



บทที่ 2

วรรณคดีที่เกี่ยวข้อง

งานวิจัยเรื่อง ผลของจำนวนตัวเลือกที่มีต่อ ค่าฟังก์ชันสารสนเทศของแบบสอบเลือกตอบแบบถูกผิด ในระดับการศึกษาที่ต่างกัน ผู้วิจัยได้ศึกษาเอกสาร หนังสือ และงานวิจัยที่เกี่ยวข้องทั้งหมดแบ่งเป็น 4 ตอน ดังต่อไปนี้

ตอนที่ 1 แนวทางในการปรับแก้ข้อบกพร่องของแบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว (multiple-choice)

ตอนที่ 2 มโนทัศน์เกี่ยวกับแบบสอบเลือกตอบแบบถูกผิด (multiple true-false)

ตอนที่ 3 มโนทัศน์เกี่ยวกับทฤษฎีการตอบสนองข้อสอบ (IRT)

ตอนที่ 4 งานวิจัยที่เกี่ยวข้อง

ตอนที่ 1 แนวทางในการปรับแก้ข้อบกพร่องของแบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว (multiple-choice)

แบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว เป็นแบบสอบปรนัยชนิดหนึ่ง ซึ่งเป็นที่นิยมมากทั้งในแบบสอบมาตรฐาน และแบบสอบวัดผลสัมฤทธิ์ที่ใช้ในการทดสอบตามสถานศึกษาต่าง ๆ ทั้งนี้เพราะมีจุดเด่นหลายประการ ได้แก่

1. สามารถวัดวัดจุดประสงค์การเรียนรู้ทางด้านพุทธิพิสัย (cognitive domain) ได้ทุกระดับตั้งแต่ระดับความรู้- ความจำ ไปจนถึงระดับการประเมินค่า
2. สามารถใช้ได้กับเนื้อหาวิชาทุกสาขาและใช้ได้กับนักเรียนทุกระดับชั้น
3. การตรวจให้คะแนนทำได้อย่างรวดเร็วถูกต้องและเป็นปรนัยแม้จะให้ผู้อื่นหรือเครื่องจักรเป็นผู้ตรวจก็ตาม
4. เมื่อเปรียบเทียบกับข้อสอบแบบถูกผิด (true-false) แบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว จะมีความแปรปรวนของคะแนน อันเนื่องมาจากการเดาน้อยกว่า เพราะว่าความน่าจะเป็นในการเดาถูกนั้นขึ้นอยู่กับจำนวนตัวเลือก
5. สามารถให้สารสนเทศสำหรับการวินิจฉัยนักเรียนได้ โดยเฉพาะอย่างยิ่งถ้าคำตอบที่ให้เลือกนั้นมีระดับความถูกต้องต่างกัน

แม้ว่าแบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว จะมีจุดเด่นอยู่หลายประการดังที่กล่าวมาแล้ว แต่ก็มีจุดอ่อนที่ทำให้ผลการวัด โดยใช้แบบสอบชนิดนี้มีความถูกต้องน้อยกว่าที่ควรจะเป็น ซึ่งจุดอ่อนดังกล่าวมี 3 ประการ คือ

1. ปัญหาเกี่ยวกับการมีคำตอบที่ถูกต้องได้หลายคำตอบ ซึ่งแบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว ไม่สามารถที่จะใช้วัด ได้ในกรณีที่ต้องการทราบว่าคุณสมบัติที่ถูกต้องทั้งหมด หรือไม่
2. การมีตัวเลือกที่ถูกต้องเพียงตัวเลือกเดียวอาจทำให้ผู้สอบสามารถตอบข้อสอบได้ถูกต้อง โดยไม่ต้องมีความรู้ในเรื่องนั้นอย่างครบถ้วนสมบูรณ์
3. การตรวจให้คะแนนมีความไม่ละเอียดถี่ถ้วน พอลที่จะแยกผู้ที่มีความรู้ถูกต้องสมบูรณ์ ครบถ้วน กับผู้ที่มีความรู้บางส่วน และผู้ที่ไม่มีความรู้ออกจากกันได้

จากจุดอ่อนดังกล่าว นักวัดผลได้ปรับปรุงแก้ไขแบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว ให้สามารถวัดได้อย่างมีประสิทธิภาพมากขึ้น ด้วยวิธีการต่าง ๆ หลายวิธี ซึ่งพอสรุปได้เป็น 3 แนวทาง คือ

1. **ปรับวิธีการตรวจให้คะแนน** วิธีการมีรูปแบบและวิธีการตอบแบบสอบจะยังคงเหมือนเดิม คือ ให้มีตัวเลือกที่เป็นคำตอบที่ถูกเพียงตัวเลือกเดียว และเลือกตอบเฉพาะตัวเลือกที่ถูกเพียงตัวเลือกเดียว แต่มีการเปลี่ยนแปลงวิธีการตรวจให้คะแนน ซึ่งมีอยู่ 2 วิธี คือ

1.1 **การให้คะแนนตามสูตร (formula scoring)** สูตรสำหรับการให้คะแนนมีอยู่หลายรูปแบบ แต่สูตรที่นิยมใช้กันมากคือ สูตรปรับแก้การเดา ซึ่งจากผลการวิจัยพบว่าการใช้สูตรปรับแก้การเดาไม่ได้ทำให้ความเที่ยง หรือความตรง ของผลการสอบแตกต่างไปจากการไม่ได้ปรับแก้การเดาเลย (เพ็ญศรี สว่างเนตร, 2520) นั่นแสดงว่า การใช้สูตรปรับแก้การเดาอย่างง่ายไม่ได้ช่วยแก้ปัญหาเรื่องการเดานี้แต่อย่างใด และดูจะเป็นการลงโทษผู้สอบที่เดาแล้วผิดเสียมากกว่า (Ebel, 1972 : 251)

1.2 **การให้น้ำหนักคะแนนของแต่ละตัวเลือกไม่เท่ากัน** วิธีการนี้จะมีการกำหนดน้ำหนักคะแนนของแต่ละตัวเลือกเอาไว้ก่อน ถ้าผู้ตอบเลือกตอบตัวเลือกใดก็จะได้คะแนนตามที่กำหนดไว้ วิธีการกำหนดน้ำหนักคะแนนทำได้หลายวิธี เช่น กำหนดค่าน้ำหนักตามระดับความถูกต้องของตัวเลือก หรือกำหนดจากผลการตอบแบบสอบคู่ขนานของกลุ่มมาตรฐาน เป็นต้น วิธีการแก้ปัญหาวิธีนี้ แม้ว่าจะทำให้การวัดมีความเที่ยงสูงขึ้น (Patnaik and Traub, 1973) แต่เมื่อคำนึงถึงการนำไปใช้

แล้วคงจะมีปัญหายุ่งยากในเรื่องการกำหนดน้ำหนักของแต่ละตัวเลือกให้เหมาะสม จึงไม่น่าจะเป็นวิธีการที่เหมาะสมกับการใช้วัดผลในชั้นเรียนตามปกติ

2. **ปรับวิธีการตอบ และวิธีการตรวจให้คะแนน** วิธีการรูปแบบของข้อสอบยังเหมือนเดิม คือ มีตัวเลือกที่เป็นคำตอบที่ถูกเพียงตัวเลือกเดียวแต่วิธีการตอบจะเปลี่ยนไป ซึ่งวิธีการตรวจให้คะแนนก็จะปรับเปลี่ยนไปตามวิธีการตอบด้วย วิธีการนี้มีอยู่ 4 วิธี คือ

