

## การปรับเทียบคะแนน O-NET ระหว่างปี โดยวิธีการเทียบ เป็นมาตราเดียวกันกับโดยใช้ทฤษฎีการตอบสนองข้อสอบ

ตำราญ มีแจ้ง ประภัสสร วงษ์ดี และ ยุพิน โภณฑา  
คณะศึกษาศาสตร์ มหาวิทยาลัยนครสวรรค์

### บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบวิธีการปรับเทียบคะแนน O-NET ระหว่างปีการศึกษา 2550 และปีการศึกษา 2551 โดยใช้วิธีการเทียบเป็นมาตราเดียวกัน กับวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ

ผลการวิจัยปรากฏว่า

- 1) ค่าเฉลี่ยคะแนน O-NET ปีการศึกษา 2550 และปีการศึกษา 2551 เดิม กับค่าเฉลี่ยคะแนน O-NET ปีการศึกษา 2550 และปีการศึกษา 2551 ที่ได้จากการปรับเทียบโดยใช้วิธีการเทียบเป็นมาตราเดียวกัน ไม่แตกต่างกันที่ระดับนัยสำคัญทางสถิติ .05
- 2) การปรับเทียบคะแนน O-NET ปีการศึกษา 2550 และ ปีการศึกษา 2551 โดยใช้ทฤษฎีการตอบสนองข้อสอบที่มี 3 พารามิเตอร์ มีค่าคะแนนจริงที่ได้จากการปรับเทียบคะแนนแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
- 3) คะแนน O-NET ที่ได้จากการปรับเทียบโดยวิธีการเทียบเป็นมาตราเดียวกัน มีความสัมพันธ์กับคะแนน O-NET เดิมสูงในทางบวก อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกรายวิชา ทั้ง 2 ปี
- 4) คะแนน O-NET ที่ได้จากการปรับเทียบโดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ มีความสัมพันธ์กับคะแนน O-NET เดิมสูงในทางบวก อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกรายวิชา ทั้ง 2 ปี
- 5) ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนน O-NET ทั้ง 2 ปี ที่ได้จากการปรับเทียบโดยวิธีการเทียบเป็นมาตราเดียวกัน กับคะแนน O-NET เดิม ไม่แตกต่าง จากค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนน O-NET ที่ได้จากการปรับเทียบโดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ กับคะแนน O-NET เดิม

**คำสำคัญ:** การปรับเทียบคะแนนระหว่างแบบทดสอบ/ การเทียบเป็นมาตราเดียวกัน/ ทฤษฎีการตอบสนองข้อสอบ

## A Comparison of Results from Equating O-NET Scores with Common Items and with a Three-Parameter IRT Model

Samran Mejang  
Prapassorn Wongdee and Upin Konta  
*Faculty of Education, Naresuan University, Thailand*

### Abstract

*The purpose of this research was to compare the O-NET scores from 2007 and 2008 equated by using two methods: common-item equating, and a 3PL IRT model. The findings can be summarized as follows:*

- 1. When using common items to equate O-NET scores, no significant differences were found between test means from 2007 and 2008 ( $p > .05$ ).*
- 2. When a 3PL IRT model was used for equating, test means were found to differ ( $p < .01$ ).*
- 3. The O-NET scores equated with the common-item method had a high positive relationship ( $p < .01$ ) with the original O-NET scores in every subject over both years.*
- 4. The O-NET scores equated with 3PL IRT also had a high positive relationship ( $p < .01$ ) with the original O-NET scores in every subject over both years.*
- 5. When comparing the correlation coefficients found in (3) with those found in (4), no differences were noted ( $p > .05$ ).*

**Keywords:** *test equating, common-item equating, item response theory*

## ความนำ

การวัดและประเมินผลการศึกษาส่วนใหญ่ นิยมใช้แบบทดสอบเป็นเครื่องมือที่ทำให้ได้ข้อมูลหรือสารสนเทศเกี่ยวกับการจัดกิจกรรมทางการศึกษา ผลที่ได้จากการทดสอบมีความสำคัญต่อการพัฒนาคุณภาพของการศึกษา ตลอดจนใช้เป็นข้อมูลพื้นฐานในการตัดสินใจของครู ผู้บริหาร และผู้ที่เกี่ยวข้องกับการจัดการศึกษาทุกระดับ นอกจากนี้ยังมีการนำแบบทดสอบไปใช้อย่างแพร่หลายในปัจจุบันตามจุดมุ่งหมายที่ต่างกัน เช่น การสอบแข่งขัน การสอบคัดเลือกบุคคลเข้าเรียน เข้าทำงาน จัดตำแหน่ง เลื่อนระดับ ซึ่งในการสอบแต่ละครั้งนั้นอาจมีการใช้แบบทดสอบในลักษณะที่แตกต่างกันไป เพื่อให้สามารถวัดได้ตรงกับความสามารถที่แท้จริงของแต่ละบุคคลและเพื่อให้เกิดความยุติธรรมแก่ทุกคน การทดสอบความรู้ในเบื้องต้นถือเป็นหน้าที่ของสถานศึกษาชั้นพื้นฐานทุกแห่ง ต้องทำการทดสอบผู้เรียน แต่ไม่สามารถรับประกันในคุณภาพของผลการวัดที่เป็นมาตรฐานเดียวกันได้ หน่วยงานที่รับผิดชอบจัดการศึกษาหลักหลายหน่วยงาน จึงได้จัดทดสอบผู้เรียนด้วยแบบทดสอบที่ได้มาตรฐานฉบับเดียวกัน สถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.) เป็นหน่วยงานระดับชาติหน่วยงานหนึ่งที่ทำหน้าที่บริหารจัดการด้านการสอบผู้เรียนซึ่งผลการทดสอบที่บริหารจัดการโดย สทศ. นั้นสามารถใช้เป็นสารสนเทศหนึ่งสำหรับพิจารณาคุณภาพในการจัดการศึกษาของสถานศึกษาด้วย เช่น สทศ. จัดการทดสอบทางการศึกษาแห่งชาติชั้นพื้นฐาน (Ordinary National Educational Test หรือ O-NET) และจัดสอบวัดความรู้อื่น ๆ ด้วย คะแนนจากการทดสอบ O-NET นี้ นอกจากจะสามารถใช้เป็นเกณฑ์ร่วมกันในการวัดคุณภาพของทุกหน่วยงาน ว่าสามารถจัดการศึกษาชั้นพื้นฐานอยู่ในระดับที่ได้มาตรฐานใกล้เคียงกัน หรือแตกต่างกันมากน้อยเพียงใดแล้ว สำหรับคะแนน O-NET ของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ยังมีความสำคัญอีกขั้นหนึ่งต่อตัวนักเรียนเอง นั่นคือ คะแนน O-NET นี้ได้ถูกใช้เป็นส่วนหนึ่งในการยื่นเข้าไปพิจารณาเพื่อคัดเลือกเข้าศึกษาต่อในระดับอุดมศึกษาต่อไป แต่เนื่องจากจำนวนในการรับนักเรียนเข้าศึกษาต่อในระดับอุดมศึกษามีจำกัด จึงทำให้มีนักเรียนบางส่วนที่อาจพลาดในการผ่านเกณฑ์การเข้าไปเรียนในระดับอุดมศึกษาได้ทันทีที่จบการศึกษาระดับชั้นมัธยมศึกษาตอนปลาย โดยจะมีผลต่อคะแนนสอบเข้า ร้อยละ 35-70 ทั้งนี้ขึ้นอยู่กับข้อกำหนดของแต่ละคณะ ดังนั้น นักเรียนจึงต้องเก็บคะแนนที่ได้จากการสอบ O-NET ไว้ เพื่อนำไปเป็นส่วนหนึ่งเพื่อร่วมพิจารณาให้ผ่านเกณฑ์การเข้าไปศึกษาต่อในระดับอุดมศึกษาต่อไป นักเรียนแต่ละคนสามารถทดสอบ O-NET เพื่อเก็บคะแนนไว้พิจารณาสอบเข้าเรียนระดับอุดมศึกษาได้เพียงครั้งเดียว คือช่วงที่กำลังศึกษาในระดับชั้นมัธยมศึกษาปีที่ 6 จึงอาจเกิดเหตุการณ์ที่มีการนำคะแนน O-NET ที่ทดสอบต่างปีกันมาร่วมพิจารณาสอบเข้าเรียนในคณะเดียวกันได้ หากนักเรียนไม่สามารถผ่านการคัดเลือกเข้าศึกษาในระดับอุดมศึกษาได้ทันทีที่จบการศึกษาระดับชั้นมัธยมศึกษาปีที่ 6 ดังนั้นแบบทดสอบที่ใช้ทดสอบแต่ละปีต้องมีมาตรฐานไม่แตกต่างกัน เพื่อไม่ให้เกิดความลำเอียงกับนักเรียนที่สอบต่างปี

การทดสอบที่ผู้สอบทำแบบทดสอบต่างฉบับกัน และสามารถเปรียบเทียบความสามารถของผู้สอบได้นั้น ผู้สอบต้องสอบจากแบบทดสอบคู่ขนาน จึงสามารถนำคะแนนดิบที่ได้จากการทดสอบฉบับต่าง ๆ มาเปรียบเทียบกันได้โดยตรง (สุรัตน์ ขวัญบุญจันทร์, 2531) แต่การสร้างแบบทดสอบให้เป็นคู่ขนานกันอย่างแท้จริงนั้นทำได้ยาก

เพราะจะต้องสร้างให้คู่ขนานกันทั้งในด้านโครงสร้างของข้อสอบ เวลาที่ใช้ในการสอบ ชนิดของข้อสอบ รูปแบบข้อสอบ และเนื้อหาในข้อสอบ (ชูศักดิ์ ชัมภลิจิต, 2529) นอกจากนี้ในการพยายามสร้างข้อสอบที่เป็นคู่ขนานจริงอาจต้องสูญเสียเวลาและไม่คุ้มค่ากับการลงทุน จึงมีผู้คิดค้นรูปแบบที่อาศัยวิธีการทางสถิติมาทำการปรับคะแนนให้เกิดความเท่าเทียมกันได้ โดยเรียกว่าวิธี “การปรับเทียบคะแนน” (Test equating) ซึ่งวิธีการนี้จะสามารถแปลงคะแนนของแบบทดสอบชุดหนึ่งไปสู่มาตรฐานเดียวกันกับคะแนนของแบบทดสอบอีกชุดหนึ่ง คะแนนของแบบทดสอบต่างชุดกันจะอยู่บนสเกลการวัดเดียวกัน จึงสามารถหาคะแนนสมมูลระหว่างแบบทดสอบต่างชุดนั้นและนำคะแนนมาเปรียบเทียบกันได้ (Angoff, 1971; Rentz & Bashaw, 1977) ทั้งนี้ไม่ว่าแบบทดสอบต่างชุดนั้นจะคู่ขนานกันหรือไม่ก็ตาม ถ้ามุ่งวัดในสิ่งเดียวกันก็จะสามารถแปลงคะแนนจากแบบทดสอบต่างชุดให้มีคุณสมบัติที่เทียบเท่ากันได้เสมอ (Lord, 1977; Kolen, 1981)

สำหรับวิธีการปรับเทียบคะแนนมีหลายรูปแบบ ถ้าแบ่งตามทฤษฎีที่สนับสนุนจะแบ่งได้สองกลุ่มใหญ่ ๆ (Flanagan, 1951) คือ 1) กลุ่มทฤษฎีการวัดแบบดั้งเดิม (Classical test theory) ประกอบด้วยรูปแบบเชิงเส้นตรง (Linear equating) รูปแบบอิกวิเปอร์เซ็นต์ไทล์ (Equipercentile equating) และรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ (Equating using the confirmatory factor analysis model) 2) กลุ่มทฤษฎีการวัดความสามารถแฝง (Latent trait theory) ประกอบด้วยรูปแบบอิงทฤษฎีการตอบสนองข้อสอบ (Item response theory equating model) ซึ่งได้จำแนกการปรับเทียบตามโมเดล 1 พารามิเตอร์ การปรับเทียบตามโมเดล 2 พารามิเตอร์ และการปรับเทียบตามโมเดล 3 พารามิเตอร์

การเลือกใช้รูปแบบการปรับเทียบคะแนนต้องพิจารณาเลือกให้เหมาะสมกับเงื่อนไขของการปรับเทียบในแต่ละรูปแบบและให้เหมาะสมกับลักษณะของข้อมูล ดังที่ อรุณี อ่อนสวัสดิ์ และ ซอบ สีสอ (2541) ได้เลือกใช้วิธีแปลงเป็นคะแนนมาตรฐานกลาง (เป็นรูปแบบภายใต้วิธีการปรับเทียบเชิงเส้นตรง) และวิธีแปลงโดยอิงตำแหน่งเปอร์เซ็นต์ไทล์ เพื่อเปรียบเทียบผลการเรียนสะสมของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2540 สำหรับคัดเลือกเข้าศึกษาในสถาบันอุดมศึกษาระบบใหม่ จากโรงเรียนสังกัดกรมสามัญศึกษา เขตการศึกษา 7 ที่มีขนาดต่างกัน และสำราญ มีแจ้ง (2542) ใช้วิธีการปรับเทียบเป็นมาตรฐานเดียวกัน (เป็นรูปแบบภายใต้วิธีการปรับเทียบเชิงเส้นตรง) เป็นหนึ่งวิธีการของการเปรียบเทียบคะแนนสอบคัดเลือกเข้าศึกษาในสถาบันอุดมศึกษาของนักเรียนต่างกลุ่มกัน คือ เป็นนักเรียนที่เรียนในโรงเรียนต่างขนาดกัน และนักเรียนที่เรียนต่าง โปรแกรมกัน

จากวิธีการที่กล่าวมานี้ มีกฎในการแปลงคะแนนต่างกันและขึ้นอยู่กับวิธีแต่ละวิธีการ จึงเกิดปัญหาว่าวิธีที่ให้ค่าการแปลงคะแนนได้ถูกต้องที่สุดหรือมีโอกาสที่จะคลาดเคลื่อนได้น้อยที่สุด เพื่อให้เกิดความยุติธรรมแก่ผู้สอบมากที่สุดคือวิธีการใด สำหรับการศึกษาครั้งนี้ผู้วิจัยพิจารณาลักษณะของข้อมูลคือคะแนน O-NET ที่จะนำมาปรับเทียบ พบว่าเป็นข้อมูลที่มาจากกลุ่มตัวอย่างต่างกลุ่มกัน นั่นคือเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2550 และนักเรียนชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2551 อย่างไรก็ตามยังถือว่าเป็นกลุ่มประชากรนักเรียนชั้นมัธยมศึกษาปีที่ 6 เช่นเดียวกัน ซึ่งเป็นนักเรียนภายใต้การจัดการศึกษาในระบบที่ไม่แตกต่างกัน อนุমানให้เป็นกลุ่มผู้สอบที่เท่าเทียมกัน ผู้วิจัยจึงได้พิจารณาเลือกวิธีการปรับเทียบคะแนนให้สอดคล้องกับลักษณะของข้อมูล และเลือก

วิธีการที่มีความยุ่งยากน้อยในด้านการปฏิบัติ คือเลือกวิธีการปรับเทียบเป็นมาตราเดียวกัน เพื่อปรับเทียบคะแนนที่มาจากกลุ่มตัวอย่างต่างกลุ่มกันแต่มาจากประชากรกลุ่มเดียวกัน และนอกจากนี้ผลการศึกษาของ สุภาภรณ์ คงทวี (2541) ที่ได้เปรียบเทียบประสิทธิภาพของวิธีปรับคะแนนเฉลี่ยสะสม 3 วิธี ได้แก่ วิธีอิกวิเปอร์เซ็นไทล์ วิธีเชิงเส้นตรงรูปแบบที่ 3 ของ Angoff (1971) และวิธีปรับโดยใช้ทฤษฎีการตอบสนองข้อสอบภายใต้ Graded Responses Model ของ Young (1995) โดยใช้แบบทดสอบเข้ามหาวิทยาลัยเป็นแบบทดสอบร่วม กลุ่มตัวอย่างเป็นนิสิตนักศึกษาชั้นปีที่ 1 ของมหาวิทยาลัยต่าง ๆ 6 แห่ง ในปีการศึกษา 2540 พบว่าเมื่อใช้คะแนนเฉลี่ยสะสมระดับมหาวิทยาลัยปี 1 เป็นเกณฑ์วิธีปรับเทียบคะแนนเชิงเส้นตรง มีประสิทธิภาพในการพยากรณ์สูงกว่าวิธีอื่น ส่วนวิธีที่สองที่ผู้วิจัยพิจารณาเลือกใช้คือวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ เพราะสามารถใช้ปรับเทียบคะแนนที่ได้จากคน 2 กลุ่มที่ต่างกัน และปรับเทียบคะแนนจากแบบทดสอบ 2 ชุดที่ไม่เป็นคู่ขนานกันได้ และนอกจากนี้ยังพบว่า เรวดี อินทสระ (2530) ได้เปรียบเทียบความคลาดเคลื่อนและความเพียงพอในการปรับเทียบคะแนน ได้ข้อสรุปว่าวิธีอิงทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ ให้ความเพียงพอในระดับที่น่าพอใจ และยังเป็นรูปแบบของการปรับเทียบที่ไม่ได้สนใจความเท่าเทียมกันของความยากของแบบทดสอบที่เกี่ยวข้องกับการปรับเทียบคะแนน และถือว่าเป็นวิธีการปรับเทียบที่อาศัยสารสนเทศที่หลากหลายกว่าวิธีอื่น มาประกอบการพิจารณาความสามารถของผู้สอบอีกด้วย

### วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบวิธีการปรับเทียบคะแนน O-NET ระหว่างปีการศึกษา 2550 และ ปีการศึกษา 2551 โดยใช้วิธีการเทียบเป็นมาตราเดียวกัน กับ วิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ

### วิธีดำเนินการวิจัย

การปรับเทียบคะแนนผู้วิจัยได้เสนอวิธีการปรับเทียบไว้ 2 วิธี คือวิธีการเทียบเป็นมาตราเดียวกัน และวิธีการปรับเทียบที่ใช้ทฤษฎีการตอบสนองข้อสอบ ซึ่งแต่ละวิธี มีขั้นตอนการดำเนินการ และวิธีการวิเคราะห์ข้อมูล ดังนี้

#### 1. วิธีการปรับเทียบเป็นมาตราเดียวกัน

การปรับเทียบคะแนนให้เป็นมาตราเดียวกันของข้อสอบ O-NET ได้พิจารณารูปแบบการเก็บรวบรวมข้อมูล พบว่า การจัดสอบข้อสอบ O-NET เป็นรูปแบบที่ผู้สอบมาจากการสุ่มจากประชากรกลุ่มเดียวกัน แล้วแบ่งเป็น 2 กลุ่มคือ กลุ่มนักเรียนชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2550 สอบแบบทดสอบ O-NET ปีการศึกษา 2550 และกลุ่มนักเรียนชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2551 สอบแบบทดสอบ O-NET ปีการศึกษา 2551 โดยนักเรียนเป็นคนละกลุ่มกัน ทำการทดสอบแบบทดสอบ O-NET คนละฉบับกัน ในการปรับเทียบคะแนนให้เป็นมาตราเดียวกัน ใช้วิธีการปรับเทียบคะแนนเชิงเส้นตรงโดยการแปลงให้เป็นคะแนนมาตรฐาน Z ดังนี้

1.1 คำนวณค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของคะแนนแบบทดสอบ O-NET ปีการศึกษา 2550 ที่สอบ โดยกลุ่มนักเรียนชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2550 และคำนวณค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของ

คะแนนแบบทดสอบ O-NET ปีการศึกษา 2551 ที่สอบ โดยกลุ่มนักเรียนชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2551 โดยดำเนินการที่ละรายวิชา แล้วคำนวณคะแนนมาตรฐานของทั้งสองกลุ่ม โดยมีวิธีการคำนวณดังนี้

$$\text{คะแนนมาตรฐาน}_{O-NET 50} = \frac{(X_{O-NET 50} - \bar{X}_{O-NET 50})}{SD_{O-NET 50}} \text{-----(1)}$$

$$\text{คะแนนมาตรฐาน}_{O-NET 51} = \frac{(X_{O-NET 51} - \bar{X}_{O-NET 51})}{SD_{O-NET 51}} \text{-----(2)}$$

1.2 ในการจัดสอบข้อสอบ O-NET เนื่องจาก สทศ. เป็นผู้พัฒนาข้อสอบและดำเนินการสอบ O-NET ทั้ง 2 ปี จึงเชื่อว่าข้อสอบทั้ง 2 ปี ต้องเป็นข้อสอบคู่ขนานที่มีความคล้ายกันทั้งในเชิงโครงสร้าง เนื้อหา ความยาก และอำนาจจำแนก ดังนั้นคะแนนมาตรฐานทั้งสองปีควรมีค่าเท่ากัน นั่นคือ สมการที่ (1) = (2) ดังนี้

$$\frac{(X_{O-NET 50} - \bar{X}_{O-NET 50})}{SD_{O-NET 50}} = \frac{(X_{O-NET 51} - \bar{X}_{O-NET 51})}{SD_{O-NET 51}} \text{-----(3)}$$

จากสมการ (3) สามารถประยุกต์เป็นสมการคำนวณคะแนนแปลงได้ ดังนี้

$$X_{T_{O-NET 50}} \text{ หรือ } X_{O-NET 51} = X_{O-NET 50} \left( \frac{SD_{O-NET 51}}{SD_{O-NET 50}} \right) - \bar{X}_{O-NET 50} \left( \frac{SD_{O-NET 51}}{SD_{O-NET 50}} \right) + \bar{X}_{O-NET 51} \text{-----(4)}$$

โดย  $X_{T_{O-NET 50}}$  หรือ  $X_{O-NET 51}$  เป็นคะแนนแปลงจากคะแนนเดิมของ O-NET ปีการศึกษา 2550 เพื่อให้สามารถปรับเทียบกับคะแนน O-NET ปีการศึกษา 2551 ได้ จึงถือว่าคะแนนแปลงนี้เสมือนเป็นคะแนน O-NET ปีการศึกษา 2551

$$X_{T_{O-NET 51}} \text{ หรือ } X_{O-NET 50} = X_{O-NET 51} \left( \frac{SD_{O-NET 50}}{SD_{O-NET 51}} \right) - \bar{X}_{O-NET 51} \left( \frac{SD_{O-NET 50}}{SD_{O-NET 51}} \right) + \bar{X}_{O-NET 50} \text{-----(5)}$$

โดย  $X_{T_{O-NET 51}}$  หรือ  $X_{O-NET 50}$  เป็นคะแนนแปลงจากคะแนนเดิมของ O-NET ปีการศึกษา 2551 เพื่อให้สามารถปรับเทียบกับคะแนน O-NET ปีการศึกษา 2550 ได้ จึงถือว่าคะแนนแปลงนี้เสมือนเป็นคะแนน O-NET ปีการศึกษา 2550

1.3 กรณีที่ทราบคะแนนนักเรียนในปีใดปีหนึ่ง เช่น ทราบคะแนน O-NET ปีการศึกษา 2550 แล้ว สมมุติว่านักเรียนที่สอบ O-NET ในปีการศึกษา 2550 ได้ยื่นคะแนนเพื่อให้พิจารณาเปรียบเทียบกับนักเรียนที่มีคะแนนสอบ O-NET ของปีการศึกษา 2551 สามารถนำคะแนนของนักเรียนที่สอบ O-NET ปีการศึกษา 2550 มาเทียบกับคะแนน O-NET ปีการศึกษา 2551 ได้จากการคำนวณคะแนนปีการศึกษา 2550 ให้เป็นคะแนนแปลง หรือ ให้เป็นคะแนนปีการศึกษา 2551 ได้จากสมการหรือสูตร (4) แล้วค่อยนำคะแนนปีการศึกษา 2550 ที่แปลงเป็นคะแนนปีการศึกษา 2551 ไปเปรียบเทียบกับคะแนน O-NET ปีการศึกษา 2551 ต่อไป

2. วิธีการเปรียบเทียบที่ใช้ทฤษฎีการตอบสนองข้อสอบ (Equating Through Item Response Theory) โดยใช้โมเดล 3 พารามิเตอร์ มีขั้นตอนการปรับเทียบดังนี้

2.1 นำคะแนนสอบรายชื่อของนักเรียนที่ได้จากการทดสอบด้วยแบบสอบทั้ง 2 ฉบับมาวิเคราะห์หาค่าความยาก ค่าอำนาจจำแนก และค่าสัมประสิทธิ์การเดาของข้อสอบเป็นรายชื่อ โดยใช้โปรแกรมการวิเคราะห์ IRT ทำการวิเคราะห์ แล้วได้ค่าพารามิเตอร์ของข้อสอบ 3 พารามิเตอร์ ประกอบด้วย ค่าความยาก ค่าอำนาจจำแนก ค่าสัมประสิทธิ์การเดาของข้อสอบทุกข้อ และได้ค่าความสามารถ (ability) ของนักเรียนแต่ละคน

2.2 ปรับเทียบคะแนนจริงที่วิเคราะห์ได้ในแต่ละปีเทียบกับระดับความสามารถ โดยสร้างตารางเทียบคะแนนตามแนวนอน ซึ่งนำเสนอค่าคะแนนจริงจำแนกตามรายวิชา เปรียบเทียบกันทั้ง 2 ปี ณ ค่าความสามารถเดียวกัน เพราะพิจารณาให้บุคคลที่มีความสามารถเท่าเทียมกัน ในสิ่งที่มุ่งศึกษา เป็นเสมือนคนเดียวกันที่ทำการทดสอบ 2 ครั้ง เมื่อจะเปรียบเทียบว่าแบบทดสอบ หรือคะแนนสอบ 2 ครั้งนั้น เป็นคู่ขนานกันหรือไม่ หรือสามารถจะนำมาปรับเทียบกันได้โดยตรงหรือไม่ จึงสามารถใช้สถิติทดสอบความแตกต่างของค่าเฉลี่ยข้อมูลที่ไม่มีอิสระต่อกัน คือ ใช้สถิติทดสอบ t-test dependent หรือ ใช้ paired t-test

กรณีที่คะแนนเต็มเท่ากันทั้ง 2 ปี สามารถใช้คะแนนจริง ณ ระดับความสามารถเดียวกันเปรียบเทียบกันได้เลย แต่ถ้าคะแนนเต็มไม่เท่ากันให้ปรับคะแนนจริงของทั้งปีการศึกษา 2550 และปีการศึกษา 2551 เป็นค่าร้อยละของคะแนนจริงแทนคะแนนจริง

3. วิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน O-NET เดิม กับคะแนนที่ปรับเทียบแล้วทั้ง 2 วิธี และวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ปรับเทียบแล้วทั้ง 2 วิธี โดยใช้การหาค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน

4. ทดสอบความแตกต่างของค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน O-NET เดิม กับคะแนน O-NET ที่แปลงด้วยวิธีการเทียบให้เป็นมาตราเดียวกัน และค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนคะแนน O-NET เดิมกับคะแนน O-NET ที่แปลงด้วยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ โดยใช้  $t^2$ -Hotelling

## ผลการวิจัย

การปรับเทียบคะแนน O-NET ในปีการศึกษา 2550 และ 2551 โดยวิธีการปรับเทียบให้เป็นมาตราเดียวกัน และโดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ ได้ผลที่สอดคล้องกันคือ คะแนน O-NET ทั้ง 2 ปี ไม่สามารถนำมาเทียบกันได้โดยตรง ต้องมีการแปลงคะแนนให้มีหน่วยเดียวกันก่อนนำมาเทียบกัน ซึ่งรายละเอียดผลการปรับเทียบคะแนนครั้งนี้ มีดังนี้

1. ผลการปรับเทียบคะแนนโดยวิธีการปรับเทียบให้เป็นมาตราเดียวกัน พบว่า ค่าเฉลี่ยคะแนน O-NET ปีการศึกษา 2550 เดิมกับค่าเฉลี่ยคะแนน O-NET ปีการศึกษา 2550 ที่ได้จากการปรับเทียบโดยใช้วิธีการเทียบเป็นมาตราเดียวกัน ไม่แตกต่างกันที่ระดับนัยสำคัญทางสถิติ .05 ทั้ง 4 รายวิชา และในทำนองเดียวกัน ค่าเฉลี่ยคะแนน

O-NET ปีการศึกษา 2551 เดิมกับค่าเฉลี่ยคะแนน O-NET ปีการศึกษา 2551 ที่ได้จากการปรับเทียบโดยใช้วิธีการเทียบเป็นมาตรฐานเดียวกัน ไม่แตกต่างกันที่ระดับนัยสำคัญทางสถิติ .05 ทั้ง 4 รายวิชา เช่นเดียวกัน

2. ผลการปรับเทียบคะแนนโดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ พบว่า ผลการเทียบคะแนน O-NET เดิมกับคะแนนแปลง O-NET ที่ปรับเทียบ โดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ ทั้ง 4 รายวิชา ที่ทำการทดสอบในปีการศึกษา 2550 และ ปีการศึกษา 2551 โดยใช้ทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์ มีค่าคะแนนจริงที่ได้จากการเทียบคะแนนแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยผู้สอบที่มีระดับความสามารถเดียวกัน ได้ค่าเฉลี่ยของคะแนนจริง (หรือค่าเฉลี่ยร้อยละของคะแนนจริง) จากการทดสอบข้อสอบ O-NET ปีการศึกษา 2550 สูงกว่าค่าเฉลี่ยของคะแนนจริง (หรือค่าเฉลี่ยร้อยละของคะแนนจริง) จากการทดสอบข้อสอบ O-NET ปีการศึกษา 2551

3. คะแนน O-NET ที่ได้จากการปรับเทียบโดยวิธีการเทียบเป็นมาตรฐานเดียวกัน มีความสัมพันธ์กับคะแนน O-NET เดิมสูงในทางบวก อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกรายวิชา ทั้ง 2 ปี

4. คะแนน O-NET ที่ได้จากการปรับเทียบโดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ มีความสัมพันธ์กับคะแนน O-NET เดิมสูงในทางบวก อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกรายวิชา ทั้ง 2 ปี

5. ผลการเปรียบเทียบคุณภาพวิธีการปรับเทียบคะแนน O-NET ระหว่างวิธีการเทียบเป็นมาตรฐานเดียวกันกับการปรับเทียบโดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ พบว่า ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนน O-NET ทั้ง 2 ปี ทั้ง 4 รายวิชา ที่ได้จากการปรับเทียบโดยวิธีการเทียบเป็นมาตรฐานเดียวกัน กับคะแนน O-NET เดิม ไม่แตกต่างจากค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนน O-NET เดิม กับค่าความสามารถที่ได้จากการปรับเทียบโดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ

## อภิปรายผล

1. จากผลการวิจัย พบว่า คะแนน O-NET ปีการศึกษา 2550 เดิม กับคะแนนแปลง O-NET ที่ได้จากการปรับเทียบมาจากคะแนน O-NET ปีการศึกษา 2551 โดยวิธีการปรับเทียบเป็นมาตรฐานเดียวกัน มีค่าเฉลี่ยของคะแนนไม่แตกต่างกัน และผลการศึกษากับคะแนน O-NET ปีการศึกษา 2551 เดิมก็ได้ผลเช่นเดียวกัน ซึ่งสอดคล้องกับสมมติฐาน ขณะเดียวกันผลการปรับเทียบคะแนนโดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ ได้ผลที่ไม่สอดคล้องกับสมมติฐาน เพราะคะแนนจริงของผู้เรียนที่ค่าความสามารถเดียวกันมีความแตกต่างกัน

จากผลการวิจัยดังกล่าวนี้ ผู้ดำเนินการจัดสอบคะแนน O-NET ควรต้องพิจารณาถึงความเหมาะสมในการเลือกวิธีการปรับเทียบในแต่ละวิธีจากข้อค้นพบดังกล่าว ถ้าต้องการใช้วิธีการปรับเทียบแบบแรก ต้องมั่นใจว่าข้อมูลคะแนนของข้อสอบ 2 ชุดนี้มีการแจกแจงเหมือนกัน แล้วจึงเลือกใช้วิธีการปรับเทียบเป็นมาตรฐานเดียวกัน ซึ่งถือว่าเป็นวิธีการที่ไม่ซับซ้อน สามารถดำเนินการได้ง่าย แต่ถ้าไม่มั่นใจในการแจกแจงข้อมูลคะแนน 2 ชุด และเป็น

การสอบจากผู้สอบคนละกลุ่มและใช้แบบสอบคนละฉบับ วิธีการเปรียบเทียบ โดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ น่าจะเป็นวิธีการที่เหมาะสม แต่อาจมีความยุ่งยากในการดำเนินการค่อนข้างมาก

2. จากผลการวิจัยที่พบว่า ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน O-NET เดิมกับคะแนนแปลงที่ได้จากการเปรียบเทียบ ทั้งโดยวิธีการเปรียบเทียบเป็นมาตราเดียวกัน และวิธีการเปรียบเทียบ โดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ มีค่าเป็นบวก ซึ่งสอดคล้องกับสมมติฐานการวิจัย ที่ตั้งไว้ ทั้งนี้อาจเป็นเพราะว่าวิธีการเปรียบเทียบคะแนนต่างวิธีกันมีความตรงในการเปรียบเทียบคะแนน และสามารถเลือกใช้วิธีการเปรียบเทียบคะแนนทั้ง 2 วิธีนี้แทนกันได้

3. จากข้อค้นพบที่ว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน O-NET เดิมกับคะแนนแปลงที่เปรียบเทียบโดยวิธีการเปรียบเทียบเป็นมาตราเดียวกัน มีค่าสัมประสิทธิ์สหสัมพันธ์สูงมาก โดยมีค่าเป็น 1.00 นั้น เพราะวิธีการแปลงคะแนนดังกล่าว มีแนวคิดมาจากการแปลงเป็นคะแนนมาตรฐาน ซึ่งฐานของการแปลงคะแนนมาตรฐานนั้นอาศัยค่าคงที่เดียวกันจากข้อมูลคะแนนเดิมมาคำนวณ จึงทำให้ได้คะแนนแปลงที่มีความสัมพันธ์กับคะแนนเดิมสูงสำหรับค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน O-NET เดิมกับคะแนนแปลงที่เปรียบเทียบ โดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบมีค่าสัมประสิทธิ์สหสัมพันธ์ที่น้อยกว่า 1.00 ทั้งนี้อาจเป็นเพราะเป็นการแปลงคะแนนที่ไม่ใช่โมเดลเชิงเส้นตรง และการหาค่าคะแนนจริงนั้นต้องอาศัยการประมาณค่าจาก 3 พารามิเตอร์ การแปลงจึงไม่ได้ขึ้นกับค่าคงที่ค่าเดียว ถ้าหากพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนเดิมกับคะแนนที่แปลงแล้วของทั้ง 2 วิธี อาจมองว่าการเปรียบเทียบคะแนนเป็นมาตราเดียวกัน มีประสิทธิภาพสูงกว่าการเปรียบเทียบคะแนน โดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ แต่เมื่อพิจารณาถึงเงื่อนไขของข้อมูลคะแนนก่อนนำมาเปรียบเทียบยังพบว่า วิธีการเปรียบเทียบคะแนนเป็นมาตราเดียวกันนั้น ต้องเปรียบเทียบคะแนนจากผู้สอบกลุ่มเดียวกัน ที่ทำการทดสอบจากแบบสอบต่างฉบับ หรือเปรียบเทียบแบบสอบที่เป็นคู่ขนานอย่างแท้จริง 2 ฉบับ และสำหรับข้อมูลคะแนนในครั้งนี้อาจเงื่อนไขดังกล่าว ดังนั้นค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนเดิมกับคะแนนเปรียบเทียบที่สูงกว่าจึงไม่ควรใช้ค่าสัมประสิทธิ์สหสัมพันธ์นี้เพียงค่าเดียวเป็นเกณฑ์ที่บ่งบอกคุณภาพหรือประสิทธิภาพของวิธีการเปรียบเทียบ แต่ควรพิจารณาเงื่อนไขและเกณฑ์ที่เกี่ยวข้องอื่น ๆ ประกอบการพิจารณาคัดเลือกวิธีการเปรียบเทียบที่เหมาะสมกับข้อมูลที่มีลักษณะเฉพาะในแต่ละสถานการณ์ และผู้สร้างแบบทดสอบควรคำนึงถึงขั้นตอนในการพัฒนาแบบทดสอบที่มุ่งให้มีความเท่าเทียมกันเพื่อใช้เปรียบเทียบกันได้ เพราะรูปแบบการเปรียบเทียบตามแนวนอนนั้นระบุเงื่อนไขว่า ถึงแม้ว่าแบบทดสอบที่จะนำมาเปรียบเทียบนั้นยากที่จะมีความเท่าเทียมกันในทุกด้าน แต่ก็ควรเป็นแบบทดสอบที่พยายามสร้างหรือมุ่งให้คู่ขนานกัน (Alternate forms) นอกจากนี้การแจกแจงความสามารถของผู้สอบที่ทำการสอบด้วยแบบทดสอบแต่ละฉบับควรมีการแจกแจงที่ประมาณได้ว่ามีระดับความสามารถที่ใกล้เคียงกัน (ศิริชัย กาญจนวาสี, 2541 อ้างอิงมาจาก Hambleton & Swaminathan, 1985) เป็นต้น

4. จากข้อค้นพบที่ว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน O-NET เดิมกับคะแนนแปลงที่เปรียบเทียบโดยวิธีการเปรียบเทียบเป็นมาตราเดียวกัน มีค่าสูงกว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน O-NET เดิมกับค่าความสามารถที่เปรียบเทียบ โดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ แต่เมื่อทำการทดสอบความแตกต่างของ

ค่าสัมประสิทธิ์สหสัมพันธ์ทั้งสองค่าโดยใช้สถิติ  $t^2$ -Hotelling แล้วพบว่า ค่าสัมประสิทธิ์ทั้ง 2 ค่าไม่แตกต่างกัน ซึ่งถ้าพิจารณาเกณฑ์ในการพิจารณาคุณภาพวิธีการเทียบคะแนนแล้ว น่าจะพอสรุปได้ว่าวิธีการเทียบคะแนน O-NET ทั้ง 2 วิธีให้คุณภาพสูง เพราะทั้ง 2 วิธีต่างก็ให้คะแนนแปลงที่มีค่าสัมประสิทธิ์สหสัมพันธ์สูงและความสัมพันธ์ทางบวกกับคะแนน O-NET เดิมทั้งคู่ และให้คุณภาพที่ไม่แตกต่างกันด้วย

## ข้อเสนอแนะ

### สำหรับสถาบันทดสอบทางการศึกษาแห่งชาติ

จากผลการวิจัยพบว่าวิธีการเทียบคะแนนทั้ง 2 วิธีนี้มีประสิทธิภาพไม่แตกต่างกัน แต่ความยากง่ายในการปรับเทียบแตกต่างกัน จึงเสนอแนะเพื่อให้เป็นทางเลือกในการพิจารณาตัดสินใจนำไปใช้ ดังนี้

1. ในกรณีที่มีการพิจารณาเปรียบเทียบคะแนน O-NET ต่างปีกัน ควรพิจารณาเปรียบเทียบความเท่ากันของค่าเฉลี่ยของคะแนนในแต่ละปีก่อน ถ้าไม่เท่ากันควรมีการปรับเทียบคะแนนก่อนแล้วจึงนำมาเปรียบเทียบกัน

2. สำหรับเกณฑ์การเลือกใช้วิธีการปรับเทียบคะแนน O-NET ที่ต่างปีกัน โดยวิธีการปรับเทียบเป็นมาตรฐานเดียวกัน คือการแจกแจงของคะแนนจากแบบสอบทั้ง 2 ชุด ควรมีลักษณะคล้ายกัน หรือหากมีความแตกต่างกันก็มีเพียงเล็กน้อย เช่น อาจมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานแตกต่างกัน หรือถ้าผู้พัฒนาข้อสอบได้ทำการตรวจสอบแล้วพบว่าแบบทดสอบทั้งสองฉบับมีความเป็นคู่ขนานกัน ก็สามารถเลือกใช้วิธีการปรับเทียบคะแนนโดยวิธีนี้ได้ ซึ่งถ้าผลการศึกษาลักษณะของข้อมูลคะแนนทั้ง 2 ชุดเป็นไปตามเกณฑ์ ขอเสนอแนะให้ใช้วิธีการนี้เป็นทางเลือกในการใช้ปรับเทียบคะแนน เนื่องจากคะแนนที่ได้จากการปรับเทียบมีค่าสัมประสิทธิ์สหสัมพันธ์สูงสุดทางบวก และเป็นวิธีการที่ง่าย มีขั้นตอนที่ไม่ยุ่งยากซับซ้อน และสามารถคำนวณกับข้อมูลคะแนนจำนวนมาก ๆ ได้ เพราะไม่มีข้อจำกัดเกี่ยวกับ โปรแกรมในการวิเคราะห์ข้อมูล

3. ในการเลือกใช้วิธีการปรับเทียบคะแนน O-NET ที่ต่างปีกัน และผู้สอบมาจากต่างกลุ่มกัน ถ้าไม่มั่นใจว่าแบบทดสอบ 2 ฉบับนั้นมีความเป็นคู่ขนานกันจริงหรือไม่ วิธีการที่ปลอดภัยที่สุดตามคำแนะนำของ Lord (1977 อ้างถึงใน กาวินี ศรีสุขวัฒนานันท์, 2529) คือ ให้จัดกระทำในฐานะที่ข้อสอบนั้นไม่เป็นคู่ขนานกัน และใช้วิธีการประเภทที่ไม่ใช่โมเดลเส้นตรง นั่นคือใช้การออกแบบเพื่อศึกษาการเทียบมาตรฐานคะแนน โดยใช้ทฤษฎีการตอบสนองข้อสอบ โดยในการใช้การปรับเทียบคะแนนโดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบนี้ต้องเริ่มต้นด้วยการใช้โปรแกรมประมาณค่าพารามิเตอร์ของข้อสอบที่มีหลากหลายโปรแกรม ซึ่งถ้าผู้ทำการปรับเทียบคะแนนมีความสนใจศึกษาค้นคว้าให้เข้าใจในวิธีการวิเคราะห์ข้อมูลให้ถ่องแท้แล้ว ผู้วิจัยเห็นว่าวิธีการปรับเทียบคะแนนโดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบนี้เหมาะสมอย่างยิ่งต่อการปรับเทียบคะแนน O-NET ที่ต่างปีกันและกลุ่มผู้สอบก็ต่างกลุ่มกัน เพราะลักษณะของรูปแบบการสอบสอดคล้องกับการปรับเทียบโดยวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบมากที่สุด

## การวิจัยต่อไป

1. ควรมีการศึกษาความแตกต่างของผลการเปรียบเทียบคะแนนในรูปแบบอื่น ๆ ด้วย เช่น การเปรียบเทียบคะแนนตามแนวตั้งในข้อสอบ O-NET วิชาเดียวกัน แต่ต่างระดับชั้นกัน
2. ควรทำการศึกษารูปแบบการเปรียบเทียบคะแนน O-NET ที่ต่างปีกัน ในรายวิชาเดียวกัน โดยพิจารณาจำนวนข้อสอบรวมภายในที่อาจแทรกอยู่ในข้อสอบแต่ละฉบับ เพื่อค้นหาคำตอบว่า ถ้าข้อสอบรวมมีความยาวแตกต่างกัน ควรใช้วิธีการเปรียบเทียบใดจึงจะเหมาะสมที่สุด
3. ควรใช้วิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบปรับเทียบคะแนน O-NET ให้เป็นค่าความสามารถที่แท้จริงของผู้สอบแล้วศึกษาวิธีกำหนดเกณฑ์มาตรฐานในการจัดกลุ่มผู้เรียนเป็น กลุ่มเก่ง ปานกลาง อ่อน ตามค่าคะแนน O-NET โดยผ่านการปรับเทียบคะแนนเป็นค่าความสามารถของผู้เรียน

## เอกสารอ้างอิง

- ชูศักดิ์ ชัมภลิจิต. (2529). *การเปรียบเทียบคะแนนทดสอบ*. ในโครงการพัฒนาศึกษาอาเซียน บทคัดสรรทางวิชาการ ทดสอบ. นนทบุรี : โรงพิมพ์มหาวิทยาลัยสุโขทัยธรรมมาธิราช.
- ภาวิณี ศรีสุขวัฒนานนท์. (2529). *การเปรียบเทียบผลจากการใช้รูปแบบการเทียบมาตรฐานที่แตกต่างกันเมื่อแบบทดสอบรวมมีความยาวต่างกัน*. วิทยานิพนธ์ครุศาสตรดุษฎีบัณฑิต สาขาวิชาการวัดและประเมินผล การศึกษา, จุฬาลงกรณ์มหาวิทยาลัย.
- เรวดี อินทสระ. (2530). *การเปรียบเทียบความคลาดเคลื่อนของการเทียบมาตรฐานระหว่างรูปแบบอิงทฤษฎีการตอบข้อสอบกับรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ*. วิทยานิพนธ์การศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา, มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- ศิริชัย กาญจนวาสี. (2541). *การปรับเทียบคะแนนระหว่างแบบทดสอบ (Test Equating)*. กรุงเทพฯ: โรงพิมพ์แห่ง จุฬาลงกรณ์มหาวิทยาลัย.
- สถาบันทดสอบทางการศึกษาแห่งชาติ ,<http://www.niets.or.th/pdf/history.pdf> สืบค้นวันที่ 7 พฤศจิกายน 2551.
- สถาบันทดสอบทางการศึกษาแห่งชาติ ,<http://www.niets.or.th/pdf/obligation.pdf> สืบค้นวันที่ 7 พฤศจิกายน 2551.
- สถาบันทดสอบทางการศึกษาแห่งชาติ ,<http://www.niets.or.th/pdf/authority.pdf> สืบค้นวันที่ 7 พฤศจิกายน 2551.
- สำนักงานคณะกรรมการการศึกษาแห่งชาติ.สำนักงานกฤษฎมนตรี. (2545). *พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 และที่แก้ไขเพิ่มเติม (ฉบับที่ 2) พ.ศ.2545*. กรุงเทพฯ: บริษัทพริกหวานจำกัด.
- สำราญ มีแจ้ง. (2542). *การเปรียบเทียบคะแนนสอบคัดเลือกเข้าศึกษาในสถาบัน อุดมศึกษาระบบใหม่ที่มีวิธีแปลงคะแนนผลการเรียนสะสมต่างกัน*. *วารสารศึกษาศาสตร์ มหาวิทยาลัยนเรศวร*.4(1), 19-41.

- สุภาภรณ์ คงทวี. (2541). การปรับแก้เฉลี่ยสะสมระดับมัธยมศึกษาตอนปลายด้วยคะแนนสอบเข้ามหาวิทยาลัย : การเปรียบเทียบระหว่างวิธีการปรับเทียบแบบอิกวิเปอร์เซน ไทล์เชิงเส้นตรงและไออาร์ที. วิทยานิพนธ์ ครุศาสตรมหาบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา, จุฬาลงกรณ์มหาวิทยาลัย.
- สุรัตน์ ขวัญบุญจันทร์. (2531). การสร้างแบบทดสอบและตารางเทียบคะแนนในแนวนอนวิชาวิทยาศาสตร์ที่วิเคราะห์ด้วยวิธีราสช์โมเดล. วิทยานิพนธ์การศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา มหาวิทยาลัย ศรีนครินทร วิโรฒประสานมิตร.
- อรุณี อ่อนสวัสดิ์ และชอบ ลีซอ (2541). ผลของวิธีการแปลงผลการเรียนสะสมในการสอบคัดเลือกบุคคลเข้าศึกษา ในสถาบันอุดมศึกษาระบบใหม่. *วารสารศึกษาศาสตร์ มหาวิทยาลัยนเรศวร*. 3(2) , 99-121.
- Angoff, W.H. (1971). *Scales, Norms, and Equivalent Scores. Educational Measurement*. 2<sup>nd</sup> ed. Washington, D.C. : American Council on Education, 508-600; In R.L. Thorndike (Ed.),
- Flanagan, T.C. (1951). *Unit, Scores and Norms*. in E.F. Lindquist (Ed.) *Educational Measurement*. Washington D.C. : American Council on Education, 695-763.
- Hambleton, R.K. & Swaminathan, H. (1985). *Item Response Theory*. Kluwer. Nijhoff, Boston.
- Kolen. M.J. (1981). Comparison of traditional and Item Response Theory Methods for Equating test. *Journal of Educational Measurement*, 18 (1), 1-11.
- Lord, F.M. ( 1977). Practical Applications of Item Characteristic Curve Theory. *Journal of Educational Measurement*, 14, 117 – 138.
- Rentz & Bashaw. (1977). The National Reference Scale for Reading : An Application of the Rasch Model. *Journal of Educational Measurement*, 14, 161 – 178.
- Young. J.W. (1995). *A Comparison of Two Adjustment Methods for Improving the Prediction of Law School*. *Educational and Psychological Measurment*, 55, 558-571.