

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของเนื้อหา

ทฤษฎีทางการวัดผลการศึกษา เป็นทฤษฎีว่าด้วยการวัดคุณลักษณะแฝง (Latent Trait) ของมนุษย์ เนื่องจากการวัดผลการศึกษา เป็นการวัดทางพฤติกรรมศาสตร์หรือลักษณะทางจิตวิทยาที่ต้องการวัดคุณสมบัติต่าง ๆ ของผู้เรียน เช่น ความสามารถ ความถนัด ความฉลาด ฯลฯ ซึ่งสิ่งเหล่านี้มีลักษณะเป็นนามธรรมและเมื่อประกอบที่สลับซับซ้อนไม่สามารถวัดได้โดยตรง จึงต้องใช้วิธีการทางอ้อมโดยการวัดสิ่งที่มีความสัมพันธ์เชิงปริมาณกับสิ่งที่ต้องการจะวัด แล้วนำผลที่ได้มาอธิบายย้อนกลับไปสู่สิ่งที่ต้องการวัดจริง ๆ นั้น เครื่องมือที่สำคัญที่ใช้ในทางการวัดผลการศึกษาคือ แบบทดสอบ โดยให้บุคคลแสดงพฤติกรรมออกมาด้วยการตอบแบบทดสอบเพื่อนำผลการตอบนั้นไปอธิบายหรือทำนายความสามารถที่แท้จริง (true ability or true score) อันเป็นคุณลักษณะแฝงของบุคคล

ในการทำนายหรืออธิบายความสามารถที่แท้จริงนั้น นักวัดผลได้พยายามสร้างเครื่องมือและพัฒนาทฤษฎีต่าง ๆ ขึ้นมา เพื่อที่จะทำนายความสามารถของบุคคลให้มีความถูกต้องและเชื่อถือได้ ทฤษฎีการวัดผลในเชิงไทยที่ใช้กันอยู่ในปัจจุบันนี้ส่วนใหญ่คือทฤษฎีเทม (The Classical Test Theory) ซึ่งมีจุดอ่อนอยู่หลายประการในการประมาณค่าความสามารถที่แท้จริงของบุคคล ซึ่งแฮมเบิลตัน (Hambleton , 1979 : 14-15) กล่าวไว้ดังนี้

1. คุณสมบัติของข้อคำถาม หรือค่าสถิติของข้อสอบ (Item Statistics) ได้แก่ ค่าความยาก (p) ค่าอำนาจจำแนก (r) เป็นค่าไม่คงที่จะแปรเปลี่ยนไปตามกลุ่มผู้สอบ ดังนั้นการใช้ค่าสถิติของข้อสอบจึงใช้ได้กับกลุ่มผู้สอบที่มีลักษณะเหมือนกับกลุ่มที่นำมาใช้ในการประมาณค่าสถิติของข้อสอบเท่านั้น

2. คะแนนของผู้สอบแต่ละคนขึ้นอยู่กับข้อสอบที่สุ่มมาในการสอบแต่ละครั้ง ดังนั้นในการเปรียบเทียบความสามารถของแต่ละบุคคลจะเปรียบเทียบกันได้ก็ต่อเมื่อใช้แบบทดสอบที่มีข้อคำถามเดียวกัน หรือเป็นแบบทดสอบคู่ขนานเท่านั้น

3. ความคลาดเคลื่อนมาตรฐานในการวัดของผู้สอบแต่ละคนเท่ากัน ซึ่งขัดแย้งกับ ลอร์ด และ โนวิค (Lord and Novick, 1968 : 183) ที่ชี้ให้เห็นว่า ความคลาดเคลื่อนมาตรฐานในการวัดของแต่ละคนมีค่าไม่เท่ากัน กล่าวคือ ความคลาดเคลื่อนมาตรฐานจะมีค่ามากสำหรับคะแนนปานกลาง และจะมีค่าน้อยสำหรับคะแนนสูง หรือคะแนนต่ำ การตีความหมายจากข้อมูลตามทฤษฎีนี้จะต้องระมัดระวังมากเพราะ อาจเกิดผิดพลาดได้ง่าย

นอกจากนี้ยังไม่สามารถที่จะให้คำตอบที่ชัดเจนต่อปัญหาหลาย ๆ อย่าง เช่น ในเรื่องของคะแนนที่ไค การเทียบคะแนน การวิเคราะห์รายข้อ และการจัดข้อสอบตามความสามารถของผู้สอบ (นิสิตศึกษานิเทศ จุฬาลงกรณ์มหาวิทยาลัย, 2525 : 2)

จากปัญหาที่กล่าวมาที่ทฤษฎีเดิม (The Classical Test Theory) ก็ยังไม่สามารถเสนอแนวทางในการแก้ไขได้ ดังนั้นนักวิจัยจึงได้พยายามเสนอแนวความคิดใหม่ หรือทฤษฎีใหม่เพื่อแก้ปัญหาดังกล่าว ทฤษฎีหนึ่งที่นักวิจัยได้ให้ความสนใจ และช่วยกันพัฒนาขึ้นมาคือ ทฤษฎีลาเทนต์เทรท หรือที่เรียกกันว่าทฤษฎีคุณลักษณะแฝง (Latent Trait Theory) หรือทฤษฎีการตอบข้อสอบ (Item Response Theory : IRT) ผู้ริเริ่มทฤษฎีนี้คือ ลอร์ด (Frederic M. Lord) โดยที่ทฤษฎีนี้อธิบายความสามารถที่แท้จริงของบุคคล และคุณสมบัติของข้อสอบตามคุณลักษณะหรือตัวแปรที่ทำการศึกษาโดยอาศัยโมเดลหรือฟังก์ชันทางคณิตศาสตร์ ซึ่งพารามิเตอร์ที่เกี่ยวข้องกับความยากง่ายของข้อสอบกับความสามารถของบุคคล จะเป็นตัวกำหนดความน่าจะเป็นที่บุคคลจะตอบข้อสอบข้อนั้นได้ถูก (นิสิตศึกษานิเทศ จุฬาลงกรณ์มหาวิทยาลัย, 2525 : 2)

ต่อมานักวิจัยหลายคนรวมทั้งลาซาเฟล (Lazarsfeld) และคณะได้พัฒนาทฤษฎีนี้ไปในรูปแบบต่าง ๆ กัน เรียกว่าแบบโครงสร้างแฝงสามัญ (The General Latent Structure Model) และโมเดลเฉพาะอีก 5 โมเดล

1. แบบเส้นตรง (The Linear Model)
2. แบบความสามารถแฝง (The Latent Distance Model)
3. แบบโลจิสติกปกติ (The Normal Ogive Model)
4. แบบโลจิสติก (The Logistic Model)
5. แบบราซช (The Rasch Model)

ปัจจุบันได้มีผู้นำทฤษฎีนี้ไปใช้กันมากขึ้น เพราะมีความเชื่อเช่นเดียวกับแฮมเบิลตัน (Hambleton, 1979 : 16-32) ว่าถ้านำทฤษฎีนี้มาใช้อย่างถูกต้องกับข้อตกลงเบื้องต้นของแต่ละโมเดลแล้วจะสามารถแก้ปัญหาที่เกิดจากทฤษฎีเดิมได้ โดยที่ทฤษฎีนี้มีจุดเด่นที่น่าสนใจคือ (สงวน ลักษณะ, 2525 : 21-22)

1. คุณสมบัติของข้อสอบ หรือ ค่าพารามิเตอร์ของข้อสอบแต่ละข้อ ได้แก่ ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าการเดา (c) เป็นค่าคงที่ไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ (Invariant of Item Parameter) นั่นคือ ถ้านำข้อสอบชุดเดิมไปใช้กับผู้สอบต่างกลุ่มกันแล้ว ปรากฏว่า ค่าพารามิเตอร์ของข้อสอบเปลี่ยนไป การเปลี่ยนนี้เป็นเพราะอิทธิพลของการมีจุดตั้งต้น (Origin) ของค่าความสามารถที่แท้จริงของผู้สอบ (θ) บนแกนราบต่างกันไปเท่านั้น หากมีการปรับ (Transform) ค่าความสามารถที่แท้จริงของผู้สอบ (θ) ให้มีสเกลเดียวกันแล้ว ค่าพารามิเตอร์ของข้อสอบจากการสอบกับผู้สอบต่างกลุ่มกันจะมีค่าคงเดิมเสมอ

2. สามารถคาดคะเนลักษณะการตอบถูกของบุคคลต่างกลุ่มใด หรือสามารถใช้ข้อสอบกับใครก็ได้ (Person-Free) แม้ว่าบุคคลนั้นจะไม่เคยทดลองสอบในข้อสอบฉบับนั้นมาก่อนเลยก็ตาม นั่นคือ โครงลักษณะเฉพาะของข้อสอบจะคงเดิม แม้ว่าจะนำไปใช้ในการสอบกับบุคคลต่างกลุ่ม

3. สามารถใช้ข้อสอบข้อใดก็ได้ (Item-Free) จำนวนไม่มากขอจากแหล่งข้อสอบ (Item pool) ที่มีข้อสอบวัดสิ่งเดียวกันเป็นจำนวนหลาย ๆ ข้อมาใช้ในการประมาณค่าความสามารถที่แท้จริงของผู้สอบ (θ) ซึ่งไรท์ (Wright, 1977 : 14) กล่าวว่าแม้จะ

ใช้ข้อสอบประมาณ 3-5 ข้อ จะเป็นข้อสอบข้อใดก็ได้ ก็สามารถคาดคะเนความสามารถที่แท้จริงของผู้สอบ (๑) ได้

จากลักษณะเด่นบางประการที่กล่าวมานี้ ทำให้สามารถนำทฤษฎีนี้ไปประยุกต์ใช้ในการแก้ปัญหาทางการวัดผลในสถานต่าง ๆ สำหรับการวัดผลในประเทศไทยนั้น นักการศึกษาและนักวัดผลหลายคนได้เริ่มให้ความสนใจและศึกษาทฤษฎีนี้ และโดยทั่วไปยังไม่มี การประยุกต์ทฤษฎีนี้มาใช้ในการวิเคราะห์ข้อสอบ ดังนั้นผู้วิจัยจึงสนใจที่จะนำโลจิสติกโมเดล (The Logistic Model) ซึ่งเป็นโมเดลหนึ่งของทฤษฎีการตอบข้อสอบ (Item Response Theory : IRT) และ เป็นโมเดลที่มีโมเดลย่อยอยู่ถึง 3 โมเดล คือ แบบพารามิเตอร์ตัวเดียว (One-parameter Logistic Model) แบบพารามิเตอร์สองตัว (Two-parameter Logistic Model) และแบบพารามิเตอร์สามตัว (Three-parameter Logistic Model) มาประยุกต์ใช้ในการวิเคราะห์ข้อสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ซึ่งผู้วิจัยสร้างแบบทดสอบขึ้นเอง โดยมีจุดมุ่งหมายที่จะเปรียบเทียบผลที่ได้จากการวิเคราะห์ของแต่ละโมเดลย่อย และเปรียบเทียบผลที่ได้จากการวิเคราะห์ของแต่ละโมเดลย่อยกับการวิเคราะห์โดยใช้วิธีแบบเดิม

จุดมุ่งหมายในการวิจัย

1. เพื่อเปรียบเทียบจำนวนข้อของข้อสอบที่คัดเลือกไว้จากการวิเคราะห์โดยวิธี
 - 1.1 แบบพารามิเตอร์ตัวเดียว สองตัว และ สามตัว
 - 1.2 แบบเดิม กับ แบบพารามิเตอร์ตัวเดียว
 - 1.3 แบบเดิม กับ แบบพารามิเตอร์สองตัว
 - 1.4 แบบเดิม กับ แบบพารามิเตอร์สามตัว
2. เพื่อเปรียบเทียบความเชื่อมั่นของข้อสอบที่คัดเลือกไว้จากการวิเคราะห์โดยวิธี
 - 2.1 แบบพารามิเตอร์ตัวเดียว สองตัว และสามตัว โดยเปรียบเทียบทีละคู่
 - 2.2 แบบเดิม กับ แบบพารามิเตอร์ตัวเดียว
 - 2.3 แบบเดิม กับ แบบพารามิเตอร์สองตัว
 - 2.4 แบบเดิม กับ แบบพารามิเตอร์สามตัว

3. เพื่อเปรียบเทียบความสอดคล้องของการคัดเลือกข้อสอบจากการวิเคราะห์โดยวิธี
 - 3.1 แบบพหุรามิเตอร์ตัวเดียว สองตัว และ สามตัว โดยเปรียบเทียบทีละคู่
 - 3.2 แบบเคิม กับ แบบพหุรามิเตอร์ตัวเดียว
 - 3.3 แบบเคิม กับ แบบพหุรามิเตอร์สองตัว
 - 3.4 แบบเคิม กับ แบบพหุรามิเตอร์สามตัว

ขอบเขตในการวิจัย

1. การศึกษารุ่นนี้เป็นการเปรียบเทียบผลการวิเคราะห์ข้อสอบชนิดเลือกตอบ 5
ตัวเลือก ด้วยวิธีโลจิสติกโมเดล กับ วิธีแบบเคิม โดยใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน
วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 รายวิชา ค 101 การเปรียบเทียบผลการวิเคราะห์ข้อสอบ
ในครั้งนี เป็นการศึกษาเปรียบเทียบในเรื่องของจำนวนข้อของข้อสอบที่คัดเลือกไว้ ความเชื่อมั่นและ
ความสอดคล้องของการคัดเลือกข้อสอบจากการวิเคราะห์ข้อสอบแต่ละวิธีเท่านั้น
2. ประชากรที่ใช้ในการศึกษารุ่นนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา
2527 ในอำเภอเมือง จังหวัดเชียงใหม่
3. การศึกษารุ่นนี้ไม่คำนึงถึงผลความแตกต่างของตัวแปรอื่น ๆ เช่น เพศ อายุ
ประเภทของโรงเรียน

ข้อตกลงเบื้องต้น

1. เกณฑ์ที่ใช้ในการคัดเลือกข้อสอบที่มีคุณภาพจากการวิเคราะห์ข้อสอบโดยวิธีแบบเคิม
(The Classical Model)

ข้อสอบที่มีคุณภาพดีและได้รับการคัดเลือกไว้ต้องมีค่าความยากระหว่าง 0.20 ถึง
0.80 และมีค่าอำนาจจำแนกตั้งแต่ 0.20 ถึง 1.00 (Ebel, 1965 : 364)

2. เกณฑ์ที่ใช้ในการคัดเลือกข้อสอบที่มีคุณภาพจากการวิเคราะห์ข้อสอบโดยวิธีโลจิสติก-
โมเดล (The Logistic Models)

2.1 ข้อสอบที่มีคุณภาพดีและได้รับการคัดเลือกไว้จะต้องเหมาะสม (Fit) กับ โมเดล กล่าวคือ เส้นถดถอยรายข้อ (Item test regression) ปรับเข้าสู่รูปของเส้น โคงของข้อสอบ (Item Characteristic Curve : ICC) โดยการทดสอบด้วย ค่าไคสแควร์ (χ^2 -Test) หรือทดสอบด้วยค่าที (t-Test) (Wright and Stone, 1979 : 71-77) หรืออาจจะทดสอบด้วยค่าเอฟ (F-Test) (Hashway, 1978 : 21) แล้วมีความแตกต่างอย่างไม่มีนัยสำคัญที่ระดับ .01

2.2 คัมพี้ของค่าพารามิเตอร์ ตามปกติค่าความสามารถในการตอบข้อสอบ (θ) ค่าความยาก (b) และ ค่าอำนาจจำแนก (a) ของข้อสอบจะมีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติ (Ree, 1979 : 372)

ค่าความสามารถในการตอบข้อสอบ (θ)	มีค่าระหว่าง -3.00 ถึง +3.00
ค่าความยากของข้อสอบ (b)	มีค่าระหว่าง -2.00 ถึง +2.00
ค่าอำนาจจำแนกของข้อสอบ (a)	มีค่าระหว่าง 0.50 ถึง 2.50
และ ค่าการแจกของข้อสอบ (c)	มีค่าระหว่าง 0.00 ถึง 0.30

สมมติฐานในการวิจัย

1. จำนวนข้อของข้อสอบที่คัดเลือกไว้จากการวิเคราะห์ข้อสอบแบบพารามิเตอร์ตัวเดียว สองตัว และ สามตัว มีจำนวนไม่แตกต่างกัน
2. จำนวนข้อของข้อสอบที่คัดเลือกไว้จากการวิเคราะห์ข้อสอบแบบเดิมกับแบบพารามิเตอร์ตัวเดียว แบบเดิมกับแบบพารามิเตอร์สองตัวและแบบเดิมกับแบบพารามิเตอร์สามตัว มีจำนวนแตกต่างกัน
3. ความเชื่อมั่นของข้อสอบที่คัดเลือกไว้จากการวิเคราะห์ข้อสอบแบบพารามิเตอร์ตัวเดียว สองตัว และ สามตัว มีค่าไม่แตกต่างกัน
4. ความเชื่อมั่นของข้อสอบที่คัดเลือกไว้จากการวิเคราะห์ข้อสอบแบบเดิมกับแบบพารามิเตอร์ตัวเดียว แบบเดิมกับแบบพารามิเตอร์สองตัว และ แบบเดิมกับแบบพารามิเตอร์สามตัว มีค่าแตกต่างกัน

5. การคัดเลือกข้อสอบจากการวิเคราะห์ข้อสอบแบบพารามิเตอร์ตัวเดียว สองตัว และ สามตัว แบบเดิมกับแบบพารามิเตอร์ตัวเดียว แบบเดิมกับแบบพารามิเตอร์สองตัว และ แบบเดิมกับแบบพารามิเตอร์สามตัว มีความสอดคล้องกัน

ข้อจำกัดในการวิจัย

การวิเคราะห์คุณภาพของข้อสอบ ตามทฤษฎีเกมและทฤษฎีการตอบข้อสอบนั้นแตกต่างกัน โดยที่ทฤษฎีเกมศึกษาในเรื่องของความเชื่อมั่น (Reliability) และความคลาดเคลื่อนมาตรฐาน (Standard Error Measurement) แต่ในทฤษฎีการตอบข้อสอบศึกษาในเรื่องของอินโฟเมชันฟังก์ชัน (Test Information Function) แทนค่าความเชื่อมั่น (Hambleton, 1977 : 64) ดังนั้นในการวิจัยจึงไม่สามารถเปรียบเทียบค่าความเชื่อมั่นได้โดยตรง และในการศึกษาทฤษฎีการตอบข้อสอบในระยะแรก ๆ นี้ ผู้วิจัยทั้งหลายประมาณค่าโดยใช้ความคงที่ภายใน (Internal Consistency Reliability) ด้วยวิธีการเดียวกันกับทฤษฎีเกม (Hashway, 1978 : 25) คือประมาณค่าโดยใช้สูตร กูเดอร์ ริชาร์ดสัน สูตรที่ 20 (Kuder-Richardson : KR 20) แอลฟา (Coefficient Alpha) และการวิเคราะห์ค่าความแปรปรวนของฮอยท์ (Hoyt's Analysis of Variance Procedure)

ในการวิจัยครั้งนี้ ผู้วิจัยจึงเลือกการประมาณค่าความเชื่อมั่นโดยใช้สูตรแอลฟา (Coefficient Alpha) โดยนำข้อสอบที่ได้รับการคัดเลือกไว้จากแต่ละโมเดลไปประมาณค่าความเชื่อมั่นโดยใช้สูตร แอลฟา (Coefficient Alpha) ดังนั้นการเปรียบเทียบค่าความเชื่อมั่นในการวิจัยครั้งนี้ จึงเป็นการเปรียบเทียบค่าความเชื่อมั่นที่ประมาณค่าโดยใช้สูตร แอลฟา (Coefficient Alpha) เท่านั้น

นิยามศัพท์เฉพาะ

1. การวิเคราะห์ข้อสอบแบบเดิม หมายถึง การวิเคราะห์ข้อสอบเพื่อหาค่าความยากของข้อสอบ โดยใช้สัดส่วนหรือร้อยละของผู้สอบทั้งหมดที่ตอบข้อสอบข้อนั้นถูก ค่าอำนาจจำแนกของข้อสอบโดยใช้สูตรสหสัมพันธ์ ไบซีเรียล (Biserial Correlation) และค่าความเชื่อมั่นของข้อสอบโดยใช้สูตรแอลฟา (Coefficient Alpha)
 - 1.1 ค่าความยากของข้อสอบ หมายถึง สัดส่วนของผู้ที่เข้าสอบทั้งหมดที่ตอบข้อสอบแต่ละข้อถูก จะมีค่าอยู่ระหว่าง 0 ถึง 1
 - 1.2 ค่าอำนาจจำแนก หมายถึง คุณลักษณะของข้อสอบที่สามารถจำแนกผู้ที่ได้คะแนนสูงและผู้ที่ได้คะแนนต่ำในแต่ละข้อ จะมีค่าตั้งแต่ -1.00 ถึง $+1.00$
 - 1.3 ค่าความเชื่อมั่น หมายถึง ความคงที่แน่นอนในการตอบของผู้สอบที่ตอบข้อสอบทุกข้อ โดยการใช้ข้อสอบเพียงครั้งเดียวและฉบับเดียว
2. การวิเคราะห์ข้อสอบตามวิธีโลจิสติกโมเดล หมายถึง การวิเคราะห์ข้อสอบเพื่อหาข้อสอบที่เหมาะสม (Fit) กับโมเดล และหาค่าพารามิเตอร์ของข้อสอบ
 - 2.1 เส้นโค้งลักษณะรายข้อ หมายถึง เส้นกราฟที่เกิดจากความสามารถที่แท้จริงกับโอกาสในการตอบข้อสอบข้อนั้นได้ถูกต้อง
 - 2.2 พารามิเตอร์ของข้อสอบ หมายถึง คุณลักษณะของข้อสอบแต่ละข้อ ขึ้นอยู่กับ การพิจารณาค่าพารามิเตอร์ในการวิเคราะห์ของแต่ละโมเดล อันได้แก่ ค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าการเดาของข้อสอบ (c)
 - 2.3 การวิเคราะห์ข้อสอบแบบพารามิเตอร์ตัวเดียว หมายถึง การวิเคราะห์ข้อสอบตามวิธีโลจิสติกโมเดล โดยพิจารณาค่าความยาก (b) ของข้อสอบเพียงค่าเดียว
 - 2.4 การวิเคราะห์ข้อสอบแบบพารามิเตอร์สองตัว หมายถึง การวิเคราะห์ข้อสอบตามวิธีโลจิสติกโมเดล โดยพิจารณาค่าความยาก (b) และค่าอำนาจจำแนก (a) ของข้อสอบ
 - 2.5 การวิเคราะห์ข้อสอบแบบพารามิเตอร์สามตัว หมายถึง การวิเคราะห์ข้อสอบตามวิธีโลจิสติกโมเดล โดยพิจารณาค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าการเดา (c) ของข้อสอบ

ประโยชน์ที่จะได้รับในการวิจัย

1. ทำให้ทราบว่าข้อสอบที่ได้รับการคัดเลือกไว้จากการวิเคราะห์โดยวิธีใดที่จะให้คุณภาพสูงที่สุด อันจะเป็นแนวทางในการเลือกใช้วิธีการวิเคราะห์ข้อสอบได้อย่างมีประสิทธิภาพ
2. เป็นการเผยแพร่วิธีการวิเคราะห์ข้อสอบตามทฤษฎีใหม่ให้เป็นที่รู้จักกันมากขึ้น
3. เป็นแนวทางแก่ผู้ที่สนใจจะศึกษาค้นคว้าในเรื่องนี้ ซึ่งจะเป็นการพัฒนาทางด้านทฤษฎี และการปฏิบัติเกี่ยวกับทางการวัดผลต่อไป

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
 Copyright© by Chiang Mai University
 All rights reserved.