



การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมีค
Detecting Differential Item Functioning Using
Multiple Indicators and Multiple Causes (MIMIC)

พิชิตชัย สุขเอก

Pichitchai Sukaek

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาศึกษาศาสตรมหาบัณฑิต
สาขาวิชาการวิจัยและประเมินผลการศึกษา
มหาวิทยาลัยสงขลานครินทร์

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of
Master of Education in Educational Research and Evaluation

Prince of Songkla University

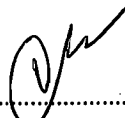
2560

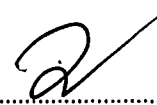
ลิขสิทธิ์ของมหาวิทยาลัยสงขลานครินทร์

ชื่อวิทยานิพนธ์ การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิค
 ผู้เขียน นายพิชิตชัย สุขเอก
 สาขาวิชา การวิจัยและประเมินผลการศึกษา

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

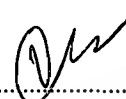
คณะกรรมการสอบ




ประธานกรรมการ

(ดร.จิระวัฒน์ ตันสกุล)

(ดร.มัยดี แวดราแม)


กรรมการ

(ดร.จิระวัฒน์ ตันสกุล)

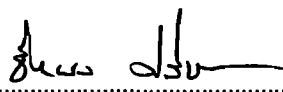

กรรมการ

(ดร.ณรงค์ศักดิ์ รอบคอบ)


กรรมการ

(ดร.เสาวรส ยิ่งวรรณะ)

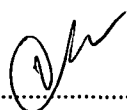
บัณฑิตวิทยาลัย มหาวิทยาลัยสงขลานครินทร์ อนุมัติให้บัณฑิตวิทยานิพนธ์ฉบับนี้
 เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาศึกษาศาสตรมหาบัณฑิต สาขาวิชาการวิจัย
 และประเมินผลการศึกษา



(รองศาสตราจารย์ ดร.ธีระพล ศรีชนะ)

คณบดีบัณฑิตวิทยาลัย

ขอรับรองว่า ผลงานวิจัยนี้มาจากการศึกษาวิจัยของนักศึกษาเอง และได้แสดงความขอบคุณบุคคลที่มี
ส่วนช่วยเหลือแล้ว

ลงชื่อ.....

(ดร.จิระวัฒน์ ต้นสกุล)

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

Prince of Songkla University
Pattani Campus

ลงชื่อ.....พิชิตชัย สุขเอก

(นายพิชิตชัย สุขเอก)

นักศึกษา

ข้าพเจ้าขอรับรองว่า ผลงานวิจัยนี้ไม่เคยเป็นส่วนหนึ่งในการอนุมัติปริญญาในระดับใดมาก่อน และ
ไม่ได้ถูกใช้ในการยื่นขออนุมัติปริญญาในขณะนี้

ลงชื่อ..... พิชิตชัย สุขเอก

(นายพิชิตชัย สุขเอก)

นักศึกษา

Prince of Songkla University
Pattani Campus

ชื่อวิทยานิพนธ์	การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิค
ผู้เขียน	นายพิชิตชัย สุขเอก
สาขาวิชา	การวิจัยและประเมินผลการศึกษา
ปีการศึกษา	2559

บทคัดย่อ

การวิจัยครั้งนี้ มีวัตถุประสงค์เพื่อวิเคราะห์คุณภาพแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) และตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิค (Multiple Indicators and Multiple Causes : MIMIC) เมื่อจำแนกตามเพศและที่ตั้งของสถานศึกษา ระหว่างปีการศึกษา 2556 – 2558 ประชากรที่ใช้ในการวิจัยครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 เครื่องมือที่ใช้ในการวิจัยคือแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2556 – 2558 โดยมีรายวิชาภาษาไทย ภาษาอังกฤษ คณิตศาสตร์ และวิทยาศาสตร์ ซึ่งเป็นข้อมูลทุติยภูมิ ผลการวิจัยพบว่า

1. ผลการประมาณค่าพารามิเตอร์ความยาก อำนาจจำแนกของข้อสอบ และระดับความสามารถของผู้สอบของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ปีการศึกษา 2556 – 2558 พบว่ารายวิชาภาษาไทย มีค่าความยากน้อยสุดเท่ากับ .01 และสูงสุดเท่ากับ 8.32 ค่าอำนาจจำแนกน้อยสุดเท่ากับ .01 และสูงสุดเท่ากับ 1.03 ค่าความสามารถของผู้สอบน้อยสุด -3.73 และสูงสุดเท่ากับ 3.41 รายวิชาภาษาอังกฤษ มีค่าความยากน้อยสุดเท่ากับ .00 และสูงสุดเท่ากับ 8.31 ค่าอำนาจจำแนกน้อยสุดเท่ากับ .00 และสูงสุดเท่ากับ 1.18 ค่าความสามารถของผู้สอบน้อยสุด -2.83 และสูงสุดเท่ากับ 5.05 รายวิชาคณิตศาสตร์ มีค่าความยากน้อยสุดเท่ากับ .05 และสูงสุดเท่ากับ 8.14 ค่าอำนาจจำแนกน้อยสุดเท่ากับ .02 และสูงสุดเท่ากับ .70 ค่าความสามารถของผู้สอบน้อยสุด -2.11 และสูงสุดเท่ากับ 3.90 รายวิชาวิทยาศาสตร์ มีค่าความยากน้อยสุดเท่ากับ -82.66 และสูงสุดเท่ากับ 494.63 ค่าอำนาจจำแนกน้อยสุดเท่ากับ -1.09 และสูงสุดเท่ากับ 1.18 ค่าความสามารถของผู้สอบน้อยสุด -2.74 และสูงสุดเท่ากับ 4.75

2. ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิค จำแนกตามเพศ พบว่ารายวิชาภาษาไทยและวิทยาศาสตร์ เพศชายมีโอกาสดอบข้อสอบถูกมากกว่าเพศหญิง และรายวิชาภาษาอังกฤษ เพศหญิงมีโอกาสดอบข้อสอบถูกมากกว่าเพศชาย เมื่อจำแนกตามที่ตั้งของสถานศึกษา พบว่ารายวิชาภาษาไทย และภาษาอังกฤษ ที่ตั้งของสถานศึกษาในเมืองมีโอกาสดอบข้อสอบถูกมากกว่านอกเมือง และรายวิชาคณิตศาสตร์และวิทยาศาสตร์ ที่ตั้งของสถานศึกษานอกเมืองมีโอกาสดอบข้อสอบถูกมากกว่าในเมือง

Thesis Title	Detecting Differential Item Functioning Using Multiple Indicators and Multiple Causes (MIMIC)
Author	Mr. Pichitchai Sukaek
Major Program	Education Research and Evaluation
Academic Year	2016

ABSTRACT

The objective of this research were to check test quality analysis of Ordinary National Educational Test (O-NET) and Differential Item Functioning (DIF) by using Multiple Indicators and Multiple Causes (MIMIC) when distinguish by genders and school locations, during 2013-2015 academic years. Parameters that used in this research were grade 12 students by using multistage stage sampling. The research instrument was test of Ordinary National Educational Test (O-NET) grade 12 students, 2013-2015 academic years, including Thai, English Mathematics and Sciences that were minor information. This research found that

1. Analysis of Parameter of Difficulty, Test Discrimination and Ability Parameters of Ordinary National Educational Test (O-NET), during 2013-2015 academic years revealed that Thai had the fewest difficulty was .01 and the highest difficulty was 8.32, the fewest test discrimination was .01 and the highest test discrimination was 1.03, the fewest ability parameters was -3.73 and the highest ability parameters was 3.41. English had the fewest difficulty was .00 and the highest difficulty was 8.31 ,the fewest test discrimination was .00 and the highest test discrimination was 1.18 , the fewest ability parameters was -2.83 and the highest ability parameters was 5.05. Mathematics had the fewest difficulty was .05 and the highest difficulty was 8.14 ,the fewest test discrimination was .02 and the highest test discrimination was .70, the fewest ability parameters was -2.11 and the highest ability parameters was 3.90. Sciences had the fewest difficulty was -82.66 and the highest difficulty was 494.63 ,the fewest test discrimination was -1.09 and the highest test discrimination was 1.18, the fewest ability parameters was -2.74 and the highest ability parameters was 4.75.

2. Analysis of Differential Item Functioning (DIF) by Multiple Indicators and Multiple Causes (MIMIC) distinguished by genders found that Thai and Sciences tests were partial male higher than female, and English test was partial female higher than male. When distinguished by school locations found that Thai and English tests were partial city higher than suburb, and Mathematics and Sciences tests were partial suburb higher than city.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี ด้วยการดูแลและให้คำแนะนำอย่างดียิ่งจาก ดร.จิระวัฒน์ ต้นสกุล อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก และคณะกรรมการสอบวิทยานิพนธ์โดยมี ดร.ม้อตี แวดราแม ประธานกรรมการสอบวิทยานิพนธ์ และ ดร.เสาวรส ยิ่งวรรณ กรรมการ ผู้ทรงคุณวุฒิภายนอก ที่กรุณาให้แนวความคิด คำแนะนำเพิ่มเติม เพื่อปรับปรุงแก้ไขวิทยานิพนธ์ฉบับนี้ มีความสมบูรณ์มากยิ่งขึ้น ผู้วิจัยขอขอบพระคุณไว้เป็นอย่างสูง ณ โอกาสนี้

ขอขอบคุณคณบดีคณะศึกษาศาสตร์ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี ได้จัดสรรเงินเป็นทุนการศึกษาปริญญาตรี-โท คณะศึกษาศาสตร์ จากเงินกองทุนวิจัย คณะศึกษาศาสตร์ที่ได้หยิบยื่นโอกาส ทำให้ผู้วิจัยสามารถพัฒนาตนเองไปอีกก้าวขั้นของความสำเร็จ

ขอขอบคุณคณาจารย์ประจำภาควิชาประเมินผลและวิจัยทางการศึกษา รวมทั้ง อาจารย์และผู้รู้หลายท่านที่มีได้เอ่ยนามไว้ ณ ที่นี้ที่ได้ให้แนวคิดและประสิทธิ์ประสาทวิชาให้ผู้วิจัยมีความรู้พื้นฐาน อันเป็นเครื่องมือสำคัญในการทำงานวิจัยครั้งนี้ได้สำเร็จ ผู้วิจัยรู้สึกซาบซึ้งในความเมตตาของท่านเป็นอย่างยิ่ง และขอขอบคุณเลขานุการประจำภาควิชาประเมินผลและวิจัยทางการศึกษา ที่ได้อำนวยความสะดวกและให้ความช่วยเหลือผู้วิจัยในการติดต่อประสานงานเรื่องต่าง ๆ เป็นอย่างดี

ขอขอบคุณคณบดีบัณฑิตวิทยาลัยที่อนุเคราะห์ทุนในการวิจัยและขอขอบคุณเจ้าหน้าที่บัณฑิตวิทยาลัยทุกท่านที่ได้ให้ความช่วยเหลือในทุกขั้นตอนเป็นอย่างดี

ขอขอบคุณสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ที่ให้ความอนุเคราะห์ในการให้ความอนุเคราะห์ข้อมูลการวิจัย เป็นอย่างดี

ขอกราบขอบพระคุณบิดามารดาที่เป็นกำลังใจและคอยช่วยเหลือสนับสนุนในทุก ๆ ด้าน อีกทั้งเพื่อน ๆ ร่วมรุ่นที่คอยให้ความช่วยเหลือมาโดยตลอด

คุณประโยชน์ใด ๆ อันพึงมีจากวิทยานิพนธ์เล่มนี้ ผู้วิจัยขอมอบแต่บิดา มารดา ครู อาจารย์ ตลอดจนจนสถาบันการศึกษาที่ได้ประสิทธิ์ประสาทให้ความรู้แก่ผู้วิจัยตลอดมา

พิชิตชัย สุขเอก

สารบัญ

	หน้า
บทคัดย่อ.....	(5)
ABSTRACT.....	(6)
กิตติกรรมประกาศ.....	(7)
สารบัญ.....	(8)
รายการตาราง.....	(10)
รายการภาพประกอบ.....	(14)
บทที่ 1 บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา.....	1
วัตถุประสงค์ของการวิจัย.....	4
ขอบเขตของการวิจัย.....	4
ข้อจำกัดของการวิจัย.....	5
นิยามศัพท์เฉพาะ.....	5
ประโยชน์ที่คาดว่าจะได้รับ.....	7
กรอบแนวคิดการวิจัย.....	7
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	8
ตอนที่ 1 มโนทัศน์ของทฤษฎีการตอบสนองข้อสอบ.....	9
ตอนที่ 2 มโนทัศน์ของการทำหน้าที่ต่างกันของข้อสอบ.....	14
ตอนที่ 3 มโนทัศน์ของการวิเคราะห์ข้อสอบด้วยวิธีมิมิค (MIMIC)	26
ตอนที่ 4 งานวิจัยที่เกี่ยวข้องกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ.....	36
บทที่ 3 วิธีการดำเนินการวิจัย.....	51
ประชากร.....	51
เครื่องมือที่ใช้ในการวิจัย.....	51
การเก็บรวบรวมข้อมูล.....	56
การวิเคราะห์ข้อมูล.....	56
สถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	57
บทที่ 4 ผลการวิจัย.....	61
ตอนที่ 1 การวิเคราะห์ค่าสถิติพื้นฐาน.....	63
ตอนที่ 2 ผลการประมาณค่าพารามิเตอร์ความยาก อำนาจจำแนกของข้อสอบ และ ระดับความสามารถของผู้สอบ.....	67
ตอนที่ 3 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบจำแนกตามเพศและที่ตั้งของ สถานศึกษา.....	86

สารบัญ (ต่อ)

	หน้า
ตอนที่ 4 ผลการวิเคราะห์แนวโน้มการทำหน้าที่ต่างกันของข้อสอบจำแนกตามเพศและที่ตั้งของสถานศึกษา.....	125
บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ.....	133
สรุปผลการวิจัย.....	133
อภิปรายผล.....	137
ข้อเสนอแนะ.....	139
บรรณานุกรม.....	140
ภาคผนวก.....	143
ภาคผนวก ก หนังสือขอความอนุเคราะห์.....	144
ภาคผนวก ข Syntax ในการวิเคราะห์ข้อมูลโดยประยุกต์ใช้โปรแกรม Mplus	148
1. ตัวอย่าง Syntax ในการวิเคราะห์การประมาณค่าระดับความสามารถของผู้สอบ ค่าพารามิเตอร์ความยากและอำนาจจำแนกของข้อสอบ โดยประยุกต์ใช้โปรแกรม Mplus.....	149
2. ตัวอย่าง Syntax ในการวิเคราะห์การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบจำแนกตามเพศและที่ตั้งของสถานศึกษา โดยประยุกต์ใช้โปรแกรม Mplus.....	150
ประวัติผู้เขียน.....	152

รายการตาราง

ตาราง		หน้า
1	วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ที่มีการตรวจให้คะแนนแบบ ทวิวิภาค (Dichotomous DIF) และพหุวิภาค (Polytomous DIF).....	23
2	ผลการสังเคราะห์ตัวแปร.....	49
3	จำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ของแบบทดสอบการ ทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558.....	52
4	จำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ของแบบทดสอบการ ทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558.....	53
5	จำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ของแบบทดสอบการ ทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558.....	54
6	จำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ของแบบทดสอบการ ทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558.....	55
7	จำนวนผู้เข้าสอบ คะแนนต่ำสุด คะแนนสูงสุด คะแนนเฉลี่ย คะแนนส่วนเบี่ยงเบน มาตรฐาน ค่าความเบ้และค่าความโด่ง จากคะแนนทดสอบทางการศึกษาระดับชาติด้าน พื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558 จำแนกตามเพศและ ที่ตั้งของสถานศึกษา.....	63
8	จำนวนผู้เข้าสอบ คะแนนต่ำสุด คะแนนสูงสุด คะแนนเฉลี่ย คะแนนส่วนเบี่ยงเบน มาตรฐาน ค่าความเบ้และค่าความโด่ง จากคะแนนทดสอบทางการศึกษาระดับชาติด้าน พื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558 จำแนกตามเพศและ ที่ตั้งของสถานศึกษา.....	64
9	จำนวนผู้เข้าสอบ คะแนนต่ำสุด คะแนนสูงสุด คะแนนเฉลี่ย คะแนนส่วนเบี่ยงเบน มาตรฐาน ค่าความเบ้และค่าความโด่ง จากคะแนนทดสอบทางการศึกษาระดับชาติด้าน พื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558 จำแนกตามเพศและ ที่ตั้งของสถานศึกษา.....	65
10	จำนวนผู้เข้าสอบ คะแนนต่ำสุด คะแนนสูงสุด คะแนนเฉลี่ย คะแนนส่วนเบี่ยงเบน มาตรฐาน ค่าความเบ้และค่าความโด่ง จากคะแนนทดสอบทางการศึกษาระดับชาติด้าน พื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558 จำแนกตามเพศและ ที่ตั้งของสถานศึกษา.....	66

รายการตาราง (ต่อ)

ตาราง		หน้า
11	ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558.....	67
12	ผลการประมาณค่าระดับความสามารถของผู้สอบที่ได้ทำแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 - 2558	70
13	ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558.....	72
14	ผลการประมาณค่าระดับความสามารถของผู้สอบที่ได้ทำแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558.....	75
15	ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558.....	77
16	ผลการประมาณค่าระดับความสามารถของผู้สอบที่ได้ทำแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558.....	78
17	ผลการประมาณค่าพารามิเตอร์ความยากของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558.....	81
18	ผลการประมาณค่าระดับความสามารถของผู้สอบที่ได้ทำแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558.....	84
19	ผลการวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 - 2558 จำแนกตามเพศและที่ตั้งของสถานศึกษา.....	86
20	ผลการวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 - 2558 จำแนกตามเพศและที่ตั้งของสถานศึกษา.....	88

รายการตาราง (ต่อ)

ตาราง		หน้า
21	ผลการวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC ของแบบทดสอบการทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 - 2558 จำแนกตามเพศและที่ตั้งของสถานศึกษา.....	90
22	ผลการวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC ของแบบทดสอบการทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 - 2558 จำแนกตามเพศและที่ตั้งของสถานศึกษา.....	92
23	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน(O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศ.....	94
24	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน(O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศ.....	97
25	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน(O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศ.....	100
26	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน(O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศ.....	102
27	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน(O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามที่ตั้งของสถานศึกษา.....	105
28	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน(O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามที่ตั้งของสถานศึกษา.....	108
29	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน(O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามที่ตั้งของสถานศึกษา.....	111
30	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน(O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามที่ตั้งของสถานศึกษา.....	113

รายการตาราง (ต่อ)

ตาราง		หน้า
31	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน(O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศและที่ตั้งของสถานศึกษา.....	116
32	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน(O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศและที่ตั้งของสถานศึกษา.....	118
33	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน(O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศและที่ตั้งของสถานศึกษา.....	121
34	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน(O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศและที่ตั้งของสถานศึกษา.....	122
35	ผลการวิเคราะห์แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศและที่ตั้งของสถานศึกษา.....	125
36	ผลการวิเคราะห์แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศและที่ตั้งของสถานศึกษา.....	127
37	ผลการวิเคราะห์แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศและที่ตั้งของสถานศึกษา.....	129
38	ผลการวิเคราะห์แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศและที่ตั้งของสถานศึกษา.....	131

รายการภาพประกอบ

ภาพประกอบ		หน้า
1	โมเดลการตอบสนองข้อสอบ (IRT).....	9
2	ข้อสอบทำหน้าที่ต่างกันแบบเอกรูป (Uniform DIF).....	16
3	ข้อสอบทำหน้าที่ต่างกันแบบไม่มีทิศทาง (Non-unidirectional DIF).....	18
4	ข้อสอบทำหน้าที่ต่างกันแบบมีทิศทางเดียวกัน (Unidirectional DIF).....	18
5	การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบและแบบสอบ โดยใช้เทคนิคการวิเคราะห์สมการถดถอย.....	20
6	โมเดลย่อยในโปรแกรม Lisrel : โมเดลมิมิค (Multiple Indicators and Multiple Causes Models).....	26
7	โมเดลย่อยในโปรแกรม M Plus: โมเดลมิมิค (Multiple Indicators and Multiple Causes Models).....	27
8	โมเดล MIMIC ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบเอกรูปบนข้อสอบข้อที่ 2.....	29
9	โมเดล MIMIC สำหรับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เมื่อมีปฏิสัมพันธ์ระหว่าง กลุ่มของตัวแปรกลุ่มและตัวแปรแฝง (θ)	31
10	กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556.....	70
11	กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2557.....	71
12	กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2558.....	71
13	กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556.....	75
14	กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2557.....	76
15	กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2558.....	76

รายการภาพประกอบ (ต่อ)

ภาพประกอบ		หน้า
16	กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556.....	79
17	กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2557.....	79
18	กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2558.....	80
19	กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556.....	84
20	กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2557.....	85
21	กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2558.....	85
22	แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558 จำแนกตามเพศ.....	126
23	แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558 จำแนกตามที่ตั้งของสถานศึกษา.....	126
24	แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558 จำแนกตามเพศ...	128
25	แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558 จำแนกตามที่ตั้งของสถานศึกษา.....	128
26	แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558 จำแนกตามเพศ....	130

รายการภาพประกอบ (ต่อ)

ภาพประกอบ	หน้า
27	แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินัยพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558 จำแนกตามที่ตั้งของสถานศึกษา..... 130
28	แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินัยพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558 จำแนกตามเพศ...132
29	แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินัยพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558 จำแนกตามที่ตั้งของสถานศึกษา..... 132

Prince of Songkla University
Pattani Campus

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การวัดผลการศึกษาเป็นการตรวจสอบผู้เรียนว่ามีความรู้หรือคุณลักษณะที่ต้องการวัดอยู่ในระดับใด ซึ่งผลที่ได้จากการวัดมีความสำคัญต่อพัฒนาการศึกษา การทดสอบเป็นวิธีการวัดผล การศึกษาวิธีหนึ่งที่ยอมรับกันมาก เครื่องมือที่ใช้ในการทดสอบที่สำคัญก็คือ ข้อสอบชนิดต่างๆ ข้อสอบเป็นเครื่องมือหลักที่ใช้ในการวัดและประเมินผลทางการศึกษาเพื่อตรวจสอบว่าผู้เข้าสอบมีคุณลักษณะแฝงหรือความสามารถอยู่ในระดับใด ในการสร้างและการตรวจสอบคุณภาพของข้อสอบ ซึ่งการสร้างข้อสอบที่ดีต้องคำนึงถึงความตรง ความเชื่อมั่น ความยาก อำนาจจำแนก ความเป็นปรนัย ด้านความตรงก็เป็นสิ่งสำคัญ ทั้งนี้เพราะความตรงเป็นคุณสมบัติของแบบทดสอบที่แสดงถึงความสามารถในการวัดได้ถูกต้องแม่นยำ (ศิริชัย กาญจนวาสี, 2555)

ในการทดสอบแต่ละครั้ง ผู้สอบอาจมีลักษณะแตกต่างกัน เช่น เชื้อชาติ ศาสนา วัฒนธรรม ภูมิฐานะ สังคม เพศ ภาษา อายุ และประสบการณ์ เป็นต้น ผู้สอบดังกล่าวอาจไม่ได้รับความยุติธรรมในการทำข้อสอบ โดยข้อสอบบางข้ออาจมีความลำเอียงเข้าข้างผู้สอบกลุ่มย่อยบางกลุ่มของผู้เข้าสอบทั้งหมด ซึ่งทำให้เกิดการได้เปรียบเสียเปรียบระหว่างผู้สอบกลุ่มย่อยด้วยกัน ทั้งๆ ที่สอบด้วยข้อสอบข้อเดียวกัน แสดงว่าข้อสอบขาดความตรง สาเหตุดังกล่าวอาจเนื่องมาจากแบบทดสอบไม่ได้วัดความสามารถเป้าหมายที่ต้องการวัดเพียงอย่างเดียว แต่ยังวัดความสามารถแทรกซ้อนที่ไม่ต้องการวัดอีกด้วย ส่วนข้อสอบที่วัดทั้งความสามารถเป้าหมายและความสามารถแทรกซ้อน เรียกว่าข้อสอบทำหน้าที่ต่างกัน ซึ่งถ้ามีผู้สอบกลุ่มย่อยใดมีความสามารถแทรกซ้อนสูงกว่าก็มีโอกาสในการตอบได้ถูกต้องมากกว่า ทั้งๆ ที่ระดับความสามารถเป้าหมายที่ต้องการวัดเท่ากัน จึงส่งผลทำให้ข้อสอบทำหน้าที่ต่างกัน (ศิริชัย กาญจนวาสี, 2555)

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบเป็นการเปรียบเทียบผลการตอบข้อสอบระหว่างกลุ่มผู้สอบอย่างน้อย 2 กลุ่มขึ้นไป โดยทั่วไปทำการเปรียบเทียบผู้สอบ 2 กลุ่ม ได้แก่ กลุ่มเปรียบเทียบ (Focal Group) เป็นกลุ่มที่คาดว่าจะจะเป็นกลุ่มที่เสียประโยชน์ในการตอบข้อสอบ และ กลุ่มอ้างอิง (Reference Group) เป็นกลุ่มที่คาดว่าจะได้ประโยชน์จากการตอบข้อสอบได้ถูกต้อง ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ จะพบว่าข้อสอบสามารถทำหน้าที่ต่างกันได้ 2 ประเภท ได้แก่ การทำหน้าที่ต่างกันของข้อสอบแบบเอกรูป (Uniform) และแบบอเนกรูป (Nonuniform) 1) ข้อสอบที่ทำหน้าที่ต่างกันแบบเอกรูป (Uniform DIF) หมายถึง ข้อสอบที่ทำให้ผู้สอบกลุ่มหนึ่งมีโอกาสในการตอบข้อสอบถูกมากกว่าผู้สอบอีกหนึ่งกลุ่มอย่างสม่ำเสมอ ในทุกระดับความสามารถเมื่อพิจารณาโค้งคุณลักษณะข้อสอบของผู้สอบ 2 กลุ่ม จะพบว่าไม่มีปฏิสัมพันธ์ระหว่างระดับความสามารถของผู้สอบกับการเป็นสมาชิกของกลุ่ม (Group membership)

2) ข้อสอบทำหน้าที่ต่างกันแบบอเนกรูป (Nonuniform DIF) หมายถึง ข้อสอบที่ทำให้โอกาสในการตอบข้อสอบถูกของผู้สอบระหว่างกลุ่มแตกต่างกันอย่างไม่สม่ำเสมอในทุกระดับความสามารถ เมื่อพิจารณาโค้งคุณลักษณะข้อสอบของผู้สอบ 2 กลุ่ม พบว่ามีปฏิสัมพันธ์ร่วมกันระหว่างระดับความสามารถของผู้สอบกับการเป็นสมาชิกของกลุ่ม เช่น ที่ระดับความสามารถหนึ่ง กลุ่มผู้สอบกลุ่ม R มีโอกาสในการตอบข้อสอบถูกมากกว่ากลุ่มผู้สอบกลุ่ม F แต่ที่ระดับความสามารถอีกระดับหนึ่ง กลุ่มผู้สอบกลุ่ม F มีโอกาสในการตอบข้อสอบถูกมากกว่ากลุ่มผู้สอบกลุ่ม R (ศิริชัย กาญจนวาสี, 2555)

วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF detection) จำแนกตามลักษณะการตรวจให้คะแนนได้เป็น 2 ประเภทคือ ข้อสอบที่มีการให้คะแนนแบบสองค่า (Dichotomous scoring) และข้อสอบที่มีการให้คะแนนแบบหลายค่า (Polytomous scoring) ซึ่งผู้วิจัยได้สนใจข้อสอบที่มีการให้คะแนนแบบสองค่า (Dichotomous scoring) เนื่องจากสามารถจำแนกความสามารถของผู้ตอบและคุณลักษณะของข้อสอบได้ โดยวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ให้คะแนนแบบสองค่า (Dichotomous scoring) จำแนกตามมิติลักษณะของตัวแปรเกณฑ์ได้เป็น 2 ประเภท คือ กลุ่มวิธีที่ใช้คะแนนที่สังเกตได้และกลุ่มวิธีที่ใช้คุณลักษณะแฝง โดยที่ผู้วิจัยได้สนใจกลุ่มวิธีที่ใช้คุณลักษณะแฝง เนื่องจากกลุ่มวิธีที่ใช้คะแนนที่สังเกตได้ วิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) หรือกลุ่มที่ไม่ใช้ทฤษฎีการตอบสนองข้อสอบ (Non-IRT approach) มีข้อจำกัดที่สำคัญได้แก่ การที่พารามิเตอร์ของข้อสอบผันแปรตามกลุ่มผู้สอบและคะแนนที่สังเกตได้หรือค่าประมาณความสามารถของผู้ตอบไม่เป็นอิสระหรือขึ้นอยู่กับข้อสอบและแบบทดสอบที่นำมาใช้ และกลุ่มนี้ใช้คุณลักษณะหรือตัวแปรแฝงซึ่งวิเคราะห์บนพื้นฐานของทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นทฤษฎีการทดสอบแนวใหม่ ซึ่งเป็นที่ยอมรับและเป็นแนวคิดของศาสตร์วัดผลในปัจจุบัน ซึ่งวิธีการตรวจสอบที่สำคัญในกลุ่มนี้ ได้แก่วิธีวัดพื้นที่ความแตกต่างระหว่างโค้งการตอบสนองข้อสอบ (IRT-D2) วิธีเคสแควร์ของลอว์รีด วิธีอัตราส่วนไลค์ลิฮูดทั่วไป (General IRT Likelihood Ratio) วิธีอัตราส่วนไลค์ลิฮูด ลอกลินีเยียร์ (Loglinear IRT Likelihood Ratio) วิธีชิปเทสต์ (SIBTEST) (Shealy & stout, 1993) และวิธีมิมีค (Multiple Indicators and Multiple Causes : MIMIC) (ศิริชัย กาญจนวาสี, 2555)

โมเดลมิมีค (Multiple Indicators and Multiple Causes : MIMIC) โมเดลที่มีตัวแปรแฝงเพียงตัวแปรเดียว โดยที่ตัวแปรแฝงนั้นได้รับอิทธิพลจากตัวแปรภายนอกสังเกตได้หลายตัวแปรและส่งอิทธิพลไปยังตัวแปรภายในสังเกตได้หลายตัวแปร กล่าวอีกอย่างหนึ่งคือ เป็นโมเดลที่มีคุณลักษณะแฝงที่มีหลายสาเหตุและวัดได้จากตัวบ่งชี้หลายตัว ซึ่งในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธีมิมีค พบว่าข้อสอบสามารถทำหน้าที่ต่างกันได้ 2 ประเภท ได้แก่ การทำหน้าที่ต่างกันของข้อสอบแบบเอกรูป (Uniform) และ แบบอเนกรูป (Nonuniform) ถ้าข้อสอบหรือข้อคำถาม ได้รับอิทธิพลอย่างมีนัยสำคัญจากตัวแปรตัวแปรภายนอกแสดงให้เห็นว่าข้อสอบหรือข้อคำถามขึ้นอยู่กับตัวแปรภายนอก ไม่ได้อธิบายตัวแปรแฝงแสดงว่าข้อสอบหรือข้อคำถาม ข้อนั้นทำหน้าที่ต่างกันของข้อสอบ ซึ่งผู้วิจัยได้สนใจการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยมิมีค เพราะ 1) โมเดลมิมีค เป็นวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบค่อนข้างใหม่ 2) มีการลดจำนวนพารามิเตอร์ในการทดสอบความแตกต่างระหว่างกลุ่ม (Ian A. Carroll, 2014) 3) มีการลดลง

ของขนาดตัวอย่างให้น้อยลงในขณะที่รูปแบบอื่นๆ เช่น การวิเคราะห์องค์ประกอบเชิงยืนยันหลายกลุ่ม (multiple-group CFA models) จำเป็นต้องใช้ขนาดตัวอย่างจำนวนมากเพื่อให้ได้ผลลัพธ์ที่มีความน่าเชื่อถือ (Woods, 2009) 4) โมเดลมิมิค มีกระบวนการในการวิเคราะห์ที่สะดวกมากยิ่งขึ้นเมื่อมีจำนวนกลุ่มตัวแปรหลายกลุ่มเมื่อเปรียบเทียบกับรูปแบบอื่นๆ เช่น การวิเคราะห์องค์ประกอบเชิงยืนยันหลายกลุ่ม (multiple-group CFA models) (Brown, 2006) 5) มีข้อตกลงเบื้องต้นเกี่ยวกับขนาดตัวอย่างที่น้อยกว่า เมื่อเปรียบเทียบกับรูปแบบอื่นๆ เช่น การวิเคราะห์องค์ประกอบเชิงยืนยันหลายกลุ่ม (multiple-group CFA models) ที่ข้อตกลงว่าต้องมีขนาดตัวอย่างใหญ่ (Brown, 2006) ปัจจุบันได้มีการพัฒนาโปรแกรม Mplus เพื่อให้นักวิจัยมีเครื่องมือสำหรับการวิเคราะห์ข้อมูลด้วยสถิติวิเคราะห์ขั้นสูง ที่ให้ผลการวิเคราะห์ข้อมูลที่มีความถูกต้องมากกว่าสถิติแบบเดิม ซึ่งในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิค ในครั้งนี้ ผู้วิจัยได้ใช้โปรแกรม Mplus เพราะ มีการพัฒนาอย่างต่อเนื่องของโปรแกรม Mplus สามารถออกแบบวิธีการวิเคราะห์ข้อมูลเป็นรูปแบบเดียวกันเกือบทุกโมเดล เพื่อลดความยุ่งยากในขั้นตอนของการวิเคราะห์สถิติขั้นสูงของโปรแกรมดั้งเดิมที่ใช้วิเคราะห์ที่กันอยู่โดยทั่วไป โปรแกรม Mplus ในการเขียนคำสั่งไม่ต้องเสียเวลาจำเมทริกซ์พารามิเตอร์ เหมือนกับโปรแกรม LISREL และนอกจากนี้ยังมีงานวิจัยที่เกี่ยวข้องของ Finch (2005) ได้ศึกษาการเปรียบเทียบประสิทธิภาพของโมเดลมิมิค กับการทดสอบโดยวิธีแมนเทล-แฮนส์เซล (Mantel & Haenszel) และวิธี SIBTEST และวิธีการทดสอบ IRT Likelihood Ratio กับความคลาดเคลื่อนประเภทที่ 1 และอำนาจการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) ได้แสดงให้เห็นว่าวิธีมิมิค มีค่าสูงขึ้นและค่าความคลาดเคลื่อนประเภทที่ 1 มีค่าลดลงกับจำนวนข้อสอบ 50 ข้อ นอกจากนี้วิธีมิมิค ยังสามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ Nonuniform DIF ได้ด้วย

ในการวิจัยครั้งนี้ ผู้วิจัยมุ่งศึกษาการวิเคราะห์คุณภาพของข้อสอบตามทฤษฎีการตอบสนอง (IRT) โดยดำเนินการวิเคราะห์วิธีมิมิค (Multiple Indicators and Multiple Causes : MIMIC) โดยประยุกต์ใช้โปรแกรม Mplus ซึ่งโปรแกรมดังกล่าวสามารถวิเคราะห์สถิติขั้นสูงได้ดี โดยศึกษาจากแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (Ordinary National Educational Test : O-NET) ระดับชั้นมัธยมศึกษาปีที่ 6 ระหว่างปีการศึกษา 2556 - 2558 โดยผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบครั้งนี้ จะเป็นประโยชน์ในการพิจารณาปรับปรุงข้อสอบให้เหมาะสมและมีประสิทธิภาพต่อไป

วัตถุประสงค์ของการวิจัย

1. เพื่อวิเคราะห์คุณภาพแบบทดสอบของการทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) ชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2556 – 2558
2. เพื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิค (Multiple Indicators and Multiple Causes : MIMIC) เมื่อจำแนกตามเพศและที่ตั้งของสถานศึกษา ระหว่างปีการศึกษา 2556 - 2558

ขอบเขตการวิจัย

1. ประชากร

ประชากรที่ใช้ในการวิจัยครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 ทั้งหมดที่เข้าสอบแบบทดสอบการทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (Ordinary National Educational Test : O-net) ระหว่างปีการศึกษา 2556 - 2558

2. ตัวแปรที่ศึกษา

2.1 ตัวแปรอิสระ ได้แก่

กลุ่มที่ศึกษาการทำหน้าที่ต่างกันของข้อสอบ ได้แก่

- 1) เพศ คือ เพศชายและหญิง
- 2) ที่ตั้งของสถานศึกษา คือ ที่ตั้งของสถานศึกษาในเมืองและนอก

เมือง

2.2 ตัวแปรตาม ได้แก่

การทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิค (Multiple Indicators and Multiple Causes : MIMIC) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (Ordinary National Educational Test : O-net) ระหว่างปีการศึกษา 2556 – 2558

ข้อจำกัดของการวิจัย

1. ผู้วิจัยวิเคราะห์แบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (Ordinary National Educational Test : O-net) ตามโมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า (Dichotomous IRT Models) โดยเลือกใช้รูปแบบปรนัย 5 ตัวเลือก 1 คำตอบ
2. ผู้วิจัยไม่เลือกแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (Ordinary National Educational Test : O-net) รายวิชาสังคมศึกษา ศาสนาและวัฒนธรรม เนื่องจากแบบทดสอบปีการศึกษา 2556 – 2558 มีจำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) แต่ละปีการศึกษาที่ไม่เหมือนกัน

นิยามศัพท์เฉพาะ

การทำหน้าที่ต่างกันของข้อสอบ (DIF) หมายถึง ข้อสอบที่ทำให้ผลการตอบของผู้สอบที่มีความสามารถเท่ากันในเรื่องที่ต้องการวัด มีโอกาสตอบข้อสอบข้อนั้นได้ถูกต้องไม่เท่ากัน เนื่องจากอยู่ในกลุ่มย่อยต่างกัน ซึ่งในการวิจัยนี้ศึกษาการทำหน้าที่ต่างของข้อสอบ กลุ่มที่ศึกษาการทำหน้าที่ต่างกันของข้อสอบ ได้แก่ เพศ และที่ตั้งของโรงเรียน ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (Ordinary National Educational Test : O-net) ระหว่างปีการศึกษา 2556 - 2558 ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

การทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิค (MIMIC) หมายถึง ถ้าข้อสอบ (Item) ได้รับอิทธิพลอย่างมีนัยสำคัญจากตัวแปรสาเหตุ แสดงว่าข้อสอบ (Item) ขึ้นอยู่กับตัวแปรสาเหตุไม่ได้ขึ้นอยู่กับตัวแปรแฝง ทำให้เกิดการทำหน้าที่ต่างกันของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (Ordinary National Educational Test : O-net) ระหว่างปีการศึกษา 2556 - 2558 ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

แนวโน้มจำนวนข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิค (MIMIC) หมายถึง แนวโน้มของการทำหน้าที่ต่างกันของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (Ordinary National Educational Test : O-net) ระหว่างปีการศึกษา 2556 - 2558 ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

เพศ หมายถึง เพศชายและเพศหญิงของนักเรียนชั้นมัธยมศึกษาปีที่ 6 เข้าสอบแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (Ordinary National Educational Test : O-net) ระหว่างปีการศึกษา 2556 – 2558

ที่ตั้งของสถานศึกษา หมายถึง เขตพื้นที่ที่สถานศึกษาตั้งอยู่ตั้งอยู่ โดยจำแนกเป็น 2 กลุ่ม ได้แก่ สถานศึกษาที่ตั้งอยู่ในเมืองและนอกเมือง ของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่เข้าสอบแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (Ordinary National Educational Test : O-net) ระหว่างปีการศึกษา 2556 – 2558 ซึ่งใช้เกณฑ์การจำแนกตามสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) (สทศ.)

วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) หมายถึง วิธีการวิเคราะห์ทางสถิติที่ใช้เพื่อบ่งบอกว่าการทำหน้าที่ต่างกันของข้อสอบระหว่างผู้สอบกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่มีความสามารถระดับเดียวกันซึ่งวิธีการในงานวิจัยครั้งนี้เป็นวิธีการตรวจสอบที่ใช้กับข้อสอบที่มีการให้คะแนนแบบสองค่า

กลุ่มอ้างอิง (Reference Group : R) หมายถึง กลุ่มผู้สอบที่คาดว่าจะได้รับประโยชน์จากการตอบข้อสอบที่ทำหน้าที่ต่างกัน คือเป็นกลุ่มที่มีความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องสูงกว่าผู้สอบอีกกลุ่มหนึ่งทั้งๆ ที่มีความสามารถเท่ากัน

กลุ่มเปรียบเทียบ (focal Group : F) หมายถึง กลุ่มผู้สอบที่คาดว่าจะเสียประโยชน์จากการตอบข้อสอบที่ทำหน้าที่ต่างกัน คือเป็นกลุ่มที่มีความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องต่ำกว่าผู้สอบอีกกลุ่มหนึ่งทั้งๆ ที่มีความสามารถเท่ากัน

โมเดล MIMIC (Multiple Indicators and Multiple Causes : MIMIC) หมายถึง โมเดลที่มีตัวแปรแฝงเพียงตัวแปรเดียว โดยที่ตัวแปรแฝงนั้นได้รับอิทธิพลจากตัวแปรภายนอกสังเกตได้หลายตัวแปร และส่งอิทธิพลไปยังตัวแปรภายในสังเกตได้หลายตัวแปร

ข้อสอบทำหน้าที่ต่างกันแบบเอกรูป (Uniform DIF) หมายถึง ข้อสอบที่ทำให้ผู้สอบกลุ่มหนึ่ง มีโอกาสในการตอบข้อสอบถูกมากกว่าผู้สอบอีกกลุ่มหนึ่งอย่างสม่ำเสมอในทุกระดับความสามารถ เมื่อพิจารณาโค้งคุณลักษณะข้อสอบของผู้สอบ 2 กลุ่ม จะพบว่าไม่มีปฏิสัมพันธ์ระหว่างระดับความสามารถของผู้สอบกับการเป็นสมาชิกของกลุ่ม (Group membership)

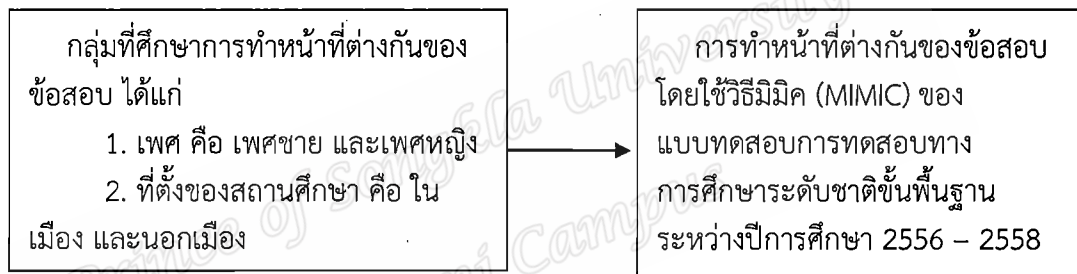
ข้อสอบทำหน้าที่ต่างกันแบบอนเอกรูป (Nonuniform DIF) หมายถึง ข้อสอบที่ทำให้โอกาสในการตอบข้อสอบถูกของผู้สอบระหว่างกลุ่มแตกต่างกันอย่างไม่สม่ำเสมอในทุกระดับความสามารถ เมื่อพิจารณาโค้งคุณลักษณะข้อสอบของผู้สอบ 2 กลุ่ม พบว่ามีปฏิสัมพันธ์ร่วมกันระหว่าง ระดับความสามารถของผู้สอบ กับการเป็นสมาชิกของกลุ่ม เช่น ที่ระดับความสามารถหนึ่งกลุ่มผู้สอบ กลุ่ม R มีโอกาสในการตอบข้อสอบถูกมากกว่ากลุ่มผู้สอบกลุ่ม F แต่ที่ระดับความสามารถอีกระดับหนึ่ง กลุ่มผู้สอบกลุ่ม F มีโอกาสในการตอบข้อสอบถูกมากกว่ากลุ่มผู้สอบกลุ่ม R

ข้อสอบ หมายถึง แบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (Ordinary National Educational Test : O-net) รายวิชาภาษาไทย ภาษาอังกฤษ คณิตศาสตร์ และวิทยาศาสตร์ ระหว่างปีการศึกษา 2556 – 2558

ประโยชน์ที่คาดว่าจะได้รับ

1. ได้ข้อเสนอแนะต่อหน่วยงานที่เกี่ยวข้องกับการสร้างข้อสอบ
2. เป็นข้อมูลในการตัดสินใจของผู้ที่จะนำวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบไปใช้ในการตรวจสอบคุณภาพของข้อสอบ ว่าสมควรที่จะตัดข้อสอบทำหน้าที่ต่างกันหรือควรปรับปรุงข้อสอบนั้นเสียใหม่
3. เป็นการแสดงวิธีการวิเคราะห์ข้อสอบแนวใหม่ ที่สามารถนำไปใช้ได้ทั้งในระดับสถานศึกษา ระดับเขตพื้นที่การศึกษา หรือระดับชาติ สามารถนำผลการสอบของผู้สอบมาวิเคราะห์และวินิจฉัยด้วยกระบวนการทางสถิติที่ได้นำไปใช้ในการวางนโยบายเพื่อยกระดับมาตรฐานการจัดการศึกษาได้

กรอบแนวคิดการวิจัย



บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยครั้งนี้มีวัตถุประสงค์ของการวิจัย คือ เพื่อวิเคราะห์คุณภาพแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2556 - 2558 และตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิค (Multiple Indicators and Multiple Causes : MIMIC) เมื่อจำแนกตามเพศและที่ตั้งของสถานศึกษา ระหว่างปีการศึกษา 2556 – 2558

ผู้วิจัยนำเสนอแนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้องกับมโนทัศน์ของการทำหน้าที่ต่างกันของข้อสอบ การวิเคราะห์ข้อสอบด้วยวิธีมิมิค (MIMIC) โดยศึกษาค้นคว้าและรวบรวมข้อมูลจากการสังเคราะห์จากหนังสือ เอกสาร บทความวิชาการ ผู้วิจัยนำเสนอเอกสารและงานวิจัยที่เกี่ยวข้อง ดังนี้

ตอนที่ 1 มโนทัศน์ของทฤษฎีการตอบสนองข้อสอบ

- 1.1 โมเดลการตอบสนองข้อสอบ (IRT Models)
- 1.2 ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ
- 1.3 พารามิเตอร์ในทฤษฎีการตอบสนองข้อสอบ

ตอนที่ 2 มโนทัศน์ของการทำหน้าที่ต่างกันของข้อสอบ

- 2.1 ความเป็นมาของการศึกษาการทำหน้าที่ต่างกันของข้อสอบ
- 2.2 ความหมายของการศึกษาการทำหน้าที่ต่างกันของข้อสอบ
- 2.3 ประเภทของการทำหน้าที่ต่างกันของข้อสอบ
- 2.4 หลักการและวิธีการตรวจสอบการตรวจสอบการทำหน้าที่ต่างกันของ

ข้อสอบ

ตอนที่ 3 มโนทัศน์ของการวิเคราะห์ข้อสอบด้วยวิธีมิมิค (MIMIC)

- 3.1 โมเดลมิมิค (MIMIC)
- 3.2 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิค (MIMIC)
- 3.3 ประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้

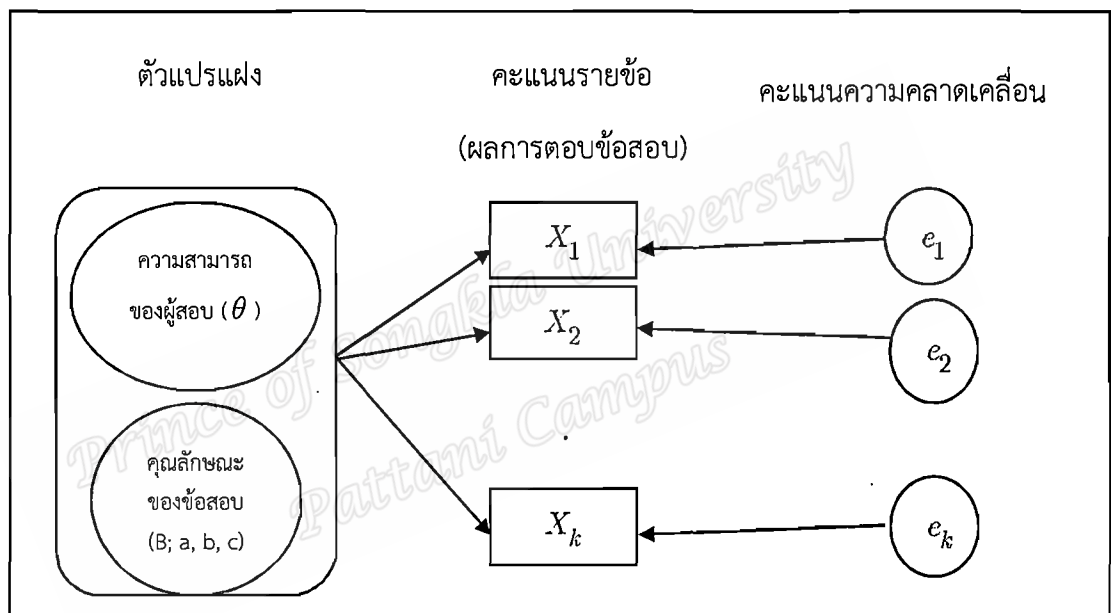
วิธีมิมิค

ตอนที่ 4 งานวิจัยที่เกี่ยวข้องกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

ตอนที่ 1 มโนทัศน์ของทฤษฎีการตอบสนองข้อสอบและการประยุกต์ใช้

1.1 โมเดลการตอบสนองข้อสอบ (IRT Models : Item Response Theory)

โมเดลการวัดเป็นระบบความสัมพันธ์ระหว่างตัวแปรอิสระที่รวมกันสำหรับทำนายตัวแปรตาม สำหรับโมเดลการตอบสนองข้อสอบ ตัวแปรอิสระประกอบด้วย ตัวแปรแฝง คือ ความสามารถที่แท้จริงของผู้สอบ (θ) และคุณลักษณะของข้อสอบ (B) หรือค่าพารามิเตอร์ (a, b, c) ส่วนตัวแปรอิสระเป็นตัวแปรที่สังเกตได้ คือ โอกาสการตอบข้อสอบได้ถูกต้อง ดังแสดงในภาพประกอบ 1



ภาพประกอบ 1 โมเดลการตอบสนองข้อสอบ (IRT).

(ที่มา : ศิริชัย กาญจนวาสี, 2555)

ทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นทฤษฎีการวัดที่อธิบายความสัมพันธ์ระหว่างความสามารถที่มีอยู่ภายในบุคคล (Latent trait or ability) กับผลการตอบข้อสอบหรือข้อคำถาม โดยใช้โค้งลักษณะข้อสอบ (Item Characteristic Curve : ICC) ซึ่งมีการกำหนดลักษณะของข้อสอบด้วยพารามิเตอร์ความยาก (b) อำนาจจำแนก (a) และโอกาสการเดาข้อสอบถูก (c) IRT จึงอยู่บนฐานความคิดที่สำคัญ 2 ประการคือ 1) ผลการตอบข้อสอบหรือข้อคำถามของผู้ตอบสามารถอธิบายได้ด้วยความสามารถที่มีอยู่ภายในของผู้ตอบ และ 2) ความสัมพันธ์ระหว่างผลการตอบข้อสอบกับความความสามารถที่มีอยู่ภายใน สามารถอธิบายได้ด้วยฟังก์ชันลักษณะข้อสอบ หรือโค้งลักษณะข้อสอบ (ICC) อันมีลักษณะเป็นฟังก์ชันทางคณิตศาสตร์ เรียกว่า ฟังก์ชันโลจิส (Logistic function) หรือใกล้เคียงกับฟังก์ชันปกติสะสม (Normalogive function)

ทฤษฎีการตอบสนองข้อสอบ พยายามอธิบายความสัมพันธ์ระหว่างคุณลักษณะภายในหรือความสามารถที่มีอยู่ภายในตัวบุคคลกับพฤติกรรมการตอบสนองข้อสอบของบุคคลนั้นว่ามีโอกาสตอบข้อสอบถูกมากน้อยเพียงไร ทฤษฎีนี้มีพื้นฐานความเชื่อว่าพฤติกรรมการตอบสนองต่อข้อสอบของผู้สอบซึ่งเป็นสิ่งที่สังเกตได้โดยตรงว่าถูกหรือผิด จะถูกกำหนดโดยคุณลักษณะภายในหรือความสามารถที่อยู่ภายในตัวบุคคล ซึ่งเป็นสิ่งที่ไม่สามารถสังเกตได้โดยตรง ทฤษฎีนี้ได้อธิบายความสัมพันธ์ดังกล่าวในรูปของฟังก์ชันคณิตศาสตร์หรือโมเดลที่แสดงความสัมพันธ์ระหว่างระดับความสามารถคุณลักษณะของข้อสอบ และโอกาสของการตอบข้อสอบได้ถูก ที่เรียกว่า ฟังก์ชันการตอบสนองข้อสอบ ซึ่งมีลักษณะความสัมพันธ์เป็นแบบฟังก์ชันโลจิสหรือฟังก์ชันปกติสะสม

ฟังก์ชันการตอบสนองข้อสอบสามารถนำมาใช้ศึกษาความสัมพันธ์ระหว่างความน่าจะเป็นในการตอบข้อสอบแต่ละข้อได้ถูก [$P_i(\theta)$] กับระดับความสามารถของผู้สอบที่วัดได้โดยแบบสอบฉบับนั้น (θ) เมื่อนำมาเขียนเป็นกราฟจะได้โค้งลักษณะข้อสอบ (Item Characteristic Curve : ICC) โค้งลักษณะข้อสอบมีหลายลักษณะขึ้นอยู่กับโมเดล (Model) หรือแบบจำลองที่ใช้อธิบายความสัมพันธ์ดังกล่าว โมเดลที่นิยมใช้กันคือ โมเดลแบบหนึ่งพารามิเตอร์ (One-Parameter Model) โมเดลแบบสองพารามิเตอร์ (Two-parameter Model) และโมเดลแบบสามพารามิเตอร์ (Three-Parameter Model) (ศิริชัย กาญจนวาสี, 2555)

1.2 ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

ฟังก์ชันทางคณิตศาสตร์ที่ใช้ใน IRT กำหนดไว้ว่า ความน่าจะเป็นของการตอบข้อสอบได้ถูกต้องขึ้นอยู่กับความสามารถของผู้สอบและคุณลักษณะของข้อสอบ โมเดลการตอบสนองข้อสอบจึงตั้งอยู่บนฐานความเชื่อหรือข้อตกลงเบื้องต้นหลายประการเกี่ยวกับลักษณะข้อมูลที่จะทำให้โมเดลสามารถนำไปใช้ได้เหมาะสม ถึงแม้ข้อตกลงเบื้องต้นบางประการจะตรวจสอบไม่ได้โดยตรงแต่เราสามารถเก็บรวบรวมข้อมูลหรือหลักฐานทางอ้อมมาช่วยสนับสนุนยืนยันได้ ข้อตกลงเบื้องต้นที่สำคัญของ IRT มีดังนี้ (ศิริชัย กาญจนวาสี, 2555)

1.2.1 ความเป็นเอกมิติ (Unidimensionality)

ข้อตกลงเบื้องต้นที่ใช้กันทั่วไปสำหรับ IRT คือข้อคำถาม/ข้อสอบทุกข้อในเครื่องมือ/แบบสอบมุ่งวัดเพียงคุณลักษณะเดียวหรือความสามารถเดียว (One ability) ซึ่งเรียกว่าความเป็นเอกมิติ (Unidimensionality)

การตรวจสอบความเป็นเอกมิติของเครื่องมือหรือแบบสอบ สามารถกระทำได้โดยใช้เทคนิคการวิเคราะห์ทางสถิติ ได้แก่ การวิเคราะห์ตัวประกอบ (Factor Analysis) เพื่อคำนวณค่าไอเกน (Eigen value) สำหรับศึกษาอัตราส่วนระหว่างค่าไอเกนของตัวประกอบแรกกับตัวประกอบถัดไป ถ้ามีอัตราส่วนที่สูงแสดงถึงเครื่องมือหรือแบบสอบวัดคุณลักษณะเด่นเดียว (Single dominant factor) หรือทำการวิเคราะห์ให้เกิดความมั่นใจยิ่งขึ้นด้วยการวิเคราะห์ตัวประกอบเชิงยืนยัน (Confirmatory Factor Analysis) เพื่อตรวจสอบยืนยันว่าเครื่องมือหรือแบบสอบมุ่งวัดเพียง

คุณลักษณะเดียวหรือความสามารถเดียว (ศิริชัย กาญจนวาสี, 2555)

1.2.2 ความเป็นอิสระ (Local Independence)

แนวคิดเกี่ยวกับ “ความเป็นอิสระระหว่างข้อสอบและผู้สอบ” มีความเกี่ยวข้องกับ และเชื่อมโยงมาจาก “ความเป็นเอกมิติของแบบสอบ” ความเป็นอิสระระหว่างข้อสอบและผู้สอบ หมายถึงเมื่อมีการควบคุมความสามารถ (θ) ที่ส่งผลต่อการตอบข้อสอบ หรือให้ θ คงที่แล้ว ผลการตอบข้อสอบแต่ละข้อจะต้องเป็นอิสระจากกันหรือกล่าวอีกนัยหนึ่งได้ว่า เมื่อควบคุมอิทธิพลของ θ แล้วผลการตอบข้อสอบรายข้อไม่มีความสัมพันธ์กัน นั่นคือโมเดลการตอบสนองข้อสอบมีเพียง θ ปัจจัยเดียวเท่านั้นที่มีอิทธิพลต่อผลการตอบรายข้อ ความเป็นอิสระสามารถจำแนกการพิจารณาเป็น ความอิสระระหว่างข้อสอบและความอิสระระหว่างผู้สอบ (ศิริชัย กาญจนวาสี, 2555)

1.2.3 โมเดลการตอบสนองข้อสอบ (Item Response Models)

IRT อยู่บนฐานความเชื่อว่า ฟังก์ชันลักษณะข้อสอบหรือโค้งลักษณะข้อสอบ (ICC) สามารถสะท้อนความสัมพันธ์จริงระหว่างความสามารถของผู้สอบกับลักษณะของข้อสอบและผลการตอบข้อสอบโมเดลการตอบสนองข้อสอบเสนอ ICC ซึ่งเป็นฟังก์ชันโลจิส ด้วยรูปลักษณะที่ แตกต่างกันตามจำนวนพารามิเตอร์ที่ใช้บรรยายลักษณะของข้อสอบ โมเดลการตอบสนองข้อสอบ แบบตรวจให้คะแนน 2 ค่า (Dichotomous Item Response Models) ที่ใช้กันแพร่หลาย ได้แก่ โมเดลการตอบสนองข้อสอบ แบบ 1-พารามิเตอร์, 2-พารามิเตอร์ และ 3-พารามิเตอร์ การเลือกใช้จึงขึ้นกับจุดมุ่งหมายของงานและธรรมชาติของข้อมูล

โมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ มีข้อตกลงเบื้องต้นว่า ข้อสอบแต่ละข้อมีพารามิเตอร์ $c = 0$ และพารามิเตอร์ a เท่ากัน แต่มีความแตกต่างกันเฉพาะ พารามิเตอร์ b เท่านั้น โมเดลนี้จึงเหมาะสำหรับใช้กับข้อสอบอิงเกณฑ์ที่ไม่สลับซับซ้อน ข้อสอบที่ค่อนข้างเรียบง่ายสำหรับพัฒนาคลังข้อสอบที่มีความเป็นเอกพันธ์

โมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ มีข้อตกลงเบื้องต้นว่าข้อสอบ แต่ละข้อมีพารามิเตอร์ $c = 0$ มีความแตกต่างกันของพารามิเตอร์ a และ b โมเดลนี้จึงเหมาะสำหรับ ใช้กับข้อสอบที่ต้องเติมคำตอบหรือข้อสอบแบบเลือกตอบที่ไม่ยากมากนักและกลุ่มผู้สอบมีความพร้อมในการตอบ

โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ มีข้อตกลงเบื้องต้นว่าข้อสอบ แต่ละข้อมีความแตกต่างกันได้ทั้ง พารามิเตอร์ a , b และ c โมเดลนี้จึงเหมาะสำหรับใช้กับข้อสอบ แบบเลือกตอบทั่วไป ข้อสอบแบบหลายตัวเลือกเนื่องจากผู้สอบสามารถเดาคำตอบได้

การตรวจสอบความเหมาะสมของโมเดลการตอบสนองข้อสอบกับข้อมูล (Model - Data Fit) ไม่ว่าจะเลือกใช้โมเดลใดก็ตาม โมเดลจะต้องสอดคล้องกับข้อมูลเพื่อให้ผลการวิเคราะห์มีความถูกต้องและน่าเชื่อถือ การตรวจสอบความสอดคล้องควรพิจารณาทั้งความไม่แปรเปลี่ยนของค่าประมาณความสามารถ (Invariance of ability parameter estimates) และ

ความไม่แปรเปลี่ยนของค่าประมาณพารามิเตอร์ของข้อสอบ (Invariance of item parameter estimates)

ความไม่แปรเปลี่ยนของค่าประมาณความสามารถตรวจสอบได้โดยการเปรียบเทียบค่าประมาณความสามารถของผู้สอบที่ได้จากกลุ่มข้อสอบที่แตกต่างกัน เช่น กลุ่มข้อสอบยาก กลุ่มข้อสอบง่าย หรือกลุ่มข้อสอบจากคลังข้อสอบเดียวกัน แต่มีความครอบคลุมเนื้อหาแตกต่างกัน เป็นต้น ค่าประมาณความสามารถจะถือว่าไม่แปรเปลี่ยน เมื่อความแตกต่างเกิดขึ้นไม่เกินความคลาดเคลื่อนมาตรฐานของการประมาณค่า (Wright, 1968 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

ความไม่แปรเปลี่ยนของค่าประมาณพารามิเตอร์ของข้อสอบตรวจสอบได้โดยการเปรียบเทียบค่าประมาณพารามิเตอร์แต่ละตัวของข้อสอบที่ได้จากตัวอย่างประชากรผู้สอบหลายกลุ่ม เช่น กลุ่มผู้สอบชาย/หญิง กลุ่มผู้สอบจำแนกตามภูมิภาค เป็นต้น ค่าประมาณพารามิเตอร์ของข้อสอบจะถือว่าไม่แปรเปลี่ยน เมื่อผลการพล็อตกราฟออกมาเป็นเส้นตรง โดยมีการกระจายไม่แตกต่างจากผลที่ได้จากตัวอย่าง 2 กลุ่ม ซึ่งเป็นกลุ่มสุ่มที่ตัดเทียมกัน (Shepard, Camilli & Williams, 1984 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555) (ศิริชัย กาญจนวาสี, 2555)

1.2.4 การสอบที่ไม่แข่งขันด้านเวลา (Nonspeeded Test Administration)

IRT ถือว่าความสามารถ (θ) เป็นปัจจัยสำคัญต่อผลการสอบความเร็วในการตอบ จะต้องไม่มีอิทธิพลต่อผลการตอบ การจัดการสอบจึงต้องไม่อยู่ในสถานการณ์ที่สอบแข่งขันกันด้วยเวลาการสอบจะต้องอยู่ในลักษณะที่ผู้สอบซึ่งมีความสามารถมีเวลาเพียงพอในการทำข้อสอบ (Power Test Administration)

การตรวจสอบถึงความเหมาะสมของมิติด้านเวลา สำหรับการดำเนินการสอบสามารถพิจารณาได้จากสัดส่วนหรือร้อยละของจำนวนผู้สอบที่ทำข้อสอบได้ครบทุกข้อ โดยผู้สอบส่วนใหญ่ (เช่น ร้อยละ 80 เป็นต้น) สามารถตอบข้อสอบได้ครบหรือเกือบครบทุกข้อ นอกจากนี้ควรพิจารณาเปรียบเทียบระหว่างความแปรปรวนของจำนวนข้อที่เว้น กับ ความแปรปรวนของจำนวนข้อที่ตอบผิด (Gulliksen, 1950 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555) ถ้าอัตราส่วนของความแปรปรวนเข้าใกล้ 0 แสดงว่าการจัดการสอบเป็นไปตามเวลา (ศิริชัย กาญจนวาสี, 2555)

1.3 พารามิเตอร์ในทฤษฎีการตอบสนองข้อสอบ

พารามิเตอร์ในทฤษฎีการตอบสนองข้อสอบ แบ่งออกเป็น 2 ชนิด คือ พารามิเตอร์ข้อสอบ (Item parameter) ประกอบด้วย ความยาก (b) อำนาจจำแนก (a) การเดา (c) และความรอบคอบ (γ) และ พารามิเตอร์ผู้สอบ (person parameter) เป็นระดับความสามารถหรือคุณลักษณะของผู้สอบ (θ) พิสัยของพารามิเตอร์ต่างๆ มีดังนี้ (Hambleton and Swaminathan, 1985: อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

พารามิเตอร์ความยาก (b) ทางทฤษฎีมีค่าตั้งแต่ $-\infty$ ถึง ∞ แต่ทางปฏิบัติจะมีค่าอยู่ระหว่าง -2.5 ถึง +2.5 ค่าที่เป็นลบแสดงว่าข้อสอบง่าย และค่าที่เป็นบวกแสดงว่าข้อสอบยาก

พารามิเตอร์อำนาจจำแนก (a) ในทางทฤษฎีมีค่าตั้งแต่ $-\infty$ ถึง ∞ ควรมีค่าเป็นบวกตามปกติมีค่าไม่เกิน +2.5 ในทางปฏิบัตินิยมใช้ข้อสอบที่มีค่าพารามิเตอร์อำนาจจำแนกอยู่ระหว่าง +0.5 ถึง +2.5 พารามิเตอร์การเดา (c) เป็นค่าแสดงความน่าจะเป็นหรือโอกาสของการตอบข้อสอบได้ถูกต้อง โดยไม่มีความรอบรู้หรือคุณลักษณะในเรื่องนั้นๆ ในทางทฤษฎีพารามิเตอร์การเดามีค่าระหว่าง 0 ถึง 1 โดยทั่วไปนิยมใช้ข้อสอบที่มีค่าพารามิเตอร์การเดาไม่เกิน 0.30

ความรอบคอบ (γ) (Barton and Lora, 1980 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555) โดยเสนอพารามิเตอร์ที่แสดงถึงความรอบคอบของผู้สอบซึ่งเป็นค่าที่บ่งชี้ว่าผู้สอบที่มีความสามารถสูงอาจตอบข้อสอบได้ไม่ถูกต้องเสมอไป เนื่องจากความไม่รอบคอบในการพิจารณาคำตอบ หรือผู้สอบอาจจะมีสารสนเทศอื่นเกี่ยวกับผู้ออกข้อสอบ จึงเลือกตอบในตัวเลือกที่ไม่ใช่คำตอบที่ถูกต้อง Barton และ Lord กล่าวว่าพารามิเตอร์ตัวนี้จะเหมาะสมในการศึกษาทางทฤษฎีเท่านั้น แต่ในทางปฏิบัติแล้วไม่สามารถพบค่าพารามิเตอร์ตัวนี้ (Hambleton and Swaminatan, 1985 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

พารามิเตอร์ผู้สอบเป็นระดับความสามารถของผู้สอบ (θ) ที่ประมาณได้จากโมเดลตามทฤษฎีการตอบสนองข้อสอบ นิยมปรับให้เป็นคะแนนมาตรฐานที่มีค่าเฉลี่ยเป็น 0 และส่วนเบี่ยงเบนมาตรฐานเป็น 1 มีค่าระหว่าง $-\infty$ ถึง ∞ แต่ส่วนใหญ่จะมีค่าอยู่ในช่วง -3.0 ถึง +3.0 ค่าที่เป็นลบแสดงว่าผู้สอบมีความสามารถต่ำและค่าที่เป็นบวกแสดงว่าผู้สอบมีความสามารถสูง (ศิริชัย กาญจนวาสี, 2555)

Prince of Songkla
Pattani Campus

ตอนที่ 2 มโนทัศน์ของการทำหน้าที่ต่างกันของข้อสอบ

2.1 ความเป็นมาของการศึกษาการทำหน้าที่ต่างกันของข้อสอบ

การศึกษาถึงคุณภาพของข้อสอบจากผลการตรวจข้อสอบของผู้สอบกลุ่มต่างๆ ในประชากรมีมานานแล้ว แต่การศึกษาคุณภาพด้านความยุติธรรมของข้อสอบหรือแบบสอบระหว่างผู้สอบกลุ่มต่างๆ เริ่มศึกษากันอย่างจริงจังในช่วงปลายทศวรรษของปี ค.ศ. 1960 มีการเสนอวิธีการต่างๆ เพื่อตรวจสอบความลำเอียงของข้อสอบ (Item bias) ความลำเอียงของแบบสอบ (Test bias) และความลำเอียงในการคัดเลือก (Selection bias) โดยนิยามความลำเอียงว่าเป็น ความคลาดเคลื่อนอย่างเป็นระบบ (Systematic error) ที่เกิดขึ้นจากการวัด ความพยายามของการตรวจสอบความลำเอียงดังกล่าว ดำเนินไปเพื่อจำแนกข้อสอบที่ทำหน้าที่ไม่เหมาะสมหรือไม่ยุติธรรมสำหรับปรับปรุงหรือตัดข้อสอบข้อนั้นออกจากแบบสอบ เป็นการขจัดข้อสอบที่ทำให้เกิดปัญหาความยุติธรรมระหว่างกลุ่มข้อสอบกลุ่มต่างๆ ที่มีลักษณะบางอย่างแตกต่างกัน เช่น เชื้อชาติ ศาสนา วัฒนธรรม ภูมิฐานะ สังคม เพศ ภาษา อายุ ประสบการณ์ เป็นต้น เพื่อพัฒนาแบบสอบให้มีคุณภาพเหมาะสมสำหรับนำไปใช้ทดสอบต่อไป (ศิริชัย กาญจนวาสี, 2555)

ในเวลาต่อมานักวัดผลการศึกษาก็ได้ทำการศึกษาค่าความลำเอียงของข้อสอบ (Item bias) กันอย่างกว้างขวาง ทำให้เกิดความสับสนของการใช้คำและความหมายมีประเด็นโต้แย้งกันว่าความลำเอียงของข้อสอบ เป็นผลการตัดสินว่าข้อสอบมีความยุติธรรมหรือไม่ อันส่งผลต่อการบรรลุจุดมุ่งหมายของการใช้แบบสอบหรือความลำเอียงของข้อสอบ เป็นสารสนเทศทางสถิติที่ได้จากข้อสอบเกี่ยวกับความสัมพันธ์ระหว่างคุณลักษณะที่ข้อสอบมุ่งวัดกับประสบการณ์ของผู้สอบกลุ่มต่างๆ ที่ทำการสอบ เมื่อกลุ่มผู้สอบต่างกลุ่มกันตอบข้อสอบข้อเดียวกัน ความแตกต่างที่เกิดขึ้นอาจมาจากความไม่เหมาะสมของข้อคำถาม ซึ่งสามารถเกิดขึ้นได้หลายลักษณะหรือประสบการณ์ของผู้สอบซึ่งอาจมีลักษณะพื้นฐานเดิมแตกต่างกันในหลายสถานการณ์จึงไม่เหมาะสมที่จะใช้คำว่า ข้อสอบลำเอียง (Biased item) เนื่องจากเป็นภาษาที่มีความหมายในเชิงลบ ประกอบกับเกณฑ์ที่ใช้สำหรับตัดสินความลำเอียงยังมีความคลุมเครือและค่อนข้างสับสน ดังนั้น จึงควรเปลี่ยนมาใช้คำว่า การทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning : DIF) ซึ่งเป็นคำที่มีความเป็นกลางและเหมาะสมกว่า (Holland & Thayer, 1988 : Holland & Wainer, 1993 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

การทำหน้าที่ต่างกันของข้อสอบ (DIF) กับความลำเอียงของข้อสอบ (Item bias) มีแนวคิดที่แตกต่างกัน สำหรับการทำหน้าที่ต่างกันของข้อสอบ เป็นกระบวนการที่เน้นการใช้วิธีการทางสถิติสำหรับตรวจสอบ เพื่อให้ได้สารสนเทศเกี่ยวกับการทำหน้าที่ของข้อสอบสำหรับกลุ่มผู้สอบกลุ่มย่อยที่มีลักษณะเฉพาะบางอย่างแตกต่างกัน ส่วนความลำเอียงของข้อสอบเป็นกระบวนการตัดสินความยุติธรรมของข้อสอบ โดยนำสารสนเทศการทำหน้าที่ต่างกันของข้อสอบมาวิเคราะห์เชิงตรรกะ (Logical analysis) โดยผู้เชี่ยวชาญพิจารณาถึงการเขียนข้อสอบเนื้อหาสาระของข้อสอบและจุดมุ่งหมายของการวัดเพื่อระบุว่าข้อสอบข้อนั้นลำเอียงเข้าข้างกลุ่มใดหรือไม่ เพราะเหตุใดจึงเป็นการตัดสินความลำเอียงของข้อสอบ (Camilli & Shapard, 1994 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

2.2 ความหมายของการศึกษาการทำหน้าที่ต่างกันของข้อสอบ

นักวิจัยทางการวัดผลหลายท่านได้ให้ความหมายของความลำเอียงของข้อสอบและการทำหน้าที่ต่างกันของข้อสอบ ไว้ดังนี้

ความลำเอียงของข้อสอบ หมายถึง ลัดส่วนของผู้สอบที่ตอบข้อสอบได้ถูกต้องไม่เท่ากัน ในแต่ละกลุ่มประชากรที่ใช้ในการศึกษา เมื่อกลุ่มผู้สอบมีคะแนนเท่ากันและข้อสอบมีความเป็นเอกพันธ์ (Scheuneman, 1979 อ้างถึงใน ศิริชัย กาญจนวาสิ, 2555)

ความลำเอียงของข้อสอบ หมายถึง ข้อสอบที่มีค่าความยากสัมพัทธ์สำหรับสมาชิกของผู้สอบกลุ่มหนึ่งมากกว่าสมาชิกของผู้สอบอีกกลุ่มหนึ่ง (Rudner, Getson & Knight, 1980 อ้างถึงใน ศิริชัย กาญจนวาสิ, 2555)

ความลำเอียงของข้อสอบ หมายถึง ความโน้มเอียงของข้อสอบที่เมื่อใช้คะแนนจากข้อสอบนั้น แล้วทำให้การตัดสินผลเป็นไปอย่างไม่ยุติธรรม (Popham, 1981 อ้างถึงใน ศิริชัย กาญจนวาสิ, 2555)

ความลำเอียงของข้อสอบ หมายถึง โอกาสในการตอบข้อสอบได้ถูกต้องแตกต่างกันสำหรับการวัดความสามารถ หรือโอกาสในการตอบข้อสอบในทางบวกแตกต่างกันสำหรับการวัดเจตคติ เมื่อผู้สอบที่มีคุณลักษณะของการวัดในปริมาณเท่ากัน แต่มาจากกลุ่มประชากรย่อยที่แตกต่างกัน (Hulin, Drasgow & Parson, 1983 อ้างถึงใน ศิริชัย กาญจนวาสิ, 2555)

ความลำเอียงของข้อสอบ หมายถึง โอกาสในการตอบข้อสอบได้ถูกต้องของผู้สอบกลุ่มหนึ่งมีค่าต่ำกว่าหรือสูงกว่าผู้สอบอีกกลุ่มหนึ่งที่มีระดับความสามารถเดียวกัน (Dorans & KuLick, 1986 อ้างถึงใน ศิริชัย กาญจนวาสิ, 2555)

ความลำเอียงของข้อสอบ หมายถึง คะแนนข้อสอบของกลุ่มผู้สอบที่มีความสามารถเท่ากัน แต่มาจากต่างกลุ่มกันมีความแตกต่างกันอย่างเป็นระบบ (Kederman, 1990 อ้างถึงใน ศิริชัย กาญจนวาสิ, 2555)

การทำหน้าที่ต่างกันของข้อสอบ หมายถึง สารสนเทศทางสถิติของข้อสอบที่ได้จากการตอบของผู้สอบต่างกลุ่มกัน และมีความสามารถเท่ากัน แต่มีโอกาสนในการตอบข้อสอบได้ถูกต้องแตกต่างกัน (Holland & Wainer, 1993 อ้างถึงใน ศิริชัย กาญจนวาสิ, 2555)

การทำหน้าที่ต่างกันของข้อสอบ หมายถึง ความเป็นพหุมิติในการวัดของข้อสอบ ซึ่งแสดงได้จากการแจกแจงความสามารถหลัก (Primary ability) ของกลุ่มผู้สอบตั้งแต่ 2 กลุ่มขึ้นไปมีความเท่ากัน แต่มีการแจกแจงความสามารถรอง (Secondary ability) แตกต่างกัน (Camilli & Shepard, 1994 อ้างถึงใน ศิริชัย กาญจนวาสิ, 2555)

การทำหน้าที่ต่างกันของข้อสอบ หมายถึง ฟังก์ชันการตอบสนองข้อสอบซึ่งคำนวณจากกลุ่มผู้สอบกลุ่มย่อยที่ต่างกันมีค่าไม่เท่ากัน (Narayanan & Swaminathan, 1996 อ้างถึงใน ศิริชัย กาญจนวาสิ, 2555)

ดังนั้นจึงสรุปได้ว่า การทำหน้าที่ต่างกันของข้อสอบ (DIF) หมายถึง การที่ข้อสอบทำให้ผู้สอบจากต่างกลุ่มกันที่มีความสามารถหรือคุณลักษณะที่มุ่งวัดเท่ากัน มีโอกาสนในการตอบข้อสอบได้ถูกต้องแตกต่างกันหรือมีฟังก์ชันการตอบสนองข้อสอบแตกต่างกัน การทำหน้าที่ต่างกันของ

ข้อสอบ เกิดขึ้นเมื่อนำข้อสอบไปทดสอบกับผู้สอบกลุ่มย่อยต่างกันที่มีความสามารถหลัก (Primary ability) ระดับเดียวกันหรือมีคุณลักษณะแฝง (Latent trait) ที่ต้องการวัดเท่ากัน แต่มีความสามารถรอง (Secondary ability) แตกต่างกัน ทำให้ผู้สอบต่างกลุ่มที่นำมาจับคู่เปรียบเทียบมีโอกาสตอบข้อสอบถูกแตกต่างกัน (ศิริชัย กาญจนวาสี, 2555)

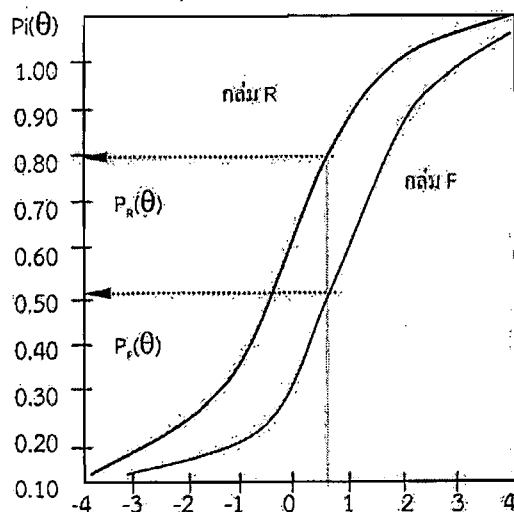
2.3 ประเภทของการทำหน้าที่ต่างกันของข้อสอบ

การทำหน้าที่ต่างกันของข้อสอบ เป็นการเปรียบเทียบผลการตอบข้อสอบระหว่างกลุ่มผู้สอบอย่างน้อย 2 กลุ่มขึ้นไป ปกตินิยมทำการเปรียบเทียบ 2 กลุ่ม ประกอบด้วยกลุ่มแรก เรียกว่า กลุ่มเปรียบเทียบ (Focal group หรือกลุ่ม F) เป็นกลุ่มที่สนใจศึกษาและคาดว่าจะจะเป็นกลุ่มที่เสียเปรียบในการตอบข้อสอบ และกลุ่มที่สอง เรียกว่า กลุ่มอ้างอิง (Reference group หรือกลุ่ม R) เป็นกลุ่มที่คาดว่าจะได้เปรียบในการตอบข้อสอบได้ถูกต้อง

ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ จะพบว่า ข้อสอบสามารถทำหน้าที่แตกต่างกันได้ 2 ประเภท (Mellenbergh, 1982 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555, น.118) ได้แก่ การทำหน้าที่ต่างกันของข้อสอบแบบเอกรูป (Uniform) และ แบบบอเนกรูป (Nonuniform) (ศิริชัย กาญจนวาสี, 2555)

2.3.1 ข้อสอบทำหน้าที่ต่างกันแบบเอกรูป (Uniform DIF)

ข้อสอบที่ทำให้ผู้สอบกลุ่มหนึ่ง มีโอกาสในการตอบข้อสอบถูกมากกว่าผู้สอบอีกกลุ่มหนึ่งอย่างสม่ำเสมอในทุกระดับความสามารถ เมื่อพิจารณาโค้งคุณลักษณะข้อสอบของผู้สอบ 2 กลุ่ม จะพบว่าไม่มีปฏิสัมพันธ์ระหว่างระดับความสามารถของผู้สอบกับการเป็นสมาชิกของกลุ่ม (Group membership) (ศิริชัย กาญจนวาสี, 2555) ดังภาพประกอบ 2



ภาพประกอบ 2 ข้อสอบทำหน้าที่ต่างกันแบบเอกรูป (Uniform DIF).

(ที่มา : ศิริชัย กาญจนวาสี, 2555)

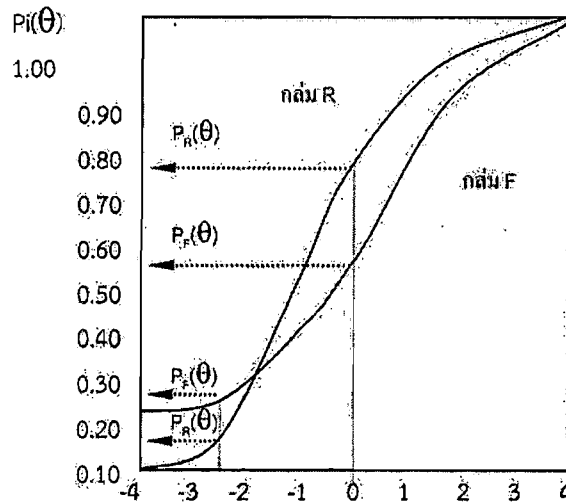
2.3.2 ข้อสอบทำหน้าที่ต่างกันแบบอเนกรูป (Nonuniform DIF)

ข้อสอบที่ทำให้โอกาสในการตอบข้อสอบถูกของผู้สอบระหว่างกลุ่มแตกต่างกันอย่างไม่สม่ำเสมอในทุกระดับความสามารถ เมื่อพิจารณาโค้งคุณลักษณะข้อสอบของผู้สอบ 2 กลุ่ม พบว่ามีปฏิสัมพันธ์ร่วมกันระหว่าง ระดับความสามารถของผู้สอบกับการเป็นสมาชิกของกลุ่ม เช่น ที่ระดับความสามารถหนึ่งกลุ่มผู้สอบกลุ่ม R มีโอกาสในการตอบข้อสอบถูกมากกว่ากลุ่มผู้สอบกลุ่ม F แต่ที่ระดับความสามารถอีกระดับหนึ่ง กลุ่มผู้สอบกลุ่ม F มีโอกาสในการตอบข้อสอบถูกมากกว่ากลุ่มผู้สอบกลุ่ม R (ศิริชัย กาญจนวาสี, 2555)

ตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) สามารถพิจารณา “ปฏิสัมพันธ์” ดังกล่าวได้จากความแตกต่างของค่าพารามิเตอร์อำนาจจำแนกของข้อสอบระหว่างผู้สอบกลุ่มย่อยสองกลุ่ม กล่าวคือ ถ้าข้อสอบทำหน้าที่ต่างกันแบบเอกรูป แล้วโค้งคุณลักษณะข้อสอบ (Item Characteristic Curves : ICCs) ระหว่างผู้สอบกลุ่มย่อยสองกลุ่มจะขนานกันหรือมีฟังก์ชันการตอบสนองข้อสอบ (Item Response Functions : IRFs) เหมือนกัน แต่ถ้าข้อสอบทำหน้าที่ต่างกันแบบอเนกรูป แล้วโค้งคุณลักษณะข้อสอบระหว่างผู้สอบกลุ่มย่อยสองกลุ่มจะไม่ขนานกันหรือมีฟังก์ชันการตอบสนองข้อสอบต่างกัน ดังนั้นความแตกต่างระหว่างโค้งคุณลักษณะข้อสอบทั้งสองแบบจะบ่งบอกถึงขนาดและทิศทางของ ข้อสอบที่ทำหน้าที่ต่างกันซึ่งสามารถคำนวณได้โดยใช้สูตรการคำนวณพื้นที่ของ Raju (Raju ,1990 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555) ข้อสอบที่ทำหน้าที่ต่างกันแบบอเนกรูปสามารถจำแนกได้เป็น 2 ลักษณะดังนี้

1) ข้อสอบทำหน้าที่ต่างกันแบบอเนกรูปโดยมีปฏิสัมพันธ์ไม่เป็นลำดับ (Disordinal interaction)

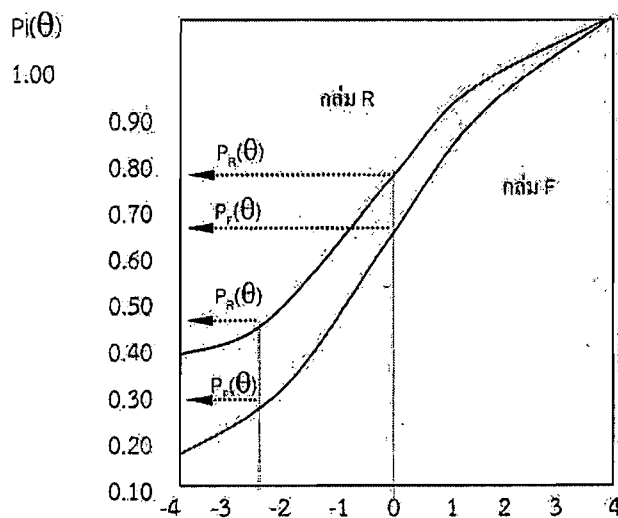
เป็นการทำหน้าที่ต่างกันสำหรับกลุ่มผู้สอบซึ่งเกิดขึ้น เมื่อโค้งคุณลักษณะข้อสอบตัดกันระหว่าง ช่วงความสามารถของผู้สอบหรือเรียกว่าข้อสอบทำหน้าที่ต่างกันแบบไม่มีทิศทาง (Non-Unidirectional DIF) (ศิริชัย กาญจนวาสี, 2555, น.119) ดังภาพประกอบ 3



ภาพประกอบ 3 ข้อสอบทำหน้าที่ต่างกันแบบไม่มีทิศทาง (Non-unidirectional DIF).
(ที่มา : ศิริชัย กาญจนวาสี, 2555)

2) ข้อสอบทำหน้าที่ต่างกันแบบอนุกรม โดยมีปฏิสัมพันธ์เป็นลำดับ (Ordinal interaction)

เป็นการทำหน้าที่ต่างกันสำหรับกลุ่มผู้สอบซึ่งเกิดขึ้น เมื่อโค้งลักษณะข้อสอบต่างกัน อย่างไม่สม่ำเสมอ แต่ไม่ตัดกัน หรืออาจตัดกันนอกช่วง ความสามารถของผู้สอบตรงปลายสุดของช่วงความสามารถต่ำหรือสูง อาจเรียกข้อสอบลักษณะนี้ว่า ข้อสอบทำหน้าที่ต่างกันแบบมีทิศทางเดียว (Unidirectional DIF) (ศิริชัย กาญจนวาสี, 2555, น.119) ดังภาพประกอบ 4



ภาพประกอบ 4 ข้อสอบทำหน้าที่ต่างกันแบบมีทิศทางเดียวกัน (Unidirectional DIF).
(ที่มา : ศิริชัย กาญจนวาสี, 2555)

2.4 หลักการและวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

2.4.1 หลักการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

การตรวจการทำหน้าที่ต่างกันของข้อสอบ (DIF detection) เป็นการเปรียบเทียบผลการตอบข้อสอบเป็นรายข้อระหว่างกลุ่มผู้สอบอย่างน้อย 2 กลุ่ม ที่มีความสามารถหลัก (primary ability) ที่มุ่งวัดเท่ากัน แต่คาดว่าจะมีความได้เปรียบหรือเสียเปรียบกันโดยกลุ่มหนึ่งถือเป็นกลุ่มอ้างอิง (Reference group) ซึ่งคาดว่าจะน่าจะได้เปรียบในการตอบข้อสอบข้อนั้น หรือมีโอกาสตอบข้อสอบได้ถูกต้องมากกว่า ส่วนอีกกลุ่มคือกลุ่มเปรียบเทียบ (Focal group) ซึ่งเป็นกลุ่มที่สนใจศึกษาและคาดว่าจะน่าจะเป็นกลุ่มที่เสียเปรียบ (ศิริชัย กาญจนวาสี, 2555)

ในการเปรียบเทียบผลการตอบข้อสอบระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบจำเป็นต้องจับคู่ (matching) ผู้สอบตามความสามารถ ซึ่งเป็นเงื่อนไขสำคัญของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เกณฑ์การจับคู่ (matching criteria) ที่นิยมใช้กันมี 2 วิธี ดังนี้

1) เกณฑ์ภายนอก (External Criterion)

การวิเคราะห์การทำหน้าที่ต่างกันโดยใช้เกณฑ์ภายนอกนี้สามารถนำไปใช้ได้ทั้งข้อสอบรายข้อและแบบสอบทั้งฉบับ โดยการใช้คะแนนจากแบบสอบอื่นเป็นเกณฑ์ภายนอกแล้วใช้เทคนิคการวิเคราะห์การถดถอย (Regression analysis) เพื่อทำการเปรียบเทียบเส้นกราฟความสัมพันธ์ระหว่างตัวแปรเกณฑ์ กับตัวแปรทำนายระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ

หลักการนี้มีจุดมุ่งหมายเพื่อสร้างสมการทำนายตัวแปรเกณฑ์ ซึ่งเป็นคะแนนของแบบสอบ อื่นจากตัวแปรทำนายซึ่งเป็นคะแนนรายข้อ หรือคะแนนแบบสอบระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบใน การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบจะใช้คะแนนรายข้อเป็นตัวแปรทำนายแต่ถ้าเป็นการวิเคราะห์การทำหน้าที่ต่างกันของแบบสอบจะใช้คะแนนรวมของแบบสอบทั้งฉบับเป็นตัวแปรทำนาย สำหรับตัวแปรเกณฑ์ที่ใช้เป็นเกณฑ์ภายนอกอาจใช้คะแนนรวมทั้งฉบับหรือเกรดเฉลี่ย หรือผลสัมฤทธิ์ในงานที่เกี่ยวข้องของผู้สอบ (Cronbach, 1970 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555) สมการทำนายสำหรับกลุ่ม อ้างอิงและกลุ่มเปรียบเทียบ แสดงได้ดังนี้

$$\text{กลุ่มอ้างอิง (R)} \quad Y_i = A_R + B_R X_i$$

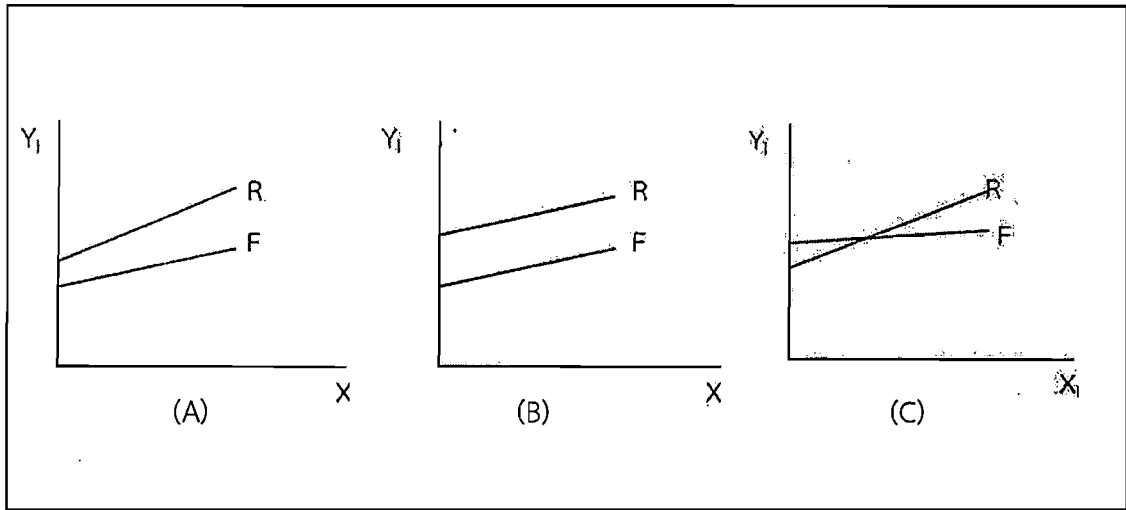
$$\text{กลุ่มเปรียบเทียบ (F)} \quad Y_i = A_F + B_F X_i$$

เมื่อ Y_i คือ คะแนนของตัวแปรเกณฑ์ภายนอก

X_i คือ คะแนนของตัวแปรทำนาย

A คือ ค่าคงที่หรือค่าตัดแกน (Intercept)

B คือ ค่าความชัน (Slope)



ภาพประกอบ 5 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบและแบบสอบ โดยใช้เทคนิคการวิเคราะห์สมการถดถอย. (ที่มา : ศิริชัย กาญจนวาสี, 2555)

จากฟังก์ชันการทำนายทั้ง 2 สมการ สามารถเปรียบเทียบค่าตัดแกน (A) และค่าความชัน (B) ของเส้นกราฟระหว่างกลุ่มอ้างอิง (R) และกลุ่มเปรียบเทียบ (F) ได้ ถ้าเส้นกราฟดังกล่าวมีค่าความชันหรือค่าตัดแกนแตกต่างกัน สำหรับข้อสอบใดหรือแบบสอบใดแสดงว่าข้อสอบหรือแบบสอบนั้นมีการทำหน้าที่ต่างกัน โดยลำเอียงเข้าข้างกลุ่มผู้สอบที่มีค่าตัดแกนตั้งสูงกว่าหรือค่าความชันที่สูงกว่า

การใช้เกณฑ์ภายนอกมีข้อดี คือเกณฑ์ที่ใช้มีความเป็นอิสระจากข้อสอบและแบบสอบที่ต้องการตรวจสอบ แต่มีจุดอ่อนตรงที่ความเหมาะสมของเกณฑ์ที่จะนำมาใช้ในทางปฏิบัติเป็นการยากที่จะหาตัวแปรเกณฑ์ภายนอกจากแบบสอบฉบับอื่นที่มีความตรงเชิงทำนายและมีความยุติธรรมสำหรับกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ ถ้าตัวแปรเกณฑ์ภายนอกขาดคุณสมบัติดังกล่าวจะทำให้ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบหรือแบบสอบขาดความแม่นยำและความสมบูรณ์ (ศิริชัย กาญจนวาสี, 2555)

2) เกณฑ์ภายใน (Internal Criterion)

การวิเคราะห์การทำหน้าที่ต่างกัน โดยใช้เกณฑ์ภายในเป็นการนำวิธีการทางสถิติมาตรวจสอบการทำหน้าที่ต่างกันของข้อสอบหรือแบบสอบ โดยเน้นการพิจารณาจากโครงสร้างภายในของแบบสอบเป็นหลักด้วยการวิเคราะห์ผลจากการตอบข้อสอบและความสามารถหรือคะแนนจริงของผู้สอบที่ได้จากแบบสอบฉบับนั้น เพื่อนำมาเปรียบเทียบระหว่างผู้สอบจากกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่มีความสามารถหรือคะแนนจริงเท่ากันว่าจะมีผลการตอบหรือโอกาสในการตอบข้อสอบได้ถูกต้องแตกต่างกันหรือไม่ เพื่อบ่งชี้ถึงการทำหน้าที่ต่างกันของข้อสอบ

การวิเคราะห์ในลักษณะนี้นิยมใช้ค่าสถิติต่างๆ เป็นตัวบ่งชี้ถึงการทำหน้าที่ต่างกันของข้อสอบ ค่าสถิติทดสอบที่นิยมนำมาใช้พอสรุปได้ดังนี้ (ศิริชัย กาญจนวาสี, 2555)

2.1) การทดสอบปฏิสัมพันธ์ (Interaction)

ในระยะเริ่มแรกของการศึกษาความลำเอียงของข้อสอบมีการใช้สถิติทดสอบ เอฟ (F-test) จากการวิเคราะห์ความแปรปรวน (ANOVA) เพื่อทดสอบปฏิสัมพันธ์ระหว่างกลุ่มผู้สอบกับข้อสอบ ถ้าการทดสอบมีนัยสำคัญเป็นสัญญาณของการทำหน้าที่ต่างกันของข้อสอบ (Cleary & Hilton, 1968: Jensen, 1974 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555) จากนั้นจึงทำการวิเคราะห์ต่อด้วยวิธีการ Post Hoc เพื่อระบุข้อสอบที่มีผลต่อการเกิดปฏิสัมพันธ์ ซึ่งเป็นข้อที่ทำหน้าที่ต่างกัน

วิธีการนี้มีข้อดีที่สามารถศึกษาผู้สอบหลายๆ กลุ่มได้สะดวกแต่มีจุดอ่อนในเรื่องการควบคุมกลุ่มต่างๆ ให้มีความสามารถที่ทัดเทียมกัน ขนาดตัวอย่างของกลุ่มต่างๆ และอัตราความคลาดเคลื่อนประเภทที่ 1 จะสูงขึ้น ถ้าจำนวนข้อสอบเพิ่มมากขึ้น (ศิริชัย กาญจนวาสี, 2555)

2.2) การวัดความเบี่ยงเบนสัมพัทธ์ (Relative Deviation)

การคำนวณค่าความยากของข้อสอบ เช่น p , b เป็นต้น เมื่อคำนวณแยกแยะระหว่างกลุ่มและแปลงให้เป็นค่าความยากมาตรฐาน (A) สามารถนำมาพล็อตเปรียบเทียบเป็นรายข้อ ถ้าข้อใดเบี่ยงเบนไปจากแกนหลักที่คาดหมายหรือเบี่ยงเบนเกินจากความคลาดเคลื่อนมาตรฐานของค่าความยากที่กำหนดย่อมแสดงถึงการทำหน้าที่ต่างกันของข้อสอบ (Cleary & Hilton, 1968 : Angoff & Ford , 1973 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555) รวมทั้งสามารถคำนวณค่าสหสัมพันธ์ระหว่างค่าความยากรายข้อระหว่างกลุ่ม เพื่อแสดงถึงการทำหน้าที่ต่างกันของแบบสอบ ถ้าสหสัมพันธ์เข้าใกล้ 1.00 แสดงว่าค่าความยากสัมพัทธ์ของข้อสอบมีค่าใกล้เคียงกันระหว่างกลุ่ม ดังนั้นแบบสอบวัดคุณลักษณะคล้ายกันระหว่างกลุ่ม

วิธีการนี้มีข้อดีและข้อเสียคล้ายการทดสอบปฏิสัมพันธ์ นอกจากนี้ค่าความยากของข้อสอบ (p) มิใช่ตัวแทนของค่าความยากจริงของข้อสอบและได้รับอิทธิพลจากค่าแทรกซ้อนอื่นได้แก่ ค่าอำนาจจำแนก และความสามารถของผู้สอบ (ศิริชัย กาญจนวาสี, 2555)

2.3) การเปรียบเทียบน้ำหนักตัวประกอบ (Factor Loading)

การวิเคราะห์ตัวประกอบ (Factor Analysis) เป็นเทคนิคทางสถิติที่นิยมใช้ในการ ตรวจสอบความตรงเชิงทฤษฎีหรือโครงสร้าง (Construct Validity) เมื่อนำการวิเคราะห์ตัวประกอบมา ใช้ในการวิเคราะห์โครงสร้างของแบบสอบแยกตามกลุ่มผู้สอบ ความไม่สอดคล้องกันระหว่างน้ำหนักตัว ประกอบบนคุณลักษณะสำคัญที่มุ่งวัดหรือความแตกต่างของค่าเฉลี่ย

คะแนน ตัวประกอบ (Factor scores) ระหว่างกลุ่มผู้สอบย่อมสะท้อนการทำหน้าที่ต่างกันของข้อสอบและแบบสอบ

การใช้เทคนิคการวิเคราะห์ตัวประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA) สำหรับศึกษาการทำหน้าที่ต่างกัน จะมีจุดอ่อนในเรื่องความไม่สอดคล้องระหว่างน้ำหนักตัว ประกอบอาจเกิดจากความแตกต่างของความสามารถระหว่างกลุ่มก็ได้

แนวทางที่เหมาะสมจึงควรใช้เทคนิคการวิเคราะห์ตัวประกอบเชิงยืนยัน (Confirmatory Factor Analysis: CFA) นอกจากนี้ยังสามารถใช้ CFA สำหรับตรวจสอบความแตกต่างระหว่างกลุ่ม ในด้านคุณลักษณะหรือความสามารถหลักและความสามารถรองได้อีกด้วย (Camilli & Shepard, 1994 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555) (ศิริชัย กาญจนวาสี, 2555)

2.4) การเปรียบเทียบโอกาสตอบข้อสอบถูก

การวิเคราะห์โอกาสตอบข้อสอบถูกของผู้สอบจากกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่มีความสามารถเท่ากัน เป็นแนวทางสำคัญที่นิยมใช้กันและเป็นที่ยอมรับในปัจจุบัน สำหรับบ่งชี้การทำหน้าที่ต่างกันของข้อสอบ มีการคำนวณค่าสถิติ 2 แนวทางดังนี้

2.4.1) เปรียบเทียบค่าสัดส่วนหรือความน่าจะเป็น

ในการตอบข้อสอบถูกของผู้สอบ ต่างกลุ่มที่มีความสามารถเท่ากัน เช่น วิธีแมนเทิล-แฮนส์เซล (MH) เป็นต้น

2.4.2) เปรียบเทียบค่าฟังก์ชันการตอบสนองข้อสอบหรือโค้งลักษณะข้อสอบระหว่าง กลุ่มที่มีระดับความสามารถเท่ากัน

เป็นวิธีที่อยู่บนพื้นฐานของทฤษฎี IRT เช่น วิธีวัดความแตกต่างของพื้นที่ วิธีวัดความแตกต่างของค่าพารามิเตอร์ความยาก วิธีการทดสอบไคสแควร์ของลอร์ด (Lord's χ^2 -test) เป็นต้น

วิธีการนี้มีข้อดีที่สำคัญได้แก่ การคำนวณค่าสถิติของข้อสอบมีความน่าเชื่อถือ มีกลไกควบคุมความสามารถของผู้สอบโดยการจับคู่กลุ่มความสามารถเพื่อทำการเปรียบเทียบ ณ ตำแหน่งต่างๆ ที่มีความสามารถเท่ากัน จึงเป็นวิธีการที่ยอมรับกันทั่วไป แต่มีข้อจำกัดในด้านความสลับซับซ้อนของแนวคิดพื้นฐาน และการวิเคราะห์ที่มีความจำเป็นต้องใช้โปรแกรมคอมพิวเตอร์โดยเฉพาะ (ศิริชัย กาญจนวาสี, 2555)

2.4.2 วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF detection) จำแนกตามลักษณะ การตรวจให้คะแนนได้เป็น 2 ประเภทคือ ข้อสอบที่มีการให้คะแนนแบบทวิภาคหรือสองค่า (Dichotomous scoring) และข้อสอบที่มีการให้คะแนนแบบพหุภาค หรือหลายค่า (Polytomous scoring) วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแต่ละประเภท ยังสามารถจำแนกได้อีก 2 มิติ ได้แก่ มิติลักษณะของตัวแปรเกณฑ์ ซึ่งแบ่งเป็นกลุ่มวิธีที่ใช้คะแนนสังเกตได้ (Observed score) และกลุ่มวิธีที่ใช้คะแนนสังเกตไม่ได้หรือคะแนนของตัวแปรแฝง (Latent variable) และมีติลักษณะของสถิติวิเคราะห์ ซึ่งแบ่งเป็นกลุ่มวิธีที่ใช้สถิติพารามตริก (Parametric approach) และกลุ่มวิธีที่ใช้สถิตินันพารามตริก (Nonparametric approach) รายชื่อวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ที่สำคัญๆ ดังแสดงในตาราง 1 (Potenza & Dorans, 1995; Feinstein, 1995 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555) (ศิริชัย กาญจนวาสี, 2555)

ตาราง 1 วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ที่มีการตรวจให้คะแนนแบบทวิภาค (Dichotomous DIF) และพหุภาค (Polytomous DIF).

ประเภทและตัวแปรเกณฑ์	พารามตริก	นั้พารามตริก
1. DIF แบบทวิภาค		
1.1 คะแนนที่สังเกตได้ (observed score)	ANOVA Logistic Regression	TID MH STND
	IRT-D ² Lord's X ² General IRTL	SIBTEST
1.2 คุณลักษณะ/ตัวแปรแฝง (latent variable)	Loglinear IRTL MIMIC	
2. DIF แบบพหุภาค		
2.1 คะแนนที่สังเกตได้ (observed score)	ANOVA Polytomous Logistic Regression	Polytomous STND GMH
2.2 คุณลักษณะ/ตัวแปรแฝง (latent variable)	General IRTL PCM	Polytomous SIBTEST GPCM

(ที่มา: ศิริชัย กาญจนวาสี, 2555)

2.4.2.1 วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ให้คะแนนแบบ ทวิวิภาค

1) กลุ่มวิธีที่ใช้คะแนนที่สังเกตได้

วิธีในกลุ่มนี้มีกี่วิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) หรือกลุ่มที่ไม่ใช้ทฤษฎีการตอบสนองข้อสอบ (Non-IRT approach) โดยใช้คะแนนรวมของผู้สอบเป็นเกณฑ์การจับคู่ของกลุ่มผู้สอบ วิธีการตรวจสอบที่สำคัญในกลุ่มนี้ได้แก่ (ศิริชัย กาญจนวาสี, 2555)

1.1) การวิเคราะห์ความแปรปรวน (ANOVA) (Cleary & Hilton, 1968 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

1.2) วิธีการวิเคราะห์การถดถอยโลจิสติก (Logistic Regression: LR) (Swaminathan & Rogers, 1990 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

1.3) วิธีแปลงค่าความยากของข้อสอบ (Transformed Item Difficulty: TID) (Cleary & Hilton: 1968; Angoff & Ford, 1973 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

1.4) วิธีแมนเทล-แฮนส์เซล (Mantel - Haenszel: MH) (Holland & Thayer, 19125: 1989 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

1.5) วิธีดัชนีมาตรฐาน (Standardization: STND) การปรับให้เป็นมาตรฐานด้วยน้ำหนักตัวประกอบ (Dorans & Kulick, 1986 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

2) กลุ่มวิธีที่ใช้คุณลักษณะแฝง

วิธีในกลุ่มนี้ใช้คุณลักษณะหรือตัวแปรแฝง ซึ่งวิเคราะห์บนพื้นฐานของทฤษฎีการตอบสนองข้อสอบ (IRT) สำหรับใช้เป็นเกณฑ์การจัดคู่กลุ่มผู้สอบ วิธีการตรวจสอบที่สำคัญในกลุ่มนี้ได้แก่ (ศิริชัย กาญจนวาสี, 2555)

2.1) วิธีวัดพื้นที่ความแตกต่างระหว่างโค้งการตอบสนองข้อสอบ (IRT-D²) (Linn et. al, 1981; Shepard et. al, 1984; Raju, 1990; Kim & Cohen, 1991 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

2.2) วิธีไคสแควร์ของลอร์ด (Lord's X²) (Lord, 1980 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

2.3) วิธีอัตราส่วนไลค์ลิฮูดทั่วไป (General IRT Likelihood Ratio) (Thissen, Steinberg & Wainer, 1993 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

2.4) วิธีอัตราส่วนไลค์ลิฮูด ลอกลิเบียร์ (Loglinear IRT Likelihood Ratio) (Thissen, Steinberg & Wainer, 1993 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

2.5) วิธีซิปเทสต์ (SIBTEST) (Shealy & stout, 1993 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

2.6) วิธีมิมิค (MIMIC) (B.O. Muthen ,1985 อ้างถึง Carol M. Woods and Kevin J. Grimm, 2011)

2.4.2.2 วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ให้คะแนนแบบ
พหุวิภาค (ศิริชัย กาญจนวาสี, 2555)

1) กลุ่มวิธีที่ใช้คะแนนที่สังเกตได้

1.1) วิธีการวิเคราะห์ความแปรปรวน (ANOVA) (Cleary & Hilton, 1968 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

1.2) วิธีการวิเคราะห์การถดถอยโลจิสติกพหุวิภาค (Polytomous Logistic Regression) (Swaminathan & Rogers, 1990 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

1.3) วิธีดัชนีมาตรฐานพหุวิภาค (Polytomous standardization) (Dorans & Kulick, 1986 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

1.4) วิธีแมนเทล - แฮนส์เซลทั่วไป (General Mantel - Haenzel: GMH) (Holland & Thayer, 1988 , 1989 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

2) กลุ่มวิธีที่ใช้คุณลักษณะแฝง (ศิริชัย กาญจนวาสี, 2555)

2.1) วิธีอัตราส่วนโลคัลลิสูดในรูปทั่วไป (General IRT Likelihood Ratio) (Thissen, Steinberg & Wainer, 1993 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

2.2) วิธีการให้คะแนนบางส่วน (Partial Credit Model: PCM) (Master, 1982 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

2.3) วิธีชิปเทสต์พหุวิภาค (Polytomous SIBTEST) (Shealy & Stout, 1993 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

2.4) วิธีการให้คะแนนบางส่วนทั่วไป (Generalized Partial Credit Model:GPCM) (Muraki, 1992 , 1993 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

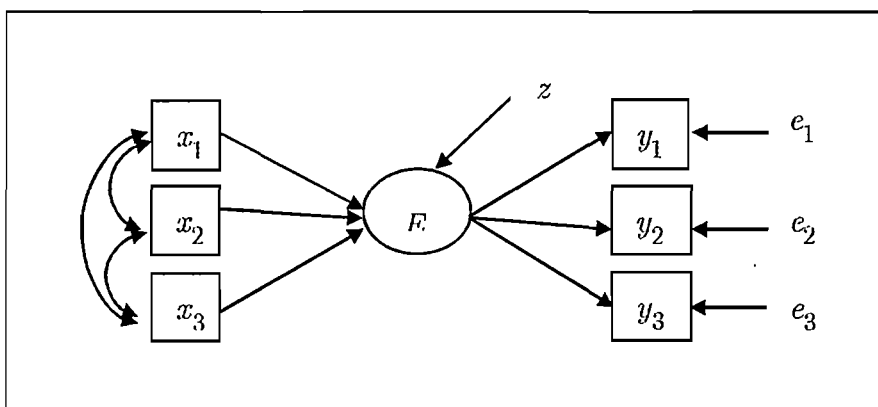
ตอนที่ 3 โมโนทัศน์ของการวิเคราะห์ข้อสอบด้วยวิธีมิมีค (MIMIC)

3.1 โมเดลมิมีค (MIMIC)

MIMIC เป็นคำที่ย่อมาจาก Multiple Indicators and Multiple Causes ซึ่งโมเดลมิมีค (MIMIC) สามารถวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบได้หลายโปรแกรมทางสถิติขั้นสูง เช่น โปรแกรม Lisrel และ โปรแกรม M Plus เป็นต้น ซึ่งผู้วิจัยได้สนใจการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดย MIMIC เพราะ 1) โมเดล MIMIC เป็นวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ค่อนข้างใหม่ 2) มีการลดจำนวนพารามิเตอร์ในการทดสอบความแตกต่างระหว่างกลุ่ม (Ian A. Carroll, 2014) 3) มีการลดลงของขนาดตัวอย่างให้น้อยลง ในขณะที่รูปแบบอื่นๆ เช่น การวิเคราะห์องค์ประกอบเชิงยืนยันหลายกลุ่ม (multiple-group CFA models) จำเป็นต้องใช้ขนาดตัวอย่างจำนวนมาก เพื่อให้ได้ผลลัพธ์ที่มีความน่าเชื่อถือ (Woods, 2009) 4) โมเดล MIMIC มีกระบวนการในการวิเคราะห์ที่สะดวกมากยิ่งขึ้น เมื่อมีจำนวนกลุ่มตัวแปรหลายกลุ่ม เมื่อเปรียบเทียบกับรูปแบบอื่นๆ เช่น การวิเคราะห์องค์ประกอบเชิงยืนยันหลายกลุ่ม (multiple-group CFA models) (Brown, 2006) 5) มีข้อตกลงเบื้องต้นเกี่ยวกับขนาดตัวอย่างที่น้อยกว่า เมื่อเปรียบเทียบกับรูปแบบอื่นๆ เช่น การวิเคราะห์องค์ประกอบเชิงยืนยันหลายกลุ่ม (multiple-group CFA models) ที่ข้อตกลงว่า ต้องมีขนาดตัวอย่างใหญ่ (Brown, 2006)

3.1.1 โมเดลมิมีค (MIMIC) ในโปรแกรม Lisrel

โมเดลมิมีค (MIMIC) หมายถึง โมเดลที่มีตัวแปรแฝงเพียงตัวแปรเดียว โดยที่ตัวแปรแฝงนั้นได้รับอิทธิพลจากตัวแปรภายนอกสังเกตได้หลายตัวแปรและส่งอิทธิพลไปยังตัวแปรภายในสังเกตได้หลายตัวแปร กล่าวอีกอย่างหนึ่งคือ เป็นโมเดลของคุณลักษณะแฝงที่มีหลายสาเหตุและวัดได้จากตัวบ่งชี้หลายตัว (นงลักษณ์ วิรัชชัย, 2538) ดังแสดงในภาพประกอบ 6

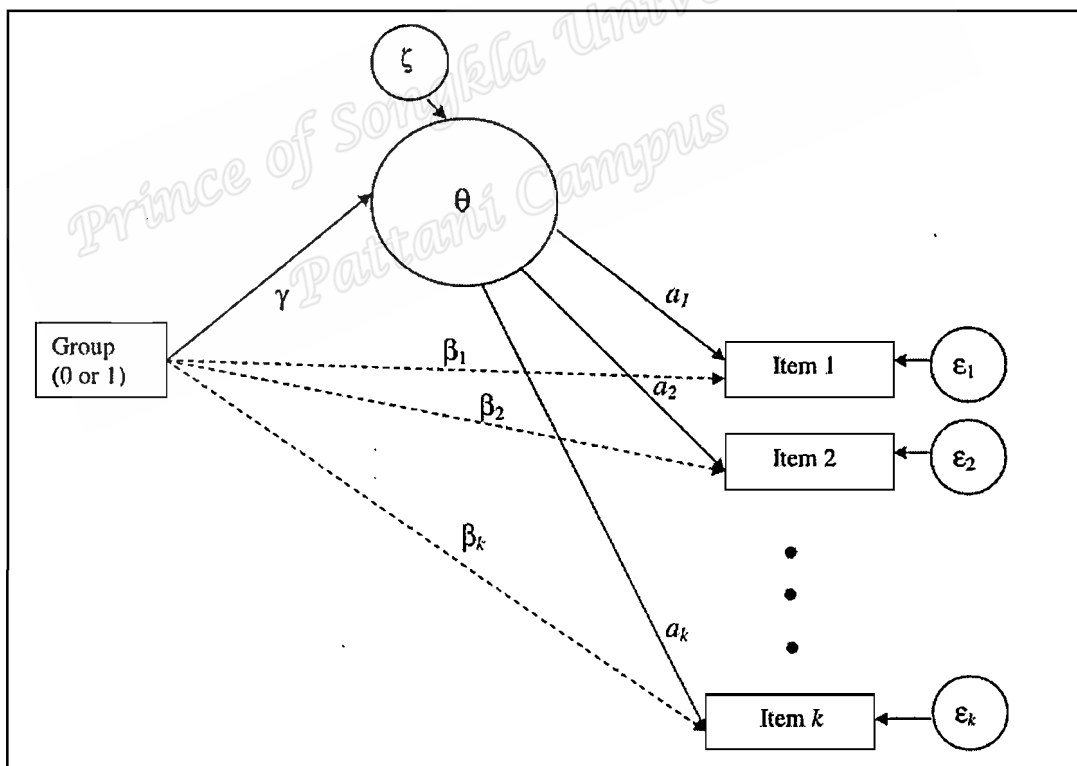


ภาพประกอบ 6 โมเดลย่อยในโปรแกรม Lisrel: โมเดลมิมีค (Multiple Indicators and Multiple Causes Models). (ที่มา: นงลักษณ์ วิรัชชัย, 2538)

ในที่นี้มีตัวบ่งชี้ 3 ตัวแปร และมีตัวแปรสาเหตุ 3 ตัวแปร ตามลักษณะโมเดลจะเห็นว่า การวัดตัวแปรภายนอกสังเกตได้ต้องมีข้อตกลงเบื้องต้นว่าไม่มีความคลาดเคลื่อนในการวัด และในการวิเคราะห์ข้อมูลจะกำหนดข้อมูลเฉพาะ เฉพาะรูปแบบและสถานะของเมทริกซ์ PH, BE, GA, PS, LY, TE เท่านั้น ส่วนเมทริกซ์ TD, LX มีค่าเป็นศูนย์ทั้งหมด โมเดลนี้มีคนี้เป็นประโยชน์มากในการตรวจสอบความเป็นเอกมิติ (Unidimensionality) ในการวิจัยสาขาการวัดผลการศึกษา (นงลักษณ์ วิรัชชัย, 2538)

3.1.2 โมเดลมิมิค (MIMIC) ในโปรแกรม M Plus

โมเดลมิมิค (MIMIC) ในโปรแกรม M Plus ซึ่งโดยปกติรูปแบบการวิเคราะห์ข้อมูลตามทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นการศึกษาคุณสมบัติอยู่บนพื้นฐานข้อตกลงความเป็นเอกมิติ (Unidimensionality) ของตัวแปรแฝงซึ่งสังเกตไม่ได้โดยตรง สำหรับตัวแปรแฝงในโมเดล IRT จะดูจากค่าเซตา (θ) ซึ่งสามารถประมาณค่าได้โดยตรงซึ่งมีอิทธิพลตรงต่อตัวชี้วัดหรือข้อคำถามที่สังเกตได้ การศึกษา (Carol M. Woods, 2009) ดังแสดงในภาพประกอบ 7



ภาพประกอบ 7 โมเดลย่อยในโปรแกรม M Plus: โมเดลมิมิค (Multiple Indicators and Multiple Causes Models). (ที่มา: Carol M. Woods, 2009)

ซึ่งในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธี MIMIC ในครั้งนี้ ผู้วิจัยได้ใช้โปรแกรม Mplus เพราะมีการพัฒนาอย่างต่อเนื่อง โปรแกรม Mplus สามารถออกแบบวิธีการวิเคราะห์ข้อมูลเป็นรูปแบบเดียวกันเกือบทุกโมเดล เพื่อลดความยุ่งยากในขั้นตอนของการวิเคราะห์สถิติขั้นสูงของโปรแกรมดั้งเดิมที่ใช้วิเคราะห์กันอยู่โดยทั่วไปโดยข้อดีของโปรแกรม Mplus ในการเขียนคำสั่งไม่ต้องเสียเวลาจำเมทริกซ์พารามิเตอร์ เหมือนกับในโปรแกรม LISREL

3.2 การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธี MIMIC

ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธี MIMIC พบว่า ข้อสอบสามารถทำหน้าที่ต่างกันได้ 2 ประเภท ได้แก่ การทำหน้าที่ต่างกันของข้อสอบแบบเอกรูป (Uniform) และแบบอเนกรูป (Nonuniform)

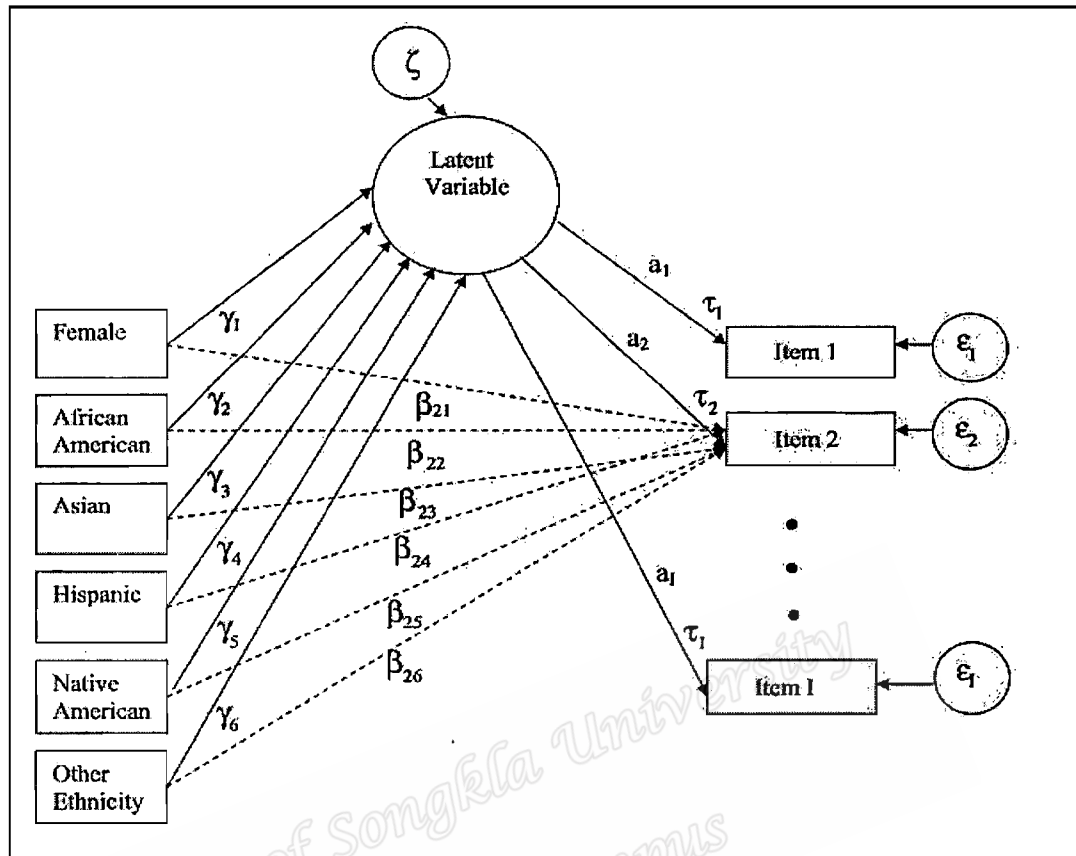
3.2.1 การทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธี MIMIC แบบเอกรูป (Uniform)

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้โมเดล MIMIC ของ B.O. Muthen (1985) เป็นการทดสอบการทำหน้าที่ต่างกันของข้อสอบแบบเอกรูปเท่านั้น เพราะไม่มีการปฏิสัมพันธ์ระหว่างตัวแปรแฝง (θ) กับกลุ่มของตัวแปรสังเกตได้ (z)

สูตรสำหรับวิธี MIMIC ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) แบบเอกรูป (Carol M. Woods and Kevin J. Grimm, 2011) คือ

$$y_i^* = \alpha_i \theta + \beta_i z + \varepsilon_i$$

โดย	y_i^*	คือ การตอบสนองการทำหน้าที่ต่างกันของข้อสอบข้อที่ i
	α	คือ อำนาจจำแนกของพารามิเตอร์
	θ	คือ ตัวแปรแฝง
	β	คือ สัมประสิทธิ์ของการถดถอยของการทดสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธี MIMIC แบบเอกรูป
	z	คือ กลุ่มตัวแปร
	ε	คือ ความผิดพลาดในการวัดของข้อสอบข้อที่ i



ภาพประกอบ 8 โมเดล MIMIC ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบเอกรูป บนข้อสอบข้อที่ 2. (ที่มา: Carol M. Woods & Thomas F. Oltmanns & Eric Turkheimer, 2009)

จากภาพประกอบ 8 โมเดล MIMIC ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบเอกรูป บนข้อสอบข้อที่ 2 ซึ่ง

γ_i คือ สัมประสิทธิ์การถดถอยของความแตกต่างค่าเฉลี่ยระหว่างตัวแปรสังเกตได้และตัวแปรแฝง (θ)

β_{ji} คือ สัมประสิทธิ์การถดถอยของความแตกต่างค่าเฉลี่ยระหว่างความยากของข้อสอบข้อที่ j และตัวแปรร่วม i

a_j คือ อำนาจจำแนกพารามิเตอร์

τ_i คือ ความยากพารามิเตอร์

ϵ_j คือ ความผิดพลาดในการวัดของข้อสอบข้อที่ j

ζ คือ เศษเหลือของตัวแปรแฝง (θ)

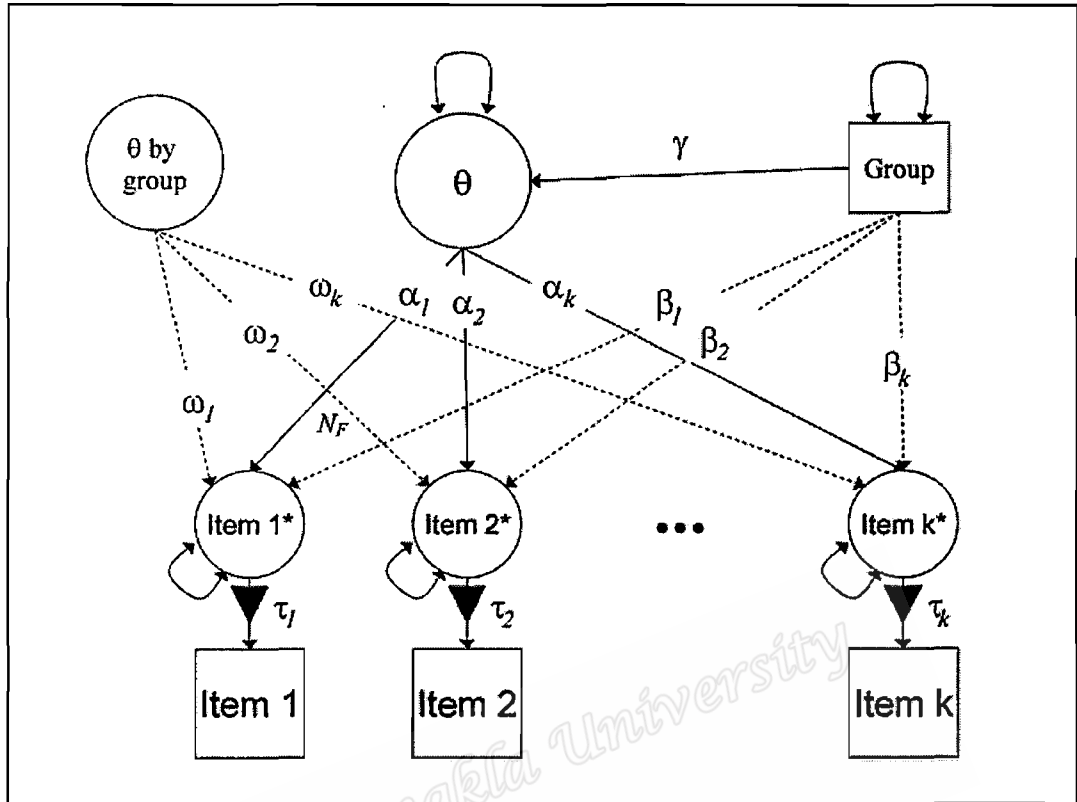
3.2.2 การทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธี MIMIC แบบอเนกรูป (Nonuniform)

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้โมเดล MIMIC ของ (B.O. Muthen ,1985 อ้างถึง Carol M. Woods and Kevin J. Grimm, 2011) เป็นการทดสอบการทำหน้าที่ต่างกันของข้อสอบแบบเอกรูปเท่านั้น เพราะมีการปฏิสัมพันธ์ระหว่างตัวแปรแฝง (θ) กับกลุ่มของตัวแปรสังเกตได้ (z)

สูตรสำหรับวิธี MIMIC ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) แบบอเนกรูป (Carol M. Woods and Kevin J. Grimm, 2011) คือ

$$y_i^* = \alpha_i\theta + \alpha_i z + \omega_i\theta z + \varepsilon_i$$

ข้อสอบข้อที่ i	โดย	y_i^*	คือ การตอบสนองการทำหน้าที่ต่างกันของ
		α	คือ อำนาจจำแนกของพารามิเตอร์
		θ	คือ ตัวแปรแฝง
		ω	คือ สัมประสิทธิ์ของการถดถอยของการทดสอบ
การทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธี MIMIC แบบอเนกรูป		z	คือ กลุ่มตัวแปร
		ε	คือ ความผิดพลาดในการวัดของข้อสอบข้อที่ i



ภาพประกอบ 9 โมเดล MIMIC สำหรับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เมื่อมีปฏิสัมพันธ์ระหว่าง กลุ่มของตัวแปรกลุ่มและตัวแปรแฝง (θ).

(ที่มา: Carol M. Woods and Kevin J. Grimm, 2011)

จากภาพประกอบ 9 โมเดล MIMIC สำหรับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เมื่อมีปฏิสัมพันธ์ระหว่าง กลุ่มของตัวแปรกลุ่มและตัวแปรแฝง (θ)

γ_j คือ ความแตกต่างของค่าเฉลี่ยระหว่างกลุ่มตัวแปรสังเกตได้และ ตัวแปรแฝง (θ)

α_i คือ อำนาจจำแนกของข้อสอบข้อที่ i

ω_{ij} คือ อิทธิพลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอนุกรมบนข้อสอบข้อที่ i

τ_i คือ ความยากของข้อสอบข้อที่ i

β_i คือ ความแตกต่างของกลุ่มตัวแปรสังเกตได้ในความยากของข้อสอบข้อที่ i

3.3 การตรวจสอบความตรงของโมเดล (Validation of the model)

ขั้นตอนที่สำคัญในการวิเคราะห์โมเดลอีกขั้นตอนหนึ่ง คือ การตรวจสอบความตรงของโมเดลที่เป็นสมมุติฐานวิจัยหรือการประเมินผลความถูกต้องของโมเดลหรือการตรวจสอบความกลมกลืนระหว่างข้อมูลเชิงประจักษ์กับโมเดล โดยมี 5 วิธีการในการตรวจสอบดังนี้

3.3.1 ความคลาดเคลื่อนมาตรฐานและสหสัมพันธ์ของค่าประมาณพารามิเตอร์ (Standard Error Correlations of Estimates)

ผลจากการวิเคราะห์ข้อมูลจะให้ค่าประมาณพารามิเตอร์ ความคลาดเคลื่อนมาตรฐาน ค่าสถิติที และสหสัมพันธ์ระหว่างค่าประมาณ ถ้าค่าประมาณที่ได้ไม่มีนัยสำคัญแสดงว่าความคลาดเคลื่อนมาตรฐานมีขนาดใหญ่และโมเดลการวิจัยอาจจะยังไม่ดีพอ ถ้าสหสัมพันธ์ระหว่างค่าประมาณมีค่าสูงมากเป็นสัญญาณแสดงว่าโมเดลการวิจัยใกล้จะไม่เป็นบวกแน่นอนและเป็นโมเดลที่ไม่ดีพอ (นงลักษณ์ วิรัชชัย, 2538)

3.3.2 สหสัมพันธ์พหุคูณและสัมประสิทธิ์การพยากรณ์ (Multiple Correlations and Coefficients of Determination)

ผลจากการวิเคราะห์ข้อมูล จะให้ค่าสหสัมพันธ์และสัมประสิทธิ์การพยากรณ์สำหรับตัวแปรสังเกตได้แยกที่ละตัวและรวมทุกตัวรวมทั้งสัมประสิทธิ์ การพยากรณ์ของสมการโครงสร้างด้วย ค่าสถิติเหล่านี้ควรมีค่าสูงสุดไม่เกินหนึ่งและค่าที่สูงแสดงว่าโมเดลมีความตรง (นงลักษณ์ วิรัชชัย, 2538)

3.3.3 ค่าสถิติวัดระดับความกลมกลืน (Goodness of fit Measures)

ค่าสถิติในกลุ่มนี้ใช้ตรวจสอบความตรงของโมเดลเป็นภาพรวมทั้งโมเดล มิใช่เป็นการตรวจสอบเฉพาะค่าพารามิเตอร์ แต่ละตัวเหมือนค่าสถิติสองประเภทแรกในทางปฏิบัตินักวิจัยควรใช้ค่าสถิติวัดระดับความกลมกลืนตรวจสอบความตรงของโมเดลทั้งโมเดล แล้วตรวจสอบความตรงของพารามิเตอร์แต่ละตัวโดยพิจารณาค่าสถิติสองประเภทแรกด้วย เพราะในบางกรณีแม้ว่าค่าสถิติวัดระดับความกลมกลืนจะแสดงว่าโมเดลกลมกลืนกับข้อมูลเชิงประจักษ์ แต่อาจจะมีพารามิเตอร์บางค่าไม่มีนัยสำคัญก็ได้ นอกจากนี้ค่าสถิติวัดระดับความกลมกลืนยังใช้ประโยชน์ในการเปรียบเทียบโมเดลที่แตกต่างกันสองโมเดลได้ด้วยว่า โมเดลใดมีความสอดคล้องกับข้อมูลเชิงประจักษ์มากกว่ากัน ค่าสถิติในกลุ่มนี้มี 4 ประเภท ดังต่อไปนี้

3.3.3.1 ค่าสถิติไค-สแควร์ (Chi-Square Statistics)

ค่าสถิติไคสแควร์เป็นค่าสถิติใช้ทดสอบสมมติฐานทางสถิติว่าฟังก์ชันความถ่วงมีค่าเป็นศูนย์ การคำนวณค่าไคสแควร์คำนวณจากผลคูณขององศาอิสระกับค่าของฟังก์ชันความถ่วง ถ้าค่าสถิติไคสแควร์มีค่าสูงมากแสดงว่าฟังก์ชันความถ่วงมีค่าแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติ นั่นคือ โมเดลไม่มีความสอดคล้องกับข้อมูลเชิงประจักษ์ ถ้าค่าสถิติไคสแควร์มีค่าต่ำมาก ยังมีค่าใกล้ศูนย์มากเท่าไร แสดงว่าโมเดลสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์

การใช้ค่าสถิติไคสแควร์เป็นค่าสถิติวัดระดับความถ่วงต้องใช้ด้วยความระวัง เพราะข้อตกลงเบื้องต้นของค่าสถิติไคสแควร์ มีอยู่ 4 ประการ คือ

- 1) ตัวแปรภายนอกสังเกตได้ต้องมีการแจกแจงปกติ ไม่มีค่าความโค้ง (ความโค้งเป็นศูนย์)
- 2) การวิเคราะห์ข้อมูลต้องใช้เมทริกซ์ ความแปรปรวนความแปรปรวนร่วมในการคำนวณ
- 3) ขนาดของตัวอย่างต้องมีขนาดใหญ่ เพราะฟังก์ชันความถ่วงจะมีการแจกแจงแบบไคสแควร์ต่อเมื่อตัวอย่างมีขนาดใหญ่เท่านั้น
- 4) ฟังก์ชันความถ่วงมีค่าเป็นศูนย์จริงตามสมมติฐานที่ใช้ทดสอบไคสแควร์ ดังนั้นในการวิเคราะห์โมเดล นักวิจัยควรตรวจสอบข้อมูลก่อนว่าตัวอย่างมีขนาดใหญ่ ตัวแปรภายในมีค่าความโค้งใกล้ศูนย์ และใช้เมทริกซ์ความแปรปรวน-ความแปรปรวนร่วม ในการวิเคราะห์ข้อมูลด้วยโปรแกรม

สำหรับเรื่องขนาดของตัวอย่าง กำหนดว่า กรณีข้อมูลในการวิเคราะห์โมเดลเป็นตัวแปรที่มีการแจกแจงปกติพหุนามทุกตัวควรใช้ตัวอย่างขนาดเท่ากับหรือมากกว่า (นงลักษณ์ วิรัชชัย, 2538)

3.3.3.2 ดัชนีวัดระดับความถ่วงกลมกลืน (Goodness-of-Fit Index: GFI)

ดัชนี GFI เป็นดัชนีที่พัฒนาขึ้นเพื่อใช้ประโยชน์จากค่าไคสแควร์ในการเปรียบเทียบระดับความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ของโมเดลสองโมเดล หลักการพัฒนา GFI คือ การนำค่าไคสแควร์มาพิจารณา ถ้าค่าไคสแควร์มีค่าสูงเมื่อเทียบกับองศาอิสระ นักวิจัยปรับโมเดลใหม่แล้ววิเคราะห์ข้อมูลอีกครั้งหนึ่ง ค่าไคสแควร์ที่ได้ใหม่นี้ถ้ามีค่าลดลงมากกว่าค่าแรก แสดงว่าโมเดลใหม่มีความสอดคล้องกับข้อมูลเชิงประจักษ์ดีขึ้น ดัชนี GFI เป็นอัตราส่วนของผลต่างระหว่างฟังก์ชันความถ่วงกลมกลืนจากโมเดลก่อนปรับและหลังปรับโมเดล กับฟังก์ชันความถ่วงกลมกลืนก่อนปรับโมเดล

ดัชนี GFI จะมีค่าอยู่ระหว่าง 0 และ 1 และเป็นค่าที่ไม่ขึ้นกับขนาดของตัวอย่าง แต่ลักษณะการแจกแจงขึ้นอยู่กับขนาดของตัวอย่าง Anderson และ Gerbing (1984) พบว่า ค่าเฉลี่ยของการแจกแจงค่าสถิติจากตัวอย่างสุ่มมีค่าเพิ่มขึ้น เมื่อขนาดของตัวอย่างมีค่าสูงขึ้น

ดัชนี GFI ที่เข้าใกล้ 1.00 แสดงว่าโมเดลมีความกลมกลืนกับข้อมูลเชิงประจักษ์ (นงลักษณ์ วิรัชชัย, 2538)

3.3.3.3 ดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (Adjusted Goodness-of-Fit Index: AGFI)

เมื่อนำดัชนี GFI มาปรับแก้โดยคำนึงถึงขนาดขององศาความอิสระ ซึ่งรวมทั้งจำนวนตัวแปรและขนาดของตัวอย่าง จะได้ค่าดัชนี AGFI และมีคุณสมบัติเช่นเดียวกับดัชนี GFI (นงลักษณ์ วิรัชชัย, 2538)

3.3.3.4 ดัชนีรากของกำลังสองเฉลี่ยของเศษ (Root Mean Squared Residual: RMF)

ดัชนี RMR เป็นดัชนีที่ใช้เปรียบเทียบระดับความกลมกลืนกับข้อมูลเชิงประจักษ์ของโมเดลสองโมเดล เฉพาะกรณีที่เป็นการเปรียบเทียบโดยใช้ข้อมูลชุดเดียวกัน ในขณะที่ดัชนี GFI และ AGFI สามารถใช้เปรียบเทียบได้ทั้งกรณีข้อมูลชุดเดียวกันและข้อมูลต่างชุดกัน ดัชนี RMR บอกขนาดของเศษที่เหลือโดยเฉลี่ยจากการเปรียบเทียบระดับความกลมกลืนของโมเดลสองโมเดลกับข้อมูลเชิงประจักษ์ และจะใช้ได้ดีต่อเมื่อตัวแปรภายนอกและตัวแปรสังเกตได้เป็นตัวแปรมาตรฐาน (standardized variable) เพราะค่าของดัชนีแปลความหมายสัมพันธ์กับขนาดของความแปรปรวนและความแปรปรวนร่วม ระหว่างตัวแปรค่าของดัชนี RMR ยิ่งเข้าใกล้ศูนย์ แสดงว่าโมเดลมีความกลมกลืนกับข้อมูลเชิงประจักษ์ (นงลักษณ์ วิรัชชัย, 2538)

3.3.4 การวิเคราะห์เศษเหลือหรือความคลาดเคลื่อน (Analysis of Residuals)

ในการใช้นักวิจัยควรวิเคราะห์เศษเหลือควบคู่กันไปกับดัชนีตัวอื่นๆ ที่กล่าวมาแล้ว ผลจากการวิเคราะห์ข้อมูลที่เกี่ยวข้องกับความคลาดเคลื่อนมีหลายแบบ แต่ละแบบใช้ประโยชน์ในการตรวจสอบความกลมกลืนของโมเดลกับข้อมูลเชิงประจักษ์ดังนี้

3.3.4.1 เมทริกซ์ความคลาดเคลื่อนในการเทียบความกลมกลืน (Fitted Residuals Matrix)

เมทริกซ์ที่เป็นผลต่างของเมทริกซ์ S และ Sigma โมเดลจะให้ค่าความคลาดเคลื่อนทั้งในรูปคะแนนดิบ และคะแนนมาตรฐาน ค่าความคลาดเคลื่อนในรูปคะแนนมาตรฐาน คือ ผลหารระหว่างความคลาดเคลื่อนกับค่าความคลาดเคลื่อนมาตรฐานของความคลาดเคลื่อน ถ้าโมเดลมีความสอดคล้องกลมกลืนกับข้อมูล ค่าความคลาดเคลื่อนในรูปคะแนนมาตรฐานไม่ควรมีค่าเกิน 2.00 ถ้ายังมีค่าเกิน 2.00 ต้องปรับโมเดล (นงลักษณ์ วิรัชชัย, 2538)

3.3.4.2 คิวพล็อต (Q-Plot)

เป็นกราฟแสดงความสัมพันธ์ระหว่างค่าความคลาดเคลื่อนกับค่าควอนไทล์ปกติ (Normal Quantiles) ถ้าได้เส้นกราฟมีความชันมากกว่าเส้นทแยงมุมอันเป็นเกณฑ์ในการเปรียบเทียบ แสดงว่าโมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ (นงลักษณ์ วิรัชชัย, 2538)

3.3.5 ดัชนีตัดแปรโมเดล (Model Modification Indices)

ดัชนีตัวนี้เป็นประโยชน์มากในการปรับโมเดล ดัชนีตัดแปรโมเดลเป็นค่าสถิติเฉพาะสำหรับพารามิเตอร์แต่ละตัวมีค่าเท่ากับค่าไคสแควร์ ที่จะลดลงเมื่อกำหนดให้พารามิเตอร์ตัวนั้นเป็นพารามิเตอร์อิสระ หรือมีการผ่อนคลายข้อกำหนดเงื่อนไขบังคับของพารามิเตอร์นั้น ข้อมูลที่ได้นี้เป็นประโยชน์มากสำหรับนักวิจัยในการตัดสินใจปรับโมเดลให้ดีขึ้น (นงลักษณ์ วิรัชชัย, 2538)

Prince of Songkla University
Pattani Campus

ตอนที่ 4 งานวิจัยที่เกี่ยวข้องกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

ชวลิต นิवासวัต (2541) ได้ศึกษาการเปรียบเทียบผลของการวิเคราะห์ความลำเอียงของข้อสอบด้วยวิธีค่าอำนาจจำแนกของข้อสอบ วิธีแมนเทิล-แฮนส์เซล และวิธีโค้งลักษณะข้อสอบ 1 พารามิเตอร์ มีวัตถุประสงค์เพื่อค้นหาความลำเอียงของข้อสอบจากแบบทดสอบวิชาความรู้พื้นฐานทั่วไป ซึ่งเป็นแบบทดสอบที่ใช้ในการสอบคัดเลือกเข้าศึกษาต่อระดับปริญญาตรีของสถาบันราชภัฏฉะเชิงเทรา ประจำปีการศึกษา 2541 จำนวน 1,875 คน โดยวิธีสุ่มอย่างง่ายจากผู้เข้าสอบทั้งสิ้นจำนวน 3,067 คน ผู้วิจัยนำผลการสอบที่ได้มาค้นหาความลำเอียงโดยการวิเคราะห์ด้วยวิธีค่าอำนาจจำแนกของข้อสอบวิธีแมนเทิล-แฮนส์เซล และวิธีโค้งลักษณะข้อสอบ 1 พารามิเตอร์ ตามตัวแปรเพศและเขตที่ตั้งของสถานศึกษา หาค่าสัมประสิทธิ์สหสัมพันธ์ของดัชนีความลำเอียงที่ได้จากการวิเคราะห์โดยใช้สูตรของเพียร์สัน โพรดัก โมเมนต์ และเปรียบเทียบจำนวนข้อสอบที่ลำเอียงที่ได้จากการวิเคราะห์ ด้วย 3 วิธีนี้ ผลการวิจัยพบว่า 1. การวิเคราะห์ความลำเอียงของข้อสอบ วิธีแมนเทิล-แฮนส์เซลมีความไวในการตรวจพบข้อสอบที่ลำเอียงได้มากกว่าวิธีค่าอำนาจจำแนกของข้อสอบ และวิธีโค้งลักษณะข้อสอบ 1 พารามิเตอร์ เป็นจำนวนข้อสอบที่ลำเอียงตามเพศ 37 ข้อ และตามเขตที่ตั้งของสถานศึกษา 12 ข้อ วิธีค่าอำนาจจำแนกของข้อสอบ พบจำนวนข้อสอบที่ลำเอียงตามเพศ 11 ข้อ และตามเขตที่ตั้งของสถานศึกษา 1 ข้อ วิธีโค้งลักษณะข้อสอบ 1 พารามิเตอร์ พบจำนวนข้อสอบที่ลำเอียงตามเพศ 24 ข้อ และตามเขตที่ตั้งของสถานศึกษา 8 ข้อ 2. ความสัมพันธ์ของดัชนีความลำเอียงของข้อสอบ จากการวิเคราะห์ความลำเอียงด้วยวิธีค่าอำนาจจำแนกของข้อสอบ วิธีแมนเทิล-แฮนส์เซล และวิธีโค้งลักษณะ ข้อสอบ 1 พารามิเตอร์ ตามตัวแปรเพศและตัวแปรเขตที่ตั้งของสถานศึกษา พบว่า ค่าสัมประสิทธิ์สหสัมพันธ์ของดัชนีที่วิเคราะห์ความลำเอียงด้วยวิธีแมนเทิล-แฮนส์เซล และวิธีโค้งลักษณะข้อสอบ 1 พารามิเตอร์มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ 0.01 3. เปรียบเทียบจำนวนข้อสอบที่ลำเอียง จากการวิเคราะห์ความลำเอียงด้วยวิธีค่าอำนาจจำแนกของข้อสอบวิธีแมนเทิล-แฮนส์เซล และวิธีโค้งลักษณะข้อสอบ 1 พารามิเตอร์ เมื่อวิเคราะห์ตามตัวแปรเพศและเขตที่ตั้งของสถานศึกษา พบว่าวิธีการวิเคราะห์ความลำเอียงของข้อสอบทั้ง 3 วิธี พบจำนวนข้อสอบที่ลำเอียงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 โดยเฉพาะวิธีแมนเทิล-แฮนส์เซลพบจำนวนข้อสอบที่ลำเอียงมากกว่าวิธีค่าอำนาจจำแนกของข้อสอบอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

สมศักดิ์ จันผ่อง (2542) ได้ศึกษาการเปรียบเทียบผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีวิเคราะห์และขนาดกลุ่มผู้สอบต่างกัน มีวัตถุประสงค์เพื่อเปรียบเทียบผลการวิเคราะห์การทำหน้าที่ต่างกันของวิธี วิเคราะห์ 3 วิธี คือ วิธีโค้งลักษณะข้อสอบที่มี 3 พารามิเตอร์ วิธีแมนเทิล-แฮนส์เซล และวิธีชิปเทสท์ กับขนาดกลุ่มผู้สอบ 2 ขนาด คือ 600 และ 1,000 คน โดยตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 โรงเรียนสังกัดกรมสามัญศึกษา ในเขตจังหวัดพิษณุโลก ภาคเรียนที่ 1 ปีการศึกษา 2541 จำนวน 2,020 คน เป็นนักเรียนชาย 1,000 คน นักเรียนหญิง 1,020 คน ซึ่งได้มาโดยการสุ่มแบบหลายขั้นตอน (Multi Stage Random Sampling) เครื่องมือที่ใช้เป็นแบบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3 จากการสร้างของชาติชาย

กุกิตติเมตรี ซึ่งเป็นแบบเลือกตอบ 5 ตัวเลือก จำนวน 40 ข้อ ผู้วิจัยใช้แบบทดสอบทั้งฉบับซึ่งมีข้อสอบที่ลำเอียงทางเพศจำนวน 7 ข้อ และข้อสอบที่ไม่ลำเอียงทางเพศจำนวน 33 ข้อ ผลการวิจัยพบว่า 1.จำนวนข้อสอบที่ทำหน้าที่ต่างกันในวิธีวิเคราะห์ทั้ง 3 วิธี พบว่าวิธีชิบเทสท์ ตรวจพบ ข้อสอบที่ทำหน้าที่ต่างกันได้มากที่สุด รองลงมาคือวิธีแมนเทล-แฮนส์เซล และวิธีไค้่งลักษณะข้อสอบ 3 พารามิเตอร์ โดยจำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้ง 3 วิธีแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ เมื่อพิจารณาภายในวิธีเดียวกันแต่ต่างขนาดสอบ พบว่าวิธีชิบเทสท์ และวิธีแมนเทลแฮนส์เซล มีจำนวนข้อสอบที่ทำหน้าที่ต่างกันแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนวิธีไค้่งลักษณะข้อสอบ 3 พารามิเตอร์ มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ 2.ค่าความสอดคล้องของวิธีวิเคราะห์ทั้ง 3 วิธีกับข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันด้วยวิธีต่าง ๆ กลุ่มผู้สอบขนาด 600 คน วิธีแมนเทล-แฮนส์เซล มีค่าความสอดคล้องสูงสุดเท่ากับ .85 รองลงมาคือ วิธีชิบเทสท์เท่ากับ .83 และวิธีไค้่งลักษณะข้อ สอบ 3 พารามิเตอร์ เท่ากับ .63 กลุ่มผู้สอบขนาด 1,000 คน วิธีไค้่งลักษณะข้อสอบ 3 พารามิเตอร์ มีค่าความสอดคล้องสูงสุด (.78) รองลงมาคือ วิธีแมนเทล-แฮนส์เซล (.63) และวิธีชิบเทสท์ (.60) ตามลำดับ 3.ความสัมพันธ์ของดัชนีการทำหน้าที่ต่างกันด้วยวิธีวิเคราะห์ 3 วิธี เมื่อพิจารณาจากกลุ่มผู้สอบขนาด 600 คน พบว่าดัชนีการทำหน้าที่ต่างกันระหว่างวิธีแมนเทลแฮนส์เซลกับวิธีชิบเทสท์ มีความสัมพันธ์เท่ากับ .9249 ส่วนวิธีไค้่งลักษณะข้อสอบ 3 พารามิเตอร์กับวิธีชิบเทสท์มีความสัมพันธ์เท่ากับ .2545 สำหรับวิธีไค้่งลักษณะข้อสอบ 3 พารามิเตอร์กับวิธีแมนเทล-แฮนส์เซล มีความสัมพันธ์เท่ากับ .2329 เมื่อพิจารณาจากกลุ่มผู้สอบขนาด 1,000 คน พบว่า ดัชนีการทำหน้าที่ต่างกันระหว่างวิธีแมนเทล - แฮนส์เซล กับวิธีชิบเทสท์ มีความสัมพันธ์เท่ากับ .6090 ส่วนวิธีไค้่งลักษณะข้อสอบ 3 พารามิเตอร์กับวิธีแมนเทล-แฮนส์เซล และวิธีไค้่งลักษณะข้อสอบ 3 พารามิเตอร์กับวิธีชิบเทสท์ มีความสัมพันธ์เท่ากับ .1372 และ .1081 ตามลำดับ

ทองอยู่ สารระ (2543) ได้ศึกษาการเปรียบเทียบอำนาจการตรวจสอบและการจำแนกผิดพลาดในการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันแบบสม่าเสมอและแบบไม่สม่าเสมอระหว่างวิธีแมนเทล-แฮนส์เซลและวิธีการถดถอยโลจิสติก โดยใช้ความยาวแบบทดสอบและขนาดตัวอย่างต่างกัน มีความมุ่งหมายเพื่อเปรียบเทียบอำนาจการตรวจสอบและการจำแนกผิดพลาดในการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันแบบสม่าเสมอและแบบไม่สม่าเสมอ ระหว่าง วิธี MH และ LR โดยใช้ความยาวแบบทดสอบและขนาดตัวอย่างต่างกัน ตัวอย่าง ที่ใช้ในการศึกษาครั้งนี้เป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2542 ในโรงเรียนสังกัดกรมสามัญศึกษา ส่วนกลาง กลุ่มที่ 3 จำนวน 3,242 คน เครื่องมือที่ใช้เป็นแบบทดสอบความสามารถทางสมองที่ผู้วิจัยสร้างขึ้นตามแนว โครงสร้างของโอติส-เลนนอน ซึ่งเป็นแบบทดสอบเลือกตอบห้าตัวเลือก จำนวน 80 ข้อ วัดความสามารถทั่วไปสามด้านคือความเข้าใจด้านภาษา เหตุผลด้านภาษา และเหตุผล ด้านภาพ ในการศึกษาครั้งนี้ผู้วิจัยได้สุ่มแบบทดสอบความยาวสามขนาด คือ 20 ข้อ, 60 ข้อ และสุ่มตัวอย่างขนาดกลุ่มอ้างอิงต่อกลุ่มเปรียบเทียบ จำนวน 10 กลุ่ม คือ 300: 100, 300: 200, 300: 300, 500: 100, 500: 200, 500: 300, 1000: 100, 1000: 200, 1000: 300 และ 1000: 1000 รวมเงื่อนไขที่ศึกษาทั้งหมด 30 เงื่อนไข (3 ขนาดความยาวแบบทดสอบ x 10 ขนาดตัวอย่าง)

ผลการศึกษา พบว่า 1. อำนาจการตรวจสอบและการจำแนกผิดพลาดในการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกัน ทั้งแบบสมำเสมอและแบบไม่สมำเสมอ ระหว่างวิธี MH และวิธี LR ภายใต้ความยาวแบบทดสอบและขนาดตัวอย่างที่ศึกษา เกือบทุกเงื่อนไข มีค่าไม่แตกต่างกัน 2. ความยาวของแบบทดสอบไม่มีผลต่ออำนาจการตรวจสอบและการจำแนกผิดพลาดในการตรวจสอบด้วยวิธี MH และวิธี LR ทั้งการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันแบบ สมำเสมอและแบบไม่สมำเสมอ 3. ขนาดของตัวอย่างมีผลต่ออำนาจการตรวจสอบในการตรวจสอบด้วยวิธี MH และวิธี LR เกือบทุกเงื่อนไขของการศึกษา ทั้งการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกัน แบบสมำเสมอและแบบไม่สมำเสมอ กล่าวคือเมื่อขนาดตัวอย่างเพิ่มขึ้น อำนาจ การตรวจสอบ จะมีค่าเพิ่มขึ้น แต่พบว่าขนาดตัวอย่างไม่มีผลต่อการจำแนกผิดพลาด ในเกือบทุกเงื่อนไขที่ศึกษา ทั้งการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกัน แบบสมำเสมอ และแบบไม่สมำเสมอ

นุชริน ไบโพธิ์ (2544) ได้ศึกษาการเปรียบเทียบความสอดคล้องของผลการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันระหว่างวิธี Lord's χ^2 วิธี Mantel – Haenszel และวิธี SIBTEST มีจุดมุ่งหมายเพื่อเปรียบเทียบจำนวนข้อสอบที่ทำหน้าที่ต่างกัน และศึกษาความสอดคล้องของผลการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันระหว่างวิธี Lord's χ^2 วิธี Mantel – Haenszel และวิธี SIBTEST ตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2543 ในเขตกรุงเทพมหานคร จำนวน 790 คน ได้จากการสุ่มแบบสองขั้นตอน เครื่องมือที่ใช้ในการศึกษา เป็นแบบทดสอบวัดความถนัดทางคณิตศาสตร์ จำนวน 60 ข้อ ผลการวิจัยสรุปได้ดังนี้ 1. ข้อสอบที่ทำหน้าที่ต่างกันที่ได้จากการตรวจสอบระหว่างวิธี Lord's χ^2 วิธี Mantel – Haenszel และวิธี SIBTEST มีจำนวนข้อที่ตรวจพบแตกต่างกัน เมื่อขนาดตัวอย่างและจำนวนข้อสอบต่างกัน ได้แก่ กรณีแบบทดสอบ 30 ข้อ ที่ใช้ตัวอย่างขนาด 300 คน และ 500 คน และแบบทดสอบ 40 ข้อ ที่ใช้ตัวอย่างขนาด 500 คนและ 700 คน 2. ความสอดคล้องของผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบเมื่อพิจารณาจากข้อสอบ จำนวน 20 ข้อและ 30 ข้อ ที่ตรวจสอบโดยวิธี Lord's χ^2 วิธี Mantel – Haenszel และวิธี SIBTEST พบว่าตัวอย่างขนาด 500 คนและ 700 คน มีจำนวนข้อสอบที่ตรวจพบสอดคล้องกันสูงสุด ส่วนข้อสอบจำนวน 40 ข้อ ซึ่งตรวจสอบด้วยวิธี Lord's χ^2 พบว่าตัวอย่างขนาด 300 คน และ 500 คน มีจำนวนข้อสอบที่ตรวจพบสอดคล้องกันสูงสุด แต่เมื่อตรวจสอบโดยวิธี Mantel – Haenszel และวิธี SIBTEST พบว่าตัวอย่างขนาด 500 คนและ 700 คน มีจำนวนข้อสอบที่ตรวจพบสอดคล้องกันสูงสุด

วรรณ รอดตัว (2544) ได้ศึกษาการเปรียบเทียบผลของการวิเคราะห์ความลำเอียงของข้อสอบที่แตกต่างกัน 3 วิธี มีวัตถุประสงค์เพื่อค้นหาความลำเอียงของแบบทดสอบวิชาคณิตศาสตร์ 1 ซึ่งเป็นแบบทดสอบที่ใช้ในการสอบคัดเลือกเข้าศึกษาต่อระดับปริญญาตรี ในสถาบันอุดมศึกษาของรัฐ และเอกชน ที่สังกัดทบวงมหาวิทยาลัย ประจำปีการศึกษา 2542 โดยวิธีวิเคราะห์ความลำเอียงของข้อสอบด้วยวิธีค่าอำนาจจำแนกของข้อ สอบ วิธีแมนเทล-แฮนส์เซล (Mantel-Haenszel) และวิธีซิปเทสท์ (SIBTEST) เปรียบเทียบผลการวิเคราะห์ความลำเอียง และ

ศึกษาความสัมพันธ์ของวิธีวิเคราะห์ความลำเอียง ทั้งสามวิธี ตัวอย่างที่ใช้เป็นผู้เข้าสอบคัดเลือกเพื่อเข้าศึกษาต่อระดับปริญญาตรี ใน สถาบันอุดมศึกษาของรัฐ และเอกชน ประจำปีการศึกษา 2542 จำนวน 2,540 คน โดยใช้วิธีสุ่มอย่างง่ายมาจากผู้เข้าสอบคัดเลือกทั้งสิ้นจำนวน 99,562 คน ผู้วิจัยนำผลการ สอบที่ได้มาค้นหาความลำเอียงด้วยวิธีค่าอำนาจจำแนกของข้อสอบ วิธีแมนเทิล-แฮนส์ เซล และวิธีชิปเทสต์ ตามตัวแปรเพศและเขตที่ตั้งของสถานศึกษา หาค่าสัมประสิทธิ์สหสัมพันธ์ของดัชนีความลำเอียงที่ได้จากการวิเคราะห์ โดยใช้สูตรเพียร์สัน โพรดักโมเมนต์ และเปรียบเทียบจำนวนข้อสอบที่ลำเอียงที่ได้จากการวิเคราะห์ด้วย 3 วิธีนี้ ผลการวิจัยพบว่า 1. ผลการวิเคราะห์ความลำเอียงของข้อสอบรายข้อด้วยวิธีวิเคราะห์ 3 วิธี ระหว่างตัวแปรเพศและเขตที่ตั้งของสถานศึกษา จากข้อสอบทั้งหมด 28 ข้อ พบว่า เขตที่ตั้งของสถานศึกษาพบจำนวนข้อสอบที่ลำเอียงสูงกว่าเพศ ดังนี้ 1) วิธีค่าอำนาจจำแนกของข้อสอบพบข้อสอบที่ลำเอียงตามเพศ จำนวน 3 ข้อ ลำเอียงตามเขตที่ตั้งของสถานศึกษา จำนวน 12 ข้อ 2) วิธีแมนเทิล-แฮนส์เซล พบข้อสอบที่ลำเอียงตามเพศ จำนวน 3 ข้อ ลำเอียงตามเขตที่ตั้งของสถานศึกษา จำนวน 8 ข้อ และ 3) วิธีชิปเทสต์พบข้อสอบที่ ลำเอียงตามเพศ จำนวน 4 ข้อ ลำเอียงตามเขตที่ตั้งของสถานศึกษา จำนวน 6 ข้อ 2. ค่าสหสัมพันธ์ของดัชนีความลำเอียงของข้อสอบ จากการวิเคราะห์ด้วยวิธีค่าอำนาจจำแนกของข้อสอบ วิธีแมนเทิล-แฮนส์ เซล และวิธีชิปเทสต์ ระหว่างตัวแปรเพศและเขตที่ตั้งของสถานศึกษา พบว่า วิธีแมนเทิล-แฮนส์เซล และวิธีชิปเทสต์ มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในตัวแปรเพศและมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ในตัวแปรเขตที่ตั้งของสถานศึกษา 3. เปรียบเทียบจำนวนข้อสอบที่ลำเอียงจากการใช้วิธีวิเคราะห์ความลำเอียงของ ข้อสอบด้วยวิธีค่าอำนาจจำแนกของข้อสอบ วิธีแมนเทิล-แฮนส์เซล และวิธีชิปเทสต์ ตามตัวแปรเพศและเขตที่ตั้งของสถานศึกษา พบว่า วิธีวิเคราะห์ความลำเอียงของข้อ สอบทั้ง 3 วิธี พบจำนวนข้อสอบที่ลำเอียงไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ ระดับ .05

สุมาลี แก้วทองค์ (2547) ได้ศึกษาสาเหตุของการทำหน้าที่ต่างกันของข้อสอบสาระการเรียนรู้ภาษาไทยและสาระการเรียนรู้สังคมศึกษา ศาสนาและวัฒนธรรม การวิจัยครั้งนี้มีวัตถุประสงค์ 2 ประการ คือ ประการแรก เพื่อศึกษาการทำหน้าที่ต่างกันของข้อสอบด้านเพศ ภาษาพูด และเชื้อชาติ ของแบบสอบสาระการเรียนรู้ภาษาไทย และสาระการเรียนรู้สังคมศึกษา ศาสนา และวัฒนธรรม และประการที่สองเพื่อตรวจสอบสาเหตุของการทำหน้าที่ต่างกันของข้อสอบด้านเพศ ภาษาพูด และเชื้อชาติของแบบสอบกลุ่มสาระการเรียนรู้ภาษาไทย และกลุ่มสาระการเรียนรู้สังคมศึกษา ศาสนาและวัฒนธรรม ตัวอย่างที่ใช้ในการวิจัย คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2546 สำนักงานเขตพื้นที่การศึกษาสงขลา พัทลุง ตรัง และสตูล จำนวน 1,320 คน ซึ่งได้จากการสุ่มตัวอย่างแบบแบ่งชั้น 3 ชั้นตอน เครื่องมือที่ใช้ในการวิจัยประกอบด้วยแบบสอบถามเพื่อสำรวจความสนใจในการอ่านของผู้เรียน แบบสอบในกลุ่มสาระการเรียนรู้ภาษาไทยและกลุ่มสาระการเรียนรู้สังคมศึกษา ศาสนา และวัฒนธรรม วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยวิธีแมนเทิล-แฮนส์เซล ด้วยโปรแกรม SIBTEST และตรวจสอบสาเหตุการทำหน้าที่ต่างกันโดยผู้เชี่ยวชาญ ผลการวิจัยพบว่า 1. แบบสอบกลุ่มสาระการเรียนรู้ภาษาไทยมี 3 ฉบับ ประกอบด้วยข้อสอบจำนวน 120 ข้อ มีข้อสอบที่ทำหน้าที่ต่างกันด้านเพศ 9 ข้อ ภาษาพูด 15 ข้อ และเชื้อชาติ 28 ข้อ 2. ข้อสอบที่

ทำหน้าที่ต่างกันด้านเพศของแบบทดสอบสาระการเรียนรู้ภาษาไทย ส่วนใหญ่มีสาเหตุมาจากการเนื้อเรื่องที่น่าสนใจและภาษาที่ใช้ในแบบสอบ โดยมีลักษณะของข้อสอบที่ออกตามความสนใจและมีการใช้ภาษาในแบบสอบที่เป็นสำนวนและคำศัพท์ และข้อสอบที่ทำหน้าที่ต่างกันด้านเพศของแบบสอบสาระการเรียนรู้สังคมศึกษา ศาสนา และวัฒนธรรม มีสาเหตุมาจากเนื้อเรื่องที่น่าสนใจและเนื้อเรื่องเกี่ยวกับวัฒนธรรม โดยมีลักษณะของข้อสอบที่มีเนื้อเรื่องที่น่าสนใจและมีเนื้อเรื่องเกี่ยวกับวัฒนธรรมของผู้สอบ

3. ข้อสอบที่ทำหน้าที่ต่างกันด้านภาษาพูดของแบบสอบสาระการเรียนรู้ภาษาไทย ส่วนใหญ่มีสาเหตุมาจาก เนื้อเรื่องที่น่าสนใจ และภาษาที่ใช้ในแบบสอบ โดยมีลักษณะข้อสอบที่ออกตามความสนใจและมีการใช้ภาษาที่เป็นคำศัพท์เฉพาะ เช่น ราชาศัพท์ และข้อสอบที่ทำหน้าที่ต่างกันด้านภาษาพูดของแบบสอบสาระการเรียนรู้สังคมศึกษา ศาสนา และวัฒนธรรม มีสาเหตุมาจากเนื้อเรื่องที่น่าสนใจและเนื้อเรื่องเกี่ยวกับวัฒนธรรม โดยมีลักษณะของข้อสอบที่ออกตามความสนใจเช่น ภูมิศาสตร์ และในเรื่องเกี่ยวกับวัฒนธรรมไทย

4. ข้อสอบที่ทำหน้าที่ต่างกันในด้านเชื้อชาติของแบบสอบสาระการเรียนรู้ภาษาไทยส่วนใหญ่มีสาเหตุมาจากเนื้อเรื่องที่น่าสนใจ และภาษาที่ใช้ในแบบสอบ โดยมีลักษณะข้อสอบที่ออกตามความสนใจและมีการใช้คำราชาศัพท์หรือคำศัพท์ยากมาออกข้อสอบ และข้อสอบที่ทำหน้าที่ต่างกันในด้านเชื้อชาติของแบบสอบสาระการเรียนรู้สังคมศึกษา ศาสนา และวัฒนธรรมมีสาเหตุมาจากเนื้อเรื่องที่น่าสนใจและเนื้อเรื่องเกี่ยวกับวัฒนธรรมและประเพณี

ศุภวัฒน์ มะลิเผือก (2548) ได้ศึกษาอิทธิพลของการทำหน้าที่ต่างกันของข้อสอบ ที่ส่งผลต่อคุณภาพของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนระดับชาติ ชั้นประถมศึกษาปีที่ 6 ซึ่งการวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาอิทธิพลของการทำหน้าที่ต่างกันของข้อสอบที่ส่งผลต่อคุณภาพแบบทดสอบ โดยเปรียบเทียบค่าความเที่ยงความตรงเชิงโครงสร้าง ความคงที่ของโครงสร้าง องค์ประกอบ และค่าสัมประสิทธิ์สหสัมพันธ์อันดับของผู้สอบระหว่างแบบทดสอบ ฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนระดับชาติ ชั้นประถมศึกษาปีที่ 6 วิชาภาษาไทย ปีการศึกษา 2546 ตัวอย่างเป็นนักเรียนสังกัดสำนักงานเขตพื้นที่ การศึกษานครศรีธรรมราช จำนวน 2,000 คน วิเคราะห์ค่าสถิติพื้นฐานโดยใช้โปรแกรม SPSS วิเคราะห์ค่าสถิติของข้อสอบและค่าความเที่ยงโดยใช้โปรแกรม LERTAP 5 ตรวจสอบความตรงเชิงโครงสร้างด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โดยใช้โปรแกรม LISREL 8.50 และตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้โปรแกรม SIBTEST และ SPSS ผลการวิจัยปรากฏว่า 1. พบข้อสอบทำหน้าที่ต่างกันแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนระดับชาติภาษาไทย เมื่อจำแนกกลุ่มผู้สอบตามตัวแปรเพศ จำนวน 12 ข้อ คิดเป็นร้อยละ 30 2. แบบทดสอบฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ต่างกัน มีค่าความความเที่ยงแตกต่างกัน 3. แบบทดสอบฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ต่างกัน มีความตรงเชิงโครงสร้างไม่แตกต่างกัน 4. แบบทดสอบฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ต่างกัน มีโครงสร้างองค์ประกอบคงที่ 5. ค่าสัมประสิทธิ์สหสัมพันธ์อันดับของผู้สอบระหว่างแบบทดสอบฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ต่างกัน มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ทิพย์วรรณ คงเนตร (2549) ได้ศึกษาการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบวิชาทักษะภาษาไทย สำหรับการวัดผลปลายภาคการศึกษาของมหาวิทยาลัยนเรศวร ปีการศึกษา 2546 มีจุดมุ่งหมายเพื่อศึกษาการทำหน้าที่ต่างกันของข้อสอบวิชาทักษะภาษาไทย สำหรับวัดผลปลายภาคการศึกษาของมหาวิทยาลัยนเรศวร ปีการศึกษา 2546 โดยวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบที่มี 3 พารามิเตอร์ ด้วยวิธีวัดความแตกต่างของ ค่าพารามิเตอร์ความยาก โดยศึกษาการทำหน้าที่ต่างกัน (Differential Item Functioning: DIF) ที่มีต่อตัวแปรเพศ กลุ่มสาขาและวิธีรับเข้าศึกษาในมหาวิทยาลัยนเรศวร ตัวอย่างที่ใช้เป็นนิสิตมหาวิทยาลัยนเรศวรชั้นปีที่ 1 ภาคการศึกษาที่ 2 ปีการศึกษา 2546 จำนวน 1,356 คน ผลการวิจัยพบว่า 1. เมื่อจำแนกผู้สอบตามเพศ มีข้อสอบที่ DIF จำนวน 71 ข้อ คิดเป็นร้อยละ 59.17 เป็นข้อที่เข้าข้างเพศหญิงจำนวน 38 ข้อ และเข้าข้างเพศชายจำนวน 33 ข้อ คิดเป็น ร้อยละ 31.67 และร้อยละ 27.50 ตามลำดับ 2.เมื่อจำแนกผู้สอบตามกลุ่มสาขามีข้อสอบที่ DIF จำนวน 69 ข้อ คิดเป็น ร้อยละ 57.50 เป็นข้อที่เข้าข้างกลุ่มสาขาวิทยาศาสตร์จำนวน 49 ข้อ และเข้าข้างกลุ่มสาขา มนุษยศาสตร์ จำนวน 20 ข้อ คิดเป็นร้อยละ 40.83 และร้อยละ 16.67 ตามลำดับ 3.เมื่อจำแนกผู้สอบตามวิธีรับเข้าศึกษา มีข้อสอบที่ DIF จำนวน 65 ข้อ คิดเป็น ร้อยละ 54.17 เป็นข้อที่เข้าข้างวิธีรับเข้าจากส่วนกลาง จำนวน 33 ข้อ และเข้าข้างวิธีรับเข้าโดยตรง จำนวน 32 ข้อ คิดเป็นร้อยละ 27.59 และร้อยละ 26.67 ตามลำดับ

ปิยะทิพย์ ดินวร (2549) ได้ศึกษาการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ: การเปรียบเทียบประสิทธิภาพระหว่างวิธีการวิเคราะห์องค์ประกอบจำกัดกับวิธีถดถอยโลจิสติก มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ ระหว่างวิธีการวิเคราะห์องค์ประกอบจำกัดกับวิธีถดถอยภายใต้เงื่อนไข 18 เงื่อนไข (3x2x2) คือ ขนาดของตัวอย่าง 3 ขนาด (2,000 คน 1,000 คน และ 300 คน) ความยาวของแบบทดสอบ 3 ขนาด (40 ข้อ 30 ข้อ และ 20 ข้อ) และเกณฑ์การจับคู่ 2 เกณฑ์ (คะแนนรวมทั้งฉบับ = TS และ คะแนนแบบทดสอบย่อย = SS) เมื่อใช้ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ โดยวิธีซิปเทสท์เป็นเกณฑ์สำหรับการเปรียบเทียบประสิทธิภาพ ตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน ปีการศึกษา 2546 ที่เข้าสอบวัดผลสัมฤทธิ์ทางการเรียนระดับชาติ วิชาภาษาไทย จำนวน 2,000 คน วิเคราะห์ค่าสถิติพื้นฐาน โดยใช้โปรแกรม SPSS ตรวจสอบความตรงเชิงโครงสร้างด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โดยใช้โปรแกรม LISREL 8.50 และตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ โดยใช้โปรแกรม SIBTEST และ SPSS ผลการวิจัยปรากฏว่า 1.วิธีถดถอยโลจิสติก (LR) มีประสิทธิภาพไม่แตกต่างกันกับวิธีวิเคราะห์องค์ประกอบจำกัด ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติจาก 2 เงื่อนไข (ขนาดตัวอย่าง X ความยาวแบบทดสอบ X เกณฑ์การจับคู่) คือ 1,000 คน X 20 ข้อ X SS และ 300 คน X 20 ข้อ X SS 2.วิธี LR มีประสิทธิภาพมากกว่าวิธี RFA ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ในแบบทดสอบพหุมิติ 16 เงื่อนไข คือ 1) 2,000 คน X 40 ข้อ X TS 2) 2,000 คน X 40 ข้อ X SS 3) 2,000 คน X 30 ข้อ X TS 4) 2,000 คน X 30 ข้อ X SS 5) 2,000 คน X 20 ข้อ X TS 6) 2,000 คน X 20 ข้อ X SS 7) 1,000 คน X 40 ข้อ X TS 8) 1,000 คน X 40 ข้อ X SS 9) 1,000 คน X 30 ข้อ X TS 10) 1,000 คน X 30 ข้อ X SS 11) 1,000 คน X 20 ข้อ X TS 12) 300 คน X 40 ข้อ

XTS 13)300 คนX40 ข้อxSS 14) 300 คนX30 ข้อxTS 15) 300 คนX30 ข้อXรร และ 16) 300 คน X20 ข้อXTร เนื่องจาก วิธี LR มีอัตราความคลาดเคลื่อนประเภทที่1 น้อยกว่าวิธี RFA อย่างมีนัยสำคัญทางสถิติที่ระดับ.05

พริญา สูงเนินและคณะ (2549) ได้ศึกษาการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ:การเปรียบเทียบระหว่างรายข้อกับรายหมวดข้อสอบโดยใช้วิธีชิปเทสท์ มีวัตถุประสงค์เพื่อเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ ระหว่างการตรวจสอบเป็นรายข้อกับรายหมวดข้อสอบ โดยใช้วิธีชิปเทสท์ ภายใต้เงื่อนไขขนาดตัวอย่างต่างกัน คือ ขนาดเล็ก (300 คน) ขนาดกลาง (1,000 คน) และขนาดใหญ่ (2,000 คน) ตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 สังกัดเขตพื้นที่การศึกษานครศรีธรรมราช ปีการศึกษา 2546 ที่เข้าสอบวัดผลสัมฤทธิ์ทางการเรียนระดับชาติ จำนวน 2,000 คน ข้อมูลทฤษฎีที่ใช้เป็นคะแนนจากแบบทดสอบวิชาภาษาไทย ชั้นประถมศึกษาปีที่ 6 จำนวน 40 ข้อ จำแนกเป็น 2 หมวดข้อสอบ คือ หมวดที่ 1 วัดด้านโครงสร้างความรู้ จำนวน 15 ข้อ และหมวดที่ 2 วัดด้าน กระบวนการ จำนวน 25 ข้อ วิเคราะห์ค่าสถิติพื้นฐานโดยใช้โปรแกรม SPSS ตรวจสอบความตรงเชิงโครงสร้างด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โดยใช้โปรแกรม LISREL 8.50 และตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ โดยใช้โปรแกรม SIBTEST ผลการวิจัยปรากฏว่า 1. ภายใต้เงื่อนไขขนาดตัวอย่างต่างกัน คือ ขนาดเล็ก ขนาดกลาง และขนาดใหญ่ ระหว่างการตรวจสอบการทำหน้าที่ต่างกันเป็นรายข้อกับรายหมวดข้อสอบ พบข้อสอบทำหน้าที่ต่างกันแตกต่างกัน 2. การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบเป็นรายข้อ เมื่อตัวอย่างขนาดเล็ก พบข้อสอบทำ หน้าที่ต่างกันจำนวน 4 ข้อ คิดเป็นร้อยละ 10 ขนาดกลางพบจำนวน 13 ข้อ คิดเป็นร้อยละ 32.5 และขนาดใหญ่พบ จำนวน 15 ข้อ คิดเป็นร้อยละ 37.5 3. การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบเป็นรายหมวดข้อสอบ เมื่อตัวอย่างขนาดเล็ก พบ ข้อสอบทำหน้าที่ต่างกันจำนวน 4 ข้อ คิดเป็นร้อยละ 10 ขนาดกลางพบ จำนวน 8 ข้อ คิดเป็นร้อยละ 20 และขนาดใหญ่ พบจำนวน 16 ข้อ คิดเป็นร้อยละ 40 4. การตรวจสอบการทำหน้าที่ต่างกันของหมวดข้อสอบ ชี้ให้เห็นว่า หมวดที่ 2 ภายใต้เงื่อนไขตัวอย่าง ขนาดกลาง มีนัยสำคัญทางสถิติที่ระดับ .05

อิทธิฤทธิ์ พงษ์ปิยะรัตน์ (2551) ได้ศึกษาการวิเคราะห์ข้อสอบและการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ: การวิเคราะห์พหุระดับ มีวัตถุประสงค์ 3 ประการ ได้แก่ 1.เพื่อวิเคราะห์ค่าพารามิเตอร์ข้อสอบ พารามิเตอร์ความสามารถของผู้สอบ และผลของตัวแปรคุณลักษณะของผู้เรียนและตัวแปรคุณลักษณะของโรงเรียนต่อโอกาสในการตอบข้อสอบได้ถูกต้อง โดยใช้วิธีการวิเคราะห์พหุระดับ ด้วยโปรแกรมโมเดลเชิงเส้นตรงระดับลดหลั่น (HLM) 2.เพื่อเปรียบเทียบผลการประมาณค่าพารามิเตอร์ข้อสอบ และพารามิเตอร์ความสามารถของผู้สอบ ระหว่างการวิเคราะห์ข้อสอบแบบพหุระดับโดยประยุกต์ใช้โปรแกรมโมเดลเชิงเส้นตรงระดับลดหลั่น (HLM) และการวิเคราะห์ข้อสอบด้วยโปรแกรม BILOG-MG และ 3.เพื่อตรวจสอบและเปรียบเทียบผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) โดยประยุกต์ใช้โปรแกรมโมเดลเชิงเส้นตรงระดับลดหลั่น (HLM) และโปรแกรม BILOG-MG ตัวอย่างที่ใช้ในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษา

ปีที่ 3 ที่ศึกษาในปีการศึกษา 2550 จำนวน 1,588 คน ผู้บริหาร 32 คน จาก 32 โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาลพบุรีเขต 1 และ 2 ซึ่งได้มาโดยการสุ่มตัวอย่างแบบสองขั้นตอน เครื่องมือที่ใช้ในการวิจัยประกอบด้วย แบบวัดความรู้วิชาคณิตศาสตร์ แบบวัดเจตคติต่อการเรียนวิชาคณิตศาสตร์ แบบประเมินความเครียดของนักเรียน แบบบันทึกข้อมูลนักเรียน แบบประเมินภาวะผู้นำทางวิชาการของผู้บริหาร และแบบบันทึกตัวแปรคุณลักษณะครูและโรงเรียน การวิเคราะห์ข้อมูลดำเนินการใน 3 ขั้นตอนดังนี้ ขั้นตอนที่ 1 ประเมินค่าพารามิเตอร์ข้อสอบ และผู้สอบด้วยโมเดล HGLM-2L และ HGLM-3L ข้อมูลแบ่งออกเป็น 3 ระดับ คือ ระดับข้อสอบ ระดับผู้สอบ และระดับโรงเรียน แล้วเปรียบเทียบผลการประมาณค่าพารามิเตอร์ข้อสอบ และผู้สอบที่ได้จากโปรแกรม HLM และ BILOG-MG ขั้นตอนที่ 2 ศึกษาผลของตัวแปรคุณลักษณะผู้เรียนและโรงเรียนที่มีต่อโอกาสในการตอบข้อสอบได้ถูกต้องทั้งในระดับผู้สอบและระดับโรงเรียน และขั้นที่ 3 เปรียบเทียบผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) สรุปผลการวิจัยได้ดังต่อไปนี้ 1. ผลการประมาณค่าพารามิเตอร์ข้อสอบด้วยโมเดล HGLM-2L และ HGLM-3L ด้วยสถิติ Empirical Bayesian มีความสัมพันธ์อย่างสมบูรณ์กับผลการประมาณค่าด้วยโปรแกรม BILOG-MG ส่วนผลการประมาณค่าพารามิเตอร์ผู้สอบด้วยโมเดล HGLM-2L มีความสัมพันธ์อย่างสมบูรณ์กับผลการประมาณค่าด้วยโปรแกรม BILOG-MG ส่วนโมเดล HGLM-3L มีระดับของค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.793 ซึ่งน้อยกว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างโมเดล HGLM-2L กับโปรแกรม BILOG-MG 2. ผลการวิเคราะห์ระดับนักเรียน (Level 2) พบว่า ตัวแปรผลการเรียนวิชาคณิตศาสตร์ในภาคเรียนที่ผ่านมาส่งผลต่อค่าเฉลี่ยของโอกาสในการตอบข้อสอบได้ถูกต้องในแต่ละโรงเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ผลการวิเคราะห์ระดับโรงเรียน (Level 3) พบว่าตัวแปรขนาดของโรงเรียน และตัวแปรความเป็นผู้นำทางวิชาการของผู้บริหารส่งผลต่อค่าเฉลี่ยของโอกาสในการตอบข้อสอบได้ถูกต้องในโรงเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 และ 0.05 ตามลำดับ 3. การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยโมเดล HGLM สามารถตรวจสอบพบข้อสอบที่ทำหน้าที่ต่างกันสอดคล้องตรงกับผลการวิเคราะห์จากโปรแกรม BILOG-M

สุภาภรณ์ แดงเพ็ง (2552) ได้ศึกษาการเปรียบเทียบความลำเอียงของแบบทดสอบคณิตศาสตร์ในการประเมินคุณภาพการศึกษาระดับท้องถิ่น ในชั้นประถมศึกษาปีที่ 5 สังกัดสำนักงานเขตพื้นที่การศึกษาปัตตานี เขต 2 ระหว่างวิธีแมนเทิล-แฮนส์เซลและโค้งลักษณะข้อสอบ 3 พารามิเตอร์ มีวัตถุประสงค์เพื่อ 1) เปรียบเทียบจำนวนข้อของแบบทดสอบที่ตรวจพบความลำเอียงจากการวิเคราะห์ระหว่างวิธีแมนเทิล -แฮนส์เซลและวิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ 2) ศึกษาความสอดคล้องของดัชนีความลำเอียงจากการวิเคราะห์ระหว่างวิธีแมนเทิล แฮนส์เซลและวิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ และ 3) เปรียบเทียบความเชื่อมั่นของ แบบทดสอบหลังจากคัดเลือกข้อสอบข้อที่มีความลำเอียงออก ระหว่างวิธีแมนเทิล -แฮนส์เซลและ วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ ตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 5 ปีการศึกษา 2550 สังกัดสำนักงานเขตพื้นที่การศึกษาปัตตานี เขต 2 จำนวน 1,122 คน ได้มาโดยการสุ่มแบบ แบ่งชั้นภูมิ (Stratified Random Sampling) โดยภาษาที่ใช้ในชีวิตประจำวันเป็นชั้น (Strata) และ โรงเรียนเป็นหน่วยการสุ่ม

(Sampling Unit) เครื่องมือที่ใช้เป็นแบบทดสอบกลุ่มสาระการเรียนรู้ คณิตศาสตร์ ชนิดเลือกตอบ 4 ตัวเลือก จำนวน 40 ข้อสร้างโดยคณะกรรมการสร้างข้อสอบสังกัด สำนักงานเขตพื้นที่การศึกษา ปัตตานี เขต 2 ผลการวิจัยพบว่า 1. วิธีแมนเทิล-แฮนส์เซลตรวจพบจำนวนข้อสอบที่มีความลำเอียง แตกต่างจากวิธีไค้กึ่งลักษณะข้อสอบ 3 พารามิเตอร์ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 2. ดัชนี ความลำเอียงจากการวิเคราะห์ด้วยวิธีแมนเทิล-แฮนส์เซลและวิธีไค้กึ่งลักษณะ ข้อสอบ 3 พารามิเตอร์ มีความสอดคล้องกันสูง อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 3. ค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกจากการวิเคราะห์ด้วยวิธีแมนเทิล-แฮนส์เซลและวิธีไค้กึ่ง ลักษณะข้อสอบ 3 พารามิเตอร์ไม่แตกต่างกัน

เรื่องเดช ศิริกิจ (2554) ได้ศึกษาการวิเคราะห์เปรียบเทียบโมเดลการประเมิน

คุณภาพการจัดการศึกษาวิชาคณิตศาสตร์: การประยุกต์ใช้โมเดลมูลค่าเพิ่มที่มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและการวิเคราะห์การทำหน้าที่ต่างกันของตัวลอง มีวัตถุประสงค์เพื่อ 1. เพื่อศึกษาผลของข้อสอบที่เกิดการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลอง (DDF) ในแบบสอบประเมินความรู้วิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 2 2. เพื่อศึกษาผลของข้อสอบที่ทำหน้าที่ต่างกัน และผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน 4 โมเดล และ 3. เพื่อเปรียบเทียบประสิทธิภาพโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษา 4 โมเดล ข้อมูลที่ใช้ในการวิจัยครั้งนี้เป็นข้อมูลทุติยภูมิจากโครงการการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์หรือ TIMSS ปี 2007 ตัวอย่างของโครงการประกอบด้วย นักเรียน ครูคณิตศาสตร์ ที่สอนนักเรียนที่เป็นตัวอย่างในการวิจัย และผู้บริหารสถานศึกษาที่นักเรียนตัวอย่างศึกษา จำนวน 150 สถานศึกษา และจำนวนนักเรียน 5,412 คน โดยศึกษาเฉพาะวิชาคณิตศาสตร์ จากแบบสอบประเมินความรู้วิชาคณิตศาสตร์ จำนวน 14 ฉบับ การวิเคราะห์ข้อมูลดำเนินการใน 3 ขั้นตอน ดังนี้ ขั้นตอนที่ 1 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบและการทำหน้าที่ต่างกันของตัวลอง ด้วยโปรแกรม DDFS 1.0 ขั้นตอนที่ 2 การวิเคราะห์มูลค่าเพิ่มในโมเดลการประเมินคุณภาพการจัดการศึกษา โดยประยุกต์ใช้โมเดลมูลค่าเพิ่ม (Value-Added Model) ด้วยโมเดลเชิงเส้นตรงแบบลดหลั่น (HLM) ด้วยการวิเคราะห์ 2 ระดับ และขั้นตอนที่ 3 การตรวจสอบผลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา ประกอบด้วย 4 โมเดลที่ต่างกัน คือ โมเดล 1 “Undetected DIF-DDF & Adjusted” โมเดล 2 “Detect DIF & Adjusted” โมเดลที่ 3 “Detect DIF-DDF & Unadjusted” โมเดล 4 “Detect DIF-DDF & Adjusted” ผลการวิจัยพบว่า 1. ตัวแปรเพศเป็นคุณลักษณะที่พบการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลองในแบบสอบวิชาคณิตศาสตร์ มากที่สุด โดยพบว่าข้อสอบส่วนใหญ่ที่เกิดการทำหน้าที่ต่างกัน (DIF) มีจำนวนข้อสอบที่เพศหญิงมีโอกาสที่จะตอบถูกมากกว่าเพศชาย ในขณะที่ข้อสอบที่เกิดการทำหน้าที่ต่างกันของตัวลอง (DDF) ข้อสอบส่วนใหญ่เพศชายมีโอกาสที่จะเลือกตอบมากกว่าเพศหญิง 2. ผลของข้อสอบที่ทำหน้าที่ต่างกันแบบสอบวิชาคณิตศาสตร์ต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ พบว่า ข้อสอบที่ทำหน้าที่ต่างกันแบบสอบวิชา

คณิตศาสตร์ไม่มีผลต่อการประเมินคุณภาพการจัดการศึกษา (จากการเปรียบเทียบ โมเดล 1 และ 2, โมเดล 1 และ 4, โมเดล 2 และ 4) ส่วนผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ พบว่าคุณลักษณะของนักเรียนและสถานศึกษามีผลต่อการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา โดยการการจัดอันดับคุณภาพการจัดการศึกษา ระหว่างโมเดล 1 และ 3, โมเดล 2 และ 3 และระหว่างโมเดล 3 และ 4 มีความแตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 3. การเปรียบเทียบประสิทธิภาพโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ พบว่าโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่มีการควบคุมอิทธิพลของตัวแปรระดับนักเรียน และระดับสถานศึกษาให้ผลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ใกล้เคียงกัน โดยโมเดล 4 มีค่าสัมประสิทธิ์การทำนาย (R^2) สูงสุดเท่ากับ 52.04% รองลงมาได้แก่โมเดล 1 และโมเดล 2 ซึ่งมีค่าสัมประสิทธิ์การทำนาย (R^2) เท่ากับ 51.96% และ 51.86% ตามลำดับ

อัญชลี ธีระวุฒิ (2555) ได้ศึกษาการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ Pre O-net วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 6 ของสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 35 มีวัตถุประสงค์ เพื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบของแบบทดสอบ Pre O-NET วิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 6 ของสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต 35 ด้วยวิธีชิปเทสต์ และวิธีแมนเทิล-แฮนส์เซล จำแนกตามกลุ่มเพศ ที่ตั้ง และขนาดของโรงเรียน และเพื่อวิเคราะห์ความสอดคล้องของผลการตรวจสอบการทำหน้าที่ต่างกันของ ข้อสอบด้วยวิธีชิปเทสต์ และวิธีแมนเทิล-แฮนส์เซล จำแนกตามกลุ่มเพศ ที่ตั้ง และขนาดของโรงเรียน ตัวอย่างในการวิจัยคือผลการตอบข้อสอบคณิตศาสตร์ 40 ข้อ ของนักเรียนจำนวน 1,500 คน ซึ่งได้จากการสุ่มอย่างง่ายทุกระดับชั้นความสามารถ ผลการวิจัยพบว่า 1. การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีชิปเทสต์ และ วิธีแมนเทิล -แฮนส์เซล จำแนกตามกลุ่มเพศ เขตที่ตั้งของโรงเรียน และขนาดของโรงเรียน พบว่า เมื่อจำแนก ตามเพศจะพบข้อสอบที่ทำหน้าที่ต่างกันสูงกว่า จำแนกตามเขตที่ตั้งโรงเรียน และขนาดโรงเรียน เมื่อจำแนกตามเพศโดยวิธีชิปเทสต์ ตรวจสอบพบว่า ข้อสอบทำหน้าที่ต่างกัน 8 ข้อ วิธีแมนเทิล-แฮนส์เซลตรวจสอบพบว่าข้อสอบทำหน้าที่ต่างกัน 7 ข้อ เมื่อจำแนกตามเขตที่ตั้งของโรงเรียนโดย วิธีชิปเทสต์ ตรวจสอบพบว่าข้อสอบทำหน้าที่ต่างกัน 1 ข้อ วิธีแมนเทิล-แฮนส์เซลตรวจสอบพบว่า ข้อสอบทำหน้าที่ต่างกัน 4 ข้อ และเมื่อจำแนกตามขนาดโรงเรียน โดยวิธีชิปเทสต์ และวิธีแมน เทล-แฮนส์เซล ตรวจสอบพบว่าข้อสอบทำหน้าที่ต่างกัน 3 ข้อ เท่ากัน 2. การวิเคราะห์ความสอดคล้องของผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่าง วิธีชิปเทสต์ และวิธีแมนเทิล - แฮนส์เซล จำแนกตามเพศ และตามขนาดโรงเรียน พบว่า วิธีชิปเทสต์ และวิธีแมนเทิล - แฮนส์เซล มีความสอดคล้องกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 แต่เมื่อจำแนกตามเขตที่ตั้งของโรงเรียน ไม่มีความสอดคล้องกันอย่างมีนัยสำคัญทางสถิติที่ ระดับ 0.05

รติพร ถึงฝั่ง (2556) ได้ศึกษาการวิเคราะห์โมเดลมิมิค: การใช้ประโยชน์จากโปรแกรม LISREL รุ่นทดลองใช้เพื่องานวิจัย โดยมีวัตถุประสงค์เพื่อนำเสนอวิธีการวิเคราะห์โมเดลมิมิคด้วยโปรแกรม LISREL รุ่นทดลองใช้ ซึ่งสามารถนำมาใช้ประโยชน์ในการวิเคราะห์ข้อมูลงานวิจัย

โมเดลมีมิคเป็นการวิเคราะห์ข้อมูลในลักษณะ ที่ตัวแปรสังเกตได้ (x-variables) หลายๆ ตัวแปร ทำนายหรือส่งผลต่อตัวแปรแฝง (Eta) ซึ่งตัวแปรแฝงนั้นวัดได้จากตัวบ่งชี้ (y-variables) หลายตัวแปร ขั้นตอนการวิเคราะห์โมเดลมีมิคเริ่มด้วยการนำเข้าข้อมูลจากโปรแกรม SPSS และตามด้วยการวิเคราะห์โมเดลมีมิคด้วยโปรแกรม LISREL ซึ่งมี 5 ขั้นตอนสำคัญคือ 1) การเตรียมข้อมูลด้วยโปรแกรม PRELIS 2) การระบุโมเดลหรือวาดภาพโมเดลการวิจัย 3) การกำหนดการแสดงผลการวิเคราะห์ 4) การวิเคราะห์โมเดล และ 5) การปรับโมเดล ผลการวิเคราะห์จะประกอบไปด้วยค่าสถิติ แสดงความสอดคล้องระหว่างโมเดลตามสมมติฐานกับข้อมูลเชิงประจักษ์ และค่าสถิติแสดงผลของตัวแปรสังเกตที่มีต่อตัวแปรแฝง

สุพัฒนา หอมบุปผา (2556) ได้ศึกษาการเปรียบเทียบการทำหน้าที่ต่างกันของ ข้อสอบ ด้วยวิธี HGLM วิธี MIMIC และวิธี BAYESIAN ซึ่งมีวัตถุประสงค์เพื่อ 1) เปรียบเทียบผลการ ประเมินค่าพารามิเตอร์ความยากของข้อสอบ พารามิเตอร์ความสามารถของผู้สอบ สำหรับผู้สอบ จำแนกตาม เพศ สถานที่ตั้งทางภูมิศาสตร์ของโรงเรียน ระหว่าง วิธี HGLM วิธี MIMIC และ วิธี BAYESIAN 2) เปรียบเทียบผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) สำหรับ ผู้สอบ จำแนกตาม เพศ สถานที่ตั้งทางภูมิศาสตร์ของโรงเรียน ระหว่างวิธี HGLM วิธี MIMIC และวิธี BAYESIAN 3) ศึกษา ลักษณะของข้อสอบที่เกิดการทำหน้าที่ต่างกันของข้อสอบ (DIF) ที่ได้จากการ วิเคราะห์การทำหน้าที่ต่างกัน โดยวิธี HGLM วิธี MIMIC และวิธี BAYESIAN ด้วยการวิเคราะห์ ลักษณะและเนื้อหาของคำหรือข้อความที่ใช้ในการเขียนข้อสอบ ข้อมูลที่ใช้ในการวิเคราะห์ครั้งนี้เป็น คะแนนการสอบวัดผลสัมฤทธิ์ทางการเรียนเพื่อประเมินคุณภาพการศึกษา ระดับชาติของนักเรียนชั้น ประถมศึกษาปีที่ 3 ในรายวิชาภาษาไทย คณิตศาสตร์และวิทยาศาสตร์ของสำนักทดสอบทาง การศึกษา กระทรวงศึกษาธิการ ปีการศึกษา 2553 ซึ่งได้มาโดยการสุ่มแบบหลายขั้นตอน (Multi-Stage Random Sampling Technique) จำนวน 1,000 คน จำแนกเป็นเพศชายและเพศหญิง ที่อยู่ใน เขตกรุงเทพมหานครและปริมณฑล และนอกเขตกรุงเทพมหานครและปริมณฑล การวิเคราะห์ ค่าพารามิเตอร์ความยากของข้อสอบ พารามิเตอร์ความสามารถ ของผู้สอบ และการทำหน้าที่ต่างกัน ของข้อสอบ (DIF) ด้วยวิธี HGLM-2L วิธี MIMIC และวิธี BAYESIAN โดยใช้โปรแกรม สำเร็จรูป 3 โปรแกรม ได้แก่ โปรแกรม HLM โปรแกรม Mplus และโปรแกรม WinBUGS ตามลำดับ เมื่อพบการ ทำหน้าที่ต่างกันของข้อสอบ ผู้วิจัยทำการวิเคราะห์เนื้อหา คำประโยค หรือข้อความ ที่ใช้ในการเขียน ข้อสอบของแต่ละวิชา ผลการวิจัยพบว่า ผลการวิเคราะห์ค่าพารามิเตอร์ความยากของข้อสอบ ค่าพารามิเตอร์ความสามารถของผู้สอบ และผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ในวิชา ภาษาไทย คณิตศาสตร์และวิทยาศาสตร์ด้วยวิธี HGLM-2L วิธี MIMIC และวิธี BAYESIAN พบว่า มีความสัมพันธ์กันในระดับสูงมาก อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 วิธี ตรวจสอบที่พบการทำ หน้าที่ต่างกันของข้อสอบมากที่สุด คือ วิธี HGLM-2L ส่วนวิธีที่ตรวจสอบการทำหน้าที่ต่างกันของ ข้อสอบน้อยที่สุด คือ วิธี MIMIC ผลการศึกษาลักษณะของข้อสอบที่ตรวจพบการทำหน้าที่ต่างกันของ ข้อสอบ (DIF) วิชา ภาษาไทย คณิตศาสตร์และวิทยาศาสตร์ลักษณะของข้อสอบที่เกิดการทำหน้าที่ ต่างกัน เมื่อจำแนกตามเพศ ส่วนใหญ่ ข้อสอบที่เกิดการทำหน้าที่ต่างกันของข้อสอบ จะมีคำศัพท์ที่ เกี่ยวข้องกับเพศนั้น จึงทำให้ข้อสอบเข้าข้างเพศนั้น และอาจ เป็นเพราะความสามารถที่แตกต่าง ระหว่างเพศชายและเพศหญิงที่มีลักษณะความสามารถ ความถนัด และความสนใจใน เรื่องนั้นๆ

ต่างกัน เมื่อจำแนกตามสถานที่ตั้งทางภูมิศาสตร์ของโรงเรียน ส่วนใหญ่สาเหตุที่ทำให้ข้อสอบเกิดการทำหน้าที่ ต่างกันอาจเป็นเพราะประสบการณ์ความคุ้นเคยเกี่ยวกับเรื่องนั้น สภาพแวดล้อมและการฝึกปฏิบัติที่แตกต่างกันระหว่าง นักเรียนในเขตกรุงเทพมหานครและปริมณฑลนักเรียนนอกเขตกรุงเทพมหานครและปริมณฑล

นำชัย ศุภฤกษ์ชัยสกุล (2557) ได้ศึกษาการประยุกต์ใช้การวิเคราะห์แบบจำลองสมการโครงสร้างในงานวิจัยเชิงทดลอง เกี่ยวกับการวิเคราะห์แบบจำลองสมการโครงสร้างเป็นเทคนิคการวิเคราะห์ทางสถิติที่สามารถผนวกแบบจำลอง การวัดตัวแปรทางทฤษฎีและแบบจำลองโครงสร้าง ความสัมพันธ์ระหว่างตัวแปรทางทฤษฎีไว้ในแบบจำลองเดียวกัน ทำให้การวิเคราะห์นี้มีอำนาจการทดสอบที่สูงกว่าในการศึกษาความสัมพันธ์ของตัวแปรและให้ค่าประมาณประชากร ได้เที่ยงตรงมากกว่าเพราะมีการปรับแก้ด้วยความคลาดเคลื่อนในการวัด การนำมาประยุกต์ใช้วิเคราะห์ข้อมูลในงานวิจัยเชิงทดลองจะช่วยเพิ่มความเที่ยงตรงภายในและความเที่ยงตรงในการสรุปผลมากขึ้น บทความนี้จะอธิบายถึงข้อดีของการวิเคราะห์ที่เหนือกว่าสถิติวิเคราะห์หลายตัวแปรแบบดั้งเดิม ทั้งความสามารถศึกษา โครงสร้างความสัมพันธ์ของตัวแปรหลายตัวได้พร้อมกัน อำนาจทดสอบทางสถิติที่เหนือกว่า รวมไปถึงความก้าวหน้า ความยืดหยุ่น และความหลากหลายในการแก้ปัญหาการละเมิดข้อตกลงเบื้องต้นของสถิติทั่วไป การกำหนด แบบจำลองสามารถทำได้ 2 แบบ คือ การกำหนดแบบจำลองให้ตัวแปรแฝงมีตัวแปรสาเหตุและตัวชี้วัดหลายตัว (MIMIC) และการกำหนดให้แบบจำลองมีโครงสร้างค่าเฉลี่ยของตัวแปร (SMM) ทั้งสองแบบล้วนมีข้อดีและข้อเสีย ที่แตกต่างกัน การเลือกใช้จึงขึ้นอยู่กับความเหมาะสมของวัตถุประสงค์ของงานวิจัย รวมทั้งความสามารถในการวิเคราะห์และอ่านแปลผลของผู้วิจัย เพื่อให้เห็นภาพในเชิงเปรียบเทียบที่ชัดเจนและเป็นรูปธรรมในเชิงปฏิบัติ

วรัญญูภรณ์ ชาลีรักษ์และคณะ (2558) ได้ศึกษาปัจจัยที่ส่งผลต่อความคิดเชิงบวกในการทำงานของข้าราชการครูและบุคลากรทางการศึกษา: การวิเคราะห์ MIMIC โดยมีวัตถุประสงค์เพื่อพัฒนาและตรวจสอบความตรงของโมเดลเชิงสาเหตุที่ส่งผลต่อความคิดเชิงบวกในการทำงานของข้าราชการครูและบุคลากรทางการศึกษาและตรวจสอบความสอดคล้องของโมเดลเชิงสาเหตุที่สร้างขึ้นกับข้อมูลเชิงประจักษ์ ตัวอย่างเป็นข้าราชการครูและบุคลากรทางการศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 7 จำนวน 400 คน โดยใช้วิธีการสุ่มแบบหลายขั้นตอน เครื่องมือที่ใช้เป็นแบบสอบถามชนิดมาตราประมาณค่า 5 ระดับ วิเคราะห์ข้อมูลโดยใช้การวิเคราะห์ไมมิก ผลการวิจัย ปรากฏว่า โมเดลไมมิกที่พัฒนาขึ้นตามสมมติฐานมีความสอดคล้องกับข้อมูลเชิงประจักษ์ดี ซึ่งความคิดเชิงบวกในการทำงาน ประกอบด้วย ความผูกพันต่อองค์กร และการทำงานอย่างมีความสุข โดยปัจจัยที่ส่งผลต่อความคิดเชิงบวกในการทำงาน ได้แก่ การมองโลกในแง่ดี ความเชื่อ ความมุ่งมั่น และความอดทน มีขนาดอิทธิพลเท่ากับ 0.41, 0.30, 0.18 และ 0.09 ตามลำดับ โดยที่ปัจจัยทั้ง 4 สามารถร่วมกันอธิบายความแปรปรวนของความคิดเชิงบวกในการทำงานได้ร้อยละ 63

Finch (2005) ได้ศึกษาการเปรียบเทียบประสิทธิภาพของโมเดล MIMIC กับการทดสอบโดยวิธีแมนเทล-แฮนส์เซล (Mantel & Haenszel) และวิธี SIBTEST และวิธีการทดสอบ IRT Likelihood Ratio กับความคลาดเคลื่อนประเภทที่ 1 และอำนาจการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) ได้แสดงให้เห็นว่าวิธี MIMIC มีค่าสูงขึ้นและค่าความคลาดเคลื่อนประเภทที่ 1 มีค่า

ลดลงกับจำนวนข้อสอบ 50 ข้อ นอกจากนี้วิธี MIMIC ยังสามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ Nonuniform DIF ได้ด้วย

Carol M. Woods (2008) ได้ศึกษา Illustration of MIMIC-Model DIF Testing with the Schedule for Nonadaptive and Adaptive Personality เกี่ยวกับตัวอย่างของการทดสอบการทำหน้าที่ต่างกันของข้อสอบ โดยใช้ตัวแปรสาเหตุและตัวบ่งชี้หลายตัว (MIMIC) ซึ่งเป็นโมเดลสมการโครงสร้าง ซึ่งข้อสอบถูกผิด 5 ระดับของ Schedule for Nonadaptive and Adaptive Personality (SNAP) ได้ทำการทดสอบการทำหน้าที่ต่างกันของข้อสอบ แบบบอเนกรูป ในตัวอย่าง โดยกำหนดตามเพศและเชื้อชาติ การทำหน้าที่ต่างกันของข้อสอบ แบบบอเนกรูป ในตัวแปรที่ศึกษา ผลการวิจัยพบว่า พบ ข้อสอบ SNAP มีนัยสำคัญทางสถิติต่อการทำหน้าที่ต่างกันของข้อสอบ และบางอิทธิพลที่มีขนาดใหญ่ การทำหน้าที่ต่างกันของข้อสอบ สามารถสร้างการวัดผลที่ค่อนข้างเบี่ยงเบน และควรศึกษาตัวอย่างที่แตกต่างกัน

จากการศึกษางานวิจัยที่เกี่ยวข้องกับการทำหน้าที่ต่างกันของข้อสอบ พบว่างานวิจัยที่เกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบ ในอดีตที่ผ่านมา มีการศึกษาโดยใช้ทั้งข้อมูลจำลองและข้อมูลเชิงประจักษ์ เป็นการเปรียบเทียบประสิทธิภาพของวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบชนิดต่างๆ ในรูปแบบที่มีความแตกต่างกัน โดยวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่พบว่าได้รับความนิยมในการนำมาใช้การทำวิจัยเป็นจำนวนมาก คือ วิธีแมนเทล – แฮนส์เซล (MH) และ วิธี SIBTEST ส่วนวิธีการตรวจสอบแบบอื่นๆ ก็มีปรากฏให้เห็นบ้าง โดยตัวแปรที่สนใจนำมาศึกษา ได้แก่ ภาษาในการสื่อสาร เพศ สถานที่ตั้งของ โรงเรียน ความสามารถของผู้สอบ ขนาดของ โรงเรียน เชื้อชาติ รายได้ของผู้ปกครอง การเรียนพิเศษของนักเรียน ผลสัมฤทธิ์ที่ผ่านมานักเรียน ระดับการศึกษาของผู้ปกครอง และความมั่งคั่งของครอบครัว เป็นต้น

สำหรับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธี MIMIC ซึ่งการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธี MIMIC เป็นวิธีที่ค่อนข้างใหม่ ซึ่งงานวิจัยในครั้งนี้ให้ความรู้เพิ่มเติมเกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบแบบเอกรูป (Uniform) และแบบบอเนกรูป (Nonuniform) ของโมเดล MIMIC ซึ่งเป็นโมเดลคณิตศาสตร์ขั้นสูง และงานวิจัยการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธี MIMIC ในประเทศไทยยังไม่ได้แพร่หลาย แต่ในระยะหลังมีผู้สนใจศึกษามากขึ้น ทำให้ผู้วิจัยเห็นถึงประโยชน์ของการศึกษา อาจจะเป็นประโยชน์แก่การพัฒนาในเรื่องของการวัดผลเป็นอย่างยิ่ง

ตาราง 2 ผลการสังเคราะห์ตัวแปร

ตัวแปร	รายชื่อผู้ผลิตผลงานวิจัยที่เกี่ยวข้อง														รวม	
	ขวริต (2541)	สมศักดิ์ (2542)	ทอง อยู่ (2543)	นุชริน (2544)	วรรณ (2544)	สุภาณี (2547)	ศุภาวัฒน์ (2548)	ทิพย์ วรรณ (2549)	ปิยะ ทิพย์ (2549)	พริญา (2546)	อิทธิฤทธิ์ (2551)	สุภา ภรณ์ (2552)	เรือง เดช (2554)	อัญชลี (2555)		สุ พัฒนา (2556)
ภาษาในการสื่อสาร						✓	✓		✓			✓				4
เพศ	✓	✓	✓	✓	✓	✓	✓	✓	✓				✓	✓	✓	14
สถานที่ตั้งของ รร	✓				✓									✓	✓	4
ความสามารถของผู้สอบ						✓			✓			✓				4
ขนาดของ รร			✓	✓										✓		4
เชื้อชาติ						✓										1
รายได้ของผู้ปกครอง																1
การเรียนรู้พิเศษ ของ นร														✓		1
ผลสัมฤทธิ์ที่ผ่านมาของ นร																1
ระดับการศึกษาของ ผู้ปกครอง														✓		1
ความมุ่งคั่งของครอบครัว														✓		1

จากตาราง 2 พบว่าในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธี MIMIC ในครั้งนี้ ผู้วิจัยได้สังเคราะห์ตัวแปรจากงานวิจัยที่เกี่ยวข้อง มีตัวแปรอิสระที่ส่งผลทำให้เกิดการทำหน้าที่ต่างกันของข้อสอบจำนวน 11 ตัวแปร ซึ่งได้แก่ ภาษาในการสื่อสาร, เพศ, สถานที่ตั้งของโรงเรียน, ความสามารถของผู้สอบ, ขนาดของโรงเรียน, เชื้อชาติ, รายได้ของผู้ปกครอง, การเรียนพิเศษของนักเรียน, ผลสัมฤทธิ์ที่ผ่านมาของนักเรียน, ระดับการศึกษาของผู้ปกครอง และความมั่งคั่งของครอบครัว

ผู้วิจัยได้สนใจตัวแปรเพศ ได้จำแนกเป็นเพศชายและเพศหญิง เนื่องจากผู้วิจัยได้ศึกษาเอกสาร บทความวิจัย และงานวิจัยที่เกี่ยวข้อง ได้ว่าตัวแปรเพศมีการทำหน้าที่ต่างกันของข้อสอบ จำนวน 14 งานวิจัย ซึ่งมีงานวิจัยที่เกี่ยวข้องของ วรณา (2544), ปิยะทิพย์ (2549), ศุภวัฒน์ (2548), ขวลิขิต (2541), สมศักดิ์ (2542), ทิพย์วรรณ (2549), พิรญา (2546), อัญชลี (2555), นุชริน (2544), สุมาลี (2547), ทองอยู่ (2543), อิทธิฤทธิ์ (2551), เรืองเดช (2554) และสุพัฒนา (2556)

ตัวแปรที่ตั้งของสถานศึกษา ได้จำแนกสถานที่ตั้งของสถานศึกษา คือ สถานที่ตั้งสถานศึกษาที่อยู่ในเมืองและนอกเมือง เนื่องจากผู้วิจัยได้ศึกษาเอกสาร บทความวิจัย และงานวิจัยที่เกี่ยวข้อง ได้ว่าตัวแปรที่ตั้งของสถานศึกษามีการทำหน้าที่ต่างกันของข้อสอบ จำนวน 4 งานวิจัย ซึ่งมีความเกี่ยวข้องของ วรณา (2544), ขวลิขิต (2541), อัญชลี (2555) และสุพัฒนา (2556) จำนวน 4 งานวิจัย ทั้งนี้ สทศ. ได้มีความอนุเคราะห์ข้อมูลเกี่ยวกับเพศและที่ตั้งของสถานศึกษาให้แก่ผู้วิจัย

ดังนั้นผู้วิจัยจึงเลือกตัวแปรเพศและที่ตั้งของสถานศึกษาในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมีค (Multiple Indicators and Multiple Causes: MIMIC) ในครั้งนี้

บทที่ 3

วิธีการดำเนินการวิจัย

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิก (Multiple Indicators and Multiple Causes : MIMIC) วัตถุประสงค์ของการวิจัยเพื่อวิเคราะห์คุณภาพแบบทดสอบของการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) ชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2556 – 2558 และตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิก (Multiple Indicators and Multiple Causes : MIMIC) เมื่อจำแนกตามเพศและที่ตั้งของสถานศึกษา ระหว่างปีการศึกษา 2556 – 2558 โดยมีวิธีการดำเนินการวิจัยดังนี้

ประชากร

นักเรียนชั้นมัธยมศึกษาปีที่ 6 ทั้งหมดที่สอบข้อสอบ O-NET (Ordinary National Educational Test) ระหว่างปีการศึกษา 2556 – 2558

เครื่องมือที่ใช้ในการวิจัย

ข้อสอบ O-NET (Ordinary National Educational Test) ระหว่างปีการศึกษา 2556 – 2558 โดยครอบคลุม 4 กลุ่มสาระการเรียนรู้ ได้แก่ ภาษาไทย ภาษาอังกฤษ คณิตศาสตร์ และวิทยาศาสตร์ ซึ่งจำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ข้อสอบ O-NET (Ordinary National Educational Test) ระหว่างปีการศึกษา 2556 – 2558 สามารถจำแนกได้ดังนี้

ตาราง 3 จำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558

รูปแบบ	แบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาไทย					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	ข้อ	จำนวน	ข้อ	จำนวน	ข้อ	จำนวน
1. ปรนัย						
1.1 5 ตัวเลือก 1 คำตอบ	11-80	70	11-80	70	11-80	70
1.2 5 ตัวเลือก 2 คำตอบ	-	-	-	-	-	-
1.3 หลายตัวเลือก 1 คำตอบ	-	-	-	-	-	-
1.4 หลายตัวเลือก มากกว่า 1 คำตอบ	-	-	-	-	-	-
2. แบบเลือกคำตอบจากแต่ละหมวดที่สัมพันธ์กัน	1-10	10	1-10	10	1-10	10
3. แบบระบายคำตอบที่เป็นค่า/ตัวเลข	-	-	-	-	-	-
รวมจำนวนแบบทดสอบ		80		80		80

จากตาราง 3 พบว่าจำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ของการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 - 2558 มีรูปแบบแบบทดสอบ ได้แก่ แบบเลือกคำตอบจากแต่ละหมวดที่สัมพันธ์กัน ตั้งแต่ ข้อที่ 1 - 10 จำนวน 10 ข้อ และแบบปรนัย 5 ตัวเลือก 1 คำตอบ ตั้งแต่ ข้อที่ 11 - 80 จำนวน 70 ข้อ ผู้วิจัยจึงนำแบบทดสอบ ตั้งแต่ ข้อที่ 11 - 80 จำนวน 70 ข้อ นำมาวิเคราะห์ เนื่องจากมีการให้คะแนนแบบสองค่า (Dichotomous) และแบบทดสอบปีการศึกษา 2556 - 2558 มีจำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ที่เหมือนกัน ทำให้ผู้วิจัยสามารถวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบต่อไป

ตาราง 4 จำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558

รูปแบบ	แบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	ข้อ	จำนวน	ข้อ	จำนวน	ข้อ	จำนวน
1. ปรนัย						
1.1 5 ตัวเลือก 1 คำตอบ	1-80	80	1-80	80	1-80	80
1.2 5 ตัวเลือก 2 คำตอบ	-	-	-	-	-	-
1.3 หลายตัวเลือก 1 คำตอบ	81-90	10	81-90	10	81-90	10
1.4 หลายตัวเลือก มากกว่า 1 คำตอบ	-	-	-	-	-	-
2. แบบเลือกคำตอบจากแต่ละ หมวดที่สัมพันธ์กัน	-	-	-	-	-	-
3. แบบระบายคำตอบที่เป็น ค่า/ตัวเลข	-	-	-	-	-	-
รวมจำนวนแบบทดสอบ		90		90		90

จากตาราง 4 พบว่าจำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ของการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 - 2558 มีรูปแบบแบบทดสอบ ได้แก่ แบบปรนัย 5 ตัวเลือก 1 คำตอบ ตั้งแต่ ข้อที่ 1 – 80 จำนวน 80 ข้อ และ แบบปรนัยหลายตัวเลือก 1 คำตอบ ตั้งแต่ ข้อที่ 81 – 90 จำนวน 10 ข้อ ผู้วิจัยจึงนำแบบทดสอบ ตั้งแต่ ข้อที่ 1 – 80 จำนวน 80 ข้อ นำมาวิเคราะห์ เนื่องจากมีการให้คะแนนแบบสองค่า (Dichotomous) และแบบทดสอบปีการศึกษา 2556 – 2558 มีจำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ที่เหมือนกัน ทำให้ผู้วิจัยสามารถวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบต่อไป

ตาราง 5 จำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558

รูปแบบ	แบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	ข้อ	จำนวน	ข้อ	จำนวน	ข้อ	จำนวน
1. ปรนัย						
1.1 5 ตัวเลือก 1 คำตอบ	1-32	32	1-32	32	1-32	32
1.2 5 ตัวเลือก 2 คำตอบ	-	-	-	-	-	-
1.3 หลายตัวเลือก 1 คำตอบ	-	-	-	-	-	-
1.4 หลายตัวเลือก มากกว่า 1 คำตอบ	-	-	-	-	-	-
2. แบบเลือกคำตอบจากแต่ละหมวดที่สัมพันธ์กัน	-	-	-	-	-	-
3. แบบบรรยายคำตอบที่เป็นค่า/ตัวเลข	33-40	8	33-40	8	33-40	8
รวมจำนวนแบบทดสอบ		40		40		40

จากตาราง 5 พบว่าจำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ของการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 - 2558 มีรูปแบบแบบทดสอบ ได้แก่ แบบปรนัย 5 ตัวเลือก 1 คำตอบ ตั้งแต่ ข้อที่ 1 – 32 จำนวน 32 ข้อ และแบบบรรยายคำตอบที่เป็นค่า/ตัวเลข ตั้งแต่ ข้อที่ 33 – 40 จำนวน 8 ข้อ ผู้วิจัยจึงนำแบบทดสอบ ตั้งแต่ ข้อที่ 1 – 32 จำนวน 32 ข้อ นำมาวิเคราะห์ เนื่องจากมีการให้คะแนนแบบสองค่า (Dichotomous) และแบบทดสอบปีการศึกษา 2556 – 2558 มีจำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ที่เหมือนกัน ทำให้ผู้วิจัยสามารถวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบต่อไป

ตาราง 6 จำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558

รูปแบบ	แบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	ข้อ	จำนวน	ข้อ	จำนวน	ข้อ	จำนวน
1. ประนัย						
1.1 5 ตัวเลือก 1 คำตอบ	1-80	80	1-80	80	1-80	80
1.2 5 ตัวเลือก 2 คำตอบ	-	-	-	-	-	-
1.3 หลายตัวเลือก 1 คำตอบ	-	-	-	-	-	-
1.4 หลายตัวเลือก มากกว่า 1 คำตอบ	81-90	10	81-90	10	81-90	10
2. แบบเลือกคำตอบจากแต่ละหมวดที่สัมพันธ์กัน	-	-	-	-	-	-
3. แบบระบายคำตอบที่เป็นค่า/ตัวเลข	-	-	-	-	-	-
รวมจำนวนแบบทดสอบ		90		90		90

จากตาราง 6 พบว่าจำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ของการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 - 2558 มีรูปแบบแบบทดสอบ ได้แก่ แบบประนัย 5 ตัวเลือก 1 คำตอบ ตั้งแต่ ข้อที่ 1 – 80 จำนวน 80 ข้อ และแบบประนัย หลายตัวเลือก มากกว่า 1 คำตอบ ตั้งแต่ ข้อที่ 81 – 90 จำนวน 10 ข้อ ผู้วิจัยจึงนำแบบทดสอบ ตั้งแต่ ข้อที่ 1 – 80 จำนวน 80 ข้อ นำมาวิเคราะห์ เนื่องจากมีการให้คะแนนแบบสองค่า (Dichotomous) และแบบทดสอบปีการศึกษา 2556 – 2558 มีจำนวนแบบทดสอบและรูปแบบแบบทดสอบ (Item Form) ที่เหมือนกัน ทำให้ผู้วิจัยสามารถวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบต่อไป

จากแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) ปีการศึกษา 2556 – 2558 ผู้วิจัยจึงได้นำแบบทดสอบรายวิชาภาษาไทย ตั้งแต่ ข้อที่ 11 – 80 จำนวน 70 ข้อ, แบบทดสอบรายวิชาภาษาอังกฤษ ตั้งแต่ ข้อที่ 1 – 80 จำนวน 80 ข้อ, แบบทดสอบรายวิชาคณิตศาสตร์ ตั้งแต่ ข้อที่ 1 – 32 จำนวน 32 ข้อและแบบทดสอบรายวิชาวิทยาศาสตร์ ตั้งแต่ ข้อที่ 1 – 80 จำนวน 80 ข้อ ไปวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบต่อไป

การเก็บรวบรวมข้อมูล

ผู้วิจัยรวบรวมข้อมูลผลการตอบของนักเรียน ซึ่งเป็นข้อมูลทุติยภูมิจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) (สทศ.) โดยมีขั้นตอนการเก็บรวบรวมข้อมูล ดังนี้

1. ขอนหนังสือจากภาควิชาประเมินผลและวิจัยทางการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยสงขลานครินทร์ เพื่อขอข้อมูลผลการตอบข้อสอบของนักเรียน จากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) (สทศ.)
2. เมื่อได้รับข้อมูลจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) (สทศ.) แล้ว ผู้วิจัยจำแนกข้อมูลตามเพศ และที่ตั้งของสถานศึกษา ระหว่างปีการศึกษา 2556 – 2558
3. ตรวจสอบความเรียบร้อยของข้อมูล แล้วนำข้อมูลมาวิเคราะห์ตามวัตถุประสงค์ของการวิจัย

การวิเคราะห์ข้อมูล

1. วิเคราะห์ ค่าสถิติพื้นฐานของข้อมูลโดยเป็นการวิเคราะห์ข้อมูลเบื้องต้นด้วยสถิติภาคบรรยาย ได้แก่ ความถี่ ค่าต่ำสุด ค่าสูงสุด ค่าเฉลี่ย ส่วนส่วนเบี่ยงเบนมาตรฐาน (SD) ค่าความเบ้ (Skewness) และค่าความโด่ง (Kurtosis)
2. การประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ และพารามิเตอร์ความยากและอำนาจจำแนกของข้อสอบ ตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory)

ในขั้นตอนนี้ ผู้วิจัยได้เขียนคำสั่งจากโปรแกรม Mplus เพื่อหาค่าพารามิเตอร์ความสามารถของผู้สอบ และพารามิเตอร์ความยากและอำนาจจำแนกของข้อสอบ
3. การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิค (Multiple Indicators and Multiple Causes : MIMIC) เมื่อจำแนกตามเพศและที่ตั้งของสถานศึกษา ระหว่างปีการศึกษา 2556 – 2558 โดยประยุกต์ใช้โปรแกรม Mplus มีขั้นตอนดังต่อไปนี้
 - 3.1 วิเคราะห์โมเดล CFA โดยไม่มีตัวแปรทำนาย ในการวิเคราะห์ข้อมูลตามทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นการศึกษาคุณสมบัติอยู่บนพื้นฐานข้อตกลงเอกมิติ (Unidimensional) ของตัวแปรแฝง
 - 3.2 การเพิ่มตัวแปรทำนาย แต่ไม่มีอิทธิพลตรงต่อข้อสอบ ซึ่งในการวิจัยครั้งนี้ ผู้วิจัยได้มีตัวแปร คือเพศ และที่ตั้งของสถานศึกษา
 - 3.3 เพิ่มอิทธิพลทางตรงต่อข้อสอบบังคับค่าให้เป็น 0 เพื่อกำหนดให้ตัวแปรต้นไม่มีผลทางตรงต่อข้อสอบ
 - 3.4 ประเมินความกลมกลืนของโมเดล และตรวจผลทางตรงที่มีนัยสำคัญทางสถิติ ถ้าตรวจพบว่ามีผลทางตรง และมีนัยสำคัญทางสถิติ แสดงว่าข้อสอบข้อนั้น คือ ข้อที่เกิดการทำหน้าที่ต่างกันของข้อสอบ (DIF)

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. วิเคราะห์ค่าสถิติพื้นฐานของข้อมูล

1.1 หาค่าเฉลี่ยของคะแนนทดสอบ (\bar{X}) โดยคำนวณจากสูตร ดังนี้

$$\bar{X} = \frac{\sum X}{n}$$

โดย \bar{X} แทน ค่าเฉลี่ย
 $\sum X$ แทน ผลรวมของคะแนนทั้งหมดในกลุ่ม
 n แทน ขนาดของตัวอย่าง

1.2 ความเบี่ยงเบนมาตรฐาน (S.D.) ของคะแนนทดสอบ ใช้สูตรดังนี้

$$S.D. = \sqrt{\frac{n\sum X^2 - (\sum X)^2}{n(n-1)}}$$

โดย $S.D.$ แทน ส่วนเบี่ยงเบนมาตรฐาน
 $\sum X$ แทน ผลรวมของคะแนน
 $\sum X^2$ แทน ผลรวมทั้งหมดของกำลังสองของคะแนน
 n แทน ขนาดของตัวอย่าง

1.3 ความเบ้ (Skewness) ของคะแนนทดสอบ วิธีของ Bowley ใช้สูตรดังนี้

$$\text{สัมประสิทธิ์ความเบ้} = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1}$$

โดย Q_1 แทน คลอไทล์ที่ 1
 Q_2 แทน คลอไทล์ที่ 2
 Q_3 แทน คลอไทล์ที่ 3

1.4 ความโค้ง (Kurtosis) ของคะแนนทดสอบ ใช้สูตรดังนี้

$$\text{สัมประสิทธิ์ความเบ้} = \frac{QD}{P_{90} - P_{10}}$$

โดย QD แทน ส่วนเบี่ยงเบนควอไทล์
 P_{90} แทน เปอร์เซนต์ไทล์ที่ 90
 P_{10} แทน เปอร์เซนต์ไทล์ที่ 10

2. การประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ และพารามิเตอร์ความยากและอำนาจจำแนกของข้อสอบ ตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory)

Models	Normal Ogive Function	Logistic Function
1 Parameter	$P_i(\theta) = \int_{-\infty}^{\theta-b_i} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$	$P_i(\theta) = \frac{1}{1 + e^{-(\theta-b_i)}}$
2 Parameter	$P_i(\theta) = \int_{-\infty}^{a_i(\theta-b_i)} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$	$P_i(\theta) = \frac{1}{1 + e^{-Da_i(\theta-b_i)}}$
3 Parameter	$P_i(\theta) = c_i + (1 - c_i) \int_{-\infty}^{a_i(\theta-b_i)} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$	$P_i(\theta) = c_i + \frac{1 - c_i}{1 + e^{-Da_i(\theta-b_i)}}$

โดย θ แทน ค่าพารามิเตอร์ความสามารถของผู้สอบ
 a_i แทน ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i
 (Discrimination parameter)
 b_i แทน ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i
 (Difficulty parameter)
 c_i แทน ค่าพารามิเตอร์โอกาสในการเดาของข้อสอบข้อ
 ที่ i ได้ถูก (Difficulty parameter)

ผู้วิจัยเลือกใช้ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) รูปแบบ 2 พารามิเตอร์ ในการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ และพารามิเตอร์ความยากและอำนาจจำแนกของข้อสอบ เนื่องจากการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมีค (Multiple Indicators and Multiple Causes : MIMIC) นั้นมีรูปแบบการวิเคราะห์ 2 พารามิเตอร์

โดยใช้โมเดลการวิเคราะห์องค์ประกอบ (Factor Analysis) รูปแบบ 2-PL IRT Model (Kamata, A and Banner, D.J., 2008)

เมื่อ	θ	แทน	ระดับความสามารถของผู้สอบ (trait level)
	λ_i	แทน	ค่าอำนาจจำแนกของข้อสอบข้อที่ i (loading: a)
	τ_i	แทน	ค่าความยากของข้อสอบข้อที่ i (threshold: $b = \frac{\tau}{\lambda}$)

3. การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธี MIMIC

3.1 การทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธี MIMIC แบบเอกรูป (Uniform)

สูตรสำหรับวิธี MIMIC ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) แบบเอกรูป (Carol M. Woods and Kevin J. Grimm, 2011, น. 342) คือ

$$y_i^* = \alpha_i \theta + \beta_i z + \varepsilon_i$$

โดย	y_i^*	คือ	การตอบสนองการทำหน้าที่ต่างกันของข้อสอบข้อที่ i
	α	คือ	อำนาจจำแนกของพารามิเตอร์
	θ	คือ	ตัวแปรแฝง
	β	คือ	สัมประสิทธิ์ของการถดถอยของการทดสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธี MIMIC แบบเอกรูป
	z	คือ	กลุ่มตัวแปร
	ε	คือ	ความผิดพลาดในการวัดของข้อสอบข้อที่ i

2.2 การทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธี MIMIC แบบอนเนกรูป (Nonuniform)

สูตรสำหรับวิธี MIMIC ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) แบบอนเนกรูป (Carol M. Woods and Kevin J. Grimm, 2011, น. 342) คือ

$$y_i^* = \alpha_i\theta + \alpha_i z + \omega_i\theta z + \varepsilon_i$$

โดย	y_i^*	คือ การตอบสนองการทำหน้าที่ต่างกันของข้อสอบข้อที่ i
	α	คือ อำนาจจำแนกของพารามิเตอร์
	θ	คือ ตัวแปรแฝง
	ω	คือ สัมประสิทธิ์ของการถดถอยของการทดสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธี MIMIC แบบอนเนกรูป
	z	คือ กลุ่มตัวแปร
	ε	คือ ความผิดพลาดในการวัดของข้อสอบข้อที่ i

Prince of Songkla University
Pattani Campus

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีวัตถุประสงค์ของการวิจัย คือ เพื่อวิเคราะห์คุณภาพแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2556 - 2558 และตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิค (Multiple Indicators and Multiple Causes : MIMIC) เมื่อจำแนกตามเพศและที่ตั้งของสถานศึกษา ระหว่างปีการศึกษา 2556 – 2558 ในแบบทดสอบที่มีการให้คะแนนแบบสองค่า (Dichotomous) วิธี MIMIC โดยประยุกต์ใช้โปรแกรม Mplus ซึ่งผลการวิจัยวิเคราะห์ข้อมูลครั้งนี้ผู้วิจัยได้เสนอเรียงตามลำดับดังนี้

ตอนที่ 1 การวิเคราะห์ค่าสถิติพื้นฐาน

ตอนที่ 2 ผลการประมาณค่าพารามิเตอร์ความยาก อำนาจจำแนกของข้อสอบ และระดับความสามารถของผู้สอบ

ตอนที่ 3 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบจำแนกตามเพศและที่ตั้งของสถานศึกษา

ตอนที่ 4 ผลการวิเคราะห์แนวโน้มการทำหน้าที่ต่างกันของข้อสอบจำแนกตามเพศและที่ตั้งของสถานศึกษา

สัญลักษณ์ที่ใช้ในการวิเคราะห์ข้อมูล

เพื่อให้การนำเสนอผลการวิเคราะห์ข้อมูลมีความเข้าใจตรงกัน ผู้วิจัยจึงได้กำหนดสัญลักษณ์และความหมายแทน ดังนี้

MIMIC	หมายถึง	วิธีการวิเคราะห์ข้อมูลคุณลักษณะแฝงที่มีหลายสาเหตุและวัดได้จากตัวบ่งชี้หลายตัว (Multiple Indicators and Multiple Causes Model)
\bar{x}	หมายถึง	ค่าเฉลี่ย (Mean)
SD	หมายถึง	ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation)
Max	หมายถึง	ค่าสูงสุด (Maximun)
Min	หมายถึง	ค่าต่ำสุด (Minimun)
$Skew$	หมายถึง	ค่าความเบ้ (Skewness)
$Kurt$	หมายถึง	ค่าความโด่ง (Kurtosis)
SE	หมายถึง	ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า (Standard Error)
a_i	หมายถึง	ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ

b_i	หมายถึง	ค่าพารามิเตอร์ความยากของข้อสอบ
θ	หมายถึง	ค่าพารามิเตอร์ความสามารถของผู้สอบ (Ability Parameter)
<i>CFA</i>	หมายถึง	การวิเคราะห์องค์ประกอบยืนยัน (Confirmatory Factor)
x^2	หมายถึง	ค่าสถิติไค-สแควร์ (Chi - Square)
<i>df</i>	หมายถึง	องศาอิสระ (Degree of freedom)
<i>CFI</i>	หมายถึง	ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (Comparative Fit Index)
<i>TLI</i>	หมายถึง	ดัชนี The Tucker lewis Index
<i>RMSEA</i>	หมายถึง	ดัชนีวัดรากที่สองของความคลื่อนในการประมาณ ค่าพารามิเตอร์ที่พอเหมาะ (Root Mean Square Error of Approximation)
<i>WRMR</i>	หมายถึง	ดัชนี Weighted root Mean Square Residual

สัญลักษณ์ที่ใช้แทนตัวแปรแฝง

<i>F1</i>	หมายถึง	แบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 - 2558
<i>F2</i>	หมายถึง	แบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 - 2558
<i>F3</i>	หมายถึง	แบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 - 2558
<i>F4</i>	หมายถึง	แบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 - 2558

สัญลักษณ์ที่ใช้แทนตัวแปรสังเกตได้

<i>T01–T70</i>	หมายถึง	แบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ตั้งแต่ข้อที่ 1 ถึงข้อที่ 70 จำนวน 70 ข้อ ปีการศึกษา 2556 - 2558
<i>E01–E80</i>	หมายถึง	แบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ตั้งแต่ข้อที่ 1 ถึงข้อที่ 80 จำนวน 80 ข้อ ปีการศึกษา 2556 - 2558
<i>M01–M32</i>	หมายถึง	แบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้น พื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ตั้งแต่ข้อที่ 1 ถึงข้อ ที่ 32 จำนวน 32 ข้อ ปีการศึกษา 2556 - 2558

S01–S80 หมายถึง แบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ตั้งแต่ข้อที่ 1 ถึงข้อที่ 80 จำนวน 80 ข้อ ปีการศึกษา 2556 – 2558

ตอนที่ 1 การวิเคราะห์ค่าสถิติพื้นฐาน

การวิเคราะห์ข้อมูลส่วนนี้ผู้วิจัยได้นำคะแนนที่ได้จากการตอบข้อสอบของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) ปีการศึกษา 2556 - 2558 โดยที่รายวิชาภาษาไทย ภาษาอังกฤษ คณิตศาสตร์ และวิทยาศาสตร์ มีจำนวนแบบทดสอบรายวิชาละ 70, 80, 32 และ 80 ตามลำดับ

ตาราง 7 จำนวนผู้เข้าสอบ คะแนนต่ำสุด คะแนนสูงสุด คะแนนเฉลี่ย คะแนนส่วนเบี่ยงเบนมาตรฐาน ค่าความเบ้และค่าความโด่ง จากคะแนนทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558 จำแนกตามเพศและที่ตั้งของสถานศึกษา

ปีการศึกษา	ตัวแปร	ค่าสถิติ							
		<i>n</i>	<i>Min</i>	<i>Max</i>	\bar{X}	<i>SD</i>	<i>Skew</i>	<i>Kurt</i>	
2556	เพศ	เพศชาย	158,948	0	67	30.99	11.05	.19	-.66
		เพศหญิง	255,641	0	69	36.10	10.32	-.04	-.48
	ที่ตั้งของสถานศึกษา	ในเมือง	158,377	0	67	33.72	11.04	.03	-.63
		นอกเมือง	256,212	0	69	34.41	10.80	-.01	-.57
	รวม		414,589	0	69	34.15	10.89	.01	-.60
2557	เพศ	เพศชาย	166,675	0	44	21.24	7.47	.00	-.70
		เพศหญิง	263,942	0	45	24.33	6.80	-.20	-.43
	ที่ตั้งของสถานศึกษา	ในเมือง	163,165	0	44	22.99	7.25	-.15	-.57
		นอกเมือง	267,452	0	45	23.23	7.22	-.18	-.56
	รวม		430,617	0	45	23.14	7.23	-.16	-.56
2558	เพศ	เพศชาย	162,544	0	68	30.93	11.35	.31	-.55
		เพศหญิง	260,887	0	68	36.01	10.79	.05	-.56
	ที่ตั้งของสถานศึกษา	ในเมือง	162,866	0	68	33.38	11.34	.17	-.59
		นอกเมือง	260,565	0	68	34.49	11.23	.09	-.62
	รวม		423,431	0	68	34.06	11.29	.12	-.62

หมายเหตุ ค่าความเบ้ และค่าความโด่งมีค่าอยู่ระหว่าง -1.96 ถึง 1.96 ลักษณะของการแจกแจงเป็นโค้งปกติ

จากตาราง 7 พบว่า คะแนนทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ในปีการศึกษา 2556 โดยภาพรวมเพศหญิงมีค่าเฉลี่ย 36.10 คะแนน และที่ตั้งของสถานศึกษานอกเมือง มีค่าเฉลี่ย 34.41 คะแนน ในปีการศึกษา 2557 โดยภาพรวมเพศหญิงมีค่าเฉลี่ย 24.33 คะแนน และที่ตั้งของสถานศึกษานอกเมือง มีค่าเฉลี่ย 23.23 คะแนน และในปีการศึกษา 2558 โดยภาพรวมเพศหญิงมีค่าเฉลี่ย 36.01 คะแนน และที่ตั้งของสถานศึกษานอกเมือง มีค่าเฉลี่ย 34.49 คะแนน

ตาราง 8 จำนวนผู้เข้าสอบ คะแนนต่ำสุด คะแนนสูงสุด คะแนนเฉลี่ย คะแนนส่วนเบี่ยงเบนมาตรฐาน ค่าความเบ้และค่าความโด่ง จากคะแนนทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558 จำแนกตามเพศและที่ตั้งของสถานศึกษา

ปีการศึกษา	ตัวแปร	ค่าสถิติ							
		<i>n</i>	<i>Min</i>	<i>Max</i>	\bar{X}	<i>SD</i>	<i>Skew</i>	<i>Kurt</i>	
2556	เพศ	เพศชาย	159,388	0	77	22.23	10.50	2.03	4.83
		เพศหญิง	256,070	3	78	23.91	10.54	1.83	3.99
	ที่ตั้งของสถานศึกษา	ในเมือง	158,732	1	78	23.17	10.76	1.97	4.55
	นอกเมือง	256,726	0	78	23.33	10.44	1.84	4.02	
	รวม	415,458	0	78	23.27	10.56	1.89	4.24	
2557	เพศ	เพศชาย	167,202	0	78	20.50	9.92	2.20	5.70
		เพศหญิง	264,376	2	79	21.85	9.66	1.98	4.91
	ที่ตั้งของสถานศึกษา	ในเมือง	163,581	1	78	21.22	9.77	2.08	5.33
	นอกเมือง	267,997	0	79	21.39	9.80	2.04	5.04	
	รวม	431,578	0	79	21.33	9.79	2.05	5.14	
2558	เพศ	เพศชาย	163,007	0	79	21.25	10.14	2.16	5.45
		เพศหญิง	261,180	0	79	22.73	9.93	1.96	4.74
	ที่ตั้งของสถานศึกษา	ในเมือง	163,203	0	79	22.26	1.29	2.03	4.87
	นอกเมือง	260,984	1	79	22.10	9.88	2.01	4.96	
	รวม	424,187	0	79	22.16	1.04	2.02	4.94	

หมายเหตุ ค่าความเบ้ และค่าความโด่งมีค่าอยู่ระหว่าง -1.96 ถึง 1.96 ลักษณะของการแจกแจงเป็นโค้งปกติ

จากตาราง 8 พบว่า คะแนนทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ในปีการศึกษา 2556 โดยภาพรวมเพศหญิงมีค่าเฉลี่ย 23.91 คะแนน และที่ตั้งของสถานศึกษานอกเมือง มีค่าเฉลี่ย 23.33 คะแนน ในปีการศึกษา 2557 โดยภาพรวมเพศหญิงมีค่าเฉลี่ย 21.85 คะแนน และที่ตั้งของสถานศึกษานอกเมือง มีค่าเฉลี่ย 21.39 คะแนน และใน

ปีการศึกษา 2558 โดยภาพรวมเพศหญิงมีค่าเฉลี่ย 22.73 คะแนน และที่ตั้งของสถานศึกษาในเมือง มีค่าเฉลี่ย 22.26 คะแนน

ตาราง 9 จำนวนผู้เข้าสอบ คะแนนต่ำสุด คะแนนสูงสุด คะแนนเฉลี่ย คะแนนส่วนเบี่ยงเบนมาตรฐาน ค่าความเบ้และค่าความโด่ง จากคะแนนทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชา คณิตศาสตร์ ปีการศึกษา 2556 – 2558 จำแนกตามเพศและที่ตั้งของสถานศึกษา

ปี การศึกษา	ตัวแปร		ค่าสถิติ						
			<i>n</i>	<i>Min</i>	<i>Max</i>	\bar{X}	<i>SD</i>	<i>Skew</i>	<i>Kurt</i>
2556	เพศ	เพศชาย	159,620	0	32	7.66	4.82	2.28	6.47
		เพศหญิง	256,125	0	32	7.27	4.16	2.28	7.59
	ที่ตั้งของ สถานศึกษา	ในเมือง	158,840	0	32	7.47	4.52	2.33	7.23
		นอกเมือง	256,905	0	32	7.40	4.38	2.30	7.29
		รวม	415,745	0	32	7.43	4.43	2.31	7.28
2557	เพศ	เพศชาย	167,520	0	32	8.17	5.06	2.08	5.16
		เพศหญิง	264,475	0	32	7.79	4.41	2.04	5.90
	ที่ตั้งของ สถานศึกษา	ในเมือง	163,728	0	32	8.02	4.80	2.08	5.51
		นอกเมือง	268,267	0	32	7.89	4.61	2.09	5.86
		รวม	431,995	0	32	7.94	4.68	2.09	5.73
2558	เพศ	เพศชาย	163,210	0	32	9.54	4.92	1.67	3.53
		เพศหญิง	261,220	0	32	9.57	4.34	1.52	3.50
	ที่ตั้งของ สถานศึกษา	ในเมือง	163,289	0	32	9.57	4.64	1.62	3.61
		นอกเมือง	261,141	0	32	9.56	4.53	1.59	3.62
		รวม	424,430	0	32	9.56	4.58	1.60	3.62

หมายเหตุ ค่าความเบ้ และค่าความโด่งมีค่าอยู่ระหว่าง -1.96 ถึง 1.96 ลักษณะของการแจกแจง เป็นโค้งปกติ

จากตาราง 9 พบว่า คะแนนทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ในปีการศึกษา 2556 โดยภาพรวมเพศชายมีค่าเฉลี่ย 7.66 คะแนน และที่ตั้งของสถานศึกษาในเมือง มีค่าเฉลี่ย 7.47 คะแนน ในปีการศึกษา 2557 โดยภาพรวมเพศชายมีค่าเฉลี่ย 8.17 คะแนน และที่ตั้งของสถานศึกษาในเมือง มีค่าเฉลี่ย 8.02 คะแนน และในปีการศึกษา 2558 โดยภาพรวมเพศหญิงมีค่าเฉลี่ย 9.57 คะแนน และที่ตั้งของสถานศึกษาในเมือง มีค่าเฉลี่ย 9.57 คะแนน

ตาราง 10 จำนวนผู้เข้าสอบ คะแนนต่ำสุด คะแนนสูงสุด คะแนนเฉลี่ย คะแนนส่วนเบี่ยงเบนมาตรฐาน ค่าความเบ้และค่าความโด่ง จากคะแนนทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558 จำแนกตามเพศและที่ตั้งของสถานศึกษา

ปีการศึกษา	ตัวแปร		ค่าสถิติ						
			<i>n</i>	<i>Min</i>	<i>Max</i>	\bar{X}	<i>SD</i>	<i>Skew</i>	<i>Kurt</i>
2556	เพศ	เพศชาย	158,963	2	73	22.90	8.24	1.60	3.78
		เพศหญิง	255,625	2	72	23.55	7.57	1.40	3.17
	ที่ตั้งของสถานศึกษา	ในเมือง	158,412	2	73	23.29	7.98	1.54	3.63
		นอกเมือง	256,176	2	73	23.31	7.75	1.45	3.31
	รวม		414,588	2	73	23.31	7.84	1.48	3.45
2557	เพศ	เพศชาย	166,639	0	75	24.17	7.90	1.32	2.78
		เพศหญิง	263,968	0	72	24.67	7.24	1.09	2.06
	ที่ตั้งของสถานศึกษา	ในเมือง	163,234	0	74	24.39	7.56	1.25	2.68
		นอกเมือง	267,373	0	75	24.54	7.48	1.16	2.23
	รวม		430,607	0	75	24.48	7.51	1.19	2.40
2558	เพศ	เพศชาย	162,628	0	71	24.23	7.46	1.19	2.45
		เพศหญิง	260,867	1	69	24.95	6.83	.96	1.82
	ที่ตั้งของสถานศึกษา	ในเมือง	162,876	0	71	24.66	7.23	1.13	2.36
		นอกเมือง	260,619	0	70	24.69	7.00	1.00	1.91
	รวม		423,495	0	71	24.68	7.09	1.05	2.10

หมายเหตุ ค่าความเบ้ และค่าความโด่งมีค่าอยู่ระหว่าง -1.96 ถึง 1.96 ลักษณะของการแจกแจงเป็นโค้งปกติ

จากตาราง 10 พบว่า คะแนนทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ในปีการศึกษา 2556 โดยภาพรวมเพศหญิงมีค่าเฉลี่ย 23.55 คะแนน และที่ตั้งของสถานศึกษานอกเมือง มีค่าเฉลี่ย 23.31 คะแนน ในปีการศึกษา 2557 โดยภาพรวมเพศหญิงมีค่าเฉลี่ย 24.67 คะแนน และที่ตั้งของสถานศึกษานอกเมือง มีค่าเฉลี่ย 24.54 คะแนน และในปีการศึกษา 2558 โดยภาพรวมเพศหญิงมีค่าเฉลี่ย 24.95 คะแนน และที่ตั้งของสถานศึกษานอกเมือง มีค่าเฉลี่ย 24.69 คะแนน

ตอนที่ 2 ผลการประมาณค่าพารามิเตอร์ความยาก อำนาจจำแนกของข้อสอบ และระดับความสามารถของผู้สอบ

การวิเคราะห์ข้อมูลส่วนนี้ผู้วิจัยได้นำคะแนนที่ได้จากการตอบข้อสอบของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ปีการศึกษา 2556 - 2558 โดยที่รายวิชาภาษาไทย ภาษาอังกฤษ คณิตศาสตร์ และวิทยาศาสตร์ มีจำนวนแบบทดสอบรายวิชาละ 70, 80, 32 และ 80 ตามลำดับ มาวิเคราะห์ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบ โดยประยุกต์ใช้โปรแกรม Mplus ซึ่งประมาณได้จากโมเดลตามทฤษฎีการตอบสนองข้อสอบ

ตาราง 11 ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 - 2558

ข้อสอบ	ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	b_i	a_i	b_i	a_i	b_i	a_i
1	.81	.11	1.39	.80	.95	.50
2	.41	.38	.89	.62	1.43	.42
3	.72	.42	1.66	.62	.09	.16
4	.64	.23	1.59	.14	1.42	.36
5	1.67	.84	.64	.41	1.57	.41
6	.68	.64	.75	.15	.70	.39
7	.26	.28	.95	.32	.23	.58
8	1.35	.25	.48	.42	1.04	.52
9	1.49	.36	.82	.68	1.36	.33
10	1.73	.46	.48	.23	.95	.33
11	.31	.45	1.28	.99	.60	.60
12	1.09	.09	.76	.64	1.13	.57
13	.64	.50	1.43	.72	.14	.29
14	.28	.53	.93	.20	1.00	.48
15	1.43	1.03	.74	.49	1.20	.43
16	1.89	.21	.65	.20	.48	.51
17	.17	.45	1.13	.29	.28	.54
18	.74	.50	.35	.40	.84	.61
19	.75	.59	.72	.70	1.58	.29

ตารางที่ 11 (ต่อ)

ข้อสอบ	ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	b_i	a_i	b_i	a_i	b_i	a_i
20	1.22	.72	.38	.27	.79	.37
21	5.64	.08	7.94	.08	1.00	.40
22	.10	.58	2.58	.23	.38	.36
23	.96	.81	1.10	.20	1.19	.31
24	.40	.38	.03	.67	.52	.44
25	.69	.37	.38	.58	.18	.24
26	.13	.38	.89	.41	.55	.41
27	2.27	.28	8.32	.10	1.02	.64
28	.28	.37	2.44	.18	.09	.26
29	1.15	.15	4.22	.18	.54	.48
30	.09	.71	.11	.48	1.07	.59
31	.68	.47	.07	.56	.12	.48
32	.26	.52	.47	.74	.02	.42
33	.19	.49	4.78	.18	.97	.43
34	1.87	.05	.21	.85	.75	.46
35	.43	.45	1.61	.83	.96	.38
36	.76	.54	1.65	.36	.50	.60
37	.17	.63	.86	.40	.89	.65
38	.35	.39	.04	.34	.85	.34
39	1.09	.76	.75	.79	.64	.59
40	.47	.59	.04	.56	.38	.42
41	1.23	1.03	.84	.63	.93	.46
42	.84	.37	.14	.50	1.24	.47
43	1.09	.88	1.05	.58	1.34	.37
44	.56	.60	.10	.63	.69	.12
45	.29	.34	.07	.66	1.89	.31
46	5.77	.01	1.54	.34	.80	.39
47	2.02	.27	.94	.40	.26	.51
48	1.30	.27	.03	.31	.20	.54
49	.72	.69	.72	.85	.05	.40
50	1.88	.06	.01	.55	.46	.56

ตารางที่ 11 (ต่อ)

ข้อสอบ	ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	b_i	a_i	b_i	a_i	b_i	a_i
51	1.52	.80	.79	.61	.03	.63
52	.11	.54	.10	.48	.26	.67
53	1.75	.55	1.00	.55	1.29	.32
54	.67	.80	.10	.66	3.25	.25
55	.02	.67	.03	.64	4.95	.22
56	4.68	.14	5.35	.07	.71	.57
57	.92	.43	.28	.60	.34	.49
58	7.60	.01	.62	.79	.48	.84
59	.67	.37	.89	.37	.06	.60
60	.51	.32	.85	.65	.58	.84
61	.15	.35	.73	.48	.07	.47
62	.28	.69	1.53	.27	.22	.52
63	.40	.55	1.59	.37	.20	.29
64	1.49	.24	.13	.67	.51	.67
65	1.51	.06	2.05	.07	.84	.39
66	.19	.64	.88	.70	.56	.95
67	.57	.32	.40	.66	.33	.45
68	.30	.63	.21	.54	.06	.59
69	.57	.41	.37	.34	.04	.63
70	.09	.55	.49	.44	.10	.34

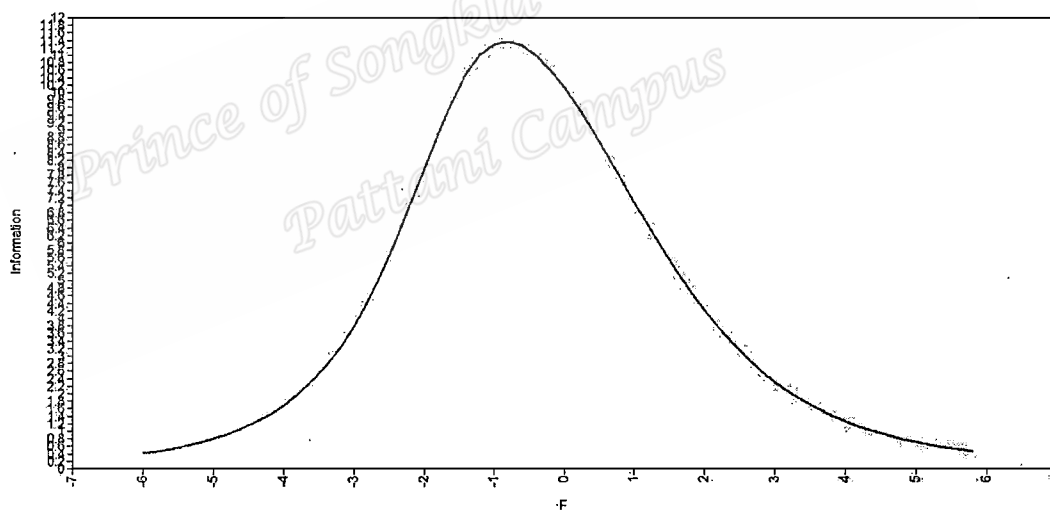
จากตาราง 11 พบว่า ค่าพารามิเตอร์ความยากของของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ในปีการศึกษา 2556 มีค่าความยากน้อยสุดเท่ากับ .02 และสูงสุดเท่ากับ 7.60 ปีการศึกษา 2557 มีค่าความยากน้อยสุดเท่ากับ .01 และสูงสุดเท่ากับ 8.32 และปีการศึกษา 2558 มีค่าความยากน้อยสุดเท่ากับ .02 และสูงสุดเท่ากับ 4.95

ค่าพารามิเตอร์อำนาจจำแนกของของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ในปีการศึกษา 2556 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ .01 และสูงสุดเท่ากับ 1.03 ปีการศึกษา 2557 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ .07 และสูงสุดเท่ากับ .99 และปีการศึกษา 2558 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ .12 และสูงสุดเท่ากับ .95

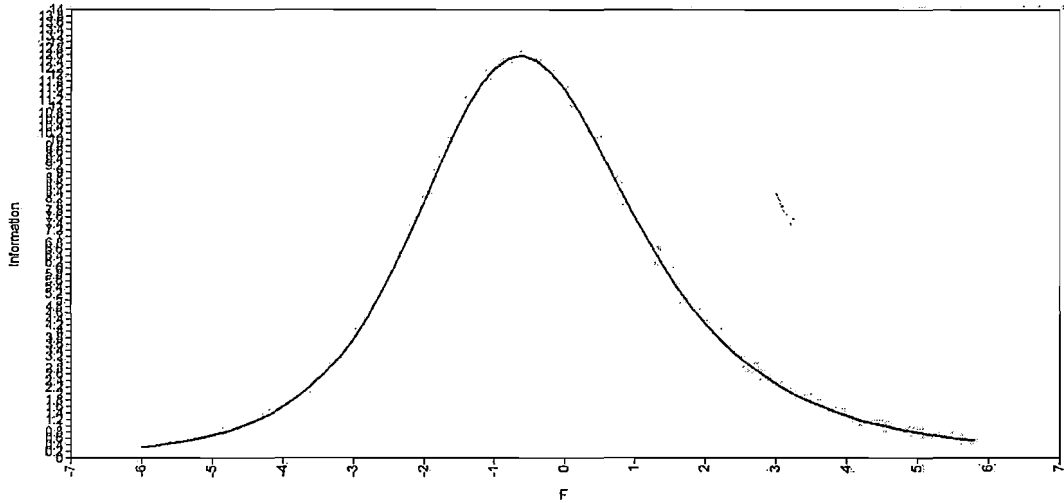
ตาราง 12 ผลการประมาณค่าระดับความสามารถของผู้สอบที่ได้ทำแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 - 2558

ปีการศึกษา	ผลการประมาณค่าระดับความสามารถของผู้สอบ		
	<i>n</i>	<i>Min</i>	<i>Max</i>
2556	414,589	-3.61	3.40
2557	430,617	-3.50	3.33
2558	423,431	-3.73	3.41

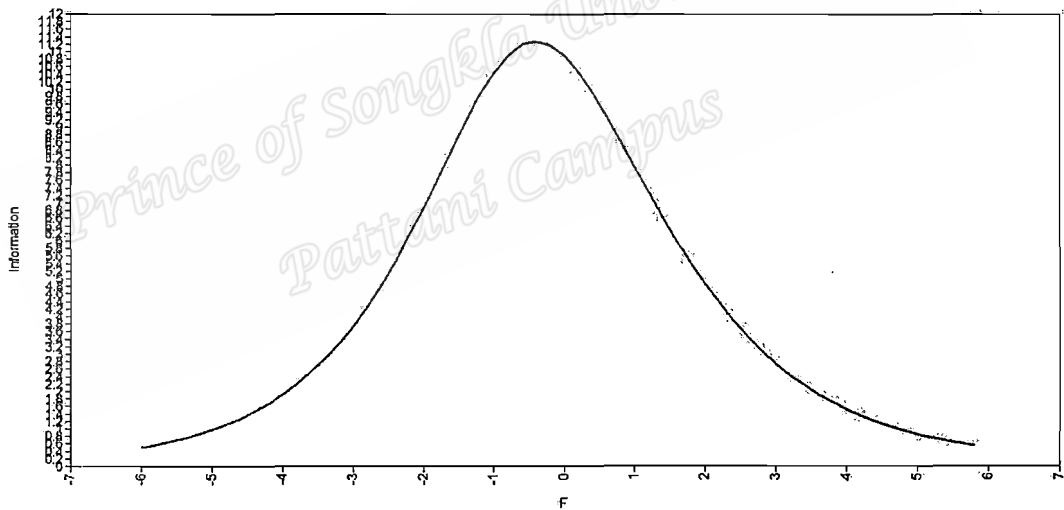
จากตาราง 12 พบว่า ค่าระดับความสามารถของผู้สอบที่ได้ทำแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 มีจำนวนผู้สอบจำนวน 414,589 คน มีค่าความสามารถของผู้สอบน้อยสุด -3.61 และสูงสุดเท่ากับ 3.40 ปีการศึกษา 2557 มีจำนวนผู้สอบจำนวน 430,617 คน มีค่าความสามารถของผู้สอบน้อยสุด -3.50 และสูงสุดเท่ากับ 3.33 และปีการศึกษา 2558 มีจำนวนผู้สอบจำนวน 423,431 คน มีค่าความสามารถของผู้สอบน้อยสุด -3.73 และสูงสุดเท่ากับ 3.41



ภาพประกอบ 10 กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556



ภาพประกอบ 11 กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2557



ภาพประกอบ 12 กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2558

ตาราง 13 ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558

ข้อสอบ	ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	b_i	a_i	b_i	a_i	b_i	a_i
1	.23	.75	3.85	.07	.26	.66
2	1.20	.40	1.15	.47	.57	.60
3	.25	.54	1.46	.50	.30	.54
4	.70	.69	2.14	.33	.18	.57
5	.87	.56	2.67	.23	1.39	.36
6	.35	.63	4.76	.08	.25	.71
7	1.15	.44	1.20	.45	.63	.56
8	.23	.54	1.38	.52	.34	.58
9	.74	.69	2.22	.31	.23	.60
10	.87	.58	3.14	.20	1.46	.37
11	.42	.81	2.89	.27	.43	.48
12	1.62	.34	2.53	.34	.50	.56
13	2.55	.29	6.07	.15	.57	.44
14	.20	.74	2.52	.35	.81	.51
15	1.16	.07	2.79	.27	1.01	.49
16	2.27	.36	1.90	.38	2.06	.38
17	1.52	.41	6.41	.16	2.34	.27
18	2.26	.36	.73	.53	4.20	.27
19	1.42	.38	2.37	.25	2.02	.41
20	3.40	.20	1.55	.45	1.85	.37
21	1.79	.35	2.16	.31	1.53	.49
22	1.18	.48	2.09	.39	4.52	.16
23	2.66	.28	2.05	.37	1.93	.40
24	1.48	.42	3.06	.33	3.43	.24
25	5.32	.01	2.22	.24	5.45	.14
26	7.06	.11	1.57	.34	1.57	.45
27	2.24	.34	2.70	.28	3.08	.27
28	2.80	.33	1.55	.46	2.37	.30
29	1.28	.44	2.64	.28	2.26	.33

ตารางที่ 13 (ต่อ)

ข้อสอบ	ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	b_i	a_i	b_i	a_i	b_i	a_i
30	3.11	.20	2.68	.33	1.57	.34
31	2.20	.32	5.53	.15	2.82	.26
32	1.02	.67	1.84	.19	6.81	.03
33	2.16	.33	2.33	.30	1.44	.09
34	1.74	.37	1.06	.42	3.27	.24
35	3.03	.30	1.76	.40	2.25	.30
36	4.82	.19	3.07	.22	1.68	.38
37	4.95	.15	.64	.74	2.54	.29
38	1.04	.59	1.15	.52	1.27	.47
39	.83	.60	1.84	.34	3.70	.18
40	.82	.70	1.04	.47	1.76	.43
41	2.38	.45	3.40	.18	1.58	.51
42	1.85	.34	2.79	.30	.76	.52
43	1.36	.54	2.41	.27	2.80	.23
44	.86	.71	3.39	.23	2.93	.24
45	.80	.38	4.41	.19	.79	.41
46	8.13	.11	2.68	.32	5.09	.00
47	.90	.44	2.30	.20	3.88	.18
48	1.70	.33	2.37	.26	.80	.53
49	2.01	.37	.44	.69	.56	.75
50	3.46	.18	3.13	.26	2.88	.31
51	.70	.99	.15	.89	.14	.74
52	.14	1.04	.00	1.13	1.86	.20
53	.01	1.14	.06	1.03	.02	.50
54	.61	1.13	.05	.94	1.37	.37
55	.40	1.03	.15	.99	.96	.26
56	.06	1.11	.15	1.18	2.48	.39
57	1.32	.41	.06	.98	2.44	.30
58	1.07	.43	.13	1.01	1.83	.44
59	.38	.76	.51	.50	1.02	.47
60	1.27	.47	3.12	.24	3.74	.21

ตารางที่ 13 (ต่อ)

ข้อสอบ	ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	b_i	a_i	b_i	a_i	b_i	a_i
61	1.00	.52	.76	.69	1.60	.41
62	2.38	.18	.57	.95	.85	.49
63	1.66	.45	.69	.62	1.66	.42
64	2.22	.40	.95	.73	1.12	.64
65	3.38	.22	2.14	.35	1.29	.54
66	7.40	.11	2.05	.38	.69	.65
67	2.58	.16	1.39	.41	2.09	.48
68	1.55	.47	1.04	.49	3.93	.22
69	2.17	.29	2.51	.32	.30	.71
70	.88	.50	1.18	.39	1.61	.36
71	2.88	.24	2.04	.32	.75	.69
72	1.78	.25	2.73	.35	3.16	.25
73	1.60	.38	2.08	.37	.92	.55
74	1.00	.40	2.17	.36	2.10	.44
75	1.05	.63	1.43	.40	4.72	.13
76	2.54	.22	1.10	.48	1.25	.44
77	2.50	.28	2.51	.31	1.74	.32
78	.82	.48	1.32	.38	2.02	.29
79	2.44	.25	2.02	.32	2.90	.24
80	1.25	.34	3.06	.30	1.91	.47

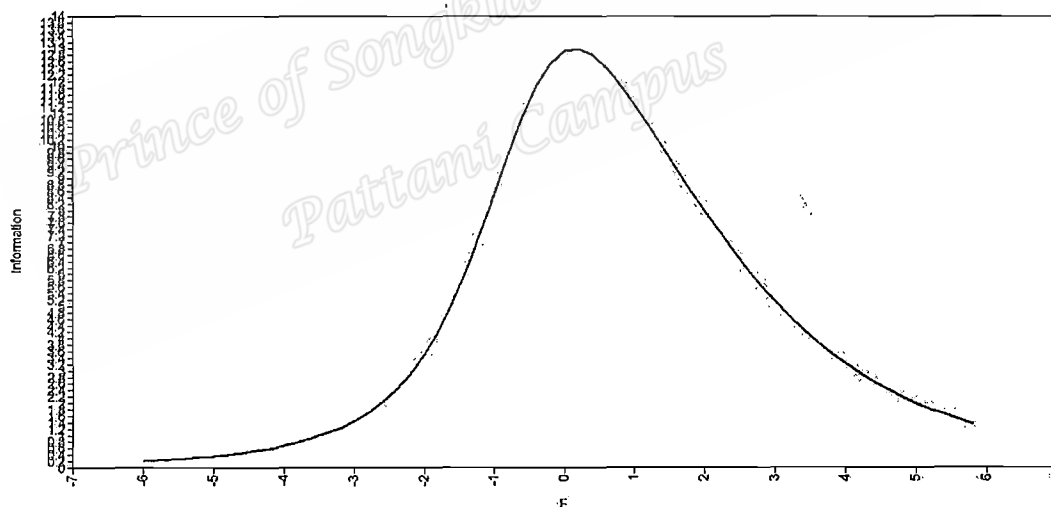
จากตาราง 13 พบว่า ค่าพารามิเตอร์ความยากของของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ในปีการศึกษา 2556 มีค่าความยากน้อยสุดเท่ากับ 18.13 และสูงสุดเท่ากับ .01 ปีการศึกษา 2557 มีค่าความยากน้อยสุดเท่ากับ .00 และสูงสุดเท่ากับ 1.18 และปีการศึกษา 2558 มีค่าความยากน้อยสุดเท่ากับ .02 และสูงสุดเท่ากับ 6.81

ค่าพารามิเตอร์อำนาจจำแนกของของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ในปีการศึกษา 2556 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ .01 และสูงสุดเท่ากับ 1.14 ปีการศึกษา 2557 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ .07 และสูงสุดเท่ากับ 1.18 และปีการศึกษา 2558 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ .00 และสูงสุดเท่ากับ .75

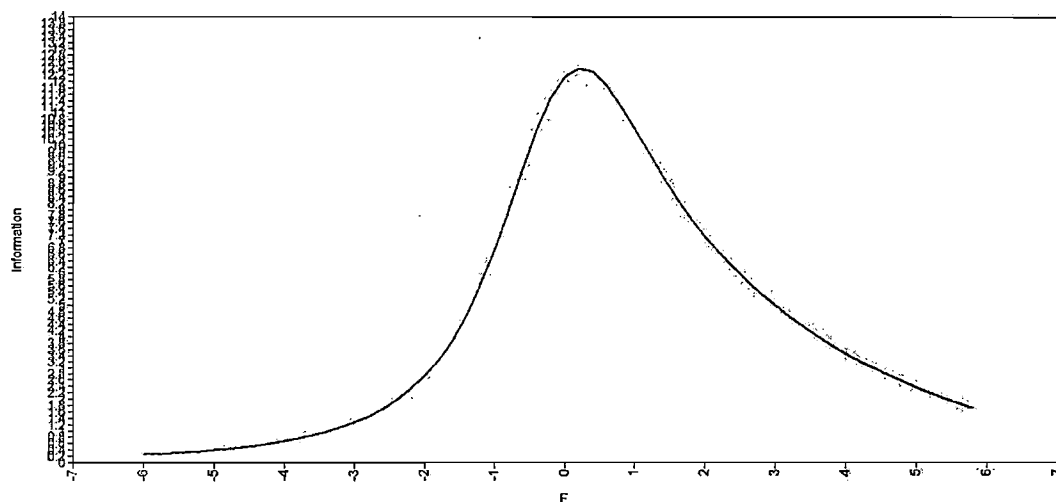
ตาราง 14 ผลการประมาณค่าระดับความสามารถของผู้สอบที่ได้ทำแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558

ปีการศึกษา	ผลการประมาณค่าระดับความสามารถของผู้สอบ		
	<i>n</i>	<i>Min</i>	<i>Max</i>
2556	415,458	-2.76	4.77
2557	431,578	-2.33	5.05
2558	424,187	-2.83	4.97

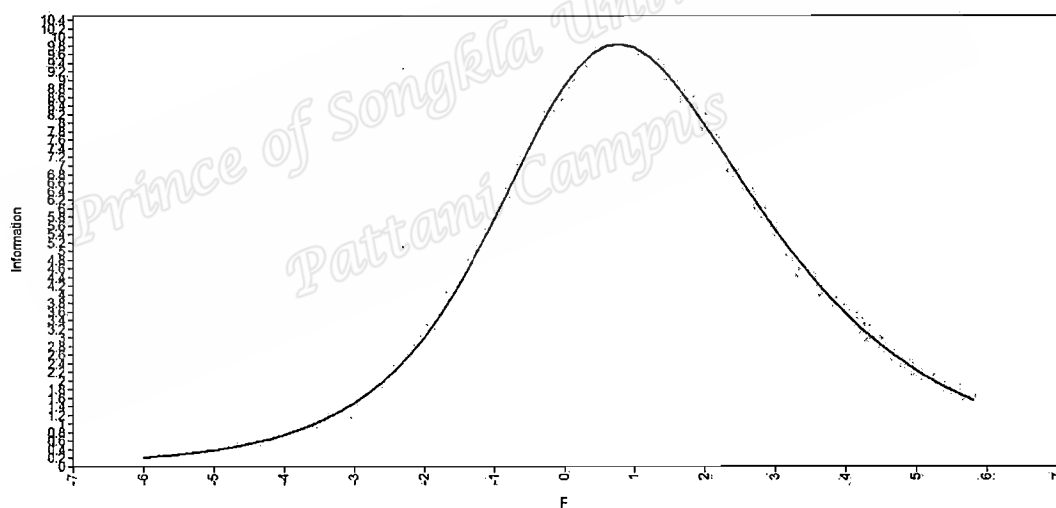
จากตาราง 14 พบว่า ค่าระดับความสามารถของผู้สอบที่ได้ทำแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 มีจำนวนผู้สอบจำนวน 415,458 คน มีค่าความสามารถของผู้สอบน้อยสุด -2.76 และสูงสุดเท่ากับ 4.77 ปีการศึกษา 2557 มีจำนวนผู้สอบจำนวน 431,578 คน มีค่าความสามารถของผู้สอบน้อยสุด -2.33 และสูงสุดเท่ากับ 5.05 และปีการศึกษา 2558 มีจำนวนผู้สอบจำนวน 424,187 คน มีค่าความสามารถของผู้สอบน้อยสุด -2.83 และสูงสุดเท่ากับ 4.97



ภาพประกอบ 13 กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556



ภาพประกอบ 14 กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2557



ภาพประกอบ 15 กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2558

ตาราง 15 ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558

ข้อสอบ	ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	b_i	a_i	b_i	a_i	b_i	a_i
1	2.04	.54	.95	.61	1.13	.40
2	2.41	.39	.98	.61	1.32	.53
3	1.88	.44	1.57	.64	1.63	.43
4	1.69	.44	1.64	.63	1.72	.47
5	2.01	.37	1.29	.77	1.36	.53
6	2.13	.47	1.93	.67	1.35	.60
7	1.11	.59	1.48	.51	2.04	.32
8	1.05	.58	2.33	.42	1.27	.57
9	1.29	.48	2.13	.32	2.64	.32
10	1.27	.47	2.22	.32	2.64	.31
11	2.16	.26	2.20	.29	1.31	.27
12	2.26	.43	1.90	.46	1.02	.67
13	2.22	.45	1.94	.45	1.46	.51
14	3.06	.24	1.05	.45	2.48	.26
15	2.33	.33	1.64	.49	1.59	.35
16	2.32	.34	1.69	.36	1.42	.45
17	2.07	.36	1.30	.40	2.20	.37
18	2.08	.39	2.23	.35	1.37	.38
19	1.03	.92	2.14	.37	2.41	.33
20	1.22	.59	2.12	.31	.05	.59
21	1.75	.40	.78	.53	.37	.59
22	1.32	.57	.48	.77	.41	.32
23	1.52	.51	.85	.46	1.93	.36
24	1.06	.65	.85	.51	.46	.23
25	1.58	.49	.84	.55	2.06	.42
26	2.31	.29	.34	.36	.42	.70
27	2.55	.34	1.71	.57	1.28	.43
28	1.28	.35	1.80	.55	1.47	.43
29	1.49	.40	2.00	.42	1.65	.41
30	.48	.66	.84	.53	8.14	.02

ตารางที่ 15 (ต่อ)

ข้อสอบ	ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	b_i	a_i	b_i	a_i	b_i	a_i
31	1.12	.51	.90	.52	.44	.62
32	1.61	.39	3.67	.22	.46	.55

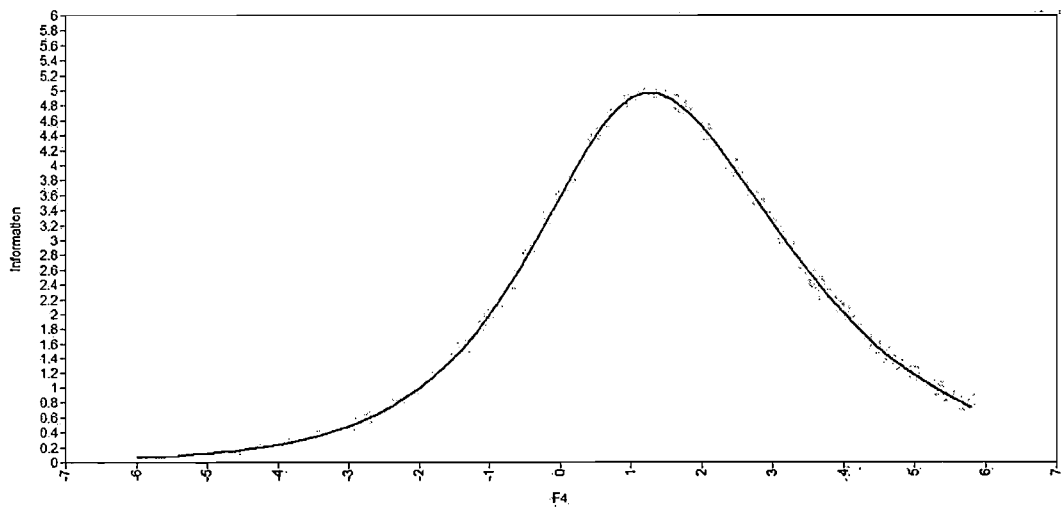
จากตาราง 15 พบว่า ค่าพารามิเตอร์ความยากของของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ในปีการศึกษา 2556 มีค่าความยากน้อยสุดเท่ากับ -0.48 และสูงสุดเท่ากับ 3.06 ปีการศึกษา 2557 มีค่าความยากน้อยสุดเท่ากับ $.34$ และสูงสุดเท่ากับ 3.67 และปีการศึกษา 2558 มีค่าความยากน้อยสุดเท่ากับ $.05$ และสูงสุดเท่ากับ 8.14

ค่าพารามิเตอร์อำนาจจำแนกของของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ในปีการศึกษา 2556 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ $.24$ และสูงสุดเท่ากับ $.92$ ปีการศึกษา 2557 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ $.22$ และสูงสุดเท่ากับ $.77$ และปีการศึกษา 2558 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ $.02$ และสูงสุดเท่ากับ $.70$

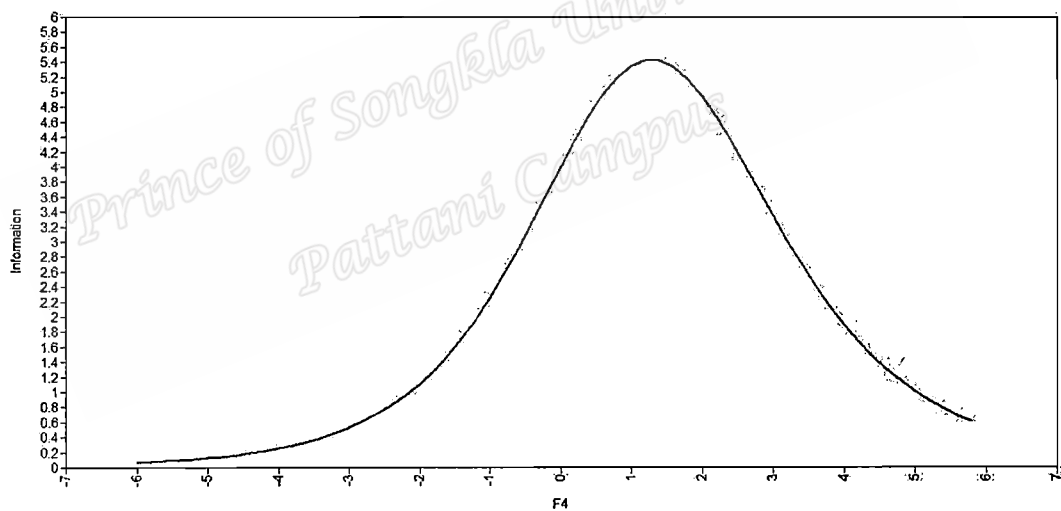
ตาราง 16 ผลการประมาณค่าระดับความสามารถของผู้สอบที่ได้ทำแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558

ปีการศึกษา	ผลการประมาณค่าระดับความสามารถของผู้สอบ		
	n	Min	Max
2556	415,745	-1.81	3.90
2557	431,995	-1.88	3.75
2558	424,430	-2.11	3.65

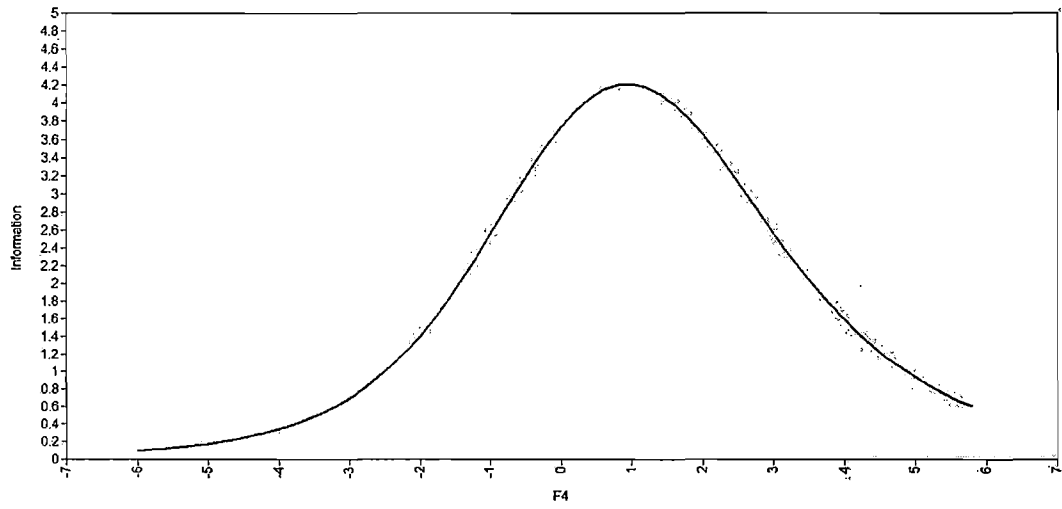
จากตาราง 16 พบว่า ค่าระดับความสามารถของผู้สอบที่ได้ทำแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 มีจำนวนผู้สอบจำนวน 415,745 คน มีค่าความสามารถของผู้สอบน้อยสุด -1.81 และสูงสุดเท่ากับ 3.90 ปีการศึกษา 2557 มีจำนวนผู้สอบจำนวน 431,995 คน มีค่าความสามารถของผู้สอบน้อยสุด -1.88 และสูงสุดเท่ากับ 3.75 และปีการศึกษา 2558 มีจำนวนผู้สอบจำนวน 424,430 คน มีค่าความสามารถของผู้สอบน้อยสุด -2.11 และสูงสุดเท่ากับ 3.65



ภาพประกอบ 16 กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาตินี้พื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556



ภาพประกอบ 17 กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาตินี้พื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2557



ภาพประกอบ 18 กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาตินี้ขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2558

Prince of Songkla University
Pattani Campus

ตาราง 17 ผลการประมาณค่าพารามิเตอร์ความยากของแบบทดสอบการทดสอบทางการศึกษา
ระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558

ข้อสอบ	ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	b_i	a_i	b_i	a_i	b_i	a_i
1	1.87	.30	4.31	.02	-3.18	-.22
2	4.63	.18	-1.03	-.41	28.09	.02
3	1.62	.43	1.17	.37	1.82	.20
4	1.95	.18	3.19	.16	-6.47	-.10
5	1.29	.46	-.13	.84	4.06	.14
6	.52	.59	2.06	.32	33.63	.03
7	1.86	.24	-.30	1.18	-5.41	-.16
8	1.53	.35	4.76	.15	-29.13	-.02
9	.69	.47	-.93	.04	-7.56	-.12
10	.65	.46	-.07	.55	-.23	.16
11	1.82	.35	1.05	.52	.27	.15
12	.91	.41	2.27	.41	2.05	.12
13	2.25	.21	1.90	.26	-.54	-.43
14	1.03	.28	-2.29	-.23	1.03	.39
15	.86	.40	1.06	.35	2.53	.14
16	1.12	.47	.12	.59	33.86	.02
17	.60	.62	-2.91	-.19	1.35	.38
18	1.50	.37	-1.04	-.85	.06	1.04
19	.67	.66	-7.88	-.08	-.27	-.38
20	2.37	.31	.85	.30	3.29	.26
21	2.54	.30	5.69	.14	13.57	.04
22	6.22	.02	4.13	.18	1.63	.05
23	1.84	.38	3.53	.26	3.63	.24
24	1.63	.38	-9.60	-.08	24.92	.04
25	1.81	.43	3.55	.21	27.04	.03
26	3.10	.26	3.96	.22	-1.05	-.22
27	.65	.32	.77	.67	-2.90	-.20
28	1.81	.38	3.03	.20	1.04	.57
29	2.10	.37	-1.43	-.37	-1.70	-.07

ตารางที่ 17 (ต่อ)

ข้อสอบ	ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	b_i	a_i	b_i	a_i	b_i	a_i
30	1.91	.29	-2.11	-.35	-.22	-.35
31	.99	.42	3.63	.23	.26	.25
32	.95	.48	5.80	.06	56.08	.01
33	4.37	.17	6.00	.12	-.65	-.29
34	5.74	.15	3.59	.19	9.76	.09
35	3.52	.22	2.69	.20	-9.25	-.08
36	2.64	.28	.53	-.15	-3.61	-.20
37	1.46	.49	4.10	.15	.56	.69
38	2.96	.08	2.25	.30	-4.49	-.20
39	3.83	.23	-1.63	-.36	-1.90	-.18
40	7.16	.08	-3.01	-.21	5.64	.12
41	3.90	.19	-.17	-1.09	-2.42	-.33
42	1.42	.27	-4.40	-.13	81.29	.01
43	4.00	.13	.07	-1.08	4.76	.19
44	3.61	.22	-3.78	-.22	.97	.28
45	.60	.44	-1.16	-.50	1.81	.42
46	4.65	.11	-.23	-.63	-.68	-.66
47	4.20	.21	-82.66	.00	1.51	.00
48	1.56	.31	2.83	.19	-.03	.27
49	2.58	.29	-2.02	-.33	2.26	.21
50	2.50	.27	2.58	.23	-1.15	-.35
51	5.75	.06	-1.39	-.34	.38	.21
52	1.65	.21	-.34	-.64	-3.29	-.03
53	1.90	.20	13.53	.05	-1.70	-.24
54	2.75	.18	.55	1.13	.14	-.64
55	3.99	.01	-35.65	-.02	-2.48	-.28
56	.25	.52	3.15	.14	2.59	.15
57	1.21	.36	-43.95	-.02	-9.65	-.06
58	3.71	.19	2.00	.13	2.36	.28
59	9.68	.06	-1.88	-.16	-5.21	-.12
60	3.55	.17	1.56	.43	2.01	.34

ตารางที่ 17 (ต่อ)

ข้อสอบ	ผลการประมาณค่าพารามิเตอร์ความยากและอำนาจจำแนกของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	b_i	a_i	b_i	a_i	b_i	a_i
61	5.33	.20	494.63	.00	-2.88	-.23
62	1.84	.04	-4.26	-.15	.40	.56
63	.56	.41	2.16	.31	-4.95	-.14
64	3.12	.17	16.42	.05	2.61	.27
65	1.38	.12	-13.15	-.04	-3.02	-.12
66	1.98	.21	23.68	.03	24.20	.03
67	4.77	.15	-3.29	-.08	-2.63	-.25
68	3.15	.03	1.87	.15	3.98	.16
69	4.88	.13	-1.77	-.38	-2.29	-.29
70	8.90	.12	3.78	.02	2.67	.24
71	.84	.07	3.20	.19	-.35	-.52
72	.60	.35	-2.14	-.35	4.65	.16
73	2.91	.24	-.07	.64	-6.34	-.11
74	.01	.61	6.12	.08	.60	.77
75	4.55	.14	1.76	.55	-.37	-.52
76	3.00	.27	-.08	-.53	-3.03	-.26
77	2.87	.19	-1.81	-.53	.38	.67
78	.41	.05	.50	.64	-.49	-.41
79	.15	.46	-.32	-.24	.90	.64
80	.49	.20	-2.15	-.29	-.92	-.61

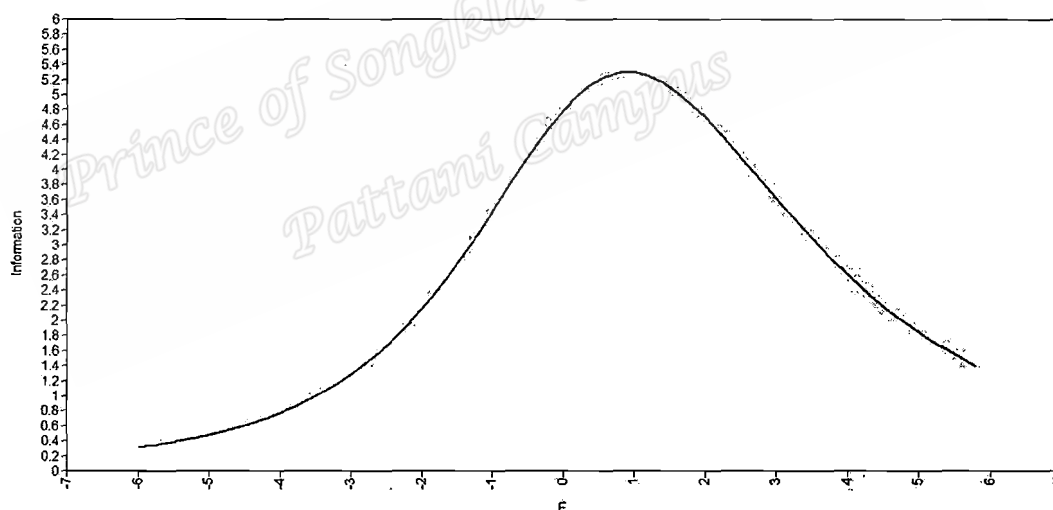
จากตาราง 17 พบว่า ค่าพารามิเตอร์ความยากของของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ในปีการศึกษา 2556 มีค่าความยากน้อยสุดเท่ากับ .01 และสูงสุดเท่ากับ 9.68 ปีการศึกษา 2557 มีค่าความยากน้อยสุดเท่ากับ -82.66 และสูงสุดเท่ากับ 494.63 และปีการศึกษา 2558 มีค่าความยากน้อยสุดเท่ากับ -30.29 และสูงสุดเท่ากับ 81.29

ค่าพารามิเตอร์อำนาจจำแนกของของแบบทดสอบการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ในปีการศึกษา 2556 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ .01 และสูงสุดเท่ากับ .66 ปีการศึกษา 2557 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ -1.09 และสูงสุดเท่ากับ 1.18 และปีการศึกษา 2558 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ -.66 และสูงสุดเท่ากับ 1.04

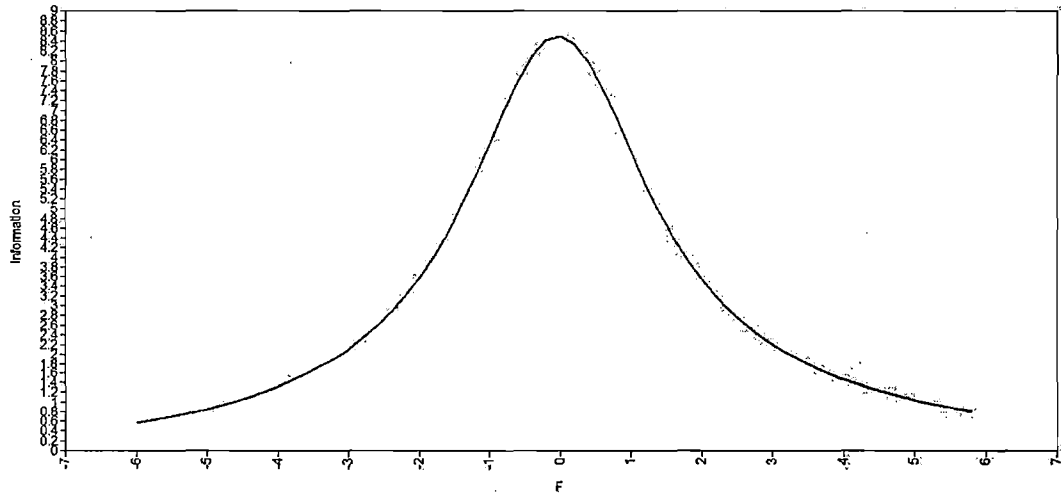
ตาราง 18 ผลการประมาณค่าระดับความสามารถของผู้สอบที่ได้ทำแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558

ปีการศึกษา	ผลการประมาณค่าระดับความสามารถของผู้สอบ		
	<i>n</i>	<i>Min</i>	<i>Max</i>
2556	414,588	-2.67	4.75
2557	430,607	-2.74	2.85
2558	423,495	-2.54	2.64

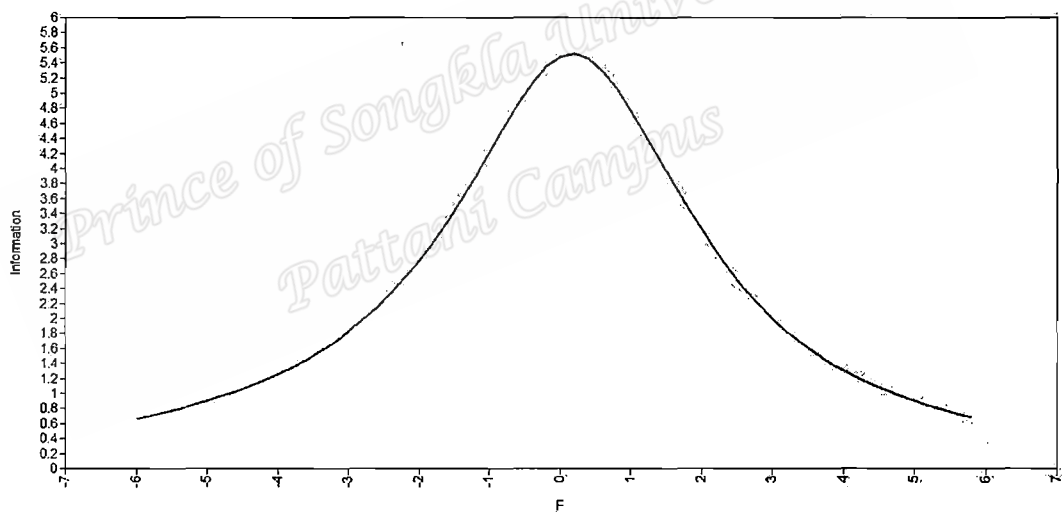
จากตาราง 18 พบว่า ค่าระดับความสามารถของผู้สอบที่ได้ทำแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 มีจำนวนผู้สอบจำนวน 414,588 คน มีค่าความสามารถของผู้สอบน้อยสุด -2.67 และสูงสุดเท่ากับ 4.75 ปีการศึกษา 2557 มีจำนวนผู้สอบจำนวน 430,607 คน มีค่าความสามารถของผู้สอบน้อยสุด -2.74 และสูงสุดเท่ากับ 2.85 และปีการศึกษา 2558 มีจำนวนผู้สอบจำนวน 423,495 คน มีค่าความสามารถของผู้สอบน้อยสุด -2.54 และสูงสุดเท่ากับ 2.64



ภาพประกอบ 19 กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556



ภาพประกอบ 20 กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2557



ภาพประกอบ 21 กราฟแสดงฟังก์ชันสารสนเทศ (Test Information) ของแบบทดสอบการทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2558

ตอนที่ 3 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบจำแนกตามเพศและที่ตั้งของสถานศึกษา

การศึกษาการทำหน้าที่ต่างกันของข้อสอบในครั้งนี้ได้จากการตอบข้อสอบของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ปีการศึกษา 2556 - 2558 โดยที่รายวิชาภาษาไทย ภาษาอังกฤษ คณิตศาสตร์ และวิทยาศาสตร์ มีจำนวนแบบทดสอบรายวิชาละ 70, 80, 32 และ 80 ตามลำดับ ผู้วิจัยได้ใช้วิธี MIMIC โดยประยุกต์ใช้โปรแกรม Mplus โดยจำแนกตามเพศและที่ตั้งของสถานศึกษา ซึ่งมีการวิเคราะห์ด้วยวิธีดังนี้

ตาราง 19 ผลการวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 - 2558 จำแนกตามเพศและที่ตั้งของสถานศึกษา

ตัวแปร	ปีการศึกษา	การวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC					
		χ^2	<i>df</i>	<i>CFI</i>	<i>TLI</i>	<i>RMSEA</i>	<i>WRMR</i>
เพศ	2556	-	2,414	.51	.48	.10	28.07
	2557	847,303.82	2,414	.80	.79	.05	16.12
	2558	826,542.01	2,414	.40	.37	.07	17.22
ที่ตั้งของสถานศึกษา	2556	-	2414	.32	.28	.08	28.33
	2557	688,252.20	2414	.67	.65	.05	16.22
	2558	-	2414	.37	.33	.08	26.72

หมายเหตุ ค่าไค-สแควร์ (Chi-square) ไม่แสดงค่าในการวิเคราะห์

จากตาราง 19 พบว่า ผลการวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC รายวิชาภาษาไทย จำแนกตามเพศ ในปีการศึกษา 2556 พิจารณาดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .51 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .48 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*) มีค่าเท่ากับ .10 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*WRMR*) มีค่าเท่ากับ 28.07

ปีการศึกษา 2557 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 847,303.82 องศาอิสระ (*df*) มีค่าเท่ากับ 2,414 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .80 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .79 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*) มีค่าเท่ากับ .05 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*WRMR*) มีค่าเท่ากับ 16.12

ปีการศึกษา 2558 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 826,542.01 องศาอิสระ (df) มีค่าเท่ากับ 2,414 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .40 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .37 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*) มีค่าเท่ากับ .07 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*WRMR*) มีค่าเท่ากับ 17.22

ในตัวแปรที่ตั้งของสถานศึกษา ปีการศึกษา 2556 พิจารณาดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .32 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .28 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*) มีค่าเท่ากับ .08 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*WRMR*) มีค่าเท่ากับ 28.33

ปีการศึกษา 2557 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 688,252.20 องศาอิสระ (df) มีค่าเท่ากับ 2414 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .67 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .65 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*) มีค่าเท่ากับ .05 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*WRMR*) มีค่าเท่ากับ 16.22

ปีการศึกษา 2558 พิจารณาจากดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .37 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .33 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*) มีค่าเท่ากับ .08 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*WRMR*) มีค่าเท่ากับ 26.72

ตาราง 20 ผลการวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 - 2558 จำแนกตามเพศและที่ตั้งของสถานศึกษา

ตัวแปร	ปีการศึกษา	การวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC					
		χ^2	<i>df</i>	<i>CFI</i>	<i>TLI</i>	<i>RMSEA</i>	<i>WRMR</i>
เพศ	2556	301,012.23	3,159	.91	.91	.03	9.42
	2557	300,528.83	3,159	.91	.90	.04	9.01
	2558	300,967.26	3,159	.90	.90	.03	8.47
ที่ตั้งของสถานศึกษา	2556	375,506.70	3,159	.91	.91	.03	9.38
	2557	301,498.53	3,159	.89	.90	.02	9.01
	2558	278,391.68	3,159	.72	.70	.03	8.69

จากตาราง 20 พบว่า ผลการวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC รายวิชาภาษาอังกฤษ จำแนกตามเพศ ในปีการศึกษา 2556 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 301,012.23 องศาอิสระ (*df*) มีค่าเท่ากับ 3,159 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .91 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .91 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*) มีค่าเท่ากับ .03 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*WRMR*) มีค่าเท่ากับ 9.42

ปีการศึกษา 2557 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 300,528.83 องศาอิสระ (*df*) มีค่าเท่ากับ 3,159 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .91 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .90 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*) มีค่าเท่ากับ .04 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*WRMR*) มีค่าเท่ากับ 9.01

ปีการศึกษา 2558 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 300,967.26 องศาอิสระ (*df*) มีค่าเท่ากับ 3,159 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .90 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .90 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*) มีค่าเท่ากับ .03 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*WRMR*) มีค่าเท่ากับ 8.47

ในตัวแปรที่ตั้งของสถานศึกษา ปีการศึกษา 2556 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 375,506.70 องศาอิสระ (*df*) มีค่าเท่ากับ 3,159 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .91 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .91 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*)

มีค่าเท่ากับ .03 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอดี (WRMR) มีค่าเท่ากับ 9.38

ปีการศึกษา 2557 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 301,498.53 องศาอิสระ (df) มีค่าเท่ากับ 3,159 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (CFI) มีค่าเท่ากับ .89 ดัชนี The Tucker Lewis Index (TLI) มีค่าเท่ากับ .90 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอดี (RMSEA) มีค่าเท่ากับ .02 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอดี (WRMR) มีค่าเท่ากับ 9.01

ปีการศึกษา 2558 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 278,391.68 องศาอิสระ (df) มีค่าเท่ากับ 3,159 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (CFI) มีค่าเท่ากับ .72 ดัชนี The Tucker Lewis Index (TLI) มีค่าเท่ากับ .70 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอดี (RMSEA) มีค่าเท่ากับ .03 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอดี (WRMR) มีค่าเท่ากับ 8.69

Prince of Songkla University
Pattani Campus

ตาราง 21 ผลการวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC ของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 - 2558 จำแนกตามเพศและที่ตั้งของสถานศึกษา

ตัวแปร	ปีการศึกษา	การวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC					
		χ^2	<i>df</i>	<i>CFI</i>	<i>TLI</i>	<i>RMSEA</i>	<i>WRMR</i>
เพศ	2556	40,060.70	495	.94	.94	.02	6.95
	2557	65,493.95	495	.92	.91	.03	8.89
	2558	80,855.94	495	.88	.86	.04	10.28
ที่ตั้งของสถานศึกษา	2556	35,316.05	495	.87	.85	.02	6.93
	2557	55,554.26	495	.83	.81	.03	8.80
	2558	72,351.26	495	.77	.74	.03	10.35

จากตาราง 21 พบว่า ผลการวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC รายวิชาคณิตศาสตร์ จำแนกตามเพศ ในปีการศึกษา 2556 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 40,060.70 องศาอิสระ (*df*) มีค่าเท่ากับ 495 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .94 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .94 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*) มีค่าเท่ากับ .02 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*WRMR*) มีค่าเท่ากับ 6.95

ปีการศึกษา 2557 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 65,493.95 องศาอิสระ (*df*) มีค่าเท่ากับ 495 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .92 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .91 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*) มีค่าเท่ากับ .03 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*WRMR*) มีค่าเท่ากับ 8.89

ปีการศึกษา 2558 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 80,855.94 องศาอิสระ (*df*) มีค่าเท่ากับ 495 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .88 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .86 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*) มีค่าเท่ากับ .04 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*WRMR*) มีค่าเท่ากับ 10.28

ในตัวแปรที่ตั้งของสถานศึกษา ปีการศึกษา 2556 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 35,316.05 องศาอิสระ (*df*) มีค่าเท่ากับ 496 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .87 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .85 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*)

มีค่าเท่ากับ .02 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พหุพหุ (WRMR) มีค่าเท่ากับ 6.93

ปีการศึกษา 2557 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 55,554.26 องศาอิสระ (df) มีค่าเท่ากับ 495 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (CFI) มีค่าเท่ากับ .83 ดัชนี The Tucker Lewis Index (TLI) มีค่าเท่ากับ .81 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พหุพหุ (RMSEA) มีค่าเท่ากับ .03 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พหุพหุ (WRMR) มีค่าเท่ากับ 8.80

ปีการศึกษา 2558 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 72,351.26 องศาอิสระ (df) มีค่าเท่ากับ 495 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (CFI) มีค่าเท่ากับ .77 ดัชนี The Tucker Lewis Index (TLI) มีค่าเท่ากับ .74 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พหุพหุ (RMSEA) มีค่าเท่ากับ .03 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พหุพหุ (WRMR) มีค่าเท่ากับ 10.35

Prince of Songkla University
Pattani Campus

ตาราง 22 ผลการวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC ของแบบทดสอบการทดสอบทางการศึกษาระดับชาตินิยมพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 - 2558 จำแนกตามเพศและที่ตั้งของสถานศึกษา

ตัวแปร	ปีการศึกษา	การวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC					
		χ^2	<i>df</i>	<i>CFI</i>	<i>TLI</i>	<i>RMSEA</i>	<i>WRMR</i>
เพศ	2556	874,424.61	3,159	.51	.48	.05	15.02
	2557	795,458.00	3,159	.52	.51	.04	14.67
	2558	700,432.98	3,159	.58	.56	.04	13.59
ที่ตั้งของสถานศึกษา	2556	707,633.98	3,159	.28	.24	.04	15.07
	2557	707,633.98	3,159	.27	.28	.03	14.16
	2558	491,223.17	3,159	.54	.51	.03	13.51

จากตาราง 22 พบว่า ผลการวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC รายวิชาวิทยาศาสตร์ จำแนกตามเพศ ในปีการศึกษา 2556 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 874,424.61 องศาอิสระ (*df*) มีค่าเท่ากับ 3,159 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .51 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .48 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*) มีค่าเท่ากับ .05 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*WRMR*) มีค่าเท่ากับ 15.02

ปีการศึกษา 2557 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 795,458.00 องศาอิสระ (*df*) มีค่าเท่ากับ 3,159 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .52 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .51 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*) มีค่าเท่ากับ .04 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*WRMR*) มีค่าเท่ากับ 14.67

ปีการศึกษา 2558 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 700,432.98 องศาอิสระ (*df*) มีค่าเท่ากับ 3,159 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .58 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .56 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*) มีค่าเท่ากับ .04 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*WRMR*) มีค่าเท่ากับ 13.59

ในตัวแปรที่ตั้งของสถานศึกษา ปีการศึกษา 2556 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 707,633.98 องศาอิสระ (*df*) มีค่าเท่ากับ 3,159 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (*CFI*) มีค่าเท่ากับ .28 ดัชนี The Tucker Lewis Index (*TLI*) มีค่าเท่ากับ .24 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (*RMSEA*)

มีค่าเท่ากับ .04 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พหุพหุ (WRMR) มีค่าเท่ากับ 15.07

ปีการศึกษา 2557 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 547,863.21 องศาอิสระ (df) มีค่าเท่ากับ 3,159 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (CFI) มีค่าเท่ากับ .35 ดัชนี The Tucker Lewis Index (TLI) มีค่าเท่ากับ .39 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พหุพหุ (RMSEA) มีค่าเท่ากับ .02 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พหุพหุ (WRMR) มีค่าเท่ากับ 13.98

ปีการศึกษา 2558 พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 491,223.17 องศาอิสระ (df) มีค่าเท่ากับ 3,159 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (CFI) มีค่าเท่ากับ .54 ดัชนี The Tucker Lewis Index (TLI) มีค่าเท่ากับ .51 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พหุพหุ (RMSEA) มีค่าเท่ากับ .03 ดัชนี Weighted Root Mean Square Residual วัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พหุพหุ (WRMR) มีค่าเท่ากับ 13.51

Prince of Songkla University
Pattani Campus

ตาราง 23 ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศ

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
1	.11*	.00	.03*	.05	-.17*	.00
2	.08*	.00	.18*	.00	-.12*	.00
3	.23*	.00	.14*	.00	-.02	.36
4	.17*	.00	-.03*	.00	-.18*	.00
5	.17*	.00	.03*	.01	-.28*	.00
6	.32*	.00	-.02*	.03	-.18*	.00
7	.17*	.00	-.09*	.00	-.21*	.00
8	.16*	.00	-.04*	.00	-.09*	.03
9	.15*	.00	.07*	.00	-.12*	.00
10	.13*	.00	.01	.63	-.12*	.00
11	.20*	.00	-.03	.08	-.21*	.00
12	-.02*	.03	.17*	.00	-.18*	.00
13	.25*	.00	.10*	.00	-.09*	.01
14	.23*	.00	-.01	.35	-.24*	.00
15	.22*	.00	.02	.15	-.31*	.00
16	.21*	.00	-.02	.09	-.21*	.00
17	.21*	.00	-.07*	.00	-.21*	.00
18	.23*	.00	-.06*	.00	-.13*	.01
19	.21*	.00	.07*	.00	-.08*	.00
20	.20*	.00	.02*	.02	-.18*	.00
21	-.01	.62	-.02	.06	-.23*	.00
22	.29*	.00	.02	.07	-.17*	.00
23	.25*	.00	.03*	.01	-.24*	.00
24	.11*	.00	.11*	.00	-.13*	.00
25	.14*	.00	.08*	.00	-.18*	.00
26	.11*	.00	.04*	.00	-.21*	.00
27	.08*	.00	.02*	.03	-.42*	.00
28	.23*	.00	.00	.75	-.14*	.00
29	.05*	.00	.06*	.00	-.24*	.00

ตารางที่ 23 (ต่อ)

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
30	.26*	.00	.06*	.00	-.39*	.00
31	.09*	.00	.02*	.04	-.27*	.00
32	.18*	.00	.05*	.00	-.30*	.00
33	.16*	.00	.07*	.00	-.16*	.00
34	-.08*	.00	.03	.07	-.28*	.00
35	.13*	.00	-.28*	.00	-.33*	.00
36	.02	.06	.16*	.00	-.26*	.00
37	.28*	.00	.06*	.00	-.36*	.00
38	.04*	.00	-.07*	.00	-.20*	.00
39	.18*	.00	.03*	.02	-.28*	.00
40	.18*	.00	.01	.39	-.24*	.00
41	.12*	.00	-.05*	.00	-.21*	.00
42	.05*	.00	-.05*	.00	-.22*	.00
43	.18*	.00	-.01	.49	-.05	.19
44	.17*	.00	.08*	.00	-.15*	.00
45	.10*	.00	.08*	.00	-.15*	.00
46	-.06*	.00	.15*	.00	-.24*	.00
47	.16*	.00	.06*	.00	-.28*	.00
48	-.09*	.00	-.07*	.00	-.18*	.00
49	.18*	.00	.02	.10	-.19*	.00
50	-.06*	.00	.02	.06	-.33*	.00
51	.19*	.00	-.04*	.00	-.27*	.00
52	.18*	.00	-.05*	.00	-.28*	.00
53	.03	.06	-.01	.42	-.13*	.00
54	.12*	.00	.08*	.00	-.15*	.00
55	.20*	.00	.09*	.00	.04	.07
56	.05*	.00	.09*	.00	-.33*	.00
57	.07*	.00	.02	.09	-.24*	.00
58	-.05*	.00	.00	.90	-.35*	.00
59	.06*	.00	-.02	.11	-.28*	.00
60	.10*	.00	-.15*	.00	-.44*	.00

ตารางที่ 23 (ต่อ)

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
61	.03*	.00	.14*	.00	-.18*	.00
62	.12*	.00	.08*	.00	-.29*	.00
63	.17*	.00	.10*	.00	-.25*	.00
64	-.02	.14	.02	.13	-.32*	.00
65	.04*	.00	.04*	.00	-.15*	.00
66	.13*	.00	-.12*	.00	-.37*	.00
67	.00	.87	.05*	.00	-.30*	.00
68	.16*	.00	-.02	.20	-.29*	.00
69	.00*	.00	.00	.83	-.31*	.00
70	.13*	.00	.00*	.00	-.26*	.00

หมายเหตุ * นัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 23 พบว่า ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย โดยวิธี MIMIC จำแนกตามเพศ ปีการศึกษา 2556 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ 65 จำนวน ข้อ ได้แก่ 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 65, 66, 68, 69, 70 ปีการศึกษา 2557 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบจำนวน 49 ข้อ ได้แก่ 2, 3, 4, 5, 6, 7, 8, 9, 12, 13, 17, 18, 19, 20, 23, 24, 25, 26, 27, 29, 30, 31, 32, 33, 35, 36, 37, 38, 39, 41, 42, 44, 45, 46, 47, 48, 51, 52, 54, 55, 56, 60, 61, 62, 63, 65, 66, 67, 70 ปีการศึกษา 2558 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ 67 จำนวน ข้อ ได้แก่ 1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70

ตาราง 24 ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศ

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
1	-.15*	.00	-.14*	.00	-.14*	.00
2	-.04*	.00	.02	.08	-.18*	.00
3	-.09*	.00	.13*	.00	-.11*	.00
4	-.07*	.00	.03	.08	-.16*	.00
5	-.05*	.00	-.03	.07	-.11*	.00
6	-.14*	.00	-.17	.09	-.15*	.00
7	-.03*	.00	.02	.10	-.13*	.00
8	-.12*	.00	.12*	.00	-.15*	.00
9	-.06*	.00	.03	.23	-.21*	.00
10	-.06*	.00	.02	.21	-.12*	.00
11	-.12*	.00	-.01	.06	-.23*	.00
12	-.07*	.00	.11*	.01	-.19*	.00
13	-.03*	.00	.00*	.01	-.16*	.00
14	-.14*	.00	.11*	.02	-.13*	.00
15	.01	.40	-.01	.15	-.11*	.00
16	.09*	.00	.02	.13	-.10*	.00
17	-.02	.18	.03	.14	-.07*	.00
18	.04*	.00	.01	.12	.06*	.00
19	-.03*	.01	-.06*	.02	-.10*	.00
20	.04*	.00	.01	.10	-.08*	.00
21	-.02	.07	.05*	.00	-.16*	.00
22	-.03*	.02	.06*	.00	-.06*	.00
23	.08*	.00	-.04*	.00	-.06*	.00
24	-.06*	.00	.04*	.00	-.11*	.00
25	.09*	.00	-.07*	.00	.01	.60
26	.00	.69	-.01	.23	-.13*	.00
27	.01	.34	.00	.36	-.06*	.00
28	.03*	.00	.06*	.02	-.09*	.00
29	-.10*	.00	.05*	.03	-.08*	.00

ตารางที่ 24 (ต่อ)

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
30	.02	.09	.02	.21	-.03*	.00
31	.01	.37	.01	.26	-.08*	.00
32	-.03	.07	-.07*	.03	-.11*	.00
33	-.03*	.00	.02	.14	.11*	.00
34	-.05*	.00	-.01	.23	.04*	.00
35	.06*	.00	.02	.23	-.05*	.00
36	.06*	.00	-.02	.21	-.06*	.00
37	-.05*	.00	-.07*	.02	.00	.71
38	-.02	.13	.03*	.03	-.11*	.00
39	-.06*	.00	.06*	.01	-.08*	.00
40	-.02	.28	-.01	.32	-.09*	.00
41	.08*	.00	-.16*	.00	-.09*	.00
42	.04*	.00	.00	.93	-.08*	.00
43	.09*	.00	-.12*	.00	-.04*	.00
44	.11*	.00	-.09*	.00	-.08*	.00
45	.08*	.00	.18*	.00	-.16*	.00
46	-.01	.18	-.12*	.00	.08*	.00
47	.05*	.00	-.05*	.00	-.04*	.00
48	.00	.81	-.08*	.00	-.08*	.00
49	.06*	.00	-.22*	.00	-.13*	.00
50	-.04*	.00	.02*	.05	.02*	.03
51	-.07*	.00	-.18*	.00	-.29*	.00
52	.03*	.05	-.32*	.00	-.09*	.00
53	.09*	.00	-.31*	.00	-.23*	.00
54	-.08*	.00	-.24*	.00	-.08*	.00
55	.02	.32	-.21*	.00	-.17*	.00
56	.09*	.00	-.35*	.00	.09*	.00
57	.02	.16	-.27*	.00	-.04*	.00
58	.01	.63	-.27*	.00	-.10*	.00
59	-.07*	.00	-.22*	.00	-.11*	.00
60	.02	.10	-.01	.24	.07*	.00

ตารางที่ 24 (ต่อ)

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
61	.02	.19	-.15*	.00	.00	.80
62	-.05	.00	-.11*	.00	-.13*	.00
63	.10	.00	-.16*	.00	-.03*	.02
64	.06	.00	-.14*	.00	-.10*	.00
65	.04	.00	-.07*	.00	-.11*	.00
66	.02	.01	-.04*	.00	-.17*	.00
67	-.02	.10	-.12*	.00	-.01	.30
68	.09	.00	-.14*	.00	.04*	.00
69	.00	.94	-.02	.12	-.15*	.00
70	-.01	.35	-.13*	.00	-.04*	.00
71	.03	.01	-.07*	.00	-.08*	.00
72	-.03	.01	-.01	.48	-.02*	.02
73	.02	.05	-.08*	.00	-.13*	.00
74	.01	.23	-.02	.14	.03*	.01
75	.06	.00	-.10*	.00	-.05*	.00
76	-.03	.00	-.13*	.00	-.07*	.00
77	.01	.22	.02	.13	-.09*	.00
78	-.03	.01	-.12*	.00	.02	.13
79	.02	.08	-.07*	.00	-.07*	.00
80	.00	.00	.00*	.00	.00*	.00

หมายเหตุ * นัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 24 พบว่า ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ โดยวิธี MIMIC จำแนกตามเพศ ปีการศึกษา 2556 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 56 ข้อ ได้แก่ 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 18, 19, 20, 22, 23, 24, 25, 28, 29, 33, 34, 35, 36, 37, 39, 41, 42, 43, 44, 45, 47, 49, 50, 51, 53, 54, 56, 59, 62, 63, 64, 65, 66, 68, 71, 72, 73, 75, 76, 78, 80 ปีการศึกษา 2557 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 52 ข้อ ได้แก่ 1, 3, 8, 12, 13, 14, 19, 21, 22, 23, 24, 25, 28, 29, 32, 37, 38, 39, 41, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 61, 62, 63, 64, 65, 66, 67, 68, 70, 71, 73, 75, 76, 78, 79,

80 ปีการศึกษา 2558 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 75 ข้อ ได้แก่ 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 62, 63, 64, 65, 66, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 79, 80

ตาราง 25 ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศ

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
1	.09*	.00	-.20*	.00	.02	.09
2	.04*	.00	-.23*	.00	.03*	.00
3	-.02	.11	-.07*	.00	.05*	.00
4	.05*	.00	-.09*	.00	.05*	.00
5	.09*	.00	-.08*	.00	.08*	.00
6	.12*	.00	-.05*	.03	.07*	.00
7	.02	.23	-.07*	.00	.08*	.00
8	.01	.28	-.06*	.00	.04*	.00
9	-.01	.49	-.08*	.00	.05*	.00
10	-.01	.29	-.08*	.00	.06*	.00
11	.04*	.00	-.15*	.00	-.05*	.00
12	.02	.07	-.09*	.00	.01	.37
13	.02	.09	-.07*	.00	.10*	.00
14	.07*	.00	-.08*	.00	.11*	.00
15	.03*	.00	-.05*	.01	.06*	.00
16	.07*	.00	-.13*	.00	.14*	.00
17	.06*	.00	-.08*	.00	.11*	.00
18	.03*	.01	-.05*	.00	.08*	.00
19	.01	.44	-.07*	.00	.08*	.00
20	-.01	.72	-.03*	.03	.03*	.01
21	.02	.09	-.13*	.00	-.07*	.00
22	.02	.21	-.14*	.00	-.01	.11
23	.05*	.00	-.08*	.00	.11*	.00
24	.09*	.00	-.06*	.00	-.07*	.00

ตารางที่ 25 (ต่อ)

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
25	.05*	.00	-.09*	.00	.10*	.00
26	-.04*	.00	-.07*	.00	-.04*	.00
27	-.02	.13	-.04	.07	.03*	.00
28	-.11*	.00	-.04	.06	-.02	.10
29	.00	1.00	.03	.11	.02*	.03
30	-.10*	.00	-.15*	.00	.00	.80
31	-.02*	.05	-.17*	.00	-.03*	.00
32	.00*	.00	.00*	.00	.00*	.00

หมายเหตุ * นัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 25 พบว่า ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ โดยวิธี MIMIC จำแนกตามเพศ ปีการศึกษา 2556 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 19 ข้อ ได้แก่ 1, 2, 4, 5, 6, 11, 14, 15, 16, 17, 18, 23, 24, 25, 26, 28, 30, 31, 32 ปีการศึกษา 2557 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 29 ข้อ ได้แก่ 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 30, 31, 32 ปีการศึกษา 2558 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 27 ข้อ ได้แก่ 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 23, 24, 25, 26, 27, 29, 31, 32

ตาราง 26 ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศ

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
1	.22*	.00	3.00*	.01	-.09*	.00
2	.22*	.00	-1.49	.06	.01	.11
3	.25*	.00	8.03	.06	-.07*	.00
4	.13*	.00	5.05	.06	-.06*	.00
5	.24*	.00	11.73	.06	.02	.08
6	.19*	.00	7.12	.06	.00	.95
7	.15*	.00	14.40	.06	.02*	.03
8	.20*	.00	6.19	.06	-.09*	.00
9	.24*	.00	4.42	.06	.03*	.00
10	.13*	.00	11.22	.06	-.09*	.00
11	.22*	.00	1.59	.06	.01	.43
12	.27*	.00	8.93	.06	-.11*	.00
13	.12*	.00	6.40	.06	-.11*	.00
14	.17*	.00	-2.95	.06	.11*	.00
15	.31*	.00	7.75	.06	-.04*	.00
16	.29*	.00	11.27	.06	-.03*	.01
17	.28*	.00	-.72*	.01	.09*	.00
18	.22*	.00	-8.03	.06	.17*	.00
19	.38*	.00	2.85	.06	-.21*	.00
20	.19*	.00	9.88	.06	.06*	.00
21	.19*	.00	3.25	.06	.03*	.00
22	.00	.75	5.04	.06	.04*	.00
23	.26*	.00	5.11	.06	.09*	.00
24	.38*	.00	-.48*	.01	.06*	.00
25	.38*	.00	7.17	.06	.02*	.01
26	.14*	.00	3.53	.06	-.20*	.00
27	.28*	.00	1.41	.06	-.06*	.00
28	.21*	.00	5.80	.06	.08*	.00
29	.23*	.00	-3.59	.06	-.12*	.00

ตารางที่ 26 (ต่อ)

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
30	.12*	.00	-1.39	.06	-.16*	.00
31	.10*	.00	4.62	.06	.12*	.00
32	.14*	.00	2.42	.06	.01	.19
33	.11*	.00	3.57	.06	-.21*	.00
34	.21*	.00	5.01	.06	.14*	.00
35	.22*	.00	5.76	.06	-.02*	.02
36	.21*	.00	.96	.00	-.07*	.00
37	.32*	.00	5.67	.06	.10*	.00
38	.13*	.00	9.34	.06	-.09*	.00
39	.23*	.00	-2.88	.06	-.05*	.00
40	.09*	.00	-1.36	.06	.04*	.00
41	.23*	.00	.00*	.00	-.06*	.00
42	.24*	.00	-.01*	.01	.05*	.00
43	.17*	.00	.00*	.01	.04*	.00
44	.25*	.00	-.02*	.01	-.03*	.00
45	.31*	.00	-.01*	.00	.00	.67
46	.05*	.00	.00*	.01	-.23*	.00
47	.25*	.00	.00*	.01	-.10*	.00
48	.24*	.00	-.02*	.01	-.02*	.02
49	.27*	.00	.02*	.01	.09*	.00
50	.16*	.00	-.01*	.01	-.11*	.00
51	.12*	.00	.00*	.01	-.02	.06
52	.25*	.00	-.01*	.01	.03*	.00
53	.18*	.00	.00*	.01	-.07*	.00
54	.18*	.00	-.01	1.00	-.30*	.00
55	-.14*	.00	.006*	.01	-.05*	.00
56	.17*	.00	.005*	.01	-.13*	.00
57	.21*	.00	.008*	.01	-.06*	.00
58	.10*	.00	.007*	.01	.13*	.00
59	-.08*	.00	-.01*	.00	-.01	.47
60	.04*	.00	.007*	.01	.21*	.00

ตารางที่ 26 (ต่อ)

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
61	.27*	.00	-.01*	.01	-.01	.49
62	.04*	.00	.00*	.00	.05*	.00
63	.12*	.00	-.01*	.01	-.01	.32
64	-.13*	.00	.002*	.01	.08*	.00
65	.12*	.00	.003*	.01	-.20*	.00
66	.18*	.00	.009*	.01	-.06*	.00
67	.10*	.00	-.01*	.00	-.02	.09
68	-.09*	.00	.008*	.01	.06*	.00
69	.01	.19	.006*	.01	.03*	.00
70	.25*	.00	-.01*	.00	.10*	.00
71	.00	.69	-.01*	.01	-.23*	.00
72	.09*	.00	.02*	.01	.08*	.00
73	.14*	.00	-.02*	.00	.00	.96
74	.01	.39	-.01*	.01	.15*	.00
75	.17*	.00	.008*	.01	-.31*	.00
76	.19*	.00	.007*	.00	.02	.11
77	.07*	.00	.01*	.01	.03*	.00
78	-.03*	.00	.008*	.00	-.16*	.00
79	.00	.88	-.02*	.01	.33*	.00
80	.007*	.00	-.02*	.00	.006*	.00

หมายเหตุ * นัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 26 พบว่า ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ โดยวิธี MIMIC จำแนกตามเพศ ปีการศึกษา 2556 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 75 ข้อ ได้แก่ 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 70, 72, 73, 75, 76, 77, 78, 80 ปีการศึกษา 2557 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 43 ข้อ ได้แก่ 1, 17, 24, 36, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68,

69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80 ปีการศึกษา 2558 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 67 ข้อ ได้แก่ 1, 3, 4, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 46, 47, 48, 49, 50, 52, 53, 54, 55, 56, 57, 58, 60, 62, 64, 65, 66, 68, 69, 70, 71, 72, 74, 75, 77, 78, 79, 80

ตาราง 27 ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามที่ตั้งของสถานศึกษา

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
1	.04*	.00	-.09*	.00	.06*	.00
2	-.06*	.00	.04*	.00	.06*	.00
3	.05*	.00	-.04*	.00	.05*	.00
4	-.01	.54	-.05*	.00	.00	.72
5	-.12*	.00	-.01	.29	-.01	.31
6	.04*	.00	-.09*	.00	.03*	.04
7	-.02*	.04	-.03*	.02	.09*	.00
8	.01	.14	-.05*	.00	.06*	.00
9	-.03*	.00	-.06*	.00	.08*	.00
10	-.03*	.02	-.14*	.00	.00	.82
11	.03*	.00	-.17*	.00	.03*	.02
12	-.07*	.00	.02	.12	.02	.15
13	.02	.06	-.06*	.00	.03*	.00
14	.01	.39	-.09*	.00	-.02	.14
15	-.12*	.00	-.04*	.00	-.03*	.02
16	.05*	.00	-.12*	.00	.02	.06
17	-.06*	.00	-.05*	.00	.09*	.00
18	.01	.43	-.05*	.00	.05*	.00
19	-.04*	.00	-.09*	.00	.08*	.00
20	-.04*	.00	-.15*	.00	-.02	.17
21	-.02	.06	-.03*	.00	.03*	.01
22	-.04*	.00	-.06*	.00	.02	.18

ตารางที่ 27 (ต่อ)

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
23	-.01	.41	-.01	.64	.00	.78
24	.05*	.00	.02	.18	.09*	.00
25	.02	.08	.01	.68	.04*	.00
26	-.04*	.00	-.04*	.00	.05*	.00
27	.01	.27	.01	.38	.05*	.00
28	-.04*	.00	-.06*	.00	.01	.28
29	-.02	.07	.04*	.00	.08*	.00
30	.06*	.00	-.02*	.05	.04*	.00
31	.02	.14	-.02	.14	.04*	.00
32	.01	.44	-.02	.10	.05*	.00
33	-.03*	.01	.05*	.00	.03*	.03
34	-.02	.12	.06*	.00	.03*	.02
35	-.07*	.00	-.18*	.00	.08*	.00
36	-.08*	.00	.08*	.00	.04*	.01
37	.10*	.00	.02	.09	-.01	.37
38	-.08*	.00	-.09*	.00	.00	.96
39	-.05*	.00	.00	.82	.04*	.01
40	.03*	.01	-.02	.18	.06*	.00
41	-.10*	.00	-.07*	.00	.08*	.00
42	-.01	.22	.03*	.03	.06*	.00
43	-.08*	.00	-.08*	.00	.07*	.00
44	.00	.84	.04*	.00	.01	.38
45	.01	.62	.02	.06	.04*	.00
46	-.08*	.00	.07*	.00	.02	.13
47	.04*	.00	.03*	.01	.03*	.01
48	-.10*	.00	-.07*	.00	.08*	.00
49	.01	.22	-.05*	.00	.03*	.02
50	-.04*	.00	.00	.87	.05*	.00
51	-.02*	.05	-.02	.15	.05*	.00
52	.04*	.00	.04*	.00	.05*	.00
53	-.08*	.00	-.05*	.00	.05*	.00

ตารางที่ 27 (ต่อ)

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
54	-.02	.08	.04*	.00	.02	.14
55	.06*	.00	.03*	.02	.04*	.00
56	.03*	.01	.01	.52	-.01	.29
57	-.01	.42	-.05*	.00	.01	.54
58	-.04*	.00	-.08*	.00	.05*	.01
59	.06*	.00	.00	.81	.04*	.01
60	-.01	.27	-.14*	.00	.03*	.04
61	-.04*	.00	-.01	.52	.01	.39
62	-.03*	.00	.00	.99	.00	.81
63	.01	.59	.04*	.00	.00	.96
64	.00	.77	-.02	.08	.00	.85
65	.00	.81	.02*	.02	.06*	.00
66	.00	.96	-.12*	.00	.03*	.04
67	.01	.22	-.09*	.00	.01	.36
68	-.02*	.02	.02	.13	.03*	.04
69	-.11*	.00	-.03*	.03	.02	.14
70	.00*	.00	.00*	.00	.00*	.00

หมายเหตุ * นัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 27 พบว่า ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมพื้นฐาน (O-NET) รายวิชาภาษาไทย โดยวิธี MIMIC จำแนกตามเพศ ปีการศึกษา 2556 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ 45 จำนวน ข้อ ได้แก่ 1, 2, 3, 5, 6, 7, 9, 10, 11, 12, 15, 16, 17, 19, 20, 22, 24, 26, 28, 30, 33, 35, 36, 37, 38, 39, 40, 41, 43, 46, 47, 48, 50, 51, 52, 53, 55, 56, 58, 59, 61, 62, 68, 69, 70 ปีการศึกษา 2557 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 50 ข้อ ได้แก่ 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 26, 28, 29, 30, 33, 34, 35, 36, 38, 41, 42, 43, 44, 46, 47, 48, 49, 52, 53, 54, 55, 57, 58, 60, 63, 65, 66, 67, 69, 70 ปีการศึกษา 2558 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 47 ข้อ ได้แก่ 1, 2, 3, 6, 7, 8, 9, 11, 13, 15, 17, 18, 19, 21, 24, 25, 26, 27, 29, 30, 31, 32, 33, 34, 35, 36, 39, 40, 41, 42, 43, 45, 47, 48, 49, 50, 51, 52, 53, 55, 58, 59, 60, 65, 66, 68, 70

ตาราง 28 ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามที่ตั้งของสถานศึกษา

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
1	.00	.78	.02	.06	.08*	.00
2	-.03*	.00	-.29	.07	.01	.64
3	.01	.51	-.37	.08	.04*	.00
4	-.02	.16	-.26	.13	.05*	.00
5	-.03*	.01	-.05*	.01	-.04*	.00
6	-.01	.37	.07*	.01	.07*	.00
7	.01	.68	-.26	.25	-.03*	.03
8	-.02*	.04	-.40	.21	.07*	.00
9	-.01	.59	-.24	.36	.07*	.00
10	-.02	.17	-.09*	.05	-.06*	.00
11	-.01	.66	-.21	.14	.11*	.00
12	-.01	.37	-.30	.12	-.03*	.03
13	-.01	.37	-.24	.13	.04*	.00
14	-.02	.18	-.30	.36	-.03*	.04
15	.01	.43	-.21	.25	.00	.80
16	.01	.52	-.22	.23	-.02	.12
17	-.01	.49	-.15*	.01	-.06*	.00
18	.00	.74	-.30	.36	-.11*	.00
19	-.02	.09	-.06*	.00	-.04*	.00
20	-.01	.32	-.17	.36	.02*	.05
21	.00	.96	-.25	.32	.06*	.00
22	-.01	.28	-.31	.32	.00	.76
23	.00	.95	-.21*	.01	-.09*	.00
24	.00	.98	-.16*	.02	-.05*	.00
25	.01	.44	-.05*	.01	-.05*	.00
26	-.01	.17	-.10*	.03	-.02	.06
27	-.02	.18	-.11*	.05	-.08*	.00
28	.01	.52	-.27	.20	-.04*	.00

ตารางที่ 28 (ต่อ)

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
29	-.01	.59	-.29	.23	-.05*	.00
30	-.02*	.03	-.21	.12	-.03*	.00
31	-.01	.19	-.24	.21	-.04*	.00
32	.00	.74	-.14*	.01	-.10*	.00
33	-.01	.42	-.15*	.05	-.04*	.00
34	-.01	.39	-.29	.23	-.08*	.00
35	-.01	.47	-.18*	.01	-.11*	.00
36	.00	.77	-.18*	.01	-.11*	.00
37	-.01	.50	-.21	.32	-.11*	.00
38	-.02	.08	-.31	.33	-.03*	.01
39	-.03*	.01	-.18*	.01	-.10*	.00
40	-.03*	.02	-.22	.33	-.08*	.00
41	.01	.23	.08*	.00	.14*	.00
42	-.01	.32	.13*	.00	.13*	.00
43	-.02	.31	.07*	.00	.11*	.00
44	-.01	.32	.00	.69	.09*	.00
45	.00	.32	.01	.50	.03*	.01
46	.00	.12	.11*	.00	.01	.61
47	-.03*	.00	.03*	.02	.04*	.00
48	-.03*	.00	.12*	.00	.13*	.00
49	-.01	.21	.19*	.00	.25*	.00
50	-.01	.23	-.10*	.00	-.03*	.01
51	.01	.23	.28*	.00	.23*	.00
52	-.08*	.05	.28*	.00	.04*	.00
53	-.09*	.01	.30*	.00	.16*	.00
54	-.06*	.02	.22*	.00	.12*	.00
55	-.03*	.01	.25*	.00	.06*	.00
56	-.07*	.02	.27*	.00	.01	.39
57	.00*	.00	.22*	.00	-.05*	.00
58	-.02*	.02	.28*	.00	.06*	.00
59	-.02*	.01	.16*	.00	.02	.10

ตารางที่ 28 (ต่อ)

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
60	-.01	.12	.02*	.14	-.02*	.04
61	-.01	.11	.22*	.00	.02	.06
62	-.01	.21	.26*	.00	.13*	.00
63	.00	.32	.21*	.00	.09*	.00
64	.00	.31	.22*	.00	.15*	.00
65	-.01	.20	.10*	.00	.09*	.00
66	.00	.12	.13*	.00	.14*	.00
67	-.02	.22	.07*	.00	.03*	.05
68	-.01	.12	.15*	.00	-.08*	.00
69	-.01	.14	.08*	.00	.15*	.00
70	.00	.36	.07*	.00	.10*	.00
71	.01	.32	.07*	.00	.09*	.00
72	-.01	.12	.03*	.02	.03*	.01
73	-.01	.32	.11*	.00	.10*	.00
74	-.02	.33	.07*	.00	.02*	.06
75	.01	.12	.08*	.00	-.07*	.00
76	.00	.12	.14*	.00	.09*	.00
77	-.01	.23	.04*	.00	.00	.76
78	-.01	.33	.03*	.01	.00	.72
79	.01	.11	.07*	.00	.03*	.01
80	.00	.36	.00*	.00	.00*	.00

หมายเหตุ * นัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 28 พบว่า ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ โดยวิธี MIMIC จำแนกตามที่ตั้งของสถานศึกษา ปีการศึกษา 2556 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 16 ข้อ ได้แก่ 2, 5, 8, 30, 39, 40, 47, 48, 52, 53, 54, 55, 56, 57, 58, 59 ปีการศึกษา 2557 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 52 ข้อ ได้แก่ 5, 6, 10, 17, 19, 23, 24, 25, 26, 27, 32, 33, 35, 36, 39, 41, 42, 43, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80 ปีการศึกษา 2558 มีข้อสอบที่ทำ

หน้าที่ต่างกันของข้อสอบ 68 จำนวน ข้อ ได้แก่ 1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 17, 18, 19, 20, 21, 23, 24, 25, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 47, 48, 49, 50, 51, 52, 53, 54, 55, 57, 58, 60, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 75, 76, 79, 80

ตาราง 29 ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามที่ตั้งของสถานศึกษา

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
1	-.01	.46	.06*	.00	-.03*	.00
2	-.04*	.00	.08*	.00	.03*	.02
3	.00	1.00	.01	.73	-.04*	.00
4	-.03*	.02	-.01	.71	-.04*	.00
5	-.01	.43	.05*	.04	-.01	.53
6	-.02	.11	.03	.28	.02	.13
7	.04*	.00	.00	.95	-.06*	.00
8	.03*	.05	-.02	.35	.04*	.00
9	.01	.36	-.02	.29	-.07*	.00
10	.03*	.01	-.01	.65	-.04*	.00
11	-.03*	.01	-.05*	.00	-.02*	.04
12	-.08*	.00	-.05*	.02	.07*	.00
13	-.06*	.00	-.05*	.01	-.03*	.01
14	-.14*	.00	-.01	.61	.00	.94
15	-.08*	.00	-.04*	.04	-.02	.07
16	-.08*	.00	-.09*	.00	-.02	.06
17	-.06*	.00	-.06*	.00	-.04*	.00
18	-.04*	.00	-.03*	.05	.02	.15
19	.15*	.00	-.04*	.01	-.06*	.00
20	.05*	.00	.01	.34	.07*	.00
21	-.03*	.02	.07*	.00	.11*	.00
22	.07	.00	.10*	.00	.05*	.00

ตารางที่ 29 (ต่อ)

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
23	.03	.06	.07*	.00	.01	.23
24	.08*	.00	.09*	.00	-.03*	.00
25	.03*	.05	.11*	.00	-.01	.51
26	.07*	.00	.18*	.00	.14*	.00
27	-.08*	.00	-.01	.81	-.01	.65
28	.17*	.00	.01	.56	-.01	.20
29	.06*	.00	.00	.92	.01	.62
30	.16*	.00	.07*	.00	.05*	.00
31	.04*	.00	.08*	.00	.09*	.00
32	.00*	.00	-.01*	.00	.00*	.00

หมายเหตุ * นัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 29 พบว่า ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ โดยวิธี MIMIC จำแนกตามที่ตั้งของสถานศึกษา ปีการศึกษา 2556 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 8 ข้อ ได้แก่ 2, 4, 7, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 26, 27, 28, 29, 30, 31, 32 ปีการศึกษา 2557 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 20 ข้อ ได้แก่ 1, 2, 5, 11, 12, 13, 15, 16, 17, 18, 19, 21, 22, 23, 24, 25, 26, 30, 31, 32 ปีการศึกษา 2558 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบจำนวน 21 ข้อ ได้แก่ 1, 2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 17, 19, 20, 21, 22, 24, 26, 30, 31, 32

ตาราง 30 ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามที่ตั้งของสถานศึกษา

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
1	-.03*	.03	.00	.99	.03*	.00
2	-.07*	.00	-.01	.44	.05*	.00
3	-.07*	.00	.01	.30	.12*	.00
4	-.07*	.00	.01	.32	.05*	.00
5	-.08*	.00	.00	.66	.05*	.00
6	-.17*	.00	.01	.56	.08*	.00
7	-.03*	.01	-.02*	.00	.04*	.00
8	-.02	.08	.01	.14	.02*	.02
9	-.06*	.00	-.01	.54	.03*	.00
10	-.03*	.04	-.01	.40	.09*	.00
11	-.05*	.00	-.01	.37	.06*	.00
12	-.03*	.04	-.01	.62	-.02*	.03
13	-.06*	.00	-.02*	.00	-.02	.09
14	.04*	.00	.00	.77	.06*	.00
15	-.15*	.00	.00	.92	.02*	.01
16	-.06*	.00	.00	.99	.07*	.00
17	-.07*	.00	.00	.90	.16*	.00
18	-.10*	.00	.01	.21	.09*	.00
19	-.09*	.00	.01	.16	-.03*	.00
20	-.09*	.00	-.01	.49	.08*	.00
21	-.11*	.00	-.02*	.00	.05*	.00
22	-.05*	.00	.00	.98	.03*	.00
23	-.10*	.00	.01	.18	.04*	.00
24	-.13*	.00	.01	.28	.07*	.00
25	-.05*	.00	-.02*	.00	.05*	.00
26	-.05*	.00	.01	.27	.00	.92
27	-.08*	.00	.00	.77	.07*	.00
28	-.09*	.00	.01	.12	.13*	.00

ตารางที่ 30 (ต่อ)

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
29	-.03	.06	.00	.67	.04*	.00
30	-.08*	.00	-.01	.44	.00	.85
31	-.15*	.00	-.01	.30	.08*	.00
32	-.07*	.00	.00	.77	.05*	.00
33	-.08*	.00	-.03*	.00	.01	.30
34	-.07*	.00	.01	.47	.10*	.00
35	-.02	.20	.00	.99	.03*	.00
36	-.08*	.00	-.03*	.00	-.02*	.02
37	-.08*	.00	.01	.12	.13*	.00
38	.02	.12	-.01	.62	-.01	.37
39	-.10*	.00	.00	.59	.03*	.00
40	-.03*	.00	-.01	.53	.07*	.00
41	-.02	.13	.02*	.00	.01	.29
42	-.10*	.00	-.01	.98	.11*	.00
43	.04*	.00	-.01	1.00	.07*	.00
44	-.04*	.00	-.01	.99	.07*	.00
45	-.10*	.00	-.02*	.00	.12*	.00
46	.05*	.00	-.01	1.00	-.02*	.02
47	-.03*	.01	-.03*	.00	.04*	.00
48	-.05*	.00	-.01	.82	.06*	.00
49	-.10*	.00	.00	1.00	.13*	.00
50	-.01	.30	-.01	.99	-.06*	.00
51	-.02*	.03	.00	1.00	.05*	.00
52	.04*	.00	-.01	1.00	.09*	.00
53	.00	.90	.00	.99	.02*	.01
54	-.04*	.00	-.03*	.00	-.11*	.00
55	-.01	.13	.00	.88	-.04*	.00
56	-.06*	.00	-.01	.97	.06*	.00
57	-.12*	.00	.02*	.00	-.04*	.00
58	-.04*	.00	.01	.89	.03*	.00
59	-.07*	.00	-.01	.99	.04*	.00

ตารางที่ 30 (ต่อ)

ข้อสอบ	ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบ					
	ปีการศึกษา 2556		ปีการศึกษา 2557		ปีการศึกษา 2558	
	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>	<i>Estimate</i>	<i>P-Value</i>
60	-.05*	.00	-.01	.99	.13*	.00
61	-.11*	.00	-.01	.88	-.05*	.00
62	.00	.96	.00	1.00	.12*	.00
63	-.09*	.00	.01	.98	.02	.06
64	.00	.83	-.01	.96	.04*	.00
65	.03*	.00	.00	1.00	.04*	.00
66	-.06*	.00	.00	.96	-.02*	.04
67	-.02*	.03	-.02	.00	-.04*	.00
68	-.03*	.01	.00	.99	.09*	.00
69	-.04*	.00	-.01	1.00	.03*	.00
70	-.05*	.00	-.01	.89	.03*	.00
71	-.05*	.00	-.02	.00	-.05*	.00
72	-.10*	.00	.03	.00	.06*	.00
73	-.07*	.00	-.02	.00	-.02*	.02
74	-.12*	.00	.01	.51	.20*	.00
75	-.04*	.00	.00	1.00	.01	.23
76	-.05*	.00	.00	1.00	.03*	.00
77	-.06*	.00	.02	.00	.13*	.00
78	.07*	.00	-.02	.00	.02*	.01
79	-.07*	.00	.00	1.00	.16*	.00
80	.00*	.00	-.02	.01	.00*	.00

หมายเหตุ * นัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 30 พบว่า ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ โดยวิธี MIMIC จำแนกตามที่ตั้งของสถานศึกษา ปีการศึกษา 2556 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 70 ข้อ ได้แก่ 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34, 36, 37, 39, 40, 42, 43, 44, 45, 46, 47, 48, 49, 51, 52, 54, 56, 57, 58, 59, 60, 61, 63, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80 ปีการศึกษา 2557 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 18 ข้อ ได้แก่ 7, 13, 21, 25, 33, 36, 41, 45, 47,

54, 57, 67, 71, 72, 73, 77, 78, 80 ปีการศึกษา 2558 มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ 72 จำนวน ข้อ ได้แก่ 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 31, 32, 34, 35, 36, 37, 39, 40, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 76, 77, 78, 79, 80

ตาราง 31 ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศและที่ตั้งของสถานศึกษา

ปีการศึกษา	การตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบโดยวิธี MIMIC			
	เพศ		ที่ตั้งของเรียน	
	ชาย (Reference)	หญิง (Focal)	ในเมือง (Reference)	นอกเมือง (Focal)
2556	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 35, 37, 38, 39, 40, 41, 42, 43, 44, 45, 47, 49, 51, 52, 54, 55, 56, 57, 59, 60, 61, 62, 63, 65, 66, 68, 70	12, 34, 46, 48, 50, 58	1, 3, 6, 11, 16, 24, 30, 37, 40, 47, 52, 55, 56, 59	2, 5, 7, 9, 10, 12, 15, 17, 19, 20, 22, 26, 28, 33, 35, 36, 38, 39, 41, 43, 46, 48, 50, 51, 53, 58, 61, 62, 68, 69

ตารางที่ 31 (ต่อ)

ปีการศึกษา	การตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบโดยวิธี MIMIC			
	เพศ		ที่ตั้งของเรียน	
	ชาย (Reference)	หญิง (Focal)	ในเมือง (Reference)	นอกเมือง (Focal)
2557	2, 3, 5, 9, 12, 13, 19, 20, 23, 24, 25, 26, 27, 29, 30, 31, 32, 33, 36, 37, 39, 44, 45, 46, 47, 54, 55, 56, 61, 62, 63, 65, 67,	4, 6, 7, 8, 17, 18, 35, 38, 41, 42, 48, 51, 52, 60, 66	2, 29, 33, 34, 36, 42, 44, 46, 47, 52, 54, 55, 63, 65	1, 3, 4, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 26, 28, 30, 35, 38, 41, 43, 48, 49, 53, 57, 58, 60, 66, 67, 69
2558	-	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70	1, 2, 3, 6, 7, 8, 9, 11, , 13, 17, 18, 19, 21, 24, 25, 26, 27, 29, 30, 31, 32, 33, 34, 35, 36, 39, 40, 41, 42, 43, 45, 47, 48, 49, 50, 51, 52, 53, 55, 58, 59, 60, 65, 66, 68	15

หมายเหตุ เพศชาย = 1 (Reference Group), เพศหญิง = 0 (Focal Group)
 ในเมือง = 1 (Reference Group), นอกเมือง = 0 (Focal Group)

จากตาราง 31 พบว่าการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ รายวิชาภาษาไทย จำแนกตามเพศ ปีการศึกษา 2556 มีจำนวน 64 ข้อ ซึ่งเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 58 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 6 ข้อ ปีการศึกษา 2557 มีจำนวน 48 ข้อ ซึ่งเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 33 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 15 ข้อ และปีการศึกษา 2558 มีจำนวน 67 ข้อ ซึ่งไม่มีข้อสอบที่เพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 67 ข้อ

ข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ รายวิชาภาษาไทย จำแนกตามสถานที่ตั้งของสถานศึกษา ปีการศึกษา 2556 มีจำนวน 44 ข้อ ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมือง จำนวน 14 ข้อ และสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมือง จำนวน 30 ข้อ ปีการศึกษา 2557 มีจำนวน 49 ข้อ ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมือง จำนวน 14 ข้อ และสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมือง จำนวน 35 ข้อ และปีการศึกษา 2558 มีจำนวน 46 ข้อ ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมือง จำนวน 45 ข้อ และสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมือง จำนวน 1 ข้อ

ตาราง 32 ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินัยขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศและที่ตั้งของสถานศึกษา

ปีการศึกษา	การตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบโดยวิธี MIMIC			
	เพศ		ที่ตั้งของเรียน	
	ชาย (Reference)	หญิง (Focal)	ในเมือง (Reference)	นอกเมือง (Focal)
2556	16, 18, 20, 23, 25, 28, 35, 36, 41, 42, 43, 44, 45, 47, 49, 53, 56, 63, 64, 65, 66, 68, 71, 73, 75,	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 19, 22, 24, 29, 33, 34, 37, 39, 50, 51, 54, 59, 62, 72, 76, 78	-	2, 5, 8, 30, 39, 40, 47, 48, 52, 53, 54, 55, 56, 58, 59

ตารางที่ 32 (ต่อ)

ปีการศึกษา	การตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบโดยวิธี MIMIC			
	เพศ		ที่ตั้งของเรียน	
	ชาย (Reference)	หญิง (Focal)	ในเมือง (Reference)	นอกเมือง (Focal)
2557	3, 8, 12, 13, 14, 21, 22, 24, 28, 29, 38, 39, 45, 50	1, 19, 23, 25, 32, 37, 41, 43, 44, 46, 47, 48, 49, 51, 52, 53, 54, 55, 56, 57, 58, 59, 61, 62, 63, 64, 65, 66, 67, 68, 70, 71, 73, 75, 76, 78, 79	6, 41, 42, 43, 46, 47, 48, 49, 51, 52, 53, 54, 55, 56, 57, 58, 59, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79	5, 10, 17, 19, 23, 24, 25, 26, 27, 32, 33, 35, 36, 39, 50
2558	18, 33, 34, 46, 50, 56, 60, 68, 74,	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 35, 36, 38, 39, 40, 41, 42, 43, 44, 45, 47, 48, 49, 51, 52, 53, 54, 55, 57, 58, 59, 62, 63, 64, 65, 66, 69, 70, 71, 72, 73, 75, 76, 77, 79,	1, 3, 4, 6, 8, 9, 11, 13, 20, 21, 41, 42, 43, 44, 45, 47, 48, 49, 51, 52, 53, 54, 55, 58, 62, 63, 64, 65, 66, 67, 69, 70, 71, 72, 73, 76, 79	5, 7, 10, 12, 14, 17, 18, 19, 23, 24, 25, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 50, 57, 60, 68, 75

หมายเหตุ เพศชาย = 1 (Reference Group), เพศหญิง = 0 (Focal Group)
 ในเมือง = 1 (Reference Group), นอกเมือง = 0 (Focal Group)

จากตาราง 32 พบว่าการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบรายวิชาภาษาอังกฤษ จำแนกตามเพศ ปีการศึกษา 2556 มีจำนวน 56 ข้อ ซึ่งเพศชายมีโอกาสดอบถูกมากกว่าเพศหญิง จำนวน 25 ข้อ และเพศหญิงมีโอกาสดอบถูกมากกว่าเพศชาย จำนวน 30 ข้อ ปีการศึกษา 2557 มีจำนวน 52 ข้อ ซึ่งมีเพศชายมีโอกาสดอบถูกมากกว่าเพศหญิง จำนวน 14 ข้อ และเพศหญิงมีโอกาสดอบถูกมากกว่าเพศชาย จำนวน 37 ข้อ และ ปีการศึกษา 2558 มีจำนวน 75 ข้อ ซึ่งเพศชายมีโอกาสดอบถูกมากกว่าเพศหญิง จำนวน 9 ข้อ และเพศหญิงมีโอกาสดอบถูกมากกว่าเพศชาย จำนวน 65 ข้อ

ข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบรายวิชาภาษาอังกฤษ จำแนกตามสถานที่ตั้งของสถานศึกษา ปีการศึกษา 2556 มีจำนวน 15 ข้อ ซึ่งไม่มีสถานที่ตั้งในเมืองมีโอกาสดอบถูกมากกว่านอกเมือง และสถานที่ตั้งนอกเมืองมีโอกาสดอบถูกมากกว่าในเมือง จำนวน 15 ข้อ ปีการศึกษา 2557 มีจำนวน 52 ข้อ ซึ่งสถานที่ตั้งในเมืองมีโอกาสดอบถูกมากกว่านอกเมือง จำนวน 36 ข้อ และสถานที่ตั้งนอกเมืองมีโอกาสดอบถูกมากกว่าในเมือง จำนวน 15 ข้อ และปีการศึกษา 2558 มีจำนวน 67 ข้อ ซึ่งสถานที่ตั้งในเมืองมีโอกาสดอบถูกมากกว่านอกเมือง จำนวน 37 ข้อ และสถานที่ตั้งนอกเมืองมีโอกาสดอบถูกมากกว่าในเมือง จำนวน 30 ข้อ

Prince of Songkla University
Pattani Campus

ตาราง 33 ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศและที่ตั้งของสถานศึกษา

ปีการศึกษา	การตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบโดยวิธี MIMIC			
	เพศ		ที่ตั้งของเรียน	
	ชาย (Reference)	หญิง (Focal)	ในเมือง (Reference)	นอกเมือง (Focal)
2556	1, 2, 4, 5, 6, 11, 14, 15, 16, 17, 18, 23, 24, 25,	26, 28, 30, 31	7, 10, 19, 20, 22, 24, 26, 28, 29, 30, 31	2, 4, 11, 12, 13, 14, 15, 16, 17, 18, 21, 27
2557	-	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 30, 31	1, 2, 5, 21, 22, 23, 24, 25, 26, 30, 31,	11, 12, 13, 15, 16, 17, 18, 19, 32
2558	2, 3, 4, 5, 6, 7, 8, 9, 10, 13, 14, 15, 16, 17, 18, 19, 20, 23, 25, 27, 29,	11, 21, 24, 26, 31	2, 8, 12, 20, 21, 22, 26, 30, 31, 32	1, 3, 4, 7, 9, 10, 11, 13, 17, 19, 24,

หมายเหตุ เพศชาย = 1 (Reference Group), เพศหญิง = 0 (Focal Group)
 ในเมือง = 1 (Reference Group), นอกเมือง = 0 (Focal Group)

จากตาราง 33 พบว่าการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ รายวิชาคณิตศาสตร์ จำแนกตามเพศ ปีการศึกษา 2556 มีจำนวน 18 ข้อ ซึ่งเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 14 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 4 ข้อ ปีการศึกษา 2557 มีจำนวน 28 ข้อ ซึ่งไม่มีเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 28 ข้อ และปีการศึกษา 2558 มีจำนวน 26 ข้อ ซึ่งเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 21 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 5 ข้อ ข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ รายวิชาคณิตศาสตร์ จำแนกตามสถานที่ตั้งของสถานศึกษา ปีการศึกษา 2556 มีจำนวน 23 ข้อ ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมือง จำนวน 11 ข้อ และสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมือง จำนวน 12 ข้อ ปี

การศึกษา 2557 มีจำนวน 20 ข้อ ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมือง จำนวน 11 ข้อ และสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมือง จำนวน 9 ข้อ และปีการศึกษา 2558 มีจำนวน 21 ข้อ ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมือง จำนวน 11 ข้อ และสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมือง จำนวน 10 ข้อ

ตาราง 34 ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศและที่ตั้งของสถานศึกษา

ปีการศึกษา	การตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบโดยวิธี MIMIC			
	เพศ		ที่ตั้งของเรียน	
	ชาย (Reference)	หญิง (Focal)	ในเมือง (Reference)	นอกเมือง (Focal)
2556	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57, 58, 60, 61, 62, 63, 65, 66, 67, 70, 72, 73, 75, 76, 77	55, 59, 64, 68, 78	14, 43, 46, 52, 65, 78	1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34, 36, 37, 39, 40, 42, 44, 45, 47, 48, 49, 51, 54, 56, 57, 58, 59, 60, 61, 63, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 79

ตารางที่ 34 (ต่อ)

ปีการศึกษา	การตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบโดยวิธี MIMIC			
	เพศ		ที่ตั้งของเรียน	
	ชาย (Reference)	หญิง (Focal)	ในเมือง (Reference)	นอกเมือง (Focal)
2557	1, 36, 43, 46, 49, 51, 57, 58, 64, 69, 72, 77,	17, 24, 41, 42, 44, 45, 47, 48, 50, 52, 53, 55, 56, 59, 61, 62, 63, 65, 66, 67, 68, 70, 71, 73, 74, 75, 76, 78, 79, 80	41, 57, 72, 77	7, 13, 21, 25, 33, 36, 47, 54, 67, 71, 73, 78, 80
2558	7, 9, 14, 17, 18, 20, 21, 22, 23, 24, 25, 28, 31, 34, 37, 40, 42, 43, 49, 52, 58, 60, 62, 64, 68, 69, 70, 72, 74, 77, 79	1, 3, 4, 8, 10, 12, 13, 15, 16, 19, 26, 27, 29, 30, 33, 35, 36, 38, 39, 41, 44, 46, 47, 48, 50, 53, 54, 55, 56, 57, 65, 66, 71, 75, 78	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 14, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 27, 28, 29, 31, 32, 34, 35, 37, 39, 40, 42, 43, 44, 45, 47, 48, 49, 51, 52, 53, 56, 58, 59, 60, 62, 64, 65, 68, 69, 70, 72, 74, 76, 77, 78, 79	12, 19, 36, 46, 50, 54, 55, 57, 61, 66, 67, 71, 73

หมายเหตุ เพศชาย = 1 (Reference Group), เพศหญิง = 0 (Focal Group)
 ในเมือง = 1 (Reference Group), นอกเมือง = 0 (Focal Group)

จากตาราง 34 พบว่าการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ รายวิชา
 วิทยาศาสตร์ จำแนกตามเพศ ปีการศึกษา 2556 มีจำนวน 74 ข้อ ซึ่งเพศชายมีโอกาสตอบถูก
 มากกว่าเพศหญิง จำนวน 69 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 5 ข้อ ปี
 การศึกษา 2557 มีจำนวน 42 ข้อ ซึ่งมีเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 12 ข้อ และ

เพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 30 ข้อ และ ปีการศึกษา 2558 มีจำนวน 66 ข้อ ซึ่งเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 31 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 35 ข้อ

ข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ รายวิชาวิทยาศาสตร์ จำแนกตามสถานที่ตั้งของสถานศึกษา ปีการศึกษา 2556 มีจำนวน 69 ข้อ ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมือง จำนวน 6 ข้อ และสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมือง จำนวน 63 ข้อ ปีการศึกษา 2557 มีจำนวน 17 ข้อ ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมือง จำนวน 4 ข้อ และสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมือง จำนวน 13 ข้อ และปีการศึกษา 2558 มีจำนวน 72 ข้อ ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมือง จำนวน 58 ข้อ และสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมือง จำนวน 13 ข้อ

Prince of Songkla University
Pattani Campus

ตอนที่ 4 ผลการวิเคราะห์แนวโน้มการทำหน้าที่ต่างกันของข้อสอบ

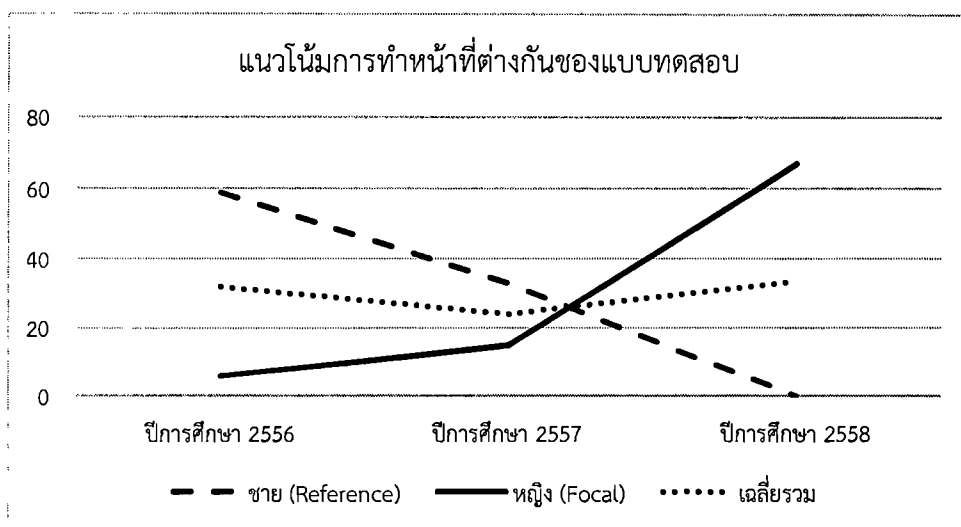
การวิเคราะห์แนวโน้มการทำหน้าที่ต่างกันของข้อสอบในครั้งนี้ได้จากการตอบข้อสอบของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ปีการศึกษา 2556 - 2558 โดยที่รายวิชาวิทยาศาสตร์ ภาษาอังกฤษ คณิตศาสตร์ และวิทยาศาสตร์ มีจำนวนแบบทดสอบรายวิชาละ 70, 80, 32 และ 80 ผู้วิจัยได้ใช้วิธี MIMIC โดยประยุกต์ใช้โปรแกรม Mplus ตามลำดับ โดยจำแนกตามเพศและที่ตั้งของสถานศึกษา ซึ่งมีการวิเคราะห์ด้วยวิธีดังนี้

ตาราง 35 ผลการวิเคราะห์แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 - 2558 โดยวิธี MIMIC จำแนกตามเพศและที่ตั้งของสถานศึกษา

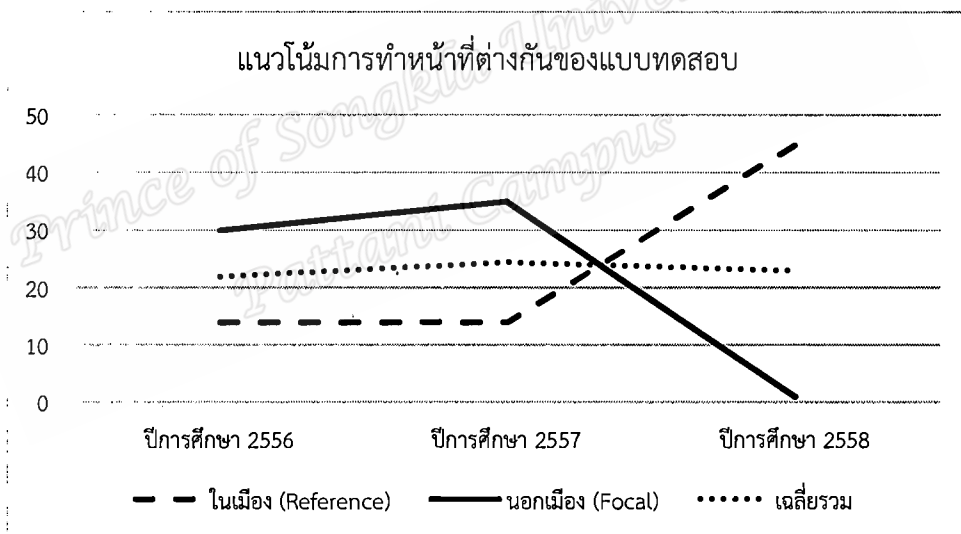
ตัวแปร	จำนวนข้อที่ทำหน้าที่ต่างกันของแบบทดสอบโดยวิธี MIMIC			แนวโน้ม
	ปีการศึกษา			
	2556	2557	2558	
เพศ				
ชาย (Reference)	59	33	0	ลดลง
หญิง (Focal)	6	15	67	เพิ่มขึ้น
เฉลี่ยรวม	32	24	33.50	ไม่แน่นอน
ที่ตั้งของเรียน				
ในเมือง (Reference)	14	14	45	เพิ่มขึ้น
นอกเมือง (Focal)	30	35	1	ไม่แน่นอน
เฉลี่ยรวม	22	24.50	23	ไม่แน่นอน

จากตาราง 35 พบว่าการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ รายวิชาภาษาไทย ปีการศึกษา 2556 - 2558 จำแนกตามเพศ โดยเฉลี่ยข้อสอบมีการทำหน้าที่ต่างกันมีแนวโน้มไม่แน่นอน ซึ่งเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิงมีแนวโน้มลดลงและเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชายมีแนวโน้มเพิ่มขึ้น

จำแนกตามสถานที่ตั้งของสถานศึกษา โดยเฉลี่ยข้อสอบมีการทำหน้าที่ต่างกันมีแนวโน้มไม่แน่นอน ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมืองมีแนวโน้มเพิ่มขึ้น และสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมือง มีแนวโน้มไม่แน่นอน



ภาพประกอบ 22 แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558 จำแนกตามเพศ



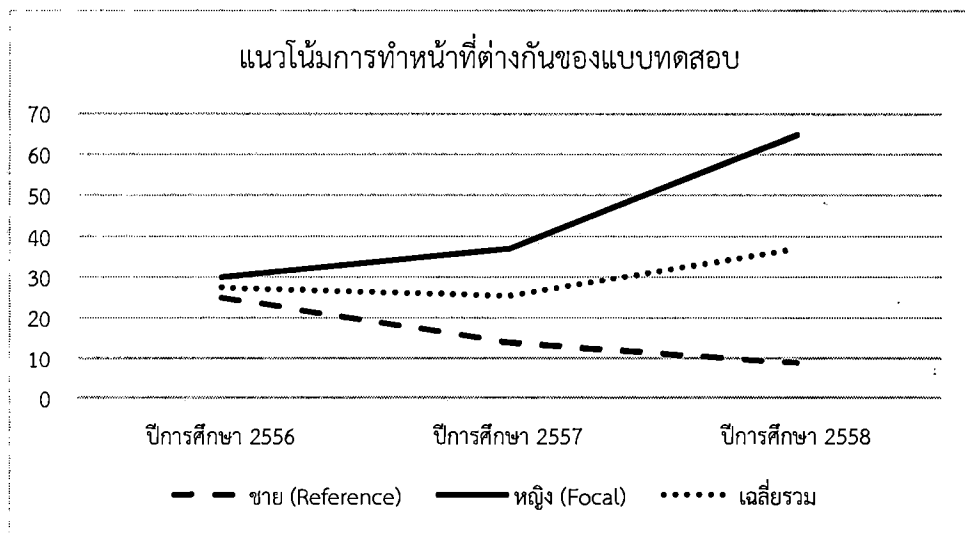
ภาพประกอบ 23 แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 – 2558 จำแนกตามที่ตั้งของสถานศึกษา

ตาราง 36 ผลการวิเคราะห์แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศ และที่ตั้งของสถานศึกษา

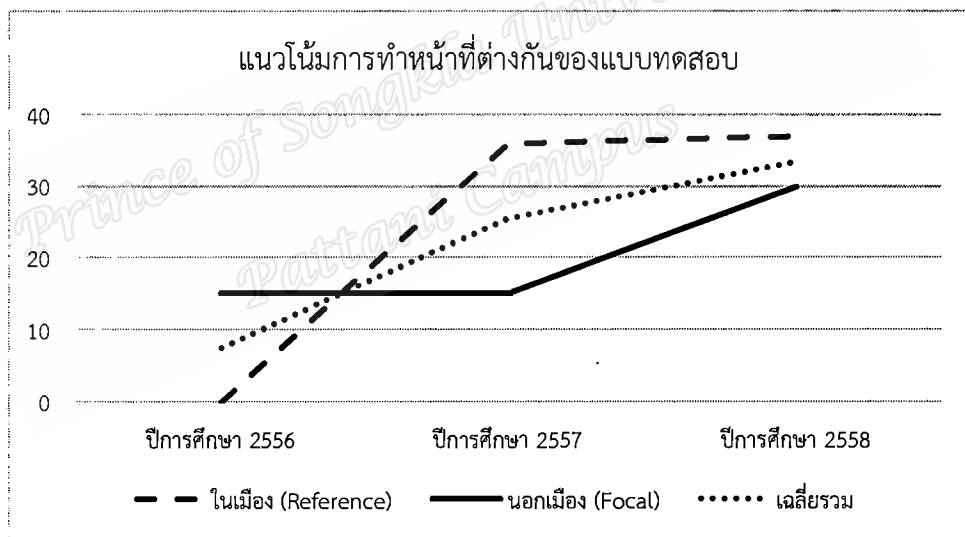
ตัวแปร	จำนวนข้อที่ทำหน้าที่ต่างกันของแบบทดสอบโดยวิธี MIMIC			
	ปีการศึกษา			แนวโน้ม
	2556	2557	2558	
เพศ				
ชาย (Reference)	25	14	9	ลดลง
หญิง (Focal)	30	37	65	เพิ่มขึ้น
เฉลี่ยรวม	27.50	25.50	37	ไม่แน่นอน
ที่ตั้งของเรียน				
ในเมือง (Reference)	0	36	37	เพิ่มขึ้น
นอกเมือง (Focal)	15	15	30	เพิ่มขึ้น
เฉลี่ยรวม	7.5	25.50	33.50	เพิ่มขึ้น

จากตาราง 36 พบว่าการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558 จำแนกตามเพศ โดยเฉลี่ยข้อสอบมีการทำหน้าที่ต่างกันมีแนวโน้มไม่แน่นอน ซึ่งเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิงมีแนวโน้มลดลงและเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชายมีแนวโน้มเพิ่มขึ้น

จำแนกตามสถานที่ตั้งของสถานศึกษา โดยเฉลี่ยข้อสอบมีการทำหน้าที่ต่างกันมีแนวโน้มเพิ่มขึ้น ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมืองมีแนวโน้มเพิ่มขึ้น และสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมือง มีแนวโน้มเพิ่มขึ้น



ภาพประกอบ 24 แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558 จำแนกตามเพศ



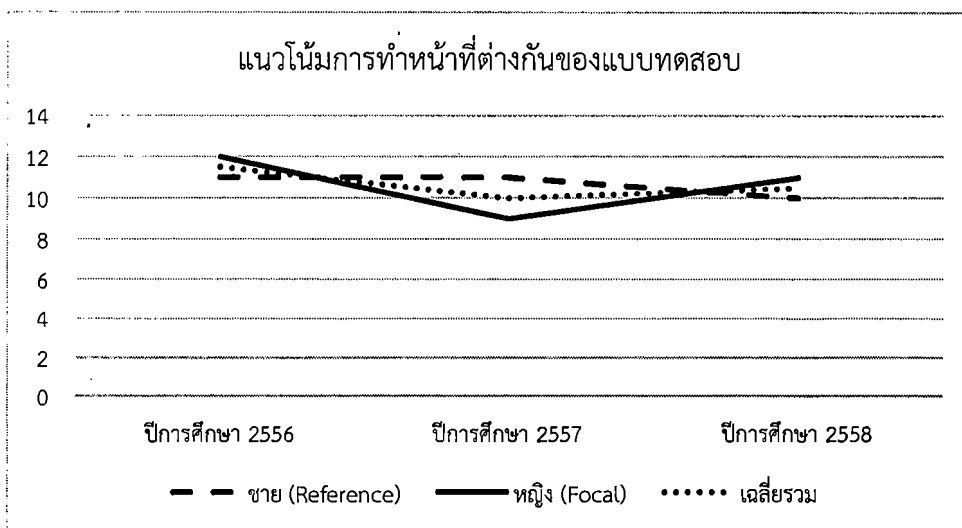
ภาพประกอบ 25 แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) รายวิชาภาษาอังกฤษ ปีการศึกษา 2556 – 2558 จำแนกตามที่ตั้งของสถานศึกษา

ตาราง 37 ผลการวิเคราะห์แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศ และที่ตั้งของสถานศึกษา

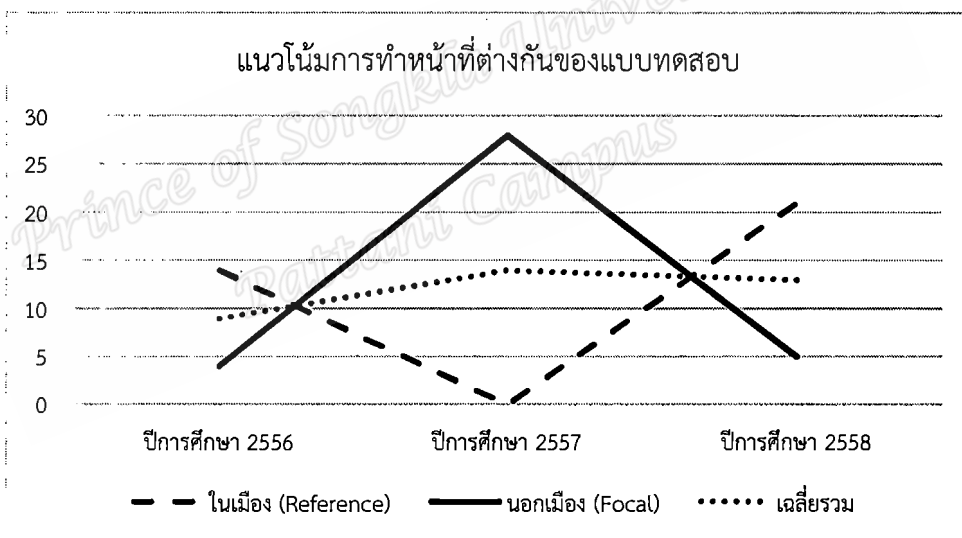
ตัวแปร	จำนวนข้อที่ทำหน้าที่ต่างกันของแบบทดสอบโดยวิธี MIMIC			แนวโน้ม
	ปีการศึกษา			
	2556	2557	2558	
เพศ				
ชาย (Reference)	11	11	10	ลดลง
หญิง (Focal)	12	9	11	ไม่แน่นอน
เฉลี่ยรวม	11.50	10	10.50	ไม่แน่นอน
ที่ตั้งของเรียน				
ในเมือง (Reference)	14	0	21	ไม่แน่นอน
นอกเมือง (Focal)	4	28	5	ไม่แน่นอน
เฉลี่ยรวม	9	14	13	ไม่แน่นอน

จากตาราง 37 พบว่าการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558 จำแนกตามเพศ โดยเฉลี่ยข้อสอบมีการทำหน้าที่ต่างกันมีแนวโน้มไม่แน่นอน ซึ่งเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิงมีแนวโน้มลดลงและเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชายมีแนวโน้มไม่แน่นอน

จำแนกตามสถานที่ตั้งของสถานศึกษา โดยเฉลี่ยข้อสอบมีการทำหน้าที่ต่างกันมีแนวโน้มไม่แน่นอน ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมืองมีแนวโน้มไม่แน่นอนและสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมืองมีแนวโน้มไม่แน่นอน



ภาพประกอบ 26 แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558 จำแนกตามเพศ



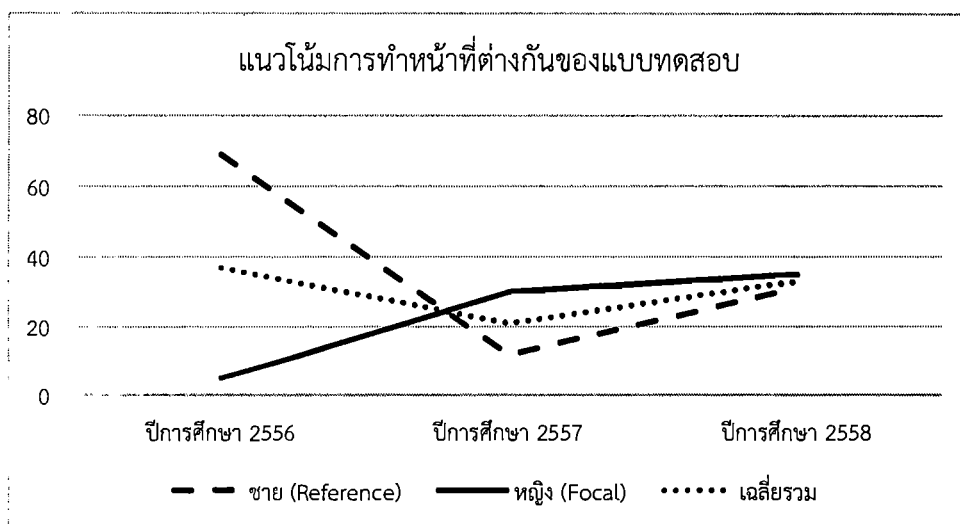
ภาพประกอบ 27 แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ ปีการศึกษา 2556 – 2558 จำแนกตามที่ตั้งของสถานศึกษา

ตาราง 38 ผลการวิเคราะห์แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินัยพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558 โดยวิธี MIMIC จำแนกตามเพศ และที่ตั้งของสถานศึกษา

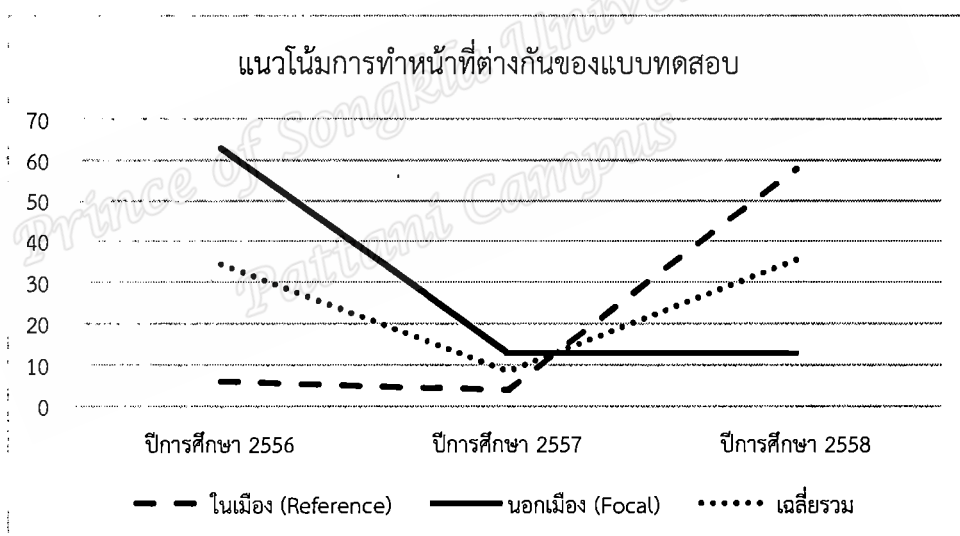
ตัวแปร	จำนวนข้อที่ทำหน้าที่ต่างกันของแบบทดสอบโดยวิธี MIMIC			แนวโน้ม
	ปีการศึกษา			
	2556	2557	2558	
เพศ				
ชาย (Reference)	69	12	31	ไม่แน่นอน
หญิง (Focal)	5	30	35	เพิ่มขึ้น
เฉลี่ยรวม	37	21	33	ไม่แน่นอน
ที่ตั้งของเรียน				
ในเมือง (Reference)	6	4	58	ไม่แน่นอน
นอกเมือง (Focal)	63	13	13	ลดลง
เฉลี่ยรวม	34.50	8.50	35.50	ไม่แน่นอน

จากตาราง 38 พบว่าการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558 จำแนกตามเพศ โดยเฉลี่ยข้อสอบมีการทำหน้าที่ต่างกันมีแนวโน้มไม่แน่นอน ซึ่งเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิงมีแนวโน้มไม่แน่นอนและเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชายมีแนวโน้มเพิ่มขึ้น

จำแนกตามสถานที่ตั้งของสถานศึกษา โดยเฉลี่ยข้อสอบมีการทำหน้าที่ต่างกันมีแนวโน้มไม่แน่นอน ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมืองมีแนวโน้มไม่แน่นอนและสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมืองมีแนวโน้มลดลง



ภาพประกอบ 28 แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558 จำแนกตามเพศ



ภาพประกอบ 29 แนวโน้มการทำหน้าที่ต่างกันของแบบทดสอบทางการศึกษาระดับชาตินิยมพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 – 2558 จำแนกตามที่ตั้งของสถานศึกษา

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์ของการวิจัย คือ เพื่อวิเคราะห์คุณภาพแบบทดสอบของการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2556 - 2558 และตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิติก (Multiple Indicators and Multiple Causes : MIMIC) เมื่อจำแนกตามเพศและที่ตั้งของสถานศึกษา ระหว่างปีการศึกษา 2556 - 2558

ประชากรที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่เข้าสอบวัดผลสัมฤทธิ์ทางการเรียนเพื่อประเมินคุณภาพการศึกษาระดับชาติปีการศึกษา 2556 - 2558 ซึ่งเป็นนักเรียนที่เข้าสอบรายวิชาภาษาไทย ภาษาอังกฤษ คณิตศาสตร์ และวิทยาศาสตร์

สรุปผลการวิจัย

ผลการวิจัยโดยมีวัตถุประสงค์เพื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิติก (Multiple Indicators and Multiple Causes : MIMIC) เมื่อจำแนกตามกลุ่มเพศและกลุ่มที่ตั้งของสถานศึกษา ระหว่างปีการศึกษา 2556 - 2558 ดังนี้

1. ผลการประมาณค่าพารามิเตอร์ความยาก อำนาจจำแนกของข้อสอบ และระดับความสามารถของผู้สอบ

1.1 ค่าพารามิเตอร์ความยากของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาไทย ในปีการศึกษา 2556 มีค่าความยากน้อยสุดเท่ากับ .02 และสูงสุดเท่ากับ 7.60 ปีการศึกษา 2557 มีค่าความยากน้อยสุดเท่ากับ .01 และสูงสุดเท่ากับ 8.32 และปีการศึกษา 2558 มีค่าความยากน้อยสุดเท่ากับ .02 และสูงสุดเท่ากับ 4.95 ค่าพารามิเตอร์อำนาจจำแนกของแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาไทย ในปีการศึกษา 2556 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ .01 และสูงสุดเท่ากับ 1.03 ปีการศึกษา 2557 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ .07 และสูงสุดเท่ากับ .99 และปีการศึกษา 2558 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ .12 และสูงสุดเท่ากับ .95

ค่าระดับความสามารถของผู้สอบที่ได้ทำแบบทดสอบการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) รายวิชาภาษาไทย ปีการศึกษา 2556 มีจำนวนผู้สอบจำนวน 414,589 คน มีค่าความสามารถของผู้สอบน้อยสุด -3.61 และสูงสุดเท่ากับ 3.40 ปีการศึกษา 2557 มีจำนวนผู้สอบจำนวน 430,617 คน มีค่าความสามารถของผู้สอบน้อยสุด -3.50 และสูงสุดเท่ากับ

เท่ากับ .66 ปีการศึกษา 2557 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ -1.09 และสูงสุดเท่ากับ 1.18 และปีการศึกษา 2558 มีค่าอำนาจจำแนกน้อยสุดเท่ากับ -.66 และสูงสุดเท่ากับ 1.04

ค่าระดับความสามารถของผู้สอบที่ได้ทำแบบทดสอบการทดสอบทางการศึกษาระดับชาตินิยมพื้นฐาน (O-NET) รายวิชาวิทยาศาสตร์ ปีการศึกษา 2556 มีจำนวนผู้สอบจำนวน 414,588 คน มีค่าความสามารถของผู้สอบน้อยสุด -2.67 และสูงสุดเท่ากับ 4.75 ปีการศึกษา 2557 มีจำนวนผู้สอบจำนวน 430,607 คน มีค่าความสามารถของผู้สอบน้อยสุด -2.74 และสูงสุดเท่ากับ 2.85 และปีการศึกษา 2558 มีจำนวนผู้สอบจำนวน 423,495 คน มีค่าความสามารถของผู้สอบน้อยสุด -2.54 และสูงสุดเท่ากับ 2.64

2. ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบจำแนกตามเพศและที่ตั้งของสถานศึกษา

2.1 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ รายวิชาภาษาไทย จำแนกตามเพศ ปีการศึกษา 2556 มีจำนวน 64 ข้อ ซึ่งเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 58 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 6 ข้อ ปีการศึกษา 2557 มีจำนวน 48 ข้อ ซึ่งเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 33 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 15 ข้อ และปีการศึกษา 2558 มีจำนวน 67 ข้อ ซึ่งไม่มีข้อสอบที่เพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 67 ข้อ

ข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ รายวิชาภาษาไทย จำแนกตามสถานที่ตั้งของสถานศึกษา ปีการศึกษา 2556 มีจำนวน 44 ข้อ ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมือง จำนวน 14 ข้อ และสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมือง จำนวน 30 ข้อ ปีการศึกษา 2557 มีจำนวน 49 ข้อ ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมือง จำนวน 14 ข้อ และสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมือง จำนวน 35 ข้อ และปีการศึกษา 2558 มีจำนวน 46 ข้อ ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมือง จำนวน 45 ข้อ และสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมือง จำนวน 1 ข้อ

2.2 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบรายวิชาภาษาอังกฤษ จำแนกตามเพศ ปีการศึกษา 2556 มีจำนวน 56 ข้อ ซึ่งเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 25 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 30 ข้อ ปีการศึกษา 2557 มีจำนวน 52 ข้อ ซึ่งมีเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 14 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 37 ข้อ และปีการศึกษา 2558 มีจำนวน 75 ข้อ ซึ่งเพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 9 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 65 ข้อ

ข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบรายวิชาภาษาอังกฤษ จำแนกตามสถานที่ตั้งของสถานศึกษา ปีการศึกษา 2556 มีจำนวน 15 ข้อ ซึ่งไม่มีข้อสอบที่สถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมือง และสถานที่ตั้งนอกเมืองมีโอกาสตอบถูกมากกว่าในเมือง จำนวน 15 ข้อ ปีการศึกษา 2557 มีจำนวน 52 ข้อ ซึ่งสถานที่ตั้งในเมืองมีโอกาสตอบถูกมากกว่านอกเมือง

อภิปรายผลการวิจัย

การวิจัยในครั้งนี้เป็นการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมีค (Multiple Indicators and Multiple Causes : MIMIC) โดยใช้โปรแกรมประยุกต์ Mplus โดยมีข้อมูลที่มีผลการตอบให้คะแนนแบบสองค่า (Dichotomous) เมื่อจำแนกตามกลุ่มเพศและที่ตั้งของสถานศึกษา ระหว่างปีการศึกษา 2556 - 2558 ผู้วิจัยนำเสนอการอภิปรายผลการวิจัยตามผลที่ได้ ดังนี้

ลักษณะข้อสอบที่เกิดจากการทำหน้าที่ต่างกัน โดยจำแนกตามเพศ เนื่องจากข้อสอบสามารถตอบได้หลายคำตอบตามความคิดเห็นของแต่ละบุคคล และข้อสอบที่เกิดการทำหน้าที่ต่างกันของข้อสอบ อาจมีคำศัพท์ที่เกี่ยวข้องกับเพศชายหรือเพศหญิง จึงทำให้ข้อสอบเข้าข้างกลุ่มนั้นๆ และทั้งนี้อาจเป็นเพราะความสามารถที่แตกต่างระหว่างเพศชายหรือเพศหญิง ที่มีความสามารถ ความถนัด และความสนใจในเรื่องนั้นๆ ต่างกัน จากการศึกษาเอกสารและงานที่เกี่ยวข้อง พบว่า Maier และ Casselman (1970 อ้างถึงใน สุพัฒน์นา หอมบุปผา, 2556) กล่าวว่าเพศชายจะมีความสามารถในการแก้ปัญหาทางคณิตศาสตร์และวิชาวิทยาศาสตร์เก่งกว่าเพศหญิง ส่วนเพศหญิงจะมีความสามารถในทางภาษาและสังคมศาสตร์เก่งกว่าเพศชาย ซึ่งสอดคล้องกับผลการวิจัยของ ศุภวัฒน์ มะลิเผือก (2548) ที่ได้ศึกษาอิทธิพลของการทำหน้าที่ต่างกันของข้อสอบ ที่ส่งผลต่อคุณภาพของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนระดับชาติ ชั้นประถมศึกษาปีที่ 6 ผลการวิจัยปรากฏว่าพบข้อสอบทำหน้าที่ต่างกันในรูปแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนระดับชาติภาษาไทย เมื่อจำแนกกลุ่มผู้สอบตามตัวแปรเพศ จำนวน 12 ข้อ คิดเป็นร้อยละ 30 และผลการวิจัยของ ทิพย์วรรณ คงเนตร (2549) ได้ศึกษาการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบวิชาทักษะภาษาไทย สำหรับการวัดผลปลายภาคการศึกษาของมหาวิทยาลัยนเรศวร ปีการศึกษา 2546 ผลการวิจัยพบว่าเมื่อจำแนกผู้สอบตามเพศ มีข้อสอบที่ DIF จำนวน 71 ข้อ คิดเป็นร้อยละ 59.17 เป็นข้อที่เข้าข้างเพศหญิงจำนวน 38 ข้อ และเข้าข้างเพศชายจำนวน 33 ข้อ คิดเป็น ร้อยละ 31.67 และร้อยละ 27.50 ตามลำดับ และยังสอดคล้องกับผลการวิจัยของ Le (2006 อ้างถึงใน ประภุติยา ทักษิณ, 2552) ที่วิเคราะห์หาข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบโดยใช้ข้อมูลจาก PISA รอบทดลองใช้ ปี 2006 ในมิติที่ต่างกัน เมื่อวิเคราะห์ในมิติของสมรรถนะทางวิทยาศาสตร์ พบข้อสอบที่ทำหน้าที่ต่างกันที่ไอออนเอียงให้เพศชายร้อยละ 13.5 และไอออนเอียงให้กับเพศหญิงร้อยละ 9.5 และเมื่อวิเคราะห์ในมิติของเป้าหมายของคำถาม พบข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบที่ไอออนเอียงให้เพศชายร้อยละ 18.1 และไอออนเอียงให้เพศหญิงร้อยละ 8.2

ลักษณะข้อสอบที่เกิดจากการทำหน้าที่ต่างกัน โดยจำแนกตามที่ตั้งของสถานศึกษา อาจเนื่องจากประสบการณ์เรื่องนั้น สภาพแวดล้อมและการฝึกปฏิบัติที่แตกต่างกันระหว่างผู้เรียนที่ตั้งในสถานศึกษาในเมืองและนอกเมือง ความไม่เท่าเทียมของการขยายโอกาสทางการศึกษา เศรษฐกิจ และฐานะทางครอบครัวระหว่างผู้เรียนที่ตั้งในสถานศึกษาในเมืองและนอกเมือง ถ้าครอบครัวที่มีฐานะทางเศรษฐกิจดีสามารถสนับสนุนให้บุตรหลานได้รับการศึกษาที่ดีได้มากกว่าครอบครัวที่มีฐานะ

ทางเศรษฐกิจไม่ดี ความไม่เท่าเทียมของผู้เรียนในเขตเมืองจะมีโอกาสจะเข้าถึงการศึกษาได้ง่ายกว่า ทั้งด้านระยะทางและคุณภาพของโรงเรียน ซึ่งสอดคล้องกับผลการวิจัยของ ขวลิต นิवासวัต (2541) ที่ได้ศึกษาการเปรียบเทียบผลของการวิเคราะห์ความลำเอียงของข้อสอบด้วยวิธีค่าอำนาจจำแนกของ ข้อสอบ วิธีแมนเทิล-แฮนส์เซล และวิธีโค้งลักษณะข้อสอบ 1 พารามิเตอร์ ผลการวิจัยพบว่า วิธีโค้ง ลักษณะข้อสอบ 1 พารามิเตอร์ เป็นจำนวนข้อสอบที่ลำเอียงตามเขตที่ตั้งของสถานศึกษา 12 ข้อ วิธี ค่าอำนาจจำแนกของข้อสอบ พบจำนวนข้อสอบที่ลำเอียงตามเขตที่ตั้งของสถานศึกษา 1 ข้อ วิธีโค้ง ลักษณะข้อสอบ 1 พารามิเตอร์ พบจำนวนข้อสอบที่ลำเอียงตามเขตที่ตั้งของสถานศึกษา 8 ข้อ และ ผลการวิจัยของ วรณา รอดด้วง (2544) ได้ศึกษาการเปรียบเทียบผลของการวิเคราะห์ความลำเอียง ของข้อสอบที่แตกต่างกัน 3 วิธี ผลการวิจัยพบว่า การวิเคราะห์ความลำเอียงของข้อสอบรายข้อด้วย วิธีวิเคราะห์ 3 วิธี ตัวแปรเขตที่ตั้งของสถานศึกษา วิธีค่าอำนาจจำแนกของข้อสอบพบข้อสอบที่ ลำเอียงตามเขตที่ตั้งของสถานศึกษา จำนวน 12 วิธีแมนเทิล-แฮนส์เซล พบข้อสอบที่ลำเอียงตามเขต ที่ตั้งของสถานศึกษา จำนวน 8 ข้อ และ วิธีชิปเทสท์พบข้อสอบที่ลำเอียงตามเขตที่ตั้งของสถานศึกษา จำนวน 6 ข้อ

Prince of Songkla University
Pattani Campus

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1.1 ผู้ที่เกี่ยวข้องในการออกข้อสอบควรนำผลการวิจัยนี้ไปใช้เป็นข้อมูลในการพิจารณาตัดสินผลการสอบโดยพิจารณาการทำหน้าที่ต่างกันของข้อสอบ

1.2 การควบคุมการทำหน้าที่ต่างกันของข้อสอบในทางปฏิบัติควรพิจารณา ทบทวนข้อสอบ กำจัดความเป็นไปได้ของสาเหตุที่จะทำให้เกิดการทำหน้าที่ต่างกันของข้อสอบ เช่น ภาษาที่ใช้ในการเขียนข้อสอบ ลักษณะเนื้อหาที่นำมาสร้างข้อสอบ ลักษณะเฉพาะของกลุ่มผู้สอบ เป็นต้น

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ให้มีการตรวจสอบรูปแบบ ก่อนว่ามีรูปแบบเอกรูป (Uniform) หรือ อนุกรุป (Nonuniform)

2.2 ควรมีการศึกษาการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous Item Response Theory)

2.3 ควรมีการทำการวิจัย เพื่อเปรียบเทียบการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบระหว่างวิธีต่างๆ ซึ่งจะตรวจสอบได้ว่าวิธีการใดมีความเหมาะสมในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ

2.2 ควรมีการศึกษาตัวแปรกลุ่มอื่นๆ เช่น ภูมิภาค ศาสนา เชื้อชาติ ภาษา อายุ เป็นต้น ว่าแบบทดสอบแต่ละวิชาจะส่งผลให้พบข้อสอบทำหน้าที่ต่างกันของข้อสอบ แตกต่างกันหรือไม่

บรรณานุกรม

- ขวลิต นิวาสวัต. (2541). การเปรียบเทียบผลของการวิเคราะห์ความลำเอียงของข้อสอบด้วยวิธีค่าอำนาจจำแนกของข้อสอบ วิธีแมนเทิล-แฮนส์เซล และวิธีโค้งลักษณะข้อสอบ 1 พารามิเตอร์ (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ). มหาวิทยาลัยรามคำแหง. กรุงเทพฯ
- ทองอยู่ สาระ. (2543). การเปรียบเทียบอำนาจการตรวจสอบและการจำแนกชนิดผลัดในการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันแบบสมมาตรและแบบไม่สมมาตร ระหว่างวิธีแมนเทิล-แฮนส์เซลและวิธีการถดถอยโลจิสติก โดยใช้ความยาวแบบทดสอบและขนาดกลุ่มตัวอย่างต่างกัน (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ). มหาวิทยาลัยศรีนครินทรวิโรฒ. กรุงเทพฯ
- ทิพย์วรรณ คงเนตร. (2549). การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบวิชาทักษะภาษาไทย สำหรับการวัดผลปลายภาคการศึกษาของมหาวิทยาลัยนเรศวร ปีการศึกษา 2546 (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ). มหาวิทยาลัยนเรศวร. พิษณุโลก.
- นงลักษณ์ วิรัชชัย. (2538). ความสัมพันธ์โครงสร้างเชิงเส้น (Lisrel) สถิติวิเคราะห์สำหรับการวิจัยทางสังคมศาสตร์และพฤติกรรมศาสตร์ (พิมพ์ครั้งที่ 1). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- นำชัย ศุภฤกษ์ชัยสกุล. (2557). การประยุกต์ใช้การวิเคราะห์แบบจำลองสมการโครงสร้างในงานวิจัยเชิงทดลอง. วารสารพฤติกรรมศาสตร์, 20(2), 206-237.
- นุชริน ไบโพธิ์. (2544). การเปรียบเทียบความสอดคล้องของผลการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันระหว่างวิธี Lord's χ^2 วิธี Mantel – Haenszel และวิธี SIBTEST (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ). มหาวิทยาลัยศรีนครินทรวิโรฒ. กรุงเทพฯ
- ประภคิตยา ทักซิณ. (2552). การประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน : การประยุกต์ใช้การทำหน้าที่ต่างกันของข้อสอบและโมเดลมูลค่าเพิ่ม (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ). จุฬาลงกรณ์มหาวิทยาลัย. กรุงเทพฯ
- ปิยะทิพย์ ดินวรร. (2549). การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ : การเปรียบเทียบประสิทธิภาพระหว่างวิธีการวิเคราะห์องค์ประกอบจำกัดกับวิธีถดถอยโลจิสติก (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ). มหาวิทยาลัยบูรพา. ชลบุรี.
- พีรญา สูงเนินและคณะ. (2549). การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ: การเปรียบเทียบระหว่างรายข้อกับรายหมวดข้อสอบโดยใช้วิธีซิปเทสท์. วิทยาการวิจัยและวิทยาการปัญญา, 6(2), 49-62.
- รติพร ถึงฝั่ง. (2556). การวิเคราะห์โมเดลมิติ : การใช้ประโยชน์จากโปรแกรม LISREL รุ่นทดลองใช้เพื่องานวิจัย. วารสารสมาคมนักวิจัย, 18(2), 128-140.

- เรืองเดช ศิริกิจ. (2554). การวิเคราะห์เปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาวิชา
คณิตศาสตร์ : การประยุกต์ใช้โมเดลมูลค่าเพิ่มที่มีการวิเคราะห์การทำหน้าที่ต่างกันของ
ข้อสอบและการวิเคราะห์การทำหน้าที่ต่างกันของตัวลวง (วิทยานิพนธ์ปริญญาโทศึกษาศาสตร์).
จุฬาลงกรณ์มหาวิทยาลัย. กรุงเทพฯ
- วรรณารอดตัว. (2544). การศึกษาเปรียบเทียบผลของวิธีวิเคราะห์ความลำเอียงของข้อสอบที่
แตกต่างกัน 3 วิธี (วิทยานิพนธ์ปริญญาโทศึกษาศาสตร์). มหาวิทยาลัยรามคำแหง. กรุงเทพฯ
- วรัญญูภรณ์ ชาลีรักษ์. (2558). ปัจจัยที่ส่งผลต่อความคิดเชิงบวกในการทำงานของข้าราชการครู
และบุคลากรทางการศึกษา : การวิเคราะห์ MIMIC. *Veridian E-Journal, Silpakorn
University*, 9(2), 1,336-1,348.
- ศิริชัย กาญจนวาสี. (2555). ทฤษฎีการทดสอบแนวใหม่ (Model test theories) (พิมพ์ครั้งที่ 4).
กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2556). ทฤษฎีการทดสอบแบบดั้งเดิม (Classical test theories)
(พิมพ์ครั้งที่ 7). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศุภวัฒน์ มะลิเผือก. (2548). อิทธิพลของการทำหน้าที่ต่างกันของข้อสอบ ที่ส่งผลต่อคุณภาพของ
แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนระดับชาติ ชั้นประถมศึกษาปีที่ 6 (วิทยานิพนธ์
ปริญญาโทศึกษาศาสตร์). มหาวิทยาลัยบูรพา. ชลบุรี.
- สมศักดิ์ จันทอง. (2542). การเปรียบเทียบผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธี
วิเคราะห์และขนาดกลุ่มผู้สอบต่างกัน (วิทยานิพนธ์ปริญญาโทศึกษาศาสตร์). มหาวิทยาลัย
นเรศวร. พิษณุโลก.
- สุพัฒนา หอมบุปผา. (2556). การเปรียบเทียบการทำหน้าที่ต่างกันของข้อสอบ ด้วยวิธี HGLM วิธี
MIMIC และวิธี BAYESIAN. *วารสารวิจัยราชภัฏพระนคร*, 8(2), 10-24.
- สุภาภรณ์ แดงเพ็ง. (2553). การเปรียบเทียบความลำเอียงของแบบทดสอบคณิตศาสตร์ในการ
ประเมินคุณภาพการศึกษาระดับท้องถิ่น ในชั้นประถมศึกษาปีที่ 5 สังกัดสำนักงานเขตพื้นที่
การศึกษาปัตตานี เขต 2 ระหว่างวิธีแมนเทิล-แฮนส์เซลและโค้งลักษณะข้อสอบ 3
พารามิเตอร์. (วิทยานิพนธ์ปริญญาโทศึกษาศาสตร์). มหาวิทยาลัยสงขลานครินทร์. ปัตตานี.
- สุมาลี แก้วทองค์. (2547). สาเหตุของการทำหน้าที่ต่างกันของข้อสอบสาระการเรียนรู้ภาษาไทย
และสาระการเรียนรู้สังคมศึกษา ศาสนาและวัฒนธรรม (วิทยานิพนธ์ปริญญาโทศึกษาศาสตร์).
จุฬาลงกรณ์มหาวิทยาลัย. กรุงเทพฯ
- อัญชลี ธีระวุฒิ. (2555). การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ Pre O-net วิชาคณิตศาสตร์
ชั้นมัธยมศึกษาปีที่ 6 ของสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 35 (วิทยานิพนธ์
ปริญญาโทศึกษาศาสตร์). มหาวิทยาลัยราชภัฏลำปาง. ลำปาง.
- อิทธิฤทธิ์ พงษ์ปิยะรัตน์. (2551). การวิเคราะห์ข้อสอบและการตรวจสอบการทำหน้าที่ต่างกันของ
ข้อสอบ : การวิเคราะห์พุทธระดับ (วิทยานิพนธ์ปริญญาโทศึกษาศาสตร์). จุฬาลงกรณ์
มหาวิทยาลัย. กรุงเทพฯ

- Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. The Guilford Press, New York, NY.
- Carroll, L. A. (2014). *MIMIC DIF Testing When the Latent Variable Variance Differs Between Group*. (Master's thesis). University of Kansas, Faculty of the Graduate School of Arts.
- Finch, Holmes. (2005). The MIMIC Model as a Method for Detecting DIF: Comparison With Mantel-Haenszel, SIBTEST, and the IRT Likelihood Ratio. *Applied Psychological Measurement*, 29, 278-295. DOI: 10.1177/0146621605275728.
- Kamata, A., & Bauer, D.J., (2008). *A Note on the Relation Between Factor Analysis and Item Response Theory Models*. *Structural Equation Modeling*. 15, p.136 - 153.
- Woods, C. M. & Grimm, K. J. (2011). *Testing for Nonuniform Differential Item Functioning With Multiple Indicator Multiple Cause Models*. *Applied Psychological Measurement*, p.339-361.
- Woods, C. M. (2009). *Evaluation of MIMIC-Model Methods for DIF Testing With Comparison to Two-Group Analysis*. *Multivariate Behavioral Research*, p.1-27.
- Woods, C. M., Oltmanns, T. F. & Turkheimer, E. (2009). *Illustration of MIMIC-Model DIF Testing with the Schedule for Nonadaptive and Adaptive Personality*. *Psychopathol Behav Assess*, p.320-330.

ภาคผนวก
Prince of Songkhla University
Pattani Campus

ภาคผนวก ก

หนังสือขอความอนุเคราะห์

Prince of Songkhla University
Pattani Campus

ที่ ศธ ๐๕๒๑.๒.๐๗๐๖/



ภาควิชาประเมินผลและวิจัยทางการศึกษา
คณะศึกษาศาสตร์ มหาวิทยาลัยสงขลานครินทร์
วิทยาเขตปัตตานี ตำบลรูสมะมิแล อำเภอเมือง
จังหวัดปัตตานี ๙๔๐๐๐

๒๑ มีนาคม ๒๕๖๐

เรื่อง ขอบความอนุเคราะห์ข้อมูลการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET)

เรียน ผู้อำนวยการสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)

ด้วยนายพิชิตชัย สุขเอก นักศึกษาปริญญาโทหลักสูตรศึกษาศาสตรมหาบัณฑิต สาขาวิชาการวิจัยและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี ได้จัดทำวิทยานิพนธ์หัวข้อ เรื่อง การทำหน้าที่ต่างกันของข้อสอบโดย วิธี MIMIC และการวิเคราะห์แนวโน้ม (Detecting Differential Item Functioning Using MIMIC and Trend Analysis) นั้น ในการนี้ภาควิชาฯ จึงขอความอนุเคราะห์ข้อมูลเบื้องต้นของนักเรียนชั้นมัธยมศึกษาปีที่ ๖ ซึ่งได้แก่ เพศ สถานศึกษา ผลการตอบ และคะแนนของนักเรียนชั้นมัธยมศึกษาปีที่ ๖ ปีการศึกษา ๒๕๕๖- ๒๕๕๘ และข้อสอบทั้ง ๕ วิชา ได้แก่ ภาษาไทย คณิตศาสตร์ วิทยาศาสตร์ สังคมศึกษา ศาสนา และวัฒนธรรม และภาษาอังกฤษ ปีการศึกษา ๒๕๕๖- ๒๕๕๘ เพื่อใช้เป็นข้อมูลปฐมภูมิในการจัดทำวิทยานิพนธ์ และจัดส่งข้อมูลผ่านทาง Email : sungkhung@hotmail.com และ โทร ๐๘ ๔๐๖๘๒๐๒๘

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์ จักเป็นพระคุณยิ่ง

ขอแสดงความนับถือ

(ดร.จิระวัฒน์ ตันสกุล)

หัวหน้าภาควิชาประเมินผลและวิจัยทางการศึกษา

สำนักงานเลขาธิการภาควิชา

โทร ๐ ๗๓๓๓ ๗๓๘๑

โทรสาร ๐ ๗๓๓๔ ๘๓๒๒

บันทึกข้อตกลง
ว่าด้วยการขอใช้ประโยชน์จากข้อมูล
ในการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET)
ระหว่าง
สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) กับ นายพิชิตชัย สุขเอก

บันทึกข้อตกลงฉบับนี้ทำขึ้น ณ สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) เมื่อวันที่ ๓๐ เมษายน ๒๕๖๐ ระหว่าง สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) โดย ดร.พรายทอง พวงสันเทียะ ตำแหน่ง รองผู้อำนวยการ ปฏิบัติงานแทนผู้อำนวยการสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) สำนักงานตั้งอยู่ที่ ๑๒๘ อาคารพญาไทพลาซ่า ชั้น ๓๖ แขวงทุ่งพญาไท เขตราชเทวี กรุงเทพมหานคร ซึ่งต่อไปนี้เรียกว่า "ผู้ให้ข้อมูล" ฝ่ายหนึ่ง กับ นายพิชิตชัย สุขเอก อยู่ที่ ๘๑ หมู่ที่ ๖ ตำบลมะกอกเหนือ อำเภอควนขนุน จังหวัดพัทลุง ซึ่งต่อไปนี้เรียกว่า "ผู้ขอข้อมูล" อีกฝ่ายหนึ่ง ทั้งสองฝ่ายได้ตกลงร่วมกันจัดทำบันทึกข้อตกลงฯ ดังนี้

๑. บันทึกข้อตกลงนี้จัดทำขึ้นระหว่าง ผู้ให้ข้อมูล กับ ผู้ขอข้อมูล โดยมีวัตถุประสงค์ที่จะให้ ผู้ขอข้อมูล ใช้ประโยชน์จากข้อมูลในการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ที่ผู้ให้ข้อมูล ได้ดำเนินการจัดการทดสอบและรวบรวมข้อมูลขึ้นโดยข้อมูลที่จะนำไปใช้ประโยชน์ ได้แก่ ข้อมูลผลการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ ๖ ปีการศึกษา ๒๕๕๖ ถึง ปีการศึกษา ๒๕๕๘ หรือข้อมูลเบื้องต้นของนักเรียนชั้นมัธยมศึกษาปีที่ ๖ ได้แก่ เพศ สถานศึกษา ผลการตอบข้อสอบ และคะแนนสอบของนักเรียน

๒. ผู้ขอข้อมูล ตกลงว่าจะนำข้อมูลไปใช้ประกอบการกิจการในการทำวิทยานิพนธ์ เรื่อง "การทำหน้าที่ต่างกันของข้อสอบโดย วิธี MIMIC และการวิเคราะห์แนวโน้ม (Detecting Differential Item Functioning Using MIMIC and Trend Analysis)"

๓. ผู้ขอข้อมูล จะต้องเป็นผู้รับผิดชอบในการจัดทาสีข้อมูลปกคลุมเพื่อใช้ในการบันทึกหรือบรรจุข้อมูลที่ขอใช้ประโยชน์ ตลอดจนรับผิดชอบต่อความเสียหายและค่าใช้จ่ายต่างๆ ที่เกิดขึ้นจากกรณีด้วย

๔. ผู้ขอข้อมูล จะต้องไม่นำข้อมูลไปใช้ผิดจากข้อตกลงหรือนำไปใช้ในลักษณะที่ส่งผลกระทบต่อ หรือก่อให้เกิดความเสียหายแก่เจ้าของข้อมูล หากฝ่าฝืนหรือเกิดความรับผิดชอบทางแพ่งหรืออาญาผู้ขอข้อมูล ต้องเป็นผู้รับผิดชอบในฐานะผู้ก่อให้เกิดความเสียหายขึ้นต้น และเป็นผู้รับผิดชอบทางแพ่งแต่เพียงฝ่ายเดียว

๕. เพื่อให้การปฏิบัติงาน เกิดความคล่องตัวและมีประสิทธิภาพ ผู้ให้ข้อมูล ไดมอบหมายให้ นางสาวสมพร สมพงษ์มิตร และ นายสำรวย ศรีนิวาศร์ เป็นผู้ประสานงานในการดำเนินการตามข้อตกลงนี้

๖. ผู้ให้ข้อมูล ขอสงวนสิทธิ์ในการระงับการใช้ประโยชน์จากข้อมูลในการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) เป็นการชั่วคราวหรือการเปลี่ยนแปลงวิธีการขอใช้ประโยชน์ข้อมูลได้ตามที่เห็นสมควร

/s/ หากผู้ขอ

๒

๗. หากผู้ขอข้อมูลไม่ปฏิบัติตามหลักเกณฑ์หรือวิธีการปฏิบัติที่ได้กำหนดไว้ในบันทึกข้อตกลงฉบับนี้ ผู้ให้ข้อมูล มีสิทธิยกเลิกหรือระงับการใช้ประโยชน์จากข้อมูลในการทดสอบทางการศึกษาระดับชาตินิยมพื้นฐาน (O-NET) ได้ในทันทีที่ตรวจสอบพบ โดยไม่ต้องแจ้งให้ทราบล่วงหน้า

๘. เมื่อการใช้ประโยชน์จากข้อมูลในการทดสอบทางการศึกษาระดับชาตินิยมพื้นฐาน (O-NET) ตามข้อ ๓ เสร็จสิ้น ผู้ขอข้อมูล ตกลงว่าจะส่งผลงานวิทยานิพนธ์/วิจัย ฉบับสมบูรณ์ ตามข้อ ๒ จำนวน ๓ ชุด ให้แก่ผู้ให้ข้อมูลต่อไป

บันทึกข้อตกลงฉบับนี้มีผลบังคับใช้ตั้งแต่วันที่มีการลงนามในบันทึกข้อตกลงเป็นต้นไป และทำขึ้นเป็นสองฉบับ มีข้อความถูกต้องตรงกันทุกประการ ทั้งสองฝ่ายได้อ่านและเข้าใจข้อความในบันทึกข้อตกลงเป็นอย่างดีแล้ว เห็นว่าถูกต้องตรงตามความประสงค์ จึงได้ลงนามไว้เป็นสำคัญคือหน้าพยาน และเก็บรักษาไว้ฝ่ายละหนึ่งฉบับ

ลงชื่อ.....

(นายพิชิตชัย สุขเอก)
ผู้ขอข้อมูล

ลงชื่อ.....

(ดร.พรายทอง พวงกลิ่นเขียว)
รองผู้อำนวยการ
ปฏิบัติงานแทนผู้อำนวยการ
สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)

ลงชื่อ.....พยาน

(.....)

ลงชื่อ.....พยาน

(.....)

ภาคผนวก ข

Syntax ในการวิเคราะห์ข้อมูลโดยประยุกต์ใช้โปรแกรม Mplus

Prince of Songkhla University
Pattani Campus

1. ตัวอย่าง Syntax ในการวิเคราะห์การประมาณค่าระดับความสามารถของผู้สอบ ค่าพารามิเตอร์ ความยากและอำนาจจำแนกของข้อสอบ โดยประยุกต์ใช้โปรแกรม Mplus

TITLE: item thai 56

DATA: FILE IS "D:\thesis\analysis\thai_2556.txt";

VARIABLE:

NAMES ARE t1 t2 t3 t4 t5 t6 t7 t8 t9 t10 t11 t12 t13 t14 t15 t16
t17 t18 t19 t20 t21 t22 t23 t24 t25 t26 t27 t28 t29 t30 t31 t32 t33
t34 t35 t36 t37 t38 t39 t40 t41 t42 t43 t44 t45 t46 t47 t48 t49 t50
t51 t52 t53 t54 t55 t56 t57 t58 t59 t60 t61 t62 t63 t64 t65 t66 t67
t68 t69 t70;

USEVARIABLES ARE t1 t2 t3 t4 t5 t6 t7 t8 t9 t10 t11 t12 t13 t14 t15 t16
t17 t18 t19 t20 t21 t22 t23 t24 t25 t26 t27 t28 t29 t30 t31 t32 t33
t34 t35 t36 t37 t38 t39 t40 t41 t42 t43 t44 t45 t46 t47 t48 t49 t50
t51 t52 t53 t54 t55 t56 t57 t58 t59 t60 t61 t62 t63 t64 t65 t66 t67
t68 t69 t70;

CATEGORICAL = t1 - t70;

ANALYSIS: ESTIMATOR = MLR;

MODEL: f1 BY t1* t2-t70;
f1@1;

SAVEDATA:

file = thetathai56.dat;

save = fscores;

PLOT: TYPE = PLOT3;

2. ตัวอย่าง Syntax ในการวิเคราะห์การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบจำแนกตามเพศ และที่ตั้งของสถานศึกษา โดยประยุกต์ใช้โปรแกรม Mplus

TITLE: sex thai 56

DATA: FILE IS "D:\thesis\analysis\sexthai_2556.txt";

VARIABLE:

NAMES ARE sex t1 t2 t3 t4 t5 t6 t7 t8 t9 t10 t11 t12 t13 t14 t15 t16
t17 t18 t19 t20 t21 t22 t23 t24 t25 t26 t27 t28 t29 t30 t31 t32 t33
t34 t35 t36 t37 t38 t39 t40 t41 t42 t43 t44 t45 t46 t47 t48 t49 t50
t51 t52 t53 t54 t55 t56 t57 t58 t59 t60 t61 t62 t63 t64 t65 t66 t67
t68 t69 t70;

USEVARIABLES ARE sex t1 t2 t3 t4 t5 t6 t7 t8 t9 t10 t11 t12 t13 t14 t15 t16
t17 t18 t19 t20 t21 t22 t23 t24 t25 t26 t27 t28 t29 t30 t31 t32 t33
t34 t35 t36 t37 t38 t39 t40 t41 t42 t43 t44 t45 t46 t47 t48 t49 t50
t51 t52 t53 t54 t55 t56 t57 t58 t59 t60 t61 t62 t63 t64 t65 t66 t67
t68 t69 t70;

CATEGORICAL = t1 - t70;

ANALYSIS:

ESTIMATOR IS wlsmv;

ITERATIONS = 1000;

CONVERGENCE = 0.00005;

MODEL: f1 by t1-t70;

f1 on sex;

t1-t70 on sex;

OUTPUT: MOD(10);

SAVEDATA: RESULTS IS D:\ thesis\analysis;

TITLE: loc thai 56

DATA: FILE IS "D:\thesis\analysis\locthai_2556.txt";

VARIABLE:

NAMES ARE loc t1 t2 t3 t4 t5 t6 t7 t8 t9 t10 t11 t12 t13 t14 t15 t16
t17 t18 t19 t20 t21 t22 t23 t24 t25 t26 t27 t28 t29 t30 t31 t32 t33
t34 t35 t36 t37 t38 t39 t40 t41 t42 t43 t44 t45 t46 t47 t48 t49 t50
t51 t52 t53 t54 t55 t56 t57 t58 t59 t60 t61 t62 t63 t64 t65 t66 t67
t68 t69 t70;

USEVARIABLES ARE loc t1 t2 t3 t4 t5 t6 t7 t8 t9 t10 t11 t12 t13 t14 t15 t16
t17 t18 t19 t20 t21 t22 t23 t24 t25 t26 t27 t28 t29 t30 t31 t32 t33
t34 t35 t36 t37 t38 t39 t40 t41 t42 t43 t44 t45 t46 t47 t48 t49 t50
t51 t52 t53 t54 t55 t56 t57 t58 t59 t60 t61 t62 t63 t64 t65 t66 t67
t68 t69 t70;

CATEGORICAL = t1 - t70;

ANALYSIS:

ESTIMATOR IS wlsmv;

ITERATIONS = 1000;

CONVERGENCE = 0.00005;

MODEL: f1 by t1-t70;

f1 on loc;

t1-t70 on loc;

OUTPUT: MOD(10);

SAVEDATA: RESULTS IS D:\ thesis\analysis;

ประวัติผู้เขียน

ชื่อ สกุล	นายพิชิตชัย สุขเอก	
รหัสประจำตัวนักศึกษา	5920120262	
วุฒิการศึกษา		
วุฒิ	ชื่อสถาบัน	ปีที่สำเร็จการศึกษา
วิทยาศาสตร์บัณฑิต (ศษ.)	มหาวิทยาลัยสงขลานครินทร์	2558
(คณิตศาสตร์)	วิทยาเขตปัตตานี	

ทุนการศึกษา

- ทุนการศึกษาปริญญาตรี-โท คณะศึกษาศาสตร์ จากเงินกองทุนวิจัย คณะศึกษาศาสตร์
- ทุนอุดหนุนการวิจัยเพื่อวิทยานิพนธ์ จากบัณฑิตวิทยาลัย มหาวิทยาลัยสงขลานครินทร์

ตำแหน่งและสถานที่ทำงาน

ครูผู้ช่วย โรงเรียนบ้านพรุบัว สังกัดสำนักงานสำนักงานเขตพื้นที่การศึกษาประถมศึกษา นครศรีธรรมราช เขต 3 จังหวัดนครศรีธรรมราช

การตีพิมพ์เผยแพร่ผลงาน

- พิชิตชัย สุขเอก และจิระวัฒน์ ต้นสกุล. (2560). การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธี MIMIC. ใน การสัมมนาวิชาการ การวัดผล ประเมินผล และวิจัยสัมพันธ์แห่งประเทศไทย ครั้งที่ 25 (น.258-272). นนทบุรี: มหาวิทยาลัยสุโขทัยธรรมาธิราช.
- พิชิตชัย สุขเอก และจิระวัฒน์ ต้นสกุล. (2560). การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) รายวิชาคณิตศาสตร์ โดยใช้วิธีมิมิค. วารสารศึกษาศาสตร์ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี, 29(3).