การเปรียบเทียบคุณลักษณะของแบบทดสอบเลือกตอบระหว่างกลุ่มตัวอย่าง ที่มีจำนวนต่างกันที่วิเคราะห์ด้วยราส์ชโมเดล

ปริญญานิพนธ์

ของ

สุพล นิลกลาง

ก.ค. 2541

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
พฤษภาคม 2541

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

คณะกรรมการควบคุมและคณะกรรมการสอบได้พิจารณาปริญญานิพนธ์ฉบับนี้แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒได้

คณะกรรมการควบคุม

.....ประธาน
(รองศาสตราจารย์อังคณา สายยศ)

.....กรรมการ
(รองศาสตราจารย์ชูศรี วงศ์รัตนะ)

คณะกรรมการสอบ

.....ประธาน
(รองศาสตราจารย์อังคณา สายยศ)

.....กรรมการ
(รองศาสตราจารย์ชูศรี วงศ์รัตนะ)

.....กรรมการที่แต่งตั้งเพิ่มเติม
(รองศาสตราจารย์อาวุธ วัฒนสิน)

บัณฑิตวิทยาลัยอนุมัติให้รับปริญญานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒ

.....คณบดีบัณฑิตวิทยาลัย
(ศาสตราจารย์ ดร. เสริมศักดิ์ วิศาลาภรณ์)

วันที่ 15 เดือน พฤษภาคม พ.ศ. 2541

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้ สำเร็จลุล่วงไปได้ด้วยความช่วยเหลือ แนะนำ ให้ข้อคิดเห็น และแก้ไขข้อบกพร่องต่าง ๆ อย่างดียิ่งจาก รองศาสตราจารย์อังคณา สายยศ และรองศาสตราจารย์ชูศรี วงศ์รัตนะ ผู้วิจัยรู้สึกซาบซึ้งและขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณผู้เชี่ยวชาญทุกท่านที่ให้ความช่วยเหลือแนะนำ ให้ข้อคิดเห็นและแก้ไขข้อบกพร่องในการสร้างเครื่องมือ

ขอขอบพระคุณผู้บริหารโรงเรียนและครู-อาจารย์ของโรงเรียนที่เป็นกลุ่มตัวอย่าง รวมทั้งนักเรียนทุกคนที่ให้ความร่วมมือในการเก็บรวบรวมข้อมูล

ขอขอบคุณอาจารย์กิตติธัช ทรงศิริ อาจารย์ถาวร กิตติวรนารถ อาจารย์สรรเพชญ ทุนมาก อาจารย์ไพริน นิลกลาง และเพื่อน ๆ ที่ให้คำแนะนำและช่วยเหลือในการเก็บรวบรวมข้อมูลและวิเคราะห์ข้อมูล

สุดท้ายนี้ผู้วิจัยขอน้อมรำลึกถึงพระคุณบิดา มารดา ตลอดจนญาติพี่น้องที่ได้สนับสนุนตลอดมา และขอน้อมรำลึกถึงพระคุณของครู-อาจารย์ที่ให้การอบรมสั่งสอนและสนับสนุนการศึกษาของผู้วิจัยตั้งแต่อดีตถึงปัจจุบัน

สุพล นิลกลาง

สารบัญ

บทที่		หน้า
1	บทนำ	1
	ภูมิหลัง	1
	ความมุ่งหมายของการศึกษาค้นคว้า	3
	ความสำคัญของการศึกษาค้นคว้า	4
	ขอบเขตของการศึกษาค้นคว้า	4
	นิยามศัพท์เฉพาะ	5
2	เอกสารงานวิจัยที่เกี่ยวข้อง	7
	1 ทฤษฎีการตอบข้อคำถาม	7
	1.1 พื้นฐานความเข้าใจในทฤษฎีการตอบข้อคำถาม	7
	1.2 ข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อคำถาม	9
	1.3 พิสัยค่าพารามิเตอร์	10
	1.4 โมเดลของทฤษฎีการตอบข้อคำถาม	11
	2 ราส์ชโมเดล	16
	3 วิธีการประมาณค่าพารามิเตอร์	20
	4 ดัชนีบ่งชี้ประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์	25
	5 งานวิจัยที่เกี่ยวข้อง	29
3	วิธีการดำเนินการศึกษาค้นคว้า	37
	ประชากร	37
	กลุ่มตัวอย่าง	37
	เครื่องมือที่ใช้ในการศึกษาค้นคว้า	38
	วิธีดำเนินการสร้างแบบทดสอบ	39
	ลักษณะของแบบทดสอบ	41
	วิธีดำเนินการเก็บข้อมูล	41

บทที่ หน้า

การวิเคราะห์ข้อมูล 42
สถิติที่ใช้วิเคราะห์ข้อมูล 42

4 ผลการวิเคราะห์ข้อมูล 44
สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล 44
การเสนอผลการวิเคราะห์ข้อมูล 44
ผลการวิเคราะห์ข้อมูล 45

5 สรุป อภิปรายผล และข้อเสนอแนะ 50
ความมุ่งหมายของการศึกษาค้นคว้า 50
กลุ่มตัวอย่าง 50
เครื่องมือที่ใช้ในการศึกษาค้นคว้า 50
วิธีดำเนินการเก็บรวบรวมข้อมูล 50
การวิเคราะห์ข้อมูล 51
สรุปผลการวิเคราะห์ข้อมูล 51
อภิปรายผล 53
ข้อเสนอแนะ 54

บรรณานุกรม 55
ภาคผนวก .....
ภาคผนวก ก 62
ภาคผนวก ข 74
ภาคผนวก ค 81
ภาคผนวก ง 88
ภาคผนวก จ 95
รายชื่อผู้เชี่ยวชาญ 97
ประวัติย่อของผู้วิจัย 98

บัญชีตาราง

ตาราง		หน้า
1	ความแตกต่างระหว่างความสามารถของคนกับความยากของข้อสอบ	19
2	แสดงค่าคุณสมบัติของแบบทดสอบ 5 ข้อ	27
3	ตารางการสุ่มตัวอย่างที่ใช้ในการวิจัย	38
4	ค่าความยากของแบบทดสอบของกลุ่มตัวอย่างจำนวน 200 คน 300 คน 50 คน 675 คน 1,015 คน และ 1,610 คน	45
5	ค่าความสามารถของผู้สอบของกลุ่มตัวอย่างจำนวน 200 คน, 300 คน 450 คน, 675 คน 1,015 คน และ 1,610 คน	46
6	ค่าสถิติพื้นฐานของความคลาดเคลื่อนมาตรฐานของแบบทดสอบของ ลุ่มผู้สอบ 200 คน, 300 คน, 450 คน, 675 คน, 1,015 คน และ 1,610 คน	47
7	การทดสอบความแตกต่างของค่าความยากของข้อสอบของกลุ่มผู้สอบ จำนวน 200 คน, 300 คน, 450 คน, 675 คน, 1,015 คน และ 1,610 คน	48
8	การทดสอบความแตกต่างของค่าความสามารถของข้อสอบของกลุ่มผู้สอบ จำนวน 200 คน, 300 คน, 450 คน, 675 คน, 1,015 คน และ 1,610 คน	48
9	การทดสอบความแตกต่างของค่าความสามารถของข้อสอบของกลุ่มจำนวน 200 คน, 300 คน, 450 คน, 675 คน, 1,015 คน และ 1,610 คน	49
10	ค่าความยากของแบบทดสอบ กลุ่มตัวอย่าง 200 คน	75
11	ค่าความยากของแบบทดสอบ กลุ่มตัวอย่าง 300 คน	76
12	ค่าความยากของแบบทดสอบ กลุ่มตัวอย่าง 450 คน	77
13	ค่าความยากของแบบทดสอบ กลุ่มตัวอย่าง 675 คน	78
14	ค่าความยากของแบบทดสอบ กลุ่มตัวอย่าง 1,015 คน	79
15	ค่าความยากของแบบทดสอบ กลุ่มตัวอย่าง 450 คน	80
16	แสดงค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของ ผู้สอบของกลุ่มตัวอย่าง 200 คน	82
17	แสดงค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของ ผู้สอบของกลุ่มตัวอย่าง 300 คน	83
18	แสดงค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของ ผู้สอบของกลุ่มตัวอย่าง 450 คน	84

ตาราง หน้า

19 แสดงค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบของกลุ่มตัวอย่าง 675 คน 85
20 แสดงค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบของกลุ่มตัวอย่าง 1,015 คน 86
21 แสดงค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบของกลุ่มตัวอย่าง 1,610 คน 87
22 แสดงค่าไอเก้นของตัวประกอบ 12 ตัว 89

บัญชีภาพประกอบ

ภาพประกอบ		หน้า
1	แสดงโค้งลักษณะการตอบรายข้อ	7
2	แสดงค่าพารามิเตอร์เป็นรายข้อ	8
3	แสดงโมเดลโค้งปกติสะสมที่มีพารามิเตอร์ตัวเดียว	11
4	แสดงโมเดลโค้งปกติสะสมที่มีพารามิเตอร์สองตัว	12
5	แสดงโมเดลโค้งปกติสะสมที่มีพารามิเตอร์สามตัว	12
6	แสดงความหมายของค่า Item Prarmeter ของ One-Paremeter Logistic Model	13
7	แสดงความหมายของค่า Item Prarmeter ของ Two-Paremeter Logistic Model	14
8	แสดงความหมายของค่า Item Prarmeter ของ Three-Paremeter Logistic Model	15
9	แสดงมโนทัศน์เกี่ยวกับราส์ชโมเดล	17
10	แสดงความแตกต่างระหว่างความสามารถของคนกับความยากของข้อสอบ	18
11	Information Curve ของแบบทดสอบความถนัดด้านภาษาของสถาบัน ETS	27

บทที่ 1
บทนำ

ภูมิหลัง

ในการจัดการศึกษาพบว่า การวัดผลการศึกษาเป็นองค์ประกอบที่สำคัญ เนื่องจากเป็นส่วนที่ทำให้ทราบว่า การจัดการศึกษาให้กับผู้เรียนนั้น ประสบความสำเร็จมากน้อยเพียงใด เพราะการวัดผลและการประเมินผลการศึกษา เป็นกระบวนการที่จำเป็นอย่างยิ่งในการจัดการศึกษาความรู้ที่ได้จากการวัดผลและประเมินผล ได้นำมาใช้ภายในโรงเรียนอยู่ตลอดเวลา ทั้งยังช่วยให้การตัดสินใจเกี่ยวกับการให้การศึกษาแก่นักเรียนเป็นไปอย่างถูกต้อง และเชื่อถือได้ (สมเกียรติ ปดิฐพร. 2525 : 1) การวัดผลการศึกษานั้นสามารถกระทำได้หลายวิธี แต่ที่นิยมกันมากคือ การวัดผลโดยการใช้แบบทดสอบ (Tyler. 1977 : 26) ซึ่งแบบทดสอบที่ใช้ จะเป็นเครื่องมือที่ใช้วัดหรือใช้ตรวจสอบคุณลักษณะภายใน (Traits) ของบุคคลที่เป็นนามธรรม และมีองค์ประกอบที่สลับซับซ้อน ไม่สามารถสังเกตได้โดยตรง การวัดต้องใช้ทางอ้อม โดยวัดสิ่งที่มีความสัมพันธ์เชิงปริมาณกับสิ่งที่ต้องการจะวัด โดยอาศัยแบบทดสอบไปกระตุ้นให้บุคคลตอบสนองออกมาเป็นพฤติกรรมภายนอก ผลที่วัดได้จะอยู่ในรูปของคะแนน (Lord. 1980 : 3) ซึ่งจะนำไปใช้ในการอธิบายหรือทำนายความสามารถที่แท้จริง อันเป็นคุณลักษณะภายใน (Latent traits) ของบุคคล ผลการวัดจะถูกต้องเชื่อถือได้ และมีประสิทธิภาพเพียงใด ย่อมขึ้นอยู่กับคุณภาพของแบบทดสอบที่ใช้วัด แบบทดสอบจึงจำเป็นต้องมีคุณภาพที่เชื่อถือได้ มีความคลาดเคลื่อนในการวัดน้อยที่สุด และวัดได้ตรงตามความสามารถที่แท้จริงที่ต้องการวัด (อัวยพร วิบูลย์กาญจน์. 2526 : 7) ด้วยความสำคัญของแบบทดสอบดังกล่าว จึงได้มีการสร้างและพัฒนาแบบทดสอบขึ้นมาเพื่อนำไปใช้ในโอกาสต่าง ๆ กัน โดยจำแนกแบบทดสอบไปตามเนื้อหาวิชาที่จะวัด และเน้นหนักคุณภาพของแบบทดสอบทางด้านค่าความยาก ค่าความเชื่อมั่น ค่าอำนาจจำแนก และค่าความเที่ยงตรง (Tyler. 1977 : 4) และสิ่งที่สำคัญอย่างยิ่งในการที่จะพัฒนาแบบทดสอบให้มีคุณภาพดังกล่าวก็คือ ทฤษฎีที่ใช้วิเคราะห์ข้อสอบ ซึ่งทฤษฎีที่ใช้อยู่ในปัจจุบันนั้นมีอยู่ 2 ทฤษฎี คือ ทฤษฎีแบบดั้งเดิม (Classical Test Theory) และทฤษฎีการตอบข้อถาม (Item Response Theory) (สุพัฒน์ สุกมลสันต์. 2530 : 1-4) ทฤษฎีแบบดั้งเดิม เป็นทฤษฎีที่ประมาณค่าความสามารถที่แท้จริง โดยใช้คะแนนของกลุ่มผู้สอบเป็นหลัก ซึ่งถูกมองว่าเป็นทฤษฎีที่ไม่เพียงพอต่อการประมาณค่าความสามารถที่แท้จริง และทฤษฎีนี้ไม่สามารถทำนายได้ว่า บุคคลจะตอบข้อสอบอย่างไร ยกเว้นแต่ได้ใช้ข้อสอบเหล่านั้นกับบุคคลที่คล้ายคลึงกันมาก่อน (ผจงจิต อินทรสุวรรณ. 2528 : 13) และถ้ากลุ่มตัวอย่างมีขนาดเล็ก ก็จะทำให้ค่าความยาก

ง่ายสูงเกินความจำเป็น และทำให้ค่าอำนาจจำแนกมีความคลาดเคลื่อน (Henryson. 1971 : 148) ซึ่ง แฮมเบลตั้น (Hambleton. 1978) ได้กล่าวถึงจุดอ่อนของทฤษฎีแบบดั้งเดิมไว้หลายประการ คือ

1. คุณสมบัติของข้อคำถามของแบบทดสอบไม่คงที่ แปรเปลี่ยนไปตามกลุ่มผู้สอบ

2. ในการเปรียบเทียบความสามารถของผู้สอบต่าง ๆ นั้น จะต้องใช้แบบทดสอบที่มีข้อคำถามเหมือนกัน หรือเป็นแบบทดสอบคู่ขนานกัน

3. ข้อมูลที่ได้จากการทดสอบตามทฤษฎีดั้งเดิมนั้น ไม่สามารถบอกได้ว่าผู้สอบมีความสามารถที่จะทำอะไรได้บ้าง หรือมีความบกพร่องตรงไหน

นอกจากนี้ สงบ ลักษณะ (2525 : 48-49) ยังได้กล่าวถึงปัญหาการพัฒนาแบบทดสอบและแปลความหมายคะแนน โดยยึดทฤษฎีดั้งเดิม ดังนี้

1. คะแนนผลการสอบ (Observed Score) ไม่สามารถประมาณค่าความสามารถที่แท้จริง หรือคะแนนจริงที่ชัดเจนได้

2. คะแนนของผู้สอบคนเดียวกันจากแบบทดสอบคู่ขนานสองชุด ที่วัดความสามารถเดียวกัน จะต่างกันทั้ง ๆ ที่สอบในเวลาใกล้เคียงกัน

3. ในการประมาณค่าคะแนนจริงตามสูตรของทฤษฎีดั้งเดิม ที่ได้จากการนำคะแนนผลการสอบลบด้วยคะแนนความคลาดเคลื่อนในการวัด ซึ่งคะแนนความคลาดเคลื่อนในการวัดเป็นค่าผันแปรไปตามฉบับของแบบทดสอบและลักษณะของผู้สอบ

ส่วน แฮมเบลตั้น (Hambleton. 1979 : 14-15) ได้กล่าวว่าความคลาดเคลื่อนมาตรฐานในการวัดตามทฤษฎีดั้งเดิมของแต่ละคนนั้น เดิมกำหนดว่าเท่ากัน แต่จริง ๆ นั้น ความคลาดเคลื่อนมาตรฐานในการวัดของผู้สอบแต่ละคน จะไม่เท่ากัน

จะเห็นว่าทฤษฎีดั้งเดิมไม่สามารถแก้ปัญหาดังกล่าวได้ จึงมีนักวัดผลและนักวิชาการได้ช่วยกันคิด และสร้างทฤษฎีขึ้นมาใหม่ เรียกว่า ทฤษฎีการตอบข้อคำถาม (Item Response Theory) (สุรศักดิ์ อมรรัตน์ศักดิ์. 2530 : 198) ซึ่งเป็นทฤษฎีที่ใช้โมเดลทางคณิตศาสตร์มาแสดงความสัมพันธ์ระหว่างความสามารถที่แท้จริง กับการตอบข้อคำถาม (สงบ ลักษณะ. 2525 :119) และถ้าได้นำทฤษฎีนี้มาใช้ให้เป็นไปตามข้อตกลงเบื้องต้น แล้วสามารถแก้ปัญหาดังกล่าวที่เกิดจากทฤษฎีดั้งเดิมได้ (Hambleton. 1979 : 16-22) และคำนวณค่าพารามิเตอร์ของข้อสอบซึ่งได้แก่ ค่าความยาก ค่าอำนาจจำแนก และสัมประสิทธิ์การเดา และจะได้ค่าที่ไม่แปรเปลี่ยนไปตามกลุ่มตัวอย่าง ไม่ว่าจะนำกลับไปสอบกับกลุ่มตัวอย่างใด และการประมาณค่าความสามารถของผู้สอบนั้น จะใช้ข้อสอบชุดใด จำนวนเท่าใดก็ได้ ไปประมาณค่าความสามารถของผู้สอบนั้น

นอกจากนี้ โกวิทย์ ประวาลพฤกษ์ (2525 : 63) ได้กล่าวถึงคุณสมบัติของข้อสอบที่วิเคราะห์ด้วยทฤษฎีการตอบข้อคำถามว่า ว่าสามารถเปรียบเทียบค่ายากง่ายได้โดยอิสระและเป็น

สถิติที่เพียงพอสำหรับการใช้อธิบายคุณภาพของแบบทดสอบ นอกจากนั้น แฮมเบลตั้น (Hambleton. 1985) เบรินบอม (Birnbaum. 1958) ลอร์ด (Lord. 1989) ราส์ช (Rash. 1982 : 4799-A) วาร์ม (Warm. 1979) ต่างก็ยืนยันว่า ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบข้อคำถามนั้นไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบที่แตกต่างกัน แต่ถึงอย่างไรก็ตามลอร์ด (Lord. 1980 : 38) ยังเน้นว่าเป็นการยากที่จะอธิบายให้เชื่อได้ง่าย ๆ ว่า ค่าพารามิเตอร์ของข้อสอบจะไม่แปรปรวนระหว่างกลุ่มผู้สอบที่มีระดับความสามารถต่างกัน และยังมีนักการศึกษาอีกหลายท่านได้ทำการศึกษาเกี่ยวกับค่าพารามิเตอร์ของข้อสอบ ดังนี้

คุก และทาฟท์ (Cook and Talf. 1988 : 31-34) ฮัมเบิร์ท (Humbert. 1986 : 3013-A) ได้ทำการศึกษาความคงที่ของพารามิเตอร์ ที่วิเคราะห์ด้วยทฤษฎีการตอบข้อคำถาม ผลการศึกษาพบว่าค่าพารามิเตอร์จะแปรเปลี่ยนเมื่อระยะเวลา หรือความยาวของแบบทดสอบที่มีความยาวต่างกัน ส่วนในประเทศไทย บุญชิด ร้อดแก้ว (2533 : 75) และสมพร บุญอิ่ม (2529 : 92) ได้ศึกษาค่าพารามิเตอร์ของข้อสอบ พบว่าค่าพารามิเตอร์ของข้อสอบจะแปรปรวน เมื่อลักษณะของกลุ่มผู้สอบต่างกัน ส่วนค๊อค (Kock. 1980 : 15-32) ทินสเลย์ และเดวิส (Tinsley and Dewis. 1975 : 325-337) ดูแกลส (Douglass. 1981 : 400-A) ต่างก็ยืนยันถึงความไม่แปรปรวนของค่าพารามิเตอร์ไปตามลักษณะของกลุ่มตัวอย่าง และในประเทศไทย ปนัดดา วัยวัฒนะ (2528 : 71) บัญชา แสนทวี (2530 : 110) ก็ยืนยันถึงความไม่แปรปรวนของค่าพารามิเตอร์เช่นกัน

จากคำกล่าวและผลการศึกษาข้างต้นจะเห็นว่า ผลการศึกษาและทฤษฎียังมีการขัดแย้งกันอยู่ และในประเทศไทยก็ยังมีการศึกษาในเรื่องนี้น้อยมาก และยังไม่มีผู้ใดทำการศึกษาในเรื่องคุณลักษณะของแบบทดสอบที่วิเคราะห์ด้วยราส์ชโมเดล เมื่อจำนวนกลุ่มตัวอย่างต่างกัน ผู้วิจัยจึงมีความสนใจที่จะทำการศึกษาในเรื่องนี้ ซึ่งผลการศึกษาจะทำให้ทราบว่า จำนวนกลุ่มผู้สอบจำนวนเท่าไรที่ทำให้คุณลักษณะของข้อสอบไม่แตกต่างกัน เพื่อเป็นแนวทางในการเลือกใช้กลุ่มตัวอย่างที่พอเหมาะในการวิเคราะห์ข้อสอบด้วยราส์ชโมเดล อันจะเป็นประโยชน์และพัฒนาในด้านวัดผลต่อไป

ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อเปรียบเทียบค่าความยากของแบบทดสอบ ที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน
2. เพื่อเปรียบเทียบค่าความสามารถของผู้สอบ ที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน

3. เพื่อเปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน

ความสำคัญของการศึกษาค้นคว้า

ผลจากการศึกษาครั้งนี้ จะทำให้ทราบว่า ค่าความยาก ค่าความคลาดเคลื่อนมาตรฐานและคะแนนความสามารถของผู้สอบ จากแบบทดสอบวิชาคณิตศาสตร์ เมื่อจำนวนผู้สอบต่างกัน จะได้ผลต่างกันหรือไม่ ถ้าต่างกันจะทำให้ทราบว่า ผู้สอบจำนวนเท่าไรที่ทำให้ค่าความยากกับคะแนนความสามารถต่างกัน หรือจำนวนผู้สอบจำนวนเท่าไรที่ให้ผลไม่แตกต่างกัน และจำนวนผู้สอบจำนวนเท่าไรที่ทำให้ค่าความคลาดเคลื่อนมาตรฐานมีค่าน้อยที่สุด อันจะเป็นประโยชน์ในการเลือกขนาดกลุ่มตัวอย่างที่เหมาะสมในการวิเคราะห์หาคุณลักษณะของข้อสอบที่ใช้วิธีราส์ชโมเดลต่อไป

ขอบเขตของการศึกษาค้นคว้า

1. ประชากร

ประชากรที่ใช้ในการศึกษาครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2540 สังกัดโรงเรียนกรมสามัญศึกษากลุ่ม 7 ในจังหวัดกรุงเทพมหานคร จำนวน 14 โรงเรียน จำนวนห้องเรียน 202 ห้องเรียน จำนวนนักเรียน 9,090 คน

2. กลุ่มตัวอย่าง

กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ปีการศึกษา 2540 ของโรงเรียนสังกัดกรมสามัญ ในกลุ่มที่ 7 ของจังหวัดกรุงเทพมหานคร จำนวน 1,610 คน โดยใช้วิธีการสุ่มแบ่งชั้น (Stratied Random Sampling)

3. ตัวแปรที่ใช้ในการศึกษาค้นคว้า

3.1 ตัวแปรอิสระ ได้แก่ จำนวนของกลุ่มตัวอย่างที่ต่างกัน โดยมีอัตราส่วนการเพิ่ม 50% คือ

กลุ่มที่ 1	200 คน
กลุ่มที่ 2	300 คน
กลุ่มที่ 3	450 คน
กลุ่มที่ 4	675 คน
กลุ่มที่ 5	1,015 คน
กลุ่มที่ 6	1,610 คน

3.2 ตัวแปรตาม คือคุณลักษณะของแบบทดสอบ ซึ่งประกอบด้วย

3.2.1 ค่าความยาก

3.2.2 ค่าความสามารถของผู้สอบ

3.2.3 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ

นิยามศัพท์เฉพาะ

1. ทฤษฎีการตอบข้อคำถาม (Item Response Theory : IRT) หมายถึง ทฤษฎีที่อธิบายความน่าจะเป็นในการตอบข้อคำถามได้ถูกต้อง ในรูปความสัมพันธ์ของความสามารถที่แท้จริงของผู้สอบ

2. การวิเคราะห์ด้วยราส์ชโมเดล หมายถึง การวิเคราะห์ข้อสอบด้วยโมเดลของราส์ชที่ใช้ทฤษฎีการตอบข้อคำถาม ซึ่งประกอบด้วย ค่าพารามิเตอร์ 1 ตัว คือค่าความยาก (b) ในการศึกษาครั้งนี้วิเคราะห์ด้วยโปรแกรม BICAL

3. คุณลักษณะของแบบทดสอบ หมายถึง ค่าที่แสดงถึงลักษณะของผู้สอบและลักษณะของข้อสอบ ในการศึกษาครั้งนี้ศึกษาค่าความยาก ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ และคะแนนความลามารถของผู้สอบ ซึ่งได้จากการวิเคราะห์ด้วยราส์ชโมเดล

4. ค่าความยากของข้อสอบ (Item Difficulty : b) หมายถึง ค่าความสามารถตรงจุดเปลี่ยนโค้ง ลักษณะของข้อสอบแต่ละข้อ (Item Characteristic Curve : ICC) ซึ่งเป็นระดับความสามารถที่จะตอบข้อนั้นถูกต้อง ด้วยความน่าจะเป็น 0.5 เมื่อไม่มีการเดา ซึ่งมีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติมีค่าอยู่ระหว่าง -3.00 ถึง 3.00 ซึ่งในการวิจัยครั้งนี้ประมาณค่าตามทฤษฎีการตอบข้อคำถามโดยใช้ราส์ชโมเดล โดยใช้โปรแกรม BICAL

5. ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า (Standard Error of Estimate) SEE (θ) หมายถึง ดัชนีที่บ่งชี้ค่าประมาณค่าความสามารถของผู้สอบ ที่มีความคลาดเคลื่อนไปจากความสามารถจริง ซึ่งงานวิจัยครั้งนี้หาค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า โดยใช้ราส์ชโมเดลโดยใช้โปรแกรม BICAL ซึ่งมีสูตรดังนี้

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

เมื่อ SEE (θ) คือ ค่าความคลาดเคลื่อนมาตรฐาน

I (θ) คือ ฟังก์ชันอินฟอร์เมชันของแบบทดสอบ

6. ค่าความสามารถของผู้สอบ หมายถึง คุณลักษณะหรือคุณสมบัติประจำตัวของผู้สอบซึ่งประมาณค่าได้ด้วย โดยประมาณค่าจากพฤติกรรมการตอบข้อสอบ โดยใช้ราส์ชโมเดลด้วยโปรแกรม BICAL ความสามารถของผู้สอบ จะมีค่าอยู่ระหว่าง -∞ ถึง +∞ แต่ในทางปฏิบัติจะมีค่าอยู่ระหว่าง -3 ถึง +3 ค่า -3 แสดงว่ามีความสามารถต่ำสุด ค่า 3 แสดงว่ามีความสามารถสูงสุด

7. ขนาดกลุ่มตัวอย่างที่ต่างกัน หมายถึง กลุ่มตัวอย่างที่ได้จากการทดสอบแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ จำนวน 6 กลุ่ม ที่ได้มาโดยการสุ่มแบบแบ่งชั้น (Stratied Random Sample) ได้แก่ กลุ่มจำนวน 200 คน 300 คน 450 คน 675 คน 1,015 คน และ 1,610 คน โดยมีขนาดเพิ่มขึ้น 50% ตามลำดับ (สุทิน บุญญาภิบาล. 2537 : 49)

8. ผู้เชี่ยวชาญ หมายถึง อาจารย์ที่สอนวิชาคณิตศาสตร์ ที่มีประสบการณ์ด้านการสอนวิชาคณิตศาสตร์ไม่น้อยกว่า 10 ปี จำนวน 5 คน

บทที่ 2

เอกสารงานวิจัยที่เกี่ยวข้อง

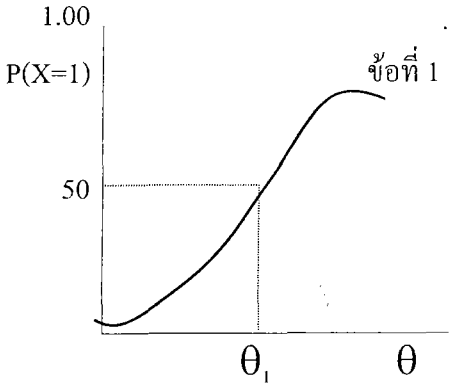
เอกสารและงานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยได้ศึกษาในเรื่องต่าง ๆ ดังนี้

1. ทฤษฎีการตอบข้อคำถาม
   - 1.1 พื้นฐานความเข้าใจในทฤษฎีการตอบข้อคำถาม
   - 1.2 ข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อคำถาม
   - 1.3 พิสัยของค่าพารามิเตอร์
   - 1.4 โมเดลของทฤษฎีการตอบข้อคำถาม
2. ราส์ชโมเดล
3. วิธีการประมาณค่าพารามิเตอร์
4. ดัชนีที่บ่งชี้ประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์
5. งานวิจัยที่เกี่ยวข้อง

1. ทฤษฎีการตอบข้อคำถาม

1.1 พื้นฐานความเข้าใจในทฤษฎีการตอบข้อคำถาม

ทฤษฎีการตอบข้อคำถามเป็นการวิเคราะห์ข้อสอบ โดยอาศัยข้อมูลจากตัวข้อสอบเอง ข้อสอบที่นำมาวิเคราะห์ในทฤษฎีการตอบข้อคำถาม ต้องเป็นข้อสอบที่มิใช่ทุกคนตอบผิดหมดหรือทุกคนตอบถูกหมด ข้อสอบทั้งสองแบบเป็นข้อสอบที่ไม่ชี้บ่งระดับโอกาสของการตอบถูกของบุคคลที่มีความสามารถแตกต่างกันอย่างใด การวิเคราะห์ข้อทดสอบที่ใช้ค่าความสามารถจริงของบุคคลนี้อาศัยข้อมูลที่เรียกว่า ลักษณะการตอบรายข้อเป็นสำคัญ โค้งลักษณะการตอบรายข้อนี้คือภาพที่แสดงความสัมพันธ์ระหว่างระดับความสามารถของผู้สอบกับโอกาสที่จะตอบถูก



ภาพประกอบ 1 แสดงโค้งลักษณะการตอบรายข้อ

จากภาพกำหนดให้แกนตั้งแทนค่าโอกาสตอบถูกและแกนนอนแทนความสามารถของผู้ตอบ ถ้าข้อสอบเป็นข้อที่มีคุณลักษณะที่ดี โอกาสของการตอบถูกของผู้ที่มีความสามารถต่ำกว่าย่อมมีค่าน้อยกว่าเสมอ บนโค้งนี้ถ้าลากเส้นจากแกนตั้งที่จุดค่า P เท่ากับ 0.50 ออกไปตัดโค้งลักษณะการตอบรายข้อ แล้วจากจุดตัดลากตั้งฉากมาตัดแกนนอน ค่าที่ได้ก็จะเป็นค่ายากง่ายของข้อทดสอบ (โกวิท ประวาลพฤกษ์. 2533 : 58 - 59) จากโค้งรายข้อนี้สามารถบ่งบอกถึงค่าพารามิเตอร์รายข้อ ซึ่งได้แก่ ค่าดังนี้

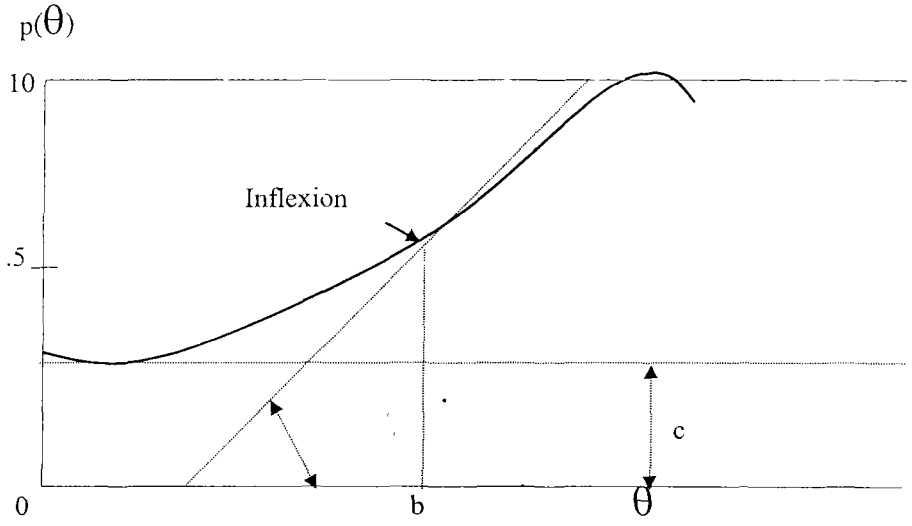
1. พารามิเตอร์ c เป็นความน่าจะเป็นที่บุคคลหนึ่งซึ่งปราศจากความสามารถโดยสิ้นเชิงจะตอบข้อนั้นได้ถูก พารามิเตอร์ตัวนี้มีชื่อเรียกว่า พารามิเตอร์ของการเดา ข้อใดที่ไม่สามารถตอบถูกได้ด้วยการเดา ข้อนั้นมีพารามิเตอร์ของการเดาเป็นศูนย์ (c = 0)

2. พารามิเตอร์ b เป็นพารามิเตอร์บอกตำแหน่งของโค้งตามมาตราความสามารถ พารามิเตอร์นี้ มีชื่อเรียกว่า ความยากของข้อสอบ (Item Difficulty) ข้อคำถามยิ่งยากโค้งยิ่งเลื่อนไปทางขวามากขึ้น โค้งโลจีสติคมีจุดเปลี่ยนโค้งอยู่ที่ θ = b เมื่อไม่มีการเดา b จะเป็นระดับความสามารถของบุคคลที่มีความน่าจะเป็นของการตอบข้อนั้นถูกเป็น .5

3. พารามิเตอร์ a เป็นสัดส่วนกับความชันของโค้ง (ความชันนี้มีค่าเป็น .425 a (1 - c)) ดังนั้น a จึงแทนอำนาจจำแนก (Discriminating Power) ซึ่งจะบอกว่าฟังก์ชันการตอบข้อคำถามแปรไปตามระดับความสามารถมากน้อยเพียงใด (ผจงจิต อินทสุวรรณ. 2528 : 16)

4. ค่าความรอบคอบ (r) เป็นค่าที่แสดงถึงความรอบคอบในการตอบข้อสอบข้อหนึ่ง ซึ่งมีค่าน้อยกว่า 1

5. ระดับความสามารถของผู้สอบ (θ) มีค่าอยู่ระหว่าง -3.0 ถึง 3.0 ค่าที่เป็นลบแสดงว่ามีความสามารถต่ำ และค่าที่เป็นบวกแสดงว่ามีความสามารถสูง



ภาพประกอบ 2 แสดงค่าพารามิเตอร์รายข้อ

ทฤษฎีการตอบข้อคำถาม (Item Response Theory : IRT) มีความเชื่อว่า ค่าพารามิเตอร์ (Parameter) ต่าง ๆ ของข้อสอบไม่ว่าจะเป็นค่าความยาก ค่าอำนาจจำแนก หรือค่าการเดาของข้อสอบแต่ละข้อเป็นคุณลักษณะที่มีอยู่ประจำ และคงที่พอสมควรในตัวข้อสอบนั้นจริง แต่เนื่องจากความสามารถของผู้สอบ (Ability) เป็นคุณลักษณะแฝง (Latent Trait) ซึ่งไม่สามารถที่จะวัดหรือสังเกตได้โดยตรง แต่จะเป็นตัวพยากรณ์ผลการสอบหรือคะแนน ซึ่งเป็นสิ่งที่เราสามารถสังเกตหรือวัดได้ (สำเริง บุญเรืองรัตน์. 2527 : 98) นักวัดผลการศึกษาจึงได้พยายามหาความสัมพันธ์ระหว่างปริมาณการกระทำข้อสอบหรือคะแนนกับปริมาณความสามารถ โดยสามารถเขียนในรูปความสัมพันธ์ได้ดังนี้ (Lord. 1982 : 141)

$$P = f(\mu I/\theta 1, \theta 2, \theta 3, \ldots, \theta k; Bj)$$

เมื่อ P แทน ผลการสอบ

f แทน ฟังก์ชัน

μi แทน ผลการตอบข้อสอบ ตอบถูกให้ 1 ตอบผิดให้ 0

$\theta 1, \theta 2, \theta 3 \ldots\ldots \theta k$ แทน ความสามารถ (Ability) ที่ 1,2.3....., θk

Bj แทน ค่าพารามิเตอร์ของข้อสอบข้อที่ j

ความสามารถดังกล่าวยังต้องอาศัยข้อตกลงเบื้องต้นต่าง ๆ ดังต่อไปนี้

1.2 ข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อคำถาม

ข้อตกลงเบื้องต้นทั่ว ๆ ไปของทฤษฎีการตอบข้อคำถาม ได้แก่

1. ความเป็นมิติเดียว (Unidimension) หมายความว่า ความสามารถของคนมีอยู่ทั้งหมด k อย่าง ซึ่งความสามารถแต่ละอย่างต่างก็ส่งผลต่อการตอบข้อสอบต่าง ๆ ที่รวมกันเป็นแบบทดสอบ ถ้าผลการตอบข้อสอบหรือคะแนนของผู้สอบนั้น สามารถอธิบายได้เพียงความสามารถเดียวก็เพียงพอแล้ว เรียกเงื่อนไขข้อนี้ว่ามีความเป็นมิติเดียว (สำเริง บุญเรืองรัตน์. 2527 : 99) ในการปฏิบัติจริงนั้นสามารถกระทำโดยการวิเคราะห์องค์ประกอบ (Factor Analysis) แล้วสังเกตค่าไอเก้น (Eigen Value) ค่าสูงสุดว่าแตกต่างจากค่าอื่นอย่างชัดเจนหรือไม่ (Hambleton and Thumb. 1973 : 195 - 211)

2. ความเป็นอิสระในการตอบข้อสอบ (Local Independence) หมายถึง การตอบข้อสอบใดข้อหนึ่งถูกจะไม่มีผลต่อการตอบข้อสอบนั้น แมคโดแนลด์ (Mcdonald. 1982 : 379 - 396) ได้มีความเห็นว่า ถ้าข้อสอบชุดนั้นมีความเป็นมิติเดียว (Unidimension) แล้ว ความเป็นอิสระในการตอบ

ข้อสอบ (Local Independence) ก็จะมีเอง ส่วนการทดสอบว่าข้อสอบแต่ละข้อเป็นไปตามข้อตกลงของ Local Independence หรือไม่นั้น ลอร์ด ได้เสนอแนะว่าให้ตรวจสอบดูแบบแผนการตอบข้อคำถามของผู้สอบที่มีความสามารถปาน ๆ กันว่า ข้อสอบแต่ละคู่ เช่น ข้อ 1. และข้อ 2. ข้อ 2.และข้อ 3. นั้น ลักษณะการตอบแตกต่างกันหรือไม่ ซึ่งสามารถตรวจสอบได้จากการคำนวณค่าไคสแคว์ เพื่อทดสอบความเป็นอิสระของการตอบคำถามเหล่านั้นได้

3. โค้งแสดงลักษณะข้อสอบ (Item Characteristic Curve : ICC) โค้งแสดงลักษณะข้อสอบเป็นฟังก์ชันทางคณิตศาสตร์ ดังนี้

$$f(x/\theta) = \sum_{g=i}^{n} \pi \, PgUg\theta_{g}^{1-Ug}$$

แสดงถึงความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบข้อนั้นถูกต้องกับระดับความสามารถที่วัดได้ โดยใช้ชุดของข้อสอบ โดยที่โอกาสที่ผู้สอบตอบข้อสอบถูกจะขึ้นอยู่กับโค้งแสดงลักษณะข้อสอบ (ICC) ในแต่ละโมเดลที่ใช้ ด้วยเหตุนี้โอกาสของการตอบถูกข้อสอบข้อนั้นจะไม่ขึ้นอยู่กับการแจกแจงความสามารถของกลุ่มประชากร (Hambleton and Swaminathan. 1985 : 25) นั่นคือ รูปร่างของเส้นโค้งข้อคำถามไม่ได้ขึ้นอยู่กับการกระจายความสามารถของกลุ่มผู้สอบ จึงเห็นได้ว่าลักษณะข้อคำถามและคุณสมบัติของคำถามนี้ไม่ขึ้นอยู่กับลักษณะของกลุ่มผู้สอบ จึงเป็นข้อคำถามที่มีคุณสมบัติคงที่ตลอดเวลา ซึ่งนับได้ว่าแก้ไขปัญหาของการวัดผลที่ยึดถือทฤษฎีดั้งเดิมได้อย่างดียิ่ง (สำเริง บุญเรืองรัตน์. 2527 : 39)

1.3 พิสัยของค่าพารามิเตอร์

ค่าพารามิเตอร์ในทฤษฎีการตอบข้อคำถามแบ่งออกเป็นค่าพารามิเตอร์ของข้อสอบ (Item Parameter) ได้แก่ ค่าความยาก (b_g) ค่าอำนาจจำแนก (a_g) ค่าสัมประสิทธิ์การเดา (c_g) และค่าพารามิเตอร์ของผู้สอบ (Examinee Parameter) ได้แก่ ระดับความสามารถของผู้สอบ (θ) พิสัยของค่าพารามิเตอร์ต่าง ๆ เหล่านี้ มีดังนี้ (Hambleten and Cook. 1977 : 81)

b_g มีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัตินั้นจะมีค่าอยู่ระหว่าง -3 ถึง +3 ค่า -3 แสดงว่าข้อสอบข้อนั้นง่ายมาก และค่า +3 แสดงว่าข้อสอบนั้นยากมาก

a_g มีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติมีค่าตั้งแต่ 0 ถึง +2 เพราะค่า a_g ที่เป็นลบแสดงว่า ข้อสอบไม่ดีใช้ไม่ได้ จึงตัดข้อสอบข้อนั้นทิ้งไป ค่า 0 แสดงว่าข้อสอบไม่มีอำนาจจำแนกหรือมีค่าอำนาจจำแนกต่ำ ค่า +2 แสดงว่าข้อสอบมีค่าอำนาจจำแนกสูง

θ มีค่าอยู่ระหว่าง -3 ถึง +3 ค่า -3 แสดงว่ามีความสามารถต่ำ +3 แสดงว่ามีความสามารถสูง

1.4 โมเดลของทฤษฎีการตอบข้อคำถาม

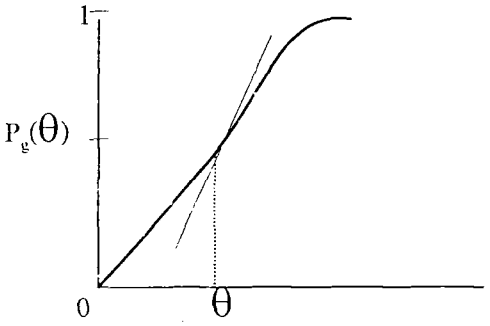
โมเดลของทฤษฎีการตอบข้อคำถามเป็นโมเดลทางคณิตศาสตร์ที่กล่าวถึงกันมาก มีอยู่ 2 โมเดล คือ (Warm. 1978 : 23 - 25 ; Hambleten and Cook. 1977 : 81 - 82)

1. โมเดลโค้งปกติสะสม (Normal Ogive Model) มี 3 รูปแบบ คือ

1.1 แบบ 1 - พารามิเตอร์ (One-Parameter) เป็นรูปแบบที่มีค่าพารามิเตอร์ตัวเดียว คือ ค่า b ที่แสดงลักษณะข้อสอบ เมื่อ θ คือ ความสามารถของผู้สอบ แสดงได้ด้วยฟังก์ชันดังนี้

$$P_g(\theta)\int_{-\infty}^{\theta-bg} \frac{1}{\sqrt{2\pi}}.e\frac{-t^2}{2}dt$$

และแสดงโค้งลักษณะข้อสอบดังนี้

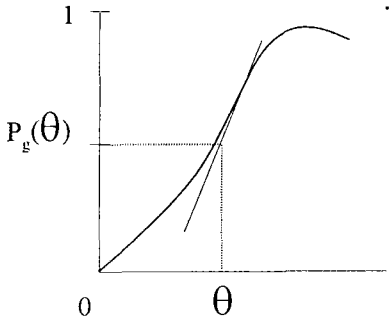


ภาพประกอบ 3 แสดงโมเดลโค้งปกติสะสมที่มีพารามิเตอร์ตัวเดียว

1.2 แบบ 2 - พารามิเตอร์ (Two-Parameter) เป็นรูปแบบที่ประกอบด้วย 2 พารามิเตอร์ คือ a และ b ที่แสดงลักษณะข้อสอบ เมื่อ θ คือ ความสามารถของผู้สอบ มีฟังก์ชันดังนี้

$$P_g(\theta)\int_{-\infty}^{\theta-bg} \frac{1}{\sqrt{2\pi}}.e\frac{-t^2}{2}dt$$

และแสดงโค้งลักษณะข้อสอบดังนี้

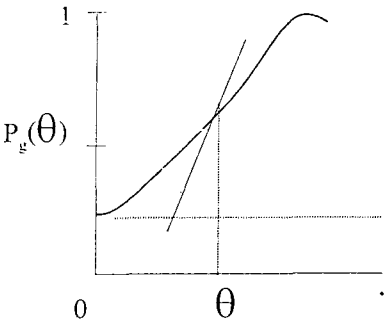


ภาพประกอบ 4 แสดงโมเดลโค้งปกติสะสมที่มีพารามิเตอร์สองตัว

1.3 แบบ 3 - พารามิเตอร์ (Three - Parameter) เป็นรูปแบบที่มีค่าพารามิเตอร์ลักษณะข้อสอบ 3 ค่า คือ a , b และ c เมื่อ θ คือ ความสามารถของผู้สอบ มีฟังก์ชันดังนี้

$$P_g(\theta) = c_g + (1 - c_g) \int_{-\infty}^{a_g(\theta - b_g)} \frac{1}{\sqrt{2\pi}} . e \frac{-t^2}{2} dt$$

และแสดงโค้งลักษณะข้อสอบดังนี้



ภาพประกอบ 5 แสดงโมเดลโค้งปกติสะสมที่มีพารามิเตอร์สามตัว

เมื่อ a เป็นค่าอำนาจจำแนกหรือสัดส่วนของความชัน ณ จุดเปลี่ยนโค้ง

b เป็นค่าความยากของข้อสอบหรือตำแหน่งของข้อสอบบนเส้นความสามารถ

c เป็นค่าโอกาสการเดาถูก

2. โมเดลโลจีสติค (Logistic Model) เป็นโมเดลที่พัฒนาเพื่อให้เหมาะสมกับสถานการณ์การทดสอบในกรณีที่ตอบโดยการเดาได้ ซึ่งจะทำให้ส่วนปลายที่สุดของโค้ง (Lower Asymptose) ของ $P_i(\theta)$ ไม่เป็นศูนย์ ดังนั้น พารามิเตอร์ของการเดาจึงจำเป็นต้องพิจารณาด้วย นอกจากนี้หากผู้สอบทำข้อสอบด้วยความสะเพร่า โมเดลนี้จะเพิ่มความรอบคอบขึ้นอีกค่าหนึ่ง ซึ่งจะเห็นได้จากปลายสูงสุดของโค้ง ICC มีค่าน้อยกว่า 1 โมเดลโค้งปกติสะสมกับโมเดลโลจีสติค จึงมีลักษณะโค้งใกล้เคียงกัน แต่โมเดลโลจีสติคมีความง่ายในการประมาณค่ามากกว่า จึงทำให้แพร่หลายมากกว่า รูปแบบโมเดลโลจีสติค มี 4 รูปแบบ คือ

2.1 แบบ 1 - พารามิเตอร์ (One-Parameter) หรือรูปแบบของราส์ชโมเดล ราส์ชได้พัฒนาโมเดลนี้ โดยถือว่าข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากันหมด และไม่มีการเดาได้ถูกต้อง จึงมีค่าพารามิเตอร์แสดงลักษณะข้อสอบเพียงตัวเดียว คือ ค่าความยากเขียนเป็นฟังก์ชันได้ดังนี้

$$P_i(\theta) = \frac{e^{Da'(\theta - bi)}}{i + e^{Da'(\theta - bi)}}; i = 1,2,3.....,n$$

หรือ

$$P_i(\theta) = \frac{e^{D(\theta - bi)}}{1 + e^{D(\theta - bi)}}; i = 1,2,3,....,n$$

เมื่อ $P_i(\theta)$ คือ ค่าความน่าจะเป็นที่ผู้สอบคนหนึ่ง ซึ่งมีระดับความสามารถจะตอบข้อสอบข้อ i ได้ถูกต้อง

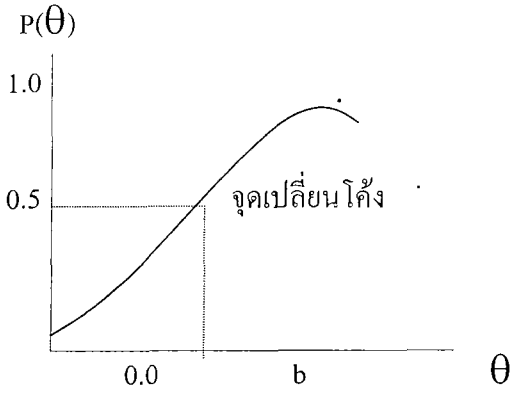
a_i คือ ค่าอำนาจจำแนก

b_i คือ ระดับความยากของข้อสอบข้อที่ i

θ คือ ระดับความสามารถของผู้สอบ

D คือ ค่าคงที่มีค่าเท่ากับ 1.702

e คือ ค่าคงที่มีค่าเท่ากับ 2.7182818



ภาพประกอบ 6 แสดงความหมายของค่า Item Parameter ของ One-Parameter Logistic Model

2.2 แบบ 2 - พารามิเตอร์ (Two-Parameter) หรือรูปแบบที่เบิร์นบอม (Birnbaum) ได้พัฒนาขึ้นในปี 1968 มีค่าพารามิเตอร์แสดงลักษณะข้อสอบ 2 ค่า คือ ค่าอำนาจจำแนก และค่าความยาก แสดงได้โดยฟังก์ชันนี้

$$P_i(\theta) = \frac{e^{Dai(\theta - bi)}}{1 + e^{Dai(\theta - bi)}}; i = 1,2,3,...., n$$

หรือ

$$P_i(\theta) = [\, 1 + e^{-Dai(\theta - bi)}] \, -1 \; ; i = 1,2,3,....,n$$

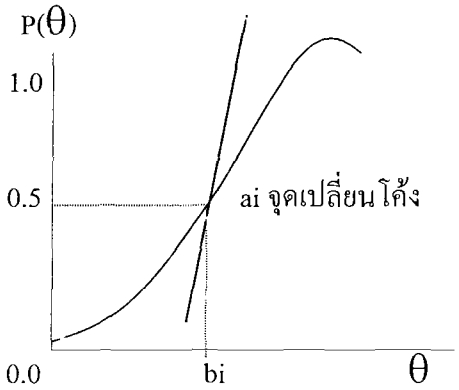
เมื่อ $P_i(\theta)$ คือ ค่าความน่าจะเป็นที่ผู้สอบคนหนึ่ง ซึ่งมีระดับความสามารถจะตอบข้อสอบข้อ i ได้ถูกต้อง

bi คือ ค่าความยากของข้อที่ i ที่แสดงระดับความสามารถที่แท้จริงที่จุดโค้งชันที่สุด

ai คือ ค่าอำนาจจำแนกข้อสอบข้อ i ที่เป็นสัดส่วนกับความชันของ $P_i(\theta)$ ณ ตำแหน่ง θ = bi

D คือ ค่าคงที่มีค่าเท่ากับ 1.702

e คือ ค่าคงที่มีค่าเท่ากับ 2.7182818



ภาพประกอบ 7 แสดงความหมายของค่า Item Parameter ของ Two-Parameter Logistic Model

2.3 แบบ 3 - พารามิเตอร์ (Three-Parameter) หรือรูปแบบที่เบิร์นบอม (Birnbaum) ได้พัฒนา โดยมีค่าพารามิเตอร์แสดงคุณลักษณะข้อสอบ 3 ค่า โดยเพิ่มค่าการเดา (C) ใน 2 - พารามิเตอร์ เขียนเป็นฟังก์ชันได้ดังนี้

$$P_i(\theta) = ci + (1 - ci)\frac{e^{Dai(\theta - bi)}}{1 + e^{Dai(\theta - bi)}}; i = 1,2,3,....,n$$

เมื่อ $P_i(\theta)$ คือ ค่าความน่าจะเป็นที่ผู้สอบคนหนึ่ง ซึ่งมีระดับความสามารถ θ จะตอบข้อสอบข้อ i ได้ถูกต้อง

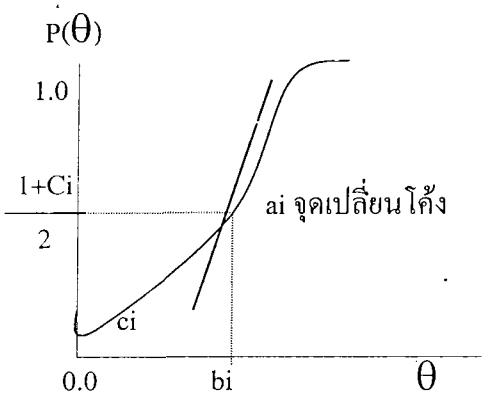
ci คือ ค่าความน่าจะเป็นของผู้สอบที่มีความสามารถต่ำมาก มีโอกาสจะทำข้อสอบข้อนี้ได้ถูกต้องหรือค่าการเดา

bi คือ ค่าความยากของข้อ i ที่แสดงระดับความสามารถที่แท้จริงที่จุดโค้งชันที่สุด คือ θ ณ จุดความน่าจะเป็น (1+C)/2

ai คือ ค่าอำนาจจำแนกข้อสอบข้อ i ที่เป็นสัดส่วนกับความชันของ $P_i(\theta)$ ณ ตำแหน่ง θ = bi

D คือ ค่าคงที่มีค่าเท่ากับ 1.702

e คือ ค่าคงที่มีค่าเท่ากับ 2.7182818



ภาพประกอบ 8 แสดงความหมายของค่า Item Parameter ของ Three-Parameter Logistic Model

2.4 แบบ 4 - พารามิเตอร์ (Four-Parameter) โมเดลนี้เสนอโดยแม็คโดแนลด์ (McDonald) ในปี ค.ศ.1967 ซึ่งลอร์ด (Lord) และบาร์ตัน (Barton) ได้พัฒนาในปี ค.ศ.1981 โดยเพิ่มพารามิเตอร์ค่าความรอบคอบ (Careless) ในโมเดลที่ใช้พารามิเตอร์ 3 ตัว เป็นค่าที่แสดงให้เห็นถึงกรณีที่ผู้สอบมีความสามารถสูงมักตอบข้อสอบผิด อาจจะเป็นเพราะความไม่รอบคอบทำให้ผู้นั้นไม่เลือกคำตอบที่ถูก จะเห็นได้จากปลายสูงสุดของโค้ง ICC ต่ำกว่า 1 เขียนเป็นสมการได้ดังนี้ (Hambleton and Swaminathan. 1985 : 48)

$$P_i(\theta) = ci + (ri - ci)\frac{e^{Dai(\theta - bi)}}{1 + e^{Dai(\theta - bi)}}; i = 1,2,3,....,n$$

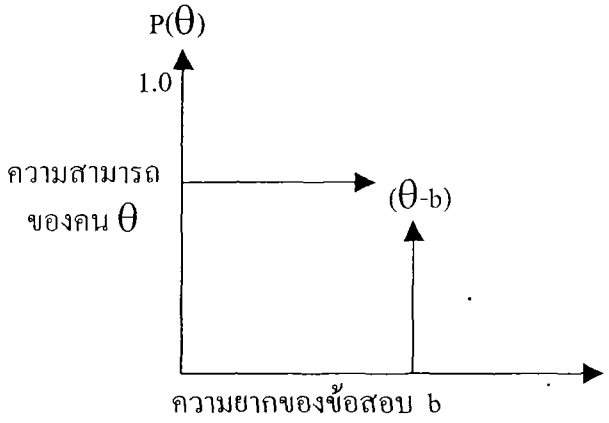
เมื่อ ri คือ ค่าความรอบคอบของข้อสอบข้อ i ซึ่งมีค่าน้อยกว่า 1

2. <u>ราส์ชโมเดล</u>

ราส์ชโมเดล เป็นโมเดลหนึ่งในโมเดลของทฤษฎีการตอบข้อคำถามที่ใช้พารามิเตอร์ตัวเดียว กล่าวคือ พิจารณาค่าความยากตัวเดียว โดยถือว่าข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากันหมดและไม่มีการเดาได้ถูกต้อง

ในการสรุปมโนทัศน์เกี่ยวกับราส์ชโมเดล อาจกล่าวได้ว่า เมื่อคน ๆ หนึ่งทำข้อสอบข้อหนึ่ง เหตุการณ์ขณะนั้นก็คือ คนพยายามใช้ความสามารถที่มีการเอาชนะข้อสอบด้วยการแก้ปัญหาโจทย์ปัญหาข้อนั้น ถ้าคนนั้นมีความสามารถมากกว่าความยากของข้อสอบข้อนั้น โอกาสที่คนจะทำข้อสอบถูกก็มีมาก ในทางตรงกันข้ามถ้าคน ๆ นั้นมีความสามารถน้อยกว่าความยากของข้อสอบ โอกาสที่จะถูกก็มีน้อยไปด้วย หรืออาจพูดอีกอย่างว่า โอกาสจะทำผิดก็มีมาก และถ้าความสามารถของคนมีพอ ๆ กันกับความยากของข้อสอบโอกาสของการทำข้อมูลก็จะมีอยู่ครึ่งต่อครึ่ง คือ อาจจะทำถูกก็ได้หรืออาจจะทำผิดก็ได้เช่นเดียวกัน

เหตุการณ์ขณะที่บุคคลใช้ความสามารถแก้ปัญหาโจทย์ข้อหนึ่ง ๆ อาจเขียนเป็นรูปได้ดังนี้



ภาพประกอบ 9 แสดงมโนทัศน์เกี่ยวกับราส์ชโมเดล

เมื่อกำหนดให้ θ คือ ความสามารถของคน และ b คือ ความยากของข้อสอบ การเปรียบเทียบความสามารถกับความยากของข้อสอบผลลัพธ์ อาจเขียนออกมาในเชิงคณิตศาสตร์ได้ดังนี้

ถ้า $\theta > b$ และ $(\theta-b) > 0$, $P > .5$

อธิบายได้ว่า เมื่อความสามารถของคนมีมากกว่าความยากของข้อสอบ คือ bi และถ้านำเอาความสามารถของคนลบด้วยความยากของข้อสอบแล้ว ผลลัพธ์จะได้มากกว่าศูนย์ โอกาสที่คนๆ นั้นจะทำข้อสอบข้อนั้นถูกก็จะมีค่ามากกว่า 0.50

ถ้า $\theta < b$ และ $(\theta-b) < 0$, $P > .5$

อธิบายได้ว่า เมื่อความสามารถของคนลบด้วยความยากของข้อสอบได้ผลลัพธ์น้อยกว่าศูนย์ โอกาสที่คน ๆ นั้นจะทำข้อสอบข้อนั้นได้ถูกต้องก็จะมีค่าน้อยกว่า 0.50

ถ้า $\theta = b$ และ $(\theta-b) = 0$, $P > .5$

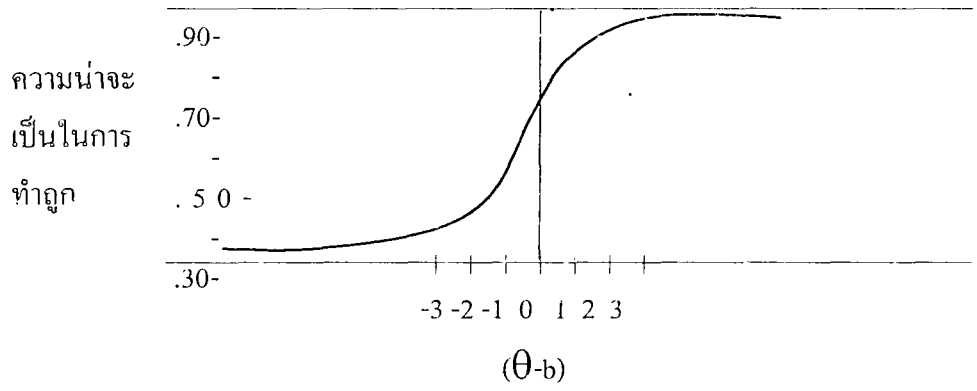
อธิบายได้ว่า เมื่อความสามารถของคนลบด้วยความยากของข้อสอบได้เท่ากับศูนย์พอดีแล้ว โอกาสที่คน ๆ นั้นจะทำข้อสอบข้อนั้นได้ก็จะมีค่าเท่ากับ .50 พอดี ซึ่งอาจจะทำถูกหรือผิดก็ได้

จะเห็นได้ว่า เมื่อคน ๆ หนึ่งทำข้อสอบข้อหนึ่ง บุคคลนั้นต้องใช้ความสามารถต่อสู้กับความยากของข้อสอบ ผลของการทำข้อสอบจึงเท่ากับความแตกต่างระหว่างความสามารถของคน

กับความยากของข้อสอบ ซึ่งก็คือความแตกต่างระหว่าง θ กับ b หรือ (θ-b) นั่นเอง แต่ค่า θ-b นี้ อาจจะให้ผลลัพธ์ได้ตั้งแต่ $-\infty$ จนถึง $+\infty$ จึงต้องคิดแปลงให้เป็นค่าความน่าจะเป็น ซึ่งจะมีค่าอยู่ระหว่าง .00 ถึง .99 หรือจาก 0 ถึง 1 เท่านั้น และเพื่อให้ค่าเปลี่ยนแปลงแต่น้อย จึงใช้ความน่าจะเป็นค่าอัตราส่วนของล็อกปกติ (Natural log) ของค่าคงที่ตัวหนึ่งคือ ค่า $e^{(\theta-b)}$ และค่าอัตราส่วนนั้นก็คือ

$$P_X = \frac{e^{(\theta-b)}}{1+e^{(\theta-b)}}$$

ซึ่งจะทำให้ได้เส้นความถี่สะสมปกติหรือ Normal Ogive ที่ราส์ชเรียก Logistic ที่อาจแสดงได้ดังต่อไปนี้



ภาพประกอบ 10 แสดงความแตกต่างระหว่างความสามารถของคนกับความยากของข้อสอบ

เพื่อให้เห็นจริงง่ายยิ่งขึ้น ตารางต่อไปนี้จะแสดงให้เห็นว่า เมื่อความยากของข้อสอบคงที่ แต่ความสามารถของคนเปลี่ยนไปเป็นหลายระดับ โอกาสที่จะทำข้อสอบนั้นถูกก็จะเปลี่ยนไป ทำนองเดียวกันถ้าความสามารถของคนคงที่ แต่ความยากของข้อสอบมีระดับแตกต่างกันหลายระดับ โอกาสที่จะทำถูกจะเปลี่ยนไปด้วยดังนี้

ตาราง 1 ความแตกต่างระหว่างความสามารถของคนกับความยากของข้อสอบ

ความสามารถของคน θ	ความยากของข้อสอบ b	ความแตกต่าง $e(\theta-b)$	โอกาสทำถูก P
5	0	148.00	.99
4	0	54.60	.98
3	0	20.10	.95
2	0	7.39	.88
1	0	2.72	.73
0	0	1.00	.50
0	1	.368	.27
0	2	.135	.12
0	3	.050	.05
0	4	.018	.02
0	5	.007	.01

จากตารางข้างบนถ้ากำหนดให้บุคคล v ทำข้อสอบข้อ i ได้คะแนน Xvi = 1 เมื่อทำถูก และ Xvi = 0 เมื่อทำผิดแล้ว สมการของราส์ชเมื่ออาศัยค่าวิเคราะห์ความสามารถของคนกับความยากของข้อสอบจะเป็นดังนี้

$$Pr\{XviBv, \delta i = eXvi(Bv-\delta i) / 1+e(Bv-\delta i)\}$$

จะเห็นได้ว่า มาตราที่ใช้วัดความสามารถของคนก็ยังใช้ข้อสอบที่นำมาใช้ทดสอบนั่นเอง และมาตราที่ใช้ประเมินความยากของข้อสอบก็คือ กลุ่มคนที่ถูกทดสอบนั่นเอง แต่โดยวิธีการทางคณิตศาสตร์ ทำให้เราได้หน่วยที่เหมาะสมที่จะใช้วัดมากกว่าเดิม หน่วยนี้เรียกว่าโลจิท (logit) หน่วยล็อคจิทของบุคคลก็คือ ค่าล็อคปกติของความสำเร็จที่คน ๆ นั้นจะทำข้อสอบถูกเมื่อข้อสอบกำหนดค่าความยากคงที่ คือ $\delta = 0$ นั่นคือ โอกาสที่คน ๆ หนึ่งซึ่งมีความสามารถเท่ากับ β ทำข้อสอบที่มีความยาก $\delta = 0$ ได้ถูกต้อง คือ $e^{\beta} / (1+e^{\beta})$ ความได้เปรียบของเขา คือ $e^{\beta} = P / (1-P)$ และค่าล็อคปกติ คือ β

ถ้าจะพิจารณาที่ตัวข้อสอบจะพบว่า หน่วยความยากของข้อสอบในรูปล็อคจิท (logit) ก็คือ ค่าล็อคปกติของความได้เปรียบที่ข้อสอบมีอยู่เหนือคน เมื่อความสามารถของคนถูกกำหนดให้

คงที่ที่ $\beta = 0$ นั่นคือ โอกาสที่คนซึ่งมีความสามารถ $\beta = 0$ จะทำข้อสอบที่มีความยาก 0 ได้สำเร็จ คือ $e^{-0} / (1+e^{0})$ เมื่อความได้เปรียบของข้อสอบ $e^{0} = (1-P)/P$ และค่าล็อคปกติของข้อสอบ คือ b

3. <u>วิธีการประมาณค่าพารามิเตอร์</u>

จากที่กล่าวมาจะเห็นได้ว่า สามารถแยกคะแนนของคนออกจากคะแนนของข้อสอบ ทำให้สามารถวิเคราะห์ค่าความสามารถของคนโดยไม่ขึ้นกับข้อสอบได้ ในทำนองเดียวกันก็อาจวิเคราะห์ค่าของข้อคำถามโดยไม่ต้องขึ้นอยู่กับคนได้ (Item) ซึ่งเป็นลักษณะพิเศษของการวิเคราะห์แบบราส์ชโมเดล ที่ทำให้สามารถหาค่าวิเคราะห์ของคน และข้อสอบได้เป็นอิสระจากกัน

ในการวิจัยครั้งนี้ ผู้วิจัยเลือกใช้โลจีสติคโมเดล แบบ 1 - พารามิเตอร์ หรือราส์ชโมเดล ที่มีจุดเด่น 3 ประการ คือ (Hambleton. 189 : 157)

1. โมเดลนี้มีจำนวนพารามิเตอร์ในข้อคำถามน้อยกว่าโมเดลอื่น ๆ จึงทำให้ง่ายต่อการนำมาใช้งาน

2. มีปัญหาอันเกิดจากการประมาณค่าพารามิเตอร์ของข้อคำถามน้อยกว่าใช้โมเดลอื่นๆ

3. คุณสมบัติในการประมาณค่าพารามิเตอร์ของข้อคำถาม (Item Parameter) และค่าพารามิเตอร์ความสามารถของผู้ตอบ (Person Parameter) แยกเป็นอิสระจากกันอย่างสมบูรณ์

และจากผลงานวิจัยหลายชิ้นด้วยกันสนับสนุนว่า การประมาณค่าความสามารถของผู้สอบโดยใช้ราส์ชโมเดล หรือแบบ 1 - พารามิเตอร์ มีคุณภาพดีเท่าเทียมกับการประมาณค่าความสามารถของผู้สอบ แบบ 3 - พารามิเตอร์ (Henly and Others. 1989 : 367)

ในทฤษฎีการตอบข้อคำถามเมื่อมีการนำไปประยุกต์ใช้ นอกจากการที่จะเลือกโมเดลต่างๆ ให้เหมาะสมแล้ว วิธีการประมาณค่าความสามารถของผู้สอบก็เป็นอีกประเด็นหนึ่งที่เป็นขั้นตอนสำคัญที่ต้องเลือกใช้ให้เหมาะสมกับสภาพแต่ละครั้ง เพื่อผลการประมาณค่าความสามารถของผู้สอบจะได้มีประสิทธิภาพสูงที่สุด วิธีการประมาณค่าในทฤษฎีการตอบข้อคำถามแบบที่มีการให้คะแนนเป็น 2 ค่า คือ ถูกให้ 1 และผิดให้ 0 มีอยู่หลายวิธี พอสรุปได้ดังนี้ (ต่าย เชี่ยงฉี. 2534 : 49 - 50 ; อ้างอิงมาจาก Hambleton and Swaminathan. 1985 : 300)

1. วิธีประมาณค่าความสามารถของผู้สอบตามค่าความยากของข้อที่ยากที่สุดที่ทำถูก (The Most Difficulty of Item Answered) เป็นการให้คะแนน โดยดูจากข้อสอบการทดสอบที่ผู้สอบทำมาทั้งหมด ข้อใดที่มีค่าความยากสูงสุดที่ผู้สอบทำถูก ถือว่าเป็นค่าความสามารถของผู้สอบคนนั้น ค่าความยากที่ว่านี้ควรจะเป็นค่าความยากที่วิเคราะห์มาจากทฤษฎีการตอบข้อคำถาม

2. วิธีประมาณค่าความสามารถของผู้สอบตามค่าเฉลี่ยความยากของข้อที่ทำถูก (Average Difficulty of All Item Answered) การประมาณค่าตามค่าความยากเฉลี่ยของข้อสอบข้อที่

ตอบถูกเป็นวิธีที่ช่วยขจัดข้อบกพร่องของการประมาณค่าตามค่าความยากเฉลี่ย โดยนำเอาเฉพาะข้อสอบที่ผู้สอบตอบถูกต้องเท่านั้นมาหาค่าเฉลี่ย ซึ่งจะให้ความหมายของคะแนนเพื่อวัดระดับความสามารถของผู้สอบได้มากขึ้น คะแนนที่ให้โดยการประมาณค่าวิธีนี้จะมากพอ ๆ กับการประมาณค่าตามผลการตอบข้อสอบข้อสุดท้ายหรือพอ ๆ กับการประมาณค่าตามค่าความยากของข้อสอบที่มากที่สุด แต่คะแนนจะมีค่าน้อยกว่า เมื่อใช้วิธีประมาณค่าตามผลการตอบข้อสอบข้อสุดท้ายได้รวมเอาค่าความยากของข้อสอบที่ใช้ในการแยกทางไว้ด้วย วิธีการประมาณค่าตามค่าความยากเฉลี่ยของข้อสอบที่ตอบถูกจะมีความคงที่ของคะแนน และให้รายละเอียดของคะแนนมากกว่าวิธีการประมาณค่าตามผลการตอบข้อสอบข้อสุดท้าย อย่างไรก็ตาม วิธีการประมาณค่าตามค่าความยากเฉลี่ยของข้อสอบที่ตอบถูกจะมีความไวต่อการเดาตอบของผู้สอบมากกว่าวิธีการประมาณค่าตามค่าความยากเฉลี่ยอาจจะมีผลทำให้มีค่าความเชื่อมั่นต่ำลง

3. วิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของทุกข้อที่ทำ (Aberage Difficulty) วิธีการประมาณค่าแบบนี้เสนอโดย ลอร์ด (Weiss. 1974 : 33 ; citing Lord. 1972) วิธีการคือ นำค่าความยากของข้อสอบทุกข้อที่ผู้สอบได้ทำ โดยไม่คำนึงว่าจะตอบถูกหรือไม่มาหาค่าเฉลี่ย ค่าเฉลี่ยนี้จะถือเป็นความสามารถของผู้สอบผู้นั้น

4. วิธีประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ทำ (Difficulty of the Final Item) นั่นคือ การประมาณค่าความสามารถวิธีนี้มักนำใช้ในการทดสอบเทเลอร์ โดยใช้ข้อสุดท้ายที่ทำ ซึ่งมีค่าความยากสูงสุด โดยไม่คำนึงว่าจะตอบถูกหรือไม่ก็ตาม จะถือว่าค่าความยากของการทดสอบเทเลอร์ขั้นสุดท้ายที่ผู้สอบได้ทำ คือ ค่าความสามารถของผู้สอบผู้นั้น

5. วิธีประมาณค่าความสามารถตามค่าความยากของข้อสอบที่ต่อจากขั้นสุดท้ายที่ทำ (Difficulty of $(N+1)^{th}$ Itme) ซึ่งมีลักษณะคล้ายกับวิธีที่ 4 แต่เป็นการตรวจสอบต่อไปอีกขั้นหนึ่ง กล่าวคือ ถ้าขั้นสุดท้ายที่ทำถูก คะแนนในขั้นต่อไปก็จะเพิ่มขึ้น และถ้าขั้นสุดท้ายที่ทำผิด ขั้นต่อไปคะแนนก็จะลดลง

6. วิธีประมาณค่าความสามารถโดยใช้หลักการของเบส์ (Bayesian Strategies) การให้หลักการของเบส์นี้มีวิธีการย่อยหลายวิธี วิธีที่ใช้กัน คือ วิธีที่นักสถิติ โอเว่น และทอร์นไดค์ (Owen. 1975 ; citing Owen. 1969, 1975 : 351 - 356 ; Thorndike. 1972 : 303) เป็นผู้เสนอขึ้น เพื่อใช้ในการประมาณค่าความสามารถของผู้สอบในการทดสอบที่ให้ชื่อว่า เบส์เซี่ยน อัพเดทติ้ง (Bayesian Updating) ซึ่งสามารถประมาณค่าความสามารถของผู้สอบได้ค่อนข้างคงที่ แต่ใช้สถิติในการวิเคราะห์ยุ่งยากซับซ้อนมาก

7. วิธีประมาณค่าความสามารถโดยใช้หลักการความเป็นไปได้สูงสุด (Maximum Likelihood) การประมาณค่าพารามิเตอร์วิธีนี้มีการพัฒนาปรับปรุงเพื่อให้สามารถนำมาใช้ในการ

ประมาณค่าพารามิเตอร์ของราส์ชโมเดลในแบบทดสอบที่ตอบถูกให้ 1 ตอบผิดให้ 0 (0,1) ได้อย่างมีประสิทธิภาพ โดย ไรท์และปันจาพาคีสัน (Wright and Panchapakesan) เป็นผู้พัฒนาการประมาณค่าพารามิเตอร์วิธีการนี้ขึ้น ต่อมา ไรท์ และดักลาสได้ปรับความลำเอียง (Bias) ของการประมาณค่าพารามิเตอร์ ค่าความยากของข้อคำถาม (b) สำหรับโมเดลที่ใช้คะแนนแบบสองค่าเพื่อลดอิทธิพลอันเนื่องมาจากความยากของข้อคำถามทั้งหมดในแบบทดสอบ

แนวความคิดของวิธีแมกซิมันไลค์ลีฮูด มีอยู่ว่า การประมาณค่าความสามารถทำโดยอาศัยการสังเกตค่าที่วัดได้จากตัวอย่างสุ่มที่เลือกมาจากการแจกแจงที่ทราบรูปแบบของฟังก์ชันความหนาแน่น แต่ไม่ทราบค่าพารามิเตอร์ จึงน่าจะใช้โอกาสที่เราเลือกตัวอย่างและวัดค่าของหน่วยตัวอย่างต่าง ๆ ได้ ($X_1 = X_1, X_2 = X_2, \ldots, X_n = X_n$) มาพิจารณาหาค่าประมาณของพารามิเตอร์ (θ) การให้คะแนนเพื่อประมาณค่าความสามารถโดยวิธีนี้มีวิธีย่อยอยู่หลายวิธี แต่วิธีที่นิยมคือ วิธีความเป็นไปได้สูงสุดแบบไม่มีเงื่อนไข (Unconditional Maximum Likelihood)

8. วิธีประมาณค่าความสามารถโดยวิธีพร็อกซ์ (Cohen's Approximation Procedure) เป็นวิธีที่แสดงให้เห็นถึงหลักการสำคัญต่าง ๆ ภายใต้การวิเคราะห์หาคุณลักษณะประจำข้อ และค่าคุณลักษณะของผู้ตอบจากราส์ชโมเดล โดยค่าทั้งสองที่ประมาณได้จากวิธีการนี้ถูกกำหนดให้อยู่บนมาตราวัดเชิงเส้นเดียวกัน และเป็นอิสระจากกลุ่มตัวอย่าง และความยาวของแบบทดสอบที่ใช้ โดยที่วิธีการประมาณค่าความสามารถวิธีนี้มีข้อตกลงเบื้องต้นอยู่ คือ ในแบบทดสอบต้องมีการกระจายในลักษณะเบ้หรือมีแนวโน้มเป็นโค้งปกติ

ในการวิจัยครั้งนี้ ผู้วิจัยได้เลือกใช้วิธีการประมาณค่าความสามารถวิธีที่ 8 เนื่องจากเป็นวิธีที่มีความคงที่ในการประมาณค่ามาก ถึงแม้จะใช้ค่าสถิติที่ซับซ้อน แต่ค่าที่ประมาณค่าออกมาได้ผลมีประสิทธิภาพ ซึ่งวิธีดังกล่าวนั้นคือ วิธีประมาณค่าความสามารถโดยใช้วิธีพร็อกซ์ (Cohen's Approximation Procedure ; PROX) ดังมีรายละเอียดดังต่อไปนี้

วิธีประมาณค่าความสามารถโดยใช้วิธีพร็อกซ์ (Cohen's Approximation Procedure ; PROX) เป็นวิธีประมาณค่าที่เริ่มจากการแปลงคะแนนรวม (Total Scores) ของผู้ตอบแต่ละคนและของข้อคำถามในเมตริกซ์ข้อมูล ผลการตอบของกลุ่มตัวอย่างให้อยู่ในรูปคะแนนโลจิท (Logit Scores) ทั้งนี้เพื่อให้คะแนนของบุคคลและของข้อคำถามอยู่ในมาตราอันตรภาค (Interval Scale) เป็นหน่วยเดียวกันบนมาตราเชิงเส้น (Linear Scale) จากนั้นจึงคำนวณหาค่าเฉลี่ยคะแนนโลจิทของแต่ละรายการ เพื่อแยกอิทธิพลของผู้ตอบออกมา และปรับขยายค่าที่คำนวณได้อีกครั้งหนึ่งเพื่อลดผลที่เกิดจากการกระจายของคะแนนข้อคำถาม และของผู้ตอบ ซึ่งหลักการของการประมาณค่าพารามิเตอร์นี้เป็นของโคเฮ็น แล้วมีการนำมาปรับปรุงโดยไรท์และมาสเตอร์ สามารถที่จะวิเคราะห์ผลได้ด้วยการใช้โปรแกรม BICAL (Wright and Mead 1979)

การวิเคราะห์ด้วยเครื่องคอมพิวเตอร์

การวิเคราะห์ด้วยเครื่องคอมพิวเตอร์ใช้โปรแกรมไบคาล (BICAL) ซึ่งเขียนโดย Wright และ Mead (Wright and Mead 1979) ในปี ค.ศ. 1978 ซึ่งมีอยู่ 2 สับโปรแกรมคือ PROX และ UCON โดยที่ข้อตกลงเบื้องต้นในการเลือกใช้โปรแกรมย่อยทั้ง 2 นี้ได้กล่าวมาแล้วในตอนต้น ขั้นตอนการวิเคราะห์โดยสังเขปมีดังนี้

ขั้นที่ 1 ตรวจข้อสอบ

โปรแกรมจะทำการตรวจสอบเพื่อหาว่าข้อสอบแต่ละข้อมีคนทำได้ถูกต้องเท่าใด และจะนับจำนวนความถี่ของคะแนนรวมในแต่ละจำนวน

ขั้นที่ 2 กำจัดข้อสอบบางข้อออก

ข้อสอบที่ไม่มีคนทำได้เลย หรือข้อสอบที่คนทำได้ทุกคนและคนที่ทำคะแนนไม่ได้เลย หรือทำได้เต็มจะถูกกำจัดออกจากการวิเคราะห์เพราะถือว่าข้อสอบหรือคะแนนดังกล่าวไม่มีลักษณะที่จำเป็นมาพิจารณาวิเคราะห์เพราะว่าเป็นข้อสอบที่ยากที่สุดหรือง่ายที่สุดและแสดงว่ามีบางคนที่มีความสามารถต่ำที่สุดและสูงที่สุด

ในทางปฏิบัติเกณฑ์ในการกำจัดข้อสอบและคนบางคนออกจากการวิเคราะหอาจใช้อัตราส่วนของการเดาได้ถูกต้องเป็นเกณฑ์ เช่น ถ้าข้อสอบชนิดเลือกตอบมีจำนวน 100 ข้อ คนที่ได้คะแนนต่ำกว่าหรือเท่ากับ 20 คะแนน และสูงกว่า 80 คะแนน จะถูกกำจัดออกจากการวิเคราะห์ถ้าเป็นความต้องการของผู้วิเคราะห์ ทั้งนี้โปรแกรมจะมีคำสั่งให้เลือกด้วย ในขั้นนี้โปรแกรมจะแบ่งกลุ่มตัวอย่างออกเป็นกลุ่ม ๆ ตามจำนวนที่ผู้วิเคราะห์ข้อสอบต้องการ โดยอาศัยคะแนนรวมในการสอบเป็นเกณฑ์แต่ไม่เกิน 6 กลุ่ม

ขั้นที่ 3 กำหนดค่าเริ่มแรกของ θ_v

ขั้นที่ 4 กำหนดค่าเริ่มแรกของ b_i

ขั้นที่ 5 กำหนดค่าของชุดของ b_i

ขั้นที่ 6 คำนวณค่าของ θ_v ของแต่ละข้อโดยวิธีการทำซ้ำ (iteration) ค่า b_h ในสมการที่หาค่ามาจากสมการ

$$P\{ x_{vi}/\theta_v, b_i\} = \exp [x_{vi} (\theta_v - b_i) / \{1 + \exp (\theta_v - b_i)\}]$$

ขั้นที่ 7 คำนวณค่าของ θ_v ของคะแนนรวมที่แตกต่างกันโดยการทำซ้ำค่าในสมการอีกสมการหนึ่งที่ derived มาจากสมการในขั้นที่ 6 โดยอาศัยค่าของ b_i ด้วย

ขั้นที่ 8 เปรียบเทียบค่า b_i ในขั้นที่ 6 และค่าที่เปลี่ยนไปในขั้นที่ 7 ว่าแตกต่างกันที่ $\alpha = 0.0001$ หรือไม่ ถ้าแตกต่างกันก็กระทำในขั้นที่ 5 6 7 และ 8 อีกหลาย ๆ ครั้ง จนกระทั่งค่า b_i ของข้อต่าง ๆ คงที่

ขั้นที่ 9 ปรับปรุงค่า b_i ของแต่ละข้อให้ถูกต้อง

ขั้นที่ 10 คำนวณค่า θ_v ของผู้สอบที่ได้คะแนนรวมแตกต่างกันโดยอาศัยค่าในขั้นที่ 9 ที่ปรับปรุงแล้ว

ขั้นที่ 11 ปรับปรุงค่า θ_v ของผู้ที่ได้คะแนนรวมแตกต่างกันให้ถูกต้อง

ขั้นที่ 12 ทำการทดสอบโค้งลักษณะของข้อสอบ (Item Characteristic Curve : ICC) ของข้อสอบแต่ละข้อว่าเหมาะสม (fit) กับ ICC ของรูปแบบหรือไม่ โดยการทดสอบ 2 ลักษณะคือ

1. ทดสอบความเหมาะสมของ ICC โดยยึดเอาจุดต่าง ๆ ที่แสดงค่าร้อยละของผู้ตอบข้อสอบแต่ละข้อได้ถูกต้อง เมื่อจุดเหล่านั้นแทนระดับความสามารถของกลุ่มตัวอย่างที่แตกต่างกัน (ซึ่งเกิดจากการแบ่งโดยอาศัยคะแนนรวมเป็นเกณฑ์ในขั้นที่ 2) ว่าเหมาะสมกับ ICC ของรูปแบบหรือไม่ ถ้าค่าความแปรปรวน = 1 แสดงว่าโค้งทั้งสองเหมาะสมกันพอดี

2. ทดสอบความเหมาะสมของ ICC โดยยึดเอาระดับความสามารถต่าง ๆ ของผู้สอบที่แตกต่างกันแต่ละคนเป็นเกณฑ์ว่าสามารถทำข้อสอบนั้นได้ร้อยละที่แตกต่างกันกระจายแล้วเป็นโค้งที่เหมาะสมกับ ICC ของรูปแบบหรือไม่ ถ้าค่าความแปรปรวน = 1 แสดงว่าโค้งทั้งสองเหมาะสมกับ ICC พอดี

ขั้นที่ 13 โปรแกรมคำนวณหาค่าอำนาจจำแนก

สืบเนื่องจากรูปแบบของราสซ์ไม่ได้คำนึงถึงเรื่องค่าอำนาจจำแนก เพราะว่าเป็นข้อตกลงเบื้องต้นแล้วว่าข้อสอบแต่ละข้อมีค่าอำนาจจำแนกเท่า ๆ กัน และไม่ได้คำนวณค่าพารามิเตอร์อำนาจจำแนกดังกล่าวด้วย แต่โปรแกรมไบคาลให้ค่าอำนาจจำแนกด้วยซึ่งเป็นการคำนวณภายหลังที่คำนวณด้วยวิธีของรูปแบบของราสซ์แล้ว ค่าอำนาจจำแนกดังกล่าวนี้บอกถึงระดับความเหมาะสมระหว่าง ICC ของข้อสอบแต่ละข้อกับ ICC ของรูปแบบ กล่าวคือ ถ้าค่าอำนาจจำแนกยิ่งใกล้ 1.00 มากเท่าใด ก็แสดงว่าโค้งทั้งสองมีความใกล้ชิดกันมาก ถ้าค่าอำนาจจำแนกน้อยกว่า 1.00 แสดงว่า ICC ของข้อสอบแบนกว่าของรูปแบบแสดงว่าข้อสอบนั้นไม่มีอำนาจจำแนกดีเท่ากับข้ออื่น ๆ ที่มีอำนาจจำแนกสูงกว่า

แต่ถ้าอำนาจจำแนกมากกว่า 1.00 แสดงว่าข้อสอบนั้นโค้งกว่า ICC ของรูปแบบและมีค่าอำนาจจำแนกมากเกินไป จึงควรนำมาพิจารณาว่าเหตุใดจึงเป็นเช่นนั้น ค่าสหสัมพันธ์แบบไบซีเรียล (Biserial Correlation) นั้นมีความหมายเหมือนกับการวิเคราะห์ข้อสอบโดยทฤษฎีมาตรฐานดั้งเดิม

ขั้นที่ 14 การเรียงลำดับข้อสอบ

การเรียงลำดับข้อสอบที่วิเคราะห์จากโปรแกรมมี 3 ลักษณะคือ

1. เรียงตามลำดับที่ของข้อสอบ

2. เรียงตามลำดับที่ความยากจากข้อที่ง่ายที่สุดไปยังข้อที่ยากที่สุด

3. เรียงตามลำดับระดับความเหมาะสมกันระหว่าง ข้อสอบกับของรูปแบบ จากข้อที่เหมาะสมมากที่สุดไปถึงข้อที่เหมาะสมน้อยที่สุด

4. ดัชนีบ่งชี้ประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์

ในการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์นั้น ต้องอาศัยดัชนีที่บ่งชี้ถึงประสิทธิภาพได้เป็นอย่างดี ในการวิจัยครั้งนี้ใช้ดัชนีที่เป็นเกณฑ์ในการเปรียบเทียบประสิทธิภาพในการประมาณค่าพารามิเตอร์ 1 ค่า คือ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า (Standard Error of Estimate : SEE(θ))

ในการวัดหรือการประมาณค่าความสามารถทางด้านการศึกษาและจิตวิทยา มักจะมีความคลาดเคลื่อนค่อนข้างสูง ถ้าความคลาดเคลื่อนในการประมาณค่ามีค่าสูง ก็จะส่งผลให้ความถูกต้องแม่นยำในการประมาณค่ามีค่าต่ำ ความคลาดเคลื่อนในการประมาณค่านี้มักนิยมนำมาแสดงในรูปของส่วนเบี่ยงเบนมาตรฐานหรือที่เรียกว่า ความคลาดเคลื่อนมาตรฐานในการประมาณค่า (Standard Error of Estimate : SEE(θ)) ซึ่งสามารถนำค่านี้ไปใช้ประโยชน์ในลักษณะคล้ายกับความคลาดเคลื่อนมาตรฐานของการวัด (Standard Error of Measurement ; S.Emeas) ซึ่งค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่ามีความสัมพันธ์กับค่าอินฟอร์เมชั่นของแบบทดสอบ (Test Information Function, I(θ) ดังนี้ (Warm. 1978 : 77 ; Hambleton. 1989 : 164)

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

ตามทฤษฎีการตอบข้อคำถาม การวิเคราะห์คุณภาพของแบบทดสอบหรือการพิจารณาความแน่นอนของการประมาณค่าความสามารถที่แท้จริงนั้น ใช้วิธีการพิจารณาจากฟังก์ชันอินฟอร์เมชั่นของแบบทดสอบ (Test Information Function) ซึ่งได้จากผลรวมของฟังก์ชันอินฟอร์เมชั่นของข้อสอบแต่ละข้อ (Item Information Function) (Lord. 1980 : 72)

ฟังก์ชันอินฟอร์เมชั่นของข้อสอบ แต่ละข้อกำหนดได้ดังนี้

$$I\{\theta,U\} = (P')^2 / P.Q. \; ; \; i = 1,2, \ldots , n$$

เมื่อ I{θ,Ui} คือ ฟังก์ชั่นอินฟอร์เมชั่นของข้อสอบแต่ละข้อ

Pi คือ ความน่าจะเป็นที่ผู้สอบที่มีความสามารถ θ จะตอบข้อสอบข้อที่ i ถูก

Qi คือ (1-Pi)

Pi คือ ความชันของโค้งลักษณะข้อสอบที่ระดับความสามารถ θ

จากสมการ จะเห็นได้ว่า ฟังก์ชันของอินฟอร์เมชั่นของข้อสอบแต่ละข้อขึ้นอยู่กับความชันของโค้งลักษณะข้อสอบ ถ้าโค้งลักษณะข้อสอบชันมากขึ้นในขณะที่ความแปรปรวนของการตอบข้อสอบถูกยิ่งน้อยลง โค้งอินฟอร์เมชั่นของข้อสอบที่ระดับความสามารถนั้นก็จะยิ่งสูงขึ้น ซึ่งความสูงของโค้งอินฟอร์เมชั่นของข้อสอบอยู่ที่ระดับความสามารถใดแสดงว่า สามารถจำแนกระดับความสามารถของผู้สอบได้ดี ณ ระดับความสามารถนั้น (ผจงจิต อินทสุวรรณ. 2525 : 64)

<u>ฟังก์ชันอินฟอร์เมชั่นของแบบทดสอบ</u>

ฟังก์ชันอินฟอร์เมชั่นของแบบทดสอบ เป็นสัดส่วนกลับกับกำลังสองของความยาวของช่วงความเชื่อมั่นเป็นผลมาจากการประมาณค่าความสามารถของผู้สอบจากการทำข้อสอบของเขา โค้งอินฟอร์เมชั่นของแบบทดสอบจึงเป็นเครื่องแสดงถึงความถูกต้องแน่นอนของค่าความสามารถที่ประมาณได้ ซึ่งแสดงในรูปสมการได้ดังนี้

$$I(\theta) = \sum_{i=1}^{n} I(e, U.)$$

$$= \sum_{i=1}^{n} \frac{(Pi)2}{PiQi}$$

เมื่อ $I(\theta)$ คือ ฟังก์ชันอินฟอร์เมชั่นของแบบทดสอบ

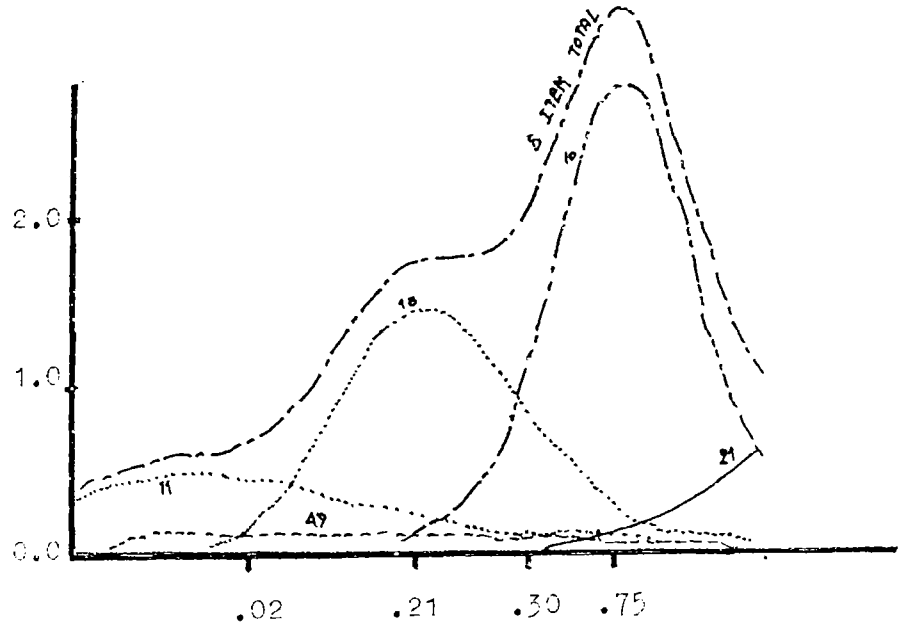
ค่า I นี้ใช้เป็นดัชนีชี้ให้เป็นคุณภาพของการวัดของแบบทดสอบ ถ้าโค้งของข้อสอบ (Information Curve หรือ Test Information Curve) สูงที่ระดับความสามารถใดก็แสดงว่า แบบทดสอบหรือข้อสอบที่คุณภาพดีมากสำหรับวัดความสามารถของบุคคลที่มีระดับความสามารถนั้น ๆ

ตาราง 2 แสดงค่าคุณสมบัติของแบบทดสอบ 5 ข้อ

ข้อสอบที่	bg	ag	cg
10	1.1	2.0	.05
11	-1.5	.9	.20
13	0.1	1.6	.16
30	2.4	1.1	.09
47	-0.4	.4	.20

ค่าต่าง ๆ ในตารางนี้เป็นคุณสมบัติของข้อสอบความถนัดทางการเรียนด้านภาษา (Lord. 1968) 5 ข้อ

เมื่อนำมาคำนวณค่า I ของแต่ละข้อ และคำนวณค่า I รวมทั้ง 5 ข้อแล้วก็สามารถกราฟได้ดังภาพประกอบ 11



ภาพประกอบ 11 แสดง Information Curve ของ แบบทดสอบความถนัดด้านภาษาของสถาบัน ETS

จากกราฟให้ความหมายว่า แบบทดสอบทั้ง 5 ข้อ เหมาะกับการทดสอบกับนักเรียนที่มีความสามารถสูง ๆ

อนึ่ง ค่าฟังก์ชันอินฟอร์เมชั่นของแบบทดสอบนอกจากจะเป็นดัชนีบ่งชี้คุณภาพของการประมาณค่า ยังสามารถนำมาใช้ในการเปรียบเทียบประสิทธิภาพในการประมาณค่าจากแบบ

ทดสอบต่างฉบับที่วัดความสามารถหรือคุณลักษณะอย่างเดียวกัน โดยการนำค่าฟังก์ชันอินฟอร์เมชั่นของแบบทดสอบแต่ละฉบับมาเปรียบเทียบกันทีละคู่ ณ ระดับความสามารถหรือคุณลักษณะเดียวกัน โดยแสดงในรูปของอัตราส่วน ซึ่งค่าอัตราส่วนของฟังก์ชันอินฟอร์เมชั่นของแบบทดสอบฉบับที่นำมาเปรียบเทียบ ณ ระดับความสามารถเดียวกันนี้เรียกว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบหรือประสิทธิภาพสัมพัทธ์ของการประมาณค่า (Relative Efficiencies of Various Test Design, Estimators อ้างอิงมาจาก Lord and Novick. 1986 : 471 - 472 ; Allen and Yen. 1979 : 266)

$$RE(\theta, X_1, X_2) = \frac{I(\theta, X_1)}{I(\theta, X_2)}$$

เมื่อ $RE(\theta, X_1, X_2)$ แทน ค่าประสิทธิภาพสัมพัทธ์ของคะแนนแบบทดสอบ X_1 กับคะแนนแบบทดสอบ X2 ณ ระดับความสามารถระดับ θ

โดยที่ X_1 และ X_2 เป็นคะแนนจากแบบทดสอบที่แตกต่างกัน 2 ฉบับ ที่วัดความสามารถอย่างเดียวกัน ถ้าผลการคำนวณค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบมีค่ามากกว่า 1 แสดงว่าแบบทดสอบ X1 มีคุรภาพสูงกว่าแบบทดสอบ X_2 ณ ความสามารถระดับนั้น แต่ถ้าค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบมีค่าน้อยกว่า 1 แสดงว่า แบบทดสอบ X_1 มีคุณภาพต่ำกว่าแบบทดสอบ X_2 ณ ระดับความสามารถนั้น แต่ถ้าค่าประสิทธิภาพสัมพัทธ์มีค่าเท่ากับ 1 แล้ว แสดงว่าคุณภาพของแบบทดสอบทั้งสองฉบับมีค่าพอกัน ณ ระดับความสามารถนั้น (Warm. 1978 : 75 - 76 ; Crocker and Algina. 1968 : 369 - 370 ; Lord and Stocking. 1988 : 270) และถ้าพิจารณาในแง่ของค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าแล้ว (Relative Efficiency of Estimator) X_1 และ X_2 จะแทนวิธีการประมาณค่าที่แตกต่างกัน 2 วิธีของแบบทดสอบฉบับเดียวกัน ค่าประสิทธิภาพสัมพัทธ์ในการประมาณค่าความสามารถนำมาเปรียบเทียบได้ โดยการใช้อัตราส่วนของค่าความแปรปรวนวิธีประมาณค่าวิธีการหนึ่ง เทียบกับวิธีการประมาณค่าอีกวิธีหนึ่ง

$$RE(\theta, X_1, X_2) = \frac{I(\theta, X_1)}{I(\theta, X_2)} = \frac{V(X_2)}{V(X_1)} = Eff(\theta, X_1, X_2)$$

จากสมการดังกล่าวนี้ จะเห็นได้ว่าการเปรียบเทียบประสิทธิภาพสัมพัทธ์ในการประมาณค่าพารามิเตอร์ ระหว่าง 2 วิธีใด ๆ สามารถคำนวณได้จากอัตราส่วนของค่าอินฟอร์เมชั่นของแบบทดสอบที่เป็นผลมาจากการประมาณค่าแต่ละวิธีในแบบทดสอบฉบับเดียวกัน ซึ่งอัตราส่วนดังกล่าวนี้มีค่าเท่ากับประสิทธิภาพสัมพัทธ์ของการประมาณค่าที่คำนวณจากการใช้อัตราส่วนของค่าความแปรปรวน

<u>งานวิจัยที่เกี่ยวข้อง</u>

งานวิจัยเกี่ยวกับต่างประเทศ มีผู้ทำการวิจัยอยู่หลายท่านดังนี้

โมเสส (Moses. 1971) ได้ศึกษาการวัดสารรูปสนิท (Goodness of Fit) ของรูปแบบของราส์ช ในการวิเคราะห์ผลสัมฤทธิ์ตามจุดประสงค์ของแบบทดสอบในโรงเรียนเขตแอฟริกาตะวันตก โดยมีจุดมุ่งหมายดังนี้

1. พิจารณาความเหมาะสมของรูปแบบของราส์ชกับข้อสอบแบบทดสอบวัดผลสัมฤทธิ์ในระดับมัธยมศึกษาตอนปลาย

2. ค้นหาขอบเขตของข้อสอบที่เหมาะสมกับรูปแบบของราส์ช เมื่อมีค่าอำนาจจำแนกเท่าๆ กัน

3. ศึกษาความไม่แปรเปลี่ยนของค่าพารามิเตอร์ที่ได้อ้างไว้

4. ค้นหาขอบเขตของการวัดคุณลักษณะมิติเดียวของแบบทดสอบวัดผลสัมฤทธิ์แต่ละฉบับ

ผลการศึกษาพบว่า ข้อสอบในวิชาคณิตศาสตร์ ภูมิศาสตร์ชีวภาพ และภาษาอังกฤษ มีความเหมาะสม รูปแบบของราส์ชดีพอใช้ที่ระดับ .01 ค่าอำนาจจำแนกของข้อสอบได้ ค่าที่ไม่เท่ากันจริง ๆ จะอยู่ในช่วง 0.5-1.25 ผู้วิจัยได้สรุปว่า

1. รูปแบบของราส์ชเหมาะสมกับแบบทดสอบวัดผลสัมฤทธิ์ทางวิชาการที่ใช้วัดในระดับมัธยมศึกษาตอนปลาย ทั้งนี้ควรจะเลือกข้อสอบที่มีความคงที่ภายในที่เหมาะสมด้วย

2. รูปแบบของราส์ช แม้มีข้อตกลงเบื้องต้นว่า ค่าจำแนกรายข้อเท่ากัน เมื่อทำการทดสอบการคัดเลือกข้อสอบ ค่าอำนาจจำแนกอาจไม่เท่ากันก็ได้

3. การพิจารณาความเป็นอิสระของบุคคล ในการคัดเลือกข้อสอบ (เมื่อปราศจากพารามิเตอร์ของข้อสอบ) สามารถใช้ได้กับรูปแบบของราส์ช

4. ในการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ (Calibration) ความเป็นอิสระของบุคคลไม่มีผลด้วย

5. ความเป็นอิสระของข้อสอบต่อบุคคล ในการวัดครั้งนี้ไม่เกิดผล และ ไรท์เลย์ และดาวิส (Whiteley and Dawis. 1974) วิเคราะห์แบบสอบอุปมาอุปไมยทางภาษา จำนวน 60ข้อ โดยใช้รูปแบบของราส์ช มีข้อสอบที่ไม่เหมาะสมกับรูปแบบประมาณ 30-40%

ส่วน เรนซ์และบาชอว์ (Rentz and Bashaw. 1975) ได้ทำการศึกษาและพบว่า ขนาดของกลุ่มตัวอย่าง 500-1,000 คน จะทำให้การประมาณค่าพารามิเตอร์ของรูปแบบของราส์ชมีความคงที่เพิ่มขึ้นอย่างรวดเร็ว และจะค่อย ๆ เพิ่มขึ้น เมื่อกลุ่มตัวอย่างที่ใช้ประมาณค่าพารามิเตอร์ มีจำนวน 2,000 คน และ 4,000 คน และเรนซ์ และรินเดอร์ (Rentz and Rinderour. 1978) ได้ใช้รูปแบบของ

ราส์ช วิเคราะห์แบบทดสอบโครงการการประเมินผลแบบทดสอบแอตแลนต้า (Atlanta Assessment Project) และแบบทดสอบวัดผลสัมฤทธิ์ของแสตนด์ฟอร์ด พบว่า จำนวนข้อสอบมีความเหมาะสมกับรูปแบบอยู่ระหว่าง 85-97% และ 92-97% ตามลำดับ สำหรับ ดูแกลส์ส (Duglass. 1980) ได้ศึกษาเปรียบเทียบการใช้รูปแบบโค้งลักษณะข้อสอบ (Item Characteristic Curve) กับการสอบในระบบห้องเรียน โดยใช้รูปแบบทั้ง 3 รูปแบบของทฤษฎีการตอบข้อคำถามประมาณค่าพารามิเตอร์ของข้อสอบ 100 ข้อ ที่เป็นข้อสอบปลายปีของ 4 ปีการศึกษา จากกลุ่มตัวอย่างระดับวิทยาลัย จำนวน 594-1,082 คน ข้อสอบแบ่งเป็น 4 ฉบับ แล้วสุ่มข้อสอบฉบับละ 43-53 ข้อ แล้วทำการสุ่มผู้สอบมา 200, 600 และ 800 คน มาวิเคราะห์ ปรากฏว่าผลการวิเคราะห์ข้อสอบด้วยรูปแบบของราส์ช มีความคงที่ของพารามิเตอร์ดีกว่ารูปแบบ 2 และ 3 พารามิเตอร์ ในกลุ่มตัวอย่าง 200, 600 และ 800 คน และเมื่อมีการเทียบมาตราของแบบทดสอบข้ามกลุ่มตัวอย่าง รูปแบบของราส์ชก็มีความคงที่ดีกว่ารูปแบบ 2 พารามิเตอร์และ 3 พารามิเตอร์ และฮัทเทน (Hutten. 1982 : 4799-A) ได้ศึกษาความเหมาะสมของข้อมูลจริงกับรูปแบบของทฤษฎีการตอบข้อคำถาม 2 รูปแบบ คือ รูปแบบของราส์ชกับรูปแบบ 3 พารามิเตอร์ ในการประมาณค่าพารามิเตอร์ความสามารถและความยากของข้อสอบใช้กลุ่มตัวอย่างนักเรียน 1,000 คน และ 250 คน กับแบบทดสอบสั้น 20 ข้อ จุดมุ่งหมายเพื่อพิจารณาความแม่นยำ (Precision) ในการประมาณค่าพารามิเตอร์ทั้ง 2 พารามิเตอร์ และผลการศึกษาพบว่า รูปแบบของราส์ชและรูปแบบ 3 พารามิเตอร์ เหมาะสมกับข้อมูล 80 เปอร์เซ็นต์ ทั้งสองรูปแบบเมื่อพิจารณาคะแนนที่ทำนายกับคะแนนที่วัดได้จากสถิติ แล้วรูปแบบราส์ชเหมาะสมกับข้อมูลทั้งหมดดีกว่ารูปแบบ 3 พารามิเตอร์ ผลยังยืนยันอีกว่าการใช้ข้อมูลเชิงประจักษ์กับรูปแบบของราส์ชกับรูปแบบ 3 พารามิเตอร์ ควรจะคำนึงถึงคุณลักษณะมิติเดียวเป็นอันดับแรก ในการประมาณค่าความสามารถจากแบบทดสอบสั้น 20 ข้อ และใช้กลุ่มตัวอย่างขนาดเล็ก จำนวน 250 คน รูปแบบของราส์ชเหมาะสมกว่ารูปแบบ 3 พารามิเตอร์ผลการศึกษาได้สนับสนุนให้ใช้กลุ่มตัวอย่างขนาด 1,000 คนขึ้นไป กับรูปแบบ 3 พารามิเตอร์ เพื่อจะได้ความแม่นยำในการประมาณค่าพารามิเตอร์ต่างๆ

ราส์ช (Rash. 1983 : 1429-A) ได้ศึกษาการใช้รูปแบบของราส์ชในแบบทดสอบหลาย ๆ คุณลักษณะ (Multidimensional Item) โดยแปลงให้เป็นคุณลักษณะเดียวกันหลาย ๆ กลุ่ม วิธีการใช้ประมาณค่าใช้คะแนนจริงจากแบบทดสอบ จากการศึกษาพบว่าค่าพารามิเตอร์ที่ได้จากการวัดแต่ละกลุ่ม มีการกระจายที่ไม่เปลี่ยนแปลงไปตามกลุ่มผู้สอบและแบบทดสอบ ส่วนแม็คเพสัน (Mcpheson. 1983) ได้ศึกษาการใช้รูปแบบของราส์ชในการประเมินข้อสอบโดยใช้แบบทดสอบ 4 ฉบับ สำหรับนักเรียนเกรด 4 และเกรด 7 ที่สอบในโปรแกรมการประเมินทางการศึกษาของมิชิแกน (Michigan Educational Assesment Program : MEAP) ซึ่งเป็นแบบทดสอบอิงเกณฑ์ในรูปแบบ การ

อ่านที่ใช้สอบในปี ค.ศ. 1976-1979 ผู้วิจัยต้องการทดสอบว่า รูปแบบของราส์ชจะเหมาะสมหรือไม่กับแบบทดสอบที่ใช้เวลามากในการที่จะให้เกิดผลสัมฤทธิ์ ผลปรากฏว่าแบบทดสอบทั้ง 4 ฉบับ เหมาะสมกับรูปแบบของราส์ชเพียง 1 ฉบับ ผู้วิจัยจึงได้เสนอแนะว่า การนำรูปแบบของราส์ชมาใช้นั้น การวัดควรจะวัดในจุดประสงค์ที่แท้จริง ทั้งยังต้องพิจารณาถึงข้อตกลงเบื้องต้นของรูปของราส์ช ในการวิเคราะห์ข้อสอบด้วย

สำหรับ ออสเชอร์ (Oescher. 1983) ได้ศึกษาเปรียบเทียบการวิเคราะห์ข้อสอบ ด้วยวิธีการคัดเลือก 3 วิธี โดยการสร้างแบบทดสอบขึ้นมา 3 ฉบับ แต่ละฉบับมี 30 ข้อ ซึ่งคัดเลือกมาจากข้อสอบ 10 ข้อ แบบทดสอบฉบับแรกคัดเลือกโดยใช้ทฤษฎีดั้งเดิมประกอบด้วย ข้อสอบมีค่าอำนาจจำแนกสูงสุด แบบทดสอบที่ 2 คัดเลือกจากการพิจารณาค่าความยาก โดยใช้แบบของราส์ช และแบบทดสอบที่ 3 ใช้การสุ่มตัวอย่างแบบแบ่งชั้น ผลปรากฏว่าไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติของการคัดเลือกตามเกณฑ์ จากการวิเคราะห์ด้วย 3 วิธีดังกล่าว เมื่อใช้จุดมุ่งหมายของการสอบเพื่อพิจารณาตำแหน่งของนักเรียนวัดความรู้ (Continuum Knowledge Scale) แต่มีความแตกต่างกันในการวัดที่มีจุดมุ่งหมายแตกต่างกัน เกณฑ์คัดเลือกแบบคลาสสิคอลหรือแบบดั้งเดิมประมาณค่าเกินความจริง ในเรื่องความแปรปรวนจริงของการกระจายของคะแนนในแบบทดสอบซึ่งเป็นสาเหตุสำคัญในการปฏิบัติของนักเรียนอยู่ในพิสัยควอไทส์ต่ำ ในการตรงกันข้ามแบบทดสอบที่คัดเลือกโดยใช้รูปแบบของราส์ช ประเมินค่าต่ำกว่าความเป็นจริง

และอาคาส (Akas. 1984) ได้ศึกษาผลการเรียงข้อสอบตามความยากและความสามารถในแบบทดสอบวัดผลสัมฤทธิ์ชนิดเลือกตอบ โดยใช้การวิเคราะห์ด้วยรูปแบบของราส์ช ซึ่งใช้กลุ่มตัวอย่างจากการสอบเด็กนักเรียน 3 ห้อง ในปี ค.ศ. 1982 ที่มหาวิทยาลัยนิวยอร์ค ลักษณะของแบบทดสอบมี 3 ชุด คือ แบบทดสอบ 2 ชุดแรก เป็นการเรียงลำดับข้อสอบต่างกัน แต่ลำดับตัวเลือกเหมือนกัน แบบทดสอบชุดที่ 3 เป็นแบบทดสอบที่มีรูปแบบเหมือน 2 ชุดแรก แต่มีการเรียงตำแหน่งตัวถูกต่างกันไป ผลการศึกษาพบว่าข้อสอบหรือผลการตอบสลับที่ ลำดับการจะเปลี่ยนความยากของข้อสอบแต่ละข้อ แต่อย่างไรก็ตามแนวโน้มค่าความยากจะแปรเปลี่ยนไป เมื่อผลการตอบของข้อสอบสลับที่ลำดับกัน เมื่อเปรียบเทียบข้ามแบบทดสอบในแต่ละชุดไม่มีความแตกต่างกัน ผู้วิจัยสรุปว่าผลการเรียงลำดับข้อมีความสำคัญในทางปฏิบัติ เมื่อใช้กับข้อสอบเดี่ยว ๆ มากกว่าที่จะใช้เป็นชุด ๆ

สำหรับ บลูม (Bloom. 1984) ได้ศึกษาความเหมาะสมของรูปแบบของราส์ชกับการใช้ข้อมูลที่ได้จากแบบทดสอบอิงเกณฑ์ในการสอบ ด้วยแบบทดสอบทักษะพื้นฐาน (Basic Skills Tests) โดยใช้การวิเคราะห์องค์ประกอบ และรูปแบบการวิเคราะห์โครงสร้าง (Structural Equation Modeling Method) ของการตอบถูกเป็น 1 ตอบผิดเป็น 0 ปรากฏว่ารูปแบบของราส์ชสามารถใช้ได้

เหมาะสมกับข้อมูลที่ได้จากแบบทดสอบอิงเกณฑ์ ที่วัดหลาย ๆ จุดประสงค์ ค่าพารามิเตอร์ของข้อสอบที่ประมาณได้ไม่เปลี่ยนแปลง ซึ่งสามารถยืนยันได้จากทั้งข้อมูลจริงและการจำลองสถานการณ์ ทั้งยังมีความแกร่ง (Robustness) ในเรื่องของการทดสอบคุณลักษณะ ณ มิติเดียว (Unidimensionality)

และทอมคลินส์ (Tompklins. 1984) ได้ใช้รูปแบบของราส์ชศึกษาเกี่ยวกับผลความยาวของข้อสอบ ต่อการประมาณค่าความสามารถในการอ่านและจำแนกความรอบรู้ของนักเรียน ส่วนคุกไอเนอร์ และทาฟต์ (Cook and Talf. 1988 : 31-34) ได้ทำการศึกษาความคงที่ของการประมาณค่าพารามิเตอร์ของข้อสอบวัดผลสัมฤทธิ์ของการเรียนวิชาชีววิทยา ทำการทดสอบกับนักเรียน 2 กลุ่ม แบ่งข้อสอบออกเป็น 4 ชุด มีข้อสอบจำนวนไม่เท่ากัน ผลการศึกษาพบว่า การเรียนในระยะเวลาต่างกัน ผลสัมฤทธิ์ที่ได้จากแบบทดสอบทั้ง 4 ฉบับ แตกต่างกัน แต่ถ้าสอบในเวลาเดียวกัน ผลทั้ง 4 ฉบับจะเหมือนกัน ส่วนจำนวนข้อสอบแต่ละชุด ทำให้ค่าพารามิเตอร์เปลี่ยนไป ผลการศึกษาทำให้เห็นว่าค่าพารามิเตอร์ไม่คงที่ แปรเปลี่ยนไปตามกลุ่มตัวอย่าง

ฟอร์สเตอร์ (Rentz. 1979 : 7 ; citing Forster) พบว่า กลุ่มตัวอย่างที่ใช้ประมาณค่าพารามิเตอร์ 200 คน ก็เพียงพอที่จะให้ค่าพารามิเตอร์ที่ให้ค่าพารามิเตอร์ที่ได้มีค่าคงที่ กลุ่มตัวอย่างที่มากกว่า 1,000 คน จะไม่เพิ่มความคงที่ในการประมาณค่าพารามิเตอร์มากนัก

โกลด์แมนและราชู (Goldman and Raju. 1988 : 11-19) ได้ศึกษาผลของขนาดกลุ่มตัวอย่างที่แตกต่างกัน 3 กลุ่ม คือ 250 คน, 500 คน และ 1,000 คน เมื่อเทียบกับจำนวนของกลุ่มตัวอย่างที่เป็นมาตรฐาน 3,000 คน ในการกะประมาณค่าพารามิเตอร์ ผลการศึกษาพบว่า กลุ่มตัวอย่าง 250 คน มีความเหมาะสมสำหรับราชส์ชโมเดล แต่ถ้าสำหรับโมเดล 2 พารามิเตอร์จะต้องใช้กลุ่มตัวอย่าง 1,000 คนขึ้นไป จึงจะสามารถกะประมาณค่าพารามิเตอร์ได้ถูกต้อง

<u>งานวิจัยภายในประเทศ</u>

งานวิจัยเกี่ยวกับเรื่องนี้ในประเทศไทย มีผู้ได้ทำการวิจัยไว้ดังนี้

อวยพร วิบูลย์กาญจน์ (2526 : 60) ได้เปรียบเทียบผลการวิเคราะห์ข้อสอบอุปมาอุปไมยด้วยคลาสสิคอล กับราส์ชโมเดล โดยใช้กลุ่มตัวอย่างเป็นนักเรียนประถมปีที่ 6 จำนวน 1,884 คน เครื่องมือที่ใช้เป็นแบบทดสอบอุปมาอุปไมย และแบบทดสอบมาตรฐานวิชาวิทยาศาสตร์ของสำนักงานทดสอบทางการศึกษา และจิตวิทยามหาวิทยาลัยศรีนครินทรวิโรฒประสานมิตร วิเคราะห์ข้อสอบด้วยคอมพิวเตอร์โปรแกรม BICAL และโปรแกรมวิเคราะห์ข้อสอบทดสอบความแตกต่างระหว่างจำนวนข้อสอบ โดยใช้ค่าสถิติทดสอบซี (Z-test) แล้วพบว่าวิธีคลาสสิคอลได้ 34 ข้อ และ

ราส์ชโมเดลได้ 26 ข้อ หาความเชื่อมั่นโดยใช้สูตร K-R20 ได้ค่าความเชื่อมั่น .85 และราส์ชโมเดลได้ ส่วนความเที่ยงตรงเชิงสภาพแบบคลาสสิคอลได้ .49 และแบบราส์ชโมเดลได้ .46

และชวลิต โพธิ์นคร (2528) ได้เปรียบเทียบผลของการกำหนดจุดตัดแบบทดสอบอิงเกณฑ์ ชนิดเลือกตอบโดยวิธีของเบอร์ก วิธีประยุกต์ราส์ชโมเดลและวิธีกำหนดจุดผ่านต่ำสุด โดยใช้นักเรียนชั้นประถมปีที่ 4 โรงเรียนสังกัดสำนักงานประถมแห่งชาติ จังหวัดนครสวรรค์ จำนวน 325 คน เครื่องมือที่ใช้ในการศึกษา คือ แบบทดสอบวัด ความถนัดทางการเรียนด้านตัวเลข 2 ฉบับ เพื่อจัดระดับความสามารถของกลุ่มตัวอย่าง แบบทดสอบอิงเกณฑ์ กลุ่มทักษะวิชาคณิตศาสตร์ ชนิดเลือกตอบ 3 ฉบับ เพื่อจัดกลุ่มรอบรู้ และแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน กลุ่มทักษะวิชาคณิตศาสตร์ เพื่อวัดผลสัมฤทธิ์ของกลุ่มรอบรู้ ผลการศึกษาพบว่า

1. คะแนนผลสัมฤทธิ์ของกลุ่มรอบรู้ที่ผ่านจุดตัด โดยวิธีของเบอร์ก วิธีประยุกต์ราส์ชโมเดล และวิธีกำหนดเกณฑ์ผ่านระดับต่ำสุด แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และพบว่าค่าเฉลี่ยของกลุ่มรอบรู้ที่ผ่านจุดตัด วิธีประยุกต์ราส์ชโมเดลต่ำกว่ากลุ่มที่ผ่านจุดตัด โดยวิธีกำหนดเกณฑ์ระดับต่ำสุดอย่างมีนัยสำคัญที่ระดับ .05

2. ค่าความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบอิงเกณฑ์ที่ได้จากกำหนดจุดตัดโดยวิธีของเบอร์ก มีแนวโน้มให้ค่าสูงสุด รองลงมา ได้แก่ วิธีประยุกต์ราส์ชโมเดล และวิธีกำหนดเกณฑ์ผ่านระดับต่ำสุด ตามลำดับ

3. จุดตัดของแบบทดสอบอิงเกณฑ์ที่คำนวณโดยวิธีประยุกต์ราส์ชโมเดล ในกลุ่มตัวอย่างที่มีความสามารถสูงปานกลางและต่ำ พบว่าในกลุ่มสูงได้คะแนนจุดตัดแตกต่างไปจากกลุ่มปานกลางและกลุ่มต่ำ ส่วนระหว่างกลุ่มปานกลางและกลุ่มต่ำ ได้คะแนนจุดเท่ากัน แต่เมื่อคำนวณร้อยละของคะแนนจุดตัดพบว่าเท่ากับร้อยละ 50 ของคะแนนทั้งหมด เมื่อคิดเฉพาะข้อที่เป็นไปตามข้อตกลงของราส์ชโมเดลในแต่ละกลุ่ม

4. จุดตัดของแบบทดสอบอิงเกณฑ์ที่คำนวณโดยวิธีของเบอร์ก ในกลุ่มที่มีความสามารถสูง ปานกลาง และต่ำ พบว่าได้คะแนนจุดตัดที่แตกต่างกันและมีแนวโน้มว่า กลุ่มสูงสุดมีจุดตัดสูงสุด รองลงมาได้แก่ กลุ่มปานกลางและกลุ่มต่ำ ตามลำดับ

สำหรับ ปนัดดา วัยวัฒนะ (2528 : 71) ได้ศึกษาผลการวิเคราะห์ข้อสอบด้วยวิธีแบบเดิมกับวิธีโลจิสติกโมเดล ทั้งสามพารามิเตอร์ คือ โมเดลที่ใช้พารามิเตอร์ตัวเดียว พารามิเตอร์ 2 ตัว และพารามิเตอร์ 3 ตัว กลุ่มตัวอย่างเป็นนักเรียน ชั้นมัธยมศึกษาปีที่ 1 จำนวน 1,687 คน เครื่องมือเป็นแบบทดสอบคณิตศาสตร์ชนิดเลือกตอบ 5 ตัวเลือก จำนวน 50 ข้อ พบว่าจำนวนของข้อสอบที่ได้รับการคัดเลือกจากการวิเคราะห์แบบโลจิสติกโมเดล ทั้ง 3 พารามิเตอร์ มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ระดับ .05 ความเชื่อมั่นของข้อสอบที่ได้รับการคัดเลือกจากวิธีโลจิสติกที่ใช้พารา

มิเตอร์ 1 ตัว และ 3 ตัว เมื่อปรับจำนวนข้อให้เท่ากันแล้ว มีค่าไม่แตกต่างกัน และเมื่อปรับจำนวนข้อของข้อสอบให้เท่ากันแล้ว ค่าความเชื่อมั่นของการวิเคราะห์แบบดั้งเดิมกับวิธีโลจิสติกไม่แตกต่างกัน ถ้าพิจารณาความสอดคล้องกันของข้อสอบพบว่า การคัดเลือกข้อสอบจากการวิเคราะห์ด้วยพารามิเตอร์ตัวเดียว จะมีความสอดคล้องกับการวิเคราะห์ข้อสอบด้วยวิธีอื่นอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05

ส่วน สมพร บุญอิ่ม (2529) ได้ศึกษาความคงที่ของค่าพารามิเตอร์ ความยากในการวิเคราะห์ข้อสอบด้วยราส์ชโมเดลกับประชากรที่มีความสามารถต่างกัน 3 กลุ่ม มีความโด่งและความแปรปรวนเท่ากัน การแจกแจงของแต่ละกลุ่มคือ

1. แจกแจงเบ้บวกสำหรับกลุ่มที่มีความสามารถต่ำ
2. แจกแจงปกติสำหรับกลุ่มที่มีความสามารถปานกลาง
3. แจกแจงเบ้ลบสำหรับกลุ่มที่มีความสามารถสูง โดยใช้เทคนิคมอนติคาร์โล ซิมูเลชันจำลองในเครื่องคอมพิวเตอร์ ศึกษากับกลุ่มตัวอย่างกลุ่มละ 500 คน ข้อสอบจำนวน 30 ข้อ ทดลองซ้ำในแต่ละสถานการณ์ 100 ครั้ง ผลการศึกษาพบว่า

1. พารามิเตอร์ความยากมีค่าคงที่ในข้อที่มีความยากไม่สูงนัก หรือต่ำมาก กลุ่มที่มีความสามารถต่ำ และกลุ่มที่มีความสามารถสูง พารามิเตอร์มีความคงที่ 19 ข้อ กลุ่มที่มีความสามารถปานกลาง มีความคงที่ 22 ข้อ

2. การกระจายของค่าความคลาดเคลื่อนมาตรฐานของความยากของข้อสอบ เมื่อผู้สอบมีความสามารถต่ำ และผู้สอบมีความสามารถสูง มีแนวโน้มไม่เป็นสมมาตร และเมื่อผู้สอบมีความสามารถปานกลาง การกระจายเป็นสมมาตร

3. ค่าความคลาดเคลื่อนมาตรฐานของค่าความยากของข้อสอบทั้ง 3 กลุ่ม มีค่าไม่แตกต่างกัน

บัญชา แสนทวี (2530 : 110) ได้ประยุกต์เอารูปแบบของราส์ชมาใช้ในการออกแบบโค้งอินฟอร์เมชัน (Information) ของแบบสอบผลสัมฤทธิ์ เพื่อใช้ตามระดับความสามารถ เครื่องมือที่ใช้เป็นแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 จำนวน 2 ฉบับ ๆละ 60 ข้อ นำไปสอบกับกลุ่มตัวอย่างนักเรียนชั้นประถมปีที่ 6 จำนวน 1,429 คน นำผลการสอบมาวิเคราะห์ด้วยโปรแกรม BICAL เพื่อหาค่าพารามิเตอร์ของข้อสอบและพารามิเตอร์ความสามารถแล้วทำการเทียบมาตรา (Eguating) ในแบบทดสอบทั้ง 2 ฉบับ จัดกลุ่มความสามารถออกเป็น 3กลุ่ม คือ กลุ่มความสามารถสูง ปานกลาง และต่ำ แล้วออกแบบทดสอบให้เหมาะสมกับความสามารถของกลุ่ม ซึ่งผลการวิจัยมีดังนี้

1. เมื่อใช้แบบทดสอบวัดความสามารถต่าง ๆ กันทั้งกลุ่ม พบว่า ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ แตกต่างกันอย่างไม่มีนัยสำคัญที่ .05

2. เมื่อใช้แบบทดสอบวัดความสามารถต่าง ๆ กัน 2 ฉบับ กับกลุ่มที่มีความสามารถสูง กลาง และต่ำ พบว่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 และมีความแตกต่างกันอย่างไม่มีนัยสำคัญที่ระดับ .05 เมื่อกลุ่มผู้สอบมีความสามารถปานกลาง ทดสอบกับแบบทดสอบทั้ง 2 ฉบับ ดังกล่าว

3. เมื่อใช้แบบทดสอบทั้ง 2 ฉบับ กับกลุ่มที่มีความสามารถต่ำ พบว่าค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 ระหว่างแบบทดสอบที่เหมาะกับความสามารถต่ำ และความสามารถปานกลาง และแบบทดสอบที่เหมาะกับความสามารถต่ำ และความสามารถสูง ส่วนค่าความเชื่อมั่นของแบบทดสอบวัดความสามารถปานกลาง และความสามารถสูง ไม่แตกต่างกัน

4. เมื่อใช้แบบทดสอบวัดความสามารถต่าง ๆ กัน 2 ฉบับ กับกลุ่มที่มีความสามารถปานกลาง พบว่าค่าความเชื่อมั่นของแบบทดสอบ แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 ระหว่างแบบทดสอบที่เหมาะสมกับความสามารถปานกลาง และความสามารถสูง ส่วนแบบทดสอบวัดความสามารถปานกลาง และความสามารถต่ำ และที่วัดความสามารถสูง และความสามารถต่ำ แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05

5. เมื่อใช้แบบทดสอบที่วัดความสามารถต่าง ๆ กัน 2 ฉบับ กับกลุ่มที่ความสามารถสูง พบว่า ค่าความเชื่อมั่นของแบบทดสอบ แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 ระหว่างแบบทดสอบที่เหมาะสมกับความสามารถต่ำและปานกลาง และแบบทดสอบวัดความสามารถต่ำและสูง แตกต่างกันอย่างไม่มีนัยสำคัญที่ระดับ .05

ส่วน วัฒนา ขัดสี (2533 : 51) ได้ศึกษาความสัมพันธ์ระหว่างค่าพารามิเตอร์ที่ได้จากการวิเคราะห์ข้อสอบ และคะแนนความสามารถโดยทฤษฎีดั้งเดิม กับทฤษฎีการตอบข้อคำถามพบว่า ค่าความยากของข้อสอบที่ได้จากทฤษฎีดั้งเดิม กับทฤษฎีการตอบข้อคำถาม มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

สำหรับ บุญชิด รอดแก้ว (2533 : 72) ได้ศึกษาความไม่แปรปรวนของค่าพารามิเตอร์ของข้อสอบที่วิเคราะห์ด้วยโมเดลโลจิสติก 3 พารามิเตอร์ กับกลุ่มตัวอย่างที่มีลักษณะต่างกันพบว่าค่าอำนาจจำแนกของข้อสอบของกลุ่ม 1,052, 2,104 และ 3,155 คน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติสำหรับค่าความยากของกลุ่ม 1,052, 2,104 และ 3,155 คน แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

จากการศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง จะเห็นได้ว่าผลการวิจัยเกี่ยวกับคุณสมบัติความคงที่ของพารามิเตอร์ของข้อสอบที่วิเคราะห์ด้วยทฤษฎีการตอบข้อคำถาม ยังมีการขัดแย้งกันอยู่ และยังมีผู้ที่ทำการศึกษาในเรื่องนี้น้อยมาก หากได้มีการศึกษาคุณลักษณะของข้อสอบที่วิเคราะห์ด้วยทฤษฎีการตอบข้อคำถาม โดยใช้รูปแบบของราส์ช เมื่อจำนวนผู้สอบต่างกัน ซึ่งผลที่ได้จากการศึกษาครั้งนี้จะได้เป็นประโยชน์ทางด้านวัดผล และจะทำให้เกิดผลดีในการนำทฤษฎีไปใช้ต่อไป ดังนั้นผู้วิจัยจึงทำการศึกษา คุณลักษณะของแบบทดสอบอันประกอบไปด้วย ค่าความยาก ค่าความคลาดเคลื่อนมาตรฐาน จะแปรเปลี่ยนไปตามจำนวนกลุ่มตัวอย่างหรือไม่

<u>สมมุติฐานการวิจัย</u>

1. ค่าความยากของข้อสอบ ของแบบทดสอบเมื่อกลุ่มตัวอย่างที่มีจำนวนต่างกัน มีค่าแตกต่างกัน

2. ค่าความสามารถของผู้สอบของแบบทดสอบ เมื่อกลุ่มตัวอย่างมีจำนวนต่างกัน มีค่าแตกต่างกัน

3. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ จากแบบทดสอบเมื่อกลุ่มตัวอย่างมีจำนวนต่างกัน มีค่าแตกต่างกัน

บทที่ 3

วิธีดำเนินการวิจัย

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2540 สังกัดโรงเรียนกรมสามัญศึกษากลุ่ม 7 ในจังหวัดกรุงเทพมหานคร จำนวน 14 โรงเรียน จำนวนห้องเรียน 202 ห้องเรียน จำนวนนักเรียน 9,090 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2540 สังกัดโรงเรียนกรมสามัญศึกษากลุ่ม 7 ในกรุงเทพมหานคร เลือกมาโดยการสุ่มแบบแบ่งชั้น (Stratified random sampling) โดยมีโรงเรียนเป็นชั้น (Strata)และมีห้องเรียนเป็นหน่วยการสุ่ม (Sampling Unit) โดยสุ่มห้องเรียนมาโรงเรียนละ 20% ได้จำนวนห้องเรียน 40 ห้องเรียน จำนวนนักเรียน 1,800 คน ซึ่งเป็นกลุ่มตัวอย่างที่เชื่อมั่นได้ที่ 95% (อย่างน้อยต้องสุ่ม 383 คน) ดังตารางต่อไปนี้

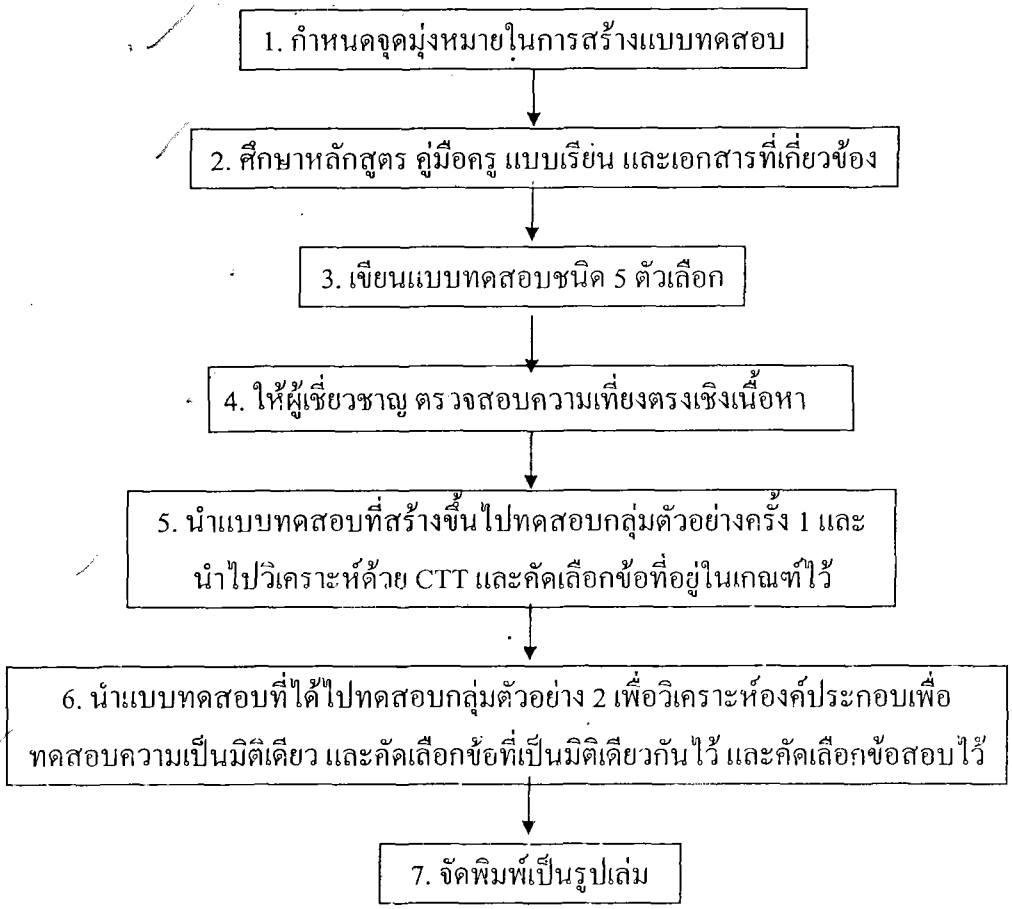
ตาราง 3 ตารางการสุ่มตัวอย่างที่ใช้ในการวิจัย

โรงเรียน	จำนวนนักเรียนทั้งหมด (ห้องเรียนทั้งหมด)		จำนวนคน (จำนวนห้องเรียน)
1. วัดอินทาราม	630	(14)	135 (3)
2. แจงร้อนวิทยา	620	(14)	135 (3)
3. ศึกษานารีวิทยา	650	(14)	135 (3)
4. บางปะกอกวิทยา	700	(16)	135 (3)
5. ศึกษานารี	680	(16)	135 (3)
6. บางมดวิทยา	630	(14)	135 (3)
7. ธนบุรีเทพีพลารักษ์	600	(12)	90 (2)
8. ราชโอรส	640	(14)	135 (3)
9. อิสลามวิทยาลัย	700	(16)	135 (3)
10.รัตนโกสินทร์สมโภช	720	(16)	135 (3)
11.มัธยมวัดดาวคะนอง	620	(16)	135 (3)
12.ทวีธาภิเษก	720	(16)	135 (3)
13.พุทธบูชา	560	(12)	90 (2)
14.วัดดาวคะนอง	620	(14)	135 (3)
รวม	9,090	(202)	1,800 (40)

<u>เครื่องมือที่ใช้ในการศึกษาค้นคว้า</u>

เครื่องมือที่ใช้ในการศึกษาครั้งนี้ เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 เรื่อง โจทย์สมการ ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521(ฉบับปรับปรุง 2533) ซึ่งแบบทดสอบที่ผู้วิจัยสร้างขึ้นตามดังแผนภูมิต่อไปนี้

ลำดับขั้นตอนการสร้างแบบทดสอบเลือกตอบ วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 เรื่อง โจทย์สมการ



วิธีดำเนินการสร้างแบบทดสอบ

ในการสร้างแบบทดสอบเรื่องโจทย์สมการ เพื่อให้เป็นเครื่องมือในการศึกษาครั้งนี้ มีขั้นตอนดำเนินการดังนี้

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบ เพื่อใช้เป็นเครื่องมือในการวิจัย
2. ศึกษาหลักสูตรคู่มือครูในการสอนวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 เรื่องโจทย์สมการ
3. เขียนแบบทดสอบเลือกตอบ จำนวน 5 ตัวเลือก เรื่องโจทย์สมการ จำนวน 60 ข้อ ดังตารางที่แสดงในภาคผนวก
4. นำแบบทดสอบที่สร้างขึ้นในข้อ 3. ไปให้ผู้เชี่ยวชาญตรวจสอบความเที่ยงตรงเชิงเนื้อหา โดยใช้เกณฑ์ในการพิจารณา 3 ใน 5 ของผู้เชี่ยวชาญ จึงถือว่าข้อสอบนั้นมีความเที่ยงตรงเชิงเนื้อหา ผลพิจารณาส่วนใหญ่ผู้เชี่ยวชาญเห็นว่ามีความเที่ยงตรงเชิงเนื้อหา มีบางข้อที่ผู้เชี่ยวชาญเห็นด้วย 4 ใน 5 และ 3 ใน 5

5. นำแบบทดสอบไปทดสอบกับกลุ่มนักเรียนในโรงเรียนศึกษานารี จำนวน 45 คน โรงเรียนบางประกอกวิทยา จำนวน 50 คน โรงเรียนอิสลามวิทยาลัย จำนวน 50 คน โรงเรียนทวีธาภิเษก จำนวน 50 คน รวมทั้งสิ้น 240 คน นำไปวิเคราะห์โดยใช้ทฤษฎีแบบดั้งเดิม ปรากฏว่าได้ข้อสอบที่มีค่าความยากตั้งแต่ 0.35 ถึง 0.95 ค่าอำนาจจำแนกตั้งแต่ 0.00 ถึง 0.52 คัดเลือกข้อสอบที่มีค่าความยากตั้งแต่ 0.20 ถึง 0.80 ไว้ ได้ข้อสอบที่อยู่ในเกณฑ์จำนวน 45 ข้อ ซึ่งมีค่าความยากอยู่ระหว่าง 0.35 ถึง 0.80 ค่าอำนาจจำแนกอยู่ระหว่าง 0.22 ถึง 0.52 (รายละเอียดดังตารางในภาคผนวก)

6. นำแบบทดสอบที่คัดเลือกไว้ในข้อ 4. ไปทดสอบกับนักเรียนในโรงเรียนวัดอินทราราม จำนวน 90 คน โรงเรียนศึกษานารี จำนวน 90 คน โรงเรียนบางปะกอกวิทยา จำนวน 95 คน โรงเรียนทวีธาภิเษก จำนวน 100 คน โรงเรียนโอรสจำนวน 95 คน รวมทั้งสิ้นจำนวน 520 คน นำผลไปวิเคราะห์องค์ประกอบ (Factor Analysis) ผลการวิเคราะห์ได้ค่าไอเก้น (Eigen Value) ดังนี้ 3.4, 2.8, 2.5, 2.3 ถึง 1.0 ซึ่งค่าไอเก้นตัวที่ 1 แตกต่างจากตัวที่ 2, 3 และองค์ประกอบอื่น ๆ ดังนั้น แบบทดสอบฉบับนี้จึงถือว่ามีความเป็นมิติเดียวและคัดเลือกข้อสอบไว้จำนวน 40 ข้อ เพื่อใช้ในการวิจัย โดยมีค่าความยากอยู่ระหว่าง 0.35 ถึง 0.75 ค่าอำนาจจำแนกอยู่ระหว่าง 0.22 ถึง 0.52 (ดังรายละเอียดในตารางภาคผนวก)

7. จัดพิมพ์เป็นรูปเล่ม

ลักษณะของแบบทดสอบ

แบบทดสอบที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ เรื่องโจทย์สมการ เป็นข้อสอบชนิดเลือกตอบ 5 ตัวเลือก ซึ่งผู้วิจัยสร้างขึ้นจำนวน 60 ข้อ มีลักษณะดังนี้

ตัวอย่างข้อสอบ

0. อ๊อดและแอ๊ด จ่ายค่าอาหารรวมกัน 72 บาท โดยที่อ๊อดจ่ายเป็น 2 เท่าของแอ๊ด สมการในข้อใดใช้แก้ปัญหานี้

1. $X^2 = 72$
2. $X = 72$
3. $2X = 72$
4. $3X = 72$ ข้อที่ถูกคือ 4
5. $2(2X+X) = 72$

00. เกษตรกรผู้หนึ่งเลี้ยงเป็ดกับหมูไว้ 40 ตัวเมื่อนับขาหมูและขาเป็ดปรากฏว่าขาหมูมากกว่าขาเป็ดอยู่ 28 ขา จงหาว่าเขาเลี้ยงสัตว์ทั้งสองชนิดต่างกันอยู่เท่าไร

1. 22 ตัว
2. 18 ตัว
3. 14 ตัว
4. 10 ตัว ข้อที่ถูกคือ 5
5. 4 ตัว

วิธีดำเนินการเก็บข้อมูล

1. ไปติดต่อกับทางโรงเรียนที่เป็นกลุ่มตัวอย่างเพื่อนัดวันเวลาในการทดสอบ โดยโรงเรียนที่เป็นกลุ่มตัวอย่างได้นัดวันสอบในช่วงวันที่ 23 ถึง 30 กุมภาพันธ์ 2541

2. เตรียมการดำเนินการสอบ ชี้แจงกรรมการคุมสอบเตรียมห้องสอบ

3. นำแบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 ที่ผู้วิจัยสร้างขึ้นไปทดสอบกับกลุ่มตัวอย่าง โดยผู้วิจัยไปเก็บเองและมีบางโรงเรียนผู้วิจัยได้ให้เพื่อนไปดำเนินการเก็บข้อมูลให้

4. นำกระดาษคำตอบไปตรวจให้คะแนน หาค่าสถิติพื้นฐาน วิเคราะห์องค์ประกอบ วิเคราะห์ด้วยราส์ชโมเดลและทดสอบสมมุติฐาน

การวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลสำหรับการวิจัยครั้งนี้ผู้วิจัยทำการวิเคราะห์ด้วยเรื่องคอมพิวเตอร์ โดยใช้โปรแกรม SPSS-X วิเคราะห์องค์ประกอบ โปรแกรมวิเคราะห์ข้อสอบ โปรแกรม BICAL โดยแบ่งข้อมูล ตามกลุ่มตัวอย่าง 6 กลุ่ม ดังนี้

กลุ่มที่ 1 200 คน

กลุ่มที่ 2 300 คน

กลุ่มที่ 3 450 คน

กลุ่มที่ 4 675 คน

กลุ่มที่ 5 1,015 คน

กลุ่มที่ 6 1,610 คน

โดยแต่ละกลุ่มได้มาจากการสุ่มอย่างง่ายแบบใส่คืน โดยการเพิ่มขนาดกลุ่มตัวอย่างขึ้นกลุ่มละ 50%

สถิติที่ใช้วิเคราะห์ข้อมูล

1. สถิติพื้นฐานของคะแนนจากแบบทดสอบ ได้แก่

   1.1 คะแนนเฉลี่ย ($\bar{X}$)

   1.2 ค่าความเบี่ยงเบนมาตรฐาน (S)

   1.3 ค่าความแปรปรวน (S^2)

2. คุณลักษณะของแบบทดสอบ

   2.1 ค่าความยาก (b)

   2.2 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถจากสูตร (Warm. 1978:77)

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

ในการวิจัยครั้งนี้ใช้โปรแกรม BICAL

โดยแต่ละกลุ่มได้มาจากการสุ่มตัวอย่างแบ่งใส่คืนโดยการเพิ่มขนาดกลุ่มตัวอย่างขึ้นกลุ่มละ 50%

3. คะแนนความสามารถจากสูตร

$$Pi(\theta) = \frac{e^{D(\theta-bi)}}{1+e^{D(\theta-bi)}} \; ; \; i = 1,2,3, \ldots ,n$$

วิเคราะห์หาค่าคุณลักษณะของแบบทดสอบโดยใช้โปรแกรม BICAL

4. วิเคราะห์องค์ประกอบด้วยโปรแกรม SPSS-X
5. สถิติที่ใช้ทดสอบสมมุติฐาน

ทดสอบความแตกต่างค่าความยากง่าย ค่าความสามารถของผู้สอบและค่าความคลาดเคลื่อนมาตรฐานในการประมานค่าของผู้สอบของกลุ่มตัวอย่างใช้ One-Way Anova ถ้าพบความแตกต่างแล้วทดสอบเป็นรายคู่ด้วยวิธีของเชฟเฟ (Scheff's test) (ล้วนสายยศและอังคณา สายยศ 2531:260)

$$F = \frac{MS_b}{MS_w}$$

เมื่อ MS_b แทนความแปรปรวนระหว่างกลุ่ม

MS_w แทนความแปรปรวนภายในกลุ่ม

$$S = \sqrt{(K-1)F_{\alpha(df1,df2)}} \sqrt{MS_e \left[\sum_{(i=1)}^{K} \frac{(C_j)2}{n_j} \right]}$$

เมื่อ S แทนค่าวิกฤตของ Scheffe'

$F\alpha(df_1, df_2)$ แทนค่าแจกแจงค่า F ที่ α มีค่า .05 และ $df_1 = df_b$ $df_2 = df_w$

MS_e ค่าความคลาดเคลื่อนของความแปรปรวนในตารางวิเคราะห์ความแปรปรวน

C_j แทนสัมประสิทธิ์ Contrast ซึ่งจะมีค่าเป็น 1,-1,1...

n_j แทนจำนวนคนในกลุ่มตัวอย่างแต่ละกลุ่ม

บทที่ 4
ผลการวิเคราะห์ข้อมูล

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

b	แทน	ค่าความยากของข้อสอบ
$\bar{b}$	แทน	ค่าความยากเฉลี่ยของข้อสอบ
SD_b	แทน	ค่าความเบี่ยงเบนมาตรฐานของความยากของข้อสอบ
N	แทน	จำนวนกลุ่มผู้สอบ
K	แทน	จำนวนข้อสอบ
SS	แทน	ผลบวกกำลังสองของคะแนนแบบทดสอบ
MS	แทน	ความแปรปรวนของคะแนนแบบทดสอบ
df	แทน	ช่วงชั้นของความเป็นอิสระ
Sc	แทน	ค่าวิกฤตของ Schffe's

การเสนอผลการวิเคราะห์ข้อมูล

ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลดังนี้

1. ค่าสถิติพื้นฐาน
   1.1 ค่าความยากของข้อสอบ
   1.2 ค่าความสามารถของผู้สอบ
   1.3 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ
2. เปรียบเทียบค่าความยากของข้อสอบระหว่างที่วิเคราะห์ได้จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน
3. เปรียบเทียบค่าความสามารถของผู้สอบระหว่างที่วิเคราะห์ได้จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน
4. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ระหว่างที่วิเคราะห์ได้จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน

ผลการวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐาน

จากการนำแบบทดสอบไปทดสอบกับนักเรียน 1,610 คน แล้วแบ่งกลุ่มตัวอย่างออกเป็น 6 กลุ่ม โดยใช้วิธีการสุ่มแบบอย่างง่าย โดยมีขนาดคนเพิ่มขึ้น 50% ของกลุ่มเดิม ซึ่งได้แก่กลุ่ม 2,00, 300, 450, 675, 1,015 และ 1,610 คน ได้ค่าสถิติพื้นฐาน จำแนกตามลักษณะตัวแปรของข้อสอบดังนี้

1.1 ค่าความยากของข้อสอบ

ผู้วิจัยได้นำค่าความยากของข้อสอบที่วิเคราะห์ได้จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน โดยใช้วิธีราส์ชโมเดล ได้ค่าความยากดังนี้

ตาราง 4 ค่าความยากของแบบทดสอบของกลุ่มตัวอย่างจำนวน 200 คน 300 คน 450 คน 675 คน 1,015 คน และ 1,610 คน

กลุ่ม	จำนวนข้อที่สอดคล้อง ICC	b	$\bar{b}$	SD_b
200 คน	28	-1.81 ถึง 3.04	0.065	0.17
300 คน	28	-1.82 ถึง 3.01	0.000	0.14
450 คน	33	-2.70 ถึง 3.31	0.001	0.13
675 คน	34	-2.25 ถึง 3.26	0.007	0.10
1,015 คน	35	-2.63 ถึง 3.28	-0.048	0.07
1,610 คน	35	-3.19 ถึง 3.19	-0.077	0.07

จากตาราง 4 ปรากฏว่ากลุ่ม 200 คน มีค่าความยากตั้งแต่ -1.81 ถึง 3 04 มีค่าความยากเฉลี่ย 0.065 มีส่วนเบี่ยงเบนมาตรฐาน 0.17 กลุ่ม 300 คน มีค่าความยากตั้งแต่ -1.82 ถึง 3.01 มีค่าความยากเฉลี่ย 0.000 มีส่วนเบี่ยงเบนมาตรฐาน 0.14 กลุ่ม 450 คน มีค่าความยากตั้งแต่ -2.70 ถึง 3.26 มีค่าความยากเฉลี่ย 0.001 มีส่วนเบี่ยงเบนมาตรฐาน 0.13 กลุ่ม 675 คน มีค่าความยากเฉลี่ย 0.007 มีส่วนเบี่ยงเบนมาตรฐาน 0.10 กลุ่ม 1,015 คน มีค่าความยากตั้งแต่ -2.63 ถึง 3.28 มีค่าความยากเฉลี่ย -0.048 กลุ่ม 1,610 คน มีค่าความยากตั้งแต่ -3.19 ถึง 3.19 มีค่าความยากเฉลี่ย -0.077 มีส่วนเบี่ยงเบนมาตรฐาน 0.07 แสดงให้เห็นว่าแบบทดสอบฉบับนี้มีค่าความยากแต่ละกลุ่มมีค่าความ

ยากใกล้ 0 ซึ่งหมายความว่ามีค่าความยากอยู่ระดับปานกลาง ส่วนการกระจายคะแนนในการสอบของแต่ละกลุ่มที่มีจำนวนต่างกัน มีค่าอยู่ระหว่าง 0.07 ถึง 0.17 แสดงว่ามีการ กระจายใกล้เคียงกัน

1.2 ค่าความสามารถของผู้สอบ

ผู้วิจัยได้นำค่าความสามารถของผู้สอบที่วิเคราะห์ได้จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน โดยใช้วิธีของราส์ชโมเดล ได้ค่าความสามารถของผู้สอบดังนี้

ตาราง 5 ค่าความสามารถของผู้สอบของกลุ่มตัวอย่างจำนวน 200 คน 300 คน 450 คน 675 คน 1,015 คน และ 1,610 คน

กลุ่มตัวอย่าง	จำนวนข้อที่สอดคล้อง ICC	θ	$\bar{\theta}$
200 คน	28	-.00 ถึง 2.24	.80
300 คน	28	-.34 ถึง 2.30	.78
450 คน	33	-.07 ถึง 2.67	1.07
675 คน	34	-0.01 ถึง 2.69	1.03
1,015 คน	35	.08 ถึง 2.82	1.06
1,610 คน	35	-.41 ถึง 2.76	.96

จากตาราง 5 ค่าความสามารถของผู้สอบของกลุ่ม 200 คน มีค่าตั้งแต่ -.33 ถึง 2.24 ค่าความสามารถเฉลี่ย .80 ค่าความสามารถของผู้สอบของกลุ่ม 300 คน มีค่าตั้งแต่ -.34 ถึง 2.30 ค่าความสามารถเฉลี่ย .78 ค่าความสามารถของผู้สอบของกลุ่ม 450 คน มีค่าตั้งแต่ -.07 ถึง 2.67 ค่าความสามารถเฉลี่ย 1.07 ค่าความสามารถของผู้สอบของกลุ่ม 675 คน มีค่าตั้งแต่ -0.01 ถึง 2.69 ค่าความสามารถเฉลี่ย 1.03 ค่าความสามารถของผู้สอบของกลุ่ม 1,015 คน มีค่าตั้งแต่ -.08 ถึง 2.82 ค่าความสามารถเฉลี่ย 1.06 ค่าความสามารถของผู้สอบของกลุ่ม 1,610 คน มีค่าตั้งแต่ -.41 ถึง 2.76 ค่าความสามารถเฉลี่ย .96 แสดงว่าความสามารถของกลุ่มตัวอย่างทั้ง 6 กลุ่มมีค่าอยู่ระหว่าง 0.80 ถึง 1.07 ซึ่งมีค่าอยู่ในช่วงคะแนนที่ระดับปานกลางถึงค่อนข้างยาก

1.3 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ

ผู้วิจัยได้นำค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบที่วิเคราะห์ได้จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน โดยใช้ราส์ชโมเดลได้ค่าดังนี้

ตาราง 6 ค่าสถิติพื้นฐานของความคลาดเคลื่อนมาตรฐานของแบบทดสอบของกลุ่มผู้สอบ 200 คน, 300 คน, 450 คน, 675 คน, 1,015 คน และ 1,610 คน

กลุ่มตัวอย่าง	จำนวนข้อที่สอดคล้อง ICC	SE	$\overline{SE}$
200 คน	28	0.40 ถึง 0.79	0.51
300 คน	28	0.40 ถึง 0.79	0.52
450 คน	33	0.38 ถึง 0.79	0.54
675 คน	34	0.37 ถึง 0.79	0.51
1,015 คน	35	0.38 ถึง 0.79	0.52
1,610 คน	35	0.39 ถึง 0.65	0.44

จากตาราง 6 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าของผู้สอบ

กลุ่ม 200 คน มีค่าตั้งแต่ 0.40 ถึง 0.79 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐาน 0.51

กลุ่ม 300 คน มีค่าตั้งแต่ 0.40 ถึง 0.79 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐาน 0.52

กลุ่ม 450 คน มีค่าตั้งแต่ 0.38 ถึง 0.79 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐาน 0.54

กลุ่ม 675 คน มีค่าตั้งแต่ 0.37 ถึง 0.79 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐาน 0.51

กลุ่ม 1,015 คน มีค่าตั้งแต่ 0.38 ถึง 0.79 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐาน 0.52

กลุ่ม 1,610 คน มีค่าตั้งแต่ 0.39 ถึง 0.65 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐาน 0.44

แสดงให้เห็นว่าแบบทดสอบฉบับนี้มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบมีค่าอยู่ระหว่าง 0.44 ถึง 0.54 หมายความว่าค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบมีค่าใกล้เคียงกัน

2. เปรียบเทียบค่าความยากของข้อสอบระหว่างที่วิเคราะห์ได้จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน

ผู้วิจัยได้นำค่าความยากของข้อสอบที่วิเคราะห์ได้จากกลุ่มตัวอย่างที่มีจำนวนไม่เท่ากันมาทดสอบความแตกต่าง โดยการวิเคราะห์ความแปรปรวน (One-Way Anova) ได้ผลดังนี้

ตาราง 7 การทดสอบความแตกต่างของค่าความยากของข้อสอบของกลุ่มผู้สอบจำนวน 200 คน, 300 คน, 450 คน, 675 คน, 1,015 คน และ 1,610 คน

แหล่งตัวแปร	SS	df	MS	F
ระหว่างกลุ่ม	0.1496	5	0.0299	0.0250
ความคลาดเคลื่อน	222.2756	186	0.1950	
รวม	222.4252	191		

จากตาราง 7 แสดงให้เห็นว่าค่าความยากของข้อสอบของกลุ่มตัวอย่างที่มีจำนวน 200 คน 300 คน 450 คน 675 คน 1,015 คน และ 1,610 คน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นคือค่าความยากของข้อสอบมีค่าคงที่ หรือไม่แปรปรวนระหว่างกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน

3. เปรียบเทียบค่าความสามารถของผู้สอบระหว่างที่วิเคราะห์ได้จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน

ผู้วิจัยได้นำค่าความสามารถของผู้สอบที่วิเคราะห์ได้จากกลุ่มตัวอย่างที่มีจำนวนไม่เท่ากันมาทดสอบความแตกต่าง โดยการวิเคราะห์ความแปรปรวน (One-Way Anova) ได้ผลดังนี้

ตาราง 8 การทดสอบความแตกต่างของค่าความสามารถของข้อสอบของกลุ่มผู้สอบจำนวน 200 คน, 300 คน, 450 คน, 675 คน, 1,015 คน และ 1,610 คน

แหล่งตัวแปร	SS	df	MS	F
ระหว่างกลุ่ม	0.0019	5	0.0004	0.0001
ความคลาดเคลื่อน	604.6327	181	3.3405	
รวม	604.6345	186		

จากตาราง 8 แสดงให้เห็นว่าค่าความสามารถของข้อสอบของกลุ่มตัวอย่างที่มีจำนวน 200 คน 300 คน 450 คน 675 คน 1,015 คน และ 1,610 คน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นคือค่าความสามารถของข้อสอบมีค่าคงที่ หรือมีค่าไม่แปรปรวนระหว่างกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน

4. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่างที่วิเคราะห์ได้จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน

ตาราง 9 การทดสอบความแตกต่างของค่าความสามารถของข้อสอบของกลุ่มจำนวน 200 คน, 300 คน, 450 คน, 675 คน, 1,015 คน และ 1,610 คน

แหล่งตัวแปร	SS	df	MS	F
ระหว่างกลุ่ม	0.0287	5	0.0057	0.1703
ความคลาดเคลื่อน	6.0925	181	0.0337	
รวม	6.1212	186		

จากตาราง 9 แสดงให้เห็นว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของข้อสอบของกลุ่มตัวอย่างที่มีจำนวน 200 คน 300 คน 450 คน 675 คน 1,015 คน และ 1,610 คน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นคือค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของข้อสอบมีค่าคงที่ หรือไม่แปรปรวนระหว่างกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อเปรียบเทียบค่าความยากของแบบทดสอบ ที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน
2. เพื่อเปรียบเทียบค่าความสามารถของผู้สอบ ที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน
3. เพื่อเปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอนที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2540 สังกัดโรงเรียนกรมสามัญศึกษากลุ่ม 7 ในกรุงเทพมหานคร เลือกมาโดยการสุ่มแบบแบ่งชั้น (Stratified random sampling) โดยมีโรงเรียนเป็นชั้น (Strata) และมีห้องเรียนเป็นหน่วยการสุ่ม (Sampling Unit) โดยสุ่มห้องเรียนมา 20% ได้จำนวนห้องเรียน 40 ห้องเรียน จำนวนนักเรียน 1,800 คน

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้านี้เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 เรื่องโจทย์สมการ เป็นแบบเลือกตอบ 5 ตัวเลือก จำนวน 40 ข้อ ที่ผู้วิจัยสร้างขึ้น

วิธีดำเนินการเก็บรวบรวมข้อมูล

ผู้วิจัยได้นำแบบทดสอบไปทดสอบกับนักเรียนกลุ่มตัวอย่างโดยดำเนินการตามขั้นตอนดังนี้

1. วางแผนดำเนินการสอบ โดยติดต่อขอความร่วมมือในการนำเครื่องมือไปทดสอบกับนักเรียนกลุ่มตัวอย่าง ตลอดจนนัดวันเวลาทำการทดสอบ
2. จัดเตรียมข้อสอบให้เพียงพอกับจำนวนนักเรียนที่สอบในแต่ละครั้ง

3. ชี้แจงนักเรียนในกลุ่มตัวอย่างให้เข้าใจและทราบวัตถุประสงค์ในการสอบและขอความร่วมมือในการสอบเพื่อให้ได้ผลตามความเป็นจริง

4. นำแบบทดสอบไปทดสอบกับกลุ่มตัวอย่าง

5. นำกระดาษคำตอบมาตรวจให้คะแนน แล้วนำผลมาวิเคราะห์หาค่าสถิติพื้นฐาน คำนวณหาค่าความยาก ค่าความสามารถของผู้สอบ และค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถด้วยโปรแกรม BICAL และทดสอบสมมุติฐานการวิจัย

การวิเคราะห์ข้อมูล

ผู้วิจัยได้ดำเนินการวิเคราะห์ข้อมูลดังนี้

1. ค่าสถิติพื้นฐาน
   1.1 ค่าความยากของข้อสอบ
   1.2 ค่าความสามารถของผู้สอบ
   1.3 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ
2. เปรียบเทียบค่าความยากของแบบทดสอบ
3. เปรียบเทียบค่าความสามารถของผู้สอบ
4. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ

สรุปผลการวิเคราะห์ข้อมูล

ผลการวิเคราะห์ข้อมูลสรุปได้ดังนี้

1. ค่าความยากของข้อสอบที่วิเคราะห์ด้วยราส์ชโมเดลที่มีจำนวนกลุ่มตัวอย่างต่างกันมีค่าอยู่ระหว่าง -3.19 ถึง 3.31 โดยกลุ่มตัวอย่างจำนวน 200 คน มีค่าความยากตั้งแต่ -1.81 ถึง 3.04 จำนวนกลุ่มตัวอย่าง 300 คน มีค่าความยากตั้งแต่ -1.82 ถึง 3.01 จำนวนกลุ่มตัวอย่าง 450 คน มีค่าความยากตั้งแต่ -2.07 ถึง 3.31 จำนวนกลุ่มตัวอย่าง 675 คน มีค่าความยากตั้งแต่ -2.25 ถึง 3.26 จำนวนกลุ่มตัวอย่าง 1,015 คน มีค่าความยากตั้งแต่ -2.63 ถึง 3.28 จำนวนกลุ่มตัวอย่างจำนวน 1,610 คน มีค่าความยากตั้งแต่ -3.19 ถึง 3.19 แสดงว่าค่าความยากของข้อสอบมีค่าความยากตั้งแต่ง่ายมากจนถึงยากมาก เมื่อนำค่าความยากของข้อสอบที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน มาเปรียบเทียบกันพบว่ามีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. ค่าความสามารถของผู้สอบที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน มีค่าอยู่ระหว่าง -0.01 ถึง 2.82 โดยกลุ่มตัวอย่าง 200 คน มีค่าความสามารถของผู้สอบตั้งแต่ -0.33 ถึง 2.82

จำนวนกลุ่มตัวอย่าง 300 คน มีค่าความยากตั้งแต่ -0.34 ถึง 2.30 จำนวนกลุ่มตัวอย่างจำนวน 450 คน มีค่าความยากตั้งแต่ -0.07 ถึง 2.67 จำนวนกลุ่มตัวอย่างจำนวน 675 คน มีค่าความยากตั้งแต่ -0.01 ถึง 2.69 จำนวนกลุ่มตัวอย่างจำนวน 1,015 คน มีค่าความยากตั้งแต่ -0.08 ถึง 2.82 จำนวนกลุ่มตัวอย่างจำนวน 1,610 คน มีค่าความยากตั้งแต่ –0.41 ถึง 2.76 แสดงว่าค่าความสามารถของผู้สอบมีค่าค่อนข้างต่ำจนถึงสูงมาก เมื่อนำค่าความสามารถของผู้สอบที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกันมาเปรียบเทียบกัน พบว่ามีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน มีค่าอยู่ระหว่าง 0.37 ถึง 0.79 โดยกลุ่มตัวอย่างจำนวน 200 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าตั้งแต่ 0.40 ถึง 0.79 จำนวนกลุ่มตัวอย่าง 300 คน มีค่าตั้งแต่ 0.40 ถึง 0.79 จำนวนกลุ่มตัวอย่าง 450 คน มีค่าตั้งแต่ 0.38 ถึง 0.79 จำนวนกลุ่มตัวอย่างจำนวน 675 คน มีค่าตั้งแต่ 0.37 ถึง 0.79 จำนวนกลุ่มตัวอย่าง 1,015 คน มีค่าตั้งแต่ 0.38 ถึง 0.79 จำนวนกลุ่มตัวอย่าง 1,610 คน มีค่าตั้งแต่ 0.39 ถึง 0.65 เมื่อนำค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกันมาเปรียบเทียบกัน พบว่ามีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

การอภิปรายผล

การศึกษาเปรียบเทียบคุณลักษณ์ของแบบทดสอบเลือกตามระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกัน ที่วิเคราะห์ด้วยราส์ชโมเดล สามารถอภิปรายผลได้ดังนี้

1. ค่าความยากของแบบทดสอบที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกันมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมุติฐานที่ตั้งไว้ ทั้งนี้เนื่องจากกลุ่มตัวอย่างที่ศึกษาในครั้งนี้เป็นตัวแทนที่ดีของประชากร เป็นไปตามทฤษฎีการสุ่มตัวอย่าง ทำให้กลุ่มตัวอย่างในแต่ละกลุ่มมีลักษณะที่เหมือนกับประชากรคือกลุ่มตัวอย่างในแต่ละกลุ่มจะประกอบด้วยนักเรียนที่มีความสามารถสูง กลาง และต่ำ ปะปนกันอยู่ ทำให้กลุ่มตัวอย่างแต่ละกลุ่มมีลักษณะความสามารถของนักเรียนคล้ายกัน จึงทำให้ค่าความยากของแบบทดสอบที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน มีค่าใกล้เคียงกันและแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ แสดงว่าค่าความยากของข้อสอบที่วิเคราะห์ได้จากกลุ่มตัวอย่างที่มีจำนวนต่างกันมีค่าคงที่ หรือไม่แปรปรวนไปตามกลุ่มตัวอย่าง ซึ่งเป็นไปตามคำกล่าวของลอร์ด (Lord. 1980 : 35) ที่กล่าวว่าค่าพารา มิเตอร์ของข้อสอบไม่แปรปรวนไปตามกลุ่มผู้สอบและผลการศึกษาครั้งนี้ยังได้ผลสอดคล้องกับการศึกษาของค็อค (Kock. 1980 : 15032) ที่ศึกษาค่าพารามิเตอร์จากการใช้แบบทดสอบวัดเจตคติตามแนวคิดของลิเกิร์ตที่ยืนยันว่าค่าพารามิเตอร์ของข้อสอบมีคุณสมบัติไม่แปรเปลี่ยนไปตามกลุ่มตัวอย่าง

นอกจากนี้ผลการศึกษาครั้งนี้ยังสอดกับผลการศึกษาของสุทิน บุญญาภิบาล (2537 : 49) ที่พบว่า ค่าความยากของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ เมื่อกลุ่มตัวอย่างต่างกันมีค่าความยากแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ และการศึกษาครั้งนี้ยังได้ผลไม่ตรงกับการศึกษาของบุญชิด รอดแก้ว (2533 : 75) ที่พบว่าค่าความยากของแบบทดสอบกลุ่มที่มีความสามารถสูงกับกลุ่มที่มีความสามารถปานกลาง กลุ่มที่มีความสามารถปานกลางและกลุ่มที่มีความสามารถต่ำ กลุ่ม 1,052 คน กับกลุ่ม 2,104 คน และกลุ่ม 1,052 คน กับกลุ่ม 3,155 คน มีค่าความยากของข้อสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2. ค่าความสามารถของผู้สอบที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมุติฐานที่ตั้งไว้ ทั้งนี้เนื่องจากกลุ่มตัวอย่างที่ศึกษาในครั้งนี้เป็นตัวแทนที่ดีของประชากร เป็นไปตามทฤษฎีการสุ่มตัวอย่าง ทำให้กลุ่มตัวอย่างในแต่ละกลุ่มมีลักษณะเหมือนกับประชากร คือกลุ่มตัวอย่างในแต่ละกลุ่มประกอบไปด้วย นักเรียนที่มีความสามารถสูง กลาง ต่ำ ปะปนกันอยู่ ทำให้กลุ่มตัวอย่างแต่ละกลุ่มมีลักษณะความสามารถของนักเรียนคล้ายกัน จึงทำให้ค่าความสามารถของผู้สอบที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน มีค่าใกล้เคียงกัน และแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งผลการศึกษาครั้งนี้สอดคล้องกับการศึกษาของค็อค (Kcck. 1980 : 15-32) ที่ศึกษาค่าพารามิเตอร์จากการใช้แบบทดสอบวัดเจตคติตามแนวของลิเกิร์ต ที่ยืนยันว่าค่าความสามารถของผู้สอบไม่แปรเปลี่ยนไปตามกลุ่มตัวอย่าง นั่นคือค่าความสามารถของผู้สอบมีค่าคงที่

3. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมุติฐานที่ตั้งไว้ ทั้งนี้เนื่องจากกลุ่มตัวอย่างที่ศึกษาในครั้งนี้เป็นตัวแทนที่ดีของประชากร เป็นไปตามทฤษฎีการสุ่มตัวอย่าง ทำให้กลุ่มตัวอย่างในแต่ละกลุ่มมีลักษณะเหมือนกับประชากร คือกลุ่มตัวอย่างในแต่ละกลุ่มประกอบไปด้วยนักเรียนที่มีความสามารถสูง กลาง และต่ำ ปะปนกันอยู่ ทำให้กลุ่มตัวอย่างแต่ละกลุ่มมีลักษณะความสามารถของนักเรียนคล้ายกัน มีค่าใกล้เคียงกัน จึงทำให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบมีค่าใกล้เคียงกัน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นคือค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่ามีค่าคงที่ไม่แปรปรวนไปตามกลุ่มตัวอย่างที่มีจำนวนต่างกัน

ดังนั้นอาจสรุปได้ว่าในการวิเคราะห์ค่าความยาก ค่าความสามารถของผู้สอบ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าของผู้สอบ ของข้อสอบด้วยวิธีของราส์ชโมเดล ควรใช้กลุ่มตัวอย่างอย่างน้อย 200 คนขึ้นไป ซึ่งจะทำให้ค่าความยาก ค่าความสามารถของผู้สอบ และค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า มีค่าคงที่หรือไม่แปรปรวน

ข้อเสนอแนะ

1. กลุ่มตัวอย่างที่เหมาะสมที่จะนำมาวิเคราะห์ข้อสอบโดยวิธีของราส์ชโมเดลนั้น ควรมีขนาดอย่างน้อยที่สุดควรจะมีขนาดกลุ่มตัวอย่างจำนวน 200 คน เพราะจะทำให้การวิเคราะห์ข้อสอบเป็นไปอย่างรวดเร็วและประหยัดค่าใช้จ่าย และเวลา

2. ควรมีการศึกษาในทำนองเดียวกันนี้ในเนื้อหาวิชาอื่น หรือในระดับชั้นที่แตกต่างกันไป

3. ควรมีการศึกษาในทำนองเดียวกัน โดยใช้ 2 พารามิเตอร์ 3 พารามิเตอร์ มาวิเคราะห์ข้อมูล และหาความสัมพันธ์กับผลการศึกษาในครั้งนี้

บรรณานุกรม

บรรณานุกรม

โกวิทย์ ประวาลพฤกษ์. "ความเข้าใจพื้นฐานเกี่ยวกับโค้งลักษณะของสอบ," วารสารการวัดผลการศึกษา. 4(2) : 1 - 21 ; กันยายน - ธันวาคม 2525.

ขรรค์ชัย คงเสน่ห์. การสร้างแบบทดสอบสองขั้นตอน โดยใช้ทฤษฎีการตอบข้อคำถาม. ปริญญานิพนธ์ กศ.ม. กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2530. อัดสำเนา.

จิราพร ไกรสรศิวเวท. ความสัมพันธ์ระหว่างรูปแบบบางประการของแบบทดสอบรูปปิรามิดกับความสามารถทางการเรียนในวิชาคณิตศาสตร์. ปริญญานิพนธ์ กศ.ม. กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2529. อัดสำเนา.

เชิดศักดิ์ โฆวาสินธุ์. การวัดทัศนคติและบุคลิกภาพ. กรุงเทพ ฯ : สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2520.

ดวงเดือน พันธุนาวิน. "อิทธิพลของสังคมต่อทัศนคติของวัยรุ่น," รายงานการวิจัย. 18 : 3 - 7 ; เมษายน 2518.

นภา กาญจนกิจโสภณ. การเปรียบเทียบคะแนนสอบระหว่างการให้น้ำหนักคะแนนรายข้อเท่ากันตามทฤษฎีมาตรฐานเติมกับการให้น้ำหนักคะแนนรายข้อต่างกัน ตามทฤษฎีการตอบสนองข้อกระทง. วิทยานิพนธ์ ค.ม. กรุงเทพ ฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2533. อัดสำเนา.

บัญชา แสนทวี. การประยุกต์รูปแบบของราส์ชในการออกแบบโค้งแสดงสารสนเทศของแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ตามระดับความสามารถของนักเรียนชั้นประถมศึกษาปีที่ 6. วิทยานิพนธ์ ค.ม. กรุงเทพ ฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2530. อัดสำเนา.

บุญชิต รอดแก้ว. การศึกษาความไม่แปรปรวนของค่าพารามิเตอร์ของข้อสอบที่วิเคราะห์ด้วยโมเดลโลจีสติคสามพารามิเตอร์กับกลุ่มตัวอย่างที่มีลักษณะต่างกัน. ปริญญานิพนธ์ กศ.ม. กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2533. อัดสำเนา.

ผจงจิต อินทสุวรรณ. ทฤษฎีการตอบข้อคำถาม. สถาบันวิจัยพฤติกรรมศาสตร์ มหาวิทยาลัย ศรีนครินทรวิโรฒ ประสานมิตร, 2528.

ธีระศักดิ์ อุรัจนานนท์. การเปรียบเทียบคุณภาพของแบบสอบถามเลือกตอบที่มีจำนวนตัวเลือกไม่เท่ากัน โดยใช้ทฤษฎีการตอบสนองข้อสอบ, วิทยานิพนธ์ ค.ม. กรุงเทพ ฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2531. อัดสำเนา.

เรวดี อินทรสะระ. การเปรียบเทียบความคลาดเคลื่อนของการเทียบมาตราระหว่างรูปแบบอิงทฤษฎีการตอบข้อสอบกับรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ. ปริญญานิพนธ์ กศ.ม. กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2530. อัดสำเนา.

สงบ ลักษณะ. "ความสามารถแท้กับการตอบข้อคำถาม," วารสารวัดผลการศึกษา. 4(1) : 47 - 54, 83 - 93 ; พฤษภาคม - สิงหาคม 2524.

สมจิตร์ ทรัพย์อัปรามไชย. การเปรียบเทียบผลการวิเคราะห์แบบทดสอบวัดเจตคติด้วยคลาสสิคอลโมเดลและราส์ชโมเดล. ปริญญานิพนธ์ กศ.ม. กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2531. อัดสำเนา.

สมพร บุญอิ่ม. การศึกษาความคงที่ของค่าพารามิเตอร์ ความยากในการวิเคราะห์ข้อกระทง วิทยานิพนธ์ ค.ม. กรุงเทพ ฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2529. อัดสำเนา.

สุชาดา กีระนันท์. การอนุมานเชิงสถิติ : ทฤษฎีขั้นต้น. พิมพ์ครั้งที่ 2. กรุงเทพ ฯ : โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย, 2530.

สุพัฒน์ สุกมลสันต์. การวิเคราะห์ข้อทดสอบแนวใหม่ด้วยเครื่องคอมพิวเตอร์. เอกสารลำดับที่ 1 ในโครงการธนาคารข้อสอบโดยใช้เครื่องคอมพิวเตอร์สถาบันภาษาจุฬาลงกรณ์มหาวิทยาลัย, 2530.

สุรัตน์ ขวัญบุญจันทร์. การสร้างแบบทดสอบและตารางเทียบคะแนนในแนวนอนวิชาวิทยาศาสตร์ที่วิเคราะห์ด้วยราส์ชโมเดล. ปริญญานิพนธ์ กศ.ม. กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2531. อัดสำเนา.

สำเริง บุญเรืองรัตน์. "การพัฒนาทฤษฎีลาเท้นเทรดเพื่อการวิเคราะห์ข้อสอบ," วารสารการวัดผลการศึกษา. 4(2) : 1 - 20 ; กันยายน - ธันวาคม 2525.

ส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี กระทรวงศึกษาธิการ, สถาบัน. หนังสือเรียนคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 กรุงเทพ ฯ : โรงพิมพ์คุรุสภา ลาดพร้าว, 2532.

องอาจ นัยพัฒน์. การศึกษาประสิทธิภาพของผลการประมาณค่าพารามิเตอร์ของพาเซียลเครดิตโมเดล ระหว่างวิธีพร็อกซ์ Prox และวูดคอน Ucon ในแบบทดสอบวัดเจตคติแบบ. ปริญญานิพนธ์ กศ.ม. กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2530. อัดสำเนา.

อวยพร วิบูลย์กาญจน์. การเปรียบเทียบผลการวิเคราะห์แบบอุปมาอุปไมยด้วยคลาสสิคคอลโมเดลกับราส์ชโมเดล. วิทยานิพนธ์ ค.ม. กรุงเทพ ฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2526. อัดสำเนา.

Anastasi, A. Psychological testing. (Fourth edition). New York : Macmillan.

Andersen, E.B. "ptiotic properties of conditional maximum likelihood estimates," Journal of the Royal Statistical Society. 32 (1): 283 - 301k ; 1976.

Baker, R.B. "Advances in item analyses," Review of Educational Research. 47(1) : 151 - 178 ; 1977.

Hejar, J.J. "A Procedure for Investigating the Unidimensionality of Achievement Tests Based on Item Parameter Estimates," Journal of Educational Asylum Measurement. 17(4) : 283 - 269 ; 1980.

Birnbaum, A. "Some Talent Trait Models. In Lord, F.M., and Novick, M.R.," Statistical theories of mental test scores. Massachusetts : Addison - Wesley, 1968.

Bloom, H.S., J.T. Hastings and G.F, Madaus. Handbook on Formative and Summation Evaluation of Student learning. New York : McGraw - Hill, 1971.

Davis, F.R. (Chairman of the Joint Committee) Standards For Educational and Psychological Test. Washington, D.C. : American Psychological Association. 1974.

Dixon, W.J., M.B. Brown. BMDP-79. Los Angeles : University of California Press, 1979.

Dodd, Babara G. "The Effect of Item Selection Procedure and (Thiepsize on the Computerized Adaptive Attitude Measurement Using the Rating Scale Model," Applied Psychological Measurement. 11(4) : 371 - 354 ; December, 1987.

Dixon, W.J., M.B. Brown. BMDP-79. Los Angeles : University of California Press, 1979.

Dodd, Babara G. "The Effect of Item Selection Procedure and (Thiepsize on the Computerized Adaptive Attitude Measurement Using the Rating Scale Model," Applied Psychological Measurement. 13(2) : 129 - 143 ; June, 1989.

Dodd, Babara G. and William R. Koch. "Effects of Variations in Item Step Values on Item and Test Information in the Partial Credit Model," Applied Psychological Measurement. 11 (4) : 371 - 384 ; December, 1987.

Drasgow. F., and C.K. Parsons. "Application Of Unidimentsional Item Response Theory Models to Multidimensional Data," Applied Psychological Measurement. 7(2) : 189 - 199 ; 1983.

Edwards, Allen L. Techinouesodd. of Attitude Scale Construction. New York : Appleton - Century - Crofts. Inc., 1957.

Engellhard, George Jr. "Thorndike, Thustone, and Rasch : A Comparision of Their Method of Scaling, Psychological and Educational Tests," Applied Psychological Measurement. 5 (1) : 21 - 38 ; Winter, 1984.

Fishbien, Martin. Readings in Attitude Theory and Measurement. New York : John Wiley and Sons, 1967.

Fishbien. Martin and I. Ajzen Belief Attitude. Intention and Behavior an Introduction to Theory and Research. Menlo Park, California : Addison - Weley Publishing Company. Inco, 1975.

Green, S.H., R.W. Lissitz and S.A.Mulaik. "Limitations of Coefficient Alpha as and Index of Test Unidimentsionality," Educational and Psychological Measurement. 37 : 827 - 838 ; 1977.

Guilford, J.P. Psychometric Methods. New York : McGraw Hill Book, 1954.

Gulliksen, H. Theory of Mental Tests. New York : Wiley, 1950.

Hambleton, R.K., and H. Swaminathan. Item Response Theory : Principles and Applications. Boston : Kluwer - Nijhoff, 1985.

Hambleton, R.K., R.J.Rovinelli. "Assessing the Dimensionality of a Set of Testtiontems," Applied Psychological Measurement. 10(3) : 287 - 302 ; 1986.

Henrysson, S. Gathering, Analysing, and Using Data on Test Items. In Thorndike, R.L. (Ed.) Educational Measurement. Washington, D.C. : American Council on Education, 1976.

Hinkle, D.E., W. Wiersma and S.G.Jurs. Applied Statistions of the Behavioral Sciences. Hoston : Houghton Mifflin, 1979.

Jensema, A Simple Technique for Estimating Leant Trait Mental Test Parameters. Educational and Psychological Measurement. 36 : 705 - 715 ; 1976.

Keats, J.A. Personal Communication. 1989.

Kingston, N. Assessing the dimensionality of the CMAT Verbal and Quantitative Measures Using Full Information Factor Analysis. Research Report 86 - 19. Princeton, C. Educational Testing Service. 1986.

Kline, P. A handbook of test construction : Introduction to Psychometric design. London : Methuen, 1986.

Lord, F.M. "An Analysis of the Verbal Scholastic Aptitude Test Using Birnbaum's Three Parameter Logistic Model," Educational and Psychological Measurement. 28 : 989 – 1020; 1968.

Lord, F.M. "Estemation of la NJ : Ability and Item Parameters When There are Omitted Responses," Psychometrika. 39(2) : 247 - 264, 1974.

Lord, F.M. Applications of Item Response Theory to Practical Testing Problems. New Jersey : Lawrence Erlbaum Associates, 1980.

Lord, F.M., & Novick, M.R. Statistical Theories of Mental Test Scores. Massachusetts : Addison Wesley, 1968.

Masters, G. "tent COT : Analyzing Classroom Tests with the Rasch Model," Educational and Psychological Measurement. 44 : 145 - 150 ; 1984.

McDonald, R.P. "The Dimensionality of Tests and Items," British Journal of Mathematical and Statistical Psychology. 34 : 100 - 117 ; 1981.

Nie, N.11. SPSS User's Guide. Illinois : SPSS, 1983.

Rasch, G. PROBABILISTIC models for Some Intelligence and Attainment tests. Copenhagen : Denmarks. Diedoggiogiske Institute, 1960.

Reckase, M.D. "Unifactor Latent Trait Models Applied to Multifactc Tests : Results and Implications," Journal of Educational Statistics. 4(3) : 207 - 230 ; 1979.

Samejima, F. "A comment of Birnbaum's three parameter logistic model in the latent trait theory," Psychometrika. 38 : 221 - 233, 1973.

SAS user's guide : statistics version 5. Paedition. North Carolina : SAS institute, 1985.

Smith, J.K. "On the Examination of Test Unidimentsionality," Educational and Psychological Measurement. 40 : 885 - 889 ; 1980.

Wainer, H. "The Future of Item Analysis," Journal of Educational Measurement. 26(2) : 191 - 208 ; 1989.

Wingersky, M.S. M.A. Barton and F.M. Lord. Logist user's Princetion, N.J. : Educational Testing Service. 1982.

Wood R.l., M.S. Wingersky and & F.M.Lord. Logist : A Computer Program for Estimating Examine Ability and Item Characteristic Curve Parameters. Research Memorandum (modified 1/78). New Jersey: Educational Testing Service, Princeton, 1976.

Wright, B.D. "Solving Measurement Problems with the Resch Model," Lougis of Educational Measurement. 14(2) : 97 - 116 ; 1977.

Wright, B.S., & Bell, S.K. Verifying the Unconditional Estimation Procedure for Rasch Item Analysis With Simulated Data. (RM26) Statistical Laboratory, Department of Education, University of Chicago, 1977.

Wright, B.D., & Douglas, G.A. "Conditional Versus Unconditional Procedures for Sample - Final Item Analysis," Educational and Psychological Measurement. 37 : 573 - 586 ; 1977.

Wemorandum 23 C, Chicago. Statistical Laboratory, Department of Education, University of Chicago.

Wright, B.D. & Panchapakesan N. A procedure for Sample Free Item Analysis. Educational and Psychological Measurement, 29 : 23 - 48 ; 1969.

Wright, B.D., & Stone, M.H. Best test design : Rasch Measurement. Chicago : MESA Press, 1979.

Yen, W.M. "Taining Maximum Likelihood Trait Estimates From Number-Correct Scores for The Three-Parameter Logistic Model," Journal of Educational Measurement. 21(2) : 93 - 111 ; 1984.

ภาคผนวก ก

แบบทดสอบเลือกตอบ 5 ตัวเลือกวิชาคณิตศาสตร์
ชั้นมัธยมศึกษาปีที่ 2 จำนวน 60 ข้อ

ข้อที่มีเครื่องหมาย * คือข้อที่ได้รับการคัดเลือก

คำชี้แจงทั่วไป

1. แบบทดสอบฉบับนี้ประกอบด้วยข้อสอบเลือกตามชนิด 5 ตัวเลือก จำนวน 60 ข้อ
2. ข้อสอบแต่ละข้อมีคำตอบถูกเพียงตัวเลือกเดียว
3. เวลาที่ใช้ในการสอบ 90 นาที

คำชี้แจงในการตอบข้อสอบ

ให้นักเรียน X ทับตัวเลือก ก ข ค ง และ จ ที่ถูกที่สุดเพียงข้อเดียว

ตัวอย่าง

O จากสมการ 2 x = 10 อยากทราบว่า x มีค่าเท่าไร

ก. 1

ข. 2

ค. 3

ง. 4

X 5

1. ผลบวกของเลข 2 จำนวน มีค่าเป็น 41 โดยที่จำนวนหนึ่งมีค่ามากกว่าอีกจำนวนหนึ่งอยู่ 5 ถ้าให้เลขจำนวนน้อยเป็น x สมการที่ให้หาเลข 2 จำนวนนี้คือข้อใด

   ก. $(x-5)+x=41$

   ข. $(x+5)+(x-5)=41$

   * ค. $(x+5)+x=41$

   ง. $(2x-41)-5=0$

   จ. $2(x-5)=41$

2. สองเท่าของเลขจำนวนหนึ่งมากกว่า 21 อยู่ 21 เลขจำนวนนั้นคืออะไร สมการที่ใช้หาเลขจำนวนนี้คือข้อใด

   ก. $2(21-x)=21$

   * ข. $2(x-21)=21$

   ค. $2(21)-x=21$

   ง. $2x-21=21$

   จ. $2x-2(21)=21$

3. จำนวน 3 จำนวนเรียงกัน (1,2,3) จำนวนน้อยรวมกับจำนวนมากมีค่า 44 ถ้าให้จำนวนที่ 2 เป็น x สมการที่ใช้หาจำนวนทั้ง 3 คือข้อใด

   ก. $x+(x-1)=44$

   ข. $x+(x-2)=44$

   ง. $(x-1)+(x-2)=44$

   * ค. $(x-1)+(x+1)=44$

   จ. $(x-2)+(x+2)=44$

4. แดงมีเงิน 3/7 ของดำ ถ้าแดงมีเงิน 210 บาท ดำจะมีเงินเท่าไร เขียนสมการตรงกับข้อใด

   ก. $\frac{3x}{7}-210=0$

   ข. $\frac{4x}{7}-210=0$

   ค. $210-\frac{3x}{7}=0$

   ง. $210-\frac{4x}{7}=0$

   จ. $\frac{3x}{7}+210=0$

5. ถ้าเอา 4 คูณเลขจำนวนหนึ่งแล้วเอา 10 มาลบผลคูณจำนวนนั้นจะได้ผลลัพธ์ 56 เลขจำนวนนั้นคือจำนวนใด เขียนเป็นสมการได้ตรงกับข้อใด

   ก. $4x+10=56$

   ข. $4x-10=56$

   ค. $4(x+10)=56$

   ง. $4(x-10)=56$

   จ. $(4x-10)+56=0$

6. เจ็ดเท่าของเงินที่ดำมีรวมกับเงินที่แม่ให้ 15 บาทเป็นเงิน 71 บาท จงหาว่าเดิมดำมีเงินเท่าไร เขียนเป็นสมการได้ตรงกับข้อใด

   ก. $7x-15=71$

   * ข. $7x-71=15$

   ค. $7x+15=71$

   ง. $7(x+15)=71$

   จ. $7(x-15)=71$

7. ถ้า 11 เท่าของผลบวกของเลขจำนวนหนึ่งกับ 7 มีค่าเท่ากับ 132 เขียนเป็นสมการได้ตรงกับข้อใด

ก. $11x + 7 = 132$

ข. $11x - 7 = 132$

ค. $11x + 132 = 7$

ง. $11(x-7) = 132$

จ. $11(x+7) = 132$

8. สนามเทนนิสมีด้านยาว ยาวกว่า 2 เท่าของด้านกว้าง 4 เมตร ถ้าสนามเทนนิสมีเส้นรอบรูปยาว 74 เมตร ถ้าด้านกว้างยาว a เมตร เขียนเป็นสมการได้ตรงกับข้อใด

ก. $2(a+2a-4) = 74$

ข. $2(2a-4) = 74$

* ค. $2[(a+4) + (a+4)] = 74$

ง. $2(2a+4+a) = 74$

จ. $2[2(a+4) + a] = 74$

9. มีเชือกยาว 90 เมตร ตัดแบ่งเป็น 2 ท่อน ถ้าท่อนหนึ่งยาวกว่าอีกท่อนหนึ่งอยู่ 18 เมตร ถ้าเส้นยาว ยาว x เมตร เขียนสมการได้ตรงกับข้อใด

ก. $2x + 18 = 90$

ข. $2(x+18) = 90$

ค. $2(x-18) = 90$

ง. $2x - 18 = 90$

จ. $2(90-x) = 18$

10. จำนวนคู่บวก 3 จำนวนเรียงกัน มีผลบวก 126 จงหาเลขจำนวนมาก ถ้าให้จำนวนมากเป็น x เขียนสมการได้ตรงกับข้อใด

ก. $x + (x+1) + (x+2) = 126$

* ข. $x + (x-1) + (x-2) = 126$

ค. $(x-2) + x + (x+2) = 126$

ง. $(x-4) + (x-2) + x = 126$

จ. $x + (x+2) + (x+4) = 126$

11. สี่เหลี่ยมด้านขนานมีด้านประกอบมุมด้านหนึ่งยาวเป็น 3 เท่าของด้านประกอบมุมด้านหนึ่ง ถ้าสี่เหลี่ยมด้านขนานรูปนี้มีเส้นรอบรูปยาว 64 นิ้ว จงหาความยาวของด้านยาว เขียนเป็นสมการได้ตรงกับข้อใด

* ก. $(3x)(x) = 64$

ข. $3x + x = 64$

ค. $0x + x = 64$

ง. $6x + 2x = 64$

จ. $(3x)(2x) = 64$

12. พี่กับน้องมีเงินรวมกัน 120 บาท คนพี่มีมากกว่า 2 เท่าของคนน้องอยู่ 30 บาท ถ้าพี่มีเงิน x บาท เขียนเป็นสมการได้ตรงกับข้อใด

ก. $(2x+30) + x = 120$

ข. $(2x-30) + x = 120$

* ค. $(2x+30) + 2x = 120$

ง. $(2x-30) + 2x = 120$

จ. $(x+30) + 2x = 120$

13. ปัจจุบันพ่อของนิดอายุ 45 ปี เมื่อ 3 ปีที่แล้วพ่อของนิดมีอายุเป็น 3 เท่าของอายุนิด ปัจจุบันนิดอายุเท่าไร เขียนเป็นสมการได้กับข้อใด

ก. $3x + 3 = 45$

ข. $3x + 3 = 48$

ค. $3(x+3) = 48$

ง. $3(x-3) = 48$

จ. $3(x-3) = 42$

14. เลขคี่ 3 จำนวนเรียงกัน ผลบวกจำนวนทั้ง 3 เป็น 69 ถ้าจำนวนที่ 2 เป็น x เขียนสมการได้ตรงกับข้อใด

ก. $x + (x+1) + (x+2) = 69$

* ข. $x + (x-1) + (x-2) = 69$

ค. $(x-1) + x + (x-2) = 69$

ง. $(x-2) + x + (x+2) = 69$

จ. $(x-4) + x + (x+14) = 69$

15. 14 น้อยกว่า 5 เท่าของของเลขจำนวนหนึ่ง เท่ากับเอา 10 บวกกับ 3 เท่าของเลขจำนวนนั้น เขียนเป็นสมการได้ดังข้อใด

ก. $5x - 14 = 10 + 3x$

ข. $14 - 5x = 10 + 3x$

* ค. $14 - 5x = 3x - 10$

ง. $5x - 14 = 3x - 10$

จ. $5(x-14) = 10 + 3x$

16. นาย A มีเงิน x บาท ให้นาย B ไป y บาท ปรากฏว่านาย B มีเงินมากกว่า A อยู่ 20 อยากทราบว่า เดิม B มีเงินอยู่เท่าไร

ก. $x + 2y + 20$

ข. $x + 2y - 20$

* ค. $x - 2y - 20$

ง. $x - 2y + 20$

จ. $y - 2x + 20$

17. 2 เท่าของผลบวกของเลขจำนวนหนึ่งกับ 0.5 มีค่ามากกว่าเลขจำนวนนั้นอยู่ 3.5 เขียนเป็นสมการได้ดังข้อใด

ก. $2x + 0.5 = x + 3.5$

* ข. $2x - 0.5 = x + 3.5$

ค. $2(x+0.5) - x = 3.5$

ง. $2(x+0.5) + x = 3.5$

จ. $2(x-0.5) - x = 3.5$

18 ผลต่างของ 3 เท่าของเลขจำนวนหนึ่งกับ 32 มีค่าเท่ากับผลบวกของ 2 เท่าของจำนวนนั้นกับ 5 จงหาเลขจำนวนนั้น เขียนเป็นสมการได้ตรงกับข้อใด

ก. $3(x-32) = 2x + 5$

ข. $3(x-32) = 2(x+5)$

* ค. $3x - 32 = 2x + 5$

ง. $3x - 32 = 2(x+5)$

จ. $3x - 32 = 5(x+2)$

19. แบ่งผลไม้ 198 ผลออกเป็น 2 กองให้ 3/4 ของกองที่หนึ่งมากกว่า 5/7 ของกองที่ 2 อยู่ 5 จงหาผลไม้กองแรก

ก. 89 ผล

* ข. 96 ผล

ค. 98 ผล

ง. 100 ผล

จ. 108 ผล

20. จิ๋วซื้อเสื้อมา 2 ตัว ตัวหนึ่งราคาแพงกว่าอีกตัวหนึ่ง 250 บาท เสื้อ 2 ตัวราคารวมกัน 990 บาท จงหาราคาเสื้อตัวที่ถูกที่สุด

ก. 250 บาท

ข. 370 บาท

* ค. 500 บาท

ง. 570 บาท

จ. 610 บาท

21. ซื้อผ้าเช็ดหน้ามา 20 ผืน ราคา 180 บาท ผ้าเช็ดหน้ามี 2 ชนิด ชนิดหนึ่งราคาผืนละ 15 บาท ชนิดที่ 2 ราคาผืนละ 5 บาท ซื้อผ้าเช็ดหน้าชนิดราคา 5 บาท มากี่ผืน

ก. 6 ผืน

* ข. 8 ผืน

ค. 10 ผืน

ง. 12 ผืน

จ. 14 ผืน

22. A พูดกับ B ว่า A มีเงินมากกว่า B อยู่ 45 บาท แต่ถ้าพ่อให้ A อีก 15 บาท A จะมีเงินเป็น 2 เท่าของ B จงหาเงินของ A และ B รวมกันมีค่าเท่าไร

ก. 45 บาท

ข. 60 บาท

* ค. 85 บาท

ง. 105 บาท

จ. 165 บาท

23. เมื่อนับจำนวนขาหมูและจำนวนหัวไก่ในฟาร์ม พบว่าจำนวนขาหมูมากกว่าจำนวนหัวไก่อยู่ 20 ขา ในฟาร์มนี้เลี้ยงไก่กี่ตัว ถ้าจำนวนหมูและไก่ รวมกันมี 20 ตัว

ก. 6 ตัว

* ข. 8 ตัว

ค. 10 ตัว

ง. 12 ตัว

จ. 15 ตัว

24. ภาพติดฝาผนังรูปสี่เหลี่ยมจัตุรัส ภาพหนึ่งทำกรอบด้านละ 1 นิ้ว พื้นที่เพิ่มขึ้นอีก 52 ตารางนิ้ว จงหาความยาวของด้านของภาพ

ก. 14 นิ้ว

ข. 13 นิ้ว

ค. 12 นิ้ว

ง. 11 นิ้ว

จ. 10 นิ้ว

25. ที่ดินรูปสามเหลี่ยมหน้าจั่ว มีฐานยาวเป็นครึ่งหนึ่งของด้านที่ประกอบมุมยอด ถ้าค่า

จ้างล้อมรั้วที่ดินเป็นเงิน 6,250 บาท โดยคิดค่าจ้างเมตรละ 125 บาท จงหาค่าด้านประกอบมุมยอด ยาวด้านละกี่เมตร

ก. 14 เมตร

ข. 16 เมตร

* ค. 18 เมตร

ง. 20 เมตร

จ. 22 เมตร

26. ปัจจุบันบิดาอายุ 55 ปี เมื่อ 7 ปีที่แล้ว บุตรมีอายุ 5/12 เท่าของบิดาปัจจุบันบุตรมีอายุเท่าไร

ก. 20 ปี

ข. 23 ปี

ค. 25 ปี

ง. 27 ปี

จ. 29 ปี

27. ถ้าแจกเงินเด็กคนละ 3 บาท จะเหลือเงินอยู่ 2 บาท แต่ถ้าแจกคนละ 4 บาทจะขาดเงินอยู่ 1 บาท จงหาว่าจำนวนเงินมีอยู่กี่บาท

ก. 11 บาท

ข. 13 บาท

* ค. 15 บาท

ง. 17 บาท

จ. 19 บาท

28. ในการเล่นเกมส์ปาเป้า โดยมีกติกาว่าผู้เล่นปาเป้าแต่ละครั้งถ้าปาถูกจะได้รับเงิน 10 บาท ถ้าปาผิดจะเสียเงิน 4 บาท ถ้าชายคนหนึ่งปาไป 20 ครั้ง เขาต้องจ่ายเงิน 10 บาท อยากทราบว่าชายคนนี้ปาถูกเป้ากี่ครั้ง

ก. 15 ครั้ง

ข. 13 ครั้ง

* ค. 10 ครั้ง

ง. 7 ครั้ง

จ. 5 ครั้ง

29. คำรง สอบวิชาคณิตศาสตร์ได้ 15 คะแนน สังคมศึกษาได้ 18 คะแนน ถ้า 5 เท่าของคะแนนวิทยาศาสตร์รวมกับคะแนนคณิตศาสตร์และสังคมศึกษา จะได้ 83 คะแนน จงหาคะแนนวิทยาศาสตร์

ก. 18 คะแนน

ข. 16 คะแนน

ค. 14 คะแนน

ง. 12 คะแนน

จ. 10 คะแนน

33. ผลต่างระหว่างจำนวนเงินของเอและบีมากกว่าผลต่างของเงินบีและซีอยู่ 5 บาท ถ้าบีมี 65 บาท และซีมี 55 บาท จงหาเงินเอ

ก. 100 บาท

ข. 90 บาท

* ค. 80 บาท

ง. 75 บาท

จ. 70 บาท

34. ด้าน ๆ หนึ่งของรูปสามเหลี่ยมด้านเท่ายาวเท่ากับด้าน ๆ หนึ่งของรูปสี่เหลี่ยมจัตุรัส ถ้าผลบวกของด้านทุกด้านของรูปสามเหลี่ยมและสี่เหลี่ยมยาวเท่ากับ 168 นิ้ว จงหาพื้นที่ของสี่เหลี่ยมจัตุรัส

* ก. 226 ตารางนิ้ว

ข. 326 ตารางนิ้ว

ค. 446 ตารางนิ้ว

ง. 576 ตารางนิ้ว

จ. 676 ตารางนิ้ว

35. ปีที่แล้ว เอกมีอายุเป็น 4 เท่าของแวว แต่ปัจจุบันเอกมีอายุเป็น 2 เท่าของแวว จงหาอายุปัจจุบันของแวว

ก. 12 เดือน

* ข. 16 เดือน

ค. 18 เดือน

ง. 20 เดือน

จ. 24 เดือน

36. มีซาลาเปา 80 ลูก ปรากฏว่าอาพลกินหมดภายใน 4 วัน แต่ละวันเขากินมากกว่าวันที่แล้วอยู่ 4 ลูก อยากทราบว่าเขากินในวันที่ 3 กี่ลูก

ก. 14 ลูก

ข. 18 ลูก

* ค. 22 ลูก

ง. 26 ลูก

จ. 30 ลูก

37. เลข 2 จำนวนรวมกันได้ 74 ถ้าเลขจำนวนมากถูกหารด้วยเลขจำนวนน้อยผลลัพธ์เป็น 3 เศษ 10 จงหาเลขจำนวนมาก

ก. 64

ข. 60

* ค. 58

ง. 56

จ. 54

38. ในการประชุมอาจารย์ของโรงเรียนแห่งหนึ่ง ถ้าอาจารย์หญิง 30 คนออกจากห้องประชุม จะมีอาจารย์ชายเป็น 2 เท่าของอาจารย์หญิง เมื่ออาจารย์หญิงออกไปแล้วอาจารย์ชายอีก 90 คนก็ออกไป ปรากฏว่ามีอาจารย์หญิงเหลืออยู่เป็น 5 เท่าของอาจารย์ชาย ที่เหลืออยู่ในห้องประชุม อยากทราบว่ามีอาจารย์หญิงกี่คน

ก. 100 คน

ข. 90 คน

ค. 80 คน

ง. 70 คน

จ. 60 คน

39. พ่อมีอายุ 41 ปี ลูกมีอายุ 9 ปี อีกกี่ปีพ่อจะมีอายุเป็น 3 เท่าของลูก

ก. 5 ปี

ข. 6 ปี

ค. 7 ปี

ง. 8 ปี

จ. 9 ปี

40. ในการเดินทางครั้งหนึ่ง ใช้เงินเป็นค่าโดยสารรถไฟมากกว่า ค่าที่พัก 140 บาท แต่ค่าที่พักน้อยกว่าค่าใช้จ่ายอย่างอื่น 100 บาท ถ้าใช้เงินทั้งหมด 1,920 บาท อยากทราบว่าจ่ายค่ารถไฟเท่าไร

ก. 560 บาท

* ข. 600 บาท

ค. 640 บาท

ง. 700 บาท

จ. 740 บาท

41. ในการเลือกตั้งหัวหน้าห้องปรากฏว่าฝ่ายชนะได้เสียงมากกว่า 9 เสียง ถ้าหากเศษหนึ่งส่วนหกของฝ่ายชนะกลับออกเสียงให้อีกฝ่ายหนึ่ง แล้วฝ่ายที่เคยแพ้จะกลับชนะ และมีเสียงมากกว่า 3 เสียง เสียงทั้งหมดมีเท่าไร

ก. 27 เสียง

* ข. 36 เสียง

ค. 45 เสียง

ง. 54 เสียง

จ. 63 เสียง

42. ปูมีเงินมากกว่ากุ้ง 50 บาท กุ้งมีเงินน้อยกว่าปลา 30 บาท คนทั้ง 3 มีเงินรวมกัน 200 บาทพอดี กุ้งมีเงินกี่บาท

ก. 90 บาท

ข. 70 บาท

* ค. 50 บาท

ง. 40 บาท

จ. 20 บาท

43. การเล่นเกม ลิงจับหลักสมมุติว่าคนเล่นเป็นลิง ถ้าลิงจับหลักละ 1 ตัวจะเหลือลิง 3 ตัวที่ไม่มีหลักจับ แต่ถ้าลิงจับหลักละ 2 ตัวจะเหลือหลัก 1 หลักที่ไม่มีลิงจับ อยากทราบว่ามีคนเล่นเป็นลิงกี่คน

ก. 4 คน

ข. 6 คน

ค. 8 คน

ง. 10 คน

จ. 12 คน

44. เลข 3 จำนวนเรียงกัน ถ้า 3 เท่าของจำนวนน้อยที่สุด รวมกับอีก 2 จำนวนมีค่าเท่ากับ 103 ดังนั้นผลบวกของจำนวนทั้ง 3 คือข้อใด

ก. 60

ข. 63

ค. 66

ง. 69

จ. 72

45. ปัจจุบันชายคนหนึ่งมีอายุเป็น 3 เท่าของลูกชายอีก 5 ปีข้างหน้า ทั้ง 2 มีอายุรวมกัน 86 ปี ปัจจุบันผลต่างของอายุคนทั้ง 2 เท่ากับกี่ปี

ก. 28 ปี

ข. 32 ปี

ค. 34 ปี

ง. 36 ปี

จ. 38 ปี

46. นิพนธ์ปั่นจักรยานไปโรงเรียน ด้วยความเร็ว 3 กิโลเมตรต่อชั่วโมง ปรากฏว่าไปสาย 5 นาที แต่ถ้านิพนธ์เดินทางด้วยความเร็วเพิ่มขึ้นอีก 1 กิโลเมตรต่อชั่วโมง จะถึงโรงเรียนก่อนโรงเรียนเข้า 5 นาที จงหาระยะทางไปกลับจากบ้านถึงโรงเรียน

ก. 2 กิโลเมตร

ข. 4 กิโลเมตร

* ค. 6 กิโลเมตร

ง. 8 กิโลเมตร

จ. 10 กิโลเมตร

47. 2 ใน 3 ของเงินจำนวนหนึ่ง เมื่อหักออก 1 ใน 4 ของจำนวนดังกล่าวจะเหลือเงิน 120 บาท จงหาว่าเงินจำนวนนั้นมีกี่บาท

ก. 240 บาท

ข. 248 บาท

* ค. 268 บาท

ง. 278 บาท

จ. 288 บาท

48. เมื่อนำ 5 คูณกับเลขจำนวนหนึ่ง และลบออกด้วย 2/3 คูณกับเลขจำนวนเดิมมีค่าเท่ากับ 26 จำนวนนั้นคือข้อใด

ก. 6

ข. 8

* ค. 10

ง. 12

จ. 14

49. ถ้ารถวิ่งจากกรุงเทพฯ ไปสระบุรี ด้วยความเร็ว 90 กิโลเมตรต่อชั่วโมง ปรากฏว่าจะถึงก่อนโดยใช้ความเร็ว 80 กิโลเมตรต่อชั่วโมง เป็นเวลา 10 นาที จงหาระยะทางกรุงเทพ-สระบุรี

ก. 80 กิโลเมตร

* ข. 90 กิโลเมตร

ค. 100 กิโลเมตร

ง. 110 กิโลเมตร

จ. 120 กิโลเมตร

50. อีก 10 ปีข้างหน้า สมศักดิ์จะมีอายุเป็น 3 เท่าของบุตร ถ้าปัจจุบันบุตรมีอายุ 18 ปี ปัจจุบันสมศักดิ์จะมีอายุเท่าไร

ก. 45 ปี

ข. 50 ปี

ค. 60 ปี

ง. 65 ปี

จ. 74 ปี

51. จำนวน ๆ หนึ่งมีค่าน้อยกว่าอีกจำนวนหนึ่งอยู่ 5 ถ้าเลขทั้ง 2 รวมกันได้ 57 จำนวนน้อยคือเลขจำนวนใด

ก. 24

ข. 26

ค. 28

ง. 30

จ. 31

52. สนามรูปสี่เหลี่ยมผืนผ้าแห่งหนึ่งมีความยาวน้อยกว่า 2 เท่าของด้านกว้างอยู่ 7 เมตร ถ้าสี่เหลี่ยมรูปนี้มีความยาวเส้นรอบรูปยาว 82 เมตร จงหาพื้นที่ของรูปสี่เหลี่ยม

ก. 400 ตารางเมตร

ข. 420 ตารางเมตร

* ค. 440 ตารางเมตร

ง. 460 ตารางเมตร

จ. 480 ตารางเมตร

53. อมรมีเงินน้อยกว่า 2 ใน 5 ของเงินของวิชัยอยู่ 20 บาท ถ้าอมรมีเงิน 120 บาท วิชัยจะมีเงินเท่าไร

ก. 300 บาท

ข. 320 บาท

ค. 330 บาท

ง. 350 บาท

จ. 370 บาท

54. อีก 5 ปีข้างหน้าบังอรจะมีอายุเป็นครึ่งหนึ่งของอุดม ถ้าปัจจุบันอุดมมีอายุ 15 ปี ปัจจุบันบังอรจะมีอายุเท่าไร

ก. 2.5 ปี

ข. 3 ปี

ค. 5 ปี

ง. 7.5 ปี

จ. 10 ปี

55. อัมพรมีเงินเหรียญบาทและเหรียญ 5 บาทรวมกันอยู่ 20 เหรียญ คิดเป็นเงิน 44 บาท อยากทราบว่าอัมพรมีเหรียญ 5 บาทกี่เหรียญ

ก. 4 เหรียญ

* ข. 5 เหรียญ

ค. 6 เหรียญ

ง. 7 เหรียญ

จ. 8 เหรียญ

56. ดำเลี้ยงไก่และหมูไว้รวมกัน 44 ตัว นับขาหมูรวมกับหัวไก่ได้ 104 อยากทราบว่าขาไก่มีทั้งหมดกี่ขา

ก. 24 ขา

ข. 36 ขา

* ค. 40 ขา

ง. 48 ขา

จ. 58 ขา

57. ในการเดินทางจากกรุงเทพฯ ไปบางแสนด้วยความเร็ว 100 กิโลเมตรต่อชั่วโมง จะถึงเร็วกว่าเดินด้วยความเร็ว 80 กิโลเมตรต่อชั่วโมง อยู่ 20 นาที จงหาระยะทางจากกรุงเทพฯ ถึงบางแสน

ก. 80 กิโลเมตร

ข. 90 กิโลเมตร

* ค. 100 กิโลเมตร

ง. 340/3 กิโลเมตร

จ. 400/3 กิโลเมตร

58. ต่อและต๋อมวิ่งแข่งกัน โดยต่อให้ต๋อมวิ่งไปก่อน 5 นาที และต่อวิ่งได้ 50 เมตรต่อนาที ต๋อมวิ่งได้ 45 เมตรต่อนาที ต่อจะวิ่งทันต๋อมในเวลากี่นาที

ก. 30 นาที

ข. 35 นาที

* ค. 40 นาที

ง. 45 นาที

จ. 50 นาที

59. กมลเลี้ยงนกไว้จำนวนหนึ่ง ต่อมา 2/5 ของนกตายไปและกมลส่งนกอีก 1/7 ของที่เหลือจากตายให้น้องชายปรากฏว่ากมลเหลือนกอยู่ 18 ตัว เดิมเขามีนกกี่ตัว

ก. 30 ตัว

ข. 33 ตัว

* ค. 35 ตัว

ง. 39 ตัว

จ. 42 ตัว

60. ผลบวกของเลขจำนวนหนึ่งกับ 11 มากกว่า 40 อยู่ 50 เลขจำนวนนั้นมีค่าเท่าใด

ก. 59

ข. 69

ค. 79

ง. 89

จ. 99

ภาคผนวก ข

แสดงค่าความยาก ค่าอำนาจจำแนกของแบบทดสอบเลือกตอบ 5 ตัวเลือก
วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 จำนวน 60 ข้อ
ข้อที่มีเครื่องหมาย * คือข้อที่ได้รับการคัดเลือก

ข้อ	ตัวเลือก	P	r
1.	ก.	0.05	.47
	ข.	0.11	.41
*	ค.	0.75	.52
	ง.	0.70	.21
	จ.	0.09	.02
2.	ก.	.09	.02
	ข.	.71	.27
*	ค.	.12	.04
	ง.	.12	.16
	จ.	.07	.06
3.	ก.	.11	.23
	ข.	.05	.26
*	ค.	.09	.30
	ง.	.71	.43
	จ.	.06	.32
4.	ก.	.95	.29
	ข.	.05	.00
	ค.	.05	.00
	ง.	.05	.00
	จ.	.05	.00
5.	ก.	.12	.14
	ข.	.39	.26
*	ค.	.18	.02
	ง.	.16	.02
	จ.	.16	.02

ข้อ	ตัวเลือก	P	r
6.	ก.	.08	.11
	ข.	.12	.16
*	ค.	.54	.22
	ง.	.15	.05
	จ.	.13	.16
7.	ก.	.05	.08
	ข.	.05	.08
	ค.	.05	.26
	ง.	.05	.26
	จ.	.82	.00
8.	ก.	.05	.30
	ข.	.07	.21
*	ค.	.11	.23
	ง.	.75	.32
	จ.	.05	.20
9.	ก.	.95	.26
	ข.	.05	.16
	ค.	.05	.16
	ง.	.05	.16
	จ.	.05	.16
10.	ก.	.11	.23
	ข.	.08	.02
*	ค.	.09	.30
	ง.	.75	.41
	จ.	.05	.26

ข้อ	ตัวเลือก	P	r
11.	ก.	.12	.16
	ข.	.09	.15
*	ค.	.09	.10
	ง.	.71	.27
	จ.	.08	.05
12.	ก.	.75	.37
	ข.	.08	.05
*	ค.	.12	.16
	ง.	.09	.12
	จ.	.09	.10
13.	ก.	.09	.10
	ข.	.09	.10
	ค.	.05	.16
	ง.	.05	.16
	จ.	.87	.20
14.	ก.	.05	.30
	ข.	.07	.21
	ค.	.68	.40
*	ง.	.76	.34
	จ.	.08	.05
15.	ก.	.11	.23
	ข.	.05	.26
*	ค.	.71	.41
	ง.	.06	.34
	จ.	.09	.30

ข้อ	ตัวเลือก	P	r
16.	ก.	.05	.08
	ข.	.05	.08
*	ค.	.05	.08
	ง.	.85	.11
	จ.	.05	.08
17.	ก.	.74	.35
	ข.	.05	.26
*	ค.	.06	.14
	ง.	.06	.10
	จ.	.08	.40
18.	ก.	.05	.26
	ข.	.11	.26
*	ค.	.10	.09
	ง.	.68	.34
	จ.	.08	.40
19.	ก.	.06	.09
*	ข.	.12	.16
	ค.	.54	.24
	ง.	.15	.04
	จ.	.12	.14
20	ก.	.12	.14
	ข.	.14	.07
*	ค.	.55	.25
	ง.	.10	.09
	จ.	.08	.13

ข้อ	ตัวเลือก	P	r
21.	ก.	.06	.14
	ข.	.09	.15
*	ค.	.09	.30
	ง.	.71	.27
	จ.	.06	.14
22.	ก.	.07	.09
	ข.	.54	.22
*	ค.	.11	.24
	ง.	.13	.16
	จ.	.15	.04
23.	ก.	.12	.14
	ข.	.14	.07
*	ค.	.12	.18
	ง.	.55	.25
	จ.	.09	.12
24.	ก.	.05	.08
	ข.	.06	.08
*	ค.	.05	.08
	ง.	.05	.08
	จ.	.80	.26
25.	ก.	.15	.11
	ข.	.20	.12
*	ค.	.22	.15
	ง.	.38	.35
	จ.	.11	.09

ข้อ	ตัวเลือก	P	r
26.	ก.	.12	.14
	ข.	.14	.07
*	ค.	.55	.23
	ง.	.12	.18
	จ.	.09	.12
27.	ก.	.05	.28
	ข.	.08	.23
*	ค.	.07	.06
	ง.	.78	.27
	จ.	.05	.26
28	ก.	.05	.26
	ข.	.05	.26
	ค.	.05	.26
	ง.	.95	.26
	จ.	.05	26
29.	ก.	.39	.19
	ข.	.18	.06
*	ค.	.18	.03
	ง.	.18	.03
	จ.	.12	.14
30.	ก.	.18	.03
	ข.	.20	.12
*	ค.	.22	.15
	ง.	.11	.09
	จ.	.36	.32

ข้อ	ตัวเลือก	P	r
31.	ก.	.05	.00
	ข.	.05	.39
	ค.	.05	.39
	ง.	.05	.39
	จ.	.95	.19
32.	ก.	.09	.03
	ข.	.09	.30
*	ค.	.75	.33
	ง.	.06	.05
	จ.	.05	.26
33.	ก.	.06	.07
	ข.	.06	.14
*	ค.	.77	.29
	ง.	.10	.19
	จ.	.09	.03
34.	ก.	.05	.00
	ข.	.05	.06
*	ค.	.06	.00
	ง.	.09	.10
	จ.	.68	.47
35.	ก.	.13	.17
	ข.	.17	.18
*	ค.	.35	.29
	ง.	.17	.03
	จ.	.22	.15

ข้อ	ตัวเลือก	P	r
36.	ก.	.05	.08
	ข.	.05	.08
	ค.	.81	.30
	ง.	.06	.00
	จ.	.05	.08
37.	ก.	.05	.26
	ข.	.05	.11
*	ค.	.10	.19
	ง.	.77	.30
	จ.	.05	.29
38.	ก.	.08	.02
	ข.	.12	.02
*	ค.	.63	.09
	ง.	.12	.16
	จ.	.07	.06
39.	ก.	.13	.16
	ข.	.14	.07
*	ค.	.57	.10
	ง.	.12	.02
	จ.	.17	.03
40.	ก.	.07	0.9
	ข.	.12	.16
*	ค.	.63	.42
	ง.	.08	.42
	จ.	.10	.31

ข้อ	ตัวเลือก	P	r
41.	ก.	.12	.14
	ข.	.17	.03
*	ค.	.14	.04
	ง.	.17	.02
	จ.	.39	.20
42.	ก.	.05	.26
	ข.	.05	.26
	ค.	.05	.26
	ง.	.95	.28
	จ.	.05	.26
43.	ก.	.13	.04
	ข.	.20	.16
*	ค.	.40	.25
	ง.	.19	.07
	จ.	.11	.09
44.	ก.	.05	.26
	ข.	.95	.16
	ค.	.05	.26
	ง.	.05	.26
	จ.	.05	.26
45.	ก.	.05	.26
	ข.	.05	.26
	ค.	.05	.26
	ง.	.05	.26
	จ.	.05	.43

ข้อ	ตัวเลือก	P	r
46.	ก.	.12	.16
	ข.	.15	.11
*	ค.	.17	.05
	ง.	.11	.29
	จ.	.46	.33
47.	ก.	.05	.26
	ข.		.27
*	ค.	.07	.06
	ง.	.05	.26
	จ.	.09	.30
48.	ก.	.09	.02
	ข.	.12	.16
*	ค.	.13	.17
	ง.	.15	.05
	จ.	.55	.21
49.	ก.	.71	.43
	ข.	.05	.14
*	ค.	.08	.42
	ง.	.09	.30
	จ.	.06	.34
50.	ก.	.05	.08
	ข.	.05	.08
	ค.	.05	.08
	ง.	.05	.08
	จ.	.86	.09

ข้อ	ตัวเลือก	P	r
51.	ก.	.05	.26
	ข.	.05	.26
	ค.	.05	.26
	ง.	.05	.26
	จ.	.95	.26
52.	ก.	.57	.18
	ข.	.12	.16
*	ค.	.15	.05
	ง.	.15	.05
	จ.	.13	.16
53.	ก.	.05	.26
	ข.	.05	.08
	ค.	.06	.00
	ง.	.89	.13
	จ.	.05	.26
54.	ก.	.05	.08
	ข.	.05	.08
	ค.	.05	.43
	ง.	.05	.26
	จ.	.05	.26
55.	ก.	.12	.14
	ข.	.15	.11
*	ค.	.46	.33
	ง.	.13	.20
	จ.	.17	.03

ข้อ	ตัวเลือก	P	r
56.	ก.	.09	.02
	ข.	.11	.23
*	ค.	.14	.09
	ง.	.55	.26
	จ.	.12	.17
57.	ก.	.12	.14
	ข.	.18	.02
	ค.	.16	.02
*	ง.	.16	.02
	จ.	.39	.32
58.	ก.	.08	.11
	ข.	.12	.16
	ก.	.15	.20
*	ง.	.46	.39
	จ.	.13	.16
59.	ก.	.12	.16
	ข.	.15	.11
*	ค.	.52	.38
	ง.	.11	.29
	จ.	.10	.12
60.	ก.	.04	.5
	ข.	.02	.26
	ก.	.90	.10
	ง.	.02	.23
	จ.	.02	.23

ภาคผนวก ค

แสดงค่าความยากของแบบทดสอบของกลุ่มตัวอย่าง 6 กลุ่ม

ตาราง 10 ค่าความยากของแบบทดสอบ กลุ่มตัวอย่าง 200 คน

ข้อที่	ค่าความยาก (b)	ข้อที่	ค่าความยาก (b)
1.	-0.30	21.	-0.20
2.	-0.62	22.	0.29
3.	*	23.	-0.05
4.	*	24.	0.02
5.	*	25.	0.46
6.	-1.81	26.	-1.01
7.	-1.18	27.	0.36
8.	.09	28.	0.61
9.	.38	29.	0.23
10.	-0.28	30.	0.00
11.	3.04	31.	0.76
12.	-0.75	32.	*
13.	0.18	33.	*
14.	0.02	34.	*
15.	-0.62	35.	*
16.	0.11	36.	*
17.	0.34	37.	*
18.	-0.05	38.	*
19.	-0.20	39.	*
20.	-0.18	40.	*

* เป็นข้อที่ไม่สอดคล้องกับโค้ง ICC จึงตัดออกในขั้นวิเคราะห์ข้อมูลในการวิจัย

ตาราง 11 ค่าความยากของแบบทดสอบ กลุ่มตัวอย่าง 300 คน

ข้อที่	ค่าความยาก (b)	ข้อที่	ค่าความยาก (b)
1.	0.79	21.	-0.25
2.	-1.03	22.	-0.05
3.	*	23.	-0.04
4.	*	24.	-0.15
5.	*	25.	1.11
6.	-1.22	26.	-1.01
7.	-1.12	27.	0.38
8.	0.11	28.	0.65
9.	0.33	29.	0.10
10.	-0.22	30.	0.20
11.	3.01	31.	0.75
12.	-0.98	32.	*
13.	0.10	33.	*
14.	0.10	34.	*
15.	-0.63	35.	*
16.	0.11	36.	*
17.	0.20	37.	*
18.	-0.01	38.	*
19.	-0.34	39.	*
20.	-0.07	40.	*

* เป็นข้อที่ไม่สอดคล้องกับโค้ง ICC จึงตัดออกในขั้นวิเคราะห์ข้อมูลในการวิจัย

ตาราง 12 ค่าความยากของแบบทดสอบ กลุ่มตัวอย่าง 450 คน

ข้อที่	ค่าความยาก (b)	ข้อที่	ค่าความยาก (b)
1.	1.80	21.	-0.21
2.	-0.42	22.	0.18
3.	-2.70	23.	0.17
4.	*	24.	-0.00
5.	*	25.	1.51
6.	-1.30	26.	-1.07
7.	-0.88	27.	0.47
8.	0.35	28.	0.99
9.	0.64	29.	0.43
10.	0.12	30.	0.25
11.	3.31	31.	1.06
12.	-0.67	32.	*
13.	-0.33	33.	*
14.	0.15	34.	*
15.	-0.47	35.	-.32
16.	0.42	36.	*
17.	0.45	37.	-1.57
18.	0.40	38.	*
19.	-0.20	39.	-1.69
20.	0.32	40.	-1.81

* เป็นข้อที่ไม่สอดคล้องกับโค้ง ICC จึงตัดออกในขั้นวิเคราะห์ข้อมูลในการวิจัย

ตาราง 13 ค่าความยากของแบบทดสอบ กลุ่มตัวอย่าง 675 คน

ข้อที่	ค่าความยาก (b)	ข้อที่	ค่าความยาก (b)
1.	2.31	21.	-0.28
2.	-0.41	22.	0.07
3.	-0.23	23.	0.18
4.	-1.25	24.	-0.18
5.	*	25.	0.95
6.	0.25	26.	0.61
7.	-0.92	27.	0.32
8.	0.28	28.	1.00
9.	0.62	29.	0.44
10.	0.41	30.	0.36
11.	3.26	31.	0.92
12.	-0.79	32.	*
13.	0.32	33.	*
14.	0.02	34.	-0.84
15.	-0.64	35.	*
16.	0.36	36.	-2.03
17.	0.34	37.	*
18.	0.35	38.	-2.13
19.	-0.35	39.	-2.25
20.	0.15	40.	*

* เป็นข้อที่ไม่สอดคล้องกับโค้ง ICC จึงตัดออกในขั้นวิเคราะห์ข้อมูลในการวิจัย

ตาราง 14 ค่าความยากของแบบทดสอบ กลุ่มตัวอย่าง 1,015 คน

ข้อที่	ค่าความยาก (b)	ข้อที่	ค่าความยาก (b)
1.	2.70	21.	-0.13
2.	-0.67	22.	0.04
3.	-0.22	23.	0.33
4.	-0.96	24.	-0.18
5.	-1.57	25.	0.70
6.	0.87	26.	-0.82
7.	-1.14	27.	0.28
8.	0.27	28.	0.95
9.	0.67	29.	0.37
10.	1.33	30.	0.70
11.	3.28	31.	0.87
12.	-0.71	32.	*
13.	0.33	33.	*
14.	0.13	34.	-1.26
15.	-0.15	35.	*
16.	0.37	36.	-2.26
17.	0.49	37.	*
18.	0.44	38.	-2.51
19.	0.17	39.	-2.63
20.	0.08	40.	*

* เป็นข้อที่ไม่สอดคล้องกับโค้ง ICC จึงตัดออกในขั้นวิเคราะห์ข้อมูลในการวิจัย

ตาราง 15 ค่าความยากของแบบทดสอบ กลุ่มตัวอย่าง 450 คน

ข้อที่	ค่าความยาก (b)	ข้อที่	ค่าความยาก (b)
1.	1.55	21.	-0.04
2.	-0.67	22.	0.02
3.	0.43	23.	0.31
4.	0.73	24.	-0.15
5.	-1.22	25.	1.44
6.	0.97	26.	-0.56
7.	0.12	27.	0.25
8.	0.15	28.	0.75
9.	0.52	29.	0.19
10.	1.00	30.	0.56
11.	3.19	31.	0.69
12.	0.46	32.	*
13.	0.23	33.	*
14.	0.04	34.	-1.85
15.	0.74	35.	*
16.	0.25	36.	-2097
17.	0.41	37.	*
18.	0.32	38.	-3.07
19.	0.04	39.	-3.19
20.	-0.13	40.	*

* เป็นข้อที่ไม่สอดคล้องกับโค้ง ICC จึงตัดออกในขั้นวิเคราะห์ข้อมูลในการวิจัย

ภาคผนวก ง

แสดงค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าของผู้สอบ
ของกลุ่มตัวอย่าง 6 กลุ่ม

ตาราง 16 แสดงค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบของกลุ่มตัวอย่าง 200 คน

ข้อที่	SE	ข้อที่	SE
1.	0.166	21	0.163
2.	0.181	22	0.150
3.	*	23	0.158
4.	*	24	0.156
5.	*	25	0.148
6.	0.280	26	0.204
7.	0.217	27	0.149
8.	0.154	28	0.147
9.	0.149	29	0.152
10.	0.165	30	0.157
11.	0.239	31	0.146
12.	0.188	32	*
13.	0.152	33	*
14.	0.156	34	*
15.	0.181	35	*
16.	0.154	36	*
17.	0.150	37	*
18.	0.158	38	*
19.	0.163	39	*
20.	0.152	40	*

* เป็นข้อที่ไม่สอดคล้องกับโค้ง ICC จึงตัดออกในขั้นวิเคราะห์ข้อมูลในการวิจัย

ตาราง 17 แสดงค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบของกลุ่มตัวอย่าง 300 คน

ข้อที่	SE	ข้อที่	SE
1.	0.119	21	0.134
2.	0.168	22	0.126
3.	*	23	0.128
4.	*	24	0.131
5.	*	25	0.121
6.	0.228	26	0.174
7.	0.172	27	0.121
8.	0.125	28	0.119
9.	0.122	29	0.125
10.	0.133	30	0.128
11.	0.195	31	0.119
12.	0.164	32	*
13.	0.125	33	*
14.	0.127	34	*
15.	0.147	35	*
16.	0.125	36	*
17.	0.123	37	*
18.	0.127	38	*
19.	0.136	39	*
20.	0.126	40	*

* เป็นข้อที่ไม่สอดคล้องกับโค้ง ICC จึงตัดออกในขั้นวิเคราะห์ข้อมูลในการวิจัย

ตาราง 18 แสดงค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบของกลุ่มตัวอย่าง 450 คน

ข้อที่	SE	ข้อที่	SE
1.	0.103	21	0.116
2.	0.123	22	0.106
3.	0.321	23	0.106
4.	*	24	0.110
5.	*	25	0.099
6.	0.169	26	0.154
7.	0.143	27	0.101
8.	0.103	28	0.097
9.	0.099	29	0.101
10.	0.107	30	0.104
11.	0.159	31	0.092
12.	0.133	32	*
13.	0.103	33	*
14.	0.106	34	0.119
15.	0.125	35	*
16.	0.102	36	0.190
17.	0.101	37	*
18.	0.102	38	0.200
19.	0.115	39	0.211
20.	0.103	40	*

* เป็นข้อที่ไม่สอดคล้องกับโค้ง ICC จึงตัดออกในขั้นวิเคราะห์ข้อมูลในการวิจัย

ตาราง 19 แสดงค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบของกลุ่มตัวอย่าง 675 คน

ข้อที่	SE	ข้อที่	SE
1.	0.95	21	0.088
2.	0.099	22	0.086
3.	0.094	23	0.093
4.	0.133	24	0.079
5.	*	25	0.105
6.	0.085	26	0.084
7.	0.117	27	0.079
8.	0.084	28	0.082
9.	0.080	29	0.083
10.	0.082	30	0.079
11.	0.130	31	0.114
12.	0.112	32	*
13.	0.084	33	*
14.	0.088	34	0.187
15.	0.106	35	*
16.	0.083	36	0.114
17.	0.083	37	*
18.	0.097	38	0.197
19.	0.086	39	0.208
20.	0.095	40	*

* เป็นข้อที่ไม่สอดคล้องกับโค้ง ICC จึงตัดออกในขั้นวิเคราะห์ข้อมูลในการวิจัย

ตาราง 20 แสดงค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบของกลุ่มตัวอย่าง 1,015 คน

ข้อที่	SE	ข้อที่	SE
1.	0.086	21	0.072
2.	0.089	22	0.068
3.	0.077	23	0.076
4.	0.099	24	0.065
5.	0.126	25	0.094
6.	0.064	26	0.069
7.	0.107	27	0.064
8.	0.069	28	0.067
9.	0.065	29	0.065
10.	0.106	30	0.065
11.	0.090	31	0.064
12.	0.068	32	*
13.	0.071	33	*
14.	0.075	34	0.113
15.	0.067	35	*
16.	0.066	36	0.185
17.	0.067	37	*
18.	0.070	38	0.195
19.	0.071	39	0.206
20.	0.075	40	*

* เป็นข้อที่ไม่สอดคล้องกับโค้ง ICC จึงตัดออกในขั้นวิเคราะห์ข้อมูลในการวิจัย

ตาราง 21 แสดงค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบของกลุ่มตัวอย่าง 1,610 คน

ข้อที่	SE	ข้อที่	SE
1.	0.053	21	0.056
2.	0.068	22	0.054
3.	0.053	23	0.059
4.	0.069	24	0.052
5.	0.083	25	0.066
6.	0.051	26	0.054
7.	0.055	27	0.054
8.	0.055	28	0.051
9.	0.052	29	0.055
10.	0.051	30	0.052
11.	0.085	31	0.051
12.	0.052	32	*
13.	0.054	33	*
14.	0.056	34	0.109
15	0.051	35	*
16.	0.054	36	0.185
17.	0.053	37	*
18.	0.056	38	0.195
19.	0.058	39	0.207
20.	0.057	40	*

* เป็นข้อที่ไม่สอดคล้องกับโค้ง ICC จึงตัดออกในขั้นวิเคราะห์ข้อมูลในการวิจัย

ภาคผนวก จ

แสดงค่าไอเก้นของแบบทดสอบ

ตาราง 22 แสดงค่าไอเก้นของตัวประกอบ 12 ตัว

ตัวประกอบ	ค่าไอเก้น
1	3.37148
2	2.88174
3	2.50352
4	2.30140
5	2.02252
6	1.70852
7	1.55944
8	1.37388
9	1.13213
10	1.09644
11	1.07658
12	1.01447

รายชื่อผู้เชี่ยวชาญ

1. รองศาสตราจารย์ เชาวนา ชวลิตธำรง	มศว.ประสานมิตร
2. รองศาสตราจารย์ ระวีวรรณ พันธ์พานิช	มศว.ประสานมิตร
3. อาจารย์ ชวลิต รวยอาจิณ	มศว.ประสานมิตร
4. อาจารย์ สมพรพงศ์ กันตามระ	โรงเรียนสวนกุหลาบวิทยาลัย
5. อาจารย์ ภานุมาศ บัวก้านทอง	โรงเรียนสวนกุหลาบวิทยาลัย

ประวัติย่อผู้วิจัย

ชื่อ ชื่อสกุล	นายสุพล นิลกลาง
วันเดือนปีเกิด	26 กรกฎาคม พ.ศ.2510
สถานที่เกิด	อำเภอประหาย จังหวัดนครราชสีมา
สถานที่อยู่ในปัจจุบัน	104/64 ถ.ทุ่งมังกร แขวงฉิมพลี ตลิ่งชัน กรุงเทพมหานคร 10170
ตำแหน่งหน้าที่การงานในปัจจุบัน	อาจารย์ 1 ระดับ 4
สถานที่ทำงานในปัจจุบัน	โรงเรียนสวนกุหลาบวิทยาลัย เขตพระนคร กรุงเทพมหานคร
ประวัติการศึกษา	ปี พ.ศ.2526 จบมัธยมศึกษาตอนต้น จากโรงเรียนวังไม้แดงพิทยาคม
	ปี พ.ศ.2530 จบมัธยมศึกษาตอนปลาย จากโรงเรียนโยธินบูรณะ
	ปี พ.ศ.2533 จบปริญญาตรี จากมหาวิทยาลัยรามคำแหง
	ปี พ.ศ.2540 จบปริญญาโท จากมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

การเปรียบเทียบคุณลักษณะของแบบทดสอบเลือกตอบระหว่างกลุ่มตัวอย่าง ที่มีจำนวนต่างกันที่วิเคราะห์ด้วยราส์ชโมเดล

บทคัดย่อ

ของ

สุพล นิลกลาง

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
พฤษภาคม 2541

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

การศึกษาครั้งนี้มีจุดมุ่งหมายเพื่อเปรียบเทียบค่าความยากของแบบทดสอบ ค่าความสามารถของผู้สอบ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าของกลุ่มตัวอย่างที่มีจำนวนไม่เท่ากัน 6 กลุ่ม คือ กลุ่ม 200 คน 300 คน 450 คน 675 คน 1,015 คน และ 1,610 คน ซึ่งวิเคราะห์ด้วยราส์ชโมเดล โดยศึกษากับกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2540 ในเขตกรุงเทพมหานคร กลุ่ม 7 จำนวน 1,610 คน ซึ่งเลือกมาโดยการสุ่มแบบแบ่งชั้น เครื่องมือที่ใช้ในการศึกษา เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 2 เรื่องโจทย์สมการ ผลการศึกษาพบว่า

ค่าความยากของแบบทดสอบ ค่าความสามารถของผู้สอบ และค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

A COMPARISON OF CHARACTERISTICS OF MULTIPLE CHOICE TEST WITH DIFFERENT SAMPLE SIZES USING RASH MODEL

AN ABSTRACT

BY

SUPHOL NILKLANG

Presented in partial fulfillment of the requirements for the Master of Education degree in Educational Measurement at Srinakharinwirot University

May 1998

The purpose of this study was to compare the difficulty indese ability of examinees, and standard error of estimation. The samples of this study composed of six samples size about 200, 300, 450, 675, 1,015 and 1,610 Matayomsuksa II students of the seventh school group in Bangkok analysed by Rash Model. This study was implemented in the second semester of 1997 academic year. The instrument used in this study was the achievement test in Mathematics of Matayomsuksa II students.

The findings of the study were as follows: The difficulty indese, ability of examinees and standard error of estimation were not statistical significant of different sample size.