

บทที่ ๔

บทสรุป

### วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์หลักเพื่อเปรียบเทียบว่า ผลการให้คะแนนรายชื่อของแบบสอบต่างกัน ๕ วิธี คือ ๑) วิธีใช้ค่าเฉลี่ย ( $\Delta$ ) ๒) วิธีใช้ค่าเบต้าแปลง ( $\beta_c$ ) ๓) วิธีใช้ค่าผลบวกความแปรปรวนร่วม ( $h^2$ ) ๔) วิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) และ ๕) วิธีประเพณีนิยม ทำให้แบบสอบการอ่านเข้าใจความภาษาอังกฤษของแบบอิงกลุ่ม มีค่าความตรงเชิงพยากรณ์และความเที่ยงแตกต่างกันหรือไม่ และในกรณีที่เป็นแบบสอบอิงเกณฑ์ ค่าความเที่ยงหรือค่าดัชนีความเชื่อถือจากการให้คะแนนแต่ละวิธีดังกล่าวแล้วแตกต่างกันหรือไม่ และเพื่อศึกษาว่า อันดับของคะแนนสอบ เมื่อคิดคะแนนต่างวิธีกัน จะมีความคงที่มากน้อยเพียงใด สัมพันธ์กับคะแนนของวิธีประเพณีนิยมมากน้อยเพียงใด รวมทั้งเพื่อศึกษาว่า โครงสร้างความรู้ของ เรื่องที่อ่านมีผลต่อความสามารถทั่วไปในการอ่านเข้าใจความภาษาอังกฤษหรือไม่

### สมมติฐานในการวิจัย

๑. ในกรณีการสอบแบบอิงกลุ่ม แบบสอบที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีค่าความตรงเชิงพยากรณ์สูงกว่าค่าความตรงเชิงพยากรณ์ของแบบสอบที่ให้คะแนนโดยวิธีอื่น ๆ รวมทั้งวิธีประเพณีนิยม อย่างมีนัยสำคัญ ไม่ว่าแบบสอบนั้นจะมีโครงสร้างความรู้แบบใดก็ตาม

๒. ในกรณีการสอบแบบอิงกลุ่ม แบบสอบที่ให้คะแนนรายชื่อโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีค่าความเที่ยงสูงกว่าค่าความเที่ยงของแบบสอบที่ให้คะแนนโดยวิธีอื่น ๆ รวมทั้งวิธีประเพณีนิยมอย่างมีนัยสำคัญ ไม่ว่าแบบสอบนั้นจะมีโครงสร้างความรู้แบบใดก็ตาม

๓. ในกรณีการสอบแบบอิงเกณฑ์ แบบสอบที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม (w) มีค่าความเที่ยง (ค่าดัชนีความเชื่อถือ) สูงกว่าค่าความเที่ยงของแบบสอบที่ให้คะแนนโดยวิธีอื่น ๆ ไม่ว่าแบบสอบนั้นจะมีโครงสร้างความรู้แบบใดก็ตาม

๔. การให้คะแนนต่างกันแต่ละวิธี ทำให้อันดับที่ของคะแนนสอบการอ่านเข้าใจ ความภาษาอังกฤษเปลี่ยนไป

๕. คะแนนที่ได้จากวิธีการให้คะแนนรายข้อทั้ง ๕ วิธี ต่างมีความสัมพันธ์ซึ่งกันและกันอย่างมีนัยสำคัญ

๖. โครงสร้างความรู้ที่เกิดจากชื่อเรื่องและภาพประกอบเรื่องที่มีคำศัพท์ ซึ่งมีความหมายสำคัญต่อเรื่องที่อ่าน และความคุ้นเคยต่อเนื้อหาสาระของเรื่องที่อ่าน ต่างมีผลต่อความสามารถทั่วไปในการอ่านเข้าใจความภาษาอังกฤษของนิสิตจุฬาลงกรณ์มหาวิทยาลัย ที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศ อย่างมีนัยสำคัญ

๗. ความคุ้นเคยต่อเรื่องที่อ่าน และความยากง่ายของเรื่องมีความสัมพันธ์กันอย่างมีนัยสำคัญ

### วิธีวิจัย

#### ๑. พลวิจัย

พลวิจัยของการวิจัยครั้งนี้ ได้แก่ นิสิตชั้นปีที่ ๑ ที่เรียนวิชาภาษาอังกฤษ พื้นฐาน เฉพาะการอ่าน ๒ ของจุฬาลงกรณ์มหาวิทยาลัย จำนวน ๑,๘๗๒ คน ซึ่งได้มาจากการสุ่มตัวอย่าง

#### ๒. เครื่องมือในการวิจัย

เครื่องมือในการวิจัย ได้แก่ แบบสอบวัดความสามารถทั่วไปในการอ่าน เข้าใจความภาษาอังกฤษแบบเลือกตอบชนิด ๔ ตัวเลือก ยาว ๕๐ ข้อ ประกอบด้วยเรื่องสำหรับอ่าน ๒ เรื่อง คือ เรื่องที่เป็นภาษาอังกฤษทั่วไป และภาษาอังกฤษเฉพาะสาขาวิชา

ซึ่งมีคำศัพท์ที่มีความหมายสำคัญต่อเรื่องทีอ่าน แบบสอบมีจำนวน ๑ ฉบับ แต่ทำให้มีโครงสร้างความรู้ต่างกัน ๒ แบบ คือ ๑) แบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง และ ๒) แบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง

### ๓. การเก็บรวบรวมข้อมูล

สร้างแบบสอบแต่ละแบบให้มี ๒ รูปแบบ (form) คือ ๑) แบบสอบที่เรียงเรื่องที่เป็นภาษาอังกฤษทั่วไปไว้เป็นลำดับแรก และเรื่องที่เป็นภาษาอังกฤษเฉพาะสาขาวิชาไว้เป็นลำดับที่สอง และ ๒) แบบสอบที่เรียงเรื่องที่เป็นภาษาอังกฤษเฉพาะสาขาวิชาไว้เป็นลำดับแรก และเรื่องที่เป็นภาษาอังกฤษทั่วไปเป็นลำดับที่สอง แล้วทำแบบสอบทั้ง ๔ รูปแบบมาจัดเรียงสลับกันให้เป็นระบบ และนำไปสอบกับผลวิจัยจากคณะต่าง ๆ ที่สุ่มมาทำการทดสอบ

### การวิเคราะห์ข้อมูล

๑. คำนวณหาค่าความตรงเชิงพยากรณ์ของแบบสอบทั้ง ๒ แบบ โดยการหาสัมประสิทธิ์สหสัมพันธ์ของคะแนนของแบบสอบที่ให้คะแนนโดยวิธีต่าง ๆ กับคะแนนรวมของคะแนนสอบกลางภาคและปลายภาคของวิชาการอ่าน แล้วจึงทดสอบนัยของความแตกต่างทางสถิติของค่าความตรงเชิงพยากรณ์ด้วยวิธี Olkin's Test of Significance

๒. คำนวณหาค่าความเที่ยงของแบบสอบถึงกลุ่มทั้ง ๒ แบบ คือ แบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง และแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง ที่ให้คะแนนโดยวิธีต่าง ๆ ด้วยวิธี Cronbach's  $\alpha$  แล้วทำการทดสอบนัยของความแตกต่างทางสถิติของค่าความเที่ยงด้วยวิธี Kristof's Test of Significance

๓. คำนวณหาค่าความเที่ยงของแบบสอบถึงเกณฑ์ (ค่าดัชนีความเชื่อถือ) ของแบบสอบทั้ง ๒ แบบ ที่ให้คะแนนโดยวิธีต่าง ๆ ด้วยวิธี  $\phi(\lambda)$  ซึ่งเสนอโดย Brennan และ Kane แล้วทำการศึกษาเปรียบเทียบขนาดของค่าความเที่ยง ณ จุดตัดต่าง ๆ

๔. คำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์ของอันดับที่ของคะแนนสอบของแบบสอบที่คิดค่านำหนักคะแนนโดยวิธีต่าง ๆ ทั้ง ๔ วิธีดังได้กล่าวมาแล้ว กับอันดับที่ของคะแนนของแบบสอบที่ให้คะแนนโดยวิธีประเพณีนิยม ด้วยวิธี Kendall's tau Correlation แล้วยกกำลังสองค่าสัมประสิทธิ์สหสัมพันธ์ เพื่อเปรียบเทียบความคงที่ของอันดับที่ของคะแนน

๕. คำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนของแบบสอบที่ให้คะแนนโดยวิธีต่าง ๆ กับคะแนนของแบบสอบที่ให้คะแนนโดยวิธีประเพณีนิยม ด้วยวิธี Pearson Product-moment Correlation เพื่อศึกษาความสัมพันธ์ซึ่งกันและกันของคะแนนสอบ

๖. ทำการวิเคราะห์ความแปรปรวน ๒ มิติ เพื่อทดสอบอิทธิพลของโครงสร้างความรู้ที่เกิดจากแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง และความรู้สึกคุ้นเคยต่อเนื้อหาสาระของเรื่องที่ย่านมากหรือน้อย ด้วยวิธี  $\chi^2$ -test และ Phi Correlation

๗. ทำการทดสอบความสัมพันธ์ระหว่างความรู้สึกคุ้นเคยมากหรือน้อยต่อเรื่องที่ย่าน กับความรู้สึกว่าเรื่องที่ย่านยากหรือง่าย ด้วยวิธี 2X2 ANOVA

การวิเคราะห์ข้อมูลโดยมากอาศัยเครื่องคอมพิวเตอร์มหภาค และเครื่องคอมพิวเตอร์จุลภาค โดยใช้โปรแกรมสำเร็จรูปของ SPSSX ITEMEX BICAL-3 LOGIST-4 ITERATE และ OLKIN เป็นต้น

### ผลการวิจัย

#### ๑. ค่าความตรงเชิงพยากรณ์ของแบบสอบอิงกลุ่ม

๑.๑ ความตรงเชิงพยากรณ์ของแบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่องที่ให้คะแนนโดยวิธีต่าง ๆ มีค่าระหว่าง .๓๕๔ ถึง .๓๗๑ ค่าความตรงเชิงพยากรณ์ของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่าผลบวกความแปรปรวนร่วม ( $h^2$ ) มีค่าสูงที่สุด และสูงกว่าค่าความตรงเชิงพยากรณ์ของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่านำหนักคะแนนที่เหมาะสม ( $w$ ) ซึ่งมีค่าต่ำที่สุดอย่างมีนัยสำคัญ

๑.๒ ความตรงเชิงพยากรณ์ของแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง ที่ให้คะแนนโดยวิธีต่าง ๆ มีค่าระหว่าง .๓๕๗ ถึง .๓๗๓ ค่าความตรงของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีค่าสูงที่สุด และสูงกว่าค่าความตรงของแบบสอบที่ให้คะแนนโดยวิธีค่าเบต้าแปลง ( $\beta_{\pm}$ ) ซึ่งมีค่าต่ำที่สุดอย่างมีนัยสำคัญ

๑.๓ ค่าความตรงของแบบสอบทั้ง ๒ แบบ ที่ให้คะแนนโดยวิธีต่าง ๆ มีลำดับที่แตกต่างกัน เช่น ค่าความเที่ยงของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีค่าสูงที่สุด สำหรับแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง แต่มีค่าต่ำที่สุดสำหรับแบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง เป็นต้น

## ๒. ค่าความเที่ยงของแบบสอบอิงกลุ่ม

๒.๑ ความเที่ยงของแบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง ที่ให้คะแนนโดยวิธีต่าง ๆ มีค่าระหว่าง .๗๔๘ ถึง .๘๘๒ ค่าความเที่ยงของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีค่าสูงที่สุด และสูงกว่าค่าความเที่ยงของแบบสอบที่ให้คะแนนโดยวิธีอื่น ๆ อย่างมีนัยสำคัญ ค่าความเที่ยงของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่าเบต้าแปลง ( $\beta_{\pm}$ ) มีค่าต่ำที่สุด

๒.๒ ความเที่ยงของแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง ที่ให้คะแนนโดยวิธีต่าง ๆ มีค่าระหว่าง .๘๐๓ ถึง .๘๘๔ ค่าความเที่ยงของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีค่าสูงที่สุด และสูงกว่าค่าความเที่ยงของแบบสอบที่ให้คะแนนโดยวิธีอื่น ๆ อย่างมีนัยสำคัญ ค่าความเที่ยงของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่าเบต้าแปลง ( $\beta_{\pm}$ ) มีค่าต่ำที่สุด

๒.๓ ลำดับที่ของค่าความเที่ยงของแบบสอบที่ให้คะแนนโดยวิธีต่าง ๆ ของแบบสอบทั้ง ๒ แบบเหมือนกัน และกระสวนของขนาดของค่าดังกล่าวสอดคล้องกันมาก

## ๓. ค่าความเที่ยงของแบบสอบอิงเกณฑ์ (ค่าดัชนีความ เชื่อถือ)

๓.๑ ความเที่ยงโดยเฉลี่ยของแบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง สำหรับจุดศักระหว่างร้อยละ ๐ ถึง ๑๐๐ มีค่าระหว่าง .๘๓๒ ถึง .๘๓๔ โดยมีค่าความเที่ยง

ของแบบสอบที่ให้คะแนนโดยวิธีประเพณีนิยม มีค่าสูงที่สุด และค่าความเที่ยงของแบบสอบ  
 ที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีค่าต่ำที่สุด แต่ถ้าจุดตัดระหว่าง  
 ร้อยละ ๔๐ ถึง ๑๐๐ ค่าความเที่ยงโดยเฉลี่ยมีค่าระหว่าง .๘๔๖ ถึง .๙๓๕ ค่าความ  
 เที่ยงของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่าเบต้าแปลง ( $\beta_c$ ) มีค่าสูงที่สุด และค่าความ  
 เที่ยงของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) ยังมีค่าต่ำที่สุด  
 เช่นเดิม

๓.๒ ความเที่ยงโดยเฉลี่ยของแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง  
 สำหรับจุดตัดระหว่างร้อยละ ๐ ถึง ๑๐๐ มีค่าระหว่าง .๙๐๘ ถึง .๙๓๗ โดยมีค่าความ  
 เที่ยงของแบบสอบที่ให้คะแนนโดยวิธีประเพณีนิยม มีค่าสูงที่สุด และค่าความเที่ยงของแบบ  
 สอบที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีค่าต่ำที่สุด แต่ถ้าจุดตัดระหว่าง  
 ร้อยละ ๕๐ ถึง ๑๐๐ ค่าความเที่ยงโดยเฉลี่ยมีค่าระหว่าง .๘๖๑ ถึง .๙๓๕ ความเที่ยง  
 ของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่าเบต้าแปลง ( $\beta_c$ ) มีค่าสูงที่สุด และความเที่ยงของแบบ  
 สอบที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีค่าต่ำที่สุดเช่นเดิม

๓.๓ ลำดับที่ของค่าความเที่ยงของแบบสอบที่ให้คะแนนโดยวิธีต่าง ๆ ของ  
 แบบสอบทั้ง ๒ แบบเหมือนกันทั้ง ๒ กรณี และกระสวนของขนาดและอันดับที่ของค่าดังกล่าว  
 สอดคล้องกันมากทั้ง ๒ กรณี

#### ๔. ความคงที่ของอันดับที่ของคะแนน

๔.๑ ความคงที่ของอันดับที่ของคะแนนสอบของแบบสอบที่ไม่มีชื่อเรื่องและ  
 ภาพประกอบเรื่อง มีค่าระหว่างร้อยละ ๘๐ ถึง ๙๐ อันดับที่ของคะแนนของแบบสอบที่ให้  
 คะแนนโดยวิธีใช้ค่าผลบวกความแปรปรวนร่วม ( $h^2$ ) มีความคงที่มากที่สุด แต่ของวิธีใช้ค่า  
 น้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีค่าความคงที่ต่ำที่สุด

๔.๒ ความคงที่ของอันดับที่ของคะแนนสอบของแบบสอบที่มีชื่อเรื่องและภาพ  
 ประกอบเรื่อง มีค่าระหว่างร้อยละ ๘๑ ถึง ๙๐ อันดับที่ของคะแนนของแบบสอบที่ให้คะแนน  
 โดยวิธีใช้ค่าผลบวกความแปรปรวนร่วม ( $h^2$ ) มีความคงที่สูงที่สุด แต่ของวิธีใช้ค่าเบต้าแปลง  
 ( $\beta_c$ ) มีค่าต่ำที่สุด

๔.๓ ภาระสวนของขนาดและอันดับที่ของความคงที่ของอันดับที่ของคะแนนของแบบสอบที่ให้คะแนนโดยวิธีต่าง ๆ ของแบบสอบทั้ง ๒ แบบ มีลักษณะใกล้เคียงกัน อันดับของร้อยละของความคงที่สลับกันเพียง เล็กน้อย

#### ๕. ความสัมพันธ์ของคะแนน

๕.๑ คะแนนของแบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง ที่ให้คะแนนโดยวิธีต่าง ๆ มีความสัมพันธ์ซึ่งกันและกันอย่างมีนัยสำคัญ และมีค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง .๙๖๕ ถึง .๙๘๘ คะแนนของแบบสอบที่ให้คะแนนโดยวิธีประเพณีนิยมสัมพันธ์กับคะแนนของแบบสอบที่ให้คะแนนโดยวิธีอื่น ๆ ระหว่าง .๙๘๑ ถึง .๙๘๕ โดยสัมพันธ์กับคะแนนของวิธีใช้ค่าผลบวกความแปรปรวนร่วม ( $h^2$ ) สูงที่สุด และกับคะแนนของวิธีใช้ค่าเบต้าแปลง ( $\beta_c$ ) ต่ำที่สุด

๕.๒ คะแนนของแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง ที่ให้คะแนนโดยวิธีต่าง ๆ มีความสัมพันธ์ซึ่งกันและกันอย่างมีนัยสำคัญ และมีค่าระหว่าง .๙๖๘ ถึง .๙๘๘ คะแนนของแบบสอบที่ให้คะแนนโดยวิธีประเพณีนิยมสัมพันธ์กับคะแนนของแบบสอบที่ให้คะแนนโดยวิธีอื่น ๆ ระหว่าง .๙๘๑ ถึง .๙๘๕ โดยสัมพันธ์กับคะแนนของวิธีใช้ค่าผลบวกความแปรปรวนร่วม ( $h^2$ ) สูงที่สุด และกับคะแนนของวิธีใช้ค่าเบต้าแปลง ( $\beta_c$ ) ต่ำที่สุด

๕.๓ ภาระสวนของขนาดและอันดับที่ของค่าสัมประสิทธิ์สหสัมพันธ์ต่าง ๆ ของแบบสอบทั้ง ๒ แบบ มีความใกล้เคียงกันมากที่สุด

#### ๖. อิทธิพลของโครงสร้างความรู้

๖.๑ โครงสร้างความรู้ที่เกิดจากชื่อเรื่องและภาพประกอบเรื่อง มีผลต่อการอ่านเข้าใจความภาษาอังกฤษไม่แตกต่างจากการที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง อย่างมีนัยสำคัญ โดยเฉลี่ยแล้วผู้สอบแบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง สอบได้คะแนนไม่แตกต่างจากผู้สอบที่สอบแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง อย่างมีนัยสำคัญ

๖.๒ โครงสร้างความรู้ที่เกิดจากความคุ้นเคยต่อสาระของเรื่องที้อ่านมากกว่าหรือน้อยกว่า มีผลต่อการอ่านเข้าใจความภาษาอังกฤษอย่างมีนัยสำคัญ โดยเฉพาะแล้ว ผู้สอบที่อ่านเรื่องที้อ่านคุ้นเคยมากกว่าสอบได้คะแนนมากกว่าคะแนนของผู้อ่านเรื่องที้อ่านคุ้นเคยน้อยกว่าหรือไม่คุ้นเคยเลยอย่างมีนัยสำคัญ

๖.๓ ปฏิกริยาร่วมระหว่างโครงสร้างความรู้ที่เกิดจากแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง กับโครงสร้างความรู้ที่เกิดจากความคุ้นเคยต่อเรื่องที้อ่าน ไม่มีผลอย่างมีนัยสำคัญ

#### ๗. ความสัมพันธ์ระหว่างความคุ้นเคยต่อเรื่องที้อ่านกับความยากง่ายของเรื่องที้อ่าน

ความคุ้นเคยต่อสาระของเรื่องที้อ่านกับความยากง่ายของเรื่องที้อ่าน มีความสัมพันธ์กันอย่างมีนัยสำคัญ โดยมากผู้ที่คุ้นเคยกับเนื้อหาสาระของเรื่องที้อ่านจะรู้สึกว่เรื่องที้อ่านนั้นง่าย แต่ผู้ที่ไม่คุ้นเคยกับเรื่องที้อ่าน หรือคุ้นเคยน้อยกว่าจะรู้สึกว่เรื่องที้อ่านนั้นยาก แม้ว่าโดยเฉลี่ยแล้ว เรื่องที้อ่านนั้นมีระดับความยากง่ายไม่แตกต่างกันอย่างมีนัยสำคัญ

#### อภิปรายผลการวิจัย

##### ๑. ค่าความตรงเชิงพยากรณ์ของแบบสอบอิงกลุ่ม

ผลการวิจัยเกี่ยวกับค่าความตรงเชิงพยากรณ์ของแบบสอบอิงกลุ่มครั้งนี้ ชัดแย้งกับสมมุติฐานในการวิจัยข้อที่ ๑ กล่าวคือ สมมุติฐานกล่าวว่า ในกรณีการสอบแบบอิงกลุ่ม แบบสอบทีให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนทีเหมาะสม (w) มีค่าความตรงเชิงพยากรณ์สูงกว่าค่าความตรงของแบบสอบทีให้คะแนนโดยวิธีอื่น รวมทั้งวิธีประเพณีนิยม อย่างมีนัยสำคัญ ไม่ว่าแบบสอบนั้นจะมีโครงสร้างความรู้แบบใดก็ตาม แต่จากการวิเคราะห์ข้อมูลพบว่า ในกรณีแบบสอบทีมีชื่อเรื่องและภาพประกอบเรื่องเท่านั้น ทีค่าความตรงเชิงพยากรณ์ของแบบสอบทีให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนทีเหมาะสม (w) มีค่าสูงทีสุด แต่ไม่แตกต่างจาก



ค่าความตรงเชิงพยากรณ์ของแบบสอบที่ให้คะแนนโดยวิธีอื่น ๆ อย่างมีนัยสำคัญ ยกเว้นค่าความตรงเชิงพยากรณ์จากการให้คะแนนโดยวิธีใช้ค่าเบต้าแปลง ( $\beta_c$ ) เท่านั้น แต่สำหรับแบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง ค่าความตรงเชิงพยากรณ์ของแบบสอบที่คิดค่าน้ำหนักคะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) กลับมีค่าต่ำที่สุด

การที่ผลเกิดขึ้นดังกล่าวแล้วข้างต้น อาจมีสาเหตุดังต่อไปนี้

ก. ความสัมพันธ์ซึ่งกันและกันของข้อสอบ

เนื่องจากคะแนนของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนัก

คะแนนที่เหมาะสม ( $w$ ) มีการกระจายมากที่สุด [ดูตารางที่ ๒๒ ภาคผนวก ข.] จึงเชื่อว่าเป็นสาเหตุหนึ่งที่ทำให้แบบสอบมีค่าความเที่ยงสูงที่สุด และแตกต่างจากค่าความเที่ยงที่ได้จากแบบสอบที่ให้คะแนนวิธีประเพณีนิยม และวิธีอื่น ๆ อย่างมีนัยสำคัญ แต่การให้คะแนนดังกล่าวกลับทำให้แบบสอบมีค่าความตรงเชิงพยากรณ์น้อยที่สุด และน้อยกว่าค่าความตรงเชิงพยากรณ์ที่ได้จากการให้คะแนนวิธีประเพณีนิยมอย่างมีนัยสำคัญ ข้อค้นพบนี้สอดคล้องกับผลการวิจัยของ Patniak และ Traub (Patniak and Traub, ๑๙๗๓:๒๘๑) ซึ่งใช้ค่ามาตรวัดจากวิธีการเปรียบเทียบรายคู่วิธีแบบ Thurstone เป็นคะแนนรายข้อของแบบสอบแล้วพบว่า แบบสอบมีค่าความเที่ยงเพิ่มขึ้นอย่างมีนัยสำคัญ แต่มีค่าความตรงลดลงอย่างมีนัยสำคัญ ผู้วิจัยทั้งสองเชื่อว่า เป็นเพราะว่าการให้คะแนนรายข้อต่างกัน ทำให้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างข้อสอบแต่ละข้อ (inter-item correlation) มีมากเพิ่มขึ้นกว่าการให้คะแนนวิธีประเพณีนิยม ความสัมพันธ์ดังกล่าวยิ่งมากขึ้น ก็ยิ่งทำให้ความสัมพันธ์ของคะแนนรวมของแบบสอบกับเกณฑ์ยิ่งต่ำลง ผู้วิจัยมีความเห็นว่า แนวคิดดังกล่าวมีเหตุผลและถูกต้องตามทฤษฎีการวัดและแนวคิดของความสัมพันธ์ระหว่างตัวแปร กล่าวคือ การที่ข้อสอบรายข้อมีคะแนนขนาดแตกต่างกันมากขึ้นกว่าคะแนนที่คิดวิธีประเพณีนิยม ทำให้ผลการสอบมีคะแนนการกระจายมากยิ่งขึ้น ดังตั้งกับว่าคะแนนมาจากกลุ่มตัวอย่างวิธีวิธีพันซ์ เมื่อคะแนนยิ่งมีการกระจายมาก ก็ยิ่งทำให้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างข้อสอบมีมากขึ้นกว่าเดิม และทำให้แบบสอบมีค่าความเที่ยงเพิ่มขึ้นจากเดิม (Mehrens and Lehmann, ๑๙๗๘:๑๐๒)

แต่ค่าสัมประสิทธิ์สหสัมพันธ์ของแบบสอบกับเกณฑ์จะยิ่งต่ำลง เพราะคะแนนตัวแปรหนึ่งมีลักษณะเป็นคะแนนจากกลุ่มตัวอย่างวิธีหนึ่ง คือ คะแนนที่ให้แก่วิธีสอบรายข้อแตกต่างกันกับคะแนนของอีกตัวแปรหนึ่งที่มีลักษณะเป็นคะแนนจากกลุ่มเอกพันธ์ คือ คะแนนที่ใช้เป็นเกณฑ์ ซึ่งเป็นการให้คะแนนวิธีประเพณีนิยม คะแนนของ ๒ ตัวแปรจึงสัมพันธ์กันต่ำ (Glass and Stanley, ๑๙๗๐:๑๑๑) อีกประการหนึ่ง การที่คะแนนข้อสอบรายข้อมีความสัมพันธ์ซึ่งกันและกันเพิ่มมากขึ้นกว่าการให้คะแนนวิธีประเพณีนิยม ค่าเหล่านี้มีลักษณะเหมือนตัวแปรต้นในสมการถดถอยพหุคูณ ย่อมจะมีความสัมพันธ์กับคะแนนเกณฑ์ที่เป็นตัวแปรตามในระดับต่ำกว่าปกติ ข้อสอบยังมีความสัมพันธ์ซึ่งกันและกันมาก ก็จะมีสัมพันธ์กับเกณฑ์ต่ำลงมากเท่านั้น (Patniak and Traub, ๑๙๗๓:๒๘๑; Kerlinger and Pedhazur, ๑๙๗๓:๘๔)

ดังนั้น การให้คะแนนรายข้อวิธีใดที่ทำให้คะแนนมีความแปรปรวนของคะแนนยิ่งมาก (ในกรณีนี้ได้แก่ คะแนนของวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม) ก็จะมีค่าความเที่ยงสูง แต่จะยิ่งมีความสัมพันธ์กับเกณฑ์ต่ำที่สุดด้วย นั่นคือ จะมีค่าความตรงเชิงพยากรณ์ต่ำที่สุด นี้อาจเป็นสาเหตุหนึ่งที่ทำให้การให้คะแนนรายข้อโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม (w) มีค่าความเที่ยงสูงสุด แต่มีค่าความตรงเชิงพยากรณ์ต่ำที่สุดด้วย เมื่อเปรียบเทียบกับวิธีการให้คะแนนรายข้อวิธีอื่น

#### ข. ลักษณะของการกระจายของคะแนนที่ใช้เป็น เกณฑ์

นอกจากนี้ ปัญหาที่น่าสนใจอีกประการหนึ่งก็คือ เหตุใดค่าความตรงเชิงพยากรณ์ที่คำนวณจากคะแนนโดยใช้วิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม (w) จึงมีค่าสูงที่สุด สำหรับแบบสอบที่มีข้อเรื่องและภาพประกอบเรื่อง แต่มีค่าต่ำที่สุดสำหรับแบบสอบที่ไม่มีข้อเรื่องและภาพประกอบเรื่อง ทั้งที่คะแนนของวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม (w) มีค่าความเที่ยงสูงที่สุดสำหรับแบบสอบทั้ง ๒ ฉบับ ผู้วิจัยคิดว่า การที่คะแนนของวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม (w) สัมพันธ์กับคะแนนเกณฑ์ (ผลรวมของคะแนนสอบการอ่านเข้าใจความประจักษ์กลางภาคและปลายภาค) สูงมาก เพราะว่าคะแนนที่ให้โดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม (w) ของแบบสอบที่มีข้อเรื่องและภาพประกอบเรื่อง มีการกระจายมากที่สุด

[โปรดดูตารางที่ ๒๑ ในภาคผนวก ข.] แต่สำหรับแบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่องนั้น แม้ว่าการกระจายของคะแนนที่ได้โดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม (w) จะมีการกระจายมากที่สุดเช่นเดียวกัน แต่การที่ค่าสัมประสิทธิ์สหสัมพันธ์มีน้อยที่สุด อาจเนื่องมาจากคะแนนสอบที่เป็นเกณฑ์ หรือคะแนนที่ได้จากวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม (w) นี้ จำนวนหนึ่งไม่คล้อยตามกัน รวมทั้งสัมพันธ์กันในทิศทางตรงกันข้าม จึงทำให้ค่าสัมประสิทธิ์สหสัมพันธ์มีค่าต่ำที่สุด แม้ว่าขนาดจะไม่แตกต่างจากค่าความตรงเชิงพยากรณ์ที่คำนวณจากคะแนนโดยวิธีอื่น [รายละเอียดแสดงไว้ในตารางที่ ๒๔-๓๐ ภาคผนวก ข.]

นอกจากนี้เป็นที่น่าสังเกตว่า ความตรงเชิงพยากรณ์ของแบบสอบทั้ง ๒ แบบที่ให้คะแนนโดยวิธีต่าง ๆ มีค่าค่อนข้างต่ำ คือ มีค่าระหว่าง .๓๖๗ - .๓๗๔ เท่านั้น แสดงว่า แบบสอบที่เป็นเครื่องมือเพื่อการวิจัยครั้งนี้ กับแบบสอบที่ใช้เป็นเกณฑ์สามารถวัดสิ่งที่ต้องการวัดรวมกันได้ ประมาณร้อยละ ๑๓-๑๔ เท่านั้น ซึ่งนับว่ามีค่าน้อย แต่สำหรับค่าความตรงเชิงพยากรณ์ของแบบสอบแล้ว ถือว่าเป็นค่าที่สูงมากพอสมควร ทั้งนี้เพราะว่า (Cronbach, ๑๙๔๘:๑๓๓-๑๓๕)

๑. ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร ๒ ตัว ที่ลุ่มมามากมีค่าไม่คงที่ ถ้ากลุ่มตัวอย่างยิ่งมีจำนวนมาก ค่าดังกล่าวก็ยิ่งไม่เป็นอิสระ และมีแนวโน้มที่จะมีค่าน้อยลง

๒. ค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปรของกลุ่มตัวอย่างที่คัดสรรมาแล้ว (เช่น ในกรณีของนิสิตชั้นปีที่ ๑ ซึ่งผ่านการสอบคัดเลือกเข้ามหาวิทยาลัยมาแล้ว) มักจะมีค่าต่ำ เพราะกลุ่มตัวอย่างมีลักษณะเอกพันธ์

๓. ตามปกติแล้ว ค่าความตรงเชิงพยากรณ์ของแบบสอบมาตรฐานที่มีชื่อเสียงทั่วไป เช่น Short Employment Test และ Iowa Test of Educational Development จะมีค่าต่ำ คือ ประมาณ ๐.๐๘ - ๐.๕๕ และโดยทั่วไปมักไม่มีแบบสอบใดที่จะมีค่าความตรงเชิงพยากรณ์มากกว่า ๐.๖๐

## ๒. ค่าความเที่ยงของแบบสอบอิงกลุ่ม

การวิจัยครั้งนี้พบว่า ผลการวิเคราะห์ข้อมูลสอดคล้องกับสมมุติฐานในการวิจัยข้อที่ ๑ กล่าวคือ ในกรณีการสอบแบบอิงกลุ่ม แบบสอบที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีค่าสูงกว่าค่าความเที่ยงของแบบสอบที่ให้คะแนนโดยวิธีอื่น ๆ รวมทั้งวิธีประเพณีนิยม อย่างมีนัยสำคัญ ทั้งแบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง และแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง

การที่ผลเกิดขึ้นดังกล่าวแล้วข้างต้น อาจมีสาเหตุดังต่อไปนี้

ก. คะแนนที่คิดจากการให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีการกระจายมากกว่าคะแนนที่คิดจากวิธีอื่น ๆ

ในการคำนวณค่าความเที่ยงของแบบสอบอิงกลุ่มครั้งนี้ ผู้วิจัยคำนวณโดยวิธี Cranbach's  $\alpha$  เนื่องจากเป็นวิธีที่ใช้กับข้อสอบที่คิดคะแนนแบบ ๐-๑ (วิธีประเพณีนิยม) หรือแบบอื่นก็ได้ (Stanley, ๑๙๗๑:๔๑๒) จึงเหมาะสมสำหรับลักษณะของข้อมูลในการวิจัยครั้งนี้ สูตรทั่วไปในการคำนวณค่าความเที่ยงดังกล่าว คือ (Guilford, ๑๙๕๕:๓๘๕)

$$\alpha = \frac{n}{n-1} \left[ 1 - \frac{\sum v_i^2}{v_t} \right]$$

ในเมื่อ  $n$  = จำนวนข้อสอบ

$v_i$  = ความแปรปรวนของข้อสอบข้อที่  $i$

$v_t$  = ความแปรปรวนของแบบสอบทั้งฉบับ

จากสูตรข้างต้นจะสังเกตได้ว่า ค่าความเที่ยงขึ้นอยู่กับค่าความแปรปรวนของแบบสอบทั้งฉบับ และความแปรปรวนของแบบสอบดังกล่าวนี้ก็ขึ้นอยู่กับค่าความแปรปรวนของข้อสอบรายข้อ คะแนนของผู้สอบยิ่งแตกต่างกันมาก ก็ยิ่งทำให้ความแปรปรวนรายข้อ และความ



แปรปรวนรวมยิ่งมากขึ้น และจะทำให้ค่าความเที่ยงมีมากเพิ่มขึ้น (Gronlund, ๑๙๗๖ : ๑๑๘) ดังนั้น ในกรณีที่ความยาวของแบบสอบคงที่ วิธีคิดคะแนนวิธีใดที่ทำให้อัตราส่วนของผลบวกของค่าความแปรปรวนรายข้อ กับความแปรปรวนรวมทั้งฉบับมีค่าน้อย (เพราะความแปรปรวนทั้งฉบับมีค่ามากกว่า) ก็จะมีค่าความเที่ยงมากกว่าวิธีอื่นที่ทำให้เกิดค่าดังกล่าวมาก (เพราะความแปรปรวนทั้งฉบับมีค่าน้อย)

จากการศึกษาค่าความแปรปรวนที่ได้จากการคิดค่าน้ำหนักคะแนนวิธีต่างกัน ปรากฏผลว่า แบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง และแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง เมื่อให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) ทำให้คะแนนมีสัมประสิทธิ์ของการกระจายมากกว่าคะแนนที่คิดจากวิธีอื่น ๆ จึงน่าจะ เป็นสาเหตุที่ทำให้ความเที่ยงที่คิดจากคะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีค่าสูงที่สุด [รายละเอียดแสดงไว้ในตารางที่ ๒๑ ภาคผนวก ข.]

แต่อย่างไรก็ตาม ค่าสัมประสิทธิ์การกระจายบอกถึงขนาดของการกระจายของค่าความแปรปรวนรวมทั้งหมดของแบบสอบทั้งฉบับเท่านั้น ไม่ได้บอกขนาดของการกระจายของความแปรปรวนรายข้อของแบบสอบ ดังนั้น ค่าดังกล่าวจึงเป็นเพียงดัชนีที่บอกขนาดเปรียบเทียบของค่าความเที่ยงได้อย่างหยاب ๆ เท่านั้น เพราะเป็นที่น่าสังเกตว่า แม้ว่าสัมประสิทธิ์ของการกระจายของคะแนนที่คิดจากวิธีใช้ค่าเบต้าแปลง ( $\beta_c$ ) มีขนาดมากกว่าค่าสัมประสิทธิ์ของการกระจายที่ได้จากการคิดคะแนนวิธีประเพณีนิยม แต่ค่าความเที่ยงที่คิดจากคะแนนโดยวิธีใช้ค่าเบต้าแปลง ( $\beta_c$ ) ก็มีค่าน้อยกว่าค่าความเที่ยงที่คิดจากคะแนนวิธีประเพณีนิยม เป็นต้น (ค่าความเที่ยงของแบบสอบอิงกลุ่ม ผู้วิจัยได้ตรวจสอบจากหลายแหล่ง เช่น จากการใช้โปรแกรม SPSSX จากการวิเคราะห์ข้อสอบวิธีประเพณีนิยม จากการวิเคราะห์ข้อสอบตามรูปแบบบราซซ์ และรูปแบบโลจิสติกส์ใช้พารามิเตอร์ ๓ ตัว และจากวิธีการคำนวณหาความเที่ยงแบบ Hoyt เพื่อทดสอบความแตกต่างของค่าความเที่ยงแต่ละวิธีให้ค่าใกล้เคียงกันมาก) ทั้งนี้เพราะเราไม่ทราบว่า ผลบวกของความแปรปรวนรายข้อจากการคิดคะแนนแต่ละวิธีมีขนาดเท่าใด

อนึ่ง ปัญหาที่ต้องพิจารณาต่อไปก็คือ เหตุใดการให้คะแนนโดยวิธีอื่น โดยมากจึงทำให้แบบสอบถามมีความแปรปรวนมากกว่าการคิดคะแนนวิธีประเพณีนิยม ผู้วิจัยคิดว่าผลดังกล่าวเกิดจากคะแนนรายข้อที่คิดโดยวิธีอื่น ๆ นั้น มีขนาดแตกต่างกันมากกว่าคะแนนรายข้อที่คิดวิธีประเพณีนิยม ซึ่งมีเพียง ๐ และ ๑ กล่าวคือ โดยทฤษฎีแล้ว คำนวณน้ำหนักคะแนนที่คิดโดยวิธีใช้ค่าเคลต้า ( $\Delta$ ) มีพิสัยระหว่าง ๑ ถึง ๒๕ คะแนนที่คิดโดยวิธีใช้ค่าเบต้าแปลง ( $\beta_t$ ) มีพิสัยระหว่าง ๓ ถึง ๒๓ คะแนนที่คิดโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีพิสัยระหว่าง .๐๐ ถึง ๑.๕ และคะแนนที่คิดโดยวิธีใช้ค่าผลบวกความแปรปรวนร่วม ( $h^2$ ) มีพิสัยระหว่าง .๐๐ ถึง ๑.๐๐ ดังนั้น ตามทฤษฎีแล้ว ความแปรปรวนของคะแนนที่คิดโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) โดยวิธีใช้ค่าผลบวกความแปรปรวนร่วม ( $h^2$ ) โดยวิธีใช้ค่าเคลต้า ( $\Delta$ ) โดยวิธีใช้ค่าเบต้าแปลง ( $\beta_t$ ) และโดยวิธีประเพณีนิยม จะมีมากเป็นอันดับที่ ๑ ๒ ๓ ๔ และ ๕ ตามลำดับ ซึ่งจะเห็นได้จากสัมประสิทธิ์ของการกระจายในตารางที่ ๒ ในทางปฏิบัติอันดับที่ตั้งกล่าวโดยมากสอดคล้องกับแนวคิดทางทฤษฎี ด้วยเหตุนี้คะแนนที่คิดโดยวิธีประเพณีนิยมจึงมีความแปรปรวนรวมน้อยกว่าคะแนนของวิธีอื่นโดยมาก ซึ่งดัชนีหยาบ ๆ อย่างหนึ่ง อาจสังเกตได้จากคะแนนรายข้อของการคิดคะแนนแต่ละวิธี ซึ่งมีสัมประสิทธิ์ของการกระจายของค่าน้ำหนักคะแนนมากกว่าการกระจายของคะแนนของวิธีประเพณีนิยม มากถึงประมาณ ๒๐-๖๓ เท่า [รายละเอียดแสดงไว้ในตารางที่ ๒๒ ภาคผนวก ข.] ดังนั้น เมื่อการคิดคะแนนแต่ละวิธีใช้คะแนนต่าง ๆ ดังกล่าวแล้ว จึงมีแนวโน้มที่จะทำให้ความแปรปรวนของข้อสอบรายข้อ หรือความแปรปรวนรวมมีขนาดมากกว่าค่าที่ได้จากการให้คะแนนวิธีประเพณีนิยม ส่วนค่าความเที่ยงของแบบสอบที่คำนวณจากคะแนนของวิธีต่าง ๆ ไม่มีลำดับที่สอดคล้องกับขนาดของค่าความแปรปรวน อาจเป็นเพราะค่าเหล่านี้เป็นเพียงดัชนีหยาบ ๆ เท่านั้น และคงทำให้อัตราส่วนของผลรวมของความแปรปรวนรายข้อ กับความแปรปรวนรวมแตกต่างกันไปจากลำดับเดิมได้

ข. คะแนนที่ได้จากวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) เป็นคะแนนที่ปรับเอาอิทธิพลของการเดาออกแล้ว

ตามทฤษฎีการวัดนั้น การคิดค่าความเที่ยงทั่วไปอาจคำนวณได้จากแนวคิดดังนี้  
(Magnuson, ๑๙๖๖:๗๕)

$$r_{tt} = 1 - \sigma_e^2 / \sigma_t^2$$

จากสมการดังกล่าวนี้จะเห็นได้ชัดเจนยิ่งขึ้นว่า ค่าความเที่ยงของแบบสอบยิ่งกลุ่ม ขึ้นอยู่กับค่าความแปรปรวนทั้งหมดของแบบสอบ ( $\sigma_t^2$ ) และค่าความแปรปรวนของความคลาดเคลื่อน ( $\sigma_e^2$ ) ดังนั้น ในกรณีที่แบบสอบที่ให้คะแนนแต่ละวิธีแล้ว มีค่าความแปรปรวนรวมทั้งหมดขนาดเท่า ๆ กัน แต่ค่าความเที่ยงมีขนาดแตกต่างกัน ย่อมแสดงว่า การให้คะแนนแต่ละวิธีก่อให้เกิดค่าความแปรปรวนของความคลาดเคลื่อน (error variance;  $\sigma_e^2$ ) แตกต่างกัน ถ้าการให้คะแนนวิธีใดก่อให้เกิดความคลาดเคลื่อนน้อยกว่า ย่อมจะมีค่าความเที่ยงมากกว่าวิธีที่ก่อให้เกิดความคลาดเคลื่อนมากกว่า และแหล่งที่ทำให้เกิดความคลาดเคลื่อนในการวัดที่สำคัญ ได้แก่ การบริหารการสอบ การคิดคะแนน และการเดา

สำหรับการให้คะแนนโดยวิธีต่าง ๆ กันนั้น อาศัยของสอบแบบปรนัยฉบับเดียวกัน แต่การคิดคะแนนแตกต่างกัน ดังนั้น แหล่งความคลาดเคลื่อนในการวัดที่เกิดจากการบริหารการสอบ และการคิดคะแนน จึงเชื่อว่าจะมีความคล้ายคลึงกัน และมีขนาดไม่แตกต่างกัน แต่เป็นที่น่าสังเกตว่า การคิดคะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) นั้น เป็นการให้คะแนนที่ละเอียดมาก และคำนึงถึงการเดาด้วย คือ คิดค่าอำนาจจำแนก ( $a$ ) ค่าความยากง่าย ( $p$ ) และค่าการเดา ( $c$ ) ของข้อสอบรายข้อ ดังนั้น คะแนนรายข้อเหล่านี้จะละเอียดมาก มีขนาดแตกต่างกันมาก และมีความคลาดเคลื่อนในการวัดน้อยกว่าการให้คะแนนโดยวิธีอื่นที่ไม่ได้คำนึงถึงความคลาดเคลื่อนในการวัดที่เกิดจากการเดาเลย นี่อาจเป็นสาเหตุหนึ่งที่ทำให้ค่าความเที่ยงของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีค่าสูงสุด ส่วนการให้คะแนนโดยวิธีใช้ค่าเตลต้า ( $\Delta$ ) และวิธีใช้ค่าเบต้าแปลง ( $\beta_t$ ) อาจทำให้เกิดความคลาดเคลื่อนในการวัด ซึ่งเกิดจากการเดามากกว่าการให้คะแนนโดยวิธีประเพณีนิยม และการให้คะแนนโดยวิธีใช้ค่าผลบวกความแปรปรวนร่วม ( $h^2$ ) นั้น คงทำให้ความคลาดเคลื่อนดังกล่าวมีมากเท่า ๆ กับการให้คะแนนวิธีประเพณีนิยม จึงทำให้มีค่าความเที่ยงขนาดเท่า ๆ กัน

อนึ่ง ผลการวิจัยเกี่ยวกับขนาดของค่าความเที่ยงของแบบสอบที่ให้คะแนนรายข้อต่างกัน มีค่าสูงกว่าค่าความเที่ยงที่ของคะแนนวิธีประเพณีนิยมนี้ ส่วนหนึ่งสอดคล้องกับผลการวิจัยของนักวิจัยบางท่าน เช่น ผลการวิจัยของ Davis และ Fifer ในปี ๑๙๕๔ และของ Henderson ในปี ๑๙๗๐ (อ้างจาก Wang and Stanley, ๑๙๗๐:๖๙๒-๓) การวิจัยของ Patnaik และ Traub ในปี ๑๙๗๓ (Patnaik and Traub, ๑๙๗๓: ๒๘๑) การวิจัยของ Diamond และ Raffeld ในปี ๑๙๗๕ (Diamond, ๑๙๗๕:๑๒๔; Raffeld, ๑๙๗๕:๒๘๕) เป็นต้น กล่าวคือ ผลการวิจัยเหล่านี้พบว่า การให้คะแนนรายข้อด้วยวิธีต่าง ๆ เช่น ใช้ค่าสัมประสิทธิ์สหสัมพันธ์ ใช้ค่ามาตรวัดแบบ Guttman ใช้ค่ามาตรวัดจากวิธีการเปรียบเทียบรายคู่แบบ Thurstone และใช้ค่าคะแนนตามความเชื่อมั่นของการตอบ เป็นต้น ต่างก็ทำให้แบบสอบมีค่าความเที่ยงสูงกว่าค่าความเที่ยงที่คำนวณจากคะแนนวิธีประเพณีนิยม

สำหรับการวิจัยครั้งนี้พบว่า ถ้าต้องการให้แบบสอบที่ให้คะแนนวิธีประเพณีนิยมมีค่าความเที่ยงเพิ่มขึ้นจากเดิม จาก .๘๕๖ (ค่าความเที่ยงของแบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง) และ .๘๖๒ (ค่าความเที่ยงของแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง) เป็น .๘๘๒ และ .๘๘๙ (ค่าความเที่ยงจากการให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม (w) ของแบบสอบทั้ง ๒ แบบ) แล้ว จะต้องเพิ่มความยาวของข้อสอบที่มีลักษณะคล้ายคลึงกับข้อสอบที่มีอยู่อีกประมาณ ๑๓-๑๕ ข้อ (Gulliksen, ๑๙๕๐:๘๓) จึงอาจนับได้ว่า การให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม (w) เป็นคะแนนรายข้อนั้น เป็นประโยชน์ต่อการทดสอบมาก เพราะทำให้ไม่ต้องสร้างข้อสอบเพิ่มขึ้นอีกประมาณ ๐.๓ เท่า ของความยาวของแบบสอบที่มีอยู่เดิม การที่เราสามารถลดจำนวนข้อสอบลงแล้ว ค่าความเที่ยงยิ่งเพิ่มขึ้นนี้ ในการพัฒนาแบบสอบนั้น แม้ว่าจะลดลงได้น้อยเพียงใดก็นับว่ามีประโยชน์และได้ผลคุ้มค่า (Lord, ๑๙๗๕:๑๐)

### ๓. ค่าความเที่ยงของแบบสอบอิงเกณฑ์ (ค่าดัชนีความเชื่อถือ)

การวิจัยครั้งนี้พบว่า ผลการวิเคราะห์ข้อมูลขัดแย้งกับสมมุติฐานการวิจัยข้อที่ ๓ กล่าวคือ สมมุติฐานกล่าวว่า ในกรณีการสอบอิงเกณฑ์ แบบสอบที่ให้คะแนนโดยวิธีใช้



ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีค่าความเที่ยง (ค่าดัชนีความเชื่อถือ) สูงกว่าค่าความเที่ยงของแบบสอบที่ให้คะแนนโดยวิธีอื่น ๆ ไม่ว่าแบบสอบนั้นจะมีโครงสร้างความรู้แบบใดก็ตาม แต่จากการวิจัยพบว่า ค่าความเที่ยงของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีค่าโดยเฉลี่ยแล้วต่ำที่สุด ไม่ว่าจุดตัดจะอยู่ระหว่างร้อยละ ๐ ถึง ๑๐๐ หรือร้อยละ ๕๐ ถึง ๑๐๐ ก็ตาม และทั้งในกรณีแบบสอบไม่มีชื่อเรื่องและภาพประกอบเรื่อง และแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง

การที่ผลเกิดขึ้นดังกล่าวแล้วข้างต้น อาจมีสาเหตุดังต่อไปนี้

ก. ความแปรปรวนของคะแนนมวลของสิ่งที่ยึดมีขนาดใหญ่ที่สุด

ปัจจัยที่สำคัญที่มีอิทธิพลต่อค่าความเที่ยง (ค่าดัชนีความเชื่อถือ) ของแบบสอบอิงเกณฑ์ มี ๒ อย่าง คือ ๑) ขนาดของความแปรปรวนของคะแนนมวลของสิ่งที่ยึด (universe score variance:  $\sigma^2(p)$ ) และ ๒) ค่าความแตกต่างระหว่างค่ามีขั้วเฉลี่ยเลขคณิตของประชากร ( $\mu$ ) กับคะแนนของจุดตัด ( $\lambda$ ) (Brennan, ๑๙๘๐:๒๐๓) ซึ่งจะสังเกตเห็นได้จากสูตรต่อไปนี้

$$\Phi(\lambda) = \frac{\sigma^2(p) + (\mu - \lambda)^2}{\sigma^2(p) + (\mu - \lambda)^2 + \sigma^2(\Delta)}$$

การศึกษาค้างนี้ ได้คำนวณหาค่าความแปรปรวนของคะแนนมวลของสิ่งที่ยึด ( $\sigma^2(p)$ ) ด้วย [โปรดดูตารางที่ ๓๑-๓๒ ภาคผนวก ข.] แต่เนื่องจากคะแนนของการคิดคะแนนแต่ละวิธีแตกต่างกัน จึงทำให้การเปรียบเทียบค่าดังกล่าวไม่ชัดเจน แต่จากการศึกษาค่าสัมประสิทธิ์การกระจายของคะแนน (ตารางที่ ๒๑) ทำให้ทราบได้ว่า คะแนนของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) มีการกระจายมากที่สุด ข้อมูลดังกล่าวนี้อาจเป็นปัจจัยเบื้องต้นที่ทำให้แบบสอบที่ให้คะแนนด้วยวิธีดังกล่าวมีแนวโน้มว่าจะมีค่าความเที่ยงน้อยที่สุด

ข. ความแตกต่างระหว่างมีชดิม เลขคณิตของประชากรกับคะแนนจุดตัดมีน้อย

การวิจัยครั้งนี้พบว่า ค่ามีชดิม เลขคณิตของประชากรโดยประมาณ ( $x_{pi}$ ) ของแบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง และแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง มีอันดับที่ของค่ามีชดิม เลขคณิตของประชากรของแบบสอบที่ให้คะแนนโดยวิธีต่าง ๆ ของแบบสอบทั้ง ๒ แบบเหมือนกัน และมีกระสวนของขนาดสอดคล้องกัน ค่ามีชดิม เลขคณิตของประชากรโดยประมาณของแบบสอบที่ให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม (w) มีค่าต่ำที่สุด [รายละเอียดแสดงไว้ในตารางที่ ๓๕ ภาคผนวก ข.]

ดังนั้น เมื่อจุดตัดของคะแนนของแบบสอบที่ให้คะแนนโดยวิธีต่าง ๆ มีค่าเท่ากัน แบบสอบที่มีการกระจายของคะแนนมวลของสิ่งที่มีขนาดใหญ่มากที่สุด แต่มีความแตกต่างระหว่างมีชดิม เลขคณิตของประชากรกับคะแนนจุดตัดน้อยที่สุด ก็น่าจะมีค่าความเที่ยง (ค่าดัชนีความ เชื่อถือ) น้อยที่สุดด้วย

อนึ่ง เป็นที่น่าสนใจ เกิดว่า เมื่อจุดตัดมีค่าระหว่างร้อยละ ๕๐ ถึง ๑๐๐ ค่าความเที่ยงของแบบสอบถึง เกณฑ์ ที่ให้คะแนนโดยวิธีต่าง ๆ (ตารางที่ ๑๐) มีลำดับที่ของขนาดเกือบจะตรงกันข้ามกับค่าความเที่ยงของแบบสอบถึงกลุ่ม (ตารางที่ ๖) ทุก ๆ ค่า

#### ๔. ความคงที่ของอันดับที่ของคะแนน

ผลของการวิจัยครั้งนี้พบว่า ผลการวิเคราะห์ข้อมูลสอดคล้องกับ สมมุติฐานในการวิจัยข้อที่ ๔ ที่กล่าวว่า "การให้คะแนนต่างกันแต่ละวิธี ทำให้อันดับที่ของคะแนนการอ่าน เข้าใจความภาษาอังกฤษเปลี่ยนไป" กล่าวคือ ผลการวิจัยพบว่า การให้คะแนนโดยวิธีต่าง ๆ ทำให้อันดับที่ของคะแนนคงที่ประมาณร้อยละ ๘๐ ถึง ๙๐ เท่านั้น ซึ่งแสดงว่า อันดับที่ของคะแนนจะแตกต่างจากอันดับที่เดิม เมื่อแบบสอบคิดค่าน้ำหนักคะแนนวิธีประเพณีนิยมประมาณร้อยละ ๑๐ ถึง ๒๐

การที่ผล เกิดขึ้นดังกล่าวข้างต้นคง เป็นเพราะว่า คะแนนของแบบสอบที่ให้คะแนนโดยวิธีประเพณีนิยม กับวิธีอื่น ๆ สัมพันธ์กันอย่างไม่สมบูรณ์ เพราะแต่ละวิธีอาศัย

แนวคิดแตกต่างกัน เช่น การให้คะแนนวิธีประเพณีนิยม กำหนดให้คะแนนข้อที่ผู้สอบตอบถูก = ๑ คะแนน และให้คะแนนข้อที่ตอบผิด = ๐ คะแนน วิธีใช้ค่าเดลต้า ( $\Delta$ ) ให้คะแนนข้อที่ผู้สอบตอบถูกมีค่ามากหรือน้อย แล้วแต่ความยากง่ายของข้อสอบนั้น และให้คะแนนข้อที่ตอบผิด = ๐ คะแนน แต่วิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) ให้คะแนนข้อที่ผู้สอบตอบถูกมีค่ามากหรือน้อย แล้วแต่ความสัมพันธ์กันของค่าความยากง่าย ( $p$ ) ค่าอำนาจจำแนก ( $a$ ) และค่าการเดา ( $c$ ) ของข้อสอบว่าจะมีลักษณะอย่างไร เป็นต้น

ด้วยเหตุดังกล่าวแล้ว คะแนนของแบบสอบที่ให้คะแนนต่างกัน ย่อมมีอันดับที่ของคะแนนไม่สมบูรณ์ ดังนั้น ค่ายกกำลังสองของค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนของแบบสอบที่ให้คะแนนวิธีประเพณีนิยม กับวิธีอื่น ๆ ซึ่งเป็นดัชนีแสดงความคงที่ของอันดับที่ของคะแนน จึงน้อยกว่า ๑.๐๐ แสดงว่า อันดับที่ของคะแนนของแบบสอบที่ให้คะแนนโดยวิธีต่าง ๆ เปลี่ยนแปลงไปจากลำดับที่ของคะแนนสอบที่ให้คะแนนวิธีประเพณีนิยม

#### ๕. ความสัมพันธ์ของคะแนน

การวิจัยครั้งนี้พบว่า ความสัมพันธ์ของคะแนนสอบที่ให้คะแนนโดยวิธีต่าง ๆ สอดคล้องกับสมมุติฐานในการวิจัยข้อที่ ๔ ที่กล่าวไว้ว่า "คะแนนที่ได้จากการให้คะแนนต่างกันแต่ละวิธี ต่างมีความสัมพันธ์ซึ่งกันและกันอย่างมีนัยสำคัญ"

การที่ผลการวิจัยสอดคล้องกับสมมุติฐานดังกล่าวแล้ว อาจเนื่องมาจากเหตุผลดังต่อไปนี้

ก. การให้คะแนนข้อสอบบางวิธีมีพื้นฐานแนวคิดบางอย่างร่วมกัน

การให้คะแนนโดยวิธีต่าง ๆ อาจแบ่งออกได้เป็น ๓ พวก ตามแนวคิดที่แตกต่างกัน กล่าวคือ

ก. วิธีให้คะแนนโดยอาศัยค่าความยากง่ายของข้อสอบเป็นปัจจัยหนึ่ง ได้แก่ วิธีใช้ค่าเดลต้า ( $\Delta$ ) วิธีใช้ค่าเบต้าแปลง ( $\beta_c$ ) และวิธีให้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ )

ข. วิธีให้คะแนนโดยอาศัยค่าความสัมพันธ์ของข้อสอบกับสิ่งที่มีง  
ทดสอบ ได้แก่ วิธีใช้ค่าผลบวกความแปรปรวนร่วม ( $h^2$ )

ค. วิธีให้คะแนนโดยวิธีอาศัยความสำเร็จในการตอบข้อสอบแต่  
ละข้อว่าประสบผลหรือไม่ ได้แก่ วิธีประเพณีนิยม

อนึ่ง เมื่อพิจารณาสูตรในการคำนวณหาคะแนนโดยวิธีต่าง ๆ ประกอบด้วย  
(โปรดดูหน้า ๔-๑๐) จะสังเกตได้ว่า สูตรในการหาค่าเฉลี่ย ( $\Delta$ ) และสูตรในการหา  
ค่าเบต้าแปลง ( $\beta_T$ ) มีความสัมพันธ์กันมากที่สุด เพราะแต่ละวิธีต่างอาศัยค่าความยาก  
ง่ายเพียงอย่างเดียวเป็นตัวประกอบที่สำคัญ จึงคาดว่าคะแนนของทั้ง ๒ วิธีนี้ จะมีความ  
สัมพันธ์กันมากที่สุด แต่สูตรในการคำนวณหาค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) นั้น ใช้  
ค่าพารามิเตอร์มากที่สุด ดังนั้น คะแนนรายข้อและคะแนนรวมจะแตกต่างกันไปจากคะแนน  
ของวิธีอื่นมาก จึงคาดว่า จะมีความสัมพันธ์กับคะแนนของวิธีอื่นน้อย

#### ข. ผลวิจัยมีขนาดใหญ่

ผลวิจัยสำหรับการวิจัยครั้งนี้มีจำนวน ๑,๘๗๒ คน มีผู้สอบแบบสอบที่  
ไม่มีชื่อเรื่องและภาพประกอบเรื่อง จำนวน ๔๔๔ คน และมีผู้สอบแบบสอบที่มีชื่อเรื่องและ  
ภาพประกอบเรื่อง จำนวน ๔๑๘ คน การที่ผลวิจัยมีขนาดใหญ่จะทำให้คะแนนของแบบ  
สอบที่ให้คะแนนต่างกันแต่ละวิธี มีการกระจายมากกว่าผลวิจัยขนาดเล็ก และเมื่อคะแนน  
มีการกระจายมากขึ้น ความสัมพันธ์ระหว่างคะแนนต่างวิธีกัน ก็ยังมีโอกาสที่จะมีความ  
สัมพันธ์กันมากยิ่งขึ้น (McCall, ๑๙๗๐: ๑๒๗)

#### ค. คะแนนของบางวิธีมีขนาดใหญ่

คะแนนของการให้คะแนนบางวิธีมีขนาดใหญ่ เช่น วิธีใช้ค่าเฉลี่ย  
( $\Delta$ ) และวิธีใช้ค่าเบต้าแปลง ( $\beta_T$ ) ดังนั้น คะแนนของวิธีเหล่านี้มีโอกาสที่จะกระจายมาก  
จึงอาจมีผลทำให้สัมพันธ์กับคะแนนของแบบสอบที่ให้คะแนนวิธีอื่นได้สูงมาก (McCall, ๑๙๗๐:  
๑๓๑)

ดังนั้น ผู้วิจัยมีความเห็นว่า การที่คะแนนของแบบสอบที่ให้คะแนนต่างวิธีกัน มีความสัมพันธ์ซึ่งกันและกันสูงอย่างมีนัยสำคัญ คงจะเกิดจากปัจจัยที่สำคัญทั้ง ๓ ประการ ดังกล่าวแล้ว ข้อค้นพบจากการวิจัยครั้งนี้สอดคล้องกับผลการวิจัยของ Corey (อ้างจาก Wang and Stanley, ๑๙๗๐:๖๘๘) ที่พบว่า สัมประสิทธิ์สหสัมพันธ์ของคะแนนสอบที่ให้คะแนนวิธีประเพณีนิยม กับคะแนนที่ให้ตามน้ำหนักความสำคัญของสิ่งที่วัด มีค่าระหว่าง .๘๓๖ ถึง .๘๖๐ ซึ่งนับว่าสูงมาก และสอดคล้องกับผลการวิจัยของ Douglas Spencer (อ้างจาก Wang and Stanley, ๑๙๗๐:๖๘๘) ที่พบว่า ความสัมพันธ์ระหว่างคะแนนสอบของวิธีประเพณีนิยม กับคะแนนที่ให้โดยวิธีต่าง ๆ ๖ วิธี (ไม่ได้ระบุวิธีใดบ้าง) มีค่าระหว่าง .๘๕๐ ถึง .๘๕๑ ซึ่งนับว่าสูงมากเช่นเดียวกับผลการวิจัยครั้งนี้ ซึ่งพบว่า ความสัมพันธ์ซึ่งกันและกันระหว่างคะแนนที่ให้โดยวิธีต่าง ๆ มีค่าระหว่าง .๘๕๑ ถึง .๘๕๕

#### ๖. อิทธิพลของโครงสร้างความรู้

ผลของการวิเคราะห์ข้อมูลเกี่ยวกับอิทธิพลของโครงสร้างความรู้ สอดคล้องกับสมมุติฐานในการวิจัยข้อที่ ๖ บางประการ กล่าวคือ สมมุติฐานกล่าวไว้ว่า "โครงสร้างความรู้ที่เกิดจากชื่อเรื่องและภาพประกอบเรื่อง ที่มีคำศัพท์ที่มีความสำคัญต่อเรื่องที่อ่าน และความคุ้นเคยต่อเนื้อหาสาระของเรื่องที่อ่าน มีผลต่อความสามารถทั่วไปในการอ่าน เข้าใจความภาษาอังกฤษของนิสิตจุฬาลงกรณ์มหาวิทยาลัย ที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศอย่างมีนัยสำคัญ" แต่ผลการวิจัยพบว่า เฉพาะความคุ้นเคยเท่านั้นที่มีผลต่อการอ่าน เข้าใจความอย่างมีนัยสำคัญ

การที่ผลการวิจัยเกิดขึ้นดังกล่าวแล้ว อาจมีสาเหตุดังต่อไปนี้

ก. ผลของโครงสร้างความรู้ที่เกิดจากชื่อเรื่องและภาพประกอบเรื่อง

การที่โครงสร้างความรู้ที่เกิดจากชื่อเรื่องและภาพประกอบเรื่อง ไม่มีผลต่อความสามารถในการอ่าน เข้าใจความอย่างมีนัยสำคัญ อาจเป็นเพราะว่า

๑. เรื่องที่อ่านทั้ง ๒ เรื่อง ของแบบสอบทั้ง ๒ แบบ

มีคำศัพท์ที่มีความหมายสำคัญต่อเรื่องที่อ่าน (transparent words) คำศัพท์ที่มีลักษณะดังกล่าวสามารถทำให้ผู้อ่านเกิดโครงสร้างความรู้เพื่อความเข้าใจต่อเรื่องที่อ่านในลักษณะ "เบื้องต่ำ-สู่-เบื้องสูง" (bottom-up processing) ได้ (Rumelhart, ๑๙๘๐:๔๑)

๒. เรื่องที่นำมาสอบได้ทำการทดสอบแล้วว่า มีระดับความยากง่ายเหมาะสมกับระดับความสามารถของผู้สอบ

ดังนั้น ผู้สอบโดยมากคงไม่มีปัญหาด้านโครงสร้างและศัพท์ในเรื่องที่อ่าน จึงทำให้ชื่อเรื่องและภาพประกอบเรื่องมีความสำคัญน้อยมาก สำหรับการอ่านเข้าใจความ แสดงว่าคำศัพท์ที่มีความสำคัญต่อเรื่องที่อ่าน ทำให้ผู้อ่านเกิดโครงสร้างความรู้ในการอ่านได้อย่างเพียงพอ โดยไม่ต้องอาศัยโครงสร้างความรู้จากชื่อเรื่องและภาพประกอบเรื่องช่วยในการเข้าใจความอีก ผลการวิจัยนี้สอดคล้องกับผลการวิจัยของ Carrell (Carrell, ๑๙๘๓:๑๙๗) ผู้พบว่าชื่อเรื่องและภาพประกอบเรื่องไม่มีผลต่อความสามารถในการอ่านเข้าใจความภาษาอังกฤษอย่างมีนัยสำคัญ สำหรับผู้เรียนภาษาอังกฤษที่ไม่ใช่เจ้าของภาษา และมีความสามารถทางภาษาในระดับสูงกว่าระดับปานกลางเล็กน้อย

ข. ผลของโครงสร้างความรู้ที่เกิดจากความคุ้นเคยต่อเนื้อหาสาระของเรื่องที่อ่าน

การที่โครงสร้างความรู้ที่เกิดจากความคุ้นเคยต่อสาระของเรื่องที่อ่าน มีผลต่อความสามารถทั่วไปในการอ่านอย่างมีนัยสำคัญ อาจเป็นเพราะว่า

๑. ความคุ้นเคยมีผลต่อความจำ และมีประโยชน์ต่อการตรวจสอบสาระของเรื่อง

ปกติแล้วคนสามารถจดจำสิ่งที่ตนเองคุ้นเคยมากกว่าได้มากกว่าสิ่งที่ตนคุ้นเคยน้อยกว่า หรือไม่คุ้นเคยเลย โครงสร้างความรู้ที่เกิดจากความคุ้นเคยทำหน้าที่ที่สำคัญอย่างหนึ่งในการอ่านเข้าใจความ คือ ทำหน้าที่ตรวจสอบสาระของเรื่องว่า

สอดคล้องกับสิ่งที่ได้รู้จักมาแล้วหรือไม่ เช่น ผู้อ่านที่ไม่คุ้นเคยหรือรู้จักขั้นตอนในการรับประทาน อาหารในภัตตาคารในต่างประเทศ จะมีปัญหาในการเข้าใจความหมายของข้อความที่เกี่ยวข้อง เช่น "Bill refused to leave a tip because the service was slow." เป็นต้น (Bransford and Others, ๑๙๘๔:๓๕) แต่สำหรับผู้ที่คุ้นเคยกับ สถานการณ์ดังกล่าวจะเข้าใจข้อความที่ยกมาข้างต้นได้ง่าย

## ๒. ความคุ้นเคยทำให้เกิดความเข้าใจในการอ่านได้รวดเร็ว

นอกจากโครงสร้างความรู้ที่เกิดจากความคุ้นเคย จะทำหน้าที่ตรวจสอบสาระของเรื่องทีอ่านแล้ว ยังทำให้เกิดความเข้าใจในการอ่านได้ รวดเร็วกว่าการอ่านสิ่งที่ไม่คุ้นเคย หรือคุ้นเคยน้อยกว่าด้วย (Bransford and Others, ๑๙๘๔:๓๕)

ผลการวิจัยครั้งนี้สอดคล้องกับผลการวิจัยของ Carrell (Carrell, ๑๙๘๔b:๑๐๓) และ Johnson (Johnson, ๑๙๘๒:๕๑๐) ที่พบว่า ผู้อ่านในระดับมหาวิทยาลัยที่ไม่ใช่เจ้าของภาษาอังกฤษ ที่มีความสามารถในการอ่าน ระดับปานกลางสามารถอ่านเรื่องที่ตนเองคุ้นเคยได้เข้าใจมากกว่า และเรียงลำดับสาระ ของเรื่องได้ถูกต้องมากกว่าเรื่องที่ตนเองไม่คุ้นเคยเลย หรือคุ้นเคยน้อยกว่าอย่างมีนัยสำคัญ

## ๓. ความสัมพันธ์ระหว่างความคุ้นเคยต่อเรื่องทีอ่านกับความง่ายของเรื่อง

การวิจัยครั้งนี้พบว่า ผลการวิจัยสอดคล้องกับสมมุติฐานในการวิจัยข้อ ที่ ๓ ที่กล่าวว่า "ความคุ้นเคยต่อเรื่องทีอ่าน และความง่ายของเรื่อง มีความสัมพันธ์ กันอย่างมีนัยสำคัญ"

การที่ความคุ้นเคยต่อเรื่องทีอ่านสัมพันธ์กับความรู้สึกว่า เรื่องทีอ่านยาก หรือง่าย อาจเป็นเพราะว่า ความคุ้นเคยต่อเรื่องทีอ่านเป็นอิสระจากความซับซ้อนของโครงสร้างผิว (surface structure complexity) ถ้าผู้อ่านมีโครงสร้างความรู้ซึ่งเกิดจาก ความคุ้นเคยต่อเรื่องทีอ่าน สอดคล้องกับโครงสร้างความรู้ของผู้เขียนหรือของเรื่องทีอ่าน จะทำให้อ่านเรื่องนั้นได้เข้าใจง่ายขึ้น และจะรู้สึกว่างเรื่องนั้นง่าย แม้ว่าโครงสร้างผิวของ

เรื่องที่อ่านจะซับซ้อนมากเพียงใด แต่ถ้าหากโครงสร้างความรู้ของผู้อ่านซึ่งเกิดจากความคุ้นเคยต่อเรื่องที่อ่าน กับโครงสร้างความรู้ของผู้เขียน หรือโครงสร้างความรู้ของเรื่องที่อ่าน มีความแตกต่างกัน จะทำให้ผู้อ่านมีปัญหาในการอ่าน และรู้สึกว่าเรื่องนั้นยาก (Frederiksen, ๑๙๘๐ อ้างจาก C. Daungploy, ๑๙๘๒:๒๔)

ผลการวิจัยนี้สอดคล้องกับผลการวิจัยของ Carrell (Carrell, ๑๙๘๓:๑๙๖) ที่พบว่า ผู้อ่านในระดับมหาวิทยาลัยที่ไม่ใช่เจ้าของภาษาอังกฤษ และมีความสามารถในการอ่านระดับสูงกว่าระดับปานกลางเล็กน้อย มักจะตอบว่าเรื่องที่ตนเองคุ้นเคยนั้นง่ายกว่าเรื่องที่ตนเองไม่คุ้นเคย ทั้ง ๆ ที่เรื่องดังกล่าวนั้นมีระดับความยากง่ายไม่แตกต่างกันอย่างมีนัยสำคัญ

#### ข้อเสนอแนะ

##### ก. เพื่อนำผลการวิจัยไปประยุกต์ใช้

๑. ถ้าเป็นไปได้ การสอบอิงกลุ่มเพื่อวัดความสามารถทั่วไปในการอ่านเข้าใจความภาษาอังกฤษ แบบสอบควรให้คะแนนโดยวิธีใช้ค่านำหนักคะแนนที่เหมาะสม (w) ซึ่งคำนวณได้จากผลการวิเคราะห์ข้อสอบด้วยรูปแบบโลจิสติกใช้พารามิเตอร์ ๓ ตัว เพราะเหตุผลต่อไปนี้

ก. ทำให้แบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่องมีค่าความตรงเชิงพยากรณ์สูงการให้คะแนนวิธีอื่น ๆ แม้ว่าทำให้แบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง มีค่าความตรงเชิงพยากรณ์ต่ำ แต่ก็ไม่ได้แตกต่างจากค่าความตรงที่คำนวณได้จากคะแนนของการให้คะแนนวิธีอื่น ๆ มากนัก

ข. ทำให้แบบสอบทั้ง ๒ แบบมีค่าความเที่ยงสูงกว่าค่าความเที่ยงที่คำนวณได้จากคะแนนของการให้คะแนนวิธีอื่น ๆ และสูงกว่าค่าความเที่ยงที่คำนวณจากคะแนนของวิธีประเพณีนิยมอย่างมีนัยสำคัญ ถ้าต้องการให้แบบสอบที่ให้คะแนนวิธีประเพณีนิยมมีค่าความเที่ยงสูงเท่ากับค่าความเที่ยงที่คำนวณจากคะแนนของวิธีใช้ค่านำหนักคะแนนที่เหมาะสม (w) จะต้องเพิ่มข้อสอบอีกจำนวนหนึ่ง เช่น ถ้ามีข้อสอบ ๕๐ ข้อ จะต้องเพิ่มขึ้นอีกประมาณ ๑๔ ข้อ หรือประมาณ ๐.๓๐ เท่าของจำนวนข้อสอบชุดเดิม เป็นต้น





ค. คะแนนรายข้อของข้อสอบของวิธีการให้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) เป็นค่าที่ไม่เปลี่ยนแปลงไปตามกลุ่มตัวอย่างของการทดสอบ จึงมีประโยชน์ในการนำแบบสอบเดิมไปสอบกับกลุ่มตัวอย่างอื่น และใช้คะแนนรายข้อที่คำนวณไว้แล้วได้อีก

๒. ในกรณีที่ไม่สามารถคิดคะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) ควรให้โดยวิธีประเพณีนิยม เพราะการให้โดยวิธีอื่น ๆ เช่น วิธีใช้ค่าผลบวกความแปรปรวนร่วม ( $h^2$ ) วิธีใช้ค่าเคลตต้า ( $\Delta$ ) วิธีใช้ค่าเบต้าแปลง ( $\beta_t$ ) ทำให้แบบสอบมีค่าความเที่ยงต่ำกว่าค่าความเที่ยงที่คำนวณจากคะแนนโดยวิธีประเพณีนิยมอย่างมีนัยสำคัญ และวิธีต่าง ๆ เหล่านี้โดยมากทำให้แบบสอบมีค่าความตรงเชิงพยากรณ์สูงเท่า ๆ กับความตรงที่คำนวณจากคะแนนโดยวิธีประเพณีนิยม

๓. ในกรณีของการสอบอิงเกณฑ์ ควรให้คะแนนของแบบสอบโดยวิธีใช้ค่าเบต้าแปลง ( $\beta_t$ ) เพราะโดยเฉลี่ยแล้วเมื่อจุดตัดมากกว่าร้อยละ ๕๐ คะแนนที่ให้โดยวิธีดังกล่าวมีค่าความเที่ยง (ค่าดัชนีความเชื่อถือ) สูงที่สุด ทั้งแบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง และแบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง แต่ถ้าหากว่าไม่สามารถให้คะแนนแบบสอบด้วยวิธีดังกล่าวได้ ก็อาจให้คะแนนโดยวิธีใช้ค่าเคลตต้า ( $\Delta$ ) หรือวิธีใช้ค่าผลบวกความแปรปรวนร่วม ( $h^2$ ) ก็ได้ เพราะคะแนนจากวิธีเหล่านี้โดยมากทำให้แบบสอบมีค่าความเที่ยง (ค่าดัชนีความเชื่อถือ) สูงกว่าค่าความเที่ยง (ค่าดัชนีความเชื่อถือ) ที่คำนวณจากคะแนนของวิธีประเพณีนิยม แต่อย่างไรก็ตาม ไม่ควรให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) สำหรับการสอบอิงเกณฑ์ เพราะวิธีนี้ทำให้แบบสอบมีค่าความเที่ยง (ค่าดัชนีความเชื่อถือ) ต่ำที่สุด ไม่ว่าจุดตัดของคะแนนสอบจะเป็นเท่าใดก็ตาม ยกเว้นเมื่อจุดตัดเท่ากับร้อยละ ๕๐ เท่านั้น

๔. ในกรณีการสอบคัดเลือก แบบสอบควรให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนนที่เหมาะสม ( $w$ ) เพราะจะทำให้แบบสอบมีความเที่ยงสูงกว่าค่าความเที่ยงที่ได้จากการให้คะแนนวิธีอื่น ๆ อย่างมีนัยสำคัญ

๕. การเรียนการสอนภาษานั้น ควรอย่างยิ่งที่จะต้องมีการสอนนำให้ ผู้เรียนมีโครงสร้างความรู้ต่อสิ่งที่จะเรียนก่อนเสมอ ไม่ว่าจะเป็นการสอนอ่าน การสอน การฟัง เข้าใจความ หรือการสอนเขียนก็ตาม เพราะโครงสร้างความรู้ที่เกิดจากความ ค้นเคยต่อเนื้อหาสาระของเรื่องที้อ่าน มีผลอย่างมีนัยสำคัญต่อการอ่านเข้าใจความ ดังนั้น โครงสร้างความรู้ดังกล่าวนี้มีส่วนทำให้ผู้เรียนเข้าใจสิ่งที่จะเรียนได้ดียิ่งขึ้นด้วย เช่น ผู้สอบอาจเล่าเรื่องย่อ ๆ ให้ผู้จะเรียนการอ่านเข้าใจความฟัง หรือสอนศัพท์บางคำที่จะก่อให้เกิดโครงสร้างความรู้ที่ถูกต้องได้ดีกว่าก่อนให้ผู้เรียนอ่านเอง เป็นต้น

#### ข. เพื่อการวิจัยต่อไป

๑. ควรทำการศึกษาว่า ในกรณีที่แบบสอบมีระดับความยากแตกต่างกัน และ/หรือมีอำนาจจำแนกแตกต่างกัน การให้คะแนนทั้ง ๕ วิธี ดังเช่นการศึกษาครั้งนี้ ทำให้แบบสอบมีค่าความตรงเชิงพยากรณ์ ค่าความเที่ยง และความคงที่ของคะแนนสอบ แตกต่างจากการให้คะแนนวิธีประเพณีนิยมมาน้อยเพียงใด

๒. ควรทำการศึกษาว่า เหตุใดการให้คะแนนโดยวิธีใช้ค่าน้ำหนักคะแนน ที่เหมาะสม (w) ทำให้แบบสอบที่มีชื่อเรื่องและภาพประกอบเรื่อง มีค่าความตรงเชิงพยากรณ์ สูงที่สุด แต่ทำให้แบบสอบที่ไม่มีชื่อเรื่องและภาพประกอบเรื่อง มีค่าความตรงเชิงพยากรณ์ ต่ำที่สุด

๓. ควรทำการศึกษาว่า การให้คะแนนโดยวิธีใช้ค่าลำดับชั้นความเข้าใจ ในการอ่าน (reading comprehension hierarchies) ทำให้แบบสอบที่มีโครงสร้าง ความรู้ต่างกัน มีค่าความตรงเชิงพยากรณ์ ค่าความเที่ยง และความคงที่ของอันดับที่ของคะแนน ต่างจากค่าดังกล่าวนี้ที่คำนวณจากคะแนนที่ได้จากการให้คะแนนวิธีต่าง ๆ ดังกล่าวแล้วหรือไม่

๔. ควรทำการศึกษาเปรียบเทียบการให้คะแนนรายข้อวิธีต่าง ๆ ว่า มีผลต่อค่าความเที่ยง ความตรง และอันดับที่ของคะแนนสอบของคะแนนปริเขต (domain score) ของสิ่งที่วัด ที่อาศัยแนวคิดของรูปแบบต่าง ๆ จากทฤษฎีความสามารถแฝงมาน้อยเพียงใด

๕. ควรทำการศึกษาค่าความตรงร่วมสมัยของแบบสอบที่ใช้ เป็นเครื่องมือสำหรับการวิจัยนี้ กับคะแนนสอบเข้ามหาวิทยาลัย หรือคะแนนสอบเพื่อจัดชั้นเรียนของสถาบันภาษา เพื่อตรวจสอบว่า ค่าความตรงดังกล่าวจะมีขนาดมากน้อยเพียงใด ในกรณีที่ให้คะแนนรายข้อต่างกันทั้ง ๕ วิธี

๖. ควรทำการศึกษาอิทธิพลของโครงสร้างความรู้ชนิดอื่น ๆ เช่น คำศัพท์ที่มีความหมายสำคัญต่อเรื่องที่อ่าน คำศัพท์ที่ไม่มีความหมายสำคัญต่อเรื่องที่อ่าน ความซับซ้อนของโครงสร้างผิว และความซับซ้อนของโครงสร้างลึก เป็นต้น ว่ามีผลต่อการอ่านเข้าใจ ความ หรือการฟังเข้าใจความของผู้อ่าน และผู้ฟังที่มีความสามารถในระดับต่าง ๆ เป็นอย่างไร