2.1 **การให้ตอบปฏิเสธตัวเลือกที่ผิด** เป็นวิธีการที่เสนอโดยคูนส์ และคณะ (Coombs, Milholand, and Womer, 1959) การตอบข้อสอบแบบนี้จะต้องบอกให้ได้ว่า ตัวเลือกใดเป็นตัวเลือกที่ผิดบ้าง ส่วนตัวเลือกที่คิดว่าเป็นตัวเลือกที่ถูก หรือไม่แน่ใจให้เว้นไว้ การตรวจให้คะแนนจะให้ทุกตัวเลือกที่ทำเครื่องหมายจากการตอบข้อสอบ และให้คะแนนโดยวิธีนี้ พบว่า ให้ผลการวัดที่มีความเที่ยงสูงกว่า แบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว แต่ไม่พบว่ามี ความแตกต่างในเรื่องของความตรง ซึ่งสอดคล้องกับการศึกษาของ อรวรรณ ดันชาเจริญรัตน์ (2517) และสำราญ มีแจ้ง (2525) ส่วนการศึกษาของ กาญจนา ศิริวิวัฒนพงษ์ (2520) กลับพบว่า วิธีการนี้มีความเที่ยงต่ำกว่าวิธีของแบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว แต่ในเรื่องของความตรงนั้นสอดคล้องกัน อย่างไรก็ตาม คูนส์ และคณะ ก็ยอมรับว่า วิธีการตรวจให้คะแนนแบบนี้มีความยุ่งยากในการตอบ และไม่สามารถแยกผู้ที่มีความรู้ถูกต้องสมบูรณ์ได้

2.2 **การให้ตอบรับตัวเลือกที่ถูกและตอบปฏิเสธตัวเลือกที่ผิด** วิธีการนี้เสนอโดยอนันต์ ศรีโสภาก (2516) วิธีการตอบนี้ผู้สอบต้องทำเครื่องหมายเพิ่มจากวิธีการที่คูนส์ และคณะเสนอ คือ ต้องทำเครื่องหมายถูกที่ตัวเลือกที่ถูกด้วย การให้คะแนนก็จะให้เป็นรายตัวเลือกเช่นกัน จากการศึกษาวิธีการตรวจให้คะแนนแบบนี้ของ อรวรรณ ดันชาเจริญรัตน์ (2517) และ กาญจนา ศิริวิวัฒนพงษ์ (2520) พบว่า ผลของการให้คะแนนแบบนี้มีความเที่ยงสูงกว่า วิธีการของคูนส์ และคณะ และสูงกว่าวิธีการให้คะแนนแบบ 0-1 ส่วนความตรงนั้นไม่แตกต่างกัน แต่จากการศึกษาของ สำราญ มีแจ้ง (2525) กลับพบว่า วิธีการนี้ให้ผลการวัดที่มีความเที่ยงไม่แตกต่างจากวิธีการให้คะแนนแบบ 0-1 แต่มีความตรงสูงกว่า ส่วนค่าอำนาจจำแนกนั้นไม่แตกต่างกัน จุดอ่อนของวิธีการนี้ก็เช่นเดียวกับวิธีการของคูนส์ และคณะ คือ ถ้าผู้สอบทราบว่าตัวเลือกใดเป็นคำตอบที่ถูกแล้ว จะสามารถทำเครื่องหมายผิดที่ตัวเลือกที่เหลือ และได้คะแนนเต็มในข้อนั้นได้ โดยไม่ต้องทราบว่า ตัวเลือกเหล่านั้นทำไมจึงผิด ด้วยเหตุนี้วิธีการดังกล่าวนี้ จึงไม่น่าจะสามารถแก้ปัญหาของการใช้แบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียวได้อย่างมีประสิทธิภาพ

2.3 **การให้เลือกคำตอบได้หลายตัวเลือก** วิธีนี้ผู้สอบต้องทำเครื่องหมายแสดงว่าตัวเลือก

นั้น ๆ ถูกได้หลายตัวเลือกเท่าที่ผู้สอบคิดได้ว่า ตัวเลือกใดควรจะเป็นคำตอบที่ถูก โดยที่คำตอบที่ถูกนั้นมิได้อยู่เพียงตัวเลือกเดียว ส่วนคะแนนที่ได้นั้นจะลดลงไปตามจำนวนตัวเลือกที่เลือกตอบเพิ่มขึ้น ผู้ใช้วิธีการนี้เชื่อว่า ยิ่งผู้สอบทำเครื่องหมายที่ตัวเลือกหลายตัว แสดงว่าผู้สอบมีความรู้ในเรื่องนั้นน้อยลงเรื่อย ๆ อย่างไรก็ตามแบบสอบเลือกตอบชนิดเลือกคำตอบได้ไม่จำกัดนี้ยังเป็นข้อสอบที่มีคำตอบที่มีตัวเลือกที่ถูกต้องอยู่เพียงตัวเลือกเดียว เมื่อผู้สอบเลือกคำตอบที่ตนคิดว่าถูกได้แล้ว จะไม่สนใจตัวเลือกอื่น เพราะจะถือว่าผิดทันทีโดยไม่สามารถบอกได้ว่า ทำไม่จึงผิด ดังนั้น การเขียนข้อสอบที่ขาดความระมัดระวัง ก็อาจช่วยให้ผู้สอบเลือกคำตอบที่ถูกต้องได้ง่ายขึ้น และการให้คะแนนก็ยังถือได้ว่าไม่ง่ายนัก

2.4 การให้ผู้สอบขอระดับความมั่นใจในการตอบ วิธีการนี้เมื่อผู้สอบเลือกคำตอบที่คิดว่าถูกแล้วผู้สอบจะต้องบอกด้วยว่ารู้สึกมั่นใจแค่ไหนว่า คำตอบที่เลือกนั้นเป็นคำตอบที่ถูกต้องจริง ๆ ซึ่งผลการศึกษาวิธีการตรวจให้คะแนนวิธีนี้ มีรายงานผลตรงกันว่าทำให้การวัดความเที่ยงสูงกว่าการให้คะแนนแบบ 0-1 หรือวิธีการใช้สูตรปรับแก้การเดา (Pugh and Brunza, 1975 ; Abu-Sayf and Diamond, 1976 ; เทัญศรี สว่างเนตร, 2520) ส่วนเรื่องความตรงนั้น Abu-Sayf and Diamond พบว่าสูงกว่าการให้คะแนนแบบ 0-1 เล็กน้อยวิธีการให้คะแนนตามระดับความมั่นใจดูจะเป็นวิธีที่น่าสนใจในการแก้ปัญหาการใช้ข้อสอบแบบเลือกตอบ แต่ก็มีข้อควรพิจารณาในการตรวจให้คะแนนแบบนี้คือ การที่ผู้สอบมีความมั่นใจในระดับสูงว่าคำตอบที่ตนเลือกเป็นคำตอบที่ถูกต้อง แต่กลับปรากฏว่าผิด ผู้สอบจะได้คะแนนน้อยกว่าผู้ที่ตอบว่าไม่มั่นใจเลยนั้น เป็นสิ่งที่ถูกต้องหรือยุติธรรมหรือไม่ เพราะผู้ที่มั่นใจมากแต่ตอบผิดนั้นถือว่าเป็นผู้ที่มีความรู้ผิด (misinformation) ส่วนผู้ที่ไม่มั่นใจนั้นถือว่าตอบโดยการเดา ผู้ที่มีความรู้ที่ไม่ถูกต้องนั้น ไม่น่าจะถือว่ามีความสามารถน้อยกว่าผู้ไม่รู้ ในแง่ของการนำไปใช้การให้คะแนนแบบนี้มีความยุ่งยากพอสมควร

3. ปรับรูปแบบข้อสอบ วิธีการตอบ และวิธีการตรวจให้คะแนน วิธีการนี้รูปแบบของแบบสอบจะถูกปรับให้มีตัวเลือกที่เป็นคำตอบที่ถูกได้หลายตัวเลือก หรือ ไม่มีเลย ส่วนวิธีการตอบและการตรวจให้คะแนนจะปรับตามรูปแบบแต่ละรูปแบบ ซึ่งวิธีนี้มีอยู่ 3 วิธี คือ

3.1 การใช้แบบสอบเลือกตอบชนิดหลายคำตอบ (multiple-answer : MA) แบบสอบนี้ผู้สอบจะต้องเลือกตัวเลือกที่เป็นคำตอบที่ถูกให้ได้ทุกตัว โดยผู้สอบไม่ทราบว่า ตัวเลือกที่ถูกมีกี่ตัว และการตอบจะให้ทำเครื่องหมายเฉพาะตัวเลือกที่คิดว่าถูกต้องเท่านั้น จากข้อมูลเชิงประจักษ์เกี่ยวกับการศึกษาถึงการใช้แบบสอบนี้ พบว่า แบบสอบนี้มีคุณภาพดีกว่าแบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียวทั้งในด้านความเที่ยงและความตรง (Hsu, Moss, and Khampalikit, 1984) อย่างไรก็ตาม

ก็ตาม แบบสอบชนิดนี้ก็มีจุดอ่อน ในด้านตัวเลือกที่เป็นคำตอบถูกต้องในแต่ละข้อมีจำนวนไม่เท่ากัน ทำให้คะแนนรวมในแต่ละข้อไม่เท่ากัน และถ้าจะปรับให้คะแนนแต่ละข้อเท่ากันจำเป็นต้องมีสูตรปรับแก้ ซึ่งเป็นการเพิ่มความยุ่งยากในการตรวจให้คะแนน และอาจทำให้เกิดความไม่สะดวกในการนำไปใช้

3.2 การใช้แบบสอบเลือกตอบชนิดเชิงซ้อน (complex multiple-choice : CMC)

แบบสอบชนิดนี้เป็นการนำเอาตัวเลือกของข้อสอบชนิดหลายคำตอบมาจัดกลุ่มใหม่ โดยจะมีกลุ่มตัวเลือกที่ประกอบด้วยตัวเลือกที่เป็นคำตอบที่ถูกต้องทั้งหมดเป็นตัวเลือกที่ถูกต้องเพียงกลุ่มเดียว ซึ่งการนำตัวเลือกมาจัดกลุ่มใหม่นั้น ถ้าจัดกลุ่มไม่เหมาะสมแล้วอาจกลายเป็นการชี้แนะคำตอบที่ถูกต้องและผู้ที่มีความรู้บ้างบางส่วนก็จะสามารถเลือกคำตอบที่ถูกต้องได้ง่ายขึ้น จึงทำให้ได้คะแนนสูงกว่าที่ควรจะเป็น ด้วยเหตุผลนี้แบบสอบชนิดนี้ จึงไม่ใช่วิธีแก้ปัญหของการใช้แบบสอบเลือกตอบที่ถูกต้องเพียงตัวเลือกเดียวที่คิติก

3.3 การใช้แบบสอบเลือกตอบแบบถูกผิด (multiple true-false : MTF) แบบสอบ

ชนิดนี้จะมีคำตอบที่ถูกต้องหลายคำตอบ หรือไม่มีเลย ผู้ตอบจะต้องตอบให้ได้ว่าตัวเลือกใดบ้างเป็นตัวเลือกที่ถูกต้อง และตัวเลือกใดบ้างเป็นตัวเลือกที่ผิด จากการศึกษาเกี่ยวกับแบบสอบชนิดนี้ของ Frisbie และ Sweeney (1982) พบว่า แบบสอบชนิดนี้มีความเที่ยงและความตรงสูงกว่าแบบสอบเลือกตอบที่ถูกต้องเพียงตัวเลือกเดียว ซึ่งสอดคล้องกับงานวิจัยของ Hill and Woods (1974 cited in Haladyna, 1994); Albanese, Kent, and Whitney (1977); Frisbie and Druva (1986); Downing, Baranowski, Grosso, and Norcini (1995) นอกจากนี้ยังพบว่า ผู้สอบส่วนใหญ่ชอบแบบสอบชนิดนี้มากกว่า แบบสอบเลือกตอบที่ถูกต้องเพียงตัวเลือกเดียว เช่นเดียวกับงานวิจัยของ จินดา โดอนันต์ (2526) ส่วนการศึกษาของ Kolstad และคณะ (Kolstad, Goaz, and Kolstad, 1982) พบว่าผู้สอนรู้สึกชอบแบบสอบชนิดนี้เพราะเขียนข้อสอบได้ง่ายขึ้น และช่วยลดสิ่งแนะที่เป็นประโยชน์ต่อการเดาในข้อสอบได้ และจักรกฤษณ์ สาราณใจ (2530) พบว่า ข้อสอบเลือกตอบชนิดตัดสินคำตอบทุกตัวเลือกที่ตรวจให้คะแนนแก่ความรู้ที่ถูกต้องสมบูรณ์ (multiple true-false questions with absolute knowledge scoring : MTF.A) และ ข้อสอบเลือกตอบชนิดตัดสินคำตอบทุกตัวเลือก ที่ตรวจให้คะแนนแก่ความรู้บางส่วน (multiple true-false questions with partial knowledge scoring : MTF.P) มีประสิทธิภาพสูงกว่า ข้อสอบเลือกตอบชนิดแบบฉบับ (the typical multiple-choice questions : T.MC) สำหรับผู้ตอบที่มีความสามารถในระดับต่ำมาก ไปจนถึงระดับสูง

จะเห็นว่าวิธีการแก้ปัญหของการใช้แบบสอบเลือกตอบที่ถูกต้องเพียงตัวเลือกเดียวแต่ละวิธีมีจุดเด่น และจุดด้อยแตกต่างกันไป วิธีที่ผู้วิจัยเห็นว่าน่าจะแก้ปัญหทั้งหมดตามประการของแบบสอบ

เลือกตอบที่ถูกเพียงตัวเลือกเดียวได้โดยไม่ยุ่งยากก็คือ การใช้แบบสอบเลือกตอบแบบถูกผิด และจากผลการศึกษาเชิงประจักษ์เกี่ยวกับแบบสอบชนิดนี้ ก็พบว่า มีประสิทธิภาพสูงกว่าเมื่อเปรียบเทียบกับ แบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว จึงเหมาะสมที่จะนำไปใช้ในการวัดผลการเรียนในชั้นเรียนเป็นอย่างมาก

ตอนที่ 2 โน้ตสั้นเกี่ยวกับแบบสอบเลือกตอบแบบถูกผิด (multiple true false : MTF)

แบบสอบเลือกตอบแบบถูกผิด (multiple true false : MTF) เป็นแบบสอบปรนัยชนิดหนึ่ง ที่สร้างขึ้นมา เพื่อให้ผลการตรวจสอบทางการศึกษาแม่นยำยิ่งขึ้น และเพื่อแก้ไขข้อบกพร่องในแบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียวเกี่ยวกับการเปิดโอกาสให้ผู้สอบตอบข้อสอบได้ถูกต้อง โดยการเอา ทำให้ผลที่ได้จากการวัดมีความถูกต้องน้อยลง โดยเฉพาะในกรณีที่คำตอบถูกต้องของข้อกระทงมีได้หลายคำตอบ และเป็นคำตอบที่ถูกต้องเท่า ๆ กัน ซึ่งข้อสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียวนั้นไม่เหมาะสมจะใช้วัด โดยเฉพาะอย่างยิ่งในกรณีที่ต้องการทราบว่าผู้ตอบรู้คำตอบที่ถูกต้องทั้งหมดหรือไม่ และการมีตัวเลือกที่ถูกเพียงตัวเลือกเดียว อาจทำให้ผู้สอบสามารถตอบข้อสอบได้ถูก โดยไม่ต้องมีความรู้ในเรื่องนั้นอย่างครบถ้วนสมบูรณ์ก็ได้ รวมทั้งการตรวจให้คะแนน 1 คะแนน สำหรับผู้เลือกคำตอบได้ถูก และ 0 คะแนน สำหรับผู้เลือกคำตอบไม่ถูกนั้น ไม่สามารถแยกผู้ที่มีความรู้ถูกต้องสมบูรณ์ กับผู้ที่มีความรู้บางส่วน และผู้ที่ไม่มีความรู้ออกจากกันได้เลย (Coombs, Milholland, and Womer, 1956 : จักรกฤษณ์ ตำราวิจัย, 2530) จากข้อบกพร่องดังกล่าวนี้ นักวัดผลได้พยายามปรับปรุงแก้ไข เพื่อให้ผลที่ได้จากการวัดมีประสิทธิภาพมากยิ่งขึ้น ด้วยวิธีการต่าง ๆ หลายวิธีดังได้กล่าวมาแล้ว ซึ่งวิธีการที่เหมาะสมที่สุด คือ การใช้แบบสอบเลือกตอบแบบถูกผิด เพราะแบบสอบนี้มีคำตอบที่ถูกต้องอยู่ได้หลายคำตอบ โดยผู้ตอบต้องสามารถจำแนกให้ได้ว่า ตัวเลือกใดถูกตัวเลือกใดผิด ทำให้การวัดความละเอียดถี่ถ้วนแม่นยำยิ่งขึ้น และสามารถประมาณค่าความสามารถที่แท้จริงของผู้สอบได้ถูกต้องรวมทั้งการนำไปใช้ในสถานการณ์จริงมีความเป็นไปได้อย่างมากเนื่องจากการเขียนข้อสอบการตอบและการตรวจให้คะแนนสามารถทำได้ง่าย เพราะคล้ายกับแบบสอบถูกผิด ด้วยเหตุนี้จึงมีผู้นำไปใช้วัดผลในระดับต่าง ๆ มากขึ้น

แบบสอบเลือกตอบแบบถูกผิด เป็นแบบสอบที่มีการพัฒนา และใช้กันมานานแล้วก่อนปี ค.ศ. 1939 ซึ่งแบบสอบนี้ในระยะแรกจะนำไปใช้ในวงการแพทยศึกษา และสาขาวิชาที่เกี่ยวข้องกับวิทยาศาสตร์สุขภาพ (Harden, 1979 ; Albanese and Sabers, 1988) เพื่อหารูปแบบของข้อสอบแบบใหม่นอกเหนือไปจากข้อสอบเลือกตอบชนิดเชิงซ้อน (complex multiple-choice : CMC) ที่ใช้กันอยู่อย่างแพร่หลายในวงการดังกล่าว และพบว่ามีประสิทธิภาพในการวัดอยู่ในขั้นที่ไม่น่าพอใจ

นอกจากนี้ได้มีการนำไปใช้ในการทดสอบ เพื่อคัดเลือกบุคคลเข้าศึกษาในมหาวิทยาลัยในไต้หวัน (Tsai and Suen, 1993) และในการทดสอบระดับประโยคมัธยมอีกด้วย (Wood, 1977 : 23)

แบบสอบเลือกตอบแบบถูกผิด เป็นแบบสอบที่มีรูปแบบต่างไปจากแบบสอบอื่น ๆ ตรงที่ตัวเลือกที่กำหนดให้ซึ่งคำตอบที่ถูกต้องอาจจะมีหลายคำตอบหรือไม่มีเลยก็ได้โดยผู้สอบจะต้องตอบให้ได้ว่า ตัวเลือกใดบ้างเป็นตัวเลือกที่ถูก และตัวเลือกใดบ้างเป็นตัวเลือกที่ผิด นั่นก็จะต้องพิจารณาตัดสินคำตอบทุกตัวเลือก และการให้ตอบในลักษณะนี้เป็นการวัดที่คล้ายกับการวัดความคิดรวบยอด ซึ่งผู้สอบจะต้องสามารถแยกแยะสิ่งที่เป็นความคิดรวบยอด กับสิ่งที่ไม่เป็นความคิดรวบยอดออกจากกันให้ได้ (Roid and Haladyna, 1982 : 150) จึงจะถือว่ารู้จริง นอกจากนี้ยังสามารถวัดเนื้อหา และพฤติกรรมเดียวกันซ้ำหลาย ๆ ครั้งได้ โดยไม่ต้องเพิ่มจำนวนข้อสอบและ ไม่ทำให้ผู้สอบเกิดความเบื่อหน่าย ในการทำแบบสอบที่คล้ายกันซ้ำหลายครั้งอีกด้วย ซึ่งทำให้แน่ใจได้ว่าสามารถวัดความรู้ความสามารถของผู้สอบได้ถูกต้องมากขึ้น

2.1 ประสิทธิภาพของแบบสอบเลือกตอบแบบถูกผิด

Frisbie (1992 cited in Haladyna, 1994 ; Popham, 1995) ได้ศึกษางานวิจัยที่ผ่านมาพบว่าประสิทธิภาพของแบบสอบเลือกตอบแบบถูกผิด มีดังนี้

2.1.1 ด้านความเที่ยง และความตรง Frisbie (1992) พบว่า แบบสอบนี้มีความเที่ยงและความตรงสูงมาก ซึ่งสอดคล้องกับงานวิจัยของ Hill and Woods (1974 cited in Haladyna, 1994) ; Albanese, Kent, and Whitney (1977) ; Frisbie and Sweeney (1982) ; Frisbie and Druva (1986) ; Downing, Baranowski, Grosso, and Norcini (1995) ที่พบว่ารูปแบบของแบบสอบเลือกตอบแบบถูกผิด สามารถประมาณค่าความเที่ยง และความตรงได้สูงกว่าเมื่อเปรียบเทียบกับแบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว (multiple-choice)

2.1.2 ด้านการพัฒนาแบบสอบ สามารถสร้างแบบสอบที่วัดเนื้อหา และพฤติกรรมเดียวกันได้ซ้ำหลาย ๆ ครั้ง โดยไม่ต้องเพิ่มจำนวนข้อสอบ (จักรกฤษณ์ สารานูใจ, 2530) ซึ่งเหมาะสมสำหรับใช้ตัดสินผลการเรียน และวินิจฉัยข้อบกพร่องในการเรียนได้เป็นอย่างดี

2.1.3 ด้านการเขียนข้อสอบ สามารถทำได้ง่าย เพราะไม่ต้องพยายามสร้างตัวเลือกให้มีคำตอบที่ถูกต้องเพียงคำตอบเดียว ซึ่งบางคำถามหรือบางวิชาสามารถทำได้ยาก และในบางครั้งอาจกลายเป็นการแนะคำตอบให้ผู้สอบตอบได้ถูกต้องมากขึ้น

2.1.4 ด้านความพึงพอใจของผู้ตอบ จากการศึกษาของ Frisbie และ Sweeney (1982) พบว่า ผู้ตอบส่วนใหญ่ชอบแบบสอบเลือกตอบแบบถูกผิดมากกว่าแบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว เพราะแบบสอบนี้เปิดโอกาสให้ผู้สอบได้แสดงว่า เขาได้เรียนรู้อะไรมา

บ้างได้มากกว่าแบบสอบเลือกตอบรูปแบบอื่น ๆ

2.1.5 ด้านความคลาดเคลื่อนในการวัด มีความคลาดเคลื่อนน้อย เพราะเป็นผลมาจากความสามารถของผู้สอบอย่างแท้จริง

2.1.6 ด้านการตรวจให้คะแนน สามารถทำได้ง่าย และสะดวกเหมือนกับการตรวจให้คะแนนข้อสอบแบบถูกผิด (true-false หรือ binary-choice item) ทั่วไป

2.2 ข้อจำกัดของแบบสอบเลือกตอบแบบถูกผิด

ข้อจำกัดของแบบสอบเลือกตอบแบบถูกผิด (Frisbie and Sweeney, 1982 : 34 และ Haladyna, 1994 : 47) มีดังนี้

2.2.1 ด้านการใช้วัดความคิดรวบยอด แบบสอบนี้ไม่เหมาะสมจะใช้วัด

2.2.2 ด้านการใช้วัดข้อความรู้ที่เป็นจริง (fact) แบบสอบนี้ไม่เหมาะสมจะใช้วัด

2.2.3 ด้านเวลาต้องกำหนดเวลาในการทำแบบสอบให้มากกว่าแบบสอบเลือกตอบแบบอื่น ๆ ถ้ากำหนดเวลาให้ไม่เพียงพอ จะทำให้ค่าความเที่ยงของแบบสอบต่ำลงด้วย

2.2.4 ความไม่เป็นอิสระของแบบสอบ มีผลต่อการประมาณค่าความเที่ยงของแบบสอบ โดยเฉพาะการกำหนดเขตของข้อสอบนั้น ไม่เหมาะสมสำหรับแบบสอบประเภทนี้

2.3 หลักของแบบสอบเลือกตอบแบบถูกผิด

หลักของแบบสอบเลือกตอบแบบถูกผิด (Haladyna, 1994 : 83) มีดังนี้

2.3.1 จำนวนตัวเลือกในแต่ละข้อกระทงของแบบสอบในแต่ละฉบับไม่เท่ากัน

2.3.2 สามารถนำหลักการของแบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว (multiple-choice : MC), แบบสอบเลือกตอบแบบตัวเลือกซ้อน (complex multiple-choice : CMC) และแบบสอบแบบถูกผิด (true-false : TF) มาพัฒนาเป็นแบบสอบแบบนี้ได้เป็นอย่างดี

2.3.3 จำนวนข้อสอบในแบบสอบขึ้นอยู่กับ ความเหมาะสมของหน้ากระดาษ ในหนึ่งหน้า อาจจะมีประมาณ 30-35 ข้อ

2.3.4 ไม่จำกัดเกี่ยวกับจำนวนข้อถูก และผิดในแต่ละข้อกระทง

2.4 การสร้างแบบสอบเลือกตอบแบบถูกผิด

การสร้างแบบสอบเลือกตอบแบบถูกผิด ใช้หลักการเกี่ยวกับการสร้างแบบสอบ

เลือกตอบที่ถูกเพียงตัวเลือกเดียว และแบบสอบแบบถูกผิด โดยมีส่วนตอนปลายของแต่ละข้อกระทงเป็นคำถาม ซึ่งเขียนโดยใช้หลักการเช่นเดียวกับ แบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว (multiple-choice) คือ ตัวคำถามอาจอยู่ในรูปประโยคคำถามที่สมบูรณ์ เช่น ใครเป็นผู้แต่ง "พระอภัยมณี" หรืออาจจะเขียนในรูปประโยคที่ไม่สมบูรณ์ก็ได้ เช่น โดเกียว เป็นเมืองหลวงของ และส่วนตัวเลือกจะมีหลายตัวเลือก ซึ่งแต่ละตัวเลือกจะต้องเป็นข้อความ หรือประโยคที่สามารถตอบได้ว่าถูก หรือผิด แต่ไม่ได้กำหนดว่าจะต้องมีตัวเลือกถูก หรือผิดในข้อกระทงนั้น ๆ จำนวนเท่าใด ทั้งนี้ขึ้นอยู่กับความเหมาะสมของเนื้อหา ในข้อกระทงหนึ่ง ๆ อาจประกอบด้วยตัวเลือกที่เป็นข้อความหรือประโยคที่ถูกทั้งหมด ผิดทั้งหมด ถูกและผิดเท่า ๆ กัน หรือไม่เท่ากันก็ได้ โดยในการเขียนใช้หลักการเดียวกันกับการเขียนข้อสอบแบบถูกผิด

2.4.1 หลักการสร้างข้อคำถามในแบบสอบเลือกตอบถูกผิดทุกตัวเลือก (Gronlund, 1981 ; Roid and Haladyna, 1982 ; Haladyna, 1994 ; Popham, 1995 ; อุทุมพร จามรมาน, 2535 และเขาวดี วิบูลย์ศรี, 2539) มีดังนี้

2.4.1.1 ควรวัดผลการเรียนรู้ที่สำคัญเพียงเรื่องเดียว

2.4.1.2 ต้องมีความชัดเจน กะทัดรัดตรงจุดมุ่งหมาย และถูกหลักไวยากรณ์ อ่านแล้วเข้าใจได้ว่าต้องการถามอะไร

2.4.1.3 ควรเขียนปัญหา หรือสร้างสถานการณ์ที่น่าสนใจขึ้นมาใหม่ ไม่ควรคัดลอกข้อความจากหนังสือมาถามโดยตรง

2.4.1.4 ควรหลีกเลี่ยงการใช้คำศัพท์ที่รู้กันเฉพาะในกลุ่ม

2.4.1.5 ควรหลีกเลี่ยงคำหรือข้อความที่ชี้แนะคำตอบที่ถูกต้อง

2.4.1.6 ควรหลีกเลี่ยงข้อความเชิงปฏิเสธ หากจำเป็นต้องใช้ ควรเน้นให้เห็นชัดเจน โดยการขีดเส้นใต้ หรือพิมพ์ให้เป็นตัวหนา ๆ หรือใช้อักษรตัวใหญ่ ๆ

2.4.1.7 ควรใช้ภาษาที่ง่าย และชัดเจนเหมาะสมกับวัยของนักเรียน คำศัพท์ที่ใช้ไม่ควรยากเกินไป และข้อความต้องไม่ซับซ้อน

2.4.1.8 พยายามให้ข้อสอบแต่ละข้อมีอิสระต่อกัน

2.4.2 หลักการสร้างตัวเลือกในแบบสอบเลือกตอบแบบถูกผิด (Gronlund, 1981 ; Roid and Haladyna, 1982 ; Popham, 1995 ; ภัทรา นิคมานนท์, 2534 ; อุทุมพร จามรมาน, 2535 และเขาวดี วิบูลย์ศรี, 2539) มีดังนี้ คือ

2.4.2.1 ควรหลีกเลี่ยงคำที่มีหลายความหมาย เลือกคำที่สามารถแยกออกเป็นสองความหมายได้อย่างแท้จริงตามธรรมชาติ

2.4.2.2 ควรหลีกเลี่ยงข้อความที่เป็นปฏิเสธ และปฏิเสธซ้อนปฏิเสธ

2.4.2.3 ควรหลีกเลี่ยงข้อความยาว ๆ ที่ยากต่อการอ่าน และทำความเข้าใจ เพราะนักเรียนจะจับใจความไม่ได้ว่าอะไรคือ สิ่งที่จะทดสอบ

2.4.2.4 ควรให้มีจำนวนตัวเลือกที่ถูกต้อง และผิด ในแต่ละข้อกระทงอยู่กระจายกันออกไป ไม่ควรให้อยู่รวมกันเป็นกลุ่ม หรือเรียงกันอย่างเป็นระบบ

2.4.2.5 ควรมีความชัดเจนในแต่ละตัวเลือก และมีความเป็นอิสระต่อกัน

ตอนที่ 3 มโนทัศน์เกี่ยวกับทฤษฎีการตอบสนองข้อสอบ (IRT)

ทฤษฎีการตอบสนองข้อสอบ (item response theory : IRT) เป็นทฤษฎีทางการวัดที่อธิบายความสัมพันธ์ระหว่างคุณลักษณะภายใน หรือ ความสามารถที่อยู่ภายในตัวบุคคลกับ พฤติกรรมการตอบสนองข้อสอบแต่ละข้อว่า มีความน่าจะเป็นในการตอบถูกมากน้อยเพียงใด ทฤษฎีนี้เชื่อว่า พฤติกรรมการตอบสนองต่อข้อสอบของผู้สอบ จะถูกกำหนดโดยคุณลักษณะภายใน หรือความสามารถที่อยู่ภายในตัวบุคคล (ศิริชัย กาญจนวาที, 2535) ซึ่งทฤษฎีนี้ตั้งอยู่บนหลักพื้นฐานที่สำคัญ 2 ประการ คือ

1. ความสามารถของบุคคลในการตอบข้อสอบได้ถูก หรือผิด สามารถอธิบายได้ด้วยคุณลักษณะภายในหรือความสามารถ (latent trait or ability) ของบุคคลนั้น
2. ความสัมพันธ์ระหว่างความสามารถในการตอบข้อสอบถูกต้อง กับความสามารถของผู้สอบที่วัดจากแบบสอบสามารถอธิบายได้ด้วยโค้งลักษณะข้อสอบ (item characteristic curve : ICC) (ศิริชัย กาญจนวาที, 2535 และชนวัฒน์ แสนสุข, 2539)

พัฒนาการของทฤษฎีเริ่มจากแนวคิดของลอร์วี (Lawley, 1943 อ้างในศิริชัย กาญจนวาที, 2535) เกี่ยวกับ โมเดลนอร์มัลโอโง (normal ogive model) จากนั้นราสช์ (Rasch) ได้เสนอ โมเดลราสช์ (rasch model) แบบ 1 พารามิเตอร์ โดยมีแนวคิดว่าคุณค่าความยากของข้อสอบ (b) เป็นสิ่งเดียวที่มีอิทธิพลต่อการตอบสนองข้อสอบ ต่อมาในปี ค.ศ. 1952 ลอร์ด (Lord) ได้เสนอฟังก์ชันนอร์มัลโอโงแบบ 2 พารามิเตอร์ โดยเพิ่มพารามิเตอร์อำนาจจำแนก (a) และในปี ค.ศ. 1968 เบิร์นบอม (Birnbaum) ได้พัฒนาโมเดลโลจิสติก (logistic model) ซึ่งเป็นฟังก์ชันที่สามารถคำนวณได้ง่ายขึ้น ในปี ค.ศ. 1974 ลอร์ด (Lord) ได้เสนอโมเดลโลจิสติกแบบ 3 พารามิเตอร์ โดยเพิ่มพารามิเตอร์โอกาสในการเดาข้อสอบในการประมาณค่าพารามิเตอร์ของผู้สอบด้วยและในปัจจุบัน บาร์ตัน และลอร์ด (Barton and Lord อ้างใน ชนวัฒน์ แสนสุข, 2539) ได้เสนอโมเดลโลจิสติกแบบ 4 พารามิเตอร์ โดยเชื่อว่านักเรียนที่มีความสามารถสูงมาก ไม่จำเป็นต้องตอบข้อสอบถูกเสมอไป แม้ข้อสอบจะง่ายหรือจะระมัดระวังอย่างมากแล้วก็ตามก็ยังไม่สามารถใช้สารสนเทศที่มีตอบข้อสอบได้ถูกต้องตามเฉลยของผู้ออกข้อสอบได้เสมอไป

นอกจากนี้โมเดลการตอบสนองข้อสอบสามารถแบ่งเป็น 4 หัวข้อดังต่อไปนี้ (Hambleton and Swaminathan, 1985 อ้างในศิริรัช กาญจนวาที, 2538)

1. ประเภทของโมเดลการตอบสนองข้อสอบ

จากแนวคิดตามทฤษฎีการตอบสนองข้อสอบ ได้มีการพัฒนาโมเดล หรือแบบจำลอง ขึ้นหลายรูปแบบ ซึ่งสามารถจำแนกได้ดังนี้ (Hambleton and Sawaminathan, 1985 อ้างในศิริรัช กาญจนวาที, 2535)

1.1 โมเดลการตอบสนองข้อสอบ ที่ใช้กับการตรวจให้คะแนนรายข้อแบบ 2 ค่า (dichotomous) เป็นข้อสอบที่ตรวจให้คะแนนแบบ 0, 1 คือ ตอบผิดให้ 0 คะแนน ตอบถูกให้ 1 คะแนน ซึ่งพัฒนาขึ้นในระยะเริ่มแรกเช่น สเกลกัทแมนสมบูรณ์ (Guttman Perfect Scale) โมเดลระยะห่างแฝง (Latent Distance Model) โมเดลเชิงเส้นตรง (Linear Model) เป็นต้น ต่อมาได้มีการพัฒนาโมเดลปกติสะสม แบบ 1, 2 และ 3 พารามิเตอร์ และโมเดลโสตติค แบบ 1, 2 และ 3 พารามิเตอร์

1.2 โมเดลการตอบสนองข้อสอบ ที่ใช้กับการตรวจให้คะแนนรายข้อแบบมากกว่า 2 ค่า (polytomous) โมเดลการตอบสนองข้อสอบแบบนามบัญญัติ (Nominal Response Model : NRM) โมเดลการตอบสนองข้อสอบ (Grade Response Model : GRM) โมเดลการให้คะแนนบางส่วน (Partial Credit Model : PCM)

1.3 โมเดลการตอบสนองข้อสอบ ที่ใช้กับการตรวจให้คะแนนรายข้อแบบต่อเนื่อง (continuous) เช่น โมเดลการตอบสนองข้อสอบแบบคะแนนต่อเนื่อง (Continuous Response Model : CRM) ซึ่งพัฒนาโดย Samejima ในปี ค.ศ. 1972 เป็นต้น

โมเดลการตอบสนองข้อสอบที่ใช้กันมาก จะเป็นโมเดลที่ตรวจให้คะแนนแบบทวิภาค และพหุภาคซึ่งโมเดลทั้ง 2 รูปแบบนี้มีลักษณะทั่วไป ข้อดีและข้อเสียแตกต่างกันไป ดังตารางที่ 1

ตารางที่ 1 การเปรียบเทียบการตรวจให้คะแนนตามโมเดลการตอบสนองข้อสอบ

แบบทวิภาค (dichotomous)	แบบพหุภาค (polytomous)
ลักษณะ กำหนดตัวเลือก หรือคำตอบที่เหมาะสม และถูกที่สุดเป็น 1 คะแนน ส่วนคำตอบอื่น ๆ เป็น 0 คะแนน	กำหนดคะแนนให้กับคำตอบในแต่ละข้อมากกว่า 2 คะแนน ซึ่งแต่ละค่าแสดงถึงระดับความสามารถของผู้สอบที่ตอบข้อกระทงนั้น
ข้อดี สะดวก รวดเร็ว และประหยัดเวลาในการตรวจให้คะแนน	ทำให้ทราบได้ว่า ผู้ที่ตอบข้อคำถามนั้นมีความสามารถอยู่ในระดับใด
ข้อเสีย ไม่สามารถทราบได้ว่า คนที่ตอบได้ 0 คะแนนนั้น เป็นคนที่มีความสามารถอยู่ในระดับใด	ในบางสถานการณ์ ไม่สามารถกำหนดให้ชัดเจนได้ว่า ในแต่ละลำดับขั้นของแบบสอบมีค่าคะแนนเป็นเท่าใด
การวิเคราะห์ ใช้โปรแกรม LOGISTIC	ใช้โปรแกรม MULTLOG ในการวิเคราะห์
ผลการวิเคราะห์ ให้ค่าอำนาจจำแนก (a), ค่าความยาก (b), ค่าการเดา (c), และค่าฟังก์ชันสารสนเทศของข้อสอบ (IIF) และแบบสอบ (TIF)	ให้ค่าอำนาจจำแนก (a), ค่าความยากในแต่ละระดับขั้นของคะแนน (b_x), และค่าฟังก์ชันสารสนเทศของข้อสอบ (IIF) และแบบสอบ (TIF)

ในงานวิจัยนี้ ผู้วิจัยตรวจให้คะแนนรายข้อแบบ 2 ค่า คือ ตอบรับ และตอบปฏิเสธตัวเลือก ทุกตัวได้ถูกต้องทุกตัวเลือกได้ 1 คะแนน และถ้าตอบรับ และตอบปฏิเสธตัวเลือกตัวหนึ่งตัวใดผิดอย่างน้อย 1 ตัวได้ 0 คะแนน ซึ่งโมเดลการตอบสนองข้อสอบ ได้แก่ โมเดลปกติสะสม และโมเดลโลจิสแบบ 1, 2 และ 3 พารามิเตอร์ ฟังก์ชันทั้งสองให้ผลลัพธ์ของการประมาณค่าใกล้เคียงกันมาก แต่ฟังก์ชันโลจิสมีลักษณะของสูตรทางคณิตศาสตร์ และวิธีการคำนวณง่ายและสะดวกกว่า นอกจากนี้โมเดลโลจิสยังมีความทนทานต่อความคลาดเคลื่อนที่เกิดขึ้นกับผู้สอบที่มีความสามารถสูงจะตอบข้อสอบได้ดีกว่า จึงทำให้โมเดลโลจิสเป็นที่นิยมกันมากในการนำไปใช้จริง

2. พารามิเตอร์ของโมเดลการตอบสนองข้อสอบ

โมเดลการตอบสนองข้อสอบประกอบด้วยพารามิเตอร์และค่าคงที่ ดังนี้

2.1 พารามิเตอร์ของผู้สอบ

θ = ระดับความสามารถของผู้สอบคนที่ p ซึ่งประมาณได้จากโมเดลตามทฤษฎีการตอบข้อสอบ ซึ่งปรับให้คะแนนมาตรฐานที่มีค่าเฉลี่ยเป็น 0 และส่วนเบี่ยงเบนมาตรฐานเป็น 1 ค่า θ มีพิสัยอยู่ระหว่าง $-\infty$ ถึง $+\infty$ แต่ผลการวิเคราะห์ส่วนมากมักให้ค่าอยู่ในช่วง -3.00 ถึง $+3.00$

2.2 พารามิเตอร์ของข้อสอบ

b_i = ค่าความยากของข้อสอบข้อที่ i ซึ่งเป็นการวัดตำแหน่งของโค้งคุณลักษณะข้อสอบ (ICC) ตามแกนนอนบนสเกลของ θ ณ จุดที่โค้งมีความชันมากที่สุด (จุดเปลี่ยนโค้ง) หรือที่ตำแหน่ง

$$b_i = \theta \text{ ที่ } p_i(\theta) = 0.50 \text{ (สำหรับ 1 และ 2 พารามิเตอร์)}$$

$$b_i = \theta \text{ ที่ } p_i(\theta) = \frac{1+c_i}{2} \text{ (สำหรับ 3 พารามิเตอร์)}$$

ค่า b_i มีพิสัยอยู่ระหว่าง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัตินิยมใช้ข้อสอบที่มีค่า b_i อยู่ระหว่าง -2.50 ถึง $+2.50$ ค่า b_i ที่อยู่ใกล้ -2.50 แสดงว่าเป็นข้อสอบที่ง่าย ส่วน b_i ที่อยู่ใกล้ $+2.50$ แสดงว่าเป็นข้อสอบที่ยาก

a_i = ค่าอำนาจจำแนกของข้อสอบข้อที่ i ซึ่งเป็นความชันของโค้ง ICC ณ จุดเปลี่ยนโค้ง หรือที่จุด $\theta = b_i$ ค่า a_i มีพิสัยอยู่ระหว่าง $-\infty$ ถึง $+\infty$ ค่า a_i ที่เป็นลบเป็นสิ่งที่ไม่พึงปรารถนา ในปฏิบัติจึงนิยมใช้ข้อสอบที่มีค่า a_i อยู่ระหว่าง $+0.50$ ถึง $+2.50$ ค่า a_i ที่สูงแสดงว่าข้อสอบนั้นมี slope ที่ชันจึงจำแนกผู้สอบที่มีความสามารถที่แตกต่างกันได้ดี

c_i = ค่าโอกาสการเดาข้อสอบถูก ซึ่งเป็นความน่าจะเป็นที่ผู้สอบที่มีความสามารถต่ำมาก ๆ จะทำข้อสอบข้อที่ i จึงเป็นค่ากำกับต่ำสุด (lower asymptote) ของ ICC ค่า c_i มีพิสัยอยู่ระหว่าง 0 ถึง 1 โดยทั่วไปนิยมใช้ข้อสอบมีค่า c_i อยู่ระหว่าง 0.00 ถึง 0.03

2.3 ค่าคงที่

e = ค่าคงที่ของลอการิทึมธรรมชาติ ซึ่งมีค่าประมาณ 2.71828

D = ค่าองค์ประกอบของการปรับสเกล ซึ่งมีค่าเท่ากับ 1.70

3. ฟังก์ชันสารสนเทศของแบบสอบ

การวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ (item response theory : IRT) จะใช้แบบแผนการตอบสนองรายข้อในการประมาณค่าความสามารถของผู้สอบ ดังนั้นการประเมินคุณภาพของแบบสอบ จึงสามารถพิจารณาจากความถูกต้องแม่นยำในการประมาณความสามารถของผู้สอบ ซึ่งมีดัชนีที่บ่งชี้ถึงความถูกต้องแม่นยำดังกล่าวเรียกว่า สารสนเทศของแบบสอบ ซึ่งเป็นค่าฟังก์ชันสารสนเทศของแบบสอบที่เกิดจาก ผลรวมเชิงพีชคณิตของค่าฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อรวมเข้าด้วยกัน ดังนั้นผลของข้อสอบแต่ละข้อจะมีผลต่อแบบสอบทั้งฉบับ โดยค่าฟังก์ชันสารสนเทศของข้อสอบ เป็นดัชนีสะสมที่สร้างจากดัชนีคุณลักษณะของข้อสอบหลายลักษณะ ได้แก่ ค่าความยาก ค่าอำนาจจำแนก และ ค่าความแปรปรวนของคะแนนรายข้อ เพื่อบ่งชี้คุณภาพของข้อสอบ (Birnbaum, 1968 อ้างถึงใน ศิริรัช กาญจนวาที, 2535) เนื่องจากค่าสารสนเทศมีความสัมพันธ์ผกผันกับ ความคลาดเคลื่อนมาตรฐานของการประมาณค่า ดังนั้นถ้าค่าสารสนเทศของแบบสอบมีค่าสูง ในช่วง θ ใดก็就会有ความถูกต้องแม่นยำสูงในการประมาณค่าความสามารถของผู้สอบในช่วง θ นั้น ๆ โดยมีความคลาดเคลื่อนมาตรฐานของการประมาณค่าต่ำ ด้วยคุณสมบัติด้านความไม่แปรเปลี่ยนไปตามกลุ่มตัวอย่างของค่าพารามิเตอร์ของข้อสอบจากการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ จึงทำให้ค่าฟังก์ชันสารสนเทศเหมาะสมที่จะใช้เป็นดัชนีบ่งบอกคุณภาพของข้อสอบและแบบสอบแทนการหาความเที่ยง และความคลาดเคลื่อนมาตรฐานของการวัดตามทฤษฎีแบบดั้งเดิม (Hambleton, 1977 อ้างในศิริรัช กาญจนวาที, 2535)

4. ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบมีข้อตกลงเบื้องต้นที่สำคัญ ดังนี้ (ศิริรัช กาญจนวาที, 2535)

4.1 แบบสอบวัดคุณลักษณะเด่นเพียงลักษณะเดียว (unidimensionality : onetrait)

4.2 การตอบข้อสอบเป็นไปอย่างอิสระ (independence : local independence) คือ การตอบข้อสอบแต่ละข้อของผู้สอบแต่ละคนมีความเป็นอิสระจากกันและกัน ประกอบด้วย

4.2.1 ความเป็นอิสระระหว่าง ข้อสอบแต่ละข้อ (item-free) หมายถึง การตอบสนองต่อข้อสอบข้อหนึ่ง ไม่มีผลต่อการตอบข้อสอบข้ออื่น ๆ

4.2.2 ความเป็นอิสระระหว่างผู้สอบ (sample-free) หมายถึง การตอบสนองข้อสอบของผู้สอบหนึ่งต้องไม่มีผลต่อการตอบสนองข้อสอบของผู้สอบคนอื่น

4.3 โค้งคุณลักษณะข้อสอบ สามารถใช้อธิบายพฤติกรรมการตอบสนองข้อสอบ (item characteristic curve : item response model) ความสัมพันธ์ระหว่างความสามารถของผู้สอบ

(๓) สามารถแสดงได้ด้วยโค้งคุณลักษณะข้อสอบ (ICC) โดย ICC เป็นฟังก์ชันคณิตศาสตร์ที่แสดงด้วยโมเดลการตอบสนองข้อสอบที่นิยมใช้กันมี 3 รูปแบบ คือ 1, 2 หรือ 3 พารามิเตอร์ ในการเลือกใช้ขึ้นอยู่กับลักษณะข้อสอบ และความเชื่อในโมเดลที่ใช้อธิบายความสัมพันธ์ดังกล่าว

4.4 ข้อสอบที่ใช้ ต้องไม่เป็นข้อสอบประเภทความเร็ว (speed test) ผู้สอบทุกคน มีโอกาสทำข้อสอบทุกข้อ เพื่อให้คะแนนรวมจากการสอบประกอบด้วยลักษณะของข้อสอบเป็น ตัวประมาณค่าความสามารถที่แท้จริงของผู้สอบ โดยไม่มีข้อจำกัดเกี่ยวกับปัจจัยด้านเวลาที่ใช้ในการสอบ

ตอนที่ 4 งานวิจัยที่เกี่ยวข้อง

การศึกษาเปรียบเทียบ ผลของจำนวนตัวเลือกที่มีต่อค่าฟังก์ชันสารสนเทศของแบบสอบ เลือกตอบแบบถูกผิด ยังไม่เคยมีการศึกษามาก่อนแต่ส่วนใหญ่จะเป็นงานวิจัยที่ศึกษา เพื่อเปรียบเทียบผลของจำนวนตัวเลือกในแบบสอบเลือกตอบที่ถูกเพียงตัวเลือกเดียว (multiple-choice) ซึ่งมีผลการวิจัยที่ได้ข้อค้นพบเป็น 2 แนวทางที่สำคัญคือ

แนวทางแรกให้ข้อสนับสนุนว่า เมื่อจำนวนตัวเลือกของข้อกระทงเพิ่มขึ้นจาก 2 ตัวเลือก เป็น 3 ตัวเลือก, 4 ตัวเลือก และ 5 ตัวเลือก จะทำให้ค่าความเที่ยงของแบบสอบสูงขึ้นตามลำดับ (Thorndike, 1971 ; Ramos and Stern, 1973 ; Budesescu and Nevo, 1985 ; Haladyna and Downing, 1993) และนอกจากนี้ยังพบว่า จำนวนตัวเลือกที่มี 4 ตัวเลือกและ 5 ตัวเลือก จะทำให้ค่าการเดาเกิดขึ้นน้อยกว่า 2 ตัวเลือก และ 3 ตัวเลือก (Ahmann and Glock, 1971; Thorndike and Hagen, 1977) ส่วนข้อสอบที่มีหลายข้ออิทธิพลของการเดามีผลต่อคะแนนสอบมาก แบบสอบที่มี 3 ตัวเลือก เป็นแบบสอบที่มีความเหมาะสม (Tversky, 1964 ; Grier, 1975) และแบบสอบที่มี 2 ตัวเลือกเป็นแบบสอบที่มีประสิทธิภาพมากที่สุดในการตัดสินใจ (Haladyna and Downing, 1993)

ผลงานวิจัยอีกแนวหนึ่งได้ข้อค้นพบว่า ค่าความเที่ยงของแบบสอบแต่ละฉบับมีค่าต่างกัน เมื่อสุ่มตัดตัวลงของแบบสอบที่มีจำนวน 5 ตัวเลือก เป็น 4 ตัวเลือก และ 3 ตัวเลือกโดยค่าความเที่ยงของ แบบสอบเลือกตอบที่มี 5 ตัวเลือกมีค่าสูงสุด และลดต่ำลงมาเป็น แบบสอบเลือกตอบที่มี 4 ตัวเลือก และ 3 ตัวเลือก ตามลำดับ (Owen and Froman, 1978 cited in Haladyna and Downing, 1993 ; วารุณี ปิตธวัชชัย, 2513 ; มาลี ชิตสวน, 2514 และ วิชัย สนิทวัฒนาพานิช, 2531) และการลดจำนวนตัวเลือกของข้อกระทงจาก 5 ตัวเลือก เป็น 4 ตัวเลือก และ 3 ตัวเลือก ไม่ทำให้ค่าความยาก และค่าอำนาจจำแนกของแบบสอบทั้งสามชุดแตกต่างกันที่ระดับนัยสำคัญ .05 (วารุณี ปิตธวัชชัย, 2513 และวิชัย สนิทวัฒนาพานิช, 2531) นอกจากนี้ยังพบว่า ในกลุ่มที่มีระดับ

ความสามารถสูง แบบสอบเลือกตอบที่มี 5 ตัวเลือก มีประสิทธิภาพสูงสุด รองลงมาคือ แบบสอบเลือกตอบที่มี 4 ตัวเลือก และ 3 ตัวเลือก ส่วนในกลุ่มที่มีระดับความสามารถปานกลาง แบบสอบเลือกตอบที่มี 3 ตัวเลือก มีประสิทธิภาพสูงสุด รองลงมา คือ แบบสอบเลือกตอบที่มี 5 ตัวเลือก และ 4 ตัวเลือก สำหรับกลุ่มที่มีระดับความสามารถต่ำ แบบสอบเลือกตอบที่มี 4 ตัวเลือก มีประสิทธิภาพสูงสุด รองลงมา คือ แบบสอบเลือกตอบที่มี 5 ตัวเลือก และ 3 ตัวเลือก ตามลำดับ (ธีระศักดิ์ อุรัจนานนท์, 2530)

จากงานวิจัยดังกล่าวข้างต้นจะเห็นได้ว่า เป็นการศึกษาเพื่อตอบคำถามว่า เมื่อจำนวนตัวเลือกในแบบสอบแตกต่างกัน จำนวนตัวเลือกเท่าใดจะมีค่าความเที่ยงสูงสุด โดยมีผลการศึกษาที่ค่อนข้างชัดเจนว่า เมื่อจำนวนตัวเลือกของข้อกระทงเพิ่มขึ้น หรือลดลง ตามลำดับ จะทำให้ค่าความเที่ยงของแบบสอบแตกต่างกันไป ซึ่งจะเห็นได้ว่า ผลการศึกษาดังกล่าวส่วนใหญ่ส่วนศึกษากับแบบสอบเลือกตอบที่ถูกต้องเพียงตัวเลือกเดียว (multiple-choice : MC) ยังไม่มีการศึกษาในแบบสอบเลือกตอบแบบถูกผิด (multiple true false : MTF) ผู้วิจัยจึงเห็นว่า การศึกษาเพื่อเปรียบเทียบผลของจำนวนตัวเลือกที่มีต่อค่าฟังก์ชันสารสนเทศของ แบบสอบเลือกตอบแบบถูกผิดในระดับการศึกษาที่ต่างกัน น่าจะเป็นประโยชน์สำหรับการศึกษาด้านการวัดผลทางการศึกษาต่อไป

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย