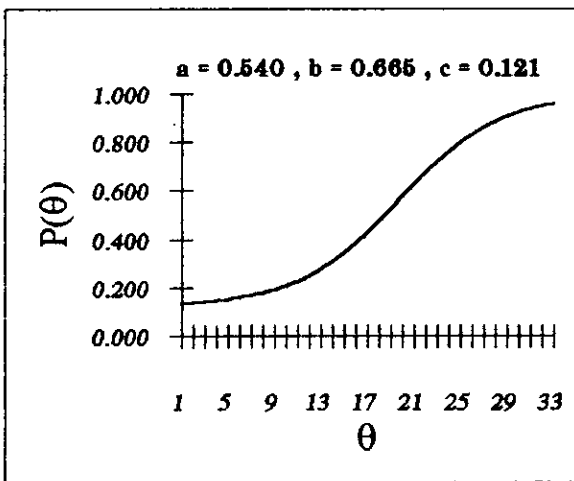
จากภาพประกอบ 70 พบว่า แบบทดสอบชนิด AX แบบเลือกตอบของนักเรียน ชั้นมัธยมศึกษาปีที่ 3 ข้อสอบข้อ 7 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 11 มีค่าความยากต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 11 และข้อ 9 มีค่าอำนาจจำแนกต่ำสุด สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 2 และข้อ 4 มีค่าการเดาต่ำสุด



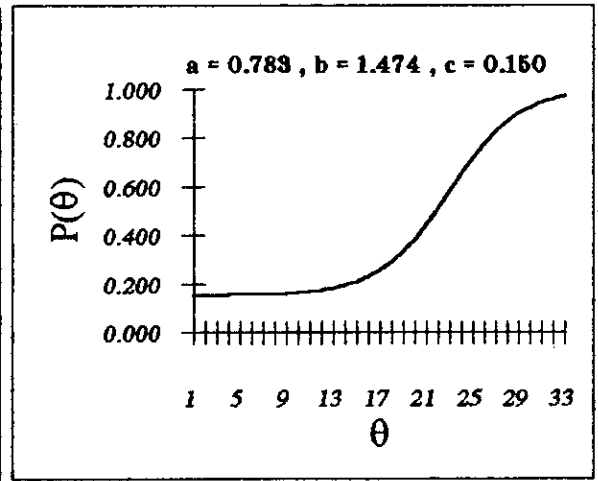
Şö 1



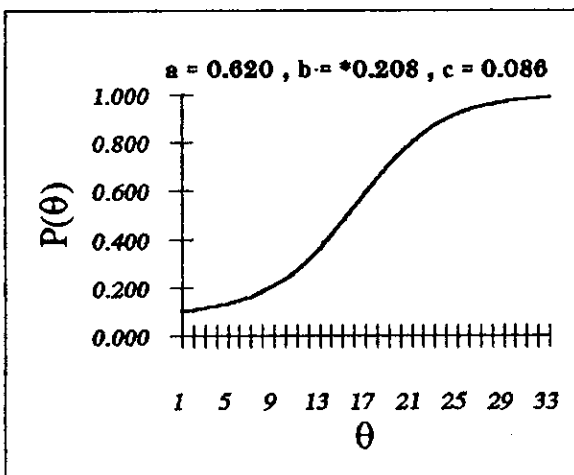
Şö 2



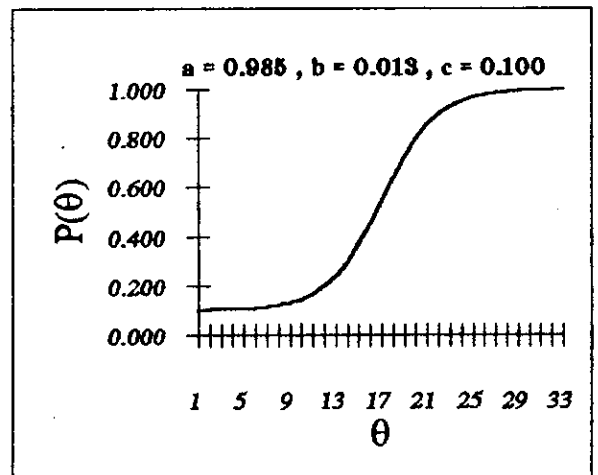
Şö 3



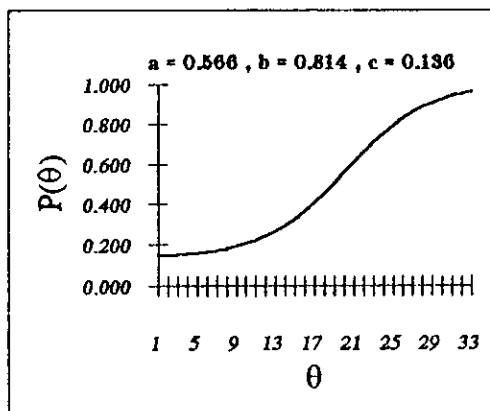
Şö 4



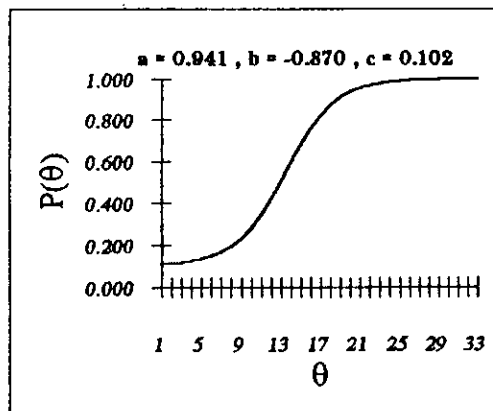
Şö 5



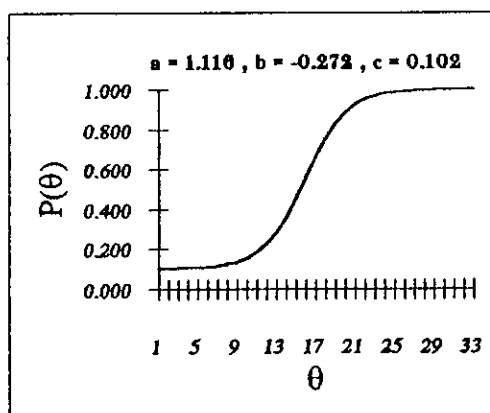
Şö 6



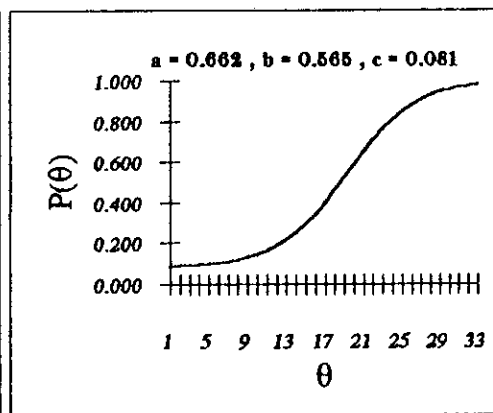
รูป 7



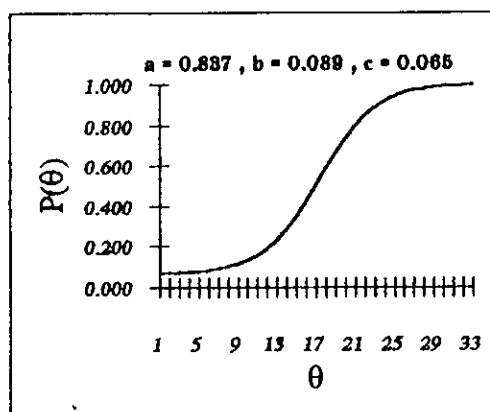
รูป 8



รูป 9



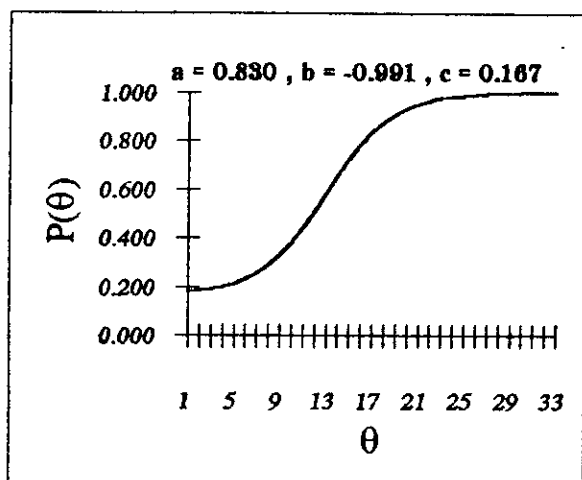
รูป 10



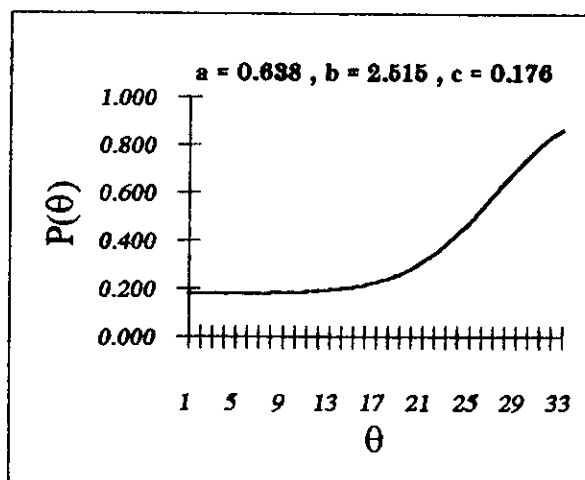
รูป 11

ภาพประกอบ 71 ได้ังลักษณะข้อสอบของข้อสอบ ชนิด CV แบบเลือกตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3

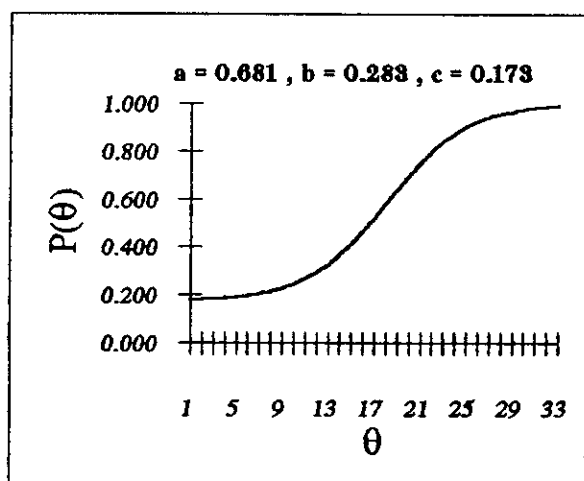
จากภาพประกอบ 71 พบว่า แบบทดสอบชนิด CV แบบเลือกตอบของนักเรียน ชั้นมัธยมศึกษาปีที่ 3 ข้อสอบข้อ 2 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 8 มีค่าความยากต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 9 และข้อ 3 มีค่าอำนาจจำแนกต่ำสุด สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 2 และข้อ 11 มีค่าการเดาต่ำสุด



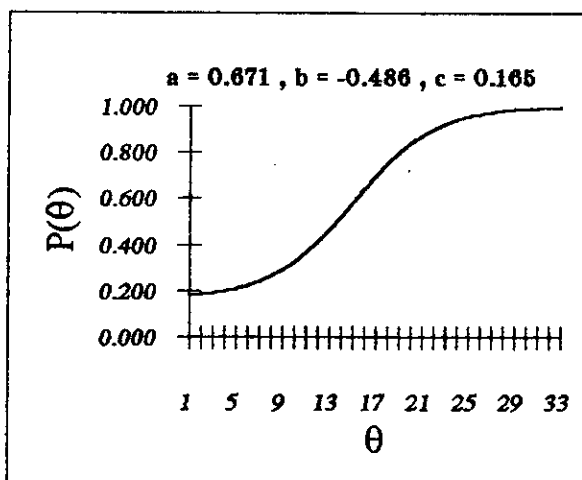
Şö 1



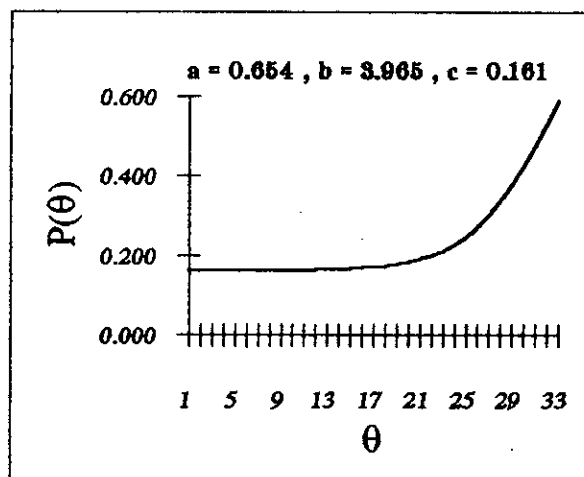
Şö 2



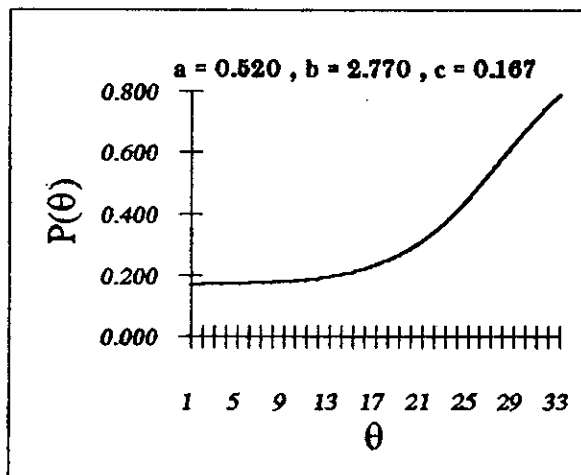
Şö 3



Şö 4



Şö 5



Şö 6

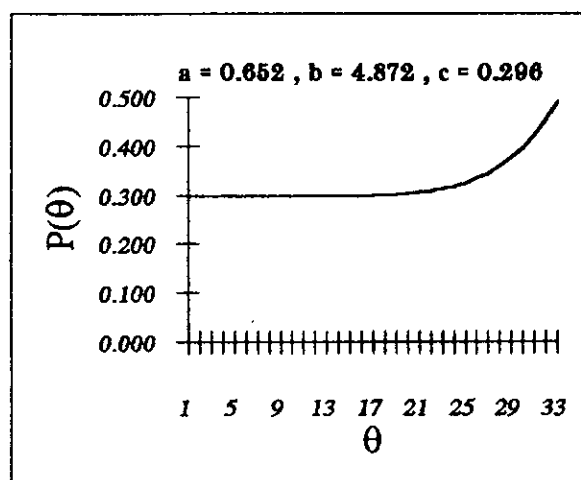


Fig 7

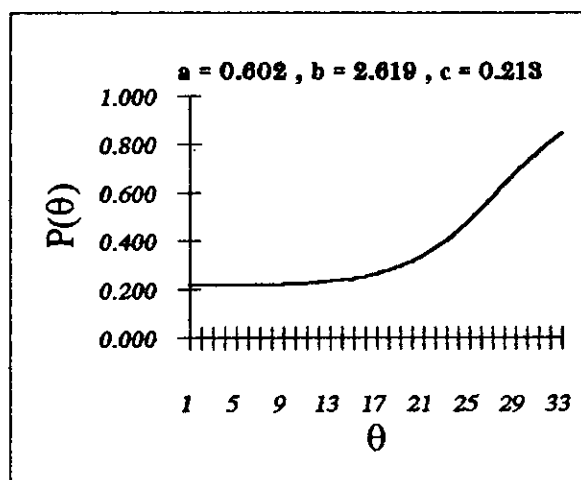


Fig 8

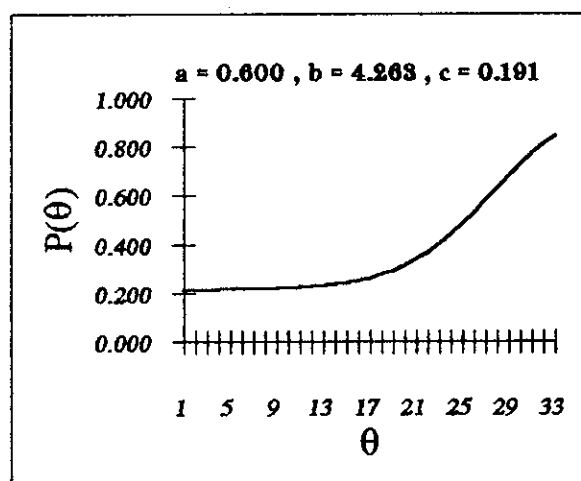


Fig 9

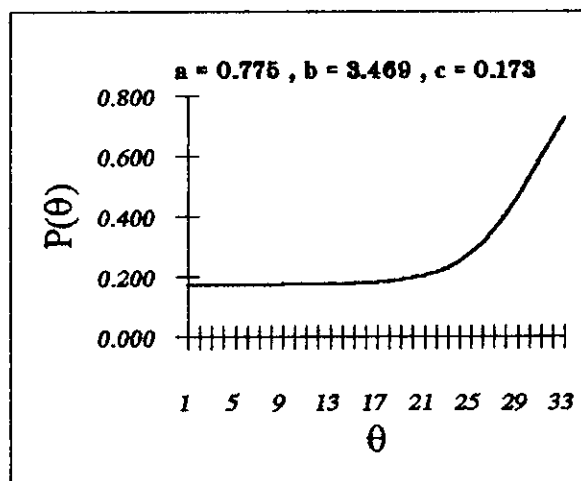


Fig 10

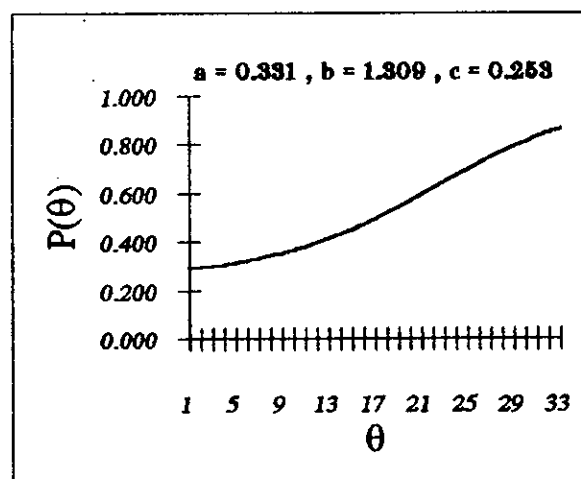


Fig 11

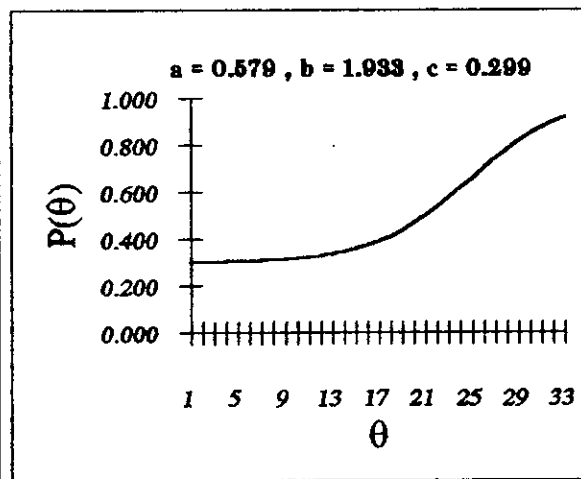
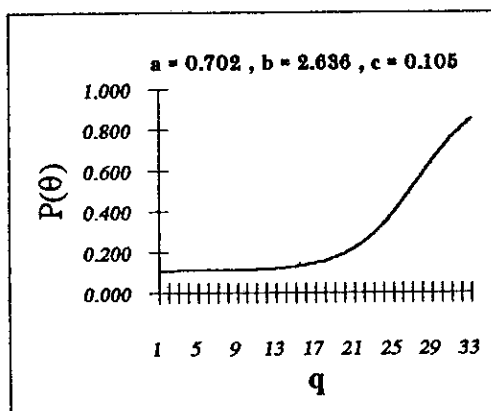
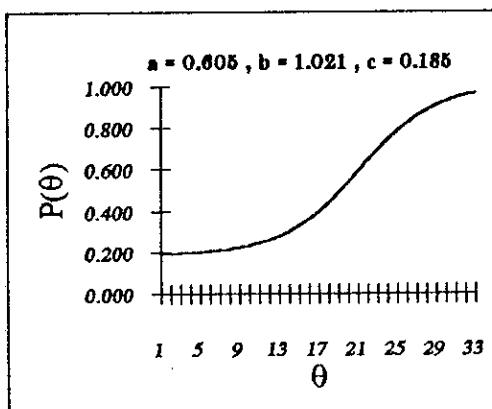


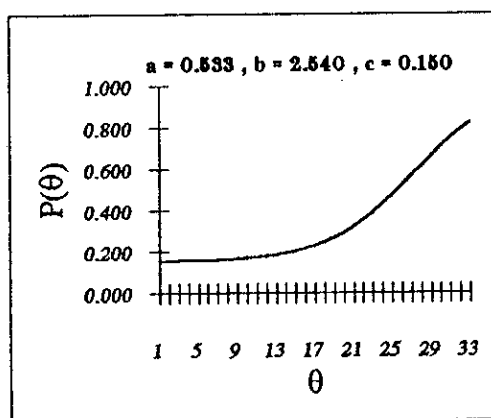
Fig 12



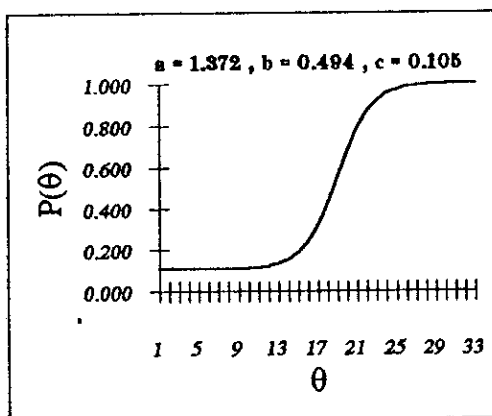
รูป 13



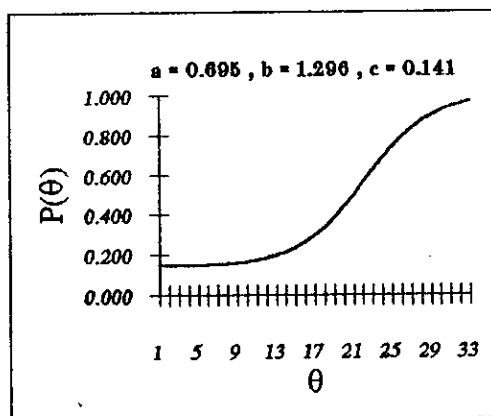
รูป 14



รูป 15



รูป 16



รูป 17

ภาพประกอบ 72 ได้ังลักษณะข้อสอบของข้อสอบ ชนิด LR แบบเลือกตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3

จากภาพประกอบ 72 พบว่า แบบทดสอบชนิด LR แบบเลือกตอบของนักเรียน ชั้นมัธยมศึกษาปีที่ 3 ข้อสอบข้อ 7 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 4 มีค่าความยากต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 16 และข้อ 11 มีค่าอำนาจจำแนกต่ำสุด สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 12 ส่วนข้อ 13 และข้อ 16 มีค่าการเดาต่ำสุด

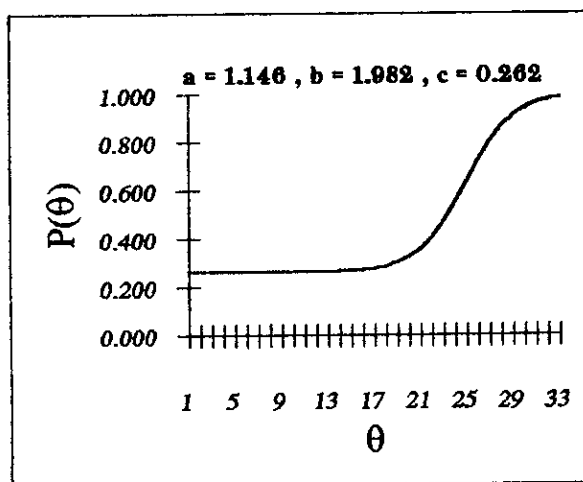


Fig 1

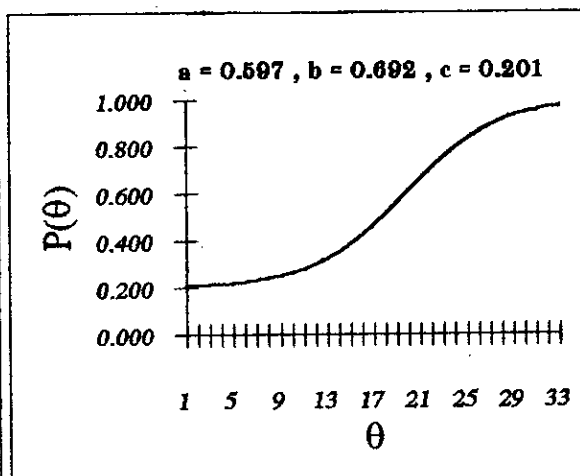


Fig 2

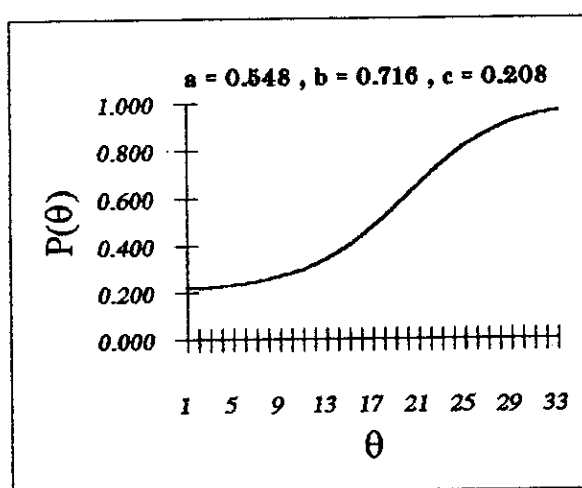


Fig 3

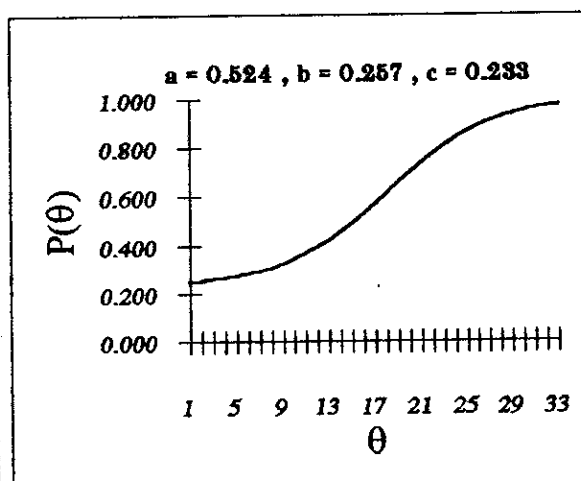


Fig 4

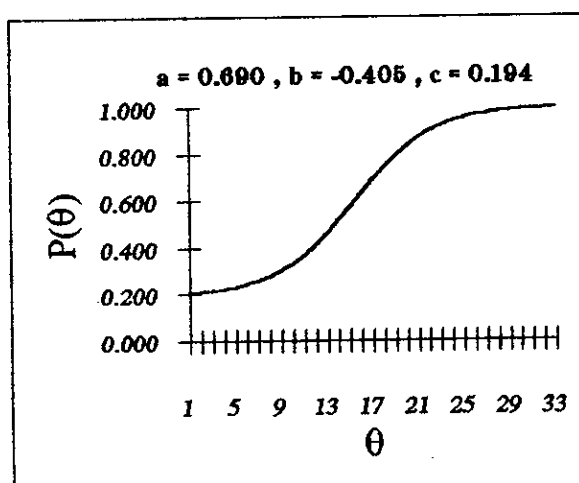


Fig 5

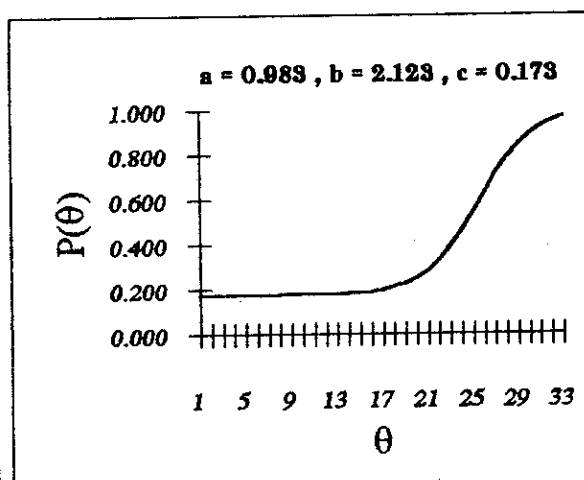
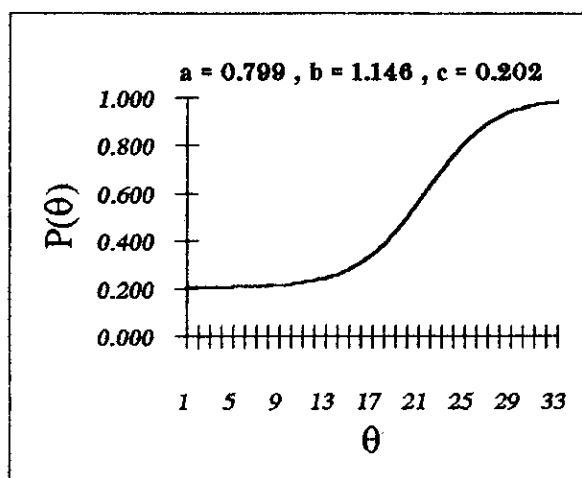
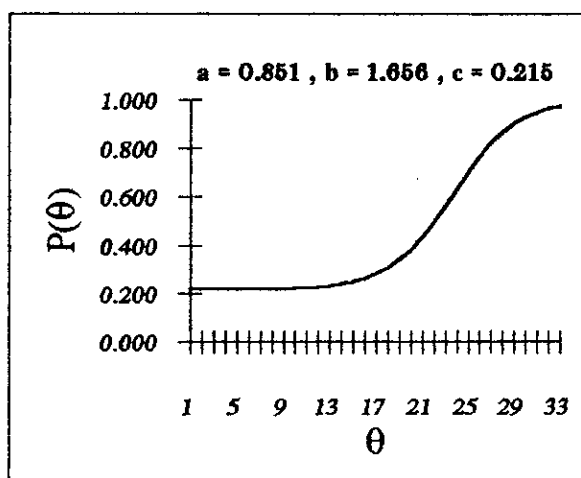


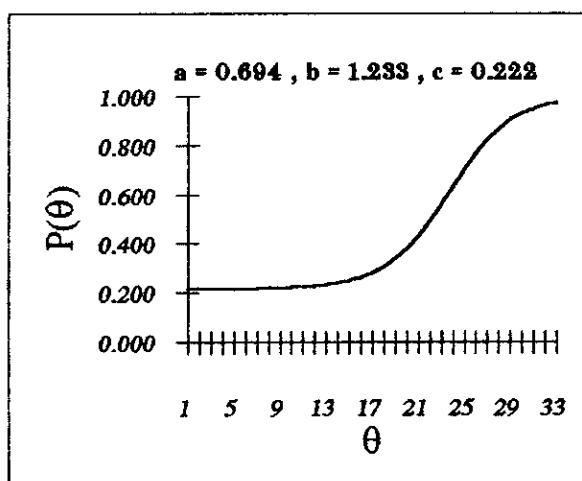
Fig 6



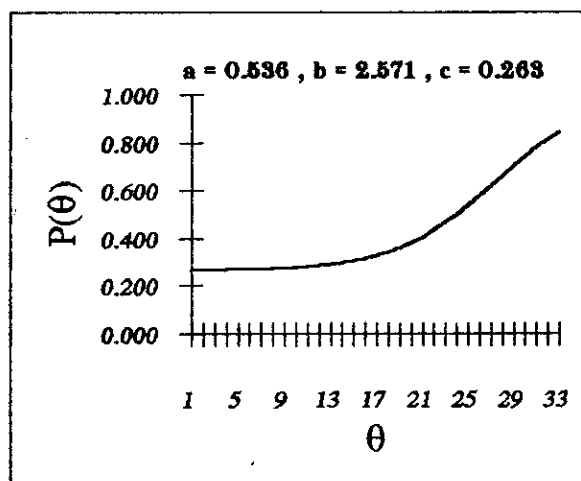
Şə 7



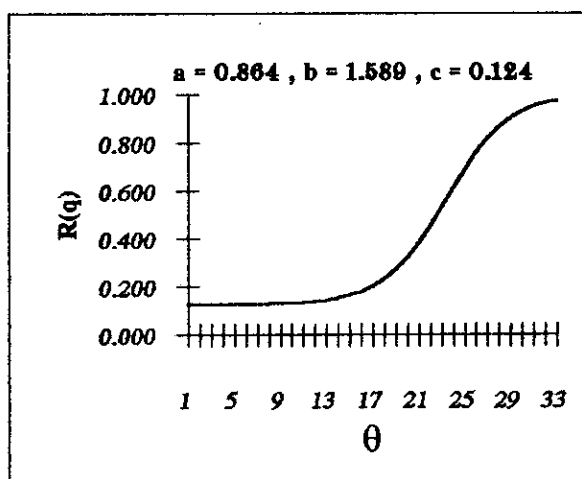
Şə 8



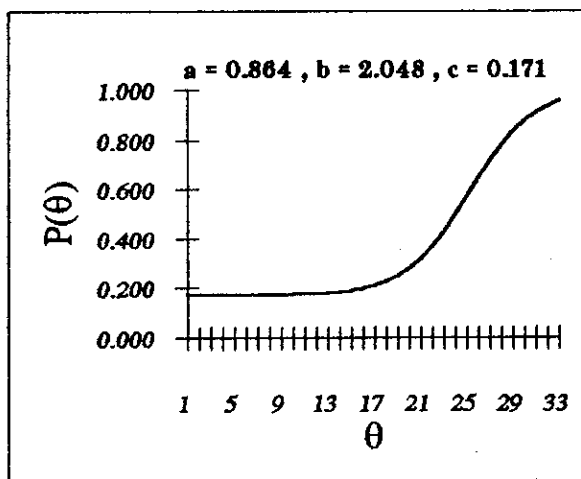
Şə 9



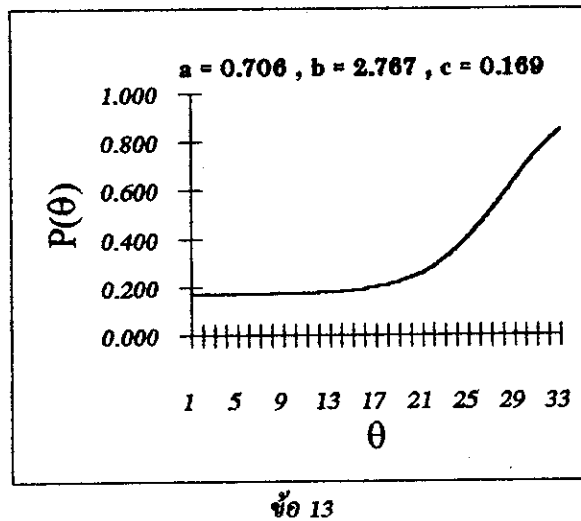
Şə 10



Şə 11

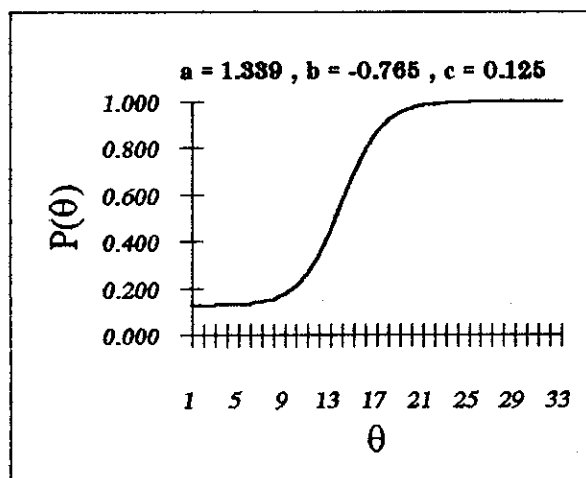


Şə 12

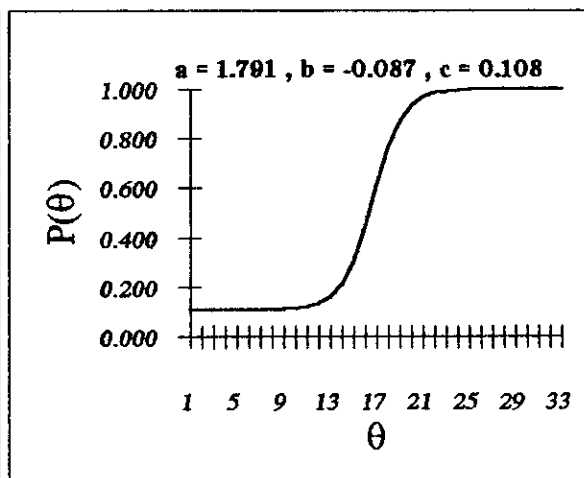


ภาพประกอบ 73 ได้ังลักษณะข้อสอบของข้อสอบ ชนิด NLR แบบเลือกตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3

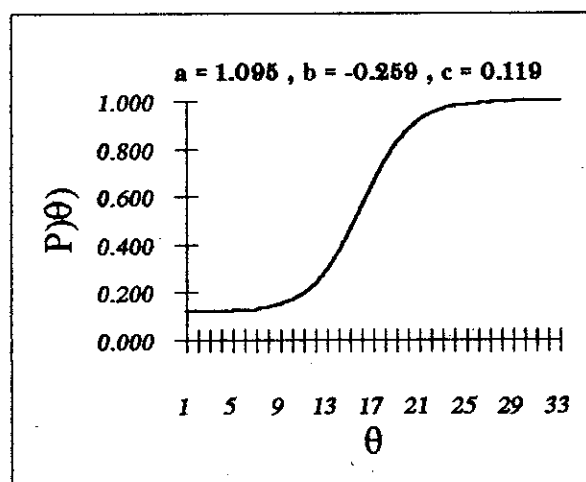
จากภาพประกอบ 73 พบว่า แบบทดสอบชนิด NLR แบบเลือกตอบของนักเรียน ชั้น
มัธยมศึกษาปีที่ 3 ข้อสอบข้อ 13 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 5 มีค่าความยาก
ต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 1 และข้อ 4 มีค่าอำนาจจำแนกต่ำสุด
สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 10 และข้อ 11 มีค่าการเดาต่ำสุด



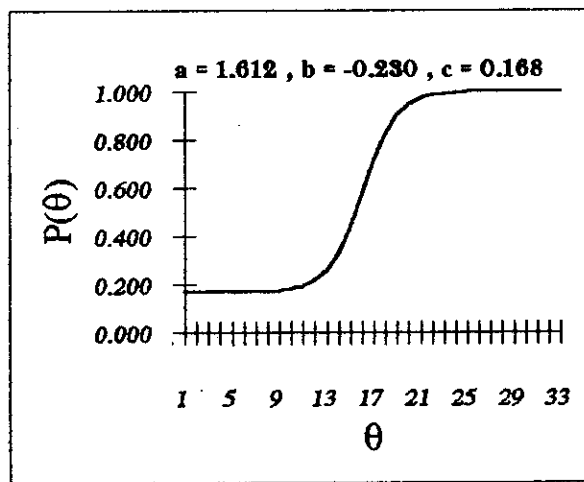
θ 1



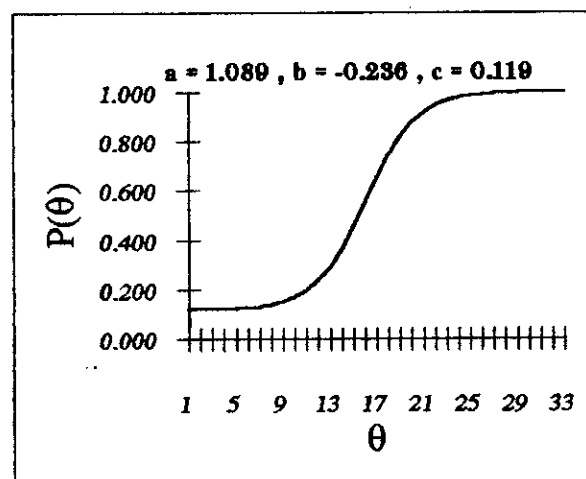
θ 2



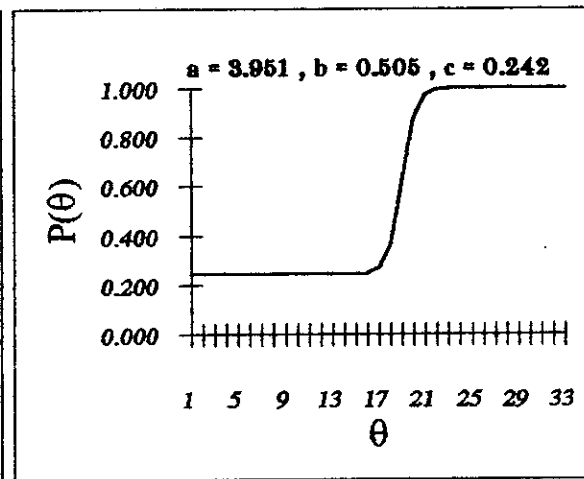
θ 3



θ 4



θ 5



θ 6

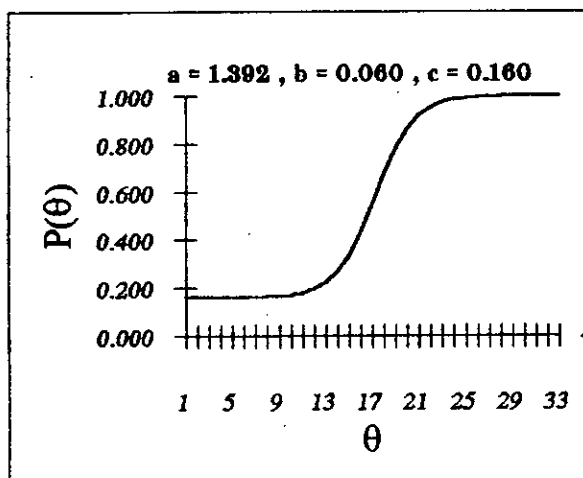


Fig 7

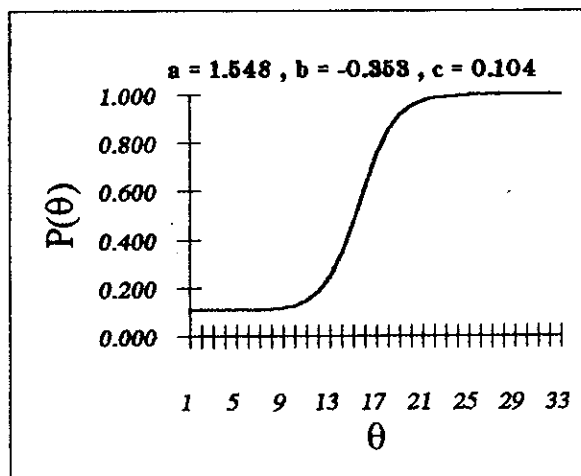


Fig 8

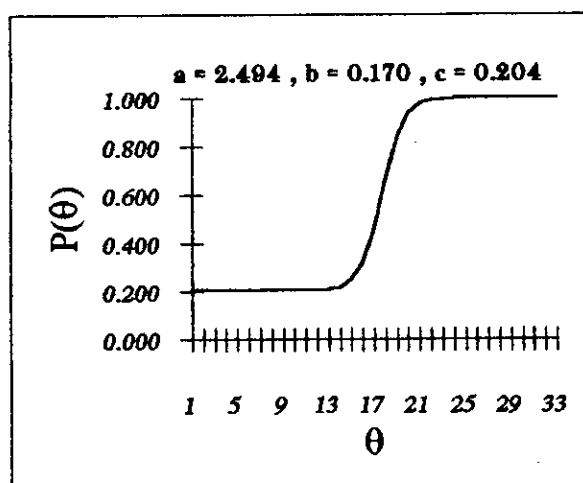


Fig 9

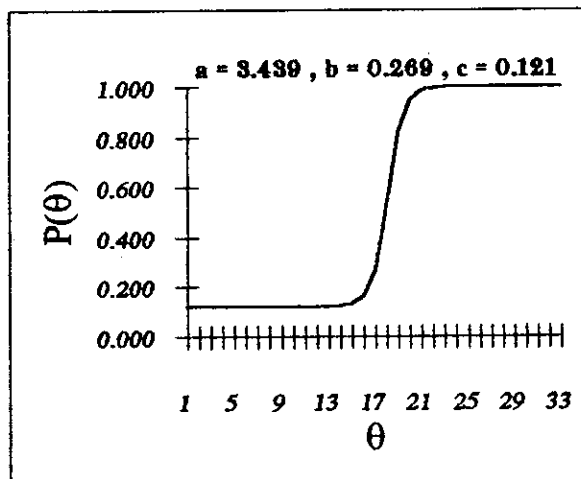


Fig 10

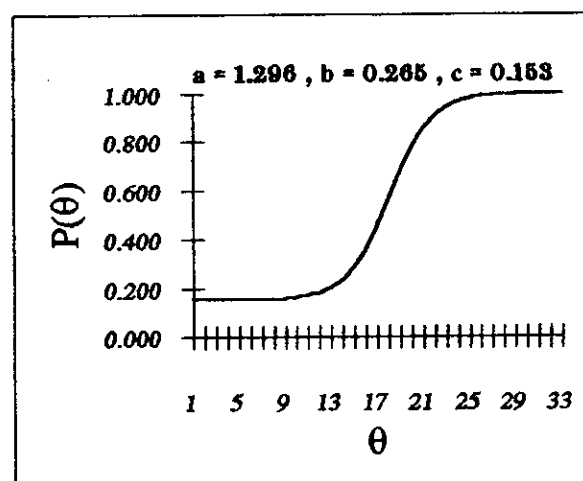


Fig 11

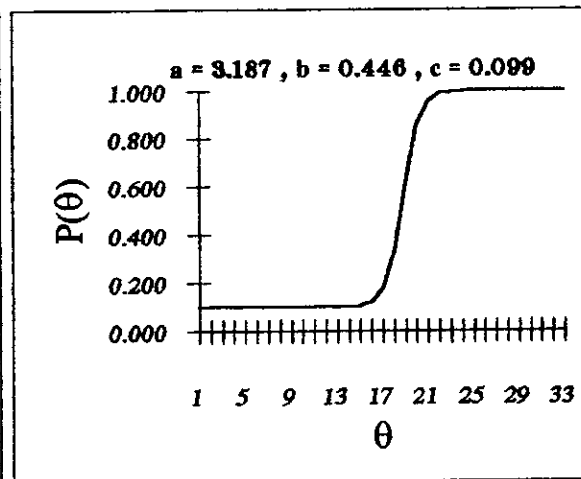


Fig 12

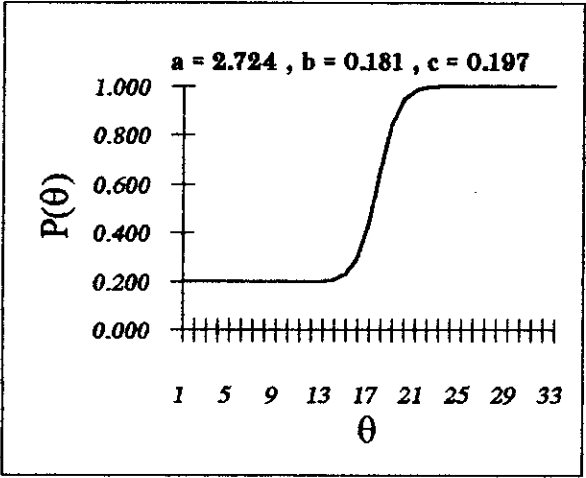


Figure 13

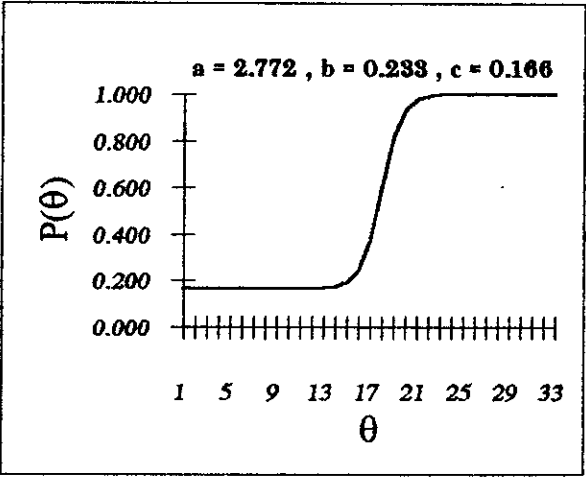


Figure 14

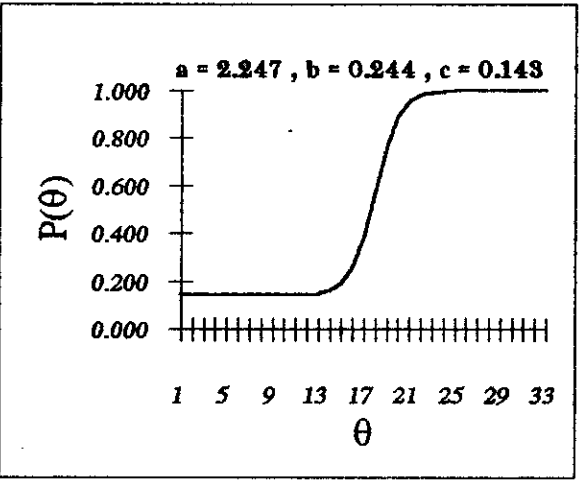


Figure 15

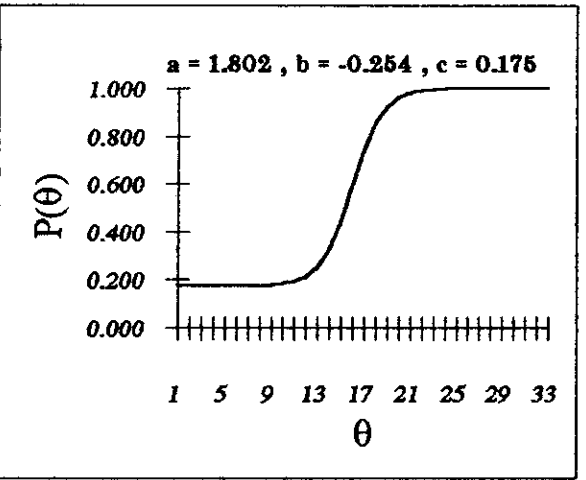


Figure 16

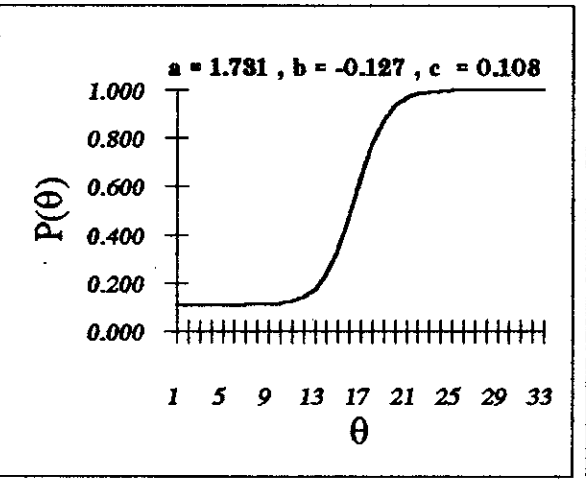


Figure 17

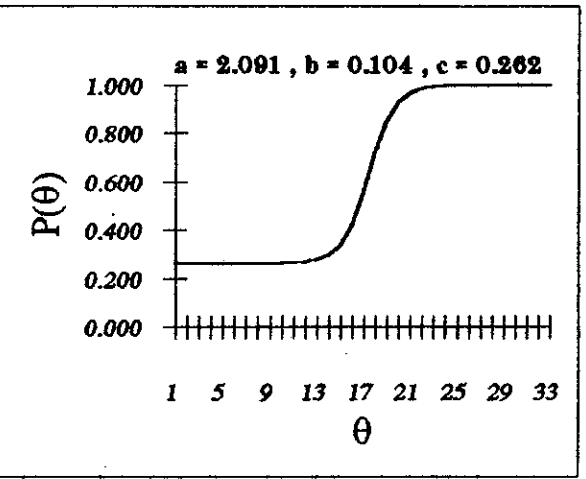
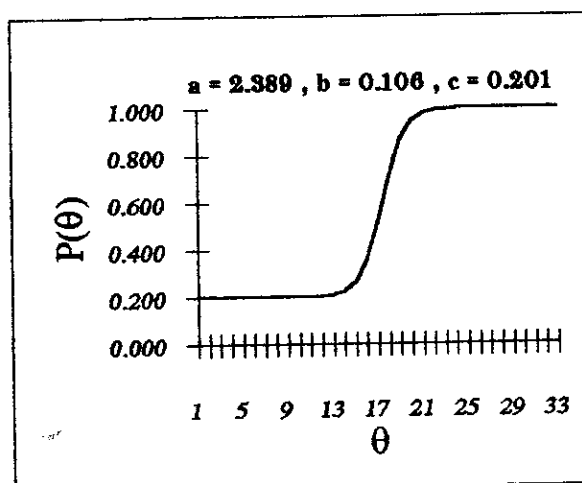


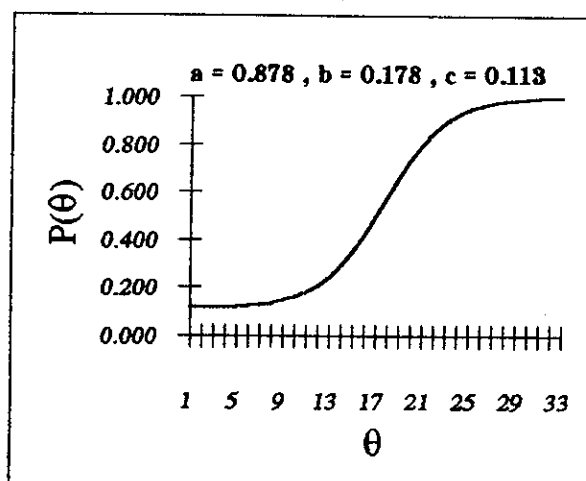
Figure 18



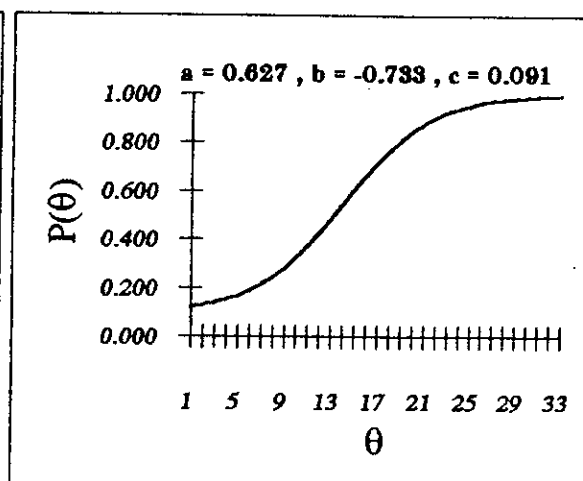
ข้อ 19

ภาพประกอบ 74 ได้ังลักษณะข้อสอบของข้อสอบ ชนิด PI แบบเลือกตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3

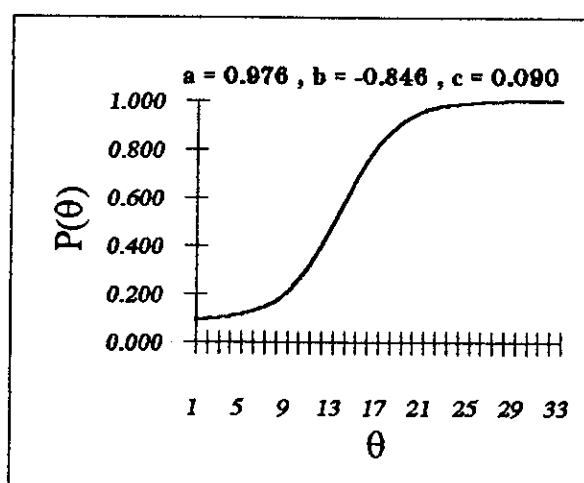
จากภาพประกอบ 74 พบว่า แบบทดสอบชนิด PI แบบเลือกตอบของนักเรียน ชั้น
มัธยมศึกษาปีที่ 3 ข้อสอบข้อ 6 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 1 มีค่าความยาก
ต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 6 และข้อ 5 มีค่าอำนาจจำแนกต่ำสุด
สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 18 และข้อ 12 มีค่าการเดาต่ำสุด



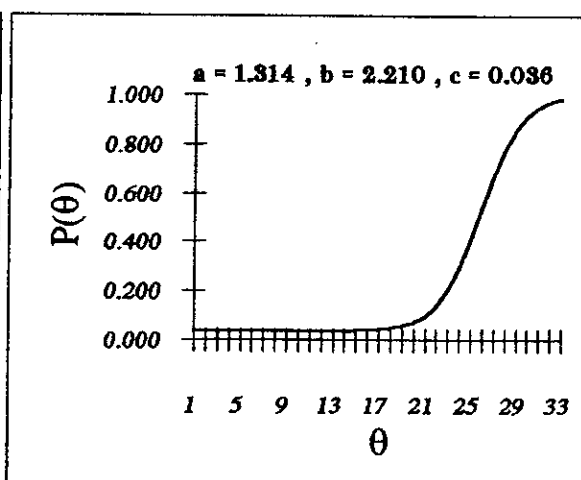
θ 1



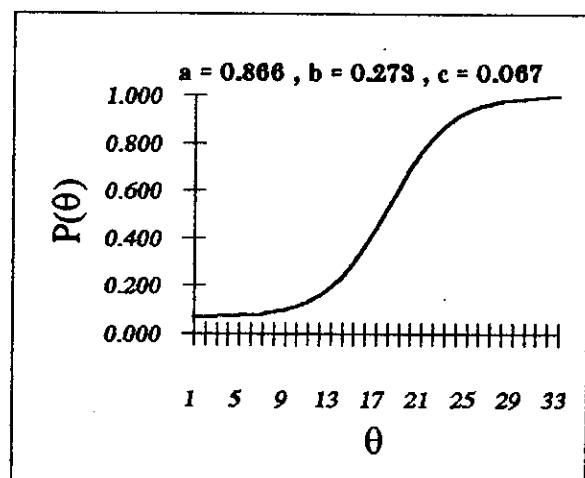
θ 2



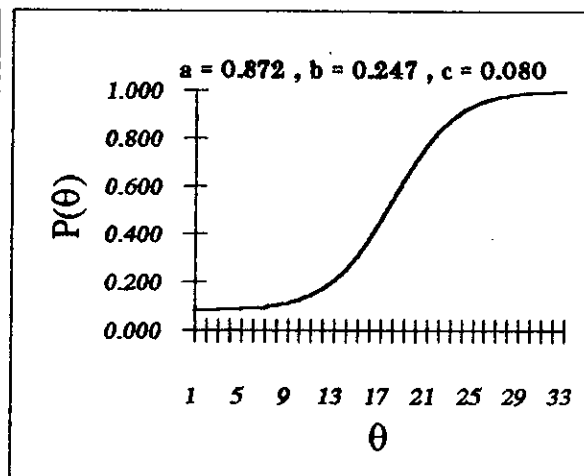
θ 3



θ 4



θ 5



θ 6

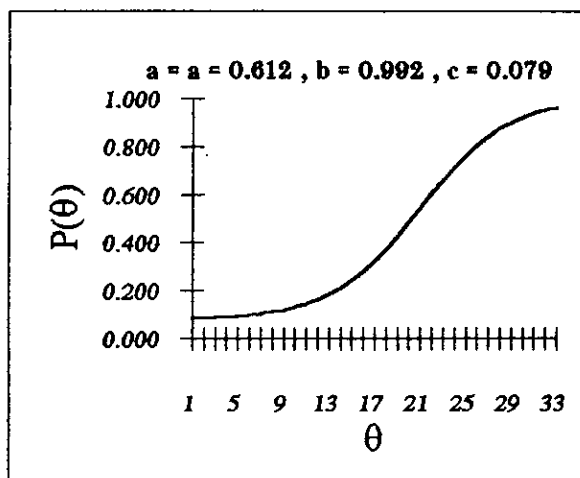


Fig 7

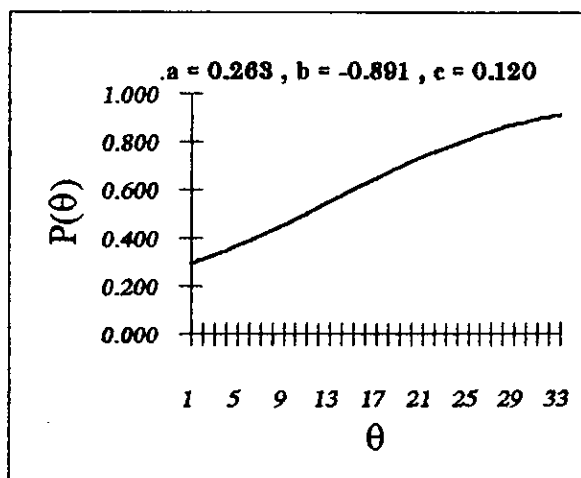


Fig 8

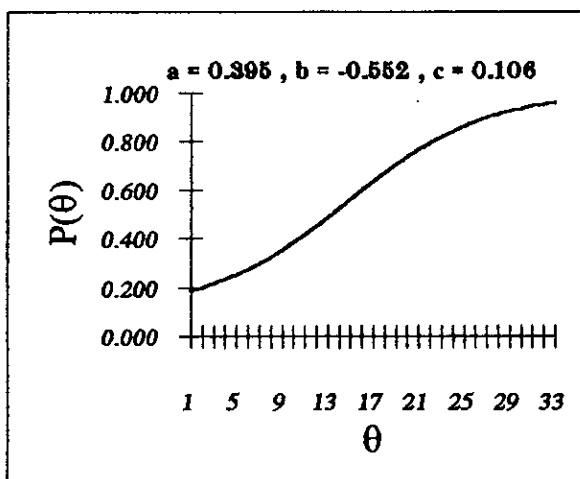


Fig 9

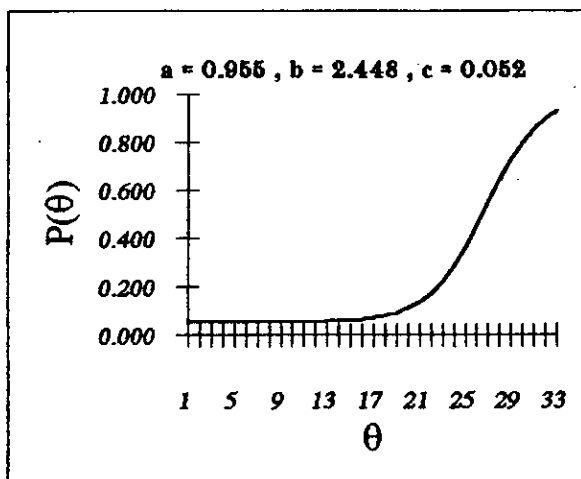


Fig 10

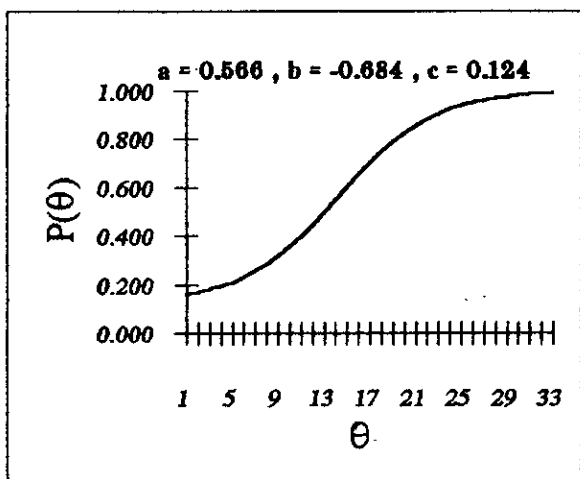


Fig 11

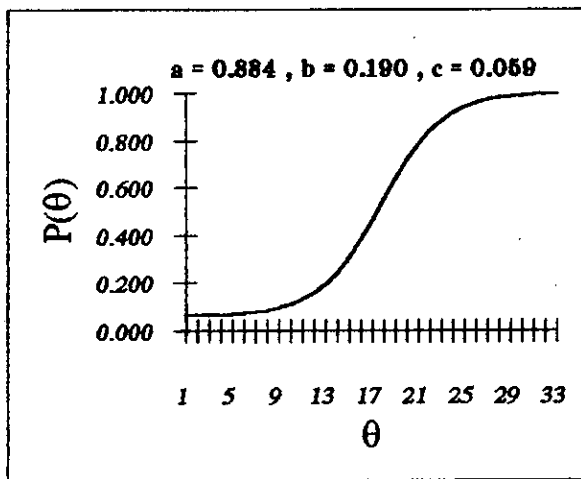


Fig 12

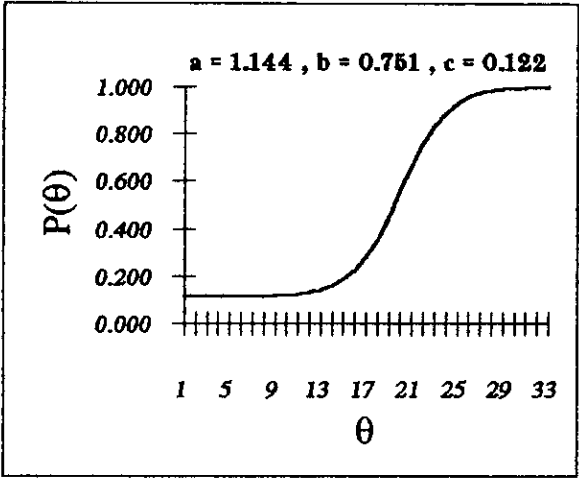


Figure 13

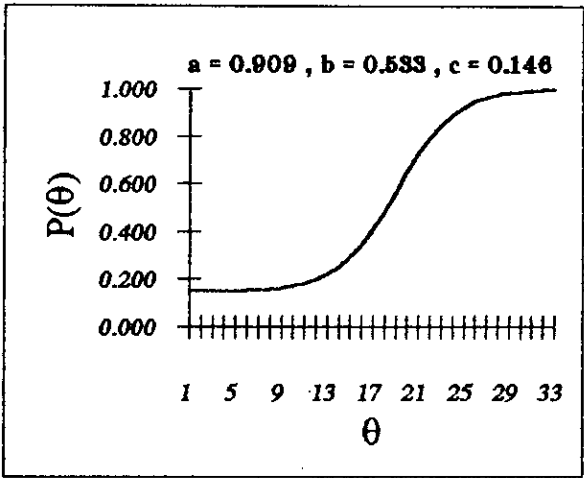


Figure 14

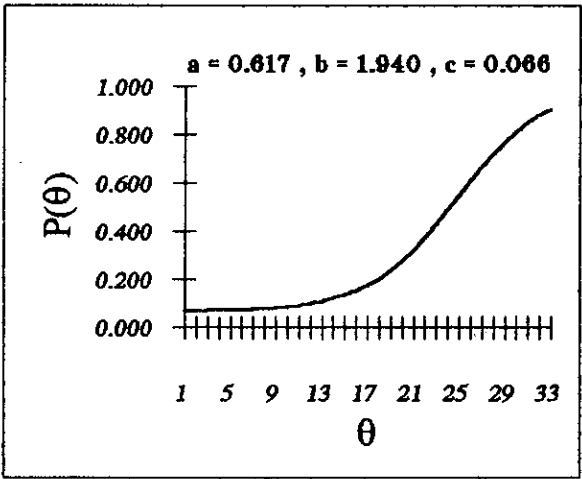


Figure 15

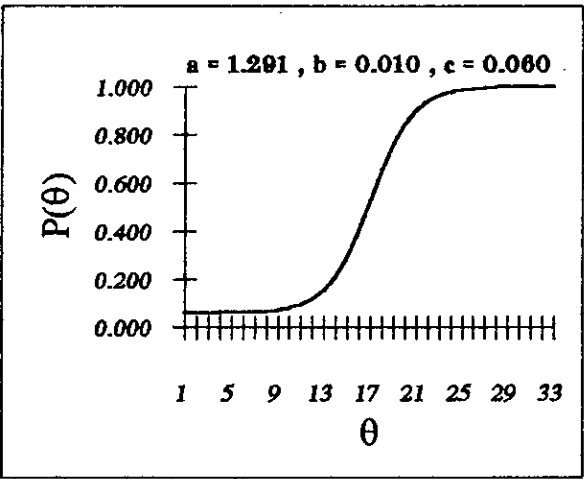


Figure 16

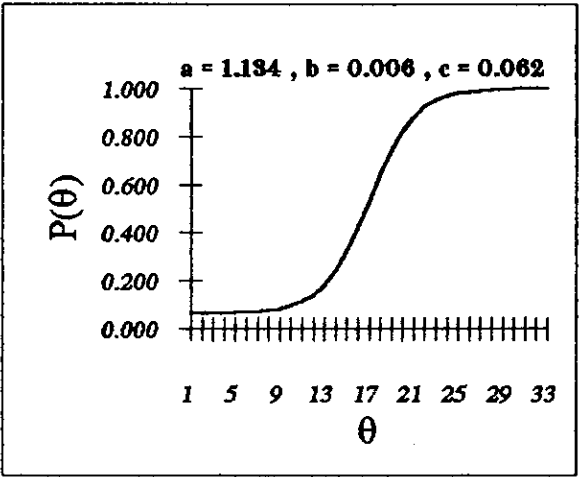


Figure 17

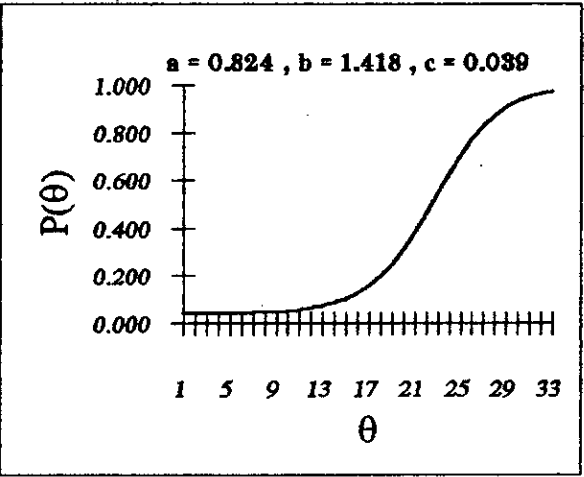
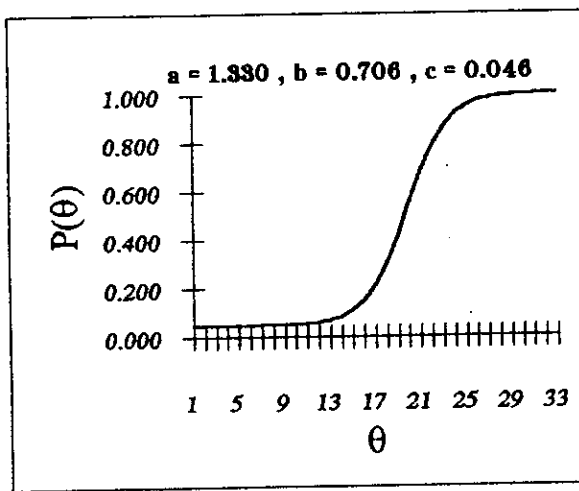
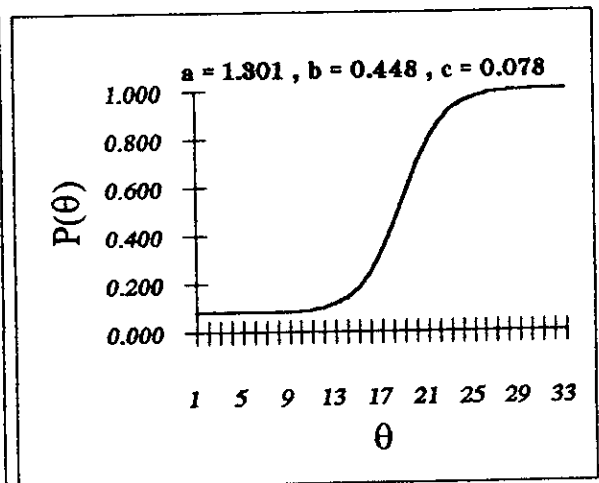


