

บทที่ 1

บทนำ



ความเป็นมาและความสำคัญของปัญหา

การวัดในทางพฤติกรรมศาสตร์ โดยเฉพาะอย่างยิ่งการวัดผลทางการศึกษา มักจะต้องวัดสิ่งที่เป็นคุณลักษณะภายใน (Latent traits) ซึ่งไม่สามารถจะสังเกตเห็นได้โดยตรง ต้องใช้วิธีวัดทางอ้อม โดยวัดสิ่งอื่นที่มีความสัมพันธ์ในเชิงปริมาณกับสิ่งที่ต้องการวัดแทน แล้วนำผลการวัดนั้นอ้างอิง อธิบายกลับเข้าสู่สิ่งที่ต้องการวัดจริง ดังนั้นทฤษฎีวัดผลทั้งหลายจึงถือเป็นทฤษฎีหนึ่งของทฤษฎีคุณลักษณะภายใน (Latent trait theory)

แบบสอบเป็นสิ่งหนึ่งที่เราสามารถใช้กระตุ้นให้บุคคลแสดงพฤติกรรมภายนอกด้วยการตอบแบบสอบ เพื่อนำเอาพฤติกรรมจำนวนข้อที่ทำถูกโยง ไปหาความสามารถที่แท้จริง ซึ่งเป็นคุณลักษณะภายในของเขา นักวัดผลทั้งหลายจึงพยายามสร้างทฤษฎีขึ้นมาเพื่ออธิบายหรือทำนายความสามารถที่แท้จริง (true ability or true score) ของบุคคลให้มีความแม่นยำเชื่อถือได้

ในปัจจุบันทฤษฎีวัดผลได้มีการพัฒนาขึ้นอย่างกว้างขวาง ทั้งนี้เพราะทฤษฎีที่เคยใช้กันมาจนถึงปัจจุบัน ที่เรียกว่าทฤษฎีคลาสสิกอล (Classical Test Theory) นั้นเน้นการประมาณค่าคะแนนจริงหรือความสามารถที่แท้จริง โดยใช้คะแนนของกลุ่มของผู้สอบเป็นสำคัญ *ถูกมองว่าเป็นทฤษฎีที่ไม่เพียงพอต่อการประมาณค่าความสามารถที่แท้จริง (สงวน ลักษณะ 2525 : 47) และมีจุดอ่อนหลายประการ ซึ่งแฮมเบิลตัน (Hambleton 1979 : 14-15) ได้กล่าวไว้ดังนี้

1. ค่าสถิติของข้อกระทง (Item statistics) ได้แก่ ค่าความยาก ค่าอำนาจจำแนกของข้อกระทงขึ้นอยู่กับกลุ่มตัวอย่างที่สอบ ดังนั้นการใช้ค่าสถิติของข้อกระทงแบบนี้ให้ได้ผลจึงจำกัดอยู่เฉพาะกลุ่มที่มีลักษณะเหมือนกับกลุ่มที่นำมาใช้ในการหาค่าสถิติของข้อกระทงเท่านั้น

2. การเปรียบเทียบความสามารถหรือคุณลักษณะ (trait) ใด ๆ ในแต่ละบุคคลมีข้อจำกัด คือจะเปรียบเทียบกันได้ก็ต่อเมื่อสอบด้วยแบบสอบฉบับเดียวกัน

หรือแบบสอบที่เป็นคู่ขนานกัน ทั้งนี้เพราะคะแนนของผู้สอบแต่ละคนขึ้นอยู่กับชอกระงที่สุ่มมาในการทดสอบแต่ละครั้ง

3. ความคลาดเคลื่อนมาตรฐานของการวัดของผู้สอบแต่ละคนเท่ากัน ซึ่ง ลอร์ดและโนวิก (Lord and Novick 1968) ได้ชี้ให้เห็นว่าความคลาดเคลื่อนมาตรฐานในการวัดของผู้สอบแต่ละคนมีค่าไม่เท่ากัน กล่าวคือ ค่าความคลาดเคลื่อนมาตรฐานในการวัดสำหรับคะแนนที่มีค่าปานกลาง จะมีขนาดเล็กกว่าค่าความคลาดเคลื่อนมาตรฐานในการวัดสำหรับคะแนนที่มีค่าสูงหรือคะแนนที่มีค่าต่ำ

จากจุดอ่อนดังกล่าว ทำให้นักวัดผลจำนวนมากสนใจและพัฒนาทฤษฎีคุณลักษณะภายในของ เฟรดเคอร์ริก เอ็ม ลอร์ด (Frederick-M. Lord) ที่เรียกว่าโค้งลักษณะของชอกระง (Item characteristic curve) หรือทฤษฎีการตอบชอกระง (Item response theory : IRT) ซึ่งลอร์ดได้เสนอในปี 1952 ซึ่งเป็นทฤษฎีที่ใช้ โมเดลทางคณิตศาสตร์ แสดงความสัมพันธ์ระหว่างความสามารถที่แท้จริงกับการตอบชอกระง (Item response) จนในปัจจุบันมีผู้นำเอาทฤษฎีใหม่นี้ไปใช้อย่างแพร่หลาย เพราะมีความเชื่อถือเกี่ยวกับแฮมเบิลตัน (Hambleton 1979 : 16-32) ว่าถ้านำทฤษฎีนี้มาใช้ตรงตามข้อตกลงเบื้องต้น จะสามารถแก้ไขจุดอ่อนทั้ง 3 ข้อของทฤษฎีคลาสสิกอลโด และช่วยความก้าวหน้าทางคอมพิวเตอร์ซึ่งช่วยเอื้ออำนวยให้การวิเคราะห์ตามทฤษฎีใหม่ง่ายและสะดวกขึ้น

สำหรับในประเทศไทยการวิเคราะห์แบบสอบยังคงใช้วิธีของทฤษฎีคลาสสิกอลโด ส่วน IRT ยังไม่มีผู้ใดศึกษาอย่างจริงจัง และยังไม่มีการประยุกต์ IRT มาใช้วิเคราะห์แบบสอบ ดังนั้นผู้วิจัยจึงสนใจที่จะนำราสซโมเดล ที่เป็นโมเดลที่ง่ายที่สุดในหลายๆ โมเดลของ IRT มาประยุกต์ใช้ในการวิเคราะห์แบบสอบอุปมาอุปไมยฉบับ ก. ของสำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒประสานมิตร ซึ่งเป็นแบบสอบที่วัดความสามารถในด้านเหตุผลโดยการหาความสัมพันธ์ของรูปภาพเพียงด้านเดียว ดังนั้นจึงมีคุณสมบัติความเป็นมิติเดียวเป็นไปตามข้อตกลงเบื้องต้นของ IRT ซึ่งเป็นไปตามที่ ลอร์ด (Lord 1980 : 20) ได้กล่าวไว้ว่าแบบสอบอุปมาอุปไมยมีคุณสมบัติความเป็นมิติเดียว ด้วยเหตุนี้จึงกล่าวผู้วิจัยจึงวิเคราะห์แบบสอบอุปมาอุปไมยด้วยวิธีราสซโมเดล โดยเปรียบเทียบผลที่ได้จากการวิเคราะห์กับวิธีคลาสสิกอลโดโมเดล

วัตถุประสงค์ของการวิจัย

1. เพื่อเปรียบเทียบและทดสอบความแตกต่างอย่างมีนัยสำคัญทางสถิติของผลที่ได้จากการวิเคราะห์แบบสอบอุปมาอุปไมย ด้วยวิธีของคลาสสิกอลโมเดล กับวิธีของราสชโมเดลในเรื่องของจำนวนข้อกระทงที่ได้รับการคัดเลือกไว้ ค่าความเที่ยง (Reliability) และค่าความตรงร่วมสมัย (Concurrent validity)
2. เพื่อหาสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ได้จากแบบสอบที่ประกอบด้วยข้อกระทงที่ได้รับการคัดเลือกไว้ด้วยคลาสสิกอลโมเดลกับราสชโมเดล

สมมุติฐานในการวิจัย

จากงานวิจัยของเร็นซ์และริเดินเออร์ (Rentz and Ridenour 1978 cited by Rentz 1979 : 3-5) พบว่าแบบสอบผลสัมฤทธิ์ที่เป็นแบบสอบมาตรฐานเมื่อนำมาวิเคราะห์ใหม่ด้วยราสชโมเดล ส่วนมากจะ fit กับโมเดล 80-97% แต่จากงานวิจัยของไวท์ลีย์และดอว์อิส (Whitely and Dawis 1974 cited by Yen 1979 : 261) พบว่าแบบสอบอุปมาอุปไมยทางภาษา เมื่อวิเคราะห์กับราสชโมเดล มีข้อกระทง 30-40% ที่ไม่ fit กับโมเดล ดังนั้นในการวิจัยครั้งนี้จำนวนข้อกระทงที่ได้รับการคัดเลือกด้วยคลาสสิกอลโมเดลน่าจะมีจำนวนมากกว่าของราสชโมเดล ซึ่งส่งผลทำให้ค่าความเที่ยงของแบบสอบที่ประกอบด้วยข้อกระทงที่ได้รับการคัดเลือกด้วยคลาสสิกอลโมเดลสูงกว่าของราสชโมเดล แต่ในการวิเคราะห์ด้วยโมเดลทั้งสองนี้ใช้แบบสอบฉบับเดียวกันจึงน่าจะมีข้อกระทงที่ได้รับการคัดเลือกจากทั้ง 2 โมเดลซ้ำกันอยู่จำนวนหนึ่งซึ่งทำให้คะแนนที่ได้รับจากทั้งสองโมเดลน่าจะมีความสัมพันธ์กัน และเป็นผลให้ความตรงร่วมสมัยที่ได้จากคลาสสิกอลโมเดลและราสชโมเดลกับเกณฑ์คือคะแนนจากแบบสอบวิทยาศาสตร์น่าจะ ไม่แตกต่างกัน ดังนั้นผู้วิจัยจึงตั้งสมมุติฐานในการวิจัยดังนี้

1. จำนวนข้อกระทงที่ได้รับการคัดเลือกไว้ด้วยคลาสสิกอลโมเดลน่าจะมีจำนวนมากกว่าของราสชโมเดลอย่างมีนัยสำคัญ
2. ค่าความเที่ยงของแบบสอบที่ประกอบด้วยข้อกระทงที่ได้รับการคัดเลือกไว้ด้วยคลาสสิกอลโมเดลน่าจะมีค่าสูงกว่าของราสชโมเดลอย่างมีนัยสำคัญ
3. ค่าความตรงร่วมสมัยของแบบสอบที่ประกอบด้วยข้อกระทงที่ได้รับการคัดเลือกไว้ด้วยคลาสสิกอลโมเดลและด้วยราสชโมเดลกับเกณฑ์คือคะแนนจากแบบสอบ

วิทยาศาสตร์ น่าจะแตกต่างกันอย่างไม่มีนัยสำคัญ

4. คะแนนที่ได้จากข้อกระทงที่ได้รับการคัดเลือกไว้ในแต่ละโมเดล น่าจะมีความสัมพันธ์กันอย่างมีนัยสำคัญ

ขอบเขตของการวิจัย

1. เนื่องจากแบบสอบถามมาอุปไมยฉบับที่ใช้ในการวิจัยครั้งนี้ เป็นแบบสอบถามมาตรฐาน ซึ่งได้มีการวิเคราะห์ด้วยคลาสสิกอลโมเดล โดยใช้กลุ่มตัวอย่างทั่วประเทศ และในการวิจัยครั้งนี้ผู้วิจัยไม่สามารถที่จะทำการวิเคราะห์แบบราสชโมเดล โดยใช้กลุ่มตัวอย่างทั่วประเทศเช่นเดียวกันได้ ดังนั้นผู้วิจัยจึงนำข้อมูลที่ได้จากกลุ่มตัวอย่างมาทำการวิเคราะห์ด้วยโมเดลทั้ง 2 โมเดล แล้วนำผลที่ได้มาเปรียบเทียบกัน ดังนั้นผลการเปรียบเทียบของทั้ง 2 โมเดลนี้จึงเป็นการเปรียบเทียบค่าพารามิเตอร์ที่ได้จากการวิเคราะห์ของผู้วิจัยเองเท่านั้น

2. กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนที่จบการศึกษาระดับประถมศึกษาปีที่ 6 และเข้ารับการคัดเลือกเข้าเรียนต่อในชั้นมัธยมศึกษาปีที่ 1 โรงเรียนบางกะปิ และโรงเรียนสตรีศรีนครบุศุครบำเพ็ญในปีการศึกษา 2526 เท่านั้น

3. การวิจัยครั้งนี้ไม่คำนึงถึงผลความแตกต่างด้านเพศอายุของกลุ่มตัวอย่าง

ข้อตกลงเบื้องต้น

1. คะแนนที่ได้จากแบบสอบถามมาอุปไมยไม่มีความเชื่อถือได้
2. แบบสอบถามมาอุปไมย วัดความสามารถทางด้านเหตุผลด้วยการหาความสัมพันธ์ของรูปภาพเพียงอย่างเดียว
3. ข้อกระทงแต่ละข้อเป็นอิสระจากกัน โอกาสที่ผู้สอบตอบคำถามข้อกระทงหนึ่งถูก ไม่มีผลต่อการตอบคำถามข้อกระทงอื่น ๆ
4. ข้อกระทงแต่ละข้อมีค่าอำนาจจำแนกใกล้เคียงกัน และค่าการเดาเป็นศูนย์ตามข้อตกลงเบื้องต้นของราสชโมเดล

ความจำกัดของการวิจัย

เนื่องจากมีโน้ตค้นและวิธีการประมาณค่าความเที่ยงของคลาสสิกอลโมเดลกับ

ราชโมเดลต่างกัน ของราชโมเดลใช้ Test Information Function
 ประมาณค่าความเที่ยง (Lord 1980 : 22) ดังนั้นผู้วิจัยจึงไม่สามารถเปรียบเทียบค่า
 ความเที่ยงของข้อกระทงที่ได้รับการคัดเลือกไว้จากทั้ง 2 โมเดลโดยตรงได้ และในการ
 ศึกษา IRT ในระยะต้น ๆ ผู้วิจัยทั้งหลายยังคงใช้วิธีประมาณค่า internal
 consistency reliability ด้วยวิธีการเดียวกับคลาสสิคัลโมเดล กล่าวคือประมาณ
 ค่าด้วยสูตร คูเจอร์ ริชาร์ดสัน กรอนบาค แอลฟา การวิเคราะห์แบบชอยท์ (Hashway
 1978 : 25) ดังนั้นผู้วิจัยจึงเลือกประมาณค่าด้วยสูตรคูเจอร์ ริชาร์ดสัน สูตรที่ 20
 โดยนำข้อกระทงที่ได้รับการคัดเลือกไว้จากแต่ละโมเดล ไปประมาณค่าความเที่ยงโดยใช้
 สูตรคูเจอร์ ริชาร์ดสัน สูตรที่ 20 ดังนั้นการเปรียบเทียบค่าความเที่ยงจากผลการวิ-
 เคราะห์ทั้ง 2 โมเดล จึงเป็นเพียงการเปรียบเทียบค่าความเที่ยงที่ประมาณค่าด้วย
 สูตรคูเจอร์ ริชาร์ดสัน สูตรที่ 20 เท่านั้น

คำจำกัดความที่ใช้ในการวิจัย

แบบสอบอุปมาอุปไมย หมายถึงแบบสอบอุปมาอุปไมย ฉบับบี ก ของสำนักทดสอบ
 ทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

แบบสอบวิทยาศาสตร์ หมายถึงแบบสอบวัดผลสัมฤทธิ์วิชาวิทยาศาสตร์ ฉบับบี ก
 ของสำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

นักเรียน หมายถึงนักเรียนที่เข้ารับการคัดเลือกเข้าเรียนต่อในระดับมัธยม-
 ศึกษาปีที่ 1 โรงเรียนบางกะปิ และโรงเรียนสตรีศรีนครินทร์บ้ำเพ็ญ

คลาสสิคัลโมเดล หมายถึงการวิเคราะห์แบบสอบที่มีโมโนทัศน์ดังนี้

ก. ข้อกระทงที่มีคุณภาพและได้รับการคัดเลือกไว้ต้องมีค่าความยาก
 ระหว่าง .2 ถึง .8 และมีค่าอำนาจจำแนกตั้งแต่ .2 ขึ้นไป

ข. ความเที่ยง หมายถึงความสม่ำเสมอของการตอบของผู้สอบต่อข้อ
 กระทงทุกข้อในแบบสอบ โดยใช้ในการสอบเพียงครั้งเดียว และใช้แบบสอบเพียงชุดเดียว
 ประมาณค่าด้วยสูตรคูเจอร์ ริชาร์ดสัน สูตรที่ 20

ค. ความตรงร่วมสมัย หมายถึงค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง
 คะแนนจากแบบสอบอุปมาอุปไมยที่ประกอบด้วยข้อกระทงที่ได้จากการคัดเลือกไว้ด้วย
 คลาสสิคัลโมเดล ราชโมเดล กับคะแนนจากแบบสอบวิทยาศาสตร์

ราสซิมอเดล หมายถึงการวิเคราะห์แบบส้อมที่มีโมโนโทนิคดังนี้

- ก. ข้อกระทงที่เลือกและได้รับการคัดเลือกไว้ จะต้องเหมาะสม (fit) กับโมเดล กล่าวคือเส้นถดถอยรายข้อ (Item test regression) ปรับเข้าสู่รูปของเส้นโค้งลักษณะรายข้อ (Item characteristic curve : ICC) โดยการทดสอบค่าที แล้วมีความแตกต่างอย่างไม่มีนัยสำคัญที่ระดับ .01
- ข. เส้นโค้งลักษณะรายข้อ หมายถึงเส้นกราฟ (curve) ที่เกิดจากการพลอต (plot) จุดค่าความสามารถที่แท้จริงกับโอกาสที่จะตอบรายข้อกระทงนั้น ๆ ได้ถูกต้อง
- ค. ความสามารถที่แท้จริง เป็นคะแนนรวมที่ได้จากข้อกระทงที่คัดเลือกไว้ควยราสซิมอเดล ซึ่งมีเกณฑ์ว่าทำถูกต้อง 1 คะแนน ทำผิดได้ 0 คะแนน
- ง. ความแกร่ง หมายถึงคุณสมบัติของการวิเคราะห์ด้วยราสซิมอเดลที่ไม่แสดงถึงความไว (Sensitive) ต่อการฝ่าฝืนข้อตกลงเบื้องต้นของราสซิมอเดล

ประโยชน์ที่ได้รับจากการวิจัย

เนื่องจาก IRT ยังเป็นเรื่องใหม่ซึ่งยังไม่เป็นที่รู้จักแพร่หลายในประเทศไทย ผู้วิจัยเชื่อว่าผลการวิจัยครั้งนี้จะช่วยให้ IRT เป็นที่รู้จักเป็นที่เข้าใจกันมากขึ้น นอกจากนี้ยังจะเป็นแนวทางให้ผู้ที่สนใจจะศึกษาค้นคว้าเกี่ยวกับ IRT ต่อไป ซึ่งจะเป็นการพัฒนาทฤษฎีวัดผลในประเทศไทยให้ก้าวหน้าขึ้นอันได้แก่ความรู้ด้านการวิเคราะห์แบบส้อมแนวการประยุกต์ใช้ทฤษฎีการวัดผลเพื่อให้มีประสิทธิภาพมากที่สุด

ศูนย์วิจัยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย