

## The Development of An Equating Method of Content Subject Areas Grade point Average (GPA) Based on Ordinary National Education Test Score (O-NET) for Senior High School Student

Sirichai Kanjanawasee

Siridej Sujiva

Aimorn Jangsiripornpakorn

Duangkamol Traiwichitkhun

### ABSTRACT

*This study was aimed at : 1) to develop an equating method of senior high school GPA using the basis of students' O-NET score ; 2) to verify the quality of such equating method ; and 3) to identify groups of schools according to learning assessment standard and learning outcomes standard, resulting from the GPA equating method. Research procedures were divided into 7 steps as following: 1) reviewed documents and research findings related to authoritative and practical equating methods. 2) proposed the selected equating methods to the O-NET score from the document research in step 1. 3) prepared database of the sampling school-Academic Year 2006, individual students' scored by schools, in five content subject areas: Thai Language, Social Studies, English Language, Mathematics, and Science. 4) prepared database of the sampling school-Academic Year 2006, individual students' O-NET score, categorized by five content subject areas: Thai Language, Social Studies, English Language, Mathematics, and Science. 5) testing at least two selected equating methods to the O-NET scores, categorized by content subject areas, yielding to individual students and schools results. 6) verified qualities of the selected equating methods to the O-NET scores categorized by content subject areas. From the selected equating methods, then, scrutinized the highest authoritative and practical one to group schools by their learning standards and learning assessment. 7) concluded the study of the equating methods and reported the development of the equating method of the Eight Content Subject Areas Grade Point Average (GPA) of Senior High Schools.*

The major findings were as following:

(1) A development of equating methods using the students' O-NET score. Four equating methods using the students' O-NET score were developed as below:

(1.1) Regression Method

$$(SG\hat{P}AX)_j = b_0 + b_1 (SONET)_j$$

(1.2) Linear Method

$$(SG\hat{P}AX)_j = \overline{SGPAX}_j + \frac{SD_{SGPAX}_j}{SD_{SONET}_j} (SONET_j - \overline{SONET}_j)$$

(1.3) Equipercentile Method

$$\text{Score of } X \text{ and } Y \text{ are equivalent when } X_{PR_k} = Y_{PR_k}$$

(1.4) Multilevel Method

First Analysis Model

Level 1: Between student within school

$$PGPAX(B/T) = \beta_{0j} + \beta_{1j} (PONET)_j + r_{ij}$$

Level 2: Between school

$$\beta_{0j} = \gamma_{00} + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + U_{1j}$$

Second Analysis Model

Level 1: Between student within school

$$PGPAX(B/T) = \beta_{0j} + r_{ij}$$

Level 2: Between school

$$\beta_{0j} = \gamma_{00} + \gamma_{01} (SONET)_j + U_{0j}$$

(2) A verification of the quality of an equating method of total and each of content subject areas Grade point Average (GPA) based on Ordinary National Education Test Score (O-NET) for senior high school students.

(2.1) Among the four selected methods: Regression Method, Linear Method, Equipercentile Method, and Multilevel Method, it was found that Linear Method and Equipercentile Method yielded more precision and empirical data fit than another two. Results from Linear and Equipercentile Methods reflected that the high quality schools trended to provide more under grading than over grading while the medium quality schools provided born under and over grading and the low quality schools provided more over grading under grading. Linear and Equipercentile were the most suitable for GPA equating but Linear Method was more practical, simple and easy to use.

(2.2) *There were high correlations between basic subjects GPA and total subjects GPA equating using all 4 methods. This indicated that both kinds of GPA were workable but total subjects GPA was more practical.*

(2.3) *Groups of school, those identified by Linear and Equipercentile Method, were different at .01 level of significant in four content subject areas: Thai Language, Mathematics, Science and Social Studies. Whereas, no different in groups of school were found in English Language. Coefficient of determination ( $R^2$ ) between O-NET score and adjusted GPA from Linear Method were higher than coefficient of determination ( $R^2$ ) calculated from Equipercentile Method. However, not much different in adjusted GPA were found when comparing between Linear and Equipercentile Method. They both scored in the same direction and related with the empirical data.*

(3) *School grading.*

(3.1) *Results of school grading characteristics by comparing actual GPA and adjusted GPA showed that most schools assigned grades higher or lower than the actual students' performance (47.43% and 47.81%). Only a few schools precisely assigned grades to students (4.76%).*

(3.2) *Results of the equating methods to each five content subject areas to the O-NET scores of the sampling schools; at about the same numbers, most schools assigned grades higher or lower than the actual students' performance. Only a few school precisely assigned grade to students. Among the five content subject areas; 49.67% schools, assigned grade in Social Studies subjects higher than the students' actual learning performance. While Thai Language subject was found grade to be given lower than students' actual learning performance in 48.20% schools. In addition, Grade, assigned in English Language Subject, was found to be precise (5.38%).*

(3.3) *Results of school grouping according to learning outcome standard indicated the significant change between before and after grade adjusting. Before grade adjusting, most schools were good (49.3%) but after grade adjusting, most schools were fair (60.8%).*

## การพัฒนาวิธีการเปรียบเทียบผลการเรียนเฉลี่ยสะสมตาม กลุ่มสาระการเรียนรู้โดยใช้คะแนน O-NET ของนักเรียน มัธยมศึกษาตอนปลาย

ศิริชัย กาญจนวาสี  
ศิริเดช สุชีวะ  
เอมอร จังศิริพรปกรณ์  
ดวงกมล ไตรวิจิตรคุณ

### บทคัดย่อ

การวิจัยเรื่องการพัฒนาวิธีการเปรียบเทียบผลการเรียนเฉลี่ยสะสมรวมและจำแนกตามกลุ่มสาระการเรียนรู้ โดยใช้คะแนน O-NET ของนักเรียนมัธยมศึกษาตอนปลาย มีวัตถุประสงค์ 3 ประการ คือ 1) เพื่อพัฒนาวิธีการเปรียบเทียบผลการเรียนเฉลี่ยสะสมรวมและจำแนกตามกลุ่มสาระการเรียนรู้โดยใช้คะแนน O-NET ของนักเรียนมัธยมศึกษาตอนปลาย 2) เพื่อตรวจสอบคุณภาพวิธีการเปรียบเทียบผลการเรียนเฉลี่ยสะสมรวมและจำแนกตามกลุ่มสาระการเรียนรู้โดยใช้คะแนน O-NET ของนักเรียนมัธยมศึกษาตอนปลาย และ 3) เพื่อจัดกลุ่มโรงเรียนตามมาตรฐานการประเมินผลการเรียนรู้และมาตรฐานการเรียนรู้ จากผลการเปรียบเทียบผลการเรียนเฉลี่ยสะสมรวมและจำแนกตามกลุ่มสาระการเรียนรู้ วิธีดำเนินการวิจัย ประกอบด้วย 7 ขั้นตอน ได้แก่ 1) ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับวิธีการเปรียบเทียบคะแนน 2) เสนอวิธีการเปรียบเทียบโดยใช้คะแนน O-NET 3) การจัดเตรียมฐานข้อมูลของคะแนนรายวิชาของ สพฐ. 4) การจัดเตรียมความพร้อมของฐานข้อมูลของคะแนน O-NET 5) ทดลองรูปแบบการเปรียบเทียบคะแนนผลการเรียนเฉลี่ยสะสม 6) การตรวจสอบคุณภาพของวิธีการเปรียบเทียบ และ 7) สรุปผลการทดลองใช้วิธีการเปรียบเทียบ และจัดทำรายงาน

ผลการวิจัยสรุปประเด็นสำคัญได้ดังนี้

(1) ผลการพัฒนาวิธีการเปรียบเทียบโดยใช้คะแนน O-NET

การวิจัยครั้งนี้พัฒนาวิธีในการเปรียบเทียบโดยใช้คะแนน O-NET จำแนกเป็น 4 วิธี คือ

(1.1) วิธีวิเคราะห์ถดถอย (Regression Method)

$$(SG\hat{P}AX)_j = b_0 + b_1 (SONET)_j$$

(1.2) วิธีเชิงเส้นตรง (Linear Method)

$$(SG\hat{P}AX)_j = \overline{SGPAX}_j + \frac{SD_{SGPAX}_j}{SD_{SONET}_j} (SONET_j - \overline{SONET}_j)$$

(1.3) วิธีอิกวิเปอร์เซ็นไทล์ (Equipercntile Method)

คะแนน  $X_i$  และ  $Y_i$  ถือเป็นคะแนนสมมูลกัน เมื่อ  $X_{PR_k} = Y_{PR_k}$

(1.4) วิธีวิเคราะห์พหุระดับ (Multilevel Method)

โมเดลการวิเคราะห์ที่ 1

สมการวิเคราะห์ระดับที่ 1: ระดับนักเรียน (Level 1: Between student within school)

$$PGPAX(B/T) = \beta_{0j} + \beta_{1j} (PONET)_j + r_{ij}$$

สมการวิเคราะห์ระดับที่ 2: ระดับผู้เรียน (Level 2: Between school)

$$\beta_{0j} = \gamma_{00} + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + U_{1j}$$

โมเดลการวิเคราะห์ที่ 2

สมการวิเคราะห์ระดับที่ 1: ระดับนักเรียน (Level 1: Between student within school)

$$PGPAX(B/T) = \beta_{0j} + r_{ij}$$

สมการวิเคราะห์ระดับที่ 2: ระดับผู้เรียน (Level 2: Between school)

$$\beta_{0j} = \gamma_{00} + \gamma_{01} (SONET)_j + U_{0j}$$

(2) การตรวจสอบคุณภาพวิธีการเปรียบเทียบผลการเรียนเฉลี่ยสะสมรวมและจำแนกตามกลุ่มสาระการเรียนรู้โดยใช้คะแนน O-NET ของนักเรียนมัธยมศึกษาตอนปลาย

(2.1) ผลการเปรียบเทียบวิธีการเปรียบเทียบผลการเรียนเฉลี่ยสะสม 4 วิธี ได้แก่ วิธีวิเคราะห์ถดถอย วิธีเชิงเส้นตรง วิธีอิคิวเปอร์เซ็นไทล์ และวิธีวิเคราะห์พหุระดับ พบว่าวิธีเชิงเส้นตรงและวิธีการอิคิวเปอร์เซ็นไทล์ให้ผลการเปรียบเทียบใกล้เคียงกัน และสอดคล้องกับข้อมูลเชิงประจักษ์มากกว่าวิธีวิเคราะห์ถดถอย และวิธีวิเคราะห์พหุระดับ กล่าวคือ โรงเรียนที่มีคุณภาพสูงมีลักษณะการให้เกรดแบบกวดเกรดมากกว่าปล่อยเกรด โรงเรียนที่มีคุณภาพปานกลางมีลักษณะการให้เกรดแบบกวดเกรดและปล่อยเกรดในจำนวนที่ใกล้เคียงกัน และโรงเรียนที่มีคุณภาพต่ำมีลักษณะการให้เกรดแบบปล่อยเกรดมากกว่ากวดเกรด ดังนั้น วิธีเชิงเส้นตรง และวิธีอิคิวเปอร์เซ็นไทล์ จึงเหมาะสำหรับการเปรียบเทียบผลการเรียน อย่างไรก็ตามในทางปฏิบัติ วิธีการวิเคราะห์เชิงเส้น มีกระบวนการในการเปรียบเทียบผลการเรียนที่ไม่ซับซ้อน สะดวกในการนำไปใช้จริงมากกว่าวิธีอิคิวเปอร์เซ็นไทล์

(2.2) ผลการเปรียบเทียบเมื่อนำคะแนน O-NET มาเปรียบเทียบผลการเรียนเฉลี่ยสะสมเฉพาะวิชาพื้นฐาน (GPAXB) และผลการเรียนเฉลี่ยสะสมรวมทุกวิชา (GPAXT) โดยใช้วิธีการเปรียบเทียบทั้ง 4 วิธี พบว่า ผลการเปรียบเทียบที่ได้มีความสัมพันธ์กันสูง และเป็นไปในทิศทางเดียวกัน ดังนั้น ในการเปรียบเทียบจึงสามารถเลือกใช้ผลการเรียนเฉลี่ยสะสมทั้ง 2 แบบ อย่างไรก็ตามในทางปฏิบัติ การนำผลการเรียนเฉลี่ยสะสมรวมทุกวิชา (GPAXT) มาเปรียบเทียบ จะสะดวกในการนำไปใช้จริงมากกว่าการเปรียบเทียบจากผลการเรียนเฉลี่ยสะสมเฉพาะวิชาพื้นฐาน (GPAXB)

(2.3) ผลการเปรียบเทียบผลการเรียนเฉลี่ยสะสมจำแนกตามกลุ่มสาระการเรียนรู้ภาษาไทย คณิตศาสตร์ วิทยาศาสตร์ และสังคมศึกษา ศาสนา และวัฒนธรรม ด้วยวิธีเชิงเส้นตรง และวิธีอิคิวเปอร์เซ็นไทล์ พบว่าการแบ่งกลุ่มโรงเรียนแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ขณะที่ผลการเปรียบเทียบผลการเรียนเฉลี่ยสะสมในกลุ่มสาระการเรียนรู้ภาษาต่างประเทศ ด้วยวิธีเชิงเส้นตรง และวิธีอิคิวเปอร์เซ็นไทล์ ไม่แตกต่างกัน การเปรียบเทียบค่าสัมประสิทธิ์การทำนาย ( $R^2$ ) ของคะแนน O-NET

ตามกลุ่มสาระการเรียนรู้ กับผลการเรียนเฉลี่ยสะสมที่ปรับแล้ว พบว่า วิธีเชิงเส้นตรงมี  $R^2$  สูงกว่าวิธี อีควิเปอร์เซ็นไทล์ อย่างไรก็ตามผลการเปรียบเทียบที่ได้จากวิธีการทั้ง 2 วิธี ไม่แตกต่างกันมากนัก และเป็นไปในทิศทางเดียวกัน สอดคล้องกับข้อมูลเชิงประจักษ์

### (3) ผลการจัดกลุ่มมาตรฐานโรงเรียน

(3.1) ผลการจัดกลุ่มโรงเรียนตามมาตรฐานการประเมินผลการเรียนรู้ในภาพรวมโดย นำผลการเรียนเฉลี่ยสะสมรวมทุกวิชา (GPAXT) ของสถานศึกษามาเปรียบเทียบกับผลการเรียนเฉลี่ยสะสมที่ปรับแก้แล้ว พบว่าโรงเรียนส่วนใหญ่ปล่อยเกรด และกดเกรด จำนวนไม่ต่างกันมากนัก มีโรงเรียนเพียงส่วนน้อยเท่านั้นที่ให้เกรดแบบปกติ ทั้งนี้พบว่ามิโรงเรียนที่ให้เกรด GPAXT แบบปล่อยเกรดคิดเป็นร้อยละ 47.43 (1,225 โรงเรียน) ให้เกรดแบบกดเกรด ร้อยละ 47.81 (1,235 โรงเรียน) และให้เกรดแบบปกติ ร้อยละ 4.76 (123 โรงเรียน) โดยเมื่อพิจารณาโรงเรียนในกลุ่มปล่อยเกรด พบว่าเป็นโรงเรียนในกลุ่มปล่อยเกรดน้อย ปล่อยเกรดค่อนข้างน้อย ปล่อยเกรดค่อนข้างมาก และปล่อยเกรดมากคิดเป็นร้อยละ 66.53 (815 โรงเรียน), 27.27 (334 โรงเรียน), 5.06 (62 โรงเรียน) และ 1.14 (14 โรงเรียน) ตามลำดับ เมื่อพิจารณาโรงเรียนในกลุ่มกดเกรด พบว่าเป็นโรงเรียนในกลุ่มกดเกรดน้อย กดเกรดค่อนข้างน้อย กดเกรดค่อนข้างมาก และกดเกรดมาก คิดเป็นร้อยละ 81.30 (1,004 โรงเรียน), 17.09 (211 โรงเรียน), 1.54 (19 โรงเรียน) และ 0.08 (1 โรงเรียน) ตามลำดับ

(3.2) ผลการจัดกลุ่มโรงเรียนตามมาตรฐานการประเมินผลการเรียนรู้ โดยนำผลการเรียนเฉลี่ยสะสมในกลุ่มสาระการเรียนรู้ ทั้ง 5 กลุ่ม ของสถานศึกษา ได้แก่ ภาษาไทย คณิตศาสตร์ วิทยาศาสตร์ สังคมศึกษา ศาสนา และวัฒนธรรม และภาษาต่างประเทศ มาเปรียบเทียบกับผลการเรียนเฉลี่ยสะสมในกลุ่มสาระการเรียนรู้ที่ปรับแก้แล้ว พบว่า กลุ่มสาระการเรียนรู้ที่มีการปล่อยเกรดมากที่สุด คือ สังคมศึกษา ศาสนา และวัฒนธรรม โดยมีโรงเรียนที่ปล่อยเกรดอยู่ร้อยละ 49.67 ส่วนกลุ่มสาระการเรียนรู้ที่มีการกดเกรดมากที่สุด คือ ภาษาไทย โดยมีโรงเรียนที่กดเกรดอยู่ร้อยละ 48.20 สำหรับกลุ่มสาระการเรียนรู้ที่ให้เกรดแบบปกติมากที่สุด คือ ภาษาต่างประเทศ โดยมีโรงเรียนที่ให้เกรดแบบปกติอยู่ร้อยละ 5.34 เมื่อพิจารณาลักษณะการประเมินผลการเรียนรู้ในกลุ่มสาระการเรียนรู้ ทั้ง 5 กลุ่ม ภายในโรงเรียนแต่ละแห่ง พบว่าลักษณะการประเมินผลการเรียนรู้ในโรงเรียนแต่ละแห่งมีความหลากหลาย โดยโรงเรียนที่ให้เกรดแบบปล่อยเกรดทุกกลุ่มสาระการเรียนรู้ คิดเป็นร้อยละ 17.27 (446 โรงเรียน) โรงเรียนที่ให้เกรดแบบกดเกรดทุกกลุ่มสาระการเรียนรู้คิดเป็นร้อยละ 16.03 (414 โรงเรียน) และโรงเรียนที่ให้เกรดแบบกดเกรดบางกลุ่มสาระการเรียนรู้และปล่อยเกรดบางกลุ่มสาระการเรียนรู้ คิดเป็นร้อยละ 66.70 (1,723 โรงเรียน)

(3.3) ผลการจัดกลุ่มโรงเรียนตามมาตรฐานการเรียนรู้ พบว่าคุณภาพโรงเรียน ซึ่งพิจารณาจากผลการเรียนเฉลี่ยสะสมก่อนและหลังการปรับเทียบผลการเรียนเฉลี่ยสะสม แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ทั้งนี้เมื่อพิจารณาคคุณภาพโรงเรียนจากผลการเรียนเฉลี่ยสะสมของโรงเรียนก่อนการปรับเทียบ พบว่าโรงเรียนส่วนใหญ่มีคุณภาพในระดับดี (B) คิดเป็นร้อยละ 49.3 แต่เมื่อพิจารณาคคุณภาพโรงเรียนจากผลการเรียนเฉลี่ยสะสมของโรงเรียนที่ปรับเทียบแล้ว พบว่าโรงเรียนส่วนใหญ่มีคุณภาพในระดับพอใช้ (C) คิดเป็นร้อยละ 60.8 นอกจากนี้ยังพบว่าหลังการปรับเทียบผลการเรียนเฉลี่ยสะสมแล้ว โรงเรียนในกลุ่มคุณภาพดีเยี่ยม (A+) ดีมาก (A) และพอใช้ (C) มีจำนวนเพิ่มขึ้น ส่วนโรงเรียนในกลุ่มคุณภาพดี (B) และปรับปรุง (D) มีจำนวนลดลง

## ความเป็นมาและความสำคัญของปัญหา

จากการประเมินสถานภาพการศึกษาไทยด้วยดัชนีพัฒนาการจัดการนานาชาติ (Institute for Management Development, IMD) พบว่าประเทศไทยอยู่ในอันดับที่ 48 จาก 61 ประเทศ (สำนักงานเลขาธิการสภาการศึกษา, 2549) เมื่อวัดคุณภาพนักเรียนโดยโครงการประเมินผลนักเรียนนานาชาติ (Programme for International Student Assessment, PISA) สำหรับประเทศสมาชิก OECD (Organisation for Economic Co-operation and Development) จำนวน 41 ประเทศ และโครงการประเมินผลการเรียนคณิตศาสตร์และวิทยาศาสตร์ (Third in International Mathematics and Science Study, TIMSS) ด้านการอ่าน พบว่านักเรียนไทยอยู่ในอันดับที่ 32 คณิตศาสตร์อยู่ในอันดับที่ 34-36 วิทยาศาสตร์อยู่ในอันดับที่ 36 ส่วนทักษะการแก้ปัญหาอยู่ในอันดับที่ 34 (Boston College, nd.) ประกอบกับการประเมินคุณภาพการศึกษาโดยสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) รอบแรก ปี 2544 - 2548 จำนวน 30,010 แห่ง พบว่าโรงเรียนที่ไม่ผ่านเกณฑ์มาตรฐานมีจำนวน 560 แห่ง (สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา, มปป.) และจากผลวิจัยของคณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย พบว่านิสิตปี 1 ที่ผ่านระบบแอดมิชชันส์มีอัตราส่วนการสอบไม่ผ่าน หรือต้องถอนรายวิชาค่อนข้างสูง (รัฐชาติ มงคลนาวิน, 2550) ซึ่งสะท้อนให้เห็นว่าทั้งผลสัมฤทธิ์ทางการเรียนและศักยภาพการแข่งขันเมื่อวัดจากคุณภาพของผู้เรียนในปัจจุบันยังอยู่ในระดับต่ำ ทั้งที่ได้มีความพยายามปรับปรุงหลักสูตร การจัดการเรียนการสอน การวัด และประเมินผล รวมทั้งระบบการสอบคัดเลือกเข้าสถาบันอุดมศึกษาที่ส่งเสริมให้นักเรียนระดับมัธยมศึกษาได้ศึกษาครบถ้วนตามหลักสูตร โดยกำหนดให้นำผลการเรียนเฉลี่ยสะสม มาเป็นส่วนประกอบในการคิดคะแนนการสอบคัดเลือกด้วยการกำหนดให้ใช้ค่าคะแนนดังกล่าวจึงเป็นความพยายามให้ผู้เรียนตั้งใจศึกษาครบทุกรายวิชาและเกิดการเรียนรู้อย่างแท้จริงทุกชั้นปีตลอดหลักสูตรการศึกษา

การที่ผลสัมฤทธิ์ของผู้เรียนอยู่ในระดับต่ำนั้นได้มีการวิเคราะห์ถึงปัจจัยที่เป็นสาเหตุหลายประการ แต่สาเหตุสำคัญที่เกี่ยวข้องกับการจัดการเรียนการสอน ก็คือ การกระจายอำนาจทางการศึกษาตามแนวปฏิรูปการศึกษา ก่อให้เกิดความหลากหลายในด้านรูปแบบการศึกษา สังเกตของสถานศึกษา บริบทของสถานศึกษา ขนาดของสถานศึกษา หลักสูตรและการสอน แผนการเรียน การวัด และการประเมินผลของสถานศึกษา ซึ่งแต่ละสถานศึกษามีศักยภาพและความพร้อมที่แตกต่างกัน จึงเป็นปัจจัยที่ไม่สามารถถือได้ว่าคะแนนเฉลี่ยสะสมของนักเรียนในแต่ละสถานศึกษา มีความเท่าเทียมกัน การนำคะแนนเฉลี่ยสะสมระดับมัธยมศึกษาตอนปลายมาใช้โดยไม่ผ่านกระบวนการจัดการทำให้คะแนนเฉลี่ยสะสมระดับมัธยมศึกษาตอนปลายเท่าเทียมกันก่อนย่อมก่อให้เกิดความไม่เป็นธรรมกับนักเรียนแต่ละคน

สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) เป็นหน่วยงานของรัฐจัดตั้งขึ้นตามพระราชกฤษฎีกาจัดตั้งสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) พ.ศ. 2548 เพื่อเป็นองค์กรกลางในการศึกษา วิจัย พัฒนา และให้บริการประเมินผลทางการศึกษาและทดสอบทางการศึกษา รวมทั้งเป็นศูนย์กลางความร่วมมือด้านการทดสอบทางการศึกษาในระดับชาติและนานาชาติ โดยเฉพาะอย่างยิ่งเป็นหน่วยงานที่ทำหน้าที่ชี้คุณภาพการศึกษาด้วยคะแนนสอบ ผลิตภัณฑ์ที่มีคุณภาพ ซึ่งการสอบความรู้รอบยอดปลายช่วงชั้นเป็นองค์ประกอบที่สำคัญ เพื่อวัดความรู้รอบยอดปลายช่วงชั้น (6 ภาคเรียน) ของชั้น ป.3 ป.6 ม.3 และ ม.6 ตามมาตรฐานการเรียนรู้ของหลักสูตรการศึกษาขั้นพื้นฐาน 2544 ทั้ง 8 กลุ่มสาระ โดยนักเรียนสามารถสอบได้ 1 ครั้ง สำหรับผู้ที่กำลังจะจบชั้น ป.3 ป.6 ม.3 และ ม.6 เพื่อจะได้คะแนนกลางที่เชื่อถือได้ สำหรับนำมาปรับปรุงการเรียนการสอนของโรงเรียนโดยเทียบกับโรงเรียนอื่น สังกัดอื่น และประเทศต่าง ๆ (อุทุมพร จามรมาน, 2551)

จากความสำคัญของปัญหาดังกล่าว ทำให้คณะผู้บริหารของสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานตระหนักถึงความจำเป็นและความสำคัญที่จะต้องมีการศึกษาวิจัยเพื่อแสวงหากระบวนการในการทำให้ผลการเรียนเฉลี่ยสะสมจากแต่ละโรงเรียนมีความเท่าเทียมกัน สามารถนำมาเปรียบเทียบกันได้อย่างยุติธรรม จึงเป็นที่มาของการวิจัยนำร่องเพื่อพัฒนาวิธีการเปรียบเทียบ (Equating) ซึ่งเป็นกระบวนการที่ใช้กำหนดความสัมพันธ์ระหว่างคะแนนจากแบบสอบต่างฉบับกันที่วัดคุณลักษณะเดียวกัน และทดสอบต่างเวลากันให้สามารถใช้คะแนนจากแบบสอบฉบับหนึ่งปรับเข้าสู่สเกล (มาตร) ของแบบสอบอีกฉบับได้ ในการวิจัยนำร่องครั้งนี้เป็นการพัฒนาวิธีการเปรียบเทียบผลการเรียนเฉลี่ยสะสมตามกลุ่มสาระการเรียนรู้ โดยใช้คะแนน O-NET ของนักเรียนมัธยมศึกษาตอนปลาย เพื่อให้ผลคะแนนเฉลี่ยสะสมของนักเรียนที่มาจากโรงเรียนที่ต่างกัน สามารถนำมาเปรียบเทียบกันได้อย่างทัดเทียมกัน ผลการวิจัยที่ได้จะเป็นข้อมูลย้อนกลับและช่วยกำหนดแนวทางในการกำหนดนโยบายการวัดและประเมินผล และพัฒนาองค์ความรู้เกี่ยวกับระบบการคัดเลือกบุคคลเข้าศึกษาต่อในระดับอุดมศึกษาของประเทศต่อไป

### วัตถุประสงค์การวิจัย

1. เพื่อพัฒนาวิธีการเปรียบเทียบผลการเรียนเฉลี่ยสะสมรวมและจำแนกตามกลุ่มสาระการเรียนรู้โดยใช้คะแนน O-NET ของนักเรียนมัธยมศึกษาตอนปลาย
2. เพื่อตรวจสอบคุณภาพของวิธีการเปรียบเทียบผลการเรียนเฉลี่ยสะสมรวมและจำแนกตามกลุ่มสาระการเรียนรู้ โดยใช้คะแนน O-NET ของนักเรียนมัธยมศึกษาตอนปลาย
3. เพื่อจัดกลุ่มโรงเรียนตามมาตรฐานการประเมินผลการเรียนรู้และมาตรฐานการเรียนรู้จากผลการเปรียบเทียบผลการเรียนเฉลี่ยสะสมรวมและจำแนกตามกลุ่มสาระการเรียนรู้



## ขอบเขตการวิจัย

การวิจัยครั้งนี้ได้กำหนดขอบเขตในการศึกษาไว้ ดังนี้

1. การเปรียบเทียบคะแนนในการวิจัยครั้งนี้ ต้องการเปรียบเทียบผลการเรียนเฉลี่ยสะสมระดับโรงเรียนและนักเรียนในภาพรวมและจำแนกตาม 5 กลุ่มสาระการเรียนรู้ จากเกรดของนักเรียนระดับมัธยมศึกษาตอนปลาย 6 ภาคการศึกษา สำหรับโรงเรียนในปีการศึกษา 2549 และใช้คะแนนจากแบบสอบถาม คือ แบบทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) ในปีการศึกษา 2549 เป็นตัวเปรียบเทียบผลการเรียนเฉลี่ยสะสม โดยกลุ่มสาระการเรียนรู้ 5 กลุ่ม ได้แก่ กลุ่มสาระการเรียนรู้ภาษาไทย กลุ่มสาระการเรียนรู้คณิตศาสตร์ กลุ่มสาระการเรียนรู้วิทยาศาสตร์ กลุ่มสาระการเรียนรู้สังคมศึกษา ศาสนา และวัฒนธรรม และกลุ่มสาระการเรียนรู้ภาษาต่างประเทศ
2. ประชากรที่ใช้ในการศึกษาครั้งนี้ คือ โรงเรียนระดับมัธยมศึกษา สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) ในปีการศึกษา 2549 มีจำนวนทั้งสิ้น 2,583 โรงเรียน

## วิธีการดำเนินงานวิจัย

การวิจัยครั้งนี้ มีขั้นตอนการดำเนินการวิจัย ดังนี้

1. ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับวิธีการเปรียบเทียบคะแนนที่เหมาะสมและมีความเป็นไปได้ในทางปฏิบัติ
2. พัฒนารูปแบบการเปรียบเทียบโดยใช้คะแนน O-NET ของนักเรียนมัธยมศึกษาตอนปลายเป็นตัวเปรียบเทียบ 2 - 3 วิธี ที่เหมาะสมจากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง
3. จัดเตรียมฐานข้อมูลของคะแนนรายวิชาของ สพฐ. จำแนกตาม 5 กลุ่มสาระการเรียนรู้ของโรงเรียนที่เป็นฐานข้อมูล ในปีการศึกษา 2549 โดยจัดเตรียมแฟ้มข้อมูลเป็น 2 แฟ้ม คือ แฟ้มข้อมูลนักเรียน (Student File) และแฟ้มข้อมูลโรงเรียน (School File)
4. จัดเตรียมฐานข้อมูลของคะแนน O-NET จำแนกตาม 5 กลุ่มสาระการเรียนรู้ของโรงเรียนที่เป็นฐานข้อมูล ในปีการศึกษา 2549 โดยจัดเตรียมแฟ้มข้อมูลเป็น 2 แฟ้ม คือ แฟ้มข้อมูลนักเรียน (Student File) และแฟ้มข้อมูลโรงเรียน (School File)
5. ทดลองใช้รูปแบบการเปรียบเทียบคะแนนผลการเรียนเฉลี่ยสะสมตามวิธีที่คัดสรรอย่างน้อย 2 วิธี โดยใช้คะแนน O-NET เป็นตัวเปรียบเทียบ จำแนกตามกลุ่มสาระการเรียนรู้ ทำให้ได้ผลการเปรียบเทียบของโรงเรียนที่เป็นกลุ่มตัวอย่างทั้งในรายบุคคล และรายโรงเรียน
6. ตรวจสอบคุณภาพของวิธีการเปรียบเทียบผลการเรียนเฉลี่ยสะสมรวมและจำแนกตามกลุ่มสาระการเรียนรู้ โดยใช้คะแนน O-NET ของนักเรียนมัธยมศึกษาตอนปลาย แล้วเลือกวิธีการเปรียบเทียบที่มีคุณภาพและมีความเป็นไปได้ในการนำไปใช้ และนำวิธีการเปรียบเทียบวิธีที่เลือกมาใช้ในจัดกลุ่มโรงเรียนตามมาตรฐานการประเมินผลการเรียนรู้และมาตรฐานการเรียนรู้

7. สรุปผลการทดลองใช้วิธีการเปรียบเทียบ และจัดทำรายงานผลการพัฒนาวิธีการเปรียบเทียบ  
 ผลการเรียนเฉลี่ยสะสมรวมและจำแนกตามกลุ่มสาระการเรียนรู้โดยใช้คะแนน O-NET ของนักเรียน  
 มัธยมศึกษาตอนปลาย

## ผลการวิจัย

สามารถสรุปผลการวิจัยประเด็นสำคัญได้ดังนี้

### 1. ผลการพัฒนาวิธีการเปรียบเทียบโดยใช้คะแนน O-NET

จากเพิ่มข้อมูลระดับนักเรียน ประกอบด้วยตัวแปรดังต่อไปนี้

$PGPAX_{ij}$  = ผลการเรียนเฉลี่ยสะสมของนักเรียนคนที่  $i$  ในโรงเรียน  $j$

$PONET_{ij}$  = คะแนนผลทดสอบ O-NET ในแต่ละกลุ่มสาระการเรียนรู้ของนักเรียน  
 คนที่  $i$  ในโรงเรียน  $j$

จากเพิ่มข้อมูลระดับโรงเรียน ประกอบด้วยตัวแปรดังต่อไปนี้

$SGPAX_j$  = ผลการเรียนเฉลี่ยสะสมของโรงเรียน  $j$

$SONET_j$  = คะแนนเฉลี่ยผลทดสอบ O-NET ของโรงเรียน  $j$

สำหรับวิธีในการเปรียบเทียบโดยใช้คะแนน O-NET ในการวิจัยครั้งนี้ พัฒนาขึ้น 4 วิธี  
 คือ

#### 1.1 วิธีวิเคราะห์ถดถอย (Regression Method)

เป็นวิธีที่ใช้เทคนิคการวิเคราะห์สมการถดถอย (Regression Equation) สำหรับการ  
 เชื่อมโยงคะแนนระหว่างแบบสอบ การเชื่อมโยงคะแนนโดยวิธีนี้เป็นการใช้สมการเส้นตรง  
 ในการทำนายตัวแปรตาม (Dependent Variable) จากตัวแปรอิสระ (Independent Variable) แสดง  
 ดังสมการ (1)

$$(SG\hat{P}AX)_j = b_0 + b_1 (SONET)_j \quad (1)$$

#### 1.2 วิธีเชิงเส้นตรง (Linear Method)

วิธีการเปรียบเทียบเชิงเส้นตรงเป็นวิธีที่ยึดถือนิยามของการเชื่อมโยงคะแนนว่า  
 คะแนนจากแบบสอบสองฉบับ  $X$  และ  $Y$  เมื่อแบบสอบ  $X$  และ  $Y$  วัดสิ่งเดียวกันและมีความเที่ยง  
 ใกล้เคียงกัน จะถือว่าคะแนนสมมูลกัน (Equivalent Score) เมื่อคะแนนของแต่ละฉบับมีคะแนน  
 มาตรฐาน (Standard Score) เท่ากัน แสดงดังสมการ (2)

$$(SG\hat{P}AX)_j = \overline{SGPAX}_j \frac{SD_{SGPAX_j}}{SD_{SONET_j}} (SONET)_j - \overline{SONET}_j \quad (2)$$

### 1.3 วิธีอิกวิเปอร์เซ็นไทล์ (Equipercntile Method)

การเปรียบเทียบคะแนนแบบอิกวิเปอร์เซ็นไทล์ใช้นิยามของการเปรียบเทียบว่า คะแนนจากแบบสอบ X และแบบสอบ Y จะมีความเท่าเทียมกันก็ต่อเมื่อคะแนนทั้งสองอยู่ที่ตำแหน่งเปอร์เซ็นไทล์เดียวกัน และลักษณะการแจกแจงของประชากรผู้สอบต้องมีลักษณะเหมือนกัน วิธีอิกวิเปอร์เซ็นไทล์เป็นการสร้างคะแนนสมมูลระหว่างแบบสอบบนพื้นฐานความเท่าเทียมกันของค่าโมเมนต์ทั้ง 3 ทางสถิติ ซึ่งได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และการแจกแจงของคะแนน จึงเป็นกระบวนการแปลงคะแนนแบบไม่ใช่เส้นตรง (Nonlinear Transformation) ยังผลให้คะแนนระหว่างคะแนนดิบและคะแนนแปลงไม่เป็นเส้นตรง

คะแนน  $X_i$  และ  $Y_i$  ถือเป็นคะแนนสมมูลกัน เมื่อ  $X_{PR_k} = Y_{PR_k}$

### 1.4 วิธีวิเคราะห์พหุระดับ (Multilevel Method)

วิธีการปรับเทียบผลการเรียนด้วยวิธีวิเคราะห์พหุระดับนั้น จะใช้โปรแกรม HLM ปรับผลการเรียนให้อยู่บนมาตรฐานเดียวกัน โดยการปรับเทียบด้วยตัวแปรระดับนักเรียน และโรงเรียนผลที่ได้จากการวิเคราะห์จะเป็นตัวประมาณค่าเฉลี่ยผลการเรียนเฉลี่ยสะสมของวิชาพื้นฐาน และวิชาเฉพาะของแต่ละโรงเรียนที่ควบคุมด้วยตัวแปรระดับนักเรียน และโรงเรียน ในการวิจัยครั้งนี้ มีโมเดลการวิเคราะห์ 2 รูปแบบ ดังนี้

โมเดลการวิเคราะห์ที่ 1

สมการวิเคราะห์ระดับที่ 1: ระดับนักเรียน (Level 1: Between student within school)

$$PGPAX(B/T) = \beta_{0j} + \beta_{1j}(PONET)_j + r_{ij}$$

สมการวิเคราะห์ระดับที่ 2: ระดับผู้เรียน (Level 2: Between school)

$$\beta_{0j} = \gamma_{00} + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + U_{1j}$$

โมเดลการวิเคราะห์ที่ 2

สมการวิเคราะห์ระดับที่ 1: ระดับนักเรียน (Level 1: Between student within school)

$$PGPAX(B/T) = \beta_{0j} + r_{ij}$$

สมการวิเคราะห์ระดับที่ 2: ระดับผู้เรียน (Level 2: Between school)

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(SONET)_j + U_{0j}$$

- ◆ การพัฒนาวิธีการเปรียบเทียบผลการเรียนเฉลี่ยสะสมตามกลุ่มสาระการเรียนรู้โดยใช้คะแนน O-NET  
ของนักเรียนมัธยมศึกษาตอนปลาย

## 2. การตรวจสอบคุณภาพวิธีการเปรียบเทียบผลการเรียนเฉลี่ยสะสมรวมและ จำแนกตามกลุ่มสาระการเรียนรู้โดยใช้คะแนน O-NET ของนักเรียนมัธยมศึกษา ตอนปลาย

### 2.1 ผลการตรวจสอบคุณภาพวิธีการเปรียบเทียบผลการเรียนเฉลี่ยสะสมเฉพาะวิชา พื้นฐาน (Basic) และผลการเรียนเฉลี่ยสะสมทุกวิชา (Total)

การเปรียบเทียบผลการเรียนเฉลี่ยในการศึกษาครั้งนี้ วิเคราะห์ผลต่างระหว่างค่า  
ผลการเรียนเฉลี่ยสะสมที่ปรับแล้วโดยใช้สมการปรับเทียบคะแนนของวิธีการวิเคราะห์ทั้ง 4 วิธีที่ศึกษา  
กับผลการเรียนเฉลี่ยสะสมเดิมที่ยังไม่ได้ปรับ (D) แล้วนำค่าที่ได้จากการคำนวณผลต่างมา  
พิจารณาจำแนกโรงเรียนออกเป็น 3 กลุ่ม คือ โรงเรียนมีลักษณะการให้เกรดแบบปล่อยเกรด (ให้เกรด  
สูงกว่าความสามารถของนักเรียน) แบบกดเกรด (ให้เกรดต่ำกว่าความสามารถของนักเรียน) และ  
แบบปกติ (ให้เกรดสอดคล้องกับความสามารถของนักเรียน) โดยใช้เกณฑ์ 3 เท่าของค่า  
ความคลาดเคลื่อนมาตรฐานของค่า D ผลการวิเคราะห์สรุปได้ดังตารางที่ 1

**ตารางที่ 1** จำนวนและร้อยละของโรงเรียน จำแนกตามวิธีการเปรียบเทียบ และลักษณะการให้เกรด

วิธีการเปรียบเทียบ	ลักษณะการให้เกรด GPAXB				ลักษณะการให้เกรด GPAXT			
	ปล่อยเกรด	ปกติ	ลดเกรด	$\chi^2$	ปล่อยเกรด	ปกติ	ลดเกรด	$\chi^2$
<b>1. วิธีวิเคราะห์ถดถอย</b>				3.681				3.359
กลุ่มสูง (n=100)	38	5	57		40	5	55	
กลุ่มกลาง (n=100)	51	3	46		47	2	51	
กลุ่มต่ำ (n=100)	50	8	42		47	6	47	
รวม รร. ทั้ง 3 กลุ่ม (n=300)	139	16	145		134	13	153	
รวม รร. ทั้งหมด (N=2,583)	1,229 (47.58%)	125 (4.84%)	1,229 (47.58%)		1,207 (46.73%)	142 (5.50%)	1,234 (47.77%)	
<b>2. วิธีเชิงสังตรง</b>				101.016**				106.794**
กลุ่มสูง (n=100)	2	1	97		1	1	98	
กลุ่มกลาง (n=100)	50	5	45		49	5	46	
กลุ่มต่ำ (n=100)	66	3	31		67	0	33	
รวม รร. ทั้ง 3 กลุ่ม (n=300)	118	9	173		117	6	117	
รวม รร. ทั้งหมด (N=2,583)	1,252 (48.47%)	119 (4.61%)	1,212 (46.92%)		1,224 (47.39%)	121 (4.68%)	1,238 (47.93%)	
<b>3. วิธีอควิเปอร์เซ็นไทล์</b>				124.306**				124.852**
กลุ่มสูง (n=100)	10	6	84		11	6	83	
กลุ่มกลาง (n=100)	42	4	54		41	4	55	
กลุ่มต่ำ (n=100)	89	2	9		89	2	9	
รวม รร. ทั้ง 3 กลุ่ม (n=300)	141	12	147		141	12	147	
รวม รร. ทั้งหมด (N=2,583)	1,144 (44.29%)	126 (4.88%)	1,313 (50.83%)		1,134 (43.90%)	152 (5.88%)	1,297 (50.22%)	
<b>4. วิธีวิเคราะห์พหุระดับ</b>								70.578**
<b>Model 1</b>				13.982**				
กลุ่มสูง (n=100)	18	11	71		59	13	28	
กลุ่มกลาง (n=100)	31	8	61		42	11	47	
กลุ่มต่ำ (n=100)	42	6	52		12	2	86	
รวม รร. ทั้ง 3 กลุ่ม (n=300)	91 (30.33%)	25 (8.33%)	184 (61.34%)		113 (37.67%)	26 (8.67%)	161 (53.63%)	
<b>Model 2</b>				9.518*				3.854
กลุ่มสูง (n=100)	34	25	41		43	12	38	
กลุ่มกลาง (n=100)	45	11	44		44	11	45	
กลุ่มต่ำ (n=100)	42	12	46		41	12	47	
รวม รร. ทั้ง 3 กลุ่ม (n=300)	121 (40.33%)	48 (16%)	131 (43.67%)		128 (42.67%)	42 (14.00%)	130 (43.33%)	

\*\*p<.01

\*p<.05

## (1) การเปรียบเทียบผลการเปรียบเทียบผลการเรียนเฉลี่ยสะสมเฉพาะวิชา พื้นฐาน (GPAXB) ด้วยวิธีการวิเคราะห์ที่ต่างกัน

โดยภาพรวมพบว่าโรงเรียนส่วนใหญ่ปล่อยเกรด (ให้เกรดสูงกว่าความสามารถ  
ของนักเรียน) และกตเกรด (ให้เกรดต่ำกว่าความสามารถของนักเรียน) จำนวนไม่ต่างกันมากนัก  
มีโรงเรียนเพียงส่วนน้อยเท่านั้นที่มีลักษณะการให้เกรด GPAXB แบบปกติ (ให้เกรดสอดคล้องกับ  
ความสามารถของนักเรียน)

เมื่อพิจารณาลักษณะการให้เกรด GPAXB โดยวิธีการเปรียบเทียบทั้ง 4 วิธี  
พบว่าไม่มีโรงเรียนที่ให้เกรดแบบกตและปล่อยเกรดในจำนวนใกล้เคียงกันโดยเฉพาะวิธีวิเคราะห์ถดถอย  
และวิธีเชิงเส้นตรง มีจำนวนโรงเรียนที่ให้เกรดแบบกตและปล่อยเกรดในจำนวนใกล้เคียงกันมาก ส่วน  
วิธีอควิเปอร์เซ็นไทล์ และวิธีวิเคราะห์พหุระดับ โดยใช้โมเดลแบบที่ 2 มีจำนวนโรงเรียนที่ให้เกรด  
แบบกตและปล่อยเกรดแตกต่างกันเล็กน้อย สำหรับวิธีวิเคราะห์พหุระดับ โดยใช้โมเดลแบบที่ 1 จำนวน  
โรงเรียนที่ให้เกรดแบบกตและปล่อยเกรดมีความแตกต่างโดยโมเดลแบบที่ 1 มีโรงเรียนที่กตเกรด  
คิดเป็นร้อยละ 61.34 และโรงเรียนที่ปล่อยเกรด คิดเป็นร้อยละ 30.33

เมื่อพิจารณาผลการเปรียบเทียบทั้ง 4 วิธี พบว่าโรงเรียนที่มีคุณภาพสูงจะมี  
ลักษณะการให้เกรดแบบกตเกรด (41% - 97%) มากกว่าปล่อยเกรด (2% - 38%)

ส่วนโรงเรียนที่มีคุณภาพปานกลางจะมีลักษณะการให้เกรด GPAXB แบบ  
ปล่อยเกรด และกตเกรด จำนวนใกล้เคียงกัน เมื่อเปรียบเทียบด้วยวิธีวิเคราะห์ถดถอย วิธีเชิงเส้นตรง  
และวิธีวิเคราะห์พหุระดับ ด้วยโมเดลแบบที่ 2 ขณะที่เมื่อเปรียบเทียบด้วยวิธีอควิเปอร์เซ็นไทล์ และวิธี  
วิเคราะห์พหุระดับ ด้วยโมเดลแบบที่ 1 พบว่ามีโรงเรียนที่มีลักษณะการให้เกรดแบบกตเกรด  
(51% - 61%) มากกว่าปล่อยเกรด (31% - 42%)

โรงเรียนที่มีคุณภาพต่ำจะมีลักษณะการให้เกรด GPAXB แบบปล่อยเกรด  
(50% - 89%) มากกว่ากตเกรด (9% - 42%) เมื่อเปรียบเทียบด้วยวิธีวิเคราะห์ถดถอย วิธีเชิงเส้นตรง  
และวิธีอควิเปอร์เซ็นไทล์ ขณะที่เมื่อเปรียบเทียบโดยวิธีวิเคราะห์พหุระดับ ด้วยโมเดลแบบที่ 1 และแบบที่  
2 พบว่ามีโรงเรียนที่มีลักษณะการให้เกรดแบบกตเกรด (46% - 52%) มากกว่าปล่อยเกรด (42%)

## (2) การเปรียบเทียบผลการเปรียบเทียบผลการเรียนเฉลี่ยสะสมทุกวิชา (GPAXT) ด้วยวิธีการวิเคราะห์ที่ต่างกัน

โดยภาพรวมพบว่าโรงเรียนส่วนใหญ่ปล่อยเกรด (ให้เกรดสูงกว่าความสามารถ  
ของนักเรียน) และกตเกรด (ให้เกรดต่ำกว่าความสามารถของนักเรียน) จำนวนไม่ต่างกันมากนัก

มีโรงเรียนเพียงส่วนน้อยเท่านั้นที่มีลักษณะการให้เกรด GPAXT แบบปกติ (ให้เกรดสอดคล้องกับความสามารถของนักเรียน)

เมื่อพิจารณาถึงลักษณะการให้เกรด GPAXT โดยวิธีการเปรียบเทียบทั้ง 4 วิธี พบว่ามีโรงเรียนที่ให้เกรดแบบกตและปล่อยเกรดในจำนวนใกล้เคียงกัน โดยเฉพาะวิธีวิเคราะห์ถดถอย และวิธีเชิงเส้นตรง มีจำนวนโรงเรียนที่ให้เกรดแบบกตและปล่อยเกรดในจำนวนใกล้เคียงกันมาก ส่วนวิธีอิกวิเปอร์เซ็นไทล์ และวิธีวิเคราะห์พหุระดับ โดยใช้โมเดลแบบที่ 2 มีจำนวนโรงเรียนที่ให้เกรดแบบกตและปล่อยเกรดแตกต่างกันเล็กน้อย สำหรับวิธีวิเคราะห์พหุระดับ โดยใช้โมเดลแบบที่ 1 จำนวนโรงเรียนที่ให้เกรดแบบกตและปล่อยเกรดมีความแตกต่าง โดยโมเดลแบบที่ 1 มีโรงเรียนที่กตเกรด คิดเป็นร้อยละ 53.63 และโรงเรียนที่ปล่อยเกรด คิดเป็นร้อยละ 37.67

เมื่อพิจารณาผลการเปรียบเทียบทั้ง 4 วิธี พบว่าโรงเรียนที่มีคุณภาพสูงจะมีลักษณะการให้เกรด GPAXT แบบปล่อยเกรด และกตเกรด จำนวนใกล้เคียงกัน เมื่อเปรียบเทียบด้วยวิธีวิเคราะห์ถดถอย และวิธีวิเคราะห์พหุระดับ ด้วยโมเดลแบบที่ 2 ขณะที่เมื่อเปรียบเทียบด้วยวิธีเชิงเส้นตรง และวิธีอิกวิเปอร์เซ็นไทล์ พบว่ามีโรงเรียนที่มีลักษณะการให้เกรดแบบกตเกรด (83% - 98%) มากกว่าปล่อยเกรด (1% - 11%) ทั้งนี้ เมื่อเปรียบเทียบโดยวิธีวิเคราะห์พหุระดับ ด้วยโมเดลแบบที่ 1 พบว่ามีโรงเรียนที่มีลักษณะการให้เกรดแบบปล่อยเกรด (59%) มากกว่ากตเกรด (28%)

ส่วนโรงเรียนที่มีคุณภาพปานกลางจะมีลักษณะการให้เกรด GPAXT แบบปล่อยเกรด และกตเกรด จำนวนใกล้เคียงกัน เมื่อเปรียบเทียบด้วยวิธีการวิเคราะห์ทั้ง 4 วิธี

โรงเรียนที่มีคุณภาพต่ำจะมีลักษณะการให้เกรด GPAXT แบบปล่อยเกรด (67% - 89%) มากกว่ากตเกรด (9% - 33%) เมื่อเปรียบเทียบด้วยวิธีเชิงเส้นตรง และวิธีอิกวิเปอร์เซ็นไทล์ ขณะที่เมื่อเปรียบเทียบโดยวิธีวิเคราะห์ถดถอย และวิธีวิเคราะห์พหุระดับ ด้วยโมเดลแบบที่ 2 พบว่ามีโรงเรียนที่มีลักษณะการให้เกรดแบบกตเกรด และปล่อยเกรด จำนวนใกล้เคียงกัน ทั้งนี้ในทางกลับกัน เมื่อเปรียบเทียบด้วยวิธีวิเคราะห์พหุระดับ ด้วยโมเดลแบบที่ 1 พบว่ามีโรงเรียนที่มีลักษณะการให้เกรดแบบกตเกรด (86%) มากกว่าปล่อยเกรด (12%)

### (3) การเปรียบเทียบวิธีการวิเคราะห์

เมื่อเปรียบเทียบวิธีการวิเคราะห์ทั้ง 4 วิธี พบว่าวิธีการเชิงเส้นตรง และวิธีอิกวิเปอร์เซ็นไทล์ จะให้ค่าการเปรียบเทียบใกล้เคียงกัน และสอดคล้องกับข้อมูลเชิงประจักษ์มากกว่าวิธีวิเคราะห์ถดถอย และวิธีวิเคราะห์พหุระดับ คือ โรงเรียนที่มีคุณภาพสูงมีลักษณะการให้เกรดแบบกตเกรดมากกว่าปล่อยเกรด โรงเรียนที่มีคุณภาพปานกลางมีลักษณะการให้เกรดแบบกตเกรดและปล่อยเกรดในจำนวนที่ใกล้เคียงกัน และโรงเรียนที่มีคุณภาพต่ำมีลักษณะการให้เกรดแบบ

ปล่อยเกรดมากกว่าตเกรด ดังนั้น วิธีเชิงเส้นตรง และวิธีอควิเปอร์เซ็นไทล์ จึงเหมาะสำหรับการเปรียบเทียบผลการเรียน อย่างไรก็ตามในทางปฏิบัติ วิธีเชิงเส้นตรง มีกระบวนการในการเปรียบเทียบผลการเรียนที่ไม่ซับซ้อน สะดวกในการนำไปใช้จริงมากกว่าวิธีอควิเปอร์เซ็นไทล์

#### **(4) การเปรียบเทียบการนำผลการเรียนเฉลี่ยสะสมเฉพาะวิชาพื้นฐาน (GPAXB) และผลการเรียนเฉลี่ยสะสมรวมทุกวิชา (GPAXT) มาเปรียบเทียบ**

เมื่อเปรียบเทียบการนำผลการเรียนเฉลี่ยสะสมเฉพาะวิชาพื้นฐาน (GPAXB) และผลการเรียนเฉลี่ยสะสมรวมทุกวิชา (GPAXT) มาเปรียบเทียบ โดยใช้วิธีการเปรียบเทียบทั้ง 4 วิธี พบว่าผลการเปรียบเทียบที่ได้มีความสัมพันธ์กันสูง และเป็นไปในทิศทางเดียวกัน ดังนั้น ในการเปรียบเทียบผลจึงสามารถเลือกใช้ผลการเรียนเฉลี่ยสะสมทั้ง 2 แบบ อย่างไรก็ตามในทางปฏิบัติ การนำผลการเรียนเฉลี่ยสะสมรวมทุกวิชามาเปรียบเทียบ (GPAXT) จะสะดวกในการนำไปใช้จริงมากกว่าการเปรียบเทียบจากผลการเรียนเฉลี่ยสะสมเฉพาะวิชาพื้นฐาน (GPAXB)

### **2.2 ผลการตรวจสอบคุณภาพวิธีการเปรียบเทียบผลการเรียนเฉลี่ยสะสมจำแนกตามกลุ่มสาระการเรียนรู้**

การเปรียบเทียบผลการเรียนเฉลี่ยสะสมในแต่ละกลุ่มสาระการเรียนรู้ ใช้วิธีการเปรียบเทียบ 2 วิธี คือ วิธีเชิงเส้นตรง และวิธีอควิเปอร์เซ็นไทล์ ซึ่งเป็นวิธีที่ให้ค่าการเปรียบเทียบใกล้เคียงกัน และสอดคล้องกับสภาพความเป็นจริงมากกว่าวิธีอื่น ๆ โดยผลการเปรียบเทียบจำแนกตามกลุ่มสาระการเรียนรู้ต่าง ๆ ได้ดังตารางที่ 2



**ตารางที่ 2** การเปรียบเทียบลักษณะการให้เกรด ด้วยวิธีเชิงเส้นตรงและวิธีอิกวิเปอร์เซ็นไทล์ จำแนกตามกลุ่มสาระการเรียนรู้

ลักษณะการให้เกรด		จำนวนโรงเรียน		Friedman Test
		จำแนกตามลักษณะการให้เกรด		
		วิธีเชิงเส้นตรง	วิธีอิกวิเปอร์เซ็นไทล์	
ภาษาไทย	ปล่อยเกรด	1,125	1,181	22.531**
	ปกติ	113	96	
	กวดเกรด	1,245	1,306	
	รวม	2,583	2,583	
คณิตศาสตร์	ปล่อยเกรด	1,247	1,172	28.636**
	ปกติ	119	108	
	กวดเกรด	1,217	1,303	
	รวม	2,583	2,583	
วิทยาศาสตร์	ปล่อยเกรด	1,252	1,165	46.615**
	ปกติ	126	150	
	กวดเกรด	1,205	1,268	
	รวม	2,583	2,583	
สังคมศึกษา ศาสนาและ วัฒนธรรม	ปล่อยเกรด	1,283	1,240	42.568**
	ปกติ	107	107	
	กวดเกรด	1,193	1,236	
	รวม	2,583	2,583	
ภาษาต่างประเทศ	ปล่อยเกรด	1,206	1,172	3.316 (p=0.069)
	ปกติ	139	96	
	กวดเกรด	1,238	1,315	
	รวม	2,583	2,583	

\*\*p < 0.01

เมื่อเปรียบเทียบผลการเรียนเฉลี่ยสะสมในกลุ่มสาระการเรียนรู้ภาษาไทยคณิตศาสตร์ วิทยาศาสตร์ และสังคมศึกษา ศาสนา และวัฒนธรรม ด้วยวิธีเชิงเส้นตรง และวิธีอิคิวเปอร์เซ็นไทล์ พบว่าการแบ่งกลุ่มโรงเรียนเป็น 3 ลักษณะ คือ ปล่อยเกรด ปกติ และกตเกรด ตามผลการเปรียบเทียบที่ได้จากทั้ง 2 วิธีมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ขณะที่การเปรียบเทียบผลการเรียนเฉลี่ยสะสมในกลุ่มสาระการเรียนรู้ภาษาต่างประเทศ ด้วยวิธีเชิงเส้นตรง และวิธีอิคิวเปอร์เซ็นไทล์ พบว่าการแบ่งกลุ่มโรงเรียนตามผลการเปรียบเทียบที่ได้จากทั้ง 2 วิธีไม่แตกต่างกัน

ในการเปรียบเทียบคุณภาพของวิธีการเปรียบเทียบในครั้งนี พิจารณาจากค่าสัมประสิทธิ์การทำนาย ( $R^2$ ) ของคะแนน O-NET ตามกลุ่มสาระการเรียนรู้ กับผลการเรียนเฉลี่ยสะสมที่ปรับแล้ว โดยมีเกณฑ์ในการตัดสินใจ คือ  $R^2$  ที่มีค่าสูงกว่า แสดงว่าผลการเปรียบเทียบสามารถทำนายค่าคะแนนเกณฑ์ได้ใกล้เคียงมากกว่า ผลการเปรียบเทียบคุณภาพของการเปรียบเทียบโดยวิธีเชิงเส้นตรง และวิธีอิคิวเปอร์เซ็นไทล์ พบว่าวิธีเชิงเส้นตรงมีคุณภาพมากกว่าวิธีอิคิวเปอร์เซ็นไทล์ อย่างไรก็ตามผลการเปรียบเทียบที่ได้จากวิธีการทั้ง 2 วิธี ไม่แตกต่างกันมากนัก และเป็นไปในทิศทางเดียวกัน สอดคล้องกับข้อมูลเชิงประจักษ์

### 3. ผลการจัดกลุ่มมาตรฐานโรงเรียน

#### 3.1 ผลการจัดกลุ่มโรงเรียนตามมาตรฐานการประเมินผลการเรียนรู้โดยภาพรวม

ผลการเปรียบเทียบผลการเรียนเฉลี่ยสะสมทุกวิชา (GPAXT) โดยใช้คะแนน O-NET ด้วยวิธีการเปรียบเทียบแบบเชิงเส้นตรง โดยภาพรวมพบว่าโรงเรียนส่วนใหญ่ปล่อยเกรด (ให้เกรดสูงกว่าความสามารถของนักเรียน) และกตเกรด (ให้เกรดต่ำกว่าความสามารถของนักเรียน) จำนวนไม่ต่างกันมากนัก มีโรงเรียนเพียงส่วนน้อยเท่านั้นที่มีลักษณะการให้เกรดแบบปกติ (ให้เกรดสอดคล้องกับความสามารถของนักเรียน) ทั้งนี้พบว่ามีโรงเรียนที่ให้เกรด GPAXT แบบปล่อยเกรด คิดเป็นร้อยละ 47.43 ให้เกรดแบบกตเกรด ร้อยละ 47.81 และให้เกรดแบบปกติ ร้อยละ 4.76

เมื่อพิจารณาโรงเรียนในกลุ่มปล่อยเกรด พบว่าเป็นโรงเรียนในกลุ่มปล่อยเกรดมากมีจำนวน 14 โรงเรียน คิดเป็นร้อยละ 1.41 กลุ่มปล่อยเกรดค่อนข้างมาก มีจำนวน 62 โรงเรียน คิดเป็นร้อยละ 5.06 กลุ่มปล่อยเกรดค่อนข้างน้อย มีจำนวน 334 โรงเรียน คิดเป็นร้อยละ 27.27 และ กลุ่มปล่อยเกรดน้อย มีจำนวน 815 โรงเรียน คิดเป็นร้อยละ 66.53 เมื่อพิจารณาโรงเรียนในกลุ่มกตเกรด พบว่าโรงเรียนในกลุ่มกตเกรดมากมีจำนวน 1 โรงเรียน คิดเป็นร้อยละ 0.08 กลุ่มกตเกรดค่อนข้างมาก มีจำนวน 19 โรงเรียน คิดเป็นร้อยละ 1.54 กลุ่มกตเกรดค่อนข้างน้อย มีจำนวน 211 โรงเรียน คิดเป็นร้อยละ 17.09 และกลุ่มกตเกรดน้อย มีจำนวน 1,004 โรงเรียน คิดเป็นร้อยละ 81.30 สามารถแสดงผลการวิเคราะห์นี้ได้ดังตารางที่ 3

**ตารางที่ 3** ผลการจัดกลุ่มโรงเรียนตามลักษณะการให้เกรด GPAXT หลังการปรับเทียบด้วยค่าเฉลี่ยคะแนน O-NET รวมทุกวิชา ของนักเรียนในโรงเรียน: วิธีเชิงเส้นตรง

ลักษณะการให้เกรด	จำนวนโรงเรียน	ระดับการปล่อยเกรด	จำนวนโรงเรียน
กลุ่มที่ 1 ปล่อยเกรด	1,225(47.43%)	ปล่อยมาก	14 (1.14%)
		ปล่อยค่อนข้างมาก	62 (5.06%)
		ปล่อยค่อนข้างน้อย	334 (27.27%)
		ปล่อยน้อย	815 (66.53%)
กลุ่มที่ 2 ปกติ	123(4.76%)		
กลุ่มที่ 3 กดเกรด	1,235(47.81%)	กดมาก	1 (0.08%)
		กดค่อนข้างมาก	19 (1.54%)
		กดค่อนข้างน้อย	211 (17.09%)
		กदन้อย	1,004 (81.30%)

**3.2 การจัดกลุ่มโรงเรียนตามมาตรฐานการประเมินผลการเรียนรู้ จำแนกตามกลุ่มสาระการเรียนรู้**

โดยภาพรวมพบว่า โรงเรียนส่วนใหญ่ปล่อยเกรด (ให้เกรดสูงกว่าความสามารถของนักเรียน) และกดเกรด (ให้เกรดต่ำกว่าความสามารถของนักเรียน) จำนวนไม่ต่างกันมากนัก มีโรงเรียนเพียงส่วนน้อยเท่านั้นที่มีลักษณะการให้เกรดแบบปกติ (ให้เกรดสอดคล้องกับความสามารถของนักเรียน) เมื่อพิจารณาเปรียบเทียบลักษณะการประเมินผลการเรียนรู้ทั้ง 3 แบบ ในกลุ่มสาระการเรียนรู้ทั้ง 5 กลุ่ม ได้แก่ ภาษาไทย คณิตศาสตร์ วิทยาศาสตร์ สังคมศึกษา ศาสนา และวัฒนธรรม และภาษาต่างประเทศ พบว่า กลุ่มสาระการเรียนรู้ที่มีการปล่อยเกรดมากที่สุด คือ กลุ่มสาระการเรียนรู้ สังคมศึกษา ศาสนา และวัฒนธรรม โดยมีโรงเรียนที่ปล่อยเกรดอยู่ร้อยละ 49.67 ส่วนกลุ่มสาระการเรียนรู้ที่มีการกดเกรดมากที่สุด คือ กลุ่มสาระการเรียนรู้ภาษาไทย โดยมีโรงเรียนที่กดเกรดอยู่ร้อยละ 48.20 สำหรับกลุ่มสาระการเรียนรู้ที่มีให้เกรดแบบปกติมากที่สุด คือ กลุ่มสาระการเรียนรู้ ภาษาต่างประเทศ โดยมีโรงเรียนที่ให้เกรดแบบปกติอยู่ร้อยละ 5.38

ทั้งนี้ เมื่อพิจารณาลักษณะการประเมินผลการเรียนรู้ในกลุ่มสาระการเรียนรู้ทั้ง 5 กลุ่ม ภายในโรงเรียนแต่ละแห่ง พบว่า ลักษณะการประเมินผลการเรียนรู้ในโรงเรียนแต่ละแห่งมีความหลากหลาย โดยโรงเรียนบางแห่งให้เกรดแบบปล่อยเกรดทุกกลุ่มสาระการเรียนรู้ บางแห่งให้เกรดแบบกดเกรดทุกกลุ่มสาระการเรียนรู้ ขณะที่บางแห่งให้เกรดแบบกดเกรดบางกลุ่มสาระการเรียนรู้และปล่อยเกรดบางกลุ่มสาระการเรียนรู้ ซึ่งผลการจัดกลุ่มโรงเรียนตามมาตรฐานการ

ประเมินผลการเรียนรู้ เมื่อวิเคราะห์ค่าแนวโน้มเข้าสู่ส่วนกลาง โดยพิจารณาจากค่าฐานนิยม (Mode) ของการประเมินผลการเรียนรู้ในโรงเรียนแต่ละแห่ง แสดงให้เห็นภาพสรุปลักษณะการประเมินผลการเรียนรู้ของโรงเรียนทั้ง 2,583 โรงเรียน พบว่าฐานนิยมของลักษณะการประเมินผลการเรียนรู้มีทั้งแบบปล่อยเกรด กดเกรด ปกติ ปล่อยเกรดและปกติ ปล่อยเกรดและกดเกรด รวมถึงปกติและกดเกรดคิดเป็นร้อยละ 48.32, 47.12, 0.31, 0.54, 3.06 และ 0.66 ตามลำดับ

ในส่วนของโรงเรียนที่มีลักษณะการประเมินผลการเรียนรู้รวมทั้ง 5 กลุ่มสาระการเรียนรู้เป็นแบบปล่อยเกรดนั้น พบว่า โรงเรียนส่วนใหญ่มีลักษณะการให้เกรดแบบปล่อยเกรดทุกวิชามากที่สุด คิดเป็นร้อยละ 17.27 สอดคล้องกันกับกลุ่มโรงเรียนที่มีลักษณะการประเมินผลการเรียนรู้รวมทั้ง 5 กลุ่มสาระการเรียนรู้เป็นแบบกดเกรด ซึ่งพบว่าโรงเรียนส่วนใหญ่มีลักษณะการให้เกรดแบบกดเกรดทุกวิชามากที่สุด คิดเป็นร้อยละ 16.03

### 3.3 การจัดกลุ่มโรงเรียนตามมาตรฐานการเรียนรู้ โดยภาพรวม

การจัดกลุ่มมาตรฐานโรงเรียนโดยการพิจารณาจากผลการเรียนเฉลี่ยสะสมทุกวิชา (SGPAXT) ทั้งก่อนและหลังการเปรียบเทียบผลการเรียน โดยใช้เกณฑ์คุณภาพโรงเรียน 5 ระดับ คือ คุณภาพของโรงเรียนระดับปรับปรุง (D) พอใช้ (C) ดี (B) ดีมาก (A) และดีเยี่ยม (A+) ปรากฏผลดังตารางที่ 4

ในตารางที่ 4 แสดงให้เห็นว่า คุณภาพโรงเรียน ก่อนและหลังการเปรียบเทียบผลการเรียนเฉลี่ยสะสม แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ทั้งนี้เมื่อพิจารณาคุณภาพโรงเรียนจากผลการเรียนเฉลี่ยสะสมของโรงเรียนก่อนการเปรียบเทียบ พบว่าโรงเรียนส่วนใหญ่มีคุณภาพในระดับดี (B) คิดเป็นร้อยละ 49.3 แต่เมื่อพิจารณาคุณภาพโรงเรียนจากผลการเรียนเฉลี่ยสะสมของโรงเรียนที่เปรียบเทียบแล้ว พบว่าโรงเรียนส่วนใหญ่มีคุณภาพในระดับพอใช้ (C) คิดเป็นร้อยละ 60.8 นอกจากนี้ยังพบว่าหลังการเปรียบเทียบผลการเรียนเฉลี่ยสะสมแล้ว โรงเรียนในกลุ่ม คุณภาพดีเยี่ยม (A+) ดีมาก (A) และพอใช้ (C) มีจำนวนเพิ่มขึ้น ส่วนโรงเรียนในกลุ่มคุณภาพดี (B) และปรับปรุง (D) มีจำนวนลดลง

**ตารางที่ 4** ผลเปรียบเทียบคุณภาพโรงเรียน ก่อนและหลังปรับเทียบผลการเรียนเฉลี่ยสะสม ด้วยวิธีเชิงเส้นตรง

คุณภาพของ โรงเรียน	ช่วงของคะแนน SGPAXT	SGPAXT ก่อนปรับ		SGPAXT หลังปรับ		Friedman Test
		จำนวน (โรงเรียน)	ร้อยละ	จำนวน (โรงเรียน)	ร้อยละ	
ปรับปรุง (D)	[0.00, 1.99]	56	2.2	4	0.2	44.260**
พอใช้ (C)	[2.00, 2.49]	1144	44.3	1572	60.8	
ดี (B)	[2.50, 2.99]	1274	49.3	846	32.8	
ดีมาก (A)	[3.00, 3.49]	107	4.1	129	5.0	
ดีเยี่ยม (A <sup>+</sup> )	[3.50, 4.00]	2	0.1	32	1.2	
	รวม	2583	100	2583	100	

\*\*p<.01

### อภิปรายผลการวิจัย

การศึกษาเรื่องการพัฒนาวิธีการปรับเทียบผลการเรียนเฉลี่ยสะสมตามกลุ่มสาระการเรียนรู้ โดยใช้คะแนน O-NET ของนักเรียนมัธยมศึกษาตอนปลายนั้นสามารถอภิปรายผลการวิจัยได้ดังนี้

1. จากผลการวิจัย พบว่า วิธีเชิงเส้นตรงและวิธีอิกวิเปอร์เซ็นไทล์เป็นวิธีการที่มีความเหมาะสมสำหรับปรับเทียบผลการเรียนเฉลี่ยสะสมตามกลุ่มสาระการเรียนรู้ โดยใช้คะแนน O-NET ของนักเรียนมัธยมศึกษาตอนปลายกว่าวิธีที่ใช้ในการวิจัยทั้ง 4 วิธี ผลการปรับเทียบจากทั้งสองวิธี มีความใกล้เคียงกัน และสอดคล้องกับสภาพความเป็นจริงมากกว่าการวิเคราะห์ถดถอย และการวิเคราะห์พหุระดับ ที่อิงกับหลักตรรกะของสถานการณ์จริงคือ โรงเรียนที่มีคุณภาพการจัดการเรียนรู้สูงจะมีธรรมชาติของการให้ผลการเรียนในลักษณะของการกตเกรด มากกว่าการปล่อยเกรด เพราะโรงเรียนสามารถจัดสอบคัดเลือกนักเรียนที่มีความรู้เข้ามาศึกษาได้ การให้เกรดของครูในกรณีของนักเรียนที่มีความสามารถสูงจึงถูกจำกัดอยู่ในช่วง (Range restriction) เข้ามาเกี่ยวข้องกับทำให้เกรดส่วนที่วิธีวิเคราะห์ถดถอย และการวิเคราะห์พหุระดับให้ผลการวิเคราะห์ที่แตกต่างจากวิธีเชิงเส้นตรงและวิธีอิกวิเปอร์เซ็นไทล์นั้น โดยเทคนิคทางสถิติที่ใช้สำหรับการวิเคราะห์มีความแตกต่างกันโดยวิธีวิเคราะห์ถดถอย และการวิเคราะห์พหุระดับจะใช้การวิเคราะห์ถดถอยพหุเหมือนกัน แต่ทั้งนี้ การวิเคราะห์พหุระดับจะมีการใช้สถิติวิเคราะห์ที่ซับซ้อนมากกว่าการวิเคราะห์ถดถอยเพราะใช้สถิติ OLS ในการวิเคราะห์ระดับที่ 1 และสถิติ EB ในการวิเคราะห์ระดับที่ 2 ค่าประมาณจึงปรับทั้งระดับผู้เรียนและระดับโรงเรียน ผลการปรับเทียบจึงแตกต่างกันออกไป (ศิริชัย กาญจนวาสิ, 2548)

ผลจากการศึกษาที่พบว่าวิธีเชิงเส้นตรง และวิธีอิกวิเปอร์เซ็นไทล์ให้ผลการวิเคราะห์ใกล้เคียงกันนั้นสอดคล้องกับการศึกษาของ Kolen (1981 อ้างถึงใน ภัทราพร เกษสังข์, 2546) ที่ประเมินและเปรียบเทียบวิธีการเชื่อมโยงคะแนนเชิงเส้นตรง วิธีการเชื่อมโยงคะแนนแบบอิกวิเปอร์เซ็นไทล์ และวิธีการใช้โค้งลักษณะข้อสอบโดยใช้แบบสอบ SAT โดยใช้โค้งลักษณะข้อสอบที่ศึกษาทั้ง 1, 2 และ 3 พารามิเตอร์ แต่ละวิธีเทียบด้วยค่าประมาณจากค่าจริง (Estimated True Score Equating) และเทียบด้วยการประมาณจากค่าสังเกต (Estimated Observed Score Equating) ใช้ข้อมูลจากโครงการ IOWA Test of Education Development (ITED) ค.ศ. 1978 ศึกษา 2 รูปแบบ คือ เทียบจากฟอร์มที่มีความยากเท่าเทียมกัน และเทียบกับฟอร์มที่มีความยากแตกต่างกัน กลุ่มตัวอย่างเป็นนักเรียนระดับที่ 1 คือ ระดับเกรด 9 และ 10 ระดับ 2 คือ ระดับเกรด 11 และ 12 จำนวนทั้งหมด 10,728 คน จากโรงเรียนในรัฐไอโอวา 34 แห่ง การศึกษานี้ใช้เกณฑ์ Cross-Validation Criterion ผลการวิจัยพบว่า การเชื่อมโยงคะแนนทั้งหมดแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และ .05 การเชื่อมโยงคะแนนโดยใช้วิธีอิกวิเปอร์เซ็นไทล์เหมาะสมที่สุดกับแบบสอบที่มีความยากแตกต่างกัน ส่วนการเชื่อมโยงคะแนนโดยใช้โค้งลักษณะข้อสอบ พบว่า กรณีใช้พารามิเตอร์เดียว ยังเป็นวิธีที่ไม่ถูกต้องสำหรับแบบสอบที่ต่างกัน เนื่องมาจากการเดาเกิดขึ้นมาก ส่วนกรณีใช้ 3 พารามิเตอร์ มีปัญหาที่เกิดขึ้นอย่างหลีกเลี่ยงไม่ได้ คือ การประมาณค่าไม่ถูกต้องของ Lower Asymptote Parameter

ส่วน Kagan และ Stock (1980) ได้ศึกษาความเท่าเทียมกันของแบบสอบชุด Miller Analogies Test และ The Graduate Record Examination โดยใช้วิธีการปรับเทียบเชิงเส้นตรง และวิธีการทำนาย กลุ่มประชากรจากมหาวิทยาลัยอริโซนา จำนวน 18,017 คน ใช้กลุ่มตัวอย่างที่ได้ทำแบบสอบทั้ง 2 ชุด ผลการวิจัยพบว่าค่าความเบี่ยงเบนมาตรฐานของคะแนนที่แปลงด้วยวิธีเชิงเส้นตรงมีค่าใกล้เคียงกับคะแนนชุดเดิม ซึ่งผลการวิจัยดังกล่าวเป็นไปในทิศทางเดียวกันกับผลการศึกษาของ Cook และ Eignor (1981) ได้ศึกษาผลการเชื่อมโยงคะแนนตามทฤษฎีตอบสนองข้อสอบเปรียบเทียบกับ การปรับเทียบแบบดั้งเดิม 2 วิธี คือ การปรับเทียบเชิงเส้นตรง และการปรับเทียบแบบอิกวิเปอร์เซ็นไทล์ การปรับเทียบตามทฤษฎี IRT โมเดล 3 พารามิเตอร์ ซึ่งประมาณค่าพารามิเตอร์โดยใช้โปรแกรม LOGIST ด้วยวิธีความเป็นไปได้สูงสุด (Maximum Likelihood Procedure) สำหรับการประมาณค่าที่อยู่ต่ำกว่าระดับการเดา ใช้วิธีเชิงเส้นตรง (Linear Interpolation) ในการปรับเทียบ ศึกษาทั้งสองกรณี คือ มีแบบสอบร่วม และไม่มีแบบสอบร่วม และผลการวิจัยของ Glowacki (1991) ที่ตรวจสอบโมเดลของการเชื่อมโยงคะแนนที่มีความเหมาะสมกับการสอบของบัณฑิตวิทยาลัยแห่งมหาวิทยาลัยอลาบามา ผลการศึกษาทั้งสองงานวิจัยพบว่าวิธีการเชื่อมโยงคะแนนแบบดั้งเดิม กับวิธีการปรับเทียบตามทฤษฎี IRT มีความสอดคล้องกันมาก เนื่องจากโดยความจริงแล้ว การแจกแจงคะแนนดิบจากแบบสอบทั้งสองชุดมีรูปร่างคล้ายคลึงกันมาก จึงทำให้รูปแบบเชิงเส้นตรงและเชิงเส้นโค้งหรืออิกวิเปอร์เซ็นไทล์ มีผลใกล้เคียงกัน

สำหรับผลการวิจัยในประเทศไทยก็ให้ผลการศึกษาคัดค้านเช่นเดียวกัน สุภาภรณ์ คงทวี (2541) ได้ทำการเปรียบเทียบประสิทธิภาพของวิธีการปรับเทียบระหว่างวิธีอิกคิวเปอร์เซ็นต์เชิงเส้นตรง และ IRT ในการปรับแก้เฉลี่ยสะสมระดับมัธยมศึกษาตอนปลาย ด้วยคะแนนสอบเข้ามหาวิทยาลัย กลุ่มตัวอย่างที่ใช้ในการวิจัยคือนิสิตนักศึกษาชั้นปีที่ 1 ปีการศึกษา 2540 ซึ่งจบชั้นมัธยมศึกษาปีที่ 6 บริบูรณ์ ในมหาวิทยาลัยของรัฐ ได้แก่ จุฬาลงกรณ์มหาวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์ มหาวิทยาลัยศิลปากร และมหาวิทยาลัยศรีนครินทรวิโรฒ โดยวิเคราะห์แยกเป็น 6 กลุ่ม ได้แก่ กลุ่มสาขาวิชาวิทยาศาสตร์ วิศวกรรมศาสตร์ สถาปัตยกรรมศาสตร์ พาณิชยศาสตร์และการบัญชี มนุษยศาสตร์-สังคมศาสตร์ (แผนการสอบศิลป์) และมนุษยศาสตร์-สังคมศาสตร์ (แผนการสอบวิทยาศาสตร์) ผลการวิจัย พบว่า เมื่อใช้คะแนนเฉลี่ยสะสมระดับมหาวิทยาลัยเป็นเกณฑ์ วิธีปรับเชิงเส้นตรงมีประสิทธิภาพในการพยากรณ์สูงกว่าวิธีอิกคิวเปอร์เซ็นต์

2. ผลจากการศึกษาพบว่าโรงเรียนให้ผลการเรียนของนักเรียนไม่สะท้อนกับผลการเรียนรู้ที่แท้จริงของผู้เรียน หากใช้หลักการของทฤษฎีการทดสอบแบบดั้งเดิม (Classical test theory) มาอธิบาย แสดงให้เห็นว่าผลการวัดมีความคลาดเคลื่อนเกิดขึ้นมาก ซึ่งสอดคล้องกับผลการวิเคราะห์เพื่อการจำแนกโรงเรียนตามลักษณะการประเมินผลการเรียนรู้ด้วยวิธีเชิงเส้นตรง พบว่าโรงเรียนมีลักษณะการประเมินผลการเรียนรู้ที่สอดคล้องกับระดับความสามารถของนักเรียนเพียงร้อยละ 4.76 นั่นหมายถึงโรงเรียนอีกร้อยละ 95.24 เป็นโรงเรียนที่อยู่ในกลุ่มมีลักษณะการประเมินผลการเรียนรู้ที่ไม่สอดคล้องกับระดับความสามารถของนักเรียน ทั้งนี้สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) และโรงเรียนควรให้ความสำคัญกับวิธีการประเมินผลการเรียนรู้ของนักเรียนให้มากขึ้น ทั้งนี้เพราะคะแนนที่ได้จะเป็นข้อมูลสำคัญเพื่อนำไปสู่การตัดสินใจทางการศึกษาทั้งระดับรายบุคคล ระดับสถาบัน รวมถึงระดับชาติ ในการนำคะแนนจากแหล่งต่างๆ มา เพื่อให้ผลที่ได้ต้องมีความเสมอภาค และสามารถเปรียบเทียบกันได้ของคะแนน โดยถือว่าเป็นประเด็นสำคัญของการทดสอบ (Cook และ Eignor, 1991) ทั้งนี้ผลจากการวิจัยยังพบว่าการให้ผลการเรียนของนักเรียนมีความหลากหลายเมื่อจำแนกตามกลุ่มสาระการเรียนรู้ มีโรงเรียนร้อยละ 17.27 ที่ปล่อยเกรดทุกวิชา และจำนวนร้อยละ 16.03 ที่ปล่อยเกรดทุกวิชา ซึ่งกรณีนี้โรงเรียนควรมีการดำเนินงานประเมินผลการเรียนรู้ให้เป็นกระบวนการที่มีมาตรฐาน อาศัยการวัดที่ครอบคลุม หลากหลาย มีผลการวัดที่แม่นยำ เชื่อถือได้ มีวิธีการแปลความหมายของการวัดที่เหมาะสม และใช้เกณฑ์การตัดสินคุณค่าที่มีความยุติธรรม

ศิริชัย กาญจนวาสี (2550) เสนอแนวทาง/วิธีการในการวัดผลและการกำหนดเกรดที่เหมาะสมไว้ดังนี้ (1) ปัจจัยที่ใช้กำหนดเกรด ควรประกอบด้วยปัจจัยหลักของความรู้ความสามารถหรือทักษะของผู้เรียนตามจุดมุ่งหมายของการเรียนรู้เป็นสำคัญ และปัจจัยเสริม เช่น การมีส่วนร่วมในกิจกรรม

การเรียนการสอน เจตคติต่อการเรียน ความรับผิดชอบ เป็นต้น ปัจจัยหลักจะต้องเป็นปัจจัยสำคัญที่สุดที่ใช้ในการให้ผลการเรียน/กำหนดเกรด ส่วนปัจจัยเสริมเป็นเพียงส่วนประกอบที่สามารถนำมาพิจารณาในกรณีที่ปัจจัยหลักขาดความสมบูรณ์ (2) เครื่องมือที่ใช้วัดผลจะต้องมีคุณภาพ วัดได้ครอบคลุมลักษณะตามน้ำหนักความสำคัญที่ต้องประเมิน ควรวัดด้วยเครื่องมือหลายอย่าง หลายครั้ง หลายเวลา และมีมาตรการในการตรวจให้คะแนนอย่างเป็นปรนัย (3) เกณฑ์ที่ใช้ในการกำหนดเกรดต้องมีความเหมาะสมกับธรรมชาติของวิชา และระบุให้ชัดเจน เกณฑ์ต้องสอดคล้องกับแผนการวัดผลและแผนการจัดกิจกรรมการเรียนรู้ และ (4) การตัดสินผลจะต้องเป็นไปอย่างยุติธรรม โปร่งใส และสามารถอธิบายให้ผู้เรียนผู้ที่เกี่ยวข้องยอมรับได้ด้วยหลักการและเหตุผล การใช้วิจารณ์ญาณจะต้องเป็นไปอย่างยุติธรรมและรับผิดชอบต่อผลที่เกิดขึ้น

## ข้อเสนอแนะ

### 1. ข้อเสนอแนะสำหรับการนำผลการวิจัยไปใช้

#### (1) การใช้ประโยชน์จากผลการวิจัย

(1.1) ในการเปรียบเทียบผลการเรียนเฉลี่ยสะสม โดยใช้คะแนน O-NET เป็นมาตรฐานในการเปรียบเทียบผลการเรียนเฉลี่ยสะสมทุกวิชาและตามกลุ่มสาระการเรียนรู้ของนักเรียนมัธยมศึกษาตอนปลาย ควรใช้ผลการเรียนเฉลี่ยสะสมรวมทุกวิชา (GPAXT) ในการเปรียบเทียบ และควรนำวิธีการเปรียบเทียบเชิงเส้นตรงไปใช้ เพราะเป็นวิธีการที่ให้ผลสอดคล้องกับข้อมูลเชิงประจักษ์ และไม่ซับซ้อน ดำเนินงานได้สะดวก (Practical)

(1.2) กระทรวงศึกษาธิการ สพฐ. และ สทศ. ควรประสานความร่วมมือสำหรับการดำเนินการจัดเตรียมฐานข้อมูลของผลการเรียนเฉลี่ยสะสม (GPAX) และ O-NET ของนักเรียนและโรงเรียนให้มีความพร้อม เป็นปัจจุบัน เพื่อประโยชน์สำหรับการเปรียบเทียบผลการเรียนเฉลี่ยสะสมในแต่ละปี

#### (2) การใช้ประโยชน์จากผลการเปรียบเทียบ

(2.1) สพฐ. ควรดำเนินการแจ้งให้โรงเรียนทราบว่าโรงเรียนแต่ละแห่งมีลักษณะการให้เกรดแบบใด เพื่อหามาตรการส่งเสริม สนับสนุน และพัฒนามาตรฐานในการประเมินผลการเรียนรู้ของผู้เรียนให้เหมาะสมสอดคล้องกับผลการเรียนรู้ที่แท้จริงของผู้เรียน และสอดคล้องกับมาตรฐานของหลักสูตรการศึกษาขั้นพื้นฐานที่เป็นไปในแนวทางเดียวกัน

(2.2) สพฐ. ควรพัฒนามาตรฐานวิธีการวัดและประเมินผลการเรียนรู้ของผู้เรียน และกำหนดหลักการและแนวทางในการวัดและประเมินผลการเรียนรู้ของผู้เรียน จัดทำคู่มือให้โรงเรียนเพื่อการวัดและประเมินผลที่เป็นแนวทางเดียวกัน



(2.3) สพฐ. ควรเพิ่มบทบาทของสำนักงานเขตพื้นที่การศึกษาในมิติด้านการกำกับติดตามและพัฒนาระบบการวัดและประเมินผลการเรียนรู้ของโรงเรียนควบคู่ไปกับการประกันคุณภาพการศึกษาในมาตรฐานด้านผู้เรียน ให้มีมาตรฐานที่สามารถเทียบเคียงกันได้ภายในเขตพื้นที่การศึกษา

(2.4) โรงเรียนที่มีลักษณะการให้เกรดแบบกดเกรด ส่วนใหญ่จะเป็นโรงเรียนคุณภาพดี นักเรียนมีความสามารถสูง ควรต้องรักษามาตรฐานการจัดการเรียนรู้ การวัดและประเมินผลที่สูงไว้ โดยหน่วยงานต้นสังกัดควรมีการนำวิธีการปรับเทียบผลการเรียนของนักเรียนเพื่อความยุติธรรมในการนำคะแนนผลการเรียนเฉลี่ยสะสมไปใช้ในการสอบคัดเลือกเข้าศึกษาต่อระดับมหาวิทยาลัย

(2.5) โรงเรียนที่มีลักษณะการให้เกรดแบบปล่อยเกรด ควรต้องสร้างมาตรการเพื่อทบทวนและส่งเสริมกระบวนการจัดการเรียนรู้ การวัดและประเมินผล ให้มีมาตรฐานที่สูงขึ้นทั้งในด้านของคุณภาพเครื่องมือ และความสอดคล้องของการวัดและประเมินผลกับมาตรฐานของหลักสูตรการศึกษาขั้นพื้นฐาน

## 2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

(1) สพฐ. ควรศึกษาวิจัยเพื่อติดตามผลการปรับเทียบผลการเรียนเฉลี่ยสะสมโดยรวมและจำแนกตามกลุ่มสาระการเรียนรู้ โดยใช้คะแนน O-NET เพื่อติดตามพัฒนาการของมาตรฐานการประเมินผลการเรียนรู้และมาตรฐานการเรียนรู้ของโรงเรียนเป็นประจำทุกปี

(2) สพฐ. ควรศึกษาเพื่อประเมินคุณภาพการจัดการศึกษาของโรงเรียนในสังกัด โดยใช้ตัวบ่งชี้พัฒนาการของมาตรฐานการประเมินผลการเรียนรู้และมาตรฐานการเรียนรู้ เป็นตัวบ่งชี้ส่วนหนึ่งในการประกันคุณภาพภายในของโรงเรียน

## เอกสารอ้างอิง

### ภาษาไทย

- ภัทรพร เกษสังข์. (2546). การศึกษาพัฒนาการความสามารถทางคณิตศาสตร์ของนักเรียน  
ชั้นมัธยมศึกษาตอนต้นโดยการเปรียบเทียบคะแนนในแนวตั้งที่ใช้วิธีการที่เหมาะสม.  
วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ  
ประสานมิตร.
- รัฐชาติ มงคลนาวิน. (2550, 14 กรกฎาคม). คณะวิทยโอเดอนิสิตออกกลางคันอ้อ. ไทยรัฐ.  
แหล่งที่มา: <http://www.thairath.co.th/news.php?section=education&content=53905>.  
[8 กุมภาพันธ์ 2551].
- รับรองมาตรฐานและประเมินคุณภาพการศึกษา, สำนักงาน. (มปพ.). สรุปผลการสังเคราะห์ผล  
การประเมินคุณภาพภายนอกสถานศึกษา ระดับการศึกษาขั้นพื้นฐาน (รอบแรก พ.ศ. 2544-  
2548) จำนวน 30,010 แห่ง. (เอกสารอัดสำเนา). แหล่งที่มา: [http://www.onesqa.or.th/  
upload/195/FileUpload/1398\\_2097.pdf](http://www.onesqa.or.th/upload/195/FileUpload/1398_2097.pdf). [8 กุมภาพันธ์ 2551].
- เลขาธิการสภาการศึกษา, สำนักงาน. (2549). สมรรถนะการศึกษาไทยในเวทีสากล พ.ศ. 2549.  
กรุงเทพมหานคร: สำนักงานเลขาธิการสภาการศึกษา.
- ศิริชัย กาญจนวาสี. (2548). การวิเคราะห์หุระดับ. พิมพ์ครั้งที่ 3. กรุงเทพฯ: โรงพิมพ์แห่ง  
จุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2550). ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theories). พิมพ์ครั้งที่  
4. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สุภาภรณ์ คงทวี. (2541). การปรับแก้เฉลี่ยสะสมระดับมัธยมศึกษาตอนปลายด้วยคะแนนสอบ  
เข้ามหาวิทยาลัย: การเปรียบเทียบระหว่างวิธีการปรับแบบอิกวิเปอร์เซ็นไทล์  
เชิงเส้นตรง และไออาร์ที. วิทยานิพนธ์ปริญญาโทบัณฑิต ภาคศึกษาวิจัยการศึกษา บัณฑิต  
วิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.
- อุทุมพร จามรมาน. (2551). เอกสารเกี่ยวกับ O-NET, A-NET และ Admissions. (เอกสาร  
อัดสำเนา). แหล่งที่มา: [http://www.niets.or.th/pdf/O-NET,A-NET,Admission.  
pdf](http://www.niets.or.th/pdf/O-NET,A-NET,Admission.pdf). [8 กุมภาพันธ์ 2551].

### ภาษาอังกฤษ

- Boston College. (nd.). *IEA's TIMSS 2003 International Report on Achievement in the Mathematics Cognitive Domains: Findings from a Developmental Project*. [online] Available from: [http://www.ldrc.ca/projects/stencils/sten\\_interp.php](http://www.ldrc.ca/projects/stencils/sten_interp.php) [2008, Feb 08].
- Cook, L., & Eignor, D. R. (1981). Using Item Response Theory in Test Score Equating. *International Journal of Education Research*, 13: 161–173.
- Cook, L., & Eignor, D. R. (1991). An NCME Instructional Module on IRT Equating Methods. *Educational Measurement: Issues and Practice*, 10: 37–45.
- Glowacki, M.L. (1991). The analysis of test equating models for the Alabama high school graduation examination. Doctoral Dissertation, university of Alabama. *Dissertation Abstracts International*, 52: 1722.