Figure 18



ข้อ 19



ข้อ 20

ภาพประกอบ 75 โค้งลักษณะข้อสอบของข้อสอบ ชนิด AR แบบเติมคำตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3

จากภาพประกอบ 75 พบว่า แบบทดสอบชนิด AR แบบเติมคำตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ข้อสอบข้อ 10 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 3 มีค่าความ
ยากต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 19 และข้อ 8 มีค่าอำนาจจำแนกต่ำ
สุด สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 14 และข้อ 4 มีค่าการเดาต่ำสุด

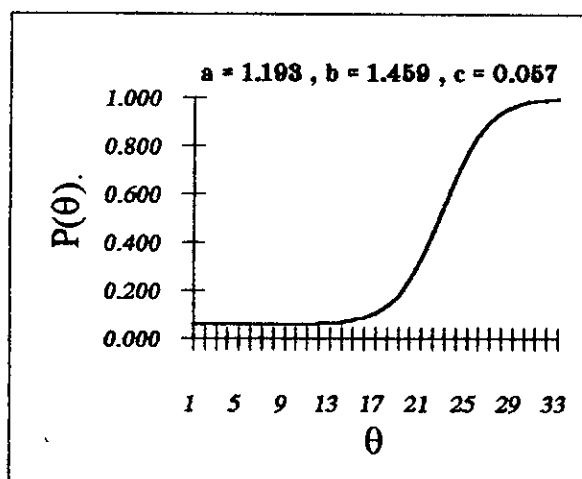


Fig 1

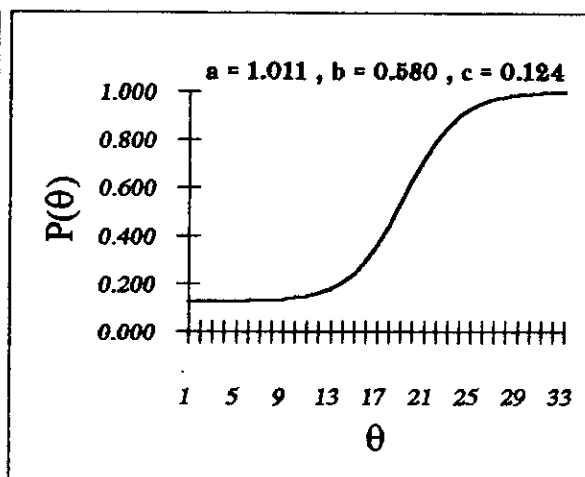


Fig 2

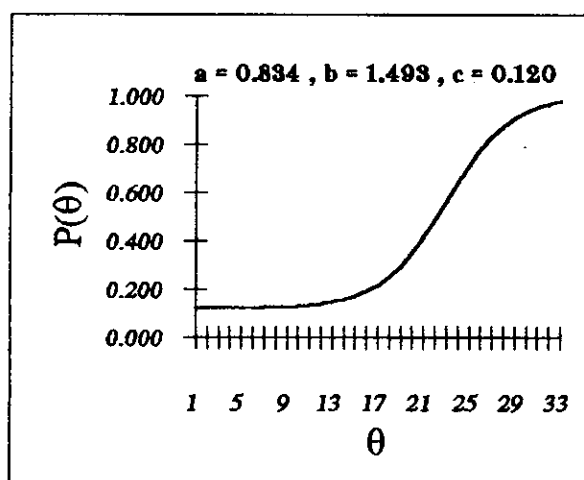


Fig 3

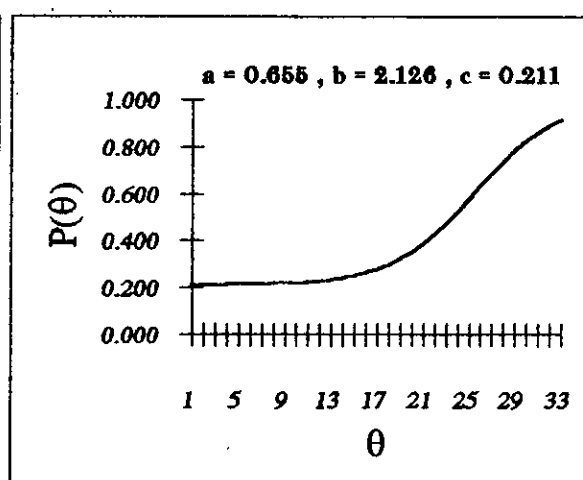


Fig 4

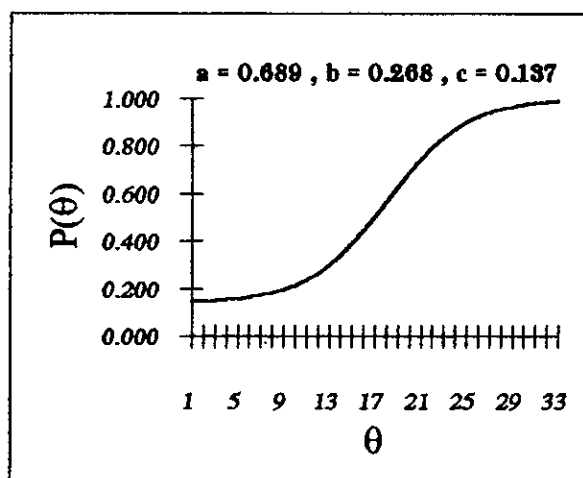


Fig 5

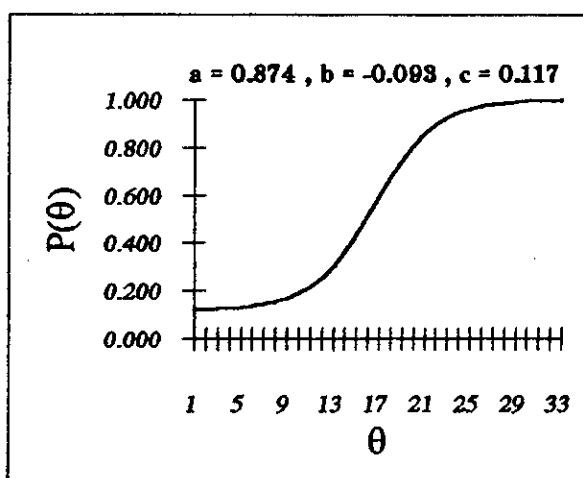


Fig 6

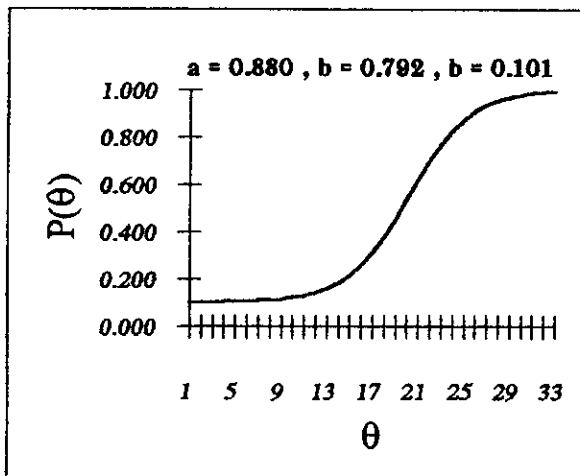


Fig 7

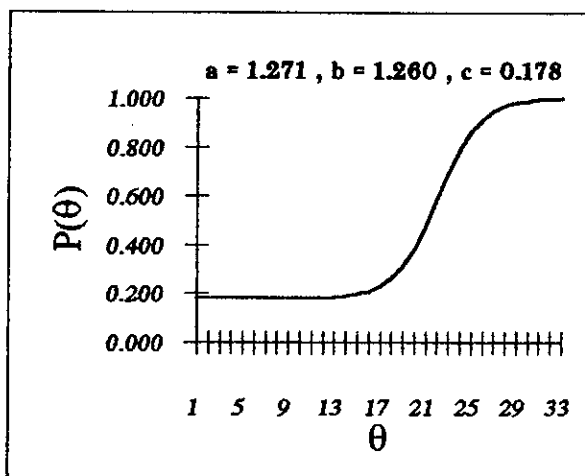


Fig 8

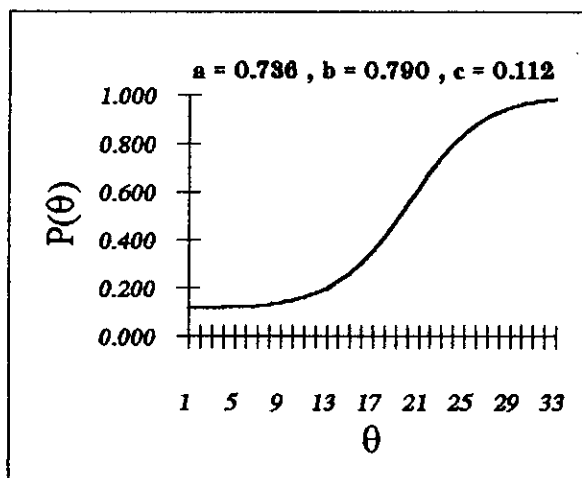


Fig 9

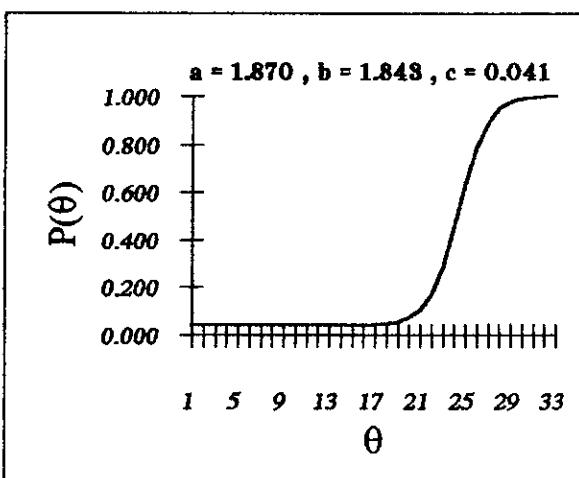


Fig 10

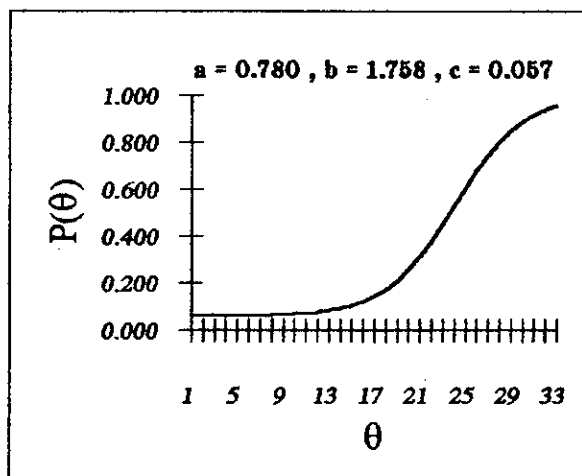


Fig 11

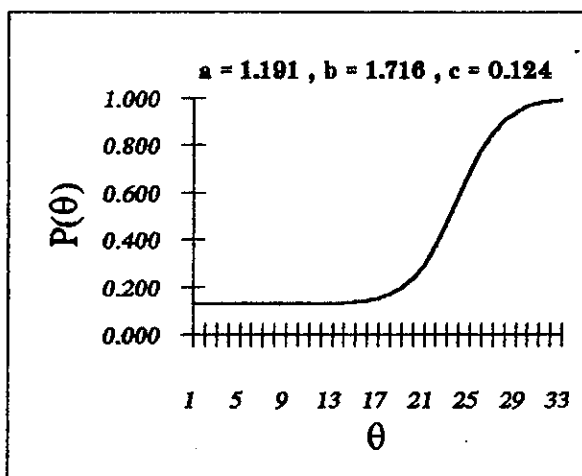
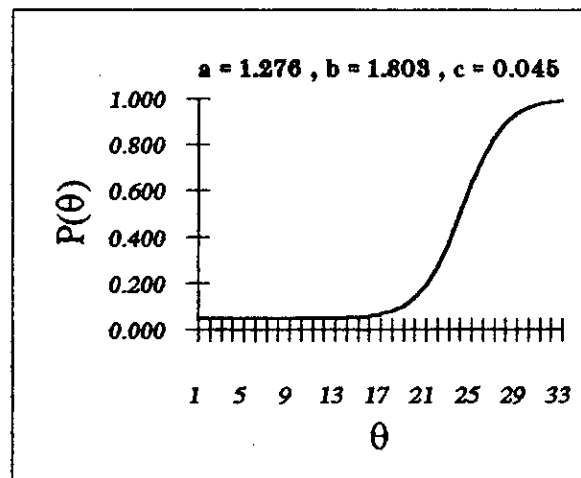


Fig 12



ข้อ 13

ภาพประกอบ 76 ได้ังลักษณะข้อสอบของข้อสอบ ชนิด AX แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3

จากภาพประกอบ 76 พบว่า แบบทดสอบชนิด AX แบบเติมคำตอบของนักเรียน ชั้น
มัธยมศึกษาปีที่ 3 ข้อสอบข้อ 4 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 6 มีค่าความยาก
ต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 10 และข้อ 4 มีค่าอำนาจจำแนกต่ำสุด
สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 4 และข้อ 10 มีค่าการเดาต่ำสุด

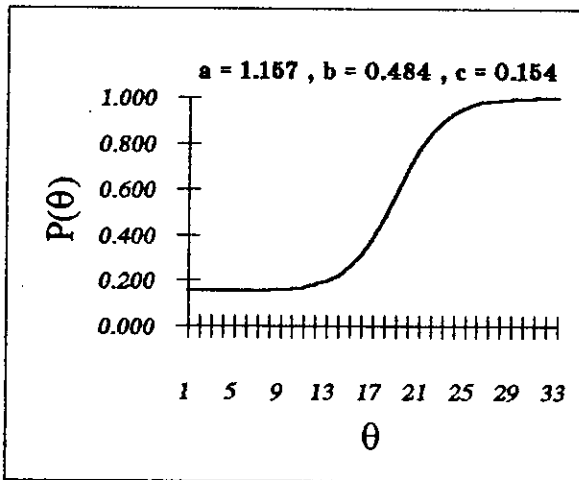


Fig 1

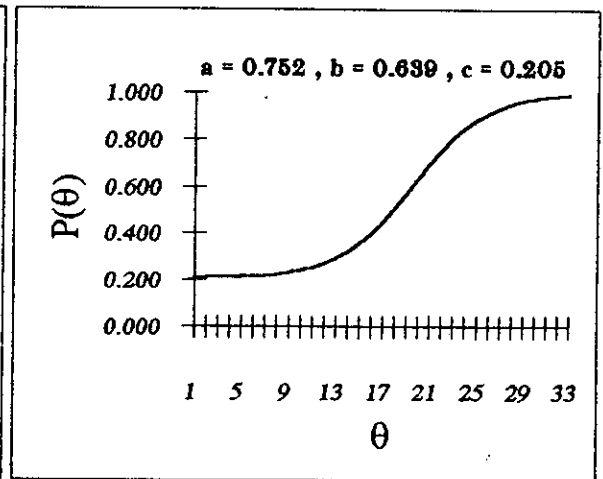


Fig 2

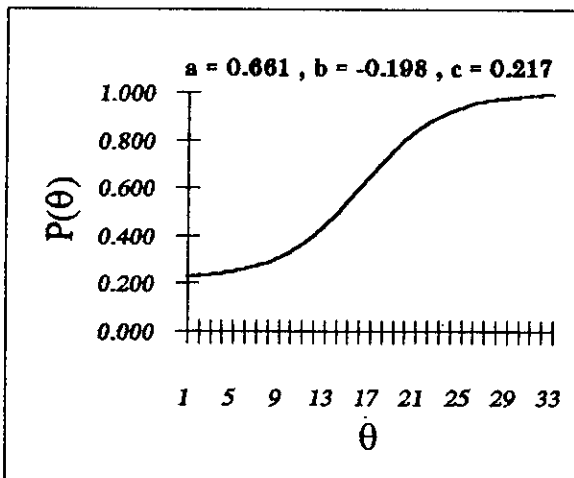


Fig 3

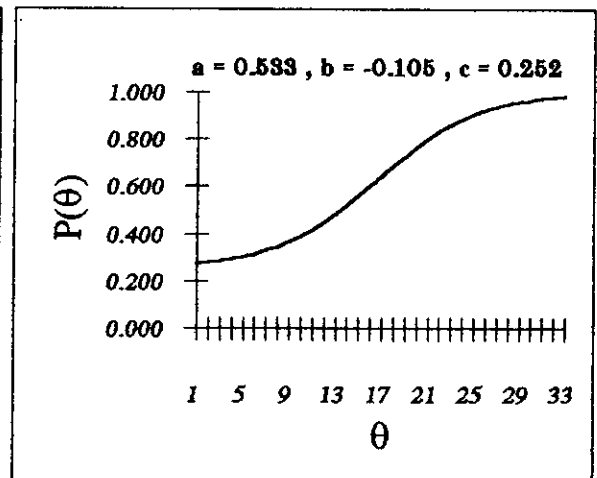


Fig 4

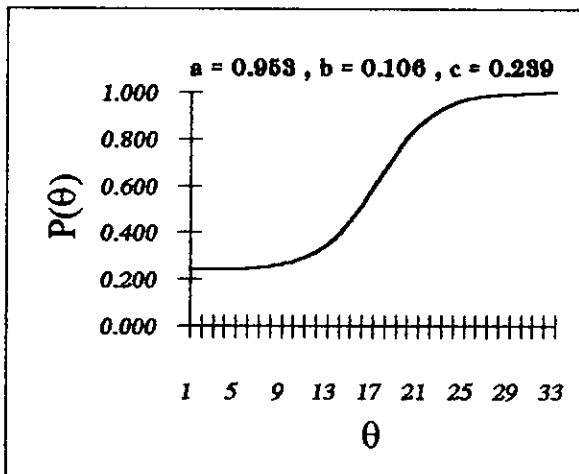


Fig 5

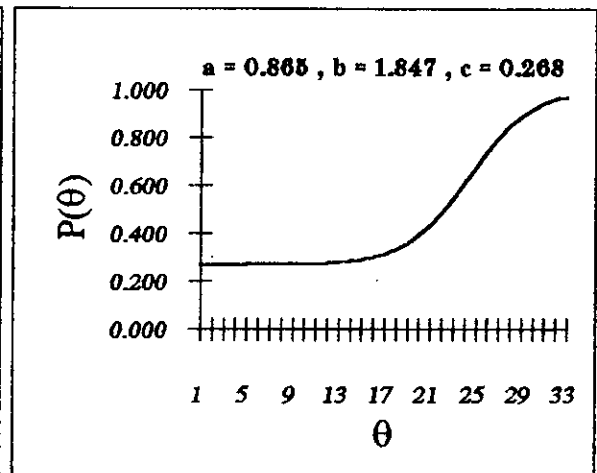
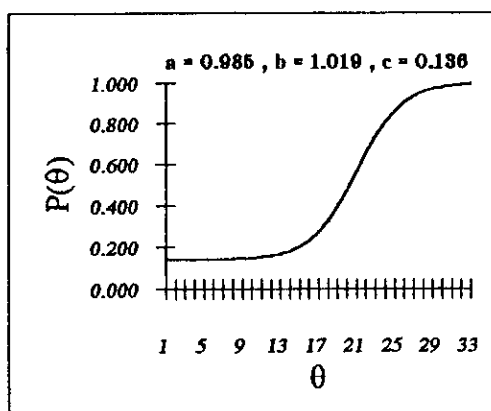
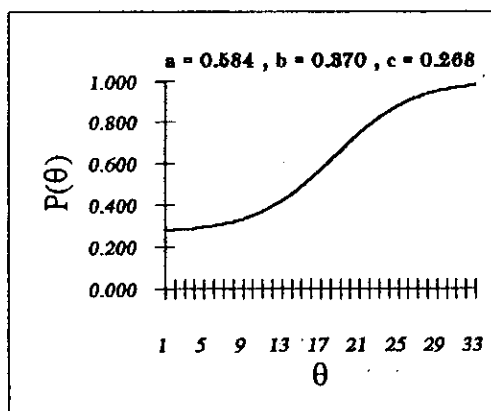


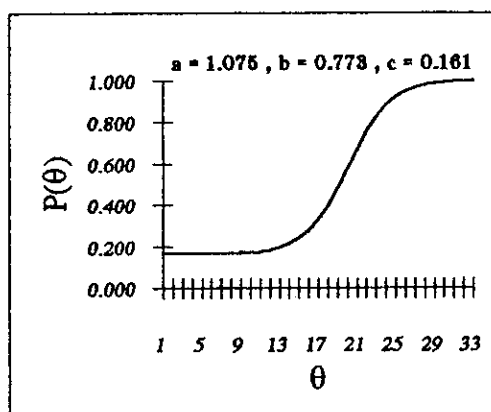
Fig 6



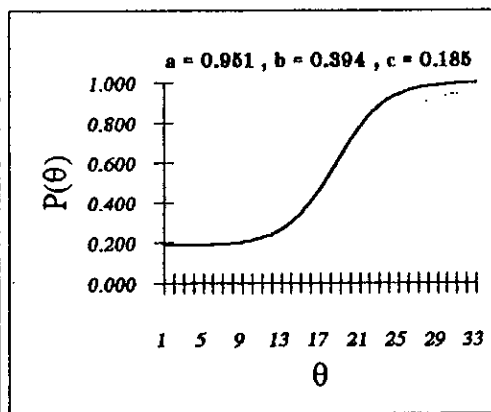
รูปที่ 7



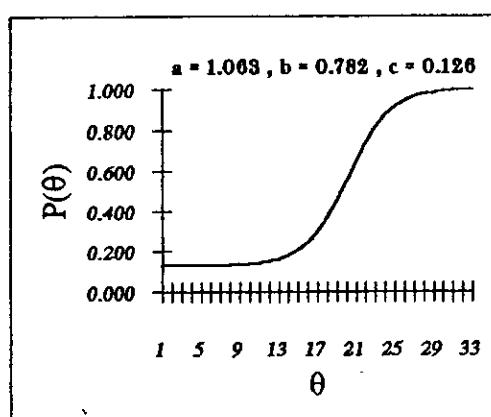
รูปที่ 8



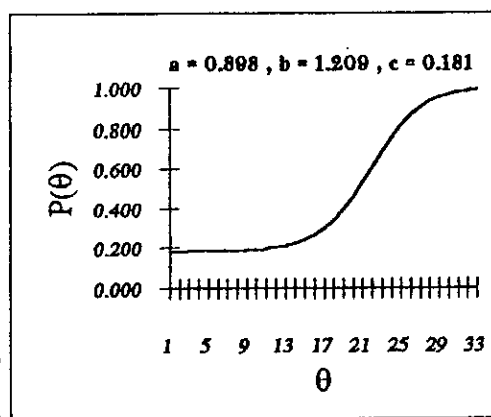
รูปที่ 9



รูปที่ 10



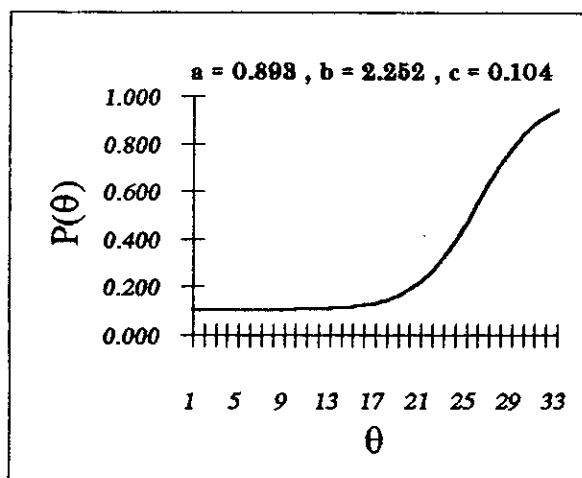
รูปที่ 11



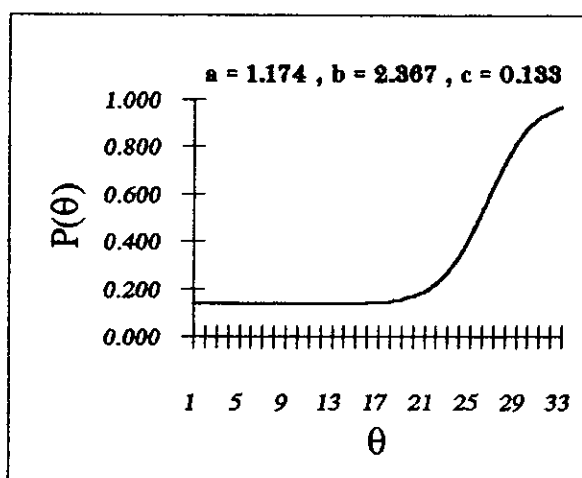
รูปที่ 12

ภาพประกอบ 77 ได้ังลักษณะข้อสอบของข้อสอบ ชนิด CV แบบเติมคำตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3

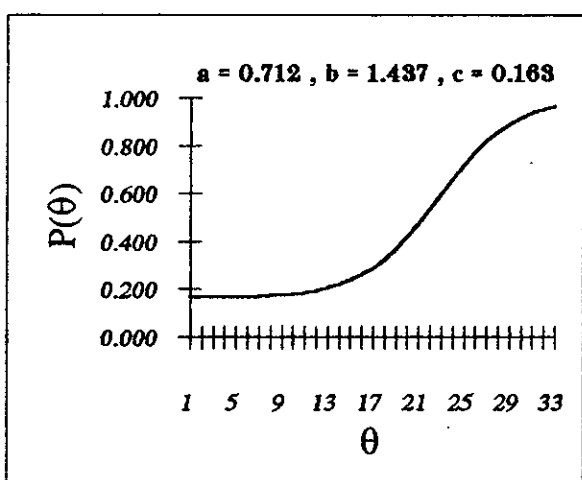
จากภาพประกอบ 77 พบว่า แบบทดสอบชนิด CV แบบเติมคำตอบของนักเรียน ชั้นมัธยมศึกษาปีที่ 3 ข้อสอบข้อ 6 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 3 มีค่าความยากต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 1 และข้อ 4 มีค่าอำนาจจำแนกต่ำสุด สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 6 และข้อ 8 ส่วนข้อ 11 มีค่าการเดาต่ำสุด



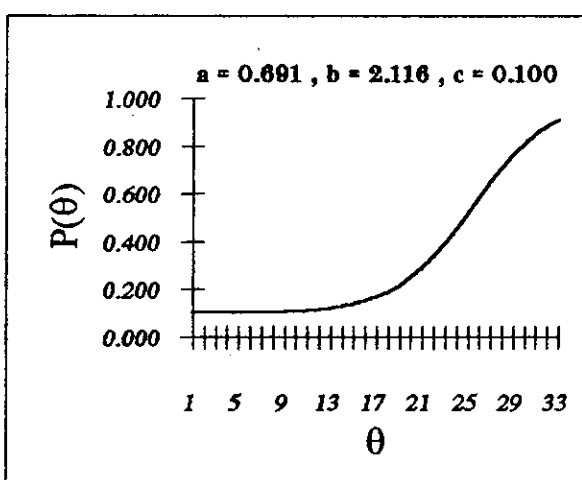
ছবি ১



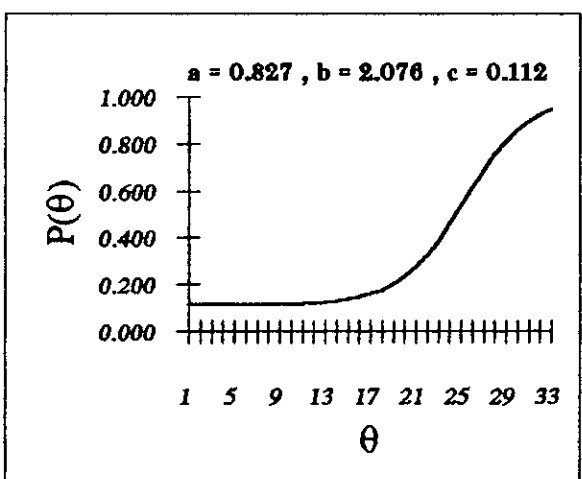
ছবি ২



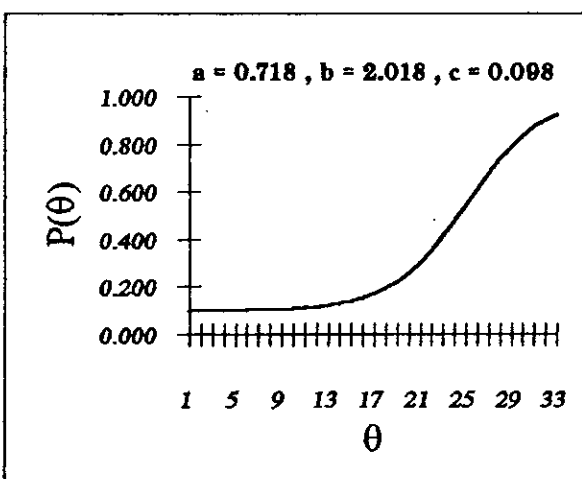
ছবি ৩



ছবি ৪



ছবি ৫



ছবি ৬

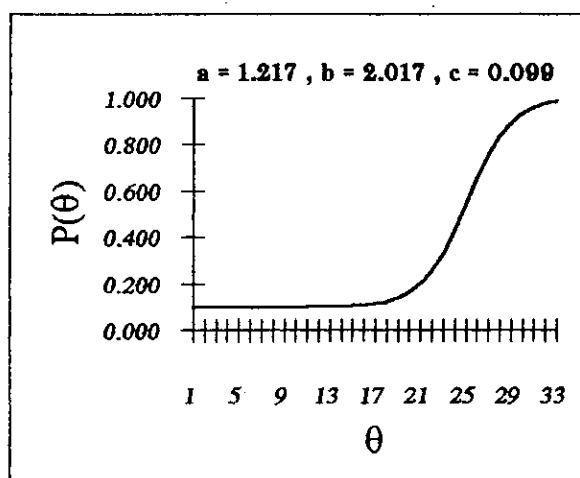


Fig 7

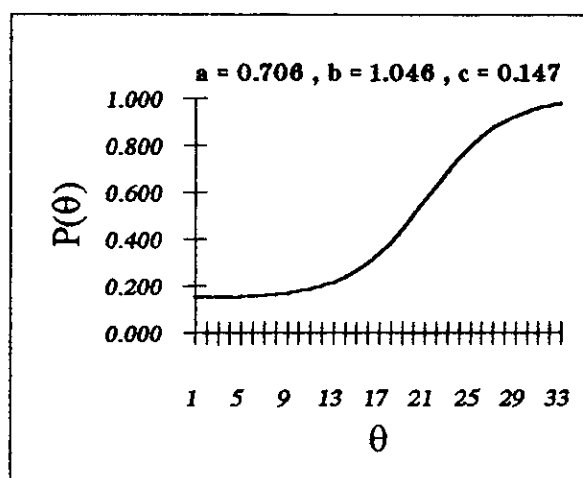


Fig 8

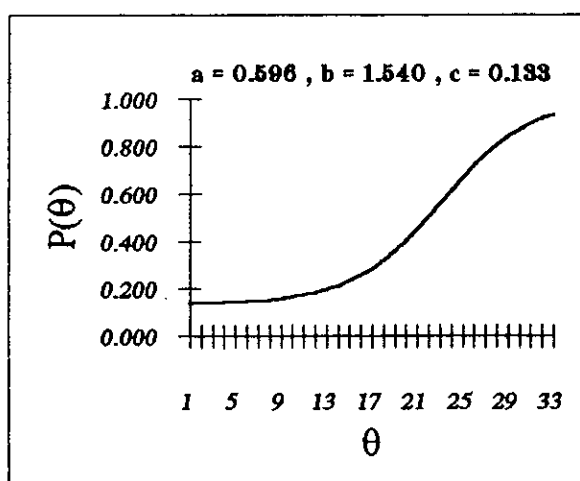


Fig 9

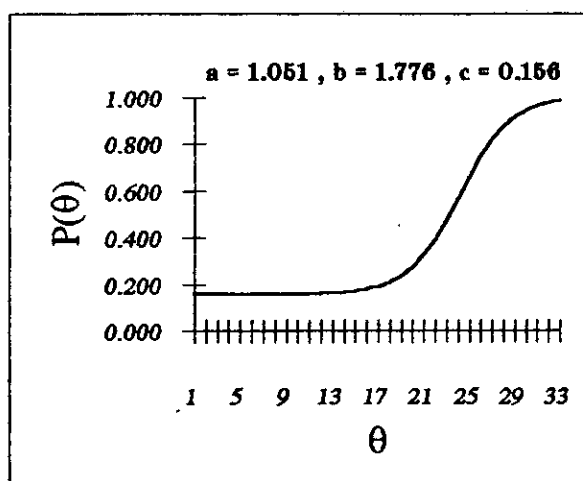


Fig 10

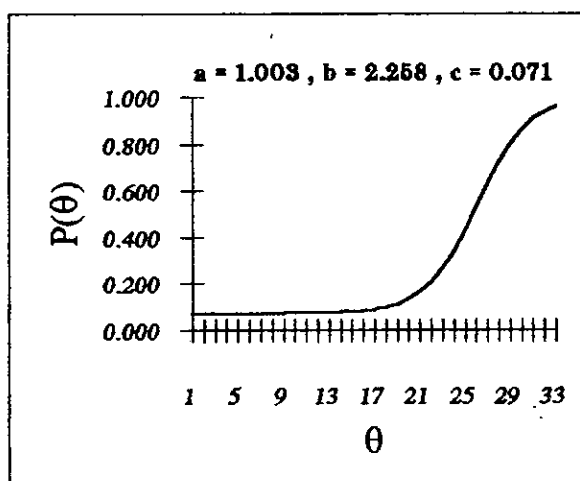


Fig 11

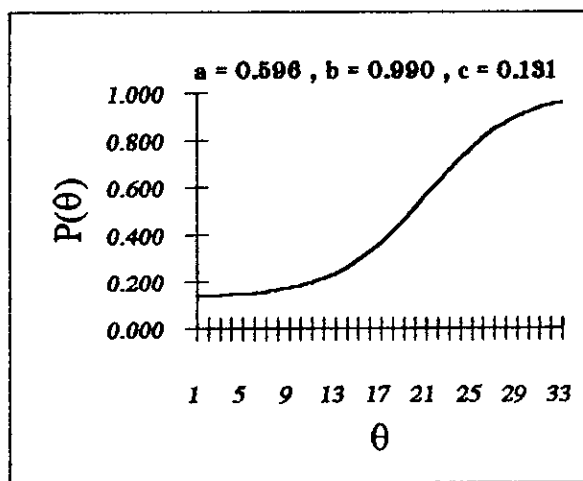
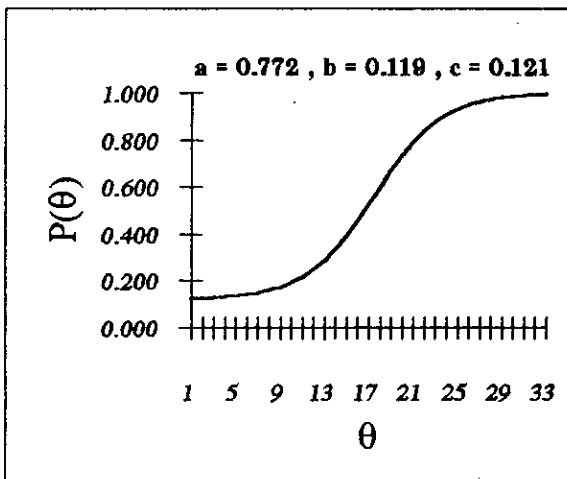
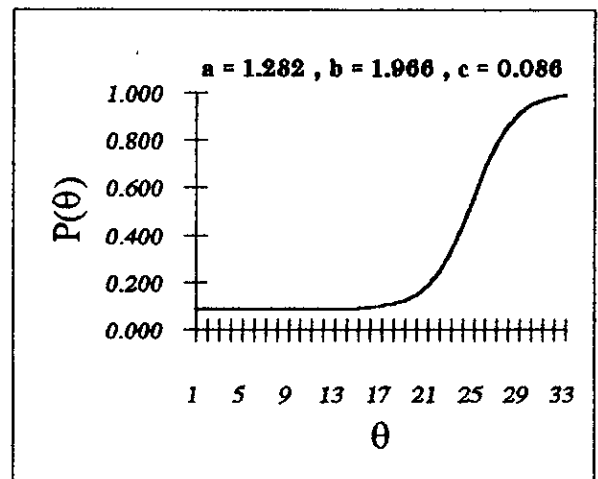


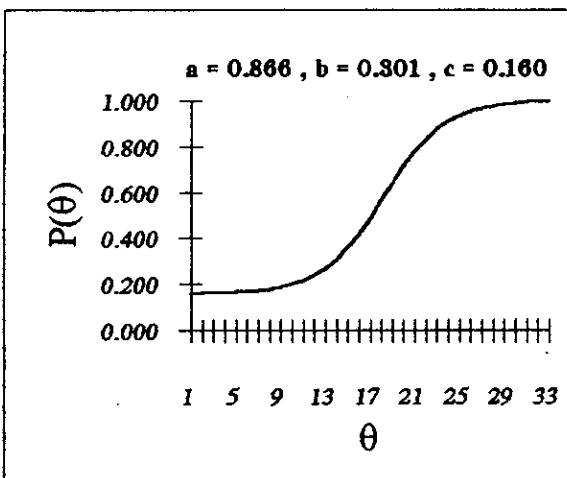
Fig 12



ข้อ 13



ข้อ 14



ข้อ 15

ภาพประกอบ 78 โค้งลักษณะข้อสอบของข้อสอบ ชนิด LR แบบเติมคำตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3

จากภาพประกอบ 78 พบว่า แบบทดสอบชนิด LR แบบเติมคำตอบของนักเรียน ชั้น
มัธยมศึกษาปีที่ 3 ข้อสอบข้อ 2 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 13 มีค่าความยาก
ต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 14 ส่วนข้อ 9 และข้อ 12 มีค่าอำนาจ
จำแนกต่ำสุด สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 3 และข้อ 11 มีค่าการเดาต่ำสุด

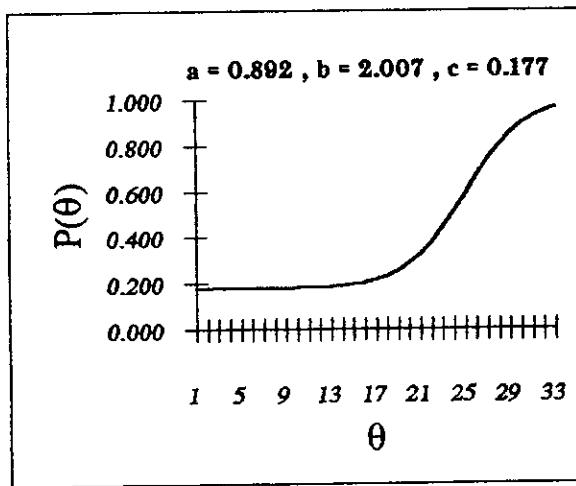


Fig 1

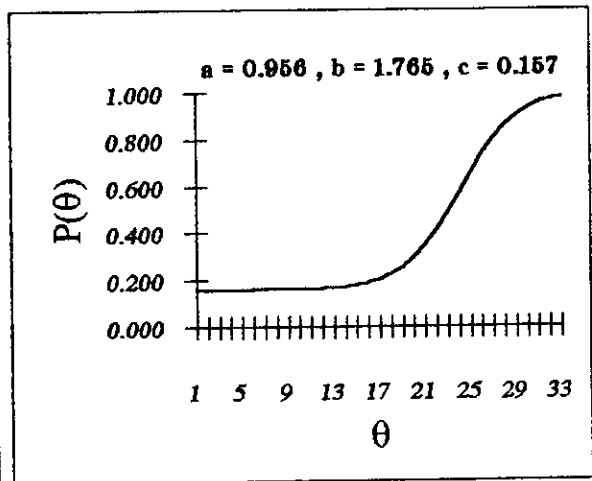


Fig 2

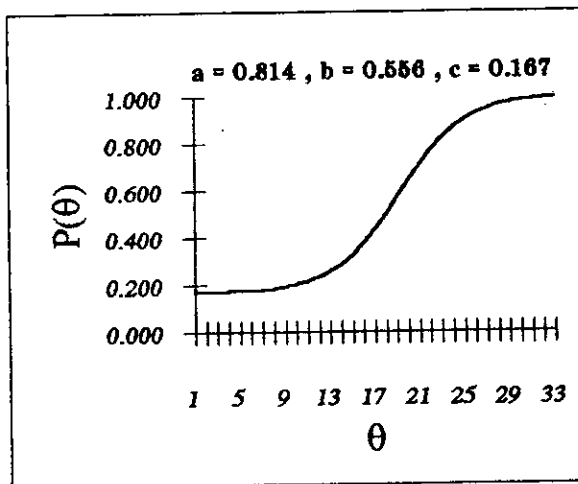


Fig 3

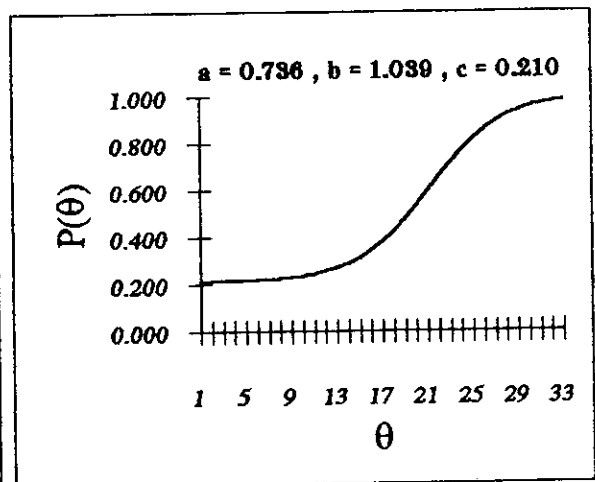


Fig 4

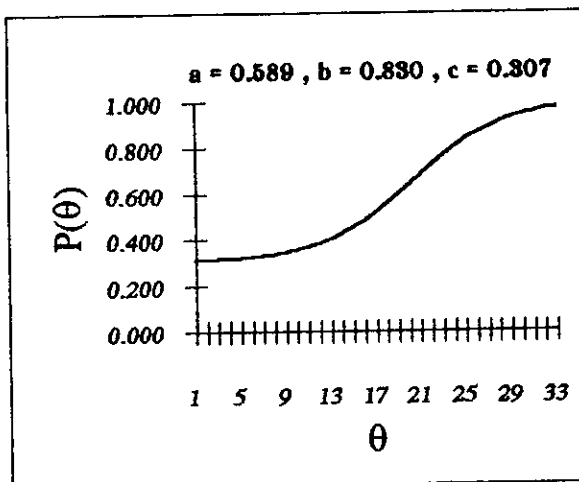


Fig 5

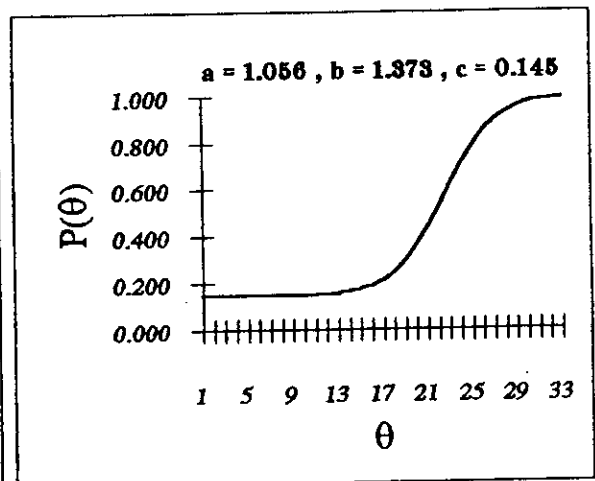
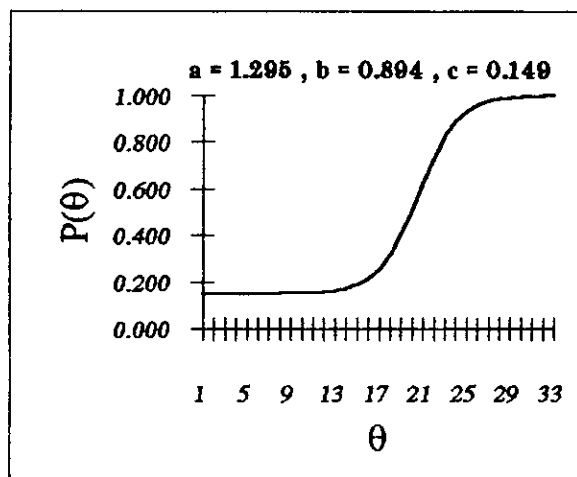
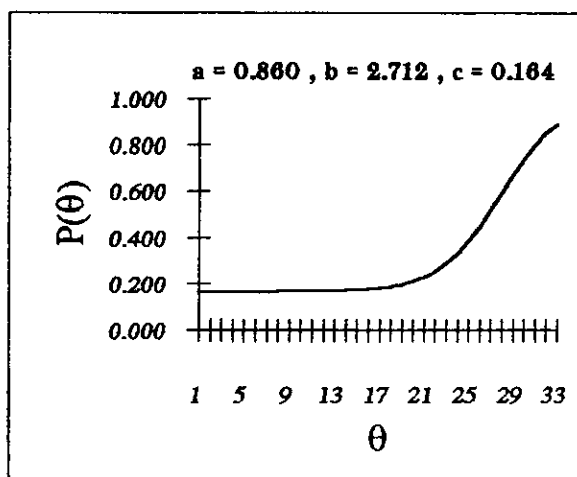


Fig 6



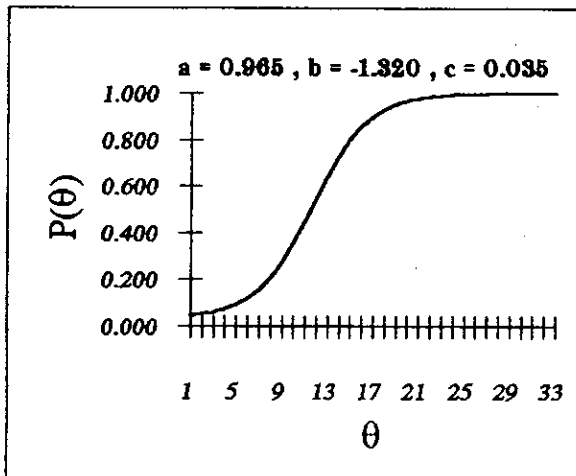
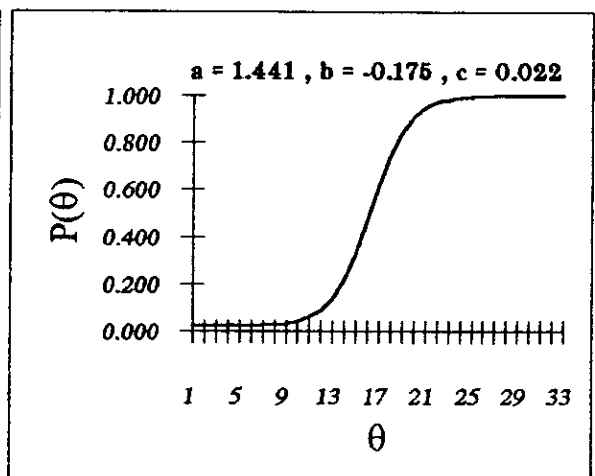
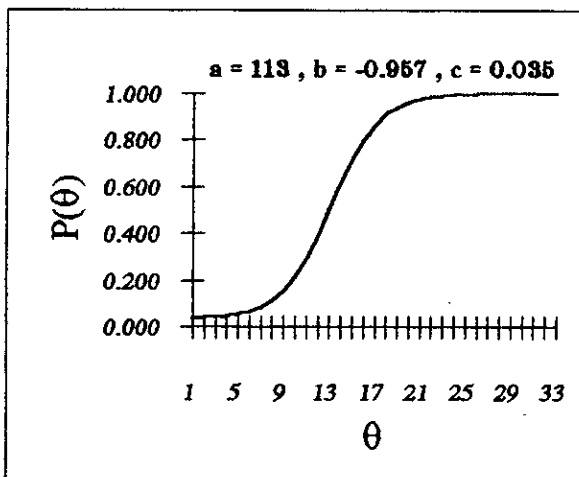
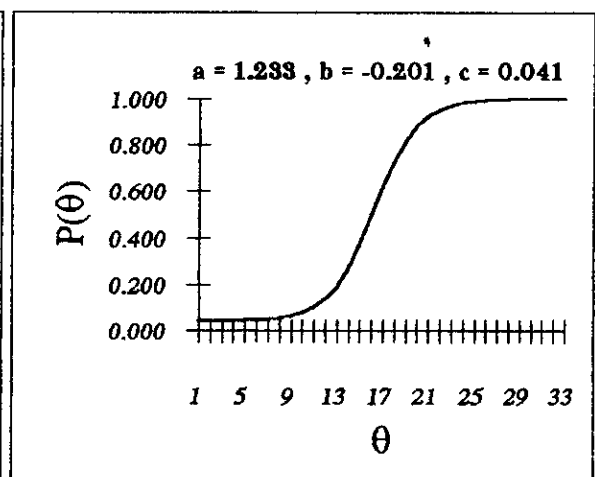
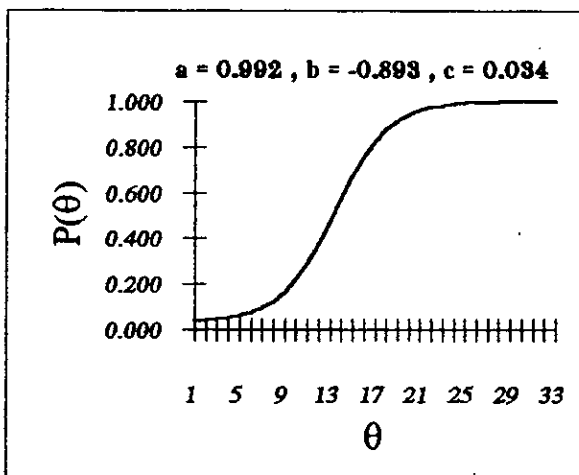
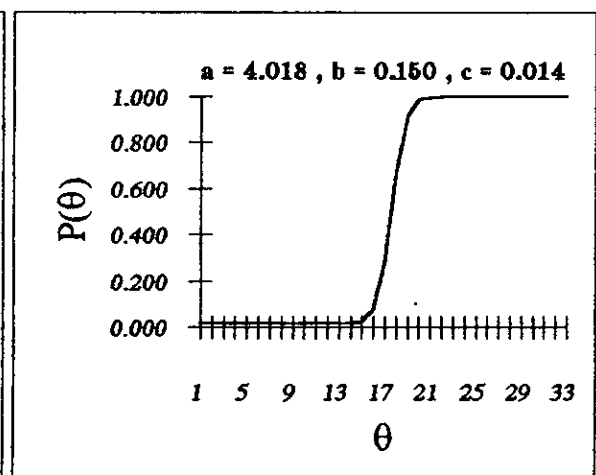
ข้อ 7

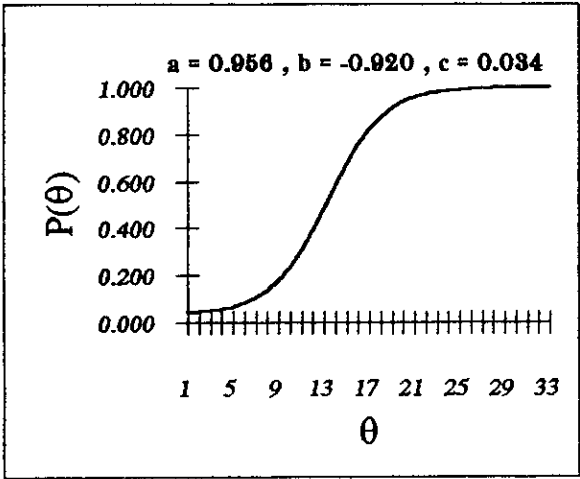


ข้อ 8

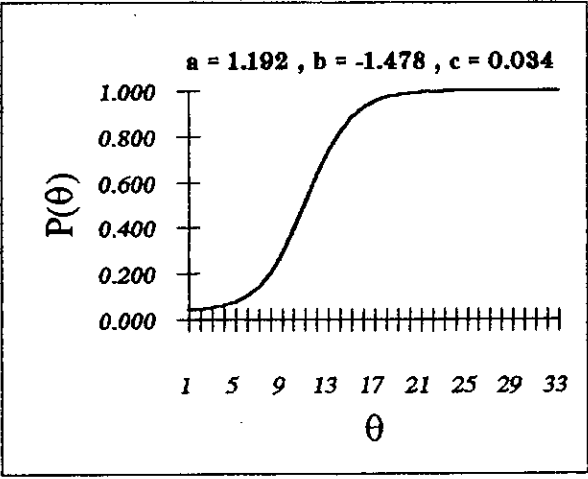
ภาพประกอบ 79 โค้งลักษณะข้อสอบของข้อสอบ ชนิด NLR แบบเติมคำตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3

จากภาพประกอบ 79 พบว่า แบบทดสอบชนิด NLR แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3 ข้อสอบข้อ 8 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 3 มีค่าความ
ยากต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 7 และข้อ 5 มีค่าอำนาจจำแนกต่ำสุด
สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 5 และข้อ 6 มีค่าการเดาต่ำสุด

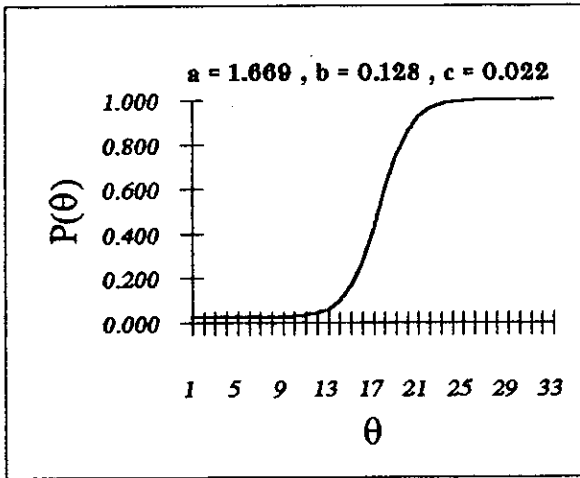
 $\theta 1$  $\theta 2$  $\theta 3$  $\theta 4$  $\theta 5$  $\theta 6$



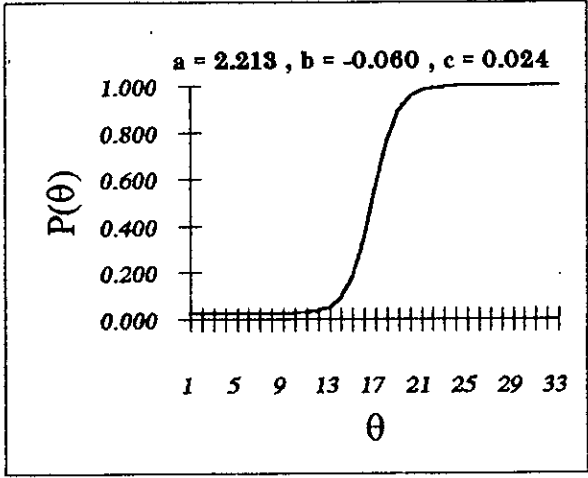
θ 7



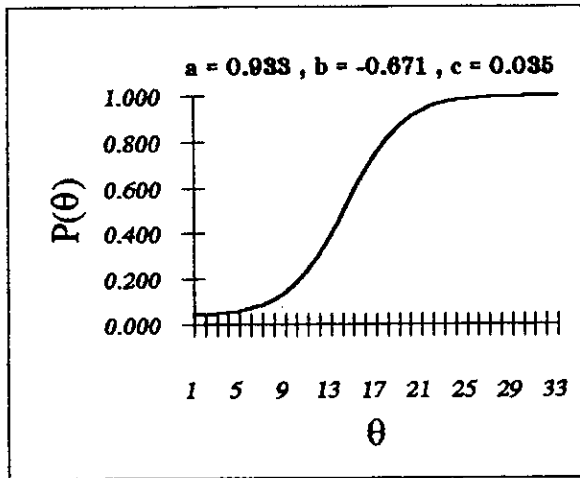
θ 8



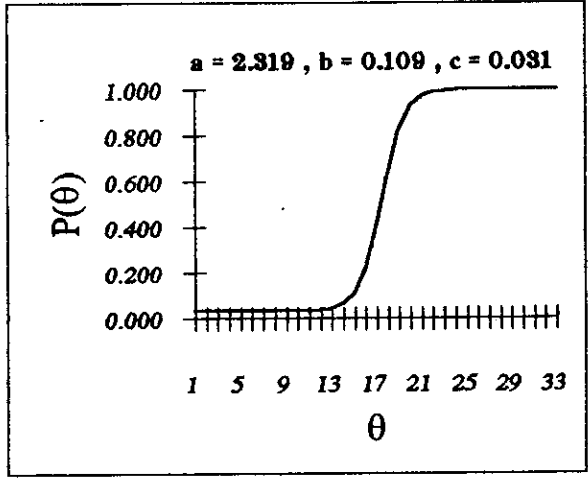
θ 9



θ 10



θ 11



θ 12

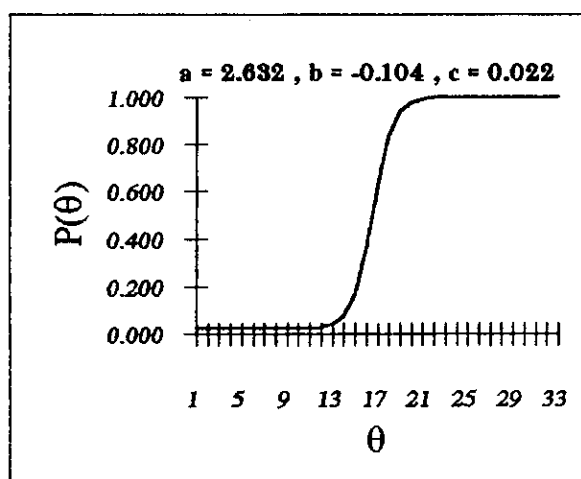


Fig 13

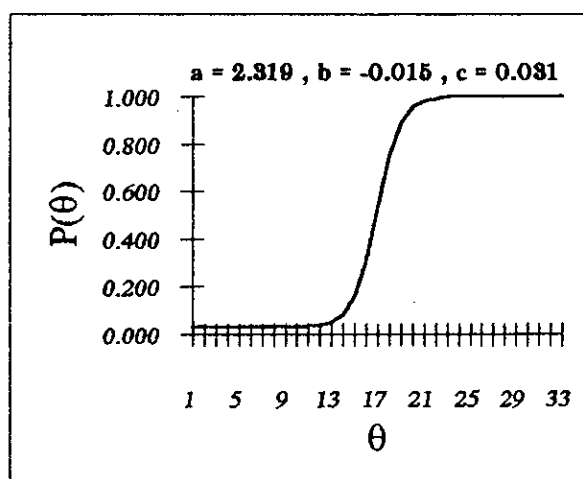


Fig 14

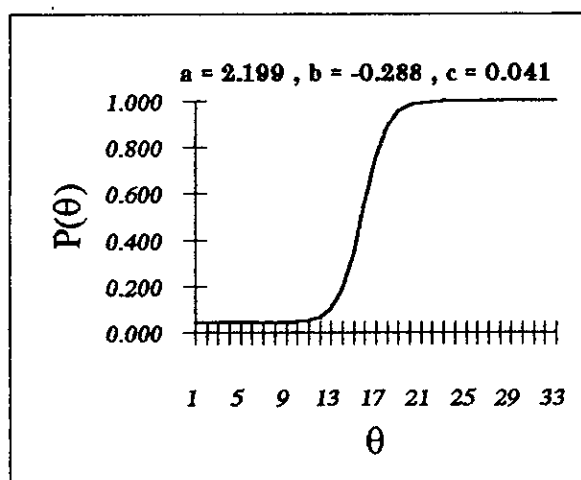


Fig 15

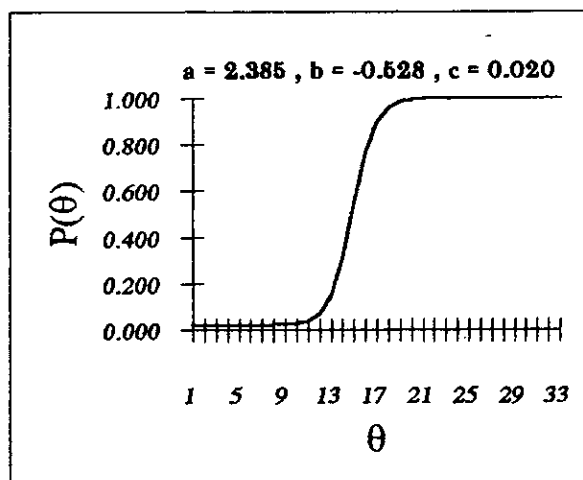


Fig 16

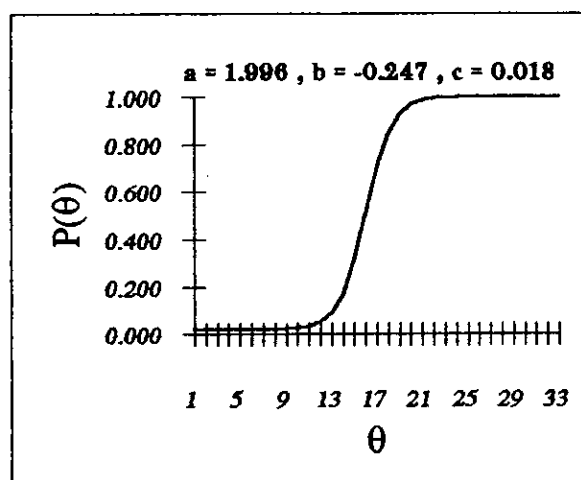


Fig 17

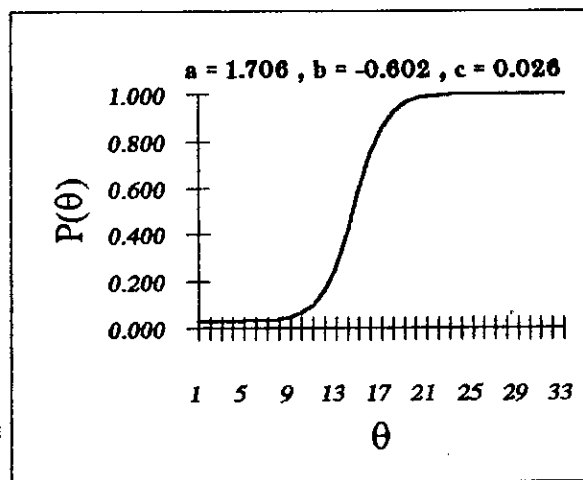
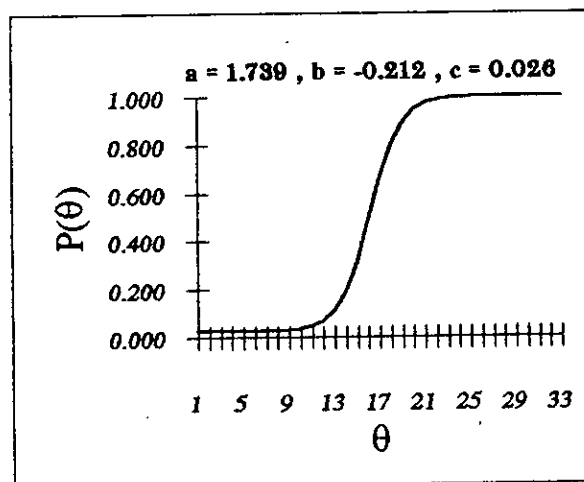


Fig 18



ข้อ 19

ภาพประกอบ 80 โด่งลักษณะข้อสอบของข้อสอบ ชนิด PI แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3

จากภาพประกอบ 80 พบว่า แบบทดสอบชนิด PI แบบเติมคำตอบของนักเรียน ชั้น
มัธยมศึกษาปีที่ 3 ข้อสอบข้อ 6 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 8 มีค่าความยาก
ต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 6 และข้อ 11 มีค่าอำนาจจำแนกต่ำสุด
สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 4 ข้อ 15 และข้อ 6 มีค่าการเดาต่ำสุด

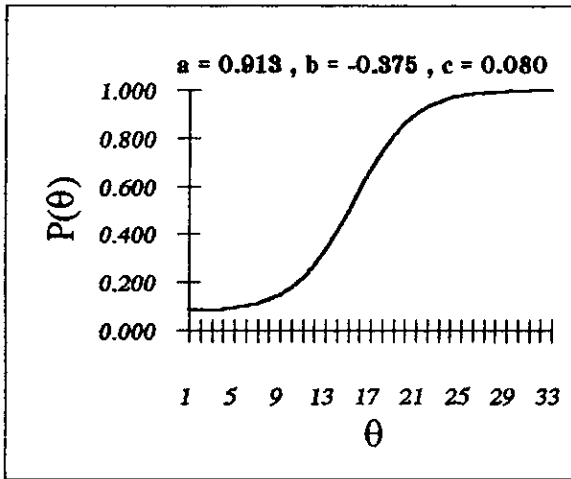


Fig 1

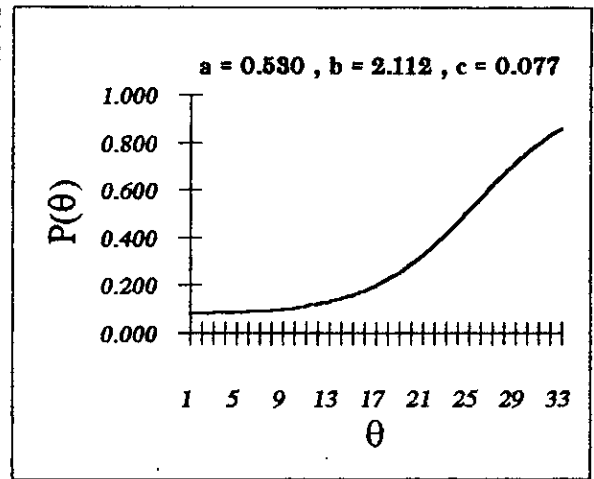


Fig 2

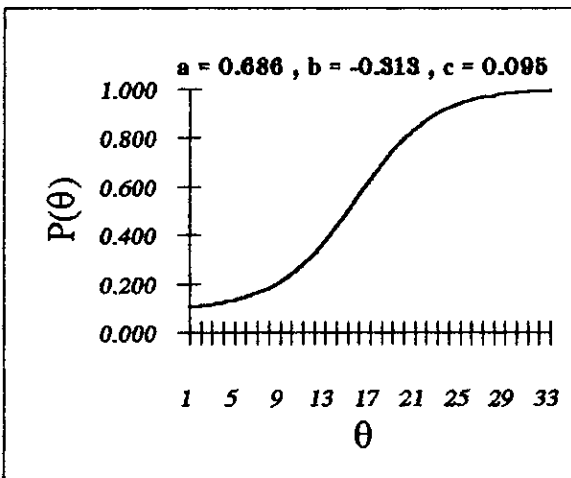


Fig 3

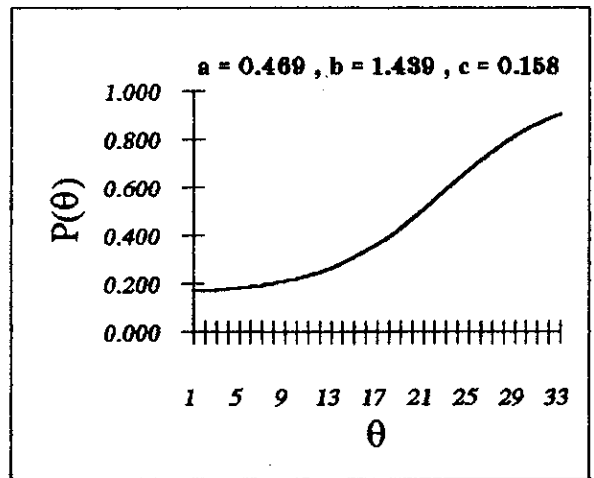


Fig 4

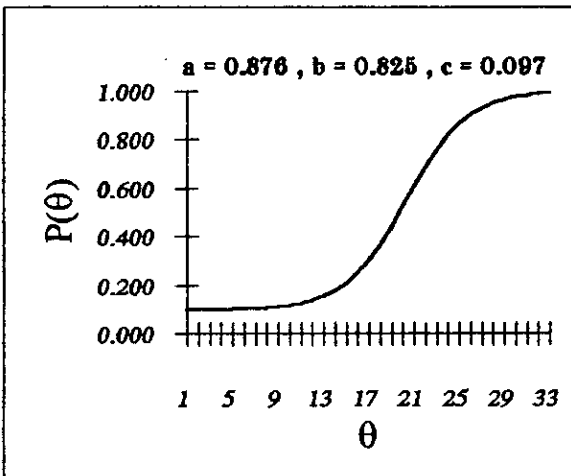


Fig 5

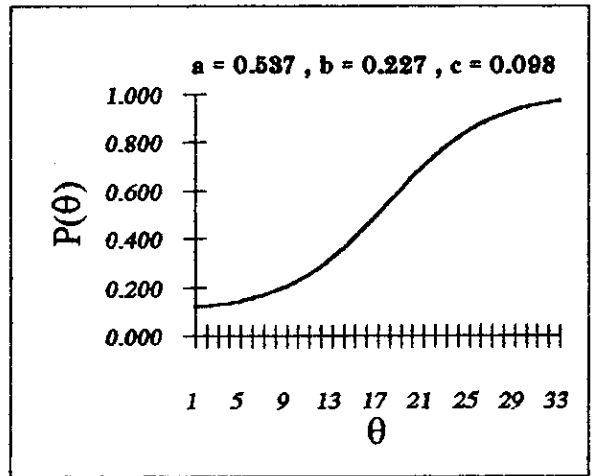


Fig 6

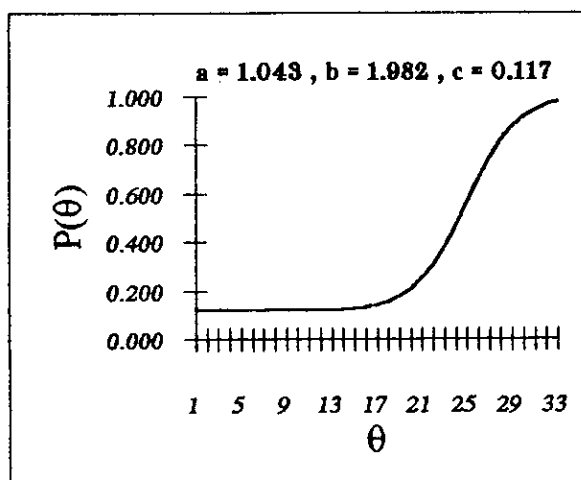


Fig 7

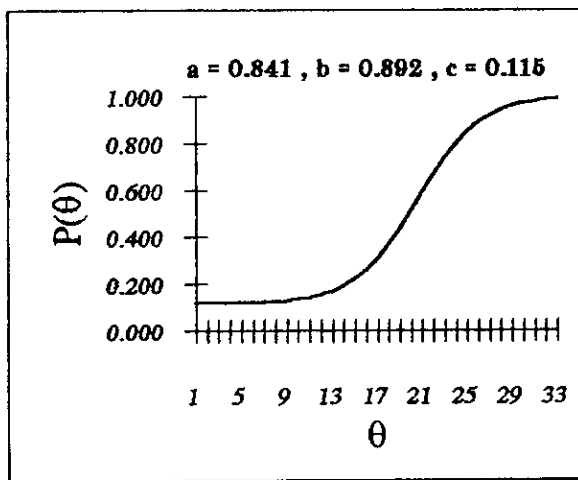


Fig 8

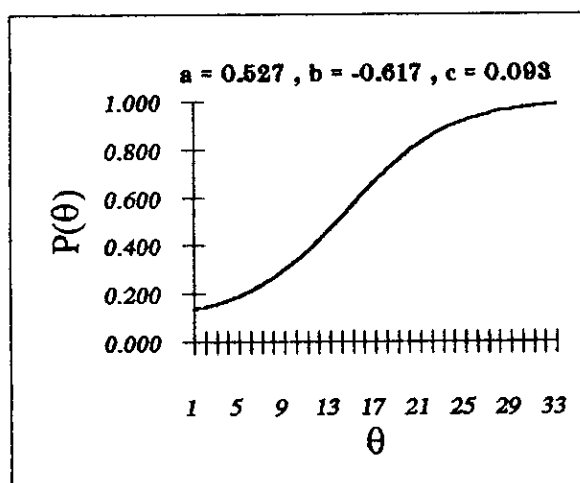


Fig 9

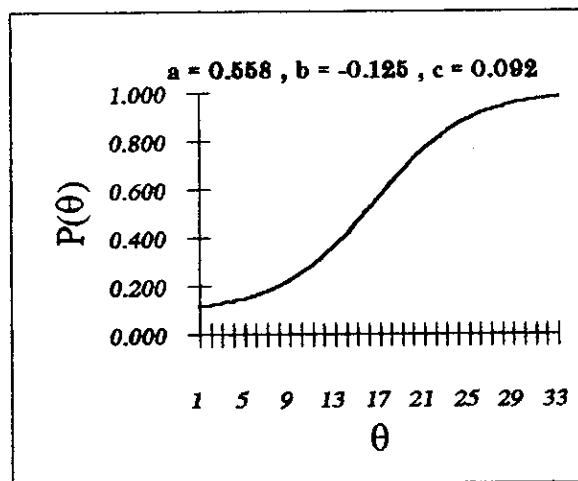


Fig 10

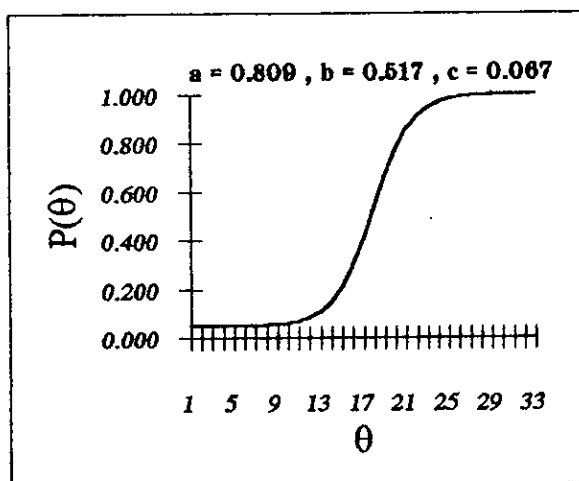


Fig 11

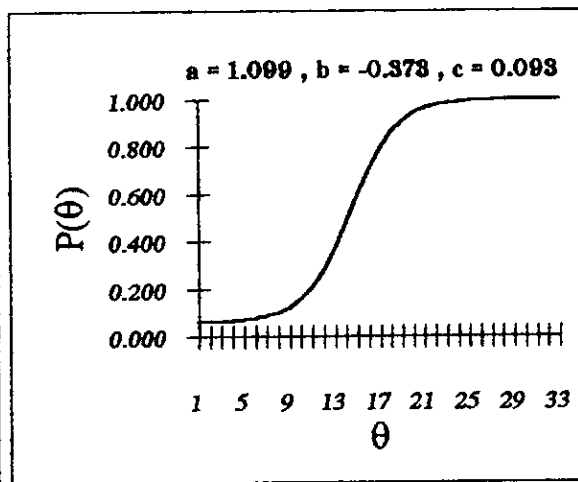
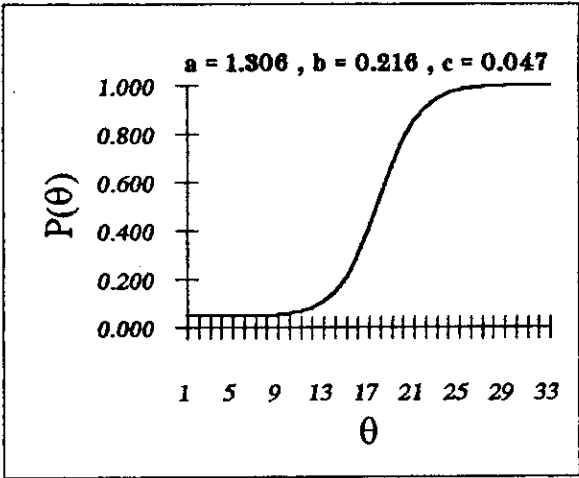
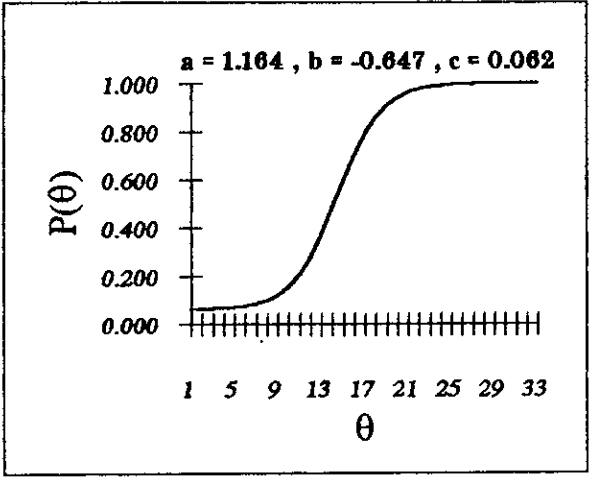


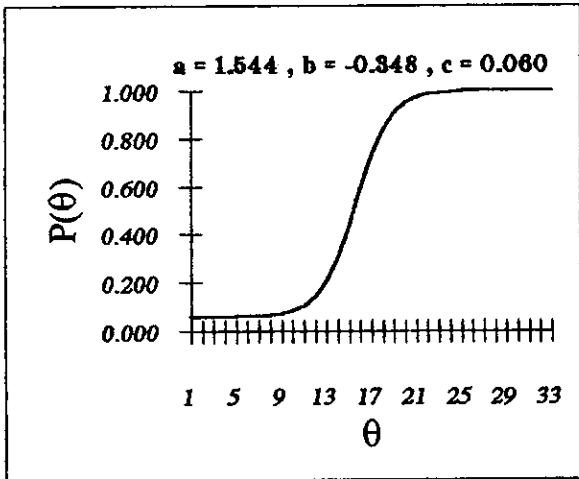
Fig 12



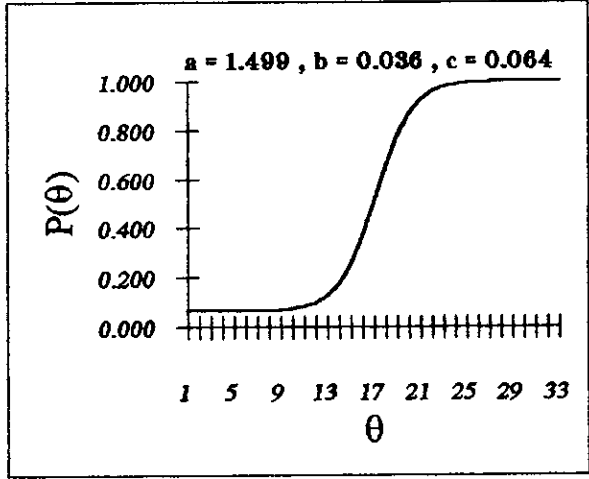
$\theta_0 13$



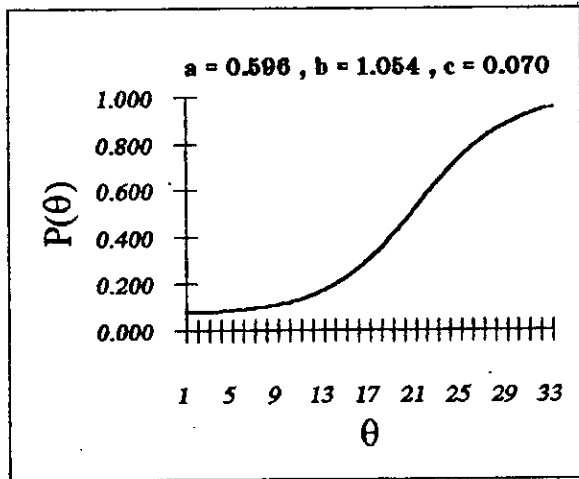
$\theta_0 14$



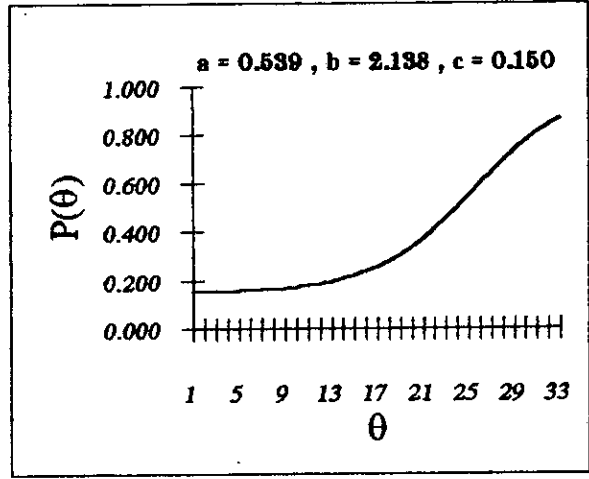
$\theta_0 15$



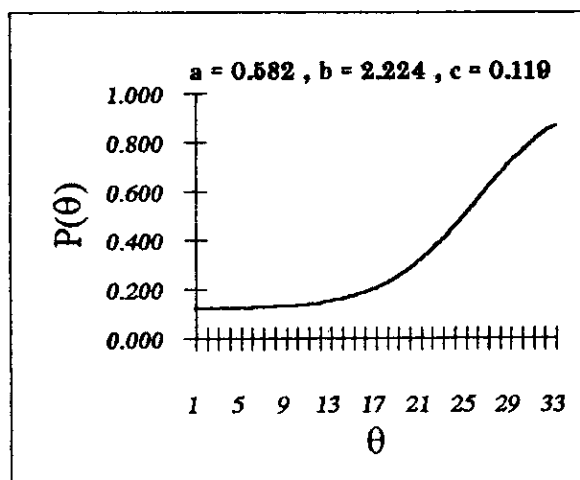
$\theta_0 16$



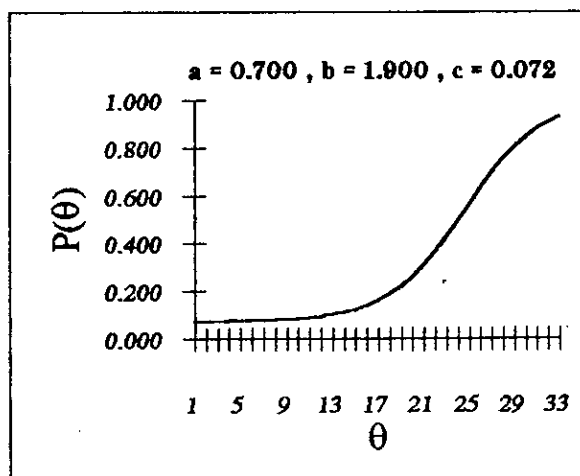
$\theta_0 17$



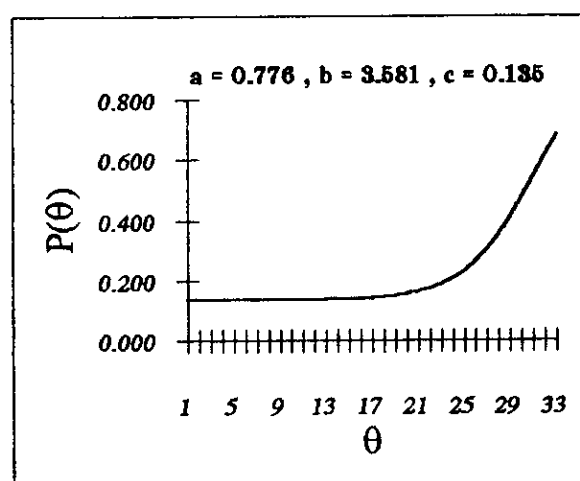
$\theta_0 18$



ข้อ 19



ข้อ 20



ข้อ 21

ภาพประกอบ 81 โต้งลักษณะข้อสอบของข้อสอบ ชนิด AR แบบเลือกตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 81 พบว่า แบบทดสอบชนิด AR แบบเลือกตอบของนักเรียนชั้นมัธยมศึกษาปีที่ 5 ข้อสอบข้อ 21 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 14 มีค่าความยากต่ำสุด โดยข้อ 5 เป็นข้อสอบที่มีค่าความยากระดับปานกลาง ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 15 และข้อ 4 มีค่าอำนาจจำแนกต่ำสุด สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 4 และข้อ 13 มีค่าการเดาต่ำสุด

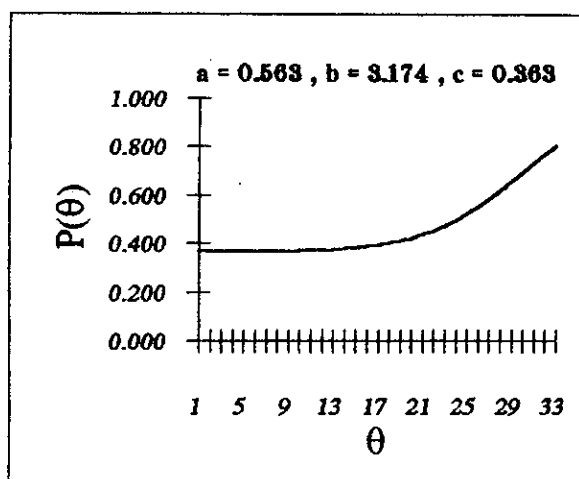


Fig 1

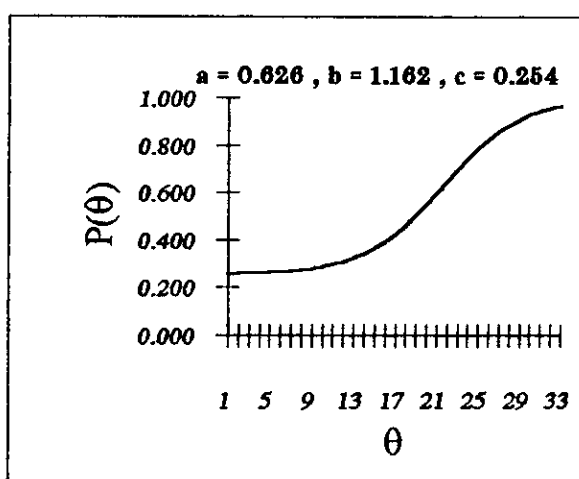


Fig 2

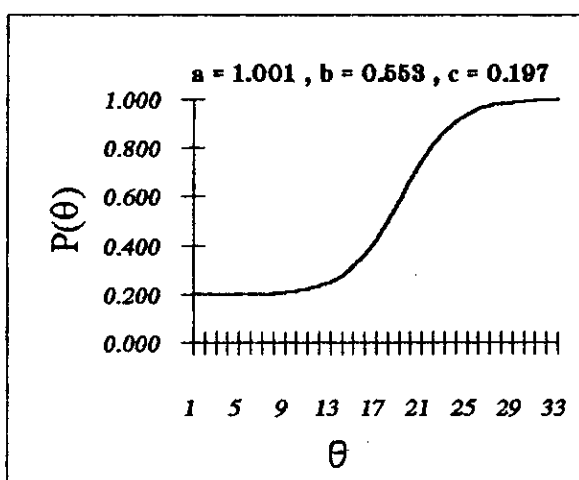


Fig 3

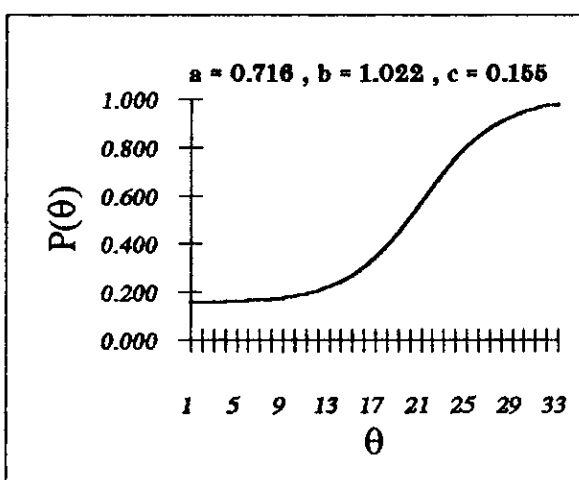


Fig 4

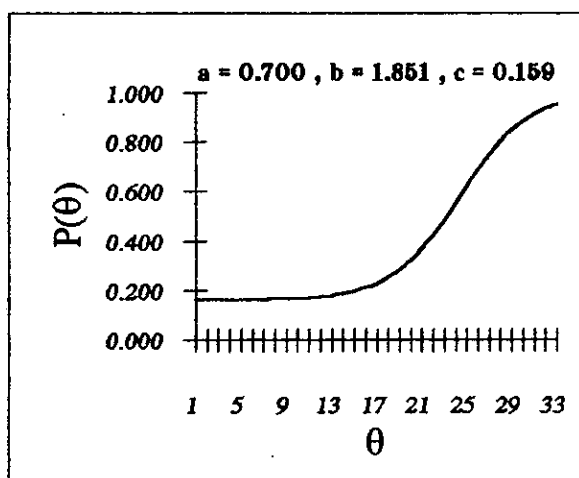


Fig 5

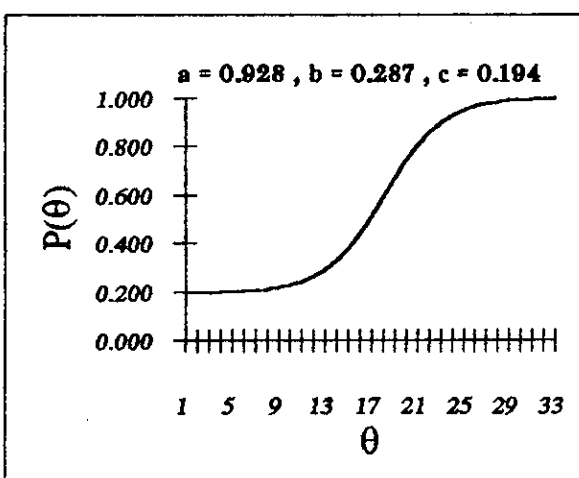


Fig 6

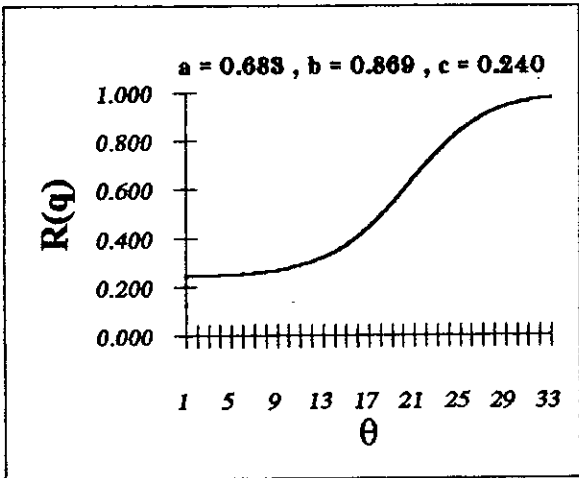


Fig 7

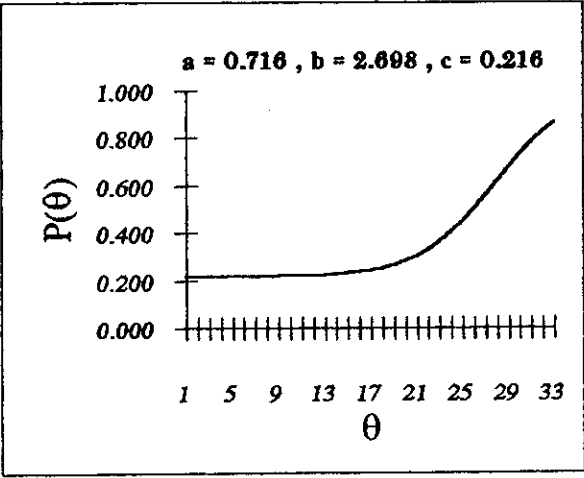


Fig 8

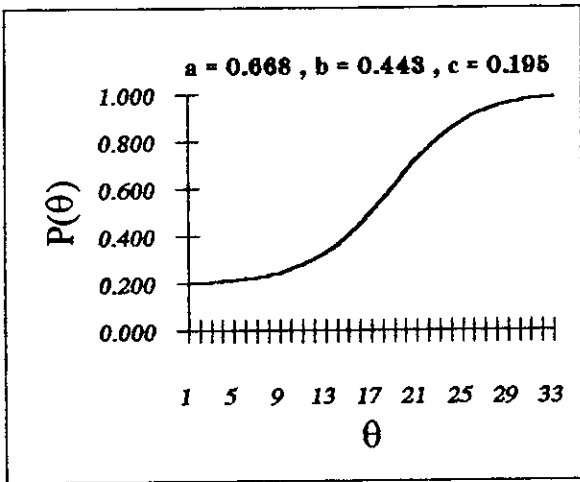


Fig 9

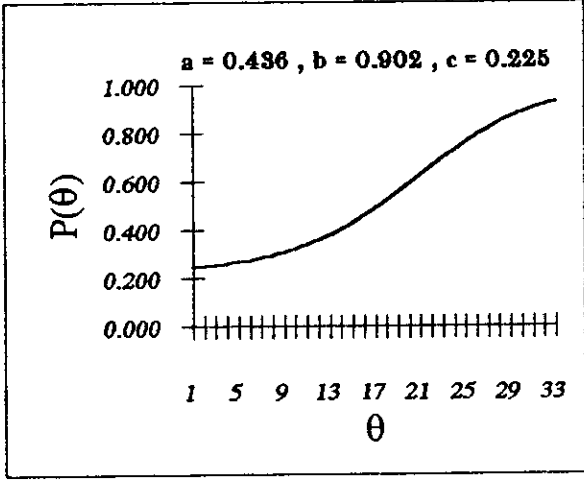


Fig 10

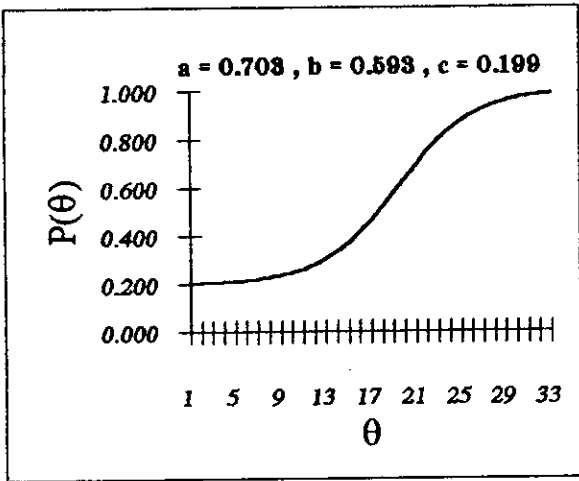


Fig 11

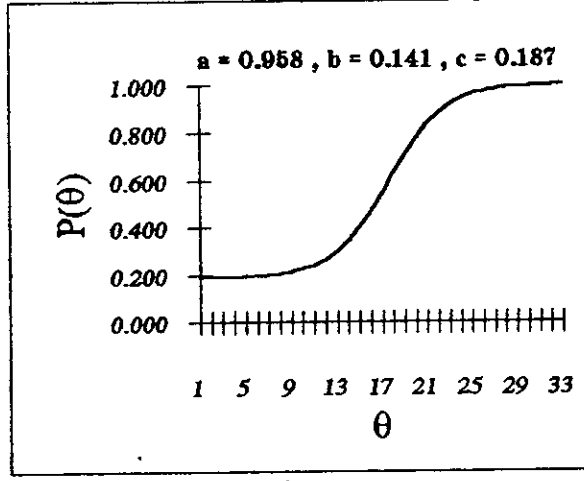
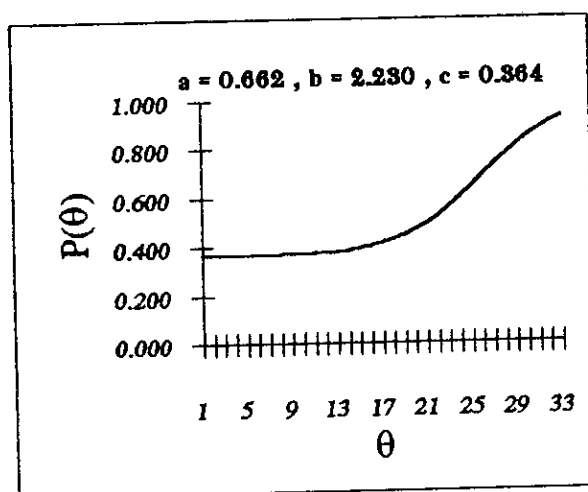


Fig 12



ข้อ 13

ภาพประกอบ 82 โค้งลักษณะข้อสอบของข้อสอบ ชนิด AX แบบเลือกตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 82 พบว่า แบบทดสอบชนิด AX แบบเลือกตอบของนักเรียนชั้นมัธยมศึกษาปีที่ 5 ข้อสอบข้อ 1 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 12 มีค่าความยากต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 3 และข้อ 10 มีค่าอำนาจจำแนกต่ำสุด สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 13 และข้อ 4 มีค่าการเดาต่ำสุด

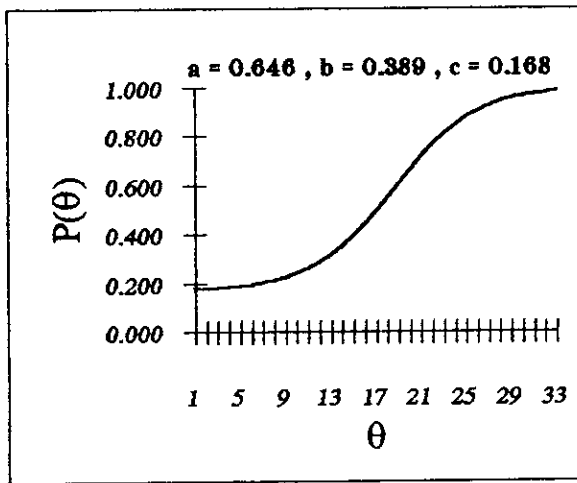


Fig 1

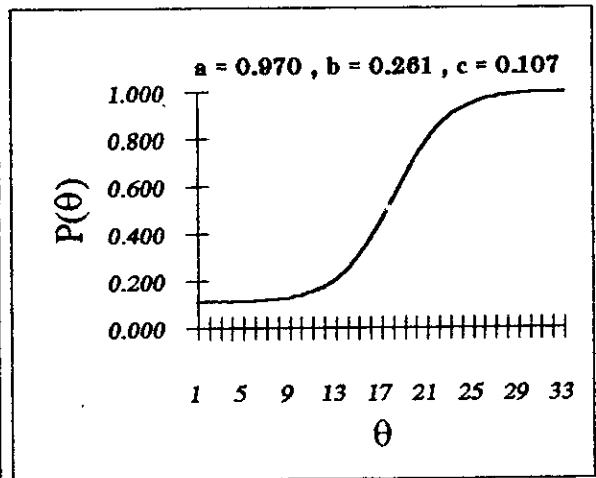


Fig 2

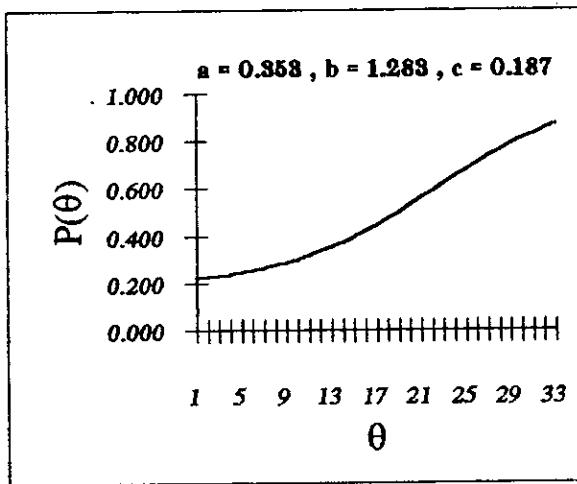


Fig 3

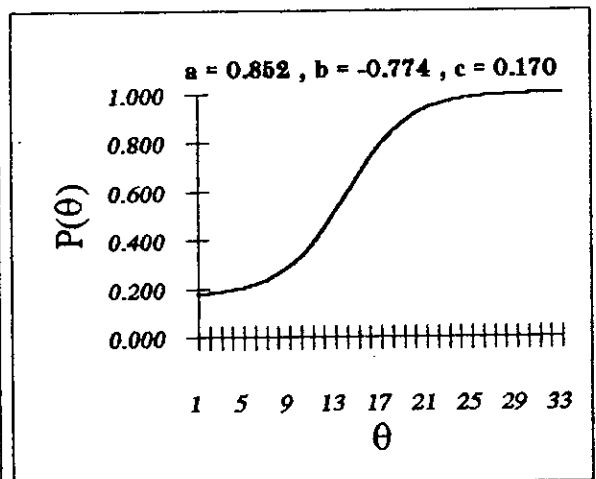


Fig 4

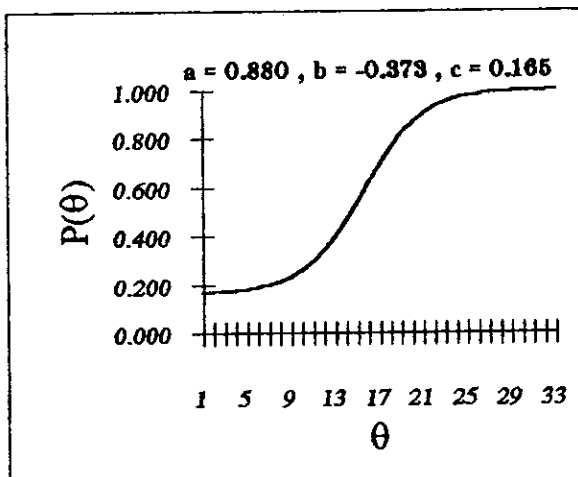


Fig 5

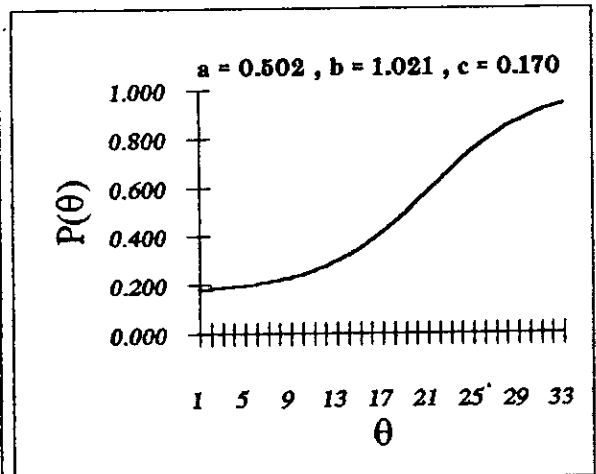
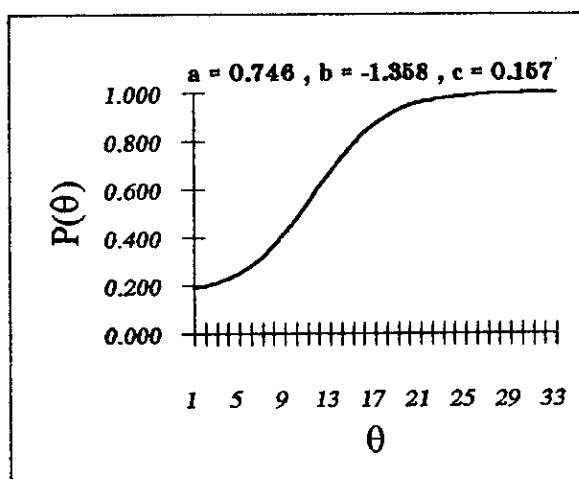
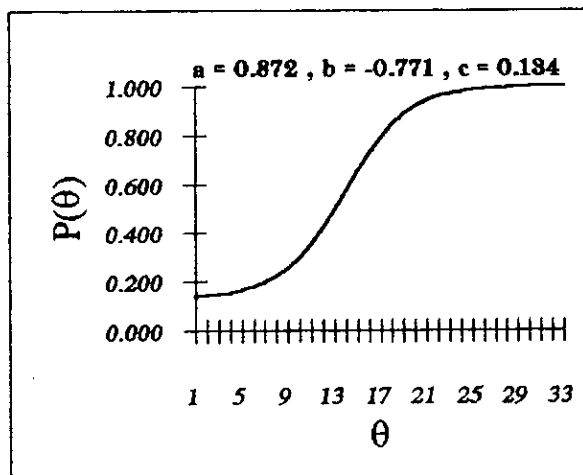


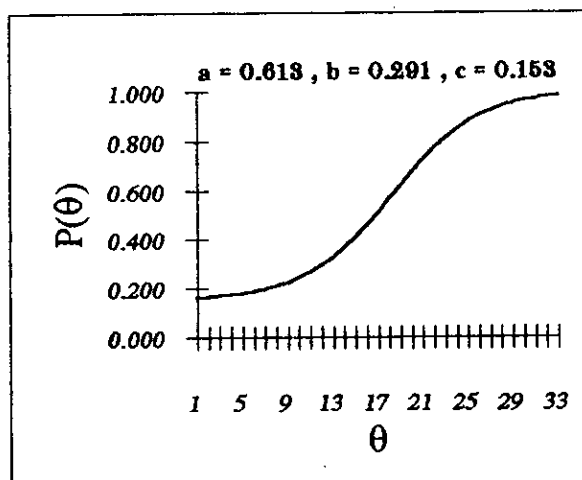
Fig 6



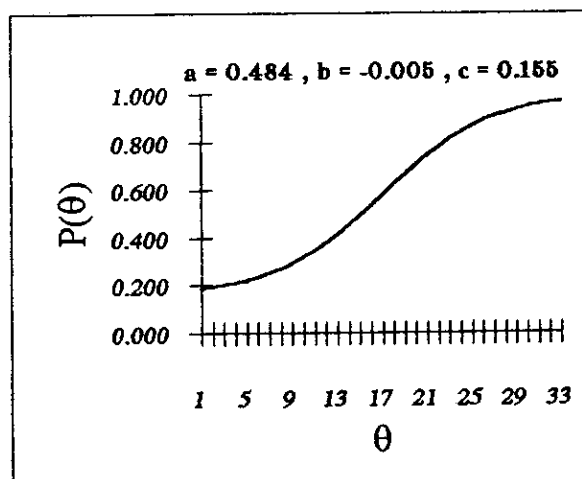
ข้อ 7



ข้อ 8



ข้อ 9



ข้อ 10

ภาพประกอบ 83 เติ้ดลักษณะข้อสอบของข้อสอบ ชนิด CV แบบเลือกตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 83 พบว่า แบบทดสอบชนิด CV แบบเลือกตอบของนักเรียน ชั้น
มัธยมศึกษาปีที่ 5 ข้อสอบข้อ 3 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 7 มีค่าความยาก
ต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 2 และข้อ 3 มีค่าอำนาจจำแนกต่ำสุด
สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 3 และข้อ 2 มีค่าการเดาต่ำสุด

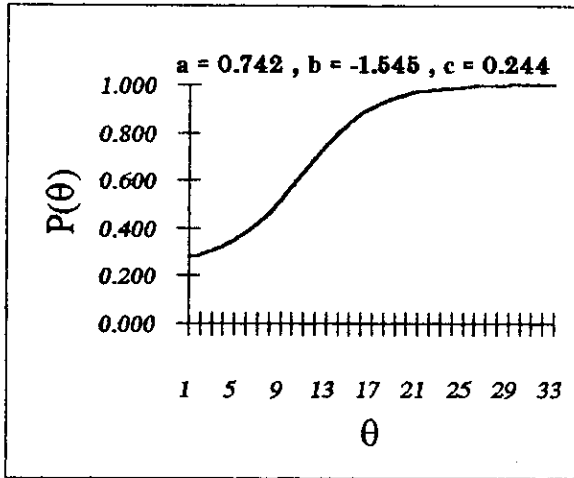


Fig 1

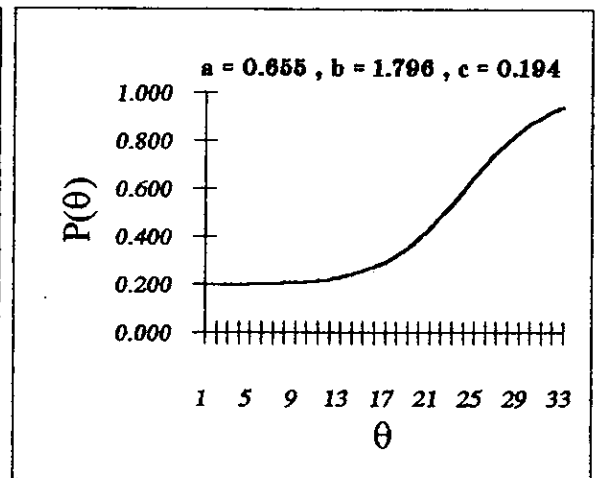


Fig 2

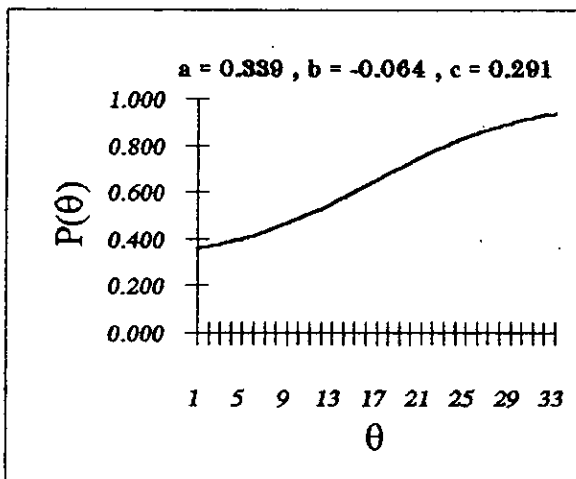


Fig 3

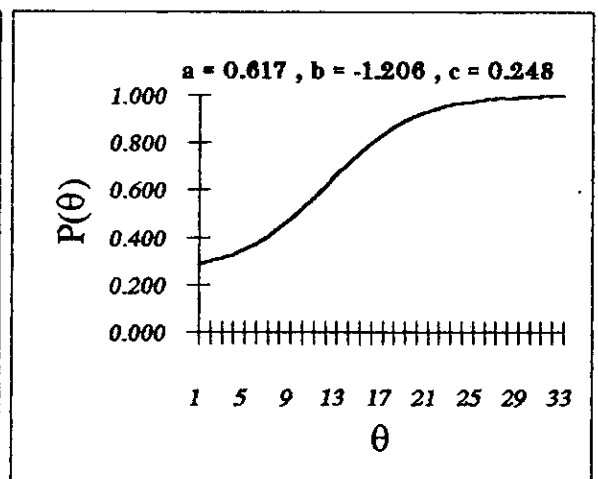


Fig 4

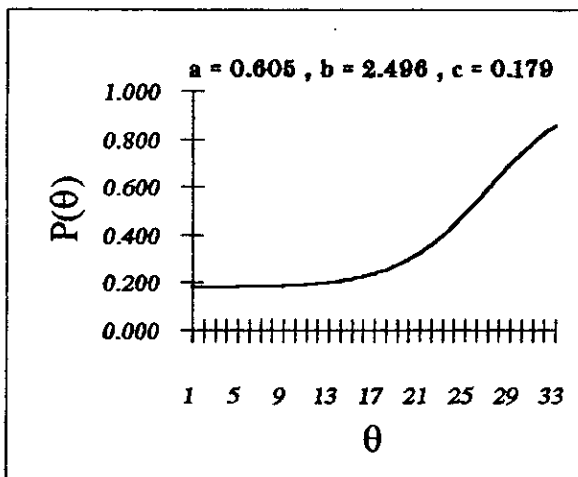


Fig 5

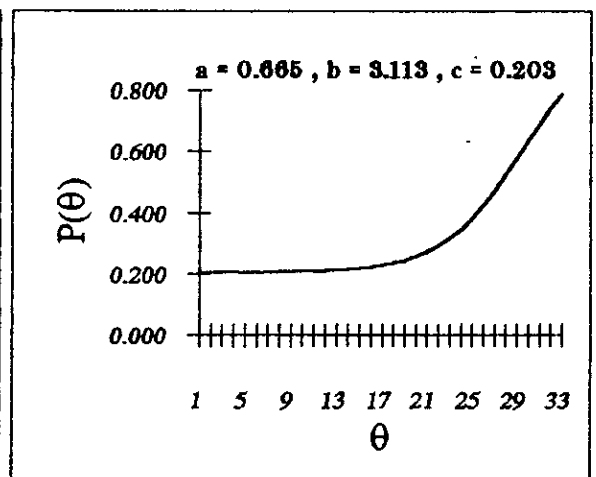
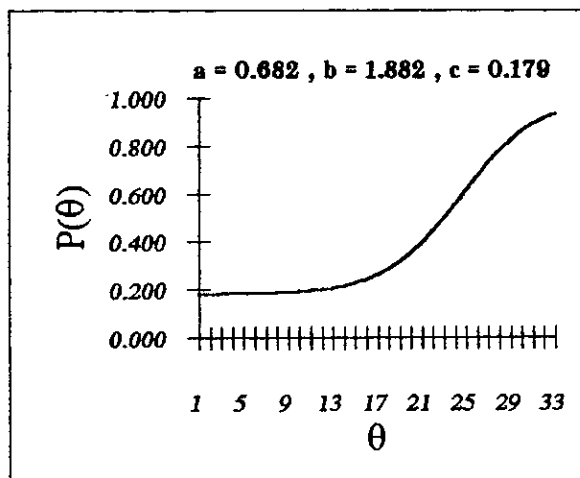
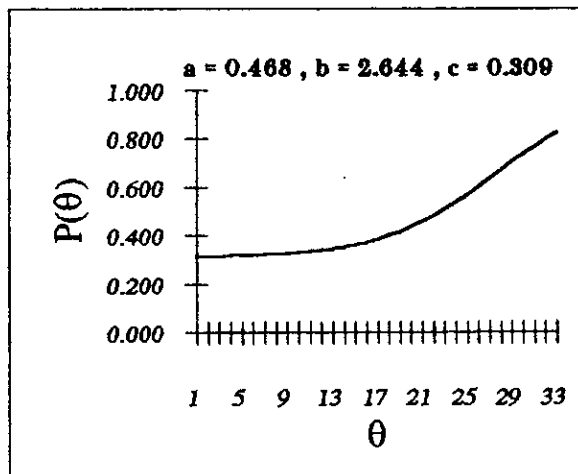


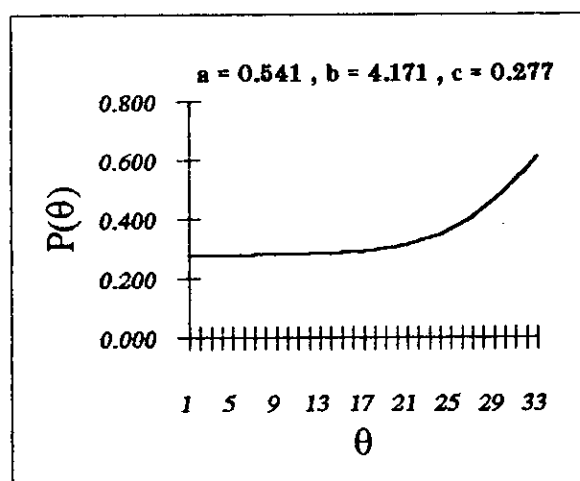
Fig 6



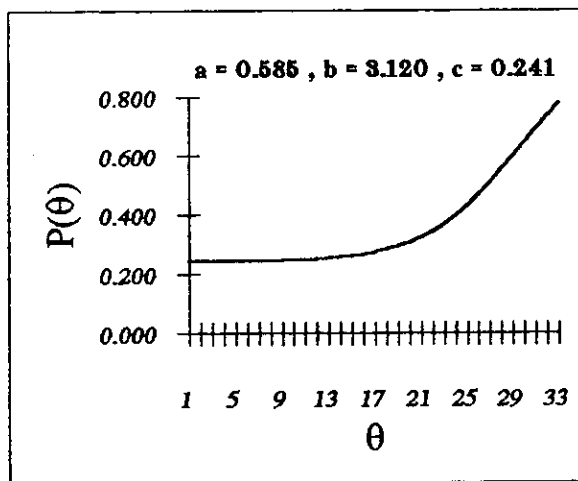
θ 7



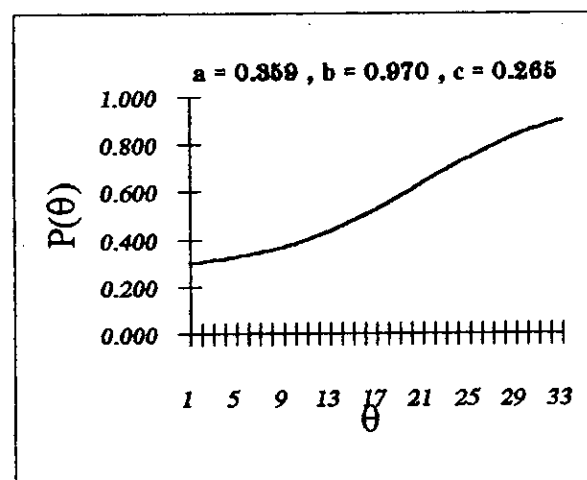
θ 8



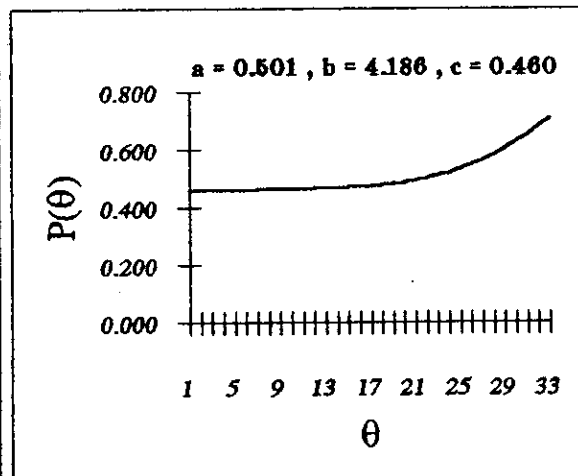
θ 9



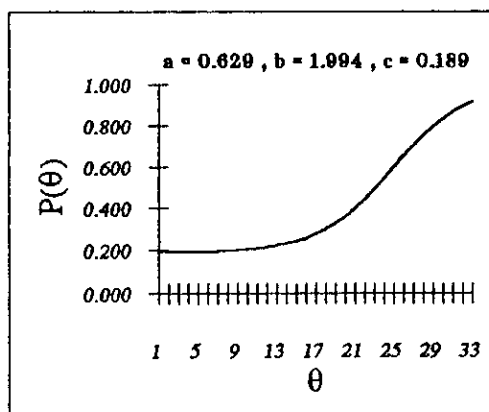
θ 10



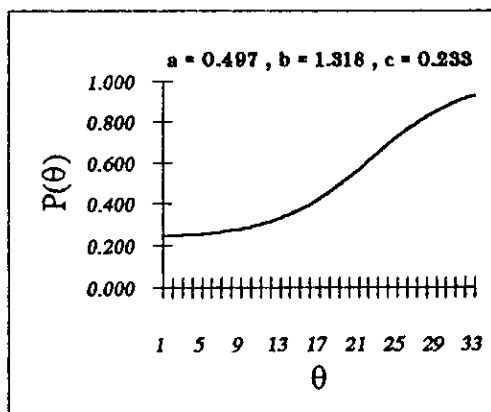
θ 11



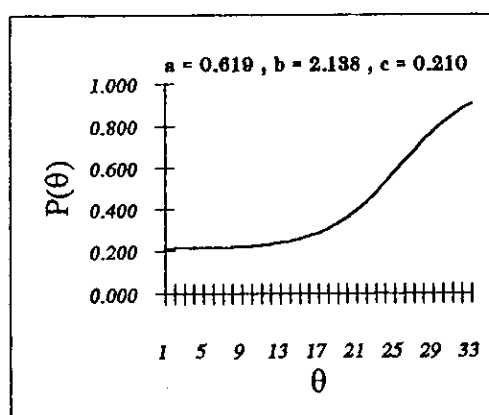
θ 12



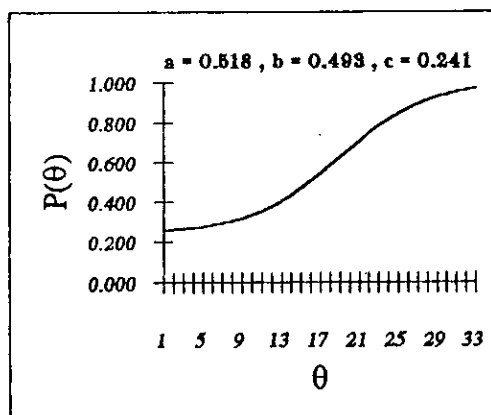
ข้อ 13



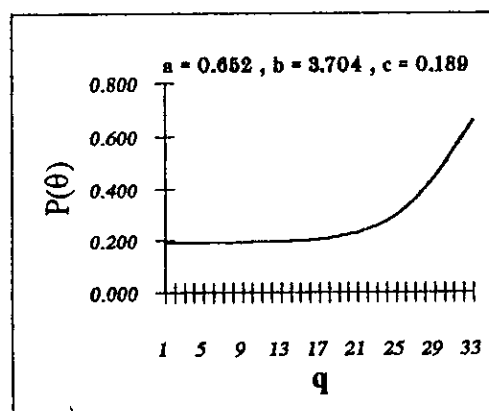
ข้อ 14



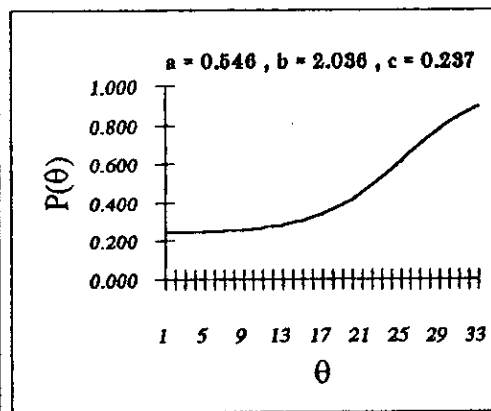
ข้อ 15



ข้อ 16



ข้อ 17



ข้อ 18

ภาพประกอบ 84 โต๊ะลักษณะข้อสอบของข้อสอบ ชนิด LR แบบเลือกตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 84 พบว่า แบบทดสอบชนิด LR แบบเลือกตอบของนักเรียน ชั้นมัธยมศึกษาปีที่ 5 ข้อสอบข้อ 12 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 1 มีค่าความยากต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 1 และข้อ 3 มีค่าอำนาจจำแนกต่ำสุด สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 12 ส่วนข้อ 5 และข้อ 7 มีค่าการเดาต่ำสุด

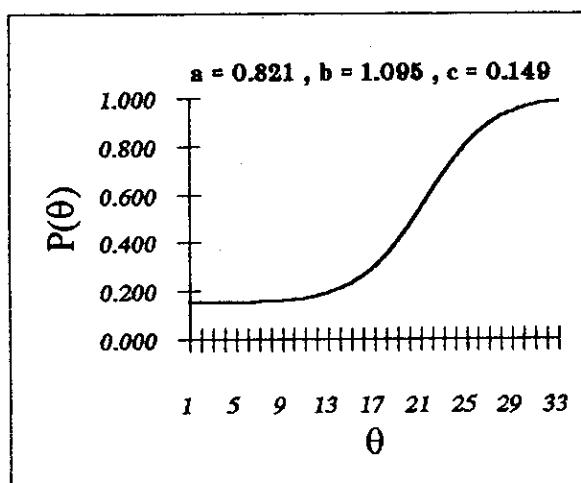


Fig 1

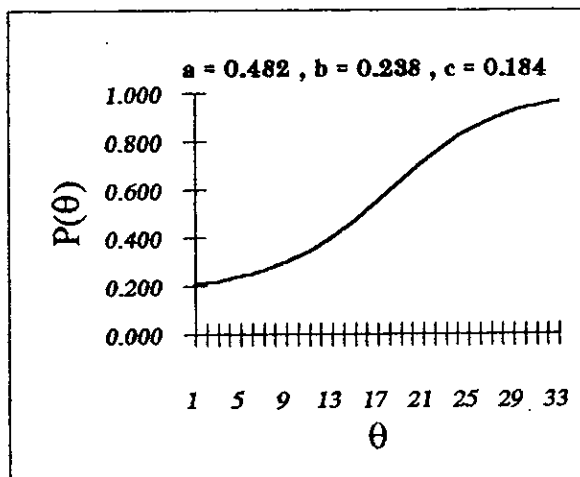


Fig 2

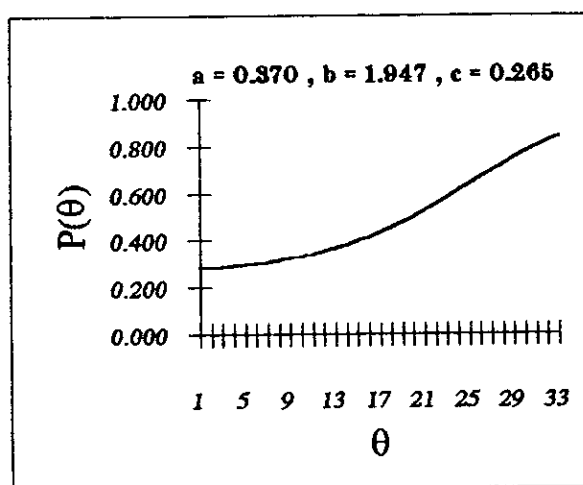


Fig 3

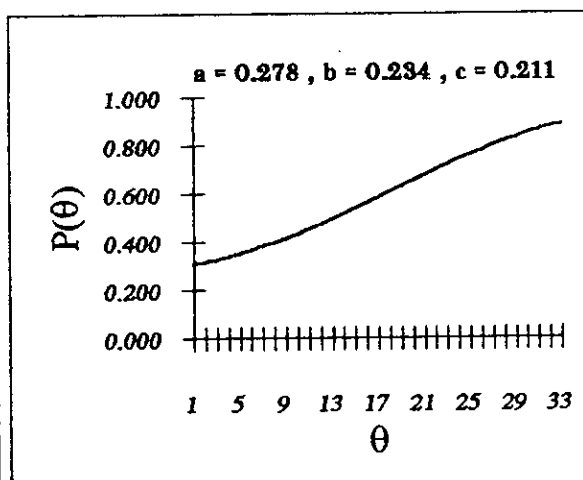


Fig 4

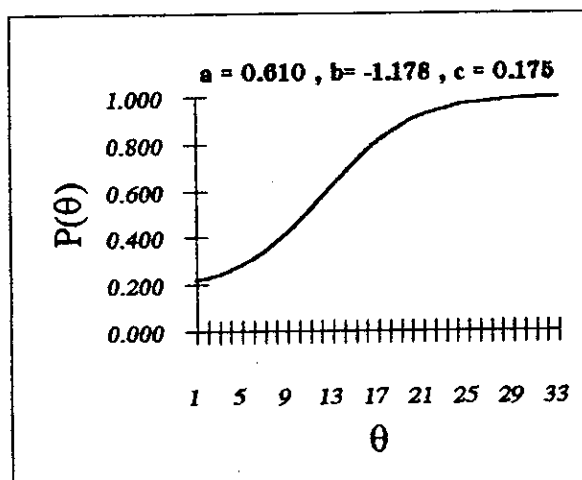


Fig 5

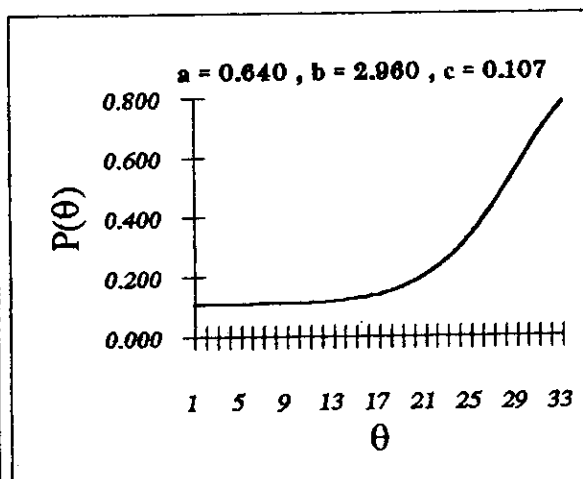
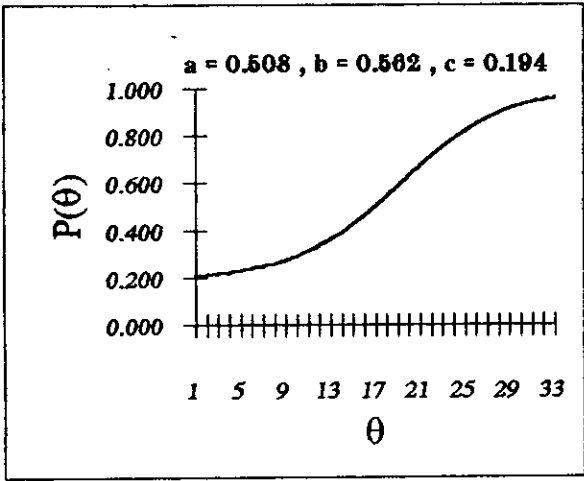
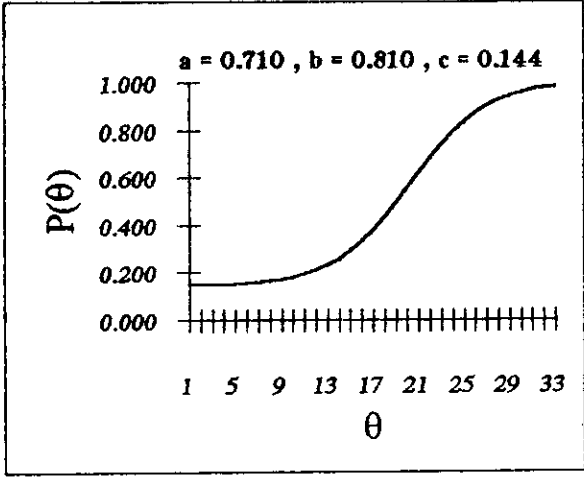


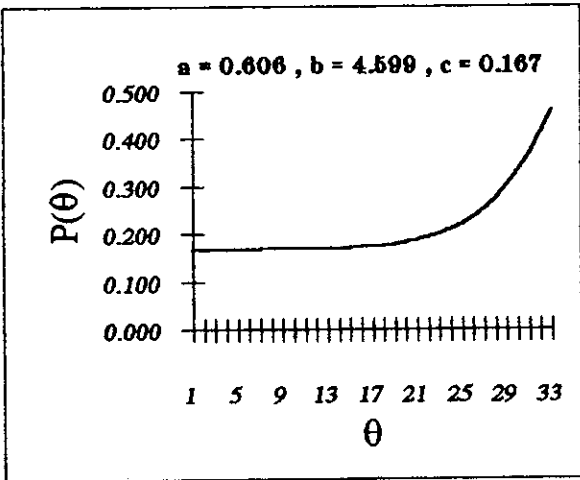
Fig 6



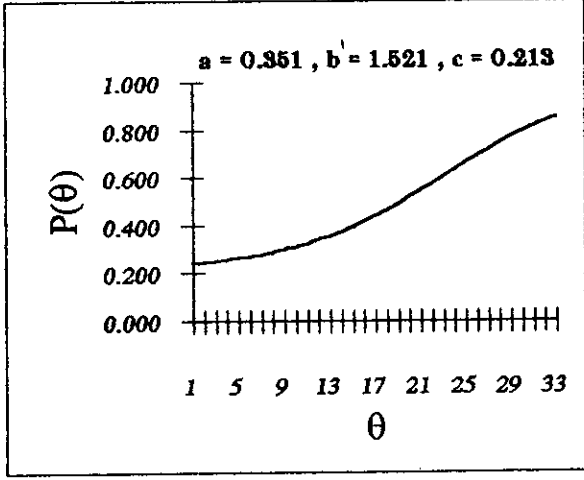
θ 7



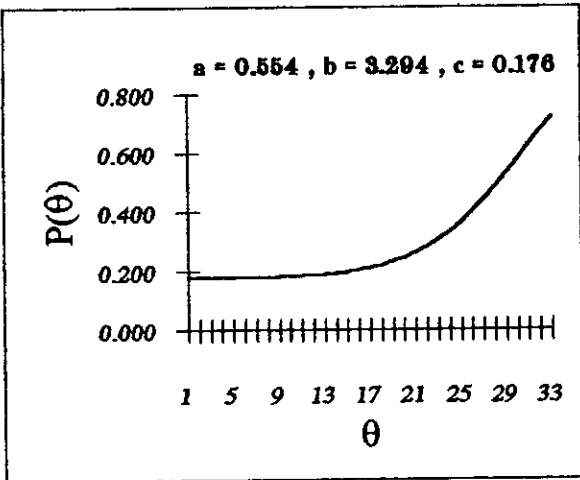
θ 8



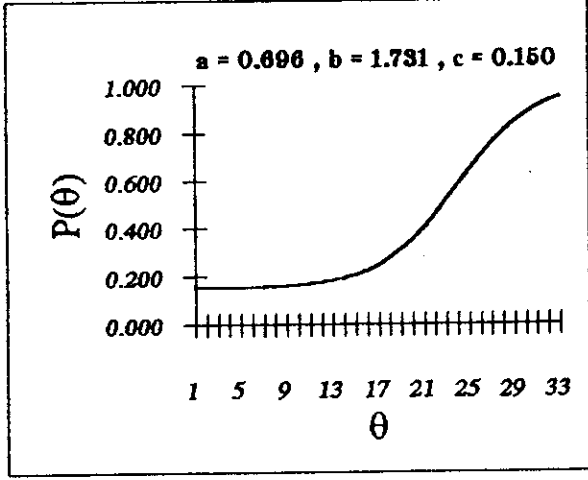
θ 9



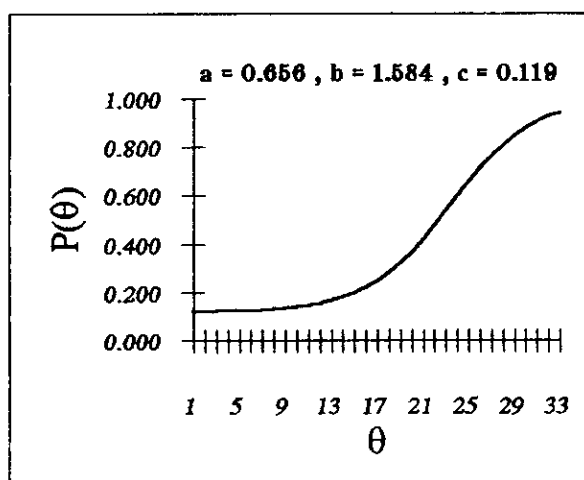
θ 10



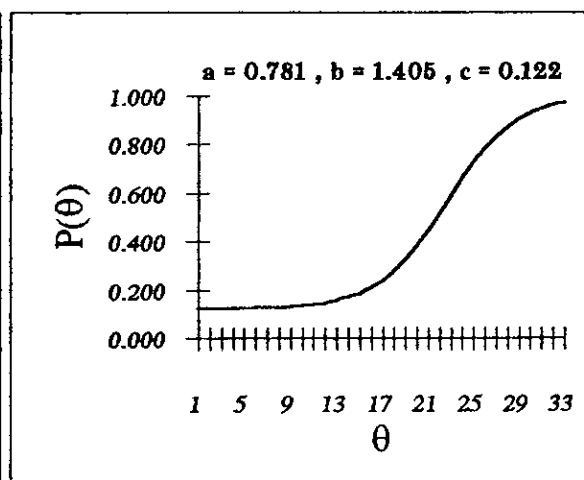
θ 11



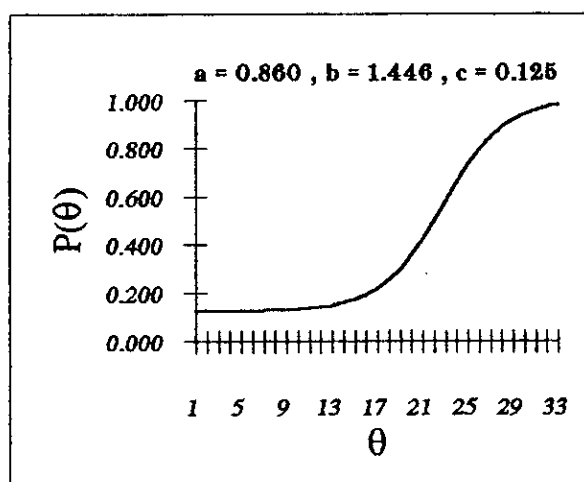
θ 12



ข้อ 13



ข้อ 14



ข้อ 15

ภาพประกอบ 85 ได้ดังลักษณะข้อสอบของข้อสอบ ชนิด NLR แบบเลือกตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 85 พบว่า แบบทดสอบชนิด NLR แบบเลือกตอบของนักเรียน ชั้น
มัธยมศึกษาปีที่ 5 ข้อสอบข้อ 9 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 5 มีค่าความยาก
ต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 15 และข้อ 4 มีค่าอำนาจจำแนกต่ำสุด
สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 3 และข้อ 6 มีค่าการเดาต่ำสุด

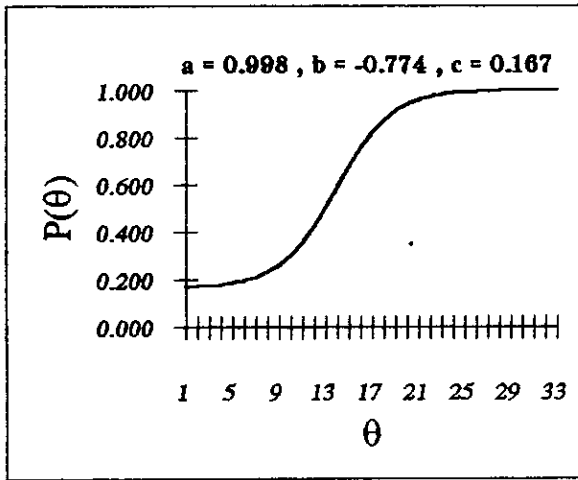


Fig 1

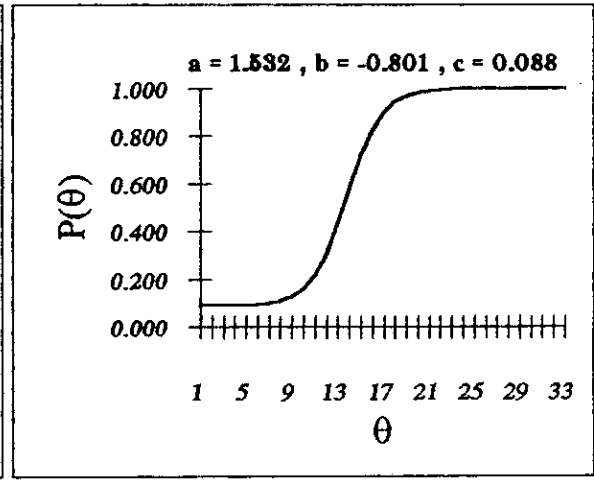


Fig 2

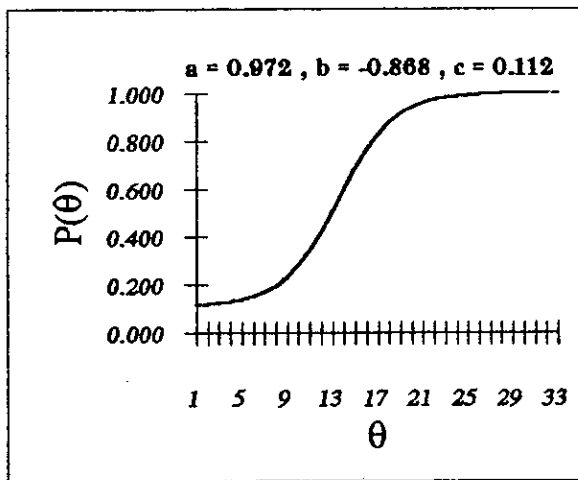


Fig 3

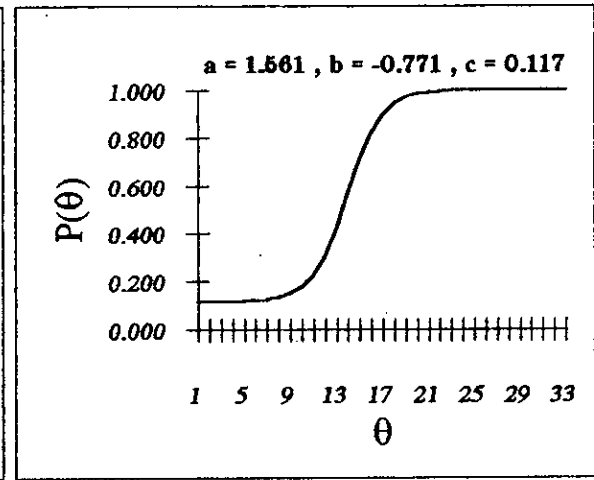


Fig 4

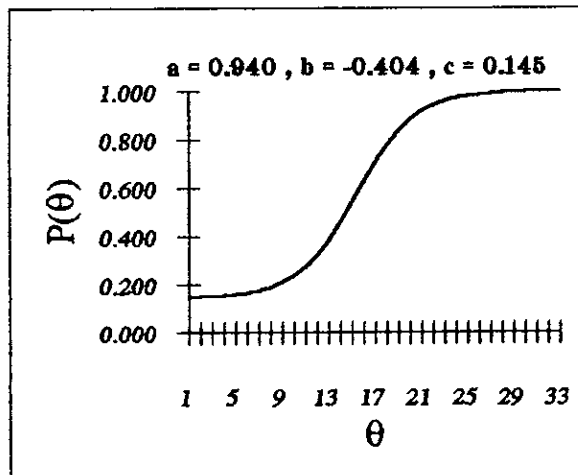


Fig 5

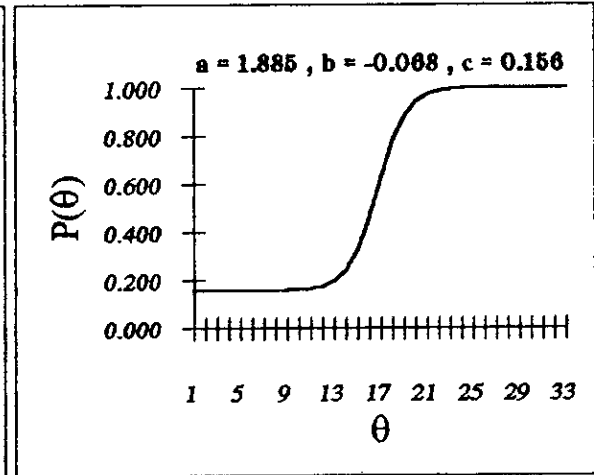
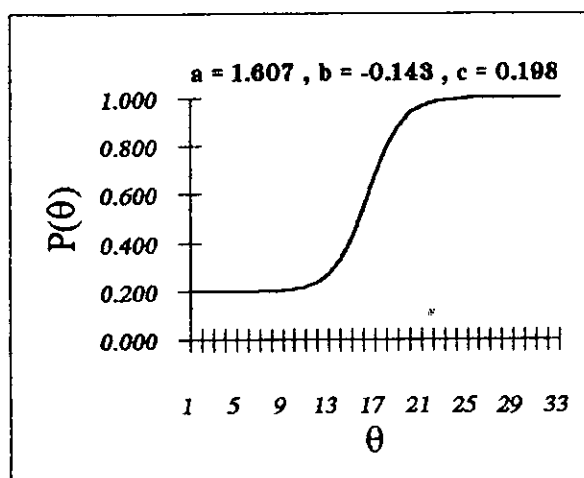
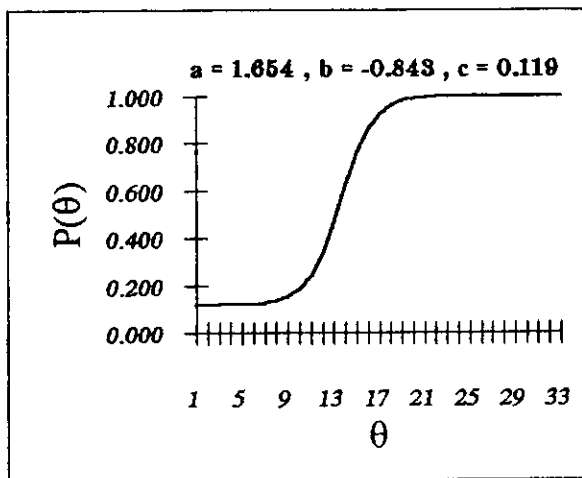


