



บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การวัดและการประเมินผลการศึกษาโดยเฉพาะการทดสอบ เป็นการนำเครื่องมือที่ใช้วัดคุณลักษณะต่าง ๆ เพื่อไปเร้าให้บุคคลนั้นตอบสนองออกมาให้สังเกตได้ ปัจจุบัน เครื่องมือที่นิยมใช้กันอย่างแพร่หลายและมีบทบาทมากในวงการศึกษาคือแบบสอบชนิดเลือกตอบ เนื่องจากมีข้อดีหลายประการ เช่น สามารถถามได้ครอบคลุมเนื้อหาและวัดสมรรถภาพในด้านต่าง ๆ ได้อย่างกว้างขวาง แต่เมื่อพิจารณาถึงวิธีการตรวจให้คะแนนจากการสอบด้วยแบบสอบประเภทนี้ตามวิธีประเพณีนิยมทั่ว ๆ ไป คือ การกำหนดน้ำหนักคะแนนรายข้อเท่ากับแบบ ๑-1 นั่นคือ ตอบถูกได้ 1 คะแนน แต่ถ้าตอบผิดจะได้ ๐ คะแนน การให้คะแนนโดยวิธีนี้อาจไม่ยุติธรรมสำหรับผู้สอบ เพราะผู้สอบบางคนอาจได้คะแนนสูงกว่าความเป็นจริงเนื่องจากการเดา ดังนั้นด้วยความพยายามในการแก้ปัญหาการให้คะแนนตามแบบประเพณีนิยมทั่ว ๆ ไปนั้น ได้มีนักวัดผลเสนอแนะวิธีการให้คะแนนที่แตกต่างกันออกไปในหลายรูปแบบ เช่น

คัมบส์ มิลฮอลแลนด์ และ โวเมอร์ (Combs, Milholland and Womer, 1956) ได้เสนอให้ผู้ตอบเลือกคำตอบที่ผิดแทนคำตอบที่ถูก ส่วนตัวเลือกที่ถูกหรือไม่แน่ใจให้เว้นไว้ว่าการให้คะแนนจะได้ 1 คะแนน เมื่อเลือกตัวเลือกที่ผิดจริง และให้ $1 - k$ คะแนน เมื่อเลือกตัวเลือกที่ถูก (k คือ จำนวนตัวเลือก) อนันต์ ศรีโสภณ (2516) เสนอให้ผู้ตอบตอบทุกตัวเลือก โดยทำเครื่องหมายให้ตรงกับความจริง การให้คะแนน ตัวเลือกที่ทำเครื่องหมายตรงกับเฉลยจะได้ 1 คะแนน แต่ถ้าทำตรงข้ามกับเฉลยจะได้ -1 คะแนน ตัวเลือกที่เว้นไว้จะได้ ๐ คะแนน ทั้งสองวิธีนี้พบว่ายังมีจุดอ่อน กล่าวคือ ถ้าผู้ตอบรู้ว่าตัวเลือกใดถูกแต่ไม่รู้ว่าตัวเลือกใดผิดเขาก็สามารถทำคะแนนข้อนั้นได้เท่า ๆ กันกับผู้รู้ที่รู้ว่าตัวเลือกใดถูกและตัวเลือกใดผิด ทั้ง ๆ ที่ผู้ตอบทั้ง 2 คน มีความรู้ไม่เท่ากัน

วิธีการให้คะแนนอีกวิธีหนึ่ง คือ การระบุความมั่นใจในการตอบ นั่นคือ นอกจากผู้ตอบจะเลือกตอบตัวเลือกที่คิดว่าถูกต้องแล้ว ยังต้องตัดสินใจอีกว่าตัวเลือกที่เป็นคำตอบที่ถูกต้องนั้นผู้ตอบมีระดับความมั่นใจในการตอบ (Confidence Weight Scoring) มากน้อยเพียงใด การได้คะแนนนั้นถ้าผู้ตอบตอบถูกต้องด้วยความมั่นใจมากที่สุดจะได้คะแนนสูงสุด แต่ถ้าผู้ตอบตอบผิดด้วยความมั่นใจมากที่สุดจะได้คะแนนติดลบมากที่สุด สิ่งที่ควรพิจารณาอีกคือ การที่ผู้ตอบมีความมั่นใจในระดับสูงว่า คำตอบที่ตนเลือกเป็นคำตอบที่ถูกต้อง แต่กลับปรากฏว่าผิด ผู้ตอบจะได้คะแนนน้อยกว่าผู้ที่ตอบว่าไม่มั่นใจเลย นั่นเป็นสิ่งที่ยุติธรรมหรือไม่ เพราะผู้ที่มั่นใจมากแต่ผิดถือว่าเป็นผู้ที่มีความรู้ที่ผิด (Misinformation) ส่วนผู้ที่ไม่มั่นใจนั้นถือว่าตอบโดยการเดา ซึ่งผู้ที่มีความรู้ไม่ถูกต้องไม่น่าจะถือว่ามีความสามารถน้อยกว่าผู้ที่ไม่รู้ (จักรกฤษณ์ สาราणीใจ, 2531)

นักวัดผลอีกกลุ่มหนึ่งเสนอวิธีตอบได้หลายตัวเลือก เช่น กิบบอนส์ โอลกิน และโซเบล (Gibbons, Olkin and Sobel, 1979) เสนอวิธีการตอบและวิธีการตรวจที่เรียกว่า Subset Selection Technique ในแบบสอบตัวเลือกถูกตัวเดียว โดยผู้ตอบสามารถเลือกตัวเลือกได้ตั้งแต่ 1 จนกระทั่งถึง $k - 1$ ตัวเลือก ที่ผู้ตอบคาดว่าจะมีตัวเลือกเลือกที่ถูกรวมอยู่ด้วย เมื่อ k คือจำนวนตัวเลือกทั้งหมดในแต่ละข้อ การได้คะแนนของผู้ตอบจะขึ้นกับจำนวนตัวเลือกที่ผู้ตอบเลือก ยิ่งจำนวนตัวเลือกเพิ่มขึ้นค่าของคะแนนจะลดลง ส่วนวิธีการตรวจให้คะแนนยึดถือหลักว่า คะแนนคาดหวัง (Expected Score) ที่ได้จากการเดาสุ่มในแต่ละข้อคำถามเป็นศูนย์ ในกรณีที่มี 5 ตัวเลือกและผู้ตอบอย่างสุ่ม 2 ตัวเลือกความน่าจะเป็นที่ตัวเลือกทั้ง 2 จะตอบถูกคือ 2 ใน 5 และความน่าจะเป็นที่เขาคงจะตอบผิดคือ 3 ใน 5 ดังนั้น คะแนนความคาดหวังจะเป็น

$$E(S) = 3(2/5) - 2(3/5) = 0$$

กรณีนี้ ถ้าตอบถูกจะได้ 3 คะแนน แต่ถ้าตอบผิดจะได้ - 2 คะแนน ซึ่งรูปแบบการให้คะแนนอาจแสดงได้ถึงตารางต่อไปนี้

ตารางการให้คะแนนตามวิธีของกิบบอนส์ และคณะ เมื่อมี 5 ตัวเลือก

จำนวนตัวเลือกที่ตอบ	คะแนนเมื่อมีตัวเลือกที่ถูกอยู่	คะแนนเมื่อไม่มีตัวเลือกที่ถูกอยู่
0	-	0
1	4	-1
2	3	-2
3	2	-3
4	1	-4
5	0	0

ซึ่งวิธีการนี้ กิบบอนส์ (Gibbons, 1979) กล่าวว่า น่าจะเป็นวิธีที่วัดส่วนของความรู้ (Partial Information) ของผู้สอบในเรื่องนั้น ๆ ได้ และยังสามารถแก้การเดาสุ่มได้อีกด้วย

อย่างไรก็ดี ปัจจุบันได้มีทฤษฎีวัดผลแนวใหม่ที่อ้างถึง ความสัมพันธ์ระหว่างความสามารถที่แท้จริงกับลักษณะการตอบข้อสอบเพื่อทำนายคุณลักษณะที่ซ่อนอยู่ภายในบุคคลที่เรียกว่า ทฤษฎีสองตอบข้อสอบ (Item Response Theory : IRT) เข้ามามีบทบาทในวงการวัดผลการศึกษาในปัจจุบันและเข้ามาแก้ปัญหาทฤษฎีการวัดแบบดั้งเดิม (Classical Test Theory) ที่พบว่ามีจุดอ่อนในบางประการ (Hambelton, 1979) คือ ประการที่หนึ่ง ค่าความยากและอำนาจจำแนกที่แปรเปลี่ยนไปตามกลุ่มผู้สอบประการที่สอง การเปรียบเทียบความสามารถระหว่างบุคคลจะเปรียบเทียบได้ต่อเมื่อใช้แบบสอบฉบับเดียวกัน และประการที่สาม ค่าความคลาดเคลื่อนมาตรฐานในการวัดของแต่ละบุคคลเท่ากัน แต่ลอร์ดและโนวิก (Lord and Novick, 1965) แสดงให้เห็นว่าความคลาดเคลื่อนมาตรฐานในการวัดของแต่ละบุคคล

ไม่เท่ากัน และภายใต้ทฤษฎีตอบสนองข้อสอบ (IRT) นี้ ลอร์ด (Lord 1980) ได้เสนอสูตรในการคำนวณค่าน้ำหนักคะแนนที่เหมาะสม (Optimal Scoring Weights) จากค่าพารามิเตอร์ต่าง ๆ ดังนี้

$$W_i = \frac{a_i}{1 - c_i} \cdot \frac{p_i - c_i}{p_i}$$

- ในเมื่อ W_i คือ ค่าน้ำหนักคะแนนที่เหมาะสมข้อที่ i
 a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i จากการวิเคราะห์ด้วย โลจิสติก 3 พารามิเตอร์
 c_i คือ ค่าการเดาของข้อสอบข้อที่ i จากการวิเคราะห์ด้วย โลจิสติก 3 พารามิเตอร์
 p_i คือ ค่าความยากง่ายของข้อสอบข้อที่ i ที่วิเคราะห์ได้จากวิธีประเพณีนิยม

วิธีการให้น้ำหนักคะแนนโดยสูตรคำนวณของลอร์ดนี้ เป็นวิธีที่ละเอียดมากเพราะคำนึงถึงความสัมพันธ์ของค่าพารามิเตอร์ที่สำคัญทั้งหมดเป็นรายข้อ และที่สำคัญอย่างยิ่ง คือ คำนึงถึงค่าการเดาอันเป็นแหล่งคลาดเคลื่อนที่สำคัญของการวัด

จากทฤษฎีและแนวคิดในเรื่องการตรวจให้คะแนนในการวัดส่วนของความรู้ (Partial Information) ของผู้สอบซึ่งน่าจะวัดได้ละเอียดกว่าวิธีการตรวจให้คะแนนแบบธรรมดาทั่ว ๆ ไป และยังสามารรถแก้การเดาลุ่มของผู้สอบ ซึ่งเสนอโดย กิบบอนส์และคณะ กับวิธีการให้คะแนนของลอร์ด ทั้ง 2 วิธีต่างเป็นวิธีการที่ต้องการให้คะแนนของข้อสอบในแต่ละข้อแตกต่างกันออกไป ทั้งนี้วัตถุประสงค์ที่สำคัญก็เพื่อต้องการที่จะแก้ปัญหาความไม่ยุติธรรมอันเกิดจากวิธีการให้น้ำหนักคะแนนรายข้อเท่ากันตามทฤษฎีการวัดแบบดั้งเดิม (Classical Test Theory) ซึ่งเมื่อพิจารณาแล้วพบว่า การให้คะแนนทั้งสองวิธีนี้ต่างก็คำนึงถึงการแก้ปัญหาเนื่องจากการเดา ซึ่งเป็นความพยายามที่จะทำให้คะแนนสอบที่ได้จาก

การวัดมีความเชื่อถือได้มากที่สุด แต่ในส่วนที่แตกต่างกันของวิธีการทั้งสองนี้ก็คือ การกำหนดน้ำหนักคะแนนโดยวิธีของกิบบอนส์ และคณะนั้นได้กำหนดขึ้นก่อนที่ผู้สอบจะทำการสอบ การที่ผู้สอบจะได้คะแนนต่างกัน ก็ขึ้นกับจำนวนตัวเลือกที่ผู้สอบจะเลือกตอบ การตรวจให้คะแนนอาจค่อนข้างยุ่งยากกว่าวิธีแบบประเพณีนิยมทั่ว ๆ ไป แต่ก็สามารถกระทำได้จากการตรวจของครูผู้สอนแบบธรรมดา ไม่ต้องผ่านกระบวนการวิเคราะห์ทางคอมพิวเตอร์แต่อย่างใด ส่วนวิธีการให้น้ำหนักคะแนนของลอร์ด การจะได้น้ำหนักคะแนนรายชื่อที่ต่างกันจะต้องคำนวณจากค่าพารามิเตอร์ของข้อสอบที่ได้จากข้อมูลเชิงประจักษ์ เมื่อกลุ่มผู้สอบได้ตอบสนองข้อสอบนั้นแล้ว และการจะได้ค่าพารามิเตอร์ก็จะต้องได้จากการวิเคราะห์ตามแนวทฤษฎีตอบสนองข้อสอบ (IRT) ด้วยเครื่องคอมพิวเตอร์ จึงทำให้น่าจะพิจารณาว่า วิธีการให้น้ำหนักคะแนนตามวิธีของกิบบอนส์และคณะซึ่งเป็นวิธีการให้น้ำหนักคะแนนรายชื่อต่างกัน ซึ่งแตกต่างไปจากวิธีการให้น้ำหนักคะแนนของ ลอร์ด จะสามารถทำให้ผลของการวัดสอดคล้องกันกับวิธีการให้น้ำหนักคะแนนของลอร์ดหรือไม่ ซึ่งผลจากการศึกษาค้นคว้าในการนิยามคะแนนสอบ จะทำให้เห็นภาพของวิธีการวัดว่าสอดคล้องกันหรือไม่ และวิธีการให้น้ำหนักคะแนนที่ต่างกัน 2 วิธี จะทำให้ค่าความเที่ยงและความตรงตามสภาพของแบบสอบต่างกันหรือไม่ วิธีใดจะให้ค่าที่สูงกว่า

ด้วยเหตุนี้ ผู้วิจัยจึงสนใจที่จะทำการศึกษาเปรียบเทียบถึงผลของวิธีการให้คะแนนที่ต่างกันระหว่างวิธีการให้น้ำหนักคะแนนของกิบบอนส์และคณะกับวิธีของลอร์ด

วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ

1. หาความสอดคล้องของคะแนนสอบระหว่างวิธีการให้น้ำหนักคะแนนที่ต่างกันระหว่างวิธีของกิบบอนส์และคณะ กับวิธีของลอร์ด
2. เปรียบเทียบค่าความเที่ยงของแบบสอบระหว่างวิธีการให้น้ำหนักคะแนนที่ต่างกัน ระหว่างวิธีของกิบบอนส์และคณะ กับวิธีของลอร์ด

3. เปรียบเทียบค่าความตรงตามสภาพของแบบสอบระหว่างวิธีการให้น้ำหนักคะแนนที่ต่างกัน ระหว่างวิธีของกิบบอนส์และคณะ กับวิธีของลอร์ด

สมมติฐานในการวิจัย

จากทฤษฎีของความเที่ยงที่พบว่า สูตรในการคำนวณหาค่าความเที่ยงของแบบสอบ คือ $r_{cc} = 1 - \frac{S^2_e}{S^2_x}$ เมื่อ r_{cc} คือค่าความเที่ยงของแบบสอบ S^2_e คือ ความแปรปรวนของความคลาดเคลื่อน และ S^2_x คือ ค่าความแปรปรวนของคะแนนสอบ นั้น แสดงว่า ค่าความเที่ยงจะมีค่าสูงขึ้น เมื่อความแปรปรวนของความคลาดเคลื่อนลดลง โดยกำหนดความแปรปรวนของคะแนนสอบให้คงที่ หรือมีค่าเพิ่มขึ้น ดังนั้น เมื่อพิจารณาถึงวิธีการให้น้ำหนักคะแนนตามวิธีของกิบบอนส์และคณะ ที่กำหนดให้ผู้ตอบสามารถเลือกตอบในแบบสอบเลือกตอบได้ตั้งแต่ 1 ถึง $k - 1$ ตัวเลือก (เมื่อ k คือ จำนวนตัวเลือกทั้งหมด) ที่ผู้ตอบคาดว่าจะมีตัวเลือกที่ถูกรวมอยู่ด้วย การจะได้คะแนนก็ขึ้นอยู่กับจำนวนตัวเลือกที่ผู้ตอบเลือกตอบ ยิ่งจำนวนตัวเลือกมากขึ้นค่าคะแนนจะลดลง โดยคำตอบถูกจะได้ $k - e$ คะแนน และถ้าตอบผิดจะได้ $-e$ คะแนน (เมื่อ e คือ จำนวนตัวเลือกที่ตอบ) เช่น ในกรณีแบบสอบนั้นเป็นแบบสอบเลือกตอบ 5 ตัวเลือก ดังนั้น ค่าของคะแนนในแต่ละข้อกระทงจึงอาจมีค่าได้ตั้งแต่ -4 ถึง $+4$ ส่วนวิธีการให้น้ำหนักคะแนนตามสูตรของลอร์ด การจะได้คะแนนในแต่ละข้อกระทง จะต้องคำนวณมาจากค่าพารามิเตอร์ต่าง ๆ ที่สำคัญเป็นรายชื่อของข้อสอบ โดยได้มาจากการวิเคราะห์ข้อมูลตามทฤษฎีตอบสนองข้อสอบ (Item Response Theory) ซึ่งเมื่อพิจารณาแล้วพบว่า ทั้งวิธีการให้น้ำหนักคะแนนตามวิธีของกิบบอนส์และคณะ กับวิธีการให้น้ำหนักคะแนนของลอร์ด ต่างก็เป็นวิธีการที่พยายามทำให้คะแนนจากผลการสอบมีการกระจายมากยิ่งขึ้นไปกว่า วิธีการให้คะแนนรายชื่อเท่ากับแบบ 0 - 1 ตามทฤษฎีการวัดแบบดั้งเดิม และจากการศึกษารายงานการวิจัยที่ศึกษาเกี่ยวกับวิธีการให้คะแนนในแบบต่าง ๆ ที่พบว่ามีความสัมพันธ์กัน เช่น สุนด์น สุกมลสันต์ (2528) ที่พบว่า วิธีการให้คะแนนที่เหมาะสมของลอร์ดมีความสัมพันธ์กับวิธีการให้คะแนนโดยใช้ค่าเดลต้า (Δ) วิธีการใช้ค่าเบต้าแปลง (β_c) วิธีการใช้ผลบวกความแปรแปรปรวนร่วม (h^2) และวิธีธรรมดาแบบประเพณีนิยมอย่างมีนัยสำคัญ งานวิจัยของพินิจ อุไรรักษ์ (2534) ที่พบว่า วิธีการให้น้ำหนักคะแนน

รายช้อตามค่าพารามิเตอร์ โดยสูตรของลอร์ด มีสหสัมพันธ์กับวิธีการให้น้ำหนักคะแนนรายช้อ
เท่ากันแบบ 0 - 1 วิธีการให้น้ำหนักคะแนน รายช้อตามค่าความสามารถของผู้สอบ สูง
มากคือ มีค่าตั้งแต่ .941 ถึง .983 ดังนั้น ผู้วิจัยจึงตั้งสมมติฐานการวิจัยดังนี้

สมมติฐานที่ 1 วิธีการให้น้ำหนักคะแนนที่ต่างกันระหว่างวิธีของกิบบอนส์และคณะ
กับวิธีของลอร์ดน่าจะมีความสัมพันธ์สอดคล้องกัน

สมมติฐานที่ 2 ค่าความเที่ยงของแบบสอบเมื่อมีการให้น้ำหนักคะแนนต่างกัน
ระหว่างตามวิธีของกิบบอนส์และคณะ กับวิธีของลอร์ดไม่น่าจะแตกต่างกัน

Mehren และ Lehman (1984) กล่าวว่าในการหาความตรงตามเกณฑ์สัมพัทธ์
นอกจากค่าความตรงจะขึ้นอยู่กับค่าความเที่ยงของตัวทำนายแล้ว ยังขึ้นกับค่าความเที่ยงของ
ตัวเกณฑ์ด้วย โดยค่าความตรงจะมีค่าเท่ากับ หรือไม่เกินรากที่สองของผลคูณของค่าความ
เที่ยงของตัวทำนาย กับค่าความเที่ยงของตัวเกณฑ์ดังสมการ $r_{xy} \leq \sqrt{(r_{xx})(r_{yy})}$
ดังนั้น ค่าความตรงตามสภาพของแบบสอบซึ่งเป็นความตรงตามเกณฑ์สัมพัทธ์แบบหนึ่ง ก็ควรจะ
ขึ้นกับค่าความเที่ยงของคะแนนสอบที่ใช้เป็นตัวทำนาย กับค่าความเที่ยงของผลการจัดลำดับ
ความสามารถของครูผู้สอนที่ใช้เป็นตัวเกณฑ์ด้วย และเนื่องจากค่าความเที่ยงของแบบสอบ
จากวิธีการให้น้ำหนักคะแนนตามวิธีของกิบบอนส์และคณะกับวิธีของลอร์ด ในสมมติฐานที่ 2
ไม่น่าจะต่างกัน อีกทั้งเกณฑ์ที่ใช้ในการศึกษาก็ใช้เกณฑ์โดยการจัดลำดับที่ความสามารถของ
ครูผู้สอนเดียวกัน ผู้วิจัยจึงตั้งสมมติฐานการวิจัยว่า

สมมติฐานที่ 3 ค่าความตรงตามสภาพของแบบสอบ เมื่อมีการให้น้ำหนักคะแนน
ต่างกันตามวิธีของกิบบอนส์และคณะ กับวิธีของลอร์ดไม่น่าจะแตกต่างกัน

ขอบเขตของการวิจัย

1. ในการวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบเกี่ยวกับวิธีการให้น้ำหนัก
คะแนนรายช้อจากการสอบที่แตกต่างกันเท่านั้น จึงไม่เน้นการศึกษาเกี่ยวกับระดับชั้น เพศ
อายุ กลุ่มตัวอย่างที่ผู้วิจัยใช้ในการวิจัยครั้งนี้จึงใช้นักเรียนมัธยมศึกษาปีที่ 3 (ม.3) ภาค

เรียนที่ 1 ปีการศึกษา 2533 ของสถานศึกษาสังกัดกองการมัธยมศึกษา กรมสามัญศึกษา กระทรวงศึกษาธิการ ในเขตจังหวัดปทุมธานีเท่านั้น

2. เครื่องมือที่ใช้ในการวิจัย เป็นแบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา วิทยาศาสตร์ ว.305 ชั้นมัธยมศึกษาปีที่ 3 เรื่องบรรยากาศ โลก...ดวงดาว และอวกาศ ที่ผู้วิจัยสร้างขึ้น

ข้อจำกัดของการวิจัย

เนื่องจากการวิจัยครั้งนี้ผู้วิจัยต้องการศึกษาเปรียบเทียบถึงผลของวิธีการให้คะแนนที่ต่างกัน 2 วิธี ระหว่างวิธีของกิบบอนส์และคณะ กับวิธีของลอร์ดต่อค่าความสอดคล้องของคะแนนสอบ ค่าความเที่ยง และค่าความตรงของแบบสอบจากผู้สอบคนเดียวกัน ผู้วิจัยจึงได้ให้นักเรียนที่เป็นกลุ่มตัวอย่างตอบแบบสอบทั้ง 2 ฉบับ โดยแบบสอบทั้ง 2 ฉบับ จะมีเนื้อหาข้อกระทงชุดเดียวกัน แต่จะต่างกันที่วิธีการตอบ และการตรวจให้คะแนน กล่าวคือ ฉบับที่ 1 ผู้สอบจะเลือกตอบได้ตั้งแต่ 1 ถึงหลายตัวเลือก ตามวิธีการตอบและการตรวจให้นำหนักคะแนนของกิบบอนส์และคณะ ส่วนฉบับที่ 2 ผู้สอบจะเลือกตอบได้เพียงตัวเลือกเดียว เพื่อนำไปวิเคราะห์คำนวณหาค่าน้ำหนักคะแนนโดยสูตรของลอร์ด ดังนั้น หลังจากผู้สอบได้ตอบแบบสอบฉบับที่ 1 ผ่านไปประมาณ 2 สัปดาห์ ผู้วิจัยจึงได้นำแบบสอบฉบับที่ 2 ไปสอบกับนักเรียนกลุ่มเดิม โดยในการสอบครั้งหลังนั้น ผู้วิจัยไม่ได้แจ้งให้นักเรียนที่เป็นกลุ่มตัวอย่างทราบมาก่อน อีกทั้งผลการสอบจะนำไปใช้ในการวิจัยเท่านั้น ไม่มีผลต่อการเรียนการสอนภายในห้องเรียน จึงคาดว่า นักเรียนไม่ได้ไปศึกษาเพิ่มเติมมาแต่อย่างใด อย่างไรก็ตาม ผู้วิจัยก็ไม่สามารถจะควบคุมตัวแปรดังกล่าวอย่างเป็นระบบได้ จึงอาจจะเป็นข้อจำกัดของการวิจัยในครั้งนี้

ข้อตกลงเบื้องต้น

ครูผู้สอนสามารถจัดลำดับที่ความสามารถด้านผลสัมฤทธิ์ทางการเรียนของนักเรียนได้ เนื่องจาก ครูผู้สอนคือผู้ที่ทำหน้าที่ประเมินผลการเรียน ทั้งผลการเรียนย่อยในระหว่างภาคกับผลการเรียนรวมในปลายภาค และสภาพการเรียนการสอนวิชาวิทยาศาสตร์ที่ครูผู้สอน

จะต้องควบคุมดูแลนักเรียนอย่างใกล้ชิด ทั้งการเรียนในภาคทฤษฎีและภาคปฏิบัติในห้องปฏิบัติการ จึงทำให้ครูมีโอกาสนใกล้ชิดและสามารถสังเกตพฤติกรรมที่คาดว่าจะมีผลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียนได้

คำจำกัดความที่ใช้ในการวิจัย

นักเรียน หมายถึง นักเรียนในชั้นมัธยมศึกษาปีที่ 3 ของสถานศึกษาสังกัดกองการมัธยมศึกษา กรมสามัญศึกษา กระทรวงศึกษาธิการ ภาคเรียนที่ 1 ปีการศึกษา 2533 ในเขตจังหวัด ปทุมธานี

แบบสอบวิชาวิทยาศาสตร์ หมายถึง แบบสอบวัดผลสัมฤทธิ์วิชาวิทยาศาสตร์ (ว.305) ชั้นมัธยมศึกษาปีที่ 3 เรื่องบรรยากาศโลก... ดวงดาวและอวกาศ ซึ่งผู้วิจัยสร้างขึ้นตามหลักสูตรของสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี เป็นแบบสอบเลือกตอบ 5 ตัวเลือก ที่มีตัวเลือกที่ถูกที่สุดเพียงตัวเลือกเดียว นอกนั้นเป็นตัวลวง

วิธีการให้น้ำหนักคะแนนของกิบบอนส์และคณะ หมายถึง วิธีการให้คะแนน ซึ่งในการตรวจให้คะแนนแต่ละข้อ ผู้ตอบสามารถเลือกตอบได้ตั้งแต่ 1 ถึง $k - 1$ ตัวเลือก โดยถ้าตอบถูกจะได้ $k - e$ คะแนน และถ้าตอบผิดจะได้ $-e$ คะแนน (เมื่อ k คือ จำนวนตัวเลือกทั้งหมด และ e คือ จำนวนตัวเลือกที่ตอบ) ส่วนข้อที่เว้นไว้ หรือเลือกทุกตัวเลือกจะได้ 0 คะแนน

คะแนนผลการสอบ = ผลรวมทางนิชคณิตของคะแนนทุกข้อ

วิธีการให้น้ำหนักคะแนนของลอร์ด หมายถึง การให้คะแนนรายข้อที่ตอบถูก โดยใช้
 ค่าน้ำหนักคะแนนที่คำนวณมาจากค่าอำนาจจำแนก (a_i)
 ค่าความยากง่าย (p_i) ที่วิเคราะห์โดยทฤษฎีมาตรฐาน
 ดั้งเดิม และค่าการเดา (c_i) จากสูตรของลอร์ด คือ

$$W_i = \frac{a_i}{1 - c_i} \cdot \frac{p_i - c_i}{p_i} \quad (\text{Lord, 1980})$$

โดยที่ค่าอำนาจจำแนก (a_i) และค่าการเดา (c_i)
 ได้จากการวิเคราะห์ข้อสอบรายข้อตามรูปแบบโลจิสติก
 3 พารามิเตอร์ และค่าความยาก (p_i) ได้จากการ
 วิเคราะห์ตามทฤษฎีการวัดแบบดั้งเดิม สำหรับข้อที่
 ตอบผิดจะได้คะแนนเป็น 0

คะแนนผลการสอบ = ผลรวมของคะแนนทุกข้อที่ตอบถูก

ความสอดคล้องของคะแนนสอบ หมายถึง ผู้สอบที่ได้คะแนนสอบจากวิธีการให้
 คะแนนสอบวิธีหนึ่งมาก จะได้คะแนนสอบจากอีก
 วิธีหนึ่งมาก หรือได้คะแนนสอบจากวิธีหนึ่งน้อย จะได้
 คะแนนสอบอีกวิธีหนึ่งน้อยด้วย

ความเที่ยงของแบบสอบ หมายถึง ค่าความคงที่ของผลการสอบที่คำนวณจากคะแนน
 ที่ได้จากวิธีให้น้ำหนัก คะแนนของกิบบอนส์และคณะ
 กับวิธีของลอร์ด โดยใช้สูตร สัมประสิทธิ์แอลฟา
 ของครอนบัค (Alpha's Cronbach Coefficient)

ความตรงตามสภาพของแบบสอบ หมายถึง ค่าความสัมพันธ์ระหว่างผลการจัดลำดับ
 ที่ของคะแนนสอบจากวิธีการให้น้ำหนักคะแนนตามวิธี
 ของกิบบอนส์และคณะ และวิธีของลอร์ด กับผลการจัด
 ลำดับที่ความสามารถของนักเรียน โดยครูผู้สอน

ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. ทำให้ทราบถึงผลของวิธีการให้น้ำหนักคะแนนสอบที่ต่างกันต่อค่าความสอดคล้องของคะแนนสอบ ค่าความเที่ยง และค่าความตรงของแบบสอบเลือกตอบ
2. เป็นแนวทางในการนำไปประยุกต์ใช้เพื่อเพิ่มคุณภาพของการให้น้ำหนักคะแนนแบบสอบเลือกตอบ เพื่อให้การวัดผลมีประสิทธิภาพมากที่สุด
3. เพิ่มพูนความรู้ทางด้านการวัดผลทางการศึกษา ซึ่งจะ เป็นประโยชน์ในวงการศึกษา และพัฒนาทฤษฎีวัดผลให้ก้าวหน้ายิ่งขึ้น



ศูนย์วิจัยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย