การประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน : การประยุกต์ใช้การทำหน้าที่ต่างกันของข้อสอบและโมเดลมูลค่าเพิ่ม

นางสาวประกฤติยา ทักษิโณ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรดุษฎีบัณฑิต
สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2552
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

A QUALITY ASSESSMENT OF SCIENCE INSTRUCTIONAL MANAGEMENT IN BASIC EDUCATION SCHOOLS: AN APPLICATION OF DIFFERENTIAL ITEM FUNCTIONING AND VALUE–ADDED MODEL

Miss.Prakittiya Tuksino

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

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 2009

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน : การประยุกต์ใช้การทำหน้าที่ต่างกันของข้อสอบและโมเดลมูลค่าเพิ่ม
โดย	นางสาวประกฤติยา ทักษิโณ
สาขาวิชา	การวัดและประเมินผลการศึกษา
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	รองศาสตราจารย์ ดร.โชติกา ภาษีผล

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้วิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาดุษฎีบัณฑิต

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

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(รองศาสตราจารย์ ดร.โชติกา ภาษีผล)

.....กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.ณัฏฐภรณ์ หลาวทอง)

.....กรรมการภายนอกมหาวิทยาลัย
(อาจารย์ ดร.บุญเรือง ศรีเหรัญ)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ประกฤติยา ทักษิโณ : การประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน : การประยุกต์ใช้การทำหน้าที่ต่างกันของข้อสอบและโมเดลมูลค่าเพิ่ม (A QUALITY ASSESSMENT OF SCIENCE INSTRUCTIONAL MANAGEMENT IN BASIC EDUCATION SCHOOLS : AN APPLICATION OF DIFFERENTIAL ITEM FUNCTIONING AND VALUE–ADDED MODEL) อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก: ศ.ดร.ศิริชัย กาญจนวาสี, อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม: รศ.ดร.โชติกา ภาษีผล, 297 หน้า

การวิจัยครั้งนี้มีจุดมุ่งหมายหลักเพื่อประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน โดยมีวัตถุประสงค์เฉพาะ 3 ข้อ ดังนี้ 1) เพื่อศึกษาผลของข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบประเมินการรู้เรื่องด้านวิทยาศาสตร์ต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน 2) เพื่อศึกษาผลของตัวแปรคุณลักษณะนักเรียนและสถานศึกษาขั้นพื้นฐานต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน และ 3) เพื่อเปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ต่อคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน ข้อมูลที่ใช้ในการวิจัยครั้งนี้เป็นข้อมูลทุติยภูมิจากโครงการประเมินผลนักเรียนระดับนานาชาติหรือ PISA ปี 2006 ประกอบด้วยนักเรียน 6,192 คน ในสถานศึกษา 212 แห่ง ซึ่งศึกษาเฉพาะวิชาวิทยาศาสตร์ จากแบบสอบประเมินการรู้เรื่องด้านวิทยาศาสตร์ จำนวน 13 ฉบับ

การดำเนินการวิจัยมีขั้นตอนหลัก 3 ขั้นตอน ดังนี้ ขั้นตอนที่ 1 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยประยุกต์ใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น (HGLM) ด้วยการวิเคราะห์ 3 ระดับ ขั้นตอนที่ 2 การวิเคราะห์มูลค่าเพิ่มในโมเดลการประเมินคุณภาพการจัดการศึกษา โดยประยุกต์ใช้โมเดลมูลค่าเพิ่ม (Value-Added Model) ด้วยโมเดลเชิงเส้นตรงแบบลดหลั่น (HLM) ด้วยการวิเคราะห์ 2 ระดับ และขั้นตอนที่ 3 การตรวจสอบผลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา ประกอบด้วย 4 โมเดลที่ต่างกัน โดยโมเดล 1 ไม่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกและไม่มีการปรับแก้คะแนนด้วยตัวแปรนักเรียนและสถานศึกษา โมเดล 2 ไม่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกและมีการปรับแก้คะแนนด้วยตัวแปรนักเรียนและสถานศึกษา โมเดล 3 ตัดข้อสอบที่ทำหน้าที่ต่างกันออกและไม่มีการปรับแก้คะแนนด้วยตัวแปรนักเรียนและสถานศึกษา โมเดล 4 ตัดข้อสอบที่ทำหน้าที่ต่างกันออกและมีการปรับแก้คะแนนด้วยตัวแปรนักเรียนและสถานศึกษา การวิเคราะห์ข้อมูลในครั้งนี้ใช้โปรแกรม HLM และ SPSS for Windows ผลการวิจัยสรุปได้ดังนี้

1) ผลของข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบวิชาวิทยาศาสตร์ต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน พบว่าข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบวิชาวิทยาศาสตร์ไม่มีผลต่อการประเมินคุณภาพการจัดการศึกษา (โมเดล 1 และ 3) และเมื่อมีการควบคุมอิทธิพลของคุณลักษณะของนักเรียนและสถานศึกษา (โมเดล 2 และ 4) พบว่าข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบวิชาวิทยาศาสตร์ทำให้การจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05

2) ผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาขั้นพื้นฐานต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน พบว่าคุณลักษณะของนักเรียนและสถานศึกษาขั้นพื้นฐานมีผลต่อการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา โดยการจัดกลุ่มคุณภาพและระดับคุณภาพการจัดการศึกษาของสถานศึกษาระหว่างโมเดล 1 และ 2 และระหว่างโมเดล 3 และ 4 มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

3) การเปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ต่อคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน พบว่าโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ที่ต่างกันทำให้การจัดคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานต่างกัน โมเดล 4 ที่มีค่าสัมประสิทธิ์การทำนาย (R^2) สูงสุดเท่ากับ 88.63% และโมเดล 2 เท่ากับ 77.75% และสหสัมพันธ์ตำแหน่งคะแนนของสเปียร์แมนของโมเดล 2 กับ 4 เท่ากับ 1.000 และโมเดล 1 กับ 3 เท่ากับ 0.981

ภาควิชา.....วิจัยและจิตวิทยาการศึกษา.....ลายมือชื่อนิสิต.....

สาขาวิชา.....การวัดและประเมินผลการศึกษา.....ลายมือชื่อ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก.....

ปีการศึกษา.....2552.....ลายมือชื่อ อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม.....

##5084461927: MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION
KEYWORDS: VALUE–ADDED MODEL / DIFFERENTIAL ITEM FUNCTIONING / QUALITY ASSESSMENT / HLM / HGLM

PRAKITTIYA TUKSINO : A QUALITY ASSESSMENT OF SCIENCE INSTRUCTIONAL MANAGEMENT IN BASIC EDUCATION SCHOOLS: AN APPLICATION OF DIFFERENTIAL ITEM FUNCTIONING AND VALUE–ADDED MODEL.
THESIS ADVISOR: PROF.SIRICHAI KANJANAWASEE, PH.D, THESIS CO-ADVISOR: ASSOC.PROF.SHOTIGA PASIPHOL, PH.D, 297 pp.

The main aim of this research is to assess the quality of science instructional management in basic education schools with the following three particular objectives: 1) to study the effects of differential item functioning in scientific knowledge tests on the quality assessment of science instructional management in basic education schools when using different quality assessment models. 2) to study the effects of student and school characteristics on the quality assessment of science instructional management in basic education schools when using different assessment models. 3) to compare the quality assessment models and the quality of science instructional management in basic education schools. The data used in this research is the secondary data from the 2006 Program for International Students Assessment (PISA), collected from 6,192 students in 212 schools, studying only science subjects. The data was based on 13 assessment tests of knowledge in science.

The research was conducted in three steps. The first step was the detection of the differential item functioning (DIF) by applying the Hierarchical Generalized Linear Models (HGLM) with three levels of analysis. In the second step, the value added in the educational quality assessment models was analyzed applying the value-added model with Hierarchical Linear Models (HLM) with two levels of analysis. Finally, in the third step, the instructional management quality resulting from assessment consisting of four models was examined. HLM and SPSS for Windows were used in analyzing the data. The research results are as follows:

1) The differential item functioning in science tests does not affect the quality assessment of science instructional management in basic education schools, (Models 1 and 3). Also, when student and school characteristics were controlled, (Models 2 and 4), it was found that the differential item functioning in science tests resulted in the ranking of schools' instructional management quality differing significantly at the 0.05 level.

2) It is found that student and school characteristics affect the quality assessment of science instructional management in basic education schools. There is a statistically significant difference at the .01 level in the rating and the ranking of instructional management quality between Models 1 and 3 and those between Models 3 and 4.

3) The comparison of the quality assessment models and the quality of science instructional management in basic education schools reveals that different models cause the rating and the ranking of the quality assessment to differ. Model 4 gives the highest proportion variance explained (R^2) of 88.63%, while that of Model 2 is 77.75%. Spearman's rank-order correlation is found to be highest at 1.000 in Model 2 and Model 4 and at 0.981 in Model 1 and Model 3.

Department:....Educational Research and Psychology..... Student's Signature Prakittiya Tuksino
Field of Study:..Educational Measurement and Evaluation... Advisor's Signature S. Kanjanawasee
Academic Year:.....2009..... Co-Advisor's Signature S. Pasiphol

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี ด้วยความเมตตากรุณาอย่างสูงยิ่งจาก ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี อาจารย์ที่ปรึกษาวิทยานิพนธ์ และรองศาสตราจารย์ ดร.โชติกา ภาษีผล ท่านอาจารย์ทั้งสองให้ความอนุเคราะห์ มีความเป็นกัลยาณมิตรต่อศิษย์เสมอมา คอยให้คำปรึกษาแนะนำ ให้ความรู้ ถ่ายทอดประสบการณ์และสละเวลาอันมีค่ายิ่ง และที่สำคัญผู้วิจัยรู้สึกซาบซึ้งใจกับโอกาสและกำลังใจที่มีให้แก่ผู้วิจัย ช่วยต่อเติมพลังใจในการพัฒนางานนี้จนทำให้วิทยานิพนธ์ฉบับนี้เสร็จสมบูรณ์ ข้าพเจ้ารู้สึกซาบซึ้งและกราบขอบพระคุณอาจารย์ทั้งสองท่านมา ณ โอกาสนี้

ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ ประธานสอบวิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.ณัฏฐภรณ์ หลาวทอง และอาจารย์ ดร.บุญเรือง ศรีเหรัญ กรรมการสอบวิทยานิพนธ์ ที่ให้คำชี้แนะที่มีคุณค่าต่องานวิทยานิพนธ์ ตรวจทานและแก้ไขงานนี้ จนทำให้งานวิจัยนี้มีความสมบูรณ์ที่สุด ตลอดจนอาจารย์ประจำภาควิชาวิจัยและจิตวิทยาการศึกษาทุกๆ ท่านที่ให้กำลังใจแก่ผู้วิจัยเสมอมา

ขอขอบพระคุณในความอนุเคราะห์ข้อมูลจากสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท.) และขอแสดงความซาบซึ้งใจ ดร.ปรีชาญ เดชศรี คุณสุชาดา ไทยแท้ ที่ช่วยชี้แนะสละเวลาในการตอบคำถาม ข้อสงสัยของผู้วิจัยได้ตลอดเวลา และ ดร.วาสนา กีรติจำเริญที่คอยเป็นกำลังใจและชี้แนะข้อมูลอันมีประโยชน์อย่างยิ่งแก่ผู้วิจัย

ขอแสดงความซาบซึ้งใจกับน้ำใจของพี่น้องสาขาวิชาการวัดและประเมินผลการศึกษา ป.เอก จุฬาฯ ทั้งในเวลาและนอกเวลาราชการ ทุกชั้นปีที่ให้ความห่วงใยต่อสุขภาพและให้กำลังใจที่ดีตลอดมา ขอบคุณ ดร.อิทธิฤทธิ์ พงษ์ปิยะรัตน์ ดร.จตุภูมิ เขตจัตุรัส ดร.ชัยวิชิต เชียรชนะ ที่ให้การแลกเปลี่ยนเรียนรู้ร่วมกันและคำแนะนำที่ดีเสมอมา ขอบคุณพี่ๆ ที่แสนดี พี่นงค์ พี่นุช จากศูนย์ทดสอบฯ ป้าน้อย พี่ต่าย และเพื่อนๆพี่ๆน้องๆ ภาควิชาวิจัยและจิตวิทยาการศึกษา จุฬาฯ ทุกคนที่ให้การช่วยเหลือ และช่วยให้การติดต่อประสานต่างๆ เป็นไปได้ด้วยดี

ข้าพเจ้ากราบขอบพระคุณ พ่อประพันธ์และแม่พอควร ทักษิโณ บิดามารดา ที่เป็นกำลังใจมอบความรัก และดูแลใส่ใจข้าพเจ้าอย่างใกล้ชิดจนทำให้เกิดแรงกายแรงใจในการพัฒนางานอยู่ตลอดเวลา น้องควรินทร์ ทักษิโณ น้องชายของข้าพเจ้า ที่คอยเป็นกำลังใจและช่วยในการจัดไฟล์ข้อมูลสำหรับการวิเคราะห์อย่างไม่เหน็ดเหนื่อย ขอบคุณแบงค์ ที่ทำให้ข้าพเจ้าเกิดแรงดลใจในการทำงานได้ตลอดเวลาทั้งกลางวันและกลางคืน จนทำให้ข้าพเจ้ามีกำลังใจทำงานได้โดยไม่รู้สึกท้อถอยและเหน็ดเหนื่อย

ท้ายนี้ข้าพเจ้าหวังเป็นอย่างยิ่งว่า งานวิทยานิพนธ์ฉบับนี้จะเป็นคุณประโยชน์ต่อการพัฒนาคุณภาพการจัดการศึกษาไทยต่อไป คุณประโยชน์อันพึงมีจากการวิจัยครั้งนี้ ผู้วิจัยขอกราบเป็นเครื่องพระคุณบิดา มารดา ที่มอบชีวิต จิตวิญญาณที่ดีแก่ผู้วิจัย ครูบาอาจารย์ทั้งในอดีตและปัจจุบัน ที่มอบแสงสว่างในการดำเนินชีวิตของผู้วิจัย

สารบัญ

		หน้า
บทคัดย่อภาษาไทย		ง
บทคัดย่อภาษาอังกฤษ		จ
กิตติกรรมประกาศ		ฉ
สารบัญ		ช
สารบัญตาราง		ฌ
สารบัญแผนภาพ		ฑ
บทที่ 1	**บทนำ**	1
	ความเป็นมาและความสำคัญของปัญหา	1
	คำถามการวิจัย	10
	วัตถุประสงค์ของการวิจัย	10
	สมมติฐานการวิจัย	11
	ขอบเขตของการวิจัย	12
	นิยามศัพท์เฉพาะ	14
	ประโยชน์ที่คาดว่าจะได้รับ	16
บทที่ 2	**เอกสารงานวิจัยที่เกี่ยวข้อง**	18
	ตอนที่ 1 มโนทัศน์เกี่ยวกับมูลค่าเพิ่มของการศึกษา	18
	ตอนที่ 2 มโนทัศน์เกี่ยวกับทฤษฎีการตอบสนองข้อสอบและการทำหน้าที่ต่างกันของข้อสอบ	30
	ตอนที่ 3 มโนทัศน์เกี่ยวกับการวิเคราะห์พหุระดับและการประยุกต์ใช้โมเดลเชิงเส้นแบบลดหลั่น	43
	ตอนที่ 4 การประเมินผลนักเรียนระดับนานาชาติ PISA 2006	63
	ตอนที่ 5 เอกสารงานวิจัยเกี่ยวกับการประเมินคุณภาพการจัดการศึกษา	79
บทที่ 3	**วิธีการดำเนินการวิจัย**	106
	3.1 ข้อมูลที่ใช้ในการวิจัย	106
	3.2 รูปแบบของแบบสอบประเมินการรู้เรื่องด้านวิทยาศาสตร์และการจัดการสอบ	107
	3.3 ปัจจัยหรือตัวแปรที่นำมาใช้ในการวิจัย	109
	3.4 วิธีการดำเนินการวิจัย	114

สารบัญ (ต่อ)

		หน้า
บทที่ 4	**ผลการวิเคราะห์ข้อมูล**……..	128
	ตอนที่ 1 ผลการวิเคราะห์ข้อมูลพื้นฐาน การตรวจสอบเงื่อนไขของการวิเคราะห์พหุระดับและการตรวจสอบความเป็นเอกมิติของแบบสอบ……..……....	128
	ตอนที่ 2 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแต่ละฉบับ และการประมาณค่าความสามารถของผู้สอบ และค่าความยากของข้อสอบ……..	144
	ตอนที่ 3 ผลการวิเคราะห์ตามโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน……..	155
	ตอนที่ 4 สรุปผลการวิเคราะห์เพื่อตอบคำถามการวิจัย…….....	163
บทที่ 5	**สรุปผลการวิจัย อภิปรายผลการวิจัย และข้อเสนอแนะ**……..	183
	5.1 สรุปผลการวิจัย……..….	184
	5.2 อภิปรายผลการวิจัย……...	189
	5.3 ข้อเสนอแนะ……..……..	204
รายการอ้างอิง	……..……..……	208
ภาคผนวก	……	219
ภาคผนวก ก	ผลการตรวจสอบเงื่อนไขเบื้องต้น…….	220
ภาคผนวก ข	ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 1 – 13……...	227
ภาคผนวก ค	ผลการประมาณค่าความยากของข้อสอบจากการวิเคราะห์ด้วยโปรแกรม HGLM-3 ของข้อสอบในแบบสอบฉบับที่ 1 - 13 ……..	262
ภาคผนวก ง	อิทธิพลคงที่ (Fixed effcet) และอิทธิพลเชิงสุ่ม (random effect) จากการวิเคราะห์ในโมเดล 1 – 4 ……	275
ภาคผนวก จ	คะแนนมูลค่าเพิ่ม ระดับที่ของคะแนนมูลค่าเพิ่ม และการจัดกลุ่มคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษา 212 แห่ง จากการวิเคราะห์ด้วยโมเดล 1 – 4……..……...	282
ภาคผนวก ฉ	ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มของสถานศึกษาจากผลการวิเคราะห์ 4 โมเดล.....	290
ประวัติผู้เขียนวิทยานิพนธ์.....		297

สารบัญตาราง

		หน้า
ตารางที่ 1	สรุปหลักการและแนวคิดพื้นฐานของวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ตรวจให้คะแนนสองค่า…….……..	41
ตารางที่ 2	สรุปหลักการและแนวคิดพื้นฐานของวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า……..……	42
ตารางที่ 3	เปรียบเทียบค่าความสามารถของผู้สอบและความยากง่ายของข้อสอบจากการประมาณค่าด้วยโมเดลราสซ์และโมเดล HGLM-3L……..……	48
ตารางที่ 4	ค่าความเที่ยงโดยเฉลี่ยของแบบสอบในโครงการประเมินผล PISA 2006…….....	69
ตารางที่ 5	สรุปความหมายของตัวแปรที่ใช้ศึกษาในโครงการประเมินผล PISA 2006 จำแนกตามระดับของข้อมูล.....	71
ตารางที่ 6	ผลการสังเคราะห์ตัวแปรทำนายที่มีอิทธิผลต่อผลสัมฤทธิ์ทางการเรียนจำแนกตามระดับตัวแปร……..……..	95
ตารางที่ 7	อันดับของตัวแปรที่นิยมนำมาศึกษาปัจจัยที่มีผลต่อคุณภาพการจัดการศึกษา และแหล่งข้อมูลในการเก็บข้อมูลของโครงการประเมินผล PISA 2006.....…..	97
ตารางที่ 8	จำนวนกลุ่มตัวอย่างนักเรียนในโครงการประเมินผล PISA 2006 จำแนกตามสังกัดของสถานศึกษา……..……	106
ตารางที่ 9	ลักษณะของข้อสอบที่ใช้ในโครงการประเมินผล PISA 2006 จำแนกตามสมรรถนะทางวิทยาศาสตร์……..……..	107
ตารางที่ 10	จำนวนข้อสอบที่ใช้ในโครงการประเมินผล PISA 2006 ของแบบสอบ 13 ฉบับ จำแนกตามลักษณะการตรวจให้คะแนนและสมรรถนะทางวิทยาศาสตร์.....	108
ตารางที่ 11	จำนวนกลุ่มตัวอย่างนักเรียนที่ทำการทดสอบแต่ละฉบับ ค่าประมาณของอิทธิพลจากแบบสอบ และค่าประมาณอิทธิพลจากแบบสอบหลังการแปลงคะแนนสเกลของ PISA.....	109
ตารางที่ 12	สเกลการวัดและความหมายของตัวแปรในระดับนักเรียนและสถานศึกษา…….....	111
ตารางที่ 13	จำนวนข้อสอบที่ตัดได้ในแต่ละสมรรถนะทางวิทยาศาสตร์ตามเกณฑ์การตัดข้อสอบในแบบสอบฉบับที่ 1 - 13…….....	120
ตารางที่ 14	ผลการเปรียบเทียบความยากเฉลี่ยของข้อสอบจากค่าอิทธิพลจากแบบสอบและค่าความยากเฉลี่ยของข้อสอบในแบบสอบฉบับที่ 1 ถึง 13	123

สารบัญตาราง(ต่อ)

หน้า

ตารางที่ 15 สรุปสมการวิเคราะห์มูลค่าเพิ่มด้วยโมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น 2 ระดับ ในโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา 4 โมเดล..... 125

ตารางที่ 16 ความถี่ ร้อยละของตัวแปรจัดประเภทจำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา..... 129

ตารางที่ 17 ค่าสถิติพื้นฐานของตัวแปรต่อเนื่อง จำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา..... 133

ตารางที่ 18 ค่าสถิติพื้นฐานของคะแนนวิชาวิทยาศาสตร์ของนักเรียนจากผลการประเมินตามโครงการประเมินผล PISA 2006 จำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา..... 135

ตารางที่ 19 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร ในระดับนักเรียน..... 139

ตารางที่ 20 ค่า Tolerance และ VIFจากผลการทดสอบความสัมพันธ์ระหว่างตัวแปรต้นระดับนักเรียน..... 140

ตารางที่ 21 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร ในระดับสถานศึกษา..... 141

ตารางที่ 22 ค่า Tolerance และ VIF จากผลการทดสอบความสัมพันธ์ระหว่างตัวแปรต้นระดับสถานศึกษา..... 142

ตารางที่ 23 ผลการตรวจข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบทั้ง 13 ฉบับ..... 145

ตารางที่ 24 จำนวนข้อสอบที่ทำหน้าที่ต่างกันที่ตัดออกไปในแบบสอบทั้ง 13 ฉบับ.. 146

ตารางที่ 25 การเปรียบเทียบโครงสร้างของแบบสอบก่อนตัดและหลังตัดข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบทั้ง 13 ฉบับ..... 147

ตารางที่ 26 สรุปผลของตัวแปรระดับสถานศึกษาต่อคุณลักษณะของนักเรียนที่ทำให้โอกาสการตอบข้อสอบที่ทำหน้าที่ต่างกันได้ถูกในแบบสอบทั้ง 13 ฉบับ..... 148

ตารางที่ 27 สรุปลักษณะของข้อสอบที่ทำหน้าที่ต่างกันในแต่ละคุณลักษณะของนักเรียน..... 150

ตารางที่ 28 ค่าสถิติพื้นฐานของค่าความยากของข้อสอบในแบบสอบฉบับที่ 1 ถึง 13..... 151

ตารางที่ 29 ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบ**ก่อนการตัด**ข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบฉบับที่ 1 ถึง 13 152

ตารางที่ 30 ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบ**หลังการตัด**ข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบฉบับที่ 1 ถึง 13..... 153

สารบัญตาราง(ต่อ)

หน้า

ตารางที่ 31 สรุปคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ ก่อนและหลังการตัดข้อสอบการทำหน้าที่ต่างกันออก..... 154

ตารางที่ 32 สรุปค่าสัมประสิทธิ์ของตัวแปรตัวแปรคุณลักษณะระดับนักเรียนและสถานศึกษาที่ได้จากการวิเคราะห์โมเดล 1 ถึง 4..... 161

ตารางที่ 33 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มวิชาวิทยาศาสตร์ของสถานศึกษาจากผลการวิเคราะห์ 4 โมเดล..... 164

ตารางที่ 34 ผลการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาจากผลการวิเคราะห์ 4 โมเดล จำแนกตามคุณลักษณะของสถานศึกษา..... 167

ตารางที่ 35 ผลการจัดกลุ่มคุณภาพการจัดการศึกษา การจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษา และค่าสถิติทดสอบ Wilcoxon Signed Ranks Test จากผลการวิเคราะห์โมเดล 1 และ โมเดล 3..... 171

ตารางที่ 36 ผลการจัดกลุ่มคุณภาพการจัดการศึกษา การจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษา และค่าสถิติทดสอบ Wilcoxon Signed Ranks Test จากผลการวิเคราะห์โมเดล 2 และ โมเดล 4..... 172

ตารางที่ 37 ผลการจัดกลุ่มคุณภาพการจัดการศึกษา การจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษา และค่าสถิติทดสอบ Wilcoxon Signed Ranks Test จากผลการวิเคราะห์โมเดล 1 และ โมเดล 2..... 174

ตารางที่ 38 ผลการจัดกลุ่มคุณภาพการจัดการศึกษา การจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษา และค่าสถิติทดสอบ Wilcoxon Signed Ranks Test จากผลการวิเคราะห์โมเดล 3 และ โมเดล 4..... 175

ตารางที่ 39 สรุปค่าความแปรปรวนของตัวแปรตามที่อธิบายได้ของโมเดล 1 ถึง 4 177

ตารางที่ 40 ผลการจัดอันดับคุณภาพการจัดการศึกษาของสถานศึกษาโดยเปรียบเทียบกับโมเดล 4..... 178

ตารางที่ 41 ผลการวิเคราะห์ความสอดคล้องของการจัดอันดับคุณภาพการจัดการศึกษาด้วยโมเดลการประเมินคุณภาพการจัดการศึกษาที่ต่างกัน 4 โมเดล..... 179

ตารางที่ 42 ค่าสหสัมพันธ์ของอันดับของคะแนนมูลค่าเพิ่มที่ได้จากโมเดลประเมินคุณภาพการจัดการศึกษาที่ต่างกัน 4 โมเดล..... 180

ตารางที่ 43 สรุปการเปรียบเทียบลักษณะโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา และแนวทางการเลือกใช้โมเดลประเมินคุณภาพการจัดการศึกษาของสถานศึกษา..... 181

สารบัญตาราง(ต่อ)

หน้า

ตารางที่ 44 สรุปความสอดคล้องของผลการประเมิน สัมประสิทธิ์การทำนายและค่าสหสัมพันธ์ของตำแหน่งคะแนนมูลค่าเพิ่มจากผลการประเมิน 4 โมเดล.. 188

ตารางที่ 45 ผลการเปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา 4 โมเดล..... 201

ตารางที่ 46 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 1..... 228

ตารางที่ 47 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 2..... 231

ตารางที่ 48 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 3..... 233

ตารางที่ 49 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 4..... 235

ตารางที่ 50 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 5..... 237

ตารางที่ 51 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 6..... 240

ตารางที่ 52 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 7..... 242

ตารางที่ 53 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 8..... 243

ตารางที่ 54 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 9..... 245

ตารางที่ 55 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 10..... 246

ตารางที่ 56 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 11..... 248

ตารางที่ 57 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 12..... 250

ตารางที่ 58 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 13..... 251

ตารางที่ 59 สรุปข้อสอบที่ทำหน้าที่ต่างกันฉบับที่ 1 ถึง 13 ในแต่ละคุณลักษณะของนักเรียน..... 253

ตารางที่ 60 ผลการวิเคราะห์อิทธิพลของตัวแปรระดับสถานศึกษาที่มีผลต่อโอกาสการตอบข้อสอบถูกในข้อสอบที่ทำหน้าที่ต่างกันแต่ละฉบับ..... 258

ตารางที่ 61 ผลการประมาณค่าความยากของข้อสอบจากการวิเคราะห์ด้วยโปรแกรม HGLM-3 จำแนกตามฉบับข้อสอบจำนวน 13 ฉบับ..... 263

ตารางที่ 62 อิทธิพลคงที่ (Fixed effcet)และอิทธิพลเชิงสุ่ม (random effect) จากการวิเคราะห์โมเดล 1..... 276

ตารางที่ 63 อิทธิพลคงที่ (Fixed effcet) และอิทธิพลเชิงสุ่ม (random effect) จากการวิเคราะห์ในโมเดล 2 ที่มีการควบคุมตัวแปรคุณลักษณะระดับนักเรียน..... 277

ตารางที่ 64 อิทธิพลคงที่ (Fixed effcet)และอิทธิพลเชิงสุ่ม (random effect) จากการวิเคราะห์โมเดล 3..... 279

สารบัญตาราง(ต่อ)

หน้า

ตารางที่ 65 อิทธิพลคงที่ (Fixed effcet) และอิทธิพลเชิงสุ่ม (random effect) จากการวิเคราะห์ในโมเดล 4 ที่มีการควบคุมตัวแปรคุณลักษณะระดับนักเรียน..... 280
ตารางที่ 66 คะแนนมูลค่าเพิ่ม ระดับที่ของคะแนนมูลค่าเพิ่ม และการจัดกลุ่มคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษา 212 แห่ง จากการวิเคราะห์ด้วยโมเดล 1 – 4..... 283
ตารางที่ 67 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มของสถานศึกษาจากผลการวิเคราะห์โมเดล 1 ที่จัดกลุ่มคุณภาพการจัดการศึกษา จำแนกตามคุณลักษณะของสถานศึกษา..... 291
ตารางที่ 68 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มของสถานศึกษาจากผลการวิเคราะห์โมเดล 2 ที่จัดกลุ่มคุณภาพการจัดการศึกษา จำแนกตามคุณลักษณะของสถานศึกษา..... 292
ตารางที่ 69 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มของสถานศึกษาจากผลการวิเคราะห์โมเดล 3 ที่จัดกลุ่มคุณภาพการจัดการศึกษา จำแนกตามคุณลักษณะของสถานศึกษา..... 293
ตารางที่ 70 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มของสถานศึกษาจากผลการวิเคราะห์โมเดล 4 ที่จัดกลุ่มคุณภาพการจัดการศึกษา จำแนกตามคุณลักษณะของสถานศึกษา..... 295

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

		หน้า
ภาพที่ 1	โครงสร้างความสัมพันธ์ของข้อมูลในการวิเคราะห์ข้อสอบแบบพหุระดับ……	6
ภาพที่ 2	โครงสร้างความสัมพันธ์ของข้อมูลในสถานศึกษาที่มีลักษณะลดหลั่น…	7
ภาพที่ 3	โค้งคุณลักษณะของข้อสอบ……	31
ภาพที่ 4	การเกิด DIF แบบเอกรูป……	39
ภาพที่ 5	การเกิด DIF แบบอเนกรูป……	40
ภาพที่ 6	กรอบความคิดเชิงทฤษฎีของตัวแปรที่เกี่ยวข้องกับผลการเรียนรู้ของนักเรียน……	99
ภาพที่ 7	โมเดลการวิเคราะห์ในโมเดล 1 (UnDetect DIF & Unadjusted)……	101
ภาพที่ 8	โมเดลการวิเคราะห์ในโมเดล 2 (UnDetect DIF & Adjusted)……	102
ภาพที่ 9	โมเดลการวิเคราะห์ในโมเดล 3 (Detect DIF & Unadjusted)……	103
ภาพที่ 10	โมเดลการวิเคราะห์ในโมเดล 4 (Detect DIF & Adjusted)……	104
ภาพที่ 11	กรอบแนวคิดการวิจัยเพื่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน……	105
ภาพที่ 12	การจำลองไฟล์การวิเคราะห์ข้อมูลระดับข้อสอบ……	115
ภาพที่ 13	การจำลองไฟล์การวิเคราะห์ข้อมูลระดับนักเรียน……	115
ภาพที่ 14	การจำลองไฟล์การวิเคราะห์ข้อมูลระดับสถานศึกษา……	116
ภาพที่ 15	สรุปขั้นตอนการดำเนินการวิจัย……	127
ภาพที่ 20	ฮีสโตแกรม (Histogram) ของคะแนนมูลค่าเพิ่มจากโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน จำนวน 212 แห่ง……	164
ภาพที่ 21	ความสัมพันธ์ของตำแหน่งคะแนนผลการประเมินคุณภาพการจัดการศึกษา 4 โมเดล……	180
ภาพที่ 22	กราฟ Q-Q plot ของ residual level 1……	221
ภาพที่ 23	กราฟ Q-Q plot ของ residual level 2……	221
ภาพที่ 24	กราฟระหว่างคะแนนความคลาดเคลื่อน (residual) กับค่าทำนาย (Expected value)……	222
ภาพที่ 25	แผนภาพสกรีของการวิเคราะห์องค์ประกอบของแบบสอบฉบับที่ 1 – 13……	223
ภาพที่ 26	แผนภาพองค์ประกอบของแบบสอบฉบับที่ 1 – 13……	225

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

มาตรฐานการศึกษาของชาติ ตามพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 และฉบับแก้ไข ปี พ.ศ. 2545 มีอุดมการณ์และหลักการจัดการศึกษาเพื่อพัฒนาสังคมไทยให้เป็นสังคมแห่งการเรียนรู้ ให้คนไทยได้รับโอกาสเท่าเทียมกันทางการศึกษา และพัฒนาคนได้อย่างต่อเนื่องตลอดชีวิต หน่วยงานที่มีหน้าที่โดยตรงคือสถานศึกษาที่จะต้องมีการจัดกระบวนการเรียนรู้ เพิ่มพูนความรู้ ทักษะให้นักเรียนเป็นบุคคลที่สามารถเรียนรู้ และสามารถพัฒนาตนเองได้ ในหมวด 6 ว่าด้วยมาตรฐานและการประกันคุณภาพการศึกษา กำหนดให้หน่วยงานต้นสังกัดและสถานศึกษาจัดทำรายงานเกี่ยวกับคุณภาพของสถานศึกษา เพื่อนำไปสู่การพัฒนาคุณภาพ (มาตรา 48) ซึ่งมุมมองของคุณภาพการศึกษามีแนวคิดที่กว้างไกลขึ้นในทุกกลุ่มที่มีความเกี่ยวข้องกับการศึกษา ทำให้เพิ่มความกดดันในการปฏิบัติที่จะทำให้เกิดประสิทธิผลของระบบการศึกษามากขึ้น และมีความต้องการประเมินคุณภาพการจัดการศึกษาให้มีความถูกต้อง

ในการตรวจสอบคุณภาพการจัดการศึกษาของสถานศึกษาจำเป็นต้องมีการดำเนินงานที่สำคัญคือ การประเมินผลเพื่อให้ได้ข้อมูลย้อนกลับที่จะสะท้อนให้เห็นถึงการบรรลุถึงเป้าหมายจากผลการดำเนินงาน รวมทั้งทำให้ทราบจุดอ่อนหรือปัญหาที่ต้องปรับปรุงแก้ไข ด้วยเหตุนี้จึงต้องให้ความสำคัญกับการประเมินผล โดยเฉพาะการประเมินคุณภาพจากหน่วยงานที่เป็นกลาง ซึ่งหน่วยงานในระดับชาติ (National) ที่ดำเนินการติดตามและประเมินผลการจัดการศึกษาของสถานศึกษาตามมาตรฐานการศึกษาของชาติคือ สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือ สทศ. และสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) หรือ สมศ. โดยทั้งสองหน่วยงานต่างมีหน้าที่ในการวัดและประเมินผลการจัดการศึกษาของประเทศ นอกจากนี้ยังมีหน่วยงานระดับนานาชาติ(International) ที่สำคัญ ได้แก่สมาคมนานาชาติเพื่อประเมินผลสัมฤทธิ์ทางการศึกษา (The International Association for the Evaluation of Educational Achievement: IEA) ในโครงการศึกษาคณิตศาสตร์และวิทยาศาสตร์ระดับนานาชาติครั้งที่ 3 (The Third International Mathematics and Science Study: TIMSS-3) และองค์กรเพื่อความร่วมมือและพัฒนาทางเศรษฐกิจ (Organisation for Economic Co-operation and Development: OECD) ในการดำเนินโครงการประเมินผลนักเรียนระดับนานาชาติ (Programme for International Student Assessment: PISA) เพื่อศึกษารูปแบบประสิทธิภาพของระบบการศึกษาของประเทศสมาชิก และป้อนข้อมูลกลับให้ประเทศสมาชิกทราบเกี่ยวกับระบบการศึกษาของประเทศ ทั้งนี้ผลการประเมินคุณภาพการจัดการศึกษาระดับนานาชาติจะเป็นตัวบ่งชี้ผลสัมฤทธิ์ของนักเรียนในประเทศเมื่อเปรียบเทียบกับประเทศอื่น และมีการจัดอันดับคุณภาพการศึกษาของประเทศ ตลอดจนได้ข้อสารสนเทศอย่างหลากหลายที่มีคุณค่าเพื่อตรวจสอบการปฏิบัติงานได้

นอกจากนี้ยังมีหน่วยงานในการประเมินคุณภาพการจัดการศึกษาโดยสถาบันระดับนานาชาติเพื่อการพัฒนาการจัดการโดยใช้ดัชนี IMD (International Institute for Management Development) และการประเมินคุณภาพการจัดการศึกษาของธนาคารโลก(World Bank) ซึ่งจากการประเมินคุณภาพการจัดการศึกษาในหน่วยงานดังกล่าวข้างต้น พบว่าผลการประเมินอยู่ในระดับต่ำ ตัวอย่างเช่นการใช้ดัชนี IMD ในปี 2550 พบว่าประเทศไทยได้คะแนน 57.76 จากคะแนนเต็ม 100 คะแนน การประเมินคุณภาพการจัดการศึกษาของธนาคารโลก(World Bank) พบว่าการจัดงบประมาณการศึกษาในช่วงปี 2550 มีสัดส่วนต่อ GDP เท่ากับร้อยละ 3.86 และ ปี 2551 มีสัดส่วนเท่ากับร้อยละ 3.32 ซึ่งยูเนสโกแนะนำงบการศึกษาต่อ GDP ที่เหมาะสมอยู่ที่ร้อยละ 8 (เลขาธิการสภาการศึกษา, 2551) การประเมินในโครงการ TIMSS-3 ในปี 2007 พบว่าประเทศไทยมีผลการประเมินอยู่ในอันดับที่ 24 จาก 49 ประเทศ (สสวท, 2552) การประเมินคุณภาพการจัดการศึกษาในโครงการประเมิน PISA พบว่าผลการประเมิน PISA ปี 2000 และ 2003 พบว่าไทยมีคะแนนเฉลี่ยในการประเมินทั้งสามด้านอยู่ในอับดับ 32 - 36 จากทั้งหมด 41 ประเทศ และการประเมิน PISA 2006 พบว่าไทยมีคะแนนอยู่ในอันดับ 46 – 47 จากทั้งหมด 57 ประเทศ (OECD, 2007) ส่วนการประเมินระดับชาติของประเทศไทยโดยหน่วยงานสำนักทดสอบแห่งชาติ หรือ สทศ. และสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) หรือ สมศ. โดยใช้ผลการประเมินจากคะแนนสอบพบว่าโดยเฉลี่ยอยู่ในระดับต่ำ โดยเฉพาะอย่างยิ่งในวิชาวิทยาศาสตร์และคณิตศาสตร์

โดยส่วนมาก การประเมินคุณภาพการจัดการศึกษาจะใช้ตัวบ่งชี้จากคะแนนจากการทำแบบสอบเป็นสำคัญเพื่อให้รู้ถึงผลสัมฤทธิ์ของหลักสูตรที่ใช้ในการเรียน และบ่งชี้ถึงผลสัมฤทธิ์การดำเนินการของสถานศึกษา (Postlethwaite, 2004) และคะแนนที่ได้จากแบบสอบระดับชาติหรือระดับนานาชาติจะนำมาใช้ในการบ่งบอกคุณภาพการจัดการศึกษาของสถานศึกษา โดยทั่วไปแล้วการใช้คะแนนที่ได้จากการสังเกตโดยใช้แบบสอบจะเป็นค่าที่ได้จากผลรวมของคะแนนที่แท้จริงกับค่าความคลาดเคลื่อนจากการวัด หรือเขียนสมการได้เป็น คะแนนที่สังเกตได้ (x) = คะแนนที่แท้จริง (T) + ความคลาดเคลื่อน (e) ดังนั้นถ้ามีแหล่งของความคลาดเคลื่อนเกิดขึ้น คะแนนที่ได้จากการสอบจึงไม่สามารถบอกได้ว่าเป็นความสามารถที่แท้จริงของนักเรียนและคุณภาพของผลลัพธ์ของกระบวนการศึกษาจึงไม่สามารถนิยามได้เชิงสัมบูรณ์ นอกจากนี้ยังมีปัจจัยหลายอย่างที่เป็นคุณลักษณะพื้นฐานของนักเรียนและการบริหารจัดการภายในสถานศึกษาต่างส่งผลกระทบต่อการทำคะแนนในแบบสอบ การประเมินคุณภาพการจัดการศึกษาที่ผ่านมาโดยส่วนใหญ่ จะใช้คะแนนจากการสอบเพื่อบ่งชี้ว่าสถานศึกษามีการปฏิบัติที่ได้กว่าสถานศึกษาอื่น โดยไม่คำนึงถึงพื้นฐานของนักเรียน ซึ่ง Hill (1995 cited in Downes and Vindurampulle, 2007) ชี้ให้เห็นว่าคะแนนดิบไม่ได้เป็นตัวบ่งชี้ถึงความแตกต่างของการดำเนินงานของสถานศึกษาได้อย่างถูกต้อง

นักการศึกษาจึงมีความสนใจในการพัฒนาวิธีการตรวจสอบสาระสำคัญที่ควรได้รับจากผลสัมฤทธิ์ที่ได้จากการวัด ด้วยเหตุผล 2 ประการ ประการแรก ต้องการพัฒนาคุณภาพการปฏิบัติงานสู่ระดับมาตรฐาน ที่ต้องการใช้ตัวบ่งชี้ผลการปฏิบัติได้อย่างชัดเจนและถูกต้อง ประการที่

สองคือเพื่อลดความลำเอียง และเพิ่มความตรง (Validity) ให้กับผลการวิเคราะห์ด้วยการพิจารณาปัจจัยอื่นที่ส่งผลต่อสัมฤทธิ์ (Lindsey and Deforgers, 1998) และตั้งแต่ปี 1996 วิธีที่ได้รับความสนใจเป็นอย่างมากเป็นวิธีการที่มีการปรับแก้ผลสัมฤทธิ์ตามคุณลักษณะที่แตกต่างของนักเรียนซึ่งเป็นปัจจัยนำเข้า โดยใช้หลักการวิเคราะห์ถดถอยและการวิเคราะห์พหุระดับ (Multilevel Analysis) ซึ่งโมเดลการวิเคราะห์ขึ้นอยู่กับความซับซ้อนของปัจจัยที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียน (Fitz-Gibbon, 1996; Heck, 2000) นอกจากนี้การเลือกใช้วิธีในการตรวจสอบคุณภาพการจัดการศึกษาควรเป็นการวิเคราะห์ง่ายต่อการทำความเข้าใจและมีความตรงตามหลักสถิติ (Fitz-Gibbon, 1995 cited in Linsey and Deforgers, 1998) และสนับสนุนให้ใช้โมเดลการวิเคราะห์ที่ง่ายต่อการทำความเข้าใจโดยการวิเคราะห์ค่าเศษเหลือ (Residual) ซึ่งเป็นค่าที่แสดงมูลค่าเพิ่ม (value-added) ของการศึกษา ในระยะต่อมาจึงมีความสนใจในการพัฒนาโมเดลมูลค่าเพิ่ม(Value-Added Model) ทางการศึกษามากขึ้น Grilli & Rampichini (2008) ได้กล่าวถึงการวัดคุณภาพของหน่วยงานการศึกษาสามารถวัดผ่านปัจจัยนำเข้า (input) และผลลัพธ์ (output) หรือเป็นการวัดมูลค่าเพิ่ม (value-added) ดังนั้นการประเมินคุณภาพการจัดการศึกษา ส่วนหนึ่งจึงเป็นการประเมินปัจจัยนำเข้าที่เกิดเนื่องจากการปฏิบัติงานของสถานศึกษา กล่าวคือมูลค่าเพิ่มมีค่าเท่ากับผลต่างของผลลัพธ์ที่เกิดขึ้นกับผลลัพธ์ที่คาดหวัง เมื่อพิจารณาปัจจัยนำเข้าหรือตัวแปรของคุณลักษณะของนักเรียนและสถานศึกษา

โมเดลมูลค่าเพิ่ม (Value-added Model) เป็นวิธีการหนึ่งที่รายงานผลเพื่อสะท้อนข้อมูลเกี่ยวกับการจัดการศึกษาว่าสถานศึกษาได้สร้างมูลค่าเพิ่มในผลการเรียนรู้ของนักเรียนได้มากน้อยเพียงใด โดยเปรียบเทียบคะแนนผลการเรียนรู้ที่เกิดขึ้นจริงหรือสังเกตได้ (Observed scores) กับคะแนนผลการเรียนรู้ที่ทำนายได้ (Predicted scores) จากตัวแปรภูมิหลัง ตัวแปรบริบทของชุมชนสังคม หรือผลสัมฤทธิ์เดิม ซึ่งคาดว่ามีผลต่อการเรียนรู้แต่เป็นปัจจัยที่อยู่นอกเหนือการจัดการหรือการควบคุมของสถานศึกษา (ศิริชัย กาญจนวาสี, 2550) Downes and Vindurampulle (2007) ชี้ให้เห็นว่าข้อมูลของการวิเคราะห์มูลค่าเพิ่มเป็นเครื่องมือหนึ่งสำหรับวิเคราะห์การปฏิบัติงานของสถานศึกษา ที่สามารถเปรียบเทียบผลการปฏิบัติงานของสถานศึกษาเชิงสัมพัทธ์กับสถานศึกษาอื่นซึ่งเป็นการวัดที่มีประสิทธิภาพและใช้เป็นข้อมูลในการพิจารณาร่วมกับปัจจัยอื่น เช่น ผลการสอบที่อยู่ในรูปคะแนนดิบ การประเมินครู การประเมินตนเองของสถานศึกษา การทบทวนการตัดสินและแนวโน้มของสถานศึกษา นอกจากนี้ OECD (2008) ได้แนะนำว่าโมเดลมูลค่าเพิ่มเป็นรูปแบบที่ดีที่สุดและใช้เป็นเครื่องมือในการส่งเสริมพัฒนาสถานศึกษามากกว่าการใช้ข้อมูลพื้นฐานเพียงอย่างเดียว การวัดผลการปฏิบัติงานของสถานศึกษาที่มีความถูกต้องจะช่วยบอกแนวทางการพัฒนาและการนำไปใช้ของนโยบายการศึกษา นอกจากนี้การวิเคราะห์มูลค่าเพิ่มยังให้ข้อมูลสำหรับการตัดสินใจในด้านต่างๆ หลายด้าน (Drury & Doran, 2003) เช่น การประเมินนโยบายและโครงการเพื่อแสดงให้เห็นถึงข้อมูลที่ถูกต้องของการทำงานว่าสิ่งใดควรดำเนินการต่อไปหรือไม่ควรดำเนินการต่อไป และใช้ข้อมูลจากการวิเคราะห์เพื่อการเปลี่ยนแปลงรูปแบบการสอนของครู ซึ่งครูจะได้รับข้อมูลย้อนกลับเกี่ยวกับวิธีการสอนที่ดี ความชัดเจนของเป้าหมายของการเรียนรู้ ครู

สามารถทราบข้อมูลสำคัญที่มีส่วนทำให้ผลสัมฤทธิ์ของนักเรียนเพิ่มขึ้นอันเป็นผลจากการปฏิบัติการสอนหรืออิทธิพลอื่นที่นอกเหนือจากสถานศึกษา นอกจากนี้การวิเคราะห์มูลค่าเพิ่มยังสามารถบอกประสิทธิภาพของครูมีผลต่อผลสัมฤทธิ์ของนักเรียน ซึ่งประสิทธิภาพของครูจะเป็นตัวบ่งชี้ที่สำคัญของความสัมพันธ์ของสถานศึกษาที่มีต่อการเรียนรู้ของนักเรียน

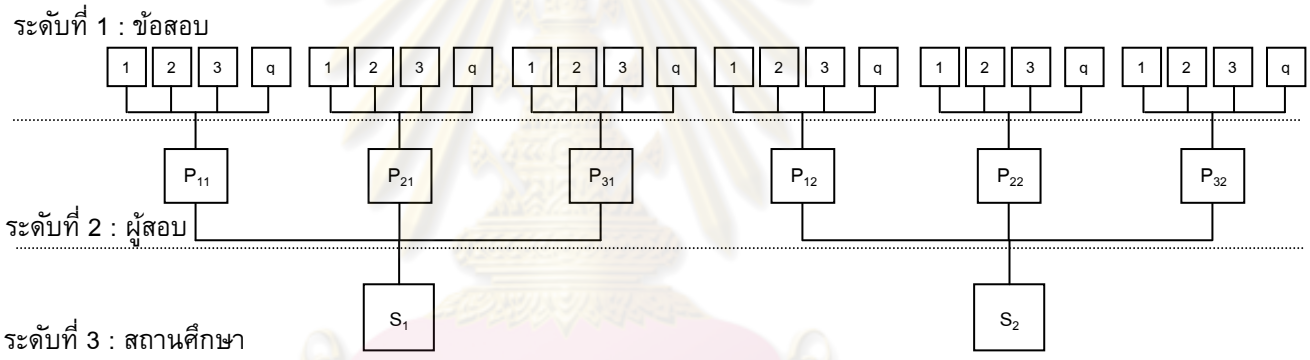
การพัฒนาโมเดลมูลค่าเพิ่มได้มีเพิ่มมากขึ้น โดยใช้ข้อมูลที่ได้จากการวัดความสามารถหรือผลสัมฤทธิ์ทางการเรียนของนักเรียน และควบคุมอิทธิพลของตัวแปรคุณลักษณะพื้นฐานของนักเรียนและสถานศึกษา เพื่อประมาณมูลค่าเพิ่มของแต่ละสถานศึกษา เช่นระบบการประเมินมูลค่าเพิ่มของรัฐเทนเนสซี (Tennessee Value-Added Assessment System: TVAAS) ซึ่งพัฒนาโดย William L. Sanders และได้ขยายการศึกษาเพิ่มเติมซึ่งเป็นชื่อที่รู้จักกันดีของการระบบการประเมินมูลค่าเพิ่มทางการศึกษา (Education Value-Added Assessment System: EVAAS) (Sanders, 1998) Ballou, Willia, Wright. (2004) ได้ศึกษาการประเมินมูลค่าเพิ่มของครูโดยการควบคุมอิทธิพลของตัวแปรพื้นฐานของนักเรียน เช่นเศรษฐานะของนักเรียนและปัจจัยบริบทอื่นๆ ที่ส่งผลต่อระดับผลสัมฤทธิ์ของนักเรียน Ronald Heck (2000) ได้ตรวจสอบผลกระทบของคุณภาพสถานศึกษาต่อผลลัพธ์ของสถานศึกษาและการพัฒนา ด้วยการประยุกต์ใช้วิธีการมูลค่าเพิ่ม โดยการเปรียบเทียบจากองค์ประกอบสิ่งแวดล้อมภายในสถานศึกษา ต่อมาในปัจจุบันมีความต้องการนำข้อมูลที่ได้มาใช้แสดงถึงความผิดชอบต่อสาธารณะมากขึ้น (Drury & Doran, 2003) โดยใช้ผลการประเมินจากคะแนนที่ได้จากการทำแบบสอบมาตรฐาน ซึ่งคะแนนจะบอกสถานะในปัจจุบันของนักเรียน และใช้ข้อมูลที่เป็นคะแนนจากการสอบของนักเรียนเพื่อศึกษาเกี่ยวกับผลการดำเนินงานของสถานศึกษาหรืออิทธิพลของสถานศึกษา (school effects) และ Haegeland and Kirkeboen (2008) ได้ศึกษาโมเดลมูลค่าเพิ่ม (Value added models: VAM) ซึ่งเป็นโมเดลที่พิจารณาเฉพาะตัวแปรความรู้เดิมของนักเรียน และโมเดลมูลค่าเพิ่มเชิงบริบท (Contextualised value-added modeling: CVA) ซึ่งเป็นโมเดลที่นำตัวแปรบริบท ได้แก่ตัวแปรเศรษฐานะเข้าร่วมพิจารณาในโมเดลด้วย และในประเทศอังกฤษ โมเดลที่มีความโดดเด่น คือโมเดลมูลค่าเพิ่มเชิงบริบท (Contextualised value-added modelling) (OECD, 2008) เป็นโมเดลที่มีการควบคุมอิทธิพลของตัวแปรอย่างหลากหลาย ซึ่งเป็นตัวแปรคุณลักษณะทางด้านเศรษฐานะที่มีผลต่อการเปลี่ยนแปลงความสามารถของนักเรียน และคุณลักษณะที่มีอิทธิพลต่อความก้าวหน้าของนักเรียนและอยู่นอกเหนือการควบคุมของสถานศึกษา เช่น เพศ เชื้อชาติ ศาสนา ความจำเป็นในการศึกษาพิเศษ ภาษาที่ใช้ การย้ายโรงเรียน การดูแลเอาใจใส่ เป็นต้น ส่วนในประเทศไทย ได้มีการศึกษารูปแบบการวิเคราะห์มูลค่าเพิ่มของผลสัมฤทธิ์ทางการเรียนของนักเรียน เช่น สุชีรา มะหิเมือง (2547) ได้ศึกษาปัจจัยที่ส่งผลต่อผลสัมฤทธิ์และพัฒนาการทางวิชาการด้วยวิธีการวิเคราะห์มูลค่าเพิ่ม โดยมีวัตถุประสงค์เพื่อวัดมูลค่าเพิ่มเนื่องจากผลสัมฤทธิ์ทางการเรียน ตัวแปรที่นำมาใช้ในการศึกษาได้แก่ ผลสัมฤทธิ์เดิมของนักเรียน เศรษฐานะ พื้นฐานทางภาษา บริบททั่วไปของสถานศึกษา การปฏิบัติงานทางวิชาการ ทำเลที่ตั้งของสถานศึกษา เป็นต้น ชูเวช ชาญสง่าเวช และสิงหา เจียมศิริ (2546) ได้ศึกษาการวัดมูลค่าเพิ่มจากการดำเนินงานของสถาบันอุดมศึกษา โดยมีวัตถุประสงค์ที่

จะพัฒนาระบบการวัดมูลค่าเพิ่มจากการดำเนินงานของสถาบันอุดมศึกษา เพื่อใช้ในกรณีศึกษาโครงการนำร่องเพื่อทดลองจัดระดับสถาบันอุดมศึกษา

อย่างไรก็ตาม ประเด็นหนึ่งที่ควรพิจารณาก่อนการใช้ผลจากการประเมิน คือแบบสอบ (test) ซึ่งการใช้วิธีมูลค่าเพิ่มต้องคำนึงแบบสอบไปใช้ต้องคำนึงถึงความตรงและความเที่ยง (Riddell, 2008) ทั้งนี้เนื่องจากคะแนนที่ได้จากการประเมินและนำมาใช้ในการวิเคราะห์มูลค่าเพิ่มอาจจะเป็นแหล่งของความคลาดเคลื่อน เนื่องจากการประเมินคุณภาพการจัดการศึกษาโดยส่วนมากเป็นการประเมินขนาดกลุ่มใหญ่ (Large scale assessment) เครื่องมือวัดผลระดับสากลส่วนใหญ่จะนิยมใช้แบบสอบมาตรฐาน ได้แก่แบบสอบทุกชนิดมีโอกาสเกิดความคลาดเคลื่อนขึ้นคะแนนจากการทำแบบสอบจึงอาจจะสะท้อนถึงผลรวมของความรู้ที่เกิดจากความสามารถของผู้สอบและโอกาสส่วนอื่นที่ช่วยให้ผู้สอบตอบถูก นอกจากนี้สิ่งหนึ่งที่หลีกเลี่ยงไม่ได้ของการเกิดความคลาดเคลื่อนคือเป็นคะแนนสอบที่เกิดจากทักษะส่วนอื่น เช่นความแตกต่างของหลักสูตรของแต่ละสถานศึกษามีผลการทำข้อสอบของนักเรียน (Emenogu & Childs, 2005) ความคลาดเคลื่อนจากการวัดที่เกิดจากกระบวนการสร้างแบบสอบและส่งผลต่อความตรงของแบบสอบ จะทำให้การแปลผลคะแนนสอบมีโอกาสที่จะได้รับผลกระทบจากแหล่งความคลาดเคลื่อนต่างๆ ส่งผลให้เกิดความผันแปรของคะแนนสอบมากกว่าความสามารถของผู้สอบที่วัดด้วยแบบสอบ แต่ถ้าแบบสอบวัดได้สมบูรณ์ทุกอย่าง คะแนนทุกคะแนนจึงจะสามารถเชื่อถือได้และสามารถแปลผลการวัดนั้นได้อย่างมีความตรง แต่ถ้ามีแหล่งของความผันแปรของความคลาดเคลื่อนที่ไม่สามารถควบคุมได้ คะแนนที่ได้จากการทำแบบสอบจะก่อให้เกิดการเบี่ยงเบนไปจากความสามารถที่แท้จริงของผู้สอบ เพื่อให้การสอบมีความเท่าเทียมกับผู้สอบทุกคน และการแปลความหมายสามารถแปลผลได้ตรงโดยปราศจากตัวแปรอื่นๆ มาแทรกซ้อน จึงควรมีการตรวจสอบเพื่อเกิดความยุติธรรมต่อกลุ่มผู้สอบ ทั้งนี้ความไม่ยุติธรรมจะเกิดขึ้นได้เมื่อกลุ่มย่อยของผู้สอบสองกลุ่มที่มีระดับความสามารถเท่ากัน แต่คะแนนที่ได้จากการสอบไม่จัดให้อยู่ในตำแหน่งเดียวกัน ดังนั้นเพื่อให้เกิดความมั่นใจในแบบสอบที่ใช้ประเมินสำหรับกลุ่มผู้สอบทุกกลุ่มที่จำนวนมากและมาจากกลุ่มที่มีคุณลักษณะต่างกันหลายกลุ่ม จึงควรจะมีการพิจารณาเกี่ยวกับตรวจสอบข้อสอบทุกข้อที่อยู่ในแบบสอบว่ามีการโอนเอียงหรือลำเอียงให้กับกลุ่มผู้สอบกลุ่มใดกลุ่มหนึ่งในการตอบข้อสอบได้ถูกต้องมากกว่าผู้สอบอีกกลุ่มหนึ่งหรือไม่ Camilli & Shepard (1994) ได้กล่าวถึงความลำเอียงนี้ว่าเป็นความไม่ตรงหรือความคลาดเคลื่อนอย่างเป็นระบบในแบบสอบที่วัดของเฉพาะกลุ่มใดกลุ่มหนึ่ง และในการระบุข้อสอบในแบบสอบที่แสดงถึงความลำเอียงต่อกลุ่มผู้สอบกลุ่มใดกลุ่มหนึ่ง Angoff (1993) ได้นิยามการทำหน้าที่ต่างกันของข้อสอบ (Differential item functioning: DIF) ซึ่งข้อสอบที่ทำหน้าที่ต่างกันจะมีความลำเอียงขึ้น ถ้าความสามารถของแต่ละคนเท่ากันและมาจากกลุ่มที่ต่างกัน แต่ไม่มีความเท่าเทียมกันในโอกาสของการตอบข้อสอบถูกต้อง กล่าวคือถ้าข้อสอบไม่มีความยุติธรรม ผู้สอบต่างกลุ่มกันจะมีโอกาสตอบข้อสอบได้ถูกต่างกันถึงแม้จะมีระดับความสามารถเท่ากัน เช่นข้อสอบที่ใช้คำศัพท์ที่มีลักษณะเฉพาะกับผู้สอบแต่ละกลุ่มอาจจะแปลความหมายของคำศัพท์ได้แตกต่างกัน

เนื่องจากความคุ้นเคยของคำศัพท์ต่างกัน ดังนั้นคะแนนที่ได้จากการทำข้อสอบจึงเกิดจากข้อจำกัดของความเข้าใจในความหมายของคำศัพท์หรือภาษาที่ใช้ในข้อสอบ (Camilli & Shepard,1994)

แบบสอบ(Test) ประกอบด้วยข้อสอบหลายข้อ (Items) ที่ถามในเนื้อหาที่ต้องการให้ผู้สอบได้แสดงสิ่งที่รู้ตามกรอบของการวัด ดังนั้นแหล่งของความยากจึงขึ้นอยู่กับการใช้เนื้อหาในข้อสอบและอาจจะมีส่วนที่ไม่มีความตรงต่อกลุ่มผู้สอบ แต่เนื่องจากการทำแบบสอบ ผู้สอบแต่ละคนจะตอบสนองข้อสอบ นั่นคือข้อสอบแต่ละข้อ (item) ที่อยู่ในแบบสอบ (test) จะสอดแทรก (nested) อยู่ในผู้สอบแต่ละคน (person) และผู้สอบจะสอดแทรกอยู่ในสถานศึกษา (school) ซึ่งมีโครงสร้างลักษณะข้อมูลเป็นแบบลดหลั่น (Hierarchy) ดังนั้นจึงสามารถนำหลักการวิเคราะห์พหุระดับ (Multi-level analysis) มาใช้ในการวิเคราะห์ข้อสอบ เมื่อจัดข้อมูลเป็นพหุระดับหรือมีลักษณะลดหลั่น (Hierarchical) ร่วมกับโมเดลการตอบสนองข้อสอบตามแนวคิดของโมเดลของราสซ์ (Rasch Model) (Kamata, 2001) สามารถเขียนเป็นความสัมพันธ์ของโครงสร้างข้อมูลได้ดังแผนภาพที่ 1 ซึ่งในการวิเคราะห์ระดับที่ 1 เป็นระดับข้อสอบที่สอดแทรกในผู้สอบ ระดับที่ 2 เป็นระดับผู้สอบที่สอดแทรกอยู่ในสถานศึกษา และระดับที่ 3 เป็นระดับสถานศึกษา



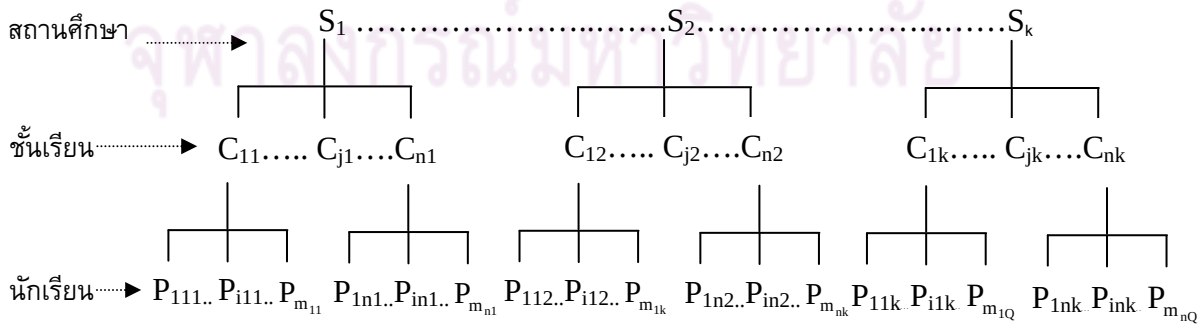
ภาพที่ 1 โครงสร้างความสัมพันธ์ของข้อมูลในการวิเคราะห์ข้อสอบแบบพหุระดับ

การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) สามารถวิเคราะห์ได้ในขั้นตอนของการวิเคราะห์ข้อสอบแบบพหุระดับ ด้วยการประยุกต์ใช้โมเดลที่วิเคราะห์ข้อมูลที่มีโครงสร้างต่างระดับ ตามลักษณะของข้อมูลด้วยโมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น (Hierarchical Generalized Linear Model: HGLM) ซึ่งโมเดล HGLM ได้เริ่มมีการศึกษาเพื่อนำมาใช้ในการวิเคราะห์ข้อสอบโดย Akihito Kamata (1998) และได้เริ่มมีการประยุกต์ใช้เพื่อการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบมากขึ้นทั้งในข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า เช่น Binici, 2007; Kamata, Chaimongkol, Genc & Bilir, 2005; Cheong, 2006; Williams & Beretvas, 2006; Vaughn, 2006 และ Chaimongkol, 2005

ผลของข้อสอบที่ทำหน้าที่ต่างกันสามารถอธิบายได้จากความไม่สอดคล้องกันของกลุ่มในการแสดงออกจากการทำข้อสอบที่เป็นตัวสะท้อนความรู้ที่แท้จริงและความแตกต่างของประสบการณ์ในโครงสร้างที่สำคัญ ดังนั้นถ้ามีความคลาดเคลื่อนเกิดขึ้นจากการทำความเข้าใจและการได้มาของคะแนนจากการทำแบบสอบมาตรฐานใดๆ อาจจะนำไปสู่การแปลความหมายจากผล

คะแนนสอบที่ไม่ถูกต้อง การนำผลคะแนนจากแบบสอบที่ใช้วัดผลการจัดการเรียนรู้ระดับสากลหรือระดับนานาชาติสำหรับการประเมินขนาดกลุ่มใหญ่ (Large scale assessment) เช่นการประเมินผลนักเรียนระดับนานาชาติในโครงการ PISA ควรคำนึงถึงความถูกต้องของผลการสอบหรือมีความตรง และปราศจากข้อสอบที่มีความลำเอียงให้กับนักเรียนกลุ่มใดกลุ่มหนึ่ง เพื่อทำให้การรายงานผลการประเมินคุณภาพการจัดการศึกษาสามารถเปรียบเทียบคุณภาพในระดับสถานศึกษา ระดับจังหวัด และระดับเขตพื้นที่การศึกษาได้ถูกต้องมากยิ่งขึ้น ถ้าข้อสอบในแบบสอบมีความโน้มเอียงต่อนักเรียนกลุ่มใดกลุ่มหนึ่งที่มีความแตกต่างกันในตัวแปรที่เป็นคุณลักษณะของผู้สอบหรือคุณลักษณะของสถานศึกษา จะส่งผลให้นักเรียนกลุ่มใดกลุ่มหนึ่งสามารถทำข้อสอบได้ถูกมากกว่าหรือมีโอกาสในการตอบข้อสอบได้ถูกต่างกัน ดังนั้นก่อนการใช้คะแนนจากแบบสอบ จึงควรตรวจสอบข้อสอบที่ทำหน้าที่ต่างกัน เพื่อให้ผลการตอบในข้อสอบข้อดังกล่าวเกิดความยุติธรรมกับผู้สอบอย่างเท่าเทียมกัน

เมื่อพิจารณาโครงสร้างของข้อมูลในระบบทางการศึกษาที่มีความเกี่ยวข้องกับบุคคลหลายฝ่ายได้แก่ นักเรียน ครู และสถานศึกษา ซึ่งข้อมูลมีลักษณะเป็นโครงสร้างแบบลดหลั่นเช่นเดียวกันกล่าวคือบุคคลที่สอดแทรกอยู่ในหน่วยงาน (Raudenbush & Bryk, 2002) ดังแสดงในแผนภาพที่ 2 ดังนั้นจึงควรพิจารณาข้อมูลที่มีลักษณะเป็นโครงสร้าง ที่สามารถจัดเป็นระดับต่างๆ ของที่หน่วยศึกษา เช่น นักเรียนที่สอดแทรกอยู่ในสถานศึกษา และสถานศึกษาสอดแทรกอยู่ในจังหวัดหรือเขตพื้นที่ เป็นต้น ความสัมพันธ์ของตัวแปรแต่ละระดับที่อยู่ในระดับเดียวกันและข้ามระดับกัน ตัวแปรที่อยู่ระดับสูงกว่ามีแนวโน้มที่จะส่งผลทางตรงหรือทางอ้อม ซึ่งอาจจะเป็นตัวแปรเชิงนโยบาย มีหน้าที่กำหนดแนวทางมาตรฐานต่อตัวแปรที่อยู่ระดับต่ำกว่า (ศิริชัย กาญจนวาสี, 2550; Raudenbush & Bryk, 2002) กล่าวคือตัวแปรที่อยู่ในระดับล่างจะมีความสัมพันธ์ซึ่งกันและกัน และได้รับอิทธิพลร่วมกันจากตัวแปรที่อยู่ในระดับเหนือขึ้นไป ตัวอย่างเช่น การศึกษาของ Berliner ที่พบว่าปัจจัยที่ส่งผลต่อประสิทธิผลของสถานศึกษาและประสิทธิผลของห้องเรียน อาจเป็นตัวแปรชุดเดียวกันแต่อยู่ต่างระดับกัน สภาวะของสถานศึกษาที่มีประสิทธิผลกับห้องเรียนที่มีประสิทธิผล มีลักษณะคล้ายกันมาก เช่นคุณลักษณะทางด้านบรรยากาศ สภาพแวดล้อมทางวิชาการ กระบวนการนิเทศติดตามผล ความคาดหวังต่อผลสัมฤทธิ์ทางการเรียน (อ้างถึงใน ศิริชัย กาญจนวาสี, 2550)



ภาพที่ 2 โครงสร้างความสัมพันธ์ของข้อมูลในสถานศึกษาที่มีลักษณะลดหลั่น

เมื่อพิจารณาโครงสร้างข้อมูลที่มีความสัมพันธ์กันในแผนภาพที่ 1 และแผนภาพที่ 2 พบว่ามีความสัมพันธ์กันอย่างต่อเนื่องในข้อมูลแต่ละระดับ กล่าวคือข้อสอบสอดแทรกอยู่ในผู้สอบหรือนักเรียน และนักเรียนสอดแทรกอยู่ในชั้นเรียน และชั้นเรียนสอดแทรกอยู่ในสถานศึกษา ซึ่งทำให้ข้อสอบที่ทำหน้าที่ต่างกันอาจจะเป็นผลที่มีความสัมพันธ์ในแต่ละระดับของข้อมูล ตัวอย่างเช่น Linn & Hirnisch (1981) พบว่าการทำหน้าที่ต่างกันของข้อสอบมีความเกี่ยวข้องกับกลุ่มเชื้อชาติที่แตกต่างกันที่อาจจะได้รับรูปแบบการสอนที่แตกต่างกัน และรูปแบบการสอนอาจจะมีลักษณะที่ผันแปรในระหว่างซึ่งมีส่วนเกี่ยวข้องกับลักษณะเฉพาะของสถานศึกษา นอกจากนี้การใช้คำศัพท์ในข้อสอบอาจจะเป็นปัญหาต่อกลุ่มผู้สอบกลุ่มใดกลุ่มหนึ่ง เช่นโจทย์ปัญหาเกี่ยวกับพีชคณิต ถ้าใช้คำศัพท์ที่มีความคุ้นเคยต่อสมาชิกบางกลุ่มมากกว่าอีกกลุ่มหนึ่งจะมีผลให้เกิดความโอนเอียงต่อการแสดงออกของผู้สอบสองกลุ่มต่างกัน และคุณลักษณะของนักเรียนอาจจะเป็นปฏิสัมพันธ์กับคุณลักษณะของข้อสอบซึ่งส่งผลต่อความสามารถในการแสดงออกของแบบสอบฉบับหนึ่ง ทำให้คุณลักษณะของข้อสอบอาจจะมีการแปรเปลี่ยนค่าได้ตามความแตกต่างของกลุ่มที่มีลักษณะต่างกัน เช่นเศรษฐานะหรือระดับความกังวล Cheong (2006) ได้ศึกษาถึงบริบทของสถานศึกษาเช่นโอกาสในการสอน แหล่งของระหว่างกลุ่มที่มีความต่างกันในหน้าที่ของข้อสอบ โดยการควบคุมคุณลักษณะเฉพาะของสถานศึกษา พบว่า DIF มีสาเหตุจากตัวแปรของบริบทสถานศึกษาและความแตกต่างด้านเชื้อชาติของนักเรียน และ Chaimongkol (2005) พบว่าแหล่งของการเกิด DIF ของตัวแปรระดับสถานศึกษา ได้แก่รายวิชาที่สอนนอกห้องเรียน จำนวนผู้สนับสนุน ทีมงานของผู้บริการ และตัวแปรบริบทของสถานศึกษา เช่น อัตราความคงที่ของนักเรียน อัตราการพักเรียนของนักเรียน

ดังนั้นการทำหน้าที่ต่างกันของข้อสอบที่เกิดขึ้นจากมีแหล่งหรือปัจจัยจากตัวแปรคุณลักษณะในแต่ละระดับที่นำมาวิเคราะห์ จึงส่งผลต่อการทำคะแนนสอบของผู้สอบแต่ละกลุ่ม ถ้าแบบสอบประกอบด้วยข้อสอบที่ทำหน้าที่ต่างกัน ย่อมทำให้เกิดความไม่เป็นธรรมแก่ผู้สอบบางกลุ่ม และส่งผลต่อการแปลความหมายคะแนนสอบเพื่อนำผลไปใช้ต่อไป กล่าวคือถ้าในแบบสอบฉบับหนึ่งประกอบด้วยข้อสอบที่ทำหน้าที่ต่างกัน และไม่มีการตรวจสอบและตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป คะแนนที่ได้จากการทำแบบสอบนักเรียนย่อมนำไปสู่การแปลความหมายของคะแนนที่คลาดเคลื่อน และส่งผลต่อคะแนนเฉลี่ยของชั้นเรียนและสถานศึกษาเป็นลำดับต่อไป อันนำไปสู่การแปลความหมายที่คลาดเคลื่อนของการประเมินคุณภาพการศึกษาของสถานศึกษาที่ประกอบด้วยนักเรียนที่ทำข้อสอบเหล่านั้น ในทางตรงข้าม ถ้ามีการตรวจสอบและตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป ก่อนนำคะแนนจากการทำแบบสอบไปใช้ในการแปลผล คะแนนที่ได้ย่อมมีความถูกต้องน่าเชื่อถือ และมีความเป็นธรรมกับกลุ่มผู้สอบ และเป็นผลการประเมินของสถานศึกษาที่ถูกต้องมากที่สุด

สิ่งที่ควรคำนึงถึงอีกประการหนึ่ง เพื่อให้ผลการประเมินคุณภาพการจัดการศึกษามีความถูกต้องและเป็นธรรมสำหรับสถานศึกษาคือการเลือกโมเดลการวิเคราะห์มูลค่าเพิ่มที่เหมาะสม และการกำหนดโมเดลการวิเคราะห์ ทั้งนี้เนื่องจากประมาณค่าผลสัมฤทธิ์ที่แสดงถึงผลการจัดการศึกษาของสถานศึกษาอาจมีผลต่างกันเนื่องจากการกำหนดโมเดลการประเมินที่ต่างกัน ซึ่งบางครั้ง

ผลลัพธ์ที่ได้อาจมีค่าต่ำกว่าผลที่เป็นจริงของสถานศึกษา รวมทั้งยังมีข้อจำกัดในการกำหนดตัวแปรภูมิหลังของนักเรียนที่ทำได้อย่างครบถ้วน นอกจากนี้ข้อมูลที่ได้จากการวัดด้วยแบบสอบที่ขาดความถูกต้อง และน่าเชื่อถืออาจมีผลต่อการประมาณค่าในระดับบุคคล Raudenbush and Willims (1996 cited in Sheerens and Bosker, 1997) และจากผลการศึกษาของนักวิชาการหลายท่าน เช่น Tekwe, Carter, Ma, Algina, Lucas, Roth, Ariet, Fisher and Resnick, 2004 ที่มีการศึกษาเปรียบเทียบโมเดลการวิเคราะห์เพื่อการประเมินคุณภาพการจัดการศึกษายังพบว่าไม่มีโมเดลใดที่แสดงให้เห็นว่ามีความถูกต้องและเพียงพอในการนำมาใช้ในสภาพแวดล้อมที่ต่างกันและมีการได้เปรียบเสียเปรียบกัน การประมาณค่าในโมเดลที่ไม่ได้รวมตัวแปรระดับนักเรียนและสถานศึกษาจะมีความลำเอียงให้กับสถานศึกษาที่มีกลุ่มประชากรมีคุณลักษณะที่ได้เปรียบกว่า และในบางกรณีสามารถเลือกใช้โมเดลอย่างง่ายในการวิเคราะห์เนื่องจากเป็นโมเดลที่ไม่ซับซ้อน แต่ผลการทดสอบพบว่ามีความสัมพันธ์กันสูงกับโมเดลที่มีความซับซ้อน

ผู้วิจัยจึงมีความสนใจประเมินคุณภาพการจัดการศึกษาของสถานศึกษา เพื่อให้เกิดความยุติธรรมสำหรับนักเรียนและสถานศึกษาที่มีคุณลักษณะพื้นฐานต่างกัน และเพื่อให้ได้ผลการประเมินที่มีความถูกต้องเชื่อถือได้ โดยใช้ข้อมูลจากการศึกษานานาชาติตามโครงการประเมินผล PISA ปี 2006 ในวิชาวิทยาศาสตร์ เนื่องจากโครงการประเมินผล PISA เป็นการรวบรวมข้อมูลจากนักเรียนจำนวนมาก ซึ่งคุณลักษณะพื้นฐานที่แตกต่างกันในนักเรียนและสถานศึกษา อาจจะเป็นปัจจัยที่ทำให้เกิดความแตกต่างในผลการประเมินผลของนักเรียน แต่จากรายงานผลประเมินในโครงการดังกล่าวเป็นเพียงการใช้ผลจากคะแนนที่ได้จากการทำแบบสอบ และไม่คำนึงถึงความแตกต่างของคุณลักษณะพื้นฐานในระดับนักเรียน ระดับสถานศึกษา ตลอดจนถึงในระดับประเทศ ในการวิจัยครั้งนี้จะคำนึงถึงลักษณะข้อมูลที่เป็นโครงสร้างแบบลดหลั่นด้วยการวิเคราะห์พหุระดับ ตลอดจนคำนึงการนำผลคะแนนสอบจากแบบสอบที่อาจจะมีความคลาดเคลื่อนเกิดขึ้นเนื่องจากแบบสอบมีข้อสอบที่ทำหน้าที่ต่างกัน จึงประยุกต์ใช้การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (Differential item functioning: DIF) ด้วยโมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น (HGLM) เพื่อให้ได้คะแนนจากแบบสอบสำหรับประเมินการรู้เรื่องด้านวิทยาศาสตร์มีความถูกต้องมากยิ่งขึ้น ดังนั้นโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาที่ศึกษาในครั้งนี้ ผู้วิจัยจึงสนใจใน 2 ประเด็นคือ ประเด็นแรกเป็นคุณภาพของแบบสอบประเมินการรู้เรื่องด้านวิทยาศาสตร์ที่มีความแตกต่างกัน 2 ลักษณะ คือไม่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันและมีการตัดข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบประเมินการรู้เรื่องด้านวิทยาศาสตร์ ทั้งนี้ในการตัดข้อสอบที่ทำหน้าที่ต่างกันจะตัดเฉพาะข้อที่จำเป็นโดยพิจารณาตามเกณฑ์การตัดข้อสอบ และรักษาโครงสร้างเดิมของแบบสอบไว้ เพื่อให้แบบสอบมียังคงมีความตรงในการประเมินและมีการเปลี่ยนแปลงความเที่ยงของแบบสอบน้อยที่สุด ประเด็นที่สองเป็นโมเดลที่ใช้ในการวิเคราะห์มูลค่าเพิ่ม(Value-added analysis) ที่มีความแตกต่างกัน 2 ลักษณะคือโมเดลที่มีการวิเคราะห์โดยไม่มีการนำตัวแปรต้นที่เป็นคุณลักษณะพื้นฐานของนักเรียนและสถานศึกษามาใช้ในการวิเคราะห์และโมเดลที่มีการวิเคราะห์โดยนำตัวแปรต้นที่เป็นคุณลักษณะพื้นฐานของนักเรียนและสถานศึกษามาใช้ในการวิเคราะห์ ดังนั้นโมเดลในการประเมินคุณภาพการ

จัดการศึกษาของสถานศึกษาขั้นพื้นฐาน ในการวิจัยครั้งนี้จึงประกอบด้วย 4 โมเดล คือ **โมเดล 1** เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออก และวิเคราะห์มูลค่าเพิ่มโดยที่ยังไม่มีการปรับแก้คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา **โมเดล 2** เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออกและวิเคราะห์มูลค่าเพิ่มโดยมีการปรับแก้คะแนนตัวแปรระดับนักเรียนและสถานศึกษา **โมเดล 3** เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกและวิเคราะห์มูลค่าเพิ่มโดยที่ยังไม่มีการปรับแก้คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา และ**โมเดล 4** เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกและวิเคราะห์มูลค่าเพิ่มโดยมีการปรับแก้คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา พร้อมทั้งเปรียบเทียบโมเดลการวิเคราะห์ที่สามารถใช้ประเมินคุณภาพการศึกษาได้อย่างถูกต้องและมีความเป็นธรรมสำหรับนักเรียน สถานศึกษา และเป็นแนวทางการประเมินคุณภาพการจัดการศึกษาโดยคำนึงถึงปัจจัยพื้นฐานที่อาจมีความแตกต่างกันในแต่ละสถานศึกษา ทำให้ผลการประเมินมีความถูกต้องน่าเชื่อถือและมีความยุติธรรมมากที่สุด

คำถามการวิจัย

1. แบบสอบประเมินการรู้เรื่องด้านวิทยาศาสตร์ที่มีข้อสอบที่ทำหน้าที่ต่างกันมีผลต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานหรือไม่ เพียงใด เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน
2. ตัวแปรคุณลักษณะที่แตกต่างของนักเรียนและสถานศึกษาขั้นพื้นฐานมีผลต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานหรือไม่ เพียงใด เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน
3. โมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ที่แตกต่างกันมีผลต่อคุณภาพการจัดการศึกษาของสถานศึกษาขั้นพื้นฐานหรือไม่ เพียงใด และเมื่อเปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาที่ต่างกันแล้ว โมเดลการประเมินคุณภาพการจัดการศึกษาที่มีคุณภาพที่สุดควรมีลักษณะโมเดลอย่างไร

วัตถุประสงค์ของการวิจัย

มีจุดมุ่งหมายหลักเพื่อประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน โดยมีวัตถุประสงค์เฉพาะดังนี้

1. เพื่อศึกษาผลของข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบประเมินการรู้เรื่องด้านวิทยาศาสตร์ต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานเมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน

2. เพื่อศึกษาผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาขั้นพื้นฐานต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน

3. เพื่อเปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ต่อคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน

สมมติฐานการวิจัย

จากการศึกษาค้นคว้าเอกสารงานวิจัย ผู้วิจัยพบว่าการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาควรเลือกใช้โมเดลการประเมินคุณภาพการจัดการศึกษาที่มีความถูกต้อง เชื่อถือได้ และมีความเป็นธรรม แต่การประเมินคุณภาพการจัดการศึกษาที่ผ่านมาโดยส่วนใหญ่เป็นการใช้คะแนนจากการทำแบบสอบเพื่อแสดงถึงผลสัมฤทธิ์ผลสัมฤทธิ์การดำเนินการของสถานศึกษา (Postlethwaite, 2004) แต่คะแนนที่ได้จากการทำแบบสอบมีโอกาสที่จะเกิดความคลาดเคลื่อน เนื่องจากการทำหน้าที่ต่างกันของข้อสอบในแบบสอบ จากผลการวิจัยของนักการศึกษาหลายท่าน เช่น Linn & Hirnisch (1981) พบว่าการทำหน้าที่ต่างกันของข้อสอบเกิดจากกลุ่มเชื้อชาติที่แตกต่างกันที่อาจจะได้รับรูปแบบการสอนที่แตกต่างกัน ซึ่งรูปแบบการสอนอาจจะแตกต่างกันในแต่ละสถานศึกษา ผลการศึกษาของ Cheong (2006) พบว่าบริบทของสถานศึกษาระหว่างกลุ่มที่มีความต่างกันในการทำหน้าที่ของข้อสอบ และ Chaimongkol (2005) พบว่าแหล่งของการเกิด DIF ของตัวแปรระดับสถานศึกษา ได้แก่รายวิชาที่สอนนอกห้องเรียน จำนวนผู้สนับสนุน ทีมงานผู้บริการ และตัวแปรบริบทของสถานศึกษา ดังนั้นข้อสอบที่ทำหน้าที่ต่างกันจะมีแหล่งเนื่องจากตัวแปรในแต่ละระดับที่นำมาวิเคราะห์ และถ้าแบบสอบประกอบด้วยข้อสอบที่ทำหน้าที่ต่างกันสำหรับผู้สอบต่างกลุ่มกัน ย่อมทำให้เกิดความไม่เป็นธรรมแก่ผู้สอบบางกลุ่ม

นอกจากนี้คุณลักษณะของนักเรียนและสถานศึกษาที่มีความแตกต่างหรือไม่เท่าเทียมกันอาจจะส่งผลต่อการประเมินคุณภาพการจัดการศึกษา มีนักการศึกษาหลายท่านได้เสนอแนะการเลือกโมเดลมูลค่าเพิ่มเพื่อการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา เช่น Tekwe, Carter, Ma, Algina, Lucas, Roth, Ariet, Fisher, and Resnick, 2004 ที่เปรียบเทียบโมเดลการวิเคราะห์เพื่อการประเมินคุณภาพการจัดการศึกษาพบว่าโมเดลการประเมินทำให้ผลการประเมินมีความแตกต่างกันระหว่างโมเดลที่มีการความตัวแปรนักเรียนและสถานศึกษาเข้ามาศึกษา ทั้งนี้การเลือกโมเดลสามารถคำนึงถึงความถูกต้อง และไม่มีการได้เปรียบเสียเปรียบกัน และในบางกรณีสามารถเลือกใช้โมเดลอย่างง่ายในการวิเคราะห์เนื่องจากเป็นโมเดลที่ไม่ซับซ้อน จากผลการศึกษาเอกสารผู้วิจัยจึงกำหนดสมมติฐานการวิจัย 3 ข้อ ดังนี้

1. การประเมินคุณภาพการจัดการศึกษาของสถานศึกษาขั้นพื้นฐานด้วยโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานที่มีการตัดและไม่ตัดข้อสอบที่ทำหน้าที่ต่างกัน ทำให้กลุ่มคุณภาพและระดับคุณภาพการจัดการศึกษาแตกต่างกัน

2. การประเมินคุณภาพการจัดการศึกษาของสถานศึกษาขั้นพื้นฐานด้วยโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานที่มีการควบคุมและไม่ควบคุมตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาขั้นพื้นฐาน ทำให้กลุ่มคุณภาพและระดับคุณภาพการจัดการศึกษาแตกต่างกัน

3. โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน ทำให้ผลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานแตกต่างกัน โดยโมเดลที่มีคุณภาพที่สุดเป็นโมเดลที่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปและมีการปรับแก้คะแนนด้วยการควบคุมอิทธิพลของคุณลักษณะของนักเรียนและสถานศึกษา

ขอบเขตของการวิจัย

1) ข้อมูลที่ใช้ในการวิจัยครั้งนี้ เป็นข้อมูลทุติยภูมิจากฐานข้อมูล (data base) ของโครงการประเมินผลนักเรียนระดับนานาชาติหรือ PISA ปี 2006 กลุ่มตัวอย่างของโครงการเป็นนักเรียนที่มีอายุ 15 ปี ซึ่งเป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 2 – 3 จำนวน 6,192 คน ในสถานศึกษา 212 แห่ง ที่ครอบคลุมสถานศึกษาในสังกัดของสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน สำนักงานการศึกษาเอกชน สังกัดของกรุงเทพมหานคร สำนักงานคณะกรรมการการอุดมศึกษา สำนักบริหารการศึกษาท้องถิ่น และสำนักงานการอาชีวศึกษา

2) แบบสอบที่ใช้ในการศึกษา เป็นแบบสอบประเมินการรู้เรื่อง(Literacy)ในโครงการประเมินผล PISA 2006 ซึ่งในการวิจัยครั้งนี้ศึกษาเฉพาะวิชาวิทยาศาสตร์ ประกอบด้วยข้อสอบแบบเลือกตอบ และแบบเขียนคำตอบ มีจำนวนข้อสอบ 108 ข้อ จัดเป็นแบบสอบย่อย 13 ฉบับ เมื่อจำแนกตามการตรวจให้คะแนน จัดเป็นข้อสอบที่ตรวจให้คะแนนสองค่ามีจำนวน 103 ข้อ และข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่ามีจำนวน 5 ข้อ

3) วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) สำหรับการวิจัยในครั้งนี้ประยุกต์ใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น (Hierarchical Generalized Linear Model: HGLM) ที่มีการวิเคราะห์พหุระดับ โดยตัวแปรระดับที่ 1 เป็นระดับข้อสอบ ระดับที่ 2 เป็นระดับผู้สอบ และระดับที่ 3 เป็นระดับสถานศึกษา

4) วิธีการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาขั้นพื้นฐานสำหรับการวิจัยในครั้งนี้ ประยุกต์ใช้การวิเคราะห์มูลค่าเพิ่ม (Value-Added Analysis) โดยตัวแปรระดับที่ 1 เป็นระดับนักเรียน ประกอบด้วยคุณลักษณะพื้นฐานของนักเรียน ได้แก่ ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา เพศ เศรษฐานะของครอบครัว แหล่งทรัพยากรที่บ้าน ความมั่งคั่งของครอบครัว ความสนใจในการเรียนรู้วิทยาศาสตร์ ความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์ ความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์ ความเพลิดเพลินทางวิทยาศาสตร์ แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์ การเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์ การใช้เวลาเรียนพิเศษวิทยาศาสตร์นอกสถานศึกษา และการใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง และตัวแปรระดับที่ 2 เป็นระดับสถานศึกษา ประกอบด้วยคุณลักษณะของสถานศึกษา ได้แก่ ขนาด

สถานศึกษา ประเภทสถานศึกษา สังกัดของสถานศึกษา ตำแหน่งที่ตั้งของสถานศึกษา ขนาดห้องเรียน แรงกดดันของผู้ปกครองต่อสถานศึกษา การขาดแคลนครูวิทยาศาสตร์ การขาดแคลนครูที่มีคุณวุฒิ งบประมาณของสถานศึกษา ดัชนีการขาดแคลนครู ดัชนีทรัพยากรการเรียน สัดส่วนนักเรียนต่อครู ดัชนีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์ ดัชนีความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากร และดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล

5) โมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาขั้นพื้นฐาน ในการวิจัยครั้งนี้พิจารณาใน 2 ประเด็น ประเด็นแรกเป็นคุณภาพของแบบสอบที่มีผลต่อความถูกต้องจากคะแนนที่ได้จากแบบสอบ จำแนกคุณภาพของแบบสอบประเมินการรู้เรื่องด้านวิทยาศาสตร์เป็น 2 ลักษณะคือ 1) แบบสอบที่ไม่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ และ 2) แบบสอบที่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ ทั้งนี้ในการตัดข้อสอบที่ทำหน้าที่ต่างกันจะตัดเฉพาะข้อที่จำเป็น โดยพิจารณาตามเกณฑ์การตัดข้อสอบ และรักษาโครงสร้างเดิมของแบบสอบไว้ เพื่อให้แบบสอบมียังคงมีความตรงในการประเมินและมีการเปลี่ยนแปลงความเที่ยงของแบบสอบน้อยที่สุด ประเด็นที่สอง เป็นโมเดลที่ใช้ในการวิเคราะห์มูลค่าเพิ่ม จำแนกเป็น 2 รูปแบบคือ 1) โมเดลที่มีปรับแก้คะแนน โดยการควบคุมปัจจัยอื่นที่เป็นคุณลักษณะของนักเรียนและสถานศึกษา และ 2) โมเดลที่ไม่มีการปรับแก้คะแนนจากตัวแปรของปัจจัยอื่นที่เป็นคุณลักษณะของนักเรียน ห้องเรียนและสถานศึกษา จากประเด็นการพิจารณาทั้งสองจึงกำหนดเป็นโมเดลการประเมินคุณภาพการจัดการศึกษา 4 โมเดล ดังนี้

5.1) โมเดล 1 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ และวิเคราะห์มูลค่าเพิ่มโดยที่ไม่มีการปรับแก้ (Unadjusted) คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา

5.2) โมเดล 2 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ และวิเคราะห์มูลค่าเพิ่มโดยที่มีการปรับแก้ (Adjusted) คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา

5.3) โมเดล 3 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ และวิเคราะห์มูลค่าเพิ่มโดยที่ไม่มีการปรับแก้ (Unadjusted) คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา

5.4) โมเดล 4 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ และวิเคราะห์มูลค่าเพิ่มโดยที่มีการปรับแก้ (Adjusted) ด้วยตัวแปรระดับนักเรียนและสถานศึกษา

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

นิยามศัพท์เฉพาะ

การรู้เรื่องวิทยาศาสตร์ (Scientific Literacy) หมายถึงความรู้สองแบบ คือความรู้วิทยาศาสตร์" และ "ความรู้ทางวิทยาศาสตร์" โดยความรู้วิทยาศาสตร์หมายถึงองค์ความรู้ซึ่งเป็นความรู้ของโลกธรรมชาติ เช่น ฟิสิกส์ เคมี ชีววิทยา ส่วนความรู้เกี่ยวกับวิทยาศาสตร์หมายถึงความรู้ในวิธีการหรือกระบวนการหาความรู้ทางวิทยาศาสตร์ หรือวิถีทางที่นำไปสู่เป้าหมายของการได้มาซึ่งความรู้ ซึ่งผลการประเมินนี้จะใช้เป็นข้อมูลที่บ่งชี้คุณภาพการจัดการศึกษาของสถานศึกษาในการเตรียมความพร้อมนักเรียน สำหรับการใช้ชีวิตในอนาคต

แบบสอบสำหรับประเมินการรู้เรื่องวิทยาศาสตร์ หมายถึง แบบสอบวิชาวิทยาศาสตร์ซึ่งวัดความรู้ด้านวิทยาศาสตร์ตามกรอบการประเมินการรู้เรื่องวิทยาศาสตร์ในโครงการประเมินผลนักเรียนระดับนานาชาติ (Programme for International Student Assessment: PISA) ปี 2006 จำนวน 13 ฉบับ โดยการวัดความรู้ทางวิทยาศาสตร์จะครอบคลุมเนื้อหาความรู้ในวิชาวิทยาศาสร์และการใช้ความรู้เกี่ยวกับวิทยาศาสตร์ ซึ่งแบบสอบวัดในสมรรถนะทางวิทยาศาสตร์สามส่วน ได้แก่ การระบุประเด็นทางวิทยาศาสตร์ การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ การใช้ประจักษ์พยานทางวิทยาศาสตร์

คะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ หมายถึง ค่าของผลการแปลงพารามิเตอร์ความสามารถของนักเรียนให้อยู่ในรูปคะแนนสเกลของ PISA ซึ่งการประมาณค่าพารามิเตอร์ความสามารถของนักเรียนได้จากผลการวิเคราะห์พหุระดับด้วยโมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น 3 ระดับ จากผลการตอบข้อสอบในแบบสอบ 13 ฉบับ

มูลค่าเพิ่มของสถานศึกษา หมายถึง ผลต่างระหว่างคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์โดยเฉลี่ยของแต่ละสถานศึกษาที่ทำนายได้จากการควบคุมอิทธิพลของปัจจัยระดับนักเรียนและระดับสถานศึกษา กับคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์โดยเฉลี่ยของสถานศึกษาที่สังเกตได้จริง ซึ่งจากผลการวิเคราะห์มูลค่าเพิ่มจะเป็นค่าส่วนที่เหลือ (residual) ของจุดตัดแกน (intercept) ด้วยหลักการวิเคราะห์ถดถอยพหุระดับ

คุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน หมายถึง คะแนนมูลค่าเพิ่มของสถานศึกษาหรือค่าส่วนที่เหลือจากการวิเคราะห์มูลค่าเพิ่มในแต่ละโมเดลการประเมินคุณภาพการจัดการศึกษา ซึ่งคุณภาพการจัดการศึกษาของสถานศึกษาจากการวิจัยในครั้งนี้แสดงถึงการดำเนินงานจัดการของแต่ละสถานศึกษาเพื่อให้นักเรียนเกิดการเรียนรู้ที่มีการควบคุมปัจจัยอื่นที่อยู่นอกเหนือการควบคุมของสถานศึกษาแล้ว

การประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน หมายถึง การตัดสินคุณค่าของสถานศึกษาขั้นพื้นฐานโดยการวิเคราะห์หาคุณภาพการจัดการศึกษาของสถานศึกษา จากการใช้ผลจากการทำแบบสอบสำหรับประเมินการรู้เรื่องด้านวิทยาศาสตร์

ตัวแปรต้นระดับนักเรียน หมายถึง คุณลักษณะของนักเรียนที่นอกเหนือจากความสามารถของผู้สอบ ได้แก่ ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา เพศ เศรษฐานะของครอบครัว แหล่งทรัพยากรที่บ้าน ความมั่งคั่งของครอบครัว ความสนใจในการเรียนรู้วิทยาศาสตร์ ความ

เชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์ ความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์ ความเพลิดเพลินทางวิทยาศาสตร์ แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์ การเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์ การใช้เวลาเรียนพิเศษวิทยาศาสตร์นอกสถานศึกษา การใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง

ตัวแปรต้นระดับสถานศึกษา หมายถึง คุณลักษณะของสถานศึกษาขั้นพื้นฐาน ได้แก่ ขนาดสถานศึกษา ประเภทสถานศึกษา สังกัดของสถานศึกษา ตำแหน่งที่ตั้งของสถานศึกษา ขนาดห้องเรียน แรงกดดันของผู้ปกครองต่อสถานศึกษา การขาดแคลนครูวิทยาศาสตร์ การขาดแคลนครูที่มีคุณวุฒิ งบประมาณของสถานศึกษา ดัชนีการขาดแคลนครู ดัชนีทรัพยากรการเรียน สัดส่วนนักเรียนต่อครู ดัชนีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์ ดัชนีความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากร ดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล

การวิเคราะห์ข้อสอบพหุระดับ หมายถึง วิธีการตรวจสอบคุณภาพของข้อสอบ โดยลักษณะข้อมูลจะมีลักษณะที่ข้อสอบสอดแทรกกับผู้สอบแต่ละคน และผู้สอบยังสอดแทรกอยู่ในสถานศึกษา การวิเคราะห์จึงจัดเป็นการวิเคราะห์ 3 ระดับ คือ ระดับที่ 1 เป็นระดับข้อสอบ ระดับที่ 2 เป็นระดับผู้สอบ และระดับที่ 3 เป็นระดับสถานศึกษา

การทำหน้าที่ต่างกันของข้อสอบ (Differential item functioning: DIF**)** หมายถึง ลักษณะของข้อสอบที่ทำให้ผู้สอบต่างกลุ่มกันที่มีระดับความสามารถของผู้สอบเท่ากัน แต่มีโอกาสในการตอบข้อสอบได้ถูกต้องแตกต่างกัน ในการวิจัยนี้ใช้การวิเคราะห์ข้อสอบที่ทำหน้าที่ต่างกันโดยใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น และตรวจสอบจากคุณลักษณะของนักเรียน 5 ตัวแปรคือ ตัวแปรความเป็นเพศชาย การเรียนพิเศษ เศรษฐานะของครอบครัว แหล่งทรัพยากรที่บ้าน และความมั่งคั่งของครอบครัว

โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น (Hierarchical Generalized Linear Model: HGLM) หมายถึง รูปแบบการวิเคราะห์ข้อมูลเชิงเส้นตรงทั่วไปที่มีการประยุกต์ใช้กับการวิเคราะห์ตามโมเดลเชิงเส้นตรงทั่วไป และใช้ฟังก์ชันโลจิสติกในการวิเคราะห์ระดับที่ 1 มีตัวแปรตามเป็นลอคออด (Log-odds) ของความน่าจะเป็นการตอบข้อสอบได้ถูกต้อง และมีตัวแปรอิสระมีมากกว่าหนึ่งตัวซึ่งเป็นผลการตอบข้อสอบแต่ละข้อ

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

โมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาขั้นพื้นฐาน หมายถึง โมเดลที่ใช้ในการวิเคราะห์หาคะแนนมูลค่าเพิ่มของสถานศึกษาขั้นพื้นฐาน โดยการประยุกต์ใช้การทำหน้าที่ต่างกันของข้อสอบและการวิเคราะห์มูลค่าเพิ่ม ประกอบด้วย 4 โมเดล ได้แก่ **โมเดล 1** เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ไม่ได้ตัดข้อสอบที่ทำหน้าที่

ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ ส่วนโมเดลการวิเคราะห์มูลค่าเพิ่มเป็นโมเดลที่ยังไม่มีการปรับแก้ (Unadjusted) คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา **โมเดล 2** เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ ส่วนโมเดลการวิเคราะห์มูลค่าเพิ่มเป็นโมเดลที่มีการปรับแก้ (Adjusted) คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา **โมเดล 3** เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ ส่วนโมเดลการวิเคราะห์มูลค่าเพิ่มเป็นโมเดลที่ยังไม่มีการปรับแก้ (Unadjusted) คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา **โมเดล 4** เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ ส่วนโมเดลการวิเคราะห์มูลค่าเพิ่มเป็นโมเดลที่มีการปรับแก้ (Adjusted) ด้วยตัวแปรระดับนักเรียนและสถานศึกษา ทั้งนี้โมเดล 3 และ 4 จะตัดข้อสอบที่ทำหน้าที่ต่างกันออกเฉพาะข้อที่จำเป็น โดยพิจารณาตามเกณฑ์การตัดข้อสอบ และรักษาโครงสร้างเดิมของแบบสอบไว้เพื่อให้แบบสอบมียังคงมีความตรงในการประเมินและมีการเปลี่ยนแปลงความเที่ยงของแบบสอบน้อยที่สุด

ประโยชน์ที่คาดว่าจะได้รับ

1) ได้โมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาขั้นพื้นฐานในวิชาวิทยาศาสตร์ตามโครงการประเมิน PISA 2006 ที่ให้ผลการประเมินมีความเป็นธรรมและถูกต้องมากที่สุด เพื่อให้ข้อสารสนเทศมีความหลากหลายและเป็นประโยชน์ต่อการวางแผนการดำเนินของสถานศึกษา ตลอดจนการปรับปรุงวิธีการดำเนินงานของสถานศึกษาเพื่อส่งเสริมให้นักเรียนมีผลการเรียนรู้ด้านวิทยาศาสตร์อยู่ในระดับสูงขึ้น

2) ได้แนวทางในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ โดยใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น เพื่อหาแหล่งของการทำหน้าที่ต่างกันของข้อสอบทั้งในตัวแปรระดับนักเรียนและใช้ตัวแปรระดับสถานศึกษาอธิบายแหล่งของการทำหน้าที่ต่างกันของข้อสอบ

3) ได้ทราบถึงอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาตามที่กำหนดในโมเดลการวิเคราะห์ ซึ่งเป็นประโยชน์ต่อการวางแผน กำหนดแนวทางเพื่อส่งเสริมหรือลดอิทธิพลของตัวแปรที่มีผลต่อคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน

4) ได้ทราบผลของการทำหน้าที่ต่างกันของข้อสอบ และปัจจัยที่ทำให้นักเรียนมีโอกาสในการตอบข้อสอบถูกได้ต่างกัน เช่น เพศ การเรียนพิเศษของนักเรียน เศรษฐานะของครอบครัว ตลอดจนตัวแปรระดับสถานศึกษาที่นำมาใช้อธิบายแหล่งของการทำหน้าที่ต่างกันของข้อสอบ เช่น ขนาดของสถานศึกษา การจัดกิจกรรมการเรียนรู้ เป็นต้น เพื่อใช้เป็นแนวทางในพัฒนาการจัดการเรียนรู้ การเลือกใช้แหล่งการเรียนรู้ การจัดกิจกรรมเสริมการเรียนรู้ เป็นต้น

5) หน่วยงานที่เกี่ยวข้องกับการจัดการศึกษาทั้งในภาครัฐ และชุมชน ได้ทราบข้อสารสนเทศจากผลการประเมินเพื่อนำไปใช้ในการตัดสินใจและกำหนดนโยบาย ตลอดจนแนวทางการส่งเสริมให้นักเรียนหรือบุตรหลานมีผลการเรียนรู้ในระดับสูงขึ้น

6) สสวท. ในฐานะหน่วยหลักที่รับผิดชอบโครงการประเมิน PISA จะได้ข้อมูลสารสนเทศในการวางแผนระดับนโยบาย เพื่อการจัดการศึกษาโดยเฉพาะด้านวิทยาศาสตร์เพื่อให้เกิดการพัฒนาที่เท่าเทียมกัน และเป็นแนวทางสำหรับสถานศึกษาที่มีข้อเสียเปรียบในการยกระดับคุณภาพการจัดการศึกษาให้สูงขึ้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

เอกสารงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้ ผู้วิจัยได้ศึกษาเอกสาร รายงานการวิจัยรวมทั้งบทความทางวิชาการทั้งในประเทศและต่างประเทศ ประมวลเป็นองค์ความรู้เกี่ยวกับการวิเคราะห์มูลค่าเพิ่ม ทฤษฎีการทดสอบโดยเฉพาะทฤษฎีการตอบสนองข้อสอบ การทำหน้าที่ต่างกันของข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า การวิเคราะห์พหุระดับ ผู้วิจัยจึงประมวลแนวคิดและทฤษฎีที่เกี่ยวข้องออกเป็น 5 ตอน ดังนี้ ตอนที่ 1 มโนทัศน์เกี่ยวกับมูลค่าเพิ่มของการศึกษา ตอนที่ 2 มโนทัศน์เกี่ยวกับทฤษฎีการตอบสนองข้อสอบและการทำหน้าที่ต่างกันของข้อสอบ ตอนที่ 3 มโนทัศน์เกี่ยวกับการวิเคราะห์พหุระดับและการประยุกต์ใช้โมเดลเชิงเส้นตรงแบบลดหลั่น ตอนที่ 4 การประเมินผลนักเรียนในระดับนานาชาติ PISA 2006 และตอนที่ 5 เอกสารงานวิจัยเกี่ยวกับการวิเคราะห์มูลค่าเพิ่มเพื่อประเมินคุณภาพการจัดการศึกษาและการสังเคราะห์สรุปเพื่อกำหนดตัวแปรที่ใช้ในการวิจัย

มโนทัศน์และแนวคิดในแต่ละตอนมีรายละเอียดดังต่อไปนี้

ตอนที่ 1 มโนทัศน์เกี่ยวกับมูลค่าเพิ่มของการศึกษา

1.1 ความหมาย

คำว่ามูลค่าเพิ่ม (value-added) ได้พัฒนามาจากเศรษฐศาสตร์ ซึ่งส่วนมากนำมาใช้เพื่ออธิบายการเพิ่มค่าในทางธุรกิจ การสนับสนุนผลผลิตหรือการบริการ มีหมายความถึงผลต่างระหว่างมูลค่าของผลผลิต (output) กับมูลค่าของปัจจัยนำเข้า (input) ส่วนทางการศึกษาส่วนมากจะนำมาใช้เพื่ออธิบายการเพิ่มคุณค่า (value) ของสถานศึกษาที่นำมาสู่มาผลการเรียนรู้ (outcome) ของนักเรียน และหมายความถึงการสนับสนุนของสถานศึกษา การดำเนินการของสถานศึกษาที่ทุ่มเทเพื่อให้เกิดการเรียนรู้ของนักเรียน (Downes & Vindurampulle, 2007)

ความหมายของคำว่ามูลค่าเพิ่ม (value-added) ในด้านการศึกษา มีความเกี่ยวข้องกับผลสำเร็จหรือผลสัมฤทธิ์ของนักเรียนที่เกิดจากการดำเนินงานของสถานศึกษา และมีการพัฒนาโมเดลมูลค่าเพิ่มหรือการประเมินมูลค่าเพิ่ม เพิ่มขึ้นอย่างมากตามการกำหนดนิยามของโมเดลหรือลักษณะเฉพาะของโมเดล นักการศึกษาได้ให้ความหมายไว้อย่างหลากหลายดังนี้

มูลค่าเพิ่ม (value-added) มีความหมายเช่นเดียวกันกับคำว่า ค่าเศษเหลือ (residual) ซึ่งเป็นส่วนของคะแนน หลังจากการนำตัวแปรหนึ่งตัวแปรหรือมากกว่าหนึ่งตัวแปรมาพิจารณาร่วมในการวิเคราะห์ด้วย (Fitz-Gibbon, 1996)

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

มูลค่าเพิ่ม คือพัฒนาการทางด้านวิชาการที่เกิดขึ้นโดยไม่รวมส่วนของผลที่เกิดจากปัจจัยที่อยู่นอกเหนือการควบคุมของสถานศึกษา (Lambert & Line, 2000)

การวัดมูลค่าเพิ่ม เป็น การทุ่มเทการดำเนินการของปัจจัยที่มีอยู่อย่างหลากหลายเพื่อให้เกิดผลต่อการเปลี่ยนแปลงของผลสัมฤทธิ์ของนักเรียน (Goldhaber & Anthony, 2003)

โมเดลมูลค่าเพิ่ม เป็นการเปลี่ยนแปลงในระดับนักเรียน โดยการเปลี่ยนปริมาณความคาดหวังของการเปลี่ยนแปลง ตลอดจนการเปลี่ยนแปลงที่เป็นพื้นฐานของนักเรียน เช่น เพศ ระดับ SES เชื้อชาติ ผลสัมฤทธิ์เดิมหรือระดับคะแนนเริ่มต้น (Lissitz. et.al, 2006)

โมเดลมูลค่าเพิ่ม เป็นการระบุเกี่ยวกับการทุ่มเทดำเนินงานของสถานศึกษาเพื่อทำให้เกิดความก้าวหน้าของนักเรียนหรือเป็นไปตามวัตถุประสงค์การศึกษา เช่นผลสัมฤทธิ์ทางด้านสติปัญญา ซึ่งการดำเนินงานของสถานศึกษายังเกี่ยวข้องกับปัจจัยอื่นหลายปัจจัยที่สนับสนุนให้เกิดความก้าวหน้าทางการศึกษาของนักเรียน (OECD, 2008)

เมื่อพิจารณาความหมายเกี่ยวกับมูลค่าเพิ่ม พบว่าในระยะแรก จะเป็นความหมายของการเพิ่มขึ้นของผลสัมฤทธิ์ของนักเรียน โดยเปรียบเทียบจากสภาวะเริ่มต้นของนักเรียนที่เป็นพื้นฐานของนักเรียน และในระยะต่อมาเป็นการเพิ่มขึ้นของผลสัมฤทธิ์ของนักเรียนที่เกิดขึ้น แต่ให้ความสำคัญไปที่ผลจากการดำเนินงานของสถานศึกษาอย่างแท้จริง ดังนั้นการวัดมูลค่าเพิ่ม จึงเริ่มมีนักการศึกษาหลายท่านสนใจ ในการนำมาใช้ประเมินคุณภาพการจัดการศึกษาของสถานศึกษาและประสิทธิผลของสถานศึกษา เนื่องจากเป็นผลที่มีความถูกต้องและยุติธรรม เพราะมีการพิจารณาปัจจัยอื่นที่เป็นคุณลักษณะของนักเรียน เช่นเดียวกับ ศิริชัย กาญจนวาสี (2550) ที่กล่าวถึงลักษณะการวัดมูลค่าเพิ่มว่าเป็นการวัดที่ปรับให้อยู่บนพื้นฐานของปัจจัยดั้งเดิมเท่าเทียมกัน เพื่อนำผลต่างมาเปรียบเทียบกันอย่างยุติธรรม ดังนั้นการนำมูลค่าเพิ่มมาใช้ประโยชน์ทางการศึกษาจึงสามารถทำให้ชัดเจนขึ้น ความหมายของมูลค่าเพิ่มจึงสรุปได้ว่าเป็นส่วนที่เพิ่มขึ้นจากสภาวะเริ่มต้นของนักเรียนที่เกิดจากการดำเนินงานของสถานศึกษา ได้พิจารณาปัจจัยอื่นที่อยู่นอกเหนือการควบคุมของสถานศึกษาหรือปรับให้อยู่ในพื้นฐานที่เท่าเทียมกัน โดยปัจจัยที่อยู่นอกเหนือการควบคุมของสถานศึกษา เช่น ตัวแปรภูมิหลังของนักเรียน สภาพความพร้อมทางด้านเศรษฐกิจของครอบครัว ทรัพยากรทางการเรียน เป็นต้น

1.2 แนวคิดพื้นฐานของมูลค่าเพิ่ม

กระบวนการศึกษามีการดำเนินการที่มีส่วนเกี่ยวข้องกับบุคคลหลายฝ่าย เช่นผู้บริหาร ครู ผู้ปกครอง คุณภาพการจัดการศึกษาจึงเป็นผลที่เกิดจากการปฏิบัติจากหลายฝ่าย และมีส่วนเกี่ยวข้องกับปัจจัยอื่นที่เอื้อประโยชน์ต่อการดำเนินการ ดังนั้นคุณภาพของการจัดการศึกษาจึงยากที่จะนิยามเชิงสัมบูรณ์ได้ ซึ่งส่วนมากจะเป็นการศึกษาถึงการมีคุณลักษณะของสถานศึกษาที่ปฏิบัติได้ดีกว่าสถานศึกษาอื่น โดยไม่คำนึงถึงพื้นฐานของนักเรียน และใช้คะแนนจากการสอบด้วยแบบสอบมาตรฐานที่แปลออกมาในรูปของคะแนนดิบเพื่อแสดงถึงผลการจัดการของสถานศึกษา Hill (1995 cited in Downes & Vindurampulle, 2007) ชี้ให้เห็นว่าคะแนนดิบไม่ได้ให้เป็นตัวบ่งชี้ความ

แตกต่างของการดำเนินงานของสถานศึกษาได้อย่างถูกต้อง และ Jimerson, Egeland, Sroufe, & Carlson (2000) ชี้ให้เห็นว่านักเรียนที่อยู่ภายในสถานศึกษามีคุณภาพที่ต่างกัน เนื่องจากอยู่บนพื้นฐานของปัจจัยหลายอย่างที่ส่งผลต่อการพัฒนาสมองของนักเรียน เช่นสุขภาวะและการดูแลตนเอง การมีปฏิสัมพันธ์กันทางบวกระหว่างการเรียนที่บ้านและสังคม

การศึกษาคุณภาพการจัดการและคุณภาพการปฏิบัติของสถานศึกษา จึงควรนำปัจจัยพื้นฐานมาพิจารณาร่วมด้วย โดยการศึกษาว่าสถานศึกษาได้สร้างมูลค่าเพิ่มในผลการเรียนรู้ได้มากน้อยเพียงใด โดยเปรียบเทียบระหว่างคะแนนผลการเรียนรู้ที่เกิดขึ้นจริงหรือสังเกตได้ (Observed scores) กับคะแนนผลการเรียนรู้ที่ทำนายได้ (Predicted scores) จากตัวแปรภูมิหลังของนักเรียน ตัวแปรบริบทของชุมชน สังคม หรือผลสัมฤทธิ์เดิม (ศิริชัย กาญจนวาสี, 2550) นอกจากนี้ Grilli & Rampichini (2008) ได้กล่าวถึงคุณภาพของหน่วยงานทางการศึกษาสามารถวัดผ่านปัจจัยนำเข้า (input) และผลลัพธ์ (output) หรือการวัดมูลค่าเพิ่ม (value-added) และคำว่าผลลัพธ์ (output) บางครั้งในการประเมินคุณภาพการศึกษาจะใช้คำว่าผลการเรียนรู้(outcome) ดังนั้นคุณภาพการจัดการศึกษาจึงเป็นการประเมินส่วนหนึ่งของปัจจัยนำเข้า มูลค่าเพิ่ม (value added) จึงเกิดจากการปฏิบัติงานของสถานศึกษา นั่นคือ มูลค่าเพิ่มมีค่าเท่ากับผลต่างของผลลัพธ์ที่เกิดขึ้นกับผลลัพธ์ที่คาดหวังเมื่อให้ปัจจัยนำเข้า ([value added] = [actual outcome] – [expected outcome given the input])

การนำโมเดลมูลค่าเพิ่มมาใช้ทางการศึกษาเป็นการรวบรวมเทคนิคทางสถิติที่ใช้คะแนนจากการสอบของนักเรียนเพื่อประมาณค่าอิทธิพลของสถานศึกษาหรือครู (MaCaffrey, Lockwood, Koretz, & Hamilton, 2003) การนำโมเดลมูลค่าเพิ่มมาใช้มีสองแนวทาง โดยแนวทางแรกเป็นการประเมินสถานศึกษาเพื่อความรับผิดชอบที่สามารถตรวจสอบได้ และแนวทางที่สองเป็นการประเมินครูที่เป็นประสิทธิผลเชิงสัมพัทธ์กับครูคนอื่น ซึ่งในบางโมเดลจะพิจารณาเฉพาะความรู้เดิมของนักเรียนหรือรวมตัวแปรอื่นเช่น เพศ ศาสนา และพื้นฐานด้านเศรษฐานะของนักเรียนเข้าร่วมพิจารณาด้วย

ดังนั้นการนำแนวคิดของการวัดมูลค่าเพิ่มมาใช้ในประเมินคุณภาพการจัดการศึกษา จึงเป็นแนวทางหนึ่งที่ช่วยให้ข้อสารสนเทศที่สะท้อนถึงการปฏิบัติงานของสถานศึกษา เพื่อทำให้การเปรียบเทียบมีความเท่าเทียมของสถานศึกษาและสามารถเปรียบเทียบผลการปฏิบัติของสถานศึกษาได้อย่างยุติธรรม เนื่องจากมีการควบคุมปัจจัยอื่นที่มีพื้นฐานต่างกันซึ่งอาจจะมีความไม่เท่าเทียมกันของปัจจัยนำเข้าของระบบการศึกษา การนำผลจากการวัดมูลค่าเพิ่มจึงมีความถูกต้องน่าเชื่อถือและเกิดประโยชน์หลายประการต่อผู้ที่เกี่ยวข้องทั้งผู้กำหนดนโยบาย ผู้บริหาร ครู ตลอดจนผู้ปกครองและนักเรียน

1.3 ความสำคัญของการประเมินมูลค่าเพิ่ม

ผลจากการประเมินมูลค่าเพิ่มมีความสำคัญกับการนำมาใช้กำหนดนโยบายที่สำคัญ 3 ด้าน (OECD, 2008) คือการริเริ่มการพัฒนาสถานศึกษา (school improvement initiatives) ระบบความสามารถในการตรวจสอบได้ของสถานศึกษา (Systems of school accountability) และการเลือกสถานศึกษา (School choice) ในส่วนของการริเริ่มการพัฒนาสถานศึกษา เป็นสิ่งสำคัญที่การพัฒนากิจกรรมหรือชุดกิจกรรมที่ต้องการพัฒนาเป็นอันดับแรก ผลจากการวิเคราะห์ทางสถิติจะให้ข้อสนเทศเกี่ยวกับความสัมพันธ์ระหว่างปัจจัยนำเข้าและตัวบ่งชี้ของความสามารถของสถานศึกษาที่สามารถบอกถึงยุทธวิธีที่ควรปฏิบัติหรือไม่ควรปฏิบัติ นำไปสู่การกำหนดนโยบายและวางแผนนโยบาย การแบ่งสรรทรัพยากร ในส่วนของความสามารถในการตรวจสอบได้ของสถานศึกษา ผลจากการวิเคราะห์มูลค่าเพิ่ม สามารถตรวจสอบการปฏิบัติงานและผลลัพธ์จากการปฏิบัติได้ทั้งหมด (McKewen,1995) เช่นการให้ข้อมูลแก่ผู้ที่เกี่ยวข้องและสาธารณะเกี่ยวกับเงินภาษีได้อย่างมีประสิทธิภาพ ในส่วนของการเลือกสถานศึกษา ข้อมูลที่ได้จากการวิเคราะห์นี้จะมีจุดมุ่งหมายไปที่ผู้ปกครองและครอบครัวของนักเรียนเกี่ยวกับความสามารถที่แตกต่างของสถานศึกษาเพื่อการตัดสินเลือกสถานศึกษา ผู้ปกครองสามารถเลือกสถานศึกษาที่ดีที่สุดและเหมาะสมกับแหล่งทรัพยากรที่สามารถเกิดขึ้นในสถานศึกษา (Hoxby, 2003)

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

1) การประเมินนโยบายและโครงการ (Policy and program evaluation) ข้อมูลจากการวัดมูลค่าเพิ่มสามารถวิเคราะห์ได้ถึงการปฏิบัติงานของสถานศึกษา ที่สามารถเปรียบเทียบผลกระทบของสถานศึกษาเชิงสัมพัทธ์กับสถานศึกษาอื่น ที่พิจารณารวมกับปัจจัยอื่น เช่นคะแนนดิบ การประเมินครู การประเมินตนเองของสถานศึกษา การทบทวนการตัดสินและแนวโน้มของสถานศึกษา ข้อมูลที่ได้สามารถช่วยในการตัดสินใจเชิงนโยบาย (Informing policy-making) ที่มีประสิทธิภาพต่อการกำหนดนโยบายในระยะเริ่มต้น เป็นประโยชน์ในการระบุสถานศึกษาที่มีค่าเฉลี่ยสูงกว่าและต่ำกว่าค่าเฉลี่ยของความสามารถได้อย่างชัดเจน (OECD, 2006) สถานศึกษาที่มีความสามารถสูงสามารถให้ตัวอย่างในการปฏิบัติที่ดี ช่วยให้วางนโยบายให้มีความสอดคล้องกันระหว่างของนโยบาย แหล่งทรัพยากร และยุทธวิธีการสอนกับคุณลักษณะของนักเรียนและสถานศึกษา (Drury & Doran, 2003)

2) เครื่องมือหนึ่งสำหรับการพัฒนาสถานศึกษา (A tool for school improvement) เนื่องจากการวัดมูลค่าเพิ่มจะนำคุณลักษณะของนักเรียนเข้าไปพิจารณาและสามารถระบุถึงวิธีการของสถานศึกษาที่สามารถทำให้ผู้เรียนมีความสามารถตามที่ความคาดหวัง โดยอิงความสำเร็จของนักเรียน การวิเคราะห์มูลค่าเพิ่มจึงเป็นสิ่งที่สะท้อนถึงสิ่งแวดล้อมและการปฏิบัติการของสถานศึกษาเพื่อแปลความถึงผลลัพธ์ของบริบทของสถานศึกษาและชั้นเรียน เพื่อเป็นแนวทางใน

การพัฒนาสถานศึกษา การวัดมูลค่าเพิ่มจะสามารถชี้ได้ถึงปัจจัยที่มีส่งผลให้นักเรียนแสดงความสามารถได้เป็นอย่างดีหรือสามารถแสดงได้ต่ำกว่าที่คาดหวัง นั่นคือสามารถช่วยในอิทธิพลโดยตรงเพื่อพัฒนาผลลัพธ์การเรียนรู้ของนักเรียน

3) ประสิทธิผลของครู (Teacher effectiveness) วิธีการวิเคราะห์มูลค่าเพิ่มมีผลต่อผลสัมฤทธิ์ของนักเรียนซึ่งนำไปใช้ในการประเมินประสิทธิภาพของครู จากการศึกษาการวิเคราะห์มูลค่าเพิ่ม พบว่าประสิทธิภาพของครูเป็นตัวบ่งชี้ที่สำคัญของความสัมพันธ์ของสถานศึกษาต่อการเรียนรู้ของนักเรียน และนอกจากนี้จากข้อสังเกตพบว่าการวิเคราะห์มูลค่าเพิ่มจะเป็นเครื่องมือในการลดช่องว่างของผลสัมฤทธิ์ของนักเรียนในข้อได้เปรียบและข้อเสียเปรียบของนักเรียน นอกจากนี้ข้อมูลที่ได้จะเป็นการให้ข้อมูลย้อนกลับแก่ครูเกี่ยวกับวิธีการสอน มีความชัดเจนในเป้าหมายของการเรียนรู้ที่ส่งผลต่อการเปลี่ยนแปลงวิธีการสอน และสามารถใช้เป็นข้อมูลสำหรับครูในการพัฒนานักเรียนต่อไป กล่าวคือสามารถระบุคุณลักษณะของนักเรียนที่ต้องมีการพัฒนาทักษะต่อไป และใช้เป็นข้อมูลต่อการวางแผนการพัฒนาของครู ตัวอย่างเช่น การวิเคราะห์มูลค่าเพิ่ม จะระบุครูที่มีประสิทธิภาพในการยกระดับความสามารถของนักเรียนที่มีจุดอ่อน (Drury & Doran, 2003)

4) เครื่องมือหนึ่งสำหรับความรับผิดชอบที่ตรวจสอบได้ (A tool for accountability) ตัวบ่งชี้การปฏิบัติงานของสถานศึกษาเป็นส่วนหนึ่งของการตรวจสอบได้ของสถานศึกษา การวัดมูลค่าเพิ่มจะเพิ่มความรอบคอบในการใช้ที่เป็นพื้นฐานสำหรับการกำหนดภาพรวมของการปฏิบัติงานสถานศึกษา จากรายงานของ OECD (2008) เกี่ยวกับวิธีการมูลค่าเพิ่มจะได้รับข้อสรุปที่คล้ายกันเกี่ยวกับกระบวนการการศึกษา โดยรายงานว่าโมเดลมูลค่าเพิ่มเป็นวิธีที่ดีที่สุดที่แนะนำได้ว่าเป็นเครื่องมือที่ส่งเสริมการพัฒนาสถานศึกษา

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

วิธีการวัดมูลค่าเพิ่มได้มีการนำมาใช้ในหลายประเทศเพื่อใช้ผลมาพิจารณาการดำเนินการของสถานศึกษาที่ส่งต่อผลสำเร็จของนักเรียน (OECD, 2008) ดังตัวอย่างของระบบการประเมินในรัฐเทนเนสซีของประเทศสหรัฐอเมริกาที่มีชื่อระบบการประเมินว่าระบบการวัดมูลค่าเพิ่มของเทนเนสซี (Tennessee Value-added Assessment System: TVAAS) ซึ่งออกแบบเพื่อวัดผลกระทบของสถานศึกษาและครูที่ส่งผลต่อความสำเร็จของนักเรียน ซึ่งพัฒนาโดย William L. Sanders และมีการศึกษาเพิ่มเติมไปสู่การดำเนินการระบบการประเมินมูลค่าเพิ่มทางการศึกษา (Education Value-Added Assessment System: EVAAS) (Sanders, 1998) นอกจากนี้ Heck (2000) ยังศึกษาการตรวจสอบผลกระทบของคุณภาพสถานศึกษาต่อผลลัพธ์ของสถานศึกษาและการพัฒนา โดยการประยุกต์ใช้วิธีการมูลค่าเพิ่ม โดยให้ความสำคัญไปที่ระดับรัฐเพื่อติดตาม

ผลสำเร็จของนักเรียน และทำความเข้าใจเกี่ยวกับความสามารถรับผิดชอบของสถานศึกษา การรายงานผลการปฏิบัติงานของสถานศึกษาแก่ผู้กำหนดนโยบาย บุคลากรของสถานศึกษาและผู้ปกครองเพื่อใช้ผลการเปรียบเทียบสถานศึกษาให้ความยุติธรรมยิ่งขึ้น

1.4 ขั้นตอนและวิธีการที่ใช้วิเคราะห์มูลค่าเพิ่ม

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

ขั้นตอนที่ 1 การกำหนดวัตถุประสงค์และการวัดผลการปฏิบัติงานของสถานศึกษา ซึ่งขั้นตอนนี้จะกำหนดวัตถุประสงค์สำหรับการนำโมเดลมูลค่าเพิ่มไปใช้ ประกอบด้วยการกำหนดลักษณะของสารสนเทศของมูลค่าเพิ่มและอธิบายถึงการอธิบายคะแนนมูลค่าเพิ่ม การพิจารณาโครงสร้างของการประเมินนักเรียนที่มีความสอดคล้องกับวัตถุประสงค์โมเดลมูลค่าเพิ่ม ตลอดจนการระบุเนื้อหาและระดับชั้นที่ประเมิน การกำหนดเป้าหมายการประเมิน และการพิจารณาศักยภาพของเครื่องมือประเมินจะต้องมั่นใจว่าสามารถใช้ในโมเดลมูลค่าเพิ่ม ซึ่งเป็นสิ่งสำคัญที่จะนำมาใช้สำหรับนำมาแปลความหมายของผลการปฏิบัติ

ขั้นตอนที่ 2 การนำเสนอและการใช้ข้อสารสนเทศของมูลค่าเพิ่ม เป็นการตัดสินและเลือกวิธีการที่เหมาะสมสำหรับการแสดงสารสนเทศของมูลค่าเพิ่ม ตลอดจนแนวทางการแปลความหมายของมูลค่าเพิ่ม ซึ่งสามารถสร้างขึ้นได้โดยผู้ที่เกี่ยวข้องในการจัดทำนโยบาย

ขั้นตอนที่ 3 การพิจารณาคุณภาพของข้อมูล เป็นขั้นตอนที่ควรจะพิจารณาระบบของข้อมูลในสถานศึกษาและสารสนเทศเพื่อรวบรวมสำหรับนำไปใช้ในโมเดลมูลค่าเพิ่ม ตลอดจนการกำหนดกลุ่มตัวอย่างนักเรียนที่เพียงพอต่อการนำมาใช้ในโมเดลมูลค่า ซึ่งอาจเกิดการสูญหายข้อมูล (missing data) ที่อาจเกิดการย้ายสถานศึกษาในช่วงที่มีการประเมิน ข้อมูลที่จำเป็นจะต้องสนับสนุนโมเดล หรือการนำไปใช้สำหรับการกำหนดนโยบายและพัฒนาโครงการต่อไป

ขั้นตอนที่ 4 การเลือกโมเดลมูลค่าเพิ่มที่เหมาะสม การกำหนดวัตถุประสงค์เชิงนโยบายและการใช้คะแนนมูลค่าเพิ่มของสถานศึกษาจะเป็นแนวทางในการเลือกโมเดล สิ่งจำเป็นที่จะต้องกำหนดคือการเลือกใช้สถิติและวิธีการวิเคราะห์ภายใต้โมเดลมูลค่าเพิ่ม ความแปรปรวนในโมเดลมูลค่าเพิ่ม เพื่อประเมินความเหมาะสมของลักษณะของโมเดล ลักษณะเฉพาะของโมเดลอาจจะทำให้

เกิดความแตกต่างของสถานศึกษา ผลกระทบของการเลือกโมเดลขึ้นอยู่กับความแตกต่างของสถานศึกษาในจุดประสงค์การใช้คะแนนมูลค่าเพิ่ม

ขั้นตอนที่ 5 กลยุทธ์กระตุ้นการสื่อสารและผู้ที่มีส่วนได้ส่วนเสีย เป้าหมายของวิธีการสื่อสารจะสอดคล้องกับวัตถุประสงค์ของนโยบาย การวัดความสามารถการปฏิบัติงานของโรงเรียนจะนำมาตัดสินควรจะมีความการอธิบายได้อย่างชัดเจน การกระตุ้นและการสื่อสารต่อผู้ที่เกี่ยวข้องจะต้องชี้แจงได้อย่างชัดเจน เพื่อนำมาใช้แปลความหมายในผลของมูลค่าเพิ่มของสถานศึกษา

ขั้นตอนที่ 6 การฝึกอบรม เป้าหมายของการฝึกอบรมมีเป้าหมายในการพัฒนากลุ่มผู้ใช้การฝึกอบรมผู้บริหารสถานศึกษาและครู จะเป็นการอธิบายเกี่ยวกับผลของมูลค่าเพิ่มและสามารถนำมาใช้ภายในสถานศึกษาเพื่อวัตถุประสงค์ของการพัฒนาสถานศึกษา รวมถึงการอบรมเกี่ยวกับการวิเคราะห์ทางสถิติและการใช้ข้อสารสนเทศที่จำเป็น นอกจากนี้การฝึกอบรมสำหรับผู้ปกครองมีเป้าหมายสำหรับเพื่อการแปลความหมายของคะแนนมูลค่าเพิ่มเพื่อความสะดวกในการเลือกสถานศึกษา

ขั้นตอนที่ 7 การนำร่องโครงการ เป็นการตัดสินใจเกี่ยวกับขนาดและคุณลักษณะของกลุ่มตัวอย่างของสถานศึกษาที่จะมีความสอดคล้องกับวัตถุประสงค์ของนโยบายการนำระบบของโมเดลมูลค่าเพิ่มไปใช้ การนำร่องโครงการควรจะใช้เพื่อประเมินกิจกรรมที่มีความเชื่อมโยงกับโมเดลมูลค่าเพิ่ม

ขั้นตอนที่ 8 การนำไปพัฒนา เป็นขั้นตอนสุดท้ายของการนำโมเดลมูลค่าเพิ่มไปใช้ ซึ่งเป็นการติดตามผลของคุณภาพ แหล่งทรัพยากร และมีการพิจารณาถึงความสอดคล้องของโมเดลมูลค่าเพิ่มกับนโยบาย

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

1) การวิเคราะห์โดยวิธีการหักออกของคะแนนดิบ (Simple Subtraction of raw score) เป็นการวัดจุดเริ่มต้นและจุดสิ้นสุดของการศึกษา ผลที่ได้ของวิธีนี้จะเรียกคะแนนพัฒนาการ (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 หรืออาจจะเป็นปัจจัยอื่นโดยพิจารณาตามเงื่อนไขหรือทฤษฎีที่เกี่ยวข้อง

