บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบวัดคุณลักษณะและแบบสอบ เมื่อใช้วิธีการตรวจให้คะแนนแบบทวิวิภาคและแบบพหุวิภาค โดยวิเคราะห์ด้วย GRM และ GPCM ซึ่งมีขั้นตอนการดำเนินการวิจัยดังนี้

1. ประชากรและกลุ่มตัวอย่าง

1.1 ประชากร

ประชากรในการศึกษาครั้งนี้แบ่งเป็น 2 กลุ่ม คือ

1.1.1 นักเรียนชั้นประถมศึกษาปีที่ 6 ของทุกเขตการศึกษาที่เรียนในปีการศึกษา 2537 และเข้ารับการทดสอบโดยการใช้แบบวัดคุณลักษณะของสำนักทดสอบ กรมวิชาการ รวมทั้งสิ้น จำนวน 102,117 คน

1.1.2 นักเรียนชั้นประถมศึกษาปีที่ 6 ในสังกัดสำนักงานการประถมศึกษาอำเภอพระนครศรีอยุธยาที่กำลังเรียนในภาคเรียนที่ 2 ปีการศึกษา 2538 จำนวน 1,567 คน

1.2 กลุ่มตัวอย่าง

ในขั้นตอนการกำหนดขนาดกลุ่มตัวอย่าง ผู้วิจัยได้กำหนดขนาดในเบื้องต้นตามสูตรของ ยามาเน (Yamane,1970: 580-581) คือ

$$n = N / (1+N e^2)$$

เมื่อ n คือ ขนาดกลุ่มตัวอย่าง

N คือ ขนาดของประชากร

e คือ ความคลาดเคลื่อนที่ยอมรับได้ (กำหนดให้คลาดเคลื่อนได้ร้อยละ 5)

ได้ขนาดกลุ่มตัวอย่างในเบื้องต้นดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

1.2.1 กลุ่มนักเรียนที่เข้ารับการทดสอบจากสำนักทดสอบกรมวิชาการ ได้ขนาดกลุ่มตัวอย่างประมาณ 399 คน

1.2.2 กลุ่มนักเรียนที่สังกัดสำนักงานการประถมศึกษาอำเภอพระนครศรีอยุธยา ได้ขนาดกลุ่มตัวอย่างประมาณ 318 คน

จากการศึกษาของแองเคนแมน และสโตน (Ankenmann and Stone,1992) เกี่ยวกับการประมาณค่าพารามิเตอร์ของผู้สอบ ด้วยวิธีมาร์จินอล แมกซิมัม ไลลิฮูด (marginal maximum likelihood :MML) โดยใช้ข้อสอบจำนวน 5, 10 และ 20 ข้อ กับกลุ่มตัวอย่างขนาด 250, 500 และ 1,000 คน พบว่าการประมาณค่าพารามิเตอร์ของผู้สอบมีความแม่นยำ (accurate) ตั้งแต่กลุ่มตัวอย่างขนาด 500 คน ขึ้นไป นอกจากนี้ ซวินเดอร์แมน และวอเลนเบอร์ก (Zwinderman and Wollenberg, 1990) ได้ทดสอบความแกร่ง (robustness) ของการประมาณค่าพารามิเตอร์แบบมาร์จินอลแมกซิมัม ไลลิฮูด (marginal maximum likelihood) พบว่าถ้าการแจกแจงของกลุ่มประชากรเบ้เล็กน้อยกลุ่มตัวอย่างที่เพียงพอสามารถใช้ได้ตั้งแต่ 500 คนขึ้นไป หากการแจกแจงมีลักษณะเบ้มากควรใช้กลุ่มตัวอย่างประมาณ 1,000 คน ขึ้นไป จากผลการศึกษาดังกล่าวผู้วิจัยจึงได้กำหนดขนาดกลุ่มตัวอย่างดังนี้

ก. กลุ่มตัวอย่างที่เป็นนักเรียนที่เข้ารับการทดสอบโดยสำนักทดสอบกรมวิชาการ จำนวน 6,300 คน เป็นข้อมูลที่สำนักงานทดสอบทางการศึกษา กรมวิชาการได้สุ่มมาเพื่อใช้วิเคราะห์ทั้งหมด การที่ผู้วิจัยกำหนดขนาดกลุ่มตัวอย่างขนาดใหญ่เพราะต้องการให้ผลการวิเคราะห์มีความแกร่ง และทำได้สะดวกเพราะข้อมูลที่ใช้วิเคราะห์มีผู้เก็บรวบรวมไว้แล้ว

ข. กลุ่มตัวอย่างที่เป็นนักเรียนสังกัดสำนักงานการประถมศึกษาอำเภอพระนครศรีอยุธยา กลุ่มตัวอย่างขนาด 500 คน แต่เนื่องจากมีนักเรียนจำนวนหนึ่งขาดเรียนในวันที่ทำการสอบ กลุ่มตัวอย่างจริงมีจำนวน 470 คน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

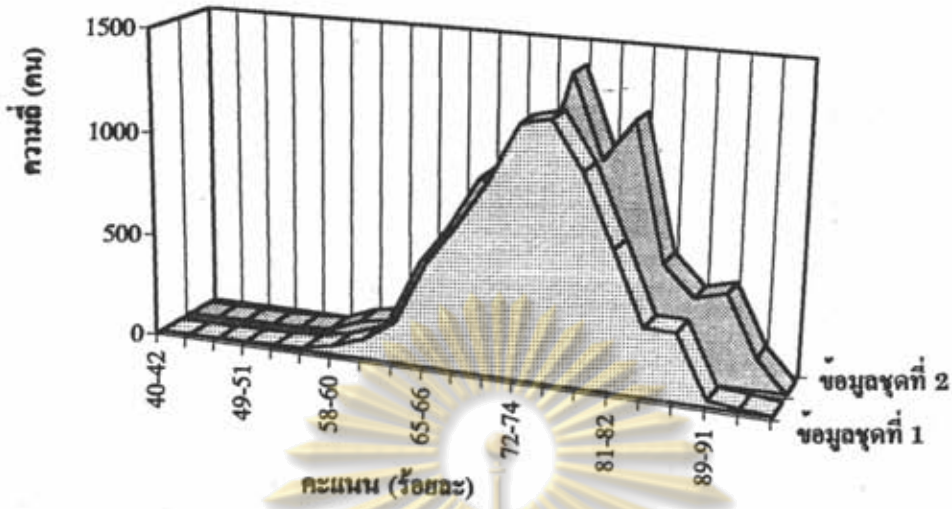
ในการวิจัยครั้งนี้ผู้วิจัยได้ศึกษาผลการใช้แบบวัดคุณลักษณะและแบบสอบคณิตศาสตร์ จากข้อมูล 2 ชุด คือ ข้อมูลชุดที่ 1 ได้จากการใช้แบบวัดและแบบสอบและนักเรียนกลุ่มตัวอย่างทั้งหมด ส่วนข้อมูลชุดที่ 2 ได้จากการใช้แบบวัดและแบบสอบที่คัดเฉพาะข้อกระทงที่เหมาะสมกับโมเดลและนักเรียนที่เหมาะสมกับโมเดล ดังจะได้นำเสนอการแจกแจงคะแนนนักเรียนดังตารางที่ 3 และภาพที่ 3 - 4

ตารางที่ 3 ค่าสถิติพื้นฐานของคะแนนสอบแบบวัดคุณลักษณะและแบบสอบคณิตศาสตร์ ชุดที่ 1 และชุดที่ 2

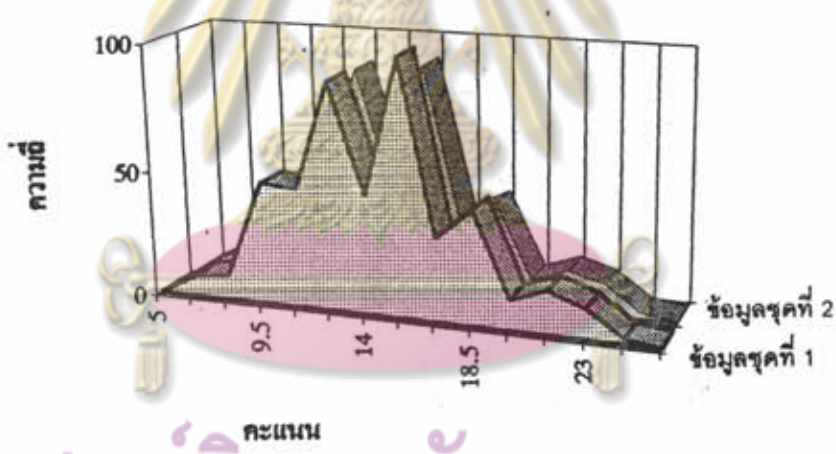
ชุดเครื่องมือ	ผู้สอบ (N)	คะแนนเต็ม	ค่าสถิติพื้นฐานของคะแนนสอบ					
			ค่าสูงสุด (MAX)	ค่าต่ำสุด (MIN)	ค่าเฉลี่ย (MEAN)	ส่วนเบี่ยงเบน (S.D.)	ความเบ้ (SKEWNESS)	ความโด่ง (KURTOSIS)
1.แบบวัดคุณลักษณะ								
1.1 ชุดที่ 1	6,300	120	115	51	89.26	7.72	- 0.14	0.22
1.2 ชุดที่ 2	6,056	88	85	42	67.06	6.99	0.01	- 0.39
2.แบบสอบคณิตศาสตร์								
2.1 ชุดที่ 1	470	24	24	7	14.39	3.65	0.37	- 0.22
2.2 ชุดที่ 2	444	24	24	7	14.36	3.66	0.40	- 0.19

จากตารางที่ 3 สามารถแสดงลักษณะการแจกแจงของคะแนนนักเรียนในการทำแบบวัดคุณลักษณะและสอบคณิตศาสตร์ทั้งชุดที่ 1 และชุดที่ 2 ได้ดังภาพที่ 3 และ 4

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



ภาพที่ 3 การแจกแจงของคะแนนจากข้อมูลการใช้แบบวัดคุณลักษณะ ชุดที่ 1 และชุดที่ 2



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาพที่ 4 การแจกแจงของคะแนนจากข้อมูลการใช้แบบสอบคณิตศาสตร์ ชุดที่ 1 และชุดที่ 2

จากกตารางที่ 3 และภาพที่ 3 และ 4 จะเห็นว่าคะแนนนักเรียนจากข้อมูลการใช้แบบวัดคุณลักษณะและแบบสอบคณิตศาสตร์ ชุดที่ 1 และ 2 มีการแจกแจงที่คล้ายคลึงกัน โดยมีข้อสังเกตว่าข้อมูลจากแบบวัดคุณลักษณะชุดที่ 2 มีลักษณะการแจกแจงที่เข้าใกล้โค้งปกติมากกว่าข้อมูลในชุดที่ 1

สำหรับข้อมูลคะแนนนักเรียนที่ทำแบบสอบคณิตศาสตร์ชุดที่ 1 และ 2 มีการแจกแจงที่คล้ายกันมากคือมีคะแนนเฉลี่ยที่เท่ากัน (14 คะแนน) การแจกแจงเป็นแบบเบ้ขวาเล็กน้อย มีความโด่งต่ำกว่าโค้งปกติ และมีความแปรปรวนของคะแนนที่ใกล้เคียงกันมาก

2 ตัวแปรที่ศึกษา

2.1 ตัวแปรต้น

2.1.1 วิธีการตรวจให้คะแนนแบ่งเป็น 2 วิธีคือ วิธีการตรวจให้คะแนนแบบทวิวิภาคและการตรวจให้คะแนนแบบพหุวิภาค เนื่องจากเป็นวิธีการตรวจให้คะแนนแบบวัดและแบบสอบที่ใช้กันอยู่ในปัจจุบัน

2.1.2 โมเดลที่ใช้วิเคราะห์เปรียบเทียบ ในการศึกษาครั้งนี้มี 3 โมเดล คือ โมเดลการวิเคราะห์ GRM , GPCM และโมเดลโลจิสติก โดย 2 โมเดลแรก เป็นโมเดลตามแนวทฤษฎี Polytomous Item Response ที่มีมโนทัศน์และพัฒนาการที่แตกต่างกัน มีการนำไปใช้มากในปัจจุบัน และมีความยืดหยุ่นในการนำไปใช้ (Muraki , 1993 ; Donoghue, 1994) และโมเดลที่สามเป็นโมเดลการวิเคราะห์ตามแนวทฤษฎี Dichotomous Item Response ที่นิยมใช้ในปัจจุบันเช่นกัน

2.2 ตัวแปรตาม

2.2.1 ค่าฟังก์ชันสารสนเทศ

2.2.2 อัตราส่วนสารสนเทศเฉลี่ย (RAI)

2.2.3 ความสอดคล้องของค่าฟังก์ชันสารสนเทศของแบบวัดระหว่างการวิเคราะห์ตาม GRM และ GPCM

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3. แบบวัดและแบบสอบ

เครื่องมือที่ใช้สำหรับเก็บรวบรวมข้อมูลในการศึกษาครั้งนี้ มี 2 ชุด คือ

3.1 แบบวัดคุณลักษณะทั่วไป

เป็นแบบวัดที่พัฒนาขึ้นโดยสำนักทดสอบกรมวิชาการ กระทรวงศึกษาธิการสำหรับใช้กับนักเรียนระดับประถมศึกษาปีที่ 6 แบบวัดดังกล่าวมีลักษณะเป็นมาตรประมาณค่าแบบสถานการณ์บังคับเลือกตอบ มี 4 ตัวเลือก กำหนดให้คะแนนตามระดับของคุณลักษณะด้านจิตพิสัย (affective domain) ในแต่ละตัวเลือกเป็น 1, 2, 3 และ 4 จากต่ำไปสูงตามลำดับขั้นของคุณลักษณะตามแนวคิดของ แครทโวล และคณะ (Krathwohl and others, 1981) คะแนนแต่ละลำดับขั้นมีความหมายดังนี้

คะแนนลำดับขั้นที่ 1 หมายถึง ระดับขั้นของคุณลักษณะในการรับรู้คุณลักษณะหรือค่านิยมนั้น

คะแนนลำดับขั้นที่ 2 หมายถึง ระดับขั้นของคุณลักษณะในการปฏิบัติตามเงื่อนไขของคุณลักษณะหรือค่านิยมนั้น

คะแนนลำดับขั้นที่ 3 หมายถึง ระดับขั้นของคุณลักษณะในด้านการเห็นประโยชน์ของคุณลักษณะหรือค่านิยมนั้น

คะแนนลำดับขั้นที่ 4 หมายถึง ระดับขั้นของคุณลักษณะในการรู้สึกยึดมั่นในคุณลักษณะหรือค่านิยมนั้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

แบบวัดคุณลักษณะฉบับนี้ประกอบด้วย 2 แบบวัดย่อย คือแบบวัดย่อยด้านการเสียสละ ซึ่งประกอบด้วยข้อกระทงที่ 1-15 และแบบวัดย่อยด้านการมุ่งพัฒนาตนเองประกอบด้วยข้อกระทงที่ 16 - 30 โดยแบบวัดดังกล่าวสำนักงานทดสอบทางการศึกษา กรมวิชาการได้เคยนำไปใช้กับนักเรียนชั้นประถมศึกษาปีที่ 6 ทั่วประเทศ ครั้งแรกในปีการศึกษา 2533 จากนั้นก็นำไปใช้ และวิเคราะห์เพื่อปรับปรุงคุณภาพมาตลอด (ปีการศึกษา 2533, 2535,และ 2537)

การวิเคราะห์คุณภาพแบบวัดดังกล่าววิเคราะห์ตามแนวทฤษฎีการทดสอบแบบดั้งเดิม (CTT) คือ หาค่าอำนาจจำแนกรายข้อโดยใช้การทดสอบที (t-test) การพิจารณาสัดส่วนของการตอบแต่ละตัวเลือก (p) หรือความเป็นตัวแทนของระดับคุณลักษณะของมาตรประมาณค่า และตรวจสอบความตรงเชิงเนื้อหา (content validity) โดยผู้เชี่ยวชาญด้านการสร้างและพัฒนาแบบสอบของกรมวิชาการ และหาความเที่ยงแบบความสอดคล้องภายในของครอนบาค (cronbach alpha : α) ในที่นี้จะเสนอเฉพาะคุณภาพด้านความเที่ยง โดยมีผลการใช้ 3 ครั้งสรุปได้ดังนี้

ครั้งที่ 1 แบบวัดมี 2 ฉบับ ๆ ละ 30 ข้อ (รวม 60 ข้อ) มีค่าความเที่ยง 0.68 และ 0.59 ตามลำดับ

ครั้งที่ 2 แบบวัดมี 2 ฉบับ ๆ ละ 30 ข้อ (รวม 60 ข้อ) มีค่าความเที่ยง 0.65 และ 0.63 ตามลำดับ

ครั้งที่ 3 แบบวัดมี 1 ฉบับ รวม 30 ข้อ (ปรับ 2 ฉบับ เข้าด้วยกัน) มีค่าความเที่ยง 0.74

3.2 แบบสอบคณิตศาสตร์

เป็นแบบสอบวัดการแก้โจทย์ปัญหาคณิตศาสตร์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 ซึ่งพัฒนาขึ้นโดย วลี เฉลยสมัย (2539) แบบสอบดังกล่าวมี 6 ข้อ แต่ละข้อแบ่งเป็น 3 ตอน คือ

ตอนที่ 1 วัดความเข้าใจวิธีการแก้โจทย์ปัญหา

ตอนที่ 2 วัดการคิดคำนวณ

ตอนที่ 3 วัดความสามารถในการวิเคราะห์ตามสภาพที่เป็นจริง

โดยทั้ง 3 ตอนมีสถานการณ์และตัวเลขเหมือนกันหมด เครื่องมือนี้ได้ทดลองใช้และปรับปรุงมาแล้ว 3 ครั้ง ก่อนใช้จริงกับนักเรียนชั้นประถมศึกษาปีที่ 6 ในสังกัดสำนักงานการประถมศึกษาอำเภอพระนครศรีอยุธยา มีผลการใช้แบบสอบดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

การใช้ครั้งที่ 1 เป็นขั้นทดลองใช้เชิงสำรวจกับนักเรียน 283 คน เพื่อสำรวจรูปแบบการคิดแก้โจทย์ปัญหาคณิตศาสตร์ โดยใช้แบบสอบปลายเปิดให้นักเรียนแสดงวิธีทำ มีจำนวน 8 ข้อ จากนั้นจึงนำแบบสอบดังกล่าว มาสร้างเป็นแบบสอบประเภทเลือกตอบ 5 ตัวเลือก มีลักษณะให้ตอบตัวเลือกที่ถูกและบางข้อนักเรียนต้องเติมคำตอบที่ถูกลงไป

การใช้ครั้งที่ 2 นำไปใช้เพื่อวิเคราะห์คุณภาพและปรับปรุง แบบสอบมี 8 ข้อ ๆ ละ 3 ตอน ใช้กับนักเรียนจำนวน 91 คน มีคุณภาพความเที่ยงแบบความสอดคล้องภายใน (KR-20) ดังนี้

ตอนที่ 1 มีความเที่ยง 0.67

ตอนที่ 2 มีความเที่ยง 0.68

ตอนที่ 3 มีความเที่ยง 0.41

รวมทั้งฉบับ (24 ข้อ) มีความเที่ยง 0.79

จากนั้นได้มีการปรับปรุงและตัดข้อกระทงบางข้อออกไป และปรับปรุงบางข้อจนเหลือข้อกระทงชุดละ 6 ข้อ เพื่อนำไปหาคุณภาพในครั้งที่ 3

การใช้ครั้งที่ 3 เป็นการนำไปใช้เพื่อหาคุณภาพก่อนใช้จริง กับนักเรียน 132 คน มีคุณภาพความเที่ยงแบบความสอดคล้องภายใน (KR-20) รวมทั้ง 3 ตอนเป็น 0.80

3. การนำไปใช้จริงกับกลุ่มตัวอย่าง 470 คน มีค่าความเที่ยงแบบ KR-20 ดังนี้ คือ

ตอนที่ 1 มีความเที่ยง 0.67

ตอนที่ 2 มีความเที่ยง 0.57

ตอนที่ 3 มีความเที่ยง 0.56

รวมทั้งฉบับ (18 ข้อ) มีความเที่ยง 0.77

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

การตรวจให้คะแนนแบบพหุวิภาคกับแบบสอบคณิตศาสตร์ในแต่ละข้อของแบบสอบทั้ง 3 ตอน กำหนดเป็น 4 ลำดับขั้น มีความหมายดังนี้

นักเรียนที่ไม่ตอบ หรือตอบไม่ถูกเลยทั้ง 3 ตอน ให้ 1 คะแนน
นักเรียนที่ตอบถูกเพียง 1 ตอน ให้ 2 คะแนน
นักเรียนที่ตอบถูก 2 ตอน ให้ 3 คะแนน
นักเรียนที่ตอบถูกทั้ง 3 ตอน ให้ 4 คะแนน

4. การวิเคราะห์เพื่อคัดเลือกผู้ตอบและข้อกระทงที่เหมาะสม

การวิเคราะห์เพื่อคัดเลือกผู้ตอบและข้อกระทงที่เหมาะสมกับโมเดล ได้เสนอผลการวิเคราะห์เป็น 2 ส่วน คือ ผลการคัดเลือกผู้ตอบและข้อกระทงที่เหมาะสม และผลการวิเคราะห์องค์ประกอบเชิงยืนยัน โดยในบทนี้เสนอเฉพาะผลการคัดเลือกผู้ตอบ และข้อกระทงที่เหมาะสม ส่วนรายละเอียดผลการวิเคราะห์ข้อกระทงที่เหมาะสม การวิเคราะห์องค์ประกอบยืนยัน ผู้วิจัยเสนอในภาคผนวก การคัดเลือกผู้ตอบที่เหมาะสม (person-fit) ข้อกระทงที่เหมาะสม (item fit) วิเคราะห์ด้วยโปรแกรมสำเร็จรูป BIGSTEPS ใช้เกณฑ์การคัดเลือกข้อกระทงที่เหมาะสมตามลำดับของความสำคัญดังนี้ (Linacre and Wright , 1994 :89)

1. ค่า standardized mean square outfit statistic (MNSQ-outfit) ของผู้ตอบหรือข้อกระทงมีค่าไม่เกิน 1.2
2. ค่าสถิติ INFIT และ OUTFIT ของผู้ตอบหรือข้อกระทงมีค่าไม่เกิน |2|
3. ถ้าค่า MNSQ-OUTFIT ของผู้ตอบหรือข้อกระทงมีค่าต่ำมาก ๆ โดยประมาณไม่เกิน 1.0 จะยอมให้ค่าสถิติ INFIT และ OUTFIT มีค่าต่ำกว่า -2 ได้
4. การคัดผู้ตอบและข้อกระทงที่เหมาะสมจะดำเนินการคัดผู้ตอบที่ไม่เหมาะสมออกก่อน จากนั้นจึงนำผลดังกล่าวมาวิเคราะห์ใหม่เพื่อคัดข้อกระทงออกในรอบที่ 2

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ผลการวิเคราะห์ผู้ตอบและข้อกระทงที่เหมาะสม แสดงดังตารางต่อไปนี้

ตารางที่ 4 ผลการวิเคราะห์เพื่อคัดเลือกผู้ตอบที่เหมาะสมจากคะแนนแบบวัดคุณลักษณะและแบบสอบคณิตศาสตร์

เครื่องมือ	ผู้ตอบทั้งหมด		ผู้ตอบที่ไม่เหมาะสม		ผู้ตอบที่เหมาะสม	
	คน	(ร้อยละ)	คน	(ร้อยละ)	คน	(ร้อยละ)
1. แบบวัดคุณลักษณะทั่วไป	6,300	(100)	244	(3.87)	6,056	(96.13)
2. แบบสอบคณิตศาสตร์	470	(100)	26	(5.53)	444	(94.47)

จากตารางที่ 4 จำนวนร้อยละของผู้ตอบที่เหมาะสมจากข้อมูลแบบวัดคุณลักษณะ มีมากกว่าข้อมูลแบบสอบคณิตศาสตร์เล็กน้อย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 5 โครงสร้างของแบบวัดคุณลักษณะหลังคัดเลือกข้อกระทงที่เหมาะสม

ข้อที่	แบบวัดย่อย (subscale)	ผลการคัดเลือกข้อกระทง ไม่เหมาะสม	เหมาะสม	ทั้งหมด ข้อ	ร้อยละ	เหลือ ข้อ	ร้อยละ
1	1. การเสียสละ		/	15	50.00	9	30.00
2			/				
3		/					
4		/					
5		/					
6		/					
7		/					
8			/				
9			/				
10			/				
11			/				
12			/				
13			/				
14		/					
15			/				
16	2. การมุ่งพัฒนาตนเอง		/	15	50.00	13	43.33
17			/				
18			/				
19			/				
20		/					
21			/				

(ต่อ หน้า 52)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 5 (ต่อ)

ข้อที่ (subscale)	แบบวัดย่อย	ผลการคัดเลือกข้อกระทง ไม่เหมาะสม	เหมาะสม	ทั้งหมด ข้อ	ทั้งหมด	เหลือ ข้อ	ร้อยละ
22			/				
23			/				
24			/				
25			/				
26		/					
27			/				
28			/				
29			/				
30			/				
รวม				30	100	22	73.33

จากตารางที่ 5 แบบวัดชุดคัดเลือกข้อกระทงที่เหมาะสม เหลือข้อกระทงทั้งหมด 22 ข้อ (ร้อยละ 73.33) มีข้อกระทงในแบบวัดย่อยที่ 1 จำนวน 9 ข้อ (ร้อยละ 30.00) และในแบบวัดย่อยที่ 2 จำนวน 13 ข้อ(ร้อยละ 43.33) เมื่อพิจารณาแล้วจะเห็นว่ามีข้อกระทงในแบบวัดย่อยที่ 1 และ 2 อยู่ถึงร้อยละ 60.00 และ 86.67 ของจำนวนข้อกระทงเดิมตามลำดับ การคัดเลือกข้อกระทงที่เหมาะสมกับโมเดลดังกล่าวทำให้โครงสร้างของแบบวัดเปลี่ยนไป ดังนั้นจึงได้วิเคราะห์องค์ประกอบเชิงยืนยัน (CFA) เพื่อตรวจสอบความตรงเชิงโครงสร้างของแบบวัดย่อยที่ 1 และ 2 จากข้อมูลชุดที่ 1 และ 2 โดยผลการวิเคราะห์พบว่าโครงสร้างของแบบวัดคุณลักษณะทั้ง 2 แบบวัดย่อย ในข้อมูลชุดที่ 2 ยังมีโครงสร้างของข้อกระทงที่สามารถวัดคุณลักษณะได้ไม่แตกต่างจากแบบวัดย่อยในข้อมูลชุดที่ 1 ดังรายละเอียดผลการวิเคราะห์ในภาคผนวก

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 6 โครงสร้างของแบบสอบคณิตศาตร์ หลังจากการคัดเลือกข้อกระทงที่เหมาะสม

ข้อที่	ผลการคัดเลือกข้อกระทง		ทั้งหมด		เหลือ	
	ไม่เหมาะสม	เหมาะสม	ข้อ	ร้อยละ	ข้อ	ร้อยละ
1	-	/	6	100.00	6	100.00
2	-	/				
3	-	/				
4	-	/				
5	-	/				
6	-	/				

จากตารางที่ 6 พบว่าแบบสอบคณิตศาสตร์ทุกข้อมีความเหมาะสม (item-fit) โครงสร้างของเนื้อหาจึงยังคงเหมือนเดิม

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

5. การตรวจให้คะแนน

การตรวจให้คะแนนแบบวัดคุณลักษณะและแบบสอบคณิตศาสตร์ในการศึกษาครั้งนี้มี 2 วิธี คือ

5.1 การตรวจให้คะแนนแบบทวิวิภาค (dichotomous) เป็นการกำหนดคะแนนของคำตอบที่ถูกที่สุดหรือแสดงถึงระดับของคุณลักษณะสูงสุดเป็น 1 ส่วนคำตอบอื่นๆ กำหนดให้เป็น 0

5.2 การตรวจให้คะแนนแบบพหุวิภาค (polytomous) เป็นการกำหนดให้คะแนนทั้ง 4 ตัวเลือก หรือผลการตอบแบบสอบเป็น 1, 2, 3, และ 4 ตามระดับของคุณลักษณะหรือความสามารถจากต่ำไปสูงตามลำดับ

6. การเก็บรวบรวมข้อมูล

ในการศึกษาครั้งนี้ผู้วิจัยศึกษาจากกลุ่มตัวอย่าง 2 กลุ่ม มีการเก็บรวบรวมข้อมูลเป็น 2 ส่วน คือ

6.1 การเก็บรวบรวมข้อมูลจากการใช้แบบวัดคุณลักษณะ

ผู้วิจัยได้เก็บรวบรวมข้อมูลจากแหล่งทุติยภูมิ ที่สำนักงานทดสอบทางการศึกษา กรมวิชาการ ได้สุ่มไว้แล้วมาทั้งหมด ขนาดกลุ่มตัวอย่างจำนวน 6,300 คน จากนั้นจึงนำข้อมูลดังกล่าวมาวิเคราะห์เพื่อคัดเลือกผู้ตอบ ที่ไม่เหมาะสม (person-misfit) และข้อกระทงที่ไม่เหมาะสม (item-misfit) ออก โดยใช้โปรแกรมสำเร็จรูป BIGSTEPS

6.2 การเก็บรวบรวมข้อมูลจากการใช้แบบสอบคณิตศาสตร์

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลด้วยตนเองโดยใช้แบบทดสอบคณิตศาสตร์ที่พัฒนาโดย วลี เฉลยสมัย (2539) ศึกษากับกลุ่มตัวอย่างซึ่งเป็นนักเรียนระดับชั้นประถมศึกษาปีที่ 6 สังกัดสำนักงานการประถมศึกษาอำเภอพระนครศรีอยุธยา จำนวน 470 คน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

7. การวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลผู้วิจัยได้ดำเนินการตามขั้นตอนดังนี้

7.1 วิเคราะห์ข้อมูลเบื้องต้น

7.1.1 วิเคราะห์ค่าสถิติบรรยายของคะแนนนักเรียนจากการใช้แบบวัดคุณลักษณะทั่วไป และแบบสอบคณิตศาสตร์โดยคำนวณค่าคะแนนเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด ค่าต่ำสุด ความเบ้ และความโด่ง

7.1.2 วิเคราะห์เพื่อคัดเลือกผู้ตอบที่เหมาะสม (person fit) และข้อกระทงที่เหมาะสม (item fit) โดยใช้โปรแกรมสำเร็จรูป BIGSTEPS ค่าสถิติที่ต้องการคำนวณ คือ

ก. ค่า Standardized Mean Square Outfit Statistic (MNSQ outfit)

ข. ค่าสถิติ INFIT และ OUTFIT ซึ่งเป็นค่าคะแนนความคลาดเคลื่อนมาตรฐาน (standardized residual) ในการคัดเลือกผู้ตอบที่เหมาะสมและข้อกระทงที่เหมาะสม ผู้วิจัยใช้เกณฑ์ในการพิจารณา ดังนี้

1 ค่า MNSQ outfit ของผู้ตอบ หรือข้อกระทงที่มีค่าไม่เกิน 1.20

2 ค่าสถิติ INFIT และ OUTFIT ของผู้ตอบ หรือข้อกระทงมีค่าไม่เกิน |2|

3 ถ้าค่า MNSQ outfit มีค่าต่ำมาก ๆ โดยประมาณไม่เกิน 1.00 จะยอมให้ค่าสถิติ INFIT และ OUTFIT มีค่าต่ำกว่า -2 ได้

4 ในขั้นตอนการคัดเลือกผู้ตอบที่ไม่เหมาะสม (person fit) และข้อกระทงที่ไม่เหมาะสม (item misfit) จะดำเนินการคัดเลือกผู้ตอบที่ไม่เหมาะสมออกก่อน จากนั้นจึงนำข้อมูลที่เหลือมาวิเคราะห์ เพื่อคัดข้อกระทงออกอีกครั้งหนึ่ง

7.1.3 การวิเคราะห์องค์ประกอบยืนยัน (confirmatory factor analysis :CFA) เป็นการวิเคราะห์เพื่อตรวจสอบความตรงเชิงโครงสร้าง (construct validity) ของแบบวัดคุณลักษณะจากข้อมูลชุดที่ 1 และ 2 ในแบบวัดย่อยที่ 1 และ 2 ตามลำดับโดยใช้สถิติทดสอบไค-สแคว์ (chi-square) วิเคราะห์ด้วยโปรแกรมสำเร็จรูป LISREL 8.10

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

7.2 การวิเคราะห์หาค่าฟังก์ชันสารสนเทศของแบบวัดคุณลักษณะและแบบสอบชุดที่ 1 และ 2 ด้วย GRM, GPCM และโมเดลโลจิสติก

7.2.1 การวิเคราะห์ด้วย GRM

$$I(\theta) = \sum_{j=1}^{m} I_j(\theta)$$

เมื่อ $I_j(\theta)$ คือ ฟังก์ชันสารสนเทศของข้อกระทง

$$I_j(\theta) = \sum_{jk=0}^{k} \frac{[P'_{jk}(\theta)]^2}{P_{jk}(\theta)}$$

m คือ จำนวนข้อกระทงทั้งฉบับ

k คือ ลำดับขั้นคะแนนที่ 0,1,...k

$P'_{jk}(\theta)$ คือ ความน่าจะเป็นของคนที่มีความสามารถ θ จะตอบข้อกระทง j ในคะแนนลำดับขั้นที่ k ได้ถูก

$P_{jk}(\theta)$ คือ ความน่าจะเป็นของคนที่มีความสามารถ θ จะตอบข้อกระทง j ได้คะแนน k

7.2.2 การวิเคราะห์ด้วย GPCM

$$I(\theta) = \sum_{j=1}^{m} I_j(\theta)$$

เมื่อ $I_j(\theta)$ คือ ฟังก์ชันสารสนเทศของข้อกระทง

$$I_j(\theta) = D^2 a_j^2 \sum_{k=1}^{mj} [T_k - \overline{T}_j(\theta)]^2 P_{jk}(\theta)$$

D คือ ค่าคงที่ของสเกลมีค่าประมาณ 1.70

a_j คือ ค่าอำนาจจำแนกข้อกระทง

$\overline{T}_j(\theta)$ **คือ** $\sum T_k P_{jk}(\theta)$

T_k คือ ค่าคะแนนใดๆ ใน T_j=(1,2,3,...,mj)

$P_{jk}(\theta)$ คือ ความน่าจะเป็นของคนที่มีความสามารถ จะตอบข้อกระทง j ได้คะแนน k

k คือ ลำดับขั้นคะแนนที่ 1,2,...k-1, k, k+1,...,mj

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

7.2.3 การวิเคราะห์ด้วยโมเดลโลจิสติก

$$I(\theta) = \frac{\sum_{j=1}^{m}[P_j(\theta)]^2}{P_j(\theta)Q_j(\theta)}$$

เมื่อ m คือ จำนวนข้อสอบ

$P_j(\theta)$ คือ ความน่าจะเป็นที่ผู้ตอบมีความสามารถ θ จะตอบข้อกระทง j ได้ถูกต้อง

$Q_j(\theta)$ คือ $1 - P_j(\theta)$

7.3 การวิเคราะห์หาอัตราส่วนสารสนเทศเฉลี่ย (ratio of average information : RAI) มีการวิเคราะห์ดังนี้

7.3.1 คำนวณค่าสารสนเทศเฉลี่ย (average information: AI) จากสูตร

$$AI(\theta, x_i) = \frac{\sum_{qpt=1}^{n} I(\theta, qpt_n)}{n}$$

เมื่อ qpt คือ Quardrature Point ใดๆ ของ θ บนโค้งฟังก์ชันสารสนเทศ

n คือ จำนวนจุดของ QPT บนโค้งฟังก์ชันสารสนเทศของแบบวัด/แบบสอบ

x_i คือ แบบวัด/แบบสอบ ชุด x ใด ๆ

7.3.2 คำนวณอัตราส่วนสารสนเทศเฉลี่ย (RAI) จากสูตร

$$RAI(\theta, x_i, y_i) = \frac{AI(\theta, x_i)}{AI(\theta, y_i)}$$

เมื่อ AI คือ ค่าเฉลี่ยสารสนเทศของแบบวัดหรือแบบสอบ

x_i คือ แบบวัดหรือแบบสอบชุด x

y_i คือ แบบวัดหรือแบบสอบชุด y

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

7.3.3 การแปลความหมายอัตราส่วนสารสนเทศเฉลี่ย

ก. ถ้าค่า $RAI(\theta, x_i, y_i) = 1$ แสดงว่าแบบวัดหรือแบบสอบที่ใช้การตรวจให้คะแนนหรือโมเดลที่วิเคราะห์ทั้ง 2 วิธี โดยเฉลี่ยมีประสิทธิภาพสัมพัทธ์เท่ากัน

ข. ถ้าค่า $RAI(\theta, x_i, y_i) > 1$ แสดงว่าแบบวัดหรือแบบสอบที่ใช้การตรวจให้คะแนนหรือโมเดลวิเคราะห์โดยวิธี X โดยเฉลี่ยมีประสิทธิภาพสัมพัทธ์สูงกว่าแบบ Y

ค. ถ้าค่า $RAI(\theta, x_i, y_i) < 1$ แสดงว่าแบบวัดหรือแบบสอบที่ใช้การตรวจให้คะแนนหรือโมเดลวิเคราะห์โดยวิธี Y โดยเฉลี่ยมีประสิทธิภาพสัมพัทธ์สูงกว่าแบบ X

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย