

บทที่ 2

วรรณคดีที่เกี่ยวข้อง

การศึกษาเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบสอบแบบตัวเลือกชั้นระหว่างประเภทที่กำหนดและไม่กำหนดสถานการณ์คำตอบ ผู้วิจัยได้แบ่งเนื้อหาที่เกี่ยวข้องออกเป็น 3 ตอน คือ

ตอนที่ 1 มโนทัศน์เบื้องต้นเกี่ยวกับแบบสอบเลือกตอบแบบตัวเลือกชั้น

ตอนที่ 2 มโนทัศน์เกี่ยวกับทฤษฎีการตอบสนองข้อสอบ

ตอนที่ 3 งานวิจัยที่เกี่ยวข้อง

ตอนที่ 1 มโนทัศน์เบื้องต้นเกี่ยวกับแบบสอบเลือกตอบแบบตัวเลือกชั้น

แบบสอบเลือกตอบแบบตัวเลือกชั้น (Double or Complex Multiple Choice Test) เป็นรูปแบบตัวเลือกที่นักวัดผลพยายามปรับปรุงแก้ไขข้อบกพร่องของการวัดที่บางเนื้อหาวิชาหรือบางพฤติกรรมที่ต้องการวัดนั้นมีคำตอบถูกหลายคำตอบ และแต่ละคำตอบนั้นต่างก็เป็นความจริงหรือมีความสำคัญเท่าเทียมกัน และเราต้องการวัดว่าผู้สอบรู้คำตอบเหล่านั้นครบถ้วนหรือไม่ หากเรานำแบบสอบเลือกตอบตัวเลือกธรรมดามาใช้แล้ว เราจะวัดสิ่งที่สำคัญเหล่านั้นได้เพียงข้อละสถานการณ์เดียวเท่านั้นหรือไม่ก็ต้องถามคำถามซ้ำ ๆ จนครบ ซึ่งเป็นสิ่งที่ไม่ควรกระทำ จึงทำให้เกิดความพยายามที่จะหารูปแบบของข้อสอบที่จะแก้ปัญหาดังกล่าว เช่น การเขียนตัวเลือกที่มีคำตอบถูกไม่จำกัด แต่ก็ต้องพบกับปัญหาในเรื่องการตรวจและการให้คะแนน ต่อมาจึงมีการดัดแปลงตัวเลือกให้อยู่ในรูปของการรวมคำตอบและเขียนเป็นแบบตัวเลือกชั้น ซึ่งมีผู้นิยมนำมาใช้อย่างแพร่หลาย เพราะสามารถวัดได้ถึงขั้นวิเคราะห์ สังเคราะห์ และความคิดรวบยอดของเนื้อหา นั้น ๆ ได้เป็นอย่างดี สามารถนำมาใช้ได้เป็นอย่างดีกับวิชาสังคมศึกษา คณิตศาสตร์ วิทยาศาสตร์ โดยเฉพาะกับเนื้อหาที่มี

คำตอบถูกได้หลายอย่างหรือมีสาเหตุที่เกี่ยวข้องกันอยู่หลายประการ (ชวาล แพรัตกุล, มปป) ซึ่งเป็นความคิดเห็นที่สอดคล้องกับ Thorndike and Hagen (1977) ที่มีความเห็นว่า แบบสอบเลือกตอบแบบตัวเลือกซ้อนเป็นแบบสอบที่มีประสิทธิภาพ เมื่อใช้วัดเกี่ยวกับการใช้ความรู้ หลักการ ความสามารถในการประยุกต์ใช้ หรือความสัมพันธ์ระหว่างเหตุผลหลาย ๆ อย่าง นอกจากนั้นแบบสอบเลือกตอบตัวเลือกซ้อนยังช่วยลดความคลาดเคลื่อนจากการใช้แบบสอบที่มีตัวเลือกแบบปิดให้หมดไปได้ โดยการเขียนข้อสอบให้อยู่ในรูปแบบตัวเลือกซ้อน (Copperud, 1979) ในด้านการตรวจให้คะแนนก็สามารถทำได้สะดวกรวดเร็วและสามารถนำผลการสอบมาวิเคราะห์เพื่อปรับปรุงข้อสอบได้เช่นเดียวกับแบบสอบเลือกตอบแบบตัวเลือกธรรมดา

แบบสอบเลือกตอบแบบตัวเลือกซ้อน ประกอบด้วย 3 ส่วน คือ ส่วนที่เป็นคำถาม ส่วนที่เป็นข้อความหรือสถานการณ์และส่วนที่เป็นตัวเลือก ส่วนที่เป็นคำถามเขียนโดยใช้หลักการเกี่ยวกับการเขียนข้อคำถามของข้อสอบแบบตัวเลือกธรรมดา ส่วนที่เป็นข้อความหรือสถานการณ์จะเขียนเช่นเดียวกับการเขียนตัวเลือกแบบธรรมดาโดยที่ข้อความหรือสถานการณ์นั้นอาจเป็นคำตอบถูกเพียง 1 สถานการณ์หรือถูกมากกว่า 1 สถานการณ์ก็ได้ (อาจถูก 2,3 สถานการณ์หรือถูกทั้งหมด) และส่วนที่เป็นตัวเลือกจะเขียนจากการรวมข้อความหรือสถานการณ์ที่เขียนไว้แล้วนั้น มาผสมกันเป็นชุด ๆ ตามความเหมาะสม จากวิธีการดังกล่าวจะทำให้สามารถเขียนตัวเลือกที่เป็นไปได้หลายตัวเลือก เช่นเมื่อกำหนดข้อความหรือสถานการณ์มาให้ 3 ข้อ จะสามารถเขียนตัวเลือกได้ 8 ตัวเลือก คือ ข้อ 1, ข้อ 2, ข้อ 3, ข้อ 1 และข้อ 2, ข้อ 1 และข้อ 3, ข้อ 2 และข้อ 3, ข้อ 1 ข้อ 2 และข้อ 3 และไม่มีข้อใดถูก ในทำนองเดียวกันเมื่อกำหนดข้อความหรือสถานการณ์มา 4 หรือ 5 ข้อ จะสามารถเขียนตัวเลือกได้ 16 หรือ 32 ตัวเลือกตามลำดับ การที่สามารถสร้างตัวเลือกได้มากมายเช่นนี้ ทำให้ผู้สร้างข้อสอบสามารถเลือกตัวเลือกที่น่าสนใจใส่ไว้ในข้อสอบได้อย่างเหมาะสม แต่มีข้อควรพิจารณาในการเลือกตัวเลือกไว้ในข้อสอบ คือ ควรเป็นตัวเลือกที่ให้ข้อความหรือสถานการณ์ปรากฏในตัวเลือกทั้งหมดด้วยจำนวนครั้งที่เท่ากัน (Whitley, 1979 อ้างถึงใน อัญชิกา ดิษเจริญ 2535) ทั้งนี้เพราะข้อความหรือสถานการณ์ปรากฏในตัวเลือกหลายครั้ง จะมีผลต่อการเดาของผู้สอบในการสร้างแบบสอบวัดผลสัมฤทธิ์ ควรระวังถึงหลักเกณฑ์ข้อนี้

แต่ไม่ควรยึดเป็นกฎเกณฑ์ตายตัว เพราะต้องพิจารณาความเป็นไปได้ของเนื้อหาที่ออกข้อสอบด้วย มิฉะนั้นตัวเลือกบางตัวจะเป็นตัวเลือกที่ผิดชัดเจนและไม่มีผู้ใดเลือกตัวเลือกนั้น นอกจากนั้น ชาวาล แพร์ตกุล ยังแนะนำว่าไม่ควรใช้ตัวเลือกประเภทปลายเปิด “ไม่มีข้อใดถูก” และตัวเลือกที่มีความหมายขัดแย้งกันเองภายในตัวเลือกเดียวกัน เช่น “1 ถูกแต่ 2 ผิด” ในแบบสอบตัวเลือกข้อนี้ Weiten (1982 อ้างถึงใน อัญติกา ดิษเจริญ, 2535) กล่าวถึงข้อจำกัดของแบบสอบเลือกตอบตัวเลือกข้อนี้ไว้ 3 ประการ ดังนี้

1. การเขียนข้อสอบเลือกตอบแบบตัวเลือกข้อนี้ ไม่เป็นไปตามกฎเกณฑ์การเขียนข้อสอบเลือกตอบที่ว่าควรเขียนให้สั้นและกระชับ ซึ่งจะมีผลต่อการตอบถูกหรือผิดของผู้สอบด้วย
2. ความยาวและความซับซ้อนของข้อสอบ มีผลให้ผู้ที่มีทักษะการอ่านและการให้เหตุผลจะตอบแบบสอบตัวเลือกข้อนี้ได้ดีกว่า ซึ่งเป็นการเพิ่มความคลาดเคลื่อนในการประมาณค่าความสามารถที่แท้จริง
3. ความสามารถในการทำแบบสอบเลือกตอบตัวเลือกข้อนี้ ขึ้นอยู่กับความชำนาญในการทำแบบสอบ (Test-Wiseness) และทักษะในการทำแบบสอบ (Test Taking Skill) ของผู้สอบแต่ละคน

การเขียนข้อสอบตัวเลือกข้อที่กำหนดและไม่กำหนดสถานการณ์คำตอบ

ข้อสอบเลือกตอบแบบตัวเลือกข้อที่กำหนดสถานการณ์คำตอบ มีวิธีการเขียนโดยเขียนข้อคำถามและเขียนข้อความหรือสถานการณ์ เขียนตัวเลือกข้อสอบแบบสอบเลือกตอบแบบตัวเลือกข้อ โดยพิจารณาข้อความหรือสถานการณ์ที่น่าจะเป็นไปได้ ซึ่งถ้าเลือกแต่ละตัวอาจมี 1, 2, 3 หรือ 4 สถานการณ์แล้วแต่ความเหมาะสม

ข้อสอบเลือกตอบแบบตัวเลือกข้อที่ไม่กำหนดสถานการณ์คำตอบ มีวิธีการเขียนโดยการเขียนข้อคำถามเหมือนกับข้อสอบเลือกตอบแบบตัวเลือกข้อที่กำหนดสถานการณ์คำตอบแต่ไม่เขียนข้อความหรือสถานการณ์ให้ โดยนำสถานการณ์ไปใช้เป็นตัวเลือกเหมือนแบบสอบตัวเลือกธรรมดา แต่มีตัวเลือกข้ออยู่ ดังตัวอย่างข้อสอบต่อไปนี้

ข้อสอบของแบบสอบเลือกตอบแบบตัวเลือกข้อที่กำหนดสถานการณ์คำตอบ
(00) ผลกระทบที่เกิดขึ้นจากการเคลื่อนที่ของเปลือกโลกคือข้อใด

1. ทำให้เกิดภูเขา
2. ทำให้เกิดภูเขาไฟ
3. ทำให้เกิดแผ่นดินไหว
4. ทำให้อุณหภูมิลูกเพิ่มขึ้น

คำตอบที่ถูกต้องคือ

- ก. ข้อ 1 และข้อ 2
- ข. ข้อ 1 และข้อ 3
- ค. ข้อ 2 และข้อ 4
- ง. ข้อ 1,2 และข้อ 3

ข้อสอบของแบบสอบเลือกตอบแบบตัวเลือกข้อที่ไม่กำหนดสถานการณ์คำตอบ
(00) ผลกระทบที่เกิดขึ้นจากการเคลื่อนที่ของเปลือกโลกคือข้อใด

- ก. ทำให้เกิดภูเขาและภูเขาไฟ
- ข. ทำให้เกิดภูเขาและเกิดแผ่นดินไหว
- ค. ทำให้เกิดภูเขาไฟและอุณหภูมิลูกเพิ่มขึ้น
- ง. ทำให้เกิดภูเขา ภูเขาไฟ และแผ่นดินไหว

ตอนที่ 2 ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory)

ทฤษฎีการทดสอบแบบประเพณีนิยม (Classical Test Theory) เป็นทฤษฎีที่มีจุดอ่อนหลายประการ อาทิเช่น ค่าพารามิเตอร์ของข้อสอบไม่คงที่แปรเปลี่ยนไปตามกลุ่มตัวอย่างของผู้สอบ (Hambleton and Swaminathan, 1985) ซึ่งไม่เหมาะสมกับการทดสอบด้วยข้อสอบเฉพาะตัว เนื่องจากผู้สอบแต่ละคนจะได้รับข้อสอบต่างชุดกันและค่าความยากง่ายของข้อสอบต่างกัน ดังนั้น ทฤษฎีที่เหมาะสมและสามารถนำมาใช้กับการทดสอบด้วยข้อสอบเฉพาะตัวได้อย่างมีประสิทธิภาพ คือ ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) โดยมีหลักการที่จะเสนอในรายละเอียดต่อไปนี้

หลักการของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ เป็นทฤษฎีที่อธิบายถึงความสัมพันธ์ระหว่าง ลักษณะหรือความสามารถที่มีอยู่ภายในตัวบุคคลกับพฤติกรรมการตอบสนองข้อสอบของ บุคคลนั้น (Lord and Novick, 1968) โดยทฤษฎีนี้มีความเชื่อว่าพฤติกรรมการตอบสนองข้อ สอบของผู้เข้าสอบเป็นสิ่งที่สังเกตได้โดยตรงนั้น จะถูกกำหนดโดยคุณลักษณะ (Trait) หรือ ความสามารถ (Ability) ที่มีอยู่ในตัวบุคคลซึ่งไม่สามารถสังเกตได้โดยตรง แต่สามารถแสดง ความสัมพันธ์ดังกล่าวได้ด้วยโมเดลทางคณิตศาสตร์ ที่กำหนดด้วยฟังก์ชันทางคณิตศาสตร์ ซึ่งเรียกว่า ฟังก์ชันลักษณะข้อสอบ (Item Characteristic Function) หรือเรียกว่าโค้งคุณ ลักษณะข้อสอบ (Item Characteristic Curve) หรือฟังก์ชันการตอบสนองข้อสอบ (Item Response Function) (Lord and Novick, 1968) ซึ่งกำหนดได้หลายรูปแบบ แต่ทั้งนี้ถ้าจะใช้ รูปแบบของทฤษฎีนี้ได้ ต้องขึ้นอยู่กับว่าข้อมูลที่ได้จากการทดสอบ เป็นไปตามข้อตกลง เบื้องต้นหรือไม่

ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

การพิจารณานำทฤษฎีการตอบสนองข้อสอบไปใช้ จะต้องคำนึงถึงว่า แบบสอบ ที่จะนำมาใช้เป็นไปตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ ดังนี้

1. แบบสอบมิติเดียว (Unidimension Tests) หมายถึง ข้อสอบแต่ละข้อใน แบบสอบจะต้องวัดความสามารถหรือคุณลักษณะเดียวกันหรือมีความเป็นเอกพันธ์กัน การ ตรวจสอบความเป็นมิติเดียวของแบบสอบส่วนใหญ่ใช้วิธีการวิเคราะห์ตัวประกอบ
2. ความเป็นอิสระในการตอบข้อสอบ (Local Independence) หมายถึง การ ตอบสนองต่อข้อสอบข้อหนึ่ง ๆ ในแบบสอบของผู้สอบคนใดคนหนึ่งมีความเป็นอิสระในทาง สถิตินั้นคือ การตอบข้อสอบข้อใดข้อหนึ่งถูกจะไม่มีผลต่อการตอบข้อสอบข้ออื่น ๆ ซึ่งเป็น ข้อตกลงเบื้องต้นที่สอดคล้องกับข้อตกลงเบื้องต้นของคุณสมบัติความเป็นแบบสอบมิติเดียว นั่นคือ ถ้ามีคุณสมบัติความเป็นแบบสอบมิติเดียว ก็จะมีความเป็นอิสระในการตอบข้อสอบ และในทางกลับกัน ถ้ามีความเป็นอิสระในการตอบข้อสอบก็จะมีความเป็นแบบสอบมิติ เดียว

3. โค้งลักษณะข้อสอบ (Item Characteristic Curve) เป็นฟังก์ชันทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างความน่าจะเป็นในการตอบข้อสอบข้อหนึ่งได้ถูกต้องกับระดับความสามารถที่วัดโดยข้อสอบข้อนั้น ซึ่งเป็นฟังก์ชันการถดถอยชนิดไม่เป็นเชิงเส้นตรง (Non-Linear Regression) ของคะแนนข้อสอบบนระดับความสามารถที่วัด โดยแบบสอบฉบับนั้น

4. ทฤษฎีการตอบสนองข้อสอบ อธิบายว่าการที่ผู้สอบจะมีโอกาสที่จะตอบถูกหรือผิดขึ้นอยู่กับความสามารถของผู้สอบเท่านั้น ดังนั้นข้อสอบที่มีลักษณะเป็นข้อสอบวัดความเร็วจะไม่สามารถประยุกต์ใช้กับทฤษฎีการตอบสนองข้อสอบนี้ได้ เพราะข้อสอบที่ผู้สอบไม่ได้ทำ ซึ่งจะได้คะแนนเป็นศูนย์ อาจเป็นเพราะไม่มีเวลาทำ ไม่ใช่เพราะทำแล้วผิด

โมเดลในทฤษฎีการตอบสนองข้อสอบ

โมเดลในทฤษฎีการตอบสนองข้อสอบ เป็นโมเดลทางคณิตศาสตร์ที่อธิบายโค้งลักษณะข้อสอบ ซึ่งมีหลายโมเดลด้วยกัน แต่ที่ได้รับการพัฒนาและมีการนำมาใช้มากที่สุด ในสถานการณ์การทดสอบ คือ โมเดลโลจิสติก (Logistic Model) เนื่องจากการนำไปใช้และการคำนวณสะดวกกว่าโมเดลอื่น โมเดลโลจิสติกนี้ถูกพัฒนาขึ้นมา 4 รูปแบบ คือ โมเดลโลจิสติกหนึ่งพารามิเตอร์ สองพารามิเตอร์ สามพารามิเตอร์ และสี่พารามิเตอร์ ดังรายละเอียดต่อไปนี้

1. โมเดลโลจิสติกหนึ่งพารามิเตอร์ (One-Parameter Logistic Model)

โมเดลนี้เบอร์นบอม (Birnbaum) ได้พัฒนาขึ้นในปี 1986 ซึ่งตรงกับโมเดลของราล์ฟ (Rasch Model) ที่พัฒนาขึ้นในปี 1960 เป็นโมเดลที่อธิบายข้อสอบด้วยพารามิเตอร์ค่าความยากเพียงค่าเดียว โดยมีข้อตกลงเบื้องต้นว่า ค่าอำนาจจำแนกของข้อสอบทุกข้อมีค่าเท่ากัน และค่าการเดาของข้อสอบ มีค่าเท่ากับ 0 มีฟังก์ชันทางคณิตศาสตร์ดังนี้

$$P_i(e) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} \quad I = 1, 2, 3, \dots$$

เมื่อ $P_i(\theta)$ คือ โอกาสที่ผู้มีความสามารถจะทำข้อสอบข้อที่ 1 ได้ถูกต้อง

θ คือ ระดับความสามารถที่แท้จริงของผู้สอบ

b_i คือ ค่าความยากของข้อสอบข้อที่ i

e คือ ค่าคงที่มีค่าเท่ากับ 2.7182818

2. โมเดลโลจิสติกสองพารามิเตอร์ (Two-Parameter Logistic Model)

โมเดลนี้อธิบายข้อสอบด้วยสองพารามิเตอร์ คือ ค่าอำนาจจำแนกของข้อสอบและค่าความยากของข้อสอบ โดยกำหนดให้ค่าการเดาของข้อสอบมีค่าเท่ากับ 0 มีฟังก์ชันทางคณิตศาสตร์ ดังนี้

$$P_i(e) = \frac{e^{D_i a_i (\theta - b_i)}}{1 + e^{D_i a_i (\theta - b_i)}} \quad i = 1, 2, 3, \dots, n$$

เมื่อ $P_i(\theta)$ คือ โอกาสที่ผู้มีความสามารถ(θ) จะทำข้อสอบข้อที่ i ได้ถูกต้อง

θ คือ ระดับความสามารถที่แท้จริงของผู้สอบ

a_i คือ ค่าอำนาจจำแนกของข้อสอบที่ i

b_i คือ ค่าความยากของข้อสอบข้อที่ i

e คือ ค่าคงที่มีค่าเท่ากับ 2.7182818

D คือ ตัวประกอบของสเกล มีค่าเท่ากับ 1.7

3. โมเดลโลจิสติกสามพารามิเตอร์ (Three-Parameter Logistic Model)

โมเดลนี้พัฒนาต่อจากโมเดลโลจิสติกสองพารามิเตอร์ เพื่อใช้กับแบบสอบที่มีอิทธิพลจากการเดาแฝงอยู่ด้วย มีฟังก์ชันทางคณิตศาสตร์ ดังนี้

$$P_i(e) = C_i + (1 - C_i) \frac{e^{D_i a_i (\theta - b_i)}}{1 + e^{D_i a_i (\theta - b_i)}} \quad i = 1, 2, 3, \dots, n$$

- เมื่อ $P_i(\theta)$ คือ โอกาสที่ผู้มีความสามารถ (θ) จะทำข้อสอบข้อที่ i ได้ถูกต้อง
- θ คือ ระดับความสามารถที่แท้จริงของผู้สอบ
- a_i คือ ค่าอำนาจจำแนกของข้อสอบที่ i
- b_i คือ ค่าความยากของข้อสอบข้อที่ i
- c_i คือ ค่าการเดาของข้อสอบข้อที่ i
- e คือ ค่าคงที่มีค่าเท่ากับ 2.7182818
- D คือ ตัวประกอบของสเกล มีค่าเท่ากับ 1.7

4. โมเดลโลจิสติกสี่พารามิเตอร์ (Four-Parameter Logistic Model)

โมเดลนี้แตกต่างจากโมเดลโลจิสติกสามพารามิเตอร์ โดยการเพิ่มค่าพารามิเตอร์ γ_i ซึ่งมีค่าต่ำกว่า 1 เล็กน้อย หมายถึง ค่าที่ผู้สอบที่มีความสามารถสูงมีความสะเพร่าตอบข้อสอบไม่ถูก ในทางปฏิบัติโมเดลนี้ยังไม่สามารถนำมาใช้ได้ มีฟังก์ชันทางคณิตศาสตร์ ดังนี้

$$P_i(e) = C_i + (Y - c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad i = 1, 2, 3, \dots, n$$

- เมื่อ γ_i คือ ค่าตัวแปรแฝงในข้อกระทง
- a_i คือ ค่าพารามิเตอร์อำนาจจำแนกข้อ i
- b_i คือ ค่าพารามิเตอร์ความยากของข้อ i
- c_i คือ ค่าพารามิเตอร์ของโอกาสการเดา
- e คือ ค่าคงที่มีค่า 2.7183
- D คือ ค่าคงที่มีค่า 1.702

การตรวจสอบคุณภาพแบบสอบ

การตรวจสอบคุณภาพของแบบสอบ ตามแนวคิดของทฤษฎีการตอบสนองข้อสอบ จะพิจารณาจากค่าฟังก์ชันสารสนเทศของแบบสอบ (Test Information Function) ซึ่งจะเป็นดัชนีที่บอกความแม่นยำในการประมาณค่าความสามารถที่แท้จริง (Lord, 1980) ซึ่งได้จากผลรวมของฟังก์ชันสารสนเทศข้อสอบ (Item Information Function) โดยที่ค่าฟังก์ชันสารสนเทศแบบสอบเป็นดัชนีผสม (Composite Index) ที่สร้างจากดัชนีบอกคุณลักษณะของข้อสอบหลายลักษณะ เช่น ค่าอำนาจจำแนก ค่าความยาก เป็นต้น มารวมเป็นดัชนีเพียงตัวเดียว เพื่อชี้ถึงคุณภาพของแบบสอบได้ จะเห็นว่าในทฤษฎีการตอบสนองข้อสอบ จะใช้ค่าฟังก์ชันสารสนเทศของแบบสอบแทนการหาความเที่ยงของคะแนนและความคลาดเคลื่อนมาตรฐานในการวัดตามแนวคิดของทฤษฎีการทดสอบแบบประเพณีนิยม (Hambleton, 1977) สำหรับโมเดลโลจิสติก 3 พารามิเตอร์ จะสามารถหาค่าฟังก์ชันสารสนเทศของข้อสอบในแต่ละระดับความสามารถ (θ) ได้จากสูตร (Lord, 1980)

$$I(\theta, U_i) = \frac{(1.7 a_i)^2 (1 - c_i)}{1.7a (\theta - b_i) - 1.7a (\theta - b_i)}$$

และสามารถกำหนดโค้งสารสนเทศของข้อสอบได้ จากสมการ

$$\bar{I}(\theta - U_i) = (p - p_i)^2 / P_i Q_i$$

เมื่อ $I(\theta, U_i)$ คือ ค่าฟังก์ชันสารสนเทศข้อสอบ

a_i คือ ความชันของ ICC ที่ระดับความสามารถ θ

P_i คือ ความน่าจะเป็นที่ผู้สอบมีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

Q_i คือ $1 - P_i$

สามารถหาค่าฟังก์ชันสารสนเทศของแบบสอบได้จากผลรวมของฟังก์ชันสารสนเทศของข้อสอบทั้งหมดในแบบสอบ เขียนสมการได้ดังนี้

$$I(\theta) = \sum I(\theta, U_i)$$

เมื่อ $I(\theta)$ คือ ค่าฟังก์ชันสารสนเทศแบบสอบ

การสอบทุกครั้งย่อมมีความคลาดเคลื่อนเกิดขึ้น ความคลาดเคลื่อนมาตรฐานในการประมาณค่า (Standard Error of Estimate : SEE) คือ ส่วนเบี่ยงเบนมาตรฐานในการประมาณค่าความสามารถ (θ) ซึ่งมีความสัมพันธ์กับค่าฟังก์ชันสารสนเทศแบบสอบ ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ คำนวณได้จากสมการ

$$SEE = \frac{1}{\sqrt{I(\theta)}}$$

จากสมการ พบว่า ค่า SEE มีค่าเท่ากับรากที่สองของส่วนกลับของฟังก์ชันสารสนเทศของแบบสอบ ดังนั้น ถ้าค่าฟังก์ชันสารสนเทศยิ่งมาก ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสมารถยิ่งน้อยลง ในการสอบทุกครั้งเราต้องการให้ค่าความคลาดเคลื่อนในการประมาณค่าความสมารถ มีค่าน้อยที่สุด

การนำเอาทฤษฎีการตอบสนองข้อสอบมาประยุกต์ใช้

คุณสมบัติที่สำคัญของทฤษฎีการตอบสนองข้อสอบในเรื่องค่าพารามิเตอร์ของข้อสอบไม่ขึ้นอยู่กับความสามารถของผู้สอบ ความสามารถของผู้สอบก็ไม่ขึ้นอยู่กับความยากหรือง่ายของข้อสอบ และค่าพารามิเตอร์ไม่เปลี่ยนแปลงไปตามตำแหน่งของข้อสอบ คุณลักษณะดังกล่าว คือ ความเป็นปรนัยของการวัดผล (Objectivity of Measurement) ที่มีความสำคัญมากในทางวัดผล (Write and Stones อ้างถึงใน สุพัฒน์ สุขมลสันต์, 2528) จึงสามารถนำเอาคุณสมบัติเด่นเหล่านี้ มาประยุกต์ใช้ในการวัดผลการศึกษาได้อย่างมีประสิทธิภาพในหลาย ๆ เรื่อง ดังนี้

1. การทดสอบด้วยข้อสอบเฉพาะตัว (Tailored Testion) หมายถึง การทดสอบที่มีการจัดข้อสอบให้มีความยากที่เหมาะสมกับระดับความสามารถของผู้สอบ นับเป็นการนำเอาทฤษฎีการตอบสนองข้อสอบ มาประยุกต์ใช้อย่างมีประสิทธิภาพ โดยการนำเอาคุณสมบัติที่ไม่แปรเปลี่ยนของค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบมาใช้ทดสอบด้วยข้อสอบเฉพาะตัวนั้น ผู้สอบแต่ละคนจะได้รับชุดของข้อสอบที่แตกต่างกันตามระดับความสามารถของผู้สอบ แต่ก็สามารถนำเอาความสามารถที่ประมาณค่าได้จากการทดสอบนี้มาเปรียบเทียบกันได้ ทั้งนี้เพราะความสามารถที่ประมาณค่าได้ยังอยู่ในมาตราวัดเดียวกัน (Common Ability)

2. การสร้างคลังข้อสอบ (Item Bank) การนำเอาทฤษฎีการตอบสนองข้อสอบมาใช้ในการวิเคราะห์และสร้างคลังข้อสอบ ทำให้ข้อสอบแต่ละข้อมีค่าสถิติต่าง ๆ เป็นค่าประจำข้อ ไม่ขึ้นอยู่กับข้ออื่นและไม่ขึ้นอยู่กับการจัดตัวอย่าง จึงทำให้เกิดความสะดวกในการคัดเลือกข้อสอบมาใช้และสามารถจัดชุดข้อสอบให้ได้คุณภาพตามเกณฑ์ที่ต้องการ

3. การหาความลำเอียงของข้อสอบ (Item Bias) สามารถใช้โค้งลักษณะข้อสอบมาตรวจสอบความลำเอียงของข้อคำถาม โดยปกติถ้าข้อคำถามนั้นไม่มีความลำเอียงโค้งลักษณะข้อสอบไม่ว่าจะได้มาจากผู้สอบกลุ่มใดก็ตามจะมีรูปร่างเหมือนกัน แต่ถ้าข้อคำถามมีความลำเอียง โค้งลักษณะข้อสอบจะมีรูปร่างแตกต่างกัน

4. การเทียบมาตรา (Equating) หมายถึง กระบวนการทางสถิติเพื่อปรับคะแนนที่ได้จากการทดสอบด้วยข้อสอบที่ต่างชุดกันแต่มีโครงสร้างเดียวกัน ให้สามารถเทียบกันได้ แต่ถ้าการทดสอบนั้นวัดความสามารถต่างกันแล้วจะนำมาเทียบกันไม่ได้ เนื่องจากความสามารถของผู้สอบที่ประมาณค่ามาจากทฤษฎีการตอบสนองข้อสอบ จะมีหน่วยการวัดเดียวกันและไม่ขึ้นอยู่กับการจัดชุดของแบบสอบ จึงสามารถนำมาเทียบกันได้

ตอนที่ 3 งานวิจัยที่เกี่ยวข้อง

การเปรียบเทียบคุณภาพของแบบสอบเลือกตอบตัวเลือกรวมดากับแบบสอบเลือกตอบตัวเลือกซ้อน ได้มีผู้ศึกษา ดังต่อไปนี้

Weiten (1982) ได้ศึกษาถึง ประสิทธิภาพสัมพัทธ์ของแบบสอบเลือกตอบแบบตัวเลือกรวมดากับแบบสอบเลือกตอบตัวเลือกซ้อน กลุ่มตัวอย่างเป็นนักศึกษาปริญญาตรีที่ลงทะเบียนเรียนวิชาจิตวิทยาทั่วไป จำนวน 47 คน เครื่องมือที่ใช้เป็นแบบสอบเลือกตอบแบบตัวเลือกซ้อน โดยดัดแปลงมาจากข้อสอบแบบตัวเลือกรวมดา ในแบบสอบแต่ละฉบับมีข้อสอบฉบับละ 40 ข้อ จัดพิมพ์แบบสอบเป็น 2 พอร์ม ดังนี้

พอร์ม A เป็นแบบสอบที่ประกอบด้วย แบบสอบย่อยที่มีข้อสอบแบบตัวเลือกรวมดา 20 ข้อ และแบบสอบย่อยที่มีข้อสอบแบบตัวเลือกซ้อน 20 ข้อ

พอร์ม B เป็นแบบสอบที่ประกอบด้วย แบบสอบย่อยที่มีข้อสอบแบบตัวเลือกซ้อน 20 ข้อ และแบบสอบย่อยที่มีข้อสอบแบบตัวเลือกรวมดา 20 ข้อ

ผลการวิจัย พบว่า แบบสอบเลือกตอบแบบตัวเลือกซ้อนยากกว่าแบบสอบเลือกตอบตัวเลือกรวมดา ค่าอำนาจจำแนกของแบบสอบตัวเลือกรวมดาสูงกว่าแบบสอบเลือกตอบตัวเลือกซ้อน การเปรียบเทียบค่าความเที่ยงได้เปรียบเทียบ 2 คู่ จาก 2 แบบสอบย่อย พบว่า ในคู่แรกแบบสอบย่อยที่ข้อสอบเป็นแบบตัวเลือกรวมดามีค่าความเที่ยงสูงกว่าแบบสอบย่อยที่ข้อสอบเป็นแบบตัวเลือกซ้อน ส่วนการเปรียบเทียบคู่สอง พบว่า แบบสอบย่อยที่มีตัวเลือกทั้ง 2 แบบมีค่าความเที่ยงไม่แตกต่างกัน สำหรับด้านความตรงได้เปรียบเทียบ 2 คู่เช่นเดียวกัน ได้ผลว่า แบบสอบที่มีตัวเลือกทั้ง 2 แบบ มีค่าความตรงไม่แตกต่างกันทั้ง 2 คู่

ปราณี ร่วมทอง (2528) ได้ศึกษาเปรียบเทียบคุณภาพของแบบสอบเลือกตอบตัวเลือกรวมดากับแบบตัวเลือกซ้อนที่ใช้วัดระดับพุทธิพิสัยต่างกัน เมื่อใช้ทดสอบกับนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนต่างกัน คือ กลุ่มนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนสูง ปานกลาง และต่ำ ในด้านความเที่ยง ความตรง ความยากและอำนาจจำแนกของแบบสอบ นอกจากนั้นยังได้เปรียบเทียบคะแนนที่ได้จากการตอบแบบสอบเลือกตอบตัวเลือกรวมดา

กับคะแนนที่ได้จากการตอบแบบสอบถามตัวเลือกซ้อน วิชาวิทยาศาสตร์ ชั้นมัธยมศึกษาปีที่ 3 (ว 305) ซึ่งมีรูปแบบของข้อสอบเป็นแบบสอบเลือกตอบตัวเลือกธรรมดาและแบบสอบเลือกตอบตัวเลือกซ้อน ที่ใช้วัดความรู้ 2 ระดับ คือ ระดับความรู้ขั้นต่ำ เป็นข้อสอบวัดการเรียนรู้ในระดับความรู้-ความจำ ความเข้าใจ การนำไปใช้ และระดับความรู้ขั้นสูง เป็นข้อสอบวัดการเรียนรู้ในระดับ การวิเคราะห์ การสังเคราะห์ และการประเมินผล กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคต้น ปีการศึกษา 2527 ในจังหวัดนครนายก จำนวน 328 คน ซึ่งแบ่งเป็น 2 กลุ่ม ๆ ละ 164 คน ผลการวิจัยพบว่า

1. การวัดระดับความรู้ขั้นต่ำ แบบสอบเลือกตอบตัวเลือกธรรมดา มีค่าความเที่ยงและค่าอำนาจจำแนกสูงกว่าแบบสอบเลือกตอบตัวเลือกซ้อน ด้านความตรงแบบสอบเลือกตอบทั้ง 2 ฉบับมีค่าความตรงไม่แตกต่างกัน และพบว่าแบบสอบเลือกตอบตัวเลือกธรรมดาง่ายกว่าแบบสอบเลือกตอบตัวเลือกซ้อน
2. การวัดระดับความรู้ขั้นสูง แบบสอบเลือกตอบตัวเลือกธรรมดาและแบบสอบเลือกตอบตัวเลือกซ้อนมีค่าความเที่ยง ความตรง ความยากและค่าอำนาจจำแนกไม่แตกต่างกัน
3. นักเรียนกลุ่มที่มีผลสัมฤทธิ์ทางการเรียนสูงและกลุ่มที่มีผลสัมฤทธิ์ทางการเรียนต่ำ ตอบแบบสอบเลือกตอบทั้ง 2 ฉบับได้ไม่แตกต่างกัน
4. นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนปานกลาง ตอบแบบสอบเลือกตอบตัวเลือกธรรมดา ได้คะแนนเฉลี่ยสูงกว่าแบบสอบเลือกตอบตัวเลือกซ้อน

วิธีสัญญา ปราบบัญญัติ (2529) ได้เปรียบเทียบคุณภาพของแบบสอบเลือกตอบตัวเลือกเดียวกับแบบสอบเลือกตอบตัวเลือกซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกเท่ากันและที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกไม่เท่ากัน ในด้านความยาก อำนาจจำแนก ความเที่ยงและเปอร์เซ็นต์การเดา แบบสอบที่ใช้วิจัย คือ แบบสอบเลือกตอบวิชาวิทยาศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 จำนวน 3 ฉบับ ๆ ละ 30 ข้อ แบบสอบฉบับที่ 1 เป็นแบบสอบเลือกตอบคำตอบเดียว ฉบับที่ 2 เป็นแบบเลือกตอบคำตอบซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกเท่ากันกับ ฉบับที่ 3 เป็นแบบสอบเลือกตอบคำตอบซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกไม่เท่ากัน กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษา

ปีที่ 2 ปีการศึกษา 2528 สังกัดกรมสามัญศึกษา จังหวัดพัทลุง จำนวน 636 ผลการวิจัยพบว่า แบบสอบเลือกตอบคำตอบข้อที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกไม่เท่ากันยากที่สุด รองลงมา ได้แก่ แบบสอบเลือกตอบคำตอบข้อที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกเท่ากัน และแบบสอบเลือกตอบคำตอบเดียวง่ายที่สุด ด้านค่าอำนาจจำแนก พบว่าแบบสอบทั้ง 3 ฉบับ มีค่าอำนาจจำแนกไม่แตกต่างกัน ส่วนด้านความเที่ยง พบว่า แบบสอบเลือกตอบคำตอบข้อที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกไม่เท่ากันมีความเที่ยงสูงกว่าแบบสอบเลือกคำตอบเดียวและแบบสอบเลือกตอบคำตอบข้อที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกเท่ากัน แต่ค่าความเที่ยงของแบบสอบเลือกตอบคำตอบเดียวกับแบบสอบเลือกตอบคำตอบข้อที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกเท่ากัน มีค่าไม่แตกต่างกัน สำหรับเปอร์เซ็นต์การเดานั้นพบว่า แบบสอบทั้ง 3 ฉบับ มีเปอร์เซ็นต์การเดาไม่แตกต่างกัน

กมลทิพย์ อธิการยานันท์ (2530) ได้ศึกษาเปรียบเทียบคุณภาพของแบบสอบเลือกตอบคำตอบถูกตัวเดียว คำตอบถูกไม่จำกัดและคำตอบรวมหรือแบบตัวเลือกข้อ ในด้านความเที่ยง ความตรง และปริมาณการเดา นอกจากนั้นยังเปรียบเทียบคะแนนที่ได้จากแบบสอบทั้งสามฉบับ ในกลุ่มนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนสูง ปานกลาง และต่ำ แบบสอบที่ใช้ในการวิจัย คือ แบบสอบวิชาวิทยาศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 (ว 203) จำนวน 3 ฉบับ ฉบับที่ 1 เป็นแบบสอบเลือกตอบคำตอบถูกตัวเดียว ฉบับที่ 2 เป็นแบบสอบเลือกตอบคำตอบถูกไม่จำกัดและฉบับที่ 3 เป็นแบบสอบเลือกตอบคำตอบรวมหรือแบบตัวเลือกข้อ มีจำนวนข้อ ฉบับละ 50 ข้อ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคต้น ปีการศึกษา 2529 ของโรงเรียนในสังกัดกรมสามัญศึกษา กรุงเทพมหานคร จำนวน 450 คน ผลการวิจัย พบว่า

ค่าความเที่ยงของแบบสอบเลือกตอบคำตอบถูกไม่จำกัดมีค่าสูงกว่าแบบสอบเลือกตอบคำตอบถูกตัวเดียว แต่ค่าความเที่ยงของสอบเลือกตอบคำตอบถูกไม่จำกัดกับแบบคำตอบรวมและค่าความเที่ยงของแบบสอบเลือกตอบคำตอบถูกตัวเดียวกับแบบสอบเลือกตอบคำตอบรวม มีค่าไม่แตกต่างกัน ด้านความตรงแบบสอบทั้ง 3 ฉบับมีค่า

ความตรงไม่แตกต่างกัน ส่วนด้านปริมาณการเดานั้นนักเรียนตอบแบบสอบทั้ง 3 ฉบับ ด้วยปริมาณการเดาที่ต่างกัน คือ ตอบแบบสอบเลือกตอบคำตอบถูกไม่จำกัดด้วยปริมาณการเดาต่ำสุด รองลงมา ได้แก่ แบบสอบเลือกตอบคำตอบรวมและแบบสอบเลือกตอบถูกตัวเดียวมีปริมาณการเดาสูงสุด

คะแนนเฉลี่ยของนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนสูง ปานกลาง ต่ำ ที่ตอบแบบสอบทั้ง 3 ฉบับ มีลักษณะที่เหมือนกัน คือ นักเรียนตอบแบบสอบเลือกตอบคำตอบถูกตัวเดียวได้คะแนนเฉลี่ยสูงสุด รองลงมาคือ แบบสอบเลือกตอบคำตอบรวมและแบบสอบเลือกตอบคำตอบถูกไม่จำกัดได้คะแนนเฉลี่ยต่ำสุดและไม่มีปฏิกริยาร่วมระหว่างลักษณะของแบบสอบกับระดับผลสัมฤทธิ์ทางการเรียนของนักเรียน

อัญชิกา ดิษเจริญ (2535) ได้ศึกษาเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบสอบเลือกตอบต่างแบบของแบบเลือกตอบ 3 แบบ คือ แบบสอบตัวเลือกธรรมดา แบบสอบตัวเลือกผสม และแบบสอบตัวเลือกซ้อน แบบสอบที่ใช้วิจัย ได้แก่ แบบสอบวิชาสร้างเสริมประสบการณ์ชีวิต ชั้นประถมศึกษาปีที่ 6 จำนวน 3 ฉบับ ฉบับละ 19 ข้อ โดยที่ฉบับที่ 1 เป็นแบบสอบตัวเลือกธรรมดา ฉบับที่ 2 เป็นแบบสอบตัวเลือกผสมและฉบับที่ 3 เป็นแบบสอบตัวเลือกซ้อน กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 สังกัดสำนักงานการประถมศึกษาอำเภอหันคา จังหวัดชัยนาท จำนวน 3,030 คน ผลการวิจัย พบว่า

ค่าฟังก์ชันสารสนเทศของข้อสอบในแบบสอบตัวเลือกธรรมดา มีค่าสารสนเทศสูงกว่าข้อสอบในแบบตัวเลือกผสมที่ระดับความสามารถปานกลางลงไป ($\theta < 0.5$) ค่าฟังก์ชันสารสนเทศของข้อสอบในแบบตัวเลือกซ้อนมีค่าฟังก์ชันสารสนเทศสูงกว่าข้อสอบในแบบสอบตัวเลือกธรรมดาที่ระดับความสามารถค่อนข้างสูง ($\theta < 0.5$) ข้อสอบในแบบสอบตัวเลือกผสมมีค่าสารสนเทศสูงกว่าข้อสอบในแบบสอบตัวเลือกซ้อนที่ระดับความสามารถปานกลางค่อนข้างสูง ($\theta = 0$ ถึง 1.5) และข้อสอบในแบบสอบตัวเลือกซ้อนมีค่าสารสนเทศสูงกว่าข้อสอบในแบบสอบตัวเลือกผสมที่ระดับความสามารถสูงขึ้นไป ($\theta > 1.5$) ข้อสอบในแบบสอบตัวเลือกธรรมดามีค่าสารสนเทศสูงกว่าข้อสอบในแบบสอบตัวเลือกผสม และข้อสอบในแบบสอบตัวเลือกซ้อนที่ระดับความสามารถปานกลางลงไป ($\theta < 1$) ส่วนที่ระดับความสามารถสูงขึ้นไป ($\theta > 1.5$) ข้อสอบในแบบสอบตัวเลือกซ้อนมีค่าสารสนเทศสูงกว่าข้อสอบในแบบสอบ

ตัวเลือกธรรมดาและข้อสอบในแบบสอบตัวเลือกผสม ค่าฟังก์ชันสารสนเทศของแบบสอบตัวเลือกธรรมดามีค่าสารสนเทศสูงกว่าแบบสอบตัวเลือกผสมและแบบสอบตัวเลือกซ้อนที่ระดับความสามารถต่ำถึงปานกลาง ($\theta = -2.0$ ถึง 0) และแบบสอบตัวเลือกซ้อนมีค่าสารสนเทศสูงกว่าแบบสอบตัวเลือกธรรมดาและแบบสอบตัวเลือกผสมที่ระดับความสามารถสูงขึ้นไป

จากการศึกษางานวิจัยเกี่ยวกับการเปรียบเทียบคุณภาพของแบบสอบเลือกตอบตัวเลือกธรรมดากับแบบสอบเลือกตอบตัวเลือกซ้อน ส่วนใหญ่ศึกษาโดยใช้ทฤษฎีการวัดแบบคลาสสิกอล ผลการวิจัยเกี่ยวกับคุณภาพของแบบสอบบางด้านยังขัดแย้งกันอยู่ เช่น ด้านความเที่ยงของแบบสอบ แต่โดยสรุปแล้วมีแนวโน้มว่า แบบสอบตัวเลือกซ้อนจะยากกว่าแบบสอบตัวเลือกธรรมดา ส่วนด้านความตรง พบว่า แบบสอบทั้งสองชนิดมีความตรงไม่แตกต่างกัน และมีแนวโน้มว่านักเรียนจะตอบแบบสอบตัวเลือกซ้อนด้วยเปอร์เซ็นต์การเดาที่ต่ำกว่า ดังนั้น ผู้วิจัยจึงสนใจที่จะเปรียบเทียบคุณภาพของแบบสอบ ระหว่างแบบสอบแบบตัวเลือกซ้อนที่กำหนดสถานการณ์คำตอบกับแบบสอบแบบตัวเลือกซ้อนที่ไม่กำหนดสถานการณ์คำตอบ โดยใช้ทฤษฎีการตอบสนองข้อสอบ (IRT)