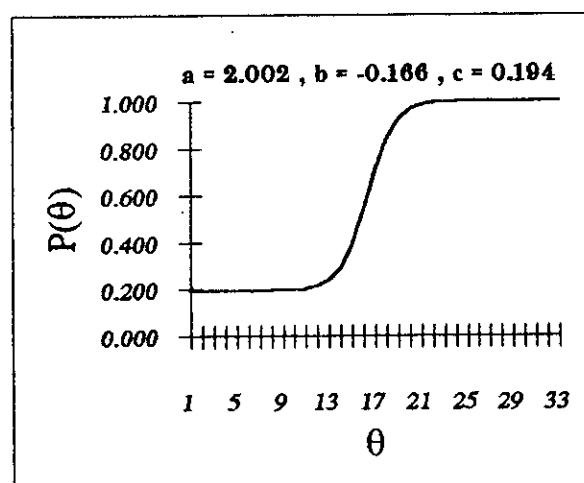
Fig 6



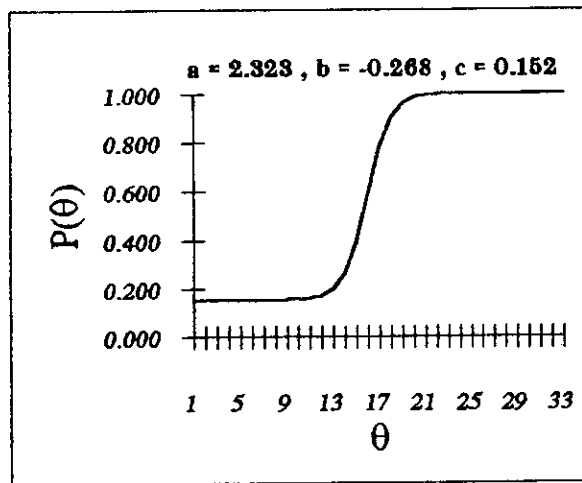
θ 7



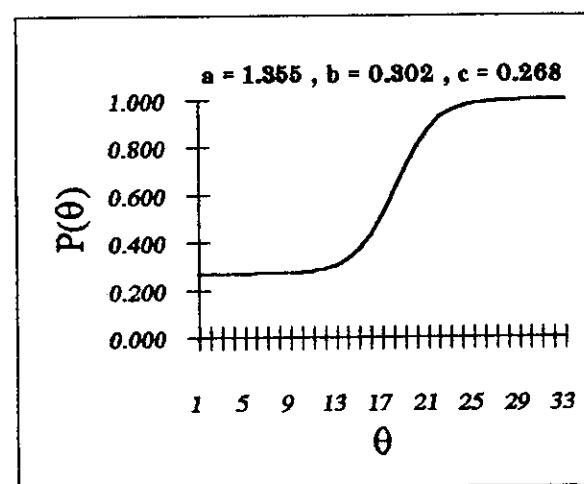
θ 8



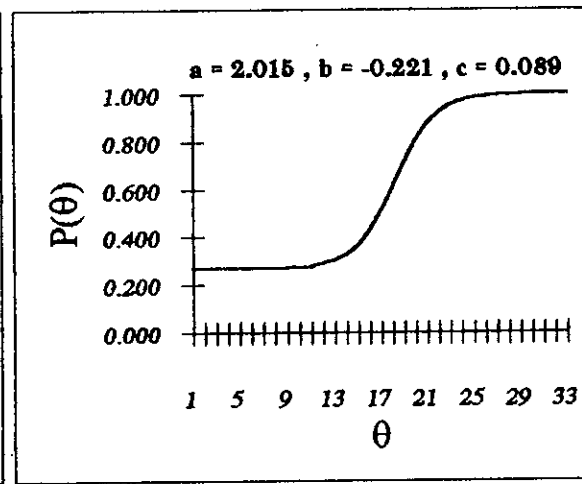
θ 9



θ 10



θ 11



θ 12

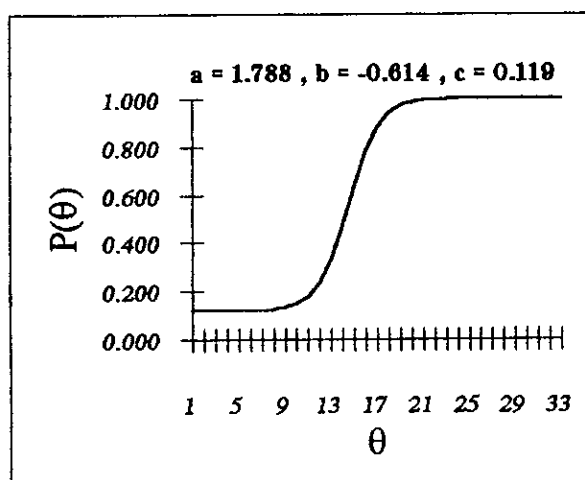


Figure 13

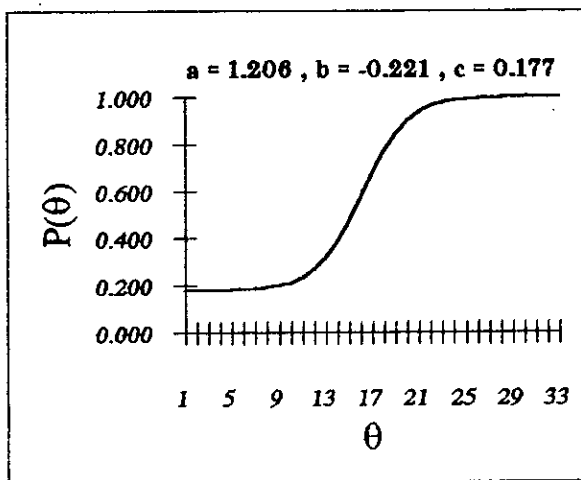


Figure 14

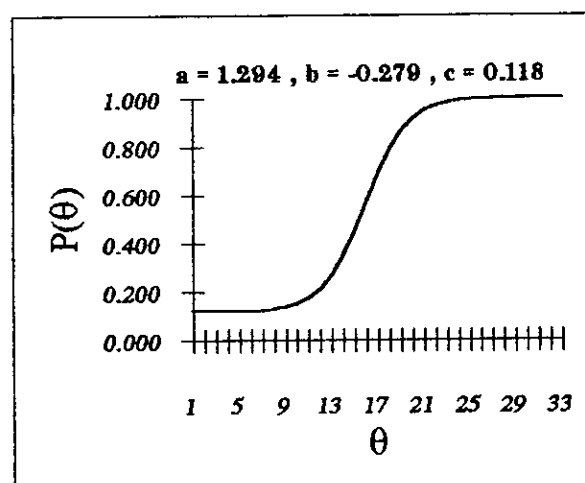


Figure 15

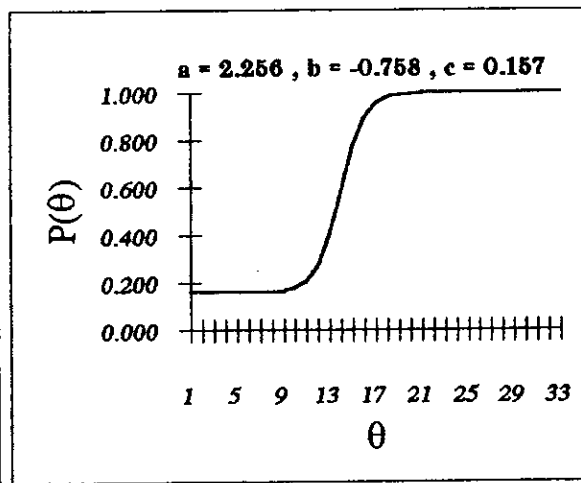


Figure 16

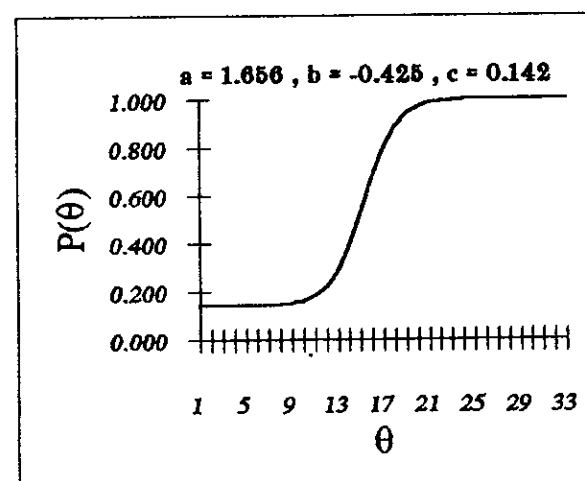


Figure 17

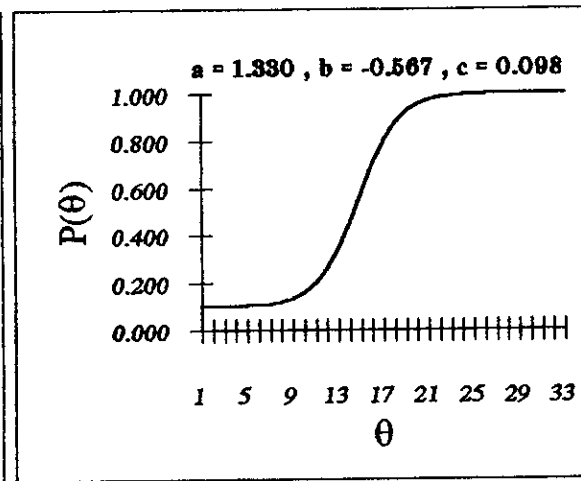
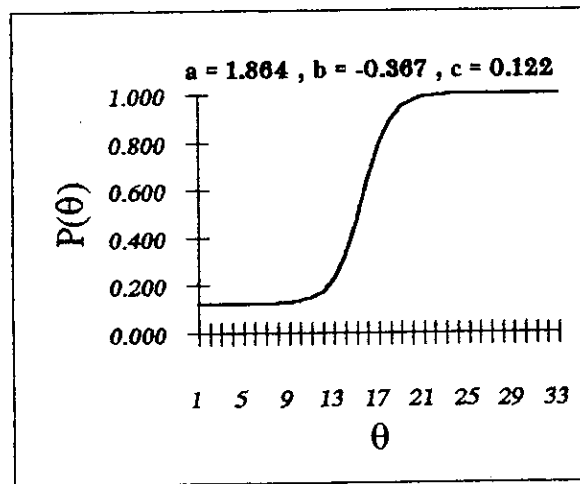


Figure 18



ข้อ 19

ภาพประกอบ 86 โค้งลักษณะข้อสอบของข้อสอบ ชนิด PI แบบเลือกตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 86 พบว่า แบบทดสอบชนิด PI แบบเลือกตอบของนักเรียน ชั้น
มัธยมศึกษาปีที่ 5 ข้อสอบข้อ 11 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 3 ค่าความยาก
ต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 10 และข้อ 5 มีค่าอำนาจจำแนกต่ำสุด
สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 11 และข้อ 2 มีค่าการเดาต่ำสุด

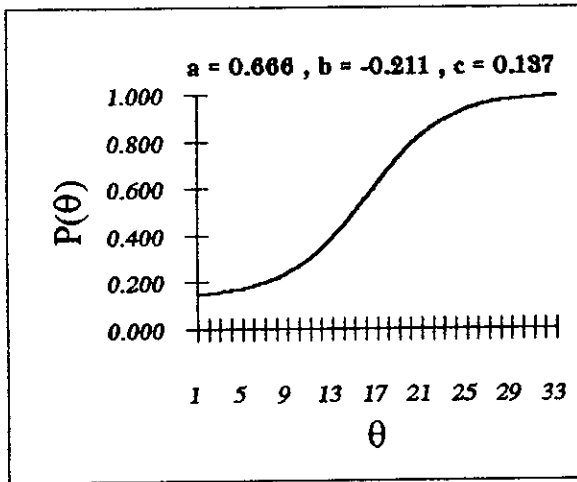


Fig 1

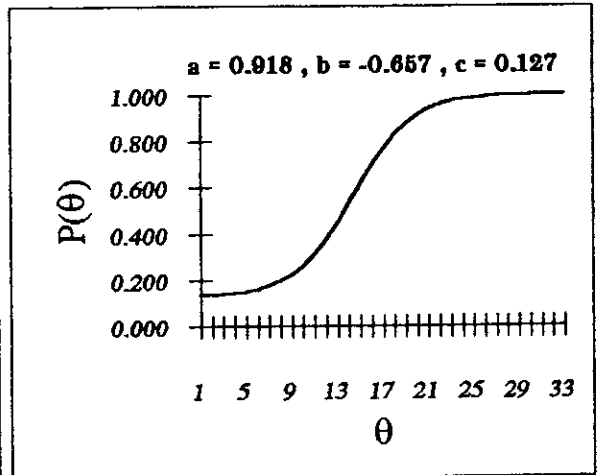


Fig 2

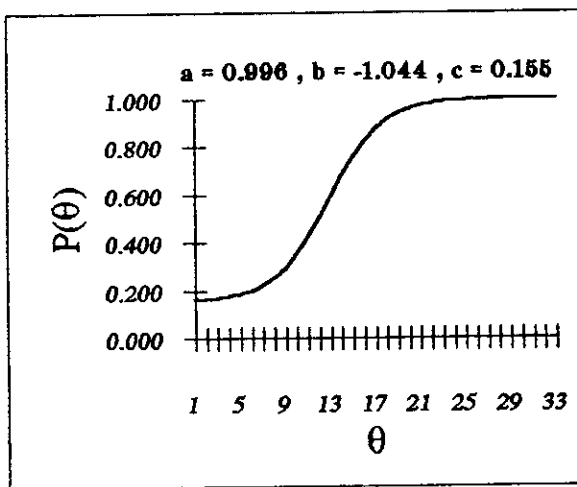


Fig 3

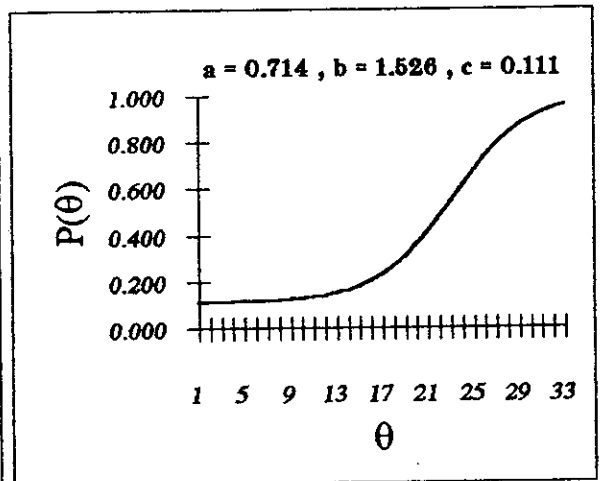


Fig 4

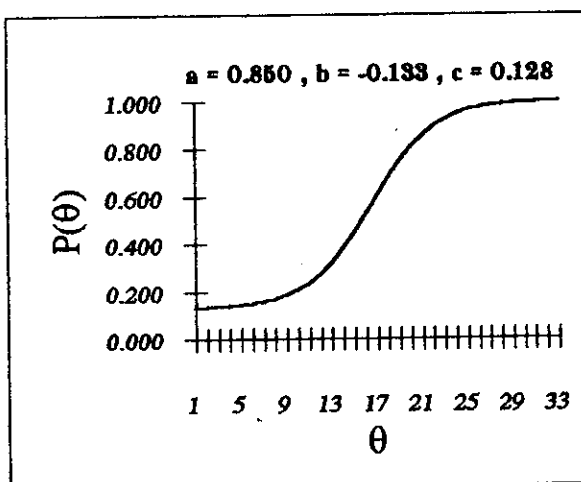


Fig 5

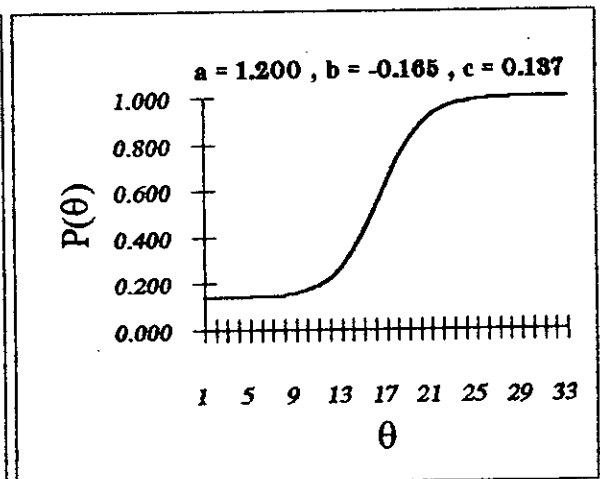


Fig 6

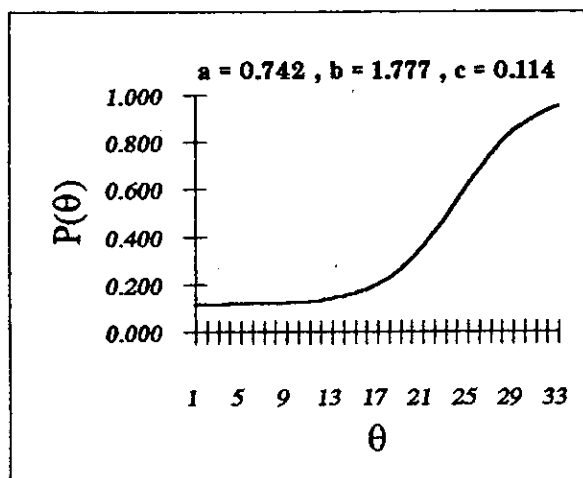


Fig 7

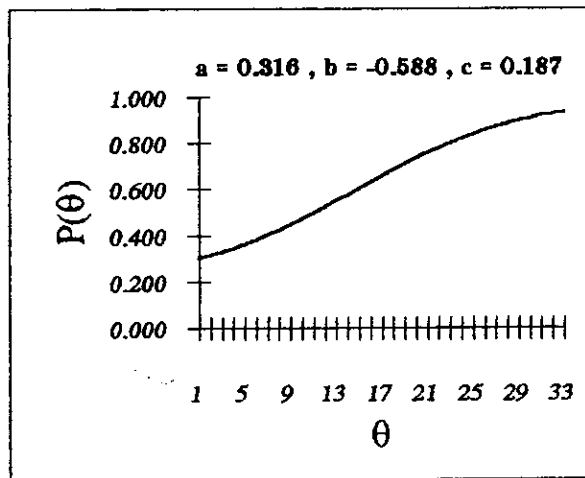


Fig 8

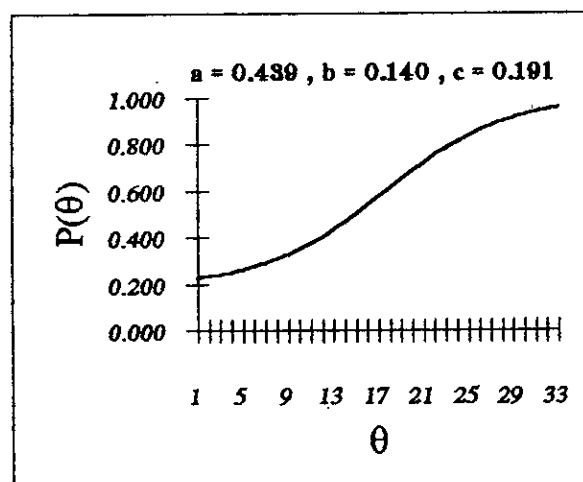


Fig 9

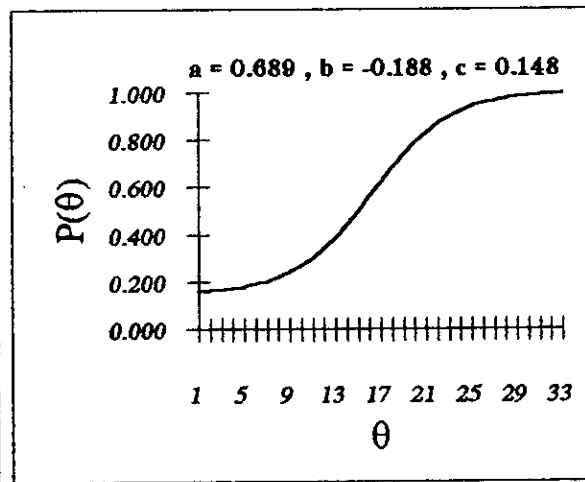


Fig 10

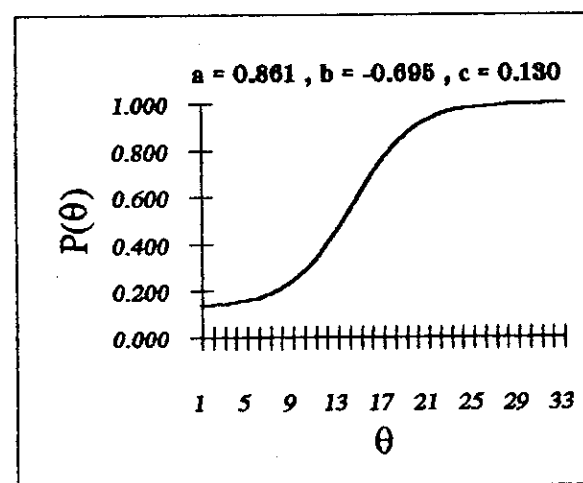


Fig 11

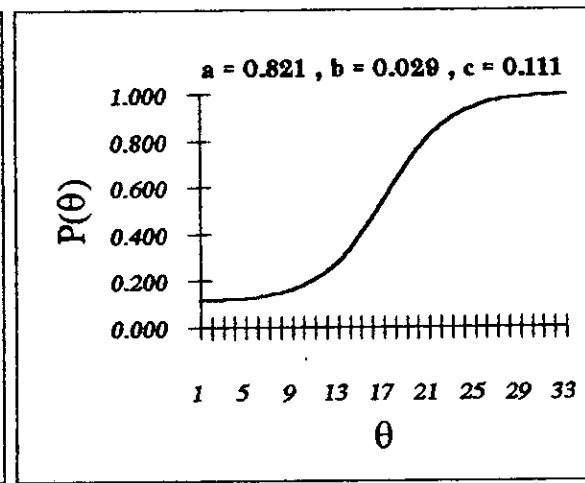


Fig 12

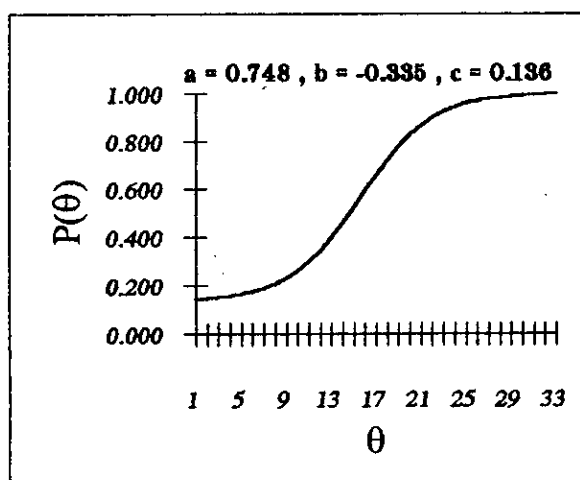


Figure 13

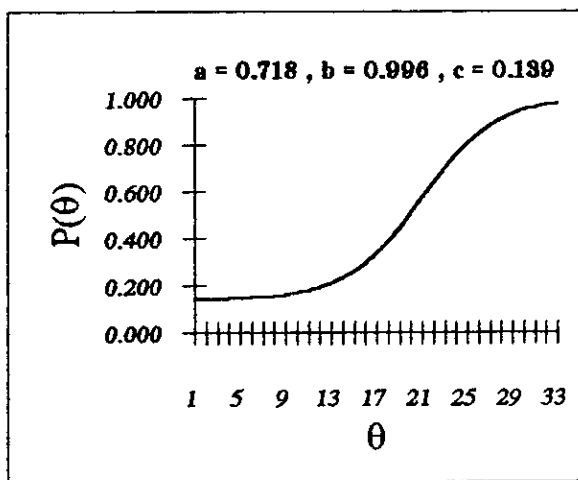


Figure 14

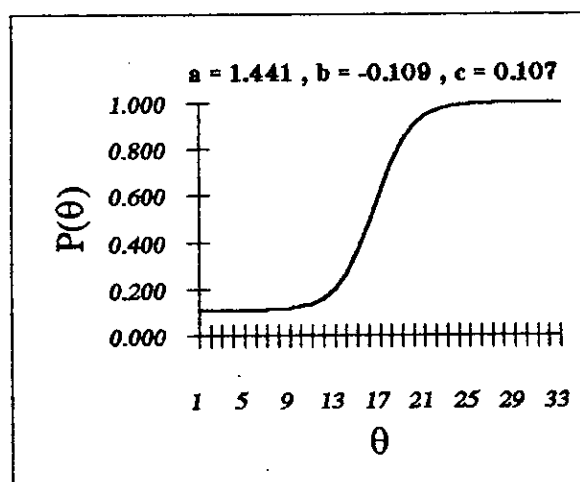


Figure 15

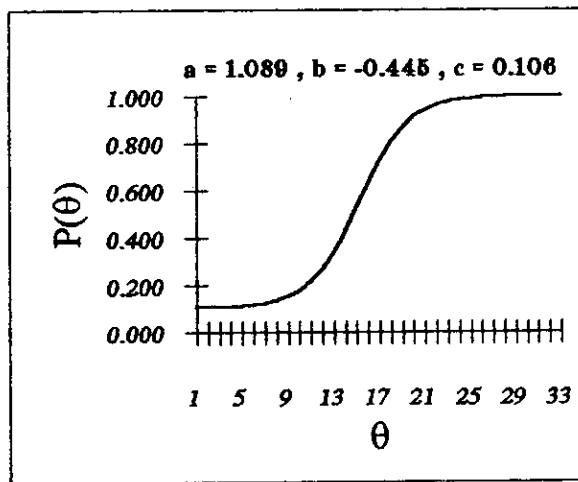


Figure 16

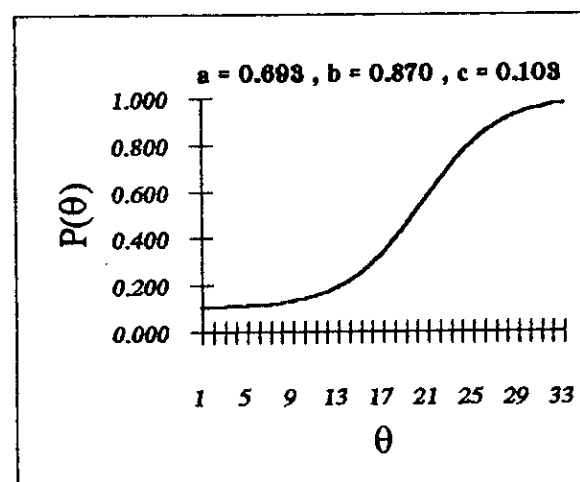


Figure 17

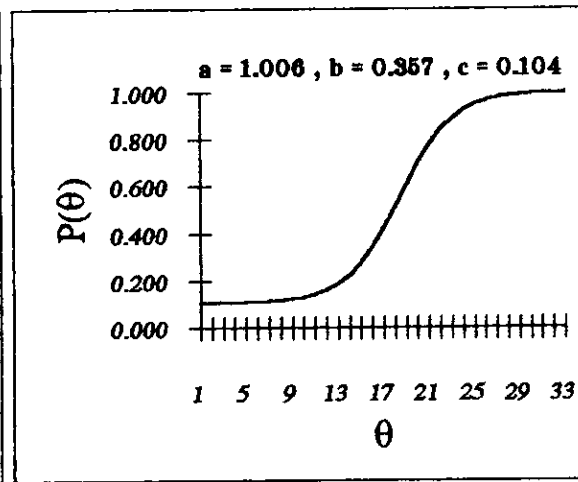
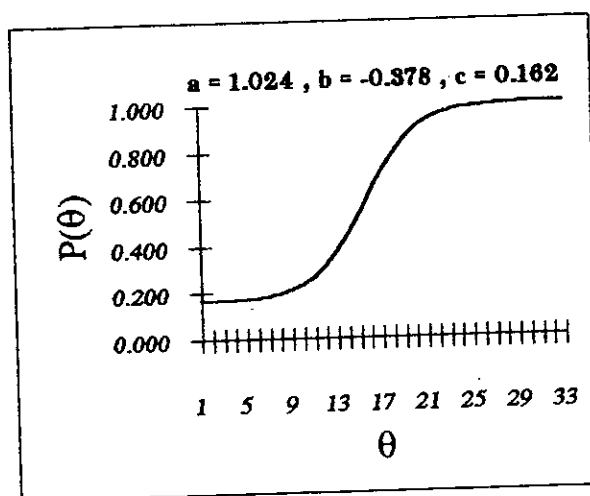


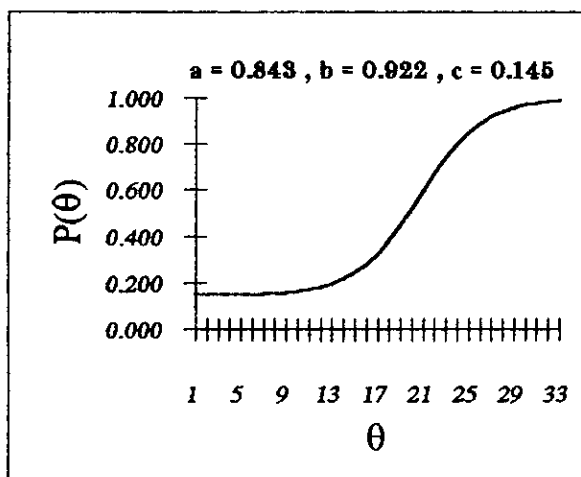
Figure 18



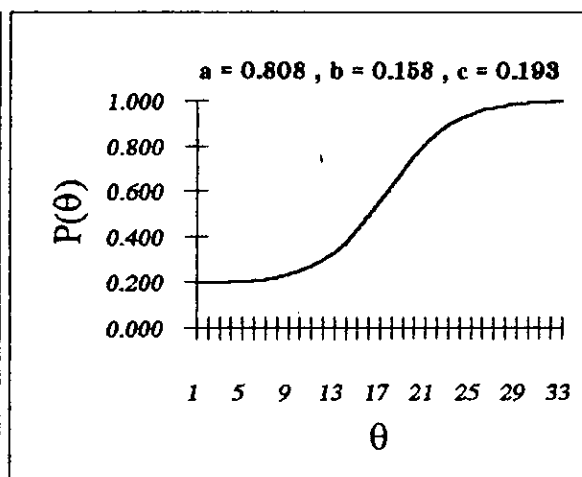
ข้อ 19.

ภาพประกอบ 87 โค้งลักษณะข้อสอบของข้อสอบ ชนิด AR แบบเติมคำตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5

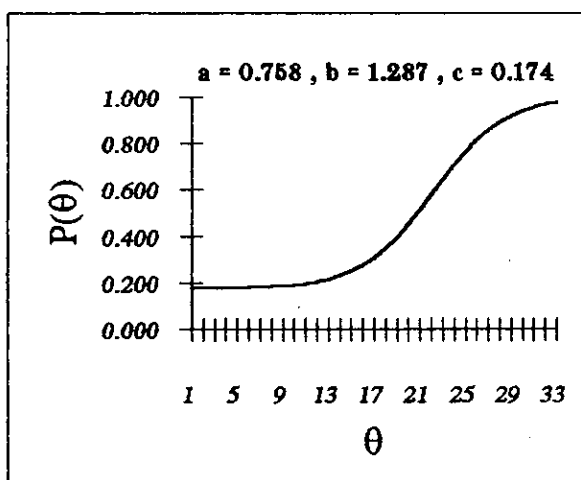
จากภาพประกอบ 87 พบว่า แบบทดสอบชนิด AR แบบเติมคำตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ข้อสอบข้อ 7 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 3 มีค่าความ
ยากต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 15 และข้อ 8 มีค่าอำนาจจำแนกต่ำ
สุด สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 9 และข้อ 17 มีค่าการเดาต่ำสุด



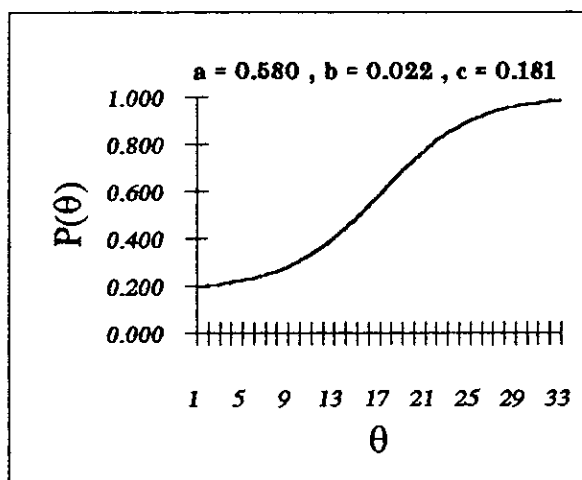
θ 1



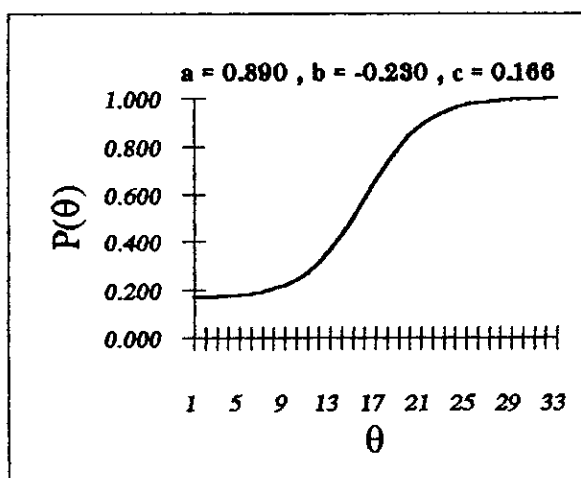
θ 2



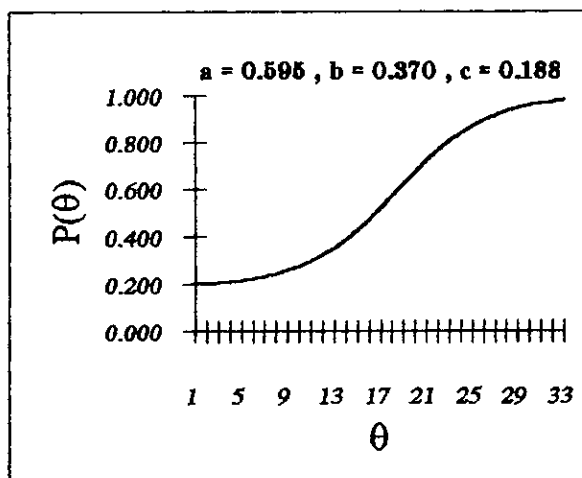
θ 3



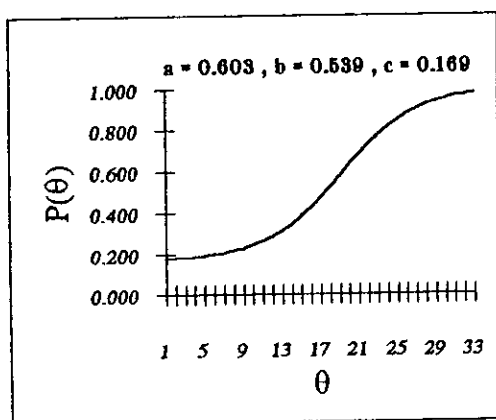
θ 4



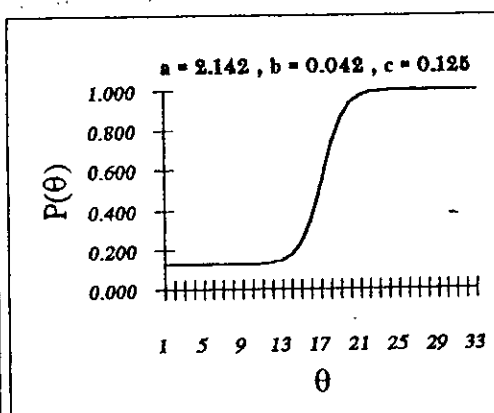
θ 5



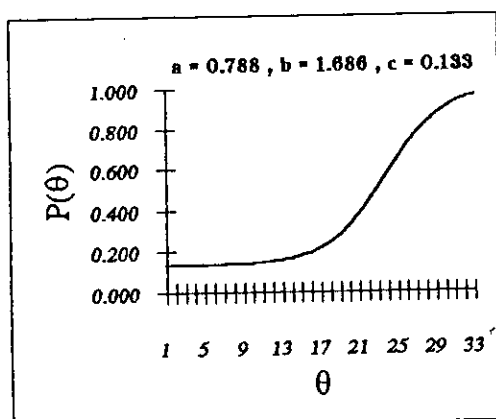
θ 6



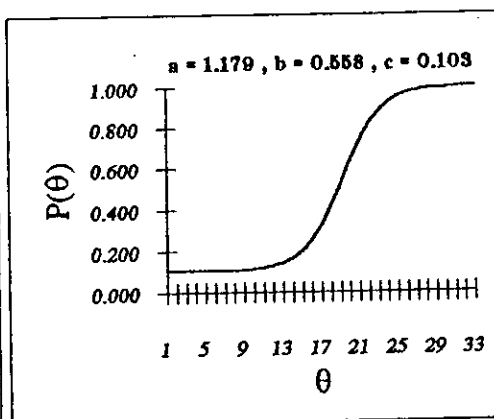
ข้อ 7



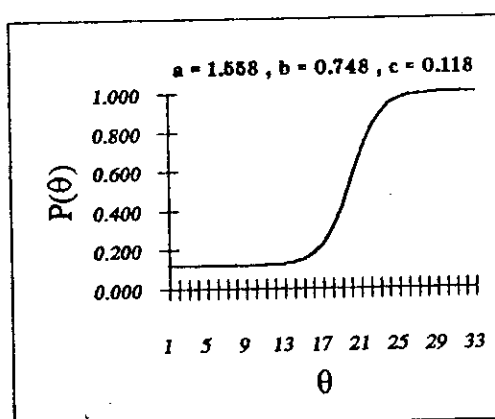
ข้อ 8



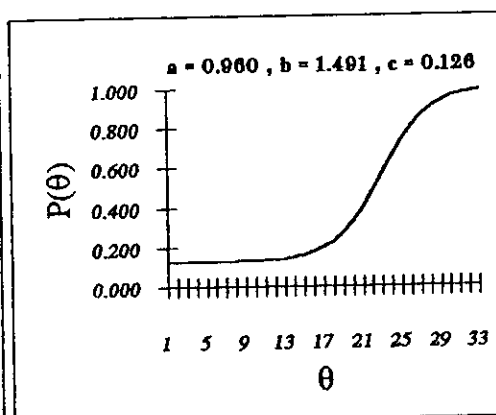
ข้อ 9



ข้อ 10



ข้อ 11



ข้อ 12

ภาพประกอบ 88 ได้ังลักษณะข้อสอบของข้อสอบ ชนิด AX แบบเดิมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 88 พบว่า แบบทดสอบชนิด AX แบบเติมคำตอบของนักเรียน ชั้นมัธยมศึกษาปีที่ 5 ข้อสอบข้อ 9 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 5 มีค่าความยากต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 8 และข้อ 4 มีค่าอำนาจจำแนกต่ำสุด สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 2 และข้อ 10 มีค่าการเดาต่ำสุด

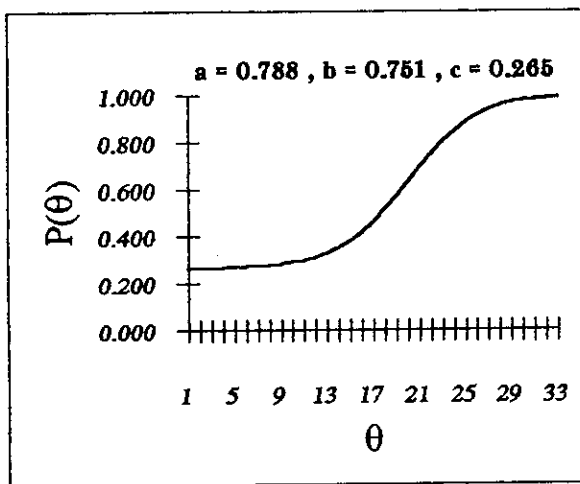


Fig 1

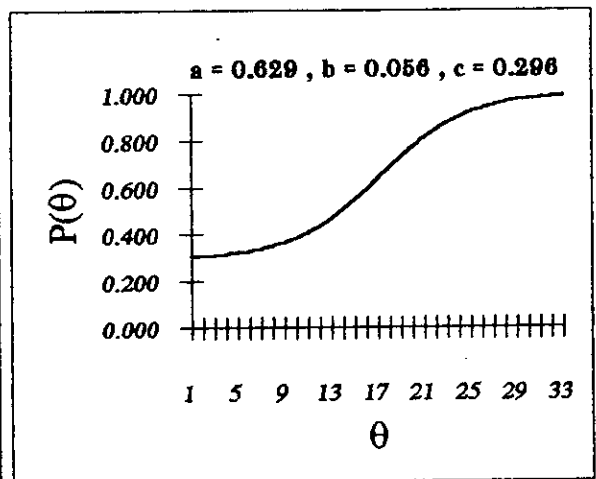


Fig 2

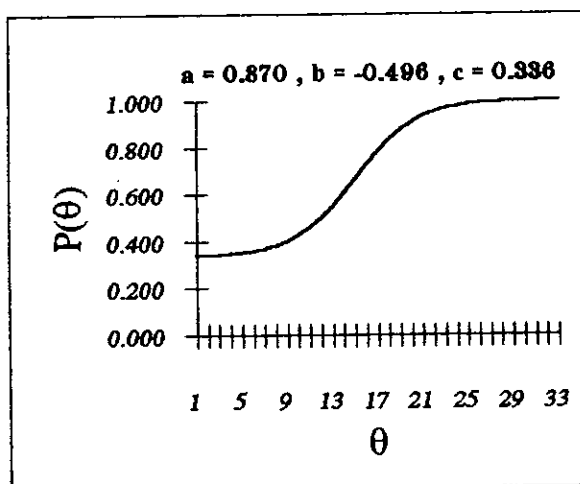


Fig 3

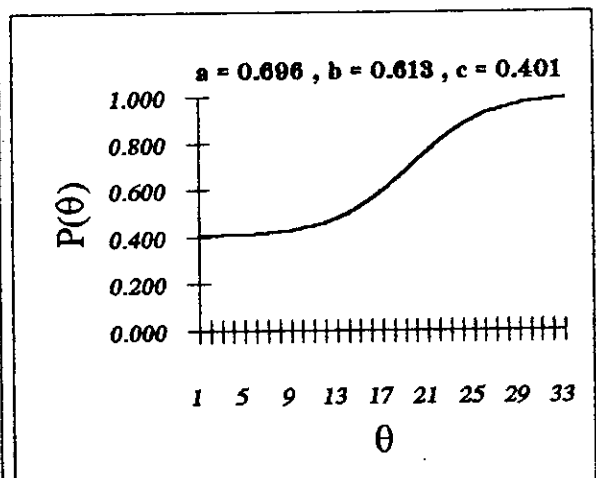


Fig 4

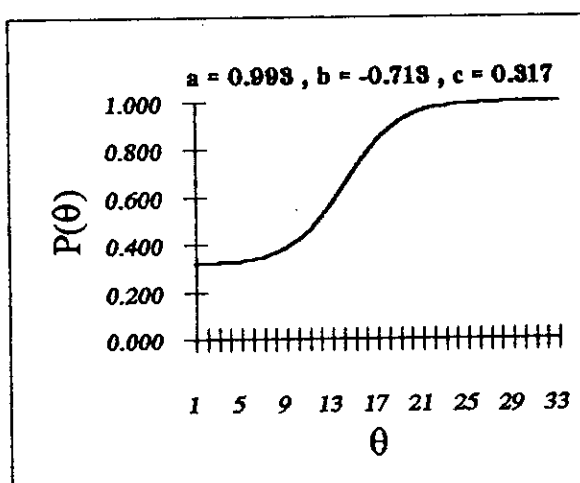


Fig 5

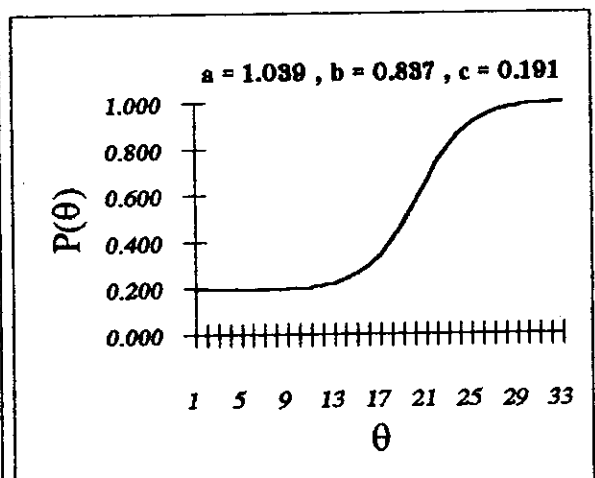
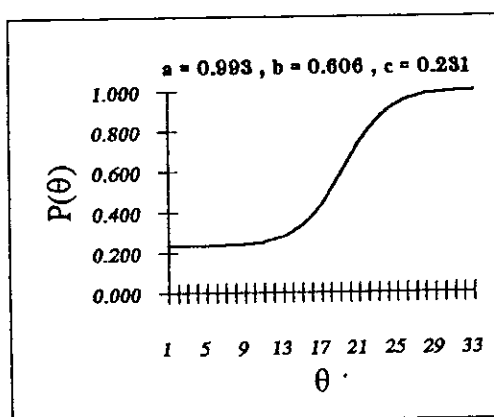
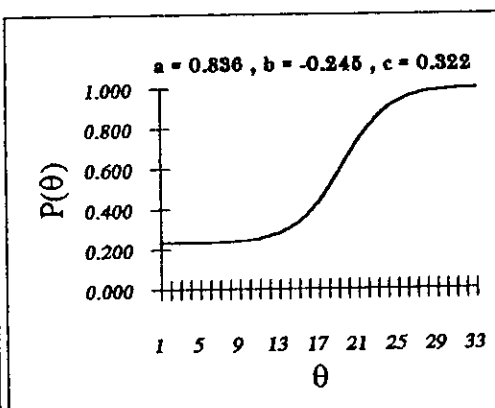


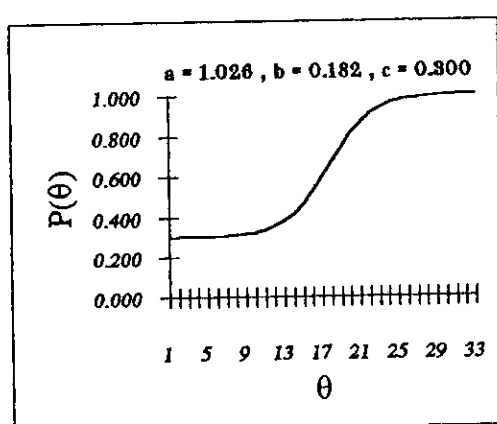
Fig 6



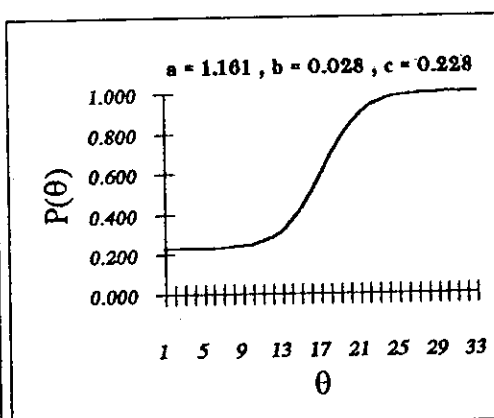
ข้อ 7



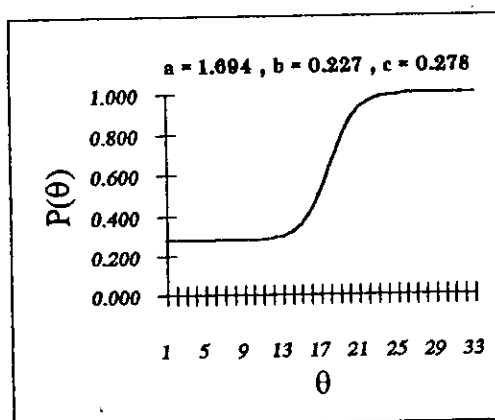
ข้อ 8



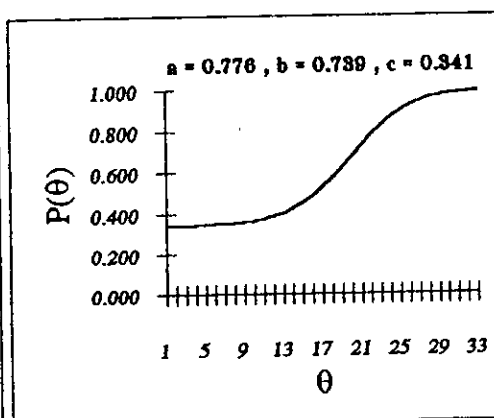
ข้อ 9



ข้อ 10



ข้อ 11



ข้อ 12

ภาพประกอบ 89 ได้ังลักษณะข้อสอบของข้อสอบ ชนิด CV แบบเดิมคำตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 89 พบว่า แบบทดสอบชนิด CV แบบเติมคำตอบของนักเรียน ชั้นมัธยมศึกษาปีที่ 5 ข้อสอบข้อ 6 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 5 มีค่าความยากต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 11 และข้อ 2 มีค่าอำนาจจำแนกต่ำสุด สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 4 และข้อ 6 มีค่าการเดาต่ำสุด

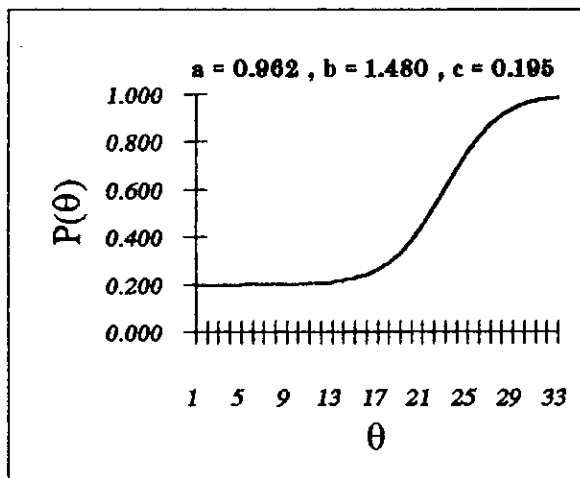


Fig 1

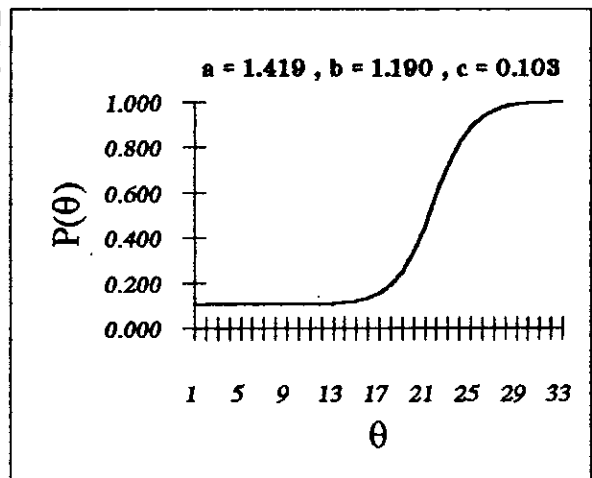


Fig 2

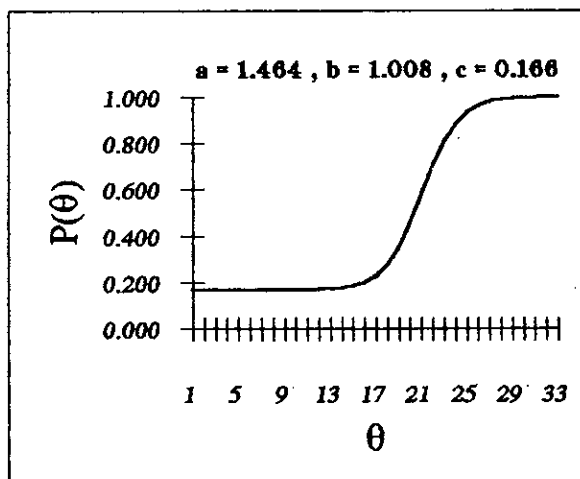


Fig 3

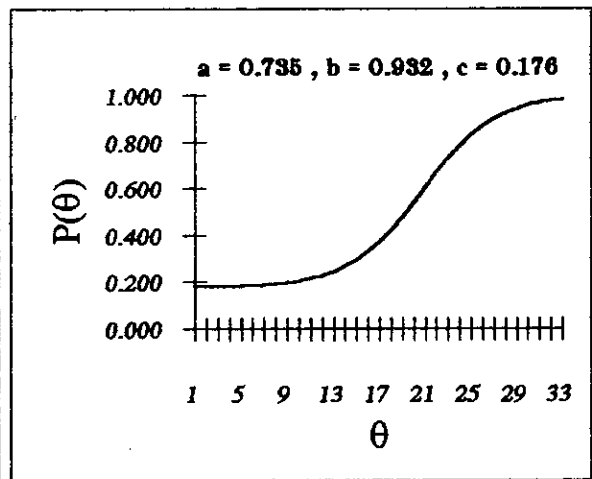


Fig 4

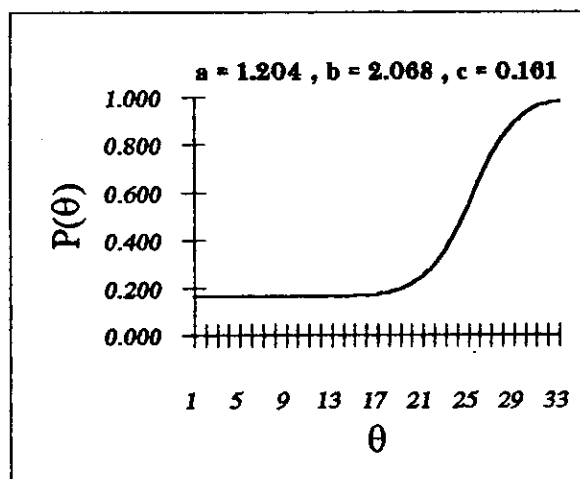


Fig 5

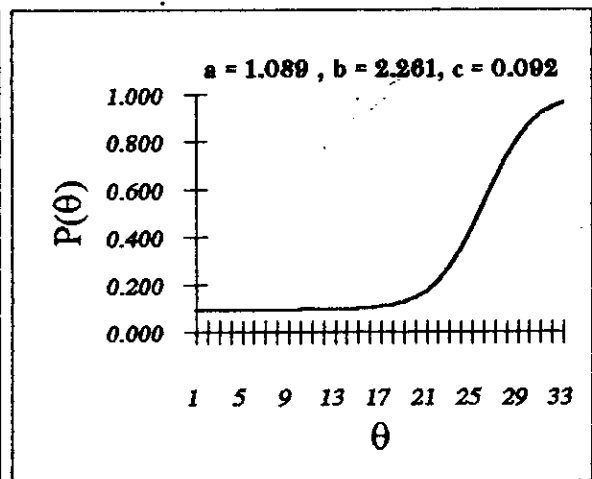


Fig 6

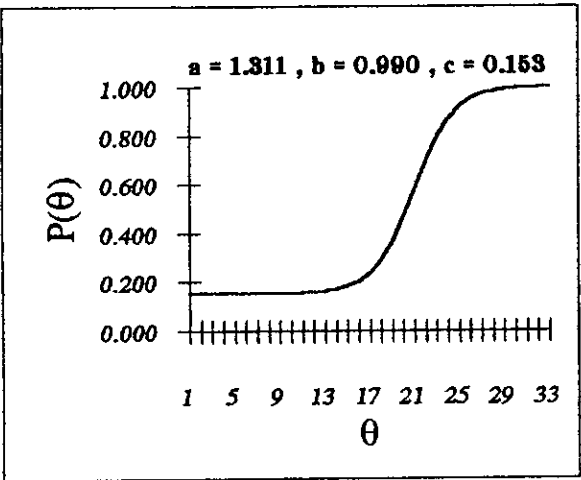


Fig 7

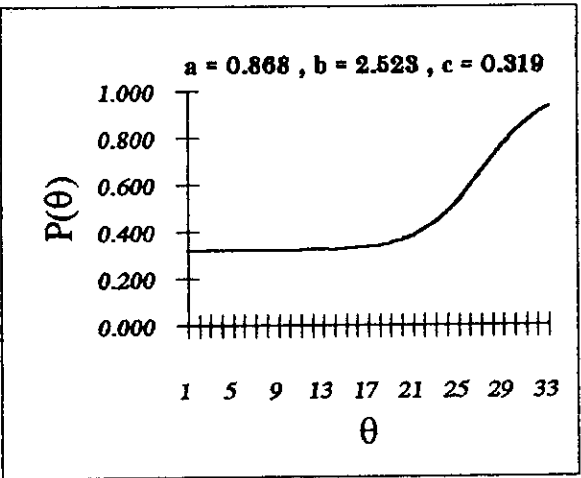


Fig 8

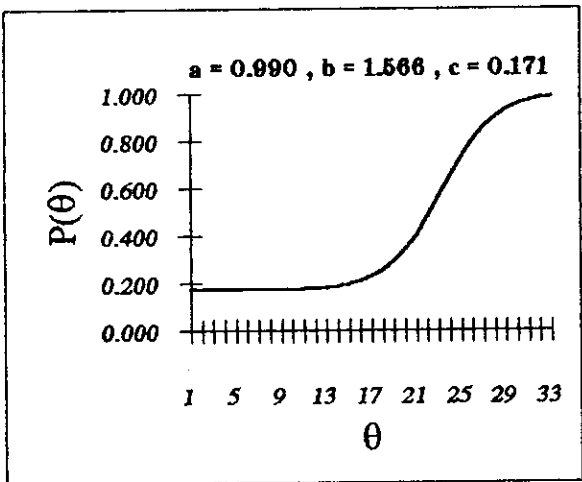


Fig 9

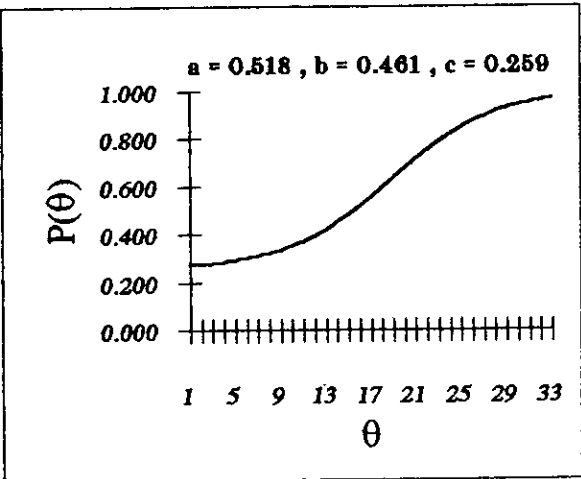


Fig 10

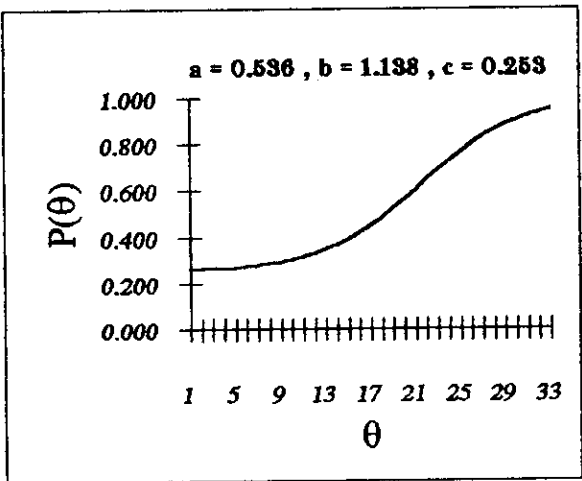


Fig 11

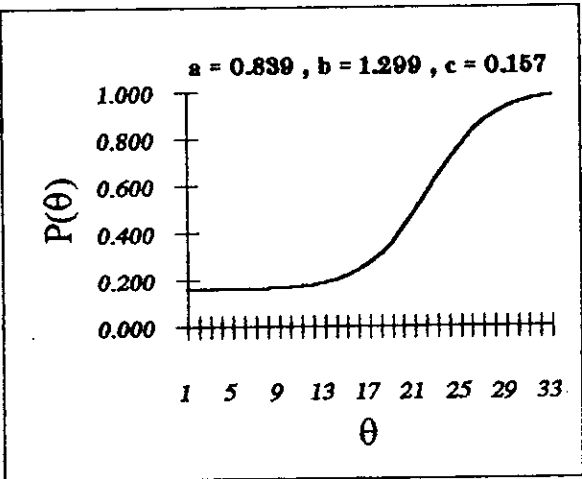
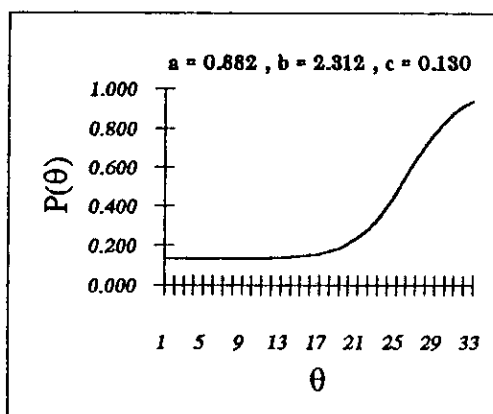
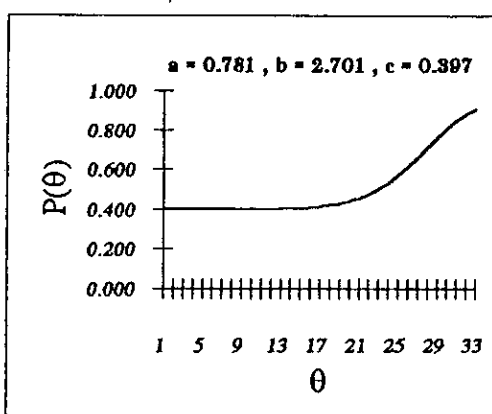


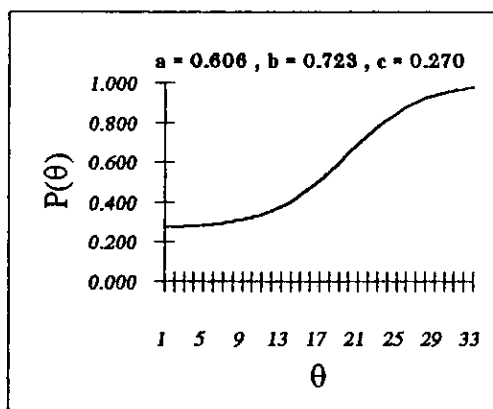
Fig 12



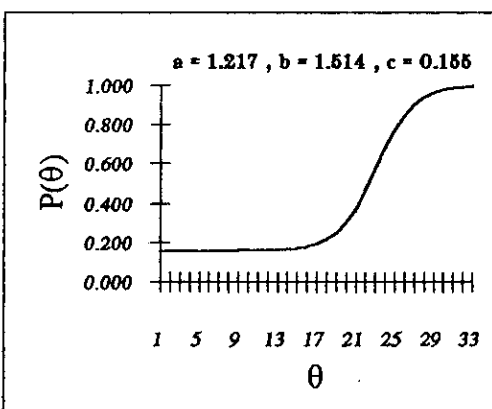
รูป 13



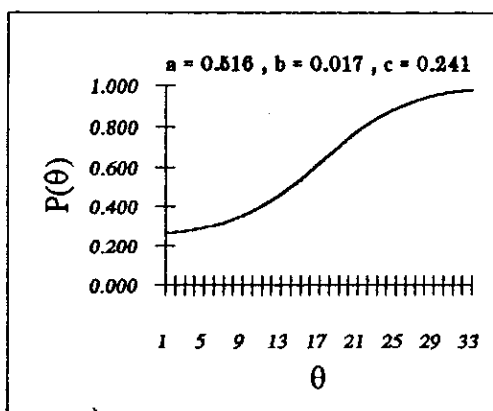
รูป 14



รูป 15



รูป 16



รูป 17

ภาพประกอบ 90 โค้งลักษณะข้อสอบของข้อสอบ ชนิด LR แบบเติมคำตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 90 พบว่า แบบทดสอบชนิด LR แบบเติมคำตอบของนักเรียน ชั้นมัธยมศึกษาปีที่ 5 ข้อสอบข้อ 14 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 17 มีค่าความยากต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 3 ส่วนข้อ 17 มีค่าอำนาจจำแนกต่ำสุด สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 14 และข้อ 6 มีค่าการเดาต่ำสุด

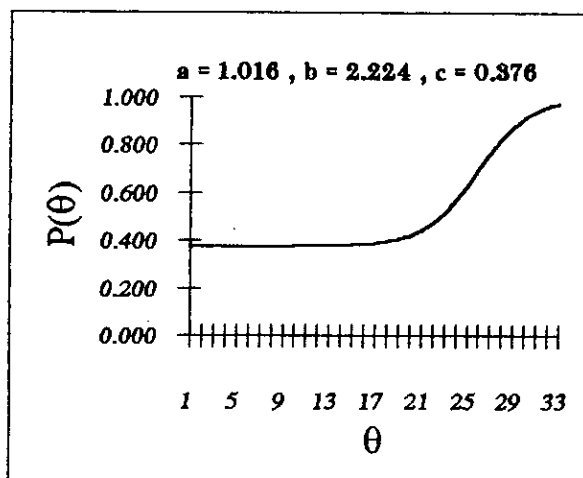


Fig 1

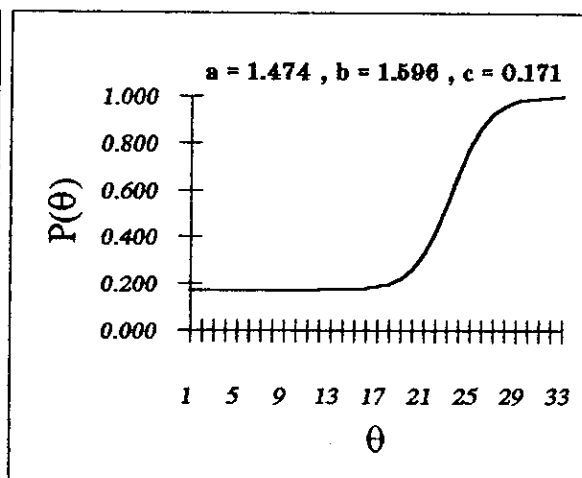


Fig 2

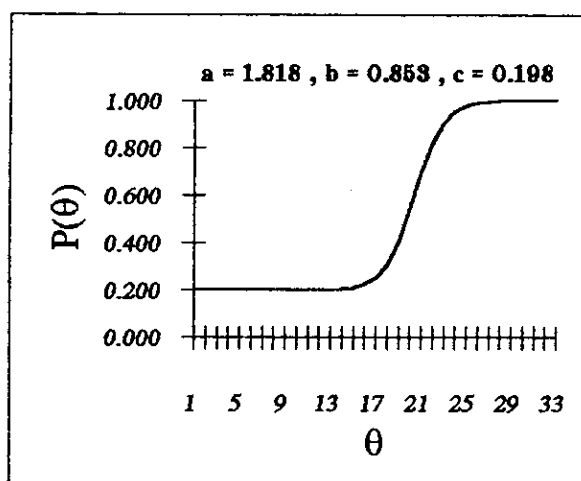


Fig 3

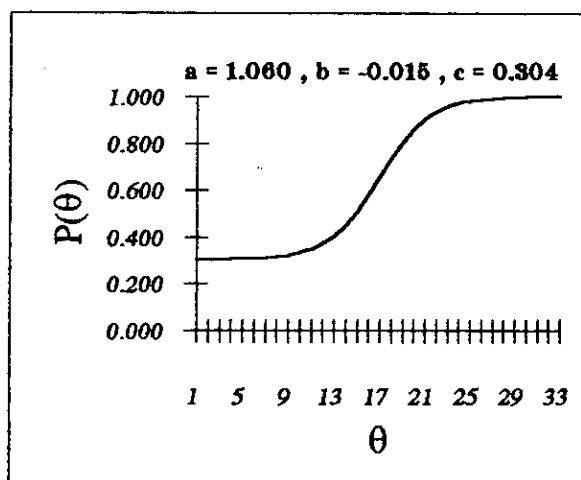


Fig 4

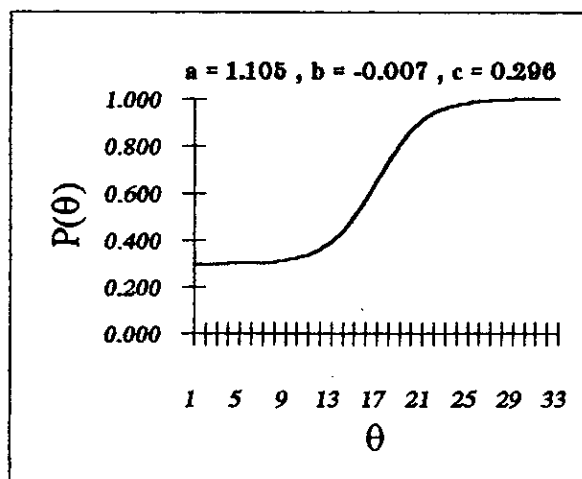


Fig 5

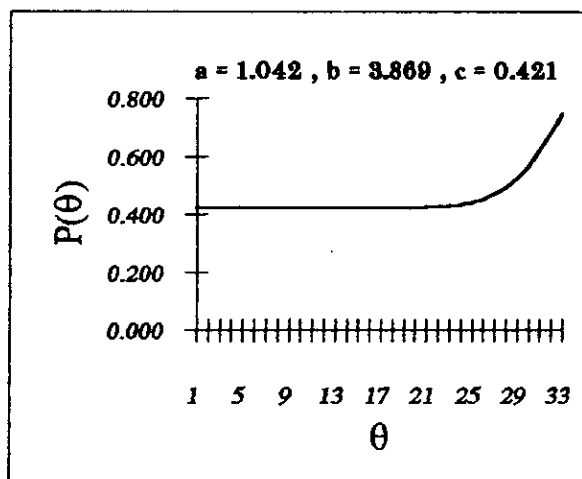
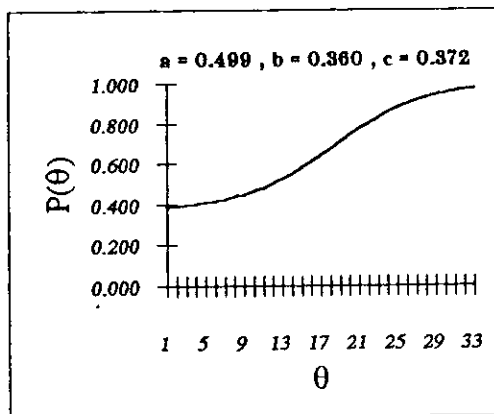
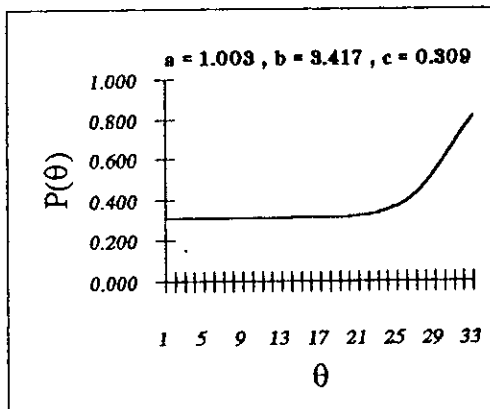


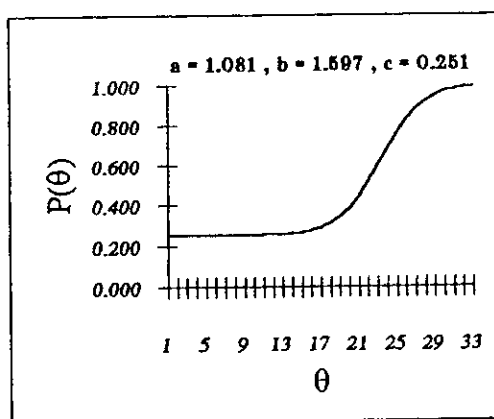
Fig 6



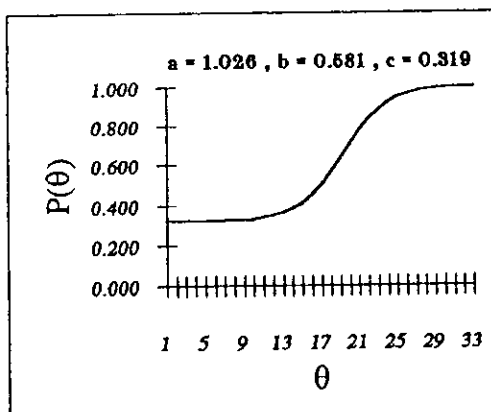
รูป 7



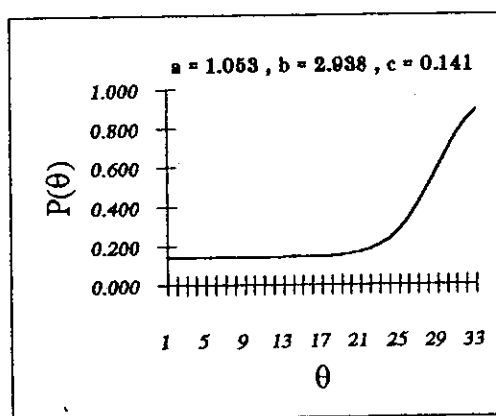
รูป 8



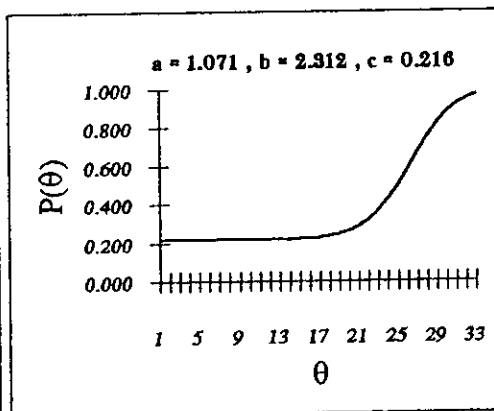
รูป 9



รูป 10



รูป 11



รูป 12

ภาพประกอบ 91 โค้งลักษณะข้อสอบของข้อสอบ ชนิด NLR แบบเดิมค่าตอบ ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 91 พบว่า แบบทดสอบชนิด NLR แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ข้อสอบข้อ 6 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 4 มีค่าความยาก
ต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 3 และข้อ 7 มีค่าอำนาจจำแนกต่ำสุด
สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 6 และข้อ 11 มีค่าการเดาต่ำสุด

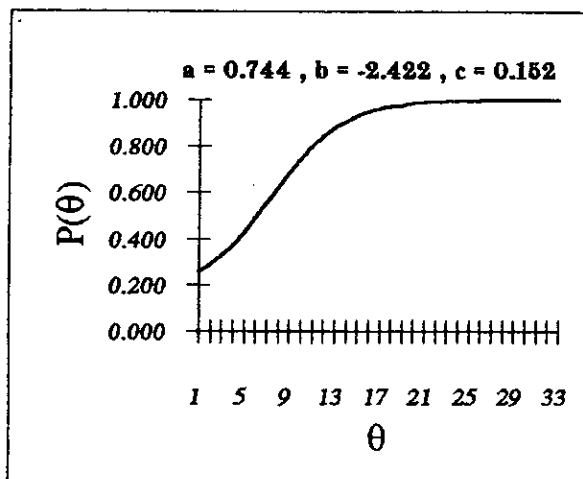


Fig 1

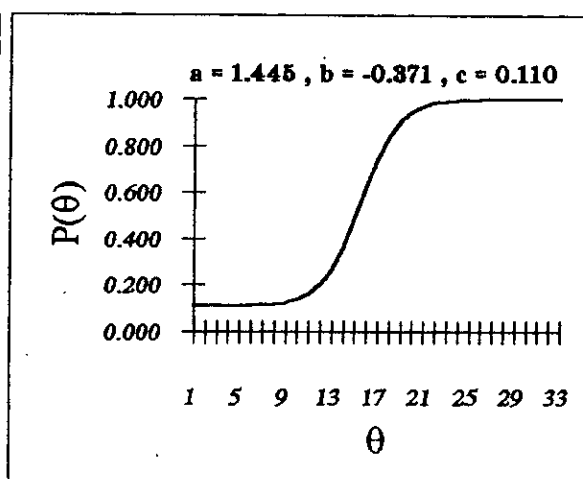


Fig 2

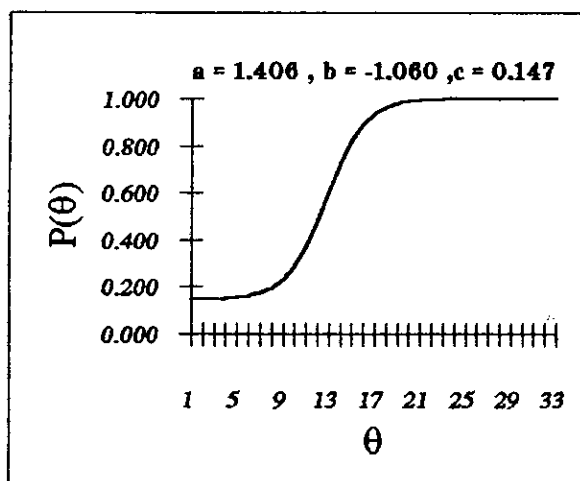


Fig 3

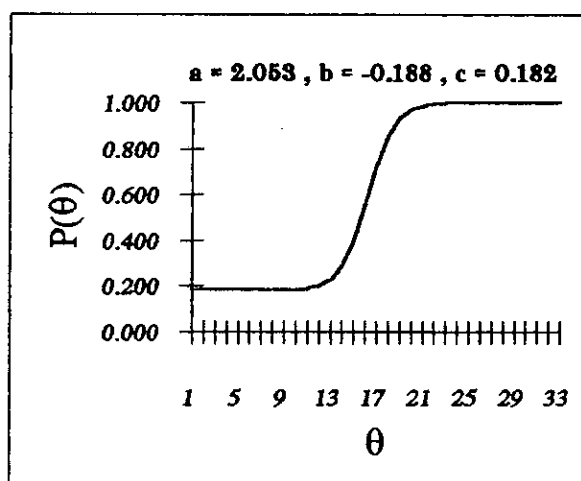


Fig 4

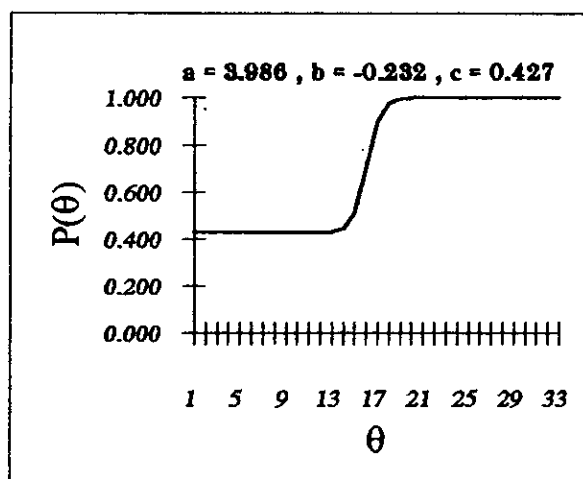


Fig 5

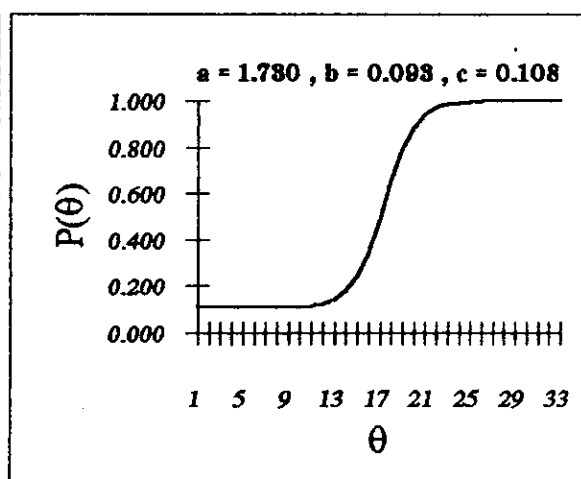


Fig 6

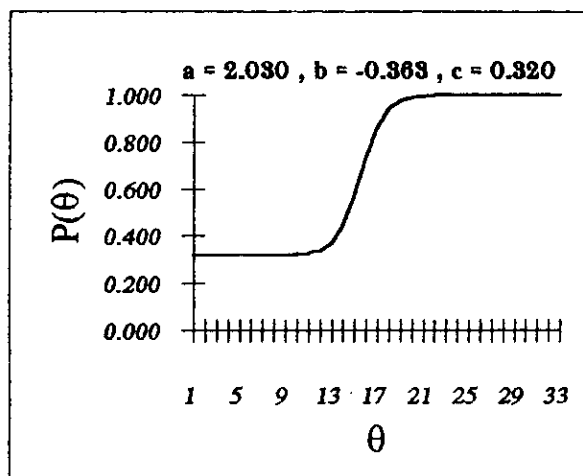


Fig 7

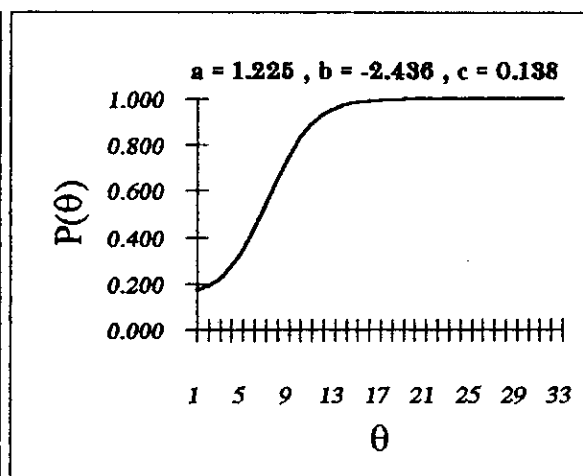


Fig 8

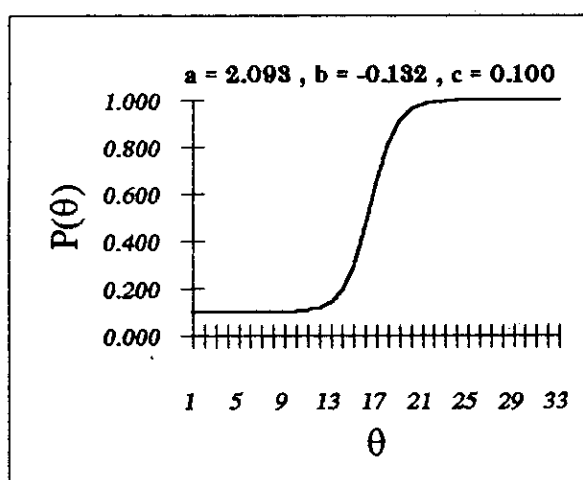


Fig 9

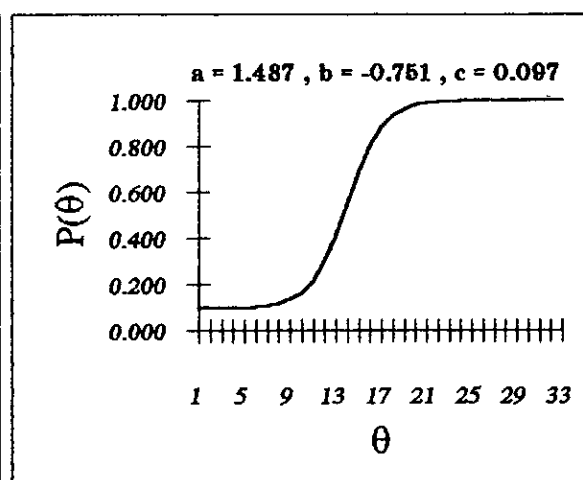


Fig 10

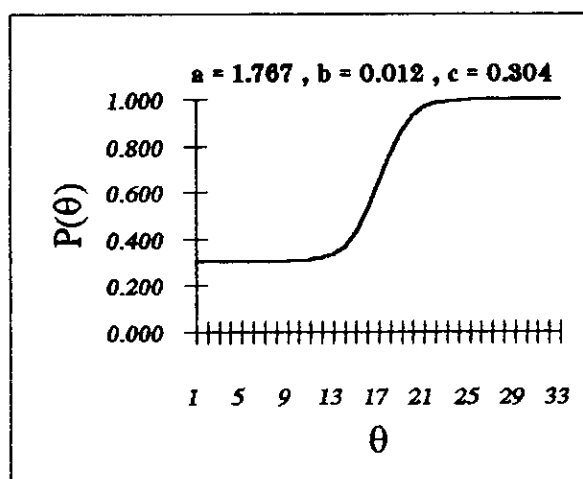


Fig 11

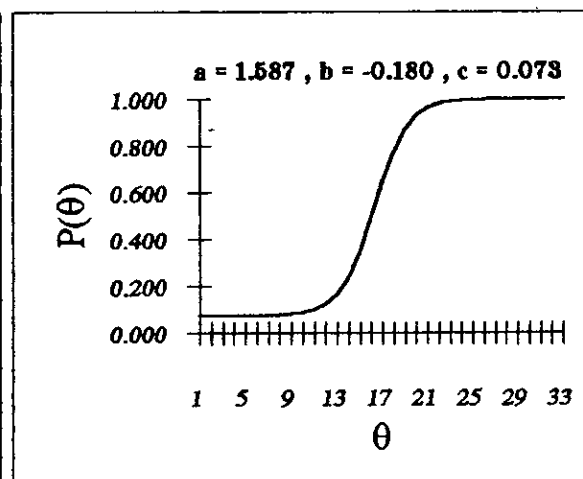


Fig 12

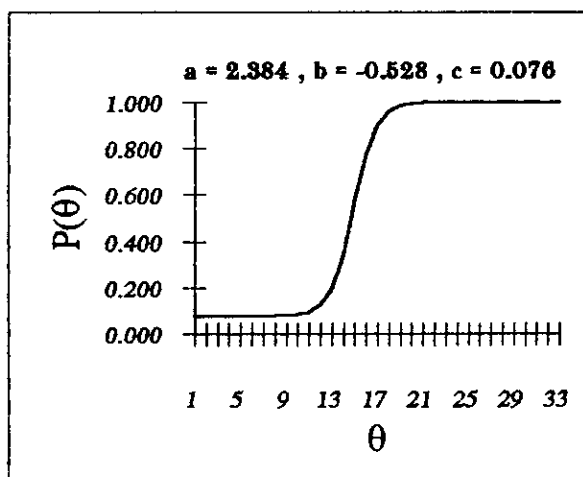


Fig 13

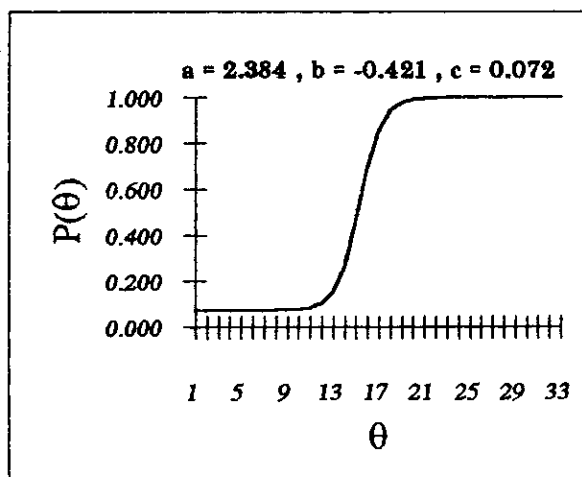


Fig 14

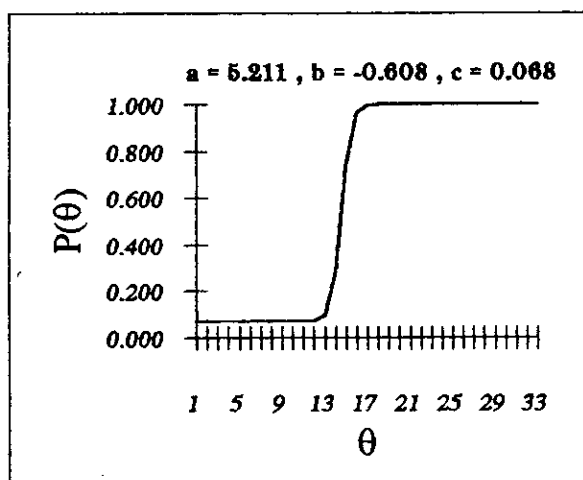


Fig 15

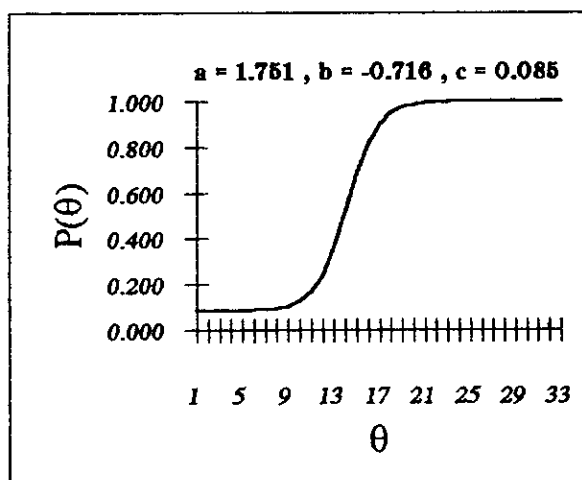


Fig 16

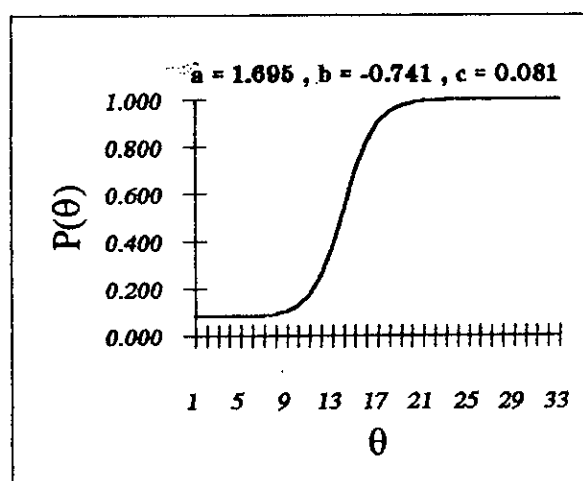


Fig 17

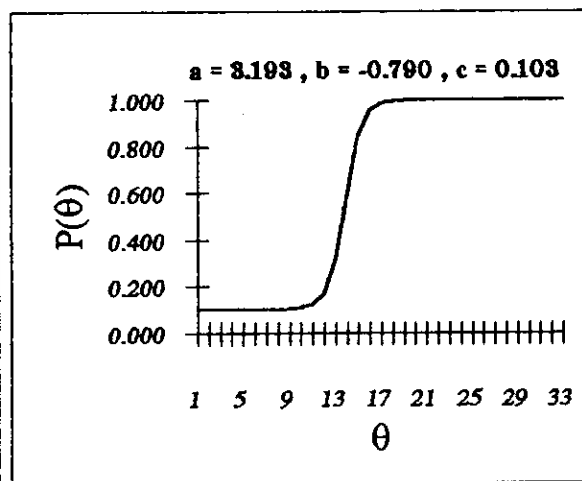
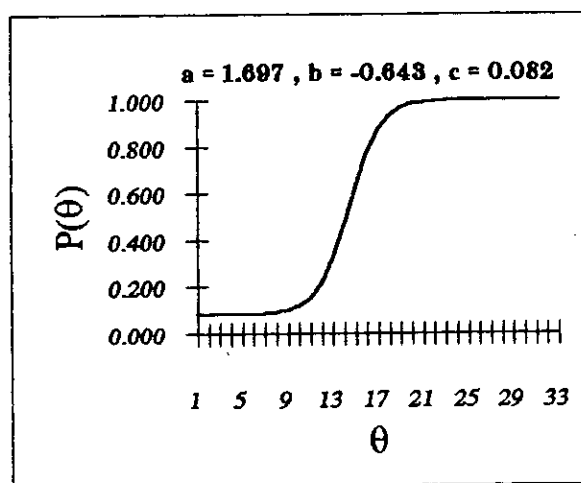


Fig 18

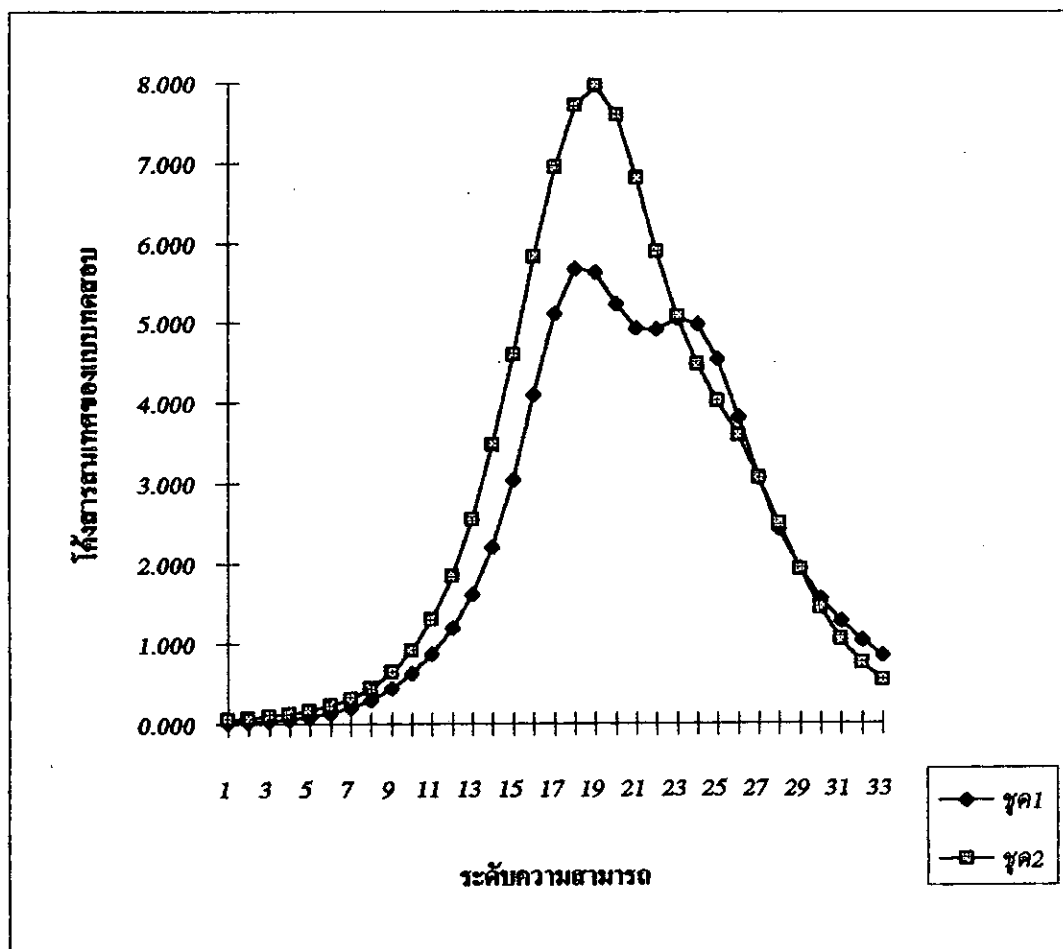


ข้อ 19

ภาพประกอบ 92 ได้ลักษณะข้อสอบของข้อสอบ ชนิด PI แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 92 พบว่า แบบทดสอบชนิด PI แบบเติมคำตอบของนักเรียน
ชั้นมัธยมศึกษาปีที่ 5 ข้อสอบข้อ 6 เป็นข้อสอบที่มีค่าความยากสูงสุด และข้อ 8 มีค่าความ
ยากต่ำสุด ส่วนข้อที่มีค่าอำนาจจำแนกสูงสุด คือ ข้อ 15 และข้อ 1 มีค่าอำนาจจำแนกต่ำ
สุด สำหรับข้อที่มีค่าการเดาสูงสุด คือ ข้อ 5 และข้อ 15 มีค่าการเดาต่ำสุด

2.4 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบแต่ละชนิด ซึ่งพัฒนาเป็น 2 รูปแบบ
คือ แบบเลือกตอบ และแบบเติมคำตอบ รูปแบบใดจะมีประสิทธิภาพสูงกว่ากัน ณ ระดับใด
ผู้วิจัยจึงได้ทำการวิเคราะห์ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบแต่ละชนิด โดยเปรียบเทียบ
ระหว่างรูปแบบของข้อคำถามแบบเลือกตอบกับแบบเติมคำตอบ แยกระดับชั้นมัธยมศึกษาปีที่ 3
และปีที่ 5 ตามลำดับ พร้อมกับนำเสนอกราฟแสดงการเปรียบเทียบโค้งสารสนเทศของแบบ
ทดสอบรูปแบบเลือกตอบ และรูปแบบเติมคำตอบ แต่ละชนิด แยกระดับชั้นมัธยมศึกษาปีที่ 3
และปีที่ 5 ตามลำดับ ปรากฏดังภาพประกอบ 93 - 116

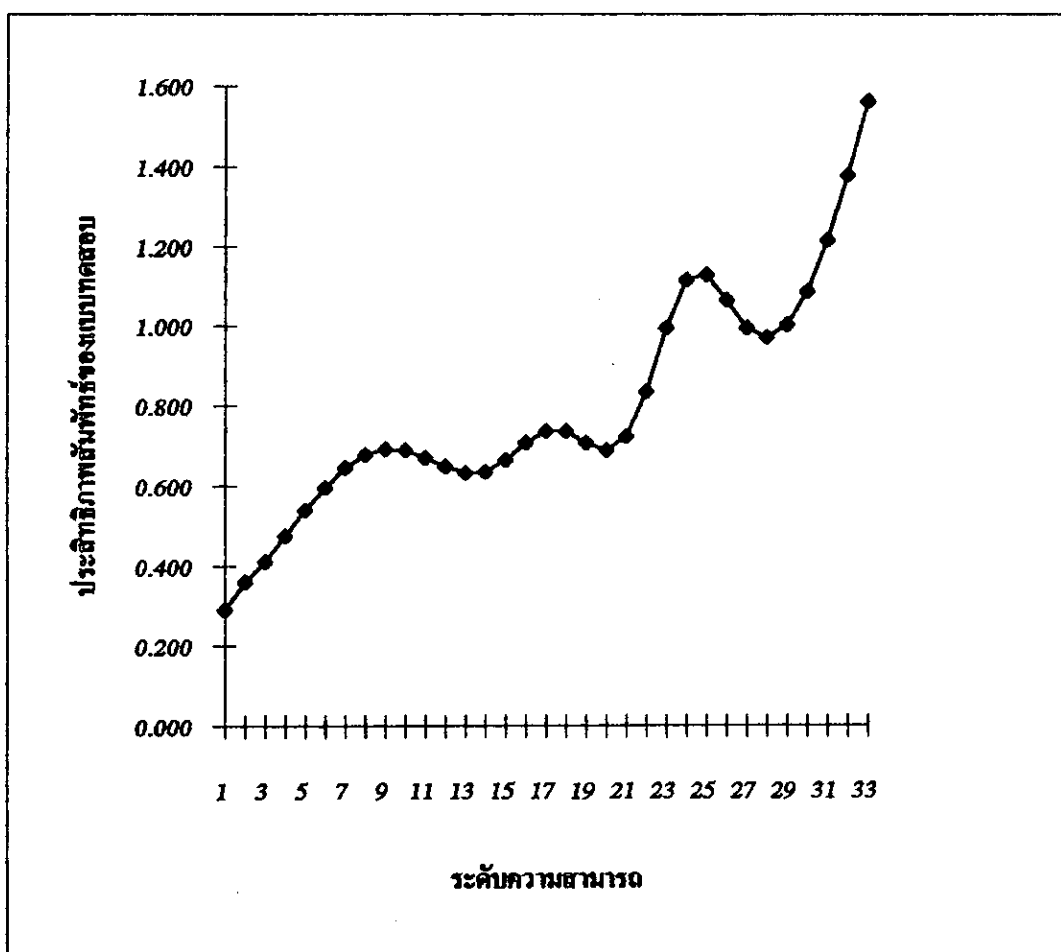


ชุด 1 AR_{31}

ชุด 2 AR_{32}

ภาพประกอบ 93 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบชนิด AR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

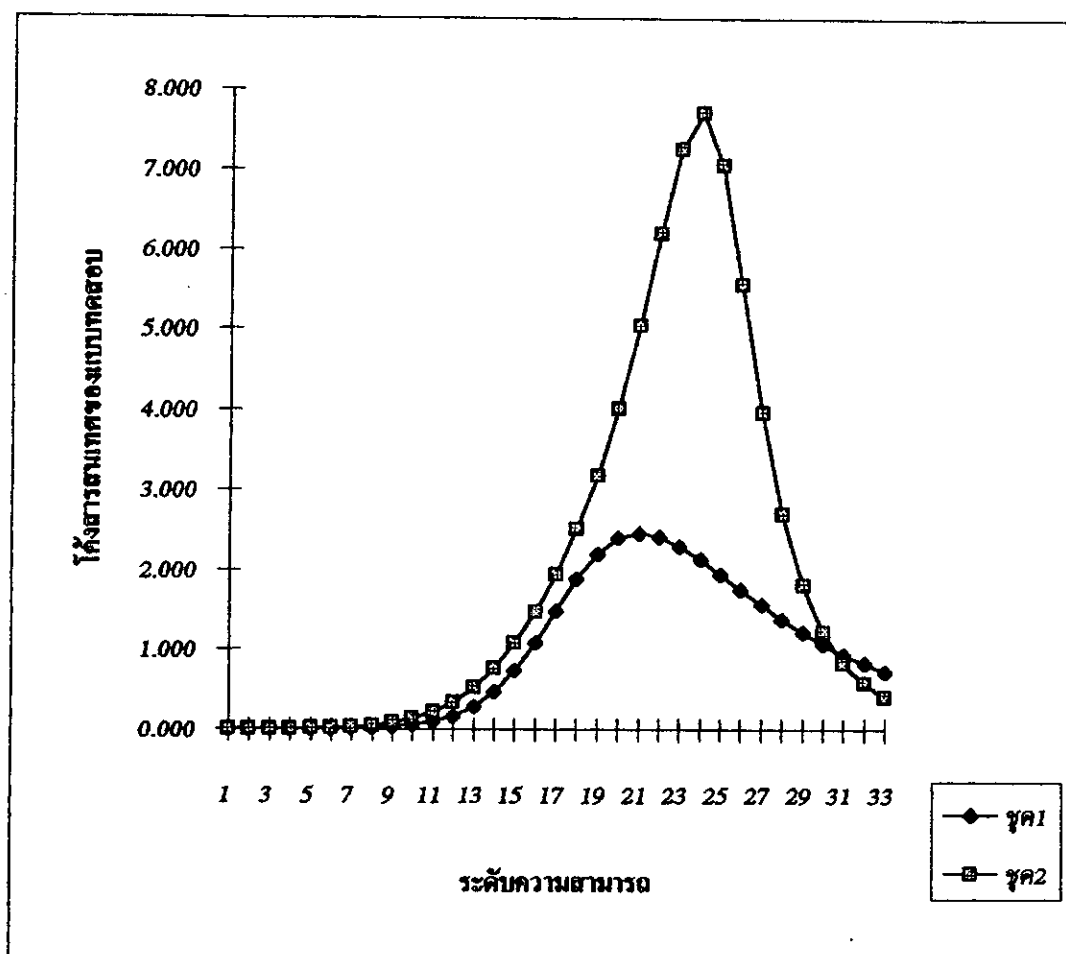
จากภาพประกอบ 93 พบว่า ค่าสารสนเทศของแบบทดสอบ ชนิด AR แบบเติมคำตอบ มีค่าฟังก์ชันสารสนเทศของแบบทดสอบ ณ ระดับความสามารถปานกลาง $\{I(\theta) = 7.954 ; \theta = 0.500\}$ สูงกว่าค่าฟังก์ชันสารสนเทศของแบบทดสอบ แบบเลือกตอบ ณ ระดับความสามารถปานกลาง $\{I(\theta) = 5.678 ; \theta = 0.250\}$



$$AR_{s_1} : AR_{s_2}$$

ภาพประกอบ 94 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด AR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

จากภาพประกอบ 94 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด AR แบบเลือกตอบกับแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีค่าต่ำสุดที่ระดับความสามารถค่า $\{RE(AR_{s_1} : AR_{s_2}) = 0.288 ; \theta = -4.000\}$ และมีค่าสูงสุดที่ระดับความสามารถสูง $\{RE(AR_{s_1} : AR_{s_2}) = 1.563 ; \theta = 4.000\}$

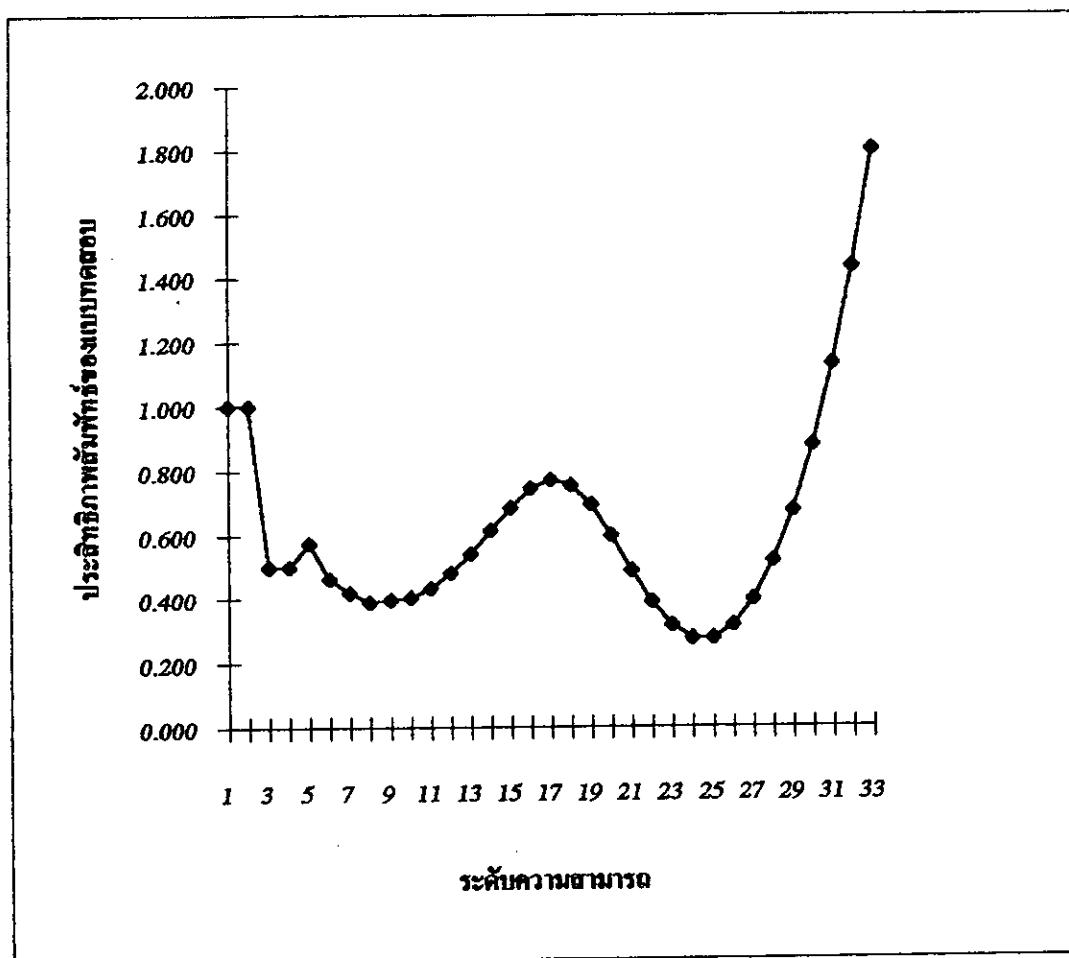


ชุด 1 AX_{s1}

ชุด 2 AX_{s2}

ภาพประกอบ 95 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบชนิด AX แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

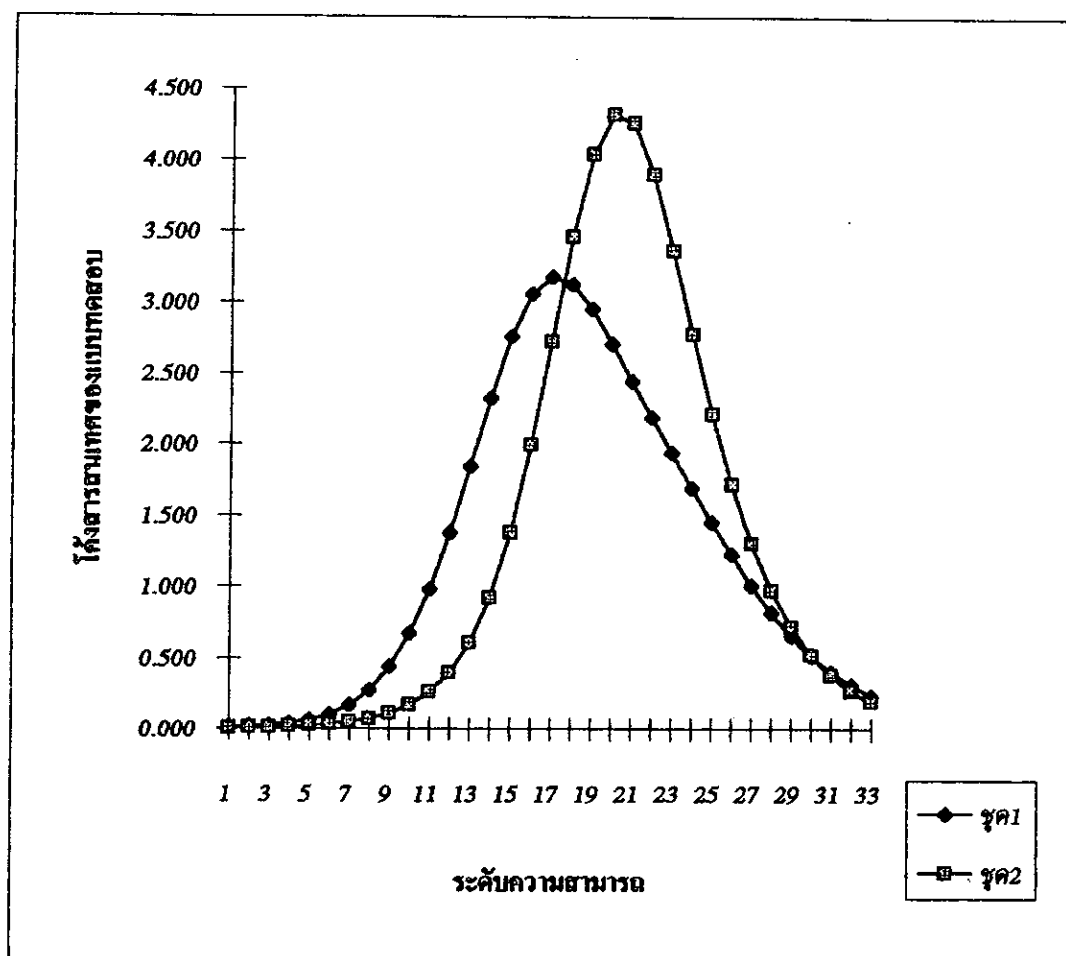
จากภาพประกอบ 95 พบว่า ค่าสารสนเทศของแบบทดสอบ ชนิด AX แบบเลือกตอบ มีค่าฟังก์ชันสารสนเทศของแบบทดสอบ ณ ระดับความสามารถปานกลาง $\{I(\theta) = 7.702 ; \theta = 1.750\}$ สูงกว่าค่าฟังก์ชันสารสนเทศของแบบทดสอบ แบบเติมคำตอบ ณ ระดับความสามารถปานกลาง $\{I(\theta) = 2.453 ; \theta = 1.000\}$



$$AX_{31} : AX_{32}$$

ภาพประกอบ 96 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด AX แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

จากภาพประกอบ 96 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด AX แบบเลือกตอบกับแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีค่าต่ำสุดที่ระดับความสามารถสูง $\{RE(AX_{31} : AX_{32}) = 0.275 ; \sigma = 2.000\}$ และมีค่าสูงสุดที่ระดับความสามารถสูง $\{RE(AX_{31} : AX_{32}) = 1.802 ; \sigma = 4.000\}$

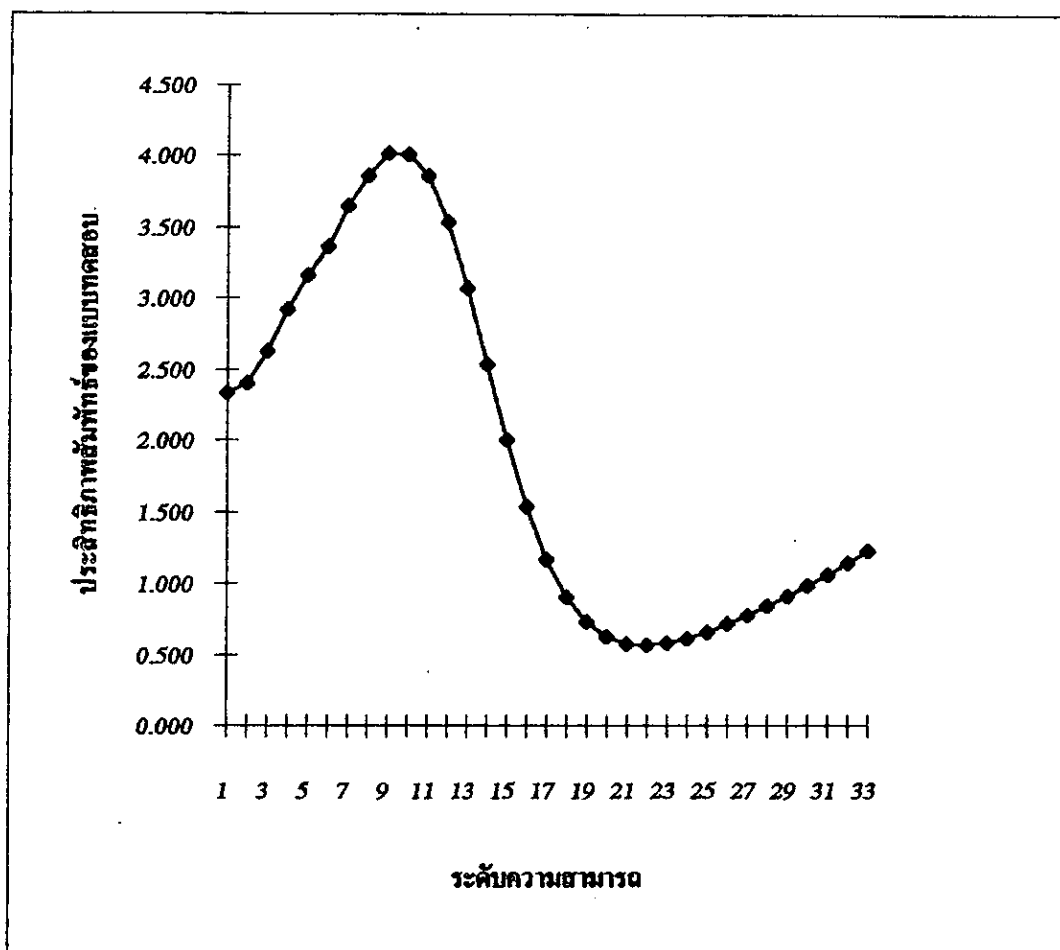


ชุด 1 CV₃₁

ชุด 2 CV₃₂

ภาพประกอบ 97 เปรียบเทียบโค้งฮิสโตแกรมของแบบทดสอบชนิด CV แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

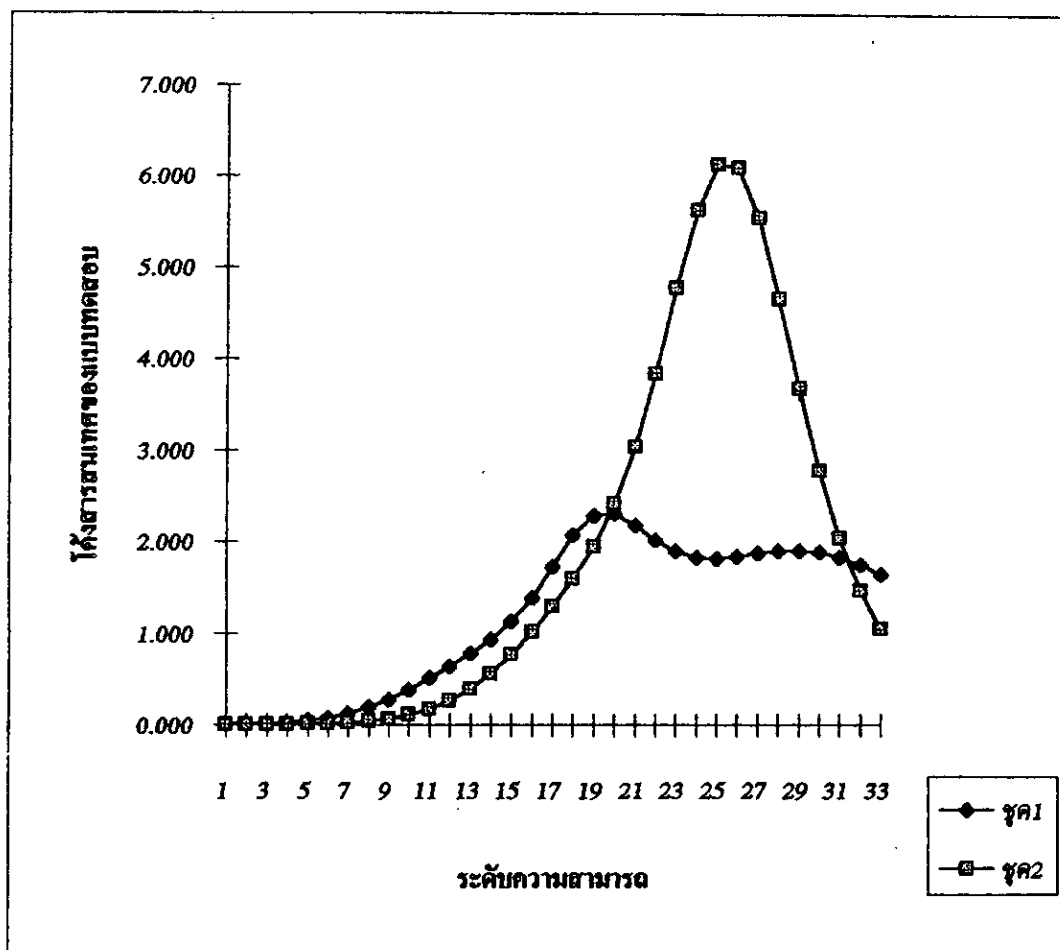
จากภาพประกอบ 97 พบว่า ค่าฮิสโตแกรมของแบบทดสอบ ชนิด CV แบบเติมคำตอบ มีค่าฟังก์ชันฮิสโตแกรมของแบบทดสอบ ณ ระดับความสามารถปานกลาง $\{I(e) = 4.322 ; e = 0.750\}$ สูงกว่าค่าฟังก์ชันฮิสโตแกรมของแบบทดสอบ แบบเลือกตอบ ณ ระดับความสามารถปานกลาง $\{I(e) = 3.184 ; e = 0.000\}$



$$CV_{s1} = CV_{s2}$$

ภาพประกอบ 98 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด CV แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

จากภาพประกอบ 98 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด CV แบบเลือกตอบกับแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีค่าต่ำสุดที่ระดับความสามารถปานกลาง $\{RE(CV_{s1} : CV_{s2}) = 0.562 ; e = 1.250\}$ และมีค่าสูงสุดที่ระดับความสามารถต่ำ $\{RE(CV_{s1} : CV_{s2}) = 4.019 ; e = -2.000\}$

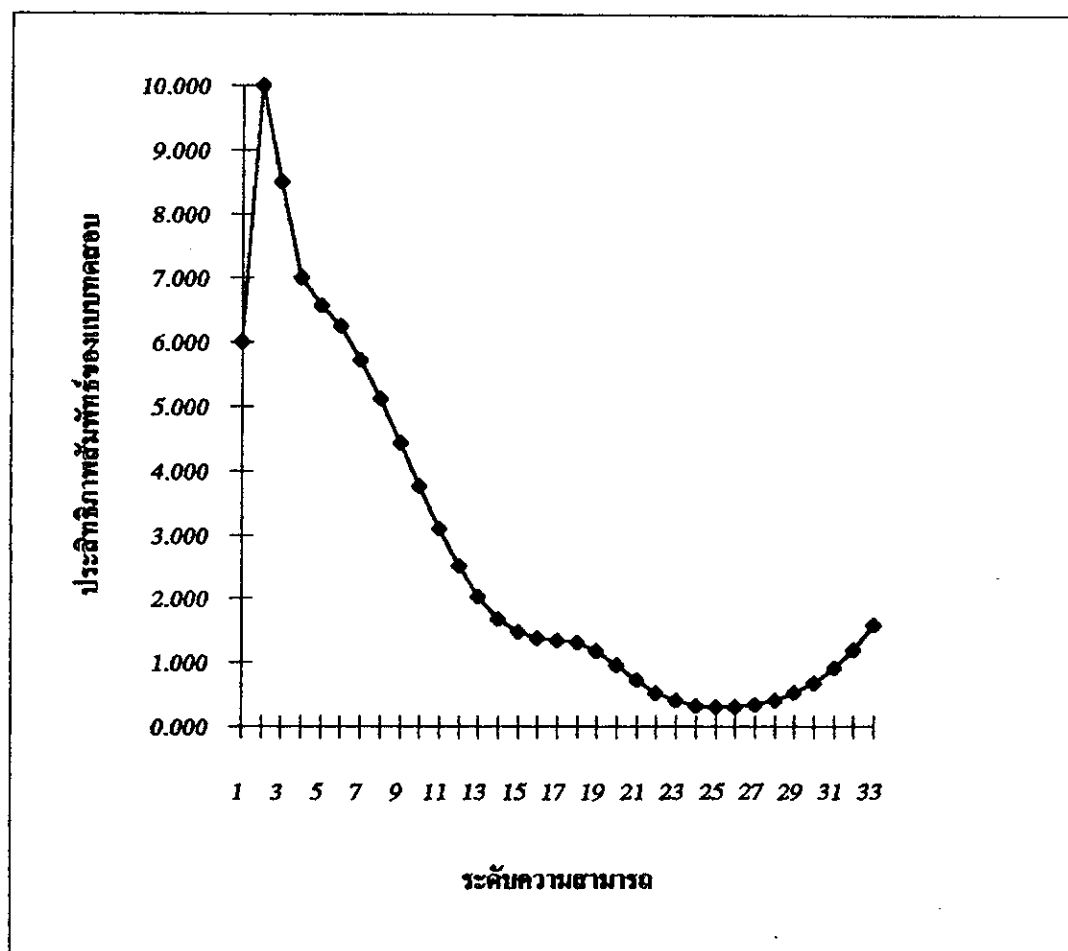


ชุด 1 $LR_{3.1}$

ชุด 2 $LR_{3.2}$

ภาพประกอบ 99 เปรียบเทียบได้จำนวนที่สอบแบบทดสอบชนิด LR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

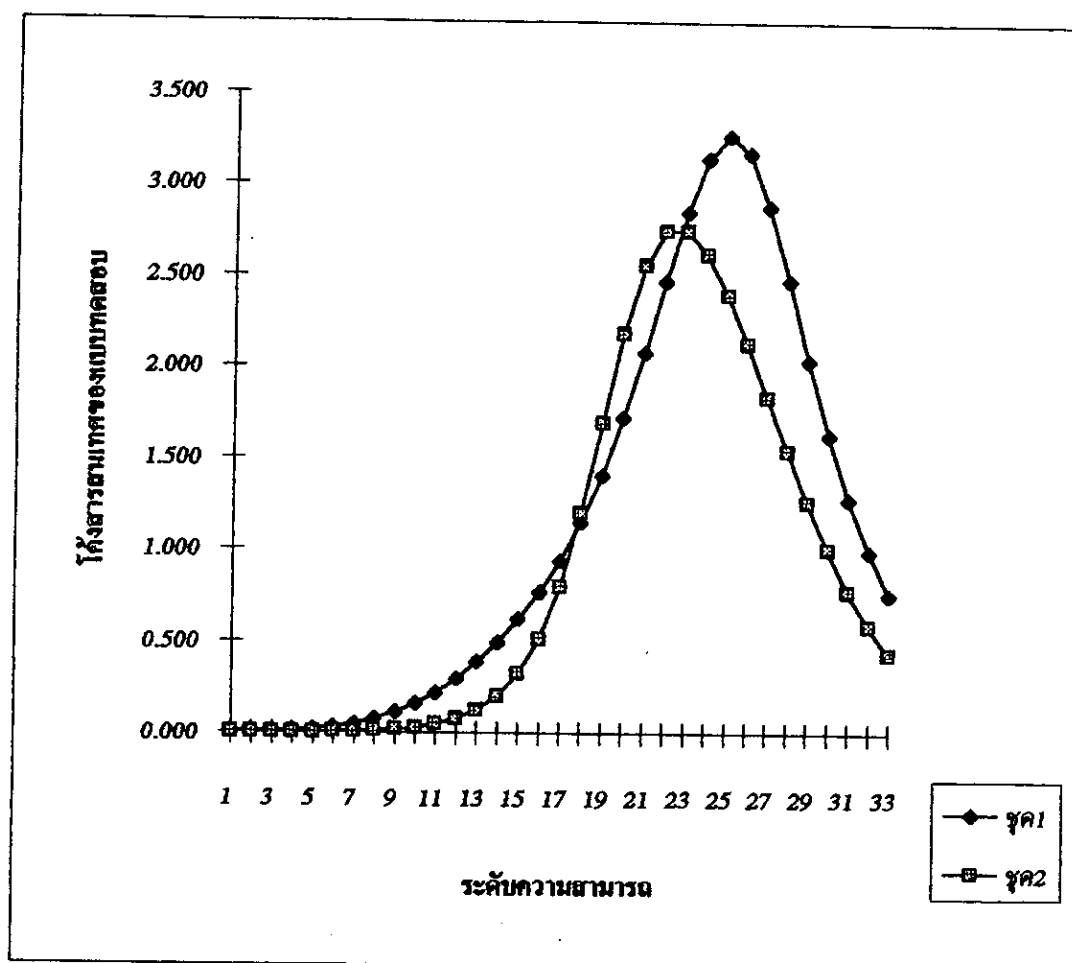
จากภาพประกอบ 99 พบว่า ค่าสารสนเทศของแบบทดสอบ ชนิด LR แบบเติมคำตอบ มีค่าฟังก์ชันสารสนเทศของแบบทดสอบ ณ ระดับความสามารถสูง $\{I(\theta) = 6.135 ; \theta = 2.000\}$ สูงกว่าค่าฟังก์ชันสารสนเทศของแบบทดสอบ แบบเลือกตอบ ณ ระดับความสามารถปานกลาง $\{I(\theta) = 2.310 ; \theta = 0.750\}$



$$LR_{31} : LR_{32}$$

ภาพประกอบ 100 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด LR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

จากภาพประกอบ 100 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด LR แบบเลือกตอบกับแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีค่าต่ำสุดที่ระดับความสามารถสูง $\{RE(LR_{31} : LR_{32}) = 0.297 ; \sigma = 2.000\}$ และมีค่าสูงสุดที่ระดับความสามารถต่ำ $\{RE(LR_{31} : LR_{32}) = 10.000 ; \sigma = -3.750\}$

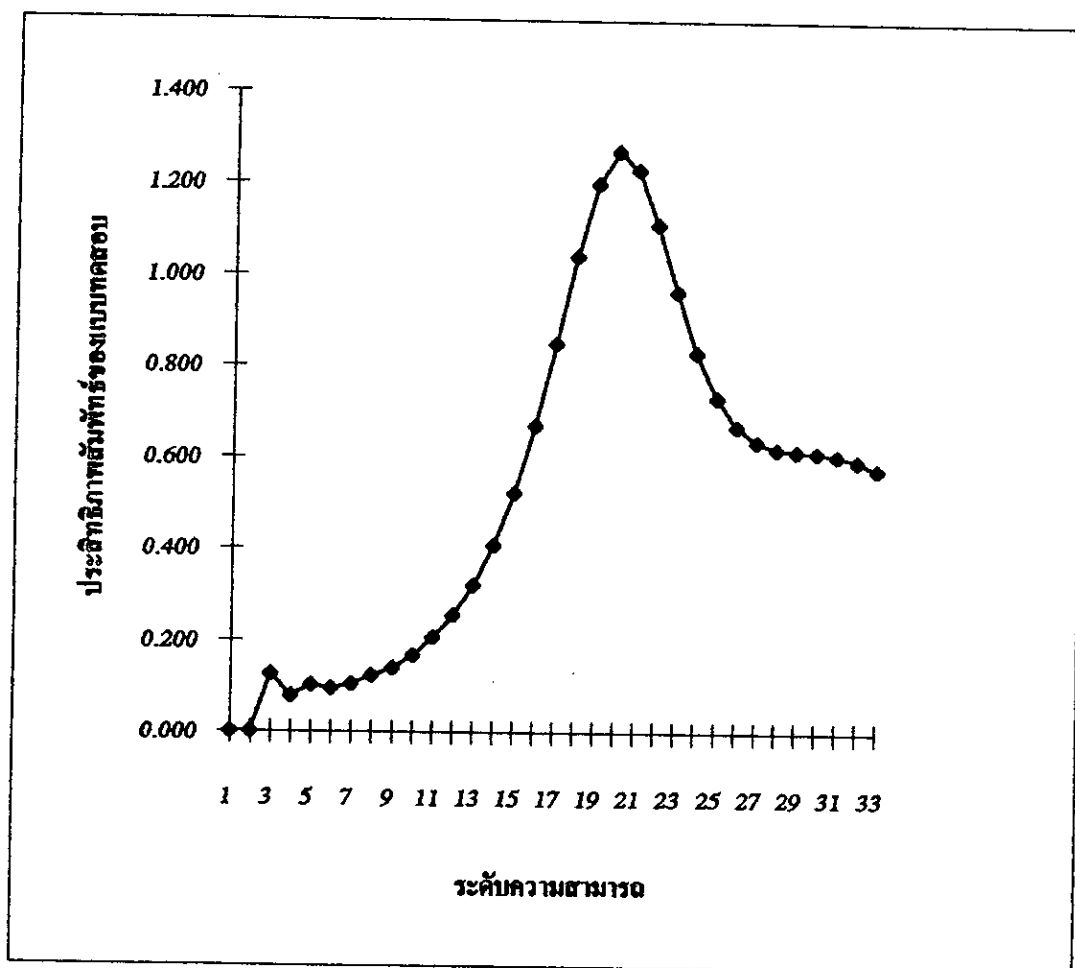


ชุด 1 NLR_{31}

ชุด 2 NLR_{32}

ภาพประกอบ 101 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบชนิด NLR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

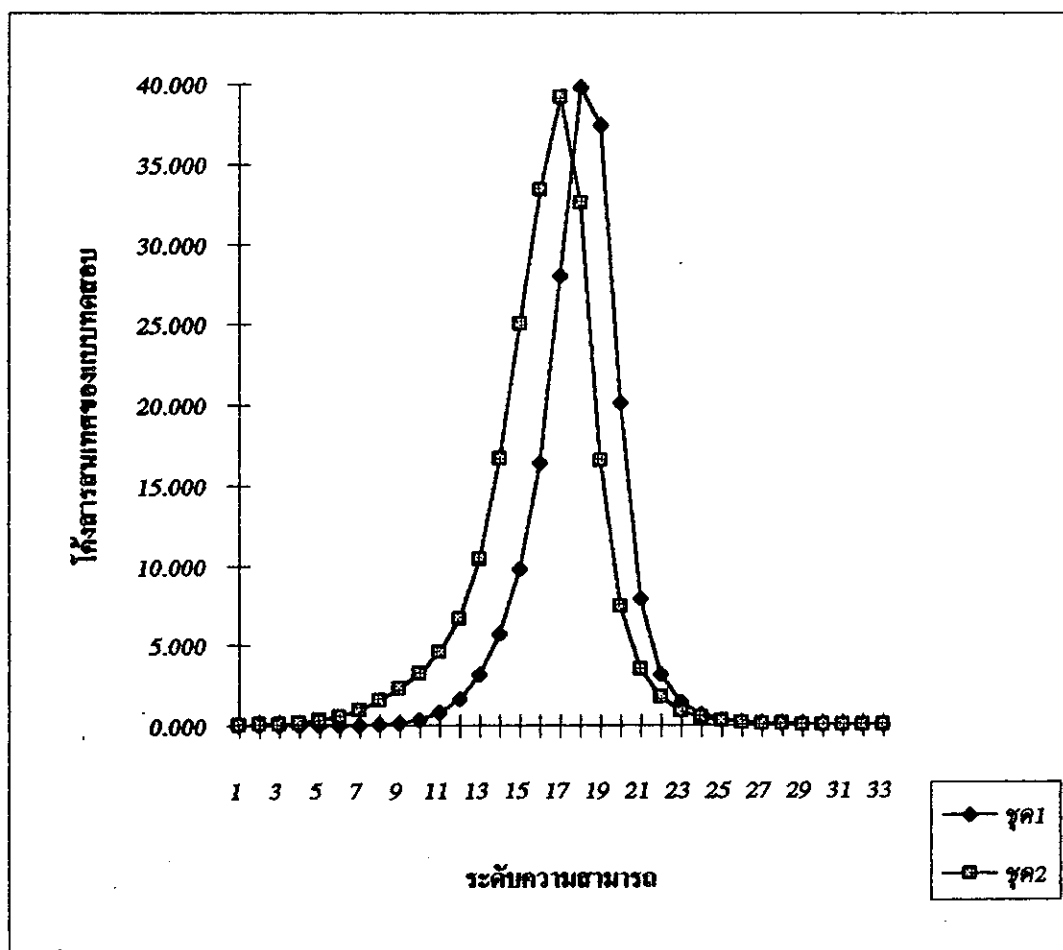
จากภาพประกอบ 101 พบว่า ค่าสารสนเทศของแบบทดสอบ ชนิด NLR แบบเลือกตอบ มีค่าฟังก์ชันสารสนเทศของแบบทดสอบ ณ ระดับความสามารถสูง $\{I(\theta) = 3.266 ; \theta = 2.000\}$ สูงกว่าค่าฟังก์ชันสารสนเทศของแบบทดสอบ แบบเติมคำตอบ ณ ระดับความสามารถสูง $\{I(\theta) = 2.747 ; \theta = 1.500\}$



$$NLR_{31} : NLR_{32}$$

ภาพประกอบ 102 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด NLR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

จากภาพประกอบ 102 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด NLR แบบเลือกตอบกับแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีค่าต่ำสุดที่ระดับความสามารถต่ำ $\{RE(NLR_{31} : NLR_{32}) = 0.000 ; \sigma = -4.000\}$ และมีค่าสูงสุดที่ระดับความสามารถปานกลาง $\{RE(NLR_{31} : NLR_{32}) = 1.271 ; \sigma = 0.750\}$

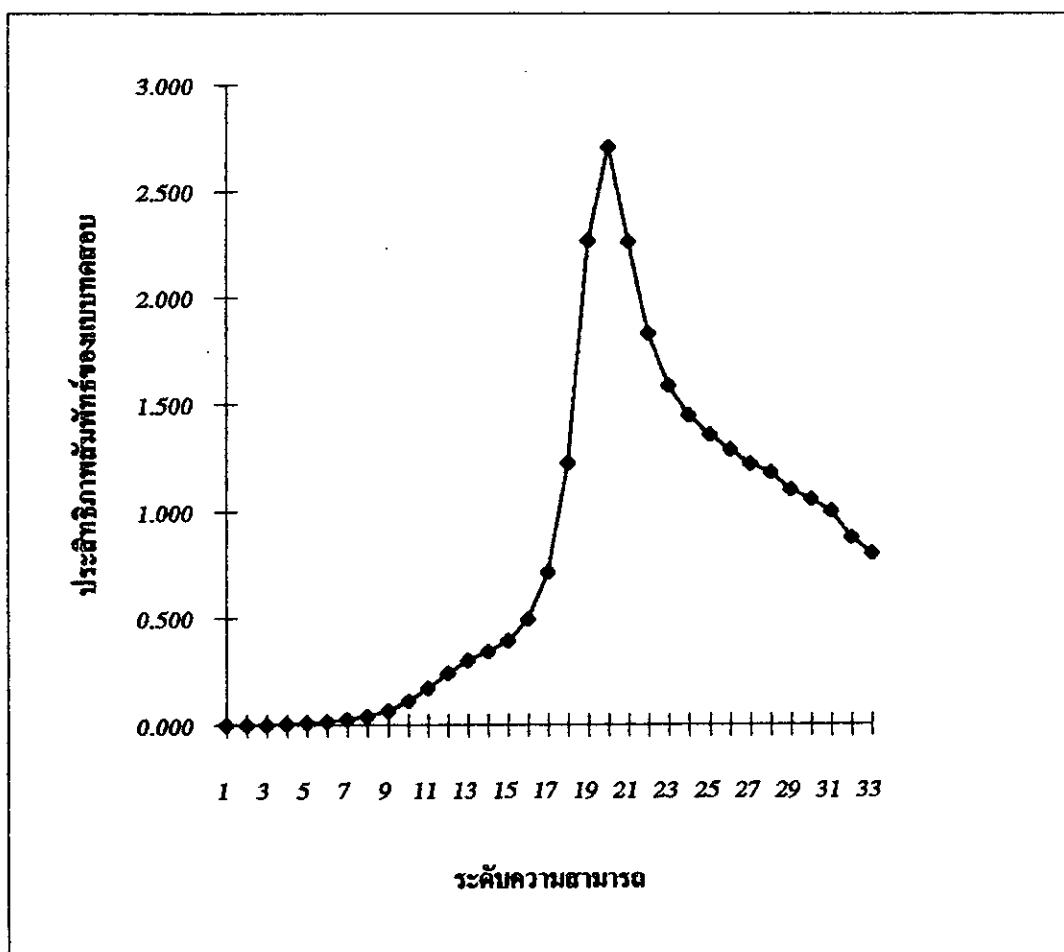


ชุด 1 PI₃₁

ชุด 2 PI₃₂

ภาพประกอบ 103 เปรียบเทียบค่าสารสนเทศของแบบทดสอบชนิด PI แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

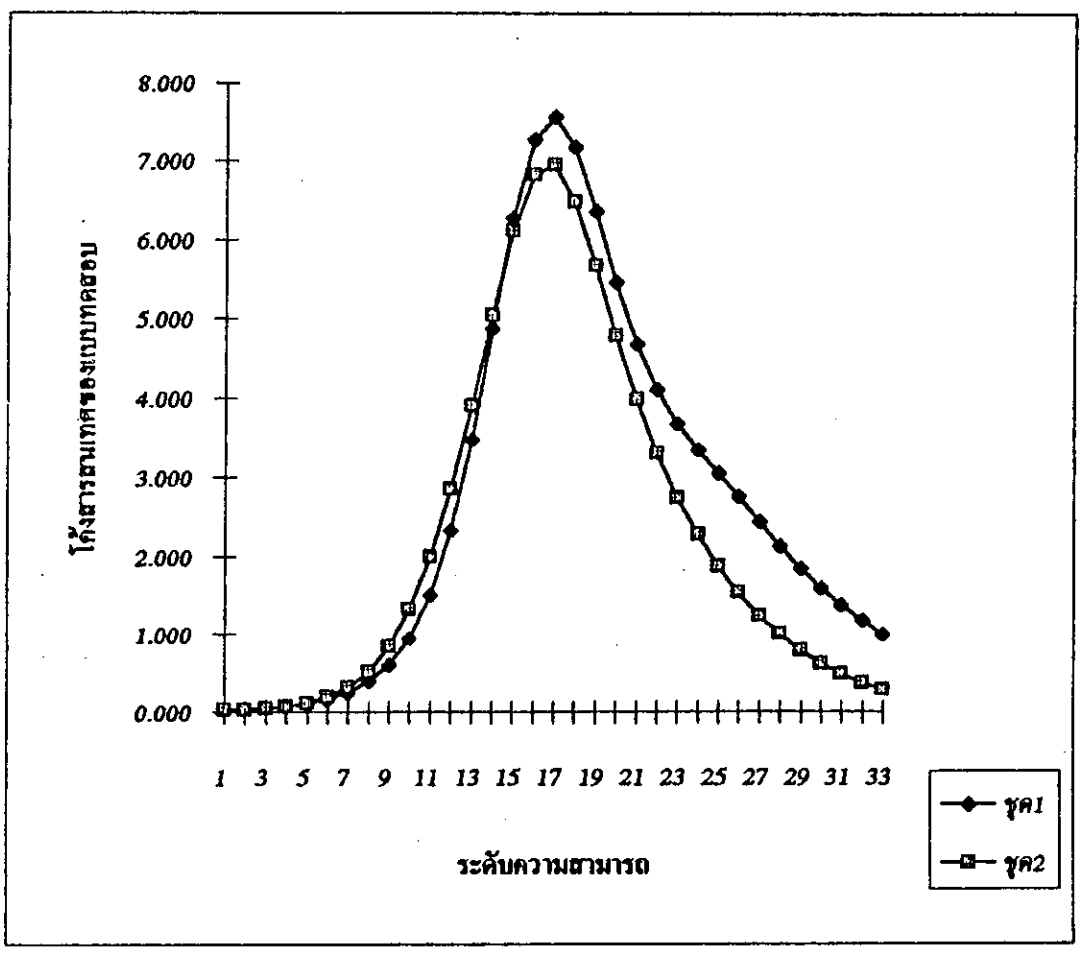
จากภาพประกอบ 103 พบว่า ค่าสารสนเทศของแบบทดสอบ ชนิด PI แบบเลือกตอบ มีค่าฟังก์ชันสารสนเทศของแบบทดสอบ ณ ระดับความสามารถปานกลาง $\{I(e) = 39.803 ; e = 0.250\}$ สูงกว่าค่าฟังก์ชันสารสนเทศของแบบทดสอบ แบบเติมคำตอบ ณ ระดับความสามารถปานกลาง $\{I(e) = 39.214 ; e = 0.000\}$



$$PI_{31} : PI_{32}$$

ภาพประกอบ 104 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด PI แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

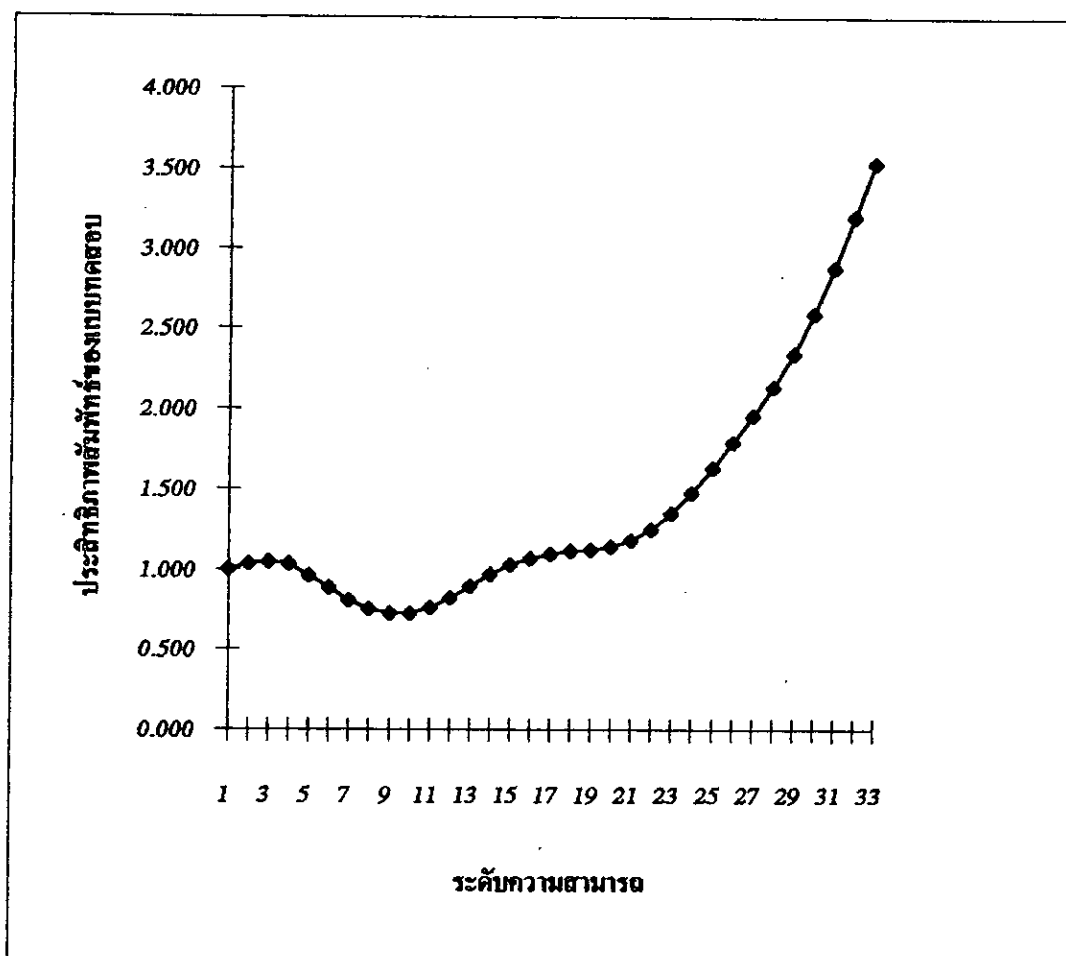
จากภาพประกอบ 104 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด PI แบบเลือกตอบกับแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีค่าต่ำสุดที่ระดับความสามารถต่ำ $\{RE(PI_{31} : PI_{32}) = 0.000 ; e = -4.000, -3.750 \text{ และ } -3.500\}$ และมีค่าสูงสุดที่ระดับความสามารถปานกลาง $\{RE(PI_{31} : PI_{32}) = 2.711 ; e = 0.750\}$



ชุด 1 AR_{51}
 ชุด 2 AR_{52}

ภาพประกอบ 105 เปรียบเทียบได้สารสนเทศของแบบทดสอบชนิด AR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

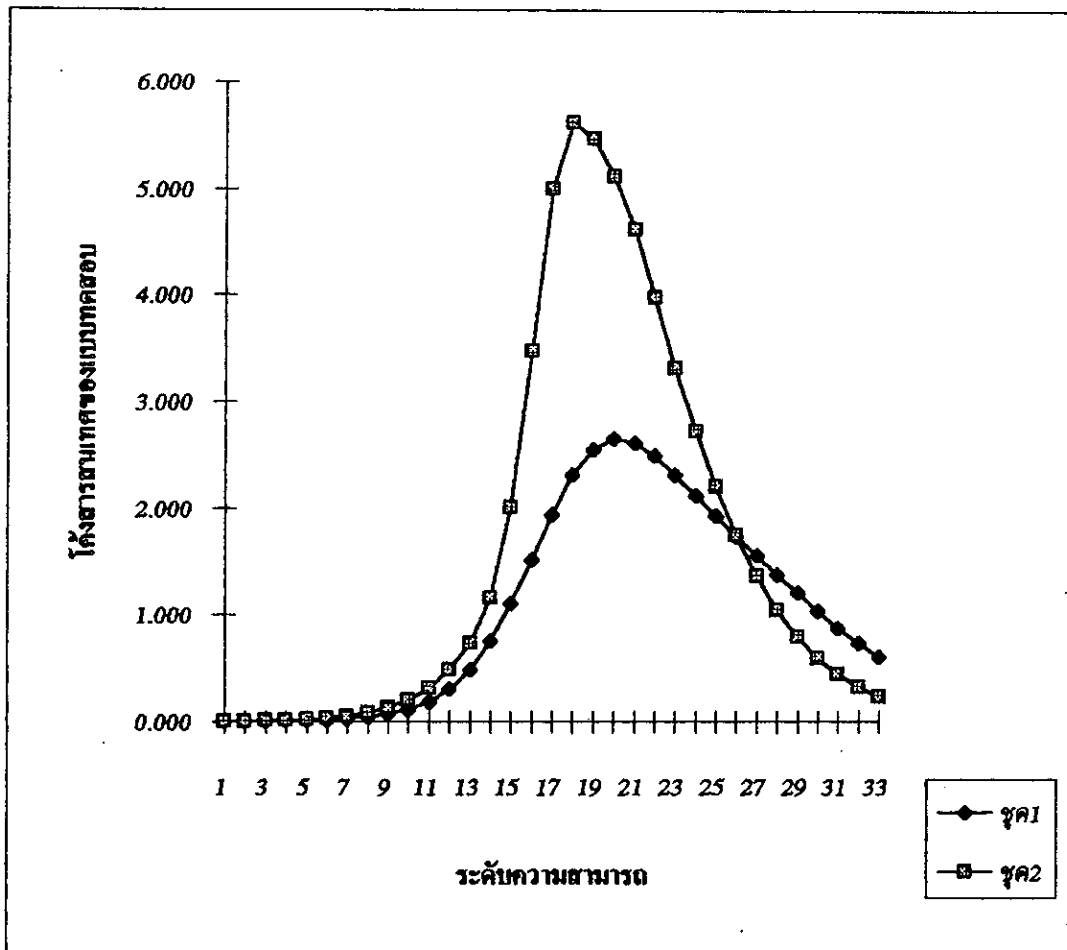
จากภาพประกอบ 105 พบว่า ค่าสารสนเทศของแบบทดสอบ ชนิด AR แบบเลือกตอบ มีค่าฟังก์ชันสารสนเทศของแบบทดสอบ ณ ระดับความสามารถปานกลาง $\{I(\theta) = 7.563 ; \theta = 0.000\}$ สูงกว่าค่าฟังก์ชันสารสนเทศของแบบทดสอบ แบบเติมคำตอบ ณ ระดับความสามารถปานกลาง $\{I(\theta) = 6.937 ; \theta = 0.000\}$



$$AR_{s_1} : AR_{s_2}$$

ภาพประกอบ 106 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด AR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 106 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด AR แบบเลือกตอบกับแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 มีค่าต่ำสุดที่ระดับความสามารถต่ำ $\{RE(AR_{s_1} : AR_{s_2}) = 0.720 ; e = -2.000\}$ และมีค่าสูงสุดที่ระดับความสามารถสูง $\{RE(AR_{s_1} : AR_{s_2}) = 3.534 ; e = 4.000\}$

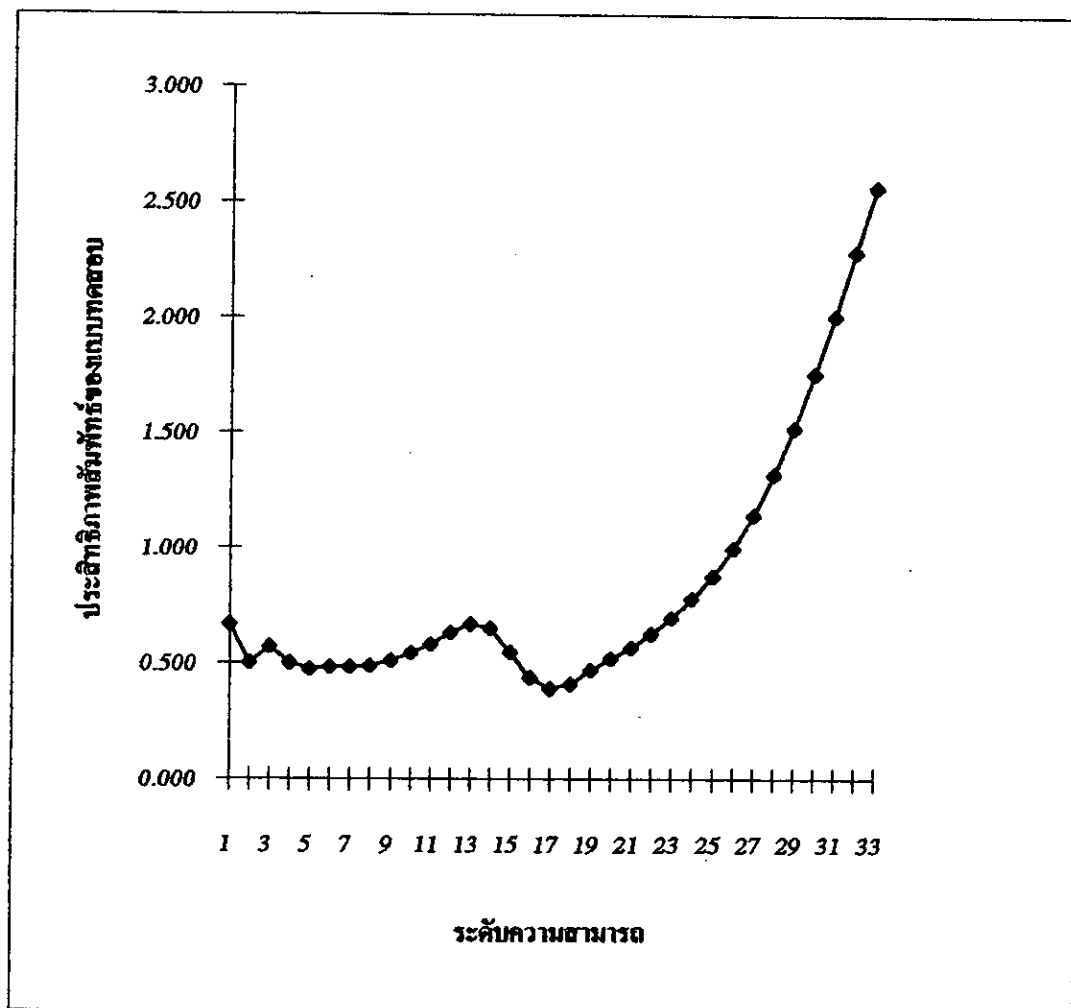


ชุด 1 AX_{51}

ชุด 2 AX_{52}

ภาพประกอบ 107 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบชนิด AX แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

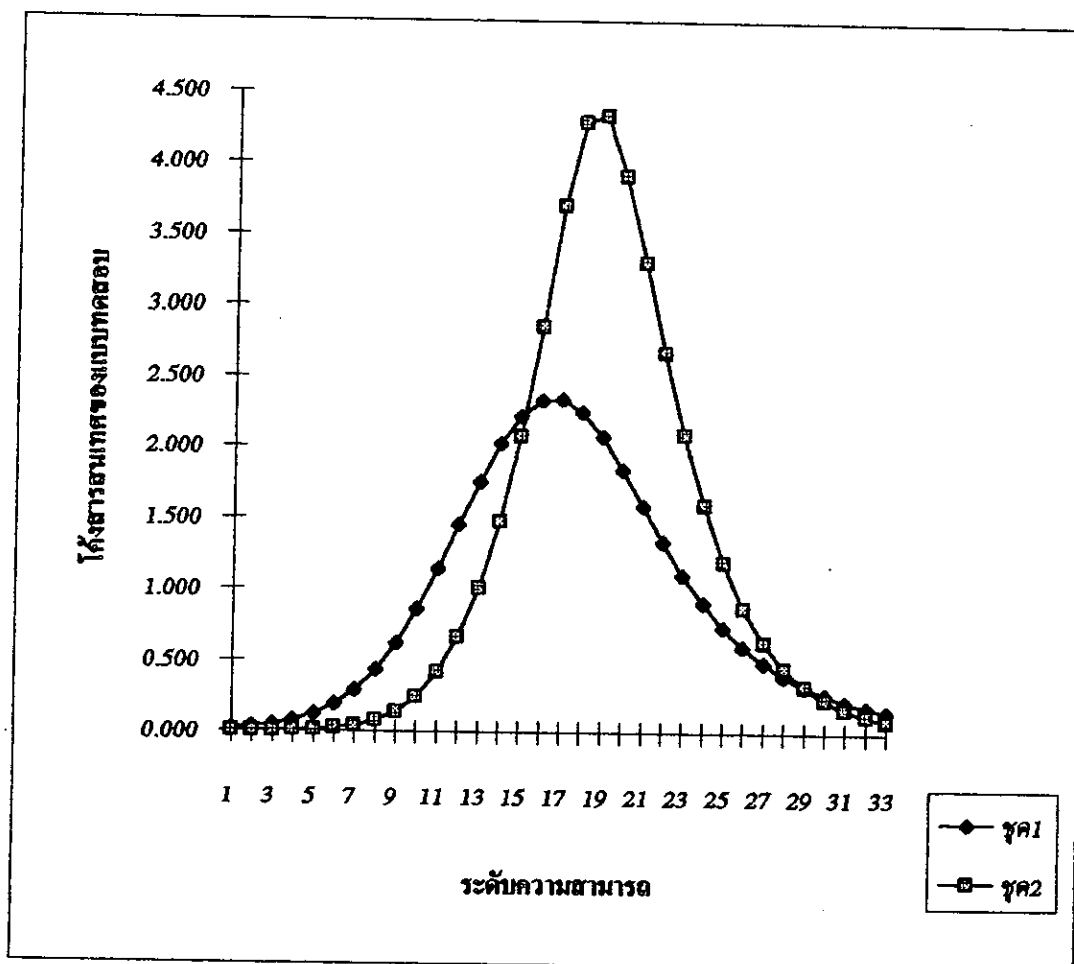
จากภาพประกอบ 107 พบว่า ค่าสารสนเทศของแบบทดสอบ ชนิด AX แบบเติมคำตอบ มีค่าฟังก์ชันสารสนเทศของแบบทดสอบ ณ ระดับความสามารถปานกลาง $\{I(\theta) = 5.617 ; \theta = 0.250\}$ สูงกว่าค่าฟังก์ชันสารสนเทศของแบบทดสอบ แบบเลือกตอบ ณ ระดับความสามารถปานกลาง $\{I(\theta) = 2.648 ; \theta = 0.750\}$



$$AX_{51} : AX_{52}$$

ภาพประกอบ 108 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด AX แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 108 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด AX แบบเลือกตอบกับแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 มีค่าต่ำสุดที่ระดับความสามารถปานกลาง $\{RE(AX_{51} : AX_{52}) = 0.388 ; e = 0.000\}$ และมีค่าสูงสุดที่ระดับความสามารถสูง $\{RE(AX_{51} : AX_{52}) = 2.566 ; e = 4.000\}$

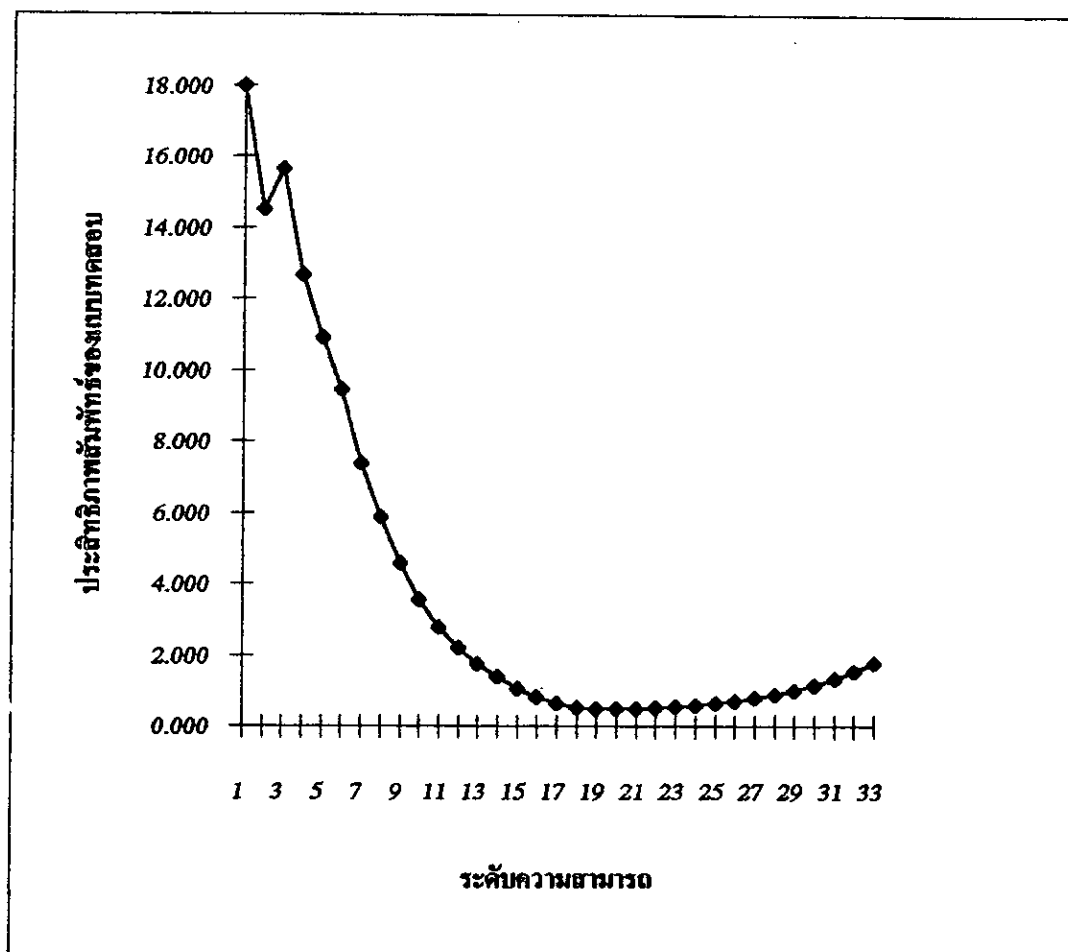


ชุด 1 CV_{S1}

ชุด 2 CV_{S2}

ภาพประกอบ 109 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบชนิด CV แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

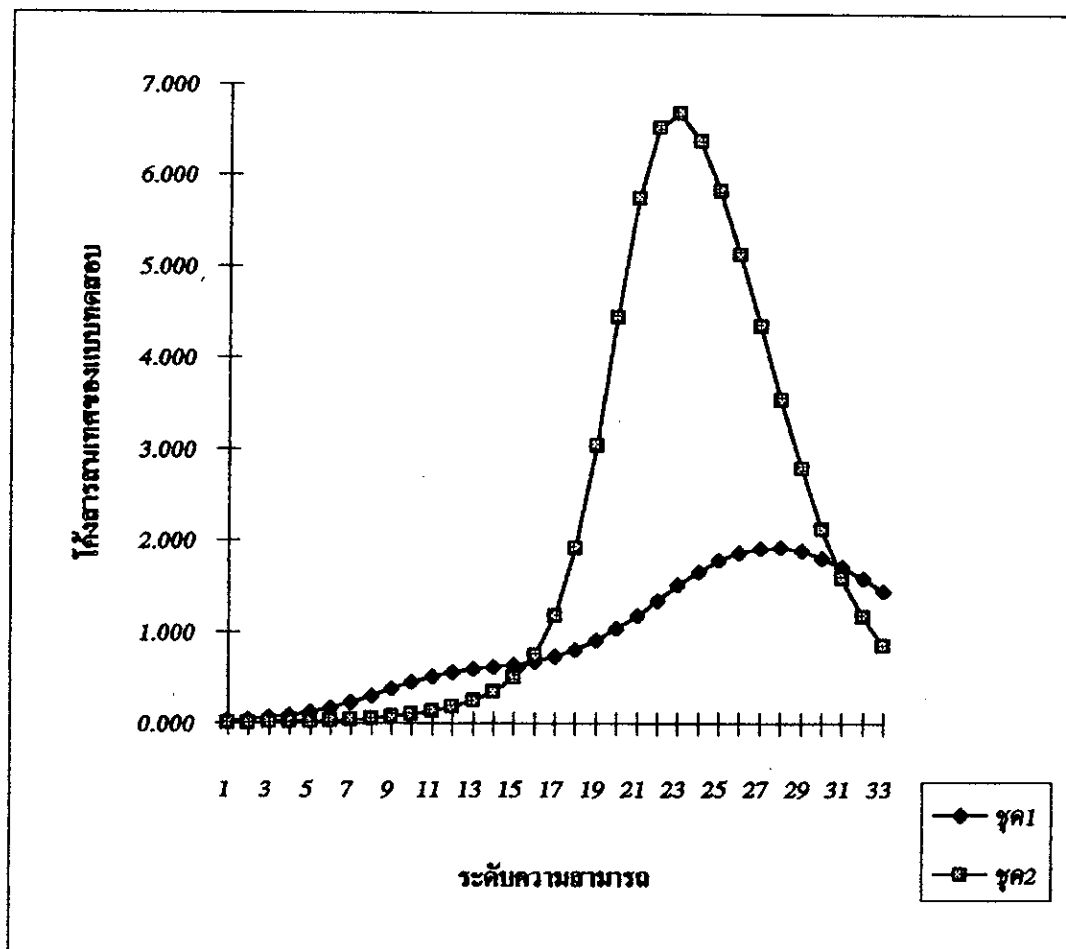
จากภาพประกอบ 109 พบว่า ค่าสารสนเทศของแบบทดสอบ ชนิด CV แบบเติมคำตอบ มีค่าฟังก์ชันสารสนเทศของแบบทดสอบ ณ ระดับความสามารถปานกลาง $\{I(\theta) = 4.336 ; \theta = 0.500\}$ สูงกว่าค่าฟังก์ชันสารสนเทศของแบบทดสอบ แบบเลือกตอบ ณ ระดับความสามารถปานกลาง $\{I(\theta) = 2.338 ; \theta = 0.000\}$



$$CV_{51} : CV_{52}$$

ภาพประกอบ 110 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด CV แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 110 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด CV แบบเลือกตอบกับแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 มีค่าต่ำสุดที่ระดับความสามารถปานกลาง $\{RE(CV_{51} : CV_{52}) = 0.471 ; e = 0.750\}$ และมีค่าสูงสุดที่ระดับความสามารถต่ำ $\{RE(CV_{51} : CV_{52}) = 18.000 ; e = -4.000\}$

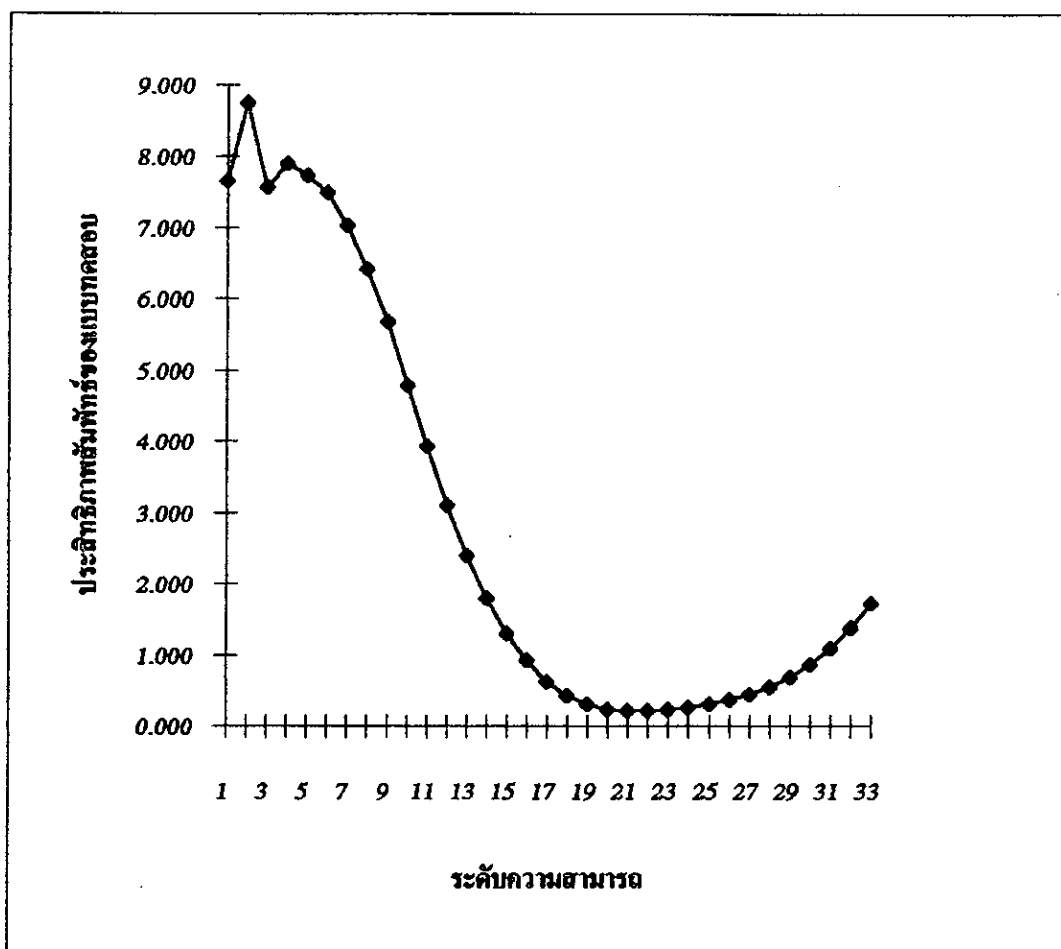


ชุด 1 LR₅₁

ชุด 2 LR₅₂

ภาพประกอบ 111 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบชนิด LR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

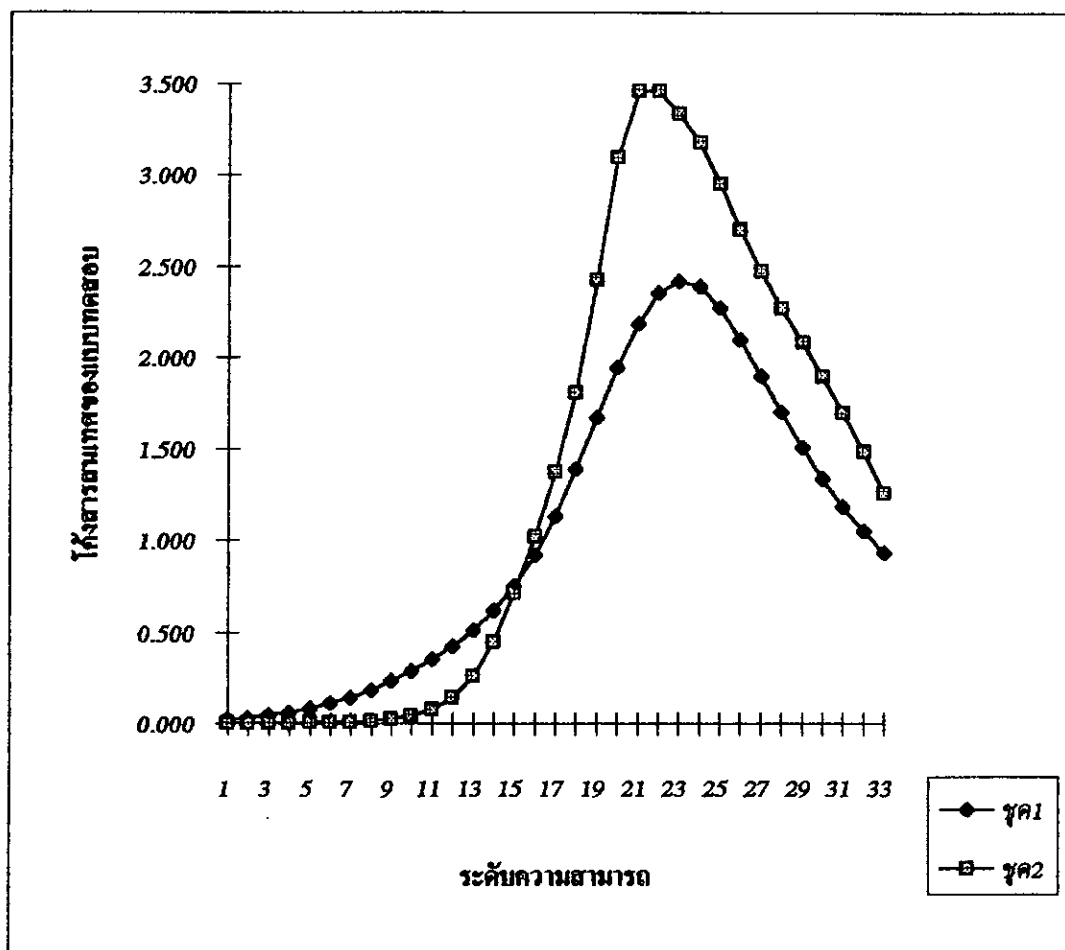
จากภาพประกอบ 111 พบว่า ค่าสารสนเทศของแบบทดสอบ ชนิด LR แบบเติมคำตอบ มีค่าฟังก์ชันสารสนเทศของแบบทดสอบ ณ ระดับความสามารถสูง $\{I(\theta) = 6.680 ; \theta = 1.500\}$ สูงกว่าค่าฟังก์ชันสารสนเทศของแบบทดสอบ แบบเลือกตอบ ณ ระดับความสามารถสูง $\{I(\theta) = 1.916 ; \theta = 2.750\}$



$$LR_{s_1} : LR_{s_2}$$

ภาพประกอบ 112 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด LR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 112 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด LR แบบเลือกตอบกับแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 มีค่าต่ำสุดที่ระดับความสามารถปานกลาง $\{RE(LR_{s_1} : LR_{s_2}) = 0.205 ; \sigma = 1.000 \text{ และ } 1.250\}$ และมีค่าสูงสุดที่ระดับความสามารถต่ำ $\{RE(LR_{s_1} : LR_{s_2}) = 8.750 ; \sigma = -3.750\}$

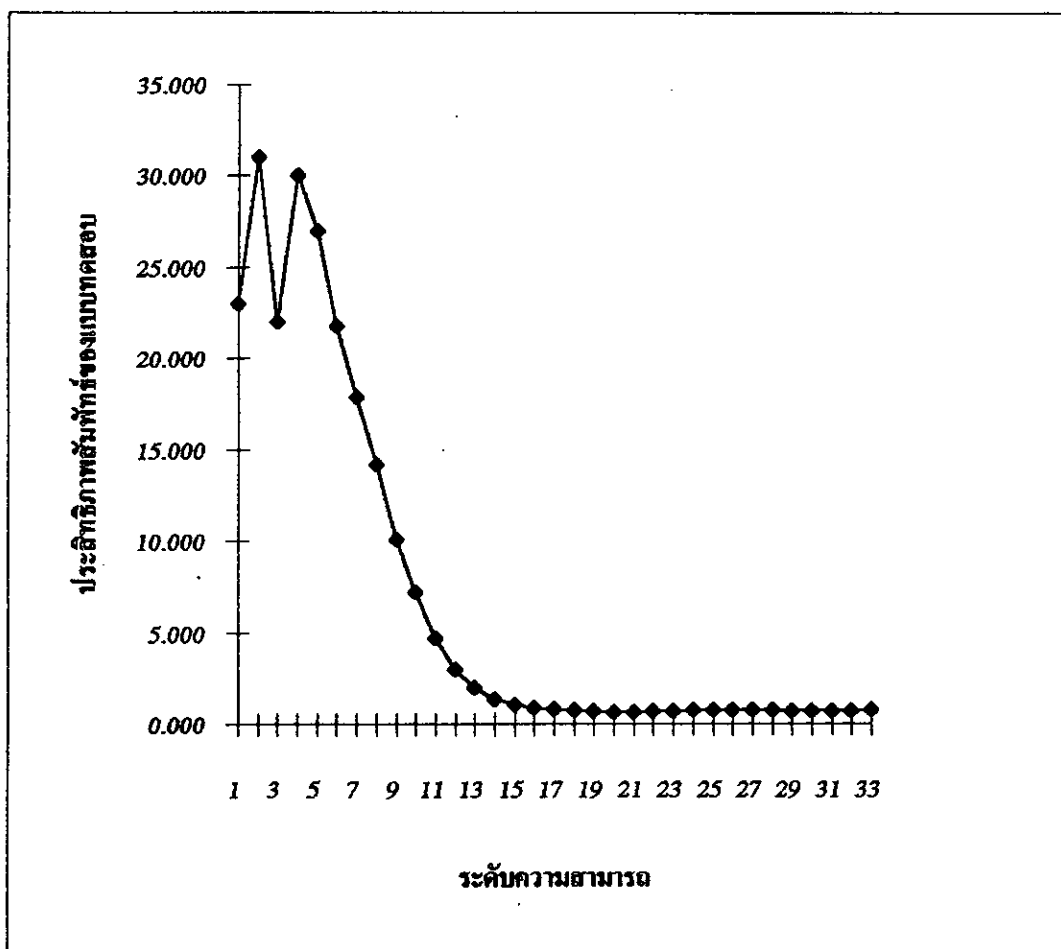


ชุด 1 NLR_{51}

ชุด 2 NLR_{52}

ภาพประกอบ 113 เปรียบเทียบค่าสารสนเทศของแบบทดสอบชนิด NLR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

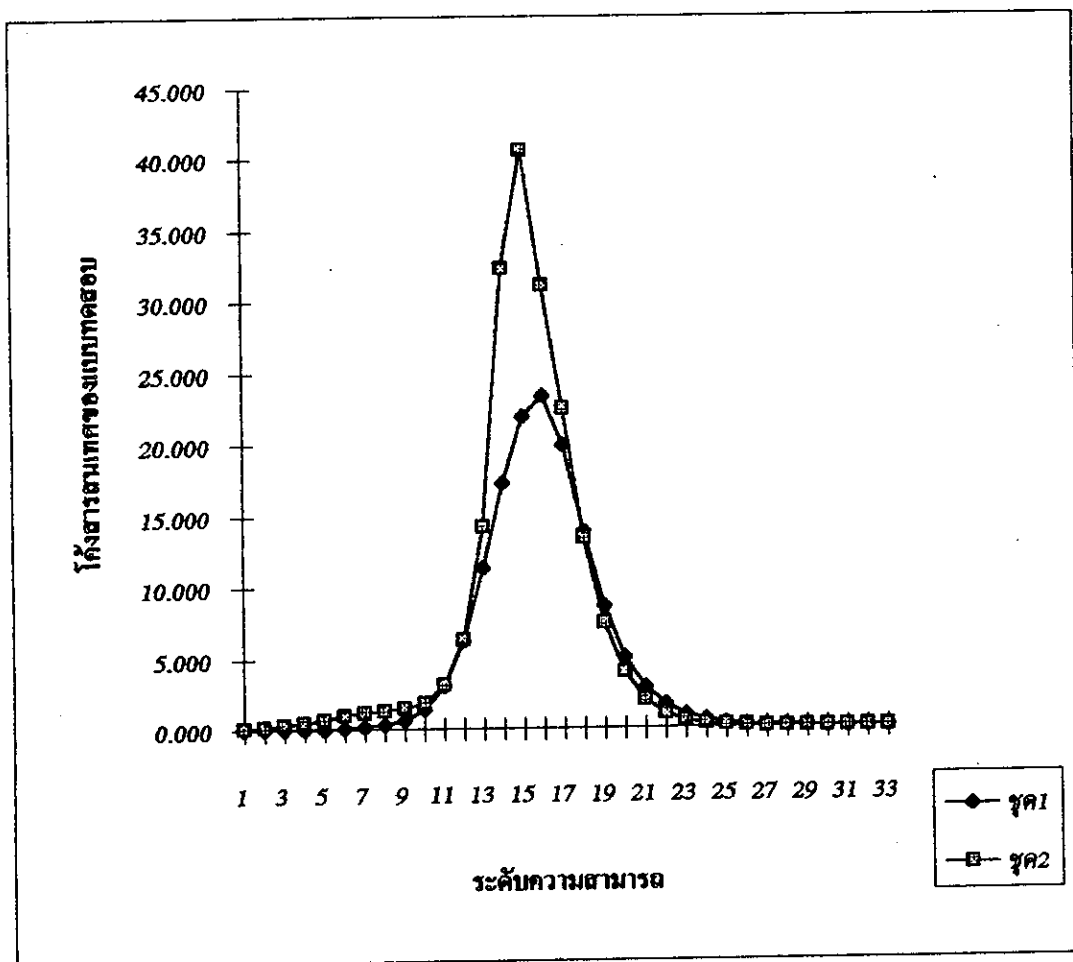
จากภาพประกอบ 113 พบว่า ค่าสารสนเทศของแบบทดสอบ ชนิด NLR แบบเติมคำตอบ มีค่าฟังก์ชันสารสนเทศของแบบทดสอบ ณ ระดับความสามารถปานกลาง $\{I(\theta) = 3.460 ; \theta = 1.250\}$ สูงกว่าค่าฟังก์ชันสารสนเทศของแบบทดสอบ แบบเลือกตอบ ณ ระดับความสามารถสูง $\{I(\theta) = 2.419 ; \theta = 1.500\}$



$$NLR_{s_1} : NLR_{s_2}$$

ภาพประกอบ 114 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด NLR แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 114 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด NLR แบบเลือกตอบกับแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 มีค่าต่ำสุดที่ระดับความสามารถปานกลาง $\{RE(NLR_{s_1} : NLR_{s_2}) = 0.628 ; \sigma = 0.750\}$ และมีค่าสูงสุดที่ระดับความสามารถต่ำ $\{RE(NLR_{s_1} : NLR_{s_2}) = 31.000 ; \sigma = -3.750\}$

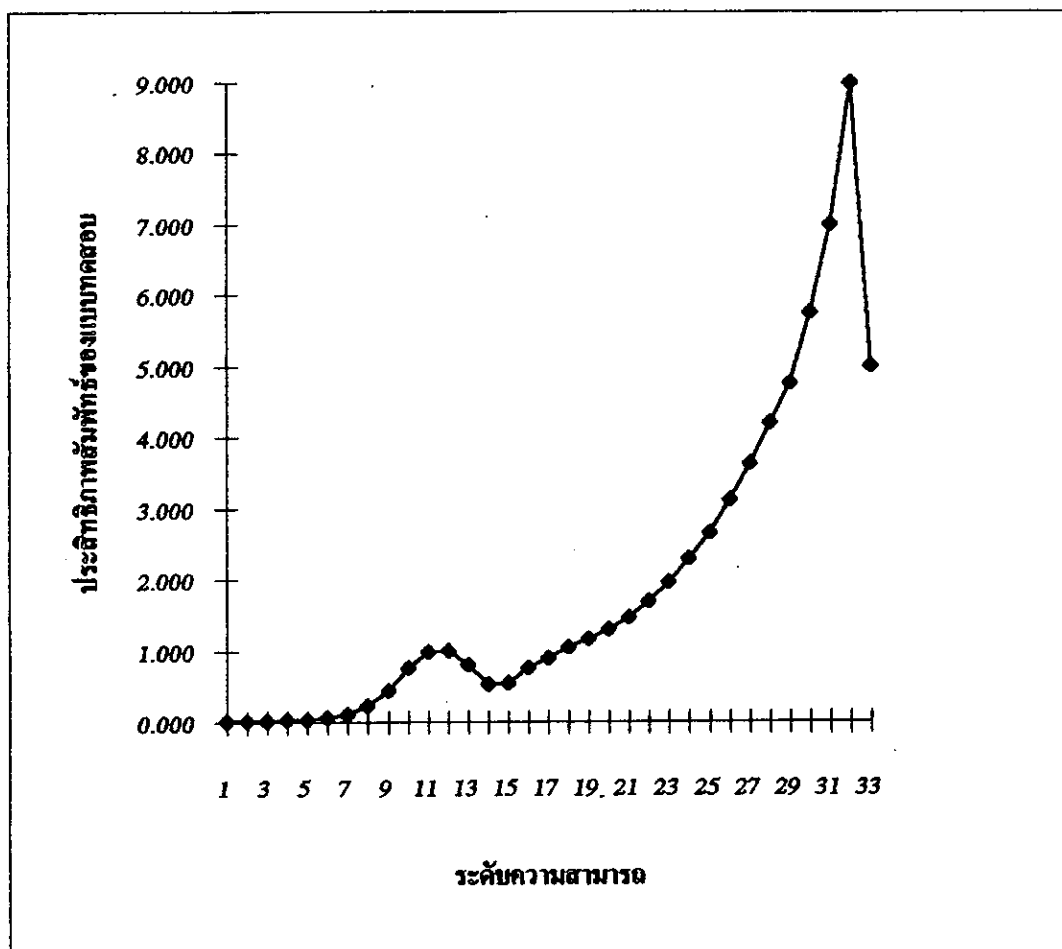


ชุด 1 PI₅₁

ชุด 2 PI₅₂

ภาพประกอบ 115 เปรียบเทียบค่าสารสนเทศของแบบทดสอบชนิด PI แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 115 พบว่า ค่าสารสนเทศของแบบทดสอบ ชนิด PI แบบเติมคำตอบ มีค่าฟังก์ชันสารสนเทศของแบบทดสอบ ณ ระดับความสามารถปานกลาง $\{I(\theta) = 40.561 ; \theta = -0.500\}$ สูงกว่าค่าฟังก์ชันสารสนเทศของแบบทดสอบ แบบเลือกตอบ ณ ระดับความสามารถปานกลาง $\{I(\theta) = 23.316 ; \theta = -0.250\}$



$$PI_{s1} : PI_{s2}$$

ภาพประกอบ 116 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด PI แบบเลือกตอบ และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

จากภาพประกอบ 116 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ชนิด PI แบบเลือกตอบกับแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 มีค่าต่ำสุดที่ระดับความสามารถต่ำ $\{RE(PI_{s1} : PI_{s2}) = 0.010 ; \sigma = -4.000\}$, และและมีค่าสูงสุดที่ระดับความสามารถสูง $\{RE(PI_{s1} : PI_{s2}) = 9.000 ; \sigma = 3.750\}$

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาการวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ของนักเรียนระดับชั้นมัธยมศึกษา ในประเด็นเกี่ยวกับการสร้างเครื่องมือวัด ทั้งนี้โดยการพัฒนาข้อคำถามชนิดต่าง ๆ ขึ้น 6 ชนิด จัดทำเป็น 2 รูปแบบ คือ รูปแบบเลือกตอบ และเติมคำตอบ และทำการศึกษาเปรียบเทียบประสิทธิภาพของข้อคำถามและแบบทดสอบที่ได้พัฒนาขึ้น โดยมีรายละเอียดในการดำเนินการวิจัย ดังนี้

วัตถุประสงค์ของการวิจัย

1. สร้างข้อคำถามเพื่อพัฒนารูปแบบการวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ซึ่งประกอบด้วยชนิดของข้อคำถาม (Item Type) 6 ชนิด และรูปแบบข้อคำถาม (Item Format) 2 รูปแบบ
2. ศึกษาคุณภาพของข้อคำถามด้วยวิธีการ
 - 2.1 การวิเคราะห์ด้วยทฤษฎีแบบดั้งเดิม เพื่อการพัฒนาข้อคำถาม
 - 2.2 การวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบ เพื่อศึกษาคุณภาพของข้อคำถามและแบบทดสอบ

ขอบเขตของการวิจัย

1. กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 และ 5 ภาคเรียนที่ 2 ปีการศึกษา 2538 ของโรงเรียนสังกัดกรมสามัญศึกษา และสำนักงานคณะกรรมการการศึกษาเอกชน ในจังหวัดกรุงเทพมหานคร นครราชสีมา และนครปฐม จำนวน 27 โรงเรียน เป็นจำนวนห้องเรียน 185 ห้องเรียน ซึ่งได้มาโดยการสุ่มแบบแบ่งชั้น มีโรงเรียนเป็นหน่วยการสุ่ม แล้วทำการสุ่มอย่างง่าย (Simple Random Sampling) อีก โดยมีห้องเรียนเป็นหน่วยการสุ่ม กลุ่มตัวอย่างเป็น 2 กลุ่ม ดังนี้
 - 1.1 กลุ่มตัวอย่างที่ใช้ในการพัฒนาข้อคำถามจำนวน 8 โรงเรียน 69 ห้องเรียน จำนวนนักเรียน 2,954 คน
 - 1.2 กลุ่มตัวอย่างที่ใช้ศึกษาคุณภาพของข้อคำถามจำนวน 19 โรงเรียน 116 ห้องเรียน จำนวนนักเรียน 4,682 คน

2. ตัวแปรที่ใช้ในการวิจัยครั้งนี้

2.1 ตัวแปรอิสระ ได้แก่ รูปแบบของข้อคำถาม และชนิดของข้อคำถามที่พัฒนาขึ้น มีรายละเอียดดังนี้

2.1.1 รูปแบบของข้อคำถาม

2.1.1.1 แบบเลือกตอบ

2.1.1.2 แบบเติมคำตอบ

2.1.2 ชนิดของข้อคำถาม

2.1.2.1 เหตุผลด้านวิเคราะห์

2.1.2.2 การวิเคราะห์คำอธิบาย

2.1.2.3 ความเห็นตรงกันข้าม

2.1.2.4 เหตุผลทางตรรก

2.1.2.5 เหตุผลทางตรรกเกี่ยวกับจำนวน

2.1.2.6 การระบุรูปแบบ

2.2 ตัวแปรตาม ได้แก่

2.2.1 ความเชื่อมั่น

2.2.2 ความเที่ยงตรง

2.2.3 ค่าฟังก์ชันสารสนเทศของแบบทดสอบ

2.2.4 ความคลาดเคลื่อนในการประมาณค่าความสามารถ ($SE(e)$)

ณ ระดับความสามารถต่าง ๆ

2.2.5 ค่าประสิทธิภาพสัมพัทธ์

2.2.6 คุณภาพของข้อคำถามรายข้อ คือ ค่าความยาก (p) ค่าอำนาจจำแนก (r) และค่าพารามิเตอร์ของข้อคำถาม คือ ค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าการเดา (c)

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ แบ่งการศึกษาเป็น 2 ขั้นตอน คือ ขั้นตอนที่หนึ่ง เป็นการสร้างและพัฒนาข้อคำถาม ขั้นตอนที่สอง เป็นขั้นตอนการศึกษาคุณภาพของข้อคำถาม

1. การดำเนินการสร้างและการพัฒนาข้อคำถามวัดความสามารถด้านเหตุผลเชิงวิเคราะห์

การดำเนินการสร้าง และพัฒนาข้อคำถามวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ผู้วิจัยดำเนินการตามลำดับขั้น ดังนี้

1.1 ศึกษา ทฤษฎี แนวคิด งานวิจัย และเอกสารที่เกี่ยวข้องกับการคิดความสามารถด้านเหตุผล และการวัด เพื่อนิยามโครงสร้าง (Construct) และกำหนดทักษะ ซึ่งเป็นคุณสมบัติ (Attribute) ที่สำคัญของโครงสร้างที่ต้องการวัด

1.2 ตรวจสอบความเที่ยงตรงของโครงสร้างที่ต้องการวัดโดยผู้เชี่ยวชาญ 3 ท่าน ตามกระบวนการเชิงเหตุผล (Logical Analysis)

1.3 วิเคราะห์พฤติกรรมข้อบ่งชี้ของแต่ละทักษะย่อย และตรวจสอบความเที่ยงตรง โดยผู้เชี่ยวชาญเช่นกัน

1.4 กำหนดลักษณะงาน ขอบเขตของงาน ซึ่งคือลักษณะของข้อคำถามที่จะใช้นำเสนอผู้เชี่ยวชาญพิจารณา จากนั้นปรับปรุงตามข้อเสนอแนะของผู้เชี่ยวชาญ

1.5 สร้างข้อคำถามที่มีลักษณะตามกำหนด 6 ชนิด คือ เหตุผลทางตรรก การวิเคราะห์ข้อถกเถียง ความเห็นตรงกันข้าม เหตุผลด้านวิเคราะห์ เหตุผลทางตรรกเกี่ยวกับจำนวน และการระบุรูปแบบ โดยจัดทำเป็น 2 รูปแบบ คือ แบบเลือกตอบ และแบบเติมคำตอบ

1.6 ตรวจสอบความเที่ยงตรงของข้อคำถามที่สร้างขึ้นในเชิงเหตุผล (Logical Analysis) และเชิงประจักษ์ (Empirical Analysis) เพื่อการนิยามปรับปรุงแก้ไขให้ข้อคำถามมีความเป็นไปได้ในการตอบสนองของผู้สอบ ความชัดเจนของข้อคำถาม สถานการณ์ คำชี้แจง และเวลาที่ต้องใช้ในการตอบ

1.7 พิจารณาปรับปรุงข้อคำถามให้ถูกต้องและเหมาะสม จากนั้นจัดฉบับเพื่อเตรียมไปทดลองใช้

1.8 นำแบบทดสอบทั้ง 12 ฉบับ ไปทดสอบกับนักเรียนชั้นมัธยมศึกษาปีที่ 3 และปีที่ 5 ซึ่งเป็นกลุ่มเป้าหมาย จากนั้นนำผลการสอบในส่วนที่เป็นแบบเติมคำตอบมาตรวจให้คะแนน

1.9 ทำการวิเคราะห์คุณภาพของข้อคำถาม ข้อมูลผลการสอบ ฉบับที่เป็นแบบเลือกตอบและแบบเติมคำตอบ เพื่อหาค่าความยาก ค่าอำนาจจำแนก และค่าความเชื่อมั่น

1.10 พิจารณาปรับปรุงข้อคำถามจากข้อมูลที่ได้ กรณีที่คิดว่าสามารถปรับปรุงแก้ไขให้ดีขึ้นได้ และมีบางข้อจำเป็นต้องตัดออก

2. การศึกษาคุณภาพของข้อคำถาม

ขั้นตอนนี้ ผู้วิจัยได้นำข้อคำถามที่ได้พิจารณาปรับปรุงแก้ไขแล้วมาจัดฉบับใหม่ นำไปทดสอบกับกลุ่มตัวอย่างเป้าหมาย นำมาตรวจให้คะแนน จากนั้นวิเคราะห์ผลการสอบด้วยทฤษฎีการวิเคราะห์แบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ

ทั้งนี้ ถือว่าการวิเคราะห์ตามวิธีทฤษฎีการตอบสนองข้อสอบ เป็นวิธีการหลักของการศึกษาคุณภาพของข้อคำถามสำหรับขั้นตอนนี้ เนื่องจากลักษณะเด่นเกี่ยวกับความไม่แปรเปลี่ยนของค่าพารามิเตอร์ ข้อมูลเกี่ยวกับโอกาสในการตอบสนองได้ถูกต้องของผู้สอบในแต่ละระดับความสามารถ จะทำให้เห็นภาพการทำหน้าที่ของข้อคำถาม ซึ่งผู้วิจัยดำเนินการตามขั้นตอน ดังนี้

2.1 วิเคราะห์องค์ประกอบ เพื่อตรวจสอบความเป็นมิติเดียวของแบบทดสอบแต่ละฉบับ ซึ่งพบว่า มีความเป็นมิติเดียว จากนั้นทำการวิเคราะห์เพื่อหาคุณภาพของข้อคำถามรายข้อ ได้แก่ ค่าความยาก ค่าอำนาจจำแนก และค่าการเดา

2.2 ข้อคำถามที่ไม่เหมาะสมกับโมเดลจะถูกตัดออกไป จากนั้น คำนวณค่าฟังก์ชันสารสนเทศของข้อสอบ (TIF) ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ (SE(๑)) ในแต่ละระดับความสามารถ

2.3 คำนวณค่าประสิทธิภาพสัมพัทธ์ เพื่อเปรียบเทียบคุณภาพของแบบทดสอบแต่ละชนิด ระหว่างรูปแบบเลือกตอบและรูปแบบเติมคำตอบ

สรุปผลการวิจัย

1. ผลการวิเคราะห์เพื่อการพัฒนาข้อคำถาม ได้ผลดังนี้

1.1 ค่าเฉลี่ยของคะแนนที่ได้จากการสอบแบบทดสอบแต่ละฉบับ พบว่า ค่าเฉลี่ยของคะแนนที่ได้จากการสอบแบบทดสอบ ชนิด AR และ PI แบบเติมคำตอบ มีค่าสูงกว่าแบบเลือกตอบ ส่วนชนิด AX และ LR แบบเลือกตอบ มีค่าสูงกว่าแบบเติมคำตอบ โดยผลดังกล่าวมีความสอดคล้องกัน ทั้งนักเรียนชั้นมัธยมศึกษาปีที่ 3 และนักเรียนชั้นมัธยมศึกษาปีที่ 5 สำหรับชนิด CV กับ NLR ให้ผลไม่สอดคล้องกัน ทั้งนี้ในชั้นมัธยมศึกษาปีที่ 3 ค่าคะแนนเฉลี่ยของแบบทดสอบ CV แบบเติมคำตอบจะสูงกว่าแบบเลือกตอบ ขณะที่ชั้นมัธยมศึกษาปีที่ 5 ชนิด CV แบบเลือกตอบจะมีคะแนนเฉลี่ยสูงกว่าแบบเติมคำตอบ และสำหรับแบบทดสอบชนิด NLR นั้น แบบเติมคำตอบจะมีค่าเฉลี่ยต่ำกว่าแบบเลือกตอบ ในระดับชั้นมัธยมศึกษาปีที่ 3 แต่ในทางกลับกัน ระดับชั้นมัธยมศึกษาปีที่ 5 ค่าคะแนนเฉลี่ยของคะแนนจากแบบทดสอบชนิด NLR

ทางกลับกัน ระดับชั้นมัธยมศึกษาปีที่ 5 ค่าคะแนนเฉลี่ยของคะแนนจากแบบทดสอบชนิด NLR แบบเติมคำตอบจะสูงกว่าแบบเลือกตอบ

1.2 ค่าเฉลี่ยสัดส่วนการตอบถูก และค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบ จากการเปรียบเทียบค่าเฉลี่ยสัดส่วนการตอบถูกของแบบทดสอบแต่ละชนิดทั้ง 2 รูปแบบ พบว่า ค่าเฉลี่ยสัดส่วนการตอบถูกของแบบทดสอบชนิด AR, AX, CV, NLR แสดงความแตกต่างกันไม่มากนักระหว่างแบบทดสอบรูปแบบเลือกตอบและเติมคำตอบ ซึ่งทั้งระดับชั้น ม.3 และ ม.5 ให้ผลสอดคล้องกัน แต่สำหรับชนิด LR มีความแตกต่างกันมากระหว่างค่าเฉลี่ยสัดส่วนการตอบถูกของแบบทดสอบ 2 รูปแบบ สำหรับชนิด PI ให้ผลไม่สอดคล้องกัน ชั้น ม.3 แบบเติมคำตอบจะแสดงสัดส่วนการตอบถูกสูงกว่าค่อนข้างมาก ขณะที่ชั้น ม.5 มีค่าที่เกือบจะไม่แตกต่างกันระหว่างแบบทดสอบ 2 รูปแบบ

ในส่วนของค่าอำนาจจำแนกเฉลี่ย ระดับชั้น ม.3 แสดงค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบ รูปแบบเติมคำตอบทุกชนิดสูงกว่าแบบเลือกตอบ เช่นเดียวกับชั้น ม.5 ค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบรูปแบบเติมคำตอบสูงกว่าแบบเลือกตอบเกือบทุกชนิด ยกเว้น ชนิด AR และ PI

ดังนั้น อาจกล่าวโดยภาพรวมได้ว่า ค่าเฉลี่ยสัดส่วนการตอบถูกของแบบทดสอบชนิดต่าง ๆ แสดงความแตกต่างกันไม่มากนักระหว่างแบบทดสอบรูปแบบเลือกตอบ และเติมคำตอบ ในขณะที่ค่าเฉลี่ยอำนาจจำแนกของข้อคำถาม รูปแบบเติมคำตอบสูงกว่าแบบเลือกตอบเกือบทุกชนิดของข้อคำถาม และถือได้ว่าค่าอำนาจจำแนกของข้อคำถามผ่านเกณฑ์การพิจารณาทั้ง 2 รูปแบบ

1.3 คุณภาพด้านความเชื่อมั่นของแบบทดสอบ

จากการศึกษาคุณภาพด้านความเชื่อมั่นของแบบทดสอบทุกชนิด พบว่า แบบทดสอบรูปแบบเติมคำตอบให้ค่าความเชื่อมั่นสูงกว่ารูปแบบเลือกตอบทุกชนิด ยกเว้น ชนิด PI ทั้งนี้ ข้อมูลที่ได้จากนักเรียนระดับชั้น ม.3 และ ม.5 ให้ผลที่สอดคล้องกัน

1.4 ผลการพิจารณาปรับปรุงข้อคำถาม

จากผลการวิเคราะห์คุณภาพของข้อคำถามรายชื่อ แบบทดสอบชนิด LR และ NLR ได้รับการพิจารณาปรับปรุงประเด็นคำถามและตัวเลือกมากที่สุด ส่วนชนิด CV มีการพิจารณาคัดข้อคำถามบางข้อออกไป เนื่องจากสถานการณ์ที่กำหนดให้เพื่อการตอบคำถามไม่เหมาะสม

2. ผลการศึกษาคูณภาพของข้อคำถามและแบบทดสอบ

2.1 ผลการวิเคราะห์คุณภาพของข้อคำถาม จากข้อมูลที่ได้จากนักเรียน ชั้น ม.3 และชั้น ม.5 โดยการวิเคราะห์ตามทฤษฎีแบบดั้งเดิม ได้ผลดังนี้

2.1.1 ค่าเฉลี่ยสัดส่วนการตอบถูกและค่าอำนาจจำแนกเฉลี่ย

การวิเคราะห์ค่าเฉลี่ยสัดส่วนการตอบถูกของแบบทดสอบแต่ละชนิด พบว่า รูปแบบคำถามแบบเลือกตอบของแบบทดสอบชนิด AX, LR และ NLR มีค่าเฉลี่ยสัดส่วนการตอบถูกสูงกว่าแบบเติมคำตอบ ขณะที่ชนิด AR, CV และ PI แบบเติมคำตอบมีค่าเฉลี่ยสัดส่วนการตอบถูกสูงกว่าแบบเลือกตอบ ทั้งนี้โดยข้อมูลที่ได้จากนักเรียนชั้น ม.3 และ ม.5 มีความสอดคล้องกัน

ดังนั้น อาจกล่าวสรุปในภาพรวมได้ว่า สำหรับแบบทดสอบชนิด AX, LR และ NLR ข้อคำถามรูปแบบเลือกตอบง่ายกว่ารูปแบบเติมคำตอบ ขณะที่ชนิด AR, CV และ PI ข้อคำถามรูปแบบเติมคำตอบง่ายกว่ารูปแบบเลือกตอบ

สำหรับค่าอำนาจจำแนกเฉลี่ยนั้น พบว่า แบบทดสอบชนิด AR, AX, CV, LR และ PI มีค่าอำนาจจำแนกเฉลี่ยของข้อสอบรูปแบบเติมคำตอบสูงกว่าแบบเลือกตอบ ส่วนชนิด NLR รูปแบบเลือกตอบมีค่าเฉลี่ยอำนาจจำแนกของข้อสอบสูงกว่าแบบเติมคำตอบ ซึ่งข้อมูลที่ได้จากนักเรียนชั้น ม.3 และ ม.5 ให้ผลที่สอดคล้องกัน จึงอาจกล่าวสรุปโดยรวมได้ว่า ข้อคำถามรูปแบบเติมคำตอบจะให้ค่าอำนาจจำแนกสูงกว่าข้อสอบรูปแบบเลือกตอบในทุกชนิดของข้อคำถาม ยกเว้น ชนิด NLR

2.1.2 ค่าความเชื่อมั่นของแบบทดสอบ

ผลการวิเคราะห์ความเชื่อมั่นจากการทดลองครั้งที่ 2 พบว่า แบบทดสอบรูปแบบเติมคำตอบทุกชนิดมีความเชื่อมั่นสูงกว่ารูปแบบเลือกตอบ ยกเว้น ชนิด NLR ทั้งนี้ข้อมูลที่ได้จากนักเรียนชั้น ม.3 และ ม.5 ให้ผลที่สอดคล้องกัน ซึ่งทำให้กล่าวสรุปโดยรวมได้ว่า แบบทดสอบรูปแบบเติมคำตอบมีความเชื่อมั่นสูงกว่าแบบเลือกตอบ

2.2 ผลการวิเคราะห์คุณภาพของข้อคำถามและแบบทดสอบ ด้วยวิธีวิเคราะห์ตามทฤษฎีตอบสนองข้อสอบ

2.2.1 ผลการศึกษาค่าพารามิเตอร์ของข้อคำถาม

2.2.1.1 ค่าอำนาจจำแนก พบว่า แบบทดสอบรูปแบบเติมคำตอบเกือบทุกชนิดข้อคำถาม มีค่าอำนาจจำแนกเฉลี่ยสูงกว่ารูปแบบเลือกตอบ และข้อมูลที่ได้จากชั้น ม.3 และ ม.5 ให้ผลที่สอดคล้องกัน ยกเว้น PI ชั้น ม.3 ที่รูปแบบเลือกตอบ

มีค่าเฉลี่ยอำนาจจำแนกสูงกว่ารูปแบบเติมคำตอบ และในระดับชั้น ม.5 ค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบชนิด AR มีค่าเท่ากันทั้ง 2 รูปแบบ ทั้งนี้อาจกล่าวได้ว่า แบบทดสอบรูปแบบเติมคำตอบมีค่าอำนาจจำแนกเฉลี่ยสูงกว่าแบบเลือกตอบ

2.2.1.2 ค่าความยาก พบว่า แบบทดสอบรูปแบบเลือกตอบเกือบทุกชนิดมีค่าความยากเฉลี่ยสูงกว่ารูปแบบเติมคำตอบ โดยข้อมูลที่ได้จากชั้น ม.3 และ ม.5 ให้ผลที่สอดคล้องกัน ยกเว้นชนิด NLR ของชั้น ม.5 ที่รูปแบบเติมคำตอบมีค่าความยากเฉลี่ยสูงกว่าแบบเลือกตอบ สำหรับชนิด CV ค่าความยากเฉลี่ยของรูปแบบเติมคำตอบสูงกว่าแบบเลือกตอบ

2.2.1.3 ค่าสัมประสิทธิ์การเดา พบว่า แบบทดสอบชนิด AR, AX แบบเลือกตอบ มีค่าสัมประสิทธิ์การเดาเฉลี่ยสูงกว่าแบบเติมคำตอบ และชนิด CV และ NLR แบบเติมคำตอบมีค่าสัมประสิทธิ์การเดาเฉลี่ยสูงกว่าแบบเลือกตอบ ทั้งนี้ผลมีความสอดคล้องกันระหว่างข้อมูลที่ได้จากนักเรียนชั้น ม.3 และชั้น ม.5 สำหรับชนิด LR และ PI รูปแบบเลือกตอบมีค่าสัมประสิทธิ์การเดาเฉลี่ยสูงกว่าแบบเติมคำตอบ ในระดับชั้น ม.3 และชั้น ม.5 รูปแบบเติมคำตอบมีค่าสัมประสิทธิ์การเดาเฉลี่ยสูงกว่า อาจกล่าวโดยภาพรวมได้ว่า แบบทดสอบแต่ละรูปแบบมีโอกาสในการเดาคำตอบได้ถูกต้องต่างกันไปตามชนิดของข้อคำถาม

2.2.2 การศึกษาค่าฟังก์ชันสารสนเทศของแบบทดสอบ และค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ

2.2.2.1 ค่าฟังก์ชันสารสนเทศของแบบทดสอบ จากการศึกษาค่าฟังก์ชันสารสนเทศของแบบทดสอบแต่ละชนิด ทั้งรูปแบบเลือกตอบและแบบเติมคำตอบ พบว่า สำหรับระดับชั้น ม.3 แบบทดสอบแบบเลือกตอบชนิด AR, AX, CV, LR, และ PI จะมีค่าฟังก์ชันสารสนเทศสูงสุดที่ระดับความสามารถปานกลาง ส่วนชนิด NLR จะมีค่าฟังก์ชันสารสนเทศสูงสุดที่ระดับความสามารถสูง ขณะที่ แบบทดสอบแบบเติมคำตอบชนิด AR, CV และ PI จะมีค่าฟังก์ชันสารสนเทศสูงสุดที่ระดับความสามารถปานกลาง สำหรับชนิด AX, LR และ NLR จะมีค่าฟังก์ชันสารสนเทศสูงสุดที่ระดับความสามารถสูง

จากกลุ่มตัวอย่างที่เป็นนักเรียน ชั้น ม.5 แบบทดสอบ แบบเลือกตอบและแบบเติมคำตอบ ชนิด AR, AX, CV, และ PI จะมีค่าฟังก์ชันสารสนเทศสูงสุด ที่ระดับความสามารถปานกลางเช่นเดียวกัน ส่วนชนิด LR ทั้ง 2 รูปแบบ และชนิด NLR แบบเลือกตอบ จะมีค่าฟังก์ชันสารสนเทศสูงสุดที่ระดับความสามารถสูง

อาจกล่าวสรุปโดยรวมได้ว่า แบบทดสอบรูปแบบเลือกตอบชนิดต่าง ๆ ส่วนใหญ่ จะให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงสุด ที่ระดับความสามารถปานกลาง ขณะที่รูปแบบเติมคำตอบจะให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงสุดที่ระดับความสามารถสูง สำหรับกลุ่มตัวอย่างที่เป็นนักเรียนชั้น ม.3 ส่วนกลุ่มตัวอย่างที่เป็นนักเรียนชั้น ม.5 แบบทดสอบทั้ง 2 รูปแบบ จะให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงสุด ที่ระดับความสามารถปานกลาง เกือบทุกชนิด เว้นแต่ชนิด LR และ NLR ที่จะให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงสุด ที่ระดับความสามารถสูง

2.2.2.2 ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ จากการศึกษาค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ ที่ระดับความสามารถต่าง ๆ ของแบบทดสอบ รูปแบบเลือกตอบและแบบเติมคำตอบ พบว่า แบบทดสอบทุกชนิดทั้งรูปแบบเลือกตอบและแบบเติมคำตอบ จะมีค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถสูงสุด ณ ระดับความสามารถต่ำสุด ($\theta = -4.000$) ยกเว้นชนิด PI แบบเติมคำตอบ มีค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถสูงสุด ณ ระดับความสามารถสูงสุด ($\theta = 4.000$)

สำหรับค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของแบบทดสอบรูปแบบเลือกตอบเกือบทุกชนิด จะมีค่าต่ำสุด ณ ระดับความสามารถปานกลาง ยกเว้นชนิด PI ที่มีค่าต่ำสุด ณ ระดับความสามารถสูง ส่วนรูปแบบเติมคำตอบชนิด AR, CV และ PI จะมีค่าต่ำสุด ณ ระดับความสามารถปานกลาง โดยชนิด AX, LR และ NLR จะมีค่าต่ำสุด ณ ระดับความสามารถสูง

กล่าวโดยรวมได้ว่า ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของแบบทดสอบทุกชนิด ทั้ง 2 รูปแบบ จะมีค่าสูงสุด ณ ระดับความสามารถต่ำ แต่จะมีค่าต่ำ ณ ระดับความสามารถปานกลาง และสูง

2.2.3 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ จากการศึกษาค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบแต่ละชนิด โดยเปรียบเทียบรูปแบบเติมคำตอบและแบบเลือกตอบ พบว่า ค่าประสิทธิภาพสัมพัทธ์ที่ได้จากนักเรียนชั้น ม.3 แบบทดสอบชนิด AR และ AX แบบเลือกตอบ จะมีประสิทธิภาพสูงกว่าแบบเติมคำตอบ ณ ระดับความสามารถสูง ส่วนแบบเติมคำตอบของแบบทดสอบชนิด AR จะมีสัมประสิทธิภาพสูงกว่าแบบเลือกตอบ ณ ระดับความสามารถต่ำ และปานกลาง ขณะที่ชนิด AX แบบเติมคำตอบจะมีประสิทธิภาพสูงกว่า ณ ระดับความสามารถสูง ชนิด CV แบบเลือกตอบจะดีกว่าแบบเติมคำตอบ ณ ระดับความสามารถต่ำ

ขณะที่แบบเติมคำจะมีประสิทธิภาพดีกว่าแบบเลือกตอบ ณ ระดับความสามารถปานกลาง แบบทดสอบชนิด LR แบบเลือกตอบจะมีประสิทธิภาพสูงกว่าแบบเติมคำตอบ ณ ระดับความสามารถต่ำ และแบบเติมคำตอบจะมีประสิทธิภาพสูงกว่าแบบเลือกตอบ ณ ระดับความสามารถสูง สำหรับชนิด NLR และ PI แบบทดสอบแบบเติมคำตอบจะมีประสิทธิภาพสูงกว่าแบบเลือกตอบ ณ ระดับความสามารถต่ำ ขณะที่แบบเลือกตอบจะมีประสิทธิภาพสูงกว่าแบบเติมคำตอบ ณ ระดับความสามารถปานกลาง

กรณีที่กลุ่มตัวอย่างเป็นนักเรียนชั้น ม.5 แบบทดสอบชนิด AR รูปแบบเติมคำตอบ จะมีประสิทธิภาพสูงกว่าแบบเลือกตอบ ณ ระดับความสามารถต่ำ ขณะที่แบบเลือกตอบจะมีประสิทธิภาพสูงกว่าแบบเติมคำตอบ ณ ระดับความสามารถสูง ชนิด AX รูปแบบเติมคำตอบ จะมีประสิทธิภาพสูงกว่าแบบเลือกตอบ ณ ระดับความสามารถปานกลาง โดยแบบเลือกตอบจะมีประสิทธิภาพสูงกว่าแบบเติมคำตอบ ณ ระดับความสามารถสูง สำหรับแบบทดสอบ ชนิด CV, LR และ NLR แบบทดสอบแบบเติมคำตอบจะมีประสิทธิภาพสูงกว่าแบบเลือกตอบ ณ ระดับความสามารถสูง

อภิปรายผล

กระบวนการพัฒนาแบบทดสอบถือว่าข้อคำถามเป็นเป้าหมายสำคัญ ทั้งนี้เพราะข้อคำถามถือว่าเป็นงานซึ่งจะทำหน้าที่เป็นสิ่งที่เร้าที่จะกระตุ้นให้ผู้ถูกสอบวัดได้แสดงความสามารถ หรือคุณลักษณะที่แท้จริงของตนเองออกมาให้ปรากฏ การวิจัยครั้งนี้ ผู้วิจัยได้ดำเนินการตามขั้นตอนมาตรฐานของการสร้างแบบทดสอบทุกประการ ทั้งนี้โดยมุ่งหวังว่าข้อคำถามหรือแบบทดสอบที่พัฒนานั้นจะมีคุณภาพเป็นที่น่าพอใจ การวิเคราะห์ข้อคำถามเป็นขั้นตอนหนึ่งซึ่งงานวิจัยครั้งนี้ได้ให้ความสำคัญค่อนข้างสูง เนื่องจากการวิเคราะห์ข้อคำถามจะช่วยสร้างความเข้าใจแบบทดสอบให้มากขึ้น การตรวจสอบข้อคำถามรายข้ออย่างใกล้ชิด จะทำให้เห็นความคลาดเคลื่อนบางประการที่มีต่อคะแนนผลการสอบ นอกจากนี้ ผลการวิเคราะห์จะช่วยชี้แนะสู่ทางในการพัฒนาการวัดลักษณะของแบบทดสอบด้วย

จากการวิเคราะห์ผลคะแนนที่ได้จากการทดสอบแบบทดสอบความสามารถด้านเหตุผลเชิงวิเคราะห์แต่ละฉบับ พบประเด็นที่น่าพิจารณาในแต่ละเรื่อง ดังต่อไปนี้

1. ค่าความยากของข้อคำถามและแบบทดสอบ พบว่า หากพิจารณาจากค่าเฉลี่ยสัดส่วนการตอบถูกโดยภาพรวม แบบทดสอบชนิด AX, LR และ NLR จะมีค่าเฉลี่ยสัดส่วนการตอบถูกค่อนข้างต่ำ ซึ่งสอดคล้องกับระดับความยากของข้อคำถามของข้อสอบชนิดดังกล่าว

เมื่อเปรียบเทียบกับชนิด AR, CV และ PI ทั้งนี้เนื่องจากลักษณะเฉพาะของข้อความ AX, LR และ NLR มีความซับซ้อนมากกว่า ผู้ตอบต้องใช้ความสามารถ (Ability) มากกว่า โดยที่แบบทดสอบชนิด (AX) มุ่งวัดความสามารถในการสร้างข้อสรุปที่ถูกต้องสอดคล้องกันในลักษณะของความสัมพันธ์เกี่ยวกับการอ้างเหตุผลหรือข้อสรุป สร้างทางเลือกหรือคำอธิบายที่เป็นไปได้จากเนื้อหาหรือสถานการณ์ที่กำหนด ส่วนแบบทดสอบชนิด (LR) ต้องการวัดทักษะการให้เหตุผลอย่างมีวิจารณญาณ การประเมินการอ้างเหตุผลจากการวิเคราะห์ข้อมูลหลักฐานที่วัดเหตุผลที่กว้างกว่าชนิด (AR) ที่มุ่งประเมินทักษะการจับหลัก และกฎ เพื่อนำไปสู่ข้อสรุปในลักษณะของการอนุมาน (Deductive) กรณีของแบบทดสอบชนิด (NLR) ผู้สอบจะต้องตอบโดยการวิเคราะห์หรือประเมินข้อค้นพบหรือข้อมูลในตาราง ต้องใช้ความสามารถในการบูรณาการเหตุผลทางตัวเลข และภาษา เพื่อนำไปสู่ทางการแก้ปัญหา ซึ่งแบบทดสอบทั้ง 3 ชนิดนี้ มีความสัมพันธ์กันค่อนข้างสูง (Emmerich and others. 1991 : 4 - 5)

