การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบวัดความสามารถด้านการรู้เรื่องการอ่าน ระหว่างโรงเรียนประถมศึกษาที่จัดการศึกษาแบบทวิภาษา และโรงเรียนปกติ*

วสันต์ นันทะเสน[1] สังวรณ์ งัดกระโทก[2] นลินี ณ นคร[3]

(รับบทความ: 10 เมษายน 2562; แก้ไขบทความ: 16 สิงหาคม 2562; ตอบรับบทความ: 23 สิงหาคม 2562)

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบวัดความสามารถด้านการรู้เรื่องการอ่าน ของการทดสอบระดับชาติ ระหว่างโรงเรียนประถมศึกษาที่จัดการศึกษาแบบทวิภาษา และโรงเรียนปกติ 2) วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ ระหว่างโรงเรียนแบบทวิภาษากลุ่ม ไทย-ม้ง ไทย-มอญ ไทย-มลายูถิ่น ไทย-กะเหรี่ยงโปว์ ไทย-เขมรถิ่นไทยและโรงเรียนปกติ และ 3) วิเคราะห์สาเหตุของการทำหน้าที่ต่างกันของข้อสอบ กลุ่มตัวอย่างได้แก่ นักเรียนชั้นประถมศึกษาปีที่ 3 ปีการศึกษา 2557 ที่เข้าร่วมโครงการทวิภาษา จำนวน 11 โรงเรียน 340 คน เป็นกลุ่มอ้างอิง และโรงเรียนปกติ จำนวน 17 โรงเรียน 339 คน เป็นกลุ่มเปรียบเทียบ เครื่องมือที่ใช้ในการวิจัย คือ แบบบันทึกคะแนนวัดผลสัมฤทธิ์ทางการเรียนระดับการศึกษาขั้นพื้นฐาน (NT) ด้านการรู้เรื่องการอ่าน (Reading Literacy) การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบใช้วิธีการถดถอยโลจิสติก

ผลการวิจัยพบว่า 1)การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ จำนวน 27 ข้อ พบข้อสอบที่ทำหน้าที่ต่างกัน จำนวน 20 ข้อ มีขนาดอิทธิพล อยู่ระหว่าง .007 ถึง .044 ซึ่งเป็นขนาดอิทธิพลขนาดเล็ก แสดงว่าข้อสอบไม่ทำหน้าที่ต่างกัน 2) เมื่อแยกกลุ่มชาติพันธุ์ พบข้อสอบที่ทำหน้าที่ต่างกันส่วนใหญ่เป็นแบบเอกรูปลำเอียงเข้าข้างกลุ่มทวิภาษา 3) เมื่อวิเคราะห์ข้อสอบที่มีอิทธิพลขนาดกลางและขนาดใหญ่ เป็นข้อสอบที่ให้บอกความหมายของคำ เครื่องหมายสัญลักษณ์ และให้ตอบคำถามจากเรื่องที่ฟังดูและอ่าน ทำหน้าที่ต่างกันแบบเอกรูปลำเอียงเข้าหากลุ่มอ้างอิง เนื่องจากกลุ่มโรงเรียนทวิภาษามีการจัดกิจกรรมการเรียนการสอนที่เชื่อมโยงภาษาถิ่นสู่ภาษาราชการ ทำให้อ่านเขียนภาษาไทยได้ กล้าพูด กล้าแสดงออก มีความคิดสร้างสรรค์ เข้าใจในหลักภาษาไทย สรุปความได้เป็นอย่างดี มีความสุขในการเรียน จึงทำข้อสอบได้ดีกว่ากลุ่มโรงเรียนปกติ

คำสำคัญ: การทำหน้าที่ต่างกันของข้อสอบ, การรู้เรื่องการอ่าน, โรงเรียนทวิภาษา, ความยุติธรรมของแบบทดสอบ

---

* วิทยานิพนธ์หลักสูตรการวัดและประเมินผลการศึกษา วิชาเอกการประเมินการศึกษา สาขาวิชาศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช, 2562

[1] นักศึกษาหลักสูตรการวัดและประเมินผลการศึกษา วิชาเอกการประเมินการศึกษา สาขาวิชาศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช, E-mail: krut1540@hotmail.co.th

[2] ผู้ช่วยศาสตราจารย์, สาขาวิชาศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช

[3] ผู้ช่วยศาสตราจารย์, สาขาวิชาศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช

วารสารวิชาการและวิจัยสังคมศาสตร์
Social Sciences Research and Academic Journal

Analyzing Differential Item Functioning in Reading Literacy between Bilingual Schools and Regular Schools *

Wasant Nantasen[1] *Sungworn Ngudgratoke*[2] *Nalinee Na Nakorn*[3]

(Received: April 10, 2019; Revised: August 16, 2019; Accepted: August 23, 2019)

Abstract

The purposes of this research were: 1) Analyzing Differential Item Functioning in the Reading Literacy national test between Bilingual Schools and Regular Schools 2) Analyzing Differential Item Functioning between Bilingual Schools Hmong-Thai Mon- Thai Mala-Thai Pwo Karen-Thai Khmer-Thai and Regular Schools 3) The cause of Analyzing Differential Item Functioning

This research used secon dary data of all population students. The sample included 340 students studying 3 in Prathomsuksa 3 of 11 bilingual elementary school. and 339 students of 20 regular schools, total are 679 students. To used purposive sampling for resemble with bilingual school.

The results of this research were as follow:

1. The differential item functioning of 27 item according by using logistic regression showed that 20 item to measure by influence test during .007 to .044 which is a small effect sizes. That doesn't differential item functioning.

2. Differential item function test between regular schools and bilingual school. Different of ethnicity so that the most differential item functioning was uniform DIF with bilingual school. 3. The cause of differential item function by medium effect size and large effect size showed that meaning words, meaning signs and answer the questions from text. Most different item functioning were uniform, biased reference group. So that the bilingual schools understood about grammar thai language, good synopsis and interested more regular schools.

Keywords: Differential Item Functioning, Reading Literacy, Bilingual Schools, Fairness in Testing

---

* Research Article From Thesis For The Master Education Degree, Educational Evaluation Program, Sukhothai Thammathirat Open University, Thailant, 2019

[1] Student in Master Education Degree, Educational Evaluation, Sukhothai Thammathirat Open University, E-mail: krut1540@hotmail.co.th

[2] Assistant Professor, Educational Evaluation Program, Sukhothai Thammathirat Open University

[3] Assistant Professor, Educational Evaluation Program, Sukhothai Thammathirat Open University

ความสำคัญและปัญหาการวิจัย

สำนักทดสอบทางการศึกษา ได้ประเมินความสามารถพื้นฐานของนักเรียนชั้นประถมศึกษาปีที่ 3 ในด้านภาษา (Literacy) ด้านคำนวณ (Numeracy) และด้านเหตุผล (Reasoning Abilities) ปีการศึกษา 2557 ทุกคน ทุกโรงเรียน ให้มีการจัดสอบพร้อมกันทั่วประเทศ และให้นำผลการประเมินไปพัฒนาและยกระดับคุณภาพการศึกษา ปรับปรุงการจัดการเรียนการสอนของโรงเรียน และเป็นแนวทางการจัดการศึกษาของเขตพื้นที่การศึกษา ตามนโยบายสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สำนักทดสอบทางการศึกษา, 2557) ความยุติธรรม (Fairness) ของแบบทดสอบ ต้องไม่ลำเอียงไปยังนักเรียนกลุ่มใดกลุ่มหนึ่ง ต้องไม่ได้เปรียบหรือเสียเปรียบกันอันเนื่องมาจากสภาพหรือบริบทที่เปลี่ยนแปลงไป เช่น เพศ วัย วัฒนธรรม (รัชนีกูล ภิญโญภานุวัฒน์, 2554) ดังนั้นการสร้างแบบทดสอบต้องคำนึงถึงความตรง (Validity) ซึ่งเป็นคุณสมบัติที่สำคัญของแบบทดสอบวัดผลสัมฤทธิ์ แบบทดสอบทุกฉบับจะต้องมีคุณภาพด้านความตรง จึงจะเชื่อถือได้ว่าเป็นแบบทดสอบที่ดี และผลที่ได้จากการวัดจะถูกต้องตามที่ต้องการ (ล้วน สายยศ, 2543) และความตรงยังเปรียบเสมือนหัวใจของการสร้างแบบทดสอบ ทำให้แบบทดสอบมีคุณภาพและเพิ่มความน่าเชื่อมากยิ่งขึ้น แต่ถ้าแบบทดสอบใดประกอบด้วยแบบสอบที่ทำหน้าที่ต่างกันเกิดความอยุติธรรมของข้อสอบหรือแบบสอบ หรือทำให้กลุ่มผู้สอบมีความได้เปรียบหรือเสียเปรียบกันทั้งที่มีความสามารถจริงเท่าเทียมกัน (ศิริชัย กาญจนวาสี, 2555) ถือได้ว่าแบบทดสอบนั้นไม่มีคุณภาพ ขาดความตรงและขาดความน่าเชื่อถือ ส่งผลถึงการเปรียบเทียบคุณภาพการจัดการศึกษาของแต่ละโรงเรียนด้วย

การสร้างแบบทดสอบ ความตรง (validity) และความยุติธรรม (Fairness) ควรจะเป็นคุณภาพของแบบทดสอบที่สำคัญยิ่ง แบบทดสอบต้องไม่ทำหน้าที่ต่างกัน ไม่มีความลำเอียง ไปยังนักเรียนกลุ่มใดกลุ่มหนึ่ง เชื้อชาติ ภาษา วัฒนธรรม เพศ แบบทดสอบต้องวัดได้ตรงตามวัตถุประสงค์ และนำไปวัดกับผู้สอบที่มีความสามารถเท่าเทียมกันให้ผลการสอบไม่ต่างกันและคะแนนที่ได้จากแบบทดสอบต้องทำนายความสามารถของนักเรียนได้อย่างแม่นยำ (ศิริชัย กาญจนวาสี, 2545) หากทำนายผลการสอบได้ยังไม่แม่นยำ จะทำให้ไม่สามารถเปรียบเทียบผลการสอบได้อย่างแท้จริงเนื่องจากยังมีลักษณะแฝงในตัวนักเรียนอีกหลายลักษณะ แบบทดสอบเกิดความลำเอียง (Item Bias) หรือแบบทดสอบไม่มีความยุติธรรม (Item Unfair) ซึ่งต่อมามีการเปลี่ยนมาใช้คำที่เหมาะสมกว่าเป็น การทำหน้าที่ต่างกันของข้อสอบหรือการทำหน้าที่ต่างกันของแบบสอบ (ศิริชัย กาญจนวาสี, 2555)

การทำหน้าที่ต่างกันของข้อสอบ (diferential item functioning; DIF) Camilli & Shepard (1994) ให้ความหมายว่า เป็นความคลาดเคลื่อนอย่างเป็นระบบหรือความไม่เที่ยงของการวัด ทำให้สมาชิกบางกลุ่มได้ประโยชน์หรือเสียประโยชน์ McCallon & Schumacker (2002) ให้ความหมายว่า เป็นลักษณะที่

แสดงถึงความไม่ยุติธรรมของข้อสอบจะเกิดขึ้นในลักษณะที่ผู้สอบมีความรู้ความสามารถเท่ากัน แต่มาจากกลุ่มที่ต่างกัน และได้รับการตอบสนองจากข้อสอบมีโอกาสตอบถูกต่างกัน สังวรณ์ งัดกระโทก (2559) ให้ความหมายว่า ข้อสอบที่กลุ่มผู้สอบที่มีความสามารถเท่ากันตอบต่างกัน และศิริชัย กาญจนวาสี (2555) ได้ให้ความหมายว่า การที่ข้อสอบทำให้ผู้สอบต่างจากกลุ่มกันที่มีความสามารถหรือคุณลักษณะที่มุ่งวัดเท่ากัน มีโอกาสในการตอบข้อสอบได้ถูกต้องแตกต่างกัน หรือมีฟังก์ชั้นการตอบสนองข้อสอบแตกต่างกัน เมื่อนำข้อสอบไปทดสอบกับผู้สอบกลุ่มย่อยต่างกัน ที่มีความสามารถหลัก (Primary ability) ระดับเดียวกันหรือมีคุณลักษณะแฝง (Latent trait) ที่ต้องการวัดเท่ากัน แต่มีความสามารถรอง (Secondary ability) แตกต่างกัน ทำให้ผู้สอบต่างกลุ่มที่นำมาจับคู่เปรียบเทียบมีโอกาสตอบข้อสอบถูกแตกต่างกัน

การศึกษางานวิจัยที่ผ่านมายังไม่พบงานวิจัยที่เกี่ยวกับการประเมินคุณภาพการศึกษาขั้นพื้นฐาน (NT) โดยที่นักเรียนมีความแตกต่างกันทางด้านภาษา ชาติพันธุ์ วัฒนธรรม ภูมิประเทศ และกิจกรรมการจัดการเรียนการสอน ซึ่งส่งผลให้มีการได้เปรียบเสียเปรียบในการสอบเกิดขึ้น

จากเหตุผลดังกล่าวข้างต้น ทำให้ผู้วิจัยสนใจที่จะศึกษาตรวจสอบการทำหน้าที่ต่างกันของข้อสอบวัดผลสัมฤทธิ์ทางการเรียนระดับการศึกษาขั้นพื้นฐาน NT ปีการศึกษา 2557 ด้านการรู้เรื่องการอ่าน (Reading Literacy) ระหว่างโรงเรียนประถมศึกษาที่จัดการศึกษาแบบทวิภาษา และโรงเรียนปกติ ด้วยวิธีถดถอยโลจิสติก (Logistic Regression) เพื่อใช้เป็นแนวทางสำหรับการปรับปรุงกระบวนการสร้างข้อสอบ อันเป็นแนวทางในการพัฒนาและปรับปรุงการสร้างข้อสอบให้มีคุณภาพต่อไป

วัตถุประสงค์ของการวิจัย

1. เพื่อวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบวัดความสามารถด้านการรู้เรื่องการอ่าน ของการทดสอบระดับชาติ ระหว่างโรงเรียนประถมศึกษาที่จัดการศึกษาแบบทวิภาษา และโรงเรียนปกติ

2. เพื่อวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบวัดความสามารถด้านการรู้เรื่องการอ่าน ระหว่างโรงเรียนประถมศึกษาที่จัดการศึกษาแบบทวิภาษากลุ่ม ไทย-ม้ง ไทย-มอญ ไทย-มลายูถิ่น ไทย-กะเหรี่ยงโปว์ ไทย-เขมรถิ่นไทยและโรงเรียนปกติ

3. เพื่อวิเคราะห์สาเหตุของการทำหน้าที่ต่างกันของข้อสอบ

ระเบียบวิธีวิจัย

การวิจัยในครั้งนี้เป็นการศึกษาการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีถดถอยโลจิสติก (Logistic Regression : LR) ใช้ข้อมูลทุติยภูมิ (Secondary data) จากสำนักทดสอบทางการศึกษาของสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) เป็นผลการตอบแบบทดสอบระดับชาติ (NT) ของนักเรียนชั้นประถมศึกษาปีที่ 3 ด้านการรู้เรื่องการอ่าน (Reading Literacy)

กลุ่มตัวอย่าง

นักเรียนร่วมโครงการพัฒนาคุณภาพการเรียนการสอนภาษาไทยโดยนำภาษาท้องถิ่นร่วมจัดการเรียนรู้ (ทวิภาษาเต็มรูป) ร่วมกับมูลนิธิภาษาศาสตร์ประยุกต์ และมหาวิทยาลัยมหิดล ปีงบประมาณ 2554 ซึ่งเลือกแบบเจาะจง (Purposive sampling) จำนวน 11 โรงเรียน นักเรียน 340 คน และโรงเรียนที่จัดการเรียนการสอนปกติตามที่มีบริบท พื้นที่ตั้ง ใกล้เคียงกับโรงเรียนทวิภาษา จำนวน 17 โรง นักเรียน 339 คน

เครื่องมือที่ใช้ในการวิจัย

แบบบันทึกผลการทดสอบวัดความสามารถด้านการรู้เรื่องการอ่าน ของนักเรียนชั้นประถมศึกษาปีที่ 3 ปีการศึกษา 2557 ของสำนักทดสอบทางการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน

การจัดเตรียมข้อมูล

กำหนดกลุ่มโรงเรียนออกเป็น 2 กลุ่ม

1. กลุ่มอ้างอิง (Reference Group; R) กำหนดให้เป็นโรงเรียนประถมศึกษาที่จัดการศึกษาแบบทวิภาษา เป็นกลุ่มที่คาดว่าจะได้ประโยชน์จากการตอบข้อสอบได้ถูกต้อง แทนด้วย 1 จำนวน 340 คน

2. กลุ่มเปรียบเทียบ (Focal Group; F) กำหนดให้เป็นโรงเรียนประถมศึกษาที่จัดการศึกษาแบบปกติเป็นกลุ่มที่เสียประโยชน์ในการตอบข้อสอบ แทนด้วย 0 จำนวน 339 คน

การวิเคราะห์ข้อมูล

1. การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีถดถอยโลจิสติก

ใช้โมเดลของซัมโบ (Zumbo, 1999) โดยใช้โปรแกรม SPSS ในการคำนวณ เพื่อเปรียบเทียบผลการตอบข้อสอบ 2 กลุ่ม (Group) โดยกลุ่มอ้างอิงคือประเภทโรงเรียนทวิภาษา (Reference group หรือกลุ่ม R แทนด้วย 1) และกลุ่มเปรียบเทียบคือประเภทโรงเรียนปกติ (Focal group หรือกลุ่ม F แทนด้วย 0) เป็นตัวแปรตาม (Dependent variable) คะแนนรวมสำหรับผู้เข้าสอบแต่ละคน เป็นตัวแปร TOT (Total test score) และปฏิสัมพันธ์ระหว่างคะแนนรวมกับประเภทโรงเรียน (TOT* GROUP) เป็นตัวแปรอิสระ (Independent variable ดังนี้

$$Y = b_0 + b_1TOT + b_2\ GROUP + b_3TOT^*\ GROUP$$

เมื่อ Y แทน คะแนนรายข้อ

b_0 แทน ค่าคะแนนจุดตัด (Intercept) ที่แกน Y รายข้อในกรณีที่ความสามารถของผู้เข้าสอบมีค่าเป็น 0

b_1 แทน สัมประสิทธิ์การถดถอยแสดงอิทธิพลต่อคะแนนรวมที่มีผลต่อคะแนนรายข้อ

b_2	แทน	สัมประสิทธิ์การถดถอยแสดงอิทธิพลต่อประเภทโรงเรียน
b_3	แทน	สัมประสิทธิ์การถดถอยแสดงอิทธิพลต่อปฏิสัมพันธ์ระหว่างระดับความสามารถของผู้เข้าสอบกับประเภทโรงเรียน
TOT	แทน	ระดับความสามารถของผู้เข้าสอบ (คะแนนรวม) (Total test score)
GROUP	แทน	ประเภทโรงเรียนทวิภาษาและประเภทโรงเรียนปกติ
TOT* GROUP	แทน	ปฏิสัมพันธ์ระหว่างระดับความสามารถของผู้สอบกับประเภทโรงเรียน

ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีถดถอยโลจิสติกแบบจัดอันดับ (OLR) จะใช้กระบวนการตรวจสอบความเหมาะสมของข้อมูลกับโมเดลจะทดสอบโมเดล 3 ขั้นตอน

ขั้นตอนที่ 1 $Y = b_0 + b_1TOT$ ..... (Step 1)

ขั้นตอนที่ 2 $Y = b_0 + b_1TOT + b_2SCHOOl$ ..... (Step 2)

ขั้นตอนที่ 3 $Y = b_0 + b_1TOT + b_2SCHOOl + b_3TOT* SCHOOl$ ..... (Step 3)

2. ตรวจสอบข้อสอบทำหน้าที่ต่างกันหรือไม่ (DIF/ NO DIF)

โดยการนำค่าไค-สแควร์ (χ^2) ที่ได้จากการคำนวณ เป็นรายข้อขั้นตอนที่ 2 (Step 2) ลบด้วยค่าไค-สแควร์ ขั้นตอนที่ 1 (Step 1) ดังนี้

$\chi^2_{step2} - \chi^2_{step1}$ นำผลที่ได้มาแปลผล ดังนี้ ($\chi^2_{วิกฤต} = 3.84$)

ถ้า $\chi^2_{คำนวณ} < \chi^2_{วิกฤต}$ ➔ ยอมรับ Step 1 หมายความว่า ค่าไค-สแควร์ ที่คำนวณได้เมื่อนำไปเทียบกับตารางค่าวิกฤตการแจกแจงความน่าจะเป็นแบบไค-สแควร์ ที่ระดับ .05 มีค่าน้อยกว่า 3.84 แสดงว่าข้อสอบข้อนั้น **ไม่ทำหน้าที่ต่างกัน (NO DIF)**

ถ้า $\chi^2_{คำนวณ} > \chi^2_{วิกฤต}$ ➔ ยอมรับ Step 2 หมายความว่า ค่าไค-สแควร์ ที่คำนวณได้เมื่อนำไปเทียบกับตารางค่าวิกฤตการแจกแจงความน่าจะเป็นแบบไค-สแควร์ ที่ระดับ .05 มีค่ามากกว่า 3.84 แสดงว่าข้อสอบข้อนั้น **ทำหน้าที่ต่างกัน (DIF)**

3. ตรวจสอบประเภทของการทำหน้าที่ต่างกันของข้อสอบ (Uniform DIF/ Nonuniform DIF)

นำข้อสอบที่ทำหน้าที่ต่างกันมาตรวจสอบประเภทของการทำหน้าที่ต่างกัน โดยนำค่าไค-สแควร์ (χ^2) ที่ได้จากการคำนวณ เป็นรายข้อขั้นตอนที่ 3 (Step 3) ลบด้วยค่าไค-สแควร์ ขั้นตอนที่ 2 (Step 2) ดังนี้

$\chi^2_{step3} - \chi^2_{step2}$ นำผลที่ได้มาแปลผล ดังนี้ ($\chi^2_{วิกฤต} = 3.84$)

ถ้า $\chi^2_{คำนวณ} < \chi^2_{วิกฤต}$ → ยอมรับ Step 2 หมายความว่า ค่าไค-สแควร์ ที่คำนวณได้เมื่อนำไปเทียบกับตารางค่าวิกฤตการแจกแจงความน่าจะเป็นแบบไค-สแควร์ ที่ระดับ .05 มีค่าน้อยกว่า 3.84 แสดงว่าข้อสอบข้อนั้นทำหน้าที่ต่างกันแบบ**เอกรูป (Uniform DIF)**

ถ้า $\chi^2_{คำนวณ} > \chi^2_{วิกฤต}$ → ยอมรับ Step 3 หมายความว่า ค่าไค-สแควร์ ที่คำนวณได้เมื่อนำไปเทียบกับตารางค่าวิกฤตการแจกแจงความน่าจะเป็นแบบไค-สแควร์ ที่ระดับ .05 มีค่ามากกว่า 3.84 แสดงว่าข้อสอบข้อนั้นทำหน้าที่ต่างกันแบบ**อเนกรูป (Nonuniform DIF)**

4. หาขนาดของอิทธิพล (Effect Size)

ขั้นตอนที่ 1 นำค่า R^2 ของ Step 2 ลบด้วย R^2 ของ Step 1 ซึ่งเป็นผลมาจากอิทธิพลของกลุ่ม (ความแตกต่างของประเภทโรงเรียน) ตรวจสอบการทำหน้าที่ต่างกันแบบเอกรูป ดังนี้

$$R^2_{uniform} = R^2_2 - R^2_1$$

ขั้นตอนที่ 2 นำค่า R^2 ของ Step 3 ลบด้วย R^2 ของ Step 2 ซึ่งเป็นผลมาจากอิทธิพลของปฏิสัมพันธ์ของประเภทโรงเรียนและความสามารถของนักเรียน ตรวจสอบการทำหน้าที่ต่างกันแบบอเนกรูป ดังนี้

$$R^2_{nonuniform} = R^2_3 - R^2_2$$

6. การวิเคราะห์ขนาดอิทธิพล (Effect Size)

การวิเคราะห์ขนาดอิทธิพล โดยที่นำผลลัพธ์ R^2 ที่ได้มาตัดสินตามเกณฑ์ (Zumbo, 1999) มีความเข้ม 3 ระดับ คือ

1) ขนาดอิทธิพลมีขนาดเล็ก (Small effect size) เมื่อมีค่าความแตกต่าง $R^2 < 0.13$

2) ขนาดอิทธิพลมีขนาดปานกลาง (Medium effect size) เมื่อมีค่าความแตกต่าง R^2 $0.13 \geq$ และ < 0.26

3) ขนาดอิทธิพลมีขนาดใหญ่ (Large effect size) เมื่อมีค่าความแตกต่าง $R^2 \geq 0.26$

สรุปผลการวิจัย

1. การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบวัดความสามารถด้านการรู้เรื่องการอ่านของการทดสอบระดับชาติ ระหว่างโรงเรียนประถมศึกษาที่จัดการศึกษาแบบทวิภาษาและโรงเรียนปกติ จำนวน 27 ข้อ เมื่อวิเคราะห์ด้วยวิธีถดถอยโลจิสติก พบข้อสอบที่ทำหน้าที่ต่างกัน จำนวน 17 ข้อ นำมาวัดขนาดอิทธิพลมีขนาดเล็ก (Small effect size : < 0.13) แสดงว่าข้อสอบ ไม่ทำหน้าที่ต่างกัน ไม่ลำเอียงเข้าข้างกลุ่มหนึ่งกลุ่มใด ข้อสอบมีความยุติธรรม

2. การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบวัดความสามารถด้านการรู้เรื่องการอ่านระหว่างโรงเรียนประถมศึกษาที่จัดการศึกษาแบบทวิภาษากลุ่ม ไทย-ม้ง ไทย-มอญ ไทย-มลายูถิ่น ไทย-กะเหรี่ยงโปว์ ไทย-เขมรถิ่นไทยและโรงเรียนปกติ เมื่อวิเคราะห์ด้วยวิธีถดถอยโลจิสติก แสดงดังตารางที่ 1

ข้อ	ประเภทโรงเรียนแบบทวิภาษา (*N*=679)	ประเภทโรงเรียนทวิภาษา กลุ่มไทย				
		ม้ง (*N*=214)	มอญ (*N*=193)	มลายูถิ่น (*N*=174)	กะเหรี่ยงโปว์ (*N*=41)	เขมรถิ่นไทย (*N*=57)
1		อเนกรูป		อเนกรูป	*เอกรูป	อเนกรูป
2	*เอกรูป	อเนกรูป		อเนกรูป		
3	*เอกรูป	อเนกรูป		อเนกรูป	*เอกรูป	*เอกรูป
4	*เอกรูป	อเนกรูป			อเนกรูป	อเนกรูป
5	อเนกรูป	อเนกรูป				
6	อเนกรูป	อเนกรูป		*เอกรูป		อเนกรูป
7						
8	*เอกรูป	อเนกรูป				*เอกรูป
9	อเนกรูป	อเนกรูป		อเนกรูป		*เอกรูป
10	อเนกรูป	อเนกรูป		*เอกรูป		อเนกรูป
11	*เอกรูป	*เอกรูป				
12	อเนกรูป	*เอกรูป				
13	อเนกรูป	*เอกรูป				
14	อเนกรูป	*เอกรูป				
15		*เอกรูป		อเนกรูป		
16	อเนกรูป	*เอกรูป				
17	*เอกรูป	*เอกรูป	อเนกรูป	อเนกรูป		
18	อเนกรูป	*เอกรูป				
19	*เอกรูป	*เอกรูป				
20	อเนกรูป	*เอกรูป				
21	อเนกรูป	*เอกรูป				
22	อเนกรูป	*เอกรูป	อเนกรูป			
23	*เอกรูป	*เอกรูป		อเนกรูป	อเนกรูป	
24		*เอกรูป				
25	อเนกรูป	*เอกรูป				
26		*เอกรูป		อเนกรูป		
27	*เอกรูป	*เอกรูป				
รวม	22	26	2	10	4	7

***** ข้อสอบลำเอียงเข้าข้างประเภทโรงเรียนทวิภาษา

3. การวิเคราะห์สาเหตุการทำหน้าที่ต่างกันของข้อสอบวัดความสามารถด้านการรู้เรื่องการอ่านระหว่างโรงเรียนประถมศึกษาที่จัดการศึกษาแบบทวิภาษา และโรงเรียนปกติ พบว่า

3.1 การทำหน้าที่ต่างกันของข้อสอบ ระหว่างโรงเรียนแบบทวิภาษา และโรงเรียนปกติ ตรวจไม่พบการทำหน้าที่ต่างกันของข้อสอบ อาจเนื่องมาจากกลุ่มตัวอย่างมีความหลากหลาย มีทุกภูมิภาค คะแนนสอบมีการกระจาย และกลุ่มตัวอย่างเกิน 200 คน และถ้ากลุ่มตัวอย่างยิ่งมากจะทำให้การแปรผลมีประสิทธิภาพมากยิ่งขึ้น (Holland and Thayer, 1988)

3.2 การวิเคราะห์สาเหตุการทำหน้าที่ต่างกันของข้อสอบวัดความสามารถด้านการรู้เรื่องการอ่าน ระหว่างโรงเรียนประถมศึกษาที่จัดการศึกษาแบบทวิภาษากลุ่มไทย-ม้ง ไทย-มอญ ไทย-มลายูถิ่น ไทย-กะเหรี่ยงโปว์ ไทย-เขมรถิ่นไทย และโรงเรียนปกติ โดยเลือกเฉพาะข้อสอบที่มีขนาดอิทธิพลขนาดกลางและขนาดใหญ่ สาเหตุเกิดจาก

3.2.1 หลักสูตรแกนกลางการศึกษาขั้นพื้นฐานเป็นหลักสูตรที่ใช้กับโรงเรียนทั่วประเทศ ไม่ว่าจะมีภูมิประเทศ ภาษา วัฒนธรรม ที่แตกต่างกัน ก็ต้องเรียนหนังสือที่กระทรวงศึกษาธิการ บังคับให้ใช้ โดยเฉพาะวิชาภาษาไทยต้องมี หนังสือวรรณคดีลำนำ ภาษาพาที แบบฝึกหัดทักษะภาษา โรงเรียนทุกโรงเรียนต้องจัดซื้อให้นักเรียนยืมเรียนทุกคน การสอนก็ต้องสอนตามที่หนังสือกำหนด ทำให้นักเรียนที่ไม่เข้าใจภาษาไทย พูดภาษาไทยไม่ได้ จึงไม่อยากเรียน ไม่ตั้งใจเรียน ครูก็เข้าสอนตามหน้าที่ สื่อสารกับนักเรียนไม่เข้าใจ ท้อแท้กับการทำงาน ถึงเวลาย้ายก็เขียนย้ายไปโรงเรียนที่ดีกว่า ทำให้เด็กที่อยู่ตามแนวชายแดนเรียนไม่จบตามหลักสูตรภาคบังคับ รัฐบาลก็จัดการศึกษาแบบศูนย์เปล่า

3.2.2 โครงการพัฒนารูปแบบการจัดการเรียนรู้ภาษาไทยโดยนำภาษาถิ่นร่วมจัดการเรียนรู้ โครงการทวิภาษา (เต็มรูป) ของสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน ร่วมกับมูลนิธิภาษาศาสตร์ประยุกต์ มหาวิทยาลัยมหิดล ได้พัฒนากระบวนการสอนภาษาไทย มีความมุ่งหมายให้เกิดความสำเร็จในการจัดการเรียนรู้ โดยจ้างครูท้องถิ่นที่จบชั้นมัธยมศึกษาปีที่ 3 และ 6 ที่มีความรู้ 2 ภาษา (ภาษาไทย-ภาษาท้องถิ่น) ให้เป็นครูช่วยสอนหรือดำเนินการสอนร่วมกับครูประจำการ มีการสร้างสื่อการเรียนรู้ที่เหมาะสมกับความสามารถทางภาษาของนักเรียนซึ่งใช้ภาษาท้องถิ่นแตกต่างจากภาษาไทย และยังมีการฝึกอบรมด้านความรู้ และกระบวนการสอนภาษาไทยเป็นภาษาที่สอง (ฟัง-พูด TPR / Total Physical Response) การฝึกจัดทำสื่อการเรียนรู้ภาษาไทย ภาษาท้องถิ่น สื่อเชื่อมโยงภาษาทั้งสอง รวมถึงการจัดทำหลักสูตรสถานศึกษา แผนการสอน และการฝึกอบรมครูให้สามารถแต่งเพลงและใช้เพลงเป็นสื่อในการจัดการเรียนรู้ เริ่มจัดการเรียนการสอนแบบทวิภาษาตั้งแต่ชั้นอนุบาล 1 พอถึงชั้นประถมศึกษาปีที่ 2 ก็สามารถอ่านเขียนภาษาไทยได้ จับใจความสำคัญ สรุปความจากเรื่องที่อ่าน จากการจัดการเรียนการสอนที่ดีมีคุณภาพ ทำให้ข้อสอบวัดความสามารถด้านภาษา (NT) ชั้นประถมศึกษาปีที่ 3 ปีการศึกษา 2557

ทำหน้าที่ต่างกัน โดยลำเอียงเข้าหากลุ่มโรงเรียนทวิภาษา โครงการพัฒนารูปแบบการจัดการเรียนรู้ภาษาไทยโดยนำภาษาถิ่นร่วมจัดการเรียนรู้ ยังได้รับรับรางวัล นวัตกรรมการรู้หนังสือ “UNESCO King Sejong Literacy Prize 2016” ซึ่งองค์การยูเนสโกมอบแด่ สถาบันวิจัยภาษาและวัฒนธรรมเอเชีย มหาวิทยาลัยมหิดล ณ สำนักงานใหญ่องค์การยูเนสโก กรุงปารีส ประเทศฝรั่งเศส และมหาวิทยาลัยราชภัฏยะลา ได้รับรางวัลนวัตกรรมเหวินฮุยด้านการส่งเสริมพัฒนาศักยภาพครู ประจำปี พ.ศ. 2560 จากองค์การยูเนสโก (UNESCO Wenhui (文晖) Award 2017) จากการฝึกอบรมและพัฒนาครูตามแบบทวิ-พหุภาษา ซึ่งเป็นมหาวิทยาลัยแรกในภูมิภาคเอเชียที่มีการสอนและฝึกอบรมครูตามแบบทวิ-พหุภาษา (องค์การทุนเพื่อเด็กแห่งสหประชาชาติ, 2561)

อภิปรายผลการวิจัย

1. ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ จำนวน 27 ข้อ พบข้อสอบที่ทำหน้าที่ต่างกัน จำนวน 17 ข้อ สอดคล้องกับ Kim, Chosen, Alagoz and Kim (2007); ธเกียรติกมล ทองงอก (2554) พบว่า เมื่อกลุ่มตัวอย่างมีขนาดใหญ่จะมีความไวในการตรวจพบการทำหน้าที่ต่างกันของข้อสอบค่อนข้างสูง เพื่อแก้ไขความคลาดเคลื่อนต่างๆ ที่อาจเกิดขึ้น จึงต้องใช้การวัดขนาดอิทธิพล (effect size measure) มาช่วยในการตัดสินประสิทธิภาพข้อสอบเพื่อให้มีความถูกต้องรัดกุมยิ่งขึ้น จึงนำข้อสอบที่ทำหน้าที่ต่างกันมาวัดขนาดอิทธิพลตามเกณฑ์ (Zumbo & Thomas, 1999) ผลปรากฏว่าข้อสอบทั้ง 17 ข้อ มีขนาดอิทธิพลขนาดเล็ก เป็นข้อสอบที่ได้มาตรฐาน ซึ่งเกิดจากสำนักทดสอบทางการศึกษาแห่งชาติ มีการพัฒนาเครื่องมือทดสอบที่ได้มาตรฐาน เป็นขั้นตอน ใช้ผู้ที่มีความรู้ มีความสามารถ มีความเข้าใจและเป็นผู้ที่รักษาความลับหรือจัดความสำคัญของการปฏิบัติงานได้อย่างเหมาะสมประกอบด้วย ผู้เชี่ยวชาญด้านเนื้อหาสาระ ผู้สอนในวิชานั้น ผู้เชี่ยวชาญด้านข้อสอบ และผู้เชี่ยวชาญด้านการวัดผล (สัมพันธ์ พันธุ์พฤกษ์. มปป.) ผลการวิจัยจึงไมสอดคล้องกับสมติฐานที่ตั้งไว้ว่าจะพบข้อสอบที่ทำหน้าที่ต่างกัน

2. ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ ระหว่างโรงเรียนทวิภาษากลุ่ม ไทย-ม้ง ไทย-มอญ ไทย-มลายูถิ่น ไทย-กะเหรี่ยงโปว์ ไทย-เขมรถิ่นไทยและโรงเรียนปกติ พบว่า ข้อสอบส่วนใหญ่ทำหน้าที่ต่างกันแบบเอกรูป (Uniform) ลำเอียงเข้าหากลุ่มอ้างอิง กล่าวคือ กลุ่มโรงเรียนทวิภาษามีโอกาสในการตอบข้อสอบถูกมากกว่า ในทุกระดับความสามารถ ซึ่งเกิดจากการเรียนรู้ภาษาของเด็กชาติพันธุ์ในหลักการสอนทวิภาษา (Dual Language Instruction : DLI) ที่นำมาใช้กับเด็กชาติพันธุ์ที่มีภาษาแม่ (Mother Language) ที่แตกต่างจากภาษาที่ใช้ในโรงเรียนจะทำให้เด็กมีกลไกทางสมองและระบบการคิดอันเกี่ยวโยงกับพัฒนาการสติปัญญาและการเรียนรู้ซึ่งอาศัยภาษาแม่-วัฒนธรรมและสิ่งแวดล้อมร่วมด้วย (ศิริพร หมั่นงาน, 2557) เมื่อทวิภาษาใช้ภาษาแม่เป็นฐานในการจัดการเรียนรู้ จึงเหมาะสมอย่างยิ่งกับผู้เรียนกลุ่มชาติพันธุ์ที่เป็นเด็กสองภาษาซึ่งมีภาษาที่ 1 (ภาษาแม่) เป็นต้นทุนเด็กติดตัวมาตั้งแต่เกิด เด็กจึงสามารถ

เรียนรู้จากการใช้ภาษาที่ 1 เป็นเครื่องมือถ่ายโอนไปสู่ภาษาที่ 2 หรือภาษาราชการได้ดี (สถาบันวิจัยภาษาและวัฒนธรรมเอเชีย. 2553; สุวิไล เปรมศรีรัตน์. 2552; สำนักเลขาธิการสภาการศึกษา. 2552; ชาตรี สำราญ. 2543)

3. เมื่อวิเคราะห์สาเหตุของการทำหน้าที่ต่างกันของข้อสอบ พบว่าเป็นข้อสอบที่ให้บอกความหมายของคำ บอกความหมายของเครื่องหมายสัญลักษณ์ และให้ตอบคำถามจากเรื่องที่ฟังดูและอ่าน ข้อสอบส่วนใหญ่ทำหน้าที่ต่างกันแบบเอกรูป ลำเอียงเข้าหากลุ่มอ้างอิง เมื่อวิเคราะห์สาเหตุพบว่าโรงเรียนทวิภาษาใช้คนกลางมาเป็นสะพานเชื่อมโยงความเข้าใจ จึงเป็นที่มาของการดึงครูท้องถิ่น มาเป็นพี่เลี้ยงร่วมกับครูไทยเริ่มจากช่วงปฐมวัย สอนการสื่อความหมาย และทำความเข้าใจกับสิ่งที่เด็กอยากจะสื่อสารให้ตรงกัน โดยใช้ชีวิตประจำวัน และประเพณีในท้องถิ่นมาอธิบาย ก่อนจะปูพื้นฐานสู่การอ่านเขียนในช่วงชั้นที่สูงขึ้น มีเทคนิคการเทียบเสียงในภาษาไทยและภาษาท้องถิ่นที่ใกล้เคียงกัน เพื่อให้เด็กเข้าใจความหมายให้ได้ก่อนด้วยอักษรภาษาไทย การเรียนทวิภาษา จึงมีเป้าหมายที่ใช้ภาษาแม่ หรือภาษาถิ่นเป็นพื้นฐาน และเป็นสะพานเชื่อมไปสู่การเรียนภาษาไทย ขณะเดียวกันก็จะทำให้เด็กคงความเป็นอัตลักษณ์ของตัวเองด้วยจึงทำให้เด็กเข้าใจในหลักภาษาไทย สรุปความได้ดี มากกว่าเด็กโรงเรียนปกติ สำนักงานเลขาธิการสภาการศึกษา (2552) ให้เหตุผลว่าโรงเรียนที่เข้าร่วมโครงการครูจะได้เข้าใจลักษณะเฉพาะของภาษา ความแตกต่างทางภาษา และสามารถนำความรู้ทางภาษามาใช้จัดการเรียนการสอน และวัดผลประเมินผลได้อย่างสอดคล้องกับโรงเรียนและพื้นที่ สอดคล้องกับ วรรณา เทียนมี (2558) กล่าวว่า กระบวนการเรียนรู้ด้านภาษานั้นต้องเริ่มต้นจากการฟังที่เริ่มจากคนใกล้ตัว หัดพูดเป็นคำ และพูดเป็นประโยคในที่สุด หากเด็กมาเริ่มต้นเรียนภาษาจากการอ่านและเขียนจะต้องเป็นการเรียนภาษาที่สอง สอดคล้องกับ อมรวิชช์ นาครทรรพ (2558) กล่าวว่า เป้าหมายของการรู้ภาษา ไม่ใช่เพียงแค่การอ่านออกเขียนได้แต่ภาษาจะเป็นเครื่องมือที่จะนำไปสู่การสื่อสารอื่นๆ เป็นการสร้างสังคมแห่งการเรียนรู้ ให้เกิดความปรองดองท่ามกลางความหลากหลายของวัฒนธรรม ดังนั้นเด็กนักเรียนโรงเรียนทวิภาษาจึงมีความได้เปรียบด้านการเรียน เข้าใจภาษาไทยได้ดีกว่าเด็กในโรงเรียนที่จัดการเรียนการสอนแบบปกติ ทำให้ได้คะแนนในการสอบ NT ปีการศึกษา 2557 สูงกว่าโรงเรียนปกติ

ข้อเสนอแนะ

1. ควรทำการศึกษาซ้ำกับกลุ่มตัวอย่าง และวิธีการเดิม แต่วัดความสามารถด้านคำนวณ และด้านเหตุผล เพื่อตรวจสอบความเห็นของผู้เชี่ยวชาญ ที่เห็นว่าการจัดการเรียนการสอนแบบทวิภาษามีผลทำให้ข้อสอบทำหน้าที่ต่างกัน โดยเฉพาะกลุ่มไทย-ม้ง

2. ควรศึกษาการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีอื่นๆ เพื่อยืนยันความถูกต้องของข้อมูล เช่น การวิเคราะห์การถดถอยโลจิสติกพหุวิภาค การวิเคราะห์ความแปรปรวน วิธีไคสแควร์ของลอร์ด วิธีดัชนีมาตรฐานพหุวิภาค วิธีซิปเทสท์ เป็นต้น

เอกสารอ้างอิง

ชาตรี สำราญ. (2543). *ครูรู้ได้อย่างไรว่าเด็กเกิดการเรียนรู้.* กรุงเทพฯ: มูลนิธิสดศรี – สกฤษดิ์วงศ์.

ธเกียรติกมล ทองงอก. (2554). *ประสิทธิภาพการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในวิธีถดถอยโลจิสติกโดยใช้เกณฑ์ขนาดอิทธิพล 2 วิธี สำหรับข้อสอบที่มีรูปแบบการตรวจให้คะแนนแบบทวิวิภาค : ข้อมูลจำลองและข้อมูลเชิงประจักษ์.* วิทยานิพนธ์ปริญญาครุศาสตรดุษฎีบัณฑิต. จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ.

รัชนีกูล ภิญโญภานุวัฒน์. (2554). *การสร้างเครื่องมือวัดด้านพุทธพิสัย.* ใน เอกสารการสอนชุดวิชาการวิจัยและสถิติ ทางการศึกษา (หน่วยที่ 6). นนทบุรี: สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมาธิราช.

ล้วน สายยศ และอังคณา สายยศ. 2543. *เทคนิคการวิจัยทางการศึกษา.* (พิมพ์ครั้งที่ 3). กรุงเทพฯ: สุวีริยาสาส์น.

วรรณา เทียนมี. (2558). *หลักการและเป้าหมายของการจัดการศึกษาแนวทางทวิภาษาเต็มรูป.* สืบค้นเมื่อ 30 ธันวาคม 2559, จาก http://www.academic.obec.go.th/web/static/d/516

ศิริพร หมั่นงาน. (2557). *ศึกษาความสามารถอ่านและเขียนภาษาไทยของนักเรียนชั้นประถมศึกษาร่วมกลุ่มชาติพันธุ์ญัฮกุรที่บกพร่องทางการเรียนรู้ จากการสอนอ่านและเขียนภาษาไทยตามแนวพหุปัญญาร่วมกับการสอนทวิภาษาและสื่อบทเรียนการ์ตูนสองภาษา (ไทย-ญัฮกุร).* ใน ชนิตา มิตรานันท์ (บรรณาธิการ), วารสารการวิจัยและพัฒนาการศึกษาพิเศษ, 80-88. กรุงเทพฯ: มหาวิทยาลัยศรีนครินทรวิโรฒ.

ศิริชัย กาญจนวาสี. (2545). *ทฤษฎีการทดสอบแนวใหม่.* (พิมพ์ครั้งที่ 3). กรุงเทพฯ: คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2555). *ทฤษฎีการทดสอบแนวใหม่ (MODERN TEST THEORIES).* (พิมพ์ครั้งที่ 4). กรุงเทพฯ: คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

สถาบันวิจัยภาษาและวัฒนธรรมเอเชีย มหาวิทยาลัยมหิดล. (2552). *หลักสูตรศิลปศาสตรมหาบัณฑิต สาขาวิชาภาษาและวัฒนธรรมเพื่อการสื่อสารและการพัฒนา.* สืบค้นเมื่อ 28 ธันวาคม 2559, จาก http://www.lc.mahidol.ac.th

สังวรณ์ งัดกระโทก. (2559). *เอกสารประกอบการอบรมเชิงปฏิบัติการ เรื่อง การวัดและประเมินผลการศึกษาด้วยทฤษฎีการวัดแบบดั้งเดิมและทฤษฎีการตอบข้อสอบ.* วันที่ 4-6 เมษายน 2559 โครงการบริการวิชาการแกสังคม ประจำปีงบประมาณ พ.ศ. 2559 มหาวิทยาลัยสุโขทัยธรรมาธิราช.

สัมพันธ์ พันธุ์พฤกษ์. (มปป.). *มาตรฐานการทดสอบทางการศึกษาแห่งชาติ.* สืบค้นเมื่อ 29 ธันวาคม 2559, จาก http://www.niets.or.th

สำนักงานเลขาธิการสภาการศึกษา. (2552). *การพัฒนานโยบายการเรียนการสอนภาษาไทยในจังหวัดชายแดนภาคใต้.* กรุงเทพฯ: สำนักงานเลขาธิการสภาการศึกษา กระทรวงศึกษาธิการ.

สำนักทดสอบทางการศึกษา. (2557). *คู่มือการจัดสอบประเมินคุณภาพการศึกษาขั้นพื้นฐานเพื่อการประกันคุณภาพผู้เรียน ปีการศึกษา 2557.* สืบค้นเมื่อ 15 มิถุนายน 2559, จาก http://www.bet.obec.go.th

สุวิไล เปรมศรีรัตน์. (2552). *โครงการพัฒนาคุณภาพการเรียนการสอนภาษาไทยโดยนำภาษาท้องถิ่นร่วมจัดการเรียนรู้.* นครปฐม : สถาบันวิจัยภาษาและวัฒนธรรมเอเชีย มหาวิทยาลัยมหิดล.

องค์การทุนเพื่อเด็กแห่งสหประชาชาติ. (2561). *สะพานเชื่อมโยงไปสู่อนาคตอันสดใส : โครงการจัดการศึกษาแบบทวิ-พหุภาษา (ภาษาไทย–มลายูถิ่น) ในโรงเรียนเขตพื้นที่จังหวัดชายแดนภาคใต้.* กรุงเทพฯ: องค์การทุนเพื่อเด็กแห่งสหประชาชาติ (ยูนิเซฟ).

อมรวิชช์ นาครทรรพ. (2558). *โจทย์ใหม่การศึกษาไทย ปฏิรูปสู่การเรียนรู้ตลอดชีวิต.* สืบค้นเมื่อ 29 ธันวาคม 2559, จากhttp://www.citizenforum.in.th/post.php?id=223

Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items.* (Vol. 4) Thousand Oaks, CA: Sage Publications.

Kim, S.H., Chosen, A.S. and Kim, S. (2007). DIF Detection and Effect Size Measures for Polytomously Scored Item. *Journal of Educational Measurement Summer,* 44 (2): 93-116.

McCallon E. L.and Schumacker, R.E. (2002). *Test bias and differential item functioning:* ELM. Metrics.

Holland, P. W., & Thayer, D.T. (1988). *Differential item performance and the Mantel-Heanszel procedure.* In Wainer, H. & Braun, H.I. {eds.}, Test Validity, 129-145. NJ: Lawrence Erlbaum Associate.

Zumbo, B.D. (1999). *A Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic regression modeling as a unitary framework for binary and likert-type (ordinal) item scores.* Ottawa, Canada: Directorate of Human Resources Research and Evaluation, Department of National Defense. Retrieved from http://www.edu.ubc.ca/faculty/zumbo/DIF/index.html

****