



บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

งานสร้างข้อสอบเป็นกระบวนการที่สำคัญและค่อนข้างลำบาก ที่ว่าสำคัญก็คือแม้ว่าจะตั้งวัตถุประสงค์การเรียนการสอนไว้ดีเพียงใดก็ตาม ถ้าข้อคำถามในแบบสอบนั้นไม่สอดคล้องกับวัตถุประสงค์ผลการสอบที่ได้ก็ไม่สามารถสรุปได้ว่านักเรียนที่ตอบข้อสอบบรรลุวัตถุประสงค์เหล่านั้นแล้วหรือยัง ส่วนที่ลำบากนั้นก็เพราะว่าต้องปฏิบัติตามขั้นตอนต่าง ๆ หลายขั้นตอนเพื่อให้ได้ข้อคำถามที่สมบูรณ์ กล่าวคือประการแรกผู้เขียนข้อสอบจำเป็นต้องระบุและนิยามทักษะหรือความรู้ที่ต้องการวัดให้ชัดเจนเพื่อจะได้ตั้งคำถามเฉพาะที่มีคุณค่า ประการต่อมาจะต้องหาทางตั้งคำถามที่แน่ใจว่าคำตอบถูกจะไม่ใช่คำตอบที่ง่ายหรือยากจนเกินไป ดังนั้น จึงต้องการความแม่นยำและศิลปะในการเขียนคำถามค่อนข้างสูง คำถามของข้อสอบแบบเลือกตอบที่ดีนั้นควรจะให้นักเรียนได้เลือก ซึ่งน้ำหนัก และใช้ความรู้ในการตอบคำถาม (Educational Testing Service 1972 : 151)

นอกจากนี้ผู้เขียนข้อสอบยังต้องใช้คำถามและตัวเลือกซึ่งผู้สอบที่รู้จริง ๆ จึงจะสามารถตอบได้ถูกต้องส่วนผู้ที่ไม่รู้จะตอบผิด ศิลปะในการเขียนข้อสอบแบบเลือกตอบต้องอาศัยความรู้ในเนื้อหาวิชาและรู้ว่าวิชานั้นสอนอย่างไรในระดับการศึกษาที่ตั้งใจจะถาม ยิ่งกว่านั้นผู้เขียนข้อสอบควรจะทราบว่าสิ่งใดที่นักเรียนควรจะรู้ สิ่งใดที่นักเรียนน่าจะไม่ว่านักเรียนจะรู้หรือไม่รู้ นั่นคือควรจะสามารถตั้งคำถามที่สามารถจำแนกนักเรียนที่ขาดสารสนเทศหรือเข้าใจผิดออกจากผู้ที่มีความรู้ แต่ละตัวเลือกในข้อสอบแบบเลือกตอบควรมีส่วนช่วยในการจำแนกผู้ที่มีความรู้ความสามารถออกจากผู้ที่มีความรู้ความสามารถน้อยกว่า อำนาจในการจำแนกเช่นนี้ขึ้นอยู่กับคำตอบที่ผิดว่าจะสามารถสร้างความสนใจของผู้สอบที่มีความสามารถต่ำได้เพียงใด (Educational Testing Service 1972 : 152)

ปัจจุบันเทคโนโลยีการเขียนข้อสอบก้าวหน้าไปจนถึงการใช้ Item Form ในการเขียนข้อสอบและสามารถใช้คอมพิวเตอร์มาช่วยในการเขียนข้อสอบแล้ว (Roid and Haladyna 1982 : 6 - 7) แต่ยังไม่ปรากฏว่าข้อสอบวัดผลสัมฤทธิ์ในการเรียนที่ใช้กันอยู่ตามโรงเรียนต่าง ๆ นั้นส่วนใหญ่เป็นข้อสอบแบบเลือกตอบธรรมดาซึ่งประกอบด้วยตัวคำถามและตัวเลือก 4 หรือ 5 ตัว เลือกให้นักเรียนเลือกมาตอบเพียงตัวเลือกเดียวที่เป็นคำตอบที่ถูกต้องหรือดีที่สุด นอกจากนี้ข้อสอบ

ดังกล่าวก็ยังพบข้อบกพร่องอยู่มาก (หน่วยศึกษานิเทศก์ กรมสามัญศึกษา 2525) โดยเฉพาะอย่างยิ่ง ในส่วนของการเขียนตัวเลือกซึ่งมีข้อบกพร่องคล้ายคลึงกับข้อสอบแบบเลือกตอบที่ Ebel (1965 : 171 - 199) ได้รวบรวมมาจากแหล่งต่างๆ ตั้งแต่ปี ค.ศ. 1950 เป็นต้นมา ข้อบกพร่องเหล่านั้น ได้แก่ การใช้คำตอบไม่เหมาะสมกับคำถาม กล่าวคือนอกจากตัวคำตอบถูกแล้วตัวเลือกอื่น ๆ เป็นคำตอบที่ไม่ตรงกับคำถาม ตัวเลือกในข้อเดียวกันขาดความเป็นเอกพันธ์ เช่น บางตัวเลือกเป็นข้อความแสดงเวลาในขณะที่บางตัวเลือกเป็นข้อความแสดงถึงสถานที่ เป็นต้น บางก็ใช้ตัวเลือกที่เหมือน ๆ กัน 2 ตัวเลือกโดยตัวเลือกตัวหนึ่งขยายความอีกตัวเลือกหนึ่ง หากตัวเลือกที่เหมือน ๆ กันนี้เป็นคำตอบถูกก็เป็นการใช้ตัวเลือกตัวหนึ่งเป็นคำตอบที่ถูกต้อง หากเป็นคำตอบที่ผิดตัวเลือกที่เหมือน ๆ กัน 2 ตัวเลือกนี้ก็มีค่าเท่ากับตัวเลือกเดียวเท่านั้น

นอกจากนี้การใช้ตัวเลือกที่ซับซ้อนและยาวมากทำให้ยากแก่การอ่าน จดจำและจำแนกคำตอบ บางข้อใช้ตัวเลือกทุกข้อและผิดทุกข้ออย่างไม่เหมาะสม ปกติตัวเลือกทั้งสองนี้ควรใช้กับข้อสอบที่ต้องการคำตอบที่ถูกต้อง (Correct Answer) แต่ผู้เขียนข้อสอบนำไปใช้กับข้อสอบที่ต้องการคำตอบที่ดีที่สุด (Best Answer)

เฉพาะตัวเลือกที่เป็นตัวลวงนั้นบางข้อตัวลวงมีระดับความยากต่างจากคำตอบถูกมากเกินไป ถ้าตัวลวงมีระดับความยากต่ำกว่าคำตอบถูกมาก ๆ ผู้สอบก็จะสามารถตอบถูกได้โดยใช้วิธีขจัดตัวเลือกที่ผิดออกไป ผู้เขียนข้อสอบบางคนก็ประมาณความสามารถของนักเรียนต่ำเกินไป จึงใช้ตัวลวงที่เห็นได้ชัดเจนว่าไม่สมเหตุผล นอกจากนี้ก็มีการเล่นคำ เช่น ใช้คำปฏิเสธ ไม่นิยามบางตัวลวงโดยไม่แสดงหรือเน้นให้เห็นชัดเจน ผู้สอบบางคนอ่านไม่ละเอียดก็จะเลือกเป็นคำตอบถูก ข้อสอบลักษณะนี้จึงเป็นการวัดความละเอียดรอบคอบมากกว่าวัดความรู้ของผู้สอบ อีกทั้งการใช้ตัวลวงที่คล้ายกันมาก ๆ ก็เป็นการทำให้คำตอบถูกเด่นชัดมากกว่าที่ควร