นอกจากคุณลักษณะเฉพาะดังกล่าวแล้ว แบบทดสอบชนิด AX, LR และ NLR มีรูปแบบการอ้างเหตุผลในลักษณะที่วิธิตดลถือว่าการอ้างเหตุผลที่ไม่มีแบบแผน (Informal Reasoning) คือ เป็นการอ้างเหตุผลที่จะประเมินในแง่ของความถูกต้อง การเป็นที่ยอมรับความเป็นไปได้ของเหตุผลที่สนับสนุน ซึ่งการอ้างเหตุผลดังกล่าวไม่เพียงแต่ได้รับอิทธิพลหรือถูกผลักดันจากความรู้ของบุคคลเท่านั้น แต่จะได้รับอิทธิพลจากความเชื่อ เจตคติของบุคคลต่อเรื่องที่ถามด้วย (Bullock. 1994 : 4948) กรณีนี้เห็นได้ชัดเจนจากผลการตอบคำถามของผู้สอบแบบทดสอบ ชนิด NLR รูปแบบการเติมคำตอบเกี่ยวกับการติดเชื้อเอชไอวี ซึ่งผู้สอบส่วนใหญ่ให้เหตุผลสนับสนุนคำตอบหรือเงื่อนไขว่า การติดเชื้อเอชไอวีมากเนื่องมาจากการเที่ยวโสเภณีเป็นส่วนใหญ่ โดยมีได้พิจารณาประเด็นหรือเหตุผลอื่นเลย ซึ่งทำให้คำอธิบายข้อมูลสนับสนุนไม่ถูกต้องสมบูรณ์ แบบทดสอบชนิด AX ก็เช่นกัน มีการนำความคิดหรือความรู้สึกที่ผู้สอบมีอยู่ มาใช้ในการสร้างข้อสรุปหรือคำอธิบายมากกว่าเงื่อนไขหรือหลักฐานที่มีอยู่ในสถานการณ์ที่กำหนดให้

ฉะนั้น จึงเห็นว่าโดยรวมแล้ว แบบทดสอบชนิด AX, LR และ NLR รูปแบบเลือกตอบจะมีค่าเฉลี่ยสัดส่วนการตอบถูกสูงกว่าแบบเติมคำตอบ เนื่องจากแบบเลือกตอบตัวเลือกแต่ละตัวเป็นตัวเลือกที่ครอบคลุมความคิดอยู่ ขณะที่แบบเติมคำตอบผู้สอบมีอิสระที่จะแสดงเหตุผล คำอธิบาย สนับสนุนการอ้างเหตุผลหรือการสร้างของตนเองอย่างเต็มที่ ความคิดความรู้สึกของผู้สอบที่มีอยู่ จึงมีโอกาสเข้ามามีอิทธิพลต่อผลการตอบได้มาก จึงทำให้ข้อสรุปหรือการอ้างเหตุผลไม่ถูกต้อง

ส่วนแบบทดสอบชนิด AR และ PI โดยลักษณะของข้อคำถามจะมีลักษณะเป็นการอ้างเหตุผลตามรูปแบบ หลักการของตรรกวิทยา (Logic) การตอบหรือคำตอบที่ถูกต้องจึงค่อนข้างจะเป็นไปตามหลักการ และกฎ จึงทำให้ง่ายต่อการพิจารณาประเมินหรือให้เหตุผลมากกว่า

สำหรับค่าความยาก หากพิจารณาในรายละเอียดจากค่าเฉลี่ยของคะแนนที่ได้ของแบบทดสอบแต่ละฉบับย่อยหรือสัดส่วนการตอบถูก จะพบว่า นักเรียนชั้นมัธยมศึกษาปีที่ 3 และ 5 มีค่าเฉลี่ยของคะแนนจากการสอบค่อนข้างต่ำ คือ ค่าเฉลี่ยต่ำกว่าครึ่งหนึ่งของคะแนนรวมทั้งฉบับทุกฉบับ ยกเว้น ฉบับ PI ซึ่งเมื่อพิจารณาในแง่ของผู้ประเมินแล้ว ถือว่าความสามารถด้านเหตุผลเชิงวิเคราะห์ของนักเรียนทั้ง 2 ระดับชั้น ยังไม่เป็นที่น่าพอใจนัก หรืออาจกล่าวได้ว่า คุณภาพของผู้เรียนในด้านนี้ยังอยู่ในระดับต่ำกว่าเป้าหมายที่ต้องการ ข้อที่ควรได้รับการพิจารณาเกี่ยวกับผลที่ปรากฏดังกล่าว อาจเป็นได้ 2 กรณี คือ กรณีที่ชนิดของคำถามของแบบทดสอบชุดนี้มีความแตกต่างไปจากลักษณะที่ผู้เรียนเคยพบ การตอบข้อคำถามผู้สอบต้องใช้ความสามารถในการคิดหลายด้าน ซึ่งผู้สอบไม่คุ้นเคยทำให้ทำข้อสอบได้ในระดับต่ำ ส่วนอีกกรณีที่ควรได้รับความสนใจอย่างยิ่ง คือ คุณภาพการเรียนการสอนที่จะใช้ในการพัฒนาผู้เรียนให้มีคุณลักษณะดังกล่าว หากถือว่าแบบทดสอบเป็นแหล่งข้อมูลที่สำคัญสำหรับครูและผู้ปกครองที่จะช่วยหาข้อมูลเพื่อประกอบการตัดสินใจเกี่ยวกับคุณภาพของหลักสูตร ความสามารถในการสอนของครู หรือคุณภาพของโปรแกรมแล้ว ระดับความสามารถด้านเหตุผลเชิงวิเคราะห์ของนักเรียนทั้ง 2 ระดับชั้น ที่ปรากฏในระดับดังกล่าว อาจบ่งชี้ให้เห็นว่าการจัดการเรียนการสอนเพื่อให้ผู้เรียนเกิดคุณลักษณะดังกล่าวให้ผลเป็นที่ไม่น่าพอใจ หรือยังไม่ประสบผลสำเร็จ ทั้งนี้เพราะเป็นที่ยอมรับว่าความสามารถด้านเหตุผลสามารถพัฒนาให้เกิดขึ้นได้จากการเรียนการสอน ดังนั้นผู้เรียนอาจได้รับการพัฒนา ผักผ่อน เพื่อส่งเสริมทักษะการคิดอย่างมีเหตุผลน้อยเกินไป ฉะนั้น จึงควรส่งเสริมให้ผู้เรียนได้รับการฝึกให้มากกว่าที่เป็นอยู่และเป็นไปอย่างจริงจัง

2. ค่าอำนาจจำแนกของข้อสอบ พบว่า หากพิจารณาจากค่าอำนาจจำแนกเฉลี่ยโดยภาพรวม แบบทดสอบรูปแบบเดิมคำตอบจะมีค่าสูงกว่าแบบเลือกตอบทุกชนิดของข้อคำถาม ซึ่งทั้งนี้อาจเป็นเพราะรูปแบบเดิมคำตอบเปิดโอกาสให้ผู้สอบได้แสดงความสามารถทางความคิด การใช้เหตุผลอย่างเต็มที่ ผู้มีความสามารถสูงมีโอกาสดำเนินการได้แสดงความสามารถของตนได้อย่างเต็มที่ ในขณะที่ผู้มีความสามารถต่ำก็จะทำได้น้อย แต่แบบเลือกตอบถูกจำกัดด้วยกรอบความคิดตามตัวเลือก อย่างไรก็ตาม ถือได้ว่าแบบทดสอบที่ได้พัฒนากันนี้ มีคุณภาพด้านอำนาจ

จำแนกเป็นที่น่าพอใจทั้ง 2 รูปแบบ สำหรับผลการทดลองทั้ง 2 ครั้ง โดยเฉพาะครั้งที่ 2 ซึ่งหลังจากได้มีการปรับปรุงแก้ไขข้อคำถาม แบบทดสอบเกือบทุกฉบับมีค่าอำนาจจำแนกเฉลี่ยสูงขึ้น

3. ค่าความเชื่อมั่นของแบบทดสอบ พบว่า รูปแบบเดิมคำตอบมีค่าความเชื่อมั่นสูงกว่าแบบเลือกตอบทุกชนิดของข้อคำถามในระดับชั้น ม.3 และ ม.5 ยกเว้น ฉบับ PI ชั้น ม.3 AR และ PI ของชั้น ม.5 ที่รูปแบบเลือกตอบมีค่าความเชื่อมั่นสูงกว่าแบบเดิมคำตอบ ซึ่งผลที่ปรากฏเช่นนี้ ก็เป็นไปตามค่าอำนาจจำแนกของข้อคำถาม

กรณีที่ค่าความเชื่อมั่นของแบบทดสอบชนิด AR และ PI รูปแบบเลือกตอบมีค่าสูงกว่าแบบเดิมคำตอบ นั้น อาจเนื่องจากลักษณะของข้อคำถามดังกล่าวมาแล้ว ที่ว่าแบบทดสอบชนิด AR, PI มีลักษณะที่เป็นการอ้างเหตุผลตามกฎ หลักทางตรรกศาสตร์ในลักษณะของ Formal Logic ดังนั้น การอ้างเหตุผลการสรุปหรือการประเมินเหตุผลดังกล่าว จึงค่อนข้างจะเป็นไปตามแบบแผน ตามกรอบของหลักการ ซึ่งมีความไม่ต่างกันมากนักระหว่างรูปแบบที่ให้ผู้สอบตัดสินใจเลือกกับแสดงความคิดของตน

หากพิจารณาจัดลำดับค่าความเชื่อมั่นตามชนิดของข้อคำถามทั้ง 6 ชนิด จะเห็นว่า ชนิด PI และ AR จะให้ค่าความเชื่อมั่นสูงอยู่ในลำดับที่ 1 และ 2 และชนิด NLR และ LR จะให้ค่าความเชื่อมั่นต่ำ ซึ่งทั้งนี้ก็เป็นผลอันเนื่องมาจากค่าความยากค่อนข้างสูง อำนาจจำแนกจึงต่ำ ส่งผลให้ค่าความเชื่อมั่นต่ำไปด้วย ข้อที่น่าพิจารณาอีกประการหนึ่งสำหรับผลการศึกษารั้งนี้ คือ ข้อคำถามของแบบทดสอบชนิด AR และ PI มีจำนวนมากกว่าแบบทดสอบชนิดอื่น ๆ คือ มี 21 ข้อ และ 19 ข้อ ตามลำดับ ขณะที่ฉบับอื่นมีประมาณ 15 ข้อ ยกเว้น ชนิด LR ที่มี 20 ข้อ อันอาจจะเป็นผลให้ค่าความเชื่อมั่นของแบบทดสอบชนิด AR และ PI มีค่าสูงกว่าชนิดอื่น ๆ

ในแง่ความคลาดเคลื่อนของการวัด ถือได้ว่าแบบทดสอบชนิด PI มีค่อนข้างต่ำกว่าชนิดอื่น ๆ ทั้งนี้อาจเป็นเพราะชนิด PI ประเด็นคำถามหรือตัวเลือก (กรณีเป็นแบบเลือกตอบ) ค่อนข้างชัดเจนในตัวของมันเอง เนื่องจากเป็นตัวเลขทั้งหมดไม่มีภาษาเข้ามาเกี่ยวข้อง การตีความคำถาม คำตอบ จึงเป็นไปได้ตรงเป้าหมาย ความคลาดเคลื่อนของการวัดจึงต่ำกว่าชนิดอื่น เนื่องจากแหล่งความคลาดเคลื่อนจากตัวแปรที่เกิดจากแบบทดสอบ หรือตัวข้อคำถามเองแทบจะไม่มี

4. ค่าฟังก์ชันสารสนเทศของแบบทดสอบ เนื่องจากค่าฟังก์ชันสารสนเทศของแบบทดสอบฉบับหนึ่ง ๆ จะมีได้หลายค่าแตกต่างกันไปตามระดับความสามารถ ซึ่ง ณ ระดับความ

สามารถได้ ที่ค่าฟังก์ชันสารสนเทศมีค่าสูงแสดงว่า แบบทดสอบฉบับนั้นมีความแม่นยำในการประมาณค่าความสามารถที่ระดับนั้นสูง แต่ถ้า ณ ระดับความสามารถที่ค่าฟังก์ชันสารสนเทศของแบบทดสอบมีค่าต่ำแสดงว่า แบบทดสอบฉบับนี้จะไม่มีความแม่นยำในการประมาณค่าความสามารถที่ระดับนั้นน้อย (Hambleton and Cook, 1977 : 84) จากผลการวิจัยครั้งนี้พบว่า โดยภาพรวมค่าฟังก์ชันสารสนเทศของแบบทดสอบ รูปแบบเลือกตอบ เกือบทุกชนิดของข้อคำถาม ยกเว้น ชนิด NLR แบบเลือกตอบให้ผลเช่นเดียวกับชนิด NLR, AX และ LR รูปแบบเติมคำตอบที่ค่าฟังก์ชันสารสนเทศมีค่าสูงสุด ณ ระดับความสามารถสูง ซึ่งสรุปได้ว่าแบบทดสอบรูปแบบเลือกตอบทุกชนิด ยกเว้น NLR และแบบเติมคำชนิด AR, CV และ PI เหมาะที่จะนำไปใช้กับผู้สอบที่มีระดับความสามารถปานกลาง ส่วนชนิด NLR, AX และ LR แบบเติมคำตอบเหมาะกับกลุ่มที่มีความสามารถระดับสูง สำหรับกลุ่มระดับความสามารถต่ำจะพบว่า มีความคลาดเคลื่อนในการประมาณค่า ($SE(e)$) สูง จึงไม่เหมาะที่จะนำไปใช้กับกลุ่มดังกล่าว และเมื่อเปรียบเทียบค่า TIF ระหว่างแบบทดสอบแต่ละฉบับ พบว่า PI และ AR มีค่า TIF สูงสุด ขณะที่ชนิด LR, NLR อยู่ลำดับท้ายสุด จากผลที่ปรากฏออกมาดังกล่าวข้างต้น คงจะเนื่องมาจากลักษณะเฉพาะของข้อคำถามแต่ละชนิดดังได้กล่าวมาแล้วในประเด็นของค่าความเชื่อมั่นของแบบทดสอบ

5. ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ พบว่า ข้อมูลระดับชั้น ม.3 และ ม.5 ให้ผลที่ค่อนข้างจะสอดคล้องกัน คือ แบบทดสอบรูปแบบเติมคำตอบเกือบทุกชนิด ยกเว้นชนิด AR และ PI จะมีคุณภาพสูงกว่าแบบเลือกตอบ ณ ระดับความสามารถสูง นั่นคือ แบบทดสอบแบบเติมคำตอบจะใช้ได้ดี หรือให้ค่าประมาณความสามารถที่แท้จริงที่แม่นยำสำหรับกลุ่มที่มีความสามารถสูงได้ดีกว่าแบบเลือกตอบ สำหรับกลุ่มความสามารถต่ำ แบบเลือกตอบจะใช้ได้เหมาะสมเป็นส่วนใหญ่ ผลดังกล่าวน่าจะเป็นไปตามผลการวิจัยทั้งหลายที่พบว่า ความแตกต่างของคะแนนที่ได้มีความเกี่ยวข้องกับรูปแบบของข้อคำถาม ซึ่งสัมพันธ์กับความสามารถในการระลึกได้ (Recall) กับการจำได้ (Recognition) การจำได้ผู้ตอบจะพิจารณาประเมินตัดสินใจจากตัวเลือกที่น่าเสนอ ไม่ต้องสร้างคำตอบออกมา แต่จะใช้ยุทธวิธีของการเปรียบเทียบและตัดออกไปเป็นพื้นฐานในการตอบคำถาม ข้อคำถามรูปแบบเลือกตอบจัดอยู่ในประเภทนี้ ส่วนแบบเติมคำตอบ ผู้สอบจำเป็นจะต้องสร้างคำตอบเอง ซึ่งลักษณะการตอบคำถามสำหรับรูปแบบเติมคำตอบนี้ กลุ่มระดับความสามารถสูงจะสามารถใช้ความคิดของตนเข้าไปอย่างเต็มที่ที่ทำให้ทำได้ดี ถ้าใช้แบบเติมคำตอบกับกลุ่มระดับความสามารถต่ำ การคิดที่ซับซ้อนอาจจะทำให้กลุ่มนี้ไม่ตอบ แต่ถ้าใช้แบบเลือกตอบผู้ตอบอาจจะพอใจกว่า เพราะมันใจ

ในคำตอบที่เลือกมากกว่าจะคิดขึ้นเอง กลุ่มระดับความสามารถต่ำจึงทำแบบเลือกตอบได้ดี ดังที่ แอมเมอร์ริช และคนอื่น ๆ (Emmerich and others. 1991 : 33) ได้ให้ความเห็นไว้เช่นกันว่า แบบทดสอบรูปแบบเลือกตอบและแบบเติมคำตอบ อาจกระตุ้นให้เกิดความแตกต่างในกระบวนการให้เหตุผล (Reasoning Process) ในสถานการณ์ของแบบเลือกตอบ ทำให้เกิดการเปรียบเทียบสถานการณ์เพื่อการตัดสินใจ ซึ่งผู้สอบรู้ว่าข้อที่ตั้งใจให้ถูกจะอยู่ในกลุ่มตัวเลือกที่ให้ ขณะที่แบบเลือกตอบ ใช่/ไม่ใช่ (Multiple-Yes/No) จะมีคำตอบที่สามารถตอบได้อย่างอิสระ มีโอกาสประเมินความถูกต้อง และหาข้อสรุปด้วยตัวของผู้อยู่สอบเอง ซึ่งให้ลักษณะที่เป็นธรรมชาติของชีวิตจริงมากกว่า แบบเลือกตอบใช่/ไม่ใช่ จะให้ความรู้สึกท้าทายทางความคิดมากกว่าแบบเลือกตอบ เชปาร์ค (Shepard. 1991 : 24) ได้แสดงความเห็นเกี่ยวกับ ผลการสอบที่ได้จากแบบทดสอบรูปแบบเลือกตอบและแบบสร้างคำตอบเอง (Constructed Response) ว่า ผู้เรียนพอใจจะสอบโดยใช้ข้อสอบแบบเลือกตอบ แม้ นักเรียนเหล่านั้นจะเห็นว่า แบบทดสอบแบบสร้างคำตอบเองเป็นดัชนีบ่งชี้ที่แม่นยำกว่า ความพอใจของนักเรียนที่มีต่อแบบเลือกตอบเกิดมาจากการรับรู้ว่าการเตรียมตัวที่น้อยกว่า และสร้างความมั่นใจในตนเองได้มากกว่า (Traub and Mac Rury. 1990) ความคิดเช่นนี้ชี้ให้เห็นว่า การวัดชนิดให้สร้างคำตอบเอง เป็นสาเหตุให้ผู้เรียนทำงานหนักขึ้น เตรียมตัวมากกว่า และเรียนหนักกว่า ซึ่งคุณลักษณะดังกล่าวผู้มีระดับความสามารถสูงจะมีเหนือกว่า

ค่าประสิทธิภาพสัมพัทธ์ ซึ่งเป็นผลที่ได้จากการเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบ 2 ฉบับ ที่วัดความสามารถเดียวกัน ณ ระดับความสามารถต่าง ๆ จากการวิจัยครั้งนี้แสดงให้เห็นว่า ณ ระดับความสามารถสูง แบบทดสอบรูปแบบเติมคำตอบให้ค่าฟังก์ชันสารสนเทศสูงกว่ารูปแบบเลือกตอบที่มีจำนวนข้อคำถามเท่ากัน แต่เนื่องจากค่าฟังก์ชันสารสนเทศของแบบทดสอบสามารถทำให้สูงขึ้นได้ โดยการเพิ่มจำนวนข้อคำถามที่มีความเทียบเท่ากับข้อคำถามที่มีอยู่ในแบบทดสอบแล้วเข้าไปในแบบทดสอบ (Hambleton and others. 1991 : 94 - 96) หรืออาจจะกล่าวได้อีกอย่างหนึ่งว่า การเพิ่มข้อคำถามเข้าไปจะทำให้ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของแบบทดสอบลดลง ซึ่งก็คือเพิ่มความแม่นยำในการประมาณค่านั่นเอง ฉะนั้น สำหรับกลุ่มความสามารถสูง หากผู้สร้างแบบทดสอบต้องการเพิ่มค่าฟังก์ชันสารสนเทศของแบบทดสอบรูปแบบเลือกตอบ ให้อยู่ในระดับที่เท่าเทียมกับแบบทดสอบรูปแบบเติมคำตอบ หรือให้มีความสามารถในการประมาณค่าความสามารถสูงสุด ณ ระดับความสามารถสูง เช่นเดียวกับแบบเติมคำตอบแล้ว จำเป็นจะต้องเพิ่มข้อคำถามที่เทียบได้กับข้อคำถามที่มีอยู่แล้วในแบบทดสอบ เพื่อให้มีความแม่นยำในการวัดเทียบ

ได้กับรูปแบบเดิมคำตอบ หรืออีกทางหนึ่งคือแบบทดสอบรูปแบบเดิมคำตอบจะต้องทำให้สั้นหรือลดจำนวนข้อลง โดยให้ค่าประมาณความสามารถ ณ ระดับนั้น มีความแม่นยำเท่ากับที่ได้จากรูปแบบเลือกตอบ ซึ่งแนวทางหลังที่ลดข้อคำถามนี้ เท่ากับเป็นการลดประสิทธิภาพในการประมาณค่าความสามารถกลุ่มระดับความสามารถสูงลง แล้วแต่ความต้องการที่จะใช้แบบทดสอบรูปแบบใด โดยการเพิ่มหรือลดข้อคำถามเพื่อขยายความยาวของแบบทดสอบให้มากขึ้นหรือเพื่อตัดให้สั้นลงนี้ ตั้งอยู่บนหลักการที่ว่าข้อคำถามหรืองานที่เพิ่มเข้าไปหรือตัดออกจะต้องเทียบกันได้ในแง่คุณภาพทางสถิติของข้อคำถามในแบบทดสอบนั้น

จะเห็นได้ว่าเมื่อใดก็ตามที่ข้อคำถามถูกเพิ่มเข้าไปในแบบทดสอบ หรือถูกตัดออกจากแบบทดสอบฉบับใดฉบับหนึ่ง ค่าฟังก์ชันสารสนเทศของแบบทดสอบ ณ ระดับ θ ต่าง ๆ จะเปลี่ยนไป ฉะนั้นในทางปฏิบัติ ข้อคำถามที่สร้างขึ้นจึงควรจัดเก็บไว้เป็นรายชื่อในลักษณะของคลังข้อสอบ (Item Bank) ข้อคำถามแต่ละข้อมีค่าพารามิเตอร์ประจำข้อ (ค่า a, b, c และค่า ICC) ซึ่งเมื่อต้องการนำแบบทดสอบไปใช้จะสามารถนำข้อคำถามเหล่านั้นไปจัดฉบับสำหรับงานทดสอบนั้น ๆ

ทั้งนี้ ลอร์ด (Hambleton and others. 1991: 100 - 101; citing Lord. 1977) ได้เสนอแนะขั้นตอนในการจัดฉบับแบบทดสอบไว้ ดังนี้

1. กำหนดฟังก์ชันสารสนเทศเป้าหมาย (Target Information Function)
2. คัดเลือกข้อความจากคลังที่มีค่าฟังก์ชันสารสนเทศของข้อคำถามที่ใช้กับฟังก์ชันสารสนเทศเป้าหมายได้
3. คำนวณค่าฟังก์ชันสารสนเทศของแบบทดสอบ สำหรับข้อคำถามที่ถูกเลือกเข้าไป
4. คัดเลือกข้อคำถามต่อไป จนกระทั่งค่าฟังก์ชันสารสนเทศมีความใกล้เคียงกับค่าฟังก์ชันสารสนเทศเป้าหมายในระดับที่น่าพอใจ

ดังนั้น ประสิทธิภาพของแบบทดสอบที่ผู้สร้างพิจารณาจัดทำขึ้นแต่ละครั้งจะแปรเปลี่ยนไปตามความต้องการ หรือกลุ่มเป้าหมายที่ต้องการใช้ ซึ่งจะเอื้อต่อการนำไปใช้ได้เหมาะสม

6. แนวคิดที่ได้จากข้อคำถามแต่ละชนิด สำหรับประเด็นเกี่ยวกับผลการตอบข้อคำถามแต่ละชนิดนี้ ผู้วิจัยขอเสนอเนื่องจากพิจารณาเห็นว่า จะเป็นประโยชน์ยิ่งต่อการศึกษาทั้งระดับความสามารถของการให้เหตุผลของผู้ตอบ และแนวคิดต่อการสร้างข้อคำถามจากการวิเคราะห์คำตอบของผู้สอบจากแบบทดสอบรูปแบบเดิมคำตอบ โดยทำในลักษณะของ

การวิเคราะห์เนื้อหา (Content Analysis) พบประเด็นที่น่าจะได้รับการพิจารณาสำหรับ
ข้อคำถามแต่ละชนิด ดังนี้

6.1 ชนิด AR ข้อคำถามชนิดนี้มุ่งประเมินทักษะในการจับหลักและกฎ เพื่อนำไป
สู่ข้อสรุป โดยพิจารณาจากเงื่อนไขที่ให้ผู้สอบส่วนใหญ่สามารถทำได้ มีบางส่วนที่ตอบตาม
ความคิดเห็น ความรู้สึก หรือนำเอาเงื่อนไขที่กำหนดให้มาตอบ มากกว่าพิจารณาให้เหตุผล
จากเงื่อนไขที่กำหนดให้

6.2 ชนิด AX ข้อคำถามชนิดนี้ ต้องการประเมินความสามารถในการสร้างข้อสรุป
ที่ถูกต้อง ความสามารถในการสร้างทางเลือกหรือคำอธิบายที่เป็นไปได้ของผู้สอบ แต่จากการ
วิเคราะห์ผลการตอบ พบว่า ผู้สอบจะตอบตามเนื้อหาของสถานการณ์ที่กำหนดในลักษณะการ
อ่านจับใจความมากกว่าการใช้เหตุผลมาตอบ และหากพิจารณาจากคะแนนผลการสอบก็พบว่า
ฉบับนี้ค่าเฉลี่ยสัดส่วนการตอบถูกต้องค่อนข้างต่ำ ค่าพารามิเตอร์ด้านความยากค่อนข้างสูง ทั้งนี้
อาจเป็นเพราะนักเรียนในระดับชั้นมัธยมศึกษาชั้นนี้ ยังมีความสามารถในการวิเคราะห์
สถานการณ์การหาเหตุผลเพื่อนำมาใช้ในการอธิบายอยู่ในระดับต่ำ

6.3 ชนิด CV ข้อคำถามชนิดนี้ มุ่งวัดความสามารถในการจับความคิดสำคัญของ
ความเห็น หรือแสดงความเห็นวิจารณ์ความเห็นที่น่าเสนอ พบว่า ผู้สอบแสดงความเห็นแบบ
สรุปมากกว่า ซึ่งไม่เป็นไปตามเป้าหมายที่ต้องการวัด อย่างไรก็ตาม ปัญหาที่พบสำหรับแบบ
ทดสอบชนิดนี้ค่อนข้างน้อย เมื่อเทียบกับบางฉบับ

6.4 ชนิด LR ข้อคำถามชนิดนี้ มุ่งวัดทักษะการให้เหตุผลอย่างมีวิจารณญาณ เช่น
การค้นหาข้อตกลงเบื้องต้น การประเมินการอ้างเหตุผล การวิเคราะห์ข้อมูลหลักฐานที่กำหนด
ให้ พบว่า ผู้สอบส่วนใหญ่ยังสรุปความ หาข้อตกลงเบื้องต้น หรือหลักการไม่เป็น นอกจากนี้
ลักษณะข้อคำถามของข้อคำถามชนิดนี้ค่อนข้างเข้าใจยากกว่าชนิดอื่น ๆ จึงทำให้ข้อคำถามมีค่า
ความยากสูง

6.5 ชนิด NLR ข้อคำถามชนิดนี้ มุ่งวัดความสามารถในการวิเคราะห์หรือประเมิน
ข้อมูลหรือข้อค้นพบที่น่าเสนอในรูปแบบตาราง แผนภูมิ หรือกราฟ ซึ่งผู้เรียนต้องใช้ความสามารถ
ในการผสมผสานเหตุผลทางตัวเลขและภาษาเพื่อการแก้ปัญหา ซึ่งพบว่า เด็กที่มีความสามารถ
สูงจะทำได้ดี และแบบทดสอบชนิดนี้จะเหมาะกับผู้มีระดับความสามารถสูง ปัญหาที่พบ คือ
ผู้สอบจะตีความ หรืออ่านแผนภูมิ ตาราง หรือกราฟ ที่ให้มากกว่าการหาเหตุผลมาสนับสนุน
ผลที่ปรากฏ

6.6 ชนิด PI ผลการวิจัยคุณภาพของข้อคำถามพบว่า ข้อคำถามชนิดนี้จะให้ค่า TIF และค่าความเชื่อมั่นสูงกว่าข้อคำถามชนิดอื่น ๆ ปัญหาจากข้อคำถามรายข้อไม่มี แต่ปัญหาที่ผู้วิจัยพบและต้องแก้ไขอย่างมาก ก็คือ ปัญหาเกี่ยวกับค่าชี้แจงการตอบ ซึ่งในการทดลองครั้งแรก พบว่า แบบเติมคำตอบ ผู้วิจัยต้องอธิบายทำความเข้าใจกับผู้สอบเอง ผู้สอบจึงจะเข้าใจและทำได้ เพราะค่าชี้แจงไม่ชัดเจน ดังนั้น ในครั้งแรกข้อมูลจากการตอบแบบทดสอบชนิด PI รูปแบบเติมคำตอบ จึงเป็นข้อมูลที่ใช้ไม่ได้มากกว่าครึ่ง ซึ่งต้องตัดทิ้งไป เมื่อได้ปรับแก้แล้ว พบว่า ครั้งที่ 2 ปัญหาในเรื่องความไม่เข้าใจวิชามีน้อยมาก

ฉะนั้น อาจกล่าวได้ว่า ผลการวิเคราะห์ข้อมูลและสภาพปัญหาที่ผู้วิจัยได้พบจากขั้นตอนการสร้างและพัฒนาข้อคำถาม จะเป็นข้อมูลที่เป็นประโยชน์ยิ่งต่อการพัฒนาการวัดผลใน ส่วนของการพัฒนาแบบทดสอบ ซึ่งผู้วิจัยจะได้ทำการวิจัยเพื่อหาข้อมูลยืนยันต่อไป

ข้อเสนอแนะ

1. การพัฒนาข้อคำถามและแบบทดสอบ ครั้งนี้ ถือได้ว่าเป็นจุดเริ่มต้นที่จะนำไปสู่การพัฒนาข้อคำถาม เพื่อวัดความสามารถในการคิดระดับสูง ซึ่งผู้วิจัยคิดว่าข้อค้นพบจากการวิจัยจะช่วยให้แนวทางในการพัฒนาข้อคำถามและแบบทดสอบ สำหรับผู้สนใจศึกษาเรื่องนี้ต่อไป ได้ ทั้งในแง่รูปแบบของข้อคำถาม ชนิดของข้อคำถาม การพิจารณาเลือกใช้ข้อคำถามแต่ละชนิด และรูปแบบ ให้เหมาะกับจุดมุ่งหมายและกลุ่มเป้าหมายที่ต้องการวัด

2. เกี่ยวกับการศึกษาวิจัยต่อ ควรมีการศึกษาวิจัยต่อเนื่องในประเด็นต่อไปนี้

2.1 ศึกษาตัวแปรต่าง ๆ ที่จะมีอิทธิพลต่อผลการตอบแบบทดสอบ รูปแบบ และชนิดต่าง ๆ อาทิ ความเหนื่อยล้า การฝึกฝน เพศ คุณสมบัติทางจิตวิทยาต่าง ๆ ที่น่าจะมีผลต่อการทำแบบทดสอบ

2.2 การศึกษาความเที่ยงตรง (Validity) ของแบบทดสอบด้านต่าง ๆ ในเชิงลึกเกี่ยวกับข้อคำถามแต่ละชนิด

2.3 กรณิรูปแบบข้อคำถามแบบเลือกตอบ แบบ 2, 3 และ 4 ตัวเลือก เพื่อจะเป็นข้อมูลประกอบการพิจารณาเลือกใช้กับกลุ่มที่มีระดับความสามารถต่ำ

3. การนำผลการวิจัยไปใช้

3.1 การพิจารณาเลือกใช้ข้อคำถาม แบบทดสอบ ข้อคำถามแต่ละชนิดที่ได้พัฒนาขึ้นเพื่อการศึกษาในครั้งนี้ หรือสามารถสะท้อนพฤติกรรมบ่งชี้ทักษะย่อยของโครงสร้างที่ต้องการวัดได้มากกว่า 1 ทักษะ ตามที่ได้แสดงไว้ในการตรวจสอบความเที่ยงตรงตาม

โครงสร้าง อย่างไรก็ตาม คุณสมบัติเฉพาะของข้อคำถามแต่ละชนิดไม่สามารถวัดทักษะข้อได้ครบทั้ง 5 ทักษะ ในทางปฏิบัติด้วยข้อจำกัดของเวลาในการสอบ หรือความสามารถในการทำข้อสอบของผู้สอบ หรืองบประมาณ อาจไม่สามารถนำแบบทดสอบฉบับย่อ ซึ่ง เป็นข้อคำถามแต่ละชนิดไปใช้ได้ทั้งหมด ผู้ใช้จำเป็นต้องพิจารณาว่าแบบทดสอบฉบับย่อใดที่สามารถวัดทักษะข้อที่เป็นเป้าหมายได้ตรงและครอบคลุมที่สุด การตีความหมายหรือสรุปผลคะแนนที่ได้ตามทักษะข้อหรือพฤติกรรมบ่งชี้ของทักษะในแบบทดสอบฉบับย่อนั้น ๆ ซึ่งจะทำได้สามารถวินิจฉัยข้อบกพร่องของผู้สอบจากผลการทำแบบทดสอบได้ค่อนข้างชัดเจน และจะเป็นประโยชน์ยิ่งต่อผู้สอนในการพิจารณาปรับปรุงการเรียนการสอน

อนึ่ง แม้ผลการศึกษารั้จะพบว่า ข้อคำถามชนิดให้ผู้สอบเติมคำตอบจะแสดงให้เห็นถึงความแม่นยำของการประมาณค่าความสามารถที่สูงกว่าแบบเลือกตอบก็ตาม แต่ผู้พัฒนาแบบทดสอบ และผู้ใช้แบบทดสอบ ควรตระหนักด้วยว่า การตัดสินใจเลือกรูปแบบ (Format) ของแบบทดสอบที่เหมาะสมตามเป้าหมายของการใช้ผลการสอบ เช่น กรณีสิ่งที่ต้องการวัดมีความซับซ้อน เช่น การให้แสดงความเห็นหรือเหตุผล รูปแบบข้อคำถามที่เอื้อต่อการแสดงความสามารถดังกล่าว จำเป็นต้องใช้รูปแบบที่มีความเป็นแบบแผน (Structure) ที่น้อยกว่า เช่น คำถามแบบปลายเปิด ขณะที่หากสิ่งที่ต้องการวัดเป็นเรื่องความรู้ความจำทั่วไป รูปแบบที่ใช้ อาจเป็นแบบแผนมากกว่า เช่น แบบเลือกตอบ (Rogers, 1995 : 656) ซึ่งตามทัศนะของโรเจอร์ถือว่า พิจารณาจากสาระของสิ่งที่ต้องการวัดเป็นสำคัญ

ลักษณะการสอบวัดเป็นองค์ประกอบหนึ่งที่จำเป็นต้องพิจารณา การประเมินในชั้นเรียน ซึ่งมีจำนวนผู้ถูกประเมินไม่มากนัก อีกทั้งครูผู้สอนเป็นผู้ประเมินผลนักเรียนด้วยตนเอง รูปแบบข้อคำถามลักษณะเติมคำตอบสามารถนำมาใช้ได้เป็นอย่างดี ทั้งนี้เพราะผู้สอนจะเป็นผู้ตรวจให้คะแนนเอง ทำให้สามารถประเมินสรุปจากผลการวัดงานลักษณะดังกล่าวได้บ่อยครั้ง จากสภาพการเรียนการสอนที่แท้จริงในลักษณะที่เรียกว่า Authentic Assessment ได้ตลอดเวลา ซึ่งจะให้ผลการประเมินที่มีความเชื่อมั่นและความเที่ยงตรงสูง นอกจากนี้ทำให้สามารถวัดความสามารถที่ซับซ้อน อาทิ การใช้เหตุผล การแก้ปัญหา การแสดงความคิดเห็นได้เป็นอย่างดีอีกประการหนึ่งด้วย

สำหรับการประเมินนักเรียนจำนวนมาก เช่น การประเมินคุณภาพการศึกษาของผู้เรียนตามหลักสูตรในระดับประเทศ การสอบคัดเลือกผู้เข้าศึกษาต่อ ข้อคำถามรูปแบบเลือกตอบจะมีความเหมาะสม ในแง่ของการดำเนินการสอบ การตรวจให้คะแนน เนื่องจากมีลักษณะเฉพาะที่ช่วยลดความคลาดเคลื่อนอันอาจเกิดจากวิธีการดำเนินการสอบ และการตรวจ

ให้คะแนน อย่างไรก็ตาม มิได้หมายความว่า กรณีผู้สอบจำนวนมากแล้ว จำเป็นต้องใช้ข้อคำถามรูปแบบเลือกตอบอย่างเดียว ทั้งนี้เพราะคุณลักษณะที่ต้องการวัดบางประการจำเป็นต้องให้ผู้ถูกวัดแสดงความสามารถของตนอย่างอิสระ ข้อคำถามรูปแบบเติมคำตอบจะเอื้อต่อการแสดงให้เห็นความสามารถดังกล่าวได้ ทางออกเพื่อลดความคลาดเคลื่อนจากการให้คะแนนอันเป็นข้อที่ได้รับการโจมตีของข้อคำถาม ลักษณะนี้ก็คือ การกำหนดเกณฑ์การให้คะแนนแบบ Rubric Scoring ซึ่งหมายถึงการให้คะแนนตามความสมบูรณ์ของคำตอบหรือขั้นตอนที่ชัดเจนตามเกณฑ์ที่กำหนดไว้

การคัดเลือกและฝึกอบรมผู้เขียนข้อสอบ เป็นทางหนึ่งที่จะช่วยทำให้ข้อคำถามรูปแบบเติมคำตอบมีความชัดเจน เพื่อป้องกันปัญหาที่อาจเกิดขึ้นกับความเที่ยงตรงของข้อคำถาม กรณีต้องดำเนินการสอบกับผู้สอบจำนวนมาก

นอกจากนี้ หากโครงสร้างของแบบทดสอบต้องการวัดความเข้มของความสามารถโดยตรง เช่นในสภาพการเรียนการสอนที่มุ่งพัฒนาผู้เรียนให้มีคุณลักษณะที่พึงประสงค์บางประการ แบบทดสอบที่สร้างจำเป็นต้องเน้นวัดความเข้มของความสามารถดังกล่าว เพื่อข้อมูลที่ได้จะเป็นประโยชน์ต่อการวินิจฉัยข้อบกพร่องของผู้เรียนได้อย่างถูกต้อง รูปแบบของข้อคำถามควรเป็นรูปแบบที่เอื้อให้เกิดความเที่ยงตรงตามโครงสร้างสูง โดยให้ผู้สอบได้แสดงความสามารถในการคิด การแก้ปัญหา การเขียน หรืออื่น ๆ อย่างเต็มที่ ขณะที่ถ้าต้องการนำผลการวัดไปใช้ในการคัดเลือกหรือประเมินคุณภาพการศึกษาระดับชาติ ซึ่งไม่ได้เน้นความสามารถหรือโครงสร้างตามทฤษฎีอย่างจริงจังนัก อาจเลือกใช้รูปแบบที่เหมาะสมและสะดวกต่อการนำไปใช้

ดังนั้น อาจกล่าวได้ว่าประสิทธิภาพของรูปแบบข้อคำถามมิได้อยู่ที่ตัวของมันเอง แต่ขึ้นอยู่กับทางเลือกนำไปใช้ โดยคำนึงถึงสาระของสิ่งที่ต้องการวัด สถานการณ์ที่ต้องการวัด ค่าใช้จ่าย ระดับความแม่นยำที่ต้องการ เพื่อการตัดสินใจเกี่ยวกับผลการประเมินเป็นสำคัญ และข้อคำถามทั้ง 6 ชนิด ที่ผู้วิจัยได้พัฒนาขึ้น ถือเป็นทางเลือกหนึ่งสำหรับการสร้างเครื่องมือเพื่อวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ของนักเรียนระดับชั้นมัธยมศึกษาชั้นนี้ ในโอกาสต่อไปอาจมีการพัฒนาชนิดของข้อคำถามเพิ่มขึ้นมากกว่านี้ เพื่อให้ครอบคลุมโครงสร้างที่ต้องการวัดมากยิ่งขึ้น โดยไม่จำกัดว่าต้องเป็น 6 ชนิดนี้ เท่านั้น

3.2 คุณภาพด้านเหตุผลเชิงวิเคราะห์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 และ 5

ความสามารถด้านเหตุผลถือเป็นกระบวนการทางสมองระดับสูง (Higher Mental Process) จากการศึกษา ผลการวิจัยด้านค่าความยากและค่าเฉลี่ยสัดส่วนการตอบถูกของ

ข้อคำถามชนิดต่าง ๆ ทั้ง 6 ชนิด พบว่า เมื่อเทียบค่าเฉลี่ยคะแนนที่ได้กับค่าเฉลี่ยตามความคาดหวังแล้ว อยู่ในระดับที่ค่อนข้างต่ำทั้ง 2 ระดับชั้น แม้ยังไม่สามารถสรุปได้ว่า นักเรียนชั้นมัธยมศึกษาปีที่ 3 และ 5 ซึ่งเป็นผลผลิตของหลักสูตรมัธยมศึกษาที่มีความสามารถด้านเหตุผลเชิงวิเคราะห์ที่อยู่ในระดับต่ำ หรือมีความสามารถทางสมองระดับสูงค่อนข้างต่ำก็ตาม แต่ข้อมูลที่ได้จากการศึกษาค้นคว้านี้ ถือเป็นสิ่งบอกเหตุประการหนึ่งที่ผู้รับผิดชอบควรให้ความเอาใจใส่ และพิจารณาหาสาเหตุว่าอะไรน่าจะส่งผลต่อสภาพที่เกิดขึ้นดังกล่าว เพื่อจะได้แก้ไขได้ตรงจุด สาเหตุมักจะอยู่ที่หลักสูตร หรือสื่อการเรียนการสอนที่มีระบุคุณลักษณะด้านนี้ไว้ให้ชัดเจน หรือการจัดกิจกรรมการเรียนการสอนอาจไม่ส่งเสริมให้ผู้เรียนได้มีโอกาสในการพัฒนาความสามารถในด้านนี้ กรมวิชาการในฐานะผู้รับผิดชอบด้านการพัฒนาหลักสูตร สื่อการเรียนการสอน อาจต้องย้อนกลับมาพิจารณาว่า จุดมุ่งหมายของหลักสูตรกำหนดความสามารถด้านเหตุผลไว้อย่างไร ชัดเจนเพียงใด สื่อการเรียนการสอนได้ให้แนวทางการจัดการเรียนการสอนที่ส่งเสริมพัฒนาการในด้านนี้เพียงใด สำหรับหน่วยงานต้นสังกัดของโรงเรียนและโรงเรียนจำเป็นจะต้องพึงเล็งไปที่กระบวนการเรียนการสอน ซึ่งควรส่งเสริมให้ผู้เรียนได้พัฒนาสมองในระดับสูงมากขึ้น โดยใช้กระบวนการสอนแบบต่าง ๆ ที่ยอมรับกันว่าเป็นวิธีสอนที่เอื้อต่อการพัฒนาสมองระดับสูง เช่น การสอนที่เน้นแนวทางกระบวนการคิด การสอนที่เน้นยุทธศาสตร์การคิด หรือแนวทางเกี่ยวกับพัฒนาการของการคิดตามทฤษฎีของเพียเจต์ ดังได้กล่าวมาแล้ว เป็นต้น

บรรณานุกรม

บรรณานุกรม

- คณะกรรมการการศึกษาแห่งชาติ, สำนักงาน. แผนการศึกษาแห่งชาติ พุทธศักราช 2535.
 กรุงเทพฯ : สำนักนายกรัฐมนตรี, 2535.
- คณะกรรมการพัฒนาเศรษฐกิจและสังคมแห่งชาติ, สำนักงาน. "แนวคิดและทิศทางการพัฒนา
 ประเทศ ในช่วงแผนพัฒนาฯ ฉบับที่ 8," เอกสารประกอบการสัมมนาระดับชาติ
ระหว่างวันที่ 3 - 4 มีนาคม พ.ศ. 2538. ม.ป.ท., ป.ป.ป.
- วิชาการ, กรม. หลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ.
2533). กรุงเทพฯ : กระทรวงศึกษาธิการ, 2533.
- ชาลี อุปกิจ. การเปรียบเทียบความสามารถในการแก้ปัญหาเชิงคณิตศาสตร์ สัมฤทธิ์ผลทาง
คณิตศาสตร์ และระดับสติปัญญาของนักเรียนระดับชั้นมัธยมศึกษาตอนต้น ที่มีแบบการ
คิดต่างกัน. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2523.
 อัดสำเนา.
- ✓ เบิดศักดิ์ โขวาสินธุ์. การฝึกสมรรถภาพสมองเพื่อพัฒนาคุณภาพการคิด. วิทยานิพนธ์
 กศ.ด. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ, 2530. อัดสำเนา.
- ผจงจิต อินทสุวรรณ. ทฤษฎีการตอบข้อคำถาม. สถาบันวิจัยพฤติกรรมศาสตร์
 มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2528.
- _____. "การวิเคราะห์แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนด้วยโมเดลโลจิสติก,"
กระบวนกรวิจัย. ฉบับที่ 45. สถาบันวิจัยพฤติกรรมศาสตร์ มหาวิทยาลัย-
 ศรีนครินทรวิโรฒ ประสานมิตร, 2534.
- พุทธทาสภิกขุ. การศึกษาคืออะไร. สำนักงานหนังสือธรรมบุชา กรุงเทพฯ : โรงพิมพ์การ
 พิมพ์พานิช, 2517.
- ✓ สมเจตน์ ไวยากรณ์. รูปแบบการสอนเพื่อพัฒนาความสามารถด้านการใช้เหตุผล. วิทยานิพนธ์
 กศ.ด. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ, 2530. อัดสำเนา.
- สำเริง บุญเรืองรัตน์. "การพัฒนาทฤษฎีลาแทนเทรทเพื่อวิเคราะห์ข้อสอบ," วารสาร
การวัดผลการศึกษา. 7(21) : 41 - 68 ; มกราคม - เมษายน, 2529.

Aiken, L.R. Psychological Testing and Assessment. 5th ed. Boston :
 Allyn and Bacon, 1985.

- Atkinson, R.C. and others. "Problem Solving and Reasoning," Steven's Handbook of Experimental Psychology : Vol 2. Learning and Cognition. New York : John Wiley and Sons, 1988.
- Ausubel, D.P. Educational Psychology : A Cognitive View. New York : Holt, Rinehart and Winston, Inc., 1968.
- Baron, I., P. Badgion, and I.W. Gaskins. "Cognitive Style and Its Improvement : A Normative Approach," Advances in The Psychology of Human Intelligence Vol 3. New Jersey : Erlbaum Hills Dales, 1986.
- Bennett, R.E. and W.C. Ward. "Item Formats for Assessment in Mathematics," Construction versus Choice in Cognitive Measurement. Hillsdale, NJ : Erlbaum, 1987.
- Bennett, R.E. "On the Meanings of Constructed Response," Research Report. E.T.S. Princeton, NJ. December. 1991.
- Bergh, Huub Van Den. "On the Construct Validity of Multiple-Choice Items for Reading Comprehension," Applied Psychological Measurement. 14(1) : 1 - 12 ; March, 1990.
- Beyer, B.K. "Developing a Scope and Sequence for Thinking Skills Instruction," Educational Leadership. 45(7) : 26 - 30 ; April, 1988.
- Birenbaum, M. and K.K. Tatsuoaka. "Open-Ended Versus Multiple-Choice Response Formats-It Does Make a Difference for Diagnostic Purposes," Applied Psychological Measurement. 11(4) : 385 - 395 ; December, 1987.
- Bridgeman, B. A Comparison of Open-Ended and Multiple Choice Question Formats for The Quantitative Section of The Graduate Record Examination. Princeton, NJ : Educational Testing Service, 1987.
- Brody, N. Intelligence. 2nd ed. San Diego : Academic Press, 1992.

Bruner, J.S., J.J. Goodnow and G.A. Austin. A study of Thinking.

New York : Wiley, 1956.

Buckel, C. "The Types, Purposes and Effects of Assessment,"

Assessment and Examination in The Secondary School. London

: Routledge, 1990.

Budescu, D.V. and B. Nevo. "Optimum Number of Options : An Inves-

tigation of The Assumption of Proportionality," Journal

of Educational Measurement. 22 : 183 - 196 ; 1985.

Bullock, M. "Causal Reasoning and Developmental Change Over the

Preschool Years," Human Development. 28(4) : 169 - 91; 1985.

Bullock, M. "Development of Reasoning in Childhood," The Interna-

tional Encyclopedia of Education. 8 (4951 - 4956) Great

Britain : BPC Wheatons Ltd., Exeter. 1994.

Coates, S. The Preschool Embedded Figures. Test-PEFT. Pala Alto

California : Consulting Psychologists Press, 1972.

Crocker, L. and J. Algina. Introduction to Classical and Modern

Test Theory. Holt, Rinehart and Winston Inc., 1986.

Downing, S.M. "True-False and Alternative-Choice Formats : A Review

of Research. Educational Measurement : Issues and Practices.

11 : 27 - 30 ; 1993.

Emmerich, W. and others. "The Development, Investigation, and

Evaluation of New Item Types for The GRE Analytical Measure,"

ETS Research Report. 91(16). Princeton, NJ : Educational

Testing Service, August, 1991.

✓ Enright, M.K., C.B. Tucker and Katz. A Cognitive Analysis of

Solutions for Verbal, Informal and Formal - Deductive

Reasoning Problems (GRE Board Professional Rep. No. 90 - 40

P, ETS Research Report No. 95-06) Princeton, N.J. : Educa-

tional Testing Service. February, 1995.

- Flavell, J.H. and E.M. Markman. "Logical Reasoning," Handbook of Child Psychology Vol. 3 : Cognitive Development. New York : Wiley, 1983.
- Frederikson, N. "The Real Test Bias : "Influences of Testing on Teaching and Learning," American Psychologist. 39 : 193 - 202 ; 1984.
- Gagne, E.D. The Cognitive Psychology of School Learning. Little, Broen and Company, 1985.
- Gardner, H. Multiple Intelligence : The Theory in Practice. New York : Basic Books, 1993.
- Goswami, U. Analogical Reasoning in Children. Lawrence Erlbanm Associates Ltd., 1992.
- Grier, J.B. "The Number of Alternatives for Optimum Test Reliability," Journal of Educational Measurement. 12(2) : 109 - 112 ; 1975
- Gronlund, N.E. and R.L. Linn. Measurement and Evaluation in Teaching. 6nd ed. New York : Macmillan Publishing Company, 1990.
- Guilford, J.P. The Nature of Human Intelligence. New York : McGraw Hill, 1967.
- Haladyna, T.M. and S.M. Downing. "A Taxonomy of Multiple Choice Item-Writing Rules," Applied Measurement in Education. 2 : 37 - 50 ; 1989.
- Haladyna, T.M. and S.M. Downing. "How many Options is Enough for A Multiple Choice Test Item?," Educational and Psychological Measurement. 53 ; 1993.
- Hambleton, R.K. and H. Swaminathan. Item Response Theory : Principles and Application. Boston : Kluwer-Nigholf Publishing, 1985.

- Hambleton, R.K., H. Swaminathan and H.J. Rogers. Fundamentals of Item Response Theory. California : Sage Publications Inc, 1991.
- Hilgard, E.R. Introduction to Psychology. New York : Harcourt, Brace and World Inc., 1962.
- Hinkle, D.E., O. Dale and A.H. Charles. "How Large Should The Sample Be? Part II-The One-Sample Case for Survey Research," Educational and Psychological Measurement. 45(2) : 271 - 280 ; Summer, 1985.
- Hogan, T.P. Relationships Between Free-Response and Choice Type Tests of Achievement : A Review of Literature. Washington DC : National Institute of Education, 1981.
- Hopkins, C.D. and R.L. Antes. Classroom Testing Construction. Illinois : F.E.Peacock Publishers, Inc., 1979.
- Hulin, C.L., F. Drasgow and C.K. Parson. Item Response Theory : Application to Psychological Measurement. Homewood, IL : Don Jones Irwin, 1983.
- Inhelder, B. The Early Growth of Logic in The Child. London : Routledge and Kegan Paul Ltd., 1964.
- Inhelder, B. and J. Piaget. The Growth of Logical Thinking from Childhood to Adolescence : An Essay on the Construction of Formal Operational Structures, Basic Book. New York : 1958.
- Johnson, D.A. and R.J. Rising. Guidelines for Teaching Mathematics. Belmont Calif, Nadelworth., 1969.
- Klausmeir, H.J. and R.E. Ripple. Learning and Human Abilities. New York : Harper and Row, Publishers, Inc., 1971.
- Krulik, S. and J.A. Rudnick. Reasoning and Problem Solving : A Handbook for Elementary School Teacher. Boston : Allyn and Bacon., 1993.

- Lord, F.M. "Optimum Number of Choices per Item - A comparison of Four Approaches," Journal of Educational Measurement. 14 : 33 - 38 ; 1977.
- Lord, F.M. Application of Item Response Theory to Practical Testing Problem. Hillsdale, New Jersey : Lawrence Erlbaum Associates, 1980.
- Lord, F.M. and M.R. Novick. Statistical Theories of Mental Test Scores. Addison - Wesley. Publishing Company, Inc., 1968.
- Lundeberg, M.A. and P.W. Fox. "Do Laboratory Finding on Test Expectancy Generalize to Classroom Outcomes?," Review of Educational Research. 61 : 94 - 106 ; 1991.
- Lumsden, J. "Test Theory," Annual Review of Psychology. Palo Alto, CA : Annual Reviews Inc., 1976.
- Lumsden, J. "The Construction of Unidimensional Tests," Psychological Bulletin. 58 : 122 - 131 ; 1961.
- McDonald, R.P. and K.S. Ahlawat. "Difficulty Factors in Binary Data," British Journal of Mathematical and Statistical Psychology. 27 : 82 - 99 ; 1974.
- Mitchell, R. Testing for Learning : How New Approaches to Evaluation Can Improve American Schools. New York : The Free Press, 1992.
- Mislevy, R. and R.D. Bock. Item Analysis and Test Scoring with Binary Logistic Models. (BILOG 3) 2nd ed. University of Chicago, 1990.
- Morante, E.A. and A. Ulesky. "Assessment of Reasoning Abilities," Educational Leadership. 42(1) : 71 - 74 ; September, 1984.
- Nickerson, Raymond S. "Kinds of Thinking Taught in Current Programs," Educational Leadership. 42(1) : 26 - 36 ; September, 1984.

- Nunnally, J.C. Educational Measurement and Evaluation. McGraw-Hill Book Company, 1972.
- Overton, W.F. "The Development of Metalogical Understanding," Overton W.F. 1990.
- Overton, W.F. "Competence and Procedures : Constraints on The Development of Logical Reasoning," Overton W.F. 1990.
- Owen, S.V. and R.D. Froman. " What's Wrong with Three-Option Multiple-Choice Items?," Educational and Psychological Measurement. 47 : 513 - 522 ; 1987.
- Park, Dong-Gun and G.J. Lantenschlager, "Improving IRT Item Bias Detection With Iterative Linking and Ability Scale Purification," Applied Psychological Measurement. 14(2) : 163 - 173 ; June, 1990.
- Perkins, D.N. "Postprimary Education has Little Impact on Informal Reasoning," Journal of Educational Psychology. 77(5) : 562 - 571; 1985.
- Phye, G.D. and T. Andre. "Cognitive Classroom Learning," Understanding Thinking and Problem Solving. Academic Press, Inc., 1986.
- Piaget, J. The Psychology of Intelligence. Totowa, NJ : Littlefield Adams, 1972.
- Powers, D.E. and S.S. Swinton. "Extending the Measurement of Graduate Admission Abilities Beyond the Verbal and Quantitative Domains," Applied Psychological Measurement. 5(2) : 141 - 158 ; 1981.
- Powers, D.E. and S.S. Swinton. "Effects of Self-Study for Coachable Test Item Types," Journal of Educational Psychology. 76 : 266 - 278 ; 1984.

- Riding, R. and S. Butterfield. Assessment and Examination in The Secondary School : A Practical Guide for Teachers and Trainers. London : Routledge, 1990.
- Roger, Tim B. The Psychological Testing Enterprise : An Introduction. California : Brooks/Cole Publishing Company, 1995.
- Rossmann, B.L. "Analytic Cognitive Style in Children," Dissertation Abstracts International. 27 : 2126 - 2131 B; 1966.
- Saljo, R. and J. Wyndhamna. "Problem Solving, Academic Performance, and Situated Reasoning : A Study of Joint Cognitive Activity in a Formal Setting," Journal of Education of Psychology. 60(3) : 245 - 254; 1990.
- Schiever, S.W. A Comprehensive Approach to Teaching Thinking. Massachusetts : Allyn and Bacon, 1991.
- Shulman, L.S. "Test Design : A View from Practices," Educational Testing and Evaluation Design, Analysis and Policy. p.63 - 73. Beverly Hills, California. Sage Publication, 1980.
- Shepard, L. "Interview on Assessment Issues with Lorries Shepard," Educational Researcher. 20(2) : 21 - 23 ; 1991a.
- Siegel, H. Educating Reason : Rationality, Critical Thinking, and Education. Great Britain : Routledge, 1988.
- Spearman, C. The Abilities of Man. New York : Macmillan, 1927.
- Smith, F. To Think in Language, Learning and Education. London : Routledge, 1992.
- Smith, B.O. "On Teaching Thinking Skills : A Conversation with B. Othanel Smith," Educational Leadership. 45(7) : April, 1988.
- Snedecor, G.W. and G.C. William. Statistical Methods. Iowa : The Iowa State University Press, 1967.

- Sternberg, R.J. Beyond IQ : A Triarchic Theory of Human Intelligence. New York Cambridge University Press, 1985.
- Sternberg, R.J. "How Can We Teach Intelligence?," Educational Leadership. 42(1) : 38 - 48 ; September, 1984.
- Swart, R.J. "Restructuring Curriculum for Critical Thinking," Educational Leadership. 43(8) : 43 - 44 ; May, 1986.
- Swinton, S.S. and D.E. Powers. "A Study of The Effects of Special Preparation on GRE Analytical Scores and Item Types," Journal of Educational Psychology. 75 : 104 - 115 ; 1983.
- Swinton, S.S., C.L. Wild and M.M. Wallmark. Investigation of Practice Effects on Types of Questions in The Graduate Record Examinations Aptitude Test. (GRE Borad Professional Rep. No.80 - 101) Princeton, NJ : Educational Test Service. 1982.
- Tatsuoka, K.K. "Item Construction and Psychometric Models Appropriate for Constructed Responses," Research Paper. Princeton, NJ : ETS August, 1991.
- Traub, R.E. and C.W. Fisher. "On The Equivalence of Constructed-Response and Multiple-Choice Tests," Applied Psychological Measurement. 1 : 355 - 369 ; 1977.
- Traub, R.E. and K. Macrury. " Multiple-Choice VS Free Response in The Testing of Scholastic Achievement," Tests and Trends 8. Jahrbuch der Padagogischen Diagnostik. Weinheim, Germany : Beltz Verlag. 128 - 159 ; 1990.
- Trevisan, M.S., G. Sax and W.B. Michael. "The Effects of The Number of Options Per Item and Student Ability on Test Validity and Reliability," Education and Psychological Measurement. 51 : 829 - 837 ; 1991.

- Tversky, A. "On The Optimal Number of Alternatives at A Choice Point," Journal of Mathematical Psychology. 1 : 386 - 391 ; 1964.
- Valett, R.E. Developing Cognitive Abilities Teaching Children to Think. Saint Louis : The C.V. Mosby Company, 1978.
- Van den Bergh, H. "On the Construct Validity of Multiple-Choice Items for Reading Comprehension," Applied Psychological Measurement. 14(1) : 1 - 12 ; March, 1990.
- Voss, J.F. "Reasoning," Encyclopedia of Education. 8 : 4948 ; 1994.
- Wallace, M. and N. Kagan. Modes of thinking in Young Children. Holt, Rinehart and Winston, Inc., 1956.
- Warm, T.A. "A Primer of Item Response Theory," Technical Report 941278. Oklahoma : U.S. Coast Guard Institute, 1978.
- Whimbey, A. "The Key to Higher Order Thinking is Precise Processing," Educational Leadership. 42(1) : 66 - 70 ; September, 1984.
- Wiggins, G. "A True Test : Toward More Authentic and Equitable Assessment," Phi Delta Kappan. 70 : 703 - 713 ; 1989.
- Witkin, H.A. and others. "Role of The Field Dependent and Field-Independent, Cognitive Style in Academic Education : A Longitudinal Study," Journal of Educational Psychology. 1977.

ภาคผนวก

ภาคผนวก ก

คุณภาพข้อคำถามรายข้อ ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อคำถาม
จากแบบทดสอบแต่ละฉบับ จากการทดลองครั้งที่ 1

ตาราง 20 ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ AR แบบเลือกตอบ
จากการทดลองครั้งที่ 1

ข้อ	ม.3		ม.5	
	p	r	p	r
1	0.717	0.716	0.676	0.802
2	0.167	0.028	0.250	0.109
3	0.732	0.336	0.556	0.686
4	0.341	0.382	0.426	0.453
5	0.362	0.605	0.491	0.424
6	0.558	0.555	0.639	0.708
7	0.130	0.367	0.167	0.078
8	0.399	0.551	0.472	0.571
9	0.638	0.597	0.667	0.483
10	0.630	0.536	0.694	0.675
11	0.413	0.620	0.481	0.622
12	0.710	0.751	0.741	0.901
13	0.558	0.603	0.648	0.896
14	0.819	0.701	0.769	0.826
15	0.623	0.676	0.676	0.975
16	0.536	0.625	0.667	0.777
17	0.130	0.041	0.157	0.216
18	0.225	0.320	0.324	0.563
19	0.246	0.600	0.315	0.317
20	0.290	0.626	0.306	0.396
21	0.167	0.361	0.167	0.386

ตาราง 21 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ AX แบบเลือกตอบ
จากการทดลองครั้งที่ 1

ข้อ ที่	ม.3		ม.5	
	p	r	p	r
1	0.369	0.332	0.545	0.405
2	0.342	0.426	0.396	0.405
3	0.101	0.281	0.158	0.374
4	0.087	0.475	0.129	0.270
5	0.396	0.444	0.327	0.406
6	0.396	0.083	0.257	0.144
7	0.255	0.468	0.267	0.654
8	0.537	0.537	0.604	0.408
9	0.483	0.473	0.465	0.084
10	0.268	0.102	0.347	0.186
11	0.456	0.583	0.535	0.499
12	0.490	0.572	0.525	0.433
13	0.497	0.603	0.495	0.591
14	0.651	0.535	0.624	0.548
15	0.309	0.365	0.465	0.107
16	0.040	0.028	0.158	0.499

ตาราง 22 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ CV แบบเลือกตอบ
จากการทดลองครั้งที่ 1

ข้อ	ม.3		ม.5	
	p	r	p	r
1	0.309	0.038	0.406	0.079
2	0.057	0.285	0.075	0.184
3	0.069	0.047	0.047	0.130
4	0.600	0.514	0.509	0.660
5	0.514	0.539	0.623	0.507
6	0.394	0.368	0.500	0.397
7	0.349	0.369	0.396	0.291
8	0.783	0.607	0.726	0.801
9	0.143	0.379	0.151	0.397
10	0.571	0.493	0.519	0.505
11	0.834	0.638	0.811	0.667
12	0.766	0.653	0.755	0.730
13	0.154	0.218	0.198	0.074
14	0.543	0.471	0.547	0.406
15	0.526	0.592	0.585	0.596

ตาราง 23 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ LR แบบเลือกตอบ
จากการทดลองครั้งที่ 1

ข้อ ที่	ม.3		ม.5	
	p	r	p	r
1	0.812	0.564	0.899	0.460
2	0.218	0.418	0.413	0.597
3	0.624	0.293	0.661	0.212
4	0.665	0.552	0.798	0.442
5	0.112	0.253	0.128	0.227
6	0.282	0.223	0.275	0.190
7	0.265	0.257	0.321	0.222
8	0.376	0.423	0.394	0.603
9	0.294	0.333	0.312	0.403
10	0.229	0.302	0.312	0.285
11	0.229	0.190	0.193	0.111
12	0.576	0.211	0.523	0.404
13	0.247	0.142	0.284	0.318
14	0.506	0.399	0.505	0.244
15	0.229	0.543	0.394	0.551
16	0.459	0.495	0.514	0.513
17	0.312	0.341	0.294	0.641
18	0.565	0.403	0.541	0.682
19	0.141	0.406	0.119	0.316
20	0.124	0.130	0.128	0.400

ตาราง 24 ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ NLR แบบเลือกตอบ
จากการทดลองครั้งที่ 1

ข้อ ที่	พ.3		พ.5	
	p	r	p	r
1	0.375	0.567	0.429	0.535
2	0.132	0.184	0.051	0.112
3	0.125	0.356	0.184	0.503
4	0.669	0.292	0.459	0.260
5	0.809	0.391	0.765	0.642
6	0.228	0.231	0.184	0.295
7	0.537	0.525	0.439	0.382
8	0.382	0.549	0.153	0.402
9	0.199	0.459	0.133	0.369
10	0.243	0.368	0.245	0.530
11	0.118	0.444	0.286	0.591
12	0.265	0.285	0.184	0.125
13	0.309	0.401	0.398	0.185
14	0.206	0.337	0.235	0.588
15	0.154	0.131	0.214	0.351

ตาราง 25 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ PI แบบเลือกตอบ
จากการทดลองครั้งที่ 1

ข้อ ที่	ม.3		ม.5	
	p	r	p	r
1	0.727	0.613	0.904	0.726
2	0.399	0.110	0.317	0.408
3	0.741	0.732	0.846	0.790
4	0.685	0.778	0.846	0.968
5	0.622	0.756	0.779	0.835
6	0.559	0.728	0.644	0.757
7	0.629	0.737	0.769	0.839
8	0.713	0.828	0.846	0.774
9	0.671	0.655	0.712	0.801
10	0.336	0.827	0.375	0.454
11	0.483	0.733	0.702	0.806
12	0.538	0.859	0.663	0.830
13	0.566	0.821	0.779	0.920
14	0.636	0.824	0.827	0.979
15	0.601	0.968	0.798	1.015
16	0.776	0.925	0.875	1.072
17	0.580	0.902	0.779	1.012
18	0.678	0.926	0.827	1.040
19	0.671	0.804	0.827	0.863

ตาราง 26 ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ AR แบบเติมคำตอบ
จากการทดลองครั้งที่ 1

ข้อ ที่	ม.3		ม.5	
	p	r	p	r
1	0.560	0.761	0.517	0.340
2	0.709	0.342	0.736	0.097
3	0.823	0.641	0.839	0.167
4	0.106	0.332	0.069	0.447
5	0.248	0.560	0.391	0.445
6	0.589	0.722	0.793	0.512
7	0.156	0.393	0.287	0.283
8	0.504	0.470	0.563	0.516
9	0.652	0.392	0.770	0.490
10	0.582	0.466	0.713	0.417
11	0.021	0.343	0.057	0.531
12	0.730	0.512	0.793	0.616
13	0.418	0.623	0.713	0.843
14	0.390	0.660	0.414	0.825
15	0.447	0.510	0.483	0.473
16	0.142	0.256	0.287	0.441
17	0.702	0.811	0.828	0.770
18	0.617	0.544	0.575	0.754
19	0.326	0.718	0.506	0.681
20	0.525	0.682	0.540	0.486
21	0.752	0.693	0.805	0.850

ตาราง 27 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบฉบับ AX แบบเติมคำตอบ
จากการทดลองครั้งที่ 1

ข้อ	ม.3		ม.5	
	p	r	p	r
1	0.236	0.668	0.218	0.469
2	0.205	0.531	0.141	0.839
3	0.465	0.690	0.462	0.714
4	0.331	0.603	0.282	0.574
5	0.228	0.437	0.244	0.597
6	0.528	0.306	0.603	0.440
7	0.528	0.376	0.608	0.555
8	0.346	0.477	0.410	0.735
9	0.528	0.772	0.385	0.640
10	0.228	0.616	0.231	0.409
11	0.386	0.761	0.372	0.815
12	0.126	0.633	0.167	0.698
13	0.197	0.785	0.244	0.772
14	0.197	0.742	0.295	0.729
15	0.205	0.607	0.141	0.839
16	0.031	0.404	0.064	0.653

ตาราง 28 ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบฉบับ CV แบบเติมคำตอบ
จากการทดลองครั้งที่ 1

ข้อ	ม.3		ม.5	
	p	r	p	r
1	0.456	0.618	0.564	0.455
2	0.478	0.444	0.515	0.443
3	0.103	0.307	0.109	0.515
4	0.426	0.683	0.386	0.505
5	0.559	0.516	0.574	0.284
6	0.471	0.423	0.555	0.386
7	0.706	0.594	0.723	0.609
8	0.662	0.397	0.683	0.495
9	0.081	0.466	0.030	0.121
10	0.382	0.592	0.327	0.399
11	0.654	0.551	0.673	0.514
12	0.559	0.619	0.485	0.639
13	0.228	0.416	0.317	0.397
14	0.485	0.605	0.594	0.725
15	0.478	0.648	0.307	0.593

ตาราง 29 ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบฉบับ LR แบบเดิมค่าตอบ
จากการทดลองครั้งที่ 1

ข้อ ข้อ	ม.3		ม.5	
	p	r	p	r
1	0.207	0.530	0.171	0.425
2	0.107	0.745	0.092	0.646
3	0.050	0.966	0.039	0.090
4	0.215	0.588	0.132	0.625
5	0.289	0.748	0.184	0.572
6	0.099	0.473	0.132	0.674
7	0.132	0.627	0.118	0.654
8	0.198	0.528	0.171	0.700
9	0.190	0.657	0.329	0.261
10	0.041	0.507	0.000	0.000
11	0.050	0.552	0.053	0.584
12	0.289	0.690	0.250	0.221
13	0.025	0.462	0.053	0.260
14	0.306	0.140	0.421	0.490
15	0.256	0.483	0.289	0.442
16	0.074	0.501	0.092	0.126
17	0.438	0.391	0.447	0.490
18	0.397	0.607	0.368	0.359
19	0.107	0.312	0.092	0.554
20	0.306	0.556	0.408	0.459

ตาราง 30 ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบฉบับ NLR แบบเดิมค่าตอบ
จากการทดลองครั้งที่ 1

ข้อ ที่	ม.3		ม.5	
	p	r	p	r
1	0.460	0.535	0.365	0.370
2	0.180	0.563	0.230	0.538
3	0.209	0.477	0.378	0.495
4	0.540	0.515	0.716	0.394
5	0.561	0.584	0.581	0.590
6	0.036	0.494	0.189	0.423
7	0.324	0.592	0.432	0.446
8	0.338	0.619	0.419	0.419
9	0.245	0.363	0.230	0.453
10	0.216	0.529	0.405	0.474
11	0.317	0.389	0.378	0.538
12	0.324	0.504	0.473	0.365
13	0.554	0.476	0.703	0.337
14	0.129	0.277	0.243	0.438
15	0.187	0.373	0.162	0.435

ตาราง 31 ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบฉบับ PI แบบเดิมค่าตอบ
จากการทดลองครั้งที่ 1

ข้อ ที่	ม.3		ม.5	
	p	r	p	r
1	0.964	0.503	0.870	0.559
2	0.909	0.541	0.826	0.651
3	0.855	0.079	0.000	0.000
4	0.836	0.891	0.826	0.764
5	0.909	0.451	0.826	0.764
6	0.800	1.112	0.739	0.859
7	0.927	0.903	0.826	0.163
8	0.982	0.392	0.826	0.389
9	0.722	0.806	0.739	0.859
10	0.927	1.196	0.826	0.839
11	0.836	0.952	0.652	0.793
12	0.727	0.978	0.522	0.933
13	0.818	0.841	0.739	0.890
14	0.745	0.839	0.696	0.881
15	0.891	1.098	0.783	0.882
16	0.909	0.660	0.783	1.015
17	0.927	1.013	0.870	0.969
18	0.927	0.903	0.957	1.050
19	0.800	0.739	0.783	0.448

ภาคผนวก ข

1. คุณภาพข้อคำถามรายข้อ ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อคำถาม จากแบบทดสอบแต่ละฉบับ จากการทดลองครั้งที่ 2
2. ค่าพารามิเตอร์ ค่าความยาก (b) ค่าอำนาจจำแนก (a) และ ค่าการเดา (c) ของข้อคำถาม จากแบบทดสอบแต่ละฉบับ
3. จำนวนข้อสอบของแบบทดสอบแต่ละฉบับ
4. ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ณ ระดับความสามารถต่าง ๆ

ภาคผนวก ข

1. คุณภาพข้อคำถามรายข้อ ค่าความยาก (p) ค่าอำนาจจำแนก (r)
ของข้อคำถาม จากแบบทดสอบแต่ละฉบับ จากการทดลองครั้งที่ 2

ตาราง 32 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ AR แบบเลือกตอบ
จากการทดลองครั้งที่ 2

ข้อ	ม.3		ม.5	
	p	r	p	r
1	0.568	0.516	0.631	0.694
2	0.187	0.323	0.223	0.379
3	0.556	0.483	0.612	0.598
4	0.378	0.485	0.388	0.322
5	0.295	0.414	0.364	0.634
6	0.477	0.568	0.510	0.530
7	0.195	0.380	0.189	0.311
8	0.378	0.483	0.369	0.585
9	0.680	0.581	0.650	0.499
10	0.643	0.457	0.568	0.513
11	0.369	0.555	0.413	0.616
12	0.676	0.640	0.646	0.750
13	0.448	0.613	0.461	0.787
14	0.701	0.546	0.704	0.756
15	0.581	0.729	0.636	0.833
16	0.452	0.768	0.524	0.847
17	0.282	0.408	0.340	0.541
18	0.274	0.327	0.286	0.335
19	0.232	0.286	0.238	0.348
20	0.191	0.511	0.199	0.463
21	0.145	0.124	0.155	0.234

ตาราง 33 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ AX แบบเลือกตอบ
จากการทดลองครั้งที่ 2

ข้อ ที่	ม.3		ม.5	
	p	r	p	r
1	0.422	0.221	0.417	0.301
2	0.451	0.427	0.457	0.434
3	0.448	0.437	0.478	0.609
4	0.108	0.181	0.091	0.149
5	0.336	0.455	0.383	0.592
6	0.368	0.114	0.391	0.110
7	0.224	0.427	0.265	0.474
8	0.462	0.498	0.535	0.623
9	0.365	0.325	0.478	0.462
10	0.202	0.337	0.265	0.324
11	0.437	0.608	0.517	0.602
12	0.448	0.433	0.500	0.405
13	0.401	0.580	0.491	0.540
14	0.513	0.615	0.561	0.615
15	0.368	0.510	0.448	0.295

ตาราง 34 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ CV แบบเลือกตอบ
จากการทดลองครั้งที่ 2

ข้อ ที่	ม.3		ม.5	
	p	r	p	r
1	0.512	0.636	0.514	0.601
2	0.334	0.441	0.486	0.719
3	0.451	0.528	0.459	0.453
4	0.311	0.504	0.395	0.282
5	0.584	0.640	0.746	0.610
6	0.549	0.713	0.665	0.618
7	0.433	0.547	0.438	0.490
8	0.751	0.703	0.822	0.640
9	0.625	0.740	0.730	0.716
10	0.133	0.047	0.162	0.200
11	0.427	0.604	0.524	0.618
12	0.512	0.685	0.578	0.461

ตาราง 35 ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ LR แบบเลือกตอบ
จากการทดลองครั้งที่ 2

ข้อ	ม.3		ม.5	
	p	r	p	r
1	0.776	0.543	0.862	0.520
2	0.253	0.437	0.321	0.405
3	0.527	0.536	0.616	0.307
4	0.671	0.492	0.803	0.468
5	0.173	0.268	0.257	0.385
6	0.253	0.422	0.239	0.315
7	0.308	0.196	0.294	0.483
8	0.165	0.041	0.161	0.081
9	0.287	0.354	0.408	0.235
10	0.211	0.198	0.303	0.266
11	0.186	0.178	0.289	0.361
12	0.502	0.214	0.537	0.385
13	0.430	0.303	0.500	0.223
14	0.160	0.498	0.303	0.431
15	0.426	0.550	0.445	0.391
16	0.245	0.462	0.312	0.422
17	0.418	0.696	0.550	0.446
18	0.190	0.534	0.206	0.278
19	0.338	0.468	0.362	0.370

ตาราง 36 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ NLR แบบเลือกตอบ
จากการทดลองครั้งที่ 2

ข้อ ที่	ม.3		ม.5	
	p	r	p	r
1	0.322	0.371	0.353	0.584
2	0.488	0.517	0.559	0.479
3	0.496	0.448	0.451	0.322
4	0.581	0.367	0.588	0.291
5	0.671	0.532	0.779	0.491
6	0.233	0.461	0.157	0.416
7	0.391	0.562	0.815	0.468
8	0.329	0.468	0.412	0.577
9	0.140	0.202	0.176	0.182
10	0.411	0.500	0.451	0.332
11	0.153	0.282	0.230	0.372
12	0.353	0.318	0.289	0.425
13	0.252	0.614	0.284	0.552
14	0.248	0.441	0.289	0.605
15	0.217	0.387	0.275	0.546

ตาราง 37 ค่าความฮาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบ ฉบับ PI แบบเลือกตอบ
จากการทดลองครั้งที่ 2

ข้อ	ม.3		ม.5	
	p	r	p	r
1	0.752	0.746	0.755	0.691
2	0.586	0.892	0.759	0.919
3	0.626	0.752	0.750	0.709
4	0.653	0.802	0.764	0.908
5	0.622	0.738	0.665	0.685
6	0.491	0.670	0.599	0.851
7	0.572	0.787	0.642	0.765
8	0.653	0.877	0.783	0.860
9	0.577	0.826	0.646	0.794
10	0.500	0.900	0.651	0.872
11	0.518	0.747	0.580	0.603
12	0.410	0.901	0.613	0.923
13	0.563	0.851	0.731	0.897
14	0.536	0.895	0.646	0.718
15	0.523	0.850	0.632	0.783
16	0.662	0.836	0.792	0.921
17	0.595	0.892	0.689	0.875
18	0.631	0.755	0.698	0.851
19	0.590	0.824	0.665	0.861

ตาราง 38 ค่าความชาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบฉบับ AR แบบเติมคำตอบ
จากการทดลองครั้งที่ 2

ข้อ	ม.3		ม.5	
	p	r	p	r
1	0.523	0.655	0.621	0.588
2	0.684	0.526	0.712	0.758
3	0.742	0.722	0.803	0.755
4	0.065	0.478	0.258	0.429
5	0.471	0.679	0.598	0.709
6	0.484	0.697	0.629	0.783
7	0.155	0.301	0.235	0.418
8	0.355	0.567	0.492	0.324
9	0.645	0.166	0.674	0.261
10	0.626	0.430	0.568	0.430
11	0.090	0.452	0.098	0.268
12	0.677	0.526	0.621	0.601
13	0.484	0.730	0.720	0.775
14	0.374	0.627	0.553	0.713
15	0.465	0.606	0.644	0.644
16	0.200	0.443	0.371	0.411
17	0.535	0.817	0.614	0.883
18	0.535	0.758	0.652	0.789
19	0.206	0.562	0.364	0.547
20	0.323	0.731	0.477	0.624
21	0.419	0.729	0.674	0.678

ตาราง 39 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบฉบับ AX แบบเดิมค่าตอบ
จากการทดลองครั้งที่ 2

ข้อ ที่	น.3		น.5	
	p	r	p	r
1	0.177	0.706	0.318	0.647
2	0.083	0.337	0.164	0.281
3	0.427	0.666	0.590	0.653
4	0.266	0.504	0.351	0.512
5	0.313	0.442	0.261	0.254
6	0.505	0.503	0.590	0.530
7	0.594	0.598	0.627	0.546
8	0.365	0.626	0.537	0.549
9	0.313	0.552	0.470	0.559
10	0.125	0.282	0.104	0.307
11	0.385	0.684	0.552	0.868
12	0.089	0.798	0.239	0.505
13	0.177	0.623	0.396	0.754
14	0.214	0.563	0.366	0.769
15	0.120	0.666	0.231	0.511

ตาราง 40 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบฉบับ CV แบบเติมคำตอบ
จากการทดลองครั้งที่ 2

ข้อ	ม.3		ม.5	
	p	r	p	r
1	0.451	0.727	0.488	0.622
2	0.477	0.588	0.626	0.528
3	0.641	0.562	0.756	0.544
4	0.641	0.446	0.618	0.425
5	0.601	0.602	0.789	0.625
6	0.353	0.436	0.390	0.688
7	0.327	0.639	0.480	0.652
8	0.582	0.502	0.707	0.606
9	0.399	0.713	0.610	0.658
10	0.497	0.699	0.593	0.786
11	0.366	0.745	0.585	0.780
12	0.340	0.545	0.553	0.525

ตาราง 41 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบกับ LR แบบเดิมค่าสอบ
จากการทดลองครั้งที่ 2

ข้อ ที่	ม.3		ม.5	
	p	r	p	r
1	0.174	0.240	0.342	0.543
2	0.164	0.544	0.241	0.663
3	0.159	0.460	0.158	0.064
4	0.174	0.396	0.361	0.669
5	0.333	0.513	0.392	0.568
6	0.189	0.430	0.215	0.418
7	0.139	0.383	0.127	0.489
8	0.194	0.507	0.348	0.684
9	0.209	0.470	0.367	0.266
10	0.114	0.225	0.133	0.433
11	0.159	0.538	0.278	0.548
12	0.388	0.538	0.576	0.375
13	0.318	0.473	0.475	0.373
14	0.244	0.453	0.323	0.514
15	0.109	0.501	0.184	0.477
16	0.383	0.453	0.462	0.315
17	0.547	0.565	0.551	0.380
18	0.139	0.502	0.241	0.561
19	0.517	0.591	0.633	0.363

ตาราง 42 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบฉบับ NLR แบบเติมคำตอบ
จากการทดลองครั้งที่ 2

ข้อ	ม.3		ม.5	
	p	r	p	r
1	0.441	0.419	0.428	0.373
2	0.256	0.426	0.239	0.510
3	0.246	0.375	0.370	0.761
4	0.462	0.586	0.659	0.477
5	0.426	0.480	0.652	0.558
6	0.082	0.266	0.101	0.749
7	0.246	0.382	0.442	0.377
8	0.549	0.419	0.659	0.353
9	0.303	0.337	0.319	0.421
10	0.236	0.473	0.326	0.313
11	0.251	0.203	0.275	0.284
12	0.277	0.513	0.341	0.486
13	0.354	0.659	0.558	0.486
14	0.092	0.274	0.138	0.424
15	0.195	0.472	0.246	0.478

ตาราง 43 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของข้อสอบฉบับ PI แบบเดิมค่าตอบ
จากการทดลองครั้งที่ 2

ข้อ	ม.3		ม.5	
	p	r	p	r
1	0.819	0.744	0.919	0.517
2	0.582	0.934	0.662	0.905
3	0.780	0.769	0.858	0.898
4	0.588	0.861	0.649	0.903
5	0.740	0.815	0.818	0.665
6	0.480	0.020	0.541	0.867
7	0.757	0.765	0.770	0.613
8	0.859	0.933	0.926	0.977
9	0.492	0.928	0.608	0.948
10	0.565	1.039	0.730	1.017
11	0.706	0.691	0.723	0.502
12	0.508	0.928	0.588	0.940
13	0.582	1.022	0.730	1.074
14	0.548	0.975	0.709	1.014
15	0.638	0.984	0.777	1.165
16	0.689	1.049	0.750	1.010
17	0.610	0.983	0.736	1.013
18	0.706	0.992	0.797	1.124
19	0.605	0.945	0.716	0.972

ภาคผนวก ข

2. ค่าพารามิเตอร์ ค่าความชุก (b) ค่าอำนาจจำแนก (a) และ
ค่าการเคา (c) ของข้อคำถาม จากแบบทดสอบแต่ละฉบับ

ตาราง 44 คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษา-
ปีที่ 3 ฉบับ AR

ข้อ	ค่าพารามิเตอร์ (แบบเลือกตอบ)			ค่าพารามิเตอร์ (แบบเติมคำตอบ)		
	b	a	c	b	a	c
1	0.069	0.584	0.157	0.178	0.878	0.113
2	2.994	0.576	0.125	-0.733	0.627	0.091
3	0.189	0.565	0.169	0.846*	0.976	0.090
4	1.197	0.561	0.139	2.210	1.314	0.036
5	1.579	1.452	0.209*	0.273	0.886	0.067
6	1.639	0.816	0.195	0.247	0.872	0.080
7	1.818	1.448	0.134	0.992	0.612	0.079
8	1.235	0.706	0.178	-0.891	0.263	0.120
9	-0.577	0.730	0.129	-0.522	0.395	0.106
10	-0.481	0.482	0.141	2.448*	0.955	0.052
11	0.943	1.004	0.151	0.684	0.566	0.124
12	-0.507	0.889	0.123	0.190	0.884	0.059
13	0.470	0.936	0.117	0.751	1.144	0.112
14	-0.651	0.751	0.140	0.533	0.909	0.146
15	-0.017	1.492	0.142	1.940	0.617	0.066
16	0.299	1.657	0.078*	0.010	1.291	0.060
17	2.053	0.538	0.140	0.006	1.134	0.062
18	2.304	0.548	0.158	1.148	-0.824	0.039*
19	2.961	0.757	0.194	0.706	1.330	0.046
20	1.790	0.868	0.082	0.448	1.301	0.078

ตาราง 45 คุณภาพของแบบสอยวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษา-
ปีที่ 3 ฉบับ AX

ข้อ	ค่าพารามิเตอร์ (แบบเลือกตอบ)			ค่าพารามิเตอร์ (แบบเติมคำตอบ)		
	b	a	c	b	a	c
1	1.353	0.532	0.253	1.459	1.193	0.057
2	1.378	0.597	0.268*	-	-	-
3	1.862	0.621	0.250	0.580	1.011	0.124
4	2.163	0.733	0.138*	1.493	0.834	0.120
5	0.870	0.750	0.227	2.126*	0.655*	0.211
6	2.484	0.491	0.259	0.268	0.689	0.137
7	3.935	0.668	0.187	0.093*	0.874	0.117
8	0.583	0.909	0.141	0.792	0.880	0.101
9	1.363	0.497	0.241	1.260	1.271	0.178
10	0.785	0.913	0.148	-	-	-
11	0.307	1.029	0.170	0.790	0.736	0.112
12	1.312	0.737	0.191	1.843	1.870*	0.041*
13				1.759	0.780	0.057
14				1.716	1.191	0.124
15				1.803	1.276	0.045

ตาราง 46 คุณภาพของแบบสัณยวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษา-
ปีที่ 3 ฉบับ CV

ข้อ	ค่าพารามิเตอร์ (แบบเลือกตอบ)			ค่าพารามิเตอร์ (แบบเติมคำตอบ)		
	b	a	c	b	a	c
1	0.230	0.642	0.107	0.484	1.157	0.154
2	1.583	0.729	0.183	0.639	0.752	0.205
3	0.665	0.540	0.121	-0.198	0.661	0.217
4	1.474	0.783	0.150	-0.105	0.533	0.252
5	-0.208	0.620	0.086	0.106	0.953	0.239
6	0.013	0.985	0.100	1.847	0.865	0.268
7	0.814	0.566	0.136	1.019	0.985	0.136
8	-0.870	0.941	0.102	0.370	0.584	0.268
9	-0.272	1.116	0.102	0.773	1.075	0.161
10	0.565	0.662	0.081	0.394	0.951	0.185
11	0.089	0.837	0.065	0.782	1.063	0.126
12				1.209	0.898	0.181

ตาราง 47 คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษา-
ปีที่ 3 ฉบับ LR

ข้อ	ค่าพารามิเตอร์ (แบบเลือกตอบ)			ค่าพารามิเตอร์ (แบบเติมคำตอบ)		
	b	a	c	b	a	c
1	-0.991	0.830	0.167	2.252	0.893	0.104
2	2.515	0.638	0.176	2.367	1.174	0.133
3	0.238	0.681	0.173	1.437	0.712	0.163
4	-0.486	0.671	0.165	2.116	0.691	0.100
5	3.965	0.654	0.161	2.076	0.827	0.112
6	2.770	0.520	0.167	2.018	0.718	0.098
7	4.872	0.652	0.296	2.017	1.217	0.099
8	2.619	0.602	0.213	1.046	0.706	0.147
9	4.263	0.600	0.191	1.540	0.596	0.133
10	3.469	0.775	0.173	1.776	1.051	0.156
11	1.304	0.331	0.253	2.258	1.033	0.071
12	1.933	0.579	0.299	0.990	0.596	0.131
13	2.636	0.702	0.105	0.119	0.772	0.121
14	1.021	0.605	0.185	1.966	1.282	0.086
15	2.540	0.533	0.150	0.301	0.866	0.160
16	0.494	1.372	0.105			
17	1.296	0.695	0.141			

ตาราง 48 คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษา-
ปีที่ 3 ฉบับ NLR

ข้อ	ค่าพารามิเตอร์ (แบบเลือกตอบ)			ค่าพารามิเตอร์ (แบบเติมคำตอบ)		
	b	a	c	b	a	c
1	1.982	1.146	0.262	2.007	0.892	0.177
2	0.692	0.597	0.201	1.765	0.956	0.157
3	0.716	0.548	0.208	0.556	0.814	0.167
4	0.257	0.524	0.233	1.039	0.736	0.210
5	-0.405	0.690	0.194	0.830	0.589	0.307
6	2.123	0.983	0.173	1.373	1.056	0.145
7	1.146	0.799	0.202	0.894	1.295	0.149
8	1.656	0.851	0.215	2.712	0.860	0.164
9	1.233	0.694	0.222			
10	2.571	0.536	0.263			
11	1.589	0.864	0.124			
12	2.048	0.864	0.171			
13	2.767	0.706	0.169			

ตาราง 49 คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษา-
ปีที่ 3 ฉบับ PI

ข้อ	ค่าพารามิเตอร์ (แบบเลือกตอบ)			ค่าพารามิเตอร์ (แบบเติมคำตอบ)		
	b	a	c	b	a	c
1	-0.765	1.339	0.125	-1.320	0.965	0.035
2	-0.087	1.791	0.108	-0.175	1.441	0.022
3	-0.259	1.095	0.119	-0.957	1.113	0.035
4	-0.230	1.612	0.168	-0.201	1.233	0.041
5	-0.236	1.089	0.119	-0.893	0.992	0.034
6	0.505	3.951	0.242	0.150	4.018	0.014
7	0.060	1.392	0.160	-0.920	0.956	0.034
8	-0.353	1.548	0.104	-1.478	1.192	0.034
9	0.170	2.494	0.204	0.128	1.669	0.022
10	0.269	3.439	0.121	-0.060	2.213	0.024
11	0.265	1.296	0.153	-0.671	0.933	0.035
12	0.446	3.187	0.099	0.109	2.318	0.031
13	0.181	2.724	0.197	-0.104	2.632	0.022
14	0.233	2.772	0.166	-0.015	2.319	0.031
15	0.244	2.247	0.143	-0.288	2.199	0.041
16	-0.254	1.802	0.175	-0.528	2.385	0.020
17	-0.127	1.731	0.108	-0.247	1.996	0.018
18	0.104	2.091	0.262	-0.602	1.706	0.029
19	0.106	2.389	0.201	-0.212	1.739	0.026

ตาราง 50 คุณภาพของแบบสอยวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษา-
ปีที่ 5 ฉบับ AR

ข้อ	ค่าพารามิเตอร์ (แบบเลือกตอบ)			ค่าพารามิเตอร์ (แบบเติมคำตอบ)		
	b	a	c	b	a	c
1	-0.375	0.913	0.080	-0.211	0.666	0.137
2	2.112	0.530	0.077	-0.657	0.918	0.127
3	-0.313	0.686	0.095	-1.044	0.996	0.155
4	1.439	0.469	0.158	1.526	0.714	0.111
5	0.825	0.876	0.097	-0.133	0.850	0.128
6	0.227	0.573	0.098	-0.165	1.200	0.137
7	1.982	1.043	0.117	1.777	0.742	0.114
8	0.892	0.841	0.115	-0.588	0.316	0.187
9	-0.617	0.527	0.093	0.140	0.439	0.191
10	-0.125	0.558	0.092	-0.188	0.689	0.148
11	0.517	0.809	0.067	-0.695	0.861	0.130
12	-0.373	1.099	0.093	0.029	0.821	0.111
13	0.216	1.306	0.047	-0.335	0.748	0.136
14	-0.647	1.164	0.062	0.996	0.718	0.139
15	-0.348	1.544	0.060	-0.109	1.441	0.107
16	0.036	1.499	0.064	-0.445	1.089	0.106
17	1.054	0.596	0.070	0.870	0.693	0.103
18	2.138	0.539	0.150	0.357	1.006	0.104
19	2.224	0.582	0.119	-0.378	1.024	0.162
20	1.900	0.700	0.072			
21	3.581	0.776	0.135			

ตาราง 51 คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษา-
ปีที่ 5 ฉบับ AX

ข้อ	ค่าพารามิเตอร์ (แบบเลือกตอบ)			ค่าพารามิเตอร์ (แบบเติมคำตอบ)		
	b	a	c	b	a	c
1	3.174	0.563	0.363	0.922	0.843	0.145
2	1.162	0.626	0.254	0.158	0.808	0.193
3	0.553	1.001	0.197	1.287	0.758	0.174
4	1.022	0.716	0.155	0.022	0.580	0.181
5	1.851	0.770	0.159	-0.230	0.890	0.166
6	0.287	0.928	0.194	0.370	0.595	0.188
7	0.869	0.683	0.240	0.539	0.603	0.169
8	2.698	0.716	0.216	0.042	2.142	0.125
9	0.443	0.668	0.195	1.686	0.788	0.133
10	0.902	0.436	0.225	0.558	1.179	0.103
11	0.593	0.703	0.199	0.748	1.558	0.118
12	0.411	0.958	0.187	1.491	0.960	0.126
13	2.230	0.662	0.364			

ตาราง 52 คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษา-
ปีที่ 5 ฉบับ CV

ข้อ	ค่าพารามิเตอร์ (แบบเลือกตอบ)			ค่าพารามิเตอร์ (แบบเติมคำตอบ)		
	b	a	c	b	a	c
1	0.389	0.646	0.168	0.751	0.788	0.265
2	0.261	0.970	0.107	0.056	0.629	0.296
3	1.283	0.353	0.187	-0.496	0.870	0.336
4	-0.774	0.852	0.170	0.613	0.696	0.401
5	-0.373	0.880	0.165	-0.713	0.993	0.317
6	1.021	0.502	0.170	0.837	1.039	0.191
7	-1.358	0.746	0.157	0.606	0.993	0.231
8	-0.771	0.872	0.134	-0.245	0.836	0.322
9	0.291	0.613	0.153	0.182	1.026	0.300
10	-0.005	0.484	0.155	0.028	1.161	0.228
11				0.227	1.694	0.278
12				0.739	0.776	0.341

ตาราง 53 คุณภาพของแบบสอบถามวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษา-
ปีที่ 5 ฉบับ LR

ข้อ	ค่าพารามิเตอร์ (แบบเลือกตอบ)			ค่าพารามิเตอร์ (แบบเติมคำตอบ)		
	b	a	c	b	a	c
1	-1.545	0.742	0.244	1.480	0.962	0.195
2	1.796	0.655	0.194	1.190	1.419	0.103
3	-0.064	0.339	0.291	1.008	1.464	0.166
4	-1.206	0.617	0.248	0.932	0.735	0.176
5	2.496	0.605	0.179	2.068	1.204	0.161
6	3.113	0.665	0.203	2.261	1.089	0.092
7	1.882	0.682	0.179	0.990	1.311	0.153
8	2.644	0.468	0.309	2.523	0.868	0.319
9	4.171	0.541	0.277	1.566	0.990	0.171
10	3.120	0.585	0.241	0.461	0.518	0.259
11	0.970	0.359	0.265	1.138	0.536	0.253
12	4.186	0.501	0.460	1.299	0.839	0.157
13	1.994	0.629	0.189	2.312	0.882	0.130
14	1.318	0.497	0.233	2.701	0.781	0.397
15	2.138	0.619	0.210	0.723	0.606	0.270
16	0.493	0.518	0.241	1.514	1.217	0.155
17	3.704	0.652	0.189	0.017	0.516	0.241
18	2.036	0.546	0.237			

ตาราง 54 คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษา-
ปีที่ 5 ฉบับ NLR

ข้อ	ค่าพารามิเตอร์ (แบบเลือกตอบ)			ค่าพารามิเตอร์ (แบบเติมคำตอบ)		
	b	a	c	b	a	c
1	1.095	0.821	0.149	2.224	1.018	0.376
2	0.238	0.482	0.184	1.596	1.474	0.171
3	1.947	0.370	0.265	0.853	1.818	0.198
4	0.234	0.278	0.211	-0.015	1.060	0.304
5	-1.178	0.610	0.175	-0.007	1.105	0.296
6	2.960	0.640	0.107	3.869	1.042	0.421
7	0.562	0.508	0.194	0.360	0.499	0.372
8	0.810	0.710	0.144	3.417	1.003	0.309
9	4.599	0.606	0.167	1.597	1.081	0.251
10	1.521	0.351	0.213	0.581	1.026	0.319
11	3.294	0.554	0.176	2.938	1.053	0.141
12	1.731	0.696	0.150	2.312	1.071	0.216
13	1.584	0.656	0.119			
14	1.403	0.781	0.122			
15	1.446	0.860	0.125			

ตาราง 55 คุณภาพของแบบสอบวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ ชั้นมัธยมศึกษา-
ปีที่ 5 ฉบับ PI

ข้อ	ค่าพารามิเตอร์ (แบบเลือกตอบ)			ค่าพารามิเตอร์ (แบบเติมคำตอบ)		
	b	a	c	b	a	c
1	-0.774	0.998	0.167	-2.422	0.744	0.152
2	-0.801	1.532	0.088	-0.371	1.445	0.110
3	-0.868	0.972	0.112	-1.060	1.406	0.147
4	-0.771	1.561	0.117	-0.188	2.053	0.182
5	-0.404	0.940	0.145	-0.232	3.986	0.427
6	-0.068	1.855	0.156	0.093	1.730	0.108
7	-0.143	1.607	0.198	-0.363	2.302	0.302
8	-0.843	1.654	0.119	-2.436	1.225	0.138
9	-0.166	2.002	0.194	-0.132	2.093	0.100
10	-0.268	2.323	0.152	-0.751	1.487	0.097
11	0.302	1.355	0.268	0.012	1.767	0.304
12	-0.244	2.015	0.089	-0.180	1.587	0.073
13	-0.614	1.788	0.119	-0.528	2.384	0.076
14	-0.221	1.206	0.177	-0.421	2.384	0.072
15	-0.279	1.294	0.118	-0.608	5.211	0.068
16	-0.758	2.256	0.157	-0.716	1.751	0.085
17	-0.425	1.656	0.142	-0.741	1.695	0.081
18	-0.567	1.330	0.098	-0.790	3.193	0.103
19	-0.367	1.864	0.122	-0.643	1.697	0.082

ภาคผนวก ข

3. จำนวนข้อสอบของแบบทดสอบแต่ละฉบับ

ตาราง 56 จำนวนข้อสอบของแบบทดสอบแต่ละฉบับ (ชั้นมัธยมศึกษาปีที่ 3)

รูปแบบ	ชนิด	จำนวน ข้อเดิม	ครั้งที่ 1 (จำนวนข้อ)	ครั้งที่ 2 (จำนวนข้อ)	ข้อที่ถูกตัดออกไปเนื่องจาก ไม่เหมาะสมกับโมเดล
เลือกตอบ	AR	21	21 (138)	20 (230)	21
	AX	16	15 (149)	12 (277)	4, 6, 10
	CV	15	12 (175)	11 (293)	10
	LR	20	19 (170)	17 (237)	8, 18
	NLR	15	15 (136)	13 (258)	9, 11
	PI	19	19 (143)	19 (222)	-
เติมคำตอบ	AR	21	21 (141)	20 (155)	7
	AX	16	15 (127)	13 (192)	2, 10
	CV	15	12 (136)	12 (153)	-
	LR	20	19 (121)	15 (201)	1, 3, 7, 10
	NLR	15	15 (139)	8 (195)	1, 6, 7, 9, 11, 14
	PI	19	19 (55)	19 (177)	-

ตาราง 57 จำนวนข้อสอบของแบบทดสอบแต่ละฉบับ (ชั้นมัธยมศึกษาปีที่ 5)

รูปแบบ	ชนิด	จำนวน ข้อเดิม	ครั้งที่ 1 (จำนวนข้อ)	ครั้งที่ 2 (จำนวนข้อ)	ข้อที่ถูกลบออกไปเนื่องจาก ไม่เหมาะสมกับโมเดล
เลือกตอบ	AR	21	21 (108)	21 (206)	-
	AX	16	15 (101)	13 (230)	4, 6
	CV	15	12 (106)	10 (185)	3, 10
	LR	20	19 (106)	18 (218)	8
	NLR	15	15 (98)	15 (204)	-
	PI	19	19 (104)	19 (212)	-
เติมคำตอบ	AR	21	21 (87)	19 (132)	8, 11
	AX	16	15 (78)	12 (134)	2, 5, 10
	CV	15	12 (101)	12 (123)	-
	LR	20	19 (76)	17 (158)	3, 10
	NLR	15	15 (74)	12 (138)	6, 10, 11
	PI	19	19 (23)	19 (148)	-

ตาราง 58 คุณภาพด้านความเชื่อมั่นของแบบทดสอบ ภายหลังจากตัดข้อคำถามบางข้อที่ไม่เหมาะสมกับโมเดลออกไป

ชนิดของ ข้อคำถาม	ชั้นมัธยมศึกษาปีที่ 3				ชั้นมัธยมศึกษาปีที่ 5			
	เลือกตอบ		เติมคำตอบ		เลือกตอบ		เติมคำตอบ	
	N	r_{cc}	N	r_{cc}	N	r_{cc}	N	r_{cc}
AR	241	0.809	155	0.844	206	0.846	132	0.835
AX	277	0.592	192	0.749	230	0.634	134	0.772
CV	293	0.751	153	0.719	185	0.656	123	0.732
LR	237	0.621	201	0.648	218	0.473	158	0.694
NLR	258	0.556	195	0.544	204	0.572	138	0.634
PI	222	0.930	177	0.938	212	0.914	148	0.928

จากตาราง 58 แสดงให้เห็นว่า ภายหลังจากตัดข้อคำถามบางข้อที่ไม่เหมาะสมกับโมเดลออกไปแล้ว ค่าความเชื่อมั่นของแบบทดสอบรูปแบบเลือกตอบที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียนชั้น มัธยมศึกษาปีที่ 3 มีค่าอยู่ระหว่าง 0.556 ถึง 0.930 ขณะที่แบบเติมคำตอบมีค่าระหว่าง 0.544 ถึง 0.938 ทั้งนี้โดยแบบเติมคำตอบชนิด AR, AX, LR และ PI มีค่าความเชื่อมั่นสูงกว่าแบบเลือกตอบ โดยเฉพาะค่า r_{cc} ของ PI แบบเลือกตอบ จะมีค่าสูงสุด ส่วนแบบทดสอบ NLR จะมีค่า r_{cc} ต่ำสุดทั้ง 2 รูปแบบ

สำหรับค่าความเชื่อมั่นที่ได้จากกลุ่มตัวอย่างที่เป็นนักเรียน ชั้นมัธยมศึกษาปีที่ 5 นั้น รูปแบบเลือกตอบมีค่าอยู่ระหว่าง 0.473 ถึง 0.914 ขณะที่แบบเติมคำตอบมีค่าระหว่าง 0.634 ถึง 0.928 ทั้งนี้โดยแบบเติมคำตอบชนิด AX, CV, LR, NLR และ PI มีค่าความเชื่อมั่นสูงกว่าแบบเลือกตอบ และ PI มีค่าสูงสุดทั้ง 2 รูปแบบ เช่นกัน

ภาคผนวก ข

4. ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ณ ระดับความสามารถต่าง ๆ

ตาราง 59 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อน
มาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ชนิด AR แบบเลือกตอบ
และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

e	AR ₃₁		AR ₃₂		AR ₃₁ : AR ₃₂
	I (e)	SE (e)	I (e)	SE (e)	
-4.000	0.015	8.078	0.052	4.401	0.288
-3.750	0.024	6.517	0.067	3.863	0.358
-3.500	0.036	5.250	0.088	3.367	0.409
-3.250	0.056	4.226	0.118	2.912	0.475
-3.000	0.086	3.404	0.160	2.497	0.538
-2.750	0.132	2.748	0.222	2.122	0.595
-2.500	0.201	2.230	0.312	1.790	0.644
-2.250	0.301	1.824	0.455	1.500	0.676
-2.000	0.440	1.508	0.637	1.253	0.691
-1.750	0.628	1.262	0.913	1.047	0.688
-1.500	0.871	1.071	1.300	0.877	0.670
-1.250	1.186	0.918	1.831	0.739	0.648
-1.000	1.605	0.789	2.544	0.627	0.631
-0.750	2.199	0.674	3.467	0.537	0.634
-0.500	3.042	0.573	4.589	0.467	0.663
-0.250	4.099	0.494	5.814	0.415	0.705
0.000	5.110	0.442	6.940	0.380	0.736
0.250	5.678	0.420	7.718	0.360	0.736
0.500	5.626	0.422	7.954	0.355	0.707
0.750	5.230	0.437	7.602	0.363	0.688

ตาราง 59 (ต่อ)

๑	AR _{๓1}		AR _{๓2}		AR _{๓1} : AR _{๓2}
	I (๑)	SE (๑)	I (๑)	SE (๑)	
1.000	4.927	0.451	6.816	0.383	0.723
1.250	4.914	0.451	5.883	0.412	0.835
1.500	5.040	0.445	5.064	0.444	0.995
1.750	4.977	0.448	4.463	0.473	1.115
2.000	4.533	0.470	4.018	0.499	1.128
2.250	3.814	0.512	3.585	0.528	1.064
2.500	3.058	0.572	3.069	0.571	0.966
2.750	2.417	0.643	2.487	0.634	0.972
3.000	1.925	0.721	1.919	0.722	1.003
3.250	1.555	0.802	1.430	0.836	1.087
3.500	1.267	0.888	1.043	0.979	1.215
3.750	1.036	0.983	0.752	1.153	1.378
4.000	0.844	1.089	0.540	1.361	1.503

ตาราง 60 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อน
มาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ชนิด AX แบบเลือกตอบ
และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

e	AX ₃₁		AX ₃₂		AX ₃₁ : AX ₃₂
	I (e)	SE (e)	I (e)	SE (e)	
-4.000	0.001	42.680	0.001	40.134	1.000
-3.750	0.001	33.903	0.001	29.451	1.000
-3.500	0.001	26.838	0.002	21.613	0.500
-3.250	0.002	21.152	0.004	15.874	0.500
-3.000	0.004	16.579	0.007	11.681	0.571
-2.750	0.006	12.907	0.013	8.625	0.462
-2.500	0.010	9.969	0.024	6.402	0.417
-2.250	0.017	7.634	0.044	4.789	0.386
-2.000	0.030	5.797	0.076	3.620	0.395
-1.750	0.052	4.372	0.130	2.773	0.400
-1.500	0.093	3.287	0.215	2.159	0.433
-1.250	0.163	2.477	0.341	1.712	0.478
-1.000	0.281	1.885	0.522	1.384	0.538
-0.750	0.468	1.461	0.766	1.142	0.611
-0.500	0.739	1.163	1.081	0.962	0.684
-0.250	1.090	0.958	1.469	0.825	0.742
0.000	1.491	0.819	1.936	0.719	0.770
0.250	1.881	0.729	2.495	0.633	0.754
0.500	2.195	0.675	3.175	0.561	0.691
0.750	2.389	0.647	4.016	0.499	0.595

ตาราง 60 (ต่อ)

e	AX ₃₁		AX ₃₂		AX ₃₁ : AX ₃₂
	I (e)	SE (e)	I (e)	SE (e)	
1.000	2.453	0.638	5.042	0.445	0.487
1.250	2.408	0.644	6.200	0.402	0.388
1.500	2.287	0.661	7.257	0.371	0.315
1.750	2.123	0.686	7.702	0.360	0.276
2.000	1.938	0.718	7.051	0.377	0.275
2.250	1.747	0.756	5.557	0.424	0.314
2.500	1.560	0.801	3.958	0.503	0.394
2.750	1.382	0.851	2.691	0.610	0.514
3.000	1.217	0.906	1.807	0.744	0.673
3.250	1.070	0.967	1.219	0.906	0.878
3.500	0.939	1.032	0.831	1.097	1.130
3.750	0.823	1.102	0.573	1.321	1.436
4.000	0.719	1.180	0.399	1.583	1.802

ตาราง 61 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อน
มาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ชนิด CV แบบเลือกตอบ
และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

e	CV ₃₁		CV ₃₂		CV ₃₁ : CV ₃₂
	I (e)	SE (e)	I (e)	SE (e)	
-4.000	0.007	11.725	0.003	18.024	2.333
-3.750	0.012	9.017	0.005	14.308	2.400
-3.500	0.021	6.925	0.008	11.377	2.625
-3.250	0.035	5.315	0.012	9.067	2.917
-3.000	0.060	4.082	0.019	7.244	3.158
-2.750	0.101	3.145	0.030	5.805	3.367
-2.500	0.168	2.439	0.046	4.666	3.652
-2.250	0.274	1.910	0.071	3.762	3.859
-2.000	0.434	1.517	0.108	3.039	4.019
-1.750	0.666	1.226	0.166	2.457	4.012
-1.500	0.980	1.010	0.254	1.985	3.858
-1.250	1.379	0.851	0.390	1.601	3.536
-1.000	1.844	0.736	0.601	1.290	3.068
-0.750	2.329	0.655	0.919	1.043	2.534
-0.500	2.760	0.602	1.380	0.851	2.000
-0.250	3.061	0.572	1.995	0.708	1.534
0.000	3.184	0.560	2.725	0.606	1.168
0.250	3.131	0.565	3.460	0.538	0.905
0.500	2.954	0.582	4.040	0.498	0.731
0.750	2.711	0.607	4.322	0.481	0.627

ตาราง 61 (ต่อ)

e	CV ₃₁		CV ₃₂		CV ₃₁ : CV ₃₂
	I (e)	SE (e)	I (e)	SE (e)	
1.000	2.449	0.639	4.257	0.485	0.575
1.250	2.191	0.676	3.899	0.506	0.562
1.500	1.940	0.718	3.366	0.545	0.576
1.750	1.695	0.768	2.777	0.600	0.610
2.000	1.456	0.829	2.214	0.672	0.658
2.250	1.226	0.903	1.718	0.763	0.714
2.500	1.011	0.995	1.305	0.875	0.775
2.750	0.818	1.106	0.973	1.014	0.841
3.000	0.651	1.239	0.715	1.183	0.910
3.250	0.511	1.399	0.519	1.388	0.985
3.500	0.397	1.587	0.373	1.636	1.064
3.750	0.306	1.808	0.267	1.934	1.146
4.000	0.234	2.066	0.191	2.290	1.225

ตาราง 62 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อน
มาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ชนิด LR แบบเลือกตอบ
และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

e	LR ₃₁		LR ₃₂		LR ₃₁ : LR ₃₂
	I (e)	SE (e)	I (e)	SE (e)	
-4.000	0.006	12.834	0.001	37.911	6.000
-3.750	0.010	10.013	0.001	28.527	10.000
-3.500	0.017	7.763	0.002	21.450	8.500
-3.250	0.028	6.007	0.004	16.129	7.000
-3.000	0.046	4.662	0.007	12.139	6.571
-2.750	0.075	3.648	0.012	9.158	6.250
-2.500	0.120	2.892	0.021	6.937	5.714
-2.250	0.184	2.334	0.036	5.286	5.111
-2.000	0.270	1.926	0.061	4.062	4.426
-1.750	0.377	1.628	0.100	3.155	3.770
-1.500	0.501	1.413	0.162	2.484	3.093
-1.250	0.634	1.256	0.253	1.987	2.506
-1.000	0.772	1.138	0.381	1.619	2.026
-0.750	0.925	1.040	0.551	1.347	1.679
-0.500	1.118	0.946	0.762	1.146	1.467
-0.250	1.385	0.850	1.008	0.996	1.374
0.000	1.725	0.761	1.283	0.883	1.345
0.250	2.067	0.696	1.589	0.793	1.301
0.500	2.287	0.661	1.948	0.716	1.174
0.750	2.310	0.658	2.408	0.644	0.959

ตาราง 62 (ต่อ)

๑	LR ₃₁		LR ₃₂		LR ₃₁ : LR ₃₂
	I (๑)	SE (๑)	I (๑)	SE (๑)	
1.000	2.183	0.677	3.029	0.575	0.721
1.250	2.017	0.704	3.839	0.510	0.525
1.500	1.892	0.727	4.771	0.458	0.397
1.750	1.830	0.739	5.628	0.422	0.325
2.000	1.820	0.741	6.135	0.404	0.297
2.250	1.840	0.737	6.102	0.405	0.302
2.500	1.870	0.731	5.546	0.425	0.337
2.750	1.895	0.726	4.661	0.463	0.407
3.000	1.901	0.725	3.681	0.521	0.516
3.250	1.880	0.729	2.780	0.600	0.676
3.500	1.829	0.739	2.037	0.701	0.898
3.750	1.747	0.757	1.465	0.826	1.192
4.000	1.638	0.781	1.044	0.979	1.569

ตาราง 63 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อน
มาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ชนิด NLR แบบเลือกตอบ
และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

๑	NLR _{๓1}		NLR _{๓2}		NLR _{๓1} : NLR _{๓2}
	I (๑)	SE (๑)	I (๑)	SE (๑)	
-4.000	0.003	18.046	0.000	72.928	0.000
-3.750	0.005	14.195	0.000	55.687	0.000
-3.500	0.008	11.188	0.001	42.436	0.125
-3.250	0.013	8.845	0.001	32.276	0.077
-3.000	0.020	7.024	0.002	24.506	0.100
-2.750	0.032	5.613	0.003	18.583	0.094
-2.500	0.049	4.521	0.005	14.082	0.102
-2.250	0.074	3.678	0.009	10.676	0.122
-2.000	0.109	3.029	0.015	8.107	0.138
-1.750	0.157	2.528	0.026	6.178	0.166
-1.500	0.218	2.141	0.045	4.733	0.206
-1.250	0.295	1.842	0.075	3.652	0.254
-1.000	0.387	1.608	0.124	2.842	0.320
-0.750	0.494	1.423	0.201	2.229	0.407
-0.500	0.619	1.271	0.323	1.760	0.522
-0.250	0.764	1.144	0.511	1.398	0.669
0.000	0.938	1.032	0.797	1.120	0.850
0.250	1.150	0.933	1.197	0.914	1.041
0.500	1.407	0.843	1.689	0.769	1.200
0.750	1.718	0.763	2.183	0.677	1.271

ตาราง 63 (ต่อ)

๑	NLR _{๓1}		NLR _{๓2}		NLR _{๓1} : NLR _{๓2}
	I (๑)	SE (๑)	I (๑)	SE (๑)	
1.000	2.078	0.694	2.559	0.625	1.231
1.250	2.470	0.636	2.745	0.604	1.111
1.500	2.851	0.592	2.747	0.603	0.964
1.750	3.144	0.564	2.615	0.618	0.832
2.000	3.266	0.553	2.395	0.646	0.733
2.250	3.170	0.562	2.125	0.686	0.670
2.500	2.881	0.589	1.832	0.739	0.636
2.750	2.476	0.636	1.538	0.806	0.621
3.000	2.039	0.700	1.257	0.892	0.616
3.250	1.629	0.783	1.000	1.000	0.614
3.500	1.276	0.885	0.775	1.136	0.607
3.750	0.987	1.007	0.587	1.305	0.595
4.000	0.756	1.150	0.436	1.515	0.577

ตาราง 64 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อน
มาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ชนิด PI แบบเลือกตอบ
และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

e	PI ₃₁		PI ₃₂		PI ₃₁ : PI ₃₂
	I (e)	SE (e)	I (e)	SE (e)	
-4.000	0.000	128.998	0.019	7.278	0.000
-3.750	0.000	78.976	0.040	4.978	0.000
-3.500	0.000	48.190	0.084	3.458	0.000
-3.250	0.001	29.324	0.166	2.452	0.006
-3.000	0.003	17.818	0.315	1.781	0.010
-2.750	0.009	10.839	0.564	1.331	0.016
-2.500	0.023	6.629	0.952	1.025	0.024
-2.250	0.059	4.104	1.511	0.814	0.039
-2.000	0.149	2.593	2.260	0.665	0.066
-1.750	0.351	1.687	3.223	0.557	0.109
-1.500	0.775	1.136	4.528	0.470	0.171
-1.250	1.599	0.791	6.601	0.389	0.242
-1.000	3.113	0.567	10.341	0.311	0.301
-0.750	5.708	0.419	16.680	0.245	0.342
-0.500	9.778	0.320	25.029	0.200	0.391
-0.250	16.382	0.247	33.349	0.173	0.491
0.000	28.004	0.189	39.214	0.160	0.714
0.250	39.803	0.159	32.545	0.175	1.223
0.500	37.439	0.163	16.508	0.246	2.268
0.750	20.153	0.223	7.435	0.367	2.711

TABLE 64 (cont)

σ	PI_{31}		PI_{32}		$PI_{31} : PI_{32}$
	$I(\sigma)$	$SE(\sigma)$	$I(\sigma)$	$SE(\sigma)$	
1.000	7.861	0.357	3.478	0.536	2.260
1.250	3.131	0.565	1.708	0.765	1.833
1.500	1.388	0.849	0.876	1.068	1.584
1.750	0.675	1.217	0.467	1.463	1.445
2.000	0.350	1.691	0.258	1.967	1.357
2.250	0.189	2.299	0.147	2.606	1.286
2.500	0.105	3.083	0.086	3.408	1.221
2.750	0.060	4.093	0.051	4.408	1.176
3.000	0.034	5.396	0.031	5.650	1.097
3.250	0.020	7.074	0.019	7.190	1.053
3.500	0.012	9.233	0.012	9.094	1.000
3.750	0.007	12.009	0.008	11.447	0.875
4.000	0.004	15.572	0.005	14.349	0.800

ตาราง 65 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อน
มาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ชนิด AR แบบเลือกตอบ
และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

e	AR _{s1}		AR _{s2}		AR _{s1} : AR _{s2}
	I (e)	SE (e)	I (e)	SE (e)	
-4.000	0.022	6.782	0.022	6.702	1.000
-3.750	0.032	5.548	0.031	5.663	1.032
-3.500	0.048	4.546	0.046	4.684	1.043
-3.250	0.072	3.727	0.070	3.785	1.029
-3.000	0.107	3.050	0.112	2.993	0.955
-2.750	0.162	2.486	0.184	2.331	0.880
-2.500	0.247	2.012	0.308	1.802	0.802
-2.250	0.383	1.615	0.513	1.396	0.747
-2.000	0.603	1.288	0.837	1.093	0.720
-1.750	0.955	1.023	1.318	0.871	0.725
-1.500	1.507	0.815	1.987	0.709	0.758
-1.250	2.334	0.655	2.856	0.592	0.817
-1.000	3.480	0.536	3.904	0.506	0.891
-0.750	4.877	0.453	5.050	0.445	0.966
-0.500	6.268	0.399	6.112	0.404	1.026
-0.250	7.261	0.371	6.817	0.383	1.065
0.000	7.563	0.364	6.937	0.380	1.090
0.250	7.174	0.373	6.479	0.393	1.107
0.500	6.360	0.397	5.677	0.420	1.120
0.750	5.457	0.428	4.793	0.457	1.139

ตาราง 65 (ต่อ)

θ	AR_{51}		AR_{52}		$AR_{51} : AR_{52}$
	$I(\theta)$	$SE(\theta)$	$I(\theta)$	$SE(\theta)$	
1.000	4.683	0.462	3.985	0.501	1.175
1.250	4.101	0.494	3.302	0.550	1.242
1.500	3.679	0.521	2.739	0.604	1.343
1.750	3.351	0.546	2.272	0.663	1.475
2.000	3.054	0.572	1.878	0.730	1.626
2.250	2.750	0.603	1.540	0.806	1.786
2.500	2.437	0.641	1.248	0.895	1.953
2.750	2.129	0.685	0.999	1.001	2.131
3.000	1.845	0.736	0.789	1.126	2.338
3.250	1.594	0.792	0.616	1.274	2.588
3.500	1.372	0.854	0.477	1.448	2.876
3.750	1.174	0.923	0.367	1.651	3.199
4.000	0.993	1.003	0.281	1.888	3.534

ตาราง 66 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อน
มาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ชนิด AX แบบเลือกตอบ
และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

e	AX ₅₁		AX ₅₂		AX ₅₁ : AX ₅₂
	I (e)	SE (e)	I (e)	SE (e)	
-4.000	0.002	24.909	0.003	19.489	0.667
-3.750	0.002	20.202	0.004	15.228	0.500
-3.500	0.004	16.298	0.007	11.906	0.571
-3.250	0.006	13.069	0.012	9.313	0.500
-3.000	0.009	10.408	0.019	7.291	0.474
-2.750	0.015	8.229	0.031	5.713	0.484
-2.500	0.024	6.459	0.050	4.483	0.480
-2.250	0.039	5.035	0.080	3.528	0.488
-2.000	0.066	3.904	0.129	2.788	0.512
-1.750	0.110	3.018	0.203	2.219	0.542
-1.500	0.183	2.336	0.315	1.782	0.581
-1.250	0.302	1.819	0.479	1.445	0.630
-1.000	0.486	1.435	0.728	1.172	0.668
-0.750	0.750	1.154	1.158	0.929	0.648
-0.500	1.100	0.953	2.009	0.706	0.548
-0.250	1.513	0.813	3.470	0.537	0.436
0.000	1.937	0.718	4.996	0.447	0.388
0.250	2.304	0.659	5.617	0.422	0.410
0.500	2.550	0.626	5.471	0.428	0.466
0.750	2.648	0.615	5.117	0.442	0.517

ตาราง 66 (ต่อ)

๑	AX_{51}		AX_{52}		$AX_{51} : AX_{52}$
	I (๑)	SE (๑)	I (๑)	SE (๑)	
1.000	2.613	0.619	4.617	0.465	0.566
1.250	2.486	0.634	3.975	0.502	0.625
1.500	2.310	0.658	3.313	0.549	0.697
1.750	2.118	0.687	2.715	0.607	0.780
2.000	1.925	0.721	2.195	0.675	0.877
2.250	1.737	0.759	1.746	0.757	0.995
2.500	1.554	0.802	1.364	0.856	1.139
2.750	1.375	0.853	1.046	0.978	1.315
3.000	1.201	0.913	0.791	1.125	1.518
3.250	1.033	0.984	0.590	1.302	1.751
3.500	0.876	1.068	0.437	1.513	2.005
3.750	0.732	1.169	0.321	1.764	2.280
4.000	0.603	1.288	0.235	2.061	2.566

ตาราง 67 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อน
มาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ชนิด CV แบบเลือกตอบ
และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

e	CV _{s1}		CV _{s2}		CV _{s1} : CV _{s2}
	I (e)	SE (e)	I (e)	SE (e)	
-4.000	0.018	7.517	0.001	34.817	18.000
-3.750	0.029	5.891	0.002	25.488	14.500
-3.500	0.047	4.620	0.003	18.565	15.667
-3.250	0.076	3.637	0.006	13.463	12.667
-3.000	0.120	2.881	0.011	9.731	10.909
-2.750	0.189	2.303	0.020	7.024	9.450
-2.500	0.288	1.862	0.039	5.080	7.385
-2.250	0.429	1.526	0.073	3.696	5.877
-2.000	0.618	1.272	0.135	2.720	4.578
-1.750	0.858	1.080	0.241	2.036	3.560
-1.500	1.141	0.936	0.411	1.561	2.776
-1.250	1.450	0.831	0.662	1.229	2.190
-1.000	1.755	0.755	1.012	0.994	1.734
-0.750	2.022	0.703	1.474	0.824	1.372
-0.500	2.221	0.671	2.080	0.693	1.068
-0.250	2.330	0.655	2.855	0.592	0.816
0.000	2.338	0.654	3.702	0.520	0.632
0.250	2.249	0.667	4.293	0.483	0.524
0.500	2.075	0.694	4.336	0.480	0.479
0.750	1.844	0.736	3.915	0.505	0.471

ตาราง 67 (ต่อ)

e	CV ₅₁		CV ₅₂		CV ₅₁ : CV ₅₂
	I (e)	SE (e)	I (e)	SE (e)	
1.000	1.589	0.793	3.302	0.550	0.481
1.250	1.336	0.865	2.672	0.612	0.500
1.500	1.106	0.951	2.095	0.691	0.528
1.750	0.907	1.050	1.599	0.791	0.567
2.000	0.741	1.162	1.193	0.916	0.621
2.250	0.603	1.287	0.875	1.069	0.689
2.500	0.491	1.426	0.634	1.256	0.774
2.750	0.401	1.580	0.455	1.482	0.881
3.000	0.327	1.748	0.325	1.753	1.006
3.250	0.268	1.933	0.232	2.076	1.155
3.500	0.219	2.136	0.165	2.460	1.327
3.750	0.180	2.357	0.118	2.915	1.525
4.000	0.148	2.600	0.084	3.452	1.762

ตาราง 68 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อน
มาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ชนิด LR แบบเลือกตอบ
และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

e	LR ₅₁		LR ₅₂		LR ₅₁ : LR ₅₂
	I (e)	SE (e)	I (e)	SE (e)	
-4.000	0.023	6.577	0.003	18.452	7.667
-3.750	0.035	5.337	0.004	14.999	8.750
-3.500	0.053	4.340	0.007	12.227	7.571
-3.250	0.079	3.552	0.010	10.002	7.900
-3.000	0.116	2.938	0.015	8.215	7.733
-2.750	0.165	2.465	0.022	6.776	7.500
-2.500	0.225	2.108	0.032	5.617	7.031
-2.250	0.295	1.841	0.046	4.681	6.413
-2.000	0.369	1.646	0.065	3.922	5.677
-1.750	0.441	1.507	0.092	3.304	4.793
-1.500	0.502	1.411	0.128	2.797	3.922
-1.250	0.550	1.348	0.177	2.376	3.107
-1.000	0.585	1.308	0.245	2.021	2.388
-0.750	0.611	1.279	0.342	1.710	1.787
-0.500	0.636	1.254	0.490	1.429	1.298
-0.250	0.670	1.222	0.735	1.166	0.912
0.000	0.721	1.178	1.165	0.926	0.619
0.250	0.795	1.122	1.904	0.725	0.418
0.500	0.896	1.056	3.035	0.574	0.295
0.750	1.025	0.988	4.441	0.475	0.231

ตาราง 68 (ต่อ)

e	LR _{s1}		LR _{s2}		LR _{s1} : LR _{s2}
	I (e)	SE (e)	I (e)	SE (e)	
1.000	1.176	0.922	5.744	0.417	0.205
1.250	1.340	0.864	6.528	0.391	0.205
1.500	1.505	0.815	6.680	0.387	0.225
1.750	1.656	0.777	6.377	0.396	0.260
2.000	1.780	0.749	5.825	0.414	0.306
2.250	1.867	0.732	5.131	0.441	0.364
2.500	1.912	0.723	4.347	0.480	0.440
2.750	1.916	0.723	3.540	0.531	0.541
3.000	1.879	0.729	2.781	0.600	0.676
3.250	1.808	0.744	2.121	0.687	0.852
3.500	1.707	0.765	1.581	0.795	1.080
3.750	1.583	0.795	1.160	0.929	1.365
4.000	1.442	0.833	0.840	1.091	1.717

ตาราง 69 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อน
มาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ชนิด NLR แบบเลือกตอบ
และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

e	NLR _{s1}		NLR _{s2}		NLR _{s1} : NLR _{s2}
	I (e)	SE (e)	I (e)	SE (e)	
-4.000	0.023	6.635	0.001	38.562	23.000
-3.750	0.031	5.637	0.001	31.483	31.000
-3.500	0.044	4.792	0.002	25.713	22.000
-3.250	0.060	4.087	0.002	20.984	30.000
-3.000	0.081	3.505	0.003	17.078	27.000
-2.750	0.109	3.029	0.005	13.812	21.800
-2.500	0.143	2.642	0.008	11.043	17.875
-2.250	0.184	2.329	0.013	8.673	14.154
-2.000	0.232	2.074	0.023	6.654	10.087
-1.750	0.287	1.866	0.040	4.978	7.175
-1.500	0.350	1.691	0.075	3.654	4.667
-1.250	0.421	1.540	0.141	2.667	2.986
-1.000	0.507	1.405	0.258	1.970	1.965
-0.750	0.612	1.278	0.445	1.498	1.375
-0.500	0.747	1.157	0.707	1.189	1.057
-0.250	0.918	1.043	1.020	0.990	0.900
0.000	1.132	0.940	1.367	0.855	0.828
0.250	1.386	0.849	1.805	0.744	0.768
0.500	1.666	0.775	2.424	0.642	0.687
0.750	1.944	0.717	3.094	0.569	0.628

ตาราง 69 (ต่อ)

σ	NLR_{s1}		NLR_{s2}		$NLR_{s1} : NLR_{s2}$
	$I(\sigma)$	$SE(\sigma)$	$I(\sigma)$	$SE(\sigma)$	
1.000	2.184	0.677	3.457	0.538	0.632
1.250	2.350	0.652	3.460	0.538	0.679
1.500	2.419	0.643	3.339	0.547	0.724
1.750	2.387	0.647	3.177	0.561	0.751
2.000	2.270	0.664	2.950	0.582	0.769
2.250	2.096	0.691	2.698	0.609	0.777
2.500	1.897	0.726	2.467	0.637	0.769
2.750	1.695	0.768	2.264	0.665	0.749
3.000	1.504	0.815	2.076	0.694	0.724
3.250	1.333	0.866	1.888	0.728	0.706
3.500	1.181	0.920	1.692	0.769	0.698
3.750	1.048	0.977	1.481	0.822	0.708
4.000	0.931	1.037	1.252	0.894	0.744

ตาราง 70 ค่าฟังก์ชันสารสนเทศ ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อน
มาตรฐานของการประมาณค่าความสามารถของแบบทดสอบ ชนิด PI แบบเลือกตอบ
และแบบเติมคำตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5

e	PI ₅₁		PI ₅₂		PI ₅₁ : PI ₅₂
	I (e)	SE (e)	I (e)	SE (e)	
-4.000	0.001	30.602	0.097	3.218	0.010
-3.750	0.002	20.386	0.167	2.444	0.012
-3.500	0.005	13.624	0.285	1.874	0.018
-3.250	0.012	9.142	0.463	1.470	0.026
-3.000	0.026	6.161	0.696	1.198	0.037
-2.750	0.058	4.167	0.943	1.030	0.062
-2.500	0.126	2.820	1.136	0.938	0.111
-2.250	0.277	1.901	1.253	0.893	0.221
-2.000	0.617	1.273	1.393	0.847	0.443
-1.750	1.381	0.851	1.832	0.739	0.754
-1.500	3.021	0.575	3.080	0.570	0.981
-1.250	6.202	0.402	6.283	0.399	0.987
-1.000	11.285	0.298	14.135	0.266	0.798
-0.750	17.201	0.241	32.277	0.176	0.533
-0.500	21.919	0.214	40.561	0.157	0.540
-0.250	23.316	0.207	31.089	0.179	0.750
0.000	19.870	0.224	22.437	0.211	0.886
0.250	13.864	0.269	13.329	0.274	1.040
0.500	8.562	0.342	7.395	0.368	1.158
0.750	4.999	0.447	3.878	0.508	1.289

ตาราง 70 (ต่อ)

e	PI ₅₁		PI ₅₂		PI ₅₁ : PI ₅₂
	I (e)	SE (e)	I (e)	SE (e)	
1.000	2.858	0.592	1.952	0.716	1.464
1.250	1.624	0.785	0.962	1.019	1.688
1.500	0.924	1.040	0.471	1.456	1.962
1.750	0.529	1.375	0.232	2.078	2.280
2.000	0.305	1.809	0.115	2.954	2.652
2.250	0.178	2.368	0.057	4.174	3.123
2.500	0.105	3.082	0.029	5.856	3.621
2.750	0.063	3.990	0.015	8.142	4.200
3.000	0.038	5.138	0.008	11.201	4.750
3.250	0.023	6.583	0.004	15.216	5.750
3.500	0.014	8.395	0.002	20.372	7.000
3.750	0.009	10.663	0.001	26.838	9.000
4.000	0.005	13.490	0.001	34.759	5.000

ภาคผนวก ค

รายชื่อผู้เชี่ยวชาญ

รายชื่อผู้เข้าร่วมงาน

1. รศ.ดร.ปรีชา ธรรมมา ข้าราชการบำนาญ มหาวิทยาลัยศรีนครินทรวิโรฒ
ประธานมิตร
2. นายเอี่ยม โตบุญเลี้ยง ข้าราชการบำนาญ กรมวิชาการ
3. ดร.บุญลือ ทองอยู่ รองเลขาธิการ ก.ค.
4. รศ.วิญญา วิศาลาภรณ์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประธานมิตร
5. ดร.สุวพร เข้มแข็ง สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี
6. ดร.ผดุงชัย ภูพัฒน์ เขตการศึกษา 7 กรมสามัญศึกษา
7. รศ.ดร.สุรศักดิ์ อมรรัตนศักดิ์ มหาวิทยาลัยรามคำแหง
8. ผศ.ดร.บุญมี พันธุ์ไทย มหาวิทยาลัยรามคำแหง
9. ผศ.ดร.สุนันท์ ศัลโกสม มหาวิทยาลัยศรีนครินทรวิโรฒ ประธานมิตร
10. ผศ.จันทิมา พรหมโชติกุล มหาวิทยาลัยศรีนครินทรวิโรฒ ประธานมิตร
11. ดร.วรรณดี แสงประทีปทอง มหาวิทยาลัยสุโขทัยธรรมาธิราช
12. นางสาวสายสมร ชำติยานนท์ สำนักงานทดสอบทางการศึกษา กรมวิชาการ
13. นางประไพ ประสาธพันธ์ สำนักงานทดสอบทางการศึกษา กรมวิชาการ
14. นางสาวประณีต เลิศไกร โรงเรียนสาธิต มหาวิทยาลัยศรีนครินทรวิโรฒ
ประธานมิตร

ประวัติย่อของผู้วิจัย

ชื่อ นางบุญชู ชลัชเรีัยร
 วัน เดือน ปีเกิด 18 มกราคม 2493
 สถานที่เกิด จังหวัดนครราชสีมา
 ที่อยู่ปัจจุบัน 19/7 หมู่ 24 ซอยประตู 1 บางระมาด เขตคลองเตย กทม.10170
 ตำแหน่งหน้าที่การงานปัจจุบัน นักวิชาการสอบ ระดับ 7
 สถานที่ทำงานปัจจุบัน สำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ
 เขตดุสิต กทม.10300
 ประวัติการศึกษา พ.ศ.2508 ม.ศ.3 โรงเรียนสุรนารีวิทยา นครราชสีมา
 พ.ศ.2510 ม.ศ.5 โรงเรียนสุรนารีวิทยา นครราชสีมา
 พ.ศ.2514 ศศ.บ.(จิตวิทยา) เกียรตินิยมดี มหาวิทยาลัยธรรมศาสตร์
 พ.ศ.2522 กศ.ม. การวัดผลและประเมินผลการศึกษา
 พ.ศ.2539 กศ.ค. การทดสอบและวัดผลการศึกษา

การพัฒนาการวัดความสามารถด้านเหตุผลเชิงวิเคราะห์
ของนักเรียนระดับชั้นมัธยมศึกษา

บทคัดย่อ

ของ

บุญชู ชัลลัษเฐียร

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรการศึกษาคณะศึกษาศาสตร์ สาขาการทดสอบและวัดผลการศึกษา
สิงหาคม 2539

การวิจัยครั้งนี้ มีวัตถุประสงค์เพื่อพัฒนาการวัดความสามารถด้านเหตุผลเชิงวิเคราะห์ของนักเรียน ระดับชั้นมัธยมศึกษา ทั้งนี้โดยการพัฒนาข้อคำถามขึ้น 6 ชนิด คือ เหตุผลด้านวิเคราะห์ (AR) การวิเคราะห์คำอธิบาย (AX) ความเห็นตรงกันข้าม (CV) เหตุผลทางตรรก (LR) เหตุผลทางตรรกเกี่ยวกับจำนวน (NLR) และการระบุรูปแบบ (PI) ข้อคำถามแต่ละชนิดจัดทำเป็น 2 แบบ คือ แบบเลือกตอบ 5 ตัวเลือก และแบบเติมคำตอบ การศึกษาคุณภาพของข้อคำถามและแบบทดสอบใช้การประมาณค่าคุณลักษณะข้อสอบ เพื่อหาค่าความยาก ค่าอำนาจจำแนก และค่าสัมประสิทธิ์การเดา คำนวณค่าฟังก์ชันสารสนเทศของแบบทดสอบ และเปรียบเทียบคุณภาพของแบบทดสอบแต่ละรูปแบบด้วยการคำนวณค่าประสิทธิภาพสัมพัทธ์ ผลการวิจัย พบว่า แบบทดสอบรูปแบบเติมคำตอบเกือบทุกชนิดข้อคำถามมีค่าอำนาจจำแนกเฉลี่ยสูงกว่าแบบเลือกตอบ ส่วนค่าความยากนั้น พบว่า แบบทดสอบชนิด AR, AX, LR และ PI รูปแบบเลือกตอบจะมีค่าความยากเฉลี่ยสูงกว่าแบบเติมคำตอบ แต่ชนิด NLR และ CV รูปแบบเติมคำตอบ ค่าความยากเฉลี่ยสูงกว่ารูปแบบเลือกตอบ และพบอีกเช่นกันว่า แบบทดสอบชนิด AR, AX, LR และ PI รูปแบบเลือกตอบ มีค่าสัมประสิทธิ์การเดาสูงกว่าแบบเติมคำตอบ ขณะที่ชนิด NLR และ CV รูปแบบเติมคำตอบมีค่าสัมประสิทธิ์การเดาสูงกว่าแบบเลือกตอบ

จากการศึกษาค่าความเชื่อมั่นของแบบทดสอบ พบว่า แบบทดสอบรูปแบบเติมคำตอบเกือบทุกชนิดให้ค่าความเชื่อมั่นสูงกว่ารูปแบบเลือกตอบ และผลการศึกษาค่าฟังก์ชันสารสนเทศของแบบทดสอบชี้ให้เห็นว่า แบบทดสอบรูปแบบเลือกตอบส่วนใหญ่ให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงสุด ที่ระดับความสามารถปานกลาง ขณะที่รูปแบบเติมคำตอบให้ค่าฟังก์ชันสารสนเทศสูงสุดที่ระดับความสามารถสูง ทั้งนี้ยังพบว่า ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของแบบทดสอบทุกชนิด ทั้ง 2 รูปแบบ มีค่าสูงสุดที่ระดับความสามารถต่ำ แต่จะมีค่าต่ำ ณ ระดับความสามารถปานกลางและสูง

ประสิทธิภาพของแบบทดสอบที่ได้จากการศึกษาค่าประสิทธิภาพสัมพัทธ์ พบว่า แบบทดสอบรูปแบบเติมคำตอบเกือบทุกชนิด มีประสิทธิภาพสูงกว่าแบบเลือกตอบเมื่อใช้กับกลุ่มที่มีความสามารถสูงและปานกลาง สำหรับแบบเลือกตอบมีประสิทธิภาพสูงกว่าแบบเติมคำตอบเมื่อใช้กับกลุ่มที่มีความสามารถต่ำ

ผลการวิจัยครั้งนี้ ให้ข้อเสนอแนะที่ควรได้รับการพิจารณา คือ ข้อคำถามชนิดต่าง ๆ ที่ได้พัฒนาขึ้นทั้ง 6 ชนิด มีลักษณะน่าสนใจ และท้าทายผู้สอบ เนื่องจากมีลักษณะที่ต่างไปจากข้อคำถาม ซึ่งผู้สอบเคยพบในการสอบปกติ ผู้สอบส่วนใหญ่จึงให้ความสนใจและทำแบบทดสอบอย่างเต็มความสามารถ ซึ่งจะมีผลกับเจตคติต่อการทำแบบทดสอบ การศึกษาคุณภาพของ

เครื่องมือและผลการสอบ สำหรับประเด็นเกี่ยวกับรูปแบบของข้อคำถามนั้น ถือได้ว่ารูปแบบ
 เติมคำตอบ เป็นรูปแบบที่มีประสิทธิภาพสูงในการประมาณค่าความสามารถของผู้สอบ เมื่อเทียบกับ
 กับรูปแบบเลือกตอบ อย่างไรก็ตาม ประสิทธิภาพที่พึงประสงค์ดังกล่าวจะเกิดขึ้นเมื่อผู้สร้าง
 แบบทดสอบหรือผู้ใช้แบบทดสอบ พิจารณานำไปใช้ให้เหมาะสมกับจุดมุ่งหมายของงานและกลุ่ม
 เป้าหมาย ทั้งนี้เพราะรูปแบบของข้อคำถาม และข้อคำถามถือเป็นงานที่กำหนดให้เพื่อผู้สอบ
 หาทางแก้ปัญหาหรือหาคำตอบ ซึ่งลักษณะงานบางอย่างอาจเหมาะกับกลุ่มผู้สอบที่มีความสามารถ
 สูง ขณะที่ลักษณะงานบางอย่าง อาจเหมาะกับกลุ่มผู้มีความสามารถระดับปานกลางหรือต่ำ
 ฉะนั้น ผู้สร้างหรือผู้ใช้แบบทดสอบจำเป็นต้องพิจารณาเลือกใช้ให้เหมาะสม

DEVELOPMENT OF THE ANALYTICAL REASONING ABILITY MEASURE
FOR THE SECONDARY SCHOOL STUDENTS

AN ABSTRACT

BY

BOONCHOO CHALASSATHIEN

Presented in partial fulfillment of the requirements for the Doctor
of Education Degree in Testing and Educational Measurement
at Srinakharinwirot University

August 1996

The research focused on the developing of analytical reasoning ability tests for the secondary school students. Six item - type ; analytical reasoning , analysis of explanation , contrasting view , logical reasoning , numerical logical reasoning and pattern identification with two item formats ; 5 option multiple choice and open-ended test items were used. IRT (3-parameter model) was applied to compute item parameters (a,b and c). The results indicated that most of the item type of supply format had higher discriminating power than multiple choice format. For difficulty, multiple choice format of AR , AX , LR and PI had higher item difficulty than the supply-type format, while NLR and CV supply-type format were more difficult than multiple choice format.

Results further indicated that guessing coefficient of AR, AX, LR and PI multiple choice format were higher than supply type format, while NLR and CV supply-type format had guessing coefficient higher than multiple choice format.

Reliabilities of most of supply-type format were higher than multiple choice format. On the other hand, it was found that multiple choice format had the highest TIF at average ability group, but showed the lowest TIF at the average and high ability group.

Finally the relative efficiency was introduced to investigate the effectiveness of the item format. The results indicated that the quality of supply-type format was higher than the multiple choice format at high ability group and for the low ability group the multiple choice format was more suitable.

The results suggested that the six item-type which were interesting and challenged for the examinees became of their feature that differ from those commonly used. Most examinees were interested and had taken the tests with their full capacity

which provided good attitude in taking the test, the investigation the quality of the test and test results.

The study also provided evidence suggesting that the supply-type items are superior in ability estimation when comparing with multiple-choice items. However, the desired efficiency will be occurred when test developers or test users have considered the application concern with work objective and the target group carefully. As item format and item are the tasks that were given to the examinees to find the solution or answer, some task may be appropriate for high ability group while the other are appropriate for average or low ability. Therefore the test developers or the test users should considered their application before making decision.