และสารสนเทศของบริบทที่กำหนดในโมเดลการวิเคราะห์ จึงจัดได้ว่าเป็นวิธีที่สามารถบ่งบอกถึงมูลค่าที่สถานศึกษาจัดให้เกิดขึ้นกับนักเรียนด้วยกระบวนการเรียนรู้ที่มีความใกล้เคียงกับหลักการของการวิเคราะห์มูลค่าเพิ่ม ผลการวิเคราะห์ที่แสดงถึงมูลค่าเพิ่มจากการวิเคราะห์ด้วยวิธีนี้คือ ค่าส่วนที่เหลือ (residual term) หรือส่วนต่างระหว่างผลสัมฤทธิ์ที่วัดได้กับค่าที่ประมาณได้จากชุดของปัจจัยหรือตัวแปรที่กำหนดในโมเดลการวิเคราะห์ว่ามีค่าเป็นบวกหรือลบในปริมาณมากน้อยเพียงใด การแปลผลว่าสถานศึกษาใดได้จัดมูลค่าเพิ่มทางการศึกษาให้กับนักเรียน คือ การพิจารณาจากค่าส่วนเหลือที่มีค่าเป็นบวก และถ้าค่าดังกล่าวมีค่าเป็นลบ หรือมีค่าเป็นศูนย์ แสดงว่าสถานศึกษาไม่สามารถจัดการศึกษาที่เป็นมูลค่าเพิ่มให้กับนักเรียน

การวิเคราะห์ถดถอยช่วยลดข้อจำกัดในการเปรียบเทียบผลสัมฤทธิ์จากการปฏิบัติระหว่างสถานศึกษาลงได้ แต่ยังมีข้อจำกัดของเทคนิคการวิเคราะห์หลายประการที่ทำให้ได้ผลการวิเคราะห์ที่ผิดพลาด ดังนี้ (1) การละเลยหลักธรรมชาติของสถานศึกษาที่มีลักษณะลดหลั่นสอดแทรกด้วยการวิเคราะห์ข้อมูลจากปัจจัยทุกระดับภายในโมเดลการวิเคราะห์เดียวกันซึ่งเป็นการปฏิบัติที่ฝ่าฝืนข้อตกลงเบื้องต้นในเรื่องความเป็นอิสระของข้อมูลจึงเป็นสาเหตุของการเกิดความคลาดเคลื่อนชนิดที่ 1 (type I error) (2) การคัดเลือกสาระของข้อมูลเกี่ยวกับผลผลิตเพื่อนำเสนอต่อสาธารณชนโดยการคัดเลือกเฉพาะปัจจัยหรือตัวแปรระดับบุคคลของนักเรียนเพียงบางส่วนเพื่อใช้ในกระบวนการวิเคราะห์ประมาณค่าผลสัมฤทธิ์ครั้งหลัง และในการเปรียบเทียบจึงอาจทำให้ขาดความชัดเจนในการสรุปผลการปฏิบัติของสถานศึกษา ซึ่งต่างจากการแสดงผลด้วยคะแนนดิบ ดังนั้นการแปลผลอาจทำได้ยาก (3) มีความเป็นไปได้ที่สถานศึกษาซึ่งมีผลสัมฤทธิ์จัดอยู่ในระดับต่ำแต่กลับได้รับการปรับแก้ในลักษณะการถดถอยเข้าหาค่าเฉลี่ยของกลุ่มนักเรียนที่มีความสามารถสูงเนื่องจากกระบวนการวิเคราะห์ได้ทำให้ผลการวิเคราะห์มีความคลาดเคลื่อน

3) วิธีวิเคราะห์ด้วยโมเดลพหุระดับ (Multilevel modeling) เป็นการวิเคราะห์ถดถอยอีกวิธีหนึ่งที่คำนึงถึงลักษณะข้อมูลแบบลดหลั่นเป็นระดับขั้น (hierarchical) ตามธรรมชาติของข้อมูลที่เกี่ยวข้องกับสถานศึกษา จุดเด่นของวิธีนี้คือปัจจัยที่ส่งผลกระทบต่อความแตกต่างของผลสัมฤทธิ์นั้นไม่ได้เกิดจากปัจจัยระดับบุคคลเพียงระดับเดียว แต่อาจมีผลที่เกิดจากปัจจัยระดับห้องเรียนและระดับสถานศึกษาด้วย จึงต้องมีการวิเคราะห์เพื่อปรับแก้ด้วยปัจจัยเหล่านั้นเป็นระดับขั้นไปโดยเริ่มจากระดับเล็กสุดคือ ระดับบุคคล ผลการวิเคราะห์ที่แสดงถึงมูลค่าเพิ่มคือ ค่าส่วนที่เหลือ (residual terms) จากโมเดลการวิเคราะห์ระดับสถานศึกษา ซึ่งแบ่งออกเป็นสองส่วน ได้แก่ ค่าส่วนที่เหลือของค่าจุดตัดหรือค่าเฉลี่ยของกลุ่มสถานศึกษา (intercept residual) และค่าส่วนที่เหลือของค่าความชัน (slope residual) ของกลุ่มตัวแปรภูมิหลังนักเรียนที่กำหนดไว้ในโมเดลการวิเคราะห์ ดังนั้นหากค่าส่วนที่เหลือเหล่านี้มีค่าเป็นบวก สามารถแปลความหมายได้ว่าสถานศึกษาได้จัดการศึกษาที่ก่อให้เกิดมูลค่าเพิ่มให้กับนักเรียนเมื่อเทียบจากผลสัมฤทธิ์ที่ได้จากการประมาณค่าบนพื้นฐานของปัจจัยหรือตัวแปรที่กำหนดในโมเดลการวิเคราะห์นั้น ในทางตรงข้ามถ้าส่วนที่เหลือมีค่าเป็นลบหรือศูนย์ แสดงว่าสถานศึกษาไม่สามารถจัดการศึกษาให้เกิดมูลค่าเพิ่มขึ้นได้

จากวิธีการวิเคราะห์ที่มีลักษณะและความซับซ้อนแตกต่างกัน การเลือกใช้วิธีวิเคราะห์สามารถเลือกโดยพิจารณาจากความเข้าใจและความซับซ้อนของหลักทางสถิติซึ่งอาจจะมีข้อจำกัดสำหรับการเลือกใช้โดยทั่วไป นอกจากนี้การพิจารณาถึงโครงสร้างข้อมูล และตัวแปรในแต่ละระดับของข้อมูล ซึ่งอาจจะสามารถเข้าถึงตัวแปรได้ยาก หรือต้องใช้ระยะเวลาในการเก็บรวบรวมข้อมูลค่อนข้างยาวนาน แต่โดยส่วนมากการวิเคราะห์ที่มีความสอดคล้องกับโครงสร้างข้อมูลทางการวิจัยทางสังคมศาสตร์มากที่สุด จะเป็นการวิเคราะห์พหุระดับ เนื่องจากตัวแปรในระดับต่างกันมีความสัมพันธ์กันและสามารถนำมาใช้ในการแปลผลและนำไปใช้ให้เกิดประโยชน์มากที่สุด

2.5 โมเดลมูลค่าเพิ่ม (Value – Added Models)

โมเดลมูลค่าเพิ่มที่รู้จักกันดีเช่น ระบบการประเมินมูลค่าเพิ่มของรัฐเทนเนสซี (Tennessee Value Added Assessment: TVAAS) (Ballou, Sanders, Wright, 2004) และระบบการประเมินมูลค่าเพิ่มทางการศึกษา (Education Value-Added Assessment System: EVAAS) ซึ่ง EVAAS เป็นกระบวนการทางสถิติสำหรับกลุ่มใหญ่ของการเปลี่ยนแปลงคะแนนผลสัมฤทธิ์ของนักเรียนเมื่อเวลาผ่านไป โดยผลจะนำมาใช้กำหนดสิ่งที่เกิดการพัฒนาและประสิทธิผลของเขตพื้นที่ สถานศึกษาและครูเมื่อเวลาผ่านไป (Sanders & Horn, 1994; Sanders, et al., 1998) ความแตกต่างของโมเดลมูลค่าเพิ่มแต่ละโมเดล ขึ้นอยู่กับการกำหนดลักษณะเฉพาะของโมเดล ทั้งนี้โมเดลที่เลือกแต่ละโมเดลขึ้นอยู่กับความสามารถในการเข้าถึงข้อมูล 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 และเครื่องมือวัดผลสัมฤทธิ์ของนักเรียนเพื่อพิจารณาผลความแตกต่างระหว่างการวัดสองครั้ง ซึ่งอาจจะมีสเกลการวัดที่ต่างกัน ตลอดจนปัจจัยพื้นฐานอื่นๆของนักเรียน ดังนั้น Haegeland and Kirkeboen (2008) ได้จัดแบ่งโมเดลมูลค่าเพิ่ม (Value added models: VAM) กับโมเดลมูลค่าเพิ่มเชิงบริบท (Contextualised value-added modeling: CVA) โดยโมเดลมูลค่าเพิ่มจะพิจารณาตัวแปรที่นำเข้าเฉพาะความรู้เดิมของนักเรียน ส่วนโมเดลมูลค่าเพิ่มเชิงบริบท จะนำตัวแปรบริบทที่เป็นตัวแปรเศรษฐานะเข้าร่วมพิจารณาในโมเดลด้วย ในประเทศอังกฤษ โมเดลที่มีความโดดเด่น คือโมเดลมูลค่าเพิ่มเชิงบริบท (Contextualised value-added modelling) เป็นโมเดลที่มีความซับซ้อนเนื่องจากมีการควบคุมอิทธิพลที่มีอยู่หลากหลายของคุณลักษณะทางด้านเศรษฐานะต่อการเปลี่ยนแปลงความสามารถของนักเรียน และคุณลักษณะหลายอย่างที่มีอิทธิพลต่อ

ความก้าวหน้าของนักเรียนที่อยู่นอกเหนือการควบคุมของสถานศึกษา เช่น เพศ เชื้อชาติ ศาสนา ความจำเป็นในการศึกษาพิเศษ ภาษาที่ใช้ การย้ายโรงเรียน การดูแลเอาใจใส่ เป็นต้น OECD (2008) ได้มีการจำแนกระหว่างโมเดลมูลค่าเพิ่ม (Value added modeling) และโมเดลความสำเร็จของบริบท (Contextualised attainment models) ซึ่งโดยทั่วไปโมเดลมูลค่าเพิ่มจะมีการวัดเกี่ยวกับความรู้เดิมของนักเรียนที่เป็นพื้นฐานสำหรับการนำมาสู่ความแตกต่างของนักเรียน ส่วน โมเดลความสำเร็จของบริบท (Contextualised attainment models) จะไม่รวมการวัดความรู้เดิมของนักเรียน แต่ปัญหาเกี่ยวกับการวัดสองครั้งอาจมีปัญหาจากสเกลการวัดที่ต่างกัน ถ้าคะแนนอยู่ในสเกลเดียวกัน จึงจะสามารถคำนวณความแตกต่างหรือคะแนนที่เพิ่มของแต่ละคน คะแนนที่เพิ่มขึ้นโดยเฉลี่ยของนักเรียนจะพิจารณาได้ว่าเป็นคะแนนมูลค่าเพิ่มของสถานศึกษา (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) โมเดลหลายตัวแปรสามารถใช้ในการให้การวัดการปฏิบัติงานของสถานศึกษา โมเดลนี้เป็นการประมาณค่าขนาดของการดำเนินการของปัจจัยต่างๆ ต่อความสามารถของนักเรียนหรือความสำเร็จในช่วงเวลาเดียว ตัวอย่างเช่นการวิเคราะห์โมเดลถดถอยของพื้นฐานทางเศรษฐานะของนักเรียนหรือคุณลักษณะทางด้านบริบทและตัวแปรอื่นๆระบุสถานศึกษาที่มีผลต่อการวัดผลสัมฤทธิ์ของนักเรียน การปรับคะแนนดิบด้วยคุณลักษณะของบริบทเป็นการวัดที่สะท้อนถึงการดำเนินงานของสถานศึกษาต่อการเรียนรู้ของนักเรียนมากกว่าการใช้คะแนนดิบในการวัดการปฏิบัติงานของสถานศึกษา ผลของโมเดลภาคตัดขวางนี้ขึ้นอยู่กับการวิเคราะห์เชิงทฤษฎีเกี่ยวกับบทบาทของผลลัพธ์ทางเศรษฐานะส่วนมากจะเป็นการศึกษาผลการดำเนินที่สำคัญต่อระดับของความสำเร็จของนักเรียนคือเศรษฐานะของผู้ปกครองของนักเรียน (OECD, 2007b; Haveman and Wolfe, 1995) ข้อสารสนเทศเกี่ยวกับพื้นฐานทางเศรษฐานะของนักเรียนต่อความสำเร็จทางการศึกษา และให้ความสำคัญในข้อสารสนเทศที่เพียงพอต่อผู้กำหนดนโยบาย โมเดลความสำเร็จทางด้านบริบท(Contextualised attainment models) เป็นโมเดลที่มีความชัดเจนต่อการใช้ผลที่มาได้ปรับและคะแนนความสำเร็จของคะแนนดิบต่อการประเมินการปฏิบัติงานของสถานศึกษา

การศึกษาเพื่อเปรียบเทียบโมเดลมูลค่าเพิ่ม ได้มีนักการศึกษาสนใจจำนวนมาก เนื่องจากความซับซ้อนของโมเดลมีความแตกต่างกัน รวมถึงตัวแปรที่นำเข้ามาพิจารณาในโมเดลอาจจะมีข้อจำกัดในการเข้าถึงได้ วัตถุประสงค์ของการศึกษาโดยส่วนมาก เป็นการเปรียบเทียบโมเดล เพื่อเสนอแนวทางที่หลากหลายในการเลือกโมเดลแต่ละโมเดลตามความเหมาะสม 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 มีการคำนวณซับซ้อนเล็กน้อยและไม่ต้องใช้ข้อสารสนเทศเกี่ยวกับปัจจัยแทรกซ้อน (covariates) 2) โมเดลอิทธิพลผสมผสานแบบเป็นชั้น (Layered mixed effects model: LMEM) โมเดลนี้มีลักษณะเดียวกับ TVAAS โดย Sanders และ Horn เป็นการใช้ข้อสารสนเทศในความแปรปรวนที่ไม่เป็นศูนย์ (non-zero covariance) ระหว่างคะแนนสอบที่เวลาต่างกัน เป็นโมเดลสำหรับการเปลี่ยนแปลงคะแนนด้วยอิทธิพลสถานศึกษาแบบสุ่มที่ไม่พยายามที่จะนำตัวแปรแทรกซ้อนเข้าไปทั้งระดับสถานศึกษาและระดับเขตพื้นที่ วิธีการนี้พยายามที่จะกำจัดอิทธิพลของตัวแปรแทรกซ้อน โมเดลนี้จะใช้วัดผลสัมฤทธิ์ของนักเรียนสองครั้ง หรืออาจจะรวมทั้งเป็นการวัดข้อมูลระยะยาว โมเดลนี้จะมีข้อดีเพิ่มขึ้นเกี่ยวกับการรวมข้อสารสนเทศจากข้อมูล และสามารถประเมินพร้อมๆกันได้ของผลที่มาจากหลายเนื้อหา 3) โมเดลเชิงเส้นแบบลดหลั่น (Hierarchical linear models: HLMM) ซึ่งกำหนดอิทธิพลสถานศึกษาเป็นแบบสุ่ม เป็นโมเดลที่ดำเนินการอิทธิพลย่อส่วน(shrunken effects) มีอยู่สองลักษณะ คือ โมเดลที่ไม่ได้ปรับแก้อย่างง่าย (Unadjusted HLMM: UHLMM) เป็นโมเดลที่ไม่ได้นำตัวแปรที่เป็นองค์ประกอบเข้าพิจารณาหรือไม่มีตัวแปรแทรกซ้อนระดับนักเรียน และ โมเดลที่มีการปรับแก้ (Adjusted HLMM: AHLMM) โดยนำตัวแปรแทรกซ้อนทั้งในระดับนักเรียนและระดับสถานศึกษา เข้าอธิบายคะแนนที่เปลี่ยนแปลง ผลจากการศึกษาของ TCMLRAFR โดยใช้ข้อมูลจำลองแสดงให้เห็นว่ามีค่าสหสัมพันธ์กันสูงระหว่างผลของ SFEM, LMEM, และ UHLMM แต่มีค่าสหสัมพันธ์ปานกลางระหว่างโมเดล AHLMM กับโมเดลอื่น การศึกษาของ TCMLRAFR สรุปว่าผลจากการวิเคราะห์ใน SFEM ได้ผลดีเช่นเดียวกันกับสองโมเดลที่ไม่ได้รวมตัวแปรที่เป็นองค์ประกอบที่เป็นตัวแปรแทรกซ้อนระดับนักเรียน และสามารถคาดหวังได้ผลเช่นเดียวกันแต่ที่มีค่าใช้จ่ายในการคำนวณต่ำกว่า ความแตกต่างระหว่าง AHLMM และโมเดลอื่นชี้ให้เห็นว่าเมื่อนำตัวแปรแทรกซ้อนระดับนักเรียนและตัวแปรองค์ประกอบใช้วิเคราะห์ในโมเดลด้วย การประมาณค่าจะเกิดการเปลี่ยนแปลง ถ้าตัวแปรแทรกซ้อนมีผลกระทบต่อการเรียนรู้ของนักเรียน ก็สามารถอธิบายได้ว่าวิธีการของ AHLMM มีการประมาณค่าที่ถูกต้องแม่นยำกว่าโมเดลอื่น เพราะมีการรวมข้อสารสนเทศและประมาณค่าได้ถูกต้องแม่นยำกว่า แต่โมเดล SFEM ก็จัดเป็นโมเดลทางเลือกได้เพราะมีวิธีการที่คำนวณอย่างง่าย

การศึกษาของ BSW (Ballou, Sanders, Wright, 2004) เป็นการประมาณ TVAAS model และเป็นโมเดลที่มีความเท่าเทียมกับโมเดล LMEM ของการศึกษาของ TCMLRAFR ผลพบว่าการนำปัจจัยด้านบริบทในโมเดล HLM ชี้ให้เห็นว่าได้ผลลัพธ์ที่มีความไวต่ออิทธิพล และชี้ให้เห็นว่า TVAAS สามารถรวมปัจจัยด้านบริบทได้ด้วย การรวมปัจจัยเหล่านี้จะมีแนวโน้มที่ทำให้ความลำเอียงจากการวัดของอิทธิพลของครูและนักเรียนมีค่าเข้าใกล้ศูนย์ การใช้ข้อมูลจากฐานข้อมูลจำนวนมากจากการศึกษา TVAAS และ BSW ที่เป็นการศึกษาข้อมูลจำลองเพื่อกำหนดว่าขนาด

อิทธิพลของครูมีการเปลี่ยนแปลงเท่าใดถ้าอิทธิพลของนักเรียนและสถานศึกษานำเข้าไปวิเคราะห์ในโมเดล และข้อสรุปว่าตัวแปรแทรกซ้อนระดับนักเรียนมีอิทธิพลเพียงเล็กน้อยต่อคะแนนผลการปฏิบัติของครู และจากการศึกษาของ BSW ได้เสนอผลของการศึกษาจากข้อมูลจำลองว่าผลกระทบของตัวแปรนักเรียนมีค่าไม่มากพอที่เห็นความแตกต่างในการประมาณค่าอิทธิพลของครู และค่าสหสัมพันธ์ที่สูงระหว่างอิทธิพลที่มีการปรับแก้และไม่มีการปรับแก้เป็นสาเหตุเนื่องจากการย่อส่วน (shrinkage)

Haegeland and Kirkeboen (2008) ได้ศึกษาเกี่ยวกับการใช้ตัวบ่งชี้ในโมเดลมูลค่าเพิ่มที่ต่างกันในการปรับแก้การปฏิบติงานของสถานศึกษา มี 2 แนวทางคือ การควบคุมความรู้เดิมหรือความสำเร็จเดิมของนักเรียน และการควบคุมเศรษฐานะของนักเรียน(SES) และใช้โมเดลที่ต่างกัน 4 โมเดลคือ โมเดลที่ไม่ได้ปรับแก้ (Unadjusted results: UR) ซึ่งพิจารณาเฉพาะตัวแปรสถานศึกษา โมเดลความสำเร็จของบริบท (Contextualised attainment models: CAM) ซึ่งประกอบด้วยตัวแปร SES และตัวแปรสถานศึกษา โมเดลมูลค่าเพิ่ม (Value added models: VAM) ประกอบด้วยตัวแปรความรู้เดิม และตัวแปรสถานศึกษา และโมเดลมูลค่าเพิ่มเชิงบริบท (Contextualised value added models: CVA) ประกอบด้วยตัวแปร SES ตัวแปรความรู้เดิม และตัวแปรสถานศึกษา การศึกษานี้จะเปรียบเทียบอิทธิพลของสถานศึกษาในโมเดลที่แตกต่างกัน ซึ่งพิจารณาผลของชุดตัวแปร SES มีผลต่อการประมาณค่าอิทธิพลของสถานศึกษาอย่างไร เมื่อรวมตัวแปร SES เข้าในโมเดล CAM และ CVA ผลจากการประมาณค่าของโมเดล SES ทำให้เกิดความแตกต่างกันของผลการปฏิบัติงานของสถานศึกษา และการใช้จำนวนตัวแปร SES มีผลต่อการประมาณค่าของลักษณะที่แตกต่างกัน เมื่อเพิ่มตัวแปร SES เข้าไปในเงื่อนไขของโมเดล CAM จะทำให้การกระจายของตัวบ่งชี้ผลการปฏิบัติของสถานศึกษาแคบลง นั่นคือตัวแปร SES มีความสัมพันธ์กับความสำเร็จแต่ละคน ดังนั้นการศึกษานี้จึงแสดงให้เห็นว่าตัวบ่งชี้การปฏิบัติของสถานศึกษาอาจจะเกิดผลกระทบเนื่องจากความแตกต่างของ SES และความรู้เดิมของนักเรียน ข้อค้นพบในครั้งนี้ ชี้ให้เห็นว่าโมเดลความสำเร็จของบริบท(contextualised attainment model) ซึ่งเป็นโมเดลที่ปราศจากตัวแปรความรู้เดิมของนักเรียน และรวม SES ในโมเดลอาจจะเกิดผลกระทบอย่างมากต่อตัวแปรการปฏิบัติงานของสถานศึกษา การรวมตัวแปร SES อาจจะเป็นกระทบอย่างมากต่อการปฏิบัติงานของสถานศึกษา

นอกจากนี้การศึกษาของ Thomas, (1996) มีการศึกษาเปรียบเทียบโมเดลมูลค่าเพิ่มของประสิทธิผลในสถานศึกษาระดับมัธยมศึกษา ด้วยการวิเคราะห์พหุระดับในโมเดลต่างกันที่ควบคุมความแตกต่างของปัจจัยพื้นฐานของนักเรียนและสถานศึกษา พบว่าความผันแปรมีค่าอย่างมากในระดับสถานศึกษา และการเปรียบเทียบระหว่างโมเดลที่แตกต่างของโมเดลมูลค่าเพิ่ม และปัจจัยระดับสถานศึกษาเป็นสิ่งที่ไม่มีนัยสำคัญในการทำนายผลลัพธ์ของนักเรียน นอกจากนี้ยังพบว่าถึงแม้ว่าจะขาดตัวแปรความรู้เดิมของนักเรียน ปัจจัยทางด้านบริบทก็เป็นสิ่งที่มีประโยชน์อย่างมาก และอาจจะมีความเพียงพอและเป็นการวัดที่เหมาะสมในการนำเข้าไปพิจารณา เนื่องจากสถานศึกษาแต่ละแห่งมีความไม่คงที่ของประสิทธิผลของหน่วยงาน

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

สิ่งที่พิจารณาการกำหนดโมเดล สรุปได้ว่ามีประเด็นที่น่าสนใจ ประเด็นแรกเป็นความซับซ้อนในการคำนวณหรือการประมาณค่าในโมเดล ประเด็นที่สองคือการให้ความสำคัญของตัวแปรแทรกซ้อนที่อาจจะส่งผลตัวแปรตามร่วมกับตัวแปรต้น ประเด็นที่สามคือตัวแปรที่นำเข้ามาในโมเดล ซึ่งการใช้ตัวแปรที่เป็นตัวบ่งชี้ของการปฏิบัติงานของสถานศึกษาอาจจะมีความแตกต่างกันในแต่ละบริบท และอาจจะมีข้อจำกัดในการเข้าถึงข้อมูลได้ ดังนั้นการเลือกโมเดลมูลค่าเพิ่มเพื่อใช้ในการประเมินคุณภาพการจัดการของสถานศึกษาจึงควรระวังในการเลือกโมเดลที่เหมาะสมเนื่องจากสภาพบริบทที่แตกต่างกัน

ตอนที่ 2 มโนทัศน์เกี่ยวกับทฤษฎีการตอบสนองข้อสอบและการทำหน้าที่ต่างกันของข้อสอบ

2.1 มโนทัศน์เกี่ยวกับทฤษฎีการตอบสนองข้อสอบ

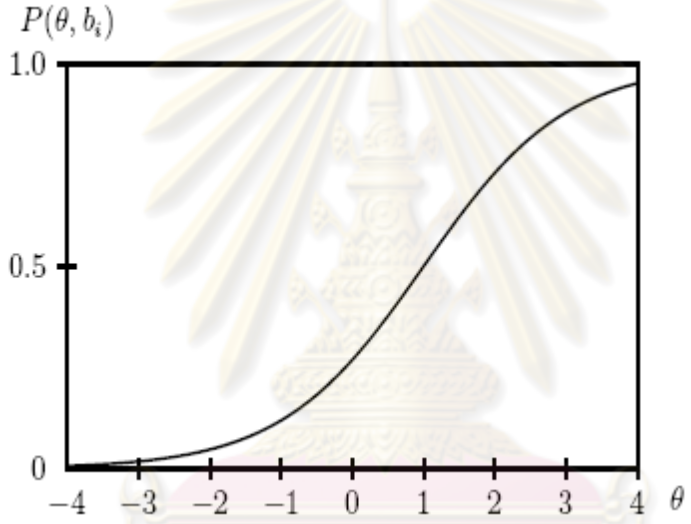
ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) เป็นโมเดลทางคณิตศาสตร์ที่มีลักษณะเฉพาะที่มีความสัมพันธ์ของคุณลักษณะแฝง (Trait) กับคุณลักษณะของข้อสอบ (Item characteristics) ของการตอบสนองของผู้สอบ (Embretson & Riese, 2000) โมเดล IRT จำแนกตามจำนวนพารามิเตอร์ที่นำมาพิจารณาถึงคุณลักษณะของข้อสอบ ประกอบด้วยโมเดลหนึ่งพารามิเตอร์ ซึ่งรู้จักกันในอีกชื่อหนึ่งคือโมเดลราสช์ จะพิจารณาเฉพาะพารามิเตอร์ความยากของข้อสอบ ส่วนโมเดลสองพารามิเตอร์ จะพิจารณาพารามิเตอร์ความยากของข้อสอบกับอำนาจจำแนกของข้อสอบ และโมเดลสามพารามิเตอร์ จะเพิ่มอีกหนึ่งพารามิเตอร์คือพารามิเตอร์ของการเดาข้อสอบ

ทฤษฎีการตอบสนองข้อสอบมีข้อได้เปรียบกว่าการทดสอบแบบดั้งเดิม เนื่องจาก โมเดล IRT จะประมาณค่าพารามิเตอร์ของข้อสอบมีข้อจำกัดในด้านของขนาดกลุ่มตัวอย่างน้อยกว่า

ทฤษฎีการวัดแบบดั้งเดิม เมื่อพิจารณาค่าสถิติที่ได้ก็สามารถอธิบายได้อย่างแม่นยำ และการแสดงคุณลักษณะของข้อสอบด้วยกราฟจะช่วยทำให้เกิดความเข้าใจได้เร็วขึ้นและสามารถทำความเข้าใจได้อย่างครอบคลุมทุกค่าพารามิเตอร์ของข้อสอบ (Camilli & Shepard, 1994: 47)

2.1.1 หลักการของทฤษฎีการตอบสนองข้อสอบ

หลักการของ IRT เป็นความสัมพันธ์ที่แสดงได้ด้วยโค้งคุณลักษณะของข้อสอบ(Item Characteristic Curve: ICC) ซึ่งเป็นฟังก์ชันที่แสดงถึงความสัมพันธ์ของโอกาสในการตอบถูกของข้อสอบบนความสามารถที่วัดได้ด้วยข้อสอบ แสดงได้ด้วยภาพที่ 3 โดยปกติคุณลักษณะแฝง (Trait) จะหมายถึงความสามารถ (Ability) จากภาพความสามารถที่แท้จริง (θ) แสดงด้วยแกน X ของกราฟและโอกาสในการตอบถูก (P(θ)) แสดงด้วยแกน Y ค่าของ P(θ) จะเพิ่มขึ้นเมื่อความสามารถเพิ่มขึ้น มีค่าอยู่ระหว่าง 0 ถึง 1 โค้ง ICC จะเป็นเส้นกราฟที่ปรับให้เรียบและต่อเนื่อง



ภาพที่ 3 โค้งคุณลักษณะของข้อสอบ

พารามิเตอร์ความสามารถของผู้สอบ (θ)

ความสามารถของผู้สอบแทนได้ด้วย θ ที่มีความหมายถึงความสามารถที่แท้จริงของผู้สอบและในบางครั้งมีความหมายถึงความสามารถแฝงที่อยู่ภายใต้การวัดด้วยแบบสอบ ค่า θ จะประมาณค่าสำหรับผู้สอบแต่ละคนในการตอบสนองข้อสอบ ดังนั้นพารามิเตอร์ความสามารถของผู้สอบจึงเปรียบได้กับคะแนนรวมของแบบสอบ ซึ่งไม่สามารถวัดได้โดยตรง การประมาณค่า θ จึงเป็นสเกลที่สร้างขึ้นและสามารถเปรียบเทียบความแตกต่างของค่า θ ได้อย่างมีความหมาย

พารามิเตอร์ความยากของข้อสอบ(b)

พารามิเตอร์ความยากของข้อสอบแทนได้ด้วย b และวัดได้ในหน่วยสเกลเดียวกันกับความสามารถของผู้สอบ ค่า b ที่มีค่าเป็นบวกและมีค่ามากจะหมายถึงข้อสอบที่ยาก ส่วนค่า b ที่เป็นลบและมีค่ามากจะหมายถึงข้อสอบที่ง่าย

พารามิเตอร์อำนาจจำแนกของข้อสอบ(a)

พารามิเตอร์อำนาจจำแนกของข้อสอบแทนด้วย a มีความหมายคล้ายกันกับค่าอำนาจจำแนกในทฤษฎีการวัดแบบดั้งเดิม นั่นคือเป็นดัชนีที่แสดงเห็นถึงการจำแนกระหว่างผู้สอบที่มีความสามารถสูงกับผู้สอบที่มีความสามารถต่ำ หรือผู้สอบที่มีความสามารถสูงควรจะมีโอกาสในการตอบถูกมากกว่าผู้สอบที่มีความสามารถต่ำ

พารามิเตอร์การเดาข้อสอบ (c)

พารามิเตอร์การเดาของข้อสอบแทนด้วย c เป็นค่าที่ไม่สามารถอธิบายได้ในทฤษฎีการทดสอบแบบดั้งเดิม ค่านี้จะแสดงให้เห็นให้ถึงผู้สอบที่มีความสามารถต่ำมีโอกาสที่จะตอบข้อสอบได้ถูก เมื่อเปรียบเทียบการประมาณค่าแล้วการประมาณค่า c จะทำได้ยากกว่าพารามิเตอร์ a และ b

2.1.2 ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ มีข้อตกลงเบื้องต้นที่สำคัญ 2 ประการคือ (Embreston & Reise, 2000)

1. โค้งลักษณะข้อสอบ(Item Characteristic curves: ICC) จะมีรูปแบบที่เป็นลักษณะเฉพาะ ซึ่งแสดงถึงการเปลี่ยนแปลงของระดับความสามารถที่สัมพันธ์กับการเปลี่ยนแปลงของความน่าจะเป็นของการตอบแต่ละข้อได้ถูกต้อง สำหรับข้อสอบที่ตรวจให้คะแนนสองค่าจะพิจารณาเป็นการตอบข้อสอบได้ถูก ส่วนข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า ลักษณะของโค้ง ICC เป็นความน่าจะเป็นของการตอบในแต่ละรายการคำตอบต่อระดับความสามารถของผู้สอบ ซึ่งโค้งลักษณะของข้อสอบโดยทั่วไปจะมีลักษณะเป็นโค้งรูปตัวเอส

2. ความเป็นอิสระของการตอบข้อสอบ (local independence) ซึ่งความเป็นอิสระของการตอบข้อสอบจะเกิดขึ้นเมื่อมีความสัมพันธ์ระหว่างข้อสอบหรือผู้สอบที่มีลักษณะที่สอดคล้องกับโมเดลการตอบสนองข้อสอบ ซึ่งมีความเป็นอิสระของการตอบข้อสอบต้องมีทั้งในแบบสอบที่เป็นเอกมิติ (unidimentional) และแบบสอบที่เป็นพหุมิติ (multidimentional)

นอกจากข้อตกลงเบื้องต้นสองประการดังกล่าวแล้ว Hambletonn& Swaminatan (1985) ได้เสนอข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกมิติของคุณลักษณะภายในที่มุ่งวัด (unidimensional latent space) ซึ่งหมายถึงคุณลักษณะภายในที่เป็นสิ่งที่กำหนดพฤติกรรมการตอบข้อสอบมีมิติเดียว (unidimentional) นั่นคือข้อสอบแต่ละข้อมีคุณสมบัติในการวัดคุณลักษณะเดียวกัน ซึ่งต่อมา Reckase (1983, 1985) และคณะ ได้พัฒนาโมเดล IRT แบบพหุมิติ (multidimentional IRT Models) ซึ่งมีข้อตกลงว่าการตอบสนองของผู้สอบแต่ละคนขึ้นอยู่กับคุณลักษณะแฝงที่มีอยู่อย่างหลากหลาย (Latent traits) จึงนำไปสู่การพัฒนาเพื่อการประมาณค่าในลักษณะพหุมิติ (multidimentional) ได้ (Weiss and Yoes, 1991)

2.1.3 โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนสองค่า (Dichotomous IRT Models)

1) โมเดลการตอบสนองแบบหนึ่งพารามิเตอร์

โมเดลนี้เป็นโมเดลที่มีพารามิเตอร์ของข้อสอบเพียงหนึ่งค่า คือพารามิเตอร์ความยากของข้อสอบ และตัวแปรตามเป็นแบบทวิภาค โมเดลนี้มีชื่อเรียกอีกอย่างว่า โมเดลราสช์ (Rasch model) นอกจากนี้ยังมีพารามิเตอร์ของผู้สอบซึ่งเป็นความสามารถของผู้สอบ โดยความสามารถของผู้สอบจะมีฟังก์ชันเชิงบวก และตัวแปรตามจะเป็นการตอบสนองข้อสอบของผู้สอบแต่ละคนต่อข้อสอบแต่ละข้อ ซึ่งจะมีค่าเป็นลอคออด (Log odds) หรือความน่าจะเป็น ส่วนตัวแปรอิสระจะเป็นคะแนนคุณลักษณะของแต่ละคน หรือความสามารถของผู้สอบแต่ละคน (θ_s) และระดับความยากของข้อสอบแต่ละข้อ(β_i) โดยค่าผลต่างของความสามารถของผู้สอบกับระดับความยากของข้อสอบจะมีค่าเป็นอัตราส่วนของลอคออด

สมการลอคออดของโมเดลราสช์จะมีตัวแปรเป็นลอกธรรมชาติของโอกาสในการตอบข้อสอบถูกต้อง โดยออดจะเท่ากับอัตราส่วนของจำนวนข้อสอบที่ตอบข้อสอบถูกต่อจำนวนข้อสอบที่ตอบผิด หรืออาจจะกำหนดให้มีค่าเท่ากับอัตราส่วนของความน่าจะเป็นของการตอบข้อสอบถูกต่อความน่าจะเป็นของการตอบข้อสอบผิด และในโมเดลราสช์ ค่าของลอคของอัตราส่วนออดจะเท่ากับผลต่างระหว่างระหว่างพารามิเตอร์ความสามารถของผู้สอบ และพารามิเตอร์ความยากของข้อสอบ ดังสมการ

$$\ln\left[\frac{p_{is}}{1-p_{is}}\right]=\theta_s-\beta_i$$

ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบถูก $(P(x_{is}=1))$ เมื่อผู้สอบมีความสามารถเป็น θ_s และความยากของข้อสอบเป็น β_i สามารถเขียนให้อยู่ในรูปฟังก์ชันโลจิสติก ได้ดังสมการ

$$P(x_{is}=1|\,\theta_s,\beta_i)=\frac{\exp(\theta_s-\beta_i)}{1+\exp(\theta_s-\beta_i)}$$

หรือสามารถเขียนได้รูปหนึ่งได้ดังสมการ

$$P_i(\theta)=\frac{1}{1+e^{-(\theta-b_i)}}$$

เมื่อ $P_i(\theta)$ แทนความน่าจะเป็นที่ผู้สอบที่มีความสามารถที่ θ จะตอบข้อสอบข้อที่ i ได้ถูก

b_i แทนค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i

e แทนลอกธรรมชาติ มีค่าเท่ากับ 2.718

2) โมเดลการตอบสนองแบบสองพารามิเ ตอร์

โมเดลการตอบสนองข้อสอบแบบสองพารามิเตอร์มีความสัมพันธ์ที่เขียนด้วยฟังก์ชันโลจิสติก ดังสมการ

$$P_i(\theta)=\frac{1}{1+e^{-Da_i(\theta-b_i)}}$$

เมื่อ $P_i(\theta)$ แทน ความน่าจะเป็นที่ผู้สอบที่มีระดับความสามารถที่ θ จะตอบข้อสอบข้อที่ i ได้ถูก

b_i แทน ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i

a_i แทน ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i

e แทนลอกธรรมชาติ มีค่าเท่ากับ 2.718

D เป็นค่าคงที่ มีค่าเท่ากับ 1.70

3) โมเดลการตอบสนองแบบสามพารามิเตอร์

โมเดลการตอบสนองข้อสอบแบบสองพารามิเตอร์มีความสัมพันธ์ที่เขียนด้วยฟังก์ชันโลจิสติก ดังสมการ

$$P_i(\theta) = c_i + \frac{(1-c_i)}{1+e^{-Da_i(\theta-b_i)}}$$

เมื่อ $P_i(\theta)$ แทน ความน่าจะเป็นที่ผู้สอบที่มีระดับความสามารถที่ θ จะตอบข้อสอบข้อที่ i ได้ถูก

b_i แทน ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i

a_i แทน ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i

c_i แทน ค่าพารามิเตอร์โอกาสในการเดาข้อสอบข้อที่ i ได้ถูก

e แทนลอกธรรมชาติ ซึ่งมีค่าเท่ากับ 2.718

D เป็นค่าคงที่ มีค่าเท่ากับ 1.70

2.1.4 โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่าสองค่า (Polytomous IRT Models)

โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่าสองค่ามีหลายโมเดล ในที่นี้จะนำเสนอโมเดลการตรวจให้คะแนนบางส่วน (Partial Credit Model) เนื่องจากเป็นโมเดลที่มีความสอดคล้องกับข้อมูลในการวิจัยในครั้งนี้ โมเดลนี้ได้พัฒนาโดยมาสเตอร์ (Master, 1982 อ้างในศิริชัย กาญจนวาสี, 2545) สำหรับใช้ในการวิเคราะห์ข้อสอบหรือข้อคำถามที่มีกระบวนการตอบหลายลำดับขั้น ซึ่งจำเป็นต้องมีการตรวจให้คะแนนการตอบถูกต้องหรือตอบถูกบางส่วนในแต่ละลำดับขั้นของกระบวนการตอบ เช่น ข้อสอบวัดผลสัมฤทธิ์ทางคณิตศาสตร์ ซึ่งมีการให้คะแนนคำตอบถูกบางส่วน โมเดล PCM มีลักษณะที่พัฒนาต่อจากโมเดลการตอบสนองข้อสอบหนึ่งพารามิเตอร์ จึงมีลักษณะพารามิเตอร์มาตรฐานคล้ายโมเดลของราสช์ (Rasch Model) และใช้หลักการคำนวณความน่าจะเป็นของการตอบแต่ละระดับขั้นการตอบโดยตรงแบบขั้นตอนเดียว (Direct IRT Method)

ลักษณะการตอบคำถามแต่ละข้อ (i) จะอธิบายได้ด้วยค่าพารามิเตอร์ระดับความยากระหว่างรายการคำตอบที่อยู่ถัดไป (item step difficulty, δ_{ij}) โดยทุกข้อ i ข้อมีค่าพารามิเตอร์

ความชันเท่ากัน (equal slope) และมีคะแนน x = 0, 1, 2,..., m_i โดยมีจำนวนรายการคำตอบ K_i = m_i + 1

ตัวอย่างการตอบที่มีรายการคำตอบเป็นคะแนน 0, 1, 2, 3 โดยการตอบจะมีขั้นตอนการตอบ 3 ขั้นตอน คือขั้นที่ 1 เป็นการตอบระหว่างรายการคำตอบ 0 และ 1 ขั้นที่ 2 เป็นการตอบระหว่างรายการคำตอบ 1 และ 2 ขั้นที่ 3 เป็นการตอบระหว่างรายการคำตอบ 2 และ 3 ค่า δ_{ij} ไม่ได้บอกถึงค่าความสามารถของผู้สอบ (θ) ที่มีโอกาส 0.50 ในการเลือกรายการคำตอบที่อยู่เหนือถัดไป แต่ค่า δ_{ij} แสดงถึงระดับความยากสัมพัทธ์ของแต่ละขั้นตอนภายในคำถามข้อเดียวกัน แต่ละขั้นตอนอาจง่ายกว่าหรือยากกว่าบางขั้นตอน δ_{ij} ยังแสดงถึงตำแหน่ง θ ที่โค้งการเลือกรายการคำตอบตัดกัน จึงบ่งบอกถึงผู้ตอบที่มีคุณลักษณะ θ จะมีโอกาสเลือกรายการคำตอบใดมากกว่ากัน

สำหรับ x = j รายการคำตอบ ฟังก์ชันการเลือกรายการคำตอบของโมเดลการตรวจให้คะแนนบางส่วน เป็นดังนี้

$$P_{ix}(\theta) = \frac{\exp[\sum_{j=0}^{x}(\theta - \delta_{ij})]}{\sum_{r=0}^{m_i}[\exp\sum_{j=0}^{r}(\theta - \delta_{ij})]}$$

เมื่อ $\sum_{j=0}^{0}(\theta - \delta_{ij}) = 0$

$P_{ix}(\theta)$ แทน ความน่าจะเป็นที่ผู้ตอบซึ่งมีความสามารถ θ จะตอบข้อ I ด้วยการเลือกหรือสามารถทำรายการคำตอบขั้นที่ x จากจำนวน m_i ขั้น

δ_{ij} แทน ค่าพารามิเตอร์ระดับความยากของขั้นการตอบที่ j ในข้อ I เมื่อ j = 1, 2, 3, ..., m_i ค่าที่สูงแสดงถึงขั้นการตอบนั้นมีความยากสัมพัทธ์สูงกว่าขั้นอื่น

2.2 มโนทัศน์เกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบ

การทำหน้าที่ต่างกันของข้อสอบ (Differential item functioning: DIF) เป็นเรื่องที่นักพัฒนาแบบสอบ นักวัดผลทางการศึกษาตลอดจนนักสถิติได้ให้ความสำคัญและพยายามที่จะพัฒนาแนวทาง เทคนิควิธีการทางสถิติ เพื่อตรวจสอบหาข้อสอบที่ทำหน้าที่ต่างกันในกลุ่มผู้สอบต่างกัน ทั้งนี้เพื่อทำให้ผลการสอบมีความยุติธรรมแก่กลุ่มผู้สอบได้มากที่สุด

2.2.1 ความเป็นมาของการทำหน้าที่ต่างกันของข้อสอบ

แบบสอบจะประกอบด้วยข้อสอบที่ถามในเนื้อหาที่ต้องการให้ผู้สอบได้แสดงสิ่งที่รู้ตามกรอบของการวัด ดังนั้นแหล่งของความยากจึงขึ้นอยู่กับการสร้างข้อสอบ และอาจจะมีส่วนที่ไม่มีความตรงต่อกลุ่มผู้สอบ เช่นการสร้างข้อสอบที่ใช้คำศัพท์ที่ผู้สอบแต่ละกลุ่มอาจจะแปลความหมายได้แตกต่างกัน ดังนั้นคะแนนที่ได้จากการทำข้อสอบจึงเกิดจากข้อจำกัดของความหมายของคำหรือภาษาที่ใช้ในข้อสอบ (Camilli & Shepard, 1994: 1) วัตถุประสงค์ของการดำเนินการเกี่ยวกับความลำเอียงของข้อสอบจะเป็นการตรวจสอบหรือการแสดงให้เห็นถึงความตรงของข้อสอบ ซึ่งสิ่งสำคัญ

ของความลำเอียงในแบบสอบคือความตรงเชิงโครงสร้าง Osterlind (1987: 10-11) ได้กล่าวถึงข้อสอบมีความตรงเชิงโครงสร้างว่าจะต้องเป็นข้อสอบที่มีโครงสร้างเดียวกันต่อผู้สอบทุกคน นั่นคือผู้สอบมีโอกาสในการทำข้อสอบถูกได้เท่ากัน โดยความลำเอียงของข้อสอบจะเปรียบเทียบความน่าจะเป็นในการตอบข้อสอบถูกของกลุ่มผู้สอบที่ต่างกันในกลุ่มประชากรเดียวกัน และข้อสอบที่ไม่มีความลำเอียง เกิดขึ้นเมื่อมีความน่าจะเป็นเท่ากันในการตอบข้อสอบถูก สำหรับผู้สอบต่างกลุ่มและมีความสามารถเท่ากัน

ความลำเอียงของแบบสอบในระยะแรกเกิดจากการสังเกตเห็นถึงความแตกต่างของค่าเฉลี่ยของเชาวน์ปัญญา(IQ) ระหว่างกลุ่มที่มีเชื้อชาติหรือศาสนาต่างกัน (Gregory, 2007: 265) ซึ่งสังเกตได้จากคนผิวดำมีคะแนนเฉลี่ยต่ำกว่าคนผิวขาว นั่นคือการสอบในแบบสอบวัด IQ ซึ่งเป็นแบบสอบมาตรฐาน พบคะแนนสอบแตกต่างกันเนื่องจากความแตกต่างของเชื้อชาติหรือศาสนา และในปี 1910 Alfred Binet จากการทดสอบเด็กที่มาจากชนชั้นทางเศรษฐกิจต่ำ พบว่าการวัดด้วยข้อสอบบางข้ออาจจะเกิดจากการฝึกฝนที่บ้านหรือสถานศึกษา มากกว่าจะเป็นความสามารถทางสติปัญญา ต่อมาในปี 1912 William Stern ได้ค้นหาความแตกต่างที่เกิดขึ้นพบว่ามีลักษณะคล้ายกับ Binet และต้นปี 1951 ได้เริ่มมีการวิจัยเกี่ยวกับความลำเอียงของข้อสอบ โดย Eells, Davis, Havighurst, Herrick , Tyler และนักวิจัยจำนวนมากรวมทั้ง Binet และ Stern ได้แสดงให้เห็นถึงความสัมพันธ์ระหว่าง IQ กับชนชั้นทางสังคม การศึกษาเกี่ยวกับสิ่งแวดล้อมที่เป็นสาเหตุของความแตกต่างที่สามารถสังเกตได้ มีวัตถุประสงค์เพื่อวัดความแตกต่างที่ไม่ได้สะท้อนความสามารถที่แท้จริง เช่น ความแตกต่างที่สังเกตได้ใน IQ ของเด็กที่ขึ้นอยู่กับเนื้อหาเฉพาะในข้อสอบ ดังนั้นวัตถุประสงค์หลักของการวิจัยจึงเป็นการอธิบายความแตกต่างของกลุ่มผู้สอบเนื่องจากการทำข้อสอบแต่ละข้อในแบบสอบ ต่อมาในปลายปี 1960 และ 1970 มีการศึกษาความลำเอียงของข้อสอบในแบบสอบเพื่อแสดงให้เห็นถึงความลำเอียงของแบบสอบวัด IQ ที่มีการใช้ในการศึกษา โดยมีวัตถุประสงค์เพื่อการจัดอันดับทางการศึกษา และการคัดเลือกเข้าทำงาน ในปี 1971 บริษัท Grigg V.Power พบว่าผลจากการสอบไม่มีความเหมาะสมเนื่องจากไม่ได้ลูกจ้างที่มีความตรงกับความสามารถในการทำงาน ในปี 1969 Jensen ได้มีการค้นหาความลำเอียงของข้อสอบ และชี้ให้เห็นว่า IQ มีความแตกต่างกันมากเนื่องจากอิทธิพลของสิ่งแวดล้อม ต่อมาจึงให้ความสำคัญต่อการนำไปสู่การปฏิบัติมากขึ้น นักวิจัยจึงสนใจที่จะแก้ปัญหาเกี่ยวกับการใช้ผลการสอบเพื่อตัดสินการจัดลำดับหรือการคัดเลือกเข้าทำงาน ตลอดจนการค้นหาวิธีการทางสถิติเกี่ยวกับความลำเอียง และการสร้างแบบสอบที่มีความเป็นอิสระทางวัฒนธรรมเพื่อไม่ให้มีผลกระทบต่อชนกลุ่มน้อย

ตั้งแต่ปี 1960 ได้มีการใช้คำว่า ความลำเอียงของข้อสอบ (Item Bias) (Angoff, 1993) มีการศึกษาเพื่อทำให้เกิดความยุติธรรมของแบบสอบและเพื่อลดความไม่เสมอภาคของแบบสอบระหว่างผู้สอบต่างกลุ่มกันที่เป็นกลุ่มย่อยของประชากร และทำความเข้าใจความแตกต่างทางวัฒนธรรมระหว่างคนผิวดำและชนกลุ่มน้อยที่แสดงถึงความไม่เท่าเทียมกันจากคะแนนแบบสอบที่เกิดจากความลำเอียงในข้อสอบของแบบสอบมากกว่าระดับความสามารถในกลุ่มย่อยของผู้สอบ เป้าหมายการศึกษาโดยทั่วไปเป็นการระบุข้อสอบในแบบสอบที่แสดงถึงความลำเอียงต่อกลุ่มผู้สอบ

กลุ่มใดกลุ่มหนึ่ง Angoff (1993) ได้กล่าวถึงความลำเอียงของข้อสอบที่นักการศึกษาได้นิยามในแนวทางเดียวกันว่า “ข้อสอบจะมีความลำเอียงถ้าความสามารถของแต่ละคนเท่ากัน จากกลุ่มที่ต่างกัน แต่ไม่มีความเท่าเทียมกันในโอกาสของการตอบข้อสอบถูกต้อง” และ Camilli & Shepard (1994: 7-8) ได้นิยามถึงความลำเอียงของแบบสอบว่าเป็นความไม่ตรงหรือความคลาดเคลื่อนอย่างเป็นระบบในแบบสอบที่ต้องการวัดของเฉพาะกลุ่มใดกลุ่มหนึ่ง และความลำเอียงนี้ไม่ได้เป็นผลเนื่องจากความคลาดเคลื่อนอย่างสุ่มของการวัด เนื่องจากความคลาดเคลื่อนอย่างสุ่มของการวัดจะมีลักษณะเดียวกันต่อกลุ่มผู้สอบทุก และการศึกษาส่วนใหญ่จะแบ่งกลุ่มผู้สอบที่แตกต่างกันในด้านเชื้อชาติ เพศ ชนชั้นทางสังคม อายุ ศาสนา ตำแหน่งที่ตั้งของสถานศึกษา

คำว่า ”ความลำเอียง” ได้นิยามเป็นสองประเด็นคือด้านสถิติและด้านสังคม โดยในด้านสถิติจะหมายถึงการโน้มเอียงของการประมาณค่าที่หันเหไปในทิศทางหนึ่งหรือไปในทิศทางอื่นที่ต่างจากค่าที่เป็นจริง ส่วนนิยามในด้านสังคม เป็นการให้ความหมายในบางอย่างที่เป็นการพิพากษ์ตัดสินและเป็นการโอนเอียงของอารมณ์อย่างไม่มีเหตุผล จากความหมายทางสังคมและสถิติที่มีความแตกต่างกันเพียงเล็กน้อย ดังนั้นเพื่อไม่ให้เกิดความสับสนของความลำเอียงที่ทำให้เกิดความไม่เท่าเทียมกันของคะแนนสอบ จึงมีข้อเสนอให้ใช้คำอื่นแทนเพื่อให้สามารถสังเกตได้โดยใช้สถิติ และใช้ในการตัดสินหรือแปลความหมายเพื่อประโยชน์ต่อการนำไปใช้ จึงใช้คำว่า “การทำหน้าที่ต่างกันของข้อสอบ” (Differential item functioning: DIF) ที่แสดงถึงความต่างกันของสมบัติทางสถิติในผู้สอบต่างกลุ่มกัน เมื่อมีการควบคุมความสามารถของผู้สอบ (Angoff, 1993: 4-5)

2.2.2 ความหมายของการทำหน้าที่ต่างกันของข้อสอบ

การให้ความหมายของการทำหน้าที่ต่างกันของข้อสอบรวมทั้งความหมายของความลำเอียงของข้อสอบ ได้มีนักวัดผลหลายท่านได้ให้ความหมายไว้ดังนี้

ความลำเอียง หมายถึง ความคลาดเคลื่อนที่เป็นระบบในกระบวนการวัดผล ส่งผลให้เกิดการเปลี่ยนแปลง โดยในบางครั้งอาจจะมีเพิ่มขึ้นหรืออาจจะลดลง ข้อสอบข้อหนึ่งจะถูกตัดสินว่าเป็นข้อสอบที่มีความลำเอียงเมื่อความยากของข้อสอบมีมากกว่าหรือน้อยกว่าต่อกลุ่มผู้สอบกลุ่มหนึ่งเมื่อเปรียบเทียบกับอีกกลุ่มหนึ่งที่มาจากประชากรเดียวกัน (OsterLind, 1987)

ความลำเอียงของข้อสอบ หมายถึง ความเท่าเทียมกันของความน่าจะเป็นในการตอบข้อสอบได้ถูกต้อง เมื่อผู้สอบที่มาจากต่างกลุ่มกันมีความสามารถเท่ากัน (Angoff, 1993)

การทำหน้าที่ต่างกันของข้อสอบ หมายถึง สิ่งที่สังเกตได้ของข้อสอบข้อหนึ่งที่แสดงถึงความแตกต่างของคุณสมบัติทางสถิติระหว่างกลุ่มผู้สอบสองกลุ่มที่กำหนดให้มีความสามารถที่ระดับเดียวกัน (Holland & Wainer, 1993)

การทำหน้าที่ต่างกันของข้อสอบ หมายถึง ความแตกต่างทางจิตวิทยาระหว่างกลุ่มผู้สอบที่จับคู่ความสามารถหรือผลสัมฤทธิ์ที่วัดด้วยข้อสอบ ข้อสอบที่ทำหน้าที่ต่างกันจะเกิดการได้เปรียบเสียเปรียบของการวัดระหว่างกลุ่มผู้สอบ เนื่องจากความแตกต่างในคุณลักษณะที่ไม่ได้ต้องการวัด เช่น การแปลงภาษาของข้อสอบ หรือเมื่อควบคุมระดับความสามารถของผู้สอบแล้ว ผู้สอบที่ต่างกลุ่มกันจะมีโอกาสในการตอบข้อสอบถูกแตกต่างกัน (Dorans & Holland, 1993)

การทำหน้าที่ต่างกันของข้อสอบ หมายถึง ความแตกต่างของกลุ่มในฟังก์ชันทางสถิติของข้อสอบระหว่างผู้สอบแต่ละคนที่มีการจับคู่ในดัชนีของความสามารถที่วัดโดยแบบสอบ (Koretz & McCaffrey, 2005)

การทำหน้าที่ต่างกันของข้อสอบ หมายถึง สิ่งที่เกิดขึ้นจากกลุ่มผู้สอบที่มาจากต่างกลุ่มกันที่มีระดับของคุณลักษณะเดียวกันมีความแตกต่างของโอกาสในการตอบถูก (Kamata, 2003)

การทำหน้าที่ต่างกันของข้อสอบ หมายถึง การที่ข้อสอบทำให้ผู้สอบจากต่างกลุ่มกันที่มีความสามารถหรือคุณลักษณะที่มุ่งวัดเท่ากัน มีโอกาสในการตอบข้อสอบได้ถูกต้องแตกต่างกัน หรือมีฟังก์ชันการตอบสนองข้อสอบแตกต่างกัน การทำหน้าที่ต่างกันของข้อสอบเกิดขึ้นเมื่อนำข้อสอบไปทดสอบกับผู้สอบกลุ่มย่อยต่างกัน ที่มีความสามารถหลัก (Primary ability) ระดับเดียวกันหรือมีคุณลักษณะแฝง (Latent trait) ที่ต้องการวัดเท่ากัน แต่มีความสามารถรอง (Secondary ability) แตกต่างกัน ทำให้ผู้สอบต่างกลุ่มที่นำมาจับคู่เปรียบเทียบมีโอกาสตอบข้อสอบถูกแตกต่างกัน (ศิริชัย กาญจนวาสี, 2545)

จากความหมายของการทำหน้าที่ต่างกันของข้อสอบจะพบว่าส่วนที่พิจารณาคล้ายกัน คือมีการควบคุมระดับความสามารถของกลุ่มผู้สอบ เพื่อใช้เปรียบเทียบผลการตอบสนองของแต่ละกลุ่มผู้สอบ ดังนั้นความหมายของการทำหน้าที่ต่างกันของข้อสอบจึงสามารถให้ความหมายได้เป็น “การแสดงออกของข้อสอบที่มีความแตกต่างกันโดยการใช้เทคนิคทางสถิติเพื่อตรวจสอบถึงผลการตอบของกลุ่มผู้สอบต่างกลุ่มกัน เมื่อมีการควบคุมความสามารถหรือคุณลักษณะแฝงของผู้สอบแต่ละกลุ่มให้เท่ากัน”

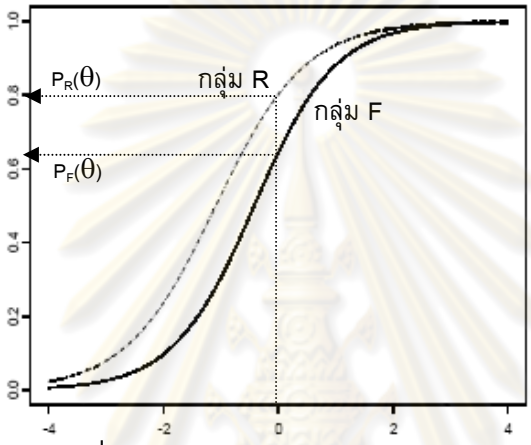
2.2.3 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบตามแนวคิดทฤษฎีการตอบสนองข้อสอบ

การใช้แนวคิดของโมเดลของทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) เพื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบมีจุดเด่นมากกว่าทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) ในหลายด้าน ประการแรก การประมาณค่าพารามิเตอร์ในโมเดล IRT มีข้อจำกัดในด้านขนาดกลุ่มตัวอย่างน้อยกว่าทฤษฎีทดสอบแบบดั้งเดิม ประการที่สอง คุณสมบัติทางสถิติสามารถอธิบายได้อย่างแม่นยำมากกว่าทฤษฎีทดสอบแบบดั้งเดิม เมื่อการทำหน้าที่ของข้อสอบมีความแตกต่างในกลุ่มผู้สอบสองกลุ่ม นอกจากนี้ยังใช้กราฟในการแปลความหมายได้ ทำให้เกิดความเข้าใจได้ในการแสดงให้เห็นถึงการทำหน้าที่ต่างกันของข้อสอบ (Camilli & Shepard, 1994: 47)

การแปลความหมายการทำหน้าที่ต่างกันของข้อสอบด้วยกราฟ ตามหลักการของ IRT เป็นการพิจารณาโค้งคุณลักษณะของข้อสอบ (Item Characteristic Curve: ICC) ซึ่งเป็นฟังก์ชันของความน่าจะเป็นการตอบข้อสอบถูกกับความสามารถของผู้สอบ ในการตรวจสอบ DIF จะเปรียบเทียบโค้ง ICC ของกลุ่มผู้สอบสองกลุ่มในข้อสอบข้อเดียวกัน ความแตกต่างของ ICC ระหว่างผู้สอบสองกลุ่มจะเป็นการแสดงให้เห็นถึงผู้สอบในกลุ่มเป้าหมาย (Focal Group: F) และ

กลุ่มอ้างอิง (Reference Group: R) ที่ระดับความสามารถเดียวกันแต่มีโอกาสไม่เท่ากันในการตอบข้อสอบถูกในข้อเดียวกัน แนวคิดของการทำหน้าที่ต่างกันของข้อสอบจึงเป็นความแตกต่างของ ICC ของผู้สอบสองกลุ่ม พิจารณา 2 รูปแบบ ดังนี้

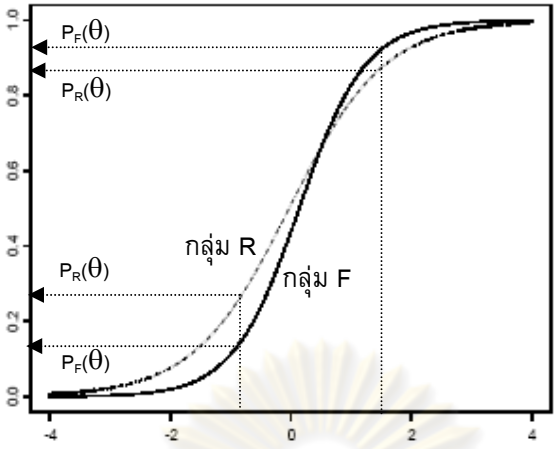
1. การเกิด DIF แบบเอกรูป (Uniform DIF) หรือความคงที่ของการเกิด DIF (Consistant DIF) การเกิด DIF ที่มีลักษณะคงที่จะเกิดขึ้นเมื่อ ICC ของสองกลุ่มมีความแตกต่างกันและเส้น ICC ทั้งสองเส้นไม่ตัดกัน หรือไม่มีปฏิสัมพันธ์กันระหว่างระดับความสามารถของผู้สอบกับกลุ่มผู้สอบ แสดงให้เห็นถึงกลุ่มผู้สอบกลุ่มหนึ่งได้เปรียบหรือมีโอกาสตอบถูกมากกว่าผู้สอบอีกกลุ่มทุกระดับความสามารถ แสดงได้ดังภาพที่ 4



ภาพที่ 4 การเกิด DIF แบบเอกรูป

จากภาพที่ 4 จะแสดงให้เห็นถึงการเกิด DIF ของกลุ่มผู้สอบสองกลุ่มระหว่างกลุ่ม R กับกลุ่ม F จากกราฟเส้น ICC ของกลุ่ม R จะอยู่เหนือเส้น ICC ของกลุ่ม F แสดงว่าผู้สอบในกลุ่ม R มีโอกาสตอบข้อสอบข้อนี้ได้ถูกมากกว่ากลุ่ม F หรือข้อสอบข้อนี้ง่ายสำหรับกลุ่ม R ส่วนค่าอำนาจจำแนก(a) ของกลุ่ม R จะมีค่าเท่ากันกับกลุ่ม F

2. การเกิด DIF แบบอเนกรูป (Nonuniform DIF) หรือความไม่คงที่ของการเกิด DIF (Inconsistant DIF) การเกิด DIF ที่มีลักษณะไม่คงที่ โค้ง ICC ของผู้สอบสองกลุ่มมีความแตกต่างกัน และเส้น ICC ทั้งสองเส้นมีจุดตัดกันบางจุดที่ระดับความสามารถหนึ่ง หรือมีปฏิสัมพันธ์ร่วมกันระหว่างระดับความสามารถของผู้สอบกับกลุ่มผู้สอบ ดังนั้นการเกิด DIF จึงเกิดขึ้นได้ทั้งส่วนได้เปรียบ และเสียเปรียบของกลุ่มผู้สอบที่ตำแหน่งความสามารถต่างๆ แสดงได้ดังภาพที่ 5



ภาพที่ 5 การเกิด DIF แบบอเนกรูป

จากภาพที่ 5 จะแสดงให้เห็นถึงการเกิด DIF ของกลุ่มผู้สอบสองกลุ่มระหว่างกลุ่ม R และกลุ่ม F จากกราฟเส้น ICC ของทั้งสองกลุ่มมีค่าพารามิเตอร์ a และ b แตกต่างกัน จากการตัดกันของเส้น ICC แสดงได้ว่าที่ระดับความสามารถต่ำ ผู้สอบกลุ่ม R มีโอกาสตอบข้อสอบถูกได้สูงกว่าผู้สอบกลุ่ม F แต่ที่ระดับความสามารถสูงขึ้นกลุ่ม R มีโอกาสตอบข้อสอบถูกมากกว่ากลุ่ม F

2.2.4 วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

ในระยะเริ่มแรก แนวความคิดของความลำเอียงของข้อสอบเกิดจากความยากที่แตกต่างกันของข้อสอบ นั่นคือโอกาสที่จะเกิดความแตกต่างในค่าเฉลี่ยของกลุ่มอาจจะเป็นผลของความลำเอียงหรือเป็นความแตกต่างในการสร้างสิ่งที่ต้องการประเมิน หรืออาจจะเกิดจากทั้งสองอย่าง ดังนั้นนักการวิจัยในระยะแรกจึงใช้วิธีการทางสถิติสำหรับการสำรวจความลำเอียงของข้อสอบโดยการพิจารณาความแตกต่างของความสามารถของกลุ่ม (Camilli & Shepard, 1994: 23) วิธีการตรวจสอบการทำหน้าที่ต่างกันองข้อสอบมีหลายวิธี แต่ยังพบว่าปัญหาและข้อจำกัดในการวิเคราะห์ในแต่ละวิธี จึงทำให้เกิดการพัฒนาวิธีการเพื่อตรวจสอบ DIF ให้มีศักยภาพยิ่งขึ้น

เมื่อพิจารณาวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยแบ่งกลุ่มตามข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า มีวิธีการต่างๆ ดังนี้

1) วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ให้ตรวจคะแนนสองค่า (Dichotomous Item)

1.1) วิธีการของการวิเคราะห์ความแปรปรวน (ANOVA)

1.2) ดัชนีการแปลงความยากของข้อสอบ (Transformed item difficultly index: TID) หรือวิธีการเดลต้าพลอต (Delta plot method)

1.3) วิธีดัชนีมาตรฐาน (Standardization: STAND)

1.4) วิธีแมนเทล-แฮนส์เซล (Mantel-Haenszel: M-H)

1.5) วิธีการถดถอยโลจิสติก (Logistic Regression: LR)

1.6) วิธี SIBTEST(Simulataneous Item Bias: SIBTEST)

1.7) วิธีวิเคราะห์ข้อสอบพหุระดับ(Multilevel item analysis method)

2) วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า (Polytomous Item)

2.1) วิธีแมนเทล-แฮนส์เซลทั่วไป(Generalized Mantel-Haenszel: GMH)

2.2) วิธีคอครัน-แมนเทล-แฮนส์เซล(Cochran-Mantel-Haenszel: CMH)

2.3) วิธีการวัดความแตกต่างของพื้นที่ (Area measures)

2.4) วิธีทดสอบ Wald statistics หรือวิธีการทดสอบไค-สแควร์ของ Lord

2.5) วิธี polytomous SIBTEST

2.6) วิธีวิเคราะห์ข้อสอบพหุระดับ(Multilevel item analysis method)

วิธีการตรวจสอบ DIF มีการศึกษาอย่างหลากหลาย ซึ่งมีหลักการและแนวคิดพื้นฐานและมีข้อจำกัดของแต่ละวิธีการตรวจสอบ ผู้วิจัยจึงสรุปแนวคิดที่ตามทฤษฎีที่เกี่ยวข้อง ผู้เสนอและปีที่พัฒนา ข้อสอบที่ใช้ศึกษา สถิติที่ใช้ การใช้สถิติทดสอบ และการแปลความของการตรวจสอบของแต่ละวิธี ดังตารางต่อไปนี้

ตารางที่ 1 สรุปหลักการและแนวคิดพื้นฐานของวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ตรวจให้คะแนนสองค่า (Dichotomous Item)

วิธีการตรวจสอบ	ผู้พัฒนาและปีที่พัฒนา	ข้อสอบที่ใช้ศึกษา	สถิติที่ใช้	การทดสอบนัยสำคัญทางสถิติ	การแปลความของการเกิด DIF
ANOVA	Cleary & Hilton(1968)	Dichotomous	พาราเมทริกซ์	สถิติทดสอบ t test	ความแตกต่างของแหล่งอิทธิพล
TID/Delta plot	Angoff(1972)	Dichotomous	นันพารา-เมทริกซ์	ไม่ใช้สถิติทดสอบ	ระยะห่างจากเส้นแกนหลัก
STAND	Dorans & Kulick (1983)	Dichotomous	นันพารา-เมทริกซ์	ไม่ใช้สถิติทดสอบ	ความแตกต่างของคะแนนที่คาดหวังของกลุ่มผู้สอบ
MH	Holland & Thayer(1988)	Dichotomous	นันพารา-เมทริกซ์	สถิติทดสอบ MH χ^2	อัตราส่วนของลอคออด
วิธีของ Lord	Lord(1980)	Dichotomous	พาราเมทริกซ์	สถิติทดสอบ d	ความแตกต่างของความยากของข้อสอบ

ตารางที่ 1 สรุปหลักการและแนวคิดพื้นฐานของวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ตรวจให้คะแนนสองค่า (Dichotomous Item)(ต่อ)

วิธีการตรวจสอบ	ผู้พัฒนาและปีที่พัฒนา	ข้อสอบที่ใช้ศึกษา	สถิติที่ใช้	การทดสอบนัยสำคัญทางสถิติ	การแปลความของการเกิด DIF
AREA	Millsap & Everson(1993)	Dichotomous	พาราเมทริกซ์	สถิติทดสอบตามวิธีที่ใช้	ความแตกต่างของพื้นที่ใต้โค้ง
Lord $\chi 2$	Lord(1980)	Dichotomous	พาราเมทริกซ์	สถิติทดสอบ $\chi 2$	พารามิเตอร์ของข้อสอบสามพารามิเตอร์
SIBTEST	Shealy & Stout (1993)	Dichotomous	นันพารา-เมทริกซ์	สถิติทดสอบของ SIB	ความแตกต่างของค่าเฉลี่ยของคะแนนของกลุ่มผู้สอบ

ตารางที่ 2 สรุปหลักการและแนวคิดพื้นฐานของวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า (Polytomous Item)

วิธีการตรวจสอบ	ผู้พัฒนาและปีที่พัฒนา	ข้อสอบที่ใช้ศึกษา	สถิติที่ใช้	การทดสอบนัยสำคัญทางสถิติ	การแปลความของการเกิด DIF
PolySIBTEST	Shealy & Stout (1993)	Polytomous	นันพารา-เมทริกซ์	สถิติทดสอบของ SIB	ความแตกต่างของค่าเฉลี่ยของคะแนนของกลุ่มผู้สอบ
Multilevel item analysis method	Kamata (1998)	Dichotomous + polytomous	พาราเมทริกซ์	สถิติทดสอบในโมเดลที่ใช้	อิทธิพลของข้อสอบในโมเดลแต่ละระดับ
PolySTAND	Dorans & Schmitt(1993)	Polytomous	นันพารา-เมทริกซ์	ไม่ใช้สถิติทดสอบ	ความแตกต่างของคะแนนที่คาดหวังของกลุ่มผู้สอบ
GMH	Holland & Thayer(1988)	Polytomous	นันพารา-เมทริกซ์	สถิติทดสอบ GMH $\chi 2$	อัตราส่วนของลอคออด
CMH	Meyer, Huynh, & Seaman,(2004)	Dichotomous +Polytomous	นันพารา-เมทริกซ์	สถิติทดสอบ CMH $\chi 2$	อัตราส่วนของลอคออด
LR	Swaminathan & Rogers(1990)	Dichotomous + Polytomous	พาราเมทริกซ์	สถิติทดสอบ $\chi 2$	อัตราส่วนของลอคออดหรืออัตราส่วนโลจิส

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบมีหลายวิธี แต่ละวิธีมีลักษณะเฉพาะและข้อจำกัดที่ต่างกัน ซึ่งเมื่อพิจารณาวิธีการวิเคราะห์โดยการประยุกต์ใช้การวิเคราะห์พหุระดับ (Multilevel item analysis) เป็นวิธีที่นำแนวคิดของโมเดล IRT เป็นโมเดลที่สามารถประมาณค่าความสามารถของผู้สอบ และค่าพารามิเตอร์ของข้อสอบ และมีจุดเด่นเกี่ยวกับความไม่แปรเปลี่ยนค่าพารามิเตอร์ข้ามกลุ่มตัวอย่าง ดังนั้นในการใช้ค่าประมาณของพารามิเตอร์จึงจะช่วยทำให้การนำผลไปใช้ในการวิเคราะห์จะได้ค่าที่ถูกต้อง รวมทั้งการวิเคราะห์ที่คำนึงถึงลักษณะของข้อมูลในการดำเนินการสอบโดยธรรมชาติที่มีลักษณะลดหลั่น นั่นคือ ข้อสอบสอดแทรกจากการทำข้อสอบของนักเรียนแต่ละคน และนักเรียนแต่ละคนก็สอดแทรกอยู่ในสถานศึกษา โดยกรอบแนวคิดของโมเดล HGLM จะกำหนดให้พารามิเตอร์ความยากของข้อสอบมีค่าคงที่สำหรับผู้สอบแต่ละกลุ่ม ซึ่งจัดเป็นจุดเด่นของโมเดลนี้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เนื่องจากข้อสอบแต่ละข้อในแบบสอบควรมีความยากคงที่หรือไม่ผันแปรของค่าพารามิเตอร์ตามกลุ่มผู้สอบ และระบุข้อสอบที่ทำหน้าที่ต่างกันได้อย่างถูกต้องแม่นยำยิ่งขึ้น นอกจากนี้การจัดข้อมูลแบบลดหลั่นจะทำให้ทราบถึงคุณลักษณะของตัวแปรแต่ละระดับ ที่จะช่วยอธิบายแหล่งของการเกิดการทำหน้าที่ของข้อสอบได้ ดังนั้นการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ในงานวิจัยนี้จึงเลือกใช้โมเดล HGLM ซึ่งจะเสนอรายละเอียดการวิเคราะห์ในตอนต่อไป

ตอนที่ 3 มโนทัศน์เกี่ยวกับการวิเคราะห์พหุระดับและการประยุกต์ใช้โมเดลเชิงเส้นแบบลดหลั่น

3.1 แนวคิดพื้นฐานของการวิเคราะห์พหุระดับ

การวิจัยทางสังคมศาสตร์ส่วนมากเกี่ยวข้องกับโครงสร้างข้อมูลแบบลดหลั่น นักวิจัยจึงศึกษาถึงคุณลักษณะในหน่วยของบุคคลและหน่วยขององค์กรหรือบริษัท ตัวแปรที่ศึกษาจึงมีลักษณะที่เป็นระดับชั้น ข้อมูลมีลักษณะเป็นโครงสร้างแบบลดหลั่นของรายบุคคลที่สอดแทรกอยู่ในหน่วยงาน (Raudenbush & Bryk, 2002) ดังนั้นจึงควรพิจารณาข้อมูลที่มีลักษณะเป็นโครงสร้าง ที่สามารถจัดเป็นระดับต่างๆ ของที่หน่วยศึกษา เช่น นักเรียนที่สอดแทรกอยู่ในสถานศึกษา และสถานศึกษาสอดแทรกอยู่ในจังหวัดหรือเขตพื้นที่ เป็นต้น

การวิเคราะห์พหุระดับเป็นวิธีการทางสถิติที่มีการใช้อย่างกว้างขวางมากขึ้นในทางสังคมศาสตร์ เนื่องจากเป็นการใช้สถิติที่มีหลักการยืดหยุ่นในการดำเนินการการวิจัย มีลักษณะที่เหมาะสมกับความเป็นจริงที่เป็นความสัมพันธ์ระหว่างตัวแปร นั่นคือเป็นโมเดลที่สามารถศึกษาได้ทั้งอิทธิพลหลัก (main effect) และอิทธิพลปฏิสัมพันธ์ (interaction effect) ของตัวแปรต้นที่ศึกษาในระดับที่ต่างกันของโครงสร้างแบบลดหลั่นของตัวแปรผลลัพธ์ ความยืดหยุ่นได้ของความสัมพันธ์ระหว่างตัวแปรที่นำไปสู่ข้อได้เปรียบหลายอย่าง เช่นโมเดลเชิงเส้นแบบลดหลั่น (HLM) จะประมาณค่าพารามิเตอร์ได้อย่างความถูกต้องของความคลาดเคลื่อนมาตรฐาน มีช่วงของความเชื่อมั่นที่ถูกต้องแม่นยำมากขึ้น

ข้อมูลที่ใช้ศึกษาในโมเดลพหุระดับมีลักษณะข้อมูลหลายระดับเป็นลำดับชั้น แสดงถึงความสัมพันธ์ของตัวแปรแต่ละระดับทั้งความสัมพันธ์ของตัวแปรที่อยู่ในระดับเดียวกัน และความสัมพันธ์ของตัวแปรที่ข้ามระดับกัน ตัวแปรที่อยู่ระดับสูงกว่ามีแนวโน้มที่จะส่งผลทางตรงหรือทางอ้อมต่อตัวแปรที่อยู่ระดับต่ำกว่า ความสัมพันธ์ระหว่างตัวแปรข้ามระดับมีหลายลักษณะ (ศิริชัย กาญจนวาสี, 2550) ดังนี้ 1) ความสัมพันธ์เชิงบริบท ตัวแปรที่อยู่ระดับสูงกว่าอาจส่งผลหรือมีความสัมพันธ์กับตัวแปรที่อย่ระดับต่ำกว่าในรูปของความสัมพันธ์เชิงสภาวะแวดล้อม 2) ความสัมพันธ์เชิงตัวแบบ ความสัมพันธ์ระหว่างตัวแปรที่อยู่ระดับสูงกว่าจะสร้างสภาวะตัวแบบแล้วสะท้อนภาพลงมายังตัวแปรระดับล่าง 3) ความสัมพันธ์เชิงถ่ายโยง โดยตัวแปรระดับบนลักษณะนี้จะสร้างเงื่อนไขหรือสภาวะแรงกระตุ้นถ่ายโยงสู่การปฏิบัติการของตัวแปรระดับล่าง เพื่อให้บรรลุผลตามเป้าหมายที่พึงปรารถนา และ 4) ความสัมพันธ์เชิงจูงใจ ตัวแปรที่อยู่ระดับสูงกว่ามีการส่งเสริมสนับสนุนในรูปของการสร้างแรงจูงใจ

3.2 การวิเคราะห์แบบพหุระดับ

3.3.1 การวิเคราะห์ตัวแปร 2 ระดับ

ในการวิเคราะห์ข้อมูล 2 ระดับ กำหนดให้มีข้อมูลจาก j กลุ่ม ซึ่งแต่ละกลุ่มมีจำนวนสมาชิกที่ต่างกัน n_i คนในแต่ละกลุ่ม กำหนดให้ Y_{ij} เป็นตัวแปรตามของคนที่ i ในกลุ่ม j และ X_{ij} เป็นตัวแปรทำนายของคนที่ i ในกลุ่ม j และ Z_j เป็นตัวแปรทำนายระดับกลุ่ม j

เขียนโมเดลการถดถอย 2 ระดับได้ดังนี้

<u>ระดับที่ 1</u> : ระดับบุคคล

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + r_{ij}$$

<u>ระดับที่ 2</u> : ระดับหน่วย

$$\beta_{oj} = G_{00} + G_{01} Z_j + u_{oj}$$

$$\beta_{1j} = G_{10} + G_{11} Z_j + u_{1j}$$

โมเดลการวิเคราะห์ 2 ระดับสามารถเขียนรวมเป็นโมเดลการถดถอยสมการเดียวได้ดังนี้

$$Y_{ij} = G_{00} + G_{10} X_{ij} + G_{01} Z_j + G_{11} X_{ij} Z_j + u_{oj} + u_{1j} X_{ij} + r_{ij}$$

จะเห็นว่าค่า Y_{ij} จะแยกเป็นสองส่วนคือส่วนที่เป็นสัมประสิทธิ์คงที่ (Fixed Coefficients) กับส่วนที่เป็นสัมประสิทธิ์สุ่ม (Random Coefficients) หรือความคลาดเคลื่อนสุ่ม สัมประสิทธิ์คงที่ประกอบด้วย $G_{00}, G_{10} X_{ij}, G_{01} Z_j, G_{11} X_{ij} Z_j$ สัมประสิทธิ์สุ่มประกอบด้วย $u_{oj}, u_{1j} X_{ij}, r_{ij}$ ส่วน $X_{ij} Z_j$ เป็นปฏิสัมพันธ์ของตัวแปรต่างระดับ X_{ij} กับ Z_j ในการส่งผลต่อ Y_{ij}

3.3.2 การวิเคราะห์ตัวแปร 3 ระดับ

การวิเคราะห์ข้อมูลตัวแปร 3 ระดับ เช่น ระดับที่ 1 วิเคราะห์ข้อมูลของตัวแปรระหว่างนักเรียนภายในหน่วยสถานศึกษา ระดับที่ 2 วิเคราะห์ข้อมูลของตัวแปรระหว่างสถานศึกษา

ภายในเขตพื้นที่การศึกษา และระดับที่ 3 วิเคราะห์ข้อมูลของตัวแปรระหว่างเขตพื้นที่การศึกษา เขียนสมการได้ดังนี้

ระดับที่ 1 : ระดับบุคคล(ระหว่างบุคคล i ภายในหน่วย j ของเขตพื้นที่ k)

$$A_{ijk} = P_{0jk} + P_{1jk}X_{ijk} + r_{ijk}$$

นำค่า P_{ojk} และ P_{1jk} มาใช้เป็นตัวแปรตามในการวิเคราะห์ระดับที่ 2

ระดับที่ 2 : ระดับหน่วย(ระหว่างหน่วย j ภายในเขตพื้นที่ k)

$$P_{ojk} = \beta_{00k} + \beta_{01k}Y_{jk} + u_{0jk}$$

$$P_{1jk} = \beta_{10k} + \beta_{11k}Y_{jk} + u_{1jk}$$

นำค่า β_{00k}, β_{01k}, β_{10k} และ β_{11k} มาใช้เป็นตัวแปรตามในการวิเคราะห์ระดับที่ 3

ระดับที่ 3 : ระดับเขต(ระหว่างเขตพื้นที่ k)

$$\beta_{00k} = G_{000} + G_{001}Z_k + v_{00k}$$

$$\beta_{01k} = G_{010} + G_{011}Z_k + v_{01k}$$

$$\beta_{10k} = G_{100} + G_{101}Z_k + v_{10k}$$

$$\beta_{11k} = G_{110} + G_{111}Z_k + v_{11k}$$

3.3 การประยุกต์ใช้โมเดลเชิงเส้นทั่วไปแบบลดหลั่น (HGLM)สำหรับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

จากหลักการวิเคราะห์พหุระดับของข้อมูลที่มีลักษณะเป็นระดับชั้น โมเดลที่ใช้ในการวิเคราะห์จะสามารถนำค่าตัวแปรที่ได้จากการวิเคราะห์ในระดับชั้นต่ำกว่าไปวิเคราะห์ต่อในระดับชั้นที่สูงขึ้น ทำให้สามารถแปลความหมายของผลที่ส่งผลต่อตัวแปรในระดับชั้นที่สูงขึ้นไป จึงสามารถนำหลักการวิเคราะห์พหุระดับ (Multi-level Analysis) ไปใช้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เมื่อจัดข้อมูลเป็นพหุระดับหรือมีลักษณะลดหลั่น (Hierarchical) ร่วมกับโมเดลการตอบสนองข้อสอบตามแนวคิดของโมเดลของราสช์ (Rasch Model)

Kamata (1998) ได้เสนอการใช้พารามิเตอร์ข้อสอบหนึ่งพารามิเตอร์ตามแนวคิดของโมเดลราชกับการใช้โมเดลเชิงเส้นโลจิสติกทั่วไป เพื่อการประยุกต์ใช้โมเดลเชิงเส้นทั่วไปแบบลดหลั่นสำหรับข้อมูลแบบสองค่าหรือสำหรับข้อสอบที่ตรวจให้คะแนนสองค่า(Dichotomous) จากนั้นได้มีการศึกษาขยายผลจากโมเดลเชิงเส้นทั่วไปแบบลดหลั่น (HGLM) ไปใช้กับข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า โดยการใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่นสำหรับข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า (Polytomous Generalized Linear Logistic Model: PHGLM) การแปลความหมายของตัวทำนายของ PHGLM จะใช้กรอบของ DIF โดยใช้การตรวจสอบด้วยวิธี PHGLM จะต้องประมาณค่าพารามิเตอร์ข้อสอบและพารามิเตอร์ของผู้สอบภายใต้สิ่งที่ต้องประเมินเป็นอันดับแรก

โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น (HGLM) สำหรบข้อสอบที่ตรวจให้คะแนนสองค่าและข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า มีหลักการและวิธีการดังนี้

3.3.1 โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่นสำหรับข้อสอบที่ตรวจให้คะแนนสองค่า

การนำหลักการพหุระดับมาใช้ในการวิเคราะห์ข้อสอบ ได้มีเริ่มมีการนำใช้โดย Kamata (2001) ซึ่งได้เสนอวิธีการวิเคราะห์ข้อสอบโดยการใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่นร่วมกับหลักการของ IRT ในโมเดลของราสซ์ โดยข้อสอบเป็นหน่วยการวิเคราะห์ในระดับที่ 1 ผู้สอบเป็นหน่วยการวิเคราะห์ในระดับที่ 2 และสถานศึกษาจะเป็นหน่วยการวิเคราะห์ในระดับที่ 3

1) ความสอดคล้องกันของโมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่นสำหรับข้อสอบที่ตรวจให้คะแนนสองค่ากับโมเดลราสซ์

โมเดลของราสซ์มีแนวคิดเช่นเดียวกันกับแนวคิดของทฤษฎีการตอบสนองข้อสอบแบบหนึ่งพารามิเตอร์ ซึ่งประกอบด้วยพารามิเตอร์ของผู้สอบหรือความสามารถของผู้สอบและพารามิเตอร์ของข้อสอบ นั่นคือค่าความยากของข้อสอบ การใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่นในการวิเคราะห์ สามารถแสดงให้เห็นถึงความสอดคล้องกันของพารามิเตอร์เมื่อมีการวิเคราะห์ 3 ระดับดังนี้

<u>การวิเคราะห์ระดับที่ 1</u> ระดับข้อสอบ

เป็นการวิเคราะห์อิทธิพลของข้อสอบที่สอดแทรกอยู่ในผู้สอบแต่ละคน และผู้สอบสอดแทรกอยู่ในแต่ละสถานศึกษา ซึ่งอิทธิพลของข้อสอบจะมีค่าคงที่สำหรับผู้สอบทุกคนที่อยู่ในสถานศึกษาเดียวกัน แต่จะมีความผันแปรแบบสุ่มระหว่างข้อสอบแต่ละข้อ แสดงสมการวิเคราะห์ดังนี้

$$\eta_{ijm} = \log\left(\frac{p_{ijm}}{1-p_{ijm}}\right) = \beta_{0jm} + \beta_{1jm}x_{1ijm} + \beta_{2jm}x_{2ijm} + \ldots + \beta_{kjm}x_{kijm}$$

$$= \beta_{0j} + \sum_{q=1}^{k}\beta_{qim}x_{qijm} \qquad = \beta_{oj} + \beta_{qjm}$$

เมื่อ p_{ijm} เป็นความน่าจะเป็นในการตอบข้อสอบที่ i ของผู้สอบ j ในสถานศึกษา m ได้ถูกต้อง

η_{ijm} เป็นค่าฟังก์ชันโลจิสกับลอคออดในการตอบข้อสอบที่ i ของผู้สอบ j ในสถานศึกษา m

x_{qijm} เป็นตัวแปรดัมมี่ตัวที่ q ในการทำข้อสอบที่ i ของผู้สอบที่ jในสถานศึกษา m โดยกำหนดให้ค่า x_{qijm} มีค่าเป็น 0 เมื่อ $q \neq i$ และ x_{qijm} มีค่าเป็น 1 เมื่อ $q = i$

$\beta_{0\,jm}$ เป็นสัมประสิทธิ์ที่เป็นค่าจุดตัดแกนและเป็นค่าที่ใช้แปลความหมายของอิทธิพลของข้อสอบที่ตัดออก หรือข้อสอบข้ออ้างอิง (reference item)

β_{qjm} เป็นอิทธิพลของตัวแปรดัมมี่ที่ q เช่นข้อสอบข้อที่ i เมื่อ i = 1, …, k –1) เมื่อเปรียบเทียบกับข้อสอบข้ออ้างอิง

การวิเคราะห์ระดับที่ 2 ระดับผู้สอบ

เป็นการวิเคราะห์ค่าความสามารถในการตอบข้อสอบระหว่างผู้สอบ และค่าอิทธิพลของข้อสอบรายข้อของผู้สอบที่อยู่ในสถานศึกษาเดียวกัน ซึ่งค่าอิทธิพลของข้อสอบ จะเป็นค่าคงที่ระหว่างสถานศึกษา แต่มีความผันแปรแบบสุ่มในข้อสอบแต่ละข้อ แสดงสมการได้ดังนี้

$$\beta_{0jm} = \gamma_{00m} + r_{0jm}$$
$$\beta_{1jm} = \gamma_{10m}$$
$$\beta_{2jm} = \gamma_{20m}$$
$$\vdots$$
$$\beta_{(k-1)} = \gamma_{(k-1)0m}$$

เมื่อ γ_{00m} เป็นค่าจุดตัดแกนของ β_{0jm} หรือค่าเฉลี่ยของอิทธิพลข้อสอบข้ออ้างอิงต่อโอกาสในการตอบข้อสอบถูกในสถานศึกษา m

r_{ojm} เป็นค่าเศษเหลือของ β_{0jm} หรือค่าที่เป็นส่วนเบี่ยงเบนของโอกาสในการตอบข้อสอบของผู้สอบ j ในสถานศึกษา m ซึ่งการแจกแจงแบบโค้งปกติ มีค่าเฉลี่ยเป็นศูนย์และมีความแปรปรวนเท่ากับ τ

การวิเคราะห์ระดับที่ 3 ระดับสถานศึกษา

เป็นการวิเคราะห์ค่าความสามารถในการตอบข้อสอบของผู้สอบในระดับสถานศึกษา และค่าเฉลี่ยรวมอิทธิพลของข้อสอบรายข้อของผู้สอบต่อโอกาสในการตอบข้อสอบถูกระหว่างสถานศึกษา แสดงสมการได้ดังนี้

$$\gamma_{00m} = \pi_{000} + u_{00m}$$
$$\gamma_{10m} = \pi_{100}$$
$$\gamma_{20m} = \pi_{200}$$
$$\vdots$$
$$\gamma_{(k-1)0m} = \pi_{(k-1)00}$$

เมื่อ π_{000} เป็นค่าจุดตัดแกนของ γ_{00m} หรือเป็นค่าเฉลี่ยรวมของอิทธิพลของข้อสอบข้ออ้างอิงต่อโอกาสในการตอบถูกของทุกสถานศึกษา

u_{00m} เป็นค่าเศษเหลือของ γ_{00m} หรือค่าที่เป็นส่วนเบี่ยงเบนของโอกาสในการตอบข้อสอบของของสถานศึกษาที่ m หรือเป็นความสามารถเฉลี่ยของนักเรียนในสถานศึกษา m ซึ่งการแจกแจงแบบโค้งปกติ มีค่าเฉลี่ยเป็นศูนย์และมีความแปรปรวนเป็น τ_{π}

π_{q00} เป็นค่าอิทธิพลของข้อสอบข้อที่ 1 ถึงข้อที่ k – 1 ต่อโอกาสในการตอบข้อสอบถูกต้องเมื่อเทียบกับข้ออ้างอิง

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อรวมสมการในระดับที่ 1 ถึงระดับที่ 3 จะเขียนสมการได้ดังนี้

$$\eta_{ijm} = \beta_{oj} + \beta_{qjm}$$
$$= \pi_{000} + r_{0jm} + u_{00m} + \pi_{q00}$$

จะได้ความน่าจะเป็นที่จะตอบข้อสอบที่ i ถูกของผู้สอบที่ j ในสถานศึกษาที่ m ดังสมการ

$$P_{ijm} = \frac{1}{1+\exp\{-\eta_{ijm}\}}$$

$$P_{ijm} = \frac{1}{1+\exp\{-(u_{00m}+r_{ojm})-(-\pi_{q00}-\pi_{000})\}}$$

เมื่อพิจารณาค่าพารามิเตอร์เทียบกับโมเดลราช ดังนี้

โมเดล HGLM-3L ; $P_{ij} = \frac{1}{1+\exp\{-(u_{00m}+r_{0jm})-(-\pi_{q00}-\pi_{000})\}}$

โมเดลราช ; $P_{ij} = \frac{1}{1+\exp\{-(\theta_j-\delta_i)\}}$

เมื่อ u_{00m} เป็นค่าความสามารถเฉลี่ยของนักเรียนในสถานศึกษาที่ m

r_{0jm} เป็นค่าความสามารถเฉพาะของนักเรียนที่ j ในสถานศึกษาที่ m

π_{000} เป็นค่าความยากของข้อสอบข้ออ้างอิง

π_{q00} เป็นค่าอิทธิพลของข้อสอบข้อที่ q (เมื่อ qเป็นข้อสอบที่ iถึง $k-1$) เมื่อเทียบกับข้อสอบข้ออ้างอิง

θ_j เป็นค่าความสามารถของผู้สอบที่ j

δ_i เป็นค่าความยากของข้อสอบข้อที่ i

ดังนั้นเมื่อเปรียบเทียบพารามิเตอร์แล้วจะพบว่าพารามิเตอร์ความสามารถของผู้สอบและความยากง่ายของข้อสอบระหว่างโมเดลราสซ์มีความสอดคล้องกับโมเดล HGLM

สรุปได้ดังตารางที่ 3

ตารางที่ 3 เปรียบเทียบค่าความสามารถของผู้สอบและความยากง่ายของข้อสอบจากการประมาณค่าด้วยโมเดลราสซ์และโมเดล HGLM-3L

โมเดล	ความสามารถของผู้สอบ	ความยากง่ายของข้อสอบ
โมเดลราสซ์ (Rasch)	θ_j	δ_i
โมเดล HGLM-3L	$u_{00m}+r_{0jm}$	$(-\pi_{q00}-\pi_{000})$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

2) การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ตรวจให้คะแนนสองค่าด้วยโมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น

การใช้หลักการวิเคราะห์พหุระดับในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ เป็นวิธีของการวิเคราะห์ถดถอยโลจิสติกมาใช้ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning : DIF) Kamata (1998) ได้เสนอวิธีการวิเคราะห์สองระดับเพื่อศึกษาเกี่ยวกับ DIF ที่เกี่ยวกับหลักการของโมเดล IRT ของราสช์ โดยมีข้อสอบเป็นระดับที่ 1 และผู้สอบเป็นระดับที่ 2 ต่อมา Binici (2007) และ Kamata, Chaimongkol, Genc & Bilir (2005) ได้ศึกษาเพิ่มเติมจากโมเดลการวิเคราะห์ 2 ระดับมาเป็นโมเดลการวิเคราะห์ 3 ระดับสำหรับการวิเคราะห์ DIF โดยโมเดลสามระดับจะมีประโยชน์สำหรับโมเดลอิทธิพลสุ่ม (random-effect model) เพื่อเป็นตัวบ่งชี้ถึงการเกิด DIF จากการจัดกลุ่มตัวแปรในโมเดลระดับที่สาม และเพื่อเป็นโมเดลของอิทธิพลการอธิบาย (explanatory effect model) ในการระบุตัวแปรคุณลักษณะโมเดลระดับที่สามที่เป็นแหล่งของการเกิดการทำหน้าที่ต่างกันของข้อสอบ

ในการวิเคราะห์ตามแนวคิดของโมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น (HGLM) ในการวิเคราะห์ระดับที่ 1 เป็นระดับข้อสอบ จะกำหนดให้ค่าพารามิเตอร์ความยากง่ายของข้อสอบเป็นค่าคงที่กับผู้สอบแต่ละกลุ่ม และการวิเคราะห์ที่ 2 เป็นระดับผู้สอบ จะใช้ตัวแปรบ่งชี้ของกลุ่มเพื่อทำนายโอกาสในการตอบข้อสอบได้ถูก ถ้าพบว่ามีนัยสำคัญทางสถิติจะแสดงได้ว่าตัวแปรบ่งชี้ของกลุ่มมีอิทธิพลต่อโอกาสในการตอบข้อสอบได้ถูก นั่นคือข้อสอบนั้นเป็นข้อสอบที่ทำหน้าที่ต่างกัน นอกจากนี้ยังสามารถวิเคราะห์ในระดับที่ 3 ซึ่งเป็นระดับสถานศึกษาที่ผู้สอบสอดแทรกอยู่ในกลุ่มระดับใหญ่ขึ้นเพื่อใช้คุณลักษณะของกลุ่มเป็นตัวแปรในการทำนายคุณลักษณะของกลุ่มผู้สอบที่ส่งผลต่อโอกาสในการตอบข้อสอบได้ถูก เพื่อระบุแหล่งของการทำหน้าที่ต่างกันของข้อสอบ

เขียนสมการแสดงการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ ตามแนวคิดของ HGLM-3L ได้ดังนี้

<u>การวิเคราะห์ระดับที่ 1</u> ระดับข้อสอบ

เป็นการวิเคราะห์อิทธิพลของข้อสอบที่สอดแทรกอยู่ในผู้สอบแต่ละคน และในแต่ละสถานศึกษา แสดงสมการวิเคราะห์ดังนี้

$$\eta_{ijm} = \beta_{0jm} + \beta_{1jm} x_{1ijm} + \beta_{2jm} x_{2ijm} + ... + \beta_{kjm} x_{kijm}$$

$$= \beta_{0j} + \sum_{q=1}^{k} \beta_{qim} x_{qijm} \qquad = \beta_{oj} + \beta_{qjm}$$

เมื่อ η_{ijm} เป็นค่าฟังก์ชันโลจิสกับลอคออดในการตอบข้อสอบที่ i ของผู้สอบ j ในสถานศึกษา m

$\beta_{0\,jm}$ เป็นสัมประสิทธิ์ที่เป็นค่าจุดตัดแกนและเป็นค่าที่ใช้แปลความหมายของอิทธิพลของข้อสอบที่ตัดออก หรือข้อสอบข้ออ้างอิง (reference item) หรือความน่าจะเป็นโดยเฉลี่ยที่ผู้สอบ j จะตอบข้อสอบข้ออ้างอิงถูกในสถานศึกษา m

β_{qjm} เป็นอิทธิพลของตัวแปรดัมมี่ที่ q เช่นข้อสอบข้อที่ i เมื่อ i = 1, …, k –1) เมื่อเปรียบเทียบกับข้อสอบข้ออ้างอิง

x_{qijm} เป็นตัวแปรอิทธิพลของข้อสอบรายข้อซึ่งเป็นตัวแปรดัมมี่ตัวที่ q ในการทำข้อสอบที่ i ของผู้สอบที่ j ในสถานศึกษา m โดยค่า x_{qijm} จะเป็น 0 เมื่อ $q \neq i$ และ x_{qijm} มีค่าเป็น 1 เมื่อ $q = i$

เพื่อให้โมเดลในสมการพหุระดับเป็นเมตริกซ์เอกลักษณ์ จึงต้องกำหนดให้ x_{qijm} ตัวใดตัวหนึ่งเป็นศูนย์ เพื่อทำให้เป็นเมตริกซ์เต็มอันดับ ซึ่งโดยทั่วไปจะตัดข้อสอบข้อสุดท้าย

การวิเคราะห์ระดับที่ 2 ระดับผู้สอบ

เป็นการวิเคราะห์ที่นำตัวแปรคุณลักษณะของกลุ่มผู้สอบเข้าไปในสมการ เพื่อทำนายโอกาสในการตอบข้อสอบถูกต้อง โดยค่าอิทธิพลของข้อสอบรายข้อของผู้สอบจะเป็นตัวบ่งชี้ถึงการทำหน้าที่ต่างกันของข้อสอบ แสดงสมการได้ดังนี้

$$\beta_{0jm} = \gamma_{00m} + \gamma_{01m}G_{jm} + r_{0jm}$$
$$\beta_{1jm} = \gamma_{10m} + \gamma_{11m}G_{jm}$$
$$\beta_{2jm} = \gamma_{20m} + \gamma_{21m}G_{jm}$$
$$\vdots$$
$$\beta_{qjm} = \gamma_{q0m} + \gamma_{q1m}G_{jm}$$

เมื่อ γ_{00m} เป็นค่าจุดตัดแกนของ β_{0jm} หรือค่าเฉลี่ยของโอกาสในการตอบข้อสอบข้ออ้างอิงของผู้สอบในสถานศึกษา m

γ_{q0m} เป็นค่าจุดตัดแกนของ β_{qjm} หรือค่าเฉลี่ยของโอกาสในการตอบข้อสอบข้อที่ q ได้ถูกต้องของผู้สอบในสถานศึกษาที่ m

γ_{q1m} เป็นค่าอิทธิพลของตัวแปรของกลุ่มผู้สอบ G_{jm} ที่มีผลต่อโอกาสในการตอบข้อสอบถูกในข้อที่ q ในสถานศึกษา m

r_{0jm} เป็นค่าเศษเหลือของ β_{0jm} หรือโอกาสของการตอบข้อสอบถูกของนักเรียน j ในสถานศึกษา m

การวิเคราะห์ระดับที่ 3 ระดับสถานศึกษา

เป็นขั้นตอนการวิเคราะห์ที่นำตัวแปรทำนายระดับสถานศึกษาเข้าไปในสมการเพื่อทำนายโอกาสในการตอบถูกของผู้สอบที่สอดแทรกในสถานศึกษาที่มีคุณลักษณะต่างกัน แสดงสมการได้ดังนี้

$$\gamma_{00m} = \pi_{000} + \pi_{001}SCH_m + u_{00m}$$
$$\gamma_{01m} = \pi_{010} + \pi_{011}SCH_m + u_{01m}$$
$$\gamma_{10m} = \pi_{100}$$
$$\gamma_{20m} = \pi_{200}$$
$$\vdots$$
$$\gamma_{q0m} = \pi_{q00}$$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อ π_{000} เป็นค่าจุดตัดแกนของ γ_{00m} หรือเป็นค่าเฉลี่ยรวมของอิทธิพลของข้อสอบข้ออ้างอิงต่อโอกาสในการตอบถูกของทุกสถานศึกษา

π_{001} เป็นค่าอิทธิพลของตัวแปรในระดับสถานศึกษาที่มีผลต่อโอกาสในการตอบข้อสอบข้ออ้างอิงได้ถูกในสถานศึกษาที่ m

π_{010} เป็นค่าจุดตัดแกนของ γ_{01m} หรือเป็นค่าเฉลี่ยรวมของอิทธิพลของตัวแปรในระดับผู้สอบที่ส่งผลต่อโอกาสในการตอบถูกของทุกสถานศึกษา

π_{011} เป็นค่าจุดตัดแกนของ γ_{00m} หรือเป็นค่าเฉลี่ยรวมของอิทธิพลของตัวแปรระดับสถานศึกษาต่อโอกาสในการตอบถูกของทุกสถานศึกษา

π_{q00} เป็นค่าเฉลี่ยของโอกาสในการตอบข้อสอบข้อ q ของผู้สอบได้ถูกต้อง

3.3.2 โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่นสำหรับข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า

การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า ได้มีผู้สนใจศึกษาถึงในการนำโมเดลเชิงเส้นตรงแบบลดหลั่นมาใช้ในการตรวจสอบ เช่น Cheong (2006) และ Williams & Beretvas (2006) ที่ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีการถดถอยโลจิสติกของข้อสอบด้วยโมเดลสามระดับ และจากการศึกษาของ Chaimongkol (2005) เพื่อตรวจสอบ DIF โดยใช้โมเดลสามระดับที่มีการสอดแทรกข้อมูลในระดับต่างๆ กล่าวคือข้อสอบสอดแทรกในนักเรียนที่ทำข้อสอบและนักเรียนสอดแทรกในสถานศึกษา โดยอิทธิพลสุ่มระดับที่สามที่แปรค่าตามหน่วยในระดับที่สาม การเกิด DIF ในการวิเคราะห์ระดับที่สาม จะแสดงถึงการเปลี่ยนแปลงในอิทธิพลของ DIF สำหรับหน่วยในระดับที่สาม

1) ความสอดคล้องกันของโมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่นสำหรับข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่ากับโมเดลราสช์

โมเดล HGLM สำหรับข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า จะมีชุดของสมการระดับที่ 1 อยู่สองสมการ สำหรับข้อสอบแต่ละข้อที่รายการคำตอบหลายคำตอบ ซึ่งอาจจะกำหนดให้คะแนนหลายค่าเช่น 1, 2, 3

<u>โมเดลระดับที่ 1</u> ระดับข้อสอบ

สมการในระดับที่ 1 มีสองสมการโดยสมการแรกเป็นสมการสำหรับรายการคำตอบแรกหรือ $m = 1$ และสมการที่สองเป็นสมการสำหรับโอกาสในการตอบของรายการคำตอบที่เหนือกว่าหรือ $1 < m < M - 1$ เขียนสมการได้เป็น

$$\eta_{1ij} = \beta_{0j} + \beta_{1j} X_{1ij} + ... + \beta_{(k-1)j} X_{(k-1)ij}$$
$$\eta_{2ij} = \beta_{0j} + \beta_{1j} X_{1ij} + ... + \beta_{(k-1)j} X_{(k-1)ij} + \delta_j$$

เมื่อ η_{mij} เป็นฟังก์ชันลอดของออดของการตอบสนองในรายการคำตอบที่ m หรือต่ำกว่า

X_{qij} เป็นตัวแปรบ่งชี้ที่ q ของข้อสอบที่ i ผู้สอบที่ j

β_{qj} เป็นค่าจุดตัดแกนหรือสัมประสิทธิ์ถดถอยของตัวแปร X_{qij} สำหรับผู้สอบ j

δ_j เป็นค่าความแตกต่างของ threshold ระหว่างรายการคำตอบ

<u>โมเดลระดับที่ 2</u> ระดับผู้สอบ

เขียนสมการในโมเดลระดับที่ 2 ได้เป็น

$$\beta_{0j} = \gamma_{00} + u_{0j}$$
$$\beta_{1j} = \gamma_{10}$$
$$\vdots$$
$$\beta_{(k-1)j} = \gamma_{(k-1)0}$$
$$\delta_j = \delta$$

เมื่อ β_{1j} ถึง $\beta_{(k-1)j}$ เป็นอิทธิพลของข้อสอบซึ่งเป็นค่าคงที่

δ_j เป็นความแตกต่างของ threshold ในโมเดลเป็นค่าคงที่ระหว่างข้อสอบ

เมื่อพิจารณาสมการในโมเดลระดับที่ 1 และโมเดลระดับที่ 2 จะได้โอกาสในการตอบรายการคำตอบเป็น 1 สัมพัทธ์กับการตอบรายการคำตอบเป็น 2 หรือ 3 สำหรับข้อสอบ i เขียนได้เป็น

$$\frac{\Pr_{ij}(X_i = 1)}{\Pr_{ij}(X_i = 2,3)} = \exp(\gamma_{00} - \gamma_{i0} + u_{0j})$$

โอกาสของการตอบสนองรายการคำตอบเป็น 1 หรือ 2 สัมพัทธ์กับโอกาสการตอบสนองรายการ คำตอบเป็น 3 สำหรับข้อสอบ i เขียนได้เป็น

$$\frac{\Pr_{ij}(X_i = 1,2)}{\Pr_{ij}(X_i = 3)} = \exp(\gamma_{00} - \gamma_{i0} + \delta + u_{0j})$$

เมื่อเปรียบเทียบค่าพารามิเตอร์ของ HGLM กับโมเดลราสช์ พบว่า

θ_j เป็นความสามารถของผู้สอบที่ j ที่สอดคล้องกับ $-u_{0j}$

$(\gamma_{00} - \gamma_{i0} + \delta)$ เป็นความยากของข้อสอบข้อ i ในการตอบในรายการคำตอบ 3

$(\gamma_{00} + \delta)$ เป็นความยากของข้อสอบของข้ออ้างอิง

$(\gamma_{00} - \gamma_{i0} - \delta_{m-2})$ เป็นความยากของข้อสอบข้อ i สำหรับรายการคำตอบ m

เมื่อ $2 < m \leq k$

2) การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่าด้วยโมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น

โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่นในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบที่ตรวจให้คะแนนสองค่ามีหลักการคล้ายกันกับข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า นั่นคือ ระดับที่ 1 เป็นการวิเคราะห์ในหน่วยของข้อสอบ ระดับที่ 2 เป็นระดับผู้สอบ ซึ่งสามารถเพิ่มตัวแปรทำนายเพื่อใช้อธิบายลักษณะของผู้สอบที่ส่งผลต่อโอกาสในการตอบข้อสอบได้ถูก และระดับที่ 3 เป็นระดับสถานศึกษา โดยการเพิ่มตัวแปรทำนายเพื่อใช้ระบุและอธิบายแหล่งที่ทำให้เกิดการทำหน้าที่ต่างกันของข้อสอบ

เขียนสมการในการวิเคราะห์ในละระดับ ได้ดังนี้

<u>โมเดลระดับที่ 1</u> ระดับข้อสอบ

การวิเคราะห์ DIF สำหรับการตอบสนองที่เป็นการจัดลำดับ m ($m=1,...,M$) ในข้อสอบแบบตรวจให้คะแนนมากกว่าสองค่า r ข้อ ที่กำหนดให้ค่าอำนาจจำแนก $\alpha_i=1$ สำหรับข้อสอบทุกข้อในกรณีที่เป็นโมเดลหนึ่งพารามิเตอร์ ของข้อสอบข้อที่ i ผู้สอบคนที่ j ในกลุ่มที่ k เขียนสมการได้ดังนี้

$$\begin{aligned}\eta_{mijk} &= \log\left(\frac{\varphi^{*}_{mijk}}{1-\varphi^{*}_{mijk}}\right)\\ &= \begin{cases}\beta_{ojk}+\beta_{1jk}X_{1ijk}+...+\beta_{(r-1)jk}X_{(r-1)ijk} & ,m=1\\ \beta_{ojk}+\beta_{1jk}X_{1ijk}+...+\beta_{(r-1)jk}X_{(r-1)ijk}+\delta_{2\mathrm{jk}}+...+\delta_{2jk} & ,1<m\le M-1\end{cases}\\ &= \beta_{0jk}+\sum_{q=1}^{r-1}\beta_{qjk}X_{qijk}+\sum_{p=2}^{m}\delta_{pjk}\end{aligned}$$

จากสมการข้างต้นแสดงในรูปของการตอบสนอง m ของข้อสอบที่ q ได้ดังนี้

$$\eta_{mqjk}=\beta_{0jk}+\beta_{qjk}+\sum_{p=2}^{m}\delta_{pjk}$$

เมื่อ η_{mijk} เป็นค่าลอคของออดของรายการคำตอบ m ของผู้สอบคนที่ j ในข้อสอบข้อที่ i ในกลุ่มที่ k

φ^{*}_{mijk} เป็นความน่าจะเป็นสะสมของรายการคำตอบ m ของผู้สอบคนที่ j ในข้อสอบข้อที่ i ในกลุ่มที่ k

β_{ojk} เป็นค่าเฉลี่ยของโอกาสในการตอบข้อสอบได้ถูกของผู้สอบคนที่ j ในกลุ่มที่ k

β_{qjk} เป็นสัมประสิทธิ์ของ X_{qijk} ซึ่งเป็นผลของข้อสอบที่ q เมื่อเทียบกับข้ออ้างอิง หรือเป็นโอกาสในการตอบข้อสอบได้ถูกของผู้สอบคนที่ j ในข้อสอบที่ q กลุ่มที่ k

X_{qijk} เป็นตัวแปรดัมมี่ของข้อสอบที่ q ของผู้สอบที่ j กำหนดให้ค่าของ X_{qijk} เท่ากับ 1 เมื่อ $q=1$ และค่าของ X_{qijk} เท่ากับศูนย์ เมื่อ $q\neq1$

δ_{pjk} เป็นค่าพารามิเตอร์ของข้อสอบ ซึ่งเป็นอิทธิพลคงที่สำหรับข้อสอบ q ในแต่ละความแตกต่างของ threshold (δ_{2iq} ถึง δ_{miq} เมื่อ $p=2,3,...,m$)

<u>โมเดลระดับที่ 2</u> ระดับผู้สอบ แสดงได้ดังนี้

โมเดลในระดับนี้เป็นเพิ่มตัวแปรทำนายที่เป็นคุณลักษณะของกลุ่มผู้สอบเพื่อใช้ในการอธิบายโอกาสในการตอบข้อสอบได้ถูกต้อง เขียนสมการได้ดังนี้

$$\beta_{0jk} = \gamma_{00k} + \gamma_{01k}G_{jk} + u_{0jk}$$
$$\beta_{1jk} = \gamma_{10k} + \gamma_{11k}G_{jk}$$
$$\vdots$$
$$\beta_{(r-1)jk} = \gamma_{(r-1)0k} + \gamma_{(r-1)1k}G_{jk}$$
$$\delta_{2jk} = \delta_{20k} + \delta_{21qk}G_{jk}$$
$$\vdots$$
$$\delta_{mjk} = \delta_{m0k} + \delta_{m1qk}G_{jk}$$

เมื่อ G_{jk} เป็นตัวแปรบ่งชี้ของกลุ่มผู้สอบ j ในการจัดกลุ่ม k โดยกำหนดให้ G_{jk} = 0 ในกลุ่มอ้างอิง (Reference Group) และ G_{jk} =1 ในกลุ่มเปรียบเทียบ (Focal Group)

u_{0jk} เป็นความสามารถของผู้สอบ j ในกลุ่ม k ซึ่งจะกำหนดให้ $u_{0jk} \sim N(\mu_G, \sigma^2_{\gamma_G})$ มีค่าเฉลี่ยและความแปรปรวนเฉพาะกลุ่ม

β_{1jk} ถึง $\beta_{(r-1)jk}$ เป็นอิทธิพลของข้อสอบ ซึ่งในโมเดลนี้จัดเป็นอิทธิพลคงที่

γ_{10k} ถึง $\gamma_{(r-1)0k}$ เป็นอิทธิพลของกลุ่มผู้สอบที่มีผลต่อ β_{1jk} ถึง $\beta_{(r-1)jk}$ ซึ่งเป็นการเปรียบเทียบระหว่างกลุ่มอ้างอิงกับกลุ่มเปรียบเทียบ แสดงถึงปริมาณของ DIF สำหรับข้อสอบ i ในการตอบรายการคำตอบที่ 1 กับรายการคำตอบที่ 2 โดยถ้าค่า DIF หรือค่าของ γ_{10k} ถึง $\gamma_{(r-1)0k}$ มีค่าเป็นบวก แสดงว่ามีผลต่อกลุ่มเปรียบเทียบ

δ_{21q} ถึง δ_{m1q} เป็นอิทธิพลคงที่ของข้อสอบข้อที่ q ในแต่ละความแตกต่าง Threshold

δ_{p1qk} เป็นค่าที่แสดงถึงการเพิ่มขึ้นหรือลดลงในอิทธิพลของ DIF ต่อการตอบในรายการคำตอบชั้นที่สูงขึ้นไปของกลุ่มเปรียบเทียบ

ถ้า δ_{p1qk} มีค่าเป็นบวก แสดงว่าค่า DIF เป็นบวก นั่นคือข้อสอบโอนเอียงต่อกลุ่มเปรียบเทียบในการตอบในรายการคำตอบที่ p (p = 2, ..., m) ในข้อสอบ q

ถ้า δ_{p1qk} มีค่าเป็นลบ แสดงว่า ค่า DIF เป็นลบ นั่นคือความลำเอียงของข้อสอบไม่เป็นผลหรือเกิดผลตรงข้ามสำหรับการตอบในรายการที่ p ของข้อสอบ q

r_{00k} เป็นค่าเฉลี่ยของการตอบข้อสอบถูกของกลุ่ม k ของข้ออ้างอิง

r_{10k} ถึง $r_{(r-1)0k}$ เป็นค่าเฉลี่ยของการตอบข้อสอบถูกของข้อที่ 1 ถึง $r-1$ ของกลุ่ม k

<u>โมเดลระดับที่ 3</u> แสดงได้ดังนี้

หลังจากการวิเคราะห์ในระดับที่ 2 แล้วพบว่าข้อสอบเกิด DIF การวิเคราะห์ในระดับที่ 3 จะเป็นการอธิบายความผันแปร ถึงคุณลักษณะอื่นของหน่วยในระดับที่ 3 ที่ช่วยในการอธิบายธรรมชาติของการ DIF

$$\gamma_{00k} = \pi_{000} + \gamma_{00k}$$
$$\gamma_{10k} = \pi_{100} + \pi_{101}W_k$$
$$\vdots$$
$$\gamma_{(r-1)0k} = \pi_{(r-1)00} + \pi_{(r-1)01}W_k$$
$$\gamma_{01k} = \pi_{010}$$
$$\gamma_{11k} = \pi_{110} + \pi_{111}W_k + r_{11k}$$
$$\vdots$$
$$\gamma_{(r-1)1k} = \pi_{(r-1)10} + \pi_{(r-1)11}W_k + r_{(r-1)k}$$
$$\delta_{20k} = \delta_{200}$$
$$\vdots$$
$$\delta_{m0k} = \delta_{m00}$$
$$\delta_{21qk} = \delta_{21q0} + \delta_{21q1}W_k + +e_{21qk}$$
$$\vdots$$
$$\delta_{m1qk} = \delta_{m1q0} + \delta_{m1q1}W_k + e_{m1qk}$$

เมื่อ W_k เป็นคุณลักษณะของตัวแปรในระดับที่ 3

π_{q11} เป็นอิทธิพลของของคุณลักษณะของกลุ่ม G_{jk} ซึ่งประมาณค่า DIF ระหว่างลำดับชั้นของการตอบสนองชั้นที่หนึ่งและชั้นที่สองจะอธิบายได้ด้วยคุณลักษณะในระดับที่ 3

δ_{p1q1} เป็นอิทธิพลของกลุ่ม G_{jk} สำหรับอิทธิพลของคุณลักษณะของ IDIF ของชั้นการตอบที่อยู่สูงขึ้นไป ค่าของ δ_{p1q1} สามารถเป็นได้ทั้งบวกและลบ ซึ่งจำกัดค่าสำหรับโมเดล DIF กำหนดค่าให้คงที่ ซึ่ง threshold ทั้งหมดจะมีค่าเป็นบวก

โมเดลการอธิบาย DIF สำหรับผลการตรวจให้คะแนนหลายค่าสามารถแสดงได้ในข้อสอบ q ได้ดังนี้

$$\eta_{mqjk} = \pi_{010}G_{jk} + \pi_{q10}G_{jk} + \pi_{q11}G_{jk}W_k + r_{q1k}G_{jk} + \pi_{q00} + \pi_{000} + \pi_{q01}W_k + u_{0jk} + r_{00k} + \sum_{p=2}^{m}\delta_{p00} + \sum_{p=2}^{m}\delta_{p1q0}G_{jk} + \sum_{p=2}^{m}\delta_{p1q1}G_{jk}W_k + \sum_{p=2}^{m}e_{p1qk}G_{jk}$$

สำหรับกลุ่มอ้างอิง ($G_j = 0$) ความยากของข้อสอบในโมเดลหนึ่งพารามิเตอร์ ซึ่งเป็นความยากที่เกี่ยวข้องกับการตอบสนองในการจัดประเภท ($m+1$) หรือสูงกว่าในข้อสอบ q จะมีค่าเท่ากับ

$$\pi_{q00} + \pi_{000} + \pi_{q01}W_k + \sum_{p=2}^{m}\delta_{p00}$$

สำหรับกลุ่มเป้าหมาย ($G_j = 1$) ความยากของข้อสอบในการตอบรายการคำตอบ ($m+1$) หรือสูงกว่าในข้อสอบ q จะมีค่าเท่ากับ

$$\pi_{q10} + \pi_{010} + \pi_{q00} + \pi_{000} + \pi_{q01}W_k + \pi_{q11}W_k +$$
$$\sum_{p=2}^{m}\delta_{p00} + \sum_{p=2}^{m}\delta_{p1q0} + \sum_{p=2}^{m}\delta_{p1q1}W_k + r_{q1k} + \sum_{p=2}^{m}e_{p1qk}$$

ถ้าพิจารณาความผันแปรโดยรวมของ DIF ระหว่างลำดับชั้นของการตอบสนองและไม่เป็นการเปลี่ยนแปลงเฉพาะของความแปรค่าได้ของ DIF อิทธิพลสุ่มจะเป็นการรวมสมการดังนี้

$$r^{*}_{p1qk} = r_{q1k} + \sum_{p=2}^{m}e_{p1qk}$$

ถ้าพิจารณาการเกิดขึ้นของ DIF แบบสุ่ม ดังนั้นการรวมอิทธิพลสุ่มจึงเป็นวิธีการที่เหมาะสม แต่ถ้าพิจารณาเฉพาะการเปลี่ยนแปลงถึงความผันแปรของ DIF ระหว่างการตอบสนองในประเภทจัดชั้น โมเดลการแยกอิทธิพลสุ่มจะใช้สมการดังนี้

$$\eta_{mqjk} = \pi_{010}G_{jk} + \pi_{q10}G_{jk} + \pi_{q11}G_{jk}W_k + r_{q1k}G_{jk} + \pi_{q00} + \pi_{000} + \pi_{q01}W_k + u_{0jk} + r_{00k} +$$
$$\sum_{p=2}^{m}\delta_{p00} + \sum_{p=2}^{m}\delta_{p1q0}G_{jk} + \sum_{p=2}^{m}\delta_{p1q1}G_{jk}W_k + \sum_{p=2}^{m}e_{p1qk}G_{jk}$$

โมเดลการทำนายจะใช้วิธีการรวมความผันแปรสำหรับข้อสอบ q เขียนสมการดังนี้

$$\eta_{mqjk} = \pi_{010}G_{jk} + \pi_{q10}G_{jk} + \pi_{q11}G_{jk}W_k + \pi_{q00} + \pi_{000} + \pi_{q01}W_k + u_{0jk} + r_{00k} +$$
$$\sum_{p=2}^{m}\delta_{p00} + \sum_{p=2}^{m}\delta_{p1q0}G_{jk} + \sum_{p=2}^{m}\delta_{p1q1}G_{jk}W_k + r^{*}_{p1qk}G_{jk}$$

อิทธิพลข้อสอบของ

$$\pi_{q00} + \pi_{000} + \pi_{q01}W_k + \sum_{p=2}^{m}\delta_{p00}$$

ซึ่งสอดคล้องกับความยากของข้อสอบในโมเดลหนึ่งพารามิเตอร์สำหรับกลุ่มอ้างอิง ($G_j = 0$) และหมายความถึงความยากของการตอบสนองในการจัดประเภท $(m+1)$ หรือชั้นสูงกว่าในข้อสอบ q สำหรับกลุ่มเป้าหมาย ($G_j = 1$) ความยากของการตอบสนองในการจัดประเภท $(m+1)$ หรือชั้นสูงกว่าในข้อสอบ q จะเป็น

$$\pi_{q10} + \pi_{010} + \pi_{q00} + \pi_{000} + \pi_{q01}W_k + \pi_{q11}W_k$$
$$+ \sum_{p=2}^{m}\delta_{00} + \sum_{p=2}^{m}\delta_{p1q0} + \sum_{p=2}^{m}\delta_{p1q1}W_k + r^{*}_{p1qk}$$

การวิเคราะห์ข้อสอบที่ใช้การวิเคราะห์พหุระดับมาประยุกต์จะช่วยให้ได้ข้อสารสนเทศมากขึ้น โดยโมเดลการวัดแบบพหุระดับเป็นโมเดลที่เหมาะสมกับโครงสร้างข้อมูลที่มีลักษณะสอดแทรกการตอบสนองข้อสอบและยังเป็นการจัดกระทำความยากของข้อสอบและความสามารถของผู้สอบเป็นพารามิเตอร์แบบสุ่ม

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3.3.3 งานวิจัยที่ศึกษาเกี่ยวกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยการประยุกต์ใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น

จากการศึกษางานวิจัยเกี่ยวกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ โดยใช้ HLM มาประยุกต์ใช้ในการตรวจสอบ การใช้โมเดล HGLM ที่ได้เริ่มมีการใช้ตั้งแต่ปี 2005 มีการพัฒนาไปสู่โมเดลสามระดับและการศึกษาเพิ่มเติมของข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า สรุปผลการศึกษาได้ดังนี้

Vaughn (2006) ได้ศึกษาการทำหน้าที่ต่างกันของข้อสอบ โดยใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น (HGLM) ของการทำหน้าที่ต่างกันของข้อสอบสำหรับข้อสอบพหุวิภาค โดยการประมาณค่าด้วยวิธีพหุระดับเบเซี่ยน การศึกษาจะพิจารณา DIF แบบสุ่มสำหรับข้อสอบแบบตรวจให้คะแนนมากกว่าสองค่า จากการถดถอยโลจิสติก 3 ระดับ ที่สอดแทรกตัวแปรเช่น คะแนนของข้อสอบของนักเรียนที่อยู่ในสถานศึกษา การใช้โมเดล HGLM ที่ปรับเข้ากับโมเดล IRT สำหรับข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่าอาจจะเกิด DIF ระหว่างระดับการตอบ นั่นคือการเกิด DIF ภายใน (inner-response DIF: IDIF) เนื่องจากผลของ DIF ที่มีการเปลี่ยนแปลงแบบสุ่ม การศึกษานี้ใช้โมเดลกับข้อมูลจริง และพิจารณาอิทธิพลสุ่ม พบว่าข้อสอบทุกข้อเกิด DIF ในระหว่างลำดับชั้นของการให้คะแนน พบข้อสอบสองข้อที่แสดงให้เห็นถึงการเกิด DIF ในชั้นที่มีคะแนนต่ำกว่าข้อสอบหนึ่งข้อที่เกิด DIF ในชั้นที่อยู่สูงกว่า ข้อสอบหนึ่งข้อมีผลต่อผู้สอบเพศหญิงและสองข้อมีผลกระทบต่อผู้สอบเพศชาย และข้อสอบข้อที่สามเป็นเพียงข้อเดียวที่แสดงให้เห็นว่ามีความผันแปรในระหว่างสถานศึกษา ดังนั้นอิทธิพลของการเกิด DIF จึงมีค่าไม่คงที่ระหว่างสถานศึกษาที่ใช้ศึกษาซึ่งแสดงให้เห็นว่ามีการเกิด DIF แบบสุ่ม

Williams & Beretvas (2006) ได้ศึกษาการระบุ DIF โดยการใช้โมเดลเชิงเส้นทั่วไปแบบลดหลั่น (HGLM) ในข้อสอบแบบให้คะแนนมากกว่าสองค่า มีการศึกษาถึงความสัมพันธ์ระหว่าง HGLM และ IRT โดยแสดงให้เห็นถึงความเท่าเทียมกันทางฟังก์ชันพีชคณิตระหว่างพารามิเตอร์จาก Muraki's rating scale model และ โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่นสำหรับข้อสอบที่ตรวจให้คะแนนหลายค่า (PHGLM) มีการจำลองข้อมูลเพื่อเปรียบเทียบการประมาณค่าพารามิเตอร์แฝงของข้อสอบและคนสอบ โดยการใช้โปรแกรม HLM 5 และ PARSCALE และเปรียบเทียบความสามารถของ PHGLM โดยสถิติในการระบุ DIF กับวิธี Generalized Mantel-Haenszel (GMH) โดยทั้งสองวิธีการดำเนินการแล้วพบว่ามีการแสดงคล้ายกัน ข้อค้นพบจากการศึกษาชี้ให้เห็นว่า PHGLM ให้ผลที่เท่าเทียมกับการใช้สถิติ GHM ทั้งสองวิธีสามารถระบุ DIF

Cheong (2006) ศึกษาอิทธิพลของบริบทสถานศึกษาที่มีต่อการทำหน้าที่ต่างกันของข้อสอบโดยใช้โมเดลเชิงเส้นทั่วไปแบบลดหลั่น โดยให้ความสนใจไปที่การทำหน้าที่ต่างกันของข้อสอบมีความเกี่ยวข้องกับกลุ่มที่แตกต่างกันทางเชื้อชาติสองกลุ่มที่อาจจะได้รับรูปแบบการสอนที่ต่างกัน โดยพิจารณารูปแบบการสอนที่ต่างกันร่วมกับความผันแปรของสถานศึกษา ผลการ

ตรวจสอบที่ได้เป็นความลำเอียงเนื่องจากการสอนที่มีอยู่ในข้อสอบ Cheong ได้ศึกษาถึงบริบทของสถานศึกษาเช่นโอกาสในการสอน แหล่งของระหว่างกลุ่มที่มีความต่างกันในหน้าที่ของข้อสอบโดยการควบคุมคุณลักษณะเฉพาะของสถานศึกษา ผลการศึกษาพบว่ายังมีการเกิด DIF และมีความแตกต่างของเชื้อชาติในการแสดงออกของข้อสอบมีความสัมพันธ์กับโอกาสในการเรียนรู้เนื่องจากครู และพบข้อสอบสองข้อเป็นข้อสอบที่มีตัวแทรกซ้อนเนื่องจากโอกาสในการสอนของครูและนักเรียนผิวขาวมีโอกาสในการเรียนการสอนมีสูงส่วนนักเรียนชาวอเมริกันมีโอกาสต่ำ

Chaimongkol (2005) ศึกษาถึงโมเดลของ DIF โดยใช้โมเดลถดถอยโลจิสติกพหุระดับ และใช้การประมาณค่าเบซี่ยน โมเดลการตรวจสอบ DIF ได้พิจารณาพารามิเตอร์ของ DIF ที่แปรค่าตามสถานศึกษา นั่นคือลักษณะเฉพาะของสถานศึกษาเป็นองค์ประกอบหนึ่งที่เป็นอิทธิพลสุ่มที่มีการแปรค่าได้ องค์ประกอบลักษณะของสถานศึกษาจะแสดงถึงการเพิ่มขึ้นของขนาดของ DIF ที่มีสถานศึกษาเป็นลักษณะของความสามารถนักเรียนและสถานศึกษา ส่วนเบี่ยงเบนมาตรฐานที่มีค่ามากแสดงให้เห็นถึงขนาดของ DIF ของข้อสอบแปรค่าข้ามสถานศึกษา และหลังจากที่ควบคุมความสามารถของนักเรียนและสถานศึกษาแล้ว การแปรค่าของ DIF จะใช้ในการอธิบายตัวแปรสถานศึกษา เช่นตัวแปรแหล่งของสถานศึกษา ในการศึกษานี้ได้อธิบายถึงการนำโมเดลไปใช้เพื่อระบุข้อสอบ DIF และอธิบายแหล่งของการเกิด DIF ของตัวแปรระดับสถานศึกษา เช่นรายวิชาที่สอนโดยการสอนนอกห้องเรียน จำนวนผู้สนับสนุน ทีมงานของผู้บริการ ตัวแปรบริบทของสถานศึกษา เช่น อัตราความคงที่ของนักเรียน อัตราการพักเรียนของนักเรียน

Binici (2007) ได้ศึกษาอิทธิพลสุ่มของการทำหน้าที่ต่างกันของข้อสอบโดยใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น(HGLM) และโมเดลผสมผสานตัวแปรแฝงเชิงเส้นทั่วไป (Generalized Linear Latent and Mixed Model: GLLAMM) โดยใช้วิธีการเปรียบเทียบการประมาณค่า ในการศึกษานี้ได้จัดการให้ DIF เป็นพารามิเตอร์แบบสุ่มที่แปรค่าตามหน่วยของกลุ่มโดยใช้หลักการของ GLLAMM และ HLM มีการเปรียบเทียบการประมาณค่า HLM และ GLLAMM ข้ามเงื่อนไขต่างๆ โดยทั้งสองหลักการมีการประมาณค่า marginal maximum likelihood ที่แตกต่างกันของวิธีการ โดย HLM จะใช้วิธี Penalized Quasi Maximum Likelihood (PQL)และ Laplace approximations ส่วน GLLAMM จะใช้วิธีการ Adaptive Gaussian Quadrature (AGQ) method ความถูกต้องของการตรวจสอบ DIF พิจารณาจากความคลาดเคลื่อนที่ประเภทที่ 1 และอำนาจการทดสอบของของการตรวจสอบ DIF อัตราความคลาดเคลื่อนประเภทที่ 1 ของวิธีการ PQL และ Laplace มีค่าต่ำ และวิธีการ Laplace จะให้ความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่า PQL ทุกเงื่อนไขวิธีการตรวจสอบ DIF ของวิธี PQL มีอำนาจการทดสอบมากกว่าวิธี Laplace โดยอำนาจการทดสอบนี้จะให้ความเร็วสำหรับวิธีการประมาณค่า ซึ่งขึ้นอยู่กับจำนวนกลุ่มตัวอย่างที่เพิ่มมากขึ้นของระดับนักเรียนและสถานศึกษา โดยขนาดกลุ่มที่มีขนาดใหญ่และจำนวนกลุ่มจะให้ผลที่ถูกต้องแม่นยำ

การศึกษาจากผลงานวิจัยหลายท่านพบว่า การใช้โมเดล HGLM สามารถระบุการเกิด DIF ได้ และสามารถระบุแหล่งของการเกิด DIF ได้ทั้งในระดับนักเรียน เช่นตัวแปรเพศ เชื้อชาติ และ

ตัวแปรระดับสถานศึกษา เช่นรายวิชาที่สอนโดยการสอนนอกห้องเรียน จำนวนผู้สนับสนุน ทีมงานของผู้บริการ อัตราความคงที่ของนักเรียน อัตราการพักเรียนของนักเรียน ดังนั้นการใช้โมเดล HGLM ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ จึงทำให้สามารถเชื่อมั่นได้ว่าเมื่อตัดข้อสอบที่เกิด DIF แล้ว จะทำให้ผลการวัดโดยใช้แบบสอบนั้นได้ค่าที่มีความตรงและถูกต้องมากยิ่งขึ้น

3.4 การประยุกต์ใช้โมเดลเชิงเส้นแบบลดหลั่น (HLM)สำหรับการวิเคราะห์มูลค่าเพิ่ม

การวิเคราะห์มูลค่าเพิ่มมีหลายวิธีดังที่กล่าวไว้ในตอนที่ 1 โดยวิธีการหักออกและการวิเคราะห์ถดถอยยังไม่มีความเหมาะสม เนื่องจากโครงสร้างของข้อมูลที่มีลักษณะลดหลั่นกล่าวคือ นักเรียนสอดแทรกอยู่ในสถานศึกษา ในขณะที่จุดประสงค์คือการวัดประสิทธิผลของหน่วยงานหรือสถานศึกษา และผลลัพธ์ (outcomes) หรือผลการเรียนรู้เป็นผลที่เกิดขึ้นจากนักเรียน ดังนั้นโมเดลที่เหมาะสมสำหรับการวิเคราะห์มูลค่า จึงควรพิจารณาโมเดลเชิงเส้นแบบลดหลั่น (HLM)

3.4.1 การตรวจสอบความเหมาะสมของโมเดล

ความเหมาะสมของโมเดล (Accuracy of a Model) พิจารณาในโมเดล 2 ระดับ ดังนี้ (ศิริชัย กาญจนวาสี, 2550)

โมเดลภายในกลุ่ม (Within-group model)

$$Y_{ij} = B0_j + B1_j X_{ij} + R_{ij}$$

โมเดลระหว่างกลุ่ม (Between-group model)

$$B0_j = G00 + G01Z_j + U0_j$$

$$B1_j = G10 + G11Z_j + U1_j$$

การตรวจสอบความเหมาะสมของโมเดลทำได้ดังนี้

1) ตัวแปร X_{ij} ส่งผลอย่างมีนัยสำคัญทางสถิติต่อ Y_{ij} หรือไม่

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

2) ค่าพารามิเตอร์ของแต่ละหน่วยมีความผันแปรระหว่างหน่วยหรือไม่

HLM จะแบ่งผลของพารามิเตอร์ของแต่ละกลุ่มออกเป็นอิทธิพลคงที่(Fixed effect) คือ $G00, G10$ และอิทธิพลสุ่ม (Random effect) คือ $U0_j, U1_j$ ดังนี้

$$B0_j = G00 + U0_j$$

$$B1_j = G10 + U1_j$$

การทดสอบค่าเฉลี่ยของอิทธิพลคงที่ว่าเป็นศูนย์หรือไม่ โดยใช้สถิติทดสอบที (t-test) มีสมมติฐานดังนี้

H_0 = G00 = 0 , H_1 = G00 $\neq$ 0

และ $H_0 = G10 = 0$, $H_0 = G10 \neq 0$

ถ้าอิทธิพลคงที่ (Fixed effect) ไม่เป็นศูนย์แสดงว่าค่าคงที่หรือค่าตัดแกน (Intercept) มีผลต่อ Y_{ij} หรือ $B1_j$ แต่ถ้ามีค่าเป็นศูนย์ แสดงว่าค่าคงที่ไม่มีผลต่อ Y_{ij} หรือ $B1_j$

การทดสอบความแปรปรวนของอิทธิพลสุ่มว่าเป็นศูนย์หรือไม่ โดยใช้สถิติทดสอบไคสแควร์ (χ^2 –test) มีสมมติฐานดังนี้

$H_0 = Var(B0_j) , Var(U0_j) = 0$, $H_1 = Var(B0_j) , Var(U0_j) > 0$

และ $H_0 = Var(B1_j) , Var(U1_j) = 0$, $H_1 = Var(B1_j) , Var(U1_j) > 0$

ถ้าอิทธิพลสุ่ม (Random effect) ไม่เป็นศูนย์แสดงว่าค่า $B0_j$ และ $B1_j$ มีความผันแปรระหว่างกลุ่ม แต่ถ้าไม่เป็นศูนย์ แสดงว่า ค่า $B0_j$ และ $B1_j$ ไม่มีความผันแปรระหว่างกลุ่ม และพบว่าค่าพารามิเตอร์ $B0_j , B1_j$ มีความผันแปรระหว่างกลุ่มจึงสามารถนำตัวแปรทำนายระหว่างกลุ่มมาทำการอธิบายความผันแปรดังกล่าว

3.4.2 ขั้นตอนการวิเคราะห์

การวิเคราะห์ใน HLM จะเริ่มจากขั้นตอนที่ 1 ที่ยังไม่มีเงื่อนไขหรือตัวแปรทำนายในระดับที่ 1 และเพื่อเป็นการทดสอบค่าเฉลี่ยของตัวแปรตามว่ามีค่าแตกต่างจากศูนย์หรือไม่ ขั้นตอนที่ 2 เป็นการนำตัวแปรทำนายระดับบุคคลเข้ามาวิเคราะห์ และขั้นตอนที่ 3 จะนำตัวแปรทำนายระดับบุคคลและระดับกลุ่มเข้ามาวิเคราะห์ร่วมด้วย ดังนี้

1) วิเคราะห์ขั้นโมเดลศูนย์ (Null model) หรือแบบไม่มีเงื่อนไขอย่างสมบูรณ์ (Fully Unconditional Models) เป็นการวิเคราะห์ขั้นแรกเพื่อให้เห็นภาพรวมของตัวแปรเกณฑ์หรือตัวแปรตาม โดยไม่นำตัวแปรทำนายเข้ามาร่วมพิจารณา และเพื่อตรวจสอบว่าตัวแปรเกณฑ์มีความแปรปรวนในหน่วยและระหว่างหน่วยเพียงพอที่จะวิเคราะห์ตัวแปรทำนายที่มีอิทธิพลในขั้นตอนหรือไม่ มีสมการวิเคราะห์ดังนี้

ระดับที่ 1 ภายในกลุ่ม

$$Y_{ij} = B0_j + R_{ij}$$

ระดับที่ 2 ระหว่างกลุ่ม

$$B0_j = G00 + U0_j$$

2) วิเคราะห์ขั้นโมเดลอย่างง่าย (Simple Model) หรือแบบไม่มีเงื่อนไข (Unconditional Models) เป็นการวิเคราะห์โดยการนำตัวแปรทำนายระดับนักเรียนเข้ามาวิเคราะห์ (SES) เพื่อดูว่าตัวแปรทำนายเหล่านั้นมีอิทธิพลต่อ b_{00} และ b_{01} หรือไม่ มีสมการวิเคราะห์ดังนี้

ระดับที่ 1 ภายในกลุ่ม

$$Y_{ij} = B0_j + B1_j X_{ij} + R_{ij}$$

ระดับที่ 2 ระหว่างกลุ่ม

$$B0_j = G00 + U0_j$$

$$B1_j = G10 + U1_j$$

3) วิเคราะห์ขั้นโมเดลสมมติฐาน (Hypothetical model) โดยการนำตัวแปรทำนายระดับนักเรียน และพิจารณาว่ามีความเหมาะสมจากการวิเคราะห์ระดับนักเรียนมาวิเคราะห์ร่วมกับตัวแปรทำนายในระดับสถานศึกษา เพื่อตรวจสอบอิทธิพลของตัวแปรระดับสถานศึกษาที่มีต่อตัวแปรเกณฑ์ มีสมการวิเคราะห์ดังนี้

ระดับที่ 1 ภายในกลุ่ม

$$Y_{ij} = B0_j + B1_j X_{ij} + R_{ij}$$

ระดับที่ 2 ระหว่างกลุ่ม

$$B0_j = G00 + G01Z_j + U0_j$$
$$B1_j = G10 + G11Z_j + U1_j$$

3.4.3 สัมประสิทธิ์การทำนาย

สัมประสิทธิ์การทำนายของโมเดลแต่ละระดับ (R^2) หรือสัดส่วนความแปรปรวนที่อธิบายได้ (Proportion variance explained) สามารถคำนวณได้จากการเทียบสัดส่วนความแปรปรวนของค่าส่วนที่เหลือที่ลดลง เมื่อมีตัวแปรทำนายกับเมื่อไม่มีตัวแปรทำนาย มีสูตรดังนี้

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

หรือ

$$R^2 = \frac{\tau_{00}(\mathit{fully\ unconditional}) - \tau_{00}(\mathit{hypothetical})}{\tau_{00}(\mathit{fully\ unconditional})}$$

เมื่อ $\tau_{00}(\mathit{fully\ unconditional})$ และ $\tau_{00}(\mathit{hypothetical})$ เป็นความแปรปรวนของค่าเศษเหลือในสมการที่ยังไม่ใส่ตัวแปรทำนาย และสมการที่ใส่ตัวแปรทำนาย ตามลำดับ

ดังนั้น ความแปรปรวนทั้งหมดของตัวแปรตามที่อธิบายได้โดยโมเดล 2 ระดับ หาได้จากผลรวมของความแปรปรวนที่อธิบายได้ในระดับที่ 1 และระดับที่ 2

3.4.4 โมเดลการวิเคราะห์มูลค่าเพิ่ม

ในการวิเคราะห์มูลค่าเพิ่มหรือค่าส่วนที่เหลือ โดยการประยุกต์ใช้โมเดลพหุระดับสำหรับควบคุมปัจจัยที่นอกเหนือการจัดการหรือควบคุมของหน่วยงาน ทั้งปัจจัยระดับบุคล เช่น ภูมิหลังคุณลักษณะของบุคคล และปัจจัยระดับหน่วยงาน เช่นบริบทชุมชน ลักษณะสถานศึกษา เพื่อให้การเปรียบเทียบผลสัมฤทธิ์ของการบริหารจัดการอยู่บนพื้นฐานของปัจจัยที่ทัดเทียมกัน โมเดลการวิเคราะห์ข้อมูลสอง ระดับ คือระดับบุคคลและระดับหน่วยองค์กร สำหรับการศึกษามูลค่าเพิ่มเป็นดังนี้

ระดับที่ 1 โมเดลระหว่างบุคคลในหน่วยงาน

$$Y_{ij} = B0_j + \sum_{p=1}^{k} Bp_j X_{pij} + R_{ij}$$

เมื่อ $B0_j$ แทนค่า intercept หรือค่าเฉลี่ยของ Y_{ij} เมื่อควบคุมผลของ X_{pij} แล้ว

Bp_j แทนค่าสัมประสิทธิ์ถดถอยที่แสดงผลของ X_{pij} ต่อ Y_{ij}

X_{pij} แทนตัวแปรควบคุมระดับบุคคล เมื่อ p=1,2,3…,k

R_{ij} แทนค่าส่วนที่เหลือหรืออิทธิพลสุ่มระดับบุคคล

ระดับที่ 2 โมเดลระหว่างหน่วยงาน

$$B0_j = G00 + \sum_{q=1}^{l} G0_q Z_j + U0_j$$

$$Bp_j = Gp0 + \sum_{q=1}^{l} G0_q Z_j + Up_j$$

เมื่อ $G00$ แทนค่า intercept หรือค่าเฉลี่ยของ $B0_j$ เมื่อควบคุมผลของ Z_j แล้ว

$G0_q$ แทนค่าสัมประสิทธิ์ถดถอยที่แสดงผลของ Z_j ต่อ $B0_j$

Z_j แทนตัวแปรควบคุมระดับหน่วยงาน เมื่อ j =1,2,3…,k

$U0_j$ แทนค่าส่วนที่เหลือ(residual terms) หรือมูลค่าเพิ่ม(value added) ของหน่วยงาน j ซึ่งแสดงผลต่างระหว่างค่าเฉลี่ยของ Y_{ij} ที่ทำนายได้จากสมการ

เมื่อพิจารณาโมเดลการวิเคราะห์ในแต่ระดับแล้ว จะเห็นว่าเป็นโมเดลที่ความเหมาะสมในการศึกษาถึงโครงสร้างของข้อมูล กล่าวคือเป็นโมเดลที่มีความเชื่อมโยง หรือศึกษาถึงความสัมพันธ์ของตัวแปรระหว่างผลลัพธ์และตัวแปรที่นำมาใช้ในการอธิบาย และนำผลที่ได้คือคะแนนมูลค่าเพิ่มมาใช้จัดอันดับสถานศึกษา เพื่อบ่งบอกถึงคุณภาพของการดำเนินงานของสถานศึกษา เมื่อมีการควบคุมตัวแปรอื่นที่อยู่นอกเหนือการควบคุมของสถานศึกษา แต่อย่างไรก็ตาม การจัดอันดับสถานศึกษาอาจมีความเสี่ยงที่เกิดจากการแปลความหมายผิดพลาด (Ladd and Walsh, 2002) การนำไปใช้โดยส่วนมากจะสามารถพิจารณาสถานศึกษาที่อยู่ในระดับล่างสุดและสูงสุด ซึ่งมีอันดับที่ความแตกต่างอย่างชัดเจน ดังนั้นการพิจาณาผลการจัดอันดับจึงต้องระวังผลที่ได้ ทั้งนี้ต้องมีการกำหนดโมเดลที่มีความเหมาะสมในการวิเคราะห์มูลค่าเพิ่ม และการเลือกตัวแปรที่มีคุณภาพ เพื่อใช้ผลที่มาแปลความหมายป้อนกลับให้แก่สถานศึกษาเพื่อเป็นแนวทางแสดงความรับผิดชอบที่ตรวจสอบได้

ตอนที่ 4 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) ในการศึกษาครั้งนี้ใช้ฐานข้อมูลจากการสำรวจและการประเมินผลของนักเรียน ผู้วิจัยจึงเสนอลักษณะของโครงการประเมินและเครื่องมือที่ใช้ในการประเมิน ดังนี้

4.1 ความเป็นมาและจุดมุ่งหมายของโครงการประเมินผลนักเรียนระดับนานาชาติ PISA

องค์กรเพื่อความร่วมมือและพัฒนาทางเศรษฐกิจ (Organisation for Economic Co-operation and Development: OECD) ถือว่าคุณภาพของการศึกษาเป็นตัวชี้วัดศักยภาพของการพัฒนาทางเศรษฐกิจ จึงมีการศึกษาวิจัยรูปแบบประสิทธิภาพของระบบการศึกษาของประเทศสมาชิก และป้อนข้อมูลกลับให้ประเทศสมาชิกทราบเกี่ยวกับระบบการศึกษาของประเทศเพื่อเตรียมพร้อมเยาวชนสำหรับอนาคต ในช่วงทศวรรษ ค.ศ. 2000 OECD ได้ดำเนินโครงการประเมินผลนักเรียนนานาชาติ หรือเรียกชื่อโครงการว่า PISA (Programme for International Student Assessment: PISA) เป็นโครงการศึกษาสำรวจความรู้และทักษะของนักเรียนกลุ่มอายุ 15 ปี ซึ่งเป็นวัยจบการศึกษาภาคบังคับ ในประเทศที่เป็นสมาชิกองค์กรเพื่อความร่วมมือและพัฒนาทางเศรษฐกิจ (OECD) และประเทศร่วมโครงการที่ไม่ใช่สมาชิกซึ่งประเทศไทยเป็นหนึ่งในประเทศร่วมโครงการที่ไม่ใช่สมาชิก โดยมีจุดมุ่งหมายเพื่อให้ข้อมูลกับประเทศสมาชิกว่าได้เตรียมความพร้อมของประชาชนให้มีศักยภาพสำหรับการแข่งขันในประชาคมโลกเพียงพอหรือไม่อย่างไร เพื่อหาตัวชี้วัดคุณภาพการศึกษาในระดับนโยบาย

โครงการประเมินผล PISA ต้องการวัดการรู้เรื่อง (Literacy) ของกลุ่มนักเรียนที่จบการศึกษาภาคบังคับของแต่ละระบบการศึกษาหรือกลุ่มนักเรียนอายุ 15 ปี สำหรับการเตรียมพร้อมในอนาคต การประเมินผลจึงเน้นเพื่อชี้อนาคต ให้ความชัดเจนกับความสามารถในการใช้ความรู้และทักษะในชีวิตจริงของผู้ใหญ่ที่ยังอยู่ในวัย 15 ปี ดังนั้น PISA จึงเลือกประเมินสิ่งที่เรียกว่า การรู้เรื่อง (Literacy) ซึ่งเน้นที่สมรรถนะของนักเรียนในการใช้ความรู้และทักษะเพื่อเผชิญกับโลกในชีวิตจริงแทนการประเมินความรู้ที่ได้เรียนตามหลักสูตรในสถานศึกษา ตัวชี้วัดที่สำคัญได้แก่ 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

ตามลำดับ ส่วนใน PISA 2006 เน้นความสำคัญการประเมินการรู้วิทยาศาสตร์ ทั้งนี้วิทยาศาสตร์ มีน้ำหนักของการประเมินเป็น 60% และการอ่านและคณิตศาสตร์มีน้ำหนักอย่างละ 20% ของการประเมิน (OECD, 2007)

จุดมุ่งหมายหลักของ PISA คือการประเมินสมรรถนะของนักเรียนที่จะใช้ความรู้และทักษะเพื่อเผชิญกับโลกในชีวิตจริงมากกว่าการประเมินความรู้ที่ได้เรียนตามหลักสูตรในสถานศึกษา โดย OECD/PISA เรียกสมรรถนะนี้ว่า “การรู้เรื่อง” (Literacy) การประเมินผลของ PISA มุ่งเน้นเพื่อให้ความชัดเจนกับความสามารถในการใช้ความรู้และทักษะในชีวิตจริงของนักเรียน โดยแนวคิดของการประเมินจะสะท้อนให้เกิดการเปลี่ยนแปลงในเป้าหมายและจุดประสงค์ของหลักสูตรในสถานศึกษา ผลการประเมินของ PISA จึงใช้เป็นข้อมูลที่บ่งชี้คุณภาพการจัดการศึกษาของชาติ นอกจากนี้ยังสามารถบอกปัจจัยที่ส่งผลต่อคุณภาพการจัดการศึกษา (สสวท, 2551) การประเมินผลในลักษณะนี้สามารถชี้แนะแนวทางสำหรับการตัดสินใจในระดับนโยบายการส่งเสริมและสนับสนุนตลอดจนการกระจายทรัพยากรทางการศึกษา และทำให้ทราบถึงปัจจัยต่างๆที่ส่งผลต่อการพัฒนาความรู้และทักษะของนักเรียน เพื่อบอกให้ทราบถึงคุณภาพและประสิทธิภาพการจัดการศึกษาของสถานศึกษาในการเตรียมความพร้อมนักเรียน สำหรับการใช้ชีวิตในอนาคต (สสวท, 2547)

4.2 ลักษณะเครื่องมือที่ใช้ในโครงการประเมินผล PISA

การสร้างเครื่องมือตามโครงการประเมินผล PISA ใช้แนวคิดตามทฤษฎีการตอบสนองของข้อสอบ(Item Response Theory) และมีวิธีดำเนินอยู่ภายใต้ข้อตกลงเบื้องต้นของโมเดลการตอบสนองข้อสอบ การสร้างเครื่องมือในแต่ละด้านอยู่ภายใต้ข้อตกลงเบื้องต้นของโมเดลราสช์ (Rasch Model) นั่นคือข้อตกลงเบื้องต้นที่สอดคล้องกับความเป็นเอกมิติของแบบสอบ (unidimensional models) ซึ่งการวิเคราะห์ความเป็นเอกมิติของแบบสอบสามารถใช้วิเคราะห์องค์ประกอบ และได้มีการนำไปใช้อย่างหลากหลาย เช่น การนำไปปฏิบัติเกี่ยวกับข้อตกลงเบื้องต้นของความเป็นเอกมติของแบบสอบของ Blum (2001 cited in Goldstein, 2004) กล่าวถึงการตรวจสอบความเป็นเอกมิติของแบบสอบว่า ถึงแม้ว่าการตรวจสอบองค์ประกอบพบว่ามีสองมิติ แต่แสดงให้เห็นว่ามีจำนวนข้อสอบในมิติหนึ่งอยู่จำนวนน้อย ดังนั้นเครื่องมือนั้นจึงจัดว่าเป็นการวัดเพียงมิติเดียว (Blum, et al, 2001 cited in Goldstein, 2004)

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

4.2.1 การรู้เรื่องวิทยาศาสตร์ (Scientific Literacy)

คำว่า วิทยาศาสตร์ จะประกอบด้วยสองสมรรถนะ ได้แก่ วิทยาศาสตร์คือองค์ความรู้ (Science as a body of knowledge) และวิทยาศาสตร์คือกระบวนการเสาะแสวงหาความรู้ (Science as enquiry) ในนิยามของ PISA การรู้เรื่องวิทยาศาสตร์ (Scientific Literacy) จึงใช้ความหมายของความรู้ทั้งสองแบบ โดยเรียกว่า “ความรู้วิทยาศาสตร์” และ “ความรู้ทางวิทยาศาสตร์” โดยความรู้วิทยาศาสตร์หมายถึงองค์ความรู้ซึ่งเป็นความรู้ของโลกธรรมชาติ เช่น ฟิสิกส์ เคมี ชีววิทยา ส่วนความรู้เกี่ยวกับวิทยาศาสตร์หมายถึง ความรู้ในวิธีการหรือกระบวนการหาความรู้ทางวิทยาศาสตร์ หรือวิถีทางที่นำไปสู่เป้าหมายของการได้มากซึ่งความรู้ (OECD, 2006) ประเด็นของ PISA 2006 สำหรับการประเมินผลวิทยาศาสตร์คือการประเมินผลต้องให้ความชัดเจนกับสมรรถนะที่เหมาะสมกับนักเรียนวัย 15 ปี ซึ่งอาจแตกต่างจากการเรียนการสอนและการประเมินผลตามปกติของสถานศึกษา นั่นคือตั้งอยู่บนพื้นฐานว่านักเรียนวัย 15 ปี ควรจะรู้อะไร ทำอะไร และให้คุณค่ากับอะไรภายในขอบเขตของบริบทส่วนตัว บริบทของสังคม และบริบทของโลก

จากนิยามของความรู้ทางวิทยาศาสตร์ของ PISA 2006 จึงกำหนดองค์ประกอบดังนี้

1) สมรรถนะทางวิทยาศาสตร์ ซึ่งประกอบด้วยสามส่วนคือ ส่วนแรกเป็นการใช้ความรู้วิทยาศาสตร์ในการระบุประเด็นทางวิทยาศาสตร์ (Identifying scientific issues)ที่ต้องการให้นักเรียนรู้ว่าประเด็นใดเป็นประเด็นที่สามารถพิจารณาในเชิงวิทยาศาสตร์ และรู้ถึงลักษณะสำคัญของการสำรวจตรวจสอบและแก้ปัญหาทางวิทยาศาสตร์ ส่วนที่สองเป็น การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ (Explain phenomena scientifically) ที่ต้องการให้ใช้ความรู้วิทยาศาสตร์มาตีความปรากฏการณ์และทำนายการเปลี่ยนแปลงที่จะเกิดขึ้นในสถานการณ์ที่กำหนด และส่วนที่สามเป็นการใช้ประจักษ์พยานทางวิทยาศาสตร์ (Using scientific evidence) ซึ่งหมายรวมถึงการแปลความประจักษ์พยาน การลงข้อสรุป อธิบายระบุข้อตกลงเบื้องต้น หาประจักษ์พยานและเหตุผลเพื่อสนับสนุน

2) ความรู้ทางวิทยาศาสตร์ ประกอบด้วยสองส่วน คือความรู้ในเรื่องโลกธรรมชาติ เรียกว่า “ความรู้วิทยาศาสตร์” กับความรู้ที่เกี่ยวข้องกับกระบวนการสืบเสาะหาความรู้วิทยาศาสตร์เรียกว่า “ความรู้เกี่ยวกับวิทยาศาสตร์”

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

4.2.2 การรู้เรื่องคณิตศาสตร์ (Mathematics Literacy)

จุดประสงค์ของการประเมินการรู้เรื่องคณิตศาสตร์ของ PISA เพื่อศึกษาเยาวชนอายุ 15 ปี ในความสามารถเกี่ยวกับการรับรู้สาระ มีข้อมูลข่าวสาร การใช้ความรู้คณิตศาสต์และความเข้าใจเพื่อช่วยให้เข้าใจประเด็นหรือความจำเป็นต่างๆ ได้อย่างมีความหมายและทำให้ภารกิจสำเร็จลุล่วงได้ โดยภารกิจการประเมินการรู้เรื่องคณิตศาสตร์ของ PISA เน้นความชัดเจนที่ต้องการให้นักเรียนเผชิญหน้ากับปัญหาทางคณิตศาสตร์ที่มีอยู่ในชีวิต ในกระบวนการนี้ต้องอาศัยทักษะหลายอย่าง เช่น ทักษะการคิดและการให้เหตุผล ทักษะการโต้แย้ง การสื่อสาร ทักษะการสร้างตัวแบบ การตั้งปัญหาและการแก้ปัญหา การดำเนินการ

กรอบการประเมินผลของ PISA ที่เน้นการประเมินผลว่านักเรียนอายุ 15 ปีรู้เรื่องคณิตศาสตร์หรือสามารถจัดการกับคณิตศาสตร์ในลักษณะที่แสดงออกว่ามีฐานความรู้ และสามารถจะเผชิญหน้ากับปัญหาในโลกของความเป็นจริง ขอบเขตของการประเมินคณิตศาสตร์วัดครอบคลุม 3 ด้านคือ สถานการณ์หรือบริบท เนื้อหาของคณิตศาสตร์ และสมรรถนะทางคณิตศาสตร์

4.2.3 การรู้เรื่องการอ่าน (Reading Literacy)

เนื่องจาก PISA เน้นความสำคัญของความรู้และทักษะที่จำเป็นเพื่อการปรับตัวเข้ากับโลกที่เปลี่ยนแปลง นั่นคือทักษะของการเรียนรู้เพื่อใช้ในชีวิต การนิยามการรู้เรื่องการอ่าน จึงมีความหมายมากกว่าการอ่านออกตามตัวอักษรและเข้าใจความหมายของคำที่อ่าน แต่ยังรวมถึงความสามารถในการค้นหาสาระ คิดวิเคราะห์ แปลความ ตีความหมาย การสะท้อนและประเมินสาระที่อ่านที่แสดงว่าเข้าใจถึงจุดประสงค์การเขียน รวมถึงความสามารถในการตีความจากโครงสร้างของเรื่องหรือจากลักษณะเด่นของการเขียน ทั้งนี้การที่นักเรียนรู้เรื่องการอ่าน แสดงให้เห็นว่ามีศักยภาพที่จะมีส่วนในการสร้างสังคมอย่างมีประสิทธิภาพในวัยที่เติบโตเป็นผู้ใหญ่ในสังคม

การประเมินการรู้เรื่องการอ่านของ PISA 2006 ต้องการที่จะประเมินความรู้และทักษะที่กว้างและลึกกว่าการอ่านตามปกติ โดยเน้นให้ความชัดเจนที่การใช้ประโยชน์ของการอ่านในสถานการณ์ต่างๆ มีหลากหลายและในวัตถุประสงค์ที่แตกต่างกัน ดังนั้นการประเมินจึงครอบคลุมถึงภารกิจของนักเรียนที่จะต้องออกไปใช้ชีวิตในโลก ทั้งในการทำงาน สถาบันการศึกษา หรือสังคมที่มีความซับซ้อน

4.3 การประเมินเจตคติทางวิทยาศาสตร์ของนักเรียน

โครงการประเมินผล PISA ได้กำหนดกรอบการประเมินเจตคติทางวิทยาศาสตร์ของนักเรียน ดังนี้

1) การเห็นคุณค่าและสนับสนุนการค้นคว้าทางวิทยาศาสตร์
   กำหนดลักษณะของนักเรียนที่สนับสนุนการค้นหาความรู้ทางวิทยาศาสตร์ ดังนี้
   1.1 รับรู้ถึงความสำคัญของการพิจารณาแง่มุมต่างๆ และโต้แย้งทางวิทยาศาสตร์
   1.2 สนับสนุนการใช้สาระจากข้อเท็จจริงและคำอธิบายที่สมเหตุสมผล
   1.3 แสดงว่าใช้ตรรกะและการคิดที่เป็นเหตุเป็นผลในการลงความเห็นหรือการสรุป
2) ความเชื่อตนเองในการเรียนวิทยาศาสตร์
   กำหนดลักษณะของนักเรียนที่มีความเชื่อตนเองในการเรียนวิทยาศาสตร์ ดังนี้
   2.1 จัดการกับภารกิจทางวิทยาศาสตร์อย่างมีประสิทธิภาพ
   2.2 เอาชนะความยากในการแก้ปัญหาเชิงวิทยาศาสตร์
   2.3 แสดงออกว่ามีความรู้สูงในทางวิทยาศาสตร์

3) ความสนใจวิทยาศาสตร์

นักเรียนที่มีความสนใจวิทยาศาสตร์ มีลักษณะดังนี้

3.1 บ่งชี้ถึงความกระตือรือล้น ใฝ่รู้ในประเด็นทางวิทยาศาสตร์หรือประเด็นที่เกี่ยวข้องกับวิทยาศาสตร์ รวมทั้งงานวิทยาศาสตร์

3.2 แสดงความตั้งใจที่จะหาความรู้ ทักษะ วิทยาศาสตร์เพิ่มเติมโดยใช้ทรัพยากรและวิธีการที่หลากหลาย

3.3 แสดงความตั้งใจที่จะใฝ่หาสาระทางวิทยาศาสตร์และแสดงความสนใจวิทยาศาสตร์อย่างต่อเนื่อง รวมทั้งตั้งใจมองหาอาชีพที่เกี่ยวข้องกับวิทยาศาสตร์

4) ความรับผิดชอบต่อทรัพยากรธรรมชาติและสิ่งแวดล้อม

ลักษณะของนักเรียนที่มีความรับผิดชอบต่อทรัพยากรธรรมชาติและสิ่งแวดล้อมมีดังนี้

4.1 แสดงการรับผิดชอบโดยส่วนตัวในการธำรงรักษาสิ่งแวดล้อม

4.2 แสดงความตระหนักถึงผลกระทบต่อสิ่งแวดล้อมอันเกิดจากการกระทำของตน

4.3 แสดงความพร้อมที่จะทำกิจกรรมด้านการรักษาสิ่งแวดล้อม

นอกจากการประเมินเจตคติทางวิทยาศาสตร์ของนักเรียน PISA 2006 ยังได้มีการสำรวจแรงจูงใจภายในการเรียนวิทยาศาสตร์ แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์ ความรู้สึกรับผิดชอบในประเด็นสิ่งแวดล้อม โดยการสำรวจใช้แบบทดสอบและแบบสอบถาม

4.4 กระบวนการพัฒนาแบบสอบและคุณภาพของแบบสอบในโครงการประเมินผล PISA 2006

กระบวนการพัฒนาเครื่องมือที่ใช้ในโครงการประเมินผล PISA มีการพัฒนาเครื่องมือจากศูนย์การพัฒนาแบบสอบ (Test Development Centres) โดยพิจารณาตามกรอบของนิยามในแต่ละองค์ประกอบหลัก (Domain) ของการประเมิน และมีการอภิปรายในด้านของเนื้อหา สมรรถนะ และความรู้ โดยกลุ่มผู้เชี่ยวชาญในเนื้อหาและนักพัฒนาแบบสอบ รวมทั้งคณะที่ปรึกษาในศูนย์ระดับชาติของแต่ละประเทศที่เข้าร่วมโครงการ ตามกำหนดระยะเวลาการดำเนินการของโครงการประเมินผล PISA การพัฒนาเครื่องมือได้กำหนดไว้ 2 ระยะ (OECD, 2009) ดังนี้

1) การพัฒนาระยะที่หนึ่ง (First phase of development) เป็นกระบวนการพัฒนาข้อสอบที่ให้ผู้พัฒนาแบบสอบซึ่งประกอบด้วยคณะทำงานในระดับประเทศ ร่วมพิจารณาประเด็นในด้านของภาษาที่ใช้ในแต่ละประเทศ การจัดรูปแบบของข้อสอบ สถานการณ์ในข้อสอบ และคำถามในข้อสอบ ในขั้นตอนนี้จะใช้ลักษณะการสัมภาษณ์นักเรียนกลุ่มเล็ก ด้วยวิธีการคิดเสียงดัง(think-aloud

methods) ทั้งการสัมภาษณ์ระดับบุคคลและระดับกลุ่ม เพื่อให้สะท้อนถึงกระบวนการคิดของนักเรียน ผลที่ได้จะทำให้เกิดความเข้าใจที่ชัดเจนของการใช้ภาษาในคำถาม การให้ข้อมูลในสถานการณ์ เพื่อให้นักเรียนตอบสนองในข้อสอบ

2) การพัฒนาในระยะที่สอง (Second phase of development) การพัฒนาในระยะนี้ เป็นการพิจารณาทบทวนข้อสอบในระดับนานาชาติ ซึ่งเป็นกระบวนการที่ให้ข้อมูลสะท้อนถึงความสัมพันธ์ของข้อสอบที่มีความแตกต่างทางด้านวัฒนธรรมและบริบทในแต่ละประเทศ ซึ่งพบว่ามีข้อสอบบางข้อที่ต้องตัดทิ้งออกไป โดยการสำรวจจะดำเนินการต่อไปในขั้นตอนการทดลองใช้ข้อสอบ pilot testing และใช้เป็นขั้นตอนหนึ่งในแต่ละศูนย์การสอบแต่ละประเทศ

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เป็นขั้นตอนหนึ่งในกระบวนการพัฒนาระยะที่สอง โดยการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในระดับประเทศ (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 แสดงให้เห็นว่าแบบสอบที่ใช้ประเมินในโครงการประเมินผล PISA มีข้อสอบที่ทำหน้าที่ต่างกัน เช่น Le (2006) ได้วิเคราะห์หาข้อสอบที่ทำหน้าที่ต่างกันทางด้านเพศ และภาษาที่ใช้ในแบบสอบ และตรวจพบ DIF ใน PISA รอบทดลองใช้ ปี 2006 และ Grisay and Monseur (2007) ได้ตรวจพบ DIF ใน PISA ที่วัดการรู้เรื่องเกี่ยวกับการอ่านในปี 2000 และ 2001 โดยใช้ตัวแปรภาษาในแบบสอบ

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

แสดงรายละเอียดดังตารางที่ 4

ตารางที่ 4 ค่าความเที่ยงโดยเฉลี่ยของแบบสอบในโครงการประเมินผล PISA 2006

องค์ประกอบหลัก	ค่าความเที่ยงของทุกประเทศ	ค่าความเที่ยงของประเทศไทย
วิทยาศาสตร์	0.920	0.88
คณิตศาสตร์	0.892	0.84
การอ่าน	0.891	0.85
เจตคติทางวิทยาศาสตร์	0.892	0.85

4.5 ข้อมูลของสถานศึกษาจากการสำรวจในโครงการประเมินผล PISA 2006

โครงการประเมินผล PISA ใช้เครื่องมือทั้งแบบสอบเพื่อวัดการรู้เรื่องและแบบสอบเพื่อเก็บข้อมูลในด้านอื่น ๆ ได้แก่ข้อมูลนักเรียน และสถานศึกษา จึงทำให้ได้ข้อมูลเพื่อศึกษาความสัมพันธ์กับการเรียนรู้ของนักเรียน ตั้งแต่ PISA 2000 จนถึง PISA 2006 โดยผู้ที่ให้ข้อมูลประกอบด้วยนักเรียน และผู้บริหาร ซึ่งนักเรียนจะให้ข้อมูลจากการตอบแบบสอบถามภายหลังจากการทำแบบสอบเสร็จ ส่วนการเก็บสำรวจข้อมูลทางสถานศึกษาจะให้ผู้บริหารเป็นผู้ให้ข้อมูลในการตอบข้อมูล สามด้านหลักคือด้านแรกเป็นตัวป้อนด้านทรัพยากรสถานศึกษา ซึ่งเน้นในข้อมูลเกี่ยวกับคุณภาพของครู อัตราส่วนของครูต่อนักเรียน ทรัพยากรการเรียน ด้านที่สองเป็นประสิทธิภาพของการเรียนการสอน โดยให้ความสำคัญกับการจัดการในห้องเรียนและกลยุทธ์การสอน และด้านที่สามเป็นประสิทธิภาพของสถานศึกษา ซึ่งให้ความสำคัญกับลักษณะการบริหารจัดการของสถานศึกษา เช่น บรรยากาศของสถานศึกษา การมีอิสระในการบริหารจัดการสถานศึกษา และความเป็นผู้นำทางการศึกษา

จากข้อมูลสามด้านหลักสามารถจำแนกข้อมูลออกเป็น 6 กลุ่ม ดังนี้

1. ทรัพยากรสถานศึกษา ประกอบด้วยข้อมูล

1.1 ทรัพยากรบุคคล เพื่อหาข้อมูลเกี่ยวกับการจัดครูวิทยาศาสตร์เพื่อทำการสอนได้ตรงวุฒิการศึกษา

1.2 ดัชนีการขาดแคลนครู ซึ่งการหาค่าดัชนีจะกำหนดค่าเฉลี่ยเป็น 0 และค่าเบี่ยงเบนมาตรฐานเป็น 1 โดยค่าดัชนีที่เป็นบวกแสดงว่ามีปัญหาอุปสรรคจากการขาดแคลนครูมากกว่าค่าเฉลี่ย และค่าดัชนีที่เป็นลบหมายถึงมีปัญหาน้อยกว่าค่าเฉลี่ย

1.3 สัดส่วนของจำนวนนักเรียนต่อครู ซึ่งแบ่งกลุ่มออกเป็นกลุ่มที่มีสัดส่วนของจำนวนนักเรียน 10 คนต่อครูหนึ่งคน สัดส่วนนักเรียน 11 – 15 คนต่อครูหนึ่งคน และกลุ่มที่มีสัดส่วนนักเรียนมากกว่า 20 คนต่อครูหนึ่งคน

1.4 ดัชนีทรัพยากรการเรียน โดยกำหนดค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายความว่ามีความเพียงพอหรือไม่ขาดแคลนหรือขาดแคลนน้อยกว่าค่าเฉลี่ย และถ้าดัชนีเป็นลบแสดงถึงความขาดแคลนทรัพยากรสูงกว่าค่าเฉลี่ย

1.5 ร้อยละของทรัพยากรการเรียนที่ขาดแคลนในแต่ละแหล่ง ประกอบด้วย อุปกรณ์ในห้องปฏิบัติการวิทยาศาสตร์ วัสดุการเรียนการสอน คอมพิวเตอร์ช่วยสอน โสตทัศนอุปกรณ์ เป็นต้น

1.6 สัดส่วนของคอมพิวเตอร์ต่อนักเรียน ซึ่งตั้งแต่นักเรียน 5 คนต่อคอมพิวเตอร์หนึ่งเครื่อง จนถึงนักเรียนมากกว่า 25 คนต่อคอมพิวเตอร์หนึ่งเครื่อง

1.7 การจัดการสอนวิทยาศาสตร์ เป็นข้อมูลเกี่ยวกับการจัดหลักสูตรรายวิชาบังคับหรือวิชาเพิ่มเติม เช่นการจัดให้นักเรียนเรียนวิทยาศาสตร์ทั่วไป วิทยาศาสตร์บูรณาการ หรือการจัดวิชาแยกเป็นเคมี ชีววิทยา ฟิสิกส์ วิทยาศาสตร์โลก

1.8 เวลาสำหรับการเรียนวิทยาศาสตร์ เป็นข้อมูลเกี่ยวกับปริมาณเวลาเรียนปกติในสถานศึกษา การเรียนเพิ่มเติมนอกเวลาเรียน การทบทวนบทเรียนด้วยตนเอง

1.9 กิจกรรมวิทยาศาสตร์นอกเวลาเรียนปกติ เช่นกิจกรรมทัศนศึกษา การแข่งขันทางวิทยาศาสตร์ สัปดาห์วิทยาศาสตร์ การทำโครงงาน การเข้าร่วมชุมนุมวิทยาศาสตร์

ความสัมพันธ์ระหว่างตัวแปรที่เป็นทรัพยากรทางการศึกษาซึ่งเป็นตัวป้อนให้ระบบการศึกษาภายในสถานศึกษาจัดเป็นตัวแปรที่ส่งผลกระทบโดยตรงต่อการเรียนของนักเรียน ทั้งนี้เนื่องจากตัวแปรของสถานศึกษาแต่ละตัวมีความเกี่ยวโยงสัมพันธ์ซึ่งกันและกัน

2. นโยบายและแนวปฏิบัติของสถานศึกษา ประกอบด้วยข้อมูล

2.1 นโยบายการรับนักเรียน เป็นข้อมูลเกี่ยวกับเกณฑ์การรับนักเรียน เช่น การพิจารณาจากเขตพื้นที่อาศัย ผลการเรียนที่ผ่านมาหรือผลการสอบ ความต้องการของผู้ปกครอง การรับรองจากสถานศึกษาเดิมเป็นต้น

2.2 การแยกสถานศึกษาตามความสามารถ ซึ่งเป็นข้อมูลเกี่ยวกับการแบ่งสายการเรียนในแต่ละโปรแกรม เช่นการแยกสายการเรียนวิทยาศาสตร์ คณิตศาสตร์หรือสายอาชีวศึกษา

2.3 การแยกกลุ่มนักเรียนภายในสถานศึกษา เป็นข้อมูลเกี่ยวกับการจัดการนักเรียนในการแยกกลุ่มเรียนในแต่ละรายวิชา หรือการแยกเรียนในทุกรายวิชา

3. รายได้ของสถานศึกษา เป็นข้อมูลสถานศึกษารัฐและเอกชน ที่มีการได้รับเงินสนับสนุนหรือรายได้ของสถานศึกษาที่ต่างกัน ที่มาของแหล่งรายได้ของสถานศึกษา

4. บทบาทของผู้ปกครอง ประกอบด้วยข้อมูล

4.1 การเลือกสถานศึกษา เป็นข้อมูลที่แสดงถึงการแข่งขันในการเข้าสถานศึกษาที่มีคุณภาพสูง

4.2 แรงกดดันของผู้ปกครองที่มีต่อสถานศึกษา เป็นความคาดหวังของผู้ปกครองที่กดดันให้สถานศึกษามีคุณภาพหรืออยู่ในมาตรฐานวิชาการที่สูงขึ้น

5. การตรวจสอบได้ของสถานศึกษา เป็นการเก็บข้อมูลเกี่ยวกับการรายงานผลการเรียนรู้ต่อผู้ปกครองและสาธารณะ ตลอดจนการสอบมาตรฐานจากภายนอก
6. การเป็นอิสระในการบริหารสถานศึกษา เป็นข้อมูลเกี่ยวกับการมีส่วนร่วมในการตัดสินใจของครูหรือบุคลากรในสถานศึกษาเกี่ยวนโยบายและการบริหารจัดการในสถานศึกษา ในการวิเคราะห์เกี่ยวกับการเป็นอิสระของสถานศึกษา PISA 2006 จะสร้างดัชนีสามตัวจากการวิเคราะห์ ได้แก่ ด้านงบประมาณ และด้านวิชาการ ได้แก่

6.1 ดัชนีชี้บอกอำนาจอิสระในด้านบุคลากร เป็นดัชนีการตัดสินใจเกี่ยวกับการบรรจุ แต่งตั้งครู หรือการกำหนดเงินเดือนครู

6.2 ดัชนีด้านงบประมาณ เป็นดัชนีในการตัดสินเกี่ยวกับการจัดสรรงบประมาณภายในสถานศึกษาเอง

6.3 ดัชนีด้านวิชาการ ซึ่งเป็นดัชนีในการตัดสินใจเกี่ยวกับการเปิดวิชาเรียนและการเลือกเนื้อหาสาระที่เรียน

สรุปข้อมูลที่การศึกษาโครงการประเมินผล PISA 2006 จำแนกตามระดับของข้อมูลเป็น 3 ระดับคือระดับนักเรียน ระดับสถานศึกษา ระดับประเทศ ได้ดังตารางที่ 5

ตารางที่ 5 สรุปความหมายของตัวแปรที่ใช้ศึกษาในโครงการประเมินผล PISA 2006 จำแนกตามระดับของข้อมูล

ตัวแปร	ความหมาย
<u>ระดับสถานศึกษา</u>	
ขนาดสถานศึกษา	จำนวนนักเรียนที่ลงทะเบียนเรียนในสถานศึกษา
ประเภทของสถานศึกษา	จำแนกเป็นสถานศึกษาของรัฐ และสถานศึกษาเอกชน
สังกัดของสถานศึกษา	จำแนกเป็นสถานศึกษาสังกัดสพฐ.,สช., สกอ., กศท.และ กทม.และ อศ.
ตำแหน่งที่ตั้งของสถานศึกษา	พิจารณาจากขนาดชุมชนที่สถานศึกษาตั้งอยู่ จำแนกเป็น ตำแหน่งหมู่บ้าน (น้อยกว่า 3,000 คน) เมืองเล็ก (3,000-15,000 คน) เมือง (15,000-100,000 คน) เมืองใหญ่ (100,000-1,000,000 คน) เมืองใหญ่มาก (มากกว่า1,000,000 คน)
ขนาดห้องเรียน	ค่าเฉลี่ยของจำนวนนักเรียนในห้องเรียน จำแนกเป็น น้อยกว่าหรือเท่ากับ 15 คน,16-20 คน,21-25 คน,26-30 คน,31-35 คน, 36-40 คน, 41-45 คน, 46-50 คน และ มากกว่า 50 คน
แรงกดดันของผู้ปกครองที่มีต่อสถานศึกษา	เป็นบทบาทของผู้ปกครองที่คาดหวังให้สถานศึกษามีคุณภาพหรืออยู่ในมาตรฐานวิชาการที่สูงขึ้น จำแนกเป็น 3 ระดับคือ ผู้ปกครองจำนวนมาก ผู้ปกครองส่วนน้อย ไม่พบเป็นส่วนใหญ่
การขาดแคลนครูวิทยาศาสตร์	เป็นระดับการขาดแคลนครูวิทยาศาสตร์ จำแนกเป็น 4 ระดับคือ ไม่ขาดแคลน ขาดแคลนน้อย ขาดแคลนเป็นบางส่วน ขาดแคลนจำนวนมาก

ตารางที่ 5 สรุปข้อมูลและความหมายของตัวแปรที่ใช้ศึกษาในโครงการประเมินผล PISA 2006 จำแนกตามระดับของข้อมูล (ต่อ)

ตัวแปร	ความหมาย/ลักษณะ
การขาดแคลนครูที่มีคุณวุฒิ	เป็นระดับการขาดแคลนครูที่มีคุณวุฒิ จำแนกเป็น 4 ระดับคือ ไม่ขาดแคลน ขาดแคลนน้อย ขาดแคลนเป็นบางส่วน ขาดแคลนจำนวนมาก
ดัชนีการขาดแคลนครู	ค่าที่แสดงถึงการขาดแคลนครู มีค่าเฉลี่ยเป็น 0 และค่าเบี่ยงเบนมาตรฐานเป็น 1 โดยค่าดัชนีที่เป็นบวกแสดงว่ามีปัญหาอุปสรรคจากการขาดแคลนครูมากกว่าค่าเฉลี่ย และค่าดัชนีที่เป็นลบหมายถึงมีปัญหาการขาดแคลนครูน้อยกว่าค่าเฉลี่ย ดัชนีการขาดแคลนครูได้มาจากมุมมองของผู้บริหารเกี่ยวกับศักยภาพในการสอนในสถานศึกษาในด้านการขาดแคลนครูวิทยาศาสตร์ การขาดแคลนครูคณิตศาสตร์ การขาดแคลนครูภาษาไทย และการขาดแคลนครูวิชาอื่นๆ
ดัชนีทรัพยากรการเรียน	ค่าที่แสดงความเพียงพอของทรัพยากรการเรียน มีค่าเฉลี่ยเป็นศูนย์ ถ้าค่าดัชนีเป็นบวกหมายความว่ามีความเพียงพอหรือไม่ขาดแคลนหรือขาดแคลนน้อยกว่าค่าเฉลี่ย และถ้าดัชนีเป็นลบแสดงถึงความขาดแคลนทรัพยากรสูงกว่าค่าเฉลี่ย หาได้จากข้อคำถาม 7 ข้อ จากมุมมองของผู้บริหารเกี่ยวกับการสอน ได้แก่ ขาดแคลนหรือไม่เพียงพอในอุปกรณ์วิทยาศาสตร์ อุปกรณ์การสอน(เช่นหนังสือเรียน) การเชื่อมต่ออินเตอร์เนต โปรแกรมสำหรับการสอน อุปกรณ์ห้องสมุด อุปกรณ์เครื่องเล่นวิดีโอและเครื่องฉายภาพ
สัดส่วนของจำนวนนักเรียนต่อครู	คิดจากจำนวนนักเรียนในสถานศึกษาหารด้วยจำนวนครูทั้งหมดในสถานศึกษา
ร้อยละงบประมาณของรัฐที่จัดสรรให้สถานศึกษา	คิดจากร้อยละงบประมาณของสถานศึกษาในส่วนเป็นงบประมาณของรัฐที่จัดสรรให้กับสถานศึกษา
ร้อยละงบประมาณจากหน่วยงานภายนอก	คิดจากร้อยละงบประมาณที่สถานศึกษาจัดหาจากหน่วยงานภายนอก เช่น เงินสนับสนุนจากชุมชน ร้านค้าเอกชน และเงินบริจาคของผู้ปกครอง
ดัชนีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์	ค่าที่แสดงการจัดกิจกรรมเพื่อการเรียนรู้ทางวิทยาศาสตร์ของสถานศึกษา มีค่าเฉลี่ยเป็นศูนย์ ถ้าค่าดัชนีเป็นบวกหมายความว่ากิจกรรมต่อการเรียนรู้ของนักเรียน และถ้าดัชนีเป็นลบแสดงถึงจัดกิจกรรมการเพื่อการเรียนรู้ต่ำกว่าค่าเฉลี่ย หาค่าจากการรายงานของผู้บริหารเกี่ยวกับกิจกรรมที่สนับสนุนการเรียนรู้เกี่ยวกับวิทยาศาสตร์ของนักเรียนที่เกิดขึ้นในสถานศึกษา เช่นชุมนุมวิทยาศาสตร์ งานแสดงนิทรรศการวิทยาศาสตร์ การแข่งขันทางวิทยาศาสตร์ การทำโครงการวิทยาศาสตร์ การศึกษานอกสถานที่

ตารางที่ 5 สรุปข้อมูลและความหมายของตัวแปรที่ใช้ศึกษาในโครงการประเมินผล PISA 2006 จำแนกตามระดับของข้อมูล (ต่อ)

ตัวแปร	ความหมาย/ลักษณะ
ดัชนีความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากร	เป็นดัชนีที่มีความสัมพันธ์กับความรับผิดชอบของบุคลากรในการจัดสรรทรัพยากร ค่าดัชนีที่มีค่าสูงหมายถึงมีระดับในการรับผิดชอบได้สูง มีค่าเฉลี่ยเป็นศูนย์ ส่วนเบี่ยงเบนมาตรานเท่ากับ 1 ถ้าค่าดัชนีเป็นบวกหมายความว่ามีความรับผิดชอบต่อการจัดสรรทรัพยากรสูงกว่าค่าเฉลี่ย ถ้าดัชนีเป็นลบแสดงถึงความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากรต่ำกว่าค่าเฉลี่ย ซึ่งคิดได้จากข้อคำถาม 6 ข้อ จากการรายงานของผู้บริหารในการพิจารณาการจัดการเกี่ยวกับการจัดสรรทรัพยากรเช่น การคัดเลือกครูอัตราจ้าง การเพิ่มเงินเดือนครู รูปแบบของงบประมาณสถานศึกษา การตัดสินใจเกี่ยวกับการจัดแบ่งงบประมาณในสถานศึกษา
ดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล	ค่าที่แสดงความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล มีค่าเฉลี่ยเป็นศูนย์ ถ้าค่าดัชนีเป็นบวกหมายความว่ามีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผลสูงกว่าค่าเฉลี่ย ถ้าดัชนีเป็นลบแสดงถึงมีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผลต่ำกว่าค่าเฉลี่ย ได้จากการคำนวณจากการตอบคำถาม 4 ข้อของผู้บริหารเกี่ยวกับความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล เช่นการกำหนดนโยบายการประเมินผล การเลือกหนังสือเรียน การกำหนดเนื้อหาหลักสูตร การตัดสินใจเกี่ยวกับหลักสูตร
สัดส่วนของนักเรียนต่อคอมพิวเตอร์	จำนวนนักเรียนต่อคอมพิวเตอร์ 1 เครื่อง ซึ่งเป็นคอมพิวเตอร์สำหรับนักเรียนอายุ 15 ปี
<u>ระดับนักเรียน</u>	
ผลการประเมินการรู้เรื่องด้านวิทยาศาสตร์	คะแนนที่จากการทำแบบสอบประเมินการรู้เรื่องด้านวิทยาศาสตร์ซึ่งแสดงถึงสมรรถนะการรู้เรื่องวิทยาศาสตร์ตามกรอบการประเมินการรู้เรื่องวิทยาศาสตร์ ในโครงการประเมิน PISA 2006
ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา	ค่าที่ได้จากการคำนวณหาผลต่างระหว่างระดับชั้นของนักเรียนกับระดับชั้นที่นักเรียนคาดหวังของนักเรียน (Chiu & Xihua, 2008)
เพศ	จำแนกเป็นนักเรียนชายและหญิง
เศรษฐานะของครอบครัว	ค่าดัชนีเศรษฐานะของครอบครัว ได้จากคะแนนองค์ประกอบในการวิเคราะห์องค์ประกอบ ซึ่งประกอบด้วยปัจจัยอาชีพของผู้ปกครอง ระดับการศึกษาของผู้ปกครอง และความเป็นเจ้าของในบ้าน
แหล่งทรัพยากรที่บ้าน	ค่าดัชนีที่บ่งบอกถึงความพร้อมของทรัพยากรในบ้าน ได้แก่ หนังสือ คอมพิวเตอร์ เป็นต้น มีค่าเฉลี่ยเป็นศูนย์ ถ้าค่าดัชนีเป็นบวกหมายความว่ามีแหล่งทรัพยากรที่บ้านสูง ถ้าดัชนีเป็นลบแสดงถึงมีแหล่งทรัพยากรที่บ้านต่ำกว่าค่าเฉลี่ย

ตารางที่ 5 สรุปข้อมูลและความหมายของตัวแปรที่ใช้ศึกษาในโครงการประเมินผล PISA 2006 จำแนกตามระดับของข้อมูล (ต่อ)

ตัวแปร	ความหมาย/ลักษณะ
ความมั่งคั่งของครอบครัว	ค่าดัชนีที่บ่งบอกถึงความเป็นอยู่ ความสะดวกสบายของครอบครัว มีค่าเฉลี่ยเป็นศูนย์ ถ้าค่าดัชนีเป็นบวกหมายความว่ามีความมั่งคั่งของครอบครัวสูง ถ้าดัชนีเป็นลบแสดงถึงมีความมั่งคั่งของครอบครัวต่ำกว่าค่าเฉลี่ย ได้จากการตอบคำถามของนักเรียนเกี่ยวกับการมีไว้ครอบครองในครอบครัว เช่นการมีห้องเป็นของตนเอง การเชื่อมต่ออินเตอร์เนต โทรศัพท์มือถือ โทรทัศน์ คอมพิวเตอร์ รถยนต์
ความสนใจในการเรียนรู้วิทยาศาสตร์	ค่าดัชนีที่บ่งบอกถึงความสนใจใฝ่รู้ในประเด็นทางวิทยาศาสตร์ ความตั้งใจที่จะหาความรู้ ทักษะทางวิทยาศาสตร์ มีค่าเฉลี่ยเป็นศูนย์ ถ้าค่าดัชนีเป็นบวกหมายความว่ามีความสนใจในการเรียนรู้วิทยาศาสตร์สูง ถ้าดัชนีเป็นลบแสดงถึงมีความสนใจในการเรียนรู้วิทยาศาสตร์ต่ำกว่าค่าเฉลี่ย หาได้จากการวัดความสนใจในเนื้อหาทางวิทยาศาสตร์ 8 คำถาม ได้แก่ เนื้อหาทางด้านฟิสิกส์ เคมี พืช สัตว์ มนุษย์ ดาราศาสตร์ โลก การออกแบบการทดลองและการอธิบายทางวิทยาศาสตร์
ความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์	ค่าดัชนีที่บ่งบอกถึงความเชื่อมั่นในตนเองในความสามารถที่จะเอาชนะความบากลำบากในการเรียนวิทยาศาสตร์ มีค่าเฉลี่ยเป็นศูนย์ ถ้าค่าดัชนีเป็นบวกหมายความว่ามีความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์สูง ถ้าดัชนีเป็นลบแสดงถึงมีความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์ต่ำ หาได้จากการตอบคำถามการคิดหาคำตอบของนักเรียน เช่นข่าวหนังสือพิมพ์ในประเด็นเกี่ยวกับสุขภาพ การอธิบายเกี่ยวกับการเกิดแผ่นไหวที่เกิดขึ้นบ่อยกว่าแหล่งอื่น การทำนายการเปลี่ยนแปลงในสิ่งแวดล้อมที่เป็นผลการอยู่รอดของสิ่งมีชีวิต การแปลความหมายเกี่ยวกับข้อมูลในป้ายรายการโภชนาการ การอภิปรายเกี่ยวกับความเป็นไปได้ในสิ่งมีชีวิตในดาวอังคาร และ การอธิบายเกี่ยวกับการเกิดฝนกรด
ความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์	ค่าดัชนีที่บ่งบอกถึงความเชื่อมั่นในความรู้ความสามารถเกี่ยวกับวิทยาศาสตร์ของตนเอง มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายความว่ามีความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์สูง ถ้าดัชนีเป็นลบแสดงถึงมีความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์ต่ำกว่าค่าเฉลี่ย หาได้จากการตอบคำถาม 6 ข้อของนักเรียน ซึ่งบ่งชี้ความเชื่อในทางบวกได้แก่ เชื่อว่าการเรียนรู้เนื้อหาขั้นสูงเป็นเรื่องที่ง่ายสำหรับตนเอง สามารถตอบคำถามได้ดี สามารถเรียนรู้ได้รวดเร็ว หัวข้อทางวิทยาศาสตร์เป็นเรื่องง่าย เมื่อเริ่มคิดแล้วจะสามารถเข้าใจเนื้อหาได้อย่างรวดเร็ว สามารถเข้าใจได้โดยง่ายในแนวคิดใหม่ทางวิทยาศาสตร์

ตารางที่ 5 สรุปข้อมูลและความหมายของตัวแปรที่ใช้ศึกษาในโครงการประเมินผล PISA 2006 จำแนกตามระดับของข้อมูล (ต่อ)

ตัวแปร	ความหมาย/ลักษณะ
ความเพลิดเพลินทางวิทยาศาสตร์	ค่าดัชนีที่บ่งบอกถึงความสนุกสนาน ความชื่นชอบในการเรียนวิทยาศาสตร์ มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายความว่ามีความเพลิดเพลินทางวิทยาศาสตร์สูง ถ้าดัชนีเป็นลบแสดงถึงมีความเพลิดเพลินทางวิทยาศาสตร์ต่ำกว่าค่าเฉลี่ย ได้จากการตอบคำถาม 4 ข้อของนักเรียน ได้แก่ มีความสนุกสนานในการเรียนวิทยาศาสตร์ ชอบอ่านเนื้อหาทางวิทยาศาสตร์ มีความสุขเมื่อได้แก้ปัญหาทางวิทยาศาสตร์ ชอบเมื่อได้รับความรู้ใหม่ทางวิทยาศาสตร์ มีความสนใจในการเรียนรู้เกี่ยวกับวิทยาศาสตร์
แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์	ค่าดัชนีที่บ่งบอกถึงแรงจูงใจจากภายนอกในการเรียนรู้ ในการทำกิจกรรมเพื่อเกิดการเรียนรู้วิทยาศาสตร์ มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายความว่ามีแรงจูงใจภายนอกในการเรียนวิทยาศาสตร์สูง ถ้าดัชนีเป็นลบแสดงถึงมีแรงจูงใจภายนอกในการเรียนวิทยาศาสตร์ต่ำกว่าค่าเฉลี่ย ได้จากการตอบคำถามของนักเรียน เช่น ความพยายามเป็นสิ่งที่มีคุณค่าเพราะมันจะช่วยในการทำงาน การเรียนในเนื้อหาทางวิทยาศาสตร์มีความสำคัญมากเพราะเป็นสิ่งจำเป็นในการเรียน ฉันศึกษาเนื้อหาทางวิทยาศาสตร์เพราะมีประโยชน์สำหรับฉัน การเรียนเนื้อหาทางวิทยาศาสตร์มีคุณค่าสำหรับฉันช่วยในการได้งาน
การเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์	ค่าดัชนีที่บ่งบอกถึงความตระหนักถึงคุณค่าของวิทยาศาสตร์ต่อชีวิตประจำวันของนักเรียน มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายความว่ามีการเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์สูง ถ้าดัชนีเป็นลบแสดงถึงมีการเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์ต่ำกว่าค่าเฉลี่ย ได้จากการตอบคำถามของนักเรียน เช่น วิทยาศาสตร์และเทคโนโลยีจะช่วยส่งเสริมชีวิตของมนุษย์ วิทยาศาสตร์เป็นสิ่งสำคัญเพื่อช่วยในการทำเข้าใจธรรมชาติ วิทยาศาสตร์และเทคโนโลยีช่วยส่งเสริมเศรษฐกิจ มีคุณค่าต่อสังคม
การใช้เวลาเรียนพิเศษวิทยาศาสตร์นอกสถานศึกษา	ระดับของการใช้เวลาเรียนในการเรียนพิเศษของนักเรียน แบ่งเป็น 5 ระดับ คือ ไม่มีเวลา น้อยกว่า 2 ชม. 2 – 4 ชม. 4 – 6 ชม. มากกว่า 6 ชม.
การใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง	ระดับของการใช้เวลาเรียนในการศึกษาความรู้ ทบทวนความรู้ด้วยตนเองของนักเรียน แบ่งเป็น 5 ระดับ คือ ไม่มีเวลา น้อยกว่า 2 ชม. 2 – 4 ชม. 4 – 6 ชม. มากกว่า 6 ชม.
จำนวนปีที่นักเรียนเรียนซ้ำชั้น	ช่วงระยะเวลาที่เรียนซ้ำชั้นของนักเรียน

การสำรวจข้อมูลตามโครงการประเมินผล PISA สามารถจัดข้อมูลได้เป็นสามระดับคือระดับนักเรียน ระดับสถานศึกษา และระดับประเทศ ทั้งนี้เพื่อใช้เป็นข้อมูลส่วนหนึ่งในการรายงานผลคุณภาพของการจัดการศึกษา ได้มีการศึกษาวิจัยเกี่ยวกับคุณภาพและประสิทธิผลของระบบการศึกษาของประเทศสมาชิกและประเทศที่เข้าร่วมโครงการ ตลอดจนการศึกษาวิจัยเกี่ยวกับรูปแบบและลักษณะของเครื่องมือที่ใช้ในการประเมินผลที่ผลต่อการตอบสนองข้อสอบเนื่องจากคุณลักษณะพื้นฐานของนักเรียนในประเทศที่เข้าร่วมโครงการ ซึ่งอาจจะเป็นส่วนหนึ่งที่อาจจะทำให้เกิดความไม่เท่าเทียมกันที่เกิดจากสาเหตุของความแตกต่างกันในด้านคุณลักษณะพื้นฐานของนักเรียนและสถานศึกษา เช่นเพศของนักเรียน ภาษาที่ใช้ และลักษณะทางเนื้อหาในหลักสูตรของแต่ละประเทศ ซึ่งกล่าวได้ว่าคุณลักษณะพื้นฐานดังกล่าวเป็นคุณลักษณะแฝงที่ส่งผลต่อการตอบสนองข้อสอบข้อสอบ

4.6 งานวิจัยที่ศึกษาเกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบในเครื่องมือที่ใช้ประเมินผลนักเรียน

การศึกษาผลการวิจัยเกี่ยวกับคุณลักษณะพื้นฐานที่แตกต่างกันของนักเรียนแต่ละประเทศที่ส่งผลต่อการตอบสนองข้อสอบ และศึกษาของสาเหตุของลักษณะของข้อสอบที่ทำหน้าที่ต่างกันในเครื่องมือที่ใช้ประเมินผลนักเรียน พบว่ามีคุณลักษณะของนักเรียนและลักษณะของเนื้อหาในเครื่องมือที่ส่งผลต่อการตอบถูกของนักเรียน ดังผลการศึกษาของงานวิจัยดังนี้

Le (2006) ได้วิเคราะห์หาข้อสอบที่ทำหน้าที่ต่างกันทางด้านเพศ และภาษาที่ใช้ในแบบสอบโดยการใช้ข้อมูลจาก PISA รอบทดลองใช้ ปี 2006 โดยใช้วิธี IRTDIF ในการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกัน และในการวิเคราะห์หาข้อสอบที่ทำหน้าที่ต่างกันได้วิเคราะห์เป็น 5 มิติ ดังนี้ มิติด้านเป้าหมายของข้อสอบ ประกอบด้วยสถานการณ์ในข้อสอบเกี่ยวกับระดับบุคคล ระดับสังคม และระดับสากล มิติด้านเนื้อหาในข้อสอบ มิติด้านสมรรถนะทางวิทยาศาสตร์ มิติด้านความรู้ทางวิทยาศาสตร์ และมิติด้านรูปแบบข้อสอบ ผลการวิเคราะห์ในแต่ละด้านพบว่าด้านเป้าหมายของข้อสอบ พบว่าส่วนใหญ่ข้อสอบโอนเอียงให้เพศชายคิดเป็นร้อยละ 18.1 และเป็นข้อสอบที่โอนเอียงให้เพศหญิงร้อยละ 8.2 ด้านเนื้อหาในข้อสอบพบว่าข้อสอบส่วนใหญ่โอนเอียงให้เพศชายคิดเป็นร้อยละ 27.9 และโอนเอียงให้เพศหญิงร้อยละ 13.3 ด้านสมรรถนะทางวิทยาศาสตร์ พบว่าข้อสอบที่วัดการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์จะมีลักษณะโอนเอียงให้เพศชายมากกว่าเพศหญิงคิดเป็นร้อยละ 9.2 ส่วนข้อสอบที่วัดการระบุประเด็นทางวิทยาศาสตร์ และการใช้ประจักษ์พยานทางวิทยาศาสตร์จะโอนเอียงให้เพศหญิงมากกว่าเพศชาย ส่วนด้านรูปแบบของข้อสอบพบข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบส่วนใหญ่เป็นข้อสอบแบบเลือกตอบ และมีลักษณะโอนเอียงให้เพศชายร้อยละ 16.2

Grisay and Monseur (2007) ได้ศึกษาความเท่าเทียมกันของความยากของข้อสอบในรูปแบบของแบบสอบในโครงการประเมินผล PISA ที่วัดการรู้เรื่องเกี่ยวกับการอ่านในปี 2000 และ 2001 ซึ่งเป็นส่วนหนึ่งของการศึกษาในโครงการ PISA ของ OECD จากผลการศึกษาชี้ให้เห็นว่ามี

ความแปรปรวนของความยากของข้อสอบที่ไม่ได้อธิบายจากปัจจัยที่เหมือนกัน แต่เป็นความแปรปรวนที่แตกต่างระหว่างประเทศที่เข้าร่วมโครงการ เมื่อเปรียบเทียบสัดส่วนของข้อสอบที่ทำหน้าที่ต่างกัน (DIF) ระหว่างกลุ่มประเทศที่มีลักษณะเหมือนกันและต่างกัน จากการวิเคราะห์เมื่อเปรียบเทียบรูปแบบของภาษาที่คล้ายกันกับต่างกัน พบว่ามีข้อสอบที่ทำหน้าที่ต่างกันร้อยละ 25 ถึงร้อยละ 30 ซึ่งคิดเป็นร้อยละที่สูงกว่ารูปแบบภาษาที่คล้ายกัน และเมื่อเปรียบเทียบรูปแบบของภาษาที่ต่างกัน พบข้อสอบที่ทำหน้าที่ต่างกันจำนวนมาก ร้อยละ 10 ถึง 23 หรือโดยเฉลี่ยคิดเป็นร้อยละ 17 นอกจากนี้ยังพบว่าแบบสอบที่มีการแปลงภาษาจากแหล่งต่างกันจะทำให้สูญเสียความเท่าเทียมกัน ซึ่งขึ้นอยู่กับคุณภาพของการแปลงภาษา เช่น เมื่อเปรียบระหว่างประเทศที่มีรูปแบบต่างกันสองรูปแบบ แต่เป็นประเทศที่มีวัฒนธรรมและภูมิศาสตร์แตกต่างกันน้อย เช่นภาษาพูดในภาษาอังกฤษกับฝรั่งเศสในประเทศแคนาดาและเยอรมันพบว่าจำนวนข้อสอบ DIF จะมีจำนวนมากกว่าเมื่อเปรียบเทียบกันในภาษาเดียวกันแต่เป็นประเทศที่อยู่ในระยะห่างออกไปอย่างประเทศไอแลนด์ นิวซีแลนด์และสหรัฐอเมริกา นอกจากนี้ยังพบว่าตัวบ่งชี้ของค่าการร่วมกันในความยากของข้อสอบมีค่าต่ำกว่าค่าที่พึงปราถนาในสองกลุ่มประเทศ คือกลุ่มประเทศเอเชียนห้าประเทศ และกลุ่มประเทศที่มีค่า GDP ต่ำ และพบว่าความเท่าเทียมกันของความยากของข้อสอบในสองกลุ่มประเทศมีค่าน้อย

Baldi, et al. (2007) ได้ศึกษาความสามารถของนักเรียนจากการประเมินผลในโครงการ PISA ปี 2006 ในด้านความแตกต่างของเพศและเชื้อชาติจากประเทศสมาชิก ผลการศึกษาในด้านเพศ พบว่าประเทศสหรัฐอเมริกาไม่มีความแตกต่างกันระหว่างเพศชายและหญิง และพบว่ามี 8 ประเทศที่พบว่าเพศชายมีคะแนนสูงกว่าเพศหญิง ส่วนประเทศที่เพศหญิงมีคะแนนสูงกว่าเพศชายมี 12 ประเทศ และคะแนนเฉลี่ยของ OECD ของเพศชายมีค่าเฉลี่ยสูงกว่าเพศหญิง เมื่อพิจารณาสมรรถนะทางวิทยาศาสตร์พบว่าเพศหญิงมีคะแนนสูงกว่าเพศชายในด้านการระบุประเด็นทางวิทยาศาสตร์ และการใช้ประจักษ์พยานทางวิทยาศาสตร์ ในขณะที่คะแนนของเพศชายสูงกว่าเพศหญิงในสมรรถนะด้านการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ แต่ในประเทศสหรัฐอเมริกาไม่พบความแตกต่างระหว่างคะแนนเพศชายและเพศหญิงในสมรรถนะทางวิทยาศาสตร์ด้านการใช้ประจักษ์พยานทางวิทยาศาสตร์ ส่วนผลการศึกษาในด้านความแตกต่างของเชื้อชาติ พบว่าในประเทศสหรัฐอเมริกานักเรียนชาวผิวดำและชาวฮิสแปนิกส์มีคะแนนต่ำกว่านักเรียนชาวผิวขาว

นอกจากการศึกษาการทำหน้าที่ต่างกันของข้อสอบในเครื่องมือที่ใช้โครงการประเมินผลนักเรียน PISA ยังพบผลการศึกษาจากงานวิจัยที่ศึกษาในโครงการประเมินผลนักเรียน TIMSS ที่สามารถระบุได้ว่าบริบททางวัฒนธรรมของแต่ละประเทศ เพศของนักเรียน และรูปแบบของเนื้อหาในแบบสอบเป็นแหล่งของการทำหน้าที่ต่างกันของข้อสอบ ดังผลการศึกษาของ Amery & Ercikan (2006) ได้ศึกษาการใช้การจับคู่ตัวแปรหลายตัวเพื่อระบุแหล่งของวัฒนธรรมเกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบ โดยใช้การวิเคราะห์ถดถอยโลจิสติกเพื่อระบุแหล่งของ DIF ระหว่างข้อสอบ

TIMSS ที่ประเทศไต้หวันและสหรัฐอเมริกา พบว่า DIF ไม่ได้มีลักษณะเฉพาะของแต่ละข้อสอบในแบบสอบ และไม่เป็นปัจจัยทางวัฒนธรรมที่เป็นแหล่งโดยธรรมชาติของ DIF ข้อจำกัดของแหล่งของ DIF มีความเชื่ออยู่บนบริบทที่มีความเฉพาะและมีการเสนอให้ศึกษาเปรียบเทียบระหว่างประเทศ นอกจากนี้ผลการศึกษาของ Hanan & Hamzeh (2006) ซึ่งได้การวิเคราะห์เนื้อหาของการทำหน้าที่ต่างกันของข้อสอบทางด้านเพศของข้อสอบวิชาคณิตศาสตร์ของ TIMSS วัตถุประสงค์ของการศึกษานี้เป็นการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบที่เกิดขึ้นที่มีความสัมพันธ์กับเพศในข้อสอบวิชาคณิตศาสตร์ของประเทศจอร์แดน โดยใช้วิธี MH ประยุกต์ใช้กับข้อสอบแบบเลือกตอบ ผลที่ได้พบว่ามีข้อสอบจำนวน 37 ข้อที่มีความสัมพันธ์กับเพศที่ส่งผลให้เกิด DIF และการวิเคราะห์เกี่ยวกับเนื้อหาที่เป็นตัวสะท้อนให้เห็นถึงรูปแบบที่จำเป็นในการศึกษาต่อไป ข้อสอบที่ทำหน้าที่ต่างกันมีเนื้อหาที่โอนเอียงให้กับนักเรียนชายในเนื้อหาทางด้านพีชคณิตและเนื้อหาที่โอนเอียงให้นักเรียนหญิงจะเป็นเนื้อหาเกี่ยวกับการวิเคราะห์ข้อมูล นอกจากนี้ยังพบว่าข้อสอบ DIF ส่วนใหญ่ที่มีผลทางลบต่อเพศหญิง จะมีลักษณะเหมือนกับการเกิดความเสี่ยงของคำตอบ เช่นการประมาณค่า การบอกค่าความคาดหวัง และข้อสอบส่วนใหญ่ที่โอนเอียงให้กับเพศหญิงจะเป็นลักษณะที่มีลักษณะที่มีข้อสอบถูกเพียงข้อเดียว จากการศึกษาได้ให้ข้อเสนอแนะว่ายังไม่มีความชัดเจนที่คุณลักษณะของข้อสอบที่ทำหน้าที่ต่างกันจะมีผลต่อความลำเอียงของข้อสอบหรือผลกระทบของข้อสอบ และการพิจารณาทบทวนเนื้อหาจะช่วยให้นักพัฒนาแบบสอบ ตลอดจนครูสามารถนำไปใช้ในการวิเคราะห์ด้านพุทธิปัญญาของนักเรียนชายและหญิงสำหรับการเกิด DIF และไม่เกิด DIF แต่การพิจารณาเนื้อหาของข้อสอบโดยอาศัยผู้เชี่ยวชาญยังมีความยากในการระบุข้อสอบที่ทำหน้าที่ต่างกัน

ส่วนของการศึกษาการทำหน้าที่ต่างกันของข้อสอบในเครื่องมือที่ใช้ประเมินผลนักเรียนในประเทศไทย พบว่าสาเหตุของการทำหน้าที่ต่างกันของข้อสอบมาจากเพศของนักเรียน เชื้อชาติ ดังผลการวิจัยของ สุมาลี แกล้วทนงค์ (2547) ศึกษาสาเหตุของการทำหน้าที่ต่างกันของข้อสอบสาระการเรียนรู้ภาษาไทย และสาระการเรียนรู้สังคมศึกษา ศาสนาและวัฒนธรรม ในด้านเพศ ภาษาพูด และเชื้อชาติ ด้วยวิธีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยวิธีแมนเทล-แฮนส์เซล ด้วยโปรแกรม SIBTEST ผลพบว่าข้อสอบที่ทำหน้าที่ต่างกันด้านเพศ ส่วนใหญ่มีสาเหตุจากเนื้อเรื่องและภาษาที่ใช้ในแบบสอบ ความสนใจในเนื้อเรื่องและวัฒนธรรม ในด้านภาษาพูดมีสาเหตุมาจากเนื้อเรื่องที่สนใจและภาษาที่ใช้ในแบบสอบตามความสนใจและผู้สอบและมีการใช้ภาษาที่เป็นคำศัพท์เฉพาะ และในด้านเชื้อชาติ มีสาเหตุมาจากเนื้อเรื่องที่สนใจและภาษาที่ใช้ในแบบสอบ นอกจากนี้จากการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบในแบบคัดเลือกบุคคลเข้าศึกษาในสถาบันอุดมศึกษา พบว่ามีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ ดังผลการศึกษาของ รักชนก ยี่สุนศรี (2544) ได้วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและแบบสอบวิชาภาษาอังกฤษและคณิตศาสตร์ สำหรับกลุ่มผู้สอบเมื่อจำแนกตามเพศและสถานที่ตั้งทางภูมิศาสตร์ของสถานศึกษา และได้เปรียบเทียบค่าความเที่ยง ความตรง และค่าฟังก์ชันสารสนเทศของแบบสอบระหว่างแบบสอบฉบับก่อนและหลังตัดข้อสอบที่พบ DIF โดยใช้ข้อมูลจากการตอบแบบสอบคัดเลือกบุคคลเข้า

ศึกษาในสถาบันอุดมศึกษา วิชาภาษาอังกฤษและคณิตศาสตร์ ปีการศึกษา 2543 ครั้งที่ 1 ผลการวิเคราะห์พบว่าคุณภาพของแบบสอบในด้านความเที่ยงของแบบสอบทั้งสองวิชา ฉบับก่อนและหลังการตัดข้อสอบที่ DIF พบว่าส่วนใหญ่ค่าความเที่ยงของแบบสอบแตกต่างกันอย่างมีนัยสำคัญ โดยถ้าตัดข้อสอบออกจากแบบสอบจำนวนมากขึ้นจะทำให้ค่าความเที่ยงของแบบสอบฉบับนั้นมีค่าลดลง

จากผลการศึกษาของงานวิจัยที่ศึกษาเกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบในเครื่องมือที่ใช้ในการประเมินผลนักเรียน พบว่าคุณลักษณะที่แตกต่างกันของนักเรียนมีผลต่อโอกาสการตอบข้อสอบถูก เช่น เพศ เชื้อชาติ ภาษาพูด แหล่งทางวัฒนธรรม และตำแหน่งที่ตั้งทางภูมิศาสตร์ ซึ่งอาจจะทำให้เกิดความไม่เท่าเทียมกันในการตอบข้อสอบถูกของนักเรียนเนื่องจากตัวแปรของคุณลักษณะพื้นฐานดังกล่าว ดังนั้นการประเมินคุณภาพการจัดการศึกษาจึงควรคำนึงถึงการทำหน้าที่ต่างกันของข้อสอบซึ่งอาจจะเป็นสาเหตุทำให้ผลการประเมินขาดความถูกต้องและยุติธรรมได้

ตอนที่ 5 เอกสารงานวิจัยเกี่ยวกับการประเมินคุณภาพการจัดการศึกษา

การศึกษาเอกสารและงานวิจัยเกี่ยวกับการประเมินคุณภาพการจัดการศึกษา พบว่าตัวบ่งชี้ที่นำมาใช้บ่งบอกคุณภาพการจัดการศึกษาจะเป็นผลสัมฤทธิ์ทางด้านวิชาการของนักเรียน ซึ่งส่วนใหญ่ได้จากการทำแบบสอบ ดังนั้นในการศึกษาครั้งนี้ ผู้วิจัยจึงสนใจปัจจัยหรือตัวแปรที่ส่งผลผลสัมฤทธิ์ทางการเรียนด้วย ผู้วิจัยจึงสนใจที่จะศึกษานำเสนอตอนนี้ออกเป็น 2 ส่วน คือส่วนแรกเป็นปัจจัยหรือตัวแปรที่เกี่ยวกับคุณภาพการศึกษาจากการทบทวนเอกสารของนักการศึกษา และส่วนที่สองเป็นรายงานวิจัยที่ศึกษาเกี่ยวกับการประเมินคุณภาพการจัดการศึกษาและผลสัมฤทธิ์การเรียนรู้ ดังนี้

5.1 ปัจจัยหรือตัวแปรที่ส่งผลต่อคุณภาพการจัดการศึกษา

คุณภาพการจัดการศึกษา โดยส่วนใหญ่พบว่าเป็นคุณภาพที่เกิดจากการดำเนินของสถานศึกษาเพื่อส่งเสริมให้นักเรียนเกิดการเรียนรู้ ดังนั้นคุณภาพการจัดการศึกษาจึงเกี่ยวข้องกับการสร้างองค์ความรู้ และการประยุกต์ใช้โดยผ่านการเรียนรู้ทำให้เกิดลักษณะเฉพาะของนักเรียนในการที่จะใช้ชีวิตของตัวเอง และการอยู่ร่วมกับผู้อื่น คุณภาพการจัดการศึกษาตามมุมมองของ OECD (OECD, 2005) ได้นิยามโดยการใช้กรอบความคิดที่แสดงถึงระบบการดำเนินการของสถานศึกษาที่แปลงปัจจัยนำเข้า(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) พบว่าให้ความสำคัญกับการเรียนรู้ (Learning) เป็นอันดับแรก รวมทั้งให้ความสำคัญกับความสัมพันธ์ระหว่างผู้เรียนและผู้สอน ปัจจัยนำเข้า (Input) กระบวนการ (Process) สิ่งแวดล้อม (Environment)

และผลผลิต (Outputs) ซึ่งส่งผลกระทบทั้งทางบวกและทางลบต่อคุณภาพการจัดการศึกษา โดยมีการเรียนรู้ประกอบด้วยองค์ประกอบที่เกี่ยวข้องกับระดับผู้เรียน (Inner learner level) และองค์ประกอบส่วนที่เกี่ยวข้องกับระบบการจัดการศึกษา (outer system level) ความแตกต่างของคุณภาพการจัดการศึกษา จึงอาจจะเกิดจากองค์ประกอบแต่ละตัว นอกจากนี้ จากการประเมินของ PISA ยังพบข้อมูลที่สำคัญต่อปัจจัยที่ส่งผลกระทบคุณภาพการจัดการศึกษา โดยแบ่งปัจจัยเหล่านี้เป็นสองระดับคือระดับสถานศึกษา ประกอบด้วยปัจจัยด้านระดับการให้งบประมาณจากรัฐ การแข่งขันของสถานศึกษา ทรัพยากรสถานศึกษา และระดับนักเรียน ประกอบด้วยตัวแปรด้านเวลาในใช้ในการเรียนวิทยาศาสตร์และคณิตศาสตร์ ภูมิหลังทางเศรษฐกิจสังคม (OECD, 2006)

ตัวแปรที่ส่งผลต่อคุณภาพการจัดการศึกษา และผลสัมฤทธิ์ทางการเรียน ที่มีอยู่หลายปัจจัยต่างมีความแตกต่างกันในแต่ละสถานศึกษา ซึ่งตัวแปรเหล่านี้จะนำมาพิจารณาเพื่อนำมาใช้กำหนดเป็นโมเดลในการวิเคราะห์หาคุณภาพการจัดการศึกษาของสถานศึกษา ในการพิจารณาตัวแปรเพื่อกำหนดในโมเดลการวิเคราะห์ต้องพิจารณาหลายส่วนประกอบกัน Salganik (1994) ได้เสนอเงื่อนไขในการพิจารณาตัวแปรที่จะต้องกำหนดในแต่ละระดับดังนี้ 1) ข้อมูลระดับบุคคลของนักเรียนจะต้องสามารถหาได้ง่าย 2) คุณลักษณะพื้นฐานของนักเรียนในโมเดลจะต้องมีความสัมพันธ์กับความสามารถของนักเรียน 3) คุณลักษณะของนักเรียนต้องเป็นสิ่งที่เกินกว่าการควบคุมของสถานศึกษา หรือเป็นสิ่งที่สถานศึกษาคาดหวังให้เกิดการเปลี่ยนแปลง 4) คุณลักษณะจะต้องเป็นที่ยอมรับทางกฎหมายที่เกี่ยวข้องกับสิ่งที่การศึกษาต้องเผชิญ เช่นความยากจน ความแตกต่างทางวัฒนธรรม

เมื่อพิจารณากระบวนการศึกษาและความสัมพันธ์ระหว่างผู้เรียนและผู้สอนที่อยู่ภายในสถานศึกษา ตลอดจนองค์ประกอบอื่นๆ ต่างส่งผลกระทบต่อคุณภาพการจัดการศึกษาได้ในสองส่วน คือ ระดับผู้เรียน ซึ่งเป็นส่วนที่เกี่ยวข้องกับการเรียนรู้แต่ละคน และปัจจัยอื่นที่เป็นสิ่งแวดล้อมที่อยู่รอบตัวนักเรียน เช่นครอบครัว ประสบการณ์หรือเหตุการณ์ที่เอื้อต่อการเรียนรู้ พื้นฐานการเรียน ทัศนคติของนักเรียน และส่วนที่สองเป็นระดับของระบบการศึกษา ซึ่งสถานศึกษามีความรับผิดชอบกับการวางแผน วางนโยบาย ที่ทำให้เกิดการสนับสนุนให้เกิดการเรียนรู้ ดังนั้นการศึกษาในครั้งนี้จึงแบ่งตัวแปรทางการศึกษาออกเป็น 2 ระดับดังนี้

5.1.1 ตัวแปรระดับสถานศึกษา

การศึกษาเกี่ยวกับปัจจัยต่อคุณภาพการจัดการศึกษาได้เพิ่มมากขึ้นตั้งแต่ปลายปี 1980 โดย Riddell (2008) ได้ศึกษาที่ผลต่อคุณภาพการจัดการศึกษาของประเทศกำลังพัฒนาจากรายงานของ Coleman (1966) และ Plowden (1967) พบว่าเป็นการศึกษาของปัจจัยนำเข้า (input) ที่นำไปสู่การผลิตเพื่อออกมาเป็นผลผลิตที่ดีหรือผลลัพธ์ (output) และข้อค้นพบในประเทศกำลังพัฒนาพบว่าอิทธิพลอันดับแรกต่อการเรียนรู้ของนักเรียนคือคุณภาพของสถานศึกษาและครู (Heyneman, 1986 cited in Riddell, 2008) และการศึกษาต่อมาจนถึงปัจจุบัน เป็นการศึกษาที่เน้นการสร้างโมเดลที่มีความซับซ้อนของปัจจัยหลายอย่างที่มีอิทธิพลต่อคุณภาพการศึกษา เป็น

การศึกษาเกี่ยวกับความสัมพันธ์ระหว่างตัวแปรที่เป็นปรากฏการณ์ในสถานศึกษา มีลักษณะที่แปรเปลี่ยนตามพื้นฐาน ความแตกต่างของการสอนในชั้นเรียน ความแตกต่างในสถานศึกษา ความแตกต่างของครู และแหล่งทรัพยากร ซึ่งเป็นการศึกษาที่นำไปสู่โมเดลการวิเคราะห์พหุระดับ (Goldstein 2004; Raudenbush and Bryk, 2002) เช่นเดียวกันกับการศึกษาของ Palardy (2008) เกี่ยวกับปัจจัยที่แตกต่างกันในผลลัพธ์ของนักเรียนระหว่างสถานศึกษา ปัจจัยเหล่านี้สามารถจัดเป็นกลุ่มได้สองกลุ่ม คือปัจจัยนำเข้าของสถานศึกษา (school inputs) ซึ่งประกอบด้วยคุณลักษณะของนักเรียน ตลอดจนแหล่งทรัพยากรของสถานศึกษาและลักษณะโครงสร้าง นโยบายและการปฏิบัติของโรงเรียน ปัจจัยนำเข้าของสถานศึกษาเป็นปัจจัยที่ให้กับสถานศึกษา เช่น บุคลากรซึ่งมีแนวโน้มที่มีอิทธิพลเหนืออิทธิพลอื่น นโยบายและการปฏิบัติของสถานศึกษาประกอบด้วยปัจจัยที่เกี่ยวกับการควบคุม ให้ความสำคัญเกี่ยวกับผู้ปฏิบัติในสถานศึกษาและผู้กำหนดนโยบาย (Fitz-Gibbon & Kochan, 2000; Good & Brophy, 1986; Palardy, 2003; Rumberger & Thomas, 2000 cite in Palardy, 2008) และ Riddell (2008) ได้ทบทวนงานวิจัยปัจจัยที่มีอิทธิพลต่อคุณภาพและประสิทธิผลทางการศึกษาในประเทศกำลังพัฒนา 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 ได้แก่บทบาทของผู้นำ นโยบายเกี่ยวกับการเรียนรู้ 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 ตัวแปรบริบทจึงเป็นสิ่งที่ช่วยอำนวยความสะดวกให้เกิดผลสำเร็จได้ดีขึ้นเมื่อเปรียบเทียบในระหว่างสถานศึกษา 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) ที่มีวัตถุประสงค์เพื่อตรวจสอบปัจจัยที่มีความสัมพันธ์กับความสามารถของนักเรียนในระหว่างประเทศ (Teddlie, & Reynolds, 2002) ได้ระบุอิทธิพลที่เป็นสากล อิทธิพลที่เป็นลักษณะของประเทศและปัจจัยที่อธิบายประสิทธิผลของบริบท ซึ่งพบว่าคุณภาพการสอน หลักสูตรและการดำเนินการมีความสัมพันธ์กันโดยตรงกับนักเรียน ตลอดจนการให้ข้อมูลย้อนกลับต่อนักเรียน โครงสร้างของแบบเรียน การตรวจสอบความเข้าใจของนักเรียน การใช้คำถามที่มีคุณภาพ และความคาดหวังสูงต่อนักเรียน

การดำเนินงานอย่างมีประสิทธิผลของสถานศึกษาขึ้นอยู่ปัจจัยหรือเงื่อนไขในระดับสถานศึกษาและห้องเรียน (Scheerens, 2001) โดยอิทธิพลในตัวแปรระดับสถานศึกษาจะเป็นตัวบ่งชี้ประสิทธิผลของสถานศึกษาโดยส่วนใหญ่ เช่นแหล่งทรัพยากรของสถานศึกษา ประกอบด้วยสัดส่วนของนักเรียนต่อครูอุปกรณ์เครื่องมือ ห้องสมุด และครูที่ผ่านการฝึกอบรม ความสัมพันธ์ของผู้ปกครองกับห้องเรียนและสถานศึกษา การมีวินัยในชั้นเรียน นอกจากนี้ Briggs & Wohlstetter (2003) ได้เสนอ 8 องค์ประกอบที่จะทำให้การบริหารจัดการของสถานศึกษามีความสำเร็จ ได้แก่ การกำหนดวิสัยทัศน์ ความเป็นอิสระของการตัดสินใจ อำนาจ ความรู้และทักษะ ข้อสารสนเทศ รางวัล ความเป็นผู้นำ และแหล่งทรัพยากร

นอกจากนี้จากการศึกษาข้อมูลของสถานศึกษาที่ส่งผลต่อการจัดการศึกษาในการประเมิน PISA ประกอบด้วยข้อมูลสามด้านหลักคือด้านแรกเป็นตัวป้อนด้านทรัพยากรสถานศึกษา ได้แก่ คุณภาพของครู อัตราส่วนของครูต่อนักเรียน ทรัพยากรการเรียน ด้านที่สองเป็นประสิทธิผลของการเรียนการสอนได้แก่การจัดการในห้องเรียนและกลยุทธ์การสอน และด้านที่สามเป็นประสิทธิผลของสถานศึกษา ได้แก่การบริหารจัดการของสถานศึกษา บรรยากาศของสถานศึกษา การมีอิสระในการบริหารจัดการสถานศึกษา และความเป็นผู้นำทางการศึกษา จากผลการประเมินของ PISA พบว่ามีความผันแปรของผลสัมฤทธิ์ของนักเรียนภายในสถานศึกษาและระหว่างสถานศึกษา พบว่ามีความแตกต่างระหว่างสถานศึกษาและอธิบายความแตกต่างระหว่างสถานศึกษาได้ ปัจจัยระดับสถานศึกษาแบ่งเป็น 3 ประเภท คือ ทรัพยากรสถานศึกษา (school resources) สภาพแวดล้อมสถานศึกษา (school climate) และนโยบายสถานศึกษา (school policies) ทรัพยากรสถานศึกษาประกอบด้วย ครูที่มีคุณภาพ, ขนาดชั้นเรียน และทรัพยากรทางการเรียน รวมทั้งคอมพิวเตอร์ สภาพแวดล้อมสถานศึกษา ประกอบด้วยสภาพความเป็นวินัย การคาดหวังต่อผลสัมฤทธิ์ ความสัมพันธ์ระหว่างครูและนักเรียน การรับรู้ในความเป็นเจ้าของนักเรียน และจริยธรรมของครู นโยบายสถานศึกษาประกอบด้วย การให้การบ้าน เวลาในการสอน ความเป็นอิสระของครู ความเป็นอิสระของสถานศึกษา นอกจากนี้ โครงสร้างขององค์กรของระบบการศึกษาในระดับชาติ ประเภทของสถานศึกษาเช่น สถานศึกษาเอกชน และสถานศึกษาของรัฐ ต่างมีความสัมพันธ์กับ

องค์ประกอบของสถานศึกษา โดยส่วนใหญ่พบว่านักเรียนในสถานศึกษาเอกชนมีความสามารถทางวิชาการสูงมาก และเมื่อควบคุมอิทธิพลของเศรษฐานะ (SES) (Matear, 2006; OECD, 2005) หรือในอีกความหมายหนึ่ง สถานศึกษาเอกชนจะมีความสามารถที่ได้เปรียบเนื่องจากคุณลักษณะของนักเรียน และความแตกต่างที่ไม่ใช่การดูแลของสถานศึกษา นักเรียนที่มีเศรษฐานะในระดับกลางในสถานศึกษาเอกชนจะทำได้ดีเช่นเดียวกับสถานศึกษาของรัฐ และในทำนองเดียวกับนักเรียนที่มีเศรษฐานะสูงหรือต่ำ ซึ่งการประเมินของโครงการประเมินผล PISA พบว่าองค์ประกอบของสถานศึกษา ประเภทของสถานศึกษาและสถานที่ตั้งสถานศึกษามีอิทธิพลอย่างมากต่อความสามารถของนักเรียน (OECD, 2005)

5.1.2 ตัวแปรระดับนักเรียน

การศึกษาเกี่ยวกับตัวแปรระดับนักเรียนที่ส่งผลต่อผลการเรียนรู้ จากการศึกษาทบทวนวิจัยของ Teddlie, Stringfield & Reynolds (2000) ได้สรุปผลที่เกิดจากตัวแปรระดับบุคคลด้านภูมิหลังของนักเรียน เกี่ยวกับขนาดของอิทธิพลที่เกิดจากเศรษฐานะสังคม (socioeconomic status: SES) ที่มีผลต่อผลสัมฤทธิ์ว่ามีความแตกต่างกันตามกลุ่มสมาชิกของแต่ละสถานศึกษา นอกจากนี้การศึกษา Chiu & Xihua (2008) พบว่าปัจจัยในด้านครอบครัวก็ส่งผลต่อผลสัมฤทธิ์ของนักเรียน สรุปได้ว่าจำนวนสมาชิกของครอบครัวสามารถตอบสนองแหล่งทรัพยากรภายนอกของนักเรียน (ผู้ให้ทรัพยากร) หรือความสมบูรณ์ของแหล่งทรัพยากร ดังนั้นจำนวนสมาชิกของครอบครัว เช่นพ่อแม่จะมีการให้แหล่งทรัพยากรทางการศึกษาจำนวนมากซึ่งส่งผลต่อโอกาสในการเรียนรู้มากและนักเรียนก็สามารถไปสู่ความสำเร็จได้สูง ตัวอย่างเช่น ครอบครัวที่มีทั้งพ่อและแม่จะมี SES สูง มีแหล่งทรัพยาการทางการศึกษามากกว่า การใช้เวลากับเด็กมากกว่า และมีความเกี่ยวข้องกันมากกว่ากับสถานศึกษาของเด็ก ในทางตรงกันข้ามผู้ปกครองที่แยกกันอยู่จะมีแหล่งทรัพยากรทางการศึกษาน้อยกว่าและมีการติดต่อพบหน้ากับเด็กน้อย ซึ่งอาจจะได้รับความสนใจใส่ใจน้อย เด็กที่มีความขัดแย้งระหว่างพ่อแม่ที่แยกกันอยู่อาจจะจำทนอยู่ในสภาวะอารมณ์ และบ่อยครั้งที่พบว่ามีแรงจูงใจทางการศึกษาต่ำและมีผลสัมฤทธิ์ทางการเรียนต่ำ อีกด้านหนึ่ง จำนวนสมาชิกที่เพิ่มขึ้นอาจจะลดการเข้าถึงแหล่งทรัยากรของเด็ก ที่ให้โอกาสในการเรียนน้อยลงและผลสัมฤทธิ์ของนักเรียนต่ำลง และนักเรียนบางคนอาจจะได้รับผลประโยชน์จากปู่ย่าตายายในด้านกายภาพ ข้อมูลข่าวสาร สังคม และแห่งทางด้านอารมณ์ และอาจจะแสดงถึงผลสำเร็จสูงขึ้น อีกด้านตรงข้ามนักเรียนที่อยู่กับปู่ย่าตายายที่ยากจนหรือป่วยจะมีข้อจำกัดในแหล่งทรัพยากรของครอบครัว

จากการศึกษาของ 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 และองค์ประกอบของชนชั้น

ทางสังคมและเศรษฐกิจและรายได้ทางวัฒนธรรม แนวทางที่วัด SES จะเป็นองค์ประกอบของรายได้ของครอบครัว อาชีพของผู้ปกครอง ระดับการศึกษาของผู้ปกครอง โดยเฉพาะอย่างยิ่งอาชีพของแม่ ดัชนีบางอย่างประกอบด้วยการวัดรายได้ทางวัฒนธรรมโดยเฉพาะเกี่ยวกับระดับของสิ่งแวดล้อมของครอบครัวที่เป็นองค์ประกอบของวัฒนธรรมชั้นสูง องค์ประกอบบางอย่างได้จากจำนวนหนังสือหรือเครื่องดนตรีในบ้าน 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 จึงมีความสำคัญในการพิจารณาตัวแปรพื้นฐานของนักเรียนด้วยเพื่อปรับค่าเฉลี่ยภายในสถานศึกษาเพื่อแสดงองค์ของตัวแปรที่มีผลต่อผลสัมฤทธิ์ 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 เช่น ผลสัมฤทธิ์เดิม หรือพื้นฐานความรู้เดิมของนักเรียน

นอกจากตัวแปรที่เป็นคุณลักษณะพื้นฐานของนักเรียนและพื้นฐานครอบครัวของนักเรียนที่ส่งผลต่อผลสัมฤทธิ์ของนักเรียน ยังพบตัวแปรแรงจูงใจและเจตคติของนักเรียน จากทฤษฎีสติปัญญาทางสังคม แรงจูงใจทางการเรียนของนักเรียน เชื่อว่ามีสามองค์ประกอบคือ คุณค่า (Value) ความคาดหวัง(expectancy) และความชื่นชอบ (affect) (Bandura, 1989) นักเรียนที่เห็นคุณค่าในการทำกิจกรรมจะมีลักษณะของความพยายามมากกว่าและได้รับความสำเร็จสูง ความคาดหวังเชื่อว่านักเรียนจะมีความเชื่อในประสิทธิภาพของตนเอง (Self – efficacy) และความเชื่อในความสามารถของตนเอง (Self – concept) นอกจากนี้ ความคาดหวังจะถูกพิจารณาว่าเป็น

ความสามารถทางการศึกษาที่สามารถมีผลกับประสบการณ์ของนักเรียน โดยเฉพาะอย่างยิ่งความสำเร็จที่ผ่านและความล้มเหลวในการทำชิ้นงานในการเรียน (Pintrich & Schunk, 2002) เนื่องจากความคาดหวังทั้งปฏิกิริยาทางอารมณ์ต่อชิ้นงานและการปฏิบัติชิ้นงานมีอิทธิพลต่อความสามารถ นักเรียนที่มีผลทางบวกเมื่อทำงานในชิ้นงานที่เรียนมีแนวโน้มที่จะมีความพยายามในการทำ และส่งผลต่อผลสัมฤทธิ์ทางการเรียนสูงขึ้น

อย่างไรก็ตาม Raudenbush and Willims (1996 cited in Sheerens and Bosker, 1997) ได้เสนอข้อพึงระวังเกี่ยวกับการกำหนดตัวแปรเพื่อการประมาณค่าผลสัมฤทธิ์ที่แสดงถึงผลการจัดการศึกษาของสถานศึกษา ไว้ดังนี้ 1) ในบางกรณี ผลลัพธ์ที่ได้อาจมีค่าต่ำกว่าผลที่เป็นจริงของสถานศึกษา เช่นการที่นักเรียนที่มีข้อได้เปรียบทางความสามารถ หรือเศรษฐานะมักจะเลือกสถานศึกษาที่มีการดำเนินการดี ดังนั้นการควบคุมสภาพเศรษฐานะจึงทำให้เกิดผลการปรับแก้ที่เกินจริงกับนักเรียนกลุ่มนี้ 2) ระบบการชดเชย (compensatory system) ประสิทธิผลของสถานศึกษา เช่นการชดเชยเพื่อแก้ไขข้อบกพร่องด้วยการจัดการศึกษาที่ดีขึ้นกว่าปกติ ซึ่งอาจจะมีผลให้ผลที่ได้มีค่าลดลง ดังนั้นผลที่เกิดขึ้นจะสามารถสังเกตได้เมื่อวิเคราะห์ผลผลิตที่มีการพยากรณ์ด้วยตัวแปรร่วมที่เกี่ยวข้อง 3) ข้อจำกัดของการกำหนดตัวแปรที่เกี่ยวข้องกับภูมิหลังของนักเรียนได้อย่างครบถ้วน นอกจากนี้ข้อมูลที่ได้จากการวัดที่ขาดความถูกต้อง น่าเชื่อถืออาจมีผลต่อการประมาณค่าในระดับบุคคล

5.2 งานวิจัยที่ศึกษาเกี่ยวกับตัวแปรที่มีอิทธิพลต่อคุณภาพการจัดการศึกษาและการใช้โมเดลมูลค่าเพิ่มเพื่อประเมินคุณภาพการจัดการศึกษา

งานวิจัยที่ศึกษาเกี่ยวกับตัวแปรหรือปัจจัยที่ส่งผลต่อคุณภาพการจัดการศึกษา และผลสัมฤทธิ์ทางการเรียนของนักเรียน พบว่าโดยส่วนใหญ่เป็นการศึกษาเกี่ยวกับอิทธิพลของสถานศึกษา ชุมชนและสังคม ข้อมูลพื้นฐานทางครอบครัว และพื้นฐานของนักเรียน และได้เริ่มมีนักการศึกษาคนสนใจศึกษามากขึ้น เนื่องจากสารสนเทศที่ได้ จะนำมาใช้กำหนดเป็นหลักการหรือนโนบายสู่การพัฒนาการศึกษาให้ทัดเทียมกับระดับนานาชาติ โดยใช้เทคนิควิเคราะห์พหุระดับรวมทั้งการวิเคราะห์มูลค่าเพิ่มเพื่อตรวจสอบประสิทธิผลของสถานศึกษา เพื่อบ่งบอกถึงคุณภาพการจัดการศึกษา เมื่อควบคุมตัวแปรในระดับต่างๆ ดังตัวอย่างงานวิจัยที่เสนอดังนี้

Chiu & Xihua (2008) ได้ศึกษาอิทธิพลทางครอบครัวและแรงจูงใจต่อผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์จากการวิเคราะห์ข้อมูลนักเรียน 41 ประเทศ โดยการศึกษาตัวแปรสามกลุ่มคือ ตัวแปรระดับประเทศ ประกอบด้วย GDP , GDP Gini, และคุณค่าทางวัฒนธรรม ตัวแปรครอบครัว ประกอบด้วย การย้ายถิ่นเข้ามาอยู่เป็นกลุ่มแรก การย้ายถิ่นเข้ามาอยู่เป็นอันดับที่สอง ภาษาที่ใช้พูดคุยในบ้าน เศรษฐานะทางสังคม มีพ่อหรือเพียงคนเดียว มีผู้ปกครองผสม ไม่ได้อยู่กับพ่อแม่ อยู่อาศัยกับปู่ย่าตายายเพียงคนเดียว จำนวนลูกพี่ลูกน้อง และลำดับการเกิด คุณลักษณะอื่นของครอบครัวประกอบด้วย จำนวนหนังสือที่มีอยู่ในบ้าน ความเป็นเจ้าของวัฒนธรรม และรูปแบบการติดต่อสื่อสารทางวัฒนธรรม และกลุ่มที่สามเป็นตัวแปรของนักเรียน ประกอบด้วย เพศ ตัวแปร

แรงจูงใจ ประกอบด้วย ความสนใจในวิชาคณิตศาสตร์ แรงจูงในการใช้เป็นเครื่องมือ ความพากเพียรพยายาม ความเชื่อในประสิทธิภาพของตนเอง ความเชื่อในความสามารถของตนเอง และปฏิสัมพันธ์ระหว่างตัวแปร

ผลจากการศึกษาพบว่าความแปรปรวนของคะแนนสอบสอบคณิตศาสตร์ของนักเรียนมีค่าเกือบครึ่งของความแปรปรวนทั้งหมด (44%) และความแปรปรวนระดับสถานศึกษาเท่ากับ 25% และที่เหลือเป็นความแปรปรวนระดับประเทศเท่ากับ 31% ผลการศึกษาในตัวแปรต่างๆ พบว่าการย้ายเข้าเป็นกลุ่มแรกและการย้ายเข้ามาเป็นกลุ่มที่สองโดยเฉลี่ยมีคะแนนคณิตศาสตร์ต่ำกว่านักเรียนที่เกิดในชาตินั้นอยู่ 12 และ 5 คะแนน ตามลำดับ นักเรียนที่พูดภาษาในบ้านแตกต่างจากที่สถานศึกษาโดยเฉลี่ยมีคะแนนคณิตศาสตร์ต่ำกว่านักเรียนกลุ่มอยู่ 11 คะแนน โดยเฉลี่ยนักเรียนมีคะแนนขึ้น 5 คะแนนในครอบครัวที่มี SES สูงขึ้น 10% การลงทุนและการมีส่วนเกี่ยวข้องของครอบครัวมีผลต่อคะแนนสอบคณิตศาสตร์ นักเรียนที่มีจำนวนหนังสือมากมีแนวโน้มที่จะมีคะแนนสูง โดยเฉลี่ยนักเรียนจะมีคะแนนสูงขึ้น 1 คะแนนในการเพิ่มขึ้น 10% ของการเป็นเจ้าของทางวัฒนธรรม และเมื่อควบคุมจำนวนหนังสือที่บ้าน สัมประสิทธิ์ถดถอยของการย้ายมาเป็นกลุ่มแรกจะลดลง 34% สัมประสิทธิ์ถดถอยของการย้ายมาอยู่เป็นกลุ่มที่สองจะลดลง 26% เศรษฐานะของครอบครัว(SES) จะลดลง 32% การอยู่กับพ่อหรือแม่อย่างเดียวจะลดลง 29% และการอยู่กับพ่อแม่ผสมจะลดลง 26% ส่วนผลของตัวแปรระดับนักเรียนพบว่า นักเรียนชายมีคะแนนสูงกว่านักเรียนหญิง 18 คะแนน เมื่อควบคุมตัวแปรเพศ จะลดสัมประสิทธิ์ถดถอยของการอยู่ผู้ปกครองผสม 25 โดยเฉลี่ยนักเรียนจะมีคะแนนสูงขึ้น 2 คะแนนในคณิตศาสตร์ที่การเพิ่มขึ้น 10% ของการสนใจในคณิตศาสตร์ อิทธิพลของผู้ปกครองผสมต่อคะแนนสอบคณิตศาสตร์จะลดลง 21% เมื่อควบคุมตัวแปรความสนใจต่อคณิตศาสตร์ โดยเฉลี่ยนักเรียนมีคะแนนเพิ่มขึ้น 1 คะแนนในการเพิ่มขึ้น 10 % ของความพากเพียรพยายาม โดยเฉลี่ยนักเรียนมีคะแนนสูงขึ้น 1 คะแนนและ 3 คะแนน คณิตศาสตร์ในการเพิ่มขึ้น 10% ของความเชื่อในตนเองและความเชื่อในมโนทัศน์ตนเองตามลำดับ เมื่อควบคุมความเชื่อในมโนทัศน์ตนเอง อิทธิพลของความสนใจวิชาคณิตศาสตร์จะลดลง 31% และอิทธิพลของผู้ปกครองผสมไม่มีนัยสำคัญ

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) จากผู้ปกครองเป็นตัวทำนายแรงจูงใจ และผลสัมฤทธิ์ในวิชาคณิตศาสตร์และวิทยาศาสตร์ในทางลบ โดยเฉพาะในเพศหญิง การได้รับการ

สนับสนุนจากผู้ปกครอง (parental psychological support) พบว่า มีอิทธิพลทางตรงต่อแรงจูงใจของนักเรียน และมีผลทางอ้อมต่อผลสัมฤทธิ์ของนักเรียน และพบว่าเศรษฐานะทางสังคม (socioeconomic) มีความสำคัญมาก เนื่องจากมีอิทธิพลทางตรงต่อความต้องการทางด้านการศึกษา (educational aspirations) แรงกดดัน การส่งเสริมทางจิตใจ และมีอิทธิพลทางอ้อมต่อความรู้เดิม ความต้องการทางด้านการศึกษา เจตคติต่อสถานศึกษา มโนภาพของตนเองเกี่ยวกับวิชาคณิตศาสตร์และวิทยาศาสตร์ และผลสัมฤทธิ์ในวิชาคณิตศาสตร์และวิทยาศาสตร์

Willms and Somers (2001) ได้ศึกษากับนักเรียนของสถานศึกษาในกลุ่มประเทศละตินอเมริกาจำนวน 13 ประเทศ ๆ ละ 100 โรง พบว่า ปัจจัยตัวแปรระดับบุคคลด้วยสถานะทางวัฒนธรรมทางสังคม (sociocultural status) ที่ส่งผลทางบวกอย่างมีนัยสำคัญต่อผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ภาษา(สเปน) และจำนวนปีในการศึกษาของนักเรียนในหลายประเทศได้แก่ การศึกษาอย่างต่อเนื่องของผู้ปกครอง และความเอาใจใส่ในแต่ละวัน (daycare) ซึ่งตัวแปรหลังนี้ยังมีผลทางบวกต่อการลดอัตราการตกซ้ำชั้นของนักเรียนและจำนวนปีในการศึกษาอย่างมีนัยสำคัญ และตัวแปรที่ส่งผลเชิงลบคือ การช่วยเหลือด้านการบ้าน สอดคล้องกับผลการศึกษาของ Ho and Willms (1996 cited in Willms and Somers,2001) ผลการศึกษายังแสดงให้เห็นถึงผลจากปัจจัยตัวแปรระดับห้องที่ส่งผลทางบวกต่อผลสัมฤทธิ์ ได้แก่ การไม่แบ่งกลุ่ม / ชั้นตามระดับความสามารถของนักเรียน มีการทดสอบอย่างปกติสม่ำเสมอ บรรยากาศในชั้นเรียน การมีส่วนร่วมที่ดีของผู้ปกครอง พบว่า มีอิทธิพลค่อนข้างสูงเมื่อพิจารณาในระดับสถานศึกษา 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 นอกจากนี้ชี้ให้เห็นว่านักเรียนชุดเดียวกันจะมีความแปรปรวนระหว่างสถานศึกษาของผลการเรียนรู้ โดยมีความแปรปรวนระหว่างสถานศึกษาเป็น 10% ถึง 20% และพิสัยของค่าเศษเหลือในคะแนนรวมการอ่านเป็น –9.9 ถึง 15.5 คะแนนคณิตศาสตร์เป็น –27.6 ถึง 33.7 และในคะแนนรวมการอ่านมีพิสัยเป็น –11.8 ถึง 11.0 ซึ่งพิสัยของค่าเศษเหลือนี้จะบ่งชี้ได้ว่ามีความผันแปรอย่างมากในวิชาคณิตศาสตร์ในสถานศึกษามากกว่าค่าเศษเหลือในวิชาการอ่านและภาษาในสถานศึกษา ความผันแปรระหว่างสถานศึกษาต่ำกว่าคะแนนที่คาดหวังในคะแนนการอ่านเป็น 57% ในคะแนนคณิตศาสตร์เป็น 51% และในคะแนนภาษาเป็น 52.5% ชี้แนะได้ว่ามีสถานศึกษาหลายสถานศึกษาในกลุ่มตัวอย่างที่มีคะแนนต่ำกว่าคะแนนที่คาดหวังในองค์ประกอบของนักเรียน โดยเฉพาะการอ่าน

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) และนักเรียนที่ย้ายเข้ามาอยู่ ใช้การวิเคราะห์ HLM เพื่อหาความผันแปรของอิทธิพลสถานศึกษาในทักษะการอ่าน ด้วยการควบคุมคุณลักษณะของนักเรียน ในโมเดล 5 โมเดลที่แตกต่างกันพบผลที่ได้ดังนี้ 1) โมเดล 1 เป็น null model ซึ่งเป็นโมเดลที่ไม่รวมตัวแปรที่ใช้ในการอธิบายทั้งในระดับบนักเรียนและระดับสถานศึกษา พบว่าแหล่งความผันแปรในระดับนักเรียน เป็น 82% และความผันแปรระหว่างสถานศึกษา มี 18% ของ 2) โมเดล 2 เป็นโมเดลที่ตรวจสอบความแตกต่างในคะแนนการอ่านระหว่าง first generation และ immigrants ที่แปรค่าในระดับนักเรียนและระดับสถานศึกษา พบว่าความแตกต่างระหว่าง first generation และชาวแคนาดาไม่มีนัยสำคัญทางสถิติที่แตกต่างจากศูนย์ในระดับนักเรียน แต่มีความผันแปรในระหว่างสถานศึกษาอย่างมีนัยสำคัญ 3) โมเดล 3 เป็นโมเดลที่นำคุณลักษณะของบุคคล 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 และการใช้โมเดลสองระดับเพื่ออธิบายความสามารถในการอ่านกับดัชนีเศรษฐานะ (ESCS) ภายในประเทศในปี 2000

กับ 2003 พบว่ามีอิทธิพลของ ESCS ในระดับนักเรียนและระดับสถานศึกษาและสามารถใช้อธิบายคุณลักษณะของระบบทางการศึกษา และเมื่อเปรียบเทียบผลในปี 2003 และปี 2000 พบว่ามีอิทธิพลคล้ายกันในหลายประเทศ

Willms (2004) ได้ศึกษาผลสัมฤทธิ์ในการอ่านของประเทศแคนาดาและประเทศสหรัฐอเมริกา เพื่อตรวจสอบการแจกแจงคะแนนในระดับนักเรียนและสถานศึกษา ประมาณค่าเศรษฐานะของครอบครัวที่เกี่ยวข้องกับความสามารถในการอ่านในแคนาดาและสหรัฐอเมริกาและการตรวจสอบความสัมพันธ์ระหว่างความสามารถในการอ่านและเศรษฐานะของครอบครัว ภายในและระหว่างสถานศึกษา 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 ส่วนสัมประสิทธิ์ปฏิสัมพันธ์พบว่าไม่มีนัยสำคัญ ดังนั้นขนาดของความแตกต่างของเพศมีลักษณะเหมือนกันในทั้งสองประเทศ ค่าประมาณของค่าเฉลี่ยของความชันของ SES เป็น 33.5 ความแตกต่างระหว่างสถานที่ที่ของนักเรียนที่เกิดในประเทศและเกิดต่างประเทศมีความแตกต่าง 5.9 3) โมเดล 3 เป็นโมเดลที่รวมชุดตัวแปรค่าเฉลี่ยของสถานศึกษา มีการประมาณค่าของอิทธิพลของบริบทของสองประเทศ การประมาณค่าของสหรัฐอเมริกาเป็น 63.8 แสดงว่านักเรียนที่ค่าเฉลี่ยในคุณลักษณะที่สามารถแสดงได้ 63.8 คะแนนและในแคนาดามีค่าประมาณเป็น 18 คะแนน ผลแสดงให้เห็นว่าทั้งสองประเทศ ได้รับประโยชน์เมื่อสถานศึกษามี SES สูงถึงแม้ว่านักเรียนแต่ละคนจะมีพื้นฐานครอบครัวที่ต่างกัน การควบคุมค่าเฉลี่ย SES ของสถานศึกษาเป็นการลดการประมาณความแตกต่างของแคนาดาและสหรัฐอเมริกาได้มากกว่า 3 คะแนน

Patrick and Tomasz (2004) ได้ศึกษาผลกระทบของของการใช้คอมพิวเตอร์ต่อผลสัมฤทธิ์ด้านการอ่านของนักเรียนอายุ 15 ปี จุดประสงค์ของการศึกษาครั้งนี้เป็นการศึกษาผลกระทบของเทคโนโลยีการสื่อสารและสารสนเทศ (ICT) ต่อความสามารถในการอ่านของนักเรียนอายุ 15 ปีของชาวแคนาดา โดยการใช้ข้อมูลจากโครงการประเมินผล PISA 2000 โดยพิจารณาความสนใจในคอมพิวเตอร์ โดยจัดลักษณะของตัวแปรเป็น 4 ตัว จะแบ่งออกตามจังหวัดและเพศ ได้แก่ การเข้าถึงคอมพิวเตอร์และอินเตอร์เน็ต ความถี่ของการใช้คอมพิวเตอร์ รูปแบบการใช้ และทัศนคติต่อ ICT จากผลการวิเคราะห์พหุระดับ เพื่อควบคุมตัวแปร ใน 5 โมเดลที่มีทั้งตัวแปรระดับบุคลและ

ระดับครอบครัว พบว่าความสามารถของการใช้คอมพิวเตอร์มีความสัมพันธ์สูงที่สุดในการศึกษาครั้งนี้ และช่วยลดช่องว่างของเพศ เพื่อพิจารณาที่ระดับบุคคล ครอบครัว จังหวัด พบว่าในโมเดลพื้นฐานอธิบายคะแนนการอ่านได้ 32.4% และเมื่อเพิ่มตัวแปรการใช้ ICT จะอธิบายได้เพิ่มเป็น 37.3% และตัวแปรที่เกี่ยวกับการใช้คอมพิวเตอร์พบว่ามีนัยสำคัญ การใช้คอมพิวเตอร์ที่สถานศึกษามีความสัมพันธ์ทางบวก แต่มีค่าไม่สูงเท่าการใช้ที่บ้าน และความสัมพันธ์เป็บค่าลบกับการใช้คอมพิวเตอร์ที่ห้องสมุดและที่อื่นๆ สัมประสิทธิ์ของดัชนีของความสนใจในคอมพิวเตอร์มีค่าเป็นลบและมีนัยสำคัญ และดัชนีความสามารถการใช้คอมพิวเตอร์มีค่าเป็นบวกสูงมากกับคะแนนการอ่าน ตัวแปรเพศมีปฏิสัมพันธ์ทางลบ ผลที่ได้ชี้ให้เห็นว่าความสามารถการใช้คอมพิวเตอร์สูงมีประโยชน์สำหรับชายมากกว่าหญิง และความถี่ของการใช้คอมพิวเตอร์สำหรับการเล่นเกมส์และการเข้าอินเตอร์เน็ตไม่มีนัยสำคัญกับทักษะการอ่าน

Opdenakker and Van Damme (2000) ศึกษาอิทธิพลที่เกิดจากสถานศึกษา การสอนของครู และบรรยากาศของชั้นเรียนที่มีต่อผลสัมฤทธิ์ทางการเรียนซึ่งวัดจากแบบสอบที่ร่วมกับคณะครูผู้สอนในสถานศึกษาเป็นผู้สร้างเพื่อให้มีความตรงตามเนื้อหาของหลักสูตรพื้นฐานของแต่ละสถานศึกษา และต่อความรู้สึกที่มีต่อสถานศึกษา กลุ่มตัวอย่างที่ใช้ในการศึกษาคือนักเรียนชั้นมัธยมศึกษาปีที่ 1 จำนวนสถานศึกษา 52 แห่ง ผลการศึกษาพบว่า ปัจจัยด้านบรรยากาศของชั้นเรียนและสถานศึกษาส่งผลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียนมากกว่าความรู้สึกต่อสถานศึกษา ในขณะที่บรรยากาศที่เป็นวิชาการ เช่น ความร่วมมือของครูในการพัฒนาการสอน การประชุม/อภิปรายทางวิชาการอย่างสม่ำเสมอ ส่งผลที่ดีต่อนักเรียน 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 การทดสอบโมเดลการจัดการศึกษาของสถานศึกษาทั้งสามรูปแบบด้วยโมเดล HLM เมื่อควบคุมเศรษฐานะของผู้ปกครอง พบว่าตัวแปรความเป็นผู้นำทางวิชาการ การสนับสนุนของผู้ปกครอง ทำให้ครูจัดการเรียนการสอนได้อย่างมีประสิทธิภาพ และนักเรียนมีผลสัมฤทธิ์ทางการเรียนสูงขึ้น

Silin and Murray-Harvey (1999) ได้ศึกษาปัจจัยที่เกี่ยวข้องกับผลการเรียนของนักเรียน ซึ่งศึกษากับนักเรียนระดับประถมศึกษา จำนวน 30 แห่ง จำแนกตามสังกัด และที่ตั้งของสถานศึกษา ตัวแปรระดับนักเรียนได้แก่ เศรษฐานะ ความเชื่อในมโนทัศน์ของตนเอง (self-concept) เจตคติและวิธีการเรียนรู้ ผลการศึกษาพว่า ตัวแปรโดยรวมสามารถอธิบายความแตกต่างของ

ผลสัมฤทธิ์ระหว่างสถานศึกษาค่อนข้างสูง และพบว่าตัวแปรด้านเศรษฐานะที่ส่งผลโดยตรงต่อผลสัมฤทธิ์ของสถานศึกษาค่อนข้างสูง และส่งผลทางอ้อมผ่านตัวแปรด้านสังกัด และขนาดของสถานศึกษาในระดับปานกลาง และส่งผ่านจากสังกัดไปสู่ขนาดสถานศึกษาระดับเล็กน้อย แสดงให้เห็นว่านักเรียนที่มีเศรษฐานะดีที่ไม่เลือกเรียนสถานศึกษาเอกชนก็จะเลือกเรียนในสถานศึกษารัฐบาลขนาดใหญ่

Hofman, Hofman and Guldemon (1999) ได้ศึกษาปัจจัยทางด้านบริบทที่ส่งผลต่อการเรียนรู้ของนักรียน โดยศึกษากับนักเรียนระดับ 8 จากสถานศึกษา 103 แห่ง พบว่านโยบายการบริหารงานที่แตกต่างกันตามสังกัดของสถานศึกษาส่งผลต่อผลสัมฤทธิ์ทางวิชาการของนักเรียนในวิชาคณิตศาสตร์มากกว่าที่ไม่ใช่ด้านวิชาการ และพบว่านักเรียนในสังกัดสถานศึกษารัฐบาลมีความรู้สึกที่ดีต่อสถานศึกษามากกว่านักเรียนในสถานศึกษาเอกชน ในขณะที่ค่าเฉลี่ยผลสัมฤทธิ์ทางการเรียนต่ำกว่า และให้ข้อคิดเห็นว่าอาจเกิดจากระบบการคัดเลือกเข้าศึกษาที่ต่างกัน เนื่องจากสถานศึกษาเอกชนมีเกณฑ์การคัดเลือกที่เข้มงวดตามคุณสมบัติเบื้องต้นที่สถานศึกษาต้องการ ซึ่งตรงข้ามกับสถานศึกษารัฐบาลที่นักเรียนเลือกเข้าเรียนตามความสมัครใจ 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

บุญเรือง ศรีเหรัญ (2542) ทำการศึกษานักเรียนชั้นมัธยมศึกษาปีที่ 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 ประสิทธิภาพการจัดการเรียนการสอนและคุณภาพของการใช้หลักสูตรส่งผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์

และวิทยาศาสตร์ เจตคติและแรงจูงใจต่อการเรียน และที่ส่งผลต่อความเป็นผู้นำ ประกอบด้วย ปัจจัย/ตัวแปรระดับนักเรียน ได้แก่ บรรยากาศและสภาพแวดล้อมในครอบครัว 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 ปัจจัยที่ส่งผลทางตรงเชิงบวกอย่างมีนัยสำคัญในระดับสูง ได้แก่ ภูมิหลังของนักเรียน รองลงมาคือ พฤติกรรมการสอนและภูมิหลังของครู และปัจจัยระดับสถานศึกษา สำหรับปัจจัยที่ส่งผลทางอ้อม ได้แก่ ภูมิหลังครูที่มีขนาดของผลที่สูงขึ้นเมื่อส่งผ่านตัวแปรพฤติกรรมการสอนของครู

ปิยะธิดา ทองอร่าม (2545) ได้ศึกษาปัจจัยที่มีอิทธิพลต่อคุณภาพของสถานศึกษา สังกัดกรมสามัญศึกษา ด้วยการวิเคราะห์พหุระดับด้วยโมเดลเชิงเส้นตรงระดับลดหลั่น 2 ระดับ คือระดับสถานศึกษา และระดับจังหวัด ผลการศึกษาพบว่า ปัจจัยระดับสถานศึกษาที่มีอิทธิพลต่อคุณภาพของสถานศึกษาอย่างมีนัยสำคัญทางสถิติ ได้แก่ สัดส่วนของนักเรียนต่อครู สัดส่วนนักเรียนต่อห้องเรียน ขวัญและกำลังใจในการปฏิบัติงาน และความเป็นผู้นำของผู้บริหารสถานศึกษา ส่วนปัจจัยระดับจังหวัดที่มีอิทธิพลต่อคุณภาพของสถานศึกษาอย่างมีนัยสำคัญทางสถิติ ได้แก่ ค่าเฉลี่ยปริมาณการให้คำแนะนำทางวิชาการจากหน่วยศึกษาอย่างมีนัยสำคัญทางสถิติ ค่าเฉลี่ยงบประมาณที่ได้รับการสนับสนุนจากจังหวัด

สุบิน ยุระรัช (2547) ได้ศึกษาโมเดลสามระดับของข้อมูลตัวแปรที่ส่งผลต่อคุณภาพนักเรียนประถมศึกษาในกรุงเทพมหานคร พบว่าตัวแปรทำนายระดับนักเรียนที่ส่งผลต่อคุณภาพนักเรียนประถมศึกษาอย่างมีนัยสำคัญทางสถิติ ได้แก่ ความคาดหวังของนักเรียนในการศึกษาต่อ ความคิดเห็นของนักเรียนที่มีต่อการเรียน พฤติกรรมการเรียนของนักเรียน แรงจูงใจใฝ่สัมฤทธิ์ของนักเรียน และพฤติกรรมของบิดามารดาในการส่งเสริมการเรียนของบุตร ตัวแปรระดับห้องเรียนพบว่าไม่มีตัวแปรใดที่ส่งผลต่อค่าเฉลี่ยคุณภาพของคุณภาพของนักเรียน และในระดับสถานศึกษาไม่พบความผันแปร

รัชนก บุญปู่ (2547) ได้ศึกษาปัจจัยที่มีอิทธิพลต่อผลการเรียนรู้ตามมาตรฐานหลักสูตรการศึกษาขั้นพื้นฐานในสถานศึกษานำร่องหลักสูตรสถานศึกษา ด้วยการวิเคราะห์พหุระดับ 2 ระดับ ผลการวิจัยพบว่า ตัวแปรระดับนักเรียนที่มีอิทธิพลต่อผลการเรียนรู้ตามมาตรฐานหลักสูตรการศึกษาขั้นพื้นฐาน อย่างมีนัยสำคัญทางสถิติ ได้แก่ เจตคติต่อการเรียน พฤติกรรมการเรียน ความคาดหวังในการศึกษาต่อ โดยชุดของตัวแปรระดับนักเรียนสามารถอธิบายความแปรปรวนของผลการเรียนรู้ได้ร้อยละ 17.4 ส่วนตัวแปรระดับสถานศึกษาที่ส่งผลต่อคุณภาพของการจัดการศึกษา

ตามหลักสูตรสถานศึกษาขั้นพื้นฐาน อย่างมีนัยสำคัญทางสถิติ ได้แก่ คุณภาพการสอน ระดับการศึกษาของผู้บริหาร โดยตัวแปรระดับสถานศึกษาสามารถอธิบายความแปรปรวนของผลการเรียนรู้ได้ร้อยละ 15.5

ศุภลักษณ์ ใจแสวงทรัพย์ (2547) ได้ศึกษาปัจจัยที่ส่งผลต่อคะแนนพัฒนาการวิชาคณิตศาสตร์ของนักเรียนในระดับมัธยมศึกษาปีที่ 3 จากผลการศึกษาพบว่า พัฒนาการทางการเรียนคณิตศาสตร์ได้รับอิทธิพลทางตรงจากฐานะทางเศรษฐกิจของผู้ปกครอง อันดับรองลงมาได้แก่ ผลสัมฤทธิ์ทางการเรียนเดิม การศึกษาของครูผู้สอน และเจตคติต่อวิชาคณิตศาสตร์

พิชิต ธรรมรักษ์ (2549) ได้ศึกษาปัจจัยที่มีผลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียนแผนการเรียนศิลป์ภาษา ในกรุงเทพมหานคร ผลการวิจัยพบว่าปัจจัยที่สำคัญที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียน ได้แก่ ประสบการณ์สอน ขนาดของสถานศึกษา ความรู้พื้นฐานเดิม อาชีพของผู้ปกครอง ความเป็นผู้นำทางวิชาการของผู้บริหาร เจตคติต่อวิชาคณิตศาสตร์ และตัวแปรทั้ง 6 ตัวแปรสามารถทำนายผลสัมฤทธิ์ทางการเรียนได้ร้อยละ 39.10

สุชีรา มะหิเมือง (2547) ได้ศึกษาปัจจัยที่ส่งผลต่อผลสัมฤทธิ์และพัฒนาการทางวิชาการด้วยวิธีการวิเคราะห์มูลค่าเพิ่ม กับกลุ่มตัวอย่างสถานศึกษาประถมศึกษา จำนวน 123 โรง ข้อมูลที่ใช้ประกอบด้วยข้อมูลทุติยภูมิของคะแนนผลสัมฤทธิ์วิชาภาษาไทย และวิชาคณิตศาสตร์ ซึ่งได้จากผลการทดสอบแห่งชาติของนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2545 และผลการทดสอบของโครงการประเมินคุณภาพและวินิจฉัยข้อบกพร่องทางการเรียนของกลุ่มตัวอย่างขณะศึกษาในชั้นประถมศึกษาปีที่ 4 และข้อมูลปฐมภูมิเกี่ยวกับภูมิหลังของกลุ่มตัวอย่าง บริบททั่วไปของสถานศึกษา และการปฏิบัติงานทางวิชาการตามตัวบ่งชี้ 6 ด้านได้จากการสำรวจด้วยแบบสอบถาม โดยใช้การวิเคราะห์โมเดลเชิงเส้นระดับลดหลั่น (HLM) และการวิเคราะห์ความสัมพันธ์โครงสร้างเชิงสาเหตุ (LISREL) ผลการศึกษาพบว่า โมเดลการวิเคราะห์มูลค่าเพิ่มของค่าเฉลี่ยผลสัมฤทธิ์และพัฒนาการทางวิชาการที่ใช้วิธีวิเคราะห์ถดถอยแบบพหุระดับด้วยโปรแกรม HLM ซึ่งกำหนดให้ตัวแปรภูมิหลังของนักเรียนด้านผลสัมฤทธิ์เดิมเป็นตัวแปรที่ส่งผลเชิงสุ่ม ในขณะที่ควบคุมอิทธิพลจากตัวแปรด้านเศรษฐานะ และพื้นฐานทางภาษา เป็นโมเดลที่มีประสิทธิภาพในการพยากรณ์(R^2) ในวิชาภาษาไทยเป็น 22.10 % และ ในวิชาคณิตศาสตร์เป็น 24.97 % นอกจากนี้สถานศึกษาที่มีมูลค่าเพิ่มฯจัดอยู่ในกลุ่มสูงเป็นสถานศึกษาที่มีบริบททั่วไปของสถานศึกษาและสภาพการปฏิบัติงานทางวิชาการดีกว่ากลุ่มสถานศึกษาที่มีมูลค่าเพิ่มฯในระดับต่ำ โมเดลเชิงสาเหตุแสดงอิทธิพลของปัจจัย/ตัวแปรระดับสถานศึกษาที่มีต่อมูลค่าเพิ่มเป็นโมเดลที่มีความตรงและความกลมกลืนกับข้อมูลเชิงประจักษ์ ตัวแปรที่มีอิทธิพลอย่างมีนัยสำคัญ ได้แก่ การปฏิบัติงานทางวิชาการที่ส่งผลเชิงบวกต่อมูลค่าเพิ่มของค่าเฉลี่ยผลสัมฤทธิ์ และเชิงลบต่อมูลค่าเพิ่มของพัฒนาการทางวิชาการ ทำเลที่ตั้งของสถานศึกษาที่ส่งผลเชิงลบต่อมูลค่าเพิ่มของพัฒนาการทางวิชาการวิชาคณิตศาสตร์ และการส่งผลเชิงบวกของมูลค่าเพิ่มของค่าเฉลี่ยผลสัมฤทธิ์ต่อมูลค่าเพิ่มของพัฒนาการทางวิชาการ ที่ระดับนัยสำคัญ 0.01

ชูเวช ชาญสง่าเวช และสิงหา เจียมศิริ (2546) ได้ทำการศึกษาการวัดมูลค่าเพิ่มจากการดำเนินงานของสถาบันอุดมศึกษา โดยมีวัตถุประสงค์ที่จะพัฒนาระบบการวัดมูลค่าเพิ่มจากการดำเนินงานของสถาบันอุดมศึกษา เพื่อใช้ในกรณีศึกษาโครงการนำร่องเพื่อทดลองจัดระดับสถาบันอุดมศึกษาไทยในสาขาวิชาครุศาสตร์/ศึกษาศาสตร์ การพัฒนาเกณฑ์ต่างๆในระบบนั้น ยึดหลักในการกำหนดชุดของเกณฑ์ซึ่งเป็นที่ยอมรับกันอย่างกว้างขวางอยู่ในแวดวงวิชาการ มูลค่าเพิ่มจากการดำเนินงานของสถาบันอุดมศึกษาในการวิจัยนี้ได้พัฒนาขึ้นเป็นเกณฑ์หนึ่งซึ่งอยู่ในสามมุมมองใหญ่ๆตามความต้องการใช้ของผู้มีส่วนได้เสียกับการดำเนินงานของสถาบันอุดมศึกษา คือ มุมมองของนักเรียนที่ประสงค์จะเข้าศึกษาในสถาบันอุดมศึกษาและผู้ปกครอง ตลอดจนศิษย์เก่าของสถาบัน ผู้ใช้บัณฑิต และคณาจารย์ประจำสถาบันอุดมศึกษานั้นๆ ผลจากการประยุกต์ในการทดลองจัดระดับสถาบันอุดมศึกษาไทย สาขาวิชาครุศาสตร์/ศึกษาศาสตร์พบว่า เกณฑ์และตัวชี้วัดที่พัฒนาขึ้นสามารถใช้ประเมินความเห็นของศิษย์เก่าผู้ใช้บัณฑิตและคณาจารย์เกี่ยวกับมูลค่าเพิ่มที่ได้รับ และสามารถนำมาใช้ในการจัดระดับสถาบันอุดมศึกษาได้เป็นอย่างดี

จากการศึกษาเอกสารเกี่ยวข้องตัวแปรทางการศึกษาที่ส่งผลต่อคุณภาพการจัดการศึกษาสามารถสรุปตัวแปรทำนายที่มีอิทธิผลต่อผลสัมฤทธิ์ทางการเรียนได้ดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 6 ผลการสังเคราะห์ตัวแปรทำนายที่มีอิทธิผลต่อผลสัมฤทธิ์ทางการเรียน จำแนกตามระดับตัวแปร

ตัวแปร	Chiu, Xihua (2008)	Gluszynski & Dhawan-Biswal (2008)	Heck (2000)	Douglas Willms. (2004)	Patrick and Thomasz (2004))	Koutsoulis และ Campbell (2001)	Willms &Somers (2001)	D'Agostino (2000)	Opdenakker & Van Damme (2000)	Hofman, Hofman and Guldemon (1999)
ระดับสถานศึกษา										
ขนาดสถานศึกษา			✓	✓				✓		
ที่ตั้งสถานศึกษา				✓				✓	✓	
บริบทสถานศึกษา/สังกัด			✓	✓			✓	✓		✓
รายได้ของหน่วยงาน	✓						✓			
แหล่งทรัพยากร อุปกรณ์							✓			
พฤติกรรมการสอนของครู									✓	
ประสบการณ์สอนของครู										
วุฒิการศึกษาของครู								✓		
ความเป็นผู้นำทางวิชาการ								✓		✓
ระดับนักเรียน										
เพศ		✓	✓	✓	✓			✓		✓
อายุ		✓	✓							
การย้ายถิ่น/ แหล่งอาศัย		✓		✓						
เชื้อชาติ/ชาติกำเนิด		✓	✓	✓						
เศรษฐานะของครอบครัว	✓	✓	✓	✓		✓	✓			✓
คุณลักษณะของครอบครัว	✓	✓		✓		✓	✓	✓		
ทรัพยากรในครอบครัว	✓							✓		
ความรู้เดิมของนักเรียน			✓	✓	✓	✓		✓	✓	✓
ความเชื่อมั่นในตนเอง	✓									
เจตคติและแรงจูงใจ	✓		✓			✓		✓	✓	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 6 ผลการสังเคราะห์ตัวแปรทำนายที่มีอิทธิผลต่อผลสัมฤทธิ์ทางการเรียน จำแนกตามระดับตัวแปร (ต่อ)

ตัวแปร	Silin and Murray-Harvey (1999)	Strand (1998)	พิชิต ธรรมรักษ์ (2549)	รัชนก บุญปู (2547)	ศุภลักษณ์ ใจแสวงทรัพย์ (2547)	สุบิน ยุระรัช (2547)	สุชีรา มะหิเมือง (2547)	ปิยะธิดา ทองอร่าม (2545)	บุญเรือง ศรีเหรัญ (2542)	นิตยา เหมือดไธสง (2543)
ระดับสถานศึกษา										
ขนาดสถานศึกษา			✓							
ที่ตั้งสถานศึกษา							✓			
บริบทสถานศึกษา/สังกัด	✓						✓			
รายได้ของหน่วยงาน									✓	
แหล่งทรัพยากร อุปกรณ์									✓	
พฤติกรรมการสอนของครู, สัดส่วน				✓				✓	✓	✓
ประสบการณ์สอนของครู			✓		✓					
วุฒิการศึกษาของครู				✓						
ความเป็นผู้นำทางวิชาการ			✓					✓		
ระดับนักเรียน										
เพศ		✓					✓			
อายุ		✓								
การย้ายถิ่น/ แหล่งอาศัย										
เชื้อชาติ/ชาติกำเนิด										
เศรษฐานะของครอบครัว	✓	✓	✓		✓		✓			
คุณลักษณะของครอบครัว						✓			✓	
ทรัพยากรในครอบครัว										
ความรู้เดิมของนักเรียน			✓		✓		✓		✓	✓
ความเชื่อมั่นในตนเอง	✓					✓				
เจตคติและแรงจูงใจ	✓		✓	✓	✓	✓			✓	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

จากผลการสังเคราะห์งานวิจัย 20 เรื่องที่ศึกษาปัจจัยที่มีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียน ซึ่งใช้เป็นตัวบ่งชี้ถึงคุณภาพการจัดการศึกษาของสถานศึกษา เมื่อนำมาพิจารณาตัวแปรที่ศึกษามากที่สุด ในตัวแปรระดับนักเรียน ได้แก่ ความรู้เดิม และเศรษฐานะของครอบครัว ส่วนตัวแปรระดับสถานศึกษาที่นำมาศึกษามากที่สุดได้แก่ สังกัดสถานศึกษา ขนาดสถานศึกษา ที่ตั้งสถานศึกษา และแหล่งข้อมูลของโครงการประเมินผล PISA 2006 ที่นำมาใช้วิจัยในครั้งนี้ สรุปได้ดังตารางที่ 7

ตารางที่ 7 อันดับของตัวแปรที่นิยมนำมาศึกษาปัจจัยที่มีผลต่อคุณภาพการจัดการศึกษา และแหล่งข้อมูลในการเก็บข้อมูลของโครงการประเมินผล PISA 2006

อันดับที่	ตัวแปร	แหล่งข้อมูลที่นำมาศึกษาใน PISA 2006
ระดับสถานศึกษา		
1	สังกัดสถานศึกษา	ประเภทสถานศึกษา
2	ขนาดสถานศึกษา	จำนวนนักเรียน
2	ที่ตั้งสถานศึกษา	ขนาดชุมชนที่ตั้งของสถานศึกษา
3	พฤติกรรมการสอนของครู	- สัดส่วนของจำนวนนักเรียนต่อครู - ขนาดห้องเรียน - ดัชนีการขาดแคลนครู
4	ความเป็นผู้นำทางวิชาการ	- นโยบายและแนวปฏิบัติของสถานศึกษา - แรงกดดันของผู้ปกครองที่มีต่อสถานศึกษา
5	รายได้ของหน่วยงาน	จำนวนงบประมาณ
6	แหล่งทรัพยากร อุปกรณ์	ดัชนีทรัพยากรการเรียน สัดส่วนของจำนวนนักเรียนต่อครู
6	วุฒิการศึกษาของครู	จำนวนครูที่จบตรงตามวิชาที่สอน
7	ประสบการณ์สอนของครู	-
ระดับนักเรียน		
1	ความรู้เดิมของนักเรียน	- ปีที่เข้าศึกษา - ระยะเวลาที่เรียนซ้ำชั้นของนักเรียน
1	เศรษฐานะของครอบครัว	เศรษฐานะของครอบครัว
2	เจตคติและแรงจูงใจ	- เจตคติในการเรียน - แรงจูงใจของนักเรียน
3	เพศ	เพศ
4	คุณลักษณะของครอบครัว	อาชีพ ระดับการศึกษา จำนวนหนังสือที่มีครอบครอง
5	อายุ	ปีเกิด
5	เชื้อชาติ/ชาติกำเนิด	เชื้อชาติ
6	การย้ายถิ่น/ แหล่งอาศัย	การอพยพย้ายถิ่น
6	ทรัพยากรในครอบครัว	- จำนวนหนังสือที่มีครอบครอง - จำนวนคอมพิวเตอร์
6	ความเชื่อมั่นในตนเอง	ความเชื่อตนเองในการเรียนวิทยาศาสตร์

5.3 การกำหนดกรอบความคิดในการวิจัย

จากการศึกษาความสัมพันธ์ระหว่างตัวแปรทางการศึกษาที่ส่งผลต่อคุณภาพการจัดการศึกษาพบว่ายูเนสโก (UNESCO,2006) โครงการประเมินนักเรียนระดับนานาชาติ(PISA, 2005) การศึกษาของ Creemers (1994), Coleman (1966), Palardy (2008) ต่างให้ความสำคัญในปัจจัยนำเข้า (Input) กระบวนการ (Process) และผลผลิต (Outputs) ซึ่งตัวแปรในแต่ละองค์ประกอบมีผลต่อผลสัมฤทธิ์ทางวิชาการของนักเรียน และจากผลการวิจัย สรุปตัวแปรแต่ละส่วนโดยแบ่งเป็นระดับนักเรียน ระดับห้องเรียน และระดับสถานศึกษา ได้ดังนี้

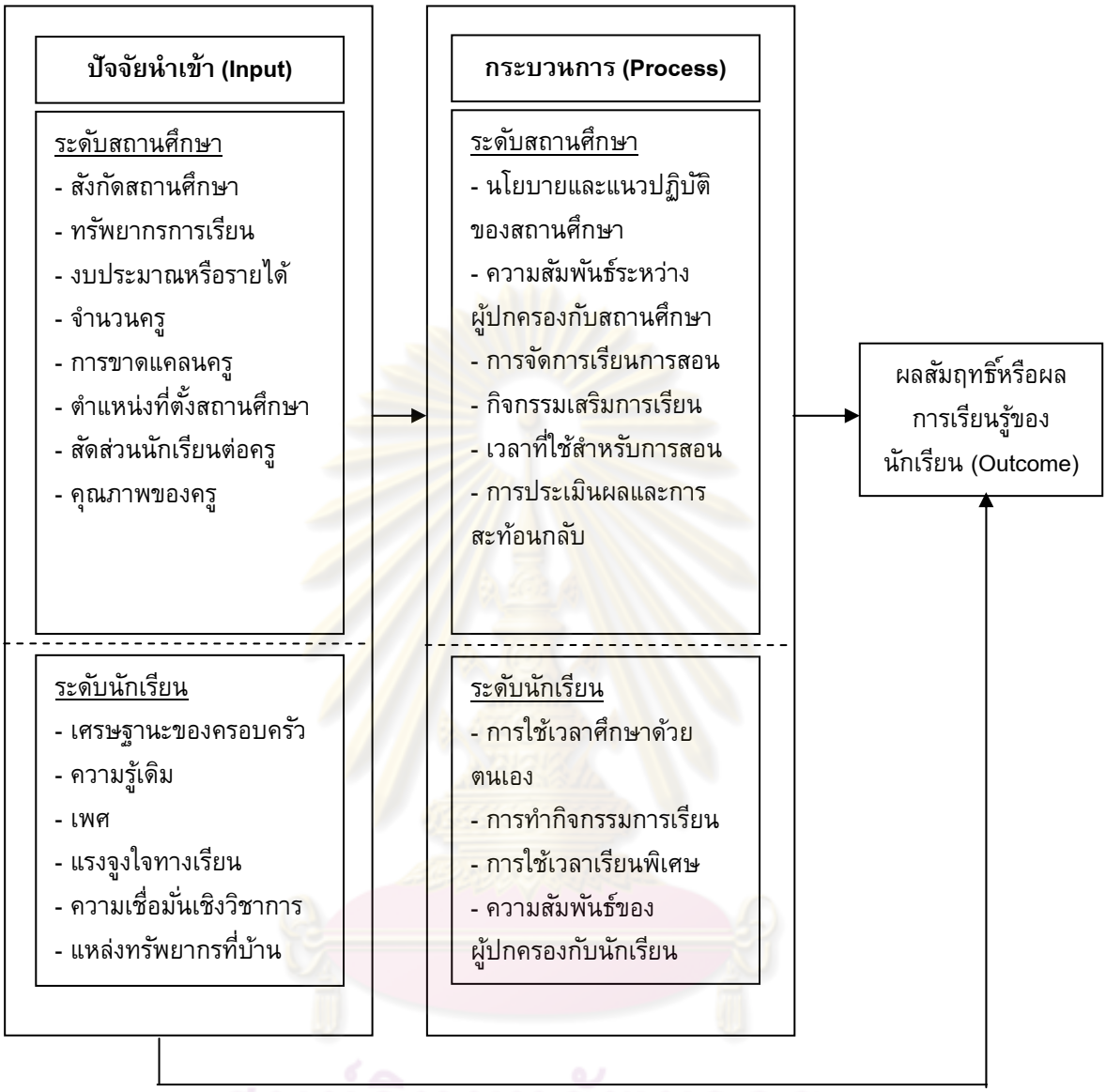
ตัวแปรระดับนักเรียน ส่วนของปัจจัยนำเข้าได้แก่เศรษฐานะของครอบครัว ความรู้เดิม เพศ แรงจูงใจและเจตคติทางเรียนของนักเรียน ความเชื่อมั่นเชิงวิชาการ จำนวนแหล่งทรัพยากรที่บ้าน ส่วนของกระบวนการได้แก่ เวลาที่ใช้ทำการบ้าน การทำกิจกรรมนอกเหนือห้องเรียน การใช้เวลาเรียน ความสัมพันธ์ของผู้ปกครองกับนักเรียน

ตัวแปรระดับสถานศึกษา ส่วนของปัจจัยนำเข้าได้แก่ทรัพยากรการเรียน งบประมาณหรือรายได้ของสถานศึกษา จำนวนครู การขาดแคลนครู ส่วนของกระบวนการได้แก่ นโยบายและแนวปฏิบัติของสถานศึกษา ความสัมพันธ์ระหว่างผู้ปกครองกับสถานศึกษา สัดส่วนของจำนวนนักเรียนต่อครู ความรู้หรือคุณภาพของครู ขนาดห้องเรียน ส่วนของกระบวนการได้แก่ การจัดการเรียนการสอน กิจกรรมเสริมการเรียน เวลาที่ใช้สำหรับการสอน การใช้หลักสูตรและตำราเรียน การประเมินผลและการสะท้อนกลับ

สามารถสรุปกรอบความคิดเชิงทฤษฎีของตัวแปรที่เกี่ยวข้องกับผลการเรียนรู้ของนักเรียน ซึ่งจัดเป็นคุณภาพการจัดการศึกษาของสถานศึกษา ดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาพที่ 6 กรอบความคิดเชิงทฤษฎีของตัวแปรที่เกี่ยวข้องกับผลการเรียนรู้ของนักเรียน



จากการศึกษาค้นคว้าเอกสารและงานวิจัย การประเมินคุณภาพการจัดการศึกษาให้มีความถูกต้องและยุติธรรม จะต้องพิจารณาทั้งแบบสอบที่ใช้ประเมินจะต้องมีความยุติธรรมสำหรับกลุ่มผู้สอบทุกกลุ่ม มีความเชื่อถือได้ ไม่โอนเอียงเข้าหากลุ่มผู้สอบกลุ่มใดกลุ่มหนึ่งเมื่อผู้สอบมีความสามารถอยู่ในระดับเท่ากันหรือเป็นแบบสอบที่ไม่มีการทำหน้าที่ต่างกันของข้อสอบ ตลอดจนการประเมินผลคะแนนการสอบควรจะต้องคำนึงถึงความแตกต่างของปัจจัยของนักเรียนและสถานศึกษา ดังนั้นผู้วิจัยจึงกำหนดกรอบการวิจัยเพื่อหาคะแนนมูลค่าเพิ่มของสถานศึกษา โดยเปรียบเทียบโมเดลการวิเคราะห์ 4 โมเดล ประกอบด้วย

โมเดล 1 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ ส่วนโมเดลการวิเคราะห์มูลค่าเพิ่มเป็นโมเดลที่ยังไม่มีการปรับแก้ (Unadjusted) คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา

โมเดล 2 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ ส่วนโมเดลการวิเคราะห์มูลค่าเพิ่มเป็นโมเดลที่มีการปรับแก้ (Adjusted) คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา

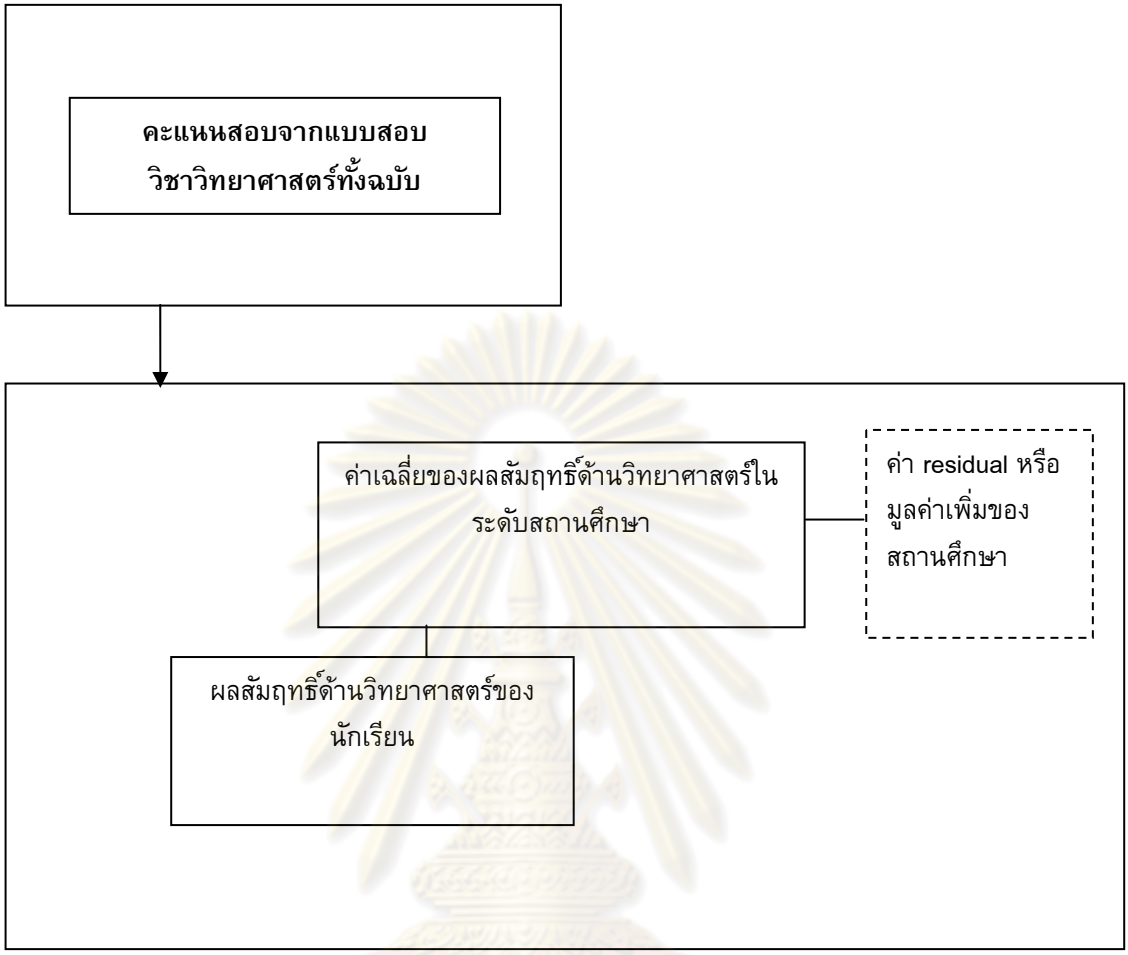
โมเดล 3 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ ส่วนโมเดลการวิเคราะห์มูลค่าเพิ่มเป็นโมเดลที่ยังไม่มีการปรับแก้ (Unadjusted) คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา

โมเดล 4 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ ส่วนโมเดลการวิเคราะห์มูลค่าเพิ่มเป็นโมเดลที่มีการปรับแก้ (Adjusted) ด้วยตัวแปรระดับนักเรียนและสถานศึกษา

กรอบการวิเคราะห์ในแต่ละโมเดล เขียนเป็นแผนภาพได้ดังนี้

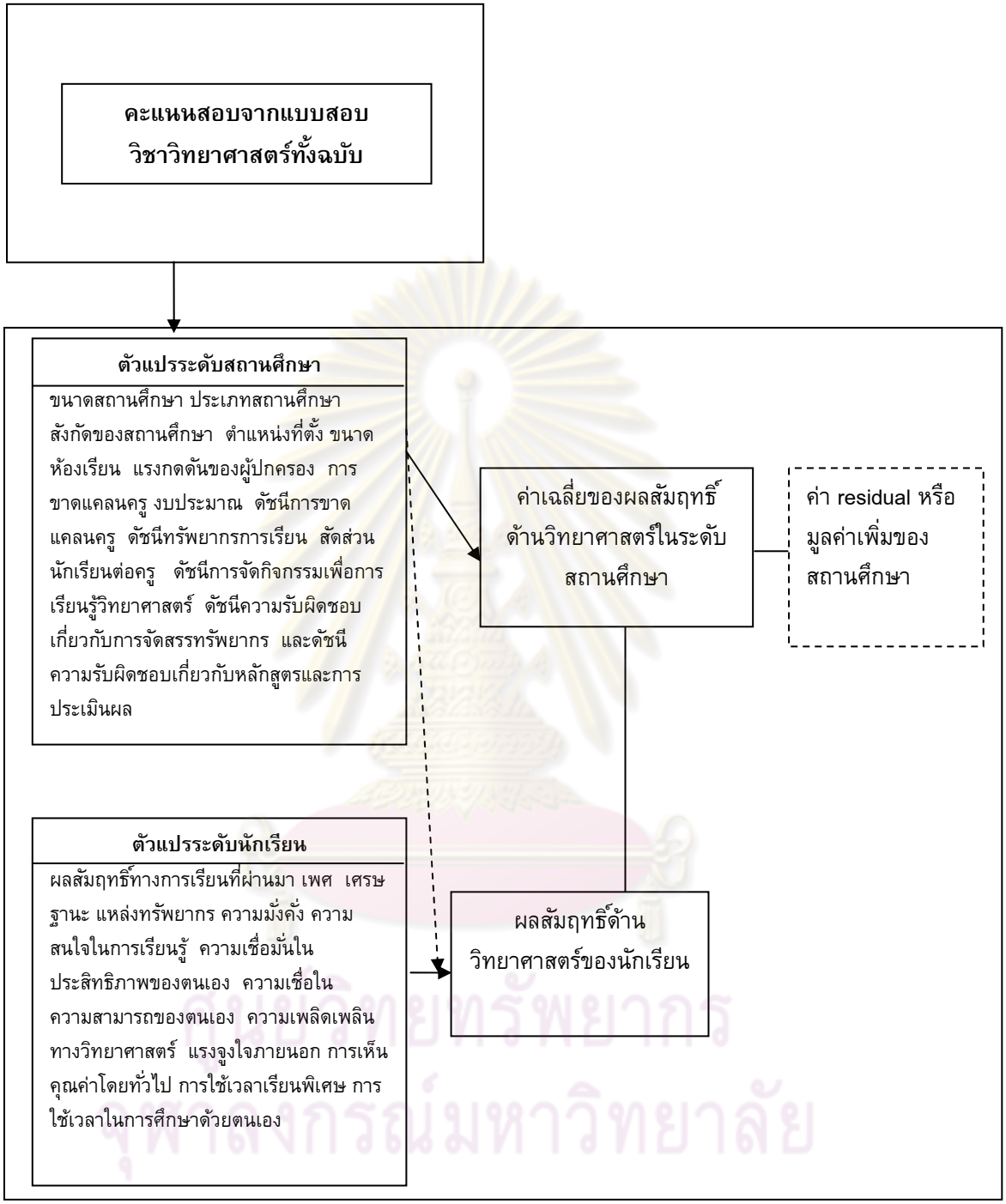
ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาพที่ 7 โมเดลการวิเคราะห์ในโมเดล 1 (UnDetect DIF & Unadjusted)

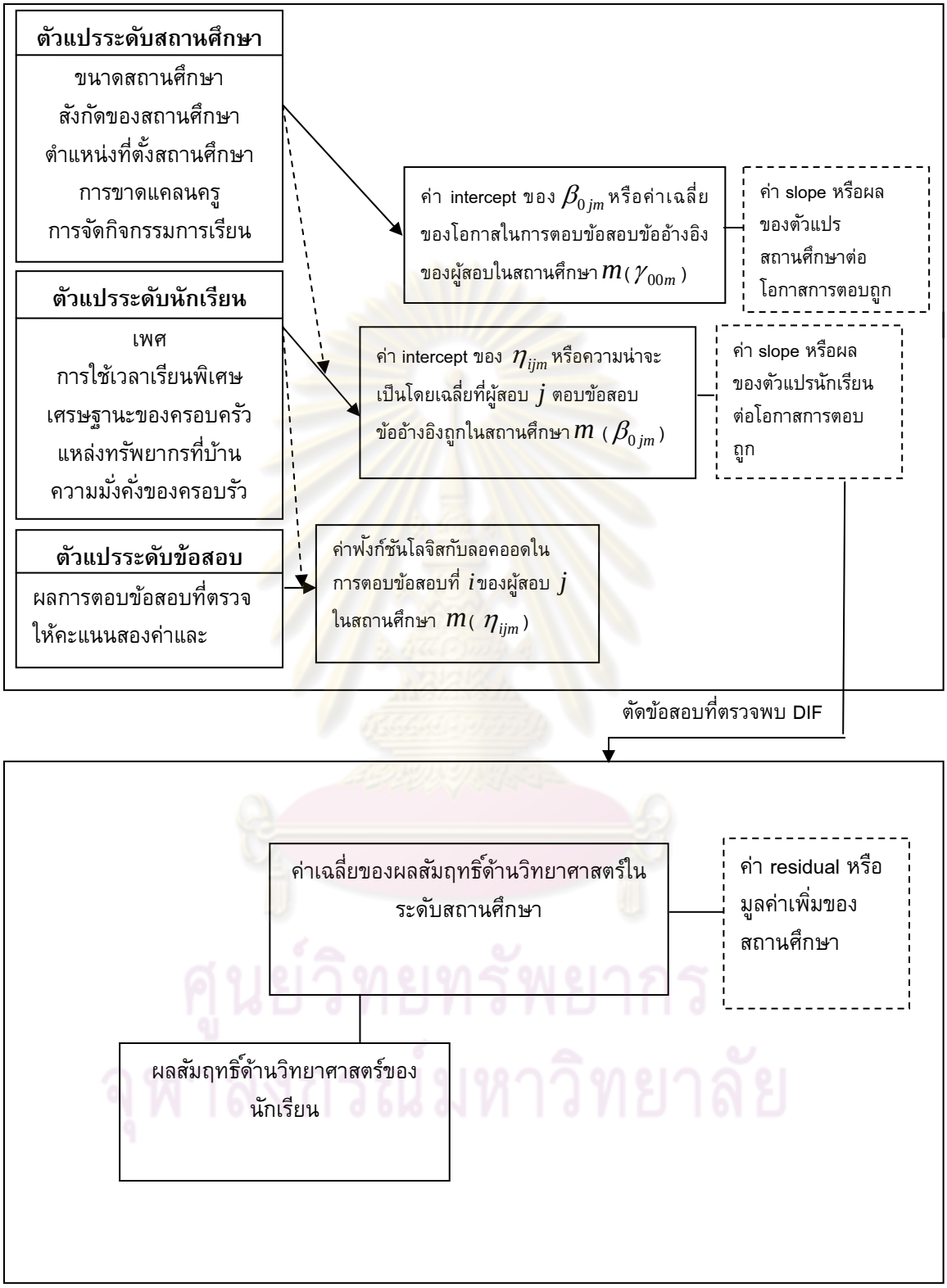


ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

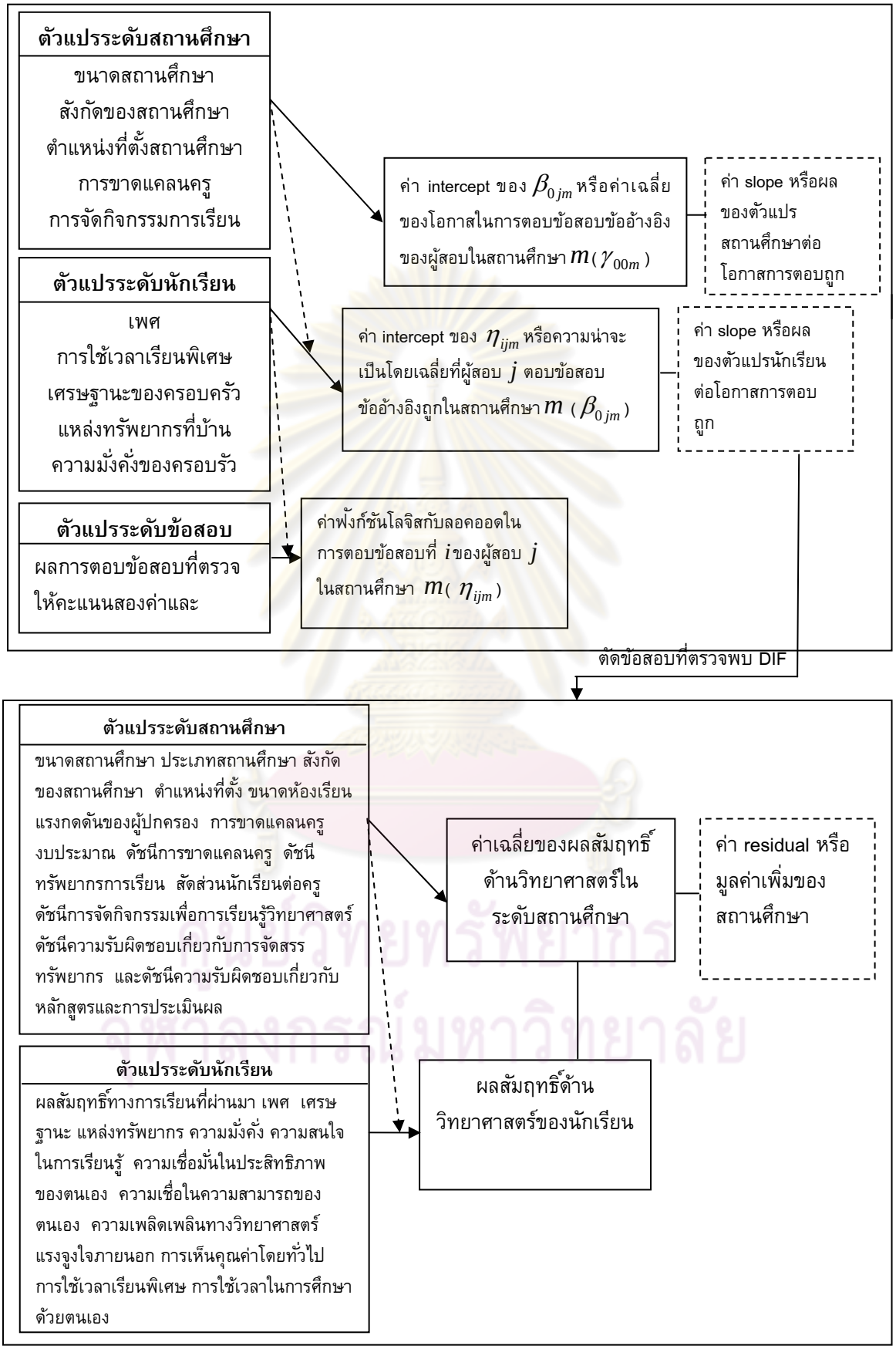
ภาพที่ 8 โมเดลการวิเคราะห์ในโมเดล 2 (UnDetect DIF & Adjusted)



ภาพที่ 9 โมเดลการวิเคราะห์ในโมเดล 3 (Detect DIF & Unadjusted)

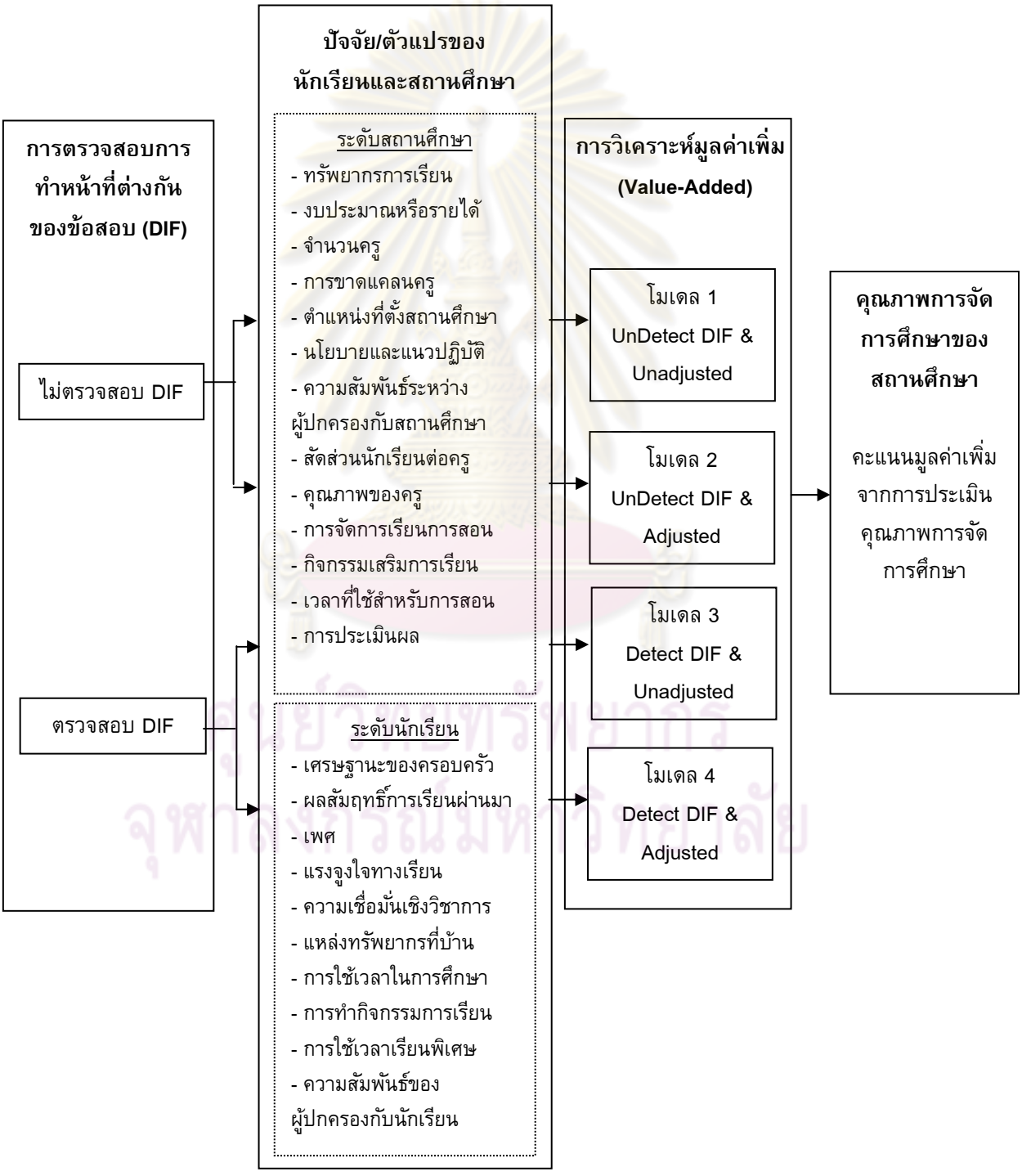


ภาพที่ 10 โมเดลการวิเคราะห์ในโมเดล 4 (Detect DIF & Adjusted)



ในการวิจัยครั้งนี้ ผู้วิจัยมีวัตถุประสงค์เพื่อประเมินคุณภาพการจัดการศึกษาของสถานศึกษาที่มีความยุติธรรมและถูกต้องเชื่อถือได้ พร้อมทั้งเปรียบเทียบโมเดลประเมินคุณภาพการจัดการศึกษา โดยการประยุกต์ใช้การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ(Differential item functioning: DIF) พร้อมกับปรับแก้คะแนนด้วยตัวแปรระดับระดับนักเรียนและสถานศึกษาด้วยการวิเคราะห์มูลค่าเพิ่ม (Value-added analysis) เขียนกรอบแนวคิดในการวิจัยดังนี้

ภาพที่ 11 กรอบแนวคิดในการวิจัยเพื่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน



บทที่ 3

วิธีการดำเนินการวิจัย

การวิจัยในครั้งนี้มีจุดประสงค์เพื่อประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน โดยการประยุกต์ใช้การวิเคราะห์มูลค่าเพิ่มและการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ข้อมูลที่ใช้ในการวิจัยเป็นข้อมูลจากโครงการประเมินผลนักเรียนระดับนานาชาติ ปี 2549 (PISA 2006) วิชาวิทยาศาสตร์ โดยใช้โปรแกรมโมเดลเชิงเส้นระดับลดหลั่น (HLM) เวอร์ชั่น 6.06 ในการวิเคราะห์ข้อมูล

ผู้วิจัยมีการดำเนินการวิจัย เสนอรายละเอียดการดำเนินการดังนี้

3.1 ข้อมูลที่ใช้ในการวิจัย

ข้อมูลที่ใช้ในการวิจัยครั้งนี้เป็นข้อมูลทุติยภูมิ ซึ่งเป็นฐานข้อมูล (data base) ของโครงการประเมินผลนักเรียนระดับนานาชาติ ปี 2549 หรือ PISA 2006 โดยกลุ่มตัวอย่างของการประเมินเป็นนักเรียนของประเทศไทยที่มีอายุ 15 ปี ซึ่งเป็นวัยที่จบการศึกษาภาคบังคับ ในทางปฏิบัติได้ใช้นักเรียนที่มีช่วงอายุ 15 ปี 3 เดือน จนถึง 16 ปี 2 เดือน สำหรับประเทศไทยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 – 3

ฐานข้อมูลของโครงการประเมินผล PISA ประกอบด้วยสถานศึกษาที่เป็นกลุ่มตัวอย่าง 212 แห่ง จำแนกตามสังกัดของสถานศึกษา และจะสุ่มนักเรียนในแต่ละสถานศึกษาประมาณ 35 คน ได้นักเรียนที่สุ่มได้ทั้งหมด 6,470 คน มีจำนวนนักเรียนที่เข้าสอบจริง 6,192 คน ดังตารางที่ 8

ตารางที่ 8 จำนวนกลุ่มตัวอย่างนักเรียนในโครงการประเมินผล PISA 2006 จำแนกตามสังกัดของสถานศึกษา

สถานศึกษาในสังกัด	สถานศึกษากลุ่มตัวอย่าง	นักเรียนกลุ่มตัวอย่าง	นักเรียนที่เข้าสอบ
สพฐ.	134	4,029	3,873
สช.	25	805	756
กทม.	15	405	349
กศท.	15	438	431
โรงเรียนสาธิต (สกอ.)	15	525	521
โรงเรียนและวิทยาลัยอาชีวศึกษา (อศ.)	8	268	262
รวม	**212**	**6,470**	**6,192**

3.2 รูปแบบของแบบสอบประเมินการรู้เรื่องด้านวิทยาศาสตร์และการจัดการสอบ

แบบสอบประเมินการรู้เรื่องด้านวิทยาศาสตร์ที่ใช้ประเมินในโครงการประเมินผล PISA 2006 จะวัดในสมรรถนะที่เหมาะสมกับนักเรียนวัย 15 ปี เกี่ยวกับความรู้ทางวิทยาศาสตร์ ซึ่งครอบคลุมทั้งความรู้วิทยาศาสตร์ และความรู้เกี่ยวกับวิทยาศาสตร์ นั่นคือวัดทั้งองค์ความรู้ของวิชาวิทยาศาสตร์ และความรู้ในวิธีการหรือกระบวนการหาความรู้ทางวิทยาศาสตร์ กำหนดกรอบการประเมินตามสมรรถนะทางวิทยาศาสตร์ประกอบด้วย 3 สมรรถนะทางวิทยาศาสตร์ คือ การระบุประเด็นทางวิทยาศาสตร์ (Identifying scientific issues: ISI) มีข้อสอบจำนวน 24 ข้อ การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ (Explain phenomena scientifically: EPS) มีข้อสอบ 53 ข้อ และการใช้ประจักษ์พยานทางวิทยาศาสตร์ (Using scientific evidence: USE) มีข้อสอบจำนวน 31 ข้อ รวมจำนวนข้อสอบ 108 ข้อ โดยในแต่ละสมรรถนะทางวิทยาศาสตร์มีการใช้รูปแบบของข้อสอบที่หลากหลาย ได้แก่ แบบเลือกตอบ แบบเลือกตอบเชิงซ้อน แบบสร้างคำตอบแบบอิสระ และสร้างคำตอบแบบปิด ดังแสดงในตารางที่ 9

ตารางที่ 9 ลักษณะของข้อสอบที่ใช้ในโครงการประเมินผล PISA 2006 จำแนกตามสมรรถนะทางวิทยาศาสตร์

สมรรถนะทางวิทยาศาสตร์	เลือกตอบ (ข้อ)	เลือกตอบเชิงซ้อน(ข้อ)	ตอบแบบอิสระ(ข้อ)	ตอบแบบปิด(ข้อ)	ข้อสอบรวม(ข้อ)
การระบุประเด็นทางวิทยาศาสตร์	9	10	5	-	24
การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์	22	11	16	4	53
การใช้ประจักษ์พยานทางวิทยาศาสตร์	7	8	15	1	31
รวม	38	29	36	5	108

เมื่อพิจารณาข้อสอบทั้งหมดในโครงการประเมินผล PISA 2006 โดยคณะกรรมการพิจารณาข้อสอบของโครงการประเมินผล ผลการพิจารณาพบว่า มีข้อสอบที่ต้องตัดออกไปจำนวน 5 ข้อ เนื่องจากจะทำให้นักเรียนเกิดความความเข้าใจผิดในเนื้อหาของข้อสอบได้ (OECD, 2009) จากข้อสอบทั้งหมด จะจัดเป็นแบบสอบ (Booklet) จำนวน 13 ฉบับ ผู้วิจัยได้พิจารณาลักษณะของข้อสอบในรูปแบบต่างๆ ในแบบสอบทั้ง 13 ฉบับ ตามการตรวจให้คะแนน จึงแบ่งออกเป็นสองลักษณะคือ ข้อสอบที่ตรวจให้คะแนนสองค่า (Dichotomous Items) และข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า (Polytomous Items) ซึ่งจะเห็นได้ว่าแบบสอบ 13 ฉบับมีจำนวนข้อสอบที่แตกต่างกัน แบบสอบที่มีทั้งข้อสอบที่ตรวจให้คะแนนสองค่าและมากกว่าสองค่า มี 8 ฉบับ ได้แก่ แบบสอบฉบับที่ 1, 2, 3, 4, 5, 9, 11 และ 12 ซึ่งมีข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่าเพียงฉบับละ 1 – 2 ข้อ ส่วนแบบสอบอีก 5 ฉบับ คือแบบสอบฉบับที่ 6, 7, 8, 10 และ 13 มีเฉพาะข้อสอบที่ตรวจให้คะแนนสองค่า ในแบบสอบแต่ละฉบับมีจำนวนข้อสอบที่แตกต่างกันในสมรรถนะทางวิทยาศาสตร์ โดยมีการจัดน้ำหนักให้ข้อสอบที่วัดในสมรรถนะด้านการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์มี

จำนวนข้อสอบสูงสุด และเมื่อพิจารณาสัดส่วนของน้ำหนักของจำนวนข้อสอบที่วัดในสมรรถนะทางวิทยาศาสตร์สามด้าน พบว่ามีสัดส่วนที่ไม่คงที่ แสดงดังรายละเอียดในตารางที่ 10

ตารางที่ 10 จำนวนข้อสอบที่ใช้ในโครงการประเมินผล PISA 2006 ของแบบสอบ 13 ฉบับ จำแนกตามลักษณะการตรวจให้คะแนนและสมรรถนะทางวิทยาศาสตร์

ฉบับ	ข้อสอบตรวจให้คะแนนสองค่า				ข้อสอบตรวจให้คะแนนมากกว่าสองค่า				รวมทั้งฉบับ
	ISI	EPS	USE	รวม	ISI	EPS	USE	รวม	
1	10	33	15	**58**	1	-	1	**2**	**60**
2	5	12	9	**26**	-	-	2	**2**	**28**
3	7	17	7	**31**	-	-	1	**1**	**32**
4	7	18	8	**33**	-	-	1	**1**	**34**
5	15	24	21	**60**	-	-	2	**2**	**62**
6	8	15	9	**32**	-	-	-	**-**	**32**
7	2	8	5	**15**	-	-	-	**-**	**15**
8	6	10	10	**26**	-	-	-	**-**	**26**
9	11	11	6	**28**	-	-	1	**1**	**29**
10	8	12	8	**28**	-	-	-	**-**	**28**
11	5	12	11	**28**	-	-	2	**2**	**30**
12	8	14	8	**30**	-	-	2	**2**	**32**
13	2	8	5	**15**	-	-	-	**-**	**15**

หมายเหตุ ISI แทน การระบุประเด็นทางวิทยาศาสตร์
EPS แทน การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
USE แทน การใช้ประจักษ์พยานทางวิทยาศาสตร์

การจัดการสอบกับนักเรียนที่อยู่ในสถานศึกษาตามกลุ่มตัวอย่างของโครงการ PISA มีการจัดการให้นักเรียนแต่ละคนได้รับแบบสอบเพียงหนึ่งฉบับด้วยวิธีการสุ่มแบบสอบให้กับนักเรียน ดังนั้นนักเรียนแต่ละคนจึงมีโอกาสเท่ากันที่จะได้รับแบบสอบฉบับใดฉบับหนึ่ง จากตารางที่ 10 จะเห็นว่าแบบสอบฉบับที่ 1 ถึงฉบับที่ 13 มีความแตกต่างกันในด้านองค์ประกอบของสมรรถนะทางวิทยาศาสตร์และความยาวของแบบสอบเนื่องจากจำนวนข้อสอบต่างกัน ดังนั้นการใช้แบบสอบ (booklet) ที่แตกต่างกัน อาจจะทำให้คะแนนของผู้สอบแต่ละคนได้รับอิทธิพลจากแบบสอบ (booklet effect) ซึ่งจะส่งผลต่อการประมาณค่าความสามารถของนักเรียน จากรายงานการประเมินผลนักเรียนตามโครงการประเมินผล PISA 2006 (OECD, 2009) ได้มีการวิเคราะห์อิทธิพลจากแบบสอบ ซึ่งมีทั้งค่าที่เป็นบวกและลบ ถ้าค่าอิทธิพลจากแบบสอบเป็นบวก แสดงว่าแบบสอบฉบับนั้นมีความยากมากกว่าค่าเฉลี่ยของแบบสอบฉบับอื่น และค่าอิทธิพลจากแบบสอบที่เป็นลบแสดงว่าแบบสอบฉบับนั้นง่ายกว่าค่าเฉลี่ยของแบบสอบฉบับอื่น ตัวอย่างเช่นแบบสอบฉบับที่ 3 มีค่าอิทธิพลจากแบบสอบเป็น – 20.5 แสดงแบบสอบฉบับที่ 3 มีความง่ายกว่าแบบสอบฉบับอื่น และแบบสอบฉบับที่ 10 มีความยากกว่าฉบับอื่น ดังนั้นในการแปลงค่าเป็นคะแนนสเกลของ PISA จะ

นำค่าอิทธิพลจากแบบสอบที่วิเคราะห์ได้มาบวกเพิ่มเข้าไปกับคะแนนสเกลของ PISA ของนักเรียนที่ทำแบบสอบแต่ละฉบับ ซึ่งภายหลังการแปลงค่าความสามารถของนักเรียนเป็นคะแนนสเกลของ PISA ในแต่ละฉบับ พบว่าค่าอิทธิพลโดยเฉลี่ยแล้วมีค่าเข้าใกล้ศูนย์ ดังนั้นการแปลงค่าเป็นคะแนนสเกลของ PISA จึงพบว่าไม่มีอิทธิพลจากแบบสอบอย่างมีนัยสำคัญ เนื่องจากมีการควบคุมอิทธิพลจากแบบสอบแล้ว (OECD, 2009) จำนวนนักเรียนที่ทำแบบสอบและค่าประมาณของอิทธิพลจากแบบสอบ แสดงผลดังตารางที่ 11

ตารางที่ 11 จำนวนกลุ่มตัวอย่างนักเรียนที่ทำการทดสอบแต่ละฉบับ ค่าประมาณของอิทธิพลจากแบบสอบและค่าประมาณอิทธิพลจากแบบสอบหลังการแปลงคะแนนสเกลของ PISA

ฉบับ	จำนวนนักเรียนที่ทำแบบสอบ	ค่าประมาณของอิทธิพลจากแบบสอบ
1	474	–3.1
2	475	–20.0
3	468	–20.5
4	466	–6.3
5	469	1.6
6	483	6.7
7	488	–19.9
8	478	17.6
9	473	0.2
10	474	21.4
11	482	12.1
12	486	20.4
13	476	–10.4

3.3 ปัจจัยหรือตัวแปรนำมาใช้ในการวิจัย

ตัวแปรที่นำมาใช้ในการวิจัยครั้งนี้ ผู้วิจัยพิจารณาถึงโครงสร้างของข้อมูลในระบบการจัดการศึกษา ซึ่งลักษณะข้อมูลมีโครงสร้างแบบลดหลั่น นั่นคือนักเรียนสอดแทรกอยู่ในสถานศึกษา ดังนั้นจึงแบ่งตัวแปรเป็นระดับสถานศึกษาและระดับนักเรียน โดยตัวแปรที่สามารถเข้าถึงฐานข้อมูลในโครงการประเมินผล PISA 2006 และนำมาใช้ในการวิจัยมีดังนี้

ตัวแปรระดับสถานศึกษา

1. ลักษณะทั่วไปของสถานศึกษา ประกอบด้วย
    1.1 ขนาดสถานศึกษา
    1.2 ประเภทของสถานศึกษา
    1.3 สังกัดของสถานศึกษา
    1.4 ตำแหน่งที่ตั้งของสถานศึกษา

1.5 ขนาดห้องเรียน

2. ทรัพยากรของสถานศึกษา ประกอบด้วย

2.1 ระดับการขาดแคลนครูวิทยาศาสตร์

2.2 ระดับการขาดแคลนครูที่มีคุณวุฒิ

2.3 ร้อยละงบประมาณของรัฐที่จัดสรรให้สถานศึกษา

2.4 ร้อยละงบประมาณจากหน่วยงานภายนอก

2.5 ดัชนีการขาดแคลนครู

2.6 สัดส่วนนักเรียนต่อครู

2.7 ดัชนีทรัพยากรการเรียน

3. การบริหารสถานศึกษา ประกอบด้วย

3.1 แรงกดดันของผู้ปกครองต่อสถานศึกษา

3.2 ดัชนีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์

3.3 ดัชนีความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากร

3.4 ดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล

ตัวแปรระดับนักเรียน

1. ลักษณะทั่วไปของนักเรียน ประกอบด้วย

1.1 คะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์

1.2 เพศ

1.3 ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา

1.4 การใช้เวลาเรียนพิเศษวิทยาศาสตร์นอกสถานศึกษา

1.5 การใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง

2. ทัศนคติในการเรียน ประกอบด้วย

2.1 ความสนใจในการเรียนรู้วิทยาศาสตร์

2.2 ความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์

2.3 ความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์

2.4 ความเพลิดเพลินทางวิทยาศาสตร์

2.5 แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์

2.6 การเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์

3. พื้นฐานครอบครัว ประกอบด้วย

3.1 เศรษฐานะของครอบครัว

3.2 แหล่งทรัพยากรที่บ้าน

3.3 ความมั่งคั่งของครอบครัว

ความหมายของตัวแปรมีรายละเอียดดังตารางที่ 12

ตารางที่ 12 สเกลการวัดและความหมายของตัวแปรในระดับนักเรียนและสถานศึกษา

ตัวแปร	สเกลการวัด / ความหมาย
ระดับสถานศึกษา	
1. ขนาดสถานศึกษา	ตัวแปรจัดประเภท แสดงจำนวนนักเรียนในสถานศึกษา จำแนกเป็น ขนาดใหญ่ ขนาดเล็กปานกลาง และขนาดเล็กมาก
2. ประเภทของสถานศึกษา	ตัวแปรจัดประเภท จำแนกเป็นสถานศึกษาของรัฐ และเอกชน
3. สังกัดของสถานศึกษา	ตัวแปรจัดประเภท จำแนกเป็นสถานศึกษาสังกัดสพฐ., สช. , สกอ., กศท.และ กทม. และ อศ.
4. ตำแหน่งที่ตั้งของสถานศึกษา	ตัวแปรจัดประเภท แสดงขนาดของชุมชนที่สถานศึกษาตั้งอยู่ จำแนกเป็น ตำแหน่งหมู่บ้าน เมืองเล็ก เมือง เมืองใหญ่ เมืองใหญ่มาก
5. ขนาดห้องเรียน	ตัวแปรจัดประเภท แสดงจำนวนนักเรียนในห้องเรียน จำแนกเป็น นักเรียนน้อยกว่าหรือเท่ากับ 15 คน, 16-20 คน, 21-25 คน, 26-30 คน, 31-35 คน, 36-40 คน, 41-45 คน, 46-50 คน และมากกว่า 50 คน
6. แรงกดดันของผู้ปกครองที่มีต่อสถานศึกษา	ตัวแปรจัดประเภท แสดงถึงบทบาทของผู้ปกครองที่กดดันให้สถานศึกษามีคุณภาพหรืออยู่ในมาตรฐานวิชาการที่สูงขึ้น จำแนกเป็น 3 ระดับคือ ผู้ปกครองจำนวนมาก ผู้ปกครองส่วนน้อย ไม่พบเป็นส่วนใหญ่
7. การขาดแคลนครูวิทยาศาสตร์	ตัวแปรจัดประเภท แสดงระดับการขาดแคลนครูวิทยาศาสตร์ จำแนกเป็น 4 ระดับคือ ไม่ขาดแคลน ขาดแคลนน้อย ขาดแคลนเป็นบางส่วน ขาดแคลนจำนวนมาก
8. การขาดแคลนครูที่มีคุณวุฒิ	ตัวแปรจัดประเภท แสดงระดับการขาดแคลนครูที่มีคุณวุฒิ จำแนกเป็น 4 ระดับคือ ไม่ขาดแคลน ขาดแคลนน้อย ขาดแคลนเป็นบางส่วน ขาดแคลนจำนวนมาก
9. ดัชนีการขาดแคลนครู	ตัวแปรต่อเนื่อง เป็นค่าที่แสดงถึงการขาดแคลนครู มีค่าเฉลี่ยเป็น 0 และค่าเบี่ยงเบนมาตรฐานเป็น 1 โดยค่าดัชนีที่เป็นบวกแสดงว่ามีปัญหาอุปสรรคจากการขาดแคลนครูมากกว่าค่าเฉลี่ย และค่าดัชนีที่เป็นลบหมายถึงมีปัญหาน้อยกว่าค่าเฉลี่ย
10. ดัชนีทรัพยากรการเรียน	ตัวแปรต่อเนื่อง เป็นค่าที่แสดงความเพียงพอของทรัพยากรการเรียน มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายความว่ามีความเพียงพอหรือไม่ขาดแคลนหรือขาดแคลนน้อยกว่าค่าเฉลี่ย และถ้าดัชนีเป็นลบแสดงถึงความขาดแคลนทรัพยากรสูงกว่าค่าเฉลี่ย
11. สัดส่วนของจำนวนนักเรียนต่อครู	ตัวแปรต่อเนื่อง ได้จากสัดส่วนของจำนวนนักเรียนต่อครู 1 คน
12. ร้อยละงบประมาณของรัฐที่จัดสรรให้สถานศึกษา	ตัวแปรต่อเนื่อง คิดจากร้อยละงบประมาณของรัฐที่จัดสรรให้กับสถานศึกษา
13. ร้อยละงบประมาณจากหน่วยงานภายนอก	ตัวแปรต่อเนื่อง คิดจากร้อยละงบประมาณที่สถานศึกษาจัดหาจากหน่วยงานภายนอก เช่น เงินสนับสนุนจากชุมชน ร้านค้าและเงินบริจาค

ตารางที่ 12 สเกลการวัดและความหมายของตัวแปรในระดับนักเรียนและสถานศึกษา (ต่อ)

ตัวแปร	สเกลการวัด / ความหมาย
14. ดัชนีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์	ตัวแปรต่อเนื่อง เป็นค่าที่แสดงการจัดกิจกรรมเพื่อการเรียนรู้ทางวิทยาศาสตร์ของสถานศึกษา มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายความว่ากิจกรรมต่อการเรียนรู้ของนักเรียน และถ้าดัชนีเป็นลบแสดงถึงจัดกิจกรรมการเพื่อการเรียนรู้ต่ำกว่าค่าเฉลี่ย
15. ดัชนีความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากร	ตัวแปรต่อเนื่อง เป็นค่าที่แสดงความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากรของสถานศึกษา มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายความว่ามีความรับผิดชอบต่อการจัดสรรทรัพยากรสูงกว่าค่าเฉลี่ย ถ้าดัชนีเป็นลบแสดงถึงความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากรต่ำกว่าค่าเฉลี่ย
16. ดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล	ตัวแปรต่อเนื่อง เป็นค่าที่แสดงความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายความว่ามีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล สูงกว่าค่าเฉลี่ย ถ้าดัชนีเป็นลบแสดงถึงมีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล ต่ำกว่าค่าเฉลี่ย
<u>ระดับนักเรียน</u>	
1. คะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์	ตัวแปรต่อเนื่อง เป็นคะแนนที่ได้จากแปลงค่าความสามารถของนักเรียนเป็นคะแนนสเกลของ PISA ในวิชาวิทยาศาสตร์
2. ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา	ค่าที่ได้จากการคำนวณหาผลต่างระหว่างระดับชั้นของนักเรียนกับระดับชั้นที่นักเรียนคาดหวังของนักเรียน (Chiu & Xihua, 2008)
3. เพศ	ตัวแปรจัดประเภท จำแนกเป็นนักเรียนชายและหญิง
4. เศรษฐานะของครอบครัว	ตัวแปรต่อเนื่อง เป็นค่าดัชนีเศรษฐานะของครอบครัว โดยพิจารณาจากตัวแปรรายได้ของครอบครัว อาชีพของผู้ปกครอง และระดับการศึกษา
5. แหล่งทรัพยากรที่บ้าน	ตัวแปรต่อเนื่อง เป็นค่าดัชนีที่บ่งบอกถึงความพร้อมของทรัพยากรในบ้าน ได้แก่ หนังสือ คอมพิวเตอร์ เป็นต้น มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายความว่ามีแหล่งทรัพยากรที่บ้านสูง ถ้าดัชนีเป็นลบแสดงถึงมีแหล่งทรัพยากรที่บ้านต่ำกว่าค่าเฉลี่ย
6. ความมั่งคั่งของครอบครัว	ตัวแปรต่อเนื่อง เป็นค่าดัชนีที่บ่งบอกถึงความเป็นอยู่ ความสะดวกสบายของครอบครัว มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายความว่ามีความมั่งคั่งของครอบครัวสูง ถ้าดัชนีเป็นลบแสดงถึงมีความมั่งคั่งของครอบครัวต่ำกว่าค่าเฉลี่ย
7. ความสนใจในการเรียนรู้วิทยาศาสตร์	ตัวแปรต่อเนื่อง เป็นค่าดัชนีที่บ่งบอกถึงความกระตือรือล้น ใฝ่รู้ในประเด็นทางวิทยาศาสตร์ ความตั้งใจที่จะหาความรู้ ทักษะ วิทยาศาสตร์เพิ่มเติม ความตั้งใจที่จะใฝ่หาสาระและแสดงความสนใจวิทยาศาสตร์อย่างต่อเนื่อง มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายถึงมีความสนใจในการเรียนรู้วิทยาศาสตร์สูง ถ้าดัชนีเป็นลบหมายถึงมีความสนใจในการเรียนรู้วิทยาศาสตร์ต่ำกว่าค่าเฉลี่ย

ตารางที่ 12 สเกลการวัดและความหมายของตัวแปรในระดับนักเรียนและสถานศึกษา (ต่อ)

ตัวแปร	สเกลการวัด / ความหมาย
8. ความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์	ตัวแปรต่อเนื่อง เป็นค่าดัชนีที่บ่งบอกถึงความเชื่อมั่นในตนเองในความสามารถที่จะเอาชนะความบากลำบากในการเรียนวิทยาศาสตร์ มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายถึงมีความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์สูง ถ้าดัชนีเป็นลบแสดงถึงมีความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์ต่ำกว่าค่าเฉลี่ย
9. ความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์	ตัวแปรต่อเนื่อง เป็นค่าดัชนีที่บ่งบอกถึงความเชื่อมั่นในความรู้ความสามารถเกี่ยวกับวิทยาศาสตร์ของตนเอง มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายความว่ามีความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์สูง ถ้าดัชนีเป็นลบแสดงถึงมีความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์ต่ำกว่าค่าเฉลี่ย
10. ความเพลิดเพลินทางวิทยาศาสตร์	ตัวแปรต่อเนื่อง เป็นค่าดัชนีที่บ่งบอกถึงความสนุกสนาน ความชื่นชอบในการเรียนวิทยาศาสตร์ของนักเรียน มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายถึงมีความเพลิดเพลินทางวิทยาศาสตร์สูง ถ้าดัชนีเป็นลบแสดงถึงมีความเพลิดเพลินทางวิทยาศาสตร์ต่ำกว่าค่าเฉลี่ย
11. แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์	ตัวแปรต่อเนื่อง เป็นค่าดัชนีที่บ่งบอกถึงแรงจูงใจจากภายนอกในการเรียนรู้และการทำกิจกรรมเพื่อเกิดการเรียนรู้วิทยาศาสตร์ของนักเรียน มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายความว่ามีแรงจูงใจภายนอกในการเรียนวิทยาศาสตร์สูง ถ้าดัชนีเป็นลบแสดงถึงมีแรงจูงใจภายนอกในการเรียนวิทยาศาสตร์ต่ำกว่าค่าเฉลี่ย
12. การเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์	ตัวแปรต่อเนื่อง เป็นค่าดัชนีที่บ่งบอกถึงความตระหนักถึงคุณค่าของวิทยาศาสตร์ต่อการดำรงชีวิตประจำวันของนักเรียน มีค่าเฉลี่ยเป็น 0 ถ้าค่าดัชนีเป็นบวกหมายความว่ามีการเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์สูง ถ้าดัชนีเป็นลบแสดงถึงมีการเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์ต่ำกว่าค่าเฉลี่ย
13. การใช้เวลาเรียนพิเศษวิทยาศาสตร์นอกสถานศึกษา	ตัวแปรจัดประเภท แสดงระดับของการใช้เวลาเรียนในการเรียนพิเศษของนักเรียนในหนึ่งสัปดาห์ แบ่งเป็น 5 ระดับ คือไม่มีเวลา น้อยกว่า 2 ชม. 2 – 4 ชม., 4 – 6 ชม. และมากกว่า 6 ชม.
14. การใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง	ตัวแปรจัดประเภท แสดงระดับของการใช้เวลาเรียนในการศึกษาความรู้ ทบทวนความรู้ด้วยตนเองของนักเรียนในหนึ่งสัปดาห์ แบ่งเป็น 5 ระดับ คือ ไม่มีเวลา น้อยกว่า 2 ชม. 2 – 4 ชม. 4 – 6 ชม.และมากกว่า 6 ชม.

3.4 วิธีการดำเนินการวิจัย

การดำเนินการวิจัยในครั้ง ผู้วิจัยมีขั้นตอนหลัก 3 ขั้นดังนี้

ขั้นที่ 1 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

ขั้นที่ 2 การวิเคราะห์มูลค่าเพิ่มในโมเดลการประเมินคุณภาพการจัดการศึกษา

ขั้นที่ 3 การตรวจสอบผลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา

แต่ละขั้นตอนมีรายละเอียด ดังนี้

ขั้นที่ 1 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ โดยการประยุกต์ใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น (HGLM) จะต้องมีการเตรียมไฟล์ข้อมูลสำหรับการวิเคราะห์พหุระดับ การวิเคราะห์ข้อมูลในส่วนนี้ จะได้ค่าพารามิเตอร์ที่สำคัญคือ ค่าความสามารถของนักเรียน ค่าความยากของข้อสอบ และอิทธิพลของตัวแปรภายนอกที่ทำให้มีโอกาสตอบข้อสอบได้ถูกต่างกันหรือตัวแปรที่ทำให้เกิดการทำหน้าที่ต่างกันของข้อสอบ การดำเนินการในขั้นตอนนี้มี 3 ขั้นตอนดังนี้

1.1 การเตรียมระบบฐานข้อมูลสำหรับการวิเคราะห์

การวิเคราะห์ข้อมูลครั้งนี้ผู้วิจัยใช้โมเดลเชิงเส้นตรงทั่วไประดับลดหลั่น(HGLM) ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และโมเดลเชิงเส้นตรงแบบลดหลั่น (HLM) สำหรับการวิเคราะห์มูลค่าเพิ่ม ผู้วิจัยจัดเตรียมข้อมูลโดยใช้โปรแกรม SPSS for Windows V.14.0 แบ่งตามระดับของการวิเคราะห์ จำนวน 3 ไฟล์ ดังนี้

<u>ไฟล์ที่ 1</u> : ระดับข้อสอบ

ผู้วิจัยจัดไฟล์ระดับข้อสอบแยกตามฉบับของแบบสอบจำนวน 13 ไฟล์ และแต่ละไฟล์ประกอบด้วยรหัสสถานศึกษา รหัสนักเรียน รหัสของแบบสอบ และผลการตอบ กำหนดรหัสดังนี้

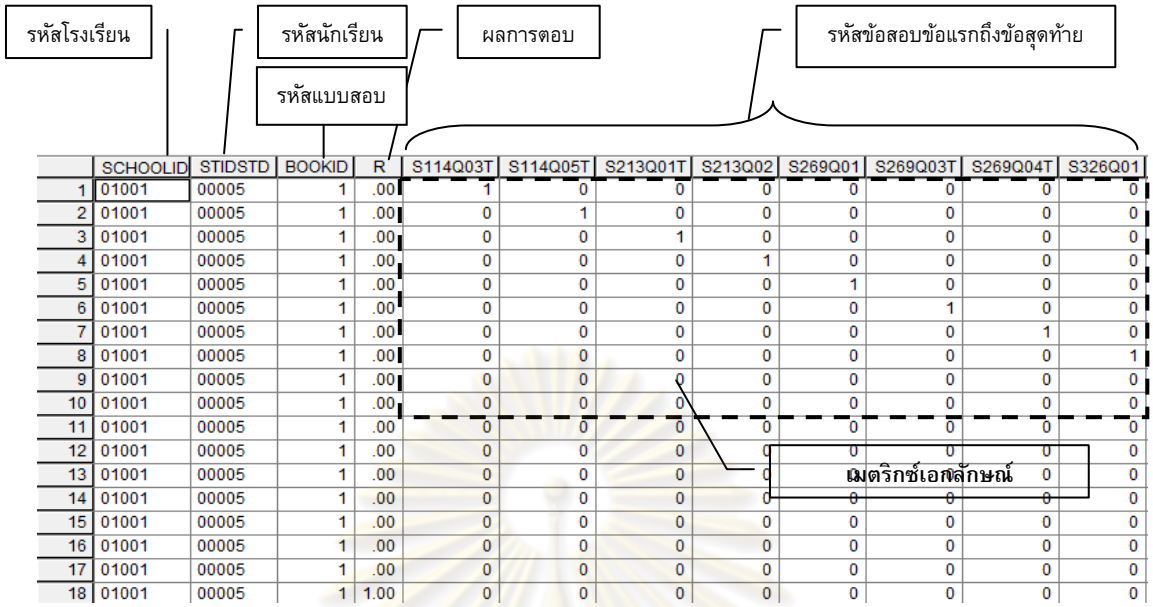
รหัสสถานศึกษา (SCHOOID) จะเรียงตามแนวแถว ตั้งแต่สถานศึกษาแรกไปถึงสถานศึกษาสุดท้าย

รหัสนักเรียน (STIDSTD) จะเรียงตามแนวแถวตั้งแต่นักเรียนคนแรกไปถึงคนสุดท้ายที่อยู่ในสถานศึกษานั้น แล้วเรียงเช่นเดียวกันต่อไปในสถานศึกษาถัดไป

ผลการตอบ (R) จะเรียงตามแนวแถวตั้งแต่ผลการตอบข้อแรกไปถึงข้อสุดท้ายของนักเรียนคนแรกโดยกำหนดให้ตอบถูกเป็น 1 และตอบผิดเป็น 0 นั้น แล้วเรียงเช่นเดียวกันต่อไปในนักเรียนคนถัดไป

รหัสข้อสอบ (S..) เรียงตามแนวสดมภ์ ตั้งแต่ข้อแรกไปถึงข้อสุดท้าย และจัดให้เป็นตัวแปรดัมมี่ เมื่อเขียนเรียงจากข้อแรกไปถึงข้อสุดท้ายจะอยู่ในรูปเมตริกซ์เอกลักษณ์

แสดงการจำลองไฟล์ข้อมูลดังแผนภาพที่ 12

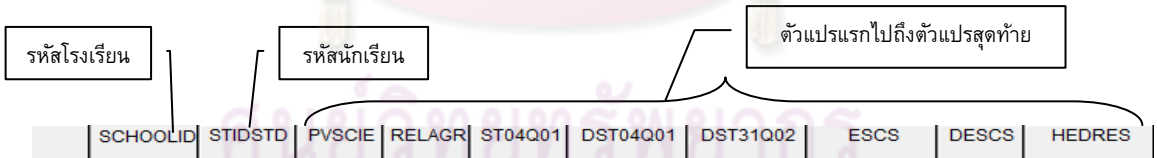


	SCHOOLID	STIDSTD	BOOKID	R	S114Q03T	S114Q05T	S213Q01T	S213Q02	S269Q01	S269Q03T	S269Q04T	S326Q01
1	01001	00005	1	.00	1	0	0	0	0	0	0	0
2	01001	00005	1	.00	0	1	0	0	0	0	0	0
3	01001	00005	1	.00	0	0	1	0	0	0	0	0
4	01001	00005	1	.00	0	0	0	1	0	0	0	0
5	01001	00005	1	.00	0	0	0	0	1	0	0	0
6	01001	00005	1	.00	0	0	0	0	0	1	0	0
7	01001	00005	1	.00	0	0	0	0	0	0	1	0
8	01001	00005	1	.00	0	0	0	0	0	0	0	1
9	01001	00005	1	.00	0	0	0	0	0	0	0	0
10	01001	00005	1	.00	0	0	0	0	0	0	0	0
11	01001	00005	1	.00	0	0	0	0	0	0	0	0
12	01001	00005	1	.00	0	0	0	0	0	0	0	0
13	01001	00005	1	.00	0	0	0	0	0	0	0	0
14	01001	00005	1	.00	0	0	0	0	0	0	0	0
15	01001	00005	1	.00	0	0	0	0	0	0	0	0
16	01001	00005	1	.00	0	0	0	0	0	0	0	0
17	01001	00005	1	.00	0	0	0	0	0	0	0	0
18	01001	00005	1	1.00	0	0	0	0	0	0	0	0

ภาพที่ 12 การจำลองไฟล์การวิเคราะห์ข้อมูลระดับข้อสอบ

<u>ไฟล์ที่ 2</u> : ระดับนักเรียน

เป็นไฟล์ที่รวมนักเรียนทั้งหมดที่เป็นกลุ่มตัวอย่างจากการทำแบบสอบฉบับที่ 1 ถึง 13 เข้าอยู่ในไฟล์เดียวกัน ประกอบด้วยรหัสสถานศึกษา รหัสนักเรียน และตัวแปรระดับนักเรียน โดยรหัสสถานศึกษา จะเรียงตามแนวแถว ตั้งแต่สถานศึกษาแรกไปถึงสถานศึกษาสุดท้าย รหัสนักเรียน จะเรียงตามแนวแถว ตั้งแต่นักเรียนคนแรกไปถึงคนสุดท้ายที่อยู่ในสถานศึกษานั้น และตัวแปรระดับนักเรียน เรียงตามแนวสดมภ์ ตั้งแต่ตัวแปรแรกไปถึงตัวแปรสุดท้าย ดังแผนภาพที่ 13

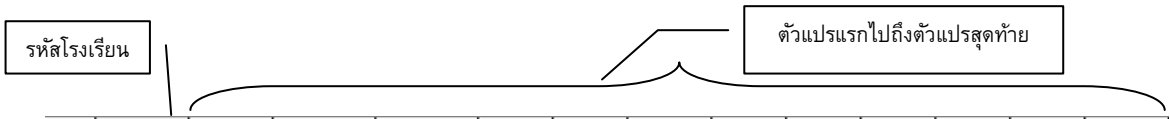


	SCHOOLID	STIDSTD	PVSCIE	RELAGR	ST04Q01	DST04Q01	DST31Q02	ESCS	DESCS	HEDRES
1	01001	00001	332.28	-.92	1	1	1	-1.3887	0	-1.1126
2	01001	00002	370.42	-1.00	1	1	1	-1.3519	0	-.4433
3	01001	00003	393.45	-1.00	1	1	1	-1.4431	0	-1.6271
4	01001	00004	384.78	-.75	1	1	0	-1.5432	0	-1.1126
5	01001	00005	454.25	-.83	1	1	1	-2.3245	0	-2.0600
6	01001	00006	443.34	-.75	1	1	0	-1.6074	0	-1.1126
7	01001	00007	392.05	-.92	1	1	1	-3.1376	0	-2.9761
8	01001	00008	366.31	-.75	1	1	1	-2.0841	0	-1.6271
9	01001	00009	334.89	-1.08	1	1	0	-1.5161	0	-1.1126
10	01001	00010	377.50	-.50	1	1	0	-3.3968	0	-2.4809
11	01001	00011	361.74	-1.08	1	1	0	.0824	1	-.4433
12	01001	00012	295.54	-1.17	1	1	1	-1.9970	0	-2.4809
13	01001	00013	389.25	-.92	1	1	0	-2.2837	0	-.4433
14	01001	00014	338.25	-1.17	2	0	0	-1.0241	0	-1.1126
15	01001	00015	336.10	-.25	2	0	1	-2.0091	0	-1.1126
16	01001	00016	368.46	-.75	2	0	0	-2.3490	0	-2.0600

ภาพที่ 13 การจำลองไฟล์การวิเคราะห์ข้อมูลระดับนักเรียน

<u>ไฟล์ที่ 3</u> : ระดับสถานศึกษา

ประกอบด้วยรหัสสถานศึกษา และตัวแปรระดับสถานศึกษา โดยรหัสสถานศึกษาจะเรียงตามแนวแถว ตั้งแต่สถานศึกษาแรกไปถึงสถานศึกษาสุดท้าย รหัสนักเรียน จะเรียงตามแนวแถว ตั้งแต่นักเรียนคนแรกไปถึงคนสุดท้ายที่อยู่ในสถานศึกษานั้น และตัวแปรระดับนักเรียน เรียงตามแนวสดมภ์ ตั้งแต่ตัวแปรแรกไปถึงตัวแปรสุดท้าย ดังแผนภาพที่ 14



	SCHOOLID	TYPESCH	RTYPESCH	RSCHSIZE	DSCHL	DSCHM	SC02Q01	PUBLIC	TYPE1	TYPE2	TYPE3	TYPE4	SC07Q01
1	01001	1	1	2	0	1	1	1	1	0	0	0	3
2	01002	1	1	3	1	0	1	1	1	0	0	0	2
3	02001	2	1	3	1	0	1	1	1	0	0	0	4
4	02002	2	1	3	1	0	1	1	1	0	0	0	2
5	02003	2	1	3	1	0	1	1	1	0	0	0	3
6	02004	2	1	3	1	0	1	1	1	0	0	0	2
7	02005	2	1	3	1	0	1	1	1	0	0	0	3
8	02006	2	1	2	0	1	1	1	1	0	0	0	2
9	02007	2	1	2	0	1	1	1	1	0	0	0	2
10	02008	2	1	1	0	0	1	1	1	0	0	0	1
11	02009	2	1	1	0	0	1	1	1	0	0	0	1
12	02010	2	1	2	0	1	1	1	1	0	0	0	3
13	02011	2	1	3	1	0	1	1	1	0	0	0	3
14	02012	2	1	3	1	0	1	1	1	0	0	0	3
15	02013	2	1	3	1	0	1	1	1	0	0	0	4
16	02014	2	1	3	1	0	1	1	1	0	0	0	5

ภาพที่ 14 การจำลองไฟล์การวิเคราะห์ข้อมูลระดับสถานศึกษา

1.2 การประมาณค่าความสามารถของนักเรียนและค่าความยากของข้อสอบ

การประมาณค่าความสามารถของนักเรียนและค่าความยากข้อสอบ ผู้วิจัยประยุกต์ใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น (HGLM) ด้วยการวิเคราะห์ 3 ระดับ มีรายละเอียดดังนี้

<u>การวิเคราะห์ระดับที่ 1</u> ระดับข้อสอบ

เป็นการวิเคราะห์อิทธิพลของข้อสอบที่สอดแทรกอยู่ในนักเรียนแต่ละคน และในแต่ละสถานศึกษา ในการวิเคราะห์จะนำตัวแปรดัมมี่ของข้อสอบข้อที่ 1 ถึง q ($x_1, x_2, x_3, \ldots, x_q$) เข้ามาวิเคราะห์โดยข้อสอบหนึ่งข้อจะเป็นข้ออ้างอิง(reference item) ซึ่งในการวิจัยนี้จะใช้ข้อสุดท้ายเป็นข้ออ้างอิง แสดงสมการวิเคราะห์ดังนี้

$$\eta_{ijm} = \beta_{0jm} + \beta_{1jm} x_{1ijm} + \beta_{2jm} x_{2ijm} + \ldots + \beta_{kjm} x_{kijm}$$

$$= \beta_{0j} + \sum_{q=1}^{k} \beta_{qim} x_{qijm}$$

<u>การวิเคราะห์ระดับที่ 2</u> ระดับนักเรียน

เป็นการวิเคราะห์เพื่อหาค่าเศษเหลือของค่าเฉลี่ยของโอกาสการตอบข้อสอบถูกของนักเรียนแต่ละคนที่อยู่ในสถานศึกษา (r_{0jm}) ซึ่งแสดงถึงค่าความสามารถของนักเรียน j ที่อยู่ในสถานศึกษา m แสดงสมการได้ดังนี้

$$\beta_{0jm} = \gamma_{00m} + r_{0jm}$$
$$\beta_{1jm} = \gamma_{10m}$$
$$\beta_{2jm} = \gamma_{20m}$$
$$\vdots$$
$$\beta_{qjm} = \gamma_{q0m}$$

<u>การวิเคราะห์ระดับที่ 3</u> ระดับสถานศึกษา

เป็นการวิเคราะห์เพื่อหาค่าเศษเหลือของ γ_{00m} หรือค่าที่เป็นส่วนเบี่ยงเบนของโอกาสในการตอบเศษเหลือของโอกาสการตอบข้อสอบถูก ซึ่งแสดงถึงค่าความสามารถของนักเรียนในสถานศึกษาที่ m(U_{00m}) แสดงสมการได้ดังนี้

$$\gamma_{00m} = \pi_{000} + u_{00m}$$
$$\gamma_{10m} = \pi_{100}$$
$$\gamma_{20m} = \pi_{200}$$
$$\vdots$$
$$\gamma_{q0m} = \pi_{q00}$$

จากสมการทั้ง 3 ระดับ จะได้ค่าพารามิเตอร์แต่ละค่าได้ดังนี้

ค่าความสามารถของนักเรียน เท่ากับ $u_{00m} + r_{0jm}$

ค่าความยากของข้อสอบ เท่ากับ $(-\pi_{q00} - \pi_{000})$

1.3 การวิเคราะห์หาอิทธิพลของตัวแปรภายนอกต่อโอกาสการตอบข้อสอบได้ถูก

เป็นการวิเคราะห์เพื่อหาผลของตัวแปรระดับนักเรียนที่มีต่อโอกาสการตอบข้อสอบถูกในการวิเคราะห์ระดับที่ 2 ซึ่งเป็นการตรวจสอบหาข้อสอบที่ทำหน้าที่ต่างกันเนื่องจากนักเรียนที่มีคุณลักษณะต่างกัน และใช้ตัวแปรระดับสถานศึกษาอธิบายผลของโอกาสการตอบข้อสอบถูกในการวิเคราะห์ระดับที่ 3 ซึ่งในการวิเคราะห์ในส่วนนี้ประกอบด้วยการวิเคราะห์ข้อสอบที่ตรวจให้คะแนนสองค่า และข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า เขียนสมการในการวิเคราะห์ดังนี้

1) การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบที่ตรวจให้คะแนนสองค่า

มีสมการการวิเคราะห์ดังนี้

<u>การวิเคราะห์ระดับที่ 1</u> ระดับข้อสอบ

เป็นการวิเคราะห์อิทธิพลของข้อสอบที่สอดแทรกอยู่ในผู้สอบแต่ละคน ในแต่ละสถานศึกษา ในการวิเคราะห์จะนำตัวแปรดัมมี่ของข้อสอบข้อที่ 1 ถึง q (x_1, x_2, x_3, …, x_q) และมีข้อสอบหนึ่งข้อเป็นข้ออ้างอิง แสดงสมการดังนี้

$$\eta_{ijm} = \beta_{0jm} + \beta_{1jm} x_{1ijm} + \beta_{2jm} x_{2ijm} + ... + \beta_{kjm} x_{kijm}$$
$$= \beta_{0j} + \sum_{q=1}^{k} \beta_{qim} x_{qijm}$$

การวิเคราะห์ระดับที่ 2 ระดับนักเรียน

เป็นการวิเคราะห์ที่นำตัวแปรคุณลักษณะของกลุ่มผู้สอบ (กำหนดเป็นกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ) เข้าไปในสมการ (G_{jm}) เพื่อทำนายโอกาสในการตอบข้อสอบได้ถูกต้อง โดยค่าอิทธิพลของข้อสอบรายข้อของผู้สอบ ($\gamma_{11m}, \gamma_{21m}, \gamma_{31m}, \ldots, \gamma_{q1m}$) จะเป็นตัวบ่งชี้ถึงการทำหน้าที่ต่างกันของข้อสอบข้อที่ 1 ถึง q แสดงสมการได้ดังนี้

$$\beta_{0jm} = \gamma_{00m} + \gamma_{01m}G_{jm} + r_{0jm}$$
$$\beta_{1jm} = \gamma_{10m} + \gamma_{11m}G_{jm}$$
$$\beta_{2jm} = \gamma_{20m} + \gamma_{21m}G_{jm}$$
$$\vdots$$
$$\beta_{qjm} = \gamma_{q0m} + \gamma_{q1m}G_{jm}$$

การวิเคราะห์ระดับที่ 3 ระดับสถานศึกษา

เป็นขั้นตอนการวิเคราะห์ที่นำตัวแปรทำนายระดับสถานศึกษาเข้าไปในสมการ (SCH_m) เพื่อทำนายโอกาสในการตอบถูกของผู้สอบที่สอดแทรกในสถานศึกษาที่มีคุณลักษณะต่างกัน แสดงสมการได้ดังนี้

$$\gamma_{00m} = \pi_{000} + \pi_{001}SCH_m + u_{00m}$$
$$\gamma_{01m} = \pi_{010} + \pi_{011}SCH_m + u_{01m}$$
$$\gamma_{10m} = \pi_{100}$$
$$\gamma_{20m} = \pi_{200}$$
$$\vdots$$
$$\gamma_{q0m} = \pi_{q00}$$

2) การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า

มีสมการการวิเคราะห์ดังนี้

การวิเคราะห์ระดับที่ 1 ระดับข้อสอบ

การวิเคราะห์ DIF สำหรับการตอบสนองที่เป็นการจัดลำดับ m ($m = 1,\ldots, M$) ในข้อสอบแบบตรวจให้คะแนนมากกว่าสองค่าจำนวน r ข้อ ที่กำหนดให้ค่าอำนาจจำแนก $\alpha_i = 1$ สำหรับข้อสอบทุกข้อในกรณีที่เป็นโมเดลหนึ่งพารามิเตอร์ i สำหรับข้อสอบข้อ i ผู้สอบ j ในสถานศึกษา k

$$\eta_{mijk} = \log\left(\frac{\varphi^{*}_{mijk}}{1-\varphi^{*}_{mijk}}\right)$$
$$= \begin{cases} \beta_{ojk} + \beta_{1jk}X_{1ijk} + \ldots + \beta_{(r-1)jk}X_{(r-1)ijk} & ,m = 1 \\ \beta_{ojk} + \beta_{1jk}X_{1ijk} + \ldots + \beta_{(r-1)jk}X_{(r-1)ijk} + \delta_{2jk} + \ldots + \delta_{2jk} & ,1 < m \le M-1 \end{cases}$$

$$= \beta_{0jk} + \sum_{q=1}^{r-1} \beta_{qjk} X_{qijk} + \sum_{p=2}^{m} \delta_{pjk}$$

การวิเคราะห์ระดับที่ 2 ระดับนักเรียน

โมเดลนี้เป็นเพิ่มตัวแปรทำนายที่เป็นคุณลักษณะของกลุ่มผู้สอบ (G_{jk}) เพื่อใช้ในการอธิบายโอกาสในการตอบข้อสอบได้ถูกต้อง โดยค่าอิทธิพลของข้อสอบรายข้อของผู้สอบ (γ_{11k}, γ_{21k}, γ_{31k}, ..., γ_{q1k}) จะเป็นตัวบ่งชี้ถึงการทำหน้าที่ต่างกันของข้อสอบข้อที่ 1 ถึง q เขียนสมการได้ดังนี้

$$\begin{aligned}
\beta_{0jk} &= \gamma_{00k} + \gamma_{01k} G_{jk} + u_{0jk} \\
\beta_{1jk} &= \gamma_{10k} + \gamma_{11k} G_{jk} \\
&\vdots \\
\beta_{(r-1)jk} &= \gamma_{(r-1)0k} + \gamma_{(r-1)1k} G_{jk} \\
\delta_{2jk} &= \delta_{20k} + \delta_{21qk} G_{jk} \\
&\vdots \\
\delta_{mjk} &= \delta_{m0k} + \delta_{m1qk} G_{jk}
\end{aligned}$$

การวิเคราะห์ระดับที่ 3 ระดับสถานศึกษา

เป็นการนำตัวแปรระดับสถานศึกษาเข้ามาวิเคราะห์ (SCH_m) เพื่ออธิบายคุณลักษณะของสถานศึกษาต่อโอกาสการตอบข้อสอบได้ถูก เขียนสมการวิเคราะห์ ดังนี้

$$\begin{aligned}
\gamma_{00k} &= \pi_{000} + \gamma_{00k} \\
\gamma_{10k} &= \pi_{100} + \pi_{101}(SCH_k) \\
&\vdots \\
\gamma_{(r-1)0k} &= \pi_{(r-1)00} + \pi_{(r-1)01}(SCH_k) \\
\gamma_{01k} &= \pi_{010} \\
\gamma_{11k} &= \pi_{110} + \pi_{111}(SCH_k) + r_{11k} \\
&\vdots \\
\gamma_{(r-1)1k} &= \pi_{(r-1)10} + \pi_{(r-1)11}(SCH_k) + r_{(r-1)k} \\
\delta_{20k} &= \delta_{200} \\
&\vdots \\
\delta_{m0k} &= \delta_{m00} \\
\delta_{21qk} &= \delta_{21q0} \\
&\vdots \\
\delta_{m1qk} &= \delta_{m1q0}
\end{aligned}$$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

1.4 การตัดข้อสอบภายหลังการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ

การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบที่ตรวจให้คะแนนสองค่าในแบบสอบฉบับที่ 1 - 13 และข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่าในแบบสอบฉบับที่ 1, 2, 5, 11 และ 12 ส่วนแบบสอบฉบับที่ 3, 4, 6, 7, 8, 9, 10 และ 13 ไม่มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบที่ตรวจให้คะแนนมากกว่าสองค่า เนื่องจากมีข้อสอบน้อยกว่า 2 ข้อ ซึ่งเป็นข้อจำกัดของการใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น (HGLM) ที่จะต้องมีข้อสอบอ้างอิงหนึ่งข้อ ดังนั้นแบบสอบที่มีข้อสอบหนึ่งข้อจึงไม่วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ เมื่อผู้วิจัยตรวจสอบแล้วพบว่ามีข้อสอบที่ทำหน้าที่ต่างกันข้อสอบในแบบสอบแต่ละฉบับ ผู้วิจัยจะตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปบางข้อโดยพิจารณาตามเกณฑ์การตัดข้อสอบ แล้วประมาณค่าความสามารถของนักเรียนใหม่ภายหลังการตัดข้อสอบที่ทำหน้าที่ต่างกัน โดยการตัดข้อสอบที่ทำหน้าที่ต่างกันจะคำนึงถึงการรักษาโครงสร้างเดิมของแบบสอบ กล่าวคือหลังจากตัดข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบออกไปแล้วโครงสร้างของแบบสอบจะยังคงมีโครงสร้างเดิม ทั้งนี้แบบสอบที่มีจำนวนข้อสอบที่น้อยลงจะส่งผลให้ความเที่ยงของแบบสอบมีค่าน้อยลง (ศิริชัย กาญจนวาสี, 2548) และจากผลการศึกษาของรักชนก ยี่สุ่นศรี (2544) พบว่าการตัดข้อสอบออกจากแบบสอบในจำนวนที่มากขึ้นจะทำให้ค่าความเที่ยงของแบบสอบมีค่าลดลง นอกจากนี้ Clauser (1993) ได้แนะนำเกณฑ์ที่ใช้ในการตัดข้อสอบที่ทำหน้าที่ต่างกันควรตัดข้อสอบไม่เกินร้อยละ 20 (cited in Naraya & Swaminathan,1994) ดังนั้นผู้วิจัยจึงกำหนดจำนวนข้อสอบที่ตัดได้ในแต่ละสมรรถนะทางวิทยาศาสตร์โดยคำนึงถึงโครงสร้างเดิมของแบบสอบ และยึดตามเกณฑ์ที่ใช้ในการตัดข้อสอบ สรุปได้ดังตารางที่ 13

ตารางที่ 13 จำนวนข้อสอบที่ตัดได้ในแต่ละสมรรถนะทางวิทยาศาสตร์ตามเกณฑ์การตัดข้อสอบในแบบสอบฉบับที่ 1 - 13

ฉบับ	โครงสร้างก่อนตัด DIF				โครงสร้างหลังตัด DIF				จำนวนข้อสอบที่ตัดได้ (ข้อ)			
	ISI	EPS	USE	รวม	ISI	EPS	USE	รวม	ISI	EPS	USE	รวม
1	11 (18%)	33 (55%)	16 (27%)	**60**	9 (19%)	26 (54%)	13 (27%)	**48**	2	7	3	12
2	5 (18%)	12 (43%)	11 (39%)	**28**	4 (18%)	10 (43%)	9 (39%)	**23**	1	2	2	5
3	7 (22%)	17 (53%)	8 (25%)	**32**	6 (23%)	14 (54%)	6 (23%)	**26**	1	3	2	6
4	7 (21%)	18 (53%)	9 (26%)	**34**	6 (22%)	14 (52%)	7 (26%)	**27**	1	4	2	7
5	15 (24%)	24 (39%)	23 (37%)	**62**	12 (24%)	19 (39%)	18 (37%)	**49**	3	5	5	13
6	8 (25%)	15 (47%)	9 (28%)	**32**	6 (24%)	12 (48%)	7 (28%)	**25**	2	3	2	7
7	2 (13%)	8 (54%)	5 (33%)	**15**	2 (17%)	6 (50%)	4 (33%)	**12**	0	2	1	3

ตารางที่ 13 จำนวนข้อสอบที่ตัดได้ในแต่ละสมรรถนะทางวิทยาศาสตร์ตามเกณฑ์การตัดข้อสอบใน แบบสอบฉบับที่ 1 - 13(ต่อ)

ฉบับ	โครงสร้างก่อนตัด DIF				โครงสร้างหลังตัด DIF				จำนวนข้อสอบที่ตัดได้ (ข้อ)			
	ISI	EPS	USE	รวม	ISI	EPS	USE	รวม	ISI	EPS	USE	รวม
8	6 (24%)	10 (38%)	10 (38%)	**26**	5 (24%)	8 (38%)	8 (38%)	**21**	1	2	2	5
9	11 (38%)	11 (38%)	7 (24%)	**29**	9 (38%)	9 (38%)	6 (24%)	**24**	2	2	1	5
10	8 (29%)	12 (42%)	8 (29%)	**28**	6 (29%)	9 (42%)	6 (29%)	**21**	2	3	2	7
11	5 (17%)	12 (40%)	13 (43%)	**30**	4 (18%)	10 (41%)	10 (41%)	**24**	1	2	3	6
12	8 (25%)	14 (44%)	10 (31%)	**32**	6 (24%)	11 (44%)	8 (32%)	**25**	2	3	2	7
13	2 (13%)	8 (54%)	5 (33%)	**15**	2 (17%)	6 (50%)	4 (33%)	**12**	0	2	1	3

หมายเหตุ ISI แทน การระบุประเด็นทางวิทยาศาสตร์
EPS แทน การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
USE แทน การใช้ประจักษ์พยานทางวิทยาศาสตร์

1.5 การประมาณค่าความสามารถของนักเรียนภายหลังการตัดข้อสอบที่ทำหน้าที่ต่างกัน

การประมาณค่าความสามารถของผู้สอบภายหลังการตัดข้อสอบที่ทำหน้าที่ต่างกัน ผู้วิจัยประยุกต์ใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น (HGLM) ด้วยการวิเคราะห์ 3 ระดับ และใช้ค่า $u_{00m} + r_{0jm}$ เป็นค่าความสามารถของนักเรียน เช่นเดียวกันกับการประมาณค่าความสามารถของผู้สอบก่อนการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป โดยการวิเคราะห์จะนำเฉพาะข้อสอบที่ไม่พบการทำหน้าที่ต่างกันเข้ามาวิเคราะห์ จึงสรุปสมการดังนี้

<u>การวิเคราะห์ระดับที่ 1</u> ระดับข้อสอบ

$$\eta_{ijm} = \beta_{0jm} + \beta_{1jm} x_{1ijm} + \beta_{2jm} x_{2ijm} + \ldots + \beta_{kjm} x_{kijm}$$
$$= \beta_{0j} + \sum_{q=1}^{k} \beta_{qim} x_{qijm}$$

<u>การวิเคราะห์ระดับที่ 2</u> ระดับผู้สอบ

$$\beta_{0jm} = \gamma_{00m} + r_{0jm}$$
$$\beta_{1jm} = \gamma_{10m}$$
$$\beta_{2jm} = \gamma_{20m}$$
$$\vdots$$
$$\beta_{qjm} = \gamma_{q0m}$$

การวิเคราะห์ระดับที่ 3 ระดับสถานศึกษา

$$\gamma_{00m} = \pi_{000} + u_{00m}$$
$$\gamma_{10m} = \pi_{100}$$
$$\gamma_{20m} = \pi_{200}$$
$$\vdots$$
$$\gamma_{q0m} = \pi_{q00}$$

1.6 การแปลงค่าคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์

คะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ตามโครงการประเมินผลนักเรียน PISA จะใช้วิธีคำนวณแปลงค่าความสามารถของนักเรียนที่ได้จากการประมาณค่าจากผลการตอบสนองของข้อสอบตามแนวคิดของโมเดลราสซ์ ที่มีฟังก์ชันโลจิสติก (Logistic Function) และใช้ค่าความสามารถที่ได้มาแปลงเป็นคะแนนสเกลของ PISA (OECD, 2009) ซึ่งอยู่ในรูปคะแนนมาตรฐานที่มีค่าเฉลี่ยเท่ากับ 500 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 100 มีสมการดังนี้

$$\text{คะแนนสเกลของ PISA} = \left(\frac{Logit(\theta) - 0.1797}{1.0724}\right) \times 100 + 500$$

ในการวิจัยครั้งนี้ผู้วิจัยวิเคราะห์ข้อสอบเพื่อตรวจหาข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบด้วย HGLM-3L โดยสมการการวิเคราะห์ระดับที่ 1 เป็นระดับข้อสอบ ระดับที่ 2 เป็นระดับนักเรียน และระดับที่ 3 เป็นระดับสถานศึกษา ซึ่งการใช้ HGLM-3L จะสามารถประมาณค่าความสามารถของผู้สอบได้จากการวิเคราะห์ในระดับที่ 2 และระดับที่ 3 จากการศึกษาของ Kamata (1998) พบว่าการวิเคราะห์ข้อสอบด้วย HGLM-3L เป็นโมเดลการวิเคราะห์ที่สอดคล้องกับแนวคิดของโมเดลราสซ์ โดยค่าความสามารถของผู้สอบจากโมเดลราสซ์ (θ_j) มีความสอดคล้องกับค่าความสามารถหรือค่าเศษเหลือที่ได้จากการวิเคราะห์ HGLM-3L หรือผลรวมของค่าเศษเหลือจากผลการวิเคราะห์ในสมการระดับที่ 2 และค่าเศษเหลือจากผลการวิเคราะห์ในระดับที่ 3 ($u_{00m} + r_{0jm}$) ดังนั้นผู้วิจัยจึงใช้ค่าความสามารถจากการวิเคราะห์ข้อสอบด้วย HGLM-3L

คะแนนสเกลของการประเมินการรู้เรื่องด้านวิทยาศาสตร์ จึงมีสมการการแปลงความสามารถของนักเรียนเช่นเดียวกับการแปลงคะแนนสเกลของ PISA มีสมการดังนี้

$$\text{คะแนนสเกลของ PISA} = \left(\frac{HGLM(u_{00m} + r_{0jm}) - 0.1797}{1.0724}\right) \times 100 + 500$$

การแปลงคะแนนสเกลของการประเมินการรู้เรื่องด้านวิทยาศาสตร์ มีการแปลงคะแนนจากค่าความสามารถของนักเรียนสองครั้งคือ ความสามารถของนักเรียนจากแบบสอบก่อนการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป และความสามารถของนักเรียนหลังการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป แต่จากการที่นักเรียนได้ทำแบบสอบต่างกัน อาจจะทำให้คะแนนของผู้สอบแต่ละคน

ได้รับอิทธิพลจากแบบสอบ เนื่องจากแบบสอบมีความยากของข้อสอบแตกต่างกัน ดังนั้นในการแปลงค่าเป็นคะแนนสเกลของ PISA จะนำค่าอิทธิพลจากแบบสอบ (ดังตารางที่ 11) มาบวกเพิ่มเข้าไปกับคะแนนสเกลของ PISA ของนักเรียนที่ทำแบบสอบแต่ละฉบับ

จากผลการวิเคราะห์ข้อสอบเพื่อประมาณค่าความยากของข้อสอบโดย HGLM-3L ผู้วิจัยพบว่าค่าความยากเฉลี่ยของข้อสอบในแบบสอบแต่ละฉบับมีความสอดคล้องกับค่าอิทธิพลจากแบบสอบจากรายงานของ PISA นั่นคือ ถ้าค่าอิทธิพลมีค่าเป็นบวก แสดงว่าแบบสอบฉบับนั้นมีความยากเฉลี่ยของข้อสอบมากกว่าแบบสอบฉบับอื่น และค่าอิทธิพลจากแบบสอบที่เป็นลบ แสดงว่าแบบสอบฉบับนั้นมีความยากเฉลี่ยของข้อสอบง่ายกว่าแบบสอบฉบับอื่น และปริมาณของค่าอิทธิพลมีความแตกต่างกันขึ้นอยู่กับค่าความยากเฉลี่ยของข้อสอบแต่ละฉบับ ดังตารางที่ 14

ตารางที่ 14 ผลการเปรียบเทียบความยากเฉลี่ยของข้อสอบจากค่าอิทธิพลจากแบบสอบ และค่าความยากเฉลี่ยของข้อสอบในแบบสอบฉบับที่ 1 ถึง 13

ฉบับ	ค่าอิทธิพลจากแบบสอบ	ค่าความยากเฉลี่ยของข้อสอบ	การแปลผลความยากของข้อสอบ
1	–3.1	.309952	ง่าย
2	–20.0	.246914	ง่าย
3	–20.5	.244974	ง่าย
4	–6.3	.294758	ง่าย
5	1.6	.460216	ยาก
6	6.7	.499478	ยาก
7	–19.9	.249310	ง่าย
8	17.6	.540691	ยาก
9	0.2	.399658	ยาก
10	21.4	.593823	ยาก
11	12.1	.502530	ยาก
12	20.4	.589036	ยาก
13	–10.4	.216778	ง่าย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ขั้นที่ 2 การวิเคราะห์มูลค่าเพิ่มในโมเดลการประเมินคุณภาพการจัดการศึกษา

ผู้วิจัยประยุกต์ใช้โมเดลเชิงเส้นตรงแบบลดหลั่นด้วยการวิเคราะห์ 2 ระดับ สำหรับการวิเคราะห์มูลค่าเพิ่มในโมเดลการประเมินคุณภาพการจัดการศึกษา 4 โมเดล โดยใช้โปรแกรม HLM for Windows Version 6.03 มีการวิเคราะห์แต่ละระดับดังนี้

ระดับที่ 1 ระดับนักเรียน

เป็นการวิเคราะห์ข้อมูลของตัวแปรระหว่างนักเรียนภายในหน่วยของสถานศึกษา (ระหว่างนักเรียน i ภายในสถานศึกษา k) โดยการวิเคราะห์ในโมเดลการประเมินคุณภาพการจัดการศึกษาโมเดล 1 และ 3 จะไม่มีการนำตัวแปรระดับนักเรียนเข้ามาวิเคราะห์ ส่วนโมเดลการประเมินคุณภาพการจัดการศึกษาโมเดล 2 และ 4 จะนำตัวแปรระดับนักเรียนเข้ามาวิเคราะห์ มีสมการดังนี้

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

เมื่อ Y_{ik} = คะแนนสเกลของการประเมินการรู้เรื่องด้านวิทยาศาสตร์ของนักเรียนที่ i ภายในสถานศึกษา k

$B0_k$ = ค่าจุดตัดแกน (intercept) หรือค่าเฉลี่ยของ Y_{ik} เมื่อควบคุมผลของ X_{pij} แล้ว ซึ่งเป็นค่าที่แสดงถึงคะแนนเฉลี่ยของคะแนนสเกลของการประเมินการรู้เรื่องด้านวิทยาศาสตร์ของนักเรียนในสถานศึกษา k

Bp_k = ค่าความชัน (slope) หรือค่าสัมประสิทธิ์ถดถอยที่แสดงผลของ X_{pik} ต่อ Y_{ik} ซึ่งเป็นที่แสดงถึงค่าอิทธิพลของการเปลี่ยนแปลงของคะแนนสเกลของการประเมินการรู้เรื่องด้านวิทยาศาสตร์ของนักเรียนเมื่อควบคุมตัวแปรคุณลักษณะของนักเรียน 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 1 และ 3 จะไม่มีการนำตัวแปรระดับสถานศึกษาเข้ามาวิเคราะห์ ส่วนโมเดลการประเมินคุณภาพการจัดการศึกษาโมเดล 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 $G00$ = ค่าจุดตัดแกน(intercept) หรือค่าเฉลี่ยของ $B0_j$ เมื่อควบคุมผลของ Z_k แล้ว ซึ่งเป็นค่าแสดงถึงคะแนนเฉลี่ยของคะแนนสเกลของการประเมินการรู้เรื่องด้านวิทยาศาสตร์ของนักเรียนในสถานศึกษาทุกโรง

$G0_q$ = ค่าความชัน(slope) ค่าสัมประสิทธิ์ถดถอยที่แสดงผลของ Z_j ต่อ $B0_j$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ซึ่งแสดงถึงค่าอิทธิพลของการเปลี่ยนแปลงของคะแนนสเกลของการประเมินการรู้เรื่องด้านวิทยาศาสตร์ของนักเรียนเมื่อมีการควบคุมตัวแปรคุณลักษณะของสถานศึกษา Y_k

Z_k = ตัวแปรควบคุมระดับหน่วยงาน

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

สรุปสมการวิเคราะห์มูลค่าเพิ่มโมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่นด้วยการวิเคราะห์ 2 ระดับ ในโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา 4 โมเดล ดังนี้

ตารางที่ 15 สรุปสมการวิเคราะห์มูลค่าเพิ่มด้วยโมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น 2 ระดับ ในโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา 4 โมเดล

ระดับการวิเคราะห์	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
ระดับที่ 1	ตัวแปรนักเรียน : ไม่มี ตัวแปรตาม : คะแนนสเกลของการประเมินการรู้เรื่องด้านวิทยาศาสตร์ของนักเรียนที่***ไม่ตัดข้อ DIF*** สมการวิเคราะห์ : $Y_{ik} = B0_k + R_{ik}$	ตัวแปรนักเรียน : มี ตัวแปรตาม : คะแนนสเกลของการประเมินการรู้เรื่องด้านวิทยาศาสตร์ของนักเรียนที่ไม่ตัดข้อ DIF สมการวิเคราะห์ : $Y_{ik} = B0_k + B1_k*(X_{ik})_{ik} + R_{ik}$	ตัวแปรนักเรียน : ไม่มี ตัวแปรตาม : คะแนนสเกลของการประเมินการรู้เรื่องด้านวิทยาศาสตร์ของนักเรียนที่***ตัดข้อ DIF*** สมการวิเคราะห์ : $Y_{ik} = B0_k + R_{ik}$	ตัวแปรนักเรียน : มี ตัวแปรตาม : คะแนนสเกลของการประเมินการรู้เรื่องด้านวิทยาศาสตร์ของนักเรียนที่***ตัดข้อ DIF*** สมการวิเคราะห์ : $Y_{ik}=B0_k+B1_k*(X_{ik})_{ik} + R_{ik}$
ระดับที่ 2	ตัวแปรสถานศึกษา : ไม่มี ตัวแปรตาม : $B0_k$ สมการวิเคราะห์ : $B0_k = G00 + U0_k$ คะแนนมูลค่าเพิ่ม : $U0_k$	ตัวแปรสถานศึกษา : มี ตัวแปรตาม : $B0_k$, Bp_k สมการวิเคราะห์ : $B0_k=G00+G01*(Y_k) + U0_k$ $B1 = G10$ $B2 = G20$ $B3 = G30$ . . . $Bp_k = Gp0$ 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 : $B0_k=G00+G01*(Y_k) + U0_k$ $B1 = G10$ $B2 = G20$ $B3 = G30$ . . . $Bp_k = Gp0$ คะแนนมูลค่าเพิ่ม : $U0_k$

ขั้นที่ 3 การตรวจสอบผลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา

โมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาในการวิจัยครั้งนี้มี 4 โมเดล ผู้วิจัยมีการตรวจสอบผลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา โดยใช้ค่าจากการวิเคราะห์ ดังนี้

5.1 คะแนนมูลค่าเพิ่มจากโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา 4 โมเดล ซึ่งเป็นค่าเศษเหลือที่ได้จากผลการวิเคราะห์ระดับที่ 2 (r_{0k}) หรือผลต่างระหว่างค่าเฉลี่ยของคะแนนสเกลของการประเมินการรู้เรื่องด้านวิทยาศาสตร์ของนักเรียนที่วัดได้จริงกับค่าเฉลี่ยของคะแนนสเกลของการประเมินการรู้เรื่องด้านวิทยาศาสตร์ของนักเรียนที่คาดหวังซึ่งประมาณค่าจากชุดของตัวแปรที่นำมาควบคุมทั้งในระดับนักเรียนและสถานศึกษา

5.2 การจัดระดับคุณภาพ(Ranking) การจัดการศึกษาของสถานศึกษา โดยการเรียงระดับคะแนนมูลค่าเพิ่มจากคะแนนสูงไปคะแนนต่ำ จากผลการวิเคราะห์ในโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา 4 โมเดล

5.3 การจัดกลุ่มคุณภาพ(Rating) การจัดการศึกษาของสถานศึกษา โดยการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษา ผู้วิจัยแบ่งกลุ่มโดยใช้ตำแหน่งเปอร์เซ็นต์ไทล์ และจัดกลุ่มเป็น 3 กลุ่ม คือ

กลุ่มที่ 1 เป็นกลุ่มต่ำ มีตำแหน่งเปอร์เซ็นต์ไทล์ 1 – 24.99
กลุ่มที่ 2 เป็นกลุ่มปานกลาง มีตำแหน่งเปอร์เซ็นต์ไทล์ 25 – 74.99
กลุ่มที่ 3 เป็นกลุ่มสูง มีตำแหน่งเปอร์เซ็นต์ไทล์ 76 – 100

5.4 ค่าความสามารถในการอธิบายความแปรปรวนของตัวแปรตามด้วยตัวแปรทำนายหรือค่าสัมประสิทธิ์การทำนายในแต่ละโมเดล (R^2) มีสมการดังนี้

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

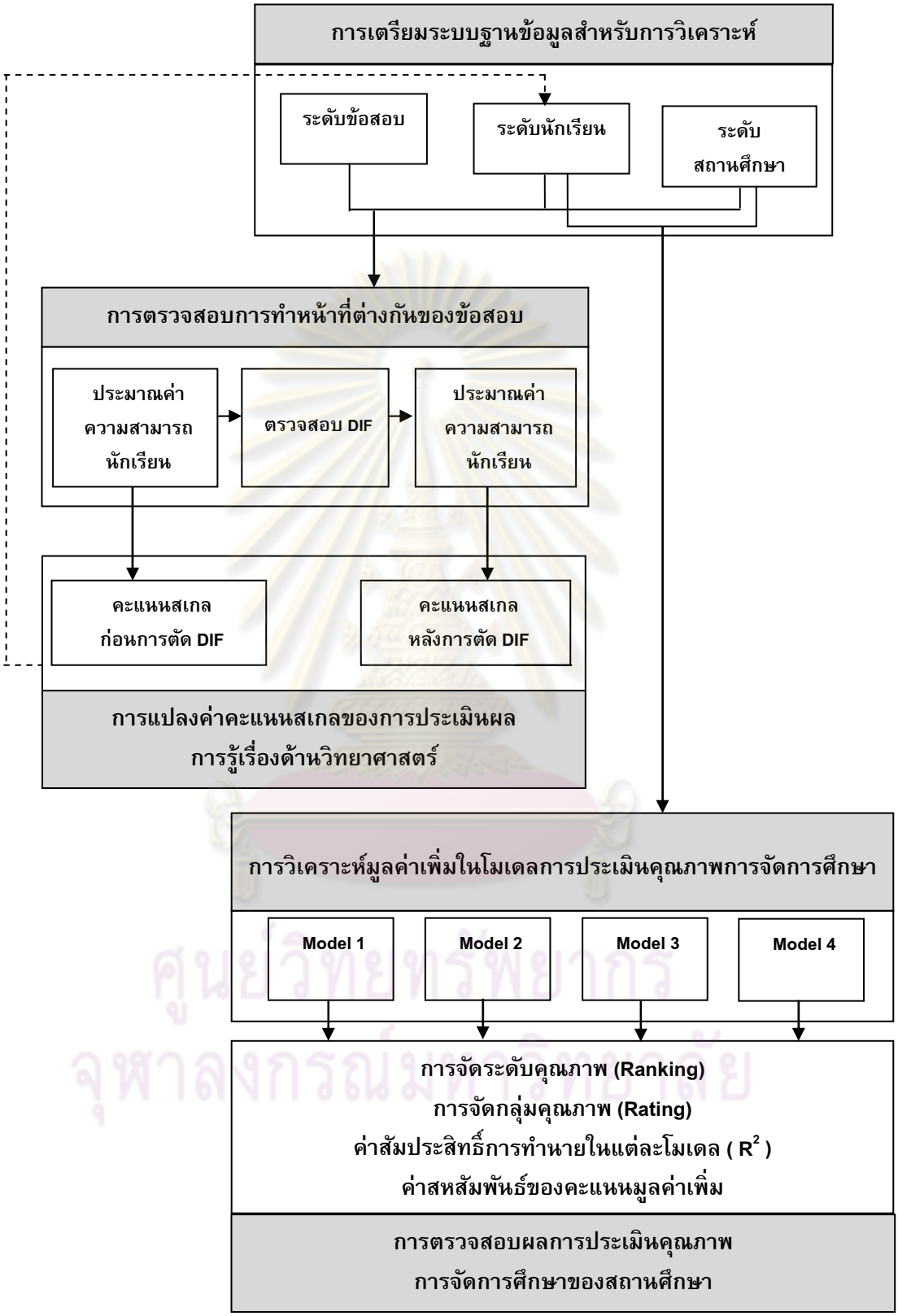
ความแปรปรวนของตัวแปรตามที่อธิบายได้ด้วยตัวแปรทำนายในแต่ละระดับนักเรียนและสถานศึกษา คำนวณได้จากผลคูณระหว่างความแปรปรวนที่อธิบายได้ตามโมเดลสมมติฐานกับความแปรปรวนทั้งหมดในระดับนั้นที่ได้จากโมเดลที่ไม่มีเงื่อนไขอย่างสมบูรณ์ ดังนี้

ความแปรปรวนที่อธิบายได้จากระดับที่ i = σ_i^2(explained) x σ_i^2

ความแปรปรวนทั้งหมดที่อธิบายได้โดยโมเดลทั้งสองระดับ เท่ากับ ผลรวมของความแปรปรวนที่อธิบายได้จากระดับที่ 1 กับระดับที่ 2

5.5 ค่าสหสัมพันธ์ของคะแนนมูลค่าเพิ่มของสถานศึกษาที่ได้จากการวิเคราะห์ในแต่ละโมเดล เพื่อใช้ค่าที่บ่งบอกถึงผลการตรวจสอบของคุณภาพการจัดการศึกษาที่มีความสอดคล้องกันในแต่ละโมเดล โดยใช้สหสัมพันธ์ตำแหน่งคะแนนของสเปียร์แมน (Spearman's Rank – Order Correlation)

สรุปขั้นตอนการดำเนินการวิจัย ได้ดังแผนภาพที่ 15



แผนภาพที่ 15 สรุปขั้นตอนการดำเนินการวิจัย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในการวิจัยครั้งนี้มีจุดประสงค์เพื่อประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน ข้อมูลที่ใช้ในการวิจัยเป็นข้อมูลจากโครงการประเมินผลนักเรียนระดับนานาชาติ ปี 2549 หรือ PISA ปี 2006 โดยการประยุกต์ใช้การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบและการวิเคราะห์มูลค่าเพิ่ม มีขั้นตอนการวิเคราะห์ที่สำคัญคือการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบภายใต้โมเดล HGLM และการวิเคราะห์มูลค่าเพิ่มภายใต้โมเดล HLM

ผู้วิจัยได้กำหนดการนำเสนอผลการวิเคราะห์ข้อมูลเพื่อตอบคำถามตามวัตถุประสงค์การวิจัย โดยแบ่งการเสนอผลการวิเคราะห์ข้อมูลเป็น 4 ตอน ดังนี้

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลพื้นฐาน การตรวจสอบเงื่อนไขการวิเคราะห์พหุระดับและการตรวจสอบความเป็นเอกมิติของแบบสอบ

ตอนที่ 2 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแต่ละฉบับ และการประมาณค่าความสามารถของผู้สอบ และค่าความยากของข้อสอบ

ตอนที่ 3 ผลการวิเคราะห์ตามโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน 4 โมเดล

ตอนที่ 4 สรุปผลการวิเคราะห์เพื่อตอบคำถามการวิจัย

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลพื้นฐาน การตรวจสอบเงื่อนไขของการวิเคราะห์พหุระดับและการตรวจสอบความเป็นเอกมิติของแบบสอบ

การนำเสนอผลการวิเคราะห์ข้อมูลพื้นฐาน แบ่งเป็น 3 ตอนคือ 1.1) ข้อมูลพื้นฐานของกลุ่มตัวอย่าง เป็นการเสนอคุณลักษณะพื้นฐานของนักเรียนและสถานศึกษา เพื่อแสดงลักษณะการแจกแจงของข้อมูล ประกอบด้วยค่าสถิติเชิงบรรยายได้แก่ ค่าความถี่ ร้อยละ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน ค่าต่ำสุด ค่าสูงสุด ค่าความเบ้ และค่าความโด่ง 1.2) การตรวจสอบเงื่อนไขของการวิเคราะห์พหุระดับ เพื่อตรวจสอบข้อตกลงเบื้องต้นเกี่ยวกับความสัมพันธ์ของตัวแปรต้นและลักษณะการแจกแจงของความคลาดเคลื่อนก่อนการใช้โมเดลพหุระดับ และการวิเคราะห์ถดถอย และ 1.3) การตรวจสอบความเป็นเอกมิติของแบบสอบ ตามข้อตกลงเบื้องต้นของโมเดลราสช์ด้วยการวิเคราะห์องค์ประกอบ

1.1 ข้อมูลพื้นฐานของกลุ่มตัวอย่าง

การวิเคราะห์ข้อมูลพื้นฐานของกลุ่มตัวอย่าง ผู้วิจัยนำตัวแปรคุณลักษณะพื้นฐานของนักเรียนและสถานศึกษาที่มีอยู่ในฐานข้อมูลของโครงการการประเมินผลนักเรียน PISA 2006 มาวิเคราะห์เพื่อให้เห็นถึงลักษณะโครงการสร้างของกลุ่มตัวอย่าง

จากการวิเคราะห์ข้อมูลพื้นฐานของกลุ่มตัวอย่างที่เป็นตัวแปรจัดประเภท จำแนกเป็นระดับนักเรียนและระดับสถานศึกษา ในระดับนักเรียน พบว่านักเรียนเป็นเพศชายคิดเป็นร้อยละ 42.1 และเป็นเพศหญิงร้อยละ 57.9 นักเรียนส่วนใหญ่ไม่มีเวลาเรียนพิเศษวิทยาศาสตร์นอกสถานศึกษา คิดเป็นร้อยละ 56.2 นักเรียนมีเวลาเรียนพิเศษวิทยาศาสตร์นอกสถานศึกษาน้อยกว่า 2 ชั่วโมง และ 2 – 4 ชั่วโมงร้อยละ 22.3 และ 15.6 ตามลำดับ เวลาที่ใช้ในการศึกษาวิทยาศาสตร์ด้วยตนเอง พบว่า นักเรียนส่วนใหญ่ใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเองน้อยกว่า 2 ชั่วโมง คิดเป็นร้อยละ 56.8 รองลงมาคือใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง 2 - 4 ชั่วโมง คิดเป็นร้อยละ 25.0

ในระดับสถานศึกษา พบว่า ขนาดของสถานศึกษาเรียงตามลำดับขนาดใหญ่ กลาง และเล็ก คิดเป็นร้อยละ 43.4 39.6 และ 17.0 ตามลำดับ สถานศึกษาส่วนใหญ่เป็นสถานศึกษาของรัฐ คิดเป็นร้อยละ 88.2 และเป็นสถานศึกษาเอกชนร้อยละ 11.8 มีตำแหน่งที่ตั้งสถานศึกษาเรียงตามหมู่บ้าน เมืองเล็ก เมือง เมืองใหญ่ คิดเป็นร้อยละ 24.1 20.8 25.5 และ 20.3 ตามลำดับ สถานศึกษาส่วนใหญ่แบ่งขนาดห้องเรียนให้มีจำนวนนักเรียน 36 – 40 คน และนักเรียน 41 – 45 คน คิดเป็นร้อยละ 29.2 และ 26.4 ตามลำดับ แรงกดดันของผู้ปกครองต่อสถานศึกษา พบว่าส่วนใหญ่ผู้ปกครองส่วนน้อยมีแรงกดดันต่อสถานศึกษา คิดเป็นร้อยละ 44.3 การขาดแคลนครูวิทยาศาสตร์ พบว่าส่วนใหญ่ขาดแคลนน้อย คิดเป็นร้อยละ 44.8 และการขาดแคลนครูที่มีคุณวุฒิ พบว่าส่วนใหญ่ขาดแคลนน้อย คิดเป็น ร้อยละ 44.8 แสดงดังตารางที่ 16

ตารางที่ 16 ความถี่ ร้อยละของตัวแปรจัดประเภทจำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา

คุณลักษณะของนักเรียนและสถานศึกษา	ความถี่	ร้อยละ
นักเรียน		
1. เพศ		
1.1 ชาย	2,608	42.1
1.2 หญิง	3,584	57.9
2. เวลาที่เรียนพิเศษวิทยาศาสตร์นอกสถานศึกษา / สัปดาห์		
2.1 ไม่มีเวลา	3,481	56.2
2.2 น้อยกว่า 2 ชั่วโมง	1,380	22.3
2.3 เวลา 2 – 4 ชั่วโมง	967	15.6
2.4 เวลา 4 – 6 ชั่วโมง	260	4.2
2.5 มากกว่า 6 ชั่วโมง	104	1.7
3. เวลาที่ใช้ในการศึกษาวิทยาศาสตร์ด้วยตนเอง / สัปดาห์		
3.1 ไม่มีเวลา	588	9.5
3.2 น้อยกว่า 2 ชั่วโมง	3,514	56.8
3.3 เวลา 2 – 4 ชั่วโมง	1,546	25.0
3.4 เวลา 4 – 6 ชั่วโมง	381	6.2
3.5 มากกว่า 6 ชั่วโมง	163	2.6
รวม	6,192	100.0

ตารางที่ 16 ความถี่ ร้อยละของตัวแปรจัดประเภทจำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา (ต่อ)

คุณลักษณะของนักเรียนและสถานศึกษา	ความถี่	ร้อยละ
สถานศึกษา		
1. ขนาดสถานศึกษา		
1.1 ขนาดเล็ก (นักเรียน 1 – 499 คน)	36	17.0
1.2 ขนาดกลาง (นักเรียน 500 – 1,500 คน)	84	39.6
1.3 ขนาดใหญ่ (นักเรียนมากกว่า 1,500 คน)	92	43.4
2. ประเภทของสถานศึกษา		
2.1 สถานศึกษาของรัฐ	187	88.2
2.2 สถานศึกษาเอกชน	25	11.8
3. สังกัดสถานศึกษา		
3.1 สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน(สพฐ.)	134	63.2
3.2 สำนักงานการศึกษาเอกชน (สช.)	25	11.8
3.3 สำนักงานคณะกรรมการการอุดมศึกษา(สาธิต)	15	7.1
3.4 สำนักบริหารการศึกษาท้องถิ่นและกรุงเทพมหานคร(กศท.และ กทม.)	30	14.2
3.5 อื่นๆ (สถานศึกษาและวิทยาลัยอาชีวศึกษา (อศ.))	8	3.8
4. ตำแหน่งที่ตั้งสถานศึกษา		
4.1 หมู่บ้าน	51	24.1
4.2 เมืองเล็ก	44	20.8
4.3 เมือง	54	25.5
4.4 เมืองใหญ่	43	20.3
4.5 เมืองใหญ่มาก	20	9.4
5. ขนาดห้องเรียน		
5.1 นักเรียนน้อยกว่าหรือเท่ากับ 15 คน	5	2.4
5.2 นักเรียน 16-20 คน	4	1.9
5.3 นักเรียน 21-25 คน	12	5.7
5.4 นักเรียน 26-30 คน	14	6.6
5.5 นักเรียน 31-35 คน	20	9.4
5.6 นักเรียน 36-40 คน	62	29.2
5.7 นักเรียน 41-45 คน	56	26.4
5.8 นักเรียน 46-50 คน	28	13.2
5.9 นักเรียนมากกว่า 50 คน	11	5.2
6. แรงกดดันของผู้ปกครองต่อสถานศึกษา		
6.1 มีจำนวนมาก	59	27.8
6.2 มีส่วนน้อย	94	44.3
6.3 ไม่พบเป็นส่วนใหญ่	59	27.8

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 16 ความถี่ ร้อยละของตัวแปรจัดประเภทจำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา (ต่อ)

คุณลักษณะของนักเรียนและสถานศึกษา	ความถี่	ร้อยละ
7. การขาดแคลนครูวิทยาศาสตร์		
7.1 ไม่ขาดแคลน	72	34.0
7.2 ขาดแคลนน้อย	95	44.8
7.3 ขาดแคลนเป็นบางส่วน	31	14.6
7.4 ขาดแคลนจำนวนมาก	14	6.6
8. การขาดแคลนครูที่มีคุณวุฒิ		
8.1 ไม่ขาดแคลน	37	17.5
8.2 ขาดแคลนน้อย	95	44.8
8.3 ขาดแคลนเป็นบางส่วน	58	27.4
8.4 ขาดแคลนจำนวนมาก	22	10.4
รวม	212	100.0

ผลการวิเคราะห์คุณลักษณะพื้นฐานของกลุ่มตัวอย่างที่เป็นตัวแปรต่อเนื่อง จำแนกเป็นระดับนักเรียนและระดับสถานศึกษา **ระดับนักเรียน** พบว่าตัวแปรคุณลักษณะของนักเรียนมีความแตกต่างของค่าของตัวแปรหรือมีการกระจายของข้อมูลค่อนข้างมาก ได้แก่ เศรษฐานะของครอบครัวมีค่าต่ำสุด -3.81 และค่าสูงสุด 2.54 มีค่าเฉลี่ย -1.23 ส่วนเบี่ยงเบนมาตรฐาน 1.18 มีลักษณะการแจกแจงเป็นแบบเบ้ขวา (skewness = .57) และค่อนข้างแบน (kurtosis = -.55) แหล่งทรัพยากรที่บ้านมีค่าต่ำสุด -3.83 และค่าสูงสุด .59 มีค่าเฉลี่ย -1.16 ส่วนเบี่ยงเบนมาตรฐาน 1.02 มีลักษณะการแจกแจงเป็นแบบเบ้ขวา (skewness = .25) และค่อนข้างแบน (kurtosis = -.53) นั่นคือนักเรียนส่วนใหญ่มีเศรษฐานะของครอบครัวและแหล่งทรัพยากรที่บ้านค่อนข้างต่ำและการกระจายของข้อมูลมีลักษณะค่อนข้างแบน ส่วนความมั่งคั่งของครอบครัวมีค่าต่ำสุด -4.29 และค่าสูงสุด 1.79 มีค่าเฉลี่ย -1.14 ส่วนเบี่ยงเบนมาตรฐาน 1.08 มีลักษณะการแจกแจงเป็นแบบเบ้ซ้าย (skewness = -.03) และแบน (kurtosis = -.07) นั่นคือนักเรียนส่วนใหญ่มีความมั่งคั่งของครอบครัวค่อนข้างสูงและมีการกระจายของข้อมูลค่อนข้างมาก ผลสัมฤทธิ์ทางการเรียนที่ผ่านมามีค่าต่ำสุด -3.08 และค่าสูงสุด 1.75 มีค่าเฉลี่ย .01 ส่วนเบี่ยงเบนมาตรฐาน .51 มีลักษณะการแจกแจงเป็นแบบเบ้ซ้าย (skewness = -.47) และค่อนข้างโด่ง (kurtosis = .65) แสดงว่านักเรียนส่วนใหญ่มีผลสัมฤทธิ์ทางการเรียนที่ผ่านมาสูง ความสนใจในการเรียนรู้วิทยาศาสตร์มีค่าต่ำสุด -3.14 และค่าสูงสุด 3.28 มีค่าเฉลี่ย .76 ส่วนเบี่ยงเบนมาตรฐาน .79 มีลักษณะการแจกแจงเป็นแบบเบ้ขวา (skewness = .14) และโด่ง (kurtosis = 2.83) แสดงว่านักเรียนส่วนใหญ่มีความสนใจในการเรียนรู้วิทยาศาสตร์ค่อนข้างต่ำ ความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์มีค่าต่ำสุด -3.76 และค่าสูงสุด 3.22 มีค่าเฉลี่ย .07 ส่วนเบี่ยงเบนมาตรฐาน .73 ลักษณะการแจกแจงเป็นแบบเบ้ขวา (skewness = .09) และโด่ง (kurtosis = 2.19) แสดงว่านักเรียนส่วนใหญ่มี

ความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์ค่อนข้างต่ำ ความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์มีค่าต่ำสุด -2.36 และค่าสูงสุด 2.24 มีค่าเฉลี่ย .66 ส่วนเบี่ยงเบนมาตรฐาน .76 ลักษณะการแจกแจงเป็นแบบเบ้ซ้าย (skewness = -.14) และค่อนข้างโด่ง (kurtosis = .38) แสดงว่านักเรียนส่วนใหญ่มีความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์ค่อนข้างสูง ซึ่งเป็นที่น่าสังเกตว่าความแตกต่างกันกับความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์ ความเพลิดเพลินทางวิทยาศาสตร์มีค่าต่ำสุด -2.15 และค่าสูงสุด 2.05 มีค่าเฉลี่ย .71 ส่วนเบี่ยงเบนมาตรฐาน .67 ลักษณะการแจกแจงเป็นแบบเบ้ซ้าย (skewness = -.16) และค่อนข้างโด่ง (kurtosis = .09) แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์มีค่าต่ำสุด -2.10 และค่าสูงสุด 1.82 มีค่าเฉลี่ย .70 ส่วนเบี่ยงเบนมาตรฐาน .65 ลักษณะการแจกแจงเป็นแบบเบ้ขวา (skewness = .04) และค่อนข้างแบน (kurtosis = -.08) การเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์มีค่าต่ำสุด -3.66 และค่าสูงสุด 2.18 มีค่าเฉลี่ย .77 ส่วนเบี่ยงเบนมาตรฐาน .91 มีลักษณะการแจกแจงเป็นแบบเบ้ซ้าย (skewness = -.01) และค่อนข้างแบน (kurtosis = -.68) แสดงว่านักเรียนส่วนใหญ่มีความเพลิดเพลินทางวิทยาศาสตร์ค่อนข้างสูงและเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์ แต่มีแรงจูงใจภายนอกในการเรียนวิทยาศาสตร์ต่ำ

ระดับสถานศึกษา พบว่างบประมาณของรัฐมีร้อยละต่ำสุด .00 และร้อยละสูงสุด 100.00 มีค่าเฉลี่ย 75.45 ส่วนเบี่ยงเบนมาตรฐาน 26.53 ลักษณะการแจกแจงเป็นแบบเบ้ซ้าย (skewness = -.90) และค่อนข้างแบน (kurtosis = -.38) งบประมาณจากหน่วยงานภายนอกมีร้อยละต่ำสุด .00 และร้อยสูงสุด 60.00 มีค่าเฉลี่ย 8.11 ส่วนเบี่ยงเบนมาตรฐาน 12.20 มีลักษณะการแจกแจงเป็นแบบเบ้ขวา (skewness = 2.04) และโด่ง(kurtosis = 4.24) แสดงให้เห็นว่าสถานศึกษาได้รับงบประมาณของรัฐเป็นส่วนใหญ่ และได้รับงบประมาณจากหน่วยงานภายนอกเป็นส่วนน้อย ดัชนีการขาดแคลนครูมีค่าต่ำสุด -1.05 และค่าสูงสุด 3.61 มีค่าเฉลี่ย .64 ส่วนเบี่ยงเบนมาตรฐาน .84 มีลักษณะการแจกแจงเป็นแบบเบ้ซ้าย (skewness = -.04) และค่อนข้างโด่ง (kurtosis = .64) แสดงว่าสถานศึกษาส่วนใหญ่มีดัชนีการขาดแคลนครูค่อนข้างสูง ดัชนีทรัพยากรการเรียนมีค่าต่ำสุด -3.43 และค่าสูงสุด 2.13 มีค่าเฉลี่ย -.63 ส่วนเบี่ยงเบนมาตรฐาน 1.09 มีลักษณะการแจกแจงเป็นแบบเบ้ขวา (skewness = .13) และค่อนข้างโด่ง (kurtosis = .62) แสดงว่าสถานศึกษาส่วนใหญ่มีดัชนีทรัพยากรการเรียนต่ำ สัดส่วนนักเรียนต่อครูมีค่าต่ำสุด 9.22 และค่าสูงสุด 36.46 มีค่าเฉลี่ย 22.07 ส่วนเบี่ยงเบนมาตรฐาน 5.55 มีลักษณะการแจกแจงเป็นแบบเบ้ซ้าย (skewness = -.17) และค่อนข้างแบน (kurtosis = -.09) แสดงว่าส่วนใหญ่ครูหนึ่งคนรับผิดชอบนักเรียนมากกว่า 22 คน ดัชนีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์มีค่าต่ำสุด -2.26 และค่าสูงสุด 1.63 มีค่าเฉลี่ย 1.35 ส่วนเบี่ยงเบนมาตรฐาน .59 มีลักษณะการแจกแจงเป็นแบบเบ้ซ้าย (skewness = -2.54) และโด่ง (kurtosis = 8.22) แสดงว่าสถานศึกษาส่วนใหญ่มีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์ค่อนข้างสูง ดัชนีความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากรมีค่าต่ำสุด -1.10 และค่าสูงสุด 2.01 มีค่าเฉลี่ย .21 ส่วนเบี่ยงเบนมาตรฐาน 1.07 มีลักษณะการแจกแจงเป็นแบบเบ้ขวา (skewness = .72) และแบน (kurtosis = -.98) ดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผลมีค่า

ต่ำสุด -1.06 และค่าสูงสุด 1.26 มีค่าเฉลี่ย .82 ส่วนเบี่ยงเบนมาตรฐาน .73 มีลักษณะการแจกแจงเป็นแบบเบ้ซ้าย (skewness = -1.29) และค่อนข้างโด่ง (kurtosis = .04) แสดงให้เห็นว่าสถานศึกษามีความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากรส่วนใหญ่อยู่ในระดับต่ำ แต่มีดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผลสูง แสดงผลดังตารางที่ 17

ตารางที่ 17 ค่าสถิติพื้นฐานของตัวแปรต่อเนื่องจำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา

คุณลักษณะของตัวแปร	N	Min	Max	$\overline{X}$	SD	Skewness	Kurtosis
นักเรียน							
1. ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา	6,192	-3.08	1.75	.01	.51	-.47	.65
2. เศรษฐานะของครอบครัว	6,192	-3.81	2.54	-1.23	1.18	.57	-.55
3. แหล่งทรัพยากรที่บ้าน	6,192	-3.83	.59	-1.16	1.02	.25	-.53
4. ความมั่งคั่งของครอบครัว	6,192	-4.29	1.79	-1.41	1.08	-.03	-.07
5. ความสนใจในการเรียนรู้วิทยาศาสตร์	6,192	-3.14	3.28	.76	.79	.14	2.83
6. ความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์	6,192	-3.76	3.22	.07	.73	.09	2.19
7. ความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์	6,192	-2.36	2.24	.66	.76	-.14	.38
8. ความเพลิดเพลินทางวิทยาศาสตร์	6,192	-2.15	2.05	.71	.67	-.16	.09
9. แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์	6,192	-2.10	1.82	.70	.65	.04	-.08
10. การเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์	6,192	-3.66	2.18	.77	.91	-.01	-.68
สถานศึกษา							
1. งบประมาณของรัฐ	212	.00	100.00	75.45	26.53	-.90	-.38
2. งบประมาณจากหน่วยงานภายนอก	212	.00	60.00	8.11	12.20	2.04	4.24
3. ดัชนีการขาดแคลนครู	212	-1.05	3.61	.64	.84	-.04	.64
4. ดัชนีทรัพยากรการเรียน	212	-3.43	2.13	-.63	1.09	.13	.62
5. สัดส่วนนักเรียนต่อครู	211	9.22	36.46	22.07	5.55	-.17	-.09
6. ดัชนีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์	212	-2.26	1.63	1.35	.59	-2.54	8.22
7. ดัชนีความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากร	212	-1.10	2.01	.21	1.07	.72	-.98
8. ดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล	212	-1.06	1.26	.82	.73	-1.29	.04

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

การวิเคราะห์คะแนนวิชาวิทยาศาสตร์ของนักเรียนจากผลการประเมินตามโครงการ PISA 2006 พบว่านักเรียนและสถานศึกษาที่มีคุณลักษณะพื้นฐานต่างกัน จะมีคะแนนเฉลี่ยวิชาวิทยาศาสตร์แตกต่างกันในแต่ละคุณลักษณะของตัวแปร เมื่อพิจารณาใน**ระดับนักเรียน** พบว่านักเรียนทั้งหมดมีคะแนนเฉลี่ยวิชาวิทยาศาสตร์เท่ากับ 429.73 คะแนน มีส่วนเบี่ยงเบนมาตรฐาน 79.21 มีลักษณะการแจกแจงของข้อมูลเป็นแบบเบ้ขวา (skewness = .30) และค่อนข้างแบน (kurtosis = -.08) แสดงว่านักเรียนส่วนใหญ่มีคะแนนต่ำกว่าค่าเฉลี่ย โดยนักเรียนหญิงมีคะแนนเฉลี่ยสูงกว่านักเรียนชาย (คะแนนเฉลี่ยนักเรียนหญิงและชาย เท่ากับ 435.73 และ 421.56 คะแนน ตามลำดับ) แต่ความแปรปรวนของคะแนนของนักเรียนชายมีมากกว่านักเรียนหญิง (ความแปรปรวนของคะแนนของนักเรียนชายและหญิง เท่ากับ 7199.43 และ 5520.75 ตามลำดับ) นักเรียนที่มีเวลาเรียนพิเศษวิทยาศาสตร์นอกสถานศึกษามากกว่า 6 ชั่วโมงต่อสัปดาห์จะมีคะแนนเฉลี่ยสูงสุด ($\overline{X}$ = 524.60, SD = 81.49) นักเรียนที่ใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเองมากกว่า 6 ชั่วโมงต่อสัปดาห์จะมีคะแนนเฉลี่ยสูงสุด ($\overline{X}$ = 478.27, SD = 84.42)

เมื่อพิจารณา**ระดับสถานศึกษา** พบว่าสถานศึกษาขนาดใหญ่มีคะแนนเฉลี่ยสูงสุด ($\overline{X}$ = 445.65, SD = 50.61) คะแนนของนักเรียนในสถานศึกษาขนาดเล็ก กลาง ใหญ่ มีลักษณะการแจกแจงของคะแนนต่างกัน โดยนักเรียนในสถานศึกษาขนาดเล็กส่วนใหญ่มีคะแนนต่ำเนื่องจากมีลักษณะการแจกแจงของข้อมูลเป็นแบบเบ้ซ้ายและค่อนข้างโด่ง (skewness = -.49, kurtosis = .47) สถานศึกษาขนาดกลางซึ่งมีความแปรปรวนสูงสุด (variance = 2989.67) มีลักษณะการแจกแจงคะแนนของนักเรียนเป็นแบบเบ้ขวา และโด่ง (skewness = 1.22, kurtosis = 1.30) สถานศึกษาขนาดใหญ่มีลักษณะการแจกแจงคะแนนของนักเรียนเป็นแบบเบ้ขวา และค่อนข้างแบน (skewness = .75, kurtosis = .35) แสดงว่านักเรียนในสถานศึกษาขนาดกลางและขนาดใหญ่มีคะแนนสูงกว่าค่าเฉลี่ย นักเรียนในสถานศึกษาของรัฐมีคะแนนเฉลี่ย($\overline{X}$ = 423.03, SD = 56.37) สูงกว่านักเรียนในสถานศึกษาเอกชน และมีความแปรปรวนสูงสุด (variance = 3178.37) เมื่อพิจารณาสังกัดสถานศึกษา พบว่าสถานศึกษาในสังกัดสำนักงานคณะกรรมการอุดมศึกษา (สกอ.) หรือโรงเรียนสาธิตมีคะแนนผลการประเมินสูงสุด ($\overline{X}$ = 540.43, SD = 28.51) รองลงมาคือสถานศึกษาในสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) ($\overline{X}$ = 416.50, SD = 47.79) และสถานศึกษาในสำนักงานการศึกษาเอกชน(สช.) ($\overline{X}$ = 416.44, SD = 48.91) ตามลำดับ ส่วนสถานศึกษาที่มีคะแนนต่ำสุดอยู่ในสังกัดสำนักงานประสานและพัฒนาการจัดการศึกษาท้องถิ่น และสำนักการศึกษากรุงเทพมหานคร(กศท. และ กทม.) ($\overline{X}$ = 397.83, SD = 34.84) ส่วนสถานศึกษาและวิทยาลัยอาชีวศึกษา (อศ.) ส่วนใหญ่มีคะแนนต่ำกว่าค่าเฉลี่ย เมื่อพิจารณาตำแหน่งที่ตั้งของสถานศึกษา พบว่าสถานศึกษาที่มีตำแหน่งที่ตั้งอยู่ในเมืองใหญ่มากจะมีคะแนนเฉลี่ยสูงสุด ($\overline{X}$ = 486.94, SD = 68.12) และมีความแปรปรวนสูงสุด (variance = 4641.57) และสถานศึกษาที่มีตำแหน่งที่ตั้งอยู่ในหมู่บ้านจะมีคะแนนเฉลี่ยต่ำสุด ($\overline{X}$ = 385.52, SD = 33.86) ขนาดของห้องเรียน พบว่าห้องเรียนที่มีจำนวนนักเรียน 46 – 50 มีคะแนนเฉลี่ยสูงสุด ($\overline{X}$ = 547.24, SD = 53.32) และขนาดห้องเรียนจำนวนนักเรียน 36 – 40 คน มีลักษณะการแจกแจงเป็นแบบเบ้ขวาและโด่ง (skewness = 1.29, kurtosis = 2.09) แสดงว่านักเรียนส่วนใหญ่มีคะแนนผลการประเมินต่ำ สถานศึกษาที่มีแรงกดดัน

ของผู้ปกครองจำนวนมากจะมีคะแนนเฉลี่ยสูงสุด ($\overline{X}$ = 449.33, SD = 55.98) และสถานศึกษาที่ไม่พบแรงกดดันของผู้ปกครองเป็นส่วนใหญ่จะมีคะแนนต่ำสุด และมีลักษณะการแจกแจงของคะแนนเป็นแบบเบ้ขวาและโด่ง (skewness = 1.11, kurtosis = 3.18) แสดงว่านักเรียนในสถานศึกษาที่ไม่มีแรงกดดันของผู้ปกครองพบว่าส่วนใหญ่จะมีคะแนนต่ำ แต่นักเรียนในสถานศึกษาที่มีแรงกดดันของผู้ปกครองจะมีคะแนนสูง สถานศึกษาที่ไม่ขาดแคลนครูวิทยาศาสตร์จะมีคะแนนเฉลี่ยสูงสุด ($\overline{X}$ = 473.04, SD = 58.10) สถานศึกษาที่ขาดแคลนครูวิทยาศาสตร์จำนวนมากจะมีคะแนนต่ำสุด สถานศึกษาที่ขาดแคลนครูวิทยาศาสตร์มากมีลักษณะการแจกแจงของคะแนนแบบโด่ง (skewness = .82, kurtosis = 1.66) นั่นคือนักเรียนในสถานศึกษาที่มีการขาดแคลนครูวิทยาศาสตร์จำนวนมากส่วนใหญ่จะมีคะแนนผลการประเมินต่ำ สถานศึกษาที่ไม่ขาดแคลนครูที่มีคุณวุฒิจะมีคะแนนเฉลี่ยสูงสุด ($\overline{X}$ = 449.01, SD = 70.30) สถานศึกษาที่ขาดแคลนครูที่มีคุณวุฒิจำนวนมากจะมีคะแนนต่ำสุด แสดงให้เห็นว่านักเรียนในสถานศึกษาที่ไม่มีการขาดแคลนครูวิทยาศาสตร์และครูที่มีคุณวุฒิจะมีคะแนนผลการประเมินวิทยาศาสตร์สูง แสดงผลดังตารางที่ 18

ตารางที่ 18 ค่าสถิติพื้นฐานของคะแนนวิชาวิทยาศาสตร์ของนักเรียนจากผลการประเมินตามโครงการประเมินผล PISA 2006 จำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา

คุณลักษณะ	N	Min	Max	range	$\overline{X}$	SD	Variance	Skew.	Kur.
นักเรียน									
1. เพศ									
1.1 ชาย	2,608	155.01	710.49	555.47	421.56	84.84	7199.43	.31	-.18
1.2 หญิง	3,584	216.84	677.39	460.55	435.66	74.30	5520.75	.36	-.03
2. เวลาที่เรียนพิเศษวิทยาศาสตร์นอกเวลาเรียน/สัปดาห์									
2.1 ไม่มีเวลา	3480	203.97	669.09	465.12	423.97	68.47	4689.19	.32	.22
2.2 น้อยกว่า 2 ชม.	1381	155.01	655.57	500.55	413.71	78.50	6163.37	.34	.05
2.3 เวลา 2 – 4 ชม.	967	194.18	710.49	516.31	448.72	92.68	8590.06	.01	-.64
2.4 เวลา 4 – 6 ชม.	260	222.25	677.39	455.14	483.20	99.24	9849.36	-.51	-.44
2.5 มากกว่า 6 ชม.	104	283.32	669.55	386.23	524.60	81.49	6640.78	-.75	.26
3. เวลาที่ใช้ในการศึกษาวิทยาศาสตร์ด้วยตนเอง/สัปดาห์									
3.1 ไม่มีเวลา	588	155.01	638.04	483.02	401.13	72.48	5254.45	.19	.18
3.2 น้อยกว่า 2 ชม.	3,460	203.97	659.39	455.42	422.54	73.68	5430.13	.36	.13
3.3 เวลา 2 – 4 ชม.	1,546	206.95	686.34	479.38	444.87	83.33	6944.91	.14	-.43
3.4 เวลา 4 – 6 ชม.	381	183.08	710.49	527.41	462.26	86.73	7522.54	.06	-.27
3.5 มากกว่า 6 ชม.	163	295.63	668.53	372.90	478.27	84.42	7128.18	-.03	-.60
รวม	6,192	155.01	710.49	555.47	429.73	79.21	6275.21	.30	-.08

ตารางที่ 18 ค่าสถิติพื้นฐานของคะแนนวิชาวิทยาศาสตร์ของนักเรียนจากผลการประเมินตาม โครงการ PISA 2006 จำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา (ต่อ)

คุณลักษณะ	N	Min	Max	range	$\overline{X}$	SD	Variance	Skew.	Kur.
สถานศึกษา									
1. ขนาดสถานศึกษา									
1.1 ขนาดเล็ก	36	290.94	425.80	134.86	374.51	31.00	961.28	-.49	.47
1.2 ขนาดกลาง	84	324.89	580.94	256.05	417.08	54.67	2989.67	1.22	1.30
1.3 ขนาดใหญ่	92	345.28	579.90	234.62	445.65	50.61	2561.66	.52	-.14
2. ประเภทสถานศึกษา									
2.1 รัฐ	187	290.94	580.94	290.00	423.03	56.37	3178.37	.733	.365
2.2 เอกชน	25	348.55	525.62	177.07	416.44	48.91	2392.47	.869	-.022
3. สังกัดสถานศึกษา									
3.1 สพฐ.	134	290.94	545.18	254.24	416.50	47.79	2284.21	.33	.20
3.2 สช.	25	348.55	525.62	177.07	416.44	48.91	2392.47	.86	-.02
3.3 รร.สาธิต	15	491.10	580.94	89.84	540.43	28.51	812.83	.13	-1.24
3.4 กศท.และ กทม.	30	326.72	479.89	153.17	397.83	34.84	1357.24	.44	.28
3.5 อศ.	8	388.69	452.19	63.50	406.66	20.72	429.53	1.88	3.43
4. ตำแหน่งที่ตั้งสถานศึกษา									
4.1 หมู่บ้าน	51	290.94	474.94	184.00	385.52	33.86	1146.81	-.20	.93
4.2 เมืองเล็ก	44	326.72	515.86	189.14	404.28	42.51	1807.52	.76	.78
4.3 เมือง	54	324.89	560.44	235.55	431.86	47.45	2252.41	.51	.43
4.4 เมืองใหญ่	43	359.89	580.94	221.05	442.06	53.08	2817.60	.69	-.14
4.5 เมืองใหญ่มาก	20	357.03	579.90	222.88	486.94	68.12	4641.57	-.42	-.92
5. ขนาดห้องเรียน									
5.1 <= 15 คน	5	341.71	461.44	119.73	386.82	46.82	2192.83	1.23	1.28
5.2 16-20 คน	4	338.39	422.20	83.81	384.55	34.56	1194.97	-.72	1.79
5.3 21-25 คน	12	290.94	428.86	137.93	375.64	40.66	1653.33	-1.03	.58
5.4 26-30 คน	14	328.03	442.86	114.84	376.22	31.44	988.528	.46	.30
5.5 31-35 คน	20	362.66	579.90	217.24	437.78	71.02	5044.31	.79	-.65
5.6 36-40 คน	62	324.89	580.94	256.05	417.71	54.00	2916.04	1.29	2.09
5.7 41-45 คน	56	357.03	529.70	172.67	427.84	45.39	2060.36	.54	-.53
5.8 46 - 50 คน	28	355.33	559.49	204.16	457.24	53.32	2844.06	.22	-.65
5.9 > 50 คน	11	376.24	525.62	149.38	441.37	53.35	2846.87	.20	-1.24

ตารางที่ 18 ค่าสถิติพื้นฐานของคะแนนวิชาวิทยาศาสตร์ของนักเรียนจากผลการประเมินตามโครงการ PISA 2006 จำแนกตามคุณลักษณะของนักเรียนและสถานศึกษา (ต่อ)

คุณลักษณะ	N	Min	Max	range	$\overline{X}$	SD	Variance	Skew.	Kur.
6. แรงกดดันของผู้ปกครองต่อสถานศึกษา									
6.1 มีจำนวนมาก	59	336.39	580.94	244.55	449.33	55.98	3134.22	.10	-.66
6.2 มีส่วนน้อย	90	311.00	579.90	268.90	417.15	55.24	3052.38	1.02	1.06
6.3 ไม่พบเป็นส่วนใหญ่	59	290.94	560.44	269.50	403.67	47.02	2211.32	1.11	3.18
7. การขาดแคลนครูวิทยาศาสตร์									
7.1 ไม่ขาดแคลน	72	324.89	573.23	248.34	437.04	58.10	3375.85	.55	-.63
7.2 ขาดแคลนน้อย	95	290.94	580.94	290.00	419.55	54.54	2974.78	.88	1.34
7.3 ขาดแคลนเป็นบางส่วน	31	326.72	529.56	202.84	410.17	49.70	2470.20	.58	.06
7.4 ขาดแคลนจำนวนมาก	14	328.03	491.10	163.08	391.26	40.88	1671.92	.82	1.66
8. การขาดแคลนครูที่มีคุณวุฒิ									
8.1 ไม่ขาดแคลน	37	311.00	580.94	269.94	449.01	70.30	4943.46	.07	-.99
8.2 ขาดแคลนน้อย	95	324.89	574.40	249.51	427.39	54.16	2934.03	.62	-.04
8.3 ขาดแคลนเป็นบางส่วน	58	326.72	579.90	253.18	402.97	43.50	1893.03	1.56	4.71
8.4 ขาดแคลนจำนวนมาก	22	290.94	472.73	181.79	406.00	40.50	1640.46	-.85	2.17
รวม	212	290.94	580.94	290.00	422.24	55.61	3093.06	.74	.34

1.2 การตรวจสอบเงื่อนไขของการวิเคราะห์พหุระดับ

การตรวจสอบเงื่อนไขเบื้องต้นของการวิเคราะห์เชิงการทำนาย สิ่งที่ต้องระวังก่อนการใช้โมเดล คือการตรวจสอบความสัมพันธ์ระหว่างตัวแปร (Measures of Collinearity) และการตรวจสอบความสอดคล้องกับสมการการวิเคราะห์ถดถอยและโมเดลพหุระดับ

เพื่อให้สะดวกต่อการนำเสนอผลการวิเคราะห์ในส่วนนี้ และให้เกิดความเข้าใจในสัญลักษณ์ของตัวแปรแต่ละตัว ผู้วิจัยจึงกำหนดสัญลักษณ์และความหมายแทนตัวแปรในระดับนักเรียนและระดับสถานศึกษา ดังนี้

ระดับตัวแปร	ชื่อตัวแปร	ความหมายของตัวแปร
ระดับนักเรียน	SCIE	คะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์
	RELAGR	ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา
	MALE	ตัวแปรดัมมี่ความเป็นเพศชายของนักเรียน
	ESCS	เศรษฐานะของครอบครัว
	HEDRES	แหล่งทรัพยากรที่บ้าน
	WEALTH	ความมั่งคั่งของครอบครัว
	INTSCIE	ความสนใจในการเรียนรู้วิทยาศาสตร์
	SCIEEFF	ความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์
	SCSCIE	ความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์
	JOYSCIE	ความเพลิดเพลินทางวิทยาศาสตร์
	INSTSCIE	แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์
	GENSCIE	การเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์
	ST31Q02	ระดับการใช้เวลาเรียนพิเศษวิทยาศาสตร์นอกสถานศึกษา
	ST31Q03	ระดับการใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง
ระดับสถานศึกษา	DSCHL	ตัวแปรดัมมี่สถานศึกษาขนาดใหญ่
	DSCHM	ตัวแปรดัมมี่สถานศึกษาขนาดกลาง
	PUBLIC	ตัวแปรดัมมี่สถานศึกษาของรัฐ
	TYPE1	ตัวแปรดัมมี่สถานศึกษาสังกัด สพฐ.
	TYPE2	ตัวแปรดัมมี่สถานศึกษาสังกัด สช.
	TYPE3	ตัวแปรดัมมี่สถานศึกษาสังกัด สกอ.
	TYPE4	ตัวแปรดัมมี่สถานศึกษาสังกัด กศท. และ กทม.
	COMMS	ตัวแปรดัมมี่สถานศึกษาที่ตั้งอยู่ในหมู่บ้าน
	COMMM	ตัวแปรดัมมี่สถานศึกษาที่ตั้งอยู่ในเมืองเล็ก
	COMML	ตัวแปรดัมมี่สถานศึกษาที่ตั้งอยู่ในเมือง
	COMMXL	ตัวแปรดัมมี่สถานศึกษาที่ตั้งอยู่ในเมืองใหญ่
	SC06Q01	ขนาดห้องเรียน
	SC16Q01	ระดับแรงกดดันของผู้ปกครองต่อสถานศึกษา
	SC14Q01	ระดับการขาดแคลนครูวิทยาศาสตร์
	SC14Q04	ระดับการขาดแคลนครูที่มีคุณวุฒิ
	SC03Q01	ร้อยละงบประมาณของรัฐที่จัดสรรให้สถานศึกษา
	SC03Q03	ร้อยละงบประมาณจากหน่วยงานภายนอก
	TCSHORT	ดัชนีการขาดแคลนครู
	SCMATEDU	ดัชนีทรัพยากรการเรียน
	STRATIO	สัดส่วนนักเรียนต่อครู
	SCIPROM	ดัชนีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์
	RESPRES	ดัชนีความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากร
	RESPURR	ดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

การตรวจสอบความสัมพันธ์ระหว่างตัวแปร (Measures of Collinearity)

การตรวจสอบความสัมพันธ์ของตัวแปรที่ใช้ในการวิจัย มุ่งตรวจสอบความสัมพันธ์ระหว่างตัวแปรต้นคู่ที่มีความสัมพันธ์กันสูงหรือภาวะร่วมเส้นตรงพหุ (multicollinearity) ถ้าหากพบตัวแปรต้นคู่ที่มีความสัมพันธ์กันในระดับสูง จะพิจารณาเลือกตัวแปรใดตัวแปรหนึ่งเข้าสู่สมการการวิเคราะห์แทนได้

จากการวิเคราะห์ความสัมพันธ์ของตัวแปรต่อเนื่องระหว่างตัวแปรระดับนักเรียน พบว่า ตัวแปรที่มีความสัมพันธ์กันในระดับสูงอย่างมีนัยสำคัญที่ระดับ .01 ได้แก่ความสัมพันธ์ระหว่างเศรษฐานะของครอบครัวกับความมั่งคั่งของครอบครัว ($r_{ESCS,WEALTH}$ = .780) ลำดับถัดมาคือความสัมพันธ์ระหว่างเศรษฐานะของครอบครัวกับแหล่งทรัพยากรที่บ้าน ($r_{ESCS,\ HEDRES}$ = .692) ความสัมพันธ์ระหว่างแหล่งทรัพยากรที่บ้านกับความมั่งคั่งของครอบครัว ($r_{HEDRES,\ WEALTH}$ = .648) และความสัมพันธ์ระหว่างความสนใจในการเรียนรู้วิทยาศาสตร์กับความเพลิดเพลินทางวิทยาศาสตร์ ($r_{INTSCIE,\ JOYSCIE}$ = .520 ส่วนตัวแปรคู่ที่ไม่มีความสัมพันธ์กัน ได้แก่ความสัมพันธ์ระหว่างผลสัมฤทธิ์ทางการเรียนที่ผ่านมากับความสนใจในการเรียนรู้วิทยาศาสตร์ ($r_{RELAGR,INTSCIE}$ = -.011) ความสัมพันธ์ระหว่างผลสัมฤทธิ์ทางการเรียนที่ผ่านมากับแรงจูงใจภายนอกในการเรียนวิทยาศาสตร์ ($r_{RELAGR,\ INSTSCIE}$ = .011) ความสัมพันธ์ระหว่างแหล่งทรัพยากรที่บ้านกับความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์ ($r_{HEDRES,\ SCSCIE}$ = .016) และความสัมพันธ์ระหว่างความมั่งคั่งของครอบครัวกับแรงจูงใจภายนอกในการเรียนวิทยาศาสตร์ ($r_{WEALTH,\ INSTSCIE}$ = -.017) นอกจากนี้ยังพบว่าตัวแปรตามหรือคะแนนผลการประเมินวิชาวิทยาศาสตร์ในโครงการ PISA 2006 มีความสัมพันธ์กับตัวแปรต้นทุกตัวในระดับนักเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 แสดงดังตารางที่ 19

ตารางที่ 19 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร ในระดับนักเรียน

	SCIE	RELAGR	ESCS	HEDRES	WEALTH	INTSCIE	SCIEEFF	SCSCIE	JOYSCIE	INSTSCIE	GENSCIE
SCIE	1.000										
RELAGR	.341**	1.000									
ESCS	.497**	.252**	1.000								
HEDRES	.448**	.230**	.692**	1.000							
WEALTH	.394**	.216**	.780**	.648**	1.000						
INTSCIE	.118**	-.011	-.022*	.052**	-.077**	1.000					
SCIEEFF	.164**	.033**	.124**	.171**	.069**	.415**	1.000				
SCSCIE	-.065**	-.098**	-.036**	.016	-.072**	.380**	.308**	1.000			
JOYSCIE	.180**	-.025*	.031**	.090**	-.049**	.520**	.390**	.388**	1.000		
INSTSCIE	.160**	.011	.024*	.104**	-.017	.450**	.304**	.450**	.465**	1.000	
GENSCIE	.239**	.062**	.099**	.139**	.041**	.333**	.296**	.231**	.381**	.407**	1.000

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

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

การตรวจสอบภาวะร่วมเส้นตรงพหุ (multicollinearity) ผู้วิจัยใช้การวิเคราะห์ถดถอยพหุ (multiple regression) มีตัวแปรตามเป็นคะแนนผลการประเมินวิชาวิทยาศาสตร์ในโครงการ PISA 2006 (PVSCIE) และตัวแปรทำนายเป็นตัวแปรต่อเนื่องในระดับนักเรียนได้แก่ ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา (RELAGR) เศรษฐานะของครอบครัว (ESCS) แหล่งทรัพยากรที่บ้าน (HEDRES) ความมั่งคั่งของครอบครัว (WEALTH) ความสนใจในการเรียนรู้วิทยาศาสตร์ (INTSCIE) ความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์ (SCIEEFF) ความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์ (SCSCIE) ความเพลิดเพลินทางวิทยาศาสตร์ (JOYSCIE) แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์(INSTSCIE) และการเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์ (GENSCIE) โดยพิจารณาจากค่า Tolerance และค่า VIF ซึ่ง Hair และคณะ (1998) เสนอว่าค่าสถิติเริ่มต้นของตัวแปรที่แสดงว่ามีภาวะร่วมเส้นตรงพหุ คือ มีค่า Tolerance ตั้งแต่ .10 ลงมา หรือ VIF มีค่า 10 ขึ้นไป

จากผลทดสอบความสัมพันธ์ระหว่างตัวแปรต้น (Measures of Collinearity) ในระดับนักเรียน โดยพิจารณาค่า Tolerance และ Variance Inflation Factor : VIF) ของตัวแปรต้น พบว่าผลการทดสอบความสัมพันธ์ระหว่างตัวแปรต้นแต่ละตัวมีค่า Tolerance สูง หรือมีค่าสูงกว่า .10 ตามข้อเสนอแนะของ Hair (Tolerance = .327 - .917) และค่า VIF ต่ำ หรือมีค่าต่ำกว่า 10 ตามข้อเสนอแนะของ Hair (VIF = 1.091 – 3.061) แสดงว่าตัวแปรต้นแต่ละตัวมีความสัมพันธ์กันต่ำ จึงไม่เกิดปัญหาที่ตัวแปรต้นมีความสัมพันธ์กันสูงหรือปัญหาภาวะร่วมเส้นตรงพหุ (multicollinearity) แสดงดังตารางที่ 20

ตารางที่ 20 ค่า Tolerance และ VIFจากผลการทดสอบความสัมพันธ์ระหว่างตัวแปรต้น ระดับนักเรียน

ตัวแปร	Collinearity Statistics	
	Tolerance	VIF
RELAGR	.917	1.091
ESCS	.327	3.061
HEDRES	.473	2.115
WEALTH	.360	2.775
INTSCIE	.618	1.619
SCIEEFF	.742	1.347
SCSCIE	.719	1.392
JOYSCIE	.606	1.649
INSTSCIE	.623	1.604
GENSCIE	.759	1.317

ผลการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรต้นที่เป็นตัวแปรต่อเนื่องในระดับสถานศึกษา พบว่า ตัวแปรที่มีความสัมพันธ์กันในระดับสูงอย่างมีนัยสำคัญที่ระดับ .01 ได้แก่ ความสัมพันธ์ทาง

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ลบระหว่างดัชนีการขาดแคลนครูกับดัชนีทรัพยากรการเรียน ($r_{TCSHORT,\ SCMATEDU}$ = -.496) ลำดับถัดมาคือความสัมพันธ์ทางลบระหว่างงบประมาณของรัฐกับดัชนีทรัพยากรการเรียน ($r_{SC03Q01,\ SCMATEDU}$ = -.468) และความสัมพันธ์ทางลบระหว่างงบประมาณของรัฐกับดัชนีความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากร ($r_{SC03Q01,\ RESPRES}$ = -.398) ส่วนตัวแปรต้นคู่ที่ไม่มีความสัมพันธ์ได้แก่ ความสัมพันธ์ระหว่างงบประมาณของรัฐกับดัชนีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์ ($r_{SC03Q01,\ SCIPROM}$ = .000) ความสัมพันธ์ระหว่างดัชนีทรัพยากรการเรียนกับสัดส่วนนักเรียนต่อครู ($r_{SCMATEDU,\ STRATIO}$ = -.107) ความสัมพันธ์ระหว่างดัชนีการขาดแคลนครู กับดัชนีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์ ($r_{TCSHORT\ ,\ SCIPROM}$ = -.017) และความสัมพันธ์ระหว่างสัดส่วนนักเรียนต่อครู กับดัชนีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์ ($r_{STRATIO\ ,\ SCIPROM}$ = .115) นอกจากนี้ดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล พบว่าส่วนใหญ่ไม่มีความสัมพันธ์กันกับตัวแปรต้นตัวอื่น ยกเว้นความสัมพันธ์คู่กับตัวแปรดัชนีความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากร แสดงดังตารางที่ 21

ตารางที่ 21 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร ในระดับสถานศึกษา

	SC03Q01	SC03Q01	TCSHORT	SCMATEDU	STRATIO	SCIPROM	RESPRES	RESPCURR
SC03Q01	1.000							
SC03Q03	-.406**	1.000						
TCSHORT	.250**	-.077	1.000					
SCMATEDU	-.468**	.188**	-.496**	1.000				
STRATIO	.244**	-.094	.149*	-.107	1.000			
SCIPROM	.000	.101	-.017	.029	.115	1.000		
RESPRES	-.398**	.009	-.186**	.188**	-.156*	-.083	1.000	
RESPCURR	-.057	.019	-.073	.017	-.097	.036	.265**	1.000

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

เมื่อทดสอบความสัมพันธ์ระหว่างตัวแปรต้นในระดับสถานศึกษา เพื่อตรวจสอบภาวะร่วมเส้นตรงพหุ (multicollinearity) โดยใช้การวิเคราะห์ถดถอยพหุที่มีตัวแปรตามเป็นคะแนนผลการประเมินวิชาวิทยาศาสตร์ในโครงการ PISA 2006 (PVSCIE) และตัวแปรทำนายเป็นตัวแปรต่อเนื่องในระดับสถานศึกษาได้แก่ งบประมาณของรัฐ (SC03Q01) งบประมาณจากหน่วยงานภายนอก (SC03Q03) ดัชนีการขาดแคลนครู (TCSHORT) ดัชนีทรัพยากรการเรียน (SCMATEDU) สัดส่วนนักเรียนต่อครู (STRATIO) ดัชนีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์ (SCIPROM) ดัชนีความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากร (RESPRES) และดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล (RESPURR) ค่า Tolerance และ Variance Inflation Factor : VIF) ของตัวแปรต้นแต่ละตัว พบว่าผลการทดสอบความสัมพันธ์ระหว่างตัวแปรต้นมีค่า Tolerance สูง (Tolerance = .655 - .974) หรือมีค่าสูงกว่า .10 และค่า VIF ต่ำ หรือมีค่าต่ำกว่า 10 (VIF = 1.090 – 1.526) แสดงว่าตัวแปรต้นแต่ละตัวมีความสัมพันธ์กันต่ำ จึงไม่เกิดปัญหาที่ตัวแปรต้นมีความสัมพันธ์กันสูงหรือปัญหาภาวะร่วมเส้นตรงพหุ (multicollinearity) แสดงดังตารางที่ 22

ตารางที่ 22 ค่า Tolerance และ VIF จากผลการทดสอบความสัมพันธ์ระหว่างตัวแปรต้น ระดับสถานศึกษา

ตัวแปร	Collinearity Statistics	
	Tolerance	VIF
SC03Q01	.552	1.811
SC03Q03	.798	1.253
TCSHORT	.733	1.363
SCMATEDU	.623	1.605
STRATIO	.909	1.100
SCIPROM	.965	1.037
RESPRES	.746	1.341
RESPCURR	.916	1.092

การตรวจสอบการแจกแจงของความคลาดเคลื่อน (Distributional assumptions)

การใช้โมเดลพหุระดับในการวิเคราะห์ จำเป็นต้องมีการตรวจสอบการแจกแจงของความคลาดเคลื่อนหรือค่าเศษเหลือ (residual) โดยความคลาดเคลื่อนจะต้องมีลักษณะการแจกแจงแบบปกติ Raudenbush et al. (2004) ได้เสนอแนะแนวทางในการตรวจสอบลักษณะการแจกแจง โดยการพล๊อตกราฟระหว่างค่าที่สังเกตได้ (Observed Value) กับค่าที่คาดหวังปกติ (Expected Normal Value) โดยค่าเศษเหลือในระดับที่ 1 สามารถประมาณค่าได้โดยการใช้โปรแกรม HLM และบันทึกค่าเศษเหลือ ระดับที่ 1 (residual level 1) ในไฟล์ SPSS หลังจากนั้นจึงเขียนกราฟ Q-Q plot ของ residual level 1

จากการเขียนกราฟ Q-Q plot ของ residual level 1 พบว่าได้กราฟเป็นเส้นตรงที่เอียงทำมุม 45 องศากับแนวระดับ แสดงให้เห็นว่าความคลาดเคลื่อนหรือค่าเศษเหลือมีลักษณะการแจกแจงแบบปกติ แสดงดังภาพที่ 22 (ภาคผนวก ก)

การตรวจสอบการแจกแจงของค่าเศษเหลือในระดับที่ 2 (residual level 2) สามารถตรวจสอบการแจกแจงแบบปกติ ได้โดยการเขียนกราฟของ Mahalanobis distance ซึ่งถ้าความคลาดเคลื่อนมีการแจกแจงแบบปกติ Mahalanobis distance จะมีการแจกแจงไคสแควร์ ด้วยการเขียน Q-Q plot ของค่า Mahalanobis distance และ ค่า Expected value นั่นคือเส้นกราฟจะมีลักษณะเป็นเส้นตรงที่เอียงทำมุม 45 องศากับแนวระดับ จากกราฟ Q-Q plot ของ ค่า Mahalanobis distance (mdist) และ Q-Q plot ของค่า Expected value (chipct) พบว่าเส้นกราฟจะมีลักษณะเป็นเส้นตรงที่เอียงทำมุม 45 องศากับแนวระดับ แสดงว่าความคลาดเคลื่อนหรือค่าเศษเหลือในระดับที่ 2 มีการแจกแจงแบบปกติ ดังภาพที่ 23 (ภาคผนวก ก)

นอกจากนี้การตรวจสอบการแจกแจงของความคลาดเคลื่อนแบบปกติแล้ว ก่อนการวิเคราะห์สมการถดถอย จะต้องมีตรวจสอบว่าความคลาดเคลื่อนหรือส่วนที่เหลือ กับค่าทำนายของตัวแปรตาม มีความสัมพันธ์เป็นเส้นตรงหรือไม่ โดยการเขียนกราฟระหว่างความคลาดเคลื่อนหรือส่วนที่เหลือ(residual) กับค่าทำนาย (Expected value) ของตัวแปรตาม พบว่ามีลักษณะของค่าที่เกิดขึ้นเป็นกระจายทั่วไปคล้ายรูปวงกลม ดังนั้นจึงมีลักษณะการกระจายเป็นไปอย่างสุม หรือไม่เป็นระบบ จึงไม่น่าจะเกิดปัญหาความสัมพันธ์ที่ไม่ใช่เชิงเส้นตรง ดังภาพที่ 24 (ภาคผนวก ก)

1.3 การตรวจสอบความเป็นเอกมิติของแบบสอบ (Unidimentionality)

การนำหลักการพหุระดับมาใช้ในการวิเคราะห์ข้อสอบ โดยการใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่นได้ มีความสอดคล้องกับหลักการของ IRT ในโมเดลของราสช์ (Rasch Model) (Kamata, 2001) ซึ่งมีแนวคิดเช่นเดียวกันกับทฤษฎีการตอบสนองข้อสอบแบบหนึ่งพารามิเตอร์ ประกอบด้วยการประมาณค่าพารามิเตอร์ของผู้สอบหรือความสามารถของผู้สอบ และพารามิเตอร์ของข้อสอบหรือค่าความยากของข้อสอบ

ดังนั้นก่อนการวิเคราะห์ข้อสอบ ผู้วิจัยจึงจำเป็นต้องตรวจสอบข้อตกลงเบื้องต้นที่สำคัญของทฤษฎี IRT นั่นคือข้อตกลงเบื้องต้นที่สอดคล้องกับความเป็นเอกมิติของแบบสอบ (unidimensional models) ซึ่งการวิเคราะห์ความเป็นเอกมิติของแบบสอบสามารถตรวจสอบได้ด้วยการวิเคราะห์องค์ประกอบ (Factor Analysis) โดยการพิจารณาจากค่าไอเกน (Eigen value) (Lord & Novick, 1968 ; Reckase, 1979) ถ้าผลการวิเคราะห์องค์ประกอบพบว่าค่าไอเกนตัวแรกมีค่ามากกว่าตัวที่สองและตัวอื่นๆ อย่างเห็นได้ชัด จึงสามารถสรุปได้ว่ามีความเป็นเอกมิติของแบบสอบ (วรนุช แหยมแสง, 2536 ; สุวิมล ติรกานันท์, 2537) นอกจากนี้ Hair และคณะ (1998) เสนอว่าน้ำหนักองค์ประกอบในแต่ละองค์ประกอบที่มีค่ามากกว่า 0.3 สามารถจัดให้อยู่ในองค์ประกอบนั้น ดังนั้นในการตรวจสอบความเป็นเอกมิติของแบบสอบ ผู้วิจัยจึงแสดงผลการวิเคราะห์องค์ประกอบโดยใช้การเขียนกราฟของค่าไอเกนหรือแผนภาพสกรี (Scree plot) และการเขียนแผนภาพองค์ประกอบ (Loading plot)

จากผลการตรวจสอบความเป็นเอกมิติของแบบสอบที่ใช้ในการวิจัยครั้งนี้ พบว่าแบบสอบทั้ง 13 ฉบับ มีความเป็นเอกมิติ เนื่องจากค่าไอเกนขององค์ประกอบที่ 1 มีค่ามากที่สุด และแตกต่างจากตัวที่สองหรือตัวอื่นๆ อย่างชัดเจน ซึ่งแสดงได้ด้วยการเขียนกราฟของค่าไอเกนหรือแผนภาพสกรี (ดังภาพที่ 25) และจากการเขียนแผนภาพองค์ประกอบ (Loading plot) พบว่ามีเพียงองค์ประกอบเดียว แสดงดังภาพที่ 26 (ภาคผนวก ก)

ตอนที่ 2 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแต่ละฉบับ และการประมาณค่าความสามารถของผู้สอบ และค่าความยากของข้อสอบ

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยโมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น (Hierarchical Generalized Linear Model: HGLM) เป็นการประยุกต์หลักการวิเคราะห์พหุระดับเพื่อวิเคราะห์ข้อสอบ ในการวิจัยครั้งนี้ ผู้วิจัยใช้โมเดลการวิเคราะห์ 3 ระดับ โดยระดับที่ 1 เป็นระดับข้อสอบ ระดับที่ 2 เป็นระดับนักเรียน และระดับที่ 3 เป็นระดับสถานศึกษา เพื่อวิเคราะห์หาอิทธิพลของตัวแปรภายนอกที่มีผลต่อโอกาสการตอบข้อสอบถูก และวิเคราะห์ค่าความสามารถของนักเรียนและค่าความยากของข้อสอบ

การนำเสนอผลการวิเคราะห์ในตอนนี้จะแบ่งเป็น 2 ตอน คือ 2.1) สรุปผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ 2.2) สรุปข้อค้นพบเกี่ยวกับลักษณะของข้อสอบที่ทำหน้าที่ต่างกันในแต่ละคุณลักษณะของนักเรียน และ 2.3) ผลการประมาณค่าความสามารถของนักเรียนและค่าความยากของข้อสอบ

2.1 สรุปผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบเป็นการวิเคราะห์หาอิทธิพลของตัวแปรระดับนักเรียนที่มีผลต่อโอกาสการตอบข้อสอบถูก ผู้วิจัยได้พิจารณาตัวแปรเพื่อนำมาจัดกลุ่มคุณลักษณะของนักเรียน ที่เป็นตัวแปรภายนอกที่อยู่นอกเหนือการควบคุมให้เกิดขึ้นได้เอง จึงกำหนดตัวแปรที่นำมาใช้ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบดังนี้

MALE	ตัวแปรดัมมี่ความเป็นเพศชายของนักเรียน กำหนดให้นักเรียนเพศชายแทนค่าเป็น 1 และนักเรียนเพศหญิงแทนค่าเป็น 0
EXTRA	ระดับการใช้เวลาเรียนพิเศษวิทยาศาสตร์นอกสถานศึกษา กำหนดให้ นักเรียนที่เรียนพิเศษแทนค่าเป็น 1 และนักเรียนที่ไม่เรียนพิเศษแทนค่าเป็น 0
ESCS	เศรษฐานะของครอบครัว กำหนดให้นักเรียนที่เศรษฐานะของครอบครัวสูงกว่าค่าเฉลี่ยของนักเรียนทั้งหมด แทนค่าเป็น 1 และมีเศรษฐานะของครอบครัวต่ำกว่าค่าเฉลี่ย แทนค่าเป็น 0
HEDRES	แหล่งทรัพยากรที่บ้าน กำหนดให้นักเรียนที่มีแหล่งทรัพยากรที่บ้านสูงกว่าค่าเฉลี่ยของนักเรียนทั้งหมดแทนค่าเป็น 1 และมีแหล่งทรัพยากรที่บ้านต่ำกว่าค่าเฉลี่ย แทนค่าเป็น 0
WEALTH	ความมั่งคั่งของครอบครัว กำหนดให้นักเรียนที่มีความมั่งคั่งของครอบครัวสูงกว่าค่าเฉลี่ยของนักเรียนทั้งหมดแทนค่าเป็น 1 และมีความมั่งคั่งของครอบครัวต่ำกว่าค่าเฉลี่ย แทนค่าเป็น 0

ในการนำเสนอส่วนนี้ ผู้วิจัยสรุปจำนวนข้อสอบที่ทำหน้าที่ต่างกันในแต่ละฉบับ และเสนอรายละเอียดของผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแต่ละฉบับในภาคผนวก ข

จากการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) ในแบบสอบทั้ง 13 ฉบับ ที่เกิดจากคุณลักษณะของนักเรียนที่ต่างกลุ่มกันในตัวแปรความเป็นเพศชาย การเรียนพิเศษ เศรษฐานะของครอบครัว แหล่งทรัพยากรที่บ้าน และความมั่งคั่งของครอบครัว ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบคิดเป็นร้อยละ 11.53 ถึง 67.85 โดยแบบสอบที่มีข้อสอบที่ทำหน้าที่ต่างกันมากกว่าร้อยละ 50 ได้แก่ แบบสอบฉบับที่ 1, 7, 8, 10 และ 12 แบบสอบที่ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันน้อยที่สุด ได้แก่ แบบสอบฉบับที่ 2, 4 และ 11 เป็นที่น่าสังเกตว่า ส่วนใหญ่ข้อสอบที่ทำหน้าที่ต่างกันจะเป็นข้อสอบที่วัดในสมรรถนะทางวิทยาศาสตร์ด้านการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ (EPS) แสดงรายละเอียดดังตารางที่ 23

ตารางที่ 23 ผลการตรวจข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบทั้ง 13 ฉบับ

ฉบับที่	จำนวนนักเรียน	สมรรถนะทางวิทยาศาสตร์			รวม	ร้อยละของ DIF
		ISI	EPS	USE		
1	474	10(7)	33(19)	15(11)	**58**(37)	63.79%
2	475	5(1)	12(2)	9(0)	**26**(3)	11.53%
3	468	7(0)	17(5)	7(2)	**31**(7)	22.58%
4	466	7(2)	18(4)	8(0)	**33**(6)	18.18%
5	469	15(6)	24(5)	21(5)	**60**(16)	26.67%
6	483	8(4)	15(5)	9(3)	**32**(12)	37.50%
7	488	2(0)	8(7)	5(3)	**15**(10)	66.67%
8	478	6(4)	10(6)	10(7)	**26**(17)	65.38%
9	473	11(5)	11(1)	6(3)	**28**(9)	32.14%
10	474	8(6)	12(7)	8(6)	**28**(19)	67.85%
11	482	5(1)	12(5)	11(0)	**28**(6)	21.43%
12	486	8(5)	14(10)	8(5)	**30**(20)	66.67%
13	476	2(1)	8(5)	5(1)	**15**(7)	46.67%

หมายเหตุ ตัวเลขในวงเล็บ เป็นจำนวนข้อสอบที่ทำหน้าที่ต่างกัน

ISI แทน การระบุประเด็นทางวิทยาศาสตร์

EPS แทน การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์

USE แทน การใช้ประจักษ์พยานทางวิทยาศาสตร์

จากผลการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบแต่ละฉบับ ที่พบว่าแบบสอบแต่ละฉบับมีข้อสอบที่ทำหน้าที่ต่างกันค่อนข้างมาก ผู้วิจัยจึงเลือกตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปเป็นบางข้อ โดยพิจารณาเลือกตัดข้อสอบตามเกณฑ์ 3 ข้อ ดังนี้ 1) เป็นข้อสอบที่ตรวจพบว่าทำหน้าที่ต่างกันเนื่องจากตัวแปรคุณลักษณะนักเรียนมากกว่าหนึ่งคุณลักษณะ 2) เลือกข้อสอบที่ทำหน้าที่ต่างกันที่มีค่าสัมประสิทธิ์สูง และ 3) คำนึงถึงโครงสร้างของแบบสอบ โดยภายหลังการตัดข้อสอบออกแล้ว ร้อยละข้อสอบในแต่ละสมรรถนะทางวิทยาศาสตร์ต้องไม่แตกต่างจากโครงสร้าง

เดิม และข้อสอบที่ตัดออกไปควรมีจำนวนเป็นร้อยละ 20 ของข้อสอบในแบบสอบนั้น ตามเกณฑ์การตัดข้อสอบที่ผู้วิจัยได้เสนอไว้ในบทที่ 3 (ตารางที่ 13) และผู้วิจัยได้เสนอรายละเอียดของข้อสอบที่ทำหน้าที่ต่างกันในแต่ละคุณลักษณะของตัวแปร และข้อสอบที่ตัดออกในภาคผนวก ข

เมื่อตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปในแบบสอบแต่ละฉบับแล้ว พบว่า จำนวนข้อสอบที่ตัดออกไปในแบบสอบแต่ละฉบับมีประมาณร้อยละ 20 แสดงรายละเอียดดังตารางที่ 24 และแบบสอบหลังตัดมีโครงสร้างคงเดิมกับแบบสอบก่อนตัด ในแต่ละสมรรถนะทางวิทยาศาสตร์ มีแบบสอบบางฉบับที่มีโครงสร้างแตกต่างจากเดิมเล็กน้อย เช่นแบบสอบฉบับที่ 2 และ 3 ทั้งนี้เนื่องจากในแบบสอบฉบับดังกล่าวไม่ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันในสมรรถนะทางวิทยาศาสตร์ จึงทำให้โครงสร้างของข้อสอบแตกต่างไปเล็กน้อย เมื่อเปรียบเทียบความเที่ยงของแบบสอบก่อนตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป ซึ่งมีค่า .734 - .891 และความเที่ยงของแบบสอบหลังตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป ซึ่งมีค่า .686 - .856 จะเห็นได้ว่าความเที่ยงของแบบสอบหลังตัดข้อสอบที่ทำหน้าต่างกันออกไปมีค่าลดลงเล็กน้อย ทั้งนี้เนื่องจากแบบสอบที่มีข้อสอบจำนวนน้อย จึงทำให้มีความเที่ยงน้อยกว่าแบบสอบที่มีข้อสอบจำนวนมาก แสดงรายละเอียดดังตารางที่ 25

ตารางที่ 24 จำนวนข้อสอบที่ทำหน้าที่ต่างกันที่ตัดออกไปในแบบสอบทั้ง 13 ฉบับ

ฉบับที่	สมรรถนะทางวิทยาศาสตร์			รวม	จำนวนข้อสอบหลังการตัด DIF	ร้อยละของข้อสอบที่ตัด
	ISI	EPS	USE			
1	10(2)	33(7)	15(3)	58(12)	46	20.68%
2	5(1)	12(2)	9(0)	26(3)	23	11.53%
3	7(0)	17(3)	7(1)	31(4)	27	12.90%
4	7(1)	18(4)	8(0)	33(5)	28	15.15%
5	15(3)	24(5)	21(4)	60(12)	48	20.00%
6	8(2)	15(3)	9(2)	32(7)	25	21.87%
7	2(0)	8(2)	5(1)	15(3)	12	20.00%
8	6(1)	10(2)	10(2)	26(5)	21	19.23%
9	11(2)	11(1)	6(1)	28(4)	24	14.28%
10	8(2)	12(2)	8(2)	28(6)	22	21.42%
11	5(1)	12(2)	11(0)	28(3)	25	10.71%
12	8(1)	14(3)	8(1)	30(5)	25	16.67%
13	2(0)	8(2)	5(1)	15(3)	12	20.00%

หมายเหตุ ตัวเลขในวงเล็บ เป็นจำนวนข้อสอบที่ทำหน้าที่ต่างกันที่ตัดออก

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 25 การเปรียบเทียบโครงสร้างของแบบสอบก่อนตัดและหลังตัดข้อสอบที่ทำหน้าที่ต่างกัน ในแบบสอบทั้ง 13 ฉบับ

ฉบับที่	ร้อยละของข้อสอบก่อนตัด DIF			ร้อยละของข้อสอบหลังตัด DIF			ความเที่ยง (r_{tt})	
	ISI	EPS	USE	ISI	EPS	USE	ก่อนตัด DIF	หลังตัด DIF
1	17%	57%	26%	17%	57%	26%	.891	.856
2	19%	46%	35%	17%	44%	39%	.751	.735
3	23%	54%	23%	26%	52%	22%	.818	.792
4	21%	55%	24%	21%	50%	29%	.798	.788
5	25%	40%	35%	25%	40%	35%	.880	.856
6	25%	47%	28%	24%	48%	28%	.836	.785
7	13%	54%	33%	17%	50%	33%	.769	.750
8	24%	38%	38%	24%	38%	38%	.781	.742
9	39%	39%	22%	38%	41%	21%	.766	.719
10	29%	42%	29%	29%	42%	29%	.856	.819
11	18%	43%	39%	16%	40%	44%	.688	.676
12	27%	46%	27%	28%	44%	28%	.791	.719
13	13%	54%	33%	17%	50%	33%	.734	.686

เมื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแต่ละฉบับที่เกิดขึ้นจากคุณลักษณะที่ต่างกันของนักเรียนดังสมการวิเคราะห์ HGLM ในระดับที่ 2 ผู้วิจัยจึงใช้ตัวแปรระดับสถานศึกษามาอธิบายตัวแปรระดับนักเรียนที่ทำให้โอกาสในการตอบข้อสอบถูกต่างกัน โดยการวิเคราะห์ผลของตัวแปรในระดับสถานศึกษาที่ทำให้นักเรียนในแต่ละกลุ่มมีโอกาสตอบข้อสอบถูกได้ต่างกัน ซึ่งทำได้ด้วยการเพิ่มตัวแปรคุณลักษณะสถานศึกษาเข้าไปในสมการวิเคราะห์ในระดับที่ 3 ในข้อสอบข้ออ้างอิง และข้อสอบที่ทำหน้าที่ต่างกัน ผู้วิจัยได้สรุปข้อค้นพบจากข้อสอบในแบบสอบทั้ง 13 ฉบับ ที่ตัวแปรระดับสถานศึกษามีผลต่อโอกาสการตอบข้อสอบได้ถูกของนักเรียนที่ต่างกลุ่มกัน โดยนำเสนอให้เห็นเฉพาะทิศทางของสัมประสิทธิ์ถดถอยที่มีนัยสำคัญทางสถิติของตัวแปรระดับสถานศึกษา เพื่ออธิบายการทำหน้าที่ต่างกันของข้อสอบในแต่ละข้อ ส่วนรายละเอียดการวิเคราะห์ที่แสดงขนาดของสัมประสิทธิ์ของตัวแปร แสดงผลในภาคผนวก ข

เมื่อเพิ่มตัวแปรระดับสถานศึกษาเข้าในสมการวิเคราะห์ระดับที่ 3 ในแบบสอบทั้ง 13 ฉบับ พบว่า ส่วนใหญ่ อิทธิพลของตัวแปรระดับสถานศึกษามีทิศทางเดียวกันต่อโอกาสในการตอบข้อสอบถูก ในค่าเฉลี่ยรวมของข้อสอบข้ออ้างอิงหรือโอกาสในการตอบข้อสอบได้ถูกของนักเรียนในทุกสถานศึกษา พบว่าตัวแปรที่มีอิทธิพลในทางบวก หรือทำให้โอกาสในการตอบข้อสอบถูกได้สูงขึ้น ได้แก่ ตัวแปรสถานศึกษาของรัฐ ดัชนีทรัพยากรการเรียน ดัชนีการจัดกิจกรรมการเรียนรู้ นั่นคือการเป็นสถานศึกษาของรัฐ การมีดัชนีทรัพยากรการเรียนสูง และการมีดัชนีการจัดกิจกรรมการเรียนรู้สูง จะส่งผลให้โอกาสในการตอบข้อสอบได้ถูก และตัวแปรที่มีอิทธิพลในทางลบ หรือทำให้

โอกาสในการตอบข้อสอบถูกต่ำลง ได้แก่ งบประมาณของรัฐ สถานศึกษาขนาดกลาง การตั้งอยู่ในเมือง หมู่บ้าน เมืองเล็ก ดัชนีการขาดแคลนครู

เมื่อพิจารณาการทำหน้าที่ต่างกันของข้อสอบที่เกิดจากตัวแปรระดับนักเรียนที่อธิบายได้ด้วยตัวแปรระดับสถานศึกษา พบว่า ตัวแปรระดับสถานศึกษาที่มีค่าอิทธิพลเป็นบวกโดยส่วนใหญ่ ได้แก่ สถานศึกษาของรัฐ สถานศึกษาตั้งในเมืองใหญ่ ดัชนีทรัพยากรการเรียน ดัชนีการจัดกิจกรรมการเรียนรู้ นั่นคือการเป็นสถานศึกษาของรัฐ ตั้งอยู่ในเมืองใหญ่ มีดัชนีทรัพยากรการเรียนสูง และดัชนีการจัดกิจกรรมเรียนรู้สูง จะมีผลต่อคุณลักษณะของนักเรียนที่ทำให้มีโอกาสสูงขึ้นในการตอบข้อสอบได้ถูก ตัวแปรระดับสถานศึกษาที่มีค่าอิทธิพลเป็นลบโดยส่วนใหญ่ ได้แก่ งบประมาณของรัฐ ตั้งอยู่ในเมือง ดัชนีการขาดแคลนครู และแรงกดดันของผู้ปกครอง นั่นคือการที่สถานศึกษาได้รับงบประมาณของรัฐสูง การตั้งอยู่ในเมือง สถานศึกษามีปัญหาการขาดแคลนครูและไม่มีแรงกดดันของผู้ปกครองจะทำให้นักเรียนมีโอกาสต่ำลงในการตอบข้อข้อสอบได้ถูก

แสดงรายละเอียดดังตารางที่ 26

ตารางที่ 26 สรุปผลของตัวแปรระดับสถานศึกษาต่อคุณลักษณะของนักเรียนที่ทำให้โอกาสการตอบข้อสอบที่ทำหน้าที่ต่างกันได้ถูกในแบบสอบทั้ง 13 ฉบับ

ตัวแปร	ตัวแปรระดับสถานศึกษา											
	1	2	3	4	5	6	7	8	9	10	11	12
ข้อสอบข้ออ้างอิง	+	-		-		-	-	-	-	+	+	-
เพศชาย		-		+	+	-			-		+	-
เพศหญิง						-				+	+	
เรียนพิเศษ					+					+		
ไม่เรียนพิเศษ	+											
เศรษฐานะสูง	+	-	+	+					-		+	-
เศรษฐานะต่ำ				-								-
แหล่งทรัพยากรสูง					+	-/+			-			
แหล่งทรัพยากรต่ำ												
ความมั่งคั่งสูง	+			-	+				-			
ความมั่งคั่งต่ำ	-			+	-/+		-			+	+	

หมายเหตุ 1. กำหนดหมายเลข 1 – 12 ด้วยตัวแปรระดับสถานศึกษา ดังนี้

(1) สถานศึกษาของรัฐ (2) งบประมาณของรัฐ (3) สถานศึกษาขนาดใหญ่ (4) สถานศึกษาขนาดกลาง (5) สถานศึกษาตั้งในเมืองใหญ่ (6) สถานศึกษาตั้งในเมือง (7) สถานศึกษาตั้งในหมู่บ้าน (8) สถานศึกษาตั้งในเมืองเล็ก (9) ดัชนีการขาดแคลนครู (10) ดัชนีทรัพยากรการเรียน (11) ดัชนีการจัดกิจกรรมเรียนรู้ (12) แรงกดดันของผู้ปกครอง

2. กำหนดเครื่องหมายแทนความหมายดังนี้

\+ แทนค่าอิทธิพลหรือสัมประสิทธิ์ของตัวแปรที่เป็นบวกซึ่งสอดคล้องกันในแบบสอบทั้ง 13 ฉบับ

– แทนค่าอิทธิพลหรือสัมประสิทธิ์ของตัวแปรที่เป็นลบซึ่งสอดคล้องกันในแบบสอบทั้ง 13 ฉบับ

+/- แทนค่าอิทธิพลหรือสัมประสิทธิ์ของตัวแปรที่เป็นทั้งบวกและลบในแบบสอบทั้ง 13 ฉบับ

2.2 สรุปข้อค้นพบเกี่ยวกับลักษณะของข้อสอบที่ทำหน้าที่ต่างกันในแต่ละคุณลักษณะของนักเรียน

จากผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบในแต่ละฉบับ ผู้วิจัยได้พิจารณาจำนวนข้อสอบที่ทำหน้าที่ต่างกันในแต่ละลักษณะของข้อสอบ 3 ด้านคือ 1) ด้านสมรรถนะทางวิทยาศาสตร์ ซึ่งประกอบด้วยการระบุประเด็นทางวิทยาศาสตร์ การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ การใช้ประจักษ์พยานทางวิทยาศาสตร์ 2) ด้านเป้าหมายการถาม ซึ่งจำแนกเป็นข้อสอบที่มีการถามในระดับบุคคล ระดับสังคม ระดับโลก และ 3) ด้านการใช้ความรู้วิทยาศาสตร์ ซึ่งประกอบด้วยข้อสอบที่วัดในเนื้อหาความรู้วิทยาศาสตร์ และความรู้ทางวิทยาศาสตร์ที่วัดเกี่ยวกับการใช้ความรู้ไปแก้ปัญหาในชีวิตประจำวัน ผลการพิจารณาพบว่าข้อสอบที่ทำหน้าที่ต่างกันเนื่องจากตัวแปรภายนอก มีลักษณะของข้อสอบ ดังนี้

ตัวแปรเพศ ข้อสอบที่ทำหน้าที่ต่างกัน ส่วนใหญ่พบว่านักเรียนชายมีโอกาสทำข้อสอบได้ถูกมากกว่านักเรียนหญิง เมื่อพิจารณาลักษณะของคำถาม พบว่าเป็นข้อสอบที่วัดในสมรรถนะทางวิทยาศาสตร์ในด้านการระบุประเด็นทางวิทยาศาสตร์ การใช้ประจักษ์พยานทางวิทยาศาสตร์ได้ดีกว่านักเรียนหญิง และลักษณะของเป้าหมายการถามเป็นการถามในระดับสังคม

ตัวแปรการเรียนพิเศษ ข้อสอบที่ทำหน้าที่ต่างกัน ส่วนใหญ่พบว่านักเรียนที่เรียนพิเศษมีโอกาสทำข้อสอบได้ถูกมากกว่านักเรียนที่ไม่เรียนพิเศษ เมื่อพิจารณาลักษณะของคำถาม พบว่าส่วนมากเป็นข้อสอบที่วัดในสมรรถนะทางวิทยาศาสตร์ในด้านการระบุประเด็นทางวิทยาศาสตร์ เป็นข้อสอบที่มีเป้าหมายการถามในลักษณะระดับสังคม และวัดในความรู้วิทยาศาสตร์ในด้านเนื้อหามากกว่าความรู้เกี่ยวกับวิทยาศาสตร์หรือความรู้ที่เกี่ยวกับกระบวนการสืบเสาะหาความรู้วิทยาศาสตร์

ตัวแปรเศรษฐานะของครอบครัว ข้อสอบที่ทำหน้าที่ต่างกัน ส่วนใหญ่พบว่านักเรียนที่มีเศรษฐานะของครอบครัวสูงมีโอกาสทำข้อสอบได้ถูกมากกว่านักเรียนที่มีเศรษฐานะของครอบครัวต่ำ เมื่อพิจารณาลักษณะของคำถาม พบว่าส่วนมากเป็นข้อสอบที่วัดในสมรรถนะทางวิทยาศาสตร์ในด้านการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ เป็นข้อสอบที่มีเป้าหมายการถามในลักษณะระดับบุคคล และสังคม และเป็นข้อสอบที่วัดความรู้วิทยาศาสตร์ที่เกี่ยวกับเนื้อหามากกว่าความรู้ทางวิทยาศาสตร์ที่มีการใช้ความรู้ในการแก้ปัญหาหรือใช้กระบวนการสืบเสาะหาความรู้วิทยาศาสตร์

ตัวแปรทรัพยากรของครอบครัว ข้อสอบที่ทำหน้าที่ต่างกัน ส่วนใหญ่พบว่านักเรียนที่มีทรัพยากรของครอบครัวต่ำมีโอกาสทำข้อสอบได้ถูกมากกว่านักเรียนที่มีทรัพยากรของครอบครัวสูง เมื่อพิจารณาลักษณะของคำถาม พบว่าส่วนมากเป็นข้อสอบที่วัดในสมรรถนะทางวิทยาศาสตร์ด้านการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ การใช้ประจักษ์พยานทางวิทยาศาสตร์ และเป็นข้อสอบที่มีเป้าหมายการถามในลักษณะระดับสังคม

ตัวแปรความมั่งคั่งของครอบครัว ข้อสอบที่ทำหน้าที่ต่างกัน ส่วนใหญ่พบว่านักเรียนที่มีความมั่งคั่งของครอบครัวสูงมีโอกาสทำข้อสอบได้ถูกมากกว่านักเรียนที่มีความมั่งคั่งของครอบครัวต่ำ เมื่อพิจารณาลักษณะของคำถาม พบว่าส่วนมากเป็นข้อสอบที่มีเป้าหมายการถามในลักษณะ

ระดับสังคม เป็นคำถามในสมรรถนะทางวิทยาศาสตร์ด้านการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ และการใช้ประจักษ์พยานทางวิทยาศาสตร์ และเป็นความรู้วิทยาศาสตร์ที่เกี่ยวกับเนื้อหามากกว่าความรู้เกี่ยวกับวิทยาศาสตร์หรือความรู้ที่เกี่ยวกับกระบวนการสืบเสาะหาความรู้วิทยาศาสตร์

แสดงรายละเอียดจำนวนข้อสอบที่ค้นพบในแต่ละคุณลักษณะของนักเรียนและลักษณะของข้อสอบ ในตารางที่ 27

ตารางที่ 27 สรุปลักษณะของข้อสอบที่ทำหน้าที่ต่างกันในแต่ละคุณลักษณะของนักเรียน

คุณลักษณะ	สมรรถนะทางวิทยาศาสตร์				เป้าหมายของคำถาม				ความรู้วิทยาศาสตร์		
	ISI	EPS	USE	รวม	PE	SO	GL	รวม	KOS	KAS	รวม
ชาย	26 (38.2%)	24 (35.3%)	18 (26.5%)	68 (100%)	19 (27.9%)	42 (61.8%)	7 (10.3%)	68 (100%)	31 (45.6%)	37 (54.4%)	68 (100%)
หญิง	-	10 (83.3%)	2 (16.7%)	12 (100%)	6 (50%)	3 (25.9%)	3 (25.9%)	12 (100%)	10 (83.3%)	2 (16.7%)	12 (100%)
เรียน	8 (44.4%)	5 (27.8%)	5 (27.8%)	18 (100%)	4 (22.2%)	11 (61.1%)	3 (16.7%)	18 (100%)	7 (38.9%)	11 (61.1%)	18 (100%)
ไม่เรียน	-	2 (66.7%)	1 (33.3%)	3 (100%)	3 (100%)	-	-	3 (100%)	2 (66.7%)	1 (33.3%)	3 (100%)
SES สูง	5 (13.9%)	21 (58.3%)	10 (27.8%)	36 (100%)	13 (36.1%)	13 (36.1%)	10 (27.8%)	36 (100%)	23 (63.9%)	13 (36.1%)	36 (100%)
SES ต่ำ	2 (12.5%)	9 (56.3%)	5 (31.3%)	16 (100%)	10 (62.5%)	3 (18.8%)	3 (18.8%)	16 (100%)	10 (62.5%)	6 (37.5%)	16 (100%)
ทรัพยากรสูง	1 (16.7%)	2 (33.3%)	3 (50%)	6 (100%)	1 (16.7%)	4 (66.7%)	1 (16.7%)	6 (100%)	4 (66.7%)	2 (33.3%)	6 (100%)
ทรัพยากรต่ำ	5 (29.4%)	6 (35.3%)	6 (35.3%)	17 (100%)	5 (29.4%)	11 (64.7%)	1 (5.9%)	17 (100%)	7 (41.2%)	10 (58.8%)	17 (100%)
มั่งคั่งสูง	2 (7.4%)	16 (59.3%)	9 (33.3%)	27 (100%)	10 (37%)	11 (40.7%)	6 (22.2%)	27 (100%)	23 (85.2%)	4 (14.8%)	27 (100%)
มั่งคั่งต่ำ	3 (15.8%)	11 (57.9%)	5 (26.3%)	19 (100%)	3 (15.8%)	12 (63.2%)	4 (21.1%)	19 (100%)	13 (68.4%)	6 (31.6%)	19 (100%)

หมายเหตุ

1) สมรรถนะทางวิทยาศาสร์ กำหนดสัญลักษณ์ดังนี้
   การระบุประเด็นทางวิทยาศาสตร์ แทนด้วย ISI : Identifying scientific issues
   การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ แทนด้วย EPS : Explain phenomena scientifically
   การใช้ประจักษ์พยานทางวิทยาศาสตร์ แทนด้วย USE : Using scientific evidence
2) เป้าหมายของคำถาม กำหนดสัญลักษณ์ดังนี้
   ระดับบุคคล แทนด้วย PE : Personal
   ระดับสังคม แทนด้วย SO : Socail
   ระดับโลก แทนด้วย GL : Global
3) ความรู้วิทยาศาสร์ กำหนดสัญลักษณ์ดังนี้
   ความรู้วิทยาศาสตร์ แทนด้วย KOS : Knowledge Of Science
   ความรู้ทางวิทยาศาสตร์ แทนด้วย KOS : Knowledge About Science

2.3 ผลการประมาณค่าความยากของข้อสอบและค่าความสามารถของผู้สอบ

การประมาณค่าพารามิเตอร์ความยากข้อสอบและความสามารถของผู้สอบ ผู้วิจัยประยุกต์ใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น ด้วยการวิเคราะห์ 3 ระดับ โดยระดับที่ 1 เป็นระดับข้อสอบ ระดับที่ 2 เป็นระดับนักเรียน และระดับที่ 3 เป็นระดับสถานศึกษา จากสมการทั้ง 3 ระดับ สามารถหาค่าพารามิเตอร์แต่ละค่าได้ดังนี้

ค่าพารามิเตอร์ของข้อสอบหรือค่าความยากของข้อสอบ (δ_i) ; $(-\pi_{q00} - \pi_{000})$

ค่าพารามิเตอร์ของผู้สอบหรือค่าความสามารถของผู้สอบ (θ_j) ; $u_{00m} + r_{0jm}$

1) ผลการประมาณค่าความยากของข้อสอบ

จากการประมาณค่าความยากของข้อสอบในแบบสอบแต่ละฉบับ โดยภาพรวมข้อสอบทั้ง 13 ฉบับ มีค่าความยากโดยเฉลี่ยเท่ากับ .475898 ส่วนเบี่ยงเบนมาตรฐาน 26.53 ค่าความยากต่ำสุดเท่ากับ -2.069487 และค่าความยากสูงสุดเท่ากับ 2.866956 ลักษณะการกระจายเป็นแบบเบ้ซ้าย (skewness = -.005) และค่อนข้างโด่ง (kurtosis = .430) เมื่อพิจารณาข้อสอบในแต่ละฉบับพบว่า ข้อสอบในแบบสอบที่มีค่าความยากโดยเฉลี่ยต่ำสุดหรือง่ายที่สุด เป็นข้อสอบในแบบสอบฉบับที่ 7 (Mean = .149310, S.D. = .891339) และข้อสอบในแบบสอบที่มีค่าความยากโดยเฉลี่ยสูงสุดหรือยากที่สุด เป็นข้อสอบในแบบสอบฉบับที่ 8 (Mean = .840691, S.D. = .729703) ดังแสดงในตารางที่ 28 ส่วนค่าความยากของข้อสอบรายข้อ แสดงรายละเอียดดังภาคผนวก ค

ตารางที่ 28 ค่าสถิติพื้นฐานของค่าความยากของข้อสอบในแบบสอบฉบับที่ 1 ถึง 13

ฉบับ	Min	Max	Range	Mean	S.E	S.D	Skewness	Kurtosis
1	-1.161613	1.973651	3.135264	.309952	.096838	.724672	.148	-.374
2	-1.420448	2.608134	4.028582	.246914	.190586	.914022	.322	.716
3	-1.507635	2.438143	3.945778	.244974	.148649	.800499	.532	1.226
4	-1.962271	2.019146	3.981417	.294758	.143586	.799458	-.248	1.469
5	-2.069487	2.713777	4.783264	.460216	.120896	.904703	-.427	.940
6	-1.335097	1.970359	3.305456	.499478	.134745	.762236	-.366	.076
7	-1.753256	1.546286	3.299542	.249310	.230142	.891339	-.697	.508
8	-.404706	2.386167	2.790873	.540691	.143106	.729703	.292	-.402
9	-1.305132	2.866956	4.172088	.399658	.195418	.977091	.670	.316
10	-.935177	2.355651	3.290828	.593823	.159156	.842174	.485	-.442
11	-1.824356	2.763548	4.587904	.502530	.188558	.942791	-.193	1.413
12	-1.401295	2.664999	4.066294	.589036	.168413	.875103	.158	.617
13	-1.570104	1.837204	3.407308	.216778	.240502	.931460	-.580	.012
รวม	-2.069487	2.866956	4.936443	.475898	.043008	.847169	-.005	.430

2) ผลการประมาณค่าความสามารถของผู้สอบ

การประมาณค่าความสามารถของผู้สอบ ผู้วิจัยได้ดำเนินการวิเคราะห์ 2 ครั้ง คือ ครั้งที่ 1 เป็นการวิเคราะห์ค่าความสามารถของผู้สอบที่ได้จากการวิเคราะห์อิทธิพลของข้อสอบทุกข้อในแบบสอบ และครั้งที่ 2 เป็นการวิเคราะห์ความสามารถของผู้สอบจากการวิเคราะห์ในระดับข้อสอบ ที่นำข้อสอบที่ทำหน้าที่ต่างกันเข้ามาวิเคราะห์ด้วย ดังนั้นผู้สอบแต่ละคนจะถูกประมาณค่าสองครั้ง นั่นคือค่าความสามารถของผู้สอบก่อนการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป และค่าความสามารถของผู้สอบหลังการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป

จากการประมาณค่าความสามารถของผู้สอบก่อนการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป พบว่า โดยภาพรวมทั้ง 13 ฉบับ มีค่าความสามารถโดยเฉลี่ยเท่ากับ .02225 ส่วนเบี่ยงเบนมาตรฐาน .781813 ผู้สอบมีค่าความสามารถต่ำสุดเท่ากับ -2.525 และค่าความสามารถสูงสุดเท่ากับ 3.333 ลักษณะการกระจายเป็นแบบเบ้ขวา (skewness = .387) และค่อนข้างโด่ง (kurtosis = .179) เมื่อพิจารณาความสามารถของกลุ่มผู้สอบที่ทำแบบสอบในแต่ละฉบับ พบว่ากลุ่มผู้สอบที่ทำแบบสอบฉบับที่ 4 มีความสามารถของผู้สอบโดยเฉลี่ยต่ำสุด (Mean = .00853 , S.D. = .687716) และกลุ่มผู้สอบที่ทำแบบสอบฉบับที่ 7 มีค่าเฉลี่ยของความสามารถของผู้สอบสูงสุด (Mean = .03531, S.D. = .910717) นอกจากนี้ยังพบว่ากลุ่มผู้สอบที่ทำแบบสอบในแต่ละฉบับมีลักษณะการกระจายค่าความสามารถแบบเบ้ขวา นั่นคือส่วนใหญ่มีความสามารถต่ำกว่าค่าความสามารถเฉลี่ย แสดงรายละเอียดดังตารางที่ 29

ตารางที่ 29 ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบ**ก่อนการตัด**ข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบฉบับที่ 1 ถึง 13

ฉบับ	n	Min	Max	Range	Mean	S.E	S.D	Skewness	Kurtosis
1	474	-2.021	2.466	4.487	.02578	.036106	.786073	.511	-.028
2	474	-1.673	2.110	3.783	.01794	.031733	.690866	.223	-.258
3	468	-1.672	2.532	4.204	.01361	.036654	.792954	.418	-.070
4	466	-1.761	1.947	3.708	.00853	.031858	.687716	.355	.021
5	469	-2.525	2.476	5.001	.03430	.037121	.803899	.264	.364
6	483	-2.048	2.503	4.551	.01795	.036509	.802368	.485	.165
7	488	-1.962	2.578	4.540	.03531	.041226	.910717	.292	-.396
8	478	-1.728	2.344	4.072	.02177	.034657	.757704	.481	.214
9	473	-1.873	2.169	4.042	.02550	.032187	.700028	.143	-.280
10	475	-2.254	3.333	5.587	.02615	.045295	.987190	.422	.096
11	482	-1.405	1.566	2.971	.01636	.025283	.555071	.257	-.081
12	486	-2.003	2.525	4.528	.01454	.036847	.812304	.314	.114
13	476	-1.720	2.729	4.449	.03122	.036436	.794939	.518	-.150
รวม	6,192	5.858	-2.525	3.333	.02225	.009935	.781813	.387	.179

การประมาณค่าความสามารถของผู้สอบหลังการตัดข้อสอบที่ทำหน้าที่ต่างกันออก พบว่าโดยภาพรวมทั้ง 13 ฉบับ มีค่าความสามารถของผู้สอบโดยเฉลี่ยเท่ากับ .02147 ส่วนเบี่ยงเบนมาตรฐาน .746666 ผู้สอบมีค่าความสามารถต่ำสุดเท่ากับ -2.338 และค่าความสามารถสูงสุดเท่ากับ 3.221 ลักษณะการกระจายเป็นแบบเบ้ขวา (skewness = .404) และค่อนข้างโด่ง (kurtosis = .175) เมื่อพิจารณาความสามารถของผู้สอบที่ทำแบบสอบแต่ละฉบับ เมื่อตัดข้อสอบที่ทำหน้าที่ต่างกันออก พบว่าความสามารถของผู้สอบที่ทำแบบสอบแต่ละฉบับมีค่าลดลง นั่นคือความคลาดเคลื่อนจากการประมาณค่าลดลง แสดงรายละเอียดดังตารางที่ 30

ตารางที่ 30 ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบ**หลังการตัด**ข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบฉบับที่ 1 ถึง 13

ฉบับที่	n	Min	Max	Range	Mean	S.E	S.D	Skewness	Kurtosis
1	474	-1.946	2.390	4.336	.02497	.034808	.757821	.516	-.016
2	474	-1.650	2.071	3.721	.01771	.031282	.681052	.215	-.268
3	468	-1.590	2.450	4.040	.01328	.035187	.761218	.445	-.031
4	466	-1.741	1.915	3.656	.00844	.031436	.678600	.347	.011
5	469	-2.338	2.343	4.681	.03225	.034770	.753002	.292	.382
6	483	-1.911	2.339	4.250	.01698	.034145	.750418	.482	.162
7	488	-1.736	2.420	4.156	.03360	.037888	.836962	.364	-.369
8	478	-1.636	2.233	3.869	.02095	.032932	.719990	.484	.216
9	473	-1.789	2.105	3.894	.02480	.031042	.675114	.165	-.276
10	475	-2.179	3.221	5.400	.02529	.043696	.952332	.419	.103
11	482	-1.320	1.513	2.833	.01595	.024212	.531553	.291	-.077
12	486	-1.851	2.391	4.242	.01388	.034509	.760773	.350	.137
13	476	-1.685	2.684	4.369	.03078	.035726	.779456	.519	-.143
รวม	6,192	-2.338	3.221	5.559	.02147	.009489	.746666	.404	.175

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3) สรุปผลการแปลงค่าคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์

คะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ตามโครงการประเมินผลนักเรียน PISA จะใช้วิธีคำนวณแปลงค่าความสามารถของนักเรียน ในการวิจัยครั้งนี้ใช้ค่าความสามารถของนักเรียนจากการประมาณของโมเดล HGLM ซึ่งมีความสอดคล้องกับโมเดลราสช์ และใช้ค่าความสามารถที่ได้มาแปลงเป็นคะแนนสเกลของ PISA ซึ่งอยู่ในรูปคะแนนมาตรฐานที่มีค่าเฉลี่ยเท่ากับ 500 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 100 (ตามรายละเอียดในบทที่ 3) ในการแปลงคะแนนสเกลของการประเมินการรู้เรื่องด้านวิทยาศาสตร์ มีการแปลงคะแนนจากค่าความสามารถของนักเรียนสองครั้งคือ ก่อนและหลังการตัดข้อสอบที่ทำหน้าที่ต่างกัน (DIF) ออกไป

พบว่า คะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์หลังการตัด DIF มีค่าเฉลี่ยต่ำกว่าคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ก่อนการตัด DIF ในทั้งแบบสอบ 13 ฉบับ คะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ก่อนการตัด DIF ในแบบสอบฉบับที่ 3 มีคะแนนเฉลี่ยต่ำสุดเท่ากับ 484.51 และในแบบสอบฉบับที่ 5 มีคะแนนเฉลี่ยสูงสุดเท่ากับ 486.44 คะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์หลังการตัด DIF ในแบบสอบฉบับที่ 9 มีคะแนนเฉลี่ยต่ำสุดเท่ากับ 482.28 และในแบบสอบฉบับที่ 10 มีคะแนนเฉลี่ยสูงสุดเท่ากับ 485.84 แสดงรายละเอียดดังตารางที่ 31

ตารางที่ 31 สรุปคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ ก่อนและหลังการตัดข้อสอบการทำหน้าที่ต่างกันออก

ฉบับที่	n	คะแนนก่อนตัด DIF				คะแนนหลังตัด DIF			
		Min	Max	Mean	S.D	Min	Max	Mean	S.D
1	474	294.79	713.19	485.64	73.30	323.41	668.90	482.41	55.15
2	474	327.24	680.00	484.82	64.38	286.67	689.32	482.98	51.93
3	468	327.33	719.35	484.51	73.94	303.74	696.78	483.72	62.87
4	466	319.03	664.80	484.03	64.12	333.86	625.17	483.08	51.75
5	469	247.79	714.13	486.44	74.96	228.02	679.53	483.20	64.31
6	483	292.27	716.64	484.91	74.81	305.04	701.35	484.82	69.97
7	488	300.29	723.64	485.53	84.92	321.36	708.91	485.37	78.04
8	478	322.11	701.82	485.27	70.65	330.69	691.47	485.19	67.13
9	473	308.59	685.50	485.62	65.27	297.21	669.27	482.28	60.72
10	475	273.06	794.04	485.93	91.98	280.05	783.60	485.84	88.73
11	482	352.23	629.27	484.76	51.75	257.86	671.33	482.51	59.21
12	486	296.47	718.70	484.59	75.74	167.73	780.26	483.12	70.24
13	476	322.86	737.72	485.15	74.12	326.12	733.52	485.11	72.68
รวม	6,192	247.79	794.04	485.33	72.88	167.73	780.26	484.29	69.32

นอกจากนี้ ผู้วิจัยได้ตรวจความสอดคล้องของคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ตามโครงการ PISA และคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ก่อนการตัด DIF ซึ่งผู้วิจัยใช้โมเดล HGLM ในการประมาณค่าความสามารถนักเรียน แต่คะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ตามโครงการของ PISA ใช้การประมาณค่าตามโมเดลราสช์ พบว่า คะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์มีค่าสหสัมพันธ์กันสูงถึง .913 และเมื่อเขียนกราฟระหว่างคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ตามโครงการ PISA คะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ก่อนการตัด DIF พบว่ามีเส้นกราฟเป็นเส้นตรง แสดงว่าคะแนนทั้งสองมีความสอดคล้องกัน

ตอนที่ 3 ผลการวิเคราะห์ตามโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน

การวิเคราะห์โมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน ประกอบด้วย 4 โมเดล ได้แก่ โมเดล 1 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์มูลค่าเพิ่มและไม่มีการปรับแก้คะแนนหรือควบคุมอิทธิพลตัวแปรระดับนักเรียนและสถานศึกษา โมเดล 2 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์มูลค่าเพิ่มและมีการปรับแก้คะแนนหรือควบคุมอิทธิพลตัวแปรระดับนักเรียนและสถานศึกษา โมเดล 3 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์มูลค่าเพิ่ม และไม่มีการปรับแก้คะแนนหรือควบคุมอิทธิพลตัวแปรระดับนักเรียนและสถานศึกษา โมเดล 4 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์มูลค่าเพิ่ม และมีการปรับแก้คะแนนหรือควบคุมอิทธิพลตัวแปรระดับนักเรียนและสถานศึกษา

การนำเสนอในตอนนี้จะแสดงให้เห็นถึงสมการวิเคราะห์ในแต่ละโมเดล และอิทธิพลของตัวแปรระดับนักเรียนและระดับสถานศึกษาที่มีต่อผลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์

1) การวิเคราะห์ตามโมเดล 1

โมเดล 1 เป็นโมเดลที่ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ และวิเคราะห์มูลค่าเพิ่มโดยไม่มีการปรับแก้ (Unadjusted) คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา นั่นคือการวิเคราะห์ในโมเดล 1 จะไม่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาเข้ามาเป็นตัวแปรในการวิเคราะห์พหุระดับเพื่อหาค่าเศษเหลือ (residual) ของสถานศึกษา

การวิเคราะห์ตามโมเดลพหุระดับที่ไม่มีตัวแปรทำนายในทุกระดับ จัดเป็นการวิเคราะห์โมเดลพหุระดับแบบไม่มีเงื่อนไขสมบูรณ์ (Fully Unconditional Models) (ศิริชัย กาญจนวาสี, 2550) หรือการวิเคราะห์โมเดลศูนย์ (Null Models) ซึ่งการวิเคราะห์ในขั้นตอนนี้จะไม่นำตัวแปรทำนายเข้ามาร่วมในการวิเคราะห์ และสามารถตรวจสอบความผันแปรของคะแนนผลการประเมินวิชาวิทยาศาสร์ในสถานศึกษาและระหว่างสถานศึกษา ซึ่งมีโมเดลในการวิเคราะห์ ดังนี้

<u>ระดับที่ 1</u> ภายในสถานศึกษา

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

<u>ระดับที่ 2</u> ระหว่างสถานศึกษา

$$B0_k = G00 + U0_k$$

2) การวิเคราะห์ตามโมเดล 2

โมเดล 2 เป็นโมเดลที่ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์มูลค่าเพิ่ม และมีการปรับแก้ (adjusted) คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา นั่นคือ มีการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา เพื่อหาคะแนนเศษเหลือ (residual) ของสถานศึกษาหรือมูลค่าเพิ่มของสถานศึกษา มีสมการวิเคราะห์ ดังนี้

ระดับที่ 1 : ระดับนักเรียน(ระหว่างนักเรียน i ภายในสถานศึกษา k)

$$SCIE_{ik} = B0_k + B1_k*(RELAGR)_{ik} + B2_k*(MALE)_{ik} + B3_k*(ST31Q02)_{ik} + B4_k*(ST31Q03)_{ik} + B5_k*(ESCS_1)_{ik} + B6_k*(HEDRES_1)_{ik} + B7_k*(WEALTH_1)_{ik} + B8_k*(INTSCI_A)_{ik} + B9_k*(SCIEEF_A)_{ik} + B10_k*(SCSCIE_1)_{ik} + B11_k*(JOYSCI_A)_{ik} + B12_k*(INSTSC_A)_{ik} + B13_k*(GENSCI_A)_{ik} + R_{ik}$$

ระดับที่ 2 : ระดับหน่วย(ระหว่างสถานศึกษา)

$$B0 = G00 + G01*(DSCHL) + G02*(DSCHM) + G03*(TYPE1) + G04*(TYPE2) + G05*(TYPE3) + G06*(TYPE4) + G07*(COMMS) + G08*(COMMM) + G09*(COMML) + G010*(COMMXL) + G011*(SC06Q01) + G012*(SC16Q01) + G013*(SC14Q01) + G014*(SC14Q04) + G015*(SC03Q01) + G016*(TCSHORT) + G017*(SCMATEDU) + G018*(STRATIO) + G019*(SCIPROM) + G020*(RESPRES) + G021*(RESPCURR) + G022*(SC03Q03) + U0$$

$$B1 = G10$$

$$B2 = G20$$

$$B3 = G30$$

$$B4 = G40$$

$$B5 = G50$$

$$B6 = G60$$

$$B7 = G70$$

$$B8 = G80$$

$$B9 = G90$$

$$B10 = G100$$

$$B11 = G110$$

$$B12 = G120$$

$$B13 = G130$$

3) การวิเคราะห์ตามโมเดล 3

โมเดล 3 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ และการวิเคราะห์มูลค่าเพิ่มเป็นโมเดลที่ยังไม่มีการปรับแก้ (Unadjusted) คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา นั่นคือคะแนนผลการประเมินวิชาวิทยาศาสร์ที่นำมาใช้เป็นตัวแปรตามในโมเดล 3 เป็นคะแนนที่ได้จากการตัด

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ข้อสอบที่ทำหน้าที่ต่างกันออกจากแบบสอบก่อนนำผลคะแนนที่ได้มาวิเคราะห์พหุระดับเพื่อหาคะแนนเศษเหลือ(residual) ของสถานศึกษา แต่ไม่มีการนำตัวแปรต้นในระดับนักเรียนและระดับสถานศึกษาเข้ามาวิเคราะห์ ซึ่งมีโมเดลในการวิเคราะห์ ดังนี้

ระดับที่ 1 ภายในสถานศึกษา

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

ระดับที่ 2 ระหว่างสถานศึกษา

$B0_k = G00 + U0_k$

4) การวิเคราะห์ตามโมเดล 4

โมเดล 4 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาในวิชาวิทยาศาสตร์ที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์ ส่วนโมเดลการวิเคราะห์มูลค่าเพิ่มเป็นโมเดลที่มีการปรับแก้ (adjusted) ด้วยตัวแปรระดับนักเรียนและสถานศึกษา นั่นคือโมเดล 4 เป็นโมเดลที่ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และตัดข้อสอบที่ทำหน้าที่ต่างกันออกจากแบบสอบก่อนนำผลคะแนนที่ได้มาวิเคราะห์พหุระดับเพื่อหาคะแนนเศษเหลือ(residual) ของสถานศึกษา และมีการนำตัวแปรต้นในระดับนักเรียนและระดับสถานศึกษาเข้ามาวิเคราะห์ มีสมการวิเคราะห์ดังนี้

ระดับที่ 1 : ระดับนักเรียน(ระหว่างนักเรียน i ภายในสถานศึกษา k)

$$SCIE_{ik} = B0_k + B1_k*(RELAGR)_{ik} + B2_k*(MALE)_{ik} + B3_k*(ST31Q02)_{ik} + B4_k*(ST31Q03)_{ik} + B5_k*(ESCS_1)_{ik} + B6_k*(HEDRES_1)_{ik} + B_k*(WEALTH_1)_{ik} + B8_k*(INTSCI_A)_{ik} + B9_k*(SCIEEF_A)_{ik} + B10_k*(SCSCIE_1)_{ik} + B11_k*(JOYSCI_A)_{ik} + B12_k*(INSTSC_A)_{ik} + B13_k*(GENSCI_A)_{ik} + R_{ik}$$

ระดับที่ 2 : ระดับหน่วย(ระหว่างสถานศึกษา)

$$B0 = G00 + G01*(DSCHL) + G02*(DSCHM) + G03*(TYPE1) + G04*(TYPE2) + G05*(TYPE3) + G06*(TYPE4) + G07*(COMMS) + G08*(COMMM) + G09*(COMML) + G010*(COMMXL) + G011*(SC06Q01) + G012*(SC16Q01) + G013*(SC14Q01) + G014*(SC14Q04) + G015*(SC03Q01) + G016*(TCSHORT) + G017*(SCMATEDU) + G018*(STRATIO) + G019*(SCIPROM) + G020*(RESPRES) + G021*(RESPCURR) + G022*(SC03Q03) + U0$$

$B1 = G10$

$B2 = G20$

$B3 = G30$

$B4 = G40$

$B5 = G50$

$B6 = G60$

$B7 = G70$

$B8 = G80$

B9 = G90

B10 = G100

B11 = G110

B12 = G120

B13 = G130

โดยที่ $SCIE_{ik}$= คะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ของนักเรียนที่ i ภายในสถานศึกษา k

$B0_k$ = ค่าจุดตัดแกน(intercept)ที่เป็นค่าแสดงถึงคะแนนเฉลี่ยของคะแนนผลการประเมินวิชาวิทยาศาสตร์ในโครงการ PISA 2006 ในสถานศึกษา k

R_{ik} = ค่าเศษเหลือหรือค่าความความคลาดเคลื่อน (residual error)

G00= ค่าจุดตัดแกน(intercept) ที่เป็นค่าแสดงถึงคะแนนเฉลี่ยของคะแนนผลการประเมินวิชาวิทยาศาสตร์ในโครงการ PISA 2006 ในสถานศึกษาทุกโรง

G01 = ค่าความชัน(slope) ที่แสดงถึงค่าอิทธิพลของการเปลี่ยนแปลงของคะแนนผลการประเมินวิชาวิทยาศาสตร์ในโครงการ PISA 2006 เมื่อมีการควบคุมตัวแปร Y_k

$U0_k$ = ค่าเศษเหลือ ซึ่งเป็นค่าที่แสดงถึงคะแนนมูลค่าเพิ่มของสถานศึกษา

ผู้วิจัยเสนอสรุปผลการวิเคราะห์ในแต่ละโมเดล ส่วนรายละเอียดผู้วิจัยได้เสนอไว้ในภาคผนวก ง สรุปผลจากโมเดล 1 ถึงโมเดล 4 ดังนี้

ผลการวิเคราะห์ในโมเดล 1 จากผลการวิเคราะห์พหุระดับในโมเดล 1 ผลการทดสอบอิทธิพลคงที่ (Fixed effcet) พบว่า ค่าเฉลี่ยของคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ทุกสถานศึกษา (G00) มีความผันแปรจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (G00 = 480.157757 ; t = 151.137; p=0.000) ผลการทดสอบอิทธิพลเชิงสุ่ม (random effect) พบว่า มีความแปรปรวนของคะแนนผลการประเมินวิชาวิทยาศาสตร์ ระหว่างสถานศึกษา ($U0_k$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (χ^2 = 4022.97893 ; p=0.000) โดยมีความแปรปรวนระหว่างสถานศึกษาเท่ากับ 2010.42084 ซึ่งอธิบายความแปรปรวนได้ร้อยละประมาณ 37.83 และความแปรปรวนภายในสถานศึกษามีค่าเท่ากับ 3304.54876 ซึ่งอธิบายความแปรปรวนได้ร้อยละประมาณ 62.17

ผลการวิเคราะห์ในโมเดล 2 จากผลการวิเคราะห์พหุระดับที่มีการควบคุมตัวแปรคุณลักษณะระดับนักเรียนและระดับสถานศึกษา ผลการทดสอบอิทธิพลคงที่ พบว่า ค่าเฉลี่ยของคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ทุกสถานศึกษา (G00) มีความผันแปรจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ค่าสัมประสิทธิ์ถดถอยของตัวแปรระดับสถานศึกษาที่มีค่าเป็นบวกสูงสุด คือ สถานศึกษาสังกัด สกอ.หรือโรงเรียนสาธิต (β = 54.316148) นั่นคือความเป็นสถานศึกษาสาธิตจะทำให้คะแนนผลการประเมินวิชาวิทยาศาสตร์มีค่าเพิ่มขึ้น 54.316148

คะแนน รองลงมาคือ สถานศึกษาขนาดใหญ่ (β = 24.891340) สถานศึกษาขนาดกลาง (β = 19.455240) ส่วนค่าสัมประสิทธิ์ถดถอยของตัวแปรระดับสถานศึกษาที่มีค่าเป็นลบสูงสุด คือ สถานศึกษาที่ตั้งอยู่ในเมืองเล็ก (β = -34.307281) อันดับต่อมาคือสถานศึกษาที่ตั้งอยู่ในหมู่บ้าน (β = -29.778053) และสถานศึกษาสังกัด สช. (β = -24.336141) ส่วนตัวแปรระดับนักเรียนที่มีอิทธิพลทางบวกสูงสุดคือ ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา (β = 15.377367) อันดับต่อมาคือ ความเพลิดเพลินทางวิทยาศาสตร์ (β = 8.992101) และระดับการใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง (β = 8.397479) ส่วนตัวแปรระดับนักเรียนที่มีอิทธิพลทางลบสูงสุดคือความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์ (β = -8.038848) รองลงมาคือความมั่งคั่งของครอบครัว (β = -4.145535)

ผลการทดสอบอิทธิพลเชิงสุ่ม พบว่า ค่าเศษเหลือระดับสถานศึกษาของคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ที่มีการควบคุมตัวแปรระดับนักเรียนและสถานศึกษา (U0) หรือมูลค่าเพิ่มของสถานศึกษามีความผันแปรระหว่างสถานศึกษาอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (χ^2 = 1323.16135 ; p=0.000) โดยมีความแปรปรวนเท่ากับ 616.57585 ซึ่งอธิบายความแปรปรวนได้ร้อยละประมาณ 16.93 และความแปรปรวนภายในสถานศึกษามีค่าเท่ากับ 3026.18439 ซึ่งอธิบายความแปรปรวนได้ร้อยละประมาณ 83.07

ผลการวิเคราะห์ในโมเดล 3 จากผลการวิเคราะห์พหุระดับในโมเดล 3 ผลการทดสอบอิทธิพลคงที่ พบว่า ค่าเฉลี่ยของคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ทุกสถานศึกษา (G00) มีความผันแปรจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (G00 = 480.283815; t = 157.594 ; p=0.000) ผลการทดสอบอิทธิพลเชิงสุ่ม พบว่า มีความแปรปรวนของคะแนนผลการประเมินวิชาวิทยาศาสตร์ ระหว่างสถานศึกษา ($U0_k$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (χ^2 = 4085.08383 ; p=0.000) โดยมีความแปรปรวนระหว่างสถานศึกษาเท่ากับ 1851.82608 ซึ่งอธิบายความแปรปรวนได้ร้อยละประมาณ 38.19 และความแปรปรวนภายในสถานศึกษามีค่าเท่ากับ 2996.81434 ซึ่งอธิบายความแปรปรวนได้ร้อยละประมาณ 61.81

ผลการวิเคราะห์ในโมเดล 4 จากผลการวิเคราะห์พหุระดับที่มีการควบคุมตัวแปรคุณลักษณะระดับนักเรียนและระดับสถานศึกษา ผลการทดสอบอิทธิพลคงที่ พบว่า ค่าเฉลี่ยของคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ทุกสถานศึกษา (G00) มีความผันแปรจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ค่าสัมประสิทธิ์ถดถอยของตัวแปรระดับสถานศึกษาที่มีค่าเป็นบวกสูงสุด คือ สถานศึกษาสังกัด สกอ.หรือโรงเรียนสาธิต (β = 52.409790) รองลงมาคือ สถานศึกษาขนาดใหญ่ (β = 23.849150) สถานศึกษาขนาดกลาง (β = 18.589401) ส่วนค่าสัมประสิทธิ์ถดถอยของตัวแปรระดับสถานศึกษาที่มีค่าเป็นลบสูงสุด คือสถานศึกษาที่ตั้งอยู่ในเมืองเล็ก (β = -32.883962) อันดับต่อมาคือสถานศึกษาที่ตั้งอยู่ในหมู่บ้าน (β = -28.640284) และสถานศึกษาสังกัด สช. (β = -23.123298) ส่วนตัวแปรระดับนักเรียนที่มีอิทธิพลทางบวกสูงสุดคือ

ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา (β = 14.586167) อันดับต่อมาคือ ความเพลิดเพลินทางวิทยาศาสตร์ (β = 8.550603) และระดับการใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง (β = 8.003062) ส่วนตัวแปรระดับนักเรียนที่มีอิทธิพลทางลบสูงสุดคือความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์ (β = -7.685855) รองลงมาคือความมั่งคั่งของครอบครัว (β = -3.944037)

ผลการทดสอบอิทธิพลเชิงสุ่ม (random effect) พบว่า ค่าเศษเหลือระดับสถานศึกษาของคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ที่มีการควบคุมตัวแปรระดับนักเรียนและสถานศึกษา (U0) หรือมูลค่าเพิ่มของสถานศึกษามีความผันแปรระหว่างสถานศึกษาอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (χ^2 = 1344.31075 ; p=0.000) โดยมีความแปรปรวนเท่ากับ 569.18444 ซึ่งอธิบายความแปรปรวนได้ร้อยละประมาณ 17.18 และความแปรปรวนภายในสถานศึกษามีค่าเท่ากับ 2744.59225 ซึ่งอธิบายความแปรปรวนได้ร้อยละประมาณ 82.82

จากการวิเคราะห์สัดส่วนความแปรปรวนทั้งหมดของตัวแปรตามที่อธิบายได้หรือสัมประสิทธิ์การทำนาย (R^2) พบว่าโมเดล 4 มีสัมประสิทธิ์การทำนายเท่ากับ 0.8863 (88.63%) และความแปรปรวนทั้งหมดของตัวแปรตามที่อธิบายได้ในโมเดล 4 เท่ากับ 0.262836 (26.28%) ส่วนโมเดล 2 ที่มีสัมประสิทธิ์การทำนายเท่ากับ 0.7775 (77.75%) และความแปรปรวนทั้งหมดของตัวแปรตามที่อธิบายได้ในโมเดล 2 เท่ากับ 0.1870537 (18.71%) นั่นคือโมเดล 4 สามารถอธิบายความแปรปรวนของตัวแปรตามหรือคะแนนผลการประเมินวิชาวิทยาศาสตร์ได้ดีกว่าโมเดล 2

สรุปผลการวิเคราะห์ในโมเดล 1 ถึงโมเดล 4 เพื่อแสดงค่าสัมประสิทธิ์ของตัวแปรคุณลักษณะระดับนักเรียนและสถานศึกษาที่ได้จากการวิเคราะห์ และความแปรปรวนในแต่ละระดับและคำนวณหาค่าสัมประสิทธิ์การทำนายได้ ดังตารางที่ 32

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 32 สรุปค่าสัมประสิทธิ์ของตัวแปรตัวแปรคุณลักษณะระดับนักเรียนและสถานศึกษาที่ได้จากการวิเคราะห์โมเดล 1 ถึง 4

ตัวแปร	Coefficient			
	Model 1	Model 2	Model 3	Model 4
ค่าเฉลี่ยรวมคะแนนผลการประเมินวิชาวิทยาศาสตร์ทุกสถานศึกษา	480.157757**	475.385969**	480.283815**	475.933434**
ระดับนักเรียน				
ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา		15.377367**		14.586167**
ความเป็นเพศชายของนักเรียน		-1.072707		-1.071667
การใช้เวลาเรียนพิเศษวิทยาศาสตร์นอกสถานศึกษา		-1.833651		-1.723489
การใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง		8.397479**		8.003062**
เศรษฐานะของครอบครัว		5.453846**		5.166525**
แหล่งทรัพยากรที่บ้าน		2.853519**		2.700539**
ความมั่งคั่งของครอบครัว		-4.145535**		-3.944037**
ความสนใจในการเรียนรู้วิทยาศาสตร์		3.835460**		3.665655**
ความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์		0.422529		0.405494
ความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์		-8.038848**		-7.685855**
ความเพลิดเพลินทางวิทยาศาสตร์		8.992101**		8.550603**
แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์		4.686307**		4.482620**
การเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์		5.975125**		5.697482**
ระดับสถานศึกษา				
สถานศึกษาขนาดใหญ่		24.891340**		23.849150**
สถานศึกษาขนาดกลาง		19.455240**		18.589401**
สถานศึกษาสังกัด สพฐ.		-0.958839		-0.698231
สถานศึกษาสังกัด สช.		-24.336141*		-23.123298*
สถานศึกษาสังกัด สกอ.		54.316148**		52.409790**
สถานศึกษาสังกัด กศท. และ กทม.		-21.993594*		-20.818182*
สถานศึกษาที่ตั้งอยู่ในหมู่บ้าน		-29.778053**		-28.640284**
สถานศึกษาที่ตั้งอยู่ในเมืองเล็ก		-34.307281**		-32.883962**
สถานศึกษาที่ตั้งอยู่ในเมือง		-16.764427*		-16.183366*
สถานศึกษาที่ตั้งอยู่ในเมืองใหญ่		-14.375449		-13.872541
ขนาดห้องเรียน		0.348630		0.298023
แรงกดดันของผู้ปกครองต่อสถานศึกษา		-7.508397*		-7.235974*
การขาดแคลนครูวิทยาศาสตร์		3.393433		3.221607
การขาดแคลนครูที่มีคุณวุฒิ		0.166243		0.096468

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 32 สรุปค่าสัมประสิทธิ์ของตัวแปรตัวแปรคุณลักษณะระดับนักเรียนและสถานศึกษาที่ได้จากการวิเคราะห์โมเดล 1 ถึง 4 (ต่อ)

ตัวแปร	Coefficient			
	Model 1	Model 2	Model 3	Model 4
งบประมาณของรัฐที่จัดสรรให้สถานศึกษา		-0.363404**		-0.351279**
ดัชนีการขาดแคลนครู		-7.621401		-7.263655
ดัชนีทรัพยากรการเรียน		5.648004**		5.349077**
สัดส่วนนักเรียนต่อครู		-0.003694**		-0.003575**
ดัชนีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์		2.066948		1.966320
ดัชนีความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากร		-1.156776		-1.123837
ดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล		5.517180*		5.226349*
งบประมาณจากหน่วยงานภายนอก		0.107041		0.101552
Varaince Component	**Model 1**	**Model 2**	**Model 3**	**Model 4**
ระดับที่ 1 (R)	3304.54876	3026.18439	2996.81434	2744.59225
ระดับที่ 2 (U0)	2010.42084	616.57585	1851.82608	569.18444
ระดับความผันแปรของคะแนนผลการประเมินวิชาวิทยาศาสตร์	**Model 1**	**Model 2-2**	**Model 3**	**Model 4-2**
ระดับภายในสถานศึกษา	62.17	83.07	61.81	82.82
ระดับระหว่างสถานศึกษา	37.83	16.93	38.19	17.18
สัดส่วนความแปรปรวนทั้งหมดของตัวแปรตามที่อธิบายได้ (R^2)				
ระดับที่ 1	-	0.0842 (8.42%)	-	0.1694 (16.94%)
ระดับที่ 2	-	0.6933 (69.33%)	-	0.7168 (71.68%)
รวมทั้งสองระดับ	-	0.7775 (77.75%)	-	0.8863 (88.63%)
ความแปรปรวนของตัวแปรตามที่อธิบายได้				
ระดับที่ 1	-	0.069886 (6.99%)	-	0.1402632 (14.03%)
ระดับที่ 2	-	0.1171677 (11.72%)	-	0.1225728 (12.26%)
รวมทั้งสองระดับ	-	0.1870537 (18.71%)	-	0.262836 (26.28%)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 4 สรุปผลการวิเคราะห์เพื่อตอบคำถามการวิจัย

การนำเสนอผลในตอนนี้เป็นผลที่ได้จากการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานโดยใช้โมเดลการประเมินคุณภาพการจัดการศึกษาที่ต่างกัน 4 โมเดล โดยจะนำเสนอผลให้เห็นคะแนนผลการประเมินในแต่ละโมเดล แล้วจึงจะนำเสนอผลเพื่อตอบคำถามและวัตถุประสงค์การวิจัย ดังนั้นจึงแบ่งออกเป็น 2 ตอน คือ 4.1) ผลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานจากโมเดลการประเมินคุณภาพการจัดการศึกษา 4 โมเดล 4.2) สรุปผลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานเพื่อตอบคำถามการวิจัย

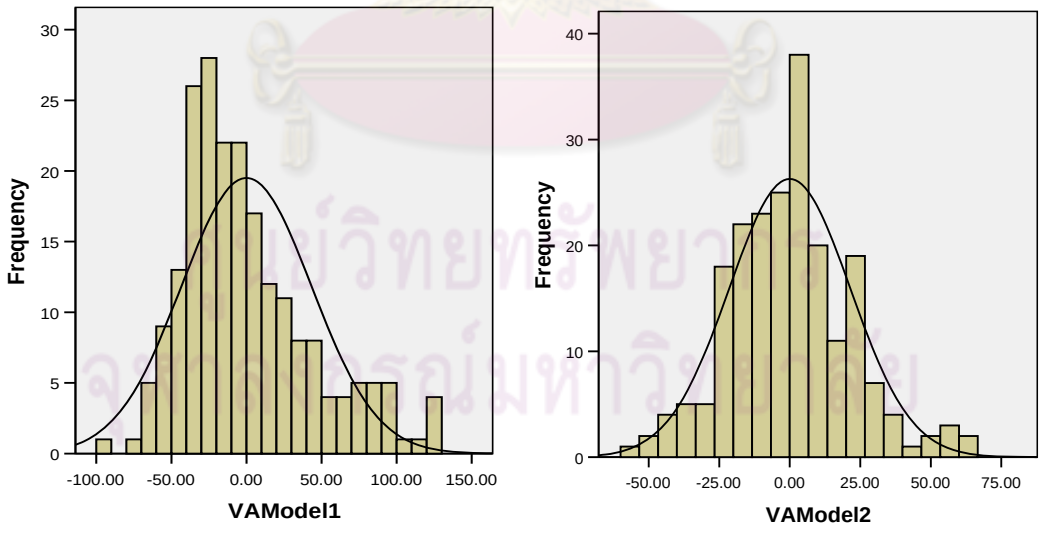
4.1 ผลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานจากโมเดลการประเมินคุณภาพการจัดการศึกษา 4 โมเดล

คุณภาพการจัดการศึกษาของการวิจัยในครั้งนี้ หมายถึงคะแนนมูลค่าเพิ่มของแต่ละสถานศึกษาที่ได้จากการวิเคราะห์ในโมเดล 1 ถึงโมเดล 4 ซึ่งได้จากการวิเคราะห์พหุระดับ คะแนนมูลค่าเพิ่มของแต่ละสถานศึกษาที่ได้จากการวิเคราะห์ 4 โมเดล ได้ผลที่ต่างกันดังนี้

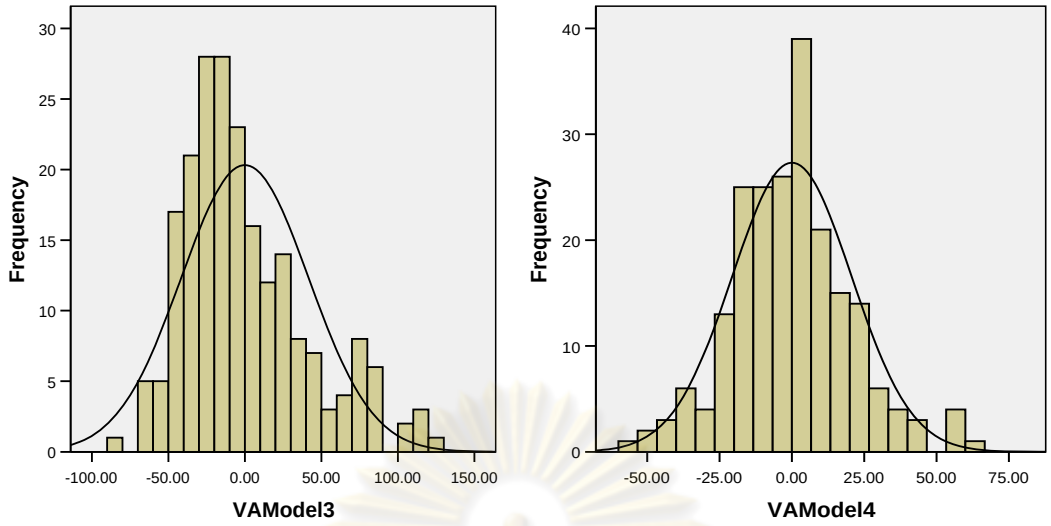
โมเดลการประเมินคุณภาพการจัดการศึกษา ได้ผลการประเมินพบว่า มีค่าเฉลี่ยของคะแนนมูลค่าเพิ่มเป็นศูนย์ ทั้งนี้เนื่องจากเป็นค่าเฉลี่ยของความคลาดเคลื่อน และเมื่อพิจารณาค่าสถิติของคะแนนมูลค่าเพิ่มของทั้ง 4 โมเดล พบว่า โมเดล 1 มีค่าใกล้เคียงกับโมเดล 3 และโมเดล 2 มีค่าใกล้เคียงกับโมเดล 4 คะแนนที่ตำแหน่งกึ่งกลางหรือคะแนนมัธยฐานของโมเดล 1 และ 3 มีค่า Med. = -9.1260 และ Med. = -9.0410 ตามลำดับ และโมเดล 2 และ 4 มีค่า Med. = .3105 และ Med. = .3240 ตามลำดับ การกระจายของคะแนนมูลค่าเพิ่มพบว่า โมเดล 1 และ 3 มีค่า S.D. = 43.35809 และ S.D. = 41.63301 ตามลำดับ และโมเดล 2 และ 4 มีค่า S.D. = 21.45819 และ S.D. = 20.64560 ตามลำดับ ลักษณะการแจกแจงของคะแนนในโมเดล 1 และ 3 มีลักษณะแบบเบ้ขวามากกว่าโมเดล 2 และ 4 คะแนนมีการแจกแจงลักษณะแบบแบนทั้ง 4 โมเดล ตำแหน่งเปอร์เซนต์ไทล์ของโมเดล 1 และ 3 มีค่าใกล้เคียงกัน (Pr_{25} = -30.7285, Pr_{75} = 25.0683 และ Pr_{25} = -29.2393, Pr_{75} = 23.9698 ตามลำดับ) และโมเดล 2 และ 4 มีค่าใกล้เคียงกัน (Pr_{25} = -14.5645, Pr_{75} = 12.0058 และ Pr_{25} = -14.1485, Pr_{75} = 11.6093 ตามลำดับ) แสดงรายละเอียดดังตารางที่ 49 และการทดสอบการแจกแจงแบบปกติด้วยสถิติทดสอบของ Kolmogorov-Smirnov พบว่า คะแนนมูลค่าเพิ่มของโมเดล 1 และ 3 ไม่มีการแจกแจงแบบปกติ ส่วนโมเดล 2 และ 4 มีการแจกแจงแบบปกติ แสดงดังภาพที่ 20

ตารางที่ 33 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มวิชาวิทยาศาสตร์ของสถานศึกษาจากผลการวิเคราะห์ 4 โมเดล

ค่าสถิติ	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
n	212	212	212	212
Mean	.0000	.0000	.0000	.0000
(Std. Error of Mean)	(2.97785)	(1.47375)	(2.85937)	(1.41795)
Median	-9.1260	.3105	-9.0410	.3240
Mode	-92.81	-58.57	-88.87	-56.16
Std. Deviation	43.35809	21.45819	41.63301	20.64560
Variance	1879.924	460.454	1733.307	426.241
Skewness	.899	.265	.905	.264
(Std. Error of Skewness)	(.167)	(.167)	(.167)	(.167)
Kurtosis	.448	.623	.457	.632
(Std. Error of Kurtosis)	(.333)	(.333)	(.333)	(.333)
Range	221.61	122.03	213.04	117.40
Minimum	-92.81	-58.57	-88.87	-56.16
Maximum	128.79	63.45	124.17	61.24
Percentiles 25	-30.7285	-14.5645	-29.2393	-14.1485
Percentiles 75	25.0683	12.0058	23.9698	11.6093
Kolmogorov-Smirnov Z	1.721	.890	1.707	.910
(Sig.)	(.005)	(.407)	(.006)	(.379)



ภาพที่ 20 ฮีสโตแกรม (Histogram) ของคะแนนมูลค่าเพิ่มจากโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานจำนวน 212 แห่ง



ภาพที่ 20 ฮีสโตแกรม (Histogram) ของคะแนนมูลค่าเพิ่มจากโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานจำนวน 212 แห่ง (ต่อ)

การจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาได้จัดเป็น 3 กลุ่ม คือ กลุ่มที่ 1 เป็นกลุ่มต่ำ มีตำแหน่งเปอร์เซ็นต์ไทล์ 1 – 24.99 มีสถานศึกษาจำนวน 53 แห่ง กลุ่มที่ 2 เป็นกลุ่มปานกลาง มีตำแหน่งเปอร์เซ็นต์ไทล์ 25 – 74.99 มีสถานศึกษาจำนวน 106 แห่ง และกลุ่มที่ 3 เป็นกลุ่มสูง มีตำแหน่งเปอร์เซ็นต์ไทล์ 75 – 100 มีสถานศึกษาจำนวน 53 แห่ง และจำแนกตามคุณลักษณะของสถานศึกษา การจัดกลุ่มคุณภาพการจัดการศึกษาในแต่ละคุณลักษณะของสถานศึกษาจากการวิเคราะห์ 4 โมเดล พบว่า

1) **ขนาดสถานศึกษา** เมื่อพิจารณาในแต่ละขนาดของสถานศึกษา การจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษา**ขนาดเล็ก** พบว่าผลการวิเคราะห์จากโมเดล 1 และ 3 จัดสถานศึกษาขนาดเล็กส่วนใหญ่ให้อยู่ในกลุ่มต่ำ (58.33% และ 55.55%) และไม่มีสถานศึกษาขนาดเล็กที่จัดอยู่ในกลุ่มสูง ส่วนผลการวิเคราะห์จากโมเดล 2 และ 4 จัดสถานศึกษาขนาดเล็กให้อยู่ในกลุ่มกลาง (58.33%) การจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษา**ขนาดกลาง** พบว่าทั้ง 4 โมเดลจัดสถานศึกษาขนาดกลางให้อยู่ในกลุ่มกลางเป็นส่วนใหญ่ การจัดกลุ่มคุณภาพการจัดการศึกษาของ**สถานศึกษาขนาดใหญ่** พบว่าทั้ง 4 โมเดลจัดสถานศึกษาขนาดใหญ่ให้อยู่ในกลุ่มกลางเป็นส่วนใหญ่

2) **ประเภทสถานศึกษา** เมื่อพิจารณาในแต่ละประเภทสถานศึกษา การจัดกลุ่มคุณภาพการจัดการศึกษาของ**สถานศึกษาของรัฐ** พบว่าทั้ง 4 โมเดล จัดสถานศึกษาของรัฐให้อยู่ในกลุ่มกลางเป็นส่วนใหญ่ การจัดกลุ่มคุณภาพการจัดการศึกษาของ**สถานศึกษาของเอกชน** พบว่าผลการวิเคราะห์จากโมเดล 1 และ 3 จัดสถานศึกษาของเอกชนส่วนใหญ่ให้อยู่ในกลุ่มต่ำ (40.00%) ส่วนผลการวิเคราะห์จากโมเดล 2 และ 4 จัดสถานศึกษาของเอกชนให้อยู่ในกลุ่มกลาง (40.00%)

3) **สังกัดของสถานศึกษา** เมื่อพิจารณาในแต่ละสังกัดสถานศึกษา การจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษา**สังกัด สพฐ.** พบว่าทั้ง 4 โมเดล จัดสถานศึกษาของรัฐให้อยู่ใน

กลุ่มกลางเป็นส่วนใหญ่ การจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษา**สังกัด สช.** พบว่าผลการวิเคราะห์จากโมเดล 1 และ 3 จัดสถานศึกษาของเอกชนส่วนใหญ่ให้อยู่ในกลุ่มต่ำ (40.00%) ส่วนผลการวิเคราะห์จากโมเดล 2 และ 4 จัดสถานศึกษาของเอกชนให้อยู่ในกลุ่มกลาง (40.00%) การจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษา**สังกัด สกอ. หรือโรงเรียนสาธิต** พบว่าทั้ง 4 โมเดล จัดสถานศึกษาของรัฐให้อยู่ในกลุ่มสูงเป็นส่วนใหญ่ และเป็นที่น่าสังเกตว่าผลการวิเคราะห์จากโมเดล 1 และ 3 จัดโรงเรียนสาธิตให้อยู่ในกลุ่มสูงทั้งหมด (100.00%) การจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษา**สังกัด กศท./กทม. และ อศ.** พบว่าทั้ง 4 โมเดล จัดสถานศึกษาของรัฐให้อยู่ในกลุ่มกลางเป็นส่วนใหญ่ และเป็นที่น่าสังเกตว่าผลการวิเคราะห์ทั้ง 4 โมเดล พบว่าสถานศึกษาในสังกัด อศ. ไม่ได้จัดอยู่ในกลุ่มสูง

4) **ตำแหน่งที่ตั้งสถานศึกษา** พบว่าผลการวิเคราะห์ทั้ง 4 โมเดล จัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาที่ตั้งอยู่ใน**หมู่บ้าน เมืองเล็ก เมือง และเมืองใหญ่** อยู่ในกลุ่มกลางเป็นส่วนใหญ่ นอกจากนี้พบว่าผลการวิเคราะห์ในโมเดล 2 และ 4 จัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาที่ตั้งอยู่ในหมู่บ้าน เมืองเล็ก ให้อยู่ในกลุ่มสูงมากกว่าจากผลการวิเคราะห์ในโมเดล 1 และ 3 และผลการวิเคราะห์ในโมเดล 2 และ 4 จัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาที่ตั้งอยู่ในเมือง เมืองใหญ่ ให้อยู่ในกลุ่มสูงน้อยกว่าจากผลการวิเคราะห์ในโมเดล 1 และ 3 การจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาที่ตั้งอยู่ใน**เมืองใหญ่มาก** พบว่าทั้ง 4 โมเดล จัดอยู่ในกลุ่มสูงเป็นส่วนใหญ่ โดยผลการวิเคราะห์จากโมเดล 1 และ 3 จัดสถานศึกษาที่ตั้งอยู่ในเมืองใหญ่ให้อยู่ในกลุ่มสูง คิดเป็น 70.00%

5) **แรงกดดันของผู้ปกครองต่อสถานศึกษา** เมื่อพิจารณาในแต่ละระดับแรงกดดันของผู้ปกครองต่อสถานศึกษา การจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาที่มี**แรงกดดันของผู้ปกครองจำนวนมาก** พบว่าผลการวิเคราะห์จากโมเดล 1 และ 3 ส่วนใหญ่จัดให้อยู่ในกลุ่มสูง (45.76%) ส่วนผลการวิเคราะห์จากโมเดล 2 และ 4 ส่วนใหญ่จัดให้อยู่ในกลุ่มกลาง (44.07%) การจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาที่มี**แรงกดดันของผู้ปกครองจำนวนน้อยและไม่พบแรงกดดัน** พบว่าทั้ง 4 โมเดล จัดอยู่ในกลุ่มกลางเป็นส่วนใหญ่

6) **การขาดแคลนครูวิทยาศาสตร์** เมื่อพิจารณาในแต่ละระดับการขาดแคลนครูวิทยาศาสตร์ พบว่าผลการวิเคราะห์ทั้ง 4 โมเดล ส่วนใหญ่จัดสถานศึกษาให้อยู่ในกลุ่มกลาง ยกเว้นสถานศึกษาที่มีระดับ**การขาดแคลนครูวิทยาศาสตร์มาก** พบว่าผลการวิเคราะห์จากโมเดล 1 และ 3 ส่วนใหญ่จัดสถานศึกษาให้อยู่ในกลุ่มต่ำ (50.00%) แต่ผลการวิเคราะห์ในโมเดล 2 และ 4 ส่วนใหญ่จัดสถานศึกษาให้อยู่ในกลุ่มกลาง (64.29%)

7) **การขาดแคลนครูที่มีคุณวุฒิ** เมื่อพิจารณาในแต่ละระดับการขาดแคลนครูที่มีคุณวุฒิการจัดกลุ่มคุณภาพการจัดการศึกษาของ**สถานศึกษาที่ไม่ขาดแคลนครูวิทยาศาสตร์** พบว่าผลการวิเคราะห์จากโมเดล 1 และ 3 ส่วนใหญ่จัดให้อยู่ในกลุ่มสูง (51.35%) ส่วนผลการวิเคราะห์จากโมเดล 2 และ 4 ส่วนใหญ่จัดสถานศึกษาให้อยู่ในกลุ่มกลาง (40.54%) ส่วนการจัดกลุ่มคุณภาพ

การจัดการศึกษาของ**สถานศึกษาที่มีการขาดแคลนครูวิทยาศาสตร์** พบว่าทั้ง 4 โมเดล ส่วนใหญ่จัดสถานศึกษาอยู่ในกลุ่มกลาง

แสดงรายละเอียดดังตารางที่ 34

ตารางที่ 34 ผลการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาจากผลการวิเคราะห์ 4 โมเดล จำแนกตามคุณลักษณะของสถานศึกษา

คุณลักษณะสถานศึกษา	รวม (212)	กลุ่มต่ำ(N = 53) (Percentile ที่ 0-24.99)				กลุ่มกลาง(N = 106) (Percentile ที่ 25-74.99)				กลุ่มสูง(N = 53) (Percentile ที่ 75 ขึ้นไป)			
		Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
1. ขนาดสถานศึกษา													
1.1 เล็ก	36	21	7	20	7	15	21	16	21	-	8	-	8
	(100.0)	(58.33)	(19.44)	(55.55)	(19.44)	(41.66)	(58.33)	(44.44)	(58.33)	(0.0)	(22.22)	(0.0)	(22.22)
1.2 กลาง	84	22	21	23	21	48	46	47	46	14	17	14	17
	(100.0)	(26.19)	(25.00)	(27.38)	(25.00)	(57.14)	(54.76)	(55.95)	(54.76)	(16.66)	(20.24)	(16.66)	(20.24)
1.3 ใหญ่	92	10	25	10	25	43	39	43	39	39	28	39	28
	(100.0)	(10.87)	(27.17)	(10.87)	(27.17)	(46.74)	(42.39)	(46.74)	(42.39)	(42.39)	(30.43)	(42.39)	(30.43)
2. ประเภทสถานศึกษา													
2.1 รัฐ	187	43	45	43	45	97	96	97	96	47	46	47	46
	(100.0)	(22.99)	(24.06)	(22.99)	(24.06)	(51.87)	(51.34)	(51.87)	(51.34)	(25.13)	(24.59)	(25.13)	(24.59)
2.2 เอกชน	25	10	8	10	8	9	10	9	10	6	7	6	7
	(100.0)	(40.00)	(32.00)	(40.00)	(32.00)	(36.00)	(40.00)	(36.00)	(40.00)	(24.00)	(28.00)	(24.00)	(28.00)
3. สังกัดสถานศึกษา													
3.1 สพฐ.	134	33	30	33	30	72	71	72	71	29	33	29	33
	(100.0)	(24.63)	(22.39)	(24.63)	(22.39)	(53.73)	(52.99)	(53.73)	(52.99)	(21.64)	(24.63)	(21.64)	(24.63)
3.2 สช.	25	10	8	10	8	9	10	9	10	6	7	6	7
	(100.0)	(40.00)	(32.00)	(40.00)	(32.00)	(36.00)	(40.00)	(36.00)	(40.00)	(24.00)	(28.00)	(24.00)	(28.00)
3.3 รร.สาธิต	15	-	4	-	4	-	5	-	5	15	6	15	6
	(100.0)	(0.0)	(26.67)	(0.0)	(26.67)	(0.0)	(33.33)	(0.0)	(33.33)	(100.0)	(40.00)	(100.0)	(40.00)
3.4 กศท/ กทม.	30	10	10	10	10	17	13	17	13	3	7	3	7
	(100.0)	(33.33)	(33.33)	(33.33)	(33.33)	(56.67)	(43.33)	(56.67)	(43.33)	(10.00)	(23.33)	(10.00)	(23.33)
3.5 อศ.	8	-	1	-	1	8	7	8	7	-	-	-	-
	(100.0)	(0.0)	(12.50)	(0.0)	(12.50)	(100.0)	(87.50)	(100.0)	(87.50)	(0.0)	(0.0)	(0.0)	(0.0)

หมายเหตุ ตัวเลขในวงเล็บแสดงจำนวนร้อยละ

ตารางที่ 34 ผลการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาจากผลการวิเคราะห์ 4 โมเดล จำแนกตามคุณลักษณะของสถานศึกษา (ต่อ)

คุณลักษณะสถานศึกษา	รวม	กลุ่มต่ำ (Percentile ที่ 0-24.99)(N = 53)				กลุ่มกลาง (Percentile ที่ 25-74.99) (N = 106)				กลุ่มสูง (Percentile ที่ 75 ขึ้นไป) (N = 53)			
		Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
4. ตำแหน่งที่ตั้งสถานศึกษา													
4.1 หมู่บ้าน	51	24	11	23	11	26	29	27	29	1	11	1	11
	(100.0)	(47.06)	(21.57)	(45.09)	(21.57)	(50.98)	(56.86)	(52.94)	(56.86)	(1.96)	(21.57)	(1.96)	(21.57)
4.2 เมืองเล็ก	44	13	9	14	9	26	26	25	26	5	9	5	9
	(100.0)	(29.54)	(20.45)	(31.82)	(20.45)	(59.09)	(59.09)	(56.81)	(59.09)	(11.36)	(20.45)	(11.36)	(20.45)
4.3 เมือง	54	8	15	8	15	31	27	31	27	15	12	15	12
	(100.0)	(14.81)	(27.77)	(14.81)	(27.77)	(57.40)	(50.00)	(57.40)	(50.00)	(27.77)	(22.22)	(27.77)	(22.22)
4.4 เมืองใหญ่	43	6	11	6	11	19	20	19	20	18	12	18	12
	(100.0)	(13.95)	(25.58)	(13.95)	(25.58)	(44.19)	(46.51)	(44.19)	(46.51)	(41.86)	(27.91)	(41.86)	(27.91)
4.5 ใหญ่มาก	20	2	7	2	7	4	4	4	4	14	9	14	9
	(100.0)	(10.00)	(35.00)	(10.00)	(35.00)	(20.00)	(20.00)	(20.00)	(20.00)	(70.00)	(45.00)	(70.00)	(45.00)
5. แรงกดดันของผู้ปกครองต่อสถานศึกษา													
5.1 มาก	59	7	16	6	16	25	26	26	26	27	17	27	17
	(100.0)	(11.86)	(27.11)	(10.17)	(27.11)	(42.37)	(44.07)	(44.07)	(44.07)	(45.76)	(28.81)	(45.76)	(28.81)
5.2 น้อย	94	27	23	28	23	48	52	47	52	19	19	19	19
	(100.0)	(28.72)	(24.47)	(29.78)	(24.47)	(51.06)	(55.32)	(50.00)	(55.32)	(20.21)	(20.21)	(20.21)	(20.21)
5.3 ไม่พบ	59	19	14	19	14	33	28	33	28	7	17	7	17
	(100.0)	(32.20)	(23.73)	(32.20)	(23.73)	(55.93)	(47.45)	(55.93)	(47.45)	(11.86)	(28.81)	(11.86)	(28.81)
6. การขาดแคลนครูวิทยาศาสตร์													
6.1 ไม่ขาดแคลน	72	13	23	14	23	33	31	32	31	26	18	26	18
	(100.0)	(18.05)	(31.94)	(19.44)	(31.94)	(45.83)	(43.05)	(44.44)	(43.05)	(36.11)	(25.00)	(36.11)	(25.00)
6.2 น้อย	95	23	22	23	22	51	47	51	47	21	26	21	26
	(100.0)	(24.21)	(23.16)	(24.21)	(23.16)	(53.68)	(49.47)	(53.68)	(49.47)	(22.10)	(27.37)	(22.10)	(27.37)
6.3 บางส่วน	31	10	5	9	5	16	19	17	19	5	7	5	7
	(100.0)	(32.26)	(16.13)	(29.03)	(16.13)	(51.61)	(61.29)	(54.84)	(61.29)	(16.13)	(22.58)	(16.13)	(22.58)
6.4 มาก	14	7	3	7	3	6	9	6	9	1	2	1	2
	(100.0)	(50.00)	(21.43)	(50.00)	(21.43)	(42.86)	(64.29)	(42.86)	(64.29)	(7.14)	(14.28)	(7.14)	(14.28)

หมายเหตุ ตัวเลขในวงเล็บแสดงจำนวนร้อยละ

ตารางที่ 34 ผลการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาจากผลการวิเคราะห์ 4 โมเดล จำแนกตามคุณลักษณะของสถานศึกษา (ต่อ)

คุณลักษณะสถานศึกษา	รวม	กลุ่มต่ำ (Percentile ที่ 0-24.99)(N = 53)				กลุ่มกลาง (Percentile ที่ 25-74.99) (N = 106)				กลุ่มสูง (Percentile ที่ 75 ขึ้นไป) (N = 53)			
		Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
7. การขาดแคลนครูที่มีคุณวุฒิ													
7.1 ไม่ขาดแคลน	37	10	11	9	11	8	15	9	15	19	11	19	11
	(100.0)	(27.03)	(29.73)	(24.32)	(29.73)	(21.62)	(40.54)	(24.32)	(40.54)	(51.35)	(29.73)	(51.35)	(29.73)
7.2 น้อย	95	21	25	20	25	50	46	51	46	24	24	24	24
	(100.0)	(22.11)	(26.31)	(21.05)	(26.31)	(52.63)	(48.42)	(53.68)	(48.42)	(25.26)	(25.26)	(25.26)	(25.26)
7.3 บางส่วน	58	18	15	20	15	34	32	32	32	6	11	6	11
	(100.0)	(31.03)	(25.86)	(34.48)	(25.86)	(58.62)	(55.17)	(55.17)	(55.17)	(10.34)	(18.96)	(10.34)	(18.96)
7.4 มาก	22	4	2	4	2	14	13	14	13	4	7	4	7
	(100.0)	(18.18)	(9.09)	(18.18)	(9.09)	(63.63)	(59.09)	(63.63)	(59.09)	(18.18)	(31.82)	(18.18)	(31.82)
รวม	212	53	53	53	53	106	106	106	106	53	53	53	53
	(100.0)	(25.00)	(25.00)	(25.00)	(25.00)	(50.00)	(50.00)	(50.00)	(50.00)	(25.00)	(25.00)	(25.00)	(25.00)

หมายเหตุ ตัวเลขในวงเล็บแสดงจำนวนร้อยละ

4.2 สรุปผลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานเพื่อตอบคำถามการวิจัย

1) ผลของข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบวิชาวิทยาศาสตร์ต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน

การเปรียบเทียบผลการประเมินคุณภาพการจัดการศึกษาระหว่างโมเดล 1 กับโมเดล 3 และระหว่างโมเดล 2 และโมเดล 4 เป็นการตอบคำถามการวิจัยและวัตถุประสงค์การวิจัยข้อที่ 1 เกี่ยวกับผลของข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบวิชาวิทยาศาสตร์ต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน ได้ผลดังนี้

โมเดล 1 และโมเดล 3 เป็นโมเดลที่ไม่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา โดยโมเดล 1 ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป แต่โมเดล 3 เป็นโมเดลที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป และจัดกลุ่มคุณภาพการจัดการศึกษาเป็น 3 กลุ่ม โดยใช้คะแนนมูลค่าเพิ่มที่ได้จากการวิเคราะห์ในแต่ละโมเดล **ผลการจัดกลุ่มคุณภาพ (Rating) การจัดการศึกษา** พบว่า โมเดล 1 และ 3 จัดกลุ่มคุณภาพการจัดการศึกษาได้สอดคล้องกันหรือจัดสถานศึกษาอยู่ในกลุ่มเดียวกันได้จำนวน 208 แห่ง คิดเป็นร้อยละ 98.2 โดยจัดเป็นกลุ่ม

ต่ำจำนวน 51 แห่ง (24.1%) กลุ่มกลาง 104 แห่ง (49.1%) และกลุ่มสูง 53 แห่ง (25.0%) และพบว่าจัดกลุ่มไม่สอดคล้องกันจำนวน 4 แห่ง คิดเป็นร้อยละ 1.8 โดยการจัดกลุ่มจากผลการวิเคราะห์ในโมเดล 1 จัดเป็นกลุ่มต่ำ แต่ผลการวิเคราะห์ในโมเดล 3 จัดเป็นกลุ่มกลาง จำนวน 2 แห่ง (.9%) และผลการวิเคราะห์ในโมเดล 1 จัดเป็นกลุ่มกลาง แต่ผลการวิเคราะห์ในโมเดล 3 จัดเป็นกลุ่มต่ำ จำนวน 2 แห่ง (.9%) **ผลการจัดระดับคุณภาพ (Ranking) การจัดการศึกษา** พบว่า สถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 ตรงกันกับระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 มีจำนวน 133 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 62.74% สถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 สูงกว่าระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 มี 43 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 20.28% และสถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 ต่ำกว่าระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 มี 36 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 16.98% เมื่อทดสอบความสอดคล้องของการจัดกลุ่มคุณภาพการจัดการศึกษาและระดับคุณภาพการจัดการศึกษาจากการประเมินด้วยโมเดล 1 และ 3 ด้วยสถิติทดสอบวิลคอกซอน (Wilcoxon Signed Ranks Test) พบว่าผลการประเมินด้วยโมเดล 1 และ 3 จัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาไม่แตกต่างกัน ($Z = -1.414$, $p = .157$) นั่นคือผลการประเมินคุณภาพการจัดการศึกษาด้วยโมเดล 1 และ 3 มีผลการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาสอดคล้องกัน และเมื่อทดสอบความสอดคล้องของการจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาพบว่าไม่แตกต่างกัน ($Z = -.121$, $p = .904$) นั่นคือผลการประเมินคุณภาพการจัดการศึกษาด้วยโมเดล 1 และ 3 มีผลการจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาสอดคล้องกัน แสดงรายละเอียดดังตารางที่ 35

โมเดล 2 และโมเดล 4 เป็นโมเดลที่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา โดยโมเดล 2 ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป แต่โมเดล 4 เป็นโมเดลที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป เมื่อจัดกลุ่มคุณภาพการจัดการศึกษาเป็น 3 กลุ่ม **ผลการจัดกลุ่มคุณภาพการจัดการศึกษา** พบว่า โมเดล 2 และ 4 จัดกลุ่มคุณภาพการจัดการศึกษาได้สอดคล้องกันหรือจัดสถานศึกษาอยู่ในกลุ่มเดียวกันได้ทั้งหมดจำนวน 212 แห่ง คิดเป็นร้อยละ 100.0 โดยจัดเป็นกลุ่มต่ำจำนวน 53 แห่ง (25.0%) กลุ่มกลาง 106 แห่ง (50.0%) และกลุ่มสูง 53 แห่ง (25.0%) **ผลการจัดระดับคุณภาพการจัดการศึกษา** พบว่า สถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 2 ตรงกันกับระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มีจำนวน 110 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 51.89% สถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 2 สูงกว่าระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มี 47 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 22.17% และสถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 2 ต่ำกว่ากับระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มี 55 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 25.94% เมื่อทดสอบความสอดคล้องของการจัดกลุ่มคุณภาพการจัดการศึกษาและระดับคุณภาพการจัดการศึกษาของสถานศึกษาจากการประเมินด้วยโมเดล 2 และ 4 ด้วยสถิติ

ทดสอบวิลคอกซอน (Wilcoxon Signed Ranks Test) พบว่าผลการประเมินด้วยโมเดล 2 และ 4 จัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาไม่แตกต่างกัน (Z = .000 , p = 1.000) นั่นคือผลการประเมินคุณภาพการจัดการศึกษาด้วยโมเดล 2 และ 4 มีผลการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาสอดคล้องกัน และเมื่อทดสอบความสอดคล้องของจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาพบว่าการจัดระดับคุณภาพการจัดการศึกษาของโมเดล 2 และ 4 แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 (Z = -2.205, p = .027) นั่นคือผลการประเมินคุณภาพการจัดการศึกษาด้วยโมเดล 2 และ 4 มีผลการจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาไม่สอดคล้องกัน แสดงรายละเอียดดังตารางที่ 36

จากผลการเปรียบเทียบการจัดกลุ่มคุณภาพการจัดการศึกษาและระดับคุณภาพการจัดการศึกษาของสถานศึกษาระหว่างโมเดล 1 และ 3 แสดงให้เห็นว่าการประเมินด้วยโมเดลที่ไม่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป และโมเดลที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป จะทำให้ผลการจัดกลุ่มคุณภาพและระดับคุณภาพการจัดการศึกษาของสถานศึกษาไม่ต่างกัน นั่นคือข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบวิชาวิทยาศาสตร์ไม่มีผลต่อการประเมินคุณภาพการจัดการศึกษา และจากผลการวิเคราะห์ของโมเดล 2 และ 4 พบว่าการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาไม่ต่างกัน แต่ระดับคุณภาพการจัดการศึกษาแตกต่างกัน แสดงว่าแบบสอบที่มีข้อสอบที่ทำหน้าที่ต่างกันจะไม่มีผลการประเมินคุณภาพการจัดการศึกษา แต่ถ้ามีการควบคุมอิทธิพลของคุณลักษณะนักเรียนและสถานศึกษาจะส่งผลให้การประเมินคุณภาพการจัดการศึกษาของสถานศึกษามีระดับคุณภาพการจัดการศึกษาที่แตกต่างกัน

ตารางที่ 35 ผลการจัดกลุ่มคุณภาพการจัดการศึกษา การจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาและค่าสถิติทดสอบ Wilcoxon Signed Ranks Test จากผลการวิเคราะห์โมเดล 1 และ โมเดล 3

	ผลการจัดกลุ่มคุณภาพ (Rating) การจัดการศึกษา				
โมเดล 1	**กลุ่มคุณภาพการจัดการศึกษา**	**โมเดล 3**			
		กลุ่มต่ำ	**กลุ่มกลาง**	**กลุ่มสูง**	**รวม**
	กลุ่มต่ำ	51 (24.1%)	2 (.9%)	0 (.0%)	53 (25.0%)
	กลุ่มกลาง	2 (.9%)	104 (49.1%)	0 (.0%)	106 (50.0%)
	กลุ่มสูง	0 (.0%)	0 (.0%)	53 (25.0%)	53 (25.0%)
	รวม	53 (25.0%)	106 (50.0%)	53 (25.0%)	212 (100.0%)
	Wilcoxon Signed Ranks Test ; Z = -1.414 , p = .157				
	ผลการจัดระดับคุณภาพ (Ranking) การจัดการศึกษา				
โมเดล 1	**ระดับสูงกว่า โมเดล 3**		**ระดับตรงกันกับโมเดล 3**		**ระดับต่ำกว่า โมเดล 3**
	43 (20.28%)		133 (62.74%)		36 (16.98%)
	Wilcoxon Signed Ranks Test ; Z = -.121 , p = .904				

ตารางที่ 36 ผลการจัดกลุ่มคุณภาพการจัดการศึกษา การจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษา และค่าสถิติทดสอบ Wilcoxon Signed Ranks Test จากการวิเคราะห์โมเดล 2 และ โมเดล 4

	ผลการจัดกลุ่มคุณภาพ (Rating) การจัดการศึกษา				
โมเดล 2	กลุ่มคุณภาพการจัดการศึกษา	โมเดล 4			
		กลุ่มต่ำ	กลุ่มกลาง	กลุ่มสูง	รวม
	กลุ่มต่ำ	53 (25.0%)	0 (.0%)	0 (.0%)	53 (25.0%)
	กลุ่มกลาง	0 (.0%)	106 (50.0%)	0 (.0%)	106 (50.0%)
	กลุ่มสูง	0 (.0%)	0 (.0%)	53 (25.0%)	53 (25.0%)
	รวม	53 (25.0%)	106 (50.0%)	53 (25.0%)	212 (100.0%)
	Wilcoxon Signed Ranks Test ; Z = .000 , p = 1.000				
	ผลการจัดระดับคุณภาพ (Ranking) การจัดการศึกษา				
โมเดล 2	ระดับสูงกว่า โมเดล 4	ระดับตรงกันกับโมเดล 4		ระดับต่ำกว่า โมเดล 4	
	47 (22.17%)	110 (51.89%)		55 (25.94%)	
	Wilcoxon Signed Ranks Test ; Z = -2.205, p = .027				

2) ผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาขั้นพื้นฐานต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน

การเปรียบเทียบผลการประเมินคุณภาพการจัดการศึกษาระหว่างโมเดล 1 กับโมเดล 2 และระหว่างโมเดล 3 และโมเดล 4 เป็นการตอบคำถามการวิจัยและวัตถุประสงค์การวิจัยข้อที่ 2 เกี่ยวกับผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาขั้นพื้นฐาน และคุณภาพของแบบสอบที่ต่างกันของวิชาวิทยาศาสตร์ต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน ได้ผลดังนี้

โมเดล 1 และโมเดล 2 เป็นโมเดลที่ไม่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปก่อนการวิเคราะห์มูลค่าเพิ่ม โดยโมเดล 1 ไม่มีการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา แต่โมเดล 2 เป็นโมเดลที่มีการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา เมื่อจัดกลุ่มคุณภาพการจัดการศึกษาออกเป็น 3 กลุ่ม **ผลการจัดกลุ่มคุณภาพ (Rating) การจัดการศึกษา** พบว่า โมเดล 1 และ 2 จัดกลุ่มคุณภาพการจัดการศึกษาได้สอดคล้องกันหรือจัดสถานศึกษาอยู่ในกลุ่มเดียวกันได้จำนวน 131 แห่ง คิดเป็นร้อยละ 61.9 โดยจัดเป็นกลุ่มต่ำจำนวน 30 แห่ง (14.2%) กลุ่มกลาง 68 แห่ง (32.1%) และกลุ่มสูง 33 แห่ง (15.6%) และพบว่าจัดกลุ่มได้ไม่สอดคล้องกันจำนวน 81 แห่ง คิดเป็นร้อยละ 38.1 โดยการจัดกลุ่มจากผลการวิเคราะห์ในโมเดล 1 จัดเป็นกลุ่มต่ำ ผลการวิเคราะห์ในโมเดล 2 จัดเป็นกลุ่มกลาง จำนวน 23 แห่ง (10.8%)

ผลการวิเคราะห์ในโมเดล 1 จัดเป็นกลุ่มกลาง ผลการวิเคราะห์ในโมเดล 2 จัดเป็นกลุ่มต่ำ จำนวน 18 แห่ง (8.5%) ผลการวิเคราะห์ในโมเดล 1 จัดเป็นกลุ่มกลาง ผลการวิเคราะห์ในโมเดล 2 จัดเป็นกลุ่มสูง จำนวน 20 แห่ง (9.4%) ผลการวิเคราะห์ในโมเดล 1 จัดเป็นกลุ่มสูง แต่ผลการวิเคราะห์ในโมเดล 2 จัดเป็นกลุ่มต่ำ จำนวน 5 แห่ง (2.4%) ผลการวิเคราะห์ในโมเดล 1 จัดเป็นกลุ่มสูง ผลการวิเคราะห์ในโมเดล 2 จัดเป็นกลุ่มกลาง จำนวน 15 แห่ง (7.1%) **ผลการจัดระดับคุณภาพ (Ranking) การจัดการศึกษา** พบว่า สถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 ตรงกันกับระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 2 มีจำนวน 2 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 0.94% สถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 สูงกว่าระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 2 มี 115 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 54.25% และสถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 ต่ำกว่าระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 2 มี 95 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 44.81% เมื่อทดสอบความสอดคล้องของการจัดกลุ่มคุณภาพการจัดการศึกษาและระดับคุณภาพการจัดการศึกษาจากการประเมินด้วยโมเดล 1 และ 2 ด้วยสถิติทดสอบวิลคอกซอน (Wilcoxon Signed Ranks Test) พบว่าผลการประเมินด้วยโมเดล 1 และ 2 จัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .01 ($Z = -6.403$, $p = .000$) และจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ($Z = -5.646$, $p = .000$) นั่นคือผลการประเมินด้วยโมเดล 1 และ 2 จัดกลุ่มคุณภาพและระดับคุณภาพการจัดการศึกษาของสถานศึกษาได้ไม่สอดคล้องกัน ดังรายละเอียดในตารางที่ 37

โมเดล 3 และโมเดล 4 เป็นโมเดลที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปก่อนการวิเคราะห์มูลค่าเพิ่ม โดยโมเดล 3 ไม่มีการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา แต่โมเดล 4 มีการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา เมื่อจัดกลุ่มคุณภาพการจัดการศึกษาเป็น 3 กลุ่ม **ผลการจัดกลุ่มคุณภาพการจัดการศึกษา** พบว่า โมเดล 3 และโมเดล 4 จัดกลุ่มคุณภาพการจัดการศึกษาได้สอดคล้องกันหรือจัดสถานศึกษาอยู่ในกลุ่มเดียวกันได้จำนวน 132 แห่ง คิดเป็นร้อยละ 62.3 โดยจัดเป็นกลุ่มต่ำจำนวน 30 แห่ง (14.2%) กลุ่มกลาง 69 แห่ง (32.5%) และกลุ่มสูง 33 แห่ง (15.6%) และพบว่าจัดกลุ่มได้ไม่สอดคล้องกันจำนวน 80 แห่ง คิดเป็นร้อยละ 37.7 โดยการจัดกลุ่มจากผลการวิเคราะห์ในโมเดล 3 จัดเป็นกลุ่มต่ำ ผลการวิเคราะห์ในโมเดล 4 จัดเป็นกลุ่มกลาง จำนวน 22 แห่ง (10.4%) ผลการวิเคราะห์ในโมเดล 3 จัดเป็นกลุ่มต่ำ ผลการวิเคราะห์ในโมเดล 4 จัดเป็นกลุ่มสูง จำนวน 1 แห่ง (0.5%) ผลการวิเคราะห์ในโมเดล 3 จัดเป็นกลุ่มกลาง ผลการวิเคราะห์ในโมเดล 4 จัดเป็นกลุ่มต่ำ จำนวน 18 แห่ง (8.5%) ผลการวิเคราะห์ในโมเดล 3 จัดเป็นกลุ่มกลาง ผลการวิเคราะห์ในโมเดล 4 จัดเป็นกลุ่มสูง จำนวน 19 แห่ง (9.0%) ผลการวิเคราะห์ในโมเดล 3 จัดเป็นกลุ่มสูง แต่ผลการวิเคราะห์ในโมเดล 4 จัดเป็นกลุ่มต่ำ จำนวน 5 แห่ง (2.4%) ผลการวิเคราะห์ในโมเดล 3 จัดเป็นกลุ่มสูง ผลการวิเคราะห์ในโมเดล 4 จัดเป็นกลุ่มกลาง จำนวน 15 แห่ง (7.1%) **ผลการจัดระดับคุณภาพการจัดการศึกษา** พบว่า สถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 ตรงกันกับระดับ

คุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มีจำนวน 2 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 0.94% สถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 สูงกว่าระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มี 115 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 54.25% และสถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 ต่ำกว่าระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มี 95 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 44.81% เมื่อทดสอบความสอดคล้องของการจัดกลุ่มคุณภาพการจัดการศึกษาและระดับคุณภาพการจัดการศึกษาจากการใช้โมเดล 3 และ 4 ด้วยสถิติทดสอบวิลคอกซอน (Wilcoxon Signed Ranks Test) พบว่าโมเดล 3 และ 4 จัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .01 (Z = -6.410 , p = .000) และ จัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 (Z = -5.646, p = .000) แสดงรายละเอียดดังตารางที่ 38

จากผลการเปรียบเทียบการจัดกลุ่มคุณภาพและระดับคุณภาพการจัดการศึกษาของสถานศึกษาระหว่างโมเดล 1 และ 2 และระหว่างโมเดล 3 และ 4 พบว่าการจัดกลุ่มคุณภาพและระดับคุณภาพการจัดการศึกษามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ แสดงให้เห็นว่าคุณลักษณะของตัวแปรระดับนักเรียนและสถานศึกษามีผลต่อการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา

ตารางที่ 37 ผลการจัดกลุ่มคุณภาพการจัดการศึกษา การจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษา และค่าสถิติทดสอบ Wilcoxon Signed Ranks Test จากผลการวิเคราะห์โมเดล 1 และ โมเดล 2

	ผลการจัดกลุ่มคุณภาพ(Rating) การจัดการศึกษา				
โมเดล 1	**กลุ่มคุณภาพการจัดการศึกษา**	**โมเดล 2**			
		กลุ่มต่ำ	**กลุ่มกลาง**	**กลุ่มสูง**	**รวม**
	กลุ่มต่ำ	30 (14.2%)	23 (10.8%)	0 (.0%)	53 (25.0%)
	กลุ่มกลาง	18 (8.5%)	68 (32.1%)	20 (9.4%)	106 (50.0%)
	กลุ่มสูง	5 (2.4%)	15 (7.1%)	33 (15.6%)	53 (25.0%)
	รวม	53 (25.0%)	106 (50.0%)	53 (25.0%)	212 (100.0%)
	Wilcoxon Signed Ranks Test ; Z = -6.403, p = .000				
โมเดล 1	**ผลการจัดระดับคุณภาพ(Ranking) การจัดการศึกษา**				
	ระดับสูงกว่า โมเดล 2		**ระดับตรงกันกับโมเดล 2**		**ระดับต่ำกว่า โมเดล 2**
	115 (54.25%)		2 (0.94%)		95 (44.81%)
	Wilcoxon Signed Ranks Test ; Z = -5.646 , p = .000				

ตารางที่ 38 ผลการจัดกลุ่มคุณภาพการจัดการศึกษา การจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษา และค่าสถิติทดสอบ Wilcoxon Signed Ranks Test จากผลการวิเคราะห์โมเดล 3 และ โมเดล 4

	ผลการจัดกลุ่มคุณภาพ(Rating) การจัดการศึกษา				
โมเดล 3	อันดับคุณภาพการจัดการศึกษา	โมเดล 4			
		กลุ่มต่ำ	กลุ่มกลาง	กลุ่มสูง	รวม
	กลุ่มต่ำ	30(14.2%)	22 (10.4%)	1 (.5%)	53 (25.0%)
	กลุ่มกลาง	18 (8.5%)	69 (32.5%)	19 (9.0%)	106 (50.0%)
	กลุ่มสูง	5 (2.4%)	15 (7.1%)	33 (15.6%)	53 (25.0%)
	รวม	53 (25.0%)	106 (50.0%)	53 (25.0%)	212 (100.0%)
	Wilcoxon Signed Ranks Test ; Z = -6.410 , p = .000				
โมเดล 3	ผลการจัดระดับคุณภาพ(Ranking) การจัดการศึกษา				
	ระดับสูงกว่า โมเดล 4		ระดับตรงกันกับโมเดล 4		ระดับต่ำกว่า โมเดล 4
	115 (54.25%)		2 (0.94%)		95 (44.81%)
	Wilcoxon Signed Ranks Test ; Z = -5.646, p = .000				

3) ผลการเปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาต่อคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน

การเปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาต่อคุณภาพการจัดการศึกษาเป็นการตอบคำถามการวิจัยและวัตถุประสงค์ข้อที่ 3 ได้ผลดังนี้

การเปรียบเทียบผลการวิเคราะห์จากโมเดลการประเมินคุณภาพการจัดการศึกษา 4 โมเดล พบว่าความแปรปรวนของคะแนนผลการประเมินวิชาวิทยาศาสตร์ระหว่างสถานศึกษาในโมเดลที่ไม่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป(โมเดล 1 และโมเดล 2) **โมเดล 1** มีค่าความแปรปรวนระหว่างสถานศึกษา (ระดับที่ 2) เท่ากับ 2010.42084 อย่างมีนัยสำคัญทางสถิติ หรือคิดเป็นความแปรปรวนภายในชั้น (Intraclass correlation) เท่ากับ 0.378 (37.83%) นั่นคือมีความผันแปรของคะแนนผลการประเมินวิทยาศาสตร์ระหว่างสถานศึกษา 37.83% และเมื่อเปรียบเทียบกับ **โมเดล 2** พบว่า เมื่อควบคุมอิทธิพลของตัวแปรทำนายด้วยตัวแปรระดับนักเรียนและสถานศึกษา พบว่าความแปรปรวนของคะแนนผลการประเมินวิชาวิทยาศาสตร์มีค่าลดลง โดยมีค่าความแปรปรวนระหว่างสถานศึกษา(ระดับที่ 2) เท่ากับ 616.57585 อย่างมีนัยสำคัญทางสถิติหรือคิดเป็นความแปรปรวนภายในชั้นเท่ากับ 0.169 (16.93%) นั่นคือมีความผันแปรของคะแนนผลการประเมินวิทยาศาสตร์ระหว่างสถานศึกษา 16.93% สัมประสิทธิ์การทำนาย (R^2) ในระดับที่ 1 เท่ากับ 8.42% นั่นคือในระดับที่ 1 สามารถอธิบายความแปรปรวนของตัวแปรตามได้ 8.42% ของความแปรปรวนภายในชั้น

(Intraclass correlation) ที่มีอยู่ในระดับที่ 1 (83.07%) ดังนั้นความแปรปรวนที่อธิบายได้ในระดับที่ 1 เท่ากับ 6.99% ในขณะที่ระดับที่ 2 มีสัมประสิทธิ์การทำนาย (R^2) เท่ากับ 69.33% นั่นคือในระดับที่ 2 สามารถอธิบายความแปรปรวนของตัวแปรตามได้ 69.33% ของความแปรปรวนภายในชั้น (Intraclass correlation) ที่มีอยู่ในระดับที่ 2 (16.93%) ดังนั้นความแปรปรวนที่อธิบายได้ในระดับที่ 2 เท่ากับ 11.72% และสัมประสิทธิ์การทำนายโดยรวมของโมเดล 2 เท่ากับ 0.7775 (77.75%) นั่นคือโมเดล 2 (ทั้งสองระดับ) สามารถอธิบายความแปรปรวนได้ 77.75% ของความแปรปรวนตัวแปรตาม โดยความแปรปรวนทั้งหมดของตัวแปรตามที่อธิบายในโมเดล 2 ทั้งสองระดับ เท่ากับ 18.71% เมื่อพิจารณาความแปรปรวนของคะแนนผลการประเมินวิชาวิทยาศาสตร์ระหว่างสถานศึกษาในโมเดลที่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป(โมเดล 3 และโมเดล 4) พบว่า **โมเดล 3** มีค่าความแปรปรวนระหว่างสถานศึกษาเท่ากับ 1851.82608 อย่างมีนัยสำคัญทางสถิติ หรือคิดเป็นความแปรปรวนภายในชั้นเท่ากับ 0.381 (38.19%) นั่นคือมีความผันแปรของคะแนนผลการประเมินวิทยาศาสตร์ระหว่างสถานศึกษา และเมื่อเปรียบเทียบกับ **โมเดล 4** พบว่า เมื่อควบคุมอิทธิพลของตัวแปรทำนายด้วยตัวแปรระดับนักเรียนและสถานศึกษา พบว่าความแปรปรวนของคะแนนผลการประเมินวิชาวิทยาศาสตร์มีค่าลดลง โดยมีค่าความแปรปรวนระหว่างสถานศึกษาเท่ากับ 569.18444 อย่างมีนัยสำคัญทางสถิติ หรือคิดเป็นความแปรปรวนภายในชั้นเท่ากับ 0.171 (17.18%) นั่นคือมีความผันแปรของคะแนนผลการประเมินวิทยาศาสตร์ระหว่างสถานศึกษา 17.18% สัมประสิทธิ์การทำนาย (R^2) ในระดับที่ 1 เท่ากับ 16.94% นั่นคือในระดับที่ 1 สามารถอธิบายความแปรปรวนของตัวแปรตามได้ 16.94% ของความแปรปรวนภายในชั้น(Intraclass correlation) ที่มีอยู่ในระดับที่ 1 (82.82%) ดังนั้นความแปรปรวนที่อธิบายได้ในระดับที่ 1 เท่ากับ 14.03% และระดับที่ 2 มีสัมประสิทธิ์การทำนาย (R^2) เท่ากับ 71.68% นั่นคือในระดับที่ 2 สามารถอธิบายความแปรปรวนของตัวแปรตามได้ 71.68% ของความแปรปรวนภายในชั้น(Intraclass correlation) ที่มีอยู่ในระดับที่ 2 (17.18%) ดังนั้นความแปรปรวนที่อธิบายได้ในระดับที่ 2 เท่ากับ 12.26% และสัมประสิทธิ์การทำนายโดยรวมของโมเดล 4 เท่ากับ 0.8863 (88.63%) นั่นคือโมเดล 4 (ทั้งสองระดับ) สามารถอธิบายความแปรปรวนได้ 88.63% ของความแปรปรวนตัวแปรตาม โดยความแปรปรวนทั้งหมดของตัวแปรตามที่อธิบายในโมเดล 4 ทั้งสองระดับ เท่ากับ 26.28%

เมื่อเปรียบเทียบโมเดลที่มีการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษาระหว่างโมเดล 2 ซึ่งเป็นโมเดลที่ไม่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป กับโมเดล 4 ซึ่งเป็นโมเดลที่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป พบว่าโมเดล 2 มีค่าความแปรปรวนระหว่างสถานศึกษาเท่ากับ 616.57585 หรือคิดเป็นความแปรปรวนภายในชั้นเท่ากับ 0.169 (16.93%) และโมเดล 4 พบว่า ความแปรปรวนของคะแนนผลการประเมินวิชาวิทยาศาสตร์มีค่าลดลง โดยมีค่าความแปรปรวนระหว่างสถานศึกษาเท่ากับ 569.18444 หรือคิดเป็นความแปรปรวนภายในชั้นเท่ากับ 0.171 (17.18%) และเมื่อเปรียบเทียบสัมประสิทธิ์การทำนายระหว่างโมเดลที่มีการควบคุมอิทธิพลของตัวแปรระดับนักรียนและสถานศึกษา พบว่าโมเดล 2 มีค่าสัมประสิทธิ์การทำนาย (R^2) มีค่าเท่ากับ 0.7775 (77.75%) ความแปรปรวนทั้งหมดของตัวแปรตามที่อธิบายในโมเดล 2 ทั้งสอง

ระดับ เท่ากับ 18.71% ในขณะที่โมเดล 4 มีค่าสัมประสิทธิ์การทำนาย (R^2) มีค่าเท่ากับ 0.8863 (88.63%) ความแปรปรวนทั้งหมดของตัวแปรตามที่อธิบายในโมเดล 4 ทั้งสองระดับ เท่ากับ 26.28%

แสดงรายละเอียดดังตารางที่ 39

ตารางที่ 39 สรุปค่าความแปรปรวนของตัวแปรตามที่อธิบายได้ของโมเดล 1 ถึง 4

องค์ประกอบความแปรปรวน	Model 1	Model 2	Model 3	Model 4
ระดับที่ 1 ความแปรปรวนภายในสถานศึกษา(R)	3304.54876	3026.18439	2996.81434	2744.59225
ระดับที่ 2 ความแปรปรวนระหว่างสถานศึกษา (U0)	2010.42084	616.57585	1851.82608	569.18444
สัดส่วนความแปรปรวนภายในชั้น (Intraclass Correlation)				
ระดับที่ 1	0.621	0.830	0.618	0.828
	(62.17%)	(83.07%)	(61.81%)	(82.82%)
ระดับที่ 2	0.378	0.169	0.381	0.171
	(37.83%)	(16.93%)	(38.19%)	(17.18%)
สัดส่วนความแปรปรวนของตัวแปรตามที่อธิบายได้ (R^2)				
ระดับที่ 1	-	0.0842	-	0.1694
		(8.42%)		(16.94%)
ระดับที่ 2	-	0.6933	-	0.7168
		(69.33%)		(71.68%)
รวมทั้งสองระดับ	-	0.7775	-	0.8863
		(77.75%)		(88.63%)
ความแปรปรวนของตัวแปรตามที่อธิบายได้				
ระดับที่ 1	-	0.069886	-	0.1402632
		(6.99%)		(14.03%)
ระดับที่ 2	-	0.1171677	-	0.1225728
		(11.72%)		(12.26%)
รวมทั้งสองระดับ	-	0.1870537	-	0.262836
		(18.71%)		(26.28%)

ดังนั้นถ้าพิจารณาความสามารถในการอธิบายความผันแปรของคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์แล้ว **โมเดล 4** เป็นโมเดลที่สามารถอธิบายความผันแปรได้สูงสุด ทั้งนี้เป็นโมเดลที่กำจัดแหล่งความคลาดเคลื่อนของการทำหน้าที่ต่างกันของข้อสอบ และอิทธิพลของคุณลักษณะของนักเรียนและสถานศึกษาเพื่อทำให้ที่อยู่บนพื้นฐานเดียวกัน จึงสามารถกล่าวได้ว่า โมเดล 4 เป็นโมเดลที่ให้ผลการประเมินที่ถูกต้องและมีความเป็นธรรมมากที่สุด

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

การเปรียบเทียบอันดับคุณภาพการจัดการศึกษาของสถานศึกษาเมื่อเทียบกับโมเดล 4

จากผลการวิเคราะห์ความสามารถในการอธิบายความผันแปรของคะแนนสเกลของการประเมินผลการรู้เรื่องด้านวิทยาศาสตร์ แสดงให้เห็นว่า **โมเดล 4** เป็นโมเดลที่สามารถอธิบายความผันแปรของตัวแปรตามได้สูงสุด จึงสามารถสรุปได้ว่าโมเดล 4 เป็นโมเดลที่ให้ผลการประเมินที่ถูกต้องและมีความเป็นธรรมมากที่สุด เนื่องจากเป็นโมเดลที่กำจัดแหล่งความคลาดเคลื่อนของการทำหน้าที่ต่างกันของข้อสอบ และอิทธิพลของคุณลักษณะของนักเรียนและสถานศึกษาเพื่อทำให้ที่อยู่บนพื้นฐานเดียวกัน ผู้วิจัยจึงเสนอผลการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาขั้นพื้นฐานโดยการเปรียบเทียบโมเดล 1 2 และ 3 กับโมเดล 4 พบว่าโมเดล 2 ให้ผลการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาได้สอดคล้องกัน หรือจัดสถานศึกษาให้อยู่ในกลุ่มคุณภาพการจัดการศึกษาได้ตรงกัน

แสดงรายละเอียดดังตารางที่ 40

ตารางที่ 40 ผลการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาโดยเปรียบเทียบกับโมเดล 4

		โมเดล 1			โมเดล 2			โมเดล 3			รวม
	กลุ่มคุณภาพการจัดการศึกษา	ต่ำ	กลาง	สูง	ต่ำ	กลาง	สูง	ต่ำ	กลาง	สูง	
โมเดล 4	ต่ำ	30 (14.2%)	23 (10.8%)	0 (.0%)	53 (25.0%)	0 (.0%)	0 (.0%)	30 (14.2%)	22 (10.4%)	1 (.5%)	53 (25.0%)
	กลาง	18 (8.5%)	68 (32.1%)	20 (9.4%)	0 (.0%)	106 (50.0%)	0 (.0%)	18 (8.5%)	69 (32.5%)	19 (9.0%)	106 (50.0%)
	สูง	5 (2.4%)	15 (7.1%)	33 (15.6%)	0 (.0%)	0 (.0%)	53 (25.0%)	5 (2.4%)	15 (7.1%)	33 (15.6%)	53 (25.0%)
รวม		53 (25.0%)	106 (50.0%)	53 (25.0%)	53 (25.0%)	106 (50.0%)	53 (25.0%)	53 (25.0%)	106 (50.0%)	53 (25.0%)	212 (100.0%)

การประเมินคุณภาพการจัดการศึกษาของสถานศึกษา ด้วยโมเดลการประเมินคุณภาพการจัดการศึกษาที่ต่างกัน จะได้ผลการประเมินคุณภาพการจัดการศึกษาที่ต่างกัน ผลดังกล่าวสรุปได้ว่า โมเดล 1 และโมเดล 2 ให้ผลการประเมินคุณภาพการจัดการศึกษาที่ต่างกันหรือไม่สอดคล้องกันมากถึง 38.1% และระหว่างโมเดล 3 และโมเดล 4 ให้ผลการประเมินคุณภาพการจัดการศึกษาที่ต่างกันหรือไม่สอดคล้องกัน 37.7% ส่วนโมเดล 2 และโมเดล 4 ให้ผลการประเมินคุณภาพการจัดการศึกษาที่ให้ผลตรงกันหรือสอดคล้องกันมากที่สุด (100.0%) และระหว่างโมเดล 1 และโมเดล 3 ให้ผลการประเมินคุณภาพการจัดการศึกษาที่ให้ผลตรงกัน 98.2% แสดงรายละเอียดดังตารางที่ 41

ตารางที่ 41 ผลการวิเคราะห์ความสอดคล้องของการจัดกลุ่มคุณภาพการจัดการศึกษาด้วยโมเดลการประเมินคุณภาพการจัดการศึกษาที่ต่างกัน 4 โมเดล

โมเดลคู่เปรียบเทียบ	ประเด็นเปรียบเทียบของโมเดล				จำนวนสถานศึกษาที่สอดคล้อง	จำนวนสถานศึกษาที่ไม่สอดคล้อง	รวม
	ไม่ตัด DIF	ตัด DIF	ไม่ควบคุมตัวแปร	ควบคุมตัวแปร			
โมเดล 1 – โมเดล 3	✓	✓	ไม่ควบคุมตัวแปร		208 (98.2%)	4 (1.8%)	212 (100.0%)
โมเดล 2 – โมเดล 4	✓	✓	ควบคุมตัวแปร		212 (100.0%)	0 (.0%)	212 (100.0%)
โมเดล 1 – โมเดล 2	ไม่ตัดข้อสอบ DIF		✓	✓	131 (61.9%)	81 (38.1%)	212 (100.0%)
โมเดล 3 – โมเดล 4	ตัดข้อสอบ DIF		✓	✓	132 (62.3%)	80 (37.7%)	212 (100.0%)

จากข้อค้นพบดังกล่าว แสดงให้เห็นประเด็นเกี่ยวกับการพิจารณาโมเดลการประเมินคุณภาพการจัดการศึกษา 2 ประเด็น ดังนี้

ประเด็นที่ 1 การมีข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบ โดยการเปรียบเทียบระหว่างโมเดล 1 และ 3 และระหว่างโมเดล 2 และ 4 พบว่าการประเมินคุณภาพการจัดการศึกษาด้วยโมเดลที่ไม่มีการตัดข้อสอบที่ทำหน้าที่กันออกไป กับโมเดลที่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป จะให้ผลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาที่สอดคล้องหรือให้ผลการประเมินที่ตรงกัน 98.2% แต่ถ้าต้องการให้ผลการประเมินตรงกันมากขึ้น ต้องมีการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา 100.0%

ประเด็นที่ 2 การควบคุมอิทธิพลของคุณลักษณะนักเรียนและสถานศึกษา โดยเปรียบเทียบระหว่างโมเดล 1 และ 2 และระหว่างโมเดล 3 และ 4 พบว่าการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาด้วยโมเดลที่ไม่มีการควบคุมตัวแปรระดับนักเรียนและสถานศึกษา กับโมเดลที่มีการควบคุมตัวแปรระดับนักเรียนและสถานศึกษา จะให้ผลการประเมินคุณภาพการจัดการศึกษาที่ตรงกัน 61.9% แต่เมื่อมีการตรวจสอบและตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปจะทำให้ผลการประเมินที่ตรงกันเพิ่มขึ้นเป็น 62.3%

เมื่อพิจารณาความสัมพันธ์ของการจัดตำแหน่งคะแนนมูลค่าเพิ่มจากการวิเคราะห์ 4 โมเดล ด้วยการตรวจสอบสหสัมพันธ์ตำแหน่งคะแนนของสเปียร์แมน (Spearman's Rank – Order Correlation) พบว่า โมเดลที่มีการควบคุมตัวแปรระดับนักเรียนและสถานศึกษา ระหว่างโมเดลไม่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป (โมเดล 2) กับโมเดลตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป (โมเดล 4) มีค่าสหสัมพันธ์กันสูงสุด เท่ากับ 1.000 และโมเดลที่ไม่มีการควบคุมตัวแปรระดับนักเรียนและสถานศึกษา ระหว่างโมเดลที่ไม่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป (โมเดล 1) กับโมเดลที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป (โมเดล 3) มีค่าสหสัมพันธ์เท่ากับ 0.981 ส่วนโมเดลที่ไม่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป ระหว่างโมเดลที่ไม่มีการควบคุมตัวแปรระดับนักเรียนและสถานศึกษา (โมเดล 1) กับโมเดลที่มีการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและ

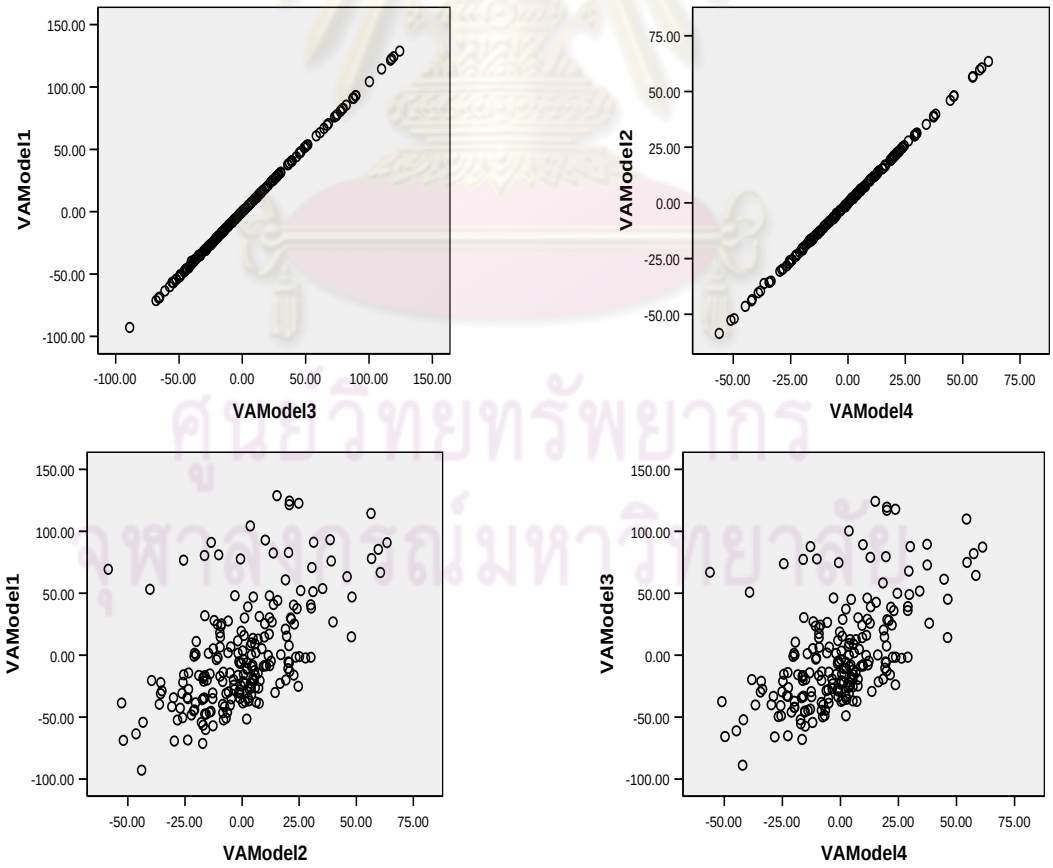
สถานศึกษา (โมเดล 2) มีค่าสหสัมพันธ์เท่ากับ 0.547 และโมเดลที่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป ระหว่างโมเดลที่ไม่มีการควบคุมตัวแปรระดับนักเรียนและสถานศึกษา (โมเดล 3) กับโมเดลที่มีการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา (โมเดล 4) มีค่าสหสัมพันธ์เท่ากับ 0.538 แสดงรายละเอียดดังตารางที่ 42

ตารางที่ 42 ค่าสหสัมพันธ์ของระดับของคะแนนมูลค่าเพิ่มที่ได้จากโมเดลประเมินคุณภาพการจัดการศึกษาที่ต่างกัน 4 โมเดล

	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
โมเดล 1	1.000			
โมเดล 2	.547**	1.000		
โมเดล 3	.981**	.538**	1.000	
โมเดล 4	.547**	1.000**	.538**	1.000

**p < .01

อย่างไรก็ตาม เมื่อพล็อตกราฟของตำแหน่งคะแนนมูลค่าเพิ่มของสถานศึกษา พบว่าโมเดล 1 กับโมเดล 3 และโมเดล 2 กับโมเดล 4 มีลักษณะของกราฟเป็นเส้นตรง แต่โมเดล 1 กับ 2 และโมเดล 3 กับ 4 ยังมีการกระจายของตำแหน่งคะแนนมูลค่าเพิ่มหรือระดับคุณภาพการจัดการศึกษาที่แตกต่างจากเชิงเส้นตรง แสดงดังภาพที่ 21



ภาพที่ 21 ความสัมพันธ์ของตำแหน่งคะแนนผลการประเมินคุณภาพการจัดการศึกษา 4 โมเดล

สรุปแนวทางการเลือกใช้โมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา

การเปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาที่ต่างกัน โดยสรุปจากผลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาที่สอดคล้องกัน ค่าสหสัมพันธ์ของตำแหน่งคะแนนมูลค่าเพิ่มที่ได้จากการวิเคราะห์ และค่าสัมประสิทธิ์การทำนายของโมเดลการวิเคราะห์สามารถสรุปลักษณะเด่นของแต่ละโมเดล และแนวทางการเลือกโมเดลการประเมินคุณภาพการจัดการศึกษาที่มีคุณภาพและให้ผลที่น่าเชื่อมากที่สุด ดังตารางที่ 43

ตารางที่ 43 สรุปการเปรียบเทียบลักษณะโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาและแนวทางการเลือกใช้โมเดลประเมินคุณภาพการจัดการศึกษาของสถานศึกษา

โมเดล	ลักษณะการวิเคราะห์	การเปรียบเทียบของโมเดล	การพิจารณาแนวทางการเลือกใช้โมเดล
โมเดล 1	โมเดลไม่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปและไม่มีปรับแก้คะแนนด้วยการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา	- ความสอดคล้องของผลการประเมินกับโมเดล 3 มากกว่าโมเดล 2 - ค่าสหสัมพันธ์ของตำแหน่งคะแนนกับโมเดล 3 มากกว่าโมเดล 2	- เป็นโมเดลการประเมินที่ง่ายต่อการนำไปใช้และใช้กันโดยทั่วไป เนื่องจากใช้ผลคะแนนจากสอบของแต่ละสถานศึกษา เป็นผลคุณภาพการจัดการศึกษาของสถานศึกษา และไม่พิจารณาหรือคำนึงถึงคุณภาพของแบบสอบและอิทธิพลของตัวแปรภายนอกระดับนักเรียน/สถานศึกษา - เป็นโมเดลที่มีการวิเคราะห์ไม่ซับซ้อน และไม่จำเป็นต้องเก็บข้อมูลคุณลักษณะของนักเรียน/สถานศึกษา
โมเดล 2	โมเดลไม่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปและมีปรับแก้คะแนนด้วยการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา	- ความสอดคล้องของผลการประเมินกับโมเดล 4 มากกว่าโมเดล 1 - ค่าสหสัมพันธ์ของตำแหน่งคะแนนกับโมเดล 4 มากกว่าโมเดล 1 - ค่าสัมประสิทธิ์การทำนาย (R^2) และความแปรปรวนของคะแนนที่สามารถอธิบายได้ของโมเดล น้อยกว่าโมเดล 4	- เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาที่ใช้ผลคะแนนจากข้อสอบทั้งฉบับเช่นเดียวกับโมเดล 1 แต่มีการควบคุมอิทธิพลคุณลักษณะของนักเรียน/สถานศึกษา - การควบคุมอิทธิพลของคุณลักษณะของนักเรียน/สถานศึกษาทำให้ผลการประเมินถูกปรับให้อยู่บนพื้นฐานของปัจจัยดั้งเดิมเท่าเทียมกัน และนำผลต่างหรือส่วนที่เพิ่มขึ้นจากการควบคุมมาเปรียบเทียบกันอย่างยุติธรรม - เป็นโมเดลที่มีการวิเคราะห์ไม่ซับซ้อนเท่ากับโมเดล 4 และให้ผลการประเมินคุณภาพการจัดการศึกษาสอดคล้องกับโมเดล 4

ตารางที่ 43 สรุปการเปรียบเทียบลักษณะโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาและแนวทางการเลือกใช้โมเดลประเมินคุณภาพการจัดการศึกษาของสถานศึกษา(ต่อ)

โมเดล	ลักษณะการวิเคราะห์	การเปรียบเทียบของโมเดล	การพิจารณาแนวทางการเลือกใช้โมเดล
โมเดล 3	โมเดลมีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปและไม่มีปรับแก้คะแนนด้วยการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา	- ความสอดคล้องของผลการประเมินกับโมเดล 1 มากกว่าโมเดล 4 - ค่าสหสัมพันธ์ของตำแหน่งคะแนนกับโมเดล 1 มากกว่าโมเดล 3	- เป็นโมเดลการประเมินที่ไม่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะนักเรียนแลสถานศึกษา แต่มีการตรวจสอบคุณภาพของแบบสอบและตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปก่อน ดังนั้นคะแนนที่จากแบบสอบ จึงมีการควบคุมโอกาสในการตอบถูกให้เท่าเทียมกันในคุณลักษณะที่ต่างกัน - เป็นโมเดลที่มีการวิเคราะห์ซับซ้อนมากกว่าโมเดล 1 แต่ให้ผลการประเมินคุณภาพการจัดการศึกษาที่ค่อนข้างสอดคล้องกันสูงกับโมเดล 1
โมเดล 4	โมเดลมีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปและมีปรับแก้คะแนนด้วยการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา	- ความสอดคล้องของผลการประเมินกับโมเดล 2 มากกว่าโมเดล 3 - ค่าสหสัมพันธ์ของตำแหน่งคะแนนกับโมเดล 2 มากกว่าโมเดล 3 - ค่าสัมประสิทธิ์การทำนาย (R^2) และความแปรปรวนของคะแนนที่สามารถอธิบายได้ของโมเดลมากกว่าโมเดล 2	- เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาที่ถูกต้องและเป็นธรรมมากที่สุด - เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาที่มีการวิเคราะห์ที่ซับซ้อนมากที่สุด มีการตรวจสอบคุณภาพของแบบสอบและตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปเช่นเดียวกับโมเดล 3 และควบคุมอิทธิลของคุณลักษณะของนักเรียน/สถานศึกษา เหมือนกับโมเดล 2 - เป็นโมเดลที่มีกำจัดข้อสอบที่ทำหน้าที่ต่างกันออกไป และควบคุมอิทธิพลของคุณลักษณะของนักเรียน/สถานศึกษาทำให้ผลการประเมินมีความเท่าเทียม และยุติธรรมสำหรับกลุ่มนักเรียนมากที่สุด - เป็นโมเดลที่มีใช้เทคนิคการวิเคราะห์ขั้นสูงและซับซ้อน แต่ให้ผลการประเมินคุณภาพการจัดการศึกษาที่สอดคล้องกับโมเดล 2

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5

สรุปผลการวิจัย อภิปรายผลการวิจัย และข้อเสนอแนะ

การวิจัยในครั้งนี้มีจุดมุ่งหมายหลักเพื่อประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน ซึ่งมีวัตถุประสงค์เฉพาะ 3 ข้อ คือ 1) เพื่อศึกษาผลของข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบวิชาวิทยาศาสตร์ต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน 2) เพื่อศึกษาผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาขั้นพื้นฐาน ต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน และ 3) เพื่อเปรียบเทียบคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานจากการประเมินคุณภาพการจัดการศึกษาเมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน โดยข้อมูลที่ใช้ในการวิจัยเป็นข้อมูลจากโครงการประเมินผลนักเรียนระดับนานาชาติวิชาวิทยาศาสตร์ ปี 2549 หรือ PISA 2006

โมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน ในการวิจัยครั้งนี้พิจารณาใน 2 ประเด็น คือคุณภาพของแบบสอบ และโมเดลที่ใช้ในการวิเคราะห์มูลค่าเพิ่ม จึงกำหนดเป็นโมเดลการประเมินคุณภาพการจัดการศึกษา 4 โมเดล คือ**โมเดล 1** เป็นโมเดลที่ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์และโมเดลการวิเคราะห์มูลค่าเพิ่มเป็นโมเดลที่ยังไม่มีการปรับแก้คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา **โมเดล 2** เป็นโมเดลที่ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์และโมเดลการวิเคราะห์มูลค่าเพิ่มเป็นโมเดลที่มีการปรับแก้คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา **โมเดล 3** เป็นโมเดลที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์และโมเดลการวิเคราะห์มูลค่าเพิ่มเป็นโมเดลที่ยังไม่มีการปรับแก้คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา และ**โมเดล 4** เป็นโมเดลที่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกก่อนนำคะแนนมาใช้ในการวิเคราะห์และโมเดลการวิเคราะห์มูลค่าเพิ่มเป็นโมเดลที่มีการปรับแก้คะแนนด้วยตัวแปรระดับนักเรียนและสถานศึกษา

การวิเคราะห์ข้อมูลเพื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบและการวิเคราะห์มูลค่าเพิ่ม ผู้วิจัยใช้โปรแกรมโมเดลเชิงเส้นระดับลดหลั่น (HLM) เวอร์ชั่น 6.06 โดยการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบประยุกต์ใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น(Hierarchical Generalized Linear Model: HGLM) 3 ระดับ (HGLM-3) โดยมีระดับที่ 1 เป็นระดับข้อสอบ ระดับที่ 2 เป็นระดับนักเรียนและระดับที่ 3 เป็นระดับสถานศึกษา ส่วนการวิเคราะห์มูลค่าเพิ่ม ผู้วิจัยใช้โมเดลเชิงเส้นตรงแบบลดหลั่น (Hierarchical Linear Model: HLM) 2 ระดับ โดยมีระดับที่ 1 เป็นระดับนักเรียน และระดับที่ 2 เป็นระดับสถานศึกษา และคะแนนมูลค่าเพิ่มของสถานศึกษาหรือค่าส่วนที่เหลือที่ได้จากการวิเคราะห์พหุระดับจากการวิเคราะห์ในแต่ละโมเดลการประเมินคุณภาพการจัดการศึกษา จะใช้เป็นคะแนนเพื่อกำหนดคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน

5.1 สรุปผลการวิจัย

สรุปผลการวิจัย จำแนกตามวัตถุประสงค์การวิจัย ดังนี้

1. **ผลของข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบวิชาวิทยาศาสตร์ต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน** สรุปผลออกเป็น 2 ข้อดังนี้

1.1 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบ 13 ฉบับ ได้นำตัวแปรระดับนักเรียนมาจัดกลุ่มเป็นสองกลุ่มในแต่ละคุณลักษณะ ได้แก่ เพศ(ชาย หญิง) การเรียนพิเศษ(เรียน ไม่เรียน) เศรษฐานะของครอบครัว (สูง ต่ำ) ทรัพยากรของครอบครัว (สูง ต่ำ) ความมั่งคั่งของครอบครัว (สูง ต่ำ) และใช้ตัวแปรระดับสถานศึกษาเพื่ออธิบายถึงปัจจัยที่ทำให้เกิดการทำหน้าที่ต่างกันของข้อสอบ พบว่าแบบสอบที่ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันมากที่สุดคือแบบสอบฉบับที่ 2 พบข้อสอบที่ทำหน้าที่ต่างกัน 3 ข้อ จากข้อสอบ 26 ข้อ คิดเป็น 11.53% ลำดับต่อมาเป็นแบบสอบฉบับที่ 4 พบข้อสอบที่ทำหน้าที่ต่างกัน 6 ข้อ จากข้อสอบ 33 ข้อ คิดเป็น 18.18% และแบบสอบฉบับที่ 11 พบข้อสอบที่ทำหน้าที่ต่างกัน 6 ข้อ จากข้อสอบ 28 ข้อ คิดเป็น 21.42% ส่วนแบบสอบที่ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันมากที่สุด คือแบบสอบฉบับที่ 10 พบข้อสอบที่ทำหน้าที่ต่างกัน 19 จากข้อสอบ 28 ข้อ คิดเป็น 67.85% และลำดับถัดมาคือแบบสอบฉบับที่ 7 พบข้อสอบที่ทำหน้าที่ต่างกัน 10 ข้อจากข้อสอบ 15 ข้อ คิดเป็น 66.67% และฉบับที่ 12 พบข้อสอบที่ทำหน้าที่ต่างกัน 20 ข้อจากข้อสอบ 30 ข้อ คิดเป็น 66.67%

เมื่อพิจารณาคุณลักษณะของตัวแปรที่มีอิทธิพลให้เกิดการทำหน้าที่ต่างกันของข้อสอบ พบว่า ข้อสอบที่โอนเอียงให้เพศชาย ตรวจพบมากที่สุดในแบบสอบฉบับที่ 1 จำนวน 23 ข้อ คิดเป็น 33.8% จากข้อสอบที่โอนเอียงให้เพศชายทั้ง 13 ฉบับ ข้อสอบที่โอนเอียงให้เพศหญิงตรวจพบมากที่สุดในแบบสอบฉบับ 7 จำนวน 5 ข้อ (41.7%) ข้อสอบที่โอนเอียงให้นักเรียนที่เรียนพิเศษ ตรวจพบมากที่สุดในแบบสอบฉบับที่ 10 จำนวน 9 ข้อ (50.0%) ข้อสอบที่โอนเอียงให้นักเรียนที่ไม่เรียนพิเศษ ตรวจพบมากที่สุดในแบบสอบฉบับที่ 3 จำนวน 3 ข้อ(66.7%) ข้อสอบที่โอนเอียงให้นักเรียนที่มีเศรษฐานะครอบครัวสูง ตรวจพบมากที่สุดในแบบสอบฉบับที่ 1 จำนวน 17 ข้อ(47.2%) ข้อสอบที่โอนเอียงให้นักเรียนที่มีเศรษฐานะของครัวต่ำ ตรวจพบมากที่สุดในแบบสอบฉบับที่ 7 จำนวน 5 ข้อ (31.3%) ข้อสอบที่โอนเอียงให้นักเรียนที่มีทรัพยากรของครอบครัวสูง ตรวจพบมากที่สุดในแบบสอบฉบับที่ 5 จำนวน 3 ข้อ (50.0%) ข้อสอบที่โอนเอียงให้นักเรียนที่มีทรัพยากรของครอบครัวต่ำ ตรวจพบมากที่สุดในแบบสอบฉบับที่ 8 จำนวน 14 ข้อ (82.4%) ข้อสอบที่โอนเอียงให้นักเรียนที่มีความมั่งคั่งของครอบครัวสูง ตรวจพบมากที่สุดในแบบสอบฉบับที่ 10 และฉบับที่ 12 จำนวน 9 ข้อ (33.3%) และข้อสอบที่โอนเอียงให้นักเรียนที่มีความมั่งคั่งของครอบครัวต่ำ ตรวจพบมากที่สุดในแบบสอบฉบับที่ 6 จำนวน 6 ข้อ (31.6%)

1.2 ผลของข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบวิชาวิทยาศาสตร์ต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน

จากการเปรียบเทียบผลการประเมินคุณภาพการจัดการศึกษาระหว่างโมเดล 1 กับโมเดล 3 และระหว่างโมเดล 2 กับโมเดล 4 พบว่า โมเดล 1 กับโมเดล 3 จัดกลุ่มคุณภาพ (Rating) การจัดการศึกษาได้สอดคล้องกันหรือจัดสถานศึกษาให้อยู่ในกลุ่มคุณภาพการจัดการศึกษาเดียวกันได้จำนวน 208 แห่ง คิดเป็นร้อยละ 98.2 และจัดกลุ่มคุณภาพได้ไม่สอดคล้องกันจำนวน 4 แห่ง คิดเป็นร้อยละ 1.8 และเมื่อพิจารณาระดับคุณภาพ (Ranking) การจัดการศึกษาของสถานศึกษาพบว่าสถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 ตรงกันกับระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 มีจำนวน 133 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 62.74% ส่วนผลการประเมินระหว่างโมเดล 2 และโมเดล 4 จัดกลุ่มคุณภาพการจัดการศึกษาได้สอดคล้องกัน 212 แห่ง คิดเป็นร้อยละ 100.0 และเมื่อพิจารณาระดับคุณภาพการจัดการศึกษาของสถานศึกษาพบว่าสถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 2 ตรงกันกับระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มีจำนวน 110 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 51.89%

จากผลของสถิติทดสอบวิลคอกซอน (Wilcoxon Signed Ranks Test) ในการทดสอบความสอดคล้องของการจัดกลุ่มคุณภาพและระดับคุณภาพการจัดการศึกษา พบว่าโมเดล 1 และ 3 จัดกลุ่มคุณภาพและระดับคุณภาพการจัดการศึกษาของสถานศึกษาไม่แตกต่างกัน และเมื่อทดสอบการจัดกลุ่มคุณภาพและระดับคุณภาพการจัดการศึกษาจากการใช้โมเดล 2 และ 4 พบว่าโมเดล 2 และ 4 จัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาไม่แตกต่างกัน และการจัดระดับคุณภาพการจัดการศึกษาของสถานศึกษาแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05

จากผลการเปรียบเทียบการจัดกลุ่มคุณภาพและระดับคุณภาพการจัดการศึกษาของสถานศึกษาระหว่างโมเดล 1 และ 3 แสดงให้เห็นว่าโมเดลที่ไม่ตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปและตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป ทำให้ผลการจัดกลุ่มคุณภาพและระดับคุณภาพการจัดการศึกษาของสถานศึกษาไม่ต่างกัน นั่นคือข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบวิชาวิทยาศาสตร์ไม่มีผลต่อการประเมินคุณภาพการจัดการศึกษา และจากผลการวิเคราะห์ของโมเดล 2 และ 4 พบว่าการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาไม่ต่างกัน แต่ระดับคุณภาพการจัดการศึกษาแตกต่างกัน แสดงว่าแบบสอบที่มีข้อสอบที่ทำหน้าที่ต่างกันไม่มีผลการประเมินคุณภาพการจัดการศึกษา แต่เมื่อมีการควบคุมอิทธิพลของคุณลักษณะนักเรียนและสถานศึกษาจะส่งผลให้การประเมินคุณภาพมีระดับคุณภาพการจัดการศึกษาของสถานศึกษาที่แตกต่างกัน

2. ผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาขั้นพื้นฐาน และคุณภาพของแบบสอบที่ต่างกันของวิชาวิทยาศาสตร์ต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน เมื่อใช้โมเดลการประเมินคุณภาพการจัดการศึกษาต่างกัน

สรุปผลออกเป็น 2 ข้อดังนี้

2.1 ตัวแปรที่ส่งผลต่อคะแนนมูลค่าเพิ่มวิชาวิทยาศาสตร์

จากผลการวิเคราะห์มูลค่าเพิ่มของคะแนนผลการประเมินวิชาวิทยาศาสตร์ พบว่า**ตัวแปรระดับสถานศึกษา** ที่มีผลต่อคะแนนผลการประเมินวิชาวิทยาศาสตร์แตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ได้แก่ สถานศึกษาขนาดใหญ่ สถานศึกษาขนาดกลาง สถานศึกษาสังกัด สกอ. สถานศึกษาที่ตั้งอยู่ในหมู่บ้าน สถานศึกษาที่ตั้งอยู่ในเมืองเล็ก ร้อยละงบประมาณของรัฐที่จัดสรรให้สถานศึกษา ดัชนีทรัพยากรการเรียน สัดส่วนนักเรียนต่อครู และตัวแปรที่ส่งผลต่อคะแนนผลการประเมินวิชาวิทยาศาสตร์อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ สถานศึกษาสังกัด สช. สถานศึกษาสังกัด กศท. และ กทม. สถานศึกษาที่ตั้งอยู่ในเมือง ดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล ส่วนตัวแปรคุณลักษณะระดับสถานศึกษาที่ไม่มีผลต่อคะแนนผลการประเมินวิชาวิทยาศาสตร์ ได้แก่ สถานศึกษาสังกัด สพฐ. สถานศึกษาที่ตั้งอยู่ในเมืองใหญ่ ขนาดห้องเรียน ระดับการขาดแคลนครูวิทยาศาสตร์ ระดับการขาดแคลนครูที่มีคุณวุฒิ ดัชนีการขาดแคลนครู ดัชนีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์ ดัชนีความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากร และร้อยละของงบประมาณที่สถานศึกษาจัดสรรจากแหล่งอื่น เมื่อพิจารณาขนาดอิทธิพลหรือค่าสัมประสิทธิ์ถดถอย พบว่า ตัวแปรที่มีค่าสัมประสิทธิ์ถดถอยที่เป็นบวกสูงสุด คือ สถานศึกษาสังกัด สกอ.หรือโรงเรียนสาธิต (β = 52.409790) รองลงมาคือ สถานศึกษาขนาดใหญ่ (β = 23.849150) ส่วนตัวแปรที่มีค่าสัมประสิทธิ์ถดถอยที่เป็นลบสูงสุด คือสถานศึกษาที่ตั้งอยู่ในเมืองเล็ก (β = -32.883962) สถานศึกษาที่ตั้งอยู่ในหมู่บ้าน (β = -28.640284)

ตัวแปรระดับนักเรียน ที่มีผลต่อคะแนนผลการประเมินวิชาวิทยาศาสตร์ที่ค่าสัมประสิทธิ์ถดถอยของตัวแปรแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 คือ ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา ระดับการใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง เศรษฐานะของครอบครัว แหล่งทรัพยากรที่บ้าน ความมั่งคั่งของครอบครัว ความสนใจในการเรียนรู้วิทยาศาสตร์ ความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์ ความเพลิดเพลินทางวิทยาศาสตร์ แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์ การเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์ ส่วนตัวแปรที่ไม่มีผลต่อคะแนนผลการประเมินวิชาวิทยาศาสตร์ ได้แก่ ความเป็นเพศชาย ระดับการใช้เวลาเรียนพิเศษวิทยาศาสตร์นอกสถานศึกษา และความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์ เมื่อพิจารณาขนาดอิทธิพลหรือค่าสัมประสิทธิ์ถดถอย พบว่าตัวแปรระดับนักเรียนที่มีอิทธิพลทางบวกสูงสุดคือ ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา (β = 14.586167) รองลงมาคือ ความเพลิดเพลินทางวิทยาศาสตร์ (β = 8.550603) และระดับการใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง (β = 8.003062) ส่วนตัวแปรระดับนักเรียนที่มีอิทธิพลทางลบสูงสุดคือความเชื่อใน

ความสามารถของตนเองในการเรียนวิทยาศาสตร์ (β = -7.685855) รองลงมาคือความมั่งคั่งของครอบครัว (β = -3.944037)

2.2 ผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาขั้นต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน

จากการเปรียบเทียบผลการประเมินคุณภาพการจัดการศึกษาระหว่างโมเดล 1 กับโมเดล 2 และระหว่างโมเดล 3 กับโมเดล 4 พบว่า โมเดล 1 กับโมเดล 2 จัดกลุ่มคุณภาพการจัดการศึกษาได้สอดคล้องกันหรือจัดอยู่สถานศึกษาในกลุ่มคุณภาพเดียวกันคิดเป็นร้อยละ 61.9 เมื่อพิจารณาระดับคุณภาพการจัดการศึกษาของสถานศึกษา พบว่าสถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 1 ตรงกันกับระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 2 มีจำนวน 2 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 0.94% และโมเดล 3 กับโมเดล 4 จัดกลุ่มคุณภาพการจัดการศึกษาได้สอดคล้องกัน คิดเป็นร้อยละ 62.3 เมื่อพิจารณาระดับคุณภาพการจัดการศึกษาของสถานศึกษา พบว่าสถานศึกษาที่มีระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 3 ตรงกันกับระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์โมเดล 4 มีจำนวน 2 แห่งจากสถานศึกษาทั้งหมด 212 แห่ง คิดเป็น 0.94% เช่นเดียวกัน

จากผลของสถิติทดสอบวิลคอกซอน (Wilcoxon Signed Ranks Test) พบว่าการจัดกลุ่มคุณภาพการจัดการศึกษาและระดับคุณภาพการจัดการศึกษาของสถานศึกษาระหว่างโมเดล 1 กับโมเดล 2 และระหว่างโมเดล 3 กับโมเดล 4 มีการจัดกลุ่มคุณภาพและระดับคุณภาพการจัดการศึกษาแตกต่างกันอย่างมีนัยสำคัญทางสถิติ แสดงให้เห็นว่าคุณลักษณะของตัวแปรระดับนักเรียนและสถานศึกษามีผลต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน

3. การเปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ต่อคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน

สรุปผลออกเป็น 2 ข้อดังนี้

3.1 ผลการเปรียบเทียบคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐานจากการใช้โมเดลการประเมินต่างกัน

<u>ความสอดคล้องของผลการประเมิน</u> จากการวิเคราะห์ด้วยโมเดลประเมินคุณภาพการจัดการศึกษา 4 โมเดล พบว่าระดับคุณภาพการจัดการศึกษาจากผลการวิเคราะห์คะแนนมูลค่าเพิ่มแต่ละโมเดลแตกต่างกัน เมื่อจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษา โดยแบ่งเป็น 3 กลุ่ม คือกลุ่มต่ำ กลุ่มกลาง กลุ่มสูง พบว่าผลการประเมินคุณภาพการจัดการศึกษาของโมเดลที่สอดคล้องกันคือ โมเดล 2 กับโมเดล 4 สอดคล้องกันมากที่สุด (100%) นั่นคือโมเดลที่มีการควบคุมอิทธิพลของคุณลักษณะของนักเรียนและสถานศึกษาจะมีผลการประเมินสอดคล้องกันมาก

ที่สุด รองลงมาคือ โมเดล 1 กับโมเดล 3 (98.2%) และ โมเดล 3 กับโมเดล 4 (62.3%) และโมเดล 1 กับ โมเดล 2 (61.9%)

<u>สัมประสิทธิ์การทำนาย</u> หรือ ความแปรปรวนของคะแนนผลการประเมินวิชาวิทยาศาสตร์ที่อธิบายได้ด้วยคุณลักษณะของนักเรียนและสถานศึกษา พบว่า โมเดล 2 มีค่าสัมประสิทธิ์การทำนาย (R^2) เท่ากับ 0.7775 (77.75%) และโมเดล 4 มีค่าสัมประสิทธิ์การทำนาย (R^2) เท่ากับ 0.8863 (88.63%)

<u>สหสัมพันธ์ของตำแหน่งคะแนนมูลค่าเพิ่ม</u> ด้วยการตรวจสอบสหสัมพันธ์ตำแหน่งคะแนนของสเปียร์แมน พบว่า โมเดล 2 กับโมเดล 4 มีค่าสหสัมพันธ์กันสูงสุด เท่ากับ 1.000 และโมเดล 1 กับโมเดล 3 มีค่าสหสัมพันธ์เท่ากับ 0.981 โมเดล 1 กับโมเดล 2 มีค่าสหสัมพันธ์เท่ากับ 0.547 และโมเดล 3 กับโมเดล 4 มีค่าสหสัมพันธ์เท่ากับ 0.538

สรุปผลได้ดังตารางที่ 44

ตารางที่ 44 สรุปความสอดคล้องของผลการประเมิน สัมประสิทธิ์การทำนายและค่าสหสัมพันธ์ของตำแหน่งคะแนนมูลค่าเพิ่มจากผลการประเมิน 4 โมเดล

โมเดล	ความสอดคล้องของผลการประเมิน			สัมประสิทธิ์การทำนาย	ค่าสหสัมพันธ์		
	โมเดล 2	โมเดล 3	โมเดล 4		โมเดล 2	โมเดล 3	โมเดล 4
โมเดล 1	61.9%	98.2%	61.9%	-	.547	.981	.547
โมเดล 2	-	62.3%	100.0%	77.75%	-	.538	1.000
โมเดล 3	-	-	62.3%	-	-	-	.538
โมเดล 4	-	-	-	88.63%	-	-	-

3.2 สรุปแนวทางการเลือกใช้โมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาขั้นพื้นฐาน

จากการเปรียบเทียบผลการประเมินคุณภาพการจัดการศึกษาของทั้ง 4 โมเดล สามารถสรุปเป็นแนวทางการพิจารณาเลือกโมเดลการประเมินคุณภาพการจัดการศึกษา แต่ละโมเดล ดังนี้

โมเดล 1 เป็นโมเดลการประเมินที่ง่ายต่อการนำไปใช้ และการวิเคราะห์ไม่ซับซ้อน สามารถใช้กันโดยทั่วไป เนื่องจากใช้ผลคะแนนจากสอบของแต่ละสถานศึกษา เป็นผลคุณภาพการจัดการศึกษาของสถานศึกษา และไม่พิจารณาหรือคำนึงถึงคุณภาพของแบบสอบและอิทธิพลของตัวแปรภายนอกระดับนักเรียน/สถานศึกษา และเป็นโมเดลที่ไม่มีความจำเป็นในการเก็บข้อมูลเกี่ยวกับคุณลักษณะของนักเรียน/สถานศึกษา

โมเดล 2 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาที่ใช้ผลคะแนนจากข้อสอบทั้งฉบับเช่นเดียวกับโมเดล 1 แต่มีการคำนึงถึงความแตกต่างของคุณลักษณะของนักเรียน/สถานศึกษา และควบคุมตัวแปรอิทธิพลของตัวแปรดังกล่าว เนื่องจากการควบคุมอิทธิพลของคุณลักษณะของนักเรียน/สถานศึกษา จึงทำให้ผลการประเมินถูกปรับให้อยู่บนพื้นฐานของปัจจัย

ดั้งเดิมเท่าเทียมกัน และนำผลต่างหรือส่วนที่เพิ่มขึ้นจากการควบคุมมาเปรียบเทียบกันอย่างยุติธรรม นอกจากนี้เมื่อเปรียบเทียบกับโมเดล 4 จะพบว่าการวิเคราะห์ไม่ซับซ้อนเท่ากับโมเดล 4 แต่ให้ผลการประเมินคุณภาพการจัดการศึกษาสอดคล้องกับโมเดลที่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปก่อนการวิเคราะห์มูลค่าเพิ่ม(โมเดล 4)

โมเดล 3 เป็นโมเดลการประเมินที่ไม่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะนักเรียนและสถานศึกษา เช่นเดียวกับโมเดล 1 แต่มีการตรวจสอบคุณภาพของแบบสอบและตัดข้อสอบที่โอนเอียงกับนักเรียนกลุ่มใดกลุ่มหนึ่งออกไป หรือข้อสอบที่ทำหน้าที่ต่างกันออกไปก่อน ดังนั้นคะแนนที่จากแบบสอบ จึงมีการควบคุมโอกาสในการตอบถูกให้เท่าเทียมกันในคุณลักษณะที่ต่างกัน เมื่อเปรียบเทียบกับโมเดล 1 จะพบว่ามีการวิเคราะห์ซับซ้อนมากกว่าโมเดล 1 แต่ให้ผลการประเมินคุณภาพการจัดการศึกษาที่ค่อนข้างสอดคล้องกับโมเดล 1

โมเดล 4 เป็นโมเดลการประเมินคุณภาพการจัดการศึกษาที่มีการวิเคราะห์ที่ซับซ้อนมากที่สุด เป็นโมเดลที่มีความน่าเชื่อถือที่สุดและมีความธรรมที่สุด เนื่องจากมีการตรวจสอบคุณภาพของแบบสอบและตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป เช่นเดียวกับโมเดล 3 และมีการควบคุมอิทธิพลของคุณลักษณะของนักเรียน/สถานศึกษา เช่นเดียวกับโมเดล 2 จึงทำให้ผลการประเมินที่ได้จึงมีการควบคุมโอกาสการตอบข้อสอบถูกของกลุ่มนักเรียนที่ต่างกัน และควบคุมอิทธิพลของคุณลักษณะของนักเรียน/สถานศึกษาทำให้ผลการประเมินมีความเท่าเทียม และยุติธรรมสำหรับกลุ่มนักเรียนมากที่สุด แต่เมื่อเปรียบเทียบกับโมเดล 2 พบว่ามีใช้เทคนิคการวิเคราะห์ขั้นสูงและซับซ้อนมากกว่าโมเดล 2 แต่ให้ผลการประเมินคุณภาพการจัดการศึกษาที่สอดคล้องกันกับโมเดล 2

5.2 อภิปรายผลการวิจัย

การวิจัยพัฒนาโมเดลมูลค่าเพิ่มของผลสัมฤทธิ์ทางวิชาการและแบบตรวจสอบรายการประเมินตนเองเพื่อเพิ่มมูลค่ากระบวนการจัดการศึกษา อภิปรายใน 3 ประเด็นสำคัญ ดังนี้

1. คุณภาพของแบบสอบต่อการประเมินคุณภาพการจัดการศึกษา

1.1 คุณภาพของแบบสอบที่ใช้ในการประเมินผลในโครงการ PISA

การวิจัยในครั้งนี้ใช้ผลการประเมินนักเรียนนานาชาติ ในโครงการ PISA ซึ่งไทยเป็นประเทศหนึ่งที่เข้าร่วมโครงการนี้จากทั้งหมด 57 ประเทศ โดยมีจุดมุ่งหมายของโครงการเพื่อให้ข้อมูลกับประเทศสมาชิกเตรียมความพร้อมของประชาชนให้มีศักยภาพสำหรับการแข่งขันในประชาคมโลกและเพื่อหาตัวชี้วัดคุณภาพของการจัดการศึกษา สิ่งสำคัญที่ควรคำนึงถึงในการนำผลการประเมินนักเรียนมาใช้คือคุณภาพของเครื่องมือหรือแบบสอบที่ใช้ในการประเมิน ทั้งนี้แบบสอบซึ่งประกอบด้วยข้อสอบที่ถามให้ผู้สอบได้แสดงสิ่งที่รู้ตามกรอบของการวัด ซึ่งเป็นไปได้ว่าอาจจะมีบางส่วนที่ไม่มีความตรงต่อคุณลักษณะของกลุ่มผู้สอบ เช่นคำศัพท์ที่ใช้ในแต่ละข้อสอบ อาจจะมีผลให้ผู้สอบแต่ละกลุ่มแปลความหมายในข้อสอบได้แตกต่างกัน (Camilli & Shepard,1994) นั่นคือเป็นผลของการทำหน้าที่ต่างกันของข้อสอบ ซึ่งหมายความว่าผู้สอบที่มีคุณลักษณะต่างกันและมี

ความสามารถเท่ากัน มีโอกาสในการทำข้อสอบถูกได้ต่างกัน ซึ่งส่งผลต่อความตรงของการประเมินในครั้งนั้น

จากผลการตรวจสอบการทำหน้าที่ต่างกันของสอบในแบบสอบที่ใช้ในโครงการ PISA ซึ่งประกอบด้วยแบบสอบ 13 ฉบับ ในภาพรวมพบว่าแบบสอบทั้ง 13 ฉบับ มีข้อสอบที่ทำหน้าที่ต่างกันค่อนข้างมาก ร้อยละ 11.54 ถึง 67.85 โดยเฉลี่ยคิดเป็นร้อยละ 41.81 และเมื่อพิจารณาทิศทางการโอนเอียงเข้าหากลุ่มผู้สอบกลุ่มใดกลุ่มหนึ่ง พบว่าส่วนใหญ่จะโอนเอียงให้เพศชายมากกว่าเพศหญิง โอนเอียงให้นักเรียนที่เรียนพิเศษมากกว่านักเรียนที่ไม่เรียนพิเศษ โอนเอียงให้นักเรียนที่มีเศรษฐานะของครอบครัวสูงกว่านักเรียนที่มีเศรษฐานะของครอบครัวต่ำ โอนเอียงให้นักเรียนที่มีทรัพยากรของครอบครัวต่ำมากกว่านักเรียนที่มีทรพยากรของครอบครัวสูง และโอนเอียงให้นักเรียนที่มีความมั่งคั่งของครอบครัวสูงมากกว่านักเรียนที่มีความมั่งคั่งของครัวต่ำ นอกจากนี้ผู้วิจัยได้วิเคราะห์ลักษณะของข้อสอบที่ทำหน้าที่ต่างกันในด้านของเป้าหมายของการถาม สมรรถนะทางวิทยาศาสตร์ และความรู้วิทยาศาสตร์ ผลพบว่าในด้านเป้าหมายของการถาม ทั้งที่เป็นการถามในระดับบุคคล ระดับสังคม และระดับโลก ผลที่คล้ายกันคือเป็นข้อสอบจะโอนเอียงให้เพศชาย ครอบครัวที่มีเศรษฐานะสูง ครอบครัวที่มีทรัพยากรต่ำและครอบครัวที่มีความมั่งคั่งสูง ในด้านสมรรถนะทางวิทยาศาสตร์พบว่า ในด้านการระบุประเด็นทางวิทยาศาสตร์ จะโอนเอียงให้เพศชาย นักเรียนที่เรียนพิเศษ ครอบครัวที่มีเศรษฐานะสูง ครอบครัวที่มีทรัพยากรต่ำ ด้านการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ ส่วนใหญ่จะโอนเอียงให้นักเรียนเพศหญิง นักเรียนที่ไม่เรียนพิเศษ ครอบครัวที่มีเศรษฐานะสูง ครอบครัวที่มีทรัพยากรต่ำ และมีความมั่งคั่งของครอบครัวสูง การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์จะโอนเอียงให้นักเรียนเพศชาย นักเรียนที่เรียนพิเศษ ครอบครัวที่มีเศรษฐานะสูง ครอบครัวที่มีทรัพยากรต่ำ และมีความมั่งคั่งของครอบครัวสูง การใช้ประจักษ์พยานทางวิทยาศาสตร์ จะโอนเอียงให้นักเรียนเพศชาย นักเรียนที่เรียนพิเศษ ครอบครัวที่มีเศรษฐานะสูง ครอบครัวที่มีทรัพยากรต่ำ และมีความมั่งคั่งของครอบครัวสูง ในด้านความรู้วิทยาศาสตร์ พบว่า ความรู้วิทยาศาสตร์ในด้านเนื้อหาจะโอนเอียงให้นักเรียนเพศชาย นักเรียนที่เรียนพิเศษ ครอบครัวที่มีเศรษฐานะสูง ครอบครัวที่มีทรัพยากรต่ำ และมีความมั่งคั่งของครอบครัวสูง และด้านความรู้ทางวิทยาศาสตร์จะโอนเอียงให้นักเรียนเพศชาย นักเรียนที่เรียนพิเศษ ครอบครัวที่มีเศรษฐานะสูง และครอบครัวที่มีทรัพยากรต่ำ

จากผลการค้นพบการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ มีประเด็นที่น่าสนใจเกี่ยวกับคุณภาพของแบบสอบในโครงการประเมินผล PISA ปี 2006 ดังนี้

ประการแรก *ข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบพบจำนวนค่อนข้างมาก* ในแบบสอบโครงการประเมินผล PISA ปี 2006 แต่ละฉบับ โดยเฉลี่ยร้อยละ 41.81 เมื่อเปรียบเทียบกับผลการวิจัยของ Le (2006) ที่วิเคราะห์หาข้อสอบที่ทำหน้าที่ต่างกันโดยใช้ข้อมูลจาก PISA รอบทดลองใช้ ปี 2006 ในมิติที่ต่างกัน เมื่อวิเคราะห์ในมิติของสมรรถนะทางวิทยาศาสตร์ พบข้อสอบที่ทำหน้าที่ต่างกันที่โอนเอียงให้เพศชายร้อยละ 13.5 และโอนเอียงให้กับเพศหญิงร้อยละ 9.5 และเมื่อวิเคราะห์ในมิติของเป้าหมายของคำถาม พบข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบที่โอนเอียงให้เพศ

ชายร้อยละ 18.1 และโอนเอียงให้กับเพศหญิงร้อยละ 8.2 และจากผลการวิจัยของ Grisay and Monseur (2007) ที่วิเคราะห์ความเท่าเทียมกันของความยากของข้อสอบและวิเคราะห์หาข้อสอบที่ทำหน้าที่ต่างกันในภาษาที่ต่างกันในแบบสอบตามโครงการประเมินผล PISA วิชาการอ่านในปี 2000 พบว่าภาษาในแบบสอบที่ต่างกัน ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันจำนวนมาก ร้อยละ 10 ถึง 23 โดยเฉลี่ยคิดเป็นร้อยละ 17 จากผลการวิจัยในครั้งนี้และนักวิจัยทั้งสองท่าน แสดงให้เห็นว่าแบบสอบที่ใช้ในโครงการ PISA พบข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบค่อนข้างมาก ทั้งนี้อาจจะมีสาเหตุมาจากลักษณะของกลุ่มตัวอย่างที่ใช้ในการวิจัยที่ได้จากการสุ่มนักเรียนที่มีความหลากหลายครอบคลุมและเป็นตัวแทนของประชากรของแต่ละประเทศ แต่ถ้าพิจารณาร้อยละของการตรวจพบข้อสอบที่ทำหน้าที่ต่างกันในการวิจัยในครั้งนี้ จะพบว่าตรวจพบข้อสอบที่ทำหน้าที่ต่างกันค่อนข้างมากกว่านักวิจัยทั้งสองท่าน ทั้งนี้วิธีการวิเคราะห์แตกต่างกัน ในการวิจัยครั้งนี้ ผู้วิจัยเลือกการประยุกต์ใช้โมเดลเชิงเส้นตรงทั่วไปแบบลดหลั่น หรือ HGLM ซึ่งวิธีการวิเคราะห์ของนักวิจัยทั้งสองท่านใช้วิธี IRTDIF ดังนั้นโอกาสที่จะตรวจพบข้อสอบที่ทำหน้าที่ต่างกันจึงต่างกัน ซึ่ง Mapuranga, Dorans & Middleton (2008) ได้จัดประเภทของวิธีการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันและได้เสนอว่าวิธีที่ต่างกันมีลักษณะการวิเคราะห์ที่ต่างกันจึงตรวจพบได้ต่างกัน นอกจากนี้ในวิจัยครั้งนี้ยังใช้ตัวแปรของคุณลักษณะของนักเรียนจำนวนมากถึง 5 แปร คือ เพศ การเรียนพิเศษของนักเรียน เศรษฐานะของครอบครัว แหล่งทรัพยากรที่บ้าน และความมั่งคั่งของครอบครัว ในขณะที่การวิจัยอื่นจะพิจารณาตัวแปรระดับนักเรียนเพียงตัวแปรเดียวคือตัวแปรเพศ จึงทำให้การวิจัยครั้งนี้ตรวจพบข้อสอบที่ทำหน้าต่างกันจำนวนมาก

ประการที่สอง *ลักษณะของข้อสอบที่ทำหน้าที่ต่างกันที่โอนเอียงเข้าหากลุ่มผู้สอบ* จากผลการตรวจสอบพบว่าข้อสอบที่ทำหน้าที่ต่างกันส่วนใหญ่โอนเอียงให้เพศชายมากกว่าเพศหญิง ซึ่งสอดคล้องกับ Le (2006) พบว่าส่วนใหญ่ข้อสอบจะโอนเอียงให้เพศชายมากกว่าเพศหญิง ไม่ว่าจะวิเคราะห์มิติในด้านของเป้าหมายของคำถาม เนื้อหาในข้อสอบและสมรรถนะทางวิทยาศาสตร์ ทั้งนี้เป็นไปได้ว่านักเรียนเพศชายจะมีโอกาสในการทำข้อสอบได้ถูกมากกว่านักเรียนเพศหญิงในลักษณะของข้อสอบที่ไม่เกี่ยวข้องกับเนื้อหาในหลักสูตร แต่เป็นการใช้ความรู้เพื่อแก้ปัญหาในชีวิตประจำวัน และจากงานวิจัยของ Hanan & Hamzeh (2006) ได้วิเคราะห์เนื้อหาในข้อสอบที่ทำหน้าที่ต่างกัน พบว่าเนื้อหาที่โอนเอียงให้กับนักเรียนเพศชายเป็นเนื้อหาเกี่ยวกับการหาความสัมพันธ์หรือพีชคณิต ส่วนเนื้อหาที่โอนเอียงให้กับนักเรียนหญิงเป็นเนื้อหาเกี่ยวกับการวิเคราะห์ข้อมูลที่ไม่เป็นลักษณะของการคาดการณ์คำตอบ แต่ต้องเป็นคำตอบที่แน่นอนเพียงหนึ่งข้อ ซึ่งลักษณะของข้อสอบในโครงการ PISA โดยส่วนใหญ่เป็นข้อสอบที่ต้องอาศัยการหาความสัมพันธ์ของข้อมูลในสถานการณ์ของข้อสอบ และตัวเลือกมีโอกาสถูกมากกว่าหนึ่งข้อ

นอกจากนี้ผลการวิจัยในครั้งนี้ยังพบว่าข้อสอบที่ทำหน้าที่ต่างกันส่วนใหญ่โอนเอียงให้กับนักเรียนที่เรียนพิเศษ มีเศรษฐานะของครอบครัวสูง และมีความมั่งคั่งของครอบครัวสูง ทั้งนี้เนื่องจากนักเรียนกลุ่มดังกล่าวมีโอกาสที่จะเรียนรู้ในเนื้อหาหรือเหตุการณ์ที่หลากหลายในชีวิตประจำวันมากกว่านักเรียนที่ไม่ได้เรียนพิเศษ และนอกจากนี้นักเรียนที่มีเศรษฐานะของ

ครอบครัวสูง และมีความมั่งคั่งของครอบครัวสูงมีโอกาสที่จะได้รับการเรียนรู้จากผู้ปกครองที่มีรายได้และการศึกษาในระดับสูง และนักเรียนที่อยู่ในครอบครัวที่ความมั่งคั่งสูงยังความเป็นอยู่ที่ดี มีความสะดวกสบายต่อการการเรียนรู้จากแหล่งค้นคว้าจากอินเตอร์เน็ต หรือมีความเป็นอยู่ที่สะดวกสะบายกว่านักเรียนในอีกกลุ่มหนึ่ง ผลที่น่าสังเกตจากการวิจัยครั้งนี้พบว่าข้อสอบที่ทำหน้าที่ต่างกันโอนเอียงให้กับนักเรียนที่มีแหล่งทรัพยากรที่บ้านต่ำ ทั้งนี้มีความเป็นไปได้ว่าผลการสำรวจแหล่งทรัพยากรที่บ้านเป็นเพียงการสำรวจของการมีอยู่ของทรัพยากร เช่น จำนวนหนังสือ พจนานุกรม เครื่องคิดเลข เครื่องคอมพิวเตอร์ แสดงว่าข้อมูลของการมีแหล่งทรัพยากรที่บ้าน ไม่ได้บ่งบอกว่านักเรียนได้ใช้ทรัพยากรเหล่านั้นเพื่อการเรียนรู้ หรือเป็นไปได้ว่าแหล่งทรัพยากรเหล่านั้นไม่ตรงกับความจำเป็นต่อการเรียนรู้ของนักเรียน อย่างไรก็ตามผลการศึกษาของ Amery & Ercikan (2006) ได้ศึกษาแหล่งของวัฒนธรรมที่ทำให้ข้อสอบทำหน้าที่ต่างกันในโครงการ TIMSS ของประเทศไต้หวันและสหรัฐอเมริกา พบว่าการทำหน้าที่ต่างกันของสอบไม่มีลักษณะเฉพาะของแต่ละข้อสอบ ปัจจัยทางวัฒนธรรมที่เป็นแหล่งของการเกิดการทำหน้าที่ต่างกันของข้อสอบในสภาพแวดล้อมหรือบริบทที่มีลักษณะเฉพาะ ซึ่งอาจจะความแตกต่างกันในแต่ละประเทศ

1.2 ผลของข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบวิชาวิทยาศาสตร์ต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน

จากการเปรียบเทียบผลการประเมินคุณภาพการจัดการศึกษาระหว่างโมเดล 1 กับโมเดล 3 ซึ่งเป็นโมเดลการประเมินคุณภาพการจัดการศึกษาที่ไม่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา แต่โมเดล 1 ไม่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปและโมเดล 3 มีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป ผลการประเมินคุณภาพการจัดการศึกษาของทั้งสองโมเดล พบว่าจัดกลุ่มคุณภาพการจัดการศึกษาได้สอดคล้องกันหรือจัดอยู่สถานศึกษาในกลุ่มเดียวกันได้จำนวนมาก (98.2%) และเมื่อพิจารณาระดับคุณภาพการจัดการศึกษาของสถานศึกษา พบว่าผลการประเมินคุณภาพการจัดการศึกษาของโมเดล 1 และโมเดล 3 ทำให้ระดับคุณภาพของสถานศึกษามีตำแหน่งตรงกัน 133 แห่ง (62.74%) ส่วนผลการประเมินคุณภาพการจัดการศึกษาระหว่างโมเดล 2 และโมเดล 4 ซึ่งเป็นโมเดลการประเมินคุณภาพการจัดการศึกษาที่มีการควบคุมอิทธิพลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษา พบว่าผลจัดกลุ่มคุณภาพการจัดการศึกษาได้สอดคล้องทุกแห่ง (100.0%) เมื่อพิจารณาระดับคุณภาพของสถานศึกษาพบว่าผลการประเมินคุณภาพการจัดการศึกษาของโมเดล 2 และโมเดล 4 ทำให้ระดับคุณภาพของสถานศึกษามีตำแหน่งตรงกัน 110 แห่ง (51.89%) จากผลทดสอบทางสถิติสรุปว่า การจัดกลุ่มคุณภาพและระดับคุณภาพการจัดการศึกษาของสถานศึกษาระหว่างการประเมินด้วยโมเดล 1 และ 3 ทำให้ผลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาไม่ต่างกัน นั่นคือข้อสอบที่ทำหน้าที่ต่างกันในแบบสอบวิชาวิทยาศาสตร์ไม่มีผลต่อการประเมินคุณภาพการจัดการศึกษา

ประเด็นที่น่าสนใจคือ ในการวิจัยครั้งนี้ การจัดกลุ่มคุณภาพการจัดการศึกษาจากโมเดลการประเมินคุณภาพการจัดการศึกษาที่มีการตัดและไม่ตัดข้อสอบที่ทำหน้าที่ต่างกันออก พบว่ามีความ

สอดคล้องกันถึงแม้จะมีระดับคุณภาพของสถานศึกษาต่างกัน แต่อย่างไรก็ตาม การมีข้อสอบทำหน้าที่ต่างกันในแบบสอบ ส่งผลต่อการได้เปรียบเสียเปรียบของนักเรียนกลุ่มใดกลุ่มหนึ่ง เนื่องจากการทำหน้าที่ต่างกันของข้อสอบเป็นผลที่เกิดขึ้นเมื่อนักเรียนที่มีคุณลักษณะย่อยต่างกันและมีความสามารถเท่ากัน แต่มีโอกาสในการตอบข้อสอบได้ถูกต่างกัน ซึ่งในการวิจัยครั้งนี้ ผลของข้อสอบที่ทำหน้าที่ต่างกัน เกิดจากคุณลักษณะของนักเรียนที่แตกต่างกันใน 5 ตัวแปร คือเพศ การเรียนพิเศษ เศรษฐานะของครอบครัว แหล่งทรัพยากรที่บ้าน และความมั่งคั่งของครอบครัว ดังนั้นการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปจึงเป็นการควบคุมหรือลดแหล่งของความคลาดเคลื่อนที่เกิดจากคุณลักษณะของนักเรียนที่ต่างกัน ที่ส่งผลต่อโอกาสในการตอบข้อสอบถูก และยังเป็นการขจัดข้อสอบที่ทำให้เกิดปัญหาความยุติธรรมระหว่างกลุ่มข้อสอบกลุ่มต่าง ๆ ที่มีคุณลักษณะต่างกัน เพื่อให้ผลการสอบสามารถนำไปใช้ได้อย่างถูกต้องยิ่งขึ้น ซึ่งในการวิจัยครั้งนี้ ผู้วิจัยได้ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันจำนวนค่อนข้างมาก แต่ไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันที่ตรวจพบทั้งหมดออกไป แต่เลือกตัดข้อสอบที่ทำหน้าที่ต่างกันออกไปเป็นบางข้อ โดยพิจารณาเลือกตัดข้อสอบตามเกณฑ์ 3 ข้อ ดังนี้ 1) เป็นข้อสอบที่ตรวจพบว่าทำหน้าที่ต่างกันเนื่องจากตัวแปรคุณลักษณะนักเรียนมากกว่าหนึ่งคุณลักษณะ 2) เลือกข้อสอบที่ทำหน้าที่ต่างกันที่มีค่าสัมประสิทธิ์สูงสุด และ 3) คำนึงถึงโครงสร้างของแบบสอบ โดยภายหลังการตัดข้อสอบออกแล้วโครงสร้างของแบบสอบจะไม่แตกต่างจากโครงสร้างเดิม และข้อสอบที่ตัดออกไปควรมีจำนวนเป็นร้อยละ 20 ของข้อสอบในแบบสอบนั้น ทั้งนี้ผู้วิจัยคำนึงถึงคุณภาพของแบบสอบในด้านความตรง (Validity) ของแบบสอบที่ต้องวัดในกรอบโครงสร้างเดิม และคุณภาพด้านความเที่ยง (Reliability) เนื่องจากการตัดข้อสอบในแบบสอบส่งผลต่อความยาวของแบบสอบ และทำให้ความเที่ยงมีค่าลดลง ซึ่งโดยทั่วไปคะแนนจากแบบสอบที่มีความยาวจะมีความน่าเชื่อมากกว่าคะแนนจากแบบสอบที่สั้นกว่า (ศิริชัย กาญจนวาสี, 2548) นอกจากนี้จากผลการวิจัยของ รักชนก ยี่สุ่นศรี (2544) พบว่าคุณภาพของแบบสอบด้านความเที่ยงของแบบสอบในฉบับก่อนและหลังการตัดข้อสอบที่ทำหน้าที่ต่างกัน มีค่าความเที่ยงของแบบสอบแตกต่างกัน โดยถ้าตัดข้อสอบออกจากแบบสอบจำนวนมากขึ้นจะทำให้ค่าความเที่ยงของแบบสอบมีค่าลดลง ดังนั้นผู้วิจัยจึงไม่ได้ตัดข้อสอบที่ทำหน้าที่ต่างกันที่ตรวจพบทั้งหมดออกไป เพราะต้องคำนึงคุณภาพของแบบสอบในด้านความเที่ยงและความตรงของแบบสอบด้วย

2. อิทธิพลของคุณลักษณะของนักเรียนและสถานศึกษาต่อการประเมินคุณภาพการจัดการศึกษา

2.1 ตัวแปรที่ส่งผลต่อคะแนนมูลค่าเพิ่มวิชาวิทยาศาสตร์

จากผลการวิเคราะห์มูลค่าเพิ่มของคะแนนผลการประเมินวิชาวิทยาศาสตร์ ที่นำตัวแปรเข้ามาศึกษาในโมเดล 2 และโมเดล 4 ตัวแปร**ระดับนักเรียน**มีที่นำเข้าศึกษา 13 ตัวแปร พบว่ามีตัวแปรที่มีอิทธิพลต่อคะแนนผลการประเมินวิชาวิทยาศาสตร์ 10 ตัวแปร ตัวแปรที่มีอิทธิพลทางบวกเรียงลำดับจากค่าสัมประสิทธิ์สูงไปต่ำได้แก่ ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา ความเพลิดเพลินทางวิทยาศาสตร์ ระดับการใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง การเห็น

คุณค่าโดยทั่วไปในวิทยาศาสตร์ เศรษฐานะของครอบครัว แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์ ความสนใจในการเรียนรู้วิทยาศาสตร์ แหล่งทรัพยากรที่บ้าน และตัวแปรที่มีอิทธิพลทางลบเรียงลำดับจากค่าสัมประสิทธิ์สูงไปต่ำ ได้แก่ ความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์ ความมั่งคั่งของครอบครัว ส่วนตัวแปรที่ไม่มีผลต่อคะแนนผลการประเมินวิชาวิทยาศาสตร์ ได้แก่ ความเป็นเพศชาย ระดับการใช้เวลาเรียนพิเศษวิทยาศาสตร์นอกสถานศึกษา และความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์ ส่วนตัวแปร**ระดับสถานศึกษา**มีตัวแปรที่นำเข้ามาศึกษามี 22 ตัวแปร พบว่ามีตัวแปรที่มีอิทธิพลต่อคะแนนผลการประเมินวิชาวิทยาศาสตร์ 13 ตัวแปร ตัวแปรที่มีอิทธิพลทางบวกเรียงลำดับจากค่าสัมประสิทธิ์สูงไปต่ำได้แก่ สถานศึกษาสังกัด สกอ. สถานศึกษาขนาดใหญ่ สถานศึกษาขนาดกลาง ดัชนีทรัพยากรการเรียน ดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล และตัวแปรที่มีอิทธิพลทางลบเรียงลำดับจากค่าสัมประสิทธิ์สูงไปต่ำ ได้แก่ สถานศึกษาที่ตั้งอยู่ในเมืองเล็ก สถานศึกษาที่ตั้งอยู่ในหมู่บ้าน สถานศึกษาสังกัด สช. สถานศึกษาสังกัด กศท. และ กทม. สถานศึกษาที่ตั้งอยู่ในเมือง ระดับแรงกดดันของผู้ปกครองต่อสถานศึกษา ร้อยละงบประมาณของรัฐที่จัดสรรให้สถานศึกษา สัดส่วนนักเรียนต่อครู ส่วนตัวแปรคุณลักษณะระดับสถานศึกษาที่ไม่มีผลต่อคะแนนผลการประเมินวิชาวิทยาศาสตร์ มีจำนวน 9 ตัวแปร ได้แก่ สถานศึกษาสังกัด สพฐ. สถานศึกษาที่ตั้งอยู่ในเมืองใหญ่ ขนาดห้องเรียน ระดับการขาดแคลนครูวิทยาศาสตร์ ระดับการขาดแคลนครูที่มีคุณวุฒิ ดัชนีการขาดแคลนครู ดัชนีการจัดกิจกรรมเพื่อการเรียนรู้วิทยาศาสตร์ ดัชนีความรับผิดชอบเกี่ยวกับการจัดสรรทรัพยากร และร้อยละของงบประมาณที่สถานศึกษาจัดสรรจากแหล่งอื่น

จากตัวแปรที่มีอิทธิพลต่อคะแนนผลการประเมินวิชาวิทยาศาสตร์ มีตัวแปรที่น่าสนใจหลายลักษณะที่ต่างกัน ดังนี้

ประการที่หนึ่ง *ลักษณะของตัวแปรระดับนักเรียนที่มีอิทธิพลทางบวก* ได้แก่ ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา ความเพลิดเพลินทางวิทยาศาสตร์ ระดับการใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง การเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์ เศรษฐานะของครอบครัว แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์ ความสนใจในการเรียนรู้วิทยาศาสตร์ แหล่งทรัพยากรที่บ้าน ซึ่งมีนักการศึกษาหลายท่านได้ศึกษาตัวแปรระดับที่ส่งผลต่อผลการเรียนรู้ของนักเรียน เช่น Teddlie, Stringfield & Reynolds (2000) สรุปว่าผลที่เกิดจากตัวแปรที่เป็นภูมิหลังของนักเรียน เศรษฐานะทางสังคม Chiu & Xihua (2008) พบว่าปัจจัยในด้านครอบครัวก็ส่งผลต่อผลสัมฤทธิ์ของนักเรียน ได้แก่ จำนวนสมาชิกของครอบครัวที่ส่งผลต่อทรัพยากรของนักเรียนหรือความเพียงพอของแหล่งทรัพยากร การมีแรงจูงใจทางการเรียน Riddell (2008) พบว่าเศรษฐานะมีผลต่อผลสัมฤทธิ์ทางวิชาการ และจากการศึกษาของ Palardy (2008) ที่สรุปผลการวิจัยของ Coleman et al., 1966; Lee & Bryk, 1989; Lee & Smith, 1993, 1995; Lee et al., 1997; McNeal, 1997; Park & Palardy, 2004 ว่าความหลากหลายของคุณลักษณะของนักเรียนมีความสัมพันธ์กับผลสัมฤทธิ์ทางการเรียน ได้แก่ คุณลักษณะทางกายภาพ เช่น เชื้อชาติ ศาสนา เพศ คุณลักษณะ

พื้นฐานของครอบครัว เช่น เศรษฐานะ (SES) และโครงสร้างของครอบครัว และพื้นฐานการศึกษา เช่นผลสัมฤทธิ์เดิม หรือพื้นฐานความรู้เดิมของนักเรียน Koutsoulis และ Campbell (2001) ที่ผลพบว่า ตัวทำนายผลสัมฤทธิ์ทางการเรียนของนักเรียนระดับมัธยมศึกษาที่ดีที่สุด คือ ความสามารถเดิม (prior ability) ของนักเรียน เศรษฐานะทางสังคม Willms and Somers (2001) : Tomasz and Dhawan-Biswal (2008) พบว่า เศรษฐานะทางสังคมส่งผลทางบวกต่อผลสัมฤทธิ์ทางการเรียน และนักวิจัยภายในประเทศไทย ได้แก่ บุญเรือง ศรีเหรัญ (2542) พบว่าปัจจัยระดับบุคคลได้แก่ ความรู้เดิมของนักเรียน บรรยากาศและสภาพแวดล้อมภายในครอบครัวส่งผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์และวิทยาศาสตร์ เจตคติและแรงจูงใจต่อการเรียน นิตยา เหมือดไธสง (2543) พบว่าปัจจัยที่ส่งผลทางตรงเชิงบวกในระดับสูง ได้แก่ ภูมิหลังของนักเรียน สุบิน ยุระรัช (2547) พบว่าแรงจูงใจใฝ่สัมฤทธิ์ของนักเรียนส่งผลต่อคุณภาพนักเรียน และ พิชิต ธรรมรักษ์ (2549) พบว่าปัจจัยที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียน ได้แก่ ความรู้พื้นฐานเดิม อาชีพของผู้ปกครอง เจตคติต่อวิชาคณิตศาสตร์

ถ้าจัดจำแนกประเภทของตัวแปรที่มีอิทธิพลในการวิจัยครั้งนี้ เป็นตัวแปรภายในซึ่งเป็นสิ่งที่นักเรียนสามารถทำให้เกิดขึ้นได้มากขึ้นหรือพัฒนาได้มากขึ้น ได้แก่ ผลสัมฤทธิ์ทางการเรียนที่ผ่านมา ความเพลิดเพลินทางวิทยาศาสตร์ ระดับการใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง การเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์ แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์ และความสนใจในการเรียนรู้วิทยาศาสตร์ ตัวแปรเหล่านี้ที่ส่งผลทางบวกต่อผลสัมฤทธิ์ทางการเรียนสามารถอธิบายได้ด้วยทฤษฎีสติปัญญาทางสังคม และแรงจูงใจทางการเรียนของนักเรียนของ Bandura (1989) ที่เชื่อว่าการเห็นคุณค่า (Value) ความคาดหวัง(expectancy) และความชื่นชอบ (affect) ของนักเรียนที่เห็นจะทำให้เกิดลักษณะของความพยายามมากกว่าและมีโอกาสที่จะได้รับความสำเร็จสูงขึ้น และตัวแปรอีกประเภทหนึ่งคือตัวแปรภายนอกที่อยู่นอกเหนือการควบคุมได้ของนักเรียนและตัวแปรภายในที่นักเรียนสามารถพัฒนาให้เกิดขึ้นมากขึ้นได้ จะเห็นได้ว่าตัวแปรภายนอก ได้แก่ เศรษฐานะของครอบครัว แหล่งทรัพยากรที่บ้าน จัดเป็นตัวแปรที่มีความสัมพันธ์ หรือสนับสนุนให้เกิดตัวแปรภายในนักเรียนได้มากขึ้น ตัวอย่างเช่นในครอบครัวที่มีเศรษฐานะสูงและมีแหล่งทรัพยากรทางการเรียนมากกว่า ย่อมสามารถส่งเสริมให้นักเรียนเกิดการเรียนรู้ได้มากขัน พ่อแม่ที่มีอาชีพการงานที่ดี มีเวลาในการช่วยเหลือและแนะนำการเรียนให้กับนักเรียน มีความพร้อมทางด้านการเงินที่จะสนับสนุนการเรียน ตลอดจนการจัดหาแหล่งทรัพยากรภายในบ้านที่สนับสนุนการเรียนรู้ของนักเรียน ส่งผลให้นักเรียนมีเวลาในการศึกษาด้วยตนเองมากขึ้น เกิดแรงจูงในในการเรียน และมีความสนใจในการเรียนมากขึ้น และส่งผลให้นักเรียนเกิดผลสัมฤทธิ์ทางการเรียนสูงขึ้น ในทางตรงกันข้าม ครอบครัวที่ไม่มีความพร้อมทางด้านการเงิน ผู้ปกครองมีฐานะทางการเงินไม่ดี มีอาชีพการงานที่ส่งผลให้ไม่มีเวลาในการดูแลนักเรียน ตลอดจนระดับการศึกษาที่ไม่สามารถให้การแนะนำการเรียนแก่นักเรียนได้ ย่อมส่งผลให้นักเรียนขาดการดูแลเอาใจใส่อย่างใกล้ชิด จึงอาจจะมีแรงจูงใจทางการเรียน และส่งผลให้ผลการเรียนต่ำลง

ประการที่สอง *ลักษณะของตัวแปรระดับนักเรียนที่มีอิทธิพลทางลบ*ได้แก่ ความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์ และความมั่งคั่งของครอบครัว ในขณะที่ความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์ไม่มีผลต่อคะแนนผลการประเมินวิชาวิทยาศาสตร์ จากแนวคิดของทฤษฎีของ Bandura (1989) ได้จัดแบ่งความเชื่อมั่นในตนเองออกเป็น ความเชื่อมั่นในประสิทธิภาพของตนเองในการเรียนวิทยาศาสตร์(Self – efficacy) และความเชื่อในความสามารถของตนเองในการเรียนวิทยาศาสตร์ (Self – concept) ซึ่งจัดเป็นความสามารถทางการศึกษาที่มีผลกับประสบการณ์ของนักเรียน นั่นคือถ้านักเรียนได้รับประสบการณ์ที่ดี ย่อมส่งให้นักเรียนเกิดความเชื่อมั่นมากขึ้น แต่อย่างไรก็ตาม โดยส่วนใหญ่ พบว่านักเรียนจะแสดงความคิดเห็นต่อการเรียนวิทยาศาสตร์ในทางบวกที่ค่อนข้างสูง และประเมินตนเองในภาพรวมว่าตนเองมีความสามารถที่จะเรียนรู้วิทยาศาสตร์ได้ จึงทำให้ขาดการระมัดระวัง และการเตรียมความพร้อมในการเรียนรู้วิทยาศาสตร์ในเนื้อหาต่างๆ ส่วนตัวแปรความมั่งคั่งของครอบครัวที่มีอิทธิพลทางลบต่อคะแนนผลการประเมิน อาจจะเป็นข้อมูลส่วนหนึ่งที่ชี้ให้เห็นว่านักเรียนที่มาจากครอบครัวที่มีฐานะการเป็นอยู่ที่สะดวกสบาย ไม่จำเป็นว่าจะได้เปรียบกว่านักเรียนที่มีฐานะการเป็นอยู่ที่ยากแค้นกว่า แต่ในทางตรงข้าม การมีชีวิตความเป็นอยู่ที่สะดวกสบายกลับทำให้นักเรียนใช้เวลาอยู่กับปัจจัยภายนอกที่ไม่เกี่ยวกับการเรียนรู้ ทำให้เวลาในการอ่านหนังสือ ทบทวนตำราน้อยลง แต่นักเรียนที่อยู่ในครอบครัวที่ไม่มีอุปกรณ์ที่อำนวยความสะดวกสบาย กลับมีผลให้นักเรียนเกิดความพยามยามที่จะเอาชนะความยากลำบากมากขึ้น

นอกจากนี้สิ่งที่น่าสนใจในตัวแปรที่พบว่าไม่มีอิทธิพลต่อคะแนนผลการประเมินวิชาวิทยาศาสตร์คือระดับการใช้เวลาเรียนพิเศษวิทยาศาสตร์นอกสถานศึกษา แสดงให้เห็นว่าการเรียนพิเศษวิทยาศาสตร์ไม่มีผลต่อคะแนนมูลค่าเพิ่มของสถานศึกษา จึงน่าจะเป็นข้อมูลหนึ่งที่ชี้ให้เห็นได้ว่าการเรียนพิเศษ หรือการกวดวิชาไม่ได้ทำให้สถานศึกษามีคุณภาพการจัดการศึกษาที่สูงขึ้น ถึงแม้ว่าในปัจจุบัน นักเรียนจะนิยมเรียนกวดวิชาเพื่อทำให้ตนเองมีความสามารถมากขึ้น และประสบความสำเร็จในการสอบคัดเลือกต่างๆ สาเหตุที่ทำให้การเรียนพิเศษของนักเรียนไม่ได้ทำให้นักเรียนทำคะแนนผลการประเมินในแบบสอบ PISA สูงขึ้น เนื่องจากลักษณะของแบบสอบเป็นการใช้ความรู้วิทยาศาสตร์มาใช้ในการแก้ปัญหา ไม่ใช่การนำความรู้ในเนื้อหาตามหลักสูตรที่สอนในสถานศึกษาหรือที่สอนในสถานที่บริการกวดวิชา

ประการที่สาม *ตัวแปรระดับสถานศึกษาที่ส่งผลต่อคะแนนผลการประเมินวิชาวิทยาศาสตร์* พบว่าตัวแปรที่มีอิทธิพลทางบวกได้แก่ สถานศึกษาสังกัด สกอ. สถานศึกษาขนาดใหญ่ สถานศึกษาขนาดกลาง ดัชนีทรัพยากรการเรียน ดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล นั่นคือเป็นตัวแปรที่ส่งผลให้นักเรียนที่อยู่ในสถานศึกษาลักษณะดังกล่าวมีผลคะแนนการประเมินสูงกว่าสถานศึกษาในลักษณะอื่น ซึ่งนักการศึกษาและนักวิจัยหลายท่านให้ข้อมูลที่สอดคล้องกัน เช่น Palardy (2008) ที่พบว่าปัจจัยนำเข้าของสถานศึกษาได้แก่แหล่งทรัพยากรของสถานศึกษา ลักษณะโครงสร้าง นโยบายและการปฏิบัติของโรงเรียนมีผลต่อคุณภาพการจัดการของสถานศึกษา และจากการศึกษาทบทวนงานวิจัยของ Heck (2000) ซึ่งศึกษาเกี่ยวกับอิทธิพลของ

สถานศึกษาใน 20 ปีที่ผ่านมาสรุปได้ว่าตัวแปรที่มีความสำคัญเป็นเงื่อนไขบริบทของสถานศึกษา ประกอบด้วยโครงสร้าง นโยบาย บุคคลและกระบวนการ เช่น ขนาดห้องเรียน คุณภาพและเงินเดือนของครู ข้อมูลพื้นฐาน ภูมิประเทศ ตัวแปรบริบทจึงเป็นสิ่งที่ช่วยอำนวยความสะดวกให้เกิดผลสำเร็จได้ดีขึ้น Ho and Willms (1996 cited in Willms and Somers, 2001) พบว่า การมีส่วนร่วมที่ดีของผู้ปกครอง พบว่า มีอิทธิพลค่อนข้างสูงเมื่อพิจารณาในระดับสถานศึกษา สำหรับตัวแปรอื่นในระดับสถานศึกษา ได้แก่ จำนวนวัสดุอุปกรณ์การสอนที่ส่งผลทางบวกอย่างมีนัยสำคัญต่อวิชาภาษา และจำนวนนักเรียนต่อห้องพบว่ามีผลทางลบ D'Agostino (2000) พบว่าขนาดสถานศึกษา ความเป็นผู้นำทางวิชาการ ระดับความยากจนของสถานศึกษา การสนับสนุนของผู้ปกครอง ทำให้ครูจัดการเรียนการสอนได้อย่างมีประสิทธิภาพ และนักเรียนมีผลสัมฤทธิ์ทางการเรียนสูงขึ้น ส่วนนักการศึกษาภายในประเทศ บุญเรือง ศรีเหรัญ (2542) พบว่าขนาดสถานศึกษา ค่าใช้จ่ายทางการศึกษา ประสิทธิภาพการจัดการเรียนการสอนและคุณภาพของการใช้หลักสูตรส่งผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์และวิทยาศาสตร์ เจตคติและแรงจูงใจต่อการเรียน ปิยะธิดา ทองอร่าม (2545) พบว่าปัจจัยระดับสถานศึกษาที่มีอิทธิพลต่อคุณภาพของสถานศึกษาอย่างมีนัยสำคัญทางสถิติ ได้แก่ สัดส่วนของนักเรียนต่อครู สัดส่วนนักเรียนต่อห้องเรียน ขวัญและกำลังใจในการปฏิบัติงานและความเป็นผู้นำของผู้บริหารสถานศึกษา พิชิต ธรรมรักษ์ (2549) พบว่าปัจจัยที่สำคัญที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียนในระดับสถานศึกษา ได้แก่ ขนาดของสถานศึกษา ความเป็นผู้นำทางวิชาการของผู้บริหาร เป็นต้น

ในขณะที่ตัวแปรที่มีอิทธิพลทางลบได้แก่ สถานศึกษาที่ตั้งอยู่ในเมืองเล็ก สถานศึกษาที่ตั้งอยู่ในหมู่บ้าน สถานศึกษาสังกัด สช. สังกัด กศท. และ กทม. สถานศึกษาที่ตั้งอยู่ในเมือง ระดับแรงกดดันของผู้ปกครองต่อสถานศึกษา ร้อยละงบประมาณของรัฐที่จัดสรรให้สถานศึกษา สัดส่วนนักเรียนต่อครู นั่นคือสถานศึกษาที่มีคุณลักษณะการตั้งอยู่ในหมู่บ้านหรือเป็นเมืองที่ไม่ใหญ่ การอยู่ในสังกัดของเอกชน และกศท.และกทม. และสัดส่วนของนักเรียนต่อครูสูงจะทำให้ผลคะแนนการประเมินต่ำกว่าสถานศึกษาในลักษณะอื่น ทั้งนี้สาเหตุของตำแหน่งที่ตั้งของสถานศึกษาที่ตั้งในหมู่บ้านหรือเมืองไม่ใหญ่ จะมีปัจจัยที่สนับสนุนการจัดการศึกษาขอสถานศึกษาได้น้อย ซึ่งสอดคล้องกับผลการศึกษาของสุภา อินทร์มณี (2542) ส่วนระดับแรงกดดันของผู้ปกครองต่อสถานศึกษาที่มีผลทางลบแปลหมายความได้ว่าสถานศึกษาที่ไม่มีแรงกดดันของผู้ปกครอง จะทำให้คะแนนผลการประเมินต่ำกว่าสถานศึกษาที่มีผู้ปกครองมีแรงกดดันต่อสถานศึกษา นั่นแสดงให้เห็นว่า การที่ผู้ปกครองเข้ามามีบทบาทต่อการปฏิบัติงานของสถานศึกษา ส่งผลดีต่อคุณภาพการจัดการศึกษา โดยลักษณะของการเข้ามามีบทบาทของผู้ปกครอง อาจจะเข้ามาในรูปแบบของคณะกรรมการสถานศึกษา การประชุมพบปะผู้ปกครอง การเข้ามาสนับสนุนทางด้านการเงิน ตลอดจนการเข้ามามีบทบาทในการตัดสินใจเกี่ยวกับหลักสูตรวิชาที่เรียนในสถานศึกษา นอกจากนี้ตัวแปรที่น่าสังเกตอีกตัวแปรหนึ่งคือร้อยละงบประงบประมาณที่รัฐจัดสรรให้แก่สถานศึกษา ซึ่งพบว่ามีอิทธิพลทางลบ นั่นแสดงว่าการเพิ่มงบประมาณทางการศึกษา ไม่ได้เป็นหลักประกันว่าคุณภาพการจัดการศึกษาจะสูงขึ้น มีข้อมูลจากประเทศสมาชิก OECD หลายประเทศ ที่พบว่าคุณภาพการจัดการศึกษาเพื่อทำ

ให้นักเรียนเกิดการเรียนรู้มากขึ้น ไม่ได้เพิ่มขึ้นตามงบประมาณค่าใช้จ่ายทางการศึกษา เช่น สหรัฐอเมริกา และลักเซมเบิร์ก ที่เสียค่าใช้จ่ายทางการศึกษาสูงมาก แต่ผลการจัดการเรียนรู้ให้กับนักเรียนไม่ได้เพิ่มขึ้นตามค่าใช้จ่ายทางการศึกษา (สสวท, 2552)

ส่วนตัวแปรคุณลักษณะระดับสถานศึกษาที่ไม่มีผลต่อคะแนนผลการประเมินวิชาวิทยาศาสตร์ ได้แก่ตัวแปรการขาดแคลนครูวิทยาศาสตร์ ระดับการขาดแคลนครูที่มีคุณวุฒิ ดัชนีการขาดแคลนครู จึงกล่าวได้ว่าตัวแปรดังกล่าวไม่มีผลต่อคะแนนมูลค่าเพิ่มที่แสดงถึงคุณภาพการจัดการศึกษาของสถานศึกษา ทั้งๆที่ประเทศไทยประสบปัญหาการขาดแคลนครูวิทยาศาสตร์จำนวนมาก แต่กลับพบว่าตัวแปรดังกล่าวไม่มีผลต่อคะแนนการประเมินวิชาวิทยาศาสตร์ ทั้งนี้อาจจะเป็นไปได้ว่าสถานศึกษาดังกล่าวสามารถชดเชยสภาวะที่ขาดแคลนครูด้วยการจัดหาแหล่งทรัพยากรทางการเรียนมากขึ้น และผู้ปกครองเข้ามามีบทบาทในการดูแล กำกับการปฏิบัติงาน ทำให้การบริการการจัดการเรียนรู้มีความเข้มงวดกับนักเรียนมากขึ้น แต่อย่างไรก็ตามสิ่งที่ควรนำมาพิจารณาอีกประการหนึ่งคือผลการสำรวจข้อมูลเกี่ยวกับระดับการขาดแคลนครูวิทยาศาสตร์ และระดับการขาดแคลนครูที่มีคุณวุฒิ ซึ่งผู้บริหารของสถานศึกษาเป็นผู้ให้ข้อมูล ซึ่งประกอบด้วยครูประจำการและครูอัตราจ้าง ดังนั้นระดับการขาดแคลนครูจึงเป็นการแสดงข้อมูลของครูที่มีอยู่ในขณะนั้น นอกจากนี้ตัวแปรหนึ่งที่พบว่าไม่มีอิทธิพลคือดัชนีการจัดกิจกรรมการเรียนรู้วิทยาศาสตร์ ทั้งนี้เป็นไปได้ ลักษณะของการจัดกิจกรรมการเรียนรู้ส่วนใหญ่ไม่ได้ส่งผลต่อการเรียนรู้วิทยาศาสตร์ เช่น การจัดทัศนศึกษา เป็นต้น จากการสำรวจข้อมูลในประเทศอื่นพบว่าดัชนีการจัดกิจกรรมการเรียนรู้ไม่ได้เป็นตัวแปรที่บ่งบอกได้ว่าผลการเรียนรู้วิทยาศาสตร์สูงขึ้น เช่น ดัชนีการจัดกิจกรรมการเรียนรู้ของฟินแลนด์และญี่ปุ่น เท่ากับ -0.6 และ -1.16 ตามลำดับ ทั้งๆที่ประเทศดังกล่าวมีระดับผลการประเมินอยู่ใน 5 อันดับแรกของประเทศสมาชิก OECD แต่ในขณะที่ดัชนีการจัดกิจกรรมการเรียนรู้ของประเทศไทยสูงถึง 1.35 แต่กลับมีผลการประเมินที่ต่ำกว่าค่าเฉลี่ยของ OECD

ประการที่สี่ *ลักษณะความสัมพันธ์ของตัวแปรข้ามระดับ* โดยตัวแปรในระดับสูงกว่าหรือระดับสถานศึกษา ซึ่งทำหน้าที่กำหนดนโยบาย แนวทางการปฏิบัติงาน นั่นคือตัวแปรระดับสถานศึกษาเป็นการสร้างเงื่อนไขหรือสภาวะกระตุ้นถ่ายโยงไปสู่การปฏิบัติในตัวแปรระดับล่างหรือระดับนักเรียน เพื่อให้บรรลุผลตามเป้าหมาย เรียกความสัมพันธ์ลักษณะนี้ว่าความสัมพันธ์เชิงถ่ายโยง (ศิริชัย กาญจนวาสี, 2550)

ตัวอย่างความสัมพันธ์ในลักษณะนี้เช่น จากผลการวิจัยในครั้งนี้ที่พบตัวแปรที่มีอิทธิพลในระดับสถานศึกษา คือการเป็นสถานศึกษาขนาดใหญ่ ดัชนีทรัพยากรการเรียน และตัวแปรที่มีอิทธิพลในระดับนักเรียนคือความเพลิดเพลินทางวิทยาศาสตร์ ระดับเวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง แรงจูงใจภายนอกในการเรียนวิทยาศาศาสตร์ เป็นต้น แสดงให้เห็นความสัมพันธ์ที่ถ่ายโยงกันได้ว่า การที่สถานศึกษามีขนาดใหญ่ หรือมีดัชนีทรัพยากรการเรียนสูง ย่อมมีบรรยากาศหรือสิ่งแวดล้อมที่ดีหรือเอื้อต่อการเรียนรู้ภายในสถานศึกษา ส่งผลให้เกิดสภาวะกระตุ้นต่อนักเรียน เนื่องจากการจัดสิ่งแวดล้อมที่ดีภายในสถานศึกษาอาจจะช่วยให้นักเรียนมีความเพลิดเพลินในการเรียน มีสถานที่ในการนั่งศึกษาความรู้ด้วยตนเอง ทำให้มีเวลาในการเรียนรู้มากขึ้นและส่งผล

นักเรียนมีผลสัมฤทธิ์ทางการเรียนสูงขึ้น หรือในด้านตรงข้าม ตัวแปรที่มีอิทธิพลทางลบ คือการที่สถานศึกษาตั้งอยู่ในหมู่บ้าน เป็นไปได้ว่าชุมชนโดยส่วนใหญ่เป็นครอบครัวหรือผู้ปกครองที่มีเศรษฐานะไม่ค่อยดี ส่งผลให้เกิดปัญหาในการสนับสนุนการศึกษาให้กับนักเรียน ทำให้ผลสัมฤทธิ์ทางการเรียนของนักเรียนไม่ดีตามไปด้วย

เช่นเดียวกันกับผลการศึกษาของ Creemers (cited in Maslowski, et al., 2008) ซึ่งแสดงให้เห็นความสัมพันธ์เกี่ยวโยงกันระหว่างตัวแปรใน 4 ขั้น คือบริบท สถานศึกษา ห้องเรียน และนักเรียน Opdenakker and Van Damme (2000) ที่ศึกษาอิทธิพลจากสถานศึกษา ที่มีต่อผลสัมฤทธิ์ทางการเรียน พบว่าปัจจัยด้านบรรยากาศของชั้นเรียนและสถานศึกษาส่งผลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียน D'Agostino (2000) ได้ศึกษาอิทธิพลของการเรียนการสอนและสถานศึกษาที่มีต่อผลสัมฤทธิ์ทางการเรียน พบว่าสถานศึกษาที่ผู้ปกครองให้การการสนับสนุนทำให้ครูจัดการเรียนการสอนได้อย่างมีประสิทธิภาพ และทำให้นักเรียนมีผลสัมฤทธิ์ทางการเรียนสูงขึ้น

2.2 ผลของตัวแปรคุณลักษณะของนักเรียนและสถานศึกษาขั้นพื้นฐานต่อการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน

จากการเปรียบเทียบผลการประเมินคุณภาพการจัดการศึกษาระหว่างโมเดล 1 กับโมเดล 2 ซึ่งเป็นโมเดลที่ไม่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ หรือไม่มีการควบคุมแหล่งความคลาดเคลื่อนที่ทำให้นักเรียนต่างกันกลุ่มมีโอกาสตอบข้อสอบถูกต่างกัน ผลการประเมินคุณภาพการจัดการศึกษาระหว่างโมเดล 1 กับโมเดล 2 พบว่าจัดกลุ่มคุณภาพการจัดการศึกษาได้สอดคล้องกันหรือจัดอยู่สถานศึกษาในกลุ่มคุณภาพเดียวกันคิดเป็นร้อยละ 61.9 เมื่อพิจารณาอันดับของสถานศึกษา พบว่าผลการประเมินคุณภาพการจัดการศึกษาของโมเดล 1 มีอันดับของสถานศึกษาตรงกันกับผลของโมเดล 2 จำนวน 2 แห่ง (0.94%) ส่วนผลการประเมินคุณภาพการจัดการศึกษาระหว่างโมเดล 3 กับโมเดล 4 ซึ่งมีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป พบว่าจัดกลุ่มคุณภาพการจัดการศึกษาได้สอดคล้องกัน คิดเป็นร้อยละ 62.3 เมื่อพิจารณาอันดับของสถานศึกษา พบว่าสถานศึกษาที่มีอันดับของคะแนนมูลค่าเพิ่มจากผลการประเมินคุณภาพการจัดการศึกษาในโมเดล 3 ตรงกันกับอันดับของสถานศึกษาจากผลของโมเดล 4 จำนวน 2 แห่ง (0.94%) เช่นเดียวกัน จากผลการวิเคราะห์จึงสรุปได้ว่า คุณลักษณะของตัวแปรระดับนักเรียนและสถานศึกษามีผลต่อการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา เนื่องจากผลการเปรียบเทียบระหว่างผลการวิเคราะห์ในโมเดลที่ไม่มีและมีการควบคุมอิทธิพลของตัวแปรระดับนักเรียนและสถานศึกษา พบว่าจัดกลุ่มคุณภาพการจัดการศึกษาต่างกันคิดเป็น 38.1% (โมเดล 1 กับโมเดล 2) และ 37.7% (โมเดล 3 กับโมเดล 4) และมีระดับที่ของสถานศึกษาต่างกันหรือเปลี่ยนตำแหน่งไปคิดเป็น 99.06% ซึ่งจัดได้ว่าเปลี่ยนอันดับไปจำนวนมาก

สาเหตุที่ทำให้ผลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา ระหว่างโมเดลที่มีและไม่มีการนำตัวแปรระดับนักเรียนและสถานศึกษาเข้ามาวิเคราะห์มูลค่าเพิ่ม เนื่องจากผลการตรวจสอบตัวแปรที่มีอิทธิพลต่อการเปลี่ยนแปลงผลคะแนนการประเมินวิชาวิทยาศาสตร์ พบตัวแปร

จำนวนมากทั้งในระดับนักเรียนและสถานศึกษา ไม่ว่าจะเป็นผลสัมฤทธิ์ทางการเรียนที่ผ่านมา ความเพลิดเพลินทางวิทยาศาสตร์ ระดับการใช้เวลาในการศึกษาวิทยาศาสตร์ด้วยตนเอง การเห็นคุณค่าโดยทั่วไปในวิทยาศาสตร์ แรงจูงใจภายนอกในการเรียนวิทยาศาสตร์ และความสนใจในการเรียนรู้วิทยาศาสตร์ ซึ่งเป็นตัวแปรระดับนักเรียน และตัวแปรระดับสถานศึกษาได้แก่ สถานศึกษาสังกัด สกอ. สถานศึกษาขนาดใหญ่ สถานศึกษาขนาดกลาง ดัชนีทรัพยากรการเรียน ดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล และแรงกดดันของผู้ปกครอง ดังนั้นเมื่อมีการควบคุมอิทธิพลของตัวแปรเหล่านี้ตามหลักการวิเคราะห์มูลค่าเพิ่ม จึงทำให้ผลการประเมินคุณภาพการจัดการศึกษาในแต่ละโมเดลการประเมินคุณภาพการจัดการศึกษาแตกต่างกัน

นักการศึกษาหลายท่านที่สนใจเปรียบเทียบผลการวิเคราะห์ของโมเดลที่ต่างกัน ซึ่งมีที่สอดคล้องกับงานวิจัยในครั้งนี้ ตัวอย่างเช่น การศึกษาของ TCMLRAFR (Tekwe, Carter, Ma, Algina, Lucas, Roth, Ariet, Fisher and Resnick, 2004) ได้ศึกษาโมเดลมูลค่าเพิ่มที่แตกต่างกัน 4 โมเดล คือ 1) โมเดลอิทธิพลคงที่ (Fixed-effects models: FEM) 2) โมเดลอิทธิพลผสมผสานแบบเป็นชั้น (layered mixed effects model: LMEM) 3) โมเดลเชิงเส้นแบบลดหลั่น (Hierarchical linear models: HLMM) และ4) โมเดลที่มีการปรับแก้ (adjusted HLMM: AHLMM) ผลการศึกษาพบว่าให้ผลการวิเคราะห์ที่แตกต่างใน 4 โมเดล และชี้ให้เห็นว่าความแตกต่างระหว่าง AHLMM ซึ่งเป็นโมเดลที่มีลักษณะคล้ายกันกับโมเดล 4 ในการวิจัยครั้งนี้ เนื่องจากมีการนำตัวแปรแทรกซ้อนระดับนักเรียนและตัวแปรองค์ประกอบใช้วิเคราะห์ในโมเดลด้วย การประมาณค่าจะเกิดการเปลี่ยนแปลง ถ้าตัวแปรแทรกซ้อนมีผลกระทบต่อการเรียนรู้ของนักเรียน และนักการศึกษาอีกท่านคือ 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

3. แนวทางการเลือกโมเดลการประเมินคุณภาพการจัดการศึกษา

3.1 สรุปลักษณะและผลของการใช้โมเดลการประเมินคุณภาพการจัดการศึกษา

จากการวิจัยในครั้งนี้ ได้ศึกษาโมเดลการประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษาขั้นพื้นฐาน ในโมเดลที่แตกต่างกัน โดยกำหนดประเด็นที่นำมาพิจารณาสองประเด็น คือ คุณภาพของแบบสอบและโมเดลการวิเคราะห์มูลค่าเพิ่ม โมเดลการประเมินคุณภาพการจัดการศึกษาทั้ง 4 โมเดล สรุปได้ดังนี้

ตารางที่ 45 ผลการเปรียบเทียบโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษา 4 โมเดล

ประเด็น	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
การควบคุมแหล่งความคลาดเคลื่อน	ไม่มีการควบคุมแหล่งความคลาดเคลื่อน	ควบคุมอิทธิพลตัวแปร	ควบคุมแหล่งของการทำหน้าที่ต่างกันของข้อสอบ	ควบคุมทั้งอิทธิพลของตัวแปรและการทำหน้าที่ต่างกันของข้อสอบ
ข้อมูลที่ใช้ในการวิเคราะห์	ใช้ข้อมูล 1 ชุด คือ 1. คะแนนการสอบ	ใช้ข้อมูล 3 ชุด คือ 1. คะแนนการสอบ 2. ตัวแปรนักเรียน 3. ตัวแปรสถานศึกษา	ใช้ข้อมูล 2 ชุด คือ 1. ผลการตอบข้อสอบแต่ละข้อ 2. คะแนนผลการสอบ	ใช้ข้อมูล 4 ชุด คือ 1. ผลการตอบข้อสอบแต่ละข้อ 2. คะแนนการสอบ 3. ตัวแปรนักเรียน 4. ตัวแปรสถานศึกษา
ขั้นตอนหลักในการวิเคราะห์	มี 2 ขั้นตอนหลัก คือ 1. เตรียมไฟล์ข้อมูลสำหรับวิเคราะห์สองไฟล์คือไฟล์นักเรียนและไฟล์สถานศึกษา 2. วิเคราะห์มูลค่าเพิ่มโดยไม่มีการนำตัวแปรเข้ามาในสมการวิเคราะห์	มี 2 ขั้นตอนหลัก คือ 1. เตรียมไฟล์ข้อมูลสำหรับวิเคราะห์สองไฟล์คือไฟล์นักเรียนและไฟล์สถานศึกษา 2. วิเคราะห์มูลค่าเพิ่มโดยนำตัวแปรเข้ามาในสมการวิเคราะห์	มี 3 ขั้นตอนหลัก คือ 1. เตรียมไฟล์ข้อมูลสำหรับวิเคราะห์สองไฟล์คือไฟล์นักเรียนและไฟล์สถานศึกษา 2. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ 3. วิเคราะห์มูลค่าเพิ่มโดยไม่มีการนำตัวแปรเข้ามาในสมการวิเคราะห์	มี 3 ขั้นตอนหลัก คือ 1. เตรียมไฟล์ข้อมูลสำหรับวิเคราะห์สองไฟล์คือไฟล์นักเรียนและไฟล์สถานศึกษา 2. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ 3. วิเคราะห์มูลค่าเพิ่มโดยนำตัวแปรเข้ามาในสมการวิเคราะห์
ความสอดคล้องของผลการประเมิน	<u>กลุ่มคุณภาพ</u> - ตรงกับโมเดล 2 (61.9%) - ตรงกับโมเดล 3 (98.2%) <u>อันดับคุณภาพ</u> - ตรงกับโมเดล 2 (0.94%) - ตรงกับโมเดล 3 (62.74%)	<u>กลุ่มคุณภาพ</u> - ตรงกับโมเดล 1 (61.9%) - ตรงกับโมเดล 4 (100.0%) <u>อันดับคุณภาพ</u> - ตรงกับโมเดล 1 (0.94%) - ตรงกับโมเดล 4 (51.89%)	<u>กลุ่มคุณภาพ</u> - ตรงกับโมเดล 1 (98.2%) - ตรงกับโมเดล 4 (62.3%) <u>อันดับคุณภาพ</u> - ตรงกับโมเดล 1 (62.74%) - ตรงกับโมเดล 4 (0.94%)	<u>กลุ่มคุณภาพ</u> - ตรงกับโมเดล 2 (100.0%) - ตรงกับโมเดล 3 (62.3%) <u>อันดับคุณภาพ</u> - ตรงกับโมเดล 2 (51.89%) - ตรงกับโมเดล 3 (0.94%)
ค่าสหสัมพันธ์ของอันดับคะแนน	<u>ค่าสหสัมพันธ์ (r)</u> - $r_{M\ 1\&\ 2}$ = 0.547 - $r_{M\ 1\&\ 3}$ = 0.981	<u>ค่าสหสัมพันธ์ (r)</u> - $r_{M\ 2\&\ 1}$ = 0.547 - $r_{M\ 2\&\ 4}$ = 1.000	<u>ค่าสหสัมพันธ์ (r)</u> - $r_{M\ 3\&\ 1}$ = 0.981 - $r_{M\ 3\ \&\ 4}$ = 0.538	<u>ค่าสหสัมพันธ์ (r)</u> - $r_{M\ 4\&\ 2}$ = 1.000 - $r_{M\ 4\&\ 3}$ = 0.538
สัมประสิทธิ์การทำนาย (R^2)	-	77.75%	-	88.63%

การเปรียบเทียบโมเดลประเมินคุณภาพการจัดการศึกษาจากตารางข้างต้น แสดงให้เห็นว่า โมเดลที่มีวิธีการวิเคราะห์ต่างกัน จะให้ผลการประเมินคุณภาพการจัดการศึกษาต่างกัน ถ้าจะพิจารณาสัมประสิทธิ์การทำนายของโมเดลหรือประสิทธิภาพในการทำนายคะแนนผลการประเมินวิทยาศาสตร์พบว่าโมเดล 2 ซึ่งเป็นโมเดลที่ไม่มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ นั่นคือไม่มีการควบคุมแหล่งความคลาดเคลื่อนที่เกิดขึ้นจากการโอนเอียงให้ผู้สอบกลุ่มใดกลุ่มหนึ่ง พบว่าโมเดล 2 มีสัมประสิทธิ์การทำนายร้อยละ 77.75 ในขณะที่การประเมินคุณภาพการจัดการศึกษาด้วยโมเดล 4 ซึ่งเป็นโมเดลที่มีขั้นตอนการวิเคราะห์เพิ่มขึ้นกว่าโมเดล 2 คือมีการวิเคราะห์การทำหน้าที่ต่างกันข้อสอบ พบว่าโมเดล 4 มีสัมประสิทธิ์การทำนายร้อยละ 88.63 ชี้ให้เห็นว่า ถ้ามีการตัดข้อสอบที่ทำหน้าที่ต่างกันออกไป นั่นคือมีการควบคุมแหล่งความคลาดเคลื่อนอันเกิดจากการที่ข้อสอบโอนเอียงเข้าหาผู้สอบกลุ่มใดกลุ่มหนึ่ง จะทำให้โมเดลการประเมินคุณภาพการจัดการศึกษามีค่าสัมประสิทธิ์การทำนายเพิ่มมากขึ้น แต่อย่างไรก็ตาม ผลการจัดกลุ่มคุณภาพการจัดการศึกษาของสถานศึกษาของทั้งสองโมเดลจะได้ผลที่สอดคล้องตรงกันทั้งหมด และการจัดอันดับของสถานศึกษาได้ตรงกันประมาณร้อยละ 50 ซึ่งเป็นการเปลี่ยนอันดับไม่มากนักเนื่องจากการจัดกลุ่มคุณภาพการจัดการศึกษาสามารถจัดได้ตรงกัน ดังนั้นจึงน่าจะแสดงให้เห็นได้ว่าการประเมินคุณภาพการจัดการศึกษาด้วยโมเดล 2 และ 4 ให้ผลการประเมินคุณภาพการจัดการศึกษาที่สอดคล้องกัน และสามารถเลือกใช้โมเดลใดโมเดลหนึ่งได้

เมื่อพิจารณาผลการประเมินคุณภาพการจัดการศึกษาของโมเดล 3 ซึ่งแตกต่างจากโมเดล 4 คือไม่มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ พบว่าโมเดล 3 ให้ผลการจัดกลุ่มคุณภาพการจัดการศึกษาที่ตรงกันกับโมเดล 4 โดยประมาณร้อยละ 60 จึงแสดงให้เห็นได้ว่าโมเดล 3 และ 4 ให้ผลการประเมินคุณภาพการจัดการศึกษาที่ค่อนข้างแตกต่างกัน และเมื่อเปรียบเทียบผลการประเมินคุณภาพการจัดการศึกษาของโมเดล 3 กับโมเดล 1 ซึ่งเป็นโมเดลที่ไม่มีการควบคุมอิทธิพลของตัวแปรเช่นเดียวกัน พบว่าผลการประเมินคุณภาพการจัดการศึกษาให้ผลการจัดกลุ่มที่ตรงกันเกือบทั้งหมด และให้ผลการจัดระดับของสถานศึกษาได้ตรงกันประมาณร้อยละ 60 ดังนั้นถ้าจะพิจารณาเลือกโมเดลใดโมเดลหนึ่งก็สามารถให้ผลการจัดกลุ่มคุณภาพการจัดการศึกษาที่ไม่แตกต่างกันมากนัก

3.2 การเลือกโมเดลการประเมินคุณภาพการจัดการศึกษา

การประเมินคุณภาพการจัดการศึกษาในประเทศไทยยังต้องอาศัยคะแนนจากการทำแบบสอบของนักเรียนซึ่งเป็นตัวชี้บ่งถึงระดับการเรียนรู้ของนักเรียน และสะท้อนถึงผลการปฏิบัติงานภายในสถานศึกษา ดังนั้นจึงต้องเลือกโมเดลการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาเพื่อให้ผลการประเมินมีความถูกต้อง น่าเชื่อถือมากที่สุด

จากการวิจัยในครั้งนี้ ผู้วิจัยใช้โมเดลการประเมินคุณภาพการจัดการศึกษาที่ต่างกัน 4 โมเดล ซึ่งพบว่าโมเดล 4 เป็นโมเดลที่น่าเชื่อถือและมีความเป็นธรรมมากที่สุด เนื่องจากมีการควบคุมแหล่งความคลาดเคลื่อนจากการทำหน้าที่ต่างกันของข้อสอบและมีการควบคุมอิทธิพลของคุณลักษณะนักเรียนและสถานศึกษา และเมื่อเปรียบเทียบผลการประเมินคุณภาพการจัดการศึกษา

กับโมเดลอื่น พบว่าให้ผลการประเมินที่ต่างกัน ทั้งนี้การเลือกใช้โมเดลได้มีนักการศึกษาที่ศึกษาโมเดลมูลค่าเพิ่มได้ให้ข้อเสนอแนะหลายประการ ตัวอย่างเช่นระบบการประเมินมูลค่าเพิ่มที่รู้จักกันดีคือ ระบบการประเมินมูลค่าเพิ่มของรัฐเทนเนสซี (Tennessee Value Added Assessment: TVAAS) และระบบการประเมินมูลค่าเพิ่มทางการศึกษา (Education Value-Added Assessment System: EVAAS) โดย Ballou, Sanders and Wright. (2004) ได้เสนอแนะเกี่ยวกับการเลือกโมเดลมูลค่าเพิ่มว่า ความแตกต่างของโมเดลมูลค่าเพิ่มแต่ละโมเดล ขึ้นอยู่กับการกำหนดลักษณะเฉพาะของโมเดล ทั้งนี้โมเดลที่เลือกแต่ละโมเดลขึ้นอยู่กับความสามารถในการเข้าถึงข้อมูล การลงความเห็นจากข้อค้นพบ ตลอดจนระดับการปฏิบัติ ที่หมายถึงการเข้าถึงโปรแกรมการวิเคราะห์ที่มีความซับซ้อน โดย Lissitz et al. (2006) ได้เสนอว่าความซับซ้อนของโมเดล สามารถพิจารณาจากตัวแปรแทรกซ้อนที่มีผลต่อตัวแปรตาม (covariates) โดยโมเดลที่การวิเคราะห์ที่ค่อนข้างง่ายคือเป็นโมเดลที่มีตัวแปรเดียว (univariate) ไปจนถึงโมเดลที่มีความซับซ้อนที่มีหลายตัวแปร (multivariate) เนื่องจากข้อจำกัดของการเลือกโมเดลที่มีความเหมาะสม จึงมีนักการศึกษาเช่น Haegeland and Kirkeboen (2008) ที่จัดประเภทของโมเดลมูลค่าเพิ่มที่ต่างกันสองรูปแบบ คือ โมเดลมูลค่าเพิ่ม (Value added models: VAM) กับโมเดลมูลค่าเพิ่มเชิงบริบท (Contextualised value-added modeling: CVA) โดยโมเดล CVA จะนำตัวแปรบริบทที่เป็นตัวแปรเศรษฐานะเข้าร่วมวิเคราะห์ในโมเดลด้วย แต่โมเดล VAM จะพิจารณาเฉพาะตัวแปรความรู้เดิมของนักเรียน

สรุปได้ว่ามีประเด็นหลายอย่างในการพิจารณาโมเดล ประเด็นแรกเป็นความซับซ้อนในการคำนวณ ประเด็นที่สองคือการให้ความสำคัญของตัวแปรแทรกซ้อนที่อาจจะส่งผลตัวแปรตามร่วมกับตัวแปรต้น ประเด็นที่สามคือตัวแปรที่นำเข้ามาในโมเดล ซึ่งการใช้ตัวแปรที่เป็นตัวบ่งชี้ของการปฏิบัติงานของสถานศึกษาอาจจะมีความแตกต่างกันในแต่ละบริบท และอาจจะมีข้อจำกัดในการเข้าถึงข้อมูลได้ ดังนั้นการเลือกโมเดลมูลค่าเพิ่มเพื่อใช้ในการประเมินคุณภาพการจัดการของสถานศึกษาจึงควรระวังในการเลือกโมเดลที่เหมาะสมเนื่องจากสภาพบริบทที่แตกต่างกัน ผู้วิจัยจึงเสนอสิ่งที่ควรพิจารณาเพื่อประกอบการตัดสินใจเลือกโมเดลประเมินคุณภาพการจัดการศึกษาในประเด็นดังนี้

วัตถุประสงค์ของการประเมิน การกำหนดวัตถุประสงค์เป็นขั้นตอนแรกสำหรับการประเมิน และใช้วัตถุประสงค์เพื่อกำหนดโมเดลการประเมินต่อไป เช่นการกำหนดกรอบการประเมินตัวแปรที่เกี่ยวข้องกับคุณภาพการจัดการศึกษา ซึ่งเป็นสิ่งสำคัญที่จะใช้สำหรับการแปลความหมายของผลการปฏิบัติหรือคุณภาพการจัดการศึกษาของสถานศึกษา การกำหนดวัตถุประสงค์การประเมินจะทำให้ผู้ประเมินสามารถกำหนดโมเดลการประเมินคุณภาพที่เหมาะสม ซึ่งอาจจะไม่มีความจำเป็นที่จะต้องใช้โมเดลที่มีความซับซ้อน

การเลือกตัวแปร การเลือกตัวแปรเพื่อนำมากำหนดโมเดลมูลค่าเพิ่มที่เหมาะสม สามารถพิจารณาได้จากวัตถุประสงค์ของการประเมิน และพิจารณาความสัมพันธ์ของตัวแปรที่ส่งผลต่อคุณภาพการจัดการศึกษา หรือเป็นตัวแปรที่ไม่เท่าเทียมกันในแต่ละสถานศึกษา ซึ่งจำเป็นที่จะต้องนำมาควบคุมอิทธิพลของตัวแปรเหล่านั้น ดังนั้นการเลือกโมเดลขึ้นอยู่กับบริบทของสถานศึกษาที่มีความแตกต่างกัน และผู้ประเมินต้องการควบคุมอิทธิพลเหล่านั้นเพื่อให้เกิดความเท่าเทียมกัน

การเก็บรวบรวมข้อมูล สิ่งที่ต้องพิจารณาคือคุณภาพของข้อมูล จึงอาจจะเป็นข้อจำกัดหนึ่งในการเก็บรวบรวมข้อมูลเพื่อที่จะให้ได้ข้อมูลที่มีคุณภาพ ควรพิจารณาระบบของข้อมูลในสถานศึกษาและสารสนเทศเพื่อใช้รวบรวมข้อมูล ตลอดจนการกำหนดกลุ่มตัวอย่างนักเรียนที่เพียงพอต่อการนำมาใช้วิเคราะห์ซึ่งอาจเกิดการสูญหายข้อมูล (missing data) ที่อาจเกิดการย้ายสถานศึกษาของนักเรียนในช่วงที่มีการประเมิน โมเดลการประเมินคุณภาพการจัดการศึกษาจะมีประสิทธิภาพได้ ส่วนหนึ่งจึงขึ้นอยู่กับคุณภาพของข้อมูลหรือการเข้าถึงข้อมูล ดังนั้นผู้ประเมินจึงควรเลือกวิธีการเก็บข้อมูลได้อย่างมีคุณภาพ เพื่อทำให้ผลการประเมินจากโมเดลการประเมินคุณภาพเป็นที่น่าเชื่อถือและถูกต้อง

ระยะเวลาและผู้เชี่ยวชาญในการวิเคราะห์ สิ่งที่ยังเป็นข้อจำกัดของการประเมินโดยทั่วไปคือความซับซ้อนในขั้นตอนหรือสมการวิเคราะห์ ซึ่งต้องอาศัยระยะเวลา ผู้เชี่ยวชาญเฉพาะ ตลอดจนโปรแกรมการวิเคราะห์ ซึ่งเป็นสิ่งหนึ่งที่ผู้ประเมินสามารถเลือกโมเดลการประเมินที่ไม่ยุ่งยากหรือซับซ้อนมากนัก แต่ยังให้ผลที่มีประสิทธิภาพเพื่อเทียบกับโมเดลการประเมินที่มีความซับซ้อน

ความถูกต้องและเป็นธรรมของผลการประเมิน การทำให้ผลการประเมินมีความถูกต้องน่าเชื่อถือ และเป็นธรรมแก่สถานศึกษาที่รับการประเมินที่ดีที่สุดคือ การควบคุมแหล่งความคลาดเคลื่อนที่อาจจะเกิดขึ้น โดยทั่วไป แหล่งความคลาดเคลื่อนอาจจะเกิดขึ้นได้การได้มาของข้อมูล และความแตกต่างที่ไม่มีความเท่าเทียมกันระหว่างสถานศึกษา ดังนั้นถ้าเลือกโมเดลการประเมินที่มีการควบคุมแหล่งความคลาดเคลื่อน จึงน่าจะทำให้ผลการประเมินมีความถูกต้องและเป็นธรรมมากที่สุด

การใช้ข้อสารสนเทศจากการประเมิน การนำเสนอและการใช้ข้อสารสนเทศของมูลค่าเพิ่ม เป็นการตัดสินและเลือกวิธีการที่เหมาะสมเพื่อแสดงสารสนเทศของมูลค่าเพิ่ม ตลอดจนแนวทางการแปลความหมายของมูลค่าเพิ่ม ซึ่งสามารถสร้างขึ้นได้โดยผู้ที่เกี่ยวข้องในการจัดทำนโยบาย ตลอดจนผู้ที่ต้องใช้สารสนเทศเพื่อนำไปใช้ในการตัดสินคุณภาพของสถานศึกษา เช่น ผู้ปกครองที่ต้องสถานศึกษาที่มีคุณภาพให้กับบุตรหลาน หรือคณะกรรมการสถานศึกษาที่ใช้สารสนเทศมากำหนดแนวทางในการพัฒนาสถานศึกษาต่อไป

5.3 ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้ แบ่งเป็น 2 ส่วน ดังนี้

1.1 ข้อเสนอแนะเชิงนโยบาย

1.1.1 ***ในระดับชาติ*** ควรใช้รูปแบบการประเมินคุณภาพการจัดการศึกษาของสถานศึกษาที่ประยุกต์ใช้โมเดลมูลค่าเพิ่ม โดยหน่วยงานภาครัฐที่เกี่ยวข้อง เช่น สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (สมศ.) สถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.) และสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท.) ควรนำโมเดลประเมินคุณภาพการจัดการศึกษาที่ประยุกต์ใช้โมเดลมูลค่าเพิ่มไปใช้ในการวิเคราะห์ ซึ่งมีหลักการปรับความไม่เท่าเทียมกัน

ของปัจจัยต่างๆในแต่ละสถานศึกษา ให้อยู่เป็นพื้นฐานที่สามารถเปรียบเทียบกันได้ ทั้งนี้เพื่อนำผลที่บ่งบอกถึงมูลค่าเพิ่มที่สถานศึกษา สามารถทำให้เกิดขึ้นจากปัจจัยที่มีอยู่

1.1.2 ***ในระดับสถานศึกษา*** ผลการวิจัยที่บ่งชี้ถึงตัวแปรระดับสถานศึกษาที่ทำให้ผลสัมฤทธิ์ทางการเรียนสูงขึ้น ได้แก่ ดัชนีทรัพยากรการเรียน ดัชนีความรับผิดชอบเกี่ยวกับหลักสูตรและการประเมินผล แรงกดดันของผู้ปกครองต่อสถานศึกษา เป็นสิ่งที่ควรได้รับการพัฒนาหรือสนับสนุนให้เกิดขึ้น นอกจากนี้ตำแหน่งที่ตั้งของสถานศึกษาที่ตั้งในหมู่บ้านหรือเมืองขนาดเล็ก สถานศึกษาสังกัด สช. สังกัด กศท. และ กทม. และสถานศึกษาที่มีสัดส่วนนักเรียนต่อครูสูงมาก ซึ่งล้วนแต่เป็นตัวแปรที่มีอิทธิพลทางลบต่อผลสัมฤทธิ์ทางการเรียน จึงควรพัฒนาหรือใช้ปัจจัยในด้านอื่นๆเพิ่มขึ้นเพื่อชดเชยปัจจัยที่สูญเสียไป เช่น การจัดแหล่งทรัพยากรทางการเรียนอื่นๆ มาทดแทนโอกาสในการศึกษาหาความรู้โดยการใช้อินเตอร์เน็ต หรือการใช้กลยุทธ์เพื่อให้ผู้ปกครองเข้ามามีบทบาทมากขึ้นต่อการปฏิบัติงานของโรงเรียน เพื่อให้ผู้ปกครองสร้างแรงกดดันต่อการปฏิบัติงานของสถานศึกษามากขึ้น ดังนั้นสถานศึกษาควรมีการวางแผนพัฒนาตัวแปรดังกล่าวให้มีคุณภาพสูงยิ่งขึ้น

1.2.3 ***ในระดับชุมชน*** จากผลงานวิจัยที่พบว่า การมีแรงกดดันของผู้ปกครองในระดับมาก จะส่งผลให้ผลสัมฤทธิ์ทางการเรียนของนักเรียนสูงขึ้น ดังนั้นผู้ปกครองหรือคณะกรรมการสถานศึกษาซึ่งเป็นบุคคลที่อยู่ในชุมชน สามารถให้การสนับสนุน ส่งเสริมสถานศึกษาได้โดยไม่จำเป็นต้องใช้เสียใช้จ่ายที่เป็นตัวเงินเสมอไป แต่สามารถกระทำได้ด้วยการเข้าไปมีส่วนร่วมในการปฏิบัติงานของสถานศึกษา ซึ่งเป็นการสร้างสภาวะกระตุ้นให้สถานศึกษาเอาใจใส่ต่อผู้รับบริการหรือผู้ปกครองมากขึ้น ทำให้เกิดกำกับ ติดตามการดำเนินของสถานศึกษา เพื่อเป็นการยกระดับผลสัมฤทธิ์ทางการเรียนของนักเรียนได้ดียิ่งขึ้น นอกจากนี้จากผลการวิจัยที่พบว่าการเรียนพิเศษนอกเวลาเรียนของนักเรียนไม่ได้ส่งเสริมให้นักเรียนมีผลสัมฤทธิ์ทางการเรียนที่เพิ่มขึ้นแต่อย่างใด แต่เวลาที่นักเรียนใช้ในการศึกษาความรู้ด้วยตนเอง เป็นสิ่งที่ช่วยให้ผลสัมฤทธิ์ทางการเรียนสูงขึ้น ดังนั้นผู้ปกครองจึงควรตระหนักถึงการจัดเวลาให้กับบุตรหลาน และปรับเปลี่ยนค่านิยมเกี่ยวกับการจัดให้บุตรการเรียนพิเศษ แต่ควรส่งเสริมหรือกระตุ้นให้นักเรียนใช้เวลาในการศึกษาด้วยตนเองมากขึ้น

อย่างไรก็ตาม ไม่ว่าจะเป็นหน่วยงานในระดับชาติ ระดับสถานศึกษา หรือระดับชุมชนต่างต้องอาศัยความร่วมมือกัน ทั้งนี้เนื่องจากปัจจัยหรือตัวแปรที่มีอิทธิพลทางบวกต่อการจัดการศึกษา ไม่ได้แสดงอิทธิพลได้เพียงตัวแปรเดียว แต่เป็นตัวแปรที่มีความสัมพันธ์เกี่ยวข้องกัน ตัวอย่างเช่น การที่มีครูที่มีคุณภาพมีผลต่อการเพิ่มขึ้นของคะแนน แต่ถ้านักเรียนมีเวลาในการศึกษาความรู้น้อย หรือผู้ปกครองมีค่านิยมเกี่ยวกับการเรียนพิเศษ คุณภาพการจัดการศึกษาย่อมไม่ได้เพิ่มขึ้นแต่อย่างใด ดังนั้นการจัดการศึกษาจึงควรวางแผนและทำให้เกิดความร่วมมือกันอย่างเป็นระบบซับซ้อน

1.2 ข้อเสนอแนะการพัฒนาเครื่องมือวัดผลสัมฤทธิ์ทางการเรียน

จากผลการวิจัยในครั้งนี้ที่พบว่ามีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบในแบบสอบสำหรับการประเมินการรู้เรื่องด้านวิทยาศาสตร์ ตามโครงการประเมิน PISA ซึ่งเป็นการประเมินนักเรียนขนาดกลุ่มใหญ่ (Large scale assessment) แสดงให้เห็นว่ามีโอกาสที่จะเกิดความคลาดเคลื่อนจากการแปลผลคะแนนสอบได้ ดังนั้นผู้พัฒนาเครื่องมือวัดผลสัมฤทธิ์ทางการเรียนจึงควรมีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และปรับปรุงข้อสอบและแบบสอบให้เป็นธรรมแก่ผู้สอบทุกกลุ่ม โดยเฉพาะอย่างยิ่งการประเมินที่มีส่วนได้ส่วนเสียสูง (High-stake Assessment) เพื่อให้ผลการประเมินมีความถูกต้องและน่าเชื่อถือที่สุด

2. ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

2.1 เนื่องจากการวิจัยในครั้งนี้เป็นการใช้ข้อมูลจากฐานข้อมูลทุติยภูมิ และพิจารณาเลือกตัวแปรที่นำมาศึกษาจากตัวแปรที่มีอยู่เพื่อนำมาทำนายความแปรปรวนของคะแนนการประเมินผู้วิจัยจึงเห็นว่าควรมีการศึกษาโมเดลมูลค่าเพิ่มที่นำมาใช้ประเมินคุณภาพการจัดการศึกษาโดยการใช้ตัวแปรที่สำคัญและมีจำนวนน้อยที่สุด แต่เป็นโมเดลที่มีประสิทธิภาพสูงสุดหรือให้ผลที่สอดคล้องกันกับการใช้ตัวแปรจำนวนมาก ทั้งนี้เพื่อประโยนช์ต่อการลดขั้นตอนการเก็บรวบรวมข้อมูล ความซับซ้อนของการวิเคราะห์ และเวลาที่ใช้ในการวิเคราะห์ และใช้เป็นแนวทางในการกำหนดตัวแปรที่ใช้ในการประเมินคุณภาพการจัดการศึกษาต่อไป

2.2 ควรมีการใช้ฐานข้อมูลของ PISA ในปีต่างๆ ทั้งนี้ประเทศไทย ยังขาดการนำข้อมูลในระดับนานาชาติมาวิเคราะห์รอบสอง (Second Analysis) เพื่อให้ได้สารสนเทศเพิ่มเติมต่อการนำไปพัฒนาการจัดการศึกษา โดยทำการวิจัยเพื่อเปรียบเทียบผลในระยะยาวหรือเปรียบเทียบผลภายในระยะเวลาเดียวกัน ทั้งนี้อาจจะศึกษาในรูปแบบของวิชาเดียวกัน แต่ต่างปีกัน หรือปีเดียวกัน แต่ใช้ข้อมูลในหลายวิชา เพื่อให้ได้ข้อสารสนเทศต่อการนำไปพัฒนาเพิ่มเพิ่มคุณภาพการจัดการศึกษาได้มากขึ้น นอกจากนี้ผู้วิจัยเห็นว่า น่าจะมีการศึกษาโมเดลมูลค่าเพิ่มโดยใช้ผลการสอบในระดับชาติ เพื่อใช้เป็นข้อมูลที่เป็นตัวแทนความรู้เดิมของนักเรียน หรือผลสัมฤทธิ์ทางการเรียนตามหลักสูตรของสถานศึกษา ซึ่งปัจจุบันมี 13 ประเทศ ที่มีการศึกษานำร่องโมเดลมูลค่าเพิ่มโดยใช้ผลการสอบภายในประเทศร่วมกับผลการประเมินตามโครงการ PISA

2.3 ควรมีการวิจัยและพัฒนาโมเดลมูลค่าเพิ่ม โดยใช้ข้อมูลระดับนักเรียน ระดับครูและระดับสถานศึกษา เพื่อศึกษาเกี่ยวกับการพัฒนาที่เกิดขึ้นของนักเรียนจากอิทธิพลของครู และประสิทธิผลครู ร่วมกับอิทธิพลของตัวแปรทำนายในระดับสถานศึกษา โดยวางแผนการวิจัยเพื่อเก็บรวบรวมข้อมูลเพิ่มเติมในตัวแปรระดับครูที่อยู่ในสถานศึกษาเดียวกัน ทั้งนี้เนื่องจากระบบการจัดการศึกษาโดยส่วนใหญเกิดการพฤติกรรมและการปฏิบัติหน้าที่การจัดกิจกรรมการเรียนรู้ของครูที่อยู่ในสถานศึกษานั้น

2.4 การวิจัยในครั้งนี้ ยังมีข้อจำกัดในกระบวนการวิจัยเกี่ยวกับข้อมูลระดับนักเรียนที่ไม่สามารถเก็บข้อมูลระดับนักเรียนที่มีการสอบในสองช่วงเวลาด้วยแบบสอบมาตรฐานได้ แต่ผู้วิจัยได้

ใช้ตัวแปรที่มีอยู่ในฐานข้อมูล และคำนวณหาค่าเพื่อเป็นตัวแปรผลสัมฤทธิ์ทางการเรียนที่ผ่านมาของนักเรียน ดังนั้นจึงควรมีการพัฒนาระบบฐานข้อมูลสำหรับการวิจัยทางด้านโมเดลมูลค่าเพิ่ม เช่น ฐานข้อมูลผลสัมฤทธิ์ทางวิชาการของผู้เรียนที่มีการเก็บข้อมูลระยะยาว จากการทดสอบที่เป็นมาตรฐาน เพื่อนำมาใช้ในการวิเคราะห์พัฒนาการความก้าวหน้าของผลการเรียนรู้ของนักเรียนต่อไป

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รายการอ้างอิง

ภาษาไทย

คณะกรรมการศึกษาแห่งชาติ, สำนักงาน. (2545). **รายงานฉบับสังเคราะห์จากวิกฤติสู่โอกาส สิ่งที่ยังท้าทายการปฏิรูปการศึกษาของไทย**. กรุงเทพ. อรุณการพิมพ์.

คณะกรรมการศึกษาแห่งชาติ, สำนักงาน. (2549). **สมรรถนะการศึกษาไทยในเวทีสากล พ.ศ. 2549**. กรุงเทพ. พริกหวานกราฟฟิค.

ชูเวช ชาญสง่าเวช และสิงหา เจียมศิริ. (2546). การวัดมูลค่าเพิ่มจากการดำเนินงานของสถาบันอุดมศึกษา. **วารสารวิธีวิทยาการวิจัย.** 16, 1 (มกราคม – มีนาคม 2546): 58-76.

นิตยา เหมือดไธสง. (2543). **การส่งผ่านตัวกลางเชิงสาเหตุของปัจจัยด้านนักเรียน ด้านครู และด้านโรงเรียนไปยังผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์: การวิเคราะห์อภิมานงานวิจัย.** วิทยานิพนธ์ปริญญามหาบัณฑิต. สาขาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

บุญเรือง ศรีเหรัญ. (2542). **การศึกษาองค์ประกอบทางการศึกษาที่สัมพันธ์และมีอิทธิพลต่อผลการเรียนรู้และประสิทธิภาพของโรงเรียนโดยใช้รูปแบบระดับชั้นลดหลั่นแรกเชิงเส้น.**วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต สาขาวิจัยและพัฒนาหลักสูตร บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒประสานมิตร.

รักชนก ยี่สุ่นศรี. (2544). **การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและแบบสอบด้วยกระบวนการ ดี เอฟ ไป ที สำหรับแบบสอบคัดเลือกบุคคลเข้าศึกษาในสถาบันอุดมศึกษา วิชาภาษาอังกฤษและวิชาคณิตศาสตร์.** วิทยานิพนธ์ปริญญามหาบัณฑิต. สาขาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

เลขาธิการสภาการศึกษา, สำนักงาน. (2551).**การพัฒนาคุณภาพการศึกษา : ระเบียบวาระแห่งชาติ.** พริกหวานกราฟฟิค. กรุงเทพ.

ศิริชัย กาญจนวาสี.(2550). **การวิเคราะห์พหุระดับ**. พิมพ์ครั้งที่ 4. กทม.สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี.(2548). **ทฤษฎีการทดสอบแบบดั้งเดิม**. พิมพ์ครั้งที่ 5. กทม.สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี.(2545). **ทฤษฎีการทดสอบแนวใหม่**. พิมพ์ครั้งที่ 2. กทม.สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

สสวท. (2552). **คุณภาพการเรียนรู้ของนักเรียนไทย สังเคราะห์การประเมินผลนักเรียนนานาชาติ PISA 2006 และ TIMSS 2007.** สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.

สสวท. (2551). **ความรู้และสมรรถนะทางวิทยาศาสตร์สำหรับโลกวันพรุ่งนี้ รายงานจากการประเมินผลนักเรียนนานาชาติ PISA 2006**. กรุงเทพ. สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.

สสวท. (2547). **ความรู้และทักษะของเยาวชนไทยสำหรับโลกวันพรุ่งนี้ รายงานจากการวิจัยโครงการประเมินผลนักเรียนนานาชาติ PISA 2000 และ PISA Plus**. กรุงเทพ. สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.

สุชีรา มะหิเมือง. (2547). **ปัจจัยที่ส่งผลผลสัมฤทธิ์และพัฒนาการทางวิชาการ : การวิเคราะห์มูลค่าเพิ่ม.** วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต. สาขาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

สุมาลี แกล้วทนงค์. (2547). **สาเหตุของการทำหน้าที่ต่างกันของข้อสอบสาระการเรียนรู้ภาษาไทยและสาระการเรียนรู้สังคมศึกษา ศาสนาและวัฒนธรรม.** วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต. สาขาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

ภาษาอังกฤษ

Aitkin, M. and Longford, N.T. (1986). Statistical Modelling Issues in School Effectiveness Studies. **Royal Statistical Society** 149(1): 1-43.

Amery, D W. & Ercikan,K.,(2006). Using Multiple-Variable Matching to Identify Cultural Sources of Differential Item Functioning **. International Journal of Testing** 6(3): 287-300.

Angoff, W.H. (1993). Perspectives on differential item functioning methodology. In P.W. Holland & H. Thayer (Eds.), **Differential item functioning** : 397-418. Hillsdale, NJ: Lawrence Erlbaum Associates.

Baldi, S. et al. (2007). **Highlights From PISA 2006: Performance of U.S.15-Year-Old Students in Science and Mathematics Literacy in an International Context**. NCES,U. S. Department of Education.

Ballou, D., Sanders, W. and Wright,P. (2004). Controlling for Student Background in Value-Added Assessment of Teachers. **Journal of Educational and Behavioral Statistics** 29(1): 37–65.

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

Binici, S.(2007). **Random-Effect Differential Item Functioning via Hierarchical Generalized Linear Model and Generalized Linear Latent Mixed Model: A Comparison of estimation method.** Unpublished doctoral dissertation, Florida State University,Department of Educational Psychology and Learning Systems.

Briggs, K.L& Wohlstetter, P.(2003). Key Elements of a Successful School-Based Management Strategy. **School Effectiveness and School Improvement** 14(3): 351-372.

Camilli, G.L., & Shepard, L.A. (1994). **Methods for identifying biased test items.** Thousand Oakes, CA: 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.

Chaimongkol,S.(2005). **Modeling Differential Item Functioning (DIF) Using Mutilevel Logistic Regression Models: A Baysain Perspective.** Unpublished doctoral dissertation, Florida State University,Department of Educational Psychology and Learning Systems.

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.

Cohen, A. S., & Bolt, D. M. (2005). A mixture model analysis of differential item functioning. **Journal of Educational Measurement** 42(2)**:** 133-148.

D'Agostino, J.V. (2000). Instructional and School Effects on Student'Longitudinal Reading and Mathematics Achievements. **School Effectiveness and School Improvement** 11: 197-235.

Dale Ballou, William Sanders, Paul Wright.(2004). Controlling for Student Background in Value-Added Assessment of Teachers . **Journal of Educational and Behavioral Statistics** 29(1): 37–65.

Dorans, N. J., & Holland, P. W. (1993). DIF detection and description: Mantel-Haenszel and standardization. In P. W. Holland & H. Wainer (Eds.), **Differential item functioning.** Hillsdale, NJ: Lawrence Erlbaum Associates.

Downes ,D.& Vindurampulle,O.(2007). **Value–added measures for school improvement.** Education Policy and Research Division . Available from http://www.education.vic.gov.au/studentlearning/research/researchpublications.htm [2008, Feb 13].

Drury, D.& Doran, H. (2003). The Value of Value-Added Analysis. **Policy Research Brief** 3(1): 1-4.

Embretson, S.E., & Reise, S.P. (2000). **Item response theory for psychologists**. Hillsdale, NJ: Erlbaum.

Emenogu, B.C., & Childs, R.a.(2005). Curriculum, Translation, and Differential Functioning of Measurement and Geometry Items. Canadian Journal of Education 28(1 & 2): 128.146

Fitz-Gibbon, C.T, & Kochan, S. (2000). School effectiveness and educational indicators. In C. Teddlie & D. Reynolds (Eds.), **The international handbook of school effectiveness research** (pp. 257–282). London/New York: Routledge/Falmer Press.

Fitz-Gibbon, C.T.(1996). **Monitoring education : indicators, quality and effectiveness**. Great Britain : Redwood Books, Trowbridge, Wiltshire.

Goldstein,H.&Thomas,S.H.,(2008).Reflectipns on the international comparative surveys debate. **Assessment in Education : Principles, Policy & Practice** 15(3): 215-222.

Goldstein, H. (1986). Multilevel Mixed Linear Model Analysis Using Iterative Generalised Least Squares. **Biometrika** 73: 43-56.

Goldstein , H.(2004). International comparisons of student attainment: some issues arising from the PISA study. **Assessment in Education: Principles, Policy & Practice** 11(3): 319 – 330.

Goldhabar, D., & Anthony, E, (2003). Teacher quality and student achievement(Urban Diversity Series No. 115). New York: ERIC Clearinghouse on Urban Education. *A*vailable from http://www.reic.ed.gov/ERICDocs/data/ ericdocs2sq1/content storagq 01/0000019b/80/1b/1e/03.pdf [2008, Feb 11].

Gregory, R J. (2007). **Psychological Testing : History, Principle, and Application**. 5th ed. Pearson Education, Inc.

Grilli, L.G.& Rampichini, C. (2008). **A review of multilevel value-added models in education.** Available from http://www.unifg.it/ev_mnf/080908-09_unirank/ dwn/Grilli_Rampichini.pdf [2008, Feb 11].

Grisay, A. and Monseur, C. (2007). Measuring the Equivalence of item difficulty in the various versions of an International test. **Studies in Educational Evaluation** 33: 69–86.

Hair, J. et al. (1998). **Multivariate Data Analysis.** 5th ed. New Jersey : Prentice-Hall PTR.

Hambleton, R.K., and Swaminathan, H. (1985). **Item Response Theory : Principles and Applications**. Boston, MA : Kluwer Academics Publishers.

Hanan, I.,Hamzeh,D. (2006). Content Analysis of Gender-Related Differential Item Functioning TIMSS Items in Mathematics in Jordan. **School Science and Mathematics** 106(8): 328.

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

Haveman, R. and Wolfe,B. (1995). The Determinants of Children's Attainmemt : A Review of Methods and Findings. **Journal of Economic Literature** 33: 1829-1878.

Heck, 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.

Helen F. Ladd , Randall P. Walsh.(2002). Implementing value-added measures of school effectiveness:getting the incentives right. **Economics of Education Review** 21: 1–17.

Hibpshman, T.(2004). **A Review of Value-Added Models**. Available from http://www.kyepsb.net/documents/Stats/Journals/Heterogeneity%20of%20regression.pdf [2008, Jan 6].

Hofman, R.H., Hofman, W.H.D., Guldemond, H. (1999). Social and cognitive outcomes : A comparison of contexts of learning. **School Effectiveness and School improvement** 10(3): 352-366.

Holland, P. W., & Wainer, H. (Eds.). (1993). **Differential item functioning**. Hillsdale, NJ:Lawrence Erlbaum Associates.

Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), **Test validity** (pp. 129-145). Hillsdale, NJ:Lawrence Erlbaum.

Hoxby, C. (2003). **The Economics of School Choice, National Bureau of Economics Research Conferencd Report.** University of Chicago Press.

Jakubowski, M.(2008). **Implementing Value Added Models of School Assessment. European University Institute**. Available from http://www.eui.eu/RSCAS/Publications/[2008, Feb 13].

Jimerson, S., Egeland, B., Sroufe, L.A. & Carlson, B.,(2000). A Prospective Longitudinal Study of High School Dropouts Examining Multiple Predictors Across Development . **Journal of School Psychology** 38(6): 525-549.

Kamata, A.,Chaimongkol,S.,Genc,E.&Bilir, K.(2005). Random-Effect Differential Item **Functioning Across Group Unites by the Hierarchical Generalized Linear Model. Florida State University**. paper was presented at the annual meeting of the American Educational Research Association, Montreal, Canada, April 2005.

Kamata, A. & Binici, S. (2003). **Random Effect DIF Analysis via Hierarchical Generalized Linear Modeling.** Paper presented at the biannual International Meeting of the Psychometric Society, Sardinia, Italy.

Kamata, A. (2001). Item analysis by the hierarchical generalized linear model. **Journal of Educational Measurement** *38*(1): 79-93.

Kamata, A. (1998). **Some generalizations of the Rasch model: An application of the hierarchical generalized linear model**. Unpublished doctoral dissertation, Michigan State University, East Lansing.

Kellaghan, T.& Greaney, V.,(2001). **Using assessment to improve the quality of education**. Paris, UNESCO. Institute for Educational Planning.

Koretz, D M. & McCaffrey,D F.(2005). **Using IRT DIF Methods to Evaluate the Validity of Score Gains .** (CSE Technical Report No. 660). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Koutsoulis, M. K., & Campbell, J. R. (2001). Family processes affect students' motivation, and science and math achievement in Cypriot high schools. **Structural Equation Modeling** *8*(1): 108-172.

Ladd, H.F. and Walsh, R.P. (2002). Implementing Value-Added Measures of Sckool Effectiveness : Getting the Incentives Right. **Economics of Education Review** 21: 1-17.

Lambert, D., and Lines, D. (2000). **Understanding assessment, purpose, perceptions, practice**, Great Britain : TJ International Ltd. Padstow Cornwall.

Le, L.T. (2006). **Investigating Gender Differential Item Functioning across Countries and Test Languages for PISA Science Items**. Paper presented at the 5th conference of International Test Commission,Brussels.

Lee, V. E. (2000). Using hierarchical linear modeling to study social context: The case of school effects. **Educational Psychologist** *35*(2): 125-141.

Lindsay, G., and Desforges, M. (1998). **Baseline assessment : practice and possibilities**. London : David Fullton.

Linn, R. L., & Hirnisch, D. L. (1981). Interactions between item content and group membership on achievement test items. **Journal of Educational Measurement** 18: 109–118.

Lissitz, R.W,(2006).Longitudinal and **Value Added Models of Student Performance**, JAM Press, Minnesota.

Lissitz, R.W,(2005). **Value-Added Models in education: Theory and Applications**, JAM Press, Minnesota.

McGraw, B.(2008). The role of the OECD in international comparative studies of achievement. **Assessment in Education : Principles, Policy & Practice** 15(3): 213-243.

Mapuranga,R., Dorans,N J.,& Middleton, K.(2008) . **A Review of Recent Developments in Differential Item Functioning .** Paper presented at the annual meeting of theNational Council on Measurement in Education (NCME). New York.

Maslowski, R. ,Doolaard, S.& Bosker, R.J.(2008). Much more than the effective classroom: a lifetime of research, evaluation,improvement, and dissemination. **School Effectiveness and School Improvement** 19(4): 353 – 363.

Matear, A. (2006). Equity in education in Chile: The tensions between policy and practice **International Journal of Educational Development** *27*(1): 101-113.

McCaffrey, D.F., Lockwood, J.R., Koretz, D.M.,&Hamilton, L. S., (2003). Evaluating value-added models for teacher accountability. Santa Monica, CA: RAND Corporation. *A*vailable from http:www.rand.org/pubs/monographs/2004/RAND_MG158.pdf [2008, Jan 3].

McKewen, N. (1995). Accountablity in Education in Canada. **Canadian Journal of Education** 20(1).

Miller, R. I. (1990). **Major American higher education issues and challenges in the 1990s**. London : Jessica Kingsley Published Ltd.

Narayanan, P., & Swaminathan, H.(1994). Performance of the Mantel-Haenszel and simultaneous item bias procedures for detecting differential item functioning. **Applied Psychological Measurement** 18(4): 315-328.

OECD. (2004). **Learning for Tomorrow's World: First Results from PISA 2003**. Paris: OECD.

OECD. (2005). **School Factors Related to Quality and Equity: Results from PISA 2000.** Paris: OECD.

OECD. (2006). Demand Sensitive Schooling? Evidence and Issues. Paris : OECD.

OECD. (2006). **Assessing Scientific, Reading and Mathematical Literacy A Framework for PISA 2006**. Paris : OECD.

OECD. (2007). **PISA 2006 Science Competencies for Tomorrow's World: Volume 1 – Analysis**. Paris: OECD.

OECD. (2007b). **Learning for Tomorrow**. Paris: OECD.

OECD.(2008). **Measuring Improvements in Learning Outcomes : Best practices to assess the Value-Added of schools**. Paris: OECD.

OECD. (2009). **PISA 2006 Technical Report**. Paris: OECD.

Opdenakker, M. C., and van Damme, J. (2000). Effect of school, teaching staff and classes on achievement and well-being in secondary education : Similarities and differences between school outcomes. **School Effectiveness and School Improvement** 11(2): 165-196.

OsterLind, S J.(1987). **Test Item Bias**. United States of America. 3rd. Sage Publications.Inc.

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).

Patrick, B., and Tomasz, G., (2004). **The Impact of Computer Use on Reading Achievement of 15-year-olds.** General enquiries regarding the documents published by the Strategic Policy and Planning. Available from http://www.hrsdc-rhdcc.gc.ca/sp-ps/arb-dgra[2008, Jan 6].

Pintrich, J.L., & Schunk, D.H.(2002). **Motivation in Education.** Englewood Cliffs, NJ,Merrill.

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

Raudenbush, S. W., & Bryk, A. S. (2002). **Hierarchical linear models: Applications and data analysis methods.** Thousand Oaks, CA: Sage Publications.

Riddell, A.,(2008). **Factors Influencing Educational Quality and Effectiveness in Developing Countries:A Review of Research.**Deutsche Gesellschaft für Technische Zusammenarbeit . Eschborn.

Salganik, L. (1994). Apples and apples: Comparing performance indicators for places with similar demographic characteristics. **Educational Evaluation and Policy Analysis** 16(2): 125-142.

Sanders, W. L., & Horn, S. (1998). Research findings from the Tennessee Value-Added Assessment System (TVAAS) database: Implications for educational evaluation and research. **Journal of Personnel Evaluation in Education** *12*(3): 247–256.

Scheerens, J.(2001). Introduction School Effectiveness in Developing Countries. **School Effectiveness and School Improvement** 12(4): 353 – 358.

Scheerens, J., and Bosker, R. J. (1997). **The foundation of educational effectiveness**. Great Britain : Redwood Books Ltd.

Shealy, R. T., & Stout, W. F. (1993). An item response theory model for test bias and differential item functioning. In P. Holland & H. Wainer (Eds.), **Differential item functioning**. Hillsdale, NJ: Erlbaum.

Schulz, W. (2005). Measuring the socio-economic background of students and its effect on achievement in PISA 2000 and PISA 2003**. Paper prepared for the Annual Meetings of the American Educational Research Association in San Francisco**, 7-11 April 2005.

Silin, H. C., and Murray-Harvey, R. (1999). Students as a central concern school, students and outcome measures, **Journal of Educational Administration** 38(3): 230-246.

Steven G. Rivkin(2007). **Value-Added Analysis and Education Policy. A***vailable from* http://www.caldercenter.org[2008, March 10].

Stephen W. Raudenbush. 2004.What Are Value-Added Models Estimating and What Does This Imply for Statistical Practice?. **Journal of Educational and Behavioral Statistics** 29(1): 121–129.

Strand, S. (1998). A value-added analysis of 1996 primary school performance table. **Education Research** 40: 123-137.

Teddlie, C. and Reynolds, D., (2000). **The International Handbook of School Effectiveness Research***,* Falmer, London.

Tekwe, C.D., Carter, R.L., Ma, C., Algina, J., Lucas, M.E., Roth, J., Ariet, M., Fisher, T., and Resnick, M.B. (TCMLRAFR) (2004) An empirical comparison of statistical models for value-added assessment of school performance. **Journal of Educational and Behavioral Statistics** 29(1): 11 – 35.

Thissen, D., Steinberg, L., & Wainer, H. (1988). Detect of Differential Item Functioning Using the Parameters of Item Response Models. In P. W. Holland & H. Wainer. (Eds.), **Differential item functioning.** Hillsdale, NJ: Lawrence Erlbaum Associates.(147-169). Hillsdale, NJ: Lawrence Erlbaum Associates.

Thomas, S .(2001). 'Dimensions of secondary school effectiveness: comparative analyses across regions', **School Effectiveness and School Improvement** 12(3).

Thomas,S.,(1996), **Comparison of value-added models for secondary-school effectiveness.Research Papers in Education** 11(1): 5 – 33.

Tomasz, G. and Dhawan-Biswal,U. (2008). Reading skills of young immigrants in Canada: the effects of duration of residency, home language exposure and schools. **General enquiries regarding the documents** ,Learning Policy Directorate

Vaughn, B K. (2006). **A Hierarchical Generalized Linear Model of random Differential Item Functioning for Polytomous Items: A Bayesian Multilevel Approach.** Unpublished doctoral dissertation, Florida State University,Department of Educational Psychology and Learning Systems.

Weiss, D.J., and Yoes, M. E. (1991). Item Response Theory. In R.K. Hambleton., and J.N. Zaal. (Eds). **Advances in Educational and Psychological Testing**. United States of America : Kluwer Academic, p. 67-95.

Wiley, D.,&Yoon, B. (1995). Teacher reports on opportunity to learn: Analyses of the 1993 California Learning Assessment System (CLAS). **Educational Evaluation and Policy Analysis** *17*(3): 355-370.

Williams, N. J., & Beretvas, S. N. (2006). DIF identification using HGLM for polytomous item. **Applied Psychological Measurement** *30*(1): 22-42.

Willms, J.D.(2004). Reading Achievement in Canada and the United States: Findings from the OECD Programme for International Student Assessment. **Learning Policy Directorate Strategic Policy and Planning Human Resources and Skills Development Canada.** Available from http://www.hrsdc-rhdcc.gc.ca/sp-ps/arb-dgra[2007, December 23].

Willms, J., and Raudenbush, S.(1989). A longitudinal hierarchical linear model for estimating school effect and their stability. **Journal of Educational Measurement** 26(3): 209-232.

Willims, J.D., and Somers, M.A. (2001). Family, classroom and school effect on children's Educational outcome in Latin American. **School Effectiveness and School Improvement** 12(4): 409-445.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

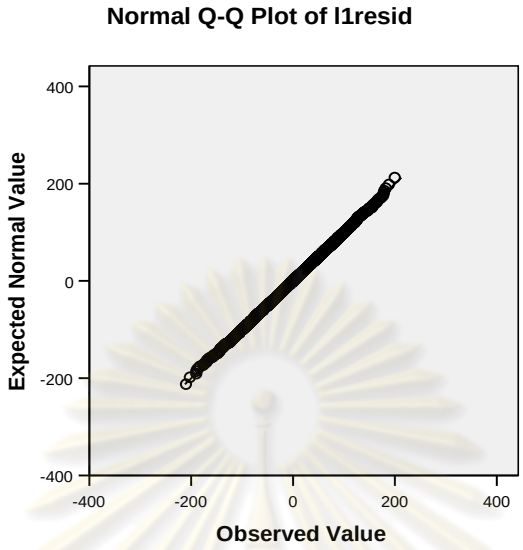
ภาคผนวก ก

ผลการตรวจสอบเงื่อนไขเบื้องต้น

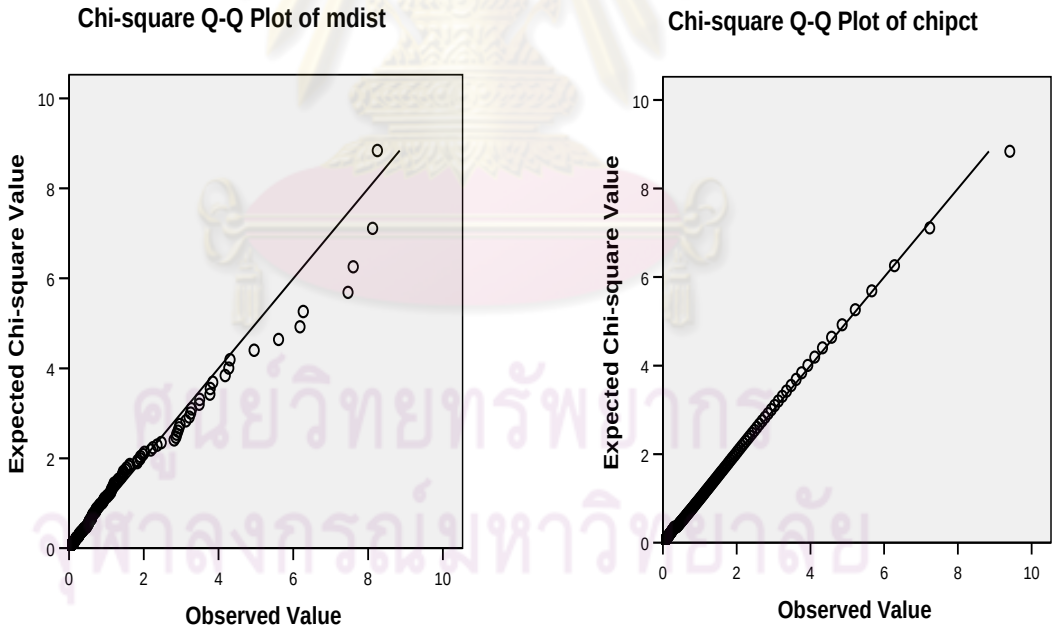
1) การตรวจสอบการแจกแจงของความคลาดเคลื่อน (Distributional assumptions)
2) การตรวจสอบความเป็นเอกมิติของแบบสอบ (Unidimentionality)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

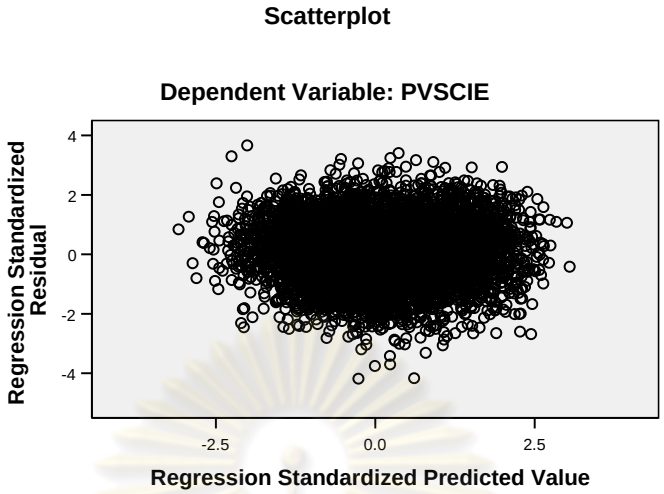
1) การตรวจสอบการแจกแจงของความคลาดเคลื่อน (Distributional assumptions)



ภาพที่ 22 กราฟ Q-Q plot ของ residual level 1



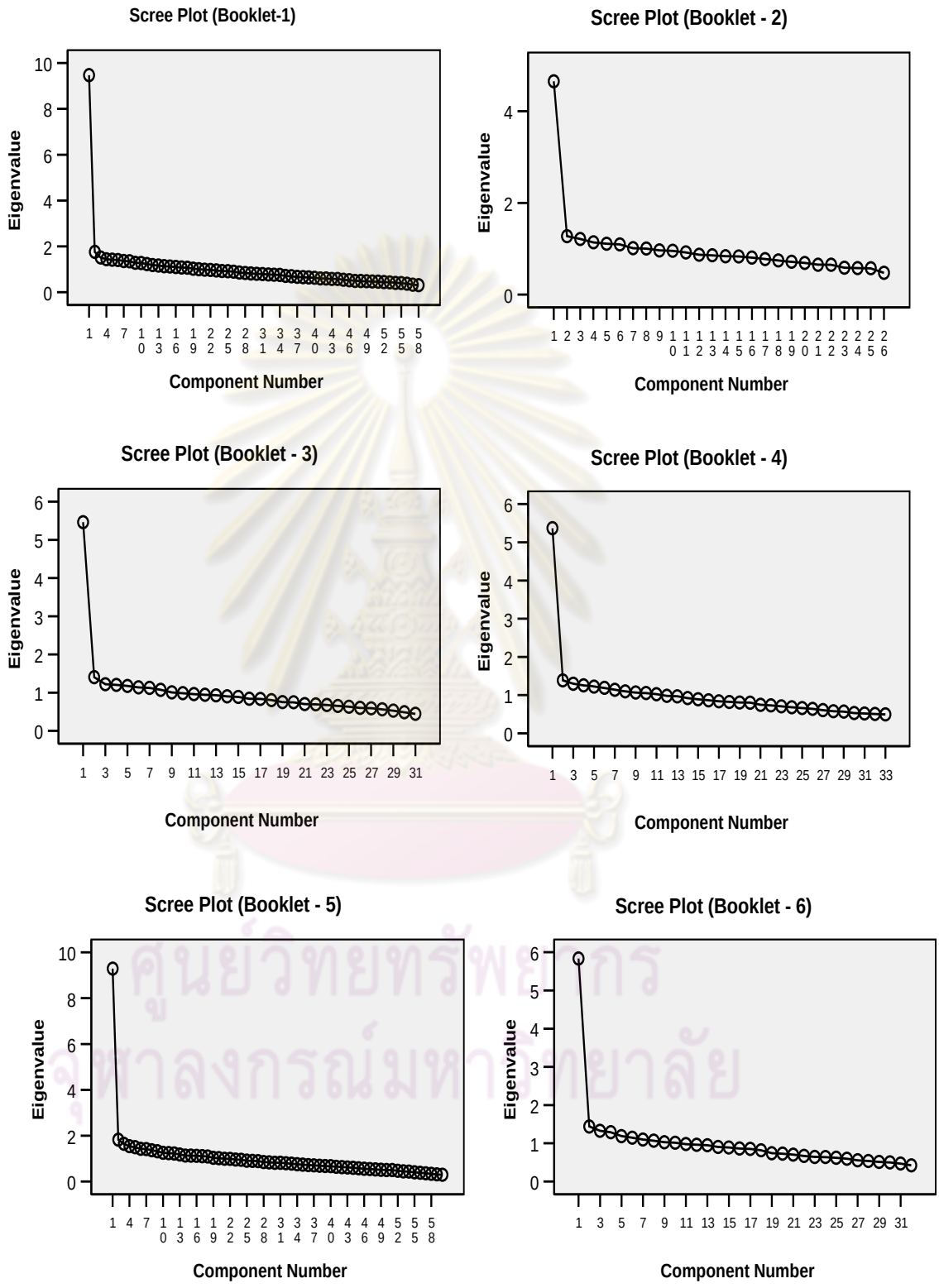
แผนภาพที่ 23 กราฟ Q-Q plot ของ residual level 2



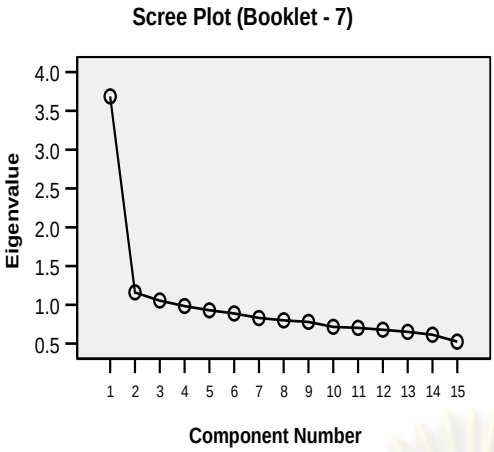
ภาพที่ 24 กราฟระหว่างคะแนนความคลาดเคลื่อน (residual) กับค่าทำนาย (Expected value)

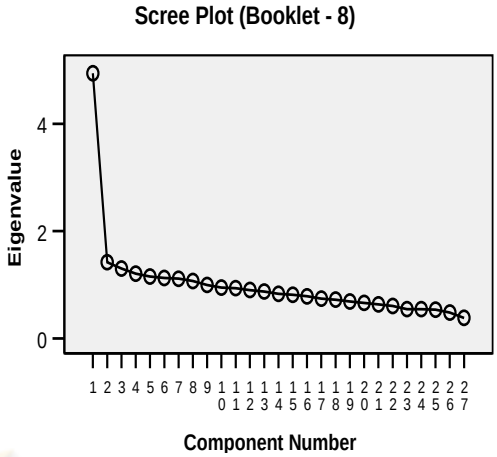
ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

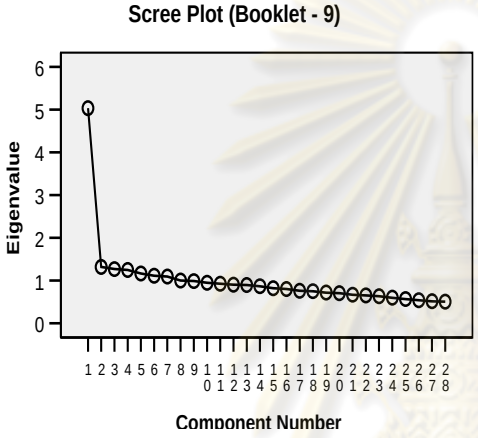
2) การตรวจสอบความเป็นเอกมิติของแบบสอบ (Unidimentionality)

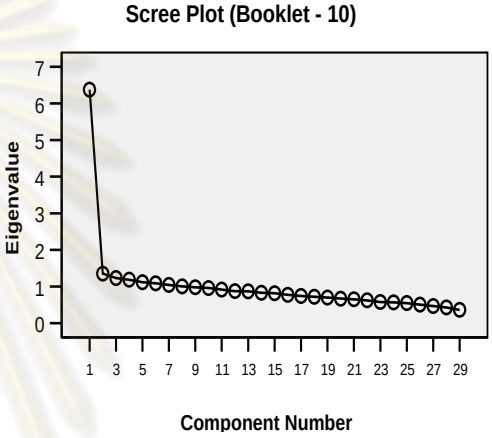


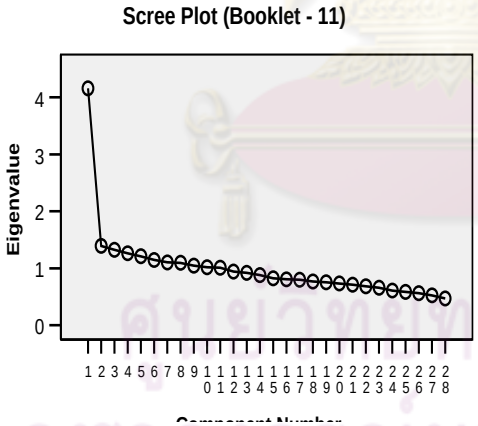
ภาพที่ 25 แผนภาพสกรีของการวิเคราะห์องค์ประกอบของแบบสอบฉบับที่ 1 – 13

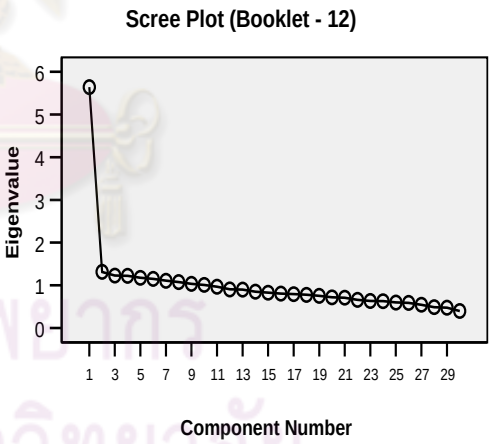


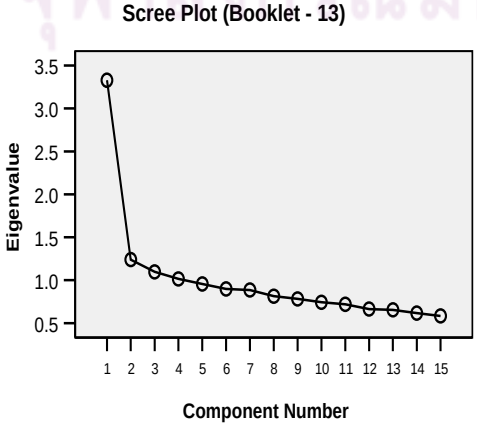






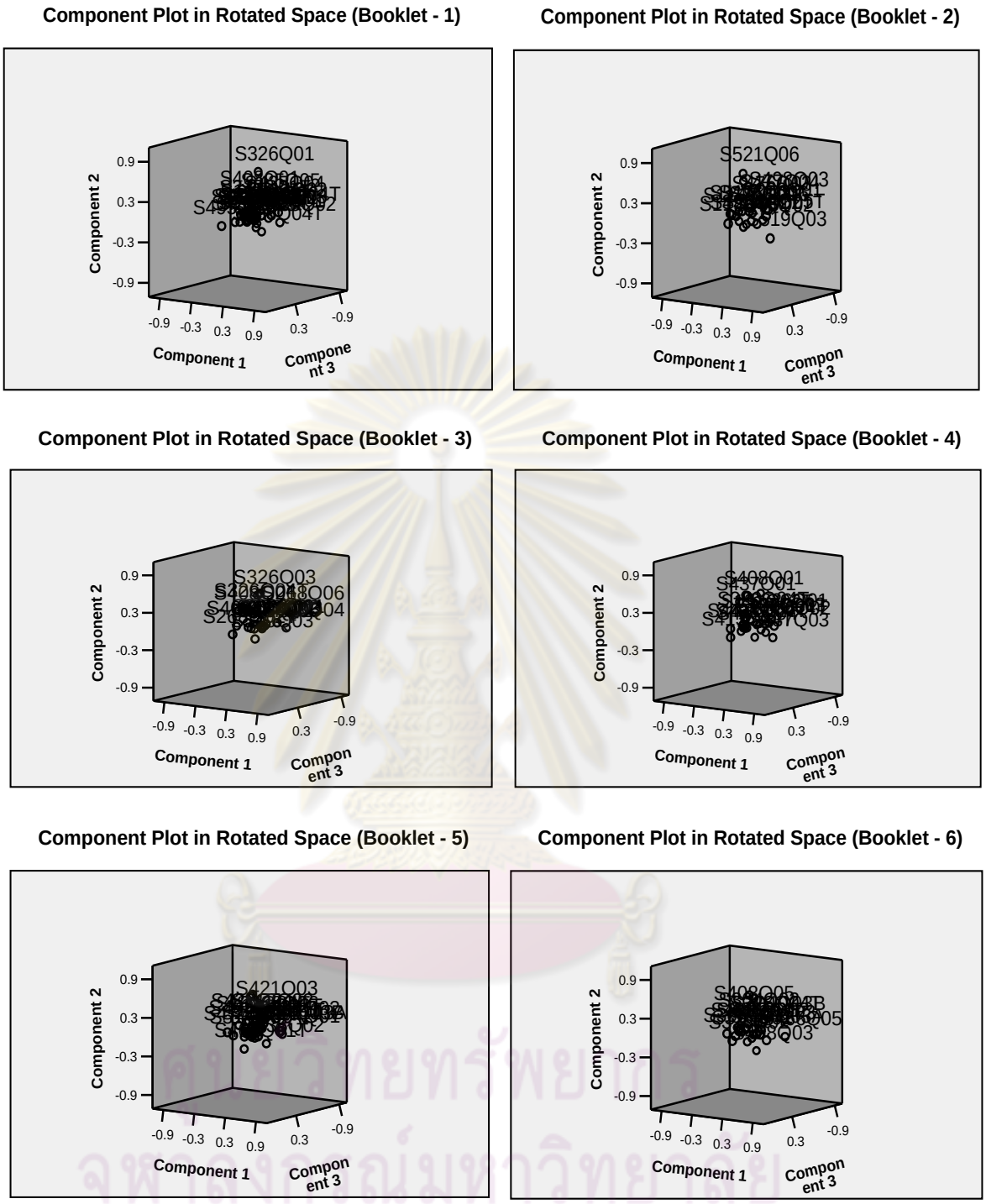






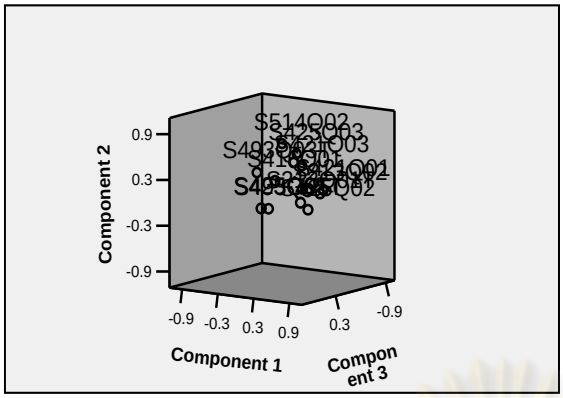
ภาพที่ 25 แผนภาพสกรีของการวิเคราะห์องค์ประกอบของแบบสอบ ฉบับที่ 1 – 13(ต่อ)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

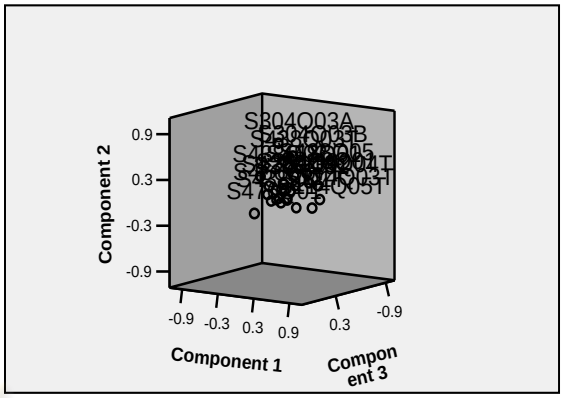


ภาพที่ 26 แผนภาพองค์ประกอบของแบบสอบฉบับที่ 1 – 13

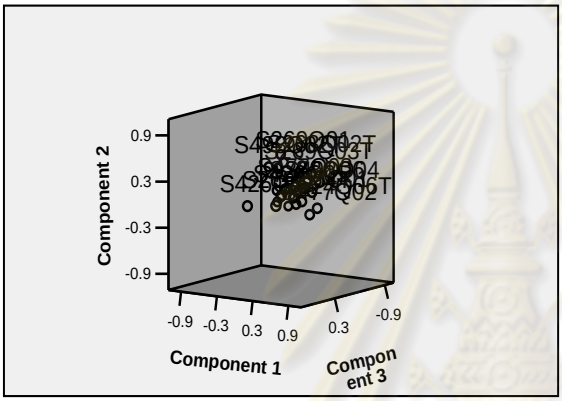
Component Plot in Rotated Space (Booklet - 7)



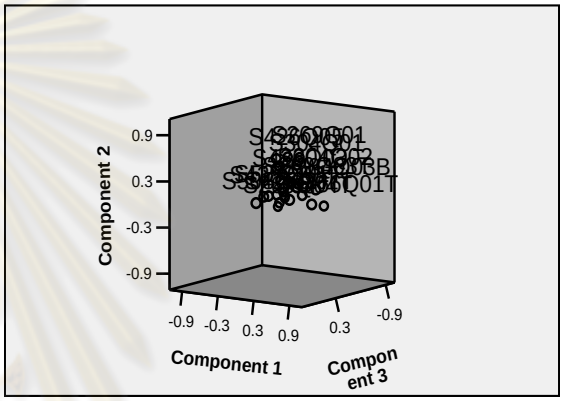
Component Plot in Rotated Space (Booklet - 8)



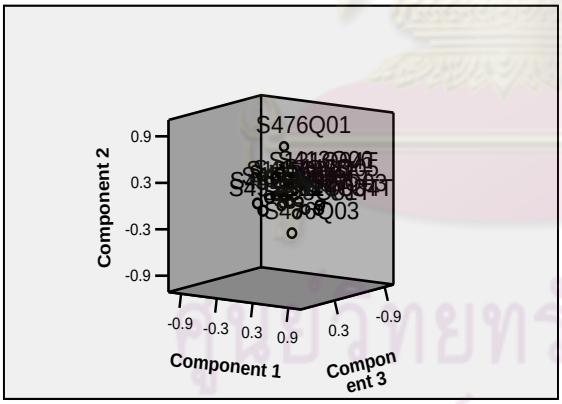
Component Plot in Rotated Space (Booklet - 9)



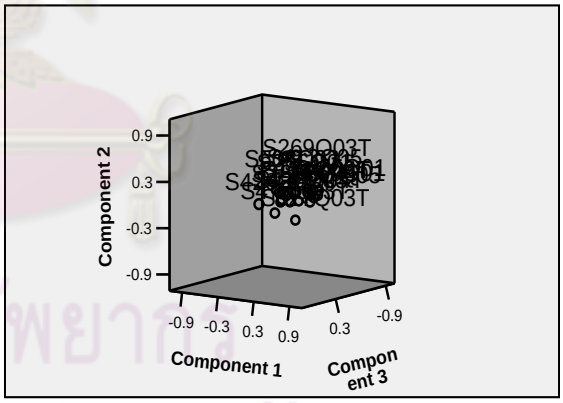
Component Plot in Rotated Space (Booklet - 10)



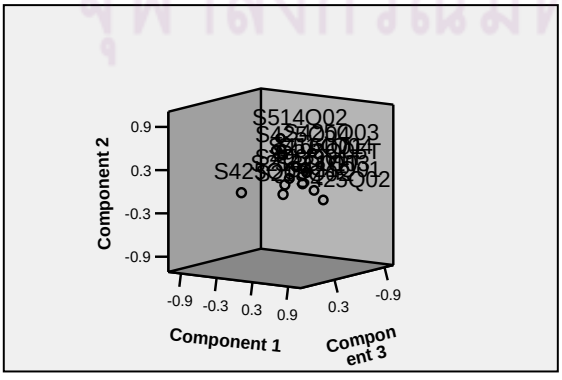
Component Plot in Rotated Space (Booklet - 11)



Component Plot in Rotated Space (Booklet - 12)



Component Plot in Rotated Space (Booklet - 13)



ภาพที่ 26 แผนภาพองค์ประกอบของแบบสอบฉบับที่ 1 – 13 (ต่อ)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ข

ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 1 – 13

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 1

ตารางที่ 46 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 1

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
Dichotomous Item										
S114Q03T	0.606435*	0.041	0.033102	0.914	0.779936	0.096	-0.214919	0.635	-0.178015	0.728
	(0.296807)		(0.30566)		(0.46885)		(0.453659)		(0.511692)	
S114Q05T	0.692626	0.053	-0.319731	0.377	1.144011*	0.030	-0.723312	0.169	-0.246559	0.665
	(0.358888)		(0.361866)		(0.5266)		(0.526348)		(0.570314)	
S213Q01T	0.939074**	0.002	0.044834	0.886	1.743805**	0.001	-0.331893	0.474	-0.937127	0.079
	(0.304651)		(0.310781)		(0.48057)		(0.463901)		(0.535133)	
S213Q02	0.452128	0.105	-0.148101	0.606	0.666528	0.152	0.222429	0.621	-0.215653	0.674
	(0.279325)		(0.287636)		(0.46475)		(0.449987)		(0.51254)	
S269Q01	0.209764	0.448	-0.001255	0.997	1.375993**	0.005	-0.41198	0.368	-0.731643	0.165
	(0.276491)		(0.285907)		(0.49027)		(0.457124)		(0.526946)	
S269Q03T	0.304652	0.298	-0.214536	0.475	0.435771	0.356	-0.140252	0.756	0.370501	0.466
	(0.292556)		(0.300696)		(0.47166)		(0.452074)		(0.508873)	
S269Q04T	-0.426090	0.221	0.116446	0.746	1.235170*	0.022	-0.941293	0.096	-0.856717	0.160
	(0.347922)		(0.360142)		(0.54098)		(0.56557)		(0.610348)	
S326Q01	1.037367**	0.000	0.191933	0.502	0.130414	0.783	0.436487	0.355	-0.516416	0.323
	(0.276343)		(0.286267)		(0.474498)		(0.471353)		(0.521655)	
S326Q02	0.801476**	0.007	-0.145808	0.627	0.335116	0.478	0.048414	0.914	-0.296937	0.562
	(0.294483)		(0.300565)		(0.472693)		(0.447497)		(0.512824)	
S326Q03	0.568710*	0.048	-0.221483	0.454	0.857512	0.063	-0.196094	0.661	-0.278656	0.583
	(0.288361)		(0.295689)		(0.462859)		(0.447085)		(0.507502)	
S326Q04T	0.701540*	0.043	-0.167837	0.632	1.178107*	0.019	0.035284	0.942	-0.251107	0.644
	(0.346819)		(0.35124)		(0.501154)		(0.481328)		(0.544417)	
S408Q01	0.632247*	0.022	-0.062609	0.826	0.494735	0.294	-0.566316	0.211	0.080434	0.879
	(0.27517)		(0.284519)		(0.471089)		(0.452652)		(0.524918)	
S408Q03	0.796614**	0.008	0.103808	0.736	-0.145684	0.770	-0.469712	0.320	-0.009238	0.986
	(0.299842)		(0.30792)		(0.498456)		(0.471601)		(0.52995)	
S408Q04T	0.298618	0.276	0.552047	0.051	-0.023273	0.960	-0.389153	0.379	-0.697132	0.167
	(0.273584)		(0.283238)		(0.46082)		(0.441947)		(0.503911)	
S408Q05	0.395257	0.165	-0.043984	0.881	0.076990	0.871	-0.222867	0.619	-0.271153	0.595
	(0.284526)		(0.293884)		(0.473123)		(0.448657)		(0.511013)	
S415Q02	0.999876**	0.001	-0.047978	0.868	0.446093	0.358	-0.52996	0.256	0.353463	0.523
	(0.278302)		(0.288324)		(0.485064)		(0.466067)		(0.554182)	
S415Q07T	0.770729**	0.005	0.072049	0.800	1.058425*	0.027	-0.836631	0.062	-0.974980	0.058
	(0.27437)		(0.283969)		(0.479333)		(0.44911)		(0.516109)	

ตารางที่ 46 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 1 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S415Q08T	0.734563**	0.008	0.126784	0.655	0.337928	0.465	-0.069866	0.876	-0.273944	0.589
	(0.274992)		(0.284362)		(0.462566)		(0.445761)		(0.508018)	
S416Q01	0.221296	0.442	-0.124524	0.674	0.948728*	0.043	0.230109	0.606	-1.126571*	0.032
	(0.287672)		(0.296846)		(0.469271)		(0.446792)		(0.524845)	
S421Q01	0.243638	0.413	-0.203750	0.505	1.069651*	0.023	-0.375397	0.414	-0.481153	0.353
	(0.297337)		(0.306068)		(0.471069)		(0.458969)		(0.517345)	
S421Q03	-0.478926	0.090	-0.063141	0.829	1.156676*	0.014	-0.083202	0.854	-0.525524	0.310
	(0.283088)		(0.291727)		(0.470385)		(0.452156)		(0.516776)	
S425Q02	0.495411	0.088	-0.479459	0.107	1.241016**	0.008	-0.767305	0.096	-0.155042	0.762
	(0.290907)		(0.297414)		(0.46731)		(0.461677)		(0.511291)	
S425Q03	0.256276	0.376	-0.351100	0.238	0.605603	0.196	-0.240023	0.593	-0.328649	0.519
	(0.289356)		(0.297238)		(0.468087)		(0.449403)		(0.5098)	
S425Q04	1.062735**	0.003	-0.258129	0.456	0.200030	0.704	-0.089015	0.857	0.127851	0.818
	(0.349957)		(0.34622)		(0.527721)		(0.491498)		(0.554583)	
S425Q05	0.434748	0.120	0.552444	0.059	0.044830	0.924	-0.190333	0.670	0.370597	0.468
	(0.280074)		(0.29344)		(0.468676)		(0.447627)		(0.510206)	
S426Q03	0.799308**	0.004	0.275276	0.340	1.228311*	0.016	-0.129816	0.783	-1.074703*	0.046
	(0.277846)		(0.288094)		(0.510543)		(0.471535)		(0.539737)	
S426Q05	0.226924	0.412	0.095892	0.736	0.303924	0.518	-0.438408	0.331	-0.477944	0.353
	(0.27631)		(0.285203)		(0.469979)		(0.450205)		(0.514611)	
S426Q07T	0.660399*	0.016	0.009164	0.974	0.437056	0.349	-0.350043	0.432	-0.843282	0.096
	(0.273404)		(0.283009)		(0.465986)		(0.445511)		(0.507607)	
S437Q01	-0.087369	0.755	0.031169	0.915	0.597157	0.200	0.449341	0.316	-0.569980	0.267
	(0.280733)		(0.290409)		(0.465465)		(0.447698)		(0.512804)	
S437Q03	0.338499	0.221	-0.140265	0.623	0.463187	0.317	-0.429664	0.333	-0.956187	0.060
	(0.276329)		(0.285298)		(0.462403)		(0.443352)		(0.508995)	
S437Q04	0.300989	0.274	0.047079	0.869	0.129725	0.779	-0.519006	0.242	-0.017131	0.973
	(0.275017)		(0.284549)		(0.461748)		(0.443542)		(0.504154)	
S437Q06	-0.065455	0.813	0.092662	0.746	0.404572	0.386	0.088436	0.845	0.113599	0.826
	(0.277184)		(0.286299)		(0.466288)		(0.452045)		(0.517210)	
S476Q01	0.325078	0.245	-0.189619	0.512	0.114960	0.811	-0.180497	0.700	-0.096147	0.859
	(0.279589)		(0.289324)		(0.480069)		(0.469791)		(0.538959)	
S476Q02	0.091442	0.740	-0.277064	0.331	0.654160	0.167	-0.330257	0.466	-0.618626	0.231
	(0.275899)		(0.285029)		(0.473544)		(0.452631)		(0.516578)	
S476Q03	0.533563	0.065	-0.224310	0.450	0.878171	0.062	-0.473893	0.299	-1.059836*	0.043
	(0.289134)		(0.296732)		(0.471527)		(0.455654)		(0.524955)	

ตารางที่ 46 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 1 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S485Q02	0.414248	0.164	-0.089911	0.769	0.939071*	0.046	0.155822	0.729	-0.678275	0.193
	(0.297915)		(0.306752)		(0.470483)		(0.449617)		(0.520720)	
S485Q03	0.607542*	0.030	-0.278631	0.339	0.190045	0.694	-0.264139	0.575	-0.179862	0.739
	(0.279907)		(0.290948)		(0.484052)		(0.471894)		(0.541051)	
S493Q01T	0.597053*	0.031	0.036592	0.899	0.454036	0.325	0.069126	0.876	-0.600846	0.236
	(0.277463)		(0.286559)		(0.461364)		(0.442415)		(0.506333)	
S493Q03T	0.590532*	0.047	0.619017	0.043	0.644886	0.238	-0.446557	0.377	-0.199807	0.737
	(0.297953)		(0.306224)		(0.545645)		(0.504893)		(0.594692)	
S493Q05T	0.086505	0.802	-0.102593	0.774	0.999150*	0.047	0.555299	0.242	-0.034498	0.950
	(0.345437)		(0.357056)		(0.504745)		(0.474270)		(0.542902)	
S495Q01T	0.715218*	0.012	-0.184448	0.526	0.592878	0.201	-0.117289	0.791	-0.765387	0.134
	(0.283435)		(0.290821)		(0.463835)		(0.443512)		(0.510626)	
S495Q02T	-0.187697	0.497	0.281969	0.327	0.326654	0.482	-0.717702	0.111	-0.025064	0.961
	(0.276602)		(0.287506)		(0.464334)		(0.450760)		(0.507428)	
S495Q03	0.757619*	0.011	0.043758	0.886	1.487183**	0.002	-0.366441	0.425	-0.289084	0.576
	(0.294828)		(0.303108)		(0.468920)		(0.459012)		(0.517564)	
S495Q04T	1.208373**	0.000	-0.192131	0.518	0.706405	0.131	-0.039050	0.931	-0.536915	0.297
	(0.293809)		(0.297057)		(0.467407)		(0.447486)		(0.513810)	
S508Q02T	0.608171*	0.027	0.268229	0.346	0.529240	0.256	-0.114269	0.797	-1.194200*	0.020
	(0.274684)		(0.284420)		(0.464997)		(0.444054)		(0.512663)	
S508Q03	0.895141**	0.002	0.193167	0.498	0.283096	0.546	-0.420791	0.353	0.295779	0.574
	(0.275710)		(0.285272)		(0.469644)		(0.452996)		(0.526903)	
S510Q01T	0.314712	0.251	-0.112059	0.692	0.501176	0.281	-0.819695	0.066	-0.194635	0.702
	(0.273823)		(0.283031)		(0.464517)		(0.446185)		(0.510087)	
S510Q04T	-0.000409	0.999	-0.268853	0.365	1.096590*	0.020	-0.021022	0.963	-1.071810*	0.040
	(0.287968)		(0.296609)		(0.469886)		(0.448629)		(0.523148)	
S514Q02	0.299499	0.280	-0.054459	0.849	0.470881	0.327	0.116246	0.805	-0.347339	0.512
	(0.277036)		(0.286177)		(0.479895)		(0.470684)		(0.529827)	
S514Q03	0.416765	0.138	-0.087737	0.762	0.996609*	0.032	-0.416607	0.351	-0.814123	0.111
	(0.280897)		(0.289790)		(0.464709)		(0.446735)		(0.511605)	
S514Q04	0.584013	0.059	-0.354748	0.260	0.896674	0.057	0.178713	0.693	0.052032	0.920
	(0.309372)		(0.314580)		(0.472280)		(0.452756)		(0.513719)	
S521Q02	0.297200	0.279	-0.225995	0.425	-0.067397	0.885	-0.432518	0.328	-0.137076	0.786
	(0.274482)		(0.283264)		(0.463377)		(0.442197)		(0.504656)	
S521Q06	0.743648**	0.009	0.193829	0.505	0.595070	0.236	0.055974	0.909	-0.570329	0.295
	(0.281019)		(0.291123)		(0.501534)		(0.488243)		(0.544592)	

ตารางที่ 46 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 1 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S527Q01T	0.441845	0.188	0.083300	0.810	0.751575	0.148	-0.309745	0.539	-0.497512	0.381
	(0.335309)		(0.346873)		(0.519456)		(0.504836)		(0.568091)	
S527Q03T	0.263532	0.339	-0.113675	0.689	0.738307	0.110	-0.908981	0.043	-0.456192	0.367
	(0.275357)		(0.284527)		(0.462924)		(0.448958)		(0.505202)	
Polytomous item										
S114Q04T	0.224192	0.407	0.483530	0.094	-1.101458*	0.011	-0.163111	0.656	-0.683343	0.130
	(0.270324)		(0.288408)		(0.430836)		(0.366239)		(0.451180)	

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 2

ตารางที่ 47 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 2

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
Dichotomous Item										
S114Q03T	-0.001196	0.997	-0.262187	0.336	0.420499	0.288	0.294180	0.465	-0.471757	0.312
	(0.277324)		(0.272312)		(0.395608)		(0.402882)		(0.465992)	
S114Q05T	0.020518	0.951	0.536490	0.102	0.911933*	0.047	0.036260	0.939	-0.692465	0.193
	(0.329229)		(0.328783)		(0.460526)		(0.473304)		(0.532079)	
S268Q01	-0.197889	0.461	-0.221789	0.416	-0.255722	0.510	-0.194576	0.623	-0.082907	0.883
	(0.268264)		(0.272413)		(0.388601)		(0.396109)		(0.559376)	
S268Q02T	-0.065587	0.825	0.117531	0.661	0.544025	0.208	-0.183239	0.689	-0.716417	0.215
	(0.295619)		(0.268088)		(0.431527)		(0.458081)		(0.577168)	
S268Q06	-0.288768	0.278	0.192335	0.485	0.114080	0.787	0.075090	0.812	-0.418797	0.389
	(0.266077)		(0.275446)		(0.421276)		(0.315754)		(0.486301)	
S476Q01	-0.089622	0.741	-0.061804	0.805	0.130204	0.793	-0.259884	0.512	-0.425971	0.454
	(0.271628)		(0.250803)		(0.497142)		(0.396637)		(0.569488)	
S476Q02	-0.374013	0.163	0.126874	0.639	0.014399	0.971	-0.343117	0.387	0.271941	0.607
	(0.268248)		(0.270587)		(0.386232)		(0.396290)		(0.528742)	
S476Q03	-0.152378	0.614	0.490156	0.103	0.477112	0.315	-0.424923	0.328	-0.513200	0.348
	(0.302207)		(0.300783)		(0.473963)		(0.434314)		(0.546546)	
S477Q02	0.522544	0.079	0.067798	0.821	-0.129100	0.808	0.238922	0.610	-0.547820	0.361
	(0.298194)		(0.299776)		(0.531265)		(0.468555)		(0.598867)	
S477Q03	0.007254	0.980	-0.086391	0.744	0.847275	0.079	-0.281785	0.539	1.300785	0.174
	(0.289089)		(0.264998)		(0.483040)		(0.458630)		(0.955562)	
S477Q04	0.068924	0.804	0.229375	0.362	-0.105041	0.799	0.322738	0.392	-0.718154	0.181
	(0.278062)		(0.251225)		(0.411559)		(0.376691)		(0.536427)	

ตารางที่ 47 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 2 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S495Q01T	0.020496	0.938	0.217427	0.395	0.116884	0.779	-0.211632	0.589	-0.415829	0.415
	(0.261739)		(0.255721)		(0.416570)		(0.392022)		(0.509465)	
S495Q02T	-0.305576	0.270	-0.411645	0.144	0.502374	0.244	-0.350019	0.370	-0.054035	0.922
	(0.276730)		(0.281909)		(0.430531)		(0.390484)		(0.551516)	
S495Q03	0.161374	0.519	0.440555	0.109	0.701843	0.061	0.179484	0.639	-0.137460	0.776
	(0.250123)		(0.275213)		(0.375571)		(0.382769)		(0.482183)	
S495Q04T	0.352625	0.216	-0.028710	0.909	0.579757	0.165	0.301414	0.467	-0.679935	0.206
	(0.284496)		(0.249630)		(0.417065)		(0.414862)		(0.536818)	
S498Q02T	-0.104454	0.725	0.344417	0.212	0.548069	0.142	0.028275	0.943	-0.914778	0.064
	(0.297164)		(0.275714)		(0.373043)		(0.390722)		(0.494613)	
S498Q03	0.099127	0.720	-0.039207	0.886	0.447041	0.330	-0.058622	0.891	-1.042722*	0.042
	(0.276650)		(0.271613)		(0.458519)		(0.427860)		(0.512600)	
S519Q02T	-0.577343*	0.045	-0.057093	0.831	-0.175227	0.635	0.089430	0.827	-0.392225	0.427
	(0.288343)		(0.267741)		(0.369318)		(0.408604)		(0.493999)	
S519Q03	-0.647684	0.112	0.361380	0.415	-0.424911	0.472	0.518921	0.380	-0.047769	0.935
	(0.407985)		(0.442718)		(0.590667)		(0.590484)		(0.585362)	
S521Q02	-0.310886	0.234	-0.057178	0.832	-0.201359	0.598	-0.176181	0.630	-0.144753	0.768
	(0.260933)		(0.269320)		(0.381953)		(0.366541)		(0.489950)	
S521Q06	-0.391611	0.171	0.082873	0.763	0.586755	0.210	-0.355047	0.372	-0.319814	0.559
	(0.286187)		(0.275358)		(0.468087)		(0.397831)		(0.547778)	
S524Q06T	-0.047261	0.858	0.168075	0.522	0.140411	0.768	0.179409	0.650	-0.683591	0.245
	(0.264206)		(0.262937)		(0.475607)		(0.395442)		(0.587194)	
Polytomous Item										
S498Q04	-0.254513	0.285	-0.244150	0.280	-0.610151	0.125	-0.111069	0.768	0.322134	0.467
	(0.237762)		(0.225583)		(0.397962)		(0.377349)		(0.442779)	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 3

ตารางที่ 48 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 3

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
Dichotomous Item										
S268Q01	0.014198	0.961	0.205704	0.484	-0.584227	0.234	0.044576	0.921	-0.195377	0.737
	(0.289961)		(0.294073)		(0.49078)		(0.444691)		(0.582375)	
S268Q02T	0.207931	0.520	-0.417864	0.201	0.205569	0.688	0.172944	0.702	-0.970059	0.094
	(0.323697)		(0.326935)		(0.513385)		(0.452943)		(0.580373)	
S268Q06	-0.050828	0.860	-0.030740	0.917	0.002536	0.996	0.226507	0.592	-0.179881	0.744
	(0.286945)		(0.291997)		(0.478036)		(0.423417)		(0.552149)	
S326Q01	0.701407*	0.015	-0.180011	0.537	0.063099	0.899	-0.134970	0.759	-0.043802	0.942
	(0.287333)		(0.291852)		(0.495615)		(0.440318)		(0.600253)	
S326Q02	0.473213	0.110	0.110213	0.714	0.205966	0.670	0.043996	0.918	-0.417146	0.447
	(0.296283)		(0.300560)		(0.484680)		(0.427213)		(0.548823)	
S326Q03	0.112625	0.709	-0.662055*	0.029	1.064419*	0.029	-0.188002	0.672	0.005107	0.993
	(0.302188)		(0.303931)		(0.486579)		(0.444368)		(0.573419)	
S326Q04T	0.276047	0.406	-0.847117*	0.012	0.177199	0.736	-0.347785	0.462	-0.468520	0.421
	(0.331763)		(0.335845)		(0.525451)		(0.472923)		(0.582351)	
S408Q01	0.202802	0.477	-0.408682	0.159	0.554326	0.272	-0.533396	0.222	-0.479684	0.415
	(0.285454)		(0.289982)		(0.503626)		(0.436330)		(0.588775)	
S408Q03	0.059101	0.846	0.176143	0.569	0.126745	0.801	-0.546623	0.228	-0.855116	0.136
	(0.302927)		(0.309359)		(0.503595)		(0.453402)		(0.574395)	
S408Q04T	-0.076472	0.786	-0.345596	0.228	0.397426	0.417	-0.525798	0.215	-1.210586*	0.029
	(0.281950)		(0.286515)		(0.489272)		(0.423853)		(0.554207)	
S408Q05	-0.121735	0.673	-0.157284	0.592	0.242474	0.611	-0.171395	0.686	-0.438840	0.421
	(0.288660)		(0.293529)		(0.477481)		(0.425139)		(0.545374)	
S415Q02	0.105164	0.716	-0.117646	0.689	0.468651	0.370	-0.034060	0.940	-0.515167	0.407
	(0.289324)		(0.293998)		(0.522638)		(0.452717)		(0.621205)	
S415Q07T	0.535498	0.058	-0.033764	0.907	0.047636	0.921	-0.196403	0.641	-0.867676	0.113
	(0.282462)		(0.286951)		(0.479375)		(0.422058)		(0.548478)	
S415Q08T	0.329681	0.245	-0.349752	0.224	0.418731	0.387	-0.412850	0.333	-0.724173	0.192
	(0.283565)		(0.287459)		(0.483755)		(0.425680)		(0.555237)	
S437Q01	-0.486126	0.090	-0.308819	0.290	0.136555	0.775	-0.074308	0.862	-0.310741	0.572
	(0.286751)		(0.291303)		(0.478020)		(0.425552)		(0.551159)	
S437Q03	0.167849	0.557	-0.560907	0.052	0.369607	0.441	-0.281424	0.507	-0.822528	0.133
	(0.285723)		(0.289043)		(0.479255)		(0.424038)		(0.548076)	

ตารางที่ 48 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 3 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S437Q04	-0.046518	0.871	-0.122315	0.674	0.913326	0.062	-0.731238	0.092	-0.859617	0.122
	(0.286081)		(0.290989)		(0.489950)		(0.434961)		(0.555501)	
S437Q06	-0.331686	0.247	0.149592	0.608	0.349487	0.471	0.131040	0.762	-0.085207	0.883
	(0.286366)		(0.292063)		(0.485048)		(0.432786)		(0.577305)	
S477Q02	0.430599	0.180	0.238384	0.465	0.406166	0.502	-0.038191	0.941	-1.422439*	0.034
	(0.321170)		(0.325947)		(0.605548)		(0.510622)		(0.672363)	
S477Q03	0.349928	0.228	-0.059676	0.840	1.166583*	0.038	0.192325	0.677	-1.179372	0.065
	(0.289786)		(0.294557)		(0.564442)		(0.462445)		(0.640804)	
S477Q04	-0.208008	0.466	0.177352	0.541	0.710995	0.166	-0.186958	0.667	-1.250325*	0.031
	(0.285274)		(0.290105)		(0.513149)		(0.434618)		(0.581204)	
S498Q02T	0.213941	0.456	-0.156581	0.591	-0.807115	0.102	0.261402	0.538	-0.545179	0.326
	(0.287131)		(0.291559)		(0.494091)		(0.424815)		(0.554833)	
S498Q03	0.347657	0.223	-0.300961	0.298	0.522894	0.282	-0.106951	0.802	-0.838378	0.133
	(0.284909)		(0.288640)		(0.485073)		(0.425488)		(0.557939)	
S510Q01T	-0.249196	0.377	-0.250135	0.383	0.058601	0.904	-0.221327	0.601	-1.272669*	0.021
	(0.282172)		(0.286650)		(0.485018)		(0.423222)		(0.551296)	
S510Q04T	0.327803	0.274	-0.126257	0.677	0.013144	0.979	-0.423825	0.336	-0.001981	0.997
	(0.299635)		0.303202)		(0.494061)		(0.440119)		(0.552933)	
S519Q02T	0.055731	0.846	-0.359683	0.215	0.870333	0.076	-0.283742	0.506	-1.293460*	0.021
	(0.285895)		(0.289939)		(0.491152)		(0.426628)		(0.559079)	
S519Q03	0.139938	0.728	-0.468861	0.252	-0.225835	0.727	0.276150	0.615	-0.639705	0.364
	(0.403419)		(0.409042)		(0.646982)		(0.549204)		(0.704891)	
S524Q06T	0.434178	0.125	-0.413088	0.150	0.198626	0.679	-0.232764	0.581	-1.086684*	0.047
	(0.283185)		(0.287067)		(0.480382)		()0.421845		(0.548685)	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 4

ตารางที่ 49 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 4

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
Dichotomous Item										
S131Q02T	0.332877	0.271	-0.292639	0.333	0.409385	0.338	0.193397	0.676	-0.690826	0.238
	(0.302014)		(0.302277)		(0.426525)		(0.463781)		(0.585402)	
S131Q04T	0.287015	0.450	-0.087393	0.814	-0.789149	0.164	0.624587	0.257	-0.785756	0.290
	(0.380323)		(0.371343)		(0.567201)		(0.550094)		(0.741856)	
S256Q01	0.265620	0.458	-0.160030	0.670	0.592024	0.378	0.613116	0.386	-0.816225	0.375
	(0.357868)		(0.376568)		(0.670974)		(0.706693)		(0.920286)	
S326Q01	0.093297	0.753	-0.394040	0.198	0.151509	0.740	0.958857	0.067	-0.958605	0.135
	(0.296535)		(0.305740)		(0.456931)		(0.523552)		(0.640794)	
S326Q02	0.288994	0.339	-0.286761	0.343	0.223054	0.601	0.081135	0.862	-0.475617	0.415
	(0.301746)		(0.302355)		(0.426692)		(0.465170)		(0.583299)	
S326Q03	-0.368509	0.216	-0.106347	0.725	0.512651	0.228	0.837149	0.069	-0.919074	0.118
	(0.297941)		(0.303065)		(0.424879)		(0.461499)		(0.587765)	
S326Q04T	0.225843	0.528	-0.262515	0.456	0.568811	0.223	0.514175	0.303	-0.387073	0.525
	(0.358563)		(0.352311)		(0.466341)		0.498710)		(0.609901)	
S408Q01	-0.433544	0.134	0.104186	0.724	-0.319651	0.453	0.786741	0.096	-1.101190	0.060
	(0.289339)		(0.294899)		(0.426386)		(0.473410)		(0.587530)	
S408Q03	-0.200315	0.504	0.055563	0.856	-0.381375	0.384	0.057645	0.904	-1.109823	0.068
	(0.299657)		(0.305159)		(0.437645)		(0.473836)		(0.609402)	
S408Q04T	-0.274190	0.342	0.093365	0.750	-0.109082	0.797	-0.046710	0.919	-0.983362	0.089
	(0.288130)		(0.293658)		(0.424035)		(0.456577)		(0.579401)	
S408Q05	-0.075908	0.798	-0.384519	0.200	0.504257	0.239	0.046036	0.921	-1.512472*	0.011
	(0.296837)		(0.299532)		(0.427692)		(0.464321)		(0.593760)	
S413Q04T	-0.544299	0.071	0.013687	0.965	-0.159360	0.713	0.308657	0.508	-0.705796	0.231
	(0.301831)		(0.308623)		(0.433524)		(0.466530)		(0.588410)	
S413Q05	0.091125	0.755	0.215850	0.468	0.264150	0.548	0.734488	0.129	-0.845955	0.168
	(0.291754)		(0.297677)		(0.440583)		(0.483937)		(0.613165)	
S413Q06	0.010435	0.977	-0.139127	0.693	0.902245*	0.047	0.639550	0.194	-0.679361	0.263
	(0.353367)		(0.352999)		(0.455635)		(0.491796)		(0.606054)	
S415Q02	-0.649526*	0.030	0.294900	0.333	0.186702	0.691	0.780672	0.137	-0.355500	0.606
	(0.300493)		(0.304622)		(0.470040)		(0.525211)		(0.690930)	
S415Q07T	-0.440379	0.130	0.029089	0.922	0.163827	0.709	0.790681	0.101	-1.398913*	0.020
	(0.290743)		(0.296423)		(0.439226)		(0.482586)		0.598800)	
S415Q08T	-0.387632	0.182	0.050400	0.865	0.344327	0.424	0.668645	0.154	-0.828034	0.166
	(0.290276)		(0.296127)		(0.430296)		(0.468802)		(0.596988)	

ตารางที่ 49 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 4 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S437Q01	-0.807571**	0.006	-0.154760	0.603	0.247778	0.560	0.741764	0.110	-0.907406	0.123
	(0.291896)		(0.297729)		(0.425786)		(0.463869)		(0.588138)	
S437Q03	-0.302601	0.296	0.066964	0.820	0.247001	0.562	0.391518	0.395	-1.128660	0.053
	(0.288970)		(0.294588)		0.426136)		(0.459968)		(0.584009)	
S437Q04	-0.347851	0.231	-0.002931	0.992	0.098990	0.815	0.294675	0.521	-0.351065	0.551
	(0.289946)		(0.295550)		(0.421759)		(0.459185)		(0.588603)	
S437Q06	-0.498515	0.085	-0.181602	0.539	0.223076	0.606	0.791624	0.096	-1.108911	0.063
	(0.290190)		(0.295826)		(0.433551)		(0.476240)		(0.597975)	
S447Q02	0.369032	0.215	0.645170*	0.034	0.035700	0.934	0.436342	0.346	-0.838115	0.154
	(0.297098)		(0.304015)		(0.427291)		(0.462517)		(0.587593)	
S447Q03	-0.195078	0.501	0.270164	0.362	0.031402	0.942	0.367960	0.432	-0.028829	0.963
	(0.290069)		(0.296068)		(0.427934)		(0.467982)		(0.619485)	
S447Q04	-0.325541	0.265	-0.254129	0.392	-0.236474	0.576	0.208408	0.648	-0.776668	0.180
	(0.292025)		(0.296475)		(0.423156)		(0.457037)		(0.578509)	
S465Q02	-0.337407	0.245	0.000462	0.999	-0.448791	0.291	0.335442	0.464	0.092821	0.875
	(0.290209)		(0.295769)		(0.424245)		(0.458262)		(0.589424)	
S465Q04	-0.155223	0.618	0.181889	0.567	0.635145	0.142	0.039352	0.935	-1.077740	0.071
	(0.311456)		(0.318036)		(0.433053)		(0.479991)		0.597254)	
S478Q01	-0.087575	0.767	0.087016	0.772	-0.837581	0.058	0.229742	0.625	-1.248231*	0.043
	(0.295767)		(0.301024)		(0.442018)		(0.470044)		(0.619154)	
S478Q02T	-0.274625	0.356	0.045234	0.882	0.320740	0.448	0.553368	0.229	-0.394389	0.499
	(0.297310)		(0.303151)		(0.422371)		(0.459644)		(0.583860)	
S478Q03T	-0.289613	0.319	0.100893	0.733	0.040675	0.924	0.827739	0.078	-0.505157	0.399
	(0.290278)		(0.296195)		(0.425965)		(0.469871)		(0.598727)	
S510Q01T	-0.797902**	0.006	-0.107809	0.715	-0.133954	0.752	-0.660215	0.159	-0.848131	0.147
	(0.289457)		(0.295446)		(0.424046)		(0.468484)		(0.584583)	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 5

ตารางที่ 50 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 5

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
Dichotomous Item										
S131Q02T	0.373339	0.296	0.030706	0.935	-0.145165	0.773	0.864861	0.103	-0.420140	0.483
	(0.357100)		(0.376162)		(0.502663)		(0.531381)		(0.599629)	
S131Q04T	0.895847*	0.041	0.030220	0.946	-0.224741	0.702	-0.014381	0.982	-0.082875	0.905
	(0.438292)		(0.445090)		(0.587856)		(0.628233)		(0.689418)	
S213Q01T	0.763206*	0.041	-0.009687	0.980	-0.447727	0.386	0.620893	0.252	-0.236827	0.696
	(0.373603)		(0.389243)		(0.516061)		(0.541834)		(0.607790)	
S213Q02	0.100834	0.776	-0.219701	0.556	0.027647	0.957	0.684868	0.199	-0.668693	0.268
	(0.355235)		(0.373396)		(0.506841)		(0.532975)		(0.603552)	
S256Q01	0.850326*	0.031	0.154123	0.709	-0.771179	0.184	0.342827	0.576	0.724922	0.369
	(0.395294)		(0.412754)		(0.580330)		(0.613824)		(0.806645)	
S268Q01	0.347464	0.328	0.068612	0.855	-0.021003	0.967	0.120335	0.822	-0.161492	0.787
	(0.355096)		(0.373975)		(0.504383)		(0.533736)		(0.597884)	
S268Q02T	0.219416	0.572	0.114392	0.781	-0.208393	0.699	0.278899	0.627	-0.102522	0.872
	(0.389099)		(0.411969)		(0.538887)		(0.574785)		(0.633450)	
S268Q06	0.023269	0.949	-0.145093	0.701	-0.542043	0.283	0.428105	0.423	0.065147	0.913
	(0.358799)		(0.378010)		(0.504206)		(0.534570)		(0.595511)	
S304Q01	0.069353	0.846	-0.225240	0.549	-0.735678	0.144	0.659212	0.215	-0.297254	0.618
	(0.357440)		(0.375965)		(0.503970)		(0.531608)		(0.596536)	
S304Q02	-0.195204	0.583	-0.234274	0.531	-0.587832	0.242	0.820227	0.122	-0.064740	0.914
	(0.355696)		(0.373863)		(0.501458)		(0.531077)		(0.598499)	
S304Q03A	0.921201*	0.019	0.153564	0.706	0.021838	0.967	0.338418	0.545	0.310599	0.611
	(0.393295)		(0.407709)		(0.527455)		(0.560137)		(0.612046)	
S304Q03B	0.410320	0.269	-0.026294	0.947	-0.263999	0.604	0.815528	0.129	0.158576	0.791
	(0.370809)		(0.389763)		(0.510215)		(0.537504)		(0.599080)	
S413Q04T	0.486563	0.181	0.456375	0.237	-0.762045	0.138	0.804068	0.135	-0.164114	0.786
	(0.363715)		(0.385353)		(0.514525)		(0.537763)		(0.603556)	
S413Q05	0.468121	0.197	0.112618	0.767	0.449110	0.422	0.420183	0.452	-0.158542	0.810
	(0.362526)		(0.380035)		(0.559494)		(0.558606)		(0.659399)	
S413Q06	0.267945	0.489	-0.340443	0.400	0.285531	0.585	0.333212	0.553	-0.484452	0.438
	(0.386961)		(0.404751)		(0.522967)		(0.562662)		(0.624825)	
S416Q01	0.031128	0.932	-0.631422	0.098	-0.310408	0.539	0.227421	0.675	0.393018	0.511
	(0.364654)		(0.381871)		(0.506180)		(0.542662)		(0.598782)	
S421Q01	0.071524	0.849	-0.393828	0.317	-0.330197	0.523	0.538201	0.327	-0.229933	0.707
	(0.374330)		(0.392809)		(0.517490)		(0.548707)		(0.611482)	

ตารางที่ 50 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 5 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S421Q03	-0.668444	0.062	-0.233375	0.535	-0.299935	0.553	0.793895	0.138	0.180161	0.767
	(0.359083)		(0.376926)		(0.505700)		(0.535719)		(0.609098)	
S425Q02	0.215913	0.550	-0.520132	0.170	-0.018634	0.971	0.867990	0.104	-1.276377*	0.038
	(0.361583)		(0.378604)		(0.506352)		(0.535126)		(0.616331)	
S425Q03	0.390061	0.275	-0.093847	0.803	-0.198328	0.692	0.377504	0.478	-0.164931	0.782
	(0.356873)		(0.375187)		(0.501171)		(0.531761)		(0.595390)	
S425Q04	1.065644**	0.008	0.203669	0.620	0.158508	0.764	1.206390*	0.028	-0.791464	0.211
	(0.400064)		(0.411522)		(0.528386)		(0.550775)		(0.632465)	
S425Q05	0.564425	0.113	-0.220147	0.555	-0.408229	0.417	-0.027130	0.960	-0.555003	0.354
	(0.356081)		(0.373438)		(0.502535)		(0.534261)		(0.598317)	
S428Q01	-0.058819	0.871	-0.112165	0.768	-0.114057	0.821	0.610000	0.255	-0.541910	0.368
	(0.360860)		(0.380602)		(0.504340)		(0.535156)		(0.602022)	
S428Q03	0.475383	0.186	-0.212601	0.573	-0.168028	0.751	1.105414	0.050	0.385822	0.567
	(0.359371)		(0.377125)		(0.530241)		(0.564387)		(0.674961)	
S428Q05	0.143200	0.713	-0.503373	0.217	-0.163660	0.756	0.940571	0.089	-0.221501	0.720
	(0.390118)		(0.407579)		(0.526614)		(0.553663)		(0.618716)	
S438Q01T	0.338283	0.339	-0.282061	0.448	-0.600915	0.232	0.238787	0.652	-0.562888	0.344
	(0.353289)		(0.371358)		(0.501902)		(0.529981)		(0.594646)	
S438Q02	0.874467*	0.015	0.419131	0.268	0.032625	0.951	0.569002	0.300	-0.046180	0.942
	(0.359341)		(0.378209)		(0.525410)		(0.548256)		(0.632224)	
S438Q03T	0.905798*	0.013	-0.125849	0.741	0.136786	0.787	0.452286	0.399	-0.203505	0.735
	(0.365353)		(0.380739)		(0.505419)		(0.535888)		(0.600848)	
S447Q02	0.567733	0.109	-0.474226	0.203	-1.090079*	0.030	0.833623	0.117	-0.193179	0.746
	(0.354560)		(0.372113)		(0.503994)		(0.531890)		(0.596742)	
S447Q03	0.451783	0.204	-0.166427	0.656	-0.273876	0.593	0.554151	0.305	-0.288535	0.638
	(0.355823)		(0.373900)		(0.513612)		(0.539803)		(0.613695)	
S447Q04	0.320040	0.368	-0.155456	0.677	-0.907883	0.071	0.519425	0.328	-0.506807	0.397
	(0.355490)		(0.373522)		(0.503510)		(0.531055)		(0.597676)	
S458Q01	0.077023	0.852	-0.160230	0.712	0.377643	0.491	0.485674	0.409	-0.241916	0.707
	(0.412344)		(0.434528)		(0.548353)		(0.587778)		(0.643818)	
S458Q02T	0.559754	0.118	0.049991	0.895	-0.686602	0.174	0.550967	0.301	-0.477300	0.426
	(0.357988)		(0.376088)		(0.504464)		(0.532104)		(0.598807)	
S465Q02	0.218680	0.538	-0.706552	0.057	-1.034708*	0.040	0.914023	0.086	-0.025473	0.967
	(0.355141)		(0.372366)		(0.504067)		(0.533241)		(0.600858)	
S465Q04	0.222778	0.540	-0.125348	0.743	-0.918801	0.075	0.458166	0.397	0.305784	0.610
	(0.363901)		(0.382917)		(0.516842)		(0.540653)		(0.600885)	

ตารางที่ 50 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 5 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S466Q01T	0.484204	0.175	0.474514	0.210	0.000858	0.999	0.457141	0.392	-0.753939	0.210
	(0.357192)		(0.378015)		(0.506386)		(0.533812)		(0.601576)	
S466Q05	-0.030550	0.932	-0.021269	0.955	-0.425670	0.396	0.452970	0.393	-0.423575	0.476
	(0.354283)		(0.373193)		(0.500901)		(0.530237)		(0.595004)	
S466Q07T	1.072840**	0.003	0.109836	0.771	-0.178435	0.732	0.194664	0.720	-0.255950	0.680
	(0.359333)		(0.378124)		(0.522519)		(0.544635)		(0.621296)	
S477Q02	0.832494*	0.024	0.147424	0.704	-0.415434	0.446	0.445610	0.434	-0.074696	0.911
	(0.369555)		(0.388004)		(0.544676)		(0.569773)		(0.664152)	
S477Q03	0.377049	0.290	-0.037786	0.920	-0.074948	0.884	0.648781	0.229	0.198910	0.749
	(0.356243)		(0.374245)		(0.511808)		(0.538668)		(0.623191)	
S477Q04	0.801435*	0.026	0.170772	0.651	-0.586700	0.246	1.068391*	0.045	-0.237054	0.692
	(0.359254)		(0.377875)		(0.505776)		(0.533063)		(0.599500)	
S478Q01	0.254207	0.475	-0.231716	0.535	-0.910200	0.071	0.314155	0.555	-0.459218	0.443
	(0.355721)		(0.373734)		(0.504296)		(0.532613)		(0.599035)	
S478Q02T	0.375462	0.292	-0.310098	0.407	-0.086594	0.864	0.540060	0.311	-0.099698	0.869
	(0.356034)		(0.373448)		(0.503739)		(0.532505)		(0.603021)	
S478Q03T	0.558792	0.115	0.270428	0.469	-0.101589	0.842	0.402956	0.450	-0.701509	0.243
	(0.354617)		(0.373545)		(0.507915)		(0.533111)		(0.599926)	
S493Q01T	0.936449*	0.010	0.369146	0.334	0.032007	0.950	-0.029812	0.957	-0.558329	0.358
	(0.363729)		(0.381713)		(0.509739)		(0.542197)		(0.606574)	
S493Q03T	0.529149	0.169	0.180639	0.653	-0.284301	0.632	1.185851	0.064	-0.649537	0.356
	(0.384791)		(0.402142)		(0.594069)		(0.640670)		(0.703042)	
S493Q05T	0.100501	0.806	-0.280097	0.515	-0.049414	0.929	0.602655	0.303	-0.110820	0.864
	(0.409431)		(0.429982)		(0.552131)		(0.584443)		(0.643185)	
S498Q02T	0.092648	0.802	-0.208456	0.591	-0.988242	0.060	0.310128	0.573	0.521287	0.391
	(0.368406)		(0.387980)		(0.527081)		(0.550764)		(0.607479)	
S498Q03	0.220564	0.549	-0.174634	0.651	-0.789009	0.129	0.553440	0.310	-0.082430	0.893
	(0.368289)		(0.387024)		(0.519857)		(0.544098)		(0.608114)	
S514Q02	1.110259**	0.003	0.028236	0.942	-0.377429	0.489	1.529075*	0.012	-0.890100	0.173
	(0.365839)		(0.385382)		(0.545595)		(0.607001)		(0.652664)	
S514Q03	0.357544	0.319	0.074733	0.843	0.032538	0.949	0.330170	0.536	-0.628267	0.296
	(0.358214)		(0.377579)		(0.504445)		(0.534532)		(0.600469)	
S514Q04	0.627515	0.097	-0.148328	0.706	0.362520	0.479	0.633201	0.245	-0.514697	0.400
	(0.378122)		(0.394242)		(0.511902)		(0.544287)		(0.611328)	
S519Q02T	0.250257	0.492	-0.391577	0.305	-0.377750	0.459	0.200178	0.713	-0.764789	0.215
	(0.364198)		(0.381636)		(0.510405)		(0.544118)		(0.616487)	

ตารางที่ 50 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 5 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S519Q03	0.620247	0.222	-0.129954	0.801	0.326012	0.611	0.471305	0.489	-0.740022	0.344
	(0.507752)		(0.514536)		(0.642154)		(0.682029)		(0.781744)	
S524Q06T	0.750792*	0.041	-0.183969	0.630	-0.604449	0.237	0.636171	0.236	-0.079607	0.895
	(0.367160)		(0.382589)		(0.510661)		(0.535842)		(0.599953)	
Polytomous Item										
S498Q04	0.187791	0.434	-0.069431	0.781	-0.409120	0.236	-0.281094	0.421	-0.204342	0.638
	(0.240063)		(0.249383)		(0.345146)		(0.349104)		(0.434257)	

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 6

ตารางที่ 51 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 6

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
Dichotomous Item										
S304Q01	0.035661	0.907	0.431859	0.157	0.063603	0.885	0.091441	0.840	-0.270189	0.622
	(0.302691)		(0.305334)		(0.439990)		(0.452989)		(0.548468)	
S304Q02	-0.363794	0.232	0.224912	0.463	0.375730	0.411	0.063136	0.892	-0.532693	0.345
	(0.304403)		(0.306346)		(0.457302)		(0.462570)		(0.564003)	
S304Q03A	0.377592	0.237	0.608224	0.059	-0.227291	0.608	0.405007	0.372	-0.767045	0.162
	(0.319285)		(0.322049)		(0.443962)		(0.453262)		(0.547809)	
S304Q03B	0.308365	0.319	0.338177	0.277	-0.086068	0.845	1.044238*	0.023	-1.261325*	0.023
	(0.308906)		(0.310794)		(0.440245)		(0.458311)		(0.553746)	
S326Q01	0.690091*	0.026	0.310835	0.316	-0.110585	0.799	0.336191	0.453	-0.460064	0.391
	(0.310026)		(0.309940)		(0.434894)		(0.447808)		(0.536571)	
S326Q02	0.589998	0.081	0.073750	0.826	0.011399	0.980	0.456329	0.322	-0.679185	0.221
	(0.338282)		(0.334581)		(0.452968)		(0.460220)		(0.554494)	
S326Q03	0.209793	0.530	-0.105994	0.752	-0.133914	0.768	0.362713	0.432	-0.574798	0.298
	(0.334332)		(0.335071)		(0.454698)		(0.462007)		(0.551201)	
S326Q04T	0.238890	0.500	-0.109437	0.758	-0.253879	0.593	0.466001	0.328	-0.152570	0.785
	(0.354208)		(0.354915)		(0.476174)		(0.475706)		(0.558334)	
S408Q01	0.236751	0.435	0.320450	0.295	-0.530976	0.220	0.350649	0.431	-0.333460	0.531
	(0.303298)		(0.305525)		(0.432634)		(0.445648)		(0.532303)	
S408Q03	0.217583	0.522	0.550465	0.108	-0.798058	0.105	0.470274	0.336	-1.471710*	0.024
	(0.340041)		(0.343055)		(0.493205)		(0.488858)		(0.652576)	
S408Q04T	0.510725	0.090	0.488412	0.107	-0.591108	0.172	0.184854	0.678	-0.764134	0.150
	(0.301580)		(0.303690)		(0.432287)		(0.445850)		(0.531022)	

ตารางที่ 51 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 6 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S408Q05	0.287580	0.377	0.064690	0.843	0.183985	0.679	-0.076405	0.869	-0.883916	0.109
	(0.325169)		(0.325769)		(0.445371)		(0.463352)		(0.551808)	
S415Q02	0.279135	0.356	0.506498	0.096	-0.071174	0.871	0.249595	0.580	-0.502671	0.354
	(0.302250)		(0.304845)		(0.438209)		(0.452119)		(0.542402)	
S415Q07T	0.408068	0.184	0.230936	0.453	-0.316858	0.464	0.196132	0.659	-0.609693	0.252
	(0.306714)		(0.307905)		(0.433143)		(0.445245)		(0.531967)	
S415Q08T	0.397998	0.198	0.145249	0.639	0.169065	0.698	0.326617	0.467	-0.688273	0.203
	(0.309076)		(0.309850)		(0.435571)		(0.448831)		(0.539783)	
S428Q01	-0.260161	0.390	-0.010517	0.973	-0.150745	0.729	0.046910	0.917	-0.587290	0.275
	(0.302367)		(0.304517)		(0.434865)		(0.448803)		(0.536999)	
S428Q03	-0.005395	0.986	0.484598	0.120	0.252886	0.613	0.899328	0.087	-1.232614*	0.038
	(0.310376)		(0.312141)		(0.499961)		(0.526522)		(0.593685)	
S428Q05	-0.230800	0.481	0.136785	0.680	0.356100	0.426	0.392395	0.395	-1.247624*	0.027
	(0.327961)		(0.331895)		(0.446828)		(0.461541)		(0.563540)	
S437Q01	0.159680	0.602	0.167265	0.587	-0.499215	0.251	0.431282	0.333	-0.418694	0.431
	(0.306230)		(0.308171)		(0.434290)		(0.445237)		(0.531968)	
S437Q03	0.095758	0.760	0.255133	0.420	-0.564834	0.204	0.247691	0.583	-0.682526	0.209
	(0.313534)		(0.315985)		(0.444323)		(0.451817)		(0.542451)	
S437Q04	0.147813	0.632	0.151680	0.626	-0.533872	0.223	0.437647	0.327	-0.424014	0.426
	(0.309406)		(0.311313)		(0.437412)		(0.445821)		(0.532710)	
S437Q06	-0.280343	0.367	0.308629	0.327	-0.035370	0.936	0.354802	0.429	-0.423654	0.429
	(0.310393)		(0.314512)		(0.434592)		(0.448601)		(0.535569)	
S438Q01T	0.331550	0.273	0.675420*	0.027	-0.694061	0.112	-0.301196	0.504	-0.162220	0.766
	(0.302205)		(0.304655)		(0.437371)		(0.450464)		(0.545653)	
S438Q02	0.438639	0.156	0.991064**	0.002	-0.025947	0.956	0.055552	0.907	-1.072310	0.055
	(0.308980)		(0.311892)		(0.471285)		(0.474039)		(0.559330)	
S438Q03T	0.812536*	0.014	0.241232	0.459	0.521516	0.239	0.642978	0.157	-0.562804	0.307
	(0.330074)		(0.325622)		(0.442377)		(0.454557)		(0.550798)	
S458Q01	0.015948	0.966	-0.264194	0.479	0.247489	0.613	0.171930	0.734	-1.319037*	0.035
	(0.369500)		(0.373415)		(0.489505)		(0.506648)		(0.625190)	
S458Q02T	0.370366	0.240	0.290838	0.357	0.090996	0.836	0.455768	0.312	-1.182842*	0.031
	(0.314717)		(0.315742)		(0.438576)		(0.449966)		(0.549503)	
S466Q01T	0.356064	0.236	0.455550	0.132	-0.166710	0.702	-0.447809	0.316	-0.798874	0.133
	(0.300303)		(0.302627)		(0.436325)		(0.446282)		(0.531503)	
S466Q05	-0.045669	0.880	0.001666	0.996	0.008208	0.985	-0.441243	0.327	-1.000806	0.061
	(0.301985)		(0.304203)		(0.436206)		(0.449368)		(0.534942)	

ตารางที่ 51 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 6 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S466Q07T	0.605891	0.052	0.774375*	0.014	-0.610663	0.190	0.098217	0.841	0.202308	0.744
	(0.311931)		(0.314917)		(0.465245)		(0.488577)		(0.619309)	
S510Q01T	0.026811	0.929	0.300565	0.321	-0.859242*	0.047	0.252663	0.571	-0.975502	0.066
	(0.300434)		(0.302804)		(0.432728)		(0.445841)		(0.531941)	

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 7

ตารางที่ 52 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 7

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
Dichotomous Item										
S213Q01T	0.231135	0.325	-0.006805	0.977	-0.365498	0.236	-0.330665	0.334	0.580879	0.150
	(0.234653)		(0.236041)		(0.308455)		(0.342130)		(0.403827)	
S213Q02	-0.600435*	0.013	0.067981	0.764	-0.282780	0.344	-0.256503	0.463	1.270125**	0.002
	(0.239882)		(0.227017)		(0.298383)		(0.349779)		(0.399818)	
S416Q01	-0.641640**	0.007	0.377482	0.136	-0.347781	0.230	-0.408796	0.260	0.182993	0.621
	(0.234107)		(0.253020)		(0.289507)		(0.362893)		(0.370143)	
S421Q01	-0.642736*	0.019	0.005294	0.986	-0.629698	0.073	0.515423	0.172	0.992007*	0.014
	(0.274650)		(0.290641)		(0.351375)		(0.377502)		(0.403518)	
S421Q03	-0.301473	0.216	0.059238	0.807	-0.755112*	0.015	-0.034497	0.925	1.577747**	0.001
	(0.243257)		(0.242096)		(0.309552)		(0.366062)		(0.462032)	
S425Q02	-0.186070	0.445	0.508195*	0.049	0.023454	0.943	-0.337658	0.376	0.212000	0.635
	(0.243420)		(0.258761)		(0.325174)		(0.380927)		(0.447562)	
S425Q03	-0.034087	0.897	0.362831	0.163	-1.241348**	0.001	-0.253488	0.501	1.141211*	0.017
	(0.261492)		(0.260239)		(0.352549)		(0.377296)		(0.476567)	
S425Q04	0.293352	0.275	0.259566	0.328	-0.205921	0.584	-0.259326	0.512	0.358964	0.425
	(0.268687)		(0.265038)		(0.376911)		(0.395898)		(0.449351)	
S425Q05	-0.080263	0.751	0.365074	0.203	-0.374231	0.224	-0.522962	0.150	0.506189	0.272
	(0.252992)		(0.286670)		(0.307774)		(0.362996)		(0.460650)	
S493Q01T	-0.207887	0.389	0.683370**	0.009	-0.889397**	0.008	-0.335419	0.358	0.409594	0.366
	(0.241166)		(0.261090)		(0.332980)		(0.364668)		(0.452777)	
S493Q03T	-0.356500	0.233	0.165313	0.593	-0.875495*	0.035	-0.382284	0.392	1.356431*	0.028
	(0.298797)		(0.309613)		(0.414924)		(0.445912)		(0.616475)	
S493Q05T	-0.614358*	0.027	-0.173667	0.523	-0.459091	0.171	-0.352283	0.333	0.712082	0.106
	(0.277899)		(0.272263)		(0.335275)		(0.363189)		(0.440271)	

ตารางที่ 52 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 7 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	P	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S514Q02	-0.187111	0.510	0.274761	0.324	-1.049044**	0.003	-0.174873	0.703	1.029816	0.056
	(0.284247)		(0.278305)		(0.345056)		(0.459824)		(0.540403)	
S514Q03	-0.544399*	0.034	0.281598	0.209	-0.505238	0.081	-0.461800	0.189	0.758259*	0.049
	(0.257786)		(0.224027)		(0.289797)		(0.351456)		(0.385774)	

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 8

ตารางที่ 53 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 8

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
Dichotomous Item										
S114Q03T	0.155526	0.570	-0.389957	0.152	-0.095741	0.825	-0.652385	0.153	0.495831	0.368
	(0.274349)		(0.272025)		(0.431797)		(0.456411)		(0.549992)	
S114Q05T	-0.239062	0.448	-0.076533	0.815	-0.264450	0.576	-0.382018	0.447	0.664349	0.242
	(0.314718)		(0.326443)		(0.473680)		(0.502496)		(0.567251)	
S304Q01	0.107859	0.692	0.343377	0.255	0.136718	0.734	-1.225124*	0.024	0.534543	0.322
	(0.272703)		(0.301371)		(0.402470)		(0.544413)		(0.539149)	
S304Q02	-0.031072	0.911	-0.269994	0.323	0.477374	0.230	-1.697136**	0.001	0.589923	0.225
	(0.276464)		(0.273166)		(0.397545)		(0.470989)		(0.486298)	
S304Q03A	0.213677	0.498	0.093441	0.777	0.220988	0.619	-0.829915	0.136	0.462288	0.423
	(0.315624)		(0.330064)		(0.445222)		(0.556543)		(0.576450)	
S304Q03B	0.095711	0.742	0.080976	0.775	0.403927	0.265	-0.942626	0.074	0.392548	0.388
	(0.291541)		(0.282757)		(0.361708)		(0.528522)		(0.454029)	
S428Q01	-0.295516	0.299	0.090817	0.757	0.056501	0.905	-0.955386*	0.040	0.954081	0.053
	(0.284164)		(0.294119)		(0.470849)		(0.466449)		(0.494574)	
S428Q03	-0.069324	0.792	0.196569	0.452	0.151711	0.720	-0.695655	0.179	1.031387*	0.021
	(0.262834)		(0.261250)		(0.423170)		(0.517657)		(0.445142)	
S428Q05	-0.325135	0.355	-0.654388	0.072	0.022309	0.959	-0.803816	0.181	1.253666*	0.039
	(0.351215)		(0.364195)		(0.433236)		(0.600464)		(0.608200)	
S438Q01T	0.291884	0.290	0.334376	0.240	-0.634586	0.146	-1.141638*	0.020	0.637195	0.191
	(0.275372)		(0.284267)		(0.436202)		(0.490718)		(0.486823)	
S438Q02	0.442868	0.078	0.008082	0.974	-0.216365	0.535	-0.585069	0.183	0.436761	0.319
	(0.251581)		(0.247413)		(0.348646)		(0.439500)		(0.438119)	
S438Q03T	0.657067*	0.027	-0.014257	0.964	0.898281*	0.032	-1.322699*	0.015	0.149714	0.799
	(0.296567)		(0.315413)		(0.418711)		(0.542116)		(0.588187)	
S458Q01	-0.108110	0.756	-0.225978	0.559	0.462953	0.354	-0.110631	0.820	-0.620132	0.372
	(0.347874)		(0.386931)		(0.498640)		(0.485948)		(0.693707)	

ตารางที่ 53 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 8 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S458Q02T	0.387847	0.177	-0.023587	0.933	-0.524065	0.241	-0.759340	0.090	1.092955**	0.008
	(0.287294)		(0.276815)		(0.446874)		(0.448962)		(0.407031)	
S466Q01T	0.086713	0.751	0.127718	0.637	0.127345	0.768	-1.605490**	0.001	0.252190	0.609
	(0.273631)		(0.271369)		(0.431212)		(0.483802)		(0.494017)	
S466Q05	-0.419457	0.110	0.190300	0.523	-0.204615	0.682	-1.318653*	0.015	0.508236	0.292
	(0.262814)		(0.297998)		(0.500426)		(0.539159)		(0.481941)	
S466Q07T	0.699780**	0.009	0.002513	0.992	-0.112142	0.794	-1.045809*	0.032	0.576274	0.237
	(0.267474)		(0.248376)		(0.428366)		(0.489074)		(0.486440)	
S476Q01	0.114626	0.662	0.106244	0.690	-0.316887	0.472	-1.036942*	0.031	0.609477	0.235
	(0.262196)		(0.266403)		(0.440436)		(0.482490)		(0.512627)	
S476Q02	-0.370479	0.148	0.040615	0.877	0.073367	0.848	-1.024514*	0.038	0.090662	0.840
	(0.256341)		(0.261897)		(0.382120)		(0.493431)		(0.448654)	
S476Q03	-0.113879	0.689	-0.133953	0.641	0.793913	0.050	-1.065752*	0.027	-0.428569	0.435
	(0.284654)		(0.287899)		(0.404958)		(0.482167)		(0.548764)	
S495Q01T	0.019594	0.941	-0.162522	0.574	-0.278028	0.477	-0.487329	0.255	0.943927	0.066
	(0.262414)		(0.289548)		(0.391320)		(0.427334)		(0.513358)	
S495Q02T	-0.441600	0.102	0.005291	0.986	0.021121	0.953	-1.030411*	0.038	0.457948	0.312
	(0.270070)		(0.297499)		(0.353270)		(0.496976)		(0.452906)	
S495Q03	0.031021	0.904	-0.048107	0.842	0.391491	0.281	-1.094248*	0.018	0.799189	0.065
	(0.255330)		(0.240591)		(0.362654)		(0.461324)		(0.433789)	
S495Q04T	0.429151	0.139	-0.465810	0.125	-0.257949	0.544	-0.564802	0.250	0.095732	0.851
	(0.289858)		(0.303782)		(0.425632)		(0.490366)		(0.509121)	
S521Q02	0.096140	0.708	0.415089	0.158	-0.187640	0.671	-1.089987*	0.021	-0.088823	0.866
	(0.256568)		(0.294267)		(0.442925)		(0.473294)		(0.525166)	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 9

ตารางที่ 54 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 9

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
Dichotomous Item										
S268Q01	0.567324*	0.021	-0.031058	0.913	0.118671	0.778	0.134784	0.707	-0.404364	0.320
	(0.245436)		(0.283133)		(0.420156)		(0.358523)		(0.405968)	
S268Q02T	-0.045995	0.884	0.179237	0.599	0.275364	0.526	0.644526	0.118	-0.310931	0.540
	(0.314683)		(0.340815)		(0.434526)		(0.412426)		(0.507201)	
S268Q06	-0.161885	0.543	-0.019901	0.947	0.223353	0.599	-0.507965	0.219	0.804853	0.091
	(0.266535)		(0.297877)		(0.425082)		(0.412821)		(0.476597)	
S269Q01	-0.043244	0.860	0.264812	0.302	0.644042	0.075	-0.351732	0.332	0.392233	0.408
	(0.243776)		(0.256052)		(0.361874)		(0.362604)		(0.474359)	
S269Q03T	0.100467	0.724	-0.025241	0.923	0.336564	0.446	0.358158	0.329	0.256895	0.621
	(0.285063)		(0.261053)		(0.441747)		(0.366308)		(0.519898)	
S269Q04T	0.128161	0.678	-0.278614	0.412	0.082927	0.846	-0.249135	0.575	0.232585	0.615
	(0.309187)		(0.339194)		(0.425617)		(0.444888)		(0.462902)	
S426Q03	0.626197*	0.012	-0.010296	0.971	0.151890	0.705	0.125758	0.732	-0.569752	0.143
	(0.248643)		(0.280238)		(0.401865)		(0.367818)		(0.389311)	
S426Q05	-0.000898	0.997	0.247598	0.357	0.087944	0.832	-0.018702	0.961	0.403958	0.372
	(0.264699)		(0.268588)		(0.414206)		(0.383553)		(0.452583)	
S426Q07T	0.159641	0.525	0.057436	0.831	-0.265000	0.504	0.336133	0.395	0.056193	0.903
	(0.251080)		(0.268181)		(0.396987)		(0.395263)		(0.458448)	
S477Q02	0.865487**	0.003	0.068406	0.818	0.261124	0.586	0.319239	0.489	-0.165199	0.745
	(0.288006)		(0.297652)		(0.479659)		(0.461319)		(0.508950)	
S477Q03	0.659039*	0.013	-0.110458	0.679	0.835731	0.047	0.320037	0.378	0.176827	0.727
	(0.263076)		(0.267167)		(0.421896)		(0.363046)		(0.507076)	
S477Q04	0.497480	0.073	-0.066506	0.807	-0.001281	0.997	0.471329	0.202	-0.137943	0.744
	(0.277535)		(0.272778)		(0.377494)		(0.368838)		(0.422533)	
S485Q02	-0.012548	0.964	-0.242101	0.340	0.273280	0.461	0.222862	0.581	-0.160047	0.705
	(0.275311)		(0.253735)		(0.370844)		(0.404216)		(0.423463)	
S485Q03	0.515131	0.056	0.133648	0.632	1.001700*	0.018	0.347702	0.382	-0.779995	0.153
	(0.270032)		(0.279269)		(0.422266)		(0.397804)		(0.545937)	
S498Q02T	-0.307503	0.283	0.034872	0.900	-0.042832	0.906	0.082643	0.831	0.177286	0.707
	(0.286177)		(0.276706)		(0.359885)		(0.386090)		(0.472331)	
S498Q03	0.267873	0.288	0.405251	0.164	0.103543	0.814	0.027581	0.942	-0.076588	0.880
	(0.251751)		(0.290979)		(0.439521)		(0.378780)		(0.505517)	
S508Q02T	0.565603	0.052	0.135378	0.605	-0.143305	0.732	-0.028823	0.936	-0.262385	0.587
	(0.291223)		(0.261869)		(0.419126)		(0.354975)		(0.483915)	

ตารางที่ 54 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 9 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S508Q03	0.611851*	0.019	-0.147190	0.597	0.369054	0.359	0.876149*	0.022	-0.318217	0.473
	(0.260619)		(0.278893)		(0.402141)		(0.382920)		(0.443107)	
S519Q02T	-0.098506	0.718	-0.257865	0.374	-0.023062	0.959	-0.005863	0.989	-0.545549	0.265
	(0.273499)		(0.290138)		(0.448085)		(0.416193)		(0.488455)	
S519Q03	-0.165802	0.743	0.604141	0.161	1.106555	0.067	-0.729610	0.255	0.738082	0.275
	(0.506772)		(0.430729)		(0.604655)		(0.640905)		(0.675180)	
S524Q06T	0.784975**	0.004	-0.068201	0.824	0.349833	0.477	-0.334123	0.450	0.108232	0.853
	(0.270153)		(0.305693)		(0.492325)		(0.442291)		(0.581192)	
S524Q07	0.168468	0.625	0.392084	0.241	-0.333389	0.464	1.062209*	0.014	-0.083955	0.866
	(0.345433)		(0.333794)		(0.455509)		(0.432421)		(0.494430)	
S527Q01T	0.165936	0.669	-0.349514	0.338	0.614297	0.210	0.231169	0.647	-0.907086	0.140
	(0.389073)		(0.364286)		(0.489933)		(0.505852)		(0.614903)	
S527Q03T	-0.047407	0.862	0.235340	0.392	-0.058508	0.876	-0.060313	0.857	0.458528	0.319
	(0.272745)		(0.275048)		(0.374742)		(0.332930)		(0.459484)	

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 10

ตารางที่ 55 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 10

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
Dichotomous Item										
S269Q01	-0.012116	0.967	0.621592*	0.033	1.346823**	0.007	0.799533	0.079	0.326683	0.576
	(0.286608)		(0.292542)		(0.495296)		(0.455348)		(0.585260)	
S269Q03T	0.067193	0.827	0.513947	0.101	0.715820	0.119	0.574417	0.193	1.172277*	0.034
	(0.307902)		(0.313886)		(0.458852)		(0.441031)		(0.552110)	
S269Q04T	0.341260	0.381	0.037248	0.924	0.347421	0.540	-0.016491	0.977	-0.410101	0.610
	(0.389584)		(0.389287)		(0.567583)		(0.568758)		(0.804844)	
S304Q01	0.438295	0.127	0.582324*	0.046	0.548902	0.225	-0.097363	0.825	1.279111*	0.020
	(0.287564)		(0.292891)		(0.451937)		(0.439218)		(0.548374)	
S304Q02	-0.225003	0.431	0.715306*	0.014	0.574703	0.209	0.683778	0.118	0.820933	0.139
	(0.285365)		(0.291142)		(0.456634)		(0.437744)		(0.555387)	
S304Q03A	0.676780*	0.032	0.291752	0.357	-0.107170	0.823	0.486242	0.279	1.578529**	0.005
	(0.316795)		(0.316302)		(0.478085)		(0.448940)		(0.556727)	
S304Q03B	0.305888	0.309	0.429572	0.159	0.672405	0.138	0.620005	0.155	1.028757	0.062
	(0.300562)		(0.305330)		(0.453989)		(0.435930)		(0.552957)	

ตารางที่ 55 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 10 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S426Q03	0.832934**	0.004	0.303000	0.291	-0.147566	0.745	0.532055	0.227	1.224103*	0.029
	(0.282323)		(0.286783)		(0.454450)		(0.439885)		(0.560535)	
S426Q05	0.246809	0.382	0.470060	0.100	-0.053339	0.908	0.370718	0.401	1.324096*	0.021
	(0.282124)		(0.286456)		(0.457543)		(0.441630)		(0.573340)	
S426Q07T	0.472049	0.095	0.129572	0.650	-0.172043	0.701	0.312287	0.466	0.849222	0.117
	(0.282909)		(0.286039)		(0.448016)		(0.428249)		(0.542274)	
S428Q01	0.161801	0.595	0.229148	0.457	0.681428	0.140	-0.251780	0.581	1.032834	0.061
	(0.304679)		(0.308095)		(0.461902)		(0.456443)		(0.551330)	
S428Q03	0.364083	0.201	0.596448*	0.039	0.384063	0.420	0.536747	0.239	1.062616	0.069
	(0.284324)		(0.289252)		(0.475835)		(0.455715)		(0.586233)	
S428Q05	0.237794	0.486	-0.252595	0.460	1.000275*	0.037	0.398382	0.391	0.775244	0.179
	(0.341539)		(0.342018)		(0.479702)		(0.463985)		(0.576221)	
S438Q01T	0.277811	0.324	0.574882*	0.044	0.106689	0.815	-0.336130	0.442	1.688753**	0.004
	(0.281550)		(0.285963)		(0.455128)		(0.437455)		(0.572915)	
S438Q02	0.622144*	0.027	0.629754*	0.028	0.191042	0.678	0.390928	0.377	0.920265	0.101
	(0.282043)		(0.287227)		(0.461057)		(0.441763)		(0.562412)	
S438Q03T	0.646701*	0.028	0.170263	0.564	0.019655	0.966	0.332835	0.443	0.677160	0.221
	(0.294131)		(0.295710)		(0.455723)		(0.434207)		(0.553300)	
S458Q01	0.096384	0.786	-0.667819	0.063	0.854881	0.081	0.681273	0.146	0.799953	0.173
	(0.354980)		(0.360309)		(0.491311)		(0.468187)		(0.586812)	
S458Q02T	0.728537*	0.013	0.399906	0.178	0.401099	0.375	0.554602	0.203	1.277797*	0.021
	(0.293126)		(0.296578)		(0.451777)		(0.435187)		(0.553925)	
S466Q01T	0.394960	0.163	-0.143268	0.615	0.307333	0.496	0.066229	0.879	0.781145	0.155
	(0.282751)		(0.285476)		(0.452064)		(0.432625)		(0.548694)	
S466Q05	-0.045071	0.874	0.477303	0.098	0.425021	0.348	0.600324	0.164	0.349543	0.525
	(0.284256)		(0.288960)		(0.452495)		(0.431239)		(0.550607)	
S466Q07T	0.674917*	0.018	0.264062	0.361	0.401258	0.402	0.455765	0.315	0.520265	0.363
	(0.283769)		(0.288966)		(0.478940)		(0.453430)		(0.571197)	
S485Q02	-0.040143	0.906	0.232803	0.493	0.398129	0.419	0.401747	0.392	1.732159**	0.002
	(0.336565)		(0.339941)		(0.492051)		(0.468688)		(0.560060)	
S485Q03	0.132425	0.640	0.573448*	0.046	0.365399	0.431	0.662325	0.140	1.336039*	0.022
	(0.283493)		(0.288491)		(0.463981)		(0.448742)		(0.582608)	
S508Q02T	0.828308**	0.007	0.354846	0.247	0.335811	0.467	0.296722	0.502	0.606822	0.283
	(0.305927)		(0.306020)		(0.462134)		(0.442411)		(0.564067)	
S508Q03	0.424972	0.130	0.618418*	0.030	0.039499	0.931	0.508364	0.243	0.888303	0.107
	(0.280937)		(0.285930)		(0.452095)		(0.434891)		(0.551001)	

ตารางที่ 55 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 10 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S527Q01T	0.090417	0.814	0.924071*	0.020	0.799037	0.146	0.430724	0.413	0.661774	0.304
	(0.383039)		(0.398647)		(0.549875)		(0.526097)		(0.642836)	
S527Q03T	0.101161	0.719	0.107622	0.705	-0.164691	0.713	0.243800	0.568	0.698209	0.197
	(0.281373)		(0.284834)		(0.447471)		(0.427702)		(0.541208)	

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 11

ตารางที่ 56 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 11

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
Dichotomous Item										
S114Q03T	0.215908	0.454	-0.048357	0.880	-0.769281	0.066	0.298220	0.497	-0.131323	0.766
	(0.288390)		(0.319308)		(0.418681)		(0.439616)		(0.441675)	
S114Q05T	0.338183	0.453	0.435923	0.337	0.096304	0.870	-1.112667	0.084	0.220785	0.775
	(0.451020)		(0.453744)		(0.585426)		(0.644962)		(0.770939)	
S131Q02T	0.350826	0.167	-0.077094	0.782	-0.618503	0.070	-0.317115	0.364	0.636079	0.195
	(0.254108)		(0.278415)		(0.341312)		(0.349251)		(0.490596)	
S131Q04T	-0.103227	0.773	-0.276215	0.426	-0.309803	0.556	-0.010572	0.985	-0.093738	0.894
	(0.358073)		(0.346804)		(0.527034)		(0.561605)		(0.702757)	
S256Q01	0.470920	0.130	0.211210	0.487	-0.573847	0.326	2.302031**	0.004	-0.316666	0.668
	(0.311136)		(0.304301)		(0.584084)		(0.782818)		(0.739390)	
S413Q04T	-0.299417	0.272	0.045995	0.869	-0.325276	0.469	0.066853	0.879	-0.163258	0.759
	(0.272229)		(0.277857)		(0.449172)		(0.437315)		(0.532498)	
S413Q05	0.232188	0.355	0.066693	0.803	0.019702	0.964	0.448771	0.347	0.274195	0.603
	(0.251090)		(0.266608)		(0.437575)		(0.477071)		(0.528336)	
S413Q06	-0.492531	0.107	-0.731662*	0.016	0.223173	0.608	0.363546	0.422	0.542452	0.282
	(0.306147)		(0.303606)		(0.435260)		(0.452379)		(0.503503)	
S447Q02	0.269641	0.326	-0.015012	0.957	-0.711355	0.082	0.296001	0.504	-0.430624	0.370
	(0.274225)		(0.279076)		(0.408914)		(0.443468)		(0.479547)	
S447Q03	-0.183932	0.455	-0.216089	0.375	-0.539060	0.210	0.052482	0.896	0.550348	0.244
	(0.246248)		(0.243657)		(0.430160)		(0.400299)		(0.472167)	
S447Q04	-0.483426	0.060	0.193568	0.481	-1.083954**	0.009	-0.123525	0.783	0.375438	0.469
	(0.257363)		(0.274890)		(0.411988)		(0.449363)		(0.518060)	
S465Q02	-0.246384	0.348	-0.164248	0.553	-0.631681	0.129	-0.187013	0.664	0.437970	0.390
	(0.262104)		(0.276905)		(0.416278)		(0.431376)		(0.508946)	

ตารางที่ 56 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 11 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S465Q04	-0.064781	0.812	-0.139654	0.584	-0.478370	0.188	0.195939	0.613	-0.049754	0.920
	(0.271920)		(0.255519)		(0.363404)		(0.387943)		(0.494305)	
S476Q01	-0.046625	0.843	-0.127559	0.589	-1.048248**	0.007	-0.172073	0.645	0.366307	0.368
	(0.234374)		(0.236522)		(0.388460)		(0.374242)		(0.406613)	
S476Q02	-0.223136	0.332	-0.182574	0.461	-1.006435*	0.011	-0.168440	0.660	0.386158	0.374
	(0.229899)		(0.247756)		(0.393886)		(0.383631)		(0.433759)	
S476Q03	-0.539622*	0.038	-0.144916	0.603	-0.591608	0.127	0.208042	0.569	-0.508654	0.273
	(0.259828)		(0.278896)		(0.387960)		(0.365471)		(0.463776)	
S478Q01	0.026355	0.921	-0.157878	0.569	-1.013697*	0.023	0.396857	0.377	0.019846	0.966
	(0.265349)		(0.277334)		(0.446040)		(0.449343)		(0.454867)	
S478Q02T	-0.310720	0.198	0.077260	0.769	0.026739	0.941	-0.005031	0.990	0.347311	0.428
	(0.241063)		(0.262716)		(0.356331)		(0.412861)		(0.438188)	
S478Q03T	0.286872	0.270	0.024046	0.928	-0.435062	0.323	0.094756	0.835	-0.219753	0.662
	(0.259718)		(0.262844)		(0.440133)		(0.453462)		(0.503485)	
S495Q01T	0.310448	0.247	0.012486	0.962	-0.368406	0.294	0.095981	0.773	-0.025147	0.954
	(0.267706)		(0.260757)		(0.351064)		(0.332520)		(0.430695)	
S495Q02T	-0.462438	0.057	0.007278	0.976	-0.586627	0.113	-0.423288	0.242	0.452891	0.326
	(0.242921)		(0.241376)		(0.370828)		(0.361342)		(0.461123)	
S495Q03	-0.217222	0.400	0.295929	0.301	0.466941	0.244	-0.157460	0.676	-0.012032	0.978
	(0.257930)		(0.285561)		(0.400381)		(0.377365)		(0.425789)	
S495Q04T	-0.045843	0.872	-0.222907	0.430	-0.681456	0.093	0.355145	0.366	-0.014648	0.975
	(0.284705)		(0.282166)		(0.406151)		(0.392909)		(0.466595)	
S521Q02	-0.278874	0.243	0.027917	0.912	-0.504400	0.194	-0.141213	0.712	-0.283704	0.570
	(0.238611)		(0.252080)		(0.388323)		(0.383368)		(0.499419)	
Polytomous Item										
S114Q04T	-0.099279	0.725	-0.087611	0.754	-0.750131	0.052	0.427086	0.282	0.185258	0.714
	(0.281960)		(0.280093)		(0.386841)		(0.396394)		(0.506353)	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 12

ตารางที่ 57 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 12

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
Dichotomous Item										
S131Q02T	0.607581*	0.012	0.111899	0.655	0.654510*	0.040	0.209217	0.511	0.299997	0.501
	(0.240551)		(0.250927)		(0.318623)		(0.318301)		(0.446223)	
S131Q04T	0.696332	0.064	-0.282676	0.483	0.680312	0.093	-0.291563	0.529	0.363054	0.470
	(0.376750)		(0.402860)		(0.405121)		(0.464024)		(0.502974)	
S256Q01	0.720335**	0.007	0.199437	0.513	0.112334	0.799	0.728148	0.173	1.682134*	0.029
	(0.266478)		(0.305235)		(0.440339)		(0.534101)		(0.771073)	
S269Q01	-0.204465	0.409	0.740191**	0.006	1.138628**	0.001	0.295962	0.398	1.055489*	0.023
	(0.247655)		(0.267829)		(0.330827)		(0.349830)		(0.464969)	
S269Q03T	0.241850	0.344	0.134796	0.636	1.173572**	0.001	-0.049789	0.882	0.853294	0.072
	(0.255205)		(0.285105)		(0.340110)		(0.334268)		(0.475450)	
S269Q04T	0.210979	0.427	0.545425	0.060	1.011630**	0.004	-0.420588	0.264	-0.343074	0.478
	(0.265520)		(0.290003)		(0.344128)		(0.376318)		(0.484045)	
S413Q04T	0.256126	0.339	-0.080425	0.774	0.085657	0.843	-0.149489	0.724	-0.389267	0.488
	(0.267549)		(0.280276)		(0.432234)		(0.424575)		(0.561473)	
S413Q05	0.137113	0.541	0.111746	0.666	0.882912**	0.008	-0.356520	0.311	1.162431*	0.024
	(0.224496)		(0.259058)		(0.331431)		(0.351484)		(0.513758)	
S413Q06	0.353022	0.244	-0.116853	0.700	1.123118**	0.004	0.014903	0.970	1.055566*	0.031
	(0.302540)		(0.303719)		(0.390549)		(0.392797)		(0.490996)	
S426Q03	0.886811**	0.001	0.294620	0.278	0.172307	0.684	-0.110512	0.787	0.772284	0.103
	(0.256252)		(0.271029)		(0.424532)		(0.409214)		(0.473650)	
S426Q05	0.581318*	0.014	0.151552	0.529	0.135510	0.665	0.294952	0.365	0.539004	0.193
	(0.237106)		(0.240872)		(0.313024)		(0.325415)		(0.413701)	
S426Q07T	0.548371*	0.020	0.602366*	0.016	0.334896	0.319	0.282693	0.360	0.562528	0.145
	(0.235282)		(0.248790)		(0.335555)		(0.308298)		(0.386245)	
S447Q02	0.646977*	0.012	0.380272	0.189	0.470010	0.111	0.282919	0.373	0.461342	0.265
	(0.257586)		(0.289432)		(0.294897)		(0.317627)		(0.413778)	
S447Q03	0.716017**	0.005	-0.089239	0.746	0.356329	0.326	-0.233238	0.509	0.967056*	0.047
	(0.254778)		(0.275958)		(0.362644)		(0.353005)		(0.487011)	
S447Q04	0.158698	0.549	0.308017	0.285	-0.242580	0.545	-0.179240	0.631	0.522708	0.218
	(0.265384)		(0.287916)		(0.400974)		(0.373412)		(0.423594)	
S465Q02	0.064754	0.790	-0.051825	0.851	0.584404	0.084	-0.239831	0.536	-0.001998	0.996
	(0.242692)		(0.275925)		(0.338549)		(0.388227)		(0.449720)	
S465Q04	-0.041407	0.885	0.300102	0.305	0.371353	0.358	-0.013315	0.974	0.138080	0.771
	(0.284854)		(0.292457)		(0.403854)		(0.403383)		(0.474997)	

ตารางที่ 57 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 12 (ต่อ)

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S478Q01	0.319644	0.191	0.336199	0.230	0.611156	0.125	-0.512836	0.188	0.380253	0.342
	(0.244435)		(0.279562)		(0.398716)		(0.389464)		(0.400157)	
S478Q02T	0.422591	0.083	-0.005558	0.984	1.107640**	0.004	-0.055921	0.862	0.642992	0.152
	(0.244047)		(0.269747)		(0.378740)		(0.320504)		(0.448838)	
S478Q03T	0.726525**	0.003	0.060355	0.828	0.568698	0.067	-0.322629	0.364	0.671423*	0.044
	(0.242372)		(0.278118)		(0.310940)		(0.355174)		(0.334041)	
S485Q02	0.549909*	0.024	0.417059	0.112	0.737481	0.052	-0.352327	0.342	1.060373*	0.017
	(0.243222)		(0.262810)		(0.380151)		(0.370699)		(0.444072)	
S485Q03	0.512836*	0.036	0.147447	0.573	1.011212**	0.006	0.262523	0.472	1.469884	0.009
	(0.244740)		(0.261538)		(0.367771)		(0.364838)		(0.560698)	
S508Q02T	0.972172**	0.000	0.193894	0.507	0.102795	0.785	0.122966	0.746	0.590613	0.137
	(0.256499)		(0.292558)		(0.376574)		(0.380547)		(0.397595)	
S508Q03	0.567910*	0.020	0.332024	0.211	0.709151*	0.046	0.071059	0.843	0.612308	0.220
	(0.244340)		(0.265482)		(0.355874)		(0.358197)		(0.498372)	
S527Q01T	0.397626	0.182	0.532365	0.114	0.929913	0.061	0.037553	0.934	1.182285*	0.016
	(0.297563)		(0.337112)		(0.497407)		(0.448359)		(0.489345)	
S527Q03T	0.536205*	0.037	0.258256	0.272	0.411203	0.220	-0.256587	0.598	-0.256587	0.598
	(0.257788)		(0.234983)		(0.334780)		(0.487604)		(0.487604)	
Polytomous Item										
S465Q01	-0.075568	0.781	0.162036	0.513	0.181852	0.598	0.215012	0.583	-0.676111	0.116
	(0.271980)		(0.247967)		(0.345040)		(0.391976)		(0.430044)	

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบฉบับที่ 13

ตารางที่ 58 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 13

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
Dichotomous Item										
S213Q01T	0.046291	0.879	0.046812	0.863	-0.288558	0.504	-0.220469	0.593	-0.443145	0.377
	(0.303204)		(0.270670)		(0.431935)		(0.413403)		(0.500867)	
S213Q02	-0.991516**	0.001	-0.153643	0.573	0.139812	0.718	-0.295288	0.407	0.024368	0.960
	(0.283578)		(0.272856)		(0.387438)		(0.356142)		(0.483095)	
S416Q01	-0.912365**	0.002	-0.129289	0.596	-0.215689	0.569	-0.317086	0.426	-0.291705	0.529
	(0.288774)		(0.244365)		(0.379138)		(0.397943)		(0.463440)	

ตารางที่ 58 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบฉบับที่ 13

ข้อสอบ	MALE		EXTRA		ECSC		HEDERS		WEALTH	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
S421Q01	-0.410688	0.195	-0.329478	0.266	-0.276754	0.523	-0.050506	0.910	-0.231452	0.646
	(0.316784)		(0.296155)		(0.433676)		(0.442926)		(0.505299)	
S421Q03	-0.998964**	0.000	0.294465	0.238	-0.044813	0.905	0.247684	0.502	-0.120506	0.798
	(0.263051)		(0.249576)		(0.372342)		(0.368671)		(0.469641)	
S425Q02	-0.487693	0.077	0.022843	0.931	-0.304687	0.449	-0.126737	0.738	-0.644741	0.101
	(0.276062)		(0.261942)		(0.402804)		(0.378714)		(0.393110)	
S425Q03	-0.180897	0.475	0.318395	0.237	-0.371305	0.318	0.003934	0.992	-0.875632*	0.049
	(0.253529)		(0.269006)		(0.371601)		(0.371196)		(0.445308)	
S425Q04	0.191917	0.513	0.204078	0.455	-0.569194	0.170	-0.133153	0.742	0.081347	0.865
	(0.293562)		(0.273419)		(0.414419)		(0.404672)		(0.476390)	
S425Q05	-0.317257	0.312	0.040832	0.893	-0.072268	0.864	-0.925200*	0.016	0.032063	0.952
	(0.313686)		(0.302376)		(0.419984)		(0.383028)		(0.525161)	
S493Q01T	-0.485420	0.081	0.221348	0.435	-0.989614*	0.023	-0.085679	0.835	-0.294047	0.535
	(0.278787)		(0.283774)		(0.434337)		(0.410628)		(0.474337)	
S493Q03T	0.318174	0.300	0.315801	0.318	-0.275704	0.610	-0.641347	0.210	-0.615199	0.280
	(0.306395)		(0.315916)		(0.541305)		(0.511699)		(0.568726)	
S493Q05T	-0.787569**	0.009	0.362996	0.240	0.293253	0.501	0.532223	0.208	-0.529643	0.230
	(0.298678)		(0.308637)		(0.435796)		(0.422803)		(0.440485)	
S514Q02	0.026294	0.931	0.236071	0.370	-0.330387	0.426	-0.225027	0.610	-0.460420	0.365
	(0.302896)		(0.263208)		(0.415173)		(0.441543)		(0.507526)	
S514Q03	-0.375623	0.185	0.217315	0.381	-0.599007	0.112	-0.036525	0.925	0.080181	0.875
	(0.283413)		(0.247810)		(0.377002)		(0.384053)		(0.508611)	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สรุปข้อสอบที่ทำหน้าที่ต่างกันในแต่ละคุณลักษณะของตัวแปร

ตารางที่ 59 สรุปข้อสอบที่ทำหน้าที่ต่างกันฉบับที่ 1 ถึง 13 ในแต่ละคุณลักษณะของนักเรียน

ฉบับ	รหัสข้อสอบ	เพศ		การเรียนพิเศษ		เศรษฐานะครอบครัว		ทรัพยากรที่บ้าน		ความมั่งคั่ง		สมรรถนะที่วัด
		ชาย	หญิง	เรียน	ไม่เรียน	สูง	ต่ำ	สูง	ต่ำ	สูง	ต่ำ	
1	S114Q03T	✓										US
	S114Q05T					✓						ES
	S213Q01T	✓				✓						**IS**
	S269Q01					✓						**ES**
	S269Q04T					✓						ES
	S326Q01	✓										US
	S326Q02	✓										US
	S326Q03	✓										US
	S326Q04T	✓				✓						**ES**
	S408Q01	✓										ES
	S408Q03	✓										ES
	S415Q02	✓										ES
	S415Q07T	✓				✓						**IS**
	S415Q08T	✓										IS
	S416Q01					✓				✓		**US**
	S421Q01					✓						ES
	S421Q03					✓						**ES**
	S425Q02					✓						**US**
	S425Q04	✓										US
	S426Q03	✓				✓				✓		**ES**
	S426Q07T	✓										IS
	S476Q03									✓		**ES**
	S485Q02					✓						ES
	S485Q03	✓										US
	S493Q01T	✓										ES
	S493Q03T	✓		✓								**ES**
	S493Q05T					✓						ES
	S495Q01T	✓										US
	S495Q03	✓				✓						**US**
	S495Q04T	✓										IS
	S508Q02T	✓								✓		IS
	S508Q03	✓										IS
	S510Q04T					✓				✓		**ES**
	S514Q03					✓						ES

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 59 สรุปข้อสอบที่ทำหน้าที่ต่างกันฉบับที่ 1 ถึง 13 ในแต่ละคุณลักษณะของนักเรียน(ต่อ)

ฉบับ	รหัสข้อสอบ	เพศ		การเรียนพิเศษ		เศรษฐานะครอบครัว		ทรัพยากรที่บ้าน		ความมั่งคั่ง		สมรรถนะที่วัด
		ชาย	หญิง	เรียน	ไม่เรียน	สูง	ต่ำ	สูง	ต่ำ	สูง	ต่ำ	
1	S521Q06	✓										ES
	S527Q03T								✓			ES
	S114Q04T*						✓					US
2	**S114Q05T**					✓						**ES**
	S498Q03										✓	**IS**
	S519Q02T		✓									**ES**
	S114Q04T*						✓					US
3	S326Q01	✓										US
	S326Q03			✓		✓						**US**
	S326Q04T			✓								ES
	S408Q04T										✓	ES
	S477Q02										✓	**ES**
	S477Q03					✓						**ES**
	S519Q02T										✓	**ES**
4	S408Q05										✓	IS
	S413Q06					✓						**ES**
	S415Q02		✓									**ES**
	S447Q02			✓								**IS**
	S478Q01										✓	**ES**
	S510Q01T		✓									**ES**
5	S131Q04T	✓										IS
	S213Q01T	✓										IS
	S256Q01	✓										**ES**
	S304Q03A	✓										**US**
	S425Q02										✓	**US**
	S425Q04	✓						✓				**US**
	S438Q02	✓										IS
	S438Q03T	✓										**IS**
	S447Q02						✓					**IS**
	S465Q02						✓					**ES**
	S466Q07T	✓										**IS**
	S477Q02	✓										**ES**
	S477Q04	✓						✓				**ES**
	S493Q01T	✓										**ES**
	S514Q02	✓						✓				**US**

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 59 สรุปข้อสอบที่ทำหน้าที่ต่างกันฉบับที่ 1 ถึง 13 ในแต่ละคุณลักษณะของนักเรียน(ต่อ)

ฉบับ	รหัสข้อสอบ	เพศ		การเรียนพิเศษ		เศรษฐานะครอบครัว		ทรัพยากรที่บ้าน		ความมั่งคั่ง		สมรรถนะที่วัด
		ชาย	หญิง	เรียน	ไม่เรียน	สูง	ต่ำ	สูง	ต่ำ	สูง	ต่ำ	
5	S524Q06T	✓										US
	S447Q05*						✓		✓			US
6	**S304Q03B**							✓			✓	**ES**
	S326Q01	✓										US
	S408Q03										✓	**ES**
	S428Q03										✓	**US**
	S428Q05										✓	ES
	S438Q01T			✓								IS
	S438Q02			✓								**IS**
	S438Q03T	✓										**IS**
	S458Q01										✓	**ES**
	S458Q02T										✓	**US**
	S466Q07T			✓								IS
	S510Q01T						✓					ES
7	S213Q02		✓							✓		ES
	S416Q01		✓									US
	S421Q01		✓							✓		ES
	S421Q03						✓			✓		ES
	S425Q02			✓								US
	S425Q03						✓			✓		**ES**
	S493Q01T			✓			✓					ES
	S493Q03T						✓			✓		**ES**
	S493Q05T		✓									ES
	S514Q02						✓					**US**
	S514Q03		✓							✓		**ES**
8	S304Q01								✓			US
	S304Q02								✓			**ES**
	S428Q01								✓			US
	S428Q03									✓		US
	S428Q05									✓		**ES**
	S438Q01T								✓			IS
	S438Q03T	✓				✓			✓			**IS**
	S458Q02T									✓		**US**
	S466Q01T								✓			IS
	S466Q05								✓			**US**

ตารางที่ 59 สรุปข้อสอบที่ทำหน้าที่ต่างกันฉบับที่ 1 ถึง 13 ในแต่ละคุณลักษณะของนักเรียน(ต่อ)

ฉบับ	รหัสข้อสอบ	เพศ		การเรียนพิเศษ		เศรษฐานะครอบครัว		ทรัพยากรที่บ้าน		ความมั่งคั่ง		สมรรถนะที่วัด
		ชาย	หญิง	เรียน	ไม่เรียน	สูง	ต่ำ	สูง	ต่ำ	สูง	ต่ำ	
8	S466Q07T	✓							✓			IS
	S476Q01								✓			ES
	S476Q02								✓			ES
	S476Q03								✓			ES
	S495Q02T								✓			US
	S495Q03								✓			US
	S521Q02								✓			ES
9	S268Q01	✓										IS
	S426Q03	✓										**ES**
	S477Q02	✓										**IS**
	S477Q03	✓										IS
	S485Q03					✓						**US**
	S508Q03	✓						✓				**IS**
	S524Q06T	✓										US
	S524Q07							✓				US
10	**S269Q01**			✓		✓						**ES**
	S269Q03T									✓		ES
	S304Q01			✓						✓		**US**
	S304Q02			✓								ES
	S304Q03A	✓								✓		US
	S426Q03	✓								✓		**ES**
	S426Q05									✓		ES
	S428Q03			✓								US
	S428Q05					✓						ES
	S438Q01T			✓						✓		**IS**
	S438Q02	✓		✓								**IS**
	S438Q03T	✓										IS
	S458Q02T	✓								✓		US
	S466Q07T	✓										IS
	S485Q02									✓		**ES**
	S485Q03			✓						✓		US
	S508Q02T	✓										IS
	S508Q03			✓								IS
	S527Q01T			✓								US

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 59 สรุปข้อสอบที่ทำหน้าที่ต่างกันฉบับที่ 1 ถึง 13 ในแต่ละคุณลักษณะของนักเรียน (ต่อ)

ฉบับ	รหัสข้อสอบ	เพศ		การเรียนพิเศษ		เศรษฐานะครอบครัว		ทรัพยากรที่บ้าน		ความมั่งคั่ง		สมรรถนะที่วัด
		ชาย	หญิง	เรียน	ไม่เรียน	สูง	ต่ำ	สูง	ต่ำ	สูง	ต่ำ	
11	**S256Q01**							✓				**ES**
	S413Q06				✓							ES
	S447Q04						✓					**IS**
	S476Q01						✓					**ES**
	S476Q03		✓									ES
	S478Q01						✓					ES
12	S131Q02T	✓				✓						US
	S256Q01	✓								✓		**ES**
	S269Q01			✓		✓				✓		**ES**
	S269Q03T					✓						ES
	S269Q04T					✓						ES
	S413Q05					✓				✓		US
	S413Q06					✓				✓		**ES**
	S426Q03	✓										ES
	S426Q05	✓										ES
	S426Q07T	✓		✓								IS
	S447Q02	✓										IS
	S447Q03	✓								✓		IS
	S478Q02T					✓						US
	S478Q03T	✓								✓		ES
	S485Q02	✓								✓		ES
	S485Q03	✓				✓				✓		**US**
	S508Q02T	✓										IS
	S508Q03	✓				✓						**IS**
	S527Q01T									✓		US
	S527Q03T	✓										ES
13	S213Q02		✓									ES
	S416Q01		✓									**US**
	S421Q03		✓									**ES**
	S425Q03										✓	ES
	S425Q05								✓			IS
	S493Q01T						✓					**ES**
	S493Q05T		✓									ES

หมายเหตุ ตัวหนังสือหนาเข้ม เป็นข้อสอบที่ถูกตัดออกไป

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ผลการวิเคราะห์อิทธิพลของตัวแปรระดับสถานศึกษาต่อโอกาสการตอบข้อสอบถูก ด้วยการวิเคราะห์ HGLM 3L

ตารางที่ 60 ผลการวิเคราะห์อิทธิพลของตัวแปรระดับสถานศึกษาที่มีผลต่อโอกาสการตอบข้อสอบถูกในข้อสอบที่ทำหน้าที่ต่างกันแต่ละฉบับ

ฉบับ	รหัสข้อสอบ	ตัวแปรระดับนักเรียนที่มีผลต่อ DIF	ตัวแปรระดับสถานศึกษา	Coef.	S.E.	p-value
1	Reference Item		สถานศึกษาของรัฐ	0.714236	0.152086	0.000
			งบประมาณของรัฐ	-0.007119	0.001727	0.000
	S213Q01T	เพศชาย เศรษฐานะสูง	แรงกดดันของผู้ปกครอง	-0.362581	0.180097	0.044
	S269Q01	เศรษฐานะสูง	-			
	S326Q04T	เพศชาย เศรษฐานะสูง	ดัชนีการขาดแคลนครู	-1.344051	0.495848	0.007
	S415Q07T	เพศชาย เศรษฐานะสูง	-			
	S416Q01	เศรษฐานะสูง มั่งคั่งต่ำ				
	S421Q03	เศรษฐานะสูง	-			
	S425Q02	เศรษฐานะสูง	สถานศึกษาขนาดใหญ่	0.739948	0.295930	0.013
			ดัชนีการจัดกิจกรรมเรียนรู้	0.565193	0.260777	0.030
	S426Q03	เพศชาย เศรษฐานะสูง มั่งคั่งต่ำ	สถานศึกษาขนาดกลาง	1.269309	0.464142	0.007
	S476Q03	มั่งคั่งต่ำ	สถานศึกษาตั้งในเมืองใหญ่	-1.058946	0.432822	0.015
	S493Q03T	เพศชาย เรียนพิเศษ	สถานศึกษาตั้งในเมืองใหญ่	1.210042	0.513814	0.019
	S495Q03	เพศชาย เศรษฐานะสูง	งบประมาณของรัฐ	-0.013530	0.005636	0.017
	S510Q04T	เศรษฐานะสูง มั่งคั่งต่ำ	-			
2	Reference Item		งบประมาณของรัฐ	-0.006702	0.001864	0.001
	S114Q05T	เศรษฐานะสูง	-			
	S498Q03	มั่งคั่งต่ำ	ดัชนีทรัพยากรการเรียน	0.273898	0.131910	0.038
	S519Q02T	เพศหญิง	ดัชนีทรัพยากรการเรียน	0.316919	0.138743	0.022

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 60 ผลการวิเคราะห์อิทธิพลของตัวแปรระดับสถานศึกษาที่มีผลต่อโอกาสการตอบข้อสอบถูกในข้อสอบที่ทำหน้าที่ต่างกันแต่ละฉบับ(ต่อ)

ฉบับ	รหัสข้อสอบ	ตัวแปรระดับนักเรียนที่มีผลต่อ DIF	ตัวแปรระดับสถานศึกษา	Coef.	S.E.	p-value
3	Reference Item		สถานศึกษาของรัฐ	0.470583	0.187812	0.013
			สถานศึกษาตั้งในหมู่บ้าน	-0.593757	0.205323	0.005
			สถานศึกษาตั้งในเมืองเล็ก	-0.675016	0.185489	0.001
			งบประมาณของรัฐ	-0.004850	0.002139	0.024
	S326Q03	ไม่เรียนพิเศษ เศรษฐานะสูง	สถานศึกษาของรัฐ	1.359637	0.558331	0.015
	S477Q02	มั่งคั่งต่ำ	-			
	S477Q03	เศรษฐานะสูง	-			
	S519Q02T	มั่งคั่งต่ำ	-			
4	Reference Item		สถานศึกษาของรัฐ	0.608525	0.167996	0.001
			งบประมาณของรัฐ	-0.004030	0.001897	0.035
	S413Q06	เศรษฐานะสูง	-			
	S415Q02	เพศหญิง	สถานศึกษาตั้งในเมือง	-1.019313	0.446498	0.022
	S447Q02	เรียนพิเศษ	ดัชนีทรัพยากรการเรียน	0.279892	0.139089	0.044
	S478Q01	มั่งคั่งต่ำ	สถานศึกษาของรัฐ	-1.051190	0.468298	0.025
	S510Q01T	เพศหญิง	-			
5	Reference Item		สถานศึกษาขนาดกลาง	-0.604076	0.177171	0.001
			สถานศึกษาของรัฐ	0.613210	0.163594	0.000
			ดัชนีทรัพยากรการเรียน	0.142593	0.048346	0.004
	S256Q01	เพศชาย				
	S304Q03A	เพศชาย	-			
	S425Q02	มั่งคั่งต่ำ	สถานศึกษาตั้งในหมู่บ้าน	-1.396597	0.538664	0.010
	S425Q04	เพศชาย ทรัพยากรสูง	สถานศึกษาตั้งในเมือง	-1.350606	0.507336	0.008
	S438Q03T	เพศชาย				
	S447Q02	เศรษฐานะต่ำ	-			
	S465Q02	เศรษฐานะต่ำ	-			
	S466Q07T	เพศชาย	-			
	S477Q02	เพศชาย	สถานศึกษาขนาดกลาง	1.156313	0.518404	0.026
	S477Q04	เพศชาย ทรัพยากรสูง	ดัชนีการขาดแคลนครู	-0.725305	0.327249	0.027
5	**S493Q01T**	เพศชาย	ดัชนีการจัดกิจกรรมเรียนรู้	0.522311	0.250154	0.037
	S514Q02	เพศชาย ทรัพยากรสูง	-			

ตารางที่ 60 ผลการวิเคราะห์อิทธิพลของตัวแปรระดับสถานศึกษาที่มีผลต่อโอกาสการตอบข้อสอบถูกในข้อสอบที่ทำหน้าที่ต่างกันแต่ละฉบับ (ต่อ)

ฉบับ	รหัสข้อสอบ	ตัวแปรระดับนักเรียนที่มีผลต่อ DIF	ตัวแปรระดับสถานศึกษา	Coef.	S.E.	p-value
6	Reference Item		สถานศึกษาของรัฐ	0.360950	0.177590	0.043
			สถานศึกษาตั้งในเมือง	-0.362764	0.178305	0.043
	S304Q03B	ทรัพยากรสูง มั่งคั่งต่ำ				
	S408Q03	มั่งคั่งต่ำ	ดัชนีการจัดกิจกรรมเรียนรู้	0.736803	0.368894	0.045
	S428Q03	มั่งคั่งต่ำ	-			
	S438Q02	เรียนพิเศษ	-			
	S438Q03T	เพศชาย	-			
	S458Q01	มั่งคั่งต่ำ	สถานศึกษาของรัฐ	-1.631519	0.597905	0.007
			สถานศึกษาตั้งในเมืองใหญ	1.697638	0.593875	0.005
	S458Q02T	มั่งคั่งต่ำ	-			
7	Reference Item		สถานศึกษาของรัฐ	0.577300	0.208407	0.007
			สถานศึกษาตั้งในเมือง	-0.548081	0.207788	0.009
			ดัชนีการขาดแคลนครู	-0.289852	0.146070	0.048
	S425Q03	เศรษฐานะต่ำ มั่งคั่งสูง	สถานศึกษาขนาดกลาง	-1.128901	0.539185	0.036
	S493Q03T	เศรษฐานะต่ำ มั่งคั่งสูง	-			
	S514Q02	เศรษฐานะต่ำ	-			
8	Reference Item		สถานศึกษาของรัฐ	0.643762	0.191872	0.001
	S304Q02	ทรัพยากรสูง				
			ดัชนีการขาดแคลนครู	-0.834788	0.339723	0.014
	S428Q05	มั่งคั่งสูง	-			
	S438Q03T	เพศชาย เศรษฐานะสูง ทรัพยากรต่ำ				
	S458Q02T	มั่งคั่งสูง	สถานศึกษาของรัฐ	1.019660	0.503295	0.042
	S466Q05	ทรัพยากรต่ำ				
9	**Reference Item**		แรงกดดันของผู้ปกครอง	-0.128594	0.056727	0.025
			งบประมาณของรัฐ	-0.006725	0.001785	0.000
			ดัชนีการจัดกิจกรรมเรียนรู้	0.158035	0.079163	0.047
	S426Q03	เพศชาย	-			
	S477Q02	เพศชาย	-			
	S485Q03	เศรษฐานะสูง	สถานศึกษาของรัฐ	-1.063787	0.497679	0.032
	S508Q03	เพศชาย ทรัพยากรสูง	-			

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 60 ผลการวิเคราะห์อิทธิพลของตัวแปรระดับสถานศึกษาที่มีผลต่อโอกาสการตอบข้อสอบถูกในข้อสอบที่ทำหน้าที่ต่างกันแต่ละฉบับ (ต่อ)

ฉบับ	รหัสข้อสอบ	ตัวแปรระดับนักเรียนที่มีผลต่อ DIF	ตัวแปรระดับสถานศึกษา	Coef.	S.E.	p-value
10	Reference Item		สถานศึกษาของรัฐ	0.636109	0.201627	0.002
			สถานศึกษาตั้งในเมือง	-0.457866	0.205042	0.027
	S269Q01	เรียนพิเศษ เศรษฐานะสูง	-			
	S304Q01	เรียนพิเศษ มั่งคั่งสูง	สถานศึกษาตั้งในเมืองใหญ่	0.981477	0.445551	0.027
			สถานศึกษาตั้งในเมืองใหญ่มาก	1.167489	0.445979	0.009
	S426Q03	เพศชาย มั่งคั่งสูง	-			
	S438Q01T	เรียนพิเศษ มั่งคั่งสูง	ดัชนีทรัพยากรการเรียน	0.405034	0.135590	0.003
	S438Q02	เพศชาย เรียนพิเศษ				
	S458Q02T	เพศชาย มั่งคั่งสูง	-			
	S485Q02	มั่งคั่งสูง	ดัชนีการขาดแคลนครู	-0.972388	0.482918	0.044
11	Reference Item		-			
	S256Q01	ทรัพยากรสูง	สถานศึกษาตั้งในเมือง	1.246983	0.545698	0.022
			สถานศึกษาตั้งในเมืองใหญ่	1.325971	0.537665	0.014
	S447Q04	เศรษฐานะต่ำ	-			
	S476Q01	เศรษฐานะต่ำ	แรงกดดันของผู้ปกครอง	-0.296744	0.150264	0.048
12	Reference Item		สถานศึกษาของรัฐ	0.481062	0.176409	0.007
	S256Q01	เพศชาย มั่งคั่งสูง	-			
	S269Q01	เรียนพิเศษ เศรษฐานะสูง มั่งคั่งสูง	-			
	S413Q06	เศรษฐานะสูง มั่งคั่งสูง	-			
	S485Q03	เพศชาย เศรษฐานะสูง มั่งคั่งสูง	-			
	S508Q03	เพศชาย เศรษฐานะสูง	-			
13	Reference Item		สถานศึกษาของรัฐ	0.637670	0.207454	0.003
	S416Q01	เพศหญิง	-			
	S421Q03	เพศหญิง	ดัชนีการจัดกิจกรรมเรียนรู้	0.639923	0.240796	0.008
	S493Q01T	เศรษฐานะต่ำ	-			

ภาคผนวก ค

ผลการประมาณค่าความยากของข้อสอบจากการวิเคราะห์ด้วยโปรแกรม HGLM-3 ของข้อสอบในแบบสอบฉบับที่ 1 - 13

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 61 ผลการประมาณค่าความยากของข้อสอบจากการวิเคราะห์ด้วยโปรแกรม HGLM-3 จำแนกตามฉบับข้อสอบจำนวน 13 ฉบับ

ฉบับที่	ข้อสอบ	δ_i	S.E
1	S114Q03T	1.049768	0.130018
	S114Q05T	1.973651	0.160310
	S213Q01T	1.038935	0.129710
	S213Q02	0.364038	0.126682
	S269Q01	-0.018954	0.113631
	S269Q03T	0.933451	0.127440
	S269Q04T	1.954423	0.165861
	S326Q01	-0.214149	0.124403
	S326Q02	0.975051	0.123728
	S326Q03	0.802991	0.122847
	S326Q04T	1.775772	0.149247
	S408Q01	-0.154434	0.130649
	S408Q03	1.149909	0.138413
	S408Q04T	0.057004	0.131452
	S408Q05	0.832508	0.131505
	S415Q02	-0.369927	0.127077
	S415Q07T	-0.035842	0.126944
	S415Q08T	0.133058	0.133674
	S416Q01	0.852372	0.123613
	S421Q01	1.115975	0.135488
	S421Q03	0.469278	0.122428
	S425Q02	0.832508	0.135971
	S425Q03	0.933451	0.136167
	S425Q04	1.775772	0.147681
	S425Q05	0.522786	0.131689
	S426Q03	-0.396304	0.136525
	S426Q05	-0.387496	0.119898
	S426Q07T	-0.145932	0.125837
	S437Q01	0.513821	0.143111
	S437Q03	0.487041	0.120509
	S437Q04	0.346692	0.120322
	S437Q06	0.175433	0.123373
	S476Q01	-0.530573	0.124778

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 61 ผลการประมาณค่าความยากของข้อสอบจากการวิเคราะห์ด้วยโปรแกรม HGLM-3 จำแนกตามฉบับข้อสอบจำนวน 13 ฉบับ (ต่อ)

ฉบับที่	ข้อสอบ	δ_i	S.E
1	S476Q02	-0.188511	0.117213
	S476Q03	0.933451	0.139069
	S485Q02	1.093668	0.123293
	S485Q03	-0.576375	0.121946
	S493Q01T	0.416390	0.128121
	S493Q03T	-1.161613	0.137570
	S493Q05T	1.775772	0.158621
	S495Q01T	0.687951	0.135475
	S495Q02T	0.381434	0.130249
	S495Q03	0.832508	0.117126
	S495Q04T	0.783492	0.131936
	S508Q02T	0.243532	0.124597
	S508Q03	-0.111980	0.118721
	S510Q01T	-0.035842	0.125713
	S510Q04T	0.862363	0.128905
	S514Q02	-0.317567	0.121314
	S514Q03	0.622793	0.118340
	S514Q04	1.196093	0.131700
	S521Q02	0.303533	0.116787
	S521Q06	-0.641553	0.133829
	S527Q01T	1.792636	0.150946
	S527Q03T	0.346692	0.129653
	S527Q04T	-0.613587	0.099905
2	S114Q03T	1.010858	0.126171
	S114Q05T	1.803315	0.170074
	S268Q01	-0.403054	0.129352
	S268Q02T	1.387018	0.13736
	S268Q06	0.507083	0.133494
	S476Q01	-0.447340	0.137228
	S476Q02	-0.042605	0.137323
	S476Q03	0.835773	0.14475
	S477Q02	-1.420448	0.150189
	S477Q03	-0.695685	0.140684
	S477Q04	0.109373	0.126827
	S495Q01T	0.444837	0.12781

ตารางที่ 61 ผลการประมาณค่าความยากของข้อสอบจากการวิเคราะห์ด้วยโปรแกรม HGLM-3 จำแนกตามฉบับข้อสอบจำนวน 13 ฉบับ (ต่อ)

ฉบับที่	ข้อสอบ	δ_i	S.E
2	S495Q02T	0.392130	0.135488
	S495Q03	0.709599	0.127921
	S495Q04T	0.937174	0.13815
	S498Q02T	0.507083	0.136786
	S498Q03	0.357281	0.130098
	S519Q02T	0.662496	0.135869
	S519Q03	2.608134	0.193015
	S521Q02	0.271002	0.133408
	S521Q06	-0.863129	0.138681
	S524Q06T	0.245315	0.126635
	S524Q07	-0.937174	0.110184
3	S268Q01	-0.690652	0.137248
	S268Q02T	1.488958	0.145963
	S268Q06	0.392716	0.128968
	S326Q01	-0.395121	0.134971
	S326Q02	0.763512	0.138965
	S326Q03	0.664754	0.129757
	S326Q04T	1.591055	0.146911
	S408Q01	-0.298031	0.12752
	S408Q03	1.133601	0.142433
	S408Q04T	-0.081963	0.137214
	S408Q05	0.582655	0.126936
	S415Q02	-0.586313	0.135092
	S415Q07T	0.012124	0.138135
	S415Q08T	0.071971	0.128012
	S437Q01	0.392716	0.129947
	S437Q03	0.357390	0.130125
	S437Q04	0.357390	0.13007
	S437Q06	0.089085	0.120749
	S477Q02	-1.507635	0.158027
	S477Q03	-0.503261	0.129576
	S477Q04	-0.324348	0.126603
	S498Q02T	0.573405	0.136673
	S498Q03	0.157679	0.135907
	S510Q01T	-0.176467	0.138837

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 61 ผลการประมาณค่าความยากของข้อสอบจากการวิเคราะห์ด้วยโปรแกรม HGLM-3 จำแนกตามฉบับข้อสอบจำนวน 13 ฉบับ (ต่อ)

ฉบับที่	ข้อสอบ	δ_i	S.E
3	S510Q04T	0.936242	0.13405
	S519Q02T	0.340190	0.127911
	S519Q03	2.438143	0.181635
	S524Q06T	0.138722	0.12574
	S524Q07	-0.814264	0.111446
4	S131Q02T	0.637359	0.120028
	S131Q04T	2.019146	0.183360
	S256Q01	-1.962271	0.152866
	S326Q01	-0.593233	0.139863
	S326Q02	0.675432	0.124887
	S326Q03	0.618488	0.125799
	S326Q04T	1.740848	0.156845
	S408Q01	-0.216177	0.11586
	S408Q03	0.842989	0.1384
	S408Q04T	-0.181473	0.130002
	S408Q05	0.609093	0.14243
	S413Q04T	0.957500	0.145571
	S413Q05	-0.303575	0.140381
	S413Q06	1.707616	0.149055
	S415Q02	-0.717580	0.13442
	S415Q07T	-0.277247	0.127629
	S415Q08T	0.059446	0.124823
	S437Q01	0.337051	0.126149
	S437Q03	0.085224	0.141893
	S437Q04	0.214674	0.113535
	S437Q06	-0.060764	0.123858
	S447Q02	0.434902	0.133458
	S447Q03	-0.112384	0.130076
	S447Q04	0.516467	0.132482
	S465Q02	0.293110	0.13279
	S465Q04	1.100821	0.138638
	S478Q01	0.637359	0.13987
	S478Q02T	0.618488	0.124974
	S478Q03T	0.076629	0.135331
	S510Q01T	0.284357	0.133302

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 61 ผลการประมาณค่าความยากของข้อสอบจากการวิเคราะห์ด้วยโปรแกรม HGLM-3 จำแนกตามฉบับข้อสอบจำนวน 13 ฉบับ (ต่อ)

ฉบับที่	ข้อสอบ	δ_i	S.E
4	S510Q04T	-0.904767	0.100127
5	S131Q02T	0.353008	0.162653
	S131Q04T	2.069487	0.198173
	S213Q01T	1.042449	0.18082
	S213Q02	0.145724	0.175312
	S256Q01	-1.676421	0.191686
	S268Q01	0.205733	0.159188
	S268Q02T	1.552429	0.173775
	S268Q06	0.650746	0.172897
	S304Q01	0.585758	0.171417
	S304Q02	0.266058	0.166875
	S304Q03A	1.374197	0.174102
	S304Q03B	0.977998	0.153685
	S413Q04T	0.775016	0.186462
	S413Q05	-0.704505	0.165653
	S413Q06	1.440462	0.166107
	S416Q01	0.824268	0.169277
	S421Q01	1.212897	0.16635
	S421Q03	0.248784	0.171239
	S425Q02	0.678984	0.17725
	S425Q03	0.423431	0.16173
	S425Q04	1.387242	0.172269
	S425Q05	0.405741	0.181089
	S428Q01	0.726614	0.179054
	S428Q03	-0.442964	0.158162
	S428Q05	1.481537	0.173818
	S438Q01T	-0.153393	0.186228
	S438Q02	-0.461033	0.171332
	S438Q03T	0.622746	0.161625
	S447Q02	0.154281	0.172113
	S447Q03	-0.362539	0.171047
	S447Q04	0.441181	0.18369
	S458Q01	1.875091	0.190124
	S458Q02T	0.530954	0.181529
	S465Q02	0.120074	0.181584

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 61 ผลการประมาณค่าความยากของข้อสอบจากการวิเคราะห์ด้วยโปรแกรม HGLM-3 จำแนกตามฉบับข้อสอบจำนวน 13 ฉบับ (ต่อ)

ฉบับที่	ข้อสอบ	δ_i	S.E
5	S465Q04	0.884609	0.174167
	S466Q01T	0.353008	0.180989
	S466Q05	0.274708	0.169447
	S466Q07T	-0.580572	0.169891
	S477Q02	-1.082322	0.176024
	S477Q03	-0.119042	0.163611
	S477Q04	0.388110	0.172966
	S478Q01	0.485842	0.190879
	S478Q02T	0.188560	0.172154
	S478Q03T	-0.076224	0.179105
	S493Q01T	0.613463	0.175447
	S493Q03T	-1.463000	0.208257
	S493Q05T	1.857107	0.17953
	S498Q02T	1.064366	0.177137
	S498Q03	1.053379	0.181522
	S514Q02	-0.875338	0.174057
	S514Q03	0.530954	0.16599
	S514Q04	1.108903	0.170958
	S519Q02T	0.905048	0.179172
	S519Q03	2.713777	0.220298
	S524Q06T	0.844224	0.161556
	S524Q07	-2.069487	0.160155
6	S304Q01	0.120973	0.140875
	S304Q02	-0.019910	0.130449
	S304Q03A	1.005736	0.134656
	S304Q03B	0.511648	0.13144
	S326Q01	0.529281	0.130411
	S326Q02	1.335097	0.143164
	S326Q03	1.385544	0.134548
	S326Q04T	1.695365	0.15009
	S408Q01	0.373027	0.128489
	S408Q03	1.532776	0.162834
	S408Q04T	0.195887	0.15363
	S408Q05	1.191879	0.140792
	S415Q02	0.079493	0.125992

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 61 ผลการประมาณค่าความยากของข้อสอบจากการวิเคราะห์ด้วยโปรแกรม HGLM-3 จำแนกตามฉบับข้อสอบจำนวน 13 ฉบับ (ต่อ)

ฉบับที่	ข้อสอบ	δ_i	S.E
6	S415Q07T	0.564784	0.157251
	S415Q08T	0.546992	0.133524
	S428Q01	0.254490	0.136693
	S428Q03	-0.638745	0.136652
	S428Q05	1.274034	0.143391
	S437Q01	0.573712	0.13838
	S437Q03	0.953979	0.154693
	S437Q04	0.738564	0.136357
	S437Q06	0.757436	0.123456
	S438Q01T	-0.389829	0.152815
	S438Q02	-0.713291	0.142198
	S438Q03T	0.933602	0.139799
	S458Q01	1.970359	0.170117
	S458Q02T	0.834242	0.138728
	S466Q01T	0.004938	0.146654
	S466Q05	0.338942	0.133639
	S466Q07T	-0.868689	0.154932
	S510Q01T	0.246095	0.149713
	S510Q04T	-1.335097	0.127628
7	S213Q01T	0.932677	0.102081
	S213Q02	0.213054	0.106111
	S416Q01	0.422324	0.113393
	S421Q01	1.087460	0.123673
	S421Q03	-0.033339	0.113091
	S425Q02	0.499531	0.117812
	S425Q03	0.122470	0.114962
	S425Q04	1.066173	0.120413
	S425Q05	0.296039	0.125263
	S493Q01T	-0.198255	0.120872
	S493Q03T	-1.753256	0.150418
	S493Q05T	1.546286	0.129396
	S514Q02	-1.387141	0.139298
	S514Q03	-0.066179	0.10268
	S514Q04	-0.508191	0.102584

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 61 ผลการประมาณค่าความยากของข้อสอบจากการวิเคราะห์ด้วยโปรแกรม HGLM-3 จำแนกตามฉบับข้อสอบจำนวน 13 ฉบับ (ต่อ)

ฉบับที่	ข้อสอบ	δ_i	S.E
8	S114Q05T	1.905112	0.153721
	S114Q03T	1.215761	0.135193
	S304Q01	1.124128	0.148368
	S304Q02	0.623391	0.12375
	S304Q03A	1.851336	0.143169
	S304Q03B	1.467505	0.131681
	S428Q01	1.157983	0.138772
	S428Q03	-0.11412	0.117823
	S428Q05	2.101745	0.166492
	S438Q01T	0.418454	0.128808
	S438Q02	0.003231	0.113187
	S438Q03T	1.192438	0.135262
	S458Q01	2.386167	0.174815
	S458Q02T	0.831624	0.132339
	S466Q01T	0.568898	0.125528
	S466Q05	0.669473	0.129897
	S466Q07T	-0.03863	0.11744
	S476Q01	-0.40471	0.121906
	S476Q02	-0.21542	0.123904
	S476Q03	0.921626	0.130249
	S495Q01T	0.861264	0.117355
	S495Q02T	0.323526	0.130516
	S495Q03	0.891258	0.113184
	S495Q04T	1.287549	0.145869
	S521Q02	0.170879	0.124343
	S521Q06	0.657514	0.099611
9	S268Q01	-0.365845	0.124240
	S268Q02T	1.791059	0.151202
	S268Q06	0.546937	0.129011
	S269Q01	0.036439	0.128593
	S269Q03T	0.675869	0.133224
	S269Q04T	1.878589	0.178053
	S426Q03	-0.252812	0.135086
	S426Q05	-0.278715	0.121655
	S426Q07T	-0.065089	0.128289

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 61 ผลการประมาณค่าความยากของข้อสอบจากการวิเคราะห์ด้วยโปรแกรม HGLM-3 จำแนกตามฉบับข้อสอบจำนวน 13 ฉบับ (ต่อ)

ฉบับที่	ข้อสอบ	δ_i	S.E
9	S477Q02	-1.305132	0.136077
	S477Q03	-0.296040	0.130126
	S477Q04	0.404568	0.134324
	S485Q02	1.102629	0.137909
	S485Q03	-0.418812	0.136179
	S498Q02T	0.537900	0.131609
	S498Q03	0.510910	0.128564
	S508Q02T	0.343540	0.129105
	S508Q03	0.078732	0.125859
	S519Q02T	0.583300	0.141446
	S519Q03	2.866956	0.240359
	S524Q06T	0.565075	0.146127
	S524Q07	1.791059	0.165628
	S527Q01T	2.093691	0.182274
	S527Q03T	0.257343	0.126730
	S527Q04T	-0.790678	0.102060
10	S269Q01	0.058661	0.142545
	S269Q03T	1.117740	0.151697
	S269Q04T	2.355651	0.18989
	S304Q01	0.524460	0.143091
	S304Q02	0.253738	0.14526
	S304Q03A	1.293831	0.168162
	S304Q03B	0.924890	0.153039
	S426Q03	-0.051083	0.141046
	S426Q05	-0.255399	0.136103
	S426Q07T	0.400550	0.133929
	S428Q01	1.151682	0.152333
	S428Q03	-0.377047	0.144054
	S428Q05	1.680463	0.159321
	S438Q01T	-0.067991	0.141312
	S438Q02	-0.221055	0.146194
	S438Q03T	0.844132	0.148986
	S458Q01	1.828907	0.170628
	S458Q02T	0.624475	0.143928
	S466Q01T	0.236655	0.133644

ตารางที่ 61 ผลการประมาณค่าความยากของข้อสอบจากการวิเคราะห์ด้วยโปรแกรม HGLM-3 จำแนกตามฉบับข้อสอบจำนวน 13 ฉบับ (ต่อ)

ฉบับที่	ข้อสอบ	δ_i	S.E
10	S466Q05	0.400550	0.145593
	S466Q07T	-0.385839	0.145786
	S485Q02	1.649443	0.166948
	S485Q03	-0.101849	0.137435
	S508Q02T	1.073356	0.141805
	S508Q03	-0.000415	0.145755
	S527Q01T	2.255384	0.182936
	S527Q03T	0.348356	0.128386
	S527Q04T	-0.935177	0.110893
11	S114Q03T	1.557246	0.140695
	S114Q05T	2.763548	0.208138
	S131Q02T	0.605025	0.128073
	S131Q04T	2.338929	0.162966
	S256Q01	-1.824356	0.159946
	S413Q04T	0.970313	0.139420
	S413Q05	-0.422441	0.129054
	S413Q06	1.571596	0.143792
	S447Q02	0.419687	0.134587
	S447Q03	-0.108675	0.119587
	S447Q04	0.506906	0.133906
	S465Q02	0.082375	0.126522
	S465Q04	0.960000	0.132188
	S476Q01	0.090684	0.113388
	S476Q02	-0.000634	0.123021
	S476Q03	1.368878	0.135751
	S478Q01	0.724695	0.131225
	S478Q02T	0.586992	0.121106
	S478Q03T	0.015963	0.134236
	S495Q01T	1.282110	0.130671
	S495Q02T	0.551192	0.120520
	S495Q03	1.557246	0.140431
	S495Q04T	1.460321	0.135171
	S521Q02	0.471816	0.122997
	S521Q06	0.033838	0.096823

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 61 ผลการประมาณค่าความยากของข้อสอบจากการวิเคราะห์ด้วยโปรแกรม HGLM-3 จำแนกตามฉบับข้อสอบจำนวน 13 ฉบับ (ต่อ)

ฉบับที่	ข้อสอบ	δ_i	S.E
12	S131Q02T	0.810266	0.129240
	S131Q04T	2.664999	0.182231
	S256Q01	-1.401295	0.135201
	S269Q01	0.203305	0.129211
	S269Q03T	0.725143	0.138064
	S269Q04T	1.714542	0.150670
	S413Q04T	1.254308	0.155054
	S413Q05	0.195004	0.129105
	S413Q06	1.730396	0.147117
	S426Q03	-0.412646	0.137285
	S426Q05	-0.404044	0.114258
	S426Q07T	0.062831	0.123263
	S447Q02	0.898227	0.128064
	S447Q03	0.430914	0.132192
	S447Q04	0.800677	0.134555
	S465Q02	0.278348	0.129038
	S465Q04	1.128550	0.149664
	S478Q01	0.888298	0.132084
	S478Q02T	0.898227	0.141356
	S478Q03T	0.244917	0.127571
	S485Q02	0.999904	0.128121
	S485Q03	-0.176430	0.127764
	S508Q02T	0.669755	0.141045
	S508Q03	-0.052429	0.124673
	S527Q01T	2.072169	0.174162
	S527Q03T	0.405199	0.133418
	S527Q04T	-0.725143	0.099831
13	S213Q01T	1.195159	0.127603
	S213Q02	0.419645	0.115745
	S416Q01	0.661999	0.116872
	S421Q01	1.160656	0.139363
	S421Q03	0.129107	0.109499
	S425Q02	0.625103	0.119151
	S425Q03	0.625103	0.118828
	S425Q04	1.230299	0.134044

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 61 ผลการประมาณค่าความยากของข้อสอบจากการวิเคราะห์ด้วยโปรแกรม HGLM-3 จำแนกตามฉบับข้อสอบจำนวน 13 ฉบับ (ต่อ)

ฉบับที่	ข้อสอบ	δ_i	S.E
13	S425Q05	0.264563	0.126443
	S493Q01T	-0.013836	0.123527
	S493Q03T	-1.570104	0.136947
	S493Q05T	1.837204	0.137019
	S514Q02	-1.025879	0.120592
	S514Q03	0.188183	0.111849
	S514Q04	-0.975521	0.103119

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ง

อิทธิพลคงที่(Fixed effcet) และอิทธิพลเชิงสุ่ม(random effect) จากการวิเคราะห์ในโมเดล 1 – 4

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ผลการวิเคราะห์ตามโมเดล 1

ตารางที่ 62 อิทธิพลคงที่ (Fixed effcet)และอิทธิพลเชิงสุ่ม (random effect) จากการวิเคราะห์**โมเดล 1**

Fixed Effect	Coefficient	SE	T-ratio	p-value
G00 (ค่าเฉลี่ยของคะแนนผลการประเมินวิชาวิทยาศาสตร์ทุกสถานศึกษา)	480.157757	3.176967	151.137	0.000
Random Effect	**SD**	**Varaince Component**	**Chi-sq**	**p-value**
U0 (ค่าส่วนที่เหลือของ $B0_k$)	44.83772	2010.42084	4022.97893	0.000
R (ค่าส่วนที่เหลือของ $SCIE_{ik}$)	57.48520	3304.54876		
Reliability of OLS Regression Coefficient			**Estimate**	
$B0_k$			0.935	
ระดับความผันแปรของคะแนนผลการประเมินวิชาวิทยาศาสตร์			**ร้อยละ**	
ระดับระหว่างสถานศึกษา			37.83	
ระดับภายในสถานศึกษา			62.17	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ผลการวิเคราะห์ตามโมเดล 2

ตารางที่ 63 อิทธิพลคงที่ (Fixed effcet) และอิทธิพลเชิงสุ่ม(random effect) จากการวิเคราะห์**โมเดล 2** ที่มีการควบคุมตัวแปรคุณลักษณะระดับนักเรียนและสถานศึกษา

Fixed Effect	Coefficient	SE	t-ratio	p-value
G00 (ค่าเฉลี่ยรวมคะแนนผลการประเมินวิชาวิทยาศาสตร์ทุกสถานศึกษา)	475.385969	19.526412	24.346	0.000
G01 (ค่าสัมประสิทธิ์ของ DSCHL)	24.891340	7.167832	3.473	0.001
G02 (ค่าสัมประสิทธิ์ของ DSCHM)	19.455240	5.552651	3.504	0.001
G03 (ค่าสัมประสิทธิ์ของ TYPE1)	-0.958839	7.155359	-0.134	0.894
G04 (ค่าสัมประสิทธิ์ของ TYPE2)	-24.336141	10.629223	-2.290	0.023
G05 (ค่าสัมประสิทธิ์ของ TYPE3)	54.316148	13.570736	4.002	0.000
G06 (ค่าสัมประสิทธิ์ของ TYPE4)	-21.993594	8.906765	-2.469	0.015
G07 (ค่าสัมประสิทธิ์ของ COMMS)	-29.778053	8.784136	-3.390	0.001
G08 (ค่าสัมประสิทธิ์ของ COMMM)	-34.307281	8.265163	-4.151	0.000
G09 (ค่าสัมประสิทธิ์ของ COMML)	-16.764427	8.293121	-2.021	0.044
G010 (ค่าสัมประสิทธิ์ของ COMMXL)	-14.375449	7.996053	-1.798	0.073
G011 (ค่าสัมประสิทธิ์ของ SC06Q01)	0.348630	1.372294	0.254	0.800
G012 (ค่าสัมประสิทธิ์ของ SC16Q01)	-7.508397	2.888110	-2.600	0.010
G013 (ค่าสัมประสิทธิ์ของ SC14Q01)	3.393433	3.392929	1.000	0.319
G014 (ค่าสัมประสิทธิ์ของ SC14Q04)	0.166243	3.355095	0.050	0.961
G015 (ค่าสัมประสิทธิ์ของ SC03Q01)	-0.363404	0.117655	-3.089	0.003
G016 (ค่าสัมประสิทธิ์ของ TCSHORT)	-7.621401	5.465649	-1.394	0.165
G017 (ค่าสัมประสิทธิ์ของ SCMATEDU)	5.648004	2.108060	2.679	0.008
G018 (ค่าสัมประสิทธิ์ของ STRATIO)	-0.003694	0.000968	-3.817	0.000
G019 (ค่าสัมประสิทธิ์ของ SCIPROM)	2.066948	3.620497	0.571	0.568
G020 (ค่าสัมประสิทธิ์ของ RESPRES)	-1.156776	2.064183	-0.560	0.575
G021 (ค่าสัมประสิทธิ์ของ RESPCURR)	5.517180	2.278434	2.421	0.017
G022 (ค่าสัมประสิทธิ์ของ SC03Q03)	0.107041	0.152846	0.700	0.484
G10 (ค่าเฉลี่ยสัมประสิทธิ์ของ RELAGR ($B1_k$) ของทุกสถานศึกษา)	15.377367	1.988742	7.732	0.000
G20 (ค่าเฉลี่ยสัมประสิทธิ์ของ MALE ($B2_k$) ของทุกสถานศึกษา)	-1.072707	1.592905	-0.673	0.500

ตารางที่ 63 อิทธิพลคงที่ (Fixed effcet) และอิทธิพลเชิงสุ่ม(random effect) จากการวิเคราะห์**โมเดล 2** ที่มีการควบคุมตัวแปรคุณลักษณะระดับนักเรียนและสถานศึกษา(ต่อ)

Fixed Effect	Coefficient	SE	t-ratio	p-value
G30 (ค่าเฉลี่ยสัมประสิทธิ์ของ ST31Q02 ($B3_k$) ของทุกสถานศึกษา)	-1.833651	1.099762	-1.667	0.095
G40 (ค่าเฉลี่ยสัมประสิทธิ์ของ ST31Q03 ($B4_k$)ของทุกสถานศึกษา)	8.397479	0.993937	8.449	0.000
G50 (ค่าเฉลี่ยสัมประสิทธิ์ของ ESCS_1 ($B5_k$) ของทุกสถานศึกษา)	5.453846	1.330609	4.099	0.000
G60 (ค่าเฉลี่ยสัมประสิทธิ์ของ HEDRES_1 ($B6_k$) ของทุกสถานศึกษา)	2.853519	1.032820	2.763	0.006
G70 (ค่าเฉลี่ยสัมประสิทธิ์ของ WEALTH_1 ($B7_k$) ของทุกสถานศึกษา)	-4.145535	1.251231	-3.313	0.001
G80 (ค่าเฉลี่ยสัมประสิทธิ์ของ INTSCI_A ($B8_k$) ของทุกสถานศึกษา)	3.835460	1.168844	3.281	0.001
G90 (ค่าเฉลี่ยสัมประสิทธิ์ของ SCIEEF_A ($B9_k$) ของทุกสถานศึกษา)	0.422529	1.086527	0.389	0.697
G100 (ค่าเฉลี่ยสัมประสิทธิ์ของ SCSCIE_1 ($B10_k$) ของทุกสถานศึกษา)	-8.038848	1.214081	-6.621	0.000
G110 (ค่าเฉลี่ยสัมประสิทธิ์ของ JOYSCI_A ($B11_k$) ของทุกสถานศึกษา)	8.992101	1.427170	6.301	0.000
G120 (ค่าเฉลี่ยสัมประสิทธิ์ของ INSTSC_A ($B12_k$) ของทุกสถานศึกษา)	4.686307	1.399529	3.348	0.001
G130 (ค่าเฉลี่ยสัมประสิทธิ์ของ GENSCI_A ($B13_k$) ของทุกสถานศึกษา)	5.975125	0.958910	6.231	0.000

Random Effect	SD	Varaince Component	Chi-sq	p-value
U0 (ค่าเฉลี่ยส่วนที่เหลือของ $B0_k$)	24.83095	616.57585	1323.16135	0.000
R (ค่าส่วนที่เหลือของ $SCIE_{ik}$)	55.01077	3026.18439		

Reliability of OLS Regression Coefficient	Estimate
$B0_j$	0.835

ระดับความผันแปรของคะแนนผลการประเมินวิชาวิทยาศาสตร์	ร้อยละ
ระดับระหว่างสถานศึกษา	16.93
ระดับภายในสถานศึกษา	83.07

ผลการวิเคราะห์ตามโมเดล 3

ตารางที่ 64 อิทธิพลคงที่(Fixed effcet)และอิทธิพลเชิงสุ่ม (random effect)จากการวิเคราะห์**โมเดล 3**

Fixed Effect	Coefficient	SE	T-ratio	p-value
G00 (ค่าเฉลี่ยของคะแนนผลการประเมินวิชาวิทยาศาสตร์ในโครงการ PISA2006 ทุกสถานศึกษา)	480.283815	3.047604	157.594	0.000
Random Effect	**SD**	**Varaince Component**	**Chi-sq**	**p-value**
U0 (ค่าส่วนที่เหลือของ $B0_k$)	43.03285	1851.82608	4085.08383	0.000
R (ค่าส่วนที่เหลือของ $PVSCIE_{ik}$)	54.74317	2996.81434		
Reliability of OLS Regression Coefficient			**Estimate**	
$B0_k$			0.936	
ระดับความผันแปรของคะแนนผลการประเมินวิชาวิทยาศาสตร์			**ร้อยละ**	
ระดับระหว่างสถานศึกษา			38.19	
ระดับภายในสถานศึกษา			61.81	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ผลการวิเคราะห์ตามโมเดล 4

ตารางที่ 65 อิทธิพลคงที่ (Fixed effcet) และอิทธิพลเชิงสุ่ม(random effect) จากการวิเคราะห์**โมเดล 4** ที่มีการควบคุมตัวแปรคุณลักษณะระดับนักเรียนและสถานศึกษา

Fixed Effect	Coefficient	SE	t-ratio	p-value
G00 (ค่าเฉลี่ยรวมคะแนนผลการประเมินวิชาวิทยาศาสตร์ทุกสถานศึกษา)	475.933434	18.754051	25.378	0.000
G01 (ค่าสัมประสิทธิ์ของ DSCHL)	23.849150	6.871621	3.471	0.001
G02 (ค่าสัมประสิทธิ์ของ DSCHM)	18.589401	5.317824	3.496	0.001
G03 (ค่าสัมประสิทธิ์ของ TYPE1)	-0.698231	6.859733	-0.102	0.919
G04 (ค่าสัมประสิทธิ์ของ TYPE2)	-23.123298	10.183352	-2.271	0.024
G05 (ค่าสัมประสิทธิ์ของ TYPE3)	52.409790	13.019698	4.025	0.000
G06 (ค่าสัมประสิทธิ์ของ TYPE4)	-20.818182	8.535226	-2.439	0.016
G07 (ค่าสัมประสิทธิ์ของ COMMS)	-28.640284	8.423554	-3.400	0.001
G08 (ค่าสัมประสิทธิ์ของ COMMM)	-32.883962	7.936628	-4.143	0.000
G09 (ค่าสัมประสิทธิ์ของ COMML)	-16.183366	7.958780	-2.033	0.043
G010 (ค่าสัมประสิทธิ์ของ COMMXL)	-13.872541	7.670663	-1.809	0.072
G011 (ค่าสัมประสิทธิ์ของ SC06Q01)	0.298023	1.317233	0.226	0.821
G012 (ค่าสัมประสิทธิ์ของ SC16Q01)	-7.235974	2.769818	-2.612	0.010
G013 (ค่าสัมประสิทธิ์ของ SC14Q01)	3.221607	3.247018	0.992	0.323
G014 (ค่าสัมประสิทธิ์ของ SC14Q04)	0.096468	3.214779	0.030	0.976
G015 (ค่าสัมประสิทธิ์ของ SC03Q01)	-0.351279	0.112932	-3.111	0.003
G016 (ค่าสัมประสิทธิ์ของ TCSHORT)	-7.263655	5.241483	-1.386	0.168
G017 (ค่าสัมประสิทธิ์ของ SCMATEDU)	5.349077	2.019204	2.649	0.009
G018 (ค่าสัมประสิทธิ์ของ STRATIO)	-0.003575	0.000929	-3.847	0.000
G019 (ค่าสัมประสิทธิ์ของ SCIPROM)	1.966320	3.469140	0.567	0.571
G020 (ค่าสัมประสิทธิ์ของ RESPRES)	-1.123837	1.978309	-0.568	0.570
G021 (ค่าสัมประสิทธิ์ของ RESPCURR)	5.226349	2.180471	2.397	0.018
G022 (ค่าสัมประสิทธิ์ของ SC03Q03)	0.101552	0.146802	0.692	0.490
G10 (ค่าเฉลี่ยสัมประสิทธิ์ของ RELAGR ($B1_k$) ของทุกสถานศึกษา)	14.586167	1.891971	7.710	0.000
G20 (ค่าเฉลี่ยสัมประสิทธิ์ของ MALE ($B2_k$) ของทุกสถานศึกษา)	-1.071667	1.518343	-0.706	0.480
G30 (ค่าเฉลี่ยสัมประสิทธิ์ของ ST31Q02 ($B3_k$) ของทุกสถานศึกษา)	-1.723489	1.045891	-1.648	0.099

ตารางที่ 65 อิทธิพลคงที่ (Fixed effcet) และอิทธิพลเชิงสุ่ม(random effect) จากการวิเคราะห์**โมเดล 4** ที่มีการควบคุมตัวแปรคุณลักษณะระดับนักเรียนและสถานศึกษา(ต่อ)

Fixed Effect	Coefficient	SE	t-ratio	p-value
G40 (ค่าเฉลี่ยสัมประสิทธิ์ของ ST31Q03 ($B4_k$)ของทุกสถานศึกษา)	8.003062	0.951445	8.411	0.000
G50 (ค่าเฉลี่ยสัมประสิทธิ์ของ ESCS_1 ($B5_k$) ของทุกสถานศึกษา)	5.166525	1.267972	4.075	0.000
G60 (ค่าเฉลี่ยสัมประสิทธิ์ของ HEDRES_1 ($B6_k$) ของทุกสถานศึกษา)	2.700539	0.986255	2.738	0.007
G70 (ค่าเฉลี่ยสัมประสิทธิ์ของ WEALTH_1 ($B7_k$) ของทุกสถานศึกษา)	-3.944037	1.189431	-3.316	0.001
G80 (ค่าเฉลี่ยสัมประสิทธิ์ของ INTSCI_A ($B8_k$) ของทุกสถานศึกษา)	3.665655	1.113902	3.291	0.001
G90 (ค่าเฉลี่ยสัมประสิทธิ์ของ SCIEEF_A ($B9_k$) ของทุกสถานศึกษา)	0.405494	1.040472	0.390	0.696
G100 (ค่าเฉลี่ยสัมประสิทธิ์ของ SCSCIE_1 ($B10_k$) ของทุกสถานศึกษา)	-7.685855	1.157414	-6.641	0.000
G110 (ค่าเฉลี่ยสัมประสิทธิ์ของ JOYSCI_A ($B11_k$) ของทุกสถานศึกษา)	8.550603	1.363186	6.273	0.000
G120 (ค่าเฉลี่ยสัมประสิทธิ์ของ INSTSC_A ($B12_k$) ของทุกสถานศึกษา)	4.482620	1.331063	3.368	0.001
G130 (ค่าเฉลี่ยสัมประสิทธิ์ของ GENSCI_A ($B13_k$) ของทุกสถานศึกษา)	5.697482	0.915154	6.226	0.000

Random Effect	SD	Varaince Component	Chi-sq	p-value
U0 (ค่าเฉลี่ยส่วนที่เหลือของ $B0_k$)	23.85759	569.18444	1344.31075	0.000
R (ค่าส่วนที่เหลือของ $SCIE_{ik}$)	52.38886	2744.59225		

Reliability of OLS Regression Coefficient	Estimate
$B0_j$	0.837

ระดับความผันแปรของคะแนนผลการประเมินวิชาวิทยาศาสตร์	ร้อยละ
ระดับภายในสถานศึกษา	17.18
ระดับระหว่างสถานศึกษา	82.82

ภาคผนวก จ

คะแนนมูลค่าเพิ่ม ระดับที่ของคะแนนมูลค่าเพิ่ม และการจัดกลุ่มคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษา 212 แห่ง จากการวิเคราะห์ด้วยโมเดล 1 – 4

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 66 คะแนนมูลค่าเพิ่ม ระดับที่ของคะแนนมูลค่าเพิ่ม และการจัดกลุ่มคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษา 212 แห่ง จากการวิเคราะห์ด้วยโมเดล 1 - 4

สถานศึกษา	คะแนนมูลค่าเพิ่มของสถานศึกษา				ระดับที่ของคะแนนมูลค่าเพิ่ม				การจัดกลุ่มคุณภาพ*			
	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
12006	128.794	15.279	124.166	15.002	1	46	1	46	3	3	3	3
12010	124.318	20.744	119.479	20.003	2	32	2	32	3	3	3	3
12009	122.632	24.79	117.74	23.691	3	22	3	22	3	3	3	3
12004	121.5	20.672	116.938	20.026	4	34	4	31	3	3	3	3
8002	114.415	56.444	109.882	54.31	5	5	5	5	3	3	3	3
12005	104.295	3.605	100.307	3.688	6	90	6	88	3	2	3	2
2091	93.189	38.538	89.533	37.323	7	11	7	11	3	3	3	3
12003	92.889	10.169	89.243	9.719	8	58	8	58	3	2	3	2
2090	91.136	31.342	87.709	30.071	9	13	9	13	3	3	3	3
8005	91.052	-13.52	87.618	-12.89	10	156	10	156	3	2	3	2
2089	90.82	63.453	87.256	61.235	11	1	11	1	3	3	3	3
2021	85.43	59.656	81.977	57.406	12	3	12	3	3	3	3	3
2014	82.844	20.377	79.52	19.569	13	36	13	36	3	3	3	3
8006	82.425	13.642	78.965	12.897	14	49	14	49	3	3	3	3
12002	80.982	-10.227	77.658	-10.117	15	147	15	147	3	2	3	2
8003	80.448	-16.464	77.322	-16.006	16	167	16	167	3	1	3	1
5001	77.995	56.724	74.923	54.546	17	4	17	4	3	3	3	3
8004	77.708	-0.658	74.663	-0.719	18	111	18	111	3	2	3	2
12008	76.671	-25.573	73.826	-24.363	19	192	19	192	3	1	3	1
5003	75.982	39.024	72.815	37.397	20	10	20	10	3	3	3	3
2045	70.676	30.55	67.91	29.467	21	15	21	15	3	3	3	3
12007	69.337	-58.574	66.897	-56.164	22	212	22	212	3	1	3	1
5002	66.749	60.57	64.338	58.318	23	2	23	2	3	3	3	3
2002	63.332	45.993	61.34	44.61	24	8	24	8	3	3	3	3
2092	60.782	19.04	58.338	18.319	25	41	25	42	3	3	3	3
2073	53.736	35.27	51.73	34.144	26	12	26	12	3	3	3	3
8001	53.137	-40.329	50.81	-39.077	27	206	27	206	3	1	3	1
2044	52.148	25.672	49.886	24.532	28	20	28	20	3	3	3	3
2071	51.252	31.062	48.926	29.682	29	14	29	14	3	3	3	3
2095	48.025	-3.141	46.145	-2.94	30	121	30	121	3	2	3	2
2093	47.922	12.005	46.005	11.544	31	54	31	54	3	2	3	2

หมายเหตุ * 1 แทน กลุ่มต่ำ 2 แทน กลุ่มกลาง 3 แทน กลุ่มสูง

ตารางที่ 66 คะแนนมูลค่าเพิ่ม ระดับที่ของคะแนนมูลค่าเพิ่ม และการจัดกลุ่มคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษา 212 แห่ง จากการวิเคราะห์ด้วยโมเดล 1 - 4 (ต่อ)

สถานศึกษา	คะแนนมูลค่าเพิ่มของสถานศึกษา				ระดับที่ของคะแนนมูลค่าเพิ่ม				การจัดกลุ่มคุณภาพ			
	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
2094	47.017	4.943	45.058	4.659	32	78	33	79	3	2	3	2
7003	46.964	48.117	45.083	46.212	33	6	32	6	3	3	3	3
2074	44.116	15.589	42.649	15.414	34	45	34	45	3	3	3	3
2096	40.909	13.792	38.998	13.024	35	48	36	48	3	3	3	3
12001	40.615	30.097	39.259	29.06	36	18	35	16	3	3	3	3
2016	40.326	22.579	38.679	21.736	37	28	37	27	3	3	3	3
2001	38.999	2.543	37.158	2.374	38	94	38	95	3	2	3	2
2012	37.99	30.312	36.02	28.998	39	16	39	18	3	3	3	3
5009	37.54	24.05	35.926	22.979	40	24	40	24	3	3	3	3
2088	31.838	-16.275	30.346	-15.751	41	164	41	165	3	1	3	1
2043	31.196	7.571	30.215	7.635	42	65	42	64	3	2	3	2
2086	30.168	12.006	28.998	11.631	43	53	43	53	3	3	3	3
2098	30.063	1.078	28.839	0.992	44	101	44	101	3	2	3	2
2004	30.005	21.252	28.599	20.115	45	30	45	30	3	3	3	3
2046	29.082	21.508	27.953	20.874	46	29	46	29	3	3	3	3
2017	27.924	-12.217	26.987	-11.66	47	151	47	151	3	2	3	2
2013	27.553	-6.092	26.437	-5.649	48	130	48	130	3	2	3	2
4009	26.791	13.104	25.791	12.789	49	50	49	50	3	3	3	3
7014	26.784	39.815	25.737	38.24	50	9	50	9	3	3	3	3
2078	25.391	9.989	24.125	9.321	51	59	53	61	3	2	3	2
2097	25.277	-8.919	24.253	-8.568	52	141	52	141	3	2	3	2
3001	25.14	22.843	24.331	22.167	53	26	51	26	3	3	3	3
2019	24.853	-10.726	23.504	-10.649	54	149	54	150	2	2	2	2
2015	23.512	-9.297	22.238	-9.159	55	142	55	143	2	2	2	2
2087	20.949	19.03	20.284	18.576	56	42	56	40	2	3	2	3
2079	19.335	-0.343	18.658	-0.282	57	108	57	109	2	2	2	2
2100	18.054	-9.695	17.437	-9.23	58	146	58	144	2	2	2	2
9006	16.946	11.854	15.984	11.184	59	55	59	55	2	2	2	2
2011	15.889	0.868	15.233	0.777	60	102	60	102	2	2	2	2
10005	15.25	19.52	14.776	18.968	61	39	61	39	2	3	2	3
11011	14.78	47.881	14.167	46.114	62	7	62	7	2	3	2	3

หมายเหตุ * 1 แทน กลุ่มต่ำ 2 แทน กลุ่มกลาง 3 แทน กลุ่มสูง

ตารางที่ 66 คะแนนมูลค่าเพิ่ม ระดับที่ของคะแนนมูลค่าเพิ่ม และการจัดกลุ่มคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษา 212 แห่ง จากการวิเคราะห์ด้วยโมเดล 1 - 4 (ต่อ)

สถานศึกษา	คะแนนมูลค่าเพิ่มของสถานศึกษา				ระดับที่ของคะแนนมูลค่าเพิ่ม				การจัดกลุ่มคุณภาพ			
	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
2099	14.766	-9.514	13.956	-9.255	63	143	64	145	2	2	2	2
2103	14.618	9.762	14.134	9.421	64	61	63	60	2	2	2	2
2075	13.543	4.763	12.645	4.193	65	80	65	81	2	2	2	2
5004	12.971	6.978	12.505	6.72	66	69	66	68	2	2	2	2
2072	11.55	-1.84	11.821	-1.07	67	120	67	115	2	2	2	2
2076	11.019	-20.06	10.44	-19.394	68	178	68	178	2	1	2	1
6002	9.886	4.142	9.351	3.838	69	83	69	85	2	2	2	2
2077	9.085	5.624	8.976	5.622	70	74	70	71	2	2	2	2
2042	7.779	3.662	7.303	3.496	71	89	72	90	2	2	2	2
2006	7.693	20.556	7.384	19.861	72	35	71	34	2	3	2	3
2039	6.721	-4.946	6.457	-4.783	73	126	73	125	2	2	2	2
2070	6.485	-9.641	6.302	-9.128	74	145	74	142	2	2	2	2
7010	5.222	8.607	4.746	8.092	75	63	76	63	2	2	2	2
2085	5.081	-12.599	5.197	-11.904	76	152	75	152	2	2	2	2
2051	3.418	0.287	3.351	0.419	77	107	77	105	2	2	2	2
2068	2.021	6.048	2.045	5.942	78	71	78	70	2	2	2	2
2005	1.91	-6.13	1.807	-5.798	79	131	79	131	2	2	2	2
2084	1.575	-1.774	1.356	-1.879	80	118	81	120	2	2	2	2
2101	1.499	-20.758	1.474	-19.94	81	181	80	183	2	1	2	1
2041	1.461	-14.673	1.209	-14.213	82	160	82	160	2	1	2	1
2003	0.406	-20.18	0.073	-19.654	83	179	84	179	2	1	2	1
4001	0.374	16.992	0.206	16.127	84	43	83	44	2	3	2	3
7006	0.275	4.104	-0.023	3.733	85	84	85	86	2	2	2	2
4005	-0.246	11.471	-0.16	11.073	86	56	86	56	2	2	2	2
6003	-0.904	-20.953	-1.084	-20.358	87	183	87	184	2	1	2	1
11007	-1.187	25.114	-1.606	23.809	88	21	88	21	2	3	2	3
2060	-1.639	23.535	-1.613	22.522	89	25	89	25	2	3	2	3
7001	-1.661	30.109	-1.618	29.019	90	17	90	17	2	3	2	3
2048	-2.069	-10.526	-2.312	-10.302	91	148	91	148	2	2	2	2
2064	-2.299	27.775	-2.491	26.332	92	19	92	19	2	3	2	3
2030	-3.487	0.535	-3.649	0.327	93	105	94	106	2	2	2	2

หมายเหตุ * 1 แทน กลุ่มต่ำ 2 แทน กลุ่มกลาง 3 แทน กลุ่มสูง

ตารางที่ 66 คะแนนมูลค่าเพิ่ม ระดับที่ของคะแนนมูลค่าเพิ่ม และการจัดกลุ่มคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษา 212 แห่ง จากการวิเคราะห์ด้วยโมเดล 1 - 4 (ต่อ)

สถานศึกษา	คะแนนมูลค่าเพิ่มของสถานศึกษา				ระดับที่ของคะแนนมูลค่าเพิ่ม				การจัดกลุ่มคุณภาพ			
	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
2102	-3.644	-11.202	-3.244	-10.58	94	150	93	149	2	2	2	2
7015	-4.852	12.729	-4.844	12.264	95	51	95	51	2	3	2	3
11009	-5.425	-1.404	-4.867	-0.984	96	116	96	113	2	2	2	2
2056	-5.519	20.202	-5.432	19.398	97	37	97	37	2	3	2	3
2066	-5.918	20.714	-5.979	19.669	98	33	99	35	2	3	2	3
9002	-6.111	2.178	-5.64	2.295	99	95	98	97	2	2	2	2
2052	-6.245	-0.376	-6.112	-0.457	100	109	100	110	2	2	2	2
4002	-6.736	-0.964	-6.847	-1.179	101	113	101	117	2	2	2	2
7005	-8.095	4.659	-7.839	4.394	102	81	104	80	2	2	2	2
11021	-8.234	9.975	-7.799	9.611	103	60	102	59	2	2	2	2
2040	-8.463	12.111	-7.836	11.857	104	52	103	52	2	3	2	3
2036	-8.584	10.627	-8.684	9.803	105	57	105	57	2	2	2	2
7004	-8.976	9.33	-8.959	8.827	106	62	106	62	2	2	2	2
4006	-9.276	4.191	-9.123	3.883	107	82	107	84	2	2	2	2
2038	-10.654	2.151	-10.823	1.601	108	96	110	98	2	2	2	2
11002	-10.876	20.115	-10.534	19.221	109	38	108	38	2	3	2	3
2029	-11.499	2.133	-10.79	2.502	110	97	109	93	2	2	2	2
6007	-12.481	5.452	-11.928	5.448	111	75	111	72	2	2	2	2
2081	-13	20.837	-12.687	19.908	112	31	112	33	2	3	2	3
2318	-13.771	-8.215	-13.512	-8.045	113	137	113	137	2	2	2	2
9004	-14.357	-23.464	-14.055	-22.652	114	187	115	189	2	1	2	1
2023	-14.471	-5.276	-13.812	-4.916	115	128	114	126	2	2	2	2
2358	-14.632	7.001	-14.211	6.614	116	68	116	69	2	2	2	2
4010	-15.131	-14.038	-14.289	-13.259	117	158	117	158	2	2	2	2
6004	-15.699	0.608	-15.183	0.597	118	103	118	103	2	2	2	2
2049	-15.915	-25.53	-15.404	-24.325	119	191	119	191	2	1	2	1
6013	-16.107	22.669	-15.842	21.402	120	27	123	28	2	3	2	3
2067	-16.141	5.712	-15.662	5.434	121	73	121	73	2	2	2	2
2033	-16.363	-18.88	-15.771	-18.283	122	177	122	177	2	1	2	1
6001	-16.376	-1.795	-15.617	-1.803	123	119	120	119	2	2	2	2
6014	-16.481	-16.292	-16.53	-16.332	124	166	125	172	2	1	2	1

หมายเหตุ * 1 แทน กลุ่มต่ำ 2 แทน กลุ่มกลาง 3 แทน กลุ่มสูง

ตารางที่ 66 คะแนนมูลค่าเพิ่ม ระดับที่ของคะแนนมูลค่าเพิ่ม และการจัดกลุ่มคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษา 212 แห่ง จากการวิเคราะห์ด้วยโมเดล 1 - 4 (ต่อ)

สถานศึกษา	คะแนนมูลค่าเพิ่มของสถานศึกษา				ระดับที่ของคะแนนมูลค่าเพิ่ม				การจัดกลุ่มคุณภาพ			
	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
2069	-17.116	-9.635	-16.403	-9.318	125	144	124	146	2	2	2	2
2020	-17.595	-16.805	-16.674	-16.062	126	170	126	168	2	1	2	1
2032	-18.53	4.8	-17.797	4.691	127	79	127	78	2	2	2	2
7008	-18.556	-17.075	-17.863	-16.269	128	173	128	171	2	1	2	1
9003	-19.82	-3.743	-19.266	-3.585	129	124	129	124	2	2	2	2
11010	-20.144	19.171	-19.359	18.417	130	40	130	41	2	3	2	3
2054	-20.185	-15.308	-19.638	-14.973	131	161	133	161	2	1	2	1
2050	-20.455	-39.574	-19.736	-38.148	132	205	135	205	2	1	2	1
2035	-20.55	5.124	-19.567	5.047	133	77	132	77	2	2	2	2
10009	-20.578	7.813	-19.431	7.569	134	64	131	65	2	2	2	2
2053	-20.807	5.308	-19.655	5.386	135	76	134	75	2	2	2	2
2083	-20.992	-16.804	-20.334	-16.186	136	169	136	170	2	1	2	1
2057	-21.243	-8.477	-20.436	-8.149	137	140	137	140	2	2	2	2
9001	-21.376	0.59	-20.711	0.475	138	104	138	104	2	2	2	2
2010	-21.936	-25.872	-21.116	-24.767	139	193	140	193	2	1	2	1
2047	-22.109	-35.585	-21.104	-34.097	140	202	139	202	2	1	2	1
9005	-22.691	3.664	-21.794	3.539	141	88	142	89	2	2	2	2
2009	-22.803	16.63	-21.57	16.376	142	44	141	43	2	3	2	3
2022	-23.004	1.245	-21.941	1.4	143	100	143	99	2	2	2	2
2034	-24.244	-3.574	-23.125	-3.264	144	123	144	122	2	2	2	2
11004	-24.84	-0.473	-23.249	-0.089	145	110	145	108	2	2	2	2
7011	-25.201	24.645	-23.974	23.685	146	23	146	23	2	3	2	3
2031	-25.972	3.791	-24.513	3.925	147	87	147	83	2	2	2	2
9007	-26.51	5.745	-25.541	5.422	148	72	150	74	2	2	2	2
2080	-26.587	7.262	-24.865	7.55	149	67	148	66	2	2	2	2
6008	-26.608	3.889	-25.176	4.047	150	86	149	82	2	2	2	2
7007	-27.529	-23.42	-26.247	-22.471	151	186	152	186	2	1	2	1
9008	-27.563	3.175	-26.242	3.32	152	91	151	91	2	2	2	2
11015	-27.886	-3.323	-27.144	-3.521	153	122	153	123	2	2	2	2
11016	-28.257	-1.079	-27.263	-1.161	154	115	154	116	2	2	2	2
4007	-28.861	-35.075	-27.525	-33.534	155	201	155	201	2	1	2	1

หมายเหตุ * 1 แทน กลุ่มต่ำ 2 แทน กลุ่มกลาง 3 แทน กลุ่มสูง

ตารางที่ 66 คะแนนมูลค่าเพิ่ม ระดับที่ของคะแนนมูลค่าเพิ่ม และการจัดกลุ่มคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษา 212 แห่ง จากการวิเคราะห์ด้วยโมเดล 1 - 4 (ต่อ)

สถานศึกษา	คะแนนมูลค่าเพิ่มของสถานศึกษา				ระดับที่ของคะแนนมูลค่าเพิ่ม				การจัดกลุ่มคุณภาพ			
	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
11018	-29.729	-5.959	-27.912	-5.274	156	129	156	129	2	2	2	2
11001	-29.97	-1.543	-28.425	-1.319	157	117	157	118	2	2	2	2
2026	-30.248	14.456	-29.259	13.538	158	47	160	47	2	3	1	3
2007	-30.529	-12.895	-29.403	-12.563	159	154	161	155	2	2	1	2
10004	-30.795	-0.989	-29.18	-0.723	160	114	159	112	1	2	2	2
2059	-30.84	-26.265	-29.56	-25.073	161	195	162	194	1	1	1	1
11012	-30.944	-7.005	-28.621	-6.097	162	134	158	132	1	2	2	2
4012	-31.076	-35.629	-29.796	-34.42	163	203	163	203	1	1	1	1
2061	-31.912	4.088	-30.875	3.726	164	85	164	87	1	2	1	2
5008	-34.289	-24.039	-32.746	-23.061	165	190	165	190	1	1	1	1
2037	-34.41	-29.992	-33.228	-28.837	166	199	167	199	1	1	1	1
2024	-34.553	-16.516	-33.681	-16.184	167	168	172	169	1	1	1	1
10608	-34.764	2.694	-33.53	2.45	168	93	169	94	1	2	1	2
2027	-34.937	-0.891	-33.663	-1.005	169	112	171	114	1	2	1	2
6005	-35.069	-23.75	-33.35	-22.531	170	188	168	187	1	1	1	1
5005	-35.089	-16.943	-33.182	-15.875	171	171	166	166	1	1	1	1
10001	-35.104	-12.898	-33.59	-12.414	172	155	170	154	1	2	1	2
11005	-35.172	-4.743	-34.707	-5.135	173	125	174	128	1	2	1	2
4008	-35.908	-17.458	-34.409	-16.72	174	175	173	174	1	1	1	1
11019	-36.922	2.948	-35.727	2.684	175	92	175	92	1	2	1	2
6010	-37.469	1.385	-36.281	1.117	176	99	176	100	1	2	1	2
2065	-37.968	6.186	-36.943	5.385	177	70	177	76	1	2	1	2
10006	-38.52	-20.34	-37.256	-19.774	178	180	178	181	1	1	1	1
4011	-38.53	0.334	-37.312	0.321	179	106	179	107	1	2	1	2
3002	-38.66	-52.696	-37.555	-51.026	180	211	181	211	1	1	1	1
11017	-38.91	7.436	-37.318	7.078	181	66	180	67	1	2	1	2
1001	-39.7	-36.181	-40.194	-36.6	182	204	186	204	1	1	1	1
4003	-39.747	-8.44	-38.22	-8.131	183	139	182	139	1	2	1	2
11014	-40.443	-5.198	-38.759	-5.043	184	127	183	127	1	2	1	2
7009	-41.556	-30.738	-39.996	-29.685	185	200	184	200	1	1	1	1
4304	-41.709	-7.636	-40.113	-7.39	186	135	185	135	1	2	1	2

หมายเหตุ * 1 แทน กลุ่มต่ำ 2 แทน กลุ่มกลาง 3 แทน กลุ่มสูง

ตารางที่ 66 คะแนนมูลค่าเพิ่ม ระดับที่ของคะแนนมูลค่าเพิ่ม และการจัดกลุ่มคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ของสถานศึกษา 212 แห่ง จากการวิเคราะห์ด้วยโมเดล 1 - 4 (ต่อ)

สถานศึกษา	คะแนนมูลค่าเพิ่มของสถานศึกษา				ระดับที่ของคะแนนมูลค่าเพิ่ม				การจัดกลุ่มคุณภาพ			
	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
2055	-42.79	-27.074	-40.763	-25.713	187	196	187	196	1	1	1	1
7013	-44.197	-20.785	-42.204	-19.889	188	182	189	182	1	1	1	1
10010	-44.947	-21.181	-42.136	-19.69	189	184	188	180	1	1	1	1
2082	-45.442	-13.598	-43.62	-12.947	190	157	190	157	1	2	1	2
7012	-46.137	-14.239	-44.564	-13.955	191	159	193	159	1	2	1	2
2008	-46.173	-6.435	-44.477	-6.473	192	132	191	133	1	2	1	2
2062	-46.459	-8.394	-44.534	-8.105	193	138	192	138	1	2	1	2
7002	-47.515	-16.277	-45.2	-15.303	194	165	194	163	1	1	1	1
11003	-47.645	-15.745	-45.908	-15.326	195	162	195	164	1	1	1	1
2025	-48.261	-22.201	-45.996	-21.068	196	185	196	185	1	1	1	1
10003	-50.461	-26.006	-49.003	-25.499	197	194	199	195	1	1	1	1
11020	-50.847	-6.999	-48.841	-6.734	198	133	197	134	1	2	1	2
2028	-51.589	2.052	-48.898	2.35	199	98	198	96	1	2	1	2
6012	-52.343	-8.122	-49.816	-7.576	200	136	201	136	1	2	1	2
6009	-52.365	-28.278	-49.587	-26.7	201	197	200	197	1	1	1	1
5006	-54.181	-43.361	-52.173	-41.738	202	207	202	207	1	1	1	1
5007	-54.289	-17.791	-52.184	-17.167	203	176	203	176	1	1	1	1
11008	-56.613	-17.051	-55.371	-17.034	204	172	205	175	1	1	1	1
6006	-57.075	-12.695	-54.372	-11.976	205	153	204	153	1	2	1	2
6011	-60.076	-15.938	-57.386	-15.129	206	163	206	162	1	1	1	1
1002	-63.565	-46.423	-61.054	-44.776	207	209	207	209	1	1	1	1
11006	-68.537	-23.797	-65.12	-22.616	208	189	208	188	1	1	1	1
10007	-68.783	-51.96	-65.697	-49.664	209	210	209	210	1	1	1	1
11013	-69.324	-29.651	-66.005	-28.267	210	198	210	198	1	1	1	1
2063	-71.262	-17.241	-68.086	-16.493	211	174	211	173	1	1	1	1
10002	-92.812	-43.965	-88.874	-42.084	212	208	212	208	1	1	1	1

หมายเหตุ

* 1 แทน กลุ่มต่ำ 2 แทน กลุ่มกลาง 3 แทน กลุ่มสูง

ภาคผนวก ฉ

ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มของสถานศึกษาจากผลการวิเคราะห์ 4 โมเดล

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 67 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มของสถานศึกษาจากผลการวิเคราะห์โมเดล 1 ที่จัดกลุ่มคุณภาพการจัดการศึกษา จำแนกตามคุณลักษณะของสถานศึกษา

คุณลักษณะสถานศึกษา	กลุ่มต่ำ (Percentile ที่ 0-24.99)(N = 53)					กลุ่มกลาง (Percentile ที่ 25-74.99) (N = 106)					กลุ่มสูง (Percentile ที่ 75 ขึ้นไป) (N = 53)				
	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max
ขนาดสถานศึกษา															
- ขนาดเล็ก	21	-48.22	15.93	-92.81	-30.80	15	-16.91	13.16	-30.25	14.78	-	-	-	-	-
- ขนาดกลาง	22	-44.51	9.97	-68.78	-30.84	48	-9.09	13.48	-30.53	19.34	14	79.10	31.42	25.39	124.32
- ขนาดใหญ่	10	-45.44	9.80	-63.57	-34.29	43	-4.07	14.97	-28.86	24.85	39	55.55	27.45	25.14	128.79
ประเภทสถานศึกษา															
- รัฐ	43	-46.63	13.34	-92.81	-30.80	97	-8.14	14.81	-30.53	24.85	47	63.06	30.73	25.28	128.79
- เอกชน	10	-40.34	7.91	-54.29	-31.08	9	-8.41	12.32	-28.86	12.97	6	51.69	24.63	25.14	78.00
สังกัดสถานศึกษา															
- สพฐ.	33	-46.40	14.65	-92.81	-30.80	72	-6.72	15.45	-30.53	24.85	29	49.28	21.79	25.28	93.19
- สช.	10	-40.34	7.91	-54.29	-31.08	9	-8.41	12.32	-28.86	12.97	6	51.69	24.63	25.14	78.00
- รร.สาธิต	-	-	-	-	-	-	-	-	-	-	15	94.70	23.24	53.14	128.79
- กศท.และ กทม.	10	-47.38	8.16	-60.08	-35.07	17	-10.83	11.04	-27.53	9.89	3	38.12	10.31	26.78	46.96
- อศ.	-	-	-	-	-	8	-15.18	14.70	-27.56	16.95	-	-	-	-	-
ตำแหน่งที่ตั้งสถานศึกษา															
- หมู่บ้าน	24	-45.77	15.89	-92.81	-30.80	26	-12.66	14.76	-30.25	15.25	1	37.54	-	37.54	37.54
- เมืองเล็ก	13	-48.07	8.99	-63.57	-35.10	26	-12.28	12.55	-30.53	11.55	5	58.08	19.91	30.01	82.43
- เมือง	8	-46.14	11.29	-68.78	-35.09	31	-2.30	13.65	-25.97	23.51	15	54.33	27.96	25.39	114.42
- เมืองใหญ่	6	-38.73	3.74	-44.20	-34.41	19	-5.82	15.53	-28.86	24.85	18	54.73	29.38	25.14	122.63
- เมืองใหญ่มาก	2	-41.72	15.05	-52.37	-31.08	4	-8.75	16.54	-27.53	9.89	14	81.86	30.30	40.62	128.79
แรงกดดันของผู้ปกครองต่อสถานศึกษา															
- มาก	7	-44.94	15.09	-68.54	-30.80	25	-3.49	17.28	-29.73	24.85	27	60.20	25.98	25.39	122.63
- น้อย	27	-44.48	10.83	-71.26	-30.84	48	-8.22	13.93	-30.53	18.05	19	64.61	34.61	25.14	128.79
- ไม่พบ	19	-46.99	14.66	-92.81	-34.29	33	-11.61	12.57	-28.86	20.95	7	60.16	36.10	26.78	114.42
การขาดแคลนครูวิทยาศาสตร์															
- ไม่ขาดแคลน	13	-44.22	10.23	-68.78	-31.08	33	-9.78	12.94	-30.25	23.51	26	63.95	29.56	25.28	124.32
- ขาดแคลนน้อย	23	-46.70	13.93	-92.81	-34.29	51	-8.43	14.92	-27.89	20.95	21	60.18	33.46	25.14	128.79
- บางส่วน	10	-43.87	13.18	-68.54	-30.80	16	-4.30	17.23	-30.53	24.85	5	58.91	25.68	37.99	91.05
- ขาดแคลนมาก	7	-45.83	14.04	-71.26	-31.91	6	-7.24	14.23	-24.84	14.78	1	53.13	-	53.14	53.14

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 67 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มของสถานศึกษาจากผลการวิเคราะห์โมเดล 1 ที่จัดกลุ่มคุณภาพการจัดการศึกษา จำแนกตามคุณลักษณะของสถานศึกษา (ต่อ)

คุณลักษณะสถานศึกษา	กลุ่มต่ำ (Percentile ที่ 0-24.99)(N = 53)					กลุ่มกลาง (Percentile ที่ 25-74.99) (N = 106)					กลุ่มสูง (Percentile ที่ 75 ขึ้นไป) (N = 53)				
	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max
การขาดแคลนครูที่มีคุณวุฒิ															
- ไม่ขาดแคลน	10	-44.43	11.55	-69.32	-30.80	8	-9.24	14.85	-29.97	19.34	19	67.01	29.95	27.55	124.32
- ขาดแคลนน้อย	21	-44.85	10.78	-68.78	-30.94	50	-3.03	14.55	-28.86	24.85	24	63.06	28.18	25.14	121.50
- บางส่วน	18	-44.59	11.93	-71.26	-30.84	34	-13.84	12.75	-30.53	13.54	6	59.67	41.53	25.39	128.79
- ขาดแคลนมาก	4	-54.91	25.98	-92.81	-37.47	14	-12.07	13.34	-26.51	15.25	4	32.36	5.32	29.08	40.33
รวม	53	-45.44	12.68	-92.81	-30.80	106	-8.16	14.58	-30.53	24.85	53	61.77	30.12	25.14	128.79

ตารางที่ 68 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มของสถานศึกษาจากผลการวิเคราะห์โมเดล 2 ที่จัดกลุ่มคุณภาพการจัดการศึกษา จำแนกตามคุณลักษณะของสถานศึกษา

คุณลักษณะสถานศึกษา	กลุ่มต่ำ (Percentile ที่ 0-24.99)(N = 53)					กลุ่มกลาง (Percentile ที่ 25-74.99) (N = 106)					กลุ่มสูง (Percentile ที่ 75 ขึ้นไป) (N = 53)				
	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max
ขนาดสถานศึกษา															
- ขนาดเล็ก	7	-64.92	15.78	-92.81	-47.65	21	-34.56	10.49	-51.59	-8.10	8	-10.77	14.26	-30.25	14.78
- ขนาดกลาง	21	-27.41	38.00	-68.78	80.45	46	-8.34	31.84	-57.08	92.89	17	38.32	50.11	-16.11	124.32
- ขนาดใหญ่	25	-14.93	31.26	-63.57	69.34	39	11.73	28.67	-38.53	104.30	28	53.23	34.61	-25.20	128.79
ประเภทสถานศึกษา															
- รัฐ	45	-24.25	38.54	-92.81	80.45	96	-5.04	33.16	-57.08	104.30	46	37.93	45.68	-30.25	128.79
- เอกชน	8	-39.04	9.83	-54.29	-28.86	10	-16.72	18.34	-41.71	12.97	7	44.36	29.69	.37	78.00
สังกัดสถานศึกษา															
- สพฐ.	30	-33.74	27.91	-92.81	31.84	71	-7.86	26.13	-51.59	48.03	33	29.83	36.88	-30.25	93.19
- สช.	8	-39.04	9.83	-54.29	-28.86	10	-16.72	18.34	-41.71	12.97	7	44.36	29.69	.37	78.00
- รร.สาธิต	4	69.89	12.08	53.14	80.45	5	89.38	10.53	77.71	104.30	6	115.68	16.94	82.43	128.79
- กศท.และ กทม.	10	-34.42	18.38	-60.08	-.90	13	-20.45	21.87	-57.08	9.89	7	9.50	28.46	-25.20	46.96
- อศ.	1	-14.35	.	-14.36	-14.36	7	-15.30	15.87	-27.56	16.95	-	-	-	-	-
ตำแหน่งที่ตั้งสถานศึกษา															
- หมู่บ้าน	11	-49.49	24.06	-92.81	-17.60	29	-28.08	16.50	-51.59	14.62	11	-2.85	19.49	-30.25	37.54
- เมืองเล็ก	9	-39.99	23.63	-63.57	11.02	26	-19.07	21.83	-57.08	47.92	9	22.43	39.63	-25.20	82.43
- เมือง	15	-18.41	39.21	-68.78	76.67	27	4.19	22.59	-39.75	77.71	12	44.79	36.29	-4.85	114.42
- เมืองใหญ่	11	-22.22	22.46	-44.20	31.84	20	14.58	33.48	-38.53	92.89	12	49.56	39.47	-16.11	122.63
- เมืองใหญ่มาก	7	3.06	51.56	-52.37	80.45	4	63.31	42.96	9.89	104.30	9	83.66	35.39	40.62	128.79

ตารางที่ 68 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มของสถานศึกษาจากผลการวิเคราะห์โมเดล 2 ที่จัดกลุ่มคุณภาพการจัดการศึกษา จำแนกตามคุณลักษณะของสถานศึกษา (ต่อ)

คุณลักษณะสถานศึกษา	กลุ่มต่ำ (Percentile ที่ 0-24.99)(N = 53)					กลุ่มกลาง (Percentile ที่ 25-74.99) (N = 106)					กลุ่มสูง (Percentile ที่ 75 ขึ้นไป) (N = 53)				
	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max
แรงกดดันของผู้ปกครองต่อสถานศึกษา															
- มาก	16	-5.36	44.54	-68.54	80.45	26	13.35	34.81	-40.44	92.89	17	56.60	34.26	-8.46	122.63
- น้อย	23	-31.19	30.63	-71.26	76.67	52	-9.91	28.20	-52.34	91.05	19	45.52	48.10	-30.25	128.79
- ไม่พบ	14	-42.87	20.95	-92.81	-15.92	28	-17.26	29.68	-57.08	104.30	17	13.44	36.95	-25.20	114.42
การขาดแคลนครูวิทยาศาสตร์															
- ไม่ขาดแคลน	23	-11.47	41.25	-68.78	80.45	31	3.86	35.78	-52.34	92.89	18	50.49	47.36	-30.25	124.32
- ขาดแคลนน้อย	22	-37.81	21.65	-92.81	1.46	47	-8.67	28.90	-51.59	104.30	26	38.41	42.90	-25.20	128.79
- บางส่วน	5	-48.30	16.17	-68.54	-30.84	19	-5.58	34.73	-50.85	91.05	7	19.20	36.44	-16.11	82.43
- ขาดแคลนมาก	3	-22.12	66.18	-71.26	53.14	9	-28.71	16.48	-57.08	-5.43	2	6.79	11.29	-1.19	14.78
การขาดแคลนครูที่มีคุณวุฒิ															
- ไม่ขาดแคลน	11	-14.06	50.47	-69.32	76.67	15	7.87	46.15	-41.71	104.30	11	71.96	30.47	37.54	124.32
- ขาดแคลนน้อย	25	-24.40	34.21	-68.78	80.45	46	-.87	29.25	-57.08	80.98	24	44.58	41.65	-16.11	121.50
- บางส่วน	15	-33.83	22.05	-71.26	1.46	32	-15.20	29.70	-51.59	91.05	11	7.14	45.65	-30.25	128.79
- ขาดแคลนมาก	2	-65.66	38.39	-92.81	-38.52	13	-18.71	18.57	-50.85	30.06	7	16.49	18.35	-8.46	40.33
รวม	53	-26.48	36.03	-92.81	80.45	106	-6.15	32.18	-57.08	104.30	53	38.78	43.73	-30.25	128.79

ตารางที่ 69 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มของสถานศึกษาจากผลการวิเคราะห์โมเดล 3 ที่จัดกลุ่มคุณภาพการจัดการศึกษา จำแนกตามคุณลักษณะของสถานศึกษา

คุณลักษณะสถานศึกษา	กลุ่มต่ำ (Percentile ที่ 0-24.99)(N = 53)					กลุ่มกลาง (Percentile ที่ 25-74.99) (N = 106)					กลุ่มสูง (Percentile ที่ 75 ขึ้นไป) (N = 53)				
	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max
ขนาดสถานศึกษา															
- ขนาดเล็ก	20	-49.06	15.87	-92.81	-30.25	16	-17.82	13.22	-30.94	14.78	-	-	-	-	-
- ขนาดกลาง	23	-43.90	10.17	-68.78	-30.53	47	-8.63	13.25	-27.89	19.34	14	79.10	31.42	25.39	124.32
- ขนาดใหญ่	10	-41.65	9.80	-63.57	-34.29	43	-4.07	14.97	-28.86	24.85	39	55.55	27.45	25.14	128.79

ตารางที่ 69 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มของสถานศึกษาจากผลการวิเคราะห์โมเดล 3 ที่จัดกลุ่มคุณภาพการจัดการศึกษา จำแนกตามคุณลักษณะของสถานศึกษา (ต่อ)

คุณลักษณะสถานศึกษา	กลุ่มต่ำ (Percentile ที่ 0-24.99)(N = 53)					กลุ่มกลาง (Percentile ที่ 25-74.99) (N = 106)					กลุ่มสูง (Percentile ที่ 75 ขึ้นไป) (N = 53)				
	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max
ประเภทสถานศึกษา															
- รัฐ	43	-46.61	13.37	-92.81	-30.25	97	-8.15	14.83	-30.94	24.85	47	63.06	30.73	25.28	128.79
- เอกชน	10	-40.34	7.91	-54.29	-31.08	9	-8.41	12.32	-28.86	12.97	6	51.69	24.63	25.14	78.00
สังกัดสถานศึกษา															
- สพฐ.	33	-46.37	14.68	-92.81	-30.25	72	-6.73	15.47	-30.94	24.85	29	49.28	21.79	25.28	93.19
- สช.	10	-40.34	7.91	-54.29	-31.08	9	-8.41	12.32	-28.86	12.97	6	51.69	24.63	25.14	78.00
- รร.สาธิต	-	-	-	-	-	-	-	-	-	-	15	94.70	23.24	53.14	128.79
- กศท.และ กทม.	10	-47.38	8.16	-60.08	-35.07	17	-10.83	11.04	-27.53	9.89	3	38.12	10.31	26.78	46.96
- อศ.	-	-	-	-	-	8	-15.18	14.70	-27.56	16.95	-	-	-	-	-
ตำแหน่งที่ตั้งสถานศึกษา															
- หมู่บ้าน	23	-46.39	15.94	-92.81	-30.25	27	-13.35	14.92	-30.94	15.25	1	37.54	.	37.54	37.54
- เมืองเล็ก	14	-46.82	9.83	-63.57	-30.53	25	-11.55	12.23	-29.97	11.55	5	58.08	19.91	30.01	82.43
- เมือง	8	-46.14	11.29	-68.78	-35.09	31	-2.30	13.65	-25.97	23.51	15	54.33	27.96	25.39	114.42
- เมืองใหญ่	6	-38.73	3.74	-44.20	-34.41	19	-5.82	15.53	-28.86	24.85	18	54.73	29.38	25.14	122.63
- เมืองใหญ่มาก	2	-41.72	15.05	-52.37	-31.08	4	-8.75	16.54	-27.53	9.89	14	81.86	30.30	40.62	128.79
แรงกดดันของผู้ปกครองต่อสถานศึกษา															
- มาก	6	-47.30	15.05	-68.54	-31.08	26	-4.54	17.76	-30.80	24.85	27	60.20	25.98	25.39	122.63
- น้อย	28	-43.96	10.98	-71.26	-30.25	47	-7.76	13.71	-30.94	18.05	19	64.61	34.61	25.14	128.79
- ไม่พบ	19	-46.99	14.66	-92.81	-34.29	33	-11.61	12.57	-28.86	20.95	7	60.16	36.10	26.78	114.42
การขาดแคลนครูวิทยาศาสตร์															
- ไม่ขาดแคลน	14	-43.22	10.52	-68.78	-30.25	32	-9.14	12.60	-29.97	23.51	26	63.95	29.56	25.28	124.32
- ขาดแคลนน้อย	23	-46.70	13.93	-92.81	-34.29	51	-8.43	14.92	-27.89	20.95	21	60.18	33.46	25.14	128.79
- บางส่วน	9	-45.28	13.16	-68.54	-30.53	17	-5.88	17.91	-30.94	24.85	5	58.91	25.68	37.99	91.05
- ขาดแคลนมาก	7	-45.83	14.04	-71.26	-31.91	6	-7.24	14.23	-24.84	14.78	1	53.13	.	53.14	53.14
การขาดแคลนครูที่มีคุณวุฒิ															
- ไม่ขาดแคลน	9	-45.94	11.14	-69.32	-34.29	9	-11.64	15.65	-30.80	19.34	19	67.01	29.95	27.55	124.32
- ขาดแคลนน้อย	20	-45.54	10.56	-68.78	-31.08	51	-3.58	14.93	-30.94	24.85	24	63.06	28.18	25.14	121.50
- บางส่วน	20	-43.17	12.10	-71.26	-30.25	32	-12.81	12.47	-29.73	13.54	6	59.67	41.53	25.39	128.79
- ขาดแคลนมาก	4	-54.91	25.98	-92.81	-37.47	14	-12.07	13.33	-26.51	15.25	4	32.36	5.32	29.08	40.33
รวม	53	-45.42	12.70	-92.81	-30.25	106	-8.17	14.58	-30.94	24.85	53	61.77	30.12	25.14	128.79

ตารางที่ 70 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มของสถานศึกษาจากผลการวิเคราะห์โมเดล 4 ที่จัดกลุ่มคุณภาพการจัดการศึกษา จำแนกตามคุณลักษณะของสถานศึกษา

คุณลักษณะสถานศึกษา	กลุ่มต่ำ (Percentile ที่ 0-24.99)(N = 53)					กลุ่มกลาง (Percentile ที่ 25-74.99) (N = 106)					กลุ่มสูง (Percentile ที่ 75 ขึ้นไป) (N = 53)				
	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max
ขนาดสถานศึกษา															
- ขนาดเล็ก	7	-64.92	15.78	-92.81	-47.65	21	-34.56	10.49	-51.59	-8.10	8	-10.77	14.26	-30.25	14.78
- ขนาดกลาง	21	-27.41	38.00	-68.78	80.45	46	-8.34	31.84	-57.08	92.89	17	38.32	50.11	-16.11	124.32
- ขนาดใหญ่	25	-14.93	31.26	-63.57	69.34	39	11.73	28.67	-38.53	104.30	28	53.23	34.61	-25.20	128.79
ประเภทสถานศึกษา															
- รัฐ	45	-24.25	38.54	-92.81	80.45	96	-5.04	33.16	-57.08	104.30	46	37.93	45.68	-30.25	128.79
- เอกชน	8	-39.04	9.83	-54.29	-28.86	10	-16.72	18.34	-41.71	12.97	7	44.36	29.69	.37	78.00
สังกัดสถานศึกษา															
- สพฐ.	30	-33.74	27.91	-92.81	31.84	71	-7.86	26.13	-51.59	48.03	33	29.83	36.88	-30.25	93.19
- สช.	8	-39.04	9.83	-54.29	-28.86	10	-16.72	18.34	-41.71	12.97	7	44.36	29.69	.37	78.00
- รร.สาธิต	4	69.89	12.08	53.14	80.45	5	89.38	10.53	77.71	104.30	6	115.68	16.94	82.43	128.79
- กศท.และ กทม.	10	-34.42	18.38	-60.08	-.90	13	-20.45	21.87	-57.08	9.89	7	9.50	28.46	-25.20	46.96
- อศ.	1	-14.35	.	-14.36	-14.36	7	-15.30	15.87	-27.56	16.95	-	-	-	-	-
ตำแหน่งที่ตั้งสถานศึกษา															
- หมู่บ้าน	11	-49.49	24.06	-92.81	-17.60	29	-28.08	16.50	-51.59	14.62	11	-2.85	19.49	-30.25	37.54
- เมืองเล็ก	9	-39.99	23.63	-63.57	11.02	26	-19.07	21.83	-57.08	47.92	9	22.43	39.63	-25.20	82.43
- เมือง	15	-18.41	39.21	-68.78	76.67	27	4.19	22.59	-39.75	77.71	12	44.79	36.29	-4.85	114.42
- เมืองใหญ่	11	-22.22	22.46	-44.20	31.84	20	14.58	33.48	-38.53	92.89	12	49.56	39.47	-16.11	122.63
- เมืองใหญ่มาก	7	3.06	51.56	-52.37	80.45	4	63.31	42.96	9.89	104.30	9	83.66	35.39	40.62	128.79
แรงกดดันของผู้ปกครองต่อสถานศึกษา															
- มาก	16	-5.36	44.54	-68.54	80.45	26	13.35	34.81	-40.44	92.89	17	56.60	34.26	-8.46	122.63
- น้อย	23	-31.19	30.63	-71.26	76.67	52	-9.91	28.20	-52.34	91.05	19	45.52	48.10	-30.25	128.79
- ไม่พบ	14	-42.87	20.95	-92.81	-15.92	28	-17.26	29.68	-57.08	104.30	17	13.44	36.95	-25.20	114.42
การขาดแคลนครูวิทยาศาสตร์															
- ไม่ขาดแคลน	23	-11.47	41.25	-68.78	80.45	31	3.86	35.78	-52.34	92.89	18	50.49	47.36	-30.25	124.32
- ขาดแคลนน้อย	22	-37.81	21.65	-92.81	1.46	47	-8.67	28.90	-51.59	104.30	26	38.41	42.90	-25.20	128.79
- บางส่วน	5	-48.30	16.17	-68.54	-30.84	19	-5.58	34.73	-50.85	91.05	7	19.20	36.44	-16.11	82.43
- ขาดแคลนมาก	3	-22.12	66.18	-71.26	53.14	9	-28.71	16.48	-57.08	-5.43	2	6.79	11.29	-1.19	14.78

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 70 ค่าสถิติพื้นฐานของคะแนนมูลค่าเพิ่มของสถานศึกษาจากผลการวิเคราะห์โมเดล 4 ที่จัดกลุ่มคุณภาพการจัดการศึกษา จำแนกตามคุณลักษณะของสถานศึกษา (ต่อ)

คุณลักษณะสถานศึกษา	กลุ่มต่ำ (Percentile ที่ 0-24.99)(N = 53)					กลุ่มกลาง (Percentile ที่ 25-74.99) (N = 106)					กลุ่มสูง (Percentile ที่ 75 ขึ้นไป) (N = 53)				
	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max	n	Mean	S.D.	Min	Max
การขาดแคลนครูที่มีคุณวุฒิ															
- ไม่ขาดแคลน	11	-14.06	50.47	-69.32	76.67	15	7.87	46.15	-41.71	104.30	11	71.96	30.47	37.54	124.32
- ขาดแคลนน้อย	25	-24.40	34.21	-68.78	80.45	46	-.87	29.25	-57.08	80.98	24	44.58	41.65	-16.11	121.50
- บางส่วน	15	-33.83	22.05	-71.26	1.46	32	-15.20	29.70	-51.59	91.05	11	7.14	45.65	-30.25	128.79
- ขาดแคลนมาก	2	-65.66	38.39	-92.81	-38.52	13	-18.71	18.57	-50.85	30.06	7	16.49	18.35	-8.46	40.33
รวม	53	-26.48	36.03	-92.81	80.45	106	-6.15	32.18	-57.08	104.30	53	38.78	43.73	-30.25	128.79

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวประกฤติยา ทักษิโณ เกิดวันที่ 27 สิงหาคม พ.ศ.2517 มีภูมิลำเนาที่บ้านห้วยผึ้ง อำเภอห้วยผึ้ง จังหวัดกาฬสินธุ์ ปี 2537 สำเร็จการศึกษาศึกษาศาสตรบัณฑิต สาขาวิชาการมัธยมศึกษา วิชาเอกฟิสิกส์ - คณิตศาสตร์ ปี 2542 สำเร็จการศึกษาศึกษาศาสตรมหาบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น และในปี 2550 ศึกษาต่อในระดับปริญญาครุศาสตรดุษฎีบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ระหว่างปี 2537 - 2543 เข้ารับราชการครู ตำแหน่งอาจารย์ ระดับ 3 โรงเรียนระดับมัธยมศึกษา จังหวัดขอนแก่น ปี 2544 เข้าทำงานตำแหน่งนักวิชาการ สาขาประเมินมาตรฐาน ที่สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท.) ซึ่งรับผิดชอบในส่วนการพัฒนาข้อสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ คณิตศาสตร์ ข้อสอบวัดความสามารถในการคิดของนักเรียน 4 ช่วงชั้น และการประเมินสถานศึกษาด้านวิทยาศาสตร์และคณิตศาสตร์ ปี 2553 เป็นอาจารย์ สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

ผู้สนใจเกี่ยวกับวิทยานิพนธ์สามารถติดต่อสอบถามผู้เขียนได้ที่
E-mail : t_prakit@hotmail.com

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย