



บทที่ ๓

วิธีการดำเนินการวิจัย

เพื่อให้การดำเนินการวิจัยบรรลุวัตถุประสงค์ของการวิจัยดังกล่าวมาแล้ว ผู้วิจัย
จึงได้ดำเนินการวิจัยดังต่อไปนี้

ประชากรและผลวิจัย

ประชากรที่ใช้ในการวิจัยครั้งนี้ ได้แก่ นิสิตชั้นปีที่ ๑ ของจุฬาลงกรณ์มหาวิทยาลัย
ที่ลงทะเบียนเรียนรายวิชาภาษาอังกฤษพื้นฐาน เฉพาะการอ่าน ๒ (Foundation English:
Reading II) ประจำภาคปลายปีการศึกษา ๒๕๒๗ ซึ่งมีจำนวนทั้งสิ้น ๒,๗๔๔ คน จาก
๑๔ คณะ และ ๔๗ ห้องเรียน (นิสิตคณะวิศวกรรมศาสตร์ ไม่ได้เรียนรายวิชาดังกล่าว)

ผลวิจัยที่ใช้ในการวิจัยครั้งนี้ ได้มาจากการสุ่มโดยวิธีการจับฉลากหมายเลขห้อง
เรียนของนิสิต ส่วนที่เหลือจากทดสอบเครื่องเพื่อการวิจัยครั้งนี้ ซึ่งมีจำนวน ๒,๔๗๗ คน
จาก ๑๔ คณะ และ ๔๓ ห้องเรียน โดยมีเป้าหมายว่า ต้องการสุ่มผลวิจัยจากแต่ละคณะ
ประมาณร้อยละ ๗๕ เพื่อให้ได้ผลวิจัยมาเพียงพอที่จะใช้เพื่อการวิเคราะห์ข้อสอบตามรูปแบบ
โลจิสติกส์ใช้พารามิเตอร์ ๓ ตัวได้ เมื่อสุ่มได้หมายเลขแล้ว ก็นำแบบสอบที่เป็นเครื่องมือเพื่อ
การวิจัยครั้งนี้ไปสอบ การสอบเพื่อเก็บข้อมูลสำหรับการวิจัยครั้งนี้ทำติดต่อกันในเวลา ๑
สัปดาห์ คือ ระหว่างวันที่ ๑๗-๒๒ ธันวาคม ๒๕๒๗ ปรากฏว่า ได้ผลวิจัยทั้งสิ้น จำนวน
๑,๔๗๒ คน จาก ๗๕ ห้องเรียน ซึ่งมีรายละเอียดดังนี้

ตารางที่ ๑ จำนวนพลวิจัย

ที่	คณะ	จำนวน นิสิต	จำนวน พลวิจัย	%ของนิสิต จากคณะ	%ของกลุ่ม ตัวอย่าง
๑	ครุศาสตร์	๓๔๑	๒๕๔	๗๕.๖๖	๑๓.๗๘
๒	ทันตแพทย์	๙๙	๖๕	๖๕.๖๕	๓.๔๗
๓	นิติศาสตร์	๑๖๕	๑๑๕	๖๙.๗๐	๖.๑๔
๔	นิเทศศาสตร์	๑๒๓	๖๗	๕๔.๔๗	๓.๕๘
๕	พาณิชยศาสตร์และการบัญชี	๔๙๓	๓๗๘	๗๖.๖๗	๒๐.๑๙
๖	แพทยศาสตร์	๑๙๐	๙๕	๕๐.๐๐	๕.๐๘
๗	เภสัชศาสตร์	๑๘๙	๑๒๗	๖๗.๒๐	๖.๗๘
๘	รัฐศาสตร์	๒๕๗	๑๖๐	๖๒.๒๕	๘.๕๕
๙	วิทยาศาสตร์	๓๔๕	๒๓๑	๖๖.๙๖	๑๒.๓๔
๑๐	เศรษฐศาสตร์	๑๕๖	๘๕	๕๗.๕๓	๔.๔๙
๑๑	ศิลปกรรมศาสตร์	๕๘	๒๕	๔๓.๑๐	๑.๓๔
๑๒	สถาปัตยกรรมศาสตร์	๙๗	๔๓	๔๔.๓๒	๒.๓๐
๑๓	สัตวแพทยศาสตร์	๖๗	๔๕	๖๗.๑๖	๒.๔๐
๑๔	อักษรศาสตร์	๒๒๔	๑๗๙	๗๙.๙๑	๙.๕๖
รวม		๒,๗๙๔	๑,๘๗๒	๖๗.๐๐	๑๐๐.๐๐

จากตารางที่ ๑ นี้แสดงว่า พลวิจัยที่ใช้ในการศึกษาคั้งนี้ ได้มาจากทุกคณะ ซึ่งคิดเป็นร้อยละจากจำนวนนิสิตแต่ละคณะ จะมีระหว่างร้อยละ ๔๓ ถึงร้อยละ ๘๐ ซึ่งเฉลี่ยแล้ว ประมาณร้อยละ ๖๗ จากประชากรทั้งหมด หรือ ๑,๘๗๒ คน ซึ่งเป็นจำนวนที่มากเพียงพอสำหรับการวิจัยต่อไป

เครื่องมือที่ใช้เพื่อการวิจัย

เครื่องมือที่ใช้เพื่อการวิจัยครั้งนี้ ได้แก่ แบบสอบวัดความสามารถทั่วไปในการอ่านเข้าใจความภาษาอังกฤษ จำนวน ๑ ฉบับ แต่ทำให้มีโครงสร้างความรู้ต่างกัน ๒ แบบ คือ ๑) แบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง แต่มีคำศัพท์ที่มีความหมายสำคัญต่อเรื่อง แบบสอบนี้เป็นแบบสอบแบบเลือกตอบ มี ๔ ตัวเลือก ยาว ๕๐ ข้อ และมีเรื่องสำหรับอ่าน ๒ เรื่อง และ ๒) แบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง **แบบสอบชนิดนี้มีลักษณะเหมือนกับแบบสอบชนิดที่ ๑ ทุกประการ ยกเว้นที่มีชื่อเรื่องและภาพประกอบเรื่องเท่านั้น**

เนื่องจากผู้วิจัยต้องการเรื่องเพื่ออ่าน ๒ เรื่อง เพื่อใช้วัดความสามารถทั่วไปในการอ่านเข้าใจความภาษาอังกฤษ และแต่ละเรื่องมีโครงสร้างความรู้ต่างกัน แต่มีระดับความยากง่ายในการอ่านไม่ต่างกัน และเหมาะสมสำหรับนิสิตชั้นปีที่ ๑ จึงดำเนินการสร้างและพัฒนาแบบสอบดังกล่าวเป็นขั้นตอนดังต่อไปนี้

ก. การเลือกสรรเรื่อง

การเลือกสรรเรื่องสำหรับสร้างแบบสอบครั้งนี้ ผู้วิจัยเลือกเรื่องที่ S. Atsawa-Ekawanit (Atsawa-Ekawanit, ๑๙๘๕) เลือกมาทำวิทยานิพนธ์ในระดับมหาบัณฑิต โดยผู้วิจัยเป็นอาจารย์ที่ปรึกษา แต่การทดสอบค่าสถิติต่าง ๆ ที่เกี่ยวข้องบางอย่าง ผู้วิจัยได้กระทำเอง การคัดเลือกเรื่องทำดังนี้

๑. คัดเลือกข้อความภาษาอังกฤษเพื่อการอ่าน ที่มีความยาวประมาณ ๓๐๐ คำ (เป็นเกณฑ์ที่กำหนดไว้ในหลักสูตรที่ต้องการให้นิสิตที่เรียนรายวิชา Foundation English : Reading II ต้องเรียน) จำนวน ๖ เรื่อง ซึ่งมีเนื้อหาเป็นภาษาอังกฤษทั่วไป (General English) ๓ เรื่อง และภาษาอังกฤษเฉพาะสาขาวิชา (English for Specific Purposes) ๓ เรื่อง จากหนังสือ ๔ เล่ม ซึ่งผู้วิจัยพิจารณาแล้วว่า เรื่องและระดับความยากง่ายของเรื่องในเชิงคำศัพท์ เนื้อหาสาระของเรื่อง และโครงสร้างทางภาษาของเรื่องอยู่ในระดับปีที่ ๑ ของมหาวิทยาลัย แหล่งดังกล่าวนี้คือ

- Ridgway, B., Exciting Stories with Exercises, ๑๙๘๑,
pp. ๓๙-๔๐

- Chaplen, F., A Course in Intermediate Scientific
English, ๑๙๘๑, pp. ๑๔, ๑๘-๑๙ (๒ เรื่อง)

- Hanson, W.J., Enquiries: Family Life, ๑๙๗๔, p. ๓๔

- Pearson, I, English in Focus: Biological Science,
๑๙๗๘, pp. ๑๑-๑๒

- Langan, J., Reading and Study Skills, ๑๙๗๘, p. ๒๘๒

๒. นำเรื่องที่คัดเลือกแล้วทั้ง ๖ เรื่อง มาสร้างเป็นแบบสอบโคลซ
(cloze test) เพื่อวัดระดับความยากง่ายในการอ่านตามวิธีของ W.L. Taylor
ผู้เป็นคนแรกที่ค้นคิดวิธีการโคลซขึ้น (Taylor, ๑๙๕๓:๔๑๕) และปัจจุบันนี้นิยมกันอย่าง
แพร่หลาย โดยการละทุกคำที่ ๘ ยกเว้นประโยคแรกและประโยคสุดท้าย ซึ่งเป็นวิธี
การทดสอบระดับความยากง่ายในการอ่านของเรื่องที่เหมาะสมสำหรับผู้อ่านปีที่ ๑
ระดับมหาวิทยาลัยในประเทศไทย (ดาราวดี นิตยศรี, ๒๕๑๘:๔๖) แต่ละเรื่องเมื่อ
ละคำแล้วมีที่ว่าง ๓๐ ช่องเท่ากัน

๓. นำแบบสอบโคลซทั้ง ๖ เรื่องมาจัดเป็น ๒ ชุด ๆ ละ ๓ เรื่อง จัดเรียง
สลับแบบสอบทั้ง ๒ ชุด อย่างเป็นระเบียบ แล้วนำไปสอบกับนักศึกษาชั้นปีที่ ๑ คณะ
วิทยาศาสตร์ จำนวน ๒ ห้องเรียน จำนวน ๖๓ คน ที่มีระดับความสามารถทั่วไปทาง
ภาษาอังกฤษที่วัดโดยแบบสอบจัดระดับความสามารถของภาควิชาภาษาต่างประเทศ
มหาวิทยาลัยมหิดล เมื่อเดือนมิถุนายน ๒๕๒๗ แล้ว มีค่ามัชฌิมเลขคณิตใกล้เคียงกับค่า
มัชฌิมเลขคณิตรวม คือ ๕๐.๙๘ มากที่สุด และค่ามัชฌิมเลขคณิตของ ๒ ห้องนั้น คือ ๔๘.๔๐
(S.D. = ๗.๓๔) และ ๔๗.๕๐ (S.D. = ๗.๑๕) เมื่อทดสอบด้วยค่าสถิติที (t-test)
แล้ว ไม่แตกต่างกันอย่างมีนัยสำคัญ ($p < .๐๕$) การสอบครั้งนี้ใช้เวลา ๑ ช.ม. ๓๐ นาที

๔. นำแบบสอบโคลงมาตรวจให้คะแนนโดยวิธีการให้คะแนนค่าที่สามารถยอมรับได้ว่า ถูกต้องตามเนื้อหาเรื่อง (Contextually acceptable scoring method) ซึ่งเป็นที่ยอมรับกันทั่วไปว่า เหมาะสมสำหรับผู้อ่านที่ไม่ใช่เจ้าของภาษา (Oller, ๑๙๗๒:๑๕๗ ; Hinoftis, ๑๙๘๐:๑๒๔) แล้วคิดคะแนนเป็นร้อยละ เพื่อเปรียบเทียบกับเกณฑ์ความยากง่ายในการอ่าน

๕. คัดเลือกเรื่องได้ ๔ เรื่อง ที่มีค่ามัชฌิมเลขคณิตระหว่างร้อยละ ๔๔-๕๓ ซึ่งเป็นค่าความยากง่ายในการอ่านของเรื่องที่มีอยู่ในเกณฑ์ที่ใช้เพื่อการเรียนการสอนได้ ซึ่งจากการวิจัยของ Anderson และ Hunt พบว่า เป็นระดับที่เหมาะสมเพื่อการเรียนการสอนสำหรับผู้อ่านที่ไม่ใช่เจ้าของภาษา และระดับดังกล่าวนี้เทียบเท่าประมาณร้อยละ ๗๔ ถ้าเรื่องนั้นทำเป็นแบบสอบแบบเลือกตอบ (Anderson and Hunt, ๑๙๗๐ อ้างจาก Soniano, ๑๙๗๓:๔๖)

๖. เพื่อต้องการตรวจสอบระดับความยากง่ายในการอ่านของเรื่องอีก และเพื่อคัดเลือกเรื่องมาใช้ในการวิจัยเพียง ๒ เรื่องเท่านั้น จึงนำแบบสอบโคลงทั้ง ๔ เรื่องไปสอบกับนักศึกษาปีที่ ๑ คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล อีกจำนวน ๖๔ คน จาก ๒ ห้องเรียน ซึ่งจากการสอบด้วยแบบสอบวัดระดับความสามารถชุดเดียวกันแล้วพบว่า ระดับความสามารถทั่วไปทางภาษาใกล้เคียงกับของนักศึกษาทั้ง ๒ กลุ่ม ดังได้กล่าวมาแล้ว (กลุ่มตัวอย่างต่างจากตัวอย่างในการสอบครั้งแรก) การสอบใช้เวลา ๒ ช.ม. ติดต่อกัน แล้วจึงคัดเลือกเรื่องมา ๒ เรื่อง โดยอาศัยเกณฑ์ดังนี้

ก. เรื่องทั้งสองมีระดับความยากง่ายในการอ่านไม่แตกต่างกันอย่างมีนัยสำคัญ

ข. เรื่องทั้งสองมีระดับความยากง่ายในการอ่านอยู่ในระดับที่ใช้เพื่อการเรียนการสอนได้ โดยอาศัยเกณฑ์ดังกล่าวแล้ว

ค. เรื่องหนึ่งเป็นเรื่องภาษาอังกฤษทั่วไป แต่อีกเรื่องหนึ่งเป็นภาษาอังกฤษเฉพาะสาขาวิชา ทั้งนี้เพื่อต้องการเรื่องที่มีโครงสร้างแนวคิดต่างกัน

ผลการคัดเลือกได้เรื่อง ๒ เรื่อง คือ The Meaning of Dreams ($\bar{X} = ๕๐.๖๔$ และ $r_{tt} = ๐.๘๒$) และเรื่อง Atomic Structure ($\bar{X} = ๕๑.๐๓$ และ $r_{tt} = ๐.๗๗$)

๗. นำแบบสอบมาปรับปรุงเป็นแบบสอบ ๒ แบบ เพื่อให้มีโครงสร้างแนวคิดต่างกัน โดยมีลักษณะดังต่อไปนี้

ก. แบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง เป็นแบบสอบฉบับเดิม ซึ่งไม่มีชื่อเรื่องและภาพประกอบเรื่องให้อ่าน แต่มีคำศัพท์ที่มีความหมายสำคัญต่อเรื่อง (No-context and Transparency) แบบสอบนี้จัดทำเป็น ๒ รูปแบบ คือ

๑. รูปแบบ ๑.๑ เป็นแบบที่เรียงลำดับเรื่องให้อ่านเป็น The Meaning of Dreams และ Atomic Structure

๒. รูปแบบ ๑.๒ เป็นแบบที่เรียงลำดับเรื่องให้อ่านเป็น Atomic Structure และ The Meaning of Dreams

ข. แบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง เป็นแบบสอบฉบับเดิม แต่มีชื่อเรื่องและภาพประกอบเรื่องให้อ่าน รวมทั้งมีคำศัพท์ที่มีความหมายสำคัญต่อเรื่อง (Context and Transparency) แบบสอบนี้จัดทำเป็น ๒ รูปแบบ คือ

๑. รูปแบบ ๒.๑ เป็นแบบสอบที่เรียงเรื่องเหมือนแบบ ๑.๑

๒. รูปแบบ ๒.๒ เป็นแบบสอบที่เรียงเรื่องเหมือนแบบ ๑.๒

อนึ่ง ในตอนท้ายของกระดาษคำตอบของแบบสอบแต่ละชนิด จะมีคำถามแบบเลือกตอบให้ผู้สอบตอบว่า ในระหว่าง ๒ เรื่องที่อ่านนั้น ผู้อ่านคุ้นเคยกับเนื้อหาสาระของเรื่องใดมากกว่า และคิดว่าเรื่องใดอ่านเข้าใจยากกว่า ทั้งนี้เพื่อใช้ข้อมูลสำหรับการศึกษาอิทธิพลของโครงสร้างแนวคิดด้านความคุ้นเคยกับเนื้อหาสาระของเรื่องให้อ่านต่อไป

ข. การสร้างแบบสอบ

๑. นำเรื่องที่คัดเลือกได้ทั้ง ๒ เรื่อง มาสร้างแบบสอบแบบเลือกตอบ เรื่องละ ๓๐ ข้อ รวมเป็น ๖๐ ข้อ ทั้งนี้เพื่อให้ได้จำนวนข้อสอบมากพอที่จะนำไปวิเคราะห์ด้วยรูปแบบโลจิสติกส์ใช้พารามิเตอร์ ๓ ตัว ซึ่งกำหนดไว้ว่า ควรมีข้อสอบมาก เช่น ประมาณ ๔๐ ข้อ เป็นต้น (Wood and Others, ๑๙๗๖:๕) และเพื่อสำรวจไว้หากจำเป็นต้องตัดข้อสอบบางข้อออกภายหลัง แบบสอบนี้สร้างตามสารระบบของ Valette และ Disick (๑๙๗๒) ซึ่งเป็นตำราการสอบทางภาษาที่ใช้กันแพร่หลายมากในปัจจุบันนี้ แบบสอบที่สร้างขึ้นนี้ แต่ละข้อมี ๔ ตัวเลือก ตามความนิยมของนักทดสอบทั่วไป และตามคำแนะนำของ Ebel ผู้เป็นนักทดสอบที่มีชื่อเสียงว่า เป็นจำนวนที่เหมาะสม (Ebel, ๑๙๗๒:๒๐๕)

๒. นำแบบสอบต้นฉบับที่สร้างขึ้นไปให้อาจารย์ชาวต่างประเทศที่สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย ตรวจสอบแก้ไขความถูกต้องทางภาษาของข้อสอบ ๑ ท่าน แล้วนำไปให้อาจารย์ที่ปรึกษาด้านการสอบทางภาษาตรวจสอบแก้ไขอีกครั้งหนึ่ง

๓. นำแบบสอบที่สร้างขึ้นไปสอบกับนิสิตชั้นปีที่ ๑ ของจุฬาลงกรณ์มหาวิทยาลัย จำนวน ๓๒๐ คน จาก ๑๔ คณะ ๆ ละ ๑ ห้องเรียน ทั้งนี้โดยการสุ่มตัวอย่างอย่างง่ายมาคณะละ ๑ ห้องเรียน การสอบใช้เวลา ๑ ชม. ในห้องเรียนพร้อมกันภายในสัปดาห์แรกของภาคปลายปีการศึกษา ๒๕๒๗ เพื่อต้องการสอบว่า เครื่องมือนี้เหมาะสมเพียงใดด้านคุณภาพของแบบสอบ และความเหมาะสมของเวลา เป็นต้น

๔. นำผลการสอบไปวิเคราะห์ข้อสอบรายข้อด้วยเครื่องคอมพิวเตอร์ของสถาบันบริการคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย โดยใช้โปรแกรมวิเคราะห์ข้อสอบวิธีประเพณีนิยม ที่ผู้วิจัยได้พัฒนาขึ้น ทั้งนี้เพื่อนำผลการวิเคราะห์มาปรับปรุงคุณภาพของแบบสอบให้มีคุณภาพดียิ่งขึ้น และผลการวิเคราะห์ข้อสอบบางส่วนมีดังนี้

ตารางที่ ๒ ผลการวิเคราะห์ข้อสอบรายข้อ ครั้งที่ ๑ (เพื่อการปรับปรุง)

$n = ๓๒๐$

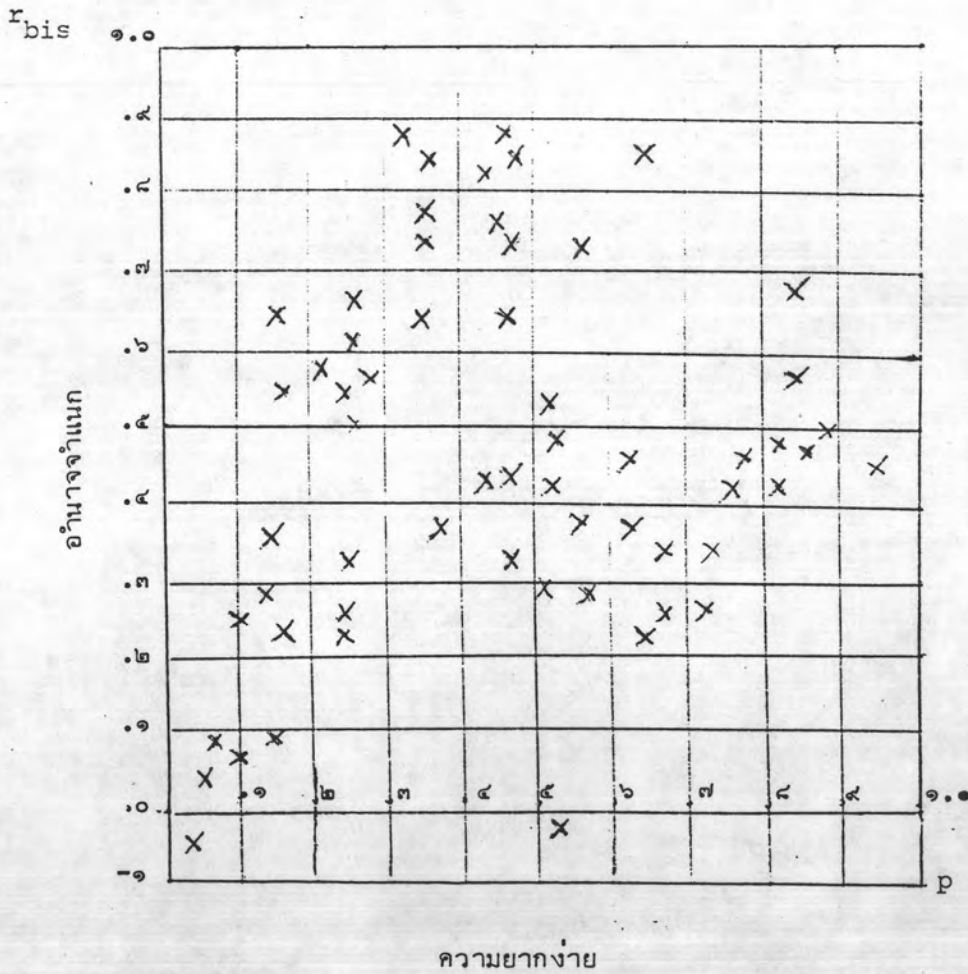
$k = ๖๐$

ค่าสถิติ การสอบ	\bar{X}	MIN	approx MDN	MAX	S.D	หมายเหตุ
คะแนนสอบ	๒๘.๓๐๐	๔.๐๐๐	๒๖.๕๐๐	๔๙.๐๐๐	๙.๙๖๐	$KR_{20} = ๐.๘๙๕$
p	๐.๔๗๒	๐.๐๖๖	๐.๕๑๑	๐.๙๕๖	๐.๕๐๕	$SE_m = ๓.๑๕๐$
Δ	๑๓.๓๓๘	๖.๑๒๕	๑๒.๖๐๐	๑๙.๐๗๕	๒.๗๖๘	
r_{bis}	๐.๔๗๑	-๐.๐๐๓	๐.๔๓๐	๐.๘๖๓	๐.๒๒๖	

จากตารางที่ ๒ นี้แสดงให้เห็นว่า โดยเฉลี่ยแบบสอบมีระดับความยากง่ายปานกลาง ($p = ๐.๔๗$) แต่มีข้อสอบบางข้อยากมาก ($p = ๐.๐๗$) และข้อสอบบางข้อง่ายมาก ($p = ๐.๙๖$) ข้อสอบมีอำนาจจำแนกโดยเฉลี่ยแล้วสูงพอสมควร ($r_{bis} = ๐.๔๗$) แต่ว่ามีข้อสอบบางข้อที่มีค่าอำนาจจำแนกติดลบ ซึ่งแสดงว่า นิสิตอ่อนทำได้ถูก แต่นิสิตเก่งทำผิด ซึ่งเป็นข้อสอบที่ไม่ดี แบบสอบนี้มีค่าความเที่ยงแบบ $KR_{20} = ๐.๘๙๕$ และ $SE_m = ๓.๑๕๐$

แผนภาพที่ ๑

ค่าความยากง่าย (p) และค่าอำนาจจำแนก (r_{bis})
 ของข้อสอบจากการวิเคราะห์ครั้งที่ ๑



จากแผนภาพที่ ๑ แสดงให้เห็นว่า มีข้อสอบอยู่ ๒ ข้อ ที่มีค่าอำนาจจำแนก ตีกลับ และอีก ๔ ข้อ ที่แม้ว่าค่าอำนาจจำแนกจะเป็นค่าบวก แต่ก็ยังมีค่าอำนาจจำแนก (r_{bis}) ต่ำกว่า ๐.๒๕ และมีข้อสอบอีก ๗ ข้อ ที่ง่ายมาก ($p > ๐.๘๐$) ข้อสอบที่เหลืออีก ๔๔ ข้อ มีแนวโน้มว่า จะเป็นข้อสอบที่จะใช้ในการวิจัยต่อไปได้ แต่จากการวิเคราะห์ผลการสอบ ผู้วิจัยพบว่า กลุ่มตัวอย่างประมาณร้อยละ ๔๐ ทำข้อสอบตั้งแต่ข้อที่ ๔๐ เป็นต้นไป ไม่ทันเวลาภายใน ๑ ช.ม. ข้อมูลเหล่านี้เป็นข้อมูลเบื้องต้นที่จะใช้ในการปรับปรุงคุณภาพข้อสอบต่อไป

๕. ใช้ผลการวิเคราะห์ข้อสอบรายข้อ เวลาในการสอบ และจุดมุ่งหมายของ เครื่องมือ ซึ่งจะใช้เป็นทั้งแบบสอบอิง เกณฑ์และแบบสอบอิงกลุ่ม ประกอบการพิจารณาในการ ปรับปรุงแบบสอบให้ดีขึ้น โดยผู้วิจัยได้ตัดข้อสอบที่มีอำนาจจำแนกติดลบ และข้อสอบที่วัดทักษะ ซ้ำกับข้อสอบอื่น แต่ควรมีคุณภาพดีต่อกว่าออกเรื่องละ ๕ ข้อ เนื่องจากค่าอำนาจจำแนก และค่าความยากง่ายจากการวิเคราะห์ข้อสอบรายข้อ เป็นเพียงค่าที่บอกแนวโน้มของอำนาจ จำแนก และความยากง่ายของแบบสอบเท่านั้น (Gronlund, ๑๙๗๖:๒๗๐-๒๗๒) ดังนั้น ข้อสอบบางข้อที่สำคัญเพราะสร้างขึ้นเพื่อวัดทักษะที่จำเป็นบางอย่างในการอ่าน เช่น วัดความ สามารถในการอ้างอิงสรุปและการตีความ แม้ว่าจะมีอำนาจจำแนกต่ำ แต่มีค่าเป็นค่าบวก ผู้วิจัยก็เลือกไว้ใช้ต่อไป หรือข้อสอบบางข้อที่ง่าย ($p > ๐.๘๐$) แต่ผู้วิจัยเห็นว่า เป็นข้อ สอบที่วัดทักษะบางอย่างที่การอ่านเข้าใจความจำเป็นต้องรู้ แม้ว่าจะเป็นทักษะพื้นฐาน เช่น การเข้าใจข้อเท็จจริงที่ระบุไว้แน่ชัดในเรื่องที่อ่าน ผู้วิจัยก็เลือกข้อสอบดังกล่าวนี้ไว้ใช้ต่อไป แต่ก็ได้พยายามเลือกข้อสอบที่เหมาะสมที่จะใช้ เป็นข้อสอบของแบบสอบอิงกลุ่ม และแบบสอบ อิงเกณฑ์ให้ได้มากที่สุด ทั้งนี้ก็เพราะว่า

ก. ข้อสอบของแบบสอบอิง เกณฑ์สามารถใช้เป็นข้อสอบของแบบสอบอิงกลุ่มได้ (Hively, ๑๙๗๔:๑๔๓)

ข. แบบสอบอิงกลุ่มที่ประกอบด้วยข้อสอบที่วัดวัตถุประสงค์เฉพาะของการสอบ อิงเกณฑ์ ก็สามารถใช้เป็นแบบสอบอิง เกณฑ์ได้ด้วย (Hambleton and Others, ๑๙๗๘: ๓) และ

ค. เกณฑ์ที่สำคัญในการแบ่งแบบสอบออกเป็นแบบ สอบอิงกลุ่มและแบบสอบอิง เกณฑ์ ที่ผู้วิจัยยึดถือคือ ลักษณะการแปรความหมาย (interpretation) ของคะแนนมากกว่าคุณลักษณะ อื่น เช่น การสร้างข้อสอบ และการเลือกข้อสอบ ซึ่งวิธีการของแบบสอบทั้ง ๒ แบบ ยังมีลักษณะ ที่ไม่แตกต่างกันอย่างชัดเจน (Brennan, ๑๙๘๔:๓๒๔; Block and Dockrell, ๑๙๘๔:๓๖)

ดังนั้น ผู้วิจัยจึงพยายามเลือกข้อสอบที่มีค่าความยากง่าย (p) ระหว่าง ๐.๒๐ ถึง ๐.๘๐ และมีค่าอำนาจจำแนก (r_{bis}) ตั้งแต่ ๐.๓๐ ขึ้นไป เพื่อให้เป็นข้อสอบของการ

สอบอิงกลุ่ม ซึ่ง เป็นเกณฑ์ที่นิยมใช้กันทั่วไป (Ebel, ๑๙๗๒:๔๐๖) แต่เนื่องจากสิ่งที่ต้องการสอบ ซึ่งได้แก่ ทักษะต่าง ๆ ที่จำเป็นในการอ่านเข้าใจความเป็นสิ่งสำคัญของการสอบครั้งนี้ ดังนั้น ผู้วิจัยจึงไม่ได้เลือกข้อสอบโดยอาศัยแต่เพียงค่าสถิติของแบบสอบที่วิเคราะห์ได้เท่านั้น (Ebel, ๑๙๗๒:๓๗๔; Mehrens and Lehmann, ๑๙๗๕:๓๓๒) แต่เลือกข้อสอบบางข้อที่มีลักษณะต่างจากเกณฑ์ดังกล่าวบ้าง เมื่อจำเป็น เพื่อให้ได้แบบสอบสำหรับการสอบอิงเกณฑ์ด้วย คือ เลือกข้อสอบที่มีค่าอำนาจจำแนกตั้งแต่ ๐.๒๕ ขึ้นไป และค่าความยากง่ายระหว่าง ๐.๑๕ ถึง ๐.๘๕ เพื่อใช้ในการปรับปรุงแก้ไขต่อไป (Ebel, ๑๙๗๒: ๓๙๙)

๖. ปรับปรุงตัวเลือกบางตัวของข้อสอบที่เลือกไว้แล้ว แต่ว่ามีค่าอำนาจจำแนกต่ำ รวมทั้งตัวเลือกของบางข้อที่มีผู้เลือกตอบจำนวนน้อย แล้วนำแบบสอบที่ปรับปรุงแล้วไปให้อาจารย์ที่ปรึกษาด้านการสอบทางภาษา และอาจารย์ชาวต่างประเทศที่เป็นเจ้าของภาษาตรวจสอบความถูกต้องอีกครั้งหนึ่ง

๗. ทำการวิเคราะห์ข้อสอบรายข้อด้วยเครื่องคอมพิวเตอร์อีกครั้งหนึ่ง เพื่อตรวจสอบค่าสถิติต่าง ๆ ที่เกี่ยวข้อง และเพื่อศึกษาแนวโน้มของค่าความเที่ยงของแบบสอบ ผลปรากฏว่า แบบสอบ ๕๐ ข้อ มีค่าสถิติต่าง ๆ ดังนี้

ตารางที่ ๓ ผลการวิเคราะห์ข้อสอบรายข้อ ครั้งที่ ๒ (เพื่อการตรวจสอบ)

n = ๓๒๐

k = ๕๐

ค่าสถิติ การสอบ	\bar{X}	MIN	approx MDN	MAX	S.D	หมายเหตุ
คะแนนสอบ	๒๓.๕๐๐	๕.๐๐๐	๒๓.๐๐๐	๔๒.๐๐๐	๘.๕๘๐	$KR_{20} = ๐.๘๘๒$
p	๐.๔๘๘	๐.๑๔๑	๐.๕๐๒	๐.๘๖๒	๐.๕๐๖	$SE_m = ๒.๕๗๐$
Δ	๑๓.๑๒๑	๘.๕๙๒	๑๒.๒๔๗	๑๗.๘๐๑	๒.๕๙๕	
r_{bis}	๐.๔๙๑	๐.๒๔๔	๐.๕๔๕	๐.๘๖๗	๐.๒๓๒	

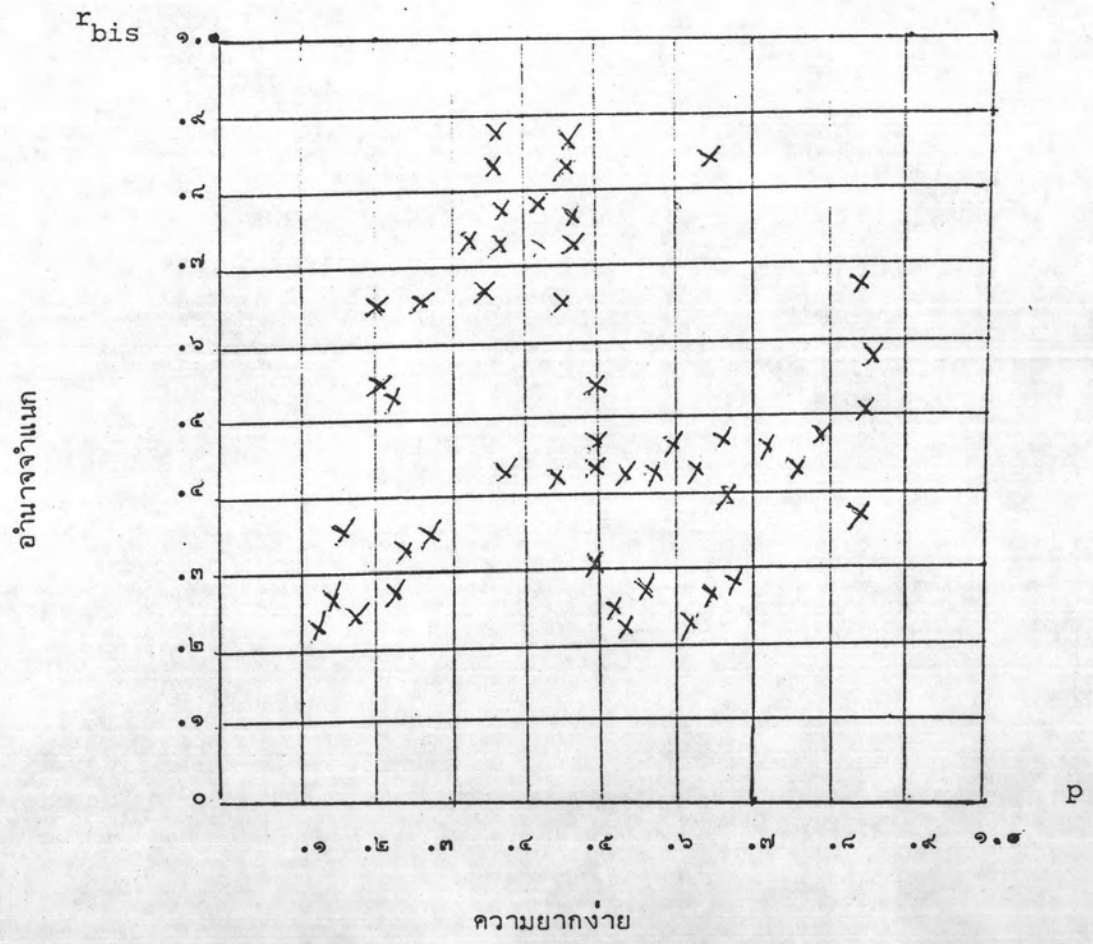
ตารางที่ ๓ แสดงว่า ค่าสถิติต่าง ๆ ของแบบสอบดีขึ้นกว่าเดิม เช่น ไม่มีข้อสอบที่อ่านง่ายจนเกินไปอีกต่อไป ข้อสอบมีค่าอ่านง่าย (r_{bis}) ตั้งแต่ ๐.๒๔๔ และมีค่าความยากง่ายโดยเฉลี่ย (p) = ๐.๔๘๘ ซึ่งนับว่าเป็นระดับที่ยากง่ายปานกลาง ค่าความยากง่ายนี้มีพิสัยระหว่าง ๐.๑๔๑ ถึง ๐.๘๖๒ และมีค่าความเที่ยงแบบ $KR_{20} = ๐.๘๘๒$ และ $SE_m = ๒.๕๗๐$ ค่าต่าง ๆ ที่เกี่ยวกับความยากง่ายและอ่านง่ายดังกล่าวนี้ แม้ว่าจะต่ำกว่าเกณฑ์ที่นิยมใช้ทั่วไปสำหรับการสอบแบบอิงกลุ่มเกี่ยวกับสัมฤทธิ์ผลในการเรียน แต่เพื่อการวิจัยครั้งนี้ ผู้วิจัยต้องการแบบสอบที่จะใช้ได้สำหรับการสอบอิงกลุ่ม และการสอบอิงเกณฑ์ เพื่อวัดความสามารถทั่วไป ดังนั้น เกณฑ์การคัดเลือกข้อสอบจึงต่างจากเกณฑ์ที่ใช้กันทั่วไป (Ebel, ๑๙๗๒:๓๔๘) อีกประการหนึ่ง ค่าสถิติต่าง ๆ เหล่านี้เป็นเพียงดัชนีชี้แนวโน้มของคุณลักษณะของแบบสอบเท่านั้น และการวิเคราะห์ข้อสอบครั้งที่ ๒ ยังอาศัยข้อมูลเดิม ดังนั้น เพื่อการศึกษาครั้งนี้ ผู้วิจัยมีความเห็นว่า ข้อสอบที่เลือกไว้ทั้ง ๕๐ ข้อ สนองจุดมุ่งหมายในการสร้างแล้ว

อนึ่ง รายละเอียดเกี่ยวกับค่าความยากง่ายและค่าอ่านง่าย มีดังต่อไปนี้

แผนภาพที่ ๒

ค่าความยากง่าย (p) และค่าอำนาจจำแนก (r_{bis})

ของข้อสอบ ๕๐ ข้อ ที่เลือกสรรแล้ว



จากแผนภาพที่ ๒ จะเห็นได้ว่า คุณลักษณะของข้อสอบด้านความยากง่ายและคำ
อำนาจำแนกดีขึ้นกว่าเดิมดังได้กล่าวมาแล้ว ส่วนด้านเนื้อหาของข้อสอบมุ่งวัดทักษะและ
ความสามารถต่าง ๆ ในขั้นที่ ๒ ๓ และ ๔ ของการอ่านเข้าใจความทางภาษา ตามที่
Valette และ Disick (Valette and Disick, ๑๙๗๒:๑๖๑) ได้เสนอไว้แล้ว
และเป็นที่ยอมรับกันแพร่หลายในหมู่นักทดสอบทางภาษา แต่ว่าแบบสอบที่สร้างขึ้นนี้ไม่มีข้อ
ทดสอบวัดทักษะหรือความสามารถทางกลไกของการอ่าน เช่น การสะกดคำหรือการเปรียบเทียบ
คำศัพท์ว่ามีตัวสะกดเหมือนกันหรือไม่ ซึ่งถือว่าเป็นทักษะหรือความสามารถขั้นที่ ๑
ในการอ่านเข้าใจความ ทั้งนี้เพราะผู้วิจัยมีความเห็นว่า ทักษะหรือความสามารถดังกล่าว
ง่ายเกินไปสำหรับผู้อ่านในระดับมหาวิทยาลัย และไม่มีการเรียนการสอนตามหลักสูตร

รายละเอียดของข้อสอบที่มุ่งวัดทักษะ หรือความสามารถในการอ่านเข้าใจความ
ภาษาอังกฤษที่ใช้ในการศึกษาครั้งนี้ มีรายละเอียดดังนี้

ตารางที่ ๔ รายละเอียดด้านเนื้อหาของข้อสอบ

ชั้นที่	ทักษะ/ความสามารถ	เรื่อง Dreams ข้อที่	รวม	เรื่อง Atom ข้อที่	รวม
๑	๑. vocab in context	๗ ๘	๒	๔ ๑๐	๒
	๒. paraphrase	๖ ๑๑	๒	-	-
	๓. reference	๔ ๑๐	๒	๗ ๘	๒
๒	๑. specific details	๑ ๒ ๓ ๔ ๑๒ ๑๔	๖	๑ ๒ ๓ ๔ ๕ ๑๒ ๑๔ ๒๑	๘
	๒. sequence	-	-	๑๓ ๑๖ ๑๗ ๑๔	๔
๓	๓. main idea	๒๔ ๒๕	๒	๒๔ ๒๕	๒
	๔. cause-effect	๑๓	๑	๖ ๑๑ ๑๔	๓
	๕. stated conclusion	๑๔ ๑๖	๒	๑๘ ๒๓	๒
	๑. point of view	-	-	-	-
Criticism	๒. inference	๕ ๑๗ ๑๘ ๑๙ ๒๐	๕	๒๐ ๒๒	๒
	๓. tone	-	-	-	-
	๔. implicated conclusion	๒๑ ๒๒ ๒๓	๓	-	-
รวม			๒๕		๒๕

ข้อมูลในตารางที่ ๔ แสดงว่า ข้อความที่เกี่ยวกับความฝัน ซึ่งมีเนื้อหาเป็นภาษาอังกฤษทั่วไปนั้น มีข้อสอบที่วัดทักษะหรือความสามารถขั้นในจำนวนที่มากเท่า ๆ กัน คือ ๖:๑๑:๘ แต่ข้อความที่เกี่ยวกับอะตอม ซึ่งมีเนื้อหาเป็นภาษาอังกฤษเฉพาะสาขาวิชา มีข้อสอบที่วัดทักษะหรือความสามารถขั้นที่ ๓ ได้เพียง ๒ ข้อ ข้อสอบโดยมากคือ ๑๙ ข้อ เป็นข้อสอบที่วัดขั้นที่ ๒ ที่เป็น เช่นนี้เพราะลักษณะทั่วไปของข้อความที่เป็นภาษาอังกฤษทั่วไป และข้อความที่เป็นภาษาอังกฤษเฉพาะสาขาวิชา โดยเฉพาะทางด้านวิทยาศาสตร์เป็นเช่นนี้ กล่าวคือ ข้อความที่มีเนื้อหาเป็นภาษาอังกฤษทั่วไปมักเน้นที่การสื่อสาร และวิพากษ์วิจารณ์ แต่ข้อความที่มีเนื้อหาทางวิทยาศาสตร์ มักเน้นที่ข้อเท็จจริงมากกว่าอย่างอื่น เป็นต้น ดังนั้น ผู้วิจัยจึงจำเป็นต้องสร้างข้อสอบให้สอดคล้องกับลักษณะของเนื้อหาของข้อความที่นำมาใช้เพื่อการสอบครั้งนี้

การเก็บรวบรวมข้อมูล

เพื่อให้ได้ข้อมูลสำหรับการศึกษาครั้งนี้ ผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลเป็นขั้นตอนดังต่อไปนี้

๑. ติดต่อกับกรรมการวิชาการของสถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย ให้ใช้ผลการสอบจากแบบสอบดังกล่าว เป็นคะแนนเก็บรายวิชาภาษาอังกฤษพื้นฐาน ๒ เฉพาะการอ่าน ได้ ๑๐% ถ้าอาจารย์เจ้าของรายวิชาต้องการ และขออนุญาตจัดสอบแบบสอบดังกล่าว กับนิสิตทุกห้องเรียนที่เรียนรายวิชาดังกล่าว และได้รับการอนุมัติให้เป็นผลวิจัย

๒. ติดต่อกับอาจารย์เจ้าของรายวิชา เพื่อขอความร่วมมือในการจัดแบบสอบดังกล่าวตามแต่อาจารย์ผู้สอนจะเห็นสมควร ภายในวันที่ ๑๗-๒๒ ธันวาคม ๒๕๒๗

๓. จัดเรื่องแบบสอบทั้ง ๔ รูปแบบ ให้เป็นระบบเดียวกัน คือ เรียงแบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง และแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง ให้เรื่องที่อ่าน ๒ เรื่อง เรียงสลับกัน ต่อเนื่องกันไป ทั้งนี้เพื่อกำจัดอิทธิพลของการสอบเรื่องที่อ่านเรื่องใดเรื่องหนึ่งก่อนหรือหลังกัน ซึ่งอาจมีผลต่อคะแนนการสอบโดยวิธีการถ่วงดุล จากการที่ได้ทดสอบแล้วว่า เรื่องที่ใช้เพื่อการสอบครั้งนี้มีความยากง่ายไม่ต่างกัน ทำให้เชื่อได้ว่าอิทธิพลดังกล่าวคงจะมีน้อยมาก นอกจากนี้ การจัดเรียงแบบสอบดังกล่าวนี้จะทำให้การสอบมีลักษณะเหมือนการสุ่ม และจะทำให้จำนวนผู้สอบของแต่ละแบบสอบใกล้เคียงกัน

๔. นับและบรรจุของแบบสอบตามจำนวนนิสิตของแต่ละห้องเรียน

๕. แจกแบบสอบแก่อาจารย์เจ้าของรายวิชา เฉพาะห้องที่ไม่ได้ใช้เพื่อการทดสอบเครื่องมือครั้งที่แล้ว เพื่อนำไปสอบกับนิสิตในความรับผิดชอบของตนตามวันและเวลาที่ได้นัดหมายกันไว้แล้ว ดังนั้น อาจารย์เจ้าของรายวิชาเป็นผู้บริหารการสอบเอง ตามคำแนะนำที่ผู้วิจัยได้เขียนแจ้งไว้แล้วในของข้อสอบ เช่น ให้นิสิตนั่งห่างกันพอสมควร และให้แบบสอบตามที่ได้จัดเรียงไว้ให้แล้ว เป็นต้น

ส่วนนิสิตนั้น ผู้วิจัยได้เขียนคำแนะนำการตอบแบบสอบไว้ในแผ่นแรกของแบบสอบแต่ละรูปแบบด้วย สำหรับปัญหาต่าง ๆ เกี่ยวกับการสอบที่อาจเกิดขึ้นในระหว่างการสอบ ผู้วิจัยเชื่อว่าอาจารย์ผู้คุมสอบทุกท่านสามารถแก้ไขได้ เพราะทุกท่านเป็นอาจารย์สอนภาษาอังกฤษ และต้องคุมสอบอยู่เป็นประจำอยู่แล้ว

๖. ตรวจสอบความสมบูรณ์ของกระดาษคำตอบที่ได้รับคืนมา โดยเฉพาะอย่างยิ่งหมายเลขประจำตัวของนิสิต เพราะต้องใช้ในการจับคู่กับคะแนนสอบกลางภาคและปลายภาคของนิสิตในภายหลัง ถ้าข้อมูลดังกล่าวไม่สมบูรณ์ ผู้วิจัยก็จะเขียนเติมให้เอง เมื่อได้ตรวจสอบชื่อของนิสิตกับทะเบียนลงวิชาเรียน (CR 52) แล้ว

จากการตรวจนับด้วยมือปรากฏว่า มีนิสิตทำแบบสอบที่ไม่มีชื่อ เรื่องและภาพประกอบเรื่อง จำนวน ๔๕๔ คน และแบบสอบที่มีชื่อ เรื่องและภาพประกอบเรื่อง จำนวน ๔๑๘ คน ดังนั้น จำนวนพลวิจัยเพื่อการวิจัยครั้งนี้มีมากเพียงพอที่จะใช้รูปแบบโลจิสติกส์ใช้พารามิเตอร์ ๓ ตัวเพื่อการวิเคราะห์ข้อสอบได้ เพราะสอดคล้องกับคำแนะนำในการใช้ที่ว่า ถ้าจะให้ผลการคำนวณถูกต้อง ควรมีผู้สอบจำนวนมาก เช่น ประมาณ ๑,๐๐๐ คน และมีข้อสอบจำนวนมาก เช่น ประมาณ ๔๐ ข้อ (Wood and Other, ๑๙๗๖:๕)

๗. นำกระดาษคำตอบที่ได้ทั้งหมด ส่งให้กับเจ้าหน้าที่เตรียมข้อมูลของสถาบันบริการคอมพิวเตอร์ เพื่อเตรียมข้อมูลลงในเทปแม่เหล็กสำหรับใช้ต่อไป

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลเพื่อการวิจัยครั้งนี้ ใช้เครื่องคอมพิวเตอร์ ของสถาบันบริการคอมพิวเตอร์ ของจุฬาลงกรณ์มหาวิทยาลัย โดยใช้โปรแกรมสำเร็จรูป ชื่อ SPSSX (Statistical Packages for the Social Sciences X, ๑๙๘๓) หรือโปรแกรมที่ผู้วิจัยได้เขียนขึ้น หรือปรับปรุงขึ้นเอง เช่น โปรแกรมสำหรับการวิเคราะห์ข้อสอบวิธีประเพณีนิยม และตามรูปแบบราสซ์ รวมทั้งโปรแกรมสำหรับการคำนวณหาค่าดัชนีความเชื่อถือ ตามแนวคิดของ Brennan และ Kane เป็นต้น นอกจากนี้ การวิเคราะห์ข้อมูลส่วนหนึ่งใช้เครื่องคอมพิวเตอร์ของสำนักงานสถิติแห่งชาติ คือ การวิเคราะห์ข้อสอบโดยรูปแบบโลจิสติกส์ใช้พารามิเตอร์ ๓ ตัว และบางส่วนใช้เครื่องคอมพิวเตอร์จุฬาลงกรณ์มหาวิทยาลัย เช่น การทดสอบความแตกต่างของค่าความเที่ยง และค่าความตรงเชิงพยากรณ์ เป็นต้น

ขั้นตอนต่าง ๆ ในการวิเคราะห์ข้อมูล มีดังนี้

๑. ใช้โปรแกรมคอมพิวเตอร์ ชื่อ ITEMEX ที่ผู้วิจัยได้ปรับปรุงขึ้นใช้ทำการวิเคราะห์ข้อสอบรายข้อของแบบสอบแต่ละแบบ เพื่อหาค่าความยากง่าย (p และ Δ) ค่าอำนาจจำแนก (r_{bis}) และค่าความเที่ยงของแบบสอบ แล้วนำค่าเดลต้า (Δ) ไปเป็นค่าน้ำหนักคะแนนของแบบสอบทั้ง ๒ แบบ ดังกล่าวแล้ว
๒. ใช้โปรแกรมคอมพิวเตอร์ ชื่อ FACTOR ของ SPSSX ทำการวิเคราะห์ตัวประกอบของแบบสอบแต่ละแบบเพื่อทดสอบว่า แบบสอบวัดเพียงมิติเดียวหรือไม่ (Lord, ๑๙๘๐:๒๑) การวิเคราะห์นี้ใช้วิธีการวิเคราะห์ตัวประกอบสำคัญ (principal component) ด้วย Subprogram ชื่อ PAL แล้วหมุนด้วยวิธีแวนซ์ และใช้ค่าสหสัมพันธ์แบบเพียร์สัน โพรดักต์โมเมนต์ (Pearson product moment) แทนค่าสหสัมพันธ์แบบเตรตาคอริก (tetrachoric) ทั้งนี้เพราะกลุ่มตัวอย่างมีขนาดใหญ่ (ประมาณ ๔๐๐-๔๕๐ คน) จึงทำให้ค่าดังกล่าวนี้มีขนาดใกล้เคียงกัน (Matzek, ๑๙๗๔:๔๐) การวิเคราะห์นี้ได้คำนวณความแปรปรวนร่วม (h^2) ของข้อสอบแต่ละข้อ และค่าน้ำหนักของตัวประกอบ

(factor loading: a_{ij}) ของแต่ละข้อด้วย แล้วนำค่าผลบวกความแปรปรวนร่วมไปเป็น คำนวณน้ำหนักคะแนนของแบบสอบทั้ง ๒ แบบ

จากการวิเคราะห์ตัวประกอบ ปรากฏว่า แบบสอบแต่ละแบบมี eigenvalue มากกว่า ๑.๐๐ อยู่ ๕ จำนวน ซึ่งถือว่าเป็นตัวประกอบร่วม (Lindeman and Others, ๑๙๘๐:๒๕) eigenvalue ของตัวประกอบที่ ๑ (หรือลำดับที่ ๑) ของแบบสอบแต่ละแบบ มีขนาดแตกต่างกันจาก eigenvalue ของตัวประกอบที่ ๒ มาก และ eigenvalue ลำดับที่ ๒ ของแบบสอบแต่ละแบบมีขนาดแตกต่างจาก eigenvalue ลำดับต่อ ๆ ไปที่เหลือน้อยไม่มากนัก จึงทำให้กล่าวได้ว่า แบบสอบแต่ละแบบวัดเพียงมิติเดียว (Lord, ๑๙๘๐:๒๑) [แสดงไว้ในแผนภาพที่ ๕-๖ ในภาคผนวก ข.]

๓. ใช้โปรแกรมคอมพิวเตอร์ ชื่อ STATPLUS ซึ่งใช้กับเครื่องคอมพิวเตอร์ จุลภาค ทำการคำนวณหาค่าสหสัมพันธ์แบบเพียร์สัน โพรดักต์โมเมนต์ ระหว่างค่าไบซีเรียล (r_{bis}) กับค่าน้ำหนักของตัวประกอบที่ ๑ ของข้อสอบแต่ละข้อ สำหรับแบบสอบแต่ละแบบ เพื่อเป็นการตรวจสอบอีกครั้งหนึ่งว่า แบบสอบเหล่านี้วัดเพียงมิติเดียวจริงหรือไม่ (McBride and Weiss, ๑๙๗๔:๓๑ อ้างจาก Warm, ๑๙๗๘:๑๐๔) ปรากฏว่า ความสัมพันธ์ของค่าดังกล่าวแล้วมีค่าสูงมาก คือ $r_{xy} = ๐.๙๗๑$ และ ๐.๙๗๖ ตามลำดับ เนื่องจากค่าดังกล่าวนี้สูงมาก และเกิน ๐.๘๐ ซึ่งถือว่าเป็นเกณฑ์ จึงกล่าวได้ว่าแบบสอบแต่ละแบบวัดเพียงมิติเดียว (Warm, ๑๙๗๘:๑๐๔) [แสดงไว้ในตารางที่ ๑๘ ในภาคผนวก ข.]

ดังนั้น จากผลการทดสอบความเป็นมิติเดียวของแบบสอบทั้ง ๒ แบบดังกล่าวแล้ว ทำให้มั่นใจได้ว่า แบบสอบแต่ละแบบวัดเพียงมิติเดียว และสามารถนำผลการสอบไปวิเคราะห์ข้อสอบรายข้อด้วยรูปแบบราสซ์ และรูปแบบโลจิสติกใช้พารามิเตอร์ ๓ ตัว ต่อไปได้ เพราะ ลักษณะแบบสอบสอดคล้องกับข้อตกลงเบื้องต้นของวิธีดังกล่าวแล้วนี้

๔. ใช้โปรแกรมคอมพิวเตอร์ ชื่อ BICAL-3 ที่ผู้วิจัยได้ปรับปรุงขึ้นใช้ทำการวิเคราะห์ข้อสอบรายข้อของแบบสอบแต่ละแบบ เพื่อหาค่าความยากง่าย (β) ของข้อสอบรายข้อ โปรแกรมนี้เป็นการวิเคราะห์ข้อสอบตามรูปแบบวิธีราสซ์

๕. ใช้โปรแกรมคอมพิวเตอร์ ชื่อ LOGIST-4 ที่สำนักงานสถิติแห่งชาติทำการวิเคราะห์ข้อสอบรายข้อของแบบสอบแต่ละแบบ เพื่อหาค่าอำนาจจำแนก (a) ค่าความยากง่าย (b) และค่าการเดา (c) ของข้อสอบรายข้อ โปรแกรมนี้เป็นการวิเคราะห์ข้อสอบตามรูปแบบโลจิสติกใช้พารามิเตอร์ ๓ ตัว

๖. ใช้เครื่องคอมพิวเตอร์จุลภาค คำนวณเปลี่ยนค่าความยากง่าย (β) จากรูปแบบราสซ์ของแบบสอบทั้ง ๒ แบบ เป็นคะแนนมาตรฐานเพื่อกำจัดค่าติดลบ โดยใช้สูตร $\beta_{t_i} = 1 + \alpha \beta_i$ เมื่อ $\beta_i =$ ค่าความยากง่าย (β)

๗. ใช้เครื่องคอมพิวเตอร์จุลภาค คำนวณเปลี่ยนค่าอำนาจจำแนก (a) ค่าความยากง่าย (p) และค่าการเดา (c) เป็นค่าน้ำหนักคะแนนที่เหมาะสม (w) ของแบบสอบทั้ง ๒ แบบ ตามวิธีของ Lord (Lord, ๑๙๕๐:๗๖) คือ

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

ในเมื่อ $a_i =$ ค่าอำนาจจำแนกของข้อที่ i

$c_i =$ ค่าการเดาของข้อที่ i

$p_i =$ ค่าความยากง่ายของข้อที่ i จากการวิเคราะห์ข้อสอบ

วิธีประเพณีนิยม

อนึ่ง คะแนนของแบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง และแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง ที่คำนวณจากการให้คะแนนทั้ง ๕ วิธี ได้แสดงไว้ในตารางที่ ๑๔ ในภาคผนวก ข.

๘. ใช้ค่าที่คำนวณได้จากการให้คะแนนแต่ละวิธี คือ ก) วิธีใช้ค่าเดลต้า (Δ) ข) วิธีใช้ค่าเบต้าแปลงแล้ว (β_{\pm}) ค) วิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม (w) ง) วิธีใช้ค่าผลบวกของความแปรปรวนร่วม (h^2) และ จ) วิธีประเพณีนิยม ไปเป็นคะแนนรายข้อของข้อสอบที่ตอบถูกแต่ละแบบสอบ แล้วใช้โปรแกรมชื่อ RELIABILITY ของ SPSSX คำนวณหาค่าความเที่ยงแบบ Cronbach's α

๙. ใช้โปรแกรมคอมพิวเตอร์ ชื่อ PEARSON CORR ของ SPSSX คำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนของแบบสอบแต่ละแบบที่ให้คะแนนรายข้อต่างกันแต่ละวิธี กับคะแนนรวมของการสอบกลางภาคและปลายภาคการศึกษารายวิชาการอ่าน ๒

๑๐. ใช้โปรแกรมคอมพิวเตอร์ ชื่อ OLKIN ซึ่งใช้กับเครื่องคอมพิวเตอร์ จุลภาคทดสอบความแตกต่างอย่างมีนัยสำคัญของค่าความตรงเชิงพยากรณ์ ตามวิธีของ Olkin (อ้างจาก Patnaik and Traub, ๑๙๗๓:๒๘๔) ด้วยโปรแกรมที่ผู้วิจัยเขียนขึ้น และได้ตรวจสอบความถูกต้องของผลการคำนวณ โดยอาศัยข้อมูลและผลลัพธ์ที่เสนอโดย Glass และ Stanley (Glass and Stanley, ๑๙๗๒:๒๓๔) [สูตรการคำนวณได้แสดงไว้ในภาคผนวก ข.]

๑๑. จัดเรียงข้อสอบของแบบสอบแต่ละแบบตามอันดับของคะแนนของแต่ละข้อ และการให้คะแนนแต่ละวิธี โดยเรียงจากข้อที่มีคะแนนมากที่สุดไปหาข้อที่มีคะแนนน้อยที่สุด แล้วแบ่ง

แบบสอบของการให้คะแนนแต่ละวิธีออกเป็น ๒ ส่วน ตามลำดับข้อคู่และคี่ และใช้โปรแกรม ชื่อ PEARSON CORR ของ SPSSX คำนวณหาค่าความแปรปรวน และค่าความแปรปรวนร่วม ของแบบสอบเหล่านี้

๑๒. ใช้เครื่องคอมพิวเตอร์จุลภาคทดสอบค่านัยสำคัญของความแตกต่างของค่า ความแปรปรวนของแบบสอบแต่ละคู่ที่แบ่งครึ่ง โดยวิธี F_{\max} (Winer, ๑๙๖๒:๒๐๖) ถ้าค่านัยของแบบสอบคู่ใดแตกต่างกันอย่างมีนัยสำคัญ ($p < .๐๕$) ก็จะทำให้การแบ่งแบบสอบ ที่ให้คะแนนวิธีนั้น ๆ ใหม่จนกว่าค่าความแปรปรวนที่ได้ไม่แตกต่างกันอย่างมีนัยสำคัญ ทั้งนี้ เพื่อให้สอดคล้องกับข้อตกลงเบื้องต้นของการทดสอบความแตกต่างของค่าความเที่ยง (Kristof, ๑๙๖๔:๑๐๕)

๑๓. ใช้โปรแกรมชื่อ ITERATE ซึ่งใช้กับเครื่องคอมพิวเตอร์จุลภาคทดสอบ ความแตกต่างอย่างมีนัยสำคัญ ($p < .๐๕$) ของค่าความเที่ยงของแบบสอบอิงกลุ่ม ที่ได้ จากการให้คะแนนข้อสอบต่างกันแต่ละวิธีของแบบสอบที่มีโครงสร้างความรู้ต่างกัน ผู้วิจัย เขียนโปรแกรมเพื่อการทดสอบนี้ขึ้นเองตามวิธีการคำนวณของ Kristof (Kristof, ๑๙๖๔:๑๐๕-๑๑๑) แล้วตรวจสอบความถูกต้องของผลการคำนวณ โดยอาศัยข้อมูลและผลลัพธ์ การคำนวณที่ Kristof เสนอไว้เป็นเกณฑ์ [สูตรการคำนวณได้แสดงไว้ในภาคผนวก ข.]

๑๔. ใช้โปรแกรมคอมพิวเตอร์ ชื่อ BRENNAN ซึ่งผู้วิจัยเขียนขึ้นเพื่อคำนวณ หาค่าความเที่ยงของแบบสอบแบบอิง เกณฑ์ สำหรับแบบสอบแต่ละแบบที่มีโครงสร้างความรู้ ต่างกัน เมื่อให้คะแนนด้วยวิธีต่าง ๆ และมีจุดตัดตั้งแต่ร้อยละ ๐-๑๐๐ โปรแกรมดังกล่าวนี้ เขียนขึ้นเพื่อการคำนวณตามวิธีของ Brennan และ Kane (Brennan and Kane, ๑๙๗๗: ๒๗๗-๒๘๓) และใช้กับเครื่องคอมพิวเตอร์ของสถาบันบริการคอมพิวเตอร์ ผู้เขียนตรวจสอบ ความถูกต้องของผลการคำนวณ โดยอาศัยข้อมูลและผลลัพธ์ที่เสนอโดย Brennan และ Kane [สูตรการคำนวณได้แสดงไว้ในภาคผนวก ข.]

๑๕. นำค่าความเที่ยง ณ จุดตัดต่าง ๆ ของแบบสอบอิงเกณฑ์ที่คำนวณได้มาเขียน เป็นกราฟเปรียบเทียบ

๑๖. ใช้โปรแกรมคอมพิวเตอร์ ชื่อ NONPAR CORR ของ SPSSX คำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างอันดับที่ของคะแนนของแบบสอบแต่ละแบบที่ให้คะแนนรายข้อต่างกัน กับอันดับที่ของคะแนนที่ได้จากการให้คะแนนวิธีประเพณีนิยม

๑๗. ใช้โปรแกรมคอมพิวเตอร์ ชื่อ PEARSON CORR ของ SPSSX คำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างกันและกันของคะแนนของแบบสอบแต่ละแบบที่ให้คะแนนรายข้อโดยวิธีต่าง ๆ ทั้ง ๕ วิธี

๑๘. ใช้โปรแกรมคอมพิวเตอร์ ชื่อ ANOVA ของ SPSSX ทำการวิเคราะห์ความแปรปรวนเพื่อทดสอบว่า โครงสร้างความรู้เกี่ยวกับชื่อเรื่องและภาพประกอบเรื่องที่มีคำศัพท์ที่มีความหมายสำคัญ และความคุ้นเคยกับเนื้อหาสาระของเรื่องที่ย่าน มีผลต่อการอ่านของนิสิตอย่างมีนัยสำคัญหรือไม่

๑๙. ใช้โปรแกรมคอมพิวเตอร์ ชื่อ CROSSTABS ของ SPSSX ทำการทดสอบความสัมพันธ์ระหว่างจำนวนผู้ตอบว่า ค้นเคยหรือไม่คุ้นเคยกับเรื่องที่ย่าน และผู้ที่มีความรู้สึกว่าย่านที่ย่านยากหรือง่าย ว่ามีมากน้อยเพียงใด