

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

สรุปผลการวิจัย

การวิจัยนี้มีวัตถุประสงค์เพื่อสร้างแบบสอบความเข้าใจในการฟังภาษาอังกฤษ สำหรับระดับเริ่มเรียน ซึ่งใช้แบบเรียน JILAP Stage I (red, green and blue series) และ Stage II (tan series) และแบบเรียน The Oxford English Course for Thailand Book II และ Book III ให้มีคุณภาพสูง เป็นที่เชื่อถือได้

ตัวอย่างประชากรที่ใช้ในการวิจัยมี ๒ กลุ่ม คือ

กลุ่มที่ ๑ ตัวอย่างประชากรที่ใช้ในการทดลองสอบ เป็นนักเรียนชั้นประถมศึกษาปีที่ ๔ - ๕ ของโรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตรศาสตร์ จำนวน ๒๕๗ คน และเป็นนักเรียนชั้นประถมศึกษาปีที่ ๖ - ๗ ของโรงเรียนวัดธาตุทอง จำนวน ๒๑๒ คน รวมตัวอย่างประชากรทั้งหมด ๔๖๙ คน

กลุ่มที่ ๒ ตัวอย่างประชากรที่ใช้ในการทดสอบจริง ได้แก่ นักเรียนชั้นประถมศึกษาปีที่ ๕, ๖ และ ๗ ในโรงเรียนต่าง ๆ ในจังหวัดขอนแก่น รวมตัวอย่างประชากรทั้งหมด ๑๑๘๖ คน

วิธีการวิจัย เนื่องจากการวิจัยนี้เป็นการสร้างแบบสอบ ซึ่งคุณภาพของแบบสอบที่สร้างขึ้นกับค่าความยาก อำนาจจำแนก และความเชื่อถือได้ของแบบสอบเป็นสำคัญ ฉะนั้น ในการดำเนินงานวิจัย ผู้วิจัยจึงได้ดำเนินการตามวิธีการต่อไปนี้

๑. ศึกษาและวิเคราะห์แบบเรียน JILAP Stage I, Stage II และ The Oxford English Course for Thailand Book II, Book III

๒. สร้างแบบสอบความเข้าใจในการฟังภาษาอังกฤษสำหรับผู้เริ่มเรียนที่ใช้แบบเรียนทั้งสอง แบบสอบที่สร้างเป็นแบบสอบคู่ขนาน ๓ ชุด คือ ชุด JIIA และ JIIB

ซึ่งยึดเนื้อหาจากแบบเรียน JILAP Stage I (red and green series) ชุด OIIA และ OIIB ซึ่งยึดเนื้อหาจากแบบเรียน The Oxford English Course for Thailand Book II และชุด JOIIIA และ JOIIIB ซึ่งยึดเนื้อหาาร่วมกันของแบบเรียน JILAP Stage I (blue series), Stage II (tan series) และแบบเรียน The Oxford English Course for Thailand Book III. ข้อทดสอบแต่ละฉบับมี ๑๕ ข้อ คัดเลือกเป็นภาพ แต่ละข้อมี ๔ คัดเลือก

๓. นำแบบสอบคู่ขนานของทั้ง ๓ ชุด ไปทดลองสอบกับตัวอย่างประชากรกลุ่มที่ ๑ จำนวน ๘๕ คน นำคะแนนที่ได้มาทำการวิเคราะห์ข้อสอบเป็นรายข้อ เพื่อคัดเลือกข้อสอบที่ดี มีระดับความยากปานกลาง และมีอำนาจจำแนกสูง จากแบบสอบคู่ขนานแล้วรวมเป็นชุดเดียวกัน นำข้อสอบที่คัดเลือกแล้วมาเรียงลำดับใหม่ตามความง่ายสำหรับนำไปทดสอบจริง หากค่าความเชื่อถือได้ของแบบสอบคู่ขนานแต่ละชุดจากผลการทดลองสอบ เพื่อให้ทราบค่าความเชื่อถือได้ของแบบสอบก่อนการปรับปรุงแก้ไข

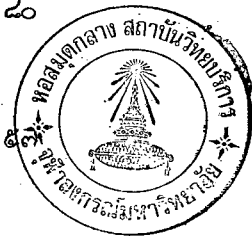
๔. นำแบบสอบที่ปรับปรุงแล้ว จำนวน ๓ ชุด คือ JII, OII และ JOIII จำนวนชุดละ ๑๕ ข้อ ไปทดสอบจริงกับตัวอย่างประชากรกลุ่มที่ ๒ จำนวน ๑๑๘ คน นำคะแนนที่ได้มาวิเคราะห์ โดยเครื่องคอมพิวเตอร์ชนิด ไอ บี เอ็ม ๓๕๐/๑๕๕ (I.B.M. ๓๕๐/๑๕๕ Computer ) ของสถาบันเทคโนโลยีแห่งเอเชีย คำนวณค่าสถิติดังต่อไปนี้

- ๔.๑ มัชฌิมเลขคณิต
- ๔.๒ ส่วนเบี่ยงเบนมาตรฐาน
- ๔.๓ ความเชื่อถือได้
- ๔.๔ ความคลาดเคลื่อนมาตรฐานในการวัด

ผลการวิจัย สรุปได้ดังนี้

ข้อทดสอบที่จะใช้ทดสอบจริง ได้พิจารณาเลือกจากข้อที่มีระดับความยากปานกลาง ( $p = .๒๐ - .๔๐$ ) และมีอำนาจจำแนกสูง ( $r = .๒๐$  ขึ้นไป) จากการทดลองสอบ ชุดละ ๑๕ ข้อ ผลการทดสอบจริงปรากฏว่า แบบสอบทั้ง ๓ มีความตรงเชิงเนื้อหา และค่าสถิติต่าง ๆ สรุปได้ดังนี้

แบบสอบ JII	ค่าระดับความยากของแบบสอบ ( $p$ ) อยู่ระหว่าง	.๒๘ - .๕๒
	เฉลี่ยค่าระดับความยาก ( $\bar{p}$ ) เท่ากับ	.๔๕
	ค่าอำนาจจำแนกของแบบสอบ ( $r_{bis}$ ) อยู่ระหว่าง	.๒๐ - .๕๔
	เฉลี่ยค่าอำนาจจำแนก ( $\bar{r}$ ) เท่ากับ	.๔๒
	ค่าความเชื่อถือได้ของแบบสอบ ( $r_{tt}$ ) เท่ากับ	.๕๖
	ค่าความคลาดเคลื่อนมาตรฐานในการวัด ( $\sigma_{meas}$ ) เท่ากับ	๑.๕๑
แบบสอบ OII	ค่าระดับความยากของแบบสอบ ( $p$ ) อยู่ระหว่าง	.๑๕ - .๔๐
	เฉลี่ยค่าระดับความยาก ( $\bar{p}$ ) เท่ากับ	.๒๗
	ค่าอำนาจจำแนกของแบบสอบ ( $r_{bis}$ ) อยู่ระหว่าง	.๑๗ - .๕๖
	เฉลี่ยค่าอำนาจจำแนก ( $\bar{r}$ ) เท่ากับ	.๔๐
	ค่าความเชื่อถือได้ของแบบสอบ ( $r_{tt}$ ) เท่ากับ	.๕๔
	ค่าความคลาดเคลื่อนมาตรฐานในการวัด ( $\sigma_{meas}$ ) เท่ากับ	๑.๔๕
แบบสอบ JOIII	ค่าระดับความยากของแบบสอบ ( $p$ ) อยู่ระหว่าง	.๑๔ - .๕๓
	เฉลี่ยค่าระดับความยาก ( $\bar{p}$ ) เท่ากับ	.๔๔
	ค่าอำนาจจำแนกของแบบสอบ ( $r_{bis}$ ) อยู่ระหว่าง	-.๐๓ - .๖๖
	เฉลี่ยค่าอำนาจจำแนก ( $\bar{r}$ ) เท่ากับ	.๔๑
	ค่าความเชื่อถือได้ของแบบสอบ ( $r_{tt}$ ) เท่ากับ	.๕๕
	ค่าความคลาดเคลื่อนมาตรฐานในการวัด ( $\sigma_{meas}$ ) เท่ากับ	๑.๕๑



(ดูรายละเอียดของแบบสอบทั้ง ๓ ชุด ได้ที่แผนกวิชามัธยมศึกษา คณะครุศาสตร์)

#### การอภิปรายผลการวิจัย

๑. จากการวิเคราะห์ผลการทดลองใช้แบบสอบ จะเห็นได้ว่า แบบสอบที่สร้างขึ้น และนำไปทดลองสอบนั้นเป็นแบบสอบที่ค่อนข้างยาก เพราะมีค่าระดับความยากต่ำ ( $\bar{p} = .๒๖ - .๓๔$ ) แต่สำหรับแบบสอบที่ใช้ทดสอบจริง มีระดับความยากและอำนาจจำแนกสูงอยู่ในเกณฑ์มาตรฐานเป็นส่วนมาก ( $\bar{p} = .๔๒ - .๔๕$ ;  $\bar{r} = .๔๐ - .๔๒$ )

๒. ค่าเฉลี่ยระดับความยากและอำนาจจำแนกของแบบสอบที่นำไปทดสอบจริงสูงกว่าที่นำไปทดลองสอบ ( $\bar{p}$  จาก .๓๘ - .๕๒ เป็น .๕๒ - .๕๘ ,  $\bar{r}$  จาก .๓๕ - .๕๑ เป็น .๕๐ - .๕๒) ทั้งนี้เพราะคะแนนจากผลการทดสอบจริง เป็นส่วนหนึ่งของคะแนนสอบปลายปีของเด็ก ทำให้ทั้งใจหายของทดสอบ และคะแนนจากผลการทดสอบไม่มีส่วนเกี่ยวข้องของการสอบปลายปี ทำให้เด็กไม่ตั้งใจทำข้อทดสอบเท่าที่ควร และการทดสอบจริงเป็นระยะที่เด็กเรียนเนื้อหาครบแล้ว ส่วนระยะที่ไปทดลองสอบ เด็กยังเรียนเนื้อหาได้ไม่ครบ นอกจากนี้ความชัดเจนของเสียงพูดจากเทป สำเนียงของบทกวีอีกเล่ม และบรรยากาศขณะสอบ มีส่วนทำให้กลุ่มตัวอย่างประมาณจริงทำข้อสอบได้ทั้งหมด

๓. ค่าความเชื่อถือได้ของแบบสอบที่นำไปทดสอบจริงสูงกว่าที่นำไปทดลองสอบ  $r_{tt}$  จาก .๓๑ - .๓๘ เป็น .๕๕ - .๕๘ ทั้งนี้เนื่องจากคุณภาพของเครื่องมือบันทึกเสียง สำเนียงพูดในเทป รวมทั้งความชัดเจนของสำเนียงคำกวีในแบบสอบที่นำไปทดสอบจริงดีกว่าแบบทดลองสอบ ทำให้นักเรียนทำแบบสอบได้มากขึ้น

๔. แบบสอบที่ใช้ในการทดสอบจริงมีความยากและอำนาจจำแนกอยู่ในเกณฑ์มาตรฐานเป็นส่วนใหญ่ ( $\bar{p} = .๕๒ - .๕๘$  ,  $\bar{r} = .๕๐ - .๕๒$ ) แสดงว่าข้อสอบมีพัฒนาการขึ้น สามารถปรับปรุงให้เป็นข้อสอบที่ดีได้ จำแนกเด็กเก่งและเด็กอ่อนได้ก็จัดเป็นข้อสอบที่มีความเชื่อถือได้ เพราะมีค่าความเชื่อถือได้เพิ่มขึ้น ( $r_{tt}$  จาก .๓๑ - .๓๘ เป็น .๕๕ - .๕๘) ค่าความเชื่อถือได้จึงจัดอยู่ในเกณฑ์ดีพอใช้ ( $r_{tt}$  .๕๕ - .๕๘) แบบสอบวัดความสามารถในกาฬพิงนี้มีตัวแปรหลายอย่างที่มีอิทธิพลด้วย เช่น คุณภาพเครื่องมือบันทึกเสียง สำเนียงพูดในเทป ความชัดเจนของประโยคเมื่อได้ฟังจากเทป ความที่ผู้ตั้งในกาฬพิงได้นั้นข้อคำถาม และความคุ้นเคยของผู้ถูกทดสอบต่อแบบสอบประเภทนี้ เป็นต้น ถ้าสามารถควบคุมตัวแปรเหล่านี้ได้มากขึ้น ผู้วิจัยเชื่อว่า จะสามารถเพิ่มค่าความเชื่อถือได้ของแบบสอบให้สูงขึ้นอีกได้ การเพิ่มจำนวนข้อทดสอบก็จะช่วยให้ค่าความเชื่อถือได้ของแบบสอบเพิ่มขึ้นด้วย

## ข้อเสนอแนะ

การวิจัยนี้ยังบกพร่อง และผลการวิจัยยังไม่สมบูรณ์ ควรต้องมีตารางแก้ไข และทำการวิจัยซ้ำอีก เพื่อปรับปรุงให้ข้อสอบที่วิเคราะห์แล้วนี้มีคุณภาพดียิ่งขึ้น ผู้วิจัยมีข้อเสนอแนะดังนี้

๑. ในการทดสอบไม่ควรทดสอบทีละหลายแบบสอบติดต่อกัน และให้เวลานานจนเกินไป ทำให้นักเรียนไม่ตั้งใจทำแบบสอบสุดท้าย ๆ ควรตัดทอนเวลาและจำนวนแบบสอบให้น้อยลง เพื่อไม่ให้เด็กเกิดความเบื่อหน่ายจนหมดความตั้งใจทำแบบสอบ

๒. การเลือกตัวอย่างประชากรในการทดสอบและในการทดสอบจริง ควรเลือกให้มีคุณสมบัติใกล้เคียงกันมากที่สุด มิฉะนั้นผลการวิจัยจะไม่ได้ดีเท่าที่ควร

๓. ควรให้นักเรียนได้เรียนเนื้อหาครบก่อนการทดสอบผลสัมฤทธิ์ทุกครั้ง เพื่อความแม่นยำและเที่ยงตรงในการวิเคราะห์ข้อทดสอบ

๔. การสร้างแบบสอบทักษะการฟังที่มีตัวเลือกเป็นภาพ ผู้สร้างข้อทดสอบจำเป็นต้องควบคุมการเขียนภาพอย่างใกล้ชิด มิฉะนั้นจะได้ภาพที่ผิดเพี้ยนจากความประสงค์ได้ง่าย รวมทั้งอาจไม่สื่อความหมายความต้องการ และอาจกำกวม เมื่อทำสำเนาออกมาจำเป็นต้องตรวจอีกครั้งว่า ภาพชัด ครบ ไม่มีสิ่งใดแหงนหาย เพราะสิ่งเหล่านี้จะมีส่วนกระทบกระเทือนต่อค่าความเชื่อถือได้ และความตรงของแบบสอบ

๕. ควรมีการสร้างแบบสอบภาษาอังกฤษในทักษะนี้และทักษะอื่น ๆ ในระดับนี้ และในระดับอื่น ๆ ด้วย เพื่อให้ได้แบบสอบที่วัดได้ตรงตามวัตถุประสงค์ แต่ในการสอบไม่จำเป็นต้องวัดทักษะเดียวในหนึ่งแบบสอบ เพราะบางทักษะมีตัวแปรต่าง ๆ ซึ่งมีอิทธิพลต่อค่าความเชื่อถือได้ของแบบสอบ และในการใช้ภาษาจริง ก็ต้องใช้ความสามารถในทุกทักษะรวมกันไปแล้ว

๖. ควรมีการสร้างแบบสอบทำนองเดียวกันนี้สำหรับวิชาอื่น ๆ ให้สามารถวัดได้ตรงตามวัตถุประสงค์ของวิชานั้น ๆ ทั้งในระดับนี้และในระดับอื่น ๆ ด้วย.