ทฤษฎีการตอบสนองต่อข้อสอบ 12

วัตถุประสงค์

เมื่อผู้อ่านได้ศึกษาเนื้อหาบทที่ 12 แล้ว ควรจะมีความสามารถดังนี้

1. อธิบายได้ว่าทฤษฎีการทดสอบแบบคลาสสิคมีจุดอ่อนอย่างไร
2. อธิบายถึงหลักการของทฤษฎีการตอบสนองต่อข้อสอบได้
3. เปรียบเทียบจุดเด่นระหว่างทฤษฎีการตอบสนองต่อข้อสอบ และทฤษฎีการทดสอบแบบคลาสสิคได้

เนื้อหา

12.1 บทนำ

การวิเคราะห์ข้อสอบเท่าที่ทำกันมาในอดีตมักจะยึดทฤษฎีการวัดแบบคลาสสิค (Classical Test Theory) กล่าวคือจะนำจำนวนผู้ตอบถูกในแต่ละข้อมาเป็นเกณฑ์ในการหาความยาก และใช้ความแตกต่างระหว่างจำนวนผู้ตอบถูกในกลุ่มที่ได้คะแนนมาก กับจำนวนผู้ตอบถูกในกลุ่มที่ได้คะแนนน้อยเป็นตัวบ่งชี้อำนาจจำแนก วิธีการวัดดังกล่าวทำในลักษณะที่นำเครื่องมือวัดมาพิจารณาโดยการสัมพันธ์กับผู้ถูกวัด และผู้ถูกวัดก็ยังพิจารณาโดยการนำไปสัมพันธ์กับเครื่อง-มือวัด แล้วทำการตัดสินว่าข้อสอบยากหรือง่าย โดยพิจารณาจำนวนคนที่ทำข้อสอบถูกแต่ละข้อว่ามีมากน้อยเพียงใด ขณะเดียวกันก็ตัดสินความสามารถมากน้อยของผู้ตอบโดยพิจารณาจำนวนข้อที่ทำได้ถูกต้อง ซึ่งเป็นที่แน่นอนว่าแบบสอบนี้หากนำไปสอบกับผู้ที่มีความสามารถมาก แบบสอบฉบับนั้นก็จะเป็นแบบสอบที่ง่าย แบบสอบฉบับเดียวกันนี้หากนำไปให้คนอ่อนทำแบบสอบนั้นก็จะเป็นแบบสอบที่ยากทันที ทำนองเดียวกันหากนำแบบสอบยาก ๆ ไปให้คน ๆ นั้นทำ คน ๆ นั้นก็จะกลายเป็นคนอ่อน และหากนำแบบสอบง่าย ๆ ไปให้ทำ คน ๆ นั้นก็จะกลายเป็นคนเก่งไป ทั้ง ๆ ที่คน ๆ นั้นก็คือคนเดิม และแบบสอบฉบับนั้นก็คือแบบสอบฉบับเดิมนั่นเอง

จะเห็นได้ว่าคุณลักษณะของคนก็ดี คุณลักษณะของแบบสอบก็ดี จะแปรเปลี่ยนไปตามคุณลักษณะของสิ่งที่เอาไปสัมพันธ์ด้วย ตัวอย่างที่มักพบเห็นกันบ่อย ๆ ก็คือกรณีที่นำแบบสอบไปทดลองใช้กับกลุ่มนักเรียนนอกเมือง แล้วนำกลับมาใช้กับนักเรียนในกรุงเทพมหานคร จะพบว่าความยากของแบบสอบนั้นแตกต่างกันมาก

ในด้านอำนาจจำแนกของข้อสอบก็เช่นกัน ค่าอำนาจจำแนกส่วนหนึ่งจะขึ้นอยู่กับลักษณะของกลุ่มตัวอย่างเป็นสำคัญ กล่าวคือหากกลุ่มตัวอย่างที่ทำแบบสอบมีลักษณะเป็นเอกพันธ์ (Homogeneous) ค่าอำนาจจำแนกก็จะต่ำ แต่หากนำไปสอบกับกลุ่มตัวอย่างที่มีลักษณะเป็นวิวิธพันธ์ (Heterogeneous) ค่าอำนาจจำแนกของข้อสอบก็จะสูง

จุดอ่อนของการวิเคราะห์ข้อสอบโดยอาศัยทฤษฎีการวัดแบบคลาสสิค ก็คือ ค่าสถิติต่าง ๆ ของข้อสอบจะแปรเปลี่ยนไปตามกลุ่มตัวอย่างที่มีระดับความสามารถแตกต่างกัน

Hambleton และคณะ (1978) ได้สรุปปัญหาที่สำคัญของการวัดที่ยึดทฤษฎีแบบคลาสสิค ไว้ 3 ประการ ดังนี้

1. คุณสมบัติของข้อคำถามจะแปรเปลี่ยนไปตามกลุ่มผู้สอบ

2. ต้องใช้แบบสอบที่มีข้อคำถามเหมือนกันหรือแบบสอบคู่ขนานเมื่อต้องการเปรียบเทียบความสามารถของผู้สอบ

3. ข้อมูลที่ได้จากการสอบ ไม่สามารถบ่งบอกระดับความสามารถของผู้สอบได้ชัดเจน

จะเห็นได้ว่าทฤษฎีการวัดแบบคลาสสิคไม่สามารถแก้ปัญหาดังที่กล่าวมาแล้วได้ นักวัดจึงได้ช่วยกันคิด และสร้างทฤษฎีขึ้นมาใหม่เพื่อแก้ปัญหาดังกล่าว ทฤษฎีหนึ่งที่นักวัดช่วยกันพัฒนาขึ้นมาก็คือทฤษฎีการตอบสนองต่อข้อสอบ (Item Response Theory) ผู้ที่ทำการบุกเบิกทฤษฎีนี้กันอย่างจริงจังก็คือ Lord, Birnbaum และ Rasch ทฤษฎีการตอบสนองต่อข้อสอบนี้บางทีก็เรียกว่าทฤษฎีความสามารถแฝง (Latent Trait Theory) หรือทฤษฎีโค้งลักษณะข้อสอบ (Item Characteristic Curve Theory)

ทฤษฎีการตอบสนองต่อข้อสอบเป็นทฤษฎีที่เข้ามามีบทบาทในการวัดทางการศึกษาและจิตวิทยาในราวปี ค.ศ. 1950-1960 แต่ทฤษฎีนี้ไม่ได้รับความสนใจจากนักวัดทางการศึกษาและจิตวิทยาเท่าที่ควร สาเหตุที่เป็นเช่นนี้อาจเนื่องมาจากความยุ่งยากซับซ้อนในการคำนวณ แต่ในปัจจุบันนักวัดทั้งหลายต่างก็หันมาสนใจทฤษฎีนี้กันมากขึ้น ทั้งนี้อาจเป็นเพราะปัจจุบันมีโปรแกรมคอมพิวเตอร์สำหรับการวิเคราะห์ข้อสอบตามทฤษฎีนี้ จึงทำให้ลดความยุ่งยากลงได้มาก และทฤษฎีนี้ยังสามารถแก้จุดอ่อนของทฤษฎีการวัดแบบคลาสสิคได้ ตลอดจนสามารถนำไปประยุกต์เพื่อแก้ปัญหาทางการวัด ซึ่งทฤษฎีการวัดแบบคลาสสิคไม่สามารถจะแก้ได้ (Lord 1977 : 117-138)

12.2 ความเป็นมาของทฤษฎีการตอบสนองต่อข้อสอบ

Ferguson (1942) และ Lawley (1943) เป็นผู้ริเริ่มทฤษฎีการตอบสนองต่อข้อสอบ โดยมีหลักการว่าผลการสอบของผู้สอบจากแบบสอบใด ๆ ขึ้นอยู่กับความสามารถ (Ability or Skill) ของผู้สอบ ต่อมาในปี ค.ศ. 1952 Lord ได้เสนอทฤษฎีนี้ใหม่ในรูปของโค้งลักษณะข้อสอบ (Item Characteristic Curve: ICC) โดย Lord ได้เสนอว่าโค้งลักษณะข้อสอบมีลักษณะเป็นโค้งปกติสะสม ต่อมาจึงเรียกว่า Normal Ogive Model ซึ่งโมเดลนี้จะกล่าวถึงพารามิเตอร์ 2 ตัวคือค่าความยากและค่าอำนาจจำแนก แต่เนื่องจากโมเดลนี้มีการคำนวณยุ่งยากมาก และขาดแคลน

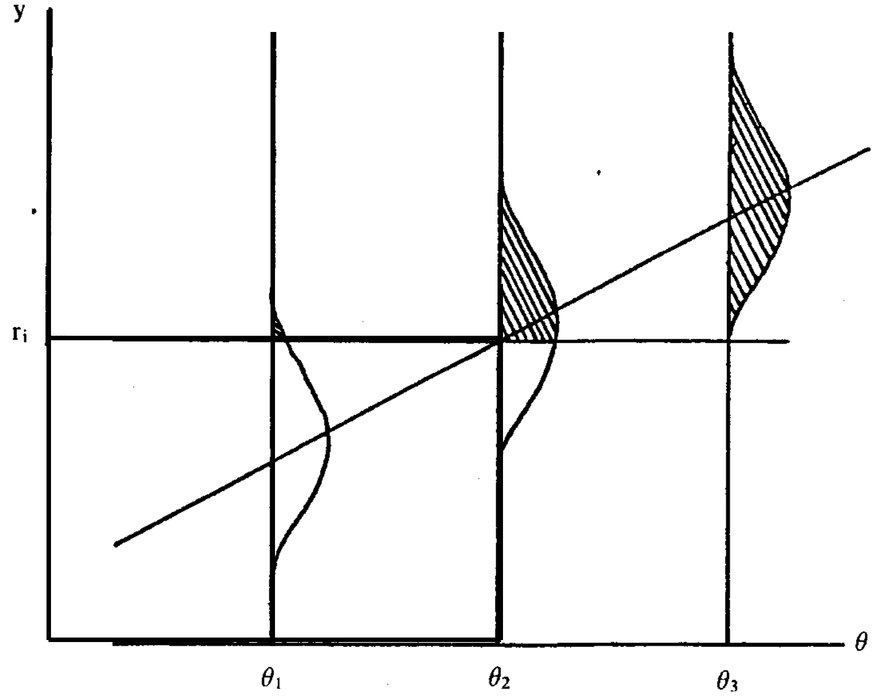
โปรแกรมคอมพิวเตอร์ที่จำเป็นต้องใช้วิเคราะห์ข้อมูลตามทฤษฎี จึงทำให้ Lord หยุดความสนใจในทฤษฎีนี้ไประยะหนึ่ง

ปี ค.ศ. 1960 Rasch ได้ศึกษาเกี่ยวกับเรื่องนี้โดยไม่ทราบแนวคิดของ Lord มาก่อน ซึ่ง Rasch ได้เสนอแนวคิดในรูปของพารามิเตอร์ตัวเดียว คือ ค่าความยาก บางครั้งจึงมีผู้เรียกแบบจำลองนี้ว่า Rasch Model และในปี ค.ศ. 1965 Lord ก็หันมาสนใจและพัฒนาทฤษฎีการตอบสนองต่อข้อสอบใหม่ (Warm 1979 : 19)

ปี ค.ศ. 1968 Birnbaum ได้เสนอ Logistic Model ที่ใช้พารามิเตอร์ 2 ตัวคือ ค่าความยากและค่าอำนาจจำแนก ซึ่งเป็นแบบจำลองที่ง่ายกว่าของ Lord จึงทำให้ Logistic Model เป็นที่นิยมแพร่หลายและมีการพัฒนาขึ้นเรื่อย ๆ จนกระทั่งสามารถใช้ได้กับพารามิเตอร์ตัวเดียว และพารามิเตอร์สามตัว (Warm 1979 : 19-21)

12.3 หลักการของทฤษฎีการตอบสนองต่อข้อสอบ

ทฤษฎีการตอบสนองต่อข้อสอบเป็นทฤษฎีที่อธิบายถึงความสัมพันธ์ระหว่างลักษณะหรือความสามารถที่มีอยู่ในตัวบุคคลกับพฤติกรรมการตอบสนองต่อข้อสอบของบุคคลนั้น โดยทฤษฎีนี้มีความเชื่อว่าพฤติกรรมการตอบสนองต่อข้อสอบของบุคคลจะถูกกำหนดโดยลักษณะหรือความสามารถที่มีอยู่ภายในตัวบุคคล (Lord and Novick 1968 : 358) ซึ่งไม่สามารถจะสังเกตได้ ทฤษฎีนี้พยายามที่จะอธิบายความสัมพันธ์ระหว่างลักษณะภายในตัวบุคคลกับพฤติกรรมที่บุคคลตอบสนองต่อข้อสอบ การอธิบายความสัมพันธ์ดังกล่าวแสดงออกมาในรูปแบบจำลองทางคณิตศาสตร์ โดยให้คะแนนที่ได้รับจากการตอบข้อสอบ (y) แทนพฤติกรรมการตอบสนองต่อข้อสอบ ให้ θ แทนลักษณะหรือความสามารถในตัวบุคคล และ r_i เป็นเกณฑ์ที่บอกว่า y แค่ไหนจึงจะทำข้อสอบข้อ i ได้ถูก ดังนั้น ถ้า $y > r_i$ แสดงว่าทำข้อสอบข้อ i ได้ถูก และถ้า $y < r_i$ แสดงว่าทำข้อสอบข้อ i ผิด ความสัมพันธ์ระหว่างความสามารถ (θ) กับพฤติกรรมการตอบสนอง (y) แสดงได้ดังภาพที่ 1



ภาพที่ 1 แสดงความสัมพันธ์ระหว่างความสามารถกับพฤติกรรมการตอบสนอง

จากภาพจะเห็นได้ว่าถ้านำโอกาสที่จะตอบถูก (พื้นที่ส่วนที่แรเงา) ในระดับความสามารถต่าง ๆ มาเขียนกราฟใหม่ จะได้โค้งลักษณะข้อสอบ (ICC) เป็นรูปต่าง ๆ ทั้งนี้ขึ้นอยู่กับฟังก์ชั่นทางคณิตศาสตร์และจำนวนพารามิเตอร์ที่จะใช้อธิบาย ฟังก์ชั่นทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างระดับความสามารถกับพฤติกรรมการตอบสนองต่อข้อสอบเรียกว่า ฟังก์ชั่นการตอบสนองต่อข้อสอบ (Item Response Function) (Lord 1980 : 12) หรือฟังก์ชั่นลักษณะข้อสอบ (Item Characteristic Function) (Lord and Novick 1968 : 360) ซึ่งสามารถเขียนฟังก์ชั่นได้ดังนี้

$P_i(\theta) = \text{Prob}(U_i = 1|\theta)$ เมื่อ $U_i = 0, 1$

จากฟังก์ชั่นข้างต้นนี้ หมายถึงโอกาสที่ผู้สอบซึ่งมีความสามารถ θ จะตอบคำถามข้อ i ได้ถูกต้อง

12.4 ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองต่อข้อสอบ

ทฤษฎีการตอบสนองต่อข้อสอบนั้น มีข้อตกลงเบื้องต้นดังนี้

1. Unidimensionality เป็นการสมมุติว่าข้อสอบในแบบสอบมีลักษณะเป็นเอกพันธ์ นั่นคือ

แบบสอบนั้นจะต้องมุ่งวัดความสามารถอย่างใดอย่างหนึ่งเพียงความสามารถเดียว หากไม่กำหนดข้อตกลงเบื้องต้นเช่นนี้จะทำให้แบบจำลองของทฤษฎีมีความยุ่งยากมาก (Hambleton and Cook 1977 : 77) ส่วนการตรวจสอบว่าข้อมูลจากการสอบเป็นไปตามข้อตกลงเบื้องต้นหรือไม่ อาจทำได้โดยการวิเคราะห์องค์ประกอบ (Factor Analysis)

2. Local Independence เป็นการกำหนดข้อตกลงเบื้องต้นเกี่ยวกับความเป็นอิสระในการตอบสนองต่อข้อสอบ กล่าวคือการตอบสนองต่อข้อสอบข้อใดข้อหนึ่งของผู้สอบไม่มีผลต่อการตอบสนองต่อข้อสอบข้ออื่น ๆ ในแบบสอบ ในการตรวจสอบว่า ข้อสอบแต่ละข้อเป็นไปตามข้อตกลงเกี่ยวกับ Local Independence หรือไม่ทำได้โดยอาศัยเทคนิคการวิเคราะห์องค์ประกอบเช่นกัน (Hambleton and Cook 1977 : 78)

3. Item Characteristic Curve เป็นข้อตกลงเกี่ยวกับโค้งลักษณะข้อสอบ กล่าวคือโค้งลักษณะข้อสอบเป็นกราฟของฟังก์ชั่นทางคณิตศาสตร์ที่แสดงถึงความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบข้อนั้นถูกกับความสามารถหรือลักษณะที่วัดโดยข้อสอบข้อนั้น โค้งลักษณะข้อสอบมีหลายรูปแบบทั้งนี้ขึ้นอยู่กับแบบจำลองที่ใช้อธิบายความสัมพันธ์ดังกล่าว

12.5 ลักษณะเด่นของทฤษฎีการตอบสนองต่อข้อสอบ

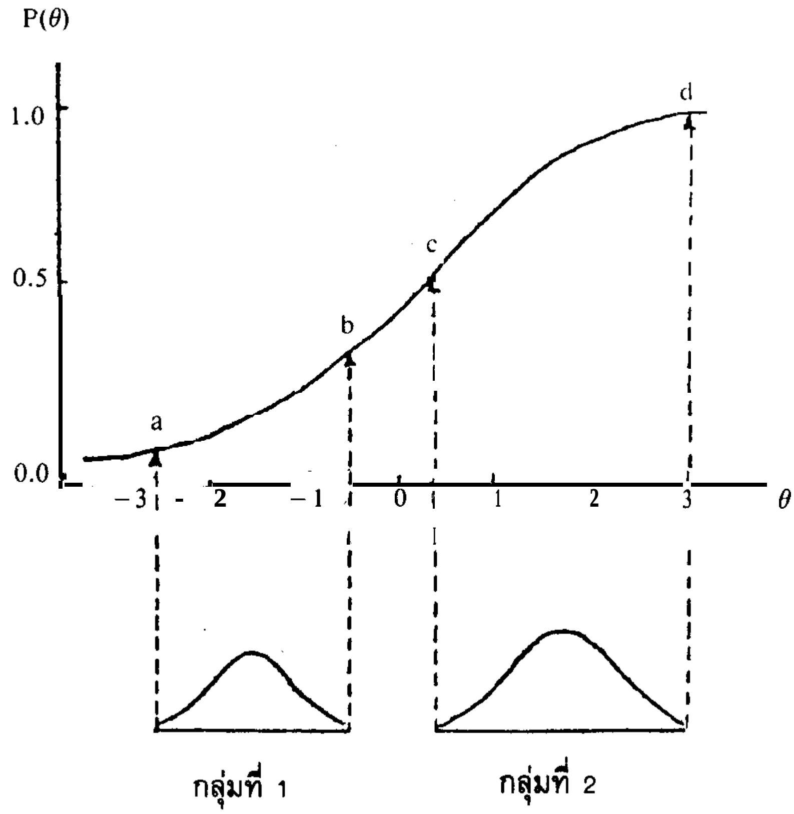
ทฤษฎีการตอบสนองต่อข้อสอบมีความเหนือกว่า (Superiority) ทฤษฎีการสอบแบบคลาสสิค ดังนี้

1. ความไม่แปรเปลี่ยนของค่าพารามิเตอร์ (Invariant of Item Parameter) กล่าวคือไม่ว่าจะประมาณค่าจากกลุ่มตัวอย่างที่มีความสามารถระดับใดก็ตาม ค่าพารามิเตอร์จะไม่แปรเปลี่ยนไปตามกลุ่มของผู้สอบ นั่นคือค่าความยาก ค่าอำนาจจำแนก และค่าโอกาสในการตอบถูกโดยการเดาจะไม่แปรเปลี่ยนไม่ว่าจะประมาณค่าจากผู้สอบกลุ่มใด

2. จะใช้ข้อสอบกับใครก็ได้ (Person-Free) นั่นคือไม่ว่าจะนำข้อสอบไปใช้สอบกับบุคคลกลุ่มใด โค้งลักษณะข้อสอบก็จะคงเดิม

3. จะใช้ข้อสอบข้อใดก็ได้ (Item-Free) ในการประมาณค่าความสามารถแท้ (θ) ของผู้สอบโดยจะใช้ข้อสอบชุดใดจำนวนเท่าใดก็ได้ บางครั้งอาจใช้ข้อสอบเพียง 3-5 ข้อ ก็สามารถประมาณค่าความสามารถแท้ของผู้สอบได้แล้ว

เกี่ยวกับความไม่แปรเปลี่ยนของค่าพารามิเตอร์นั้นสามารถอธิบายได้ด้วยภาพที่ 2



ภาพที่ 2 ความไม่แปรเปลี่ยนของค่าพารามิเตอร์ของข้อสอบที่ประมาณค่าจากกลุ่มที่แตกต่างกัน

จากภาพแสดงให้เห็นว่าในการประมาณค่าจากกลุ่มผู้สอบที่มีความสามารถต่ำ (กลุ่มที่ 1) เส้นโค้งลักษณะข้อสอบที่ได้จะอยู่ในช่วง a-b เส้นโค้งลักษณะข้อสอบในช่วงนี้จะอธิบายถึงโอกาสในการตอบข้อสอบข้อนี้ถูกของผู้สอบที่มีความสามารถต่ำ และในทำนองเดียวกันเส้นโค้งลักษณะข้อสอบในช่วง c-d ก็จะแสดงถึงโอกาสที่จะตอบข้อสอบนี้ถูกของผู้สอบที่มีความสามารถสูง (กลุ่มที่ 2) ถ้าข้อสอบข้อนี้วัดความสามารถอันเดียวกันของกลุ่มตัวอย่างทั้งสองกลุ่มแล้วเส้นโค้งลักษณะข้อสอบย่อมมีเส้นเดียว เส้นโค้งในช่วง a-b และช่วง c-d ย่อมเป็นส่วนหนึ่งของเส้นโค้งลักษณะข้อสอบข้อนี้ เมื่อเส้นโค้งลักษณะข้อสอบมีได้เส้นเดียว ตัวพารามิเตอร์ซึ่งเป็นตัวกำหนดรูปร่างลักษณะของเส้นโค้งก็ย่อมมีได้ค่าเดียวเช่นกัน หรืออาจกล่าวได้ว่าถ้าพารามิเตอร์ของข้อสอบไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบแล้วข้อสอบที่นำไปสอบกับผู้สอบที่มีความสามารถแตกต่างกัน เมื่อนำข้อมูลที่ได้จากแต่ละกลุ่มมาวิเคราะห์จะมีค่าพารามิเตอร์ของข้อสอบเป็นค่าเดียวกัน

Lord (1980 : 36) ได้อธิบายให้เห็นว่าการไม่แปรเปลี่ยนของค่าพารามิเตอร์นั้น ไม่ได้หมายความว่าค่าพารามิเตอร์ที่ได้จากการประมาณค่าโดยใช้กลุ่มตัวอย่างที่แตกต่างกันจะมีค่าเท่ากันเสมอ ค่าพารามิเตอร์ของข้อสอบจะเท่ากันหรือไม่ขึ้นอยู่กับเงื่อนไขในการทดสอบบางประการเช่น ถ้าเลือกมาตรวัดที่มีจุดเริ่มต้นเดียวกันและหน่วยในการวัดหน่วยเดียวกันแล้ว ค่าพารามิเตอร์ที่ประมาณค่าจากกลุ่มตัวอย่างที่แตกต่างกันจะมีค่าเท่ากัน ในทำนองกลับกันหากเลือกมาตรวัดที่มีจุดเริ่มต้นและมีหน่วยในการวัดแตกต่างกันแล้ว การไม่แปรเปลี่ยนของค่าพารามิเตอร์ หมายความว่าค่าพารามิเตอร์ที่ประมาณได้จากกลุ่มตัวอย่างที่ต่างกันของข้อสอบชุดหนึ่งจะมีความสัมพันธ์กันในเชิงเส้นตรง

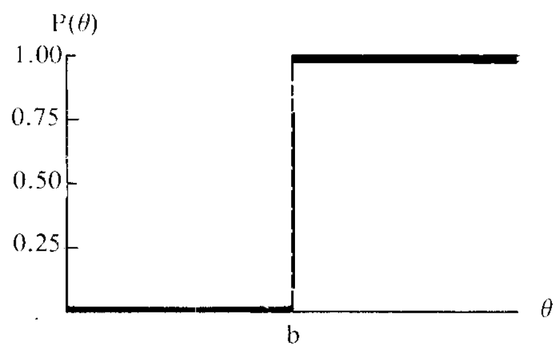
นอกจากนี้ Baker (1977 : 170) ยังได้อธิบายเพิ่มเติมเกี่ยวกับการไม่แปรเปลี่ยนของค่าพารามิเตอรนี้ว่าขึ้นอยู่กับเงื่อนไข 2 ประการคือ ความสามารถที่กล่าวถึงต้องสามารถนิยามได้ชัดเจนและวัดได้ด้วยข้อสอบ กับความสามารถที่วัดนั้นจะต้องมีความคงที่ภายในช่วงระยะเวลาหนึ่ง

แบบจำลองที่ใช้ในทฤษฎีการตอบสนองต่อข้อสอบ

ทฤษฎีการตอบสนองต่อข้อสอบได้มีการพัฒนาแบบจำลองขึ้นมาหลายแบบด้วยกัน โดยแบบจำลองแต่ละแบบนั้นจะต่างกันในรูปของฟังก์ชั่นทางคณิตศาสตร์และจำนวนพารามิเตอร์ที่จะอธิบาย ซึ่งในที่นี้จะขอเสนอแบบจำลองทางคณิตศาสตร์ในทฤษฎีการตอบสนองต่อข้อสอบดังนี้

Guttman Perfect Scale

ลักษณะของฟังก์ชั่นเป็นแบบขั้นบันได (Step Function) กล่าวคือโอกาสในการตอบข้อสอบข้อใดข้อหนึ่งถูกมีความสัมพันธ์กับความสามารถของบุคคล โดยโอกาสในการตอบข้อสอบถูกจะเป็น 0 หรือ 1 เท่านั้น เส้นโค้งลักษณะข้อสอบจะมีลักษณะดังภาพที่ 3



ภาพที่ 3 เส้นโค้งลักษณะข้อสอบที่เป็น Guttman Perfect Scale

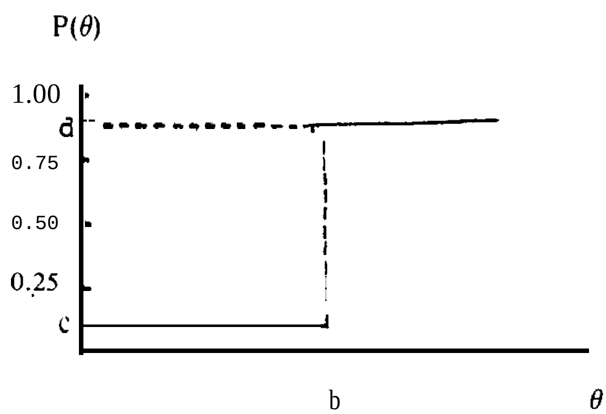
จากภาพผู้ที่มีความสามารถต่ำกว่า b ($\theta < b$) มีโอกาสที่จะตอบข้อสอบถูกเท่ากับ 0 ส่วนผู้ที่มีความสามารถมากกว่า b ($\theta > b$) มีโอกาสที่จะตอบข้อสอบถูกเป็น 1 สมการที่แสดงความสัมพันธ์ระหว่าง θ กับโอกาสที่จะตอบข้อสอบถูก (P (θ)) เป็นดังนี้

$$P(\theta) = \begin{cases} 1 \text{ ถ้า } \theta > b \\ 0 \text{ ถ้า } \theta < b \end{cases} \qquad \text{..........} (1)$$

แบบจำลอง Guttman Perfect Scale มีลักษณะเป็น Deterministic Model และมีการนำไปประยุกต์ใช้น้อยมากทั้งนี้เพราะเป็นไปได้ยากที่ข้อมูลจากการสอบจะสอดคล้องกับแบบจำลองนี้

Latent Distance Model

ลักษณะของฟังก์ชั่นยังคงเป็นแบบขั้นบันไดเช่นเดียวกันกับ Guttman Perfect Scale แต่โอกาสในการตอบข้อสอบถูกจะแปรเปลี่ยนระหว่าง 0 ถึง 1 โค้งลักษณะข้อสอบ มีลักษณะดังภาพที่ 4



ภาพที่ 4 เส้นโค้งลักษณะข้อสอบใน Latent Distance Model

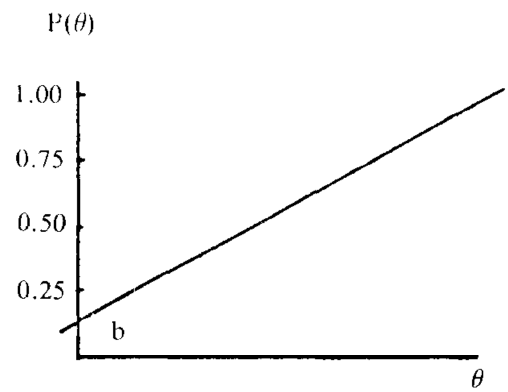
จากภาพแสดงว่าผู้ที่มีความสามารถต่ำกว่า b มีโอกาสที่จะตอบข้อสอบข้อนี้ถูกเท่ากับ c ส่วนผู้ที่มีความสามารถมากกว่า b มีโอกาสที่จะตอบข้อสอบข้อนี้ถูกเท่ากับ d โดยที่ $0 < c < d < 1.00$ ซึ่งสามารถเขียนสมการแสดงความสัมพันธ์ระหว่าง θ กับโอกาสที่จะตอบข้อสอบถูก (P (θ)) ได้ดังนี้

$$P(\theta) = \begin{cases} c \text{ ถ้า } \theta < b \\ d \text{ ถ้า } \theta > b \end{cases} \qquad \text{..........} (2)$$

แบบจำลองนี้มีลักษณะเป็น Stochastic Model อย่างไรก็ตามโอกาสที่ข้อมูลจะสอดคล้องกับแบบจำลองนี้ก็เป็นไปได้ยาก เพราะบุคคลที่มีความสามารถมากกว่า หรือน้อยกว่า b จะแสดงพฤติกรรมการตอบที่คงเส้นคงวาเช่นนี้คงเป็นไปได้ยาก

Linear Model

โค้งลักษณะข้อสอบเป็นเส้นตรง โดยโอกาสในการตอบข้อสอบถูกจะเป็นสัดส่วนโดยตรงกับความสามารถของบุคคล เส้นโค้งลักษณะข้อสอบจะมีลักษณะดังภาพที่ 5



ภาพที่ 5 เส้นโค้งลักษณะข้อสอบตาม Linear Model

สมการที่แสดงความสัมพันธ์ระหว่าง θ กับโอกาสที่จะตอบข้อสอบถูก ($P(\theta)$) เป็นดังนี้

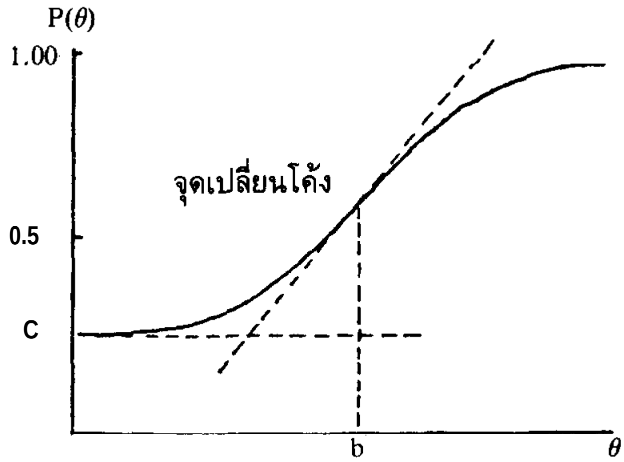
$$P(\theta) = b + a\theta \qquad \ldots\ldots\ldots\ (3)$$

จากสมการ (3) ถ้า $a \neq 0$ แล้วผู้มีความสามารถในระดับต่ำมาก ๆ มีโอกาสที่จะตอบข้อสอบถูกเป็นลบ ส่วนผู้ที่มีความสามารถสูงมาก ๆ ก็มีโอกาสที่จะตอบข้อสอบถูกมีค่าเกิน 1 ได้ Lazarsfeld ผู้เสนอแบบจำลองนี้ได้กำหนดข้อตกลงเบื้องต้นเพิ่มเติมว่าการแจกแจงของความสามารถ (θ) จะต้องอยู่ในช่วงที่ทำให้ $P(\theta)$ มีค่าอยู่ระหว่าง 0 ถึง 1 เท่านั้น (Torgerson 1958: 368) แบบจำลองดังกล่าวนี้มีข้อจำกัดในการนำไปใช้มาก ทั้งนี้เนื่องจากข้อตกลงเบื้องต้นที่กำหนดเพิ่มเติมนั่นเอง

แบบจำลองทั้งหมดที่กล่าวมานี้เป็นแบบจำลองที่ไม่ค่อยจะมีความสอดคล้องกับข้อมูลจากการสอบในสถานการณ์จริงมากนัก แต่อย่างไรก็ตามแบบจำลองดังกล่าวเป็นแบบจำลองที่ให้แนวคิดเชิงทฤษฎีที่มีคุณค่าต่อการพัฒนาแบบจำลองที่เหมาะสมในระยะเวลาต่อมา

Normal Ogive Model

โค้งลักษณะข้อสอบตามแบบจำลองนี้จะเป็นรูปตัว S ดังภาพที่ 6



ภาพที่ 6 เส้นโค้งลักษณะข้อสอบของ ***Normal Ogive Model***

จากภาพแสดงให้เห็นว่าโค้งลักษณะข้อสอบเป็นกราฟที่แสดงความสัมพันธ์ระหว่างโอกาสที่จะตอบข้อสอบข้อใดข้อหนึ่งถูกกับระดับความสามารถของผู้สอบ โดยผู้สอบที่มีระดับความสามารถสูงมีโอกาสที่จะตอบข้อสอบถูกมากกว่าผู้สอบที่มีความสามารถต่ำ ค่าความสามารถและค่าความยากของข้อสอบจะอยู่บนสเกลเดียวกันโดยมีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติแล้วระดับความสามารถจะมีค่าอยู่ระหว่าง -3 ถึง $+3$ ค่า -3 หมายถึงผู้สอบมีความสามารถต่ำมาก ค่า $+3$ หมายถึงผู้สอบมีความสามารถสูงมากค่าความยากจะมีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$ เช่นเดียวกันกับค่าอำนาจจำแนกของข้อสอบ (Hambleton and Cook 1977) ส่วนค่าการเดานั้นหมายถึงค่าที่ปลายต่ำสุด (Lower Tail) ของโค้งลักษณะข้อสอบ ซึ่งมีค่าอยู่ระหว่าง 0 ถึง 1

ความสัมพันธ์ดังกล่าวนี้จะอยู่ในรูปฟังก์ชั่นการแจกแจงสะสมแบบปกติ (Cumulative Normal Ogive) ดังนี้

$$P, (\theta) \quad = \quad c_i \; t \; (1 - c_i) \int_{-\infty}^{a_i(\theta - b_i)} \frac{1}{\sqrt{2\pi}} e^{-t^{2/2}} \, dt \qquad \text{.........} \quad (4)$$

เมื่อ $P_i(\theta)$ แทนโอกาสที่ผู้สอบซึ่งมีความสามารถ θ จะตอบคำถามข้อ i ได้ถูกต้อง

a_i แทนค่าอำนาจจำแนกของข้อสอบข้อ i

b_i แทนค่าความยากของข้อสอบข้อ i

c_i แทนค่าการเดา

ค่าพารามิเตอร์ของข้อสอบ (a_i, b_i และ c_i) จะเป็นตัวที่ทำให้รูปร่างของโค้งลักษณะข้อสอบแต่ละข้อแตกต่างกันไป ซึ่งสามารถอธิบายรายละเอียดของค่าพารามิเตอร์ต่าง ๆ เหล่านี้ได้ดังนี้

a_i เป็นสัดส่วนโดยตรงกับความชัน (Slope) ของโค้งลักษณะข้อสอบ ณ จุดเปลี่ยนโค้ง

$$\text{ความชัน ณ จุดเปลี่ยนโค้ง} = \begin{cases} \dfrac{a(1-c)}{\sqrt{2\pi}} & \text{กรณีที่มีการเดา} \\[2em] \dfrac{a}{\sqrt{2\pi}} & \text{กรณีที่ไม่มีการเดา} \end{cases}$$

ค่า a_i นี้จะมีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$

b_i เป็นค่าของระดับ θ ที่ทำให้โค้งลักษณะข้อสอบมีการเปลี่ยนโค้ง (จุดเปลี่ยนโค้งจะอยู่ที่ b, P (b)) ค่าพารามิเตอร์ b_i นี้จะบอกตำแหน่งของโค้งลักษณะข้อสอบ ณ $\theta = b$ ซึ่งถ้าไม่มีการเดา ($c = 0$) แล้วโอกาสที่จะตอบข้อสอบข้อนั้นถูกจะมีค่าเท่ากับ $\frac{1-c}{2}$ หรือมีค่าเท่ากับ 0.5

ค่า b_i นี้มีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$

c_i เป็นโอกาสของผู้สอบที่มีความสามารถในระดับต่ำจะตอบคำถามข้อ i ได้ถูกต้อง ซึ่งก็คือค่าโอกาสที่จะตอบถูกโดยการเดานั่นเอง c_i เป็น lower assymtote ของโค้งลักษณะข้อสอบ

ค่า c_i นี้จะมีค่าอยู่ระหว่าง 0 ถึง 1

สมการ (4) เป็นสมการที่มีพารามิเตอร์ 3 ตัว จึงอาจเรียกว่า Three-parameter Normal Ogive Model แต่หากในสถานการณ์การสอบที่ผู้สอบมีโอกาสตอบถูกโดยการเดามีน้อยมาก หรือไม่มีเลย สมการ (4) ก็จะเหลือค่าพารามิเตอร์เพียง 2 ตัว ดังนี้

$$P_i\ (\theta) = \int_{-\infty}^{a_i(\theta - b_i)} \frac{1}{\sqrt{2\pi}}\ e^{-t^2/2}\ dt \qquad (5)$$

Normal Ogive Model นั้นมีปัญหาบางประการทั้งในด้านทฤษฎีและคณิตศาสตร์ จึงได้

มีผู้พยายามคิดหาแบบจำลองที่มีรูปร่างคล้ายคลึงกันกับ Normal Ogive Model แต่มีความสะดวกกว่าในทางคณิตศาสตร์ แบบจำลองดังกล่าวนี้ก็คือ Logistic Model

Logistic Model

เป็นแบบจำลองที่มีลักษณะใกล้เคียงกับ Normal Ogive Model มากที่สุด โค้งลักษณะข้อสอบมีลักษณะเป็นรูปตัว S เช่นเดียวกับ Normal Ogive Model ส่วนความสัมพันธ์ระหว่างโอกาสในการตอบถูกกับระดับความสามารถอยู่ในรูปฟังก์ชั่นการแจกแจงสะสมแบบโลจิสต์ (Logistic Cumulative Distribution Function) ดังนี้

$$P_i(\theta) = c_i + (1-c_i)\left[1+e^{-1.7a_i(\theta-b_i)}\right]^{-1} \qquad \text{.......... (6)}$$

โดยที่ a_i, b_i และ c_i เป็นค่าพารามิเตอร์ของข้อสอบ และมีความหมายเช่นเดียวกับ Normal Ogive Model ส่วน e เป็นค่าคงที่มีค่าประมาณ 2.71828... และเมื่อมีการปรับค่าด้วย Scaling Factor (มีค่า 1.7) แล้วจะทำให้ค่าฟังก์ชั่น Normal Ogive กับค่าฟังก์ชั่น Logist แตกต่างกันเพียงเล็กน้อย แต่ในแง่ของการคำนวณแล้ว Logistic Model มีความง่ายและสะดวกกว่ามาก นอกจากนี้ในการสอบจริงอาจมีผู้ที่มีความสามารถสูงตอบผิดด้วยความเลินเล่อ กรณีเช่นนี้ Logistic Model มีความแกร่ง (Robustness) ต่อข้อมูลมากกว่า Normal Ogive Model จึงทำให้ Logistic Model เป็นที่นิยมกันมากในทางปฏิบัติ (Lord 1980 : 14)

สมการ (6) นี้อาจเรียกว่า Three-parameter Logistic Model เพราะมีตัวพารามิเตอร์ในสมการ 3 ตัว

Haebara (1979) ได้อธิบายโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว ไว้ดังนี้

$$\lim_{\theta \to -\infty} P_i(\theta) = c_i \qquad \text{..........(7)}$$

$$P_i(\theta/\theta = b_i) = \frac{1}{2} + \frac{c_i}{2} \qquad \text{. . (8)}$$

$$\left(\frac{\partial P_i(\theta)}{\partial \theta}\right)_{\theta=b_i} = \frac{D}{4} a_i (1-c_i) \qquad \text{........ (9)}$$

$$\frac{\partial^2 P_i(\theta)}{\partial \theta^2} \begin{cases} = 0, & \text{ถ้า } \theta = b_i \\ > 0, & \text{ถ้า } \theta < b_i \\ < 0, & \text{ถ้า } \theta > b_i \end{cases} \qquad \text{......... (10)}$$

จากสมการ (7) จะเห็นได้ว่าผู้สอบที่มีความสามารถต่ำสุด $(-\infty)$ มีโอกาสที่จะตอบข้อสอบข้อ i ถูก ด้วยความน่าจะเป็นเท่ากับ c_i และในกรณีที่ผู้สอบมีระดับความสามารถเท่ากับความยากของข้อสอบ $(\theta = b_i)$ แล้วโอกาสที่จะตอบข้อสอบข้อนั้นถูกจะเท่ากับ $\frac{1}{2} + \frac{c_i}{2}$ ดังสมการ (8) และที่ระดับความสามารถของผู้สอบ ณ จุดนี้ข้อสอบจะมีอำนาจจำแนกสูง โดยโค้งลักษณะข้อสอบจะมีความชันเท่ากับ $\frac{D}{4}\, a_i\,(1 - c_i)$ ดังสมการ (9) ค่าความชันนี้จะค่อย ๆ เพิ่มขึ้นเมื่อระดับความสามารถของผู้สอบมีค่าน้อยกว่าความยาก $(\theta < b_i)$ และค่าความชันนี้จะค่อย ๆ ลดลงเมื่อระดับความสามารถมีค่ามากกว่าค่าความยากของข้อสอบ $(\theta > b_i)$

ในกรณีที่ผู้สอบมีโอกาสตอบถูกโดยการเดาน้อยมากหรือไม่มีเลย อาจกำหนดให้ค่าการเดา (c_i) เป็น 0 สมการก็จะเหลือค่าพารามิเตอร์เพียง 2 ตัวคือ a_i และ b_i ดังนี้

$$P_i(\theta) = \frac{1}{1 + e^{-1.7\, a_i\, (\theta - b_i)}} \qquad \text{.......... (11)}$$

สมการ (11) นี้ก็คือ Two-parameter Logistic Model

ในกรณีที่กำหนดให้ $c_i = 0$ และถือว่าข้อสอบทุกข้อมีอำนาจจำแนกเท่ากัน $(a_i = \bar{a})$ สมการก็จะเหลือค่าพารามิเตอร์เพียงตัวเดียวคือ b_i ดังนี้

$$P_i(\theta) = \frac{1}{1 + e^{-1.7\, \bar{a}\, (\theta - b_i)}} \qquad \text{.......... (12)}$$

สมการ (12) นี้ก็คือ One-parameter Logistic Model

ในขณะที่ Lord พัฒนา Normal Ogive Model และ Birnbaum ได้เสนอ Logistic Model นั้น Rasch นักคณิตศาสตร์ชาวเดนมาร์กได้เสนอแบบจำลองการตอบสนองต่อข้อสอบ ดังนี้

$$P_i(\theta^*) = \frac{\theta^*}{\theta^* + b_i^*} \qquad \text{.......... (13)}$$

เมื่อ $\theta^* = e^{1.7\, \bar{a}\, \theta}$

$b_i^* = e^{1.7\, \bar{a}\, b_i}$

เมื่อพิจารณาแล้วจะเห็นว่าสมการ (13) ของ Rasch ก็คือสมการ (12) ของ One-parameter Logistic Model นั่นเอง บางครั้งจึงมีผู้เรียกแบบจำลองนี้ว่า Rasch Model

นอกจาก Logistic Model แล้ว Hambleton และคณะ (1978 : 476) ได้เสนอแบบจำลองอื่น ๆ อีก ดังนี้

Nominal Response Model

จากแบบจำลองที่กล่าวมาข้างต้นนั้นใช้ได้กับการให้คะแนนแบบ 2 ค่า (Binary Item) ส่วน Nominal Response Model นั้นเป็นแบบจำลองที่สามารถประยุกต์ใช้ได้กับข้อสอบที่มีการตรวจให้คะแนนมากกว่า 2 ค่า (Multichotomously Scored) ผู้เสนอแบบจำลองนี้คือ Bock จุดมุ่งหมายของแบบจำลองนี้ก็เพื่อเพิ่มความแม่นยำในการประมาณค่าความสามารถของผู้สอบ แบบจำลองนี้จะมีโค้งลักษณะของตัวเลือกแต่ละตัว (Item-option Characteristic Curve) โค้งของตัวเลือกที่เป็นคำตอบถูกจะมีลักษณะเป็นฟังก์ชั่นที่มีค่าเพิ่มขึ้นโดยตลอด (Monotonic Increasing Function) ส่วนโค้งของตัวเลือกที่ผิดจะมีลักษณะอย่างไรขึ้นอยู่กับการเลือก รูปแบบทางคณิตศาสตร์ของโค้งลักษณะของตัวเลือกมีได้หลายแบบ ตัวอย่างเช่น รูปแบบของ Bock ที่สมมุติว่าโอกาสที่ผู้มีความสามารถเท่ากับ θ จะเลือกตอบตัวเลือก k (จากตัวเลือกทั้งหมด m ตัว) ของข้อสอบข้อที่ i จะเป็นดังนี้

$$P_{ik}(\theta) = \frac{e^{b_{ik}^{*} + a_{ik}^{*}\theta}}{\sum_{k=1}^{m} e^{b_{ik}^{*} + a_{ik}^{*}}} \qquad \text{..........} (14)$$

เมื่อ b_{ik}^{*} และ a_{ik}^{*} คือค่าพารามิเตอร์ของตัวเลือกที่ k สำหรับระดับความสามารถ θ ใด ๆ ผลรวมของโอกาสในการเลือกตัวเลือกทุกตัวจะเท่ากับ 1 ในกรณีที่ข้อสอบมี 2 ตัวเลือก $(m = 2)$ แบบจำลองนี้ก็คือ Two-parameter Logistic Model นั่นเอง

Grade Response Model

เป็นกรณีพิเศษของ Nominal Response Model แบบจำลองนี้จะใช้ในกรณีที่พฤติกรรมการตอบสนองสามารถจัดเรียงลำดับ (Order) ได้ เช่นกรณีคะแนนจากแบบสอบวัดทัศนคติเป็นต้น Samejima ได้เสนอแบบจำลองนี้ โดยสมมุติว่าพฤติกรรมการตอบสนองต่อข้อสอบสามารถจัดแบ่งออกได้เป็น $m_i + 1$ ประเภท และให้คะแนน $x_i = 0, 1, \ldots, m_i$ ตามลำดับ โอกาสที่ผู้มีความสามารถระดับ θ จะได้คะแนนข้อนี้เป็น x_i คือ

$$P_{x_i}(\theta) = P_{x_i}^{*}(\theta) - P_{(x_i+1)}^{*}(\theta) \qquad \text{..........} (15)$$

เมื่อ $P_{x_i}^{*}(\theta)$ คือ Item Response Function ของการให้คะแนนแบบ 2 ค่า โดยผู้ที่ได้คะแนนต่ำกว่า x_i ถือเป็น 0 และผู้ที่ได้คะแนนเท่ากับ x_i หรือมากกว่าถือเป็น 1

สรุปเนื้อหาบทที่ 12

1. จุดอ่อนของทฤษฎีการวัดแบบคลาสสิคก็คือ ค่าสถิติต่าง ๆ ของข้อสอบจะแปรเปลี่ยนไปตามกลุ่มตัวอย่างที่มีระดับความสามารถแตกต่างกัน
2. ทฤษฎีการตอบสนองต่อข้อสอบเป็นทฤษฎีที่อธิบายความสัมพันธ์ระหว่างลักษณะภายในตัวบุคคลกับพฤติกรรมที่บุคคลตอบสนองต่อข้อสอบ การอธิบายความสัมพันธ์ดังกล่าวแสดงออกมาในรูปแบบจำลองทางคณิตศาสตร์
3. ทฤษฎีการตอบสนองต่อข้อสอบเป็นทฤษฎีที่มีข้อตกลงเบื้องต้นที่ค่อนข้างจะเข้มงวดกว่าทฤษฎีการทดสอบแบบคลาสสิค ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองต่อข้อสอบ คือ Unidimensionality, Local Independent และ Item Characteristic Curve
4. จุดเด่นของทฤษฎีการตอบสนองต่อข้อสอบ คือ
    1. ความไม่แปรเปลี่ยนของค่าพารามิเตอร์ (Invariant of Item Parameter)
    2. จะใช้ข้อสอบกับใครก็ได้ (Person-Free)
    3. จะใช้ข้อสอบข้อใดก็ได้ (Item-Free)

คำถามท้ายบทที่ [illegible]

1. เหตุใดการวิเคราะห์ข้อสอบโดยยึดทฤษฎีการวัดแบบคลาสสิค (Classical Test Theory) **จึง**ไม่เป็นที่ยอมรับในปัจจุบัน
2. ทฤษฎีการตอบสนองต่อข้อสอบ มีจุดเด่นอย่างไร
3. ทฤษฎีการตอบสนองต่อข้อสอบ มีข้อตกลงเบื้องต้นที่สำคัญ ๆ อะไรบ้าง