บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การวัดทางจิตวิทยาหรือทางพฤติกรรมศาสตร์ไม่สามารถวัดคุณลักษณะหรือความสามารถได้โดยตรงเหมือนการวัดทางกายภาพ ทั้งนี้เพราะลักษณะพฤติกรรมของบุคคล มีความซับซ้อนและเป็นพฤติกรรมที่แอบแฝงที่มีความแปรเปลี่ยนไปตามสถานการณ์ที่แวดล้อม ทำให้ไม่มีโอกาสวัดในสถานการณ์เดิมได้ (Lord and Novick , 1974 : 13) ค่าที่ได้จากการวัดจึงมีโอกาสแตกต่างไปจากความจริงของสิ่งที่จะวัด ดังนั้นการวัดทางด้านจิตวิทยาหรือทางพฤติกรรมศาสตร์จึงต้องอาศัยการวัดทางอ้อม โดยการกำหนดสิ่งเร้า (Stimuli) ขึ้นมาเร้าคุณลักษณะของสิ่งที่ต้องการจะวัดแล้วให้สิ่งที่ถูกวัดตอบสนอง (Response) ออกมา ผลจากการตอบสนองจะเป็นตัวแทนของคุณลักษณะของสิ่งที่ถูกวัดนั้น

ในการวัดคุณลักษณะหรือความสามารถนั้น นักจิตวิทยาและนักวัดผลได้พยายามสร้างทฤษฎีต่าง ๆ ขึ้นมาเพื่อให้ผลการวัดเป็นตัวแทนของคุณลักษณะ หรือความสามารถของบุคคลอย่างแท้จริง ทฤษฎีการวัดที่นับว่าสำคัญและเป็นทฤษฎีที่กำลังเป็นที่สนใจในการศึกษาปัจจุบันมีสองรูปแบบ คือ ทฤษฎีการวัดแบบดั้งเดิม (Classical test theory) และทฤษฎีคุณลักษณะแฝง (Latent trait theory) ซึ่งภายหลังนิยมเรียกว่า ทฤษฎีการตอบสนองข้อสอบ (Item response theoty)

ทฤษฎีการวัดแบบดั้งเดิม นับเป็นทฤษฎีการวัดที่สำคัญที่ใช้กันอย่างแพร่หลาย ทั้งในอดีตและปัจจุบัน และเป็นทฤษฎีที่มีลักษณะง่ายต่อการใช้ประโยชน์ แต่ยังมีจุดอ่อนในการวัดความสามารถที่แท้จริง แนวคิดของทฤษฎีการวัดแบบดั้งเดิมนี้อาศัยแบบจำลองเชิงบวก (Additive model) โดยอาศัยความสัมพันธ์ของคะแนนที่สังเกตมาได้ (Observed score : X_i) กับผลบวกของคะแนนจริง (True score : T_i) กับคะแนนความคลาดเคลื่อน (Error score : E_i) แสดงได้โดยสมการ $X_i = T_i + E_i$ และมีข้อตกลงเบื้องต้นว่าค่าความคลาดเคลื่อน (E) นั้นมีลักษณะเกิดอย่างสุ่ม (Randomness) ซึ่งทำให้ค่าความคลาดเคลื่อนมีลักษณะของการแจกแจงเป็นแบบโค้งปกติที่มีค่าเฉลี่ยเป็นศูนย์และค่าความแปรปรวนเป็น 6^2_e และในการประมาณค่าความสามารถจริงนั้นอาศัยคะแนนของกลุ่มผู้สอบเป็นสำคัญ โดยให้นิยามว่า ค่าเฉลี่ยของความสามารถจริงจะเท่ากับค่าเฉลี่ยของคะแนนที่สังเกตได้ของกลุ่ม ซึ่งข้อตกลงดังกล่าวนี้ แอเลนและเยน (Allen and Yen , 1979 : 56-65) กล่าวว่าเป็นทฤษฎีที่มีข้อตกลงคะแนนจริงที่อ่อน (Weak true-score) และเป็นทฤษฎีที่ไม่เพียงพอต่อการประมาณค่าความสามารถจริง

ส่วนทฤษฎีการวัดแบบคุณลักษณะแฝง หรือทฤษฎีการตอบสนองข้อสอบ (Item response theory) เป็นทฤษฎีที่อธิบายลักษณะข้อสอบแต่ละข้อด้วยพารามิเตอร์ของข้อสอบและอธิบายความสามารถของผู้สอบด้วยพารามิเตอร์ของผู้สอบ โดยค่าพารามิเตอร์เหล่านั้นพิจารณาจากความสัมพันธ์ระหว่างความสามารถจริงกับลักษณะการตอบสนองข้อสอบ ซึ่งลักษณะรูปแบบของทฤษฎีแบบคุณลักษณะแฝงนี้ ได้มีการพัฒนาขึ้นหลายแบบ โดยแต่ละแบบมีความแตกต่างกันไปตามฟังก์ชั่นคณิตศาสตร์และจำนวนพารามิเตอร์ที่ใช้อธิบาย คือแบบ 1 พารามิเตอร์เป็นรูปแบบที่มีค่าพารามิเตอร์แสดงลักษณะความยากของข้อสอบ แบบ 2 พารามิเตอร์เป็นรูปแบบที่มีค่าพารามิเตอร์แสดงลักษณะจำแนกของข้อสอบและลักษณะความยากของข้อสอบ และแบบ 3 พารามิเตอร์เป็นรูปแบบที่มีค่าพารามิเตอร์แสดงลักษณะจำแนกของข้อสอบ ลักษณะความยากของข้อสอบและลักษณะของการเดาของข้อสอบ ซึ่งวิธีการในการประมาณค่าพารา

มิเตอร์และค่าความสามารถจริงจะให้ค่าที่คงที่ ไม่ขึ้นอยู่กับลักษณะความสามารถของกลุ่มผู้สอบ นอกจากนั้นการประมาณค่าแบบคุณลักษณะแฝงที่ใช้วิธีการแบบ maximum likelihood นั้น แฮมเบิลตันและสวามินาธาน (Hambleton and Swaminathan , 1985 : 88) กล่าวว่าลักษณะการแจกแจงของค่าที่ประมาณได้จะเป็นแบบปกติ

ความแตกต่างของทฤษฎีทั้งสอง สรุปได้ดังนี้

1) จุดมุ่งหมาย

แบบดั้งเดิม มุ่งหาคะแนนจริงของผู้สอบ

แบบคุณลักษณะแฝง มุ่งหาค่าที่แท้จริงที่บอกลักษณะจริงของข้อสอบและความสามารถของผู้สอบ

2) ฟังก์ชั่นในการประมาณค่า

แบบดั้งเดิม อาศัยแบบจำลองเชิงบวก (Additive model)

$$X_i = T_i + E_i$$

โดยที่ X_i หมายถึงคะแนนของบุคคลที่ i

T_i หมายถึงคะแนนจริงของบุคคลที่ i

E_i หมายถึงคะแนนความคลาดเคลื่อนของบุคคลที่ i

แบบคุณลักษณะแฝง อาศัยคุณสมบัติความน่าจะเป็นที่แสดงความสัมพันธ์ระหว่างความสามารถจริงกับการตอบสนองข้อสอบแสดงได้โดยฟังก์ชั่น $P(\theta)$ ของแต่ละพารามิเตอร์

ก. แบบ 1 พารามิเตอร์

$$P_g(\theta) = \frac{EXP(\theta - b_g)}{1 + EXP(\theta - b_g)}$$

ข. แบบ 2 พารามิเตอร์

$$P_g(\theta) = \frac{EXP(Da_g(\theta - b_g))}{1 + EXP(Da_g(\theta - b_g))}$$

ค. แบบ 3 พารามิเตอร์

$$P_g(\theta) = c_g + (1-c_g)\frac{EXP(Da_g(\theta - b_g))}{1 + EXP(Da_g(\theta - b_g))}$$

เมื่อ $P_g(\theta)$ หมายถึงโอกาสที่ผู้สอบที่มีความสามารถ θ จะทำข้อสอบข้อ g ได้ถูกต้อง

θ หมายถึงระดับความสามารถที่แท้จริงที่คำนวณจากคะแนนจริงแต่ปรับหน่วยเป็นคะแนนมาตรฐาน

D หมายถึงค่าคงที่ มีค่าเท่ากับ 1.7

a_g หมายถึงค่าอำนาจจำแนก ซึ่งเป็นสัดส่วนโดยตรงกับค่าความชัน (slope) ของโค้ง ณ จุดเปลี่ยนโค้ง โดยความชันมีค่า = $a/2$

b_g หมายถึงค่าความยากที่แสดงระดับความสามารถ(θ)ที่จุดโค้งความชันสูงสุดหรือกรณีที่ไม่มีการเดา โดยค่า b_g อาจเป็นบวกหรือลบ ถ้าเป็นลบแสดงว่าข้อสอบง่าย และถ้าเป็นบวกแสดงว่าข้อสอบยาก

c_g หมายถึงโอกาสการเดา ซึ่งเป็นค่าโอกาสที่ผู้สอบที่มีความสามารถต่ำจะทำข้อสอบถูก

3) ข้อตกลงเบื้องต้น

<u>แบบดั้งเดิม</u> (Allen and Yen , 1979 : 57)

1. คาที่สังเกตได้เป็นผลบวกเชิงเส้นตรงของคาคะแนนจริงกับคาความคลาดเคลื่อน ($X_i = T_i + E_i$)
2. คาที่คาดหวังของคะแนนที่สังเกตมาได้ (Expected score) จะเท่ากับคาคะแนนจริง ($E(X) = T$)
3. ความสัมพันธ์ระหว่างคาความคลาดเคลื่อนกับคาคะแนนจริงเป็นศูนย์ ($\rho_{ET} = 0$)
4. ความสัมพันธ์ระหว่างคาความคลาดเคลื่อนของแบบสอบที่คู่ขนานกันมีค่าเป็นศูนย์ ($\rho_{E_1E_2} = 0$)
5. ความคลาดเคลื่อนของแบบสอบชุดหนึ่ง ไม่สัมพันธ์กับคาคะแนนจริงของแบบสอบอีกชุดหนึ่ง ($\rho_{E_1T_2} = 0$)

<u>แบบคุณลักษณะแฝง</u> (Hambleton and Swaminathan , 1985 : 15)

1. ข้อสอบแต่ละข้อมีลักษณะเป็นมิติเดียว (Unidimension)
2. การตอบข้อสอบในแต่ละข้อเป็นอิสระกัน (Local independence)
3. พฤติกรรมการตอบสนองข้อสอบและระดับความสามารถของผู้สอบแสดงความสัมพันธ์ได้ด้วยโค้งลักษณะข้อสอบ (Item characteristic curve)
4. การสอบไม่อยู่ภายใต้เงื่อนไขของ speeded test

4) คาที่ได้รับ

<u>แบบดั้งเดิม</u> แปรเปลี่ยนตามลักษณะความสามารถของกลุ่มผู้สอบ

<u>แบบคุณลักษณะแฝง</u> ไม่แปรเปลี่ยน (Invariant)

จากการทบทวนงานวิจัยที่เกี่ยวข้องกับการวิจัยนี้ สรุปได้ว่า

1) ในการประมาณค่าความสามารถจริงที่ได้จากทฤษฎีทั้งสองนี้ ยังไม่สามารถบอกความแตกต่างของผลการประมาณค่าอย่างเห็นได้ชัด นอกจากเสนอแนะให้เห็นหลักการและแนวทางของทฤษฎีนั้น ๆ เพื่อประโยชน์ในการเลือกใช้ นอกจากนั้น ผลงานการศึกษาส่วนใหญ่เป็นการเปรียบเทียบในบางคุณลักษณะของแต่ละวิธีการและผลจากการศึกษาเกี่ยวกับค่าประมาณความสามารถจริงก็ได้ผลทั้งที่สอดคล้องกันและขัดแย้งกัน ดังเช่น ลอร์ดและวินเจอร์สกี้ (Lord and Wingersky , 1981 : 453) ได้ศึกษาเปรียบเทียบคะแนนจริงของทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์กับคะแนนจริงที่ได้จากการประมาณโดยวิธีอีควิเปอร์เซนไทล์ ปรากฏว่าผลการศึกษาเกี่ยวกับค่าคะแนนจริงของทั้งสองวิธีให้ผลที่ใกล้เคียงกัน

เบเกอร์ (Baker , 1987 : 134) กล่าวว่า จากเอกสารที่เสนอเกี่ยวกับคุณสมบัติการประมาณค่าของแบบคุณลักษณะแฝงนั้น แบบ 1 และ 2 พารามิเตอร์นั้นดีกว่าแบบ 3 พารามิเตอร์ เพราะจากหลักฐานทางทฤษฎีและข้อมูลเชิงประจักษ์แสดงให้เห็นว่า ค่าความคลาดเคลื่อนของการประมาณค่าแบบ 1 พารามิเตอร์และ 2 พารามิเตอร์นั้นมีค่าน้อยกว่าแบบ 3 พารามิเตอร์ นอกจากนั้น โคเลนและวิทนีย์ (Kolen and Whitney , 1982 : 279) ได้ศึกษาการประมาณค่าคะแนนแบบเชิงเส้นตรง เปอร์เซนไทล์ ทฤษฎีการตอบสนองข้อสอบ 1 พารามิเตอร์และ 3 พารามิเตอร์ โดยใช้แบบสอบ 12 ฉบับ พบว่า การประมาณค่าคะแนนแบบเปอร์เซนไทล์และแบบ 3 พารามิเตอร์ได้ผลยังไม่คงที่ ส่วนแบบเชิงเส้นตรงและแบบ 1 พารามิเตอร์นั้นได้ผลที่คงที่ และจากการศึกษาสรุปว่าความถูกต้องของค่าการประมาณขึ้นอยู่กับลักษณะของแบบสอบ วิธีการประมาณค่าและขนาดกลุ่มตัวอย่าง

2) ในการหาค่าความสัมพันธ์ของค่าประมาณความสามารถจริงที่ได้จากทฤษฎีทั้งสองนี้ ได้มีผู้ศึกษาโดยการหาค่าความสัมพันธ์ของค่าความสามารถจริงกับค่าพารามิเตอร์ที่ประมาณได้และค่าความสามารถจริงกับค่าคะแนนที่ประมาณจากวิธีการต่างๆ ดังเช่น

เบเกอร์ (Baker , 1987 : 118 - 119) ได้สรุปประเด็นและแนวคิดจากการศึกษาของลอร์ด (Lord , 1975) และ ฮูลิน (Hulin , 1982) ว่าค่าประมาณพารามิเตอร์ a , b และ c กับค่าความสามารถจริง (Θ) มีความสัมพันธ์กัน โดย ลอร์ด ได้ simulated ข้อมูลจากแบบสอบ 90 ข้อ จำนวนตัวอย่าง 2,995 คน พบว่า ค่าอำนาจจำแนกที่ประมาณได้ (a') กับค่าพารามิเตอร์ a มีความสัมพันธ์กัน ($r_{a'a} = 0.920$) และค่าความยากที่ประมาณได้ (b') กับค่าพารามิเตอร์ b มีความสัมพันธ์กัน ($r_{b'b} = 0.988$) ส่วน ฮูลิน ได้ศึกษาความสัมพันธ์ของค่าพารามิเตอร์แบบ 2 และ 3 พารามิเตอร์ พบว่า ความสัมพันธ์ของค่าพารามิเตอร์ที่ประมาณได้กับค่าพารามิเตอร์ ($r_{a'a}$ และ $r_{b'b}$) ของแบบ 2 พารามิเตอร์สูงกว่าแบบ 3 พารามิเตอร์ และจากการศึกษาของ ชูศักดิ์ ขัมภลิขิต (ชูศักดิ์ ขัมภลิขิต , 2525 : 33) ในการเปรียบเทียบการให้คะแนนแบบให้น้ำหนักของลอร์ด (W) แบบคะแนนดิบ (X) แบบคะแนนมาตรฐาน (Z) และคะแนนความสามารถตามแบบทฤษฎีการตอบสนองข้อสอบ (Θ) ปรากฏว่าการให้คะแนนแบบคะแนนดิบมีความสัมพันธ์กับการให้คะแนนแบบอื่น ๆ อยู่ในเกณฑ์สูง และการให้คะแนนแบบให้น้ำหนัก แบบคะแนนมาตรฐาน และคะแนนความสามารถมีความสัมพันธ์กันสูงด้วย นอกจากนั้นจากการศึกษาของ สุนันท์ ศลโกสุม (สุนันท์ ศลโกสุม , 2530 . 118 - 130) ในการหาความสัมพันธ์ของวิธีวิเคราะห์ผลการสอบด้วยทฤษฎีแบบดั้งเดิม ดัชนีชี้นำของซาโต้ ทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และดัชนีชี้นำของทาซูโอกะ พบว่า คะแนนสอบ (X) ที่ได้จากทฤษฎีแบบดั้งเดิมมีความสัมพันธ์กับคะแนนความสามารถ (Θ) ที่ได้จากทฤษฎีการตอบสนองข้อสอบอยู่ในเกณฑ์สูง (0.829) ส่วนการวิเคราะห์ผลการสอบเพื่อบอกความสามารถจริงของผู้สอบนั้น แบบดั้งเดิมให้ผลใกล้เคียงกับวิธีอื่น ๆ

ในการศึกษาการประมาณค่าความสามารถจริงจากแบบสอบที่มีความยากง่ายต่างกันนั้น โคเลน (Kolen , 1981 : 3) ได้สรุปผลงานวิจัยต่าง ๆ จากการศึกษาการประมาณค่าคะแนนจากแบบสอบที่มีความยากง่ายต่างกันนั้น วิธีแบบเชิงเส้นตรง อีควิเปอร์เซนไทล์และแบบของราส์ชนั้นเป็นวิธีที่ยังไม่ดีพอ แต่วิธี

แบบ 3 พารามิเตอร์เหมาะกับแบบสอบที่มีความยากต่างกัน และจากการศึกษาวิธีการประมาณค่าแบบประเพณีนิยมหรือแบบดั้งเดิมกับทฤษฎีตอบสนองข้อสอบหรือแบบคุณลักษณะแฝง จากแบบสอบที่มีความยากเท่าเทียมกันและความยากต่างกันนั้น พบว่า วิธีการแบบเปอร์เซนไทล์เหมาะกับแบบสอบที่มีความยากต่างกัน แบบ 1 พารามิเตอร์ไม่เหมาะกับแบบสอบที่มีความยากต่างกัน และแบบ 3 พารามิเตอร์ยังมีปัญหาในการประมาณค่าเกี่ยวกับ lower asymptote parameter นอกจากนั้นลอร์ด (Lord , 1983) ยังได้ศึกษาเกี่ยวกับความคงที่ (Consistency) และความลำเอียง (Bias) ของการประมาณค่าพารามิเตอร์แบบ 3 พารามิเตอร์จากแบบสอบที่มีจำนวนข้อต่างกัน คือ 10 ข้อ 15 ข้อ และ 20 ข้อ และกลุ่มตัวอย่างที่มีขนาดต่างกัน คือ 50 คน 200 คน และ 1,000 คน ผลการศึกษาพบว่า จำนวนข้อสอบและจำนวนตัวอย่างเพิ่มขึ้นการประมาณค่าความยากและค่าอำนาจจำแนกของแบบสอบมีค่าคงที่ ส่วนความลำเอียงนั้นพบว่า ถ้าค่าอำนาจจำแนกน้อยการประมาณค่าจะได้ค่าที่มากกว่าปกติ (Overestimate) แต่ถ้าค่าอำนาจจำแนกสูง จะได้ค่าที่คงที่ และในการศึกษาค่าความลำเอียงของทั้งหมดของการประมาณค่าอำนาจจำแนกมีค่าต่ำ แต่ในกรณีความยากของข้อสอบนั้นพบว่าถ้าค่าความยากเป็นลบจะได้ค่าที่ต่ำกว่าปกติ (Underestimate) แต่ถ้าค่าความยากเป็นบวกแล้วค่าที่ได้จะใกล้เคียงกัน นอกจากนั้นในการประมาณค่าอำนาจจำแนก (a) ค่าความยาก (b) ค่าการเดา (c) และค่าความสามารถ (θ) นั้นพบว่า ในกรณีข้อสอบง่ายและค่าความยากง่ายปานกลางนั้นความลำเอียงจะเป็นลบ และจะเป็นบวกเมื่อข้อสอบมีความยาก ส่วนความลำเอียงในการประมาณค่าอำนาจจำแนกข้อสอบนั้นมีค่าเป็นบวก และการประมาณค่าการเดาความลำเอียงมีค่าเป็นลบ แต่โดยทั่วไปการประมาณค่าพารามิเตอร์นั้นค่าความคลาดเคลื่อนมาตรฐาน (Standard error) จะมีค่ามากและความลำเอียงจะมีค่ามากด้วย

ส่วนในการศึกษาผลกระทบของค่าการประมาณอันเนื่องมาจากผู้สอบมีระดับความสามารถต่างกันนั้น ถ้าพิจารณาจากความเชื่อของทฤษฎีแบบดั้งเดิมและแบบคุณลักษณะแฝงที่ว่า ค่าพารามิเตอร์ที่ประมาณตามแบบคุณลักษณะแฝงนั้นไม่ขึ้นอยู่กับกลุ่มความสามารถของผู้สอบ แต่แบบดั้งเดิมนั้นแปรเปลี่ยนไปตามกลุ่มความสามารถของผู้สอบ ดังนั้นค่าประมาณความสามารถจริงตามแบบคุณลักษณะแฝง ซึ่งอาศัยค่าความยาก ค่าอำนาจจำแนก และค่าการเดา ยอมจะให้ค่าคะแนนจริงที่แตกต่างกัน และน่าจะได้ค่าที่แตกต่างจากแบบดั้งเดิมด้วย

จากหลักการของทฤษฎีแบบดั้งเดิมและทฤษฎีแบบคุณลักษณะแฝง ประกอบกับการศึกษาจากข้อมูลเชิงประจักษ์ที่กล่าวมา สรุปได้ว่าทฤษฎีทั้งสองต่างมุ่งหาค่าความสามารถจริง (True ability value) หรือคะแนนจริง (True score) ของการสอบวัดและต่างพยายามอ้างอิงความสามารถจริงไปสู่ตัวผู้สอบและลักษณะของข้อสอบเหมือนกัน จะต่างกันก็แต่ความเชื่อ รูปแบบ และวิธีการในการประมาณค่า โดยวิธีการประมาณค่าแบบดั้งเดิมมีลักษณะง่าย เป็นรูปแบบที่สะดวกต่อการใช้และได้รับการพัฒนามาโดยตลอด ส่วนวิธีการประมาณค่าแบบคุณลักษณะแฝงนั้นกำลังพัฒนาเพื่อแก้ไขจุดบกพร่องบางประการของทฤษฎีแบบดั้งเดิมและกำลังเป็นที่สนใจ แต่การนำไปใช้ยังไม่กว้างขวางเหมือนทฤษฎีแบบดั้งเดิม ประกอบกับรูปแบบของทฤษฎีแบบคุณลักษณะแฝงนั้นมีทั้งแบบ 1 พารามิเตอร์ 2 พารามิเตอร์ และ 3 พารามิเตอร์ โดยแต่ละแบบการประมาณค่าต่างมีฟังก์ชั่นในการประมาณโอกาสในการตอบข้อสอบถูกเป็นของแต่ละแบบ และเมื่อพิจารณาจากงานการวิจัยที่เกี่ยวข้องกับทฤษฎีแบบดั้งเดิมกับแบบคุณลักษณะแฝงประกอบแล้ว ผลการศึกษามีทั้งที่สอดคล้องกันและขัดแย้งกัน และการศึกษาส่วนใหญ่มุ่งในการเปรียบเทียบคะแนนจริงในเชิงของความสัมพันธ์เส้นตรง ซึ่งเป็นคุณสมบัติเฉพาะค่าความสามารถจริง โดยมิได้เปรียบเทียบความแตกต่างในด้านอื่น ๆ โดยเฉพาะลักษณะการแจกแจงว่าวิธีการแบบใหม่มีลักษณะที่แตกต่างไปจากแบบดั้งเดิมหรือไม่ ทั้งนี้โดยพิจารณาจากการนำผลการประมาณค่าไปใช้ในงาน

วิจัยต่าง ๆ ที่ผ่านมานั้นยังคงใช้วิธีแบบดั้งเดิม ซึ่งเป็นการยอมรับข้อตกลงเบื้องต้นเกี่ยวกับลักษณะการแจกแจงของแบบดั้งเดิม แต่ถ้าจะนำทฤษฎีแบบคุณลักษณะแฝงมาใช้อธิบายความสามารถของผู้สอบแล้วควรได้ศึกษาเกี่ยวกับลักษณะการแจกแจงค่าความสามารถจริงที่ประมาณได้ ทั้งนี้เพราะวิธีการประมาณค่าความสามารถจริงของแบบดั้งเดิมนั้นแตกต่างจากแบบคุณลักษณะแฝง โดยแบบดั้งเดิมอาศัยการประมาณจากคะแนนสอบ (E(X)) ส่วนแบบคุณลักษณะแฝงอาศัยการประมาณค่าจากผลรวมของโอกาสในการตอบข้อสอบถูก (P(θ)) ซึ่งถ้านำทฤษฎีแบบคุณลักษณะแฝงมาใช้โดยอาศัยข้อตกลงเกี่ยวกับลักษณะการแจกแจงแบบทฤษฎีดั้งเดิมแล้ว อาจได้ผลสรุปที่คลาดเคลื่อนไปถ้าทฤษฎีแบบคุณลักษณะแฝงนั้นมีรูปลักษณะการแจกแจงที่ต่างไปจากทฤษฎีแบบดั้งเดิม ฉะนั้นการศึกษารูปแบบ ลักษณะการแจกแจง และค่าที่อธิบายลักษณะการแจกแจงของทฤษฎีแบบคุณลักษณะแฝงเทียบกับทฤษฎีแบบดั้งเดิม จึงเป็นสิ่งจำเป็นต่อการนำผลของทฤษฎีแบบคุณลักษณะแฝงไปใช้ ตลอดจนการพัฒนารูปแบบการนำไปใช้ที่เหมาะสมกับทฤษฎีแบบคุณลักษณะแฝงต่อไป

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาค่าที่อธิบายลักษณะการแจกแจงความถี่ของความสามารถจริงที่ประมาณตามแบบดั้งเดิมและแบบคุณลักษณะแฝง โดยใช้คะแนนสอบทั้งหมด คะแนนสอบที่ได้จากกลุ่มข้อสอบง่าย กลุ่มข้อสอบยาก กลุ่มผู้สอบที่มีความสามารถต่ำ และกลุ่มผู้สอบที่มีความสามารถสูง

2. เพื่อเปรียบเทียบความแตกต่างของค่าเฉลี่ยและค่าความแปรปรวนของความสามารถจริงที่ประมาณตามแบบดั้งเดิมกับแบบคุณลักษณะแฝง โดยใช้คะแนนสอบทั้งหมด คะแนนสอบที่ได้จากกลุ่มข้อสอบง่าย กลุ่มข้อสอบยาก กลุ่มผู้สอบที่มีความสามารถต่ำ และกลุ่มผู้สอบที่มีความสามารถสูง

3. เพื่อหาค่าความสัมพันธ์ของความสามารถจริงที่ประมาณตามแบบดั้งเดิมกับแบบคุณลักษณะแฝง โดยใช้คะแนนสอบทั้งหมด คะแนนสอบที่ได้จากกลุ่มข้อสอบง่าย กลุ่มข้อสอบยาก กลุ่มผู้สอบที่มีความสามารถต่ำ และกลุ่มผู้สอบที่มีความสามารถสูง

สมมุติฐานของการวิจัย

1. การศึกษาค่าที่อธิบายลักษณะการแจกแจงความถี่ของค่าความสามารถจริง ที่ประมาณตามแบบดั้งเดิมและแบบคุณลักษณะแฝง 1 พารามิเตอร์ 2 พารามิเตอร์ และ 3 พารามิเตอร์ ในแต่ละลักษณะกลุ่มของแต่ละแบบการประมาณค่า

เนื่องจากค่าความสามารถจริงที่นำมาศึกษาเป็นค่าที่ประมาณได้จากกลุ่มเฉพาะที่มีลักษณะกลุ่มไม่เหมือนกัน โดยต่างกันตามกลุ่มความสามารถของผู้สอบและตามระดับความยากง่ายของข้อสอบ ซึ่งทำให้การแจกแจงคะแนนสอบมีลักษณะโค้งที่แตกต่างกัน เพราะการได้คะแนนมากหรือน้อยนั้นขึ้นอยู่กับความยากและง่ายของข้อสอบ และความสามารถสูงและต่ำของผู้สอบ โดยกลุ่มความสามารถสูงและกลุ่มข้อสอบง่ายนั้นค่าคะแนนสูงจะมีมากกว่าค่าคำแนนต่ำ ส่วนกลุ่มข้อสอบยากและกลุ่มความสามารถต่ำนั้น ค่าคะแนนต่ำจะมีมากกว่าคะแนนสูง นอกจากนั้นการประมาณค่าความสามารถจริงใช้วิธีการและสมการในการคิดคำนวณค่าที่แตกต่างกัน โดยแบบดั้งเดิมนั้นประมาณค่าคะแนนจริงจากค่าคะแนนสอบ (E (X)) ส่วนแบบคุณลักษณะแฝงประมาณค่าคะแนนจริงจากผลรวมของโอกาสในการตอบข้อสอบถูก (P (θ)) และใช้จำนวนพารามิเตอร์ที่ต่างกัน โดยแบบ 1 พารามิเตอร์อาศัยค่าความยากของข้อสอบ (b) แบบ 2 พารามิเตอร์อาศัยค่าอำนาจจำแนก (a) และค่าความยาก (b) ของข้อสอบ ส่วนแบบ 3 พารามิเตอร์นั้นอาศัยค่าอำนาจจำแนก (a) ค่าความยาก (b) และ ค่าโอกาสการเดา (c) ในการประมาณค่า P (O) ประกอบกับการประมาณค่าแบบคุณลักษณะแฝงที่ใช้ในการศึกษานี้อาศัยวิธีการแบบ maximum likelihood ซึ่ง แฮมเบิลตันและสวามินาธาน (Hambleton and Swaminathan , 1985 : 88) กล่าวว่าการประมาณค่าที่ใช้ความน่าจะเป็นที่แสดงโอกาสการเกิดสูงสุดในการประมาณค่านี้ ฐานโค้งของการแจกแจงจะเป็นแบบปกติ (Asymptotically normally distributed) และค่าคะแนนมีการกระจายและปรับเปลี่ยนค่าการแจกแจงที่ละเอียดมากขึ้น

จากเหตุผลดังกล่าว สรุปเป็นสมมุติฐานของค่าอธิบายลักษณะการแจกแจงในแต่ละประเด็นได้ว่า

1.1 ค่าการวัดแนวโน้มเข้าสู่ส่วนกลาง ซึ่งได้แก่ค่าเฉลี่ย ค่ามัธยฐานและค่าฐานนิยม ในแต่ละลักษณะกลุ่มของแต่ละแบบการประมาณค่า น่าจะมีค่าไม่เท่ากัน

1.2 ค่าการวัดการกระจาย ซึ่งได้แก่ค่า พิสัย ส่วนเบี่ยงเบนควอไทล์ ความเบี่ยงเบนมาตรฐาน และสัมประสิทธิ์การกระจายหรืออัตราการกระจาย ตลอดจนค่าความคลาดเคลื่อนมาตรฐาน ในแต่ละลักษณะกลุ่มของแต่ละแบบการประมาณค่าน่าจะมีค่าไม่เท่ากัน

1.3 ค่าความเบ้และค่าความโด่งของการแจกแจงในแต่ละลักษณะกลุ่มของแต่ละแบบการประมาณค่า น่าจะมีค่าไม่เท่ากัน โดยกลุ่มคะแนนสอบทั้งหมด กลุ่มข้อสอบง่าย และกลุ่มความสามารถสูง น่าจะมีค่าความเบ้เป็นลบ ลักษณะเบ้ไปทางซ้าย ส่วนกลุ่มข้อสอบยากและกลุ่มความสามารถต่ำนั้น น่าจะมีค่าความเบ้เป็นบวก ลักษณะเบ้ไปทางขวา

1.4 รูปลักษณะการแจกแจงในแต่ละลักษณะกลุ่มของแต่ละแบบประมาณค่าน่าจะมีลักษณะรูปทรงที่ต่างกัน และต่างจากการแจกแจงแบบปกติเกือบทุกลักษณะกลุ่ม

2. การเปรียบเทียบค่าเฉลี่ยและค่าความแปรปรวนของค่าความสามารถจริง ที่ประมาณตามแบบดั้งเดิมและแบบคุณลักษณะแฝง 1 พารามิเตอร์ 2 พารามิเตอร์ และ 3 พารามิเตอร์ ในแต่ละลักษณะกลุ่มของแต่ละแบบการประมาณค่า

เนื่องจากการประมาณค่าในแต่ละวิธี ต่างมุ่งประมาณค่าความสามารถจริงของผู้สอบ โดยแบบดั้งเดิมประมาณค่าคะแนนจริงจากค่าคาดหวังของคะแนนสอบ ($E(X) = T$) และในการประมาณค่าอาศัยคะแนนของกลุ่มผู้สอบ โดยเชื่อว่าค่าเฉลี่ยของค่าความสามารถจริงเท่ากับค่าเฉลี่ยของคะแนนที่สุ่งเกตได้ ($\mu_T = \mu_X$)

ตามคุณสมบัติของคะแนนจริงที่ กัลลิคเซนได้เสนอไว้ (Gulliksen , 1967) ประกอบกับในการประมาณค่าคะแนนจริงของแบบดั้งเดิมของแต่ละบุคคลใช้ค่าความคลาดเคลื่อนในการวัดที่เหมือนกันและเท่ากัน ส่วนแบบคุณลักษณะแฝงนั้นพยายามลดค่าความคลาดเคลื่อนในการประมาณค่าโดยการนำค่าพารามิเตอร์ซึ่งเป็นค่าที่คงที่เข้ามาคิดคำนวณ ซึ่งน่าจะทำให้ค่าความสามารถจริงที่ประมาณตามแบบคุณลักษณะแฝงมีความคลาดเคลื่อนของค่าเฉลี่ยตัวอย่าง (Sample mean) น้อยกว่าแบบดั้งเดิม นอกจากนั้น รุดเนอร์ (Rudner , 1983 : 952) ได้สรุปว่า ลักษณะโค้งของข้อสอบ (ICC) แบบคุณลักษณะแฝงมีลักษณะไม่แปรผันถ้าข้อสอบได้จัดกระทำกับกลุ่มที่มีระดับความสามารถ (θ) ที่เหมือนกันแล้ว ค่าพารามิเตอร์ a , b และ c จะมีลักษณะที่คงที่ ดังนั้นถ้าค่า a , b และ c มีความคงที่เมื่อมีความสามารถเท่ากันแล้ว การประมาณค่า $P(\theta)$ จากฟังก์ชั่นของแต่ละแบบคุณลักษณะแฝงนั้นจะมีค่าไม่เท่ากัน

จากเหตุผลดังกล่าว สรุปเป็นสมมุติฐานได้ว่า

2.1 การประมาณค่าความสามารถจริงตามแบบดั้งเดิมกับแบบคุณลักษณะแฝง 1 พารามิเตอร์ 2 พารามิเตอร์ และ 3 พารามิเตอร์ น่าจะมีค่าเฉลี่ยของค่าความสามารถจริงต่างกัน ทั้งจากการประมาณค่าโดยใช้คะแนนสอบทั้งหมด คะแนนสอบที่ได้จากกลุ่มข้อสอบง่าย กลุ่มข้อสอบยาก กลุ่มผู้สอบที่มีความสามารถต่ำและกลุ่มผู้สอบที่มีความสามารถสูง

2.2 การประมาณค่าความสามารถจริงตามวิธีแบบดั้งเดิมกับแบบคุณลักษณะแฝง 1 พารามิเตอร์ 2 พารามิเตอร์ และ 3 พารามิเตอร์ น่าจะมีค่าความแปรปรวนของค่าความสามารถจริงที่แตกต่างกัน ทั้งจากการประมาณค่าโดยใช้คะแนนสอบทั้งหมด คะแนนสอบที่ได้จากกลุ่มข้อสอบง่าย กลุ่มข้อสอบยาก กลุ่มผู้สอบที่มีความสามารถต่ำและกลุ่มผู้สอบที่มีความสามารถสูง

3. การหาค่าความสัมพันธ์ของค่าความสามารถจริงที่ประมาณตามแบบดั้งเดิมกับแบบคุณลักษณะแฝง 1 พารามิเตอร์ 2 พารามิเตอร์ และ 3 พารามิเตอร์ ในแต่ละลักษณะกลุ่มของแต่ละแบบการประมาณค่า

เนื่องจากการประมาณค่าแบบดั้งเดิมและแบบคุณลักษณะแฝงต่างมุ่งหาค่าความสามารถจริง ประกอบกับในการศึกษาครั้งนี้ใช้แบบสอบฉบับเดียวกันและกลุ่มตัวอย่างเดียวกัน ดังนั้นการประมาณค่าความสามารถจริงในแต่ละแบบการประมาณค่าของแต่ละลักษณะกลุ่มน่าจะให้ค่ามากและน้อยตามกัน นอกจากนั้นในการศึกษาข้อมูลเชิงประจักษ์เกี่ยวกับค่าความสามารถจริงและค่าพารามิเตอร์ของแบบสอบในบางประเด็นและบางลักษณะมีความสัมพันธ์กัน เช่น การศึกษาของ ชูศักดิ์ ขัมภลิขิต (ชูสักดิ์ ขัมภลิขิต , 2525 : 33) ที่ศึกษาเปรียบเทียบการให้คะแนนแบบต่าง ๆ พบว่า การให้คะแนนแบบคะแนนดิบ (X) มีความสัมพันธ์กับแบบการให้น้ำหนักของลอร์ด (W) แบบคะแนนมาตรฐาน (Z) และแบบทฤษฎีการตอบสนองข้อสอบ (Θ) อยู่ในเกณฑ์สูง

จากเหตุผลของประการที่กล่าวมา สรุปเป็นสมมุติฐานของการวิจัยได้ว่า ค่าความสามารถจริงที่ประมาณตามแบบดั้งเดิมกับแบบคุณลักษณะแฝง น่าจะมีความสัมพันธ์กันในทางบวกและมีนัยสำคัญทางสถิติ ทั้งจากคะแนนสอบทั้งหมด คะแนนสอบที่ได้จากกลุ่มข้อสอบง่าย กลุ่มข้อสอบยาก กลุ่มผู้สอบที่มีความสามารถต่ำ และกลุ่มผู้สอบที่มีความสามารถสูง

ขอบเขตของการวิจัย

1. การศึกษาครั้งนี้ใช้คะแนนสอบคัดเลือกของนักเรียนชั้นมัธยมปีที่ 3 ที่สมัครเข้าศึกษาต่อระดับชั้นมัธยมปีที่ 4 ปีการศึกษา 2531 จากโรงเรียนที่ขอใช้แบบสอบของสำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เมื่อเดือนเมษายน 2531 จำนวน 9 โรงเรียน

2. ค่าความสามารถจริง ใช้วิธีการประมาณค่าจากการคิดคำนวณโดยวิธีการสองแบบคือ แบบดั้งเดิมและแบบคุณลักษณะแฝง

2.1 แบบดั้งเดิม ประมาณค่าโดยสมการเชิงเส้นตรง ซึ่งเป็นสมการที่แอเลนและเยน (Allen and Yen , 1979 : 232) ลอร์ดและโนวิค (Lord and Novick , 1974 : 64) และกัลลิคเซน (Gulliksen , 1967 : 305) ที่เสนอไว้เหมือนกัน คือ

$$Y_i = X_i r_{tt} + (1 - r_{tt}) \bar{X}$$

เมื่อ Y_i หมายถึงค่าคะแนนจริงที่ประมาณได้ของบุคคลที่ i

X_i หมายถึงค่าคะแนนดิบของบุคคลที่ i

r_{tt} หมายถึงค่าความเที่ยงของแบบสอบชุดนั้น

$\bar{X}$ หมายถึงค่าคะแนนเฉลี่ยของคะแนนสอบกลุ่มนั้น

เมื่อประมาณค่าได้แล้วจะปรับเปลี่ยนค่าคะแนนจริงเป็นค่าคะแนนมาตรฐานซี (Z-score) คือ

$$Z_i = \frac{Y_i - \bar{Y}}{S}$$

เมื่อ Z_i หมายถึงค่าคะแนนซีของบุคคลที่ i

Y_i หมายถึงค่าคะแนนจริงที่ประมาณได้ของบุคคลที่ i

$\bar{Y}$ หมายถึงค่าคะแนนเฉลี่ยของค่าคะแนนจริงของกลุ่มนั้น

S หมายถึงค่าเบี่ยงเบนมาตรฐานของค่าคะแนนจริงของกลุ่มนั้น

016653

2.2 แบบคุณลักษณะแฝง ประมาณค่าจาก $\mathcal{G} = \sum P_i(\theta)$ โดยใช้วิธีการสามวิธีคือ แบบ 1 พารามิเตอร์ ใช้รูปแบบของราสช์ คำนวณโดยใช้โปรแกรม BICAL, แบบ 2 พารามิเตอร์ คำนวณโดยใช้โปรแกรม LOGIST และแบบ 3 พารามิเตอร์ คำนวณโดยใช้โปรแกรม LOGIST

3. คะแนนสอบทั้งหมด ใช้คะแนนจากการตอบข้อสอบของนักเรียนที่เลือกมาศึกษาจำนวน 1,850 คน ในรายวิชาวิทยาศาสตร์ระดับชั้นมัธยมปีที่ 3 จำนวน 50 ข้อ

4. กลุ่มข้อสอบง่าย เป็นกลุ่มข้อสอบที่พิจารณาจากค่าพารามิเตอร์ b ที่ได้จากการวิเคราะห์แบบคุณลักษณะแฝง 1 พารามิเตอร์ 2 พารามิเตอร์และ 3 พารามิเตอร์ โดยค่า b เป็นลบทั้ง 3 แบบ หรือเป็นค่าลบ 2 ใน 3 แบบ ซึ่งมีจำนวนข้อสอบ 25 ข้อ

5. กลุ่มข้อสอบยาก เป็นกลุ่มข้อสอบที่พิจารณาจากค่าพารามิเตอร์ b ที่ได้จากการวิเคราะห์แบบคุณลักษณะแฝง 1 พารามิเตอร์ 2 พารามิเตอร์และ 3 พารามิเตอร์ โดยค่า b เป็นบวกทั้ง 3 แบบ หรือเป็นบวก 2 ใน 3 แบบ ซึ่งมีจำนวนข้อสอบ 25 ข้อ

6. กลุ่มผู้สอบที่มีความสามารถต่ำ เป็นกลุ่มผู้สอบที่ได้คะแนนต่ำ มีค่าเปอร์เซนไทล์ที่ 16 ลงมา ซึ่งมีจำนวนผู้สอบ 307 คน

7. กลุ่มผู้สอบที่มีความสามารถสูง เป็นกลุ่มผู้สอบที่ได้คะแนนสูง มีค่าเปอร์เซนไทล์ที่ 84 ขึ้นไป ซึ่งมีจำนวนผู้สอบ 372 คน

<u>นิยามศัพท์ที่ใช้ในการวิจัย</u>

1. ลักษณะการแจกแจงความถี่ หมายถึงการนำเสนอรูปลักษณะของการแจกแจงความถี่ที่แสดงด้วยกราฟความสัมพันธ์ระหว่างค่าความสามารถจริงที่ประมาณได้กับความถี่ของค่าประมาณนั้น และการนำเสนอค่าการวัดแนวโน้มเข้าสู่ส่วนกลาง ค่าการวัดการกระจาย ค่าความเบ้และค่าความโด่งของการแจกแจงของค่าคะแนนที่ประมาณได้ในแต่ละวิธี

2. ค่าประมาณความสามารถจริง หมายถึงค่าที่เป็นผลจากการคำนวณคะแนนสอบตามวิธีการแบบดั้งเดิมและแบบคุณลักษณะแฝง รวม 4 วิธี คือ แบบดั้งเดิม 1 วิธี และแบบคุณลักษณะแฝง 3 วิธี

3. วิธีการแบบดั้งเดิม หมายถึงวิธีการประมาณค่าความสามารถจริงจากคะแนนสอบโดยสมการเชิงเส้นตรง

4. วิธีการแบบคุณลักษณะแฝง หมายถึงวิธีการประมาณค่าความสามารถจริงจากคะแนนสอบ โดยทฤษฎีการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ 2 พารามิเตอร์และ 3 พารามิเตอร์ โดยใช้โปรแกรมคอมพิวเตอร์

5. ความสัมพันธ์ของค่าความสามารถจริง หมายถึงค่าสัมประสิทธิ์สหสัมพันธ์ที่คำนวณจากอันดับที่ของค่าความสามารถจริงที่ประมาณได้ในแต่ละแบบเป็นรายคู่

ข้อตกลงเบื้องต้นของการวิจัย

1. คะแนนสอบที่นำมาศึกษาเป็นคะแนนจากการสอบคัดเลือกนักเรียนชั้นมัธยมปีที่ 3 ที่เข้าศึกษาต่อระดับชั้นมัธยมปีที่ 4 ที่แสดงความสามารถของผู้สอบอย่างเต็มที่

2. คะแนนที่ใช้ในการประมาณค่า ได้มาจากแบบสอบเดียวกันและบุคคลที่สอบชุดเดียวกัน ค่าประมาณที่ได้จึงเป็นผลที่แปรเปลี่ยนไปตามวิธีการคิดคำนวณของแต่ละแบบในการประมาณค่า

3. การประมาณค่าความสามารถจริงโดยเฉพาะแบบคุณลักษณะแฝงนั้นดำเนินภายใต้ระบบคอมพิวเตอร์โปรแกรม BICAL และ LOGIST ที่บ่งบอกค่าการประมาณความสามารถจริงได้อย่างถูกต้อง

ประโยชน์ที่ได้รับจากการวิจัย

การศึกษาครั้งนี้จะได้ข้อสรุปเกี่ยวกับค่าที่อธิบายลักษณะการแจกแจงความถี่ของความสามารถจริงที่ประมาณตามแบบดั้งเดิมและแบบคุณลักษณะแฝง 1 พารามิเตอร์ 2 พารามิเตอร์และ 3 พารามิเตอร์ ทั้งที่จำแนกตามลักษณะความยากและความง่ายของข้อสอบ และตามลักษณะความสามารถสูงและต่ำของผู้สอบ อันจะเป็นประโยชน์ในการตัดสินใจเลือกวิธีการประมาณค่าความสามารถจริงที่สอดคล้องกับสภาพของกลุ่มประชากรและธรรมชาติของแบบสอบ นอกจากนั้นยังได้ข้อมูลเชิงประจักษ์เกี่ยวกับการประมาณค่าความสามารถจริงและความสัมพันธ์ของค่าความสามารถจริงที่ประมาณได้ในแต่ละแบบการประมาณค่าของแต่ละลักษณะกลุ่ม

ความไม่สมบูรณ์ของการวิจัย

เนื่องจากการศึกษาครั้งนี้ใช้คะแนนสอบคัดเลือกซึ่งเป็นสถานการณ์สอบจริง จึงไม่สามารถควบคุมการดำเนินการสอบของแต่ละโรงเรียนที่เป็นกลุ่มตัวอย่างได้ตามขั้นตอนของการวิจัย