

371.26013

๑4๗3๖

๖๒

จำนวนข้อและความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบด้วยวิธีต่างกัน

ปริญญานิพนธ์

ของ

ภาวณา วงศ์เพ็ญทักษ์

ว 2 ต.ค. 2541

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

ตุลาคม 2541

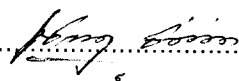
ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

๖๖๖

คณะกรรมการควบคุมและคณะกรรมการสอบได้พิจารณาปฏิญานีพนธ์ฉบับนี้แล้ว
เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาการศึกษามหาบัณฑิต
วิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒได้

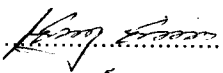
คณะกรรมการควบคุม

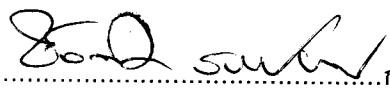

.....ประธาน
(รองศาสตราจารย์ ดร.บุญเชิด ปิญโญนนตพงษ์)

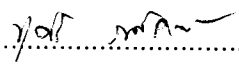

.....กรรมการ
(ผู้ช่วยศาสตราจารย์ เชาวนา ชวลิตธำรง)

คณะกรรมการสอบ

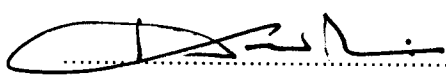

.....ประธาน
(รองศาสตราจารย์ ดร.บุญเชิด ปิญโญนนตพงษ์)


.....กรรมการ
(ผู้ช่วยศาสตราจารย์ เชาวนา ชวลิตธำรง)


.....กรรมการที่แต่งตั้งเพิ่มเติม
(อาจารย์ชวลิต รวยอาจิณ)


.....กรรมการที่แต่งตั้งเพิ่มเติม
(รองศาสตราจารย์ ชุติรี วงศ์รัตนะ)

บัณฑิตวิทยาลัยอนุมัติให้รับปฏิญานีพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒ


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

วันที่ 3 เดือน ตุลาคม พ.ศ. 2541

ปริญญาพนธ์ฉบับนี้ได้รับทุนสนับสนุนการวิจัย

ทุน “งบประมาณแผ่นดิน”

มหาวิทยาลัยศรีนครินทรวิโรฒ ประจำปีการศึกษา 2540

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้ สำเร็จลุล่วงได้ด้วยคามอนุเคราะห์และความกรุณาอย่างสูง
จากรองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์ และผู้ช่วยศาสตราจารย์เชาวนา ชวลิตธำรง ที่ได้กรุณา
ให้คำปรึกษาแนะนำ ชี้แนะแนวทาง ตรวจ แก้ไขข้อบกพร่องต่าง ๆ จนงานวิจัยฉบับนี้สำเร็จลุล่วงไปด้วยดี
ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงมา ณ โอกาสนี้

ขอกราบขอบพระคุณอาจารย์ชวลิต รวยอาจิณ และรองศาสตราจารย์ชูศรี วงศ์รัตน์ รวมทั้ง
อาจารย์ภาควิชาการวัดผลการศึกษาทุกท่าน ที่ได้กรุณาให้คำปรึกษา แนะนำ แก้ไข ปรับปรุงทำให้ปริญญานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น

ขอกราบขอบพระคุณผู้เชี่ยวชาญทุกท่านที่ให้ความช่วยเหลือ แนะนำ ให้ข้อคิดเห็นและแก้ไข
ข้อบกพร่องต่าง ๆ ในการสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า

ขอกราบขอบพระคุณมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตรที่ได้ให้ทุน"งบประมาณแผ่นดิน"
ในการทำวิจัยฉบับนี้

ขอขอบพระคุณอาจารย์พรณพิศ วาณิชยการ ภาควิชาการแนะแนวและจิตวิทยาการศึกษา ที่
กรุณาตรวจและแก้ไขบทคัดย่อภาษาอังกฤษ ของปริญญานิพนธ์ฉบับนี้ให้ถูกต้องสมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณผู้อำนวยการ คณะครู อาจารย์ทุกท่าน ของโรงเรียนในกลุ่มตัวอย่างรวมทั้งนักเรียน
ทุกคนที่ได้ให้ความร่วมมือในการเก็บรวบรวมข้อมูลเป็นอย่างดี

ขอขอบคุณโรงเรียนมัธยมศึกษาที่เอื้อเฟื้อเครื่องคอมพิวเตอร์ในการพิมพ์ปริญญานิพนธ์ฉบับนี้
ขอขอบคุณคุณโสภณ เต็มอุดม ที่ให้คำแนะนำในการพิมพ์คอมพิวเตอร์ ขอขอบคุณคุณมิ่ง คุณธิตกมล
เทพครเมือง ที่ให้คำแนะนำในการวิเคราะห์ข้อมูล ขอขอบคุณคุณพัชรา ทศนวิจิตรวงศ์ คุณราณี เพชรคุ้ม
คุณอัมพร วิชัยศรี คุณลัดดา เสือสืบพันธุ์ คุณทองอยู่ สาระ คุณอรอนงค์ บำรุง และคุณพรรณิ เจริญสุขบุตร
ที่ช่วยเหลือในการเก็บข้อมูล ขอขอบคุณเพื่อน ๆ วิชาเอกการวัดผลการศึกษา ภาควิเศษ รุ่น 7 และ
คุณกนกวรรณ สุตานิช ที่เป็นกำลังใจและช่วยเหลือในการทำปริญญานิพนธ์

คุณค่าและประโยชน์อันพึงมีจากปริญญานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบเป็นเครื่องบูชาพระคุณบิดา
มารดา และครูอาจารย์ ตลอดจนผู้มีพระคุณที่ได้ให้การอบรมสั่งสอนตลอดมา

ภาวนา วงศ์เพ็ญทักษ์

สารบัญ

บทที่	หน้า
1 บทนำ	1
ภูมิหลัง	1
ความมุ่งหมายของการศึกษาค้นคว้า	3
ความสำคัญของการศึกษาค้นคว้า	4
ขอบเขตของการศึกษาค้นคว้า	4
นิยามศัพท์เฉพาะ	4
2 เอกสารและงานวิจัยที่เกี่ยวข้อง	8
ความเชื่อมั่นของแบบทดสอบ	8
ทฤษฎีความเชื่อมั่น	8
การตีความหมายสัมประสิทธิ์ความเชื่อมั่นแบบต่าง ๆ	9
วิธีประมาณค่าความเชื่อมั่น	10
องค์ประกอบที่มีอิทธิพลต่อค่าความเชื่อมั่น	14
คุณภาพของแบบทดสอบ	15
การวิเคราะห์ข้อสอบ	15
เทคนิคการวิเคราะห์ข้อสอบ	16
ความหมายของความถนัด	23
ทฤษฎีสมรรถภาพสมองและความถนัด	23
ประโยชน์ของความถนัด	25
งานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้า	26
สมมติฐานในการศึกษาค้นคว้า	27
3 วิธีดำเนินการศึกษาค้นคว้า	29
ประชากร	29
กลุ่มตัวอย่าง	29
เครื่องมือที่ใช้ในการศึกษาค้นคว้า	30
ขั้นตอนการสร้างแบบทดสอบวัดความถนัดทางการเรียน	30
ตัวอย่างแบบทดสอบ	31
วิธีดำเนินการรวบรวมข้อมูล	32
ขั้นตอนการวิเคราะห์ข้อมูล	32
สถิติที่ใช้ในการวิเคราะห์ข้อมูล	34
4 ผลการวิเคราะห์ข้อมูล	38
การนำเสนอผลการวิเคราะห์ข้อมูล	38

บทที่	หน้า
สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล	38
ผลการวิเคราะห์ข้อมูล	39
5 สรุป อภิปรายผล และข้อเสนอแนะ	60
ความมุ่งหมายของการศึกษาค้นคว้า	60
กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้า	60
เครื่องมือที่ใช้ในการศึกษาค้นคว้า	60
การเสนอผลวิเคราะห์ข้อมูล	61
สรุปผลการวิจัย	61
อภิปรายผล	63
ข้อเสนอแนะ	66
บรรณานุกรม	68
ภาคผนวก	72
ประวัติย่อของผู้วิจัย	101

บัญชีตาราง

ตาราง		หน้า
1	จำนวนกลุ่มตัวอย่าง จำแนกตามขนาดโรงเรียน	30
2	ค่าสถิติพื้นฐานของคะแนนที่ได้จากการทดสอบ	39
3	ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบทดสอบที่วิเคราะห์ข้อสอบ วิธีหาค่าอำนาจจำแนกและค่าความยาก ที่คัดเลือกได้ตามเกณฑ์	41
4	ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบทดสอบที่วิเคราะห์ข้อสอบ วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ ที่คัดเลือกได้ตามเกณฑ์	41
5	ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบทดสอบที่วิเคราะห์ข้อสอบ วิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ ที่คัดเลือกได้ตามเกณฑ์	42
6	เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากการวิเคราะห์ข้อสอบ วิธีต่างกัน 3 วิธี จากกลุ่มตัวอย่าง 30 คน	44
7	เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากการวิเคราะห์ข้อสอบ วิธีต่างกัน 3 วิธี จากกลุ่มตัวอย่าง 60 คน	44
8	เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากการวิเคราะห์ข้อสอบ วิธีต่างกัน 3 วิธี จากกลุ่มตัวอย่าง 90 คน	45
9	เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากการวิเคราะห์ข้อสอบ วิธีต่างกัน 3 วิธี จากกลุ่มตัวอย่าง 120 คน	45
10	เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากการวิเคราะห์ข้อสอบ วิธีต่างกัน 3 วิธี จากกลุ่มตัวอย่าง 150 คน	46
11	เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากการวิเคราะห์ข้อสอบ วิธีต่างกัน 3 วิธี จากกลุ่มตัวอย่าง 300 คน	46
12	เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากวิธีหาค่าอำนาจจำแนก และค่าความยาก จากกลุ่มตัวอย่าง 6 ขนาด	47
13	เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากวิธีใช้โปรแกรมสำเร็จ โดยลดจำนวนข้อ จากกลุ่มตัวอย่าง 6 ขนาด	48
14	เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากวิธีใช้ดัชนีความเชื่อมั่น ของข้อสอบแต่ละข้อ จากกลุ่มตัวอย่าง 6 ขนาด	48
15	ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบทดสอบที่วิเคราะห์ข้อสอบ โดยการหาค่าอำนาจจำแนกและค่าความยาก จำนวน 50 ข้อ	49
16	ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบทดสอบที่วิเคราะห์ข้อสอบ ด้วยโปรแกรมสำเร็จโดยลดจำนวนข้อ จำนวน 50 ข้อ	50
17	ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบทดสอบที่วิเคราะห์ข้อสอบ โดยใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ จำนวน 50 ข้อ	50

18	ค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด	52
19	เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ จำนวน 50 ข้อ ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 30 คน	53
20	เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ จำนวน 50 ข้อ ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 60 คน	54
21	เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ จำนวน 50 ข้อ ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 90 คน	55
22	เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ จำนวน 50 ข้อ ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 120 คน	55
23	เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ จำนวน 50 ข้อ ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 150 คน	56
24	เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ จำนวน 50 ข้อ ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาดใหญ่ 300 คน	56
25	เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ จำนวน 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบวิธีหาค่าอำนาจจำแนกและ ค่าความยาก ระหว่างกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด	57
26	เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ จำนวน 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบวิธีใช้โปรแกรมสำเร็จ โดยลดจำนวนข้อ ระหว่างกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด	58
27	เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ จำนวน 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบวิธีใช้ดัชนีความเชื่อมั่น ของข้อสอบแต่ละข้อ ระหว่างกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด	58
28	จำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์และค่าความเชื่อมั่นของแบบทดสอบ ที่วิเคราะห์ข้อสอบ 3 วิธี จากกลุ่มตัวอย่าง 6 ขนาด	61
29	สรุปค่าความเชื่อมั่นของแบบทดสอบที่วิเคราะห์ข้อสอบ 3 วิธี ในแต่ละขนาดกลุ่มตัวอย่าง	62
30	สรุปค่าความเชื่อมั่นของแบบทดสอบที่วิเคราะห์ข้อสอบวิธีเดียวกัน จากกลุ่มตัวอย่าง 6 ขนาด	63
31	แสดงข้อสอบ 100 ข้อ ที่วิเคราะห์ข้อสอบวิธีหาค่าอำนาจจำแนกและค่าความยาก จากกลุ่มตัวอย่าง 6 ขนาด	74
32	แสดงข้อสอบ 100 ข้อ ที่วิเคราะห์ข้อสอบวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ จากกลุ่มตัวอย่าง 6 ขนาด	78
33	แสดงข้อสอบ 100 ข้อ ที่วิเคราะห์ข้อสอบวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบ แต่ละข้อ จากกลุ่มตัวอย่าง 6 ขนาด	82

บัญชีภาพประกอบ

ภาพประกอบ	หน้า
1 ค่าคะแนนเฉลี่ยของแบบทดสอบจำนวน 100 ข้อ จากกลุ่มตัวอย่าง 6 ขนาด	40
2 ค่าคะแนนเฉลี่ยของแบบทดสอบที่คัดเลือกได้ตามเกณฑ์ ที่วิเคราะห์ข้อสอบ 3 วิธี จากกลุ่มตัวอย่าง 6 ขนาด	43
3 ค่าคะแนนเฉลี่ยของข้อสอบ 50 ข้อแรกที่ดีที่สุดที่วิเคราะห์ข้อสอบ 3 วิธี จากกลุ่มตัวอย่าง 6 ขนาด	51
4 แผนภาพค่าความเชื่อมั่นของแบบทดสอบที่วิเคราะห์ข้อสอบวิธีต่างกัน 3 วิธี จากกลุ่มตัวอย่าง 6 ขนาด	53

บทที่ 1

บทนำ

ภูมิหลัง

สิ่งที่มีความจำเป็นและสำคัญประการหนึ่งของระบบการจัดการศึกษา คือ การวัดและการประเมินผลความก้าวหน้าของนักเรียน เพราะผลจากการวัดและการประเมินความก้าวหน้าของนักเรียนจะช่วยให้ครูนักเรียนและผู้ปกครองนำไปใช้ประโยชน์ในการพัฒนาคุณภาพของการศึกษาให้ดีขึ้นและเป็นไปตามจุดหมายที่กำหนดไว้ (อนันต์ ศรีโสภณ. 2524 : คำนำ) นอกจากนี้ยังนำผลของการวัดไปใช้หลายทางเพื่อพัฒนาบุคคลแต่ละคนให้บรรลุถึงความสามารถสูงสุดตามศักยภาพ (Potentiality) ที่แต่ละบุคคลมี (สวัสดี ประทุมราช. 2517 : 33) อย่างไรก็ตาม การวัดผลการศึกษาก็เกิดคุณค่าดังกล่าวข้างต้นได้ จำเป็นอย่างยิ่งที่จะต้องอาศัย “เครื่องมือ” ในการวัดที่เหมาะสมและมีคุณภาพสูง ทั้งนี้เพราะถ้าได้เครื่องมือที่ไม่ดีย่อมส่งผลให้การวัดก็ไม่ถูกต้องแน่นอน ทำให้ไม่สามารถนำผลการวัดมาใช้ประโยชน์ได้ (สวัสดี ประทุมราช. 2517 : 32) คุณสมบัติที่มีความสำคัญที่สุดของเครื่องมือในการวัดผลการศึกษา คือ ค่าความเชื่อมั่นและค่าความเที่ยงตรงของแบบทดสอบ (อนันต์ ศรีโสภณ. 2524 : 42) ในการเลือกใช้แบบทดสอบสำหรับการวัดผลการศึกษาจำเป็นต้องศึกษาถึงคุณภาพของแบบทดสอบว่ามีคุณภาพสูงหรือต่ำเพียงใด โดยการนำแบบทดสอบไปทดลองใช้และตรวจสอบคุณภาพ นั้นนอลลี กล่าวว่า สิ่งที่เป็นประโยชน์และจำเป็นต้องทราบอย่างหนึ่งของแบบทดสอบใด ๆ ก็คือค่าความเชื่อมั่นของแบบทดสอบ เพราะค่าความเชื่อมั่นของแบบทดสอบเป็นดัชนีชี้ว่าแบบทดสอบนั้นมีคุณภาพหรือไม่ (Nunnally. 1964 : 59) ในการสร้างแบบทดสอบจึงจำเป็นต้องหาค่าความเชื่อมั่นของแบบทดสอบทุกครั้ง (บุญเชิด ภิญโญนต์พงษ์. 2527 : 149) ความเชื่อมั่น (Reliability) ของแบบทดสอบขึ้นอยู่กับลักษณะของข้อสอบ ซึ่งสามารถสร้างให้แบบทดสอบมีความเชื่อมั่นสูงได้ด้วยการวิเคราะห์ข้อสอบ การวิเคราะห์ข้อสอบจะช่วยทำให้แบบทดสอบสั้นลงและขณะเดียวกันก็เพิ่มความเชื่อมั่นของแบบทดสอบนั้น (Anastasi . 1968 : 158)

การวิเคราะห์ข้อสอบ คือ วิธีการที่ใช้พิจารณาคุณภาพของข้อสอบ สิ่งสำคัญในการพิจารณาคุณภาพของข้อสอบได้แก่ ระดับความยากของข้อสอบ ค่าอำนาจจำแนก และความเที่ยงตรงเชิงเนื้อหาของข้อสอบ (Content Validity) โดยดูว่าข้อสอบมีความเกี่ยวข้องกับเนื้อหาวิชาเพียงใด ส่วนระดับความยากและอำนาจการจำแนกของข้อสอบพิจารณาจากตัวเลขที่ได้จากการนำข้อสอบนั้นไปทดสอบ (อนันต์ ศรีโสภณ. 2524 : 148) แอลเลน และเยน (Allen and Yen. 1979 : 120 - 121) กล่าวว่า การใช้ระดับความยากและค่าอำนาจจำแนกของข้อสอบหรือสหสัมพันธ์แบบพอยท์ไบซีเรียล (Point - Biserial) ระหว่างคะแนนแต่ละข้อกับคะแนนรวมของแบบทดสอบนั้นเป็นการวิเคราะห์ข้อสอบที่เหมาะสมกับครูและนักพัฒนาแบบทดสอบ การใช้ระดับความยากของข้อสอบเป็นประโยชน์ในการประเมินว่าความยากของข้อสอบแต่ละข้อเหมาะสมกับระดับของผู้สอบเพียงใด/ค่าอำนาจจำแนกของข้อสอบหรือสหสัมพันธ์พอยท์ไบซีเรียล ระหว่างคะแนนแต่ละข้อกับคะแนนรวมของแบบทดสอบนี้แสดงให้เห็นว่าข้อสอบแต่ละข้อได้จำแนกผู้สอบที่ทำคะแนนทั้งฉบับได้ดีกับผู้สอบที่ทำแบบทดสอบทั้งฉบับได้ไม่ดี

ความยากของข้อที่ i , (p_i) คือ สัดส่วนของผู้สอบที่ทำข้อสอบข้อนั้นถูก ข้อสอบที่มีค่าความยากเท่ากับ .3 จะยากกว่าข้อสอบที่มีค่าความยากเท่ากับ .8 เพราะว่ามีผู้ตอบถูกน้อยกว่า ความยากของ

ข้อสอบนั้นมิมีประโยชน์ในการวิเคราะห์ข้อสอบ ถ้า p_i มีค่าใกล้ 0 หรือ 1 ข้อสอบข้อนั้นควรได้รับการปรับปรุงหรือตัดทิ้งเพราะไม่สามารถบอกความแตกต่างระดับความสามารถของผู้สอบได้ ถ้า $p_i = 0$ แสดงว่า ไม่มีผู้ใดตอบข้อนั้นถูก ข้อสอบข้อนี้ยากเกินไปและไม่มีประโยชน์แต่อย่างใด ถ้า $p_i = 1$ แสดงว่า ผู้สอบทุกคนตอบถูก ซึ่งก็ไม่สามารถแสดงถึงความแตกต่างของผู้สอบได้ ความยากของข้อสอบที่มีค่าประมาณ .3 ถึง .7 จะให้ข้อมูลว่าแบบทดสอบแสดงถึงความแตกต่างระหว่างผู้สอบได้สูงสุด

ค่าอำนาจจำแนกที่ดีที่สุดในระหว่างผู้สอบในทุกระดับความสามารถแล้ว ระดับความยากที่ดีที่สุดควรมีพิสัยประมาณ .3 ถึง .7 เคลลี (Allen and Yen. 1979 : 122 ; citing Kelley. 1939) กล่าวว่า "อำนาจจำแนกของข้อสอบข้อที่ i , (d_i) คือ ความแตกต่างระหว่างสัดส่วนของผู้สอบที่ได้คะแนนสูงที่ทำข้อสอบได้ถูกต้องและสัดส่วนของผู้สอบที่ได้คะแนนต่ำและทำข้อสอบได้ถูกต้อง ช่วงสูงและช่วงต่ำโดยทั่วไปจะหมายถึง สูงและต่ำ 10% - 33% ของกลุ่มตัวอย่าง ถ้าคะแนนรวมจากการทดสอบมีการแจกแจงเป็นโค้งปกติ (Normally Distributed) ควรใช้กลุ่มสูงและกลุ่มต่ำ 27%" เคอร์ตัน (Allen and Yen. 1979 : 122 ; citing Cureton. 1957) กล่าวว่า "ถ้าคะแนนรวมจากแบบทดสอบแจกแจงเป็นโค้งที่แบนกว่าโค้งปกติแล้ว จำนวนเปอร์เซ็นต์ที่ดีที่สุดควรจะมีค่าใกล้เคียงกับ 33% การนำไปใช้ส่วนใหญ่ใช้ 25% ถึง 33% จะให้ค่าประมาณ d_i ที่ใกล้เคียงกัน การหาอำนาจจำแนกอีกวิธีหนึ่ง คือ การหาสหสัมพันธ์แบบพอยท์ไบเซเรียล (Point Biserial) ระหว่างคะแนนแต่ละข้อและคะแนนรวมของแบบทดสอบ (r_{ix}) ซึ่งหมายถึง สหสัมพันธ์ระหว่างคะแนนข้อที่ i และคะแนนรวมของแบบทดสอบ x "

ดัชนีอำนาจจำแนกของข้อสอบและสหสัมพันธ์ระหว่างข้อสอบและแบบทดสอบทั้งฉบับ (r_{ix}) เป็นสิ่งที่มีคุณค่า สำหรับข้อสอบที่ดีนั้น d_i และ r_{ix} ควรเป็นบวก ผู้สอบที่คะแนนรวมสูงควรตอบคำถามรายข้อได้ถูกต้องมากกว่าผู้ที่ได้คะแนนรวมต่ำ ข้อสอบที่มีค่า d_i หรือ r_{ix} ที่เป็นลบ จะไม่สามารถวัดสิ่งที่แบบทดสอบต้องการจะวัด ข้อสอบข้อนั้นควรได้รับการปรับปรุงหรือตัดทิ้ง ดังที่ กานดา พูนลาภทวี (2528 : 160-161) กล่าวว่า "ค่าอำนาจจำแนกของข้อสอบมีค่าตั้งแต่ -1.00 ถึง 1.00 ข้อสอบที่ r มีค่าเป็นบวก เป็นข้อสอบที่คนเก่งมีแนวโน้มที่จะตอบถูกมาก คนอ่อนมีแนวโน้มที่จะตอบผิดมาก ยิ่ง r เป็นค่าบวกสูงมากเท่าใด คนเก่งจะมีแนวโน้มตอบถูกมากเท่านั้น ส่วนคนอ่อนจะมีแนวโน้มตอบผิดมากขึ้นด้วย ถ้าข้อสอบที่มีนักเรียนกลุ่มเก่งตอบถูกหมดกลุ่มอ่อนตอบผิดหมด r มีค่าเป็น +1.00 ข้อสอบที่ r มีค่าเป็นลบ เป็นข้อสอบที่คนเก่งมีแนวโน้มที่จะตอบผิดมาก คนอ่อนมีแนวโน้มที่จะตอบถูกมาก ถ้าข้อสอบที่มีนักเรียนกลุ่มเก่งตอบผิดหมด กลุ่มอ่อนตอบถูกหมด r มีค่าเป็น -1.00 ซึ่งเป็นข้อสอบที่ไม่ดี ข้อสอบที่ r มีค่าเป็นศูนย์ เป็นข้อสอบที่คนเก่งหรืออ่อนอาจจะตอบถูกหรือผิดก็ได้ไม่แน่นอนไม่สามารถแยกคนเก่งและคนอ่อนออกจากกันได้ "

แอลเลน และเยน (Allen and Yen. 1979 : 124 - 126) กล่าวว่า สถิติที่ใช้ในการวิเคราะห์ข้อสอบมี 4 ค่า คือ

- 1) ความยากของข้อสอบ (p_i)
- 2) ส่วนเบี่ยงเบนมาตรฐานของคะแนนข้อสอบ ($S_i = \sqrt{p_i(1-p_i)}$)
- 3) ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ ($S_i r_{ix^*}$)
- 4) ดัชนีความเที่ยงตรงของข้อสอบ ($S_i r_{iy}$)

ถ้าต้องการเลือกข้อสอบที่ดีที่สุด k ข้อ จากแบบทดสอบทั้งฉบับ N ข้อ เพื่อให้ได้ความเชื่อมั่นสอดคล้องภายใน (Internal Consistency Reliability) $r_{xx'}$ สูงสุด เราสามารถทำได้โดยกำหนดจุดส่วนเบี่ยงเบนมาตรฐานของคะแนนแต่ละข้อ (S_i) และดัชนีความเชื่อมั่นของข้อสอบ (Reliability Index :

S_{ix}^2) แบบทดสอบที่ดีที่สุดจะมีค่าสหสัมพันธ์ระหว่างข้อสอบแต่ละข้อและคะแนนรวมของแบบทดสอบสูง นั่นคือ ถ้าเลือกข้อสอบ k ข้อ จากข้อสอบทั้งหมด N ข้อ โดยแต่ละข้อมีความสัมพันธ์กับคะแนนรวมสูง ก็จะได้แบบทดสอบที่มีความเชื่อมั่นสอดคล้องภายในสูงด้วย (Allen and Yen. 1979 : 125 ; citing Gulliksen. 1950)

ธวัชชัย งามสันติวงศ์ (2540 : 410 - 415) กล่าวถึง วิธีการวิเคราะห์ข้อสอบด้วยโปรแกรมสำเร็จ โดยตัดข้อสอบออกทีละข้อ เพื่อต้องการปรับปรุงแบบทดสอบให้มีค่าความเชื่อมั่นสูงยิ่งขึ้น โดยการตัดข้อสอบออกทีละข้อแล้วหาค่าความเชื่อมั่น ซึ่งการตัดข้อสอบบางข้อที่ไม่มีคุณภาพออก จะทำให้ค่าความเชื่อมั่นเพิ่มขึ้นกว่าเดิม

เนื่องจากการวิเคราะห์ข้อสอบมีหลายวิธีดังกล่าวข้างต้น ซึ่งอาจส่งผลให้แบบทดสอบที่คัดเลือกจากวิธีการวิเคราะห์ข้อสอบต่างกัน ได้จำนวนข้อและค่าความเชื่อมั่นแตกต่างกัน ด้วยเหตุผลดังกล่าว ผู้วิจัยจึงมีความสนใจว่าถ้าคัดเลือกข้อสอบด้วยการวิเคราะห์ข้อสอบต่างวิธีกัน จะมีผลทำให้ค่าความเชื่อมั่นของแบบทดสอบที่คัดเลือกจากแต่ละวิธีต่างกันหรือไม่ โดยศึกษาจากแบบทดสอบความถนัดด้านคณิตศาสตร์ แบบเลือกตอบ ตอบถูกได้ 1 ตอบผิดได้ 0 จากการวิเคราะห์ข้อสอบต่างกัน 3 วิธี คือ 1) วิธีหาค่าอำนาจจำแนก (Discrimination) และหาค่าความยาก (Item Difficulty) 2) วิธีใช้โปรแกรมสำเร็จรูปโดยลดจำนวนข้อ และ 3) วิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ ซึ่งสอบจากกลุ่มตัวอย่างขนาดเล็กและขนาดใหญ่ เพื่อศึกษาว่าค่าความเชื่อมั่นที่ได้จะเป็นอย่างไร

ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อเปรียบเทียบจำนวนข้อสอบของแบบทดสอบความถนัดด้านคณิตศาสตร์ ที่คัดเลือกได้ตามเกณฑ์ จากวิธีการวิเคราะห์ข้อสอบต่างกัน 3 วิธี คือ 1) วิธีหาค่าอำนาจจำแนกแบบสหสัมพันธ์พอยท์ไบซีเรียล (Point Biserial Correlation) และหาค่าความยาก (Item Difficulty) 2) วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ 3) วิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ โดยศึกษาแต่ละขนาดกลุ่มตัวอย่าง คือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน

2. เพื่อเปรียบเทียบจำนวนข้อสอบของแบบทดสอบความถนัดด้านคณิตศาสตร์ ที่คัดเลือกได้ตามเกณฑ์ ระหว่างกลุ่มตัวอย่าง คือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน ที่คัดเลือกข้อสอบด้วยวิธีการวิเคราะห์ข้อสอบ 3 วิธี คือ 1) วิธีหาค่าอำนาจจำแนกแบบสหสัมพันธ์พอยท์ไบซีเรียล (Point Biserial Correlation) และหาค่าความยาก (Item Difficulty) 2) วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ 3) วิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ

3. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ จำนวน 50 ข้อ ที่คัดเลือกข้อสอบด้วยวิธีการวิเคราะห์ข้อสอบต่างกัน 3 วิธี คือ 1) วิธีหาค่าอำนาจจำแนกแบบสหสัมพันธ์พอยท์ไบซีเรียล (Point Biserial Correlation) และหาค่าความยาก (Item Difficulty) 2) วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ 3) วิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ โดยศึกษาแต่ละขนาดกลุ่มตัวอย่าง คือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน

4. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ จำนวน 50 ข้อ ระหว่างกลุ่มตัวอย่าง คือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน ที่คัดเลือกข้อสอบด้วยวิธีการวิเคราะห์ข้อสอบ 3 วิธี คือ 1) วิธีหาค่าอำนาจจำแนกแบบสหสัมพันธ์พอยท์ไบซีเรียล

(Point Biserial Correlation) และหาค่าความยาก (Item Difficulty) 2) วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ 3) วิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ

ความสำคัญของการศึกษาค้นคว้า

ผลจากการศึกษาค้นคว้าครั้งนี้ เป็นประโยชน์ในการส่งเสริมความรู้ด้านของการวิเคราะห์ข้อสอบ และเป็นแนวทางในการเลือกใช้วิธีการวิเคราะห์ข้อสอบความถนัดในแต่ละขนาดกลุ่มตัวอย่าง เพื่อให้ได้แบบทดสอบที่มีค่าความเชื่อมั่นสูง

ขอบเขตการศึกษาค้นคว้า

ประชากรและกลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้า

ประชากรที่ใช้ในการศึกษาค้นคว้า

ประชากรที่ใช้ในการศึกษาครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2540 ของโรงเรียนสังกัดกรมสามัญศึกษา เขตกรุงเทพมหานคร จำนวน 59,744 คน

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้า

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2540 จำนวน 750 คน ซึ่งได้มาโดยวิธีการสุ่มแบบสองขั้นตอน (Two-Stage Random Sampling)

ตัวแปรที่ศึกษา

1. ตัวแปรต้น คือ

1.1 วิธีวิเคราะห์ข้อสอบซึ่งมี 3 วิธี ได้แก่

1.1.1 วิธีหาค่าอำนาจจำแนกแบบสหสัมพันธ์พอยท์ไบซีเรียล

(Point Biserial Correlation) และค่าความยาก (Item Difficulty)

1.1.2 วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ

1.1.3 วิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ

1.2. ขนาดกลุ่มตัวอย่าง คือ ขนาดเล็ก 30 , 60 , 90 , 120 , 150 คน และ

ขนาดใหญ่ 300 คน

2. ตัวแปรตาม ได้แก่ ค่าความเชื่อมั่นของแบบทดสอบและจำนวนข้อสอบที่คัดเลือกได้

ตามเกณฑ์

นิยามศัพท์เฉพาะ

1. ความเชื่อมั่น หมายถึง อัตราส่วนระหว่างความแปรปรวนของคะแนนจริงต่อความแปรปรวนของคะแนนที่สอบได้ (σ_T^2 / σ_x^2) ซึ่งประมาณได้จากค่า s_T^2 และ s_x^2 ตามลำดับ

2. การวิเคราะห์ข้อสอบ หมายถึง การตรวจหาคุณภาพของข้อสอบว่าข้อสอบแต่ละข้อมีคุณภาพหรือไม่ ซึ่งแบ่งเป็น 3 วิธี คือ

2.1 วิธีหาค่าอำนาจจำแนกและค่าความยากของข้อสอบ

- ค่าอำนาจจำแนกของข้อสอบ หมายถึง ดัชนีที่บ่งบอกว่าข้อสอบข้อนั้นสามารถจำแนกนักเรียนออกเป็นสองกลุ่มคือ กลุ่มสูงกับกลุ่มต่ำ ซึ่งเป็นสหสัมพันธ์ระหว่างความสามารถกับการตอบถูก ผิด ในการวิจัยครั้งนี้หาค่าอำนาจแบบสหสัมพันธ์พอยท์ไบเซรียล (Point Biserial Correlation)

- ค่าความยากของข้อสอบ หมายถึง ดัชนีบ่งบอกความยากของข้อสอบ ซึ่งเป็นสัดส่วนหรือร้อยละของจำนวนคนที่ตอบข้อนั้นถูก

2.2 วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ หมายถึง เทคนิคในการเพิ่มความเชื่อมั่นของคะแนนรวมให้มีค่าสูงสุด โดยการคำนวณความเชื่อมั่นของแบบทดสอบจากจำนวนข้อทั้งหมดแล้วพิจารณาจากการลดจำนวนข้อลงทีละข้อ โดยมีขั้นตอนดังนี้

2.2.1 หาค่าความเชื่อมั่นของแบบทดสอบทั้งหมด 100 ข้อ

2.2.2 ตัดข้อสอบข้อ 1 ออก จะเหลือข้อสอบทั้งหมด 99 ข้อ และหาค่าความเชื่อมั่น ต่อไปนำข้อ 1 เข้ามาแล้วตัดข้อ 2 ออกเพียงหนึ่งข้อ จะเหลือข้อสอบทั้งหมด 99 ข้อ เช่นเดิม และหาค่าความเชื่อมั่น ตัดข้อสอบออกทีละข้อเช่นนี้ไปจนถึงข้อสุดท้ายและหาค่าความเชื่อมั่นทุกครั้งหลังจากตัดข้อสอบแต่ละข้อ

2.2.3 พิจารณาว่า ถ้าตัดข้อสอบข้อใดแล้ว ทำให้ค่าความเชื่อมั่นเพิ่มขึ้น ก็จะตัดข้อนั้นออกไปเลย ข้อสอบจะลดลง 1 ข้อ

2.2.4 จากข้อสอบที่เหลือ 99 ข้อ ขั้นตอนไปตัดข้อสอบออกตามวิธีการในข้อ 2.2.2 จนครบ 99 ข้อ และพิจารณาค่าความเชื่อมั่นว่า ถ้าตัดข้อสอบข้อใดแล้วทำให้ค่าความเชื่อมั่นของแบบทดสอบเพิ่มขึ้น ก็จะตัดข้อนั้นออกไปเลย

2.2.5 ทำการตัดข้อสอบออกทีละข้อและดำเนินการเช่นขั้น 2.2.3 จนกว่าค่าความเชื่อมั่นจะคงที่แล้วเริ่มลดลงจึงหยุดตัดข้อสอบ จำนวนข้อสอบที่เหลือก็จะเป็นแบบทดสอบที่มีค่าความเชื่อมั่นสูงสุด

2.2.6 คัดเลือกข้อสอบจำนวน 50 ข้อ ถ้าแบบทดสอบที่ได้จากข้อ 2.2.5 มีจำนวนมากกว่า 50 ข้อ ให้ดำเนินการตัดข้อสอบตามวิธีการข้างต้น โดยพิจารณาค่าความเชื่อมั่นของแบบทดสอบในแต่ละรอบว่าถ้าตัดข้อสอบข้อใดออกแล้วทำให้มีค่าความเชื่อมั่นสูงสุดในรอบนั้น ก็จะตัดข้อนั้นทิ้ง คัดเลือกข้อสอบให้เหลือ 50 ข้อ เพื่อศึกษาค่าความเชื่อมั่นของแบบทดสอบ

2.3 วิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ หมายถึง การเลือกข้อสอบที่ดีที่สุด k ข้อ จากข้อสอบทั้งหมด N ข้อ โดยการตัดข้อที่ไม่มีคุณภาพทั้งด้วยวิธีการคัดเลือกข้อสอบที่มีสหสัมพันธ์กับคะแนนรวมของแบบทดสอบสูงสุด แล้วนำไปคำนวณค่าดัชนีความเชื่อมั่นของข้อสอบ (Item Reliability Index : $S_i r_{ix^*}$) เพื่อให้ได้ความเชื่อมั่นสอดคล้องภายใน (Internal Consistency Reliability) สูงสุด เมื่อ

S_i แทน ส่วนเบี่ยงเบนมาตรฐานของข้อสอบข้อที่ i

r_{ix^*} แทน สหสัมพันธ์แบบพอยท์ไบเซรียลระหว่างคะแนนสอบแต่ละข้อกับ

คะแนนรวมของแบบทดสอบ

ถ้าดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ (Item Reliability Index) มีค่าสูง ค่าความเชื่อมั่นจะมีค่าสูงด้วย การหาดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อมีวิธีการ ดังนี้

2.3.1 หาค่าความเชื่อมั่นของแบบทดสอบทั้งหมด 100 ข้อ

2.3.2 ตัดข้อสอบข้อ 1 ออก จะเหลือข้อสอบทั้งหมด 99 ข้อ หากค่าดัชนีความเชื่อมั่นของข้อสอบ ($S_i r_{ix^*}$) ข้อ 2 จนถึงข้อ 100 แล้วนำดัชนีความเชื่อมั่นของข้อสอบที่ได้ทั้ง 99 ค่าไปคำนวณหาค่าความเชื่อมั่น ($r_{xx'}$) ต่อไปตัดข้อสอบข้อ 2 ออกเพียงหนึ่งข้อ จะเหลือข้อสอบทั้งหมด 99 ข้อเช่นเดิม หากค่าดัชนีความเชื่อมั่นของข้อสอบ ($S_i r_{ix^*}$) ข้อ 1 ข้อ 3 จนถึงข้อ 100 แล้วนำดัชนีความเชื่อมั่นของข้อสอบที่ได้ทั้ง 99 ค่าไปคำนวณค่าความเชื่อมั่น ($r_{xx'}$) ตัดข้อสอบออกทีละข้อจนถึงข้อ 100 และหากค่าดัชนีความเชื่อมั่นของข้อสอบ ($S_i r_{ix^*}$) แล้วนำไปคำนวณค่าความเชื่อมั่น ($r_{xx'}$) ทุกครั้ง

2.3.3 พิจารณาว่า ถ้าตัดข้อสอบข้อใดแล้ว ทำให้ค่าความเชื่อมั่นเพิ่มขึ้น ก็จะตัดข้อนั้นออกไปเลย ข้อสอบจะลดลง 1 ข้อ

2.3.4 จากแบบทดสอบที่เหลือ 99 ข้อ ขั้นตอนต่อไปตัดข้อสอบตามวิธีการในข้อ 2.3.2 จนครบ 99 ข้อ และพิจารณาค่าความเชื่อมั่นว่าถ้าตัดข้อใดแล้วทำให้ค่าความเชื่อมั่นเพิ่มขึ้น ก็จะตัดข้อนั้นออกไปเลย

2.3.5 ทำการตัดข้อสอบออกทีละข้อและดำเนินการเช่นขั้น 2.3.3 จนกว่าค่าความเชื่อมั่นจะคงที่หรือลดลงจึงหยุดตัดข้อสอบ จำนวนข้อสอบที่เหลือจะเป็นแบบทดสอบที่มีค่าความเชื่อมั่นสูงสุด

2.3.6 คัดเลือกข้อสอบจำนวน 50 ข้อ ถ้าแบบทดสอบที่ได้จากข้อ 2.3.5 มีจำนวนมากกว่า 50 ข้อ ให้ดำเนินการตัดข้อสอบตามวิธีการข้างต้น โดยพิจารณาค่าความเชื่อมั่นของแบบทดสอบในแต่ละรอบว่าถ้าตัดข้อสอบข้อใดออกแล้วทำให้ค่าความเชื่อมั่นสูงสุดในรอบนั้น ก็จะตัดข้อนั้นทิ้ง คัดเลือกข้อสอบให้เหลือ 50 ข้อ เพื่อศึกษาค่าความเชื่อมั่นของแบบทดสอบ

3. ข้อสอบที่คัดเลือกได้ตามเกณฑ์ หมายถึง ข้อสอบที่คัดเลือกตามเกณฑ์จากการวิเคราะห์ 3 วิธี คือ

3.1 วิธีหาค่าอำนาจจำแนกและค่าความยาก คือ ข้อสอบทุกข้อที่มีอำนาจจำแนกตั้งแต่ .20 ขึ้นไป และมีค่าความยากตั้งแต่ .20 - .80

3.2 วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ คือ ข้อสอบที่ได้จากการวิเคราะห์เมื่อตัดข้อสอบออกทีละข้อแล้วคำนวณค่าความเชื่อมั่นของข้อสอบที่เหลือ จนกระทั่งค่าความเชื่อมั่นเริ่มลดลง จึงหยุดตัดข้อสอบ จำนวนข้อสอบที่เหลือจะเป็นข้อสอบที่คัดเลือกได้ตามเกณฑ์

3.3 วิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ คือ ข้อสอบที่ได้จากการวิเคราะห์ข้อสอบโดยตัดข้อสอบออกทีละข้อแล้วคำนวณค่าความเชื่อมั่นจากดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ จนกระทั่งค่าความเชื่อมั่นเริ่มลดลง จึงหยุดตัดข้อสอบ จำนวนข้อสอบที่เหลือจะเป็นข้อสอบที่คัดเลือกได้ตามเกณฑ์

4. แบบทดสอบความถนัดด้านคณิตศาสตร์ หมายถึง แบบทดสอบวัดความสามารถในด้านสมรรถภาพทางสมองที่เกี่ยวกับความคิดและวิธีการทางคณิตศาสตร์ ซึ่งอาศัยสมรรถภาพพื้นฐานทางสมองที่เกี่ยวข้องกับการเรียนคณิตศาสตร์ ซึ่งแบ่งเป็น 4 ประเภท(ตามแนวทฤษฎีของเทอร์สโตน) ประเภทละ 25 ข้อ คือ

4.1 แบบทดสอบประเภททักษะ (Skill) เป็นคำถามประเภทการคำนวณตัวเลขขั้นพื้นฐานทั่ว ๆ ไป

4.2 แบบทดสอบประเภทความคิดรวบยอด (Concept) เป็นคำถามประเภทความเข้าใจเกี่ยวกับหลักการ วิธีการ รวมทั้งการแปลความ ตีความ ขยายความ

4.3 แบบทดสอบประเภทโจทย์ปัญหา (Problem Solving) เป็นคำถามประเภทโจทย์ปัญหาทั่วไป

4.4 แบบทดสอบประเภทเรียงอันดับ (N-series) เป็นคำถามประเภทให้หาคำตอบที่อยู่ถัดไป โดยพิจารณาจากที่โจทย์กำหนดมาให้

5. ผู้เชี่ยวชาญ หมายถึง อาจารย์ผู้สอนวิชาวัดผลการศึกษา หรือผู้ที่มีคุณวุฒิตั้งแต่ปริญญาโททางการวัดผลการศึกษาขึ้นไป จำนวน 5 คน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การศึกษาครั้งนี้ ผู้วิจัยได้รวบรวมรายละเอียดต่าง ๆ ซึ่งประกอบด้วย ความเชื่อมั่นของแบบทดสอบ ทฤษฎีความเชื่อมั่น การตีความหมายสัมประสิทธิ์ความเชื่อมั่นแบบต่าง ๆ วิธีประมาณค่าความเชื่อมั่น องค์ประกอบที่มีอิทธิพลต่อความเชื่อมั่น คุณภาพของแบบทดสอบ การวิเคราะห์ข้อสอบ เทคนิคการวิเคราะห์ข้อสอบ ความหมายของความถนัด ทฤษฎีสมรรถภาพสมองและความถนัด ประโยชน์ของแบบทดสอบความถนัด รวมทั้งงานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้า

ความเชื่อมั่นของแบบทดสอบ

นักการศึกษาและนักจิตวิทยาได้ให้ความหมายของ “ความเชื่อมั่น” ดังนี้

นันทาสลี (Nunnally, 1964 : 59) กล่าวว่า ความเชื่อมั่นเป็นส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่ได้จากการทดสอบ

ลอร์ด และโนวิก (Lord and Novick, 1967 : 46) ให้ความหมายว่าเป็นความคงที่ของคะแนนที่ได้จากการตอบแบบทดสอบซ้ำ และคะแนนที่ได้จากการตอบแบบทดสอบทั้งสองครั้งเป็นอิสระไม่ขึ้นกับความคลาดเคลื่อนของการวัดใด ๆ

อนาสตาซี (Anastasi, 1968 : 105) ให้ความหมายว่า เป็นความคงที่ของคะแนนที่ได้จากการสอบบุคคลเดียวกันแต่ต่างเวลาและต่างโอกาสกัน

กรอนลันด์ (Gronlund, 1976 : 105) ให้ความหมายว่า เป็นความคงที่ของคะแนนในการทดสอบหรือความคงที่จากการประเมินครั้งแรกและครั้งอื่น ๆ

บุญเชิด ภิญโญนนตพงษ์ (2521 : 269) กล่าวว่า ความเชื่อมั่นหมายถึงความคงที่แน่นอนของคะแนนซึ่งได้จากการวัดนักเรียนกลุ่มเดียวกันด้วยแบบทดสอบฉบับเดียวกันหลาย ๆ ครั้ง หรือด้วยแบบทดสอบสองฉบับที่มีลักษณะเสมอเหมือนกัน หรือภายใต้เงื่อนไขของตัวแปรอื่น ๆ ในการวัดนั้น

อนันต์ ศรีโสภณ (2525 : 61) กล่าวว่า ความเชื่อมั่น หมายถึง อัตราส่วนระหว่างความแปรปรวนของคะแนนจริง และความแปรปรวนของคะแนนผลการสอบ

สำเร็จ บุญเรืองรัตน์ (2529 : 35) กล่าวว่า ความเชื่อมั่นของแบบทดสอบ หมายถึง ความสามารถของเครื่องมือที่สามารถให้ผลการวัดที่ถูกต้องแน่นอน (Accuracy) คงเส้นคงวา (Consistency) เป็นที่เชื่อถือหรือมั่นใจในผลที่ได้ออกมาอย่างแท้จริง แม้จะมีการวัดซ้ำอีก ผลที่ได้ย่อมไม่เปลี่ยนแปลงไปจากเดิม

จากความหมายของความเชื่อมั่นที่ได้กล่าวมาแล้วข้างต้น สามารถสรุปได้ว่าความเชื่อมั่นเป็นความคงที่ของคะแนนในการสอบจากผู้สอบกลุ่มเดียวกันด้วยแบบทดสอบฉบับเดียวกันหลาย ๆ ครั้ง

ทฤษฎีความเชื่อมั่น (Theory of Reliability)

ความเชื่อมั่นในทฤษฎีมาตรฐานเดิม (Reliability in Classical Theory) กล่าวว่า การวัดที่เชื่อถือได้ต้องให้ผลที่แม่นยำ แต่ผลจากการวัดล้วนแต่ขาดความแม่นยำเสมอไม่มากก็น้อยซึ่งเป็นไปตามข้อ

สมมติฐานของสเปียร์แมน (Spearmen, 1910) ที่ว่า คะแนนที่ได้จากการสอบ (Observed Scores) ประกอบด้วย คะแนนจริง (True Scores) และคะแนนความคลาดเคลื่อน (Error Scores) ดังสมการ

$$X = T + E$$

เมื่อ

$$\begin{aligned} X &= \text{คะแนนที่ได้จากการทดสอบ} \\ T &= \text{คะแนนจริง} \\ E &= \text{คะแนนความคลาดเคลื่อน} \end{aligned}$$

คะแนนความคลาดเคลื่อน คือ คะแนนที่เกิดจากความคลาดเคลื่อนในการวัด แบ่งออกเป็น 2 ประเภท คือ ความคลาดเคลื่อนที่เกิดจากการสุ่ม (Random Error) กับความคลาดเคลื่อนอย่างมีระบบ (Systematic Error) สำหรับความคลาดเคลื่อนประเภทหลังจะไม่มีผลกระทบต่อความเชื่อมั่น ส่วนความคลาดเคลื่อนที่เกิดจากการสุ่มจะมีค่าทั้งบวกและลบจึงหักล้างกันหมด และเนื่องจากความคลาดเคลื่อนไม่มีความสัมพันธ์กับคะแนนจริง จึงเขียนสมการความแปรปรวนของแบบทดสอบได้ดังนี้

$$\sigma_x^2 = \sigma_T^2 + \sigma_E^2$$

เมื่อ

$$\begin{aligned} \sigma_x^2 &= \text{ความแปรปรวนของคะแนนสอบ} \\ \sigma_T^2 &= \text{ความแปรปรวนของคะแนนจริง} \\ \sigma_E^2 &= \text{ความแปรปรวนของคะแนนความคลาดเคลื่อน} \end{aligned}$$

จากทฤษฎีความเชื่อมั่น ($\rho_{xx'}$) คือ อัตราส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่สอบ ซึ่งเขียนเป็นสมการได้ดังนี้

$$\rho_{xx'} = \frac{\sigma_T^2}{\sigma_x^2}$$

ซึ่งสามารถเขียนได้อีกอย่างหนึ่งดังนี้

$$\rho_{xx'} = 1 - \frac{\sigma_E^2}{\sigma_x^2}$$

(Mehrens and Lehmann, 1984 : 268 - 269)

การตีความหมายสัมประสิทธิ์ความเชื่อมั่นแบบต่าง ๆ

แอลเลน และเยน (Allen and Yen, 1979 : 73 - 76) ได้เสนอการพิจารณาค่าสัมประสิทธิ์ความเชื่อมั่น ($\rho_{xx'}$) ไว้ 6 ประการดังนี้

1. $\rho_{xx'}$ คือ สหสัมพันธ์ระหว่างคะแนนสอบที่ได้จากแบบทดสอบคู่ขนานของผู้สอบแต่ละคน ถ้าคะแนนสอบของแต่ละคนได้เหมือนกัน เมื่อใช้แบบทดสอบคู่ขนานความสัมพันธ์จะเท่ากับ 1.00 แบบทดสอบ

ก็มีความเชื่อมั่นที่สมบูรณ์ ถ้าคะแนนจากแบบทดสอบคู่ขนานไม่มีความสัมพันธ์กันแบบทดสอบก็ไม่มี ความเชื่อมั่น

2. ρ_{xx} . เท่ากับ สัดส่วนของความแปรปรวนใน X ที่อธิบายโดยความสัมพันธ์เชิงเส้นตรงกับ X เป็นการตีความหมายแบบมาตรฐานสำหรับสัมประสิทธิ์สหสัมพันธ์แบบ เพียร์สัน ซึ่งกำลังสองของสหสัมพันธ์ สามารถตีความหมายได้ว่าเป็นสัดส่วนของความแปรปรวนในตัวแปรหนึ่งที่อธิบายได้ด้วยความสัมพันธ์เชิงเส้นตรงกับตัวแปรอื่นๆ ดังนั้น ρ_{xx} . เป็นสัดส่วนของความแปรปรวนในคะแนนในแบบทดสอบคู่ขนาน

3. ρ_{xx} . คือ ความแปรปรวนของคะแนนจริง (σ_T^2) หารด้วยความแปรปรวนของคะแนนสอบ (σ_X^2) ถ้าค่าความแปรปรวนของคะแนนสอบ เท่ากับความแปรปรวนของคะแนนจริงแล้ว ρ_{xx} . = 1 และ ความแปรปรวนของความคลาดเคลื่อน (σ_E^2) เท่ากับ 0 จะสรุปได้ว่า การวัดปราศจากความคลาดเคลื่อน ถ้า ความเชื่อมั่นน้อยกว่า 1 แสดงว่าการวัดมีความคลาดเคลื่อน ถ้าความแปรปรวนของคะแนนสอบเท่ากับความแปรปรวนของความคลาดเคลื่อน จะได้ ρ_{xx} . = 0 ดังนั้น ถ้าความแปรปรวนของความคลาดเคลื่อนมีน้อย ความเชื่อมั่นจะมีค่าสูง ถ้าความแปรปรวนของความคลาดเคลื่อนมีมาก ความเชื่อมั่นจะมีค่าต่ำ

4. ρ_{xx} . คือ กำลังสองสหสัมพันธ์ระหว่างคะแนนสอบกับคะแนนจริง ρ_{XT}^2 เช่น ถ้าความเชื่อมั่นมีค่าเท่ากับ .64 แล้วความสัมพันธ์ระหว่างคะแนนที่สอบกับคะแนนจริงมีค่าเท่ากับ .80 ซึ่งสรุปได้ว่า

$$\rho_{xx} < \rho_{XT}$$

5. ρ_{xx} . = 1 - ρ_{XE}^2 หมายถึง สัมประสิทธิ์แห่งความเชื่อมั่นมีค่าเท่ากับ 1 ลบด้วยกำลังสอง ของความสัมพันธ์ระหว่างคะแนนสอบกับคะแนนความคลาดเคลื่อน ถ้า $\rho_{XE} = 0$ จะได้ ρ_{xx} . = 1

6. ρ_{xx} . = 1 - $\frac{\sigma_E^2}{\sigma_X^2}$ หมายถึง ค่าความเชื่อมั่นมีค่าเท่ากับ 1 ลบด้วยอัตราส่วนระหว่างความแปรปรวนคลาดเคลื่อนกับความแปรปรวนของคะแนนสอบ

$$\text{ถ้า } \rho_{xx} = 1 \text{ แล้ว } \rho_E^2 = 0$$

$$\rho_{xx} < \rho_{XT}$$

วิธีประมาณค่าความเชื่อมั่น

เฟอร์กูสัน (Ferguson. 1966 : 365 - .366) ได้กล่าวว่าวิธีหาค่าสัมประสิทธิ์ความเชื่อมั่น 4 วิธีดังนี้

1. วิธีสอบซ้ำ (Test - Retest Method) หรือเรียกว่าสัมประสิทธิ์ของความคงที่ (Coefficient of Stability) เป็นการนำแบบทดสอบฉบับเดียวกันไปทดสอบกับคน ๆ เดียวหรือคนกลุ่มเดียวกันซ้ำสองครั้งใน ช่วงเวลาที่แตกต่างกันพอสมควร คะแนนที่ได้จากการทดสอบทั้งสองครั้งนี้มีความสัมพันธ์กัน ค่าสหสัมพันธ์ที่ได้จะเป็นค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ

2. วิธีใช้แบบทดสอบคู่ขนาน (Parallel - Forms Method) เป็นการนำแบบทดสอบที่มีลักษณะคู่ขนานกัน หรือเท่าเทียมกัน โดยมีเนื้อหา ค่าเฉลี่ย และความแปรปรวนเท่ากันไปทดสอบในเวลาเดียวกัน หรือ ในเวลาที่แตกต่างกัน คะแนนที่ได้จากแบบทดสอบทั้งสองฉบับมีสหสัมพันธ์และค่าสหสัมพันธ์ที่ได้คือค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ

3. วิธีแบ่งครึ่งแบบทดสอบ (Split – Half Method) เป็นการนำแบบทดสอบฉบับเดียวไปทดสอบกับบุคคลกลุ่มเดียว แล้วแบ่งครึ่งแบบทดสอบเป็นชุดคะแนนของข้อดีและข้อคู่ นำคะแนนทั้ง 2 ชุดไปหาความสัมพันธ์ จากนั้นปรับขยายด้วยสูตรของสเปียร์แมน - บราวน์ ค่าที่ได้คือ ค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบทั้งฉบับ

4. วิธีวัดความสอดคล้องภายในของแบบทดสอบ (Internal – Consistency Method) เป็นการนำแบบทดสอบฉบับเดียวไปทดสอบกับบุคคลกลุ่มหนึ่งและนำไปหาสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบโดยวิธีของ Kuder - Richardson

แอลเลน และเยน (Allen and Yen . 1979 : 88) กล่าวว่า วิธีประมาณค่าความเชื่อมั่นที่แตกต่างกันจะได้ค่าความเชื่อมั่นที่ต่างกัน สำหรับแบบทดสอบที่อาศัยความเร็ว ควรใช้แบบสอบซ้ำ หรือแบบทดสอบคู่ขนาน เพราะการวัดความสอดคล้องภายในอาจทำให้ค่ามากกว่าความจริง ส่วนการใช้สัมประสิทธิ์แอลฟา (α) และวิธีของ Kuder-Richardson จะให้ค่าที่ต่ำกว่าความเป็นจริง และใช้กับแบบทดสอบที่มีลักษณะเป็นเอกพันธ์ (Homogeneous) เท่านั้น เพราะสูตรเหล่านี้มีพื้นฐานมาจากความเป็นเอกพันธ์ของข้อสอบ ถ้าเป็นข้อสอบที่วัดคุณลักษณะที่แตกต่างกันการหาความเชื่อมั่นแบบสัมประสิทธิ์แอลฟา และคูเดอร์-ริชาร์ดสัน จะไม่เหมาะสมเพราะได้ค่าต่ำกว่าที่ควร

เมห์เรนส์ และเลห์แมนน์ (Mehrens and Lehmann. 1984 : 271 - 272) กล่าวว่า การประมาณค่าความเชื่อมั่น มีวิธีการดังนี้

1. วิธีสอบซ้ำ (Measures of Stability)
2. วิธีใช้แบบทดสอบคู่ขนาน (Measures of Equivalence)
3. วิธีใช้แบบทดสอบคู่ขนานและสอบซ้ำ (Measures of Equivalence and Stability)
4. วิธีวัดความสอดคล้องภายใน (Measures of Internal Consistency)
 - 4.1 วิธีแบ่งครึ่งข้อสอบ (Split - Half)
 - 4.2 วิธีของ Kuder-Richardson (Kuder - Richardson Estimates)
 - 4.3 วิธีสัมประสิทธิ์แอลฟา (Coefficient Alpha)
 - 4.4 วิธีวิเคราะห์ความแปรปรวนของฮอยท์ (Hoyt's Analysis of Variance)
5. ความเชื่อมั่นของผู้ให้คะแนน (Score (Judge) Reliability)

ลวัน สายยศ (2519 : 78-79) กล่าวว่า การหาสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบใด ต้องดูข้อตกลงเบื้องต้นของแต่ละวิธีเสียก่อน พร้อมทั้งกล่าวถึงข้อตกลงเบื้องต้นสำหรับการหาสัมประสิทธิ์ความเชื่อมั่นแต่ละวิธีดังนี้

1. การหาสัมประสิทธิ์ความเชื่อมั่นแบบสอบซ้ำมีข้อตกลงว่า พฤติกรรมที่วัดจะต้องคงที่ นั่นคือในช่วงเวลาที่เว้นก่อนสอบซ้ำไม่มีผลทำให้พฤติกรรมเปลี่ยนแปลง ดังนั้นแบบทดสอบที่ใช้วัดพฤติกรรมบางอย่างที่เปลี่ยนแปลงเร็ว เช่น เจตคติ ความสนใจ ไม่นิยมใช้วิธีนี้
2. การหาสัมประสิทธิ์ความเชื่อมั่นโดยใช้แบบทดสอบคู่ขนาน แบบนี้จำเป็นต้องสร้างแบบทดสอบให้คู่ขนานกัน แบบทดสอบทั้งสองฉบับที่คู่ขนานกันนี้ มีเนื้อหาเหมือนกันคะแนนเฉลี่ยเท่ากันความแปรปรวนและความยากของแบบทดสอบเท่ากัน แต่การสร้างเครื่องมือให้มีคุณสมบัติเหมือนกันไม่ใช่ของง่าย ต้องมีเวลาและเงินมากจึงจะทำได้
3. การหาสัมประสิทธิ์ความเชื่อมั่นแบบแบ่งครึ่งแบบทดสอบมีข้อตกลงว่า เมื่อแบ่งครึ่งแบบ

ทดสอบแล้ว แบบทดสอบทั้งสองฉบับนั้นจะต้องมีคุณสมบัติเหมือนแบบทดสอบคู่ขนานทุกประการ แต่โดยทั่วไปมักจะใช้แบบทดสอบเป็นข้อคู่และข้อคู่

4. การหาสัมประสิทธิ์ความเชื่อมั่นแบบคูเดอร์ - ริชาร์ดสัน มีข้อตกลงว่าจะแนบให้แต่ละข้อมีลักษณะเป็น 0 - 1 และถ้าใช้สูตร KR-21 ความยากง่ายของข้อสอบแต่ละข้อในแบบทดสอบนี้ต้องเท่ากัน

บุญเชิด ภิญญอนันตพงษ์ (2521 : 277 - 291) ได้กล่าวถึงวิธีประมาณค่าความเชื่อมั่นของแบบทดสอบไว้ 2 แนวทางคือ

1. แบบสัมประสิทธิ์ของความคงตัว (Coefficient of Stability) เป็นการคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนสองชุด ซึ่งได้จากแบบทดสอบคนละฉบับกันหรือฉบับเดียวกัน แต่เป็นการสอบต่างเวลากัน ซึ่งจำแนกเป็น 2 วิธีคือ

1.1 วิธีสอบซ้ำ (Test-Retest Method) เป็นวิธีที่หาสัมประสิทธิ์ของความคงตัว (Coefficient of Stability) ของคะแนนที่ได้จากการสอบวัดนักเรียนกลุ่มเดียวกันสองครั้ง โดยทิ้งช่วงห่างของเวลาการสอบพอประมาณ โดยทั่วไปถ้าเป็นแบบทดสอบวัดสมรรถนะทางการเรียนจะเว้นระยะห่างกันประมาณหนึ่งเดือน ซึ่งถือเป็นระยะเวลาที่นักเรียนได้ลืมข้อคำถามที่ผ่านการสอบครั้งแรกไปแล้ว และนักเรียนคงจะเกิดพัฒนาการขึ้นจากเดิมไม่มากนัก แล้วนำคะแนนที่สอบวัดได้ในแต่ละครั้งไปคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์โดยใช้ Product Moment Correlation ของ Pearson

1.2 วิธีคู่ขนาน (Parallel - Forms Method) เป็นการคำนวณหาสัมประสิทธิ์ของความเสมอเหมือนกัน (Coefficient of Equivalence) ของคะแนนแบบทดสอบตั้งแต่สองฉบับ การประมาณค่าวิธีนี้อาศัยแนวคิดที่ว่า แบบทดสอบต่าง ๆ จะมีข้อคำถามซึ่งถือเป็นส่วนหนึ่งและเป็นตัวแทนของคุณลักษณะที่จะวัดทั้งหมด การหาค่าความเชื่อมั่นโดยวิธีนี้จึงใช้วิธีการสร้างแบบทดสอบขึ้นมาสองฉบับโดยให้แต่ละฉบับเป็นตัวแทนของลักษณะที่ต้องการวัด แล้วนำแบบทดสอบทั้งสองฉบับไปสอบกับนักเรียนกลุ่มเดียวกันในเวลาพร้อม ๆ กันแล้วคำนวณสัมประสิทธิ์สหสัมพันธ์ของคะแนนทั้งสองฉบับโดยใช้สูตรของเพียร์สัน

2. แบบสัมประสิทธิ์ของความสอดคล้องภายใน (Coefficient of Internal Consistency) แนวคิดของวิธีนี้ กำหนดว่าแบบทดสอบที่ดีจะต้องมีเอกภาพภายในการวัด (Functional Unity) กล่าวคือ ส่วนย่อยของแบบทดสอบหนึ่ง ๆ จะต้องมีความสัมพันธ์ซึ่งกันและกัน เพื่อให้เกิดความเป็นเอกพันธ์ในการที่จะวัดคุณลักษณะหนึ่ง ๆ สัมประสิทธิ์ของความสอดคล้องภายในนี้หมายความว่าข้อคำถามแต่ละข้อหรือส่วนย่อยของแบบทดสอบแต่ละส่วนจะมีความเสมอเหมือนกันทุกข้อหรือทุกส่วน แบบนี้จึงจำแนกเป็น 2 วิธีคือ

2.1 วิธีแบ่งครึ่งแบบทดสอบ (Split - Half Method) วิธีนี้นำแบบทดสอบไปสอบกับนักเรียนกลุ่มหนึ่ง แล้วตรวจให้คะแนน นำคะแนนที่ได้นี้มาวิเคราะห์โดยแบ่งคะแนนออกเป็นสองส่วนเท่า ๆ กัน โดยให้ทั้งสองส่วนมีข้อคำถามที่ถามในเนื้อหาที่คล้ายคลึงกัน และความยากง่ายของแต่ละข้อคำถามของสองส่วนมีความยากง่ายเท่าเทียมกัน วิธีที่ดีที่สุดคือ จัดแบ่งเป็นข้อคู่กับข้อคู่ แต่แบบทดสอบที่ประมาณค่าความเชื่อมั่นโดยวิธีนี้จะต้องมีจำนวนข้อไม่น้อยกว่า 60 ข้อ นำคะแนนสองส่วนไปวิเคราะห์โดยอาศัยสูตร Spearman-Brown , Horst , Rulon

2.2 วิธีวิเคราะห์ส่วนย่อย เนื่องจากวิธีแบ่งครึ่งแบบทดสอบไม่สามารถคำนวณค่าความสอดคล้องภายในได้อย่างแท้จริง เพราะการแบ่งครึ่งแบบทดสอบเพียงสองส่วนนั้น ลักษณะของความเชื่อมั่นจะเป็นความเสมอเหมือนกันระหว่างคะแนนข้อคู่กับข้อคู่มากกว่า ดังนั้นจึงได้มีผู้คิดวิธีวิเคราะห์คะแนนแบบทดสอบจากส่วนย่อย ๆ โดยจำแนกเป็นรายข้อ ซึ่งจะได้ค่าความเชื่อมั่น ซึ่งจะเป็นค่าสอดคล้องภายในของแบบทดสอบที่แท้จริง วิธีนี้จะใช้แบบทดสอบฉบับเดียวไปทดสอบกับนักเรียนกลุ่มเดียว แล้วนำคะแนน

ที่ได้จากการทดสอบมาคำนวณค่าความเชื่อมั่นด้วยวิธีต่าง ๆ เช่น วิธีของคูเดอร์ - ริชาร์ดสัน , วิธีสัมประสิทธิ์แอลฟา , วิธีวิเคราะห์ความแปรปรวนของฮอยท์ เป็นต้น

สุนีย์ เหมประสิทธิ์ (2536 : 397 - 404) ได้กล่าวถึงวิธีการประมาณค่าความเชื่อมั่นมีวิธีหลักดังนี้

1. วิธีทดสอบซ้ำ (Test - Retest Methods) เป็นวิธีที่ใช้เครื่องมือหรือแบบทดสอบฉบับหนึ่งทดสอบคนกลุ่มเดียว 2 ครั้ง โดยทิ้งช่วงระยะเวลาหนึ่ง นำคะแนนที่ได้จากการทดสอบ 2 ครั้งไปคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่าย

2. วิธีใช้แบบทดสอบสลับฟอร์ม (Alternate Forms) เป็นวิธีใช้แบบทดสอบที่มีลักษณะคล้ายคลึงหรือเท่าเทียมกัน (Equivalent) หรือคู่ขนาน (Parallel) นั่นก็คือมีคุณลักษณะที่วัดพฤติกรรมหรือคุณลักษณะเดียวกันไปทดสอบกลุ่มเดียวกัน 2 ครั้ง ด้วยแบบทดสอบคนละฉบับที่มีลักษณะคู่ขนานกัน โดยเว้นระยะห่างกัน 1 สัปดาห์ ในการทดสอบ 2 ครั้ง จากนั้นนำข้อมูลหรือคะแนนที่ได้จากการทดสอบ 2 ครั้งมาหาสัมประสิทธิ์สหสัมพันธ์อย่างง่าย วิธีนี้มีข้อเด่น คือ ขจัดปัญหาเรื่องการจำแบบทดสอบ หากใช้แบบทดสอบฉบับเดียวไปทดสอบ 2 ครั้ง ดังวิธีที่ 1 แต่อย่างไรก็ตามปัญหาของการใช้วิธีนี้ยังมีคือหาแบบทดสอบที่มีลักษณะเท่าเทียมกันยาก

3. วิธีวัดความสอดคล้องภายใน (Internal Consistency) เป็นวิธีที่ใช้แบบทดสอบครั้งเดียวกับคนกลุ่มหนึ่งโดยมีแนวคิดพื้นฐานว่าถ้าแบบทดสอบฉบับหนึ่ง ๆ วัดมโนคติเดียวหรือมโนคติที่มีลักษณะเอกพันธ์ (A Single Basic or Homogeneous Concept) ก็มีความสมเหตุสมผลที่ผู้ตอบถูกข้อหนึ่ง ๆ ควรจะตอบข้ออื่นๆ ถูกด้วย หรือกล่าวอย่างง่าย ๆ ก็คือข้อคำถามทั้งหลายควรจะมีความสัมพันธ์ในเชิงบวกต่อกัน และแบบทดสอบควรจะมีความสอดคล้องภายใน วิธีการประมาณค่าความสอดคล้องภายในมีหลายวิธีได้แก่

3.1 วิธีแบ่งครึ่ง (Split - Half Methods) คือการแบ่งแบบทดสอบออกเป็น 2 ส่วนเท่า ๆ กัน โดยมีแต่ละส่วนมีลักษณะเท่าเทียมกันและสัมพันธ์กัน วิธีนี้เหมาะสมกับแบบทดสอบที่มีความยาก (Difficulty) กระจายกันทั่ว ในทางปฏิบัตินิยมแบ่งครึ่งแบบทดสอบตามข้อคู่-ข้อคี่ ดังนั้น ความเชื่อมั่นวิธีนี้จึงเรียกว่า ความเชื่อมั่นแบบข้อคู่ - ข้อคี่ (Odd-Even Reliability) เป็นความเชื่อมั่นแบบวัดเพียงครึ่งฉบับ โดยหาสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบครึ่งฉบับทั้งสองส่วน ฉะนั้นค่าความเชื่อมั่นจึงมีค่าประมาณต่ำกว่าค่าที่ควรจะเป็น ดังนั้น Spearman & Brown จึงคิดค่าปรับแก้ค่าความเชื่อมั่น วิธีนี้มีข้อเด่น คือ ใช้แบบทดสอบเพียงครั้งเดียว ดังนั้นจึงขจัดปัญหาในเรื่องของอิทธิพลเนื่องจากความจำและการฝึกฝน ซึ่งเป็นจุดอ่อนของวิธีประมาณค่าความเชื่อมั่นวิธีที่ 1 และ 2 ดังนั้นจึงนิยมใช้หาความเชื่อมั่นของแบบทดสอบในชั้นเรียน จุดอ่อนของวิธีนี้ก็คือการแบ่งแบบทดสอบเป็น 2 ส่วน ซึ่งต้องมีลักษณะเท่าเทียมกัน

3.2 วิธีของคูเดอร์และริชาร์ดสัน (Kuder - Richardson Methods) มีแนวคิดพื้นฐานว่าข้อคำถามของแบบทดสอบต่างสอดคล้องกันคือวัดมโนคติเดียว โดยปกติวิธีนี้จะให้ค่าประมาณความเชื่อมั่นต่ำกว่าวิธีแบ่งครึ่ง แต่ให้ค่าสูงกว่าวิธีทดสอบซ้ำ และวิธีใช้แบบทดสอบสลับฟอร์ม

3.3 วิธีของครอนบาคหรือสัมประสิทธิ์แอลฟา (Cronbach Method or Alpha Coefficient) วิธีนี้มีประโยชน์มาก คือ สามารถใช้ได้กับแบบทดสอบที่มีการให้คะแนนเช่นใดก็ได้คืออาจเป็นการให้คะแนน 0, 1 หรืออาจเป็นการให้คะแนนจากแบบประเมินค่า (Rating Scale) หรือแบบทดสอบอัตนัย (Essay Test) หากการให้คะแนนเป็น 0, 1 ค่าความเชื่อมั่นที่ได้จากวิธีนี้จะมีค่าเท่ากับที่ใช้ KR20 ดังนั้นวิธีนี้จึงเหมาะที่จะใช้กับแบบวัดมโนคติเดียว (A Single , Unified Concept)

องค์ประกอบที่มีอิทธิพลต่อค่าความเชื่อมั่น (Factors affecting reliability)

ความเชื่อมั่นมีค่าเปลี่ยนแปลงขึ้นอยู่กับองค์ประกอบหลายประการ ดังเห็นได้จากการศึกษาค้นคว้าดังนี้

เมห์เรนส์และเลห์แมน (สุทิน บุญญาภิบาล, 2537:17; อ้างอิงมาจาก Mehrens and Lehmann, 1984 : 278 - 281) ได้กล่าวไว้ว่าองค์ประกอบที่มีผลต่อความเชื่อมั่นมีดังนี้

1. ความยาวของแบบทดสอบ เมื่อแบบทดสอบมีความยาวเพิ่มขึ้น จะทำให้ข้อสอบมีความเชื่อมั่นสูงขึ้น นั่นคือ แบบทดสอบที่มีจำนวนข้อมากจะมีความเชื่อมั่นสูงกว่าข้อสอบที่มีจำนวนข้อน้อย
 2. เวลาในการทำแบบทดสอบ ถ้าเวลาทำแบบทดสอบไม่เหมาะสม เช่น เมื่อให้เวลาน้อยเกินไป ผู้สอบทำไม่ทันก็อาจเดาคำตอบ ซึ่งจะทำให้การกระจายของคะแนนแตกต่างจากแบบทดสอบที่ให้เวลาเพียงพอ แต่ถ้าให้เวลามากเกินไปจะทำให้คนเรียนเก่งและเรียนอ่อนได้คะแนนเท่า ๆ กัน การกระจายของคะแนนมีน้อยจะมีผลต่อความเชื่อมั่นของแบบทดสอบเช่นกัน
 3. ความเป็นเอกพันธ์ของกลุ่มตัวอย่างที่มีความสามารถใกล้เคียงกัน (เป็นเอกพันธ์กัน) จะทำให้มีการกระจายของคะแนนน้อยส่วนกลุ่มที่มีความสามารถแตกต่างกัน (เป็นวิวิธพันธ์) จะมีช่วงการกระจายของคะแนนมากกว่า ซึ่งจะมีผลทำให้ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มที่มีความสามารถเป็นเอกพันธ์ต่ำกว่ากลุ่มที่มีความสามารถเป็นวิวิธพันธ์
 4. ความยากของแบบทดสอบ ข้อสอบที่ยากเกินไปหรือง่ายเกินไป จะทำให้ค่าความเชื่อมั่นเปลี่ยนแปลงไป ทั้งนี้ เนื่องจากค่าความเชื่อมั่นของแบบทดสอบนั้น ประเมินจากความแปรปรวนของคะแนนผลการสอบ ถ้าแบบทดสอบยากเกินไปเด็กทุกคนทำผิดหมดหรือถ้าแบบทดสอบง่ายเกินไป จะทำให้ทุกคนทำถูกหมดความแปรปรวนของคะแนนมีน้อยค่าความเชื่อมั่นลดลงด้วย และถ้าความยากของแบบทดสอบทำให้คะแนนกระจายมากจะทำให้ค่าความเชื่อมั่นของแบบทดสอบสูงขึ้น
 5. ความเป็นปรนัย แบบทดสอบที่มีความเป็นปรนัยสูงจะทำให้ความเชื่อมั่นของแบบทดสอบสูงขึ้น โดยเฉพาะอย่างยิ่งความเป็นปรนัยในการให้คะแนน ซึ่งถ้าแบบทดสอบให้คะแนนไม่เป็นปรนัยค่าความเชื่อมั่นก็จะไม่แน่นอน
 6. วิธีการคิดค่าความเชื่อมั่น การหาค่าความเชื่อมั่นมีหลายวิธี แต่ละวิธีให้ผลต่างกัน และเหมาะสมสำหรับจุดมุ่งหมายเฉพาะอย่าง
- บุญเชิด ภิญโญนันตพงษ์ (2521 : 312 - 317) การประมาณค่าความเชื่อมั่นแต่ละวิธีจะมีค่าสูงหรือไม่นั้นขึ้นอยู่กับแหล่งความคลาดเคลื่อน ซึ่งทำให้ค่าความเชื่อมั่นเปลี่ยนแปลงไป ดังนั้นในการคำนวณหาค่าความเชื่อมั่นจึงต้องคำนึงถึงความเหมาะสมของแต่ละวิธีด้วย นอกจากนั้นยังต้องพิจารณาองค์ประกอบอื่น ๆ ที่มีอิทธิพลต่อความเชื่อมั่นดังต่อไปนี้
1. ความสามารถของกลุ่มตัวอย่าง ถ้ากลุ่มตัวอย่างเป็นเอกพันธ์ (Homogeneous Group) จะได้ค่าความเชื่อมั่นต่ำกว่ากลุ่มตัวอย่างที่เป็นวิวิธพันธ์ (Heterogeneous Group)
 2. ระดับความสามารถของนักเรียนในกลุ่ม สิ่งที่มีอิทธิพลต่อค่าความเชื่อมั่นที่เกิดจากตัวนักเรียนเองอีกแบบหนึ่ง คือ ระดับความสามารถของนักเรียน ถ้านักเรียนในกลุ่มมีความสามารถเฉลี่ยสูงจะสามารถทำข้อสอบได้ถูกต้องแน่นอนทุกครั้ง แต่ถ้านักเรียนในกลุ่มมีความสามารถเฉลี่ยต่ำก็มักจะตอบโดยการเดาเป็นส่วนใหญ่ จึงทำให้คะแนนจากการสอบไม่แน่นอน ซึ่งทำให้ค่าความเชื่อมั่นต่ำลง
 3. ระดับความยากของแบบทดสอบ ความยากของแบบทดสอบมีอิทธิพลต่อค่าความเชื่อมั่นในแง่ที่ทำให้การกระจายของคะแนนมีมากน้อยต่างกัน การกระจายของคะแนนมากจะทำให้ค่าความเชื่อมั่นสูง

กว่าคะแนนที่มีการกระจายน้อย เนื่องจากข้อสอบที่ง่ายมาก นักเรียนส่วนใหญ่ทำคะแนนได้ คะแนนแต่ละคนก็ไม่แตกต่างกัน ทำนองเดียวกันข้อสอบที่ยากเกินไป นักเรียนส่วนใหญ่ต่างทำกันไม่ได้ คะแนนแต่ละคนก็ไม่แตกต่างกัน ซึ่งเป็นเหตุให้ค่าความเชื่อมั่นต่ำลง เพราะคะแนนมีการกระจายแคบ

4. ความยาวแบบทดสอบ จำนวนข้อของแบบทดสอบมีอิทธิพลต่อค่าความเชื่อมั่น แบบทดสอบใดมีจำนวนข้อน้อยจะมีค่าความเชื่อมั่นต่ำ ถ้ามีจำนวนมากจะมีค่าความเชื่อมั่นสูง

5. ความคล้ายคลึงของเนื้อหาที่ออกแบบทดสอบ แบบทดสอบที่วัดลักษณะเดียวกันทั้งฉบับย่อมจะมีค่าความเชื่อมั่นสูงกว่าข้อสอบที่มีเนื้อหาที่แตกต่างกันมาก ๆ

6. แบบทดสอบแบบเร่งรีบ (Speed Test) ข้อสอบประเภทนี้เป็นข้อสอบที่ง่าย ๆ แต่มีมากข้อ ผู้ตอบต้องอาศัยความรวดเร็วในการตอบคำถาม ส่วนมากแล้วนักเรียนสามารถตอบถูกทุกข้อที่ทำทัน หมายความว่า ทำถึงข้อใดก็มักจะได้คะแนนเท่านั้นเสมอ ดังนั้น การสอบแต่ละครั้งจึงมีคะแนนคงเดิมเสมอซึ่งทำให้ข้อสอบประเภทนี้มีความเชื่อมั่นสูง

7. ตัวอย่างประชากรที่ใช้ทดลอง จะมีผลกระทบต่อค่าความเชื่อมั่น ถ้าตัวอย่างที่นำมาทดลองมีจำนวนน้อยเกินไปหรือไม่เป็นตัวแทนของประชากรในสิ่งที่จะวัด นอกจากนั้นความไม่คุ้นเคยกับแบบทดสอบ อารมณ์ การเจ็บป่วย ความวิตกกังวล สิ่งเหล่านี้มีผลกระทบต่อความเชื่อมั่นทั้งสิ้น

8. ความเป็นปรนัย ข้อสอบใดมีความแม่นยำในการให้คะแนนมาก ข้อสอบนั้นก็จะมีค่าความเชื่อมั่นสูง ข้อสอบอัตนัยจะขาดความเป็นปรนัยในการให้คะแนนดังนั้นจึงมีความเชื่อมั่นต่ำกว่าข้อสอบปรนัยชนิดต่าง ๆ

คุณภาพของแบบทดสอบ

แบบทดสอบที่ใช้จะมีคุณภาพดีเพียงใด เหมาะที่จะนำมาใช้งานได้หรือไม่ จำเป็นที่จะต้องมีการตรวจสอบโดยการนำเกณฑ์ต่าง ๆ มาเป็นหลักในการพิจารณา

บุญเชิด ภิญโญนนตพงษ์ (2519 : 140 - 149) ได้กล่าวถึงเกณฑ์ที่ใช้ในการตรวจสอบคุณภาพของแบบทดสอบไว้ดังนี้

1. การวิเคราะห์รายข้อ (Item Analysis) คือ เทคนิคสำหรับตรวจหาคุณภาพแบบทดสอบเป็นรายข้อว่าแต่ละข้อมี ระดับความยาก (Difficulty) เท่าใด ค่าอำนาจจำแนก (Discrimination) เท่าใด ระดับความยากง่ายมาตรฐานเท่าใด

2. การวิเคราะห์ทั้งฉบับ (Test Analysis) เป็นการตรวจหาคุณภาพแบบทดสอบทั้งฉบับว่ามีความเชื่อมั่น (Reliability) และความเที่ยงตรง (Validity) มากน้อยเพียงใด แบบทดสอบใดมีค่าอำนาจจำแนกรายข้อสูงเป็นส่วนมากแล้ว คุณภาพทั้งฉบับก็จะมีแนวโน้มสูงตามด้วย

การวิเคราะห์ข้อสอบ

นักการศึกษาได้ให้ความหมายของ “การวิเคราะห์ข้อสอบ” ดังนี้

สุภาพ วาดเขียน และอรพินธ์ โกชนดา (2518 : 63) กล่าวว่า การวิเคราะห์ข้อสอบ เป็นการพิจารณาถึงระดับความยากง่ายของข้อสอบที่เหมาะสมกับการนำมาใช้ทดสอบ และเป็นการพิจารณาถึงอำนาจในการจำแนกเด็กเก่งและไม่เก่งออกเป็น ส่วน ๆ ตามต้องการ

บุญเชิด ภิญญอนันตพงษ์ (2519 : 140) กล่าวว่า การวิเคราะห์ข้อสอบ เป็นขบวนการในการตรวจหาคุณภาพข้อสอบว่าข้อสอบแต่ละข้อมีคุณภาพใช้ได้หรือไม่ได้ และตรวจหาคุณภาพทั้งฉบับว่ามีความเที่ยงตรง และเชื่อมั่นเพียงใด

บุญชม ศรีสะอาด (2520 : 116) กล่าวว่า การวิเคราะห์ข้อสอบ คือเทคนิคสำหรับตรวจคุณภาพของข้อสอบเป็นรายข้อ เพื่อทราบระดับความยาก และอำนาจจำแนกของแต่ละข้อนั้นเป็นสำคัญ

อุทุมพร ทองอุไทย (2520 : 34 - 35) กล่าวว่า การวิเคราะห์ข้อสอบเป็นการวิเคราะห์ว่าข้อสอบข้อใดยาก ข้อใดง่าย ข้อใดจำแนกคนเก่งออกจากกลุ่มได้ และข้อใดจำแนกไม่ได้ การวิเคราะห์ข้อสอบประกอบด้วย ความยาก (Difficulty) และความจำแนก (Discrimination) ดรรชนีที่บอกความยากของข้อสอบรายข้อ คือ จำนวนที่คนตอบข้อนั้น ๆ ได้ถูกต้องน้อย ยิ่งมีคนตอบได้เพียงไม่กี่คนจะยิ่งมีความยากมาก ถ้าข้อสอบใดมีคนตอบได้มาก ก็เป็นข้อสอบที่ง่าย ยิ่งมีคนทั้งหมดตอบได้ก็เป็นข้อสอบที่ง่ายมาก ดรรชนีที่บอกความจำแนกได้แก่จำนวนข้อที่คนเก่งทำได้ ด้วยจำนวนข้อที่คนไม่เก่งทำได้

อนันต์ ศรีโสภะ (2524 : 148) กล่าวว่า การวิเคราะห์ข้อสอบเป็นกรรมวิธีของการตรวจสอบคำตอบของนักเรียนในแต่ละข้อ เพื่อพิจารณาว่าข้อสอบแต่ละข้อมีระดับความยากง่าย และค่าอำนาจการจำแนกความสามารถของนักเรียนได้มากน้อยเพียงใด ตลอดจนการพิจารณาประสิทธิภาพในการลงของตัวเลือกต่าง ๆ

สมบุญ ภู่นวล (2525 : 295) กล่าวว่า การวิเคราะห์ข้อสอบ หมายถึง การค้นหาคุณลักษณะคุณภาพ ของคำถามและตัวเลือก ซึ่งเป็นส่วนที่ประกอบกันขึ้นมาเป็นข้อสอบว่าเป็นอย่างไร ดีหรือไม่ ง่ายหรือยาก

กานดา พูนลาภทวี (2528 : 159) กล่าวว่า การวิเคราะห์ข้อสอบเป็นการพิจารณาคำตอบของผู้ตอบแบบทดสอบเป็นรายข้อ เพื่อตัดสินว่าข้อสอบแต่ละข้อนั้นมีความเหมาะสมมากน้อยเพียงใดมีข้อบกพร่องด้านใดบ้าง จุดมุ่งหมายของการวิเคราะห์ข้อสอบก็เพื่อให้ได้แบบทดสอบที่มีคุณภาพสูงสุด

เทคนิคการวิเคราะห์ข้อสอบ

อนันต์ ศรีโสภะ (2515:32-49) กล่าวถึงการวิเคราะห์ข้อสอบ(Item Analysis)มีหลายวิธีคือ ความยากของข้อสอบ (Item Difficulty) คือ เปอร์เซนต์ของนักเรียนที่ตอบถูก คำนวณจากสูตร

$$P = \frac{R}{T}$$

ถ้า P = ความยากของข้อสอบ
R = จำนวนนักเรียนที่ตอบถูก
T = จำนวนนักเรียนที่นำมาวิเคราะห์

การวิเคราะห์ระดับความยากของข้อสอบนั้นขึ้นอยู่กับความมุ่งหมายของการทดสอบ ในการทดสอบแบบ Achievement Test จะหาค่าอำนาจที่เชื่อถือได้ระหว่างนักเรียนในทุกระดับของความสามารถ ดังนั้นการกระจายความยากของข้อสอบจึงเป็นสิ่งที่เราต้องการ เพราะจะทำให้ทั้งนักเรียนเก่งและอ่อนได้แสดงตามความสามารถของเขา

อำนาจจำแนกของข้อสอบ (Item Discrimination Power) หมายถึง ระดับในการจำแนกที่มีผลสัมฤทธิ์ในการเรียนสูงและต่ำ การคำนวณค่าอำนาจจำแนกของข้อสอบโดยการเอาจำนวนนักเรียนใน

กลุ่มต่ำไปลบออกจากนักเรียนในกลุ่มสูงที่ตอบถูก แล้วหารด้วยครึ่งหนึ่งของจำนวนนักเรียนที่นำมาวิเคราะห์ เขียนเป็นสูตรได้ดังนี้

$$D = \frac{R_U - R_L}{T/2}$$

เมื่อ D = ค่าอำนาจการจำแนกของข้อสอบ
 R_U = จำนวนนักเรียนที่ตอบถูกในกลุ่มสูง
 R_L = จำนวนนักเรียนที่ตอบถูกในกลุ่มต่ำ
 T = จำนวนนักเรียนที่นำมาวิเคราะห์

ถ้าข้อสอบมีอำนาจการจำแนกทางบวกสูงสุด (Maximum Positive) หรือมีค่าเป็น +1 หมายความว่า นักเรียนในกลุ่มสูงทำข้อสอบถูกหมดและนักเรียนในกลุ่มต่ำทำผิดหมดทุกคน ส่วนข้อสอบที่ไม่มีอำนาจการจำแนก หมายถึง นักเรียนในกลุ่มสูงและกลุ่มต่ำทำข้อสอบถูกหมดหรือผิดหมด หรือถูกเท่ากัน ค่าอำนาจจำแนกจะเท่ากับ 0.00 สำหรับข้อสอบที่มีอำนาจการจำแนกเป็นลบ (Negative Discrimination Power) หมายถึง มีนักเรียนในกลุ่มต่ำทำข้อสอบถูกมากกว่านักเรียนในกลุ่มสูง ข้อสอบประเภทนี้ไม่มีคุณค่า ต้องแก้ไขปรับปรุง วิธีหาอำนาจการจำแนกของข้อสอบมีหลายวิธี เช่น หาสหสัมพันธ์ Biserial Correlation Coefficient, Tetrachoric Correlation Coefficient, Product Moment Correlation Coefficient, Phi Coefficient, Upper 27 Percent Method

บุญเชิด ภิญญอนันตพงษ์ (2519 : 140 - 142) กล่าวถึง เทคนิคสำหรับตรวจหาคุณภาพ ข้อสอบเป็นรายข้อว่าแต่ละข้อจะมี

1. ระดับความยาก (Difficulty) หมายถึง จำนวนเปอร์เซ็นต์ (Percentage) หรือสัดส่วน (Proportion) ของนักเรียนที่ตอบข้อคำถามนั้นถูก ซึ่งใช้อักษรย่อว่า P หรือ p ตามลำดับ ถ้าข้อคำถามใดมีจำนวนคนเลือกตอบถูกมากก็หมายความว่า คนส่วนใหญ่ตอบข้อสอบนี้ได้ถูกต้อง แสดงว่า ข้อสอบนี้ง่าย หรือค่า P มาก ข้อสอบข้อนั้นก็ง่าย ถ้าข้อคำถามใดมีจำนวนคนเลือกตอบถูกน้อยก็หมายความว่าคนส่วนใหญ่ตอบข้อสอบนี้ไม่ถูกหรือตอบผิด แสดงว่า ข้อสอบนี้ยาก หรือค่า P น้อย ๆ ข้อสอบข้อนั้นก็ยาก ถ้าค่า P น้อย ข้อสอบจะยาก และถ้าค่า P มาก ข้อสอบจะง่าย ตามทฤษฎีวัดผล ข้อสอบที่ดีควรมีค่า $P = 50$ ซึ่งหมายความว่า มีคนตอบถูกครึ่งหนึ่ง ตอบผิดครึ่งหนึ่ง แต่การที่จะสร้างให้ข้อสอบแต่ละข้อมีค่า $P = 50$ นั้นเราไม่สามารถจะทำได้ ดังนั้นในทางปฏิบัติ P อยู่ในช่วง 20- 80

2. ค่าอำนาจจำแนก (Discrimination) หมายถึง ความสามารถของข้อคำถามนั้นสามารถจำแนกเด็กออกได้เป็น 2 กลุ่ม คือ เก่ง กับอ่อนได้ ค่าอำนาจจำแนกรายข้อนี้คำนวณจากค่า Biserial Correlation (r_{bis}) เนื่องจาก r มีค่าตั้งแต่ -1 ถึง 0 หรือ 0 ถึง 1 ซึ่งแยกเป็น 3 ลักษณะ คือ r ที่เป็นลบ เป็นศูนย์ และเป็นบวก ค่า r ที่คำนวณได้นี้เป็นค่าสัมประสิทธิ์สหสัมพันธ์แบบ Biserial ซึ่งเกิดจากความสัมพันธ์ระหว่างของสองสิ่ง ว่าสัมพันธ์สอดคล้องกันเพียงใด ข้อสอบใดที่คนเก่งตอบถูกมากกว่าคนอ่อนหรือคนเก่งตอบถูกคนไม่เก่งตอบผิด ค่า r จะเป็นบวก ข้อสอบใดคนเก่งกับคนอ่อนตอบถูกเท่ากัน ค่า r จะเป็นศูนย์ ข้อสอบใดคนเก่งตอบถูกน้อยกว่าคนอ่อน หรือคนเก่งตอบผิดคนอ่อนตอบถูกแล้ว ค่า r จะเป็นลบ

3. ค่าความยากง่ายมาตรฐาน (Δ) คือ ค่าความยากที่ดัดแปลงค่า P ให้เป็นคะแนนมาตรฐาน เนื่องจากค่าความยากง่าย (P) นั้น ถ้า P มีค่าน้อยข้อสอบจะยาก ถ้า P มีค่ามากข้อสอบจะง่าย ซึ่งขัดกับความรู้สึก นักวัดผลจึงแปลงความยากนี้ให้เป็นความยากง่ายมาตรฐาน โดยมี $S = 4$ และ $\bar{X} = 13$

$$\text{ดังนั้น } \Delta = 4 Z_p + 13$$

Z_p คือ คะแนนมาตรฐานที่แปลงมาจาก P

ถ้า $P = 50$ $Z_p = 0$ จะได้ $\Delta = 13$ นั่นคือถ้าข้อสอบใดมี $\Delta = 13$ แสดงว่าข้อสอบนั้นยากพอเหมาะ และถ้า Δ มีค่ามากกว่า 13 ข้อสอบนั้นจะยาก และถ้า Δ มีค่าน้อยกว่า 13 ข้อสอบนั้นจะง่ายลงตามลำดับ

วิญญา วิชาลาภรณ์ (2521 : 113 - 119) กล่าวว่า การวิเคราะห์ข้อสอบนั้นทำได้ 2 ทาง

1. วิเคราะห์ข้อสอบก่อนที่จะทำการทดสอบ โดยการตรวจดูว่าข้อสอบแต่ละข้อนั้นได้วัดในสิ่งที่ต้องการจะวัดหรือไม่ คำสั่งชัดเจนดีแล้วหรือยัง จำนวนข้อของข้อสอบพอเหมาะกับความยากง่ายของข้อสอบหรือไม่ อาจทำได้โดยตรวจสอบความเหมาะสมของข้อสอบ

2. วิเคราะห์ข้อสอบหลังจากทดสอบแล้ว การวิเคราะห์ข้อสอบลักษณะนี้เป็นการวิเคราะห์เพื่อปรับปรุงข้อสอบให้ดีขึ้น และเพื่อนำข้อสอบไปใช้อีก การวิเคราะห์ในลักษณะนี้เป็นการวิเคราะห์เพื่อหา

2.1 ความง่ายของข้อสอบ (Item Difficulty) หมายถึง จำนวนนักเรียนที่ตอบข้อสอบถูกคำนวณโดยสูตร

$$D = \frac{R}{N}$$

D คือ ความง่ายของข้อสอบ

R คือ จำนวนคนตอบถูก (เฉพาะที่เอามาวิเคราะห์)

N คือ จำนวนนักเรียนที่เอามาวิเคราะห์ (รวมทั้งกลุ่มสูงและกลุ่มต่ำ)

2.2 ประสิทธิภาพของตัวลวง (Attractiveness of Distractors) ตัวลวงหมายถึง ตัวเลือกที่ไม่ใช่ข้อถูก ตัวลวงที่ดีคือเมื่อมองโดยผิวเผินแล้วจะเห็นว่าน่าจะเป็นคำตอบถูก ตัวลวงใดที่มีนักเรียนเลือกน้อยกว่า 5 เปอร์เซ็นต์ ควรจะต้องปรับปรุงแก้ไขใหม่

2.3 อำนาจจำแนกของข้อสอบ (Item Discriminating Power) อำนาจจำแนกของข้อสอบหมายถึง ความสามารถของข้อสอบที่จะแยกคนเก่งออกจากคนอ่อน ข้อสอบที่มีอำนาจจำแนกนั้น คนเก่งควรจะต้องถูก และคนอ่อนควรจะต้องผิด วิธีหาอำนาจจำแนกของข้อสอบโดยวิธีที่ง่ายที่สุดคือเอาจำนวนคนตอบถูกในกลุ่มสูงลบด้วยจำนวนคนตอบถูกในกลุ่มต่ำ ผลลัพธ์ที่ได้จะออกมา 3 ลักษณะ คือ

1. คำตอบเป็นศูนย์ หมายความว่า ข้อสอบไม่มีค่าอำนาจจำแนก

2. คำตอบเป็นลบ เป็นข้อสอบที่ไม่ดี เพราะคนอ่อนตอบถูกมากกว่าคนเก่ง

3. คำตอบเป็นบวก เป็นข้อสอบที่ดี เพราะคนเก่งตอบถูกมากกว่าคนไม่เก่ง

วิธีหาอำนาจจำแนกอีกวิธีหนึ่งคือ เอาจำนวนคนตอบถูกในกลุ่มสูงลบด้วยจำนวนคนตอบถูกในกลุ่มต่ำ แล้วหารด้วยจำนวนคนในกลุ่มสูงหรือกลุ่มต่ำ คือ

$$\text{อำนาจจำแนก} = \frac{R_H - R_L}{N/2}$$

เมื่อ R_H คือ จำนวนคนตอบถูกในกลุ่มสูง

R_L คือ จำนวนคนตอบถูกในกลุ่มต่ำ

$N/2$ คือ ครึ่งหนึ่งของจำนวนคนที่เอามาวิเคราะห์ ซึ่งเท่ากับจำนวนคนในกลุ่มสูงหรือกลุ่มต่ำ

ค่าอำนาจจำแนกโดยวิธีนี้จะอยู่ระหว่าง 0.00 ถึง 1.00 Ebel ได้ให้แนวคิดเกี่ยวกับอำนาจจำแนกของข้อสอบดังนี้

- ต่ำกว่า .19 มีอำนาจจำแนกน้อยข้อสอบยังใช้ไม่ได้ จำเป็นต้องเขียนข้อสอบใหม่
- .20 ถึง .29 พอใช้ได้ แต่ต้องปรับปรุงข้อสอบใหม่
- .30 ถึง .39 ค่อนข้างดี ถ้าเป็นไปได้ควรปรับปรุงให้ดีขึ้นอีก
- .40 ขึ้นไป เป็นข้อสอบที่ดี

วิธีหาค่าอำนาจจำแนกของข้อสอบสามารถหาได้โดยวิธีอื่น ๆ เช่นหาสัมประสิทธิ์สหสัมพันธ์แบบ Biserial หรือ Tetrachoric เป็นต้น

กานดา พูลลาภทวี (2528 : 160 - 161) กล่าวว่า ในการวิเคราะห์ข้อสอบต้องพิจารณาดังนี้

1. ระดับความยาก (Level of Difficulty) หรือสัดส่วนหรือร้อยละของจำนวนคนที่ตอบข้อสอบนั้นถูก คำนวณโดยสูตร

$$P = \frac{R}{N}$$

เมื่อ P คือ ระดับความยากของข้อสอบ
 R คือ จำนวนคนที่ตอบข้อสอบนั้นถูกต้อง
 N คือ จำนวนคนทั้งหมดที่ตอบข้อสอบข้อนั้น

ระดับความยากของข้อสอบมีค่าตั้งแต่ 0 ถึง 1 ข้อสอบใดที่มีค่า P สูง แสดงว่า ข้อสอบง่าย เป็นข้อสอบที่มีคนตอบถูกเป็นจำนวนมาก ข้อสอบใดที่มีค่า P ต่ำ แสดงว่า ข้อสอบยาก เป็นข้อสอบที่มีคนตอบถูกเป็นจำนวนน้อย โดยทั่วไปข้อสอบที่มีความยากตั้งแต่ .20 ถึง .80 ถือว่าเป็นข้อสอบที่อยู่ในเกณฑ์ดี

2. อำนาจจำแนก (Discrimination Power) หมายถึง ประสิทธิภาพของข้อคำถามที่จำแนกหรือแยกให้เห็นความแตกต่างระหว่างผู้ที่มีผลสัมฤทธิ์ต่างกัน เช่น จำแนกคนเก่งกับคนอ่อน จำแนกคนที่มีความสามารถพิเศษกับคนที่ไม่มีความสามารถออกจากกันได้ โดยถือว่าคนเก่งหรือคนที่มีความสามารถจะต้องทำข้อสอบข้อนั้นได้ ส่วนคนอ่อนหรือคนที่ไม่มีความสามารถจะต้องทำข้อสอบข้อนั้นไม่ได้

$$r = \frac{R_H - R_L}{N_H \text{ หรือ } N_L}$$

เมื่อ r คือ ค่าอำนาจจำแนก
 R_H คือ จำนวนคนที่ตอบถูกในกลุ่มสูง
 R_L คือ จำนวนคนที่ตอบถูกในกลุ่มต่ำ

ค่าอำนาจจำแนกของข้อสอบมีค่าตั้งแต่ -1 ถึง 1 กรณีที่ควรพิจารณามีดังนี้

ข้อสอบที่มีค่า r เป็นบวก เป็นข้อสอบที่คนเก่งมีแนวโน้มที่จะตอบถูกมาก คนอ่อนมีแนวโน้มที่จะตอบผิดมาก ยิ่ง r เป็นค่าบวกสูงมากเท่าใด คนเก่งจะมีแนวโน้มตอบถูกมากเท่านั้น ส่วนคนอ่อนก็จะมีแนวโน้มตอบผิดมากขึ้นด้วย ข้อสอบที่มีค่า r เป็นลบ เป็นข้อสอบที่คนเก่งมีแนวโน้มที่จะตอบผิดมาก คนอ่อนมีแนวโน้มที่จะตอบถูกมาก ซึ่งเป็นข้อสอบที่ไม่ดี เพราะเป็นข้อสอบที่ลวงนักเรียนเก่ง ข้อสอบที่ r มีค่าเป็นศูนย์ เป็นข้อสอบที่คนเก่งหรือคนอ่อนอาจจะตอบถูกหรือผิดก็ได้ไม่แน่นอนไม่สามารถแยกคนเก่งและคนอ่อนออกจากกันได้

อีเบล (กานดา พูลลาภทวี. 2528 : 161 ; อ้างอิงมาจาก Ebel . 1965) ได้ให้ข้อเสนอแนะเกณฑ์ที่ใช้ในการพิจารณาค่าอำนาจจำแนกของข้อสอบดังนี้

ค่าอำนาจจำแนก	คุณภาพของข้อสอบ
.40 ขึ้นไป	ดีมาก
.30 - .39	ดี แต่ควรนำไปปรับปรุง
.20 - .29	พอใช้ได้ ควรจะนำไปปรับปรุงใหม่
ต่ำกว่า .19	ไม่ดี ต้องพิจารณาปรับปรุงใหม่หรือตัดทิ้ง

แอลเลน และเยน (Allen and Yen. 1979 : 118 - 147) กล่าวถึง เทคนิคการวิเคราะห์ข้อสอบ ซึ่งแต่ละเทคนิคของการวิเคราะห์ข้อสอบจะเป็นวิธีการเพื่อเลือกข้อสอบที่ดีที่สุดสำหรับการพิมพ์ครั้งสุดท้ายของแบบทดสอบ วิธีวิเคราะห์ข้อสอบ 4 วิธีคือ

1. ดัชนีความยากและอำนาจจำแนกของข้อสอบหรือสหสัมพันธ์แบบพอยท์ไบเซรียล (Point - Biserial Correlation) ระหว่างคะแนนของข้อสอบและคะแนนรวมของแบบทดสอบ
 2. ดัชนีความเชื่อมั่นของแบบทดสอบและดัชนีความเที่ยงตรงของแบบทดสอบ
 3. โคลงลักษณะของข้อสอบ
 4. การวิเคราะห์องค์ประกอบ
1. การวิเคราะห์ข้อสอบโดยใช้ดัชนีความยากและอำนาจจำแนกของข้อสอบ ถ้าต้องการค่าอำนาจจำแนกที่ดีที่สุดระหว่างผู้สอบในทุกระดับความสามารถแล้ว ระดับความยากที่ดีที่สุดควรมีพิสัยประมาณ .3 ถึง .7 ถ้าต้องการค่าอำนาจจำแนกที่ดีที่สุด ณ คะแนนจุดตัดหนึ่ง ๆ แล้วนั้น ระดับความยากจะต้องสัมพันธ์กับคะแนนจุดตัดนี้

ดัชนีอำนาจจำแนกของข้อสอบสำหรับข้อสอบ i , d_i คำนวณได้จากสูตร

$$d_i = \frac{u_i}{n_{iu}} - \frac{L_i}{n_{il}}$$

เมื่อ u_i คือ จำนวนของผู้สอบที่ได้คะแนนรวมจากแบบทดสอบทั้งหมดอยู่ในช่วงสูงของคะแนนรวมของแบบทดสอบและทำข้อสอบ i ได้ถูกต้องด้วย

L_i คือ จำนวนของผู้สอบที่ได้คะแนนรวมจากแบบทดสอบทั้งหมดอยู่ในช่วงต่ำของคะแนนรวมของแบบทดสอบและทำข้อสอบ i ได้ถูกต้องด้วย

n_{iu} คือ จำนวนของผู้สอบที่ได้คะแนนรวมจากแบบทดสอบอยู่ในช่วงสูงของคะแนนรวมของแบบทดสอบ

n_{il} คือ จำนวนของผู้สอบที่ได้คะแนนรวมจากแบบทดสอบอยู่ในช่วงต่ำของคะแนนรวมของแบบทดสอบ

ถ้า $n_{iu} = n_{il} = n_i$ แล้วสูตรอาจเป็น

$$d_i = \frac{u_i - L_i}{n_i}$$

เคลลี (Allen and Yen. 1979 : 122; citing Kelley. 1939) กล่าวว่า ถ้าคะแนนรวมจากแบบทดสอบมีการแจกแจงแบบปกติ (Normally Distributed) ควรใช้ 27% ของผู้สอบที่ได้คะแนนรวมจากแบบ

เคลลี (Allen and Yen. 1979 : 122; citing Kelley. 1939) กล่าวว่า ถ้าคะแนนรวมจากแบบทดสอบมีการแจกแจงแบบปกติ (Normally Distributed) ควรใช้ 27% ของผู้สอบที่ได้คะแนนรวมจากแบบทดสอบสูงสุดเป็นช่วงที่สูง และใช้ 27% ของผู้สอบที่ได้คะแนนต่ำสุดเป็นช่วงที่ต่ำ เคอร์ตัน (Allen and Yen. 1979 : 122; citing Cureton. 1957) กล่าวว่า ถ้าคะแนนรวมจากแบบทดสอบแจกแจงเป็นโค้งที่แบนกว่าโค้งปกติแล้ว จำนวนเปอร์เซ็นต์ที่ดีที่สุดควรจะมากและใกล้เคียงกับ 33%

การหาค่าอำนาจจำแนกของข้อสอบอีกแบบหนึ่งคือ การหาสหสัมพันธ์แบบพอยท์ไบเซเรียล (Point - Biserial) ระหว่างคะแนนแต่ละข้อและคะแนนรวมของแบบทดสอบ (r_{ix}) ซึ่งหมายถึงสหสัมพันธ์ระหว่างคะแนนข้อสอบที่ i และคะแนนรวมของแบบทดสอบ X

$$r_{ix} = \frac{\bar{X}_i - \bar{X}}{s_x} \sqrt{\frac{p_i}{1 - p_i}}$$

เมื่อ	$\bar{X}_i$	คือ	ค่าเฉลี่ยของคะแนน X ของผู้สอบที่ทำข้อสอบ i ได้ถูกต้อง
	$\bar{X}$	คือ	ค่าเฉลี่ยของผู้สอบทั้งหมด
	s_x	คือ	ส่วนเบี่ยงเบนมาตรฐานของคะแนน X
	p_i	คือ	ความยากของข้อสอบ

ดัชนีอำนาจจำแนกของข้อสอบ d_i และสหสัมพันธ์ระหว่างข้อสอบและแบบทดสอบทั้งฉบับ (r_{ix}) เป็นสารสนเทศ ที่มีคุณค่า สำหรับข้อสอบที่ได้นั้น d_i และ r_{ix} ควรเป็นบวก ผู้สอบที่คะแนนรวมสูงสุดควรตอบคำถามรายข้อได้ถูกต้องมากกว่าพวกที่ได้คะแนนรวมต่ำ ข้อสอบที่มีค่า d_i หรือ r_{ix} ที่เป็นลบจะวัดสิ่งที่ตรงกันข้ามกับสิ่งที่แบบทดสอบต้องการจะวัดค่า d_i หรือ r_{ix} ที่เป็นลบอาจให้ข้อเสนอแนะว่า อาจมีความผิดพลาดในการให้คะแนนข้อสอบหรือข้อสอบนั้นใช้ภาษาไม่ดี ข้อสอบที่มีค่า d_i หรือ r_{ix} ต่ำคือเป็นลบ ควรได้รับการปรับปรุงหรือตัดทิ้ง

2. การวิเคราะห์ข้อสอบซึ่งใช้ดัชนีความเชื่อมั่นและความเที่ยงตรงของข้อสอบ ในการเลือกข้อสอบที่ดีที่สุดนั้นจำเป็นต้องใช้สถิติ 4 อย่าง สำหรับแต่ละข้อสอบ (i) คือ

- 1) ความยากของข้อสอบ (p_i)
- 2) ส่วนเบี่ยงเบนมาตรฐานของคะแนนข้อสอบ

$$s_i = \sqrt{p_i(1 - p_i)}$$

- 3) ดัชนีความเชื่อมั่นของข้อสอบ ($s_i r_{ix*}$)

$$r_{xx'} = \frac{K}{K-1} \left[1 - \frac{\sum_{i=1}^k s_i^2}{\left(\sum_{i=1}^k s_i r_{ix*} \right)^2} \right]$$

เมื่อ $r_{xx'}$ คือ ค่าความเชื่อมั่น (คูเดอร์ - ริชาร์ดสัน) ของคะแนนจากแบบทดสอบ x

$s_i r_{ix*}$ คือ ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ

r_{ix^*} คือ สหสัมพันธ์แบบพอยท์ไบซีเรียล ระหว่างคะแนนข้อสอบ และคะแนนจากแบบทดสอบที่มี N ข้อ

X^* คือ คะแนนแบบทดสอบที่ประกอบด้วยข้อสอบ N ข้อ ซึ่ง $N > K$

4) ดัชนีความเที่ยงตรงของข้อสอบ S_i, r_{iy} , เมื่อ r_{iy} หมายถึง สหสัมพันธ์แบบพอยท์ไบซีเรียล ระหว่างคะแนนข้อสอบแต่ละข้อและคะแนนเกณฑ์

$$r_{xy} = \frac{\sum_{i=1}^k S_i r_{iy}}{\sum_{i=1}^k S_i r_{ix^*}}$$

การประมาณค่าความเชื่อมั่น $r_{xx'}$ จะพบว่าความเชื่อมั่นแบบความสอดคล้องภายใน $r_{xx'}$ จะมีค่าสูงสุด เมื่อสหสัมพันธ์ระหว่างข้อสอบแต่ละข้อและคะแนนรวมของแบบทดสอบสูงสุด (r_{ix^*}) มีค่าสูงสุด เมื่อดัชนีความเชื่อมั่น ($S_i r_{ix^*}$) นั้นคำนวณโดยใช้ข้อสอบ k ข้อ ซึ่งคัดเลือกมาจากข้อสอบทั้งหมด N ข้อ

3. การวิเคราะห์ข้อสอบโดยใช้โค้งลักษณะของข้อสอบ โค้งลักษณะของข้อสอบ

(Item – Characteristic Curve, ICC) เป็นกราฟที่แสดงให้เห็นถึงความสัมพันธ์ระหว่างความน่าจะเป็นของการผ่านของข้อสอบเฉพาะและตำแหน่งของผู้สอบในคุณลักษณะที่ต้องการจะวัดโดยแบบทดสอบ เนื่องจากคะแนนภายใต้ลักษณะนั้นหาไม่ได้ จึงจำเป็นต้องใช้คะแนนที่ได้จากการวัดของแบบทดสอบ (Observed Test Scores) เป็นตัวประมาณค่าคุณลักษณะในแต่ละข้อสอบนั้น ICC ประมาณโดยการกำหนดจุด (Plot) โดยใช้คะแนนรวมจากแบบทดสอบอยู่ในแกนนอนและสัดส่วนของผู้สอบที่สอบผ่านข้อสอบนั้นในแกนตั้ง ICC สำหรับข้อสอบที่ดีควรจะมีค่าชันเป็นบวก และความยากในระดับปานกลาง

4. การใช้การวิเคราะห์องค์ประกอบสำหรับการวิเคราะห์ข้อสอบและการพัฒนาแบบทดสอบ วิธีวิเคราะห์องค์ประกอบสามารถนำมาใช้ในการพัฒนาแบบทดสอบได้ เพื่อช่วยในการกำหนดว่ากลุ่มของข้อสอบนั้นมีความเป็นเอกพันธ์หรือไม่ นั่นคือมีองค์ประกอบหลักตัวหนึ่งที่มีอิทธิพลต่อการกำหนดสิ่งดังกล่าว และยังช่วยในการเลือกข้อสอบที่มีลักษณะเป็นเอกพันธ์ การวิเคราะห์องค์ประกอบอาจใช้ประเมินความเที่ยงตรงเชิงโครงสร้าง การเลือกข้อสอบด้วยน้ำหนักขององค์ประกอบเฉพาะ จะทำให้ได้ความเที่ยงตรงเชิงเนื้อหาที่เหมาะสม เมื่อองค์ประกอบต่าง ๆ มีความสัมพันธ์กัน เมตริกซ์ของสหสัมพันธ์ระหว่างองค์ประกอบสามารถนำมาหาองค์ประกอบได้ นั่นคือองค์ประกอบสามารถนำมาวิเคราะห์องค์ประกอบต่อไปเพื่อให้ได้ตัวประกอบลำดับที่สอง (Second - Order Factors) ผลการวิเคราะห์องค์ประกอบสามารถนำไปสู่การพัฒนาแบบทดสอบแบบใหม่

ในการวิจัยครั้งนี้ ใช้การวิเคราะห์ข้อสอบ 3 วิธี คือ

1. การหาค่าอำนาจจำแนกของแบบทดสอบด้วยค่าสหสัมพันธ์แบบพอยท์ไบซีเรียล (Point - Biserial Correlation) และหาค่าความยาก (Item Difficulty)
2. การใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ
3. การใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ

ความหมายของความถนัด

นักการศึกษา นักจิตวิทยา และนักวัดผลการศึกษา ได้ให้ความหมายของความถนัด (Aptitude) ไว้แตกต่างกันดังนี้

ครอนบาช (Cronbach. 1970 : 38) กล่าวว่า ความถนัดเป็นประสบการณ์การเรียนรู้ที่กว้างขวาง และอ้างอิงถึงสถานการณ์อนาคต

ชวาล แพร์ตกุล (2517 : 1) กล่าวว่า ความถนัด หมายถึง สมรรถวิสัย (Capacity) และทิศทางแห่งความองกงามของสมองหรือกล่าวให้ง่ายขึ้นก็หมายถึง ชีตความสามารถสูงสุดของบุคคลที่เขาอาจมีได้ต่อการเรียนรู้และการฝึกฝนในวิทยาการต่าง ๆ และทักษะทั้งปวง ถ้าเขาได้รับการฝึกฝนและประสบการณ์ที่เหมาะสม

สมบุรณ์ ชิตพงศ์ และสำเริง บุญเรืองรัตน์ (2518 : 17) กล่าวว่า ความถนัด หมายถึง ศักยภาพที่มีอยู่ในตัวบุคคลอันเป็นผลมาจากการฝึกฝนความรู้และประสบการณ์ทั้งสิ้นทั้งปวง

ล้วน สายยศ และอังคณา สายยศ (2527 : 27) กล่าวว่า ความถนัด หมายถึง ความสามารถที่บุคคลได้รับประสบการณ์ ฝึกฝนตนเอง และสั่งสมไว้มากจนเกิดเป็นทักษะเด่นชัดด้านใดด้านหนึ่ง พร้อมทั้งจะปฏิบัติกิจกรรมด้านนั้นได้อย่างดี

จากความหมายของความถนัดดังกล่าว อาจสรุปได้ว่า ความถนัด หมายถึง ความสามารถสมองของแต่ละบุคคลที่เกิดจากการเรียนรู้ ฝึกฝน และประสบการณ์ ซึ่งบุคคลนั้นสามารถแสดงออกได้อย่างมีประสิทธิภาพ

ทฤษฎีสมรรถภาพสมองและความถนัด

ทฤษฎีเกี่ยวกับสมรรถภาพสมองและความถนัดมีหลายทฤษฎี แต่ที่สำคัญมีดังต่อไปนี้ (ล้วน สายยศ และอังคณา สายยศ. 2527 : 27 - 34)

1. ทฤษฎีองค์ประกอบเดียว (Uni - Factor Theory) บางทีทฤษฎีนี้เรียกว่า Global Theory ผู้คิดทฤษฎีนี้คือ บิเนท์ และ ซีมอน (Binet and Simon. 1905) ทฤษฎีนี้เสนอโครงสร้างเชาว์ปัญญาเป็นลักษณะอันหนึ่งอันเดียวไม่แบ่งออกเป็นส่วนย่อยคล้ายกับเป็นความสามารถทั่วไป (General Ability) นั้นเอง ในปี ค.ศ.1905 หรือ พ.ศ. 2448 บิเนท์ และ ซีมอน ได้สร้างข้อสอบวัดตามแนวคิดของเขาเป็นครั้งแรก ข้อสอบฉบับนี้สร้างวัดเชาว์ปัญญาเป็นแบบ Global Measure คือ วัดออกมาเป็นคะแนนเดียวแล้วแปลความหมายว่าใครมีปัญญาระดับใด

2. ทฤษฎีสององค์ประกอบ (Bi - Factor Theory) ทฤษฎีนี้ได้นำโดยนักจิตวิทยาชาวอังกฤษ ชื่อ สเปียร์แมน (Charles Spearman) ในปี ค.ศ. 1927 ทฤษฎีนี้เกิดจากการวิเคราะห์คุณลักษณะโดยขบวนการทางสถิติ พบว่าสมรรถภาพทางสมองของมนุษย์นั้นมียุทธปัจจัยประกอบอยู่สองประการคือ องค์ประกอบทั่วไป (General Factor) เรียกย่อ ๆ ว่า G - Factor และองค์ประกอบเฉพาะ (Specific Fator) เรียกย่อ ๆ ว่า S - Factor แต่ละองค์ประกอบมีกิจกรรมเฉพาะตัวชนิดหนึ่งของมัน สมรรถภาพทั่วไปที่เรียกว่า G - Factor นั้นจะสอดแทรกอยู่ในทุกอริยาบถของความคิดและการกระทำของมนุษย์ และแต่ละคนก็มีสมรรถภาพสมองชนิดทั่วไปนี้แตกต่างกันออกไปเล็กน้อยตามแต่ละบุคคล ส่วนองค์ประกอบเฉพาะหรือ S - Factor นั้นเป็นองค์ประกอบที่สำคัญที่ทำให้มนุษย์มีความแตกต่างกัน และเป็นความสามารถพิเศษที่มีอยู่ในแต่ละบุคคล เช่น ความสามารถพิเศษด้านดนตรี ด้านศิลปะ การวาดเขียน ทางเครื่องยนต์กลไก และทางช่างต่าง ๆ เป็นต้น

3. ทฤษฎีหลายองค์ประกอบ (Multiple - Factor Theory) ทฤษฎีนี้เป็นที่ยอมรับอย่างกว้างขวางของนักจิตวิทยาชาวอเมริกัน ผู้นำในการสร้างทฤษฎีนี้คือ เทอร์สโตน (L.L. Thurstone) เสนอทฤษฎีนี้เมื่อปี ค.ศ. 1933 โดยทำการวิจัยโครงสร้างทางสมองออกได้หลายอย่าง แต่ที่เห็นเด่นชัดและสำคัญ ๆ มีอยู่ 7 ประการ คือ

3.1 องค์ประกอบด้านภาษา (Verbal Factor ใช้ตัวย่อว่า V.) องค์ประกอบของสมองส่วนนี้จะส่งผลให้รู้ถึงความสามารถด้านความเข้าใจในภาษาและการสื่อสารทั่วไป ผู้มีองค์ประกอบด้านนี้สูง จะมีความสามารถในการอ่านเรื่อง อ่านแบบเข้าใจความหมาย รู้ความสัมพันธ์ของคำ รู้ความหมายของศัพท์ได้เป็นอย่างดี

3.2 องค์ประกอบด้านความคล่องแคล่วในการใช้ถ้อยคำ (Word Fluency Factor ใช้ตัวย่อว่า W.) เป็นความสามารถที่จะใช้คำได้มากในเวลาจำกัด เช่น ให้หาคำที่ขึ้นต้นด้วย “ด” มากที่สุด ในเวลาที่จำกัด

3.3 องค์ประกอบด้านจำนวน (Number Factor ใช้ตัวย่อว่า N.) เป็นความสามารถที่มองเห็นความสัมพันธ์ และความหมายของจำนวน และมีความแม่นยำคล่องแคล่วในการบวก ลบ คูณ หาร ในวิชาเลขคณิตเป็นอย่างดี

3.4 องค์ประกอบด้านมิติสัมพันธ์ (Space Factor ใช้ตัวย่อว่า S.) เป็นความสามารถที่จะเข้าใจถึงขนาดและมิติต่าง ๆ ได้แก่ ความสั้น ยาว ไกล ใกล้ และพื้นที่หรือทรวดทรงที่มีขนาดและปริมาตรที่ต่างกัน สามารถสร้างจินตนาการให้เห็นส่วนย่อยและส่วนผสมของวัตถุต่าง ๆ ที่นำมาซ้อนทับกัน สามารถรู้ความสัมพันธ์ของรูปทรงเรขาคณิตเมื่อเปลี่ยนแปลงที่อยู่

3.5 องค์ประกอบด้านความจำ (Memory Factor ใช้ตัวย่อว่า M.) เป็นความสามารถในด้านความทรงจำเรื่องราว และมีสติระลึกจูนสามารถถ่ายทอดได้

3.6 องค์ประกอบด้านสังเกตพิจารณา (Perceptual Speed Factor ใช้ตัวย่อว่า P.) เป็นความสามารถในการมองเห็นรายละเอียดความคล้ายคลึง หรือความแตกต่างระหว่างสิ่งของต่าง ๆ อย่างรวดเร็วถูกต้อง

3.7 องค์ประกอบด้านเหตุผล (Reasoning Factor ใช้ตัวย่อว่า R.) เป็นความสามารถในการจัดประเภทอุปมาอุปไมย และสรุปความได้อย่างรวดเร็วและถูกต้อง

4. ทฤษฎีไฮราคิคอล (Hierarchical Theory) ผู้นำทฤษฎี คือ เวอร์นอน (Vernon) เบิร์ต (Burt) ชาวอังกฤษ และฮัมฟรีย์ (Humphreys) ชาวอเมริกาโดยเฉพาะเวอร์นอน ได้เสนอโครงสร้างของเชาวันปัญญา ในปี ค.ศ.1960 โดยเริ่มต้นอธิบายตามแบบของสเปียร์แมน นั่นคือ เวอร์นอนเริ่มจุดแรกด้วย G - Factor ขึ้นต่อไปแบ่งออกเป็น 2 องค์ประกอบใหญ่ ๆ คือ ความถนัดทางภาษา (Verbal Education ; V : ed) ความถนัดทางช่าง (Practical Mechanical ; K : m) ซึ่งองค์ประกอบใหญ่ทั้งสองรวมเรียกว่า Major Group Factors องค์ประกอบใหญ่ทั้งสองยังแบ่งองค์ประกอบย่อยลงไปอีกได้เรียกว่า Verbal - Educational V ; ed ยังแบ่งย่อยเป็นองค์ประกอบด้านภาษา (Verbal) องค์ประกอบด้านตัวเลข (Number) และอื่น ๆ อีก ส่วนด้าน K ; m ได้แบ่งย่อยออกเป็นความรู้เชิงกล (Mechanical Infomation) มิติสัมพันธ์ (Spatial) ความสามารถในการใช้กลไกของกล้ามเนื้อ (Psychomotor Abilities) และอื่น ๆ อีก ซึ่งแต่ละองค์ประกอบย่อยยังแบ่งเป็นองค์ประกอบย่อย ๆ ลงไปอีกเป็นองค์ประกอบต่ำที่สุด เรียกว่าองค์ประกอบเฉพาะ (Specific Factors)

5. ทฤษฎีโครงสร้างสามมิติของปัญญา (Three Faces of Intellect Model) ทฤษฎีนี้สร้างขึ้นโดย กิลฟอร์ด (Guidford) ได้วิเคราะห์ความสัมพันธ์ของคุณลักษณะของเชาวันปัญญาเป็น 3 มิติ คือ

มิติที่ 1 ด้านกระบวนการหรือวิธีการของการคิด (Operations) มีส่วนประกอบย่อย 5 ส่วน คือ การรู้การเข้าใจ (Cognition) ความจำ (Memory) การคิดออกแนก (Divergent Production) การคิดแบบเอกนัย (Convergent Production) และการคิดแบบประเมินค่า (Evaluation)

มิติที่ 2 ด้านเนื้อหา (Content) เป็นด้านที่ประกอบด้วยสิ่งเร้าและข้อมูลต่าง ๆ แบ่งออกเป็น 4 อย่าง ดังนี้ ภาพ (Figural) สัญลักษณ์ (Symbolic) ภาษา (Semantic) และพฤติกรรม (Behavioral)

มิติที่ 3 ผลการคิด (Products) เป็นผลของกระบวนการจัดกระทำของความคิดกับข้อมูลจากเนื้อหา แบ่งออกเป็น 6 อย่าง ดังนี้ หน่วย (Units) จำพวก (Classes) ความสัมพันธ์ (Relations) ระบบ (Systems) การแปลงรูป (Transformations) และ การประยุกต์ (Implications)

6. ทฤษฎีความสามารถทางสมองสองระดับ (Two - Level Theory of Mental Ability) ทฤษฎีนี้เสนอโดยเจนเซน (Jensen) เจนเซนเสนอทฤษฎีว่าความสามารถทางสมองมีอยู่สองระดับ ระดับ I (Level I) เป็นความสามารถด้านการเรียนรู้และจำอย่างนกแก้ว นั่นคือ เป็นความสามารถที่จะสะสมหรือเก็บข้อมูลไว้ได้ และพร้อมที่จะระลึกนึกออกได้ ระดับ II (Level II) เป็นระดับของการจัดกระทำทางสมอง เป็นขั้นสร้างมโนภาพเหตุผล และแก้ปัญหา ระดับ II นี้ดูไปแล้วก็เหมือนกับองค์ประกอบทั่วไป (G - factor) นั่นเอง

7. ทฤษฎีปัญญาของแคทเทลล์ ทฤษฎีนี้คิดโดย อาร์ บี แคทเทลล์ (R.B. Cattell) เขาเสนอทฤษฎีปัญญาวา โครงสร้างเชาว์ปัญญา ประกอบด้วย 2 ส่วน คือ

7.1 ส่วนที่เป็นฟลูอิด (Fluid Component) เป็นความสามารถทั่วไป ผู้ที่มีปริมาณความสามารถด้านนี้สูง จะสามารถทำงานชนิดต่าง ๆ ได้ดี เช่น ความสามารถด้านเหตุผลเชิงอุปมาน และอนุมานเหตุผลสัมพันธ์ ความสามารถเข้าใจการเปลี่ยนแปลงของอนุกรมภาพ เป็นต้น

7.2 ส่วนที่เป็นคริสตอลไลซ์ (Crystallized Component) เป็นความสามารถที่เชื่อมโยงกับวัฒนธรรมและสิ่งแวดล้อมอย่างใกล้ชิด พูดง่าย ๆ ว่าความสามารถที่จะเข้าใจภาษาความสามารถที่จะประเมินคุณค่า เป็นต้น

ประโยชน์ของแบบทดสอบวัดความถนัด

ลัวัน สายยศ และอังคณา สายยศ (2517 : 119 - 124) กล่าวถึงประโยชน์ของแบบทดสอบความถนัดดังนี้

1. ใช้ในการสอบคัดเลือก
2. ใช้ในการแยกประเภทนักเรียน
3. ใช้ในการวินิจฉัยความสามารถ
4. ใช้ในการพยากรณ์ความสำเร็จ
5. ใช้ในการวัดพัฒนาการ
6. ใช้ในการเปรียบเทียบสติปัญญา
7. ใช้ในการประเมินผลการศึกษา
8. ใช้ในการวิจัย

งานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้า

ซิมเมอร์แมน และฮัมฟรี (วารุณี ปิครวัชชัย. 2513:9; อ้างอิงมาจาก Zimmermann and Humphrey. 1953:460-461) ได้ศึกษาถึงค่าความเชื่อมั่นของแบบทดสอบเลือกตอบพบว่า การกำจัดตัวลวงที่มีค่าอำนาจจำแนกต่ำสุดออกจากตัวเลือกของแบบทดสอบที่มี 5 ตัวเลือก จะทำให้ค่าความเชื่อมั่นของแบบทดสอบสูงขึ้น และทำให้เวลาในการตอบแต่ละข้อลดลง

บูเดสคู และเนโว (Budescu and Nevo. 1985: 183-196) ได้ศึกษาตัวเลือกที่เหมาะสมเพื่อหาข้อสันนิษฐานเกี่ยวกับสัดส่วน โดยใช้กลุ่มตัวอย่าง 1018 คน ที่สอบเข้ามหาวิทยาลัยไฮฟา (The University of Haifa) เครื่องมือในการวิจัยเป็นแบบทดสอบ 3 ฉบับ ประกอบด้วย แบบทดสอบคำศัพท์ 60 ข้อ แบบทดสอบความเข้าใจในการอ่าน 24 ข้อ แบบทดสอบเหตุผลทางคณิตศาสตร์ 25 ข้อ ครั้งแรกใช้แบบทดสอบ 5 ตัวเลือก แล้วนำมาแบ่งเป็นแบบทดสอบสภาพการณ์ต่างๆ 6 แบบ คือ แบบทดสอบที่มี 4, 3 และ 2 ตัวเลือก โดยการสุ่มตัวลวงตัวสุดท้ายออกและแบบทดสอบที่มี 4, 3 และ 2 ตัวเลือก โดยการคัดตัวลวงที่มีอำนาจจำแนกต่ำสุดออก ผลการศึกษาพบว่า แบบทดสอบทั้ง 3 ฉบับที่มีจำนวนตัวเลือกมากจะให้ค่าความแปรปรวนมาก ค่าความเชื่อมั่นและค่าอำนาจจำแนกสูงกว่าแบบทดสอบที่มีจำนวนตัวเลือกน้อย

ผดุงชาติ สุวรรณวงศ์ (2516 : 42 - 43) ได้ทำการวิจัยเรื่องความเชื่อมั่นและความเที่ยงตรงของข้อสอบที่เพิ่มความยาว โดยให้ความคิดเห็นว่า ถ้าขยายข้อสอบให้ยาวขึ้นความเชื่อมั่นจะสูงขึ้น ผลการวิจัยพบว่า ข้อสอบบางฉบับเมื่อเพิ่มความยาวขึ้นแล้วมีค่าความเชื่อมั่นต่ำลง ทั้งนี้เป็นเพราะข้อสอบที่เพิ่มบางข้อมีความสัมพันธ์ทางลบกับข้อเดิมจึงทำให้สหสัมพันธ์เฉลี่ยลดลง

บุญเรือน จรัสวิมล (2521 : 40 - 41) ได้ศึกษาการเพิ่มข้อสอบที่มีผลค่าความเชื่อมั่นของแบบทดสอบ โดยเพิ่มจำนวนข้อสอบในแบบทดสอบวัดผลสัมฤทธิ์วิชาวิทยาศาสตร์ จากเดิม 30 ข้อ เป็น 60 ข้อ ข้อสอบที่เพิ่มในแต่ละฉบับมีช่วงระดับความยากต่างกันคือ .40 - .60 , .30 - .70 และ .20 - .80 ผลการศึกษาพบว่า การเพิ่มจำนวนข้อสอบด้วยข้อสอบที่มีความยากต่างกันไม่ทำให้ค่าความเชื่อมั่นต่างกัน เมื่อนำแบบทดสอบที่เพิ่มความยาวไปเปรียบเทียบกับค่าความเชื่อมั่นฉบับเดิมพบว่า แบบทดสอบที่เพิ่มความยาวที่มีช่วงระดับความยาก .30 - .70 และ .20 - .80 มีความเชื่อมั่นสูงกว่าฉบับเดิมอย่างมีนัยสำคัญทางสถิติที่ .01 และ .05 ตามลำดับ ส่วนแบบทดสอบที่เพิ่มจำนวนข้อใน ช่วงระดับความยาก .40 - .60 ค่าความเชื่อมั่นสูงกว่าฉบับเดิมอย่างไม่มีนัยสำคัญทางสถิติ

ยศอารีย์ รวยธนพานิช (2534 : 56-58) ได้เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบปรนัยชนิดเลือกตอบวิชาคณิตศาสตร์ที่มีความยากแตกต่างกันคือ .40 - .60 , .30 - .70 และ .20 - .80 และเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวนข้อต่างกันคือ 30 , 40 และ 50 ข้อ ผลการศึกษาพบว่า แบบทดสอบที่มีช่วงความยากต่างกันเมื่อมีจำนวนข้อเท่ากันมีค่าความเชื่อมั่นไม่ต่างกัน และแบบทดสอบที่มีความยาวต่างกันเมื่อมีช่วงความยากเท่ากัน มีค่าความเชื่อมั่นไม่ต่างกัน ยกเว้นที่มีช่วงความยาก .40 - .60 แบบทดสอบที่มี 50 , 40 ข้อ มีค่าความเชื่อมั่นสูงกว่าแบบทดสอบที่มี 30 ข้อ ส่วนแบบทดสอบที่มี 50 และ 40 ข้อ มีค่าความเชื่อมั่นไม่ต่างกัน

สุทิน บุญญาภิบาล (2537 : 82) ได้ศึกษาการเปรียบเทียบคุณภาพของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ โดยทฤษฎีคลาสสิกอล เมื่อกลุ่มผู้สอบมีจำนวนต่างกัน โดยศึกษากับกลุ่มตัวอย่างขนาด 100 คน 150 คน 225 คน 338 คน 507 คน 761 คน 1,142 คน 1,713 คน 2,570 คน พบว่า

1. ค่าความยากของแบบทดสอบมีค่าความยากแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. เมื่อกลุ่มตัวอย่างต่างกันมีค่าอำนาจจำแนกแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3. เมื่อกลุ่มตัวอย่างต่างกันมีค่าความเชื่อมั่น เมื่อหาโดยสูตรครุเดอร์-ชาร์ตสัน 20 และโดยวิธีแบ่งครึ่งฉบับแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

อรนุช ธิติรักษ์พาณิชย์ (2538 : 91 -94) ได้เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดบุคลิกภาพเยาวชนไทย เมื่อลดความยาวให้สั้นลง โดยมีความยาว 300 ข้อ แล้วลดลงเหลือ 210 , 180 , 150 และ 120 ข้อ นั้นความยาวของแบบทดสอบควรเป็นเท่าใดจึงจะเหมาะสมหรือยังมีค่าความเชื่อมั่นไม่แตกต่างไปจากฉบับเดิม ผลการศึกษา พบว่า แบบทดสอบที่ลดความยาวลงเหลือ 120 ข้อ มีค่าความเชื่อมั่นแตกต่างไปจากเดิม และความยาวของแบบทดสอบวัดบุคลิกภาพเยาวชนไทยที่เหมาะสม คือ 150 ข้อ

พิมพ์อุษา เจริญยศ (2539 : 51 -54) ได้เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อโดยใช้เกณฑ์ค่าสถิติกับใช้เกณฑ์พินิจของครู โดยมีแบบทดสอบ 40 ข้อ แล้วลดลงเหลือ 35 , 30 , 25 , 20 และ 15 ข้อ โดยใช้เกณฑ์ค่าสถิติพิจารณา พบว่าแบบทดสอบที่มีจำนวน 25 ข้อ มีค่าความเชื่อมั่นสูงสุด ส่วนแบบทดสอบที่มีจำนวน 15 ข้อ มีค่าความเชื่อมั่นต่ำสุด และจากการเปรียบเทียบค่าความเชื่อมั่นแต่ละฉบับ พบว่า แบบทดสอบที่มีจำนวน 35 ข้อ และ 15 ข้อ แตกต่างไปจากแบบทดสอบที่มีจำนวน 40 ข้อ ส่วนที่เหลือไม่พบความแตกต่างอย่างมีนัยสำคัญทางสถิติ ค่าความเชื่อมั่นของแบบทดสอบระหว่างฉบับที่มีจำนวน 40 ข้อ กับฉบับที่ลดจำนวนข้อลงเหลือ 35 , 30 , 25 , 20 และ 15 ข้อ โดยใช้เกณฑ์ดุลพินิจของครูพิจารณา แบบทดสอบที่มีจำนวน 25 ข้อ มีค่าความเชื่อมั่นสูงสุด ส่วนแบบทดสอบที่มีจำนวน 15 ข้อ มีค่าความเชื่อมั่นต่ำสุดและจากการเปรียบเทียบค่าความเชื่อมั่นแต่ละฉบับ พบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติเป็นส่วนใหญ่ ยกเว้นค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 25 ข้อ เพียงฉบับเดียวเท่านั้นที่ให้ค่าความเชื่อมั่นไม่แตกต่างไปจากแบบทดสอบที่มีจำนวน 40 ข้อ ค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติกับใช้เกณฑ์ดุลพินิจของครูที่มีจำนวนข้อเท่ากัน พบว่า ความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติทุกคู่ โดยสรุปแล้วแบบทดสอบเลือกตอบวิชาคณิตศาสตร์ ฉบับโจทย์ปัญหาที่มีจำนวน 40 ข้อ สามารถลดจำนวนข้อลงเหลือเป็น 25 ข้อ ได้โดยยังคงค่าความเชื่อมั่นไม่เปลี่ยนแปลง

จากการศึกษาค้นคว้าเอกสารและงานวิจัยที่เกี่ยวข้อง ผู้วิจัยมีความสนใจในการศึกษาจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์และค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ด้วยเทคนิควิธีการวิเคราะห์ข้อสอบต่างกันและขนาดกลุ่มตัวอย่างต่างกันว่าจะได้จำนวนข้อและค่าความเชื่อมั่นต่างกันหรือไม่ และวิธีใดจะให้ค่าความเชื่อมั่นสูงสุด ซึ่งจะเป็นแนวทางในการเลือกวิธีการวิเคราะห์ข้อสอบและขนาดกลุ่มตัวอย่างเพื่อให้ได้แบบทดสอบที่มีค่าความเชื่อมั่นให้สูงขึ้นอีกวิธีหนึ่ง

สมมติฐานในการศึกษาค้นคว้า

1. จำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ จากการวิเคราะห์ข้อสอบวิธีหาค่าอำนาจจำแนกและความยาก วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ และวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ มีจำนวนข้อแตกต่างกันในแต่ละขนาดกลุ่มตัวอย่าง คือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน

2. จำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ ระหว่างกลุ่มตัวอย่างคือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน มีจำนวนข้อแตกต่างต่างกัน จากการวิเคราะห์ข้อสอบ คือ วิเคราะห์ค่าอำนาจจำแนกและค่าความยาก วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ และวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ

3. แบบทดสอบจำนวน 50 ข้อ ที่วิเคราะห์ด้วยวิธีหาค่าอำนาจจำแนกและค่าความยาก วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ วิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ มีค่าความเชื่อมั่นแตกต่างกันในแต่ละขนาดกลุ่มตัวอย่าง คือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน

4. แบบทดสอบจำนวน 50 ข้อ ที่วิเคราะห์จากกลุ่มตัวอย่างขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน มีค่าความเชื่อมั่นแตกต่างกัน ในแต่ละวิธีวิเคราะห์ข้อสอบ คือ วิเคราะห์ค่าอำนาจจำแนกและค่าความยาก วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ วิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2540 ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดกรุงเทพมหานคร ทั้งหมด 109 โรงเรียน จำนวนนักเรียน 56,823 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2540 ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดกรุงเทพมหานคร จำนวน 12 โรงเรียน จำนวน 750 คน โดยวิธีการสุ่มแบบสองขั้นตอน (Two Stage Random Sampling) โดยมีขั้นตอนการดำเนินการดังนี้

ประมาณกลุ่มตัวอย่างที่ใช้ในการศึกษา ด้วยความเชื่อมั่น 95% ($\alpha = .05$) เมื่อเทียบจากตารางขนาดของกลุ่มตัวอย่างที่ควรเลือกจากประชากรพบว่า ต้องใช้กลุ่มตัวอย่างอย่างน้อย 398 คน (ลัวน สายยศ และอังคณา สายยศ.2536 ; อ้างอิงมาจาก Yamane. 1967) แต่ในการวิจัยครั้งนี้ใช้กลุ่มตัวอย่าง 750 คน ซึ่งทำการสุ่ม ดังนี้

ขั้นที่ 1 สุ่มโรงเรียนตามขนาดโรงเรียน ด้วยวิธีการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) มีขนาดโรงเรียนเป็นชั้น (Strata) มีโรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit) โดยแบ่งโรงเรียนออกเป็น 3 ขนาดคือ โรงเรียนขนาดใหญ่พิเศษ โรงเรียนขนาดใหญ่ และโรงเรียนขนาดกลางคือ

1. โรงเรียนขนาดใหญ่พิเศษ มีนักเรียนมากกว่า 2,500 คน มี 48 โรงเรียน มีจำนวนห้องเรียน 645 ห้องเรียน
2. โรงเรียนขนาดใหญ่ มีนักเรียน 1,500 ถึง 2,500 คน มี 47 โรงเรียน มีจำนวนห้องเรียน 522 ห้องเรียน
3. โรงเรียนขนาดกลาง มีนักเรียน 500 ถึง 1,500 คน มี 14 โรงเรียน มีจำนวนห้องเรียน 112 ห้องเรียน

สุ่มโรงเรียนมาประมาณ 10 % ของแต่ละขนาด ได้โรงเรียนขนาดใหญ่พิเศษ 5 โรงเรียน โรงเรียนขนาดใหญ่ 5 โรงเรียน และโรงเรียนขนาดกลาง 2 โรงเรียน

ขั้นที่ 2 สุ่มห้องเรียนจากโรงเรียนแต่ละขนาดในขั้นที่ 1 ด้วยวิธีการสุ่มอย่างง่าย โดยสุ่มโรงเรียนขนาดใหญ่พิเศษ 12 ห้องเรียน จำนวนนักเรียน 384 คน โรงเรียนขนาดใหญ่ 10 ห้องเรียน จำนวนนักเรียน 293 คน โรงเรียนขนาดกลาง 2 ห้องเรียน จำนวนนักเรียน 73 คน รวมนักเรียนที่ใช้เป็นกลุ่มตัวอย่างทั้งสิ้น 750 คน ดังแสดงในตาราง 1

ตาราง 1 จำนวนกลุ่มตัวอย่าง จำแนกตามขนาดโรงเรียน

โรงเรียน	จำนวนนักเรียนทั้งหมด (ห้องเรียน)	จำนวนนักเรียนกลุ่มตัวอย่าง (ห้องเรียน)
ขนาดใหญ่พิเศษ		
พระโขนงพิทยาล	738 (16)	97 (3)
สวนกุหลาบวิทยาลัย	662 (14)	87 (3)
ศรีอยุธยา	543 (12)	71 (2)
ชิโนรสวิทยาลัย	523 (12)	67 (2)
วัดราชบพิธ	477 (10)	62 (2)
ขนาดใหญ่		
มัธยมวัดธาตุทอง	502 (11)	66 (2)
เทพศิลา	487 (9)	64 (2)
ธนบุรีวรเทพิพลารักษ์	465 (12)	61 (2)
สุวรรณารามวิทยาคม	428 (9)	56 (2)
โยธินบูรณะ	351 (8)	46 (2)
ขนาดกลาง		
มักกะสันพิทยาล	292 (8)	38 (1)
มัธยมวัดเบญจมบพิตร	264 (10)	35 (1)
รวม	6161 (138)	750 (24)

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ คือแบบทดสอบวัดความถนัดด้านคณิตศาสตร์ ซึ่งสร้างตามแนวทฤษฎีของเทอร์สโตน เป็นแบบทดสอบปรนัยชนิดเลือกตอบ 5 ตัวเลือกจำนวน 100 ข้อ ซึ่งผู้วิจัยสร้างขึ้นเอง

ขั้นตอนการสร้างแบบทดสอบวัดความถนัดทางการเรียน

เครื่องมือที่ใช้ในการวิจัยครั้งนี้มีขั้นตอนการสร้าง ดังนี้

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบวัดความถนัดด้านคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2
2. ศึกษาทฤษฎีและเอกสารงานวิจัยที่เกี่ยวข้องคือแบบทดสอบวัดความถนัดตามแนวทฤษฎีของเทอร์สโตน เพื่อนำมาเป็นแนวทางในการสร้างแบบทดสอบวัดความถนัดด้านคณิตศาสตร์
3. เขียนนิยามของแบบทดสอบวัดความถนัดทางด้านคณิตศาสตร์ แล้วสร้างข้อสอบตามที่นิยามไว้ ซึ่งแบ่งเป็น 4 ประเภท ประเภทละ 35 ข้อ จำนวนทั้งหมด 140 ข้อ

4. นำแบบทดสอบความถนัดด้านคณิตศาสตร์ที่สร้างขึ้น ให้ผู้เชี่ยวชาญทางด้านวัดผลการศึกษา ตรวจสอบและวิจารณ์เกี่ยวกับความชัดเจนของคำถามและตัวเลือกปรับปรุงแก้ไขตามที่คุณเชี่ยวชาญแนะนำ และคัดเลือกแบบทดสอบทั้ง 4 ประเภท ประเภทละ 25 ข้อ

5. พิมพ์แบบทดสอบจำนวน 100 ข้อ เป็นรูปเล่ม เพื่อนำไปใช้ในการวิจัย

ตัวอย่างแบบทดสอบ

แบบทดสอบวัดความถนัดด้านคณิตศาสตร์

คำชี้แจง

1. แบบทดสอบฉบับนี้เป็นแบบปรนัยชนิดเลือกตอบ มีจำนวน 100 ข้อ ใช้เวลาทำ 60 นาที ฉะนั้นนักเรียนควรรีบตอบโดยเร็วและให้ครบทุกข้อจึงจะได้คะแนนดี
2. แบบทดสอบฉบับนี้เป็นแบบทดสอบวัดความถนัดด้านคณิตศาสตร์ การตอบของนักเรียนไม่มีผลต่อการเรียนวิชาคณิตศาสตร์แต่อย่างใด โปรดทำให้เต็มความสามารถ

ประเภททักษะ (Skill) เป็นคำถามประเภทการคำนวณด้านตัวเลขขั้นพื้นฐาน ทั่วไป

(0) $6241 + 6400 + 6561 = ?$

ก 19,020

ข 19,022

ค 19,200

(ง) 19,202

จ 19,220

ประเภทความคิดรวบยอด (concept) เป็นคำถามประเภทความเข้าใจเกี่ยวกับหลักการ วิธีการ รวมทั้งการแปลความ ตีความ ขยายความ

(00) ชื้อมะม่วงมา 1 ชะลอม กับ 10 ผล เป็นเงิน 100 บาท เมื่อขายไปได้กำไร 10 บาท ขายมะม่วงไปผลละเท่าไร จากโจทย์หาคำตอบได้หรือไม่เพราะเหตุใด

ก หาไม่ได้เพราะไม่ทราบราคาทุน

ข หาไม่ได้เพราะไม่ทราบราคาขาย

(ค) หาไม่ได้เพราะไม่ทราบจำนวนมะม่วงที่ซื้อมา

ง หาไม่ได้เพราะไม่ทราบจำนวนมะม่วงที่ขาย

จ หาได้

ประเภทโจทย์ปัญหา (Problem Solving) เป็นคำถามประเภทโจทย์ปัญหาทั่วไป

(000) มีเชือกยาว $5\frac{1}{4}$ เมตร ต้องการตัดเชือกให้ยาวเส้นละ 90 เซนติเมตร

จะได้เชือกทั้งหมดกี่เส้น

ก 9

ข 8

ค 7

ง 6

(จ) 5

ประเภทเรียงลำดับ (N - series) เป็นคำถามประเภทให้หาคำตอบที่อยู่ถัดไป โดยพิจารณาจากที่โจทย์กำหนดมาให้

(0000) 1 , 2 , 4 , 7 , 11 , ?

- ก 14
- ข 15
- (ค) 16
- ง 17
- จ 18

3. ขอให้กาเครื่องหมาย X ลงในกระดาษคำตอบเพียงข้อละหนึ่งคำตอบเท่านั้น ถ้าตอบเกินกว่าหนึ่งคำตอบจะถือว่าข้อนั้นไม่ได้คะแนน

4. อย่าขีดเขียนหรือทำเครื่องหมายใด ๆ ลงในแบบทดสอบฉบับนี้ ถ้าต้องการขีดเขียนใด ๆ ให้ทำด้านหลังกระดาษคำตอบ

วิธีดำเนินการรวบรวมข้อมูล

1. ขอหนังสือจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อขอความอนุเคราะห์ในการเก็บรวบรวมข้อมูล
2. ติดต่อโรงเรียนที่ใช้เป็นกลุ่มตัวอย่าง ขออนุญาตผู้บริหารโรงเรียน เพื่อกำหนด วัน เวลา ที่ทำการทดสอบ
3. วางแผนในการดำเนินการสอบโดยผู้วิจัยเป็นผู้ดำเนินการสอบเอง จัดเตรียมแบบทดสอบให้เพียงพอกับจำนวนนักเรียนในแต่ละครั้ง
4. นำแบบทดสอบไปทดสอบกับนักเรียนที่เป็นกลุ่มตัวอย่างทั้งหมด ตามวัน เวลาที่นัดหมายไว้
5. ชี้แจงให้นักเรียนในกลุ่มตัวอย่างเข้าใจ และทราบถึงวัตถุประสงค์ในการสอบและขอความร่วมมือในการสอบ
6. อธิบายให้นักเรียนทุกคนเข้าใจวิธีการทำแบบทดสอบ
7. หลังจากดำเนินการสอบกลุ่มตัวอย่างครบทุกโรงเรียนแล้ว นำกระดาษคำตอบมาตรวจให้คะแนน ดังนี้ ถ้าตอบถูกให้คะแนน 1 คะแนน ถ้าตอบผิดหรือเว้นว่างให้ 0 คะแนน และนำผลมาวิเคราะห์ทางสถิติ

ขั้นตอนการวิเคราะห์ข้อมูล

1. หาค่าสถิติพื้นฐานของแบบทดสอบ
 - 1.1 คะแนนเฉลี่ย ($\bar{X}$)
 - 1.2 ความเบี่ยงเบนมาตรฐาน (SD)
2. วิเคราะห์ข้อสอบด้วยวิธีที่ต่างกัน 3 วิธี กับกลุ่มตัวอย่างขนาดเล็ก 5 ขนาดคือ 30 คน 60 คน 90 คน 120 คน 150 คน และกลุ่มตัวอย่างขนาดใหญ่ 300 คน คือ

2.1 คัดเลือกข้อสอบตามเกณฑ์

2.1.1 วิเคราะห์ข้อสอบโดยหาค่าอำนาจจำแนกแบบสหสัมพันธ์พอยท์ไบซีเรียล (Point Biserial Correlation) และหาค่าความยาก คัดข้อสอบทุกข้อที่มีค่าอำนาจจำแนกตั้งแต่ .20 ขึ้นไป และมีค่าความยากตั้งแต่ .20 - .80

2.1.2 วิเคราะห์ข้อสอบโดยใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ จากข้อสอบทั้งหมด 100 ข้อ มีขั้นตอนดังนี้

2.1.2.1 หาค่าความเชื่อมั่นของข้อสอบทั้งหมด 100 ข้อ

2.1.2.2 ตัดข้อสอบข้อ 1 ออก จะเหลือข้อสอบทั้งหมด 99 ข้อ และหาค่าความเชื่อมั่นโดยใช้สูตรคูเดอร์ - ริชาร์ดสัน (Kuder - Richardson) ต่อกำหนดข้อ 1 เข้ามา แล้วตัดข้อ 2 ออกเพียงหนึ่งข้อ จะเหลือข้อสอบทั้งหมด 99 ข้อเช่นเดิม และหาค่าความเชื่อมั่น ตัดข้อสอบออกทีละข้อเช่นนี้ไปจนถึงข้อ 100 และหาค่าความเชื่อมั่นทุกครั้งหลังจากตัดแต่ละข้อ

2.1.2.3 พิจารณาว่า ถ้าตัดข้อสอบข้อใดแล้ว ทำให้ค่าความเชื่อมั่นเพิ่มขึ้น ก็จะตัดข้อนั้นออกไปเลย ข้อสอบจะลดลง 1 ข้อ

2.1.2.4 จากข้อสอบที่เหลือ 99 ข้อ ขึ้นต่อไปตัดข้อสอบออกตามวิธีการในข้อ

2.1.2.2 จนครบ 99 ข้อ และพิจารณาค่าความเชื่อมั่นว่าถ้าตัดข้อสอบข้อใดแล้วทำให้ค่าความเชื่อมั่นของแบบทดสอบเพิ่มขึ้นก็จะตัดข้อนั้น

2.1.2.5 ทำการตัดข้อสอบออกทีละข้อ เมื่อถึงจุดจุดหนึ่งค่าความเชื่อมั่นจะคงที่หรือลดลงก็จะหยุดตัดข้อสอบ จำนวนข้อสอบที่เหลือจะเป็นข้อสอบที่คัดเลือกได้ตามเกณฑ์

2.1.3 วิเคราะห์ข้อสอบโดยใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ จากข้อสอบทั้งหมด 100 ข้อ มีขั้นตอนดังนี้

2.1.3.1 หาค่าความเชื่อมั่นของข้อสอบทั้งหมด 100 ข้อ

2.1.3.2 ตัดข้อสอบข้อ 1 ออก จะเหลือข้อสอบทั้งหมด 99 ข้อ หาค่าดัชนีความเชื่อมั่นของข้อสอบ (S_{rix}) ข้อ 2 จนถึงข้อ 100 แล้วนำดัชนีความเชื่อมั่นของข้อสอบที่ได้ทั้ง 99 ค่า ไปคำนวณหาค่าความเชื่อมั่น ($r_{xx'}$) โดยใช้สูตรคูเดอร์ - ริชาร์ดสัน (Kuder - Richardson) ต่อกำหนดข้อ 1 เข้ามาแล้วตัดข้อสอบข้อ 2 ออกเพียงหนึ่งข้อ จะเหลือข้อสอบทั้งหมด 99 ข้อเช่นเดิม หาค่าดัชนีความเชื่อมั่นของข้อสอบ (S_{rix}) ข้อ 1 ข้อ 3 จนถึงข้อ 100 แล้วนำดัชนีความเชื่อมั่นของข้อสอบที่ได้ทั้ง 99 ค่า ไปคำนวณค่าความเชื่อมั่น ($r_{xx'}$) ตัดข้อสอบออกทีละข้อจนถึงข้อ 100 และหาค่าดัชนีความเชื่อมั่นของข้อสอบ (S_{rix}) แล้วนำไปคำนวณค่าความเชื่อมั่น ($r_{xx'}$) ทุกครั้ง

2.1.3.3 พิจารณาว่า ถ้าตัดข้อสอบข้อใดแล้ว ทำให้ค่าความเชื่อมั่นเพิ่มขึ้น ก็จะตัดข้อนั้น ออกไปเลย ข้อสอบจะลดลง 1 ข้อ

2.1.3.4 จากข้อสอบที่เหลือ 99 ข้อ ขึ้นต่อไปตัดข้อสอบตามวิธีการในข้อ 2.1.3.2 จนครบ 99 ข้อ และพิจารณาค่าความเชื่อมั่นว่าถ้าตัดข้อใดแล้วทำให้ค่าความเชื่อมั่นเพิ่มขึ้น ก็จะตัดข้อนั้น

2.1.3.5 ทำการตัดข้อสอบออกทีละข้อ เมื่อถึงจุดจุดหนึ่งค่าความเชื่อมั่นจะคงที่หรือลดลง ก็จะหยุดตัดข้อสอบ จำนวนข้อสอบที่เหลือจะเป็นข้อสอบที่คัดเลือกได้ตามเกณฑ์

2.2 คัดเลือกข้อสอบจำนวน 50 ข้อ เพื่อศึกษาค่าความเชื่อมั่น

2.2.1 วิเคราะห์ข้อสอบโดยหาค่าอำนาจจำแนกแบบสหสัมพันธ์พอยท์ไบซีเรียล (Point Biserial Correlation) และหาค่าความยาก คัดข้อที่มีค่าอำนาจจำแนกตั้งแต่ .20 ขึ้นไป และมีค่าความยากตั้งแต่ .20 - .80 โดยคัดเลือกข้อที่มีค่าความยากปานกลาง และมีค่าอำนาจจำแนกสูง จำนวน 50 ข้อ

2.2.2 วิเคราะห์ข้อสอบโดยใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ จากการวิเคราะห์ข้อสอบ ข้อ 2.1.2.5 ให้ดำเนินการตัดข้อสอบต่อตามวิธีการข้างต้น โดยพิจารณาค่าความเชื่อมั่นของแบบทดสอบในแต่ละรอบว่าถ้าตัดข้อสอบข้อใดออกแล้วทำให้มีค่าความเชื่อมั่นสูงสุดในรอบนั้น ก็จะตัดข้อนั้นทิ้ง คัดเลือกข้อสอบให้เหลือ 50 ข้อ

2.2.3 วิเคราะห์ข้อสอบโดยใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ จากการวิเคราะห์ข้อสอบข้อ 2.1.3.5 ให้ดำเนินการตัดข้อสอบต่อตามวิธีการข้างต้น โดยพิจารณาค่าความเชื่อมั่นของแบบทดสอบในแต่ละรอบว่าถ้าตัดข้อสอบข้อใดออกแล้วทำให้มีค่าความเชื่อมั่นสูงสุดในรอบนั้น ก็จะตัดข้อนั้นทิ้ง คัดเลือกข้อสอบให้เหลือ 50 ข้อ

3. นำจำนวนข้อที่คัดเลือกได้ตามเกณฑ์และค่าความเชื่อมั่นของแบบทดสอบที่ได้จากแต่ละวิธี วิเคราะห์ข้อสอบ มาศึกษาดังนี้

3.1 เขียนกราฟเส้นภาพ (Profile)

3.2 ทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ จากการวิเคราะห์ข้อสอบต่างวิธี ในกลุ่มตัวอย่างเดียวกัน โดย Q-test และจากการวิเคราะห์ข้อสอบวิธีเดียวกัน ในกลุ่มตัวอย่างขนาดต่างกัน โดย χ^2 - test

3.3 ทดสอบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบที่วิเคราะห์ข้อสอบ ด้วยวิธีต่างกันโดยใช้ UX₁

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. หาค่าสถิติพื้นฐาน

1.1 คะแนนเฉลี่ย ($\bar{X}$)

1.2 ความเบี่ยงเบนมาตรฐาน (SD)

2. หาค่าอำนาจจำแนกโดยวิธีหาค่าสหสัมพันธ์แบบพอยท์ไบซีเรียล (Point Biserial Correlation) (Allen and Yen 1979 : 122)

$$r_{pbis} = \frac{\bar{X}_i - \bar{X}}{S_x} \cdot \sqrt{\frac{p_i}{1 - p_i}}$$

เมื่อ	r_{pbis}	แทน	ค่าสหสัมพันธ์
	$\bar{X}_i$	แทน	คะแนนเฉลี่ยของกลุ่มที่ทำข้อนั้นได้
	$\bar{X}$	แทน	คะแนนเฉลี่ยของผู้สอบทั้งหมด
	S_x	แทน	คะแนนเบี่ยงเบนมาตรฐานของแบบทดสอบทั้งฉบับ
	p_i	แทน	สัดส่วนของคนที่ทำข้อนั้นได้

3. หาค่าความยาก (ล้วน สายยศและอังคณา สายยศ. 2531 : 179)

$$P = \frac{R}{N}$$

เมื่อ	P	แทน	ค่าความยาก
	R	แทน	จำนวนคนที่ทำข้อนั้นถูก
	N	แทน	จำนวนคนที่ทำข้อนั้นทั้งหมด

4. วิเคราะห์ข้อสอบโดยใช้ดัชนีความเชื่อมั่นของข้อสอบ และหาค่าความเชื่อมั่นด้วยสูตร กูเดอร์ - ริชาร์ดสัน (Allen and Yen. 1979 : 124 - 125)

$$r_{xx} = \frac{k}{k-1} \left[1 - \frac{\sum_{i=1}^k S_i^2}{\left(\sum_{i=1}^k S_i r_{ix} \right)^2} \right]$$

เมื่อ	r_{xx}	แทน	ค่าความเชื่อมั่นของแบบทดสอบ
	k	แทน	จำนวนข้อสอบ
	S_i^2	แทน	ความแปรปรวนของคะแนนข้อที่ i
	$S_i r_{ix}$	แทน	ดัชนีความเชื่อมั่นของข้อสอบข้อที่ i

ค่าดัชนีความเชื่อมั่นของข้อสอบ หาโดย (Allen and Yen. 1979 : 124)

$$S_i = \sqrt{p_i (1 - p_i)}$$

เมื่อ	S_i	แทน	ส่วนเบี่ยงเบนมาตรฐานของคะแนนข้อที่ i
	p_i	แทน	ความยากของข้อสอบข้อที่ i

$$r_{ix} = \frac{\bar{X}_i - \bar{X}}{S_x} \sqrt{\frac{p_i}{1 - p_i}}$$

เมื่อ	r_{ix}	แทน	สหสัมพันธ์แบบพอยท์ไบซีเรียลระหว่างคะแนนข้อสอบแต่ละข้อ และคะแนนรวมของแบบทดสอบ
	$\bar{X}_i$	แทน	ค่าเฉลี่ยของคะแนน X ของผู้สอบที่ทำข้อสอบ i ได้ถูกต้อง
	$\bar{X}$	แทน	ค่าเฉลี่ยของผู้สอบทั้งหมด
	S_x	แทน	ส่วนเบี่ยงเบนมาตรฐานของคะแนน X
	p_i	แทน	ความยากของข้อสอบ

5. หาค่าความเชื่อมั่นของแบบทดสอบโดยวิธีใช้สูตร KR-20 ของคูเดอร์และริชาร์ดสัน (บุญเชิด ภิญโญนนตพงษ์. 2523 : 291 ; อ้างอิงมาจาก Kuder and Richardson. 1937)

$$r_{tt} = \frac{k}{k-1} \left[\frac{s_t^2 - \sum pq}{s_t^2} \right]$$

เมื่อ	r_{tt}	แทน	ค่าความเชื่อมั่นแบบทดสอบ
	k	แทน	จำนวนข้อสอบ
	p	แทน	อัตราส่วนของคนที่ทำถูกในแต่ละข้อ
	q	แทน	$1 - p$
	s_t^2	แทน	ความแปรปรวนของคะแนนที่สอบได้

6. สถิติทดสอบนัยสำคัญของความแตกต่างจำนวนข้อสอบ ที่ได้จากการวิเคราะห์ข้อสอบวิธีต่างกัน ในกลุ่มตัวอย่างเดียวกัน โดยใช้สูตร Q-Test (บุญเชิด ภิญโญนนตพงษ์. 2527 : 77 ; อ้างอิงมาจาก Hambleton et.al.1978 : 37 ; citing Brennan and Stolurow. 1971)

$$Q = \frac{(n-1)(n\sum p^2 - T^2)}{nT - \sum x^2}$$

Q	แทน	การทดสอบแบบคิว
n	แทน	จำนวนวิธีวิเคราะห์ข้อสอบ
p	แทน	จำนวนข้อสอบที่ถูกคัดเลือกแต่ละวิธีวิเคราะห์ข้อสอบ
x	แทน	คะแนนของข้อสอบที่ถูกคัดเลือก
T	แทน	ผลรวมของข้อสอบทุกข้อ

7. สถิติทดสอบนัยสำคัญของความแตกต่างจำนวนข้อสอบ ที่ได้จากการวิเคราะห์ข้อสอบวิธีเดียวกัน และกลุ่มตัวอย่างขนาดต่างกัน โดยใช้สูตร χ^2 - test (ชูศรี วงศ์รัตน์. 2537 : 219)

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

χ^2	แทน	การทดสอบแบบไค-กำลังสอง
O	แทน	จำนวนข้อสอบที่คัดเลือกได้
E	แทน	จำนวนข้อสอบที่คาดหวังที่จะคัดเลือกได้

8. สถิติทดสอบนัยสำคัญของความแตกต่างของค่าความเชื่อมั่นหลายค่า ซึ่งประมาณจากกลุ่มตัวอย่างกลุ่มเดียว ใช้สูตร UX_1 (บุญเชิด ภิญโญนนตพงษ์. 2537 : 57 ; อ้างอิงมาจาก Woodruff and Feldt . 1986 : 393 - 413)

$$UX_1 = \frac{\sum_{i=1}^m (u_i - \bar{u})^2}{s_u^2 - c_u}$$

$$\bar{u} = \frac{\sum_{i=1}^m u_i}{m}$$

$$u_i = \frac{1}{(1-r_i)^{\frac{1}{3}}}$$

$$S_u^2 = \frac{2}{9m(N_c - 1)} \sum_{i=1}^m u_i^2$$

$$C_u = \frac{4}{9m(m-1)(N_c - 1)} \sum_{i=2}^m \sum_{j=1}^{i-1} r_{ij}^2 u_i u_j$$

$$N_c = N \left(\frac{\bar{n}_h - 1}{\bar{n}_h + 1} \right)$$

$$n_h = \frac{M}{\sum \frac{1}{n_i}}$$

UX_1	แทน	สถิติทดสอบที่มีการแจกแจงแบบ χ^2 df = m - 1
r	แทน	ค่าความเชื่อมั่นที่คำนวณได้
r_{ij}	แทน	ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร
S_u^2	แทน	ค่าความแปรปรวนของ u_i
C_u	แทน	ค่าความแปรปรวนร่วม u_j

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การนำเสนอผลการวิเคราะห์ข้อมูล

การศึกษาวิจัยครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูล ตามลำดับดังนี้

1. ค่าสถิติพื้นฐานของแบบทดสอบจำนวน 100 ข้อ จากกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด คือขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน
2. ค่าสถิติพื้นฐานของแบบทดสอบที่คัดเลือกได้ตามเกณฑ์ จากการวิเคราะห์ข้อสอบต่างกัน 3 วิธี คือ วิธีหาค่าอำนาจจำแนกและค่าความยาก วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ และวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ จากกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด คือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน
3. การทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์
 - 3.1 การทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ จากการวิเคราะห์ข้อสอบ 3 วิธี ในแต่ละกลุ่มตัวอย่าง 6 ขนาด
 - 3.2 การทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ จากการวิเคราะห์ข้อสอบ 3 วิธี และกลุ่มตัวอย่างต่างกัน 6 ขนาด
4. ค่าสถิติพื้นฐานของแบบทดสอบจำนวน 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบ 3 วิธี ในแต่ละกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด
5. ค่าความเชื่อมั่นของแบบทดสอบจำนวน 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบต่างกัน 3 วิธี ในแต่ละกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด
6. การทดสอบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบจำนวน 50 ข้อ
 - 6.1 การทดสอบความแตกต่างระหว่างค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบ 3 วิธี ในแต่ละกลุ่มตัวอย่าง 6 ขนาด
 - 6.2 การทดสอบความแตกต่างระหว่างค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบ 3 วิธี และกลุ่มตัวอย่างต่างกัน 6 ขนาด

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลและการแปลความหมายของผลการวิเคราะห์ข้อมูล ผู้วิจัยได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล ดังนี้

$\bar{X}$ แทน คะแนนเฉลี่ย

SD แทน ความเบี่ยงเบนมาตรฐาน

วิธี 1 แทน วิธีวิเคราะห์ข้อสอบโดยการหาค่าอำนาจจำแนกและค่าความยาก

วิธี 2 แทน วิธีวิเคราะห์ข้อสอบด้วยโปรแกรมสำเร็จโดยลดจำนวนข้อ

วิธี 3 แทน วิธีวิเคราะห์ข้อสอบโดยใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ

Q แทน สถิติทดสอบที่มีการแจกแจงแบบไค-กำลังสอง

χ^2 แทน สถิติทดสอบที่มีการแจกแจงแบบไค-กำลังสอง

UX_i แทน สถิติทดสอบที่มีการแจกแจงแบบไค-กำลังสอง

P แทน ค่าระดับนัยสำคัญ

ผลการวิเคราะห์ข้อมูล

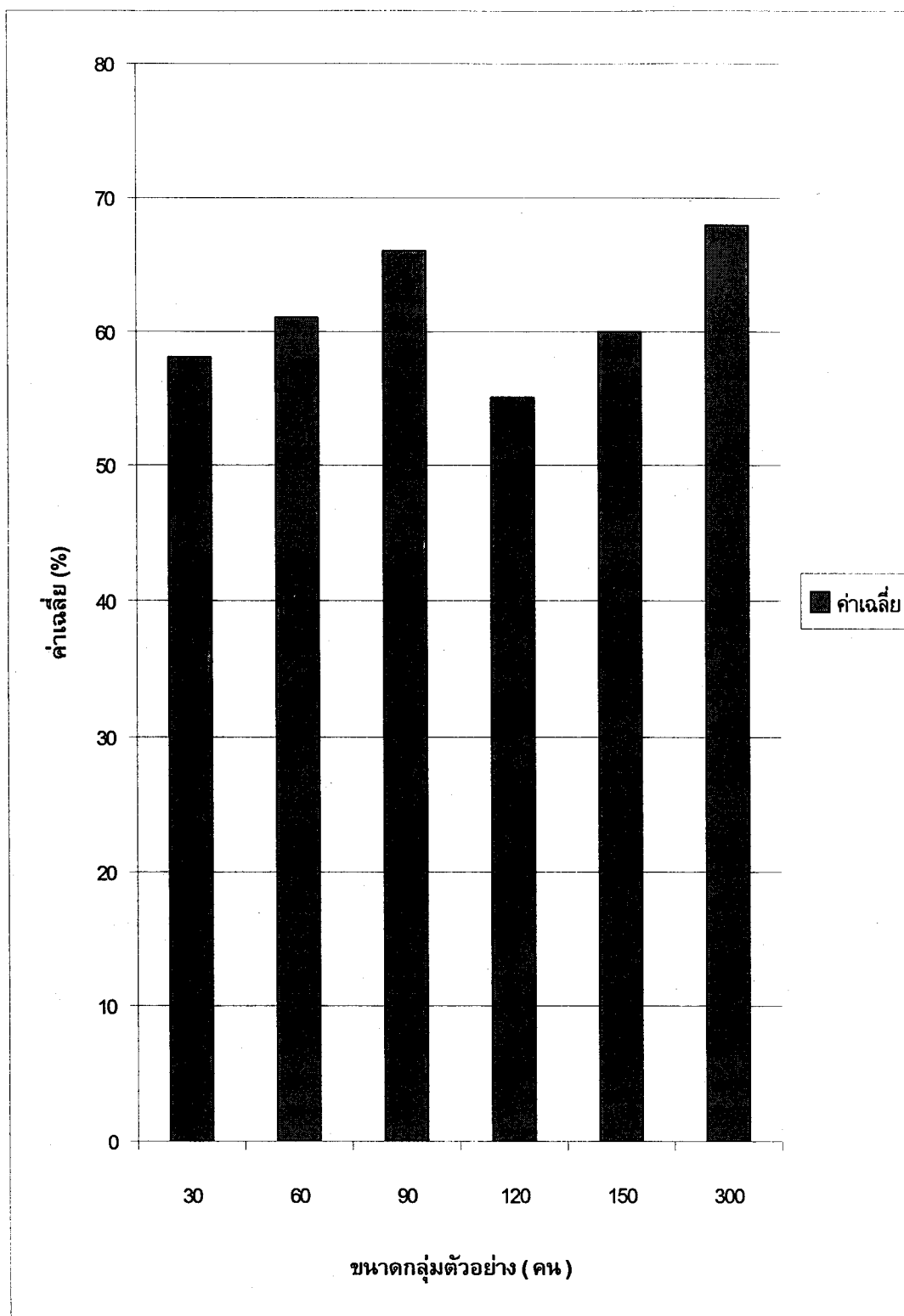
1. ค่าสถิติพื้นฐานของคะแนนแบบทดสอบจำนวน 100 ข้อ ที่ทดสอบจาก 6 กลุ่มตัวอย่าง

การวิเคราะห์ครั้งนี้ ได้จากการนำคะแนนผลการสอบด้วยแบบทดสอบที่สร้างขึ้นจำนวน 100 ข้อ กับกลุ่มตัวอย่าง 6 ขนาด คือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน มาหาค่าสถิติพื้นฐาน คือ ค่าเฉลี่ยและความเบี่ยงเบนมาตรฐานของคะแนนแบบทดสอบในกลุ่มตัวอย่างแต่ละขนาด ปรากฏดังแสดงในตาราง 2 และนำเสนอเป็นแผนภูมิแท่งดังแสดงในภาพประกอบ 1 ตามลำดับ

ตาราง 2 ค่าสถิติพื้นฐานของคะแนนที่ได้จากการทดสอบ

ขนาดกลุ่มตัวอย่าง (คน)	ความยาวของแบบทดสอบ (ข้อ)	$\bar{X}$	SD
30	100	57.967	19.352
60	100	60.767	19.671
90	100	65.478	17.903
120	100	55.050	17.248
150	100	60.033	17.080
300	100	67.630	17.276

ผลการวิเคราะห์ตามตาราง 2 พบว่า ค่าเฉลี่ยของแบบทดสอบจากกลุ่มตัวอย่างขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน มีค่าตั้งแต่ 55.050 ถึง 67.630 และค่าความเบี่ยงเบนมาตรฐานมีค่าตั้งแต่ 17.080 ถึง 19.671



ภาพประกอบ 1 ค่าคะแนนเฉลี่ยของแบบทดสอบจำนวน 100 ข้อ จากกลุ่มตัวอย่าง 6 ขนาด

2. ค่าสถิติพื้นฐานของแบบทดสอบที่คัดเลือกได้ตามเกณฑ์

การวิเคราะห์ครั้งนี้จะนำเสนอจากแบบทดสอบ จำนวน 100 ข้อ ที่ทดสอบจากกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด มาวิเคราะห์ข้อสอบด้วยวิธีการต่างกัน 3 วิธี แล้วคัดเลือกข้อที่มีคุณภาพตามเกณฑ์และคำนวณค่าสถิติปรากฏดังแสดงในตาราง 3 – 5 และนำเสนอเป็นแผนภูมิแท่งดังแสดงในภาพประกอบ 2 ตามลำดับ

ตาราง 3 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบทดสอบที่วิเคราะห์ข้อสอบวิธีหาค่าอำนาจจำแนกและค่าความยาก ที่คัดเลือกได้ตามเกณฑ์

ขนาดกลุ่มตัวอย่าง (คน)	ความยาวของแบบทดสอบ (ข้อ)	$\bar{X}$	SD
30	78	40.6333	17.2056
60	79	43.8500	17.5681
90	66	35.6000	15.4874
120	74	34.3750	14.7557
150	65	31.5933	13.5101
300	67	37.7700	15.1475

ผลการวิเคราะห์ตามตาราง 3 พบว่า ค่าเฉลี่ยของแบบทดสอบจากกลุ่มตัวอย่างขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน มีค่าตั้งแต่ 31.5933 ถึง 43.8500 และค่าความเบี่ยงเบนมาตรฐานมีค่าตั้งแต่ 13.5101 ถึง 17.5681

ตาราง 4 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบทดสอบที่วิเคราะห์ข้อสอบวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ ที่คัดเลือกได้ตามเกณฑ์

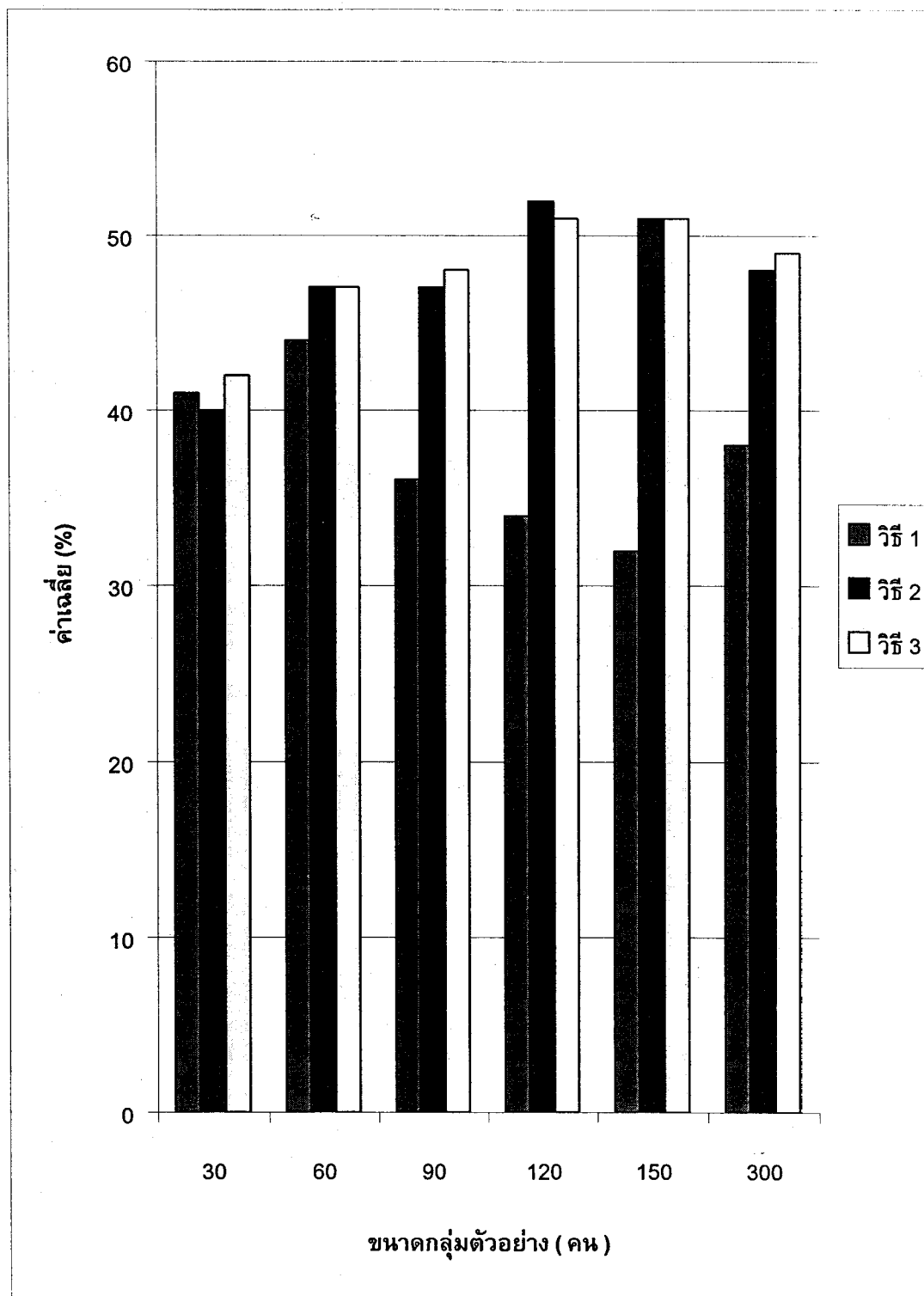
ขนาดกลุ่มตัวอย่าง (คน)	ความยาวของแบบทดสอบ (ข้อ)	$\bar{X}$	SD
30	71	40.3009	16.3309
60	77	46.9833	17.5010
90	78	47.0000	16.6881
120	95	52.8083	16.8888
150	85	50.6400	15.9652
300	77	47.6233	16.2089

ผลการวิเคราะห์ตามตาราง 4 พบว่า ค่าเฉลี่ยของแบบทดสอบจากกลุ่มตัวอย่างขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน มีค่าตั้งแต่ 40.3009 ถึง 52.8083 และค่าความเบี่ยงเบนมาตรฐานมีค่าตั้งแต่ 15.9652 ถึง 17.5010

ตาราง 5 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบทดสอบที่วิเคราะห์ข้อสอบวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ ที่คัดเลือกได้ตามเกณฑ์

ขนาดกลุ่มตัวอย่าง (คน)	ความยาวของแบบทดสอบ (ข้อ)	$\bar{X}$	SD
30	73	41.7667	16.5606
60	76	46.2500	17.3641
90	78	47.5444	16.4761
120	91	50.5083	16.4740
150	85	50.5800	16.0666
300	78	48.5133	16.2788

ผลการวิเคราะห์ตามตาราง 5 พบว่า ค่าเฉลี่ยของแบบทดสอบจากกลุ่มตัวอย่างขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน มีค่าตั้งแต่ 41.7667 ถึง 50.5800 และค่าความเบี่ยงเบนมาตรฐานมีค่าตั้งแต่ 16.0666 ถึง 17.3641



ภาพประกอบ 2 ค่าคะแนนเฉลี่ยของแบบทดสอบที่คัดเลือกได้ตามเกณฑ์ ที่วิเคราะห์ข้อสอบ 3 วิธี
จากกลุ่มตัวอย่าง 6 ขนาด

3. การทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์

การวิเคราะห์ครั้งนี้ได้นำจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ จากวิธีวิเคราะห์ข้อสอบต่างกัน 3 วิธี จากแต่ละกลุ่มตัวอย่าง 6 ขนาด มาทดสอบความแตกต่างจำนวนข้อที่ได้ ซึ่งแยกวิเคราะห์เปรียบเทียบระหว่างวิธีจากกลุ่มตัวอย่างเดียวกันโดยใช้ Q – test ทดสอบปรากฏผลดังตาราง 6 – 11 และวิเคราะห์เปรียบเทียบระหว่างกลุ่มโดยใช้ χ^2 ทดสอบ ปรากฏผลดังตาราง 12 – 14

3.1 การทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ จากการวิเคราะห์ข้อสอบ 3 วิธี ในกลุ่มตัวอย่างเดียวกัน

ตาราง 6 เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากการวิเคราะห์ข้อสอบวิธีต่างกัน 3 วิธี จากกลุ่มตัวอย่าง 30 คน

ข้อสอบที่ (ข้อ)	วิธี 1	วิธี 2	วิธี 3	Q
คัดเลือกไว้	78	71	73	2.437
ตัดออก	22	29	27	

ผลการวิเคราะห์ตามตาราง 6 พบว่า ข้อสอบที่วิเคราะห์ด้วยวิธีหาค่าอำนาจจำแนกและค่าความยาก คัดเลือกข้อสอบตามเกณฑ์ได้ 78 ข้อ ตัดออก 22 ข้อ ข้อสอบที่วิเคราะห์ด้วยวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ คัดเลือกข้อสอบตามเกณฑ์ได้ 71 ข้อ ตัดออก 29 ข้อ และข้อสอบที่วิเคราะห์ด้วยวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ คัดเลือกข้อสอบตามเกณฑ์ได้ 73 ข้อ ตัดออก 27 ข้อ เมื่อทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ด้วยสูตร Q – Test พบว่า แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า จำนวนข้อสอบที่ได้จากการวิเคราะห์ข้อสอบ 3 วิธี ไม่แตกต่างกัน

ตาราง 7 เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากการวิเคราะห์ข้อสอบวิธีต่างกัน 3 วิธี จากกลุ่มตัวอย่าง 60 คน

ข้อสอบที่ (ข้อ)	วิธี 1	วิธี 2	วิธี 3	Q
คัดเลือกไว้	79	77	76	0.5
ตัดออก	21	23	24	

ผลการวิเคราะห์ตามตาราง 7 พบว่า ข้อสอบที่วิเคราะห์ด้วยวิธีหาค่าอำนาจจำแนกและค่าความยาก คัดเลือกข้อสอบตามเกณฑ์ได้ 79 ข้อ ตัดออก 21 ข้อ ข้อสอบที่วิเคราะห์ด้วยวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ คัดเลือกข้อสอบตามเกณฑ์ได้ 77 ข้อ ตัดออก 23 ข้อ และข้อสอบที่วิเคราะห์ด้วยวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ คัดเลือกข้อสอบตามเกณฑ์ได้ 76 ข้อ ตัดออก 24 ข้อ เมื่อทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ด้วยสูตร Q – Test พบว่า แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า จำนวนข้อสอบที่ได้จากการวิเคราะห์ข้อสอบ 3 วิธี ไม่แตกต่างกัน

ตาราง 8 เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากการวิเคราะห์ข้อสอบวิธีต่างกัน 3 วิธี
จากกลุ่มตัวอย่าง 90 คน

ข้อสอบที่ (ข้อ)	วิธี 1	วิธี 2	วิธี 3	Q
คัดเลือกไว้	66	78	78	12.00*
ตัดออก	34	22	22	

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 8 พบว่า ข้อสอบที่วิเคราะห์ด้วยวิธีหาค่าอำนาจจำแนกและค่าความยาก คัดเลือกข้อสอบตามเกณฑ์ได้ 66 ข้อ ตัดออก 34 ข้อ ข้อสอบที่วิเคราะห์ด้วยวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ คัดเลือกข้อสอบตามเกณฑ์ได้ 78 ข้อ ตัดออก 22 ข้อ และข้อสอบที่วิเคราะห์ด้วยวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ คัดเลือกข้อสอบตามเกณฑ์ได้ 78 ข้อ ตัดออก 22 ข้อ เมื่อทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ด้วยสูตร Q – Test พบว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า จำนวนข้อสอบที่ได้จากการวิเคราะห์ข้อสอบ 3 วิธี แตกต่างกัน

ตาราง 9 เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากการวิเคราะห์ข้อสอบวิธีต่างกัน 3 วิธี
จากกลุ่มตัวอย่าง 120 คน

ข้อสอบที่ (ข้อ)	วิธี 1	วิธี 2	วิธี 3	Q
คัดเลือกไว้	74	95	91	27.6296*
ตัดออก	26	5	9	

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 9 พบว่า ข้อสอบที่วิเคราะห์ด้วยวิธีหาค่าอำนาจจำแนกและค่าความยาก คัดเลือกข้อสอบตามเกณฑ์ได้ 74 ข้อ ตัดออก 26 ข้อ ข้อสอบที่วิเคราะห์ด้วยวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ คัดเลือกข้อสอบตามเกณฑ์ได้ 95 ข้อ ตัดออก 5 ข้อ และข้อสอบที่วิเคราะห์ด้วยวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ คัดเลือกข้อสอบตามเกณฑ์ได้ 91 ข้อ ตัดออก 9 ข้อ เมื่อทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ด้วยสูตร Q – Test พบว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า จำนวนข้อสอบที่ได้จากการวิเคราะห์ข้อสอบ 3 วิธี แตกต่างกัน

ตาราง 10 เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากการวิเคราะห์ข้อสอบวิธีต่างกัน 3 วิธี
จากกลุ่มตัวอย่าง 150 คน

ข้อสอบที่ (ข้อ)	วิธี 1	วิธี 2	วิธี 3	Q
คัดเลือกไว้	65	85	85	25.8064*
ตัดออก	35	15	15	

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 10 พบว่า ข้อสอบที่วิเคราะห์ด้วยวิธีหาค่าอำนาจจำแนกและค่าความยาก คัดเลือกข้อสอบตามเกณฑ์ได้ 65 ข้อ ตัดออก 35 ข้อ ข้อสอบที่วิเคราะห์ด้วยวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ คัดเลือกข้อสอบตามเกณฑ์ได้ 85 ข้อ ตัดออก 15 ข้อ และข้อสอบที่วิเคราะห์ด้วยวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ คัดเลือกข้อสอบตามเกณฑ์ได้ 85 ข้อ ตัดออก 15 ข้อ เมื่อทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ด้วยสูตร Q – Test พบว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า จำนวนข้อสอบที่ได้จากการวิเคราะห์ข้อสอบ 3 วิธี แตกต่างกัน

ตาราง 11 เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากการวิเคราะห์ข้อสอบวิธีต่างกัน 3 วิธี
จากกลุ่มตัวอย่าง 300 คน

ข้อสอบที่ (ข้อ)	วิธี 1	วิธี 2	วิธี 3	Q
คัดเลือกไว้	67	77	78	13.0588*
ตัดออก	33	23	22	

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 11 พบว่า ข้อสอบที่วิเคราะห์ด้วยวิธีหาค่าอำนาจจำแนกและค่าความยาก คัดเลือกข้อสอบตามเกณฑ์ได้ 67 ข้อ ตัดออก 33 ข้อ ข้อสอบที่วิเคราะห์ด้วยวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ คัดเลือกข้อสอบตามเกณฑ์ได้ 77 ข้อ ตัดออก 23 ข้อ และข้อสอบที่วิเคราะห์ด้วยวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ คัดเลือกข้อสอบตามเกณฑ์ได้ 78 ข้อ ตัดออก 22 ข้อ เมื่อทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ด้วยสูตร Q – Test พบว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า จำนวนข้อสอบที่ได้จากการวิเคราะห์ข้อสอบ 3 วิธี แตกต่างกัน

3.2 การทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ จากวิธีวิเคราะห์เดียวกัน จากกลุ่มตัวอย่างต่างกัน 6 ขนาด

ตาราง 12 เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากวิธีหาค่าอำนาจจำแนกและค่าความยาก จากกลุ่มตัวอย่าง 6 ขนาด

ข้อสอบที่ (ข้อ)	ขนาดกลุ่มตัวอย่าง (คน)						χ^2
	30	60	90	120	150	300	
คัดเลือกไว้	78	79	66	74	65	67	59.42*
ตัดออก	22	21	34	26	35	33	

* มีนัยสำคัญที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 12 พบว่า ข้อสอบที่วิเคราะห์ด้วยวิธีหาค่าอำนาจจำแนกและค่าความยาก จากกลุ่มตัวอย่าง 30 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 78 ข้อ ตัดออก 22 ข้อ จากกลุ่มตัวอย่าง 60 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 79 ข้อ ตัดออก 21 ข้อ จากกลุ่มตัวอย่าง 90 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 66 ข้อ ตัดออก 34 ข้อ จากกลุ่มตัวอย่าง 120 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 74 ข้อ ตัดออก 26 ข้อ จากกลุ่มตัวอย่าง 150 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 65 ข้อ ตัดออก 35 ข้อ และจากกลุ่มตัวอย่าง 300 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 67 ข้อ ตัดออก 33 ข้อ เมื่อทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ด้วย χ^2 – test พบว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า จำนวนข้อสอบที่ได้จากการวิเคราะห์ข้อสอบวิธีหาค่าอำนาจจำแนกและค่าความยากจากกลุ่มตัวอย่าง 6 ขนาด แตกต่างกัน

ตาราง 13 เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ
จากกลุ่มตัวอย่าง 6 ขนาด

ข้อสอบที่ (ข้อ)	ขนาดกลุ่มตัวอย่าง (คน)						χ^2
	30	60	90	120	150	300	
คัดเลือกไว้	71	77	78	95	85	77	118.66*
ตัดออก	29	23	22	5	15	23	

* มีนัยสำคัญที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 13 พบว่า ข้อสอบที่วิเคราะห์ด้วยวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ จากกลุ่มตัวอย่าง 30 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 71 ข้อ ตัดออก 29 ข้อ จากกลุ่มตัวอย่าง 60 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 77 ข้อ ตัดออก 23 ข้อ จากกลุ่มตัวอย่าง 90 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 78 ข้อ ตัดออก 22 ข้อ จากกลุ่มตัวอย่าง 120 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 95 ข้อ ตัดออก 5 ข้อ จากกลุ่มตัวอย่าง 150 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 85 ข้อ ตัดออก 15 ข้อ และจากกลุ่มตัวอย่าง 300 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 77 ข้อ ตัดออก 23 ข้อ เมื่อทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ด้วย χ^2 - test พบว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า จำนวนข้อสอบที่ได้จากการวิเคราะห์ข้อสอบวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อจากกลุ่มตัวอย่าง 6 ขนาด แตกต่างกัน

ตาราง 14 เปรียบเทียบจำนวนข้อสอบที่คัดเลือกตามเกณฑ์จากวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ
จากกลุ่มตัวอย่าง 6 ขนาด

ข้อสอบที่ (ข้อ)	ขนาดกลุ่มตัวอย่าง (คน)						χ^2
	30	60	90	120	150	300	
คัดเลือกไว้	73	76	78	91	85	78	113.58*
ตัดออก	27	24	22	9	15	22	

* มีนัยสำคัญที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 14 พบว่า ข้อสอบที่วิเคราะห์ด้วยวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ จากกลุ่มตัวอย่าง 30 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 73 ข้อ ตัดออก 27 ข้อ จากกลุ่มตัวอย่าง 60 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 76 ข้อ ตัดออก 24 ข้อ จากกลุ่มตัวอย่าง 90 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 78 ข้อ ตัดออก 22 ข้อ จากกลุ่มตัวอย่าง 120 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 91 ข้อ ตัดออก 9 ข้อ จากกลุ่มตัวอย่าง 150 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 85 ข้อ ตัดออก 15 ข้อ และจากกลุ่มตัวอย่าง 300 คน คัดเลือกข้อสอบตามเกณฑ์ได้ 78 ข้อ ตัดออก 22 ข้อ เมื่อทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ด้วย χ^2 -test พบว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า จำนวนข้อสอบที่ได้จากการวิเคราะห์ข้อสอบวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อจากกลุ่มตัวอย่าง 6 ขนาดแตกต่างกัน

4. ค่าสถิติพื้นฐานของคะแนนจากแบบทดสอบ 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบต่างกัน 3 วิธี ในแต่ละกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด

การวิเคราะห์ครั้งนี้ ได้นำข้อสอบจำนวน 50 ข้อแรกที่ดีที่สุดซึ่งได้จากการวิเคราะห์ข้อสอบ 3 วิธี คือ วิธีหาค่าอำนาจจำแนกและค่าความยาก วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ และวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ จากกลุ่มตัวอย่าง 6 ขนาด มาหาค่าสถิติพื้นฐาน คือ ค่าเฉลี่ย และความเบี่ยงเบนมาตรฐาน ปรากฏผลดังแสดงในตาราง 15 – 17 และนำเสนอเป็นแผนภูมิแท่งดังแสดงในภาพประกอบ 3 ตามลำดับ

ตาราง 15 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบทดสอบที่วิเคราะห์ข้อสอบโดยการหาค่าอำนาจจำแนก และค่าความยาก จำนวน 50 ข้อ

ขนาดกลุ่มตัวอย่าง (คน)	ความยาวของแบบทดสอบ (ข้อ)	$\bar{X}$	SD
30	50	24.6000	12.8937
60	50	27.8166	13.0650
90	50	27.0000	12.4313
120	50	23.3083	11.0907
150	50	25.5866	11.2972
300	50	27.1966	12.1723

ผลการวิเคราะห์ตามตาราง 15 พบว่า ค่าเฉลี่ยของแบบทดสอบจากกลุ่มตัวอย่างขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน มีค่าตั้งแต่ 23.3083 ถึง 27.8166 และค่าความเบี่ยงเบนมาตรฐานมีค่าตั้งแต่ 11.0907 ถึง 13.0650

ตาราง 16 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบทดสอบที่วิเคราะห์ด้วยโปรแกรมสำเร็จโดยลดจำนวนข้อ
จำนวน 50 ข้อ

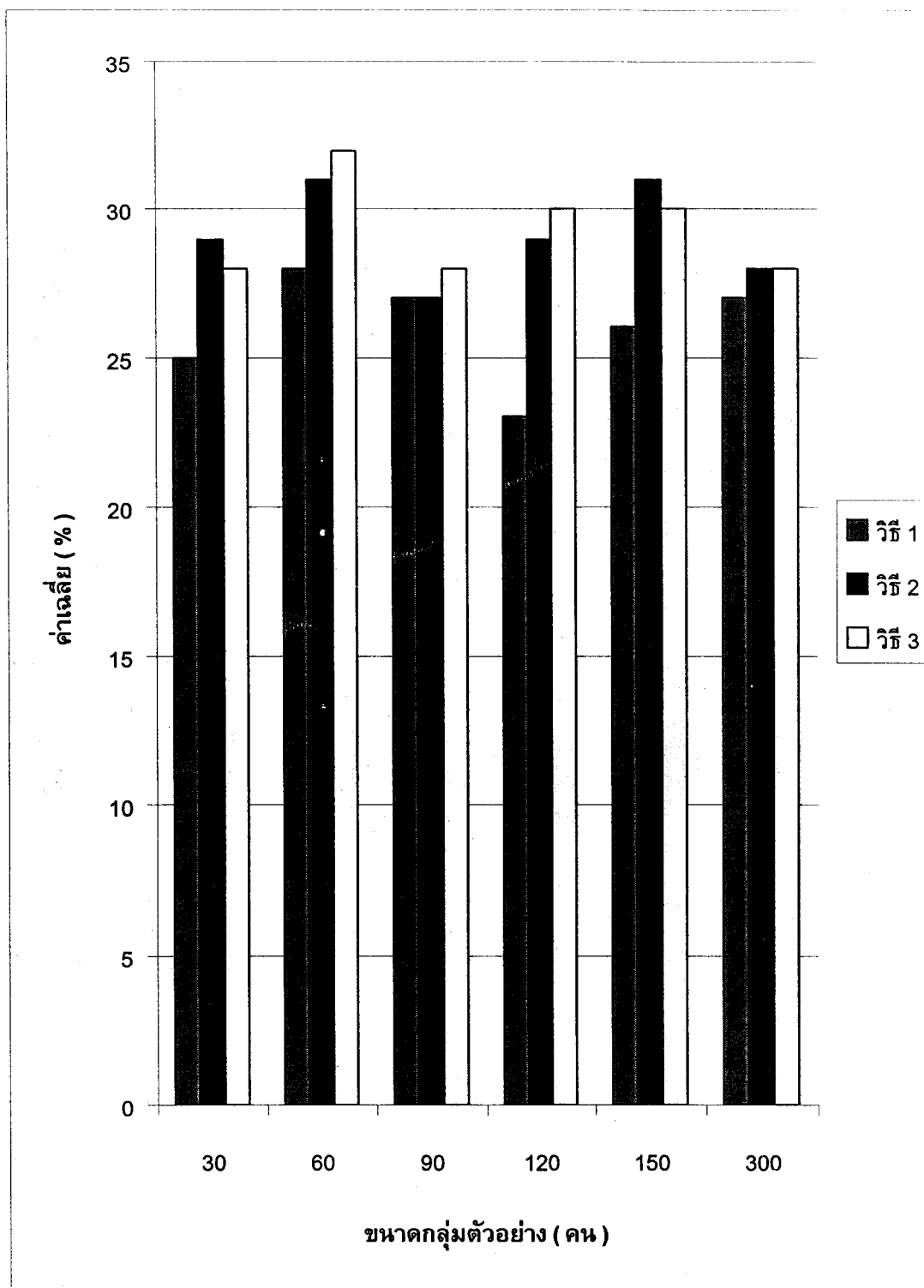
ขนาดกลุ่มตัวอย่าง (คน)	ความยาวของแบบทดสอบ (ข้อ)	$\bar{X}$	SD
30	50	28.9000	13.1394
60	50	31.0666	12.9965
90	50	26.9777	12.6659
120	50	28.5000	11.0765
150	50	30.5133	10.7770
300	50	28.2066	12.2530

ผลการวิเคราะห์ตามตาราง 16 พบว่า ค่าเฉลี่ยของแบบทดสอบจากกลุ่มตัวอย่างขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน มีค่าตั้งแต่ 26.9777 ถึง 31.0666 และค่าความเบี่ยงเบนมาตรฐานมีค่าตั้งแต่ 10.7770 ถึง 13.1394

ตาราง 17 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบทดสอบที่วิเคราะห์โดยใช้ดัชนีความเชื่อมั่นของข้อสอบ
แต่ละข้อ จำนวน 50 ข้อ

ขนาดกลุ่มตัวอย่าง (คน)	ความยาวของแบบทดสอบ (ข้อ)	$\bar{X}$	SD
30	50	27.8000	13.1945
60	50	31.6166	12.9157
90	50	27.9000	12.5358
120	50	30.2416	10.7429
150	50	29.7066	10.8445
300	50	28.0033	12.3555

ผลการวิเคราะห์ตามตาราง 17 พบว่า ค่าเฉลี่ยของแบบทดสอบจากกลุ่มตัวอย่างขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน มีค่าตั้งแต่ 27.8000 ถึง 31.6166 และค่าความเบี่ยงเบนมาตรฐานมีค่าตั้งแต่ 10.7429 ถึง 13.1945



ภาพประกอบ 3 ค่าคะแนนเฉลี่ยของข้อสอบ 50 ข้อแรกที่ดีที่สุดในวิเคราะห์ข้อสอบ 3 วิธี จากกลุ่มตัวอย่าง 6 ขนาด

5. ค่าความเชื่อมั่นของแบบทดสอบจำนวน 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบต่างกัน 3 วิธี ในแต่ละกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด

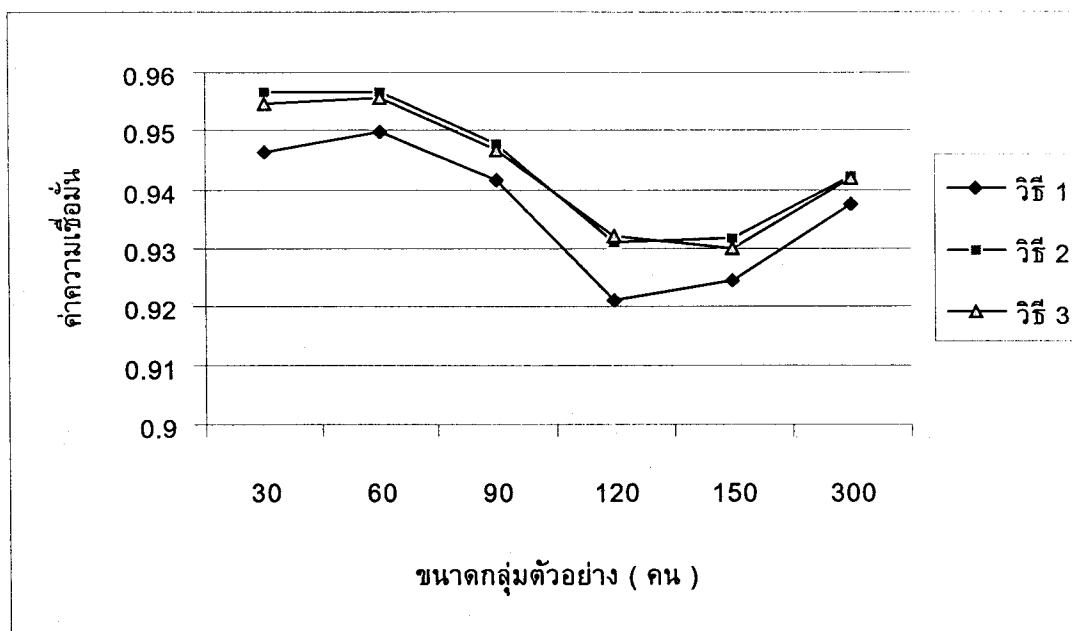
การวิเคราะห์ครั้งนี้ เป็นการวิเคราะห์หาความเชื่อมั่นของแบบทดสอบจำนวน 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบ 3 วิธี จากกลุ่มตัวอย่างที่มีขนาดต่างกัน 6 ขนาด ปรากฏผลดังแสดงในตาราง 18

ตาราง 18 ค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ที่วิเคราะห์ข้อสอบ 3 วิธี จากกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด

วิธีวิเคราะห์ข้อสอบ	กลุ่มตัวอย่าง (คน)					
	30	60	90	120	150	300
1	.9462	.9498	.9415	.9211	.9247	.9374
2	.9565	.9565	.9476	.9309	.9316	.9423
3	.9545	.9555	.9467	.9320	.9299	.9419

ผลการวิเคราะห์ตามตาราง 18 พบว่า ค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ ที่ได้จากการวิเคราะห์ข้อสอบทั้ง 3 วิธี จากกลุ่มตัวอย่าง 6 ขนาด มีผลดังนี้ กลุ่มตัวอย่างขนาดเล็ก 30 คน ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบวิธีหาอำนาจจำแนกและค่าความยากได้ค่าต่ำสุด .9462 ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อได้ค่าสูงสุด .9565 กลุ่มตัวอย่างขนาดเล็ก 60 คน ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบวิธีหาอำนาจจำแนกและค่าความยากได้ค่าต่ำสุด .9498 ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อได้ค่าสูงสุด .9565 กลุ่มตัวอย่างขนาดเล็ก 90 คน ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบวิธีหาอำนาจจำแนกและค่าความยากได้ค่าต่ำสุด .9415 ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อได้ค่าสูงสุด .9476 กลุ่มตัวอย่างขนาดเล็ก 120 คน ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบวิธีหาอำนาจจำแนกและค่าความยากได้ค่าต่ำสุด .9211 ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อได้ค่าสูงสุด .9320 กลุ่มตัวอย่างขนาดเล็ก 150 คน ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบวิธีหาอำนาจจำแนกและค่าความยากได้ค่าต่ำสุด .9247 ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ ได้ค่าสูงสุด .9316 กลุ่มตัวอย่างขนาดใหญ่ 300 คน ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบวิธีหาอำนาจจำแนกและค่าความยากได้ค่าต่ำสุด .9374 ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อได้ค่าสูงสุด .9423

เพื่อให้เห็นภาพที่ชัดเจน ผู้วิจัยได้นำค่าความเชื่อมั่นของแบบทดสอบจากตาราง 18 มานำเสนอเป็นเส้นภาพ ดังแสดงในภาพประกอบ 4



ภาพประกอบ 4 เส้นภาพค่าความเชื่อมั่นของแบบทดสอบที่วิเคราะห์ข้อสอบวิธีต่างกัน 3 วิธี จากกลุ่มตัวอย่าง 6 ขนาด

6. การทดสอบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบ

6.1 นำค่าความเชื่อมั่นของแบบทดสอบจำนวน 50 ข้อ ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี ในแต่ละขนาดกลุ่มตัวอย่าง 6 ขนาด มาทดสอบความแตกต่างปรากฏผลดังแสดงในตาราง 19 - 24

ตาราง 19 เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์จำนวน 50 ข้อ ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 30 คน

วิธีวิเคราะห์ข้อสอบ	ค่าความเชื่อมั่น	UX ₁ (dependent)	P
1	.9462	.3424	< .64693
2	.9565		
3	.9545		

ผลการวิเคราะห์ตามตาราง 19 พบว่า ความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี ให้ค่าเรียงลำดับจากน้อยไปหามากดังนี้ วิวิธีวิเคราะห์ข้อสอบโดยหาค่าอำนาจจำแนกและค่าความยาก เท่ากับ .9462 วิวิธีวิเคราะห์ข้อสอบโดยใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ เท่ากับ .9545 และวิวิธีวิเคราะห์ข้อสอบด้วยโปรแกรมสำเร็จโดยลดจำนวนข้อ เท่ากับ .9565 ตามลำดับ เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่นแต่ละค่าพร้อมกันด้วยสูตร UX_1 พบว่า แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .01 ($UX_1 = .3424$, $P = .64693$) แสดงว่า ความเชื่อมั่นของแบบทดสอบไม่แตกต่างกัน ไม่ว่าจะคัดเลือกข้อสอบเข้าฉบับด้วยวิวิธีวิเคราะห์ข้อสอบแบบใด

ตาราง 20 เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์จำนวน 50 ข้อ ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 60 คน

วิวิธีวิเคราะห์ข้อสอบ	ค่าความเชื่อมั่น	UX_1 (dependent)	P
1	.9498	.3311	< .64107
2	.9565		
3	.9555		

ผลการวิเคราะห์ตามตาราง 20 พบว่า ค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี ให้ค่าเรียงลำดับจากน้อยไปหามากดังนี้ วิวิธีวิเคราะห์ข้อสอบโดยหาค่าอำนาจจำแนกและค่าความยาก เท่ากับ .9498 วิวิธีวิเคราะห์ข้อสอบโดยใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ เท่ากับ .9555 และวิวิธีวิเคราะห์ข้อสอบด้วยโปรแกรมสำเร็จโดยลดจำนวนข้อ เท่ากับ .9565 ตามลำดับ เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่นแต่ละค่าพร้อมกันด้วยสูตร UX_1 พบว่า แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .01 ($UX_1 = .3311$, $P = .64107$) แสดงว่า ความเชื่อมั่นของแบบทดสอบไม่แตกต่างกัน ไม่ว่าจะคัดเลือกข้อสอบเข้าฉบับด้วยวิวิธีวิเคราะห์ข้อสอบแบบใด

ตาราง 21 เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์จำนวน 50 ข้อ
ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 90 คน

วิธีวิเคราะห์ข้อสอบ	ค่าความเชื่อมั่น	UX ₁ (dependent)	P
1	.9415	.2971	< .62019
2	.9476		
3	.9467		

ผลการวิเคราะห์ตามตาราง 21 พบว่า ค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี ให้ค่าเรียงลำดับจากน้อยไปหามากดังนี้ วิธีวิเคราะห์ข้อสอบโดยหาค่าอำนาจจำแนกและค่าความยาก เท่ากับ .9415 วิธีวิเคราะห์ข้อสอบโดยใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ เท่ากับ .9467 และวิธีวิเคราะห์ข้อสอบด้วยโปรแกรมสำเร็จโดยลดจำนวนข้อ เท่ากับ .9476 ตามลำดับ เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่นแต่ละค่าพร้อมกันด้วยสูตร UX₁ พบว่า แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .01 (UX₁ = .2971 , P = .62019) แสดงว่า ความเชื่อมั่นของแบบทดสอบไม่แตกต่างกัน ไม่ว่าจะคัดเลือกข้อสอบเข้าฉบับด้วยวิธีวิเคราะห์ข้อสอบแบบใด

ตาราง 22 เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์จำนวน 50 ข้อ
ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 120 คน

วิธีวิเคราะห์ข้อสอบ	ค่าความเชื่อมั่น	UX ₁ (dependent)	P
1	.9211	.7491	< .66482
2	.9309		
3	.9320		

ผลการวิเคราะห์ตามตาราง 22 พบว่า ค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี ให้ค่าเรียงลำดับจากน้อยไปหามากดังนี้ วิธีวิเคราะห์ข้อสอบโดยหาค่าอำนาจจำแนกและค่าความยาก เท่ากับ .9211 วิธีวิเคราะห์ข้อสอบด้วยโปรแกรมสำเร็จโดยลดจำนวนข้อ เท่ากับ .9309 และวิธีวิเคราะห์ข้อสอบโดยใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ เท่ากับ .9320 ตามลำดับ เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่นแต่ละค่าพร้อมกันด้วยสูตร UX₁ พบว่า แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .01 (UX₁ = .7491, P = .66482) แสดงว่า ความเชื่อมั่นของแบบทดสอบ

ไม่แตกต่างกัน ไม่ว่าจะคัดเลือกข้อสอบเข้าฉบับด้วยวิธีวิเคราะห์ข้อสอบแบบใด

ตาราง 23 เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์จำนวน 50 ข้อ
ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 150 คน

วิธีวิเคราะห์ข้อสอบ	ค่าความเชื่อมั่น	UX ₁ (dependent)	P
1	.9247	.3542	< .65255
2	.9316		
3	.9299		

ผลการวิเคราะห์ตามตาราง 23 พบว่า ค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี ให้ค่าเรียงลำดับจากน้อยไปหามากดังนี้ วิธีวิเคราะห์ข้อสอบโดยหาค่าอำนาจจำแนกและค่าความยาก เท่ากับ .9247 วิธีวิเคราะห์ข้อสอบโดยใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ เท่ากับ .9299 และวิธีวิเคราะห์ข้อสอบด้วยโปรแกรมสำเร็จโดยลดจำนวนข้อ เท่ากับ .9316 ตามลำดับ เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่นแต่ละค่าพร้อมกันด้วยสูตร UX₁ พบว่า แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .01 (UX₁ = .3542, P = .65255) แสดงว่า ความเชื่อมั่นของแบบทดสอบไม่แตกต่างกัน ไม่ว่าจะคัดเลือกข้อสอบเข้าฉบับด้วยวิธีวิเคราะห์ข้อสอบแบบใด

ตาราง 24 เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์จำนวน 50 ข้อ
ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาดใหญ่ 300 คน

วิธีวิเคราะห์ข้อสอบ	ค่าความเชื่อมั่น	UX ₁ (dependent)	P
1	.9374	.5817	< .68746
2	.9423		
3	.9419		

ผลการวิเคราะห์ตามตาราง 24 พบว่า ค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ที่วิเคราะห์ข้อสอบต่างกัน 3 วิธี ให้ค่าเรียงลำดับจากน้อยไปหามากดังนี้ วิธีวิเคราะห์ข้อสอบโดยหาค่าอำนาจจำแนกและค่าความยาก เท่ากับ .9374 วิธีวิเคราะห์ข้อสอบโดยใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ

เท่ากับ .9419 และวิธีวิเคราะห์ข้อสอบด้วยโปรแกรมสำเร็จโดยลดจำนวนข้อ เท่ากับ .9423 ตามลำดับ เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่นแต่ละค่าพร้อมกันด้วยสูตร UX_1 พบว่า แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .01 ($UX_1 = .5817$, $P = .68746$) แสดงว่า ความเชื่อมั่นของแบบทดสอบ ไม่แตกต่างกัน ไม่ว่าจะคัดเลือกข้อสอบเข้าฉบับด้วยวิธีวิเคราะห์ข้อสอบแบบใด

6.2 นำค่าความเชื่อมั่นของแบบทดสอบจำนวน 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบ 3 วิธี จากกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด มาทดสอบความแตกต่างปรากฏผลดังแสดงในตาราง 25 - 27

ตาราง 25 เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์จำนวน 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบวิธีหาค่าอำนาจจำแนกและค่าความยาก ระหว่างกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด

ขนาดกลุ่มตัวอย่าง (คน)	ค่าความเชื่อมั่น	UX_1 (Independent)	P
30	.9462	5.4426	< .36422
60	.9498		
90	.9415		
120	.9211		
150	.9247		
300	.9374		

ผลการวิเคราะห์ตามตาราง 25 พบว่า ค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ ที่วิเคราะห์ข้อสอบวิธีหาค่าอำนาจจำแนกและค่าความยาก จากกลุ่มตัวอย่างขนาดต่างกัน ให้ค่าเรียงลำดับจากน้อยไปหามาก ดังนี้ กลุ่มตัวอย่างขนาดเล็ก 120 คน เท่ากับ .9211 กลุ่มตัวอย่างขนาดเล็ก 150 คน เท่ากับ .9247 กลุ่มตัวอย่างขนาดใหญ่ 300 คน เท่ากับ .9374 กลุ่มตัวอย่างขนาดเล็ก 90 คน เท่ากับ .9415 กลุ่มตัวอย่างขนาดเล็ก 30 คน เท่ากับ .9462 และกลุ่มตัวอย่างขนาดเล็ก 60 คน เท่ากับ .9498 ตามลำดับ เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่นแต่ละค่าพร้อมกันด้วยสูตร UX_1 พบว่า แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .01 ($UX_1 = 5.4426$, $P = .36422$) แสดงว่า ความเชื่อมั่นของแบบทดสอบไม่แตกต่างกัน ไม่ว่าจะใช้กลุ่มตัวอย่างขนาดใด

ตาราง 26 เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์จำนวน 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ ระหว่างกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด

ขนาดกลุ่มตัวอย่าง (คน)	ค่าความเชื่อมั่น	UX_1 (Independent)	P
30	.9565	7.0564	< .21538
60	.9565		
90	.9476		
120	.9309		
150	.9316		
300	.9423		

ผลการวิเคราะห์ตามตาราง 26 พบว่า ค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ที่วิเคราะห์ข้อสอบวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ จากกลุ่มตัวอย่างขนาดต่างกัน ให้ค่าเรียงลำดับจากน้อยไปหามากดังนี้ กลุ่มตัวอย่างขนาดเล็ก 120 คน เท่ากับ .9309 กลุ่มตัวอย่างขนาดเล็ก 150 คน เท่ากับ .9316 กลุ่มตัวอย่างขนาดใหญ่ 300 คน เท่ากับ .9423 กลุ่มตัวอย่างขนาดเล็ก 90 คน เท่ากับ .9476 กลุ่มตัวอย่างขนาดเล็ก 30 คน เท่ากับ .9565 และกลุ่มตัวอย่างขนาดเล็ก 60 คน เท่ากับ .9565 ตามลำดับ เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่นแต่ละค่าพร้อมกันด้วยสูตร UX_1 พบว่า แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .01 ($UX_1 = 7.0564$, $P = .21538$) แสดงว่า ความเชื่อมั่นของแบบทดสอบไม่แตกต่างกัน ไม่ว่าจะใช้กลุ่มตัวอย่างขนาดใด

ตาราง 27 เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์จำนวน 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อระหว่างกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด

ขนาดกลุ่มตัวอย่าง (คน)	ค่าความเชื่อมั่น	UX_1 (Independent)	P
30	.9545	6.2153	< .28495
60	.9555		
90	.9467		
120	.9320		
150	.9299		
300	.9419		

ผลการวิเคราะห์ตามตาราง 27 พบว่า ค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ ที่วิเคราะห์ข้อสอบวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ จากขนาดกลุ่มตัวอย่างต่างกันให้ค่าเรียงลำดับจากน้อยไปหามาก ดังนี้ กลุ่มตัวอย่างขนาดเล็ก 150 คน เท่ากับ .9299 กลุ่มตัวอย่างขนาดเล็ก 120 คน เท่ากับ .9320 กลุ่มตัวอย่างขนาดใหญ่ 300 คน เท่ากับ .9419 กลุ่มตัวอย่างขนาดเล็ก 90 คน เท่ากับ .9467 กลุ่มตัวอย่างขนาดเล็ก 30 คน เท่ากับ .9545 และกลุ่มตัวอย่างขนาดเล็ก 60 คน เท่ากับ .9555 ตามลำดับ เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่นแต่ละค่าพร้อมกันด้วยสูตร UX_1 พบว่า แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .01 ($UX_1 = 6.2153, P = .28495$) แสดงว่า ความเชื่อมั่นของแบบทดสอบไม่แตกต่างกัน ไม่ว่าจะใช้กลุ่มตัวอย่างขนาดใด

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อเปรียบเทียบจำนวนข้อของแบบทดสอบความถนัดด้านคณิตศาสตร์ ที่คัดเลือกได้ตามเกณฑ์จากวิธีวิเคราะห์ข้อสอบต่างกัน 3 วิธี คือ 1) วิธีหาค่าอำนาจจำแนกแบบสหสัมพันธ์พอยท์ไบซีเรียล (Point Biserial Correlation) และหาค่าความยาก (Item Difficulty) 2) วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ 3) วิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ โดยศึกษาแต่ละขนาดกลุ่มตัวอย่างคือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน

2. เพื่อเปรียบเทียบจำนวนข้อของแบบทดสอบความถนัดด้านคณิตศาสตร์ ที่คัดเลือกได้ตามเกณฑ์ ระหว่างกลุ่มตัวอย่าง คือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน ที่คัดเลือกข้อสอบ 3 วิธี คือ 1) วิธีหาค่าอำนาจจำแนกแบบสหสัมพันธ์พอยท์ไบซีเรียล (Point Biserial Correlation) และหาค่าความยาก (Item Difficulty) 2) วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ 3) วิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ

3. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ จำนวน 50 ข้อ ระหว่างวิธีคัดเลือกข้อสอบด้วยการวิเคราะห์ข้อสอบต่างกัน 3 วิธี คือ 1) วิธีหาค่าอำนาจจำแนกแบบสหสัมพันธ์พอยท์ไบซีเรียล (Point Biserial Correlation) และหาค่าความยาก (Item Difficulty) 2) วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ 3) วิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ ในแต่ละขนาดกลุ่มตัวอย่าง คือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และกลุ่มตัวอย่างขนาดใหญ่ 300 คน

4. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความถนัดด้านคณิตศาสตร์ จำนวน 50 ข้อ ระหว่างกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด คือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน ในแต่ละวิธีคัดเลือกข้อสอบ 3 วิธี คือ 1) วิธีหาค่าอำนาจจำแนกแบบสหสัมพันธ์พอยท์ไบซีเรียล (Point Biserial Correlation) และหาค่าความยาก (Item Difficulty) 2) วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ 3) วิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้า

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2540 จำนวน 750 คน ซึ่งได้มาโดยวิธีการสุ่มแบบสองขั้นตอน (Two-Stage Random Sampling)

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าเป็นแบบทดสอบความถนัดด้านคณิตศาสตร์ ซึ่งสร้างตามแนวของเรอร์สโตน แบ่งเป็น 4 ประเภท คือ ทักษะ ความคิดรวบยอด โจทย์ปัญหา และเรียงอันดับ ประเภทละ 25 ข้อ รวม 100 ข้อ เป็นแบบทดสอบปรนัยชนิดเลือกตอบ 5 ตัวเลือก ซึ่งผู้วิจัยสร้างขึ้นเอง

การเสนอผลวิเคราะห์ข้อมูล

การศึกษาวิจัยครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูล ตามลำดับดังนี้

1. ค่าสถิติพื้นฐานของแบบทดสอบจำนวน 100 ข้อ จากขนาดกลุ่มตัวอย่างต่างกัน 6 ขนาด
2. ค่าสถิติพื้นฐานของแบบทดสอบที่คัดเลือกได้ตามเกณฑ์ จากการวิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างต่างกัน 6 ขนาด
3. การทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์
 - 3.1 การทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ จากการวิเคราะห์ข้อสอบต่างกัน 3 วิธี ในแต่ละขนาดกลุ่มตัวอย่าง 6 ขนาด
 - 3.2 การทดสอบความแตกต่างของจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ ระหว่างกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด ที่วิเคราะห์ข้อสอบ 3 วิธี
4. ค่าสถิติพื้นฐานของแบบทดสอบจำนวน 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด
5. ค่าความเชื่อมั่นของแบบทดสอบจำนวน 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด
6. การทดสอบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบจำนวน 50 ข้อ
 - 6.1 การทดสอบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบที่วิเคราะห์ข้อสอบ 3 วิธี ในแต่ละขนาดกลุ่มตัวอย่าง 6 ขนาด
 - 6.2 การทดสอบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบระหว่างกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด ที่วิเคราะห์ข้อสอบ 3 วิธี

สรุปผลการวิจัย

ผลการวิจัยสรุปได้ดังตาราง

ตาราง 28 จำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์และค่าความเชื่อมั่นของแบบทดสอบที่วิเคราะห์ข้อสอบ 3 วิธี จากกลุ่มตัวอย่าง 6 ขนาด

ขนาดกลุ่มตัวอย่าง (คน)	จำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ (ข้อ) / ค่าความเชื่อมั่น		
	วิธี 1	วิธี 2	วิธี 3
30	78 / .9556	71 / .9569	73 / .9550
60	79 / .9613	77 / .9600	76 / .9551
90	66 / .9509	78 / .9536	78 / .9522
120	74 / .9353	95 / .9438	91 / .9437
150	65 / .9342	85 / .9456	85 / .9442
300	67 / .9507	77 / .9503	78 / .9497

1. เปรียบเทียบจำนวนข้อสอบ ที่คัดเลือกได้ตามเกณฑ์ จากการวิเคราะห์ข้อสอบ 3 วิธี คือ วิธีหาค่าอำนาจจำแนกและค่าความยาก วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ และวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ ในแต่ละขนาดกลุ่มตัวอย่างโดยใช้ Q-Test ทดสอบ สรุปว่า วิธีการวิเคราะห์ข้อสอบ 3 วิธี คัดเลือกข้อสอบได้จำนวนไม่แตกต่างกัน เมื่อวิเคราะห์จากกลุ่มตัวอย่างขนาดเล็ก 30 คน และ 60 คน แต่คัดเลือกข้อสอบได้จำนวนแตกต่างกันเมื่อวิเคราะห์จากกลุ่มตัวอย่างขนาดเล็ก 90 คน 120 คน 150 คน และกลุ่มตัวอย่างขนาดใหญ่ 300 คน

2. เปรียบเทียบจำนวนข้อสอบ ที่คัดเลือกได้ตามเกณฑ์ จากการวิเคราะห์ข้อสอบวิธีเดียวกันในแต่ละขนาดกลุ่มตัวอย่าง คือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน โดยใช้ χ^2 - Test ทดสอบ สรุปว่า ข้อสอบที่วิเคราะห์วิธี 1 หาค่าอำนาจจำแนกและค่าความยากจากกลุ่มตัวอย่างขนาดต่างกัน ได้จำนวนข้อแตกต่างกันอย่างมีนัยสำคัญทางสถิติ วิธี 2 ใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ และวิธี 3 ใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ ก็ได้ผลทำนองเดียวกัน

3. เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบจำนวน 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบ 3 วิธี คือ วิธีหาค่าอำนาจจำแนกและค่าความยาก วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ และวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ ในแต่ละขนาดกลุ่มตัวอย่างโดยใช้ UX_1 ทดสอบ สรุปว่า วิธีการวิเคราะห์ข้อสอบทั้ง 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 30 คน ได้ค่าความเชื่อมั่นไม่แตกต่างกัน กลุ่มตัวอย่างขนาดเล็ก 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน ก็ได้ผลทำนองเดียวกัน

4. เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบจำนวน 50 ข้อ จากการวิเคราะห์ข้อสอบวิธีเดียวกันในแต่ละขนาดกลุ่มตัวอย่าง คือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน โดยใช้ UX_1 ทดสอบ สรุปว่า ข้อสอบที่วิเคราะห์วิธี 1 หาค่าอำนาจจำแนกและค่าความยากจากกลุ่มตัวอย่างขนาดต่างกัน ได้ค่าความเชื่อมั่นไม่แตกต่างกัน วิธี 2 ใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ และวิธี 3 ใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ ก็ได้ผลทำนองเดียวกัน

5. สรุปวิธีวิเคราะห์ข้อสอบ / ขนาดกลุ่มตัวอย่าง ที่ให้ค่าเชื่อมั่นของแบบทดสอบจำนวน 50 ข้อ สูงสุดและต่ำสุด ได้ผลดังแสดงในตาราง 29 – 30

ตาราง 29 สรุปค่าความเชื่อมั่นของแบบทดสอบจำนวน 50 ข้อ ที่วิเคราะห์ข้อสอบ 3 วิธี
ในแต่ละขนาดกลุ่มตัวอย่าง

ขนาดกลุ่มตัวอย่าง (คน)	ค่าความเชื่อมั่น	
	สูงสุด	ต่ำสุด
30	วิธี 2	วิธี 3
60	วิธี 2	วิธี 1
90	วิธี 2	วิธี 1
120	วิธี 3	วิธี 1
150	วิธี 2	วิธี 1
300	วิธี 2	วิธี 1

ตาราง 30 สรุปค่าความเชื่อมั่นของแบบทดสอบที่วิเคราะห์ข้อสอบวิธีเดียวกัน จากกลุ่มตัวอย่าง 6 ขนาด

วิธีวิเคราะห์ข้อสอบ	ค่าความเชื่อมั่น	
	สูงสุด	ต่ำสุด
วิธี 1	60 คน	120 คน
วิธี 2	30 คน 60 คน	120 คน
วิธี 3	60 คน	150 คน

อภิปรายผล

การวิจัยครั้งนี้ มีจุดมุ่งหมายสำคัญเพื่อศึกษาผลการคัดเลือกข้อสอบจากการวิเคราะห์ข้อสอบ 3 วิธี และกลุ่มตัวอย่างขนาดต่างกัน 6 ขนาด ว่าจะได้ข้อสอบจำนวนข้อแตกต่างกันหรือไม่ และข้อสอบที่คัดเลือกเข้าฉบับจะให้ค่าความเชื่อมั่นแตกต่างกันหรือไม่ การวิเคราะห์ข้อสอบวิธีใดและใช้ขนาดกลุ่มตัวอย่างขนาดเท่าไรจึงจะให้ค่าความเชื่อมั่นสูงสุด ผลการวิเคราะห์สามารถนำมาอภิปรายได้ดังนี้

1. การศึกษาจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ จากการวิเคราะห์ข้อสอบ 3 วิธี ในแต่ละขนาดกลุ่มตัวอย่าง 6 ขนาด พบว่า ขนาดเล็ก 30 คน และ 60 คน จำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์แตกต่างกันอย่างไม่มีนัยสำคัญ แต่ในกลุ่มตัวอย่างขนาดเล็ก 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน แตกต่างกันอย่างมีนัยสำคัญ ซึ่งสอดคล้องกับสมมติฐานเป็นส่วนใหญ่ การที่กลุ่มตัวอย่างขนาดเล็ก 30 คน และ 60 คน จำนวนข้อสอบที่คัดเลือกได้ในแต่ละวิธีวิเคราะห์ข้อสอบไม่แตกต่างกันอาจเป็นเพราะว่าจำนวนคนในกลุ่มตัวอย่างน้อยทำให้มีความแปรปรวนน้อยจึงส่งผลให้ได้จำนวนข้อสอบที่คัดเลือกได้ไม่ต่างกัน จากการศึกษาสะท้อนให้เห็นว่า การวิเคราะห์ข้อสอบวิธีหาค่าอำนาจจำแนกและค่าความยากจะมีความไวในการคัดเลือกข้อสอบได้มากกว่าวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อและวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อเพราะข้อสอบที่ใช้วิธีหาค่าอำนาจจำแนกและค่าความยากจะคัดข้อสอบที่ไม่มีคุณภาพออกได้มากกว่าวิธีอื่น

2. การศึกษาจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ ในระหว่างกลุ่มตัวอย่าง 6 ขนาด จากการวิเคราะห์ข้อสอบ 3 วิธี พบว่า จำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์แตกต่างกันอย่างมีนัยสำคัญ ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ จากการศึกษาสะท้อนให้เห็นว่าขนาดของกลุ่มตัวอย่างมีผลต่อจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์คือ ขนาดกลุ่มตัวอย่างต่างกันจำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ก็จะต่างกัน

3. การศึกษาค่าความเชื่อมั่นของแบบทดสอบจำนวน 50 ข้อ ที่ได้จากการวิเคราะห์ข้อสอบต่างกัน 3 วิธีคือ วิธีหาค่าอำนาจจำแนกและค่าความยาก วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ และวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ ในแต่ละขนาดกลุ่มตัวอย่าง 6 ขนาด คือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน ได้ค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ แต่สอดคล้องกับการศึกษาของรัชกร สุวรรณจรัส (2540:100) ได้ศึกษาเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบฉบับเดียวกันที่คำนวณโดยสูตรต่างกัน 2 สูตร พบว่า ค่าความเชื่อมั่นของแบบทดสอบมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ การที่ผลการวิจัยเป็นเช่นนี้เนื่องจากวิธีการวิเคราะห์ข้อสอบแต่ละวิธีได้คัดข้อสอบที่ไม่มีคุณภาพออก ซึ่งส่วนใหญ่เป็นข้อเดียวกันแสดงให้เห็นว่าวิธีวิเคราะห์ข้อสอบทั้ง 3 วิธี คัดข้อสอบที่ไม่มีคุณภาพออกได้เหมือน ๆ กัน ทั้งนี้เมื่อพิจารณาข้อที่คัดออกจาก

การวิเคราะห์ข้อสอบวิธีหาค่าอำนาจจำแนกและค่าความยาก วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ และวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ ในแต่ละขนาดกลุ่มตัวอย่างมีดังนี้

3.1 ในกลุ่มตัวอย่างขนาดเล็ก 30 คน พบว่าข้อสอบที่ถูกคัดออกจากการวิเคราะห์ข้อสอบทั้ง 3 วิธี คือ ข้อ 1, 2, 7, 8, 10, 11, 13, 16, 19, 24, 29, 31, 37, 40, 43, 45, 46, 52, 55, 67, 68, 72, 73, 75, 76, 84, 85, 86, 87, 91, 92 และ 100 ซึ่งข้อ 1, 7, 10, 13, 16, 19, 31, 37, 40, 43, 46, 52, 55, 67 และ 73 เป็นแบบทดสอบประเภททักษะ ข้อ 2, 8, 11, 29 และ 68 เป็นแบบทดสอบประเภทโจทย์ปัญหา ข้อ 24, 45, 72 และ 75 เป็นแบบทดสอบประเภทความคิดรวบยอด ข้อ 76, 84, 85, 86, 87, 91, 92 และ 100 เป็นแบบทดสอบประเภทเรียงอันดับ สะท้อนให้เห็นว่าข้อสอบที่ถูกคัดออกจากการวิเคราะห์ข้อสอบทั้ง 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 30 คน ส่วนมากจะเป็นแบบทดสอบประเภททักษะ ซึ่งอาจเป็นเพราะว่าข้อสอบประเภทนี้ค่อนข้างง่ายเด็กเก่งและเด็กอ่อนสามารถทำได้พอ ๆ กัน จึงเป็นข้อที่ไม่มีคุณภาพซึ่งเป็นผลให้ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบทั้ง 3 วิธีได้ค่าใกล้เคียงกัน

3.2 ในกลุ่มตัวอย่างขนาดเล็ก 60 คน พบว่าข้อสอบที่ถูกคัดออกจากการวิเคราะห์ข้อสอบทั้ง 3 วิธี คือ ข้อ 1, 2, 3, 4, 7, 10, 11, 13, 15, 16, 17, 18, 19, 21, 22, 24, 25, 26, 27, 30, 33, 42, 50, 51, 63, 70, 74, 75, 81, 83, 85, 86, 87, 97, 99 และ 100 ซึ่งข้อ 1, 4, 7, 10, 13, 16, 19, 22, 25 และ 70 เป็นแบบทดสอบประเภททักษะ ข้อ 2, 11, 17, 26, 50 และ 74 เป็นแบบทดสอบประเภทโจทย์ปัญหา ข้อ 3, 15, 18, 21, 24, 27, 30, 33, 42, 51, 63 และ 75 เป็นแบบทดสอบประเภทความคิดรวบยอด ข้อ 81, 83, 85, 86, 87, 97, 99 และ 100 เป็นแบบทดสอบประเภทเรียงอันดับ สะท้อนให้เห็นว่าข้อสอบที่ถูกคัดออกจากการวิเคราะห์ข้อสอบทั้ง 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 60 คน ส่วนมากจะเป็นแบบทดสอบประเภททักษะ เช่นเดียวกับกลุ่มตัวอย่างขนาดเล็ก 30 คน ซึ่งเป็นผลให้ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบทั้ง 3 วิธีได้ค่าใกล้เคียงกัน

3.3 ในกลุ่มตัวอย่างขนาดเล็ก 90 คน พบว่าข้อสอบที่ถูกคัดออกจากการวิเคราะห์ข้อสอบทั้ง 3 วิธี คือ ข้อ 1, 4, 6, 7, 10, 11, 13, 15, 16, 18, 19, 22, 24, 25, 28, 31, 32, 33, 34, 35, 37, 40, 43, 46, 49, 52, 55, 58, 64, 67, 73, 76, 77, 80, 85, 87 และ 100 ซึ่งข้อ 1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, 34, 37, 40, 43, 46, 49, 52, 55, 58, 64, 67 และ 73 เป็นแบบทดสอบประเภททักษะ ข้อ 11, 32 และ 35 เป็นแบบทดสอบประเภทโจทย์ปัญหา ข้อ 6, 15, 18, 24 และ 33 เป็นแบบทดสอบประเภทความคิดรวบยอด ข้อ 76, 77, 80, 85, 87 และ 100 เป็นแบบทดสอบประเภทเรียงอันดับ จะเห็นว่าข้อสอบที่ถูกคัดออกจากการวิเคราะห์ข้อสอบทั้ง 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 90 คน ส่วนมากจะเป็นแบบทดสอบประเภททักษะเช่นเดียวกับกลุ่มตัวอย่างขนาดเล็ก 30 คน ซึ่งเป็นผลให้ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบทั้ง 3 วิธีได้ค่าใกล้เคียงกัน

3.4 ในกลุ่มตัวอย่างขนาดเล็ก 120 คน พบว่าข้อสอบที่ถูกคัดออกจากการวิเคราะห์ข้อสอบทั้ง 3 วิธี คือ ข้อ 1, 3, 4, 7, 8, 10, 11, 12, 15, 16, 17, 21, 22, 24, 25, 27, 29, 30, 37, 41, 43, 46, 48, 50, 63, 72, 75, 85, 87 และ 99 ซึ่งข้อ 1, 4, 7, 10, 16, 22, 25, 37, 43 และ 46 เป็นแบบทดสอบประเภททักษะ ข้อ 8, 11, 17, 29, 41 และ 50 เป็นแบบทดสอบประเภทโจทย์ปัญหา ข้อ 3, 12, 15, 21, 24, 27, 30, 48, 63, 72 และ 75 เป็นแบบทดสอบประเภทความคิดรวบยอด ข้อ 85, 87 และ 99 เป็นแบบทดสอบประเภทเรียงอันดับ จะเห็นว่าข้อสอบที่ถูกคัดออกจากการวิเคราะห์ข้อสอบทั้ง 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 120 คน ส่วนมากจะเป็นแบบทดสอบประเภททักษะเช่นเดียวกับกลุ่มตัวอย่างขนาดเล็ก 30 คน ซึ่งเป็นผลให้ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบทั้ง 3 วิธีได้ค่าใกล้เคียงกัน

3.5 ในกลุ่มตัวอย่างขนาดเล็ก 150 คน พบว่าข้อสอบที่ถูกคัดออกจากการวิเคราะห์ข้อสอบ ทั้ง 3 วิธี คือ ข้อ 1, 2, 3, 4, 7, 10, 11, 13, 15, 16, 18, 19, 21, 24, 25, 31, 37, 43, 48, 51, 61, 63, 65, 72, 79, 87, 94, 95, 97 และ 99 ซึ่งข้อ 1, 4, 7, 10, 13, 16, 19, 25, 31, 37, 43 และ 61 เป็นแบบทดสอบประเภททักษะ ข้อ 2, 11 และ 65 เป็นแบบทดสอบประเภทโจทย์ปัญหา ข้อ 3, 15, 18, 21, 24, 48, 51, 63 และ 72 เป็นแบบทดสอบประเภทความคิดรวบยอด ข้อ 79, 87, 94, 95, 97 และ 99 เป็นแบบทดสอบประเภทเรียงอันดับ จะเห็นว่าข้อสอบที่ถูกคัดออกจากการวิเคราะห์ข้อสอบทั้ง 3 วิธี จากกลุ่มตัวอย่างขนาดเล็ก 150 คน ส่วนมากจะเป็นแบบทดสอบประเภททักษะเช่นเดียวกับกลุ่มตัวอย่างขนาดเล็ก 30 คน ซึ่งเป็นผลให้ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบทั้ง 3 วิธีได้ค่าใกล้เคียงกัน

3.6 ในกลุ่มตัวอย่างขนาดใหญ่ 300 คน พบว่าข้อสอบที่ถูกคัดออกจากการวิเคราะห์ข้อสอบ ทั้ง 3 วิธี คือ ข้อ 1, 2, 4, 5, 6, 7, 10, 11, 13, 15, 16, 18, 19, 20, 22, 23, 24, 25, 26, 28, 31, 32, 33, 34, 35, 37, 40, 43, 46, 49, 52, 55, 58, 61, 64, 70, 73, 76, 87 และ 100 ซึ่งข้อ 1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, 34, 37, 40, 43, 46, 49, 52, 55, 58, 61, 64, 70 และ 73 เป็นแบบทดสอบประเภททักษะ ข้อ 2, 5, 11, 20, 23, 26, 32 และ 35 เป็นแบบทดสอบประเภทโจทย์ปัญหา ข้อ 6, 15, 18, 24 และ 33 เป็นแบบทดสอบประเภทความคิดรวบยอด ข้อ 76, 87 และ 100 เป็นแบบทดสอบประเภทเรียงอันดับ จะเห็นว่าข้อสอบที่ถูกคัดออกจากการวิเคราะห์ข้อสอบทั้ง 3 วิธี จากกลุ่มตัวอย่างขนาดใหญ่ 300 คน ส่วนมากจะเป็นแบบทดสอบประเภททักษะเช่นเดียวกับกลุ่มตัวอย่างขนาดเล็ก 30 คน ซึ่งเป็นผลให้ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบทั้ง 3 วิธีได้ค่าใกล้เคียงกัน

4. ค่าความเชื่อมั่นของแบบทดสอบจำนวน 50 ข้อ ระหว่างกลุ่มตัวอย่าง 6 ขนาดคือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน ที่วิเคราะห์ข้อสอบแต่ละวิธี คือ วิธีหาค่าอำนาจจำแนกและค่าความยาก วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ และวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ พบว่า ค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ แต่สอดคล้องกับการศึกษาของสุทิน บุญญาภิบาล (2537 : 50) ซึ่งพบว่าค่าความเชื่อมั่นของแบบทดสอบ เมื่อกลุ่มผู้สอบมีจำนวนต่างกันมีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งผลการวิจัยเป็นเช่นนี้เนื่องจากกลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นตัวแทนที่ดีของประชากร เป็นไปตามทฤษฎีการสุ่มตัวอย่างคือ กลุ่มตัวอย่างแต่ละกลุ่มประกอบด้วยนักเรียนที่มีความสามารถสูง กลาง ต่ำ ปั่นกันทำให้การกระจายของคะแนนมีมากพอ ๆ กัน ซึ่งการกระจายของคะแนนในแต่ละกลุ่มใกล้เคียงกันทำให้ค่าความเชื่อมั่นของแบบทดสอบมีค่าใกล้เคียงกัน และเป็นดังคำกล่าวของเมเรนส์และลีแมน (Mehrens and Lehmann. 1984 : 278 – 281) ที่ว่ากลุ่มตัวอย่างในแต่ละขนาดเป็นวิวิธพันธ์ มีช่วงการกระจายของคะแนนมาก ทำให้ค่าความเชื่อมั่นสูงกว่ากลุ่มตัวอย่างที่เป็นเอกพันธ์ซึ่งมีช่วงการกระจายของคะแนนน้อยกว่า ทั้งนี้เมื่อพิจารณาจากผลการศึกษา

4.1 การวิเคราะห์ข้อสอบวิธีหาค่าอำนาจจำแนกและค่าความยาก จากกลุ่มตัวอย่าง 6 ขนาดคือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน พบว่าข้อสอบที่ถูกคัดออกในทุก ๆ ขนาดกลุ่มตัวอย่างคือข้อ 1, 2, 4, 7, 10, 11, 13, 15, 16, 19, 22, 24, 28, 31, 40, 41, 43, 46, 49, 52, 55, 58, 70, 76 และ 87 ซึ่งข้อ 1, 4, 7, 10, 13, 16, 19, 22, 28, 31, 40, 43, 46, 49, 52, 55, 58 และ 70 เป็นแบบทดสอบประเภททักษะ ข้อ 2, 11 และ 41 เป็นแบบทดสอบประเภทโจทย์ปัญหา ข้อ 15 และ 24 เป็นแบบทดสอบประเภทความคิดรวบยอด ข้อ 76 และ 87 เป็นแบบทดสอบประเภทเรียงอันดับ จะเห็นว่าข้อที่ถูกคัดออกส่วนใหญ่เป็นแบบทดสอบประเภททักษะ อาจเป็นเพราะว่าแบบทดสอบประเภทนี้ค่อนข้างง่าย จึง

ทำให้ข้อสอบประเภทนี้ไม่มีคุณภาพจึงถูกคัดออก ซึ่งเป็นผลให้ค่าความเชื่อมั่นของแบบทดสอบในแต่ละขนาดกลุ่มตัวอย่างได้ค่าใกล้เคียงกัน

4.2 การวิเคราะห์ข้อสอบวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ จากกลุ่มตัวอย่าง 6 ขนาด คือขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน พบว่า ข้อสอบที่ถูกคัดออกในทุก ๆ ขนาดกลุ่มตัวอย่างคือ ข้อ 1, 7, 10, 11, 13, 16, 21, 24, 31, 85 และ 87 ซึ่งข้อ 1, 7, 10, 13, 16 และ 31 เป็นแบบทดสอบประเภททักษะ ข้อ 11 เป็นแบบทดสอบประเภทโจทย์ปัญหา ข้อ 21 และ 24 เป็นแบบทดสอบประเภทความคิดรวบยอด ข้อ 85 และ 87 เป็นแบบทดสอบประเภทเรียงอันดับ จะเห็นว่าข้อที่ถูกคัดออกส่วนใหญ่เป็นแบบทดสอบประเภททักษะ อาจเป็นเพราะว่าแบบทดสอบประเภทนี้ค่อนข้างง่าย จึงทำให้ข้อสอบประเภทนี้ไม่มีคุณภาพจึงถูกคัดออก ซึ่งเป็นผลให้ค่าความเชื่อมั่นของแบบทดสอบในแต่ละขนาดกลุ่มตัวอย่างได้ค่าใกล้เคียงกัน

4.3 การวิเคราะห์ข้อสอบวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ จากกลุ่มตัวอย่าง 6 ขนาด คือขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน พบว่า ข้อสอบที่ถูกคัดออกในทุก ๆ ขนาดกลุ่มตัวอย่างคือ ข้อ 1, 7, 10, 11, 16, 24, 85 และ 87 ซึ่งข้อ 1, 7, 10 และ 16 เป็นแบบทดสอบประเภททักษะ ข้อ 11 เป็นแบบทดสอบประเภทโจทย์ปัญหา ข้อ 24 เป็นแบบทดสอบประเภทความคิดรวบยอด ข้อ 85 และ 87 เป็นแบบทดสอบประเภทเรียงอันดับ จะเห็นว่าข้อที่ถูกคัดออกส่วนใหญ่เป็นแบบทดสอบประเภททักษะ อาจเป็นเพราะว่าแบบทดสอบประเภทนี้ค่อนข้างง่าย จึงทำให้ข้อสอบประเภทนี้ไม่มีคุณภาพจึงถูกคัดออก ซึ่งเป็นผลให้ค่าความเชื่อมั่นของแบบทดสอบในแต่ละขนาดกลุ่มตัวอย่างได้ค่าใกล้เคียงกัน

จากการวิจัยพบว่า การวิเคราะห์ข้อสอบทั้ง 3 วิธี คือวิธีหาค่าอำนาจจำแนกและค่าความยาก วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ และวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ ไม่ว่าจะใช้กลุ่มตัวอย่างขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน ให้ค่าความเชื่อมั่นใกล้เคียงกัน และวิธีหาค่าอำนาจจำแนกและค่าความยากเป็นวิธีที่สะดวกและรวดเร็วกว่าวิธีอื่น

ข้อเสนอแนะ

ผลจากการวิจัยครั้งนี้ สามารถสรุปเป็นข้อเสนอแนะในการนำไปใช้และในการวิจัยต่อไปนี้

1. ข้อเสนอแนะในการนำไปใช้

1.1 ในการวิเคราะห์ข้อสอบความถนัดด้านคณิตศาสตร์ควรเลือกใช้วิธีหาค่าอำนาจจำแนกและค่าความยาก

1.2 ในการเลือกขนาดกลุ่มตัวอย่างเพื่อใช้ในการวิเคราะห์ข้อสอบความถนัดด้านคณิตศาสตร์สามารถเลือกกลุ่มตัวอย่างขนาดเล็ก 30 คนได้

2. ข้อเสนอแนะในการวิจัย

2.1 ควรมีการศึกษาค่าความเชื่อมั่นของแบบทดสอบที่ได้จากวิธีวิเคราะห์ข้อสอบวิธีอื่น และความยาวของแบบทดสอบต่างกัน เพื่อเป็นแนวทางในการเลือกวิธีวิเคราะห์ข้อสอบและความยาวของแบบทดสอบที่เหมาะสม

2.2 ควรมีการศึกษากับแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนในวิชาต่าง ๆ และขนาดกลุ่มตัวอย่างหลาย ๆ ขนาด

2.3 ควรมีการศึกษาโปรแกรมการวิเคราะห์รายข้อ โดยใช้สถิติที่มีอยู่ในโปรแกรมสำเร็จรูป

เช่น Factor Analysis และ Regression Analysis มาประยุกต์ใช้ในการวิเคราะห์รายชื่อ

2.4 ในการคัดเลือกข้อสอบจากการวิเคราะห์ข้อสอบ วิธีหาค่าอำนาจจำแนกและค่าความยาก
เมื่อกลุ่มตัวอย่างขนาดเล็ก ควรศึกษาโดยใช้เกณฑ์การคัดเลือกข้อสอบด้วยค่าอำนาจจำแนกที่มีนัยสำคัญ
ทางสถิติ แทนค่าอำนาจจำแนก .20 ขึ้นไป

บรรณานุกรม

บรรณานุกรม

- กานดา พุนลาภทวี. การประเมินผลการศึกษา. กรุงเทพฯ : ภาควิชาครุศาสตร์เทคโนโลยี คณะครุศาสตร์
อุตสาหกรรมและวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้า พระนครเหนือ, 2528.
- ✓ ขวาล แพร่ตฤกล. "แบบทดสอบเพื่อจำแนกอาชีพ," พัฒนาวัดผล 10. กรุงเทพฯ : สำนักทดสอบทางการศึกษา
และจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2517.
- ✓ ชูศรี วงศ์รัตน์. สถิติเพื่อการวิจัย. กรุงเทพฯ : คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
ประสานมิตร, 2537.
- ✓ รัชกร สุวรรณจรัส. การประมาณค่าความเชื่อมั่นของแบบทดสอบเอ็มอีคิว วัดความสามารถในการแก้ปัญหา
วิชาบัญชีกิจการบริหาร ที่มีผู้ตรวจให้คะแนน 1 คน. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ
ประสานมิตร, 2540.
- ✓ รัชชัย งามสันติวงศ์. หลักการและวิธีใช้คอมพิวเตอร์ในงานสถิติเพื่อการวิจัย SPSS FOR WINDOWS.
พิมพ์ครั้งที่ 5. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2540.
- บุญชม ศรีสะอาด. การวัดและประเมินผลการศึกษา. กรุงเทพฯ : คณะศึกษาศาสตร์ มหาวิทยาลัย
ศรีนครินทรวิโรฒ มหาสารคาม, 2520.
- บุญเชิด ภิญโญนันทพงษ์. การวัดและประเมินผลการศึกษา. กรุงเทพฯ : ภาควิชาพื้นฐานทางการศึกษา
คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2519.
- _____ การวัดและประเมินผลการศึกษา : ทฤษฎีและการประยุกต์. พิมพ์ครั้งที่ 2. กรุงเทพฯ :
ภาควิชาพื้นฐานการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร,
2521.
- _____ การทดสอบแบบอิงเกณฑ์ : แนวคิดและวิธีการ. กรุงเทพฯ : ภาควิชาพื้นฐานทางการศึกษา
คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2527.
- _____ การประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งส่วนย่อยตามแบบจำลองคะแนนจริงสัมพัทธ์.
ปริญญาณิพนธ์ กศ.ด. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2537. อัดสำเนา.
- ✓ บุญเรือน จรัสวิมล. การเพิ่มข้อสอบที่มีผลต่อค่าความเชื่อมั่นของแบบทดสอบ. ปริญญาณิพนธ์ กศ.ม.
กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2521. อัดสำเนา.
- ✓ ผดุงชาติ สุวรรณวงศ์. "Reliability and Validity ของข้อสอบที่เขียนเพิ่มให้ยาวขึ้น," ในพัฒนาวัดผล 9. หน้า
42- 43 . กรุงเทพฯ : กรุงเทพมหานครพิมพ์, 2516.
- ✓ พิมพ์อุษา เจริญยศ. การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อโดยใช้
เกณฑ์ค่าสถิติกับใช้เกณฑ์ดุลพินิจของครู. ปริญญาณิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัย
ศรีนครินทรวิโรฒ ประสานมิตร , 2539. อัดสำเนา.
- ✓ ยศอารีย์ รวยธนพานิช. การเปรียบเทียบความเชื่อมั่นของแบบทดสอบปรนัยชนิดเลือกตอบที่มีช่วงความยาก
และความยาวของแบบทดสอบต่างกันในวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 3. ปริญญาณิพนธ์
กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2534. อัดสำเนา.
- ✓ ล้วน สายยศ. "ความเชื่อมั่นอีกรูปหนึ่ง" พัฒนาวัดผล 12. กรุงเทพฯ : สำนักทดสอบทางการศึกษาและ
จิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2519.

ล้วน สายยศ และอังคณา สายยศ. หลักการวิจัยทางการศึกษา. พิมพ์ครั้งที่ 2. กรุงเทพฯ : ศึกษาพร, 2531.
 _____. หลักการสร้างแบบทดสอบความถนัดทางการเรียน. พิมพ์ครั้งที่ 1. กรุงเทพฯ : วัฒนาพานิช,
 2527.

วารุณี ปิครวัชชัย. การศึกษาเปรียบเทียบค่าสถิติของแบบทดสอบ ชนิดเลือกตอบที่มีจำนวนตัวเลือก
 ไม่เท่ากัน. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย , 2513. อัดสำเนา.

วิญญา วิศาลาภรณ์. วัดผล 401 การประเมินผลทางการศึกษา. กรุงเทพฯ : ภาควิชาพื้นฐานการศึกษา
 มหาวิทยาลัยศรีนครินทรวิโรฒ พิษณุโลก, 2521.

สมบุญ ภู่นวล. การประเมินผลและการสร้างแบบทดสอบ. พระนครศรีอยุธยา : วิทยาลัยครู
 พระนครศรีอยุธยา , 2525.

สมบูรณ์ หิตพงศ์ และสำเริง บุญเรืองรัตน์. การวัดความถนัด. กรุงเทพฯ : ไทยวัฒนาพานิช, 2518.

สวัสดิ์ ประทุมราช. “แนวการปรับปรุงการวัดผลและประเมินผลการศึกษา,” วารสารแนะแนว 8(34) : 32-34;
 สิงหาคม – กันยายน 2517.

สุนีย์ เหมะประสิทธิ์. สถิติประยุกต์เพื่อการวิจัย. กรุงเทพฯ : คณะครุศาสตร์อุตสาหกรรม สถาบันเทคโนโลยี
 พระจอมเกล้าธนบุรี, 2536.

✓ สุนิ บุญญาภิบาล. การเปรียบเทียบคุณภาพของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์
 โดยทฤษฎีคลาสสิกอล เมื่อกลุ่มผู้สอบมีจำนวนต่างกัน. วิทยานิพนธ์ กศ.ม. กรุงเทพฯ :
 มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2537. อัดสำเนา.

สุภาพ วาดเขียน และอรพินธ์ โภชนาดา. การประเมินการเรียนการสอน. กรุงเทพฯ : แผนกวิชาวิจัย
 การศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2518.

✓ สำเริง บุญเรืองรัตน์. ทฤษฎีการวัดและการประเมินผลการศึกษา. พิมพ์ครั้งที่ 2. กรุงเทพฯ : สยามศึกษา,
 2529.

✓ อรุณช ธิติรักษ์พาณิชย์. การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ วัดบุคลิกภาพของเยาวชนไทย
 เมื่อลดความยาวให้สั้นลง. วิทยานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ
 ประสานมิตร, 2538. อัดสำเนา.

อุทุมพร ทองอุไทย. การประเมินทางการศึกษา. กรุงเทพฯ : คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย,
 2520.

อนันต์ ศรีโสภณ. การพัฒนาการทดสอบมัธยมศึกษา. กองส่งเสริมและวัดผลการศึกษา กรมวิสามัญศึกษา
 กระทรวงศึกษาธิการ นครหลวงกรุงเทพธนบุรี, 2515.

_____. การวัดและประเมินผลการศึกษา. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร,
 2524.

✓ _____. การวัดผลการศึกษา. พิมพ์ครั้งที่ 3. กรุงเทพฯ : ไทยวัฒนาพานิช, 2525.

Allen, Mary J. and Wendy M.Yen. Introduction to Measurement Theory. California : Brooks/COLE,
 1979.

✓ Anastasi, Anne. Psychological Testing. 5 th ed. New York : Macmillan, Co., Inc., 1968.

Budescu, Davic V. and Barch Nevo. “ Optional Number of Options : An Investigation of the
 Assumption of Proportionalify,” Journal of Educational Measurement. 22 : 183-196; Fall,
 1985.

- ✓ ✓ Cronbach, Lee J. Essentials of Psychological Testing. 3 rd ed. New York : Harper and Row Publishers, 1970.
- Edward, Allen Louis. Techniques of Attitude Scale Construction. Englewood Clifts, N.J. : Prentice – Hall, 1957.
- ✓ ✓ Ferguson, George A. Statistical Analysis in Psychology Education. 3 rd ed. New York : Mc Graw – Hill, Book Company , 1966.
- Good, Carter Victor. Dictionary of Education. 3 6 rd ed. New York : Mc Graw – Hill, 1973.
- ✓ Gronlund, Norman E. Measurement and Evaluation in Teaching . 3 rd ed. New York : Macmilan, 1976.
- ✓ Lord, Frederic M. and Melvin R. Novick. Statistical Theories of Mental Test Scores . California : Harcount Brace Jovanovich, Inc., 1967.
- ✓ ✓ Mehren , William A. and Irvin J. Lehmann . Measurement and Evaluation in Education and Psychology. New York : Rinehart and Winston , 1984.
- ✓ ✓ Nunnally Jum C. Education Measurement and Evaluation. New York : Mc. Graw Hill Book Company, 1964.
- Snedecor and Cochran. Statistical Methods. Iowa State University Press, 1967.

ภาคผนวก

ภาคผนวก ก

- ข้อสอบที่คัดเลือกไว้และตัดออกที่วิเคราะห์หาค่าความยากและค่าอำนาจจำแนกจากกลุ่มตัวอย่าง 6 ขนาด
- ข้อสอบที่คัดเลือกไว้และตัดออกที่วิเคราะห์หาวัดโปรแกรมสำเร็จโดยลดจำนวนข้อจากกลุ่มตัวอย่าง 6 ขนาด
- ข้อสอบที่คัดเลือกไว้และตัดออกที่วิเคราะห์หาวัดดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อจากกลุ่มตัวอย่าง 6 ขนาด

ตาราง 31 แสดงข้อสอบ 100 ข้อ ที่วิเคราะห์ข้อสอบวิธีหาค่าอำนาจจำแนกและค่าความยาก
จากกลุ่มตัวอย่าง 6 ขนาด

ข้อที่	ขนาดกลุ่มตัวอย่าง (คน)					
	30	60	90	120	150	300
1	0	0	0	0	0	0
2	0	1	1	1	1	0
3	1*	0	1*	1	0	1*
4	0	0	0	0	0	0
5	1*	1*	0	1*	1*	1
6	1*	1*	1	1*	1*	1
7	0	0	0	0	0	0
8	1	1*	1*	1	1*	1*
9	1*	1*	1*	1*	1*	1*
10	0	0	0	0	0	0
11	1	0	0	1	0	0
12	1*	1*	1*	1	1*	1*
13	1	0	0	0	0	0
14	1*	1*	1*	1*	1*	1*
15	0	0	0	0	0	0
16	0	1	0	0	0	0
17	1*	1*	1*	1	1*	1*
18	1*	1	1	1*	1	1
19	0	0	0	0	0	0
20	1*	1*	1*	1*	1*	1
21	1	1	1*	1	1	1*
22	0	0	0	0	0	0
23	1*	1*	1	1*	1*	1
24	0	0	0	0	0	1
25	1*	0	0	0	0	0
26	1	1	1	1*	0	0
27	1*	1	1*	0	1	1*
28	0	0	0	0	0	0
29	0	1*	1*	1	1*	1*
30	1*	1	1*	1	1*	1*

ตาราง 31 (ต่อ)

ข้อที่	ขนาดกลุ่มตัวอย่าง (คน)					
	30	60	90	120	150	300
31	0	0	0	0	0	0
32	1*	1	0	1	0	0
33	0	1	0	1*	0	0
34	0	1*	0	0	0	0
35	1*	1*	0	1*	0	0
36	1*	1*	1*	1*	1*	1*
37	0	1*	0	0	0	0
38	1*	1*	1*	1*	1*	1*
39	1*	1*	1*	1*	1*	1*
40	1	1	0	0	0	0
41	0	1	1	0	1	1
42	1*	1	1*	1*	1*	1*
43	1	1	0	0	0	0
44	1*	1*	1*	1*	1*	1*
45	1	1*	1*	1*	1*	1*
46	0	0	0	0	0	0
47	1*	1*	1*	1*	1*	1*
48	1*	1*	1*	1	1	1*
49	1	0	0	1	0	0
50	1*	1	1	1	0	1*
51	1*	1	1*	1*	1	1*
52	1	1	0	0	0	0
53	1	1*	1*	1*	1*	1
54	1*	1*	1*	1*	1*	1*
55	1	0	0	0	0	0
56	1*	1*	1*	1*	1*	1*
57	1*	1*	1*	1*	1*	1*
58	1	0	0	0	0	0
59	1*	1*	1*	1*	1*	1*
60	1*	1*	1*	1*	1*	1*

ตาราง 31 (ต่อ)

ข้อที่	ขนาดกลุ่มตัวอย่าง (คน)					
	30	60	90	120	150	300
61	1	1*	0	1	0	0
62	1*	1*	1*	1*	1*	1*
63	1*	1	1*	1	1	1*
64	1*	1*	0	1	1*	1
65	1	1*	1*	1*	1	1*
66	1	1*	1*	1*	1*	1*
67	1	1*	0	1	0	0
68	0	1*	1*	1*	1*	1*
69	1*	1	1*	1*	1*	1*
70	0	1	0	0	0	0
71	1*	1*	1*	1*	1*	1*
72	1	1*	1	1	1	1
73	1	1*	0	0	0	0
74	1*	1	1*	1	1*	1*
75	1	1	1	1	1*	1*
76	0	0	0	0	0	0
77	1*	0	0	1	1*	0
78	1*	1	1	1*	1*	1
79	1*	1*	1*	1*	1	1*
80	1*	1*	1	1*	1*	1
81	0	1	1*	1*	1*	1*
82	1*	1*	1*	1*	1*	1*
83	1*	1	1*	1*	1*	1
84	1	1*	1*	1*	1*	1*
85	1	1	1	1	1*	1*
86	1	1	1*	1*	1*	1*
87	1	0	1	1	0	1
88	1*	1*	1*	1*	1*	1*
89	1*	1*	1*	1*	1*	1*
90	1*	1*	1*	1*	1*	1*

ตาราง 31 (ต่อ)

ข้อที่	ขนาดกลุ่มตัวอย่าง (คน)					
	30	60	90	120	150	300
91	1	1*	1	1*	1*	1
92	1	1*	1*	1*	1*	1*
93	1*	1*	1*	1*	1*	1*
94	1*	1*	1*	1*	1	1*
95	1*	1*	1*	1*	1	1*
96	1*	1*	1*	1*	1*	1*
97	1	1	1*	1*	1	1*
98	1*	1*	1*	1*	1*	1*
99	1*	1	1	1	1	1
100	1	1	1	1*	1*	1

0 คือ ข้อที่ถูกตัดออก , 1 คือ ข้อที่คัดเลือกได้ตามเกณฑ์ , * คือ ข้อที่คัดเลือกเพื่อศึกษาค่าความเชื่อมั่น

ตาราง 32 แสดงข้อสอบ 100 ข้อ ที่วิเคราะห์ข้อสอบวิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ
จากกลุ่มตัวอย่าง 6 ขนาด

ข้อที่	ขนาดกลุ่มตัวอย่าง (คน)					
	30	60	90	120	150	300
1	0	0	0	0	1	0
2	1	1	1	1*	1	0
3	1*	0	1	1	0	1
4	1*	0	0	1	0	0
5	1*	1*	1*	1*	1*	1
6	1*	1*	1	1*	1*	1
7	1	1	1	1	1*	0
8	0	1	1	1	1	1*
9	1*	1*	1*	1	1*	1*
10	0	0	0	1	0	0
11	0	1	0	1	0	0
12	1*	1*	1*	1	1*	1*
13	0	0	0	1	0	0
14	1*	1*	1*	1*	1*	1*
15	1*	1	0	1	1	0
16	0	0	0	1	0	0
17	1*	1	1	0	1	1
18	1*	0	0	1	1	0
19	0	0	0	1*	1	1
20	1*	1	1*	1*	1*	1
21	1	0	1	0	0	1
22	1*	1	0	1	1	0
23	1*	1	1*	1	1	1
24	0	0	0	0	0	0
25	1*	0	1	1	1	0
26	1*	1	1*	1	1*	1
27	1*	1	1*	0	1	1*
28	1*	1*	0	1	1*	0
29	0	1*	1	1	1	1
30	1*	0	1*	1	1*	1*

ตาราง 32 (ต่อ)

ข้อที่	ขนาดกลุ่มตัวอย่าง (คน)					
	30	60	90	120	150	300
31	0	1	0	1	1	0
32	1*	1*	1	1*	1*	1
33	1*	1	1	1*	1*	0
34	1*	1*	0	1*	1*	0
35	1*	1*	1	1*	1*	1
36	1*	1*	1*	1*	1*	1*
37	0	1*	0	1	1	0
38	1*	1*	1*	1*	1*	1*
39	1*	1*	1*	1*	1*	1*
40	1	1*	1	1*	1*	0
41	1	1*	1*	1	1*	1*
42	1*	0	1	1*	1*	1*
43	1	1*	0	1	0	0
44	1*	1*	1*	1*	1*	1*
45	0	1	1	1*	1*	1
46	1	1*	1	1	1*	0
47	1*	1*	1*	1*	1*	1*
48	1*	1*	1*	1	1	1*
49	1*	1*	1	1*	1*	0
50	1*	1	1*	1	1*	1*
51	1*	1	1*	1	0	1*
52	0	1*	0	1	1*	1
53	1*	1*	1*	1*	1*	1*
54	1*	1*	1*	1*	1*	1*
55	0	1*	0	1*	1*	1
56	0	1	1*	1	1*	1*
57	0	1*	1*	1*	1*	1
58	1*	1*	1	1*	1*	1
59	1*	1*	1*	1*	1*	1*
60	1*	1*	1*	1*	1*	1*

ตาราง 32 (ต่อ)

ข้อที่	ขนาดกลุ่มตัวอย่าง (คน)					
	30	60	90	120	150	300
61	1*	1*	1*	1*	1	1
62	1*	1	1*	1	1*	1*
63	1	0	1*	1	0	1
64	1*	1*	0	1*	1*	1*
65	1*	1*	1*	1	0	1
66	1*	1	1*	1	1*	1*
67	0	1*	1	1*	1*	1
68	0	1*	1*	1*	1*	1*
69	1*	0	1*	1	1	1*
70	1*	1	1*	1*	0	1
71	1*	1*	1*	1	1	1*
72	0	1*	1*	1	1	1*
73	0	1*	1	1*	1*	1
74	1	0	1*	1*	1*	1*
75	0	0	1*	1	1	1
76	1	1*	1	1*	1	1
77	1	1*	1	1*	1*	1*
78	0	1*	1*	1*	1*	1*
79	1	1	1*	1*	0	1*
80	1	1*	1	1*	1*	1*
81	0	1	1	1*	1*	1*
82	1*	1*	1	1	1*	1*
83	1	0	1*	1*	1*	1*
84	0	1*	1*	1	1	1*
85	0	0	0	1	1	1
86	0	0	1*	1*	1	1*
87	1	0	1	1	0	0
88	1	1	1*	1*	1	1*
89	1*	1*	1*	1*	1*	1*
90	1	1*	1*	1*	1*	1*

ตาราง 32 (ต่อ)

ข้อที่	ขนาดกลุ่มตัวอย่าง (คน)					
	30	60	90	120	150	300
91	0	1*	1*	1	1	1*
92	0	1	0	1	1	1*
93	1	1*	1*	1*	1	1*
94	1*	1	1*	1	1	1*
95	1*	1	1*	1*	1	1*
96	1	1*	1*	1*	1	1*
97	0	0	1	1*	1	1*
98	1	1*	1	1*	1	1*
99	1*	1	1*	1	1	1*
100	1	0	0	1*	1	1

0 คือ ข้อที่ถูกตัดออก , 1 คือ ข้อที่คัดเลือกได้ตามเกณฑ์ , * คือ ข้อที่คัดเลือกเพื่อศึกษาค่าความเชื่อมั่น

ตาราง 33 แสดงข้อสอบ 100 ข้อ ที่วิเคราะห์ข้อสอบวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ
จากกลุ่มตัวอย่าง 6 ขนาด

ข้อที่	ขนาดกลุ่มตัวอย่าง (คน)					
	30	60	90	120	150	300
1	1	0	0	0	1	0
2	1	1	1*	1	1	1
3	1*	0	1	1	0	1
4	1*	0	0	1	0	0
5	1*	1*	1*	1*	1	1
6	1*	1*	1	1*	1*	1
7	1	1	1	1	0	0
8	0	1	1	1	1	1*
9	1*	1*	1*	1	1*	1*
10	0	1	0	1	0	0
11	0	1	0	1	0	0
12	1*	1	1*	1	1*	1*
13	0	0	0	1*	0	0
14	1*	1*	1*	1*	1*	1*
15	1*	0	0	1	1	0
16	0	0	0	1	0	0
17	1*	1	0	0	1	1
18	1*	0	0	1*	1*	0
19	0	0	0	1*	1	1
20	1*	1	1*	1*	1*	1
21	1*	0	1	0	0	1
22	1*	1	0	1	1	0
23	1*	1	1*	1	1*	1
24	0	0	0	0	0	0
25	1*	0	1	1	1	0
26	1*	1	1*	1*	1*	1
27	1*	1	1*	0	1*	1*
28	1*	1*	1	1*	1*	0
29	0	1*	0	1	1*	1*
30	1*	0	1*	1	1*	1*

ตาราง 33 (ต่อ)

ข้อที่	ขนาดกลุ่มตัวอย่าง (คน)					
	30	60	90	120	150	300
31	0	1*	1	1*	1	0
32	1*	1*	1	1*	1*	1
33	1*	1	1	1*	1*	0
34	1*	1*	0	1*	1	0
35	1*	1*	1	1*	1*	1
36	1*	1*	1*	1	1*	1*
37	0	1*	0	0	1	0
38	1*	1*	1*	1*	1*	1*
39	1*	1*	1*	1	1*	1*
40	1	1*	1	1*	1*	0
41	1*	1*	1*	1	1*	1*
42	1*	0	1*	1*	1*	1*
43	1	1*	0	0	0	0
44	1*	1*	1*	1*	1*	1*
45	0	1	1	1	1*	1
46	1	1*	1	1	1*	0
47	1*	1*	1*	1*	1*	1*
48	1*	1*	1*	1	1	1*
49	1*	1*	1	1*	1*	0
50	1*	1	1*	1	1*	1*
51	1*	1	1	1	0	1*
52	0	1*	0	1*	1*	1
53	1*	1*	1*	1*	1*	1*
54	1*	1*	1*	1*	1*	1*
55	0	1*	0	1*	1*	1
56	1	1	1*	1	1*	1*
57	0	1*	1*	1*	1	1
58	1*	1*	1	1*	1*	1
59	1*	1*	1*	1*	1*	1*
60	1*	1*	1*	1*	1*	1*

ตาราง 33 (ต่อ)

ข้อที่	ขนาดกลุ่มตัวอย่าง (คน)					
	30	60	90	120	150	300
61	1*	1*	1*	1*	1	1
62	1*	1	1*	1	1*	1*
63	1	0	1*	0	0	1
64	1	1*	0	1*	1*	1
65	1*	1*	1*	1	0	1
66	1*	1	1*	0	1*	1*
67	0	1*	1	1*	1*	1
68	0	1*	1*	1*	1*	1*
69	1*	0	1*	1	1	1*
70	1	0	1*	1*	1	1
71	1*	1*	1*	1	1	1*
72	0	1*	1	1	1	1*
73	0	1*	1	1*	1*	1
74	1	0	1*	1*	1	1*
75	0	0	1*	1	1	1
76	1	1*	1	1*	1	1
77	1	1*	1	1*	1*	1*
78	0	1*	1*	1*	1*	1*
79	1	1	1*	1	0	1*
80	1	1*	1	1*	1*	1*
81	0	1	1*	1	1*	1*
82	1*	1*	1*	1	1*	1*
83	1	0	1*	1*	1*	1*
84	0	1*	1*	1	1	1*
85	0	0	0	1	1	1
86	0	0	1*	1*	1	1*
87	1	0	1	1	0	0
88	1	1	1	1*	1*	1*
89	1*	1*	1*	1*	1*	1*
90	1	1*	1*	1*	1*	1*

ตาราง 33 (ต่อ)

ข้อที่	ขนาดกลุ่มตัวอย่าง (คน)					
	30	60	90	120	150	300
91	0	1*	1*	1	1	1*
92	0	1	0	1	1	1*
93	1	1*	1*	1*	1	1*
94	1*	1	1*	1	1	1*
95	1*	1	1*	1*	1	1*
96	1	1*	1*	1*	1	1*
97	0	0	1	1*	1	1*
98	1	1*	1	1*	1	1*
99	1*	1	1	1	1	1*
100	1	0	0	1*	1	1

0 คือ ข้อที่ถูกตัดออก , 1 คือ ข้อที่คัดเลือกได้ตามเกณฑ์ , * คือ ข้อที่คัดเลือกเพื่อศึกษาค่าความเชื่อมั่น

ภาคผนวก ข

แบบทดสอบที่ใช้ในการวิจัย

แบบทดสอบวัดความถนัดด้านคณิตศาสตร์

คำชี้แจง

1. แบบทดสอบฉบับนี้เป็นแบบปรนัยชนิดเลือกตอบ มีจำนวน 100 ข้อ ใช้เวลาทำ 60 นาที ฉะนั้นนักเรียนควรรีบตอบโดยเร็วและให้ครบทุกข้อจึงจะได้คะแนนดี
2. แบบทดสอบฉบับนี้เป็นแบบทดสอบวัดความถนัดด้านคณิตศาสตร์ การตอบของนักเรียนไม่มีผลต่อการเรียนวิชาคณิตศาสตร์แต่จะบอกถึงความสามารถของนักเรียน โปรดทำให้เต็มความสามารถ

ตัวอย่างที่ 1

(0) ฉันมีปากกา 7 ด้าม พี่ให้อีกครึ่งโหล ฉันมีปากกาทั้งหมดกี่ด้าม

- ก 10
- ข 11
- ค 12
- ง 13
- จ 14

คำตอบ ข้อที่ถูก คือ ข ง

ตัวอย่างที่ 2

(00) เครื่องหมาย ? คือเลขจำนวนใด

2 4 6 8 ?

- ก 9
- ข 10
- ค 11
- ง 12
- จ 13

คำตอบ ข้อที่ถูก คือ ข

3. ให้กาเครื่องหมาย X ลงในกระดาษคำตอบเพียงข้อละหนึ่งคำตอบเท่านั้น ถ้าตอบเกินกว่าหนึ่งคำตอบจะถือว่าข้อนั้นผิด

4. อย่าขีดเขียนหรือทำเครื่องหมายใด ๆ ลงในแบบทดสอบฉบับนี้

1. $1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 = ?$
 - ก. 33
 - ข. 34
 - ค. 35
 - ง. 36
 - จ. 37
2. ไข่ 1 โหล แลกถ้วยได้ 3 ใบ ถ้วย 2 ใบ แลกไข่ได้กี่ใบ
 - ก. 4
 - ข. 6
 - ค. 8
 - ง. 10
 - จ. 12
3. "ซื้อมะม่วงมา 5 กิโลกรัม กับ 5 ผล เป็นเงิน 100 บาท เมื่อขายไปได้กำไร 20 บาท ขายมะม่วงไปผลละเท่าไร" จากโจทย์หาคำตอบได้หรือไม่เพราะเหตุใด
 - ก. หาไม่ได้เพราะไม่ทราบราคาทุน
 - ข. หาไม่ได้เพราะไม่ทราบราคาขาย
 - ค. หาไม่ได้เพราะไม่ทราบจำนวนมะม่วงที่ซื้อ
 - ง. หาไม่ได้เพราะไม่ทราบจำนวนมะม่วงที่ขาย
 - จ. หาได้
4. $21 + 12 + 31 + 13 = ?$
 - ก. 73
 - ข. 74
 - ค. 75
 - ง. 76
 - จ. 77
5. พัดลมปิดป้ายไว้ราคา 350 บาท แต่ลดให้ผู้ซื้อเงินสดร้อยละ 6 เมื่อค้าซื้อด้วยเงินสดต้องจ่ายเงินกี่บาท
 - ก. 315
 - ข. 319
 - ค. 325
 - ง. 329
 - จ. 330
6. สี่เหลี่ยมผืนผ้ารูปหนึ่งพื้นที่จะเป็นอย่างไร ถัดด้านกว้างเพิ่มขึ้นเป็นสองเท่าของรูปเดิม และด้านยาวลดลงเป็นครึ่งหนึ่งของรูปเดิม
 - ก. พื้นที่เท่าเดิม
 - ข. พื้นที่ลดลง 2 เท่า
 - ค. พื้นที่ลดลง 1 เท่า
 - ง. พื้นที่เพิ่มขึ้น 1 เท่า
 - จ. พื้นที่เพิ่มขึ้น 2 เท่า
7. $125 - 85 - 23 = ?$
 - ก. 17
 - ข. 40
 - ค. 63
 - ง. 187
 - จ. 233
8. ในการวิ่งรอบสนามวงกลมของกิ้ง ตุ่ม และ นาด ปรากฏว่าใช้เวลาวิ่งรอบสนามหนึ่งรอบใช้เวลา 2, 3 และ 4 นาที ตามลำดับ ถ้าเริ่มออกวิ่งพร้อมกันจะใช้เวลากี่นาที จึงจะมาถึงจุดเริ่มต้นพร้อมกันอีก
 - ก. 12
 - ข. 24
 - ค. 30
 - ง. 60
 - จ. 120

9. ข้อใดมีค่ามากที่สุด

- ก. a^1
- ข. $a + 1$
- ค. $a - 1$
- ง. $a \div 1$
- จ. $a \times 1$

10. $(33 + 27 + 18) - 52 = ?$

- ก. 24
- ข. 26
- ค. 62
- ง. 96
- จ. 130

11. มีน้ำมันอยู่ 24 ลิตร ใช้ไป 8.08 ลิตร และเติมอีก 4.75 ลิตร จะมีน้ำมันกี่ลิตร

- ก. 11.17
- ข. 15.92
- ค. 20.67
- ง. 27.33
- จ. 36.83

12. อีก 5 ปีข้างหน้าออยมีอายุ X ปี เมื่อ 4 ปีที่แล้วออยมีอายุเท่าไร

- ก. 9 ปี
- ข. $X + 4$ ปี
- ค. $X - 4$ ปี
- ง. $X - 5$ ปี
- จ. $X - 9$ ปี

13. $6241 + 6400 + 6561 = ?$

- ก. 19,020
- ข. 19,022
- ค. 19,200
- ง. 19,202
- จ. 19,220

14. สี่เหลี่ยมจัตุรัสรูปหนึ่งมีพื้นที่ 16 ตารางเมตร จะมีเส้นรอบรูปยาวกี่เมตร

- ก. 8
- ข. 12
- ค. 16
- ง. 32
- จ. 64

15. ข้อใดมีค่ามากที่สุด

- ก. 0.1
- ข. 0.01
- ค. 0.001
- ง. 0.0001
- จ. 0.00001

16. $(300 - 152 - 42) + 2 = ?$

- ก. 102
- ข. 104
- ค. 106
- ง. 108
- จ. 110

17. เลข 3 จำนวนเรียงกัน ผลบวกของเลขจำนวนน้อยกับจำนวนมากเท่ากับ 88 จงหาผลรวมของเลข 3 จำนวนนั้น

- ก. 126
- ข. 129
- ค. 132
- ง. 147
- จ. 162

18. สี่เหลี่ยมใดที่เส้นทแยงมุมทั้ง 2 จะไม่แบ่งครึ่งพื้นที่สี่เหลี่ยมนั้น

- ก. สี่เหลี่ยมผืนผ้า
- ข. สี่เหลี่ยมด้านขนาน
- ค. สี่เหลี่ยมขนมเปียกปูน
- ง. สี่เหลี่ยมคางหมู
- จ. สี่เหลี่ยมจัตุรัส

19. $(10+11+12+9+8+10+9+11) \div 8 = ?$

- ก. 7.5
- ข. 10
- ค. 10.5
- ง. 11
- จ. 11.5

20. อ้อยมีเงินเป็น 2 เท่าของอิม อิมมีเงินเป็น $\frac{3}{4}$ ของอิม ถ้าอิมมีเงิน 300 บาท อ้อยมีเงินกี่บาท

- ก. 115
- ข. 225
- ค. 300
- ง. 375
- จ. 450

21. $(ก + 12)$ จะมีค่าเป็นลบเมื่อใด

- ก. มีค่ามากกว่า 12
- ข. มีค่าน้อยกว่า 12
- ค. มีค่ามากกว่าลบ 12
- ง. มีค่าน้อยกว่าลบ 12
- จ. มีค่าเท่ากับ 12

22. $3 \times 5 \times 15 = ?$

- ก. 225
- ข. 222
- ค. 252
- ง. 255
- จ. 522

23. ร้านค้าคิดราคาหมีไฟฟ้า 800 บาท ถ้าซื้อเงินสดจ่ายในราคา 720 บาท ทางร้านลดให้กี่เปอร์เซ็นต์

- ก. 9
- ข. 10
- ค. 11
- ง. 12
- จ. 13

24. $\frac{5}{12}(-60 + 24) = \frac{5}{12}(24) + \frac{5}{12}(-60)$

ใช้คุณสมบัติข้อใดต่อไปนี้บ้าง

- | |
|--|
| 1. กฎการจัดหมู่
2. กฎการสลับที่
3. กฎการกระจาย |
|--|

- ก. ข้อ 1
- ข. ข้อ 2 และข้อ 3
- ค. ข้อ 1 และข้อ 3
- ง. ข้อ 1 และข้อ 2
- จ. ข้อ 1, 2 และข้อ 3

25. $9 \times 9 \times 9 = ?$

- ก. 18
- ข. 27
- ค. 81
- ง. 243
- จ. 729

26. ถ้านาฬทรายาว 32 เซนติเมตร มาตัดเป็นสี่เหลี่ยมจัตุรัส จะมีพื้นที่กี่ตารางเซนติเมตร

- ก. 36
- ข. 49
- ค. 64
- ง. 81
- จ. 100

27. ถ้า $a > b$, $b > c$, $c = d$ และ $d > e$ ข้อใดมีค่ามากที่สุด

- ก. $\frac{a}{b}$
- ข. $\frac{a}{d}$
- ค. $\frac{a}{c}$
- ง. $\frac{a}{a}$
- จ. $\frac{a}{e}$

28. $101 \times 101 = ?$
 ก. 1,021
 ข. 10,021
 ค. 10,200
 ง. 10,201
 จ. 12,001
29. มีเชือกอยู่ 3 เส้นยาวเส้นละ 35 , 49 และ 56 เซนติเมตร ต้องการตัดให้ยาวที่สุดและเท่าๆ กัน จะได้เชือกยาวเส้นละกี่เซนติเมตร โดยไม่เหลือเศษเลย
 ก. 5
 ข. 7
 ค. 9
 ง. 10
 จ. 12
30. ถ้าฐานและส่วนสูงของสามเหลี่ยมรูปหนึ่ง เพิ่มขึ้นจากเดิมเป็นสองและสามเท่าตามลำดับพื้นที่ใหม่จะเป็นกี่เท่าของพื้นที่เดิม
 ก. 2 เท่า
 ข. 3 เท่า
 ค. 4 เท่า
 ง. 5 เท่า
 จ. 6 เท่า
31. $(36 \times 2) + (61 \times 3) = ?$
 ก. 255
 ข. 256
 ค. 257
 ง. 258
 จ. 259
32. เชือกเส้นหนึ่งยาว 22 เมตร ตัดขายไป 5 เมตร ตัดส่วนที่เหลือมัดสิ่งของอีก 4.5 เมตร จะยังเหลือเชือกยาวกี่เมตร
 ก. 10.5
 ข. 12.5
 ค. 14.5
 ง. 16.5
 จ. 21.5
33. ดาวบอกคนขายว่าต้องการริบบิ้นยาว 250 เซนติเมตร คนขายวัดให้ 2 เมตร ครึ่ง ดาวจะอย่างไร
 ก. ให้เงินเพิ่ม
 ข. ต่อว่าคนขาย
 ค. จ่ายเงินน้อยลง
 ง. จ่ายเงินตามราคา
 จ. ไม่ยอมรับริบบิ้น
34. $(321 \times 3) - 564 = ?$
 ก. 395
 ข. 396
 ค. 397
 ง. 398
 จ. 399
35. แอนทำขนมคุกกี้ใช้แป้ง 3 ส่วน ต่อน้ำตาล 1 ส่วน ถ้าเขาใช้น้ำตาล 4 ถ้วย จะต้องใช้แป้งกี่ถ้วย
 ก. 6
 ข. 8
 ค. 9
 ง. 10
 จ. 12
36. ข้อใดได้กำไรคิดเป็นเปอร์เซ็นต์สูงสุด
 ก. ลงทุน 10 บาท กำไร 10 บาท
 ข. ลงทุน 30 บาท กำไร 15 บาท
 ค. ลงทุน 100 บาท กำไร 80 บาท
 ง. ลงทุน 200 บาท กำไร 100 บาท
 จ. ลงทุน 250 บาท กำไร 200 บาท
37. $(1024 + 2014) - 565 = ?$
 ก. 2,443
 ข. 2,453
 ค. 2,463
 ง. 2,473
 จ. 2,483

38. โรงเรียนแห่งหนึ่งมีนักเรียน 1,899 คน
เป็นนักเรียนหญิง $\frac{4}{9}$ ดังนั้น จะมี
นักเรียนชายกี่คน
ก. 633
ข. 844
ค. 1,055
ง. 1,266
จ. 1,477
39. ถ้า $a + b$ มีค่าเป็นบวก ดังนั้น a และ b
ควรมีได้หลายค่ายกเว้นข้อใด
ก. a และ b เป็นจำนวนเต็มลบ
ข. a เป็นจำนวนเต็มบวก b เป็นจำนวน
เต็มลบ
ค. a เป็นจำนวนเต็มลบ b เป็นจำนวน
เต็มบวก
ง. a เป็นจำนวนเต็มศูนย์ b เป็นจำนวน
เต็มบวก
จ. b เป็นจำนวนเต็มศูนย์ a เป็นจำนวน
เต็มบวก
40. $(12 \times 24) - (10 \times 11) = ?$
ก. 146
ข. 158
ค. 166
ง. 169
จ. 178
41. สระน้ำสี่เหลี่ยมผืนผ้ากว้าง 20 เมตร ยาว
25 เมตร ทำถนนโดยรอบถนนกว้าง 2 เมตร
คิดเป็นพื้นที่ถนนกี่ตารางเมตร
ก. 80
ข. 94
ค. 100
ง. 164
จ. 196
42. ถ้า A เป็นจำนวนคี่แล้ว จำนวนที่ถัดไปคือ
อะไร
ก. $A + 1$
ข. $A + 2$
ค. $A + 3$
ง. $A + 4$
จ. $A + 5$
43. $(112 \times 2) + 125 = ?$
ก. 99
ข. 224
ค. 239
ง. 349
จ. 439
44. เลี้ยงสุนัขและแมวรวมกัน 460 ตัว ถ้าเลี้ยง
สุนัข 60 เปอร์เซนต์จะเลี้ยงแมวกี่ตัว
ก. 160
ข. 184
ค. 205
ง. 276
จ. 295
45. 5.2×5.2 ถ้าเลื่อนจุดทศนิยมของตัวตั้งไป
ข้างขวา 1 ตำแหน่ง และเลื่อนจุดทศนิยม
ของตัวคูณไปข้างซ้าย 1 ตำแหน่ง ผลคูณจะ
เป็นอย่างไร
ก. มีค่าเพิ่มขึ้น 10 เท่า
ข. มีค่าลดลง 10 เท่า
ค. มีค่าเท่าเดิม
ง. มีค่าลดลง 2 เท่า
จ. มีค่าเป็นจำนวนเต็ม
46. $(2 \times 22) - 9 = ?$
ก. 34
ข. 35
ค. 36
ง. 37
จ. 38

47. ปักเสาต้นหนึ่งยาว 15 ฟุต ลงไปในน้ำ
ปรากฏว่าปักอยู่ในดิน $\frac{1}{3}$ ของความยาวทั้งหมด โผล่พ้นน้ำ 4 ฟุต น้ำลึกกี่ฟุต
ก. 7
ข. 6
ค. 5
ง. 4
จ. 3
48. เมื่อ 4 ปีที่แล้ว แดงอายุ ก ปี อีก จ ปี
แดงจะมีอายุเท่าไร
ก. $4ก + จ$ ปี
ข. $4ก จ$ ปี
ค. $4 + ก จ$ ปี
ง. $ก - 4 + จ$ ปี
จ. $4 + ก + จ$ ปี
49. $(156 \div 13) + 17 = ?$
ก. 19
ข. 29
ค. 92
ง. 91
จ. 160
50. นำลูกเต๋าที่มีขนาดกว้างด้านละ 1 เมตร
มาทำลูกเต๋าลึก ๆ ขนาดกว้างด้านละ 1
เซนติเมตรได้กี่ลูก
ก. 100
ข. 1,000
ค. 10,000
ง. 100,000
จ. 1,000,000
51. การแบ่งครึ่งมุมให้เท่า ๆ กันในข้อใดไม่
สามารถทำได้โดยวิธีแบ่งครึ่งมุมด้วย
วงเวียนและสันตรง
ก. การแบ่ง 2 มุม
ข. การแบ่ง 4 มุม
ค. การแบ่ง 6 มุม
ง. การแบ่ง 8 มุม
จ. การแบ่ง 16 มุม
52. $(85 \div 5) + 46 = ?$
ก. 61
ข. 62
ค. 63
ง. 64
จ. 65
53. แบ่งเงิน 150 บาทให้สมศรี สมปองและ
สมพล ให้สมศรีได้ $\frac{1}{6}$ ส่วน สมปองได้
 $\frac{2}{6}$ ส่วน และสมพลได้ $\frac{3}{6}$ ส่วน อยาก
ทราบว่าสมศรีได้รับเงินกี่บาท
ก. 25
ข. 50
ค. 75
ง. 100
จ. 125
54. "สามเท่าของเลขจำนวนหนึ่งมากกว่าสี่
อยู่แปด" เขียนเป็นประโยคสัญลักษณ์
ได้ดังข้อใด
ก. $3(X - 4) = 8$
ข. $3X - 4 = 8$
ค. $3X = 8 - 4$
ง. $3X - 8 = -4$
จ. $3X + 4 = 8$
55. $(242 \times 12) \div 4 = ?$
ก. 267
ข. 276
ค. 627
ง. 726
จ. 762

56. อาภาพรอ่านหนังสือเล่มหนึ่ง วันแรกอ่าน

ได้ $\frac{1}{3}$ ของจำนวนหน้าทั้งหมด วันที่สอง
อ่านได้ 10 หน้า ปรากฏว่าอ่าน 2 วันรวม
กันได้ครึ่งเล่มพอดี อยากทราบว่าหนังสือ
เล่มนี้มีกี่หน้า

- ก. 60
ข. 100
ค. 120
ง. 140
จ. 180

57. ถ้า C เป็นเลขโดดที่ไม่ใช่ศูนย์
C ในข้อใดมีค่ามากที่สุด

- ก. 64,32C
ข. 54,C23
ค. 43,5C2
ง. 3C,423
จ. 21,C42

58. $(3 \times 3 \times 3) \div 3 = ?$

- ก. 3
ข. 7
ค. 9
ง. 27
จ. 81

59. ปัจจุบันพี่ชายมนต์สิทธิ์มีอายุ 28 ปี เมื่อ 3
ปีที่แล้ว มนต์สิทธิ์มีอายุเป็น $\frac{1}{5}$ ของอายุ
พี่ชาย ปัจจุบันมนต์สิทธิ์อายุกี่ปี

- ก. 5
ข. 6
ค. 7
ง. 8
จ. 9

60. ค่าของ n ที่ทำให้ $n + 2 = n + 12$ คือข้อใด

- ก. 0
ข. เลขลบทั้งหมด
ค. เลขบวกทั้งหมด
ง. ไม่มีเลขใดที่ทำได้
จ. ค่าใดก็ได้

61. $(1028 \times 6) \div 12 = ?$

- ก. 510
ข. 512
ค. 514
ง. 516
จ. 518

62. นกกระจอกตัวหนึ่งหนักประมาณ 5 กรัม
นกกระจอกเพศตัวหนึ่งหนักประมาณ 25
กิโลกรัม อยากทราบว่าต้องใช้นกกระจอก
กี่ตัวจึงจะหนักเท่ากับนกกระจอกเพศ
หนึ่งตัว

- ก. 500
ข. 5,50
ค. 5,000
ง. 5,500
จ. 6,000

63. $\frac{a-b}{a+b} - \frac{b}{a+b} = ?$

- ก. $\frac{a}{a+b}$
ข. $\frac{a-b}{a}$
ค. $\frac{a}{2(a+b)}$
ง. $\frac{a-2b}{a+b}$
จ. $\frac{a-2b}{2(a+b)}$

64. $(50 \times 27) \div 25 = ?$

ก. 50

ข. 51

ค. 52

ง. 53

จ. 54

65. ดอกกุหลาบ 100 ดอก ราคา 75 บาท
ถ้ามีเงิน 126 บาท จะซื้อดอกกุหลาบ
ได้กี่ดอก

ก. 95

ข. 150

ค. 168

ง. 175

จ. 180

66. ถ้า a และ b เป็นความยาวและความกว้าง
ของรูปสี่เหลี่ยมผืนผ้าแล้ว ab คืออะไร

ก. ด้านสองด้าน

ข. เส้นตั้งฉาก

ค. เส้นรอบรูปสี่เหลี่ยมผืนผ้า

ง. พื้นที่ของรูปสี่เหลี่ยมผืนผ้า

จ. เส้นทแยงมุมของสี่เหลี่ยมผืนผ้า

67. $(235 \times 7) \div 5 = ?$

ก. 289

ข. 323

ค. 327

ง. 329

จ. 397

68. สามเหลี่ยมมุมฉากรูปหนึ่งมีด้านประกอบ
มุมฉากยาว 24 และ 32 เซนติเมตร และ
ด้านตรงข้ามมุมฉากยาว 40 เซนติเมตร
สามเหลี่ยมรูปนี้มีพื้นที่ตารางเซนติเมตร

ก. 384

ข. 480

ค. 640

ง. 768

จ. 960

69. ถ้า $A = B + C$ แล้วข้อใดเป็นจริง

1. $C = A - B$

2. $A - C = B$

3. $C + B = A$

ก. ข้อ 1 ถูก

ข. ข้อ 3 ถูก

ค. ข้อ 1 และข้อ 2 ถูก

ง. ข้อ 2 และข้อ 3 ถูก

จ. ข้อ 1, 2 และ 3 ถูก

70. $(7 \times 8) + (7 \times 30) = ?$

ก. 265

ข. 266

ค. 267

ง. 268

จ. 269

71. สี่เหลี่ยมคางหมูรูปหนึ่งมีพื้นที่ 128 ตาราง
เซนติเมตร มีส่วนสูง 8 เซนติเมตร มีด้านคู่
ขนานด้านหนึ่งยาว 15 เซนติเมตร ความ
ยาวของด้านคู่ขนานอีกด้านหนึ่งยาวกี่
เซนติเมตร

ก. 15

ข. 17

ค. 32

ง. 34

จ. 47

72. จุดคู่อันดับใดอยู่ห่างจาก $(a, 4)$ ขึ้นไปแนว
ตั้งเป็นระยะทาง 3 หน่วย

ก. $(a + 3, 4)$

ข. $(a - 3, 4)$

ค. $(a, 3)$

ง. $(a, 4)$

จ. $(a, 7)$

73. $12 \times (60 + 4) = ?$

- ก. 765
- ข. 766
- ค. 767
- ง. 768
- จ. 769

74. สี่เหลี่ยมผืนผ้ารูปหนึ่งมีด้านยาวยาวเป็น 2 เท่าของด้านกว้าง ถ้าวัดรอบรูปสี่เหลี่ยมผืนผ้ายาว 60 เมตร สี่เหลี่ยมผืนผ้านี้มีพื้นที่กี่ตารางเมตร

- ก. 100
- ข. 150
- ค. 200
- ง. 260
- จ. 300

75. เลขจำนวนหนึ่งมีสองหลัก เลขหลักหน่วยคือ a หลักสิบคือ b เลขจำนวนนั้นเป็นเท่าไร

- ก. $10a + b$
- ข. $10b + a$
- ค. $a + b$
- ง. ab
- จ. $10ab$

คำชี้แจง : ข้อ 76 - 100

เครื่องหมาย $?$ คือเลขจำนวนใด

76. $-2 \quad -4 \quad -6 \quad -8 \quad ?$

- ก. -10
- ข. -11
- ค. -12
- ง. -13
- จ. -14

77. $2.2 \quad 3.0 \quad ? \quad 4.6 \quad 5.4$

- ก. 3.3
- ข. 3.4
- ค. 3.5
- ง. 3.6
- จ. 3.8

78. $105 \quad 79 \quad 53 \quad 27 \quad ?$

- ก. 0
- ข. 1
- ค. 3
- ง. 5
- จ. 7

79. $1 \quad 0 \quad -2 \quad -5 \quad -9 \quad ?$

- ก. -13
- ข. -14
- ค. -15
- ง. -16
- จ. -17

80. $2 \quad 3 \quad 5 \quad 8 \quad 12 \quad ?$

- ก. 13
- ข. 14
- ค. 15
- ง. 16
- จ. 17

81. $1 \quad 4 \quad 9 \quad 16 \quad 25 \quad ?$

- ก. 32
- ข. 34
- ค. 36
- ง. 38
- จ. 40

82. $6 \quad 10 \quad 18 \quad 34 \quad ?$

- ก. 46
- ข. 56
- ค. 66
- ง. 76
- จ. 86

83. 3 7 13 21 ?

ก. 30

ข. 31

ค. 32

ง. 33

จ. 34

84. 640 160 40 10 ?

ก. 5

ข. 4

ค. 3

ง. 2.5

จ. 1.5

85. 64 ? 16 4

ก. 25

ข. 36

ค. 49

ง. 64

จ. 81

86. 60 56 66 62 ?

ก. 68

ข. 70

ค. 72

ง. 74

จ. 46

87. 3 9 4 16 5 ?

ก. 18

ข. 20

ค. 22

ง. 23

จ. 25

88. 1 22 4 33 9 44 ?

ก. 8

ข. 10

ค. 12

ง. 14

จ. 16

89. 5 7 4 6 3 ?

ก. 1

ข. 2

ค. 3

ง. 4

จ. 5

90. 1 8 5 12 ? 16 13

ก. 5

ข. 6

ค. 7

ง. 8

จ. 9

91. 30 13 33 13 36 13 ?

ก. 38

ข. 39

ค. 40

ง. 41

จ. 42

92. 8 12 10 16 12 ?

ก. 18

ข. 20

ค. 22

ง. 24

จ. 26

93. 5 6 1 2 7 8 3 ?

ก. 4

ข. 5

ค. 6

ง. 7

จ. 8

94. 21 7 3 ? 5 6 24 3 8

ก. 11

ข. 15

ค. 30

ง. 33

จ. 35

95. 25 7 20 9 15 11 ?

ก. 10

ข. 11

ค. 12

ง. 13

จ. 14

96. 28 33 31 36 34 ?

ก. 35

ข. 36

ค. 37

ง. 38

จ. 39

97. 4 7 11 6 9 13 8 ?

ก. 9

ข. 10

ค. 11

ง. 12

จ. 13

98. 1009 3007 5005 7003 ?

ก. 9009

ข. 9007

ค. 9005

ง. 9003

จ. 9001

99. 54327 43275 32754 27543 ?

ก. 25437

ข. 35427

ค. 47532

ง. 52734

จ. 75432

100. 3 6 11 18 27 ?

ก. 37

ข. 38

ค. 39

ง. 40

จ. 41

ภาคผนวก ค

รายชื่อผู้เชี่ยวชาญที่ตรวจสอบเครื่องมือ

รายชื่อผู้เชี่ยวชาญที่ตรวจสอบเครื่องมือ

รองศาสตราจารย์นิภา ศรีไพโรจน์

อาจารย์ชวลิต รวยอาจิน

อาจารย์ยุพา มานะจิตต์

ผู้ช่วยศาสตราจารย์ประพนธ์ จำยเจริญ

อาจารย์นพดล กองศิลป์

ภาควิชาการวัดผลและวิจัยการศึกษา

คณะศึกษาศาสตร์ มศว. ประสานมิตร

ภาควิชาการวัดผลและวิจัยการศึกษา

คณะศึกษาศาสตร์ มศว. ประสานมิตร

ข้าราชการบำนาญ

โรงเรียนสาธิตมหาวิทยาลัยศรีนครินทรวิโรฒ

ประสานมิตร (ฝ่ายประถม)

โรงเรียนสาธิตมหาวิทยาลัยศรีนครินทรวิโรฒ

ประสานมิตร (ฝ่ายประถม)

ประวัติย่อผู้วิจัย

ชื่อ	นางสาวภาวนา วงศ์เพ็ญทักษ์
เกิด	10 ธันวาคม 2510
สถานที่เกิด	อำเภอโคกสำโรง จังหวัดลพบุรี
สถานที่อยู่ปัจจุบัน	161 / 83 หมู่ 9 ถนนวิภาวดีรังสิต แขวงสีกัน เขตดอนเมือง กรุงเทพมหานคร
สถานที่ทำงานปัจจุบัน	โรงเรียนมัธยมศึกษาสันพิทยา แขวงมัธยมศึกษา เขตราษฎร์บุรี กรุงเทพมหานคร
ตำแหน่ง	อาจารย์ 2 ระดับ 5
ประวัติการศึกษา	
พ.ศ. 2529	ชั้นมัธยมศึกษาปีที่ 6 โรงเรียนพิบูลวิทยาลัย จังหวัดลพบุรี
พ.ศ. 2533	คบ. (วิชาเอกคอมพิวเตอร์ศึกษา) วิทยาลัยครูเทพสตรี จังหวัดลพบุรี
พ.ศ. 2537	บธ.บ (วิทยาการจัดการทั่วไป) มหาวิทยาลัยสุโขทัยธรรมมาธิราช
พ.ศ. 2541	กศ.ม. (การวัดผลการศึกษา) มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

จำนวนข้อและความเชื่อมั่นของแบบทดสอบที่ได้จากการวิเคราะห์ข้อสอบด้วยวิธีต่างกัน

บทคัดย่อ
ของ
ภาวณา วงศ์เพ็ญทักษ์

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
ตุลาคม 2541

การศึกษาครั้งนี้ มีจุดมุ่งหมายเพื่อศึกษาผลการคัดเลือกข้อสอบตามเกณฑ์และค่าความเชื่อมั่นของแบบทดสอบที่วิเคราะห์ข้อสอบ 3 วิธีคือ วิธีหาค่าอำนาจจำแนกและค่าความยาก วิธีใช้โปรแกรมสำเร็จโดยลดจำนวนข้อ และวิธีใช้ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อ จากกลุ่มตัวอย่าง 6 ขนาด คือ ขนาดเล็ก 30 คน 60 คน 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน จะได้จำนวนข้อแตกต่างกันหรือไม่ และค่าความเชื่อมั่นแตกต่างกันหรือไม่ และการวิเคราะห์ข้อสอบวิธีใด ใช้กลุ่มตัวอย่างขนาดเท่าไรจึงจะให้ค่าความเชื่อมั่นสูงสุด

กลุ่มตัวอย่างเป็นนักเรียนระดับมัธยมศึกษาปีที่ 2 ของโรงเรียนสังกัดกรมสามัญศึกษา ในกรุงเทพมหานคร จำนวน 750 คน ที่ได้จากการสุ่มแบบ 2 ขั้นตอน เครื่องมือที่ใช้ในการศึกษาคือ แบบทดสอบความถนัดด้านคณิตศาสตร์ สร้างตามแนวของเธอร์สโตน จำนวน 100 ข้อ แบ่งเป็น 4 ประเภท คือ ทักษะ ความคิดรวบยอด โจทย์ปัญหา และเรียงอันดับ

ผลการศึกษาพบว่า 1) จำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ จากการวิเคราะห์ข้อสอบ 3 วิธี ในกลุ่มตัวอย่างขนาดเล็ก 30 คน และ 60 คน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนกลุ่มตัวอย่างขนาดเล็ก 90 คน 120 คน 150 คน และขนาดใหญ่ 300 คน แตกต่างกันอย่างมีนัยสำคัญทางสถิติ 2) จำนวนข้อสอบที่คัดเลือกได้ตามเกณฑ์ ระหว่างกลุ่มตัวอย่าง 6 ขนาด ที่วิเคราะห์ข้อสอบ 3 วิธี แตกต่างกันอย่างมีนัยสำคัญทางสถิติ 3) ค่าความเชื่อมั่นของแบบทดสอบที่วิเคราะห์ข้อสอบ 3 วิธี ในแต่ละกลุ่มตัวอย่าง 6 ขนาด แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ 4) ค่าความเชื่อมั่นของแบบทดสอบ ระหว่างกลุ่มตัวอย่าง 6 ขนาด ที่วิเคราะห์ข้อสอบ 3 วิธี แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

จากการศึกษาสรุปได้ว่า ในการคัดเลือกข้อสอบจะเลือกใช้วิธีวิเคราะห์ข้อสอบใดก็ได้ แต่วิธีหาค่าอำนาจจำแนกและค่าความยากเป็นวิธีที่สะดวกรวดเร็วและสิ้นเปลืองเวลาและแรงงานน้อยที่สุด ส่วนกลุ่มตัวอย่างที่ใช้วิเคราะห์สำหรับแบบทดสอบความถนัดด้านคณิตศาสตร์สามารถใช้ขนาดเล็ก 30 คน ได้

**NUMBER OF ITEMS AND RELIABILITY OF A TEST
FROM DIFFERENT ITEM ANALYSIS METHODS**

**AN ABSTRACT
BY
PAWANA WONGPENTUX**

**Presented in partial fulfillment of the requirements for the
Master of Education degree in Educational Measurement
At Srinakharinwirot University
October 1998**

The purpose of the research was to study the selection of the number of items by criterion and test reliability from three methods of item analysis which were item discrimination and difficulty, computer program to reduce the number of items and reliability index : whether the number of the items and test reliability would be different and which would be the most suitable method of item analysis and sample sizes in order to get the highest reliability. The six different sizes of the samples were 30, 60, 90, 120, 150 and the large sizes was 300.

The sample group consisted of 750 Mathayom Suksa II students in the academic year 2540 in Bangkok by using the 2 stage random sampling. The research tool was the aptitude mathematics test, based on Thurstone conception, which divided into four types: skill, concept, problem solving and number-series.

The findings were as follows :

- 1 The number items selected by criterion from three methods of item analysis in the small size samples of 30 and 60 students were not significantly different but they were significantly different in the samples of 90, 120, 150 and 300 students.

- 2 The number items from selected items by criterion among the 6 group sizes from three methods of item analysis were significantly different.

- 3 The reliability of the test from the three methods of item analysis in each size of the samples was significantly different

- 4 The reliability of the test and the six different sample sizes from three methods of item analysis was significantly different.

The findings were concluded that any one of the three methods could be used for the selection of the number of items, but the item discrimination and difficulty method would save time and energy . The sample group which is tested by the aptitude mathematics test could be the small size of 30 students.