การวิเคราะห์เปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ : การประยุกต์ใช้โมเดลมูลค่าเพิ่มที่มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและการวิเคราะห์การทำหน้าที่ต่างกันของตัวลวง

นายเรืองเดช ศิริกิจ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรดุษฎีบัณฑิต
สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2554
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ที่ส่งผ่านทางบัณฑิตวิทยาลัย
The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the Graduate School.

A COMPARATIVE ANALYSIS OF THE MODEL OF A QUALITY ASSESSMENT OF MATHEMATICS SUBJECT: AN APPLICATION OF DIFFERENTIAL ITEM FUNCTIONING AND DIFFERENTIAL DISTRACTOR FUNCTIONING

Mr. Ruangdech Sirikit

A Dissertation Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy Program in Educational Measurement and Evaluation

Department of Educational Research and Psychology

Faculty of Education

Chulalongkorn University

Academic Year 2011

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การวิเคราะห์เปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์: การประยุกต์ใช้โมเดลมูลค่าเพิ่มที่มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและการวิเคราะห์การทำหน้าที่ต่างกันของตัวลวง
โดย	นายเรืองเดช ศิริกิจ
สาขาวิชา	การวัดและประเมินผลการศึกษา
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร.โชติกา ภาษีผล
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาดุษฎีบัณฑิต

…….. คณบดีคณะครุศาสตร์
(ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี)

คณะกรรมการสอบวิทยานิพนธ์

……ประธานกรรมการ
(รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ)

…….. อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(รองศาสตราจารย์ ดร.โชติกา ภาษีผล)

…… อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี)

…… กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.กมลวรรณ ตังธนกานนท์)

…… กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.ณัฏฐภรณ์ หลาวทอง)

…… กรรมการภายนอกมหาวิทยาลัย
(อาจารย์ ดร.สุวิมล กฤชคฤหาสน์)

เรืองเดช ศิริกิจ : การวิเคราะห์เปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ : การประยุกต์ใช้โมเดลมูลค่าเพิ่มที่มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ และการวิเคราะห์การทำหน้าที่ต่างกันของตัวลวง (A COMPARATIVE ANALYSIS OF THE MODEL OF A QUALITY ASSESSMENT OF MATHEMATICS SUBJECT: AN APPLICATION OF DIFFERENTIAL ITEM FUNCTIONING AND DIFFERENTIAL DISTRACTOR FUNCTIONING) อ. ที่ปรึกษาวิทยานิพนธ์หลัก ศ.ดร.โชติกา ภาษีผล อ. ที่ปรึกษาวิทยานิพนธ์ร่วม ศ.ดร.ศิริชัย กาญจนวาสี, 350 หน้า.

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1. เพื่อศึกษาผลของข้อสอบที่เกิดการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) ในแบบสอบประเมินความรู้วิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 2 2. เพื่อศึกษาผลของข้อสอบที่ทำหน้าที่ต่างกัน และผลของตัวแปรคุณลั กษณะของนักเรียนและสถานศึกษาต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน 4 โมเดล และ 3. เพื่อเปรียบเทียบประสิทธิภาพโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ เมื่อใช้โมเดลการประเมินคุณภาพการจัดการ ศึกษา 4 โมเดล ข้อมูลที่ใช้ในการวิจัยครั้งนี้เป็นข้อมูลทุติยภูมิจากโครงการการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์ หรือ TIMSS ปี 2007 กลุ่มตัวอย่างของโครงการประกอบด้วยนักเรียน ครูคณิตศาสตร์ ที่สอนนักเรียนที่เป็นกลุ่มตัวอย่างในการวิจัยและผู้บริหารสถานศึกษาที่นักเรียนกลุ่มตัวอย่างศึกษา จำนวน 150 สถานศึกษา และจำนวนนักเรียน 5,412 คน โดยศึกษาเฉพาะวิชาคณิตศาสตร์ จากแบบสอบประเมินความรู้วิชาคณิตศาสตร์ จำนวน 14 ฉบับ การวิเคราะห์ข้อมูลดำเนินการใน 3 ขั้นตอน ดังนี้ ขั้นตอนที่ 1 การตรวจสอบการทำ หน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง 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 และการทำหน้าที่ต่างกันของตัวลวงในแบบสอบวิชาคณิตศาสตร์ มากที่สุด โดยพบว่าข้อสอบส่วนใหญ่ที่เกิดการทำหน้าที่ต่างกัน (DIF) มีจำนวนข้อสอบที่ เพศหญิงมีโอกาสที่จะตอบถูกมากกว่าเพศชาย ในขณะที่ข้อสอบที่เกิดการทำ หน้าที่ต่างกันของตัวลวง (DDF) ข้อสอบส่วนใหญ่เพศชายมีโอกาสที่จะเลือกตอบมากกว่าเพศหญิง

2. ผลของข้อสอบที่ทำหน้าที่ต่างกัน ในแบบสอบวิชาคณิตศาสตร์ต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ พบว่า ข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบวิชาคณิตศาสตร์ไม่มีผลต่อการประเมิ นคุณภาพการจัดการศึกษา (จากการเปรียบเทียบ โมเดล 1 และ 2, โมเดล 1 และ 4, โมเดล 2 และ 4) ส่วนผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาต่อการประเมินคุณภาพการจัดการศึกษาวิชา คณิตศาสตร์ พบว่าคุณลักษณะของนักเรียนและสถานศึกษา มีผลต่อการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา โดยการการจัดอันดับคุณภาพการจัดการศึกษา ระหว่าง โมเดล 1 และ 3, โมเดล 2 และ 3 และระหว่างโมเดล 3 และ 4 มีความแตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3. การเปรียบเทียบ ประสิทธิภาพ โมเดลการประเมินคุณภาพการจัดการศึกษาวิชา คณิตศาสตร์ พบว่าโมเดลการประเมิ นคุณภาพการจัดการศึกษาวิชา คณิตศาสตร์ที่ มีการควบคุมอิทธิพล ของตัวแปร ระดับนักเรียน และ ระดับสถานศึกษา ให้ผลการประเมินคุณภาพการจัดการศึกษาวิชำ คณิตศาสตร์ใกล้เคียงกัน โดยโมเดล 4 มีค่าสัมประสิทธิ์การทำนาย (R^2) สูงสุดเท่ากับ 52.04% รองลงมาได้แก่โมเดล 1 และโมเดล 2 ซึ่งมีค่าสัมประสิทธิ์การทำนาย (R^2) เท่ากับ 51.96% และ 51.86% ตามลำดับ

ภาควิชา วิจัยและจิตวิทยาการศึกษา ลายมือชื่อนิสิต.....

สาขาวิชา การวัดและประเมินผลการศึกษา ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....

ปีการศึกษา 2554 ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม.....

5184497627 : MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION
KEYWORDS : DIFFERENTIAL DISTRACTOR FUNCTIONING / DIFFERENTIAL ITEM FUNCTIONING / VALUE–ADDED MODEL / QUALITY ASSESSMENT / HLM

RUANGDECH SIRIKIT: A COMPARATIVE ANALYSIS OF THE MODEL OF A QUALITY ASSESSMENT OF MATHEMATICS SUBJECT: AN APPLICATION OF DIFFERENTIAL ITEM FUNCTIONING AND DIFFERENTIAL DISTRACTOR FUNCTIONING. ADVISOR: ASSOC.PROF.SHOTIGA PASIPHOL, Ph.D., CO-ADVISOR: PROF.SIRICHAI KANJANAWASEE, Ph.D., 350 pp.

The purposes of this study were 1. to study the effects of differential item functioning and differential distractor functioning on mathematics test in grade 8. 2. to study the effects of DIF - DDF and effects of student-school characteristics on a quality assessment of mathematics subject when using 4 different assessment models. 3. to compare the efficiency of the model of a quality assessment of mathematics subject. The data used in this research is the secondary data from the 2007 Trends in International Mathematics and Science Study (TIMSS), collected from 5,412 students in 150 schools from Thailand, studying only mathematics subjects. The data were based on 14 assessment tests of knowledge in mathematics. The data analysis was analyzed in three steps. The first step was the detection of the differential item functioning (DIF) and differential distractor functioning (DDF) by using the DDFS program. In the second step, to controlled characteristics of student and school in the quality assessment models of mathematics and was analyzed by applying the value-added model with Hierarchical Linear Models (HLM) with two levels of analysis. Finally in the third step, to compared the quality assessment models of mathematics subject. HLM and SPSS for Windows were used in analyzing the data.

The research results were as follows:

1. Sex was the characteristic that found DIF and DDF more than other characteristics in the mathematics test and for the item that found DIF, most of items that, favors the Female group and for DDF, most of the distractor was relatively more attractive for the male than female.
2. The differential item functioning and differential distractor functioning in mathematics tests did not affect on a quality assessment of mathematics subject. (to compare model 1 and 2, model 1 and 4, model 2 and 4) and the student and school characteristics affect a quality assessment of mathematics subject. There was a statistically significant difference at the .05 (to compare model 1 and 3, model 2 and 3, model 3 and 4)
3. The comparison of the efficiency of quality assessment models of mathematics subject reveal that different models cause the result of the quality assessment to differ. Model 4 gives the highest proportion variance explained (R^2) of 52.03%, while that of Model 1 is 51.96% and Model 2 is 51.86%.

Department : Educational Research and Psychology Student's Signature .....

Field of Study : Educational Measurement and Evaluation Advisor's Signature .....

Academic Year : 2011 Co-advisor's Signature .....

กิตติกรรมประกาศ

การวิจัยในครั้งนี้สำเร็จลงได้ด้วยดีจากความอนุเคราะห์ของท่านผู้มีพระคุณหลายท่าน ซึ่งผู้วิจัยขออนุญาตกล่าวถึงเพื่อแสดงการขอบคุณ ในความกรุณาของทุกท่านดังนี้

ผู้วิจัยขอกราบขอบพระคุณรองศาสตราจารย์ ดร .โชติกา ภาษีผล อาจารย์ที่ปรึกษาวิทยานิพนธ์เป็นอย่างสูง เพราะตลอดระยะเวลาที่ผู้วิจัยศึกษาอยู่ในจุฬาลงกรณ์มหาวิทยาลัย ท่านคอยให้คำปรึกษา ถ่ายทอดความรู้ ให้ความช่วยเหลือ เสียสละเวลา คอยเคี่ยวเข็ญ ให้คำแนะนำที่เป็นประโยชน์เสมอ ท่านมีเทคนิคดีๆมาคอยกระตุ้นให้ผู้วิจัยมีความอดทน และพยายามฟันฝ่าอุปสรรคต่างๆ จนสำเร็จได้

ขอกราบของพระคุณศาสตราจารย์ ดร .ศิริชัย กาญจนวาสี อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม ในความกรุณาที่ท่านได้มอบให้กับผู้วิจัยเสมอ ท่าน ได้ให้ความรู้ มุมมองดีๆ แบ่งปันประสบการณ์มากมาย ทุกครั้งที่ผู้วิจัยได้พูดคุยกับท่าน ทำให้ผู้วิจัยรู้สึกมีพลังในการ เดินหน้าทำงานเสมอๆ ผู้วิจัยจะยึดท่านเป็นต้นแบบในการทำงาน และในการเรียนรู้ตลอดไป

ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ ประธานสอบวิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร .ณัฏฐภรณ์ หลาวทอง ผู้ช่วยศาสตราจารย์ ดร.กมลวรรณ ตังธนกานนท์ และอาจารย์ ดร.สุวิมล กฤชคฤหาสน์ กรรมการสอบวิทยานิพนธ์ ที่ให้คำชี้แนะที่มีคุณค่าต่องานวิทยานิพนธ์ ตรวจทานและแก้ไขงานนี้ จนทำให้งานวิจัยนี้มีความสมบูรณ์ที่สุด ตลอดจนท่านรองศาสตราจารย์ ดร .อวยพร เรืองตระกูล ศาสตราจารย์ ดร .สุวิมล ว่องวาณิช ตลอดจนอาจารย์ประจำ ภาควิชาวิจัยและจิตวิทยาการศึกษาทุกๆ ท่านที่ให้กาลังใจแก่ผู้วิจัยเสมอมา รวมไปถึง Associate Professor Randall David Penfield, Ph.D. สำหรับโปรแกรม DDFS 1.0 และคำปรึกษาที่ส่งมาอย่างรวดเร็วทันใจเสมอ

ขอกราบขอบพระคุณ โรงเรียนเซนต์คาเบรียล โดยเฉพาะอย่างยิ่งท่าน ภราดา ดร .อนุศักดิ์ นิธิภัทราภรณ์ สำหรับมุมมองการในการปฏิบัติงานที่ดี และขอบคุณในความเมตตาของท่านตลอดระยะเวลาที่ทำงานที่โรงเรียนเซนต์คาเบรียล และขอขอบพระคุณคณะศึกษาศาสตร์ มหาวิทยาลัยราชภัฏจันทรเกษมที่สนับสนุน และให้เวลาในการทำวิทยานิพนธ์เรื่องนี้จนสำเร็จ

ขอขอบคุณสำหรับคำแนะนำดีๆ และกำลังใจจากเพื่อน พี่ และน้องภาควิชาวิจัยทุกท่านทั้งในเวลาและนอกเวลา โดยเฉพาะท่านอาจารย์ ดร .ประกฤติยา ทักษิโณ ที่มีคำแนะนำดีๆให้น้องคนนี้เสมอๆ

และที่สำคัญที่สุดผู้วิจัยขอขอบพระคุณ คุณแม่จุฑามาศ ศิริกิจ ผู้ซึ่งมีความรักให้กับผู้วิจัยคอยเป็นกำลังใจให้ เป็นแรงผลักดันให้เดินหน้าอยู่เสมอ

ท้ายที่สุดนี้ผู้วิจัยหวังเป็นอย่างยิ่งว่างานวิจัยเรื่องนี้จะเป็นประโยชน์ในหลายๆทางทั้งเรื่องของการพัฒนาการจัดการศึกษา หรือการเป็นประโยชน์ในการศึกษาสำหรับน้องๆต่อไปในอนาคต

สารบัญ

หน้า

บทคัดย่อภาษาไทย..... ง
บทคัดย่อภาษาอังกฤษ..... จ
กิตติกรรมประกาศ..... ฉ
สารบัญ..... ช
สารบัญตาราง..... ฎ
สารบัญภาพ..... ฒ

บทที่ 1 บทนำ 1
ความเป็นมาและความสำคัญของปัญหา..... 1
วัตถุประสงค์ของการวิจัย 8
สมมติฐานของวิจัย..... 8
ขอบเขตของการวิจัย 10
คำจำกัดความที่ใช้ในการวิจัย 11
ประโยชน์ที่คาดว่าจะได้รับ 14

บทที่ 2 แนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 16

ตอนที่ 1 ความลำเอียงของแบบสอบ 16
1.1 ความเป็นมาของความลำเอียงของแบบสอบ 16
1.2 ความหมายของการทำหน้าที่ต่างกันของข้อสอบและการทำหน้าที่ต่างกันของตัวลวง 19
1.3 ประเภทของการทำหน้าที่ต่างกันของข้อสอบ 21
1.4 หลักการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ 22
1.5 วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง 26
1.6 เกณฑ์สำหรับเปรียบเทียบคุณภาพของวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ 30
1.7 งานวิจัยที่เกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง 32

หน้า

ตอนที่ 2 แนวคิดเกี่ยวกับมูลค่าเพิ่มของการจัดการศึกษา 52
2.1 นิยามและความหมายของมูลค่าเพิ่ม 52
2.2 ขั้นตอนและวิธีการที่ใช้วิเคราะห์มูลค่าเพิ่ม 56
2.3 โมเดลมูลค่าเพิ่ม 60
ตอนที่ 3 การวิเคราะห์พหุระดับ 65
3.1 แนวคิดเกี่ยวกับการวิเคราะห์พหุระดับ 66
3.2 การวิเคราะห์พหุระดับด้วยโมเดลเชิงเส้นตรงระดับลดหลั่น 73
ตอนที่ 4 การประเมินผลนักเรียนระดับนานาชาติ 78
4.1 การศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ พ.ศ.2550 (TIMSS2007) 78
4.2 การประเมินผลนักเรียนระดับนานาชาติ PISA 83
ตอนที่ 5 เอกสารงานวิจัยเกี่ยวกับการประเมินคุณภาพการจัดการศึกษา 87
5.1 ปัจจัยหรือตัวแปรที่ส่งผลต่อคุณภาพการจัดการศึกษา 88
5.2 งานวิจัยที่ศึกษาเกี่ยวกับตัวแปรที่มีอิทธิพลต่อคุณภาพการจัดการศึกษาและการใช้โมเดลมูลค่าเพิ่มเพื่อประเมินคุณภาพการจัดการศึกษา 94
5.3 การกำหนดกรอบแนวคิดในการวิจัย 109

บทที่ 3 วิธีดำเนินการวิจัย 111

ตอนที่ 1 ข้อมูลที่ใช้ในการวิจัย 111
ตอนที่ 2 ตัวแปรที่ใช้ในการวิจัย 120
ตอนที่ 3 การวิเคราะห์ข้อมูล 125

บทที่ 4 ผลการวิเคราะห์ข้อมูล 133
ตอนที่ 1 ผลการวิเคราะห์ข้อมูลพื้นฐาน การตรวจสอบเงื่อนไขการวิเคราะห์พหุระดับและการตรวจสอบความเป็นเอกมิติของแบบสอบ 134
1.1 ข้อมูลพื้นฐานของกลุ่มตัวอย่าง 134
1.2 การตรวจสอบเงื่อนไขของการวิเคราะห์พหุระดับ 144

หน้า

ตอนที่ 2 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวงแต่ละฉบับ 150
2.1 สรุปผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง 150
2.2 ผลการวิเคราะห์ลักษณะของข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวงในแต่ละคุณลักษณะของนักเรียน. 157
ตอนที่ 3 ผลการวิเคราะห์ตามโมเดลกรประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ของสถานศึกษาขั้นพื้นฐาน..... 160
3.1 ลักษณะโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์4 โมเดล 161
3.2 ผลการประมาณค่าความสามารถของผู้สอบในแต่ละโมเดล 171
3.3 ผลของข้อสอบที่ทำหน้าที่ต่างกันต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน 174
3.4 ผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน 180
ตอนที่4 ผลการเปรียบเทียบประสิทธิภาพโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ 185
4.1 ผลการวิเคราะห์คุณลักษณะทั่วไปของโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ 4 โมเดล 185
4.2 ผลการเปรียบเทียบประสิทธิภาพโมเดลการประเมินคุณภาพการจัดการศึกษาต่อคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ 193

บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ 199
5.1 สรุปผลการวิจัย 200
5.2 อภิปรายผลการวิจัย 210
5.3 ข้อเสนอแนะ 219

หน้า

รายการอ้างอิง 222

ภาคผนวก 228

ภาคผนวก ก ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง ฉบับที่ 1 – 14 229

ภาคผนวก ข ผลการวิเคราะห์ของโมเดลการวิเคราะห์มูลค่าเพิ่มด้วยโปรแกรม Hierarchical linear model (HLM) (แสดงเฉพาะส่วนที่สำคัญ) 325

ประวัติผู้เขียนวิทยานิพนธ์ 350

สารบัญตาราง

ตารางที่		หน้า
1	คุณภาพของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ……..	31
2	แสดงตัวแปร 3 ระดับ ……	75
3	แสดงโมเดลหลักสูตรของ TIMSS 2007 ……..	79
4	แสดงร้อยละของเนื้อหาในวิชาคณิตศาสตร์ของการประเมิน TIMSS 2007 …..	79
5	สรุปความหมายของตัวแปรที่ใช้ในโครงการประเมินผล TIMSS 2007 ……..	80
6	ผลการสังเคราะห์ตัวแปรทำนายที่มีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียน จำแนกตามระดับของตัวแปร ……	106
7	อันดับของตัวแปรที่นิยมนามาศึกษาปัจจัยที่มีผลต่อคุณภาพการจัดการศึกษา และแหล่งข้อมูลในการเก็บข้อมูลของโครงการประเมินผล TIMSS 2007 …….	108
8	กรอบประชากรนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำแนกตามสังกัด ……	112
9	จำนวนกลุ่มตัวอย่างสำหรับการเก็บรวบรวมข้อมูลเพื่อการประเมิน …..……	113
10	แสดงการจำแนกคะแนนของแบบทดสอบวิชาคณิตศาสตร์โครงการ TIMSS 2007 ตามด้านเนื้อหา (Content Domains) และด้านพฤติกรรมการเรียนรู้ (Cognitive Domain) ……..…	114
11	จำนวนข้อสอบวิชาคณิตศาสตร์แต่ละฉบับ จำแนกตาม Cluster ……..	114
12	ร้อยละของเนื้อหาที่ใช้ในการประเมินโครงการ TIMSS 2007 ……..	115
13	หัวข้อการเรียนรู้จำแนกแต่ละเนื้อหาที่ใช้ในการประเมินโครงการTIMSS 2007 .	115
14	ร้อยละของพฤติกรรมการเรียนรู้ที่ใช้ในการประเมินโครงการ TIMSS 2007 ……	118
15	สเกลการวัดและความหมายของตัวแปรในระดับนักเรียนและสถานศึกษา	122
16	สรุปสมการวิเคราะห์มูลค่าเพิ่มด้วยโมเดลเชิงเส้นแบบลดหลั่นด้วยการวิเคราะห์ 2 ระดับ ในโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา 4 โมเดล ……..……	130
17	แสดงความถี่ ร้อยละ ของตัวแปรจัดประเภทจำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา ……..	135
18	แสดงค่าสถิติพื้นฐานของตัวแปรต่อเนื่องจากการศึกษาแนวโน้มผลสัมฤทธิ์ทางการ เรียนคณิตศาสตร์นานาชาติ (TIMSS 2007) จำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา ….……	139

ตารางที่ **หน้า**

19 แสดงค่าสถิติพื้นฐานของคะแนนความสามารถรายบุคคลในวิชาคณิตศาสตร์จากการศึกษาแนวโน้มผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์นานาชาติ (TIMSS 2007) จำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา 142

20 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรในระดับนักเรียน 147

21 ค่า Tolerance และ VIF จากผลการทดสอบความสัมพันธ์ระหว่างตัวแปรต้นระดับนักเรียน 148

22 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรในระดับสถานศึกษา 148

23 ค่า Tolerance และ VIF จากผลการทดสอบความสัมพันธ์ระหว่างตัวแปรต้นระดับสถานศึกษา 149

24 จำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง..... 152

25 จำนวนข้อสอบที่ทำหน้าที่ต่างกันที่ตัดออกไปในแบบสอบทั้ง 14 ฉบับ 155

26 การเปรียบเทียบโครงสร้างของร้อยละของแบบสอบก่อนตัดและหลังตัดข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวงในแบบสอบทั้ง 14 ฉบับ 156

27 สรุปลักษณะของข้อสอบที่ทำหน้าที่ต่างกันในแต่ละคุณลักษณะของนักเรียน... 160

28 สรุปค่าสัมประสิทธ์ของตัวแปรคุณลักษณะระดับนักเรียนและสถานศึกษาที่ได้จากการวิเคราะห์โมเดล 1 ถึง 4 169

29 แสดงค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบก่อนการตัดข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบฉบับที่ 1 ถึง 14 172

30 แสดงค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบหลังการตัดข้อสอบที่ทำหน้าที่ต่างกัน (DIF) พิจารณาเฉพาะตัวเลือกที่เป็นตัวถูก ในแบบสอบฉบับที่ 1 ถึง 14 173

31 แสดงค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบหลังการตัดข้อสอบที่ทำหน้าที่ต่างกันของตัวลวง (DDF) พิจารณาการเกิดการทำหน้าที่ต่างทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง ในแบบสอบฉบับที่ 1 ถึง 14 174

ตารางที่ **หน้า**

32 เปรียบเทียบผลการจัดระดับคุณภาพการจัดการศึกษา การจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษา ระหว่างโมเดล1 (Undetected DIF-DDF&Adjusted) กับโมเดล 2 (Detect DIF&Adjusted) และค่าสถิติทดสอบWilcoxon Signed Ranks Test 176

33 เปรียบเทียบผลการจัดระดับคุณภาพการจัดการศึกษา การจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษา ระหว่างโมเดล1 (Undetected DIF-DDF&Adjusted) กับโมเดล4 (Detect DIF-DDF&Adjusted) และค่าสถิติทดสอบWilcoxon Signed Ranks Test 178

34 เปรียบเทียบผลการจัดระดับคุณภาพการจัดการศึกษา การจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษา ระหว่างโมเดล 2 (Detect DIF&Adjusted) กับโมเดล 4 (Detect DIF-DDF&Adjusted) และค่าสถิติทดสอบWilcoxon Signed Ranks Test 179

35 เปรียบเทียบผลการจัดระดับคุณภาพการจัดการศึกษา การจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษา ระหว่างโมเดล1 (Undetected DIF-DDF&Adjusted) กับโมเดล3 (Detect DIF-DDF&Unadjusted) และค่าสถิติทดสอบWilcoxon Signed Ranks Test 182

36 เปรียบเทียบผลการจัดระดับคุณภาพการจัดการศึกษา การจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษา ระหว่างโมเดล 2 (Detect DIF & Adjusted) กับโมเดล 3 (Detect DIF-DDF & Unadjusted) และค่าสถิติทดสอบ Wilcoxon Signed Ranks Test 183

37 เปรียบเทียบผลการจัดระดับคุณภาพการจัดการศึกษา การจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษา ระหว่างโมเดล3 (Detect DIF-DDF & Unadjusted) กับโมเดล4 (Detect DIF-DDF & Adjusted) และค่าสถิติทดสอบ Wilcoxon Signed Ranks Test 185

38 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มวิชาคณิตศาสตร์ของสถานศึกษาจากผลการวิเคราะห์ 4 โมเดล 186

39 ผลการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาจากผลการวิเคราะห์ 4 โมเดล จำแนกตามคุณลักษณะของสถานศึกษา 190

ตารางที่ **หน้า**

40 สรุปค่าความแปรปรวนของตัวแปรตามที่อธิบายได้ของโมเดลที่ 1 ถึง 4 194

41 แสดงผลการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาโดย
เปรียบเทียบกับโมเดล 4 196

42 ผลการวิเคราะห์ความสอดคล้องของการจัดกลุ่มคุณภาพการจัดการศึกษา
ด้วยโมเดลการประเมินคุณภาพการจัดการศึกษาที่ต่างกัน 4 โมเดล 197

43 ค่าสหสัมพันธ์ของอันดับของคะแนนมูลค่าเพิ่มที่ได้จากโมเดลประเมิน
คุณภาพการจัดการศึกษาที่ต่างกัน 4 โมเดล 198

44 สรุปความสอดคล้องของผลการประเมิน สัมประสิทธิ์การทำนายและค่า
สหสัมพันธ์ของตำแหน่งคะแนนมูลค่าเพิ่มจากผลการประเมิน 4 โมเดล 208

45 ตารางที่ 45 จำนวนข้อสอบที่ถูกตัดและช่วงค่า Z(LOR) ที่ใช้ในการพิจารณา
การเกิด DIF และ DDF 232

สารบัญภาพ

ภาพที่ **หน้า**

1 โครงสร้างโมเดล 3 ระดับและแนวทางการวิเคราะห์ 75
2 กรอบแนวคิดในการวิจัยเพื่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสต 110
3 ฮีสโตแกรม (Histogram) ของคะแนนมูลค่าเพิ่มจากโมเดลการประเมิน
คุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ของสถานศึกษาขั้นพื้นฐาน
จำนวน 150 แห่ง 187

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

คุณภาพและมาตรฐาน ทางการศึกษาเป็นสิ่งที่ทุก ฝ่ายที่เกี่ยวข้อง ล้วน แล้วแต่ต้องการให้เกิดขึ้นกับการจัดการศึกษาของ สถานศึกษาใน ประเทศไทย ดังจะเห็นได้จากในพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 (ฉบับปรับปรุงแก้ไข พ.ศ. 2545) ในมาตราที่ 4 ให้คำจำกัดความของ “มาตรฐานการศึกษา” ว่าหมายถึงข้อกำหนดเกี่ยวกับคุณลักษณะ คุณภาพที่พึงประสงค์และมาตรฐานที่ต้องการให้เกิดขึ้นในสถานศึกษาทุกแห่ง และเพื่อใช้เป็นหลักในการเทียบเคียงสำหรับการส่งเสริมและกำกับดูแล การตรวจสอบ การประเมินผล และการป ระกันคุณภาพทางการศึกษา จากความหมายดังกล่าวเป็นการยืนยันที่ชัดเจนถึงเจตนารมณ์ของพระราชบัญญัติการศึกษาแห่งชาติที่ชัดเจนที่ต้องการเห็นสถานศึกษาทั่วประเทศไทยสามารถจัดการศึกษาได้อย่างมีคุณภาพ และเป็นมาตรฐานเดียวกัน

สำหรับการตรวจสอบคุณภาพ การจัดการศึกษาของสถานศึ กษานั้น คะแนนของนักเรียนจากการทำแบบสอบเป็นตัวบ่งชี้ สำคัญตัวหนึ่ง เพื่อให้ทราบถึงผลสัมฤทธิ์ของหลักสูตรที่ใช้ในการจัดการเรียนการสอน รวมไปถึงผลสัมฤทธิ์ของการดำเนินการของสถานศึกษา (Postlethwaite, 2004) ดังนั้นหลังจากที่นักเรียนผ่านการทดสอบแล้ว การประเมินผลหลัง การทดสอบถือว่าเป็นเรื่องที่มีความสำคัญ เพื่อให้ได้ข้อมูลการประเมินที่บ่งบอกถึงสารสนเทศที่ สำคัญในการจัดการศึกษาของสถานศึกษา รวมไปถึงจุดอ่อนหรือจุดแข็งที่ควรต้องมีการพัฒนาหรือปรับปรุงแก้ไข จากความสำคัญดังกล่าว ทำให้ผู้ประเมินที่มีหน้าที่ ในการประเมินผลการทดสอบ นั้นต้องเป็นหน่วยงานที่ไม่มีส่วนได้ส่วนเสียกับผลการประเมิน หรือไม่เกี่ยวข้องกับ สถานศึกษาใดๆ และเป็นหน่วยงานหรือบุคคลที่ได้รับการยอมรับจากทุกฝ่ายที่เกี่ยวข้อง เพราะจะต้องทำหน้าที่ในการตัดสินคุณภาพของ สถานศึกษา ต่างๆอย่าง เป็นกลางและมีความ ยุติธรรม สำหรับในประเท ศไทยหน่วยงานที่ทำหน้าที่ดังที่กล่าวมาที่เห็นได้อย่างชัดเจนที่สุด คือสถาบันทดสอบทางการศึกษาแห่งชาติ “สทศ.” (องค์การมหาชน) และสำนักงานรับรองมาตรฐานและประเมินคุณภาพทางการศึกษา “สมศ.” (องค์การมหาชน) เป็นต้น ส่วนในต่างประเทศนั้น มีมากมายหลายหน่วยงาน

เช่น สถาบันทดสอบทางการศึกษา (Educational Testing Service; ETS) องค์กรเพื่อความร่วมมือและพัฒนาทางเศรษฐกิจ (Organization for Economic Co-operation and Development) เป็นหน่วยงานที่รับผิดชอบดำเนินการ จัดโครงการประเมินผลนักเรียนระดับนานาชาติ (Program for International Student Assessment; PISA) หรือสมาคมการประเมินผลการศึกษานานาชาติ (The International Association for the Evaluation of Educational Achievement; IEA) เป็นหน่วยงานที่รับผิดชอบดำเนินการจัด โครงการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ (Trends in International Mathematics and Science Study; TIMSS) เป็นต้น

จากการที่หน่วยงานที่เกี่ยวข้องกับการทดสอบ และการประเมินผลทางการศึกษาต่างๆ ได้เกิดขึ้นอย่างมากมาย จึงสามารถสรุปได้ว่า ปัจจุบันการทดสอบเป็นสิ่งที่สำคัญอย่างยิ่งในสังคมยุคใหม่ โดย ถูกนำมาใช้ในตัดสินผลในประเด็ นหรือ สถานการณ์ที่ สำคัญ ๆ อย่างมากมาย และหลากหลาย เช่น การคัดเลือกนักเรียนเข้าสู่มหาวิทยาลัย หรือการคัด เลือกคนเข้าทำงาน เป็นต้น ดังนั้นการทดสอบที่มีมาตรฐานจึงเป็นสิ่ง จำเป็นที่ผู้เกี่ยวข้องกับผลกา รทดสอบทุกฝ่ายต้องการให้เกิดขึ้นเพื่อนำไปสู่ผลของการทดสอบที่ออกม าให้ที่ยอมรับแก่ ผู้ใช้ผลการทดสอบรวมไปถึง คนในสังคมได้ องค์ประกอบที่สำคัญมากสำหรับการทดสอบนั่นคือเครื่องมือที่ใช้ซึ่งส่วนใหญ่แล้วเครื่องที่ใช้ในการทดสอบก็คือแบบสอบ

คุณภาพของแบบสอบ จึงเป็นสิ่งที่ต้องคำนึงถึงเป็นอันดับแรก การวิเคราะห์ ความเที่ยง และความตรง ก็เป็นวิธีการที่สามารถให้สารสนเทศในเรื่องคุณภาพของแบบสอบได้ การใช้แบบสอบเพื่อวัดลักษณะภายในของบุคคลนั้นคุณ ลักษณะสำคัญของแบบสอบจึงอยู่ที่คำถามสำคัญ 2 ข้อ ได้แก่ คะแนนจากแบบสอบมีความคงเส้นคงวามากน้อยเพียงไร และคะแนนจากแบบสอบสามารถใช้แสดงถึงลักษณะหรือสิ่งที่มุ่งวั ดได้ดีเพียงไร ซึ่งทั้งสองคำถามเ ป็นเรื่องที่เกี่ยวกับความเที่ยงและความตรงของแบบสอบ โดยที่ ความสัมพันธ์ด้านหนึ่งระหว่าง ความเที่ยงกับความตรงนั่นคือ สัมประสิทธิ์ของความเที่ยงที่สูงบ่งบอกถึงความคงเส้นคงวาของคะแนนสอบที่ได้จากการทำการสอบเมื่อทำการสอบซ้ำๆ แต่มิได้ป ระกันว่าคะแนนสอบสามารถใช้สรุปอ้างอิงถึงลักษณะหรือสิ่งที่มุ่งวัดได้อย่างถูกต้อง (ศิริชัย กาญจนวาสี, 2548) ดังนั้นจากความหมายดังกล่าว อาจจะสรุปได้ว่าความตรงเป็นสิ่งที่สำคัญ ที่ต้องคำนึงถึงเป็นอันดับแรก โดย สามารถนิยามความสำคัญของความตรงว่า “ความตรง” เป็นหัวใจข องแบบสอบ การตรวจสอบความตรงของแบบสอบมี 3

ประเภทหลัก คือ 1. ความตรงตามเนื้อเรื่อง (Content validity) 2. ความตรงตามเกณฑ์สัมพันธ์ (Criterion-related validity) 3. ความตรงเชิงทฤษฎี (Construct validity) และการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (differential item functioning; DIF) ก็เป็นอีกประเภทหนึ่งที่ใช้ตรวจสอบคุณภาพด้านความตรงของข้อสอบ (Camilli & Shepard, 1994) โดยจะเห็นได้ว่าการศึกษาเกี่ยวกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบของนักวัดผล ในปัจจุบัน มีผู้ที่ให้ความสนใจเป็นอย่างมาก ทั้งนี้ส่วนหนึ่งเป็นเพราะใน สังคมปัจจุบันมีการแข่งขันโดยใช้แบบสอบมาตรฐานเป็นเครื่องมือในการตัดสินค่อนข้างมาก ดังนั้นการตรวจสอบว่าแบบสอบที่ใช้ในสถานการณ์ต่างๆ มีความยุติธรรมหรือไม่ต่อผู้ที่มีส่วนได้ส่วนเสียจึงถือว่าเป็นเรื่องที่มีความสำคัญ

เครื่องมือที่ถูกใช้ในการทดสอบอย่างแพร่หลาย คือข้อสอบแบบหลายตัวเลือกหรือข้อสอบปรนัย โดยถูกนำมาใช้สำหรับสถานการณ์การทดสอบที่สำคัญต่างๆ เช่นการสอบวัดผลการเรียนรู้ในสถานศึกษา หรือการสอบคัดเลือกนักเรียนเข้าศึกษาต่อในระดับมหาวิทยาลัย เป็นต้น สาเหตุที่แบบสอบแบบหลายตัวเลือกได้รับความนิยมในการนำไปใช้ มาจากลักษณะเด่นของข้อสอบชนิดนี้ นั่นคือ สามารถนำไปใช้ได้กับการ ทดสอบที่มีผู้สอบเป็นจำนวนมาก ภายในเวลาที่จำกัด และสามารถตรวจสอบคุณลักษณะแฝง (Latent trait) ของผู้สอบได้ตรงตามความต้องการของผู้ทดสอบ โดยจากการที่ข้อสอบแบบหลายตัวเลือกถูกใช้ในการทดสอบในสถานการณ์ต่างๆ อย่ าง มากมาย ดังนั้นการควบคุมคุณภาพของแบบสอบให้ปราศจากความลำเอียงจึงเป็นเรื่องที่มีสำคัญอย่างมาก

การศึกษาเรื่องความลำเอียงของข้อสอบ (Item bias) เริ่มต้นตั้งแต่ในช่วง ค.ศ. 1960 โดยเป็นการออกแบบเพื่อหาวิธีที่จะพัฒนาการทดสอบเพื่อให้เหมาะสมสำหรับความแตกต่างในด้านต่างๆของสังคม เช่น วัฒนธรรม สีผิว สังคม เพศ หรืออายุ เป็นต้น (Angoff, 1993) ตั้งแต่อดีตที่ผ่านมามีการเรียกร้องมากมายเพื่อที่จะต่อต้านความไม่เป็นธรรมของแบบสอบในสถานการณ์ต่างๆ และนำมาซึ่งการฟ้องร้องไปยังองค์กรต่างๆที่เกี่ยวข้อง เช่น NAACP เพื่อที่ให้หยุดหรือยกเลิกการใช้แบบสอบชนิดนั้นๆ (Rudner, Getson and Knight cited in Hambleton, Swaminathan and Rogers, 1991) ในปัจจุบันวิธีการตรวจสอบว่าข้อสอบมีการทำหน้าที่ต่างกันหรือไม่ มีการพัฒนาไปมากวิธีการใหม่ๆถูกคิดค้นและ นำมาใช้ในการตรวจสอบ โดยแต่ละวิธี จะมีลักษณะของเงื่อนไขการนำไปใช้ที่แตกต่างกัน เช่น ใช้กับข้อสอบแบบตรวจให้คะแนนสองค่า หรือ

แบบตรวจให้คะแนนหลายค่า เป็นต้น สำหรับประเทศไทยการทำหน้าที่ต่างกันของข้อสอบ ยังเป็นปัญหาที่เกิดขึ้นไม่มาก ผู้สอบส่วนใหญ่ไม่ได้สนใจเรื่องนี้มากนักนั่นอาจเป็นเพราะบริบทของประเทศไทยไม่ได้มีค วามแตกต่างทางด้านสังคมมากนัก แต่ในความเป็นจริงแล้วการทำหน้าที่ต่างกันของข้อสอบไม่ได้เกิดขึ้นเพราะความแตกต่างทางสังคมเท่านั้น แต่ยังมีประเด็นอื่นๆที่สามารถส่งผลให้ข้อสอบมีการทำหน้าที่แตกต่างกัน เช่น เพศ, อายุ หรือประสบการณ์ เป็นต้น

จากการที่แนวคิดในเรื่องข องการทำหน้าที่ต่างกันของข้อสอบได้รับการพัฒนามาอย่างต่อเนื่อง นักวัดผลส่วนหนึ่งได้มีการขยายองค์ความรู้โดยมีวัตถุประสงค์เพื่อให้แบบสอบมีมาตรฐานมากที่สุด โดยทั่วไปนักวัดผลจะให้ความสำคัญกับตัวเลือกที่ถูกต้องมากกว่าตัวเลือกที่เป็นตัวลวง แต่ก็มีนักวัดผลบางส่วนที่ เชื่อว่าการศึกษาทั้งตัวเลือกที่เป็นข้อถูก เพียงอย่างเดียวอาจจะทำให้นักวัดผลได้สารสนเทศที่ยังไม่ครบถ้วน โดยหากมีการเพิ่มการศึกษาเกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบในอีกมิติหนึ่ง คือทำการวิเคราะห์ ตัวเลือกที่เป็นตัวลวงจะช่วยเพิ่มสารสนเทศที่สำคัญ แต่ทำให้ผลกา รวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบมีความสำคัญมากยิ่งขึ้น ซึ่งจะทำให้การทดสอบเกิดความยุติธรรมกับผู้สอบอย่างสูงสุด

การมุ่งศึกษาเพิ่มเติมในเรื่องของการทำหน้าที่ต่างกันของตัวลวง (Differential Distractor Functioning; DDF) เป็นการขยายแนวคิดของการทำหน้าที่ต่ างกันของข้อสอบจากการศึกษาตัวถูกไปสู่การศึกษาตัวลวง (Bank, 2009) โดยประโยชน์ของการศึกษาในเรื่องการทำหน้าที่ต่างกันของตัวลวง ทำให้ช่วยเพิ่มมิติในการตรวจสอบการทำหน้าที่ของข้อสอบให้มีความหลากหลาย และเกิดความยุติธรรมกับผู้สอบมากที่สุด ประสิทธิภาพของตัวลวงสำหรับ แบบสอบประเภทหลายตัวเลือก (Multiple Choices) ยังคงเป็นสิ่งที่ต้องคำนึงถึง ตามนัยทางทฤษฎีผู้ที่เลือกตัวลวงซึ่งถือว่าเป็นคำตอบที่ผิด จะเป็นเครื่องแสดงว่าผู้นั้นยังไม่สัมฤทธิ์ผลตามเป้าหมายของการวัด (ศิริชัย กาญจนวาสี, 2548) ซึ่งสอดคล้องกับ Dorans, Schmitt และ Bleistein (อ้างถึงใน Penfield, 2008) ได้กล่าวถึงการวิเคราะห์การทำหน้าที่ของข้อสอบที่สมบูรณ์นั้นต้องประกอบไปด้วยการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง

การประเมินคุณภาพการศึกษาจากผลสัมฤทธิ์นั้นมีความสอดคล้องกับโมเดลคะแนนจริ งแบบดั้งเดิม กล่าวคือคะแนนที่ได้จากการวัดหรือคะแนนสังเกตได้ (Observed score, X) = คะแนน

จริง (True score, T) + คะแนนความคลาดเคลื่อน (Error score, E) ดังนั้นหากมีคะแนนส่วนที่เป็นคะแนนความคลาดเคลื่อนเกิดขึ้น คะแนนที่ได้จากการสอบก็ไม่สามารถบอกได้ว่าเป็นความสามารถที่แท้จริงของนักเรียนทั้งหมด โดยอาจจะเกิดจากปัจจัยหลายๆด้านที่ส่งผลกระทบต่อการทำคะแนนสอบของนักเรียน และจากสาเหตุดังกล่าวจึงทำให้เกิดการศึกษาเกี่ยวกับวิธีการตรวจสอบสาระสำคัญที่ควรได้รับจากผลคะแนนที่ได้จากการวัดเกิดขึ้นโดยมีความต้องการที่จะพัฒนาคุณภาพการปฏิบัติงานสู่ระดับมาตรฐาน ที่ต้องการใช้ตัวบ่งชี้ผลการปฏิบัติงานได้อย่างมีความชัดเจนและถูกต้อง รวมไปถึงเพื่อให้ผลการวัดมีความตรงในการวิเคราะห์เกิดขึ้นมากที่สุด

โมเดลมูลค่าเพิ่ม (Value-added Model) เป็นวิธีการที่ช่วยรายงานผลเพื่อสะท้อนข้อมูลเกี่ยวกับการจัดการ ศึกษา ว่าสถานศึกษา ได้สร้างมูลค่าเพิ่มในผลการเรียนรู้ได้มากน้อยเพียงใด โดยเปรียบเทียบระหว่างคะแนนผลการเรียนรู้ที่เกิดขึ้นจริงหรือสังเกตได้ (Observed scores) กับคะแนนผลการเรียนรู้ที่ทำนายได้ (Predicted scores) จากตัวแปรภูมิหลังของนักเรียน ตัวแปรบริบทของชุมชน สังคม หรือผลสัมฤทธิ์เดิม (ศิริชัย กาญจนวาสี, 2550) การนำโมเดลมูลค่าเพิ่มมาใช้ทางการศึกษาเป็นการรวบรวมเทคนิคทางสถิติที่ใช้คะแนนจากการสอบของนักเรียนเพื่อประมาณค่าอิทธิพลของสถานศึกษาหรือครู (MaCaffrey, Lockwood, Koretz, & Hamilton, 2003) การนำโมเดลมูลค่าเพิ่มมาใช้มีสองแนวทาง โดยแนวทางแรกเป็นการประเมิน สถานศึกษาเพื่อความรับผิดชอบที่สามารถตรวจสอบได้ และแนวทางที่สองเป็นการประเมินครูที่เป็นประสิทธิผลเชิงสัมพัทธ์กับครูคนอื่น ซึ่งในบางโมเดลจะพิจารณาเฉพาะความรู้เดิมของนักเรียนหรือรวมตัวแปรอื่นเช่น เพศ ศาสนา และพื้นฐานด้านเศรษฐานะของนักเรียนเข้าร่วมพิจารณาด้วย

ดังนั้นการนำแนวคิดของการวิเคราะห์มูลค่าเพิ่มมาใช้ในประเมินคุณภาพการจัดการศึกษาจึงเป็นแนวทางหนึ่งที่ช่วยให้ข้อสารสนเทศที่สะท้อนถึงการปฏิบัติงานของสถานศึกษา และยังสามารถเปรียบเทียบผลการปฏิบัติของสถานศึกษาได้อย่างยุติธรรม เนื่องจากมีการควบคุมปัจจัยอื่นที่มีพื้นฐานต่างกันซึ่งอาจจะมีความไม่เท่าเทียมกันของปัจจัยนำเข้าของระบบการศึกษา การนำผลจากการวัดมูลค่าเพิ่มจึงมีความถูกต้องน่าเชื่อถือและเกิดประโยชน์หลายประการต่อผู้ที่เกี่ยวข้องทั้งผู้กำหนดนโยบาย ผู้บริหาร ครู ตลอดจนผู้ปกครองและนักเรียน(ประกฤติยา ทักษิโณ, 2552)

ผู้วิจัยจึงมีความสนใจ ที่จะพัฒนาการประเมินคุณภาพการจัดการศึกษา วิชาคณิตศาสตร์เพื่อให้เกิดความยุติธรรมสำหรับนักเรียนและสถานศึกษาที่มีคุณลักษณะพื้นฐานที่แตกต่างกัน

และเพื่อให้ได้ผลการประเมินที่มีความถูกต้องเชื่อถื อได้ โดยใช้ข้อมูลจากการศึกษานานาชาติตาม โครงการประเมินผล TIMSS 2007 ในวิชาคณิตศาสตร์ เนื่องจากโครงการประเมินผล TIMSS เป็น การรวบรวมข้อมูลจากนักเรียนจำนวนมาก ซึ่งคุณลักษณะพื้นฐานที่แตกต่างกันในนักเรียนและ สถานศึกษา อาจจะเป็นปัจจัยที่ทำให้เกิดความแตกต่างในผลกา รประเมินผลของนักเรียน แต่จาก รายงานผลประเมินในโครงการดังกล่าวเป็นเพียงการใช้ผลจากคะแนนที่ได้จากการทำแบบสอบ และไม่คำนึงถึงความแตกต่างของคุณลักษณะพื้นฐานในระดับนักเรียน ระดับสถานศึกษา ตลอด จนถึงในระดับประเทศ ในการวิจัยครั้งนี้จะคำนึงถึงลักษณะข้อมูลที่เป็นโครงสร้างแบบลดหลั่นด้วย การวิเคราะห์พหุระดับ และจากแนวคิดของ Dorans, Schmitt และ Bleistein (อ้างถึงใน Penfield, 2008) ได้กล่าวถึงว่าการวิเคราะห์การทำหน้าที่ของข้อสอบที่สมบูรณ์นั้นต้องประกอบไปด้วยการ วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (Differential item functioning; DIF) และการทำหน้าที่ ต่างกันของตัวลวง (Differential Distractor Functioning; DDF) ทำให้ผู้วิจัย คำนึงการนำผล คะแนนสอบ จากแบบสอบที่อาจจะมีความคลาดเคลื่อนเกิดขึ้นเนื่องจาก แบบสอบมีข้อสอบที่ทำ หน้าที่ต่างกัน ทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง จึงประยุกต์ใช้การวิเคราะห์การทำ หน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) เพื่อให้ได้คะแนนจาก แบบสอบสำหรับประเมิน คุณภาพการจัดการศึกษาในวิชาคณิตศาสตร์ที่ มีความถูกต้องมากยิ่งขึ้น นอกจากนี้ผู้วิจัยยังสนใจที่จะพัฒนาโมเดลการประเมินคุณภาพก ารจัดการศึกษาให้เหมาะสม สำหรับการนำไปใช้ได้จริงอีกด้วย เนื่องจากการศึกษาเกี่ยวกับการประเมินคุณภาพการจัด การศึกษาที่ผ่านมานั้นยังไม่เป็นที่แพร่หลายมากนัก เหตุผลหนึ่งเนื่องจากความซับซ้อนของโมเดล ที่ใช้ในการวิเคราะห์ ผู้วิจัยจึงมีความต้องการที่จะพัฒนาโมเดลการประ เมินให้มีความซับซ้อน น้อยลงเพื่อให้สามารถง่ายต่อการนำไปใช้ ดังนั้นโมเดลการประเมินคุณภาพการจัดการศึกษาของ สถานศึกษาที่ทำการศึกษาในครั้งนี้ ผู้วิจัยจึงสนใจใน 2 ประเด็น คือ

ประเด็นแรกเป็นคุณภาพของแบบสอบ ที่วัดผลสัมฤทธิ์ในวิชาคณิตศาสตร์ ที่มีความ แตกต่างกัน 2 ลักษณะ คือมีการตัดข้อสอบทำหน้าที่ต่างกัน โดยพิจารณาเฉพาะตัวถูก และมีการ ตัดข้อสอบที่ทำหน้าที่ต่างกัน โดยพิจารณาทั้งตัวถูกและตัวลวง ในแบบ วัดผลสัมฤทธิ์ในวิชา คณิตศาสตร์ ทั้งนี้ในการตัดข้อสอบที่ทำหน้าที่ต่างกันจะตัดเฉพาะข้อที่จำเป็นโดยพิจารณาตาม เกณฑ์การตัดข้อสอบ และรัก ษาโครงสร้างเดิมของแบบสอบไว้ เพื่อให้แบบสอบยังคงมีความตรง

ในการประเมินและมีการเปลี่ยนแปลงความเที่ยงของแบบสอบน้อยที่สุด ประเด็นที่สองเป็นโมเดลที่ใช้ในการวิเคราะห์มูลค่าเพิ่ม (Value-added analysis) ที่มีความแตกต่างกัน 2 ลักษณะคือ โมเดลการวิจัยเดิมของประกฤติยา ทัฬกษิโน (2552) ซึ่งทำการวิจัยและศึกษาเกี่ยวกับ การประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน ที่มีการวิเคราะห์โดยนำตัวแปรต้นที่เป็นคุณลักษณะพื้นฐานของนักเรียนและสถานศึกษามาใช้ในการวิเคราะห์ และโมเดลที่มีการวิเคราะห์ความสัมพันธ์โครงสร้าง เชิงสาเหตุ เพื่อคัดเลือกตัวแปรที่มีอิทธิพลต่อผลสัมฤทธิ์การสอบของนักเรียน ซึ่งพบว่ามี 2 โมเดลที่มีสัมประสิทธิ์การทำนายอยู่ในระดับสูง คือโมเดลที่ไม่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกและมีการปรับแก้คะแนนด้วยตัวแปรนักเรียนและสถานศึกษา และโมเดล ที่ตัดข้อสอบที่ทำ หน้าที่ต่างกันออกและมีการปรับแก้คะแนนด้วยตัวแปรนักเรียนและสถานศึกษา

งานวิจัยในครั้งนี้ผู้วิจัยต้องการขยายแนวคิด และเทคนิควิธีการในการวิเคราะห์ข้อมูลในเรื่องของ การประเมินคุณภาพการจัดการศึกษาของสถานศึกษา โดยประยุกต์ใช้การวิเคราะห์มูลค่าเพิ่มที่มีการวิเคราะห์ 2 ระดับ คือ ระดับนักเรียน และระดับสถานศึกษา พร้อมทั้งมีเทคนิคการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวงเพื่อเพิ่มความน่าเชื่อถือ และความยุติธรรมของผลการวัด โดยในแต่ละส่วนประกอบไปด้วยโมเดลการวัด 4 โมเดล คือ โดยมีรายละเอียดดังนี้

โมเดล 1 “Undetected DIF-DDF & Adjusted” เป็นโมเดลการประเมินคุณภาพการจัดการศึกษา วิชาคณิตศาสตร์ที่ ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออก และวิเคราะห์มูลค่าเพิ่มโดยมีการควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา (Adjusted)

โมเดล 2 “Detect DIF & Adjusted” เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ตัวถูกออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา (Adjusted)

โมเดล 3 “Detect DIF-DDF & Unadjusted” เป็นโมเดลเป็นการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ทั้งตัวถูกและตัวลวงออก แต่ไม่ได้ทำการวิเคราะห์มูลค่าเพิ่ม (Unadjusted)

โมเดล 4 “Detect DIF-DDF & Adjusted” เป็นโมเดลเป็นการประเมินคุณภาพการจัดการศึกษา วิชาคณิตศาสตร์ที่ตัดข้อสอบ ที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ทั้งตัวถูกและตัวลวงออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการ ควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา (Adjusted)

โดยหลังจากที่ทำการวิเคราะห์ข้อมูล ผลการสอบของโครงการการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์ หรือ TIMSS 2007 เรียบร้อยแล้วนั้น ก็ดำเนินการเปรียบเทียบโมเดลการวิเคราะห์ที่สามารถใช้ประเมินคุณภาพการศึกษาได้อย่างถูกต้องและมีความเป็นธรรมสำหรับนักเรียน สถานศึกษา เพื่อเป็นแนวทางการประเมินคุณภาพการจัดการศึกษา วิชาคณิตศาสตร์ได้อย่างมีความถูกต้อง และสมบูรณ์มากที่สุด

วัตถุประสงค์ในการวิจัย

1. เพื่อศึกษา ผลการตรวจสอบ การทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) ในแบบสอบประเมินความรู้วิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 2
2. เพื่อ ศึกษาผลของข้อสอบที่ทำหน้าที่ต่างกัน และผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา ต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน 4 โมเดล
3. เพื่อเปรียบเทียบ ประสิทธิภาพ ของโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษา 4 โมเดล

สมมติฐานการวิจัย

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องเกี่ยวกับ โมเดลการประเมินคุณภาพการจัดการศึกษา ผู้วิจัยพบว่า ควรเลือกใช้โมเดลที่มีความถูกต้อง เชื่อถือได้ และมีความ เป็นธรรมมากที่สุด โดยวิธีการวิเคราะห์มูลค่าเพิ่มเป็นวิธีการที่ มีข้อดีคือ เป็นโมเดลการวิเคราะห์พหุระดับ สามารถควบคุมตัวแปรหรือปัจจัยในระดับนักเรียนและระดับสถานศึกษาซึ่งมีผลต่อการเรียนรู้ของนักเรียนได้ (ศิริชัย กาญจนวาสี, 2554) นอกจากนี้ Chaimongkol (2005) พบว่าแหล่งของการเกิดการทำหน้าที่ต่างกันของข้อสอบ (DIF) ของตัวแปรระดับสถานศึกษา ได้แก่รายวิชาที่สอนนอก

ห้องเรียน จำนวนผู้สนับสนุน ทีมงานผู้บริการ และตัวแปรบริบทของสถานศึกษา ดังนั้นข้อสอบที่ทำหน้าที่ต่างกันจะมีแหล่ง เนื่องจากตัวแปรในแต่ละระดับที่นำ มาวิเคราะห์ และถ้าแบบสอบประกอบด้วยข้อสอบที่ทำหน้าที่ต่างกันสำหรับผู้สอบต่างกลุ่มกัน ย่อมทำให้เกิดความไม่เป็นธรรมแก่ผู้สอบบางกลุ่ม พร้อมทั้งDorans, Schmitt และ Bleistein (อ้างถึงใน Penfield, 2008) ได้กล่าวถึงการวิเคราะห์การทำหน้าที่ของข้อสอบที่สมบูรณ์นั้นต้องประกอบไปด้วยการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง ดังนั้น โมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่มีประสิทธิภาพ และมีความยุติธรรมที่สุด น่าจะเป็นโมเดลที่มีการวิเคราะห์มูลค่าเพิ่มเพื่อควบคุมตัวอิทธิพลของตัวแปรในระดับนักเรียนและระดับสถานศึกษา พร้อมทั้งพิจารณาถึงการเกิดการทำหน้าหน้าที่ต่างกันของข้อสอบและการทำหน้าที่ต่างกันของตัวลวง และจากการศึกษาเอกสาร และงานวิจัยที่เกี่ยวข้องเกี่ยวกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง พบว่ามีการศึกษาการเกิด DIF และ DDF ของตัวแปรคุณลักษณะที่หลากหลายแต่ตัวแปรที่พบในเอกสาร และงานวิจัยมากที่สุดคือ ตัวแปรเพศ ด้วยเหตุผลต่างๆที่ ผู้วิจัยได้ศึกษามาจึงตั้งสมมติฐานการวิจัย 3 ข้อ ดังนี้

1. คุณลักษณะเพศน่าจะเป็นคุณลักษณะที่พบการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวงในแบบสอบวิชาคณิตศาสตร์มากที่สุด

2. การประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ด้วยโมเดลการประเมินคุณภาพการจัดการศึกษาที่มีเงื่อนไขการตัดข้อสอบและไม่ตัดข้อสอบที่เกิดการทำหน้าที่ต่างกันของข้อสอบและการทำหน้ าที่ต่างกันของตัวลวง ทำให้ ระดับคุณภาพ และอันดับคุณภาพการจัดการศึกษาแตกต่างกัน และการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ด้ วยโมเดลการประเมินคุณภาพการจัดการศึกษาที่มีเงื่อนไขการควบคุมและไม่ควบคุมตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาทำให้ระดับคุณภาพและอันดับคุณภาพการจัดการศึกษาแตกต่างกัน

3. โมเดลการประเมิน คุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่แตกต่างกันทำให้ผลการประเมินคุณภาพการจัด การศึกษาแตกต่างกัน โดยโมเดลที่มีประสิทธิภาพที่สุด น่าจะเป็นโมเดลที่มีการตัดข้อสอบที่เกิดการทำหน้าที่ต่างกันของข้อสอบและการทำหน้าที่ต่างกันของตัวลวง พร้อมทั้งมีการควบคุมอิทธิพลของตัวแปรคุณลักษณะระดับนักเรียนและระดับสถานศึกษา

ขอบเขตของการวิจัย

1. ข้อมูลที่ใช้ในการวิจัยครั้งนี้ เป็นข้อมูลทุติยภูมิจากฐานข้อมูล (data base) ของโครงการการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์ หรือ TIMSS 2007 จัดโดย IEA (The International Association for the Evaluation of Educational Achievement) เป็นสมาคมนานาชาติที่ทำหน้าที่ประเมิน ผลด้านการศึกษา กลุ่มตัวอย่างของโครงการ ประกอบด้วย นักเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน 5,412 คน ครูคณิตศาสตร์ที่สอนนักเรียนที่เป็นกลุ่มตัวอย่างในการวิจัย จำนวน 150 คน และผู้บริหารสถานศึกษาที่นักเรียนกลุ่มตัวอย่างศึกษา จำนวน 150 สถานศึกษา

2. แบบสอบที่ใช้ในการศึกษา เป็นแบบสอบวัดผลสัมฤทธิ์ ในโครงการการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์ หรือ TIMSS 200 7 ซึ่งในการวิจัยครั้งนี้ศึกษาเฉพาะวิชาคณิตศาสตร์ ซึ่งประกอบด้วยข้อสอบแบบเลือกตอบ และแบบเขียนคำตอบ มีข้อสอบทั้งหมด 14 ฉบับ โดยมีจำนวนข้อสอบฉบับละ 28 – 33 ข้อ

3. โมเดล การประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ในงานวิจัยเรื่องนี้ ประกอบด้วย 4 โมเดล คือ

โมเดล 1 “Undetected DIF-DDF & Adjusted” เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา (Adjusted)

โมเดล 2 “Detect DIF & Adjusted” เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ตัวถูกออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา (Adjusted)

โมเดล 3 “Detect DIF-DDF & Unadjusted” เป็นโมเดลเป็นการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ทั้งตัวถูกและตัวลวงออก แต่ไม่ได้ทำการวิเคราะห์มูลค่าเพิ่ม (Unadjusted)

โมเดล 4 “Detect DIF-DDF & Adjusted” เป็นโมเดลเป็นการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ทั้งตัวถูกและตัวลวงออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการ ควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา (Adjusted)

คำจำกัดความที่ใช้ในการวิจัย

โมเดลการประเมินคุณภาพการจัดการศึกษา หมายถึง โมเดลในการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ ซึ่งในการวิจัยครั้งนี้ประกอบด้วยการวิเคราะห์ 2 ระดับ คือระดับนักเรียน และระดับสถานศึกษา โดยประกอบไปด้วยโมเดลการวัด 4 โมเดล โดยมีรายละเอียดดังนี้

โมเดล 1 “Undetected DIF-DDF & Adjusted” เป็นโมเดลการประเมินคุณภาพการจัดการศึกษา วิชาคณิตศาสตร์ที่ ไม่ได้ ตัดข้อสอบที่ทำหน้าที่ต่างกันออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา (Adjusted)

โมเดล 2 “Detect DIF & Adjusted” เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ตัวถูกออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา (Adjusted)

โมเดล 3 “Detect DIF-DDF & Unadjusted” เป็นโมเดลเป็นการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ทั้งตัวถูกและตัวลวงออก แต่ไม่ได้ทำการวิเคราะห์มูลค่าเพิ่ม (Unadjusted)

โมเดล 4 “Detect DIF-DDF & Adjusted” เป็นโมเดลเป็นการประเมินคุณภาพการจัดการศึกษา วิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่า งกันโดยการวิเคราะห์ทั้งตัวถูกและตัวลวงออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการ ควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา (Adjusted)

ทั้งนี้โมเดลที่มีการตัดข้อสอบที่ทำหน้าที่ต่างกัน ทั้งตัวถูกและตัวลวงออกนั้นจะเป็น เฉพาะข้อที่จำเป็น โดยพิจารณาเลือกตัดข้อสอบตามเกณ ฑ์ 2 ข้อ ดังนี้ 1. เลือกข้อสอบที่ทำหน้าที่ต่างกันที่มีค่าอิทธิพลสูงสุด และ 2. คำนึงถึงโครงสร้างของแบบสอบ โดยภายหลังการตัด

ข้อสอบออกแล้วโครงสร้างของแบบสอบจะไม่แตกต่างจากโครงสร้างเดิม และข้อสอบที่ตัดออกไปควรมีจำนวนไม่เกินร้อยละ 20 (Clauser (1993), cited in Naraya & Swaminathan,1994) ของข้อสอบในแบบสอบนั้น โดยการตัดข้อสอบที่ทำหน้าที่ต่างกันจะคำนึงถึงการรักษาโครงสร้างเดิมโดยจำแนกตามสัดส่วนเนื้อหาของแบบสอบ เพื่อให้แบบสอบมียังคงมีความตรงในการประเมินและมีการเปลี่ยนแปลงความเที่ยง ของแบบสอบน้อยที่สุด

มูลค่าเพิ่มของการ จัดการศึกษาวิชาคณิตศาสตร์ ในสถานศึกษา หมายถึง ค่าส่วนต่างระหว่างค่าพารามิเตอร์ความสามารถโดยเฉลี่ยของสถานศึกษาที่วัดได้จริงกับค่าพารามิเตอร์ความสามารถโดยเฉลี่ยของสถานศึกษาที่ได้จากการประมาณค่าหรือการพยากรณ์จากปัจจัยหรือตัวแปรของสถานศึกษา

ผลการวิเคราะห์ แสดงได้ จากค่าส่วนที่เหลือ (residual term) จากโมเดลการวิเคราะห์ระดับสถานศึกษา ด้วยหลักการวิเคราะห์พหุระดับ ของโปรแกรมการวิเคราะห์เชิงเส้นระดับลดหลั่น (HLM)

โมเดลเชิงเส้นตรงแบบลดหลั่น (Hierarchical Linear Model: HLM) หมายถึง รูปแบบการวิเคราะห์ข้อมูลที่มีตัวแปรอิสระหรือตัวแปรทำนายมากกว่าหนึ่งตัว ที่จัดลักษณะของตัวแปรได้เป็นระดับชั้นอย่างน้อย 2 ระดับ และมีตัวแปรตามหรือตัวแปรเกณฑ์อยู่ในระดับที่ 1 โดยผลการวิเคราะห์ในระดับต่ำ กว่าสามารถนำมาใช้ในการวิเคราะห์ ในระดับที่สูงขึ้น ทำให้สามารถอธิบายอิทธิพลที่ส่งผลของตัวแปรที่อยู่ในระดับต่างกันได้

คุณภาพการจัดการศึกษา วิชา คณิตศาสตร์ของสถานศึกษาขั้นพื้นฐาน หมายถึง คะแนนมูลค่าเพิ่มของสถานศึกษาหรือค่าส่วนที่เหลือจากการวิเคราะห์ มูลค่าเพิ่มในแต่ละโมเดลการประเมินคุณภาพการจัดการศึกษา ซึ่งคุณภาพการจัดการศึกษาของสถานศึกษาจากการวิจัยในครั้งนี้แสดงถึงการดำเนินงานจัดการของแต่ละสถานศึกษา เพื่อให้นักเรียนเกิดการเรียนรู้ที่มีการควบคุมปัจจัย ตัวแปรระดับนักเรียน และระดับสถานศึกษา ที่อยู่นอกเหนือการควบคุมของสถานศึกษาแล้ว

การเปรียบเทียบโมเดล การประเมินคุณภาพการจัดการศึกษา หมายถึง การเปรียบเทียบผลการจั ดระดับคุณภาพและการจัดอันดับคุณภาพของสถานศึกษาจากคะแนนมูลค่าเพิ่ม โดยการจัดระดับคุณภาพ จัดแบ่งออกเป็นกลุ่ม 3 กลุ่ม คือ กลุ่มที่ 1 เป็นกลุ่มต่ำ มี

ตำแหน่งเปอร์เซ็นต์ไทล์ 1 – 24.99 กลุ่มที่ 2 เป็นกลุ่มกลาง มีตำแหน่งเปอร์เซ็นต์ไทล์ 25 – 74.99 และ กลุ่มที่ 3 เป็นกลุ่มสูง มีตำแหน่งเปอร์เซ็นต์ไทล์ 75 ขึ้นไป และการจัดอันดับคุณภาพเป็นการเรียงอันดับคะแนนมูลค่าเพิ่มของสถานศึกษาจากมากที่สุดไปน้อยที่สุด

ประสิทธิภาพ ของโมเดลการประเมินคุณภาพการจัดการศึกษา หมายถึง ค่าสัมประสิทธิ์การทำนาย (R^2) จากโมเดลการประเมินคุณภาพการจัดการศึ กษา ซึ่งเป็นโมเดลการวิเคราะห์พหุระดับที่มีการควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา ซึ่งมีผลต่อการเรียนรู้ของนักเรียน

ค่าความสามารถของนักเรียน จากการวัดผลสัมฤทธิ์ในวิชาคณิตศาสตร์ หมายถึง ค่าพารามิเตอร์ความสามารถของนักเรียน จากการทดสอบ TIMSS 2007 ซึ่งการประมาณค่าพารามิเตอร์ความสามารถของนักเรียนได้จากผลการวิเคราะห์ด้วยโปรแกรม MULTILOG

การประเมินคุณภาพการจัดการศึกษา วิชาคณิตศาสตร์ หมายถึง การตัดสินคุณค่าของสถานศึกษาขั้นพื้นฐานโดยการวิเคราะห์หาคุณภาพการจัดการศึกษาของสถานศึกษา จากการใช้ผลจากการทำแบบสอบวัดผลสัมฤทธิ์ในวิชาคณิตศาสตร์

ตัวแปรต้นระดับนักเรียน หมายถึง คุณลักษณะของนักเรียนที่นอกเหนือจากความสามารถของผู้สอบ ได้แก่ ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา เพศ ระดับการศึกษาของผู้ปกครอง แหล่งทรัพยากรที่บ้าน ความมั่งคั่งของครอบครัว กิจกรรมการฝึกปฏิบัติในวิชาคณิตศาส ตร์ ดัชนีระดับเวลาที่ใช้ในการทำการบ้านวิชาคณิตศาสตร์ การใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเอง ดัชนีของความเชื่อมั่นในตนเองเกี่ยวการเรียนวิชาคณิตศาสตร์ ดัชนีความเพลิดเพลินในการเรียนวิชาคณิตศาสตร์ แรงจูงใจภายนอกในการเรียนคณิตศาสตร์ การเห็นคุณค่าในวิชาคณิตศาสต ร์ ดัชนีของนักเรียนในความรู้สึกปลอดภัยในการมาเรียนหนังสือภายในสถานศึกษา

ตัวแปรต้นระดับสถานศึกษา หมายถึง คุณลักษณะของสถานศึกษาขั้นพื้นฐาน ได้แก่ ขนาดของสถานศึกษา ที่ตั้งสถานศึกษา ระดับการขาดแคลนครูคณิตศาสตร์ในสถานศึกษา การให้การสนับสนุนของผู้ปกครอง ดัชนีการให้การสนับสนุนด้านทรัพยากรที่ใช้ในการจัดการเรียนการ

สอนคณิตศาสตร์ การจัดกิจกรรมที่สนับสนุนการเรียนรู้ในวิชาคณิตศาสตร์ภายใน สถานศึกษา ดัชนีการรับรู้ถึงบรรยากาศในที่ดีในที่ทำงาน การพัฒนาครูในสถานศึกษา ประสบการณ์การทำงานของครูในสถานศึกษา ดัชนีความพึงพอใจของครูที่มีต่อการจัดการภายในสถานศึกษา ดัชนีการให้ความสำคัญกับการให้การบ้านคณิตศาสตร์ของครูใน สถานศึกษา ขนาดของห้องเรียน ระดับความถี่ในการทำงานร่วมกันของครูคณิตศาสตร์ในสถานศึกษา

การทำหน้าที่ต่างกันของข้อสอบ (Differential Item functioning: DIF) หมายถึง ลักษณะของข้อสอบที่ทำให้ผู้สอบต่างกลุ่มกันที่มีระดับความสามารถของผู้สอบเท่ากัน แต่มีโอกาสในการตอบข้อสอบได้ถูกต้องแตกต่างกัน ในการวิจัยนี้ใช้การวิเคราะห์ข้อสอบที่ทำหน้าที่ต่างกัน โดยใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น และตรวจสอบจากคุณลักษณะของนักเรียน 5 ตัวแปร คือตัวแปรเพศ การใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ ระดับการศึกษาของผู้ปกครองแหล่งทรัพยากรที่บ้าน และ ความมั่งคั่งของครอบครัว

การทำหน้าที่ต่างกันของตัวลวง (Differential Distractor functioning: DDF) หมายถึง การที่สมาชิกในกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่มีความสามาร ถเท่ากัน มีโอกาสในการเลือกตัวลวงแทนที่จะเลือกตัวถูกได้แตกต่างกัน และตรวจสอบจากคุณลักษณะของนักเรียน 5 ตัวแปร คือ ตัวแปรเพศ การใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ ระดับการศึกษาของผู้ปกครอง แหล่งทรัพยากรที่บ้าน และ ความมั่งคั่งของครอบครัว

แบบวัดผลสัมฤทธิ์ในวิช าคณิตศาสตร์ หมายถึง แบบสอบวิชา คณิต ศาสตร์ ซึ่งวัดความรู้ด้าน คณิต ศาสตร์ตาม หลักสูตรของประเทศไทย ในโครงการ การศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์ หรือ TIMSS 2007

ประโยชน์ที่คาดว่าจะได้รับ

1. **ประโยชน์ด้านวิชาการ**

1.1 เป็นการขยายแนวคิดในเรื่องของการ ตรวจสอบ การทำหน้าที่ต่า งกันของข้อสอบร่วมกับการทำหน้าที่ต่างกันของตัวลวง ของข้อสอบไปสู่การปฏิบัติจริง ซึ่งจะทำให้ได้สารสนเทศที่มากยิ่งขึ้นเกี่ยวกับคุณภาพของข้อสอบที่ใช้ อีกทั้งเป็นการช่วยเพิ่มมิติในการเข้าใจเรื่องของการทำ

หน้าที่ต่างกันของข้อสอบได้อย่างสมบูรณ์ และ เป็นการเพิ่มความน่ าเชื่อถือของการนำผลการทดสอบไปใช้ในการตัดสิน

1.2 ผลการ ศึกษาสามารถแสดงให้เห็นถึงความเชื่อมโยงของแนวคิดเกี่ยวกับการวิเคราะห์มูลค่าเพิ่ม (value added) โดยการประยุกต์ใช้การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง ซึ่งทำให้เกิดการขยายองค์ ความรู้เกี่ยวกับการวิเคราะห์โดยใช้โมเดลพหุระดับ

2 ประโยชน์ด้านการนำไปใช้

2.1 จากผลการวิจัยทำให้ทราบถึงอิทธิพลของตัวแปรคุณลักษณะระดั บนักเรียน และระดับสถานศึกษา ส่งผลต่อคุณภาพการจัดการศึกษาในวิชาคณิตศาสตร์ของนักเรียน ซึ่งจะเป็นประโยชน์ต่อการวางแผน การกำหนดแนวทางในการพัฒนาสถานศึกษา

2.2 วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง ดำเนินการโดยใช้โปรแกรมที่ถูกสร้างขึ้นมาใหม่ คือโปรแกรม DDFS 1.0 ซึ่งง่ายต่อการใช้งาน ดังนั้นหน่วยงานที่เกี่ยวข้องกับการทดสอบ หรือสถานศึกษาที่มีความสนใจในเรื่องขอ งความยุติธรรมของแบบสอบ สามารถที่จะนำไปใช้ในหน่วยงาน หรือสถานศึกษาได้ไม่ยาก และจะเป็นเป็นการเพิ่มความน่าเชื่อถือของคุณภาพเครื่องมืออีกด้วย

บทที่ 2

แนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยแบ่งการนำเสนอแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง ออกเป็น 5 ตอน คือ ตอนที่ 1 ความลำเอียงของแบบสอบ ตอนที่ 2 แนวคิดเกี่ยวกับมูลค่าเพิ่มของการจัดการศึกษา ตอนที่ 3 การวิเคราะห์พหุระดับ ตอนที่ 4 การศึกษาแนวโน้ มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ พ .ศ.2550 (TIMSS 2007) ตอนที่ 5 เอกสารงานวิจัยเกี่ยวกับการประเมินคุณภาพการจัดการศึกษา โดยแต่ละตอนมีรายละเอียด ดังนี้

ตอนที่ 1 ความลำเอียงของแบบสอบ

1.1 ความเป็นมาของความลำเอียงของแบบสอบ

งานวิจัยเกี่ยวกับความยุติธรรมข องข้อสอบปรากฏครั้งแรกในปี ค .ศ. 1910 โดย Binet (Benet & Simon อ้างถึงใน Camilli & Shepard, 1994) และมาได้รับความสนใจมากขึ้นในช่วงปี ค.ศ. 1960 ซึ่งเป็นผลมาจาก ความต้องการ ที่จะเข้าไปศึกษาถึงความไม่เท่าเทียมกันของการทดสอบความสามารถนักเรียน โดยใช้แบบสอบ โดยความแต กต่างที่เกิดขึ้นนั้นอาจจะเป็นผลเนื่องมาจากความแตกต่างในด้านต่างๆ โดยเฉพาะทางด้านวัฒนธรรม หรือสีผิว ข้อสันนิษฐานที่เกี่ยวกับการทำหน้าที่ของข้อสอบบางข้ออาจจะเกิดความลำเอียงเข้าข้างนักเรียนบางกลุ่ม โอกาสที่ผู้สอบที่มีความสามารถเท่ากันจะตอบข้อสอบข้อนั้นได้ถูกต้ องมีไม่เท่ากัน อันมีผลเนื่องมาจากข้อสอบข้อนั้นต้องใช้ประสบการณ์ที่ชนกลุ่มน้อยขาดโอกาสในการเรียนรู้ การศึกษาในครั้งนั้นจึงทำเพื่อค้นหา แล้วคัดข้อสอบข้อที่มีความลำเอียงออกไปเพื่อให้แบบสอบมีความยุติธรรมแก่ผู้เข้าสอบทุกคน (Angoff, 1993) นอกจากนั้นมีการเสนอวิธี การต่างๆ เพื่อตรวจสอบความลำเอียงของข้อสอบ (Item bias) ความลำเอียงของแบบสอบ (Test bias) และความลำเอียงในการคัดเลือก (Selection bias) โดยนิยามความลำเอียงว่าเป็นความคลาดเคลื่อนอย่างเป็นระบบ (Systematic error) (ศิริชัย กาญจนวาสี , 2550) ที่เกิดขึ้นจากการวัด ซึ่ งความพยายาม ของนักวัดผลในการศึกษาเกี่ยวกับความลำเอียงของแบบสอบนั้น เพื่อต้องการพัฒนาแบบสอบให้มีคุณภาพ และมีความเหมาะสมที่จะนำไปใช้ให้มากที่สุด

การศึกษาในเรื่องของ ความลำเอียงของข้อสอบ (Item bias) เกิดขึ้นอย่างกว้างขวาง จนกระทั่งเกิดประเด็นความเข้าใจที่ไม่ต รงกันกับการใช้คำและความหมาย มีการให้ความเห็นว่า ความลำเอียงของข้อสอบ เป็นผลการตัดสินว่าข้อสอบมีความยุติธรรมหรือไม่ อันส่งผลต่อการบรรลุจุดมุ่งหมายของการใช้แบบสอบหรือความลำเอียงของข้อสอบ เป็นสารสนเทศทางสถิติที่ได้จากข้อสอบเกี่ยวกับความสัมพันธ์ระหว่างคุณลักษณะ ที่ข้อสอบมุ่งวัด กับประสบการณ์ของผู้สอบกลุ่มต่างๆที่ทำการสอบเมื่อกลุ่มผู้สอบต่างกลุ่มกันตอบข้อสอบข้อเดียวกัน ความแตกต่างที่เกิดขึ้นอาจมาจากความไม่เหมาะสมของข้อคำถาม ซึ่งสามารถเกิดขึ้นได้หลายลักษณะ หรือประสบการณ์ของผู้สอบซึ่งอาจมีลักษณะพื้นฐานเดิมแตกต่างกัน ในหลายสถานการณ์จึงไม่เหมาะสมที่จะใช้คำว่า ข้อสอบลำเอียง (Biased item) เนื่องจากเป็นภาษาที่มีความหมายในเชิงลบ ประกอบกับเกณฑ์ที่ใช้สำหรับตัดสินความลำเอียงยังมีความคลุมเครือ และค่อนข้างสับสน ดังนั้นจึงควรเปลี่ยนมาใช้คำว่า การทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning; DIF) ซึ่งเป็นคำที่มีความเป็นกลางและเหมาะสมมากกว่า (Halland and Wainer, 1993)

การทำหน้าที่ต่างกันของข้อสอบ (DIF) กับความลำเอียงของข้อสอบ (Item bias) มีแนวคิดที่แตกต่างกัน โดยสำหรับการทำหน้าที่ต่างกันของข้อสอบ เป็นกระบวนการที่เน้นการใ ช้วิธีการทางสถิติสำหรับการตรวจสอบ เพื่อให้ได้สารสนเทศสำหรับการทำหน้าที่ของข้อสอบสำหรับกลุ่มผู้สอบกลุ่มย่อยที่มีลักษณะเฉพาะบางอย่างแตกต่างกัน ส่วนความลำเอียงของข้อสอบเป็นกระบวนการตัดสินความยุติธรรมของข้อสอบ โดยนำสารสนเทศการทำหน้าที่ต่างกันของข้อสอบมาวิเคราะห์เชิงตรรกะ (Logical analysis) โดยผู้เชี่ยวชาญพิจารณาถึงการเขียนข้อสอบ เนื้อหาสาระของข้อสอบและจุดมุ่งหมายของการวัด เพื่อระบุว่าข้อสอบข้อนั้นลำเอียงเข้าข้างกลุ่มใดหรือไม่ เพราะเหตุใดจึงเป็นการตัดสินความลำเอียงของข้อสอบ (Camilli and Shapard อ้างถึงในศิริชัย กาญจนวาสี, 2550)

การประเมินเกี่ยวกับ การทำหน้าที่ต่างกันของข้อสอบไม่หยุดแค่ เพียงเป้าหมายที่ว่าเป็นการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยวิธีทางสถิติ โดยเป้าหมายต่อไปของการศึกษาคือการพยายามที่จะเข้าใจถึงเหตุผลของการเกิดการทำหน้าที่ต่างกันของข้อสอบ (Green, Crone และ Folk, 1989) การพัฒนาวิธีลอกลิเนียร์สำหรับใช้ในการประเมินการทำหน้าที่ต่างกันของตัวลวง ให้เป็นวิธีมาตรฐานเพื่อใช้ในการวิเคราะห์ตัวลวงซึ่งจะทำให้ได้ประโยชน์จากการวิเคราะห์

การศึกษาเกี่ยวกับการทำหน้าที่ของตัวลวงนั้นยังไม่ได้มีผู้ทำการศึกษามากนัก โด ยการศึกษาส่วนมากของกลุ่มที่มีความแตกต่างในการตอบข้อสอบจะมุ่งเน้นไปที่การตอบข้อถูกเป็นหลัก คะแนนแบบสอบให้ความหมายที่เหมือนกันหรือไม่สำหรับกลุ่มผู้สอบที่มีความแตกต่างกัน โดยถ้าความสัมพันธ์ของคะแนนแบบสอบมาจากสมรรถนะที่เกิดขึ้นในโรงเรียนหรือในการทำงานคือมีความเหมือนสำหรับแต่ละกลุ่มย่อย การศึกษาเกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบมักถูกเรียกได้อีกแบบหนึ่งว่าความลำเอียงของข้อสอบ

เนื่องจากตัวเลือกที่เป็นตัวลวงไม่มีผลต่อคะแนนในแบบสอบ ซึ่งทำให้ดูเหมือนว่าการวิเคราะห์การทำหน้าที่ของตัวลวงไม่ มีความจำเป็นที่จะต้องควบ คุมสำหรับ คะแนนรวมของแบบสอบ โดยการทำหน้าที่ต่างกันของตัวลวงไม่ได้เป็นการนำเอาความสามารถของผู้สอบไปทำการพิจารณา แต่เกี่ยวข้องกับการวิเคราะห์การตอบข้อสอบแบบ 2 ทางของแต่ละกลุ่มย่อยในคำตอบที่เลือกของข้อสอบแต่ละข้อ อย่างไรก็ตามบ่อยครั้งที่การวิเคราะห์ การทำหน้าที่ ต่างกันของตัวลวงเป็นการหาตัวลวงที่ถูกเลือกโดยผู้สอบซึ่งมีระดับความสามารถแตกต่างกัน ตัวลวงบางข้อดึงดูดให้ผู้สอบที่ไม่มีความรู้เกี่ยวกับเนื้อหาที่ต้องการวัด และใช้เหตุผลอื่นในการเลือกแต่เป็นคำตอบที่ไม่ถูกต้อง สิ่งหนึ่งที่เห็นได้ชัดเจนจากตัวอย่างปฏิสัมพันธ์ข องความสามารถและการเลือกตัวเลือกความสามารถต้องคงที่ด้วยเหตุผลบางอย่างเมื่อการวิเคราะห์ความแตกต่างของกลุ่มผู้สอบในการเลือกตัวเลือกที่เป็นตัวลวงเพื่อว่าการตรวจสอบความแตกต่างของกลุ่มผู้สอบจะเป็นความแตกต่างของความสามารถที่ถูกซ่อนเร้น (Osterlind อ้างถึงใน Green, Crone และ Folk, 1989)

วิธีการวิเคราะห์การตอบข้อสอบเป็นการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยผ่านตัวเลือกที่เป็นตัวลวงซึ่งผู้สอบตอบผิด การวิเคราะห์ในลักษณะนี้เรียกว่าการทำหน้าที่ต่างกันของตัวลวง (Differential Distractor Functioning) หรือ DDF การศึกษาเ กี่ยวกับตัวลวงกำลังกลายเป็นสิ่งที่มีความสำคัญมาก นั่นเพราะทำให้ได้ข้อมูลสารสนเทศ ที่มากยิ่งขึ้น เกี่ยวกับสมรรถนะของนักเรียนโดยเฉพาะส่วนที่เป็นการเริ่มต้นของนักเรียนในการพิจารณาที่จะตัดตัวเลือกที่เป็นตัวลวงทิ้งซึ่งเป็นอีกทางที่ใช้เพื่อพัฒนาการประเมินการเปลี่ย นแปลง (Barton อ้างถึงใน Middleton และ Laitusis, 2007) ผลการวิเคราะห์ การทำหน้าที่ต่างกันของตัวลวง จะทำให้ผู้พัฒนาแบบสอบได้ตัดสินใจ ซึ่งข้อสอบหลายข้ออาจต้องการการสังเกตต่อไป เนื่องจากข้อสอบเหล่านั้นแสดงการทำหน้าที่ต่างกันของตัวลวงโดยมีตัวลวงทำหน้าที่ต่างกันสำ หรับกลุ่ม

ผู้สอบที่มีความแตกต่าง ด้วยเหตุนี้ตัวลวงอาจจะสามารถแยกเหตุผลที่เป็นไปได้ว่าทำไมข้อสอบจึงมีการทำหน้าที่ต่างกัน

จากการที่การศึกษาถึงการทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning; DIF) มีการศึกษามาอย่างยาวนาน โดยมีวัตถุประสงค์เพื่อให้ ข้อสอบที่ใช้มีความยุติธรรมสำหรับผู้เข้าสอบมากที่สุด โดยในอดีตที่ผ่านมางานวิจัยส่วนใหญ่เป็นการศึกษาผ่านตัวเลือกที่เป็นตัวถูก ส่วนการศึกษาที่เกี่ยวกับการทำหน้าที่ต่างกันของตัวลวง (Differential Distractor Functioning; DDF) นั้นยังมีไม่มากนัก ดังนั้นการศึกษาใ นเรื่องการทำหน้าที่ต่างกันของตัวลวงให้มากยิ่งขึ้นจึงเป็นสิ่งที่มีความจำเป็นเพื่อเป็นการพัฒนาในเรื่องเกี่ยวกับความยุติธรรมของข้อสอบให้มีความสมบูรณ์มากยิ่งขึ้น

1.2 ความหมายของการทำหน้าที่ต่างกันของข้อสอบและการทำหน้าที่ต่างกันของตัวลวง

1.2.1 การทำหน้าที่ต่างกันของข้อสอบ

ความลำเอียงของข้อสอบ (Item Bias) หรือที่ในปัจจุบันเรียกว่า การทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning; DIF) เป็นอีกกระบวนการหนึ่งในการตรวจสอบคุณภาพของข้อสอบ ซึ่งมีนักวัดผลหลายท่านได้ทำการศึกษา และได้ให้ความหมายของความลำเอียง ของข้อสอบและการทำหน้าที่ต่างกันของข้อสอบ ไว้ดังนี้

ความลำเอียงของข้อสอบ หมายถึง สัดส่วนของผู้สอบที่ตอบข้อสอบได้ถูกต้องไม่เท่ากันในแต่ละกลุ่มประชากรที่ใช้ในการศึกษา เมื่อกลุ่มผู้สอบมีคะแนนเท่ากันและข้อสอบมีความเป็นเอกพันธ์ (Scheuneman, 1979)

ความลำเอียงของข้อสอบ หมายถึง เป็นข้อสอบซึ่งผู้สอบที่มีความสามารถเท่ากันแต่อยู่ต่างกลุ่มกันมีโอกาสที่จะตอบถูกไม่เท่ากัน (Lord, 1980)

การทำหน้าที่ต่างกันของข้อสอบ หมายถึง สารสนเทศทางสถิติของข้อสอบที่ได้จากผลการตอบของผู้สอบต่างกลุ่มกัน และมีความสามารถเท่ากัน แต่มีโอกาสในการ ตอบข้อสอบได้ถูกต้องแตกต่างกัน (Halland and Wainer, 1993)

การทำหน้าที่ต่างกันของข้อสอบ หมายถึง ข้อสอบที่เข้าข้างผู้สอบกลุ่มหนึ่งมากกว่าอีกกลุ่มหนึ่งที่นำมาจับคู่เปรียบเทียบกัน ซึ่งทำให้ผู้สอบกลุ่มหนึ่งได้ประโยชน์ แต่อีกกลุ่มหนึ่งเสียประโยชน์ (Shealy and Stout, 1993)

การทำหน้าที่ต่างกันของข้อสอบ หมายถึง ความเป็นพหุมิติในการวัดของข้อสอบ ซึ่งแสดงได้จากการแจกแจงความสามารถหลัก (Primary ability) ของกลุ่มผู้สอบตั้งแต่ 2 กลุ่มขึ้นไปมีความเท่ากันแต่มีการแจกแจงความสามารถรอง (Secondary ability) แตกต่างกัน (Camilli and Shepard, 1994)

การทำหน้าที่ต่างกันของข้อสอบ หมายถึง การที่ข้อสอบให้ผลการตอบข้อสอบของกลุ่มที่นำมาเปรียบเทียบกันมีความแตกต่าง (potenza and Doran, 1995)

การทำหน้าที่ต่างกันของข้อสอบ หมายถึง ฟังก์ชันการตอบสนองข้อสอบซึ่งคำนวณจากกลุ่มผู้สอบกลุ่มย่อยที่ต่างกัน มีค่าไม่เท่ากัน (Narayanan and Sawaminathan, 1996)

การทำหน้าที่ต่างกันของข้อสอบ หมายถึง การที่ข้อสอบทำให้ผู้สอบจากต่างกลุ่มกันที่มีความสามารถหรือคุณลักษณะที่มุ่งวัดเท่ากัน มีโอกาสในการตอบข้อสอบได้ถูกต้องแตกต่างกันหรือมีฟังก์ชันการตอบสนองข้อสอบแตกต่างกัน การ ทำหน้าที่ต่างกันของข้อสอบเกิดขึ้นเมื่อนำข้อสอบไปทดสอบกับผู้สอบกลุ่มย่อยต่างกันที่มีความสามารถหลัก (Primary ability) ระดับเดียวกันหรือมีคุณลักษณะแฝง (Latent trait) ที่ต้องการวัดเท่ากัน แต่มีความสามารถรอง (Secondary ability) แตกต่างกัน ทำให้ผู้สอบต่างกลุ่มที่ นำมาจับคู่เปรียบเทียบมีโอกาสตอบข้อสอบถูกแตกต่างกัน (ศิริชัย กาญจนวาสี, 2550)

ดังนั้น ความหมายของการทำหน้าที่ต่างกันของตัวลวงจากนักวัดผลหลายๆท่าน จึงสามารถสรุปได้ว่า การทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning) หมายถึง การที่ข้อสอบทำให้ผู้ต อบจากต่างกลุ่มกันที่มีความสามารถหรือคุณลักษณะที่มุ่งวัดเท่ากัน มีโอกาสในการตอบข้อสอบได้ถูกแตกต่างกัน

1.2.2 การทำหน้าที่ต่างกันของตัวลวง

การทำหน้าที่ต่างกันของตัวลวง หมายถึง การวิเคราะห์การตอบข้อสอบโดยเป็นการตรวจสอบความแตกต่างของการทำหน้าที่ของข้อสอบผ่านตัวเลือกที่เป็นตัวลวงเมื่อผู้สอบ ต่างกลุ่มกัน แต่มีความสามารถเท่ากันเลือกตัวเลือกที่เป็นตัวลวง (Green, Crone & Folk, 1989)

การทำหน้าที่ต่างกันของตัวลวง เป็นการวิเคราะห์ความแตกต่างในการเลือกตอบตัวเลือกที่เป็นตัวลวงของข้อสอบ โดยกลุ่มผู้สอบที่มีความแตกต่างกัน จะมีโอกาสในการเลือกตัวลวงได้แตกต่างกัน (Kato, Moen & Thurlow, 2009)

การทำหน้าที่ต่างกันของตัวลวง (DDF) หมายถึง การที่สมาชิกในกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่มีความสามารถเท่ากัน มีโอกาสในการเลือกตัวลวงแทนที่จะเลือกตัวถูกได้แตกต่างกัน (Penfield, 2010)

ดังนั้นจากความหมายของการทำหน้าที่ต่างกันของตัวลวงจากนักวัดผลหลายๆท่าน จึงสามารถสรุปได้ว่าการทำหน้าที่ต่างกันของตัวลวง (Differential Distractor Functioning) หมายถึง การที่สมาชิกในกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่มี ความสามารถหรือคุณลักษณะที่มุ่งวัดเท่ากัน มีโอกาสในการเลือกตัวลวงแทนที่จะเลือกตัวถูกได้แตกต่างกัน

1.3 ประเภทของการทำหน้าที่ต่างกันของข้อสอบ

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

ในการวิเคราะห์การทำหน้าที่ ต่างกันของข้อสอบ จะพบว่า ข้อสอบสามารถทำหน้าที่แตกต่างกันได้ 2 ประเภท (Mellenbergh, 1982) ดังนี้

1. ข้อสอบทำหน้าที่ต่างกันแบบเอกรูป (Uniform DIF) หมายถึง ข้อสอบที่ทำให้ผู้สอบกลุ่มหนึ่งมีโอกาสในการตอบข้อสอบถูกมากกว่าผู้สอบอีกกลุ่มหนึ่งอย่างสม่ำเสมอกันในทุกระดั บความสามารถเมื่อพิจารณาโค้งคุณลักษณะของข้อสอบของผู้สอบ 2 กลุ่ม จะพบว่าไม่มีปฏิสัมพันธ์ระหว่างระดับความสามารถของผู้สอบกับการเป็นสมาชิกของกลุ่ม (Group membership)

2. ข้อสอบทำหน้าที่ต่างกันแบบอเนกรูป (Non-uniform DIF) หมายถึง ข้อสอบที่ทำให้โอกาสในการตอบข้อสอบถู กของผู้สอบระหว่างกลุ่มแตกต่างกันอย่างไม่สม่ำเสมอกันในทุกระดับความสามารถ เมื่อพิจารณาโค้งคุณลักษณะข้อสอบของผู้สอบ 2 กลุ่ม พบว่ามีปฏิสัมพันธ์ร่วมกันระหว่างระดับความสามารถของผู้สอบกับการเป็นสมาชิกของกลุ่ม เช่น ที่ระดับ

ความสามารถหนึ่ง กลุ่มผู้สอบกลุ่ม R มีโอกาสในการตอบข้อสอบถูกมากกว่ากลุ่มผู้สอบกลุ่ม F แต่ที่ระดับความสามารถอีกระดับหนึ่งกลุ่มผู้สอบกลุ่ม F มีโอกาสในการตอบข้อสอบถูกมากกว่ากลุ่มผู้สอบกลุ่ม R

ตามทฤษฎีการตอบสนองของข้อสอบ (Item Response Theory; IRT) สามารถพิจารณา “ปฏิสัมพันธ์” ดังกล่าวได้จากความแตกต่าง ของค่าพารามิเตอร์อำนาจจำแนกของข้อสอบระหว่างผู้สอบกลุ่มย่อยสองกลุ่ม กล่าวคือ ถ้าข้อสอบทำหน้าที่ต่างกันแบบเอกรูป แล้วโค้งลักษณะข้อสอบ (Item Characteristic Curves; ICCs) ระหว่างผู้สอบกลุ่มย่อยสองกลุ่มจะขนานกัน หรือมีฟังก์ชันการตอบสนองข้อสอบ (Item Response Functions; IRFs) เหมือนกัน แต่ถ้าข้อสอบทำหน้าที่ต่างกันแบบอเนกรูปแล้วโค้งลักษณะข้อสอบระหว่างผู้สอบกลุ่มย่อยทั้งสองจะไม่ขนานกัน หรือมีฟังก์ชันการตอบสนองข้อสอบต่างกัน ดังนั้นความแตกต่างระหว่างโค้งลักษณะข้อสอบทั้งสองแบบจะบ่งบอกถึงขนาดและทิศทางของข้อสอบที่ทำหน้าที่ ต่างกัน ซึ่งสามารถคำนวณได้โดยใช้สูตรการคำนวณพื้นที่ของ Raju (1990)

ข้อสอบที่ทำหน้าที่ต่างกันแบบอเนกรูป สามารถจำแนกได้เป็น 2 ลักษณะ (Swaminathan and Roger, 1990) ดังนี้

1. ข้อสอบทำหน้าที่ต่างกันแบบอเนกรูปโดยมีปฏิสัมพันธ์ไม่เป็นลำดับ (Disordinal interaction) เป็นการทำหน้าที่ต่างกันสำหรับกลุ่มผู้สอบซึ่งเกิดขึ้นเมื่อโค้งลักษณะข้อสอบตัดกัน
2. ข้อสอบทำหน้าที่ต่างกันแบบอเนกรูปโดยมีปฏิสัมพันธ์เป็นลำดับ (Ordinal interaction) เป็นการทำหน้าที่ต่างกันสำหรับกลุ่มผู้สอบซึ่งเกิดขึ้นเมื่อโค้งลักษณะข้อสอบต่างกันอย่างไม่สม่ำเสมอ แต่ไม่ตัดกัน หรืออาจตัดกันนอกช่วง

1.4 หลักการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF detection) เป็นการเปรียบเทียบผลการตอบข้อสอบเป็นรายข้อระหว่างกลุ่มผู้สอบอย่างน้อย 2 กลุ่ม ที่มีความสามารถหลัก (primary ability) ที่มุ่งวัดเท่ากัน แต่คาดว่าจะมีความได้เปรียบหรือเสียเปรียบกัน โดยกลุ่มหนึ่งถือเป็น กลุ่มอ้างอิง (Reference group) ซึ่งคาดว่าน่าจะได้เปรียบในการตอบข้อสอบข้อนั้น หรือมีโอกาสตอบข้อสอบได้ถูกต้องมากกว่า ส่วนอีกกลุ่มคือ กลุ่มเสียเปรียบ (Focal group) ซึ่งเป็นกลุ่มที่สนใจศึกษา และคาดว่าน่าจะเป็นกลุ่มที่เสียเปรียบ

การเปรียบเทียบผลการตอบข้อสอบระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบจำเป็นต้องจับคู่ (matching) ผู้สอบตามความสามารถ ซึ่งเป็นเงื่อนไขสำคัญของการตรวจสอบการทำหน้าที่ต่างกันข้อสอบ เกณฑ์การจับคู่matching criteria) ที่นิยมใช้กันมี2 วิธี (ศิริชัย กาญจนวาสี 2550)ดังนี้

1. เกณฑ์ภายนอก (external criteria)

การวิเคราะห์การทำหน้าที่ต่างกันโดยใช้เกณฑ์ภายนอกนี้ สามารถนำไปใช้ทั้งข้อสอบรายข้อและแบบสอบทั้งฉบับ โดยการใช้คะแนนจากแบบสอบอื่นเป็นเกณฑ์ภายนอกแล้วใช้เทคนิคการวิเคราะห์การถดถอย (regression analysis) เพื่อทำการเปรียบเทียบเส้นกราฟความสัมพันธ์ระหว่างตัวแปรเกณฑ์ กับตัวแปรทำนายระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ

หลักการนี้มีจุดมุ่งหมาย เพื่อสร้างสมการทำนายตัวแปรเกณฑ์ ซึ่ งจะเป็นคะแนนของแบบสอบอื่นจากตัวแปรทำนายซึ่งเป็ นคะแนนรายข้อ หรือคะแนนแบบสอบระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ จะใช้คะแนนรายข้อเป็นตัวแปรทำนาย แต่ถ้าเป็นการวิเคราะห์การทำหน้าที่ต่างกันของแบบสอบ จะใช้คะแนนรวมของแบบสอบทั้งฉบับเป็นตัวแปรทำนายสำหรับตัวแปรเกณฑ์ที่ใช้เป็นเกณฑ์ภายนอก อาจใช้คะแนนรวมทั้งฉบับ หรือเกรดเฉลี่ย หรือผลสัมฤทธิ์ในงานที่เกี่ยวข้องของผู้สอบ (Cronbach, 1970) สมการทำนายสำหรับกลุ่มอ้างอิงและกลุ่มเปรียบเทียบแสดงได้ดังนี้

กลุ่มอ้างอิง (R) $Y_i = A_R + B_R X_i$

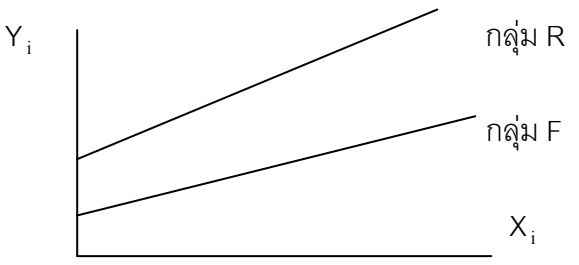
กลุ่มเปรียบเทียบ (F) $Y_i = A_F + B_F X_i$

เมื่อ Y_i = คะแนนของตัวแปรเกณฑ์ภายนอก

X_i = คะแนนของตัวแปรทำนาย

A = ค่าคงที่หรือค่าตัดแกน (intercept)

B = ค่าความชัน (slope)



การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบและแบบสอบโดยใช้เทคนิคการวิเคราะห์สมการถดถอย

จากฟังก์ชันการทำนายทั้ง 2 สมการ สามารถเปรียบเทียบค่าตัดแกน (A) และค่าความชัน (B) ของเส้นกราฟระหว่างกลุ่มอ้างอิง และกลุ่มเปรียบเทียบได้ ถ้าเส้นกราฟดังกล่าวมีค่า ความชัน หรือค่าตัดแกน แตกต่างกัน สำหรับข้อสอบใด หรือแบบสอบใดแสดงว่าข้อสอบหรือแบบสอบนั้นมีการทำหน้าที่ต่างกัน โดยเข้าข้างกลุ่มผู้สอบที่มีค่าตัดแกนหรือค่าความชันที่สูงกว่า

การใช้เกณฑ์ภายนอกมีข้อดี คือเกณฑ์ที่ใช้มีความเป็นอิสระจากข้อสอบ และแบบสอบที่ต้องการตรวจสอบ แต่มีจุดอ่อนตรงที่ความเหมาะสมของเกณฑ์ที่จะนำมาใช้ ในทางปฏิบัติเป็นการยากที่จะหาตัวแปรเกณฑ์ภายนอกจากแบบสอบฉบับอื่นที่มีความตรงเชิงทำนาย และมีความยุติธรรมสำหรับกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ ถ้าตัวแปรเกณฑ์ภายนอกขาดคุณสมบัติดังกล่าวจะทำให้ผลการวิเคราะห์ การทำหน้าที่ต่างกันของข้อสอบหรือแบบสอบขาดความแม่นยำ และความสมบูรณ์

2. เกณฑ์ภายใน (Internal Criterion)

การวิเคราะห์การทำหน้าที่ต่างกัน โดยใช้เกณฑ์ภายในเป็นการนำวิธีการทางสถิติมาตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ หรือแบบสอบ โดยเน้นการพิจารณาจากโครงสร้างภายในของแบบสอบเป็นหลัก ด้วยการวิเคราะห์ผลจากการตอบข้อสอบและความสามารถหรือคะแนนจริงของผู้สอบที่ได้จากแบบสอบฉบับนั้น เพื่อนำมาเปรียบเทียบระหว่างผู้สอบจากกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่มีความสามารถหรือคะแนนจริงเท่ากัน ว่าจะมีผลการตอบหรือโอกาสในการตอบข้อสอบได้ถูกต้องแตกต่างกันหรือไม่ เพื่อบ่งชี้ถึงการทำหน้าที่ต่างกันของข้อสอบ การวิเคราะห์ในลักษณะนี้นิยมใช้ค่าสถิติต่าง ๆ เป็นตัวบ่งชี้ถึงการทำหน้าที่ต่างกันของข้อสอบ ค่าสถิติทดสอบที่นิยมนำมาใช้พอสรุปได้ดังนี้

- การทดสอบปฏิสัมพันธ์ (Interaction)

ในระยะเริ่มแรกของกา รศึกษาความลำเอียงของข้อสอบ มีการใช้สถิติทดสอบเอฟ (F-test) จากการวิเคราะห์ความแปรปรวน (ANOVA) เพื่อทดสอบปฏิสัมพันธ์ระหว่างกลุ่มผู้สอบกับข้อสอบ ถ้าการทดสอบมีนัยสำคัญเป็นสัญญาณของการทำหน้าที่ต่างกันของข้อสอบ (Cleary & Hilton, 1968; Jensen, 1974) จากนั้นจึงทำก ารวิเคราะห์ต่อด้วยวิธีการ Post Hoc เพื่อระบุข้อสอบที่มีผลต่อการเกิดปฏิสัมพันธ์ ซึ่งเป็นข้อที่ทำหน้าที่ต่างกัน

วิธีการนี้มีข้อดีที่สามารถศึกษาผู้สอบหลาย ๆ กลุ่มได้สะดวก แต่มีจุดอ่อนในเรื่องการควบคุมกลุ่มต่าง ๆ ให้มีความสามารถที่ทัดเทียมกัน ขนาดกลุ่มตัวอย่า งของกลุ่มต่าง ๆ และอัตราความคลาดเคลื่อนประเภทที่ 1 จะสูงขึ้น ถ้าจำนวนข้อสอบเพิ่มมากขึ้น

- การวัดความเบี่ยงเบนสัมพันธ์ (Relative Deviation)

การคำนวณค่าความยากของข้อสอบ เช่น p, b เป็นต้น เมื่อคำนวณแยกระหว่างกลุ่มและแปลงให้เป็นค่าความยากมาตรฐาน (Δ) สามารถนำมาพล๊อตเปรียบเทียบเป็นรายข้อ ถ้าข้อใดเบี่ยงเบนไปจากแกนหลักที่คาดหมาย หรือเบี่ยงเบนเกินจากความคลาดเคลื่อนมาตรฐานของค่าความยากที่กำหนด ย่อมแสดงถึงการทำหน้าที่ต่างกันของข้อสอบ (Cleary & Hilton, 1968 ; Angoff & Ford, 1973) รวมทั้งสามารถคำนวณค่าสหสัมพันธ์ระหว่างค่าความยากรายข้อระหว่างกลุ่ม เพื่อแสดงถึงการทำหน้าที่ต่างกันของแบบสอบ ถ้าสหสัมพันธ์เข้าใกล้ 1.00 แสดงว่าค่าความยากสัมพัทธ์ของข้อสอบมีค่าใกล้เคียงกันระหว่างกลุ่ม ดังนั้นแบบสอบวัดคุณลักษณะคล้ายกันระหว่างกลุ่ม

วิธีการนี้มีข้อดีและข้อเสียค ล้ายการทดสอบปฏิสัมพันธ์ นอกจากนี้ค่าความยากของข้อสอบ (p) มิใช่ตัวแทนของค่าความยากจริงของข้อสอบ และได้รับอิทธิพลจากค่าแทรกซ้อนอื่น ได้แก่ ค่าอำนาจจำแนก และความสามารถของผู้สอบ

- การเปรียบเทียบน้ำหนักตัวประกอบ (Factor Loading)

การวิเคราะห์ตัวประกอบ (Factor Analysis) เป็นเทคนิคทางสถิติที่นิยมใช้ในการตรวจสอบความตรงเชิงทฤษฎีหรือโครงสร้าง (Construct Validity) เมื่อนำการวิเคราะห์ตัวประกอบมาใช้ในการวิเคราะห์โครงสร้างของแบบสอบแยกตามกลุ่มผู้สอบ ความไม่สอดคล้องกันระหว่างน้ำหนักตัวประกอบบนคุณลักษณะสำคัญที่มุ่งวัด หรือ คว ามแตกต่างของค่าเฉลี่ยคะแนนตัวประกอบ (Factor scores) ระหว่างกลุ่มผู้สอบ ย่อมสะท้อนการทำหน้าที่ต่างกันของข้อสอบและแบบสอบ

การใช้เทคนิคการวิเคราะห์ตัวประกอบเชิงสำรวจ (Exploratory Factor Analysis; EFA) สำหรับศึกษาการทำหน้าที่ต่างกัน ของข้อสอบมีจุดอ่อนในเรื่องคว ามไม่สอดคล้องระหว่างน้ำหนัก

ตัวประกอบอาจเกิดจากความแตกต่างของความสามารถระหว่างกลุ่มก็ได้ แนวทางที่เหมาะสมจึงควรใช้เทคนิคการวิเคราะห์ตัวประกอบเชิงยืนยัน (Confirmatory Factor Analysis; CFA) นอกจากนี้ยังสามารถใช้ CFA สำหรับตรวจสอบความแตกต่างระหว่างกลุ่ม ในด้าน คุณลักษณะหรือความสามารถหลัก และความสามารถรองได้อีกด้วย (Camilli & Shepard, 1994)

- การเปรียบเทียบโอกาสตอบข้อสอบถูก

การวิเคราะห์โอกาสตอบข้อสอบถูกของผู้สอบจากกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่มีความสามารถเท่ากัน เป็นแนวทางสำคัญที่นิยมใช้กันและเป็นที่ยอมรับในปัจจุบัน สำหรับบ่งชี้การทำหน้าที่ต่างกันของข้อสอบ มีการคำนวณค่าสถิติ 2 แนวทางดังนี้

1. เปรียบเทียบค่าสัดส่วนหรือความน่าจะเป็นในการตอบข้อสอบถูกของผู้สอบต่างกลุ่มที่มีความสามารถเท่ากัน เช่น วิธีแมนเทล – แฮนส์เซล (MH) เป็นต้น
2. เปรียบเทียบค่าฟังก์ชันการตอบสนองข้อส อบ หรือโค้งลักษณะข้อสอบระหว่างกลุ่มที่มีระดับความสามารถเท่ากัน เป็นวิธีที่อยู่บนพื้นฐานของทฤษฎี IRT เช่น วิธีวัดความแตกต่างของพื้นที่ วิธีวัดความแตกต่างของค่าพารามิเตอร์ความยาก วิธีการทดสอบไค – สแควร์ของ ลอร์ด (Lord's χ^2 - test) เป็นต้น

วิธีการนี้มีข้อดีที่สำคัญได้แก่ การคำนวณค่าสถิติของข้อสอบมีความน่าเชื่อถือ มีกลไกลควบคุมความสามารถของผู้สอบโดยการจับคู่กลุ่มความสามารถ เพื่อทำการเปรียบเทียบ ณ ตำแหน่งต่างๆ ที่มีความสามารถเท่ากัน จึงเป็นวิธีการที่ยอมรับกันทั่วไป แต่มีข้อจำกัดในด้านความสลับ ซับซ้อนของแนวคิดพื้นฐาน และการวิเคราะห์มีความจำเป็นต้องใช้โปรแกรมคอมพิวเตอร์โดยเฉพาะ

1.5 วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบและการทำหน้าที่ต่างกันของตัวลวง

วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF detection) มีหลายวิธีด้วยกัน ทั้งนี้เพราะมีวิ ธีการศึกษาและการคิดค้นวิธีการต่างๆ เพื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้อย่างมีประสิทธิภาพมากที่สุด ซึ่งสามารถแบ่งตามประเภทการวิเคราะห์ได้เป็น 2 กลุ่ม คือ (Potenza and Dorans อ้างถึงใน อิทธิฤทธิ์ พงษ์ปิยรัตน์, 2551) โดยวิธีการตรวจสอบการทำ

หน้าที่ของตั วลวง ส่วนใหญ่ก็ใช้แนวคิดเดียวกับการตรวจสอบการทำหน้าที่ของตัวถูกด้วย เช่นเดียวกัน (Koon, 2010) โดยวิธีการที่ใช้สามารถสรุปได้ดังนี้

1. กลุ่มวิธีที่ใช้คะแนนสังเกตได้ (Observed score) เป็นกลุ่มที่ใช้คะแนนรวมของแบบสอบเป็นเกณฑ์ในการจับคู่ผลสอบสองกลุ่มตามความรู้ หรื อความสามารถที่แท้จริงของผู้สอบ วิธีการในกลุ่มนี้ได้แก่ วิธีแมนเทล -แฮนส์เซล วิธีถดถอยโลจิสติก และวิธีทำให้เป็นมาตรฐาน จุดเด่นของวิธีการในกลุ่มนี้คือ กลุ่มตัวอย่างขนาดเล็ก การวิเคราะห์ไม่ยุ่งยากซับซ้อน ส่วนจุดด้อยของวิธีการในกลุ่มนี้คือค่าสถิติเปลี่ยนไปตามขนาดกลุ่มตัวอย่าง เมื่อกลุ่มตัวอย่างที่ศึกษาเปลี่ยนไป ผลการศึกษาก็อาจเปลี่ยนแปลงไปด้วย

2. กลุ่มวิธีที่ใช้คะแนนสังเกตไม่ได้หรือคะแนนของตัวแปรแฝง (Latent variable) เป็นกลุ่มที่มีทฤษฎีการทดสอบเป็นพื้นฐาน มุ่งประมาณค่าคุณลักษณะแฝง หรือใช้คะแนนจริงของผู้สอบเป็นเกณฑ์ในการจับคู่เปรียบเทียบผู้สอบ วิธีการในกลุ่มนี้ได้แก่ วิธีการตอบสนองข้อสอบ (IRT) และ วิธีซิบเทสท์ (SIBTEST) เป็นต้น

นอกจากนี้ ศิริชัย กาญจนวาสี (2550) ได้แบ่งกลุ่มจำแนกวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบตามลักษณะการตรวจให้คะแนนได้เป็น 2 ประเภทคือ ข้อสอบที่มีการให้คะแนนแบบทวิวิภาค หรือสองค่า (Dichotomous scoring) และข้อสอบที่มีการให้คะแนนแบบพหุวิภาค หรือหลายค่า (Polytomous scoring) โดยมีรายละเอียดดังต่อไปนี้

1. วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ให้คะแนนแบบทวิวิภาค

1.1 กลุ่มวิธีที่ใช้คะแนนที่สังเกตได้

วิธีในกลุ่มนี้มักวิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) หรือกลุ่มที่ไม่ใช้ทฤษฎีการตอบสนองข้อสอบ (Non – IRT approach) โดยใช้คะแนนรวมของผู้สอบเป็นเกณฑ์การจับคู่ของกลุ่มผู้สอบ วิธีการตรวจสอบที่สำคัญในกลุ่มนี้ ได้แก่

- การวิเคราะห์ความแปรปรวน (ANOVA)
- วิธีการวิเคราะห์การถดถอยโลจิสติก (Logistic Regression ; LR)
- วิธีแปลงค่าความยากของข้อสอบ (Transformed Item Difficulty : TID)
- วิธีแมนเทล – แฮนส์เซล (Mantel – Haenszel ; MH)
- วิธีดัชนีมาตรฐาน (Standardization ; STND)

1.2 กลุ่มวิธีที่ใช้คุณลักษณะแฝง

วิธีในกลุ่มนี้ใช้คุณลักษณะหรือตัวแปรแฝง(trait) ซึ่งวิเคราะห์บนพื้นฐานของทฤษฎีการตอบสนองข้อสอบ (IRT) สำหรับใช้เป็นเกณฑ์การจัดคู่กลุ่มผู้สอบ วิธีการตรวจสอบที่สำคัญในกลุ่มนี้ได้แก่

- วิธีวัดพื้นที่ความแตกต่างระหว่างโค้งการตอบสนองข้อสอบ (IRT - D^2)
- วิธีไค – สแควร์ของลอร์ด (Lord's χ^2)
- วิธีอัตราส่วนไลค์ลิฮูดทั่วไป (General IRT Likelihood Ratio)
- วิธีอัตราส่วนไลค์ลิฮูด ลอกลิเนียร์ (Loglinear IRT Likelihood Ratio)
- วิธีซิปเทสท์ (SIBTEST)

2. วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ให้คะแนนแบบพหุวิภาค

2.1 กลุ่มวิธีที่ใช้คะแนนที่สังเกตได้

- วิธีการวิเคราะห์ความแปรปรวน (ANOVA)
- วิธีการวิเคราะห์การถดถอยโลจิสติกพหุวิภาค (Polytomous Logistic Regression)
- วิธีดัชนีมาตรฐานพหุวิภาค (Polytomous Standardization)
- วิธีแมนเทล – แฮนส์เซลทั่วไป (General Mantel – Haenzel ; GMH)

2.2 กลุ่มวิธีที่ใช้คุณลักษณะแฝง

- วิธีอัตราส่วนไลค์ลิฮูดในรูปทั่วไป (General IRT Likelihood Ratio)
- วิธีการให้คะแนนบางส่วน (Partial Credit Model ; PCM)
- วิธีซิปเทสท์พหุวิภาค (Polytomous SIBTEST)
- วิธีการให้คะแนนบางส่วนทั่วไป (Generalized Partial Credit Model; GPCM)

การศึกษาเกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning; DIF) เป็นเรื่องที่ได้รับความสนใจมาตั้งแต่ในอดีตจนถึงปัจจุบัน โดยเป็นลักษณะของการตรวจสอบข้อสอบที่ทำให้ผู้สอบที่มีความสามารถเท่ากันมีโอก าสทำข้อสอบถูกต้องแตกต่างกัน โดยการทำหน้าที่ต่างกันของข้อสอบเกิดขึ้นได้ 2 ลักษณะ คือการทำหน้าที่ต่างกันแบบเอกรูป

หมายถึงข้อสอบที่ทำให้ผู้สอบกลุ่มหนึ่งมีโอกาสในการตอบข้อสอบถูกมากกว่าผู้สอบอีกกลุ่มหนึ่งอย่างสม่ำเสมอกันในทุกระดับความสามารถ และการทำหน้าที่ต่างกันแบบอเนกรูป หมายถึงข้อสอบที่ทำให้โอกาสในการตอบข้อสอบถูกของผู้สอบระหว่างกลุ่มแตกต่างกันอย่างไม่สม่ำเสมอกันในทุกระดับความสามารถ โดยวิธีการที่ใช้สำหรับตรวจสอบการทำหน้าที่ต่างกันของข้อสอบก็ได้รับการพัฒนามาอย่างต่อเนื่องเช่นเดียวกัน โดยกลุ่มวิธีที่นิยมใช้กันอย่ างกว้างขวางในปัจจุบันคือ กลุ่มวิธีที่ใช้โมเดล IRT เช่น วิธีวัดพื้นที่ความแตกต่างระหว่างโค้งการตอบสนองข้อสอบ (IRT - D^2) วิธีไค – สแควร์ของลอร์ด (Lord's χ^2) วิธีอัตราส่วนไลค์ลิฮูดทั่วไป (General IRT Likelihood Ratio) วิธีอัตราส่วนไลค์ลิฮูด ลอกลิเนียร์ (Loglinear IRT Likelihood Ratio) หรือวิธีซิปเทสท์ (SIBTEST) เป็นต้น

สำหรับในการศึกษาครั้งนี้ผู้วิจัยเลือกใช้วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) โดยวิธีการวิเคราะห์แมนเทล-แฮนส์เซลร่วมกับอัตราส่วนแต้มต่อ (odds ratio) แสดงสูตรในการวิเคราะห์ได้ดังนี้

$$\hat{\alpha}_{MH_i} = \frac{\sum_{s=1}^{s} R_{1s} F_{0s} / n_s}{\sum_{s=1}^{s} R_{0s} F_{1s} / n_s} \qquad (1)$$

และแมนเทล-แฮนส์เซลร่วมกับลอกของอัตราส่วนแต้มต่อ (odds ratio)แสดงสูตรในการวิเคราะห์ได้ดังนี้

$$\hat{\lambda}_j = \ln(\hat{\alpha}_j) \qquad (2)$$

เนื่องจากการวิเคราะ ห์แมนเทล- แฮนส์เซลร่วมกับลอกของอัตราส่วนออด โดยเป็นการเปรียบเทียบอัตราส่วนการตอบของกลุ่มอ้างอิง (Reference) กับกลุ่มเปรียบเทียบ (Focal) ดังนั้นการแปลผลสามารถทำได้ดังนี้

การแปลผลสำหรับการทำหน้าที่ต่างกันของข้อสอบ (DIF)

$\ln(\hat{\alpha}_j) = 0$ หมายถึง ไม่พบการทำหน้าที่ต่างกันของข้อสอบ

$\ln(\hat{\alpha}_j) > 0$ หมายถึง กลุ่มอ้างอิง (Reference) มีโอกาสที่จะตอบถูกมากกว่ากลุ่มเปรียบเทียบ (Focal)

$\ln(\hat{\alpha}_j) < 0$ หมายถึง กลุ่มเปรียบเทียบ (Focal) มีโอกาสที่จะตอบถูกมากกว่ากลุ่มอ้างอิง (Reference)

การแปลผลสำหรับการทำหน้าที่ต่างกันของตัวลวง (DDF)

$\ln(\hat{\alpha}_j)$ = 0 หมายถึง ไม่พบการทำหน้าที่ต่างกันของตัวลวง

$\ln(\hat{\alpha}_j)$ > 0 หมายถึง กลุ่มเปรียบเทียบ (Focal) มีโอกาสที่จะเลือกตอบมากกว่ากลุ่มอ้างอิง (Reference)

$\ln(\hat{\alpha}_j)$ < 0 หมายถึง กลุ่มอ้างอิง (Reference) มีโอกาสที่จะเลือกตอบมากกว่ากลุ่มเปรียบเทียบ (Focal)

1.6 เกณฑ์สำหรับเปรียบเทียบคุณภาพของวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

คุณภาพของวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบนิยมพิจารณาจากดัชนีบ่งชี้สำคัญ 2 ตัว ได้แก่ อำนาจการทดสอบ หรืออัตราความถูกต้องของการตรวจพบข้อที่ทำหน้าที่ต่างกัน (Power rate) และความคลาดเคลื่อนของการตรวจสอบ หรืออัตราตามความคลาดเคลื่อนประเภทที่ 1 (Type 1 error rate) ซึ่งเป็นโอกาสของการเกิดความคลาดเคลื่อนในลักษณะที่ตรวจพบว่า ข้อสอบทำหน้าที่ต่างกันทั้งๆ ที่ความเป็นจริงข้อสอบไม่ได้ทำหน้าที่ต่างกัน

ในการคำนวณค่าสถิติตามวิธีตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ มีจุดมุ่งหมายเพื่อทดสอบนัยสำคัญของผลการตรวจสอบ โดยมีสมมติฐานศูนย์ของการทดสอบ คือ ข้อสอบไม่ได้ทำหน้าที่ต่างกัน หรือ Ho : No DIF ผลการทดสอบสมมุติฐานของวิธีการตรวจสอบ DIF วิธีต่าง ๆ นำสู่การตัดสินใจว่า ยอมรับสมมุติฐานศูนย์ (Accept Ho) หรือปฏิเสธสมมุติฐานศูนย์ (Reject Ho) ผลของการตัดสินใจมีโอกาสเกิดเหตุการณ์ 4 ลักษณะ รายละเอียดดังตารางที่ 1

ตารางที่ 1 คุณภาพของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

การตัดสินใจตามผลการตรวจสอบ	Ho: No DIF **ความเป็นจริง**	
	Ho ถูก	**Ho ผิด**
Accept Ho	• ตัดสินใจถูก (True negative) • ระดับความเชื่อมั่น (1-α)	• ตัดสินผิด (Type II Error, β) • False negative
Reject Ho	• ตัดสินผิด (Type I Error, α) • False positive	• ตัดสินถูก (True positive) • อำนาจการทดสอบ (1-β)

จากค่าสถิติที่คำนวณตามวิธีการตรวจสอบการทำหน้าที่ต่ างกันของข้อสอบ จะนำไปสู่การตัดสินใจสรุปผลการตรวจสอบดังนี้

1) ตัดสินใจถูก มีโอกาสเกิดขึ้น 2 ลักษณะคือ การสรุปถูกว่า
   - ข้อสอบ No DIF ตามความเป็นจริง (True negative)
   - ข้อสอบ DIF ตามความเป็นจริง (True positive)
2) ตัดสินใจผิด มีโอกาสเกิดขึ้น 2 ลักษณะคือ การสรุปผิดว่า
   - ข้อสอบ DIF ทั้ง ๆ ที่ความจริงข้อสอบ No DIF (False positive)
   - ข้อสอบ No DIF ทั้ง ๆ ที่ความจริงข้อสอบ DIF (False negative)

เนื่องจากอำนาจการทดสอบ (1-β) กับ β เป็นค่าดัชนีมีสเกลที่ผกผันกันและความคลาดเคลื่อนประเภทที่ 1 (α) กับ 1-α ต่างก็เป็นดัชนีมีสเกลที่ผกผันเช่นเดียวกัน ดังนั้นการพิจารณาดัชนีบ่งชี้คุณภาพ 2 ตัว คือ อัตราความถูกต้องของการตรวจพบ DIF (Power rate) และความคลาดเคลื่อนประเภทที่ 1 (Type I error rate) ก็เพียงพอที่จะให้สารสนเทศครบทั้ง 4 เหตุการณ์

จากแนวคิดที่ผู้วิจัยได้ศึกษา พบว่า การทำหน้าที่ต่างกันของตัวลวง (Differential Distractor Functioning; DDF) เป็นประเด็นที่นักวัดผลยังมีการศึกษาไม่มาก โดยเป็นการเข้าไปศึกษาถึงตัวเลือกที่ไม่ใช่คำตอบที่ถูกต้อง กระบวนการที่ใช้ใน การศึกษามีหลักการที่มีความคล้ายคลึงกับการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning; DIF)

ที่เป็นการวิเคราะห์ตัวเลือกที่เป็นคำตอบที่ถูกต้อง ซึ่งผลจากการศึกษา ในเรื่องของการทำหน้าที่ต่างกันของตัวลวงนั้น จะทำให้ผู้วิจัยได้ รับสารสนเทศที่ สำคัญในเรื่องที่เกี่ยวกับการทำหน้า ที่ของข้อสอบข้อนั้นๆมากขึ้น ดังนั้นการขยายความรู้เกี่ยวกับการทำหน้าที่ต่างกันของตัวลวง จึงเป็นการช่วยเพิ่มมิติในการเข้าใจเรื่องของการทำหน้าที่ต่างกันของข้อสอบได้อย่างสมบูรณ์

1.7 งานวิจัยที่เกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง

การศึกษาในครั้งนี้ ผู้วิจัยต้องการเข้าไปศึกษา เกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบ โดยเน้นไปที่วิธีที่ใช้สำหรับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และต้องการขยายความรู้ในเรื่องของการทำหน้าที่ต่างกันของตัวลวง ดังนั้นการทบทวน ศึกษาเอกสาร และงานวิจัยที่เกี่ยวข้องผู้วิจัยจึงนำเสนองานที่เกี่ยวข้องกับการทำหน้าที่ต่างกันของข้อสอบทั้งที่เป็นการวิเคราะห์ตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง และโมเดลการวิเคราะห์มูลค่าเพิ่ม โดยมีรายละเอียดดังนี้

1.7.1 การทำหน้าที่ต่างกันของข้อสอบ

ทัศนีย์ พีรมนตรี (2529) ได้ทำการวิเคราะห์ความลำเอียงของแบบสอบวิชาคณิตศาสตร์ในโครงการตรวจสอบคุณภาพการศึกษา ซึ่งพัฒนาแบบสอบโดยสำนักงานทดสอบทางการศึกษา กรมวิชาการ โดยใช้วิธีวิเคราะห์ 3 วิธี คือ วิธีกำหนดค่าเดลต้า วิธีทดสอบความแตกต่างระหว่างกลุ่มด้วยสถิติไค – สแควร์ ในโมเดลลอกลิเนีย 2 โมเดล คือโมเดลที่ไม่มีพารามิเตอร์ผลร่วมระหว่างระดับคะแนนกับกลุ่ม และโมเดลที่ไม่มีพารามิเตอร์ของผลหลักที่เกิดจากกลุ่ม และวิธีการตอบสนองข้อสอบ 3 พารามิเตอร์ โดยกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2526 ทั่วทุกภาคของประเทศ ได้แก่ ภาคกลาง ภาคเหนือ ภาคใต้ ภาคตะวันออก และภาคตะวันออกเฉียงเหนือ

ผลการศึกษาพบว่าวิธีการตอบสนองข้อสอบ 3 พารามิเตอร์ พบข้อสอบที่มีความลำเอียงในแต่ละภาคจำนวนข้อมากที่สุด วิธีกำหนดจุดค่าเดลต้าและวิธีการตอบสนองข้อสอบ 3

พารามิเตอร์ มีข้อลำเอี ยงซ้ำกันมากที่สุดระหว่างกรุงเทพมหานครและภาคตะวันออกเฉียงเหนือ ส่วนวิธีไค – สแควร์มีจำนวนข้อที่ซ้ำกันมากที่สุดระหว่างกรุงเทพมหานครและภาคตะวันออก

สุรศักดิ์ อมรรัตนศักดิ์ (2530) ได้ทำการศึกษาเปรียบเทียบผลของวิธีวิเคราะห์ความลำเอียงของข้อสอบ 4 วิธี คือ วิธีวิเคราะห์ความแปรปรวน วิธีแปลงค่าความยากของข้อสอบและวิธีโค้งลักษณะข้อสอบ ที่มีพารามิเตอร์ 2 ตัวและ 3 ตัว โดยใช้ข้อมูลจากแบบสอบแข่งขันเพื่อบรรจุเข้ารับราชการเป็นข้าราชการครู สังกัดสำนักงานคณะกรรมการประถมศึกษาแห่งชาติ ปี พ .ศ. 2529 กลุ่มตัวอย่างคือผู้เข้าสอบแข่ งขัน ตัวแปรที่ศึกษาคือ ตัวแปรเพศ โดยทำการเปรียบเทียบความแตกต่างของผลการคัดเลือกก่อนและหลังการศึกษาความลำเอียงของข้อสอบในด้าน จำนวนผู้ที่ได้รับการคัดเลือก สัดส่วนของชาย : หญิง ที่ได้รับการคัดเลือกและความเที่ยงของแบบสอบ

ผลการวิจัยพบว่า วิธีวิเคราะห์ความลำเอี ยงทั้ง 4 วิธี พบข้อสอบที่มีความลำเอียงแตกต่างกันโดยวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว พบข้อสอบที่ลำเอียงจำนวนมากที่สุด รองลงมาคือ วิธีวิเคราะห์ความแปรปรวน และวิธีที่พบข้อสอบที่ลำเอียงจำนวนน้อยที่สุดคือ วิธีแปลงค่าความยากของข้อสอบ วิธีทั้ง 4 วิธี มีความสัมพันธ์กันทางบวกอย่างมีนัยสำคัญที่ระดับ 0.001 และเมื่อใช้การคิดคะแนนดิบและคะแนนรวมที่แตกต่างกัน 5 วิธี พบว่ามีจำนวนผู้ที่ได้รับการคัดเลือกแตกต่างกันประมาณร้อยละ 4-24 และเมื่อใช้คะแนนมาตรฐานที่ปกติรวมกับคะแนนแปลงแบบอื่น ๆ 4 วิธี พบว่าจำนวนผู้ที่ได้รับก ารคัดเลือกแตกต่างกันร้อยละ 4-23 และเมื่อตัดข้อสอบที่มีความลำเอียงออกแล้วพบว่าจำนวนหญิงและชายที่ได้รับการคัดเลือกมีจำนวนใกล้เคียงกันมากขึ้นกว่าก่อนตัดข้อสอบที่มีความลำเอียง และค่าความเที่ยงของแบบสอบหลังการตัดข้อสอบที่ลำเอียงออกจากแบบสอบ ลดลงเล็กน้อย

สุพัฒน์ สุกมลสันต์ (2534) ได้ทำการศึกษาความลำเอียงของข้อสอบภาษาอังกฤษ ชุด กข และ กขค เข้ามหาวิทยาลัย ปี พ .ศ. 2531 – 2533 จำนวน 6 ฉบับ ฉบับละ 100 ข้อ โดยวิเคราะห์ความลำเอียงต่อเพศและต่อภาคภูมิศาสตร์ของผู้สมัครสอบ เปรียบเทียบวิธีวิเคราะห์ 3 วิธี คือ วิธีกำหนดจุดค่าเดลต้า วิธีไค – สแควร์ชนิดแบ่งความสามารถของผู้สอบเป็น 3 ระดับ ได้แก่ กลุ่มผู้ที่มีความสามารถระดับต่ำ (ผู้ที่ได้คะแนน 0 – 40 คะแนน) กลุ่มผู้ที่มีความสามารถระดับปานกลาง (ผู้ที่ได้คะแนนรวม 41 – 70 คะแนน) และระดับสูง (ผู้ที่ได้คะแนนรวม 71 – 100

คะแนน) และวิธีการวัดพื้นที่ความแตกต่างระหว่างโค้งลักษณะข้อสอบที่วิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ กลุ่มตัวอย่างที่ใช้คือ ผู้สมัครสอบคัดเลือกเข้ามหาวิทยาลัย ซึ่งสอบแบบสอบภาษาอังกฤษเข้ามหาวิทยาลัย ชุด กข และ ชุด กขค ตัวแปรที่ศึกษาความลำเอียง คือ เพศ และภาคภูมิศาสตร์ของผู้สอบซึ่งแยกออกออกเป็น 5 ภาคตามภูมิลำเนาของผู้สอบ ได้แก่ ภาคกลาง ภาคตะวันออก ภาคตะวันออกเฉียงเหนือ ภาคเหนือ และภาคใต้

ผลการวิจัยพบว่าแบบสอบวิชาภาษาอังกฤษ ชุด กข และชุด กขค ปี 2531 – 2533 มีความลำเอียงต่อเพศ โดยเฉลี่ยประมาณฉบับละ ร้อยละ 7 – 28 และ 4 – 41 ตามลำดับและมีความลำเอียงต่อภาคภูมิศาสตร์โดยเฉลี่ยประมาณฉบับละ ร้อยละ 6 – 45 และ 5 – 43 ตามลำดับ วิธีการวิเคราะห์ทั้ง 3 วิธี พบข้อสอบที่มีความลำเอียงต่อเพศและต่อภาคภูมิศาสตร์ของผู้สอบจำนวนแตกต่างกันอย่างมีนัยสำคัญแต่จำนวนข้อที่ลำเอียงแต่ละวิธีมี ความสัมพันธ์กันอย่างไม่มีนัยสำคัญ วิธีโลจิสติก 3 พารามิเตอร์จะพบข้อสอบที่ทำหน้าที่ต่างกันจำนวนมากที่สุด รองลงมาคือ วิธีไค – สแควร์ และวิธีกำหนดจุดเดลต้าพบข้อสอบที่ลำเอียงจำนวนน้อยที่สุด

กาญจนา วัธนสุนทร (2537) ได้พัฒนาเกณฑ์การตัดสินข้อสอบลำเอียงทางเพศด้วยข้อมูลเชิงประจักษ์ ด้วย 3 วิธีการตรวจสอบคือ วิธีทฤษฎีการตอบสนองข้อสอบโมเดล 2 พารามิเตอร์ วิธีแมนเทล – แฮนส์เซล ดัชนีที่พัฒนาเพื่อเกณฑ์ในการตัดสินข้อสอบลำเอียงคือ SA , UA, α_{MH} , β_{SIB} ตามลำดับ โดยใช้แบบสอบวิชาภาษาอังกฤษคว ามยาว 50 60 70 และ 80 และแบบสอบวิชาคณิตศาสตร์ความยาว 20 30 และ 40 กับกลุ่มตัวอย่าง 6 ขนาด คือ 100 200 400 600 800 และ 1,000 คน

ผลการวิจัยพบว่าเกณฑ์ที่พัฒนาจากข้อมูลเชิงประจักษ์ เพื่อใช้เป็นเกณฑ์ในการตัดสินความลำเอียงของข้อสอบระหว่างเพศชายและหญิง คือ

1) $|SA| > .80$ และ $UA > .50$ กรณีความยาวแบบสอบต่ำกว่า 50 ข้อ และ

2) $|SA| > .40$ และ $UA > 1.20$ กรณีความยาวแบบสอบ 50 ข้อขึ้นไป

3) $\alpha_{MH} < .60$ และ $\alpha_{MH} > 1.40$ สำหรับทุกขนาดของผู้สอบและความยาวแบบสอบ

4) $\beta_{SIB} > .06$ สำหรับทุกขนาดของผู้สอบและความยาวแบบสอบ

การใช้ดัชนี SA หรือ UA ควรใช้ผู้สอบขนาด 800 คนขึ้นไป ส่วนดัชนี α_{MH} และ β_{SIB} ควรใช้ขนาดผู้สอบอย่างน้อย 600 คนขึ้นไป

จิติมา วรรณศรี (2539) ได้เปรียบเทียบประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่างวิธีแมนเทล – แฮนส์เซลกับวิธีซิบเทสท์ โดยศึกษาจากข้อมูลจำลอง ปัจจัยที่ศึกษาได้แก่ ความยาวของแบบสอบ 3 ขนาด คือ 30 60 และ 90 ข้อ ขนาดกลุ่มตัวอย่าง 3 ขนาด คือ 200 600 และ 1,000 คน โดยแต่ละขนาดมีอัตราส่วนระหว่างผู้สอบกลุ่มอ้างอิงและกลุ่มเปรียบเทียบแตกต่างกันคือ 1:1 1:0.9 1:0.75 และ 1:0.5

ผลการศึกษาพบว่าวิธีแมนเทล – แฮนส์เซลกับวิธีซิบเทสท์ มีประสิทธิภาพเท่าเทียมกันในการตรวจสอบการทำหน้าที่ต่างกันที่ทุกขนาดกลุ่มตัวอย่างและอัตราส่วนภายใต้ค วามยาวแบบสอบเดียวกันและเมื่อใช้แบบสอบที่มีความยาวปานกลาง (60 ข้อ) ทั้งสองวิธีมีประสิทธิภาพที่สุด นอกจากนี้เมื่อใช้ขนาดกลุ่มตัวอย่างมากขึ้นจะสามารถตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันได้ถูกต้องมากขึ้น โดยส่วนมากวิธีซิบเทสท์มีอัตราความคลาดเคลื่อนประเภทที่ 1 มากกว่าวิธีแมนเทล – แฮนส์เซล เล็กน้อย

เกษร หว่างจิตร (2539) ได้ทำการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบสำหรับแบบสอบคัดเลือกระดับบัณฑิตศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ด้วยวิธีแมนเทล – แฮนส์เซล กลุ่มตัวอย่างที่ใช้คือผู้สมัครสอบคัดเลือกเข้าศึกษาระดับบั ณฑิตศึกษา โดยกลุ่มผู้สอบจำแนกตามเพศ ภูมิลำเนา ประสบการณ์ในการสอบและสังกัดสถานศึกษา เครื่องมือที่ใช้เป็นแบบสอบคัดเลือกวิชาภาษาอังกฤษ จำนวน 80 ข้อและแบบสอบวิชาภาษาไทย จำนวน 95 ข้อ

ผลการวิจัยพบว่า ข้อสอบที่พบการทำหน้าที่ต่างกันส่วนใหญ่เป็นแบบอเนกรูป ข้อสอบที่มีค่าอำนาจจำแนกค่อนข้างต่ำ (0.20 – 0.39) สอดคล้องกันทั้งวิชาภาษาไทยและภาษาอังกฤษ สำหรับข้อสอบที่ทำหน้าที่ต่างกันในวิชาภาษาไทยส่วนใหญ่เป็นข้อสอบที่ง่ายมาก (0.80 – 1.00) แต่ในวิชาภาษาอังกฤษส่วนใหญ่เป็นข้อสอบที่ยาก (0.00 – 0.19) ลักษณะเนื้อหาของข้อสอบที่ทำหน้าที่ต่างกันในวิชาภาษาไทยเป็นเนื้อหาด้านการใช้ภาษาไทย 1 และวิชาภาษาอังกฤษเป็นข้อสอบที่มีลักษณะการให้รายละเอียดมากที่สุด ส่วนค่าความเที่ยงและความตรงตามทฤษฎีของ

แบบสอบกรณีก่อนและหลังตัดข้อสอบที่ทำหน้าที่ต่างกันออกจากแบบสอบ ไม่แตกต่างกันอย่างมีนัยสำคัญทั้งสองวิชา

ญาณภัทร สีหะมงคล (2540) ได้ทำการเปรียบเทียบความสอดคล้องขอผลการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกัน ระหว่างวิธี Lord's X^2 และวิธี Raju's Area Measurement และวิธี Closed Interval Area โดยใช้นักเรียนประถมศึกษาปีที่ 4 จำนวน 11,404 คนทำแบบสอบคณิตศาสตร์ 80 ข้อ ผลการวิจัยพบว่า จำนวนข้อสอบที่ทำหน้าที่ต่างกันที่ได้จากการตรวจสอบระหว่างวิธี Lord's X^2 วิธี Raju's Area Measurement และวิธี Closed Interval Area แตกต่างกันตามขนาดกลุ่มตัวอย่างและความยาวของแบบทดสอบแตกต่างกัน โดยมีความสัมพันธ์ของผลการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันทั้ง 3 วิธีมีในระดับสูง และมีนัยสำคัญทางสถิติเกือบทุกเงื่อนไข

เสรี ชัดแช้ม (2540) ศึกษาเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอเนกรูประหว่างวิธีแมนเทล -แฮนส์เซลแบบปกติกับวิธีแมนเทล -แฮนส์เซลแบบแบ่งกลุ่มควมสามารถของผู้สอบและความยากของข้อสอบโดยใช้วิธี IRT เป็นเกณฑ์โดยศึกษาจากข้อมูลผลการตอบแบบสอบวัดความสามารถในการอ่านภาษาไทยของนักเรียนชั้นมัธยมศึกษาปีที่ 1 สังกัดกรมสามัญศึกษาจังหวัดชลบุรี จำนวน 1,200 คน โดยกลุ่มผู้สอนจำแนกตามเพศ

ผลการศึกษาพบว่า วิธี MH แบบแบ่งกลุ่มความสามารถของผู้สอบและ ความยากของข้อสอบ สามารถตรวจพบข้อสอบทำหน้าที่ต่างกันแบบอเนกรูปได้สอดคล้องกับวิธี IRT และตรวจพบข้อสบทำหน้าที่ต่างกันมากกว่าวิธี MH แบบปกติ ข้อสอบที่ตรวจพบส่วนใหญ่เป็นข้อสอบยากปานกลางและข้อสอบง่าย ซึ่งมีโค้งลักษณะข้อสอบของกลุ่มผู้สอบสองกลุ่มตัดกันบริเวณใกล้ ๆ จุดกลางของช่วงความสามารถ

นพมาศ พิพัฒนสุข (2541) ได้ศึกษาการเปรียบเทียบประสิทธิภาพระหว่างวิธีแมนเทล -แฮนส์เซลกับวิธีถดถอยโลจิสติกในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เมื่อใช้เกณฑ์จับคู่เปรียบเทียบแตกต่างกันในแบบสอบพหุมิติ กลุ่มตัวอย่างเป็นนักเรียนชั้น ประถมศึกษาปีที่ 6 จำนวน 1,076 คน ผลการวิจัยพบว่า วิธีแมนเทล -แฮนส์เซลเมื่อใช้คะแนนรวมเป็นเกณฑ์จับคู่เปรียบเทียบตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 15 ข้อ และเมื่อใช้คะแนนแบบสอบย่อยเป็นเกณฑ์การจับคู่ตรวจข้อสอบทำหน้าที่ต่างกันจำนวน 14 ข้อ สำหรับวิธีถดถอยโลจิสติ กเมื่อใช้

คะแนนรวมเป็นเกณฑ์จับคู่เปรียบเทียบตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 17 ข้อ และเมื่อใช้คะแนนแบบสอบย่อยเป็นเกณฑ์การจับคู่ตรวจข้อสอบทำหน้าที่ต่างกันจำนวน 13 ข้อ โดยที่วิแมนเทล-แฮนส์เซลมีประสิทธิภาพดีกว่าวิธีถดถอยโลจิสติกในการตรวจสอบการทำหน้าที่ต่าง กันของข้อสอบในแบบสอบชนิดพหุมิติเมื่อใช้เกณฑ์จับคู่แบบคะแนนรวม และมีประสิทธิภาพไม่แตกต่างกันเมื่อใช้เกณฑ์จับคู่คะแนนสอบย่อย แต่อย่างไรก็ตามวิธีถดถอยโลจิสติกเมื่อใช้เกณฑ์จับคู่เปรียบเทียบคะแนนหลายแบบสอบย่อยมีความเหมาะสมในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบชนิดพหุมิติ

วลีมาศ แซ่อึ้ง (2543) ทำการเปรียบเทียบอำนาจการทดสอบและอัตราความคลาดเคลื่อนประเภทที่ 1 ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอเนกรูประหว่างวิธีซิปเทสท์ปรับใหม่ วิธีซิปเทสท์ วิธีแมนเทล -แฮนส์เซลและวิธีถดถอยโลจิสติก โดยมีเงื่อนไขที่ทำการศึกษา 324 เงื่อนไข ผลการวิจัยพบว่า อำนาจการทดสอบในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอเนกรูปของวิธีซิปเทสท์ปรับใหม่ และวิธีถดถอยโลจิสติกมีค่าเท่าเทียมกันในทุกเงื่อนไข และทั้งสองวิธีมีอำนาจการทดสอบสูงกว่าวิธีซิปเทสท์และวิธีแมนเทล -แฮนส์เซลภายใต้เกือบทุกเงื่อนไข ส่วนอัตราความคลาดเคลื่อนประเภทที่ 1 ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบอเนกรูปของวิธีซิปเทสท์ปรับใหม่ วิธีซิปเทสท์ วิธีแมนเทล -แฮนส์เซล และวิธีถดถอยโลจิสติกมีค่าอยู่ในเกณฑ์ของอัตราความคลาดเคลื่อนประเภทที่ 1 ที่ระดับ 10% ทุกเงื่อนไข

อารี วัชรโสตถิกุล (2543) ได้เปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ศึกษาเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้รูปแบบต่างกันคือ รูปแบบคะแนนรวมทั้งฉบับแยกตามเนื้อหา และแยกตามระดับพฤติกรรมด้วยวิธีการตรวจสอบต่างกัน คือ วิธี SIBTEST และวิธีการถดถอยโลจิสติก แล้วทำการคัดเลือกข้อสอบที่ทำหน้าที่ต่างกันออกจากแบบทดสอบ เพื่อเปรียบเทียบค่าความเชื่อมั่น ผลการศึกษาพบว่า จำนวนข้อสอบที่ทำหน้าที่ต่างกันโดยใช้วิธีการตรวจสอบต่างกัน แตกต่างกันในรูปแบบรวมทั้งฉบับ ส่วนรูปแบบแยกตามเนื้อหาและแยกตามระดับพฤติกรรมไม่แตกต่างกัน

รักชนก ยี่สุ่นศรี (2544) ทำการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและแบบสอบสำหรับกลุ่มผู้สอบเมื่อจำแนกตามเพศและสถานที่ตั้งทางภูมิศาสตร์และโรงเรียนที่จบการศึกษา เพื่อเปรียบเทียบความเที่ยง ความตรง และฟังก์ชันสารสนเทศของแบบสอบระหว่า งแบบสอบฉบับก่อนและหลังตัดข้อสอบที่พบ DIF โดยใช้ข้อมูลจากการตอบแบบสอบคัดเลือกบุคคลเข้าศึกษาในสถาบันอุดมศึกษา วิชาภาษาอังกฤษและคณิตศาสตร์ ปีการศึกษา 2543 ครั้งที่ 1 และเลือกศึกษาในส่วนที่เป็นข้อสอบแบบหลายตัวเลือกจากกลุ่มตัวอย่างที่เป็นผู้เข้าสอบจำนวน 4,000 คน และ 3,600 คน ตามลำดับ

ผลวิจัยพบว่า แบบสอบวิชาภา ษาอังกฤษทำหน้าที่ต่างกันตามเพศและสถานที่ตั้งตามภูมิศาสตร์ ส่วนแบบสอบวิชาคณิตศาสตร์ทำหน้าที่ต่างกันตามเพศของผู้สอบ และข้อสอบทำหน้าที่ต่างกันตามเพศของผู้สอบมากที่สุดทั้งสองวิชา นอกจากนี้เมื่อเปรียบเทียบคุณภาพข องแบบสอบฉบับก่อนและหลังการตัดข้อสอบที่พบว่าทำหน้าที่ต่างกัน พบว่า ในด้านความตรงจะไม่ต่างกัน แต่แบบสอบฉบับหลังจากตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันแล้ว ส่วนใหญ่จะมีค่าความเที่ยงลดลง และมีค่าฟังก์ชันสารสนเทศมากขึ้น ส่วนความสัมพันธ์ระหว่างตำแหน่งของคะแนนรวมของผู้สอบก่อนและหลังจากการตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันทั้งในกรณีที่ตัดทุกข้อและตัดในบางข้อ พบว่าทุกกรณีมีความสัมพันธ์ในทางบวกซึ่งกันและกันอย่างมีนัยสำคัญ

นิคม กีรติวรางกูร (2545) ได้ทำการเปรียบเทียบประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระห ว่างวิธีการวิเคราะห์องค์ประกอบจำกัด แมนเทล -แฮนส์เซล และการตอบสนองข้อสอบ โดยเทียบกับเกณฑ์ที่กำหนดคือความยาวของแบบสอบ ขนาดกลุ่มตัวอย่างค่าความยากของข้อสอบและค่าอำนาจจำแนกของข้อสอบ ผลการศึกษาพบว่า โดยภาพรวมวิธี RFA มีประสิทธิภาพในการการตรวจสอบการทำหน้าที่ต่าง กันของข้อสอบได้สูงที่สุด รองลงมาคือวิธีแมนเทล-แฮนส์เซล และวิธี IRT-2PL โดยวิธี IRT-2PL มีความคลาดเคลื่อนประเภทที่ 1 สูงที่สุด สูงกว่าวิธีแมนเทล -แฮนส์เซล และวิธี RFA ตามลำดับ และยังพบว่าวิธีแมนเทล -แฮนส์เซลมีประสิทธิภาพในการการตรวจสอบการทำหน้าที่ต่างกันของข้ อสอบสูงภายใต้เงื่อนไขแบบสอบที่มีความยากต่ำ อำนาจจำแนกสูง ที่ความยาวแบบสอบ 75 ข้อ และขนาดกลุ่มตัวอย่าง 1,000 คน

ส่วนวิธี IRT-2PL มีประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบสูงภายใต้เงื่อนไขแบบสอบที่มีความยากต่ำ ที่ความยาวแบบสอบ 75 ข้อ และขนาดกลุ่มตัวอย่าง 1,000 คน

ชุติมา แสงดารารัตน์ (2545) ได้เปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนตามความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจในข้อสอบด้วยวิธีการตรวจแตกต่างกัน มีกลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนในกรุงเทพมหานคร จำนวน 585 คน ผลการศึกษาที่สำคัญพบว่า จำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลด้วยวิธีการตรวจแตกต่างกัน สามารถบ่งชี้ข้อที่ทำหน้าที่ต่างกัน จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกสนใจ ได้จำนวนข้อแตกต่ างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนกลุ่มอ้างอิงมีความรู้สึกคุ้นเคย ความรู้สึกพอใจ ได้จำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

สุมาลี แกล้วทนงค์ (2547) ได้ศึกษาสาเหตุของการทำหน้าที่ต่างกันของข้อสอบสาระการเรียนรู้ภาษาไทย และสาระการเรียนรู้สั งคมศึกษา ศาสนาและวัฒนธรรม กับนักเรียนชั้นมัธยมศึกษาปีที่ 1 จำนวน 1,320 คน ผลการศึกษาพบว่า ข้อสอบที่ทำหน้าที่ต่างกันด้านเพศของแบบสอบสาระการเรียนรู้ภาษาไทยส่วนใหญ่มีสาเหตุมาจากความสนใจในเนื้อเรื่องและภาษาที่ใช้ในแบบสอบ ส่วนข้อสอบที่ทำห น้าที่ต่างกันด้านเพศของ แบบสอบสา มารถเรียนรู้สังคมศึกษาศาสนาและวัฒนธรรมนั้นมีสาเหตุมาจากเนื้อเรื่องที่สนใจและเนื้อเรื่องที่เกี่ยวกับวัฒนธรรมประเพณี

อิทธิฤทธิ์ พงษ์ปิยรัตน์ (2551) ได้ทำการวิเคราะห์ข้อสอบและตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และเป็นการเปรียบเทียบผลการวิเคราะห์กา รทำหน้าที่ต่างกันของข้อสอบโดยการประยุกต์ใช้โปรแกรมโมเดลเชิงเส้นตรงระดับลดหลั่น (HLM) และ BILOG-MG โดยใช้เครื่องมือความรู้วิชาคณิตศาสตร์ เจตคติของนักเรียนต่อวิชาคณิตศาสตร์ ภาวะผู้นำทางวิชาการของผู้บริหาร กลุ่มตัวอย่างคือนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2550 จำนวน 1,588 คน ผู้บริหาร 32 คน ผลการศึกษาพบว่า ผลการวิเคราะห์ระดับนักเรียน (Level 2) ตัวแปรผลการเรียนวิชาคณิตศาสตร์ในภาคเรียนที่ผ่านมาส่งผลต่อค่าเฉลี่ยของโอกาสในการตอบข้อสอบได้ถูกต้องอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ผลการวิเคราะห์ระดับโรงเรียน (Level 3) พบว่าตัวแปร

ขนาดของโรงเรียน และตัวแปรความเป็นผู้นำทางด้านวิชาการของผู้บริหารส่งผลต่อค่าเฉลี่ยของโอกาสในการตอบข้อสอบได้ถูกต้องอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และ .05 ตามลำดับ โดยโมเดล HGLM สามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบสอดคล้องตรงกับ ผลการวิเคราะห์จากโปรแกรม BILOG-MG

Shealy และ Stout (1993) ได้เสนอวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบและแบบสอบคือวิธีซิบเทสส์ (SIBTEST) ซึ่งเป็นวิธีการที่อยู่บนพื้นฐานของทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ เป็นวิธีการที่สามารถตรวจสอบการทำหน้าที่ต่างกันข องข้อสอบได้ทั้งกรณีที่แบบสอบมีความลำเอียงข้อเดียวและกรณีที่มีความลำเอียงหลายข้อ ได้ทำการศึกษาประสิทธิภาพในการตรวจสอบโดยการเปรียบเทียบกับวิธีแมนเทล – เฮนส์เซล ในกรณีที่มีข้อสอบลำเอียงข้อเดียว โดยใช้ข้อมูลผลการตอบจากแบบสอบวิชาคณิตศาสตร์ของ ACT และแบบสอบของ ASVAB ขนาดกลุ่มตัวอย่างของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบต่างกัน คือ (3000, 3000), (3000, 1000), (1500, 1500), (1000, 500), (500, 500) และ (500, 250)

ผลการศึกษาพบว่าในกรณีที่มีข้อสอบลำเอียงข้อเดียวทั้งวิธีแมนเทล -เฮนส์เซลและวิธีซิลเทสส์มีอำนาจในการตรวจสอบดีทียบเท่ากันได้และวิธีซิบเทสส์มีประสิทธิภาพดีในการตรวจค้นกรณีที่มีข้อสอบที่ลำเอียงหลาย ๆ ข้อ แม้ว่ามีขนาดความลำเอียงค่อนข้างน้อยและผลการศึกษพบว่าทั้งวิธีซิบเทสส์และแมนเทล– เฮนส์เซลจะมีประสิทธิภาพดีเมื่อใช้กับแบบสอบที่ยาวพอควร($\geq$ 25)

Cohen และ Kim (1993) ศึกษาเปรียบเทียบการใช้ทฤษฎีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ โดยใช้สถิติทดสอบไค – สแควร์ ของลอร์ดและสถิติทดสอบพื้นที่ความแตกต่างระหว่างโค้งลักษณะข้อสอบชนิดคิดเครื่องหมายและไม่คิดเครื่องหมาย อันได้แก่ สถิติ Z ของ Raju โดยใช้ข้อมูลจำ ลองตามทฤษฎีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ จำลองการตอบข้อสอบจากแบบสอบ 2 ขนาด คือ 20 ข้อ และ 60 ข้อ จำนวนผู้สอบ 100 และ 500 คน และมีจำนวนข้อสอบที่ลำเอียงในแบบสอบจำนวน ร้อยละ 0 10 และ 20

ผลการวิจัยพบว่า แบบสอบยิ่งมีความยาว จะทำให้การระบุข้อสอบลำเอียงผิดพล าด (EP) และข้อสอบที่ไม่ถูกระบุว่าลำเอียงผิดพลาด (FN) จะมากขึ้นด้วย และถ้ามีจำนวนข้อสอบที่ลำเอียงอยู่ในแบบสอบน้อย จำนวนข้อสอบที่ถูกระบุผิดพ ลาดก็จะยิ่งมากขึ้น องค์ประกอบที่มี

อิทธิพลต่อการระบุข้อสอบลำเอียงผิดพลาด (FP) อีกประการคือ ระดับความมีนัยสำคัญ (α) ค่า α ยิ่งมาก FP จะเกิดมาก กรณีที่จำนวนผู้สอบ 100 คน พบว่าความแตกต่างในการระบุว่าข้อสอบไม่ลำเอียงผิดพลาด สำหรับ Z และ ไคว์สแควร์ ไม่แตกต่างกันมากเท่ากับกรณีกลุ่มผู้สอบจำนวน 500 คน ซึ่งมีการระบุว่าไม่ลำเอียงผิดพลาด สำ หรับไคว์สแควร์ น้อยกว่า Z โดยเฉพาะอย่างยิ่งในกรณีที่มีข้อสอบลำเอียงในแบบสอบมากที่สุดคือ ร้อยละ 20

Narayanan และ Swaminathan (1994) ได้ศึกษาผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทล -เฮนส์เซลและวิธีซิบเทสส์ ใช้ข้อมูลจำลองความยาวแบบสอบ 40 ข้อ ตัวแปรที่ศึกษาคือ (1) ขนาดกลุ่มตัวอย่าง โดยกลุ่มอ้างอิงมี 3 ขนาดได้แก่ 300 500 และ 1,000 คน (2) การกระจายของความสามารถ 2 แบบ (3) อัตราส่วนของข้อสอบที่ทำหน้าที่ต่างกันที่มีภายในแบบสอบ 2 ขนาด (4) ขนาดพื้นที่ระหว่างโค้งคุณลักษณะข้อสอบของผู้สอบสองกลุ่ม 4 ขนาด (5) ค่าความยากและค่าอำนาจจำแนกของแบบสอบ 6 ระดับ

ผลการศึกษาพบว่าขนาดของกลุ่มตัวอย่าง อัตราส่วนของข้อสอบที่ทำหน้าที่ต่างกันของพื้นที่ระหว่างโค้งคุณลักษณะ ค่าความยาก และค่าอำนาจจำแนกเป็นตัวแปรที่มีผลกระทบต่ออัตราการตรวจสอบของทั้งสองวิธีอย่างมีนัยสำคัญ วิธีแมนเ ทล - เฮนส์เซลและวิธีซิบเทสส์มีประสิทธิภาพเท่ากันในการตรวจสอบการทำหน้าที่ต่างกันแบบเอกรูปเมื่อการกระจายความสามารถเท่ากันระหว่างกลุ่ม แต่เมื่อการกระจายความสามารถไม่เท่ากันระหว่างกลุ่ม วิธีซิบเทสส์จะมีประสิทธิภาพมากกว่า จึงสรุปได้ว่าการกระจายความสามารถไม่มีผลกระทบต่อการตรวจสอบด้วยวิธีซิบเทสส์แต่มีผลกระทบต่อวิธีแมนเทล – เฮนส์เซลอย่างมีนัยสำคัญ ในกรณีการกระจายความสามารถต่างกันเพิ่มขึ้นระหว่างผู้สอบทั้งสองกลุ่มจะทำให้อัตราความคลาดเคลื่อนชนิดที่ 1 เพิ่มมากขึ้น

Welch และ Miller (1995) ได้ทำการศึกษาปัญหาของการตรวจสอ บการทำหน้าที่ต่างกันของแบบสอบประเมินการเขียน โดยมีจุดมุ่งหมายเพื่อศึกษาว่าปัญหาใดบ้างที่เป็นอุปสรรคต่อการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ประเมินผลการปฏิบัติจริงและเพื่อค้นหาเกณฑ์การจับคู่ที่เหมาะสม โดยใช้สถิติGMH (Generalized Mantel-Haenszel) และ การวิเคราะห์ฟังก์ชันโลจิสติก เครื่องมือที่ใช้ศึกษาเป็นแบบประเมินทักษะการเขียนโดยตรงและแบบสอบวัดทักษะการ

เขียนแบบเลือกตอบ จำนวน 40 ข้อ กลุ่มตัวอย่างเป็นนักเรียนเกรด 8 ข้อมูลที่นำมาใช้ได้จากแบบประเมินทักษะการเขียนโดยตรง มี 3 ฟอร์ม คือ 01A, 01B และ 01C ตัวแปรที่ใช้แบ่งกลุ่มผู้สอบคือ ตัวแปรเพศและตัวแปรสีผิว ผลการศึกษาโดยภาพรวมแล้วพบว่าเกณฑ์การจับคู่โดยใช้เกณฑ์ภายนอกไม่เหมาะสมและยังบ่งชี้อีกว่าเกณฑ์การจับคู่ที่ไม่เหมาะสมจะเป็นปัญหาในการตรวจสอบการทำหน้าที่ต่างกัน

Raju และคณะ (1995) ได้ศึกษาประสิทธิภาพของวิธีการตรวจสอ บการทำหน้าที่ต่างกันของข้อสอบและแบบสอบ ด้วยวิธี DFIT ซึ่งประกอบด้วยดัชนี CDIF และ NCDIF สำหรับตรวจสอบการทำหน้าที่ต่างกันที่ระดับข้อสอบ และดัชนี DTF สำหรับตรวจสอบการทำหน้าที่ ต่างกันที่ระดับแบบสอบ โดยจะทำการศึกษาเปรียบเทียบกับวิธีวิเคราะห์อีก3วิธีคือวิธีทดสอบไคสแควร์ของลอร์ด วิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบแบบคิดเครื่องหมาย (ESA) และแบบไม่คิดเครื่องหมาย (EUA) ใช้ข้อมูลจำลองจากโมเดลโลจิสติกแบบ 2 พารามิเตอร์ (2PLM) เปรียบเทียบประสิทธิภาพวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เมื่อใช้ขนาดกลุ่มตัวอย่าง (500, 1000) อัตราส่วนของข้อที่ทำหน้าที่ต่างกันในแบบสอบ (0%, 5%, 10%, 20%) และรูปแบบการทำหน้าที่ต่างกันของข้อสอบ (uniform, non-uniform)

ผลการวิจัยพบว่าในภาพรวมดัชนี LC ของวิธีทดสอบไค – สแควร์ของลอร์ด ดัชนี EUA ของวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบแบบไม่คิดเครื่องหม าย มีประสิทธิภาพเท่าเทียมกันในการตรวจสอบที่ระดับข้อสอบ โดยสามารถระบุข้อที่ทำหน้าที่ต่างกันได้ถูกต้องสูงในทุกเงื่อนไข ในขณะที่ดัชนี CDIF และดัชนี ESA มีประสิทธิภาพต่ำ เมื่อกลุ่มตัวอย่างมีขนาดเล็ก(N=500) และอัตราส่วนของข้อที่ทำหน้าที่ต่างกันสูง (20%) ดัชนี NCDIF จะมีประสิทธิภาพสูงที่สุดแต่ดัชนี CDIF มีประสิทธิภาพต่ำที่สุด เมื่อขนาดกลุ่มตัวอย่างมีขนาดใหญ่ (N=1000) และอัตราส่วนการทำหน้าที่ต่างกันสูง (20%) ดัชนี ESA มีประสิทธิภาพต่ำที่สุด

Giray (1995) ได้ทำการศึกษาวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ สำหรับแบบสอบคัดเลือกบุคคลเข้าศึกษาในมหาวิทยาลัย ซึ่งเป็นข้อสอบด้านการคำนวณ คำศัพท์ และเรขาคณิต โดยแบ่งกลุ่มผู้สอบตามตัวแปรที่ศึกษาคือ เพศและเศรษฐานะ กลุ่มตัวอย่างเป็นนักเรียนของโรงเรียนมัธยมศึกษาใน Ankara ประเทศตุรกี จากผลการวิจัยพบว่า ข้อสอบที่ทำ

หน้าที่ต่างกันเป็นข้อสอบในวิชาคณิตศาสตร์ ทั้งในกรณีแบ่งกลุ่มผู้สอบตามตัวแปรเพศและเศรษฐานะ สำหรับตัวแปรเพศ ผลการวิเคราะห์พบว่าผู้ชายจะมีข้อได้เปรียบในการทำข้อสอบด้านการคำนวณ ในขณะที่เพศหญิงจะมีข้อได้เปรียบในการทำสอบด้านคำศัพท์และเรขาคณิต ซึ่งชี้ให้เห็นว่าผู้หญิงจะมีความสามาร ถในด้านการใช้ภาษาหรือเกี่ยวกับการใช้คำศัพท์ดีกว่าและเพศชายจะมีทักษะด้านการคำนวณดีกว่าเพศหญิง

Chang และคณะ (1996) ได้ทำการศึกษาผลของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่มีการให้คะแนนแบบหลายค่า โดยประยุกต์วิธีซิบเทสท์มาใช้ เปรียบเทียบกับวิธีแมนเทลและวิธี SMD แบ่งการศึกษาออกเป็น 2 ตอน ตอนที่ 1 ใช้ข้อมูลจำลองภายใต้เงื่อนไขเดียวกับงานวิจัยของ Zwick และคณะ (1993) เพื่อทำการเปรียบเทียบวิธีซิบเทสท์แบบประยุกต์กับวิธีแมนเทลและวิธี SMD ผลที่ได้ก็คือวิธีซิบเทสท์มีประสิทธิภาพในการตรวจสอบค่อนข้างดีแต่วิธีแมนเทลและวิธี SMD มีประสิทธิภาพค่อนข้างดีกว่า ในการศึกษาตอนที่ 2 ใช้ข้อมูลจำลองคือ ข้อสอบที่ศึกษามีอำนาจการจำแนกแตกต่างกัน 11 ค่า ตั้งแต่ .15 ถึง 2.00 ขนาดกลุ่มตัวอย่างต่างกัน คือ 500 1,000 และ 2,000 คน ข้อสอบมีความยาว 24 ข้อสำหรับวิธีซิบเทสท์ และ 25 ข้อสำหรับวิธีแมนเทลและวิธี SMD ผลการศึกษาสรุปว่าทั้งวิธีแมนเทลและวิธี SMD มีค่าความคลาดเคลื่อนชนิดที่ 1 ค่อนข้างสูงเมื่อพารามิเตอร์ของค่าอำนาจจำแนกของข้อสอบที่ศึกษามีค่าต่างจากค่าเฉลี่ยของอำนาจจำแนกของแบบสอบที่มีความตรง และอัตราการปฏิเสธ (rejection rates) ของทั้งสามวิธีการจะมีค่าสูงขึ้นเมื่อค่าอำนาจจำแนกสูงขึ้น

Katherine และ Meichu (1996) ได้ทำการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบสำหรับแบบสอบวิชาคณิตศาสตร์ ซึ่งเป็นแบบสอบแบบหลายตัวเลือก โดยแบ่งกลุ่มผู้สอบตามตัวแปรเพศ กลุ่มตัวอย่างเป็นนักเรียนชั้นปีที่ 3 ของโรงเรียนมัธยม ศึกษา ข้อมูลที่ใช้ในการวิจัยรวบรวมมาจาก The Second International Mathematics Study (1985) ในการวิจัยครั้งนี้ได้ศึกษาความสัมพันธ์ระหว่างดัชนีที่ใช้วิเคราะห์การทำหน้าที่ต่างกันและลักษณะเฉพาะของแบบสอบแบบหลายตัวเลือก โดยมีสมมติฐานว่าข้อสอบด้านพีชคณิตและการคำนว ณจะเข้าข้างผู้สอบเพศหญิง ส่วนข้อสอบด้านเรขาคณิตและเลขคณิตจะเข้าข้างเพศชาย ผลการวิจัยพบว่าข้อสอบด้านพีชคณิตและการคำนวณจะง่ายสำหรับผู้สอบเพศหญิงและข้อสอบด้านเรขาคณิตจะเข้าข้าง

เพศชาย ในขณะที่ข้อสอบด้านเลขคณิตจะเข้าข้างเพศหญิง ซึ่งเป็นสมมติฐานเดียวที่ถูกปฏิเสธ จากผลการวิจัยในภาพรวม พบว่าข้อสอบด้านการประยุกต์จะทำหน้าที่ต่างกันโดยจะเป็นข้อสอบที่ค่อนข้างยากสำหรับเพศหญิง

Ning และ Suzanne (1996) ได้ทำการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ในแบบสอบการประเมินสมรรถภาพในวิชาคณิตศาสตร์ การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ โดยแบ่งกลุ่มผู้สอบตามเพศ และระบุองค์ประกอบที่อาจเกี่ยวข้องสัมพันธ์กับการทำหน้าที่ต่างกัน เช่น เนื้อหา กระบวนการรับรู้ และความแตกต่างในการกระจายของค่าความสามารถ เป็นต้น โดยใช้แบบสอบ QUASAR (Quantitative Understanding : Amplifying Student Achievement and Reasoning) Cognitive Assessment Instrument (QCAI) ในการทดสอบทักษะการคิดและการใช้เหตุผลของนักเรียน ซึ่งประกอบด้วยข้อสอบแบบปลายเปิด โดยผู้สอบต้องแสดงวิธีทำและอธิบายคำตอบของเขาด้วย แบบสอบมี 4 ฟอร์ม 33 ข้อ ข้อสอบจะเป็นแบบให้ค ะแนนแบบหลายค่า (polytomously scored items) กลุ่มตัวอย่างเป็นนักเรียนเกรด 6 และเกรด 7 จาก 6 โรงเรียนที่ร่วมโครงการ QUASAR โดยใช้กลุ่มตัวอย่างจำนวน 1,782 คน ผลการวิจัยพบว่าข้อสอบจำนวน 31 ข้อ จากข้อสอบทั้งหมด 33 ข้อ คิดเป็นร้อยละ 94 ไม่ทำหน้าที่ต่างกันหรือไม่ทำหน้าที่ต่างกันในระดับที่มากนัก ส่วนข้อสอบอีก 2 ข้อที่เหลือ ทำหน้าที่ต่างกัน เมื่อแบ่งกลุ่มผู้สอบตามเพศ

Zwick และคณะ (1997) ได้ทำการศึกษาผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่เป็นชิ้นงาน ด้วยวิธีแมนเทล (Mantel approach) และวิธี GMH (Generalized of the Mantel-Haenszel approach) ใช้ข้อมูลจำลองความยาวของแบบสอบ 25 ข้อ 24 ข้อแรกใช้เป็นตัวแปรที่ใช้จับคู่และข้อที่ 25 เป็นข้อสอบที่ศึกษา กลุ่มเปรียบเทียบมีการกระจาย2 ระดับ คือ N(0,1) และ N(-1,1) และลักษณะเฉพาะของข้อสอบที่ศึกษามี 27 ระดับ แตกต่างกันตามค่าความยา ก รูปแบบของ DIF มี 4 รูปแบบคือ แบบคงที่ (constant) แบบสมดุล (balance) แบบมีผลเฉพาะคะแนนต่ำ (shift low) และแบบมีผลเฉพาะคะแนนสูง (shift high) และขนาดของ DIF มี 2 ขนาด คือ .10 และ .25 ข้อสอบ ข้อที่ 1-20 เป็นข้อสอบที่ให้คะแนนแบบสองค่าและข้อที่ 21 – 24 เป็นข้อสอบที่ให้คะแนนแบบ4 ค่า ซึ่งข้อที่ 24 เป็นข้อสอบที่ทำหน้าที่ต่างกัน

ผลการศึกษาพบว่าทั้งสองวิธีมีประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในการวัดความสามารถในการปฏิบัติงานในกรณีที่มีการให้คะแนนแบบหลายค่า วิธี GMH จะมีประสิทธิภาพมากกว่าวิธีแมนเทล เมื่ อการกระจายของข้อมูลในกลุ่มผู้สอบแตกต่างกัน ผลการศึกษาแสดงให้เห็นว่าควรรวมข้อสอบที่ศึกษาเข้าไปในตัวแปรที่ใช้ในการจับคู่ด้วยเหมือนในกรณีของการให้คะแนนแบบสองค่า ในเงื่อนไขของ DIF แบบคงที่ วิธีแมนเทลจะมีประสิทธิภาพมากกว่าวิธี GMH ส่วนในเงื่อนไข DIF แบบสมดุล วิธี GMH จะมีความสามารถเหนือกว่าวิธีแมนเทลอย่างเห็นได้ชัดที่ขนาดของ DIF เท่ากับ .25 ส่วนในกรณี DIF แบบมีผลเฉพาะที่ระดับคะแนนต่ำและระดับสูง ทั้งสองวิธีมีประสิทธิภาพการตรวจสอบใกล้เคียงกัน

Oshima และคณะ (1997) ได้ทำการศึกษาวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบและแบบสอบ (DFIT Framework) สำหรับแบบสอบชนิดพหุมิติ ซึ่งวิธีการนี้ได้ขยายมาจากวิธี DFIT แบบมิติเดียวที่ Raju และคณะเสนอไว้ในปี 1995 ในการศึกษาใช้ข้อมูลจำลองจากโมเดลโลจิสติกแบบ 2 พารามิเตอร์ (M2PL) จำนวน 40 ข้อ และข้อมูลเป็นแบบ 2 มิติ องค์ประกอบที่สนใจศึ กษาคือ รูปแบบการทำหน้าที่ต่างกัน (เอกรูปและแบบอเนกรูป) ทิศทางการทำหน้าที่ต่างกัน (แบบทิศทางเดียว , แบบสองทิศทางที่สมดุลกัน) และการกระจายของค่าความสามารถแตกต่างกันระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ (N(0, 1), N(-1, 1))

ผลการศึกษาพบว่าในทุกเงื่อนไขข้อสอบที่ถูกจำลองไม่ทำหน้าที่ต่างกัน ดัชนี CDIF และดัชนี NCDIF ที่คำนวณได้มีค่าเป็น 0.00 สำหรับเงื่อนไขที่ 1 (การทำหน้าที่ต่างกันแบบเอกรูปและมีทิศทางเดียว) ข้อสอบ 4 ข้อที่จำลองให้ทำหน้าที่ต่างกัน ดัชนี CDIF ของทั้ง 4 ข้อมีค่าประมาณ 0.40 และดัชนี DTF มีนัยสำคัญเมื่อทดสอบด้วยสถิติไค -สแควร์ และค่าดัชนี NCDIF มีค่าเป็น 0.10 ซึ่งระบุว่าทำหน้าที่ต่างกัน ในเงื่อนไขที่ 2 (เอกรูปและสองทิศทางแบบสมดุล) ค่าดัชนี CDIF ทั้ง 4 ข้อมีค่าเป็น 0.00 และ DTF มีค่าเป็น 0.00 แต่ค่าดัชนี NCDIF เท่ากับ 0.10 ซึ่งมีนัยสำคัญทางสถิติอีกเช่นกัน ส่วนในเงื่อนไขเกี่ยวกับการทำหน้าที่ต่างกันแบบอเนกรูป พบว่าเมื่อค่าความยากของทั้งสองมิติแตกต่างกันจะทำให้ค่า CDIF และ NCDIF มีค่ามากขึ้น ถ้าค่าความยากของทั้งสองมิติแตกต่างกันในทิศทางตรงข้าม ดัชนี CDIF จะมีค่าเท่ากันแต่ NCDIF จะมีค่ามากกว่า เมื่อค่าความยากทั้ง 2 มิติต่างกันแต่เป็นไปในทิศทางเดียวกัน

Flowers และคณะ (1997) ได้ศึกษาความสัมพันธ์ระหว่างวิธี DFIT แบบพหุวิภาคและวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบและแบบสอบวิธีอื่น ๆ โดยทำการเปรียบเทียบประสิทธิภาพในการตรวจสอบระหว่างวิธี DFIT วิธีซิบเทสส์ และวิธีทดสอบค่าไค-สแควร์ของลอร์ด และใช้ข้อมูลจำลอง ศึกษาผลของขนาดกลุ่มตัวอย่างต่างกัน (500 และ 1000) ค่าการกระจายของค่าความสามารถของกลุ่มเปรียบเทียบ (N(0, 1), N(-1, 1)) จำนวนข้อสอบที่ทำหน้าที่ต่างกัน (0%, 10%, 20%) ผลการวิจัยพบว่าในภาพรวมวิธี DFIT มีประสิทธิภาพในการตรวจสอบดี จำนวนของข้อที่ทำหน้าที่ต่างกัน ขนาดของการทำหน้าที่ต่างกันและค่าความยากมีผลต่ออัตราความคลาดเคลื่อนชนิดที่ 1

Oshima และคณะ (2006) ได้ทำการศึกษาสาเหตุของการทำหน้าที่ต่างกันของกลุ่มข้อสอบ (Differential Bunble Functiong : DBF) โดยทำการแบ่งข้อสอบออกเป็นกลุ่ม ๆ แตกต่างกัน คือแบ่งตามลำดับขั้นของการเรียนรู้ วัตถุประสงค์ที่ต้องการวัดและแบ่งตามบทความที่อยู่ในแบบสอบ แล้วทำการคำนวณค่าดัชนี bundle-CDIF, bundle NCDIF และ bundle-DTF ซึ่งมีพื้นฐานมาจากทฤษฎีของวิธี DFIT เครื่องมือที่ใช้เป็นแบบวัดทักษะการอ่านจา ก Metropolitan Achievement Tests กลุ่มตัวอย่างที่ใช้เป็นนักเรียนเกรด 4 จำนวน 1,000 คน แบบสอบยาว 55 ข้อ ตัวแปรที่ใช้แบ่งกลุ่มตัวอย่างคือตัวแปรเพศและฐานะเศรษฐสังคม

ผลการศึกษาพบว่า สำหรับการเปรียบเทียบระหว่างเพศชายและเพศหญิง เมื่อตัดข้อสอบที่ตรวจพบว่าทำหน้าที่ ต่างกันออกไป 2 ข้อ คือข้อที่ 22 และ 25 แบบสอบฉบับนั้นก็ไม่ทำหน้าที่ต่างกันอีก นั่นคือดัชนี DTF ไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญ สำหรับการเปรียบเทียบระหว่างฐานะเศรษฐสังคม ตรวจไม่พบข้อสอบที่ทำหน้าที่ต่างกัน และเมื่อแบ่งวิเคราะห์ตามกลุ่มของข้อสอบพบว่าความถี่ 5 มีค่าดัชนี NCDIF สูงที่สุดโดยจะเข้าข้างเพศชายมากกว่าเพศหญิงแต่ค่าดัชนี bundle NCDIF มีค่าไม่แตกต่างกันเมื่อแบ่งกลุ่มตามฐานะเศรษฐสังคม

Larry (1997) ได้ทำการศึกษาเปรียบเทียบประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบการแปล โดยทำการเปรียบเทียบระหว่างกระบวนการ DFIT และวิธีแมนเทล – เฮนส์เซล (MH) ใช้ข้อมูลจากแบบสอบการแปล ซึ่งมีข้อสอบทั้งหมดจำนวน 50 ข้อ แต่

ในการวิจัยครั้งนี้ใช้เฉพาะข้อสอบแบบเลือกตอบ จำนวน 30 ข้อ โดยมีวัตถุประสงค์เพื่อศึกษาจำนวนและรูปแบบที่สอดคล้องกันระหว่างทั้ง 2 วิธีดังกล่าว

ผลการศึกษาพบว่ากระบวนการ DFIT มีความไวในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ดีกว่า นั่นคือกระบวนการ DFIT สามารถตรวจพบข้อที่ทำหน้าที่ต่างกันอย่างมีนัยสำคัญจำนวน 10 ข้อ จากข้อสอบทั้งหมด จำนวน 30 ข้อ ในขณะที่วิธี แมนเทล – เฮนส์เซล (MH) ระบุข้อสอบที่ DIF อย่างมีนัยสำคัญจำนวน 2 ข้อ เท่านั้น ในภาพรวมพบว่าทั้งสองวิธีระบุข้อสอบที่ DIF สอดคล้องกันเพียงร้อยละ 20 ข้อสอบที่กระบวนการ DFIT ระบุว่าเกิด DIF เป็นแบบเอกรูป จำนวน 4 ข้อและเป็นแบบอเนกรูป จำนวน 6 ข้อ ในขณะที่วิธี MH ระบุว่าข้อสอบที่เกิด DIF เป็นแบบเอกรูปและอเนกรูปอย่างละ 1 ข้อ

Rozowski และ Reith (1999) ได้ทำการตรวจสอบคุณภาพในการวัดและความตรงของแบบสอบที่ประกอบด้วยข้อสอบที่ DIF และไม่ DIF โดยใช้วิธีแมนเทล – เฮนส์เซล (MH) ในการตรวจสอบ DIF และแบ่งกลุ่มผู้สอบตามตัวแปรเพศและเชื้อชาติ ในการตรวจสอบคุณภาพของแบบสอบนั้น ได้ ใช้ค่าสัมประสิทธิ์แอลฟ่า ตรวจสอบความตรงของแบบสอบโดยเทียบกับเกณฑ์ภายนอกคือ ผลการสอบ American College Testing (ACT), Scholastic Aptitude Test (SAT) verbal (SATV), quantitative (SATQ), composite (SATC) และแรงค์เปอร์เซ็นต์ไทล์ของเกรดเฉลี่ย โดยใช้ข้อมูลของผู้สอบจากกลุ่มข้อมูลของ High School and Beyond (HSB) ที่จัดทำโดย National Educational Longitudinal Studies Program

ผลการวิจัยพบว่าแบบสอบที่ประกอบด้วยข้อสอบที่ทำหน้าที่ต่างกันนั้น ไม่ได้ทำให้แบบสอบมีคุณภาพในการวัดหรือมีความตรงต่ำลงมากนัก และเมื่อพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์แบบแรงค์ระหว่างคะแนนสอบของแบบสอบที่ไม่มีข้อสอบที่ DIF กับแบบสอบที่ประกอบด้วยข้อสอบที่ DIF เข้าข้างเฉพาะกลุ่มใดกลุ่มหนึ่ง และเข้าข้างกลุ่มใดกลุ่มหนึ่งในจำนวนที่เท่า ๆ กัน พบว่าค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้มีค่าสูง โดยแบบสอบที่มีข้อสอบที่ เข้าข้างทั้งสองกลุ่มในจำนวนที่สมดุลกันจะมีความสัมพันธ์กันแบบสอบที่ไม่มีข้อสอบที่ DIF เลย สูงที่สุด

Flowers และคณะ (1999) ได้พัฒนากระบวนการ DFIT สำหรับแบบสอบเอกมิติ (unidimensional) ที่มีข้อสอบที่มีการให้คะแนนแบบหลายค่า (polytomous) โดยกระบวนการนี้

ได้รับการพัฒนามาจากกระบวนการ DFIT สำหรับแบบสอบเอกมิติที่มีการให้คะแนนแบบ 2 ค่า (dichotomous) ของ Raju และคณะ (1995) ข้อมูลที่ใช้ในการศึกษาเป็นข้อมูลจำลอง ข้อสอบให้มีการให้คะแนนเป็น 0, 1, 2, 3 และ 4 องค์ประกอบที่สนใจศึกษาประกอบด้วยการกระจายค่าความสามารถของกลุ่มเปรี ยบเทียบ (ค่าความสามารถกระจายเหมือนกลุ่มอ้างอิง โดยทำการสุ่มกลุ่มตัวอย่างที่มีค่าความสามารถกระจายแบบ N(0,1) และค่าความสามารถกระจายไม่เหมือนกลุ่มอ้างอิง โดยทำการสุ่มกลุ่มตัวอย่างที่มีความสามารถกระจายแบบ N(-1,1) ความยาวของแบบสอบ (20 ข้อ และ 40 ข้อ) จำนวนข้อสอบที่ DIF (0%, 5%, 10% และ 20%) และทิศทางการเกิด DIF (แบบทิศทางเดียว และแบบสมดุลสองทิศทาง)

ผลการวิจัยพบว่ากระบวนการ DFIT มีประสิทธิภาพในการระบุข้อสอบที่ DIF และแบบสอบที่ DTF สำหรับข้อมูลที่มีการให้คะแนนแบบหลายค่า ความยาวของแบบสอบ การกระจายของค่าความสามารถข องกลุ่มเปรียบเทียบ จำนวนข้อสอบที่ DIF และทิศทางของการเกิด DIF มีผลกระทบเพียงเล็กน้อยต่ออัตราการตรวจสอบที่ถูกต้องหรืออัตราความผิดพลาดในทุก ๆ เงื่อนไข และข้อสอบที่มีค่า DIF มากจะถูกตรวจพบ ในขณะที่ข้อสอบที่มีค่า DIF น้อย ๆ จะไม่ถูกตรวจพบ ในภาพรวมดัชนี NCDIF มีความคงที่มากกว่าดัชนี CDIF มีอยู่ 2 เงื่อนไขที่ดัชนี CDIF ไม่เป็นไปตามที่คาดไว้ คือเงื่อนไขที่ 1 เมื่อมีแบบสอบที่มีความยาว 20 ข้อ และการกระจายของค่าความสามารถของกลุ่มอ้างอิงเหมือนกับกลุ่มเปรียบเทียบ ดัชนี CDIF ระบุข้อที่ไม่ DIF ว่าเกิด DIF ผิดพลาด ร้อยละ 18 เงื่อนไขที่ 2 เมื่อแบบสอบมีความยาว 40 ข้อ และการกระจายความสามารถของทั้งสองกลุ่มต่างกัน ดัชนี CDIF ระบุข้อที่ DIF ได้เพียงร้อยละ 50 จากข้อสอบที่ DIF ทั้งหมด

Ellis และ Mead (2000) ได้ทำการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและแบบสอบ โดยใช้กระบวนการ DFIT สำหรับแบบสอบถาม Sixteen Personality Factor (16PF) ที่แปลจากภาษาอังกฤษเป็นภาษาสเปน โดยแบ่งกลุ่มผู้สอบเป็นกลุ่มย่อยตามเชื้อชาติ 3 กลุ่ม ประกอบด้วย กลุ่ม Anglo Americans, Spanish speakers และ Hispanic Americans

ผลการวิจัยพบว่า ดัชนี CDIF/DTF ตรวจพบข้อสอบที่ทำ หน้าที่ต่างกันน้อยกว่า เมื่อใช้ดัชนี NCDIF ในการตรวจสอบ ในการวิเคราะห์พบว่า เมื่อแบ่งกลุ่มผู้สอบเปรียบเทียบระหว่างกลุ่ม Anglo Americans และกลุ่ม Spanish speakers ตรวจพบข้อสอบที่ DIF มากที่สุด และเมื่อ

แบ่งกลุ่มผู้สอบเปรียบเทียบระหว่างกลุ่ม Anglo Americans และกลุ่ม Hispanic Americans ตรวจพบข้อสอบที่ DIF จำนวนน้อยที่สุด

1.7.2 การทำหน้าที่ต่างกันของตัวลวง

Marshall (1993) ทำการตรวจสอบความแตกต่างของเพศในตัวเลือกที่เป็นตัวลวง โดยใช้แบบสำรวจแคลิฟอเนียร์ เวอร์ชั่น 6 ของทักษะพื้นฐาน ผู้วิจัยสนใจในความแตกต่างของเพศตัวเลือกที่เป็นตัวลวง และปีที่สร้างแบบสอบขึ้นมาโดยใช้โมเดลลอกซ์ลิเนีย ซึ่งยอมให้อิทธิพลหลักแยกปฏิกิริยาของตัวแปรปรับเทียบเป็นแต่ละประเภท ข้อสอบมีจำนวน 30 ข้อในแบบสอบ 13% แสดงให้เห็นถึงการทำหน้าที่ต่างกันของตัวลวงระหว่างเพศ รวมถึงรูปแบบของความคลาดเคลื่อนขึ้นอยู่กับความแตกต่างของปีที่ใช้แบบสอบ และ 20% ของข้อสอบแสดงถึงการทำหน้าที่ต่างกันของตัวลวง การศึกษาในครั้งนี้ยังใช้ตัวเลือกที่เป็นตัวลวงในการแยกชนิดของความคลาดเคลื่อนที่เกิดขึ้นโดยเพศชายและเพศหญิง พบว่าความแตกต่างของเพศเป็นชนิดของความคลาดเคลื่อนที่มีขนาดใหญ่

Green et al. (1989) เป็นการศึกษาที่ใช้โมเดลลอกซ์ลิเนียในการประเมินการทำหน้าที่ต่างกันของตัวลวง ข้อสอบที่ใช้ในการศึกษาคือข้อสอบ SAT ในปี 1986 ในส่วนที่เกี่ยวกับคำซึ่งประกอบไปด้วยข้อสอบ 85 ข้อ การศึกษาครั้งนี้ใช้การสุ่มนักเรียนจำนวนทั้งสิ้น 2,000 คน โดยแยกกลุ่มผู้สอบเป็นกลุ่มคนผิวขาว ผิวดำ และชาวละติน ผลจากการวิเคราะห์ 16% ของข้อสอบทั้งหมดแสดงการทำหน้าที่ต่างกันของตัวลวง ตามระดับความมีนัยสำคัญทางสถิติ และในการเลือกคำตอบดูเหมือนจะเป็นการเลือกที่มีความแตกต่างมากของนักเรียนที่เป็นชาวละติน อย่างไรก็ตามข้อสอบไม่ มากที่แสดงถึงความแตกต่างถึง 10% โดยที่ข้อสอบส่วนใหญ่แสดงความแตกต่างระหว่าง 2% และ 3% ระหว่างตัวลวงในแต่ละข้อ

Bank (2006) เป็นการศึกษาเพื่อตรวจสอบการทำหน้าที่ต่างกันของตัวลวงระหว่างกลุ่มผู้สอบที่มีผิวดำ ชาวละติน และผิวขาว จำนวน 4,765 คน โดยเป็นนักเรียนเกรด 5 ข้อสอบเป็นการอ่านจำนวน 20 ข้อและส่วนที่เป็นศิลปทางภาษาของแบบสอบเทอร่าโนวา ใช้ลักษณะทางวัฒนธรรมของคนผิวดำในการเลือกตัวลวง หลังจากการเลือกข้อสอบนั้น ลักษณะของวัฒนธรรมคนผิวดำในการเลือกตัวเลือกที่ผิด และการตรวจสอบกลุ่มของข้อสอบสำหรับการทำหน้าที่ต่างกัน

แบบ Bundle (Differential Bundle Functioning; DBF) พบว่าข้อสอบ 11 ข้อ แสดง DBF ซึ่งเป็นการขยายการศึกษาการทำหน้าที่ต่างกันของข้อสอบ และปรากฏเมื่อผู้สอบมีระดับคุณลักษณะแฝงที่เหมือนกัน มีความแตกต่างที่เป็นไปได้ของคำตอบกลุ่มเดียวกันของข้อสอบที่ถูกต้อง กลุ่มนี้จะมีรูปแบบของเกณฑ์มาตรฐานที่กำหนดไว้ล่วงหน้า เช่นเนื้อหาของข้อสอบหรือการวิเคราะห์เพื่อจัดกลุ่มข้อมูลที่มีความสัมพันธ์กัน (Cluster Analysis) การตรวจสอบต่อไปด้วยข้อสอบ 11 ข้อสำหรับการทำหน้าที่ต่างกันของตัวลวงโดยใช้โมเดล loglinear และข้อสอบชุดนี้จำนวน 8 ข้อแสดงถึงการทำหน้าที่ต่างกันของตัวลวงระหว่างผู้สอบผิวดำและผิวขาว โดยผลที่ได้รวมไปถึงแบบสอบเทอร่าโนวาด้วย ตัวลวงส่วนมากเหมาะสมที่จะแสดงการทำหน้าที่ต่างกันของตัวลวงมากกว่าตัวเลือกที่ถูกซึ่งเหมาะสมที่จะแสดงการทำหน้าที่ต่างกันของข้อสอบ และดังนั้นข้อแนะนำที่เป็นการรวมกา รวิเคราะห์การทำหน้าที่ต่างกันของตัวลวงในข้อสอบจะสามารถให้เข้าใจเรื่องข้อสอบมากยิ่งขึ้น

Penfield (2008) เป็นการศึกษาเกี่ยวกับการทำหน้าที่ต่างกันของตัวลวงที่ให้สารสนเทศที่มีคุณค่าเกี่ยวกับแหล่งและสาเหตุที่เป็นไปได้ของความไม่แปรเปลี่ยนในข้อสอบแบบหลายตัวเลือก โดยเป็นการนำเสนอการประมาณค่าออด เรโชว์ของอิทธิพลของการทำหน้าที่ต่างกันของตัวลวงภายใต้ nominal response model ผลการวิจัยจากการจำลองข้อมูลชี้ให้เห็นว่า การประมาณค่าอิทธิพลของการทำหน้าที่ต่างกันของตัวลวงสามารถรักษาคุณสมบัติทางสถิติที่ดีภายใต้ความหลากหลายของเงื่ อนไข

Koon (2010) เป็นการศึกษาการเปรียบเทียบวิธีการที่ใช้ในการตรวจสอบการทำหน้าที่ต่างกันของตัวลวง ผลการวิจัยพบว่า วิธีอัตราส่วนออด และการตรวจสอบด้วยวิธีถดถอยแบบโลจิสติกให้ผลการตรวจสอบการทำหน้าที่ต่างกันของตัวลวงที่มีสอดคล้องกัน ส่วนวิธีการตรวจสอบแบบถดถอยแบบพหุให้ผลการตรวจสอบทีสอดคล้องกับวิธีการวิเคราะห์ตัวลวงแบบมาตรฐาน

จากการศึกษางานวิจัยที่เกี่ยวข้องกับการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง พบว่า งานวิจัย เกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบ ในอดีตที่ผ่านมานั้น มีการศึกษาโดยใช้ทั้งข้อมูลจำลองและข้อมูลเชิงประจักษ์ โดยส่วนที่เป็นการใช้ข้อมูลจำลองส่วนใหญ่จะเป็นการเปรียบประสิทธิภาพของวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบชนิดต่างๆใน

เงื่อนไขหรือในรูปแบบที่มีความแตกต่างกัน โดยวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่พบว่าได้รับความนิยมในการนำมาใช้ในการทำวิจัยเป็นจำนวนมากคือ วิธีซิบเทสท์ (SIBTEST) และวิธีแมนเทล-แฮนส์เซล (MH) ส่วนวิธีการตรวจสอบแบบอื่นก็มีปรากฏให้เห็นบ้าง โดยในการศึกษานั้นนักวิจัยหลายๆท่านมีการวิเคราะห์ประสิทธิภาพของวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยให้เงื่อนไขที่แตกต่างกันในหลายๆเงื่อนไข ได้แก่ ขนาดของกลุ่มตัวอย่าง ความยาวของแบบสอบ ลักษณะของข้อสอบ การแจกแจงค่าความสามารถ สัดส่วนของการทำหน้าที่ต่างกันของข้อสอบ และขนาดของอิทธิพลของข้อสอบที่ทำหน้าที่แตกต่างกัน ส่วนในงานวิจัยโดยใช้ข้อมูลเชิงประจักษ์ตัวแปรที่สนใจนำมาศึ กษา เช่น เพศ สีผิว เชื้อชาติ วัฒนธรรม สภาพทางเศรษฐสังคม ประสบการณ์ในการสอบ ภูมิศาสตร์ ศาสนา ระดับสติปัญญา เป็นต้น

จากงานวิจัยที่เกี่ยวข้องกับการทำหน้าที่ต่างกันของข้อสอบ ทั้งภายในประเทศ และจากต่างประเทศจะเห็นว่าส่วนใหญ่ลักษณะการทำวิจัยจะมี 2 ประเด็นหลัก นั่นคือประเด็นแรกที่พบมากเกี่ยวกับการเปรียบเทียบประสิทธิภาพของวิธีตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยที่ในช่วงแรกวิธีการตรวจสอบที่ผลการวิจัยส่วนใหญ่พบว่ามีประสิทธิภาพกว่าวิธี อื่นๆที่นำมาเปรียบเทียบ นั่นคือ วิธีวัดพื้นที่ความแตกต่างระหว่างโค้งการตอบสนองข้อสอบ (IRT-D^2) แต่หลังจากนั้นวิธีแมลเทล-แฮนส์เซล และวิธีซิบเทสท์กลับได้รับความนิยมในการนำไปใช้ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ประเด็นที่สองพบว่าเป็นการศึกษาเกี่ยวกับ สาเหตุของการทำหน้าที่ต่างกันของข้อสอบ เช่น เพศ หรือภูมิภาคที่แตกต่าง เป็นต้น

สำหรับการศึกษาเกี่ยวกับการทำหน้าที่ต่างกันของตัวลวง นั้นงานวิจัยในประเทศไทยยังไม่พบว่ามีการศึกษาแล้วเผยแพร่ ส่วนในต่างประเทศก็ยังมีการศึกษากันน้อยมาก แต่ในระยะหลังมีแนวโน้มที่มีผู้สนใจศึกษามากขึ้นกว่าในอดีต และจากผล การศึกษาดังกล่าว ทำให้ผู้วิจัยเห็นถึงประโยชน์ของการศึ กษาเรื่องของการทำหน้าที่ของตัวลวง หากมีการศึกษาอย่างจริงจัง ก็จะเป็นประโยชน์แก่การพัฒนาในเรื่องของการวัดผลเป็นอย่างยิ่ง

ตอนที่ 2 แนวคิดเกี่ยวกับมูลค่าเพิ่มของการจัดการศึกษา

2.1 นิยามและความหมายของมูลค่าเพิ่ม

ความหมายของคำว่า มูลค่าเพิ่ม (Value-added) ในด้านการศึกษา มีความเกี่ยวข้องกับผลสำเร็จหรือผลสัมฤทธิ์ของนักเรียนที่เกิดจากการดำเนินงานของสถานศึกษา และมีการพัฒนาโมเดลมูลค่าเพิ่มหรือการประเมินมูลค่าเพิ่ม เพิ่มขึ้นอย่างมากตามการกำหนดนิยามของโมเดลหรือลักษณะเฉพาะของโมเดล นักการศึกษาได้ให้ความหมายไว้อย่างหลากหลายดังนี้

การวัดมูลค่าเพิ่ม เป็นความพยายามเพื่อระบุว่าคุณค่า (Value) ทางการศึกษาที่สถานศึกษาได้ปฏิบัติให้เพิ่มมากขึ้น ซึ่งทำนายจากพื้นฐานของนักเรียนและความสำเร็จเดิมของนักเรียน (Hill, 1995 cite in Downes & Vindurampulle, 2007)

โมเดลมูลค่าเพิ่ม เป็นการ เปลี่ยนแปลงในระดับนักเรียนโดยการเปลี่ยนปริมาณความคาดหวังของการเปลี่ยนแปลง ตลอดจนการเปลี่ยนแปลงที่เป็นพื้นฐานของนักเรียน เช่น เพศ ระดับ SES เชื้อชาติ ผลสัมฤทธิ์เดิมหรือระดับคะแนนเริ่มต้น (Lissitz. Et.al, 2006)

การวัดมูลค่าเพิ่ม เป็นการทุ่มเทการดำเนินการของปั จจัยที่มีอยู่อย่างหลากหลายเพื่อให้เกิดผลต่อการเปลี่ยนแปลงของผลสัมฤทธิ์ของนักเรียน (Goldhaber & Anthony, 2003)

โมเดลมูลค่าเพิ่ม เป็นการระบุเกี่ยวกับการทุ่มเทดำเนินงานของสถานศึกษาเพื่อทำให้เกิดความก้าวหน้าของนักเรียนหรือเป็นไปตามวัตถุประสงค์ของการศึกษา เช่น ผลสัมฤทธิ์ทางด้านสติปัญญาซึ่งการดำเนินงานของสถานศึกษายังเกี่ยวข้องกับปัจจัยอื่นหลายปัจจัยที่สนับสนุนให้เกิดความก้าวหน้าทางการศึกษาของนักเรียน (OECD, 2008)

โมเดลมูลค่าเพิ่ม เป็นการเปลี่ยนแปลงในระดับนักเรียนโดยการเปลี่ยนปริมาณความคาดหวังของการเปลี่ยนแปลง ตลอดจนการเปลี่ยนแปลงที่เป็นพื้นฐานของนักเรียน เช่น เพศ ระดับ SES เชื้อชาติ ผลสัมฤทธิ์เดิมหรือระดับคะแนนเริ่มต้น (Lissitz. Et.al, 2006)

โมเดลมูลค่าเพิ่มเป็นส่วนที่เพิ่มขึ้นจากสภาวะเริ่มต้นของนักเรียนที่เกิดจากการดำเนินงานของสถานศึกษา ได้พิจารณาปัจจัยอื่นที่อยู่ นอกเหนือการควบคุมของสถานศึกษาหรือปรับให้อยู่ในพื้นฐานที่เท่าเทียมกัน โดยปัจจัยที่อยู่นอกเหนือการควบคุมของสถานศึกษา เช่น ตัวแปรภูมิหลังของนักเรียน สภาพความพร้อมทางด้านเศรษฐกิจของครอบครัว ทรัพยากรทางการเรียน เป็นต้น (ประกฤติยา ทักษิโณ, 2552)

โดยเมื่อพิจารณา ความหมายเกี่ยวกับมูลค่าเพิ่มจึงสรุปได้ว่า โมเดลมูลค่าเพิ่มเป็นการเปลี่ยนแปลงในระดับนักเรียนที่เกิดจากการดำเนินงานของสถานศึกษา โดยสามารถพิจารณาได้จากตัวแปรระดับนักเรียนและสถานศึกษา เช่นเจตคติทางวิทยาศาสตร์ของนักเรียน การแสดงความตั้งใจที่จะใฝ่หาสาระทางวิทยาศ าสตร์ เพศ ระดับ SES เชื้อชาติ ผลสัมฤทธิ์เดิมหรือระดับคะแนนเริ่มต้น เป็นต้น

ผลจากการประเมินมูลค่าเพิ่มมีความสำคัญกับการนำมาใช้กำหนดนโยบายที่สำคัญ 3 ด้าน (OECD, 2008) คือการริเริ่มการพัฒนาสถานศึกษา (school improvement initiatives) ระบบความสามารถในการตรวจสอบไ ด้ของสถานศึกษา (Systems of school accountability) และการเลือกสถานศึกษา (School choice) ในส่วนของการริเริ่มการพัฒนาสถานศึกษา เป็นสิ่งสำคัญที่การพัฒนากิจกรรมหรือชุดกิจกรรมที่ต้องการพัฒนาเป็นอันดับแรก ผลจากการวิเคราะห์ทางสถิติจะให้ข้อสนเทศเกี่ยวกับความสัมพันธ์ ระหว่างปัจจัยนำเข้าและตัวบ่งชี้ของความสามารถของสถานศึกษาที่สามารถบอกถึงยุทธวิธีที่ควรปฏิบัติหรือไม่ควรปฏิบัตินำไปสู่การกำหนดนโยบายและวางแผ นนโยบาย การแบ่งสรรทรัพยากรในส่วนของความสามารถในการตรวจสอบได้ของสถานศึกษา ผลจากการวิเคราะห์มูลค่าเพิ่ม สามารถตรวจสอบการปฏิบัติงานและผลลัพธ์จากการปฏิบัติได้ทั้งหมด (McKewen, 1995) เช่นการให้ข้อมูลแก่ผู้ที่เกี่ยวข้องและสาธารณะเกี่ยวกับเงินภาษีได้อย่างมีประสิทธิภาพ ในส่วนของการเลือกสถานศึกษา ข้อมูลที่ได้จากการวิเคราะห์นี้จะมีจุดมุ่งหมายไปที่ผู้ปกครองและครอบครัวของนักเรียนเกี่ ยวกับความสามารถที่แตกต่างของสถานศึกษาเพื่อการตัดสินเลือกสถานศึกษา ผู้ปกครองสามารถเลือกสถานศึกษาที่ดีที่สุดและเหมาะสมกับแหล่งทรัพยากรที่สามารถเกิดขึ้นในสถานศึกษา (Hoxby, 2003)

ข้อมูลที่ได้จากการวิเคราะห์มูลค่าเพิ่มมีบทบาทสำคัญต่อนโยบายการจัดสรรแหล่งทรัพยาก ร และกลยุทธ์การสอนที่จำเป็นในแต่ละสถานศึกษา การวัดมูลค่าเพิ่มมีความสำคัญหลายประการสำหรับใช้ในการตัดสินใจแก่ผู้ที่เกี่ยวข้องหลายฝ่าย (Drury & Doran, 2003, Downes & Vindurampulle, 2007 และ OECD, 2008) ดังนี้

1. การประเมินนโยบายและโครงการ (Policy and program evaluation) ข้อมูลจากการวัดมูลค่าเพิ่มสามารถวิเคราะห์ได้ถึงการปฏิบัติงานของสถานศึกษาที่สามารถเปรียบเทียบผลกระทบของสถานศึกษาเชิงสัมพัทธ์กับสถานศึกษาอื่นที่พิจารณารวมกับปัจจัยอื่น เช่นคะแนน

ดิบ การประเมินครู การประเมินตนเองของสถานศึกษา การทบทวนการตัดสินและแนว โน้มของสถานศึกษา ข้อมูลที่ได้สามารถช่วยในการตัดสินใจเชิงนโยบาย (Informing policy-making) ที่มีประสิทธิภาพต่อการกำหนดนโยบายในระยะเริ่มต้นเป็นประโยชน์ในการระบุสถานศึกษาที่มีค่าเฉลี่ยสูงกว่าและต่ำกว่าค่าเฉลี่ยของความสามารถได้อย่างชัดเจน (OECD, 2006) สถานศึกษาที่มีความสามารถสูงสามารถให้ตัวอย่างในการปฏิบัติที่ดี ช่วยให้วางนโยบายให้มีความสอดคล้องกันระหว่างของนโยบายแหล่งทรัพยากรและยุทธวิธีการสอนกับคุณลักษณะของนักเรียนและสถานศึกษา (Drury & Doran, 2003)

2. เครื่องมือหนึ่งสำหรับการพัฒนาสถานศึกษา (A tool for school improvement) เนื่องจากการวัดมูลค่าเพิ่มจะนำคุณลักษณะของนักเรียนเข้าไปพิจารณาและสามารถระบุถึงวิธีการของสถานศึกษาที่สามารถทำให้ผู้เรียนมีความสามารถตามที่ความคาดหวัง โดยอิงความสำเร็จของนักเรียน การวิเคราะห์มูลค่าเพิ่มจึงเป็นสิ่งที่สะท้อนถึงสิ่งแวดล้อมและการปฏิบัติการของสถานศึกษาเพื่อแปลความถึงผลลัพธ์ของบริบทของสถานศึกษาและชั้นเรียน เพื่อเป็นแนวทางในการพัฒนาสถานศึกษา การวัดมูลค่าเพิ่มจะสามารถชี้ได้ถึงปัจจัยที่ส่งผลให้นักเรียนแสดงความสามารถได้เป็นอย่างดีหรือสามารถแสดงได้ต่ำกว่าที่คาดหวัง นั่นคือสามารถช่วยในอิทธิพลโดยตรงเพื่อพัฒนาผลลัพธ์การเรียนรู้ของนักเรียน

3. ประสิทธิผลของครู (Teacher effectiveness) วิธีการวิเคราะห์มูลค่าเพิ่มมีผลต่อผลสัมฤทธิ์ของนักเรียนซึ่งนำไปใช้ในการประเมินประสิทธิภาพของครู จากการศึกษาการวิเคราะห์มูลค่าเพิ่ม พบว่าประสิทธิภาพของครูเป็นตัวบ่งชี้ที่สำคัญของความสัมพันธ์ของสถานศึกษาต่อการเรียนรู้ของนักเรียน และนอกจากนี้จากข้อสังเกตพบว่าการวิเคราะห์มูลค่าเพิ่มจะเป็นเครื่องมือในการลดช่องว่างของผลสัมฤทธิ์ของนักเรียนในข้อได้เปรียบและข้อเสียเปรียบของนักเรียนนอกจากนี้ข้อมูลที่ได้จะเป็นการให้ข้อมูลย้อนกลับ แก่ครูเกี่ยวกับวิธีการสอน มีความชัดเจนในเป้าหมายของการเรียนรู้ที่ส่งผลต่อการเปลี่ยนแปลงวิธีการสอน และสามารถใช้เป็นข้อมูลสำหรับครูในการพัฒนานักเรียนต่อไป กล่าวคือสามารถระบุคุณลักษณะของนักเรียนที่ต้องมีการพัฒนาทักษะต่อไปและใช้เป็นข้อมูลต่อการวางแผนการพัฒนาของครู ตัวอย่างเช่น การวิเคราะห์มูลค่าเพิ่ม จะระบุครูที่มีประสิทธิภาพในการยกระดับความสามารถของนักเรียนที่มีจุดอ่อน (Drury & Doran, 2003)

4. เครื่องมือหนึ่งสำหรับความรับผิดชอบที่ตรวจสอบได้ (A tool for accountability) ตัวบ่งชี้การปฏิบัติงานของสถานศึกษาเป็นส่ว นหนึ่งของการตรวจสอบได้ของสถานศึกษา การวัดมูลค่าเพิ่มจะเพิ่มความรอบคอบในการใช้ที่เป็นพื้นฐานสำหรับการกำหนดภาพรวมของการปฏิบัติงานสถานศึกษาจากรายงานของ OECD (2008) เกี่ยวกับวิธีการมูลค่าเพิ่มจะได้รับข้อสรุปที่คล้ายกันเกี่ยวกับกระบวนการการศึกษา โดยรายงานว่าโมเ ดลมูลค่าเพิ่มเป็นวิธีที่ดีที่สุดที่แนะนำได้ว่าเป็นเครื่องมือที่ส่งเสริมการพัฒนาสถานศึกษา

5. การรายงานต่อผู้ปกครองและชุมชน (Reporting to parents and community) การวัดมูลค่าเพิ่มจะให้ข้อสารสนเทศกับผู้ปกครองได้อย่างถูกต้องเกี่ยวกับวิธีการที่สถานศึกษาได้พัฒนาผลการเรียนรู้ของนักเรียน โดยรายงานให้ทราบถึงการปฏิบัติของสถานศึกษาในบริบทของสมรรถนะของนักเรียนและสิ่งแวดล้อมในชุมชนของพวกเขา การวัดมูลค่าเพิ่มจะให้ข้อสารสนเทศแก่ผู้ปกครองต่อผลสำเร็จของครูในการยกระดับผลสัมฤทธิ์ของนักเรียน

วิธีการวัดมูลค่าเพิ่มได้มีการนำมาใช้ ในหลายประเทศเพื่อใช้ผลมาพิจารณาการดำเนินการของสถานศึกษาที่ส่งต่อผลสำเร็จของนักเรียน (OECD, 2008) ดังตัวอย่างของระบบการประเมินในรัฐเทนเนสซีของประเทศสหรัฐอเมริกาที่มีชื่อระบบการประเมินว่าระบบการวัดมูลค่าเพิ่มของเทนเนสซี (Tennessee Value-added Assessment System: TVAAS) ซึ่งออกแบบเพื่อวัดผลกระทบของสถานศึกษา และครูที่ส่งผลต่อความสำเร็จของนักเรียน ซึ่งพัฒนาโดย William L. Sanders และมีการศึกษาเพิ่มเติมไปสู่การดำเนินการระบบการประเมินมูลค่าเพิ่มทางการศึกษา (Education Value-Added Assessment System: EVAAS) (Sanders, 1998) นอกจากนี้ Heck (2000) ยังศึกษาการตรวจสอบผลกระทบของคุณภาพสถานศึกษาต่อผลลัพธ์ของสถานศึกษาและการพัฒนาโดยการประยุกต์ใช้วิธีการมูลค่าเพิ่ม โดยให้ความสำคัญไปที่ระดับรัฐเพื่อติดตามผลสำเร็จของนักเรียน และทำความเข้าใจเกี่ยวกับความสามารถรับผิดชอบของสถานศึกษา กา รรายงานผลการปฏิบัติงานของสถานศึกษาแก่ผู้กำหนดนโยบายบุคคลากรของสถานศึกษาและผู้ปกครองเพื่อใช้ผลการเปรียบเทียบสถานศึกษาให้ความยุติธรรมยิ่งขึ้น

2.2 ขั้นตอนและวิธีการที่ใช้วิเคราะห์มูลค่าเพิ่ม

ขั้นตอนที่สำคัญของการวิเคราะห์มูลค่าเพิ่มแบ่งเป็น 5 ขั้นตอน (Scheerens and Bosker, 1997; Hanson, 1988 cited in Miller, 1990) คือ 1) การกำหนดเป้าหมายของการวัดผลผลิตที่ต้องการเปรียบเทียบ 2) กำหนดชนิดของข้อมูลและการรวบรวมข้อมูล 3) การสร้างโมเดลการวิเคราะห์ซึ่งประกอบด้วยปัจจัยที่มีความสัมพันธ์อย่างมีเหตุผลกับผลผลิตที่เกิดขึ้น 4) การกำหนดหรือการสร้างเครื่องมือการวัดสถานะของนักเรียนตามปัจจัยที่กำหนดขึ้น 5) การเลือกใช้เทคนิคทางสถิติที่เหมาะสม นอกจากนี้ OECD (2008) ได้เสนอขั้นตอนที่สำคัญในการวิเคราะห์มูลค่าเพิ่ม เพื่อสามารถนำไปปฏิบัติได้อย่างรวดเร็วและง่ายต่อการทำความเข้าใจ จำนว น 8 ขั้นตอน ซึ่งมีแนวทางเดียวกันกับขั้นตอนที่เสนอไปข้างต้น แต่มีการนำเสนอขั้นตอนที่ละเอียดเพิ่มมากขึ้นเกี่ยวกับแนวทางในการนำโมเดลมูลค่าเพิ่มไปดำเนินการใช้ซึ่งเสนอในขั้นที่ 5, 6, 7 และ 8 ดังนี้

ขั้นตอนที่ 1 การกำหนดวัตถุประสงค์และการวัดผลการปฏิบัติงานของสถา นศึกษา ซึ่งขั้นตอนนี้จะกำหนดวัตถุประสงค์สำหรับการนำโมเดลมูลค่าเพิ่มไปใช้ประกอบด้วยการกำหนดลักษณะของสารสนเทศของมูลค่าเพิ่มและอธิบายถึงการอธิบายคะแนนมูลค่าเพิ่ม การพิจารณาโครงสร้างของการประเมินนักเรียนที่มีความสอดคล้องกับวัตถุประสงค์โมเดลมูลค่าเพิ่ม ตลอดจนการระบุเนื้อหาและระดับชั้นที่ประเมิน การกำหนดเป้าหมายการประเมิน และการพิจารณาศักยภาพของเครื่องมือประเมินจะต้องมั่นใจว่าสามารถใช้ในโมเดลมูลค่าเพิ่ม ซึ่งเป็นสิ่งสำคัญที่จะนำมาใช้สำหรับนำมาแปลความหมายของผลการปฏิบัติ

ขั้นตอนที่ 2 การนำเสนอและการใช้ข้อสารสนเทศของ มูลค่าเพิ่ม เป็นการตัดสินและเลือกวิธีการที่เหมาะสมสำหรับการแสดงส ารสนเทศของมูลค่าเพิ่ม ตลอดจนแนวทางการแปลความหมายของมูลค่าเพิ่ม ซึ่งสามารถสร้างขึ้นได้โดยผู้ที่เกี่ยวข้องในการจัดทำนโยบาย

ขั้นตอนที่ 3 การพิจารณาคุณภาพของข้อมูล เป็นขั้นตอนที่ควรจะพิจารณาระบบขอ งข้อมูลในสถานศึกษาและสารสนเทศเพื่อรวบรวมสำหรับนำไปใช้ในโมเดลมูลค่าเพิ่ม ตลอดจนการกำหนดกลุ่มตัวอย่างนักเรียนที่เพียงพอต่อการนำมาใช้ในโมเดลมูลค่า ซึ่งอาจเกิดการสูญหายข้อมูล (missing data) ที่อาจเกิดการย้ายสถานศึกษาในช่วงที่มีการประเมินข้อมูลที่จำเป็นจะต้องสนับสนุนโมเดล หรือการนำไปใช้สำหรับการกำหนดนโยบายและพัฒนาโครงการต่อไป

ขั้นตอนที่ 4 การเลือกโมเดลมูลค่าเพิ่มที่เหมาะสม การกำหนดวัตถุประสงค์เชิงนโยบายและการใช้คะแนนมูลค่าเพิ่มของสถานศึกษาจะเป็นแนวทางในการเลือกโมเดล สิ่งจำเป็นที่จะต้องกำหนดคือการเลือกใช้สถิติและ วิธีการวิเคราะห์ภายใต้โมเดลมูลค่าเพิ่ม ความแปรปรวนในโมเดลมูลค่าเพิ่ม เพื่อประเมินความเหมาะสมของลักษณะของโมเดล ลักษณะเฉพาะของโมเดลอาจจะทำให้เกิดความแตกต่างของสถานศึกษา ผลกระทบของการเลือกโมเดลขึ้นอยู่กับความแตกต่างของสถานศึกษาในจุดประสงค์การใช้คะแนนมูลค่าเพิ่ม

ขั้นตอนที่ 5 กลยุทธ์กระตุ้นการสื่อสาร และผู้ที่มีส่วนได้ส่วนเสีย เป้าหมายของวิธีการสื่อสารจะสอดคล้องกับวัตถุประสงค์ของนโยบาย การวัดความสามารถการปฏิบัติงานของโรงเรียนจะนำมาตัดสินควรจะมีการอธิบายได้อย่างชัดเจน การกระตุ้นและการสื่อสารต่อผู้ที่เกี่ยวข้องจะต้องชี้แจงได้อย่างชัดเจน เพื่อนำมาใช้แปลความหมายในผลของมูลค่าเพิ่มของสถานศึกษา

ขั้นตอนที่ 6 การฝึกอบรม เป้าหมายของการฝึกอบรม มีเป้าหมายในการพัฒนากลุ่มผู้ใช้ การฝึกอบรมผู้บริหารสถานศึกษาและครู จะเป็นการอธิบายเกี่ยวกับผลของมูลค่าเพิ่มและสามารถนำมาใช้ภายในสถานศึกษาเพื่อวัตถุประสงค์ของการพัฒนาสถานศึกษา รวมถึงการอบรมเกี่ยวกับการวิเคราะห์ทางสถิติและการใช้ข้อสารสนเทศที่จำเป็น นอกจากนี้การฝึกอบรมสำหรับผู้ปกครองมีเป้าหมาย สำหรับเพื่อการแปลความหมายของคะแนนมูลค่าเพิ่มเพื่อความสะดวกในการเลือกสถานศึกษา

ขั้นตอนที่ 7 การนำร่องโครงการ เป็นการตัดสินใจเกี่ยวกับขนาดและคุณลักษณะของกลุ่มตัวอย่างของสถานศึกษาที่จะมีความสอดคล้องกับวัตถุประสงค์ของนโยบายการนำระบบของโมเดลมูลค่าเพิ่มไปใช้ การนำร่องโครงการ ควรจะใช้เพื่อประเมินกิจกรรมที่มีความเชื่อมโยงกับโมเดลมูลค่าเพิ่ม

ขั้นตอนที่ 8 การนำไปพัฒนา เป็นขั้นตอนสุดท้ายของการนำโมเดลมูลค่าเพิ่มไปใช้ ซึ่งเป็นการติดตามผลของคุณภาพ แหล่งทรัพยากร และมีการพิจารณาถึงความสอดคล้องของโมเดลมูลค่าเพิ่มกับนโยบาย

สถิติที่นำมาใช้สำหรับการวัดมูลค่าเพิ่มมีอย่างหลากหลายวิธีที่มีระดับความซับซ้อนต่างกัน และผลการวิเคราะห์ที่แตกต่างกัน (Fitz-Gibbon, 1996; สุชีรา มะหิเมือง, 2547; ศิริชัย กาญจนวาสี, 2550) สรุปได้ดังนี้

1. การวิเคราะห์โดยวิธีการหักออกของคะแนนดิบ (Simple Subtraction of raw score) เป็นการวัดจุดเริ่มต้นและจุดสิ้นสุดของการศึกษา ผลที่ได้ของวิธีนี้จะเรียกคะแนน พัฒนาการ (gain score) หรืออิทธิพลรวม (gross effects) เป็นวิธีที่เหมาะสมเมื่อมีการวัดผลสัมฤทธิ์เริ่มต้นและผลสัมฤทธิ์ในเวลาต่อมาด้วยวิธีเดียวกัน ผลที่ได้ลักษณะนี้จึงเป็นข้อมูลดิบ (raw data) (มูลค่าเพิ่ม = ผลสัมฤทธิ์ครั้งหลัง – ผลเริ่มต้น) ที่ทำให้เห็นภาพรวม แต่ยากในการแปลความหมายของมูลค่าเพิ่มที่เกิดขึ้น นอกจากนี้การแสดงผลรวม (gross) จะแสดงความหมายในเชิงเปรียบเทียบว่าผ่านหรือไม่ผ่านเกณฑ์ได้ก็ต่อเมื่อมีตัวบ่งชี้คุณภาพกำกับไว้ หรือมีการเปรียบเทียบกับกลุ่มเหมือน โดยทั่วไปวิธีนี้มักเป็นที่รู้จักกันในรูปของค่าเฉลี่ยผลสัมฤทธิ์ (mean achievement) ของนักเรียนทุกคนในสถานศึกษาที่มีการควบคุมเรื่องความคลาดเคลื่อนจากการสุ่มตัวอย่างแล้ว

2. การวิเคราะห์ถดถอย (Regression analysis) เป็นวิธีที่คำนึงถึงความเท่าเทียมกันของพื้นฐานความรู้เดิมของนักเรียน หรืออาจจะเป็นปัจจัยอื่นโดยพิจ ารณาตามเงื่อนไขหรือทฤษฎีที่เกี่ยวข้อง และสารสนเทศของบริบทที่กำหนดในโมเดลการวิเคราะห์ จึงจัดได้ว่าเป็นวิธีที่สามารถบ่งบอกถึงมูลค่าที่สถานศึกษาจัดให้เกิดขึ้นกับนักเรียนด้วยกระบวนการเรียนรู้ที่มีความใกล้เคียงกับหลักการของการวิเคราะห์มูลค่าเพิ่ม ผลการวิเคราะห์ที่แสดงถึงมูลค่าเพิ่มจากการวิเคราะห์ด้วยวิธีนี้คือ ค่าส่วนที่เหลือ (residual term) หรือส่วนต่างระหว่างผลสัมฤทธิ์ที่วัดได้กับค่าที่ประมาณได้จากชุดของปัจจัยหรือตัวแปรที่กำหนดในโมเดลการวิเคราะห์ว่ามีค่าเป็นบวกหรือลบในประมาณมากน้อยเพียงใด การแปลผลว่าสถานศึกษาใ ดได้จัดมูลค่าเพิ่มทางการศึกษาให้กับนักเรียน คือการพิจารณาจากค่าส่วนเหลือที่มีค่าเป็นบวก และถ้าค่าดังกล่าวมีค่าเป็นลบ หรือมีค่าเป็นศูนย์แสดงว่าสถานศึกษาไม่สามารถจัดการศึกษาที่เป็นมูลค่าเพิ่มให้กับนักเรียน

การวิเคราะห์ถดถอยช่วยลดข้อจำกัดในการเปรียบเทียบผลสั มฤทธิ์จากการปฏิบัติระหว่างสถานศึกษาลงได้ แต่ยังมีข้อจำกัดของเทคนิคการวิเคราะห์หลายประการที่ทำให้ได้ผลการวิเคราะห์ที่ผิดพลาด ดังนี้ (1) การละเลยหลักธรรมชาติของสถานศึกษาที่มีลักษณะลดหลั่นสอดแทรกด้วยการวิเคราะห์ข้อมูลจากปัจจัยทุกระดับภายในโมเดลการวิเคราะห์เดี ยวกันซึ่งเป็นการปฏิบัติที่ฝ่าฝืนข้อตกลงเบื้องต้นในเรื่องความเป็นอิสระของข้อมูลจึงเป็นสาเหตุของการเกิดความคลาดเคลื่อนชนิดที่ 1 (type I error) (2) การคัดเลือกสาระของข้อมูลเกี่ยวกับผลผลิตเพื่อนำเสนอต่อสาธารณชน โดยการคัดเลือกเฉพาะปัจจัยหรือตัวแปรระดับบุคคลของนั กเรียนเพียง

บางส่วนเพื่อใช้ในการกระบวนการวิเคราะห์ประมาณค่าผลสัมฤทธิ์ครั้งหลัง และในการเปรียบเทียบจึงอาจทำให้ขาดความชัดเจนในการสรุปผลการปฏิบัติของสถานศึกษา ซึ่งต่างจากการแสดงผลด้วยคะแนนดิบ ดังนั้นการแปลผลอาจทำได้ยาก (3) มีความเป็นไปได้ที่สถานศึกษาซึ่งมีผลสัมฤทธิ์จัดอยู่ในระดับต่ำแต่กลับได้รับการปรับแก้ในลักษณะการถดถอยเข้าหาค่าเฉลี่ยของกลุ่มนักเรียนที่มีความสามารถสูงเนื่องจากกระบวนการวิเคราะห์ได้ทำให้ผลวิเคราะห์มีความคลาดเคลื่อน

3. วิธีวิเคราะห์ด้วยโมเดลพหุระดับ (Multilevel modeling) เป็นการวิเคราะห์ถดถอยอีกวิธีหนึ่งที่คำนึงถึงลักษณะ ข้อมูลแบบลดหลั่นเป็นระดับขั้น (hierarchical) ตามธรรมชาติของข้อมูลที่เกี่ยวข้องกับสถานศึกษา จุดเด่นของวิธีนี้คือปัจจัยที่ส่งผลกระทบต่อความแตกต่างของผลสัมฤทธิ์นั้นไม่ได้เกิดจากปัจจัย ระดับบุคคลเพียงระดับเดียว แต่อาจมีผลที่เกิดจากปัจจัยระดับห้องเรียนและระดับสถานศึกษาด้วย จึงต้องมีการวิเคราะห์เพื่อปรับแก้ด้วยปัจจัยเหล่านั้นเป็นระดับขั้นไป โดยเริ่มจากระดับเล็กสุดคือ ระดับบุคคล ผลการวิเคราะห์ที่แสดงถึงมูลค่าเพิ่มคือ ค่าส่วนที่เหลือ (residual terms) จากโมเดลการวิเคราะห์ระดับสถานศึกษา ซึ่ง แบ่งออกเป็นสองส่วน ได้แก่ ค่าส่วนที่เหลือของค่าจุดตัดหรือค่าเฉลี่ยของกลุ่มสถานศึกษา (intercept residual) และค่าส่วนที่เหลือของค่าความชัน (slope residual) ของกลุ่มตัวแปรภูมิหลังนักเรียนที่กำหนดไว้ในโมเดลการวิเคราะห์ ดังนั้นหากค่าส่วนที่เหลือเหล่านี้มีค่าเป็น นบวก สามารถแปลความหมายได้ว่าสถานศึกษาได้จัดการศึกษาที่ก่อให้เกิดมูลค่าเพิ่มให้กับนักเรียนเมื่อเทียบจากผลสัมฤทธิ์ที่ได้จากการประมาณค่าบนพื้นฐานของปัจจัยหรือตัวแปรที่กำหนดในโมเดลการวิเคราะห์นั้น ในทางตรงข้าม ถ้าส่วนที่เหลือมีค่าเป็นลบหรือศูนย์ แสดงว่าสถานศึกษาไม่สามารถจัดการศึกษาให้เกิดมูลค่าเพิ่มขึ้นได้

จากวิธีการวิเคราะห์ที่มีลักษณะและความซับซ้อนแตกต่างกัน การเลือกใช้วิธีวิเคราะห์สามารถเลือกโดยพิจารณาจากความเข้าใจและความซับซ้อนของหลักทางสถิติซึ่งอาจจะมีข้อจำกัดสำหรับการเลือกใช้โดยทั่วไป นอกจากนี้การพิจารณา ถึงโครงสร้างข้อมูล และตัวแปรในแต่ละระดับของข้อมูล ซึ่งอาจจะสามารถเข้าถึงตัวแปรได้ยาก หรือต้องใช้ระยะเวลาในการเก็บรวบรวมข้อมูลค่อนข้างยาวนาน แต่โดยส่วนมากการวิเคราะห์ที่มีความสอดคล้องกับโครงสร้างข้อมูลทางการวิจัยทางสังคมศาสตร์มากที่สุด จะเป็นการวิเคราะห์พหุ ระดับ เนื่องจากตัวแปรใน

ระดับต่างกันมีความสัมพันธ์กันและสามารถนำมาใช้ในการแปลผลและนำไปใช้ให้เกิดประโยชน์มากที่สุด

2.3 โมเดลมูลค่าเพิ่ม (Value-Added Models)

โมเดลมูลค่าเพิ่มที่รู้จักกันดีเช่น ระบบการประเมิน มูลค่าเพิ่มของรัฐเทนเนสซี (Temmessee Value Added Assessment: TVAAS) (Ballou, Sanders, Wright, 2004) และระบบการประเมินมูลค่าเพิ่มทางการศึกษา (Education Value-Added Assessment System: EVAAS) ซึ่ง EVAAS เป็นกระบวนการทางสถิติสำหรับกลุ่มใหญ่ของการเปลี่ยนคะแนนผลสัมฤทธิ์ของนักเรียนเมื่อเวลาผ่านไป โดยผลจะนำมาใช้กำหนดสิ่งที่เกิดการพัฒนาและประสิทธิผลของเขตพื้นที่ สถานศึกษาและครูเมื่อเวลาผ่านไป (Sanders & Horn, 1994; Sanders, et al., 1998) ความแตกต่างของโมเดลมูลค่าเพิ่มแต่ละโมเดลขึ้นอยู่กับการกำหนดลักษณะเฉพาะของโมเดล ทั้งนี้โมเดลที่เลือกแต่ละโมเดลขึ้นอยู่กับความสามารถในก ารเข้าถึงข้อมูล การลงความเห็นจากข้อค้นพบ ตลอดจนระดับการปฏิบัติ และความสามารถในการเข้าถึงโปรแกรมการวิเคราะห์ที่มีความซับซ้อน กลุ่มของโมเดลการวิเคราะห์มูลค่าเพิ่ม Lissitz, et al (2006) ได้จัดกลุ่มโมเดลการวิเคราะห์โดยพิจารณาจากโมเดลที่มีอิทธิพลคงที่ (fixed effect) ซึ่งเป็นการวัดปัจจัยที่ร่วมกันของสถานศึกษา และอิทธิพลสุ่ม (random effect) ซึ่งเป็นการวัดปัจจัยที่มีการแปรค่าในบุคคลหรือสถานศึกษา (Raudenbush & Bryk, 2002) นอกจากนี้ความซับซ้อนของโมเดลยังพิจารณาจากตัวแปรแทรกซ้อนที่มีผลต่อตัวแปรตาม (covariates) โดยโมเดลที่การวิเคราะห์ที่ค่อนข้างง่ายคือเป็นโมเดลที่มีตัวแปรเดียว (univariate) ไปจนถึงโมเดลที่มีความซับซ้อนที่มีหลายตัวแปร (multivariate)

เนื่องจากข้อจำกัดของการเลือกโมเดลที่มีความเหมาะสม ขึ้นอยู่กับความสามารถที่เข้าถึงได้ของข้อมูล และเครื่องมือวัดผลสัมฤทธิ์ของนั กเรียน เพื่อพิจารณาผลความแตกต่างระหว่างการวัดสองครั้ง ซึ่งอาจจะมีสเกลการวัดที่ต่างกัน ตลอดจนปัจจัยพื้นฐานอื่น ๆ ของนักเรียน ดังนั้น Haegeland and Kirkeboen (2008) ได้จัดแบ่งโมเดลมูลค่าเพิ่ม (Value added models: VAM) กับโมเดลมูลค่าเพิ่มเชิงบริบท (Contextualised value-added modeling: CVA) โดยโมเดลมูลค่าเพิ่มจะพิจารณาตัวแปรที่นำเข้าเฉพาะความรู้เดิมของนักเรียน ส่วนโมเดลมูลค่าเพิ่มเชิงบริบท จะนำตัวแปรบริบทที่เป็นตัวแปรเศรษฐานะเข้าร่วมพิจารณาในโมเดลด้วย ในประเทศ

อังกฤษ โมเดลที่มีความโดดเด่น คือโมเดลมูลค่าเพิ่มเชิ งบริบท (Contextualised value-added modelling) เป็นโมเดลที่มีความซับซ้อน เนื่องจากมีการควบคุมอิทธิพลที่มีอยู่หลากหลายของคุณลักษณะทางด้านเศรษฐานะต่อการเปลี่ยนแปลงความสามารถของนักเรียน และคุณลักษณะหลายอย่างที่มีอิทธิพลต่อความก้าวหน้าของนักเรียนที่อยู่นอกเหนื อการควบคุมของสถานศึกษา เช่น เพศ เชื้อชาติ ศาสนา ความจำเป็นในการศึกษาพิเศษ ภาษาที่ใช้ การย้ายโรงเรียน การดูแลเอาใจใส่ เป็นต้น OECD (2008) ได้มีการจำแนกระหว่างโมเดลมูลค่าเพิ่ม (Value added modeling) และโมเดลความสำเร็จของบริบท (Contextualised attainment models) ซึ่งโดยทั่วไปโมเดลมูลค่าเพิ่มจะมีการวัดเกี่ยวกับความรู้เดิมของนักเรียนที่เป็นพื้นฐานสำหรับการนำมาสู่ความแตกต่างของนักเรียน ส่วนโมเดลความสำเร็จของบริบท (Contextualised attainment models) จะไม่รวมการวัดความรู้เดิมของนักเรียน แต่ปัญหาเกี่ยวกับการวัดสองครั้ งอาจมีปัญหาจากสเกลการวัดที่ต่างกัน ถ้าคะแนนอยู่ในสเกลเดียวกัน จึงจะสามารถคำนวณความแตกต่าง หรือคะแนนที่เพิ่มของแต่ละคน คะแนนที่เพิ่มขึ้นโดยเฉลี่ยของนักเรียน จะพิจารณาได้ว่าเป็นคะแนนมูลค่าเพิ่มของสถานศึกษา (OECD, 2008)

โมเดลการศึกษาภาคตัดขวาง (Cross sectional models) ได้มีการศึกษาเพิ่มมากขึ้นและนำมาใช้กับโครงสร้างลดหลั่นของระบบสถานศึกษา ซึ่งนักเรียนสอดแทรกในชั้นเรียน ชั้นเรียนสอดแทรกในสถานศึกษา และสถานศึกษาสอดแทรกในเขตพื้นที่หรืออำเภอ (Aitkin and Longford, 1986: Goldstein, 1986; Willms and Raudenbush, 1989) การประมาณค่านี้จะได้โดยการวิเคราะห์การศึกษาระหว่างประเทศสมาชิกของ OECD การประมาณค่าภาคตัดขวางจัดเป็นโมเดลความสำเร็จด้านบริบท (Contextualised attainment models) โมเดลหลายตัวแปรสามารถใช้ในการวัดการปฏิบัติงานของสถานศึกษา โมเดลนี้เป็นการประมาณค่าขนาดของการดำเนิ นการของปัจจัยต่าง ๆ ต่อความสามารถของนักเรียนหรือความสำเร็จในช่วงเวลาเดียว ตัวอย่างเช่นการวิเคราะห์โมเดลถดถอยของพื้นฐานทางเศรษฐานะของนักเรียนหรือคุณลักษณะทางด้านบริบทและตัวแปรอื่น ๆ ระบุสถานศึกษาที่มีผลต่อการวัดผลสัมฤทธิ์ของนักเรียน การปรับคะแนนดิบด้วยคุณลั กษณะของบริบทเป็นการวัดที่สะท้อนถึงการดำเนินงานของสถานศึกษาต่อการเรียนรู้ของนักเรียนมากกว่าการใช้คะแนนดิบในการวัดการปฏิบัติงานของสถานศึกษา ผลของโมเดลภาคตัดขวางนี้ขึ้นอยู่กับการวิเคราะห์เชิงทฤษฎีเกี่ยวกับบทบาทของผลลัพธ์ทางเศรษฐานะ

ของผู้ปกครองของนักเรียน (OECD, 2007b; Haveman and Wolfe, 1995) ข้อสารสนเทศเกี่ยวกับพื้นฐานทางเศรษฐานะของนักเรียนต่อความสำเร็จทางการศึกษา และให้ความสำคัญในข้อสารสนเทศที่เพียงพอต่อผู้กำหนดนโยบาย โมเดลความสำเร็จทางด้านบริบท (Contextualised attainment models) เป็นโมเดลที่มีความชัดเจนต่อการใช้ผลที่มาได้ปรับและคะแนนความสำเร็จของคะแนนดิบต่อการประเมินการปฏิบัติงานของสถานศึกษา

การศึกษาเพื่อเปรียบเทียบโมเดลมูลค่าเพิ่ม ได้มีนักการศึกษาสนใจจำนวนมาก เนื่องจากความซับซ้อนของโมเดลมีความแตกต่างกัน รวมถึงตัวแปรที่นำเข้ามาพิจารณาในโมเดลอาจจะมีข้อจำกัดใน การเข้าถึงได้ วัตถุประสงค์ของการศึกษาโดยส่วนมาก เป็นการเปรียบเทียบโมเดลเพื่อเสนอแนวทางที่หลากหลายในการเลือกโมเดลแต่ละโมเดลตามความเหมาะสม ดังตัวอย่างดังนี้

การศึกษาของ TCMLRAFR (Tekwe, Carter, Ma, Algina, Lucas, Roth, Ariet, Fisher, and Resnick, 2004) ได้ศึกษาโมเดลมูลค่าเพิ่มที่แตกต่างกัน 4 โมเดล คือ 1) โมเดลอิทธิพลคงที่ (Fixed-effects models: FEM) ซึ่งจัดเป็นอิทธิพลสถานศึกษา (school effect) เช่นการพัฒนาผลสัมฤทธิ์นักเรียนเนื่องจากอิทธิพลของครูหรือสถานศึกษา เป็นโมเดลที่ง่ายที่สุด มีการคำนวณซับซ้อนเล็กน้อยและไ ม่ต้องใช้ข้อสารสนเทศเกี่ยวกับปัจจัยแทรกซ้อน (covariates) 2) โมเดลอิทธิพลผสมผสานแบบเป็นชั้น (Layered mixed effects model: LMEM) โมเดลนี้มีลักษณะเดียวกับ TVAAS โดย Sanders และ Horn เป็นการใช้ข้อสารสนเทศในความแปรปรวนที่ไม่เป็นศูนย์ (non-zero covariance) ระหว่างคะแนนสอบที่เวลาต่างกัน เป็นโมเดลสำหรับการเปลี่ยนแปลงคะแนนด้วยอิทธิพลสถานศึกษาแบบสุ่มที่ไม่พยายามที่จะกำจัดอิทธิพลของตัวแปรแทรกซ้อนเข้าไปทั้งระดับ สถานศึกษาและระดับเขตพื้นที่ วิธีการนี้พยายามที่จะกำจัดอิทธิพลของตัวแปรแทรกซ้อน โมเดลนี้จะมีข้อดีเพิ่มขึ้นเกี่ ยวกับการรวมข้อสารสนเทศจากข้อมูล และสามารถประเมินพร้อม ๆ กันได้ของผลที่มาจากหลายเนื้อหา 3) โมเดลเชิงเส้นแบบลดหลั่น (Hierarchical linear models: HLMM) ซึ่งกำหนดอิทธิพลสถานศึกษาเป็นแบบสุ่ม เป็นโมเดลที่ดำเนินการอิทธิพลย่อส่วน (shrunken effects) มีอยู่สองลักษณะคือ โมเดลที่ไม่ได้ปรับแก้อย่างง่าย (Unadjusted HLMM: UHLMM) เป็นโมเดลที่ไม่ได้นำตัวแปรที่เป็นองค์ประกอบเข้าพิจารณาหรือไม่มีตัวแปรแทรกซ้อนระดับนักเรียน และโมเดลที่

มีการปรับแก้ (Adjusted HLMM: AHLMM) โดยนำตัวแปรแทรกซ้อนทั้งในระดับนักเรียนและระดับสถานศึกษาเข้ าอธิบายคะแนนที่เปลี่ยนแปลง ผลจากการศึกษาของ TCMLRAFR โดยใช้ข้อมูลจำลองแสดงให้เห็นว่ามีค่าสหสัมพันธ์กันสูงระหว่างผลของ SFEM, LMEM, และ UHLMM แต่มีค่าสหสัมพันธ์ปานกลางระหว่างโมเดล AHLMM กับโมเดลอื่น การศึกษาของ TCMLRAFR สรุปว่าผลจากการวิเคราะห์ใน SFEM ได้ผลดีเช่นเดียวกันกับสองโมเดลที่ไม่ได้รวมตัวแปรที่เป็นองค์ประกอบที่เป็นตัวแปรแทรกซ้อนระดับนักเรียน และสามารถคาดหวังได้ผลเช่นเดียวกันแต่ที่มีค่าใช้จ่ายในการคำนวณต่ำกว่า ความแตกต่างระหว่าง AHLMM และโมเดลอื่นชี้ให้เห็นว่าเมื่อนำตัวแปรแทรกซ้อนระดับนักเรียนและตัวแป รองค์ประกอบใช้วิเคราะห์ในโมเดลด้วยการประมาณค่าจะเกิดการเปลี่ยนแปลง ถ้าตัวแปรแทรกซ้อนมีผลกระทบต่อการเรียนรู้ของนักเรียนก็สามารถอธิบายได้ว่าวิธีการของ AHLMM มีการประมาณค่าที่ถูกต้องแม่นยำกว่าโมเดลอื่น เพราะมีการรวมข้อสารสนเทศและประมาณค่าได้ถูกต้องแม่นยำกว่า แต่โมเดล SFEM ก็จัดเป็นโมเดลทางเลือกได้ เพราะมีวิธีการที่คำนวณอย่างง่าย

การศึกษาของ BSW (Ballou, Sanders, Wright, 2004) เป็นการประมาณ TVAAS model และเป็นโมเดลที่มีความเท่าเทียมกับโมเดล LMEM ของการศึกษาของ TCMLRAFR ผลพบว่าการนำปัจจัยด้านบริบทในโมเดล HLM ชี้ให้เห็นว่าได้ผลลัพธ์ที่มีความไวต่ออิทธิพลและชี้ให้เห็นว่า TVAAS สามารถรวมปัจจัยด้านบริบทได้ด้วยการรวมปัจจัยเหล่านี้จะมีแนวโน้มที่ทำให้ความลำเอียงจากการวัดของอิทธิพลของครูและนักเรียนมีค่าเข้าใกล้ศูนย์ การใช้ข้อมูลจากฐานข้อมูลจำนวนมากจากการศึกษา TVAAS และ BSW ที่เป็นการศึกษาข้อมูลจำลองเพื่อกำหนดว่าขนาดอิทธิพลของครูมีการเปลี่ยนแปลงเท่าใดถ้าอิทธิพลของนักเรียนและสถานศึกษานำเข้าไปวิเคราะห์ในโมเดล และข้อสรุปว่าตัวแปรแทรกซ้อนระดับนักเรียน มีอิทธิพลเพียงเล็กน้อยต่อคะแนนผลการปฏิบัติของครูและจากการศึกษาของ BSW ได้เสนอ ผลของการศึกษาจากข้อมูลจำลองว่าผลกระทบของตัวแปรนักเรียนมีค่าไม่มากพอที่เห็นความแตกต่างในการประมาณค่าอิทธิพลของครู และค่าสหสัมพันธ์ที่สูงระหว่างอิทธิพลที่มีการปรับแก้และไม่มีการปรับแก้เป็นสาเหตุเนื่องจากการย่อส่วน (shrinkage)

Haegeland and Kirkeboen (2008) ได้ศึกษาเกี่ยวกับการใช้ตัวบ่งชี้ในโมเดลมูลค่าเพิ่มที่ต่างกันในการปรับแก้การปฏิบัติงานของสถานศึกษา มี 2 แนวทางคือ การควบคุมความรู้เดิม

หรือความสำเร็จเดิมของนักเรียน และการควบคุมเศรษฐานะของนักเรียน (SES) และใช้โมเดลที่ต่างกัน 4 โมเดลคือ โมเดลที่ไม่ได้ปรับแก้ (Unadjusted results: UR) ซึ่งพิจารณาเฉพาะตัวแปรสถานศึกษา โมเดลความสำเร็จของบริบท (Contextualised attainment models: CAM) ซึ่งประกอบด้วยตัวแปรความรู้เดิม และตัวแปรสถานศึกษา และโมเดลมูลค่าเพิ่มเชิงบริบท (Contextualised value added models: CVA) ประกอบด้วยตัวแปร SES ตัวแปรความรู้เดิม และตัวแปรสถานศึกษา การศึกษานี้จะเปรียบเทียบอิทธิพลของสถานศึกษาในโมเดลที่แตกต่างกัน ซึ่งพิจารณาผลของชุดตัวแปร SES มีผลต่อการประมาณค่า อิทธิพลของสถานศึกษาอย่างไรเมื่อรวมตัวแปร SES เข้าในโมเดล CAM และ CVA ผลจากการประมาณค่าของโมเดล SES ทำให้เกิดความแตกต่างกันของผลการปฏิบัติงานของสถานศึกษา และการใช้จำนวนตัวแปร SES มีผลต่อการประมาณค่าของลักษณะที่แตกต่างกันเมื่อเพิ่มตัวแปร SES เข้าไปในเงื่อนไขของโมเดล CAM จะทำให้การกระจายของตัวบ่งชี้ผลการปฏิบัติของสถานศึกษาแคบลง นั่นคือตัวแปร SES มีความสัมพันธ์กับความสำเร็จแต่ละคน ดังนั้นการศึกษานี้จึงแสดงให้เห็นว่าตัวบ่งชี้การปฏิบัติของสถานศึกษาอาจจะเกิดผลกระทบเนื่องจากความแตกต่างของ SES และความรู้เดิมของนักเรียน ข้อค้นพบในครั้งนี้ชี้ให้เห็นว่าโมเดลความสำเร็จของบริบท (contextualized attainment model) ซึ่งเป็นโมเดลที่ปราศจากตัวแปรความรู้เดิมของนักเรียน และรวม SES ในโมเดลอาจจะเกิดผลกระทบอย่างมากต่อตัวแปรการปฏิบัติงานของสถานศึกษา การรวมตัวแปร SES อาจจะเป็นผลกระทบอย่างมากต่อการปฏิบัติงานของสถานศึกษา

นอกจากนี้การศึกษาของ Thomas, (1996) มีการศึกษาเปรียบเทียบโมเ ดลมูลค่าเพิ่มของประสิทธิผลในสถานศึกษาระดับมัธยมศึกษาด้วยการวิเคราะห์พหุระดับในโมเดลต่างกันที่ควบคุมความแตกต่างของปัจจัยพื้นฐานของนักเรียนและสถานศึกษา พบว่าความผันแปรมีค่าอย่างมากในระดับสถานศึกษา และการเปรียบเทียบระหว่างโมเดลที่แตกต่างของโมเดลมูลค่าเพิ่มและปัจจัยระดับสถานศึกษาเป็นสิ่งที่ไม่มีนัยสำคัญในการทำนายผลลัพธ์ของนักเรียน นอกจากนี้ยังพบว่าถึงแม้ว่าจะขาดตัวแปรความรู้เดิมของนักเรียน ปัจจัยทางด้านบริบทก็เป็นสิ่งที่มีประโยชน์อย่างมากและอาจจะมีความเพียงพอและเป็นการวัดที่เหมาะสมในการนำเข้าไปพิจารณาเนื่องจากสถานศึกษาแต่ละแห่งมีความไม่คงที่ของประสิทธิผลของหน่วยงาน

เมื่อพิจารณาผลการศึกษาเกี่ยวกับโมเดลมูลค่าเพิ่มที่มีอยู่หลากหลาย สามารถสรุปผลการศึกษาได้ดังนี้ 1) ไม่มีโมเดลใดที่แสดงให้เห็นว่ามีความถูกต้องและเพียงพอในการนำมาใช้ในสภาพแวดล้อมที่มีการได้เปรียบเสียเป รียบกันสูง 2) การประมาณค่าในโมเดลที่ไม่ได้รวมตัวแปรแทรกซ้อนระดับนักเรียนและสถานศึกษา จะสามารถเชื่อถือได้เฉพาะตำแหน่งที่อยู่ปลายสุดการจัดอันดับของสถานศึกษา และจะมีความลำเอียงให้กับสถานศึกษาในกลุ่มประชากรที่ได้เปรียบกว่า เช่น ในผลการศึกษาของ TVAAS 3) โมเดลเชิงเส้นลดหลั่น Hierarchical linear models ที่มีการรวมตัวแปรแทรกซ้อน เช่น โมเดล AHLMM ของการศึกษา TCMLRAFR จะมีการประมาณค่าของอิทธิพลของครูและสถานศึกษาที่มีความถูกต้องมากกว่าและมีความลำเอียงน้อยกว่า แต่ยังมีข้อจำกัดเกี่ยวกับความซับซ้อนของการคำนวณ 4) โมเดลอย่างง่ายเช่น โมเดล FEM ของการศึกษา TCMLRAFR ถึงแม้จะมีโอกาสเกิดความลำเอียงมาก แต่ก็เป็นปัญหาโดยทั่วไปของการประมาณค่า 5) ถ้าไม่ได้รวมตัวแปรแทรกซ้อนเข้าไปในโมเดล เช่น โมเดลอย่างง่าย (SFEM) อาจจะมีประสิทธิภาพเท่าเทียมกับโมเดลที่มีความซับซ้อนอย่าง LMEM

สิ่งที่พิจารณาการกำหนดโมเดล สรุปได้ว่ามีประเด็นที่น่าสนใจประเด็นแรกเป็นความซับซ้อนในการคำนวณหรือการประมาณค่าในโมเดล ประเด็นที่สองคือการให้ความสำคัญของตัวแปรแทรกซ้อนที่อาจจะส่งผลตัวแปรตามร่วมกับตัวแปรต้น ประเด็นที่สามคือตัวแปรที่นำเข้ามาในโมเดลซึ่งการใช้ตัวแปรที่เป็นตัวบ่งชี้ของการปฏิบัติงานของสถานศึกษาอาจจะมีความแตกต่างกันในแต่ละบริบท และอาจจะมีข้อจำกัดในการเข้าถึงข้อมูลได้ ดังนั้นการเลือกโมเดลมูลค่าเพิ่มเพื่อใช้ในการประเมินคุณภาพการจัดการของสถานศึกษาจึงควรระวังในการเลือกโมเดลที่เหมาะสมเนื่องจากสภาพบริบทที่แตกต่างกัน

ตอนที่ 3 การวิเคราะห์พหุระดับ

หัวใจสำคัญของการวิเคราะห์ข้อมูลนั้น คือ การเลือกใช้สถิติให้เหมาะสมกับข้อคำถามการวิจัย จึงจะสามารถมีผลการวิจัยที่มีความน่าเชื่อถือ การวิเคราะห์พหุระดับเป็นอีกวิธีการหนึ่งที่ช่วยในการวิเคราะห์ข้อมูลทางสถิติขั้นสูง โดยมีรายละเอียดดังนี้

3.1 แนวคิดเกี่ยวกับการวิเคราะห์พหุระดับ

โครงสร้างและธรรมชาติของข้อมูลการวิจัยทางสังคมศาสตร์ส่วนใหญ่มีลักษณะเป็นพหุระดับ ในการวิเคราะห์ข้อมูลที่ให้ผลการวิเคราะห์ที่ถูกต้องจึงต้องคำนึงถึงระดับข้อมูล ซึ่งการวิเคราะห์ที่เหมาะสมสำหรับข้อมูลดังก ล่าวคือการวิเคราะห์พหุระดับ โดยผู้วิจัยได้นำเสนอแนวคิดการวิเคราะห์พหุระดับ และการวิเคราะห์พหุระดับด้วยโมเดลเชิงเส้นตรงระดับลดหลั่น การวิจัยทางสังคมศาสตร์ส่วนใหญ่เกี่ยวข้องกับโครงสร้างของข้อมูลที่เป็นพหุระดับ เนื่องจากโครงสร้างและธรรมชาติของข้อมูลที่ศึกษามี ความเกี่ยวข้องกับข้อมูลหลายระดับ โดยมีความสัมพันธ์ในลักษณะต่าง ๆ (ศิริชัย กาญจนวาสี, 2548) ดังนี้

1. การจัดองค์กรและข้อมูลขององค์กร

การจัดองค์กรโดยทั่วไปและการแบ่งส่วนการบริหารงานภายในองค์กรมีลักษณะการจัดตั้งเป็นลำดับขั้น (hierarchy) หน่วยงานหรือการบริหา รระดับสูงย่อมต้องมีความรับผิดชอบหรือมีอิทธิพลต่อหน่วยงาน หรือการบริหารงานในระดับรองลงมาเป็นลำดับขั้น การจัดองค์กรและการดำเนินงานทางการศึกษาก็เช่นกัน ตัวแปรในระดับบนย่อมส่งผลกระทบต่อตัวแปรในระดับล่างเป็นทอด ๆ ไป เช่น นโยบายของกระทรวงศึกษาธิการย่อมส่งผลต่อกา รบริหารการศึกษาระดับเขตการศึกษา – การบริหารการศึกษาระดับโรงเรียน – การเรียนการสอนในชั้นเรียน – ผลสัมฤทธิ์ทางการเรียนของนักเรียน เป็นต้น สมมติว่าผู้วิจัยสนใจศึกษาผลสัมฤทธิ์ทางการเรียนของนักเรียน สิ่งนี้ย่อมมีความเกี่ยวข้องกับตัวแปรภูมิหลังของนักเรียน ตัวแปร เกี่ยวกับการเรียนการสอนระดับชั้นเรียน ตัวแปรเกี่ยวกับนโยบายระดับโรงเรียน และตัวแปรเกี่ยวกับสภาพแวดล้อมของชุมชนอันเป็นที่ตั้งของโรงเรียน เป็นต้น

2. ปฏิสัมพันธ์ระหว่างตัวแปรภายในองค์กร

เนื่องจากการจัดองค์กรและธรรมชาติของข้อมูลในองค์กรมีลักษณะเป็นพหุระดับมีค วามสลับซับซ้อนและไม่หยุดนิ่ง (Dynamic) ดังนั้นตัวแปรต่าง ๆ ที่อยู่ในระดับเดียวกันและต่างระดับกันจึงมีปฏิสัมพันธ์กัน (interaction) ซึ่งกันและกันตลอดเวลา การศึกษาเพื่อให้เกิดความเข้าใจในกระบวนการภายในองค์กรอย่างชัดเจนลึกซึ้ง จึงต้องทำความเข้าใจกับความสัมพันธ์ ระหว่างตัวแปรในระดับเดียวกันและปฏิสัมพันธ์ระหว่างตัวแปรต่างระดับ

3. การรวมกลุ่มของสมาชิกในองค์กร

สมาชิกในองค์กรมักรวมกันเป็นหน่วย แผนก ฝ่าย โดยมีจุดมุ่งหมายใดจุดมุ่งหมายหนึ่ง หน่วยทางการศึกษาอาจเป็นชั้นเรียน โรงเรียน กลุ่มโรงเรียน โรงเรียนในเขตพื้นที่การศึกษา เป็นต้น ดังนั้นหน่วยทางการศึกษา จึงประกอบด้วยสมาชิก (ผู้เรียน ผู้สอน ผู้บริหาร ฯลฯ) มาอยู่รวมกันด้วยจุดมุ่งหมายเฉพาะอย่างไม่ได้มาอยู่รวมกันอย่างสุ่ม (Random) ดังนั้น การศึกษาเปรียบเทียบระหว่างหน่วยทางการศึกษาจะต้องมั่นใจได้ว่ามีการควบคุมความแตกต่างเ บื้องต้นของหน่วย (pre-existing difference) และคัดเลือกตัวแปรในระดับต่าง ๆ ที่เกี่ยวข้องได้อย่างครอบคลุม

ข้อมูลทางสังคมศาสตร์มักมีโครงสร้างของข้อมูลเป็นระดับที่สอดแทรกลดหลั่น เช่น ข้อมูลเกี่ยวกับประชาชน ซึ่งรวมตัวกันอยู่ในตำบล ตำบลรวมอยู่ในอำเภอ อำเภอรวมอยู่ ในจังหวัด เป็นต้น ดังนั้นการวิเคราะห์ข้อมูลของตัวแปรจึงควรทราบถึงความแปรปรวนของค่าตัวแปรในระดับต่าง ๆ ว่ามีส่วนประกอบย่อยแต่ละส่วนแตกต่างกันอย่างไรตามระดับข้อมูล การวิเคราะห์ความแปรปรวน (Variance component analysis) เป็นเทคนิคการวิเคราะห์ทางสถิติอย่างหนึ่งที่จะช่วยตอบคำถามว่า “ความแปรปรวนของค่าในตัวแปรมีส่วนประกอบย่อย ๆ แต่ละส่วนแตกต่างกันอย่างไรตามระดับข้อมูล” (ศิริชัย กาญจนวาสี, 2548)

วิธีประมาณค่าความแปรปรวน สามารถกระทำได้ด้วยวิธีหลัก 3 วิธี (Corbeil and Searle, 1976 อ้างถึงใน ศิริชัย กาญจนวาสี, 2548) คือ

1. การใช้วิธีการวิเคราะห์ความแปรปรวน (ANOVA) โดยการเลือกโมเดลวิเคราะห์ที่เหมาะสมกับข้อมูล เพื่อคำนวณค่าคาดหมายของกำลังสองเฉลี่ย (Expected mean square) ของแต่ละแหล่งความแปรปรวนจากนั้นจึงหาค่าความแปรปรวนแต่ละส่วนที่ต้องการ

2. การใช้วิธีความเป็นไปได้สูงสุด (Maximum likelihood estimation, ML) เพื่อประมาณค่าส่วนประกอบความแปรปรวนแต่ละส่วนที่มีความเป็นไปได้สูงสุด ตามข้อมูลเชิงประจักษ์ที่เก็บรวบรวมมาได้

3. การใช้วิธีกำลังสองที่ไม่ลำเอียงต่ำสุด (Minimum norm quadratic unbiases estimation, MNQUE) เพื่อประมาณค่าส่วนที่ให้ค่าความคลาดเคลื่อนกำลังสองต่ำสุด

เนื่องจากข้อมูลทางสังคมศาสตร์มักมีโครงสร้างของข้อมูลเป็นพหุระดับที่สอดแทรกลดหลั่น การวิเคราะห์ข้อมูลแบบประเพณีนิยมจึงไม่สามารถให้ผลสรุปได้อย่างถูกต้อง เนื่องจาก

1. การวิเคราะห์ข้อมูลแบบประเพณีนิยมเป็นการวิเคราะห์ข้อมูลระดับเดียว (Single level approach) เช่น การวิเคราะห์การถดถอยแบบสมการเดียว หากนำมาใช้ศึกษาอิทธิพลของตัวแปรจากข้อมูลต่างระดับ จะทำให้ละเลยต่อโครงสร้างของระดับข้อมูล ผลที่ตามมาก็คือเกิดความผิดพลาดในการสรุปผลระหว่างระดับ (Aggregation bias) เนื่องจากมีคว ามผิดพลาดในการประมาณค่าสัมประสิทธิ์การทำนาย ตลอดจนความคลาดเคลื่อนของการทำนาย (Raudenbush and Bryk, 1986; Kanjanawasee, 1989 อ้างถึงใน ศิริชัย กาญจนวาสี, 2548)

2. โมเดลการวิเคราะห์ข้อมูลแบบประเพณีนิยมเป็นการวิเคราะห์ที่บีบบังคับให้ตัวแปรต่างระดับให้เสมือนอ ยู่ในระดับเดียวกัน และทำการวิเคราะห์รวมโดยไม่สนใจความแต กต่างระหว่างหน่วยของการวิเคราะห์ วิธีดังกล่าวจึงไม่สามารถคำนวณค่าความแปรปรวนภายในหน่วยหรือกลุ่ม (Within group variability) จึงเป็นการละเลยต่อการศึกษาปฏิสัมพันธ์ระหว่างตัวแปรที่อยู่ต่างระดับ (Raudenbush and Bryk, 1989; Kanjanawasee, 1989 อ้างถึงใน ศิริชัย กาญจนวาสี, 2548)

3. โมเดลการวิเคราะห์ข้อมูลแบบประเพณีนิยมใช้วิธีการควบคุมความแตกต่างเบื้องต้นระหว่างหน่วยที่ศึกษาด้วยค่าคงที่เหมือนกันทุกหน่วย จึงค่อนข้างที่จะขัดแย้งกับแนวคิดที่ว่า หน่วยทางการศึกษาแต่ ละแห่งน่าจะมีธรรมชาติและลักษณะของการรวมกลุ่มที่แตกต่างกัน (Kanjanawasee, 1989 อ้างถึงใน ศิริชัย กาญจนวาสี, 2548)

จากปัญหาของการวิเคราะห์ข้อมูลแบบประเพณีนิยมซึ่งเป็นการวิเคราะห์ข้อมูลระดับเดียวดังกล่าว ทำให้นักวิธีวิทยาการวิจัยหลายท่านได้เสนอเทคนิควิธีการออ กแบบการวิจัยและการวิเคราะห์พหุระดับ (Multi-Level Analysis) ซึ่งเหมาะสมกับโครงสร้างและธรรมชาติของข้อมูลทางการศึกษา เนื่องจากการวิเคราะห์พหุระดับเป็นเทคนิควิธีทางสถิติสำหรับใช้วิเคราะห์ข้อมูลที่มีตัวแปรอิสระหลายตัวและตัวแปรอิสระเหล่านั้น สามารถจัดเป็นระดับได้ อย่างน้อย 2 ระดับขึ้นไป โดยตัวแปรระดับเดียวกันต่างมีความสัมพันธ์ซึ่งกันและกัน และได้ผลร่วมกันจากตัวแปรระดับอื่น ๆ (ศิริชัย กาญจนวาสี , 2548)

จากที่กล่าวมา สรุปได้ว่าข้อมูลทางสังคมศาสตร์มักมีโครงสร้างของข้อมูลเป็นระดับที่สอดแทรกลดหลั่น การวิเคราะห์ข้อมูลของ ตัวแปรจึงควรทราบถึงความแปรปรวนของค่าตัวแปรใน

ระดับต่าง ๆ ว่ามีส่วนประกอบย่อยแต่ละส่วนแตกต่างกันอย่างไรตามระดับข้อมูล แต่การวิเคราะห์ข้อมูลแบบประเพณีนิยมซึ่งเป็นการวิเคราะห์ข้อมูลระดับเดียวไม่สามารถให้ผลสรุปได้อย่างถูกต้อง ดังนั้นในการวิเคราะห์ความแปรปรวนในแ ต่ละระดับข้อมูลจึงควรใช้การวิเคราะห์พหุระดับ (Multi-level analysis) ซึ่งเป็นการวิเคราะห์ข้อมูลที่มีความสอดคล้องกับลักษณะโครงสร้าง และธรรมชาติของข้อมูล อีกทั้งยังสามารถสรุปผลได้อย่างถูกต้อง

- **การวิเคราะห์พหุระดับด้วยโมเดลเชิงเส้นตรงระดับลดหลั่น**

การวิเคราะห์ส่วนประกอบความแปรปรวน (Variance component analysis) ของตัวแปรที่สนใจ จะทำให้ทราบว่าความแปรปรวนของตัวแปรที่สนใจนั้นสามารถแยกส่วนเป็นความแปรปรวนภายในหน่วย ตามระดับชั้นของโครงสร้างที่ลดหลั่นกันมากน้อยเพียงใด การทราบธรรมชาติความผันแปรของตัวแปรต้น และตัวแปรตามที่ สนใจ จะเป็นประโยชน์อย่างยิ่งต่อการศึกษาความสัมพันธ์ระหว่างตัวแปร

การสร้างโมเดลส มการทำนายตัวแปรตามที่มีประสิทธิภาพ โดยเฉพาะอย่างยิ่งการวิเคราะห์สมการทำนายแบบพหุระดับจะนำไปสู่ข้อค้นพบหรือสารสนเทศที่มีความถูกต้อง (Multi-level analysis) เป็นเทคนิคทางสถิติที่ใช้ในการวิเคราะห์อิทธิพลของตัวแปรทำนายหลายระดับที่มีต่อตัวแปรตาม ซึ่งตัวแปรทำนายมีโครงสร้างเป็นระดับลดหลั่น (Hierarchical) อย่างน้อย 2 ระดับ โดยตัวแปรทำนายและตัวแปรตามที่อยู่ระดับล่างต่างมีความสัมพันธ์ซึ่งกันและกัน และได้รับอิทธิพลร่วมกันจากตัวแปรทำนายที่อยู่ระดับบน (ศิริชัย กาญจนวาสี, 2548)

การวิเคราะห์ประมาณค่าส่วนประกอบความแปรปรวน (Analysis of variance component Estimation) ในการวิเคราะห์ข้อมูลพหุระดับนั้นนงลักษณ์ วิรัชชัย (2535 อ้างถึงใน ศิริชัย กาญจนวาสี , 2548) กล่าวไว้ว่าตัวแปรที่วัดได้ในระดับนักเรียนมี ความแปรปรวน ซึ่งแยกส่วนประกอบได้ตามระดับที่ลดหลั่นกัน เช่น กรณีมีสามระดับ คือระดับนักเรียน ระดับห้องเรียน และระดับโรงเรียน และแสดงส่วนประกอบความแปรปรวนได้ดังนี้

$$\sigma^2_y = \sigma^2_{pupil} + \sigma^2_{class} + \sigma^2_{school}$$

เมื่อ σ^2_y แทน ความแปรปรวนของตัวแปรตามที่สนใจศึกษา

σ^2_{pupil} แทน ความแปรปรวนระหว่างนักเรียนภายในโรงเรียน

σ^2_{class} แทน ความแปรปรวนระหว่างห้องเรียนภายในโรงเรียน

σ^2_{school} แทน ความแปรปรวนระหว่างโรงเรียน

- **การประมาณค่าพารามิเตอร์ในโมเดลพหุระดับ**

การประมาณค่าพารามิเตอร์ในโมเดลพหุระดับสามารถกระทำได้หลายวิธี โดยวิธีที่สำคัญ 2 วิธี ได้แก่ วิธีกำลังสองน้อยที่สุด (OLS) และวิธีตามโมเดล HLM (Hierarchical Linear Approach) (ศิริชัย กาญจนวาสี, 2548)

1. OLS (Ordinary Least Square) approach

เป็นการใช้เทคนิคกำลังสองน้อยที่สุดแบบธรรมดา ซึ่งนิยมใช้กันทั่วไปในโปรแกรมคอมพิวเตอร์มาตรฐานแบบประเพณีนิยมทั้งหลายที่มีอยู่ในปัจจุบัน แต่มีข้อเสียในด้านความเหมาะสมของโมเดลที่ใช้วิเคราะห์และผลการวิเคราะห์ที่ได้รับ ตลอดจนมีความยุ่งยากในการเตรียมข้อมูลสำหรับการวิเคราะห์

2. โมเดล HLM (Hierarchical Linear Model) approach

เป็นการใช้เทคนิคการวิเคราะห์ความแปรปรวนแบบผสม ใช้หลักการสัมประสิทธิ์แบบสุ่มและการประมาณค่าพารามิเตอร์ด้วยวิธีของเบยส์ (Bayesian estimation) เทคนิค HLM พัฒนาโดย Raudenbush and Bryk ที่จะให้ผลการวิเคราะห์พหุระดับที่มีความคงเส้นคงวาและน่าเชื่อถือมากกว่าวิธี OLS (Raudenbush and Bryk, 1986 : Kanjanawasee, 1989 อ้างถึงใน ศิริชัย กาญจนวาสี , 2548)

- **จุดเด่นของเทคนิควิธีของ HLM**

เทคนิควิธีของ HLM มีจุดเด่นมีสำคัญ ดังนี้

1) สามารถตรวจสอบความเหมาะสมของโมเดล (Accuracy of a model) เมื่อพิจารณาถึงโมเดล 2 ระดับ ได้แก่

- Within – Group Model

$$Y_{ij} = B0_j + B1_j\left(X_{ij}\right) + R_{ij}$$

- Between – Group Model

$$B0_j = G00+G01(Z_j)+U0_j$$

$$B1_j = G10+G11(Z_j)+U1_j$$

โมเดลดังกล่าวสามารถมีตัวทำนายระดับบุคคล (X_{ij}) และตัวแปรทำนายระดับกลุ่ม (Z_j) ได้หลายตัว และการทำนายตัวแปรตามระหว่างหน่วยไม่จำเป็นต้องใช้ Z_j ชุดเดียวกันก็ได้ จากโมเดล 2 ระดับดังกล่าว HLM สามารถตรวจสอบความเหมาะสมของโมเดลได้ ดังนี้

1.1) ตัวแปร X_{ij} ส่งผลอย่างมีนัยสำคัญทางสถิติต่อ Y_{ij} หรือไม่ ?

HLM จะคำนวณผลเฉลี่ยของ X_{ij} ที่มีต่อ Y_{ij} จากทุกกลุ่มและทดสอบนัยสำคัญทางสถิติ

1.2) ค่าพารามิเตอร์ของแต่ละหน่วยมีความผันแปรระหว่างหน่วยหรือไม่?

HLM จะแบ่งผลของพารามิเตอร์ของแต่ละกลุ่มออกเป็นอิทธิพลคงที่ (Fixed effect) และอิทธิพลสุ่ม (Random effect) ดังนี้

$$B0_j = G00+U0_j$$

$$B1_j = G10+U1_j$$

(Fixed effect) (Random effect)

(ค่าเฉลี่ย) (ค่าความคลาดเคลื่อน)

HLM จะทำการทดสอบค่าเฉลี่ยของ Fixed effect ว่าเป็น 0 หรือไม่โดยใช้สถิติการทดสอบที (T-test)

$$H_0: G00=0$$

$$H_1: G00 \neq 0$$

และ $H_0: G10=0$

$$H_1: G10 \neq 0$$

ถ้า Fixed effect ไม่เป็น 0 แสดงว่า ค่าคงที่ (Intercept) มีผลต่อ Y_{ij} หรือ $B1_j$ แต่ถ้ามีค่าเป็น 0 แสดงว่า ค่าคงที่ (Intercept) ดังกล่าวไม่มีผลต่อ Y_{ij} หรือ $B1_j$ นอกจากนี้ HLM ยังทำการทดสอบความแปรปรวนของ Random effect หรือ parameter Variance ว่าเป็น 0 หรือไม่? โดยใช้สถิติทดสอบไคสแควร์ $(\chi^2\text{-test})$

$$H_0: Var(B0_j) \text{ or } Var(U0_j)=0$$

$$H_1: Var(B0_j) \text{ or } Var(U0_j)>0$$

และ $$H_0: Var(B1_j) \text{ or } Var(U1_j)=0$$

$$H_1: Var(B1_j) \text{ or } Var(U1_j)>0$$

ถ้า Random effect ไม่เป็น 0 แสดงว่า พารามิเตอร์ $(B0_j)$ และ $(B1_j)$ มีความผันแปรระหว่างกลุ่ม แต่ถ้ามีค่าเป็น 0 แสดงว่า พารามิเตอร์ดังกล่าวไม่มีความผันแปรระหว่างกลุ่ม ซึ่งสามารถตั้งข้อจำกัดให้เป็นค่าคงที่ ในการวิเคราะห์ได้ในกรณีที่ค่าพารามิเตอร์ของแต่ละกลุ่มมีความผันแปรระหว่างกลุ่ม จึงสมเหตุสมผลที่จะหาตัวแปรทำนายระหว่างกลุ่มมาอธิบายความผันแปรดังกล่าว

2) สามารถประมาณค่าพารามิเตอร์ได้อย่างคงเส้นคงวาและน่าเชื่อถือ

จากสมการระดับที่ 1 (Within – Group Model) และระดับที่ 2 (Between – Group Model) การแจกแจงของค่าความคลาดเคลื่อนหรือค่าส่วนที่เหลือ เป็นดังนี้

$$R_{ij} \sim N(0, \tau^2)$$

$$U0_j \sim N(0, \tau_0)$$

$$U1_j \sim N(0, \tau_1)$$

<u>ตัวอย่าง</u>การประมาณค่า $B1_j = G10 + U1_j$

- ถ้าใช้เทคนิค OLS จะประมาณค่าได้ $B1_j'$ ซึ่งเป็นค่าความชันที่สังเกตได้ (Observed slope)

โดย $B1_j' = B1_j + E_j$

เมื่อ $E_j \sim N(0, V_j)$

ดังนั้น $Var(B1_j') = Var(B1_j) + Var(E_j)$

$= \tau_1 + V_j$

ทำให้สามารถประมาณค่าความเที่ยง (Reliability) (W_j) ได้

$$W_j = \frac{\tau_1}{\tau_1 + V_j}$$

- HLM จึงใช้วิธีการของเบส์มาช่วยในการประมาณค่าพารามิเตอร์ดังนี้

$$G10^* = \frac{\sum W_j B1_j}{\sum W_j}$$

$$B1_j^* = B1_j + (1 - W_j)G10^*$$

จากที่กล่าวมา สรุปได้ว่าวิธีกา รประมาณค่าพารามิเตอร์โดยใช้วิธีตามโมเดล HLM (Hierarchical Linear Approach) ให้ค่าพารามิเตอร์ที่มีความน่าเชื่อถือมากกว่าวิธีกำลังสองน้อยที่สุด (OLS)

3.2 การวิเคราะห์พหุระดับด้วยโมเดลเชิงเส้นตรงระดับลดหลั่น

สำหรับการวิเคราะห์พหุระดับด้วยโมเดลเชิงเส้นตรงระดับลดหลั่นที่ใช้ในการวิเคราะห์ข้อมูลคุณภาพของหลักสูตรเพื่อการจัดอันดับและระดับแบบลดหลั่น (Hierarchical ranking and rating) นั้น มีโครงสร้างของข้อมูลเป็นระดับที่สอดแทรกลดหลั่น โดยหลักสูตรแทรกอยู่ใน (nested) คณะ และคณะแทรกอยู่ใน (nested) มหาวิทยาลัย ซึ่งเป็นข้อมูล 3 ระดับ ผู้วิจัยจึงศึกษาการวิเคราะห์ข้อมูลพหุระดับด้วยโมเดลเชิงเส้นตรงระดับลดหลั่นสำหรับข้อมูล 3 ระดับ ดังนี้

- **การวิเคราะห์ข้อมูล 3 ระดับ**

ในหัวข้อการวิเคราะห์ข้อมูล 3 ระดับโดยใช้เทคนิควิธีของ HLM ประกอบด้วย โมเดลการวิเคราะห์ข้อมูล การพัฒนาโมเดลการวิเคราะห์ข้อมูล การวิเคราะห์ข้อมูล และสัมประสิทธิ์การทำนาย โดยมีรายละเอียด ดังนี้

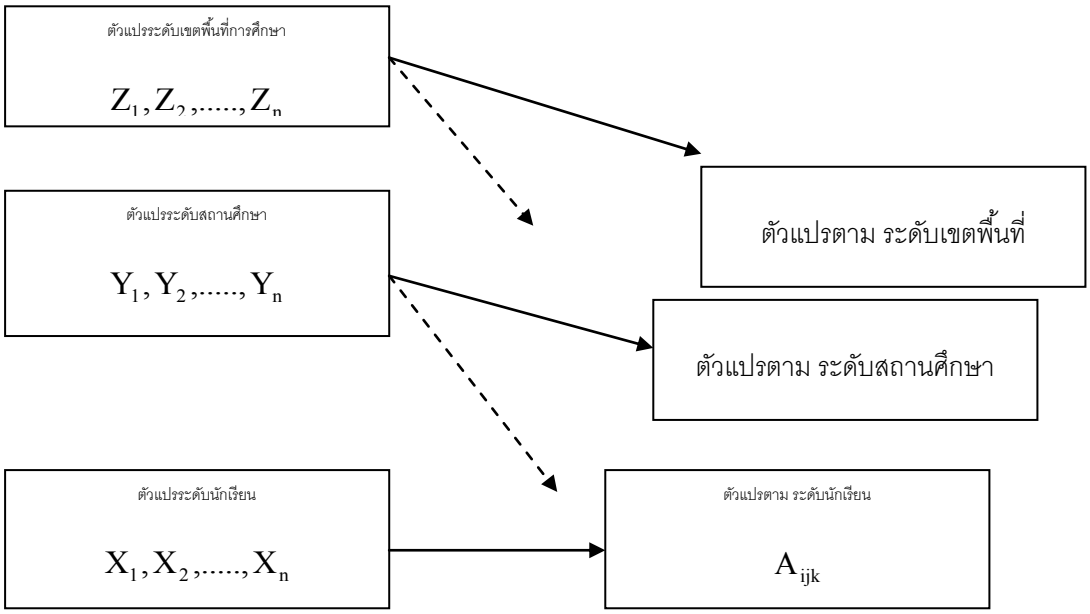
1. **โมเดลการวิเคราะห์ข้อมูล 3 ระดับ**

โมเดลการวิเคราะห์ข้อมูล 3 ระดับ โดยทั่วไปประกอบไปด้วย โมเดล ระดับที่ 1 (Level – 1 Model) ซึ่งเป็นการวิเคราะห์ระหว่างสมาชิกภายในกลุ่ม โดยสร้างสมการถดถอยระหว่างตัวแปรตามและตัวแปรทำนายระดับสมาชิกที่อยู่ภายในกลุ่มเดียวกัน ทำให้ได้ค่าสัมประสิทธิ์ของกลุ่มซึ่งนำไปใช้เป็นตัวแปรตามสำหรับการวิเคราะห์ภายในระดับที่ 2 โมเดลระดับที่ 2 (Level – 2 Model) ซึ่งเป็นการวิเคราะห์ระหว่างกลุ่มภายในหน่วย โดยการสร้างสมการถดถอยระหว่างตัวแปรตามระดับกลุ่มและตัวแปรทำนายระดับกลุ่มที่อยู่ภายในหน่วยเดียวกัน ทำให้ได้ค่าสัมประสิทธิ์ของหน่วยซึ่งนำไปใช้เป็นตัวแปรตามสำหรับการวิเคราะห์ในระดับที่ 3 และโมเดลระดับที่ 3 (Level – 3 Model) เป็นการวิเคราะห์ระหว่างหน่วย โดยการสร้างสมการถดถอยระหว่างตัวแปรตามระดับหน่วยกับตัวแปรทำนายระดับหน่วย จึงทำให้สามารถศึกษาผลของตัวแปรทำนายทั้ง 3 ระดับ ต่อตัวแปรตามที่สนใจในระดับสมาชิก , กลุ่ม และหน่วย รวมทั้งสามารถศึกษาปฏิสัมพั นธ์ระหว่างตัวแปรภายในระดับเดียวกันและต่างระดับได้ด้วย (ศิริชัย กาญจนวาสี, 2548)

2. **การพัฒนาโมเดลข้อมูล 3 ระดับ**

เมื่อผู้วิจัยให้ความสนใจต่อการทำนายหรืออธิบายตัวแปรตามด้วยตัวแปรทำนาย 3 ระดับ ภายใต้บริบทของคำถามวิจัยที่มีความเฉพาะเจาะจง โครงสร้างตัวแปรทำนาย 3 ระดับ ไม่ว่าจะอยู่ในลักษณะของข้อมูลภาคตัดขวาง (Cross-section) หรือข้อมูลติดตามระยะยาว (Longitudinal) จะช่วยอธิบายหรือทำนายความผันแปรของตัวแปรตามที่สนใจศึกษาได้อย่างลุ่มลึกและครอบคลุม ผู้วิจัยต้องศึกษาแนวคิดของทฤษฎีและค้นคว้างานวิจัยที่เกี่ยวข้อง เพื่อคัดเลือก ตัวแปรสำคัญและสร้างโมเดลตามสมมติฐานของการวิจัยตัวอย่างลักษณะข้อมูล เช่น ระดับที่ 1 ข้อมูลเกี่ยวกับนักเรียน ซึ่งนักเรียนแต่ละคนรวมตัวอยู่ในโรงเรียนใดโรงเรียนหนึ่ง หรือนักเรียนสอดแทรก (nested) อยู่ในโรงเรียน ระดับที่ 2 ข้อมูลเกี่ยวกับโรงเรียน ซึ่งโรงเรียนแต่ ละโรงรวมตัว

กันอยู่ในเขตพื้นที่การศึกษา หรือสอดแทรกอยู่ในเขตพื้นที่การศึกษา เป็นต้น (ศิริชัย กาญจนวาสี, 2548) ดังแผนภาพที่ 1



ภาพที่ 1 โครงสร้างโมเดล 3 ระดับและแนวทางการวิเคราะห์

ตัวอย่างการสร้างโมเดลการวิเคราะห์ข้อมูล สมมติว่าผู้วิจัย คัดเลือกตัวแปร 3 ระดับ สำหรับทำนายตัวแปรตามระดับนักเรียนที่สนใจ คือ A_{ijk} โดยตัวแปรทำนาย 3 ระดับ ประกอบด้วยตัวแปรทำนายระดับนักเรียน , สถานศึกษา และเขตพื้นที่การศึกษา ระดับละ 1 ตัวแปร ได้แก่ $X1_{ijk}$, $Y1_{jk}$ และ $Z1_k$ ตามลำดับ จากตัวอย่างดังกล่าว สามารถเขียนผังแสดงตัวแปร 3 ระดับ (ศิริชัย กาญจนวาสี, 2548) ได้ดังนี้

ตารางที่ 2 แสดงตัวแปร 3 ระดับ

ระดับ	หน่วยการวิเคราะห์	ตัวแปรตาม	ตัวแปรทำนาย
1	นักเรียน	A_{ijk}	$X1_{ijk}$
2	สถานศึกษา		$Y1_{jk}$
3	เขตพื้นที่การศึกษา		$Z1_k$

โมเดลการวิเคราะห์ข้อมูล 3 ระดับของกรณีตัวอย่าง เป็นดังนี้

โมเดลระดับที่ 1 (Level – 1 Model) (โมเดลระหว่างนักเรียน ภายในโรงเรียนและเขตพื้นที่การศึกษา)

$$A_{ijk} = P0_{jk} + P1_{jk}(X1_{ijk}) + E_{ijk}$$

โมเดลระดับที่ 2 (Level – 2 Model) (โมเดลระหว่างโรงเรียน, ภายในเขตพื้นที่การศึกษา)

$$P0_{jk} = B00_{k} + B01_{k}(Y1_{jk}) + R0_{jk}$$

$$P1_{jk} = B10_{k} + B11_{k}(Y1_{jk}) + R1_{jk}$$

โมเดลระดับที่ 3 (Level – 3 Model) (โมเดลระหว่างเขตพื้นที่การศึกษา)

$$B00_{k} = G000 + G001(Z1_{k}) + U00_{k}$$

$$B01_{k} = G010 + G011(Z1_{k}) + U01_{k}$$

$$B10_{k} = G100 + G101(Z1_{k}) + U10_{k}$$

$$B11_{k} = G110 + G111(Z1_{k}) + U11_{k}$$

3. การวิเคราะห์ข้อมูล

ในการวิเคราะห์โมเดล 3 ระดับ ผู้วิเคราะห์สามารถเริ่มทำการวิเคราะห์จากขั้นพื้นฐาน คือ การวิเคราะห์โมเดล 3 ระดับแบบไม่มีเงื่อนไขอย่างสมบูรณ์ (Fully unconditional model) ซึ่งเป็นตัวแปรที่ไม่มีตัวแปรทำนายใด ๆ ทั้งสิ้น เพื่อศึกษาความแปรปรวนเฉพาะของตัวแปรตามว่ ามีความผันแปรมากน้อยเพียงใดในแต่ละระดับของการวิเคราะห์ ขั้นต่อมาทำการวิเคราะห์โมเดล 3 ระดับแบบไม่มีเงื่อนไข (Unconditional model) โดยเพิ่มตัวแปรทำนายเฉพาะในโมเดลระดับที่ 1 รวมทั้งอิทธิพลสุ่มหรือความแปรปรวนของตัวแปรตาม และสัมประสิทธิ์การถดถอยสำหรับระดับที่ 2 และ 3 ต่อจากนั้นจึงทำการวิเคราะห์โมเดล 3 ระดับตามสมมติฐาน (Hypothetical model) เพื่อศึกษาผลของตัวแปรทำนายทั้ง 3 ระดับต่อตัวแปรตามที่สนใจ (ศิริชัย กาญจนวาสี, 2548) ดังนี้

1. การวิเคราะห์ขั้นโมเดลศูนย์ (Null model) เพื่อให้เห็นภาพรวมของตัวแปรตามทั้งสามระดับโดยไม่มีตัวแปรอิสระเข้าร่วมพิจารณา และเพื่อตรวจสอบว่าตัวแปรตามระดับที่ 3 ว่ามีความ

ผันแปรเพียงพอที่จะวิเคราะห์ตัวแปรอิสระที่มีอิทธิพลในขั้นต่อไปหรือไม่ โดยใช้ t-test ทดสอบ fixed effect และใช้ χ^2 – test ทดสอบ random effect

2. การวิเคราะห์ขั้นโมเดลแบบง่าย (Simple model) เป็นการวิเคราะห์โดยนำตัวแปรระดับที่ 1 และระดับที่ 2 เข้ามาวิเคราะห์ทีละตัว เพื่อศึกษาว่าตัวแปรอิสระตัวนั้นมีอิทธิพลต่อตัวแปรตามระดับที่สามอย่างมีนัยสำคัญหรือไม่ และเพื่อตรวจสอบว่าตัวแปรอิสระตัวนั้นทำให้ตัวแปรตามดังกล่าวเกิดความผันแปรระหว่างหน่วยของระดับที่ 3 หรือไม่ โดยใช้ t – test ทดสอบ fixed effect และใช้ χ^2 – test ทดสอบ random effect และ conterlized ตัวแปรอิสระในกระบวนการวิเคราะห์โดยใช้คำสั่ง centering around grand mean

3. การวิเคราะห์โมเดล ตามสมมติฐาน (Hypothetical model) เป็นการวิเคราะห์เพื่อตรวจสอบอิทธิพลของตัวแปรอิสระ ที่มีต่อตัวแปรตามทั้ง 3 ระดับตามที่สนใจ โดยใช้ t – test ทดสอบ fixed effect และใช้ χ^2 – test ทดสอบ random effect และ conterlized ตัวแปรอิสระในกระบวนการวิเคราะห์ โดยใช้คำสั่ง centering around grand mean

4. สัมประสิทธิ์การทำนาย

สัมประสิทธิ์การทำนายของโมเดลแต่ละระดับ (R^2) หรือ สัดส่วนความแปรปรวนที่อธิบายได้ (Proportion variance explained) สามารถคำนวณได้จากเทียบสัดส่วนความแปรปรวนของค่า ส่วนที่เหลือที่ลดลง เมื่อมีตัวแปรทำนายกับเมื่อไม่มีตัวแปรทำนาย ซึ่งมีสูตรโดยทั่วไปดังนี้ (ศิริชัย กาญจนวาสี, 2548)

$$R^2 = \frac{\tau_{00}(\text{unconditional}) - \tau_{00}(\text{conditional})}{\tau_{00}(\text{unconditional})}$$

เมื่อ R^2 แทน สัมประสิทธิ์การทำนายของโมเดล

$\tau_{00}(\text{unconditional})$ แทน ความแปรปรวนของค่าส่วนที่เหลือของโมเดลแบบไม่มีเงื่อนไข (โมเดลที่ไม่มีตัวแปรทำนาย)

$\tau_{00}(\text{conditional})$ แทน ความแปรปรวนของค่าส่วนที่เหลือของโมเดลแบบมีเงื่อนไข (โมเดลที่มีตัวแปรทำนาย)

ตอนที่ 4 การประเมินผลนักเรียนระดับนานาชาติ

4.1 การศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ พ.ศ.2550 (TIMSS2007)

IEA (The International Association for the Evaluation of Educational Achievement) เป็นสมาคมนานาชาติที่ทำหน้าที่ประเมินผลด้านการศึกษา ได้จัดทำโครงการเพื่อประเมินผลด้านคณิตศาสตร์และวิทยาศาสตร์มาแล้วในช่วง 30 ปีที่ผ่านมา และประเทศไทยได้เข้าร่วมโครงการวิจัยเพื่อประเมินผลดังกล่าวมาเป็นระยะ ในปี ค.ศ. 2007 IEA ได้จัดทำโครงการประเมินผลร่วมกับนานาชาติในชื่อว่า TIMSS2007 (Trends in International Mathematics and Science Study – 2007) เพื่อประเมินผลการจัดการศึกษาวิชาวิทยาศาสตร์และคณิตศาสตร์ของนักเรียนระดับ ป.4 และ ม.2 (Grade 4 and Grade 8) โดยมีประเทศต่างๆ ร่วมโครงการประมาณ 70 ประเทศการประเมินผลดังกล่าวมีเป้าหมายเพื่อศึกษาความรู้ความสามารถนักเรียนในระดับ ป.4 และม.2 ของประเทศต่างๆ ทั่วโลก ในวิชาวิทยาศาสตร์และคณิตศาสตร์ และประเทศไทยได้เข้าร่วมการประเมินในช่วงชั้น ม.2 ในส่วนของการรู้ (Literacy) และการเรียนรู้ (Leaning) ที่ครอบคลุมเนื้อหาด้านคณิตศาสตร์และวิทยาศาสตร์ที่เรียนรู้ถึงชั้น ม.2 โดยเปรียบเทียบกับประเทศที่เข้าร่วมโครงการรวมทั้งเปรียบเทียบเพื่อดูแนวโน้มของผลสัมฤทธิ์ในวิชาดังกล่าวของนักเรียนของประเทศไทยกับการประเมินในครั้งก่อนด้วย เพื่อให้สามารถเห็นแนวโน้มด้านการจัดการเรียนรู้การสอนคณิตศาสตร์และวิทยาศาสตร์ และสามารถใช้ข้อมูลข้อสนเทศจากการวิจัยครั้งนี้ ในการพัฒนาการเรียนการสอนต่อไป

4.1.1 กรอบแนวคิด ในข้อสอบวิชาคณิตศาสตร์ ที่ใช้ในโครงการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์ พ.ศ.2550 (TIMSS2007)

กรอบแนวคิดที่ใช้ในการประเมินของข้อสอบ TIMSS 2007 ได้เป็นการนำเอาสารสนเทศจากการสอบในปี 2003 มาใช้เป็นแนวในการปรับปรุงการประเมินให้มีความครอบคลุมมากยิ่งขึ้น โดยโมเดลหลักสูตรการประเมิน TIMSS 2007 ถูกสร้างมาจากหลักสูตร นิยามแบบกว้างๆ หรือแนวคิดหลักซึ่งใช้ในการพิจารณาถึงโอกาสทางการศึกษาที่มอบให้กับนักเรียน และองค์ประกอบซึ่งมีอิทธิพลต่อนักเรียน

โมเดลหลักสูตรของTIMSS 2007 มีองค์ประกอบที่สำคัญ3 ส่วน คือจุดมุ่งหมายของหลักสูตร กระบวนการของหลักสูตร และความสำเร็จของหลักสูตร โดยสามารถแสดงได้ดังตารางที่3

ตารางที่ 3 แสดงโมเดลหลักสูตรของ TIMSS 2007

องค์ประกอบ	รายละเอียด
1. จุดมุ่งหมายของหลักสูตร	เนื้อหาของหลักสูตรการศึกษาชาติ และความต้องการของสังคม
2. กระบวนการของหลักสูตร	โรงเรียน, ครู และการจัดห้องเรียน
3. ความสำเร็จของหลักสูตร	ผลที่เกิดขึ้น และคุณลักษณะของผู้เรียน

การประเมินการจัดการศึกษาคณิตศาสตร์ของ TIMSS 2007 โดยการประเมินในระดับชั้นมัธยมศึกษาปีที่ 2 ประกอบไปด้วยส่วนที่สำคัญ 2 ส่วนคือ ส่วน ของเนื้อหาและ ส่วนที่เน้นกระบวนการคิด โดยในส่วนของเนื้อหาประกอบด้วย 4 ส่วนที่สำคัญ และส่วนที่เน้นกระบวนการคิดจะมีอยู่ด้วยกัน 3 ส่วน ซึ่งมีรายละเอียดดังตารางที่ 4

ตารางที่ 4 แสดงร้อยละของเนื้อหาในวิชาคณิตศาสตร์ของการประเมิน TIMSS 2007

ส่วนประกอบ	ร้อยละ
ส่วนเนื้อหา (Content Domain)	
1. จำนวน	30%
2. พีชคณิต	30%
3. เรขาคณิต	20%
4. ข้อมูล และโอกาส	20%
รวม	**100%**
ส่วนกระบวนการคิด (Conitive Domain)	
1. ความรู้	35%
2. การประยุกต์	40%
3. การให้เหตุผล	25%
รวม	**100%**

ขอบเขตของส่วนประกอบของการประเมิน การจัดการศึกษาคณิตศาสตร์ ของ TIMSS 2007 ในแต่ละเนื้อหา มีความหลากหลายของเนื้อ หาที่ใช้ ซึ่งเป็นการนำเสนอรายการของจุดประสงค์การเรียนรู้ซึ่งมีความครอบคลุมเนื้อหา และมีความเหมาะสมสำหรับประเทศต่างๆ

4.1.2 กรอบแนวคิดเกี่ยวกับบริบทในโครงการ ศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์ พ.ศ.2550 (TIMSS2007)

ภายใต้บริบท ของการจัดการศึกษาในประเทศ ต่าง ๆ ซึ่ง มีจำนวนองค์ประกอบที่หลากหลายที่ส่งผลไปยัง การเรียนรู้ของผู้เรียน ตัวอย่างเช่น ลักษณะของโรงเรียน สถานที่ตั้งของโรงเรียน กระบวนการจัดการเรียนการสอน คุณลักษณะของครู เป็นต้น โดยสำหรับการประเมินการจัดการศึกษาคณิตศาสตร์ของ TIMSS 2007 นั้นสามารถกำหนดกรอบแนวคิดได้กว้างๆ 5 ประเด็นด้วยกัน คือ หลักสูตร โรงเรียน ครูและการเตรียมการสอน การจัดกิจกรรมการเรียนรู้ และนักเรียน

จากการกำหนดกรอบแนวคิดดังกล่าวจึงสามารถสรุปข้อมูลที่ใช้ในโครงการประเมินผล TIMSS 2007 จำแนกตามระดับของข้อมูลได้เป็น 2 ระดับ คือระดับนักเรียน ระดับสถานศึกษา ซึ่งมีรายละเอียดดังตารางที่ 5 ต่อไปนี้

ตารางที่ 5 สรุปความหมายของตัวแปรที่ใช้ในโครงการประเมินผล TIMSS 2007

ตัวแปร	ความหมาย
ระดับสถานศึกษา	
ขนาดสถานศึกษา	จำนวนนักเรียนที่ลงทะเบียนเรียนในสถานศึกษา
ตำแหน่งที่ตั้งของสถานศึกษา	พิจารณาจากขนาดชุมชนที่ สถานศึกษาตั้งอยู่ จำแนกเป็น ตำแหน่งหมู่บ้าน (3,000 คน หรือน้อยกว่า) ตำบล (3,001-15,000 คน) เมืองเล็ก (15,001-50,000 คน) เมือง (50,001-100,000 คน) เมืองใหญ่ (100,001-500,000 คน) เมืองใหญ่มาก (มากกว่า500,000 คน)
การใช้ภาษาอังกฤษในสถานศึกษา	พิจารณาจากร้อยละของนักเรียนที่พูดภาษาอังกฤษเป็นธรรมชาติ จำแนกเป็น น้อยมาก (ร้อยละ 0-10) น้อย (ร้อยละ 11-25) ปานกลาง (ร้อยละ 26-50) มาก (มากกว่าร้อยละ 50)
เวลาในการเรียนของสถานศึกษา	พิจารณาจากเวลาที่สถานศึกษาจัดการเรียนการสอนรายวัน รายสัปดาห์
การพัฒนาสถานศึกษา	คิดจากร้อยละของเวลาที่ทางโรงเรียนจัดกิจกรรมการอบรม สัมมนาเกี่ยวกับการจัดการเรียนการสอน ความรู้เกี่ยวกับหลักสูตร การให้คำแนะนำ การประเมินครู การจัดกิจกรรมสร้างความสัมพันธ์กับชุมชน และการระดมทุน
การมีส่วนร่วมของผู้ปกครอง	คิดจากร้อยละของผู้ปกครองที่มีส่วนร่วมในการจัดกิจกรรมพิเศษของโรงเรียนเช่น งานวันวิชาการ งานคอนเสิร์ต งานกีฬาสีเป็นต้น และการบริจาคเงินเข้าโรงเรียน

ตารางที่ 5 สรุปความหมายของตัวแปรที่ใช้ในโครงการประเมินผล TIMSS 2007 (ต่อ)

ตัวแปร	ความหมาย
บุคลากรในสถานศึกษา	พิจารณาจากระดับความพึงพอใจในการทำงานของ ครู ความเข้าใจในเป้าหมายของโรงเรียน โดยจำแนกเป็น 5 ระดับ คือน้อยมาก น้อย ปานกลาง มาก และมากที่สุด
ผลการปฏิบัติงานของครู	พิจารณาจากระดับความความสำเร็จในการดำเนินงานตามหลักสูตร ความคาดหวังของครูที่มีต่อผลสัมฤทธิ์ทางการเรียน โดยจำแนกเป็น 5 ระดับ คือน้อยมาก น้อย ปานกลาง มาก และมากที่สุด
การพัฒนาบุคลากร	คิดจากร้อยละของครูที่ได้รับการสนับสนุนให้ได้รับโอกาสในการพัฒนาเกี่ยวกับคณิตศาสตร์ จำแนกเป็น น้อยมาก (ร้อยละ 25 หรือน้อยกว่า) น้อย (ร้อยละ 25-50) มาก (ร้อยละ 51-75) มากที่สุด (มากกว่าร้อยละ 75)
เป้าหมายของโรงเรียน	พิจารณาจากประสิทธิผลของโรงเรียนในด้านต่างๆ เช่น ทางด้านวิชาการ พัฒนาการของนักเรียน หรือระเบียบวินัยในตนเองของนักเรียน เป็นต้น
บทบาทของผู้บริหาร	พิจารณาจากความเป็นผู้นำในการบริหารงานโรงเรียนให้มีประสิทธิผล
แหล่งทรัพยากรณ์ อุปกรณ์	พิจารณาจากการให้ความส นับสนุนการจัดการเรียนการสอนในวิชาคณิตศาสตร์ เช่น ความเพียงพอขอสถานที่เรียน งบประมาณ ระบบไฟ เป็นต้น
เทคโนโลยี การสนับสนุน และอุปกรณ์	พิจารณาจากความพร้อมในเรื่องของอุปกรณ์ สื่อการจัดการเรียนการสอน เช่นระบบคอมพิวเตอร์ การพัฒนาบุคลากร การขึ้นเงินเดือนครู เป็นต้น
บริบทสถานศึกษา	พิจารณาถึงคุณค่า วัฒนธรรม ความปลอดภัย และโครงสร้างของชุมชน
ความร่วมมือของผู้ปกครอง	พิจารณาจากการมีส่วนร่วมของผู้ปกครองในการดูแลนักเรียน
การเตรียมความพร้อมด้านวิชาการ	พิจารณาจากความรู้ในเรื่องหลักสูตรของครู ความรู้ในเนื้อหาในวิชาคณิตศาสตร์ และประสบการณ์การสอนในวิชาคณิตศาสตร์
คุณลักษณะครู	พิจารณาจากเพศ ประสบการณ์ครู
หลักสูตรสถานศึกษา	พิจารณาจากการจัดการเรียนในวิชาคณิตศาสตร์ว่ามีความครอบคลุม รวมไปถึงความเข้มข้นในเนื้อหาวิชา โดยพิจารณาจากกรอบการประเมินของ TIMSS

ตารางที่ 5 สรุปความหมายของตัวแปรที่ใช้ในโครงการประเมินผล TIMSS 2007 (ต่อ)

ตัวแปร	ความหมาย
การจัดกิจกรรมการเรียนการสอน	พิจารณาจากกิจกรรมที่ใช้ในการจัดการเรียนการสอนของครู
การประเมินและการบ้าน	เวลาที่ใช้ในการประเมินผลของครูในห้องเรียน รวมไปถึงการให้การบ้านสำหรับนักเรียนในการเรีย นการสอนปกติ เพื่อฝึกให้นักเรียนเกิดทักษะในบทเรียน
การใช้ computer และ internet	พิจารณาจากการใช้เทคโนโลยีในการจัดการเรียนการสอน
การใช้เครื่องคิดเลข	การใช้เครื่องเลขในการจัดการเรียนการสอน
ความสำคัญของการสืบค้น	พิจารณาจากการให้ความสำคัญการเรียนรู้ของผู้เรียนที่เน้นการค้นคว้า
ระดับนักเรียน	
ผลการสอบ TIMSS	คะแนนที่จากการทำ แบบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ซึ่ง แสดงถึงสมรรถนะ ทางการเรียนในวิชาคณิตศาสตร์ ตามกรอบการประเมินการรู้วิชาคณิตศาสตร์ ในโครงการประเมิน TIMSS 2007
เพศ	จำแนกเป็นนักเรียนชายและหญิง
แหล่งทรัพยากรที่บ้าน	พิจารณาเกี่ยวกับข้อมูลส่วนตัวของนักเรียน เช่น จำนวนหนังสือเรียนที่บ้าน เป็นต้น
ระดับการศึกษาของครอบครัว	ได้จากค่าคะแนน ซึ่งประกอบด้วยปัจจัยระดับการศึกษาของผู้ปกครอง
ความมั่งคั่งของครอบครัว	ค่าที่บ่งบอกถึงความเป็นอยู่ ความสะดวกสบายของครอบครัว ได้จากการตอบคำ ถามของนักเรียนเกี่ยวกับการมีไว้ครอบครองในครอบครัว เช่น การมีการเชื่อมต่ออินเตอร์เนต โทรศัพท์มือถือ โทรทัศน์ คอมพิวเตอร์ เครื่องเล่นเกมรถยนต์
เจตคติของนักเรียนต่อวิชาคณิตศาสตร์	พิจารณาจากระดับการมีเจตคติที่ดีต่อวิชาคณิตศาสตร์ โดยจำแนกเป็น 4 ระดับ คือน้อยมาก น้อย มาก และมากที่สุด
การใช้เวลาในการศึกษาคณิตศาสตร์เพิ่มเติม	พิจารณาจากระดับการมีกิจกรรมในวิชาคณิตศาสตร์ในช่วงเวลานอกห้องเรียน เช่น การทำการบ้าน หรือการเรียนพิเศษ เป็นต้น โดยจำแนกเป็น 3 ระดับ คือ ระดับต่ำ ระดับปานกลาง และระดับสูง
การเห็นคุณค่าในวิชาคณิตศาสตร์	พิจารณาจากระดับการเห็นคุณค่าวิชาคณิตศาสตร์ โดยจำแนกเป็น 3 ระดับ คือ ระดับต่ำ ระดับปานกลาง และระดับสูง

การสร้างกรอบแนวคิดของโครงการประเมินผลนักเรียน TIMSS 2007 นั้นสามารถจัดข้อมูลได้เป็นสามระดับคือระดับนักเรียน ระดับสถานศึกษา และระดับประเทศ ทั้งนี้ข้อมูล TIMSS 2007 สามารถนำมาใช้เพื่อเป็นข้อมูลส่วนหนึ่งในการรายงานผลคุณภาพของการจัดการศึกษาเกี่ยวกับคุณภาพและประสิทธิผลของระบบการศึกษาของประเทศสมาชิกและประเทศที่เข้าร่วมโครงการ ตลอดจนการศึกษาวิจัยเกี่ยวกับรูปแบบและลักษณะของเครื่องมือที่ใช้ใน การประเมินผลที่ผลต่อการตอบสนองข้อสอบเนื่องจากคุณลักษณะพื้นฐานของนักเรียนในประเทศที่เข้าร่วมโครงการ ซึ่งอาจจะเป็นส่วนหนึ่งที่อาจจะทำให้เกิดความไม่เท่าเทียมกันที่เกิดจากสาเหตุของความแตกต่างกันในด้านคุณลักษณะพื้นฐานของนักเรียนและสถานศึกษา เช่น เพศของนักเรียน ภาษาที่ใช้ และลักษณะทางเนื้อหาในหลักสูตรของแต่ละประเทศ ซึ่งกล่าวได้ว่าคุณลักษณะพื้นฐานดังกล่าวเป็นคุณลักษณะแฝงที่ส่งผลต่อการตอบสนองข้อสอบข้อสอบ โดยเป็นการศึกษาโดยอาศัยวิธีการวิเคราะห์ที่มีความทันสมัย และละเอียดมากยิ่งขึ้น

4.2 การประเมินผลนักเรียนระดับนานาชาติ PISA

องค์การในระดับนานาชาติที่เกี่ยวข้องกับการสำรวจเพื่อประเมินผลนักเรียนนานาชาติและเปรียบเทียบผลสัมฤทธิ์ของนักเรียนมีสององค์การหลักคือสมาคมนานาชาติเพื่อประเมินผลสัมฤทธิ์ทางการศึกษา (The International Association for the Evaluation of Educational Achievement : IEA) และองค์กรเพื่อความร่วมมือและพัฒนาทางเศรษฐกิจ (Organisation for Economic Co-operation and Development: OECD) ในการดำเนินโครงการประเมินผลนักเรียนระดับนานาชาติ (Programme for International Student Assessment: PISA) ในการศึ กษาครั้งนี้ใช้ฐานข้อมูลจาก การสำ รวจและการประเมินผลของนักเรียน ผู้วิจัยจึงเสนอลักษณะของโครงการประเมินและเครื่องมือที่ใช้ในการประเมิน ดังนี้

4.2.1 ความเป็นมาและจุดมุ่งหมายของโครงการประเมินผลนักเรียนระดับนานาชาติ PISA

องค์กรเพื่อความร่วมมือและพัฒนาทางเศรษฐกิจ (Organisation for Economic Co-operation and Development: OECD) ถือว่าคุณภาพของการศึกษาเป็นตัวชี้วัดศักยภาพของการพัฒนาทางเศรษฐกิจ จึงมีการศึกษาวิจัยรูปแบบประสิทธิภาพของระบบการศึกษาของประเทศ

สมาชิก และป้อนข้อมูลกลับให้ประเทศสมาชิกทราบเกี่ยวกับระบบการศึกษาข องประเทศเพื่อเตรียมพร้อมเยาว ชนสำหรับอนาคต ในช่วงทศวรรษ ค .ศ. 2000 OECD ได้ดำเนินโครงการประเมินผลนักเรียนนานาชาติ หรือเรียกชื่อโครงการว่า PISA (Programme for International Student Assessment: PISA) เป็นโครงการศึกษาสำรวจความรู้และทักษะของนักเรียนกลุ่มอายุ 15 ปี ซึ่งเป็นวัยจบการศึกษาภาคบั งคับ ในประเทศที่เป็นสมาชิกองค์กรเพื่อความร่วมมือและพัฒนาทางเศรษฐกิจ (OECD) และประเทศร่วมโครงการที่ไม่ใช่สมาชิกซึ่งประเทศไทยเป็นหนึ่งในประเทศร่วมโครงการที่ไม่ใช่สมาชิก โดยมีจุดมุ่งหมายเพื่อให้ข้อมูลกับประเทศสมาชิกว่าได้เตรียมความพร้อมของประชาชนให้มีศักยภาพสำหรับการแข่งขันในประชาคมโลกเพียงพอหรือไม่อย่างไร เพื่อหาตัวชี้วัดคุณภาพการศึกษาในระดับนโยบาย

โครงการประเมินผล PISA ต้องการวัดการรู้เรื่อง (Literacy) ของกลุ่มนักเรียนที่จบการศึกษาภาคบังคับของแต่ละระบบการศึกษาหรือกลุ่มนักเรียนอายุ 15 ปี สำ หรับการเตรียมพร้อม ในอนาคต การประเมินผลจึงเน้นเพื่อชี้อนาคต ให้ความชัดเจนกับความสามารถในการใช้ความรู้และทักษะในชีวิตจริงของผู้ใหญ่ที่ยังอยู่ในวัย 15 ปี ดังนั้น PISA จึงเลือกประเมินสิ่งที่เรียกว่า การรู้เรื่อง (Literacy) ซึ่งเน้นที่สมรรถนะของนักเรียนในการใช้ความรู้และทักษะเพื่ อเผชิญกับโลกในชีวิตจริงแทนการประเมินความรู้ที่ได้เรียนตามหลั กสูตรในสถานศึกษา ตัวชี้วัดที่สำ คัญได้แก่ การรู้เรื่องวิทยาศาสตร์ (Scientific Literacy)การรู้เรื่องคณิตศาสตร์ (Mathematics Literacy) การรู้เรื่องการอ่าน (Reading Literacy) โครงการ PISA 2006 มีความแตกต่างจาก PISA 2003 และ 2000 คือการศึกษาของ PISA ในสองครั้งแรก ได้แก่ PISA 2000 และ PISA 2003 เน้นการอ่าน คณิตศาสตร์ ตามลำ ดับ ส่วนใน PISA 2006 เน้นความสำ คัญการประเมินการรู้วิทยาศาสตร์ ทั้งนี้วิทยาศาสตร์ มีน้ำหนักของการประเมินเป็น 60% และการอ่านและคณิตศาสตร์มีน้ำหนักอย่างละ 20% ของการประเมิน (OECD, 2007)

จุดมุ่งหมายหลักของ PISA คือการประเมินสมรรถนะของนักเรียนที่จะใช้ความรู้และทักษะเพื่อเผชิญกับโลกในชีวิตจริงมากกว่าการประเมินความรู้ที่ได้เรียนตามหลักสูตรในสถานศึกษา โดย OECD/PISA เรียกสมรรถนะนี้ว่า “การรู้เรื่อง” (Literacy) การประเมินผลของ PISA มุ่งเน้นเพื่อให้ความชัดเจนกับความสามารถในการใช้ความรู้และทักษะในชีวิตจริงของนักเรียน โดยแนวคิดของการประเมินจะสะท้อนให้เกิดการเปลี่ยนแปลงในเป้าหมายและ

จุดประสงค์ของหลักสูตรในสถานศึกษา ผลการประเมินของ PISA จึงใช้เป็นข้อมูลที่บ่งชี้คุณภาพการจัดการศึกษาของชาติ นอกจากนี้ยังสามารถบอกปัจจัยที่ส่งผลต่อคุณภาพการจัดการศึกษา (สสวท, 2551) การประเมินผลในลักษณะนี้สามารถชี้แนะแนวทางสำ หรับการตัดสินใจในระดับนโยบายการส่งเสริมและสนับสนุนตลอดจนการกระจายทรัพยากรทางการศึกษา และทำ ให้ทราบถึงปัจจัยต่างๆที่ส่งผลต่อการพัฒนาความรู้และทักษะของนักเรียน เพื่อบอกให้ทราบถึงคุณภาพและประสิทธิภาพการจัดการศึกษาของสถานศึกษาในการเตรียมความพร้อมนักเรียน สำ หรับการใช้ชีวิตในอนาคต (สสวท, 2547)

4.2.2 **ลักษณะเครื่องมือที่ใช้ในโครงการประเมินผล** PISA

การสร้างเครื่องมือต ามโครงการประเมินผล PISA ใช้แนวคิดตามทฤษฎีการตอบสนองของข้อสอบ(Item Response Theory) และมีวิธีดำ เนินอยู่ภายใต้ข้อตกลงเบื้องต้นของโมเดลการตอบสนองข้อสอบ การสร้างเครื่องมือในแต่ละด้านอยู่ภายใต้ข้อตกลงเบื้องต้นของโมเดลราสช์ (Rasch Model) นั่นคือข้อตกลงเบื้องต้นที่ สอดคล้องกับความเป็นเอกมิติของแบบสอบ (unidimensional models) ซึ่งการวิเคราะห์ความเป็นเอกมิติของแบบสอบสามารถใช้วิ เคราะห์องค์ประกอบ และได้มีการนำ ไปใช้อย่างหลากหลาย เช่น การ นำ ไปปฏิบัติเกี่ยวกั บข้อตกลงเบื้องต้นของความเป็นเอกมิติของแบบสอบของ Blum (2001 cited in Goldstein, 2004) กล่าวถึงการตรวจสอบความเป็นเอกมิติของแบบสอบว่า ถึงแม้ว่าการตรวจสอบองค์ประกอบพบว่ ามีสองมิติ แต่แสดงให้เห็นว่ามีจำ นวนข้อสอบในมิติหนึ่งอยู่จำ นวนน้อย ดังนั้นเครื่องมือนั้นจึงจัดว่าเป็นการวัดเพียงมิติเดียว (Blum, et al, 2001 cited in Goldstein, 2004)

เครื่องมือที่ใช้ในโครงการประเมินผล PISA ประกอบด้วยแบบสอบวัดการรู้เรื่องทั้งสามด้าน คือ การรู้เรื่องวิทยาศาสตร์ (Scientific Literacy) การรู้เรื่องคณิตศาสตร์ (Mathematics Literacy) การรู้เรื่องการอ่าน (Reading Literacy) และแบบประเมินเจตคติทางวิทยาศาสตร์ของนักเรียน

4.2.3 **กระบวนการพัฒนาแบบสอบและคุณภาพของแบบสอบในโครงการประเมินผล** PISA

กระบวนการพัฒนาเครื่องมือที่ใช้ในโครงการประเมินผล PISA มีการพัฒนาเครื่องมือจากศูนย์การพัฒนาแบบสอบ (Test Development Centres) โดยพิจารณาตามกรอบของนิยาม

ในแต่ละองค์ประกอบหลัก (Domain) ของการประเมิน และมีการอภิปรายในด้านของเนื้อหา สมรรถนะ และความรู้ โดยกลุ่มผู้เชี่ยวชาญในเนื้อหาและนักพัฒนาแบบสอบ รวมทั้งคณะที่ปรึกษาในศูนย์ระดับชาติของแต่ละประเทศที่เข้าร่ วมโครงการ ตามกาหนดระยะเวลาการดำเนินการของโครงการประเมินผล PISA การพัฒนาเครื่องมื อได้กำหนดไว้ 2 ระยะ (OECD, 2009) ดังนี้

1. การพัฒนาระยะที่หนึ่ง (First phase of development) เป็นกระบวนการพัฒนาข้อสอบที่ให้ผู้พัฒนาแบบสอบซึ่งประกอบด้วยคณะทำ งานในระดับประเทศ ร่วมพิจารณาประเด็นในด้านของภาษาที่ใช้ในแต่ละประเทศ การจัดรูปแบบของข้อสอบ สถานการณ์ ในข้อสอบ และคำ ถามในข้อสอบ ในขั้นตอนนี้จะใช้ลักษณะการสัมภาษณ์นักเรียนกลุ่มเล็ก ด้วยวิธีการคิดเสียงดัง (think-aloud methods) ทั้งการสัมภาษณ์ระดับบุคคลและระดับกลุ่ม เพื่อให้สะท้อนถึงกระบวนการคิดของนักเรียนผลที่ได้จะทำ ให้เกิดความเข้าใจที่ชัดเจนของการใช้ภาษาในคำ ถาม การให้ข้อมูลในสถานการณ์เพื่อให้นักเรียนตอบสนองในข้อสอบ

2. การพัฒนาในระยะที่สอง (Second phase of development) การพัฒนาในระยะนี้เป็นการพิจารณาทบทวนข้อสอบในระดับนานาชาติ ซึ่งเป็นกระบวนการที่ให้ข้อมูลสะท้อนถึงความสัมพันธ์ของข้อสอบที่มีความแตกต่างทางด้านวั ฒนธรรมและบริบทในแต่ละประเทศ ซึ่งพบว่ามีข้อสอบบางข้อที่ต้องตัดทิ้งออกไป โดยการสำรวจจะดำเนินการต่อไปในขั้นตอนการทดลองใช้ข้อสอบ pilot testing และใช้เป็นขั้นตอนหนึ่งในแต่ละศูนย์การสอบแต่ละประเทศ

การตรวจสอบการทำ หน้าที่ต่างกันของข้อสอบ เป็นขั้นตอนหนึ่งในกระบวนกา รพัฒนาระยะที่สอง โดยการตรวจสอบการทำ หน้าที่ต่างกันของข้อสอบในระดับประเทศ (Country DIF) โดยใช้โมเดล IRT ในการประมาณค่าในแต่ละประเทศ จากกลุ่มตัวอย่าง 51 ประเทศ ด้วยการวิเคราะห์ข้อมูลเกี่ยวกับปฏิสัมพันธ์ของข้อสอบกับประเทศ (item-by-country Interactions) หรือเรียกว่า “การทำหน้าที่ต่างของข้อสอบในระดับประเทศ ” ซึ่งแสดงถึงความแตกต่างของวัฒนธรรมแต่ละประเทศ จากการตรวจสอบการทำหน้าที่ต่างของข้อสอบของ OECD (2009) ได้รายงานว่าพบข้อสอบที่ทำหน้าที่ต่างกันหรือมีปฏิสัมพันธ์ระหว่างข้อสอบกับประเทศจำนวน 22 ข้อจากข้อสอบที่วัดด้านทัศนคติ และเป็นข้อสอบที่ทำหน้าที่ต่างกันเนื่องจากตัวแปรเพศจำนวน 15 ข้อ ซึ่งเป็นจำนวนที่น้อยกว่าข้อสอบที่วัดด้านความรู้(cognitive items) แต่ไม่รายงานจำนวนข้อสอบที่ตรวจพบ แสดงให้

เห็นว่าแบบสอบที่ใช้ประเมินในโครงการประเมินผลPISA มีข้อสอบที่ทำหน้าที่ต่างกัน เช่น Le (2006) ได้วิเคราะห์หาข้อสอบที่ทำหน้าที่ต่างกันทางด้านเพศ และภาษาที่ใช้ในแบบสอบ และตรวจพบ DIF ใน PISA รอบทดลองใช้ ปี 2006 และ Grisay and Monseur (2007) ได้ตรวจพบ DIF ใน PISA ที่วัดการรู้เรื่องเกี่ยวกับการอ่านในปี2000 และ 2001 โดยใช้ตัวแปรภาษาในแบบสอบ

คุณภาพของแบบสอบที่ใช้ในโครงการประเมินผล PISA ได้รายงานคุณภาพของแบบสอบในด้านความเที่ยง (Reliability) ของแบบสอบ (OECD, 2009) ในแบบสอบวัดการรู้เรื่องด้านวิทยาศาสตร์ คณิตศาสตร์ การอ่าน และแบบประเมินเจตคติทางวิทยาศาสตร์ของนักเรียน พบว่าค่าความเที่ยงของเครื่องมือจากผลทุกประเทศมีค่าอยู่ในช่วง 0.892 – 0.920 และแบบสอบการรู้เรื่องด้านวิทยาศาสตร์มีค่าความเที่ยงเท่ากับ 0.920 ส่วนค่าความเที่ยงในแต่ละประเทศจะมีค่าแตกต่างกัน สาหรับประเทศไทยมีค่าเที่ยงอยู่ในช่วง 0.84 – 0.88 และแบบสอบการรู้เรื่องด้านวิทยาศาสตร์มีค่าความเที่ยงเท่ากับ 0.88

ดังนั้นผู้วิจัยสามารถสรุปได้ว่าการ ทั้งโครงการประเมินผล TIMSS และการประเมินผลนักเรียนระดับนานาชาติ PISA ต่างก็เป็นโครงการประเมินผลที่มีเป้าหมายในการจัดเดียวกัน นั่นคือต้องการทราบถึงผลสัมฤทธิ์ทางการเรียนในวิชาคณิตศาสตร์ และคณิตศาสตร์ของนักเรี ยนแต่มีความแตกต่างกันในส่วนของลักษณะของการทดสอบ คือ โครงการประเมินผล TIMSS ต้องการวัดผลสัมฤทธิ์ของนักเรียน ส่วน การประเมินผลนักเรียนระดับนานาชาติ PISA มีวัตถุประสงค์เพื่อประเมินถึงความสามารถในการเรียนรู้ในวิชาคณิตศาสตร์ และวิทยาศาสตร์

ตอนที่ 5 เอกสารงานวิจัยเกี่ยวกับการประเมินคุณภาพการจัดการศึกษา

จากการศึกษาเอกสารและงานวิจัยเกี่ยวกับการประเมินคุณภาพการจัดการศึกษา พบว่าตัวบ่งชี้ที่นำมาใช้บ่งบอกคุณภาพการจัดการศึกษาจะเป็นผลสัมฤทธิ์ทางด้านวิชาการของนักเรียน ซึ่งส่วนใหญ่ได้จากการทำแบบสอบ ดังนั้นในการศึกษาครั้งนี้ ผู้วิจัยจึงสนใจปัจจัยหรือตัวแปรที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนด้วย ผู้วิจัยจึงนำเสนอตอนนี้ออกเป็น 2 ส่วน คือส่วนแรกเป็นปัจจัยหรือตัวแปรที่เกี่ยวกับคุณภาพการศึกษาจากการทบทวนเอกสารของนักการศึกษา และส่วนที่สองเป็นรายงานวิจัยที่ศึกษาเกี่ยวกับการประเมินคุณภาพก ารจัดการศึกษาและผลสัมฤทธิ์การเรียนรู้ ดังนี้

5.1 ปัจจัยหรือตัวแปรที่ส่งผลต่อคุณภาพการจัดการศึกษา

ตัวแปรที่ส่งผลต่อคุณภาพการจัดการศึกษา และผลสัมฤทธิ์ทางการเรียน ที่มีอยู่หลายปัจจัยต่างมีความแตกต่างกันในแต่ละสถานศึกษา ซึ่งตัวแปรเหล่านี้จะนามาพิจารณาเพื่อนำมาใช้กำหนดเป็นโมเดลในการวิเคราะห์หาคุณภาพการจัดการศึกษาของสถานศึกษา ในการพิจารณาตัวแปรเพื่อกำหนดในโมเดลการวิเคราะห์ต้องพิจารณาหลาย ๆ ส่วนประกอบกัน Salganik (1994) ได้เสนอเงื่อนไขในการพิจารณาตัว แปรที่จะต้องกำหนดในแต่ละระดับดังนี้ 1) ข้อมูลระดับบุคคลของนักเรียนจะต้องสามารถหาได้ง่าย 2) คุณลักษณะพื้นฐานของนักเรียนในโมเดลจะต้องมีความสัมพันธ์กับความสามารถของนักเรียน 3) คุณลักษณะของนักเรียนต้องเป็นสิ่งที่เกินกว่าการควบคุมของสถานศึกษา หรือเป็นสิ่งที่สถานศึกษาคาดหวังให้เกิดการเปลี่ยนแปลง 4) คุณลักษณะจะต้องเป็นที่ยอมรับทางกฎหมายที่เกี่ยวข้องกับสิ่งที่การศึกษาต้องเผชิญ เช่นความยากจน ความแตกต่างทางวัฒนธรรม

คุณภาพการจัดการศึกษา โดยส่วนใหญ่พบว่าเป็นคุณภาพที่เกิดจากการดำ เนินงานของสถานศึกษาเพื่อส่งเสริมให้นักเรียนเกิดการเรียนรู้ ดังนั้นคุณภาพการจัดการศึกษาจึงเกี่ยวข้องกับการสร้างองค์ความรู้ และการประยุกต์ใช้โดยผ่านการเรียนรู้ทำ ให้เกิดลักษณะเฉพาะของนักเรียนในการที่จะใช้ชีวิตของตัวเอง และการอยู่ร่วมกับผู้อื่น คุณภาพการจัดการศึกษาตามมุมมองของ OECD (OECD, 2005) ได้นิยามโดยการใช้กรอบความคิดที่แสดงถึงระบบการดำ เนินการของสถานศึกษาที่ แปลงปัจจัยนำ เข้า(input) มาสู่ผลลัพธ์ทางการศึกษา (outcome) มีมุมมองที่แตกต่างกัน 6 ด้านคือ 1) มุมมองทางด้านผลผลิต (productivity) 2) มุมมองเกี่ยวกับประสิทธิผลการสอน (instrumental effectiveness) 3) มุมมองเกี่ยวกับความเหมาะสม (adaptation) 4) มุมมองด้านความเท่าเทียม (equity) 5) มุมเกี่ยวกับประสิทธิภาพ (efficiency) 6) มุมมองด้านความไม่สอดคล้องต่อเนื่องกัน (disjointed) และจากการศึกษาเป้าหมายของยูเนสโก (UNESCO) พบว่าให้ความสำ คัญกับการเรียนรู้ (Learning) เป็นอันดับแรก รวมทั้งให้ความสำ คัญกับความสัมพันธ์ระหว่างผู้เรียนและผู้สอน ปัจจัยนำเข้า (Input) กระบวนการ (Process) สิ่งแวดล้อม (Environment) และผลผลิต (Outputs) ซึ่งส่งผลกระทบทั้งทางบวกและทางลบต่อคุณภาพการจัดการศึกษา โดยมีการเรียนรู้ประกอบด้วยองค์ประกอบที่เกี่ยวข้องกับระดับผู้เรียน (Inner learner level) และองค์ประกอบส่วนที่เกี่ยวข้องกับระบบการจัดการศึกษา (outer system level) ความแตกต่างของคุณภาพการจัดการศึกษา จึงอาจจะเกิดจากองค์ประกอบแต่ละตัว นอกจากนี้ จากการประเมินของ PISA ยังพบข้อมูลที่สำคัญต่อปัจจัยที่ส่งผลกระทบคุณภาพการจัดการศึกษา โดยแบ่งปัจจัยเหล่านี้เป็นสองระดับ คือระดับสถานศึกษา ประกอบด้วยปัจจัยด้านระดับการให้งบประมาณจากรัฐ การแข่งขันของสถานศึกษา ทรัพยากรสถานศึกษา และระดับนักเรียน

ประกอบด้วยตัวแปรด้านเวลาที่ใช้ในการเรียนวิทยาศาสตร์และคณิตศาสตร์ ภูมิหลังทางเศรษฐกิจสังคม (OECD, 2006)

เมื่อพิจารณากระบวนการศึกษาและค วามสัมพันธ์ระหว่างผู้เรียนและผู้สอนที่อยู่ภายในสถานศึกษา ตลอดจนองค์ประกอบอื่นๆ ต่างส่งผลกระทบต่อคุณภาพการจัดการศึกษาได้ในสองส่วนคือ ระดับผู้เรียน ซึ่งเป็นส่วนที่เกี่ยวข้องกับการเรียนรู้แต่ละคน และปัจจัยอื่นที่เป็นสิ่งแวดล้อมที่อยู่รอบตัวนักเรียน เช่นครอบครัว ประสบการณ์หรือเหตุการณ์ที่เอื้อต่อการเรียนรู้ พื้นฐานการเรียน ทัศนคติของนักเรียน และส่วนที่สองเป็นระดับของระบบการศึกษา ซึ่งสถานศึกษามีความรับผิดชอบกับการวางแผน วางนโยบาย ที่ทำให้เกิดการสนับสนุนให้เกิดการเรียนรู้ ดังนั้นการศึกษาในครั้งนี้จึงแบ่งตัวแปรทางการศึกษาออกเป็น 2 ระดับดังนี้

5.1.1 ตัวแปรระดับสถานศึกษา

การดำ เนินงานอย่างมีประสิทธิผลของสถานศึกษาขึ้นอยู่ปัจจัยหรือเงื่อนไขในระดับสถานศึกษาและห้องเรียน (Scheerens, 2001) โดยอิทธิพลในตัวแปรระดับสถานศึกษาจะเป็นตัวบ่งชี้ประสิทธิผลของสถานศึกษาโดยส่วนใหญ่ เช่น แหล่งทรัพยากรของสถานศึกษา ประกอบด้วยสัดส่วนของนักเรียนต่อครูอุปกรณ์เครื่องมือ ห้องสมุด และครูที่ผ่านการฝึกอบรม ความสัมพันธ์ของผู้ปกครองกับห้องเรียนและสถานศึกษา การมีวินัยในชั้นเรียน นอกจากนี้ Briggs & Wohlstetter (2003) ได้เสนอ 8 องค์ประกอบที่จะทำให้การบริหารจัดการของสถานศึกษามีความสำ เร็จ ได้แก่ การกำหนดวิสัยทัศน์ ความเป็นอิสระของการตัดสินใจ อำนาจ ความรู้และทักษะ ข้อสารสนเทศ รางวัล ความเป็นผู้นำ และแหล่งทรัพยากร

นอกจากนี้จากการศึกษาข้อมูลของสถานศึกษาที่ส่งผลต่อการจัดการศึกษาในการประเมิน PISA ประกอบด้วยข้อมูลสามด้านหลักคือด้านแรกเป็นตัวป้อนด้านทรัพยากรสถานศึกษา ได้แก่คุณภาพของครู อัตราส่วนของครูต่อนักเรียน ทรัพยากรการเรียน ด้านที่สองเป็นประสิทธิผลของการเรียนการสอนได้แก่การจัดการในห้องเรียนและกลยุทธ์การสอน และด้านที่สามเป็นประสิทธิผลของสถานศึกษา ได้แก่การบริหารจัดการของสถานศึกษา บรรยากาศของสถานศึกษา การมีอิสระในการบริหารจัดการสถานศึกษา และความเป็นผู้นำทางการศึกษา จากผลการประเมินของ PISA พบว่ามีความผันแปรของผลสัมฤทธิ์ของนักเรียนภายในสถานศึกษาและระหว่างสถานศึกษา พบว่ามีความแตกต่างระหว่างสถานศึกษาและอธิบายความ แตกต่างระหว่างสถานศึกษาได้ ปัจจัยระดับสถานศึกษาแบ่งเป็น 3 ประเภท คือ ทรัพยากรสถานศึกษา (school resources) สภาพแวดล้อมสถานศึกษา (school climate) และนโยบายสถานศึกษา (school policies) ทรัพยากรสถานศึกษาประกอบด้วย ครูที่มีคุณภาพ, ขนาดชั้นเรียน และทรัพยากร

ทางการเรียน รวมทั้งคอมพิวเตอร์ สภาพแวดล้อมสถานศึกษา ประกอบด้วยสภาพความเป็นวินัย การคาดหวังต่อผลสัมฤทธิ์ ความสัมพันธ์ระหว่างครูและนักเรียน การรับรู้ในความเป็นเจ้าของนักเรียน และจริยธรรมของครู นโยบายสถานศึกษาประกอบด้วย การให้การบ้าน เวลาในการสอน ความเป็นอิสระของครู ความเป็นอิสระของสถานศึกษา นอกจากนี้ โครงสร้างขององค์กรของระบบการศึกษาในระดับชาติ ประเภทของสถานศึกษาเช่น สถานศึกษาเอกชน และสถานศึกษาของรัฐต่างมีความสัมพันธ์กับองค์ประกอบของสถานศึกษา โดยส่วนใหญ่พบว่านักเรียนในสถานศึกษาเอกชนมีความสามารถทางวิชาการสูงมาก และเมื่อควบคุมอิทธิพลของเศรษฐานะ (SES) (Matear, 2006; OECD, 2005) หรือในอีกความหมายหนึ่ง สถานศึกษาเอกชนจะมีความสามารถที่ได้เปรียบเนื่องจากคุณลักษณะของนักเรียน และความแตกต่างที่ไม่ใช่การดูแลของสถานศึกษา นักเรียนที่มีเศรษฐานะในระดับกลางในสถานศึกษาเอกชนจะทำ ได้ดีเช่นเดียวกับสถานศึกษาของรัฐ และในทานองเดียวกับนักเรียนที่มีเศรษฐานะสูงหรือต่ำ ซึ่งการประเมินของโครงการประเมินผล PISA พบว่าองค์ประกอบของสถานศึกษา ประเภทของสถานศึกษาและสถานที่ตั้งสถานศึกษามีอิทธิพลอย่างมากต่อความสามารถของนักเรียน (OECD, 2005)

การศึกษาเกี่ยวกับปัจจัยต่อคุณภาพการจัดการศึกษาได้เพิ่มมากขึ้นตั้งแต่ปลายปี 1980 โดย Riddell (2008) ได้ศึกษาที่ผลต่อคุณภาพการจัดการศึกษาของประเทศกำ ลังพัฒนาจากรายงานของ Coleman (1966) และ Plowden (1967) พบว่าเป็นการศึกษาของปัจจัยนำ เข้า (input) ที่นำไปสู่การผลิตเพื่อออกมาเป็ นผลผลิตที่ดีหรือผลลัพธ์ (output) และข้อค้นพบในประเทศกำ ลังพัฒนาพบว่าอิทธิพลอันดับแรกต่อการเรียนรู้ของนักเรียนคือคุณภาพของสถานศึกษาและครู (Heyneman, 1986 cited in Riddell, 2008) และการศึกษาต่อมาจนถึงปัจจุบัน เป็นการศึกษาที่เน้นการสร้างโมเดลที่มีความซับซ้อนข องปัจจัยหลายอย่างที่มีอิทธิพลต่อคุณภาพการศึกษา เป็นการศึกษาเกี่ยวกับความสัมพันธ์ระหว่างตัวแปรที่เป็นปรากฏการณ์ในสถานศึกษา มีลักษณะที่แปรเปลี่ยนตามพื้นฐาน ความแตกต่างของการสอนในชั้นเรียน ความแตกต่างในสถานศึกษา ความแตกต่างของครู และแหล่งทรัพยากร ซึ่งเป็นการศึกษาที่นำไปสู่โมเดลการวิเคราะห์พหุระดับ (Goldstein 2004; Raudenbush and Bryk, 2002) เช่นเดียวกันกับการศึกษาของ Palardy (2008) เกี่ยวกับปัจจัยที่แตกต่างกันในผลลัพธ์ของนักเรียนระหว่างสถานศึกษา ปัจจัยเหล่านี้สามารถจัดเป็นกลุ่มได้สองกลุ่ม คือปัจจัยนำ เข้าของสถา นศึกษา (school inputs) ซึ่งประกอบด้วยคุณลักษณะของนักเรียน ตลอดจนแหล่งทรัพยากรของสถานศึกษาและลักษณะโครงสร้าง นโยบายและการปฏิบัติของโรงเรียน ปัจจัยนำ เข้าของสถานศึกษาเป็นปัจจัยที่ให้กับสถานศึกษา เช่นบุคลากรซึ่งมีแนวโน้มที่มีอิทธิพลเหนืออิทธิพลอื่น นโยบายและการปฏิบัติของสถานศึกษา ประกอบด้วยปัจจัยที่เกี่ยวกับการควบคุม ให้ความสาคัญ

เกี่ยวกับผู้ปฏิบัติในสถานศึกษาและผู้กำหนดนโยบาย (Fitz-Gibbon & Kochan, 2000; Good & Brophy, 1986; Palardy, 2003; Rumberger & Thomas, 2000 cite in Palardy, 2008) และ Riddell (2008) ได้ทบทวนงานวิจัยปัจจัยที่มีอิทธิพลต่อคุณภาพและประสิทธิผลทางการศึกษาในประเทศกาลังพัฒนา ได้แบ่งออกเป็น 4 กลุ่มคือ กลุ่มที่ 1 กลุ่มสถานศึกษา ได้แก่ ภาวะผู้นำที่เข้มแข็ง (Strong leadership) สิ่งแวดล้อมในชั้นเรียนและสถานศึกษา (Orderly school and classroom environment) ครูที่มีจุดเน้นในหลักสูตรพื้นฐาน ความคาดหวังต่อศักยภาพและความสามารถของนักเรียน ความถี่ของการประเมินและการส่งผลย้อนกลับ โครงสร้างการสอน ความเพียงพอของหนังสือและสื่ออุปกรณ์ คุณภาพของครู กลุ่มที่ 2 กลุ่มครู ได้แก่ การติดตามผลความก้าวหน้าของนักเรียน การจัดกลุ่มความสามารถของนักเรียน ความเชี่ยวชาญและทักษะด้านภาษาของครู ความคาดหวังของครูต่อนักเรียน กลุ่มที่ 3 การสอน ได้แก่ วิธีการสอน เป้าหมายการเรียนรู้ การใช้สื่ออุปกรณ์ใหม่ การอธิบายที่ชัดเจน การตรวจสอบความเข้าใจ การจัดเวลาเพื่อให้นักเรียนฝึกทักษะใหม่ ความสมบูรณ์ของชิ้นงานเพื่อการเรียนรู้ ความถี่ของการสอบและการส่งผลย้อนกลับ กลุ่มที่ 4 การกำหนดนโยบาย ได้แก่บทบาทของผู้นำนโยบายเกี่ยวกับการเรียนรู้ จากการศึกษาทบทวนงานวิจัยของ Heck (2000) ซึ่งศึกษาเกี่ยวกับอิทธิพลของสถานศึกษาใน 20 ปีที่ผ่านมาของ Creemers, 1994; Hallinger & Murphy, 1986; Hallinger & Heck, 1998; Heck & Marcoulides, 1996; Leithwood, 1994; McDonnell, 1995; Mortimore, 1993; Reynolds & Packer, 1992; Sammons, et al., 1995; Witte & Walsh, 1990 สรุปได้ว่าตัวแปรที่มีความสำคัญเป็นเงื่อนไขบริบทของสถานศึกษา (Contextual conditions) ประกอบด้วยโครงสร้าง นโยบาย บุคคลและกระบวนการ เช่น ขนาดห้องเรียน คุณภาพและเงินเดือนของครู ข้อมูลพื้นฐาน ภูมิประเทศ ตัวแปรบริบทจึงเป็นสิ่งที่ช่วยอำนวยความสะดวกให้เกิดผลสาเร็จได้ดีขึ้นเมื่อเปรียบเทียบในระหว่างสถานศึกษา โดยตัวแปรสถานศึกษา (School variables) ได้แก่ข้อสารสนเทศเกี่ยวกับการจัดการของสถานศึกษา การกำหนดสัดส่วนของแหล่งทรัพยากร รูปแบบของชั้นเรียน และการสอนนักเรียน นอกจากนี้จากเอกสารเกี่ยวกับประสิทธิผลของการศึกษาของ Creemers ในปี 1994 (cited in Maslowski, et al., (2008) ในเอกสารนี้ได้พัฒนาโมเดลเกี่ยวกับประสิทธิผลการศึกษาอยู่ 4 ขั้น คือบริบท สถานศึกษา ห้องเรียน และชั้นเรียน ตามแนวคิดของ Carroll (1963, 1989) โมเดลของ Creemers เริ่มจากสามองค์ประกอบพื้นฐาน คือ คุณภาพ เวลา และโอกาสในการเรียน ซึ่งโมเดลจะแสดงถึง บริบท (context) ปัจจัยนำเข้า(input) กระบวนการ(process) และผลลัพธ์ (output) ต่างเป็นพื้นฐานของกระบวนอย่างเป็นระบบ และจากโครงการวิจัยประสิทธิผลของสถานศึกษาระดับนานาชาติ (The International School Effectiveness Research Project (ISERP) ที่มีวัตถุประสงค์เพื่อตรวจสอบปัจจัยที่มีความสัมพันธ์กับ ความสามารถของนักเรียนใน

ระหว่างประเทศ (Teddlie, & Reynolds, 2002) ได้ระบุอิทธิพลที่เป็นสากล อิทธิพลที่เป็นลักษณะของประเทศและปัจจัยที่อธิบายประสิทธิผลของบริบท ซึ่งพบว่าคุณภาพการสอน หลักสูตรและการดำเนินการมีความสัมพันธ์กันโดยตรงกับนักเรียน ตลอดจนการให้ข้อมูลย้ อนกลับต่อนักเรียน โครงสร้างของแบบเรียน การตรวจสอบความเข้าใจของนักเรียน การใช้คำถามที่มีคุณภาพ และความคาดหวังสูงต่อนักเรียน

5.1.2 **ตัวแปรระดับนักเรียน**

การศึกษาเกี่ยวกับตัวแปรระดับนักเรียนที่ส่งผลต่อผลการเรียนรู้ จากการศึกษาทบทวนวิจัยของ Raudenbush and Willims (1996 cited in Sheerens and Bosker, 1997) ได้เสนอข้อพึงระวังเกี่ยวกับการกำ หนดตัวแปรเพื่อการประมาณค่าผลสัมฤทธิ์ที่แสดงถึงผลการจัดการศึกษาของสถานศึกษา ไว้ดังนี้ 1) ในบางกรณี ผลลัพธ์ที่ได้อาจมีค่าต่ำกว่าผลที่เป็นจริงของสถานศึกษา เช่นการที่นักเรียนที่มีข้อได้เปรียบทางความสามารถ หรือเศรษฐานะมักจะเลือกสถานศึกษาที่มีการดำเนินการดี ดังนั้นการควบคุมสภาพเศรษฐานะจึงทำ ให้เกิดผลการปรับแก้ที่เกินจริงกับนักเรียนกลุ่มนี้ 2) ระบบการชดเชย (compensatory system) ประสิทธิผลของสถานศึกษา เช่นการชดเชยเพื่อแก้ไขข้อบกพร่องด้วย การจัดการศึกษาที่ดีขึ้นกว่าปกติ ซึ่งอาจจะมีผลให้ผลที่ได้มีค่าลดลง ดังนั้นผลที่เกิดขึ้นจะสามารถสังเกตได้เมื่อวิเคราะห์ผลผลิตที่มีการพยากรณ์ด้วยตัวแปรร่วมที่เกี่ยวข้อง 3) ข้อจำกัดของการกำหนดตัวแปรที่เกี่ยวข้องกับภูมิหลังของนักเรียนได้อย่างครบถ้วน นอกจากนี้ข้อมูลที่ได้จากการวัดที่ขาดความถูกต้อง น่าเชื่อถืออาจมีผลต่อการประมาณค่าในระดับบุคคล

Riddell (2008) พบว่าเศรษฐานะมีผลต่อผลสัมฤทธิ์ทางวิชาการอย่างมีนัยสำคัญ (Sirin, 2005 cited in Riddell, 2008) และจากการศึกษาของ Park (2008) โดยใช้ข้อมูล PISA พบว่าความเกี่ยวข้องระหว่างผู้ปกครอง การสื่อสารระหว่างนักเรียนกับผู้ปกครองมีผลต่อผลสัมฤทธิ์ของนักเรียน นอกจากนี้ จากผลการประเมิน PISA พบว่าผลสัมฤทธิ์ของนักเรียนยังสามารถอธิบายได้ด้วยพื้นฐานทางครอบครัวของนักเรียนแต่ละคนหรือเศรษฐานะทางสังคม (socioeconomic status: SES) ซึ่ง SES สามารถวัดได้หลายแนวทาง และองค์ประกอบของชนชั้นทางสังคมและเศรษฐกิจและรายได้ทางวัฒนธรรม แนวทางที่วัด SES จะเป็นองค์ประกอบของรายได้ของครอบครัว อาชีพของผู้ปกครอง ระดับการศึกษาของผู้ปกครอง โดยเฉพาะอย่างยิ่งอาชีพของแม่ ดัชนีบางอย่างประกอบด้วยการวัดรายได้ทางวัฒน ธรรมโดยเฉพาะเกี่ยวกับระดับของสิ่งแวดล้อมของครอบครัวที่เป็นองค์ประกอบของวัฒนธรรมชั้นสูง องค์ประกอบบางอย่างได้จากจำ นวนหนังสือหรือเครื่องดนตรีในบ้าน การเดินทางพิพิธภัณฑ์ โรงละคร แหล่งหรือสถานที่ฟังเพลง

(OECD, 2004) นอกจากนี้ข้อมูลจากผลการประเมิน PISA ในปี 2003 พบว่า SES มีผลกระทบต่อผลสัมฤทธิ์ของนักเรียน แต่มีอยู่ในระดับที่ต่างกันในแต่ละประเทศ PISA ได้ใช้รวบรวมดัชนี SES จากอาชีพของผู้ปกครอง ระดับการศึกษาของผู้ปกครอง และความเป็นเจ้าของครอบครัวในทางวัฒนธรรม ในประเทศที่เข้าร่วมกับองค์กร OECD ในการประเมิน PISA 2003 ค่าเฉลี่ยของความแปรปรวนของความสามารถของนักเรียนในวิชาคณิตศาสตร์สามารถอธิบายได้โดย SES ประมาณ 20% (OECD, 2004)

ปัจจัยพื้นฐานของนักเรียน (Student background) มีปัจจัยหลายปัจจัย ได้แก่ สิ่งแวดล้อมของบ้านและ SES ที่มีผลต่อคุณภาพของสถานศึกษาต่อความสามารถของนักเรียน (Wiley & Yoon, 1995) ในการพัฒนาการเปรียบเทียบสถานศึกษา จึงมีความสาคัญในการพิจารณาตัวแปรพื้นฐานของนักเรียนด้วยเพื่อปรับค่าเฉลี่ยภายในสถานศึกษาเพื่อแสดงองค์ประกอบของตัวแปรที่มีผลต่อผลสัมฤทธิ์ ในขณะที่ Muthén และเพื่อนได้ให้ความสำคัญในการควบคุมความรู้เดิมหรือความสามารถเดิม (cited in Heck, 2000) และจากการตรวจสอบผลกระทบของคุณภาพสถานศึกษาต่อผลลัพธ์ของสถานศึกษาพบตัวแปรองค์ประกอบของนักเรียนที่พบว่ามีผลกระทบต่อผลสัมฤทธิ์ได้แก่ ผลสัมฤทธิ์เดิม (prior achievement) เพศ (gender) เชื้อชาติ (ethnicity) เศรษฐานะสังคม (SES) พื้นฐานทางภาษา (language background) สถานะทางการศึกษาพิเศษ (special education status) (Sammons et al., 1994 cited in OECD, 2008) เช่นเดียวกันกับ Palardy (2008) พบว่าจากงานวิจัยของ Coleman et al., 1966; Lee & Bryk, 1989; Lee & Smith, 1993, 1995; Lee et al., 1997; McNeal, 1997; Park & Palardy, 2004 พบความหลากหลายของคุณลักษณะของนักเรียน (Student characteristics) มีความสัมพันธ์กับผลสัมฤทธิ์หรือการเรียนรู้ ซึ่งประกอบด้วย คุณลักษณะทางกายภาพ เช่น เชื้อชาติ ศาสนา เพศ คุณลักษณะพื้นฐานของครอบครัว เช่น เศรษฐานะสังคม (SES) โครงสร้างของครอบครัว และพื้นฐานการศึกษา เช่นผลสัมฤทธิ์เดิม หรือพื้นฐานความรู้เดิมของนักเรียน

Teddlie, Stringfield & Reynolds (2000) ได้สรุปผลที่เกิดจากตัวแปรระดับบุคคลด้านภูมิหลังของนักเรียน เกี่ยวกับขนาดของอิทธิพลที่เกิดจากเศรษฐานะสังคม (socioeconomic status: SES) ที่มีผลต่อผลสัมฤทธิ์ว่ามีความแตกต่างกันตามกลุ่มสมาชิกของแต่ละสถานศึกษา นอกจากนี้การศึกษา Chiu & Xihua (2008) พบว่าปัจจัยในด้านครอบครัวก็ส่งผลต่อผลสัมฤทธิ์ของนักเรียน สรุปได้ว่าจำนวนสมาชิกของครอบครัวสามารถตอบสนองแหล่งทรัพยากรภายนอกของนักเรียน (ผู้ให้ทรัพยากร) หรือความสมบูรณ์ของแหล่งทรัพยากร ดังนั้นจำนวนสมาชิกของครอบครัว เช่นพ่อแม่จะมีการให้แหล่งทรัพยากรทางการศึกษาจำ นวนมากซึ่งส่งผลต่อโอกาสในการเรียนรู้มากและนักเรียนก็สามารถไปสู่ความสำ เร็จได้สูง ตัวอย่างเช่น ครอบครัวที่มีทั้งพ่อและ

แม่จะมี SES สูง มีแหล่งทรัพยาการทางการศึกษามากกว่า การใช้เวลากับเด็กมากกว่า และมีความเกี่ยวข้องกันมากกว่ากับสถานศึกษาของเด็ก ในทางตรงกันข้ามผู้ปกครองที่แยกกันอยู่จะมีแหล่งทรัพยากรทางการศึกษาน้อยกว่าและมีการติดต่อพบหน้ากับเด็กน้อย ซึ่งอาจจะได้รับความสนใจใส่ใจน้อย เด็กที่มีความขัดแย้งระหว่างพ่อแม่ที่แยกกันอยู่อาจจะจำ ทนอยู่ในสภาวะอารมณ์และบ่อยครั้งที่พบว่ามีแรงจูงใจทางการศึกษาต่ำ และมีผลสัมฤทธิ์ทางการเรียนต่ำ อีกด้านหนึ่งจำนวนสมาชิกที่เพิ่มขึ้น อาจจะลดการเข้าถึงแหล่งทรั พยากรของเด็ก ที่ให้โอกาสในการเรียนน้อยลงและผลสัมฤทธิ์ของนักเรียนต่ำลง และนักเรียนบางคนอาจจะได้รับผลประโยชน์จากปู่ย่าตายายในด้านกายภาพ ข้อมูลข่าวสาร สังคม และแห่งทางด้านอารมณ์ และอาจจะแสดงถึงผลสำ เร็จสูงขึ้น อีกด้านตรงข้ามนักเรียนที่อยู่กับปู่ย่าตายายที่ยากจนหรือป่วยจะมีข้อจำ กัดในแหล่งทรัพยากรของครอบครัว

นอกจากตัวแปรที่เป็นคุณลักษณะพื้นฐานของนักเรียนและพื้นฐานครอบครัวของนักเรียนที่ส่งผลต่อผลสัมฤทธิ์ของนักเรียน ยังพบตัวแปรแรงจูงใจและเจตคติของนักเรียน จากทฤษฎีสติปัญญาทางสังคม แรงจูงใจทางการเรียนของนักเรียน เชื่อว่ามีสามองค์ประกอบคือ คุณค่า (Value) ความคาดหวัง (expectancy) และความชื่นชอบ (affect) (Bandura, 1989) นักเรียนที่เห็นคุณค่าในการทำ กิจกรรมจะมีลักษณะของ ความพยายามมากกว่าและได้รับความสำ เร็จสูง ความคาดหวังเชื่อว่านักเรียนจะมีความเชื่อในประสิทธิภาพของตนเอง (Self – efficacy) และความเชื่อในความสามารถของตนเอง (Self – concept) นอกจากนี้ ความคาดหวังจะถูกพิจารณาว่าเป็นความสามารถทางการศึกษาที่สามารถมีผลกับประสบการณ์ของนักเรียน โดยเฉพาะอย่างยิ่ง ความสำเร็จที่ผ่านและความล้มเหลวในการทำ ชิ้นงานในการเรียน (Pintrich & Schunk, 2002) เนื่องจากความคาดหวังทั้งปฏิกิริยาทางอารมณ์ต่อชิ้นงานและการปฏิบัติชิ้นงานมีอิทธิพลต่อความสามารถ นักเรียนที่มีผลทางบวกเมื่อทำ งานในชิ้นงานที่เรียนมีแนวโน้มที่จะมีความพยายามในการทำ และส่งผลต่อผลสัมฤทธิ์ทางการเรียนสูงขึ้น

5.2 งานวิจัยที่ศึกษาเกี่ยวกับตัวแปรที่มีอิทธิพลต่อคุณภาพการจัดก ารศึกษาและการใช้โมเดลมูลค่าเพิ่มเพื่อประเมินคุณภาพการจัดการศึกษา

งานวิจัยที่ศึกษาเกี่ยวกับตัวแปรหรือปัจจัยที่ส่งผลต่อคุณภาพการจัดการศึกษา และผลสัมฤทธิ์ทางการเรียนของนักเรียน พบว่าโดยส่วนใหญ่เป็นการศึกษาเกี่ยวกับอิทธิพลของสถานศึกษา ชุมชนและสังคม ข้อมูลพื้นฐานทางครอบครัว และพื้นฐานของนักเรียน และได้เริ่มมีนักการศึกษาคนสนใจศึกษามากขึ้น เนื่องจากสารสนเทศที่ได้ จะนำมาใช้กำหนดเป็นหลักการหรือนโนบายสู่การพัฒนาการศึกษาให้ทัดเทียมกับระดับนานาชาติ โดยใช้เทคนิควิเคราะห์พหุ

ระดับ รวมทั้งการวิเคราะห์มูลค่าเพิ่มเพื่อตรวจส อบประสิทธิผลของสถานศึกษา เพื่อบ่งบอกถึงคุณภาพการจัดการศึกษา เมื่อควบคุมตัวแปรในระดับต่างๆ ดังตัวอย่างงานวิจัยที่เสนอดังนี้ Koutsoulis and Campbell (2001) ศึกษาอิทธิพลของพฤติกรรมและเศรษฐกิจของครอบครัว ที่มีต่อผลสัมฤทธิ์ในวิชาคณิตศาสตร์และวิทยาศาสตร์ และแรงจูงใจของนักเรียนระดับมัธยมศึกษา ที่ประกอบด้วยมโนภาพเกี่ยวกับตนเอง (self-concepts) เจตคติต่อสถานศึกษา (attitude toward school) และความปรารถนา (aspiration) โดยการศึกษานี้อยู่บนพื้นฐานโมเดลของ Walberg (1984a,1984b,1986 cited in Koutsoulis & Campbell, 2001) จากนั้นวิเคราะห์ข้อมูลโดยใช้การวิเคราะห์เส้นทาง (path analysis) ผลการวิจัยนี้ชี้ให้เห็นว่า ตัวทำนายผลสัมฤทธิ์ทางการเรียนของนักเรียนระดับมัธยมศึกษาที่ดีที่สุด คือ ความสามารถเดิม (prior ability) ของนักเรียน มโนทัศน์ของตนเองเกี่ยวกับวิชาคณิตศาสตร์และวิทยาศาสตร์ เป็นตัวทำนายผลสัมฤทธิ์ในวิชาคณิตศาสตร์และวิทยาศาสตร์ที่ดี แรงกดดัน(pressure) จากผู้ปกครองเป็นตัวทำ นายแรงจูงใจ และผลสัมฤทธิ์ในวิชาคณิตศาสตร์และวิทยาศาสตร์ในทางลบ โดยเฉพาะในเพศหญิง การได้รับการสนับสนุนจากผู้ปกครอง (parental psychological support) พบว่า มีอิทธิพลทางตรงต่อแรงจูงใจของนักเรียน และมีผลทางอ้อมต่อผลสัมฤทธิ์ของนักเรียน และพบว่าเศรษฐานะทางสังคม (socioeconomic) มีความสำคัญมาก เนื่องจากมีอิทธิพลทางตรงต่อความต้องการทางด้านการศึกษา (educational aspirations) แรงกดดัน การส่งเสริมทางจิตใจ และมีอิทธิพลทางอ้อมต่อความรู้เดิม ความต้องการทางด้านการศึกษา เจตคติต่อสถานศึกษา มโนภาพของตนเองเกี่ยวกับวิชาคณิตศาสตร์และวิทยาศาสตร์ และผลสัมฤทธิ์ในวิชาคณิตศาสตร์และวิทยาศาสตร์

Chiu & Xihua (2008) ได้ศึกษาอิทธิพลทางครอบครัวและแรงจูงใจต่อผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์จากการวิ เคราะห์ข้อมูลนักเรียน 41 ประเทศ โดยการศึกษาตัวแปรสามกลุ่มคือตัวแปรระดับประเทศ ประกอบด้วย GDP , GDP Gini, และคุณค่าทางวัฒนธรรม ตัวแปรครอบครัว ประกอบด้วย การย้ายถิ่นเข้ามาอยู่เป็นกลุ่มแรก การย้ายถิ่นเข้ามาอยู่เป็นอันดับที่สอง ภาษาที่ใช้พูดคุยในบ้าน เศรษฐานะทางสังคม มีพ่อหรือแม่เพียงคนเดียว มีผู้ปกครองผสม ไม่ได้อยู่กับพ่อแม่ อยู่อาศัยกับปู่ย่าตายายเพียงคนเดียว จำนวนลูกพี่ลูกน้อง และลำ ดับการเกิด คุณลักษณะอื่นของคร อบครัวประกอบด้วย จำนวนหนังสือที่มีอยู่ในบ้าน ความเป็นเจ้าของวัฒนธรรม และรูปแบบการติดต่อสื่อสารทางวั ฒนธรรม และกลุ่มที่สามเป็นตัวแปรของนักเรียน ประกอบด้วย เพศ ตัวแปร แรงจูงใจ ประกอบด้วย ความสนใจในวิชาคณิตศาสตร์ แรงจูงในการใช้เป็นเครื่องมือ ความพากเพียรพยายาม ความเชื่อในประสิทธิภาพของตนเอง ความเชื่อในความสามารถของตนเอง และปฏิสัมพันธ์ระหว่างตัวแปร ผลจากการศึกษาพบว่าความแปรปรวนของคะแนนสอบคณิตศาสตร์ของนักเรียนมีค่าเกือบครึ่งของความแปรปรวนทั้งหมด (44%) และ

ความแปรปรวนระดับสถานศึกษาเท่ากับ 25% และที่เหลือเป็นความแปรปรวนระดับประเทศเท่ากับ 31% ผลการศึกษาในตัวแปรต่างๆ พบว่าการย้ายเข้าเป็นกลุ่มแรกและการย้ายเข้ามาเป็ นกลุ่มที่สองโดยเฉลี่ยมีคะแนนคณิตศาสตร์ต่ำ กว่านักเรียนที่เกิดในชาตินั้นอยู่ 12 และ 5 คะแนนตามลำดับ นักเรียนที่พูดภาษาในบ้านแตกต่างจากที่สถานศึกษาโดยเฉลี่ยมีคะแนนคณิตศาสตร์ต่ำกว่านักเรียนกลุ่มอยู่ 11 คะแนน โดยเฉลี่ยนักเรียนมีคะแนนขึ้น 5 คะแนนในครอบครัวที่มี SES สูงขึ้น 10% การลงทุนและการมีส่วนเกี่ยวข้องของครอบครัวมีผลต่อคะแนนสอบคณิตศาสตร์ นักเรียนที่มีจานวนหนังสือมากมีแนวโน้มที่จะมีคะแนนสูง โดยเฉลี่ยนักเรียนจะมีคะแนนสูงขึ้น 1 คะแนนในการเพิ่มขึ้น 10% ของการเป็นเจ้าของทางวัฒนธรรม และเมื่อควบคุมจำนวนหนังสือที่บ้าน สัมประสิทธิ์ถดถอยของการย้ายมาเป็นกลุ่มแรก จะลดลง 34% สัมประสิทธิ์ถดถอยของการย้ายมาอยู่เป็นกลุ่มที่สองจะลดลง 26% เศรษฐานะของครอบครัว(SES) จะลดลง 32% การอยู่กับพ่อหรือแม่อย่างเดียวจะลดลง 29% และการอยู่กับพ่อแม่ผสมจะลดลง 26% ส่วนผลของตัวแปรระดับนักเรียนพบว่า นักเรียนชายมีคะแนนสูงกว่านักเรียนหญิง 18 คะแนน เมื่อควบคุมตัวแปรเพศจะลดสัมประสิทธิ์ถดถอยของการอยู่ผู้ปกครองผสม 25 โดยเฉลี่ยนักเรียนจะมีคะแนนสูงขึ้น 2 คะแนนในคณิตศาสตร์ที่การเพิ่มขึ้น 10% ของการสนใจในคณิตศาสตร์ อิทธิพลของผู้ปกครองผสมต่อคะแนนสอบคณิตศาสต ร์จะลดลง 21% เมื่อควบคุมตัวแปรความสนใจต่อคณิตศาสตร์ โดยเฉลี่ยนักเรียนมีคะแนนเพิ่มขึ้น 1 คะแนนในการเพิ่มขึ้น 10 % ของความพากเพียรพยายาม โดยเฉลี่ยนักเรียนมีคะแนนสูงขึ้น 1 คะแนนและ 3 คะแนนคณิตศาสตร์ในการเพิ่มขึ้น 10% ของความเชื่อในตนเองและความเชื่อในมโนทัศน์ต นเอง ตามลาดับ เมื่อควบคุมความเชื่อในมโนทัศน์ตนเอง อิทธิพลของความสนใจวิชาคณิตศาสตร์จะลดลง 31% และอิทธิพลของผู้ปกครองผสมไม่มีนัยสำคัญ

Willms and Somers (2001) ได้ศึกษากับนักเรียนของสถานศึกษาในกลุ่มประเทศละตินอเมริกาจำนวน 13 ประเทศๆ ละ 100 โรง พบว่า ปัจจัยตัวแปรระดับบุคคลด้วยสถานะทางวัฒนธรรมทางสังคม (sociocultural status) ที่ส่งผลทางบวกอย่างมีนัยสำ คัญต่อผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ภาษา(สเปน) และจำนวนปีในการศึกษาของนักเรียนในหลายประเทศได้แก่ การศึกษาอย่างต่อเนื่องของผู้ปกครอง และความเอาใจใส่ในแต่ละวัน (daycare) ซึ่งตัวแปรหลังนี้ยังมีผลทางบวกต่อการลดอัตราการตกซ้ำ ชั้นของนักเรียนและจำ นวนปีในการศึกษาอย่างมีนัยสำ คัญ และตัวแปรที่ส่งผลเชิงลบคือ การช่วยเหลือด้านการบ้าน สอดคล้องกับผลการศึกษาของ Ho and Willms (1996 cited in Willms and Somers,2001) ผลการศึกษายังแสดงให้เห็นถึงผลจากปัจจัยตัวแปรระดับห้องที่ส่งผลทางบวกต่อผลสัมฤทธิ์ ได้แก่ การไม่แบ่งกลุ่ม / ชั้นตามระดับความสามารถของนักเรียน มีการทดสอบอย่างปกติสม่ำ เสมอ บรรยากาศในชั้นเรียน

การมีส่วนร่วมที่ดีของผู้ปกครอง พบว่า มีอิทธิพลค่อนข้างสูงเมื่อพิจารณาในระดับสถานศึ กษา สำหรับตัวแปรอื่นในระดับสถานศึกษา ได้แก่ จำนวนวัสดุอุปกรณ์การสอนที่ส่งผลทางบวกอย่างมีนัยสำคัญต่อวิชาภาษา และจำนวนนักเรียนต่อห้องพบว่ามีผลทางลบ

Heck (2000) ได้ศึกษาปัจจัยในด้านบริบท และการปฏิบัติของสถานศึกษาที่มีต่อผลสัมฤทธิ์และความก้าวหน้าด้านวิชาการ กับนักเรียนจาก 122 สถานศึกษา โดยผลสัมฤทธิ์ที่นำมาศึกษาเป็นผลสัมฤทธิ์ความสามารถวิชาคณิตศาสตร์ การอ่าน และการใช้ภาษาในเกรด 3 และเกรด 6 ผลพบว่าค่าสหสัมพันธ์ intra class มีขนาดใหญ่ ซึ่งอธิบายร้อยละของความแปรปรวนระหว่างกลุ่ม ค่าสหสัมพันธ์ intra class correlations ของนักเรียนเกรด 6 ของการอ่านเป็น .194 คณิตศาสตร์เป็น .231 การอ่านเป็น .190 สำหรับนักเรียนคนเดียวกันในเกรด 3 ค่าสหสัมพันธ์ intra class correlations ของการอ่านเป็น .210 คณิตศาสตร์เป็น .201 ภาษาเป็น .179 นอกจากนี้ชี้ให้เห็นว่านักเรียนชุดเดียวกันจะมีความแปรป รวนระหว่างสถานศึกษาของผลการเรียนรู้ โดยมีความแปรปรวนระหว่างสถานศึกษาเป็น 10% ถึง 20% และพิสัยของค่าเศษเหลือในคะแนนรวมการอ่านเป็น –9.9 ถึง 15.5 คะแนนคณิตศาสตร์เป็น –27.6 ถึง 33.7 และในคะแนนรวมการอ่านมีพิสัยเป็น –11.8 ถึง 11.0 ซึ่งพิสัยของค่าเศษเหลือนี้จะบ่งชี้ได้ว่ามีความผันแปรอย่างมากในวิชาคณิตศาสตร์ในสถานศึกษามากกว่าค่าเศษเหลือในวิชาการอ่านและภาษาในสถานศึกษา ความผันแปรระหว่างสถานศึกษาต่ำ กว่าคะแนนที่คาดหวังในคะแนนการอ่านเป็น 57% ในคะแนนคณิตศาสตร์เป็น 51% และในคะแนนภาษาเป็น 52.5% ชี้แนะได้ว่ามีสถานศึกษาหลายสถานศึกษาในกลุ่มตัวอย่างที่มีคะแนนต่ำ กว่าคะแนนที่คาดหวังในองค์ประกอบของนักเรียนโดยเฉพาะการอ่าน

Tomasz and Dhawan-Biswal (2008) ได้ตรวจสอบทักษะในการพัฒนานักเรียน โดยใช้ข้อมูลจากการประเมินโครงการประเมินผล PISA 2000 และการสำรวจเด็กเข้าเมือง (YITS) เพื่อศึกษาว่า อายุของนักเรียนที่ย้ายถิ่นในประเทศแคนาดาที่มีผลต่อความสามารถทางวิชาการในสถานศึกษา ความแตกต่างของภาษาที่ใช้ในบ้านที่มีผลต่อทักษะการอ่านในภาษาอังกฤษและภาษาฝรั่งเศสของนักเรียนย้ายถิ่น ตรวจสอบหาตัวแปรครอบครัวและตัวแปรในสถานศึกษาที่เป็นอิทธิพลต่อความสามารถทางวิชาการของนักเรียนแต่ละคน และศึกษาถึงความสามารถในการอ่านของนักเรียนที่มีความผันแปรในระหว่างสถานศึกษา แบ่งกลุ่มประชากรออกเป็น 3 กลุ่มดังนี้ กลุ่มนักเรียนที่มีถิ่นเดิม (native born) นักเรียนที่ย้ายถิ่นรุ่นที่ 1 (first generation) และนักเรียนที่ย้าย

เข้ามาอยู่ ใช้การวิเคราะห์ HLM เพื่อหาความผันแปรของอิทธิพลสถานศึกษาในทักษะการอ่าน ด้วยการควบคุมคุณลักษณะของนักเรียน ในโมเดล 5 โมเดลที่แตกต่างกันพบผลที่ได้ดังนี้ 1) โมเดล 1 เป็น null model ซึ่งเป็นโมเดลที่ไม่รวมตัวแปรที่ใช้ในการอธิบายทั้งในระดับบนักเรียน และระดับสถานศึกษา พบว่าแหล่งความผันแปรในระดับนักเรียน เป็น 82% และความผันแปรระหว่างสถานศึกษา มี 18% ของ 2) โมเดล 2 เป็นโมเดลที่ตรวจสอบความแตกต่างในคะแนนการอ่านระหว่าง first generation และ immigrants ที่แปรค่าในระดับนักเรียนและระดับสถานศึกษา พบว่าความแตกต่างระหว่าง first generation และชาวแคนาดาไม่มีนัยสำคัญทางสถิติที่แตกต่างจากศูนย์ในระดับนักเรียน แต่มีความผันแปรในระหว่างสถานศึกษาอย่างมีนัยสำ คัญ 3) โมเดล 3 เป็นโมเดลที่นำ คุณลักษณะของบุคคล ประกอบด้วยคุณลักษณะของครอบครัว (family characteristics) เศรษฐานะสังคม (SES) และเพศ ซึ่งเป็นความผันแปรเชิงสุ่มในระดับสถานศึกษา เมื่อมีการควบคุมคุณลักษณะแล้ว ความแตกต่างของการอ่านของ first generation และ immigrants ยังคงมีเหมือนเดิมและมีความผันแปรอย่างมีนัยสำคัญในระหว่างสถานศึกษา 4) โมเดล 4 เป็นโมเดลความแตกต่างเฉลี่ยใน first generation และ immigrants ในระดับสถานศึกษา โดยรวมตัวแปรศาสนาในระดับสถานศึกษา พบว่ามีความแตกต่างระหว่างคะแนนการอ่านระหว่าง first generation และชาวแคนาดา และมีความผันแปรระหว่างศาสนา และ 5) โมเดล 5 เป็นการเพิ่มการควบคุมค่าเฉลี่ยของเศรษฐานะสังคมของสถานศึกษา เพื่ออธิบายความแตกต่างในระดับสถานศึกษา พบว่าค่าเฉลี่ย SES ของสถานศึกษามีนัยสำคัญทางสถิติ ในความแตกต่างของคะแนนการอ่านระหว่างสถานศึกษา

Schulz (2005) เป็นการศึกษาอิทธิพลของเศรษฐานะทางสังคมของครอบครัวนักเรียนต่อผลการเรียนรู้ โดยการเปรียบเทียบในระดับนานาชาติ PISA เพื่อใช้อธิบายความแตกต่างโครงสร้างของระบบทางการศึกษาที่สามารถส่งผ่านผลกระทบของพื้นฐาน SES ของครอบครัว โดยการวัด SES ในบริบทของการศึกษาของ OECD/ PISA ใช้ดัชนีของเศรษฐานะทางสังคมจากอาชีพของผู้ปกครอง ระดับการศึกษาของผู้ปกครองและความเป็นเจ้าของ จากการศึกษาแสดงให้เห็นว่าความแตกต่างของความสามารถของนักเรียน เนื่องจากพื้นฐาน SES โดยการวิเคราะห์พหุระดับ ในการวิเคราะห์อิทธิพลของ SES ต่อความสามารถในการอ่านในปี 2000 กับ 2003 ของ PISA ด้วยโมเดลถดถอยหนึ่งระดับพบว่ามีลักษณะรูปแบบคล้ายคลึงกันภายในประเทศ และการใช้โมเดลสองระดับเพื่ออธิ บายความสามารถในการอ่านกับดัชนีเศรษฐานะ (ESCS) ภายในประเทศในปี 2000 กับ 2003 พบว่ามีอิทธิพลของ ESCS ในระดับนักเรียนและระดับสถานศึกษาและสามารถใช้อธิบายคุณลักษณะของระบบทางการศึกษา และเมื่อเปรียบเทียบผลในปี 2003 และปี 2000 พบว่ามีอิทธิพลคล้ายกันในหลายประเทศ

Willms (2004) ได้ศึกษาผลสัมฤทธิ์ในการอ่านของประเทศแคนาดาและประเทศสหรัฐอเมริกา เพื่อตรวจสอบการแจกแจงคะแนนในระดับนักเรียนและสถานศึกษา ประมาณค่าเศรษฐานะของครอบครัวที่เกี่ยวข้องกับความสามารถในการอ่านในแคนาดาและสหรัฐอเมริกาและการตรวจสอบความสัมพันธ์ระหว่างความสาม ารถในการอ่านและเศรษฐานะของครอบครัวภายในและระหว่างสถานศึกษา และการเปรียบเทียบปัจจัยทางด้านครอบครัวและโรงเรียนและความสามารถในการอ่านของความแตกต่างในสถานศึกษาประเทศแคนาดาและสหรัฐอเมริกา โดยการใช้ข้อมูลจากการศึกษา PISA ปี 2000 การศึกษาได้กำหนดคุณลักษณะความแตกต่างในพื้นฐานครอบครัวของนักเรียน อิทธิพลของบริบท และปัจจัยเกี่ยวกับแหล่งทรัพยากรของสถานศึกษาและนโยบายและแนวปฏิบัติของสถานศึกษาและชั้นเรียน ใช้การวิเคราะห์ HLM สองระดับในโมเดลที่แตกต่างกัน พบผลดังนี้ 1) โมเดล 1 เป็น null model เป็นโมเดลที่ยังไม่นำตัวแปรนักเรียนและสถานศึกษาเข้าพิจารณา เพื่อประเมินความแตกต่างระหว่างประเทศแคนาดาและสหรัฐอเมริกา มีค่าประมาณความแตกต่างเป็น 31.7 คะแนน คะแนนความเคลาดเคลื่อนมาตรฐาน 4.1 2) โมเดล 2 เป็นการรวมตัวแปรคุณลักษณะนักเรียนสามตัวแปรได้แก่ เพศ เศรษฐานะของครอบครัว (SES) และแหล่งภูมิลำเนา(foreign-born) รวมทั้งปฏิสัมพันธ์ของตัวแปรด้วย โดยอิทธิพลของความแตกต่างระหว่างสองประเทศมีค่าสัมประสิทธิ์ของหญิงเป็น 29.3 ซึ่งมีค่าเฉลี่ยของหญิงอยู่สูงกว่าชาย ส่วนสัมประสิทธิ์ปฏิสัมพันธ์พบว่าไม่มีนัยสำ คัญ ดังนั้นขนาดของความแตกต่างของเพศมีลักษณะเหมือน กันในทั้งสองประเทศ ค่าประมาณของค่าเฉลี่ยของความชันของ SES เป็น 33.5 ความแตกต่างระหว่างสถานที่ของนักเรียนที่เกิดในประเทศและเกิดต่างประเทศมีความแตกต่าง 5.9 3) โมเดล 3 เป็นโมเดลที่รวมชุดตัวแปรค่าเฉลี่ยของสถานศึกษามีการประมาณค่าของอิทธิพลของบริบทของสองประเทศ การประมาณค่าของสหรัฐอเมริกาเป็น 63.8 แสดงว่านักเรียนที่ค่าเฉลี่ยในคุณลักษณะที่สามารถแสดงได้ 63.8 คะแนนและในแคนาดามีค่าประมาณเป็น 18 คะแนน ผลแสดงให้เห็นว่าทั้งสองประเทศ ได้รับประโยชน์เมื่อสถานศึกษามี SES สูงถึงแม้ว่านักเรียนแต่ละคนจะมีพื้นฐานครอบครัวที่ต่าง กัน การควบคุมค่าเฉลี่ย SES ของสถานศึกษาเป็นการลดการประมาณความแตกต่างของแคนาดาและสหรัฐอเมริกาได้มากกว่า 3 คะแนน

Patrick and Tomasz (2004) ได้ศึกษาผลกระทบของของการใช้คอมพิวเตอร์ต่อผลสัมฤทธิ์ด้านการอ่านของนักเรียนอายุ 15 ปี จุดประสงค์ของการศึกษาครั้งนี้เป็นก ารศึกษาผลกระทบของ เทคโนโลยีการสื่อสารและสารสนเทศ (ICT) ต่อความสามารถในการอ่านของนักเรียนอายุ 15 ปีของชาวแคนาดา โดยการใช้ข้อมูลจากโครงการประเมินผล PISA 2000 โดยพิจารณาความสนใจในคอมพิวเตอร์ โดยจัดลักษณะของตัวแปรเป็น 4 ตัว จะแบ่งออกตามจังหวัด

และเพศ ได้แก่ การเข้าถึงคอมพิวเตอร์และอินเตอร์เน็ต ความถี่ของการใช้คอมพิวเตอร์ รูปแบบการใช้ และทัศนคติต่อ ICT จากผลการวิเคราะห์พหุระดับ เพื่อควบคุมตัวแปร ใน 5 โมเดลที่มีทั้งตัวแปรระดับบุคคลและระดับครอบครัว พบว่าความสามารถของการใช้คอมพิวเตอร์มีความสัมพันธ์สูงที่สุดในการศึกษาครั้งนี้ และช่วยลดช่องว่างของเพศ เพื่อพิจารณาที่ระดับบุคคล ครอบครัว จังหวัด พบว่าในโมเดลพื้นฐานอธิบายคะแนนการอ่านได้ 32.4% และเมื่อเพิ่มตัวแปรการใช้ ICT จะอธิบายได้เพิ่มเป็น 37.3% และตัวแปรที่เกี่ยวกับการใช้คอมพิวเตอร์พบว่ามีนัยสำ คัญ การใช้คอมพิวเตอร์ที่ส ถานศึกษามีความสัมพันธ์ทางบวก แต่มีค่าไม่สูงเท่าการใช้ที่บ้าน และความสัมพันธ์เป็บค่าลบกับการใช้คอมพิวเตอร์ที่ห้องสมุดและที่อื่นๆ สัมประสิทธิ์ของดัชนีของความสนใจในคอมพิวเตอร์มีค่าเป็นลบและมีนัยสำคัญ และดัชนีความสามารถการใช้คอมพิวเตอร์มีค่าเป็นบวกสูงมากกับคะแ นนการอ่าน ตัวแปรเพศมีปฏิสัมพันธ์ทางลบ ผลที่ได้ชี้ให้เห็นว่าความสามารถการใช้คอมพิวเตอร์สูงมีประโยชน์สำ หรับชายมากกว่าหญิง และความถี่ของการใช้คอมพิวเตอร์สำหรับการเล่นเกมส์และการเข้าอินเตอร์เน็ตไม่มีนัยสำคัญกับทักษะการอ่าน

Opdenakker and Van Damme (2000) ศึกษาอิทธิพลที่เกิดจากสถานศึกษาการสอนของครู และบรรยากาศของชั้นเรียนที่มีต่อผลสัมฤทธิ์ทางการเรียนซึ่งวัดจากแบบสอบที่ร่วมกับคณะครูผู้สอนในสถานศึกษาเป็นผู้สร้างเพื่อให้มีความตรงตามเนื้อหาของหลักสูตรพื้นฐานของแต่ละสถานศึกษา และต่อความรู้สึกที่มีต่อสถานศึกษา กลุ่มตัวอย่างที่ใช้ในการศึกษาคือนักเรียนชั้นมัธยมศึกษาปีที่ 1 จานวนสถานศึกษา52 แห่ง ผลการศึกษาพบว่าปัจจัยด้านบรรยากาศของชั้นเรียนและสถานศึกษาส่งผลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียนมากกว่าความรู้สึกต่อสถานศึกษา ในขณะที่บรรยากาศที่เป็นวิชาการ เช่น ความร่วมมือของครูในการพัฒนาการสอน การประชุม/อภิปรายทางวิชาการอย่างสม่าเสมอส่งผลที่ดีต่อนักเรียน ยกเว้นนักเรียนที่มีแรงจูงใจใฝ่สัมฤทธิ์ต่ำ

D'Agostino (2000) ได้ศึกษาอิทธิพลของการเรียนการสอน และสถานศึกษาที่มีต่อผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ และการอ่านในระยะยาว ซึ่งมีวัตถุประสงค์เพื่อศึกษาอิทธิพลของรูปแบบการจัดการเรียนการสอน และการบริหารจัดการสถานศึกษาที่มีต่อผลสัมฤทธิ์ทางการเรียนในวิชาคณิตศาสตร์และการอ่านของนักเรียนระดับ 1 ถึง 3 มีตัวแปรที่เข้าร่วมในการศึกษาคือตัวแปรระดับนักเรียนได้แก่ เพศ ระดับการศึกษา อาชีพ รายได้ของครอบครัว แหล่งข้อมูลสนับสนุนการศึกษา ตัวแปรระดับครูได้แก่การจัดการสอนคณิตศาสตร์และการอ่าน ตัวแปรระดับสถานศึกษาได้แก่ ขนาดสถานศึกษา ความเป็นผู้นำทางวิชาการ ระดับความยากจนของสถานศึกษา การสนับสนุนของผู้ปกครอง ผลการศึกษาพบว่า การทดสอบโมเดลการจัดการศึกษา

ของสถานศึกษาทั้งสามรูปแบบด้วยโมเดล HLM เมื่อควบคุมเศรษฐานะของผู้ปกครอง พบว่าตัวแปรความเป็นผู้นำทางวิชาการ การสนับสนุนของผู้ปกครอง ทาให้ครูจัดการเรียนการสอนได้อย่างมีประสิทธิภาพ และนักเรียนมีผลสัมฤทธิ์ทางการเรียนสูงขึ้น

Silin and Murray-Harvey (1999) ได้ศึกษาปัจจัยที่เกี่ยวข้องกับผลการเรียนของนักเรียน ซึ่งศึกษากับนักเรียนระดับประถมศึกษา จำนวน 30 แห่ง จาแนกตามสังกัด และที่ตั้งของสถานศึกษา ตัวแปรระดับนักเรียนได้แก่ เศรษฐานะ ความเชื่อในมโนทัศน์ของตนเอง (self-concept) เจตคติและวิธีการเรียนรู้ ผลการศึกษาพบว่า ตัวแปรโดยรวมสามารถอธิบายความแตกต่างของผลสัมฤทธิ์ระหว่างสถานศึกษาค่อนข้างสูง และพบว่าตัวแปรด้านเศรษฐานะที่ส่งผลโดยตรงต่อผลสัมฤทธิ์ของสถานศึกษาค่อนข้างสูง และส่งผลทางอ้อมผ่านตัวแปรด้านสังกัด และขนาดของสถานศึกษาในระดับปานกลาง และส่งผ่านจากสังกัดไปสู่ขนาดสถานศึกษา ระดับเล็กน้อย แสดงให้เห็นว่านักเรียนที่มีเศรษฐานะดีที่ไม่เลือกเรียนสถานศึกษาเอกชนก็จะเลือกเรียนในสถานศึกษารัฐบาลขนาดใหญ่

Hofman, Hofman and Guldemon (1999) ได้ศึกษาปัจจัยทางด้านบริบทที่ส่งผลต่อการเรียนรู้ของนักรียน โดยศึกษากับนักเรียนระดับ 8 จากสถานศึกษา 103 แห่ง พบว่านโยบายการบริหารงานที่แตกต่างกันตามสังกัดของสถานศึกษาส่งผลต่อผลสัมฤทธิ์ทางวิชาการของนักเรียนในวิชาคณิตศาสตร์มากกว่าที่ไม่ใช่ด้านวิชาการ และพบว่านักเรียนในสังกัดสถานศึกษารัฐบาลมีความรู้สึกที่ดีต่อสถานศึกษามากกว่านักเรียนในสถานศึกษาเอกชน ในขณะที่ค่าเฉลี่ยผลสัมฤทธิ์ทางการเรียนต่ำ กว่า และให้ข้อคิดเห็นว่าอาจเกิดจากระบบการคัดเลือกเข้าศึกษาที่ต่างกัน เนื่องจากสถานศึกษาเอกชนมีเกณฑ์การคัดเลือกที่เข้มงวดตามคุณสมบัติเบื้องต้นที่สถานศึกษาต้องการ ซึ่งตรงข้ามกับสถานศึกษารัฐบาลที่นักเรียนเลือกเข้าเรียนตามความ สมัครใจ ผลก็คือนักเรียนในสถานศึกษารัฐบาลจึงค่อนข้างมีความสามารถที่หลากหลายเป็นสาเหตุให้ผลสัมฤทธิ์โดยเฉลี่ยต่ำกว่าสถานศึกษาเอกชน แต่มีความรู้สึกที่ดีต่อสถานศึกษามากกว่า

Strand (1998) ได้ทำการวิเคราะห์มูลค่าเพิ่มในสถานศึกษาระดับประถมศึกษา นักเรียนระดับ 6 จำนวนสถานศึกษา 57 แห่ง โดยใช้ผลการทดสอบวัดความสามารถทางเหตุผลที่ไม่ใช้ภาษา ซึ่งมีการสอบก่อนและระหว่างเรียนระดับ 6 ของกลุ่มตัวอย่างเพื่อวัดระดับความสามารถในการได้รับการพัฒนา และใช้เป็นข้อมูลผลสัมฤทธิ์เดิมของนักเรียน และผลสัมฤทธิ์ครั้งหลังเป็นผลการสอบ KS-2 ของนักเรียนระดับ 6 ซึ่งข้อจำกัดของการศึกษาคือขนาดกลุ่มตัวอย่างมีความแตกต่างค่อนข้างมาก ผลการศึกษาพบว่าที่ระดับความสามารถเชิงเหตุผลเดียวกัน นักเรียนที่มีอายุเฉลี่ยมากกว่ามีผลสัมฤทธิ์สูงกว่านักเรียนที่มีอายุน้อยที่เรียนในสถานศึกษาที่มีอายุเฉลี่ยของนักเรียนมากกว่า ซึ่งแสดงให้เห็นถึงสภาพความสัมพันธ์ของกลุ่มนักเรียนที่อยู่ห้องเรียนเดียวกัน

และพบว่านักเรียนที่อยู่ในสถานะได้เปรียบด้านเศรษฐานะ และระดับสติปัญญามีผลสัมฤทธิ์ของนักเรียนดีกว่าอย่างมีนัยสำคัญทางสถิติ

Corbo and Nicholas (1984) ได้ศึกษาทัศนคติในวิชาคณิตศาสตร์ และผลสัมฤทธิ์ทางการเรียนของนักเรียนระดับ 5-7 ในโรงเรียนของรัฐเพนซิลวาเนีย จำนวนทั้งสิ้น 600 คน การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อตรวจสอบ สำรวจความสัมพันธ์ของ 3 ตัวแปร ได้แก่ เพศ ระดับชั้น และผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์กับทัศนคติทางคณิตศาสตร์ เครื่องมือที่ใช้ในการวิจัยประกอบด้วย The Iowa Tests of Basic Skills และ A Semantic Differential Attitude Scale) ผลการวิจัยพบว่า 1. เพศเป็นปัจจัยที่ไม่มีนัยสำคัญทางสถิติ ในการพัฒนาทัศนคติทางวิชาคณิตศาสตร์ของความสัมพันธ์ระหว่างทัศนคติ กับผลสัมฤทธิ์ทางการเรียน 2. ทัศนคติโดยทั่วไปสัมพันธ์เพียงเล็กน้อย กับผลสัมฤทธิ์ทางการเรียนของนักเรียน 3. ระดับชั้นของนักเรียนเป็นปัจจัยที่มีนัยสำคัญทางสถิติทั้งในการพัฒนาทัศนคติของนักเรียน และพลังอำนาจของความสำพันธ์ระหว่างทัศนคติ-ผลสัมฤทธิ์ทางการเรียน

Trautwein, Ludtke, Schnyder and Niggli (2006) ได้ศึกษางานวิจัยมีวัตถุประสงค์เพื่อสำรวจความแตกต่างทางเพศ ที่มี ต่อผลสัมฤทธิ์ทางการเรียนของนักเรียนระดับ 5-6 และเพื่อจำแนกปัจจัยที่สามารถอธิบายความแปรปรวนได้ ผลการวิจัยพบว่ามีความแตกต่างกันระหว่างเพศชายและเพศหญิงในระดับ 6 นอกจากนั้น ผลจากสมการถดถอยได้พบว่า มี 3 ตัวแปรที่เป็นตัวแปรทำนายที่สำคัญของผลสัมฤทธิ์ทางการเรียนของเด็กชายและเด็กหญิง ตัวแปรเหล่านี้ได้แก่ เจตคติในวิชาคณิตศาสตร์ โรงเรียนที่นักเรียนเข้าศึกษา ความคาดหวังของเพื่อน ความคาดหวังของผู้ปกครอง และความถี่ของการทำแบบฝึกหัด

บุญเรือง ศรีเหรัญ (2542) ทางการศึกษานักเรียนชั้นมัธยมศึกษาปีที่ 3 ของสถานศึกษาขนาดกลางขึ้นไป จำนวนสถานศึกษา 54 แห่ง เพื่อศึกษาปัจจัยที่ส่งผลต่อผลสัมฤทธิ์เชิงวิชาการรายวิชาวิทยาศาสตร์ และคณิตศาสตร์ (ผลการเรียนเฉลี่ยระดับชั้น ม. 3) และผลสัมฤทธิ์ที่ไม่ใช่ด้านวิชาการ ได้แก่ เจตคติและแรงจูงใจในการเรียน ลักษณะความเป็นผู้นำ และการปฏิบัติงานรวมกับผู้อื่น รวบรวมข้อมูลด้วยแบบสอบถาม ผลการศึกษาพบว่าปัจจัยระดับบุคคลซึ่งประกอบด้วย ความรู้เดิมของนักเรียน (คะแนน GPA ระดับ ม. 1-2) บรรยากาศและสภาพแวดล้อมภายในครอบครัว ระดับห้องเรียน ได้แก่ ขนาดห้องเรียน ความสัมพันธ์ระหว่างนักเรียน และประสบการณ์การสอนของครู และระดับสถานศึกษา ได้แก่ ขนาดสถานศึกษา ค่าใช้จ่ายทางการศึกษา ประสิทธิภาพการจัดการเรียนการสอนและคุณภาพของการใช้หลักสูตรส่งผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์และวิทยาศาสตร์ เจตคติและแรงจูงใจต่อการเรียน และที่ส่งผลต่อความเป็นผู้นำ ประกอบด้วย ปัจจัย/ตัวแปรระดับนักเรียน ได้แก่ บรรยากาศและสภาพแวดล้อมใน

ครอบครัว และความรู้เดิมของนักเรียนระดับห้องเรียน ได้แก่ บรรยากาศและสิ่งแวดล้อมของห้องเรียน ความสัมพันธ์ระหว่างนักเรียนปัจจัยตัวแปรระดับสถานศึกษา ได้แก่ ขนาดของสถานศึกษา ค่าใช้จ่ายทางการเรียน การบริหารสถานศึกษา ความเป็นผู้นำทางวิชาการของผู้บริหาร การใช้ทรัพยากรทางการศึกษา และการใช้เทคโนโลยีทางการศึกษา และตัวแปรที่ส่งผลต่อการปฏิบัติงานร่วมกับผู้อื่น ประกอบด้วย ระดับนักเรียน ได้แก่ เพศ และผลสัมฤทธิ์เดิมของนักเรียน ระดับห้องเรียน และระดับสถานศึกษา

นิตยา เหมือดไธสง (2543) ศึกษาความสัมพันธ์ระหว่างปัจจัยระดับนักเรียน ครู และสถานศึกษาที่ส่งผลต่อผลสัมฤทธิ์วิชาคณิตศาสตร์ ทรงตรงและทางอ้อม ด้วยวิธีวิเคราะห์ถดถอยพหุคูณ และการวิเคราะห์ลิสเรล ผลการวิเคราะห์ความสัมพันธ์เชิงสาเหตุด้วยลิสเรล พบว่า ปัจจัยที่ส่งผลทางตรงเชิงบวกอย่างมีนัยสำ คัญในระดับสูง ได้แก่ ภูมิหลังของนักเรียน รองลงมาคือพฤติกรรมการสอนและภูมิหลังของครู และปัจจัยระดับสถานศึกษา สำหรับปัจจัยที่ส่งผลทางอ้อม ได้แก่ ภูมิหลังครูที่มีขนาดของผลที่สูงขึ้นเมื่อส่งผ่านตัวแปรพฤติกรรมการสอนของครู

ปิยะธิดา ทองอร่าม (2545) ได้ศึกษาปัจจัยที่มีอิทธิพลต่อคุณภาพของสถานศึกษา สังกัดกรมสามัญศึกษา ด้วยการวิเคราะห์พหุระดับด้วยโมเดลเชิงเส้นตรงระดับลดหลั่น 2 ระดับ คือระดับสถานศึกษา และระดับจังหวัด ผลการศึกษาพบว่า ปัจจัยระดับสถานศึกษาที่มีอิทธิพลต่อคุณภาพของสถานศึกษาอย่างมีนัยสำ คัญทางสถิติ ได้แก่ สัดส่วนของนักเรียนต่อครู สัดส่วนนักเรียนต่อห้องเรียน ขวัญและกาลังใจในการปฏิบัติงาน และความเป็นผู้นำ ของผู้บริหารสถานศึกษา ส่วนปัจจัยระดับจังหวัดที่มีอิทธิพลต่อคุณภาพของสถานศึกษาอย่างมีนัยสำ คัญทางสถิติ ได้แก่ ค่าเฉลี่ยปริมาณการให้คำ แนะนำทางวิชาการจากหน่วยศึกษาอย่างมีนัยสำ คัญทางสถิติ ค่าเฉลี่ยงบประมาณที่ได้รับการสนับสนุนจากจังหวัด

สุบิน ยุระรัช (2547) ได้ศึกษาโมเดลสามระดับของข้อมูลตัวแปรที่ส่งผลต่อคุณภาพนักเรียนประถมศึกษาในกรุงเทพมหานคร พบว่าตัวแปรทำนายระดับนักเรียนที่ส่งผลต่อคุณภ าพนักเรียนประถมศึกษาอย่างมีนัยสำ คัญทางสถิติ ได้แก่ ความคาดหวังของนักเรียนในการศึกษาต่อ ความคิดเห็นของนักเรียนที่มีต่อการเรียน พฤติกรรมการเรียนของนักเรียน แรงจูงใจใฝ่สัมฤทธิ์ของนักเรียน และพฤติกรรมของบิดามารดาในการส่ง เสริมการเรียนของบุตร ตัวแปรระดับห้องเรียน พบว่าไม่มีตัวแปรใดที่ส่งผลต่อค่าเฉลี่ยคุณภาพของคุณภาพของนักเรียน และในระดับสถานศึกษาไม่พบความผันแปร

รัชนก บุญปู่ (2547) ได้ศึกษาปัจจัยที่มีอิทธิพลต่อผลการเรียนรู้ตามมาตรฐานหลักสูตรการศึกษาขั้นพื้นฐานในสถานศึกษานำ ร่องหลักสูตรสถานศึกษา ด้วยการวิเคราะห์พหุระดับ 2 ระดับ ผลการวิจัยพบว่า ตัวแปรระดับนักเรียนที่มีอิทธิพลต่อผลการเรียนรู้ตามมาตรฐานหลักสูตร

การศึกษาขั้นพื้นฐาน อย่างมีนัยสำคัญทางสถิติ ได้แก่ เจตคติต่อการเรียน พฤติกรรมการเรียน ความคาดหวังในการศึกษาต่อ โดยชุดของตัวแปรระดับนักเรียนสามารถอธิบายความแปรปรวนของผลการเรียนรู้ได้ร้อยละ 17.4 ส่วนตัวแปรระดับสถานศึกษาที่ส่งผลต่อคุณภาพของการจัดการศึกษาตามหลักสูตรสถานศึกษาขั้นพื้นฐาน อย่างมีนัยสำคัญทางสถิติ ได้แก่ คุณภาพการสอน ระดับการศึกษาของผู้บริหาร โดยตัวแปรระดับสถานศึกษาสามารถอธิบายความแปรปรวนของผลการเรียนรู้ได้ร้อยละ 15.5

ศุภลักษณ์ ใจแสวงทรัพย์ (2547) ได้ศึกษาปัจจัยที่ส่งผลต่อคะแนนพัฒนาการวิชาคณิตศาสตร์ของนักเรียนในระดับมัธยมศึกษาปีที่ 3 จากผลการศึกษาพบว่า พัฒนาการทางการเรียนคณิตศาสตร์ได้รับอิทธิพลทางตรงจากฐานะทางเศรษฐกิจ ของผู้ปกครอง อันดับรองลงมาได้แก่ผลสัมฤทธิ์ทางการเรียนเดิม การศึกษาของครูผู้สอน และเจตคติต่อวิชาคณิตศาสตร์

พิชิต ธรรมรักษ์ (2549) ได้ศึกษาปัจจัยที่มีผลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียนแผนการเรียนศิลป์ภาษา ในกรุงเทพมหานคร ผลการวิจัยพบว่าปัจจัยที่สำ คัญที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียน ได้แก่ ประสบการณ์สอน ขนาดของสถานศึกษา ความรู้พื้นฐานเดิม อาชีพของผู้ปกครอง ความเป็นผู้นำ ทางวิชาการของผู้บริหาร เจตคติต่อวิชาคณิตศาสตร์ และตัวแปรทั้ง 6 ตัวแปรสามารถทานายผลสัมฤทธิ์ทางการเรียนได้ร้อยละ 39.10

สุชีรา มะหิเมือง (2547) ได้ศึกษาปัจจัยที่ส่งผลต่อผลสัมฤทธิ์และพัฒนาการทางวิชาการด้วยวิธีการวิเคราะห์มูลค่าเพิ่ม กับกลุ่มตัวอย่างสถานศึกษาประถมศึกษาจำนวน 123 โรง ข้อมูลที่ใช้ประกอบด้วยข้อมูลทุติยภูมิของคะแนนผลสัมฤทธิ์วิชาภาษาไทยและวิชาคณิตศาสตร์ซึ่งได้จากผลการทดสอบแห่งชาติของนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2545 และผลการทดสอบของโครงการประเมินคุณภาพและวินิจฉัยข้อบกพร่องทางการเรียนของกลุ่มตัวอย่างขณะศึกษาในชั้นประถมศึกษาปีที่ 4 และข้อมูลปฐมภูมิเกี่ยวกับภูมิหลังของกลุ่มตัวอย่าง บริบททั่วไปของสถานศึกษา และการปฏิบัติงานทางวิชาการตามตัวบ่งชี้ 6 ด้านได้จากการสำรวจด้วยแบบสอบถาม โดยใช้การวิเคราะห์โมเดลเชิงเส้นระดับลดหลั่น(HLM) และการวิเคราะห์ความสัมพันธ์โครงสร้างเชิงสาเหตุ(LISREL) ผลการศึกษาพบว่า โมเดลการวิเคราะห์มูลค่าเพิ่มของค่าเฉลี่ยผลสัมฤทธิ์และพัฒนาการทางวิชาการที่ใช้วิธีวิเคราะห์ถดถอยแบบพหุระดับด้วยโปรแกรม HLM ซึ่งกาหนดให้ตัวแปรภูมิหลังของนักเรียนด้านผลสัมฤทธิ์เดิมเป็นตัวแปรที่ส่งผลเชิงสุ่ม ในขณะที่ควบคุมอิทธิพลจากตัวแปรด้านเศรษฐานะ และพื้นฐานทางภาษาเป็นโมเดลที่มีประสิทธิภาพในการพยากรณ์(R^2) ในวิชาภาษาไทยเป็น 22.10 % และในวิชาคณิตศาสตร์เป็น 24.97 % นอกจากนี้สถานศึกษาที่มีมูลค่าเพิ่มฯจัดอยู่ในกลุ่มสูงเป็นสถานศึกษาที่มีบริบททั่วไปของสถานศึกษาและสภาพการปฏิบัติงานทางวิชาการดีกว่ากลุ่มสถานศึกษา

ที่มีมูลค่าเพิ่มฯในระดับต่ำ โมเดลเชิงสาเหตุแสดงอิทธิพลของปัจจัยตัวแปรระดับสถานศึกษาที่มีต่อมูลค่าเพิ่มเป็นโมเดลที่มีความตรงและความกลมกลืนกับข้อมูลเชิงประจักษ์ตัวแปรที่มีอิทธิพลอย่างมีนัยสำคัญ ได้แก่ การปฏิบัติงานทางวิชาการที่ส่งผลเชิงบวกต่อมูลค่าเพิ่มของค่าเฉลี่ยผลสัมฤทธิ์และเชิงลบต่อมูลค่าเพิ่มของพัฒนาการทางวิชาการทำเลที่ตั้งของสถานศึกษาที่ส่งผลเชิงลบต่อมูลค่าเพิ่มของพัฒนาการทางวิชาการวิชาคณิตศาสตร์ และการส่งผลเชิงบวกของมูลค่าเพิ่มของค่าเฉลี่ยผลสัมฤทธิ์ต่อมูลค่าเพิ่มของพัฒนาการทางวิชาการที่ระดับนัยสำคัญ 0.01

ชูเวช ชาญสง่าเวช และสิงหา เจียมศิริ (2546) ได้ทำการศึกษาการวัดมูลค่าเพิ่มจากการดำเนินงานของสถาบันอุดมศึกษา โดยมีวัตถุประสงค์ที่จะพัฒนาระบบการวัดมูลค่าเพิ่มจากการดำเนินงานของสถาบันอุดมศึกษา เพื่อใช้ในกรณีศึกษาโครงการนำ ร่องเพื่อทดลองจัดระดับสถาบันอุดมศึกษาไทยในสาขาวิชาครุศาสตร์/ศึกษาศาสตร์การพัฒนาเกณฑ์ต่างๆในระบบนั้นยึดหลักในการกำหนดชุดของเกณฑ์ซึ่งเป็นที่ยอมรับกันอย่างกว้างขวางอยู่ในแวดวงวิชาการมูลค่าเพิ่มจากการดำเนินงานของสถาบันอุดมศึกษาในการวิจัยนี้ได้พัฒนาขึ้นเป็นเกณฑ์หนึ่งซึ่งอยู่ในสามมุมมองใหญ่ๆตามความต้องการใช้ของผู้มีส่วนได้เสียกับการดำเนินงานของสถาบันอุดมศึกษา คือ มุมมองของนักเรียนที่ประสงค์จะเข้าศึกษาในสถาบันอุดมศึกษาและผู้ปกครองตลอดจนศิษย์เก่าของสถาบันผู้ใช้บัณฑิต และคณาจารย์ประจำสถาบันอุดมศึกษานั้นๆ ผลจากการประยุกต์ในการทดลองจัดระดับสถาบันอุดมศึกษาไทยสาขาวิชาครุศาสตร์/ศึกษาศาสตร์พบว่าเกณฑ์และตัวชี้วัดที่พัฒนาขึ้นสามารถใช้ประเมินความเห็นของศิษย์เก่าผู้ใช้บัณฑิตและคณาจารย์เกี่ยวกับมูลค่าเพิ่มที่ได้รับและสามารถนามาใช้ในการจัดระดับสถาบันอุดมศึกษาได้เป็นอย่างดี

อนงค์ อินตาพรหม(2552) ได้ทำการศึกษาการวิเคราะห์พหุระดับของปัจจัยระดับครูและนักเรียนที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียน ผลการวิจัยพบว่า รูปแบบการเรียนรู้แบบนักทฤษฎีมีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ของนักเรียนอย่างมีนัยสำคัญที่ระดับ .05 โดยตัวแปรระดับโรงเรียนพบว่า ขนาดห้องเรียน และคณะครูที่สำเร็จการศึกษามีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ของนักเรียนอย่างมีนัยสำคัญที่ระดับ .01 และความพึงพอใจของครูต่องานสอนมีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ของนักเรียนอย่างมีนัยสำคัญที่ระดับ.05

จากการศึกษาเอกสารเกี่ยวข้องตัวแปรทางการศึกษาที่ส่งผลต่อคุณภาพการจัดการศึกษา สามารถสรุปตัวแปรทำนายที่มีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนได้ดังนี้

ตารางที่ 6 ผลการสังเคราะห์ตัวแปรทำนายที่มีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียน จำแนกตามระดับของตัวแปร

ตัวแปร	Chiu Xihua (2008)	Gluszynski & Dhawan Biswal (2008)	Douglas Willms (2004)	Koutsoulis และ Campbell (2001)	Willms &Somers (2001)	Patrick and Thomasz (2001)	Heck (2000)	D'Agostino (2000)	Opdenakker& Van Damme (2000)	Hofman, Hofman and Guldemon (1999)	TIMSS 2007	รวม
ระดับสถานศึกษา												
ขนาดสถานศึกษา			✓				✓	✓	✓		✓	5
ที่ตั้งสถานศึกษา			✓					✓		✓	✓	4
บริบทสถานศึกษา/สังกัด					✓		✓	✓				3
แหล่งทรัพยากร อุปกรณ์					✓						✓	2
พฤติกรรมการสอนของครู									✓	✓	✓	3
ประสบการณ์สอนของครู											✓	1
วุฒิการศึกษาของครู								✓				1
ความเป็นผู้นำทางวิชาการ								✓				1
ระดับนักเรียน												
เพศ		✓	✓			✓	✓	✓			✓	6
อายุ		✓	✓				✓					3
การย้ายถิ่น/ แหล่งอาศัย		✓										1
เชื้อชาติ/ชาติกาเนิด		✓	✓				✓					3
เศรษฐานะของครอบครัว	✓	✓	✓	✓	✓		✓			✓	✓	8
คุณลักษณะของครอบครัว	✓	✓		✓	✓			✓			✓	6
ทรัพยากรในครอบครัว	✓							✓			✓	3
ความรู้เดิมของนักเรียน	✓		✓	✓		✓	✓	✓	✓	✓	✓	9
ความเชื่อมั่นในตนเอง	✓		✓	✓							✓	4
เจตคติและแรงจูงใจ	✓						✓	✓	✓	✓	✓	6

ตารางที่ 6 ผลการสังเคราะห์ตัวแปรทำนายที่มีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียน จำแนกตามระดับของตัวแปร (ต่อ)

ตัวแปร	อิทธิฤทธิ์ พงษ์ปิยรัตน์ (2551)	ประกฤติยา (2552)	พิชิต ธรรมรักษ์ (2549)	รัชนก บุญปู่ (2547)	ศุภลักษณ์ ใจแสวงทรัพย์ (2547)	สุบิน ยุระรัช (2547)	สุชีรา มะหิเมือง (2547)	ปิยะธิดา ทองอร่าม (2545)	บุญเรือง ศรีเหรัญ (2542)	นิตยา เหมือดไธสง (2543)	TIMSS 2007	รวม
ระดับสถานศึกษา												
ขนาดสถานศึกษา	✓	✓	✓			✓					✓	5
ที่ตั้งสถานศึกษา		✓				✓	✓				✓	4
บริบทสถานศึกษา/สังกัด		✓					✓		✓			3
แหล่งทรัพยากร อุปกรณ์		✓							✓		✓	3
พฤติกรรมการสอนของครู		✓		✓				✓	✓	✓	✓	6
ประสบการณ์สอนของครู		✓	✓								✓	3
วุฒิการศึกษาของครู		✓		✓								2
ความเป็นผู้นำทางวิชาการ	✓		✓					✓				3
ระดับนักเรียน												
เพศ		✓				✓	✓				✓	4
อายุ		✓										1
การย้ายถิ่น/ แหล่งอาศัย		✓										1
เชื้อชาติ/ชาติกาเนิด		✓										1
เศรษฐานะของครอบครัว		✓	✓			✓	✓				✓	5
คุณลักษณะของครอบครัว		✓			✓				✓		✓	4
ทรัพยากรในครอบครัว		✓									✓	2
ความรู้เดิมของนักเรียน	✓	✓	✓			✓	✓		✓	✓	✓	8
ความเชื่อมั่นในตนเอง		✓			✓						✓	3
เจตคติและแรงจูงใจ		✓	✓	✓	✓				✓		✓	6

ตัวแปรที่ใช้ในงานวิจัยเรื่องนี้ได้จากผลการสังเคราะห์งานวิจัย 20 เรื่องที่ศึกษาปัจจัยที่มีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียน และจากฐานข้อมูลของ TIMSS 2007 โดยผลจากการสังเคราะห์ตัวแปรซึ่งใช้เป็นตัวบ่งชี้ถึงคุณภาพการจัดการศึกษาของสถานศึกษา มีความสอดคล้องกับตัวแปรต่างๆใน TIMSS 2007 และเมื่อนำมาพิจารณาตัวแปรที่ศึกษามากที่สุด ในตัวแปรระดับนักเรียน ได้แก่ ความรู้เดิมและเศรษฐานะของครอบครัว ส่วนตัวแปรระดับสถานศึกษาที่นำมาศึกษามากที่สุดได้แก่ สังกัดสถานศึกษา ขนาดสถานศึกษา ที่ตั้งสถานศึกษา

ตารางที่ 7 อันดับของตัวแปรที่นิยมนำมาศึกษาปัจจัยที่มีผลต่อคุณภาพการจัดการศึกษา และแหล่งข้อมูลในการเก็บข้อมูลของโครงการประเมินผล TIMSS 2007

อันดับที่	**ตัวแปร**	**แหล่งข้อมูลที่นำมาศึกษาใน** TIMSS 2007
ระดับสถานศึกษา		
1	ขนาดสถานศึกษา	จำนวนนักเรียนในสถานศึกษา
2	พฤติกรรมการสอนของครู	- กิจกรรมการเรียนการสอน - การใช้สื่อการสอนในห้องเรียน - การเน้นการให้การบ้าน
3	ตำแหน่งที่ตั้งสถานศึกษา	ขนาดของชุมชนที่สถานศึกษาตั้งอยู่
4	สังกัดสถานศึกษา	ไม่มีในฐานข้อมูล
5	แหล่งทรัพยากร อุปกรณ์	- แหล่งการเรียนรู้และสนับสนุนทางการเรียน - การมีส่วนร่วมสนับสนุนการเรียนการสอนของผู้ปกครอง
6	ประสบการณ์สอนของครู	จำนวนระยะเวลาการทำงานของครู
ระดับนักเรียน		
1	ความรู้เดิมของนักเรียน	- ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา - การใช้เวลาเรียนคณิตศาสตร์ในสถานศึกษา
2	เศรษฐานะของครอบครัว	- ความมั่งคั่งของครอบครัว
3	คุณลักษณะของครอบครัว	- ความสนใจในการเรียนรู้คณิตศาสตร์ - แรงจูงใจภายนอกในการเรียนคณิตศาสตร์
4	เพศ	เพศ
4	คุณลักษณะของครอบครัว	ระดับการศึกษาผู้ปกครอง
5	ความเชื่อมั่นในตนเอง	ความเชื่อตนเองในการเรียนวิทยาศาสตร์
6	ทรัพยากรในครอบครัว	- แหล่งทรัพยากรที่บ้าน

5.3 การกำหนดกรอบแนวคิดในการวิจัย

จากการศึกษาค้นคว้าเอกสารและงานวิจัย การประเมินคุณภาพการ จัดการศึกษาให้มีความถูกต้องและยุ ติธรรม จะต้องพิจารณา จากแบบสอบที่ใช้ประเมิน ซึ่งจะต้องมีความยุติธรรมสำ หรับกลุ่มผู้สอบทุกกลุ่ม มีความเชื่อถือได้ ไม่โอนเอียงเข้าหากลุ่มผู้สอบกลุ่มใดกลุ่มหนึ่งเมื่อผู้สอบมีความสามารถอยู่ในระดับเท่ากันหรือเป็นแบบสอบที่ไม่มีการทำ หน้าที่ต่างกันของข้อสอบ ทั้งตัวเลือ กที่เป็นตัวถูก และตัวเลือกที่เป็นตัวลวง ตลอดจนการประเมินผลคะแนนการสอบควรจะต้องคำนึงถึงความแตกต่างของปัจจัยของนักเรียนและสถานศึกษา ดังนั้นผู้วิจัยจึงกำหนดกรอบการวิจัยเพื่อหาคะแนนมูลค่าเพิ่มของสถานศึกษา โดยเปรียบเทียบโมเดล การวิเคราะห์ 2 ระดับ คือระดับนักเรียน และระดับสถานศึกษา โดยในแต่ละส่วนประกอบไปด้วยโมเดลการวัด 4 โมเดล คือ โดยมีรายละเอียดดังนี้

โมเดล 1 “Undetected DIF-DDF & Adjusted” เป็นโมเดลการประเมินคุณภาพการจัดการศึกษา วิชาคณิตศาสตร์ที่ ไม่ได้ ตัดข้อสอบที่ทำหน้าที่ต่างกันออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา (Adjusted)

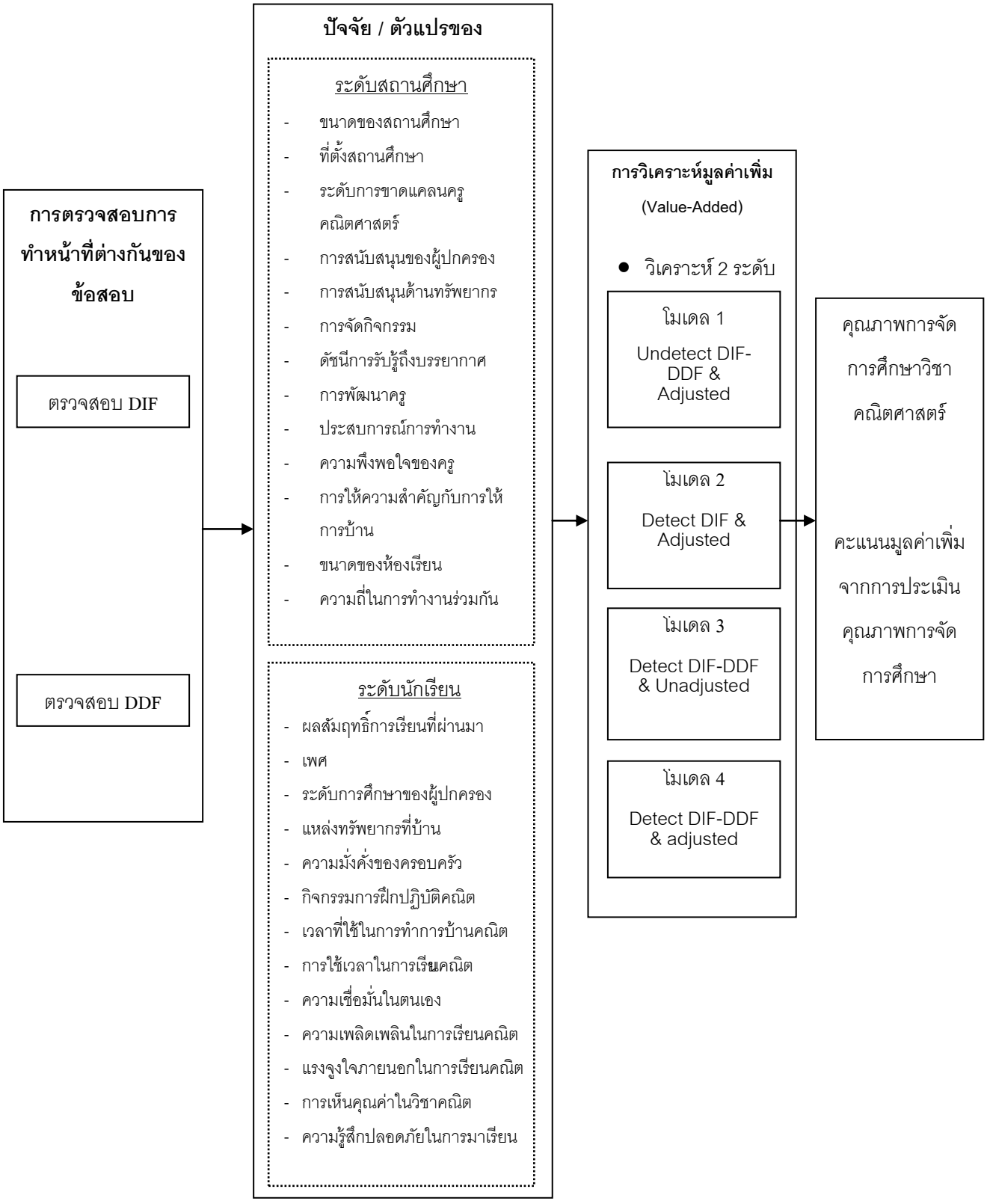
โมเดล 2 “Detect DIF & Adjusted” เป็นโมเดลการประเมินคุณภาพการจัดการศึกษา วิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ตัวถูกออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา (Adjusted)

โมเดล 3 “Detect DIF-DDF & Unadjusted” เป็นโมเดลเป็นการประเ มินคุณภาพการจัดการศึกษา วิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ทั้งตัวถูกและตัวลวงออก แต่ไม่ได้ทำการวิเคราะห์มูลค่าเพิ่ม (Unadjusted)

โมเดล 4 “Detect DIF-DDF & Adjusted” เป็นโมเดลเป็น การประเมินคุณภาพการจัดการศึกษา วิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันโ ดยการวิเคราะห์ทั้งตัวถูกและตัวลวงออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา (Adjusted)

ในการวิจัยครั้งนี้ ผู้วิจัยมีวัตถุประสงค์เพื่อประเมินคุณภาพการจัดการศึกษา วิชาคณิตศาสตร์ที่มีความยุติธรรมและถูกต้องเชื่อถือได้ พร้อมทั้งเปรียบเทียบโมเดลประเมินคุณภาพการจัดการศึกษา โดยการประยุกต์ใช้การวิเคราะห์การทำ หน้าที่ต่างกันของข้อสอบ ที่ตรวจสอบทั้งตัวถูกและตัวลวง (Differential item functioning: DIF & Differenrtial Distractor Functioning: DDF) พร้อมกับปรับแก้คะแนนด้วยตัว

แปรระดับนักเรียนและสถานศึกษา ด้วยการวิเคราะห์มูลค่าเพิ่ม (Value-added analysis) เขียนกรอบแนวคิดในการวิจัยดังนี้



ภาพที่ 2 กรอบแนวคิดในการวิจัยเพื่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์

บทที่ 3

วิธีดำเนินการวิจัย

การวิเคราะห์เปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ : การประยุกต์ใช้โมเดลมูลค่าเพิ่มที่มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ และการวิเคราะห์การทำหน้าที่ต่างกันของตัวลวง มีวัตถุประสงค์คือ 1. เพื่อศึกษาผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) ในแบบสอบประเมินความรู้วิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 2 2. เพื่อศึกษาผลของข้อสอบที่ทำหน้าที่ต่างกัน และผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาต่อการประเมินคุณภ าพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน 4 โมเดล และ 3. เพื่อเปรียบเทียบประสิทธิภาพของโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษา 4 โมเดล โดยวิธีดำเนินการวิจัยแบ่งออกเป็น 3 ตอน คือ ตอนที่ 1 ข้อมูลที่ใช้ในการวิจัย ตอนที่ 2 การเก็บรวบรวมข้อมูล ตอนที่ 3 การวิเคราะห์ข้อมูล โดยมีรายละเอียด ดังนี้

ตอนที่ 1 ข้อมูลที่ใช้ในการวิจัย

การวิจัยครั้งนี้ผู้วิจัยใช้ข้อมูลทุติยภูมิซึ่งได้จาก การศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ระดับนานาชาติ 2550 หรือ TIMSS 2007 ซึ่งเป็นโครงการที่จัดขึ้นเพื่อประเมินผลสัมฤทธิ์ทางการเรียนของนักเรียนในวิชาคณิตศาสตร์ โดยทำการศึกษานักเรียนในระดับชั้นมัธยมศึกษาปีที่ 2 (Grade 8) จากหน่วยงานที่จัดการศึกษาในระดับมัธยมศึกษาตอนต้นและหน่วยงานที่เกี่ยวข้องกับการจัดเก็บฐานข้อมูลสถิติทางการศึกษา ประกอบด้วย

- สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.)
- สำนักคณะกรรมการส่งเสริมการศึกษาเอกชน (สช.)
- สำนักการศึกษากรุงเทพมหานคร (สนศ.)
- สำนักบริหารการศึกษาท้องถิ่น (สถ.)
- สำนักงานคณะกรรมการอุดมศึกษา (สกอ.)

- ศูนย์รวบรวมข้อมูลการศึกษา กระทรวงศึกษาธิการ

ฐานข้อมูลสถิติทางการศึกษา ของประเทศไทยปีการศึกษา 2549 มีสถานศึกษาที่เปิดสอนระดับมัธยมศึกษาตอนต้น 10,550 สถานศึกษา ห้องเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน 23,952 ห้อง และนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน 855,365 คน ผู้วิจัยได้ตัดสถานศึกษาที่มีจำนวนนักเรียนชั้นมัธยมศึกษาปีที่ 2 น้อยกว่า 15 คน ออกจากกรอบการวิจัย คงเหลือ สถานศึกษา ที่เป็นกลุ่มตัวอย่าง จำนวน 9,481 สถานศึกษา ห้องเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน 22,883 ห้อง และนักเรียนชั้นมัธยมศึกษาปีที่2 จำนวน 844,336 คน รายละเอียดดังตารางที่ 8

ตารางที่ 8 กรอบประชากรนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำแนกตามสังกัด

สังกัด	จำนวน สถานศึกษา	จำนวน ห้องเรียน	จำนวน นักเรียน
สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน	8,563	20,578	752,639
สำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน	607	1,610	64,524
สำนักการศึกษากรุงเทพมหานคร	62	150	6,009
สำนักบริหารการศึกษาท้องถิ่น	227	449	17,629
สำนักงานคณะกรรมการการอุดมศึกษา	22	96	3,895
รวม	**9,481**	**22,883**	**844,336**

คณะ ผู้วิจัย ในประเทศไทยคือสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท.) ได้เก็บรวบรวมข้ อมูลเพื่อการประเมินในเดือนมกราคม 2550 โดยมีขั้นตอนในการวิจัยเช่นเดียวกับการเก็บรวบรวมข้อมูลเบื้องต้น ดังรายละเอียดต่อไปนี้

1) กลุ่มตัวอย่าง

คณะผู้วิจัย ในประเทศไทยคือสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท.) ได้ส่งข้อมูลจำนวนสถานศึกษาในประเทศไทยให้กับ Statistics Canada สำหรับการสุ่มเลือกกลุ่มตัวอย่างสำหรับการเก็บรวบรวมข้อมูลเพื่อการประเมิน โดยกลุ่มตัวอย่างที่ใช้ในการเก็บรวบรวมข้อมูลเพื่อการประเมิน ประกอบด้วย นักเรียน ครูคณิตศาสตร์ ที่สอนนักเรียนที่เป็นกลุ่มตัวอย่างในการวิจัยและผู้บริหาร สถานศึกษา ที่นัก เรียนกลุ่มตัวอย่างศึกษา จำนวน 150 สถานศึกษา ซึ่งเมื่อดำเนินการ เก็บรวบรวมข้อมูลเสร็จเรียบร้อยพบ ว่าจำนวนนักเรียนที่เป็นกลุ่ม

ตัวอย่างลดลงจากเป้าหมาย แต่สามารถนำข้อมูลมาวิเคราะห์ผลการศึกษาได้ รายละเอียด ดังตารางที่ 9

ตารางที่ 9 จำนวนกลุ่มตัวอย่างสำหรับการเก็บรวบรวมข้อมูลเพื่อการประเมิน

กลุ่มตัวอย่าง	จำนวนสถานศึกษา	จำนวนห้องเรียน	จำนวนครู	จำนวนนักเรียน
เป้าหมาย	150	150	150	5,579
เก็บได้จริง	150	150	150	5,412

2) **เครื่องมือในการเก็บรวบรวมข้อมูล** ประกอบด้วย

2.1) **แบบสอบ** แบบสอบโครงการ TIMSS 2007 ประกอบด้วยข้อสอบวิชาคณิตศาสตร์ จำนวน 430 ข้อ ซึ่งมีจำนวนมาก ดังนั้นเพื่อให้นักเรียนได้ทำข้อสอบทุกข้อตามเวลาที่กำหนดในการทดสอบ คือ 1 ชั่วโมง 30 นาที (ใช้เวลาในการทดสอบวิชาละ 45 นาที) จึงแบ่งข้อสอบแต่ละวิชาออกเป็น 14 Cluster ประกอบด้วยข้อสอบแบบเลือกตอบ (Multiple – Choice Item) และข้อสอบแบบเขียนตอบ (Constructed – Response Item) ซึ่งการสร้างข้อสอบนั้นได้ถูกสร้างมาจากการสังเคราะห์เนื้อหา หลักสูตรจากประเทศต่างๆที่เข้าร่วมสอบ โดยมีรายละเอียดตามที่แทรกไว้ในภาคผนวก โดยข้อสอบแต่ละ Cluster มีสัดส่วนของเนื้อหาวิชาและพฤติกรรมการเรียนรู้ ตามกรอบในกา รประเมินผลสัมฤทธิ์ ทางการเรียนคณิตศาสตร์ของโครงการ โดยในการเก็บรวบรวมข้อมูลเพื่อการประเมินมีการจัดชุด แบบสอบย่อยเป็น 14 ฉบับ แต่ละฉบับมีข้อสอบ อยู่ในช่วง 28 – 30 ข้อ ประกอบด้วยเนื้อหาวิชาคณิตศาสตร์จำนวน 4 เรื่อง โดยรายละเอียดการให้คะแนนของแบบสอบใน วิชาคณิตศาสตร์โครงการ TIMSS 2007 มีรายละเอียดดังตารางที่ 10 ดังนี้

ตารางที่ 10 แสดงการจำแนกคะแนนของแบบสอบวิชาคณิตศาสตร์โครงการTIMSS 2007 ตามด้านเนื้อหา(Content Domains) และด้านพฤติกรรมการเรียนรู้(Cognitive Domain)

ด้านเนื้อหา (Content Domains)	ด้านพฤติกรรมการเรียนรู้ (Cognitive Domain)			คะแนนรวม	คิดเป็นร้อยละ
	ความรู้	การประยุกต์	การให้เหตุผล		
จำนวน	28	33	11	72	30%
พีชคณิต	33	15	21	69	29%
เรขาคณิต	8	29	13	50	21%
ข้อมูลและโอกาส	14	21	12	47	20%
คะแนนรวม	83	98	57	238	100%
ร้อยละ	35%	41%	24%	100%	

การจัดแบบสอบให้กับนักเรียนโครงการมีหลักเกณฑ์ในการสุ่ม แบบสอบโดยการสุ่มแบบเป็นระบบ ดังนั้นนักเรียนที่มีลำดับการสอบที่ติดกันจึงไม่มีโอกาสได้ แบบสอบฉบับเดียวกัน โดยนักเรียนทุกคนต้องเริ่มทำแบบสอบในแต่ละส่วนพร้อมกัน ดังรายละเอียดตารางที่ 11

ตารางที่ 11 จำนวนข้อสอบวิชาคณิตศาสตร์แต่ละฉบับ จำแนกตาม Cluster

ฉบับ	Cluster														รวม
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
1	13	17	-	-	-	-	-	-	-	-	-	-	-	-	30
2	-	17	15	-	-	-	-	-	-	-	-	-	-	-	32
3	-	-	15	17	-	-	-	-	-	-	-	-	-	-	32
4	-	-	-	17	12	-	-	-	-	-	-	-	-	-	29
5	-	-	-	-	12	17	-	-	-	-	-	-	-	-	29
6	-	-	-	-	-	17	15	-	-	-	-	-	-	-	32
7	-	-	-	-	-	-	15	18	-	-	-	-	-	-	33
8	-	-	-	-	-	-	-	18	14	-	-	-	-	-	32
9	-	-	-	-	-	-	-	-	14	17	-	-	-	-	31
10	-	-	-	-	-	-	-	-	-	17	15	-	-	-	32
11	-	-	-	-	-	-	-	-	-	-	15	17	-	-	32
12	-	-	-	-	-	-	-	-	-	-	-	17	11	-	28
13	-	-	-	-	-	-	-	-	-	-	-	-	11	17	28
14	13	-	-	-	-	-	-	-	-	-	-	-	-	17	30

ขอบเขตการประเมินในวิชาคณิตศาสตร์ ประกอบด้วย ด้านเนื้อหา (Content Domains) และด้านพฤติกรรมการเรียนรู้ (Cognitive Domain) มีรายละเอียดดังนี้

1. **ด้านเนื้อหา (Content Domains)**

การประเมินผลด้านเนื้อหาคณิตศาสตร์ของโครงการ TIMSS 2007 ครอบคลุมเนื้อหาเรื่องจำนวน พีช คณิต เรขาคณิต และข้อมูลและโอกาส โดยการกำหนดน้ำหนักความสำคัญของเนื้อหาที่ใช้ในการประเมิน มีรายละเอียดดังตาราง ที่ 12

ตารางที่ 12 ร้อยละของเนื้อหาที่ใช้ในการประเมินโครงการ TIMSS 2007

เนื้อหา	ร้อยละ
1. จำนวน	30
2. พีชคณิต	30
3. เรขาคณิต	20
4. ข้อมูลและโอกาส	20

หัวข้อการเรียนรู้ของเนื้อหาที่ใช้ในการประเมินโครงการ TIMSS 2007 ครอบคลุมเนื้อหาตามหลักสูตรของแต่ละประเทศที่เข้าร่วมประเมินในระดับชั้นมัธยมศึกษาปีที่ 2 โดยหัวข้อที่ใช้ในการประเมินในแต่ละเนื้อหามีรายละเอียดดังตารางที่ 13 ดังนี้

ตารางที่ 13 หัวข้อการเรียนรู้จำแนกแต่ละเนื้อหาที่ใช้ในการประเมินโครงการ TIMSS 2007

เนื้อหา	หัวข้อ
จำนวน	● จำนวนเต็ม (ค่าประจำหลัก การแยกตัวประกอบและการบวก ลบ คูณ หาร) ● การคำนวณ การคาดคะเนหรือการประมาณเกี่ยวกับจำนวนเต็ม ● เศษส่วนอย่างง่าย (การเท่ากันของเศษส่วน และการเรียงลำดับเศษส่วน) ● ทศนิยม (ค่าประจำหลัก การเรียงลำดับ การประมาณค่า และการแปลงทศนิยมและเศษส่วน) ● การแสดงทศนิยมและเศษส่วนโดยใช้คำพูด จำนวน หรือแบบรูป (รวมทั้งเส้นจำนวน)

ตารางที่ 13 หัวข้อการเรียนรู้จำแนกแต่ละเนื้อหาที่ใช้ในการประเมินโครงการ TIMSS 2007 (ต่อ)

เนื้อหา	หัวข้อ
จำนวน (ต่อ)	● การคำนวณเศษส่วน ● การคำนวณทศนิยม ● การแสดงการเปรียบเทียบ การเรียงลำดับ และการคำนวนจำนวนเต็ม ● อัตราส่วน (การเท่ากัน การหารจำนวนด้วยอัตราส่วนที่กำหนดให้) ● การแปลงเปอร์เซ็นต์ให้อยู่ในรูปของเศษส่วนหรือทศนิยมและการแปลงในทางกลับกัน
พีชคณิต	● แบบรูปหรื อลำดับทางเลขคณิต และเรขาคณิต (การขยาย ส่วนที่หายไป การสร้างแบบรูป) ● ผลรวม ผลคูณ และกำลังของนิพจน์ที่มีตัวแปร ● การประเมินค่านิพจน์ด้วยค่าของจำนวนที่กำหนดให้ ● การทำให้อยู่ในรูปอย่างง่ายหรือการเปรียบเทียบนิพจน์เชิงพีชคณิต ● สร้างแบบรูปสถานการณ์โดยใช้นิพจน์ ● การประเมินค่าฟังก์ชัน/สูตรด้วยค่าของตัวแปรที่กำหนดให้ ● สมการเชิงเส้นและอสมการอย่างง่าย และสมการเชิงเส้นสองตัวแปร ● การแสดงฟังก์ชันในรูปของคู่อันดับ ตาราง กราฟ คำพูด หรือสมการ
เรขาคณิต	● มุม (มุมแหลม มุมฉาก มุมตรง มุมป้าน และมุมสะท้อน) ● ความสัมพันธ์ของมุ มที่จุด มุมบนเส้นตรง มุมตรงข้าม มุมที่เกิดจากการลากเส้นตรงตัดเส้นขนาน และการตั้งฉาก ● สมบัติของรูปเรขาคณิต (รูปสามเหลี่ยม รูปสี่เหลี่ยม และรูปหลายเหลี่ยมอื่นๆ) ● สร้างหรือวาดรูปสามเหลี่ยมและรูปสี่เหลี่ยมผืนผ้าตามขนาดที่กำหนดให้ ● รูปร่างที่มีขนาดสอดคล้องกัน (รูปสามเหลี่ยม รูปสี่เหลี่ยม) และการวัดสิ่งที่สอดคล้องกัน ● สามเหลี่ยมคล้ายและการทบทวนสมบัติของสามเหลี่ยมคล้าย ● ความสัมพันธ์ระหว่างรูปเรขาคณิตสองมิติและสามมิติ ● ใช้ทฤษฎีปีทาโกรัส (โดยไม่ต้องพิสูจน์) เพื่อหาความยาวของด้าน

ตารางที่ 13 หัวข้อการเรียนรู้จำแนกแต่ละเนื้อหาที่ใช้ในการประเมินโครงการ TIMSS 2007 (ต่อ)

เนื้อหา	หัวข้อ
เรขาคณิต (ต่อ)	● การวัด การวาด และการประมาณค่าขนาดของมุม ความยาวของเส้น พื้นที่ และปริมาตร ● สูตรที่ใช้ในการหาความยาวรอบรูป เส้นรอบวง พื้นที่ของวงกลม พื้นที่ผิว และปริมาตร ● การวัดพื้นที่ที่ไม่ปกติหรือพื้นที่เชิงประกอบ (เช่น ปกคลุมด้วยกริด หรือภาคตัดและชิ้นที่ไม่มีการจัดเรียง) ● ใช้หลักคาร์ทีเซียน-คู่อันดับ สมการ ระยะตัดแกน อินเตอร์เซคชัน และเกรเดียนต์ ● เส้นและแกนสมมาตรของรูปเรขาคณิตสองมิติ ● การเลื่อนขนาน การสะท้อน และการหมุน
ข้อมูลและโอกาส	● การอ่านข้อมูลจากตาราง กราฟรูปภาพ กราฟแท่ง แผนภูมิรูปวงกลม และกราฟเส้น ● การจัดการและการแสดงข้อมูลโดยใช้ตาราง กราฟรูปภาพ กราฟแท่ง แผนภูมิรูปวงกลม และกราฟเส้น ● ลักษณะของชุดข้อมูลประกอบด้วยค่าเฉลี่ย ค่ามัธยฐาน ค่าพิสัย และลักษณะการกระจาย (ในแบบทั่ว) ● การแปลความหมายชุดข้อมูล (เช่น การสรุป การทำนาย และการประมาณค่าระหว่างจุดข้อมูลที่กำหนดให้และค่าที่นอกเหนือออกไป) ● การแสดงข้อมูลซึ่งอาจนำไปสู่การแปลความหมายที่ผิดพลาด (เช่น การจัดกลุ่มที่ไม่เหมาะสม และการทำให้ไขว้เขวหรือการบิดเบือนของสเกล) ● การใช้ข้อมูลจากการทดลองเพื่อประมาณโอกาสของสิ่งที่จะเกิดขึ้นในอนาคต ● การใช้โอกาสของสิ่งที่เกิดขึ้นเฉพาะอย่างเพื่อแก้ปัญหา

2. **ด้านพฤติกรรมการเรียนรู้** (Cognitive Domain)

การประเมินผลด้านพฤติกรรมการเรียนรู้ของโครงการ TIMSS 2007 พฤติกรรมการเรียนรู้ 3 ด้าน คือ ความรู้/ความเข้าใจ การประยุกต์ใช้ความรู้ และการบูรณาการความรู้และการให้เหตุผล การกำหนดน้ำหนักของพฤติกรรมการเรียนรู้ที่ใช้ในการประเมิน ดังรายละเอียดตารางที่ 14

ตารางที่ 14 ร้อยละของพฤติกรรมการเรียนรู้ที่ใช้ในการประเมินโครงการ TIMSS 2007

พฤติกรรมการเรียนรู้	ร้อยละ
- ความรู้/ความเข้าใจ	35
- การประยุกต์ใช้ความรู้	40
- การบูรณาการความรู้และการให้เหตุผล	25

สำหรับแต่ละพฤติกรรมการเรียนรู้ที่ใช้ในการประเมินโครงการ TIMSS 2007 ครอบคลุมเนื้อหาตามหลักสูตรของประเทศสมาชิกและระดับชั้นมัธยมศึกษาปีที่ 2 ซึ่งเป็นกลุ่มตัวอย่างของการประเมินในครั้งนี้ รายละเอียดพฤติกรรมการเรียนรู้ที่ใช้ในการประเมินในแต่ละด้าน เป็นดังนี้

ความรู้/ความเข้าใจ ครอบคลุมถึงข้อเท็จจริง กระบวนการ และแนวคิดทางคณิตศาสตร์ที่นักเรียนต้องรู้ ความรู้/ความเข้าใจ ตามนิยามของโครงการ TIMSS มีพฤติกรรมการเรียนรู้ ดังนี้

1. ความรู้ (Recall)	รู้บทนิยาม ศัพท์เฉพาะ และสัญกรณ์
2. ความจำ (Recognize)	จำเนื้อหาเรื่องรูปทรง จำนวน และนิพจน์ (ทศนิยมและเศษส่วนที่เท่ากัน ความแตกต่างของรูปทรงทางเรขาคณิต)
3. การคำนวณ (Compute)	ใช้กระบวนการบวก ลบ คูณ หาร ในการคำนวณจำนวนจริง
4. การเรียกใช้ข้อมูล (Retrieve)	การเรียกใช้ข้อมูลจากกราฟ ตาราง หรือแหล่งข้อมูลอื่น ๆ
5. การวัด (Measure)	ใช้เครื่องมือในการวัดและประมาณค่า
6. การจำแนก (Classify/Order)	จำแนก/จัดกลุ่มรูปทรง จำนวน และนิพจน์

การประยุกต์ใช้ความรู้ เน้นให้นักเรียนประยุกต์ใช้ความรู้และเข้าใจในการแก้ปัญหาหรือตอบคำถาม การประยุกต์ใช้ความรู้ตามนิยามของโครงการ TIMSS มีพฤติกรรมการเรียนรู้ ดังนี้

1. การเลือก (Select)	เลือกกระบวนการในการนำมาใช้ในการแก้ปัญหา
2. การอธิบาย (Represent)	แสดงข้อมูลทางคณิตศาสตร์ เพื่อมาใช้ในอ้างอิง

3. แบบรูป (Model)	รูปแบบหรือสมการที่ใช้ในการแก้ปัญหา
4. การนำไปปฏิบัติ (Implement)	การแก้ปัญหาคณิตศาสตร์ที่ให้นักเรียนแสดงโดยการวาดภาพ
5. การแก้ปัญหาทั่วไป (Solve Routine Problems)	การแก้ปัญหาทั่วไป เช่น การเปรียบเทียบข้อมูลในตาราง

การบูรณาการความรู้และการให้เหตุผล เป็นการแก้ปัญหาโดยการรวมเอาปัญหา ที่ไม่คุ้นเคยปัญหาที่ซับซ้อนและการแก้ปัญหาหลายขั้นตอน การบูรณาการความรู้และการให้เหตุผล ตามนิยามของโครงการ TIMSS มีพฤติกรรมการเรียนรู้ ดังนี้

1. การวิเคราะห์ (Analyze)	กำหนดและบรรยาย หรือใช้ความสัมพันธ์ระหว่างตัวแปรหรือปัญหาทางคณิตศาสตร์
2. การสรุป (Generalize)	สรุปผลการคิดทางคณิตศาสตร์และการแก้ปัญหา
3. การสังเคราะห์ (Synthesize/Integrate)	รวบรวมกระบวนการในการหาคำตอบ
4. การยืนยัน (Justify)	ยืนยันคำตอบทั้งที่เป็นคำตอบที่ถูกและที่ผิด
5. การแก้ปัญหาที่ไม่เคยพบ (Solve Non – routine Problems)	การแก้ปัญหาที่มีการประยุกต์ใช้ในชีวิตประจำวัน

2.2 แบบสอบถามนักเรียน นักเรียนที่เป็นกลุ่มตัวอย่างในการวิจัยทุกคนจะต้องตอบแบบสอบถาม จำนวน 1 ฉบับ ภายหลังจากการทำ แบบสอบ ใช้เวลาประมาณ 30 นาที มีประเด็นดังนี้

- ข้อมูลทั่วไปเกี่ยวกับตัวนักเรียน
- การเรียนวิชาคณิตศาสตร์ในสถานศึกษา
- การใช้คอมพิวเตอร์เพื่อการเรียนวิชาคณิตศาสตร์และทั่วไป
- สถานศึกษาของนักเรียน
- กิจกรรมที่นักเรียนทำนอกสถานศึกษา
- การทำการบ้านวิชาคณิตศาสตร์

2.3 **แบบสอบถามครูคณิตศาสตร์** ครูคณิตศาสตร์ที่สอนนักเรียนที่เป็นกลุ่มตัวอย่าง จำนวน 1 คน ต้องตอบแบบสอบถามจำนวน 1 ฉบับ มีประเด็น ดังนี้

- ข้อมูลทั่วไปเกี่ยวกับตัวครู
- การเตรียมเพื่อการสอน
- การพัฒนาวิชาชีพ
- สถานศึกษาที่สอน
- นักเรียนที่เป็นกลุ่มตัวอย่าง
- การสอนคณิตศาสตร์
- การใช้เครื่องคิดเลขและคอมพิวเตอร์สำหรับการสอนวิชาคณิตศาสตร์
- การบ้านของนักเรียน
- การประเมินผลการเรียนของนักเรียน

2.4 **แบบสอบถามผู้บริหาร สถานศึกษา** ผู้บริหารสถานศึกษาที่นักเรียนที่เป็นกลุ่มตัวอย่าง จำนวน 1 คน ต้องตอบแบบสอบถามจำนวน 1 ฉบับ มีประเด็น ดังนี้

- ลักษณะของสถานศึกษา
- การปฏิบัติงานในฐานะผู้บริหารสถานศึกษา
- การมีส่วนร่วมของผู้ปกครองในกิจกรรมของสถานศึกษา
- บรรยากาศการเรียนรู้ของสถานศึกษา
- การจัดการสอนวิชาคณิตศาสตร์ในสถานศึกษา
- ครูที่สอนวิชาคณิตศาสตร์ในระดับชั้นมัธยมศึกษาปีที่ 2 ของสถานศึกษา
- พฤติกรรมนักเรียน
- แหล่งการเรียนรู้และเทคโนโลยี

ตอนที่ 2 ตัวแปรที่ใช้ในการวิจัย

ตัวแปรที่นำมาใช้ในการวิจัยครั้งนี้ เป็นข้อมูลที่มาจากฐานข้อมูลในโครงการประเมินผล TIMSS 2007 ผู้วิจัยพิจารณาถึงโครงสร้างของข้อมูลในระบบการจัดการศึกษา ซึ่งลักษณะข้อมูลมีโครงสร้างแบบลดหลั่น นั่นคือนักเรียนสอดแทรกอยู่ในสถานศึกษา ดังนั้นจึงแบ่งตัวแปรเป็นระดับ

สถานศึกษาและระดับนักเรียน โดยตัวแปรที่ใช้ในการวิจัยอาจมีชื่อที่แตกต่างออกไปในฐานข้อมูลของ TIMSS เพื่อให้มีความเหมาะสมและสอดคล้องกับการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง ดังนั้นตัวแปรที่นำมาใช้ในงานวิจัยมีดังนี้

ตัวแปรระดับสถานศึกษา

1. ลักษณะทั่วไปของสถานศึกษา ประกอบด้วย
    - 1.1 ขนาดสถานศึกษา
    - 1.2 ตำแหน่งที่ตั้งของสถานศึกษา
    - 1.3 จำนวนนักเรียนในชั้นเรียน
2. ทรัพยากรของสถานศึกษา ประกอบด้วย
    - 2.1 ระดับการขาดแคลนครูคณิตศาสตร์ในสถานศึกษา
    - 2.2 การให้การสนับสนุนของผู้ปกครอง
    - 2.3 ดัชนีการให้การสนับสนุนด้านทรัพยากรที่ใช้ในการจัดการเรียนการสอนคณิตศาสตร์
    - 2.4 การจัดกิจกรรมที่สนับสนุนการเรียนรู้ในวิชาคณิตศาสตร์ภายในสถานศึกษา
3. การบริหารสถานศึกษา ประกอบด้วย
    - 3.1 ดัชนีการรับรู้ถึงบรรยากาศในที่ดีในที่ทำงาน
    - 3.2 การพัฒนาครูในสถานศึกษา
    - 3.3 ประสบการณ์การทำงานของครูในสถานศึกษา
    - 3.4 ดัชนีความพึงพอใจของครูที่มีต่อการจัดการภายในสถานศึกษา
    - 3.5 ดัชนีการให้ความสำคัญกับการให้การบ้านคณิตศาสตร์ของครูในสถานศึกษา
    - 3.6 ระดับความถี่ในการทำงานร่วมกันของครูคณิตศาสตร์ในสถานศึกษา

ตัวแปรระดับนักเรียน

1. ลักษณะทั่วไปของนักเรียน ประกอบด้วย
    - 1.1 ค่าความสามารถของนักเรียน
    - 1.2 เพศ
    - 1.3 ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา
    - 1.4 กิจกรรมการฝึกปฏิบัติในวิชาคณิตศาสตร์

1.5 การใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเอง

1.6 ดัชนีระดับเวลาที่ใช้ในการทำการบ้านวิชาคณิตศาสตร์

2. ทัศนคติในการเรียน ประกอบด้วย

2.1 แรงจูงใจภายนอกในการเรียนคณิตศาสตร์

2.2 ดัชนีของความเชื่อมั่นในตนเองเกี่ยวการเรียนวิชาคณิตศาสตร์

2.3 ดัชนีความเพลิดเพลินในการเรียนวิชาคณิตศาสตร์

2.4 การเห็นคุณค่าในวิชาคณิตศาสตร์

2.5 ดัชนีของนักเรียนในความรู้สึกปลอดภัยในการมาเรียนหนังสือภายในสถานศึกษา

3. พื้นฐานครอบครัว ประกอบด้วย

3.1 ระดับการศึกษาของผู้ปกครอง

3.2 แหล่งทรัพยากรที่บ้าน

3.3 ความมั่งคั่งของครอบครัว

ความหมายของตัวแปรมีรายละเอียด ดังตารางที่ 15

ตารางที่ 15 สเกลการวัดและความหมายของตัวแปรในระดับนักเรียนและสถานศึกษา

ตัวแปร	สเกลการวัด / ความหมาย
<u>ระดับสถานศึกษา</u>	
ขนาดสถานศึกษา	ตัวแปรจัดประเภท แสดงจำนวนนักเรียนในสถานศึกษา จำแนกเป็น ขนาดใหญ่พิเศษ ขนาดใหญ่ ขนาดปานกลาง และขนาดเล็ก
ที่ตั้งสถานศึกษา	ตัวแปรจัดประเภท แสดงขนาดของชุมชนที่สถานศึกษาตั้งอยู่ จำแนกเป็น ตำแหน่งหมู่บ้าน ตำบล เมืองเล็ก เมือง เมืองใหญ่ เมืองใหญ่มาก
ระดับการขาดแคลนครูคณิตศาสตร์ในสถานศึกษา	ตัวแปรจัดประเภท แสดงระดับการขาดแคลนครูคณิตศาสตร์ จำแนกเป็น ไม่ขาดแคลน ขาดแคลนน้อย ขาดแคลนเป็นบางส่วน และขาดแคลนจำนวนมาก

ตารางที่ 15 สเกลการวัดและความหมายของตัวแปรในระดับนักเรียนและสถานศึกษา (ต่อ)

ตัวแปร	สเกลการวัด / ความหมาย
การให้การสนับสนุนของผู้ปกครอง	ตัวแปรต่อเนื่อง คิดจากกิจกรรมการให้การสนับสนุนของผู้ปกครอง
ดัชนีการให้การสนับสนุนด้านทรัพยากรที่ใช้ในการจัดการเรียนการสอนคณิตศาสตร์	ตัวแปรจัดประเภท แสดงระดับ การสนับสนุนด้านทรัพยากรที่ใช้ในการจัดการเรียนการสอนคณิตศาสตร์ จำแนกเป็นระดับต่ำ ปานกลาง และสูง
การจัดกิจกรรมที่สนับสนุนการเรียนรู้ในวิชาคณิตศาสตร์ภายใน	ตัวแปรต่อเนื่อง คิดจาก อัตรา การจัดกิจกรรมที่สนับสนุนการเรียนรู้ในวิชาคณิตศาสตร์ภายใน
ดัชนีการรับรู้ถึงบรรยากาศในที่ดีในที่ทำงาน	ตัวแปรจัดประเภท แสดงระดับการรับรู้ถึงบรรยากาศในที่ดีในที่ทำงาน จำแนกเป็นระดับต่ำ ปานกลาง และสูง
การพัฒนาครูในสถานศึกษา	ตัวแปรต่อเนื่อง คิดจากอัตราการจัดการพัฒนาครูในสถานศึกษา
ประสบการณ์การทำงานของครูในสถานศึกษา	ตัวแปรต่อเนื่อง คิดจากจำนวนระยะเวลาการทำงานของครู
ดัชนีความพึงพอใจของครูที่มีต่อการจัดการภายในสถานศึกษา	ตัวแปรจัดประเภท แสดงระดับความพึงพอใจของครูที่มีต่อการจัดการภายในสถานศึกษา จำแนกเป็นระดับต่ำ ปานกลาง และสูง
ดัชนีการให้ความสำคัญกับการให้การบ้านคณิตศาสตร์ของครูในสถานศึกษา	ตัวแปรจัดประเภท แสดงระดับ การให้ความสำคัญกับการให้การบ้านคณิตศาสตร์ของครูในสถานศึกษา จำแนกเป็นระดับต่ำ ปานกลาง และสูง
ขนาดของห้องเรียน	ตัวแปรจัดประเภท แสดงขนาดของห้องเรียน จำแนกเป็นนักเรียน 1 – 24 คน นักเรียน 25 – 40 คน และนักเรียน มากกว่า 41 คน
ระดับความถี่ในการทำงานร่วมกันของครูคณิตศาสตร์ในสถานศึกษา	ตัวแปรจัดประเภท แสดงระดับความถี่ในการทำงานร่วมกันของครูคณิตศาสตร์ในสถานศึกษา จำแนกเป็นไม่เคยเลย 2-3 ครั้งต่อเดือน และเป็นประจำ

ตารางที่ 15 สเกลการวัดและความหมายของตัวแปรในระดับนักเรียนและสถานศึกษา (ต่อ)

ตัวแปร	สเกลการวัด / ความหมาย
ระดับนักเรียน	
ค่าความสามารถจากการประเมินความรู้ทางคณิตศาสตร์	ตัวแปรต่อเนื่อง ค่าพารามิเตอร์ความสามารถของนักเรียนจากการทดสอบ TIMSS 2007
เพศ	ตัวแปรจัดประเภท จำแนกเป็นนักเรียนชายและหญิง
ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา	ค่าที่ได้จากการคำ นวณหาผลต่างระหว่างระดับชั้นของนักเรียนกับระดับชั้นที่นักเรียนคาดหวังของนักเรียน (Chiu & Xihua, 2008)
ระดับการศึกษาของผู้ปกครอง	ตัวแปรต่อเนื่อง คิดจากระดับการศึกษาของผู้ปกครอง
แหล่งทรัพยากรที่บ้าน	ตัวแปรต่อ เนื่อง คิดจาก ความพร้อมของทรัพยากรในบ้าน ได้แก่ หนังสือ คอมพิวเตอร์ เป็นต้น
ความมั่งคั่งของครอบครัว	ตัวแปรต่อเนื่อง คิดจากความเป็นอยู่ ความสะดวกสบายของครอบครัว
กิจกรรมการฝึกปฏิบัติในวิชาคณิตศาสตร์	ตัวแปรต่อเนื่อง คิดจากกิจกรรมที่หลากหลายของกิจกรรมการฝึกปฏิบัติในวิชาคณิตศาสตร์ในห้องเรียน
ดัชนีระดับเวลาที่ใช้ในการทำการบ้านวิชาคณิตศาสตร์	ตัวแปรจัดประเภท แสดงระดับเวลาที่ใช้ในการทำการบ้านวิชาคณิตศาสตร์ จำแนกเป็นระดับต่ำ ปานกลาง และสูง
การใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเอง	ตัวแปรต่อเนื่อง คิดจากการใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเอง
ดัชนีของความเชื่อมั่นในตนเองเกี่ยวการเรียนวิชาคณิตศาสตร์	ตัวแปรจัดประเภท แสดงระดับของความเชื่อมั่นในตนเองเกี่ยวการเรียนวิชาคณิตศาสตร์ จำแนกเป็นระดับต่ำ ปานกลาง และสูง
ดัชนีความเพลิดเพลินในการเรียนวิชาคณิตศาสตร์	ตัวแปรจัดประเภท แสดงระดับของความเพลิดเพลินในการเรียนวิชาคณิตศาสตร์ จำแนกเป็นระดับต่ำ ปานกลาง และสูง
แรงจูงใจภายนอกในการเรียนคณิตศาสตร์	ตัวแปรต่อเนื่อง คิดจากค่าเฉลี่ยการเกิดสถานการณ์ที่แสดงการเกิดแรงจูงใจภายนอกในการเรียนคณิตศาสตร์

ตารางที่ 15 สเกลการวัดและความหมายของตัวแปรในระดับนักเรียนและสถานศึกษา (ต่อ)

ตัวแปร	สเกลการวัด / ความหมาย
การเห็นคุณค่าในวิชาคณิตศาสตร์	ตัวแปรต่อเนื่อง คิดจากค่าเฉลี่ยการเกิดสถานการณ์ที่แสดงการเห็นคุณค่าในวิชาคณิตศาสตร์
ดัชนีของนักเรียนในความรู้สึกปลอดภัยในการมาเรียนหนังสือภายในสถานศึกษา	ตัว แปรจัดประเภท แสดงระดับความรู้สึกปลอดภัยในการมาเรียนหนังสือภายในสถานศึกษา จำแนกเป็นระดับต่ำ ปานกลาง และสูง

ตอนที่ 3 การวิเคราะห์ข้อมูล

งานวิจัยเรื่อง การวิเคราะห์เปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์: การประยุกต์ใช้โมเดลมูลค่าเพิ่มที่ มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและการวิเคราะห์การทำหน้าที่ต่างกันของตัวลวงมีการวิเคราะห์โดยแบ่งออกเป็น 4 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 การจัดกระทำข้อมูลตามปัจจัยที่ศึกษา

ขั้นตอนที่ 2 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ(DIF) และการทำหน้าที่ต่างกันของตัวลวง(DDF) และการประมาณค่าความสามารถของนักเรียน

ขั้นตอนที่ 3 การวิเคราะห์มูลค่าเพิ่มในโมเดลการประเมินคุณภาพการจัดการศึกษา

ขั้นตอนที่ 4 การตรวจสอบผลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา

ขั้นตอนที่ 1 การจัดกระทำข้อมูลตามปัจจัยที่ศึกษา

ผู้วิจัยดำเนินการวิเคราะห์ ค่าสถิติพื้นฐานของข้อมูลด้วยโปรแกรมการวิเคราะห์ SPSS 11.5 for windows โดยเป็นการวิเคราะห์ข้อมูลเบื้องต้นด้วยสถิติภาคบรรยาย ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ความเบ้ ความโด่ง ค่าต่ำสุด และค่าสูงสุด

ขั้นตอนที่ 2 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF)

2.1 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF)

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) โดยใช้โปรแกรม DDFS 1.0 (Penfield, 2010) ซึ่งเป็นโปรแกรมที่ใช้การวิเคราะห์แมนเทล-แฮนส์เชลร่วมกับอัตราส่วนแต้มต่อ (odds ratio) แสดงสูตรในการวิเคราะห์ได้ดังนี้

$$\hat{\alpha}_{MH_i} = \frac{\sum_{s=1}^{s} R_{1s} F_{0s} / n_s}{\sum_{s=1}^{s} R_{0s} F_{1s} / n_s} \qquad (1)$$

และแมนเทล-แฮนส์เชลร่วมกับลอกของอัตราส่วนแต้มต่อ (odds ratio) แสดงสูตรในการวิเคราะห์ได้ดังนี้

$$\hat{\lambda}_j = \ln(\hat{\alpha}_j) \qquad (2)$$

เนื่องจากการวิเคราะห์แมนเทล-แฮนส์เชลร่วมกับลอกของอัตราส่วนแต้มต่อ (odds ratio) เป็นการเปรียบเทียบอัตราส่วนการตอบของกลุ่มอ้างอิง (Reference) กับกลุ่มเปรียบเทียบ (Focal) ดังนั้นการแปลผลสามารถทำได้ดังนี้

การแปลผลสำหรับการทำหน้าที่ต่างกันของข้อสอบ (DIF)

$\ln(\hat{\alpha}_j) = 0$ หมายถึง ไม่พบการทำหน้าที่ต่างกันของข้อสอบ

$\ln(\hat{\alpha}_j) > 0$ หมายถึง กลุ่มอ้างอิง (Reference) มีโอกาสที่จะตอบถูกมากกว่ากลุ่มเปรียบเทียบ (Focal)

$\ln(\hat{\alpha}_j) < 0$ หมายถึง กลุ่มเปรียบเทียบ (Focal) มีโอกาสที่จะตอบถูกมากกว่ากลุ่มอ้างอิง (Reference)

การแปลผลสำหรับการทำหน้าที่ต่างกันของตัวลวง (DDF)

$\ln(\hat{\alpha}_j)$ = 0 หมายถึง ไม่พบการทำหน้าที่ต่างกันของตัวลวง

$\ln(\hat{\alpha}_j)$ > 0 หมายถึง กลุ่มเปรียบเทียบ (Focal) มีโอกาสที่จะเลือกตอบมากกว่ากลุ่มอ้างอิง (Reference)

$\ln(\hat{\alpha}_j)$ < 0 หมายถึง กลุ่มอ้างอิง (Reference) มีโอกาสที่จะเลือกตอบมากกว่ากลุ่มเปรียบเทียบ (Focal)

2.2 การตัดข้อสอบภายหลังการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF)

จากผลการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) แต่ละฉบับ ผู้วิจัยจึงเลือกตัดข้อสอบที่ทำหน้าที่ ต่างกันออกไปเป็นบางข้อ ตามเงื่อนไขของโมเดลในการศึกษา โดยพิจารณาเลือกตัดข้อสอบตามเกณฑ์ 2 ข้อ ดังนี้ 1. เลือกข้อสอบที่ทำหน้าที่ต่างกันที่มีค่า อิทธิพล สูงสุด และ 2. คำนึงถึงโครงสร้างของแบบสอบ โดยภายหลังการตัดข้อสอบออกแล้วโครงสร้างของแบบสอบจะไม่แตกต่างจากโครงส ร้างเดิม และข้อสอบที่ตัดออกไปควรมีจำนวนไม่เกินร้อยละ 20 (Clauser (1993), cited in Naraya & Swaminathan,1994) ของข้อสอบในแบบสอบนั้น โดย การตัดข้อสอบที่ทำหน้าที่ต่างกันจะคำนึงถึงการรักษาโครงสร้างเดิม โดยจำแนกตามสัดส่วนเนื้อหา ของแบบสอบ

2.3 การประมาณค่าความสามารถของนักเรียน

การประมาณค่าความสามารถของนักเรียนจากการประเมินผลความสามารถทางคณิตศาสตร์จากการทดสอบนานาชาติ TIMSS 2007 ผู้วิจัยใช้โปรแกรม MULTILOG ในการวิเคราะห์ประมาณค่าความสามารถของผู้สอบ ผู้วิจัย ได้ดำเนินการวิเคราะห์ 3 ครั้ง คือ ครั้งที่ 1 เป็นการวิเคราะห์ค่าความสามารถของผู้สอบ ก่อนการตัดข้อสอบที่ทำหน้าที่ต่างกัน ครั้งที่ 2 เป็นการวิเคราะห์ ค่าความสามารถของผู้สอบ หลังจากการตัดข้อสอบที่ทำหน้าที่ต่างกันพิจารณาเฉพาะตัวเลือกที่เป็นตัวถูก และครั้งที่ 3 เป็นการวิเคราะห์ค่าความสามารถของผู้สอบหลังจากการตัดข้อสอบที่ทำหน้าที่ต่างกันพิจารณาทั้งการเกิดกับตัวเลือกที่เป็นตัวถูก และตัวเลือกที่เป็นตัวลวง

ขั้นตอน ที่ 3 การวิเคราะห์มูลค่าเพิ่มในโมเดลการประเมินคุณภาพการจัดการศึกษา

สำหรับขั้นตอนการวิเคราะห์มูลค่าเพิ่มในโมเดลการประเมินคุณภาพการจัดการศึกษานั้น ผู้วิจัยประยุกต์ใช้ โมเดลเชิงเส้นแบบลดหลั่นด้วยการวิเคราะห์ 2 ระดับ สำหรับการ วิเคราะห์มูลค่าเพิ่มในโมเดลการประเมินคุณภาพการจัดการศึกษา 4 โมเดล โดยใช้โปรแกรม HLM for Windows Version 6.03 มีการวิเคราะห์แต่ละระดับ ดังนี้

ระดับที่ 1 ระดับนักเรียน

เป็นการวิเคราะห์ข้อมูลของตัวแปร ระหว่างนักเรียนภายในหน่วยของสถานศึกษา (ระหว่างนักเรียน *i* ภายในสถานศึกษา *k)* โดยการวิเคราะห์ในโมเดลการประเมินคุณภาพการจัดการศึกษาโมเดล 3 จะไม่มีการนำตัวแปรระดับนักเรียนเข้ามาวิเคราะห์ ส่วนโมเดลการประเมินคุณภาพการจัดการศึกษาโมเดล 1, 2 และ 4 จะนำตัวแปรระดับนักเรียนเข้ามาวิเคราะห์ มีสมการดังนี้

$$Y_{ik} = B0_k + \sum_{p=1}^{m} Bp_k X_{pik} + R_{ik}$$

Y_{ik} = ค่าความสามารถจากการประเมินวิชาคณิตศาสตร์ของนักเรียนที่ *i* ภายในสถานศึกษา *k*

$B0_k$ = ค่าจุดตัดแกน (intercept) หรือค่าเฉลี่ยของ Y_{ik} เมื่อควบคุมผลของ X_{pik} แล้วซึ่งเป็นค่าที่แสดงถึงคะแนนเฉลี่ยของค่าความสามารถของนักเรียนจากการประเมินวิชาคณิตศาสตร์ของนักเรียนภายในสถานศึกษา *k*

Bp_k = ค่าความชัน (slope) หรือค่าสัมประสิทธิ์ถดถอยที่แสดงผลของ X_{pik} ต่อ Y_{ik} ซึ่งเป็นที่แสดงถึงค่าอิทธิพลของการเปลี่ยนแปลงของค่าความสามารถของนักเรียนจากการประเมินวิชาคณิตศาสตร์มื่อควบคุมตัวแปรคุณลักษณะของนักเรียน X_{ik}

X_{pik} = ตัวแปรควบคุมระดับบุคคล เมื่อ p = 1,2,3,…,m

R_{ik} = ค่าส่วนที่เหลือหรืออิทธิพลสุ่มระดับบุคคล

โดยค่า $B0_k$ และ Bp_k ที่ได้จากการวิเคราะห์ระดับที่ 1 จะใช้เป็นตัวแปรตามในการวิเคราะห์ระดับที่ 2

ระดับที่ 2 ระดับสถานศึกษา

เป็นการวิเคราะห์ข้อมูลของตัวแปรระหว่างสถานศึกษา โดยการวิเคราะห์ในโมเดลการประเมินคุณภาพการจัดการศึกษาโมเดล 3 จะไม่มีการนำตัวแปรระดับสถานศึกษาเข้ามาวิเคราะห์ ส่วนโมเดลการประเมินคุณภาพการจัดการศึกษาโมเดล 1, 2 และ 4 จะนำตัวแปรระดับสถานศึกษาเข้ามาวิเคราะห์ มีสมการดังนี้

$$B0_k = G00 + \sum_{q=1}^{n} G0_q Z_k + U0_k$$

$$Bp_k = Gp0 + \sum_{q=1}^{n} G0_q Z_k + Up_k$$

โดยที่ $B0_k$ = ค่าเฉลี่ยของ Y_{ik} เมื่อควบคุมผลของ X_{pik} แล้วซึ่งเป็นค่าที่แสดงถึงคะแนนเฉลี่ยของค่าความสามารถของนักเรียนจากการประเมินวิชาคณิตศาสตร์ของนักเรียนภายในสถานศึกษา *k*

G00 = ค่าจุดตัดแกน(intercept) หรือค่าเฉลี่ยของ $B0_j$ เมื่อควบคุมผลของ Z_k แล้วซึ่งเป็นค่าแสดงถึงคะแนนเฉลี่ยของค่าความสามารถของนักเรียนจากการประเมินวิชาคณิตศาสตร์ในสถานศึกษาทุกแห่ง

$G0_q$ = ค่าความชัน (slope) ค่าสัมประสิทธิ์ถดถอยที่แสดงผลของ Z_j ต่อ $B0_j$ ซึ่งแสดงถึงค่าอิทธิพลของการเปลี่ยนแปลงของ ค่าความสามารถของนักเรียนจากการประเมินวิชาคณิตศาสตร์ เมื่อมีการควบคุมตัวแปรคุณลักษณะของสถานศึกษา Y_k

Z_k = ตัวแปรควบคุมระดับหน่วยงาน

$U0_k$ = ค่าส่วนที่เหลือ (residual terms) หรือมูลค่าเพิ่ม (value added) ของสถานศึกษา k

สรุปสมการวิเคราะห์มูลค่าเพิ่มโมเดลเชิงเส้นแบบลดหลั่นด้วยการวิเคราะห์ 2 ระดับ ในโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา 4 โมเดล ดังตารางที่ 16

ตารางที่ 16 สรุปสมการวิเคราะห์มูลค่าเพิ่มด้วยโมเดลเชิงเส้นแบบลดหลั่นด้วยการวิเคราะห์ 2 ระดับ ในโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา 4 โมเดล

ระดับการวิเคราะห์	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
ระดับที่ 1	**ตัวแปรนักเรียน** : มี **ตัวแปรตาม** : ค่าความสามารถจากการประเมินวิชาคณิตศาสตร์ของนักเรียนโดยไม่ตัดข้อสอบที่เกิด DIF และ DDF **สมการวิเคราะห์** : $Y_{ik} = B0_k+B1_k*(X_{ik})_{ik} + R_{ik}$	**ตัวแปรนักเรียน** : มี **ตัวแปรตาม** : ค่าความสามารถจากการประเมินวิชาคณิตศาสตร์ของนักเรียนโดยตัดข้อสอบที่เกิดเฉพาะ DIF เท่านั้น **สมการวิเคราะห์** : $Y_{ik} = B0_k+B1_k*(X_{ik})_{ik} + R_{ik}$	**ตัวแปรนักเรียน** : ไม่มี **ตัวแปรตาม** : ค่าความสามารถจากการประเมินวิชาคณิตศาสตร์ของนักเรียนโดยตัดข้อสอบที่เกิดทั้ง DIF และ DDF **สมการวิเคราะห์** : $Y_{ik} = B0_k + R_{ik}$	**ตัวแปรนักเรียน** : มี **ตัวแปรตาม** : ค่าความสามารถจากการประเมินวิชาคณิตศาสตร์ของนักเรียนโดยตัดข้อสอบที่เกิดทั้ง DIF และ DDF **สมการวิเคราะห์** : $Y_{ik} = B0_k+B1_k*(X_{ik})_{ik} + R_{ik}$
ระดับที่ 2	**ตัวแปรสถานศึกษา** : มี **ตัวแปรตาม** : $B0_k$,Bp_k สมการวิเคราะห์ : $B0_k=G00+G01*(Y_k) + U0_k$ $B1 = G10$ $B2 = G20$ . . . $Bp_k = Gp0$ คะแนนมูลค่าเพิ่ม : $U0_k$	**ตัวแปรสถานศึกษา** : มี **ตัวแปรตาม** : $B0_k$, Bp_k สมการวิเคราะห์ : $B0_k=G00+G01*(Y_k) + U0_k$ $B1 = G10$ $B2 = G20$ . . . $Bp_k = Gp0$ คะแนนมูลค่าเพิ่ม : $U0_k$	**ตัวแปรสถานศึกษา** : ไม่มี **ตัวแปรตาม** : $B0_k$ สมการวิเคราะห์ : $B0_k = G00 + U0_k$ คะแนนมูลค่าเพิ่ม : $U0_k$	**ตัวแปรสถานศึกษา** : มี **ตัวแปรตาม** : $B0_k$,Bp_k สมการวิเคราะห์ : $B0_k=G00+G01*(Y_k) + U0_k$ $B1 = G10$ $B2 = G20$ . . . $Bp_k = Gp0$ คะแนนมูลค่าเพิ่ม : $U0_k$

ขั้นตอน ที่ 4 การตรวจสอบผลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา

โมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาในการวิจัยครั้งนี้มี 4 โมเดล ผู้วิจัยมีการตรวจสอบผลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาโดยใช้ค่าจากการวิเคราะห์ ดังนี้

4.1 คะแนนมูลค่าเพิ่มจากโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา 4 โมเดลซึ่งเป็นค่าเศษเหลือที่ได้จากการวิเคราะห์ระดับที่ 2 (r_{0k}) หรือผลต่างระหว่างค่าเฉลี่ยของค่าความสามารถ จากการประเมิน วิชาคณิตศาสตร์ของนักเรียน ที่วัดได้จริงกับค่าเฉลี่ยของ ค่าความสามารถจากการประเมินวิชาคณิตศาสตร์ของนักเรียน ที่คาดหวัง ซึ่งประมาณค่าจากชุดของตัวแปรที่นำมาควบคุมทั้งในระดับนักเรียนและสถานศึกษา

4.2 การจัดอันดับคุณภาพ (Ranking) การจัดการศึกษาของสถานศึกษาโดยการเรียงระดับคะแนนมูลค่าเพิ่มจากคะแนนสูงไปคะแนนต่ำ จากผลการวิเคราะห์ในโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา 4 โมเดล

4.3 การจัดระดับคุณภาพ (Rating) การจัดการศึกษาของสถานศึกษา โดยการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษา ผู้วิจัยแบ่งกลุ่มโดยใช้ตำแหน่งเปอร์เซ็นต์ไทล์ และจัดกลุ่มเป็น 3 กลุ่ม คือ (ประกฤติยา ทักษิโณ, 2552)

กลุ่มที่ 1 เป็นกลุ่มต่ำ มีตำแหน่งเปอร์เซ็นต์ไทล์ 1 – 24.99

กลุ่มที่ 2 เป็นกลุ่มปานกลาง มีตำแหน่งเปอร์เซ็นต์ไทล์ 25 – 74.99

กลุ่มที่ 3 เป็นกลุ่มสูง มีตำแหน่งเปอร์เซ็นต์ไทล์ 75 – 100

4.4 ค่าความสามารถในการอธิบายความแปรปรวน ของตัวแปรตามด้วยตัวแปรทำนาย หรือค่าสัมประสิทธิ์การทำนายในแต่ละโมเดล (R^2) มีสมการดังนี้

$$R^2 = \frac{\text{ความแปรปรวนของค่า residual ที่ลดลงเมื่อมีตัวแปรทำนาย}}{\text{ความแปรปรวนของค่า residual ที่ลดลงเมื่อไม่มีตัวแปรทำนาย}}$$

ความแปรปรวน ของตัวแปรตามที่อธิบายได้ด้วยตัวแปรทำนายในแต่ละระดับทั้งระดับนักเรียนและระดับสถานศึกษา คำนวณได้จากผลคูณระหว่างความแปรปรวนที่อธิบายได้ตามโมเดลสมมติฐานกับความแปรปรวนพื้นหมดในระดับนั้นที่ได้จากโมเดลที่ไม่มีเงื่อนไขอย่างสมบูรณ์ ดังนี้

ความแปรปรวนที่อธิบายได้จากระดับที่ i = $\sigma_i^2(\exp lained) \times \sigma_i^2$

ความแปรปรวนทั้งหมดที่อธิบายได้โดยโมเดลทั้งสองระดับ เท่ากับ ผลรวมของความแปรปรวนที่อธิบายได้จากระดับที่ 1 กับระดับที่ 2

4.5 ค่าสหสัมพันธ์ของมูลค่าเพิ่ม ของสถานศึกษาที่ได้จากการวิเคราะห์ในแต่ละโมเดล เพื่อใช้ค่าที่บ่งบอกถึงผลการตรวจสอบคุณภาพการจัดการศึกษาที่มีความสอดคล้องกันในแต่ละโมเดล โดยใช้สหสัมพันธ์ตำแหน่งคะแนนของสเปียร์แมน (Spearman's Rank-Order Correlation)

บทที่ 4

ผลการวิเคราะห์ข้อมูล

งานวิจัยเรื่องการวิเคราะห์เปรียบเทียบโมเดลการประเมินคุณภาพ การจัดการศึกษาวิชาคณิตศาสตร์ : การประยุกต์ใช้โมเดลมูลค่าเพิ่มที่มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและการวิเคราะห์การทำหน้าที่ต่างกันของตัวลวง โดยมีวัตถุประสงค์เ พื่อ 1. เพื่อศึกษาผลการตรวจสอบ การทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) ในแบบสอบประเมินความรู้วิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 2 2. เพื่อศึกษาผลของข้อสอบที่ทำหน้าที่ต่างกัน และผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษ าต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน 4 โมเดล และ 3. เพื่อเปรียบเทียบประสิทธิภาพของโมเดลการประเมินคุณภาพการจัดการศึกษาวิชา คณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษา 4 โมเดล โดยมีขั้นตอ นขั้นตอนที่ใช้ในการวิเคราะห์ที่สำคัญ นั่นคือการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและการทำหน้าที่ต่างกันของตัวลวง ด้วยวิธีการวิเคราะห์แมนเทล- แฮนส์เซลร่วมกับอัตราส่วนแต้มต่อ (odds ratio) โดยใช้โปรแกรม DDFS (Penfield, 2010), การหาค่าความสามารถของผู้สอบโดยใช้ โปรแกรม MULTILOG และการวิเคราะห์มูลค่าเพิ่มภายใต้โมเดล HLM

ผู้วิจัยได้กำหนดการนำเสนอผลการวิเคราะห์ข้อมูลเพื่อตอบคำถามตามวัตถุประสงค์ที่ใช้ในการวิจัย โดยแบ่งการนำเสนอผลการวิเคราะห์ข้อมูลเป็น 4 ตอน ดังนี้

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลพื้นฐาน การตรวจสอบเงื่อนไขการวิเคราะห์พหุระดับ

ตอนที่ 2 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบและการตรวจสอบการทำหน้าที่ต่างกันของตัวลวงของข้อสอบแต่ละฉบับ

ตอนที่ 3 ผลของข้อสอบที่ทำหน้าที่ต่างกัน และผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาต่อการประเมินคุณภาพการจัด การศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน 4 โมเดล

ตอนที่4 ผลการเปรียบเทียบประสิทธิภาพโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลพื้นฐาน การตรวจสอบเงื่อนไขการวิเคราะห์พหุระดับ

การนำเสนอผลการวิเคราะห์ข้อมูลพื้นฐาน แบ่งเป็น 3 ตอนคือ 1.1) ข้อมูลพื้นฐานของกลุ่มตัวอย่างเป็นการเสนอคุณลักษณะพื้นฐานของนักเรียนและสถานศึกษาเพื่อแสดงลักษณะการแจกแจงของข้อมูล ประกอบด้วยค่าสถิติเชิงบรรยายได้แก่ ค่าความถี่ ร้อยละ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน ค่าต่ำสุด ค่าสูงสุด ค่าความเบ้ และค่าความโด่ง 1.2) การตรวจสอบเงื่อนไขของการวิเคราะห์พหุระดับ เพื่อตรวจสอบข้อตกลงเบื้องต้นเกี่ยวกับความสัมพันธ์ของตัวแปรต้นและลักษณะการแจกแจงของความคลาดเคลื่อนก่อนการใช้โมเดลพหุระดับ

1.1 ข้อมูลพื้นฐานของกลุ่มตัวอย่าง

การวิเคราะห์ข้อมู ลพื้นฐานของกลุ่มตัวอย่าง ผู้วิจัยนำ ตัวแปรคุณลักษณะพื้นฐานของนักเรียนและสถานศึกษาที่มีอยู่ในฐานข้อมูลของโครงการการประเมินผลนักเรียน TIMSS 2007 มาวิเคราะห์เพื่อให้เห็นถึงลักษณะโครงการสร้างของกลุ่มตัวอย่าง

จากการวิเคราะห์ข้อมูลพื้นฐานของกลุ่มตัวอย่างที่เป็น ตัวแปรจัดประเภท จำแนกเป็นระดับนักเรียนและระดับสถานศึกษา ในระดับนักเรียน พบว่าเป็นเพศชาย คิดเป็นร้อยละ 45.40 และเพศหญิงร้อยละ 54.60 นักเรียนส่วนใหญ่มีระดับเวลาที่ใช้ในการทำการบ้านวิชาคณิตศาสตร์อยู่ในระดับปานกลาง คิดเป็นร้อยละ 45.73 รองลงมาคือนักเรียน มีระดับ เวลาที่ใช้ในการทำการบ้านวิชาคณิตศาสตร์อยู่ใน ระดับสูง คิดเป็นร้อยละ 39.62 และน้อยที่สุดคือนักเรียน มีระดับเวลาที่ใช้ในการทำการบ้านวิชาคณิตศาสตร์อยู่ในระดับ อยู่ในระดับต่ำ คิดเป็นร้อยละ 14.65 ในเรื่องของดัชนีความเชื่อมั่นในตนเองเกี่ยวกับการเรียนวิชาคณิตศาสตร์ ส่วนใหญ่อยู่ในระดับปานกลาง คิดเป็นร้อยละ 58.43 รองลงมาคือในระดับสูงคิดเป็นร้อยละ 22.67 และระดับต่ำคิดเป็นร้อยละ 18.90 ด้านความเพลิดเพลินในการเรียนวิชาคณิตศาสตร์ ส่วนใหญ่อยู่ในระดับสูง คือคิดเป็นร้อยละ 56.30 รองลงมาคือมีความเพลิดเพลินในการเรียนวิชาคณิตศาสตร์ ในระดับปานกลาง คิดเป็นร้อยละ 30.56 และระดับต่ำร้อยละ 13.14 ดัชนีของนักเรียนในความรู้สึกปลอดภัยในการมาเรียนหนังสือในสถานศึกษา ส่วนใหญ่อยู่ในระดับปานกลาง คิดเป็นร้อยละ 47.45 ส่วนรองลงมาอยู่ในระดับสูง คิดเป็นร้อยละ 30.64 และร้อยละ 21.91 ตามลำดับ

ในระดับสถานศึกษา พบว่าสถานศึกษาขนาดเล็ก กลาง ใหญ่ และใหญ่พิเศษ คิดเป็นร้อยละ 28.67, 37.33, 15.33 และ18.67 ตามลำดับ ตำแหน่งที่ตั้งของสถานศึกษาอยู่ในหมู่บ้าน ตำบล เมืองเล็ก เมือง เมืองใหญ่ และเมืองใหญ่มาก คิดเป็นร้อยละ 11.33, 18.68, 27.33, 19.33, 13.33 และ 10 ตามลำดับ ในขณะที่ขนาดของห้องเรียนมีขนาด นักเรียน 1-24 คน นักเรียน 25-40 คน และนักเรียนมากกว่า 41 คน คิดเป็นร้อยละ 19, 73 และ 58 ตามลำดับ ระดับการขาดแคลนครูคณิตศาสตร์ในสถานศึกษา ไม่ขาดแคลน ขาดแคลนน้อย ขาดแคลนบางส่วน และขนาดแคลนจำนวนมาก คิดเป็นร้อยละ 48, 6, 49 และ 47 ตามลำดับ ดัชนีการให้การสนับสนุนด้านทรัพยากรที่ใช้ในการจัดการเรียนการสอนคณิตศาสตร์ มีระดับเรียงตาม ระดับต่ำ ระดับปานกลาง และระดับสูง คิดเป็นร้อยละ 12.67, 66.67 และ20.67 ตามลำดับ ดัชนีการรับรู้ถึงบรรยากาศที่ดีในที่ทำงาน มีระดับเรียงตาม ระดับต่ำ ระดับปานกลาง และระดับสูง คิดเป็นร้อยละ 18, 68.67 และ 13.33 ตามลำดับ ดัชนีความพึงพอใจของครูที่มีต่อการจัดการภายในสถานศึกษา มีระดับเรียงตาม ระดับต่ำ ระดับปานกลาง และระดับสูง คิดเป็นร้อยละ 16, 68 และ 16 ตามลำดับ ดัชนีการให้ความสำคัญกับการให้การบ้านคณิตศาสตร์ของครู ในสถานศึกษา มีระดับเรียงตาม ระดับต่ำ ระดับปานกลาง และระดับสูง คิดเป็นร้อยละ 15, 71 และ 64 ตามลำดับ และระดับความถี่ในการทำงานร่วมกันของครูคณิตศาสตร์ในสถานศึกษา ไม่เคยเลย 2-3 ครั้งต่อเดือน และเป็นประจำ คิดเป็นร้อยละ 5.33, 55.33 และ 39.33 ตามลำดับ แสดงดังตารางที่ 17 ต่อไปนี้

ตารางที่ 17 แสดงความถี่ ร้อยละ ของตัวแปรจัดประเภทจำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา

คุณลักษณะของนักเรียนและสถานศึกษา	ความถี่	ร้อยละ
ระดับนักเรียน		
1. เพศ		
1.1 หญิง	2955	54.60
1.2 ชาย	2457	45.40

ตารางที่ 17 แสดงความถี่ ร้อยละ ของตัวแปรจัดประเภทจำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา (ต่อ)

คุณลักษณะของนักเรียนและสถานศึกษา	ความถี่	ร้อยละ
2. ดัชนีระดับเวลาที่ใช้ในการทำการบ้านวิชาคณิตศาสตร์		
2.1 ระดับต่ำ	793	14.65
2.2 ระดับปานกลาง	2475	45.73
2.3 ระดับสูง	2144	39.62
3. ดัชนีของความเชื่อมั่นในตนเองเกี่ยวการเรียนวิชาคณิตศาสตร์		
3.1 ระดับต่ำ	1023	18.90
3.2 ระดับปานกลาง	3162	58.43
3.3 ระดับสูง	1227	22.67
4. ดัชนีความเพลิดเพลินในการเรียนวิชาคณิตศาสตร์		
4.1 ระดับต่ำ	711	13.14
4.2 ระดับปานกลาง	1654	30.56
4.3 ระดับสูง	3047	56.30
5. ดัชนีของนักเรียนในความรู้สึกปลอดภัยในการมาเรียนหนังสือภายในสถานศึกษา		
5.1 ระดับต่ำ	1186	21.91
5.2 ระดับปานกลาง	2568	47.45
5.3 ระดับสูง	1658	30.64
รวม	**5,412**	**100**
<u>ระดับสถานศึกษา</u>		
1. ขนาดของสถานศึกษา		
1.1 ขนาดเล็ก (นักเรียน น้อยกว่า 500 คน)	43	28.67
1.2 ขนาดกลาง (นักเรียน 500 – 1,499 คน)	56	37.33
1.3 ขนาดใหญ่ (นักเรียน 1500 – 2,499 คน)	23	15.33
1.4 ขนาดใหญ่พิเศษ (นักเรียน ตั้งแต่ 2,500 คนขึ้นไป)	28	18.67
2. ตำแหน่งที่ตั้งสถานศึกษา		
2.1 หมู่บ้าน	17	11.33
2.2 ตำบล	28	18.68
2.3 เมืองเล็ก	41	27.33

ตารางที่ 17 แสดงความถี่ ร้อยละ ของตัวแปรจัดประเภทจำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา (ต่อ)

คุณลักษณะของนักเรียนและสถานศึกษา	ความถี่	ร้อยละ
2.4 เมือง	29	19.33
2.5 เมืองใหญ่	20	13.33
2.6 เมืองใหญ่มาก	15	10.00
3. ขนาดของห้องเรียน		
3.1 นักเรียน 1 – 24 คน	19	12.67
3.2 นักเรียน 25 – 40 คน	73	48.67
3.3 นักเรียน มากกว่า 41 คน	58	38.67
4. ระดับการขาดแคลนครูคณิตศาสตร์ในสถานศึกษา		
4.1 ไม่ขาดแคลน	48	32.00
4.2 ขาดแคลนน้อย	6	4.00
4.3 ขาดแคลนเป็นบางส่วน	49	32.67
4.3 ขาดแคลนจำนวนมาก	47	31.33
5. ดัชนีการให้การสนับสนุนด้านทรัพยากรที่ใช้ในการจัดการเรียนการสอนคณิตศาสตร์		
5.1 ระดับต่ำ	19	12.67
5.2 ระดับปานกลาง	100	66.67
5.3 ระดับสูง	31	20.67
6. ดัชนีการรับรู้ถึงบรรยากาศที่ดีในที่ทำงาน		
6.1 ระดับต่ำ	27	18.00
6.2 ระดับปานกลาง	103	68.67
6.3 ระดับสูง	20	13.33
7. ดัชนีความพึงพอใจของครูที่มีต่อการจัดการภายในสถานศึกษา		
7.1 ระดับต่ำ	24	16
7.2 ระดับปานกลาง	102	68
7.3 ระดับสูง	24	16

ตารางที่ 17 แสดงความถี่ ร้อยละ ของตัวแปรจัดประเภทจำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา (ต่อ)

คุณลักษณะของนักเรียนและสถานศึกษา	ความถี่	ร้อยละ
8. ดัชนีการให้ความสำคัญกับการให้การบ้านคณิตศาสตร์ของครูในสถานศึกษา		
8.1 ระดับต่ำ	15	10.00
8.2 ระดับปานกลาง	71	47.33
8.3 ระดับสูง	64	42.67
9. ระดับความถี่ในการทำงานร่วมกันของครูคณิตศาสตร์ในสถานศึกษา		
9.1 ไม่เคยเลย	8	5.33
9.2 2-3 ครั้งต่อเดือน	83	55.33
9.3 เป็นประจำ	59	39.33
รวม	**150**	**100**

ผลการวิเคราะห์คุณลักษณะพื้นฐานของกลุ่มตัวอย่างที่เป็นตัวแปรต่อเนื่อง จำแนกเป็นระดับนักเรียนซึ่งประกอบด้วยตัวแปรทั้งหมด 8 ตัวแปรและระดับสถานศึกษา ประกอบด้วยตัวแปร 4 ตัวแปร ระดับนักเรียน พบว่า ส่วนใหญ่มีลักษณะการแจกแจงเป็นแบบเบ้ซ้าย (skewness < 0) คือตัวแปรผลสัมฤทธิ์ทางการเรียนที่ผ่านมา ความมั่งคั่งของครอบครัว การใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเอง แรงจูงใจภายนอกในการเรียนคณิตศาสตร์ และการเห็นคุณค่าในวิชาคณิตศาสตร์ แสดงให้เห็นว่า ส่วนใหญ่ค่าเฉลี่ยผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ ของตัวแปรระดับนักเรียน จำแนกตามคุณลักษณะต่างๆ มีค่าสูงกว่าค่าเฉลี่ย ส่วนลักษณะการแจกแจงโดยพิจารณาจากค่า kurtosis พบว่าส่วนใหญ่มีการแจกแจงที่สูงโด่งเกินกว่าโค้งปกติ (kurtosis > 0) คือตัวแปรแหล่งทรัพยากรที่บ้าน กิจกรรมการฝึกปฏิบัติในวิชาคณิตศาสตร์ การใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเอง แรงจูงใจภายนอกในการเรียนคณิตศาสตร์ และการเห็นคุณค่าในวิชาคณิตศาสตร์ เมื่อพิจารณาสัมประสิทธิ์ของความแปรผัน (%cv) พบว่าตัวแปรระดับนักเรียนที่มีการกระจายสูงที่สุด คือ ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา ($\overline{X}$ = 2.84, %cv = 57.39) รองลงมาคือระดับการศึกษาของผู้ปกครอง ($\overline{X}$ = 2.97, %cv = 54.88) ส่วนตัวแปรระดับนักเรียนที่มีการ

กระจายน้อยที่สุด คือ ความมั่งคั่งของครอบครัว ($\overline{X}$ = 14.97, %cv = 68.74) และตามด้วยการเห็นคุณค่าในวิชาคณิตศาสตร์ ($\overline{X}$ = 1.58, %cv = 22.15)

ระดับสถานศึกษา พบว่า ส่วนใหญ่มีลักษณะการแจกแจงเป็นแบบเบ้ซ้าย (skewness < 0) คือตัวแปร การให้การสนับสนุนของผู้ปกครอง การจัดกิจกรรมที่สนับสนุนการเรียนรู้ในวิชาคณิตศาสตร์ภายใน และการพัฒนาครูในสถานศึกษา แสดงให้เห็นว่าส่วนใหญ่ค่าเฉลี่ยผลสัมฤทธิ์ ทางการเรียนคณิตศาสตร์ ของตัวแปรระดับสถานศึกษาจำแนกตามคุณลักษณะต่างๆ มีค่าสูงกว่าค่าเฉลี่ย ส่วนลักษณะการแจกแจงโดยพิจารณาจากค่า kurtosis พบว่าส่วนใหญ่มีการแจกแจงที่ แบนกว่าโค้งปกติ (kurtosis < 0) คือตัวแปร การจัดกิจกรรมที่สนับสนุนการเรียนรู้ในวิชาคณิตศาสตร์ภายใน การพัฒนาครูในสถานศึกษา และประสบการณ์การทำงานของครูในสถานศึกษา เมื่อพิจารณาสัมประสิทธิ์ของความแปรผัน (%cv) พบว่าตัวแปรระดับสถานศึกษาที่ มีการกระจายสูงที่สุด คือ ประสบการณ์การทำงานของครูในสถานศึกษา ($\overline{X}$ = 3.49, %cv = 14.04) รองลงมาคือการจัดกิจกรรมที่สนับสนุนการเรียนรู้ในวิชาคณิตศาสตร์ภายใน ($\overline{X}$ = 3.49, %cv = 14.04) ส่วนตัวแปรระดับสถานศึกษาที่มีการกระจายน้อยที่สุด คือการให้การสนับสนุนของผู้ปกครอง ($\overline{X}$ = 1.86, %cv = 10.75) และตามด้วยการพัฒนาครูในสถานศึกษา($\overline{X}$ = 3.81, %cv = 17.59) แสดงรายละเอียด ดังตารางที่ 18

ตารางที่ 18 แสดงค่าสถิติพื้นฐานของตัวแปรต่อเนื่องจากการศึกษาแนวโน้มผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์นานาชาติ (TIMSS 2007) จำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา

คุณลักษณะของตัวแปร	N	Min	Max	$\overline{X}$	SD	%CV	Sk	Ku
ระดับนักเรียน								
1. ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา	5412	0	5	2.84	1.63	57.39	-0.56	-0.91
2. ระดับการศึกษาของผู้ปกครอง	5412	1	7	2.97	1.63	54.88	0.63	-0.42
3. แหล่งทรัพยากรที่บ้าน	5412	1	5	2.17	0.99	45.62	0.84	0.57
4. ความมั่งคั่งของครอบครัว	5412	1	2	1.66	0.19	11.45	-0.35	-0.17
5. กิจกรรมการฝึกปฏิบัติในวิชาคณิตศาสตร์	5412	1.12	4	2.72	0.40	14.71	0.07	0.04
6. การใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเอง	5412	1	4	3.27	0.74	22.63	-0.77	0.19

ตารางที่ 18 แสดงค่าสถิติพื้นฐานของตัวแปรต่อเนื่องจากการศึกษาแนวโน้มผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์นานาชาติ (TIMSS 2007) จำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา (ต่อ)

คุณลักษณะของตัวแปร	N	Min	Max	$\overline{X}$	SD	%CV	Sk	Ku
7. แรงจูงใจภายนอกในการเรียนคณิตศาสตร์	5412	1	4	3.61	0.54	14.96	-1.41	1.85
8. การเห็นคุณค่าในวิชาคณิตศาสตร์	5412	1	4	3.49	0.49	14.04	-1.07	1.60
สถานศึกษา								
1. การให้การสนับสนุนของผู้ปกครอง	150	1.2	2	1.86	0.20	10.75	-1.46	1.40
2. การจัดกิจกรรมที่สนับสนุนการเรียนรู้ในวิชาคณิตศาสตร์ภายใน	150	1	2	1.58	0.35	22.15	-0.24	-0.96
3. การพัฒนาครูในสถานศึกษา	150	2	5	3.81	0.67	17.59	-0.26	-0.11
4. ประสบการณ์การทำงานของครูในสถานศึกษา	150	0	36	14.97	10.29	68.74	0.26	-1.25

การวิเคราะห์คะแนนความสามารถรายบุคคล (θ) ในวิชาคณิตศาสตร์ของนักเรียนจากการประเมินตามโครงการ TIMSS 2007 เป็นการวิเคราะห์โดยใช้โมเดลการตอบสนองข้อสอบ (Item Response Theory) ดังนั้นผลการวิเคราะห์ส่วนใหญ่มักให้ค่าอยู่ในช่วง -3 ถึง 3 โดยผลการวิเคราะห์ข้อมูล พบว่าค่าระดับความสามารถของผู้สอบ พบว่านักเรียนและสถานศึกษาที่มีคุณลักษณะพื้นฐานต่างกัน จะมีค่าความสามารถเฉลี่ยวิชาคณิตศาสตร์แตกต่างกันในแต่ละคุณลักษณะของตัวแปร เมื่อพิจารณาในระดับนักเรียน พบว่า านักเรียนทั้งหมดมีค่าความสามารถ (θ) เฉลี่ยวิชาคณิตศาสตร์เท่ากับ -0.55 มีส่วนเบี่ยงเบนมาตรฐาน 0.84 มีลักษณะการแจกแจงของข้อมูลเป็นแบบเบ้ขวา แสดงว่านักเรียนส่วนใหญ่มีค่าความสามารถต่ำกว่าค่าเฉลี่ย (skewness = 0.64) และมีการแจกแจงที่ สูงกว่าโค้งปกติ (kurtosis = 0.17) โดยเมื่อจำแนกตามคุณลักษณะของตัวแปร เฉพาะที่มีค่าความสามารถ (θ) เฉลี่ยวิชาคณิตศาสตร์ สูงที่สุด พบว่านักเรียนหญิงมีค่าเฉลี่ยความสามารถสูงกว่านักเรียนชาย (ค่าเฉลี่ยความสามารถนักเรียนหญิงและชาย เท่ากับ -0.47 และ -0.65 ตามลำดับ) นักเรียนที่มีระดับเวลาใน การทำการบ้านวิชาคณิตศาสตร์อยู่ในระดับสูงจะมีค่าเฉลี่ยความสามารถสูงที่สุด ($\overline{X}$ = -0.38, SD = 0.84) นักเรียนที่มีระดับความเชื่อมั่นในตนเองเกี่ยวกับการเรียนวิชาคณิตศาสตร์อยู่ในระดับสูง จะมีค่าเฉลี่ย

ความสามารถสูงที่สุด ($\overline{X}$ = -0.11, SD = 0.96) นักเรียนที่มีความเพลิดเพลินในการเรียนวิชาคณิตศาสตร์อยู่ในระดับสูง จะมีค่าเฉลี่ยความสามารถสูงที่สุ ด ($\overline{X}$ = -0.41, SD = 0.88) นักเรียนที่มี ความรู้สึกปลอดภัยในการมาเรียนหนังสือภายในสถานศึกษา สูง จะมีค่าเฉลี่ยความสามารถสูงที่สุด ($\overline{X}$ = -0.47, SD = 0.85)

ระดับสถานศึกษา พบว่า สถานศึกษา ทั้งหมดมีค่าความสามารถ (θ) เฉลี่ยวิชาคณิตศาสตร์เท่ากับ -0.63 มีส่วนเบี่ยงเบนมาตรฐาน 0.57 มีลักษณะการแจกแจงของข้อมูลเป็นแบบเบ้ขวา แสดงว่าสถานศึกษาส่วนใหญ่มีค่าความสามารถต่ำกว่าค่าเฉลี่ย (skewness = 1.20) และมีการแจกแจงที่ สูงกว่าโค้งปกติ (kurtosis = 1.11) โดยเมื่อจำแนกตามคุณลักษณะของตัวแปรเฉพาะที่มีค่าความสามารถ (θ) เฉลี่ยวิชาคณิตศาสตร์สูงที่สุด พบว่า สถานศึกษาเล็กจะมีค่าเฉลี่ยความสามารถนักเรียนสูงที่สุด ($\overline{X}$ = -0.52, SD = 0.63) สถานศึกษาที่ตั้งอยู่ในตำบลจะมีค่าเฉลี่ยความสามารถนักเรียนสูงที่สุด ($\overline{X}$ = -0.53, SD = 0.68) สถานศึกษาที่มีขนาดห้องเรียน มีนักเรียน 1-24 คน จะมีค่าเฉลี่ยความสามารถนักเรี ยนสูงที่สุด ($\overline{X}$ = -0.51, SD = 0.67) สถานศึกษาที่มีระดับการขาดแคลนครูคณิตศาสตร์ เป็นไม่ขาดแคลน จะมีค่าเฉลี่ยความสามารถนักเรียนสูงที่สุด ($\overline{X}$ = -0.50, SD = 0.64) สถานศึกษาที่มีการให้การสนับสนุนด้านทรัพยากรที่ใช้ในกา รจัดการเรียนการสอนคณิตศาสตร์ ในระดับต่ำ จะมีค่าเฉลี่ยความสามารถนักเรียนสูงที่สุด ($\overline{X}$ = -0.61, SD = 0.65) สถานศึกษาที่มี การรับรู้ถึงบรรยากาศในที่ดีในที่ทำงาน ของครู อยู่ในระดับปานกลาง และระดับสูง จะมีค่าเฉลี่ยความสามารถนักเรียนสูงที่สุด ($\overline{X}$ = -0.62, SD = 0.61 และ 0.52 ตามลำดับ) สถานศึกษาที่มี ระดับ ความพึงพอใจของครูที่มีต่อการจัดการภายในสถานศึกษา อยู่ในระดับสูง จะมีค่าเฉลี่ยความสามารถนักเรียนสูงที่สุด ($\overline{X}$ = -0.60, SD = 0.55) สถานศึกษาที่มีการให้ความสำคัญกับการให้การบ้านคณิตศาสตร์ของครูในสถานศึกษาอยู่ในระดับต่ำ จะมีค่าเฉลี่ยความสามารถนักเรียนสูงที่สุด ($\overline{X}$ = -0.40, SD = 0.78) สถานศึกษาที่มี ระดับความถี่ในการทำงานร่วมกันของครูคณิตศาสตร์ในสถานศึกษา ในระดับไม่เคยเลย จะมีค่าเฉลี่ยความสามารถนัก เรียนสูงที่สุด ($\overline{X}$ = -0.49, SD = 0.57) แสดงรายละเอียดดังตารางที่ 19

ตารางที่ 19 แสดงค่าสถิติพื้นฐานของคะแนนความสามารถรายบุคคลในวิชาคณิตศาสตร์จากการศึกษาแนวโน้มผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์นานาชาติ (TIMSS 2007) จำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา

คุณลักษณะของตัวแปร	N	Range	Min	Max	$\overline{X}$	SD	Var	Sk	Ku
ระดับนักเรียน									
1. เพศ									
1.1 หญิง	2955	4.87	-2.44	2.43	-0.47	0.84	0.71	0.57	0.03
1.2 ชาย	2457	4.92	-2.51	2.41	-0.65	0.83	0.69	0.75	0.43
2. ดัชนีระดับเวลาที่ใช้ในการทำการบ้านวิชาคณิตศาสตร์									
2.1 ระดับต่ำ	793	4.88	-2.51	2.38	-0.74	0.80	0.64	0.94	0.72
2.2 ระดับปานกลาง	2475	4.87	-2.44	2.43	-0.63	0.83	0.69	0.76	0.46
2.3 ระดับสูง	2144	4.86	-2.43	2.43	-0.38	0.84	0.70	0.45	-0.04
3. ดัชนีของความเชื่อมั่นในตนเองเกี่ยวการเรียนวิชาคณิตศาสตร์									
3.1 ระดับต่ำ	1023	4.34	-2.41	1.93	-0.66	0.73	0.53	0.62	0.33
3.2 ระดับปานกลาง	3162	4.88	-2.51	2.38	-0.68	0.77	0.59	0.64	0.36
3.3 ระดับสูง	1227	4.83	-2.40	2.43	-0.11	0.96	0.93	0.24	-0.58
4. ดัชนีความเพลิดเพลินในการเรียนวิชาคณิตศาสตร์									
4.1 ระดับต่ำ	711	4.30	-2.41	1.89	-0.67	0.74	0.55	0.56	0.19
4.2 ระดับปานกลาง	1654	4.88	-2.51	2.38	-0.75	0.75	0.57	0.77	0.72
4.3 ระดับสูง	3047	4.86	-2.43	2.43	-0.41	0.88	0.78	0.53	-0.12
5. ดัชนีของนักเรียนในความรู้สึกปลอดภัยในการมาเรียนหนังสือภายในสถานศึกษา									
5.1 ระดับต่ำ	1186	4.83	-2.43	2.40	-0.65	0.84	0.70	0.81	0.58
5.2 ระดับปานกลาง	2568	4.84	-2.41	2.43	-0.55	0.83	0.69	0.66	0.22
5.3 ระดับสูง	1658	4.93	-2.51	2.43	-0.47	0.85	0.73	0.51	-0.09
รวม	5412	4.93	-2.51	2.43	-0.55	0.84	0.71	0.64	0.17

ตารางที่ 19 แสดงค่าสถิติพื้นฐานของคะแนนความสามารถรายบุคคลในวิชาคณิตศาสตร์จากการศึกษาแนวโน้มผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์นานาชาติ (TIMSS 2007) จำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา (ต่อ)

คุณลักษณะของตัวแปร	N	Range	Min	Max	$\overline{X}$	SD	Var	Sk	Ku
สถานศึกษา									
1. ขนาดของสถานศึกษา									
1.1 ขนาดเล็ก	43	2.68	-1.38	1.30	-0.52	0.63	0.40	1.07	1.01
1.2 ขนาดกลาง	56	2.02	-1.32	0.70	-0.68	0.51	0.26	1.22	1.01
1.3 ขนาดใหญ่	23	2.21	-1.32	0.89	-0.66	0.60	0.36	1.38	1.22
1.4 ขนาดใหญ่พิเศษ	28	2.38	-1.39	0.99	-0.65	0.55	0.31	1.24	1.81
2. ตำแหน่งที่ตั้งสถานศึกษา									
2.1 หมู่บ้าน	17	1.71	-1.25	0.46	-0.67	0.45	0.20	0.93	1.03
2.2 ตำบล	28	2.63	-1.33	1.30	-0.53	0.68	0.46	1.23	1.17
2.3 เมืองเล็ก	41	2.27	-1.38	0.89	-0.68	0.54	0.29	1.24	1.69
2.4 เมือง	29	2.31	-1.32	0.99	-0.58	0.63	0.39	1.22	0.52
2.5 เมืองใหญ่	20	1.95	-1.23	0.72	-0.56	0.56	0.31	0.80	-0.19
2.6 เมืองใหญ่มาก	15	1.78	-1.39	0.39	-0.83	0.46	0.21	1.29	2.21
3. ขนาดของห้องเรียน									
3.1 นักเรียน 1 – 24 คน	19	2.43	-1.35	1.08	-0.51	0.67	0.45	0.92	0.45
3.2 นักเรียน 25 – 40 คน	73	2.68	-1.38	1.30	-0.65	0.55	0.30	1.41	2.06
3.3 นักเรียน มากกว่า 41 คน	58	2.38	-1.39	0.99	-0.63	0.56	0.31	1.09	0.64
4. ระดับการขาดแคลนครูคณิตศาสตร์ในสถานศึกษา									
4.1 ไม่ขาดแคลน	48	2.38	-1.39	0.99	-0.50	0.64	0.40	0.81	-0.22
4.2 ขาดแคลนน้อย	6	0.36	-1.14	-0.78	-1.02	0.13	0.02	1.62	2.96
4.3 ขาดแคลนเป็นบางส่วน	49	1.68	-1.38	0.30	-0.79	0.39	0.15	0.61	-0.24
4.4 ขาดแคลนจำนวนมาก	47	2.62	-1.32	1.30	-0.54	0.63	0.40	1.21	0.99
5. ดัชนีการให้การสนับสนุนด้านทรัพยากรที่ใช้ในการจัดการเรียนการสอนคณิตศาสตร์									
5.1 ระดับต่ำ	19	2.31	-1.32	0.99	-0.61	0.65	0.42	1.44	1.34
5.2 ระดับปานกลาง	100	2.69	-1.39	1.30	-0.63	0.55	0.30	1.18	1.39
5.3 ระดับสูง	31	2.21	-1.32	0.89	-0.62	0.59	0.35	1.18	0.79

ตารางที่ 19 แสดงค่าสถิติพื้นฐานของคะแนนความสามารถรายบุคคลในวิชาคณิตศาสตร์จากการศึกษาแนวโน้มผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์นานาชาติ (TIMSS 2007) จำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา (ต่อ)

คุณลักษณะของตัวแปร	N	Range	Min	Max	$\overline{X}$	SD	Var	Sk	Ku
6. ดัชนีการรับรู้ถึงบรรยากาศในที่ดีในที่ทำงาน									
6.1 ระดับต่ำ	27	1.78	-1.32	0.46	-0.66	0.44	0.19	0.53	0.35
6.2 ระดับปานกลาง	103	2.69	-1.39	1.30	-0.62	0.61	0.37	1.21	0.87
6.3 ระดับสูง	20	2.11	-1.22	0.89	-0.62	0.52	0.27	1.49	2.49
7. ดัชนีความพึงพอใจของครูที่มีต่อการจัดการภายในสถานศึกษา									
7.1 ระดับต่ำ	24	2.65	-1.35	1.30	-0.67	0.59	0.35	1.95	4.65
7.2 ระดับปานกลาง	102	2.47	-1.39	1.08	-0.62	0.57	0.33	1.15	0.80
7.3 ระดับสูง	24	2.08	-1.38	0.70	-0.60	0.55	0.30	0.75	0.26
8. ดัชนีการให้ความสำคัญกับการให้การบ้านคณิตศาสตร์ของครูในสถานศึกษา									
8.1 ระดับต่ำ	15	2.24	-1.25	0.99	-0.40	0.78	0.60	0.66	0.58
8.2 ระดับปานกลาง	71	2.47	-1.39	1.08	-0.69	0.49	0.24	1.24	0.28
8.3 ระดับสูง	64	2.65	-1.35	1.30	-0.62	0.59	0.34	1.21	0.30
9. ระดับความถี่ในการทำงานร่วมกันของครูคณิตศาสตร์ในสถานศึกษา									
9.1 ไม่เคยเลย	8	1.63	-1.17	0.46	-0.49	0.57	0.32	0.38	-0.80
9.2 2-3 ครั้งต่อเดือน	83	2.65	-1.35	1.30	-0.62	0.59	0.35	1.28	1.23
9.3 เป็นประจำ	59	2.47	-1.39	1.08	-0.66	0.54	0.29	1.19	1.39
รวม	150	2.69	-1.39	1.3	-0.63	0.57	0.32	1.20	1.11

1.2 การตรวจสอบเงื่อนไขของการวิเคราะห์พหุระดับ

การตรวจสอบเงื่อนไขเบื้องต้นของการวิเคราะห์เชิงการทำ นาย สิ่งที่ต้องระวังก่อนการใช้โมเดล คือการตรวจสอบ ภาวะร่วมเส้นตรงพหุ (Multicollinearity) และการตรวจสอบความสอดคล้องกับสมการการวิเคราะห์ถดถอยและโมเดลพหุระดับ

เพื่อให้สะดวกต่อการนำ เสนอผลการวิเคราะห์ในส่วนนี้และให้เกิดความเข้าใจในสัญลักษณ์ ของตัวแปรแต่ละตัว ผู้วิจัยจึงกำหนดสัญลักษณ์และความหมายแทนตัวแปรในระดับนักเรียนและระดับสถานศึกษา ดังนี้

ระดับตัวแปร	ชื่อตัวแปร	ความหมายของตัวแปร
ระดับนักเรียน	THETA	ค่าความสามารถจากการประเมินความรู้ทางคณิตศาสตร์
	FEMALE	ตัวแปรดัมมี่ความเป็นเพศหญิง
	RELAGR	ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา
	PAEDU	ระดับการศึกษาของผู้ปกครอง
	LEARES	แหล่งทรัพยากรที่บ้าน
	WEALTH	ความมั่งคั่งของครอบครัว
	PRASCH	กิจกรรมการฝึกปฏิบัติในวิชาคณิตศาสตร์
	TIMEHOM	ดัชนีระดับเวลาที่ใช้ในการทำการบ้านวิชาคณิตศาสตร์
	INTMAT	การใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเอง
	SELFLEA	ดัชนีของความเชื่อมั่นในตนเองเกี่ยวการเรียนวิชาคณิตศาสตร์
	JOYMATH	ดัชนีความเพลิดเพลินในการเรียนวิชาคณิตศาสตร์
	INSTMAT	แรงจูงใจภายนอกในการเรียนคณิตศาสตร์
	VALMATH	การเห็นคุณค่าในวิชาคณิตศาสตร์
	SAFESCH	ดัชนีของนักเรียนในความรู้สึกปลอดภัยในการมาเรียนหนังสือภายในสถานศึกษา
ระดับสถานศึกษา	DSCHSL	ตัวแปรดัมมี่สถานศึกษาขนาดใหญ่พิเศษ
	DSCHL	ตัวแปรดัมมี่สถานศึกษาขนาดใหญ่
	DSCHM	ตัวแปรดัมมี่สถานศึกษาขนาดกลาง
	COMMXXL	ตัวแปรดัมมี่สถานศึกษาที่ตั้งอยู่ในเมืองใหญ่มาก
	COMMXL	ตัวแปรดัมมี่สถานศึกษาที่ตั้งอยู่ในเมืองใหญ่
	COMML	ตัวแปรดัมมี่สถานศึกษาที่ตั้งอยู่ในเมือง
	COMMM	ตัวแปรดัมมี่สถานศึกษาที่ตั้งอยู่ในเมืองเล็ก
	COMMS	ตัวแปรดัมมี่สถานศึกษาที่ตั้งอยู่ในตำบล
	SCQ21AC	ระดับการขาดแคลนครูคณิตศาสตร์ในสถานศึกษา

ระดับตัวแปร	ชื่อตัวแปร	ความหมายของตัวแปร
ระดับสถานศึกษา	SCQ20AE	การให้การสนับสนุนของผู้ปกครอง
	SCMAEDU	ดัชนีการให้การสนับสนุนด้านทรัพยากรที่ใช้ในการจัดการเรียนการสอนคณิตศาสตร์
	MATPROM	การจัดกิจกรรมที่สนับสนุนการเรียนรู้ในวิชาคณิตศาสตร์ภายใน
	PPSC	ดัชนีการรับรู้ถึงบรรยากาศในที่ดีในที่ทำงาน
	DVTESCH	การพัฒนาครูในสถานศึกษา
	EXPTEA	ประสบการณ์การทำงานของครูในสถานศึกษา
	ADEQ	ดัชนีความพึงพอใจของครูที่มีต่อการจัดการภายในสถานศึกษา
	HOMEWORK	ดัชนีการให้ความสำคัญกับการให้การบ้านคณิตศาสตร์ของครูใน
	CLASSSIZ	ขนาดของห้องเรียน
	TEACOLLA	ระดับความถี่ในการทำงานร่วมกันของครูคณิตศาสตร์ในสถานศึกษา

การตรวจสอบภาวะร่วมเส้นตรงพหุ (Multicollinearity)

การตรวจสอบความสัมพันธ์ของตัวแปรที่ใช้ในการวิจัย มุ่งตรวจสอบความสัมพันธ์ระหว่างตัวแปรต้นคู่ที่มีความสัมพันธ์กันสูงหรือภาวะร่วมเส้นตรงพหุ (multicollinearity) ถ้าหากพบตัวแปรต้นคู่ที่มีความสัมพันธ์กันในระดับสูง จะพิจารณาเลือกตัวแปรใดตัวแปรหนึ่งเข้าสู่สมการการวิเคราะห์แทนได้

จากการวิเคราะห์ความสัมพันธ์ของตัวแปรต่อเนื่องระหว่างตัวแปรระดับนักเรียน พบว่าตัวแปรที่มีความสัมพันธ์กันในระดับสูงอย่างมีนัยสำคัญที่ระดับ .01 ได้แก่ความสัมพันธ์ระหว่างความมั่งคั่งของครอบครัวกับแหล่งทรัพยากรที่บ้าน ($r_{WEALTH,\ LEARES} = 0.464$) รองลงมาคือการเห็นคุณค่าในวิชาคณิตศาสตร์ และแรงจูงใจภายนอกในการเรียนคณิตศาสตร์ ($r_{VALMATH,\ INSTMAT} = 0.420$) และความมั่งคั่งของครอบครัวกับระดับการศึกษาของผู้ปกครอง ($r_{WEALTH,\ PAEDU} = 0.362$) ส่วนตัวแปรคู่ที่ไม่มีความสัมพันธ์กัน ได้แก่ **การใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเองกับแหล่งทรัพยากรที่บ้าน** ($r_{INTMAT,\ LEARES} = 0.014$) **การใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเองกับความมั่งคั่งของครอบครัว** ($r_{INTMAT,\ WEALTH} = -0.020$) **และแรงจูงใจภายนอกในการเรียนคณิตศาสตร์** กับระดับ

การศึกษาของผู้ปกครอง ($r_{VALMATH,\ INSTMAT}$ = 0.026) นอกจากนี้ยังพบว่าตัวแปรตามหรือค่าความสามารถของนักเรียนจากการประเมินความรู้วิชาคณิตศาสตร์ในโครงการ TIMSS 2007 มีความสัมพันธ์กับตัวแปรต้นทุกตัวในระดับนักเรียนอย่างมีนัยสำคัญทางสถิติ แสดงดังตารางที่ 20

ตารางที่ 20 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรในระดับนักเรียน

	THETA	RELAGR	PAEDU	LEARES	WEALTH	PRASCH	INTMAT	INSTMAT	VALMATH
THETA	1.000								
RELAGR	0.222**	1.000							
PAEDU	0.237**	0.198**	1.000						
LEARES	0.328**	0.191**	0.325**	1.000					
WEALTH	0.352**	0.235**	0.362**	0.464**	1.000				
PRASCH	0.237**	0.103**	0.077**	0.230**	0.235**	1.000			
INTMAT	0.034*	0.046**	-0.065**	0.014	-0.020	0.204**	1.000		
INSTMAT	0.165**	0.115**	0.026	0.089**	0.128**	0.262**	0.298**	1.000	
VALMATH	0.094**	0.036**	-0.041**	0.024**	0.041**	0.257**	0.316**	0.420**	1.000

*p < .05 , **p < .01

การตรวจสอบภาวะร่วมเส้นตรงพหุ (multicollinearity) ผู้วิจัยใช้การวิเคราะห์ถดถอยพหุ (multiple regression) มีตัวแปรตามเป็นค่าความสามารถของนักเรียนจากการประเมินวิชาคณิตศาสตร์ในการทดสอบ TIMSS 2007 และตัวแปรทำนายเป็นตัวแปรต่อเนื่องในระดับนักเรียน ได้แก่ ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา (RELAGR) ระดับการศึกษาของผู้ปกครอง (PAEDU) แหล่งทรัพยากรที่บ้าน (LEARES) ความมั่งคั่งของครอบครัว (WEALTH) กิจกรรมการฝึกปฏิบัติในวิชาคณิตศาสตร์ (PRASCH) การใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเอง(INTMAT) แรงจูงใจภายนอกในการเรียนคณิตศาสตร์ (INSTMAT) และการเห็นคุณค่าในวิชาคณิตศาสตร์ (VALMATH) โดยพิจารณาจากค่า Tolerance และค่า VIF ซึ่ง Hair และคณะ (1998) เสนอว่าค่าสถิติเริ่มต้นของตัวแปรที่แสดงว่ามีภาวะร่วมเส้นตรงพหุ คือ มีค่า Tolerance ตั้งแต่ .10 ลงมา หรือ VIF มีค่า 10 ขึ้นไป

จากผลทดสอบภาวะร่วมเส้นตรงพหุ (Multicollinearity) ในระดับนักเรียน โดยพิจารณาค่า Tolerance และ Variance Inflation Factor: VIF) ของตัวแปรต้น พบว่า ผลการทดสอบความสัมพันธ์ระหว่างตัวแปรต้นแต่ละตัวมีค่า Tolerance สูง หรือมีค่าสูงกว่า .10 ตามข้อเสนอแนะของ Hair (Tolerance = .700 - .916) และค่า VIF ต่ำ หรือมีค่าต่ำกว่า 10 ตามข้อเสนอแนะของ Hair (VIF = 1.091 – 1.429) แสดงว่าตัวแปรต้นแต่ละตัวมีความสัมพันธ์กันต่ำ

จึงไม่เกิดปัญหาที่ตัวแปรต้นมีความสัมพันธ์กันสูงหรือปัญหาภาวะร่วมเส้นตรงพหุ (multicollinearity) แสดงรายละเอียด ดังตารางที่ 21

ตารางที่ 21 ค่า Tolerance และ VIF จากผลการทดสอบความสัมพันธ์ระหว่างตัวแปรต้นระดับนักเรียน

ตัวแปร	Collinearity Statistics	
	Tolerance	VIF
RELAGR	0.916	1.091
PAEDU	0.821	1.218
LEARES	0.737	1.357
WEALTH	0.700	1.429
PRASCH	0.833	1.200
INTMAT	0.848	1.179
INSTMAT	0.763	1.310
VALMATH	0.765	1.307

ผลการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรต้นที่เป็นตัวแปรต่อเนื่องในระดับสถานศึกษา พบว่า ตัวแปรที่มีความสัมพันธ์กันในระดับสูงอย่างมีนัยสำคัญที่ระดับ .01 ได้แก่ การพัฒนาครูในสถานศึกษากับ การให้การสนับสนุนของผู้ปกครอง ($r_{DVTESCH,\ SCQ20AE} = 0.264$) ส่วนตัวแปรต้นคู่ที่ไม่มีความสัมพันธ์ได้แก่ การจัดกิจกรรมที่สนับสนุนการเรียนรู้ในวิชาคณิตศาสตร์ภายในกับการให้การสนับสนุนของผู้ปกครอง ($r_{MATPROM,\ SCQ20AE} = 0.079$) การพัฒนาครูในสถานศึกษากับการจัดกิจกรรมที่สนับสนุนการเรียนรู้ในวิชาคณิตศาสตร์ภายใน ($r_{DVTESCH,\ MATPROM} = 0.117$) และประสบการณ์การทำงานของครูในสถานศึกษา (EXPTEA) ไม่มีความสำพันธ์กับตัวแปรใดตัวใดเลย รายละเอียด แสดงในตารางที่ 22

ตารางที่ 22 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรในระดับสถานศึกษา

	SCQ20AE	MATPROM	DVTESCH	EXPTEA
SCQ20AE	1.000			
MATPROM	0.079	1.000		
DVTESCH	.264**	.117	1.000	
EXPTEA	.016.	.110	-.045	1.00

**p < .01

การตรวจสอบภาวะร่วมเส้นตรงพหุ (multicollinearity) ผู้วิจัยใช้การวิเคราะห์ถดถอยพหุ (multiple regression) มีตัวแปรตามเป็นค่าความสามารถของนักเรียนจากการประเมินวิชาคณิตศาสตร์ในการทดสอบ TIMSS 2007 และตัวแปรทำนายเป็นตัวแปรต่อเนื่องในระดับสถานศึกษา ได้แก่ การให้การสนับสนุนของผู้ปกครอง (SCQ20AE) การจัดกิจกรรมที่สนับสนุนการเรียนรู้ในวิชาคณิตศาสตร์ภายใน (MATPROM) การพัฒนาครูในสถานศึกษา (DVTESCH) และประสบการณ์การทำงานของครูในสถานศึกษา (EXPTEA)) โดยพิจารณาจากค่า Tolerance และค่า VIF ซึ่ง Hair และคณะ (1998) เสนอว่าค่าสถิติเริ่มต้นของตัวแปรที่แสดงว่ามีภาวะร่วมเส้นตรงพหุ คือ มีค่า Tolerance ตั้งแต่ .10 ลงมา หรือ VIF มีค่า 10 ขึ้นไป

จากผลทดสอบ ภาวะร่วมเส้นตรงพหุ (Multicollinearity) ในระดับสถานศึกษา โดยพิจารณาค่า Tolerance และ Variance Inflation Factor: VIF) ของตัวแปรต้น พบว่า ผลการทดสอบความสัมพันธ์ระหว่างตัวแปรต้นแต่ละตัวมีค่า Tolerance สูง หรือมีค่าสูงกว่า .10 ตามข้อเสนอแนะของ Hair (Tolerance = 0.921 - 0.984) และค่า VIF ต่ำ หรือมีค่าต่ำกว่า 10 ตามข้อเสนอแนะของ Hair (VIF = 1.016 – 1.086) แสดงว่าตัวแปรต้นแต่ละตัวมีความสัมพันธ์กันต่ำ จึงไม่เกิดปัญหาที่ตัวแปรต้นมีความสัมพันธ์กันสูงหรือปัญหาภาวะร่วมเส้นตรงพหุ (multicollinearity) แสดงรายละเอียด ดังตารางที่ 23

ตารางที่ 23 ค่า Tolerance และ VIF จากผลการทดสอบความสัมพันธ์ระหว่างตัวแปรต้นระดับสถานศึกษา

ตัวแปร	Collinearity Statistics	
	Tolerance	VIF
SCQ20AE	0.928	1.078
MATPROM	0.984	1.017
DVTESCH	0.921	1.086
EXPTEA	0.984	1.016

ตอนที่ 2 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวงแต่ละฉบับ

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง ดำเนินการโดย ใช้โปรแกรม DDFS (Penfield, 2010) ซึ่งเป็นโปรแกรมที่ใช้การตรวจสอบการ ตรวจสอบการทำหน้าที่ต่า งกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวงด้วย วิธีการวิเคราะห์แมนเทล- แฮนส์เซลร่วมกับ ลอกของอัตราส่วน แต้มต่อ (odds ratio) เพื่อนำผลการวิเคราะห์ที่ได้มาใช้ในการพิจารณาเพื่อตัดข้อสอบตามลักษณะของโมเดลต่างๆ จากนั้นนำคะแนนผลการสอบที่ได้ไปคำนวณหาค่าความสามารถขอ งผู้สอบ โดยใช้โปรแกรม MULTILOG การนำเสนอผลการวิเคราะห์ในตอนนี้จะแบ่งเป็น 2 ตอน ดังนี้

2.1 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบและการทำหน้าที่ต่างกันของตัวลวง

2.2 สรุปข้อค้นพบเกี่ยวกับลักษณะของข้อสอบที่ทำ หน้าที่ต่างกันในแต่ละคุณลักษณะขอ งนักเรียน

2.1 สรุปผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง

การตรวจสอบการทำ หน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลว งเป็นการตรวจสอบการทำหน้าที่ของข้อสอบที่ทำให้ผู้สอบที่มีความสามารถ หรือคุณลักษณะที่มุ่งวัดเท่ากัน มีโอกาสในการตอบข้อสอบข้อเดียวกันได้แตกต่างกัน ผู้วิจัยได้พิจารณาตัวแปรเพื่อนำมาจัดกลุ่มคุณลักษณะของนักเรียน โดยกำหนดตัวแปรที่นำ มาใช้ตรวจสอบการทำ หน้าที่ต่างกันของข้อสอบ ดังนี้

FEMALE ตัวแปรดัมมี่ความเป็นเพศหญิงของนักเรียน กำหนดให้นักเรียนเพศหญิงเป็น 1 และนักเรียนเพศชายเป็น0

TIMEHOM การใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ กำหนดให้นักเรียนที่ใช้เวลาสูงกว่าค่าเฉลี่ยของนักเรียนทั้งหมด เป็น1 และนักเรียนที่ใช้เวลาต่ำกว่าค่าเฉลี่ยของนักเรียนทั้งหมด เป็น0

PAEDU ระดับการศึกษาของผู้ปกครอง กำหนดให้นักเรียนที่มี ระดับการศึกษาของผู้ปกครอง สูงกว่าค่าเฉลี่ยของนักเรียนทั้งหมด เป็น1 และนักเรียนที่มีระดับการศึกษาของผู้ปกครองต่ำกว่าค่าเฉลี่ยของนักเรียนทั้งหมด เป็น0

LEARES แหล่งทรัพยากรที่บ้าน กำ หนดให้นักเรียนที่มีแหล่งทรัพยากรที่บ้านสูงกว่าค่าเฉลี่ยของนักเรียนทั้งหมด มีค่าเป็น 1 และมีแหล่งทรัพยากรที่บ้านต่ำกว่าค่าเฉลี่ย มีค่าเป็น 0

WEALTH ความมั่งคั่งของครอบครัว กำหนดให้นักเรียนที่มีความมั่งคั่งของครอบครัวสูงกว่าค่าเฉลี่ยของนักเรียนทั้งหมด มีค่าเป็น 1 และมีความมั่งคั่งของครอบครัวต่ำกว่าค่าเฉลี่ย มีค่าเป็น 0

การนำเสนอส่วนนี้ ผู้วิจัยสรุปจำ นวนข้อสอบที่ทำ หน้าที่ต่างกันในแต่ละฉบับและเสนอรายละเอียดของผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวงแต่ละฉบับในภาคผนวก

ผลการตรวจสอบ ข้อสอบที่เกิดการทำหน้าที่ต่างกัน โดยการจำแนกตามลักษณะการตรวจพบ คือ 1. ข้อสอบที่พบการทำหน้าที่ต่างกันของข้อสอบ (DIF) เท่านั้น 2. ข้อสอบที่พบการทำหน้าที่ต่างกันของตัวลวง (DDF) เท่านั้น และ 3. ข้อสอบที่พบทั้งการทำหน้าที่ต่างกันของข้อสอบและการทำหน้าที่ต่างกันของตัวลวง (DIF-DDF) ในแบบสอบทั้ง 14 ฉบับ ที่เกิดจากคุณลักษณะของนักเรียนที่ต่า งกลุ่มกันในตัวแปรความเป็นเพศ หญิง การใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ ระดับการศึกษาของผู้ปกครองแหล่งทรัพยากรที่บ้าน และความมั่งคั่งของครอบครัว ภาพรวมพบว่าสามารถ ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันในทุกลักษณะ ในแบบสอบ แต่ละฉบับคิดเป็นร้อยละ 31.25 ถึง 64.29 และเมื่อจำแนกรายกรณีพบว่า แบบสอบแต่ละฉบับตรวจพบการทำหน้าที่ต่างกันของข้อสอบ (DIF) เท่านั้น คิดเป็นร้อยละ 6.25 ถึง 25 ในขณะที่แบบสอบแต่ละฉบับตรวจพบการทำหน้าที่ต่างกันของ ตัวลวง (DDF) เท่านั้น คิดเป็นร้อยละ 6.06 ถึง 22.58 และแบบสอบแต่ละฉบับการตรวจพบการทำหน้าที่ต่ างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) ในข้อเดียวกัน คิดเป็นร้อยละ 6.25 ถึง 34.48

โดยแบบสอบที่มีข้อสอบทำหน้าที่ต่างกันรวมทุกลักษณะมากที่สุดคือ แบบสอบฉบับที่ 12 คิดเป็นร้อยละ 64.29 (DIF 25%, DDF 17.86% และ DIF-DDF 21.43%) รองลงมาคือแบบสอบฉบับที่ 1 คิดเป็นร้อยละ 58.62 (DIF 13.79%, DDF 10.34% และ DIF-DDF 34.48%) และแบบสอบฉบับที่ 11 คิดเป็นร้อยละ 57.38 (DIF 21.88%, DDF 12.50% และ DIF-DDF 25%) ส่วนแบบสอบที่มีข้อสอบทำหน้าที่ต่างกันรวมทุกคุณลักษณะน้อยที่สุดคือ แบบสอบฉบับที่ 6 คิดเป็นร้อยละ 31.25 (DIF 6.25%, DDF 21.88% และ DIF-DDF 6.25%) รองลงมาคือแบบสอบฉบับ

ที่ 13 คิดเป็นร้อยละ35.71 (DIF 7.14%, DDF 14.29% และ DIF-DDF 14.29%) และแบบสอบฉบับที่ 7 คิดเป็นร้อยละ 36.36 (DIF 15.15%, DDF 6.06% และ DIF-DDF 15.15%)

โดยเมื่อจำแนกรายเนื้อหาใน แบบสอบ แต่ละฉบับพ บว่าเรื่องจำนวน เป็นเนื้อหาที่ พบข้อสอบทำหน้าที่ต่างกันรวมทุกคุณลักษณะมากที่สุด โดยพบมากที่สุดในแบบสอบ ถึง 9 ฉบับ คือ แบบสอบฉบับที่ 1, 2, 3, 4, 6, 7, 12, 13 และ 14 แสดงรายละเอียดดังตารางที่ 24

ตารางที่24 จำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูและตัวเลือกที่เป็นตัวลวง

ฉบับที่	เนื้อหา	จำนวนข้อสอบ	DIF	DDF	DIF-DDF	รวม
1.	1.1 จำนวน	11 (37.93%)	-	2 (6.90%)	4 (13.79%)	6 (20.69%)
	1.2 พีชคณิต	6 (20.69%)	1 (3.45%)	1 (3.45%)	1 (3.45%)	3 (10.34%)
	1.3 เรขาคณิต	8 (27.59%)	2 (6.90%)	-	3 (10.34%)	5 (17.24%)
	1.4 ข้อมูลและโอกาส	4 (13.79%)	1 (3.45%)	-	2 (6.90%)	3 (10.34%)
	รวม	**29 (100%)**	**4** (13.79%)	3 (10.34%)	**10** (34.48%)	**17** (58.62%)
2.	2.1 จำนวน	14 (45.16%)	2 (6.45%)	2 (6.45%)	3 (9.68%)	7 (22.58%)
	2.2 พีชคณิต	5 (16.13%)	2 (6.45%)	-	-	2 (6.45%)
	2.3 เรขาคณิต	7 (22.58%)	-	-	3 (9.68%)	3 (9.68%)
	2.4 ข้อมูลและโอกาส	5 (16.13%)	-	2 (6.45%)	-	2 (6.45%)
	รวม	**31 (100%)**	**4** (12.90%)	**4** (12.90%)	**6** (19.35%)	**14** (45.16%)
3.	3.1 จำนวน	15 (46.88%)	2 (6.25%)	3 (9.38%)	3 (9.38%)	8 (25%)
	3.2 พีชคณิต	5 (15.63%)	-	1 (3.13%)	1 (3.13%)	2 (6.25%)
	3.3 เรขาคณิต	7 (21.88%)	1 (3.13%)	1 (3.13%)	2 (6.25%)	4 (12.50%)
	3.4 ข้อมูลและโอกาส	5 (15.63%)	1 (3.13%)	1 (3.13%)	1 (3.13%)	3 (9.38%)
	รวม	**32 (100%)**	**4** (12.50%)	**6** (18.75%)	**7** (21.88%)	**17** (53.13%)
4.	4.1 จำนวน	8 (27.59%)	1 (3.45%)	1 (3.45%)	1 (3.45%)	3 (10.34%)
	4.2 พีชคณิต	8 (27.59%)	-	1 (3.45%)	2 (6.90%)	3 (10.34%)
	4.3 เรขาคณิต	6 (20.69%)	-	2 (6.90%)	1 (3.45%)	3 (10.34%)
	4.4 ข้อมูลและโอกาส	7 (24.14%)	1 (3.45%)	1 (3.45%)	-	2 (6.90%)
	รวม	**29** (100%)	2 (6.90%)	**5** (17.24%)	**4** (13.79%)	**11** (37.93%)
5	5.1 จำนวน	6 (20.69%)	1 (3.45%)	2 (6.90%)	1 (3.45%)	**4** (13.79%)
	5.2 พีชคณิต	9 (31.03%)	3 (10.34%)	1 (3.45%)	1 (3.45%)	5 (17.24%)
	5.3 เรขาคณิต	6 (20.69%)	1 (3.45%)	-	1 (3.45%)	2 (6.90%)
	5.4 ข้อมูลและโอกาส	8 (27.59%)	2 (6.90%)	-	1 (3.45%)	3 (10.34%)
	รวม	**29** (100%)	**7** (24.14%)	3 (10.34%)	**4** (13.79%)	**14** (48.28%)

ตารางที่ 24 จำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง (ต่อ)

ฉบับที่	เนื้อหา	จำนวนข้อสอบ	DIF	DDF	DIF-DDF	รวม
6.	6.1 จำนวน	8 (25%)	-	1 (3.13%)	2 (6.25%)	3 (9.38%)
	6.2 พีชคณิต	9 (28.13%)	1 (3.13%)	2 (6.25%)	-	3 (9.38%)
	6.3 เรขาคณิต	7 (21.88%)	-	3 (9.38%)	-	3 (9.38%)
	6.4 ข้อมูลและโอกาส	8 (25%)	1 (3.13%)	1 (3.13%)	-	2 (6.25%)
	รวม	**32 (100%)**	**2 (6.25%)**	**7 (21.88%)**	**2 (6.25%)**	**10 (31.25%)**
7.	7.1 จำนวน	9 (27.27%)	-	-	3 (9.09%)	3 (9.09%)
	7.2 พีชคณิต	11 (33.33%)	1 (3.03%)	1 (3.03%)	1 (3.03%)	3 (9.09%)
	7.3 เรขาคณิต	7 (21.21%)	1 (3.03%)	1 (3.03%)	1 (3.03%)	3 (9.09%)
	7.4 ข้อมูลและโอกาส	6 (18.18%)	3 (9.09%)	-	-	3 (9.09%)
	รวม	**33 (100%)**	**5 (15.15%)**	**2** (6.06%)	**5 (15.15%)**	**12 (36.36%)**
8.	8.1 จำนวน	8 (25%)	-	-	2 (6.25%)	2 (6.25%)
	8.2 พีชคณิต	11 (34.37%)	2 (6.25%)	1 (3.13%)	1 (3.13%)	4 (12.50%)
	8.3 เรขาคณิต	9 (28.13%)	1 (3.13%)	1 (3.13%)	**3** (9.38%)	5 (15.63%)
	8.4 ข้อมูลและโอกาส	4 (12.50%)	1 (3.13%)	1 (3.13%)	-	2 (6.25%)
	รวม	**32 (100%)**	**4** (12.50%)	**3** (9.38%)	**6** (18.75%)	**13** (40.63%)
9.	9.1 จำนวน	7 (22.58%)	-	1 (3.23%)	**3** (9.68%)	4 (12.90%)
	9.2 พีชคณิต	9 (29.03%)	-	1 (3.23%)	**3** (9.68%)	4 (12.90%)
	9.3 เรขาคณิต	9 (29.03%)	1 (3.23%)	**3** (9.68%)	1 (3.23%)	5 (16.13%)
	9.4 ข้อมูลและโอกาส	6 (19.35%)	2 (6.45%)	2 (6.45%)	-	4 (12.90%)
	รวม	**31 (100%)**	**3** (9.68%)	**7** (22.58%)	**7** (22.58%)	**17** (54.84%)
10.	10.1 จำนวน	5 (15.63%)	1 (3.13%)	-	1 (3.13%)	2 (6.25%)
	10.2 พีชคณิต	12 (37.50%)	2 (6.25%)	1 (3.13%)	3 (9.38%)	**6** (18.75%)
	10.3 เรขาคณิต	7 (21.88%)	-	1 (3.13%)	1 (3.13%)	2 (6.25%)
	10.4 ข้อมูลและโอกาส	8 (25%)	3 (9.38%)	1 (3.13%)	-	4 (12.50%)
	รวม	**32** (100%)	**6** (18.75%)	3 (9.38%)	5 (15.63%)	**14** (43.75%)
11.	11.1 จำนวน	**6** (18.75%)	3 (9.38%)	-	-	3 (9.38%)
	11.2 พีชคณิต	12 (37.50%)	1 (3.13%)	2 (6.25%)	4 (12.50%)	7 (21.88%)
	11.3 เรขาคณิต	7 (21.88%)	-	1 (3.13%)	2 (6.25%)	3 (9.38%)
	11.4 ข้อมูลและโอกาส	7 (21.88%)	3 (9.38%)	1 (3.13%)	2 (6.25%)	**6** (18.75%)
	รวม	**32 (100%)**	7 (21.88%)	4 (12.50%)	8 (25%)	**19 (57.38%)**

ตารางที่ 24 จำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง (ต่อ)

ฉบับที่	เนื้อหา	จำนวนข้อสอบ	DIF	DDF	DIF-DDF	รวม
12.	12.1 จำนวน	8 (28.57%)	3 (10.71%)	2 (7.14%)	3 (10.71%)	8 (28.57%)
	12.2 พีชคณิต	12 (42.86%)	3 (10.71%)	1 (3.57%)	1 (3.57%)	5 (17.86%)
	12.3 เรขาคณิต	4 (14.29%)	1 (3.57%)	-	-	1 (3.57%)
	12.4 ข้อมูลและโอกาส	4 (14.29%)	-	2 (7.14%)	2 (7.14%)	4 (14.29%)
	รวม	**28 (100%)**	**7** (25%)	**5** (17.86%)	**6** (21.43%)	**18** (64.29%)
13.	13.1 จำนวน	9 (32.14%)	-	1 (3.57%)	3 (10.71%)	4 (14.29%)
	13.2 พีชคณิต	12 (42.86%)	1 (3.57%)	2 (7.14%)	1 (3.57%)	4 (14.29%)
	13.3 เรขาคณิต	3 (10.71%)	-	1 (3.57%)	-	1 (3.57%)
	13.4 ข้อมูลและโอกาส	4 (14.28%)	1 (3.57%)	-	-	1 (3.57%)
	รวม	**28** (100%)	2 (7.14%)	4 (14.29%)	4 (14.29%)	**10** (35.71%)
14.	14.1 จำนวน	12 (40%)	2 (6.67%)	3 (10%)	1 (3.33%)	6 (20%)
	14.2 พีชคณิต	7 (23.33%)	-	1 (3.33%)	3 (10%)	4 (13.33%)
	14.3 เรขาคณิต	7 (23.33%)	-	1 (3.33%)	2 (6.67%)	3 (10%)
	14.4 ข้อมูลและโอกาส	4 (13.33%)	-	-	-	-
	รวม	**30** (100%)	**2** (6.67%)	**5** (16.67%)	**6** (20%)	**13** (43.33%)

จากผลการตรวจสอบข้อ สอบที่ทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) แต่ละฉบับ ที่พบว่าแบบสอบแต่ละฉบับมีข้อสอบที่ทำหน้าที่ต่างกันค่อนข้างมาก ผู้วิจัยจึงเลือกตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปเป็นบางข้อ โดยพิจารณาเลือกตัดข้อสอบตามเกณฑ์ 2 ข้อ ดังนี้ 1. เลือกข้อสอบที่ทำหน้าที่ต่างกันที่มีค่าสัมประสิทธิ์สูงสุด และ 2. คำนึงถึงโครงสร้างของแบบสอบ โดยภายหลังการตัดข้อสอบออกแล้วโครงสร้างของแบบสอบจะไม่แตกต่างจากโครงสร้างเดิม และข้อสอบที่ตัดออกไปควรมีจำนวนไม่เกินร้อยละ 20 ของข้อสอบในแบบสอบนั้น โดยพิจารณาตามสัดส่วนของเนื้อหาในแบบสอบแต่ละฉบับ ตามเกณฑ์การตัดข้อสอบที่ผู้วิจัยได้เสนอไว้ในบทที่ 3 และเมื่อตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปในแบบสอบแต่ละฉบับแล้ว พบว่า จำนวนข้อสอบที่ตัดออกไปในแบบสอบแต่ละฉบับมีไม่เกินร้อยละ 20 แสดงรายละเอียดดังตารางที่ 25

ตารางที่ 25 จำนวนข้อสอบที่ทำหน้าที่ต่างกันที่ตัดออกไปในแบบสอบทั้ง 14 ฉบับ

ฉบับ	ก่อนตัด	จำนวนข้อสอบ			
		หลังตัด DIF	ร้อยละของข้อสอบที่ตัด DIF	หลังตัด DIF+DDF	ร้อยละของข้อสอบที่ตัด DIF+DDF
1	29	26	10.34	24	17.24
2	31	28	9.68	27	12.90
3	32	28	12.50	28	12.50
4	29	27	6.90	25	13.79
5	29	23	20.69	25	13.79
6	32	30	6.25	30	6.25
7	33	30	9.09	29	12.12
8	32	28	12.50	27	15.63
9	31	28	9.68	26	16.13
10	32	27	15.63	28	12.50
11	32	29	9.38	27	15.63
12	28	23	17.86	24	14.29
13	28	26	7.14	24	14.29
14	30	28	6.67	26	13.33

การพิจารณาเพื่อตัดข้อสอบที่เกิดการทำหน้าที่ต่างกันของข้อสอบ(DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) นั้น สัดส่วนของเนื้อหาของข้อสอบแต่ละฉบับเป็นสิ่งหนึ่งที่ผู้วิจัยคำนึงถึง โดยแบบสอบวิชาคณิตศาสตร์ ในโครงการการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์ หรือ TIMSS 2007 ประกอบด้วยเนื้อหาวิชา 4 เรื่องสำคัญ คือจำนวน, พีชคณิต, เรขาคณิต และข้อมูลและโอกาส โดยเนื้อหาที่อยู่ในแบบสอบแต่ละฉบับก็จะมีสัดส่วนที่แตกต่างกันออกไป ดังนั้นผู้วิจัยจึงตัดข้อสอบตามเงื่อนไขในการวิจัย โดยให้จำนวนข้อสอบหลังการตัดข้อสอบตามเงื่อนไขแล้วนั้นมีสัดส่วนเท่าเดิมหรือใกล้เคียงกับสัดส่วนเนื้อหาเดิมมากที่สุด นอกจากนี้ผู้วิจัยวิเคราะห์ค่าความเที่ยงของแบบสอบตามเงื่อนไขของการตัดข้อสอบ และหาค่าพารามิเตอร์ความยากของข้อสอบในแบบสอบแต่ละฉบับโดยแสดงเป็นช่วงพารามิเตอร์ความยากจากน้อยที่สุดถึงมากที่สุด สำหรับการวิเคราะห์พารามิเตอร์ความยากนั้นเนื่องจากจำนวนนักเรียนที่สอบในแต่ละฉบับมีจำนวนไม่มาก ผู้วิจัยจึงวิเคราะห์โดยใช้โมเดลการวิเคราะห์แบบ 1 พารามิเตอร์ รายละเอียด ดังตารางที่ 26

ตารางที่ 26 การเปรียบเทียบโครงสร้างของร้อยละของแบบสอบก่อนตัด และหลังตัดข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวงในแบบสอบทั้ง14 ฉบับ

ฉบับที่	ร้อยละของข้อสอบก่อนการตัดDIF และ DDF				ร้อยละของข้อสอบหลังการตัด DIF				ร้อยละของข้อสอบหลังการตัด DIF+DDF				ความเที่ยง/ช่วงพารามิเตอร์ความยาก		
	Nu	Al	Ge	Da	Nu	Al	Ge	Da	Nu	Al	Ge	Da	ก่อนตัด	DIF	DIF+DDF
1.	38	21	28	13	42	19	27	12	38	21	29	12	0.8286 (-.72) – 2.5	0.8277 (-.66) – 2.88	0.8051 (-.68) – 2.13
2.	45	16	23	16	43	14	25	18	41	18	22	19	0.8462 (-.77) – 2.36	0.8342 (-.93) – 2.06	0.8239 (-.74) – 2.08
3.	46	16	22	16	46	18	21	15	45	16	23	16	0.8104 (-.94) – 1.66	0.8268 (-.85) – 1.48	0.8087 (-.82) – 1.81
4.	28	27	21	24	26	30	22	22	28	24	20	28	0.8095 (-.69) – 2.45	0.8109 (-.69) – 2.15	0.8131 (-.69) – 2.46
5.	21	31	21	27	22	30	22	26	20	32	20	28	0.8536 (-.26) – 1.95	0.8195 (-.69) – 2.09	0.8333 (-.69) – 1.82
6.	25	28	22	25	27	27	23	23	20	30	23	27	0.8476 (-.83) – 1.99	0.8503 (-.35) – 2.98	0.8555 (-.33) – 2.75
7.	27	33	22	18	30	33	20	17	24	34	21	21	0.8779 (-.45) – 2.04	0.8711 (-.43) – 2.23	0.8641 (-.43) – 1.76
8.	25	34	28	13	28	32	29	11	22	37	26	15	0.8768 (-.69) – 1.60	0.8605 (-.69) – 1.61	0.8574 (-.69) – 1.59
9.	23	29	29	19	25	32	29	14	23	27	31	19	0.8527 (-.68) – 2.23	0.8493 (-.68) – 1.28	0.8399 (-.64) – 2.11
10.	15	38	22	25	15	37	26	22	14	36	21	29	0.8489 (-.53) – 1.85	0.8045 (-.56) – 1.99	0.8282 (-.69) – 1.85
11.	19	38	22	21	17	38	24	21	22	34	22	22	0.8267 (-.69) – 2.28	0.8158 (-.58) – 2.26	0.8181 (-.55) – 2.23
12.	29	43	14	14	27	43	13	17	25	45	17	13	0.8465 (-.69) – 2.07	0.8061 (-.69) – 2.11	0.8361 (-.69) – 2.08
13.	32	43	11	14	31	42	12	15	29	42	13	16	0.8348 (-.69) – 2.36	0.8190 (-.69) – 2.22	0.8129 (-.69) – 2.27
14.	41	23	23	13	36	25	25	14	39	23	23	15	0.8661 (-.25) – 2.13	0.8533 (-.25) – 2.18	0.8421 (-.08) – 2.16

2.2 ผลการวิเคราะห์ลักษณะของข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวงในแต่ละคุณลักษณะของนักเรียน

จากผลการวิเคราะห์การทำ หน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) ในแต่ละฉบับ ผู้วิจัยได้พิจารณาจำ นวนข้อสอบที่ทำ หน้าที่ต่างกันในแต่ละเนื้อหา โดย ลักษณะของข้อสอบ 4 เรื่องคือ 1. จำนวน 2. พีชคณิต 3. เรขาคณิต และ 4. ข้อมูลและโอกาส ผลการพิจารณาพบว่าข้อสอบที่ทำ หน้าที่ต่างกันเนื่องจากตัวแปรภายนอก มีลักษณะของข้อสอบ ดังนี้

จากผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) โดยภาพรวมพบว่าข้อสอบที่ทำหน้าที่ต่างกันจะพบในตัวแปรเพศมากที่สุด (DIF 57 ข้อ, DDF 65 ข้อ) รองลงมาคือการศึกษาของผู้ปกครอง(DIF 37 ข้อ, DDF 39 ข้อ) และความมั่งคั่งของครอบครัว (DIF 31 ข้อ, DDF 36 ข้อ) ส่วนที่พบข้อสอบที่ทำหน้าที่ต่างกันน้อยที่สุดคือในตัวแปรแหล่งทรัพยากรที่บ้าน (DIF 26 ข้อ, DDF 24 ข้อ) รองลงมาคือการใช้เวลาในการทำการบ้าน (DIF 28 ข้อ, DDF 35 ข้อ) โดยผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง จำแนกตามตัวแปรแต่ละตัว นอกจากนี้ยังพบว่าการเกิดการทำหน้าที่ต่างกันของข้อสอบ (DIF) ส่วนใหญ่กลุ่มอ้างอิง (Reference group) ซึ่งเป็นกลุ่มที่คาดว่าน่าจะได้เปรียบในการสอบครั้งนี้ มีโอกาสในการที่จะตอบถูกมากกว่ากลุ่มเปรียบเทียบ (Focal group) โดยมีเพียงตัวแปรความมั่งคั่งตัวเดียวเท่านั้นที่พบว่าจำนวนข้อสอบที่เกิดการทำหน้าที่ต่างกัน (DIF) มีจำนวนข้อเท่ากันทั้ง 2 กลุ่ม ส่วนการทำหน้าที่ต่างกันของตัวลวง (DDF) จากผลการวิเคราะห์ข้อมูลพบว่ามีความสอดคล้องกับการทำหน้าที่ต่างกันของข้อสอบ นั่นคือส่วนใหญ่กลุ่ม เปรียบเทียบ (Focal group) มีโอกาสที่จะเลือกตอบตัวเลือกที่เป็นตัวลวงมากกว่ากลุ่มอ้างอิง (Reference group) โดยมีเพียงตัวแปรแหล่งทรัพยากรที่บ้านเท่านั้นที่มีโอกาสเท่ากัน โดยจำแนกรายละเอียดตัวแปรแต่ละตัวได้ ดังนี้

ตัวแปรเพศ จากการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) พบว่านักเรียนหญิงมีโอกาสทำข้อสอบได้ถูกมากกว่านักเรียน ชาย และเมื่อพิจารณาลักษณะข องเนื้อหาสำหรับข้อสอบเรื่องพีชคณิต นักเรียนหญิงมีโอกาสทำข้อสอบได้มากกว่า ในขณะที่เรื่องเรขาคณิตนักเรียนชายมีโอกาสตอบถูกมากกว่า ส่วนข้อสอบที่เกิดการทำหน้าที่ต่างกันของตัวลวง (DDF) นั้นพบว่าเพศชายมีโอกาสในการที่จะเลือกตัวลวงได้มากกว่าเพศหญิงเล็กน้อย และเมื่อพิจารณาลักษณะของ

เนื้อหาพบว่าเป็นข้อสอบ เรื่องจำนวน , พีชคณิต , และข้อมูลและโอกาส เพศชายมีโอกาสที่จะถูกลวงได้มากกว่าเพศหญิง ในขณะที่เรื่องเรขาคณิตเพศหญิงมีโอกาสที่จะถูกลวงได้มากกว่าเพศชาย

ตัวแปรการใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ พบว่าการทำหน้าที่ต่างกันของข้อสอบ (DIF) นั้นนักเรียนที่มีใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ สูงมีโอกาสทำข้อสอบได้ถูกมากกว่านักเรียนที่ใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ ต่ำเป็นส่วนใหญ่ โดยเฉพาะเมื่อพิจารณาลักษณะของ เนื้อหาที่พบอย่างชัดเจนคือ ข้อสอบเรื่องจำนวนและพีชคณิต นักเรียนที่ ใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ สูงมีโอกาสทำข้อสอบได้มากกว่า ส่วนข้อสอบที่เกิดการทำหน้าที่ต่างกันของตัวลวง (DDF) นั้นพบว่า นักเรียนที่ ใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ ต่ำ มีโอกาสในการที่จะเลือกตัวลวงได้มากกว่า นักเรียนที่ใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ สูง และเมื่อพิจารณาลักษณะของ เนื้อหาพบว่าข้อสอบเรื่องจำนวน และพีชคณิต นักเรียนที่ใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ ต่ำ มีโอกาสที่จะถูกลวงได้มากกว่า อีกกลุ่ม

ตัวแปรระดับการศึกษาของผู้ปกครองจากการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ(DIF) พบว่านักเรียนที่มีระดับการศึกษาของผู้ปกครองสูงมีโอกาสทำข้อสอบได้ถูกมากกว่านักเรียนที่มีระดับการศึกษาของผู้ปกครองต่ำเล็กน้อย และเมื่อพิจารณาลักษณะของเนื้อหา พบว่าข้อสอบเรื่องพีชคณิตและเรขาคณิตนักเรียนที่มีระดับการศึกษาของผู้ปกครองสูงมีโอกาสตอบถูกมากกว่า ในขณะที่เรื่องข้อมูลและโอกาสนักเรียนที่มีระดับการศึกษาของผู้ปกครองต่ำมีโอกาสตอบถูกมากกว่า ส่วนข้อสอบที่เกิดการทำหน้าที่ต่างกันของตัวลวง (DDF) นั้นพบว่านักเรียนที่มีระดับการศึกษาของผู้ปกครองต่ำ มีโอกาสในการที่จะเลือกตัวลวงได้มากกว่า นักเรียนที่มีระดับการศึกษาของผู้ปกครอง สูง และเมื่อพิจารณาลักษณะของเนื้อหาพบว่าข้อสอบเรื่องจำนวน, เรขาคณิต และข้อมูลและโอกาสนักเรียนที่มีระดับการศึกษาของผู้ปกครองต่ำ มีโอกาสที่จะถูกลวงได้มากกว่าอีกกลุ่มหนึ่ง ในขณะที่เรื่องพีชคณิตนักเรียนที่มีระดับการศึกษาของผู้ปกครองสูงมีโอกาสที่จะถูกลวงได้มากกว่ากลุ่มตรงข้าม

ตัวแปรแหล่งทรัพยากรที่บ้าน จากการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) พบว่านักเรียนที่มีแหล่งทรัพยากรที่บ้านที่บ้านสูงมีโอกาสในการทำข้อสอบถูกต้องมากกว่านักเรียนที่มีแหล่งทรัพยากรที่บ้านต่ำ และ เมื่อพิจารณาลักษณะของ เนื้อหาพบว่าเรื่องพีชคณิต กับข้อมูลและโอกาสนักเรียนที่มี แหล่งทรัพยากรที่บ้าน ที่บ้านสูงมีโอกาสในการทำข้อสอบถูกต้องมากกว่านักเรียนที่มีแหล่งทรัพยากรที่บ้านต่ำ ส่วนเรื่องจำนวน และเรขาคณิตมีโอกาสในการตอบถูกเท่ากัน

ส่วนข้อสอบที่เกิดการทำหน้าที่ต่างกันของตัวลวง (DDF) นั้นพบว่าทั้งกลุ่ม นักเรียนที่มีแหล่งทรัพยากรที่บ้านที่บ้านสูง และกลุ่มของ นักเรียนที่มีแหล่งทรัพยากรที่บ้านต่ำ มีโอกาสในการที่จะเลือกตอบตัวเลือกที่เป็นตัวลวงได้เท่ากัน และเมื่อ พิจารณาลักษณะของ เนื้อหา พบว่า เนื้อหาที่มีความแตกต่างคือพีชคณิตและเรขาคณิตโดยนักเรียนที่มีแหล่งทรัพยากรที่ บ้านสูงมีโอกาสในการที่จะเลือกตอบตัวเลือกที่เป็นตัวลวงได้มากกว่ากลุ่มของ นักเรียนที่มีแหล่งทรัพยากรที่บ้าน ต่ำ ส่วนเรื่องข้อมูลและโอกาส นักเรียนที่มีแหล่งทรัพยากรที่บ้านที่บ้าน ต่ำมีโอกาสในการที่จะเลือกตอบตัวเลือกที่เป็นตัวลวงได้มากกว่ากลุ่มของนักเรียนที่มีแหล่งทรัพยากรที่บ้านสูง

ตัวแปรความมั่งคั่งของครอบครัว จากการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) พบว่านักเรียนที่มีความมั่งคั่งของครอบครัวสูงมีโอกาสในการทำข้อสอบถูกต้องเท่ากับนักเรียนที่มีความมั่งคั่งของครอบครัวต่ำ และเมื่อพิจารณาลักษณะของเนื้อหาพบว่าเรื่องจำนวน กับข้อมูลและโอกาส นักเรียนที่มีความมั่งคั่งของครอบครัวสูงมีโอกาสในการทำข้อสอบถูกต้องมากกว่านักเรียนที่มีความมั่งคั่งของครอบครัวต่ำ ส่วนเรื่องพีชคณิตและเรขาคณิต นักเรียนที่มีความมั่งคั่งของครอบครัวต่ำมีโอกาสในการทำข้อสอบถูกต้องมากกว่านักเรียนที่มีความมั่งคั่งของครอบครัวสูง ส่วนข้อสอบที่เกิดการทำหน้าที่ต่างกันของตัวลวง(DDF) นั้นพบว่าทั้งกลุ่มนักเรียนที่มีแหล่งทรัพยากรที่บ้านที่บ้านสูงและกลุ่มของนักเรียนที่มีความมั่งคั่งของครอบครัวต่ำมีโอกาสในการที่จะเลือกตอบตัวเลือกที่เป็นตัวลวงมากกว่ากลุ่มของนักเรียนที่มีความมั่งคั่งของครอบครัวสูง และเมื่อพิจารณาลักษณะของเนื้อหาพบว่า เนื้อหาที่มีความแตกต่างคือจำนวน และพีชคณิตโดยนักเรียนที่มีความมั่งคั่งของครอบครัวต่ำมีโอกาสในการที่จะเลือกตอบตัวเลือกที่เป็นตัวลวงได้มากกว่ากลุ่มของนักเรียนที่มีความมั่งคั่งของครอบครัวสูง ส่วนเรื่องเรขาคณิตนักเรียนที่มีความมั่งคั่งของครอบครัวสูงมีโอกาสในการที่จะเลือกตอบตัวเลือกที่เป็นตัวลวงได้มากกว่ากลุ่มของนักเรียนที่มีแหล่งทรัพยากรที่บ้านต่ำ รายละเอียดดังตารางที่27

ตารางที่ 27 สรุปลักษณะของข้อสอบที่ทำหน้าที่ต่างกันในแต่ละคุณลักษณะของนักเรียน

คุณลักษณะ	เนื้อหาคณิตศาสตร์								รวม	
	จำนวนและตัวเลข		พีชคณิต		เรขาคณิต		ข้อมูลและโอกาส			
	DIF	DDF	DIF	DDF	DIF	DDF	DIF	DDF	DIF	DDF
หญิง	9 (30%)	8 (24%)	13 (43%)	7 (21%)	4 (13%)	14 (42%)	4 (13%)	4 (12%)	30 (100%)	33 (100%)
ชาย	10 (37%)	11 (31%)	3 (11%)	12 (34%)	10 (37%)	6 (17%)	4 (15%)	6 (17%)	27 (100%)	35 (100%)
การใช้เวลาในการทำการบ้านสูง	7 (47%)	1 (8%)	5 (33%)	4 (31%)	1 (7%)	5 (38%)	2 (13%)	3 (23%)	15 (100%)	13 (100%)
การใช้เวลาในการทำการบ้านต่ำ	3 (27%)	4 (29%)	2 (18%)	6 (43%)	3 (27%)	2 (14%)	3 (27%)	2 (14%)	11 (100%)	14 (100%)
การศึกษาของผู้ปกครองสูง	6 (30%)	7 (41%)	7 (35%)	6 (35%)	4 (20%)	1 (6%)	3 (15%)	3 (18%)	20 (100%)	17 (100%)
การศึกษาของผู้ปกครองต่ำ	6 (32%)	9 (45%)	6 (32%)	3 (15%)	2 (11%)	4 (20%)	5 (26%)	4 (20%)	19 (100%)	20 (100%)
แหล่งทรัพยากรสูง	3 (30%)	5 (31%)	4 (40%)	6 (38%)	1 (10%)	4 (25%)	2 (20%)	1 (6%)	10 (100%)	16 (100%)
แหล่งทรัพยากรต่ำ	3 (38%)	5 (31%)	3 (38%)	5 (31%)	1 (13%)	3 (19%)	1 (13%)	3 (19%)	8 (100%)	16 (100%)
ความมั่งคั่งสูง	3 (19%)	4 (27%)	5 (31%)	4 (27%)	4 (25%)	6 (40%)	4 (25%)	1 (7%)	16 (100%)	15 (100%)
ความมั่งคั่งต่ำ	2 (13%)	8 (40%)	6 (38%)	7 (35%)	6 (38%)	4 (20%)	2 (13%)	1 (5%)	16 (100%)	20 (100%)

ตอนที่ 3 ผลของข้อสอบที่ทำหน้าที่ต่างกัน และผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา ต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน 4 โมเดล

การประเมินคุณภาพการจัดการศึกษาในงานวิจัยเรื่องนี้ เงื่อนไขที่สำคัญของแต่ละโมเดลประกอบด้วย 2 ส่วนด้วยกัน นั่นคือ ส่วนแรกเป็นลักษณะของข้อสอบที่มีการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวงเกิดขึ้น ส่วนที่สองเป็นคุณลักษณะของตัวแปรในระดับนักเรียนและสถานศึกษา ดังนั้นในตอนที่ 3 นี้ประกอบไปด้วยประเด็นที่สำคัญ 3 ประเด็นด้วยกัน คือ 1. โมเดลที่ใช้ในการประเมินคุณภาพการจัดการศึกษา 4 โมเดล โดยทั้ง 4 โมเดลมีความแตกต่างกันตามเงื่อนไขที่ต่างกันออกไป 2. ผลการประมาณค่าความสามารถของผู้สอบในแต่ละโมเดล

3. ผลของข้อสอบที่ทำหน้าที่ต่างกันต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน และ 4. ผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน โดยมีรายละเอียดที่สำคัญ ดังนี้

3.1 ลักษณะโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์โมเดล

การวิเคราะห์โมเดลการประเมิน คุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ของสถานศึกษาขั้นพื้นฐาน ประกอบด้วย 4 โมเดล ได้แก่ โมเดล 1 “Undetected DIF-DDF & Adjusted” เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ ไม่ได้ ตัดข้อสอบที่ทำหน้าที่ต่างกันออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการปรับแก้ (Adjusted) โมเดล 2 “Detect DIF” เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ตัวถูกออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการปรับแก้ (Adjusted) โมเดลที่ 3 “Detect DIF-DDF & Unadjusted” เป็นโมเดลเป็นการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ทั้งตัวถูกและตัวลวงออก แต่ไม่ได้ทำการวิเคราะห์มูลค่าเพิ่ม (Unadjusted) โมเดล 4 “Detect DIF-DDF & Adjusted” เป็นโมเดลเป็นการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ตัดข้อส อบที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ทั้งตัวถูกและตัวลวงออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการปรับแก้ (Adjusted)

การนำเสนอในตอนนี้ จะแสดงให้เห็นถึงสมการวิเคราะห์ในแต่ละโมเดล และอิทธิพลของตัวแปรระดับนักเรียนและระดับสถานศึกษาที่มีผลต่อการประเมินคุณภาพการจัดการศึ กษาวิชาคณิตศาสตร์

1. การวิเคราะห์ตามโมเดลที่ 1 (Undetected DIF-DDF & Adjusted)

โมเดล 1 “Undetected DIF-DDF & Adjusted” เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาคณิตศาสตร์ที่ ไม่ได้ ตัดข้อสอบที่ทำหน้าที่ต่างกัน ออก ก่อนการนำค่าความสามารถ มาใช้ในการ วิเคราะห์มู ลค่าเพิ่ม และมีการปรับแก้ (Adjusted) ค่าความสามารถ

ด้วยตัวแปรระดับนักเรียนและสถานศึกษา เพื่อหาคะแนนเศษเหลือ (residual) ของสถานศึกษา หรือมูลค่าเพิ่มของสถานศึกษา มีสมการวิเคราะห์ดังนี้

ระดับที่ 1 ระดับนักเรียน (ระหว่างนักเรียน i ภายใต้สถานศึกษา k)

$$THETA1_{ik} = B0_k + (.016)_k*(RELAGR1)_{ik} + (.042)*(SEX)_{ik} + (.006)*(PAEDU1)_{ik} + (.044)*(LEARES)_{ik} + (.072)*(WEALTH)_{ik} + (.075)*(PRASCH)_{ik} + (.034)*(TIMEHOM)_{ik} + (-.040)*(INTMAT)_{ik} + (.175)*(SELFLEA)_{ik} + (.072)*(JOYMATH)_{ik} + (.010)*(INSTMAT)_{ik} + (.032)*(VALMATH)_{ik} + (.009)*(SAFESCH)_{ik} + R_{ik}$$

ระดับที่ 2 ระดับหน่วย (ระหว่างสถานศึกษา)

$$B0_k = -1.90 + (.746)*(DSCHSL) + (.285)*(DSCHL) + (.035)*(DSCHM) + (.178)*(COMMXXL) + (-.021)*(COMMXL) + (.322)*(COMML) + (.046)*(COMMM) + (.093)*(COMMS) + (.003)*(SCQ21AC) + (-.053)*(SCQ20AE) + (.002)*(SCMAEDU) + (-.034)*(MATPROM) + (.034)*(PPSC) + (.031)*(DVTESCH) + (-.005)*(EXPTEA) + (-.027)*(ADEQ) + (.106)*(HOMEWORK) + (-.006)*(CLASSSIZ) + (.030)*(TEACOLLA) + U0_k$$

2. การวิเคราะห์ตามโมเดลที่ 2 (Detect DIF & Adjusted)

โมเดล 4 “Detect DIF & Adjusted” เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันเฉพาะที่เป็นตัวเลือกที่เป็นตัวถูกออกก่อนการนำค่าความสามารถมาใช้ในการ วิเคราะห์มูลค่าเพิ่ม และมีการปรับแก้ (Adjusted) ค่าความสามารถด้วยตัวแปรระดับนักเรียนและสถานศึกษา เพื่อหาคะแ นนเศษเหลือ (residual) ของสถานศึกษา หรือมูลค่าเพิ่มของสถานศึกษา มีสมการวิเคราะห์ดังนี้

ระดับที่ 1 ระดับนักเรียน (ระหว่างนักเรียน i ภายใต้สถานศึกษา k)

$$THETA2_{ik} = B0_k + (.005)_k*(RELAGR1)_{ik} + (.018)*(SEX)_{ik} + (.006)*(PAEDU1)_{ik} + (.010)*(LEARES)_{ik} + (.058)*(WEALTH)_{ik} + (.023)*(PRASCH)_{ik} + (.013)*(TIMEHOM)_{ik} + (.013)*(INTMAT)_{ik} + (.015)*(SELFLEA)_{ik} + (.014)*(JOYMATH)_{ik} + (.018)*(INSTMAT)_{ik} + (.019)*(VALMATH)_{ik} + (.012)*(SAFESCH)_{ik} + R_{ik}$$

ระดับที่ 2 ระดับหน่วย (ระหว่างสถานศึกษา)

$$B0_k = -1.784 + (.716)*(DSCHSL) + (.282)*(DSCHL) + (.043)*(DSCHM) + (.165)*(COMMXXL) + (-.026)*(COMMXL) + (.309)*(COMML) + (.029)*(COMMM) + (.077)*(COMMS) + (.003)*(SCQ21AC) + (-.048)*(SCQ20AE) + (.007)*(SCMAEDU) + (-.034)*(MATPROM) + (.033)*(PPSC) + (.030)*(DVTESCH) + (-.005)*(EXPTEA) + (-.028)*(ADEQ) + (.106)*(HOMEWORK) + (-.008)*(CLASSSIZ) + (.027)*(TEACOLLA) + U0_k$$

3. การวิเคราะห์ตามโมเดลที่ 3 (Detect DIF-DDF & Unadjusted)

โมเดลที่ 3 “Detect DIF-DDF & Unadjusted” เป็นโมเดลเป็นการประเมินคุณภาพการจัดการศึกษาในวิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกัน ทั้งตัวเลือกที่เป็นตัวถูก และตัวเลือกที่เป็นตัวลวงออก ก่อนการนำค่าควา มสามารถมาใช้ในการวิเคราะห์มูลค่าเพิ่ม โดยไม่มีการปรับแก้ (Unadjusted) คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา นั่นคือค่าความสามารถจากผลการประเมินในวิชาคณิตศาสตร์ที่นำมาใช้เป็นตัวแปรตามในโมเดลนี้ เป็นค่าที่ได้จากการตัดข้อสอบที่เกิดการหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวงก่อนนำค่าความสามารถที่ได้มาวิเคราะห์พหุระดับเพื่อหาคะแนนเศษเหลือ(residual) ของสถานศึกษา แต่ไม่มีการนำตัวแปรต้นในระดับนักเรียนและระดับสถานศึกษาเข้ามาวิเคราะห์ในโมเดลในการวิเคราะห์ดังนี้

ระดับที่ 1 ภายในสถานศึกษา

$$THETA3_{ik} = B0_k + R_{ik}$$

ระดับที่ 2 ระหว่างสถานศึกษา

$$B0_k = -.694 + U0_k$$

4. การวิเคราะห์ตามโมเดลที่ 4 (Detect DIF-DDF & Adjusted)

โมเดล 4 “Detect DIF-DDF & Adjusted” เป็นโมเดลเป็นการประเมินคุณภาพการจัดการศึกษาในวิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกัน ทั้งตัวเลือกที่เป็นตัวถูก และตัวเลือกที่เป็นตัวลวงออก ก่อนการนำค่าความสามารถมาใช้ในการวิเคราะห์มูลค่าเพิ่ม โดยมีการปรับแก้ (adjusted) คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา นั่นคือค่าความสามารถจากผลการประเมินในวิชาคณิตศาสตร์ที่นำมาใช้เป็นตัวแปรตามในโมเดลที่4 เป็นค่าที่ได้จากการตัดข้อสอบที่เกิดการหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวงก่อนนำค่าความสามารถที่ได้มาวิเคราะห์พหุระดับเพื่อหาคะแนนเศษเหลือ(residual) ของสถานศึกษา พร้อมทั้งมีการนำตัวแปรต้นในระดับนักเรียนและระดับสถานศึกษาเข้ามาวิเคราะห์โดยมีสมการในการวิเคราะห์ดังนี้

ระดับที่ 1 ระดับนักเรียน (ระหว่างนักเรียน i ภายใต้สถานศึกษา k)

$$THETA4_{ik} = B0_k + (.017)_k*(RELAGR1)_{ik} + (.049)*(SEX)_{ik} + (.008)*(PAEDU1)_{ik} + (.048)*(LEARES)_{ik} + (.069)*(WEALTH)_{ik} + (.069)*(PRASCH)_{ik} + (.030)*(TIMEHOM)_{ik} + (-.044)*(INTMAT)_{ik} + (.164)*(SELFLEA)_{ik} + (.067)*(JOYMATH)_{ik} + (.011)*(INSTMAT)_{ik} + (.046)*(VALMATH)_{ik} + (.011)*(SAFESCH)_{ik} + R_{ik}$$

ระดับที่ 2 ระดับหน่วย (ระหว่างสถานศึกษา)

$$B0_k = -1.936 + (.766)*(DSCHSL) + (.310)*(DSCHL) + (.047)*(DSCHM) + (.160)*(COMMXXL) + (-.038)*(COMMXL) + (.312)*(COMML) + (.037)*(COMMM) + (.102)*(COMMS) + (.0003)*(SCQ21AC) + (-.070)*(SCQ20AE) + (.017)*(SCMAEDU) + (-.042)*(MATPROM) + (.029)*(PPSC) + (.034)*(DVTESCH) + (-.005)*(EXPTEA) + (-.027)*(ADEQ) + (.107)*(HOMEWORK) + (-.027)*(CLASSSIZ) + (.043)*(TEACOLLA) + U0_k$$

โดยที่ $THETA_{ik}$ = ค่าความสามารถของผลการประเมินความสามารถในวิชาคณิตศาสตร์ของนักเรียนคนที่i ภายใต้สถานศึกษาที่k

$B0_k$ = ค่าจุดตัดแกน (intercept) ที่เป็นค่าแสดงถึงคะแนนเฉลี่ยของค่าความสามารถของผลการประเมินความสามารถในวิชาคณิตศาสตร์ในโครงการ TIMSS 2007 ในสถานศึกษา k

R_{ik} = ค่าเศษเหลือหรือค่าความคลาดเคลื่อน (residual error)

G00 = ค่าจุดตัดแกน (intercept) ที่เป็นค่าแสดงถึงคะแนนเฉลี่ยของค่าความสามารถของผลการประเมินความสามารถในวิชาคณิตศาสตร์ในโครงการ TIMSS 2007 ในสถานศึกษาทุกโรงเรียน

G01 = ค่าความชัน (slope) ที่แสดงถึงค่าอิทธิพลของการเปลี่ยนแปลงของคะแนนความสามารถในวิชาคณิตศาสตร์ในโครงการ TIMSS 2007 เมื่อมีการควบคุมตัวแปร Y_k

$U0_k$ = ค่าเศษเหลือ ซึ่งเป็นค่าที่แสดงถึงมูลค่าเพิ่มของสถานศึกษา

ผู้วิจัยเสนอสรุปผลการวิเคราะห์ในแต่ละโมเดล ส่วนรายละเอียดผู้วิจัยได้เสนอไว้ในภาคผนวก สรุปผลจากโมเดลที่ 1 ถึงโมเดลที่ 4 ดังนี้

ผลการวิเคราะห์ในโมเดล 1 (Undetected DIF-DDF & Adjusted) จากผลการวิเคราะห์พหุระดับในโมเดล 1 ผลการทดสอบอิทธิพลคงที่ (Fixed effect) พบว่า ค่าเฉลี่ยของความสามารถของนักเรียนจากการประเมินความสามารถทางคณิตศาสตร์ทุกสถานศึกษา (G00) มีความผันแปรจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (G00 = -1.900923, t = -7.204, p = 0.000) ค่าสัมประสิทธิ์ถดถอยของตัวแปรระดับสถานศึกษาที่มีค่าเป็นบวกสูงที่สุด คือ สถานศึกษาที่มีขนาดใหญ่พิเศษ โดยเป็นสถานศึกษาที่มีนักเรียนตั้งแต่ 2500 คนขึ้นไป (β = 0.746488) นั่นคือความเป็นสถานศึกษาขนาดใหญ่พิเศษจะทำให้ค่าความสามารถของนักเรียนจากการประเมินความสามารถทางคณิตศาสตร์มีค่าเพิ่มขึ้น 0.746488 รองลงมาคือ สถานศึกษาที่ตั้งอยู่ในเมือง (β = 0.322065) สถานศึกษาขนาดใหญ่ (β = 0.284704) ส่วนค่าสัมประสิทธิ์ถดถอยของตัวแปรระดับสถานศึกษาที่มีค่าเป็นลบสูงที่สุดคือ การให้การสนับสนุนของผู้ปกครอง (β = -0.052732) รองลงมาคือการจัดกิจกรรมเสริมทางคณิตศาสตร์ภายในสถานศึกษา (β = -0.034328) และ ดัชนีความพึงพอใจของครูที่มีต่อการจัดการภายในสถานศึกษา (β = -0.026689) ส่วนตัวแปรระดับนักเรียนที่มีอิทธิพลทางบวกสูงที่สุดคือความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียน

คณิตศาสตร์ (β = 0.174597) รองลงมาคือการฝึกปฏิบัติทักษะทางด้านคณิตศาสตร์ในชั่วโมงเรียน (β = 0.074926) และความมั่งคั่งของครอบครัว (β = 0.072858) ส่วนตัวแปรระดับนักเรียนที่มีอิทธิพลน้อยสูงที่สุดคือการใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเอง (β = -0.040604) รองลงมาคือ ความรู้สึกเกี่ยวความปลอดภัยในการมาเรียนภายในสถานศึกษา (β = 0.008864)

ผลการทดสอบอิทธิพลเชิงสุ่ม (Random effect) พบว่า ค่าเศษเหลือระดับสถานศึกษาของค่าความสามารถของนักเรียนจากการประเมินความสามารถทางคณิตศาสตร์ที่มีการควบคุมตัวแปรระดับนักเรียนและสถานศึกษา (U0) หรือมูลค่าเพิ่มของสถานศึกษามีความผันแปรระหว่างสถานศึกษาอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (χ^2 = 2791.16706, p = 0.000) โดยมีความแปรปรวนระหว่างสถานศึกษาเท่ากับ 0.17316 ซึ่งอธิบายความแปรปรวนได้ประมาณร้อยละ 34.05 และความแปรปรวนภายในสถานศึกษามีค่าเท่ากับ 0.33539 ซึ่งอธิบายความแปรปรวนได้ประมาณร้อยละ 65.95

ผลการวิเคราะห์ในโมเดล 2 (Detect DIF & Adjusted) จากผลการวิเคราะห์พหุระดับในโมเดล 2 ผลการทดสอบอิทธิพลคงที่ (Fixed effect) พบว่า ค่าเฉลี่ยของความสามารถของนักเรียนจากการประเมินความสามารถทางคณิตศาสตร์ทุกสถานศึกษา (G00) มีความผันแปรจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (G00 = -1.784011, t = -5.281, p = 0.000) ค่าสัมประสิทธิ์ถดถอยของตัวแปรระดับสถานศึกษาที่มีค่าเป็นบวกสูงที่สุดคือ สถานศึกษาที่มีขนาดใหญ่พิเศษ โดยเป็นสถานศึกษาที่มีนักเรียนตั้งแต่ 2500 คนขึ้นไป (β = 0.715612) นั่นคือความเป็นสถานศึกษาขนาดใหญ่พิเศษจะทำให้ค่าความสามารถของนักเรียนจากการประเมินความสามารถทางคณิตศาสตร์มีค่าเพิ่มขึ้น 0.715612 รองลงมาคือ สถานศึกษาที่ตั้งอยู่ในเมือง (β = 0.309145) สถานศึกษาขนาดใหญ่ (β = 0.165491) ส่วนค่าสัมประสิทธิ์ถดถอยของตัวแปรระดับสถานศึกษาที่มีค่าเป็นลบสูงที่สุดคือ การให้การสนับสนุนของผู้ปกครอง สถานศึกษาที่ตั้งอยู่ในเมืองใหญ่ (β = -0.047561) รองลงมาคือ การจัดกิจกรรมเสริมทางคณิตศาสตร์ภายในสถานศึกษา (β = -0.034449) และดัชนีความพึงพอใจของครูที่มีต่อการจัดการภายในสถานศึกษา (β = -0.028297) ส่วนตัวแปรระดับนักเรียนที่มีอิทธิพลทางบวกสูงที่สุดคือความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนคณิตศาสตร์ (β = 0.170069) รองลงมาคือความมั่งคั่งของครอบครัว (β = 0.071467) และ การฝึกปฏิบัติทักษะทางด้านคณิตศาสตร์ในชั่วโมง เรียน (β = 0.068224) ส่วนตัวแปรระดับนักเรียนที่

มีอิทธิพลน้อยสูงที่สุดคือการใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเอง (β = -0.037190) รองลงมาคือ ความรู้สึกเกี่ยวความปลอดภัยในการมาเรียนภายในสถานศึกษา (β = 0.005854)

ผลการทดสอบอิทธิพลเชิงสุ่ม (Random effect) พบว่า ค่าเศษเหลือระดับสถานศึกษาของค่าความสามารถของนักเรียนจากการประเมินความสามารถทางคณิตศาสตร์ที่มีการควบคุมตัวแปรระดับนักเรียนและสถานศึกษา (U0) หรือมูลค่าเพิ่มของสถานศึกษามีความผันแปรระหว่างสถานศึกษาอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (χ^2 = 2519.05496, p = 0.000) โดยมีความแปรปรวนระหว่างสถานศึกษาเท่ากับ 0.16014 ซึ่งอธิบายความแปรปรวนได้ประมาณร้อยละ 31.63 และความแปรปรวนภายในสถานศึกษามีค่าเท่ากับ 0.34614 ซึ่งอธิบายความแปรปรวนได้ประมาณร้อยละ 68.37

ผลการวิเคราะห์ในโมเดล 3 (Detect DIF-DDF & Unadjusted) จากผลการวิเคราะห์พหุระดับในโมเดลที่ 3 ผลการทดสอบอิทธิพลคงที่ พบว่าค่าเฉลี่ยของความสามารถของนักเรียนจากการประเมินความสามารถทางคณิตศาสตร์ ทุกสถานศึกษา (G00) มีมีความผันแปรจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (G00 = -0.694172, t = -15.279, p = 0.000) ผลการทดสอบอิทธิพลเชิงสุ่ม พบว่า มีความแปรปรวนของค่าความสามารถนักเรียนจากการประเมินวิชาคณิตศาสตร์ระหว่างสถานศึกษา($U0_k$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (χ^2 = 5066.81360, p = 0.000) โดยมีความแปรปรวนระหว่างสถานศึกษาเท่ากับ 0.29852 ซึ่งอธิบายความแปรปรวนได้ประมาณร้อยละ 44.75 และความแปรปรวนภายในสถานศึกษามีค่าเท่ากับ 0.36862 ซึ่งอธิบายความแปรปรวนได้ประมาณร้อยละ 55.25

ผลการวิเคราะห์ในโมเดล 4 (Detect DIF-DDF & Adjusted) จากผลการวิเคราะห์พหุระดับในโมเดล 4 ผลการทดสอบอิทธิพลคงที่ (Fixed effect) พบว่า ค่าเฉลี่ยของความสามารถของนักเรียนจากการประเมินความสามารถทางคณิตศาสตร์ทุกสถานศึกษา (G00) มีความผันแปรจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (G00 = -1.936307, t = -7.537, p = 0.000) ค่าสัมประสิทธิ์ถดถอยของตัวแปรระดับสถานศึกษาที่มีค่าเป็นบวกสูงที่สุดคือ สถานศึกษาที่ มีขนาดใหญ่พิเศษ โดยเป็นสถานศึกษาที่มีนักเรียนตั้งแต่2500 คนขึ้นไป (β = 0.765686) นั่นคือความเป็นสถานศึกษาขนาดใหญ่พิเศษจะทำให้ค่าความสามารถของนักเรียนจากการประเมินความสามารถทาง

คณิตศาสตร์มีค่าเพิ่มขึ้น 0.765686 รองลงมาคือ สถานศึกษาที่ตั้งอยู่ในเมือง(β = 0.311936) สถานศึกษาขนาดใหญ่ (β = 0.309957) ส่วนค่าสัมประสิทธิ์ถดถอยของตัวแปรระดับสถานศึกษาที่มีค่าเป็นลบสูงที่สุดคือ การให้การสนับสนุนของผู้ปกครอง (β = -0.069615) รองลงมาคือ การจัดกิจกรรมเสริมทางคณิตศาสตร์ภายในสถานศึกษา (β = -0.041858) และสถานศึกษาที่มีขนาดใหญ่(β = -0.037516) ส่วนตัวแปรระดับนักเรียนที่มีอิทธิพลทางบวกสูงที่สุดคือความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนคณิตศาสตร์ (β = 0.164069) รองลงมาคือ การฝึกปฏิบัติทักษะทางด้านคณิตศาสตร์ในชั่วโมง เรียน (β = 0.069308) และ ความมั่งคั่งของครอบครัว(β = 0.068723) ส่วนตัวแปรระดับนักเรียน ที่มีอิทธิพลน้อยสูงที่สุดคือการใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเอง (β = -0.044132) รองลงมาคือการศึกษาของผู้ปกครองนักเรียน (β = 0.007753)

ผลการทดสอบอิทธิพลเชิงสุ่ม(Random effect) พบว่า ค่าเศษเหลือระดับสถานศึกษาของค่าความสามารถของนักเรียนจากการประเมินความสามารถทางคณิตศาสตร์ที่มีการควบคุมตัวแปรระดับนักเรียนและสถานศึกษา(U0) หรือมูลค่าเพิ่มของสถานศึกษามีความผันแปรระหว่างสถานศึกษาอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (χ^2 = 2548.51324, p = 0.000) โดยมีความแปรปรวนระหว่างสถานศึกษาเท่ากับ0.16232ซึ่งอธิบายความแปรปรวนได้ประมาณร้อยละ 31.99 และความแปรปรวนภายในสถานศึกษามีค่าเท่ากับ 0.34501 ซึ่งอธิบายความแปรปรวนได้ประมาณร้อยละ 68.01

จากการวิเคราะห์สัดส่วนความแปรปรวนทั้งหมดของตัวแปรตามที่อธิบายได้หรือสัมประสิทธิ์การทำนาย (R^2) พบว่าโมเดลที่ 1 มีสัมประสิทธิ์การทำนายเท่ากับ 0.51813 (51.81%) โมเดลที่ 2 มีสัมประสิทธิ์การทำนายเท่ากับ 0.51611 (51.61%) โมเดลที่ 4 มีสัมประสิทธิ์การทำนายเท่ากับ 0.51846 (51.85%) นั่นคือทั้ง 3 โมเดล สามารถอธิบายความแปรปรวนของตัวแปรตามหรือค่าความสามารถของนักเรียนจากการประเมินความสามารถทางคณิตศาสตร์ได้ใกล้เคียงกัน

สรุปผลการวิเคราะห์ในโมเดล1 ถึง โมเดล4 เพื่อแสดงค่าสัมประสิทธิ์ของตัวแปรคุณลักษณะระดับนักเรียนและสถานศึกษาที่ได้จากการวิเคราะห์ และความแปรปรวนในแต่ละระดับ และคำนวณหาค่าสัมประสิทธิ์การทำนายได้ ดังตารางที่28

ตารางที่ 28 สรุปค่าสัมประสิทธ์ของตัวแปรคุณลักษณะระดับนักเรียนและสถานศึกษาที่ได้จากการวิเคราะห์โมเดล 1 ถึง 4

ตัวแปร	Coefficient			
	Model 1	Model 2	Model 3	Model 4
ค่าเฉลี่ยความสามารถของนักเรียน	-1.900923 **	-1.784011**	-0.694172**	-1.936307**
ระดับนักเรียน				
ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา	0.016077 **	0.005484**	NA	0.017101**
ระดับความเป็นเพศหญิง	0.042964 *	0.017711*	NA	0.049061**
ระดับการศึกษาของผู้ปกครอง	0.006838	0.005808	NA	0.007753
แหล่งทรัพยากรที่บ้าน	0.044089 **	0.009935**	NA	0.04771**
ความมั่งคั่งของครอบครัว	0.072858	0.058337	NA	0.068723
การฝึกปฏิบัติทักษะทางด้านคณิตศาสตร์ในชั่วโมงเรียน	0.075259 **	0.023233**	NA	0.069308**
ดัชนีระดับเวลาที่ใช้ในการทำการบ้านวิชาคณิตศาสตร์	0.034006 **	0.012703*	NA	0.030171*
การใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเอง	-0.040604 **	0.012725**	NA	-0.044132**
ความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนคณิตศาสตร์	0.174597 **	0.014711**	NA	0.164069**
ความเพลิดเพลินทางคณิตศาสตร์	0.072164 **	0.014038**	NA	0.067318**
แรงจูงใจภายนอกในการเรียนคณิตศาสตร์	0.010592	0.01761	NA	0.0109
การเห็นคุณค่าโดยทั่วไปในคณิตศาสตร์	0.031805	0.019036*	NA	0.046444*
ความรู้สึกเกี่ยวความปลอดภัยในการมาเรียนภายในสถานศึกษา	0.008864	0.011722	NA	0.011288
ระดับสถานศึกษา			NA	
สถานศึกษาขนาดใหญ่พิเศษ	0.746488 **	0.715612**	NA	0.765686**
สถานศึกษาขนาดใหญ่	0.284704	0.281919	NA	0.309957*

ตารางที่ 28 สรุปค่าสัมประสิทธ์ของตัวแปรคุณลักษณะระดับนักเรียนและสถานศึกษาที่ได้จากการวิเคราะห์โมเดล 1 ถึง 4 (ต่อ)

ตัวแปร	Coefficient			
	Model 1	Model 2	Model 3	Model 4
สถานศึกษาขนาดกลาง	0.035227	0.043076	NA	0.047342
สถานที่ตั้งสถานศึกษาอยู่ใน เมืองใหญ่มาก	0.177807	0.165491	NA	0.160355
สถานที่ตั้งสถานศึกษาอยู่ใน เมืองใหญ่	-0.020505	-0.025896	NA	-0.037516
สถานที่ตั้งสถานศึกษาอยู่ในเมือง	0.322065 *	0.309145**	NA	0.311936**
สถานที่ตั้งสถานศึกษาอยู่ใน เมืองเล็ก	0.046254	0.028721	NA	0.037242
สถานที่ตั้งสถานศึกษาอยู่ในตำบล	0.093346	0.077362	NA	0.101627
การขาดแคลนครูคณิตศาสตร์	0.003038	0.002975	NA	0.000395
การจัดกิจกรรมเพื่อการเรียนรู้คณิตศาสตร์	-0.052732	-0.047561	NA	-0.069615
ดัชนีการสนับสนุนการจัดสรรทรัพยากร	0.002421	0.007441	NA	0.017139
การจัดกิจกรรมเพื่อการเรียนรู้คณิตศาสตร์	-0.034328	-0.034449	NA	-0.041858
ดัชนีการรับรู้ถึงบรรยากาศที่ดี ในการทำงานภายในสถานศึกษา	0.033867	0.033164	NA	0.028842
การพัฒนาครูในสถานศึกษา	0.030584	0.029958	NA	0.033932
ประสบการณ์การทำงานของครูในสถานศึกษา	-0.005452	-0.005342	NA	-0.004737
ดัชนีความพึงพอใจของครูที่มีต่อการจัดการภายในสถานศึกษา	-0.026689	-0.028297	NA	-0.026939
ดัชนีการให้ความสำคัญกับการให้การบ้านคณิตศาสตร์ข องครูในสถานศึกษา	0.105993 *	0.106145 *	NA	0.107073*
ขนาดของห้องเรียน	-0.005824	-0.008604	NA	-0.027418
ระดับความถี่ในการทำงานร่วมกันของครูคณิตศาสตร์ในสถานศึกษา	0.030774	0.02676	NA	0.042742

ตารางที่ 28 สรุปค่าสัมประสิทธ์ของตัวแปรคุณลักษณะระดับนักเรียนและสถานศึกษาที่ได้จากการวิเคราะห์โมเดล 1 ถึง 4 (ต่อ)

ตัวแปร	Coefficient			
	Model 1	Model 2	Model 3	Model 4
variance components	Model 1	Model 2	Model 3	Model 4
ระดับที่ 1 (R)	0.33539	0.34614	0.36862	0.34501
ระดับที่ 2 (U0)	0.17316	0.16014	0.29852	0.16232
ระดับความผันแปรของ ค่าความสามารถของนักเรียนจากการประเมินความสามารถวิชาคณิตศาสตร์ที่	Model 1	Model 2	Model 3	Model 4
ระดับภายในสถานศึกษา	65.95	68.37	55.25	68.01
ระดับภายระหว่างสถานศึกษา	34.05	31.63	44.75	31.99
สัดส่วนของความแปรปรวนทั้งหมดของตัวแปรตามที่อธิบายได้ (R^2)				
ระดับที่ 1	0.07040 (7.04%)	0.06484 (6.48%)	NA	0.06405 (6.41%)
ระดับที่ 2	0.44922 (44.92%)	0.45376 (45.38%)	NA	0.45625 (45.63%)
รวมทั้งสองระดับ	0.51962 (51.96%)	0.51860 (51.86%)	NA	0.52030 (52.03%)

3.2 ผลการประมาณค่าความสามารถของผู้สอบในแต่ละโมเดล

การประมาณค่าความสามารถของนักเรียนจากการประเมินผลความสามารถทางคณิตศาสตร์จากการทดสอบนานาชาติ TIMSS 2007 ผู้วิจัยใช้โปรแกรม MULTILOG ในการวิเคราะห์ประมาณค่าความสามารถของผู้สอบ ผู้วิจัยได้ดำเนินการวิเคราะห์ 3 ครั้ง คือ ครั้งที่ 1 เป็นการวิเคราะห์ค่าความสามารถของผู้สอบก่อนการตัดข้อสอบที่ทำหน้าที่ต่างกัน ครั้งที่ 2 เป็นการวิเคราะห์ค่าความสามารถของผู้สอบหลังจากการตัดข้อสอบที่ทำหน้าที่ต่างกันพิจารณาเฉพาะตัวเลือกที่เป็นตัวถูก และครั้งที่ 3 เป็นการวิเคราะห์ค่าความสามารถของผู้สอบหลังจากกา รตัดข้อสอบที่ทำหน้าที่ต่างกันพิจารณาทั้งการเกิดกับตัวเลือกที่เป็นตัวถูก และตัวเลือกที่เป็นตัวลวง

ผลการประมาณค่าความสามารถของผู้สอบก่อนการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป พบว่า โดยภาพรวมทั้ง14 ฉบับ มีค่าความสามารถโดยเฉลี่ยเท่ากับ0.54822 ส่วนเบี่ยงเบนมาตรฐาน

0.842 ผู้สอบมีค่าความสามารถต่ำสุดเท่ากับ-2.505 และค่าความสามารถสูงสุดเท่ากับ2.429 เมื่อพิจารณาความสามารถของกลุ่มผู้สอบที่ทำแบบสอบในแต่ละฉบับ พบว่ากลุ่มผู้สอบที่ทำแบบสอบฉบับที่ 4 มีความสามารถของผู้สอบโดยเฉลี่ยต่ำสุด (Mean = -0.83703, S.D. = 0.727) และกลุ่มผู้สอบที่ทำแบบสอบฉบับที่2 มีค่าเฉลี่ยของความสามารถของผู้สอบสูงสุด (Mean = -0.24098, S.D. = 0.814) นอกจากนี้ยังพบว่ากลุ่มผู้สอบที่ทำแบบสอบในแต่ละฉบับมีลักษณะการกระจายค่าความสามารถแบบเบ้ขวา นั่นคือส่วนใหญ่มีความสามารถต่ำกว่าค่าความสามารถเฉลี่ย แสดงรายละเอียดดังตารางที่ 29

ตารางที่ 29 แสดงค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบก่อนการตัดข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบฉบับที่ 1 ถึง 14

ฉบับ	N	Range	Min	Max	Mean	SD	Var	Sk	Ku
1	390	4.115	-2.027	2.088	-0.51631	0.768	0.590	0.772	0.377
2	396	4.233	-2.088	2.145	-0.24098	0.814	0.663	0.335	-0.287
3	384	3.897	-2.058	1.839	-0.37088	0.701	0.491	0.396	0.047
4	385	4.252	-2.505	1.747	-0.83703	0.727	0.529	0.593	0.429
5	390	4.853	-2.439	2.414	-0.62455	0.901	0.812	0.820	0.427
6	383	4.781	-2.404	2.377	-0.49792	0.846	0.715	0.580	0.186
7	382	4.492	-2.088	2.404	-0.64368	0.897	0.805	1.035	0.899
8	388	4.858	-2.429	2.429	-0.61188	0.928	0.862	0.920	0.437
9	386	4.022	-1.995	2.027	-0.38095	0.826	0.682	0.493	-0.408
10	380	4.727	-2.377	2.350	-0.49582	0.806	0.650	0.499	0.096
11	388	3.865	-2.058	1.807	-0.39495	0.757	0.574	0.464	0.055
12	387	4.321	-2.231	2.090	-0.52091	0.842	0.709	0.680	0.341
13	390	4.225	-2.225	2.000	-0.80340	0.811	0.657	0.888	0.510
14	383	4.407	-2.317	2.090	-0.74236	0.901	0.812	0.911	0.461
รวม	**5412**	**4.934**	**-2.505**	**2.429**	**-0.54822**	**0.842**	**0.708**	**0.643**	**0.166**

ผลการประมาณค่าความสามารถของผู้สอบหลังการตัดข้อสอบที่ทำหน้าที่ต่างกัน (DIF) พิจารณาเฉพาะตัวเลือกที่เป็นตัวถูกพบว่า โดยภาพรวมทั้ง 14 ฉบับ มีค่าความสามารถโดยเฉลี่ยเท่ากับ -0.51754 ส่วนเบี่ยงเบนมาตรฐาน 0.834 ผู้สอบมีค่าความสามารถต่ำสุดเท่ากับ -2.283 และค่าความสามารถสูงสุดเท่ากับ 2.322 เมื่อพิจารณาความสามารถของกลุ่มผู้สอบที่ทำแบบสอบในแต่ละฉบับ พบว่ากลุ่มผู้สอบที่ทำแบบสอบฉบับที่ 13 มีความสามารถของผู้สอบโดยเฉลี่ยต่ำสุด (Mean = -0.90650, S.D. = 0.794) และกลุ่มผู้สอบที่ทำแบบสอบฉบับที่ 2 มีค่าเฉลี่ยของความสามารถของผู้สอบสูงสุด (Mean = -0.18240, S.D. = 0.806) นอกจากนี้ยังพบว่ากลุ่มผู้สอบ

ที่ทำแบบสอบในแต่ละฉบับมีลักษณะการกระจายค่าความสามารถแบบเบ้ขวา นั่นคือส่วนใหญ่มีความสามารถต่ำกว่าค่าความสามารถเฉลี่ย แสดงรายละเอียดดังตารางที่ 30

ตารางที่ 30 แสดงค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบหลังการตัดข้อสอบที่ทำหน้าที่ต่างกัน (DIF) พิจารณาเฉพาะตัวเลือกที่เป็นตัวถูก ในแบบสอบฉบับที่ 1 ถึง 14

ฉบับ	N	Range	Min	Max	Mean	SD	Variance	Ske	Kur
1	390	4.292	-2.164	2.128	-0.58944	0.793	0.629	1.062	1.157
2	396	4.429	-2.198	2.231	-0.18240	0.806	0.650	0.320	-0.033
3	384	4.124	-1.926	2.198	-0.26735	0.773	0.597	0.570	0.191
4	385	4.064	-2.283	1.781	-0.65147	0.692	0.479	0.380	0.264
5	390	3.970	-2.051	1.919	-0.67452	0.828	0.686	0.641	0.019
6	383	4.494	-2.263	2.231	-0.49882	0.868	0.754	0.632	0.097
7	382	4.585	-2.263	2.322	-0.58860	0.895	0.800	0.974	0.630
8	388	4.462	-2.231	2.231	-0.59223	0.878	0.771	0.925	0.481
9	386	4.017	-1.889	2.128	-0.24330	0.866	0.749	0.593	-0.270
10	380	4.429	-2.231	2.198	-0.49250	0.742	0.551	0.551	0.196
11	388	3.921	-1.995	1.926	-0.37277	0.764	0.583	0.441	0.018
12	387	3.970	-2.051	1.919	-0.49082	0.772	0.596	0.753	0.425
13	390	4.186	-2.278	1.908	-0.90650	0.794	0.630	1.001	0.767
14	383	4.388	-2.231	2.157	-0.69816	0.871	0.758	0.915	0.557
รวม	**5412**	**4.605**	**-2.283**	**2.322**	**-0.51754**	**0.834**	**0.695**	**0.663**	**0.191**

ผลการประมาณค่าความสามารถของผู้สอบหลังการตัดข้อสอบที่ทำหน้าที่ต่างกันของตัวลวง (DDF) พิจารณาการเกิดการทำหน้าที่ต่างทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง พบว่า โดยภาพรวมทั้ง 14 ฉบับ มีค่าความสามารถโดยเฉลี่ยเท่ากับ -0.62037 ส่วนเบี่ยงเบนมาตรฐาน 0.836 ผู้สอบมีค่าความสามารถต่ำสุดเท่ากับ -2.263 และค่าความสามารถสูงสุดเท่ากับ 2.293 เมื่อพิจารณาความสามารถของกลุ่มผู้สอบที่ทำแบบสอบในแต่ละฉบับ พบว่ากลุ่มผู้สอบที่ทำแบบสอบฉบับที่ 13 มีความสามารถของผู้สอบโดยเฉลี่ยต่ำสุด (Mean = -0.95472, S.D. = 0.789) และกลุ่มผู้สอบที่ทำแบบสอบฉบับที่ 2 มีค่าเฉลี่ยของความสามารถของผู้สอบสูงสุด (Mean = -0.33141, S.D. = 0.802) นอกจากนี้ยังพบว่ากลุ่มผู้สอบที่ทำแบบสอบในแต่ละฉบับมีลักษณะการกระจายค่าความสามารถแบบเบ้ขวา นั่นคือส่วนใหญ่มีความสามารถต่ำกว่าค่าความสามารถเฉลี่ย แสดงรายละเอียดดังตารางที่ 31

ตารางที่ 31 แสดงค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบหลังการตัดข้อสอบที่ทำหน้าที่ต่างกันของตัวลวง (DDF) พิจารณาการเกิดการทำหน้าที่ต่างทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง ในแบบสอบฉบับที่ 1 ถึง 14

ฉบับ	N	Range	Min	Max	Mean	SD	Var	Sk	Ku
1	390	4.141	-2.09	2.051	-0.60736	0.775	0.601	0.986	1.149
2	396	4.362	-2.164	2.198	-0.33141	0.802	0.643	0.522	0.141
3	384	4.254	-2.164	2.09	-0.41374	0.746	0.557	0.589	0.281
4	385	4.016	-2.246	1.77	-0.80201	0.720	0.519	0.506	0.361
5	390	4.219	-2.219	2	-0.75165	0.859	0.738	0.736	0.202
6	383	4.395	-2.231	2.164	-0.49583	0.890	0.792	0.505	-0.237
7	382	4.556	-2.263	2.293	-0.66618	0.873	0.762	0.944	0.753
8	388	4.396	-2.198	2.198	-0.72729	0.888	0.789	0.983	0.534
9	386	3.979	-2.128	1.851	-0.46382	0.824	0.679	0.352	-0.512
10	380	4.494	-2.263	2.231	-0.57482	0.792	0.628	0.534	0.134
11	388	4.053	-2.164	1.889	-0.46876	0.779	0.608	0.504	0.105
12	387	4.056	-2.09	1.966	-0.62067	0.851	0.725	0.702	0.204
13	390	4.016	-2.213	1.803	-0.95472	0.789	0.623	0.960	0.657
14	383	4.111	-2.192	1.919	-0.80970	0.856	0.732	1.006	0.591
รวม	**5412**	**4.556**	**-2.263**	**2.293**	**-0.62037**	**0.836**	**0.698**	**0.664**	**0.168**

3.3 ผลของข้อสอบที่ทำหน้าที่ต่างกันต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน

โมเดลการประเมินคุณภาพการจัดการศึกษา ในงานวิจัยเรื่องนี้ หมายถึง การเปรียบเทียบผลการจัดระดับคุณภาพและการจัดอันดับคุณภาพของสถานศึกษาจากคะแนนมูลค่าเพิ่ม โดยการจัดระดับคุณภาพจัดแบ่งออกเป็นกลุ่ม 3 กลุ่ม คือ กลุ่มที่ 1 เป็นกลุ่มต่ำ มีตำแหน่งเปอร์เซ็นต์ไทล์ 1 – 24.99 กลุ่มที่ 2 เป็นกลุ่มกลาง มีตำแหน่งเปอร์เซ็นต์ไทล์ 25 – 74.99 และ กลุ่มที่ 3 เป็นกลุ่มสูง มีตำแหน่งเปอร์เซ็นต์ไทล์ 75 ขึ้นไป และการจัดอันดับคุณภาพเป็นการเรียงอันดับคะแนนมูลค่าเพิ่มของสถานศึกษาจากมากที่สุดไปน้อยที่สุด

จากการศึกษาผลของข้อสอบที่ทำหน้าที่ต่างกันต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน พบว่า ข้อสอบที่เกิดการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวงในแบบสอบ

ประเมินความรู้วิชาคณิตศาสตร์ของสถานศึกษาขั้นพื้นฐานจะทำการเปรียบเทียบผลการประเมินคุณภาพการจัดการศึกษาระหว่างโมเดล 1 กับ โมเดล 2, โมเดล 1 กับ โมเดล 4 และ โมเดล 2 กับ โมเดล 4 โดยทั้ง 3 โมเดลที่มีการนำมาเปรียบเทียบล้วนแต่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา แต่มีความแตกต่างกันในลักษณะของการข้อสอบที่ใช้ในการวิเคราะห์ พบว่า

โมเดล 1 กับโมเดล 2 ทั้งสองโมเดลมีความแตกต่างกันคือโมเดลที่ 1 เป็นโมเดลที่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา แต่ไม่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป ส่วนโมเดล 2 เป็นโมเดลที่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา แต่ตัดข้อสอบที่ทำหน้าที่ต่างกันเฉพาะการเกิดขึ้นกับตัวเลือกที่เป็นตัวถูกออกไป

ผลการจัดระดับคุณภาพการจัดการศึกษา พบว่า โมเดล 1 และโมเดล 2 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาได้สอดคล้องกัน 148 แห่ง คิดเป็นร้อยละ 98.67 โดยจัดเป็นกลุ่มต่ำจำนวน 37 แห่ง (24.67%) กลุ่มกลาง 75 แห่ง (50%) และกลุ่มสูง 36 แห่ง (24%) และพบว่าจัดกลุ่มไม่สอดคล้องกันจำนวน 2 แห่ง คิดเป็นร้อยละ 1.34 โดยผลการวิเคราะห์ในโมเดล 1 จัดเป็นกลุ่มกลางแต่ในโมเดล 2 จัดเป็นกลุ่มสูง จำนวน 1 แห่ง คิดเป็นร้อยละ 0.67 และผลการวิเคราะห์ในโมเดล 1 จัดเป็นกลุ่มสูงแต่ในโมเดล 2 จัดเป็นกลุ่มกลาง จำนวน 1 แห่ง คิดเป็นร้อยละ 0.67

ผลการจัดอันดับคุณภาพการจัดการศึกษา พบว่า สถานศึกษาที่มีอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 ตรงกับ อันดับคุณภาพ การจัด การศึกษาจากผลการวิเคราะห์โมเดล 2 มีจำนวน 28 แห่ง คิดเป็นร้อยละ 18.67 สถานศึกษาที่มีอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะ ห์ โมเดล 1 สูงกว่าอันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์โมเดล 2 มี 53 แห่งจากสถานศึกษาทั้งหมด 150 แห่ง คิดเป็น 39.33% และสถานศึกษาที่มีอันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 ต่ำกว่าอันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์โมเดล 2 มี 63 แห่งจากสถานศึกษาทั้งหมด 150 แห่ง คิดเป็น 42% เมื่อทดสอบความสอดคล้องของการจัด ระดับคุณภาพการจัดการศึกษาและ อันดับคุณภาพการจัดการศึกษาจากการใช้โมเดล 1 และ 2 ด้วยสถิติทดสอบวิลคอกซอน (Wilcoxon Signed Ranks Test) พบว่าโมเดล 1 และ 2 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาไม่แตกต่างกัน ($Z = -.248$, $p = .804$) นั่นคือผลการประเมินคุณภาพการจัดการศึกษาด้วยโมเดล 1 และ 2 มีผลการจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาสอดคล้องกัน และเมื่อทดสอบความสอดคล้อง

ของการจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษาพบว่าไม่แตกต่างกัน (Z = -.384 , p = .701) นั่นคือผลการประเมินคุณภาพการจัดการศึกษาด้วยโมเดล 1 และ 2 มีผลการจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษาสอดคล้องกัน แสดงรายละเอียดดังตารางที่ 32

ตารางที่ 32 เปรียบเทียบผลการจัดระดับคุณภาพการจัดการศึกษา การจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษา ระหว่างโมเดล 1 (Undetected DIF-DDF&Adjusted) กับโมเดล 2 (Detect DIF&Adjusted) และค่าสถิติทดสอบ Wilcoxon Signed Ranks Test

	ผลการจัดระดับคุณภาพการจัดการศึกษา				
	ระดับคุณภาพการจัดการศึกษา	**โมเดล 2**			
		กลุ่มต่ำ	**กลุ่มกลาง**	**กลุ่มสูง**	**รวม**
โมเดล 1	กลุ่มต่ำ	37 (24.67%)	-	-	37 (24.67%)
	กลุ่มกลาง	-	75 (50%)	1 (0.67%)	76 (50.67%)
	กลุ่มสูง	-	1 (0.67%)	36 (24%)	37 (24.67%)
	รวม	37 (24.67%)	76 (50.67%)	37 (24.67%)	150 (100%)
	Wilcoxon Signed Ranks Test : Z = -.248 , p = .804				
โมเดล 1	**ผลการจัดอันดับคุณภาพการจัดการศึกษา**				
	ระดับสูงกว่าโมเดล 2		**ระดับตรงกันกับโมเดล 2**		**ระดับต่ำกว่าโมเดล 2**
	59 (39.33%)		28 (18.67%)		63 (42%)
	Wilcoxon Signed Ranks Test : Z = -.384 , p = .701				

โมเดล 1 กับโมเดล 4 ทั้งสองโมเดลมีความแตกต่างกันคือโมเดลที่ 1 เป็นโมเดลที่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา แต่ไม่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป ส่วนโมเดล 4 เป็นโมเดลที่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา พร้อมทั้งตัดข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวงออกไป

ผลการจัด ระดับคุณภาพ การจัดการศึกษา พบว่า โมเดล 1 และโมเดล 4 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษา ได้สอดคล้องกัน 148 แห่ง คิดเป็นร้อยละ 98.67 โดยจัดเป็นกลุ่มต่ำจำนวน 37 แห่ง (24.67%) กลุ่มกลาง 75 แห่ง (50%) และกลุ่มสูง 36 แห่ง (24%) และพบว่า จัดกลุ่มไม่สอดคล้องกันจำนวน 2 แห่ง คิดเป็นร้อยละ 1.34 โดยผลการวิเคราะห์ในโมเดล 1 จัดเป็นกลุ่มกลางแต่ในโมเดล 4 จัดเป็นกลุ่มสูง จำนวน 1 แห่ง คิดเป็นร้อย

ละ 0.67 และผลการวิเคราะห์ในโมเดล 1 จัดเป็นกลุ่มสูงแต่ในโมเดล 4 จัดเป็นกลุ่มกลาง จำนวน 1 แห่ง คิดเป็นร้อยล ะ 0.67

ผลการจัดอันดับคุณภาพการจัดการศึกษา พบว่า สถานศึกษาที่มีอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 ตรงกับอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มีจำนวน 18 แห่ง คิดเป็นร้อยละ 12 สถานศึกษาที่มีอันดับคุณภาพการจัดการศึกษาจากผลการวิเค ราะห์ โมเดล 1 สูงกว่าอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มี 61 แห่งจากสถานศึกษาทั้งหมด 150 แห่ง คิดเป็น 40.67% และสถานศึกษาที่มีอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 ต่ำกว่าอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มี 71 แห่งจากสถานศึกษาทั้งหมด 150 แห่ง คิดเป็น 47.33% เมื่อทดสอบความสอดคล้องของการจัดระดับคุณภาพการจัดการศึกษาและอันดับคุณภาพการจัดการศึกษาจากการใช้โมเดล 1 และ 4 ด้วยสถิติทดสอบวิลคอกซอน (Wilcoxon Signed Ranks Test) พบว่าโมเดล 1 และ 4 จัดระดับคุณภาพการจัดการศึก ษาของสถานศึกษาไม่แตกต่างกัน ($Z = -.527$, $p = .598$) นั่นคือผลการประเมินคุณภาพการจัดการศึกษาด้วยโมเดล 1 และ 4 มีผลการจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาสอดคล้องกัน และเมื่อทดสอบความสอดคล้องของการจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษาพบว่าไม่แตกต่างกัน ($Z = -.134$, $p = .893$) นั่นคือผลการประเมินคุณภาพการจัดการศึกษาด้วยโมเดล 1 และ 4 มีผลการจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษาสอดคล้องกัน แสดงรายละเอียดดังตารางที่ 33

ตารางที่ 33 เปรียบเทียบผลการจัดระดับคุณภาพการจัดการศึกษา การจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษาระหว่างโมเดล1 (Undetected DIF-DDF&Adjusted) กับโมเดล 4 (Detect DIF-DDF&Adjusted) และค่าสถิติทดสอบ Wilcoxon Signed Ranks Test

	ผลการจัดระดับคุณภาพการจัดการศึกษา				
	ระดับคุณภาพการจัดการศึกษา	**โมเดล 4**			
		กลุ่มต่ำ	**กลุ่มกลาง**	**กลุ่มสูง**	**รวม**
โมเดล 1	กลุ่มต่ำ	37 (24.67%)	-	-	37 (24.67%)
	กลุ่มกลาง	-	75 (50%)	1 (0.67%)	76 (50.67%)
	กลุ่มสูง	-	1 (0.67%)	36 (24%)	37 (24.67%)
	รวม	37 (24.67%)	76 (50.67%)	37 (24.67%)	150 (100%)
	Wilcoxon Signed Ranks Test : Z = -.527 , p = .598				
โมเดล 1	**ผลการจัดอันดับคุณภาพการจัดการศึกษา**				
	ระดับสูงกว่าโมเดล 4		**ระดับตรงกันกับโมเดล 4**	**ระดับต่ำกว่าโมเดล 4**	
	61 (40.67%)		18 (12%)	71 (47.33%)	
	Wilcoxon Signed Ranks Test : Z = -.134, p = .893				

โมเดล 2 กับโมเดล 4 ทั้งสองโมเดลมีความแตกต่างกันคือโมเดลที่ 2 เป็นโมเดลที่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา พร้อมทั้งมีการตัดข้อสอบที่ทำหน้าที่ต่างกันเฉพาะที่เป็นตัวเลือกที่เป็นตัวถูกออกไป ส่วนโมเดล 4 เป็นโมเดลที่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา พร้อมทั้งตัดข้อสอบที่ทำ หน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวงออกไป

ผลการจัด ระดับคุณภาพ การจัดการศึกษา พบว่า โมเดล 2 และโมเดล 4 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษา ได้สอดคล้องกัน 148 แห่ง คิดเป็นร้อยละ 98.67 โดยจัดเป็นกลุ่มต่ำจำนวน 37 แห่ง (24.67%) กลุ่มกลาง 75 แห่ง (50%) และกลุ่มสูง 36 แห่ง (24%) และพบว่าจัดกลุ่มไม่สอดคล้องกันจำนวน 2 แห่ง คิดเป็นร้อยละ 1.34 โดยผลการวิเคราะห์ในโมเดล 2 จัดเป็นกลุ่มกลางแต่ในโมเดล 4 จัดเป็นกลุ่มสูง จำนวน 1 แห่ง คิดเป็นร้อยละ 0.67 และผลการวิเคราะห์ในโมเดล 2 จัดเป็นกลุ่มสูงแต่ในโมเดล 4 จัดเป็นกลุ่มกลาง จำนวน 1 แห่ง คิดเป็นร้อยละ 0.67

ผลการจัดอันดับคุณภาพการจัดการศึกษา พบว่า สถานศึกษาที่มีอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล2 ตรงกับอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มีจำนวน 15 แห่ง คิดเป็นร้อยละ 10 สถานศึกษาที่มีอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 2 สูงกว่าอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล4 มี 71 แห่งจากสถานศึกษาทั้งหมด 150 แห่ง คิดเป็น 47.33% และสถานศึกษาที่มีอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล2 ต่ำกว่าอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มี 64 แห่งจากสถานศึกษาทั้งหมด 150 แห่ง คิดเป็น 42.67% เมื่อทดสอบความสอดคล้องของการจัดระดับคุณภาพการจัดการศึกษาและอันดับคุณภาพการจัดการศึกษาจากการใช้โมเดล 2 และ 4 ด้วยสถิติทดสอบวิลคอกซอน (Wilcoxon Signed Ranks Test) พบว่าโมเดล 2 และ 4 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาไม่แตกต่างกัน (Z = -.697, p = .486) นั่นคือผลการประเมินคุณภาพการจัดการศึกษาด้วยโมเดล 2 และ 4 มีผลการจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาสอดคล้องกัน และเมื่อทดสอบความสอดคล้องของการจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษาพบว่าไม่แตกต่างกัน (Z = -.154, p = .878) นั่นคือผลการประเมินคุณภาพการจัดการศึกษาด้วยโมเดล 2 และ 4 มีผลการจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษาสอดคล้องกัน แสดงรายละเอียดดังตารางที่34

ตารางที่ 34 เปรียบเทียบผลการจัดระดับคุณภาพการจัดการศึกษา การจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษา ระหว่างโมเดล2 ((Detect DIF&Adjusted)) กับโมเดล 4 (Detect DIF-DDF&Adjusted) และค่าสถิติทดสอบ Wilcoxon Signed Ranks Test

	ผลการจัดระดับคุณภาพการจัดการศึกษา				
	ระดับคุณภาพการจัดการศึกษา	**โมเดล 4**			
		กลุ่มต่ำ	**กลุ่มกลาง**	**กลุ่มสูง**	**รวม**
โมเดล 2	กลุ่มต่ำ	37 (24.67%)	-	-	37 (24.67%)
	กลุ่มกลาง	-	75 (50%)	1 (0.67%)	76 (50.67%)
	กลุ่มสูง	-	1 (0.67%)	36 (24%)	37 (24.67%)
	รวม	37 (24.67%)	76 (50.67%)	37 (24.67%)	150 (100%)
	Wilcoxon Signed Ranks Test : Z = -.697, p = .486				
โมเดล 2	**ผลการจัดอันดับคุณภาพการจัดการศึกษา**				
	ระดับสูงกว่าโมเดล 4		**ระดับตรงกันกับโมเดล 4**		**ระดับต่ำกว่าโมเดล 4**
	71 (47.33%)		15 (10%)		64 (42.67%)
	Wilcoxon Signed Ranks Test : Z = -.154, p = .878				

จากผลการเปรียบเทียบการจัด ระดับคุณภาพ การจัดการศึกษาและ อันดับคุณภาพ การจัดการศึกษาของสถานศึกษาระหว่างโมเดล 1 และ 2 ซึ่งเป็นการ ประเมินด้วยโมเดลที่ไม่ตัดข้อสอบที่ทำ หน้าที่ต่างกันออกไป และโมเดลที่ตัดข้อสอบที่ทำ หน้าที่ต่างกัน เฉพาะตัวเลือกที่เป็นตัวถูกออกไป โมเดล 1 และ 4 เป็นโมเดล ที่ไม่ตัดข้อสอบที่ทำ หน้าที่ต่างกันออกไป และโมเดลที่ตัดข้อสอบที่ทำ หน้าที่ต่างกัน ทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง ออกไป และโมเดล 2 และ 4 ซึ่งเป็นการเปรียบเทียบระหว่างโมเดลที่ ตัดข้อสอบที่ทำ หน้าที่ต่างกัน เฉพาะตัวเลือกที่เป็นตัวถูก ออกไป กับโมเดลที่ตัดข้อส อบที่ทำ หน้าที่ต่างกัน ทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง ออกไป พบว่าผลการประเมิน ทั้ง 3 แบบ จะทำ ให้ผลการจัด ระดับคุณภาพ และอันดับคุณภาพ การจัดการศึกษาของสถานศึกษาไม่ต่างกัน นั่นคือ การทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) ในแบบสอบวิชา คณิตศาสตร์ ไม่มีผลต่อการประเมินคุณภาพการจัดการศึกษา

3.4 ผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน

การสรุปผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาขั้นพื้นฐานที่มีผลต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ จะทำการเปรียบเทียบผลการประเมินคุณภาพการจัดการศึกษาระหว่างโมเดล 1 กับ โมเดล 3, โมเดล 2 กับ โมเดล 3 และ โมเดล 3 กับโมเดล 4 โดยทั้ง 3 โมเดลที่มีการนำมาเปรียบเทียบล้วนแต่มีความแตกต่างกันตามลักษณะของโมเดลแต่ละแบบ พบว่า

โมเดล 1 กับโมเดล 3 ทั้งสองโมเดลมีความแตกต่างกันคือโมเดลที่ 1 เป็นโมเดลที่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา แต่ไม่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป ส่วนโมเดล 3 เป็นโมเดลที่ไม่มีการควบคุมอิ ทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา แต่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวงออกไป

ผลการจัด ระดับคุณภาพ การจัดการศึกษา พบว่า โมเดล 1 และโมเดล 3 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาอยู่ในกลุ่มเดียวกั นได้ 96 แห่ง คิดเป็นร้อยละ 64 โดย

จัดเป็นกลุ่มต่ำจำนวน 21 แห่ง (14%) กลุ่มกลาง 49 แห่ง (32.67%) และกลุ่มสูง 26 แห่ง 17.33%) และพบว่าจัดกลุ่มไม่สอดคล้องกันจำนวน 54 แห่ง คิดเป็นร้อยละ 36 โดยผลการวิเคราะห์ในโมเดล 1 จัดเป็นกลุ่มกลางแต่ในโมเดล 3 จัดเป็นกลุ่มต่ำ จำนวน 16 แห่ง (10.67%), โมเดล 1 จัดเป็นกลุ่มกลางแต่ในโมเดล 3 จัดเป็นกลุ่มสูงจำนวน 11 แห่ง (7.33%), โมเดล 1 จัดเป็นกลุ่มต่ำแต่ในโมเดล 3 จัดเป็นกลุ่มกลางจำนวน 16 แห่ง (10.67%) และโมเดล 1 จัดเป็นกลุ่มสูงแต่ในโมเดล 3 จัดเป็นกลุ่มกลางจำนวน 11 แห่ง (7.33%)

ผลการจัดอันดับคุณภาพการจัดการศึกษา พบว่า สถานศึกษาที่มี อันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 ตรงกับ อันดับคุณภาพ การจัด การศึกษาจากผลการวิเคราะห์โมเดล 3 มีจำนวน 3 แห่ง คิดเป็นร้อยละ 2 สถานศึกษาที่มี อันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์ โมเดล 1 สูงกว่าอันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 มี 70 แห่งจากสถานศึกษาทั้งหมด 150 แห่ง คิดเป็น 46.67% และสถานศึกษาที่มีอันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 ต่ำกว่าอันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 มี 77 แห่งจากสถานศึกษาทั้งหมด 150 แห่ง คิดเป็น 51.33% เมื่อทดสอบความสอดคล้องของการจัด ระดับคุณภาพ การจัดการศึกษาและ อันดับคุณภาพการจัดการศึกษาจากการใช้โมเดล 1 และ 3 ด้วยสถิติทดสอบวิลคอกซอน (Wilcoxon Signed Ranks Test) พบว่าโมเดล 1 และ 3 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาไม่แตกต่างกัน ($Z = -.261$, $p = .794$) แต่จัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษาแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($Z = -2.239$, $p = .025$) นั่นคือผลการประเมินด้วยโมเดล 1 และ 3 จัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษาได้ไม่สอดคล้องกัน แสดงรายละเอียดดังตารางที่ 35

ตารางที่ 35 เปรียบเทียบผลการจัดระดับคุณภาพการจัดการศึกษา การจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษา ระหว่างโมเดล 1 (Undetected DIF-DDF&Adjusted) กับโมเดล 3 (Detect DIF-DDF&Unadjusted) และค่าสถิติทดสอบ Wilcoxon Signed Ranks Test

	ผลการจัดระดับคุณภาพการจัดการศึกษา				
	ระดับคุณภาพการจัดการศึกษา	โมเดล 3			
		กลุ่มต่ำ	กลุ่มกลาง	กลุ่มสูง	รวม
โมเดล 1	กลุ่มต่ำ	21 (14%)	16 (10.67%)	0	37 (24.67%)
	กลุ่มกลาง	16 (10.67%)	49 (32.67%)	11 (7.33%)	76 (50.67%)
	กลุ่มสูง	0	11 (7.33%)	26 (17.33%)	37 (24.67%)
	รวม	37 (24.67%)	76 (50.67%)	37 (24.67%)	150 (100%)
	Wilcoxon Signed Ranks Test : Z = -.261 , p = .794				
โมเดล 1	ผลการจัดอันดับคุณภาพการจัดการศึกษา				
	ระดับสูงกว่าโมเดล 3		ระดับตรงกันกับโมเดล 3	ระดับต่ำกว่าโมเดล 3	
	70 (46.67%)		3 (2%)	77 (51.33%)	
	Wilcoxon Signed Ranks Test : Z = -2.239, p = .025				

โมเดล 2 กับโมเดล 3 ทั้งสองโมเดลมีความแตกต่างกันคือโมเดลที่ 2 เป็นโมเดลที่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา และมีการตัดข้อสอบที่ทำหน้าที่ต่างกันเฉพาะที่เป็นตัวถูกออกไป ส่วนโมเดล 3 เป็นโมเดลที่ไม่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา แต่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวงออกไป

ผลการจัดระดับคุณภาพการจัดการศึกษา พบว่า โมเดล 2 และโมเดล 3 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาอยู่ในกลุ่มเดียวกันได้ 96 แห่ง คิดเป็นร้อยละ 64 โดยจัดเป็นกลุ่มต่ำจำนวน 21 แห่ง (14%) กลุ่มกลาง 49 แห่ง (32.67%) และกลุ่มสูง 26 แห่ง (17.33%) และพบว่าจัดกลุ่มไม่สอดคล้องกันจำนวน 54 แห่ง คิดเป็นร้อยละ 36 โดยผลการวิเคราะห์ในโมเดล 2 จัดเป็นกลุ่มกลางแต่ในโมเดล 3 จัดเป็นกลุ่มต่ำ จำนวน 16 แห่ง (10.67%), โมเดล 2 จัดเป็นกลุ่มกลางแต่ในโมเดล 3 จัดเป็นกลุ่มสูงจำนวน 11 แห่ง (7.33%), โมเดล 2 จัดเป็นกลุ่มต่ำแต่ในโมเดล 3 จัดเป็นกลุ่มกลางจำนวน 16 แห่ง (10.67%) และโมเดล 2 จัดเป็นกลุ่มสูงแต่ในโมเดล 3 จัดเป็นกลุ่มกลางจำนวน 11 แห่ง (7.33%)

ผลการจัดอันดับคุณภาพ การจัดการศึกษา พบว่า สถานศึกษาที่มี อันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์โมเดล2 ตรงกับอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 มีจำนวน 2 แห่ง คิดเป็นร้อยละ 1.33 สถานศึกษาที่มีอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์ โมเดล 2 สูงกว่าอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล3 มี 72 แห่งจากสถานศึกษาทั้งหมด 150 แห่ง คิดเป็น 48% และสถานศึกษาที่มีอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล2 ต่ำกว่าอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 มี 76 แห่งจากสถานศึกษาทั้งหมด 150 แห่ง คิดเป็น 50.67% เมื่อทดสอบความสอดคล้องของการจัดระดับคุณภาพการจัดการศึกษาและอันดับคุณภาพการจัดการศึกษาจากการใช้โมเดล 2 และ 3 ด้วยสถิติทดสอบวิลคอกซอน (Wilcoxon Signed Ranks Test) พบว่าโมเดล 2 และ 3 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาไม่แตกต่างกัน ($Z = -.101$, $p = .919$) แต่จัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษาแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($Z = -2.271$, $p = .023$) นั่นคือผลการประเมินด้วยโมเดล2 และ 3 จัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษาได้ไม่สอดคล้องกัน แสดงรายละเอียดดังตารางที่36

ตารางที่ 36 เปรียบเทียบผลการจัดระดับคุณภาพการจัดการศึกษา การจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษา ระหว่างโมเดล2 (Detect DIF & Adjusted) กับโมเดล 3 (Detect DIF-DDF & Unadjusted) และค่าสถิติทดสอบ Wilcoxon Signed Ranks Test

	ผลการจัดระดับคุณภาพการจัดการศึกษา				
	ระดับคุณภาพการจัดการศึกษา	**โมเดล 3**			
		กลุ่มต่ำ	**กลุ่มกลาง**	**กลุ่มสูง**	**รวม**
โมเดล 2	กลุ่มต่ำ	21 (14%)	16 (10.67%)	0	37 (24.67%)
	กลุ่มกลาง	16 (10.67%)	49 (32.67%)	11 (7.33%)	76 (50.67%)
	กลุ่มสูง	0	11 (7.33%)	26 (17.33%)	37 (24.67%)
	รวม	37 (24.67%)	76 (50.67%)	37 (24.67%)	150 (100%)
	Wilcoxon Signed Ranks Test : $Z = -.101$, $p = .919$				
โมเดล 2	**ผลการจัดอันดับคุณภาพการจัดการศึกษา**				
	ระดับสูงกว่าโมเดล 3		**ระดับตรงกันกับโมเดล 3**		**ระดับต่ำกว่าโมเดล 3**
	72 (48%)		2 (1.33%)		76 (50.67%)
	Wilcoxon Signed Ranks Test : $Z = -2.271$, $p = .023$				

โมเดล 3 กับโมเดล 4 ทั้งสองโมเดลมีความแตกต่างกันคือโมเดลที่ 3 เป็นโมเดลที่ไม่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา แต่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวงออกไป ส่วนโมเดล 4 เป็นโมเดลที่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา และมีการตัดข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวงออกไป

ผลการจัด ระดับ คุณภาพ การจัดการศึกษา พบว่า โมเดล 3 และโมเดล 4 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาอยู่ในกลุ่มเดียวกันได้ 98 แห่ง คิดเป็นร้อยละ 65.33 โดยจัดเป็นกลุ่มต่ำจำนวน 21 แห่ง (14%) กลุ่มกลาง 50 แห่ง (33.33%) และกลุ่มสูง 27 แห่ง (18%) และพบว่าจัดกลุ่มไม่สอดคล้องกัน จำนวน 52 แห่ง คิดเป็นร้อยละ 34.67 โดยผลการวิเคราะห์ในโมเดล 3 จัดเป็นกลุ่มกลางแต่ในโมเดล 4 จัดเป็นกลุ่มต่ำ จำนวน 16 แห่ง (10.67%), โมเดล 3 จัดเป็นกลุ่มกลางแต่ในโมเดล 4 จัดเป็นกลุ่มสูงจำนวน 10 แห่ง (6.67%), โมเดล 3 จัดเป็นกลุ่มต่ำแต่ในโมเดล 4 จัดเป็นกลุ่มกลางจำนวน 16 แห่ง (10.67%) และโมเดล 3 จัดเป็นกลุ่มสูงแต่ในโมเดล 4 จัดเป็นกลุ่มกลางจำนวน 10 แห่ง (6.67%)

ผลการจัดอันดับคุณภาพการจัดการศึกษา พบว่า สถานศึกษาที่มี อันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 ตรงกับ อันดับคุณภาพ การจัด การศึกษาจากผลการวิเคราะห์โมเดล 4 มีจำนวน 6 แห่ง คิดเป็นร้อยละ 4 สถานศึกษาที่มี อันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์ โมเดล 3 สูงกว่าอันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มี 73 แห่งจากสถานศึกษาทั้งหมด 150 แห่ง คิดเป็น 48.67% และสถานศึกษาที่มีอันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 ต่ำกว่าอันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มี 71 แห่งจากสถานศึกษาทั้งหมด 150 แห่ง คิดเป็น 47.33% เมื่อทดสอบความสอดคล้องของการจัด ระดับคุณภาพ การจัดการศึกษาและ อันดับคุณภาพการจัดการศึกษาจากการใช้โมเดล 3 และ 4 ด้วยสถิติทดสอบวิลคอกซอน (Wilcoxon Signed Ranks Test) พบว่าโมเดล 3 และ 4 จัดระดับคุณภาพ การจัดการศึกษาของสถานศึกษาไม่แตกต่างกัน ($Z = -.273$, $p = .785$) แต่จัดอันดับคุณภาพ การจัดการศึกษาของสถานศึกษาแตกต่างกันอย่างมีนัยสำ คัญทางสถิติที่ระดับ .05 ($Z = -2.434$, $p = .015$) นั่นคือผลการประเมินด้วยโมเดล 3 และ 4 จัดอันดับคุณภาพ การจัดการศึกษาของสถานศึกษา ได้ไม่สอดคล้องกัน แสดงรายละเอียดดังตารางที่ 37

ตารางที่ 37 เปรียบเทียบผลการจัดระดับคุณภาพการจัดการศึกษา การจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษา ระหว่างโมเดล 3 (Detect DIF-DDF & Unadjusted) กับโมเดล 4 (Detect DIF-DDF & Adjusted) และค่าสถิติทดสอบ Wilcoxon Signed Ranks Test

	ผลการจัดระดับคุณภาพการจัดการศึกษา				
	ระดับคุณภาพการจัดการศึกษา	**โมเดล 4**			
		กลุ่มต่ำ	**กลุ่มกลาง**	**กลุ่มสูง**	**รวม**
โมเดล 3	กลุ่มต่ำ	21 (14%)	16 (10.67%)	0	37 (24.67%)
	กลุ่มกลาง	16 (10.67%)	50 (33.33%)	10 (6.67%)	76 (50.67%)
	กลุ่มสูง	0	10 (6.67%)	27 (18%)	37 (24.67%)
	รวม	37 (24.67%)	76 (50.67%)	37 (24.67%)	150 (100%)
	Wilcoxon Signed Ranks Test : Z = -.273 , p = .785				
โมเดล 3	**ผลการจัดอันดับคุณภาพการจัดการศึกษา**				
	ระดับสูงกว่าโมเดล 4		**ระดับตรงกันกับโมเดล 4**		**ระดับต่ำกว่าโมเดล 4**
	73 (48.67%)		6 (4%)		71 (47.33%)
	Wilcoxon Signed Ranks Test : Z = -2.434, p = .015				

จากผลการเปรียบเทียบการจัด ระดับคุณภาพ และอันดับคุณภาพ การจัดการศึกษาของสถานศึกษาระหว่างโมเดล 1 และ3, โมเดล 2 และ3 และระหว่างโมเดล 3 และ 4 พบว่าการจัดระดับคุณภาพของสถานศึกษาไม่แตกต่างกัน แต่อันดับคุณภาพการจัดการศึกษามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ แสดงให้เห็นว่าคุณลักษณะของตัวแปรระดับนักเรียนและสถานศึกษามีผลต่อการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา

ตอนที่ 4 ผลการเปรียบเทียบประสิทธิภาพโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์

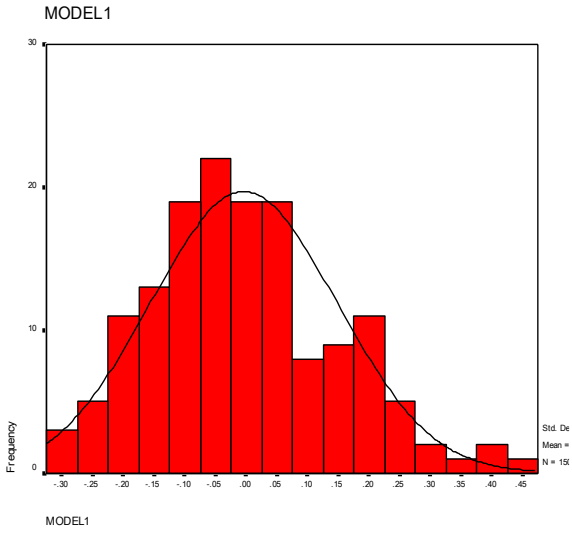
4.1 ผลการวิเคราะห์คุณลักษณะทั่วไปของโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ 4 โมเดล

การประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ พิจารณาจากโมเดลการประเมินคุณภาพการจัดการศึกษา 4 โมเดล โดยคุณภาพการจัดการศึกษาของการวิจัยในครั้งนี้ หมายถึงคะแนนมูลค่าเพิ่มซึ่งเป็นค่าความสามารถในวิชาคณิตศาสตร์ของแต่ละสถานศึกษาที่ได้จากการวิเคราะห์ในโมเดลที่ 1 ถึง โมเดลที่ 4 ซึ่งได้จากการวิเคราะห์พหุระดับ คะแนนมูลค่าเพิ่มของแต่ละ

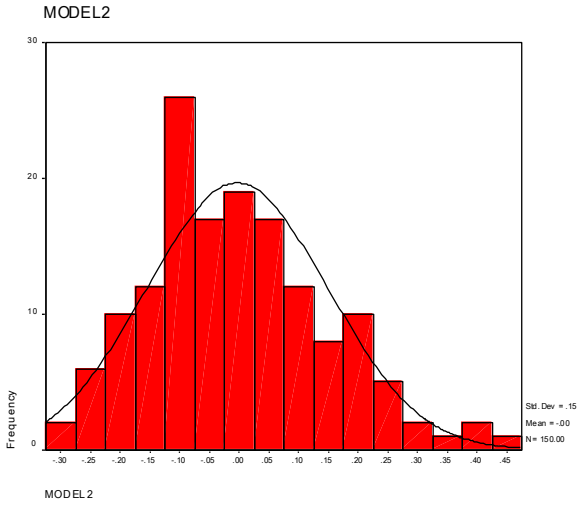
สถานศึกษาที่ได้จากการวิเคราะห์ 4 โมเดลได้ผลที่แตกต่างกัน ดังนี้โมเดลการประเมินคุณภาพการจัดการศึกษา ได้ผลการประเมินพบว่า ค่าสถิติของคะแนน มูลค่าเพิ่มของโมเดลที่ 1, 2 และ 4 มีค่าใกล้เคียงกัน ทั้งค่าเฉลี่ยและค่ามัธยฐานคือ โมเดล 1, 2 และ 4 มีค่า $\overline{X}$ = -0.00362, -0.00312 , -0.00363 และ Med. = -0.01627, -0.01906, -0.02431 ตามลำดับ ส่วนโมเดลที่ 3 มี $\overline{X}$ = -0.02285 และ Med. = -0.04088 ตำแหน่งเปอร์เซ็นต์ไทล์ของทั้ง4 ที่ P_{25} คือ -0.11283, -0.10643, -0.14292 และ -0.09729 ตามลำดับ ส่วน P_{50} โมเดลที่ 1, 2 และ 4 มีค่าใกล้เคียงกันคือ P_{50} = -0.01627, -0.01906 และ -0.02431 ตามลำดับ ส่วน ส่วนโมเดลที่3 P_{50} = -0.04088 และใน P_{75} โมเดลที่ 1, 2 และ 4 มีค่าใกล้เคียงกัน เช่นเดียวกัน คือโมเดลที่ 1 P_{75} = 0.08166, 0.08612 และ 0.09075 ตามลำดับ ส่วนโมเดลที่ 3 P_{75} = 0.09075 แสดงรายละเอียด ดังตารางที่38

ตารางที่ 38 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มวิชาคณิตศาสตร์ของสถานศึกษาจากผลการวิเคราะห์4 โมเดล

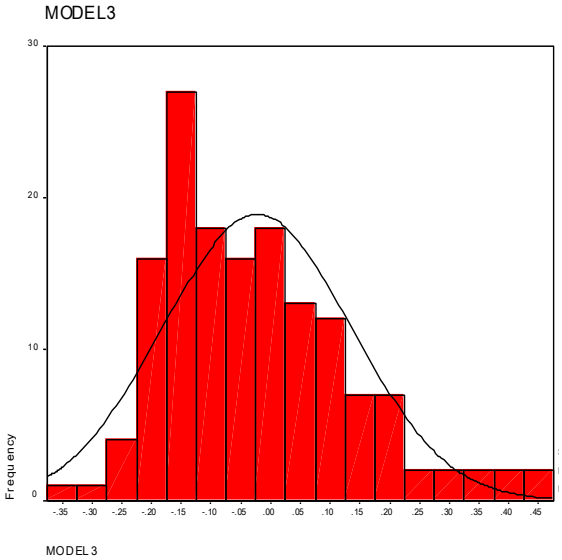
ค่าสถิติ	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
N	150	150	150	150
Mean	-0.00362	-0.00312	-0.02285	-0.00363
(Std. Error of Mean)	0.01237	0.01240	0.01291	0.01229
Median	-0.01627	-0.01906	-0.04088	-0.02431
Mode	0.05508	0.12817	-0.34450	-0.12844
Std. Deviation	0.15146	0.15181	0.15806	0.15051
Variance	0.02294	0.02305	0.02498	0.02265
Skewness	0.47455	0.49767	0.78313	0.52322
(Std. Error of Skewness)	0.19804	0.19804	0.19804	0.19804
Kurtosis	0.08130	0.12326	0.40299	0.27722
(Std. Error of Kurtosis)	0.39358	0.39358	0.39358	0.39358
Range	0.77421	0.77722	0.78882	0.77823
Minimum	-0.32355	-0.32082	-0.34450	-0.32268
Maximum	0.45066	0.45640	0.44432	0.45555
Percentiles 25	-0.11283	-0.10643	-0.14292	-0.09729
Percentiles 50	-0.01627	-0.01906	-0.04088	-0.02431
Percentiles 75	0.08166	0.08612	0.07183	0.09075



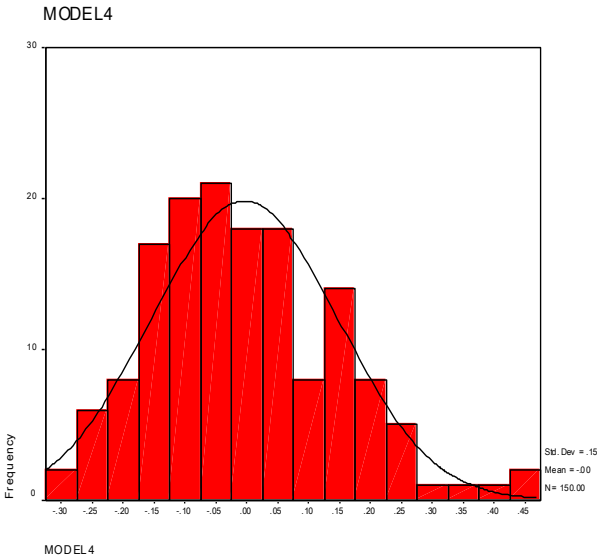
Value-Added Model 1



Value-Added Model 2



Value-Added Model 3



Value-Added Model 4

ภาพที่ 3 ฮีสโตแกรม (Histogram) ของคะแนนมูลค่าเพิ่มจากโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ของสถานศึกษาขั้นพื้นฐาน จำนวน 150 แห่ง

การจัดระดับคุณภาพ การจัดการศึกษาของสถานศึกษาได้จัดเป็นกลุ่ม 3 กลุ่ม คือ กลุ่มที่ 1 เป็นกลุ่มต่ำ มีตำแหน่งเปอร์เซ็นต์ไทล์ 1 – 24.99 โดยในกลุ่มนี้ โมเดลที่ 1, 2 และ 4 มีจำนวนโรงเรียนเท่ ากัน คือ จำนวน 23 โรงเรียน ในขณะที่โมเดลที่ 3 มีจำนวน 32 โรงเรียน กลุ่มที่ 2 เป็นกลุ่มกลาง มีตำแหน่งเปอร์เซ็นต์ไทล์ 25 – 74.99 โดยในกลุ่มนี้ โมเดลที่ 1 มีจำนวน 99

โรงเรียน โมเดลที่ 2 มีจำนวน 94 โรงเรียน 94 โมเดลที่ 3 มีจำนวน 83 โรงเรียน และ โมเดลที่ 4 มีจำนวน 95 โรงเรียน กลุ่มที่ 3 เป็นกลุ่มสูง มีตำแหน่งเปอร์เซ็นต์ไทล์ 75 ขึ้นไป โดยในกลุ่มนี้ โมเดล 1 มีจำนวน 28 โรงเรียน โมเดล 2 มีจำนวน 33 โรงเรียน 94 โมเดล 3 มีจำนวน 35 โรงเรียน และ โมเดล 4 มีจำนวน 32 โรงเรียน

1. ขนาดสถานศึกษา การจัด ระดับคุณภาพ สถานศึกษาจำแนกโด ยขนาดสถานศึกษา พบว่า โมเดล 1, 2 และ 4 มีจำนวนโรงเรียนใกล้เคียงกัน ในทุกขนาดโรงเรียน โดยเฉพาะขนาดใหญ่และ ขนาดใหญ่พิเศษมีจำนวนสถานศึกษาในแต่ละ อันดับคุณภาพ จำนวนเท่ากันทั้ง 3 โมเดล โดยมีรายละเอียดดังนี้ การจัดคุณภาพของสถานศึกษาขนาดเล็ก ขนาดกลาง และขนาดใหญ่ ทั้ง 4 โมเดลจัดสถานศึกษาทั้ง 3 ขนาดคือขนาดเล็ก , ขนาดกลาง และขนาดใหญ่ ถูกจัดให้อยู่ในกลุ่มกลางเป็นส่วนใหญ่ ยกเว้นการจัดคุณภาพของสถานศึกษาขนาดใหญ่พิเศษ ทั้ง 4 โมเดลจัดสถานศึกษาทั้ง 3 ขนาดให้อยู่ในกลุ่ม สูงเป็นส่วนใหญ่

2. ตำแหน่งที่ตั้งสถานศึกษา การจัด ระดับคุณภาพ สถานศึ กษาจำแนกโดยขนาดสถานศึกษาพบว่า โมเดลที่ 1, 2 และ 4 มีจำนวนโรงเรียน เท่ากันในทุกขนาดโรงเรียน โดยมีรายละเอียดดังนี้ การจัดคุณภาพของสถานศึกษา ของโรงเรียนที่ตั้งอยู่ใน หมู่บ้าน, ตำบล, เมืองเล็ก, เมือง และเมืองใหญ่ ทั้ง 4 โมเดลจัดสถานศึกษาทั้ง 5 ตำแหน่งที่ตั้งให้อ ยู่ในกลุ่มกลางเป็นส่วนใหญ่ ยกเว้น โรงเรียนที่ตั้งอยู่ใน เมืองใหญ่มากโมเดลที่ 1, 2 และ 4 จัดระดับคุณภาพสถานศึกษาอยู่ในกลุ่มทุก ระดับคุณภาพ เท่ากัน ส่วนโมเดลที่ 3 จัดระดับคุณภาพ สถานศึกษาอยู่ในกลุ่มกลางเป็นส่วนใหญ่

3. ขนาดของห้องเรียน การจัด ระดับคุณภาพ สถานศึกษาจำแน กโดย ขนาดของห้องเรียน คือนักเรียน 1 – 24 คน, นักเรียน 25 – 40 คน และนักเรียนมากกว่า 41 คน พบว่า โมเดล 1, 2 และ 4 มีจำนวนโรงเรียนใกล้เคียงกันในทุกขนาดห้องเรียน โดยเฉพาะห้องเรียนที่นักเรียนมากกว่า 41 คน มีจำนวนสถานศึกษาในแต่ละ อันดับคุณภาพ จำนวนเท่ากันทั้ง 3 โมเดล และทั้ง 4 โมเดลขนาดห้องเรียนทั้ง 3 ขนาดให้อยู่ในกลุ่มกลางเป็นส่วนใหญ่

4. ระดับการขาดแคลนครูคณิตศาสตร์ในสถานศึกษา การจัด ระดับคุณภาพสถานศึกษาจำแนกโดย ระดับการขาดแคลนครูคณิตศาสตร์ในสถานศึกษา พบว่า โมเดล 1, 2 และ 4 มีจำนวนโรงเรียนใกล้เคียงกันในทุก ระดับ โดย การจัดคุ ณภาพของสถานศึกษาของโรงเรียนที่มี ระดับการขาดแคลนครูคณิตศาสตร์ในสถานศึกษ าเป็นขาดแคลนจำนวนมาก , ขาดแคลนเป็นบางส่วน และไม่ขาดแคลน ทั้ง 4 โมเดลจัด ระดับการขาดแคลนครูคณิตศาสตร์ในสถานศึกษา ทั้ง 3 ระดับให้อยู่ในกลุ่มกลางเป็นส่วนใหญ่

5. การให้การสนับสนุนด้านทรัพยากรที่ใ ช้ในการจัดการเรียนการสอนคณิตศาสตร์ พบว่าโมเดล 1, 2 และ 4 มีจำนวนโรงเรียนใกล้เคียงกันในทุก ระดับ โดยการจัดระดับคุณภาพสถานศึกษาจำแนกโดย การให้การสนับสนุนด้านทรัพยากรที่ใช้ในการจัดการเรียนการสอนคณิตศาสตร์ พบว่า การจัดคุณภาพของสถานศึกษา ทั้ง 4 โมเดลจัดสถานศึกษาที่ ให้การสนับสนุนทุกระดับคือ ระดับต่ำ, ระดับปานกลาง และระดับสูง ให้อยู่ในกลุ่มกลางเป็นส่วนใหญ่

6. การรับรู้ถึงบรรยากาศ ที่ดีในที่ทำงาน การจัดระดับคุณภาพ สถานศึกษาจำแนกโดยการการรับรู้ถึงบรรยากาศ ที่ดีในที่ทำงาน พบว่าพบว่าโมเดล 1, 2 และ 4 มีจำนวนโรงเรียนใกล้เคียงกัน ในทุกระดับ โดย การจัดคุณภาพของสถานศึกษา ทั้ง 4 โมเดลจัดสถานศึกษาทุกระดับการรับรู้ถึงบรรยากาศ ที่ดีในที่ทำงาน ให้อยู่ในกลุ่มกลางเป็นส่วนใหญ่

7. ความพึงพอใจของครูที่มีต่อการจัดการภายในสถานศึกษา การจัดระดับคุณภาพสถานศึกษาจำแนกโดยความพึงพอใจของครูที่มีต่อการจัดการ ภายในสถานศึกษา แบ่งออกเป็น 3 ระดับ คือ ระดับต่ำ, ระดับปานกลาง และระดับสูง พบว่าโมเดล 1, 2 และ 4 มีจำนวนโรงเรียนใกล้เคียงกันในทุกระดับ โดยการจัดคุณภาพของสถานศึกษา ทั้ง 4 โมเดลจัดสถานศึกษาทุกระดับให้อยู่ในกลุ่มกลางเป็นส่วน ใหญ่

8. ดัชนีการให้ความสำคัญกับการให้กา รบ้านคณิตศาสตร์ของครูในสถานศึกษา การจัดระดับคุณภาพ สถานศึกษาจำแนกโดย การให้ความสำคัญกับการให้การบ้านคณิตศาสตร์ของครูในสถานศึกษา แบ่งออกเป็น 3 ระดับ คือ ระดับต่ำ, ระดับปานกลาง และระดับสูง พบว่าโมเดล 1, 2 และ 4 มีจำนวนโรงเรียนใกล้เคียงกันในทุกระดับโดย การจัดคุณภา พของสถานศึกษา ทั้ง 4 โมเดลจัดสถานศึกษาทุกระดับให้อยู่ในกลุ่มกลางเป็นส่วนใหญ่

9. ระดับความถี่ในการทำงานร่วมกันของครูคณิตศาสตร์ในสถานศึกษา การจัดระดับคุณภาพ สถานศึกษาจำแนกโดยระดับความถี่ในการทำงานร่วมกันของครูคณิตศาสตร์ในสถานศึกษา แบ่งออกเป็น 3 ระดับ คือ ไม่เคยเลย, 2-3 ครั้งต่อเดือน และ เป็นประจำ พบว่าโมเดล 1, 2 และ 4 มีจำนวนโรงเรียนใกล้เคียงกันในทุกระดับโดย การจัดคุณภาพของสถานศึกษา ทั้ง 4 โมเดลจัดสถานศึกษาทุกระดับให้อยู่ในกลุ่มกลางเป็นส่วนใหญ่

รายละเอียดดังตารางที่ 39

ตารางที่ 39 ผลการจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาจากผลการวิเคราะห์ 4 โมเดล จำแนกตามคุณลักษณะของสถานศึกษา

คุณลักษณะสถานศึกษา	รวม (150)	กลุ่มต่ำ (N = 37) (Percentile ที่ 0 – 24.99)				กลุ่มกลาง (N = 76) (Percentile ที่ 25 – 74.99)				กลุ่มสูง (N = 37) (Percentile ที่ 75 ขึ้นไป)			
		M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4
1.ขนาดสถานศึกษา													
1.1 ขนาดเล็ก	43	9	9	17	9	27	26	22	26	7	8	4	8
	(100)	(20.93)	(20.93)	(39.53)	(20.93)	(62.79)	(60.47)	(51.16)	(60.47)	(16.28)	(18.60)	(9.30)	(18.60)
1.2 ขนาดกลาง	56	13	13	18	13	29	30	28	30	14	13	10	13
	(100)	(23.21)	(23.21)	(32.14)	(23.21)	(51.79)	(53.57)	(50.00)	(53.57)	(25.00)	(23.21)	(17.86)	(23.21)
1.3 ขนาดใหญ่	23	6	6	2	6	11	11	15	11	6	6	6	6
	(100)	(26.09)	(26.09)	(8.70)	(26.09)	(47.83)	(47.83)	(65.22)	(47.83)	(26.09)	(26.09)	(26.09)	(26.09)
1.4 ขนาดใหญ่พิเศษ	28	9	9	0	9	9	9	11	9	10	10	17	10
	(100)	(32.14)	(32.14)		(32.14)	(32.14)	(32.14)	(39.29)	(32.14)	(35.71)	(35.71)	(60.71)	(35.71)
2.ตำแหน่งที่ตั้งสถานศึกษา													
2.1 หมู่บ้าน	17	3	3	10	3	11	11	5	11	3	3	2	3
	(100)	(17.65)	(17.65)	(58.82)	(17.65)	(64.71)	(64.71)	(29.41)	(64.71)	(17.65)	(17.65)	(11.76)	(17.65)
2.2 ตำบล	28	6	6	8	6	18	18	16	19	4	4	4	3
	(100)	(21.43)	(21.43)	(28.57)	(21.43)	(64.29)	(64.29)	(57.14)	(67.86)	(14.29)	(14.29)	(14.29)	(10.71)
2.3 เมืองเล็ก	41	9	9	11	9	20	20	23	20	12	12	7	12
	(100)	(21.95)	(21.95)	(26.83)	(21.95)	(48.78)	(48.78)	(56.10)	(48.78)	(29.27)	(29.27)	(17.07)	(29.27)
2.4 เมือง	29	8	8	3	8	14	14	14	13	7	7	12	8
	(100)	(27.59)	(27.59)	(10.34)	(27.59)	(48.28)	(48.28)	(48.28)	(44.83)	(24.14)	(24.14)	(41.38)	(27.59)
2.5 เมืองใหญ่	20	6	6	4	6	8	8	10	8	6	6	6	6
	(100)	(30)	(30)	(20)	(30)	(40)	(40)	(50)	(40)	(30)	(30)	(30)	(30)
2.6 เมืองใหญ่มาก	15	5	5	1	5	5	5	8	5	5	5	6	5
	(100)	(33.33)	(33.33)	(6.67)	(33.33)	(33.33)	(33.33)	(53.33)	(33.33)	(33.33)	(33.33)	(40.00)	(33.33)
3. ขนาดของห้องเรียน													
3.1นักเรียน1 – 24 คน	19	4	4	9	4	11	11	8	10	4	4	2	5
	(100)	(21.05)	(21.05)	(47.37)	(21.05)	(57.89)	(57.89)	(42.11)	(52.63)	(21.05)	(21.05)	(10.53)	(26.32)
3.2 นักเรียน	73	16	16	23	16	41	41	37	42	16	16	13	15
25 – 40 คน	(100)	(21.92)	(21.92)	(31.51)	(21.92)	(56.16)	(56.16)	(50.68)	(57.53)	(21.92)	(21.92)	(17.81)	(20.55)
3.3 นักเรียน	58	17	17	5	17	24	24	31	24	17	17	22	17
มากกว่า 41 คน	(100)	(29.31)	(29.31)	(8.62)	(29.31)	(41.38)	(41.38)	(53.45)	(41.38)	29.31)	(29.31)	(37.93)	(29.31)

ตารางที่ 39 ผลการจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาจากผลการวิเคราะห์ 4 โมเดล จำแนกตามคุณลักษณะของสถานศึกษา (ต่อ)

คุณลักษณะสถานศึกษา	**รวม** (150)	**กลุ่มต่ำ** (N = 37) (Percentile **ที่** 0 – 24.99)				**กลุ่มกลาง** (N = 76) (Percentile **ที่** 25 – 74.99)				**กลุ่มสูง** (N = 37) (Percentile **ที่** 75 **ขึ้นไป**)			
		M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4
4.ระดับการขาดแคลนครูคณิตศาสตร์ในสถานศึกษา													
4.1 ไม่ขาดแคลน	48	15	15	15	15	23	23	18	22	10	10	15	11
	(100)	(31.25)	(31.25)	(31.25)	(31.25)	(47.92)	(47.92)	(37.50)	(45.83)	(20.83)	(20.83)	(31.25)	(22.92)
4.2 ขาดแคลนน้อย	6	1	1	3	1	4	4	2	4	1	1	1	1
	(100)	(16.67)	(16.67)	(50.00)	(16.67)	(66.67)	(66.67)	(33.33)	(66.67)	(16.67)	(16.67)	(16.67)	(16.67)
4.3 ขาดแคลนเป็นบางส่วน	49	12	12	6	12	22	22	30	22	15	15	13	15
	(100)	(24.49)	(24.49)	(12.24)	(24.49)	(44.90)	(44.90)	(61.22)	(44.90)	(30.61)	(30.61)	(26.53)	(30.61)
4.4 ขาดแคลนจำนวนมาก	47	9	9	13	9	27	27	26	28	11	11	8	10
	(100)	(19.15)	(19.15)	(27.66)	(19.15)	(57.45)	(57.45)	(55.32)	(59.57)	(23.40)	(23.40)	(17.02)	(21.28)
5.ดัชนีการให้การสนับสนุนด้านทรัพยากรที่ใช้ในการจัดการเรียนการสอนคณิตศาสตร์													
5.1 ระดับต่ำ	19	1	1	1	1	13	13	9	13	5	5	9	5
	(100)	(5.26)	(5.26)	(5.26)	(5.26)	(68.42)	(68.42)	(47.37)	(68.42)	(26.32)	(26.32)	(47.37)	(26.32)
5.2 ระดับปานกลาง	100	28	28	29	28	50	51	52	51	22	21	19	21
	(100)	(28.00)	(28.00)	(29.00)	(28.00)	(50.00)	(51.00)	(52.00)	(51.00)	(22.00)	(21.00)	(19.00)	(21.00)
5.3 ระดับสูง	31	8	8	7	8	13	12	15	12	10	11	9	11
	(100)	(25.81)	(25.81)	(22.58)	(25.81)	(41.94)	(38.71)	(48.39)	(38.71)	(32.26)	(35.48)	(29.03)	(35.48)
6.ดัชนีการรับรู้ถึงบรรยากาศที่ดีในที่ทำงาน													
6.1 ระดับต่ำ	27	4	4	6	4	15	15	12	14	8	8	9	9
	(100)	(14.81)	(14.81)	(22.22)	(14.81)	(55.56)	(55.56)	(44.44)	(51.85)	(29.63)	(29.63)	(33.33)	(33.33)
6.2 ระดับปานกลาง	103	28	28	29	28	50	50	52	51	25	25	22	24
	(100)	(27.18)	(27.18)	(28.16)	(27.18)	(48.54)	(48.54)	(50.49)	(49.51)	(24.27)	(24.27)	(21.36)	(23.30)
6.3 ระดับสูง	20	5	5	2	5	11	11	12	11	4	4	6	4
	(100)	(25.00)	(25.00)	(10.00)	(25.00)	(55.00)	(55.00)	(60.00)	(55.00)	(20.00)	(20.00)	(30.00)	(20.00)

ตารางที่ 39 ผลการจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาจากผลการวิเคราะห์ 4 โมเดล จำแนกตามคุณลักษณะของสถานศึกษา (ต่อ)

คุณลักษณะสถานศึกษา	รวม (150)	กลุ่มต่ำ (N = 37) (Percentile ที่ 0 – 24.99)				กลุ่มกลาง (N = 76) (Percentile ที่ 25 – 74.99)				กลุ่มสูง (N = 37) (Percentile ที่ 75 ขึ้นไป)			
		M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4
7. ดัชนีความพึงพอใจของครูที่มีต่อการจัดการภายในสถานศึกษา													
7.1 ระดับต่ำ	24	7	7	6	7	11	11	8	11	6	6	10	6
	(100)	(29.17)	(29.17)	(25.00	(29.17)	(45.83)	(45.83)	(33.33)	(45.83)	(25.00)	(25.00)	(41.67)	(25.00)
7.2 ระดับปานกลาง	102	23	23	23	23	54	54	56	55	25	25	23	24
	(100)	(22.55)	(22.55)	(22.55)	(22.55)	(52.94)	(52.94)	(54.90)	(53.92)	(24.51)	(24.51)	(22.55)	(23.53)
7.3 ระดับสูง	24	7	7	8	7	11	11	12	10	6	6	4	7
	(100)	(29.17)	(29.17)	(33.33)	(29.17)	(45.83)	(45.83)	(50.00)	(41.67)	(25.00)	(25.00)	(16.67)	(29.17)
8. ดัชนีการให้ความสำคัญกับการให้การบ้านคณิตศาสตร์ของครูในสถานศึกษา													
8.1 ระดับต่ำ	15	0	0	4	0	10	11	8	11	5	4	3	4
	(100)			(26.67)		(66.67)	(73.33)	(53.33)	(73.33)	(33.33)	(26.67)	(20.00)	(26.67)
8.2 ระดับปานกลาง	71	26	26	23	26	29	29	33	29	16	16	15	16
	(100)	(36.62)	(36.62)	(32.39)	(36.62)	(40.85)	(40.85)	(46.48)	(40.85)	(22.54)	(22.54)	(21.13)	(22.54)
8.3 ระดับสูง	64	11	11	10	11	37	36	35	36	16	17	19	17
	(100)	(17.19)	(17.19)	(15.63)	(17.19)	(57.81)	(56.25)	(54.69)	(56.25)	(25.00)	(26.56)	(29.69)	(26.56)
9. ระดับความถี่ในการทำงานร่วมกันของครูคณิตศาสตร์ในสถานศึกษา													
9.1 ไม่เคยเลย	8	0	0	3	0	6	6	4	6	2	2	1	2
	(100)			(37.50)		(75.00)	(75.00)	(50.00)	(75.00)	(25.00)	(25.00)	(12.50)	(25.00)
9.2 2-3 ครั้งต่อเดือน	83	20	20	19	20	44	45	45	45	19	18	19	18
	(100)	(24.10)	(24.10)	(22.89)	(24.10)	(53.01)	(54.22)	(54.22)	(54.22)	(22.89)	(21.69)	(22.89)	(21.69)
9.3 เป็นประจำ	59	17	17	15	17	26	25	27	25	16	17	17	17
	(100)	(28.81)	(28.81)	(25.42)	(28.81)	(44.07)	(42.37)	(45.76)	(42.37)	(27.12)	(28.81)	(28.81)	(28.81)

หมายเหตุ ตัวเลขในวงเล็บแสดงจานวนร้อยละ

4.2 ผลการเปรียบเทียบประสิทธิภาพโมเดลการประเมินคุณภาพการจัดการศึกษาต่อคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์

การพิจารณาเพื่อสรุปผลการ เปรียบเทียบ ประสิทธิภาพ โมเดลการประเมินคุณภาพการจัดการศึกษาทั้ง 4 โมเดลต่อคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ของสถานศึกษา ได้ผลดังนี้

การเปรียบเทียบ ประสิทธิภาพ ผลการวิเคราะห์จากโมเดลก ารประเมินคุณภาพการจัดการศึกษา 4 โมเดล พบว่าความแปรปรวนของคะแนนผลการประเมินวิชาคณิตศาสตร์ระหว่างสถานศึกษาในโมเดล 3 ซึ่งเป็นโมเดลที่มีการตัดข้อสอบที่ทำ หน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง ออกไป แต่ไม่ได้ ควบคุมอิทธิพลของตัวแปรทำ นายด้วยตัวแปรระดับนักเรียนและสถานศึกษา มีค่าความแปรปรวนระหว่างสถานศึกษา (ระดับที่ 2) เท่ากับ 0.29852 อย่างมีนัยสำคัญทางสถิติ หรือคิดเป็นความแปรปรวนภายในชั้น (Intraclass correlation) เท่ากับ 0.4475 (44.75%) นั่นคือมีความผันแปรของคะแนนผลการประเมิน คณิตศาสตร์ระหว่า งสถานศึกษา 44.75% และเมื่อเปรียบเทียบกับ โมเดลที่ 4 ซึ่งเป็นโมเดลที่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวงออกไปเหมือนโมเดลที่ 3 แต่แตกต่างตรงมีการ ควบคุมอิทธิพลของตัวแปรทำ นายด้วยตัวแปรระดับนักเรียนและสถานศึกษา พบว่าความแปรปรวนของคะแนนผลการประเมินวิชา คณิตศาสตร์ มีค่าลดลง โดยมีค่าความแปรปรวนระหว่างสถานศึกษา (ระดับที่ 2) เท่ากับ 0.16232 อย่างมีนัยสำ คัญทางสถิติหรือคิดเป็นความแปรปรวนภายในชั้นเท่ากับ 0.3199 (31.99%) นั่นคือมีความผันแปรของคะแนนผลการประเมิน คณิตศา สตร์ระหว่าง สถานศึกษา 31.99% สัมประสิทธิ์การทำ นาย (R^2) ในระดับที่ 1 เท่ากับ 6.41% นั่นคือในระดับที่ 1 สามารถอธิบายความแปรปรวนของตัวแปรตามได้ 6.41% ของความแปรปรวนภายในชั้น (Intraclass correlation) ที่มีอยู่ในระดับที่ 1 (68.01%) ในขณะที่ระดับที่ 2 มีสัมประสิทธิ์การทำ นาย (R^2) เท่ากับ 45.63% นั่นคือในระดับที่ 2 สามารถอธิบายความแปรปรวนของตัวแปรตามได้ 45.63% ของความแปรปรวนภายในชั้น (Intraclass correlation) ที่มีอยู่ในระดับที่ 2 (31.99%) และสัมประสิทธิ์การทำ นายโดยรวมของโมเดล 4 เท่ากับ 0.5203 (52.03%) นั่นคือโมเดล 4 (ทั้งส องระดับ) สามารถอธิบายความแปรปรวนได้ 52.03%

เมื่อเปรียบเทียบ โมเดลที่มีการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา ระหว่างโมเดล 1 ซึ่งเป็นโมเดลที่ไม่มีการตัดข้อสอบที่เกิดการทำหน้าที่ต่างกันออกเลย โมเดล 2 ที่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันเฉพาะที่เ ป็นตัวเลือกที่เป็นตัวถูก ออกไป และโมเดล 4 ที่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูก และตัวเลือกที่เป็นตัวลวงออกไป พบว่าโมเดล 1 มีค่าความแปรปรวนระหว่างสถานศึกษาเท่ากับ 0.17316 หรือคิดเป็น

ความแปรปรวนภายในชั้นเท่ากับ 0.3405 (34.05%) โมเดล 2 มีค่าความแปรปรวนระหว่างสถานศึกษาเท่ากับ 0.16014 หรือคิดเป็นความแปรปรวนภายในชั้นเท่ากับ 0.3163 (31.63%) และโมเดล 4 มีค่าความแปรปรวนระหว่างสถานศึกษาเท่ากับ 0.16232 หรือคิดเป็นความแปรปรวนภายในชั้นเท่ากับ 0.3199 (31.99%) และเมื่อเปรียบเทียบสัมประสิทธิ์การทำนายระหว่างโมเดลที่มีการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา พบว่าโมเดล 1 มีค่าสัมประสิทธิ์การทำนาย (R^2) มีค่าเท่ากับ 0.51962 (51.96%) โมเดล 2 มีค่าสัมประสิทธิ์การทำนาย (R^2) มีค่าเท่ากับ 0.5186 (51.86%) และโมเดล 4 มีค่าสัมประสิทธิ์การทำนาย (R^2) มีค่าเท่ากับ 0.5203 (52.03%) แสดงรายละเอียด ดังตารางที่ 40

ตารางที่ 40 สรุปค่าความแปรปรวนของตัวแปรตามที่อธิบายได้ของโมเดลที่ 1 ถึง 4

variance components	Model 1	Model 2	Model 3	Model 4
ระดับที่ 1 ความแปรปรวนภายในสถานศึกษา(R)	0.33539	0.34614	0.36862	0.34501
ระดับที่ 2 ความแปรปรวนระหว่างสถานศึกษา(U0)	0.17316	0.16014	0.29852	0.16232
ระดับความผันแปรของ ค่าความสามารถของนักเรียนจากการประเมินความสามารถวิชาคณิตศาสตร์ที่				
ระดับภายในสถานศึกษา (ระดับที่ 1)	0.6595 (65.95%)	0.6837 (68.37%)	0.5525 (55.25%)	0.6801 (68.01%)
ระดับระหว่างสถานศึกษา (ระดับที่ 2)	0.3405 (34.05%)	0.3163 (31.63%	0.4475 (44.75%)	0.3199 (31.99%)
สัดส่วนของความแปรปรวนทั้งหมดของตัวแปรตามที่อธิบายได้ (R^2)				
ระดับที่ 1	0.07040 (7.04%)	0.06484 (6.48%)	NA	0.06405 (6.41%)
ระดับที่ 2	0.44922 (44.92%)	0.45376 (45.38%)	NA	0.45625 (45.63%)
รวมทั้งสองระดับ	0.51962 (51.96%)	0.51860 (51.86%)	NA	0.52030 (52.03%)

ผลการวิเคราะห์จากตารางที่ 40 หากพิจารณาความสามารถในการอธิบายความผันแปรของค่าความสามารถของนักเรียนที่ได้จากการสอบประเมินผลในวิชาคณิตศาสตร์แล้วโมเดล 1, 2

และ4 เป็นโมเดลที่มีสามารถอธิบายความผันแปรได้ ใกล้เคียงกันมาก ทั้ง 3 โมเดลกำจัดแหล่งความคลาดเคลื่อน จากอิทธิพลของคุณลักษ ณะของนักเรียนและสถานศึกษาเพื่อทำ ให้อยู่บนพื้นฐานเดียวกัน จึงสามารถกล่าวได้ว่าทั้ง 3 โมเดล เป็นโมเดลที่ให้ผลการประเมินที่ถูกต้ องและมีความเป็นธรรม แต่หากพิจารณาถึงค่าสัมประสิทธิ์การทำ นาย (R^2) ก็จะพบว่า ถึงแม้ทั้ง 3 โมเดลจะมีความใกล้เคียงกัน แต่โมเดลที่ 4 ก็มีค่าค่าสัมประสิทธิ์การทำนาย (R^2) สูงที่สุด คือ 52.03% ทั้งนี้ โมเดล 4 นอกจากจะมีการกำ จัดแหล่งความคลาดเคลื่อน จากอิทธิพลของคุณ ลักษณะของนักเรียนและสถานศึกษา แล้วยังมีการกำ จัดแหล่งความคลาดเคลื่อน จากอิทธิพลของข้อสอบที่ทำ หน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง จึงอาจกล่าวได้ว่าโมเดล 4 เป็นโมเดลที่มีประสิทธิภาพดีที่สุด โดยมีแนวโน้มให้ผลการประเมินที่ถูกต้องและเป็นธรรมที่สุด

ผลการเปรียบเทียบคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อเทียบกับโมเดล 4

ผลการวิเคราะห์ความสามารถในการอธิบายความผันแปรของคะแนนสเกล ของ ค่าความสามารถของนักเรียนที่ได้จากการสอบ ประเมินผลในวิชาคณิตศาสตร์ แสดงให้เห็นว่า โมเดล 1, 2 และ 4 เป็นโมเดลที่สามารถอธิบายความผันแปรของตัวแปรตามได้ ใกล้เคียงกัน แต่โมเดล 4 มีค่า**ค่าสัมประสิทธิ์การทำนาย** (R^2) **สูง**ที่สุด จึงสรุปได้ว่าโมเดล 4 เป็นโมเดลที่ให้ผลการประเมินที่ถูกต้องและมีความเป็นธรรมมากที่สุด เนื่องจากเป็นโมเดลที่มีการกำจัดแหล่งความคลาดเคลื่อนจากอิทธิพลของข้อ สอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง และอิทธิพลของคุณลักษ ณะของนักเรียนและสถานศึกษา ทำให้อยู่บนพื้นฐานเดียวกัน ผู้วิจัยจึงเสนอผลการ เปรียบเทียบ คุณภาพ โมเดลการประเมิน การจัดการศึกษา วิชาคณิตศาสตร์ โดยการเปรียบเทียบโมเดล 1, 2 และ 3 กับโมเดล 4 พบว่าโมเดลให้ผลการจัด ระดับคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ ได้สอดคล้องกัน หรือจัดสถานศึกษาให้อยู่ใน ระดับคุณภาพ การจัดการศึกษาได้ตรงกัน แสดงรายละเอียดดังตารางที่ 41

ตารางที่41 แสดงผลการจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาโดยเปรียบเทียบกับโมเดล4

ระดับคุณภาพการจัดการศึกษา		โมเดล 1			โมเดล 2			โมเดล 3			รวม
		ต่ำ	กลาง	สูง	ต่ำ	กลาง	สูง	ต่ำ	กลาง	สูง	
โมเดล 4	ต่ำ	37 (24.67%)	-	-	37 (24.67%)	-	-	21 (14%)	16 (10.67%)	-	37 (24.67%)
	กลาง	-	75 (50%)	1 (0.67%)	-	75 (50%)	1 (0.67%)	16 (10.67%)	50 (33.33%)	10 (6.67%)	76 (50.67%)
	สูง	-	1 (0.67%)	36 (24%)	-	1 (0.67%)	36 (24%)	-	10 (6.67%)	27 (18%)	37 (24.67%)
	รวม	37 (24.67%)	76 (50.67%)	37 (24.67%)	37 (24.67%)	76 (50.67%)	37 (24.67%)	37 (100%)	76 (50.67%)	37 (24.67%)	150 (100%)

การประเมินคุณภาพการจัดการศึกษาของสถานศึกษา ด้วยโมเดลการประเมินคุณภาพการจัดการศึกษาที่ต่างกัน ได้ผลการประเมินคุณภาพการจัดการศึกษา สามารถสรุปได้ดังนี้ โมเดล 1 โมเดล 2 และโมเดล 4 ซึ่งต่างก็เป็นโมเดลที่มีการควบคุมอิทธิพลของตัวแปรในระดับนักเรียนและระดับสถานศึกษาผลการเปรียบกันรายคู่ระหว่างโมเดล 1 กับ โมเดล 2, โมเดล 1 กับ โมเดล 4 และ โมเดล 2 กับ โมเดล 4 พบว่าต่างก็ให้ผลการประเมินคุณภาพการจัดการศึกษาที่สอดคล้องกันมากถึง 98.67% ในขณะที่การเปรียบเทียบระหว่างโมเดล 1 กับ โมเดล 3, โมเดล 2 กับ โมเดล 3 และ โมเดล 3 กับ โมเดล 4 ซึ่งเป็นการเปรียบเทียบระหว่างโมเดล 3 ซึ่งไม่มีการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา กับโมเดลอื่นๆที่มีการควบคุม ให้ผลการประเมินคุณภาพการจัดการศึกษาที่ต่างกันหรือไม่สอดคล้องกัน ค่อนข้างมาก โดยโมเดล 1 กับ โมเดล 3 ให้ผลการประเมินคุณภาพการจัดการศึกษาไม่สอดคล้องกันมากถึง 36% ส่วนโมเดล 2 กับ โมเดล 3 ให้ผลการประเมินคุณภาพการจัดการศึกษาไม่สอดคล้องกัน 36% และโมเดล 3 กับ โมเดล 4 ให้ผลการประเมินคุณภาพการจัดการศึกษาไม่สอดคล้องกัน 34.67% แสดงรายละเอียดดังตารางที่ 42

ตารางที่ 42 ผลการวิเคราะห์ความสอดคล้องของการจัดระดับคุณภาพการจัดการศึกษาด้วยโมเดลการประเมินคุณภาพการจัดการศึกษาที่ต่างกัน 4 โมเดล

โมเดลคู่เปรียบเทียบ	ประเด็นเปรียบเทียบของโมเดล					จำนวนสถานศึกษา		รวม
	ไม่ตัด DIF	ตัด DIF	ตัด DIF+DDF	ไม่ควบคุมตัวแปร	ควบคุมตัวแปร	สอดคล้อง	ไม่สอดคล้อง	
โมเดล1–โมเดล2	✓	✓		ควบคุมตัวแปร		148 (98.67%)	2 (1.33%)	150 (100%)
โมเดล1–โมเดล4	✓		✓	ควบคุมตัวแปร		148 (98.67%)	2 (1.33%)	150 (100%)
โมเดล2–โมเดล4		✓	✓	ควบคุมตัวแปร		148 (98.67%)	2 (1.33%)	150 (100%)
โมเดล1–โมเดล3	✓		✓	✓	✓	96 (64%)	54 (36%)	150 (100%)
โมเดล2–โมเดล3		✓	✓	✓	✓	96 (64%)	54 (36%)	150 (100%)
โมเดล3–โมเดล4	ตัด DIF+DDF			✓	✓	98 (65.33%)	52 (34.67%)	150 (100%)

จากข้อค้นพบดังกล่าว แสดงให้เห็นประเด็นเกี่ยวกับการพิจารณาโมเดลการประเมินคุณภาพการจัดการศึกษา 2 ประเด็น คือ

ประเด็นที่ 1 การมีข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบ การเปรียบเทียบระหว่างโมเดล 1 กับ โมเดล 2, โมเดล 1 กับ โมเดล 4 และ โมเดล 2 กับ โมเดล 4 พบว่าการประเมินคุณภาพการจัดการศึกษาด้วยโมเดลที่มีความแตกต่างกันเกี่ยวการควบคุมอิทธิพล แหล่งความคลาดเคลื่อนที่เกิดขึ้นจากการโอนเอียงให้ ผู้สอบกลุ่มใดกลุ่มหนึ่ง นั่นคือโมเดล 1 เป็นโมเดลที่ไม่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป โมเดล 2 เป็นโมเดลที่ตัดเฉพาะข้อสอบที่เกิดการทำหน้าที่ต่างกันของตัวเลือกที่เป็นตัวถูก (DIF) เท่านั้น และโมเดล 4 เป็นโมเดลที่มีการตัดข้อสอบที่เกิดการทำหน้าที่ต่างกั นของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) ในข้อเดียวกัน โดยทั้ง 3 โมเดลมีการควบคุมอิทธิพลของตัวแปรในระดับนักเรียนและระดับสถานศึกษา ซึ่งผลที่ได้จากการเปรียบเทียบ ให้ผลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาที่สอดคล้องหรือให้ผลการประเมินที่ตรง กัน 98.67% แต่ถ้าต้องการให้ผลการประเมินตรงกันมากขึ้น ต้องมีการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา 100.0%

ประเด็นที่ 2 การควบคุมอิทธิพลของคุณลักษณะนักเรียนและสถานศึกษา การเปรียบเทียบระหว่างโมเดล 1 กับ โมเดล 3, โมเดล 2 กับ โมเดล 3 และ โมเดล 3 กับ โมเดล 4 ซึ่งเป็นการเปรียบเทียบระหว่างโมเดล 3 ซึ่งไม่มีการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา กับโมเดลอื่นๆที่มีการควบคุม พบว่าการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาด้วยโมเดลที่ไม่มีการควบคุมตัวแปรระดับนักเรียนและสถานศึกษา กับโมเดลที่มีการควบคุมตัวแปรระดับนักเรียนและสถานศึกษา จะให้ผลการประเมินคุณภาพการจัดการศึกษา ที่สอดคล้องกันน้อยกว่าในประเด็นที่ 1 มาก คือโมเดล 1 กับ โมเดล 3 ให้ผลการประเมินคุณภาพการจัดการศึกษาไม่สอดคล้องกันมากถึง 36% ส่วนโมเดล 2 กับ โมเดล 3 ให้ผลการประเมินคุณภาพการจัดการศึกษาไม่สอดคล้องกัน 36% และโมเดล 3 กับ โมเดล 4 ให้ผลการประเมินคุณภาพการจัดการศึกษาไม่สอดคล้องกัน 34.67%

เมื่อพิจารณาความสัมพันธ์ของการจัดตำแหน่งคะแนนมูลค่าเพิ่มจากการวิเคราะห์ 4 โมเดล ด้วยการตรวจสอบสหสัมพันธ์ตำแหน่งคะแนนของสเปียร์แมน (Spearman's Rank – Order Correlation) พบว่า กลุ่มของโมเดลที่มีการควบคุมอิทธิพลตัวแปรระดับนักเรียนและระดับสถานศึกษามีค่าสหสัมพันธ์ที่ค่อนข้างสูงนั่นคือ โมเดล 1 กับโมเดล 2 มีค่าสหสัมพันธ์กันสูงสุดเท่ากับ .995 ส่วนโมเดล 1 กับโมเดล 4 เท่ากับ .992 และโมเดล 2 กับโมเดล 4 เท่ากับ .987 ในขณะที่เมื่อหาค่าสหสัมพันธ์ระหว่างโมเดลที่ 3 ซึ่งเป็นโมเดลเดียวที่ไม่มีการควบคุมอิทธิพลตัวแปรระดับนักเรียนและระดับสถานศึกษา กับโมเดลอื่นๆ พบว่ามีค่าสหสัมพันธ์น้อยลงกว่าการเปรียบเทียบเทียบในตอนแรก นั่นคือ สุดเป็น โมเดล 1 กับ โมเดล 3 เท่ากับ 0.721, โมเดล 2 กับ โมเดล 3 เท่ากับ .708 และโมเดล 3 กับ โมเดล 4 เท่ากับ .733 แสดงรายละเอียด ดังตารางที่ 43

ตารางที่ 43 ค่าสหสัมพันธ์ของอันดับของคะแนนมูลค่าเพิ่มที่ได้จากโมเดลประเมินคุณภาพการจัดการศึกษาที่ต่างกัน 4 โมเดล

โมเดล	ค่าสหสัมพันธ์			
	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
โมเดล 1	-	0.995	0.721	0.992
โมเดล 2	-	-	0.708	0.987
โมเดล 3	-	-	-	0.733
โมเดล 4	-	-	-	-

บทที่ 5

สรุปผลงานวิจัย อภิปรายผล และข้อเสนอแนะ

การวิเคราะห์เปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ : การประยุกต์ใช้โมเดลมูลค่าเพิ่มที่มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและการวิเคราะห์การทำหน้าที่ต่างกันของตัวลวง มีวัตถุประสงค์ในการดำเนินการวิจัย 3 ข้อ คือ 1. เพื่อศึกษาผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) ในแบบสอบประเมินความรู้วิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 2 2. เพื่อศึกษาผลของข้อสอบที่ทำหน้าที่ต่างกัน และผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน 4 โมเดล และ 3. เพื่อเปรียบเทียบประสิทธิภาพของโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษา 4 โมเดล โดยข้อมูลที่ใช้ในการวิจัยเป็นข้อมูลทุติยภูมิที่ได้จาก โครงการ การศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์ หรือ TIMSS ปี 2007

การประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ ของสถานศึกษาขั้นพื้นฐาน ประกอบด้วยโมเดลที่ใช้ในการวิจัย 4 โมเดล คือ โมเดล 1 “Undetected DIF-DDF & Adjusted” เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ไม่ตัดข้อสอบที่ทำหน้าที่ต่างกันออก แต่มีการวิเคราะห์มูลค่าเพิ่ม โดยมีการ ควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา (Adjusted) โมเดล 2 “Detect DIF & Adjusted” เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ตัวถูกออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการ ควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา (Adjusted) โมเดล 3 “Detect DIF-DDF & Unadjusted” เป็นโมเดลเป็นการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ทั้งตัวถูกและตัวลวงออก แต่ไม่ได้ทำการวิเคราะห์มูลค่าเพิ่ม (Unadjusted) และโมเดล 4 “Detect DIF-DDF & Adjusted” เป็นโมเดลเป็นการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ทั้งตัวถูกและตัวลวงออก และวิเคราะห์มูลค่าเพิ่ม โดยมีการ ควบคุมตัวแปรระดับนักเรียนและระดับสถานศึกษา (Adjusted)

การวิเคราะห์ข้อมูลเพื่อตรวจสอบการทำ หน้าที่ต่างกันของข้อสอบ (DIF)และและการทำหน้าที่ต่างกันของตัวลวง (DDF) ผู้วิจัยใช้วิธีการวิเคราะห์แมนเทล-แฮนส์เซลร่วมกับอัตราส่วนออด ด้วยโปรแกรม DDFS (Penfield, 2010) และทำการวิเคราะห์มูลค่าเพิ่ม โดยใช้โปรแกรมโมเดลเชิงเส้นระดับลดหลั่น (Hierarchical Linear Model: HLM) เวอร์ชั่น 6.03 โดยมี ระดับที่ 1 เป็นระดับนักเรียนและระดับที่ 2 เป็นระดับสถานศึกษา และคะแนนมูลค่าเพิ่มของสถานศึกษาหรือค่าส่วนที่เหลือที่ได้จากการวิเคราะห์พหุระดับจากการวิเคราะห์ในแต่ละโมเดลการประเมินคุณภาพการจัดการศึกษา จะใช้เป็นคะแนนเพื่อกำ หนดคุณภาพการจัดการศึกษาวิชา คณิตศาสตร์ของสถานศึกษาขั้นพื้นฐาน

5.1 สรุปผลการวิจัย

สรุปผลการวิจัย จำแนกตามวัตถุประสงค์การวิจัย ดังนี้

1. การศึกษาผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) ในแบบสอบประเมินความรู้วิชาคณิตศาสตร์

ผลการตรวจสอบ ข้อสอบที่เกิดการทำหน้าที่ต่างกัน โดยการจำแนกตามลักษณะการตรวจพบ นั่นคือ 1. ข้อสอบที่พบการทำหน้าที่ต่างกันของข้อสอบ (DIF) เท่านั้น 2. ข้อสอบที่พบการทำหน้าที่ต่างกันของตัวลวง (DDF) เท่านั้น และ 3. ข้อสอบที่พบทั้งการทำหน้าที่ต่างกันของข้อสอบและการทำหน้าที่ต่างกันของตัวลวง (DIF-DDF) ในแบบสอบทั้ง 14 ฉบับ ที่เกิดจากคุณลักษณะของนักเรียนที่ต่างกลุ่มกันในตัวแปรความเป็นเพศ (หญิง, ชาย) การใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ (สูง, ต่ำ) ระดับการศึกษาของผู้ปกครอง(สูง, ต่ำ) แหล่งทรัพยากรที่บ้าน (สูง, ต่ำ) และความมั่งคั่งของครอบครัว (สูง, ต่ำ) ภาพรวมพบว่าสามารถ ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันในทุกลักษณะ ในแบบสอบ แต่ละฉบับ คิดเป็นร้อยละ 31.25 ถึง 64.29 และเมื่อจำแนกรายกรณีพบว่าแบบสอบแต่ละฉบับตรวจพบการทำหน้าที่ต่างกันของ ข้อสอบ (DIF) เท่านั้น คิดเป็นร้อยละ 6.25 ถึง 25 ในขณะที่แบบสอบแต่ละฉบับ ตรวจพบการทำหน้าที่ต่างกันของ ตัวลวง (DDF) เท่านั้น คิดเป็นร้อยละ 6.06 ถึง 22.58 และแบบสอบแต่ละฉบับการตรวจพบการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) ในข้อเดียวกัน คิดเป็นร้อยละ 6.25 ถึง 34.48

โดยแบบสอบที่มีข้อสอบทำหน้าที่ต่างกันรวมทุกลักษณะมากที่สุดคือ แบบสอบฉบับที่ 12 คิดเป็นร้อยละ 64.29 (DIF 25%, DDF 17.86% และ DIF-DDF 21.43%) รองลงมาคือแบบสอบฉบับที่ 1 คิดเป็นร้อยละ 58.62 (DIF 13.79%, DDF 10.34% และ DIF-DDF 34.48%) และแบบสอบฉบับที่ 11 คิดเป็นร้อยละ 57.38 (DIF 21.88%, DDF 12.50% และ DIF-DDF 25%) ส่วนแบบสอบที่มีข้อสอบทำหน้าที่ต่างกันรว มทุกคุณลักษณะน้อยที่สุดคือ แบบสอบฉบับที่ 6 คิดเป็นร้อยละ 31.25 (DIF 6.25%, DDF 21.88% และ DIF-DDF 6.25%) รองลงมาคือแบบสอบฉบับที่ 13 คิดเป็นร้อยละ35.71 (DIF 7.14%, DDF 14.29% และ DIF-DDF 14.29%) และแบบสอบฉบับที่ 7 คิดเป็นร้อยละ 36.36 (DIF 15.15%, DDF 6.06% และ DIF-DDF 15.15%)

เมื่อพิจารณาคุณลักษณะของตัวแปรที่มีอิทธิพลทำให้เกิดการทำหน้าที่ต่างกันของข้อสอบ โดยจากผลการวิเคราะห์การทำ หน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) ในแต่ละฉบับผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) โดยภาพรวมพบว่าข้อสอบที่ทำหน้าที่ต่างกันจะพบในตัวแปรเพศมากที่สุด (DIF 57 ข้อ, DDF 65 ข้อ) รองลงมาคือการศึกษาของผู้ปกครอง (DIF 37 ข้อ, DDF 39 ข้อ) และความมั่งคั่งของครอบครัว (DIF 31 ข้อ, DDF 36 ข้อ) ส่วนที่พบข้อสอบที่ทำหน้าที่ต่างกันน้อยที่สุดคือในตัวแปรแหล่งทรัพยากรที่บ้าน (DIF 26 ข้อ, DDF 24 ข้อ) รองลงมาคือการใช้เวลาในการทำการบ้าน (DIF 28 ข้อ, DDF 35 ข้อ) โดยผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง จำแนกตามตัวแปรแต่ละตัว นอกจากนี้ยังพบว่าการเกิดการทำหน้าที่ต่างกันของข้อสอบ (DIF) ส่วนใหญ่กลุ่มอ้างอิง (Reference group) ซึ่งเป็นกลุ่มที่คาดว่าน่าจะได้เปรียบในการสอบครั้งนี้ มีโอกาสในการที่จะตอบถูกมากกว่ากลุ่มเปรียบเทียบ (Focal group) โดยมีเพียงตัวแปรความมั่งคั่งตัวเดียวเท่านั้นที่พบว่าจำนวนข้อสอบที่เ กิดการทำหน้าที่ต่างกัน (DIF) มีจำนวนข้อเท่ากันทั้ง 2 กลุ่ม ส่วนการทำหน้าที่ต่างกันของตัวลวง (DDF) จากผลการวิเคราะห์ข้อมูลพบว่ามีความสอดคล้องกับการทำหน้าที่ต่างกันของข้อสอบ นั่นคือส่วนใหญ่กลุ่ม เปรียบเทียบ (Focal group) มีโอกาสที่จะเลือกตอบตัวเลือกที่เป็นตัวลวงมากกว่ากลุ่มอ้างอิง (Reference group) โดยมีเพียงตัวแปรแหล่งทรัพยากรที่บ้านเท่านั้นที่มีโอกาสเท่ากัน โดยจำแนกรายละเอียดตัวแปรแต่ละตัวได้ ดังนี้

สำหรับ ตัวแปร เพศ พบว่า นักเรียน หญิง มีโอกาสทำ ข้อสอบได้ถูกมากกว่านักเรียน ชาย และเมื่อพิจารณาลักษณะของเนื้อหาสำหรับข้อสอบเรื่องพีชคณิต นักเรียนหญิงมีโอกาสทำข้อสอบได้มากกว่า ในขณะที่เรื่องเรขาคณิตนักเรียนชายมีโอกาสตอบถูกมากกว่า ส่วนข้อสอบที่เกิดการทำหน้าที่ต่างกันของตัวลวง (DDF) นั้นพบว่าเพศชายมีโอกาสในการที่จะเลือกตัวลวงได้มากกว่าเพศหญิงเล็กน้อย และเมื่อพิจารณาลักษณะข องเนื้อหาพบว่าเป็นข้อสอบ เรื่องจำนวน, พีชคณิต, และข้อมูลและโอกาส เพศชายมีโอกาสที่จะถูกลวงได้มากกว่าเพศหญิง ในขณะที่เรื่องเรขาคณิต เพศหญิงมีโอกาสที่จะถูกลวงได้มากกว่าเพศชาย

ตัวแปรการใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ พบว่าการทำหน้าที่ต่างกันของข้อสอบ (DIF) นั้นนักเรียนที่มีใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ สูงมีโอกาสทำ ข้อสอบได้ถูกมากกว่านักเรียนที่ใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ ต่ำเป็นส่วนใหญ่ โดยเฉพาะ เมื่อพิจารณาลักษณะของ เนื้อหาที่พบอย่างชัดเจนคือ ข้อสอบ เรื่องจำนวนและพีชคณิต นักเรียนที่ ใช้เวลาในการทำกา รบ้านวิชาคณิตศาสตร์ สูงมีโอกาสทำข้อสอบได้มากกว่า ส่วนข้อสอบที่เกิดการทำหน้าที่ต่างกันของตัวลวง (DDF) นั้นพบว่านักเรียนที่ ใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ ต่ำ มีโอกาสในการที่จะเลือกตัวลวงได้มากกว่านักเรียนที่ ใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ สูง และเมื่อพิจารณาลักษณะของ เนื้อหาพบว่าข้อสอบ เรื่องจำนวน และพีชคณิต นักเรียนที่ใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ ต่ำ มีโอกาสที่จะถูกลวงได้มากกว่าอีกกลุ่ม

ตัวแปรระดับการศึกษาของผู้ปกครองพบว่านักเรียนที่มี ระดับการศึกษาของผู้ปกครองสูงมีโอกาสทำข้อสอบได้ถูกมากกว่านักเรีย นที่มี ระดับการศึกษาของผู้ปกครอง ต่ำเล็กน้อย และเมื่อพิจารณาลักษณะของเนื้อหา พบว่าข้อสอบเรื่องพีชคณิตและเรขาคณิตนักเรียนที่มี ระดับการศึกษาของผู้ปกครองสูงมีโอกาสตอบถูกมากกว่า ในขณะที่เรื่องข้อมูลและโอกาสนักเรียนที่มีระดับการศึกษาของผู้ปกครองต่ำมีโอกาสตอบถูกมา กกว่า ส่วนข้อสอบที่เกิดการทำหน้าที่ต่างกันของตัวลวง (DDF) นั้นพบว่านักเรียนที่มี ระดับการศึกษาของผู้ปกครองต่ำ มีโอกาสในการที่จะเลือกตัวลวงได้มากกว่านักเรียนที่มี ระดับการศึกษาของผู้ปกครอง สูง และเมื่อพิจารณาลักษณะของเนื้อหาพบว่าข้อสอบเรื่องจำนวน , เรขาคณิต และ ข้อมูลและโอกาส นักเรียนที่มี ระดับการศึกษาของผู้ปกครองต่ำ มีโอกาสที่จะถูกลวงได้มากกว่าอีกกลุ่มหนึ่ง ในขณะที่เรื่องพีชคณิตนักเรียนที่มีระดับการศึกษาของผู้ปกครองสูงมีโอกาสที่จะถูกลวงได้มากกว่ากลุ่มตรงข้าม

ตัวแปรแหล่งทรัพยากรที่บ้าน พบว่านักเรียนที่มีแหล่งทรั พยากรที่บ้านที่บ้านสูงมีโอกาสในการทำข้อสอบถูกต้องมากกว่านักเรียนที่มีแหล่งทรัพยากรที่บ้านต่ำ และเมื่อพิจารณาลักษณะของเนื้อหาพบว่า เรื่องพีชคณิต กับข้อมูลและโอกาสนักเรียนที่มีแหล่งทรัพยากรที่บ้านที่บ้านสูงมีโอกาสในการทำข้อสอบถูกต้องมากกว่านักเรียนที่มีแหล่งท รัพยากรที่บ้านต่ำ ส่วนเรื่องจำนวน และเรขาคณิตมีโอกาสในการตอบถูกเท่ากัน ส่วนข้อสอบที่เกิดการทำหน้าที่ต่างกันของตัวลวง (DDF) นั้นพบว่าทั้งกลุ่มนักเรียนที่มีแหล่งทรัพยากรที่บ้านที่บ้านสูงและกลุ่มของนักเรียนที่มีแหล่งทรัพยากรที่บ้านต่ำมีโอกาสในการที่จะเลือกตอบ ตัวเลือกที่เป็นตัวลวงได้เท่ากัน และเมื่อพิจารณาลักษณะของเนื้อหาพบว่า เนื้อหาที่มีความแตกต่างคือพีชคณิตและเรขาคณิตโดยนักเรียนที่มีแหล่งทรัพยากรที่บ้านสูงมีโอกาสในการที่จะเลือกตอบตัวเลือกที่เป็นตัวลวงได้มากกว่ากลุ่มของนักเรียนที่มีแหล่งทรัพยากรที่บ้านต่ำ ส่ว นเรื่องข้อมูลและโอกาสนักเรียนที่มีแหล่งทรัพยากรที่บ้านที่บ้านต่ำมีโอกาสในการที่จะเลือกตอบตัวเลือกที่เป็นตัวลวงได้มากกว่ากลุ่มของนักเรียนที่มีแหล่งทรัพยากรที่บ้านสูง

ตัวแปรความมั่งคั่งของครอบครัว พบว่านักเรียนที่มี ความมั่งคั่งของครอบครัวสูงมีโอกาสในการทำข้ อสอบถูกต้องเท่ากับนักเรียนที่มีความมั่งคั่งของครอบครัวต่ำ และเมื่อพิจารณาลักษณะของเนื้อหาพบว่าเรื่องจำนวน กับข้อมูลและโอกาสนักเรียนที่มีความมั่งคั่งของครอบครัวสูงมีโอกาสในการทำข้อสอบถูกต้องมากกว่านักเรียนที่มีความมั่งคั่งของครอบครัวต่ำ ส่วนเรื่องพีชคณิตและ เรขาคณิต นักเรียนที่มีความมั่งคั่งของครอบครัวต่ำมีโอกาสในการทำข้อสอบถูกต้องมากกว่านักเรียนที่มีความมั่งคั่งของครอบครัวสูง ส่วนข้อสอบที่เกิดการทำหน้าที่ต่างกันของตัวลวง (DDF) นั้นพบว่า ทั้งกลุ่มนักเรียน ที่มีแหล่งทรัพยากรที่บ้านที่บ้านสูงและกลุ่มของนักเรียนที่ มีความมั่งคั่งของครอบครัวต่ำมีโอกาสในการที่จะเลือกตอบตัวเลือกที่เป็นตัวลวงมากกว่ากลุ่มของนักเรียนที่มีความมั่งคั่งของครอบครัวสูง และเมื่อพิจารณาลักษณะของเนื้อหาพบว่า เนื้อหาที่มีความแตกต่างคือจำนวน และพีชคณิตโดยนักเรียนที่มีความมั่งคั่งของครอบครัวต่ำมีโอ กาสในการที่จะเลือกตอบตัวเลือกที่เป็นตัวลวงได้มากกว่ากลุ่มของนักเรียนที่มีความมั่งคั่งของครอบครัว สูง ส่วนเรื่องเรขาคณิตนักเรียนที่มีความมั่งคั่งของครอบครัว สูงมี โอกาสในการที่จะเลือกตอบตัวเลือกที่เป็นตัวลวงได้มากกว่ากลุ่มของนักเรียนที่มีแหล่งทรัพยากรที่บ้านต่ำ

2. ผลของข้อสอบที่ทำหน้าที่ต่างกัน และผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาต่อการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน

2.1 ผลของข้อสอบที่ทำหน้าที่ต่างกัน ในแบบสอบประเมินความรู้วิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาที่ต่างกัน

การเปรียบเทียบผลการประเมินคุณภาพการจัดการศึกษาด้วยการเปรียบเทียบการจัดระดับคุณภาพ (Rating) การจัดการศึกษา และอันดับคุณภาพ (Ranking) การจัดการศึกษาของสถานศึกษาระหว่างโมเดล 1 กับ โมเดล 2, โมเดล 1 กับ โมเดล 4 และโมเดล 2 กับ โมเดล 4 พบว่าผลการจัดระดับคุณภาพการจัดการศึกษา พบว่า โมเดล 1 และโมเดล 2 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาอยู่ในกลุ่มเดียวกันได้ 148 แห่ง คิดเป็นร้อยละ 98.67 และผลการจัดอันดับคุณภาพการจัดการศึกษา พบว่าสถานศึกษาที่มีอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 ตรงกับอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 2 มีจำนวน 28 แห่ง คิดเป็นร้อยละ 18.67 ส่วนโมเดล 1 และโมเดล 4 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาอยู่ในกลุ่มเดียวกันได้ 148 แห่ง คิดเป็นร้อยละ 98.67 และการจัดอันดับคุณภาพการจัดการศึกษา พบว่าสถานศึกษาที่มีอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 ตรงกับอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มีจำนวน 18 แห่ง คิดเป็นร้อยละ 12 และผลการจัดระดับคุณภาพการจัดการศึกษา พบว่า โมเดล 2 และโมเดล 4 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาอยู่ในกลุ่มเดียวกันได้ 148 แห่ง คิดเป็นร้อยละ 98.67 โดยผลการจัดอันดับคุณภาพการจัดการศึกษา พบว่าสถานศึกษาที่มีอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 2 ตรงกับอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มีจำนวน 15 แห่ง คิดเป็นร้อยละ 10

จากผลของสถิติทดสอบวิลคอกซอน (Wilcoxon Signed Ranks Test) ในการทดสอบความสอดคล้องของการจัดระดับคุณภาพและอันดับคุณภาพการจัดการศึกษา พบว่าโมเดล 1 และ 2, โมเดล 1 และ 4 และโมเดล 2 และ 4 ทั้ง 3 คู่จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษา และจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษาไม่แตกต่างกัน

จากผลการเปรียบเทียบการจัดระดับคุณภาพการจัดการศึกษาและอันดับคุณภาพการจัดการศึกษาของสถานศึกษาระหว่างโมเดล 1 และ 2 ซึ่งเป็นการประเมินด้วยโมเดลที่ไม่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป และโมเดลที่ตัดข้อสอบที่ทำหน้าที่ต่างกันเฉพาะตัวเลือกที่เป็นตัวถูกออกไป โมเดล 1 และ 4 เป็นโมเดลที่ไม่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป และโมเดลที่ตัดข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวงออกไป และโมเดล 2 และ 4 ซึ่งเป็นการเปรียบเทียบระหว่างโมเดล ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันเฉพาะตัวเลือกที่เป็นตัวถูกออกไปกับโมเดลที่ตัดข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวงออกไป พบว่าผลการประเมินทั้ง 3 แบบ จะทำให้ผลการจัดระดับคุณภาพและอันดับคุณภาพการจัดการศึกษาของสถานศึกษาไม่ต่างกัน แสดงให้เห็นว่าการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) ในแบบสอบวิชาคณิตศาสตร์ไม่มีผลต่อการประเมินคุณภาพการจัดการศึกษา

2.2 ผลของ ตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาเมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาที่แตกต่างกัน

การเปรียบเทียบผลการประเมินคุณภาพการจัดการศึกษาด้วยการเปรียบเทียบการจัดระดับคุณภาพ (Rating) การจัดการศึกษา และ อันดับคุณภาพ (Ranking) การจัดการศึกษาของสถานศึกษาระหว่างโมเดล 1 กับ โมเดล 3, โมเดล 2 กับ โมเดล 3 และโมเดล 3 กับ โมเดล 4 พบว่าผลการจัดระดับคุณภาพการจัดการศึกษา พบว่า โมเดล 1 และโมเดล 3 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาอยู่ในกลุ่มเดียวกันได้ 96 แห่ง คิดเป็นร้อยละ 64 และผลการจัดอันดับคุณภาพการจัดการศึกษา พบว่าสถานศึกษาที่มีอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 ตรงกับอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 มีจำนวน 3 แห่ง คิดเป็นร้อยละ 2 ส่วน ส่วนผลการจัดระดับคุณภาพการจัดการศึกษา พบว่า โมเดล 2 และโมเดล 3 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาอยู่ในกลุ่มเดียวกันได้ 96 แห่ง คิดเป็นร้อยละ 64 และผลการจัดอันดับคุณภาพการจัดการศึกษา พบว่าสถานศึกษาที่มีอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 2 ตรงกับอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 มีจำนวน 2 แห่ง คิดเป็นร้อยละ 1.33 โดยผลการจัดระดับคุณภาพการจัดการศึกษา พบว่า โมเดล 3 และโมเดล 4 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาอยู่ในกลุ่มเดียวกันได้ 98 แห่ง คิดเป็นร้อยละ 65.33 และผลการจัดอันดับคุณภาพการจัดการศึกษา พบว่าสถานศึกษาที่มี อันดับ

คุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 ตรงกับอันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มีจำนวน 6 แห่ง คิดเป็นร้อยละ 4

จากผลของสถิติทดสอบวิลคอกซอน (Wilcoxon Signed Ranks Test) ในการทดสอบความสอดคล้องของการจัด ระดับคุณภาพ และอันดับคุณภาพการจัดการศึกษา พบว่าโมเดล 1 และ 3, โมเดล 2 และ 3 และโมเดล 3 และ 4 ทั้ง 3 คู่จัดระดับคุณภาพ การจัดการศึกษาของสถานศึกษา ไม่แตกต่างกัน แต่โมเดล 1 และ 3, โมเดล 2 และ 3 และโมเดล 3 และ 4 ทั้ง 3 คู่จัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษาแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

จากผลการเปรียบเทียบการจัด ระดับคุณภาพ และอันดับคุณภาพ การจัดการศึกษาของสถานศึกษาระหว่างโมเดล 1 และ3, โมเดล 2 และ3 และระหว่างโมเดล 3 และ 4 พบว่าการจัดระดับคุณภาพ ของสถานศึกษาไม่ แตกต่างกัน แต่อันดับคุณภาพ การจัดการศึกษามีความแตกต่างกันอย่างมีนัยสำ คัญทางสถิติ แสดงให้เห็นว่าคุณลักษณะของตัวแปรระดับนักเรียนและสถานศึกษามีผลต่อการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา

โดยรายละเอียด ผลของตัวแปรที่ส่งผลต่อผล การวิเคราะห์มูลค่าเพิ่มของคะแนนผลการประเมินวิชาคณิตศาสตร์ พบว่าตัวแปรระดับสถานศึกษาที่มีผลต่อคะแนนผลการประเมินวิชาคณิตศาสตร์แตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ได้แก่ สถานศึกษาขนาดใหญ่พิเศษ และ สถานที่ตั้งสถานศึกษาอยู่ใน เมือง ส่วนตัวแปรที่ส่งผลต่อคะแนนผลการประเมินวิชาคณิตศาสตร์อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ สถานศึกษาขนาดใหญ่ และดัชนีการให้ความสำคัญกับการให้การบ้านคณิตศาสตร์ของครูในสถานศึกษา ส่วนตัวแปรคุณลักษณะระดับสถานศึกษาที่ไม่มีผลต่อคะแนนผลการประเมินวิชาคณิตศาสตร์ ได้แก่ ดัชนีการสนับสนุนการจัดสรรทรัพยากร การจัดกิจกรรมเพื่อการเรียนรู้คณิตศาสตร์ ดัชนีการรับรู้ถึงบรรยากาศที่ดี ในการทำงานภายในสถานศึกษา การพัฒนาครูในสถานศึกษา ประสบการณ์การทำงานของครูในสถานศึกษา ดัชนีความพึงพอใจของครูที่มีต่อการจัดการภายในสถานศึกษา ขนาดข องห้องเรียน ระดับความถี่ในการทำงานร่วมกันของครูคณิตศาสตร์ในสถานศึกษา เมื่อพิจารณาขนาดอิทธิพลหรือค่าสัมประสิทธิ์ถดถอย พบว่า ค่าสัมประสิทธิ์ถดถอยของตัวแปรระดับสถานศึกษาที่มีค่าเป็นบวกสูงที่สุดคือ สถานศึกษาที่ มีขนาดใหญ่พิเศษ โดยเป็นสถานศึกษาที่มีนักเรียนตั้งแต่2500 คนขึ้นไป (β = 0.765686) รองลงมาคือ สถานศึกษาที่ตั้งอยู่ในเมือง(β = 0.311936) ส่วนตัวแปรที่มีสัมประสิทธิ์ถดถอยของตัวแปรระดับสถานศึกษาที่มีค่าเป็นลบสูงที่สุดคือ การให้การสนับสนุนของ

ผู้ปกครอง (β = -0.069615) รองลงมาคือ การจัดกิจกรรมเสริมทางคณิตศาสตร์ภายในสถานศึกษา (β = -0.041858)

ตัวแปรระดับนักเรียน ที่มีผลต่อคะแนนผลการประเมินวิชาคณิตศาสตร์ที่มี ค่าสัมประสิทธ์ถดถอยแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 คือ ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา ระดับความเป็นเพศหญิง แหล่งทรัพยากรที่บ้าน การฝึกปฏิบัติทักษะทางด้านคณิตศาสตร์ในชั่วโมงเรียน การใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเอง ความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนคณิตศาสตร์ ความเพลิดเพลินทางคณิตศาสตร์ ส่วนตัวแปรที่ส่งผลต่อคะแนนผลการประเมินวิชาคณิตศาสตร์อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ ดัชนีระดับเวลาที่ใช้ในการทำการบ้าน วิชาคณิตศาสตร์ และ การเห็นคุณค่าโดยทั่วไปในคณิตศาสตร์ ส่วนตัวแปรที่ไม่มีผลต่อคะแนนผลการเมินวิชาคณิตศาสตร์ ได้แก่ ระดับการศึกษาของผู้ปกครอง ความมั่งคั่งของครอบครัว แรงจูงใจภายนอกในการเรียนคณิตศาสตร์ และความรู้สึกเกี่ยวความ ปลอดภัยในการมาเรียนภายในสถานศึกษา เมื่อพิจารณาขนาดอิทธิพลหรือค่าสัมประสิทธิ์ถดถอย พบว่าตัวแปรระดับนักเรียนที่มีอิทธิพลทางบวกสูงที่สุดคือ ความเชื่อมั่นในประสิทธิภาพของตนเอง ในการเรียนคณิตศาสตร์ (β = 0.164069) รองลงมาคือ การฝึกปฏิบัติทักษะทางด้านคณิตศาสตร์ในชั่วโมง เรียน (β = 0.069308) และความมั่งคั่งของครอบครัว (β = 0.068723) ส่วนตัวแปรระดับนักเรียนที่มีอิทธิพลน้ อยที่สุดคือ การใช้เวลาในการเรียนคณิตศาสตร์ด้วยตนเอง(β = -0.044132) รองลงมาคือ ระดับการศึกษาของผู้ปกครอง (β = -0.007753)

3. ผลการเปรียบเทียบประสิทธิภาพโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์

3.1 ผลการเปรียบเทียบประสิทธิภาพคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ขอ งสถานศึกษาขั้นพื้นฐานจากการใช้โมเดลการประเมินต่างกัน

ความสอดคล้องของผลการประเมิน จากการวิเคราะห์ด้วยโมเดลประเมินคุณภาพการจัดการศึกษา 4 โมเดล พบว่าอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์คะแนนมูลค่ าเพิ่มแต่ละโมเดลแตกต่างกัน เมื่อจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษา โดยแบ่งเป็น 3 กลุ่ม คือ กลุ่มต่ำ กลุ่มกลาง กลุ่มสูง พบว่าผลการประเมินคุณภาพการจัดการศึกษาของโมเดลที่สอดคล้องกัน

คือ โมเดล 1 กับ โมเดล 2 สอดคล้องกันมากที่สุดคือ 100% และโมเดล 1 กับ โมเดล 4 และโมเดล 2 กับ โมเดล 4 ทั้งสองโมเดลมีความสอดคล้องกันมากถึง 98.67% นั่นคือโมเดลที่มีการควบคุมอิทธิพลของคุณลักษณะของนักเรียนและสถานศึกษาจะมีผลการประเมินสอดคล้องกันมาก

สัมประสิทธิ์การทำนาย หรือความแปรปรวนของคะแนนผลการประเมินวิชาคณิตศาสตร์ที่อธิบายได้ด้วยคุณลักษณะของนักเรียนและสถานศึกษา พบว่าโมเดล 4 โมเดล 1 และโมเดล 2 มีความใกล้เคียงกัน โดยโมเดล 4 มีค่าสัมประสิทธิ์การทำนาย (R^2) มีค่าเท่ากับ 0.5203 (52.03%) โมเดล 1 มีค่าสัมประสิทธิ์การทำนาย (R^2) มีค่าเท่ากับ 0.51962 (51.96%) โมเดล 1 มีค่าสัมประสิทธิ์การทำนาย (R^2) มีค่าเท่ากับ 0.51962 (51.96%) และโมเดล 2 มีค่าสัมประสิทธิ์การทำนาย (R^2) มีค่าเท่ากับ 0.5186 (51.86%)

สหสัมพันธ์ของตำแหน่งคะแนนมูลค่าเพิ่ม ด้วยการตรวจสอบสหสัมพันธ์ตำแหน่งคะแนนของสเปียร์แมน พบว่า โมเดล 1 กับโมเดล 2 มีค่าสหสัมพันธ์กันสูงสุด เท่ากับ .995 ส่วนโมเดล 1 กับโมเดล 4 เท่ากับ .992 และโมเดล 2 กับโมเดล 4 เท่ากับ .987 ต่ำที่สุดเป็น โมเดล 2 กับ โมเดล 3 เท่ากับ .708

ตารางที่ 44 สรุปความสอดคล้องของผลการประเมิน สัมประสิทธิ์การทำนายและค่าสหสัมพันธ์ของตำแหน่งคะแนนมูลค่าเพิ่มจากผลการประเมิน 4 โมเดล

โมเดล	ความสอดคล้องของผลการประเมิน				สปส.	ค่าสหสัมพันธ์			
	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4	การทำนาย	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
โมเดล 1	-	100%	64%	98.67%	51.96%	-	0.995	0.721	0.992
โมเดล 2	-	-	64%	98.67%	51.86%	-	-	0.708	0.987
โมเดล 3	-	-	-	65.33%	-	-	-	-	0.733
โมเดล 4	-	-	-	-	52.03%	-	-	-	-

3.2 สรุปแนวทางการเลือกใช้โมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาขั้นพื้นฐาน

จากการเปรียบเทียบผลการประเมินคุณภาพการจัดการศึกษาของทั้ง 4 โมเดล สามารถสรุปเป็นแนวทางการพิจารณาเลือกโมเดลการประเมินคุณภาพการจัดการศึกษา แต่ละโมเดล ดังนี้

โมเดล 1 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาที่ใช้ผลคะแนนจากข้อสอบทั้งฉบับ แต่มีการคำนึงถึงความแตกต่างของคุณลักษณะของนักเรียน และสถานศึกษา พร้อมทั้งมีการ ควบคุมตัวแปรอิทธิพลของตัวแปรดังกล่าว เนื่องจากการควบคุมอิทธิพล ของคุณลักษณะของนักเรียน และสถานศึกษา จึงทำให้ผลการประเมินถูกปรับให้อยู่บนพื้นฐานของปัจจัยดั้งเดิมเท่าเทียมกัน และนำผลต่างหรือส่วนที่เพิ่มขึ้นจากการควบคุมมาเปรียบเทียบ

กันอย่างยุติธรรม นอกจากนี้เมื่อเปรียบเทียบกับโมเดล 2 และ 4 จะพบว่าการวิเคราะห์ไม่ซับซ้อนเท่ากับโมเดลทั้ง 2 แบบนั้น แต่ให้ผลการประเมินคุณภาพการจัดการศึกษาสอดคล้องกับโมเดลที่มีการตัดข้อสอบ เกิดการทำหน้าที่ต่างกันทั้งเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง

โมเดล 2 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาที่มีการวิเคราะห์ที่ซับซ้อนมากกว่าโมเดล 2 และ โมเดล 3 เป็นโมเดลที่มีความน่าเชื่อถือและมีความธรรม เนื่องจากมีการตรวจสอบคุณภาพของแบบสอบและตัดข้อสอบที่ทำหน้าที่ต่างกันเฉพาะที่เป็นตัวเลือกที่เป็นตัวถูกออกไป (DIF) และมีการควบคุมอิทธิพลของคุณลักษณะของนักเรียน และสถานศึกษาเช่นเดียวกับโมเดล 1 จึงทำให้ผลการประเมินที่ได้จึงมีการควบคุมโอกาสการตอบข้อสอบถูกของกลุ่มนักเรียนที่ต่างกัน และควบคุมอิทธิพลของคุณลักษณะของนักเรียน และสถานศึกษาทำให้ผลการประเมินมีความเท่าเทียม และยุติธรรมสำหรับกลุ่มนักเรียนมากขึ้น แต่เมื่อเปรียบเทียบกับโมเดล 1 พบว่ามีใช้เทคนิคการวิเคราะห์ขั้นสูงและซับซ้อนมากกว่าโมเดล 1 แต่ให้ผลการประเมินคุณภาพการจัดการศึกษาที่สอดคล้องกันกับโมเดล 1

โมเดล 3 เป็นโมเดลการประเมินที่ไม่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะนักเรียนและสถานศึกษา แต่มีการตรวจสอบคุณภาพของแบบสอบและตัดข้อสอบ ที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูก และตัวเลือกที่เป็นตัวลวงออกไป ดังนั้นคะแนนที่ได้จากแบบสอบจึงมีการควบคุมโอกาสในการตอบถูกให้เท่าเทียมกันในคุณลักษณะที่ต่างกัน เป็นโมเดลที่มีการวิเคราะห์ไม่ซับซ้อนเท่าโมเดล 1, 2 และ 4 ตามลำดับ แต่ให้ผลการประเมินคุณภาพการจัดการศึกษาที่ค่อนข้างไม่สูงมาก

โมเดล 4 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาที่มีการวิเคราะห์ที่ซับซ้อนมากที่สุด เป็นโมเดลที่มีความน่าเชื่อถือที่สุดและมีความธรรมที่สุด เนื่องจากมีการตรวจสอบคุณภาพของแบบสอบทั้งการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวงพร้อมทั้งตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป เช่นเดียวกับโมเดล 2 กับ โมเดล 3 และมีการควบคุมอิทธิพลของคุณลักษณะของนักเรียนและสถานศึกษา เช่นเดียวกับโมเดล 1 กับ โมเดล 2 จึงทำให้ผลการประเมินที่ได้จึงมีการควบคุมโอกาสการตอบข้อสอบถูกของกลุ่มนักเรียนที่ต่างกัน และควบคุมอิทธิพลของคุณลักษณะของนักเรียน และสถานศึกษา ทำให้ผลการประเมินมีความเท่าเทียม และยุติธรรมสำหรับกลุ่มนักเรียนมากที่สุด แต่เมื่อเปรียบเทียบกับโมเดล 1 พบว่ามีการใช้เทคนิคการวิเคราะห์ขั้นสูงและซับซ้อนมากกว่าโมเดล 1 กับ โมเดล 2 แต่ให้ผลการประเมินคุณภาพการจัดการศึกษาที่สอดคล้องกันกับโมเดล 1 กับ โมเดล 2

5.2 อภิปรายผลการวิจัย

การวิจัยเรื่อง การวิเคราะห์เปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ : การประยุกต์ใช้โมเดลมูลค่าเพิ่มที่มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและการวิเคราะห์การทำหน้าที่ ต่างกันของตัวลวง สามารถอภิปรายผลการวิจัยได้ใน 3 ประเด็นต่อไปนี้

1. ข้อสอบที่เกิดการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) ในแบบสอบประเมินความรู้วิชาคณิตศาสตร์

ในการวิจัยครั้งนี้ ผลของข้อสอบที่ทำหน้าที่ต่างกัน เกิดจากคุณลักษณ ะของนักเรียนที่แตกต่างกันใน 5 คุณลักษณะ คือความเป็นเพศ (หญิง, ชาย) การใช้เวลาในการทำการบ้านวิชาคณิตศาสตร์ (สูง, ต่ำ) ระดับการศึกษาของผู้ปกครอง(สูง, ต่ำ) แหล่งทรัพยากรที่บ้าน (สูง, ต่ำ) และความมั่งคั่งของครอบครัว (สูง, ต่ำ) โดยผลจากการวิจัยพบว่าตัวแปรเพศเป็ นเป็นตัวแปรที่พบการทำหน้าที่ต่างกันของข้อสอบมากที่สุด นั่นอาจเป็นเพราะคุณลักษณะที่แตกต่างกันระหว่างผู้หญิงและผู้ชาย นั่นอาจเป็นเพราะคุณลักษณะของเพศมีการแบ่งแยกได้อย่างชัดเจนที่สุด โดยผลการวิจัยที่ได้มีความ สอดคล้องกับงานวิจัยของ สุพัฒน์ สุกมลสันต์ (2534) ได้ทำการศึกษาความลำเอียงของข้อสอบภาษาอังกฤษ ชุด กข และ กขค เข้ามหาวิทยาลัย ปี พ .ศ. 2531 – 2533 จำนวน 6 ฉบับ ฉบับละ 100 ข้อ โดยวิเคราะห์ความลำเอียงต่อเพศและต่อภาคภูมิศาสตร์ของผู้สมัครสอบ พบข้อสอบที่มีความลำเอียงต่อเพศและต่อภาคภูมิศาสตร์ของผู้สอบจำนวนแตกต่างกันอย่างมีนัยสำคัญ และงานวิจัยของ รักชนก ยี่สุ่นศรี (2544) ทำการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและแบบสอบสำหรับกลุ่มผู้สอบเมื่อจำแนกตามเพศและสถานที่ตั้งทางภูมิศาสตร์และโรงเรียนที่จ บการศึกษา ผลวิจัยพบว่า แบบสอบวิชาภา ษาอังกฤษทำหน้าที่ต่างกันตามเพศและสถานที่ตั้งตามภูมิศาสตร์ ส่วนแบบสอบวิชาคณิตศาสตร์ทำหน้าที่ต่างกันตามเพศของผู้สอบ และข้อสอบทำหน้าที่ต่างกันตามเพศของผู้สอบมากที่สุดทั้งสองวิชา งานวิจัยของ Marshall (1993) ทำการตรวจสอบความแต กต่างของเพศในตัวเลือกที่เป็นตัวลวง โดยใช้แบบสำรวจแคลิฟอเนียร์ เวอร์ชั่น 6 ของทักษะพื้นฐาน ผู้วิจัยสนใจในความแตกต่างของเพศ ตัวเลือกที่เป็นตัวลวง และปีที่สร้างแบบสอบขึ้นมาโดยใช้โมเดลลอกซ์ลิเนีย ซึ่งยอมให้อิทธิพลหลักแยกปฏิกิริยาของตัวแปรปรับเทียบเป็นแต่ละประเ ภท ข้อสอบมีจำนวน 30 ข้อในแบบสอบ 13%

แสดงให้เห็นถึงการทำหน้าที่ต่างกันของตัวลวงระหว่างเพศ รวมถึงรูปแบบของความคลาดเคลื่อนขึ้นอยู่กับความแตกต่างของปีที่ใช้แบบสอบ และ 20% ของข้อสอบแสดงถึงการทำหน้าที่ต่างกันของตัวลวง การศึกษาในครั้งนี้ยังใช้ตัวเลือกที่เป็นตัวลว งในการแยกชนิดของความคลาดเคลื่อนที่เกิดขึ้นโดยเพศชายและเพศหญิง พบว่าความแตกต่างของเพศเป็นชนิดของความคลาดเคลื่อนที่มีขนาดใหญ่

ประเด็นที่น่าสนใจอีกประเด็นหนึ่ง คือผลการวิจัยยังพบว่าเนื้อหา สาระวิชาคณิตศาสตร์ที่แตกต่างกันทำให้โอกาสในการตอบข้อสอบได้ถูกต้อง แล ะโอกาสในการเลือกตัวลวงแต่ละตัวเลือกไม่เท่ากันระหว่างเพศชายและเพศหญิงที่มีความสามารถเท่ากัน โดยจากผลการวิจัยพบว่าเพศหญิงมีโอกาสในการทำข้อสอบวิชาคณิตศาสตร์ในเนื้อหาเรื่องพีชคณิตได้ถูกต้องได้มากกว่าเพศชาย ในขณะที่เพศชายมีโอกาสทำข้อสอบวิชาคณิตศาสตร์ เรื่องเรขา คณิตได้ถูกต้องมากกว่าเพศหญิง และสำหรับตัวลวง ข้อสอบเรื่องจำนวน , พีชคณิต , และข้อมูลและโอกาส เพศชายมีโอกาสที่จะถูกลวงได้มากกว่าเพศหญิง ในขณะที่เรื่องเรขาคณิตเพศหญิงมีโอกาสที่จะถูกลวงได้มากกว่าเพศชาย ซึ่งจากข้อค้นพบที่ว่าลักษณะเนื้อหาที่แตกต่างส่งผลให้โอกาสตอบ ข้อสอบได้ถูกต้องของเพศหญิง และเพศชายต่างกัน มีความสอดคล้องกับงานวิจัยของ Katherine และ Meichu (1996) ได้ทำการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ สำหรับแบบสอบวิชาคณิตศาสตร์ ซึ่งเป็นแบบสอบแบบหลายตัวเลือก โดยแบ่งกลุ่มผู้สอบตามตัวแปรเพศ กลุ่มตัวอย่าง ผลการวิจัยพบว่าข้อสอบด้านพีชคณิตและการคำนวณจะง่ายสำหรับผู้สอบเพศหญิงและข้อสอบด้านเรขาคณิตจะเข้าข้างเพศชาย ในขณะที่ข้อสอบด้านเลขคณิตจะเข้าข้างเพศหญิง หรืองานวิจัยของ Giray (1995) ได้ทำการศึกษาวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ สำหรับแบบสอบคัดเลือกบุคคลเข้าศึกษาใน มหาวิทยาลัย ซึ่งเป็นข้อสอบด้านการคำนวณ คำศัพท์ และเรขาคณิต โดยแบ่งกลุ่มผู้สอบตามตัวแปรที่ศึกษาคือ เพศและเศรษฐานะ จากผลการวิจัยพบว่า ข้อสอบที่ทำหน้าที่ต่างกันเป็นข้อสอบในวิชาคณิตศาสตร์ ทั้งในกรณีแบ่งกลุ่มผู้สอบตามตัวแปรเพศและเศรษฐานะ สำหรับตัวแปรเพศ ผลการวิ เคราะห์พบว่าผู้ชายจะมีข้อได้เปรียบในการทำข้อสอบด้านการคำนวณ ในขณะที่เพศหญิงจะมีข้อได้เปรียบในการทำสอบด้านคำศัพท์และเรขาคณิต ซึ่งชี้ให้เห็นว่าผู้หญิงจะมีความสามารถใน ด้านการใช้ภาษาหรือเกี่ยวกับการใช้คำศัพท์ดีกว่าและเพศชายจะมีทักษะด้านการคำนวณดีกว่าเพศหญิง

นอกจากนี้ข้อค้นพบที่ว่าเนื้อหาวิชาทำให้โอกาสในการตอบข้อสอบได้ถูกต้องต้องกันของเพศชายและเพศหญิงยังมีความสอดคล้องกับเนื้อหาในรายวิชาอื่นๆอีกด้วย ดังจะพบได้จากงานวิจัยของสุมาลี แกล้วทนงค์ (2547) ได้ศึกษาสาเหตุของการทำหน้าที่ต่างกันของข้อสอบสาระการเรียนรู้ภาษาไ ทย และสาระการเรียนรู้สังคมศึกษา ศาสนาและวัฒนธรรม กับนักเรียนชั้นมัธยมศึกษาปีที่ 1 จำนวน 1,320 คน ผลการศึกษาพบว่า ข้อสอบที่ทำหน้าที่ต่างกันด้านเพศของแบบสอบสาระการเรียนรู้ภาษาไทยส่วนใหญ่มีสาเหตุมาจากความสนใจในเนื้อเรื่องและภาษาที่ใช้ในแบบสอบ ส่วนข้อสอบที่ทำห น้าที่ต่างกันด้านเพศของแบบสอบสามารถเรียนรู้สังคมศึกษาศาสนาและวัฒนธรรมนั้นมีสาเหตุมาจากเนื้อเรื่องที่สนใจและเนื้อเรื่องที่เกี่ยวกับวัฒนธรรมประเพณี

2. เพื่อศึกษาผลของข้อสอบที่ทำหน้าที่ต่างกัน และผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาต่อการประเมินคุณภาพก ารจัดการศึกษาวิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน

2.1 ข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบประเมินความรู้วิชาคณิตศาสตร์เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาที่ต่างกัน

จากผลการเปรียบเทียบผลการประเมินคุณภาพการจัดการศึกษาด้วยก ารเปรียบเทียบการจัดระดับคุณภาพ (Rating) การจัดการศึกษา และอันดับคุณภาพ (Ranking) การจัดการศึกษาของสถานศึกษาของโมเดล 1 โมเดล 2 และ โมเดล 4 โดยทั้ง 3 โมเดลต่างเป็นโมเดลที่มีการควบคุมอิทธิพลของคุณลักษณะของนักเรียน และสถานศึกษา แต่มีความแตกต่างกันในการส่วนที่เกี่ยวกับจำนวนข้อสอบที่ใช้ในการวิเคราะห์ คือโมเดล 1 จะไม่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป โมเดล 2 จะตัดข้อสอบที่ทำหน้าที่ต่างกันเฉพาะตัวเลือกที่เป็นตัวถูกออกไป และโมเดล 4 จะตัดข้อสอบข้อที่เกิดการทำหน้าที่ต่างกันทั่งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง โดยผลจากการวิเคราะห์ พบว่าผลการจัดระดับคุณภาพการจัดการศึกษา พบว่า

โมเดล 1, โมเดล 2 และโมเดล 4 จัดระดับคุณภาพและอันดับคุณภาพการจัดการศึกษาของสถานศึกษ าได้สอดคล้องกัน มาก โดยผลการทดสอบทางสถิติพบว่า ในการทดสอบความสอดคล้องของการจัด ระดับคุณภาพ และอันดับคุณภ าพการจัดการศึกษา พบว่าโมเดล 1 และ 2, โมเดล 1 และ 4 และโมเดล 2 และ 4 ทั้ง 3 คู่จัดระดับคุณภาพ การจัดการศึกษาของสถานศึกษา

และจัด อันดับคุณภาพ การจัดการศึกษาของสถานศึกษาไม่แตกต่างกัน แสดงให้เห็นว่าการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลว ง (DDF) ในแบบสอบวิชาคณิตศาสตร์ไม่มีผลต่อการประเมินคุณภาพการจัดการศึกษา ซึ่งสอดคล้องกับผลการวิจัยของประกฤติยา ทักษิโณ (2552) พบว่าผลของข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบวิชาวิทยาศาสตร์ต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน พบว่าข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบวิชาวิทยาศาสตร์ไม่มีผลต่อการประเมินคุณภาพการจัดการศึกษา โดยผลการวิจัยได้จากการนำโมเดลการประเมินที่ไม่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกและไม่มีการปรับแก้คะแนนด้วยตัวแปรนักเรียนและสถานศึกษา เปรียบเทียบกับ โมเดลที่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกและไม่มีการปรับแก้คะแนนด้วย ตัวแปรนักเรียนและสถานศึกษา

อย่างไรก็ตามถึงแม้ว่าการวิจัยครั้งนี้ทั้งการจัด ระดับคุณภาพ และอันดับคุณภาพ การจัดการศึกษาจากโมเดลการประเมินคุณภาพการจัดการศึกษาทั้งโมเดล 1 คือไม่มีการตัดข้อสอบออกเลย โมเดล 2 ตัดข้อสอบข้อที่ทำหน้าที่ต่างกันเฉพาะตัวเลือกที่เป็นตัวถูก (DIF) และโมเดล 4 ตัดข้อสอบข้อที่เกิดทั้งการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) พบว่ามีความสอดคล้องกัน แต่เนื่องจากการเกิดการทำหน้าที่ต่างกันของข้อสอบ หรือการทำหน้าที่ต่างกันของตัวลวงในข้อสอบแต่ละข้อนั้น ทำให้เกิดการได้เปรียบเสียเปรียบของนักเรียนกลุ่มใดกลุ่มหนึ่ง เพราะ การทำหน้าที่ต่างกันของข้อสอบเป็นผลที่เกิดขึ้นเมื่อนักเรียนที่มีคุณลักษณะย่อยต่างกันและมีความสามารถหรือคุณลักษณะที่มุ่งวัดเท่ากัน แต่มีโอกาสในการตอบข้อสอบได้ถูกต้องต่างกัน ดังนั้นหากในการสถานการณ์การสอบ หากมีการตัดข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูก และตัวเลือกที่เป็นตัวลวงออกไปเพื่อให้ข้อสอบมีคุณภาพ และเป็นธรรมมากขึ้น แต่หากตัดข้อสอบมากจนเกินไปก็อาจจะทำให้คุณภาพของข้อสอบลดลงได้เช่นกัน โดยงานวิจัยครั้งนี้ผู้วิจัยได้ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูก และตัวเลือกที่เป็นตัวลวงค่อนข้างมาก แต่เนื่องจากข้อสอบที่ใช้ในโครงการ การศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์ หรือ TIMSS ปี 2007 ซึ่งมีทั้งหมด 14 ฉบับ ในแต่ละฉบับมีจำนวนข้อสอบ ที่ไม่มากเท่าไรนัก ดังนั้นผู้วิจัยจึงจำเป็นต้องเลือกตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปเป็นบางข้อ โดยพิจารณาเลือกตัดข้อสอบตามเกณฑ์ 2 ข้อ ดังนี้ 1. เลือกข้อสอบที่ทำหน้าที่ต่างกันที่มีค่าสัมประสิทธิ์สูงสุด และ 2. คำนึงถึงโครงสร้างของแบบสอบ โดยภายหลังการตัดข้อสอบออก แล้ว

โครงสร้างของแบบสอบจะไม่แตกต่างจากโครงสร้างเดิม และข้อสอบที่ตัดออกไปควรมีจำนวนไม่เกินร้อยละ 20 ของข้อสอบในแบบสอบนั้น ทั้งนี้ผู้วิจัยคำนึงถึงคุณภาพของแบบสอบในด้านความตรง (Validity) ของแบบสอบที่ต้องวัดในกรอบโครงสร้างเดิม และคุณภาพด้านความเที่ยง (Reliability) เนื่องจากการตัดข้อสอบในแบบสอบส่งผลต่อความยาวของแบบสอบ และทาให้ความเที่ยงมีค่าลดลง ซึ่งโดยทั่วไปคะแนนจากแบบสอบที่มีความยาวจะมีความน่าเชื่อมากกว่าคะแนนจากแบบสอบที่สั้นกว่า (ศิริชัย กาญจนวาสี, 2548) นอกจากนี้จากผลการวิจัยของ รักชนก ยี่สุ่นศรี (2544) พบว่า คุณภาพของแบบสอบด้านความเที่ยงของแบบสอบในฉบับก่อนและหลังการตัดข้อสอบที่ทำหน้าที่ต่างกัน มีค่าความเที่ยงของแบบสอบแตกต่างกัน โดยถ้าตัดข้อสอบออกจากแบบสอบจำนวนมากขึ้นจะทำให้ค่าความเที่ยงของแบบสอบมีค่าลดลง ดังนั้นผู้วิจัยจากการที่ผู้วิจัยคำนึงถึงคุณภาพของแบบสอบในด้านความเที่ยงและความตรงจึงไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูก และตัวเลือกที่เป็นตัวลวงออกไปทั้งหมด

2.2 ตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาเมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาที่แตกต่างกัน

จากการเปรียบเทียบผลการประเมินคุ ณภาพการจัดการศึกษาด้วยการจัด ระดับคุณภาพ (Rating) การจัดการศึกษา และอันดับคุณภาพ (Ranking) การจัดการศึกษาของสถานศึกษา ระหว่างโมเดล 1 กับโมเดล 3, โมเดล 2 กับ โมเดล 3 และโมเดล 3 กับ โมเดล 4 ซึ่งจะเห็นได้ว่าในการเปรียบเทียบทั้งโดยทั้ง 3 คู่นั้น ใช้โมเดล 3 เป็นโมเดลหลักในการเปรียบเทียบ เนื่องจากโมเดล 3 เป็นโมเดลเดียวที่ไม่มีการควบคุมแหล่งความคลาดเคลื่อนในระดับนักเรียนและระดับสถานศึกษา พบว่าโมเดล 1 และโมเดล 3 จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาอยู่ในกลุ่มเดียวกันได้ คิดเป็นร้อยละ 64 และผลการจัดอันดับคุณภาพการจัดการศึกษา พบว่าสถานศึกษาที่มี อันดับคุณภาพการจัดการศึกษา ในโมเดล 1 ตรงกับอันดับคุณภาพการจัดการศึกษาจากการวิเคราะห์โมเดล 3 คิดเป็นร้อยละ 2 เท่านั้น ส่วนการเปรียบเทียบระหว่างโมเดล 2 และโมเดล 3 พบว่าจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาอยู่ในกลุ่มเดียวกัน คิดเป็นร้อยละ 64 และผลการจัดอันดับคุณภาพ พบว่าสถานศึกษาที่มี อันดับคุณภาพ การจัดการศึกษาจากผลการวิเคราะห์โมเดล 2 ตรงกับอันดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 มีจำนวนน้อยมาก โดยคิดเป็นร้อยละ 1.33 และสุดท้ายเป็นการเปรียบเทียบระหว่าง โมเดล 3 และโมเดล 4 พบว่าจัดระดับคุณภาพ การจัดการศึกษาของสถานศึกษาอยู่ในกลุ่มเดียวกันได้ คิดเป็นร้อยละ 65.33

ส่วนผลการจัด อันดับคุณภาพ การจัดการศึกษา พบว่าโมเดล 3 จัดอันดับคุณภาพ ได้สอดคล้องตรงกับการวิเคราะห์ใน โมเดล 4 คิดเป็นเพียงแค่ร้อยละ 4 เท่านั้น และจากผลการวิเคราะห์ทางสถิติถึงแม้ว่าจากผลการเปรียบเทียบการจัดระดับคุณภาพและอันดับคุณภาพการจัดการศึกษาของสถานศึกษาระหว่างโมเดล 1 และ3, โมเดล 2 และ3 และระหว่างโมเดล 3 และ 4 พบว่าการจัดระดับคุณภาพของสถานศึกษาไม่แตกต่างกัน แต่การวิเคราะห์อันดับคุณภาพการจัดการศึกษามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ผลจากการวิเคราะห์ทำให้ทราบว่า เมื่อมีการควบคุมอิทธิพลของตัวแปร ในระดับนักเรียน และระดับสถานศึกษา ตามหลักการวิเคราะห์มูลค่าเพิ่ม ทำให้ผลการประเมินคุณภาพการจัดการศึกษาในแต่ละโมเดล ให้ผลการประเมินที่มีความแตกต่างกัน โดยผลจากการวิจัยพบว่าตัวแปรที่มีอิทธิพลต่อการเปลี่ยนแปลงผลคะแนนการประเมินวิชาคณิตศาสตร์ มีหลายตัวด้วยกันเช่นในระดับนักเรียน ความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนคณิตศาสตร์ การฝึกปฏิบัติทักษะทางด้านคณิตศาสตร์ในชั่วโมงเรียน ความเพลิดเพลินทางคณิตศาสตร์ ระดับความเป็นเพศหญิง แหล่งทรัพยากรที่บ้าน การเห็นคุณค่าโดยทั่วไปในคณิตศาสตร์ ดัชนีระดับเวลาที่ใช้ในการ ทำการบ้าน วิชาคณิตศาสตร์ และ ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา และในระดับสถานศึกษา เช่น สถานศึกษาขนาดใหญ่พิเศษ สถานที่ตั้งสถานศึกษาอยู่ใน เมือง สถานศึกษาขนาดใหญ่ และดัชนีการให้ความสำ คัญกับการให้การบ้านคณิตศาสตร์ของครูในสถานศึกษา แสดงให้เห็นว่าคุณลักษณะของตัวแปรระดับนักเรียนและสถานศึกษามีผลต่อก ารประเมินคุณภาพการจัดการศึกษา ซึ่งก็มีความสอดคล้องกับความเป็นจริง ดังตัวอย่างเช่นคุณภาพของสถานศึกษาในประเทศไทยในปัจจุบันที่มีความแตกต่างกันค่อน ข้างมาก ส่วนหนึ่งเป็นเพราะเรื่องของความแตกต่างทางด้านทรัพยากรที่สำคัญเช่น เงิน บุคลากร หรือสถานที่ตั้งของสถานศึกษา เป็นต้น ความแตกต่างเหล่านี้เสมือนเป็นมูลค่าเพิ่มที่แตกต่างกันของแต่ละสถานศึกษา

ดังนั้น จากข้อค้นพบดังกล่าวมีความ สอดคล้องกับงานวิจัยของ ประกฤติ ยา ทักษิโณ (2552) พบว่า พบว่าผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาขั้นพื้นฐานต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน พบว่าคุณลักษณะของนักเรียนและสถานศึกษาขั้นพื้นฐานมีผลต่อการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา โดยการจัด ระดับคุณภาพ และอันดับคุณภาพ การจัดการศึกษาของสถานศึกษา

เปรียบเทียบระหว่าง โมเดลที่มีการควบคุมอิทธิพลของคุณลักษณะของนักเรียนและสถานศึกษา กับระหว่างโมเดลที่ไม่มีการควบคุมอิทธิพลของคุณลักษณะของนักเรียนและสถานศึกษา มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เช่นเดียวกับ Haegeland and Kirkeboen (2008) ได้ศึกษาโมเดลมูลค่าเพิ่มที่ต่างกันคือ 1. โมเดลที่ไม่ได้ปรับแก้ (Unadjusted results: UR) 2. โมเดลความสำเร็จของบริบท (Contextualised attainment models: CAM) 3. โมเดลมูลค่าเพิ่ม (Value added models: VAM) 4) โมเดลมูลค่าเพิ่มเชิงบริบท (Contextualised value added models: CVA) พบว่าแต่ละโมเดลได้ผลต่างกัน เนื่องจากตัวแปรเศรษฐานะทางสังคม ความรู้เดิม และตัวแปรสถานศึกษามีผลต่อการประมาณค่าอิทธิพลของสถานศึกษา

นอกจากนี้ Chiu & Xihua (2008) พบว่าปัจจัยในด้านครอบครัวก็ส่งผลต่อผลสัมฤทธิ์ของนักเรียน สรุปได้ว่าจำนวนสมาชิกของครอบครัวสามารถตอบสนองแหล่งทรัพยากรภายนอกของนักเรียน (ผู้ให้ทรัพยากร) หรือความสมบูรณ์ของแหล่งทรัพยากร ดังนั้นจำนวนสมาชิกของครอบครัว เช่นพ่อแม่จะมีการให้แหล่งทรัพยากรทางการศึกษาจำนวนมากซึ่งส่งผลต่อโอกาส ในการเรียนรู้มาก และนักเรียนก็สามารถไปสู่ความสำเร็จได้สูง เช่นเดียวกับTeddlie, Stringfield & Reynolds (2000) ได้สรุปผลที่เกิดจากตัวแปรระดับบุคคลด้านภูมิหลังของนักเรียน เกี่ยวกับขนาดของอิทธิพลที่เกิดจากเศรษฐานะสังคม (socioeconomic status: SES) ที่มีผลต่อผลสัมฤทธิ์ว่ามีความแตกต่างกันตามกลุ่มสมาชิกของแต่ละสถานศึกษา งานวิจัยของ Pintrich & Schunk (2002) พบว่านอกจากตัวแปรที่เป็นคุณลักษณะพื้นฐานของนักเรียนและพื้นฐานครอบครัวของนักเรียนที่ส่งผลต่อผลสัมฤทธิ์ของนักเรียน ยังพบตัวแปรแรงจูงใจและเจตคติของนักเรียน จากทฤษฎีส ติปัญญาทางสังคม แรงจูงใจทางการเรียนของนักเรียน เชื่อว่ามีสามองค์ประกอบคือ คุณค่า (Value) ความคาดหวัง (expectancy) และความชื่นชอบ (affect)

ข้อสรุปของ Bandura (1989) พบว่านักเรียนที่เห็นคุณค่าในการทำกิจกรรมจะมีลักษณะของความพยายามมากกว่าและได้รับความสำ เร็จสูง ความคาดหวังเชื่อว่านักเรียนจะมีความเชื่อในประสิทธิภาพของตนเอง (Self – efficacy) และความเชื่อในความสามารถของตนเอง (Self – concept) นอกจากนี้ ความคาดหวังจะถูกพิจารณาว่าเป็นความสามารถทางการศึกษาที่สามารถมีผลกับประสบการณ์ของนักเรียน โดยเฉพาะอย่างยิ่ง ความสำ เร็จที่ผ่านและความล้มเหลวในการทำชิ้นงานในการเรียน และจากการศึกษาของ Park & Palardy (2004) ว่าความหลากหลายของคุณลักษณะของนักเรียนมีความสัมพันธ์กับผลสัมฤทธิ์ทางการเรียน ได้แก่ คุณลักษณะทางกายภาพ เช่น เชื้อชาติ ศาสนา เพศ คุณลักษณะพื้นฐานของครอบครัว เช่น เศรษฐ านะและโครงสร้างของ

ครอบครัว และพื้นฐานการศึกษา เช่นผลสัมฤทธิ์เดิม หรือพื้นฐานความรู้เดิมของนักเรียน นอกจากงานวิจัยในต่างประเทศแล้วนั้น

นักวิจัยในประเทศไทยหลายท่านก็มีข้อค้นพบที่น่าสนใจ คือ พิชิต ธรรมรักษ์ (2549) พบว่าปัจจัยที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียน วิชาคณิตศาสตร์ ของนักเรียน ได้แก่ ประสบการการสอนของครู ขนาดโรงเรียน ความรู้พื้นฐานเดิม ความเป็นผู้นำทางวิชาการของผู้บริหาร อาชีพของผู้ปกครอง และ เจตคติต่อวิชาคณิตศาสตร์ และ บุญเรือง ศรีเหรัญ (2542) พบว่าปัจจัยระดับบุคคลได้แก่ความรู้เดิมของนักเรียน บรรยากาศและ สภาพแวดล้อมภายในครอบครัวส่ง ต่อผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์และวิทยาศาสตร์ เจตคติและแรงจูงใจต่อการเรียน ข้อสังเกตของ ตัวแปรระดับนักเรียนที่มีอิทธิพลทางบวก คือส่วนใหญ่เป็นตัวแปรที่นักเรียนสามารถพัฒนาความสามารถของตัวเองให้เกิดได้มากขึ้น

3. เปรียบเทียบโมเดลก ารประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ต่อคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์

ผลการวิจัยในครั้งนี้ ได้ จากการ ศึกษาโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ของสถานศึกษาขั้นพื้นฐานในโมเดลที่แตกต่างกัน โดยกำหนดประเด็นที่นำมาพิจารณาสองประเด็น คือ คุณภาพของแบบสอบและโมเดลการวิเคราะห์มูลค่าเพิ่ม โมเดลการประเมินคุณภาพการจัดการศึกษาทั้ง4 โมเดล สรุปได้ดังนี้

จาก การเปรียบเทียบ โมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ ได้อย่างชัดเจน โดยแสดงให้เห็นว่า โมเดลที่มีการควบคุม ควบคุมอิทธิพลตัวแปร ในระดับนักเ รียนและระดับสถานศึกษาจะให้ผลการประเมินคุณภาพการจัดการศึกษาที่ใกล้เคียงกัน โดยพิจารณาได้จาก สัมประสิทธิ์การทำนาย (R^2) ของแต่ละโมเดล พบว่าโมเดล 1 ซึ่งเป็นโมเดลที่ไม่มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง นั่นคือเป็นโมเดลที่ไม่มีการควบคุมแหล่งความคลาดเคลื่อนที่เกิดขึ้นจากการโอนเอียงให้ผู้สอบกลุ่มใดกลุ่มหนึ่ง มีสัมประสิทธิ์การทำนายเป็น ร้อยละ 51.96 ในขณะที่ การประเมินคุณภาพการจัดการศึกษาด้วยโมเดล 2 ซึ่งเป็นโมเดลที่มีขั้นตอนในการวิเคราะห์เพิ่มมากขึ้นกว่าโมเดล 1 คือมีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) พบว่าโมเดล 2 มีสัมประสิทธิ์การทำนายใกล้เคียงกับโมเดล 1 โดยคิดเป็นร้อย ละ 51.86% และสุดท้าย การประเมินคุณภาพการจัดการศึกษาด้วยโมเดล 4 ซึ่งเป็นโมเดลที่มีขั้นตอนในการวิเคราะห์เพิ่มมากขึ้นกว่าโมเดล 1 และโมเดล 2 คือมีการวิเคราะห์การทำหน้าที่ต่างกันของตัวลวง (DDF) พบว่าโมเดล 4 ถึงแม้ว่าจะมีสัมประสิทธิ์การทำนายใกล้เคียงกับ

โมเดล 1 และโมเดล 2 แต่ก็เป็นโมเดลที่มี สัมประสิทธิ์การทำนาย สูงที่สุดโดย ทำนายได้ร้อยละ 52.03 ซึ่งอาจจะแสดงให้เห็นว่า ถ้ามีการตัดข้อสอบที่ทำ หน้าที่ต่างกัน ทั้งตัวเลือ กที่เป็นตัวถูกและตัวเลือกที่เป็นตัวลวง ออกไป นั่นคือมีการควบคุมแหล่งความคลาดเคลื่อนอันเกิดจากการที่ข้อสอบโอนเอียงเข้าหาผู้สอบกลุ่มใดกลุ่มหนึ่ง จะทำ ให้โมเดลการประเมินคุณภาพการจัดการศึกษามีค่าสัมประสิทธิ์การทำ นายเพิ่มมากขึ้น แต่อย่างไรก็ตาม ผลการจัด ระดับคุณภา พและการจัดอันดับคุณภาพ การจัดการศึกษาของสถานศึกษาของทั้ง 3 โมเดล มีความ สอดคล้องตรงกันทั้งหมด ดังนั้นจึงน่าจะแสดงให้เห็นได้ว่าการประเมิ นคุณภาพการจัดการศึกษาด้วยโมเดล 1 โมเดล 2 และโมเดล 4 ให้ผลการประเมินคุณภาพการจัดการศึกษาที่ มีความ สอดคล้องกัน และสามารถเลือกใช้ โมเดลใด ก็ได้ในการประเมิน โดยขึ้นอยู่กับองค์ประกอบความพร้อมของสถานศึกษา

โดยในการเลือกใช้โมเดลได้มีการนักการศึกษาและนักวิจัยที่ศึกษาเกี่ยวกับการเลือกใช้โมเดลมูลค่าเพิ่ม เช่น Ballou, Sanders and Wright. (2004) ได้เสนอแนะเกี่ยวกับการเลือกโมเดลมูลค่าเพิ่มว่า ความแตกต่างของโมเดลมูลค่าเพิ่มแต่ละโมเดล ขึ้นอยู่กับการกำ หนดลักษณะเฉพาะของโมเดล ทั้งนี้ การเลือกใช้โมเดลในการวิเคราะห์แต่ ละโมเดล ก็ขึ้นอยู่กับความสามารถในการเข้าถึงข้อมูล ของผู้วิจัย และประกฤติยา ทักษิโณ (2552) สรุปประเด็นในการเลือกใช้โมเดลในการวิเคราะห์มูลค่ าเพิ่มดังนี้ การเลือกโมเดลมูลค่าเพิ่มเพื่อใช้ในการประเมินคุณภาพการจัดการของสถานศึกษาจึงควรระวังในการเลือกโมเดลที่เหมาะสมเนื่องจากสภาพบริบทที่แตกต่างกัน โดยควรพิจารณาประเด็นต่างๆ ดังนี้ วัตถุประสงค์ของการประเมิน การเลือกตัวแปร การเก็บรวบรวมข้อมูล ระยะเวลาและ ผู้เชี่ยวชาญในการวิเคราะห์ ความถูกต้องและเป็นธรรมของผลการประเมิน และการใช้ข้อสารสนเทศจากการประเมิน

ดังนั้นการประเมินคุณภาพการจัดการศึกษาโดยการวิเคราะห์มูลค่าเพิ่มได้อย่างมีประสิทธิภาพสูงสุดนั้น จำเป็นที่จะต้องคำนึงถึงรายละเอียดต่างๆ ของทั้งตัวแปรในระดับนัก เรียนและสถานศึกษา อีกทั้งหากมีการประยุกต์ใช้การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวงก็จะช่วยเพิ่มความน่าเชื่อถือของผลการประเมินที่ได้รับอีกด้วย

5.3 ข้อเสนอแนะ

จากงานวิจัยเรื่อง การวิเคราะห์เปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์: การประยุกต์ใช้โมเดลมูลค่าเพิ่มที่มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและการวิเคราะห์การทำหน้าที่ต่างกันของตัวลวง ได้ผลการวิจัยที่สามารถนำไปประโยชน์ได้ในหลายๆ สถานการณ์ โดยผู้วิจัยได้นำเสนอ **ข้อเสนอแนะในการนำผลการวิจัยไปใช้** และข้อเสนอแนะสำหรับงานวิจัยครั้งต่อไป ดังนี้

1. **ข้อเสนอแนะในการนำผลการวิจัยไปใช้** แบ่งเป็น 3 ส่วน ดังนี้

1.1 ระดับชาติ

จากผลการวิจัยพบว่า โมเดล 4 ที่มีการตัดข้อสอบที่มีการทำหน้าที่ต่างกันของข้อสอบและการทำหน้าที่ต่างกันของตัวลวงออกไป พร้อมทั้งมีการควบคุมอิทธิพลของตัวแปรระดับนักเรียน และระดับสถานศึกษา เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ เป็นโมเดลที่มีความน่าเชื่อถือที่สุด หน่วยงานหลักของชาติที่มีหน้าที่ในการประเมินสถานศึกษา ควรนำรูปแบบการประเมิน ของโมเดล 4 ไปใช้ เพราะ โมเดลการประเมิน นี้มีหลักการปรับความไม่เท่าเทียมกันของปัจจัยต่างๆในแต่ละสถานศึกษา ให้อยู่เป็นพื้นฐานที่สามารถเปรียบเทียบกันได้ ทั้งนี้เพื่อนำผลที่บ่งบอกถึงมูลค่าเพิ่มที่สถานศึกษาสามารถทำให้เกิดขึ้นจากปัจจัยที่มีอยู่ โดยหน่วยงานหลักที่ควรต้องนำไปใช้คือ สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา(องค์การมหาชน) หรือ สมศ. ซึ่งเกี่ยวข้องกับการประเมินมาตรฐานการจัดการศึกษาโดยตรง ซึ่งจะช่วยเพิ่มมิติในการประเมินได้เป็นอย่างดี

1.2 ระดับเขตพื้นที่

หน้าที่ของเขตพื้นที่คือ การกำกับ ดูแล ติดตาม และประเมินผลสถานศึกษาขั้นพื้นฐาน ในเขตพื้นที่การศึกษา ที่รับผิดชอบ เพื่อพัฒนาสถานศึกษาในเขตของตนเองให้มีมาตรฐาน เป็นเรื่องปกติที่ในเขตพื้นที่การศึกษาจะมีทั้งโรงเรียนที่มีมาตรฐานสูงและต่ำคละกันไป จากผลการวิจัยเรื่องนี้ เขตพื้นที่ควรมีการศึกษาวิธีการวิเคราะห์มูลค่าเพิ่ม พร้อมทั้ง นำโมเดลการประเมินไปใช้ในการตรวจสอบคุณภาพของสถานศึกษาในเขตพื้นที่ของหน่วยงานตนเอง โดยเฉพาะหากสามารถศึกษาจนพบว่าสถานศึกษาในกำกับของตนเองที่ใดที่มีมูลค่าเพิ่มในการจัดการศึกษามากที่สุด ก็จะสร้างประโยชน์ให้กับเขตพื้นที่เป็นอย่างมาก เพราะผลที่เขตพื้นที่จะได้รับในเบื้องต้นนั้น คือจะทำให้ทราบถึงตัวแปรระดับนักเรียน และสถานศึกษาที่มีทำให้สถานศึกษามีคุณภาพดียิ่งขึ้น พร้อมทั้ง

น่าจะมีความเหมาะสมกับการพัฒนาสถานศึกษาในบริบทของตนเอง นั่นจะช่วยให้เขตพื้นที่มีผลการดำเนินงานตามเป้าหมายที่วางไว้ได้เป็นอย่างดี

1.3 ระดับสถานศึกษา

คลังข้อสอบจัดเป็นการพัฒนางานทางด้านการวัดและประเมินผลอีกท างหนึ่งที่สถานศึกษาขั้นพื้นฐานหลายๆแห่ง กำลังพัฒนาคลังข้อสอบของตนเอง ซึ่งโดยปกติจะคัดเลือกข้อสอบเพื่อนำมาเก็บไว้ในคลังข้อสอบด้วยค่าความยากง่าย และค่าอำนาจจำแนกเป็นหลัก การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง เป็นอีกวิธีการหนึ่งที่จะช่วยสร้างความน่าเชื่อถือให้กับคลังข้อสอบของสถานศึกษา ในอดีตที่ผ่านมาสถานศึกษาไม่นิยมทำการตรวจสอบความลำอียงของข้อสอบเท่าใดนัก สาเหตุหนึ่งเนื่องจากขาดความชำนาญในการตรวจสอบ แต่จากงานวิจัยเรื่องนี้ โปรแกรมการวิเคราะห์ DDFS 1.0 (Penfield, 2010) เป็นโปรแกรมที่ใช้ง่ายโดยสามารถ (วิธีการใช้นำเสนอในภาคผนวก) สถานศึกษาสามารถนำไปใช้ในการวิเคราะห์ข้อสอบภายในสถานศึกษาได้ง่ายขึ้น จะเป็นการช่วยเพิ่มความน่าเชื่อถือของเครื่องมือที่ใช้ในการวัดผลภายในสถานศึกษาได้เป็นอย่างดี นอกจากนี้ยังเป็นมิติใหม่ของการวัดและประเมินผลในสถานศึกษาขั้นพื้นฐานอีกด้วย

Download from website: http://www.education.miami.edu/facultysites/penfield/index.html

2. ข้อเสนอแนะสำหรับงานวิจัยครั้งต่อไป

2.1 จากการที่งานวิจัย เรื่องนี้ใช้ผลการสอบ TIMSS 2007 ในการวิเคราะห์ข้อมูล ปัญหาที่พบอย่างหนึ่ง คือการที่ข้อสอบ TIMSS 2007 มีจำนวนข้อสอบ ในแต่ละฉบับ ที่ค่อนข้างน้อย แบบสอบทั้งหมด TIMSS 2007 ทั้งหมด 14 ฉบับ ในแต่ละฉบับจะมี **จำนวนข้อสอบฉบับละ** 28 – 33 **ข้อ**เท่านั้น ซึ่งจากลักษณะของโมเดล การประเมิน จำนวนข้อสอบเพียงเท่านี้ส่งผลต่อเงื่อนไขของโมเดลในการวิจัย ซึ่งในบางโมเดลจะมีก ารตัดข้อสอบที่ตรวจพบว่ามีการทำหน้าที่ต่างกันของข้อสอบ หรือการทำหน้าที่ต่างกันของตัวลวง ดังนั้นจึงส่งผลให้การเลือกตัดข้อสอบมีข้อจำกัดเนื่องจากข้อสอบน้อย จากข้อจำกัดดังกล่ าวข้อเสนอแนะสำหรับงานวิจัยครั้งต่อไปคือควรมีการศึกษากับแบบสอบซึ่งมีจำนวนข้อสอบที่มากขึ้นกว่าในงานวิจัยเรื่องนี้ ซึ่งจะช่วยเพิ่มสารสนเทศที่มีประโยชน์เกี่ยวกับการวิเคราะห์ได้มากยิ่งขึ้น

2.2 การวิจัยในครั้ง ต่อไปควรมีการพิจารณาการเลือกตัดข้อสอบด้วยเทคนิควิธีการพิจารณาจาก ขนาดอิทธิพล (Effect Size) โดยขนาดอิทธิพล เป็นค่าสถิติที่ใช้วัดเพื่อบอกระดับ

ความสัมพันธ์ระหว่างอิทธิพลของตัวแปรต้นที่มีต่อตัวแปรตาม เพื่อความน่าเชื่อถือของเกณฑ์ในการเลือกตัดข้อสอบที่เกิดการทำหน้าที่ต่างกันทั้งตัวเลือกที่เป็นตัวถูกและตัวเลือกที่เป็นตัว ลวง โดยสำหรับในการวิจัยเรื่องนี้ผู้วิจัยมีก ารเลือกตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปเป็นบางข้อตามเงื่อนไขของโมเดลในการศึกษา โดยพิจารณาเลือกตัดข้อสอบตามเกณฑ์ 2 ข้อ ดังนี้ 1. เลือกข้อสอบที่ทำหน้าที่ต่างกันที่มีค่า อิทธิพล สูงสุด และ 2. คำนึงถึงโครงสร้างของแบบสอ บ โดยภายหลังการตัดข้อสอบออกแล้วโครงสร้างของแบบสอบจะไม่แตกต่างจากโครงสร้างเดิม และข้อสอบที่ตัดออกไปควรมีจำนวนไม่เกินร้อยละ 20 (Clauser (1993), cited in Naraya & Swaminathan,1994) ของข้อสอบในแบบสอบนั้น โดยการตัดข้อสอบที่ทำหน้าที่ต่างกันจะคำนึงถึงการรักษาโครงสร้างเดิมของแบบสอบ

2.3 ข้อมูลที่ใช้ในการวิเคราะห์ในงานวิจัยเรื่องนี้เป็นข้อมูลทุติยภูมิที่ได้จาก โครงการการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์ หรือ TIMSS ปี 2007 ซึ่งข้อจำกัดเรื่องหนึ่งที่ผู้วิจัยพบนั่นคือ ข้อจำกัดในเรื่องของตัวแปรทั้งในระดับสถานศึกษาและระดับนั กเรียน เนื่องจากเป็นข้อมูลที่มีการเก็บรวบรวมไว้ ดังนั้นตัวแปรบางตัวที่มีความน่าสนใจจากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง จึงจะไม่มีในฐานข้อมูล ข้อเสนอแนะสำหรับงานวิจัยในอนาคต ผู้วิจัยควรศึกษาโดยการเก็บข้อมูลด้วยตัวผู้วิจัยเอง โดยผู้วิจัยอาจทำการสร้างข้ อสอบพร้อมทั้งสังเคราะห์ตัวแปรให้เข้ากับบริบทของประเทศไทยอย่างแท้จริง จะช่วยสร้างคุณค่าของงานวิจัยได้มากยิ่งขึ้น

2.4 การเปรียบเทียบเทียบประสิทธิภาพ ของโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาคณิตศาสตร์ในงานวิจัยเรื่องนี้ มีเกณฑ์ในการตัดสินคุณภาพเป็นค่าสัมประสิทธิ์การทำนาย (R^2) จากโมเดลการประเมินคุณภาพการจัดการศึกษา วิชาคณิตศาสตร์ ข้อเสนอแนะสำหรับงานวิจัยในครั้งต่อไปผู้วิจัยควรพิจารณาถึงองค์ประกอบอื่นๆ ที่เกี่ยวข้องกับการประเมินคุณภาพของโมเดลการจัดการศึกษาวิชาคณิตศาสตร์ เพื่อที่จะเป็นเกณฑ์ในการพิจารณาตัดสินคุณภาพกา รจัดการศึกษาของโมเดลที่แตกต่างกัน เพื่อให้ได้มุมมองที่มากขึ้นนอกเหนือไปจากค่าสัมประสิทธิ์การทำนาย (R^2)

รายการอ้างอิง

ภาษาไทย

กาญจนา วัธนสุนทร . (2537). การพัฒนาเกณฑ์ตัดสินข้อสอบลำเอียงทางเพศ . วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาวิจัยและจิตวิทยาการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

เกษร หว่างจิตร. (2539). การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบสำ หรับแบบสอบคัดเลือกระดับบัณฑิตศึกษาวิชาภาษาไทยและภาษาอังกฤษด้วยวิธีแมนเทล -แฮนส์เซล . วิทยานิพนธ์ปริญญามหาบัณฑิต , ภาควิชาวิจัยและจิตวิทยาการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

จิติมา วรรณศรี. (2539). การเปรียบเทียบประสิทธิภาพในการตรวจสอบการทำหน้าที่ ต่างกันของข้อสอบด้วยวิธีแมนเทล -แฮนส์เซลกับวิธีซิบเทสท์ เมื่อความยาวของแบบทดสอบ ขนาดกลุ่มตัวอย่างและอัตราส่วนกลุ่มอ้างอิงและกลุ่มเปรียบเทียบต่างกัน. วิทยานิพนธ์ปริญญามหาบัณฑิต , ภาควิชาวิจัยและจิตวิทยาการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

ญาณภัทร สีหะมงคล. (2540). การเปรียบเทียบความสอดคล้องของผลการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันระหว่างวิธี Lord's^2 วิธี Raju's Area Measures และวิธี Closed Interval Area. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต, ภาควิชาการทดสอบและวัดผลทางการศึกษา บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.

ทัศนีย์ พีรมนตรี . (2529). การวิเคราะห์ความลำเอียงของแบบสอบวิชาคณิตศาสตร์โครงการตรวจสอบคุณภาพการศึกษาชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2526. วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาวิจัยและจิตวิทยาการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

นพมาศ พิพัฒนสุข. (2541). การเปรียบเทียบประสิทธิภาพระหว่างวิธีแมนเทล -แฮนส์เซลกับวิธีถดถอยโลจิสติกในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบเมื่อใช้เกณฑ์จับคู่เปรียบเทียบแตกต่างกันในแบบสอบชนิดพหุมิติ . วิทยานิพนธ์ปริญญามหาบัณฑิต , ภาควิชาวิจัยและจิตวิทยาการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

นิตยา เหมือดไธสง. (2543). การส่งผ่านตัวกลางเชิงสาเหตุของปัจจัยด้านนักเรียนด้านครูและด้านโรงเรียนไปยังผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ : การวิเคราะห์อภิมานงานวิจัย . วิทยานิพนธ์ปริญญามหาบัณฑิต , สาขาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

บุญเรือง ศรีเหรัญ. (2542). การศึกษาองค์ประกอบทางการศึกษาที่สัมพันธ์และมีอิทธิพลต่อผลการเรียนรู้และประสิทธิภาพของโรงเรียนโดยใช้รูปแบบระดับชั้นลดหลั่นแรกเชิงเส้น . วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต , สาขาวิจัยและพัฒนาหลักสูตร บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

ประกฤติยา ทักษิโณ . (2552). การประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน : การประยุกต์ใช้การทำหน้าที่ต่างกันของข้อสอบและโมเดลมูลค่าเพิ่ม . วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต , ภาควิชาวิจัยและจิตวิทยากา รศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

พิชิต ธรรมรักษ์. (2549). ปัจจัยที่มีผลต่อผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียนแผนการเรียนศิลป์ภาษาในกรุงเทพมหานคร. วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาวิจัยและจิตวิทยาทางการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

รักชนก ยี่สุ่นศรี . (2543). การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและแบบสอบด้วยกระบวนการ ดี เอฟ ไอ ที สำหรับแบบสอบคัดเลือกบุคคลเข้าศึกษาในสถาบันอุดมศึกษาวิชาภาษาอังกฤษและวิชาคณิตศาสตร์ . วิทยานิพนธ์ปริญญามหาบัณฑิต , ภาควิชาวิจัย บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

วลีมาศ แซ่อึ่ง. (2543). การเปรียบเทียบอำนาจการทดสอบและอัตราความคลาดเคลื่อนประเภทที่ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบบนเนกรูประหว่างวิธีซิปเทสท์ปรับใหม่วิธีซิปเทสท์ วิธีแมนเทล-แฮนส์เซล และวิธีการถอดถอยโลจิสติก. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต, ภาควิชาวิจัยและจิตวิทยาการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี.(2550). การวิเคราะห์พหุระดับ. พิมพ์ครั้งที่ 4. กรุงเทพมหานคร: สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี . (2548). ทฤษฎีการทดสอบแบบดั้งเดิม . พิมพ์ครั้งที่ 5. กรุงเทพมหานคร: โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี . (2550). ทฤษฎีการทดสอบแนวใหม่ . พิมพ์ครั้งที่ 3. กรุงเทพมหานคร : โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

สสวท. (2552). การศึกษาแนวโน้มผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์นานาชาติ. กรุงเทพมหานคร: สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.

สสวท. (2552). ตัวอย่างข้อสอบที่ใช้ในการประเมินผลการเรียนรู้วิชาคณิตศาสตร์. กรุงเทพมหานคร: สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.

สุชีรา มะหิเมือง. (2547). ปัจจัยที่ส่งผลผลสัมฤทธิ์และพัฒนาการทางวิชาการ: การวิเคราะห์มูลค่าเพิ่ม. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต, สาขาวิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

สุรศักดิ์ อมรรัตนศักดิ์. (2530). การศึกษาเปรียบเทียบผลของวิธีวิเคราะห์ความลำเอียงของข้อสอบที่แตกต่างกัน 4 วิธี. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิจัยและจิตวิทยาการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

เสรี ชัดแช้ม. (2540). การเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันแบบไม่สม่ำเสมอของข้อสอบระหว่างวิธีแมนเทล -แฮนส์เซลแบบปกติและวิธีแมนเทล -แฮนส์เซลแบบแบ่งกลุ่มความสามารถผู้สอบและความยาวของข้อสอบ. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต, ภาควิชาวิจัยและจิตวิทยาการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

อนงค์ อินตาพรหม. (2552). การวิเคราะห์พหุระดับของปัจจัยระดับครูและนักเรียนที่ส่งผลต่อผลสัมฤทธิ์การเรียนคณิตศาสตร์. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต, ภาควิชาวิจัยและจิตวิทยาการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

อิทธิฤทธิ์ พงษ์ปิยรัตน์. (2551). การวิเคราะห์ข้อสอบและการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ: การวิเคราะห์พหุระดับ. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต, ภาควิชาวิจัยและจิตวิทยาการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

ภาษาอังกฤษ

Angoff W. H. (1993). Perspectives on differential item functioning methodology. In Holland P. W. and Wainer H. (eds.). Differential Item Functioning, pp. 3-23. New Jersey: Lawrence Erlbaum Association.

Bandura, A. (1989). Human agency in social cognitive theory. American Psychologist 44: 1175 – 1184.

Camilli G., and Shepard L. A. (1994). Method for identifying biased test items. California: SAGE.

Cheong, Y. F. (2006). Analysis of school context effects on differential item functioning using hierarchical generalized linear models. International Journal of Testing 6(1): 57-79.

Chiu,M.M., Xihua,Z. (2008). Family and motivation effect on mathematics achievement : Analyses of students in 41 countries. Learning and Instruction 18: 321-336.

D'Agostino, J.V. (2000). Instructional and school effects on student' longitudinal reading and mathematics achievements. School Effectiveness and School Improvement 11: 197-235.

Green B. F., Crone C. R., and Folk V. G. (1989). A Method for studying differential distractor functioning. Journal of Educational Measurement 26(2): 147-160.

Hack, R.H. (2000). Examining the impact of school quality on school outcomes and improvement: A value-added approach. Educational Administration Quarterly 36(4): 513-552.

Hægeland, T., Kirkebøen, L.J. (2008). School performance and value added indicators what is the effect of controlling for socioeconomic background?. Research Department. [Online]. Available from http://www.ssb.no/english/subjects/04/90/doc_200808_ en/doc_200808_en.pdf [2008, March 3]

Hambleton K. R., Swaminathan H., and Rogers H J. (1991). Fundamentals of item response theory. California: SAGE.

Holland P. W., and Wainer H. (eds.). (1993). Differential item functioning. New Jersey: Lawrence Erlbaum Association.

Kato K., Moen R. E., and Thurlow M. L. (2009). Differentials of a state reading assessment: item functioning, distractor functioning, and omission frequency for disability categories. Educational Measurement: Issues and Practice 28(2): 28-40.

Koon S. (2010). A Comparison of methods for detecting differential distractor functioning. Doctoral dissertation, Florida State University.

Lord F. M. (1980). Applications of item response theory to practical testing problems. hillsdale. New Jersey: Lawrence Erlbaum Association.

Mellenbergh G. J. (1994). Generalized linear item response theory. Psychological Bullentin 115(2): 300-307.

Middleton K., and Laitusis C. C. (2007). Examining test items for differential distractor functioning among students with learning disabilities. Princeton, NJ.: Education Testing Service.

Monahan O. Patrick, and co. (2007). Odds Ratio, Delta, ETS classification, and standardization measures of DIF magnitude for binary logistic regression. Journal of Educational and Behavioral Statistics. 32(1): 92–109.

Narayanan P. and Swaminathan H. (1994). Performance of the Mantel-Haenzel and simultaneous Item bias proceders for detecting differential item functioning. Applied Psychological Measurement 18: 15-328.

Oshima, T. C., Raju, N. S., and Nanda, A. O. (2006). A new method for assessing the statistical significance in the differential functioning of items and tests (DFIT) framework. Journal of Educational Measurement 43: 1-17.

Palardy,G.J.(2008). Differential school effects among low, middle, and high social class composition schools: a multiple group, multilevel latent growth curve analysis. School Effectiveness and School Improvement 19(1).

Penfield R. D. (2008). An Odds Ratio approach for assessing differential distractor functioning effects under the nominal response model. Journal of Educational Measurement. 45(3): 247–269.

Penfield, R. D. (2007). DIFAS 4.0 Differential item functioning analysis system: User's manual. Unpublished manuscript.

Penfield R. D. (2010). DDFS 1.0, Differential Distractor Functioning Software User's Manual. Unpublished manuscript.

Postlethwaite, T.N.,(2004). Monitoring educational achievement. Paris, UNESCO. Institute for Educational Planning.

Potenza M. T., and Dorans N. J. (1995, March). DIF assessment for polytomously scored item: A framework for classification and evaluation. Applied Psychological Measurement 19(1): 23-37.

Raju N. S. (1990). Determining the significance of estimated signed and unsigned areas between two Item response functions. Applied Psychological Measurement. 19(4): 353-368.

Raju N. S., Fortmann-Johnson K. A., Kim W., Morris S. B., Nering M. L., and Oshima T.C. (2009). The item parameter replication method for detecting differential functioning in the polytomous DFIT framework. Applied Psychological Measurement 33: 133-147.

Shealy R., and Stout W. T. (1993). A Model-base standardization approach that separates true Bias/DIF as well as item Bias/DIF. Psychometrika 58(2): 159-194.

TIMSS. (2007). User guide for the International database. Publisher: TIMSS & PIRLS International Study Center. Lynch School of Education, Boston College.

Trautwein, Ludtke, Schnyder and Niggli. (2006). Predicting homework effort: Support for a domain-specific, multilevel homework model. Journal of Educational Psychology 98: 438-456.

Zwick R., Thayer D.T., and Mazzeo J., (1997). Descriptive and inferential procedures for assessing differential item functioning in polytomous items. Journal of Applied Measurement in Education 10: 321-344.

ภาคผนวก

ภาคผนวก ก

ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ และการทำหน้าที่ต่างกันของตัวลวง ฉบับที่ 1 – 14

การใช้โปรแกรม DDFS 1.0 (Differential Distractor Functioning Software)

By Randall D. Penfield (2010) University of Miami School of Education

ความหมายของ DIF และ DDF

การทำหน้าที่ต่างกันของข้อสอบ (DIF) หมายถึง การที่ข้อสอบทำให้ผู้ตอบจากต่างกลุ่มกันที่มีความสามารถหรือคุณลักษณะที่มุ่งวัดเท่ากัน มีโอกาสในการตอบข้อสอบได้ถูกแตกต่างกัน (ศิริชัย, 2550)

การทำหน้าที่ต่างกันของตัวลวง (DDF) หมายถึง การที่สมาชิกในกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่มีความสามารถเท่ากัน มีโอกาสในการเลือกตัวลวงแทนที่จะเลือกตัวถูกได้แตกต่างกัน (Penfield, 2010)

การใช้โปรแกรม DDFS

- System Requirements

The following hardware and software is required for running DDFS:

· Microsoft Windows 95 or later, or Microsoft Windows NT 3.51 or later

· 24 MB RAM for Windows 95 or later, 32 MB RAM for Microsoft NT or later

· Pentium 90MHz or higher

- **การเตรียมข้อมูลก่อนการวิเคราะห์**

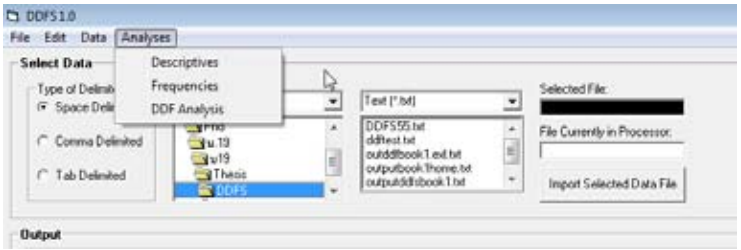
1. ข้อมูลต้องเป็นตัวเลขทั้งหมด และต้องเป็นตัวเลขจำนวนเต็ม 0-9 โดยแทนตำแหน่งของ MISSING ด้วย “.”
2. แถวแรกของข้อมูลเป็นส่วนของตัวเลือกที่ถูกต้อง(เฉลย) โดยผลการตอบของผู้สอบแต่ละคนเริ่มที่ 2

.	3	2	3	1	4	(this row contains the correct options)
1	2	4	3	2	3	(this row contains responses of examinee 1)
1	3	2	3	2	2	(this row contains responses of examinee 2)
2	3	1	3	2	3	(this row contains responses of examinee 3)
2	3	2	3	4	1	(this row contains responses of examinee 4)

3. ไฟล์ข้อมูลที่ใช้ต้องเป็น “.txt” หรือ “.dat” เท่านั้น โดยสามารถเตรียมจาก SPSS หรือ EXCEL

- **การวิเคราะห์ข้อมูล**

1. นำเข้าข้อมูลที่ได้เตรียมไว้ จากนั้นเลือก Import Selected Data File



2. เมนูการวิเคราะห์ประกอบด้วย 3 ส่วน

→ Descriptives
→ Frequencies
→ DDF Analysis

- **Descriptives** : จะคำนวณให้ค่า mean, standard deviation, minimum, maximum, and number of valid cases

DESCRIPTIVES

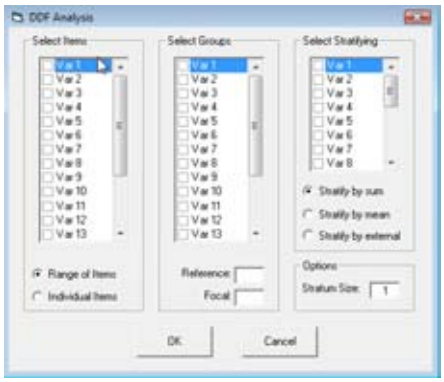
Name	Mean	SD	Min	Max	N
Var 1	1.5	0.501	1	2	400
Var 2	0.522	0.5	0	1	400
Var 3	0.507	0.501	0	1	400
Var 4	0.5	0.501	0	1	400

- **Frequencies** : จะคำนวณให้ค่า frequency, percentage of the total sample

FREQUENCIES AND PERCENTAGES FOR: Var 1

Value	Freq	% Total	% Pres	Cum %
1	200	50	50	50
2	200	50	50	100
Present	400	100		
Missing	0	0		
Total	400	100		

- DDF Analyses : **มีขั้นตอนในการวิเคราะห์ดังนี้**



1. เลือกฟังก์ชัน DDF Analyses จะปรากฏเมนูให้เลือก 3 ส่วน คือ

- "Select Items" เป็นการเลือกข้อสอบที่จะใช้วิเคราะห์
- "Select Group" เป็นการเลือกกลุ่มที่เป็นกลุ่มอ้างอิง กับกลุ่มเปรียบเทียบ
- "Select Stratify" เป็นการเลือกกรณีในการวิเคราะห์

2. ผลการวิเคราะห์ข้อมูลจะปรากฏอยู่ในช่อง output โดยจะแสดงทั้งผลการวิเคราะห์ DIF และ DDF โดยใช้สถิติดังนี้

- LOR (The Mantel-Haenszel common log-odds ratio)

- **DIF** : ค่าเป็นบวก กลุ่มอ้างอิงมีโอกาสตอบถูกมากกว่า และค่าเป็นลบ กลุ่มเปรียบเทียบมีโอกาสตอบถูกมากกว่า
- **DDF** : ค่าเป็นบวก กลุ่มเปรียบเทียบมีโอกาสเลือกมากกว่า และค่าเป็นลบ กลุ่มอ้างอิงมีโอกาสเลือกมากกว่า
  - SE(LOR) (The standard error of the Mantel-Haenszel common log-odds ratio)
    - Z(LOR) : Z(LOR) $\geq$ 2.0 หรือ Z(LOR) $\leq$ - 2.0 เกิด DIF หรือ DDF อย่างมีนัยสำคัญ
  - BD (The Breslow-Day chi-square test) ใช้ตรวจสอบความเป็น nonuniform DIF
- **ตัวอย่างการวิเคราะห์และแปลผล**

Variable: 3 Correct Option: 4

Option	LOR	SE(LOR)	Z(LOR)	BD	N-REF	N-FOC
DIF	0.4444	0.0354	12.5537	0.683	.	.
Correct	.	.	.	.	8759	6453
1	0.2889	0.0784	3.6849	.	400	430
2	0.318	0.0738	4.3089	.	415	442
3	0.4096	0.1322	3.0983	.	110	130
5	0.5115	0.0399	12.8195	.	1769	2327

กรณี DIF : ตัวเลือก 4 เป็นตัวเลือกที่ถูกต้อง เกิด DIF กลุ่มอ้างอิงมีโอกาสในการตอบถูกมากกว่า อย่างมีนัยสำคัญทางสถิติ(LOR = .444, Z(LOR) = 12.5537) แต่ไม่เกิด nonuniform DIF (BD < 3.84)

กรณี DDF : เกิด DDF ทุกตัวเลือกที่เป็นตัวลวง และตัวลวงทุกตัวกลุ่มเปรียบเทียบมีโอกาสในการเลือกตอบสูงกว่า

รายละเอียดการตัดข้อสอบในแบบสอบสอบแต่ละฉบับ

การตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) แต่ละฉบับ ผู้วิจัยจึงเลือกตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปเป็นบางข้อ ตามเงื่อนไขของโมเดลในการศึกษา โดยพิจารณาเลือกตัดข้อสอบตามเกณฑ์ 2 ข้อ ดังนี้ 1. เลือกข้อสอบที่ทำหน้าที่ต่างกันที่มีค่าอิทธิพลสูงสุด และ 2. คำนึงถึงโครงสร้างของแบบสอบ โดยภายหลังการตัดข้อสอบออกแล้วโครงสร้างของแบบสอบจะไม่แตกต่างจากโครงสร้างเดิม และข้อสอบที่ตัดออกไปควรมีจำนวนไม่เกินร้อยละ 20 (Clauser (1993), cited in Naraya & Swaminathan,1994) ของข้อสอบในแบบสอบนั้น โดยการตัดข้อสอบที่ทำหน้าที่ต่างกันจะคำนึงถึงการรักษาโครงสร้างเดิมโดยจำแนกตามสัดส่วนเนื้อหาของแบบสอบ โดยในงานวิจัยครั้งนี้ผู้วิจัยใช้โปรแกรม DDFS 1.0 (Penfield, 2010) ซึ่งเป็นโปรแกรมที่ใช้การวิเคราะห์ แมนเทล-แฮนส์เซลร่วมกับอัตราส่วนแต้มต่อ (odds ratio) โดยมีการพิจารณาค่าที่สำคัญคือค่า Z(LOR) ซึ่งเป็นค่าที่บอกถึงระดับการเกิดการทำหน้าที่ต่างกันของข้อสอบ (DIF) และการทำหน้าที่ต่างกันของตัวลวง (DDF) คือ $Z(LOR) \geq 2.0$ หรือ $Z(LOR) \leq -2.0$ เกิด DIF หรือ DDF อย่างมีนัยสำคัญ รายละเอียดดังตารางที่ 45

ตารางที่ 45 จำนวนข้อสอบที่ถูกตัดและช่วงค่า Z(LOR) ที่ใช้ในการพิจารณาการเกิด DIF และ DDF

ฉบับที่	การตัดข้อที่เกิด DIF		การตัดข้อที่เกิด DIF + DDF		
	จำนวนข้อ	Z(LOR)	จำนวนข้อ	Z(LOR)_DIF	Z(LOR)_DDF
1	3	(-2.76) – 2.63	5	(-3.77) – 2.82	(-2.78) – 2.95
2	3	(-3.60) – 3.23	4	(-2.42) – 2.53	(-2.04) – 2.23
3	3	2.06 – 2.59	4	(-3.78) – 2.66	(-3) – 3.04
4	2	(-2.99) – 2.31	4	(-2.41) – 3.39	(-2.83) – 3.24
5	6	(-2.72) – 2.79	4	(-3.22) – 2.06	(-3.09) - 2.17
6	2	(-2.06) – 2.41	2	(-2.01) – 2.20	(-2.17) – 2.75
7	3	2.06 – 2.49	4	(-2.63) – 2.08	(-3.12) – 2.79
8	4	(-2.09) – 3.24	5	(-3.29) – 2.36	(-3.12) – 2.79
9	3	2.08 – 3.23	5	(-2.83) – 2.74	(-2.66) – 3.24
10	5	(-2.73) – 2.75	4	(-2.38) – 2.88	(-2.47) – 2.89
11	3	(-4.52) – 2.46	5	(-2.50) – 2.79	(-3.58) – 3.02
12	5	(-2.39) – 2.89	4	(-3.31) - 2.76	(-3.07) – 2.41
13	2	(-2.01) – 2.16	4	(-2.24) – 2.31	(-2.29) – 2.02
14	2	(-2.05) – 2.28	4	(-3.26) – 2.64	(-3.07) – 2.26

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 1

ข้อ	ข้อสอบ	Option	Response	GIRL		HOMEWK		SES		EQUIP		RICH	
				LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)
1	m022043 (Number)	DIF	4	-0.25	-0.98	0.1107	0.4092	0.1461	0.5248	-.468	-1.6444	0.2722	1.0469
		DDF**	1	0.51	1.56	0.2002	0.6076	0.373	1.0987	-0.6622	-1.9267	0.7161**	2.0513
			2	-0.5	-0.84	-0.1996	-0.3234	0.9766	1.4837	-1.4985**	-2.2389	-0.1945	-0.3388
			3	0.34	0.77	6904	1.4024	-1.4017	-1.8521	0.0421	0.0788	-0.4333	-0.967
			5	-0.37	-0.48	-0.6467	-0.553	0.1701	0.2	0.8763	0.7201	1.0673	1.1963
2	m022046 (Number)	DIF		.3862	1.5718	-.0293	-.1185	.2977	1.1075	-.267	-1.0064	-.3228	-1.285
		DDF											
3	m022049 (Geometry)	DIF**	4	-0.8534**	-3.7678	0.1975	0.8751	0.6302**	2.537	-0.6534**	-2.5494	0.2266	1.0226
		DDF**	1	-0.7767**	-2.1129	-0.327	-0.7893	0.6766	1.7805	-1.1119**	-2.7763	-0.2949	-0.8302
			2	-0.797**	-2.776	-0.0491	-0.1643	0.5457	1.6863	-0.2445	-0.7384	0.3299	1.1345
			3	-0.8788**	-2.7549	0.7179**	2.374	0.6353	1.9118	-0.8631**	-2.5326	0.3771	1.2271
4	m022050 (Algebra)	DIF	5	0.3673	1.4992	-0.0618	-0.2472	0.1737	0.6879	0.0272	0.1023	0.1688	0.676
		DDF**	1	0.9731	1.4136	0.2319	0.3131	0.1043	0.1466	-0.3751	-0.4997	0.0326	0.0499
			2	0.2552	0.8187	-0.4053	-1.2106	0.1394	0.4206	-0.3458	-0.9974	0.2464	0.7549
			3	0.6214**	2.0327	0.1203	0.3916	0.1882	0.6046	0.1896	0.5591	0.0113	0.037
			4	-0.0677	-0.1281	-0.4623	-0.91	0.1013	0.2083	0.1153	0.2316	0.4115	0.7429

5	m022055 (Geometry)	DIF		-.6261	-1.7262	-.5585	-1.593	-.1179	-.331	-.1481	-.4066	.1222	.3301
		DDF											
6	m022057 (Number)	DIF**	3	-0.1574	-0.682	-0.2631	-1.0832	0.5124	1.9753	0.4321	1.6184	0.4926**	2.0611
		DDF**	1	0.0213	0.0572	-0.2928	-0.7589	0.4709	1.0951	0.0254	0.0651	0.9537**	2.1494
			2	-0.4937	-1.2708	-0.9254**	-2.0693	0.3279	0.7835	0.7243	1.4867	-0.0649	-0.1796
			4	0.0924	0.2484	0.0989	0.2531	0.355	0.841	0.6785	1.4281	0.4314	1.1395
			5	-0.7721	-1.0955	-1.1161	-1.3698	0.4915	0.605	0.3523	0.4435	0.218	0.3182
7	m022257 (Data and Chance)	DIF**	3	0.4728**	2.1004	0.2596	1.0898	-0.6066**	-2.4759	0.1836	0.7251	-0.0791	-0.3459
		DDF**	1	-0.2258	-0.5959	0.5345	1.359	0.4861	1.2092	-0.1754	-0.4409	-0.1196	-0.3181
			2	0.6888**	2.4786	0.416	1.453	-1.0386**	-3.183	0.4794	1.4315	-0.209	-0.7456
			4	0.6124	1.6361	-0.2398	-0.545	-0.944**	-2.0793	-0.087	-0.2027	-0.1373	-0.3606
			5	0.4374	0.673	-0.3086	-0.4493	-0.0169	-0.0234	-0.803	-1.0797	0.3919	0.6396
8	m022062 (Geometry)	DIF**	2	-0.8712**	-2.6872	-0.0662	-0.2102	-0.3955	-1.2132	-0.106	-0.3324	0.0171	0.0536
		DDF**	1	-0.8796**	-2.6842	-0.0756	-0.2379	-0.3683	-1.1253	-0.1355	-0.4174	0.024	0.075
			3	-0.1653	-0.2479	-0.5804	-0.8256	-0.6694	-0.8663	0.8784	1.2044	4027	0.5748
			4	-2.5177	-1.6605			0.3646	0.3668	0.5046	0.5787	1.3218	1.0527
			5	-0.07	-0.0914	-0.494	-0.5933	-1.5023	-1.24	3.1427**	2.0981	0.6931	0.9126

9	m022066 (Number)	DIF**	4	0.8605**	2.4978	-0.3054	-0.9482	-0.6425	-1.9048	0.1881	0.5393	-0.3377	-0.99
		DDF**	1	1.0732**	2.9565	-0.2909	-0.8676	-0.6354	-1.8212	0.1714	0.4677	-0.3568	-1.0022
			2	-0.5436	-0.6482	0.0804	0.1006	0.7572	1.0469	-0.0147	-0.0194	0.4163	0.5274
			3	0.5563	0.7406	-1.1456	-1.562	-1.0727	-1.3117	0.0873	0.1105	-0.7133	-1.0599
10	m022232 (Number) (Polytomous)	DIF		-.033	-.077	-.009	-.022	-.068	-.162	-.33	-.782	.388	.858
		DDF											
11	m022234a (Geometry) (Polytomous)	DIF		.142	.275	-.421	-.877	.16	.337	-.217	-.395	-.378	-.744
		DDF											
12	m022234b (Number) (Polytomous)	DIF		.201	.389	-.045	-.094	-.179	-.339	-.266	-.468	.513	.955
		DDF											
13	m022243 (Geometry)	DIF**		.9187**	2.6294	.1744	.5413	-.226	-.6706	.1568	.4762	-.9113**	-2.7252
		DDF											
14	m042003 (Number)	DIF	4	-0.1513	-0.5688	0.1058	0.3756	0.121	0.3992	0.1427	0.4874	-0.0938	-0.3481
		DDF**	1	0.8386	1.4308	-0.968	-1.3876	-0.7472	-0.9259	1.055	1.5197	-1.2084**	-2.0983
			2	-0.3814	-1.1586	-0.047	-0.1359	0.1909	0.5061	-0.0845	-0.2455	0.022	0.0659
			3	-0.4792	-1.0583	0.5136	1.1095	0.2361	0.4746	0.9279	1.5553	0.1645	0.3677

15	m042079 (Number)	DIF	3	0.335	1.4052	0.1026	0.4051	-0.2149	-0.8128	0.0929	0.3373	0.0435	0.1819
		DDF	1	1.3076	2.1039	0.7985	1.396	-0.4726	-0.6686	0.5677	0.6969	-0.5726	-0.9595
			2	0.3704	1.0243	0.3354	0.9019	0.3789	0.9693	-0.2463	-0.5984	0.5071	1.3271
			4	0.2984	1.1019	0.0079	0.0268	-0.5359	-1.7067	0.1314	0.4193	-0.055	-0.2042
16	m042018 (Number)	DIF		-.1933	-.6712	.4296	1.4282	.2826	.9026	-.0863	-.2767	.3953	1.3501
		DDF											
17	m042055 (Number)	DIF**	2	0.3868	1.7614	0.0502	0.2206	0.7176**	2.8174	-0.3705	-1.4982	0.1666	0.7431
		DDF**	1	0.0698	0.1559	0.3098	0.6704	0.393	0.702	0.0486	0.0899	-0.288	-0.6695
			3	0.4438	1.4933	0.0413	0.1319	0.6034	1.7505	-0.0503	-0.144	0.0102	0.0338
			4	0.4673	1.7307	-0.0734	-0.2612	0.8874**	2.9541	-0.6963**	-2.3953	0.3102	1.1063
18	m042039 (Number)	DIF**	1	0.0571	0.2342	0.1953	0.7582	-0.5422**	-2.0951	0.063	0.2322	-0.4579	-1.8156
		DDF**	2	0.4678	1.2852	0.2488	0.6302	-0.2442	-0.613	-0.0003	-0.0007	-0.0285	-0.0735
			3	0.0324	0.11	0.2693	0.8335	-1.15**	-3.3943	0.4403	1.2244	-0.9707**	-3.2005
			4	0.117	0.3322	0.3966	1.1566	-0.0191	-0.0534	-0.4704	-1.3217	0.4331	1.1145
19	m042199 (Algebra)	DIF**	4	0.4878**	2.0461	-0.0886	-0.3521	-0.1232	-0.4644	0.3737	1.3644	-0.1758	-0.7319
		DDF**	1	0.2248	0.5186	0.3596	0.7958	-0.231	-0.4664	1.2244	1.9306	-0.4522	-1.0923
			2	0.3069	1.0102	0.0196	0.06	0.08	0.2293	-0.1427	-0.4097	0.0035	-0.0108
			3	0.6411**	2.1207	-0.355	-1.0738	-0.1153	-0.338	0.3896	1.1003	0.0711	0.2373

20	m042301a (Algebra)	DIF**		-.2046	-.7872	.218	.841	.2732	1.0052	.0179	.0655	.6091**	2.3002
		DDF											
21	m042301b (Algebra)	DIF		-.0694	-.2143	.2252	.6757	.0615	.1795	-.223	-.6421	.5542	1.5962
		DDF											
22	m042301c (Algebra)	DIF		.3142	.5486	.1821	.3485	-.2618	-.4459	.3363	.6232	.4942	.7711
		DDF											
23	m042263 (Algebra)	DIF		-.4309	-1.0528	.2371	.5789	-.1642	-.3829	-.067	-.1577	-.0154	.389
		DDF											
24	m042265 (Geometry)	DIF**	3	-0.2019	-0.8025	0.2852	1.0961	-0.5888**	-2.2508	0.0592	0.2202	-0.4083	-1.5495
		DDF**	1	0.3034	0.6956	0.2188	0.4777	-0.3566	-0.7401	-0.3138	-0.7187	-0.1157	-0.2515
			2	-0.3578	-0.9686	0.3292	0.8261	-0.0994	-0.2557	0.1228	0.2983	-0.0822	-0.203
			4	-0.0705	-0.2524	0.2868	0.9795	-0.8711**	-2.8645	0.025	0.08	-0.649**	-2.2249
25	m042137 (Geometry)	DIF	3	0.0283	0.1159	0.067	0.2566	-0.1528	-0.5866	0.0339	0.1265	-0.089	-0.3516
		DDF	1	0.3811	1.0174	0.5354	1.2368	-0.0218	-0.0527	0.5921	1.3045	-0.4784	-1.2273
			2	0.1279	0.4505	0.0194	0.0648	-0.6638	-2.0763	-0.0403	-0.13	-0.1359	-0.479
			4	-0.9015	-1.6191	-0.7289	-1.1453	-0.0222	-0.0445	-0.2041	-0.3791	0.0225	0.041

26	m042148 (Geometry)	DIF**	2	-0.5333**	-2.1478	-0.4058	-1.5889	0.0103	0.0376	-0.0459	-0.1639	-0.1906	-0.7805
		DDF	1	-0.5859	-1.7386	-0.7058	-1.9395	-0.007	-0.0178	0.1601	0.417	-0.1359	-0.4287
			3	-0.5678	-1.7133	-0.0768	-0.2233	0.4326	1.198	-0.1411	-0.3748	0.0438	0.1333
			4	-0.4149	-0.9556	-0.009	-0.0191	-0.6055	-1.0859	0.0396	0.0744	-0.1053	. -.2368
27	m042254 (Data and Chance)	DIF**	1	-0.0163	-0.0579	-0.5396	-1.691	0.7306**	2.2487	0.3401	0.9762	0.1685	0.5873
		DDF**	2	0.9326	1.6236	-0.3149	-0.5582	0.8551	1.4425	0.0336	0.055	0.1242	0.2271
			3	-0.205	-0.4698	-0.5555	-1.1451	1.0997**	2.2963	0.1408	0.2733	0.6908	1.4639
			4	-0.1487	-0.2929	-0.589	-0.9676	0.4444	0.69	0.4542	0.6791	-0.0378	-0.0768
28	m042250 (Data and Chance)	DIF**		-.2159	-.7001	-.9594**	-2.7593	-.8494**	-2.2596	.6968	1.8155	-.5808	1.217
		DDF											
29	m042220 (Data and Chance) (Polytomous)	DIF		-.164	-.432	.29	.738	.056	.141	.579	1.385	-.363	-.931
		DDF											

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 2

ข้อ	ข้อสอบ	Option	Response	GIRL		HOMEWK		SES		EQUIP		RICH	
				LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)
1	m042003 (Number)	DIF	4	-0.1854	-0.7189	0.3827	1.3967	-0.5862	-1.8177	-0.2344	-0.7304	-0.0469	-0.1807
		DDF	1	0.2466	0.527	0.4197	0.9391	-0.105	-0.1831	0.2875	0.4547	0.0042	0.0093
			2	-0.0508	-0.1471	0.4271	1.159	-0.9853	-1.8981	0.0006	0.0013	0.2777	0.7842
			3	-0.2656	-0.719	0.1999	0.4909	-0.6679	-1.3667	-0.3293	-0.7375	-0.4772	-1.2395
2	m042079 (Number)	DIF	3	0.0758	0.3436	0.1704	0.7106	-0.2456	-0.9403	0.5531	1.9924	-0.176	-0.7554
		DDF**	1	0.6278	1.1611	-0.0439	-0.0684	-1.646	-1.4716	0.2951	0.3588	0.0032	0.0058
			2	0.6912	1.9313	0.109	0.2753	-0.5742	-1.2367	0.353	0.7949	-0.8626**	-2.3126
			4	-0.2423	-0.9399	0.3079	1.151	-0.0377	-0.1304	0.5805	1.7895	0.0617	0.2267
3	m042018 (Number)	DIF**		-1.0158**	-3.4539	-.1975	-.713	-1.0332**	-3.6025	.4101	1.3629	-.1975	-.6822
		DDF											
4	m042055 (Number)	DIF	2	-0.1285	-0.601	0.1305	0.5828	-0.0431	-0.1748	0.0409	0.1651	0.3409	1.5328
		DDF	1	-0.3658	-0.8346	-0.4073	-0.8537	-0.4277	-0.7244	-0.7927	-1.6287	0.4712	1.0867
			3	-0.2984	-0.9937	0.3211	1.0107	-0.0498	-0.1402	0.1636	0.4336	0.2704	0.8977
			4	0.0592	0.2166	0.0754	0.2662	0.0583	0.1905	0.0511	0.1617	0.2477	0.858

5	m042039 (Number)	DIF	1	0.1702	0.7127	-0.0832	-0.3324	0.068	0.2508	-0.3023	-1.0941	-0.2867	-1.1579
		DDF**	2	0.5061	1.264	0.2822	0.6079	-0.1218	-0.2468	0.1686	0.3231	-0.3778	-0.9502
			3	0.0697	0.2312	-0.216	-0.6866	-0.4634	-1.3039	-0.2596	-0.7388	-0.5178	-1.6994
			4	0.0324	0.0887	0.2099	0.5668	1.0606**	2.4961	-0.7976**	-2.0901	-0.0363	-0.0954
6	m042199 (Algebra)	DIF	4	0.4714	1.8744	-0.1049	-0.4061	0.1192	0.4211	0.263	0.9326	-0.1657	-0.6485
		DDF	1	0.6147	1.5929	-0.0313	-0.0787	0.6634	1.5657	0.8174	1.5753	-0.3156	-0.8283
			2	0.2327	0.7013	-0.1569	-0.4653	0.2073	0.5389	-0.3096	-0.8177	-0.0321	-0.0941
			3	0.3997	1.2017	-0.1331	-0.3843	-0.2812	-0.7361	0.3205	0.8314	-0.113	-0.3353
7	m042301a (Algebra)	DIF**		.3905	1.5558	.4318	1.7074	.9581**	3.2292	-.1047	-.3485	.0587	.2348
		DDF											
8	m042301b (Algebra)	DIF		-.0464	-.1385	-.0735	-.2259	.7261	1.9817	-.7984**	-2.1445	.4125	1.1709
		DDF											
9	m042301c (Algebra)	DIF		-.8716	-1.3672	.2351	.3992	-.3228	-.5702	-.7656	-1.2976	-.2357	-.3687
		DDF											
10	m042263 (Algebra)	DIF**		.094	.2785	-.3897	-1.1455	-.0681	-.1967	.6594	1.9515	-.9219**	-2.3105
		DDF											

11	m042265 (Geometry)	DIF	3	-0.4185	-1.6855	-0.0461	-0.1799	-0.2985	-1.1019	0.29	1.0368	-0.2723	-1.0857
		DDF	1	-0.4537	-0.948	0.4295	0.9211	0.3545	0.6685	-0.6692	-1.2643	0.897	1.6259
			2	-0.0306	-0.0863	-0.1095	-0.2828	-0.4998	-1.1403	0.4755	1.0281	-0.2769	-0.7316
			4	-0.5457	-1.8574	-0.0936	-0.3158	-0.2861	-0.9085	0.4689	1.3926	-0.5478	-1.935
12	m042137 (Geometry)	DIF**	3	0.522**	2.1914	-0.3301	-1.3462	0.3585	1.2444	-0.0115	-0.0426	0.3545	1.4317
		DDF**	1	0.6651	1.5507	-0.908	-1.9344	0.0111	0.0188	-0.1454	-0.2822	-0.2005	-0.4544
			2	0.5508**	2.0816	-0.2786	-0.9915	0.5084	1.5355	0.1601	0.5097	0.4322	1.5768
			4	0.1186	0.2661	-0.4628	-1.054	0.6313	1.2728	0.2281	0.4715	0.6243	1.2549
13	m042148 (Geometry)	DIF**	2	-0.1303	-0.5614	0.1623	0.6799	-0.1525	-0.5899	-0.3421	-1.2652	0.5457**	2.2411
		DDF**	1	-0.7752**	-2.04	0.0429	0.1138	-0.2213	-0.5374	0.0083	0.0184	0.6735	1.7719
			3	0.0784	0.2596	0.1854	0.618	-0.7284	-1.93	-0.6555	-1.8697	0.2968	0.9706
			4	0.2365	0.6845	0.0875	0.2395	0.2419	0.6334	-0.2715	-0.6837	1.1124**	2.6993
14	m042254 (Data and Chance)	DIF	1	0.1506	0.4931	0.0082	0.0243	0.1407	0.3611	-0.079	-0.1948	-0.1383	-0.44
		DDF**	2	1.698**	2.346	-0.0903	-0.1392	0.0123	0.015	-0.2685	-0.3634	0.1601	0.2864
			3	-1.185	-2.0811	-0.3901	-0.6621	-0.0763	-0.1278	-0.211	-0.3326	-0.2796	-0.5957
			4	0.1862	0.4186	0.455	1.0149	0.0399	0.0712	-0.1917	-0.3218	-0.5312	-1.1257

15	m042250 (Data and Chance)	DIF		.2786	.8309	-.3124	-.8743	-.1402	-.3343	-.7337	-1.6491	-.534	-1.5803
		DDF											
16	m042220 (Data and Chance) (Polytomous)	DIF											
		DDF											
17	m022097 (Number)	DIF**	3	-0.7702**	-2.9397	0.317	1.2136	-0.0334	-0.1168	-0.3268	-1.1154	0.1193	0.4801
		DDF**	1	-0.8266**	-2.9147	0.2651	0.9509	-0.0593	-0.1938	-0.2986	-0.9482	0.072	0.2751
			2	-1.4031	-1.8308	0.7781	1.0588	-1.3918	-1.2899	0.1732	0.1802	-0.0351	-0.046
			4	-0.4233	-0.8289	0.54	0.9705	-0.0349	-0.0574	-0.2262	-0.3576	-0.1055	-0.1966
18	m022101 (Data and Chance)	DIF	1	-0.0235	-0.1019	0.0555	0.2283	0.3966	1.4776	0.1208	0.4536	0.0208	0.0869
		DDF**	2	1.0739**	2.0039	-0.6642	-1.1551	1.1727	1.8275	-0.6963	-1.2095	0.0403	0.0823
			3	0.107	0.3501	-0.2484	-0.7578	0.513	1.423	0.2516	0.7067	0.1573	0.4741
			4	-0.0931	-0.3277	0.2056	0.679	0.1765	0.5233	0.2831	0.8031	-0.2077	-0.7142
19	m022104 (Number)	DIF**	4	-0.6452**	-2.2552	-0.3328	-1.1877	-0.2116	-0.6977	0.1614	0.5382	0.0929	0.3196
		DDF	1	-0.5738	-1.845	-0.2007	-0.6652	-0.3747	-1.1256	0.038	0.1157	-0.0467	-0.1523
			2	1.0721	0.9197	-0.3695	-0.3784	0.2464	0.2166	1.2694	0.8622	0.2578	0.2748
			3	-0.815	-1.8298	-0.942	-1.8802	0.1614	0.3364	0.5495	1.0687	0.6326	1.2554

20	m022105 (Geometry)	DIF**	4	-0.5541**	-2.4249	0.1106	0.4645	-0.0109	-0.0429	-0.2867	-1.0959	-0.1246	-0.5348
		DDF**	1	-0.5626	-1.7274	0.057	0.1766	-0.2006	-0.4968	-0.066	-0.1736	-0.1752	-0.5366
			2	-0.3578	-1.0852	-0.1225	-0.3543	0.6334	1.8046	-0.5017	-1.2917	0.281	0.8006
			3	-0.117	-0.3012	0.2704	0.6025	-0.786	-1.3531	-0.0108	-0.022	-0.1841	-0.4605
			5	-0.8295**	-2.0002	0.4504	1.0856	0.0972	0.2398	-0.9284**	-2.0324	-0.1057	-0.2714
21	m022106 (Number)	DIF		-.2008	-.7664	.1698	.6446	-.0656	-.2292	-.1067	-.3611	.2804	1.0086
		DDF											
22	m022108 (Geometry)	DIF	3	0.0668	0.2969	0.0346	0.1488	0.1535	0.5913	-0.1732	-0.6578	-0.0023	-0.0099
		DDF	1	-0.659	-0.8641	-1.6082	-1.5627			1.1556	0.8792	0.192	0.2598
			2	-0.0525	-0.1822	-0.0715	-0.2369	0.0105	0.0315	-0.3534	-1.0693	-0.0903	-0.3019
			4	0.4031	1.1866	0.1565	0.4555	0.2574	0.6781	0.1001	0.2438	-0.2281	-0.6737
			5	0.302	0.611	0.0715	0.138	-0.1005	-0.1543	-0.0589	-0.0949	0.0735	0.1411
23	m022110 (Number)	DIF		.4192	1.6413	-.0641	-.246	-.1527	-.5467	-.346	-1.1068	.3605	1.365
		DDF											
24	m022181 (Data and Chance)	DIF	2	0.1283	0.4667	-0.0517	-0.1793	0.5616	1.6378	-0.2366	-0.691	-0.0737	-0.2642
		DDF	1	0.0355	0.0867	-0.6398	-1.2822	0.7475	1.4706	-0.9236	-1.8498	-0.5312	-1.1841
			3	0.1692	0.3865	0.1409	0.3146	0.7452	1.3076	-0.3403	-0.6528	0.0235	0.0521
			4	0.556	1.1032	0.0959	0.1947	0.9503	1.5409	0.4905	0.589	0.5899	1.126

25	m032307 (Number)	DIF		.7348	1.4397	.5617	1.1644	.4255	.8042	.7104	1.3714	.4567	.9058
		DDF											
26	m032523 (Number)	DIF	2	-0.2595	-0.9485	-0.0502	-0.1821	-0.1878	-0.6256	-0.0804	-0.2649	-0.2571	-0.9078
		DDF	1	-0.4304	-1.1986	-0.0504	-0.1293	-0.403	-0.9184	-0.1539	-0.3351	-0.4342	-1.1249
			3	-0.404	-1.1868	0.0428	0.1275	0.0797	0.2221	-0.191	-0.5462	-0.078	-0.2311
			4	0.1059	0.3107	-0.0627	-0.1876	-0.244	-0.6604	-0.2086	-0.5576	-0.2015	-0.5774
27	m032701 (Number)	DIF**	3	0.7777**	2.5258	-0.0176	-0.0567	-0.1663	-0.4279	0.0745	0.1866	0.3191	1.013
		DDF**	1	1.4314**	2.2331	-0.4275	-0.7305			0.2766	0.3345	-0.1851	-0.3098
			2	0.5663	1.2432	0.1192	0.2474	-0.2373	-0.4002	0.719	0.9782	0.7206	1.3922
			4	0.0911	0.1778	0.015	0.0286	0.1756	0.269	-0.7373	-1.1479	0.1959	0.3895
28	m032704 (Number)	DIF	2	0.435	1.8201	-0.0453	-0.182	-0.2516	-0.924	0.1118	0.3972	-0.1927	-0.8043
		DDF	1	0.4931	1.4427	0.29	0.8098	-0.2723	-0.6577	0.2674	0.5924	0.0122	0.035
			3	0.3209	1.0123	-0.1631	-0.4834	-0.146	-0.407	0.0565	0.1526	-0.4792	-1.5266
			4	0.1149	0.2411	0.3218	0.6971	-0.3417	-0.5571	0.2799	0.4937	0.2873	0.5985

29	m032525 (Number)	DIF**	1	0.027	0.107	-0.065	-0.2581	0.642**	2.116	0.0122	0.0408	0.0689	0.2729
		DDF**	2	-0.4159	-1.1407	-0.205	-0.5786	0.1057	0.2326	0.4295	0.9205	-0.3542	-0.9847
			3	-0.0595	-0.1511	-0.2417	-0.6029	0.6746	1.3725	-0.6312	-1.1851	0.5925	1.4024
			4	0.1978	0.6139	0.2981	0.9264	0.775**	2.1163	-0.2673	-0.7396	0.2308	0.7217
30	m032579 (Geometry)	DIF	2	0.4091	1.7117	-0.3448	-1.3836	-0.1576	-0.5694	0.357	1.2705	-0.2026	-0.8414
		DDF	1	0.4198	1.5519	-0.4337	-1.4747	-0.1391	-0.4378	0.304	0.9488	-0.1439	-0.5165
			3	-0.0684	-0.1822	0.021	0.0555	0.1746	0.3943	0.4966	1.067	-0.3303	-0.8995
			4	0.8393	1.1251	-0.0615	-0.0892	-0.4656	-0.5239			-0.293	-0.4364
31	m032691 (Geometry)	DIF		-.0776	-.2674	-.1974	-.6978	-.1849	-.5963	.6106	1.8627	-.0634	-.216
		DDF											

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 3

ข้อ	ข้อสอบ	Option	Response	GIRL		HOMEWK		SES		EQUIP		RICH	
				LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)
1	m022097 (Number)	DIF	3	-0.4572	-1.8854	0.1733	0.6924	-0.0116	-0.043	-0.4327	-1.5481	0.1921	0.7822
		DDF**	1	-0.5661**	-2.1756	0.2664	1.0011	-0.199	-0.6862	-0.377	-1.2784	0.0953	0.3698
			2	-1.0853	-1.197	-0.4298	-0.5263	1.4727	1.7422	-1.2234	-1.1765	0.3285	0.4134
			4	0.6009	0.988	-0.495	-0.7129	0.276	0.3981	-0.5064	-0.7478	0.9017	1.3414
2	m022101 (Data and Chance)	DIF**	1	0.0728	0.3064	0.5155**	2.0604	-0.0252	-0.0976	0.0303	0.1117	0.0739	0.3087
		DDF	2	0.4864	0.962	0.2935	0.5516	-1.1654	-1.5694	0.132	0.201	-0.3653	-0.729
			3	-0.0147	-0.0487	0.5014	1.6396	0.1441	0.4541	0.1106	0.3144	0.1879	0.6183
			4	0.1636	0.5187	0.5793	1.6904	-0.0594	-0.1661	-0.0647	-0.1832	-0.0361	-0.1133
3	m022104 (Number)	DIF**	4	-0.3954	-1.4644	-1.0389**	-3.7806	-0.1667	-0.5711	0.027	0.0916	-0.3991	-1.3989
		DDF**	1	-0.2569	-0.8908	-0.7926**	-2.7322	-0.2061	-0.6495	0.0404	0.1253	-0.2967	-0.9683
			2	-0.5153	-0.7582	-1.2573	-1.4776	-1.4571	-1.2402	1.345	1.2944	-0.9623	-1.1625
			3	-0.3534	-0.7598	-1.7223**	-3.0386	0.5389	1.1229	-0.2907	-0.6299	-0.35	-0.8061

4	m022105 (Geometry)	DIF	4	-0.1912	-0.8368	0.3647	1.5499	0.3688	1.4554	-0.0389	-0.155	-0.0783	-0.3323
		DDF	1	-0.3231	-0.9682	0.4927	1.4787	0.3776	1.0498	-0.3132	-0.9494	-0.1316	-0.3975
			2	0.0658	0.2035	0.3238	0.9945	0.2829	0.7698	-0.2126	-0.5519	0.0369	0.1067
			3	0.4647	0.8863	-0.8254	-1.2364	-0.0581	-0.0895	1.8822	1.6381	-1.0201	-1.8403
			5	-0.3365	-0.906	0.495	1.2249	0.3442	0.8367	0.1205	0.285	0.0665	0.1767
5	m022106 (Number)	DIF		.1977	.7309	.1201	.4326	-.0243	-.0871	.2574	.9038	.5135	1.8379
		DDF											
6	m022108 (Geometry)	DIF**	3	-0.4766**	-2.094	-0.0333	-0.1446	0.0185	0.0742	-0.0896	-0.3549	-0.1615	-0.7093
		DDF**	1	-2.1846	-1.7751	0.0625	0.0721	0.6643	0.7353	-2.1054	-1.9917	-0.2627	-0.3055
			2	-0.6801**	-2.2211	-0.1647	-0.557	0.0397	0.125	0.023	0.0697	-0.039	-0.1318
			4	-0.1571	-0.504	0.1699	0.5111	-0.152	-0.4163	0.1481	0.3988	0.0102	0.0325
			5	-0.2052	-0.3776	-0.4571	-0.7919	-0.8008	-1.1312	-0.6293	-1.1186	-1.0813	-1.9416
7	m022110 (Number)	DIF		.2529	.9898	-.0587	-.2278	-.1342	-.5013	.4051	1.4325	-.3049	-1.1855
		DDF											

8	m022181 (Data and Chance)	DIF**	2	0.8421**	2.6598	0.1607	0.4911	0.1657	0.4605	-0.2583	-0.7357	0.1065	0.3433
		DDF**	1	0.5367	1.1347	0.1154	0.21	0.221	0.3519	0.2458	0.3846	-0.1204	-0.2385
			3	1.5156**	3.0409	-0.0817	-0.1631	-0.2751	-0.4753	-0.9039	-1.8194	-0.1008	-0.2237
			4	0.7907	1.417	0.7197	1.3081	0.2351	0.4133	0.6625	1.0152	0.6422	1.1628
9	m032307 (Number)	DIF**		.6815	1.2986	1.4407**	2.3339	-.27	-.5618	.4734	.9941	-.6866	-1.1545
		DDF											
10	m032523 (Number)	DIF**	2	-0.91**	-3.3505	-0.2864	-1.0848	-0.2205	-0.792	0.5457	1.9587	-0.3933	-1.4214
		DDF**	1	-0.7057**	-2.0043	-0.4476	-1.2163	-0.2814	-0.7363	1.0137**	2.473	-0.0576	-0.1604
			3	-1.0397**	-3.0032	-0.4733	-1.4116	-0.296	-0.8433	0.5679	1.612	-0.3984	-1.1843
			4	-0.9499**	-2.9629	-0.0997	-0.3209	-0.0411	-0.1272	0.1712	0.5245	-0.5277	-1.5996
11	m032701 (Number)	DIF	3	-0.0483	-0.1291	-0.4972	-1.1754	0.1119	0.244	-0.5873	-1.3333	-0.1137	-0.3094
		DDF	1	-1.5141	-1.2676	-0.1364	-0.1345	0.5798	0.4522	-1.3643	-1.434	0.387	0.428
			2	0.5192	0.8907	-3.1623	-1.9189	-1.3848	-1.1584	0.3703	0.4592	-0.4137	-0.7129
			4	-0.0762	-0.1642	0.0192	0.0403	0.5334	1.0295	-0.7367	-1.3848	-0.0527	-0.1145

12	m032704 (Number)	DIF**	2	0.7068**	2.6492	0.1302	0.4837	0.3747	1.2885	-0.0206	-0.0691	0.2597	1.0105
		DDF**	1	0.7876**	2.2671	0.1237	0.3475	0.5671	1.5107	0.1574	0.4003	0.1172	0.3501
			3	0.4172	1.0659	-0.1642	-0.4009	-0.1726	-0.367	-0.4307	-0.951	0.2653	0.6882
			4	0.5699	1.1816	0.5375	1.1749	0.4163	0.814	0.0024	0.0044	0.2964	0.6932
13	m032525 (Number)	DIF	1	0.3951	1.561	0.3029	1.1539	-0.2407	-0.8593	0.0393	0.138	-0.2139	-0.8481
		DDF**	2	-0.1333	-0.3925	0.3798	1.092	-0.3449	-0.8864	0.1806	0.4526	-0.1975	-0.5917
			3	0.3938	0.8252	0.1141	0.2091	-0.5559	-0.8685	-0.2004	-0.3381	0.0274	0.0569
			4	0.7535**	2.3488	0.3138	0.9582	0.0893	0.2518	0.1211	0.3306	-0.3706	-1.1509
14	m032579 (Geometry)	DIF	2	0.0011	0.0041	0.5712	1.9187	0.2428	0.7681	-0.093	-0.3044	0.5152	1.8858
		DDF**	1	0.022	0.0738	0.6858**	2.0564	0.2193	0.6155	0.1076	0.3139	0.3269	1.1051
			3	0.1404	0.3595	0.5917	1.4172	0.5777	1.3064	-0.3306	-0.7712	0.7977	1.9657
			4	-1.0475	-1.5807	0.1303	0.2115	1.2095	1.6384	-0.7376	-1.071	1.3098**	2.0285
15	m032691 (Geometry)	DIF		-.3982	-1.367	-.4947	-1.7225	.0484	.1694	-.3225	-1.0921	.0492	.1723
		DDF											
16	m042001 (Number)	DIF	2	-0.2143	-0.6221	0.1908	0.5366	0.1907	0.4886	0.6464	1.4403	-0.4739	-1.41
		DDF	1	0.261	0.2342	1.6003	1.3522			-0.3474	-0.2493		
			3	0.1191	0.2806	0.1732	0.4008	0.1934	0.4302	0.6534	1.1912	-0.535	-1.2929
			4	-0.8096	-1.5638	0.34	0.6793	0.3728	0.6271	0.6072	0.8981	-0.2909	-0.6148

17	m042022 (Number)	DIF	3	0.1251	0.498	0.0252	0.099	-0.0998	-0.3545	-0.2783	-1.0029	0.3604	1.4211
		DDF**	1	0.1778	0.4921	-0.0521	-0.1392	-0.8195	-1.7311	0.2209	0.497	0.4478	1.2116
			2	-0.0905	-0.2531	0.221	0.6146	0.4414	1.1043	-0.4409	-1.1464	0.83**	2.0881
			4	0.1474	0.4092	-0.3003	-0.8118	-0.2785	-0.6677	-0.4415	-1.135	-0.0589	-0.1741
18	m042082 (Algebra)	DIF	1	0.1187	0.4072	-0.2309	-0.7908	-0.2665	-0.8848	0.0387	0.1206	0.3987	1.3286
		DDF	2	0.404	0.7976	-0.1338	-0.2624	0.7857	1.3181	0.4364	0.6847	0.4229	0.8002
			3	-0.2145	-0.5579	-0.4809	-1.1945	-0.1132	-0.2709	-0.389	-0.8965	0.8423	1.974
			4	0.1712	0.5524	-0.1396	-0.4454	-0.3574	-1.1204	0.1032	0.3053	0.2341	0.7408
19	m042088 (Algebra)	DIF	1	0.26	1.0023	-0.0929	-0.3475	0.0257	0.0876	-0.2318	-0.8128	0.1913	0.7194
		DDF	2	0.1375	0.441	-0.2439	-0.738	-0.289	-0.8329	-0.0845	-0.2498	0.0104	0.0332
			3	0.4297	1.3283	0.4447	1.3007	.1948	0.4967	0.0609	0.1607	0.4851	1.4197
			4	-0.0519	-0.1194	-0.2922	-0.6421	0.3381	0.6952	-0.264	-0.6177	0.2091	0.4739
20	m042304a (Number)	DIF		-.1753	-.5908	.067	.2209	-.0606	-.1934	-.3474	-1.0302	-.5133	-1.7749
		DDF											
21	m042304b (Number) (Polytomous)	DIF**		-.205	-.876	-.533**	-2.184	-.285	-1.122	-.05	-.196	.364	1.411
		DDF											

22	m042304c (Number)	DIF		.0309	.0962	.1031	.3209	-.1952	-.6162	.0022	.0067	-.1091	-.326
		DDF											
23	m042304d (Number) (Polytomous)	DIF		-.093	-.351	.076	.287	.626	2.189	-.027	-.093	-.05	-.187
		DDF											
24	m042267 (Algebra)	DIF**	4	0.7643**	2.5468	-0.3466	-1.2627	-0.0041	-0.0136	0.4173	1.4311	-0.3885	-1.295
		DDF**	1	1.098**	2.6985	-0.4415	-1.0789	0.127	0.2984	-0.0004	-0.001	-0.4486	-0.9754
			2	0.9918**	2.5301	-0.3402	-0.9692	-0.1168	-0.2986	0.3898	1.0395	-0.4855	-1.3607
			3	0.7671**	2.0129	-0.1655	-0.4789	-0.0916	-0.2387	0.7607**	2.0493	-0.5445	-1.4788
25	m042239 (Algebra)	DIF	1	0.3234	1.067	0.2858	0.921	-0.3584	-1.1472	-0.2628	-0.7635	0.1919	0.6325
		DDF	2	0.5379	1.5071	0.2734	0.7427	-0.1961	-0.556	-0.3898	-1.0277	-0.0751	-0.2203
			3	0.6577	1.6085	0.3497	0.7771	0.2185	0.4596	-0.669	-1.2293	0.7687	1.7623
			4	0.1772	0.495	0.3925	1.0973	-0.4665	-1.2202	-0.3931	-0.8983	0.5256	1.4027
26	m042238 (Algebra)	DIF	3	0.0073	0.029	0.1338	0.5335	-0.015	-0.0564	-0.2238	-0.7942	0.4819	1.9424
		DDF**	1	-0.2654	-0.8779	0.2716	0.9346	-0.1186	-0.376	-0.3083	-0.9382	0.4787	1.6377
			2	-0.0471	-0.144	0.3172	0.9218	0.1529	0.3967	-0.1606	-0.4203	0.8171**	2.3279
			4	0.7949**	2.0129	0.0579	0.1513	0.5545	1.3459	-0.1418	-0.3279	0.4933	1.2961

27	m042279 (Geometry)	DIF**	2	-0.5509**	-2.1896	-0.4349	-1.6824	-0.127	-0.4555	0.3591	1.2315	-0.3329	-1.321
		DDF**	1	-0.1083	-0.2764	-0.6199	-1.5179	-0.1339	-0.3054	-0.1991	-0.4672	-0.3912	-0.98
			3	-0.8798**	-2.9104	-0.3511	-1.1649	-0.2543	-0.7727	0.7149	1.9683	-0.4555	-1.5739
			4	0.4395	0.7953	-0.5091	-0.8717	-0.1805	-0.2614	0.8289	1.0086	0.347	0.5619
28	m042036 (Geometry)	DIF	4	-0.1556	-0.5987	-0.0001	-0.0004	0.1714	0.6052	-0.016	-0.0541	-0.2973	-1.095
		DDF	1	-0.4388	-1.1882	-0.0266	-0.0754	0.2194	0.5674	-0.1294	-0.3124	-0.6062	-1.5672
			2	-0.201	-0.5841	0.0042	0.0121	0.4682	1.2328	-0.6143	-1.5431	-0.1167	-0.3273
			3	0.0773	0.2326	-0.1315	-0.3963	-0.2109	-0.5621	0.2962	0.7599	-0.4086	-1.2281
29	m042130 (Geometry)	DIF**		-.4584	-1.7934	.0108	.0417	.4177	1.5539	-.3815	-1.3903	.6976**	2.5856
		DDF											
30	m042303a (Data and Chance)	DIF		.6377	1.5599	-.2123	-.5549	-.0954	-.2479	.2528	.6626	-.5518	-1.3268
		DDF											
31	m032691 (Data and Chance) (Polytomous)	DIF		.283	1.189	.421	1.711	-.276	-1.091	.068	.264	-.271	-1.067
		DDF											
32	m042222 (Data and Chance)	DIF	1	0.3985	1.5141	-0.215	-0.8052	-0.05	-0.1832	0.4008	1.4198	-0.2036	-0.73
		DDF**	2	0.6999	1.7393	-0.318	-0.7831	-0.0338	-0.0752	0.1249	0.2924	-0.2194	-0.5502
			3	0.2321	0.6781	0.4792	1.3378	0.2847	0.7766	0.0146	0.0387	0.044	0.1162
			4	0.4197	1.3788	-0.6705**	-2.043	-0.244	-0.7644	0.7078	2.0947	-0.3312	-1.0402

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 4

ข้อ	ข้อสอบ	Option	Response	GIRL		HOMEWK		SES		EQUIP		RICH	
				LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)
1	m042001 (Number)	DIF	2	-0.0271	-0.0854	-0.0038	-0.0115	-0.1362	-0.3665	0.1508	0.3648	-0.2044	-0.6452
		DDF	1	0.6315	0.5409	1.0152	0.9677						
			3	-0.1514	-0.4123	-0.2336	-0.5905	-0.0017	-0.0041	-0.2791	-0.6366	0.2002	0.5431
			4	0.0966	0.1788	0.1781	0.3318	-0.0225	-0.033	1.2856	1.1792	-0.7668	-1.3448
2	m042022 (Number)	DIF	3	0.319	1.2699	0.0551	0.2082	0.2117	0.782	-0.2544	-0.868	-0.0589	-0.2306
		DDF**	1	0.441	1.3223	-0.0478	-0.1291	-0.2701	-0.6917	-0.0031	-0.0073	-0.2153	-0.6063
			2	0.3005	0.7418	0.0267	0.0618	0.0029	0.0065	0.6473	1.0841	-0.763	-1.8193
			4	0.3065	0.8639	0.2777	0.7742	0.5546	1.4881	-0.7755**	-2.0429	0.4719	1.3353
3	m042082 (Algebra)	DIF	1	-0.2373	-0.8898	0.2181	0.7934	-0.1432	-0.5087	0.2743	0.9317	-0.1857	-0.6599
		DDF	2	0.239	0.5566	-0.177	-0.413	0.1207	0.2775	0.1823	0.3742	0.1278	0.2985
			3	-0.5545	-1.2033	0.3985	0.8624	-0.6412	-1.2873	0.4857	0.9306	-0.1847	-0.4165
			4	-0.33	-1.1735	0.2573	0.8818	-0.096	-0.3193	0.1713	0.5504	-0.2855	-0.9593
4	m042088 (Algebra)	DIF	1	0.192	0.7	0.0399	0.1413	0.0661	0.2322	-0.213	-0.7138	0.3891	1.3696
		DDF	2	0.2707	0.8412	0.123	0.3651	-0.0857	-0.253	0.0271	0.0753	0.3623	1.0681
			3	0.1635	0.4425	-0.2232	-0.5968	0.4088	1.0197	-0.5391	-1.3334	0.4749	1.2319
			4	-0.0093	-0.0221	0.1869	0.4034	-0.1678	-0.3737	-0.4656	-0.9348	0.342	0.7844

5	m042304a (Number)	DIF		.1376	.4876	-.2986	-.9953	-.3136	-1.0309	.3548	1.085	.0036	.0126
		DDF											
6	m042304b (Number) (Polytomous)	DIF		-.079	-.325	-.089	-.352	-.049	-.199	.289	1.138	.1	.372
		DDF											
7	m042304c (Number)	DIF		.1713	.5748	-.002	-.0067	-.0589	-.192	-.3047	-.9636	-.046	-.1487
		DDF											
8	m042304d (Number) (Polytomous)	DIF		-.122	-.434	-.042	-.143	.003	.01	-.145	-.487	.319	1.022
		DDF											
9	m042267 (Algebra)	DIF	4	0.152	0.5283	-0.146	-0.5089	-0.5155	-1.7831	-0.3205	-1.0292	-0.2642	-0.8902
		DDF**	1	0.681	1.4468	0.2792	0.5738	-0.7681	-1.5663	-0.4653	-0.9156	-0.6589	-1.4757
			2	0.8484**	2.1743	-0.766	-1.9392	-0.0729	-0.185	-0.7884	-1.7984	0.1806	0.4522
			3	-0.2697	-0.7319	0.1078	0.3062	-0.6879	-1.9237	-0.4356	-1.1076	-0.5525	-1.4997
10	m042239 (Algebra)	DIF	1	-0.4759	-1.4706	-0.2066	-0.6234	-0.059	-0.1755	0.0581	0.1674	-0.5637	-1.6046
		DDF	2	-0.6915	-1.724	-0.1822	-0.4459	-0.33	-0.7837	0.2431	0.5983	-0.5187	-1.2861
			3	-0.2863	-0.7064	0.0138	0.0312	0.2563	0.5653	-0.58	-1.1503	-0.5087	-1.1429
			4	-0.4518	-1.2175	-0.1767	-0.4794	0.1013	0.2609	-0.0462	-0.1099	-0.5046	-1.2733

11	m042238 (Algebra)	DIF**	3	-0.4613	-1.8012	0.9484**	3.3871	-0.3122	-1.2138	0.1787	0.6423	0.231	0.8678
		DDF**	1	-0.5499	-1.8113	0.9725**	3.0543	-0.272	-0.914	-0.0658	-0.2033	0.3073	0.9951
			2	-0.7709**	-2.3311	0.4739	1.2899	-0.658	-1.9017	0.5723	1.5459	0.0812	0.2425
			4	-0.0011	-0.003	1.3914**	3.5154	-0.3023	-0.8085	0.6049	1.3782	-0.2238	-0.5889
12	m042279 (Geometry)	DIF	2	0.0533	0.222	0.0804	0.3192	-0.0967	-0.3757	0.414	1.4754	-0.4289	-1.7863
		DDF**	1	-0.2439	-0.6407	0.0287	0.0717	0.3506	0.8796	-0.1652	-0.4113	-0.1458	-0.3726
			3	0.0931	0.3114	-0.0502	-0.1586	-0.3771	-1.1287	0.6993	1.7885	-0.6992**	-2.3815
			4	0.1403	0.2959	0.6579	1.3166	0.0907	0.1799	0.4978	0.8413	-0.2789	-0.5949
13	m042036 (Geometry)	DIF	4	0.1013	0.3962	0.0932	0.357	0.3929	1.3967	-0.4511	-1.5401	0.0972	0.3709
		DDF	1	-0.5883	-1.4169	-0.1958	-0.4772	0.0729	0.1675	0.0623	0.1299	0.214	0.5331
			2	0.2259	0.7515	0.0808	0.2567	0.4697	1.4264	-0.6197	-1.8178	0.3783	1.1915
			3	0.157	0.4922	0.1594	0.489	0.3369	0.9432	-0.4259	-1.1458	-0.1634	-0.5204
14	m042130 (Geometry)	DIF		.0922	.3964	.0068	.0278	.4892	1.9662	-.0822	-.3162	.2287	.9323
		DDF											
15	m042303a (Data and Chance)	DIF		-.3022	-.8558	.145	.3917	.0626	.1707	.0249	.0671	-.4304	-1.0384
		DDF											

16	m042303b (Data and Chance) (Polytomous)	DIF		.468	1.983	.315	1.324	.12	.511	-.459	-1.897	.196	.794
		DDF											
17	m042222 (Data and Chance)	DIF	1	0.0131	0.0488	-0.0442	-0.1646	-0.4806	-1.7419	0.1144	0.3968	-0.194	-0.6821
		DDF**	2	0.2033	0.4511	-0.0453	-0.1002	-0.8449	-1.8982	0.4554	0.9063	0.2006	0.4599
			3	0.0838	0.2125	0.1499	0.3813	-1**	-2.3725	0.0354	0.0818	-0.1469	-0.3699
			4	-0.1112	-0.3889	0.0595	0.2045	-0.2805	-0.9441	0.0945	0.3088	-0.0557	-0.1823
18	m032142 (Number)	DIF**	2	-0.1499	-0.6713	0.0121	0.0519	0.379	1.5476	0.0513	0.2048	0.4774**	2.0341
		DDF**	1	-0.1028	-0.2851	-0.5431	-1.3666	0.7204	1.9576	-0.442	-1.1675	0.2749	0.7763
			3	-0.1841	-0.6542	-0.0237	-0.082	0.0386	0.123	0.3311	1.0039	0.6739**	2.3254
			4	-0.5642	-1.5394	0.5848	1.5273	0.7211	1.7502	0.1029	0.2452	0.3125	0.7876
19	m032198 (Algebra)	DIF**	4	0.6926**	2.5122	-0.6202**	-2.4104	-0.076	-0.2815	0.2845	1.0234	0.1002	0.3624
		DDF**	1	1.3314**	3.241	-1.0698**	-2.8332	-0.1003	-0.2747	0.1397	0.3579	-0.2374	-0.6469
			2	0.5552	1.3925	-0.6211	-1.5744	0.2952	0.691	0.5441	1.2038	0.412	0.9701
			3	0.6283	1.9097	-0.4038	-1.3309	-0.3512	-1.063	0.2998	0.8944	0.0539	0.1648
20	m032640 (Algebra) (Polytomous)	DIF		-.496	-1.797	-.279	-1.018	.458	1.63	-.223	-.782	.391	1.348
		DDF											
21	m032344 (Geometry)	DIF		-.4397	-1.3442	-.6576	-1.9677	-.501	-1.5004	.2924	.8618	-.1913	-.5389
		DDF											

22	m032754 (Geometry)	DIF		-.3167	-1.1575	-.4181	-1.4916	.0091	.033	-.1153	-.3942	-.074	-.2692
		DDF											
23	m032755 (Number) (Polytomous)	DIF**		-.424	-.885	.179	.369	-.297	-.629	.592	1.257	-2.408**	-2.991
		DDF											
24	m032753a (Data and Chance) (Polytomous)	DIF		.262	.804	-.02	-.062	-.674	-1.982	.286	.846	-.67	-1.701
		DDF											
25	m032753b (Data and Chance) (Polytomous)	DIF		.023	.06	-.155	-.413	.162	.427	.349	.923	-.416	-.97
		DDF											
26	m032753c (Data and Chance)	DIF**		.5015	1.6063	.7441**	2.3066	-.128	-.4148	-.4718	-1.4211	-.2385	-.7414
		DDF											
27	m032756 (Data and Chance)	DIF		.0124	.0464	.4577	1.6099	.0851	.3067	.3353	1.1852	.3101	1.1236
		DDF											
28	m032205 (Geometry)	DIF**	2	-0.2633	-1.1548	0.048	0.2048	0.5172**	2.1127	-0.1624	-0.6465	-0.1112	-0.4754
		DDF**	1	0.0849	0.2243	0.2953	0.7664	-0.1067	-0.2441	-0.1746	-0.422	-0.0312	-0.082
			3	-0.5237	-1.5084	0.2331	0.6493	-0.0302	-0.0797	0.7669	1.6873	-0.0286	-0.0803
			4	-0.1711	-0.499	-0.444	-1.2013	0.8936**	2.4961	-0.4821	-1.3054	-0.1677	-0.4783

29	m032163 (Algebra)	DIF	3	0.1023	0.4302	-0.4726	-1.9385	-0.0892	-0.3535	-0.0702	-0.2669	0.3345	1.3149
		DDF**	1	-0.125	-0.3731	-0.1691	-0.4809	-0.5245	-1.3494	0.0202	0.0503	0.23	0.6077
			2	0.0078	0.0262	-0.6511	-2.1058	-0.1288	-0.3989	-0.0486	-0.1453	0.4198	1.3468
			4	0.224	0.5632	-0.9983**	-2.2944	0.0042	0.0101	-0.3154	-0.7633	0.671	1.5818

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 5

ข้อ	ข้อสอบ	Option	Response	GIRL		HOMEWK		SES		EQUIP		RICH	
				LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)
1	m032142 (Number)	DIF	2	-0.0058	-0.0256	0.2137	0.9109	0.2133	0.8481	0.3	1.1886	0.3842	1.6252
		DDF**	1	0.2634	0.8656	0.1681	0.531	0.4196	1.2389	-0.2433	-0.7454	0.7137**	2.1324
			3	-0.1019	-0.3271	0.0318	0.0944	0.3702	1.0644	0.98**	2.4186	0.3252	1.0015
			4	0.0326	0.0892	0.6954	1.9042	-0.1333	-0.3319	0.4767	1.163	0.004	0.0109
2	m032198 (Algebra)	DIF**	4	0.5833**	2.283	-0.465	-1.8401	-0.6866**	-2.5619	0.5108	1.8302	-0.8662**	-3.2249
		DDF**	1	0.7358	1.8857	-0.4895	-1.2386	-0.6556	-1.5591	0.5146	1.1574	-0.7149	-1.8641
			2	1.1106	2.507	-0.3093	-0.7307	0.0616	0.1388	0.5275	1.0834	-0.6851	-1.5627
			3	0.4885	1.7478	-0.5506	-1.9532	-0.9117**	-2.9419	0.4945	1.5875	-0.9265**	-3.0925
3	m032640 (Algebra) (Polytomous)	DIF**		-.087	-.341	-.404	-1.591	-.369	-1.357	.684**	2.533	.072	.256
		DDF											
4	m032344 (Geometry)	DIF**		-.5591	-1.5526	.1767	.4989	-.4013	-1.1049	.3592	.9301	.7881**	2.0422
		DDF											
5	m032754 (Geometry)	DIF		-.106	-.4005	-.1748	-.6474	.4504	1.5949	-.0989	-.3378	.3615	1.3652
		DDF											

6	m032755 (Number) (Polytomous)	DIF		-.154	-.373	.063	.15	-.372	-.877	.506	1.177	-.895	-1.593
		DDF											
7	m032753a (Data and Chance) (Polytomous)	DIF		.501	1.653	.326	1.022	-.45	-1.415	-.722	-1.915	-.199	-.573
		DDF											
8	m032753b (Data and Chance) (Polytomous)	DIF		-.012	-.038	-.154	-.494	-.528	-1.567	-.48	-1.287	-.479	-1.261
		DDF											
9	m032753c (Data and Chance)	DIF**		.214	.8243	.3257	1.2503	-.2118	-.785	-.1013	-.3585	-.7146**	-2.7202
		DDF											
10	m032756 (Data and Chance)	DIF**		-.7016**	-2.6053	.3568	1.3289	-.0398	-.1437	-.2884	-.9547	-.2196	-.8094
		DDF											
11	m032205 (Geometry)	DIF**	2	0.0886	0.36	0.1297	0.5039	0.5855**	2.0573	-0.2395	-0.8496	0.0558	0.2136
		DDF**	1	0.236	0.6927	-0.1024	-0.2898	0.2863	0.7206	0.5047	1.1037	-0.3414	-0.9434
			3	0.1127	0.3024	-0.2309	-0.545	0.5984	1.3788	-0.7529	-1.879	-0.0713	-0.179
			4	-0.2418	-0.5311	1.0004**	2.171	0.7859	1.538	-0.7117	-1.4416	0.8353	1.5996
12	m032163 (Algebra)	DIF	3	-0.0961	-0.3971	0.0289	0.1166	-0.0548	-0.2111	0.0113	0.0415	0.4368	1.7382
		DDF**	1	0.3373	0.969	0.1322	0.3641	0.2054	0.5629	-0.6375	-1.5839	0.7364**	2.0354
			2	-0.1934	-0.6836	-0.0996	-0.3469	-0.2038	-0.6711	0.2377	0.7403	0.2676	0.9406
			4	-0.3745	-1.0252	0.1388	0.3613	-0.5455	-1.2835	-0.0369	-0.0887	0.4784	1.202

13	m042032 (Number)	DIF**	2	0.0741	0.2947	-0.0309	-0.1187	-0.6269**	-2.0862	0.0776	0.2625	0.0159	0.061
		DDF**	1	0.0995	0.3839	-0.0688	-0.255	-0.8249**	-2.546	0.1569	0.5061	-0.099	-0.3691
			3	-0.302	-0.4526	0.3775	0.598	0.1794	0.2501	-0.3794	-0.5334	0.5064	0.6729
			4			-0.0364	-0.0319			-1.4973	-1.2084		
14	m042031 (Number)	DIF	1	0.1225	0.5123	0.4103	1.6046	-0.0104	-0.0391	-0.4549	-1.6137	0.0686	0.2792
		DDF**	2	-0.5026	-1.6968	0.3293	1.0761	0.0032	0.0101	-0.5497	-1.6008	0.3158	1.0687
			3	0.6806**	2.1905	0.6169	1.8694	0.0809	0.2386	-0.4435	-1.2682	-0.0068	-0.0214
			4	0.6177	1.1384	0.523	0.993	-0.2771	-0.3948	-0.5862	-1.0109	0.1186	0.2234
15	m042186 (Number)	DIF		-.1748	-.5921	.062	.2105	.0516	.1652	-.2936	-.8362	-.0514	-.1723
		DDF											
16	m042059 (Number) (Polytomous)	DIF**		-.523**	-2.004	.357	1.373	.533	1.96	.082	.296	.157	.618
		DDF											
17	m042236 (Algebra)	DIF**	3	0.2049	0.8329	0.1542	0.6105	0.6142**	2.1634	-0.1067	-0.3741	0.0047	0.019
		DDF	1	-0.0441	-0.0612	0.4806	0.6937	0.9214	1.1606	-0.7313	-1.0456	0.5057	0.7227
			2	0.1143	0.3537	0.3582	1.0878	0.3748	0.9861	-0.3585	-1.0211	-0.1653	-0.5057
			4	0.261	0.9328	-0.0832	-0.2857	0.4025	1.2442	0.1155	0.3306	-0.1384	-0.5

18	m042226 (Algebra)	DIF**		1.1101**	2.7857	-.4165	-1.1993	.064	.1729	-.271	-.7039	.1404	.3864
		DDF											
19	m042103 (Algebra)	DIF		.6026	1.0201	.0491	.0888	-.2282	-.4242	.0556	.1042	.0309	.0517
		DDF											
20	m042086 (Algebra)	DIF		.1154	.3462	-.37	-1.1205	.0895	.264	.2592	.7174	-.0641	-.1792
		DDF											
21	m042228 (Algebra)	DIF		.2834	1.0039	.3536	1.2588	-.1279	-.4478	.2954	.9688	-.1414	-.4941
		DDF											
22	m042245 (Algebra)	DIF	4	-0.377	-1.3186	-0.0142	-0.0475	0.5536	1.673	-0.2317	-0.7008	-0.2313	-0.7485
		DDF	1	-0.1574	-0.4679	-0.1202	-0.3371	0.2449	0.6438	0.1172	0.3075	-0.5293	-1.4549
			2	-0.6146	-1.5881	-0.1851	-0.467	0.6728	1.5734	-0.191	-0.4364	0.0375	0.0933
			3	-0.1116	-0.3322	0.0589	0.1659	0.407	0.9864	-0.3477	-0.853	-0.3801	-1.0477
23	m042270 (Geometry)	DIF		.0161	.0595	-.4484	-1.6632	.0646	.2273	.0264	.0846	.2156	.7789
		DDF											
24	m042201 (Geometry)	DIF		-.014	-.0459	-.1209	-.3948	.049	.1477	.2563	.742	.1582	.5126
		DDF											

25	m042152 (Geometry)	DIF	4	-0.3873	-1.7422	-0.0226	-0.101	0.2649	1.0566	-0.0872	-0.347	0.0032	0.0138
		DDF	1	-0.531	-1.8298	0.0323	0.1145	0.5776	1.796	-0.2788	-0.8905	0.1331	0.4516
			2	-0.5306	-1.6535	-0.4447	-1.3587	-0.0114	-0.0331	0.0792	0.2214	-0.4206	-1.2796
			3	0.18	0.4686	-0.4858	-1.1968	-0.0104	-0.021	0.1711	0.3601	0.2969	0.7161
26	m042269 (Data and Chance)	DIF	3	0.3882	1.6967	-0.0635	-0.2715	-0.0287	-0.117	0.2688	1.0529	-0.0625	-0.2677
		DDF	1	0.4264	1.1432	-0.0963	-0.254	-0.3455	-0.8228	0.4897	1.0868	0.0399	0.1053
			2	0.5738	1.9366	-0.1649	-0.5424	-0.2133	-0.6655	0.2886	0.8516	-0.3638	-1.2378
			4	0.3967	1.2686	0.0388	0.1184	0.2382	0.6963	0.0188	0.0524	0.1889	0.5567
27	m042179 (Data and Chance)	DIF	4	-0.102	-0.4353	-0.1246	-0.5229	-0.1458	-0.5756	0.0676	0.2552	0.2535	1.0545
		DDF	1	-0.0797	-0.2461	-0.4416	-1.3345	-0.1657	-0.4951	0.2343	0.6456	0.1547	0.4763
			2	-0.2665	-0.8907	-0.1041	-0.3324	-0.3223	-0.9264	0.0186	0.0535	0.4777	1.5184
			3	0.012	0.0327	0.0199	0.0538	-0.1578	-0.3754	0.0563	0.1329	-0.1691	-0.4536
28	m042177 (Data and Chance)	DIF**	3	-0.5201**	-2.1108	-0.2251	-0.8979	-0.5906**	-2.2046	-0.2743	-0.9482	-0.2791	-1.1106
		DDF**	1	-0.5815	-1.6988	-0.3472	-0.9966	-0.6233	-1.6152	-0.3942	-1.0077	-0.3937	-1.1441
			2	-0.6959**	-2.0625	-0.2333	-0.69	-0.4631	-1.2943	0.0427	0.1022	-0.1014	-0.3034
			4	-0.1607	-0.3947	-0.6967	-1.5028	-0.8671	-1.7398	-0.8635	-1.9104	-0.3223	-0.7478

29	m042207 (Data and Chance) (Polytomous)	DIF		.313	1.098	-.151	-.532	.561	1.759	-.188	-.588	.272	.858
		DDF											

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 6

ข้อ	ข้อสอบ	Option	Response	GIRL		HOMEWK		SES		EQUIP		RICH	
				LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)
1	m042032 (Number)	DIF**	2	0.5207**	2.072	-0.2793	-1.0701	-0.5892**	-2.0095	0.384	1.2728	0.0199	0.0795
		DDF**	1	0.4761	1.834	-0.3642	-1.3229	-0.6818**	-2.1734	0.4339	1.3576	-0.055	-0.2141
			3	1.1815	1.7548	0.1157	0.1903	0.374	0.5794	0.2277	0.3332	0.5937	0.8295
			4										
2	m042031 (Number)	DIF	1	-0.1885	-0.7471	0.0329	0.1283	0.1531	0.5521	-0.0478	-0.1685	0.1222	0.4809
		DDF	2	-0.2772	-0.9819	0.1514	0.5275	0.1866	0.6002	-0.1753	-0.5409	0.1585	0.554
			3	0.2146	0.583	-0.2185	-0.5909	0.0312	0.078	-0.1107	-0.2782	0.0224	0.063
			4	-0.2812	-0.5869	-0.7715	-1.3254	0.2577	0.4148	0.4416	0.6447	-0.4699	-0.8785
3	m042186 (Number)	DIF		-.1337	-.4718	-.2922	-1.0249	.4265	1.4034	.4922	1.618	-.2302	-.7811
		DDF											
4	m042059 (Number) (Polytomous)	DIF		-.323	-.724	-.305	-.673	.071	.154	-.027	-.055	-.193	-.433
		DDF											
5	m042236 (Algebra)	DIF	3	0.0271	0.1122	-0.1797	-0.7347	-0.083	-0.3196	-0.0939	-0.3426	0.1986	0.8021
		DDF**	1	-0.2902	-0.5453	-1.7924**	-2.3403	0.9274	1.659	-0.6463	-1.1973	1.1472	1.6938
			2	0.3521	1.0332	-0.4222	-1.1816	0.0619	0.1647	-0.083	-0.2107	0.4614	1.3683
			4	-0.217	-0.7622	0.1129	0.4045	-0.1642	-0.5407	0.0561	0.1718	0.0162	0.0564

6	m042226 (Algebra)	DIF		.2884	.9188	-.1952	-.623	-.5248	-1.6858	.4205	1.236	-.5308	-1.552
		DDF											
7	m042103 (Algebra)	DIF**		1.7306**	2.4056	-.3398	-.5745	.3062	.51	-1.209	-1.4955	.3747	.5624
		DDF											
8	m042086 (Algebra)	DIF		-.1649	-.5243	.4249	1.265	-.6354	-1.8945	.1625	.4657	-.3148	-.9114
		DDF											
9	m042228 (Algebra)	DIF		-.1296	-.4689	.0256	.0921	-.1094	-.3702	.0945	.3237	-.2842	-.9889
		DDF											
10	m042245 (Algebra)	DIF	4	0.1956	0.7307	-0.2259	-0.8173	0.3824	1.2571	-0.4726	-1.4672	-0.1305	-0.4684
		DDF	1	-0.3267	-1	-0.2283	-0.6971	0.0808	0.2127	-0.1076	-0.2509	-0.2791	-0.8386
			2	0.4798	1.2492	-0.5764	-1.3763	0.6039	1.4283	-0.7796	-1.7746	-0.0345	-0.0901
			3	0.1878	0.5564	-0.2143	-0.6183	0.4017	1.0943	-0.5032	-1.284	0.0299	0.0889
11	m042270 (Geometry)	DIF		-.1352	-.4835	.2241	.7924	.2021	.6631	.1297	.4043	-.0618	-.217
		DDF											
12	m042201 (Geometry)	DIF		-.1856	-.6633	-.1799	-.6199	.0777	.2569	-.3166	-.9594	.1453	.5206
		DDF											

13	m042152 (Geometry)	DIF	4	-0.16	-0.707	0.1252	0.5378	-0.1113	-0.458	-0.2512	-0.9755	0.2514	1.0827
		DDF**	1	0.2348	0.8383	-0.0004	-0.0014	0.1167	0.4031	-0.4087	-1.3137	0.2437	0.8424
			2	-0.8249**	-2.4155	0.3771	1.1448	-0.3421	-0.9411	-0.289	-0.7701	0.3334	1.0455
			3	-0.6033	-1.3152	-0.022	-0.0487	-0.8122	-1.4611	0.7746	1.2655	0.0964	0.2107
14	m042269 (Data and Chance)	DIF	3	0.1319	0.5949	-0.2375	-1.0331	0.2371	1.0064	-0.1475	-0.5902	0.1705	0.7531
		DDF**	1	0.5254	1.3168	-0.2611	-0.6379	-0.3583	-0.8203	1.3454**	2.0391	0.0059	0.0145
			2	-0.3572	-1.2455	-0.4637	-1.5462	0.4124	1.3728	-0.276	-0.8835	0.3937	1.3699
			4	0.377	1.2357	-0.0399	-0.1271	0.2538	0.7703	-0.4295	-1.2478	-0.1554	-0.4951
15	m042179 (Data and Chance)	DIF	4	-0.1052	-0.4596	0.3305	1.3777	0.2868	1.1363	0.0601	0.2304	0.4196	1.7571
		DDF	1	-0.3321	-0.9994	0.5302	1.524	0.3948	1.0746	0.7528	1.716	0.6186	1.7604
			2	0.2658	0.8655	0.1539	0.4814	-0.0272	-0.079	-0.1844	-0.5409	0.3948	1.2658
			3	0.0242	0.0733	0.1741	0.5063	0.5933	1.686	-0.26	-0.7123	0.3099	0.8977
16	m042177 (Data and Chance)	DIF**	3	-0.0796	-0.3141	-0.2773	-1.0653	-0.5897**	-2.0576	0.2972	0.9861	-0.394	-1.5785
		DDF	1	-0.4255	-1.2037	-0.2774	-0.7617	-0.7864	-1.9185	0.7339	1.5858	-0.4215	-1.226
			2	-0.1322	-0.3785	-0.1808	-0.496	-0.5582	-1.3265	-0.1741	-0.4096	-0.2615	-0.7393
			4	0.3017	0.7168	-0.5157	-1.1407	-0.4937	-1.0409	0.1689	0.348	-0.6355	-1.4893

17	m042207 (Data and Chance) (Polytomous)	DIF		.323	.727	.305	.663	-.071	-.153	.027	.055	.193	.422
		DDF											
18	m032381 (Number)	DIF		-.1296	-.5041	-.3338	-1.2668	.2651	.9325	-.159	-.5487	.1199	.4403
		DDF											
19	m032416 (Number)	DIF	4	0.2025	0.6885	-0.4752	-1.5989	-0.1916	-0.6344	-0.2775	-0.8361	-0.4494	-1.4749
		DDF	1	0.1324	0.359	-0.5126	-1.2889	-0.4528	-1.1275	-0.0825	-0.1672	-0.1572	-0.3866
			2	0.2346	0.6488	-0.0523	-0.1474	-0.1985	-0.5428	-0.3474	-0.8694	-0.6132	-1.8073
			3	0.0705	0.1786	-0.8504	-1.9244	-0.2729	-0.6743	-0.3109	-0.6659	-0.0841	-0.1866
20	m032160 (Number)	DIF	4	-0.4569	-1.7203	0.2968	1.0812	0.2954	1.0186	-0.0202	-0.0719	-0.1489	-0.5486
		DDF**	1	0.0167	0.0356	-0.1719	-0.3411	-0.3908	-0.7442	0.043	0.0828	-0.5561	-1.187
			2	-0.4277	-1.4597	0.7511**	2.4363	0.4632	1.4066	0.1021	0.3225	-0.1459	-0.4803
			3	-0.7598**	-2.1616	-0.4259	-1.1467	0.2908	0.7687	-0.1894	-0.5203	-0.3467	-1.0114
21	m032273 (Algebra)	DIF	2	-0.0541	-0.2236	0.2126	0.8632	-0.2181	-0.8165	0.4399	1.5627	-0.0562	-0.2329
		DDF	1	-0.4051	-1.0616	0.6254	1.7041	-0.5739	-1.3124	0.5017	1.1134	-0.0394	-0.1073
			3	0.1354	0.4495	0.358	1.1738	0.2295	0.7127	0.3476	0.9621	0.0982	0.3238
			4	.3514 .	0.9209	-0.3525	-0.8484	-0.4849	-0.9329	0.3098	0.6648	-0.3715	-0.9494

22	m032540 (Algebra)	DIF	4	0.1544	0.6556	-0.2115	-0.8868	0.2197	0.8595	0.0504	0.1908	-0.0219	-0.093
		DDF	1	0.1014	0.3219	-0.5594	-1.6639	-0.473	-1.2441	-0.0104	-0.0289	-0.1792	-0.582
			2	0.481	1.1339	-0.5357	-1.2084	0.3786	0.8599	0.7275	1.2172	0.9217	1.9134
			3	0.3825	1.2537	0.0653	0.215	0.45	1.4032	0.0457	0.1376	-0.0995	-0.3229
23	m032698 (Algebra)	DIF	1	-0.003	-0.012	0.4167	1.5754	0.2777	1.018	0.1311	0.4634	0.3622	1.3772
		DDF**	2	0.0809	0.2703	0.8311**	2.6251	0.2454	0.7374	-0.3646	-1.0965	1.1735**	3.2588
			3	0.1007	0.268	0.4015	0.986	0.7658	1.9188	0.4356	0.8707	0.056	0.1455
			4	-0.1009	-0.2849	-0.281	-0.7289	-0.2507	-0.6232	0.6903	1.5378	-0.2655	-0.7338
24	m032097 (Geometry)	DIF	3	-0.4336	-1.72	-0.3571	-1.3911	0.0047	0.0175	-0.5195	-1.77	0.1993	0.7725
		DDF**	1	-0.3532	-1.1876	-0.5414	-1.7297	0.0459	0.1428	-0.5995	-1.7922	0.1109	0.3638
			2	-0.2746	-0.9126	0.0046	0.0151	-0.233	-0.7461	-0.323	-0.8992	0.1165	0.3855
			4	-1.0611**	-2.0285	-1.2373**	-2.0028	0.3911	0.7397	-0.8699	-1.5657	0.0754	0.1548
25	m032575 (Geometry)	DIF	2	0.0074	0.0303	0.2217	0.8691	-0.2593	-0.9927	-0.0993	-0.3554	-0.0444	-0.1844
		DDF	1	0.6116	1.7055	0.4405	1.1747	-0.0318	-0.0826	-0.6465	-1.5478	-0.3402	-0.9801
			3	0.1636	0.5082	0.0543	0.1634	-0.3895	-1.1056	0.1365	0.36	0.0545	0.178
			4	-0.3817	-1.1745	0.3236	0.9759	-0.1389	-0.4038	-0.1101	-0.3017	0.1245	0.385

26	m032414 (Geometry)	DIF		.2075	.6594	.42	1.2567	-.2983	-.9221	-.4241	-1.2211	-.4629	-1.407
		DDF											
27	m032294 (Geometry)	DIF	1	0.3749	1.5621	0.0795	0.321	-0.0182	-0.0691	-0.0014	-0.005	-0.3488	-1.4431
		DDF**	2	0.918**	2.6772	-0.2008	-0.5705	0.4877	1.3839	-0.1896	-0.4823	-0.6878**	-2.0194
			3	0.3161	0.9068	-0.0105	-0.0278	-0.1089	-0.2667	0.0946	0.2131	0.2241	0.6086
			4	-0.1386	-0.3833	0.4513	1.3036	-0.4693	-1.1102	-0.0163	-0.0403	-0.4876	-1.3939
28	m032688 (Data and Chance)	DIF		-.5315	-1.97	-.0543	-.1987	.0064	.0228	.2257	.7775	.0778	.2871
		DDF											
29	m032529 (Number)	DIF**	2	-0.1793	-0.6466	0.515	1.6952	0.0633	0.2126	0.6528**	2.2009	-0.5156	-1.7934
		DDF**	1	-0.1943	-0.5399	0.6739	1.7391	0.3337	0.8448	0.2079	0.5528	-0.3889	-1.0842
			3	1.1208	1.9996	0.5348	1.0679	0.2274	0.456	0.5205	0.9734	-0.4946	-1.0194
			4	-0.387	-1.293	0.471	1.4123	0.0159	0.0484	0.9467**	2.7472	-0.4421	-1.4013
30	m032637a (Data and Chance)	DIF		.3521	1.3884	.1057	.4131	-.1284	-.4705	-.0312	-.1107	.3889	1.4776
		DDF											
31	m032637b (Data and Chance)	DIF		.3396	1.0788	-.1004	-.3193	.5209	1.4883	-.2496	-.7226	.268	.8383
		DDF											

32	m032637c (Data and Chance)	DIF		-.0104	-.0397	.1801	.6825	.0381	.1375	-.1388	-.4705	.3368	1.2451
		DDF											

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 7

ข้อ	ข้อสอบ	Option	Response	GIRL		HOMEWK		SES		EQUIP		RICH	
				LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)
1	m032381 (Number)	DIF		-.2963	-1.1633	.1637	.6385	-.1367	-.5188	.1186	.4313	-.3517	-1.3879
		DDF											
2	m032416 (Number)	DIF**	4	0.6173**	2.0645	-0.14	-0.4741	0.1618	0.5349	0.2732	0.8908	-0.1284	-0.4338
		DDF**	1	0.5692	1.5018	-0.0676	-0.1742	-0.4937	-1.1441	0.7044	1.6294	-0.4417	-1.1602
			2	0.783**	2.3023	-0.0899	-0.2665	0.1509	0.4467	0.4345	1.2137	-0.1216	-0.3585
			3	0.0714	0.1647	-0.2909	-0.6734	0.8657	1.8852	-0.4686	-1.0696	0.5708	1.2542
3	m032160 (Number)	DIF**	4	-0.7554**	-2.6302	-0.44	-1.5619	-0.1544	-0.5257	-0.55	-1.7427	0.4541	1.5451
		DDF**	1	-0.1563	-0.3288	-0.296	-0.6059	-0.616	-1.1344	-0.0699	-0.1242	0.1586	0.3169
			2	-0.8932**	-2.8683	-0.5	-1.6469	-0.1594	-0.5159	-0.6098	-1.809	0.5204	1.6354
			3	-0.5221	-1.3408	-0.4001	-1.0752	-0.1851	-0.438	-0.5848	-1.3109	0.1472	0.3829
4	m032273 (Algebra)	DIF	2	0.2445	1.0137	-0.2227	-0.9275	0.0604	0.2292	-0.358	-1.3318	0.0136	0.0542
		DDF	1	-0.0238	-0.0764	-0.5448	-1.6321	0.2314	0.6684	-0.1214	-0.3224	-0.1844	-0.5536
			3	0.3084	0.9097	-0.0489	-0.1523	0.0171	0.0486	-0.2129	-0.5751	0.3552	1.0392
			4	0.8803	1.8818	-0.1648	-0.3717	-0.0401	-0.0731	-0.5216	-1.1193	-0.2713	-0.5951

5	m032540 (Algebra)	DIF	4	0.2003	0.8589	-0.1363	-0.5773	0.3092	1.1734	-0.3158	-1.1881	0.4445	1.7472
		DDF**	1	0.0859	0.2642	-0.2321	-0.6863	0.4248	1.1807	-0.3226	-0.8921	0.7806**	2.2095
			2	0.946**	2.3986	-0.0774	-0.2021	-0.2899	-0.6053	-0.2323	-0.5345	0.0469	0.1183
			3	0.1831	0.6071	-0.1287	-0.4205	0.4314	1.2393	-0.2426	-0.6737	0.4448	1.345
6	m032698 (Algebra)	DIF	1	0.177	0.7108	-0.1409	-0.556	-0.0781	-0.3012	-0.3317	-1.1775	0.1949	0.7676
		DDF	2	0.1407	0.4448	-0.2765	-0.8518	0.0382	0.1204	-0.4347	-1.2711	0.2991	0.9214
			3	0.3781	0.956	-0.0068	-0.0174	-0.8158	-1.765	-0.4292	-0.932	-0.1469	-0.3908
			4	0.0859	0.2807	-0.1124	-0.3546	0.0571	0.1743	-0.242	-0.6643	0.278	0.8535
7	m032097 (Geometry)	DIF	3	-0.0194	-0.0817	0.2018	0.827	0.2093	0.8141	0.0518	0.2014	-0.026	-0.1064
		DDF	1	0.1069	0.3681	0.0845	0.2796	0.2082	0.629	0.0565	0.1725	0.0612	0.2049
			2	0.2796	0.9456	0.1492	0.4868	0.2606	0.8305	0.0059	0.0184	-0.0035	-0.0114
			4	-0.7915	-1.5544	0.29	0.5814	0.471	0.9156	-0.5333	-0.9732	-0.479	-0.9679
8	m032575 (Geometry)	DIF	2	-0.3915	-1.4966	0.2401	0.9245	-0.0567	-0.2143	0.2093	0.7605	-0.1457	-0.5652
		DDF**	1	0.8726**	2.0269	0.6332	1.4417	-0.0356	-0.0852	0.1673	0.3763	0.0131	0.0316
			3	-0.4609	-1.3313	0.2561	0.74	-0.4669	-1.2261	-0.1379	-0.3688	-0.1694	-0.4795
			4	-0.666	-2.0826	0.0116	0.0376	0.0834	0.264	0.3649	1.0723	-0.4555	-1.5173

9	m032414 (Geometry)	DIF		-.206	-.6598	-.0797	-.2534	.5302	1.5494	-.2303	-.6828	-.1814	-.5724
		DDF											
10	m032294 (Geometry)	DIF**	1	0.3343	1.4117	0.5154**	2.0799	-0.2915	-1.1329	0.1844	0.6875	-0.304	-1.2784
		DDF**	2	1.3617**	3.7093	0.6006	1.7886	-0.085	-0.2332	0.1599	0.4247	0.0288	0.0866
			3	-0.0416	-0.1112	0.2604	0.6564	-0.1117	-0.2711	0.7405	1.4957	-0.6072	-1.6063
			4	-0.0844	-0.2344	0.4983	1.3486	-0.7719	-1.7778	-0.1777	-0.4636	-0.5541	-1.5264
11	m032688 (Data and Chance)	DIF**		.7075**	2.3814	-.6566**	-2.3375	-.359	-1.2071	-.4838	-1.4964	-.06	-.2058
		DDF											
12	m032529 (Number)	DIF**	2	-0.7297**	-2.2508	-0.2528	-0.8224	0.126	0.3685	0.2078	0.6435	-0.059	-0.1809
		DDF**	1	-0.426	-0.9265	-0.2132	-0.4786	0.3171	0.629	1.011	1.8989	0.2102	0.4304
			3	-0.629	-1.2338	-0.3673	-0.7705	0.2447	0.4749	0.0164	-0.0346	0.3357	0.6197
			4	-0.8025**	-2.2902	-0.2789	-0.838	0.2122	0.565	0.0259	0.0733	-0.2856	-0.8257
13	m032637a (Data and Chance)	DIF		-.2271	-.821	-.1004	-.3678	-.1546	-.519	-.1714	-.5661	-.0507	-.1837
		DDF											
14	m032637b (Data and Chance)	DIF**		.7808**	2.489	.1377	.4412	-.1735	-.5094	-.0408	-.113	.4813	1.5093
		DDF											

15	m032637c (Data and Chance)	DIF		-.3308	-1.1558	-.1515	-.5448	.2082	.6887	-.1246	-.414	.2011	.7221
		DDF											
16	m042183 (Number)	DIF	4	0.1841	0.7382	-0.4381	-1.716	0.0435	0.1629	-0.2109	-0.7706	0.0166	0.0647
		DDF	1	0.5314	1.7666	-0.4581	-1.5179	0.0536	0.1684	-0.0004	-0.0012	-0.1978	-0.6651
			2	-0.2946	-0.7443	-0.4131	-1.024	0.385	0.9108	-0.0366	-0.08	0.1201	0.2997
			3	-0.458	-1.1912	-0.6168	-1.5185	-0.0984	-0.2262	-0.7106	-1.5802	0.2571	0.6323
17	m042060 (Number)	DIF	2	-0.324	-1.2222	0.1379	0.518	-0.1805	-0.6673	0.0266	0.0969	0.005	0.0187
		DDF	1	-0.4188	-1.4058	0.2073	0.708	-0.312	-1.0355	-0.0991	-0.3234	0.1201	0.3982
			3	0.3052	0.749	-0.0419	-0.0918	-0.2512	-0.5531	0.5727	1.2311	-0.4719	-1.0796
			4	-0.2352	-0.5953	0.2634	0.6336	0.0026	0.0066	0.0119	0.0281	-0.1418	-0.3771
18	m042019 (Number)	DIF		-.1679	-.531	-.2439	-.7825	.1426	.447	.2846	.8789	.395	1.254
		DDF											
19	m042023 (Number)	DIF		.4697	1.3168	-.0356	-.1082	.1385	.402	-.0669	-.1957	-.1306	-.3856
		DDF											
20	m042197 (Number)	DIF		.2946	.7219	-.1848	-.4713	-.1119	-.2704	-.0074	-.0184	-.3579	-.8059
		DDF											

21	m042234 (Algebra)	DIF	3	-0.0987	-0.4249	0.2087	0.8624	-0.2337	-0.9307	0.0039	0.0147	-0.0894	-0.369
		DDF	1	-0.5344	-1.6773	-0.2061	-0.6549	0.0971	0.2923	-0.1079	-0.3116	-0.1043	-0.3403
			2	-0.0757	-0.2087	0.1146	0.3014	-0.7591	-1.6948	-0.1226	-0.2816	-0.0817	-0.2192
			4	-0.0054	-0.0191	0.4437	1.4639	-0.2288	-0.7388	-0.0563	-0.1698	0.0641	0.2087
22	m042066 (Algebra)	DIF		.178	.7319	.1342	.5387	-.1232	-.4708	-.2724	-1.0059	.0105	.0424
		DDF											
23	m042243 (Algebra)	DIF**	1	0.0545	0.1616	-0.0085	-0.0254	-0.4223	-1.2587	0.7068**	2.0764	-0.4078	-1.1769
		DDF**	2	0.3468	0.7318	-0.5214	-1.1129	-0.6483	-1.3384	1.0822**	2.0232	-0.0604	-0.1258
			3	0.3831	0.9778	0.1018	0.2536	-0.0132	-0.0321	0.0375	0.0972	-0.2575	-0.6033
			4	-0.5552	-1.2615	-0.0014	-0.0034	-0.8509**	-2.0415	1.3624**	3.0316	-0.6268	-1.5816
24	m042248 (Algebra)	DIF		.0101	.0229	-.1446	-.3403	-.2946	-.6999	.2536	.6021	-.1245	-.2841
		DDF											
25	m042229a (Algebra)	DIF		.5069	1.0705	.509	1.0916	-.2774	-.5783	-.0907	-.1946	.0212	.0426
		DDF											
26	m042229b (Algebra)	DIF		.0353	.0621	-.8635	-1.3388	.0659	.113	.117	.2128	.0084	.0147
		DDF											

27	m042080a (Algebra)	DIF**		-.2309	-.9247	.1865	.7484	.7054**	2.4324	-.2033	-.7361	.5325**	2.0584
		DDF											
28	m042080b (Algebra)	DIF		.4958	.6897	.199	.2964	.0975	.1268	.794	1.0901	-.0517	-.059
		DDF											
29	m042120 (Geometry)	DIF	1	0.3744	1.5912	0.0722	0.3002	0.1384	0.5348	0.3568	1.3676	-0.3353	-1.4059
		DDF	2	0.2435	0.9365	0.1325	0.5008	0.1037	0.3625	0.2028	0.7212	-0.1082	-0.413
			3	0.7183	1.7793	-0.4116	-0.9561	0.5599	1.3115	0.5537	1.1309	-0.6741	-1.6723
			4	1.3653	1.7446	0.4058	0.6125			1.5819	1.3631	-0.5166	-0.6665
30	m042203 (Geometry)	DIF	3	0.09	0.3584	0.1139	0.4374	0.1158	0.4205	0.1045	0.3741	0.3115	1.1675
		DDF	1	0.5326	1.1528	-0.1905	-0.4036	-0.2667	-0.5179	-0.6523	-1.2745	0.0121	0.0256
			2	-0.0682	-0.2314	-0.0016	-0.0053	0.3318	0.9922	0.3037	0.8924	0.4881	1.604
			4	0.0372	0.094	0.3745	0.9035	-0.0322	-0.076	-0.25	-0.5927	-0.0298	-0.0673
31	m042264 (Geometry)	DIF**		-.2261	-.6552	.6292	1.7386	-.2609	-.7602	.7317**	2.1345	.0727	.2082
		DDF											

32	m042255 (Data and Chance)	DIF	2	-0.314	-1.3505	0.0537	0.2281	0.1443	0.5632	-0.0752	-0.2989	0.0221	0.0936
		DDF	1	-0.2961	-0.8583	0.1469	0.4219	0.1225	0.3117	0.058	0.1531	-0.045	-0.129
			3	-0.5256	-1.6696	-0.0621	-0.1952	0.1991	0.5875	0.3292	0.9211	0.2428	0.7592
			4	-0.3655	-0.7476	0.2415	0.5022	-0.2505	-0.4477	-0.8324	-1.6995	-0.8104	-1.5943
33	m042224 (Data and Chance)	DIF**		-.5491**	-2.0099	.2179	.7847	-.2127	-.7676	.3652	1.2659	-.6235**	-2.2086
		DDF											

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 8

ข้อ	ข้อสอบ	Option	Response	GIRL		HOMEWK		SES		EQUIP		RICH	
				LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)
1	m042183 (Number)	DIF	4	-0.0131	-0.0551	-0.1936	-0.8104	0.0631	0.2519	-0.217	-0.807	-0.1798	-0.7439
		DDF	1	0.1082	0.3603	-0.2652	-0.8861	0.0972	0.3161	-0.1054	-0.3131	-0.2412	-0.8019
			2	-0.4279	-1.201	-0.1809	-0.5222	-0.4781	-1.215	-0.5155	-1.2836	-0.4955	-1.4113
			3	0.1505	0.4361	-0.3547	-0.9869	0.2586	0.698	-0.1873	-0.4817	-0.0441	-0.127
2	m042060 (Number)	DIF**	2	-0.65**	-2.3611	0.03	0.1094	0.0219	0.0767	0.1938	0.6492	0.2513	0.8994
		DDF**	1	-0.5145	-1.6991	-0.0218	-0.0698	-0.0949	-0.2912	0.5387	1.5343	0.3896	1.2114
			3	-0.5331	-1.2927	0.5304	1.3263	0.9102**	2.0977	-0.7449	-1.6528	0.4284	1.0118
			4	-0.8228	-1.8256	0.2613	0.5742	-0.4109	-0.8585	0.4463	0.8921	-0.1853	-0.4356
3	m042019 (Number)	DIF		-.0019	-.006	.3845	1.2423	-.441	-1.4226	.5308	1.65	.1158	.3559
		DDF											
4	m042023 (Number)	DIF		.1175	.3296	.1894	.5416	-.2273	-.6715	.2662	.7496	-.0628	-.1731
		DDF											
5	m042197 (Number)	DIF		.1894	.4696	.17	.451	.2005	.4923	-.3829	-.9417	.0426	.0906
		DDF											

6	m042234 (Algebra)	DIF	3	0.3931	1.596	0.1133	0.4457	-0.1538	-0.5839	-0.4896	-1.6435	-0.1279	-0.5053
		DDF**	1	0.2653	0.8704	-0.2584	-0.7956	-0.0181	-0.0542	-1.1382**	-2.8412	0.0824	0.2502
			2	-0.4115	-0.999	0.2102	0.5131	0.6376	1.4718	-0.5895	-1.1669	0.457	1.0786
			4	0.7553**	2.1899	0.3617	1.0743	-0.9252**	-2.3482	-0.1714	-0.4432	-0.4316	-1.3333
7	m042066 (Algebra)	DIF		-.522	-1.9773	.0828	.3257	.1266	.4579	-.2511	-.9118	-.0801	-.3069
		DDF											
8	m042243 (Algebra)	DIF**	1	-0.3414	-1.1858	-0.3752	-1.2987	-0.6263**	-2.1332	0.4755	1.5458	-0.0272	-0.0905
		DDF**	2	-0.7577	-1.7382	-0.5581	-1.3282	-0.6621	-1.5452	-0.0152	-0.0358	-0.1223	-0.2979
			3	-0.3835	-1.0733	-0.0916	-0.2405	-0.2839	-0.7702	0.5852	1.397	0.3856	0.9389
			4	-0.1967	-0.5716	-0.3832	-1.1508	-0.9404**	-2.5293	0.3889	1.0327	-0.3277	-0.9471
9	m042248 (Algebra)	DIF		-.1641	-.4263	.0285	.0777	.1378	.3566	-.7168	-1.6918	.6719	1.6793
		DDF											
10	m042229a (Algebra)	DIF		.5442	1.1106	.1652	.3256	-.1386	-.2883	-.1598	-.3085	-.2748	-.4594
		DDF											
11	m042229b (Algebra)	DIF		.1095	.2473	-.3007	-.6462	.375	.7991	.0469	0981	-.4189	-.805
		DDF											

12	m042080a (Algebra)	DIF**		.0662	.2689	.5026	1.9383	1.0474**	3.4296	-.7241**	-2.4638	.5962**	2.2414
		DDF											
13	m042080b (Algebra)	DIF		.3503	.5664	-.9165	-1.5497	.7149	1.1977	-.1843	-.3103	1.3129	1.4926
		DDF											
14	m042120 (Geometry)	DIF	1	0.2137	0.9051	-0.1111	-0.4649	-0.1006	-0.3953	0.0833	0.3153	-0.0941	-0.3892
		DDF	2	0.0809	0.3152	-0.0582	-0.2245	-0.0411	-0.1475	0.1544	0.5207	0.0159	0.0602
			3	0.7615	1.6398	-0.0607	-0.1337	0.1524	0.3171	-0.4412	-0.9395	0.2563	0.5512
			4	0.8674	1.1652	-0.3358	-0.4393	0.2737	0.3541	0.118	0.1366	-0.4408	-0.5955
15	m042203 (Geometry)	DIF	3	-0.0123	-0.0462	-0.0746	-0.2743	0.1647	0.5681	0.2487	0.8326	0.3404	1.2092
		DDF	1	0.2981	0.7056	-0.4825	-1.1092	0.0052	0.0112	0.7157	1.2137	0.1508	0.3485
			2	-0.2559	-0.7814	-0.231	-0.6981	0.2321	0.6624	-0.1535	-0.4332	0.5123	1.5256
			4	-0.1056	-0.2829	0.3159	0.7911	0.0534	0.126	0.2287	0.5172	0.3243	0.786
16	m042264 (Geometry)	DIF**		-.0053	-.0164	-.6593**	-2.0871	-.1981	-.5949	.2887	.8983	-.4385	-1.2017
		DDF											

17	m042255 (Data and Chance)	DIF	2	0.0039	0.0164	-0.0306	-0.1258	0.0365	0.1392	-0.008	-0.029	0.0807	0.3328
		DDF	1	0.0561	0.165	-0.2636	-0.7361	0.3122	0.8381	0.1104	0.2702	0.3948	1.0593
			3	-0.1935	-0.627	0.0271	0.0854	-0.2473	-0.7094	0.1539	0.418	0.1098	0.3648
			4	0.204	0.4293	0.41	0.8558	0.3273	0.6245	-0.2995	-0.5807	-0.4286	-0.8857
18	m042224 (Data and Chance)	DIF**		-.0117	-.0433	-.2269	-.87	-.5387	-1.9546	.7409**	2.648	-.2963	-1.0485
		DDF											
19	m032094 (Number)	DIF**	1	0.2031	0.8537	0.3359	1.3649	0.6312**	2.3587	-0.3302	-1.1964	0.3195	1.2836
		DDF**	2	0.0498	0.159	0.4737	1.458	1.0292**	2.7884	-0.3403	-0.9659	0.6871**	2.0365
			3	-0.0931	-0.2307	0.286	0.6642	0.1167	0.2375	0.1639	0.3165	-0.1672	-0.4075
			4	0.5243	1.6004	0.0633	0.1908	0.8388**	2.4483	-0.3035	-0.7764	0.2568	0.7791
20	m032662 (Number)	DIF	4	-0.2994	-0.9284	0.1369	0.4179	0.1971	0.5685	0.4092	1.1816	-0.011	-0.0305
		DDF	1	-0.2273	-0.6419	0.1347	0.3712	0.3753	0.9802	0.4059	1.0416	0.1422	0.368
			2	-0.2272	-0.5753	0.0853	0.221	0.2562	0.6071	0.4377	1.0673	-0.1452	-0.3475
			3	-0.5743	-1.1167	0.7947	1.4216	0.4855	0.8602	-0.0951	-0.1658	-0.2563	-0.4293
21	m032064 (Number)	DIF		-.5265	-1.3797	-.3048	-.9228	-.0209	-.0572	.0175	.0494	.5636	1.4247
		DDF											

22	m032419 (Algebra)	DIF	3	0.2366	0.9773	-0.1793	-0.7502	0.0102	0.0401	-0.1432	-0.5272	0.2441	0.9641
		DDF	1	0.1697	0.5	-0.0899	-0.2681	0.2363	0.6476	-0.4515	-1.1718	0.5426	1.4761
			2	0.3278	1.1311	-0.3433	-1.1858	0.0121	0.0396	-0.0394	-0.1193	0.2925	0.9845
			4	-0.2196	-0.605	0.0184	0.0519	-0.2809	-0.7449	-0.0009	-0.0022	-0.0395	-0.1068
23	m032477 (Algebra)	DIF	1	-0.1006	-0.3893	0.2743	1.0615	0.1863	0.6765	-0.2455	-0.8976	0.0829	0.311
		DDF	2	0.0521	0.1762	0.2168	0.7455	0.3816	1.2196	-0.02	-0.0632	-0.0566	-0.1871
			3	-0.1251	-0.3704	0.6869	1.8923	0.1269	0.3267	-0.2211	-0.5774	0.1804	0.497
			4	0.1174	0.2552	-0.9064	-1.7167	0.4367	0.9015	-0.4819	-1.0636	0.598	1.298
24	m032538 (Algebra)	DIF**		.7411**	2.0523	-.1113	-.3448	-.1389	-.4203	.1061	.3042	-.4189	-1.2675
		DDF											
25	m032324 (Geometry)	DIF	4	-0.0965	-0.3602	-0.3388	-1.2623	0.1436	0.5019	0.3203	1.1246	0.0225	0.0784
		DDF	1	0.0974	0.2978	-0.5584	-1.6356	0.2221	0.627	0.3706	1.0043	0.0429	0.1202
			2	-0.4296	-1.1347	-0.0777	-0.206	0.3144	0.8045	0.3306	0.8558	0.115	0.2943
			3	-0.1057	-0.314	-0.0042	-0.0122	-0.1058	-0.2799	0.1935	0.5204	-0.2225	-0.6355

26	m032116 (Geometry)	DIF	2	0.0476	0.2089	0.3413	1.4617	0.15	0.6112	-0.0582	-0.2273	0.2703	1.1248
		DDF**	1	-0.1911	-0.7436	0.3118	1.1833	0.2795	1.0257	0.0007	0.0024	0.6557**	2.2871
			3	0.1097	0.3161	0.158	0.4481	-0.3378	-0.8458	-0.2277	-0.5582	0.2395	0.6829
			4	0.4312	1.1474	0.6626	1.804	0.4015	1.0305	0.1006	0.2454	-0.2883	-0.7982
27	m032100 (Geometry)	DIF**	4	-0.6854**	-2.735	-0.1132	-0.4666	-0.1513	-0.598	-0.1671	-0.6221	0.0999	0.395
		DDF**	1	-0.7548**	-2.4223	-0.2757	-0.9019	-0.181	-0.5764	-0.0141	-0.042	-0.1173	-0.3882
			2	-0.6841**	-2.1075	-0.099	-0.3182	0.0559	0.1714	-0.4301	-1.2085	0.1988	0.5915
			3	-0.2132	-0.5321	-0.2245	-0.5608	-0.4811	-1.0768	-0.1421	-0.3179	0.3043	0.7525
28	m032402 (Geometry)	DIF**	1	0.4896**	2.0257	-0.1182	-0.4872	0.2934	1.1298	0.0254	0.0929	-0.0423	-0.1754
		DDF**	2	0.5751	1.5552	-0.379	-0.9963	-0.1249	-0.2975	0.4506	0.9787	-0.5057	-1.4122
			3	0.2526	0.8361	-0.141	-0.4588	0.4684	1.4497	-0.2379	-0.6852	0.245	0.7808
			4	1.3127**	2.3631	0.4203	0.8465	0.9904	1.8912	0.2663	0.4427	0.4238	0.8347
29	m032734 (Geometry)	DIF		-.0148	-.0621	.084	.351	-.2177	-.8601	.0085	.0337	-.291	-1.1834
		DDF											

30	m032397 (Geometry)	DIF**	2	-0.2289	-0.9642	-0.0051	-0.0211	-0.4666	-1.8516	0.2823	1.0713	-0.8198**	-3.2858
		DDF**	1	-0.2643	-0.8956	0.0678	0.2246	-0.4353	-1.3745	-0.2289	-0.6989	-0.5145	-1.677
			3	0.1474	0.4105	-0.3989	-1.0629	-0.418	-1.0484	1.112**	2.177	-1.0244**	-2.6698
			4	-0.6228	-1.7091	-0.0864	-0.2383	-0.1815	-0.4959	0.5792	1.3813	-1.1282**	-3.1183
31	m032695 (Data and Chance) (Polytomous)	DIF											
		DDF											
32	m032132 (Data and Chance)	DIF	4	0.333	1.5488	0.1662	0.7684	-0.2971	-1.2957	-0.0382	-0.1596	-0.2246	-0.9907
		DDF**	1	0.9235	1.9994	0.0616	0.1362	-0.3105	-0.6334	0.3352	0.6097	-1.0453**	-2.2518
			2	1.1072**	2.9369	-0.1103	-0.2984	-1.2007**	-2.5716	-0.1083	-0.2612	0.1686	0.4424
			3	0.0746	0.2944	0.1957	0.7738	0.0187	0.0703	-0.0682	-0.241	-0.1162	-0.436

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 9

ข้อ	ข้อสอบ	Option	Response	GIRL		HOMEWK		SES		EQUIP		RICH	
				LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)
1	m032094 (Number)	DIF	1	0.3034	1.2737	0.2146	0.8709	-0.1147	-0.4094	-0.3985	-1.477	0.1966	0.8121
		DDF	2	0.2482	0.824	0.4129	1.2935	-0.1889	-0.5105	-0.5246	-1.5434	0.2354	0.7589
			3	0.1402	0.342	-0.0308	-0.0764	0.6332	1.4801	-0.352	-0.7924	0.8312	1.8763
			4	0.5368	1.5641	-0.0347	-0.0943	-0.8031	-1.6624	0.0031	0.0078	-0.2006	-0.585
2	m032662 (Number)	DIF**	4	0.0509	0.1521	0.0917	0.2591	1.2052**	2.3706	0.1953	0.544	0.1113	0.3047
		DDF**	1	-0.0134	-0.0371	0.0873	0.2359	1.2976**	2.3849	0.2295	0.5975	0.0322	0.0837
			2	0.1266	0.3195	-0.1159	-0.2637	0.9186	1.4562	0.4474	0.9707	0.1225	0.2819
			3	-0.0816	-0.1708	0.3772	0.7265	1.2733	1.7756	0.4366	0.8706	-0.265	-0.5183
3	m032064 (Number)	DIF		-.1562	-.4346	.0473	.1383	-.5927	-1.7185	.2119	.6254	-.3027	-.7706
		DDF											
4	m032419 (Algebra)	DIF**	3	0.4378	1.7877	0.5482**	2.1215	0.4084	1.4472	-0.2313	-0.8849	0.0613	0.2472
		DDF**	1	0.2273	0.6562	0.453	1.228	0.734	1.7623	0.1086	0.2903	0.0471	0.1359
			2	0.5395	1.9323	0.7858**	2.636	0.2923	0.8947	-0.1824	-0.6185	0.1098	0.3845
			4	0.6656	1.7177	0.0019	0.0047	0.0115	0.0254	0.0354	0.0846	-0.4935	-1.2573

5	m032477 (Algebra)	DIF**	1	-0.8508**	-2.8275	-0.281	-0.9894	0.1411	0.411	0.0419	0.14	-0.3379	-1.1215
		DDF**	2	-0.5662	-1.6886	-0.2878	-0.8782	0.2106	0.5438	0.0724	0.2183	-0.2281	-0.6508
			3	-1.0006**	-2.6619	-0.0156	-0.0428	0.4123	0.9559	-0.4541	-1.0563	-0.4081	-1.146
			4	-1.2096**	-2.3643	-0.4856	-1.0598	-0.8092	-1.0053	-0.0182	-0.0354	-0.5192	-1.1241
6	m032538 (Algebra)	DIF		.5668	1.8963	.0519	.1757	-.1058	-.3295	.3537	1.1825	.178	.583
		DDF											
7	m032324 (Geometry)	DIF	4	0.1167	0.4759	0.0779	0.3106	0.261	0.9223	0.0504	0.1994	-0.1267	-0.5038
		DDF	1	0.1066	0.331	0.2423	0.7778	0.5737	1.6345	-0.0803	-0.2514	-0.0625	-0.191
			2	0.1935	0.5581	0.0038	0.0099	-0.0723	-0.169	0.5015	1.2658	-0.3913	-1.0809
			3	-0.0409	-0.1306	0.0015	0.0044	-0.2155	-0.5471	0.2917	0.8648	-0.4017	-1.2573
8	m032116 (Geometry)	DIF	2	0.0655	0.281	0.0269	0.239	0.4967	1.7187	-0.2327	-0.909	-0.0304	-0.1271
		DDF**	1	-0.0914	-0.329	0.0712	0.2789	0.4859	1.4418	-0.2947	-0.9919	0.1863	0.6503
			3	0.4545	1.3706	0.1068	0.3429	1.1946**	2.9431	-0.2419	-0.6626	-0.2116	-0.6385
			4	-0.1829	-0.5285	-0.4649	0.375	-0.0887	-0.1828	0.0028	0.0069	-0.1644	-0.4688

9	m032100 (Geometry)	DIF**	4	-0.5452**	-2.114	-0.0341	0.2525	0.3787	1.2657	-0.5596	-1.9656	0.0508	0.2002
		DDF**	1	-0.36	-1.162	0.1679	0.3133	0.54	1.4803	-0.5619	-1.6273	0.4331	1.3237
			2	-0.7004**	-2.0624	-0.2761	0.3356	-0.0007	-0.0017	-0.4296	-1.1018	-0.0155	-0.0474
			3	-0.4062	-0.865	-0.3301	0.4458	0.6865	1.3716	-0.7282	-1.4494	-0.0336	-0.0771
10	m032402 (Geometry)	DIF	1	0.0899	0.3677	0.267	0.2624	0.0361	0.123	0.0799	0.2859	0.4083	1.6267
		DDF**	2	0.1451	0.384	0.6379	0.3865	0.229	0.5558	0.4459	0.9975	0.1867	0.494
			3	-0.0282	-0.0923	0.3292	0.3257	0.0801	0.21	-0.1173	-0.3311	0.8159**	2.4852
			4	1.1072	1.549			-0.9321	-0.8407	0.0389	0.0566	-0.4092	-0.6412
11	m032734 (Geometry)	DIF		-.3597	-1.3856	-.2014	-.7983	.3272	1.1389	-.1013	-.3804	.4586	1.7517
		DDF											
12	m032397 (Geometry)	DIF	2	-0.2605	-1.0313	-0.5049	-1.9517	0.1399	0.468	0.0419	0.152	-0.0715	-0.2772
		DDF**	1	-0.1903	-0.5984	-0.5655	-1.7283	0.0606	0.1586	-0.1335	-0.3805	0.2706	0.8136
			3	0.1254	0.3388	-0.3467	-0.8885	0.4012	0.8637	0.4187	0.9182	-0.0262	-0.068
			4	-0.5099	-1.402	-0.8515**	-2.1901	0.2546	0.5891	-0.0975	-0.2472	-0.4048	-1.138
13	m032695 (Data and Chance) (Polytomous)	DIF											
		DDF											

14	m032132 (Data and Chance)	DIF	4	0.1702	0.7729	0.0169	0.0731	-0.21	-0.8232	0.4606	1.8754	-0.3828	-1.6991
		DDF**	1	-0.0401	-0.0788	0.0683	0.1184	-0.2629	-0.3831	-0.2042	-0.3687	-0.0651	-0.1186
			2	0.2827	0.6773	0.1036	0.2309	-0.1098	-0.2136	0.4098	0.8613	-0.64	-1.4655
			3	0.1707	0.7062	0.0035	0.0138	-0.2047	-0.7425	0.5594**	2.0166	-0.3759	-1.5425
15	m042041 (Number)	DIF**	3	0.1541	0.585	0.0439	0.1572	-0.8214**	-2.4681	0.435	1.4146	-0.2011	-0.7658
		DDF**	1	0.4003	1.1457	-0.081	-0.2106	-0.9886**	-2.1183	0.5496	1.246	-0.5029	-1.4024
			2	0.2451	0.6869	0.3386	0.8845	-0.621	-1.3123	0.0465	0.1168	0.0113	0.0319
			4	-1.0693	-1.5234	0.285	0.4596	-0.2138	-0.3103	0.3902	0.5193	0.678	1.0928
16	m042024 (Number)	DIF**	2	-0.3522	-1.3731	-0.0737	-0.2793	-0.3251	-1.0722	0.7865**	2.7357	0.2067	0.791
		DDF**	1	-0.3333	-1.1513	-0.0213	-0.0686	-0.1248	-0.3647	1.1746**	3.2412	0.3563	1.1674
			3	-0.7196	-1.8556	-0.0354	-0.0904	-0.5469	-1.1575	0.2503	0.6188	0.2034	0.543
			4	0.1004	0.1852	0.2471	0.4634	-0.6828	-0.9309	0.9763	1.5331	-0.0076	-0.0138
17	m042016 (Number)	DIF	1	0.0141	0.0599	-0.2312	-0.9649	0.3298	1.1728	-0.5074	-1.9011	0.3607	1.4936
		DDF**	2	-0.2484	-0.7057	-0.5781	-1.5241	-0.0094	-0.0209	-0.5863	-1.4023	0.7784**	2.1373
			3	0.6148	1.6918	-0.5909	-1.4543	0.7242	1.63	-0.8306	-1.9701	0.5854	1.5026
			4	-0.2083	-0.7233	-0.0866	-0.3054	0.4799	1.4609	-0.4613	-1.4672	0.0012	0.0042

18	m042002 (Number)	DIF		-.5589	-1.8247	-.2337	-.8042	.4342	1.336	-.3409	-1.124	-.3029	-.9854
		DDF											
19	m042198a (Algebra)	DIF		-.0864	-.2381	.1512	.4046	-.7844	-1.8314	.0661	.1679	-.5825	-1.661
		DDF											
20	m042198b (Algebra)	DIF		-.1871	-.598	.1946	.6401	-.2848	-.8306	-.5077	-1.5251	.3107	.978
		DDF											
21	m042198c (Algebra)	DIF		-.404	-.8733	-.8241	-1.8366	-.612	-1.4431	.8113	1.8182	-.4896	-.9703
		DDF											
22	m042077 (Algebra)	DIF**	3	-0.4929	-1.7454	0.2966	1.0099	-0.8603**	-2.7959	0.4622	1.606	-0.2959	-1.0456
		DDF**	1	-0.508	-1.5214	0.501	1.5159	-1.0674**	-2.7746	0.3244	0.9658	-0.4301	-1.3006
			2	-0.5951	-1.4176	-0.0397	-.0896.	-1.0913**	-2.2381	1.0005**	2.185	-0.0374	-0.0962
			4	-0.8203**	-2.2554	3409	0.8508	-0.5	-1.2814	0.5636	1.4522	-0.3019	-0.8162
23	m042235 (Algebra)	DIF	4	0.1713	0.6066	-0.4201	-1.4956	-0.5221	-1.6718	0.0018	0.0058	-0.0015	-0.0053
		DDF**	1	0.8531**	2.0552	-0.4497	-1.0527	-0.8494	-1.6035	-0.3105	-0.645	-0.6603	-1.5755
			2	0.3177	0.8645	-0.4624	-1.1992	-0.4151	-0.9714	0.1954	0.4656	0.3852	0.9563
			3	0.184	0.5303	-0.179	-0.5282	-0.5895	-1.5579	-0.0434	-0.1143	-0.0067	-0.0198

24	m042067 (Algebra)	DIF	2	-0.0436	-0.1888	0.2353	0.9924	0.4067	1.4752	-0.2139	-0.8342	0.1145	0.4817
		DDF	1	-0.2631	-0.9599	0.4253	1.5233	0.3896	1.2298	-0.2107	-0.6949	0.0535	0.1958
			3	0.2538	0.6479	0.0752	0.1943	0.0606	0.1266	0.039	0.0938	0.1594	0.3978
			4	0.3055	0.8766	0.0108	0.0279	0.5973	1.3826	-0.3596	-0.8688	0.2453	0.6443
25	m042150 (Geometry)	DIF	2	0.3084	1.2516	-0.109	-0.4369	-0.0937	-0.3327	0.0228	0.0872	-0.3021	-1.188
		DDF	1	-0.6674	-1.181	-0.317	-0.6002	1.217	1.8646	0.2889	0.4608	0.2461	0.4562
			3	0.4921	1.8898	-0.0533	-0.204	-0.0274	-0.091	-0.0595	-0.217	-0.2409	-0.9006
			4	0.3384	0.6893	-0.5245	-0.9536	-0.039	-0.0663	-0.0046	-0.0079	-0.8022	-1.526
26	m042300a (Geometry)	DIF**		1.0257**	3.2356	.0398	.1326	.175	.5115	-.0606	-.1913	-.7228**	-2.3598
		DDF											
27	m042300b (Geometry)	DIF		-.0351	-.1181	-.2596	-.9017	.1374	.4071	.5188	1.7731	-.3739	-1.2938
		DDF											
28	m042260 (Data and Chance)	DIF	2	0.1994	0.8918	-0.185	-0.8036	-0.4856	-1.8591	-0.0976	-0.3951	0.1159	0.5112
		DDF**	1	0.0156	0.0626	-0.1109	-0.4346	-0.4078	-1.4329	-0.0918	-0.3418	0.2327	0.9125
			3	0.8715**	2.0859	-0.4397	-0.9624	-0.7746	-1.4621	-0.2334	-0.4781	-0.3349	-0.8275
			4	1.2474	1.553	-0.0183	-0.0247	-0.3133	-0.3919	-0.4936	-0.646	0.7052	0.9845

29	m042169a (Data and Chance)	DIF		-.3227	-.9607	.576	1.7669	-.1871	-.5482	-.1124	-.3341	.2767	.8287
		DDF											
30	m042169b (Data and Chance)	DIF**		.2566	.5426	.2974	.658	1.3632**	2.3023	-.8172	-1.6068	1.6548**	3.2045
		DDF											
31	m042169c (Data and Chance)	DIF**				1.3042	1.2317	1.7538	1.5081	-1.3407	-1.1986	3.1715**	2.0843
		DDF											

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 10

ข้อ	ข้อสอบ	Option	Response	GIRL		HOMEWK		SES		EQUIP		RICH	
				LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)
1	m042041 (Number)	DIF**	3	0.2394	0.913	-0.1158	-0.4134	-0.6355**	-2.0586	0.1605	0.5187	-0.3346	-1.2397
		DDF**	1	0.5008	1.4579	-0.2097	-0.5674	-1.1147**	-2.4055	0.1717	0.4179	-0.5217	-1.5043
			2	0.4349	1.2447	-0.0131	-0.0345	-0.0889	-0.227	-0.0434	-0.1073	-0.17	-0.4592
			4	-1.4785	-1.4516	0.099	0.1159	-0.6211	-0.6567	0.0139	0.0148	0.7219	0.7972
2	m042024 (Number)	DIF	2	-0.2881	-1.0245	-0.0884	-0.3174	0.3581	1.1921	-0.3144	-1.0342	0.4914	1.7068
		DDF	1	-0.6866	-1.9792	-0.1164	-0.3559	0.6124	1.7537	-0.4866	-1.3979	0.602	1.7505
			3	0.1525	0.4121	0.0066	0.0174	0.3706	0.9385	-0.2871	-0.6875	0.4921	1.3311
			4	0.0626	0.1003	0.4921	0.707	-1.379	-1.2782	-0.0154	-0.0214	0.1508	0.2643
3	m042016 (Number)	DIF	1	-0.1475	-0.5907	0.2202	0.8731	-0.2326	-0.8811	-0.0316	-0.1174	0.107	0.4131
		DDF	2	-0.1658	-0.4413	0.2266	0.619	-0.379	-0.8886	0.1204	0.3043	-0.0559	-0.1478
			3	0.5147	1.3327	0.16	0.3903	-0.1785	-0.4191	0.389	0.8396	0.2307	0.6126
			4	-0.4877	-1.5876	0.2161	0.7296	-0.1061	-0.3455	-0.073	-0.2317	0.1056	0.3379
4	m042002 (Number)	DIF**		-.071**	-.2588	.36	1.2672	-.2967	-1.027	-.2122	-.7142	-.0131	-.0466
		DDF											

5	m042198a (Algebra)	DIF**		.9843**	2.7487	-.5235	-1.466	-.1361	-.3548	-.0525	-.141	-.241	-.7143
		DDF											
6	m042198b (Algebra)	DIF**		-.0102	-.0353	-.2326	-.7858	-.5098	-1.6308	.267	.853	-.818**	-2.7285
		DDF											
7	m042198c (Algebra)	DIF		-.5051	-1.2121	-.0402	-.1046	-.1925	-.4663	.3305	.8691	-.0152	-.0339
		DDF											
8	m042077 (Algebra)	DIF	3	0.0499	0.1774	-0.303	-1.0943	-0.3682	-1.2834	-0.06	-0.2007	0.018	0.0644
		DDF	1	0.1586	0.507	-0.2455	-0.793	-0.4417	-1.3825	0.2297	0.678	0.0718	0.2254
			2	0.6051	1.3549	-0.3097	-0.7328	-0.2488	-0.5434	-0.0786	-0.1654	-0.0142	-0.0339
			4	-0.561	-1.3784	-0.3245	-0.8141	-0.1729	-0.4079	-0.3958	-0.9689	0.0656	0.1771
9	m042235 (Algebra)	DIF**	4	0.6947**	2.3881	0.1675	0.576	-0.4579	-1.6135	-0.1774	-0.5717	-0.1118	-0.3898
		DDF**	1	0.6788	1.4763	0.5078	1.0606	-0.9453	-1.7961	-0.4018	-0.8085	-0.3575	-0.8388
			2	0.9035**	2.185	-0.5964	-1.2535	-0.5396	-1.2667	-0.5665	-1.3053	0.1162	0.2803
			3	0.72**	2.1518	0.4293	1.2982	-0.2853	-0.869	-0.014	-0.0386	-0.1291	-0.3878

10	m042067 (Algebra)	DIF	2	0.0895	0.3772	0.0544	0.226	0.4272	1.6355	-0.0847	-0.3301	-0.0481	-0.1951
		DDF	1	0.0149	0.0539	0.2178	0.7497	0.4134	1.3812	0.2394	0.759	0.0533	0.1834
			3	-0.0069	-0.0208	0.0311	0.0913	0.4928	1.3201	-0.396	-1.0924	-0.1696	-0.4933
			4	0.1102	0.2861	-0.151	-0.4127	0.1211	0.2866	-0.3887	-0.9949	-0.0057	-0.0154
11	m042150 (Geometry)	DIF	2	-0.1363	-0.5532	-0.4011	-1.6371	0.0858	0.3277	-0.4869	-1.7066	0.0253	0.0944
		DDF**	1	-0.1873	-0.3861	-0.3555	-0.6684	0.2611	0.4634	0.265	0.4123	0.1574	0.2521
			3	-0.2542	-0.9684	-0.4366	-1.6922	0.1108	0.4074	-0.668**	-2.2476	0.066	0.239
			4	0.4416	0.8296	-0.4282	-0.8545	-0.481	-0.7143	-0.204	-0.306	-0.0997	-0.185
12	m042300a (Geometry)	DIF		.1544	.5564	-.0591	-.2054	.0611	.2049	-.3493	-1.1479	.3469	1.2744
		DDF											
13	m042300b (Geometry)	DIF		-.0197	-.0666	.014	.0449	-.431	-1.3948	.0004	.0012	.2515	.8118
		DDF											
14	m042260 (Data and Chance)	DIF	2	0.121	0.5105	0.0115	0.0479	-0.3584	-1.4268	0.1431	0.5529	0.2453	0.9984
		DDF	1	0.0093	0.037	0.0377	0.1515	-0.3138	-1.1913	-0.0082	-0.0306	0.3098	1.1911
			3	0.7663	1.5572	0.0013	0.0024	-0.2233	-0.4258	0.5075	0.8347	0.0869	0.1782
			4	0.9185	1.5301	0.2708	0.3748	-0.9491	-1.0308	0.6733	0.7575	0.1767	0.2972

15	m042169a (Data and Chance)	DIF**		-.1506	-.477	-.0825	-.2578	.2595	.7765	.1631	.5078	.695**	2.0611
		DDF											
16	m042169b (Data and Chance)	DIF**		-.0304	-.0588	.3069	.5826	2.6446**	2.5703	-1.8832**	-2.5274	.8814	1.42
		DDF											
17	m042169c (Data and Chance)	DIF		.7124	.7817	-1.0356	-1.2559	.3685	.3052	-.943	-1.0892		
		DDF											
18	m032352 (Algebra)	DIF	3	-0.1376	-0.5405	-0.2154	-0.8301	-0.2711	-1.0086	0.369	1.3433	-0.2919	-1.1279
		DDF**	1	-1.4609**	-2.9406	-0.6368	-1.5356	-0.0419	-0.1071	0.3644	0.927	-0.355	-0.9842
			2	0.1488	0.4494	-0.3356	-0.9685	-0.6191	-1.6678	0.3886	1.0623	-0.4434	-1.3003
			4	0.5746	1.4551	0.443	1.0739	-0.2851	-0.6273	0.3757	0.8121	-0.228	-0.5601
19	m032725 (Number)	DIF		.3339	.9973	-.2171	-.6647	-.24	-.722	.607	1.7916	-.1626	-.4571
		DDF											
20	m032683 (Algebra) (Polytomous)	DIF											
		DDF											
21	m032738 (Algebra)	DIF**	4	0.5345**	2.0246	0.1315	0.4885	0.0195	0.0666	0.5461	1.8264	-0.3216	-1.2369
		DDF**	1	1.7669**	2.8819	-0.3536	-0.5931	0.8404	1.4911	-0.3319	-0.644	0.2787	0.5238
			2	0.5078	1.169	-0.1058	-0.2143	0.3361	0.7163	0.1801	0.3701	0.09	0.2058
			3	0.4798	1.5956	0.3678	1.2031	-0.2888	-0.7956	0.9607**	2.5176	-0.3919	-1.3107

22	m032295 (Algebra)	DIF**	1	0.1779	0.6209	-0.4357	-1.4855	-0.1649	-0.506	0.2032	0.6154	-0.7012**	-2.3778
		DDF**	2	0.0882	0.2794	-0.2112	-0.6633	-0.0396	-0.11	0.0571	0.1618	-0.7979**	-2.4672
			3	0.1766	0.3249	-1.0757	-1.7208	-0.4422	-0.67	0.3904	0.6227	-0.4373	-0.8311
			4	1.1905	1.8763	-1.7081**	-2.1359	0.0717	0.1151	-0.0539	-0.0795	0.1482	0.2385
23	m032331 (Geometry)	DIF	2	0.2867	1.0035	0.2248	0.7503	0.0235	0.0799	0.1116	0.3609	0.3856	1.2042
		DDF	1	0.231	0.617	0.2155	0.5583	-0.4396	-1.1559	0.4226	1.0499	0.1868	0.4844
			3	0.4184	1.211	0.1965	0.5592	-0.177	-0.4908	0.1361	0.3494	0.5612	1.4926
			4	0.079	0.2375	0.4617	1.2458	0.1922	0.563	-0.2959	-0.7922	0.387	0.9765
24	m032623 (Geometry)	DIF	4	0.0449	0.1651	0.1557	0.5456	0.3072	0.9846	-0.2662	-0.8601	0.367	1.2177
		DDF	1	-0.3663	-1.0633	-0.0687	-0.1924	0.4118	1.0724	-0.542	-1.4188	0.4267	1.2437
			2	0.1134	0.295	0.541	1.3586	-0.1929	-0.4382	0.1774	0.3916	0.3585	0.8767
			3	0.3663	1.108	-0.0306	-0.0898	0.5964	1.5129	-0.2157	-0.5709	0.3863	1.0068
25	m032679 (Geometry)	DIF	1	-0.0197	-0.0801	-0.0836	-0.3439	-0.0943	-0.3616	0.5131	1.9032	-0.2075	-0.8404
		DDF	2	-0.0323	-0.0931	-0.4374	-1.2726	-0.0506	-0.1415	0.7636	1.9302	0.033	0.0984
			3	0.0329	0.0908	-0.2022	-0.5549	0.1861	0.4769	0.3065	0.7713	-0.3178	-0.8867
			4	0.1372	0.3841	0.2889	0.8182	0.0068	0.0175	0.4319	1.0792	-0.1513	-0.4239

26	m032047 (Algebra)	DIF	2	-0.2944	-1.319	0.2374	1.0211	0.124	0.5259	-0.3447	-1.4291	-0.0801	-0.342
		DDF	1	-0.3528	-1.3298	0.2651	0.9775	0.3687	1.3451	-0.3496	-1.2778	0.0868	0.3123
			3	-0.3287	-0.6474	-0.1402	-0.2471	0.1198	0.225	-0.431	-0.8741	-0.4647	-0.9001
			4	-0.112	-0.3187	0.3143	0.8449	-0.1559	-0.4061	0.0408	0.0967	-0.0242	-0.0673
27	m032398 (Geometry)	DIF**	2	-0.0957	-0.3953	0.0678	0.2759	0.7613**	2.8783	-0.2392	-0.9175	0.6266**	2.3971
		DDF**	1	0.0005	0.0018	-0.0255	-0.0889	1.0175**	3.3548	-0.3979	-1.3228	0.9054**	2.8964
			3	-0.1336	-0.3819	0.3273	0.9296	0.3893	1.0334	-0.0002	-0.0005	0.3793	1.0415
			4	-0.1942	-0.4457	-0.5139	-1.0785	0.4649	1.0043	-0.3098	-0.6982	0.424	1.0062
28	m032507 (Data and Chance)	DIF	3	-0.5611	-1.7723	0.6634	1.9196	0.6009	1.6495	0.1145	0.3367	0.1749	0.5067
		DDF**	1	-0.4829	-1.2648	0.325	0.729	0.483	1.0985	-0.1891	-0.4575	0.3725	0.8128
			2	0.0432	0.097	0.9821	1.8255	0.4089	0.7908	-0.5922	-1.1291	0.4247	0.8664
			4	-0.6578	-1.8047	0.864**	2.2154	0.7357	1.7702	0.3748	0.972	-0.0182	-0.0501
29	m032424 (Algebra)	DIF	3	-0.4068	-1.6149	-0.106	-0.4232	-0.1307	-0.5	0.0223	0.0813	-0.163	-0.6415
		DDF	1	-0.4501	-1.1734	-0.1883	-0.4934	-0.0385	-0.0911	-0.0996	-0.2291	-0.2564	-0.6593
			2	-0.3105	-0.8715	-0.6141	-1.5955	0.35	0.9597	-0.408	-1.0249	0.1506	0.4138
			4	-0.4586	-1.4417	0.2532	0.8131	-0.395	-1.2017	-0.0817	-0.248	-0.0621	-0.2052

30	m032681a (Data and Chance)	DIF		.1516	.5576	.1839	.6717	.4577	1.5909	.1557	.5235	.2181	.7862
		DDF											
31	m032681b (Data and Chance)	DIF		-.1643	-.6575	.0559	.22	.0741	.2763	-.061	-.2262	-.2627	-.9939
		DDF											
32	m032681c (Data and Chance)	DIF**		-.6846**	-2.2744	.2607	.8873	.1071	.3509	.1096	.3616	-.1262	-.4157
		DDF											

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 11

ข้อ	ข้อสอบ	Option	Response	GIRL		HOMEWK		SES		EQUIP		RICH	
				LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)
1	m032352 (Algebra)	DIF	3	-0.057	-0.2275	0.0548	0.2093	-0.202	-0.7495	-0.056	-0.1928	0.2899	1.1069
		DDF**	1	-0.1083	-0.2788	-0.0281	-0.0659	-0.0727	-0.1693	-0.1813	-0.4154	0.1514	0.3682
			2	-0.0649	-0.2207	0.0415	0.1368	-0.1652	-0.5266	-0.2194	-0.6491	0.2855	0.9382
			4	0.0806	0.1663	1.1933**	2.2127	-0.8887	-1.5483	0.8025	1.1269	0.3213	0.655
2	m032725 (Number)	DIF**		.0755	.2292	-.8845 **	-2.657	-.4004	-1.2256	-.0347	-.1011	-.4316	-1.2012
		DDF											
3	m032683 (Algebra) (Polytomous)	DIF											
		DDF											
4	m032738 (Algebra)	DIF**	4	0.6996**	2.7917	-0.443	-1.6889	0.2668	1.0222	-0.339	-1.2021	0.1885	0.7492
		DDF**	1	0.6176	1.0756	-0.5247	-0.8667	-0.152	-0.2147	-0.6907	-1.1669	0.054	0.0986
			2	0.893	1.7771	-0.6801	-1.2354	0.4896	1.0314	-1.1244**	-2.2601	0.4354	0.9084
			3	0.6401**	2.2618	-0.4291	-1.457	0.3048	1.051	0.0061	0.0183	0.1226	0.4329
5	m032295 (Algebra)	DIF**	1	0.4068	1.5934	-0.2538	-0.9467	-0.3653	-1.3107	0.6596**	2.1762	-0.747**	-2.8643
		DDF**	2	0.3716	1.3702	-0.184	-0.6493	-0.4696	-1.5488	0.7316**	2.2069	-0.6512**	-2.3467
			3	-0.3512	-0.6426	-0.752	-1.1263	-0.1751	-0.3093	0.3315	0.5614	-1.5119**	-2.5127
			4	1.6129	1.8556	0.4092	0.5544	-0.3366	-0.4522	1.2039	1.3057	-0.9521	-1.4616

6	m032331 (Geometry)	DIF**	2	-0.687**	-2.3887	-0.0942	-0.3252	-0.3775	-1.3255	0.292	0.9226	-0.1707	-0.5595
		DDF**	1	-0.8084**	-2.0744	0.0807	0.2172	-0.1218	-0.3454	0.4651	1.0874	-0.0597	-0.1494
			3	-0.3717	-1.1219	-0.0919	-0.2644	-0.4655	-1.3373	0.416	1.1008	0.0083	0.0221
			4	-1.3319**	-3.5833	0.0219	0.0619	-0.4794	-1.376	-0.146	-0.3802	-0.1554	-0.4543
7	m032623 (Geometry)	DIF	4	-0.2551	-0.8803	-0.3784	-1.2866	-0.3396	-1.1197	-0.0123	-0.0363	-0.0642	-0.2124
		DDF	1	-0.0235	-0.0646	-0.2426	-0.6965	-0.3772	-0.9569	-0.1389	-0.3137	-0.1729	-0.4545
			2	-0.3279	-0.7998	-0.4715	-1.1094	-0.5026	-1.1829	-0.1481	-0.3436	0.1204	0.3094
			3	-0.3325	-0.9869	-0.3777	-1.0385	-0.2706	-0.7679	-0.1132	-0.2775	-0.2707	-0.7768
8	m032679 (Geometry)	DIF**	1	0.3158	1.3513	-0.3579	-1.523	-0.4099	-1.6765	0.052	0.1985	-0.5815**	-2.4989
		DDF**	2	0.0755	0.2449	-0.2929	-1.523	-0.5989	-1.7651	0.1106	0.3061	-0.8186**	-2.6466
			3	0.8506**	2.1996	-0.8206	-1.523	-0.4891	-1.2301	-0.1012	-0.2411	-0.8129**	-2.295
			4	0.3615	1.1199	-0.2894	-1.523	-0.2785	-0.8356	-0.0557	-0.1576	-0.1701	-0.5192
9	m032047 (Algebra)	DIF	2	-0.0346	-0.1571	0.1191	-1.523	0.0565	0.2367	-0.1514	-0.5979	0.3366	1.4409
		DDF**	1	0.071	0.2574	0.1504	-1.523	0.2346	0.7878	-0.073	-0.2273	0.1736	0.5857
			3	-1.0483**	-2.1042	-0.0596	-1.523	-0.5469	-1.0558	0.2832	0.5113	0.6839	1.4573
			4	0.2159	0.5704	0.4356	-1.523	0.48	1.2223	-0.7437	-1.8431	1.3555**	3.0557

10	m032398 (Geometry)	DIF	2	-0.0755	-0.3419	-0.1069	-1.523	0.1223	0.5209	-0.4108	-1.5941	-0.0619	-0.2688
		DDF**	1	-0.3657	-1.2927	0.1378	-1.523	0.1195	0.4132	-0.7011**	-2.2536	0.3068	1.034
			3	0.1484	0.4444	-0.0267	-1.523	0.1023	0.289	-0.414	-1.0944	-0.3165	-0.9465
			4	0.5086	1.4143	-1.0884	-1.523	0.3567	0.8812	0.1815	0.3921	-0.3131	-0.8367
11	m032507 (Data and Chance)	DIF	3	-0.2514	-0.9192	0.3436	1.1735	0.1574	0.5469	-0.0382	-0.1263	0.4494	1.5768
		DDF	1	-0.3435	-0.9865	0.5951	1.6489	0.1393	0.3869	-0.2786	-0.7127	0.381	1.0861
			2	-0.4237	-0.9792	0.7084	1.6199	0.1111	0.2394	0.0024	0.0053	0.6092	1.4641
			4	-0.3101	-1.0181	0.0744	0.2247	0.273	0.8675	0.0194	0.0576	0.3413	1.0692
12	m032424 (Algebra)	DIF**	3	-0.3412	-1.3938	-0.0703	-0.2858	-0.6162**	-2.3148	0.715**	2.5328	-0.4469	-1.7699
		DDF**	1	-0.1634	-0.4563	-0.6398	-1.5739	-0.8307**	-2.0119	1.4107**	2.7737	-0.5855	-1.6414
			2	-0.3701	-1.0041	-0.2536	-0.6863	-0.4451	-1.1155	0.2958	0.7438	-0.46	-1.2014
			4	-0.5152	-1.6191	0.1417	0.4572	-0.3759	-1.1079	0.8536**	2.2763	-0.3296	-1.0329
13	m032681a (Data and Chance)	DIF**		.655**	2.2688	.276	.9375	.2856	.9539	-.0522	-.1591	-.5513	-1.865
		DDF											
14	m032681b (Data and Chance)	DIF**		.1879	.7723	.5831**	2.3408	-.0419	-.1665	-.0096	-.0362	.1801	.7099
		DDF											

15	m032681c (Data and Chance)	DIF**		-.0679	-.249	.1277	.462	-.6281**	-2.2384	.4603	1.5916	-.232	-.7951
		DDF											
16	m042015 (Number)	DIF**	3	0.5809**	2.2997	-0.5775**	-2.1941	-0.3139	-1.1635	0.3132	1.0822	-0.1801	-0.7181
		DDF	1	0.1129	0.1272	-1.9201	-1.2495	-0.4864	-0.4352	-0.9349	-0.9137	-0.2399	-0.2284
			2	0.5035	1.9306	-0.4602	-1.7114	-0.298	-1.0735	0.39	1.2808	-0.1679	-0.6475
			4	1.3831	1.8385	-2.0417	-1.7413	-0.0696	-0.0927	-0.0836	-0.1146	-0.207	-0.3201
17	m042196 (Number)	DIF	3	-0.3818	-1.4534	0.2634	0.9799	-0.3826	-1.3989	0.3436	1.1419	-0.2282	-0.876
		DDF	1	-0.4887	-1.6023	0.2261	0.7452	-0.2761	-0.8773	0.4678	1.2616	-0.2188	-0.7183
			2	-0.4597	-1.3236	0.5165	1.3888	-0.759	-1.9807	0.2624	0.6789	-0.1777	-0.5234
			4	-0.0725	-0.0642	0.6444	0.4865			-2.6551	-1.6644		
18	m042194 (Number)	DIF		-.3141	-1.1716	-.3715	-1.3263	-.114	-.3997	.0448	.1431	-.2879	-1.0727
		DDF											
19	m042114a (Number)	DIF**		-.6909**	-2.3239	.5581	1.8523	.8004**	2.4582	-.1633	-.5206	.646**	2.0994
		DDF											
20	m042114b (Number)	DIF		-.5595	-1.6934	.3686	1.1384	.8434	2.3	-.4804	-1.3884	.6125	1.7764
		DDF											

21	m042112 (Algebra)	DIF	4	0.3586	1.4649	0.0646	0.2569	0.3587	1.4078	-0.392	-1.4198	-0.1053	-0.4033
		DDF	1	0.0506	0.1146	-0.0511	-0.112	0.0142	0.0301	-0.6271	-1.3252	-0.4095	-0.8398
			2	0.6561	1.6329	-0.0077	-0.0198	0.1626	0.3892	-0.3411	-0.7878	-0.0312	-0.0762
			3	0.5179	1.5341	0.0041	0.0117	0.5564	1.5651	-0.3571	-0.941	-0.1609	-0.4428
22	m042109 (Algebra)	DIF**	4	-0.2525	-0.9852	0.199	0.7719	0.7658**	2.551	0.126	0.4471	0.4302	1.5689
		DDF**	1	-0.0198	-0.0413	-2.2183**	-2.5758	0.3051	0.5011	0.5557	0.9741	-0.0236	-0.0434
			2	-0.1158	-0.374	0.2628	0.8109	0.6946	1.9381	0.0893	0.2612	0.6159	1.7562
			3	-0.2347	-0.8241	0.4054	1.429	0.9719**	3.0202	0.1142	0.3691	0.4229	1.4538
23	m042050 (Algebra)	DIF**		.7765	1.802	-2.2338**	-4.5173	-.9412**	-2.4491	.3213	.848	-.855	-1.7872
		DDF											
24	m042074a (Algebra)	DIF		-.0339	-.1047	.5235	1.5902	.2614	.8009	-.2796	-.8161	.5008	1.4272
		DDF											
25	m042074b (Algebra)	DIF		.0389	.0911	.1958	.4836	.1651	.3922	-.3641	-.8332	.7161	1.5554
		DDF											
26	m042074c (Algebra)	DIF		.3926	.6887	-.2711	-.5266	.4344	.8015	-.6104	-1.0809	-.211	-.3382
		DDF											

27	m042151 (Geometry)	DIF		.072	.28	-.2005	-.7847	-.1778	-.6883	.4043	1.4286	-.1144	-.441
		DDF											
28	m042132 (Geometry)	DIF	1	0.0287	0.1077	0.4627	1.6472	0.1583	0.5584	-0.295	-0.9634	0.2552	0.9338
		DDF	2	-0.0276	-0.0858	0.418	1.247	0.4793	1.4147	-0.396	-1.0735	0.5107	1.557
			3	0.6056	1.6916	0.0367	0.0973	-0.0719	-0.1917	-0.4592	-1.1631	-0.0655	-0.1909
			4	-0.1721	-0.4884	0.4226	1.1553	0.0297	0.0763	-0.3216	-0.741	0.3524	0.9458
29	m042257 (Geometry)	DIF	3	0.2263	1.0203	0.2102	0.9264	0.3081	1.2978	-0.4298	-1.6789	0.3693	1.5715
		DDF	1	0.4035	1.3004	0.434	1.3268	0.1898	0.5679	-0.14	-0.3816	0.4379	1.3408
			2	0.1841	0.6905	0.0513	0.1886	0.177	0.6281	-0.5185	-1.6758	0.2729	0.9674
			4	-0.0645	-0.1195	-0.3673	-0.6379	0.3417	0.615	-1.0473	-1.7132	0.2943	0.5514
30	m042158 (Data and Chance)	DIF**	2	0.2946	1.3018	0.6102**	2.5867	0.3197	1.3547	-0.4862	-1.8896	0.1331	0.58
		DDF**	1	0.3263	1.1005	0.7966**	2.5763	0.0774	0.2494	-0.743**	-2.2413	0.1018	0.3374
			3	0.7185	1.9175	-0.082	-0.1973	0.5562	1.4757	-0.2981	-0.7027	-0.0357	-0.1016
			4	0.0255	0.0647	0.8049**	2.0133	0.8684**	2.0845	-0.5523	-1.2572	0.5083	1.1983

31	m042252 (Data and Chance)	DIF**	1	-0.6815**	-2.2884	-0.3543	-1.2035	-0.0044	-0.0141	0.3838	1.2146	0.4016	1.2338
		DDF**	2	-0.661**	-2.1071	-0.1879	-0.5999	-0.105	-0.3225	0.5843	1.7333	0.4653	1.3637
			3	-0.4204	-0.9203	-0.4831	-1.0781	0.2882	0.5882	-0.3713	-0.7968	0.4734	0.8435
			4	-0.5276	-0.7159	-0.8045	-1.2154	-0.3418	-0.4036	2.3116**	2.079	0.2478	0.3872
32	m042261 (Data and Chance)	DIF	3	-0.387	-1.7607	0.0641	0.289	0.1954	0.8347	0.1963	0.7868	-0.0594	-0.2634
		DDF**	1	1.0952	1.2	0.3558	0.3894	-0.0256	-0.027	0.456	0.4176	-0.619	-0.6944
			2	-0.5985**	-2.3921	0.2316	0.95	0.2346	0.92	0.1611	0.585	-0.0871	-0.3528
			4	-0.0706	-0.1854	-0.4522	-1.0434	0.1392	0.3214	0.2361	0.5088	0.0968	0.2358

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 12

ข้อ	ข้อสอบ	Option	Response	GIRL		HOMEWK		SES		EQUIP		RICH	
				LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)
1	m042015 (Number)	DIF**	3	0.5975**	2.2641	-0.5616**	-2.0846	-0.3031	-1.0669	0.0292	0.0996	-0.1323	-0.4816
		DDF**	1	0.8108	1.1013	-0.4608	-0.6101	-0.058	-0.0753	0.2007	0.2375	-0.1699	-0.2317
			2	0.5906**	2.189	-0.5301	-1.9326	-0.3658	-1.2438	0.034	0.1129	-0.2063	-0.7357
			4	0.3773	0.2496			0.3959	0.3067	-0.3567	-0.2734		
2	m042196 (Number)	DIF	3	-0.0923	-0.3637	0.0985	0.3837	0.4077	1.4401	0.1776	0.6532	-0.0167	-0.062
		DDF**	1	-0.4268	-1.4021	0.1861	0.6175	0.9391**	2.7532	0.0839	0.2608	-0.2712	-0.8681
			2	0.2659	0.8089	0.0703	0.2072	-0.1049	-0.2706	0.2281	0.6138	0.3242	0.9163
			4			-0.1975	-0.193	0.5033	0.4645	0.3542	0.3048	-0.8781	-0.9377
3	m042194 (Number)	DIF**		-.1834	-.639	-.2925	-1.0072	-.0751	-.2438	.6411**	2.0574	-.4148	-1.4565
		DDF											
4	m042114a (Number)	DIF**		.1787	.5888	.5916	1.9461	-.0011	-.0035	-.6337**	-2.0003	.0069	.0229
		DDF											
5	m042114b (Number)	DIF**		-.6881**	-2.245	.1718	.5736	-.21	-.677	-.7772**	-2.3892	.0698	.2297
		DDF											

6	m042112 (Algebra)	DIF	4	0.2743	1.1608	-0.1548	-0.6523	0.0145	0.0561	-0.3912	-1.4994	0.254	1.0043
		DDF**	1	0.3367	0.9149	0.0406	0.1079	0.5822	1.4647	-0.5716	-1.4215	0.9671**	2.2697
			2	0.51	1.2798	-0.1462	-0.3578	0.0315	0.0702	-0.4392	-0.956	0.8052	1.7413
			3	0.2189	0.6511	-0.3748	-1.0902	-0.1378	-0.3717	-0.2656	-0.7349	-0.3229	-0.9455
7	m042109 (Algebra)	DIF**	4	-0.84**	-3.3097	0.1025	0.4056	0.5966**	2.1845	-0.782**	-2.618	0.7257**	2.5535
		DDF**	1	-0.7548	-1.7509	-0.0989	-0.2409	-0.1869	-0.3926	-0.3125	-0.6021	-0.0321	-0.0729
			2	-0.9357**	-2.7028	-0.0388	-0.1162	0.3715	1.0057	-0.8494**	-2.1262	0.7679**	2.0107
			3	-0.8232**	-2.9956	0.1917	0.6839	0.7308**	2.5357	-0.9851**	-3.0698	0.9084**	2.8423
8	m042050 (Algebra)	DIF		.7141	1.6182	-.7026	-1.7552	.4429	1.0091	.2516	.6229	-.2778	-.5879
		DDF											
9	m042074a (Algebra)	DIF		-.0338	-.1049	-.4053	-1.297	.0828	.244	.2478	.771	.4021	1.235
		DDF											
10	m042074b (Algebra)	DIF		-.4081	-1.2635	.1238	.3833	.0737	.2196	.4378	1.3793	-.1256	-.3606
		DDF											
11	m042074c (Algebra)	DIF		.8363	1.5241	.0053	.0105	.263	.4991	.3634	.706	-.3185	-.4996
		DDF											

12	m042151 (Geometry)	DIF**		.0911	.3395	-.3478	-1.3115	-.5928**	-2.0793	.0609	.2203	-.4534	-1.6694
		DDF											
13	m042132 (Geometry)	DIF	1	0.1548	0.5873	0.2766	1.0271	-0.173	-0.6321	0.49	1.7903	-0.4027	-1.3934
		DDF	2	0.1568	0.5285	0.539	1.7144	0.0294	0.0929	0.3388	1.0873	-0.1017	-0.3182
			3	0.6415	1.5627	0.018	0.044	0.0265	0.0644	0.5574	1.2378	-0.5148	-1.1412
			4	0.1113	0.3253	0.1171	0.3414	-0.702	-1.8973	0.2696	0.7554	-0.7208	-1.9614
14	m042257 (Geometry)	DIF	3	-0.3917	-1.6957	0.2202	0.9434	0.2771	1.1408	0.148	0.5985	0.49	1.9307
		DDF	1	-0.4635	-1.4557	0.223	0.698	0.2272	0.6967	0.1202	0.3581	0.5754	1.6649
			2	-0.3913	-1.4271	0.3527	1.2957	0.2365	0.8238	0.1391	0.4827	0.3981	1.3504
			4	-0.5033	-0.9703	-0.1208	-0.2277	0.9204	1.7465	0.7794	1.2614	0.3043	0.558
15	m042158 (Data and Chance)	DIF**	2	0.0947	0.3964	-0.1645	-0.6823	-0.118	-0.4547	0.0603	0.2285	0.5229**	2.0554
		DDF**	1	0.2832	0.9529	0.0295	0.0969	-0.1128	-0.343	0.3157	0.8828	0.7681**	2.2826
			3	-0.2406	-0.4819	-0.7982	-1.5775	-0.1314	-0.2451	-0.0709	-0.1327	0.3148	0.6671
			4	-0.1917	-0.5031	-0.1574	-0.4276	-0.0019	-0.0048	-0.4804	-1.278	0.4996	1.3127
16	m042252 (Data and Chance)	DIF	1	-0.0758	-0.273	0.5633	1.8928	-0.3189	-1.1012	0.4502	1.5578	-0.0162	-0.051
		DDF**	2	-0.1377	-0.4729	0.6819**	2.219	-0.2528	-0.8256	0.5318	1.7882	0.0165	0.0507
			3	0.1223	0.3145	0.5665	1.3077	-0.2821	-0.6742	-0.4725	-1.057	0.0454	0.0968
			4	-0.2184	-0.3518	-1.3932	-1.1991	0.1897	0.297	1.6119	1.5295	0.5464	0.9169

17	m042261 (Data and Chance)	DIF	3	0.2308	1.0035	0.2255	0.9519	0.1996	0.8	0.0244	0.0979	-0.1378	-0.5666
		DDF**	1	0.8097	1.0467	-0.441	-0.4747	2.4446**	2.417	-0.0273	-0.033	-0.2914	-0.3517
			2	0.1106	0.452	0.2058	0.8272	0.0639	0.2391	0.0862	0.3275	-0.2103	-0.8228
			4	0.9095**	2.0274	0.603	1.3134	0.3706	0.7755	-0.0235	-0.0473	0.1436	0.3126
18	m032166 (Number)	DIF**	2	-0.0039	-0.0144	0.0866	0.3213	-0.3903	-1.3023	0.3153	1.0559	-0.5701**	-2.0252
		DDF**	1	0.0985	0.2534	-0.4622	-1.1555	-1.1666**	-2.2469	0.6574	1.431	-0.4515	-1.1148
			3	0.3226	0.7198	0.4227	0.9699	0.2531	0.5405	0.3301	0.642	-0.8654	-1.8354
			4	-0.384	-1.0512	0.5198	1.4339	-0.1981	-0.5174	-0.0531	-0.1415	-0.4283	-1.1601
19	m032721 (Data and Chance)	DIF	2	-0.1636	-0.7307	-0.1594	-0.693	-0.0272	-0.1132	-0.0963	-0.3904	-0.1799	-0.7337
		DDF**	1	-0.0809	-0.3063	0.1148	0.4153	0.1613	0.5706	-0.1323	-0.4564	-0.0938	-0.3148
			3	-0.8008**	-2.1063	-0.9917**	-2.4468	-0.5904	-1.4764	0.0579	0.1549	-0.6949	-1.8163
			4	-0.1603	-0.4562	0.103	0.2882	-0.1202	-0.3083	-0.1176	-0.3087	0.1231	0.3414
20	m032757 (Algebra) (Polytomous)	DIF		-.476	-.824	-.692	-1.201	-.265	-.473	-.559	-1.083	.029	.055
		DDF											
21	m032760a (Algebra) (Polytomous)	DIF**		-.072	-.146	-.027	-.053	1.34**	2.384	-.236	-.463	.601	1.099
		DDF											

22	m032760b (Algebra)	DIF		.1023	.3076	.4267	1.293	.4043	1.1203	.011	.0335	.3679	1.0262
		DDF											
23	m032760c (Algebra)	DIF**		.48	1.169	.4813	1.2563	.3748	.9195	-.5263	-1.2884	1.3165**	2.8896
		DDF											
24	m032761 (Algebra) (Polytomous)	DIF**		.506	1.063	-.707	-1.437	-.781	-1.454	.689	1.354	-1.245**	-2.021
		DDF											
25	m032692 (Geometry) (Polytomous)	DIF		.026	.068	.722	1.91	-.129	-.339	.098	.259	.066	.163
		DDF											
26	m032626 (Number)	DIF**	4	0.7201**	2.7601	-0.2819	-1.0429	0.0688	0.2417	-0.0457	-0.1597	-0.3425	-1.2866
		DDF**	1	0.7592**	2.4132	-0.3666	-1.0943	0.0284	0.0818	-0.0115	-0.0335	-0.3892	-1.2091
			2	0.4561	0.9762	-0.0431	-0.0897	0.2214	0.4404	0.7038	1.1339	-0.5905	-1.1672
			3	0.8091	1.6881	-0.5488	-1.0433	0.2455	0.4708	-0.7671	-1.5181	-0.1315	-0.2745
27	m032595 (Number)	DIF	3	0.0645	0.2555	-0.1429	-0.5565	-0.1861	-0.6785	-0.1013	-0.3744	-0.2972	-1.1404
		DDF**	1	-0.1982	-0.6753	-0.1473	-0.5002	-0.3338	-1.0143	0.0231	0.0729	-0.5022	-1.7128
			2	0.9479**	2.1776	-0.1937	-0.4542	-0.5656	-1.1793	0.046	0.1035	-0.406	-0.9407
			4	-0.5531	-1.2066	-0.5882	-1.1919	1.2175**	2.5291	-0.1363	-0.2927	0.2806	0.5909

28	m032673 (Algebra)	DIF	3	-0.3933	-1.6442	-0.026	-0.107	-0.2566	-1.0195	-0.1581	-0.6171	-0.0644	-0.2494
		DDF	1	-0.3548	-1.0951	0.2018	0.611	-0.4885	-1.3502	0.1896	0.5298	-0.2668	-0.8017
			2	-0.396	-1.3594	-0.1468	-0.4892	0.0638	0.2075	-0.1774	-0.5637	-0.2229	-0.7211
			4	0.0281	0.0675	-0.1875	-0.4273	-0.6944	-1.4828	0.0913	0.2055	0.4414	0.912

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 13

ข้อ	ข้อสอบ	Option	Response	GIRL		HOMEWK		SES		EQUIP		RICH	
				LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)
1	m032166 (Number)	DIF	2	-0.1286	-0.5061	-0.398	-1.5397	0.1216	0.4382	0.0535	0.1966	-0.1773	-0.6861
		DDF	1	-0.0251	-0.0777	-0.364	-1.1034	-0.4845	-1.2794	0.5902	1.5815	-0.5119	-1.5947
			3	0.0119	0.029	-0.6868	-1.5245	0.2579	0.5447	-0.5545	-1.2525	0.0427	0.1039
			4	-0.0772	-0.2038	-0.7169	-1.7567	0.4489	1.077	0.0979	0.2364	-0.0579	-0.1458
2	m032721 (Data and Chance)	DIF**	2	-0.1562	-0.7036	0.5099**	2.1643	0.1023	0.4191	-0.147	-0.6192	0.3706	1.5872
		DDF	1	-0.0201	-0.0769	0.328	1.1841	0.0729	0.2521	0.1166	0.4073	0.1811	0.6595
			3	-0.1922	-0.5978	1.0597	3.068	0.4391	1.2614	-0.2871	-0.8494	0.5087	1.5369
			4	0.022	0.0559	0.5518	1.316	-0.4104	-0.8422	-0.6235	-1.497	0.533	1.3118
3	m032757 (Algebra) (Polytomous)	DIF		.58	1.443	.253	.652	.289	.726	-.646	-1.564	.604	1.514
		DDF											
4	m032760a (Algebra) (Polytomous)	DIF		.119	.304	-.163	-.426	.179	.429	-.262	-.637	.239	.583
		DDF											
5	m032760b (Algebra)	DIF		.5803	1.4443	.2533	.6564	.2885	.7751	-.6459	-1.6091	.6038	1.5462
		DDF											

6	m032760c (Algebra)	DIF		.119	.3046	-.1629	-.4296	.1787	.4548	-.2622	-.6522	.2391	.5901
		DDF											
7	m032761 (Algebra) (Polytomous)	DIF		-.286	-.775	-.414	-1.14	-.325	-.881	.039	.106	-.108	-.282
		DDF											
8	m032692 (Geometry) (Polytomous)	DIF		-.272	-.649	.172	.413	.037	.085	.46	1.065	-.882	-1.747
		DDF											
9	m032626 (Number)	DIF**	4	0.5745**	2.3072	-0.3115	-1.1702	0.4605	1.5814	-0.149	-0.5486	-0.1552	-0.624
		DDF**	1	0.5675**	2.0225	-0.3978	-1.3081	0.227	0.6914	-0.0022	-0.0071	-0.2212	-0.7917
			2	0.8068	1.5356	-0.4013	-0.6571	0.2797	0.4212	-0.2661	-0.4663	-0.2236	-0.4129
			3	0.5189	1.0008	0.3027	0.5573	1.4921**	2.2846	-1.0997	-1.9498	0.1203	0.2293
10	m032595 (Number)	DIF**	3	0.064	0.2417	-0.2142	-0.7898	-0.6588**	-2.2439	0.3769	1.3248	-0.2834	-1.04
		DDF**	1	0.0233	0.0746	-0.1872	-0.5772	-0.8053**	-2.2923	0.4736	1.4062	-0.283	-0.905
			2	-0.1137	-0.2305	-0.4777	-0.912	-0.7757	-1.3353	-0.3247	-0.6107	-0.5457	-1.029
			4	-0.0556	-0.1361	0.1192	0.2911	0.0099	0.0216	0.6529	1.4123	-0.5967	-1.4382

11	m032673 (Algebra)	DIF**	3	-0.5102**	-2.0142	0.088	0.3442	0.2698	0.9477	-0.1167	-0.4487	-0.056	-0.2255
		DDF	1	-0.5255	-1.7179	-0.1058	-0.339	-0.015	-0.043	0.1312	0.418	-0.148	-0.5178
			2	-0.3709	-1.1995	0.243	0.7766	0.498	1.4544	-0.427	-1.3457	-0.079	-0.2482
			4	-0.7662	-1.4642	0.4935	0.9581	0.731	1.2338	0.203	0.3589	0.2365	0.4372
12	m042182 (Number)	DIF**	3	-0.5171	-1.958	0.2601	0.9745	-0.6073**	-2.1705	0.3218	1.1651	-0.3323	-1.2791
		DDF**	1	-0.2984	-0.828	0.5778	1.6014	-0.9023**	-2.1366	-0.0034	-0.009	-0.6639	-1.8534
			2	-0.5225	-1.5408	0.0857	0.2369	-0.3806	-1.0581	0.812	1.9472	0.1941	0.5582
			4	-0.5291	-1.1618	0.0094	0.0202	-0.671	-1.2629	-0.2085	-0.4768	-1.1281**	-2.3936
13	m042081 (Number)	DIF		-.2861	-.7674	-.4137	-1.1488	-.3249	-.8669	.039	.105	-.1078	-.2812
		DDF											
14	m042049 (Algebra)	DIF	3	-0.1702	-0.7267	-0.0124	-0.0508	0.4335	1.6032	-0.2493	-1.0184	-0.1085	-0.4557
		DDF**	1	-0.2824	-0.6623	-0.0005	-0.0011	0.9355**	2.0497	0.0398	0.0923	-0.2127	-0.5212
			2	-0.2141	-0.6752	-0.2454	-0.714	0.5642	1.4222	-0.0722	-0.2102	0.1122	0.3398
			4	-0.2146	-0.6563	0.194	0.5743	0.3525	0.9335	-0.1853	-0.5573	-0.3704	-1.1083

15	m042052 (Number)	DIF	1	0.4118	1.5435	0.0288	0.1055	-0.1649	-0.5457	0.0052	0.0176	-0.0509	-0.1914
		DDF**	2	1.0602**	2.144	0.5223	1.1479	-0.4238	-0.7789	-0.5053	-1.0466	-0.3269	-0.6815
			3	-0.0095	-0.0182	0.4196	0.7636	0.6085	0.9648	-0.6328	-0.9802	0.7285	1.1801
			4	0.242	0.8259	-0.3039	-0.9819	-0.1427	-0.4231	0.0911	0.2753	-0.074	-0.2561
16	m042076 (Algebra)	DIF	4	0.24	1.0105	0.2641	1.071	0.3214	1.1988	-0.4879	-1.858	0.364	1.4577
		DDF**	1	0.1338	0.4775	0.3537	1.2363	0.2429	0.7394	-0.5579	-1.7177	0.3141	1.0562
			2	-0.0533	-0.1519	0.0378	0.1041	0.6268	1.5969	-0.7638**	-2.0488	0.5752	1.6009
			3	0.4443	1.0962	0.2667	0.6182	0.6915	1.4421	-0.9203	-1.9894	0.6271	1.4419
17	m042302a (Number) (Polytomous)	DIF		-.018	-.052	.271	.863	.154	.427	-.253	-.735	.211	.588
		DDF											
18	m042302b (Number) (Polytomous)	DIF		-.187	-.491	.038	.106	-.454	-1.244	.658	1.732	.049	.124
		DDF											
19	m042302c (Number) (Polytomous)	DIF		.087	.209	-.185	-.45	.143	.356	.09	.216	-.521	-1.198
		DDF											

20	m042100 (Algebra)	DIF	3	0.0815	0.3331	0.2979	1.1573	-0.1788	-0.6642	-0.2983	-1.1274	0.2253	0.9162
		DDF	1	-0.0492	-0.1091	0.1065	0.2219	-0.8004	-1.3298	-0.411	-0.8462	0.0528	0.1213
			2	0.2153	0.7924	0.2731	0.9549	-0.1872	-0.6142	-0.2199	-0.7414	0.2374	0.8636
			4	-0.2811	-0.6521	0.3322	0.774	0.0999	0.217	-0.829	-1.8777	0.3947	0.9118
21	m042202 (Algebra)	DIF**	2	-0.2637	-1.0768	-0.5027**	-2.0377	0.1722	0.6459	0.1599	0.6303	0.1026	0.4051
		DDF**	1	0.3904	0.6902	-0.8737	-1.6234	0.261	0.463	-0.1054	-0.2056	0.2551	0.4765
			3	1.0017	1.6149	-1.3804**	-2.0925	0.5743	0.9486	-0.2366	-0.4449	0.4227	0.7556
			4	-0.4207	-1.6459	-0.444	-1.7243	0.1852	0.6659	0.2591	0.9661	0.0309	0.1171
22	m042240 (Algebra)	DIF	4	0.2462	0.9793	0.0953	0.3621	-0.0721	-0.2591	0.2073	0.7706	0.083	0.3256
		DDF	1	0.3418	1.0077	0.1328	0.3742	0.1265	0.338	0.1148	0.3109	0.4759	1.3316
			2	-0.0256	-0.0783	0.2368	0.6888	0.0139	0.0391	0.0807	0.2362	-0.0472	-0.1442
			3	0.674	1.6536	-0.103	-0.238	-0.5104	-1.0372	0.4639	1.0293	-0.0566	-0.1379
23	m042093 (Algebra)	DIF		-.272	-.6358	.1718	.4119	.037	.0837	.4595	1.0546	-.8824	-1.7501
		DDF											

24	m042271 (Geometry)	DIF	3	-0.0263	-0.1038	-0.029	-0.1101	-0.1939	-0.6891	0.1063	0.4007	0.1403	0.5128
		DDF	1	0.1527	0.414	-0.2578	-0.6849	0.2139	0.5366	-0.001	-0.0028	0.3852	0.9391
			2	-0.353	-1.0825	-0.2048	-0.5987	-0.2773	-0.7531	0.4363	1.1993	0.2864	0.8218
			4	-0.0551	-0.1842	0.2241	0.6964	-0.3424	-0.9991	0.1723	0.5408	-0.2192	-0.6974
25	m042268 (Geometry)	DIF	2	0.2145	0.8165	-0.2249	-0.8455	-0.1643	-0.5793	0.4213	1.5773	-0.31	-1.1127
		DDF**	1	0.3734	1.172	-0.2364	-0.714	-0.2771	-0.7938	0.642	1.9687	-0.5528	-1.7009
			3	-0.0712	-0.195	-0.3456	-0.9381	-0.2551	-0.5994	0.8902**	2.2446	-0.1569	-0.4181
			4	0.5525	1.6089	-0.3185	-0.9061	-0.0928	-0.2543	0.2883	0.8207	-0.329	-0.844
26	m042159 (Data and Chance)	DIF		-.0183	-.053	.2713	.8214	.1542	.4285	-.2534	-.7134	.2108	.5901
		DDF											
27	m042164 (Data and Chance)	DIF		-.1873	-.4948	.0376	.108	-.4545	-1.2123	.6581	1.7653	.0489	.1271
		DDF											
28	m042167 (Data and Chance)	DIF		.0866	.2118	-.185	-.4648	.1429	.3319	.0901	.219	-.5212	-1.1508
		DDF											

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 14

ข้อ	ข้อสอบ	Option	Response	GIRL		HOMEWK		SES		EQUIP		RICH	
				LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)	LOR	Z(LOR)
1	m022043 (Number)	DIF**	4	0.0442	0.1839	-0.6247**	-2.3645	0.2864	1.0816	-0.9162**	-3.264	0.3971	1.5244
		DDF**	1	0.4105	1.2697	-0.7128	-1.9663	0.3954	1.1356	-1.1609**	-3.0663	0.2265	0.6483
			2	-0.4665	-1.0179	-0.3739	-0.8069	-0.5442	-0.9814	-1.1705**	-2.3087	0.0514	0.1088
			3	0.1473	0.3397	-0.9745	-1.8225	0.3706	0.7671	-0.7823	-1.5699	0.3152	0.7045
			5	-1.5447**	-2.2847	-0.3073	-0.5736	-0.0609	-0.1019	-0.0702	-0.113	0.1034	0.202
2	m022046 (Number)	DIF		-.0809	-.2879	-.3496	-1.2314	.2418	.7886	.3584	1.1565	-.1708	-.5877
		DDF											
3	m022049 (Geometry)	DIF**	4	-0.6046**	-2.7296	0.1102	0.4762	-0.1666	-0.7077	0.1729	0.7127	0.29	1.2404
		DDF**	1	-0.41	-1.1943	0.139	0.3756	-0.2665	-0.6881	-0.2462	-0.6797	0.2765	0.7477
			2	-0.1491	-0.5085	0.2215	0.732	-0.3792	-1.1686	0.7695**	2.0697	0.0826	0.2643
			3	-0.9608**	-3.1815	0.0817	0.268	-0.0155	-0.0501	-0.0124	-0.0389	0.5258	1.7405
4	m022050 (Algebra)	DIF	5	0.0966	0.3772	-0.0998	-0.3824	-0.063	-0.2356	0.4065	1.4728	-0.2107	-0.7804
		DDF	1	0.2023	0.4246	-0.3132	-0.6577	-0.5128	-1.0029	-0.1027	-0.2207	-0.6739	-1.3167
			2	-0.096	-0.2882	-0.0512	-0.1489	0.2336	0.6801	0.5917	1.5579	0.2633	0.733
			3	0.0814	0.2647	-0.2147	-0.6745	-0.1954	-0.6149	0.4321	1.291	-0.4078	-1.2732
			4	0.8256	1.3866	-0.4917	-0.8306	0.1132	0.1797	0.1808	0.2976	0.8963	1.4565

5	m022055 (Geometry)	DIF		.2763	.7061	.3399	.8704	.2781	.7109	-.526	-1.2653	.0323	.0738
		DDF											
6	m022057 (Number)	DIF	3	-0.1989	-0.8907	-0.0661	-0.2834	0.1723	0.6911	0.1826	0.7097	0.3541	1.4656
		DDF**	1	-0.3852	-0.8185	-0.3131	-0.6375	0.2925	0.601	-0.0932	-0.1709	1.9324**	2.5034
			2	-0.0868	-0.2702	-0.2188	-0.6441	0.447	1.1945	-0.1229	-0.343	0.4639	1.3632
			4	-0.2656	-0.7888	0.4314	1.2121	-0.1128	-0.2924	0.6741	1.5179	-0.0091	-0.0253
			5	0.6764	1.1466	-2.4472**	-2.0427	0.8239	1.3155	-0.1171	-0.1926	-0.2639	-0.4398
7	m022257 (Data and Chance)	DIF	3	0.4182	1.8073	0.2635	1.0782	-0.0917	-0.3683	0.0468	0.1821	-0.0622	-0.2584
		DDF	1	0.0849	0.2408	0.4829	1.2543	-0.1261	-0.3292	0.3371	0.8178	-0.48	-1.3625
			2	0.4621	1.5151	0.1487	0.4758	0.1811	0.5466	-0.2688	-0.821	0.3449	1.0525
			4	0.3484	1.0113	0.2008	0.5365	-0.3713	-0.9012	0.2141	0.5063	-0.0254	-0.071
			5	1.005	1.9387	0.3689	0.6806	-0.7588	-1.2657	0.2442	0.4289	-0.3106	-0.6149
8	m022062 (Geometry)	DIF	2	-0.5044	-1.9084	-0.0791	-0.2931	0.4187	1.4468	-0.4904	-1.5912	0.1547	0.5396
		DDF**	1	-0.6729**	-2.4319	-0.024	-0.0857	0.4156	1.3914	-0.4939	-1.5372	0.1252	0.424
			3	0.1311	0.229	0.6662	1.1328	0.0203	0.0336	0.1986	0.2922	-0.4168	-0.7125
			4	-0.1133	-0.2222	-1.0069	-1.4331	0.348	0.624	-0.7654	-1.193	0.193	0.3295
			5	0.324	0.6194	-0.2812	-0.5107	0.2373	0.4018	-0.1693	-0.2881	-0.4711	-0.8752

9	m022066 (Number)	DIF	4	0.0971	0.3028	0.0412	0.1231	0.0881	0.2606	0.6355	1.8807	-0.382	-1.0611
		DDF**	1	0.3023	0.8402	0.0116	0.03	-0.1941	-0.5063	0.9394**	2.3794	-0.6659	-1.6948
			2	0.3766	0.6239	0.3396	0.5091	0.2616	0.331	0.0776	0.1125	0.3352	0.4951
			3	-0.3119	-0.6834	0.3425	0.7795	0.4043	0.9073	0.5033	1.0706	-0.2474	-0.4827
10	m022232 (Number) (Polytomous)	DIF		-.358	-.865	.038	.09	.487	1.143	-.06	-.134	-.121	-.276
		DDF											
11	m022234a (Geometry) (Polytomous)	DIF		-.208	-.413	.051	.1	-.125	-.27	.136	.281	-.308	-.557
		DDF											
12	m022234b (Number) (Polytomous)	DIF		.215	.335	.655	1.031	-.581	-1.037	1.253	1.673	-.721	-1.17
		DDF											
13	m022243 (Geometry)	DIF		-.2442	-.7411	-.1804	-.5602	-.5303	-1.5963	.2388	.7245	.1127	.3222
		DDF											
14	m042182 (Number)	DIF	3	-0.3328	-1.3227	0.256	0.9903	0.1185	0.4165	-0.0837	-0.2891	0.1112	0.422
		DDF**	1	-0.1489	-0.4438	0.1688	0.4793	-0.0988	-0.2461	0.1325	0.3407	0.1606	0.4496
			2	-0.2286	-0.643	-0.0239	-0.0671	0.4251	1.0458	0.2475	0.5712	0.0444	0.1213
			4	-1.0607**	-2.2153	0.5515	1.2086	0.5069	1.0594	-0.5498	-1.0917	-0.0812	-0.1769

15	m042081 (Number)	DIF**		-.1302	-.354	.0372	.1019	-.7706**	-2.0522	-.3021	-.7982	.3936	.9526
		DDF											
16	m042049 (Algebra)	DIF	3	-0.1104	-0.4619	0.0461	0.186	-0.149	-0.5599	0.0657	0.247	0.1593	0.6462
		DDF	1	-0.1929	-0.5022	-0.0418	-0.1027	-0.2609	-0.6067	-0.1472	-0.361	-0.2776	-0.724
			2	-0.0773	-0.2334	-0.1426	-0.4092	0.3885	0.9643	0.2209	0.5605	0.4648	1.3809
			4	-0.1979	-0.6089	0.1273	0.3841	0.1146	0.3308	0.0071	0.0199	0.37	1.0437
17	m042052 (Number)	DIF	1	-0.0937	-0.3543	-0.2586	-0.9302	-0.5675	-1.9211	0.5277	1.7399	-0.2011	-0.7521
		DDF	2	0.0565	0.1215	-0.0249	-0.0534	-0.2243	-0.4327	0.5035	0.9395	-0.0481	-0.0997
			3	-0.1781	-0.3791	-0.0061	-0.0122	-1.3582	-1.9733	1.2483	1.7785	-0.2796	-0.6149
			4	-0.1624	-0.5528	-0.4431	-1.4175	-0.5875	-1.7585	0.4218	1.2762	-0.2496	-0.8516
18	m042076 (Algebra)	DIF**	4	0.5668**	2.323	-0.048	-0.1935	0.3377	1.2438	-0.246	-0.9051	0.173	0.6739
		DDF**	1	0.6765**	2.2762	-0.1356	-0.4505	0.3585	1.098	-0.3813	-1.1586	0.378	1.197
			2	0.7169**	2.0993	-0.2278	-0.6257	0.17	0.4339	0.1134	0.2816	0.225	0.6469
			3	0.0304	0.0767	-0.2868	-0.7413	0.5518	1.2448	-0.5203	-1.2254	-0.1647	-0.4131
19	m042302a (Number) (Polytomous)	DIF**		.031	.095	.73**	2.281	.179	.509	-.131	-.389	-.514	-1.503
		DDF											

20	m042302b (Number) (Polytomous)	DIF		-.134	-.38	-.071	-.213	-.347	-.925	.145	.385	.276	.713
		DDF											
21	m042302c (Number) (Polytomous)	DIF		.316	1.039	-.683	-2.331	.068	.228	-.321	-1.059	.542	1.652
		DDF											
22	m042100 (Algebra)	DIF	3	-0.0642	-0.2651	-0.1714	-0.6884	0.2995	1.1285	-0.3025	-1.0697	0.3415	1.3345
		DDF**	1	0.5108	1.1278	-0.9222	-1.7046	0.3472	0.7293	-1.0224**	-2.0383	0.7818	1.581
			2	-0.191	-0.692	0.0296	0.106	0.2717	0.8926	-0.1476	-0.4492	0.2231	0.7937
			4	0.0088	0.021	-0.4439	-0.976	0.3525	0.763	-0.149	-0.3037	0.3635	0.803
23	m042202 (Algebra)	DIF**	2	0.5613**	2.291	0.2927	1.1506	-0.4309	-1.6453	0.0425	0.1579	-0.318	-1.2432
		DDF**	1	1.0991**	2.2583	0.1853	0.3642	-0.4539	-0.7321	0.6339	1.0319	0.0404	0.0746
			3	0.2678	0.5846	0.5185	1.0865	-0.5943	-1.027	-0.0755	-0.1473	-0.3175	-0.6438
			4	0.5568**	2.1665	0.2864	1.0679	-0.3242	-1.191	-0.0503	-0.1774	-0.2683	-1.0209
24	m042240 (Algebra)	DIF**	4	0.4715	1.7246	-0.6354**	-2.1918	0.4006	1.2693	-0.0713	-0.221	-0.3424	-1.2255
		DDF**	1	0.7066	1.9149	-0.737	-1.9019	0.6604	1.5845	-0.3701	-0.9016	0.0219	0.0571
			2	0.3905	1.0941	-0.7241	-1.8235	-0.0776	-0.1809	0.6243	1.3224	-0.4076	-1.1031
			3	0.2391	0.5886	-0.7099	-1.5953	0.5002	1.0528	-0.37	-0.7505	-0.8329**	-2.0128

25	m042093 (Algebra)	DIF		-.1387	-.297	-.0105	-.023	.494	1.0125	-.0188	-.0427	-.025	-.0441
		DDF											
26	m042271 (Geometry)	DIF	3	0.0094	0.0373	0.1242	0.46	-0.2238	-0.8222	0.357	1.2972	-0.3368	-1.2488
		DDF	1	0.4821	1.2171	-0.0577	-0.1417	-0.2674	-0.6136	0.4755	1.1035	-0.7555	-1.9064
			2	-0.1709	-0.5284	0.0923	0.2754	-0.028	-0.0804	0.1984	0.5686	-0.0128	-0.0395
			4	-0.0608	-0.2095	0.2666	0.8389	-0.3885	-1.2006	0.3909	1.1845	-0.4201	-1.3006
27	m042268 (Geometry)	DIF**	2	-0.2148	-0.818	0.7871**	2.6369	-0.3455	-1.237	-0.2238	-0.7406	-0.6444**	-2.2406
		DDF**	1	-0.6984**	-2.0632	0.8203**	2.1384	-0.2106	-0.6003	0.1812	0.4572	-1.0158**	-3.0514
			3	0.1514	0.4007	0.6764	1.7255	-0.1515	-0.3651	-0.721	-1.6334	-0.5111	-1.2131
			4	-0.0905	-0.2552	1.0367**	2.6081	-0.3319	-0.882	-0.2413	-0.6178	-0.3024	-0.7925
28	m042159 (Data and Chance)	DIF		.331	1.0172	.1752	.5239	.5676	1.5032	-.6105	-1.6129	.0704	.2113
		DDF											
29	m042164 (Data and Chance)	DIF		.0069	.0202	.3469	1.0017	-.2286	-.6498	.3128	.894	-.2993	-.7746
		DDF											
30	m042167 (Data and Chance)	DIF		-.0277	-.0613	-.1746	-.3962	.3126	.7085	.2065	.4993	.0065	.0133
		DDF											

ภาคผนวก ข

ผลการวิเคราะห์ของโมเดลการวิเคราะห์มูลค่าเพิ่มด้วยโปรแกรม
Hierarchical linear model (HLM) (แสดงเฉพาะส่วนที่สำคัญ)

โมเดล 1

Program: HLM 6 Hierarchical Linear and Nonlinear Modeling

Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon

Publisher: Scientific Software International, Inc. (c) 2000

techsupport@ssicentral.com

www.ssicentral.com

Module: HLM2.EXE (6.03.26284.1)

Date: 10 April 2012, Tuesday

Time: 15:58: 5

SPECIFICATIONS FOR THIS HLM2 RUN

Problem Title: no title

The data source for this run = total_addfile.mdm

The command file for this run = whlmtemp.hlm

Output file name = E:\All Work\Phd\บ.19\บ19\Thesis\ข้อมูลวิเคราะห์\Multilog_หาค่าพารามิเตอร์\เตรียมวิเคราะห์มูลค่าเพิ่ม\hlm2.txt

The maximum number of level-1 units = 5412

The maximum number of level-2 units = 150

The maximum number of iterations = 100

Method of estimation: restricted maximum likelihood

Weighting Specification

	Weight	Variable	Normalized?
Level 1	no		
Level 2	no		
Precision	no		

The outcome variable is THETA1

The model specified for the fixed effects was:

Level-1 Coefficients	Level-2 Predictors

---------- ----------

INTRCPT1, B0 INTRCPT2, G00

DSCHSL, G01

DSCHL, G02

DSCHM, G03

COMMXXL, G04

COMMXL, G05

COMML, G06

COMMM, G07

COMMS, G08

SCQ21AC, G09

$ SCQ20AE, G010

SCMAEDU, G011

$ MATPROM, G012

PPSC, G013

$ DVTESCH, G014

$ EXPTEA, G015

ADEQ, G016

HOMEWORK, G017

CLASSSIZ, G018

TEACOLLA, G019

#% RELAGR1 slope, B1 INTRCPT2, G10

SEX slope, B2 INTRCPT2, G20

#% PAEDU1 slope, B3 INTRCPT2, G30

#% LEARES slope, B4 INTRCPT2, G40

#% WEALTH slope, B5 INTRCPT2, G50

#% PRASCH slope, B6 INTRCPT2, G60

TIMEHOM slope, B7 INTRCPT2, G70

#% INTMAT slope, B8 INTRCPT2, G80

SELFLEA slope, B9 INTRCPT2, G90

JOYMATH slope, B10 INTRCPT2, G100

#% INSTMAT slope, B11 INTRCPT2, G110

#% VALMATH slope, B12 INTRCPT2, G120

SAFESCH slope, B13 INTRCPT2, G130

'#' - The residual parameter variance for this level-1 coefficient has been set

to zero.

'%' - This level-1 predictor has been centered around its grand mean.

'$' - This level-2 predictor has been centered around its grand mean.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions

INTRCPT1

Summary of the model specified (in equation format)

Level-1 Model

Y = B0 + B1*(RELAGR1) + B2*(SEX) + B3*(PAEDU1) + B4*(LEARES) + B5*(WEALTH) + B6*(PRASCH) + B7*(TIMEHOM) + B8*(INTMAT) + B9*(SELFLEA) + B10*(JOYMATH) + B11*(INSTMAT) + B12*(VALMATH) + B13*(SAFESCH) + R

Level-2 Model

B0 = G00 + G01*(DSCHSL) + G02*(DSCHL) + G03*(DSCHM) + G04*(COMMXXL) + G05*(COMMXL) + G06*(COMML) + G07*(COMMM) + G08*(COMMS) + G09*(SCQ21AC) + G010*(SCQ20AE) + G011*(SCMAEDU) + G012*(MATPROM) + G013*(PPSC) + G014*(DVTESCH) + G015*(EXPTEA) + G016*(ADEQ) + G017*(HOMEWORK) + G018*(CLASSSIZ) + G019*(TEACOLLA) + U0

B1 = G10

B2 = G20

B3 = G30

B4 = G40

B5 = G50

B6 = G60

B7 = G70

B8 = G80

B9 = G90

B10 = G100

B11 = G110

B12 = G120

B13 = G130

Iterations stopped due to small change in likelihood function

******* ITERATION 7 *******

Sigma_squared = 0.33539

Tau

INTRCPT1,B0 0.17316

Tau (as correlations)

INTRCPT1,B0 1.000

Random level-1 coefficient Reliability estimate

INTRCPT1, B0 0.945

The value of the likelihood function at iteration 7 = -5.005242E+003

The outcome variable is THETA1

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	-1.900923	0.349671	-5.436	130	0.000
DSCHSL, G01	0.746488	0.170003	4.391	130	0.000
DSCHL, G02	0.284704	0.166280	1.712	130	0.089
DSCHM, G03	0.035227	0.109999	0.320	130	0.749
COMMXXL, G04	0.177807	0.192259	0.925	130	0.357
COMMXL, G05	-0.020505	0.180986	-0.113	130	0.910
COMML, G06	0.322065	0.166400	1.935	130	0.055
COMMM, G07	0.046254	0.147110	0.314	130	0.754
COMMS, G08	0.093346	0.143549	0.650	130	0.516
SCQ21AC, G09	0.003038	0.029116	0.104	130	0.918
SCQ20AE, G010	-0.052732	0.189591	-0.278	130	0.781
SCMAEDU, G011	0.002421	0.070666	0.034	130	0.973
MATPROM, G012	-0.034328	0.114410	-0.300	130	0.765
PPSC, G013	0.033867	0.069398	0.488	130	0.626
DVTESCH, G014	0.030584	0.060901	0.502	130	0.616
EXPTEA, G015	-0.005452	0.003626	-1.504	130	0.135
ADEQ, G016	-0.026689	0.065726	-0.406	130	0.685

HOMEWORK, G017	0.105993	0.058965	1.798	130	0.074
CLASSSIZ, G018	-0.005824	0.085330	-0.068	130	0.946
TEACOLLA, G019	0.030774	0.065431	0.470	130	0.638

For RELAGR1 slope, B1

INTRCPT2, G10	0.016077	0.005400	2.977	5379	0.003

For SEX slope, B2

INTRCPT2, G20	0.042964	0.017438	2.464	5379	0.014

For PAEDU1 slope, B3

INTRCPT2, G30	0.006838	0.005719	1.196	5379	0.232

For LEARES slope, B4

INTRCPT2, G40	0.044089	0.009783	4.507	5379	0.000

For WEALTH slope, B5

INTRCPT2, G50	0.072858	0.057453	1.268	5379	0.205

For PRASCH slope, B6

INTRCPT2, G60	0.075259	0.022874	3.290	5379	0.001

For TIMEHOM slope, B7

INTRCPT2, G70	0.034006	0.012508	2.719	5379	0.007

For INTMAT slope, B8

INTRCPT2, G80	-0.040604	0.012528	-3.241	5379	0.002

For SELFLEA slope, B9

INTRCPT2, G90	0.174597	0.014482	12.056	5379	0.000

For JOYMATH slope, B10

INTRCPT2, G100	0.072164	0.013822	5.221	5379	0.000

For INSTMAT slope, B11

INTRCPT2, G110	0.010592	0.017337	0.611	5379	0.541

For VALMATH slope, B12

INTRCPT2, G120	0.031805	0.018740	1.697	5379	0.089

For SAFESCH slope, B13

INTRCPT2, G130	0.008864	0.011541	0.768	5379	0.443

The outcome variable is THETA1

Final estimation of fixed effects

(with robust standard errors)

Fixed Effect	Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	-1.900923	0.263888	-7.204	130	0.000
DSCHSL, G01	0.746488	0.172973	4.316	130	0.000
DSCHL, G02	0.284704	0.161712	1.761	130	0.080
DSCHM, G03	0.035227	0.087822	0.401	130	0.689
COMMXXL, G04	0.177807	0.168598	1.055	130	0.294
COMMXL, G05	-0.020505	0.139128	-0.147	130	0.883
COMML, G06	0.322065	0.122078	2.638	130	0.010
COMMM, G07	0.046254	0.102132	0.453	130	0.651
COMMS, G08	0.093346	0.087173	1.071	130	0.287
SCQ21AC, G09	0.003038	0.027496	0.110	130	0.913
SCQ20AE, G010	-0.052732	0.167081	-0.316	130	0.753
SCMAEDU, G011	0.002421	0.070471	0.034	130	0.973
MATPROM, G012	-0.034328	0.111103	-0.309	130	0.758
PPSC, G013	0.033867	0.066070	0.513	130	0.609
DVTESCH, G014	0.030584	0.056116	0.545	130	0.586
EXPTEA, G015	-0.005452	0.003839	-1.420	130	0.158
ADEQ, G016	-0.026689	0.058964	-0.453	130	0.651
HOMEWORK, G017	0.105993	0.053010	1.999	130	0.047
CLASSSIZ, G018	-0.005824	0.076524	-0.076	130	0.940
TEACOLLA, G019	0.030774	0.061768	0.498	130	0.619
For RELAGR1 slope, B1					
INTRCPT2, G10	0.016077	0.005634	2.853	5379	0.005
For SEX slope, B2					
INTRCPT2, G20	0.042964	0.022593	1.902	5379	0.057
For PAEDU1 slope, B3					
INTRCPT2, G30	0.006838	0.006281	1.089	5379	0.277
For LEARES slope, B4					
INTRCPT2, G40	0.044089	0.010549	4.179	5379	0.000
For WEALTH slope, B5					
INTRCPT2, G50	0.072858	0.059509	1.224	5379	0.221

For PRASCH slope, B6

INTRCPT2, G60 0.075259 0.023517 3.200 5379 0.002

For TIMEHOM slope, B7

INTRCPT2, G70 0.034006 0.014572 2.334 5379 0.020

For INTMAT slope, B8

INTRCPT2, G80 -0.040604 0.013284 -3.057 5379 0.003

For SELFLEA slope, B9

INTRCPT2, G90 0.174597 0.017010 10.264 5379 0.000

For JOYMATH slope, B10

INTRCPT2, G100 0.072164 0.015225 4.740 5379 0.000

For INSTMAT slope, B11

INTRCPT2, G110 0.010592 0.017399 0.609 5379 0.542

For VALMATH slope, B12

INTRCPT2, G120 0.031805 0.020454 1.555 5379 0.120

For SAFESCH slope, B13

INTRCPT2, G130 0.008864 0.011215 0.790 5379 0.429

Final estimation of variance components:

Random Effect Standard Variance df Chi-square P-value

Deviation Component

INTRCPT1, U0 0.41613 0.17316 130 2791.16706 0.000

level-1, R 0.57913 0.33539

Statistics for current covariance components model

Deviance = 10010.483693

Number of estimated parameters = 2

โมเดล 2

Program: HLM 6 Hierarchical Linear and Nonlinear Modeling

Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon

Publisher: Scientific Software International, Inc. (c) 2000

techsupport@ssicentral.com

www.ssicentral.com

Module: HLM2.EXE (6.03.26284.1)

Date: 10 April 2012, Tuesday

Time: 15:58:39

SPECIFICATIONS FOR THIS HLM2 RUN

Problem Title: no title

The data source for this run = total_addfile.mdm

The command file for this run = whlmtemp.hlm

Output file name = E:\All Work\Phd\บ.19\บ19\Thesis\ข้อมูลวิเคราะห์\Multilog_หาค่าพารามิเตอร์\เตรียมวิเคราะห์มูลค่าเพิ่ม\hlm2.txt

The maximum number of level-1 units = 5412

The maximum number of level-2 units = 150

The maximum number of iterations = 100

Method of estimation: restricted maximum likelihood

Weighting Specification

	Weight Variable		
	Weighting?	Name	Normalized?
Level 1	no		
Level 2	no		
Precision	no		

The outcome variable is THETA2

The model specified for the fixed effects was:

Level-1 Coefficients	Level-2 Predictors
INTRCPT1, B0	INTRCPT2, G00
	DSCHSL, G01
	DSCHL, G02
	DSCHM, G03

COMMXXL, G04

COMMXL, G05

COMML, G06

COMMM, G07

COMMS, G08

SCQ21AC, G09

$ SCQ20AE, G010

SCMAEDU, G011

$ MATPROM, G012

PPSC, G013

$ DVTESCH, G014

$ EXPTEA, G015

ADEQ, G016

HOMEWORK, G017

CLASSSIZ, G018

TEACOLLA, G019

#% RELAGR1 slope, B1 INTRCPT2, G10

SEX slope, B2 INTRCPT2, G20

#% PAEDU1 slope, B3 INTRCPT2, G30

#% LEARES slope, B4 INTRCPT2, G40

#% WEALTH slope, B5 INTRCPT2, G50

#% PRASCH slope, B6 INTRCPT2, G60

TIMEHOM slope, B7 INTRCPT2, G70

#% INTMAT slope, B8 INTRCPT2, G80

SELFLEA slope, B9 INTRCPT2, G90

JOYMATH slope, B10 INTRCPT2, G100

#% INSTMAT slope, B11 INTRCPT2, G110

#% VALMATH slope, B12 INTRCPT2, G120

SAFESCH slope, B13 INTRCPT2, G130

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

'%' - This level-1 predictor has been centered around its grand mean.

'$' - This level-2 predictor has been centered around its grand mean.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions

INTRCPT1

Summary of the model specified (in equation format)

Level-1 Model

Y = B0 + B1*(RELAGR1) + B2*(SEX) + B3*(PAEDU1) + B4*(LEARES) + B5*(WEALTH) + B6*(PRASCH) + B7*(TIMEHOM) + B8*(INTMAT) + B9*(SELFLEA) + B10*(JOYMATH) + B11*(INSTMAT) + B12*(VALMATH) + B13*(SAFESCH) + R

Level-2 Model

B0 = G00 + G01*(DSCHSL) + G02*(DSCHL) + G03*(DSCHM) + G04*(COMMXXL) + G05*(COMMXL) + G06*(COMML) + G07*(COMMM) + G08*(COMMS) + G09*(SCQ21AC) + G010*(SCQ20AE) + G011*(SCMAEDU) + G012*(MATPROM) + G013*(PPSC) + G014*(DVTESCH) + G015*(EXPTEA) + G016*(ADEQ) + G017*(HOMEWORK) + G018*(CLASSSIZ) + G019*(TEACOLLA) + U0

B1 = G10

B2 = G20

B3 = G30

B4 = G40

B5 = G50

B6 = G60

B7 = G70

B8 = G80

B9 = G90

B10 = G100

B11 = G110

B12 = G120

B13 = G130

Iterations stopped due to small change in likelihood function

******* ITERATION 7 *******

Sigma_squared = 0.34614

Tau

INTRCPT1,B0 0.16014

Tau (as correlations)

INTRCPT1,B0 1.000

Random level-1 coefficient Reliability estimate

INTRCPT1, B0 0.939

The value of the likelihood function at iteration 7 = -5.083349E+003

The outcome variable is THETA2

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	-1.784011	0.337814	-5.281	130	0.000
DSCHSL, G01	0.715612	0.164040	4.362	130	0.000
DSCHL, G02	0.281919	0.160413	1.757	130	0.081
DSCHM, G03	0.043076	0.106167	0.406	130	0.685
COMMXXL, G04	0.165491	0.185522	0.892	130	0.374
COMMXL, G05	-0.025896	0.174665	-0.148	130	0.883
COMML, G06	0.309145	0.160639	1.924	130	0.056
COMMM, G07	0.028721	0.142040	0.202	130	0.840
COMMS, G08	0.077362	0.138593	0.558	130	0.577
SCQ21AC, G09	0.002975	0.028092	0.106	130	0.916
SCQ20AE, G010	-0.047561	0.182920	-0.260	130	0.795
SCMAEDU, G011	0.007441	0.068164	0.109	130	0.914
MATPROM, G012	-0.034449	0.110366	-0.312	130	0.755
PPSC, G013	0.033164	0.066943	0.495	130	0.621

DVTESCH, G014 0.029958 0.058748 0.510 130 0.610

EXPTEA, G015 -0.005342 0.003498 -1.527 130 0.129

ADEQ, G016 -0.028297 0.063409 -0.446 130 0.656

HOMEWORK, G017 0.106145 0.056896 1.866 130 0.064

CLASSSIZ, G018 -0.008604 0.082356 -0.104 130 0.917

TEACOLLA, G019 0.026760 0.063140 0.424 130 0.672

For RELAGR1 slope, B1

INTRCPT2, G10 0.016211 0.005484 2.956 5379 0.004

For SEX slope, B2

INTRCPT2, G20 0.037617 0.017711 2.124 5379 0.033

For PAEDU1 slope, B3

INTRCPT2, G30 0.007059 0.005808 1.215 5379 0.225

For LEARES slope, B4

INTRCPT2, G40 0.048147 0.009935 4.846 5379 0.000

For WEALTH slope, B5

INTRCPT2, G50 0.071467 0.058337 1.225 5379 0.221

For PRASCH slope, B6

INTRCPT2, G60 0.068224 0.023233 2.937 5379 0.004

For TIMEHOM slope, B7

INTRCPT2, G70 0.028806 0.012703 2.268 5379 0.023

For INTMAT slope, B8

INTRCPT2, G80 -0.037190 0.012725 -2.922 5379 0.004

For SELFLEA slope, B9

INTRCPT2, G90 0.170069 0.014711 11.561 5379 0.000

For JOYMATH slope, B10

INTRCPT2, G100 0.060920 0.014038 4.340 5379 0.000

For INSTMAT slope, B11

INTRCPT2, G110 0.026722 0.017610 1.517 5379 0.129

For VALMATH slope, B12

INTRCPT2, G120 0.040607 0.019036 2.133 5379 0.033

For SAFESCH slope, B13

INTRCPT2, G130 0.005854 0.011722 0.499 5379 0.617

The outcome variable is THETA2

Final estimation of fixed effects

(with robust standard errors)

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	-1.784011	0.256060	-6.967	130	0.000
DSCHSL, G01	0.715612	0.166554	4.297	130	0.000
DSCHL, G02	0.281919	0.156055	1.807	130	0.073
DSCHM, G03	0.043076	0.084685	0.509	130	0.611
COMMXXL, G04	0.165491	0.162001	1.022	130	0.309
COMMXL, G05	-0.025896	0.133843	-0.193	130	0.847
COMML, G06	0.309145	0.116675	2.650	130	0.009
COMMM, G07	0.028721	0.098562	0.291	130	0.771
COMMS, G08	0.077362	0.084845	0.912	130	0.364
SCQ21AC, G09	0.002975	0.026635	0.112	130	0.912
SCQ20AE, G010	-0.047561	0.161356	-0.295	130	0.769
SCMAEDU, G011	0.007441	0.067992	0.109	130	0.913
MATPROM, G012	-0.034449	0.108103	-0.319	130	0.750
PPSC, G013	0.033164	0.064209	0.516	130	0.606
DVTESCH, G014	0.029958	0.053923	0.556	130	0.579
EXPTEA, G015	-0.005342	0.003663	-1.458	130	0.147
ADEQ, G016	-0.028297	0.057554	-0.492	130	0.623
HOMEWORK, G017	0.106145	0.051357	2.067	130	0.040
CLASSSIZ, G018	-0.008604	0.073867	-0.116	130	0.908
TEACOLLA, G019	0.026760	0.060325	0.444	130	0.658
For RELAGR1 slope, B1					
INTRCPT2, G10	0.016211	0.005735	2.827	5379	0.005
For SEX slope, B2					
INTRCPT2, G20	0.037617	0.022697	1.657	5379	0.097
For PAEDU1 slope, B3					
INTRCPT2, G30	0.007059	0.005972	1.182	5379	0.238
For LEARES slope, B4					

INTRCPT2, G40 0.048147 0.011106 4.335 5379 0.000

For WEALTH slope, B5

INTRCPT2, G50 0.071467 0.059517 1.201 5379 0.230

For PRASCH slope, B6

INTRCPT2, G60 0.068224 0.024509 2.784 5379 0.006

For TIMEHOM slope, B7

INTRCPT2, G70 0.028806 0.014103 2.043 5379 0.041

For INTMAT slope, B8

INTRCPT2, G80 -0.037190 0.013119 -2.835 5379 0.005

For SELFLEA slope, B9

INTRCPT2, G90 0.170069 0.017285 9.839 5379 0.000

For JOYMATH slope, B10

INTRCPT2, G100 0.060920 0.015752 3.868 5379 0.000

For INSTMAT slope, B11

INTRCPT2, G110 0.026722 0.017625 1.516 5379 0.129

For VALMATH slope, B12

INTRCPT2, G120 0.040607 0.019890 2.042 5379 0.041

For SAFESCH slope, B13

INTRCPT2, G130 0.005854 0.010784 0.543 5379 0.587

Final estimation of variance components:

Random Effect Standard Variance df Chi-square P-value

Deviation Component

INTRCPT1, U0 0.40018 0.16014 130 2519.05496 0.000

level-1, R 0.58834 0.34614

Statistics for current covariance components model

Deviance = 10166.697160

Number of estimated parameters = 2

โมเดล 3

Program: HLM 6 Hierarchical Linear and Nonlinear Modeling

Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon

Publisher: Scientific Software International, Inc. (c) 2000

techsupport@ssicentral.com

www.ssicentral.com

Module: HLM2.EXE (6.03.26284.1)

Date: 10 April 2012, Tuesday

Time: 15:55: 9

SPECIFICATIONS FOR THIS HLM2 RUN

Problem Title: no title

The data source for this run = total_addfile.mdm

The command file for this run = whlmtemp.hlm

Output file name = E:\All Work\Phd\บ.19\บ19\Thesis\ข้อมูลวิเคราะห์\Multilog_หาค่าพารามิเตอร์\เตรียมวิเคราะห์มูลค่าเพิ่ม\hlm2.txt

The maximum number of level-1 units = 5412

The maximum number of level-2 units = 150

The maximum number of iterations = 100

Method of estimation: restricted maximum likelihood

Weighting Specification

	Weight Variable		
	Weighting?	Name	Normalized?
Level 1	no		
Level 2	no		
Precision	no		

The outcome variable is THETA34

The model specified for the fixed effects was:

Level-1	Level-2

```
  Coefficients            Predictors
---------------------   ---------------
      INTRCPT1, B0      INTRCPT2, G00
The model specified for the covariance components was:
-----------------------------------------------------------
      Sigma squared (constant across level-2 units)
      Tau dimensions
           INTRCPT1
Summary of the model specified (in equation format)
-----------------------------------------------------
Level-1 Model
           Y = B0 + R
Level-2 Model
           B0 = G00 + U0
Iterations stopped due to small change in likelihood function
******* ITERATION 3 *******
 Sigma_squared =      0.36862
 Tau
 INTRCPT1,B0      0.29852
Tau (as correlations)
 INTRCPT1,B0  1.000
 ----------------------------------------------------
  Random level-1 coefficient   Reliability estimate
 ----------------------------------------------------
  INTRCPT1, B0                      0.964
 ----------------------------------------------------
The value of the likelihood function at iteration 3 = -5.232274E+003
 The outcome variable is  THETA34
 Final estimation of fixed effects:
 ---------------------------------------------------------------------------
                                Standard            Approx.
   Fixed Effect         Coefficient   Error      T-ratio   d.f.     P-value
 ---------------------------------------------------------------------------
 For       INTRCPT1, B0
   INTRCPT2, G00          -0.694172   0.045434   -15.279       149     0.000
```

The outcome variable is THETA34

Final estimation of fixed effects

(with robust standard errors)

Fixed Effect	Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	-0.694172	0.045281	-15.330	149	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	0.54637	0.29852	149	5066.81360	0.000
level-1, R	0.60714	0.36862			

Statistics for current covariance components model

Deviance = 10464.548637

Number of estimated parameters = 2

โมเดล 4

Program: HLM 6 Hierarchical Linear and Nonlinear Modeling

Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon

Publisher: Scientific Software International, Inc. (c) 2000

techsupport@ssicentral.com

www.ssicentral.com

Module: HLM2.EXE (6.03.26284.1)

Date: 10 April 2012, Tuesday

Time: 15:57: 3

SPECIFICATIONS FOR THIS HLM2 RUN

Problem Title: no title

The data source for this run = total_addfile.mdm

The command file for this run = whlmtemp.hlm

Output file name = E:\All Work\Phd\บ.19\บ19\Thesis\ข้อมูลวิเคราะห์\Multilog_หาค่าพารามิเตอร์\เตรียมวิเคราะห์มูลค่าเพิ่ม\hlm2.txt

The maximum number of level-1 units = 5412

The maximum number of level-2 units = 150

The maximum number of iterations = 100

Method of estimation: restricted maximum likelihood

Weighting Specification

	Weight	Variable	Normalized?
Level 1	no		
Level 2	no		
Precision	no		

The outcome variable is THETA34

The model specified for the fixed effects was:

Level-1 Coefficients	Level-2 Predictors
INTRCPT1, B0	INTRCPT2, G00
	DSCHSL, G01
	DSCHL, G02
	DSCHM, G03
	COMMXXL, G04
	COMMXL, G05
	COMML, G06
	COMMM, G07
	COMMS, G08
	SCQ21AC, G09
$	SCQ20AE, G010

SCMAEDU, G011
$ MATPROM, G012
PPSC, G013
$ DVTESCH, G014
$ EXPTEA, G015
ADEQ, G016
HOMEWORK, G017
CLASSSIZ, G018
TEACOLLA, G019
#% RELAGR1 slope, B1 INTRCPT2, G10
SEX slope, B2 INTRCPT2, G20
#% PAEDU1 slope, B3 INTRCPT2, G30
#% LEARES slope, B4 INTRCPT2, G40
#% WEALTH slope, B5 INTRCPT2, G50
#% PRASCH slope, B6 INTRCPT2, G60
TIMEHOM slope, B7 INTRCPT2, G70
#% INTMAT slope, B8 INTRCPT2, G80
SELFLEA slope, B9 INTRCPT2, G90
JOYMATH slope, B10 INTRCPT2, G100
#% INSTMAT slope, B11 INTRCPT2, G110
#% VALMATH slope, B12 INTRCPT2, G120
SAFESCH slope, B13 INTRCPT2, G130

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

'%' - This level-1 predictor has been centered around its grand mean.

'$' - This level-2 predictor has been centered around its grand mean.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions

INTRCPT1

Summary of the model specified (in equation format)

Level-1 Model

Y = B0 + B1*(RELAGR1) + B2*(SEX) + B3*(PAEDU1) + B4*(LEARES) + B5*(WEALTH) + B6*(PRASCH) + B7*(TIMEHOM) + B8*(INTMAT) + B9*(SELFLEA) + B10*(JOYMATH) + B11*(INSTMAT) + B12*(VALMATH) + B13*(SAFESCH) + R

Level-2 Model

B0 = G00 + G01*(DSCHSL) + G02*(DSCHL) + G03*(DSCHM) + G04*(COMMXXL) + G05*(COMMXL) + G06*(COMML) + G07*(COMMM) + G08*(COMMS) + G09*(SCQ21AC) + G010*(SCQ20AE) + G011*(SCMAEDU) + G012*(MATPROM) + G013*(PPSC) + G014*(DVTESCH) + G015*(EXPTEA) + G016*(ADEQ) + G017*(HOMEWORK) + G018*(CLASSSIZ) + G019*(TEACOLLA) + U0

B1 = G10

B2 = G20

B3 = G30

B4 = G40

B5 = G50

B6 = G60

B7 = G70

B8 = G80

B9 = G90

B10 = G100

B11 = G110

B12 = G120

B13 = G130

Iterations stopped due to small change in likelihood function

******* ITERATION 7 *******

Sigma_squared = 0.34501

Tau

INTRCPT1,B0 0.16232

Tau (as correlations)

INTRCPT1,B0 1.000

Random level-1 coefficient Reliability estimate

INTRCPT1, B0 0.940

The value of the likelihood function at iteration 7 = -5.075619E+003

The outcome variable is THETA34

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	-1.936307	0.339857	-5.697	130	0.000
DSCHSL, G01	0.765686	0.165064	4.639	130	0.000
DSCHL, G02	0.309957	0.161420	1.920	130	0.057
DSCHM, G03	0.047342	0.106825	0.443	130	0.658
COMMXXL, G04	0.160355	0.186679	0.859	130	0.392
COMMXL, G05	-0.037516	0.175751	-0.213	130	0.832
COMML, G06	0.311936	0.161630	1.930	130	0.055
COMMM, G07	0.037242	0.142912	0.261	130	0.795
COMMS, G08	0.101627	0.139446	0.729	130	0.467
SCQ21AC, G09	0.000395	0.028268	0.014	130	0.989
SCQ20AE, G010	-0.069615	0.184065	-0.378	130	0.706
SCMAEDU, G011	0.017139	0.068594	0.250	130	0.803
MATPROM, G012	-0.041858	0.111060	-0.377	130	0.707
PPSC, G013	0.028842	0.067365	0.428	130	0.669
DVTESCH, G014	0.033932	0.059117	0.574	130	0.567
EXPTEA, G015	-0.004737	0.003520	-1.346	130	0.181
ADEQ, G016	-0.026939	0.063806	-0.422	130	0.673
HOMEWORK, G017	0.107073	0.057251	1.870	130	0.063
CLASSSIZ, G018	-0.027418	0.082867	-0.331	130	0.741
TEACOLLA, G019	0.042742	0.063534	0.673	130	0.502
For RELAGR1 slope, B1					
INTRCPT2, G10	0.017101	0.005475	3.123	5379	0.002
For SEX slope, B2					
INTRCPT2, G20	0.049061	0.017683	2.774	5379	0.006
For PAEDU1 slope, B3					
INTRCPT2, G30	0.007753	0.005799	1.337	5379	0.181

For LEARES slope, B4

INTRCPT2, G40	0.047710	0.009919	4.810	5379	0.000

For WEALTH slope, B5

INTRCPT2, G50	0.068723	0.058247	1.180	5379	0.239

For PRASCH slope, B6

INTRCPT2, G60	0.069308	0.023196	2.988	5379	0.003

For TIMEHOM slope, B7

INTRCPT2, G70	0.030171	0.012683	2.379	5379	0.018

For INTMAT slope, B8

INTRCPT2, G80	-0.044132	0.012705	-3.474	5379	0.001

For SELFLEA slope, B9

INTRCPT2, G90	0.164069	0.014687	11.171	5379	0.000

For JOYMATH slope, B10

INTRCPT2, G100	0.067318	0.014016	4.803	5379	0.000

For INSTMAT slope, B11

INTRCPT2, G110	0.010900	0.017582	0.620	5379	0.535

For VALMATH slope, B12

INTRCPT2, G120	0.046444	0.019006	2.444	5379	0.015

For SAFESCH slope, B13

INTRCPT2, G130	0.011288	0.011703	0.965	5379	0.335

The outcome variable is THETA34

Final estimation of fixed effects

(with robust standard errors)

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	-1.936307	0.256915	-7.537	130	0.000
DSCHSL, G01	0.765686	0.166904	4.588	130	0.000
DSCHL, G02	0.309957	0.156476	1.981	130	0.049
DSCHM, G03	0.047342	0.085212	0.556	130	0.579
COMMXXL, G04	0.160355	0.161506	0.993	130	0.323
COMMXL, G05	-0.037516	0.134052	-0.280	130	0.780

COMML, G06	0.311936	0.116919	2.668	130	0.009
COMMM, G07	0.037242	0.099210	0.375	130	0.708
COMMS, G08	0.101627	0.084187	1.207	130	0.230
SCQ21AC, G09	0.000395	0.026976	0.015	130	0.988
SCQ20AE, G010	-0.069615	0.164693	-0.423	130	0.673
SCMAEDU, G011	0.017139	0.068481	0.250	130	0.803
MATPROM, G012	-0.041858	0.107621	-0.389	130	0.698
PPSC, G013	0.028842	0.064470	0.447	130	0.655
DVTESCH, G014	0.033932	0.054805	0.619	130	0.537
EXPTEA, G015	-0.004737	0.003718	-1.274	130	0.205
ADEQ, G016	-0.026939	0.057333	-0.470	130	0.639
HOMEWORK, G017	0.107073	0.051663	2.073	130	0.040
CLASSSIZ, G018	-0.027418	0.072572	-0.378	130	0.706
TEACOLLA, G019	0.042742	0.060518	0.706	130	0.481
For RELAGR1 slope, B1					
INTRCPT2, G10	0.017101	0.005607	3.050	5379	0.003
For SEX slope, B2					
INTRCPT2, G20	0.049061	0.022616	2.169	5379	0.030
For PAEDU1 slope, B3					
INTRCPT2, G30	0.007753	0.006141	1.262	5379	0.207
For LEARES slope, B4					
INTRCPT2, G40	0.047710	0.011109	4.295	5379	0.000
For WEALTH slope, B5					
INTRCPT2, G50	0.068723	0.061147	1.124	5379	0.262
For PRASCH slope, B6					
INTRCPT2, G60	0.069308	0.024459	2.834	5379	0.005
For TIMEHOM slope, B7					
INTRCPT2, G70	0.030171	0.014995	2.012	5379	0.044
For INTMAT slope, B8					
INTRCPT2, G80	-0.044132	0.013292	-3.320	5379	0.001
For SELFLEA slope, B9					
INTRCPT2, G90	0.164069	0.016989	9.657	5379	0.000
For JOYMATH slope, B10					
INTRCPT2, G100	0.067318	0.016267	4.138	5379	0.000
For INSTMAT slope, B11					

INTRCPT2, G110 0.010900 0.018274 0.596 5379 0.551

For VALMATH slope, B12

INTRCPT2, G120 0.046444 0.020277 2.290 5379 0.022

For SAFESCH slope, B13

INTRCPT2, G130 0.011288 0.011302 0.999 5379 0.318

Final estimation of variance components:

Random Effect Standard Variance df Chi-square P-value
Deviation Component

INTRCPT1, U0 0.40289 0.16232 130 2548.51324 0.000

level-1, R 0.58738 0.34501

Statistics for current covariance components model

Deviance = 10151.237163

Number of estimated parameters = 2

ประวัติผู้เขียนวิทยานิพนธ์

นายเรืองเดช ศิริกิจ เกิดเมื่อวันที่ 10 พฤศจิกายน พ.ศ. 2522 อยู่บ้านเลขที่ 45/150 ตำบลวัดชลอ อำเภอบางกรวย จังหวัดนนทบุรี สำเร็จการศึกษาปริญญาตรีครุศาสตรปริญญาบัณฑิต วิชาเอกคณิตศาสตร์ ภาควิชามัธยมศึกษา จากคณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เมื่อปีการศึกษา 2544 และเข้าศึกษาต่อในหลักสูตรครุศาสตรมหาบัณฑิต สาขาวิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2545 และสำเร็จการศึกษาในปีการศึกษา 2546 และในปีการศึกษา 2551 ได้เข้าศึกษาต่อในระดับปริญญา ครุศาสตรดุษฎีบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย