

บทที่ 2

วรรณคดีที่เกี่ยวข้อง

การศึกษาเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบสอบเลือกตอบตัวเลือกธรรมดา ตัวเลือกผสม และตัวเลือกซ้อนนี้ ผู้วิจัยได้แบ่งเนื้อหาที่เกี่ยวข้องเป็น 3 ตอน คือ ตอนที่ 1 เป็น มโนทัศน์เบื้องต้นแบบสอบเลือกตอบตัวเลือกธรรมดา ตัวเลือกผสม และตัวเลือกซ้อน ตอนที่ 2 เป็นมโนทัศน์เกี่ยวกับทฤษฎีการตอบสนองข้อสอบ และตอนที่ 3 เป็นงานวิจัยที่เกี่ยวข้อง

ตอนที่ 1 มโนทัศน์เบื้องต้นเกี่ยวกับแบบสอบเลือกตอบตัวเลือกธรรมดา ตัวเลือกผสม และตัวเลือกซ้อน

แบบสอบเลือกตอบเป็นแบบสอบที่นิยมใช้กันอย่างแพร่หลาย เพราะมีส่วนดีหลายประการ คือ คำถามมีความเฉพาะเจาะจงจึงไม่ทำให้เกิดความกำกวมของข้อคำถาม ถามคำถามได้จำนวนมาก ข้อการสอบวัดจึงทำได้ครอบคลุมเนื้อหา สามารถวัดสมรรถภาพสมองได้แทบทุกระดับ ตั้งแต่วัดความจำที่ผิวเผินไปจนถึงความสามารถระดับสูงที่ลึกซึ้ง การตรวจให้คะแนนได้ผลคงที่ ประหยัดเวลาและแรงงานในการตรวจ นอกจากนี้ยังสามารถวิเคราะห์ได้ว่าข้อสอบข้อใดดีหรือไม่ดี เพื่อนำผลไปปรับปรุงข้อสอบได้อีกด้วย ดังนั้นความสนใจในเรื่องการพัฒนาแบบสอบเลือกตอบเพื่อใช้เป็นเครื่องมือวัดและประเมินผลการศึกษาจึงเกิดขึ้นอย่างกว้างขวาง และได้มีการพัฒนาวิธีเขียนข้อสอบแบบเลือกตอบ ให้มีรูปแบบข้อคำถามและตัวเลือกแตกต่างกันออกไป เพื่อให้สามารถนำไปใช้ในการวัดผลการเรียนรู้ได้อย่างเหมาะสม

แบบสอบเลือกตอบตัวเลือกธรรมดา

ข้อสอบเลือกตอบโดยทั่วไปจะประกอบด้วย 2 ส่วน คือ ส่วนที่เป็นตอนนำหรือตัวคำถาม (stem) กับส่วนที่เป็นตัวเลือก (choice or option or alternative) ในตัวเลือกประกอบไปด้วยคำตอบถูก (corrector or answer) และตัวลวง (distracters) วิธีสร้างข้อสอบแบบเลือกตอบจึงมีปัญหายุ่งยากอยู่ที่การเขียนตัวคำถามและตัวเลือกเป็นสำคัญ

คำถามของแบบสอบเลือกตอบที่กำหนดให้ผู้สอบเลือกตัวเลือกนั้น สามารถเขียนได้ 2 แบบ คือ คำถามที่ให้ผู้สอบเลือกตัวเลือกที่ถูกจากตัวเลือกอื่น ๆ ที่ผิด กับคำถามที่ให้ผู้สอบเลือกตัวเลือกที่ถูกที่สุดจากตัวเลือกอื่น ๆ ที่ถูกต้องน้อยกว่า ซึ่งข้อคำถามประเภทหลังนี้จะยากกว่า ข้อคำถามประเภทแรก (Gronlund, 1967:140-141)

การเขียนตัวเลือกของแบบสอบเลือกตอบสามารถเขียนได้หลายแบบ ทั้งนี้ต้องเขียนให้สอดคล้องกับคำถามด้วย คำถามที่ให้ผู้สอบเลือกตัวเลือกที่ถูกจากตัวเลือกอื่น ๆ ที่ผิด จะใช้ได้กับตัวเลือกทุกรูปแบบ แต่คำถามที่ให้ผู้สอบเลือกตัวเลือกที่ถูกที่สุดจากตัวเลือกอื่น ๆ ที่ถูกต้องน้อยกว่า จะใช้ได้เหมาะสมในบางสถานการณ์เท่านั้น กล่าวคือ ใช้ได้กับตัวเลือกธรรมดาทั่วไป แต่ไม่ควรใช้กับตัวเลือกปลายปิด "ไม่มีข้อใดถูก" และตัวเลือกปลายปิด "ถูกทุกข้อ" (Thorndike, 1971: 117) เนื่องจากผู้สอบจะไม่เลือกตัวเลือกแบบปลายเปิดเมื่อเห็นว่าตัวเลือกใดตัวหนึ่งถูก หรือผู้สอบจะเลือกตัวเลือกแบบปลายปิดเมื่อพบว่ามียุทธวิธีมากกว่าหนึ่งตัว แต่จะไม่พิจารณาว่าตัวเลือกใดถูกมากหรือน้อยกว่ากัน จากที่กล่าวมาจะเห็นได้ว่าเมื่อใช้คำถามที่มีความซับซ้อนมากขึ้น จะทำให้การเขียนตัวเลือกมีข้อจำกัดมากขึ้น

อย่างไรก็ตามนักวิจัยหลายท่านได้แนะนำเกี่ยวกับวิธีเขียนข้อสอบแบบเลือกตอบ โดยทั่วไปไว้พอสรุปได้ ดังนี้ (วิเชียร เกตุสิงห์, 2520; เอนก พ.อนุกุลบุตร, 2522; อุทุมพร จามรมาน, 2535; Ebel, 1965; Mehrens and Lehmann, 1984)

1. ตัวคำถามต้องมีความชัดเจน อ่านแล้วเข้าใจได้ว่าการถามอะไร
2. พยายามทำให้ข้อความในตัวเลือกเป็นข้อความสั้น ๆ โดยตัดคำที่ไม่จำเป็นออกไป หรือนำคำ/ข้อความที่ซ้ำ ๆ กันในทุกตัวเลือกไปใส่ไว้ในตัวคำถาม
3. ให้แต่ละตัวคำถามเป็นข้อความที่คมีสาระ หลีกเลี่ยงข้อความที่ไม่เกี่ยวข้องกับเรื่องที่ต้องการถาม
4. หากจำเป็นต้องใช้คำถามแบบปฏิเสธ ควรเน้นให้เห็นชัดเจนโดยการขีดเส้นใต้ คำว่า "ไม่" หรือพิมพ์ให้เป็นตัวหนา ๆ หรือใช้อักษรตัวใหญ่ ๆ
5. พยายามใช้คำถามที่วัดความสามารถของสมองชั้นสูง เช่น วัดความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประมาณค่า ไม่ใช่ถามแต่ความจำเพียงอย่างเดียว โดยการเขียนปัญหาหรือสร้างสถานการณ์ที่น่าสนใจขึ้นใหม่ ไม่ลอกข้อความในหนังสือมาถามโดยตรง
6. ข้อสอบแต่ละข้อต้องมีคำตอบที่ถูกต้อง หรือคำตอบที่ถูกที่สุดเพียงคำตอบเดียว



7. ควรทำให้ตัวเลือกต่าง ๆ สอดคล้องเป็นเอกพันธ์กับคำถามและตัวเลือกอื่น ๆ ทั้งทางเนื้อหา ไวยากรณ์ และแบบฟอร์ม

8. ตัวเลือกที่เป็นคำตอบถูกต้องไม่ใช่ชัดเจนผู้ที่ไม่รู้คำตอบก็เดาออก การเขียนตัวเลือกถูกต้องมีลักษณะเป็นคู่ขนานกับตัวเลือกอื่น คือความยาวเท่ากัน เป็นข้อความที่มีลักษณะคล้ายกัน มีเนื้อหาเกี่ยวข้องกัน

9. ตัวเลือกที่เป็นตัวลวง ต้องไม่ผิดจนเห็นได้ชัดเจนและควรมีเหตุผลที่อาจจะเป็นไปได้ทำให้ผู้ผิดใจแล้วน่าจะเป็นตัวถูกทุกตัว สามารถดึงดูดความสนใจของผู้ตอบที่ไม่มีความรู้จริงให้มาเลือกตอบได้

10. การใช้ตัวเลือกประเภท "ไม่มีคำตอบ" หรือ "ถูกทุกข้อ" ควรใช้อย่างระมัดระวังและเลือกใช้ให้เหมาะสมกับลักษณะของคำถาม และเนื้อหาวิชา

11. พยายามอย่าให้คำถามชี้แนะคำตอบ โดยที่คำถามและคำตอบมีเสียงคล้ายกัน หรือซ้ำกัน หรือรูปแบบไวยากรณ์ชี้แนะกัน

12. เขียนตัวเลือกไม่ให้เหลื่อมล้ำ (overlap) กัน

13. จำนวนตัวเลือกควรมีไม่น้อยกว่า 3 ตัวเลือก แต่จะใช้ 4 หรือ 5 ตัวเลือกนั้นขึ้นอยู่กับอายุของผู้ตอบ เนื้อหาวิชาที่สอบ และเวลาที่ทำการสอบ

14. ภาษาที่ใช้ในการเขียนข้อสอบควรเป็นภาษาที่ง่ายและชัดเจน คำศัพท์ที่ใช้ไม่ควรยากเกินไปและควรใช้ภาษาที่เหมาะสมกับเด็ก

15. ตำแหน่งของคำตอบ ควรอยู่ในลักษณะแบบสุ่ม กระจายกันออกไป ไม่ควรอยู่ซ้ำ ๆ กัน หรือเรียงกันอย่างเป็นระบบ

16. ข้อสอบแต่ละข้อควรเป็นอิสระแก่กัน อย่าให้ตัวปัญหาของข้อคำถามคาบเกี่ยวกัน เช่น จะตอบข้อ 2 ได้ต้องตอบข้อ 1 ได้เสียก่อน หรือคำถามข้อแรก ๆ แนะนำคำตอบข้อหลัง ๆ เป็นต้น

17. ตัวเลือกควรเรียงลำดับตามเกณฑ์ใดเกณฑ์หนึ่ง เช่น เรียงตามลำดับเหตุการณ์ เรียงจากมากไปหาน้อย หรือเรียงจากน้อยไปหามาก เป็นต้น

18. ควรใช้แบบฟอร์มของข้อสอบที่มีคุณภาพ

แบบสอบเลือกตอบตัวเลือกผสม

แบบสอบเลือกตอบตัวเลือกผสม มีรูปแบบเหมือนกับแบบสอบเลือกตอบตัวเลือกธรรมดา เพียงแต่ตัวเลือกตัวสุดท้ายจะเขียนในลักษณะการรวมคำตอบ เช่น "ถูกต้องทั้งข้อ ก และข้อ ข" "ถูกต้องทั้งข้อ ก และข้อ ค" หรือ "ถูกต้องทั้งข้อ ข และข้อ ค" เป็นต้น การเขียนตัวเลือกแบบผสมใน

รูปของการรวมคำตอบนี้ ผู้ออกข้อสอบจำนวนมากนิยมนำมาใช้ ซึ่งตัวเลือกในลักษณะนี้มีทั้งข้อดีและข้อเสีย หากเลือกใช้ให้เหมาะสมแล้วก็จะมีคุณค่ามิใช่น้อย ตัวเลือกผสมนี้เมื่อทำหน้าที่เป็นตัวลวงแล้วจะสามารถดึงดูดผู้สอบที่ไม่รู้จักให้ไปเลือกเป็นคำตอบมากขึ้น เพราะเป็นตัวเลือกที่เขียนขึ้นจากการรวมตัวเลือกอื่นที่เป็นไปได้ถึง 2 ตัวไว้ด้วยกัน ทำให้โอกาสในการเดาคำตอบได้ถูกลดลง และทำให้ข้อสอบยากขึ้น การศึกษาถึงคุณภาพของข้อสอบที่ใช้ตัวเลือกแบบผสมจึงน่าจะได้รับการศึกษาให้ชัดเจนยิ่งขึ้น เพื่อจะได้นำไปใช้ได้อย่างถูกต้องเหมาะสม

แบบสอบเลือกตอบตัวเลือกซ้อน

แบบสอบเลือกตอบตัวเลือกซ้อน (double or complex multiple choice) เป็นรูปแบบตัวเลือกที่นักจิตพลพชาสามปรับปรุงแก้ไขข้อบกพร่องของการวัด ที่บางเนื้อหาวิชาหรือบางพฤติกรรมที่ต้องการวัดนั้นมีคำตอบถูกหลายคำตอบ และแต่ละคำตอบนั้นต่างก็เป็นความจริง หรือมีความสำคัญเท่าเทียมกัน และเราต้องการวัดว่าผู้สอบรู้คำตอบเหล่านั้นครบถ้วนหรือไม่ หากเรานำแบบสอบเลือกตอบตัวเลือกธรรมดามาใช้แล้ว เราจะวัดสิ่งที่สำคัญเหล่านั้นได้เพียงข้อละอย่างเดี๋ยวนั้น หรือไม่ก็ต้องถามคำถามซ้ำ ๆ จนครบ ซึ่งเป็นสิ่งที่ไม่ควรกระทำ จึงทำให้เกิดความพยายามที่จะหารูปแบบของข้อสอบที่จะแก้ปัญหาดังกล่าว เช่น การเขียนตัวเลือกที่มีคำตอบถูกไม่จำกัด แต่ก็ต้องพบกับปัญหาในเรื่องการตรวจและการให้คะแนน ต่อมาจึงมีการดัดแปลงตัวเลือกให้อยู่ในรูปของการรวมคำตอบและเขียนเป็นแบบตัวเลือกซ้อน ซึ่งมีผู้นิยมนำมาใช้อย่างแพร่หลาย เพราะสามารถวัดได้ถึงขั้นวิเคราะห์ สังเคราะห์ และความคิดรวบยอดของเนื้อหา นั้น ๆ ได้เป็นอย่างดี สามารถนำมาใช้ได้อย่างดีกับวิชาสังคมศึกษา คณิตศาสตร์ วิทยาศาสตร์ โดยเฉพาะกับเนื้อหาที่มีคำตอบถูกได้หลายอย่าง หรือมีสาเหตุที่เกี่ยวข้องกันอยู่หลายประการ (ชวาล แพร์ตกุล, มปป. : 148) ซึ่งเป็นความคิดเห็นที่สอดคล้องกับ Thorndike and Hagen (1977: 240) ที่มีความเห็นว่า แบบสอบเลือกตอบตัวเลือกซ้อนใช้ได้อย่างมีประสิทธิภาพเมื่อใช้วัดเกี่ยวกับการใช้ความรู้หลักการหรือความสามารถในการประยุกต์ใช้ หรือความสัมพันธ์ระหว่างเหตุและผลหลาย ๆ อย่าง นอกจากนั้นแบบสอบเลือกตอบตัวเลือกซ้อนยังใช้ลดความคลาดเคลื่อนจากการใช้ข้อสอบที่ใช้ตัวเลือกแบบปลายปิดให้หมดไปได้ โดยการเขียนข้อสอบให้อยู่ในรูปแบบตัวเลือกซ้อน (Copperud, 1979: 56) ในด้านการตรวจให้คะแนนก็สามารถทำได้สะดวกรวดเร็ว และสามารถนำผลการสอบมาวิเคราะห์ เพื่อปรับปรุงข้อสอบได้เช่นเดียวกับแบบสอบเลือกตอบตัวเลือกธรรมดา

แบบสอบเลือกตอบตัวเลือกซ้อนประกอบด้วย 3 ส่วน คือ ส่วนที่เป็นคำถาม ส่วนที่เป็นข้อความหรือสถานการณ์และส่วนที่เป็นตัวเลือก

ส่วนที่เป็นคำถาม เขียนโดยใช้หลักการเกี่ยวกับการเขียนข้อความของข้อสอบแบบตัวเลือกธรรมดา ส่วนที่เป็นข้อความหรือสถานการณ์จะเขียนเช่นเดียวกับการเขียนตัวเลือกแบบธรรมดา โดยที่ข้อความหรือสถานการณ์นั้นอาจเป็นคำตอบถูกเพียง 1 สถานการณ์หรือถูกมากกว่า 1 สถานการณ์ก็ได้ (อาจถูก 2, 3 สถานการณ์หรือถูกทั้งหมด) และส่วนที่เป็นตัวเลือกจะเขียนจากการรวมข้อความหรือสถานการณ์ที่เขียนไว้แล้วนั้น มาผสมกันเป็นชุด ๆ ตามความเหมาะสมจากวิธีการดังกล่าวจะทำให้สามารถเขียนตัวเลือกที่เป็นไปได้หลายตัวเลือก เช่น เมื่อกำหนดข้อความหรือสถานการณ์มาให้ 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 และไม่มีข้อใดถูกในทำนองเดียวกันเมื่อกำหนดข้อความหรือสถานการณ์มา 4 หรือ 5 ข้อ จะสามารถเขียนตัวเลือกได้ 16 หรือ 32 ตัวเลือกตามลำดับ การที่สามารถสร้างตัวเลือกได้มากมายเช่นนี้ ทำให้ผู้สร้างข้อสอบสามารถเลือกตัวเลือกที่น่าสนใจใส่ไว้ในข้อสอบได้อย่างเหมาะสม แต่มีข้อควรพิจารณาในการเลือกตัวเลือกไว้ในข้อสอบ คือ ควรเป็นตัวเลือกที่ทำให้ข้อความหรือสถานการณ์ปรากฏในตัวเลือกทั้งหมดด้วยจำนวนครั้งที่เท่ากัน (Whitley, 1979 อ้างถึงใน Kolsted and Other, 1983:45) ทั้งนี้เพราะข้อความหรือสถานการณ์ที่ปรากฏในตัวเลือกหลายครั้งจะมีผลต่อการเดาของผู้สอบ ในการสร้างแบบสอบวัดผลสัมฤทธิ์ควรระวังถึงหลักเกณฑ์ข้อนี้ แต่ไม่ควรยึดเป็นกฎเกณฑ์ตายตัว เพราะต้องพิจารณาความเป็นไปได้ของเนื้อหาที่ออกข้อสอบด้วย มิฉะนั้นตัวเลือกบางตัวจะเป็นตัวเลือกที่ผิดชัดเจน และไม่มีผู้ใดเลือกตัวเลือกนั้นเลย นอกจากนั้น ชวาล แพริตกุล (มปป., 154) ยังแนะนำว่าไม่ควรใช้ตัวเลือกประเภทปลายเปิด "ไม่มีข้อใดถูก" และตัวเลือกที่มีความหมายขัดแย้งกันเองภายในตัวเลือกเดียวกัน เช่น "1 ถูกแต่ 2 ผิด" ในแบบสอบตัวเลือกซ้อนนี้

Weiten (1982:47) กล่าวถึงข้อจำกัดของแบบสอบเลือกตอบตัวเลือกซ้อนไว้

3 ประการ ดังนี้

1. การเขียนข้อสอบเลือกตอบตัวเลือกซ้อนไม่เป็นไปตามกฎเกณฑ์การเขียนข้อสอบเลือกตอบที่ว่าควรเขียนให้สั้นและกระชับ ซึ่งจะมีผลต่อการตอบถูกหรือผิดของผู้สอบด้วย
2. ความยาวและความซับซ้อนของข้อสอบ มีผลให้ผู้ที่มักชะลอการอ่านและการให้เหตุผลจะตอบแบบสอบตัวเลือกซ้อนได้ดีกว่า ซึ่งเป็นการเพิ่มความแปรปรวนคลาดเคลื่อน (error variance)

3. ความสามารถในการทำแบบสอบเลือกตอบตัวเลือกซ้อนขึ้นอยู่กับความชำนาญในการทำแบบสอบ (test-wisness) และทักษะในการทำแบบสอบ (test taking skill) ของผู้สอบแต่ละคน

ตอนที่ 2 ทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบเป็นทฤษฎีที่อธิบายถึงความสัมพันธ์ระหว่างลักษณะหรือความสามารถที่มีอยู่ภายในตัวบุคคลกับพฤติกรรมการตอบสนองต่อข้อสอบของบุคคลนั้น โดยมีความเชื่อว่า พฤติกรรมการตอบสนองข้อสอบของผู้เข้าสอบซึ่งเป็นสิ่งที่สังเกตได้โดยตรง จะถูกกำหนดโดยคุณลักษณะ (Trait) หรือความสามารถ (Ability) ที่มีอยู่ในตัวบุคคลซึ่งไม่สามารถสังเกตได้ ทฤษฎีนี้พยายามอธิบายความสัมพันธ์ดังกล่าวให้แสดงออกมาในรูปแบบจำลองทางคณิตศาสตร์ โดยให้คะแนนที่ได้รับจากการตอบข้อสอบ (y) แทนพฤติกรรมการตอบสนองข้อสอบ และ θ แทนคุณลักษณะหรือความสามารถภายในตัวบุคคล ฟังก์ชันการถดถอยของ y บน θ จะเป็นฟังก์ชันทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างพฤติกรรมการตอบสนองข้อสอบกับระดับความสามารถ เรียกว่า ฟังก์ชันการตอบสนองข้อสอบ (Item Response Function) (Lord and Novick, 1968:360) ซึ่งกำหนดได้หลายรูปแบบ ทั้งนี้ขึ้นอยู่กับข้อตกลงเบื้องต้นเกี่ยวกับข้อมูลจากการทดสอบ (Hambelton and cook, 1977:75)

ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

1. มิติของลาเทนท์สเปซ (Dimensionality of Latent Space)

ลาเทนท์สเปซ หมายถึง ลักษณะหรือความสามารถที่เป็นตัวกำหนดพฤติกรรมการตอบสนองข้อสอบข้อใดข้อหนึ่งอาจมีได้หลายลักษณะ ลักษณะทั้งหมดนี้รวมเรียกว่า ลาเทนท์สเปซ (Latent Space) จำนวนลักษณะทั้งหมดในลาเทนท์สเปซก็คือ มิติของลาเทนท์สเปซ ในทฤษฎีการตอบสนองข้อสอบโดยทั่วไปมักจะถือว่า ลักษณะหรือความสามารถที่เป็นตัวกำหนดพฤติกรรมการตอบสนองข้อสอบมีเพียงลักษณะเดียว และยอมรับเป็นข้อตกลงเบื้องต้นที่สำคัญประการหนึ่งของแบบจำลองหลายแบบในทฤษฎีนี้ เรียกว่า Unidimensionality Assumption การกำหนดข้อตกลงเช่นนี้ อาจขัดแย้งกับความจริงที่ทราบกันดีว่า ผลการตอบข้อสอบนั้นขึ้นอยู่กับการปัจจัยหลายอย่างประกอบกัน แต่อย่างไรก็ตาม ข้อตกลงนี้เพียงแต่ต้องการว่า มีความสามารถหรือคุณลักษณะที่เด่นและสำคัญ (dominant factor) เพียงลักษณะเดียวที่เป็นตัวกำหนดพฤติกรรมการตอบสนองข้อสอบ นั่นคือ ข้อสอบในแบบสอบมีลักษณะเป็นเอกพันธ์ (Homogeneous) การถือว่า

ลาเทนท์สเปซที่สนใจมีเพียงมิติเดียวจะช่วยให้แบบจำลองของทฤษฎีมีความซับซ้อนน้อยลง และง่ายแก่การแปลความหมายของคะแนนจากแบบสอบ อย่างไรก็ตามมีผู้พยายามที่จะพัฒนาแบบจำลองที่ไม่ต้องมีข้อตกลงเบื้องต้นที่ว่า ลาเทนท์สเปซมีเพียงมิติเดียวซึ่งเรียกว่า Multidimension Item Response Models ซึ่งแม้ว่าจะเป็นแบบจำลองที่สมเหตุสมผลมากกว่า แต่การพัฒนาแบบจำลองเหล่านี้ยังไม่ถึงระดับที่จะนำไปประยุกต์ใช้ได้ทางปฏิบัติ (Hambleton and Swaminathan, 1985:17)

2. ความเป็นอิสระในการตอบสนองข้อสอบ (Local Independence) หมายความว่า การตอบสนองข้อสอบต่าง ๆ ในแบบสอบของบุคคลใดบุคคลหนึ่งมีความเป็นอิสระในเชิงสถิติ นั่นคือ การตอบสนองข้อสอบข้อหนึ่งไม่มีผลต่อการตอบสนองข้อสอบข้ออื่น ๆ ในแบบสอบ เช่น เนื้อหาของข้อคำถามข้อหนึ่งต้องไม่ขึ้นหาค่าตอบให้ข้อสอบข้ออื่น ๆ ถ้าการตอบข้อสอบมีความเป็นอิสระในการตอบข้อสอบแล้ว ความน่าจะเป็นร่วมของคะแนนรายข้อของผู้เข้าสอบแต่ละคน คือ (Hambleton and Swaminathan, 1985:23)

$$\text{Prob } [U_1 = u_1, U_2 = u_2, \dots, U_n = u_n / \theta] = \prod_{i=1}^n P_i(\theta)^{u_i} Q_i(\theta)^{1-u_i}$$

เมื่อ $U_i, i=1, 2, \dots, n$ เป็นผลการตอบข้อสอบ n ข้อ (เมื่อตอบถูก $u_i=1$, เมื่อตอบผิด $u_i=0$)

P_i = ความน่าจะเป็นของผู้สอบคนหนึ่งที่ทำข้อสอบข้อที่ i ถูก

Q_i = $1 - P_i$

ข้อตกลงดังกล่าวเป็นข้อตกลงเบื้องต้นอีกประการหนึ่งของทฤษฎีการตอบสนองข้อสอบในทุกแบบจำลอง ซึ่งเป็นประโยชน์ต่อการประมาณค่าพารามิเตอร์ของข้อสอบและการประมาณค่าความสามารถของผู้สอบ

3. โค้งลักษณะข้อสอบ (Item Characteristic Curve : ICC) เป็นฟังก์ชันแสดงความสัมพันธ์ระหว่างระดับความสามารถของผู้ตอบข้อสอบกับความน่าจะเป็นในการตอบข้อสอบข้อนั้นถูก ซึ่งเป็นฟังก์ชันการถดถอยชนิดไม่เป็นเส้นตรง (Non-Linear Regression) โค้งลักษณะข้อสอบมีหลายรูปแบบ ทั้งนี้ขึ้นอยู่กับความเชื่อในแบบจำลองที่ใช้อธิบายความสัมพันธ์ดังกล่าว

แบบจำลองที่ใช้ในทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบได้มีการพัฒนาแบบจำลองขึ้นมาหลายแบบด้วยกัน โดยแต่ละแบบจะแตกต่างกันในรูปของฟังก์ชันทางคณิตศาสตร์และจำนวนพารามิเตอร์ที่กล่าวถึง และสามารถจำแนกได้เป็น 3 ประเภทใหญ่ๆ คือ (Hambleton and Swaminathan, 1985: 35-52)

1. แบบจำลองที่ประยุกต์ใช้กับการให้คะแนนแบบ 2 ค่า (Binary Item) คือ ตอบถูกได้ 1 คะแนน ตอบผิดได้ 0 คะแนน สามารถแบ่งตามลำดับขั้นการพัฒนาได้ 2 ระยะ คือ

1.1 แบบจำลองในระยะเริ่มแรก (ค.ศ. 1943-1968) ได้แก่ Guttman Perfect Scale, Latent Distance Model และ Linear Model แบบจำลองเหล่านี้เป็นแบบจำลองที่ไม่สู้จะมีความสอดคล้องกับข้อมูลจากการสอบในสถานการณ์จริงมากนัก แต่ก็ยังเป็นแบบจำลองที่มีคุณค่าต่อการพัฒนาแบบจำลองที่เหมาะสมในระยะต่อมา

1.2 แบบจำลองที่ได้รับการพัฒนาในระยะหลัง (ค.ศ. 1952-1982) ได้แก่ One-, Two-, Three - Parameter Normal Ogive Model และ One-, Two-, Three -, Four-Parameter Logistic Model เป็นแบบจำลองที่ได้รับการปรับปรุงให้มีความสอดคล้องกับข้อมูลจากการสอบในสถานการณ์จริงมากขึ้น สามารถนำไปปฏิบัติได้ง่าย และมีผู้สนใจนำไปใช้อย่างแพร่หลาย

2. แบบจำลองที่ประยุกต์ใช้กับการให้คะแนนที่มากกว่า 2 ค่า (Multichotomously Scored) ได้แก่ Nominal Response Model, Grade Response Model และ Partial Credit Model ซึ่งแบบจำลองเหล่านี้มีจุดมุ่งหมายเพื่อเพิ่มความแม่นยำในการประมาณค่าความสามารถของผู้เข้าสอบ โดยใช้ประโยชน์จากสารสนเทศทั้งจากการตอบข้อสอบถูกหรือผิด

3. แบบจำลองที่ประยุกต์ใช้กับการให้คะแนนแบบต่อเนื่อง (Continuous) ได้แก่ Continuous Response Model ที่พัฒนาโดย Samejima ในปี 1972 ซึ่งจะ เป็นประโยชน์ต่อการประมาณค่าความสามารถของผู้เข้าสอบทางจิตวิทยา หรือผู้ที่สนใจศึกษาทางด้านที่เกี่ยวกับทัศนคติ (Attitude) ซึ่งการตอบสนองต่อข้อสอบเป็นมาตราที่ต่อเนื่อง (Continuous Scale)

ในที่นี้ ผู้วิจัยจะนำเสนอรายละเอียดเฉพาะแบบจำลองที่มีการประยุกต์ใช้กับการให้คะแนนแบบ 2 ค่า ที่ได้พัฒนาขึ้นในระยะหลัง กำลังเป็นที่นิยมในปัจจุบันและเหมาะสมกับข้อมูลจากผลการข้อสอบแบบเลือกตอบที่ให้คะแนนเป็น 0-1 คือ Normal Ogive Model และ Logistic Model ซึ่งแบบจำลอง Three Parameter Logistic Model ที่ผู้วิจัยนำมาเป็นกรอบในการศึกษาถึงคุณภาพของแบบสอบนี้เป็นแบบจำลองหนึ่งใน Logistic Model และจะได้กล่าวถึงรายละเอียดต่อไป

Normal Ogive Model แบบจำลองนี้แสดงความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบถูกกับระดับความสามารถอยู่ในรูปของฟังก์ชันการแจกแจงสะสมปกติ (cumulative normal ogive) ซึ่งแสดงได้ด้วยสมการดังนี้ (Lord and Novick, 1968:366)

$$P_j(\theta) = \int_{-\infty}^{a_j(\theta - b_j)} \phi(t) dt$$

เมื่อ $P_j(\theta)$ คือโอกาสที่ผู้สอบซึ่งมีระดับความสามารถ θ จะตอบข้อสอบข้อที่ j ได้ถูก $\phi(t)$ แทน Normal frequency function

a_j และ b_j เป็นค่าพารามิเตอร์อำนาจจำแนก และความยากที่แสดงถึงลักษณะของข้อสอบ

เส้นโค้งลักษณะของข้อสอบจะมีลักษณะคล้ายรูปตัว S โดยที่ค่าโอกาสของการตอบข้อสอบถูกจะเข้าใกล้ 0 เมื่อ θ มีค่าน้อย ๆ และจะมีค่าเพิ่มขึ้นเรื่อย ๆ โดยจะเข้าใกล้ 1 เมื่อ θ มีค่ามาก ๆ ในสมการข้างต้นมีค่าพารามิเตอร์ของข้อสอบ 2 ตัว คือ a_j กับ b_j จึงเรียกแบบจำลองที่อธิบายด้วยสมการนี้ว่า Two-Parameter Normal Ogive Model แต่ในการสอบที่ใช้ข้อสอบแบบเลือกตอบซึ่งผู้เข้าสอบที่มีความสามารถน้อย ๆ ก็มีโอกาสดตอบข้อสอบถูกได้โดยการเดา แบบจำลองที่มีพารามิเตอร์ 2 ตัว จะไม่เหมาะที่จะใช้อธิบายสถานการณ์การสอบเช่นนั้น เพราะในสถานการณ์ดังกล่าว เมื่อ θ มีค่าน้อย ๆ ค่า $P_j(\theta)$ จะไม่เป็น 0 แต่จะมีค่าเข้าใกล้ค่า ๆ หนึ่งที่มากกว่า 0 จึงได้มีการเพิ่มค่าพารามิเตอร์ของข้อสอบเข้าไปอีก 1 ตัว คือ c_j และมีสมการดังนี้

$$P_j(\theta) = c_j + (1 - c_j) \int_{-\infty}^{a_j(\theta - b_j)} \phi(t) dt$$

แบบจำลองนี้เรียกว่า Three - Parameter Normal Ogive Model และ c_j ก็คือค่าพารามิเตอร์ที่แสดงถึงโอกาสที่บุคคลที่มีความสามารถต่ำมากจะตอบข้อสอบถูก เรียกว่าค่าการเดา

อย่างไรก็ตามการนำ Normal Ogive Model มาใช้จริงในทางปฏิบัติเป็นไปได้ยาก เนื่องจากความสลับซับซ้อนทางคณิตศาสตร์ที่เป็นพื้นฐานของทฤษฎี ในระยะต่อมาได้มีการพัฒนา

Logistic Model ขึ้น และได้รับความนิยมอย่างแพร่หลาย เพราะเป็นแบบจำลองที่มีลักษณะคล้ายกับ Normal Ogive Model มาก แต่มีวิธีการคำนวณค่าพารามิเตอร์ต่าง ๆ ง่ายกว่า (Lord and Novick, 1968:366) แต่แบบจำลองดังกล่าวก็นับว่ามีประโยชน์มากสำหรับการพัฒนา Logistic Model ในระยะต่อมา

Logistic model

ผู้ที่มีบทบาทสำคัญในการพัฒนาแบบจำลองนี้ คือ Birnbaum (Hambelton and Swaminathan, 1985:36) ซึ่งแสดงความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบถูกต้องกับระดับความสามารถอยู่ในรูปของฟังก์ชันการแจกแจงสะสมแบบโลจิสต์ (Logistic cumulation distribution function) มีเส้นโค้งลักษณะของข้อสอบเป็นรูปตัว S เช่นเดียวกับ Normal Ogive Model และเมื่อปรับค่าตัวแปรด้วย Scaling factor ซึ่งมีค่าเท่ากับ 1.7 แล้ว โค้งลักษณะข้อสอบของแบบจำลองทั้งสองมีความแตกต่างกันไม่เกิน .01 สำหรับทุก ๆ ค่าของความสามารถ แต่ในแง่ของการคำนวณ Logistic Model มีความง่ายและสะดวกกว่ามาก จึงเป็นที่นิยมในการนำไปใช้ปฏิบัติงานจริง

Logistic model ได้รับการพัฒนาขึ้นมาใช้ปฏิบัติงานอยู่ในปัจจุบัน 3 รูปแบบ คือ One-, Two-, Three-Parameter Logistic model ดังนี้

1. One-Parameter Logistic model เป็นแบบจำลองที่ถือว่าข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากันหมด ($a_j = \bar{a}$) และกำหนดให้ค่าพารามิเตอร์การเดา $c_j = 0$ คงมีค่าพารามิเตอร์ความยากเท่านั้นที่กำหนดคุณลักษณะของข้อสอบ มีสมการดังนี้

$$P_j(e_j) = \frac{1}{1 + e^{-1.7 a_j (e_j - b_j)}}$$

2. Two - Parameter logistic model เป็นแบบจำลองที่เชื่อว่าผู้สอบมีโอกาสตอบข้อสอบถูกต้องด้วยการเดาน้อยมาก จึงกำหนดให้ค่าพารามิเตอร์การเดา (c_j) เป็น 0 และมีค่าพารามิเตอร์อำนาจจำแนก (a_j) และค่าความยาก (b_j) 2 ตัว เป็นตัวกำหนดคุณลักษณะข้อสอบ มีสมการดังนี้

$$P_j(e_j) = \frac{1}{1 + e^{-1.7 a_j (e_j - b_j)}}$$

3. Three - Parameter logistic model เป็นรูปแบบที่เชื่อว่าโอกาสที่ผู้สอบจะตอบข้อสอบถูกต้องขึ้นอยู่กับพารามิเตอร์ 3 ตัว คือ ค่าอำนาจจำแนก (a_j) ค่าความยาก (b_j) และค่าการเดา (c_j) มีสมการดังนี้

$$-1.7 a_j (e_1 - b_j)$$

$$P_j(e_1) = c_j + (1-c_j) / [1 + e^{-1.7 a_j (e_1 - b_j)}]$$

$P_j(e_1)$ คือ โอกาสที่ผู้สอบซึ่งที่ระดับความสามารถ (e_1) ตอบข้อสอบข้อที่ j ได้ถูกต้อง

a_j คือ ค่าอำนาจจำแนกของข้อที่ j โดยทฤษฎีแล้ว a จะมีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$ ข้อสอบที่ใช้ได้ควรมีค่า a เป็นบวก โดยปกติจะมีค่าไม่เกิน 2.00

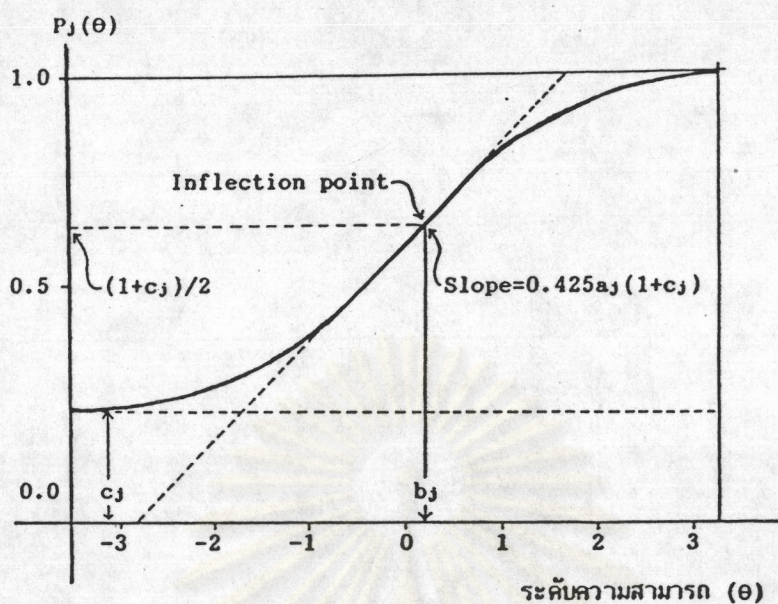
b_j คือ ค่าความยากของข้อสอบข้อที่ j เป็นค่าของระดับความสามารถ e ระดับหนึ่ง ที่ทำให้เส้นโค้งลักษณะของข้อสอบมีการเปลี่ยนโค้ง เป็นค่าพารามิเตอร์ที่บอกตำแหน่งของเส้นโค้งลักษณะของข้อสอบ ณ $e = b$ โอกาสที่จะตอบข้อสอบถูกต้องจะมีค่า $(1-c_j)/2$ หรือเท่ากับ 0.5 ถ้า $c_j = 0$ ค่า b มีค่าที่เป็นไปได้ตั้งแต่ $-\infty$ ถึง $+\infty$ ในทางปฏิบัติมักจะพิจารณากันในช่วง -2.5 ถึง $+2.5$

c_j คือ ค่าการเดา (guessing parameter or pseudo - chance score level) เป็นค่ากำกับเส้นโค้งที่ต่ำสุด (Lower asymptote) ของเส้นโค้งลักษณะข้อสอบ แสดงถึงโอกาสที่ผู้สอบที่มีความสามารถต่ำสุดจะตอบข้อสอบถูกต้อง มีค่าที่เป็นไปได้อยู่ระหว่าง 0 ถึง 1 ข้อสอบที่มีค่าการเดาสูงเป็นข้อสอบที่ไม่ดี และข้อสอบที่ใช้ควรมีค่าการเดาไม่เกิน 0.30 (Urry, 1977 : 184)

e คือ ระดับความสามารถของผู้สอบ โดยทฤษฎีแล้วจะมีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$ ในทางปฏิบัติจะมีค่าอยู่ระหว่าง -3.00 ถึง $+3.00$

e คือ ค่าเอ็กซ์โปเนนเชียล เป็นค่าคงที่มีค่าเท่ากับ 2.7182

แบบจำลองโลจิสติกที่มีพารามิเตอร์ 3 ตัวนี้ เขียนเป็นเส้นโค้งลักษณะของข้อสอบจะมีลักษณะดังนี้ (Hambleton and Swaminathan, 1985:39)



ภาพที่ 1 เส้นโค้งลักษณะข้อสอบ (ICC) แบบ 3 พารามิเตอร์ จะเห็นว่า

1. ค่าอำนาจจำแนก (a_j) จะเป็นสัดส่วนกับความชัน (Slope) ของโค้ง ณ จุดเปลี่ยนโค้ง หรือจุดที่มีความชันมากที่สุด [ความชัน ณ จุดนี้มีค่า = $.425 a_j (1 - c_j)$]
2. ค่าความยาก (b_j) จะอยู่บนแกนความสามารถ (θ) ตรงจุดตัดกับเส้นที่ลากมาจากจุดเปลี่ยนโค้งมาตั้งฉากกับแกนความสามารถ θ
3. ค่าการเดา (c_j) เป็นส่วนตัดบนแกนตั้งของโค้งลักษณะข้อสอบ

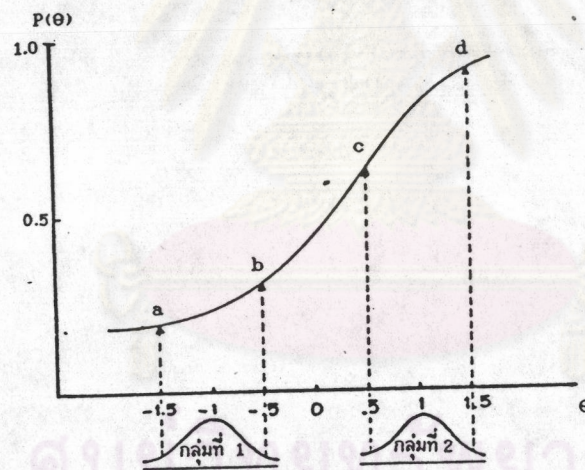
ลักษณะเด่นของทฤษฎีการตอบสนองข้อสอบ

หากข้อมูลจากผลการสอบกับแบบจำลองการตอบสนองข้อสอบมีความเหมาะสมกันแล้ว ทฤษฎีการตอบสนองข้อสอบ จะมีความเหนือกว่า (Superiority) ทฤษฎีการสอบแบบคลาสสิก (Classical Test Theory) ดังนี้ (Hambleton and Swaminathan, 1985:10-13)

1. ความเป็นอิสระจากกลุ่มผู้เข้าสอบ (Sample - free) นั่นคือ ไม่ว่าจะนำข้อสอบไปใช้กับบุคคลใด โค้งลักษณะข้อสอบก็จะคงเดิม
2. ความเป็นอิสระจากกลุ่มข้อสอบ (Item - free) ในการประมาณค่าความสามารถ (θ) ของผู้เข้าสอบ นั่นคือ จะใช้ข้อสอบชุดใดก็ได้ จำนวนเท่าใดก็ได้ ซึ่งบางครั้งอาจใช้ข้อสอบเพียง 3-5 ข้อ ก็สามารถประมาณค่าความสามารถที่แท้จริงของผู้สอบได้แล้ว ทั้งนี้ข้อสอบต้องได้รับการคัดเลือกมาจากคลังข้อสอบขนาดใหญ่ ที่ข้อสอบแต่ละข้อระบุคุณลักษณะที่วัดได้

3. ความไม่แปรเปลี่ยนของค่าพารามิเตอร์ (Invariant of Parameter)

กล่าวคือ ไม่ว่าจะประมาณค่าพารามิเตอร์จากกลุ่มตัวอย่างที่มีความสามารถระดับใดก็ตาม ค่าพารามิเตอร์จะไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ ทั้งนี้เพราะ ฟังก์ชันการตอบสนองต่อข้อสอบซึ่งเป็นฟังก์ชันที่อธิบายความสัมพันธ์ระหว่างพฤติกรรมการตอบสนองข้อสอบกับระดับความสามารถนั้น เป็นฟังก์ชันการถดถอย (Regression function) ของคะแนนจากการสอบบนความสามารถ ซึ่งตามทฤษฎีสถิติ ฟังก์ชันการถดถอยจะไม่แปรเปลี่ยนไปตามการแจกแจงความถี่ของตัวแปรทำนาย ดังนั้น พารามิเตอร์ต่าง ๆ ในสมการถดถอยซึ่งเป็นตัวกำหนดรูปร่างของเส้นถดถอยจะไม่เปลี่ยนแปลงด้วย (Lord, 1980:34) ไม่ว่าจะประมาณค่าจากกลุ่มผู้สอบกลุ่มใดก็ตาม ความเข้าใจในเรื่องดังกล่าว Baker (1977:170-172) ได้เขียนกราฟและอธิบายเกี่ยวกับค่าพารามิเตอร์ของข้อสอบข้อเดียวกัน ระหว่างกลุ่มผู้สอบที่มีความสามารถต่างกัน ดังนี้



ภาพที่ 2 แสดงความไม่แปรเปลี่ยนของค่าพารามิเตอร์

จากภาพที่ 2 แสดงให้เห็นถึงการประมาณค่าจากกลุ่มที่ 1 ซึ่งเป็นกลุ่มที่มีความสามารถต่ำ เส้นโค้งลักษณะของข้อสอบที่ได้จะอยู่ในช่วง a - b ซึ่งจะตกอยู่ในช่วงพิสัยของความสามารถของกลุ่มที่ 1 เส้นโค้งลักษณะของข้อสอบในช่วงนี้จะบอกถึงโอกาสที่ผู้สอบกลุ่มที่ 1 ซึ่งเป็นกลุ่มที่มีความสามารถต่ำจะตอบข้อสอบข้อนั้นถูก ส่วนเส้นโค้งลักษณะข้อสอบในช่วง c - d ก็แสดงถึงโอกาสที่ผู้สอบกลุ่มที่ 2 ซึ่งเป็นกลุ่มที่มีความสามารถระดับสูงจะตอบข้อสอบข้อนั้นถูก ถ้าข้อสอบข้อนี้

วัดความสามารถเดียวกันของกลุ่มตัวอย่างทั้งสองกลุ่มแล้ว เส้นโค้งลักษณะข้อสอบย่อมจะต้องมีเพียงเส้นเดียว เส้นโค้งช่วง $a - b$ และช่วง $c - d$ ย่อมเป็นส่วนหนึ่งของเส้นโค้งลักษณะของข้อสอบข้อนี้ เมื่อเส้นโค้งลักษณะข้อสอบมีได้เส้นเดียว ตัวพารามิเตอร์ที่กำหนดรูปร่างลักษณะของเส้นโค้งนี้ย่อมมีค่าได้ตัวละ 1 ค่าเท่านั้น นั่นคือ การประมาณค่าพารามิเตอร์ของข้อสอบที่ได้จากกลุ่มตัวอย่างทั้งสองควรจะเป็นค่าเดียวกัน

แต่อย่างไรก็ตามการไม่แปรเปลี่ยนของค่าพารามิเตอร์นี้ไม่ได้หมายความว่า ค่าพารามิเตอร์ที่ได้จากการประมาณค่าโดยใช้กลุ่มตัวอย่างที่ต่างกันจะมีค่าเท่ากันพอดี ค่าพารามิเตอร์ที่ประมาณโดยการใช้กลุ่มตัวอย่างที่ต่างกันจะมีค่าเท่ากันหรือไม่ขึ้นอยู่กับเงื่อนไขการทดสอบบางประการ เช่น ถ้าเลือกมาตรวัดที่มีจุดเริ่มต้นเดียวกัน และหน่วยในการวัดหน่วยเดียวกันแล้ว ค่าพารามิเตอร์ที่ประมาณได้จากกลุ่มตัวอย่างที่แตกต่างกันจะมีค่าเท่ากัน ในทางตรงกันข้าม ถ้าเลือกมาตรวัดที่มีจุดเริ่มต้น และมีหน่วยในการวัดแตกต่างกันแล้ว การไม่แปรเปลี่ยนของค่าพารามิเตอร์จะหมายความว่า ค่าพารามิเตอร์ที่ประมาณได้จากกลุ่มตัวอย่างที่ต่างกันของข้อสอบชุดหนึ่ง จะมีความสัมพันธ์กันในเชิงเส้นตรง (Lord, 1980:36)

ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ (Item and Test Information Function)

สารสนเทศนิยามได้จากความไม่แน่นอนของเหตุการณ์บางประการ ถ้าความไม่แน่นอนในการเกิดเหตุการณ์หนึ่งมีน้อย แสดงว่ามีสารสนเทศเกี่ยวกับเหตุการณ์นั้นมาก จากนิยามนี้ได้ถูกนำไปใช้เป็นมโนทัศน์ทั่วไปในทฤษฎีการตอบสนองข้อสอบ โดยพิจารณาจากกระบวนการอ้างอิงเชิงสถิติ ในกรณีที่ไม่มีสารสนเทศใด ๆ เกี่ยวกับประชากรเลย ความไม่แน่นอนในการประมาณค่าพารามิเตอร์ของประชากรในเชิงทฤษฎีอาจจะต้องประมาณเป็นค่าใด ๆ ที่มีค่าอยู่ระหว่าง $-\alpha$ ถึง $+\alpha$ แต่หากมีสารสนเทศเกี่ยวกับประชากรบ้างช่วงของการประมาณค่าจะแคบลง นั่นคือ ความถูกต้องแม่นยำในการประมาณค่าจะเริ่มมีมากขึ้น ตามปกติความแม่นยำในการประมาณค่าพารามิเตอร์จะแสดงได้ด้วยค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า กล่าวคือ ถ้าความคลาดเคลื่อนมาตรฐานของการประมาณค่ามีมาก ความแม่นยำถูกต้องของการประมาณค่าก็จะมีน้อย เพราะช่วงของค่าที่ประมาณจะกว้าง ในทางตรงกันข้ามถ้าความคลาดเคลื่อนมาตรฐานของการประมาณค่ามีน้อย ความแม่นยำในการประมาณค่าก็จะมีมาก ช่วงของค่าประมาณจะแคบ แสดงว่าค่าสารสนเทศมีความสัมพันธ์กับความคลาดเคลื่อนมาตรฐานของการประมาณค่า โดยมีความสัมพันธ์ในทิศทางกลับกัน ดังสมการ (warm, 1978:77)

$$SE(\theta) = 1 / \sqrt{I(\theta)}$$

$SE(\theta)$ คือความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ θ ระดับความสามารถ θ

Birnbaum (1968:418) ได้นิยามฟังก์ชันสารสนเทศของแบบสอบว่า เป็นดัชนีที่บ่งชี้ถึงความถูกต้องแม่นยำในการประมาณค่าความสามารถ มีค่าเท่ากับอัตราส่วนกำลังสองของความชันของเส้นถดถอยบนคะแนนความสามารถจริงต่อความแปรปรวนของคะแนน

ถ้าให้ y เป็นคะแนนที่ได้จากการตอบแบบสอบ ค่าสารสนเทศสำหรับคะแนน y แสดงได้ดังสมการต่อไปนี้

$$I(\theta, y) = [d_{y/\theta}]^2 / \text{var}(y/\theta)$$

จากสูตรจะเห็นว่าค่าสารสนเทศสำหรับคะแนน y ก็คือ กำลังสองของอัตราส่วนระหว่างความชันของเส้นถดถอยของ y บน θ กับความคลาดเคลื่อนมาตรฐานของการวัด y θ ระดับความสามารถ θ นั้นเอง ดังนั้นค่าสารสนเทศของแบบสอบฉบับหนึ่งจึงมีได้หลายค่าแตกต่างกันไปตามระดับของ θ และที่ระดับความสามารถใด ๆ ค่าสารสนเทศจะมากหรือน้อยขึ้นอยู่กับปัจจัยต่อไปนี้ (Lord, 1980:67-68)

1. ค่าความคลาดเคลื่อนมาตรฐานของการวัด y θ ระดับความสามารถนั้น ซึ่งถ้ามีความคลาดเคลื่อนน้อยค่าสารสนเทศจาก y เกี่ยวกับ θ ก็จะมีมาก
2. ความชันของเส้นถดถอยของ y บน θ ซึ่งถ้ามีความชันมาก ค่าสารสนเทศจาก y เกี่ยวกับ θ ก็จะมีมาก

นอกจากนั้น Warm (1978:69-70) ได้ศึกษาและอธิบายถึงความสัมพันธ์ระหว่างฟังก์ชันการตอบสนองข้อสอบกับฟังก์ชันสารสนเทศของข้อสอบไว้ 3 ประการ ดังนี้

1. ฟังก์ชันสารสนเทศของข้อสอบมีค่าสูงสุดเมื่อเข้าใกล้ค่าซึ่งโค้งฟังก์ชันการตอบสนองข้อสอบมีค่าชันที่สุด
2. พื้นที่ใต้โค้งทั้งหมดของฟังก์ชันสารสนเทศของข้อสอบเพิ่มมากขึ้น ขณะที่ค่าอำนาจจำแนกเพิ่มขึ้น
3. พื้นที่ใต้โค้งทั้งหมดของฟังก์ชันสารสนเทศของข้อสอบลดลง ขณะที่ค่าการเดาเพิ่มขึ้น

การหาค่าสารสนเทศของแบบสอบนั้น สามารถหาได้จากสมการต่อไปนี้

$$I(\mathbf{e}_i) = \frac{\sum_{i=1}^n [P'_i(\mathbf{e}_i)]^2}{P_i(\mathbf{e}_i) Q_i(\mathbf{e}_i)}$$

เมื่อ $[P'_i(\mathbf{e}_i)]^2$ เป็นค่าอนุพันธ์ของโค้งลักษณะของข้อสอบ (ICC) ข้อที่ i
ณ ระดับความสามารถ e

และฟังก์ชันสารสนเทศของข้อสอบมีสูตรเป็น

$$I(\mathbf{e}_i) = \frac{[P'_i(\mathbf{e}_i)]^2}{P_i(\mathbf{e}_i) Q_i(\mathbf{e}_i)}$$

ซึ่ง

$$P'_i(\mathbf{e}_i) = \frac{1.7 a_i (1 - c_i)}{e^{1.7 a_i (e_j - b_i)} + 2 + e^{-1.7 a_i (e_j - b_i)}}$$

$$P_i(\mathbf{e}_i) = (c_i + e^{1.7 a_i (e_j - b_i)}) / (1 + e^{1.7 a_i (e_j - b_i)})$$

$$Q_i(\mathbf{e}_i) = 1 - c_i / (1 + e^{1.7 a_i (e_j - b_i)})$$

จากสมการจะเห็นได้ว่า ค่าพารามิเตอร์ของข้อสอบเป็นตัวกำหนดค่าสารสนเทศของข้อสอบ จึงอาจกล่าวได้ว่า ค่าสารสนเทศของข้อสอบและแบบสอบเป็นดัชนีผสม (composite index) ที่สร้างจากดัชนีบอกคุณลักษณะของข้อสอบหลายลักษณะ รวมเป็นดัชนีเพียงตัวเดียวที่ชี้ถึงคุณภาพของข้อสอบและแบบสอบได้ และด้วยคุณลักษณะความไม่แปรเปลี่ยนของค่าพารามิเตอร์ของข้อสอบ ค่าสารสนเทศจึงเหมาะที่จะใช้เป็นดัชนีชี้ถึงคุณภาพของข้อสอบและแบบสอบได้ดีกว่าค่าสถิติหรือค่าดัชนีอื่น ๆ ตามแนวคิดของทฤษฎีการวัดแบบคลาสสิก

ตอนที่ 3 งานวิจัยที่เกี่ยวข้อง

ในส่วนของงานวิจัยที่เกี่ยวข้องนี้ ผู้วิจัยจะเสนอการวิจัยเกี่ยวกับการเปรียบเทียบคุณภาพของแบบสอบเลือกตอบตัวเลือกธรรมดา กับแบบสอบเลือกตอบตัวเลือกซ้อน และการวิจัยเกี่ยวกับการเปรียบเทียบคุณภาพของแบบสอบเลือกตอบตัวเลือกธรรมดา กับแบบสอบเลือกตอบตัวเลือกผสม ตามลำดับ

การเปรียบเทียบคุณภาพของแบบสอบเลือกตอบตัวเลือกธรรมดา กับแบบสอบเลือกตอบตัวเลือกซ้อน

Weiten (1982:46-50) ได้ศึกษาประสิทธิภาพสัมพัทธ์ของแบบสอบเลือกตอบตัวเลือกธรรมดา กับแบบสอบเลือกตอบตัวเลือกซ้อน กลุ่มตัวอย่างเป็นนักศึกษาปริญญาตรีที่ลงทะเบียนเรียนวิชาจิตวิทยาทั่วไป จำนวน 47 คน เครื่องมือที่ใช้เป็นแบบสอบเลือกตอบตัวเลือกธรรมดาและแบบสอบเลือกตอบตัวเลือกซ้อน โดยข้อสอบแบบตัวเลือกซ้อนดัดแปลงมาจากข้อสอบแบบตัวเลือกธรรมดา ในแบบสอบแต่ละฉบับมีข้อสอบฉบับละ 40 ข้อ จัดพิมพ์แบบสอบเป็น 2 รูปแบบ ดังนี้

ฟอร์ม A เป็นแบบสอบที่ประกอบด้วยแบบสอบย่อยที่มีข้อสอบแบบตัวเลือกธรรมดา 20 ข้อ และแบบสอบย่อยที่มีข้อสอบแบบตัวเลือกซ้อน 20 ข้อ

ฟอร์ม B เป็นแบบสอบที่ประกอบด้วยแบบสอบย่อยที่มีข้อสอบแบบตัวเลือกซ้อน 20 ข้อ และแบบสอบย่อยที่มีข้อสอบแบบตัวเลือกธรรมดา 20 ข้อ

ผลการวิจัยพบว่า แบบสอบเลือกตอบตัวเลือกซ้อนยากกว่าแบบสอบเลือกตอบตัวเลือกธรรมดา ค่าอำนาจจำแนกของแบบสอบตัวเลือกธรรมดาสูงกว่าแบบสอบเลือกตอบตัวเลือกซ้อน การเปรียบเทียบค่าความเที่ยงได้เปรียบเทียบ 2 คู่ จาก 2 แบบสอบย่อย พบว่า ในคู่แรกแบบสอบย่อยที่ข้อสอบเป็นแบบตัวเลือกธรรมดามีค่าความเที่ยงสูงกว่าแบบสอบย่อยที่ข้อสอบเป็นแบบตัวเลือกซ้อน ส่วนการเปรียบเทียบอีก 1 คู่ แบบสอบย่อยที่มีตัวเลือกทั้ง 2 แบบ มีค่าความเที่ยงไม่แตกต่างกัน สำหรับด้านความตรงได้เปรียบเทียบ 2 คู่ เช่นเดียวกัน ได้พบว่า แบบสอบที่มีตัวเลือกทั้ง 2 แบบ มีค่าความตรงไม่แตกต่างกันทั้ง 2 คู่

ปราณี ร่วมทอง (2528:71-75) ได้ศึกษาเปรียบเทียบคุณภาพของแบบสอบเลือกตอบตัวเลือกธรรมดา กับแบบตัวเลือกซ้อนที่ใช้วัดระดับพุทธิพิสัยต่างกัน เมื่อใช้ทดสอบกับนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนต่างกัน คือ กลุ่มนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนสูง ปานกลาง และต่ำ ในด้านความเที่ยง ความตรง ความยาก และอำนาจจำแนกของแบบสอบ นอกจากนั้นยังได้เปรียบเทียบคะแนนที่ได้จากการตอบแบบสอบเลือกตอบตัวเลือกธรรมดา กับคะแนนที่ได้จากการ

ตอบแบบสอบตัวเลือกซ้อนด้วย แบบสอบที่ใช้ในการวิจัยเป็นแบบสอบวิชาวิทยาศาสตร์ชั้นมัธยมศึกษาปีที่ 3 (ว 305) ซึ่งมีรูปแบบของข้อสอบเป็นแบบสอบเลือกตอบตัวเลือกธรรมดา และแบบสอบเลือกตอบตัวเลือกซ้อน ที่วัดความรู้ 2 ระดับ คือ ระดับความรู้ขั้นต่ำ เป็นข้อสอบวัดการเรียนในระดับ ความรู้-ความจำ ความเข้าใจ การนำไปใช้ และระดับความรู้ขั้นสูง ซึ่งเป็นข้อสอบวัดการเรียนในระดับ การวิเคราะห์ การสังเคราะห์ และการประเมินผล กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคต้น ปีการศึกษา 2527 ในจังหวัดนครนายก จำนวน 328 คน ซึ่งแบ่งเป็น 2 กลุ่ม ๆ ละ 164 คน ผลการวิจัยพบว่า

1. ในการวัดระดับความรู้ขั้นต่ำ แบบสอบเลือกตอบตัวเลือกธรรมดา มีค่าความเที่ยงและค่าอำนาจจำแนกสูงกว่าแบบสอบเลือกตอบตัวเลือกซ้อน ด้านความตรงแบบสอบเลือกตอบทั้ง 2 ฉบับ มีค่าความตรงไม่แตกต่างกัน และพบว่าแบบสอบเลือกตอบตัวเลือกธรรมดาง่ายกว่าแบบสอบเลือกตอบตัวเลือกซ้อน

2. ในการวัดระดับความรู้ขั้นสูง แบบสอบเลือกตอบตัวเลือกธรรมดา และแบบสอบเลือกตอบตัวเลือกซ้อนมีค่าความเที่ยง ความตรง ความยาก และอำนาจจำแนกไม่แตกต่างกัน

3. นักเรียนกลุ่มที่มีผลสัมฤทธิ์ทางการเรียนสูง และกลุ่มที่มีผลสัมฤทธิ์ทางการเรียนต่ำ ตอบแบบสอบเลือกตอบทั้ง 2 ฉบับได้ไม่แตกต่างกัน

4. นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนปานกลาง ตอบแบบสอบเลือกตอบตัวเลือกธรรมดา ได้คะแนนเฉลี่ยสูงกว่าแบบสอบเลือกตอบตัวเลือกซ้อน

วัญญา ปรานปัญจะ (2529:42-47) ได้เปรียบเทียบคุณภาพของแบบสอบเลือกตอบตัวเลือกเดี่ยว กับแบบสอบเลือกตอบตัวเลือกซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกเท่ากัน และที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกไม่เท่ากัน ในด้านความยาก อำนาจจำแนก ความเที่ยง และเปอร์เซ็นต์การเดา แบบสอบที่ใช้วิจัยได้แก่ แบบสอบเลือกตอบวิชาวิทยาศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 จำนวน 3 ฉบับ ๆ ละ 30 ข้อคือ ฉบับที่ 1 เป็นแบบสอบเลือกตอบคำตอบเดี่ยว ฉบับที่ 2 เป็นแบบสอบเลือกตอบคำตอบซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกเท่ากัน กับ ฉบับที่ 3 เป็นแบบสอบเลือกตอบคำตอบซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกไม่เท่ากัน กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ปีการศึกษา 2528 สังกัดกรมสามัญศึกษา จังหวัดพัทลุง จำนวน 636 คน ผลการวิจัยพบว่า

แบบสอบเลือกตอบคำตอบซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกไม่เท่ากัน ยากที่สุด รองลงมาได้แก่ แบบสอบเลือกตอบคำตอบซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือก

เท่ากัน และแบบสอบเลือกตอบคำตอบเดียวง่ายที่สุด ด้านค่าอำนาจจำแนกพบว่า แบบสอบทั้ง 3 ฉบับ มีค่าอำนาจจำแนกไม่แตกต่างกัน ส่วนด้านความเที่ยงพบว่า แบบสอบเลือกตอบคำตอบซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกไม่เท่ากันมีความเที่ยงสูงกว่าแบบสอบเลือกตอบคำตอบเดียว และแบบสอบเลือกตอบคำตอบซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกเท่ากัน แต่ค่าความเที่ยงของแบบสอบเลือกตอบคำตอบเดียวกับแบบสอบเลือกตอบคำตอบซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกเท่ากัน มีค่าไม่แตกต่างกัน สำหรับเปอร์เซ็นต์การเดานั้นพบว่า แบบสอบทั้ง 3 ฉบับ มีเปอร์เซ็นต์การเดาไม่แตกต่างกัน

กมลทิพย์ อธิการชานนท์ (2530 :68-78) ได้ศึกษาเปรียบเทียบคุณภาพของแบบสอบเลือกตอบคำตอบถูกตัวเดียว คำตอบถูกไม่จำกัด และคำตอบรวมหรือแบบตัวเลือกซ้อน ในด้านความเที่ยง ความตรง และปริมาณการเดา นอกจากนี้ยังเปรียบเทียบคะแนนที่ได้จากแบบสอบทั้งสามฉบับ ในกลุ่มนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนสูง ปานกลาง และต่ำ แบบสอบที่ใช้ในการวิจัยได้แก่ แบบสอบวิชาวิทยาศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 (ว 203) จำนวน 3 ฉบับ ฉบับที่ 1 เป็นแบบสอบเลือกตอบคำตอบถูกตัวเดียว ฉบับที่ 2 เป็นแบบสอบเลือกตอบคำตอบไม่จำกัด และฉบับที่ 3 เป็นแบบสอบเลือกตอบคำตอบรวมหรือแบบตัวเลือกซ้อน มีจำนวนข้อสอบฉบับละ 50 ข้อ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคต้น ปีการศึกษา 2529 ของโรงเรียนในสังกัดกรมสามัญศึกษา กรุงเทพมหานคร จำนวน 450 คน ผลการวิจัยพบว่า

ค่าความเที่ยงของแบบสอบเลือกตอบคำตอบถูกไม่จำกัดมีค่าสูงกว่าแบบสอบเลือกตอบคำตอบถูกตัวเดียว แต่ค่าความเที่ยงของแบบสอบเลือกตอบคำตอบถูกไม่จำกัดกับแบบคำตอบรวม และค่าความเที่ยงของแบบสอบเลือกตอบคำตอบถูกตัวเดียวกับแบบสอบเลือกตอบคำตอบรวม มีค่าไม่แตกต่างกัน ด้านความตรงแบบสอบทั้ง 3 ฉบับมีค่าความตรงไม่แตกต่างกัน ส่วนด้านปริมาณการเดานั้นนักเรียนตอบแบบสอบทั้ง 3 ฉบับด้วยปริมาณการเดาที่ต่างกัน คือ ตอบแบบสอบเลือกตอบคำตอบถูกไม่จำกัดด้วยปริมาณการเดาต่ำสุด รองลงมาได้แก่แบบสอบเลือกตอบคำตอบรวม และแบบสอบเลือกตอบคำตอบเดียวมีปริมาณการเดาสูงสุด

คะแนนเฉลี่ยของนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนสูง ปานกลาง ต่ำ ที่ตอบแบบสอบทั้ง 3 ฉบับ มีลักษณะที่เหมือนกัน คือ นักเรียนตอบแบบสอบเลือกตอบคำตอบถูกตัวเดียวได้คะแนนเฉลี่ยสูงสุด รองลงมาคือแบบสอบเลือกตอบคำตอบรวม และแบบสอบเลือกตอบคำตอบถูกไม่จำกัดได้คะแนนเฉลี่ยต่ำสุด และไม่มีปฏิสัมพันธ์ระหว่างลักษณะของแบบสอบกับระดับผลสัมฤทธิ์ทางการเรียนของนักเรียน

จากการศึกษางานวิจัยเกี่ยวกับการเปรียบเทียบคุณภาพของแบบสอบเลือกตอบตัวเลือก
 ธรรมดา กับแบบสอบเลือกตอบตัวเลือกซ้อน ทั้งหมดศึกษาโดยใช้ทฤษฎีการวัดแบบคลาสสิก โดย
 ผลการวิจัยในคุณภาพบางด้านของแบบสอบยังขัดแย้งกันอยู่ เช่น ด้านความเที่ยงของแบบสอบ
 แต่โดยสรุปแล้วมีแนวโน้มว่า แบบสอบตัวเลือกซ้อนจะยากกว่าแบบสอบตัวเลือกธรรมดา ส่วนด้าน
 ความตรงพบว่าแบบสอบทั้งสองชนิดมีความตรงไม่แตกต่างกัน และมีแนวโน้มว่านักเรียนจะตอบ
 แบบสอบตัวเลือกซ้อนด้วยเปอร์เซ็นต์การเดาที่ต่ำกว่า

การเปรียบเทียบคุณภาพของแบบสอบเลือกตอบตัวเลือกธรรมดากับแบบสอบเลือกตอบ ตัวเลือกผสม

Hughes and Trimble (1965:117-125) ได้ศึกษาถึงคุณภาพของแบบสอบเลือกตอบ
 ที่มีตัวเลือกซับซ้อน (Complex Alternatives) แตกต่างกันในวิชาจิตวิทยาทั่วไป กลุ่มตัวอย่าง
 เป็นนักศึกษาปริญญาตรีชั้นปีที่ 1 จำนวน 63 คน แบ่งออกเป็น 3 กลุ่ม คือ กลุ่มที่ 1 ให้ทำแบบสอบ
 เลือกตอบที่ใช้ตัวเลือกธรรมดา กลุ่มที่ 2 ให้ทำแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบปลายเปิด
 "ไม่มีข้อใดถูก" และแบบปลายปิด "ถูกทุกข้อ" กลุ่มที่ 3 ให้ทำแบบสอบเลือกตอบที่ใช้ตัวเลือกผสม
 "ถูกทั้งข้อ 1 และข้อ 2" แบบสอบแต่ละฉบับมีข้อสอบฉบับละ 55 ข้อ แบ่งเป็น 3 ส่วน คือ
 ส่วนที่ 1 เป็นข้อสอบควบคุมซึ่งเป็นข้อสอบที่ใช้ตัวเลือกธรรมดาเหมือนกันทั้ง 3 ฉบับ ละ 15 ข้อ
 ส่วนที่ 2 เป็นข้อสอบที่ใช้เปรียบเทียบ ซึ่งเป็นข้อสอบที่ใช้ตัววางสุดท้ายเป็นตัวเลือกแบบต่าง ๆ
 จำนวน 35 ข้อ ส่วนที่ 3 เป็นข้อสอบที่ใช้ตัวเลือกแบบต่าง ๆ เป็นคำตอบที่ถูกต้อง จำนวน 5 ข้อ
 โดยจัดเรียงข้อสอบในแบบสอบแต่ละฉบับให้มีข้อสอบทุก ๆ 10 ข้อ ใน 50 ข้อแรก ประกอบด้วย
 ข้อสอบแบบเลือกตอบที่ใช้ตัวเลือกธรรมดา 3 ข้อ ข้อสอบแบบเลือกตอบที่ใช้ตัววางสุดท้ายเป็น
 ตัวเลือกแบบต่าง ๆ 6 ข้อ และข้อสอบแบบเลือกตอบที่ใช้ตัวเลือกแบบต่าง ๆ เป็นคำตอบที่ถูก 1
 ข้อ สำหรับ 5 ข้อสุดท้ายเป็นข้อสอบแบบเลือกตอบที่ใช้ตัววางสุดท้ายเป็นตัวเลือกแบบต่าง ๆ ซึ่ง
 ผลการวิจัยพบว่า แบบสอบเลือกตอบที่ใช้ตัวเลือกซับซ้อนทั้ง 2 ฉบับ ยากกว่าแบบสอบที่ใช้ตัวเลือก
 แบบธรรมดา ส่วนค่าอำนาจจำแนกของแบบสอบทั้ง 3 ฉบับไม่แตกต่างกัน ผลการเปรียบเทียบ
 คะแนนเฉลี่ยของนักเรียนทั้ง 3 กลุ่ม ปรากฏว่ามีความแตกต่างกัน โดยคะแนนเฉลี่ยของนักเรียนที่
 ทำแบบสอบเลือกตอบตัวเลือกธรรมดา มีค่าสูงกว่าคะแนนเฉลี่ยของนักเรียนที่ทำแบบสอบที่ใช้
 ตัวเลือกซับซ้อนทั้งสองกลุ่ม และในข้อสอบที่นักเรียนทั้งสองกลุ่มตอบผิดนั้นพบว่านักเรียนตอบผิด
 เนื่องจากเลือกตัววางซับซ้อนแบบผสม "ถูกทั้งข้อ 1 และข้อ 2" และตัววางซับซ้อนแบบปลายเปิด
 "ไม่มีข้อใดถูก" กับแบบปลายปิด "ถูกทุกข้อ"

ต่อมา Mueller (1975 : 135-141) ได้ศึกษาในเรื่องเดียวกันนี้อีก โดยศึกษากับ นักศึกษาในมหาวิทยาลัยอินเดียนา จำนวน 2000 คน ในวิชาการประกอบอาชีพเป็นพนักงานขาย (Salesman) ผลการวิจัยพบว่า แบบสอบที่ใช้ตัวเลือกผสมมากที่สุด รองลงมาได้แก่แบบสอบที่ใช้ตัวเลือกแบบปลายเปิดและแบบปลายปิด แบบสอบที่ใช้ตัวเลือกแบบธรรมดาง่ายที่สุด สำหรับค่าอำนาจจำแนกของแบบสอบสอบทั้ง 3 ฉบับนั้นพบว่าไม่แตกต่างกัน เมื่อพิจารณาประสิทธิภาพของ ตัวลวง ปรากฏว่านักเรียนเลือกตัวลวงแบบผสมมากที่สุด รองลงมาเป็นตัวลวงแบบปลายเปิดและ ตัวลวงแบบปลายปิด ส่วนตัวลวงแบบธรรมดา นักเรียนเลือกน้อยที่สุด

นวนน้อย แต่บรรพกุล (2520: 55-62) ได้ศึกษาสมรรถนะในการวัดของแบบสอบ เลือกตอบที่ใช้ตัวเลือกซับซ้อนแตกต่างกัน ในวิชาภาษาไทย ระดับชั้นประถมศึกษาปีที่ 7 โดยศึกษา แบบสอบเลือกตอบที่มีตัวเลือกแตกต่างกัน 4 แบบ คือ ตัวเลือกแบบธรรมดา ตัวเลือกแบบปลายเปิด "ผิดทุกข้อ" ตัวเลือกแบบปลายปิด "ถูกทุกข้อ" และตัวเลือกแบบผสม "ถูกทั้งข้อ ก และข้อ ข" หรือ "ถูกทั้งข้อ ก และข้อ ค" เป็นต้น แบบสอบที่ใช้ตัวเลือกแต่ละแบบยังแบ่งย่อยตามเนื้อหาเป็น 3 ฉบับ คือ แบบสอบสะกดคำ หลักภาษา และความเข้าใจในการอ่าน แบบสอบแต่ละฉบับ ประกอบด้วย 2 ส่วน ส่วนแรกเป็นข้อสอบที่ใช้เปรียบเทียบจริง จำนวน 25 ข้อ ซึ่งเป็นแบบสอบ เลือกตอบที่ใช้ตัวเลือกธรรมดา หรือตัวเลือกซับซ้อนแบบใดแบบหนึ่งเป็นตัวลวงตัวสุดท้าย ส่วนที่ 2 เป็นข้อสอบควบคุม ซึ่งเป็นข้อสอบที่ใช้ตัวเลือกธรรมดา จำนวน 10 ข้อ เป็นข้อสอบง่าย ๆ เพื่อไม่ให้มีผลต่อข้อสอบที่ใช้เปรียบเทียบจริง ผลการวิจัยพบว่า

1. ค่าความเที่ยงของแบบสอบที่ใช้ตัวเลือกต่างกันทั้ง 4 ชุด ไม่แตกต่างกัน ทั้งใน แบบสอบสะกดคำ หลักภาษา และความเข้าใจในการอ่าน
2. ค่าความตรงของแบบสอบที่ใช้ตัวเลือกต่างกันทั้ง 4 ชุด เป็นดังนี้
 - 2.1 แบบสอบสะกดคำ แบบสอบเลือกตอบที่ใช้ตัวเลือกแตกต่างกันทั้ง 4 ชุด มี ค่าความตรงไม่แตกต่างกัน
 - 2.2 แบบสอบหลักภาษา แบบสอบเลือกตอบที่ใช้ตัวเลือกธรรมดามีค่าความตรง สูงกว่าแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบปลายเปิด และตัวเลือกแบบผสม แบบสอบเลือกตอบ ตัวเลือกแบบปลายเปิดมีค่าความตรงสูงกว่าแบบสอบเลือกตอบตัวเลือกแบบผสม ส่วนการ เปรียบเทียบค่าความตรงของแบบสอบเลือกตอบที่ใช้ตัวเลือกต่างกันคู่อื่น ๆ พบว่ามีค่าไม่แตกต่างกัน

2.3 แบบสอบความเข้าใจในการอ่าน แบบสอบเลือกตอบที่ใช้ตัวเลือกธรรมดา มีค่าความตรงสูงกว่าแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบปลายปิด และตัวเลือกแบบผสม ส่วนการเปรียบเทียบค่าความตรงของแบบสอบเลือกตอบที่ใช้ตัวเลือกต่างกันคู่อื่น ๆ ไม่แตกต่างกัน

3. ค่าอำนาจจำแนกของแบบสอบเลือกตอบที่ใช้ตัวเลือกต่างกันทั้ง 4 ชุด ไม่แตกต่างกัน ทั้งในแบบสอบสะกดคำ หลักภาษา และความเข้าใจในการอ่าน

4. ค่าความยากมาตรฐานของแบบสอบที่ใช้ตัวเลือกต่างกัน 4 ชุด แตกต่างกันเฉพาะในแบบสอบความเข้าใจในการอ่าน โดยที่แบบสอบเลือกตอบที่ใช้ตัวเลือกแบบผสมมีค่าความยากมาตรฐานสูงสุด สำหรับแบบสอบที่ใช้ตัวเลือกธรรมดา แบบปลายเปิด และแบบปลายปิด มีค่าไม่แตกต่างกัน

5. การศึกษาเกี่ยวกับการเดา พบว่า แบบสอบเลือกตอบที่ใช้ตัวเลือกแบบผสมมีเปอร์เซ็นต์การเดาน้อยที่สุด ทั้งในแบบสอบสะกดคำ หลักภาษา และความเข้าใจในการอ่าน

6. คะแนนผลการสอบของนักเรียนที่ทำแบบสอบที่ใช้ตัวเลือกต่างกันทั้ง 4 ชุด มีความแตกต่างกันโดยเฉพาะอย่างยิ่ง คะแนนผลการสอบของนักเรียนที่ทำแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบผสมมีค่าต่ำสุด ทั้งในแบบสอบสะกดคำ หลักภาษา และความเข้าใจในการอ่าน

จากงานวิจัยเกี่ยวกับการเปรียบเทียบคุณภาพของแบบสอบเลือกตอบตัวเลือกธรรมดา กับแบบสอบเลือกตอบตัวเลือกผสม ปรากฏว่า เป็นการศึกษาโดยใช้ทฤษฎีการวัดแบบคลาสสิก ทั้งหมดเช่นเดียวกัน และพบว่าแบบสอบตัวเลือกผสมยากกว่าแบบสอบตัวเลือกธรรมดา แต่ด้านค่าอำนาจจำแนกไม่พบว่ามีค่าแตกต่างกัน