ในการตอบข้อสอบแบบเลือกตอบนั้นมักวัดผลเชื่อว่า ผู้ที่มีความฉลาดในการตอบข้อสอบ (Test-wise) อยู่บ้างจะไม่ตอบแบบเดาสุ่มแม้ว่าจะไม่แน่ใจในคำตอบที่ถูกต้องก็ตาม แต่จะใช้วิธีขจัดตัวเลือกที่คิดว่าไม่ถูกต้องออกไปทีละตัวเลือก จนกระทั่งได้คำตอบคือตัวเลือกที่เหลืออยู่เป็นตัวสุดท้าย วิธีนี้จะทำได้ง่ายและประสบความสำเร็จอย่างสูงหากตัวเลือกในข้อสอบเหล่านั้นเป็นตัวเลือกที่ด้อยประสิทธิภาพอันได้แก่ตัวเลือกที่มีข้อบกพร่องดังกล่าวข้างต้น ฉะนั้นจึงกล่าวได้ว่าประสิทธิภาพของข้อสอบแบบเลือกตอบนอกจากจะขึ้นอยู่กับคุณภาพของตัวคำถามแล้วก็ยังขึ้นอยู่กับคุณภาพของตัวเลือกเป็นอย่างมาก ผู้วิจัยจึงสนใจศึกษาผลของตัวเลือกลักษณะต่าง ๆ ของข้อสอบแบบเลือกตอบ

จากการศึกษาค้นคว้าตำราที่ว่าด้วยการเขียนข้อสอบวัดผลสัมฤทธิ์ ในส่วนของการเขียนข้อสอบแบบเลือกตอบประกอบกับประสบการณ์ในงานวัดผลการศึกษาในระดับมัธยมศึกษาของผู้วิจัยเอง

ตลอดจนการที่ได้พูดคุยกับครูผู้สอนในระดับมัธยมศึกษาหลายคนเกี่ยวกับลักษณะของข้อสอบที่ใช้อยู่นอ จะสรุปได้ว่า ในการเขียนข้อสอบแบบเลือกตอบเพื่อวัดผลสัมฤทธิ์ของนักเรียนนั้นมักเขียนตัวเลือก ในลักษณะต่อไปนี้ คือ

1. ใช้ความเข้าใจผิด ๆ หรือข้อผิดพลาดต่างๆ ของนักเรียนที่พบในชั้นเรียนมาเขียนเป็นตัวลวง(Scannell and Tracy 1975 : 134 ; Tuckman 1975 : 93) บางครั้งอาจใช้วิธีตั้งคำถามให้นักเรียนตอบสั้น ๆ หรือเติมคำ แล้วรวบรวมเอาความเข้าใจผิด ๆ หรือข้อผิดพลาดที่มีความถี่มาก ๆ มาใช้เป็นตัวลวง(Marshall and Hales 1971 : 106 - 107) ตัวเลือกที่มีลักษณะเช่นนี้จะเรียกว่าตัวเลือกตก(ไพบูลย์ จิตรโรต 2514) การเขียนตัวเลือกโดยวิธีนี้ครูผู้สอนยอมรับว่าเป็นสิ่งที่เหมาะสมหากมีเวลาพอก็จะเขียนตัวเลือกในลักษณะนี้
2. ในวิชาคณิตศาสตร์นอกจากวิธีแรกทีกล่าวแล้วครูมักใช้วิธีเขียนตัวลวงให้ใกล้เคียงกับคำตอบถูก ตัวอย่างเช่นมีโจทย์ว่า ดินสอราคาแท่งละ 3 บาท ถ้าซื้อหนึ่งโหลต้องจ่ายเงินกี่บาท คำตอบที่ถูกต้องคือ 36 บาท ตัวลวงก็อาจจะเป็น 32 บาท 34 บาท 38 บาท เป็นต้น ตัวเลือกที่ได้มาในลักษณะนี้จะเรียกว่าตัวเลือกเจียด(ไพบูลย์ จิตรโรต 2514)
3. ในบางครั้งผู้เขียนข้อสอบไม่สามารถหาตัวเลือกที่เหมาะสมได้อีก ตัวเลือกที่มีอยู่ก็ยังไม่ครบตามจำนวนที่ต้องการจึงใช้ตัวเลือกปลายเปิด **ไม่มีข้อใดถูก** และหรือตัวเลือกปลายปิด **ทุกข้อ** การใช้ตัวเลือก **ไม่มีข้อใดถูก** และ **ตัวเลือกถูกทุกข้อ** มีข้อควรระวัง คือ ต้องใช้กับข้อคำถามที่มีคำตอบที่ถูกผิดชัดเจน(Correct Answer) **ไม่ควร** ใช้กับข้อคำถามที่ต้องการคำตอบที่ดีที่สุด(Best Answer) (Ebel and Frisbie 1986 : 189 ; Scannell and Tracy 1975 : 131 ; Marshall and Hales 1971 : 103 ; Ebel 1965 : 159, 182) ตัวเลือกปลายเปิด **ไม่มีข้อใดถูก** นี้เหมาะกับข้อสอบคณิตศาสตร์ คำศัพท์ การสะกดคำ และไวยากรณ์ (Mehrens and Lehman 1984 : 162 ; Thorndike 1971 : 117) และนิยมเขียนไว้เป็นตัวเลือกในลำดับสุดท้าย(Stanley and Hopkins 1972 : 252) แต่บางคนก็มีความเห็นว่าตัวเลือก **ไม่มีข้อใดถูก** และ **ถูกทุกข้อ** นี้ไม่ควรใช้(Nunnally 1964 : 126 - 127)
4. นอกจากตัวเลือกปลายเปิด **ไม่มีข้อใดถูก** และตัวเลือกปลายปิด **ทุกข้อ** แล้วข้อสอบบางข้อที่มีคำตอบถูกได้มากกว่าหนึ่งคำตอบก็มีการใช้ตัวเลือกผสม เช่น ถูกทั้ง ก และ ข หรือ ถูกทั้ง ข และ ค เป็นต้น ตัวเลือกทำนองนี้บางคนก็เห็นว่าไม่ควรใช้(Nunnally 1964 : 126 - 127)

5. ครูบางคนสอนหลายระดับชั้น เช่นสอนทั้งระดับ ม.1 ม.2 และ ม.3 บางครั้งจึงอาจเผลอเขียนข้อสอบที่เกินระดับความรู้ของนักเรียนที่จะสอบ แต่บางคนก็ตั้งใจใช้ศัพท์แปลก ๆ หรือข้อความรู้ที่นักเรียนยังไม่เคยเรียนมาเป็นตัวลวง เนื่องจากตัวเลือกที่เกินระดับความรู้ของผู้สอบดังกล่าวนี้ นักเรียนยังไม่เคยมีประสบการณ์มาก่อนเป็นสิ่งแปลกปลอมสำหรับ เขาจึงเรียกตัวเลือกทำนองนี้ว่า **ตัวเลือกแปลก**

จากลักษณะการเขียนตัวเลือกของข้อสอบแบบเลือกตอบดังกล่าวนี้ ผู้วิจัยพิจารณาแล้วมีความเห็นว่าลักษณะตัวเลือกที่น่าสนใจนำมาศึกษาได้แก่ตัวเลือกดังต่อไปนี้

1. **ตัวเลือกดัก** การเขียนตัวเลือกในลักษณะตัวเลือกดักเป็นสิ่งพึงประสงค์และควรสนับสนุนให้ครูเขียนตัวเลือกในลักษณะนี้ เพราะตัวเลือกที่ผิดซึ่งนักเรียนเลือกนั้นจะให้สารสนเทศของความบกพร่องของนักเรียนอันจะนำไปสู่การสอนซ่อมเสริมที่เหมาะสมต่อไปได้ ฉะนั้นผู้วิจัยจึงเลือกตัวเลือกดักเป็นตัวเลือกลักษณะหนึ่งที่ศึกษา

2. **ตัวเลือกเจียด** แม้ว่าการเขียนตัวเลือกในลักษณะตัวเลือกเจียดเป็นที่นิยมมากในการเขียนข้อสอบวิชาคณิตศาสตร์แต่ก็เป็นลักษณะที่พบได้ในวิชาอื่น ๆ ซึ่งมักเขียนตัวลวงโดยการล้อเลียนตัวเลือกที่เป็นคำตอบถูก เนื่องจากตัวเลือกเจียดนี้เขียนค่อนข้างง่ายจึงเป็นที่นิยมพอสมควรด้วยเหตุนี้ผู้วิจัยจึงเลือกมาศึกษา

3. **ตัวเลือกแปลก** การเขียนตัวเลือกในลักษณะของตัวเลือกแปลกนี้มีกพบอยู่สองกรณีคือ กรณีที่ผู้เขียนข้อสอบคนนั้นสอนหลายรายวิชาและต่างระดับชั้นจึงอาจเผลอเขียนตัวเลือกที่เกินระดับความรู้ของผู้สอบไปบ้างโดยมิได้ตั้งใจ ส่วนอีกกรณีหนึ่งนั้นผู้เขียนข้อสอบจงใจเขียนตัวเลือกให้แปลกไปจากสิ่งที่นักเรียนเคยเรียนมาเพื่อลวงผู้สอบที่รู้ไม่จริง แม้ว่าตัวเลือกในลักษณะตัวเลือกแปลกนี้จะพบน้อยแต่ก็เป็นลักษณะที่น่าสนใจนำมาศึกษา

4. **ตัวเลือกปลายเปิด** แม้ว่านักวัดผลบางท่านไม่สนับสนุนให้ใช้ตัวเลือกปลายเปิด **ไม่มีข้อใดถูก** แต่กลับพบว่า เป็นตัวเลือกที่นิยมใช้กันมากในทุกวิชา โดยเฉพาะอย่างยิ่งเมื่อไม่สามารถหาตัวลวงที่เหมาะสมได้ ฉะนั้นตัวเลือกปลายเปิดจึงเป็นตัวเลือกอีกลักษณะหนึ่งที่ผู้วิจัยสนใจนำมาศึกษา

5. **ตัวเลือกให้เติม** สืบเนื่องจากการใช้ตัวเลือกปลายเปิด **ไม่มีข้อใดถูก** ในกรณีที่ตัวเลือกปลายเปิดนี้เป็นคำตอบที่ถูกต้องบางครั้งก็ไม่สามารถแน่ใจได้ว่าผู้สอบที่เลือกตัวเลือกนี้จะรู้คำตอบที่แท้จริงของคำถามข้อนี้ คำตอบที่เขาคิดได้อาจไม่ตรงกับตัวเลือกอื่น ๆ ที่มีอยู่แต่ก็เป็นคำตอบที่

ไม่ถูก จะเน้นเพื่อให้แน่ใจว่าคำตอบที่ผู้สอบคิดได้นั้นถูกต้องจริงหรือไม่จึงน่าจะให้ผู้เลือกตัวเลือก ปลายเปิดนี้ เขียนคำตอบที่เขาคิดได้ลงไปด้วย ส่วนที่ใช้ อาจเป็นคำตอบที่ถูกต้องคือ...ตัวเลือก ลักษณะนี้ผู้วิจัยขอเรียกว่า **ตัวเลือกให้เติม** แม้ว่าตัวเลือกลักษณะนี้ยังไม่เป็นที่นิยมแต่ผู้วิจัยเชื่อว่า ข้อสอบที่ใช้ตัวเลือกให้เติมนี้จะวัดความสามารถของผู้สอบได้แม่นยำกว่าตัวเลือกลักษณะอื่น จึงเลือกมาศึกษา

สำหรับตัวเลือกปลายปิดทุกข้อ และตัวเลือกผสมทุกทั้ง ก และ ข นั้นใช้ได้กับคำถาม ในบางเรื่องที่มีคำตอบถูกได้หลายคำตอบ แต่โดยทั่วไปเฉพาะอย่างยิ่งในวิชาคณิตศาสตร์ คำถาม ส่วนใหญ่ก็มีคำตอบที่ถูกต้องเพียงคำตอบเดียว ตัวเลือกรทั้งสองลักษณะนี้จึงแทบไม่มีโอกาสที่จะเป็น คำตอบที่ถูกต้องเลยผู้วิจัยจึงมิได้นำมาศึกษา

กล่าวโดยสรุป ลักษณะตัวเลือกของข้อสอบแบบเลือกตอบที่ผู้วิจัยศึกษามีอยู่ 5 แบบ คือ ตัวเลือกตัด ตัวเลือกเจียด ตัวเลือกปลายเปิด ตัวเลือกแปลก และตัวเลือกให้เติม

เนื่องจากแบบสอบวินิจฉัยวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 เรื่องสมการที่ประดิษฐ์ เรื่องตระกูล สร้างขึ้นในปี 2529 นั้นเป็นแบบสอบวัดผลสัมฤทธิ์เชิงวินิจฉัยซึ่งมีลักษณะตัวเลือกเป็น แบบตัวเลือกตัดตามที่ผู้วิจัยสนใจศึกษาอีกทั้งผ่านการทดลองใช้และวิเคราะห์คุณภาพมาแล้ว พบว่ามี คุณภาพในระดับที่ยอมรับได้ ผู้วิจัยจึงเลือกแบบสอบดังกล่าวเป็นตัวแทนของแบบสอบที่มีลักษณะตัวเลือกเป็นตัวเลือกตัดและนำแบบสอบฉบับนี้มาปรับตัวเลือกเป็นแบบสอบอีก 4 ฉบับ ซึ่งมีตัวเลือกใน ลักษณะของตัวเลือกเจียด ตัวเลือกปลายเปิด ตัวเลือกแปลก และตัวเลือกให้เติมตามลำดับ

ในการเขียนข้อสอบแบบเลือกตอบนั้นให้ความสำคัญกับการเขียนตัวเลือกเป็นอย่างมาก โดยเฉพาะอย่างยิ่งการเขียนตัวลวงดังกล่าวแล้วข้างต้น แต่เมื่อนำแบบสอบไปทดสอบแล้วกลับไม่ได้สนใจสารสนเทศจากตัวลวงเลย ทั้งในการใช้ผลการสอบและในกระบวนการประมาณค่าความสามารถของผู้สอบ ไม่ว่าจะเป็นการประมาณด้วยทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) หรือทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) ดังจะเห็นได้จากการที่ ไปเน้นคะแนนรวมที่ได้จากจำนวนข้อที่ตอบถูกซึ่งมีจุดอ่อนอยู่ว่าจะคะแนนรวมตามจำนวนข้อที่ตอบถูกที่ เท่ากันนั้นมาจากรูปแบบการตอบสนอง (Response Patterns) ต่าง ๆ กันตัวอย่างเช่นแบบสอบ ที่มีจำนวนข้อสอบ 10 ข้อจะมีรูปแบบการตอบสนองต่าง ๆ ที่เป็นไปได้ซึ่งเกิดจากลักษณะของการ ตอบถูกและการตอบผิดในข้อต่างๆถึง 252 รูปแบบสำหรับกรณีคะแนนรวม 5 คะแนน (Harnisch 1983 : 191) แต่ละรูปแบบย่อมให้ความหมายแตกต่างกัน ดังนั้น การตีความหมายผลการสอบ จากคะแนนรวมเพียงอย่างเดียวจึงอาจนำไปสู่การตีความหมายผิดๆ ได้ นักวิจัยหลายท่านสนใจ สารสนเทศที่ได้จากการวิเคราะห์รูปแบบการตอบเพิ่มเติมจากสารสนเทศที่ได้จากคะแนนรวม ซึ่ง

สารสนเทศดังกล่าวจะเป็นประโยชน์ในการวิเคราะห์ความแตกต่างระหว่างบุคคล ความไม่สอดคล้องกันของหลักสูตรระหว่างโรงเรียนต่าง ๆ ในเขตการศึกษาเดียวกัน ตลอดจนความแตกต่างทางวัฒนธรรม เป็นต้น (Tatsuoka 1984 : 95)

การหันไปสนใจรูปแบบการตอบสนองทำให้พบว่า คะแนนของผู้สอบบางครั้งก็อาจจะไม่สามารถเป็นตัวแทนของการวัดความสามารถหรือผลสัมฤทธิ์ได้ (Drasgow and Levine 1986 : 59) คะแนนของผู้สอบอาจจะสูงกว่าที่ควรจะเป็น (Spuriously High) เมื่อผู้สอบลอกคำตอบจากเพื่อนข้าง ๆ ที่มีความสามารถสูงกว่า หรือมีคนบอกคำตอบบางข้อให้ หรือรู้ข้อสอบล่วงหน้า หรือมีผู้เติมคำตอบที่ถูกต้องให้ในข้อที่เว้นว่างไว้ไม่ได้ตอบ (Drasgow and Levine 1986 : 59 ; Hulin and others 1983 : 113) ผู้สอบบางคนอาจจะได้คะแนนต่ำกว่าที่ควรจะเป็น (Spuriously Low) เนื่องมาจากการตอบไม่ตรงข้อ หรือผู้สอบตีความหมายคำถามลึกซึ้งเกินกว่าที่ผู้สร้างแบบสอบตั้งใจ หรือการมีปัญหาทางภาษา หรือการได้รับการศึกษามาผิดแผกไปจากคนอื่น ๆ หรือเนื่องจากยุทธวิธีในการทำข้อสอบหรือความไม่คุ้นเคยกับรูปแบบของแบบสอบ (Test Format) หรือการไม่ตั้งใจทำข้อสอบ และอื่น ๆ (Drasgow and Levine 1986 : 59 ; Hulin and others 1983 : 112 - 113 ; Levine and Rubin 1979 : 270) สาเหตุต่าง ๆ เหล่านี้ทำให้เกิดรูปแบบการตอบสนองที่ผิดปกติขึ้น (Unusual Response Patterns) ผู้สอบที่มีรูปแบบการตอบสนองที่ผิดปกตินี้หมายความว่า รูปแบบการตอบสนองของเขาแตกต่างไปจากรูปแบบการตอบสนองของคนอื่น ๆ ที่มีความสามารถระดับเดียวกัน คะแนนที่ได้จึงเป็นคะแนนที่ไม่เหมาะสม (Inappropriate) กับความสามารถของผู้สอบคนนั้น

จะเห็นว่า ผู้สอบที่มีความสามารถระดับเดียวกันอาจจะทำคะแนนสอบได้แตกต่างกันขึ้นอยู่กับรูปแบบของการตอบสนอง (Response Patterns) และรูปแบบของการตอบสนองที่ต่างกักันนั้นก็เนื่องมาจากนิสัยในการตอบข้อสอบ (Response Styles) (Stanley and Hopkins 1972 : 141) ความฉลาดในการตอบข้อสอบ (Test-wise) และลักษณะตัวเลือกของข้อสอบ ผู้วิจัยสนใจประเด็นที่ว่าตัวเลือกลักษณะต่าง ๆ ของข้อสอบแบบเลือกตอบที่ใช้ในการทดสอบอยู่ขณะนี้มีผลต่อรูปแบบการตอบสนองข้อสอบหรือไม่

อนึ่งจากการศึกษาค้นคว้างานวิจัยที่เกี่ยวข้องพบว่า การวิจัยภายในประเทศได้เคยมีการศึกษาเกี่ยวกับลักษณะตัวเลือกแบบต่าง ๆ ของข้อสอบแบบเลือกตอบในแง่ของ ความยาก อำนาจ จำแนก ความเที่ยงและความตรง (ไพบูลย์ จิตรวิโต 2514 , นวลน้อย แต่บรรพกุล 2520 และ นิรมล บุญตะรัตน์ 2524) แต่ยังไม่พบว่ามีการศึกษาในแง่ของรูปแบบการตอบสนอง (Response Patterns) และการวิจัยที่พบนั้นเป็นการศึกษาโดยอาศัยทฤษฎีการวัดแบบดั้งเดิม (Classical Test Theory) ซึ่งในปัจจุบันทฤษฎีนี้ถูกมองว่าเป็นทฤษฎีที่มีข้อบกพร่องหลายประการ ประการ

แรก คือ การใช้ทฤษฎีนี้ไปศึกษาคุณภาพของแบบสอบนั้นค่าสถิติที่ได้ของข้อสอบจะแปรเปลี่ยนไปตามกลุ่มตัวอย่างผู้สอบที่ใช้คำนวณค่าสถิตินั้น **ประการที่สอง** ในการเปรียบเทียบความสามารถของผู้สอบจะต้องใช้แบบสอบฉบับเดียวกันหรือแบบสอบคู่ขนาน **ประการที่สาม** นิยามความเที่ยงของแบบสอบในรูปของการวัดคู่ขนานซึ่งเป็นไปได้ยากในทางปฏิบัติ **ประการที่สี่** ไม่มีพื้นฐานในการที่จะบอกได้ว่าผู้สอบคนหนึ่งจะมีโอกาสตอบข้อสอบข้อหนึ่งได้ถูกต้องเพียงใด ข้อบกพร่อง**ประการที่ห้า** ของทฤษฎีการวัดแบบดั้งเดิมก็คือ การที่กำหนดข้อตกลงเบื้องต้นว่าความแปรปรวนของความคลาดเคลื่อนในการวัดของผู้สอบทุกคนเท่ากัน (Hambleton and Swaminathan 1985 : 1 - 3)

เนื่องจากทฤษฎีการวัดแบบดั้งเดิมมีข้อบกพร่องหลายประการดังกล่าวและเราต้องการรูปแบบการวัดที่ให้สารสนเทศเกี่ยวกับความแม่นยำของคะแนนสอบที่ใช้ประมาณความสามารถซึ่งก็คือรูปแบบการวัดตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) นั้นเอง รูปแบบการวัดนี้มีลักษณะดังนี้คือ (Hambleton and Swaminathan 1985 : 9) **ประการแรก** ผลการทำแบบสอบของผู้สอบสามารถทำนายหรืออธิบายได้ ในรูปของคุณลักษณะแฝงหรือความสามารถบางอย่างในตัวผู้สอบ **ประการที่สอง** รูปแบบนี้ชี้ถึงความสัมพันธ์ระหว่างผลการทำข้อสอบที่สังเกตได้กับคุณลักษณะแฝงหรือความสามารถที่อยู่ภายใต้ผลการทำแบบสอบนั้น **ประการที่สาม** เป็นรูปแบบที่ใช้ประมาณความสามารถหรือคุณลักษณะแฝงของผู้สอบ **ประการที่สี่** คุณลักษณะแฝงหรือความสามารถของผู้สอบต้องประมาณหรืออ้างอิงไปจากผลการทำแบบสอบที่สังเกตได้

นอกจากลักษณะ 4 ประการนี้แล้วรูปแบบการวัดตามทฤษฎีการตอบสนองข้อสอบยังมีจุดเด่นอีก 3 ประการ (Hambleton and Swaminathan 1985 : 11) คือ **ประการแรก** ค่าพารามิเตอร์ของข้อสอบเป็นอิสระจากกลุ่มผู้สอบที่ใช้ประมาณค่า **ประการที่สอง** ความสามารถของผู้สอบเป็นอิสระจากข้อสอบที่ใช้ประมาณค่านั้น **ประการที่สาม** สามารถทราบความแม่นยำของการประมาณความสามารถของผู้สอบ

จากทฤษฎีการตอบสนองข้อสอบดังกล่าวจะเห็นว่าทฤษฎีนี้เชื่อว่าผลการตอบข้อสอบนั้นขึ้นอยู่กับลักษณะของข้อสอบและความสามารถของผู้สอบ ความสัมพันธ์ระหว่างผลการตอบข้อสอบแต่ละข้อกับความสามารถของผู้สอบนั้นแสดงได้ด้วยฟังก์ชันของการตอบสนองข้อสอบ (Item Response Function) หรือโค้งลักษณะของข้อสอบ (Item Characteristic Curve) ฟังก์ชันนี้คือความน่าจะเป็น P หรือ $P(\theta)$ ของการตอบข้อสอบถูกโดยสมมุติ (assume) ว่าความน่าจะเป็นในการตอบข้อสอบถูก ($P(\theta)$) จะเพิ่มขึ้นเมื่อความสามารถ (θ) เพิ่มขึ้นและความน่าจะเป็นนี้สามารถแทนด้วยฟังก์ชันโลจิสติกที่มี 3 พารามิเตอร์ (Three - Parameter Logistic Function) (Lord 1980 : 12) ดังนี้

$$P = P(\theta) = c + (1 - c) / [1 + e^{-1.7a(\theta-b)}] \quad (1-1)$$

เมื่อ a , b และ c เป็นค่าพารามิเตอร์ของข้อสอบ e คือค่าคงที่มีค่า 2.71828

พารามิเตอร์ c นั้นคือความน่าจะเป็นที่ผู้สอบที่ไม่มีความสามารถเลย ($\theta = -\infty$) จะตอบข้อสอบข้อนั้นได้ถูกต้องจึงเรียก c ว่าเป็นพารามิเตอร์การเดาถ้าข้อสอบข้อนั้นไม่สามารถเดาถูกต้องได้เลยแล้ว c จะเป็น 0

พารามิเตอร์ b แสดงตำแหน่งของโค้งลักษณะของข้อสอบบนสเกล (Scale) ของความสามารถและเรียกว่าความยากของข้อสอบ ถ้าข้อสอบยิ่งยากโค้งก็ยิ่งจะอยู่ค่อนข้างไปทางขวามาก ในโค้งโลจิสติก (Logistic Curve) นั้นจุดเปลี่ยนโค้งจะอยู่ที่ $\theta = b$ ในกรณีที่ไม่มี การเดาค่า b เป็นระดับความสามารถที่ความน่าจะเป็นในการตอบข้อสอบข้อนั้นถูกคือ 0.5 แต่ถ้ามีการเดาแล้วค่า b จะเป็นระดับความสามารถที่ความน่าจะเป็นในการตอบข้อสอบข้อนั้นถูก คือจุดกลางระหว่าง c และ 1.0

พารามิเตอร์ a เป็นสัดส่วนของความชัน (Slope) ของโค้งที่จุดเปลี่ยนโค้ง จึงแทนค่าอำนาจจำแนกของข้อสอบซึ่งก็คือขนาด (Degree) ที่การตอบข้อสอบแปรเปลี่ยนไปตามระดับความสามารถ

เนื่องจากจุดมุ่งหมายของการทดสอบก็เพื่อนำผลการสอบอ้างอิงไปยังความสามารถของผู้สอบโดยใช้ข้อสอบเป็นสิ่งที่เร้าหรือเครื่องมือในการประมาณความสามารถของผู้สอบ จากลักษณะเด่นของทฤษฎีการตอบสนองข้อสอบ ที่ว่าสามารถทราบความแม่นยำของการประมาณความสามารถได้นั้นก็คือ เราสามารถทราบว่าถ้าใช้ข้อสอบข้อหนึ่งในการประมาณความสามารถของผู้สอบเราจะทราบความแม่นยำได้จากฟังก์ชันสารสนเทศของข้อสอบ (Item Information Function) ($I(\theta, u_1)$) ซึ่งคำนวณได้จากสูตรต่อไปนี้

$$I(\theta, u_1) = P_1'^2 / P_1 Q_1 \quad (1-2)$$

เมื่อ $P_1 = P_1(\theta)$ คือฟังก์ชันการตอบสนองข้อสอบ

$$Q_1 = 1 - P_1$$

P_1' คืออนุพันธ์ของ P_1 เมื่อพิจารณาตาม θ

ถ้าฟังก์ชันสารสนเทศมีค่าสูงแสดงว่ามีสารสนเทศมาก ความแม่นยำในการประมาณค่าความสามารถก็มีมาก ปริมาณของสารสนเทศนี้จะแปรเปลี่ยนไปในระดับความสามารถต่าง ๆ ค่าสารสนเทศของข้อสอบที่ระดับความสามารถใด ๆ จะแปรผันตรงกับกำลังสองของค่าอำนาจจำแนก (Lord 1980 : 21)

เมื่อข้อสอบหลาย ๆ ข้อรวมกันเข้าเป็นแบบสอบก็จะสามารถทราบความแม่นยำของแบบสอบฉบับนั้นในการประมาณความสามารถของผู้สอบได้ จากฟังก์ชันสารสนเทศของแบบสอบ (Test Information Function) ($I(\theta)$) ซึ่งเป็นสัดส่วนกลับกับกำลังสองของช่วงความเชื่อมั่นที่เป็นเอซิมโทติก (Asymptotic Confidence Interval) ในการประมาณค่าความสามารถของผู้สอบจากการตอบสนองแบบสอบ ฟังก์ชันสารสนเทศของแบบสอบนี้สามารถแสดงในรูปผลรวมของฟังก์ชันสารสนเทศของข้อสอบในแบบสอบฉบับนั้นได้ดังนี้ (Lord 1980 : 22)

$$I(\theta) = \sum_1 I(\theta, u_1) \quad (1-3)$$

จะเห็นว่า ฟังก์ชันสารสนเทศของข้อสอบและแบบสอบนั้นขึ้นอยู่กับฟังก์ชันการตอบสนองข้อสอบ เช่นเดียวกับกับรูปแบบการตอบสนองข้อสอบ และฟังก์ชันการตอบสนองข้อสอบเองนั้นก็ขึ้นอยู่กับพารามิเตอร์ของข้อสอบซึ่งจะแปรเปลี่ยนไปตามลักษณะของข้อสอบ ฉะนั้นถ้าลักษณะของข้อสอบเปลี่ยนแปลงไปก็หน้าจะมีผลกระทบต่อค่าพารามิเตอร์ของข้อสอบ ฟังก์ชันการตอบสนองข้อสอบ รูปแบบการตอบสนองข้อสอบ ฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ

จากประเด็นเกี่ยวกับรูปแบบของตัวเลือกและแนวคิดเกี่ยวกับทฤษฎีการวัดผลดังกล่าว ผู้วิจัยจึงทำการศึกษาเพื่อตอบคำถามที่ว่า ข้อสอบแบบเลือกตอบที่มีตัวเลือกลักษณะต่าง ๆ ดังที่ใช้ในการทดสอบอยู่ขณะนี้ซึ่งได้แก่ ตัวเลือกตัก ตัวเลือกเฉียด ตัวเลือกปลายเปิด ตัวเลือกแปลกและตัวเลือกให้เติมเต็มมีผลต่อโค้งลักษณะของข้อสอบ (Item Characteristic Curve) ฟังก์ชันสารสนเทศของข้อสอบ (Item Information Function) ฟังก์ชันสารสนเทศของแบบสอบ (Test Information Function) และรูปแบบการตอบสนองข้อสอบ (Response Patterns) หรือไม่

คำตอบที่ได้จากการศึกษาครั้งนี้ ในส่วนที่เกี่ยวกับรูปแบบการตอบสนองข้อสอบหากพบว่าตัวเลือกลักษณะใดมีแนวโน้มที่จะให้รูปแบบการตอบสนองข้อสอบที่ผิดปกติ อันจะยังให้ได้คะแนนสอบที่ไม่เหมาะสมกับความสามารถของผู้สอบแล้วก็จะ เป็นข้อมูลให้ผู้เขียนข้อสอบตระหนักในการใช้ตัวเลือกลักษณะนั้นต่อไป

ส่วนคำตอบเกี่ยวกับโค้งลักษณะของข้อสอบ ตลอดจนฟังก์ชันสารสนเทศของข้อสอบและฟังก์ชันสารสนเทศของแบบสอบ หากพบว่าลักษณะของตัวเลือกมีผลกระทบต่อโค้งลักษณะของข้อสอบฟังก์ชันสารสนเทศของข้อสอบและฟังก์ชันสารสนเทศของแบบสอบแล้ว ข้อมูลที่ได้ก็จะ เป็นประโยชน์แก่ผู้เขียนข้อสอบในการเลือกใช้ตัวเลือกในลักษณะที่จะให้ค่าสารสนเทศของข้อสอบและสารสนเทศของแบบสอบที่สูงกว่าซึ่งจะให้ความแม่นยำในการวัดมากกว่า

วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์ เพื่อใช้โมเดลโลจิสติกที่มีพารามิเตอร์ 3 ตัว (Three Parameter Logistic Model) ในการศึกษาผลของตัวเลือกลักษณะต่าง ๆ ของข้อสอบแบบเลือกตอบที่ใช้อยู่ในสภาพการณ์ของการทดสอบทั่ว ๆ ไปอันได้แก่ ตัวเลือกตัด ตัวเลือกเฉียด ตัวเลือกปลายเปิด ตัวเลือกแปลกและตัวเลือกให้เติม ที่มีต่อ

1. โค้งลักษณะของข้อสอบ (Item Characteristic Curves)
2. ค่าฟังก์ชันสารสนเทศของข้อสอบ (Item Information Functions)
3. ค่าฟังก์ชันสารสนเทศของแบบสอบ (Test Information Functions)
4. รูปแบบการตอบสนองข้อสอบ (Response Patterns)

สมมติฐานของการวิจัย

ลักษณะอย่างหนึ่งของข้อสอบแบบเลือกตอบคือเปิดโอกาสให้ผู้ตอบสามารถตอบถูกต้องด้วยการเดาและโอกาสในการเดาแต่ละข้อ ได้ถูกต้องนั้น เป็นสัดส่วนกลับกับจำนวนตัวเลือกในข้อนั้น นั่นคือ โอกาสในการเดาถูกขึ้นอยู่กับจำนวนตัวเลือก ถ้าจำนวนตัวเลือกมีมากโอกาสในการเดาถูกก็น้อย แต่ถ้าจำนวนตัวเลือกน้อยลงโอกาสในการเดาถูกก็จะมากขึ้น ฉะนั้นผู้สอบที่มีความฉลาดในการตอบข้อสอบ (Test-wise) รู้จักใช้วิธีขจัดตัวเลือกที่ไม่สมเหตุผลออกไปทำให้จำนวนตัวเลือกลดลงโอกาสที่จะเดาถูกก็จะมากขึ้น เหตุการณ์นี้ผู้เขียนข้อสอบสามารถป้องกันมิให้เกิดขึ้นได้โดยการเขียนตัวลวงให้ดูสมเหตุผลมีความเป็นไปได้ที่จะเป็นคำตอบถูก ตัวลวงเช่นนี้มักใช้ความเข้าใจผิด ๆ หรือข้อผิดพลาดของนักเรียนมาเป็นตัวลวงดังเช่นแบบสอบฉบับตัวเลือกตัดที่ใช้ในการศึกษาคั้งนี้

ข้อสอบในแบบสอบฉบับตัวเลือกตัด ฉบับตัวเลือกเฉียด ฉบับตัวเลือกปลายเปิด และฉบับตัวเลือกแปลกนั้นแต่ละข้อมีคำตอบมาให้เลือก 4 คำตอบ ในการตอบข้อสอบผู้ตอบใช้ความสามารถในการจำได้ (Recognition) เลือกคำตอบมาคำตอบหนึ่ง ส่วนข้อสอบในแบบสอบฉบับตัวเลือกให้

เดิมนั้นมีคำตอบมาให้เลือก 3 คำตอบ คำตอบที่ 4 ผู้ตอบเป็นผู้เขียนเองหากไม่เลือกคำตอบใด ๆ ในจำนวน 3 คำตอบที่ให้มา ในกรณีที่คำตอบที่ให้มานั้น ไม่มีคำตอบถูกต้องเลย ผู้ตอบต้องใช้ความสามารถนึก (Recall) หาคำตอบเอง ฉะนั้นย่อมต้องใช้ความสามารถสูงกว่าการตอบข้อสอบชนิดตัวเลือกตัก ตัวเลือกเจียด ตัวเลือกปลายเปิดและตัวเลือกแปลก ผู้ตอบที่มีความสามารถสูงเท่านั้นก็จะมีโอกาสตอบข้อสอบถูกต้องได้มาก ผู้ตอบที่มีความสามารถปานกลางและความสามารถต่ำมีโอกาสตอบถูกต้องน้อย สำหรับกรณีที่คำตอบที่ให้มามีคำตอบถูกต้องรวมอยู่ด้วยข้อสอบชนิดตัวเลือกให้เติมจึงเสมือนข้อสอบที่มี 3 ตัวเลือก โอกาสที่จะตอบข้อสอบนี้ได้ถูกน่าจะมากกว่าข้อสอบตัวเลือกตักตัวเลือกเจียดตัวเลือกปลายเปิดและตัวเลือกแปลก

เนื่องจากผู้วิจัยมีความเห็นว่า ข้อสอบชนิดตัวเลือกตักเป็นข้อสอบที่มีลักษณะที่พึงประสงค์ หากปรับข้อสอบชนิดตัวเลือกตักเป็นตัวเลือกชนิดอื่น ๆ แล้วผลจะเป็นอย่างไร เมื่อพิจารณาลักษณะการตอบและโอกาสในการตอบถูกแล้วผู้วิจัยจึงตั้งสมมุติฐานของการวิจัย ดังนี้

1. โฉงลักษณะของข้อสอบในแบบสอบจับตัวเลือกตักแตกต่างจากโฉงลักษณะของข้อสอบในแบบสอบจับตัวเลือกให้เติม แต่ไม่แตกต่างจากโฉงลักษณะของข้อสอบในแบบสอบจับตัวเลือกเจียด จับตัวเลือกปลายเปิดและจับตัวเลือกแปลก
2. ข้อสอบและแบบสอบจับตัวเลือกตักมีความแม่นยำเท่ากับข้อสอบ และแบบสอบจับตัวเลือกเจียด จับตัวเลือกปลายเปิด และจับตัวเลือกแปลกในการประมาณความสามารถของผู้สอบทุกระดับความสามารถเมื่อพิจารณาจากค่าดัชนีประสิทธิภาพเชิงสัมพัทธ์ของข้อสอบและแบบสอบ
3. ข้อสอบและแบบสอบจับตัวเลือกให้เติมมีความแม่นยำสูงกว่าข้อสอบและแบบสอบจับตัวเลือกตักเมื่อใช้ประมาณความสามารถของผู้สอบที่มีความสามารถระดับสูงขึ้นไป ($\theta > 1$) แต่มีความแม่นยำต่ำกว่า เมื่อใช้ประมาณความสามารถของผู้สอบที่มีความสามารถระดับปานกลางลงไป ($\theta < 1$) เมื่อพิจารณาจากค่าดัชนีประสิทธิภาพเชิงสัมพัทธ์ของข้อสอบและแบบสอบ
4. รูปแบบการตอบสนองข้อสอบในแบบสอบจับตัวเลือกตักมีความเหมาะสมกับระดับความสามารถของผู้สอบเท่ากับรูปแบบการตอบสนองข้อสอบในแบบสอบจับตัวเลือกเจียด จับตัวเลือกปลายเปิดและจับตัวเลือกแปลก แต่มีความเหมาะสมกับระดับความสามารถของผู้สอบน้อยกว่ารูปแบบการตอบสนองข้อสอบในแบบสอบจับตัวเลือกให้เติม สำหรับผู้สอบที่มีความสามารถระดับสูงขึ้นไป ($\theta > 1$) เมื่อพิจารณาจากค่าดัชนีความเหมาะสมของรูปแบบการตอบสนองข้อสอบ

ขอบเขตของการวิจัย

ตัวแปรที่ศึกษา ได้แก่ตัวแปรดังต่อไปนี้

1. **ตัวแปรอิสระ** คือ รูปแบบตัวเลือกของข้อสอบในแบบสอบซึ่งมี 5 แบบ ดังนี้
 - 1.1 ตัวเลือกตัดก
 - 1.2 ตัวเลือกเจียด
 - 1.3 ตัวเลือกปลายเปิด
 - 1.4 ตัวเลือกแปลก
 - 1.5 ตัวเลือกให้เติม

2. **ตัวแปรตาม** มี 6 ตัวแปร ดังนี้
 - 2.1 โด่งลักษณะของข้อสอบ
 - 2.2 ค่าฟังก์ชันสารสนเทศของข้อสอบ
 - 2.3 ค่าฟังก์ชันสารสนเทศของแบบสอบ
 - 2.4 ค่าดัชนีประสิทธิภาพเชิงสัมพัทธ์ของข้อสอบ
 - 2.5 ค่าดัชนีประสิทธิภาพเชิงสัมพัทธ์ของแบบสอบ
 - 2.6 ค่าดัชนีความเหมาะสมของรูปแบบการตอบสนองข้อสอบ

ผลวิจัยที่ใช้ในการศึกษารั้งนี้ คือ นักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2531 โรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ในจังหวัดนครราชสีมา ผู้วิจัยได้แบ่งเป็น 2 กลุ่มคือ กลุ่มที่ทดลองใช้แบบสอบจำนวน 173 คนเป็นนักเรียนโรงเรียนกุดจิกวิทยา และกลุ่มที่ใช้รวบรวมข้อมูลเพื่อการวิเคราะห์ตามวัตถุประสงค์ของการวิจัยจำนวน 1965 คน เป็นนักเรียนโรงเรียนบุญวัฒนา โรงเรียนสุรธรรมพิทักษ์ โรงเรียนปัทมราชวิทยาลัยประชานรินทร์ โรงเรียนสูงเนินและโรงเรียน ขามทะเลสอวิทยา

แบบสอบที่ใช้ในการศึกษารั้งนี้มี 5 ฉบับ แต่ละฉบับประกอบด้วยข้อสอบแบบเลือกตอบ ชนิด 4 ตัวเลือกจำนวน 60 ข้อ แบบสอบฉบับที่มีรูปแบบตัวเลือกเป็นตัวเลือกตัดกใช้แบบสอบวัดผลสัมฤทธิ์เชิงวินิจฉัย วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 เรื่องสมการ ที่ประดิษฐ์ เรื่องตระกูล (2529) สร้างขึ้น ส่วนแบบสอบอีก 4 ฉบับคือ แบบสอบที่มีตัวเลือกเจียด ตัวเลือกปลายเปิด ตัวเลือกแปลกและตัวเลือกให้เติมนั้นผู้วิจัยดัดแปลงมาจากแบบสอบฉบับแรก

ในการศึกษาคั้งนี้ใช้โมเดลโลจิสติกที่มีพารามิเตอร์ 3 ตัว (Three - Parameter Logistic Model) เป็นกรอบในการศึกษา

ข้อตกลงเบื้องต้น

1. การศึกษาคั้งนี้ถือว่าการสอบในเวลา สถานที่ และมีผู้ดำเนินการสอบแตกต่างกัน ไม่มีผลต่อการตอบสนองข้อสอบของนักเรียนที่เป็นพลวิชัย
2. พลวิชัยตอบแบบสอบอย่างเต็มความสามารถและอย่างเป็นอิสระเมื่อได้รับรู้ว่าเป็นการสอบเพื่อประเมินคุณภาพของโรงเรียนและเป็นการสอบวัดความสามารถทางคณิตศาสตร์เพื่อจัดนักเรียนเข้าแผนการเรียน

ความไม่สมบูรณ์ของการวิจัย

ในการประมาณค่าพารามิเตอร์ a ตามความเห็นของลอร์ด (Lord) ค่าประมาณที่ได้จะมีความแม่นยำสูงเมื่อประมาณจากแบบสอบที่มีข้อสอบจำนวน 50 ข้อขึ้นไป และจากการตอบของผู้สอบ 1000 คนขึ้นไป อย่างไรก็ตามลอร์ดไม่ได้แสดงข้อมูลเชิงประจักษ์ยืนยันเห็นดังกล่าว แต่มีการศึกษาของสวามินาธานและกิฟฟอร์ด (Swaminathan and Gifford 1979 อ้างถึงใน Hulin and others 1983 : 99 - 100) ซึ่งเป็นการศึกษาจากข้อมูลที่สร้างขึ้นด้วยเครื่องคอมพิวเตอร์ โดยศึกษาถึงความแม่นยำของค่าพารามิเตอร์ของข้อสอบที่ประมาณจากแบบสอบซึ่งมีความยาว 10, 15, 20 และ 80 ข้อ กับผู้สอบจำนวน 50, 200 และ 1000 คน โดยใช้โมเดลโลจิสติกที่มีพารามิเตอร์ 3 ตัว ความแม่นยำพิจารณาจากค่าสหสัมพันธ์แบบเพียร์สันระหว่างค่าพารามิเตอร์ที่แท้จริงกับค่าที่ประมาณได้

ผลการศึกษาพบว่าในการประมาณค่าพารามิเตอร์ a นั้นเมื่อแบบสอบมีความยาวเพิ่มขึ้น และจำนวนผู้สอบมากขึ้น ค่าสหสัมพันธ์ระหว่างค่าที่แท้จริงกับค่าประมาณจะเพิ่มขึ้นด้วย เช่นเดียวกับการประมาณค่าพารามิเตอร์ b แต่ในการประมาณค่า b พบข้อสังเกตว่าเมื่อใช้ข้อสอบเพียง 10 ข้อกับผู้สอบเพียง 50 คน ค่าสหสัมพันธ์ระหว่าง b และ \hat{b} สูงกว่า .80 ซึ่งจะเห็นว่าในการประมาณค่า b แม้จะใช้จำนวนผู้สอบน้อยความแม่นยำของค่าประมาณที่ได้ก็อยู่ในระดับค่อนข้างสูง

ส่วนการประมาณค่า c นั้นข้อมูลที่ได้จากการศึกษาวิจัยที่ผ่านมา ยังไม่ได้ข้อสรุปที่ชัดเจนเกี่ยวกับจำนวนผู้สอบ

ดังนั้น ในการวิจัยครั้งนี้ผู้วิจัยใช้แบบสอบซึ่งมีข้อสอบฉบับละ 60 ข้อ และใช้ผู้สอบฉบับละประมาณ 400 คนนั้น ในแง่ของจำนวนผู้สอบตามแนวคิดของ ลอร์ด อาจจะยังไม่สมบูรณ์ แต่จากข้อค้นพบบางส่วนที่กล่าวมาข้างต้นก็ยังไม่ได้แสดงอย่างชัดเจน ว่าต้องเป็น 1000 คนเสมอไป ดังเช่น ในการประมาณค่าพารามิเตอร์ b ซึ่งใช้ข้อสอบเพียง 10 ข้อกับผู้สอบเพียง 50 คนก็ยังให้ค่าสหสัมพันธ์ระหว่างค่าที่แท้จริงกับค่าที่ประมาณได้สูงดังกล่าวแล้ว

นิยามศัพท์ที่ใช้ในการวิจัย

1. **แบบสอบเลือกตอบธรรมดา** หมายถึง แบบสอบที่ประกอบด้วยข้อสอบแบบเลือกตอบที่มี 4 ตัวเลือกและมีตัวเลือกที่เป็นคำตอบที่ถูกต้องเพียงตัวเลือกเดียว
2. **แบบสอบเลือกตอบกับตัวเลือกตัด** หมายถึง แบบสอบเลือกตอบธรรมดาที่มีตัวลวงซึ่งเป็นคำตอบจากความเข้าใจผิด วิธีการคิดคำนวณที่ผิด ความสะเพร่าในการคิดคำนวณ การจำคลาดเคลื่อน คำตอบเหล่านี้ผู้สร้างแบบสอบสำรวจมาก่อนที่จะสร้างแบบสอบ
3. **แบบสอบเลือกตอบกับตัวเลือกเจ็ด** หมายถึง แบบสอบเลือกตอบธรรมดาที่มีตัวลวงซึ่งมีค่าใกล้เคียงกับคำตอบที่ถูกต้อง หรือเป็นประโยชน์คณิตศาสตร์ที่ใช้สัญลักษณ์สื่อเลียนคำตอบที่ถูกต้อง
4. **แบบสอบเลือกตอบกับตัวเลือกปลายเปิด** หมายถึง แบบสอบเลือกตอบธรรมดาที่มีตัวเลือกในลำดับสุดท้ายของแต่ละข้อเป็นตัวเลือกปลายเปิด **ไม่มีข้อใดถูก** ตัวเลือกนี้ทั้งที่เป็นคำตอบที่ถูกต้องและตัวลวง
5. **แบบสอบเลือกตอบกับตัวเลือกแปลก** หมายถึงแบบสอบเลือกตอบธรรมดาที่ข้อสอบแต่ละข้อมีตัวเลือกตัวหนึ่งแปลกไปจากตัวเลือกตัวอื่น ๆ เนื่องจากใช้สัญลักษณ์หรือความรู้ทางคณิตศาสตร์ที่สูงกว่าความรู้ในระดับชั้นมัธยมศึกษาปีที่ 1 หรือใช้ข้อความที่ไม่ค่อยพบบ่อยนักในแบบสอบวิชาคณิตศาสตร์ทั่ว ๆ ไปได้แก่ **ไม่พบคำตอบที่ต้องการ** **ไม่สามารถหาค่าได้** **มีคำตอบถูกต้องมากกว่า 1 ข้อ** **มีได้หลายค่า** ตัวเลือกแปลกเหล่านี้เป็นได้ทั้งคำตอบที่ถูกต้องและตัวลวง

6. แบบสอบเลือกตอบฉบับตัวเลือกให้เติม หมายถึง แบบสอบเลือกตอบธรรมดาที่แต่ละข้อมีตัวเลือกในลำดับสุดท้ายคือ คำตอบที่ถูกต้องคือ..... ผู้สอบที่เลือกตัวเลือกนั้นจะต้องเติมคำตอบที่ตนคิดว่าถูกต้องลงในที่ว่างที่เว้นไว้ให้ ตัวเลือกนี้เป็นได้ทั้งคำตอบที่ถูกต้องและตัวลวง ในกรณีที่ตัวเลือกนี้เป็นคำตอบที่ต้องการให้เลือก จะถือว่าผู้ตอบตอบถูกต้องเมื่อคำตอบที่เติมนั้นเป็นคำตอบที่ถูกต้อง

7. โด่งลักษณะของข้อสอบ(Item Characteristic Curve) หมายถึง กราฟของฟังก์ชันการตอบสนองข้อสอบซึ่งแสดงความสัมพันธ์ระหว่างระดับความสามารถของผู้สอบกับความน่าจะเป็นที่ผู้สอบซึ่งมีความสามารถระดับนั้นจะตอบข้อสอบข้อหนึ่ง ๆ ได้ถูกต้อง

8. ค่าฟังก์ชันสารสนเทศของข้อสอบ(Item Information Function) หมายถึง ความแม่นยำในการประมาณค่าความสามารถจริงของผู้สอบด้วยผลการตอบสนองข้อสอบข้อนั้นๆ ซึ่งมีค่าเท่ากับอัตราส่วนของกำลังสองของความชันของเส้นโด่งลักษณะของข้อสอบต่อค่าความแปรปรวนของข้อสอบข้อนั้น เขียนเป็นประโยคคณิตศาสตร์ได้ดังนี้

$$I(\theta, u_1) = [P_1'(\theta)]^2 / P_1(\theta) Q_1(\theta)$$

เมื่อ $P_1(\theta)$ คือฟังก์ชันของเส้นโด่งลักษณะของข้อสอบที่ระดับความสามารถ θ

$Q_1(\theta)$ คือ $1 - P_1(\theta)$

$P_1'(\theta)$ คืออนุพันธ์ของ $P_1(\theta)$

9. ค่าฟังก์ชันสารสนเทศของแบบสอบ(Test Information Function) หมายถึง ผลรวมของค่าฟังก์ชันสารสนเทศของข้อสอบในแบบสอบฉบับนั้น เขียนเป็นประโยคคณิตศาสตร์ได้ดังนี้

$$I(\theta) = \sum_{i=1}^n \{ [P_i'(\theta)]^2 / P_i(\theta) Q_i(\theta) \}$$

10. ประสิทธิภาพเชิงสัมพัทธ์ของข้อสอบ หมายถึง อัตราส่วนระหว่างค่าฟังก์ชันสารสนเทศของข้อสอบที่นำมาเปรียบเทียบกับกัน

11. ประสิทธิภาพเชิงสัมพัทธ์ของแบบสอบ หมายถึง อัตราส่วนระหว่างค่าฟังก์ชันสารสนเทศของแบบสอบที่นำมาเปรียบเทียบกับกัน

12. รูปแบบการตอบสนองข้อสอบ(Response Pattern) หมายถึง ลักษณะการตอบข้อสอบแต่ละข้อถูกหรือผิดในแบบสอบฉบับหนึ่ง ๆ ของผู้สอบแต่ละคน

13. ค่าดัชนีความเหมาะสมของรูปแบบการตอบสนองข้อสอบ หมายถึง ค่าที่ชี้ว่าผู้สอบคนหนึ่ง ๆ ทำแบบสอบนั้นได้เช่นเดียวกับผู้สอบคนอื่น ๆ ที่มีความสามารถระดับเดียวกันหรือไม่ ในการศึกษาครั้งนี้หมายถึงค่าของ Standardized l_0 หรือ Z_3 ซึ่งคำนวณได้จากสมการต่อไปนี้

$$Z_3 = [l_0 - M(\hat{\theta})] / [S(\hat{\theta})]^{1/2}$$

เมื่อ l_0 คือ ล็อกการิทึมของโลจิสติกไลค์ลิฮูดฟังก์ชันกรณีมี 3 พารามิเตอร์ (Three - Parameter Logistic Likelihood Function) ที่เป็นค่าประมาณไลค์ลิฮูดสูงสุดของ θ

$$l_0 = \sum_{i=1}^n [u_i \log P_i(\hat{\theta}) + (1 - u_i) \log Q_i(\hat{\theta})]$$

u_i คือคะแนนข้อที่ i (ตอบถูก = 1, ตอบผิด = 0)

$M(\hat{\theta})$ คือ Conditional Expectation ของ l_0 เมื่อกำหนด $\theta = \hat{\theta}$

$$M(\hat{\theta}) = \sum_{i=1}^n [P_i(\hat{\theta}) \log P_i(\hat{\theta}) + Q_i(\hat{\theta}) \log Q_i(\hat{\theta})]$$

$S(\hat{\theta})$ คือ Conditional Variance

$$S(\hat{\theta}) = \sum_{i=1}^n P_i(\hat{\theta}) Q_i(\hat{\theta}) \{\log [P_i(\hat{\theta}) / Q_i(\hat{\theta})]\}^2$$

14. ระดับความสามารถ หมายถึง ระดับความรู้ความเข้าใจในเนื้อหาสาระวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1 เรื่อง สมการ ซึ่งวัดในหน่วยของ θ ที่มีค่าเฉลี่ยเท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 ในการศึกษาครั้งนี้แบ่งระดับความสามารถออกเป็น 5 ระดับ คือ

ระดับต่ำมาก	$\theta < -2.5$
ระดับต่ำ	$-2.5 < \theta < -1.0$
ระดับปานกลาง	$-1.0 < \theta < 1.0$
ระดับสูง	$1.0 < \theta < 2.5$
ระดับสูงมาก	$2.5 < \theta$

ประโยชน์ที่คาดว่าจะได้รับ

1. ข้อความรู้อันได้จากการศึกษาคั้งนี้จะเป็นประโยชน์สำหรับผู้เขียนข้อสอบในการพิจารณาเลือกใช้ลักษณะตัวเลือกในการเขียนข้อสอบแบบเลือกตอบให้เหมาะสมกับวัตถุประสงค์ของการใช้แบบสอบนั้น กล่าวคือ หากต้องการข้อสอบที่จะให้ค่าพารามิเตอร์ของข้อสอบตามที่ต้องการ เช่น ค่าอำนาจจำแนกที่สูงกว่า ค่าการเดาที่น้อยกว่า และค่าความยากมากน้อยตามที่ต้องการนั้น ควรใช้ตัวเลือกลักษณะใดก็สามารถพิจารณาเลือกได้จากโค้งลักษณะของข้อสอบตัวเลือกแบบต่าง ๆ ถ้าต้องการใช้แบบสอบนั้นเพื่อการประมาณค่าความสามารถของผู้สอบก็เลือกใช้ลักษณะตัวเลือกที่ให้ค่าฟังก์ชันสารสนเทศสูงที่สุดในระดับความสามารถนั้น ๆ แต่ถ้าเป็นแบบสอบที่ใช้ในการคัดเลือกเพื่อจุดมุ่งหมายอย่างใดอย่างหนึ่งก็ควรใช้ลักษณะตัวเลือกที่ให้ค่าดัชนีความเหมาะสมของรูปแบบการตอบสนองข้อสอบที่สูงกว่า

2. ผลจากการศึกษาคั้งนี้จะก่อให้เกิดแนวคิดและช่องทางในการศึกษาวิจัย เรื่องของข้อสอบและการประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบให้กว้างขวางยิ่งขึ้นต่อไปอีก ในแง่ของการพัฒนาแบบสอบให้เหมาะสมกับสถานการณ์ต่าง ๆ ของการสอบ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย