

บทที่ 2

วรรณคดีที่เกี่ยวข้อง

การวิจัยเพื่อพัฒนาแบบสอบความเข้าใจในการอ่านภาษาอังกฤษของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยการตรวจให้คะแนนแบบพหุวิภาค ในครั้งนี้ ผู้วิจัยได้ศึกษาเอกสาร หนังสือ และงานวิจัยที่เกี่ยวข้องกับ แบบสอบโคลซ (cloze test) และการประยุกต์ใช้แบบสอบแบบสอบโคลซ (cloze test) เป็นโมดิฟายด์โคลซ (modified cloze test) เพื่อใช้ในการวัดความเข้าใจในการอ่าน การเลือกเรื่อง จำนวนคำที่เหมาะสม และการตัดคำในแบบสอบโคลซ (cloze test) การอ่านและความเข้าใจในการอ่าน สมรรถภาพทางด้านทักษะการอ่านวิชาภาษาอังกฤษของนักเรียนระดับประถมศึกษา และการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบที่ตรวจให้คะแนนแบบพหุวิภาค (Polytomous IRT Model) ปรากฏรายละเอียดดังนี้

- ตอนที่ 1 แนวคิดเกี่ยวกับแบบสอบโคลซ (cloze test) และการประยุกต์ใช้แบบสอบโคลซ (cloze test) เป็นแบบสอบโมดิฟายด์โคลซ (modified cloze test) เพื่อใช้ในการวัดความเข้าใจในการอ่าน
- ตอนที่ 2 แนวคิดเกี่ยวกับการเลือกเรื่อง จำนวนคำที่เหมาะสม และการตัดคำในแบบสอบโคลซ (cloze test)
- ตอนที่ 3 แนวคิดเกี่ยวกับการอ่านและความเข้าใจในการอ่าน
- ตอนที่ 4 แนวคิดเกี่ยวกับสมรรถภาพทางด้านทักษะการอ่านวิชาภาษาอังกฤษ ของนักเรียนระดับประถมศึกษา
- ตอนที่ 5 แนวคิดเกี่ยวกับการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบที่ตรวจให้คะแนนแบบพหุวิภาค (Polytomous IRT Model)
- ตอนที่ 6 งานวิจัยที่เกี่ยวข้องกับการตรวจให้คะแนนแบบต่างๆ
- ตอนที่ 7 งานวิจัยที่เกี่ยวข้องกับการตรวจให้คะแนนแบบพหุวิภาค (Polytomous Scoring)
- ตอนที่ 8 งานวิจัยที่เกี่ยวข้องกับความเข้าใจในการอ่านที่ใช้แบบสอบโคลซ ในระดับชั้นประถมศึกษา

ตอนที่ 1 แนวคิดเกี่ยวกับแบบสอบโคลซ (cloze test) และการประยุกต์ใช้แบบสอบ โมดิฟายด์โคลซ (modified cloze test) เพื่อใช้ในการวัดความเข้าใจในการอ่าน

สุพัฒน์ สุขมลสันต์ (2530) ได้กล่าวถึงทฤษฎีและแนวคิดของ cloze test ว่ามีแนวคิดเชิงทฤษฎีสนับสนุน 3 แนวคิดคือ

1. หลักการของการปิดของนักจิตวิทยากลุ่มเกสตัลท์ (Gestalt principle of Closure)

โคลซ (cloze) เป็นวิธีการที่วิลสัน แอล เทย์เลอร์ คิดขึ้นในปี 1953 โดยยึดหลักจิตวิทยาแบบ Gestalt ซึ่งใช้ทฤษฎีที่เรียกว่า “closure” คือ การเพิ่มสิ่งที่ขาดลงไปให้สมบูรณ์ได้ในมโนภาพ เช่น เราคุ้นเคยกับภาพใดภาพหนึ่ง เมื่อเรามาดูที่เคยเห็นขาดตอนไป เราก็สามารถเติมให้สมบูรณ์ได้ เทย์เลอร์ได้นำหลักการนี้มาใช้ทำแบบทดสอบ เพื่อวัดความเข้าใจในการอ่านและทดสอบว่า ผู้ทำสามารถใช้คำเติมใส่ในช่องว่างที่ถูกตัดตอนไปให้เหมาะสมกับเนื้อหาโดยใช้คำชี้แนะ (Clues) ที่มีอยู่ ถ้าผู้อ่านสามารถเติมคำลงในช่องว่างได้ตรงกับคำเดิม หรือความหมายเติมแสดงว่าผู้อ่านเข้าใจเรื่องที่อ่าน ทั้งนี้เพราะการที่ผู้อ่านจะเติมคำได้ถูกต้องนั้นนอกจากจะเข้าใจความหมายของคำที่จะเติม ลักษณะและแบบแผนของภาษาแล้ว ผู้อ่านจะต้องเข้าใจความหมายของข้อความที่ต้องตัดคำนั้นทั้งหมดด้วย จึงจะสามารถเติมคำได้ถูกต้องและสอดคล้องกับข้อความนั้น

2. ทฤษฎีสารสนเทศ

การเติมคำใน cloze test คำที่สรรหามาเติมจะเป็นคำที่คาดหมายได้โดยง่ายหรือยากขึ้นอยู่กับช่องว่างที่จะเติมว่าสามารถเติมคำต่างๆได้มากหรือน้อย ถ้าเป็นช่องว่างที่เติมคำต่างๆกันได้มาก ก็แสดงว่าช่องว่างนั้นให้สารสนเทศมาก ส่วนช่องว่างใดที่มีข้อความเชิงประกอบเป็นเชิงบังคับ หรือมีส่วนบอกล่วงหน้าว่าคำที่ตามมาควรเป็นคำอะไรหรือคำชนิดใด ช่องว่างนั้นจะถือว่าให้สารสนเทศน้อย (อมรรรัตน์ แกมทอง, 2530)

3. ทฤษฎีลำดับชั้นการตอบสนอง (Response Hierarchy)

การเลือกคำมาเติมใน cloze test จะเลือกจากลำดับชั้นความคุ้นเคย หรือจากประสบการณ์ต่อแบบแผนในการใช้คำนั้นๆ ทั้งนี้ขึ้นอยู่กับความรู้ของแต่ละคนที่มีต่อคำนั้นในหลักการทางไวยากรณ์และความเคยชินในการใช้คำคำนั้นกับเนื้อหาที่อ่าน

cloze test เป็นแบบสอบที่สามารถวัดความเข้าใจในการอ่านได้ดีที่สุดวิธีหนึ่งเพราะ cloze test มีคุณสมบัติเด่นๆ ดังนี้ (John W.Oller, 1979 อ้างถึงใน สุพัฒน์ สุขมลสันต์, 2530)

1. cloze test ไม่เหมือนกับแบบทดสอบที่ให้ผู้สอบเติมคำลงในประโยคแต่ละประโยคตามที่เคยเห็นทั่วไปแต่ประโยคต่างๆในแบบสอบโคลสนี้ จะต้องมีใจความต่อเนื่องกันจนหมดข้อความ
2. การเติมคำในช่องว่างให้ได้ความนั้น จะทำได้โดยอาศัยคำซ้ำซ้อน (Redundancy) จากเนื้อความที่ต่อเนื่องกันในข้อความที่ให้และความเข้าใจในระบบของภาษานั้นๆของผู้สอบ
3. cloze test สามารถทดสอบความเข้าใจและการใช้ภาษาเกี่ยวกับเวลา(Temporal) การลำดับข้อความและการวางตำแหน่งอันเหมาะสมของคำที่เติมในข้อความที่ให้อ่าน
4. วิธีการ cloze เป็นวิธีการวัดความเข้าใจในการอ่านที่มีความตรงและความเชื่อถือได้สูงกว่าวิธีการอื่น
5. คะแนนจากการทดสอบโดยใช้ cloze test สามารถใช้ประเมินผลประสิทธิภาพในการเรียนการสอนได้
6. คะแนนจากการทดสอบโดยใช้ cloze test สามารถประเมินพฤติกรรมในการเรียนของการเรียนได้

เนื่องจากการที่จะให้นักเรียนในระดับชั้นประถมศึกษา ทดสอบโดยใช้แบบสอบโคลซซึ่งต้องเขียนเติมคำที่ขาดหายไปด้วยตัวเอง ซึ่งเรียกว่าเป็นการทำแบบสอบโคลซนั้นเป็นสิ่งที่ยากและยากลำบากในการตรวจให้คะแนนด้วย ดังนั้นผู้วิจัยจึงดัดแปลงแบบสอบโคลซ (cloze test) เป็นแบบ สอบโมดิฟายด์โคลซ (modified cloze test) ซึ่งมีลักษณะต่างไปจากแบบสอบโคลซเดิม โดยที่แบบสอบนี้มีการให้สิ่งชี้แนะในลักษณะต่าง ๆ กันไว้แทนที่ช่องว่างของคำที่ถูกตัดทิ้งไป ซึ่งรูปแบบของแบบสอบ โดยทั่ว ๆ ไปจะมี 4 รูปแบบด้วยกันคือ แบบสอบโมดิฟายด์โคลซ ที่มีตัวอักษรชี้แนะ แบบสอบโมดิฟายด์โคลซที่มีช่องว่างเท่าจำนวนตัวอักษรที่ต้องเติม (e——) แบบสอบโมดิฟายด์โคลซแบบจับคู่ และแบบสอบโมดิฟายด์โคลซแบบเลือกตอบ แบบสอบโมดิฟายด์โคลซเป็นแบบสอบที่เหมาะสมกับการนำมาใช้วัดความเข้าใจในการอ่านภาษาอังกฤษของนักเรียนระดับประถมศึกษาตอนปลาย และระดับมัธยมศึกษาตอนต้น เนื่องจากนักเรียนทั้งสองระดับนี้มีพื้นฐานในการเรียนภาษาอังกฤษยังไม่มากพอ และยังด้อยประสบการณ์ในการใช้ภาษา แบบสอบที่จะนำมาใช้จึงควรมีลักษณะยั่วยุและจูงใจให้นักเรียนแสดงความสามารถออกมา นอกจากนี้แบบสอบโมดิฟายด์โคลซยังสามารถนำมาใช้ในการเป็นแบบฝึกหัด เพื่อฝึกฝนการอ่านภาษาอังกฤษให้กับนักเรียนได้อีกด้วย (อมรรัตน์ แกมทอง, 2530)ในการนำแบบสอบโมดิฟายด์โคลซ (modified cloze test) มาวัดความเข้าใจในการอ่านวิชาภาษาอังกฤษของนักเรียนระดับชั้นประถมศึกษาปีที่ 6 นั้น ผู้วิจัยยึดหลักการการสร้างเหมือนการสร้างแบบสอบโคลซแต่ดัดแปลงในส่วนที่นำคำที่ขาดหายไป

สร้างเป็นตัวเลือก เพื่อให้ให้นักเรียนเลือกแทนที่จะให้นักเรียนเขียนเติมเอง และจำนวนคำจะมีน้อย เป็นเรื่องหรือข้อความสั้นๆ อันเนื่องมาจากนักเรียนในระดับนี้มีพื้นฐานความรู้ด้านคำศัพท์และ โครงสร้างประโยคจำกัด อีกทั้งมีช่วงความสนใจในการอ่านที่สั้น เรื่องที่นำมาสร้างเป็นแบบสอบ modified cloze test จึงเป็นเรื่องสั้นๆเหมาะสมกับวัยและความสนใจของนักเรียนระดับชั้นประถมศึกษาปีที่ 6

ตอนที่ 2 แนวคิดเกี่ยวกับการเลือกเรื่อง จำนวนคำที่เหมาะสม และการตัดคำในแบบ สอบโคลซ

พิกุลสกี และ โทบิน (Pikulsiki and Tobin, 1982 อ้างถึงใน อมรรัตน์ แกมทอง, 2530) เสนอ แนวการเลือกเรื่องหรือข้อความ และจำนวนคำที่จะนำมาสร้างเป็นแบบสอบโคลซ ไว้ดังนี้คือ

1. เลือกเรื่องหรือข้อความที่มีความยาว 200-300 คำ เรื่องหรือข้อความที่เลือกมานั้นควรมี ลักษณะเป็นตัวแทนที่ดีของหนังสือที่เลือกมา ชนิดของเรื่องหรือข้อความที่เลือกมานั้นขึ้นอยู่กับ วัตถุประสงค์ของการทดสอบ และคำนึงถึงเป้าหมายในการเรียนของนักเรียนและของรายวิชาด้วยการเลือกเรื่องหรือข้อความเพื่อสร้างแบบทดสอบโคลซ ครูต้องระวังไม่ให้นักเรียนมีความรู้สึกว่าจะ ต้องใช้ความรู้เกินความสามารถที่มีอยู่จึงจะเข้าใจข้อความนั้นได้

2. ถ้าข้อความนั้นมีคำจำพวก anaphoric word (เช่น it this) เป็นจำนวนมากก็ควรเลือก ข้อความอื่น

(อ้อยทิพย์ กรมกุล, 2524 อ้างถึงในนิคม ทองบุรี, 2535) กล่าวถึงวิธีการเลือกเรื่องหรือ ข้อความที่จะนำมาสร้างเป็นแบบสอบโคลซ

เลือกเรื่องหรือข้อความที่มีความยาว 250-375 คำ ชนิดและลักษณะของบทความที่เลือกขึ้น อยู่กับวัตถุประสงค์ของการทดสอบ และเนื้อหาวิชาว่าจะทดสอบอะไร เรื่องหรือข้อความต้องมีความที่สมบูรณ์ชัดเจนในตัวเองมีความยาวพอที่จะใช้ตัดคำได้ตามวัตถุประสงค์ของผู้ทดสอบว่าจะ วัดอะไรและต้องการจำนวนคำที่ถูกตัดออกเท่าใด แต่ไม่ควรเกิน 50 คำ เรื่องหรือข้อความที่เลือก ควรจะเป็นเนื้อหาเป็นกลาง ไม่เน้นเนื้อหาเฉพาะเจาะจงในสาขาวิชาใดเกินไป เนื่องจากผู้สอบจะ ไม่ได้รับความยุติธรรม เพราะพื้นฐานและความรู้ของแต่ละคนในสาขาวิชาต่างๆไม่เท่ากัน การใช้ cloze test ที่เน้นเนื้อหาเฉพาะสาขาวิชาจะได้ก็ต่อเมื่อใช้ทดสอบผู้ทำที่มีประสบการณ์ในสาขาวิชา นั้นๆเท่านั้น ในการเลือกเรื่องหรือข้อความไม่ควรเลือกในเรื่องที่คลุมเครือ หรือดึงความสนใจของผู้

สอบออกจากเรื่องที่จะทดสอบ หรือเรื่องที่มีข้อขัดแย้ง หรือเนื้อหาที่ผู้ทำอาจจะไม่เห็นด้วยเป็น อย่างมาก เช่น ด้านการเมือง

การตัดคำในแบบสอบโคลซ พิกุลสกี และ โทบิน (Pikulski and Tobin, 1982 อ้างถึงใน อมรรัตน์ แกมทอง, 2530)) เสนอแนะไว้ดังนี้

1. ประโยคแรกและประโยคสุดท้ายของเรื่องหรือข้อความจะคงไว้ไม่ต้องตัดคำ
2. ตัดคำออกจากเรื่องหรือข้อความอย่างเป็นระบบ เช่น ทุกคำที่ 5 โดยเริ่มจากประโยค ที่สองของเรื่องหรือข้อความไปจนกระทั่งได้ 50 ช่อง จำนวนเลข เช่น 2500 ถือว่าเป็นเพียงคำเดียว ส่วนคำที่มียัติภังค์ (hyphen) มาเชื่อมนั้นจะต้องเป็นสองคำ ยกเว้นคำที่ไม่สามารถมีความหมายอยู่ ตามลำพังได้ เช่น Co-operation เป็นต้น

(อ้อยทิพย์ กรมกุล, 2524 อ้างถึงใน นิคม ทองบุรี, 2535) กล่าวถึงวิธีการตัดคำในการสร้าง แบบสอบโคลซ ไว้ดังนี้

1. เมื่อเลือกเรื่องหรือข้อความได้แล้ว ให้นับจำนวนคำทั้งหมดแล้วหารด้วยจำนวนคำที่ คิดว่าจะตัดในเรื่องหรือข้อความ เช่น จะตัดทั้งหมด 40 คำ จากเรื่องหรือข้อความที่มีคำทั้งหมด 320 คำ ได้ผลลัพธ์จากการหารเท่ากับ 8 ดังนั้นให้ตัดคำออกทุกๆคำที่ 8 หรือถ้าเรื่องหรือข้อความที่ ยาวประมาณ 280 คำ ก็ตัดทุกคำที่ 7 แล้วแทนที่คำที่ถูกตัดด้วยเส้นตรงที่มีความยาวเท่ากัน (Uniform length) โดยทั่วไปแล้วประโยคแรก หรือ 2-3 ประโยคแรกและประโยคสุดท้ายของเรื่องหรือ ข้อความจะไม่มี การตัดคำ เพื่อให้ผู้อ่านมีโอกาสที่จะคุ้นเคยกับเรื่องที่จะดำเนินต่อไป

เนื่องมาจากนักเรียนในระดับนี้มีพื้นฐานความรู้ด้านคำศัพท์และโครงสร้างประโยคจำกัดอีก ทั้งมีช่วงความสนใจในการอ่านที่สั้น เรื่องหรือข้อความที่นำมาสร้างเป็นแบบสอบ modified cloze test จึงเป็นข้อความสั้นๆเหมาะสมกับวัยและความสนใจของนักเรียนระดับชั้นประถมศึกษาปีที่ 6 โดยเรื่องที่เลือกมาให้อ่านนั้นเป็นเรื่องที่อยู่ในความสนใจของนักเรียนในระดับชั้นประถมศึกษาปีที่ 6 ทั้ง 3 เรื่อง ได้แก่ นิทาน (A farmer and a snake) เรื่องจากประสบการณ์ทั่วไป (Something for my best friend) และ ข้าว (Tata wonderful pockets) ซึ่งแต่ละเรื่องมีจำนวนคำ 80-85 คำ และเว้น ทุกคำที่ 8

ตอนที่ 3 แนวคิดเกี่ยวกับการอ่านและความเข้าใจในการอ่าน

ความหมายของการอ่าน

การอ่านเป็นเครื่องมือในการแสวงหาความรู้และเป็นทักษะที่มีประโยชน์อย่างยิ่งทั้งในด้านการเรียนและการดำรงชีวิตประจำวันของเรา ดังนั้นนักการศึกษาหลายท่านจึงได้ให้ความสนใจในการอ่านและได้ให้ความหมายของการอ่านไว้ต่าง ๆ กันดังนี้

การอ่านเป็นกระบวนการในการแปลความหมายของตัวอักษร หรือสัญลักษณ์ที่มีการจัดบันทึกไว้ออกมาเป็นคำพูด หรือความคิดอย่างมีเหตุผล (จรรยา เรื่องมาลัย, 2525)

การอ่าน คือ การรับรู้ความหมายจากถ้อยคำที่ตีพิมพ์อยู่ในสิ่งพิมพ์หรือหนังสือ โดยผู้อ่านรับรู้ว่าคุณเขียนได้สื่ออะไรมายังผู้อ่านทั้งในด้านความคิด ความรู้ ความหมาย ความสัมพันธ์กับสิ่งอื่นรับรู้ว่าคุณเขียนมีความตั้งใจจะแสดงความคิดอะไร (เขาวลี นาคสุขศรี, 2534)

จากความหมายของการอ่านที่มีผู้ให้ไว้หลายท่านได้ให้ไว้จึงพอสรุปได้ว่า การอ่านเป็นการแปลความหมายของตัวอักษรออกมาเป็นความคิด เข้าใจความหมายของตัวอักษร เข้าใจความต้องการของผู้เขียน และนำความรู้หรือความคิดที่ได้ไปใช้ให้เป็นประโยชน์

ความเข้าใจในการอ่าน

จุดประสงค์ที่สำคัญของการอ่านก็คือความเข้าใจเพราะการอ่านโดยทั่วไปมีเป้าหมายเพื่อให้เกิดความเข้าใจในสิ่งที่อ่าน เพราะถ้าผู้อ่านสามารถรวบรวมความคิดจากสิ่งที่อ่าน นำไปใช้ในชีวิตประจำวันได้ การอ่านที่ปราศจากความเข้าใจจึงไม่ใช่การอ่านที่แท้จริง

ความเข้าใจในการอ่าน หมายถึง ความสามารถในการทำความเข้าใจข้อมูลซึ่งเป็นภาษาเขียนได้ตรงกับความตั้งใจของผู้เขียน ซึ่งความเข้าใจในการอ่านที่วัดได้จากแบบสอบที่ผู้วิจัยสร้างขึ้น (ศิริพร ฉันทานนท์, 2527)

ความเข้าใจในการอ่านเป็นจุดหมายหลักของการอ่านทั้งปวง การอ่านโดยจับใจความอะไรไม่ได้ไม่ควรนับว่าเป็นการอ่านเพราะการอ่านในลักษณะดังกล่าว ไม่ก่อให้เกิดประโยชน์ใดๆแก่ผู้อ่าน การอ่านเพื่อความเข้าใจจึงเป็นการอ่านที่สำคัญที่สุด

ความเข้าใจในการอ่าน (Comprehension Reading) คือความสามารถที่จะอนุมานข้อสนเทศหรือความหมายอันพึงประสงค์จากสิ่งที่อ่านมาแล้วได้ อย่างมีประสิทธิภาพมากที่สุดเท่าที่ทำได้ (สมุทร เชนเชาวนิช, 2537)

การอ่านเป็นกระบวนการที่ซับซ้อนและต้องใช้ความสามารถภายในตัวบุคคลหลายอย่าง เพราะการอ่านเป็นกระบวนการของความคิด ดังที่ ธอร์นไดค์ (Thorndike, 1971 อ้างถึงใน ศิริพร ฉันทานนท์, 2527) ได้กล่าวว่า การอ่าน หมายถึงการคิดและรวมถึงการเข้าใจความหมายของสัญลักษณ์ที่มองเห็นและเข้าใจความคิดของผู้เขียนด้วย การอ่านต้องใช้กลไก (Mechanics) อยู่

ตลอดเวลา ผู้วิจัยครั้งหลังๆ นี้จึงได้กล่าวว่า “การอ่านเป็นกระบวนการของความคิดซึ่งรวมถึงการเข้าใจตัวอักษรที่ปรากฏในแต่ละบรรทัด การเชื่อมโยงความคิด การใช้วิจารณญาณ และการประเมินเรื่องราวที่อ่านอย่างพิเคราะห์”

เนื่องจากแบบสอบ modified cloze test เป็นแบบสอบที่ผู้วิจัยสร้างขึ้นโดยวิธีตัดคำออกจากเรื่องหรือข้อความโดยตัดทุกคำที่ 8 จากการศึกษาผลงานวิจัยในเรื่องผลของการละหรือตัดคำในตำแหน่งที่ต่างกันแบบสอบโคลซที่มีผลต่อคะแนนสอบของนักเรียนในระดับประถมศึกษาปรากฏว่าการละหรือตัดคำในตำแหน่งที่ต่างกันแบบสอบโคลซมีผลต่อคะแนนสอบ การละหรือตัดคำในตำแหน่งที่ 5 7 8 เป็นตำแหน่งที่เหมาะสม (นิคม ทองบุรี, 2535) ผู้วิจัยจึงเลือกตัดคำที่ 8 เพราะเป็นตำแหน่งที่เหมาะสมและไม่ยุ่งยากต่อการตัดคำสามารถตัดคำได้ลงตัวและไม่มีผลกระทบต่อข้อความในเรื่อง ในการทำแบบสอบผู้สอบจะต้องอ่านเรื่องหรือข้อความแล้วคิดหาคำที่ถูกตัดออกไปเพื่อจะได้เลือกคำที่ให้มา (ซึ่งเป็นคำที่ถูกตัดออก) ได้อย่างถูกต้อง

การที่จะเข้าใจเรื่องหรือข้อความที่อ่านได้ตลอด จนสามารถคิดหาคำที่ถูกตัดออกไปได้นั้น จากการวิจัยของในเรื่ององค์ประกอบของความเข้าใจในการอ่าน พบว่า ขึ้นอยู่กับพื้นฐานสำคัญ 3 ประการ คือ คำศัพท์ (Eskey, 1975) โครงสร้างทางไวยากรณ์ (Williams, 1986) และ ความหมายทางวัฒนธรรม (Littlewood, 1983)

องค์ประกอบของความเข้าใจในการอ่าน

1. องค์ประกอบทางด้านคำศัพท์ จากผลการวิจัยพบว่า ความหมายของคำศัพท์เป็นรากฐานที่สำคัญของความเข้าใจในการอ่าน เมื่อนักเรียนเข้าใจความหมายของคำศัพท์ชัดเจนและกว้างขวางพอแล้วย่อมทำให้นักเรียนเกิดความคิดจะนำมาใช้ในการอ่านทำให้เข้าใจประโยคข้อความที่อ่านได้ดียิ่งขึ้น คำศัพท์ถือเป็นกุญแจสำคัญในการอ่านเอาความ จะมองข้ามไม่ได้ (Eskey, 1975) ผู้เรียนที่มีปัญหาทางด้านคำศัพท์จะมีปัญหาในการอ่านและการทำความเข้าใจข้อความที่อ่านดังนั้นคำศัพท์ที่นำมาใช้ในการแต่งข้อความที่จะนำมาให้นักเรียนอ่านนั้นจะต้องมีความง่ายเหมาะสมกับวัยของนักเรียนซึ่งในการวิจัยครั้งนี้ผู้วิจัยได้นำคำศัพท์จากประมวลคำศัพท์ของนักเรียนชั้นประถมศึกษาปีที่ 6 มาแต่งเป็นเรื่องหรือข้อความให้นักเรียนอ่าน

2. องค์ประกอบทางด้านโครงสร้างทางไวยากรณ์ เป็นองค์ประกอบสำคัญอีกประการหนึ่งที่มีความสัมพันธ์ต่อการอ่านและความเข้าใจในการอ่านนักเรียนที่มีปัญหาทางด้านไวยากรณ์ก็จะมีปัญหาด้านการอ่านและความเข้าใจในการอ่านเช่นกัน (Williams, 1986) โครงสร้างทางไวยากรณ์ที่คำนึงในการนำมาใช้ในการอ่านและทำความเข้าใจในการอ่านนั้นก็คือความถูกต้อง

ด้านไวยากรณ์ คือวางคำถูกต้องลงในตำแหน่งที่ถูกต้องโดยอาศัยหน้าที่ของคำ(parts of speech) เป็นหลักคำในภาษาอังกฤษแบ่งได้ดังนี้คือ noun,pronoun,verb,adjective,adverb,preposition, conjunction, interjection ต้องใช้คำเหล่านี้ในรูปที่ถูกต้องและวางในตำแหน่งที่ถูกต้อง (เชาวน์ เหวงเดช, 2534)

3. องค์ประกอบทางด้านสังคมและวัฒนธรรม ความถูกต้องในการใช้เป็นที่ยอมรับเชิง บริบทสังคมและวัฒนธรรม เป็นการพิจารณาความหมายโดยรวมของคำที่นำมาเติมว่ากลมกลืน สอดคล้องกับเนื้อหาทั้งเรื่องหรือไม่และมองไปถึงการที่เป็นที่ยอมรับหรือนิยมกระทำกันในสังคม หรือเป็นการขัดแย้งต่อวัฒนธรรม บรรทัดฐานของสังคมหรือไม่ โดยคำนึงถึงการใช้ภาษาที่ถูกต้อง เป็นที่ยอมรับของเจ้าของภาษา (Littlewood, 1983; วรรณภา วิมลเจลา, 2538)

กล่าวโดยสรุป ผู้อ่านที่มีความสามารถเข้าใจสิ่งที่ตนอ่าน การที่จะเข้าใจข้อความที่อ่านได้ ตลอด จนสามารถคิดทนายคำที่ถูกตัดออกไปได้นั้น จากการวิจัยในเรื่ององค์ประกอบของความเข้าใจในการอ่าน พบว่า ขึ้นอยู่กับพื้นฐานสำคัญ 3 ประการ คือ คำศัพท์ โครงสร้างทางไวยากรณ์ และ ความหมายทางวัฒนธรรม

ตอนที่ 4 แนวคิดเกี่ยวกับสมรรถภาพทางด้านทักษะการอ่านวิชาภาษาอังกฤษของ นักเรียนระดับชั้นประถมศึกษา

ในการวิจัยครั้งนี้เป็นการวิจัยที่เน้นการสร้างแบบสอบถามที่เกี่ยวข้องกับทักษะทางการอ่าน ผู้วิจัยจึงนำคะแนนที่นักเรียนสอบได้จากแบบสอบถามผลสัมฤทธิ์ที่ครูสร้างขึ้นในส่วนที่เกี่ยวข้องเฉพาะ คะแนนด้านการอ่านมาเป็นคะแนนเกณฑ์ในการหาความตรงตามเกณฑ์สัมพัทธ์ซึ่งแบบสอบถามวัดผลสัมฤทธิ์ที่ครูสร้างขึ้นที่เกี่ยวข้องกับการอ่านมีรายละเอียดดังนี้ (กระทรวงศึกษาธิการ,2537)

สมรรถภาพพื้นฐานตามหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) ของนักเรียนชั้นประถมศึกษาปีที่ 6 ซึ่งเป็นกรอบเนื้อหาของการสร้างแบบสอบถามผลสัมฤทธิ์ปลายปีวิชาภาษาอังกฤษด้านทักษะการอ่าน ดังตารางที่ 1

ตารางที่ 1 สมรรถภาพพื้นฐานสำหรับการประเมินคุณภาพนักเรียนกลุ่มประสบการณ์ พิเศษ (วิชาภาษาอังกฤษ) ด้านทักษะการอ่าน

พฤติกรรม	ขอบเขตการวัด	เนื้อหา
การอ่าน	ความเข้าใจ(จับใจความ แปลความ ดีความ ขยายความ เปรียบเทียบ	เรื่องจากประสบการณ์ทั่วไป นิทาน เพลง การ์ตูน ข่าว ประกาศต่างๆ จดหมาย คำแนะนำ เป็นต้น

ตอนที่ 5 แนวคิดเกี่ยวกับการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบที่ตรวจให้คะแนนแบบพหุวิภาค (Polytomous IRT Model)

การวิเคราะห์ข้อสอบมีจุดประสงค์ที่สำคัญ 2 ประการ คือ การวิเคราะห์เพื่อหาคุณภาพรายข้อกระทง แล้วนำผลที่ได้ไปปรับปรุงข้อสอบ และวิเคราะห์เพื่อหาคุณภาพแบบสอบทั้งฉบับด้านความตรงและความเที่ยง การวิเคราะห์ดังกล่าวแบ่งได้เป็น 2 แนวทาง คือ การวิเคราะห์ตามทฤษฎีการทดสอบดั้งเดิม (Classical Test Theory : CTT) วิธีนี้มีข้อจำกัดที่ค่าสถิติที่ได้จะแปรเปลี่ยนไปตามกลุ่มตัวอย่าง อีกแนวทางหนึ่งคือการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) การวิเคราะห์ตามแนวทางหลังนี้ให้ค่าพารามิเตอร์ที่สามารถอ้างอิงได้กับกลุ่มตัวอย่างทั่วไปจึงมีการนำมาใช้กันมากในปัจจุบัน IRT Model ดังกล่าว จำแนกได้เป็น 2 ประเภท ตามลักษณะการให้คะแนน คือ (1) Dichotomous IRT Model ซึ่งใช้กับข้อสอบที่ตรวจให้คะแนนแบบทวิวิภาค และ (2) Polytomous IRT Model ซึ่งใช้กับข้อสอบที่ตรวจให้คะแนนแบบพหุวิภาค การวิเคราะห์ตามโมเดลแรกมีการนำมาใช้และรู้จักกันดี ส่วนการวิเคราะห์ตามแนวทางที่สองนั้น มีการพัฒนามากว่า 25 ปี (ธนวัฒน์ แสนสุข, 2538) โดยกำลังเป็นที่รู้จักในการนำมาใช้ในวงการวัดผลทางการศึกษา โมเดลสำหรับการวิเคราะห์ตามแนวทางดังกล่าวมีด้วยกันหลายโมเดล เช่น Graded Response Model (GRM), Rating Scale Model (RSM) และ Generalized Partial Credit Model (GPCM) เป็นต้น โมเดลที่รู้จักกันดีในปัจจุบัน คือ GRM, RSM, GPCM เป็นต้น

ในขณะที่มีการพัฒนาโมเดลก็มีการพัฒนาโปรแกรมคอมพิวเตอร์ เพื่อใช้วิเคราะห์ตามโปรแกรมดังกล่าวมากขึ้น ตัวอย่างโปรแกรมคอมพิวเตอร์ที่มีการพัฒนาขึ้นและเป็นที่ยอมรับกันดี ได้แก่ MULTILOG, BIGSTEPS, และ PARSCALE เป็นต้น

โมเดลที่มีผู้ศึกษากันมากในปัจจุบัน ไม่เข้มงวดเกี่ยวกับข้อตกลงเบื้องต้น และสามารถใช้ได้กับแบบสอบและแบบวัดหลายลักษณะ คือ GRM และ GPCM (Donoghue, 1994; Deayaia, 1994; Muraki, 1992) ในที่นี้จะกล่าวถึงเฉพาะโมเดล GRM ซึ่งเป็นโมเดลที่ใช้ในการวิจัยครั้งนี้

ซาเมจิมา (Samejima, 1976) ได้พัฒนา GRM จากโมเดลโลจิสติก 2 พารามิเตอร์ เพื่อใช้กับแบบสอบและแบบวัดที่ตรวจให้คะแนนแบบพหุวิภาค (Polytomous) รูปแบบทั่วไปของโมเดลนี้พัฒนามาจากสูตร

$$P_i(u_j = 1/\theta) = P_i(u_j = 1/\theta) = \frac{\exp^{[a_j(\theta - b_j)]}}{1 + \exp^{[a_j(\theta - b_j)]}}$$

เมื่อ $P_i(u_j = 1/\theta)$ คือ ความน่าจะเป็นของคนที่มีความสามารถ θ สามารถตอบข้อสอบข้อ j ($U=1$) ได้ถูกต้อง

a_j คือ ค่าพารามิเตอร์อำนาจจำแนก

b_j คือ ค่าพารามิเตอร์ของความยากของข้อสอบ

exp คือ ค่าคงที่ 2.7183

ถ้ามีข้อตกลงว่าข้อสอบทุกข้อมีอำนาจจำแนกแตกต่างกัน โมเดลนี้ก็คือโมเดลโลจิสติกแบบ 2 พารามิเตอร์

GRM มีข้อตกลงว่าการตอบสนองข้อกระทง j ของผู้ตอบสามารถแบ่งคะแนนออกเป็น m_j+1 ลำดับชั้น (categories) คะแนนแต่ละลำดับ (j_k) กำหนดให้เป็น 0, 1, 2, ..., m_j ตามลำดับ (Donoghue, 1994; Koch, 1983) เขียนสมการทั่วไปได้ดังนี้

$$P_{jk}(\theta) = \frac{\exp^{[Da_j(\theta - b_{jk})]}}{1 + \exp^{[Da_j(\theta - b_{jk})]}}$$

เมื่อ $P_{jk}(\theta)$ คือ ความน่าจะเป็นที่ผู้มีคุณลักษณะ θ จะตอบข้อกระทง j ได้คะแนน k

D คือ ค่าคงที่ของสเกล (scaling constant) เมื่อปรับโด่งฟังก์ชันของโมเดลโลจิสติก และโมเดล นอร์มอล ออร์โจีพี มีค่าเท่ากับ 1.70

a_j คือ ค่าอำนาจจำแนกของข้อความ j

θ คือ ความสามารถหรือคุณลักษณะภายในของผู้ตอบ

b_{jk} คือ ค่าความยากของข้อความ j ในลำดับชั้นที่ jk เมื่อ $jk = 0, 1, \dots, k$

exp คือ ค่าคงที่ มีค่าเท่ากับ 2.7183

ในโมเดลนี้แสดงถึงความสัมพันธ์ของคนที่มีคุณลักษณะภายใน (θ) สูง มีความน่าจะเป็นในการตอบลำดับชั้นคะแนนที่สูงกว่าคนที่มีคุณลักษณะภายในต่ำ

b_{jk} คือค่าความยากในแต่ละลำดับชั้นคะแนน เช่น มีการตรวจให้คะแนนเป็น 1, 2, และ 3 ก็จะมีค่าความยากตามลำดับชั้นคะแนนเป็น 2 ลำดับชั้น จากคะแนน 0 เป็น คะแนน 1 ไม่เป็นลำดับชั้นคะแนนหรือถือว่าเป็น 0 จากลำดับชั้นคะแนน 1 เป็นคะแนน 2 เป็นลำดับชั้นคะแนนที่ 1 จากคะแนน 2 เป็นคะแนน 3 เป็นลำดับชั้นคะแนนที่ 2

ลักษณะเฉพาะฟังก์ชันของโมเดล GRM สรุปได้ดังนี้

1. ฟังก์ชันสารสนเทศของแบบวัด (TIF) เขียนแทนได้ด้วยสมการดังนี้

$$I(\theta) = \sum_{j=1}^m I_j(\theta)$$

เมื่อ $I(\theta)$ คือ ฟังก์ชันสารสนเทศของข้อกระทงตั้งแต่ข้อที่ 1 ถึงข้อที่ m

2. ฟังก์ชันสารสนเทศของข้อกระทง (IIF) สามารถแทนด้วยสมการดังนี้

$$I_j(\theta) = \sum_{jk=0}^k \frac{[P'_{jk}(\theta)]^2}{P_{jk}(\theta)}$$

ข้อตกลงเบื้องต้นของโมเดลนี้คือ ค่าความยากตามลำดับขั้นคะแนนต้องเรียงจากค่าน้อยไปหาค่ามาก เช่นถ้ามีคะแนนในการตอบเป็น 1, 2 และ 3 ค่าความยากตามลำดับขั้นคะแนนจะมี 2 ลำดับขั้น คือ ค่าความยากในลำดับขั้นที่ 1 (b1) และค่าความยากในลำดับขั้นที่ 2 (b2) ดังนั้น ค่า b1 ต้องน้อยกว่าค่า b2

ตอนที่ 6 งานวิจัยที่เกี่ยวข้องกับการตรวจให้คะแนนแบบต่างๆ

ในการวัดผลการศึกษา นั้น วิธีที่นิยมกันทั่วไปในการให้คะแนน ได้แก่ การให้คะแนนข้อที่ถูกได้ 1 คะแนน และข้อที่ผิดได้ 0 คะแนน แต่เพื่อต้องการให้ค่าความตรงและค่าความเที่ยงสูงขึ้น จึงมีผู้พยายามคิดค้นหาวิธีที่จะให้คะแนนอีกหลายวิธี และมีงานวิจัยอีกเป็นจำนวนมากที่แสดงผลของความพยายามนี้มีทั้งประสบความสำเร็จและล้มเหลวซึ่งจะกล่าวต่อไปนี้

1. การให้คะแนนตัวเลือกแตกต่างกัน

การให้คะแนนตัวเลือกมีหลายวิธี แต่ละวิธีมีผลต่อความตรงและความเที่ยงของแบบสอบแตกต่างกัน คือ

1.1 วิธีอาศัยความเห็นของผู้เชี่ยวชาญ

การให้คะแนนตัวเลือกของข้อสอบโดยผู้เชี่ยวชาญมี 2 วิธี คือ 1) อาศัยความเห็นของผู้เชี่ยวชาญว่าตัวเลือกแต่ละตัวมีความถูกต้องมากน้อยเพียงใด และ 2) ใช้ค่าเฉลี่ยของลำดับที่ของความถูกต้องของตัวเลือกแต่ละตัวที่ตัดสินโดยผู้เชี่ยวชาญจากการศึกษาของ Nedelsky ในปี 1954 (อ้างจาก Wang and Stanley, 1970:671) ที่ให้นักศึกษาที่สอบผ่านการสอบรบยอดแล้ว 306 คน เป็นกลุ่มตัวอย่าง โดยกำหนดค่าน้ำหนักคะแนนโดยวิธีที่ 1 ปรากฏว่า แบบสอบแบบเลือกตอบมีค่าความเที่ยงสูงกว่าค่าความเที่ยงที่คำนวณจากคะแนนวิธีประเพณีนิยม (0,1) และในปี 1964 Jacobs และ Vanderverter (อ้างจาก Wang and Stanley, 1970:693) ได้ใช้วิธีการให้คะแนนวิธีเดียวกัน ปรากฏว่า การให้คะแนนดังกล่าวทำให้แบบสอบแบบเลือกตอบมีค่าความเที่ยงจากการสอบซ้ำ ค่าความตรงเชิงพยากรณ์ และค่าความตรงร่วมสมัย (Concurrent Validity) มีค่าสูงเกินกว่าค่าความเที่ยงและความตรงชนิดเดียวกันที่คำนวณจากคะแนนที่ให้โดยวิธีประเพณีนิยม

ส่วนการให้คะแนนวิธีที่ 2 นั้น Hambleton, Roberts และ Traub ได้ทำการศึกษาในปี 1970 (อ้างจาก Wang and Stanley, 1970:693) ปรากฏว่าการให้คะแนนวิธีดังกล่าวทำให้แบบสอบแบบเลือกตอบมีค่าความตรงเชิงพยากรณ์และค่าความเที่ยงสูงเกินกว่าค่าความตรงและค่าความเที่ยงที่คำนวณจากคะแนนที่ให้โดยวิธีประเพณีนิยมแต่ไม่มีนัยสำคัญทางสถิติ

2. วิธีใช้ค่าความสัมพันธ์ระหว่างตัวเลือกกับสิ่งที่มุ่งทดสอบ

การให้คะแนนโดยอาศัยแนวคิดเกี่ยวกับความสัมพันธ์ระหว่างตัวเลือกกับสิ่งที่มุ่งทดสอบ เท่าที่พบมี 3 วิธีคือ 1) ใช้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวเลือกกับคะแนนรวม 2) ใช้ค่ามาตรวัดที่คำนวณจากวิธีของ Guttman's Scale Analysis และ 3) ใช้มาตรวัดที่คำนวณจากวิธีการเปรียบเทียบรายคู่ของ Thurstone (Thurstone's Paired Comparison) จากการศึกษาวรรณคดีที่เกี่ยวข้องพบว่า Davis และ Fifer ได้ใช้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวเลือกแต่ละตัวกับคะแนนรวม เป็นคะแนนแบบสอบคู่ขนานวิชาคณิตศาสตร์ ซึ่งยาว 45 ข้อ ในปี 1959 (อ้างจาก Wang and Stanley, 1970:692) ผลปรากฏว่า การให้คะแนนวิธีดังกล่าวทำให้แบบสอบมีค่าความเที่ยงสูงมากกว่าค่าความเที่ยงที่คำนวณจากคะแนนที่ให้โดยวิธีประเพณีนิยมอย่างมีนัยสำคัญ คือเพิ่มจาก 0.68 เป็น 0.76 ซึ่งถ้าทำให้แบบสอบนั้นค่าความเที่ยงเท่ากับค่าที่เพิ่มขึ้น เมื่อให้คะแนนวิธีประเพณีนิยมจะต้องเพิ่มข้อสอบอีก 45 - 67 ข้อ

ส่วนการให้คะแนนโดยใช้ค่ามาตรวัดจากวิธี Guttman's Scale Analysis นั้น Hendrickson ได้ทำการศึกษาเกี่ยวกับกลุ่มตัวอย่างที่เป็นเพศชาย 5,000 คน และเพศหญิง 5,000 คน ที่สอบ Scholastic Aptitude Test ในปี 1970 และ Raffeld ได้ทำการศึกษาเกี่ยวกับกลุ่มตัวอย่างที่เป็นนักเรียน Grade 9 จำนวน 1,354 คน ที่สอบวิชาการอ่านเข้าใจความ ศัพท์ ไวยากรณ์ การสะกดคำ และคณิตศาสตร์ (อ้างจาก Wang and Stanley, 1970:693; Raffeld, 1975:285) จากการศึกษา Hendrickson พบว่าความเที่ยงของแบบสอบสูงเกินกว่าค่าของความเที่ยงที่คำนวณจากคะแนนที่ให้โดยวิธีประเพณีนิยมอย่างมีนัยสำคัญ ซึ่งถ้าให้คะแนนโดยวิธีประเพณีนิยม และต้องการให้แบบสอบนั้นมีค่าความเที่ยงสูงเท่ากับค่าความเที่ยงที่เพิ่มขึ้น จะต้องเพิ่มข้อสอบอีกประมาณ 49 % ในทำนองเดียวกัน Raffeld พบว่า ความเที่ยงและความตรงเชิงพยากรณ์ของแบบสอบที่ให้คะแนนโดยใช้มาตรวัดวิธี Guttman's Scale Analysis มีค่าสูงกว่าค่าความเที่ยงและความตรงเชิงพยากรณ์ที่คำนวณจากคะแนนที่ให้โดยวิธีประเพณีนิยมอย่างมีนัยสำคัญ เมื่อใช้คะแนนปลายปีเป็นเกณฑ์สำหรับการพยากรณ์

แต่อย่างไรก็ตาม เมื่อ Patnaik และ Traub (Patnaik and Traub, 1973:281) ทำการศึกษาผลของการให้คะแนนต่อค่าของความเที่ยงและความตรงเชิงพยากรณ์ในปี 1973 กับนักศึกษาระดับ

บัณฑิตศึกษา จำนวน 61 คน โดยใช้แบบสอบวัดเชาวน์ปัญญาที่เป็นแบบสอบแบบเลือกตอบ ชนิดมี 5 ตัวเลือก และใช้มาตรวัดจากวิธีการเปรียบเทียบรายคู่ของ Thurstone เป็นคะแนนของตัวเลือกแต่ละตัว ผลปรากฏว่า การให้คะแนนดังกล่าวทำให้แบบสอบมีค่าความเที่ยงแบบแบ่งครึ่งสูงกว่าค่าความเที่ยงที่คำนวณจากคะแนนที่ให้โดยวิธีประเพณีนิยมอย่างมีนัยสำคัญ ($p < .05$) แต่มีค่าความตรงเชิงพยากรณ์ที่ใช้ผลการสอบสัมฤทธิ์ผลวิชาต่างๆ เป็นเกณฑ์ต่ำกว่าค่าความตรงชนิดดังกล่าวอย่างมีนัยสำคัญ ($p < .05$) การที่ค่าความเที่ยงของแบบสอบเพิ่มขึ้น แต่ค่าความตรงเชิงพยากรณ์ลดลง จากค่าความตรงเชิงพยากรณ์ที่คำนวณจากคะแนนที่ให้โดยวิธีประเพณีนิยมนั้น Patnaik และ Traub อธิบายว่า เป็นเพราะการให้คะแนนที่ไม่ใช่ประเพณีนิยมทำให้แบบสอบมีค่าสหสัมพันธ์ระหว่างข้อสอบเพิ่มขึ้น เพราะคะแนนมีการกระจายเพิ่มขึ้นมากกว่าเดิมด้วยเหตุนี้จึงทำให้คะแนนจากแบบสอบมีความสัมพันธ์กับเกณฑ์ต่ำเช่นเดียวกับกรณีค่าสัมประสิทธิ์สหสัมพันธ์พหุคูณของตัวแปรที่มีความสัมพันธ์ซึ่งกันและกันสูง มีความสัมพันธ์กับเกณฑ์ต่ำ (Patnaik and Traub, 1977:285)

3. วิธีอาศัยความเชื่อมั่นของผู้สอบต่อตัวเลือกของแบบสอบ

วิธีการให้คะแนนโดยอาศัยความเชื่อมั่นของผู้สอบต่อตัวเลือกของแบบสอบ อาจแบ่งออกได้ 3 วิธีคือ 1) วิธีคิดคะแนนตามความมั่นใจในการตอบ 2) วิธีคิดคะแนนโดยการหักค่าการเดา และ 3) วิธีคิดคะแนนโดยการกำจัดตัวเลือก สำหรับการคิดคะแนนวิธีที่ 1 นั้น แต่ละข้อถ้าผู้สอบตอบถูกในการเลือกตอบครั้งแรกจะได้ 3 คะแนน ถ้าเลือก 2 ครั้ง ตอบถูกได้ 2 คะแนน ถ้าเลือก 3 ครั้ง ตอบถูกได้ 1 คะแนน และถ้าเลือก 4 ครั้ง ตอบถูกจะได้ 0 คะแนน (ข้อสอบมี 4 ตัวเลือก) วิธีที่ 2 ให้คะแนนโดยอาศัยสูตรการคำนวณที่หักค่าการเดา และวิธีที่ 3 นั้น ผู้สอบเลือกตอบได้หลายๆตัวเลือกในแต่ละข้อ ถ้าตอบถูกจะได้ 1 คะแนน แต่ถ้าตอบผิด จะได้ 1 คะแนน (ข้อสอบมี 4 ตัวเลือก)

จากการศึกษาของ Diamond (Diamond, 1976:129) ในปี 1975 เกี่ยวกับการให้คะแนนทั้ง 3 วิธี กับกลุ่มตัวอย่าง 84 คน ในระดับบัณฑิตศึกษา ด้วยแบบสอบวิชาวัดผลที่ค่อนข้างง่าย และมีอำนาจจำแนกต่ำ ผลปรากฏว่า แบบสอบที่ให้คะแนนทั้ง 3 วิธีดังกล่าวมีค่าความเที่ยงแบบ Flanagan สูงกว่าค่าความเที่ยงที่คำนวณจากคะแนนที่ให้โดยวิธีประเพณีนิยม ผลการศึกษาครั้งนี้มีความคล้ายคลึงกับผลการศึกษาของ Hanna ในปีเดียวกัน (Hanna, 1975:145) ซึ่งได้ให้คะแนนตัวเลือกของแบบสอบแบบเลือกตอบชนิด 4 ตัวเลือก 2 ฉบับ ซึ่งยาว 10 ข้อ และ 30 ข้อ และใช้กับกลุ่มตัวอย่าง 85 คน โดยวิธีให้คะแนนตามความมั่นใจในการตอบ ผลการศึกษาพบว่า แบบสอบที่ให้คะแนนโดยวิธีดังกล่าว มีค่าความเที่ยงแบบแบ่งครึ่งสูงกว่าค่าความเที่ยงที่คำนวณจากคะแนน

ที่ให้โดยวิธีประเพณีนิยมเล็กน้อยแต่ค่าความตรงเชิงพยากรณ์ซึ่งใช้คะแนนสอบการตีความ และการเขียนรายงานเป็นเกณฑ์กลับมีค่าต่ำกว่าค่าความตรงเชิงพยากรณ์ที่ได้จากคะแนนที่ให้โดยวิธีประเพณีนิยม ดังนั้น Hanna จึงสรุปว่า การให้คะแนนวิธีประเพณีนิยมให้ผลการสอบดีกว่าวิธีการให้คะแนนตามความมั่นใจในการตอบ เพราะการสอบนั้นควรคำนึงถึงความตรงของแบบสอบมากกว่าความเที่ยง

นอกจากนี้ในปี 1975 Hakstian และ Kansup (Hakstian and Kansup 1975:231) ยังทำการศึกษาผลการให้คะแนนโดยวิธีการคะแนนตามความเชื่อมั่นในการตอบ และวิธีการให้คะแนนโดยวิธีการกำจัดตัวเลือก โดยใช้นักเรียนระดับ Grade 9 จำนวน 1,023 คน เป็นกลุ่มตัวอย่าง ผลการศึกษาพบว่า สอดคล้องกับผลการศึกษาของ Diamond และ Hanna กล่าวคือ การให้คะแนนทั้ง 2 วิธี ทำให้แบบสอบมีค่าความเที่ยงสูงขึ้นเล็กน้อย แต่ให้ค่าความตรงเชิงพยากรณ์ต่ำกว่าค่าความตรงที่คำนวณจากคะแนนที่ให้โดยวิธีประเพณีนิยม

เพื่อเป็นการตรวจสอบผลการศึกษาที่ทำไว้ในปี 1975 Hanna ได้ทำการศึกษาเกี่ยวกับผลการให้คะแนน โดยวิธีการให้คะแนนตามความมั่นใจในการตอบอีกในปี 1977 ครั้งนี้ Hanna (Hanna, 1977:643) ใช้แบบสอบแบบเลือกตอบ ยาว 18 ข้อ 2 ฉบับ คือ วิชาวิทยาศาสตร์และคณิตศาสตร์ และใช้นักเรียน grade 5 และ 6 จำนวน 1,391 คน เป็นกลุ่มตัวอย่าง ผลการศึกษาปรากฏว่ามีความสอดคล้องกับผลการศึกษาในปี 1975 คือ แบบสอบมีค่าความเที่ยงสูงกว่าค่าความเที่ยงที่คำนวณจากคะแนนที่ให้โดยวิธีประเพณีนิยมอย่างมีนัยสำคัญ แต่ให้ค่าความตรงไม่แตกต่างจากค่าความตรงที่คำนวณจากคะแนนที่ให้โดยวิธีประเพณีนิยม

ตอนที่ 7 งานวิจัยที่เกี่ยวข้องกับการตรวจให้คะแนนแบบพหุวิภาค (Polytomous Scoring)

ธนวัฒน์ แสนสุข (2538) ได้ศึกษาการใช้ GRM, GPCM และโมเดลโลจิสติกในการเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบวัดที่มีการตรวจให้คะแนนแบบทวิวิภาค (0,1) และแบบพหุวิภาค (1,2,3,4) ข้อมูลจากการตรวจให้คะแนนแบบทวิวิภาควิเคราะห์ตามโมเดลโลจิสติก 1,2 และ 3 พารามิเตอร์ส่วนการตรวจให้คะแนนแบบพหุวิภาควิเคราะห์ตาม GRM และ GPCM การวิจัยครั้งนี้ทำการศึกษาเกี่ยวกับข้อมูล 2 ชุด (1) ข้อมูลผลการวัดคุณลักษณะทั่วไปของนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 6300 คน จากสำนักทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ (2) ข้อมูลผลสัมฤทธิ์ทางคณิตศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 470 คน ที่ผู้วิจัย

เก็บเอง ทำการตรวจสอบความสมบูรณ์ของข้อมูลโดยวิเคราะห์ความเหมาะสมของผู้ตอบและข้อ
 กระทบตาม RSM ด้วยโปรแกรม BIGSTEPS และวิเคราะห์เพื่อหาค่าฟังก์ชันสารสนเทศของแบบ
 สอบโดยใช้โปรแกรม MULTILOG และ PARSCALE จากนั้นตรวจสอบความตรงเชิงโครงสร้างของ
 แบบวัดคุณลักษณะทั่วไปโดยใช้วิธีวิเคราะห์องค์ประกอบเชิงยืนยัน ด้วยโปรแกรม LISREL

ผลการวิจัยพบว่าการตรวจให้คะแนนแบบพหุวิภาคเมื่อวิเคราะห์ตาม GRM ได้ค่าฟังก์ชัน
 สารสนเทศสูงกว่าการตรวจให้คะแนนแบบทวิวิภาค สำหรับการตรวจให้คะแนนแบบพหุวิภาค วิ
 เคราะห์ตาม GPCM และแบบทวิวิภาคยังไม่สามารถสรุปได้ว่าวิธีใดให้ค่าฟังก์ชันสารสนเทศสูงกว่า
 กัน

ปรมินทร์ อริเดช (2539) ได้ศึกษาการใช้ GRM,GPCM และโมเดลโลจิสติกในการเปรียบเทียบ
 ฟังก์ชันสารสนเทศของมาตรฐานค่าที่มีวิธีการตรวจให้คะแนนแบบทวิวิภาคและพหุวิภาค
 โดยการตรวจให้คะแนนแบบทวิวิภาควิเคราะห์ตามโมเดล 1,2 และ 3พารามิเตอร์ ส่วนการตรวจให้
 คะแนนแบบพหุวิภาควิเคราะห์ตาม GRM และGPCM กลุ่มตัวอย่างที่ใช้ในการวิจัย คือนักเรียนชั้น
 ประถมศึกษาปีที่ 6 สังกัดสำนักงานการประถมศึกษาจังหวัดเชียงใหม่ จำนวน 800 คน เครื่องมือที่
 ใช้ในการวิจัยคือมาตรฐานค่าแบบลิเคิร์ท และมาตรฐานค่าแบบตัวเลือกบังคับตอบ ทำ
 การตรวจสอบความตรงเชิงโครงสร้างของมาตรฐานค่าโดยวิธีวิเคราะห์องค์ประกอบเชิงยืนยัน
 ด้วยโปรแกรม LISREL และวิเคราะห์เพื่อหาค่าฟังก์ชันสารสนเทศของมาตรฐานค่าโดยใช้
 โปรแกรม MULTILOG และ PARSCALE ผลการวิจัยพบว่า ในมาตรฐานค่าแบบลิเคิร์ทการ
 ตรวจให้คะแนนแบบทวิวิภาคที่วิเคราะห์ตามโมเดลโลจิสติกให้ค่าฟังก์ชันสารสนเทศสูงกว่าการ
 ตรวจให้คะแนนแบบพหุวิภาคที่วิเคราะห์ตาม GRM และGPCM และในมาตรฐานค่าแบบตัว
 เลือกบังคับตอบการให้คะแนนแบบพหุวิภาคที่วิเคราะห์ตาม GRM ให้ค่าสารสนเทศสูงกว่าการให้
 คะแนนแบบทวิวิภาค และการตรวจให้คะแนนแบบพหุวิภาคที่วิเคราะห์ตาม GPCM ตามลำดับ

ตอนที่ 8 งานวิจัยที่เกี่ยวข้องกับความเข้าใจในการอ่านที่ใช้แบบสอบโคลซ ในระดับชั้น ประถมศึกษา

จรรยา เรืองมาลัย (2525) ได้ศึกษาเพื่อเปรียบเทียบสัมฤทธิ์ผลความเข้าใจในการอ่านภาษา
 อังกฤษของนักเรียนชั้นประถมศึกษาปีที่ 5 ที่เรียนโดยฝึกทำแบบฝึกโคลซและไม่ได้ฝึกทำแบบฝึก
 โคลซ ผู้วิจัยได้ดำเนินการวิจัยกับตัวอย่างประชากร 2 กลุ่ม ซึ่งเลือกมาจากนักเรียนชั้นประถม
 ศึกษาปีที่ 5 ห้อง 2 และ ห้อง 7 โรงเรียนวัดป้อมมิเชียรโชติการาม จำนวน 60 คน โดยแบ่งเป็น 2
 กลุ่ม ตัวอย่างประชากรนี้มีพื้นฐานการเรียนรู้อยู่ในระดับใกล้เคียงกันและไม่มีพื้นฐานความรู้ทาง

ภาษาอังกฤษมาก่อนเลย ผู้วิจัยได้ดำเนินการสอนตัวอย่างประชากรทั้ง 2 กลุ่ม โดยให้กลุ่มทดลองเรียนโดยทำแบบฝึกโคลซ ส่วนกลุ่มควบคุมไม่ได้ทำแบบฝึกโคลซ ใช้เวลาในการสอน 14 คาบ หลังจากสอนจบเนื้อหาที่กำหนดแล้ว ผู้วิจัยได้ทำการทดสอบสัมฤทธิ์ผลด้านความเข้าใจในการอ่านของตัวอย่างประชากรกลุ่มทดลองและกลุ่มควบคุม โดยใช้แบบสอบที่สร้างขึ้น ซึ่งได้นำไปหาอำนาจจำแนกค่าความยากง่ายและความเชื่อถือได้ก่อนนำมาใช้กับตัวอย่างประชากรและนำคะแนนที่ได้มาทดสอบความมีนัยสำคัญ โดยใช้ค่าที่ ผลการวิจัยสรุปได้ว่านักเรียนที่เรียนโดยทำแบบฝึกโคลซ มีสัมฤทธิ์ผลด้านความเข้าใจในการอ่านสูงกว่านักเรียนที่เรียนโดยไม่ได้ทำแบบฝึกโคลซอย่างมีนัยสำคัญที่ระดับ .05

นิคม ทองบุรี (2535) ได้ศึกษาผลของการละคำต่างตำแหน่ง 6 ตำแหน่ง ตั้งแต่ทุกคำที่ 5 ถึง 10 กับระดับการศึกษา 3 ระดับ คือ ประถมศึกษาปีที่ 6 มัธยมศึกษาปีที่ 3 และ 6 ที่มีต่อคะแนนจากแบบสอบโคลซ กลุ่มตัวอย่างจำนวน 444 คน ได้จากการสุ่มตัวอย่างจากประชากร 115,433 คน จัดนักเรียนเข้ากลุ่มทดลอง 6 กลุ่ม ๆ ละ 74 คน แต่ละกลุ่มมีนักเรียนชั้นประถมศึกษาปีที่ 6 มัธยมศึกษาปีที่ 3 และ 6 จำนวน 24 35 15 ตามลำดับ เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบโคลซที่ผู้วิจัยสร้างขึ้นจำนวน 6 ฉบับ ผลการวิจัยพบว่า

1. ค่าเฉลี่ยจากการทดสอบจากแบบทดสอบโคลซของกลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 6 มีค่าสูงกว่าค่าเฉลี่ยของกลุ่มตัวอย่างนักเรียนชั้นประถมศึกษาปีที่ 6 และ ชั้นมัธยมศึกษาปีที่ 3 ในทุกตำแหน่งของการละ

2. ค่าเฉลี่ยจากการทดสอบจากแบบทดสอบโคลซของกลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีค่าสูงกว่าค่าเฉลี่ยของกลุ่มตัวอย่างนักเรียนชั้นประถมศึกษาปีที่ 6 ในทุกตำแหน่งของการละ

3. การละคำ 6 ตำแหน่ง (คือตำแหน่งที่ 5 ถึง 10) กับระดับการศึกษาของนักเรียน(ประถมศึกษาปีที่ 6 และ มัธยมศึกษาปีที่ 3 มัธยมศึกษาปีที่ 6) ต่างร่วมกันส่งผลต่อคะแนนสอบจากแบบทดสอบโคลซอย่างมีนัยสำคัญทางสถิติที่ระดับ .05