

บทที่ 5

บทสรุป

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาวิธีการปรับเทียบผลการเรียน โดยใช้การปรับสเกล และทำการตรวจสอบคุณภาพของวิธีการปรับเทียบผลการเรียนที่พัฒนาขึ้นกับข้อมูลจำลอง โดยพิจารณาความสามารถของวิธีการปรับเทียบผลการเรียนที่ผู้วิจัยพัฒนาขึ้นในการลดความเอนเอียงของผลการเรียนเดิม จากค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผลการเรียนกับคะแนน O-NET และค่าดัชนีความแตกต่าง (Discrepancy Index) ของผลการเรียนที่ปรับเทียบแล้วของแต่ละวิธีกับผลการเรียนเกณฑ์ โดยมีรายละเอียดดังนี้

1. การวิจัยครั้งนี้เป็นการศึกษาและพัฒนาวิธีการปรับเทียบผลการเรียนแบบปรับสเกลขั้นเดียว และวิธีการปรับเทียบผลการเรียนแบบปรับสเกลสองขั้น เพื่อใช้ในการปรับเทียบผลการเรียนระหว่างโรงเรียนให้มีมาตรฐานเดียวกัน และสามารถเปรียบเทียบกันได้ คะแนนที่นำมาใช้ในการปรับเทียบ เป็นผลการเรียนเฉลี่ยสะสมตลอดหลักสูตร (GPAX: Cumulative GPA) และผลการสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (Ordinary National Educational Test: O-NET) ซึ่งเป็นองค์ประกอบของระบบการรับบุคคลเข้าศึกษาในสถาบันอุดมศึกษาในระบบกลาง (Admissions)

2. ข้อมูลที่ใช้ในการวิจัยครั้งนี้เป็นข้อมูลจำลอง ซึ่งได้จากการจำลองสถานการณ์ (Simulation) ด้วยเครื่องมือโครคอมพิวเตอร์ ตามผลการสอบในระบบกลางการรับบุคคลเข้าศึกษาในสถาบันอุดมศึกษา (Admissions) ประจำปีการศึกษา 2550 จำนวนทั้งหมด 2,177 โรงเรียน นักเรียนจำนวน 65,991 คน ซึ่งประกอบด้วย ผลการเรียนเฉลี่ยสะสมตลอดหลักสูตร (GPAX) และผลการสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) รวมทั้ง 5 รายวิชา คือ วิชาภาษาไทย สังคมศึกษา ภาษาอังกฤษ คณิตศาสตร์ และวิทยาศาสตร์

3. ตัวแปรที่ศึกษามีดังต่อไปนี้

ตัวแปรที่ใช้ในการตรวจสอบคุณภาพของวิธีการปรับเทียบผลการเรียนที่พัฒนาขึ้น ประกอบด้วย

3.1 ตัวแปรอิสระ ได้แก่ วิธีการปรับเทียบผลการเรียน ซึ่งจำแนกได้ดังนี้

- 1) วิธีการปรับเทียบผลการเรียนแบบปรับสเกลขั้นเดียว (One-step Scaling Calibration : OSC)
- 2) วิธีการปรับเทียบผลการเรียนแบบปรับสเกลสองขั้น (Two-Step Scaling Calibration : TSC)

- 3) วิธีการปรับเทียบผลการเรียนแบบเชื่อมโยงส่วนเดียว (One-Section Linking Grade Calibration: OLC)
- 4) วิธีการปรับเทียบผลการเรียนแบบเชื่อมโยงเสริมกันทั้งสองส่วน (Two-Section Additive Linking Grade Calibration: TALC)
- 5) วิธีการปรับเทียบผลการเรียนแบบเชื่อมโยงร่วมกันทั้งสองส่วน (Two-Section Interactive Linking Grade Calibration: TILC)

3.2 ตัวแปรตาม ได้แก่ คุณภาพของการปรับเทียบผลการเรียน ซึ่งพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ และความคลาดเคลื่อน (Error) ของการปรับเทียบ โดยการหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผลการเรียนกับคะแนน O-NET และดัชนีความแตกต่าง (Discrepancy Index) ระหว่างผลการเรียนกับผลการเรียนเกณฑ์ โดยใช้ค่าดัชนีความแตกต่าง RMSD (Root Mean Square Difference) และ MAD (Mean of Absolute Difference)

สรุปผลการวิจัย

การวิจัยครั้งนี้สรุปผลได้ตามวัตถุประสงค์ของการวิจัยดังนี้

ตอนที่ 1 ผลการพัฒนาวิธีการปรับเทียบผลการเรียน

การพัฒนาวิธีการปรับเทียบผลการเรียน พิจารณาจากกรอบแนวคิดที่ได้จากการศึกษาปัญหามาตรฐานของผลการเรียนระหว่างโรงเรียน และการเอนเอียง (Bias) ของผลการเรียนภายในโรงเรียน ทำให้ไม่สามารถนำผลการเรียนมาเปรียบเทียบกันได้ ซึ่งจากผลการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง ได้กรอบแนวคิดการพัฒนาวิธีการปรับเทียบผลการเรียนดังนี้

1. ผลการเรียนของนักเรียนแต่ละคน เป็นผลจากการกำหนดระดับผลสัมฤทธิ์หรือความรู้ความสามารถ อันเป็นผลมาจากประสบการณ์การเรียนรู้ของผู้เรียน การกำหนดระดับผลสัมฤทธิ์ดังกล่าวขึ้นอยู่กับส่วนประกอบที่สำคัญ 3 ประการ คือ ส่วนแรก เป็นความรู้ความสามารถของนักเรียน โดยส่วนประกอบนี้เป็นสิ่งที่เราต้องการประเมินโดยอาศัยการวัดที่ครอบคลุม การวัดที่แม่นยำเชื่อถือได้ การแปลความหมายที่เหมาะสม และใช้วิธีการตัดสินคุณค่าที่ยุติธรรม (ศิริชัย กาญจนวาสี, 2548 : 247) ส่วนที่สอง เป็นระบบการตัดสินผลการเรียนแบบกำหนดระดับผลสัมฤทธิ์เป็นตัวอักษร (Letter grade system) ซึ่งเป็นกระบวนการที่ใช้ในการกำหนดผลการเรียน โดยระบบการตัดสินผลการเรียนแบบกำหนดระดับผลสัมฤทธิ์เป็นตัวอักษรนี้ เป็นส่วนที่ทำให้ผลการเรียนของนักเรียนเอนเอียงไปจากความสามารถที่ควรจะเป็น (Cohen, 2000 : 1-4) และส่วนที่สาม เป็นบริบทของโรงเรียนหรือเป็นบริบทของการประเมินผล ซึ่งต่างโรงเรียนก็จะมีบริบทต่างกันไป โดยส่วนประกอบนี้จะประกอบไปด้วย นโยบายของโรงเรียน คุณลักษณะของครูผู้สอน

คุณลักษณะของนักเรียน เป็นต้น ซึ่งปัจจัยเหล่านี้ส่งผลให้มาตรฐานของผลการเรียนของแต่ละโรงเรียนแตกต่างกันไป เนื่องจากปัจจัยเหล่านี้มีผลต่อนักเรียนทั้งหมดในแต่ละโรงเรียน ซึ่งถือว่าเป็นประชากรหรือกลุ่มผู้สอบ (population) โดยตามหลักของการวัดและประเมินผลทางการศึกษาแล้ว นักวัดผลการศึกษาต่างยอมรับกันว่า การวัดและประเมินผลการเรียนรู้ นั้น ขึ้นอยู่กับกลุ่มประชากรหรือกลุ่มผู้สอบ (Flanagan. 1951; Angoff. 1971) ดังนั้นผลการเรียนของแต่ละโรงเรียนจึงมีมาตรฐานแตกต่างกันไปตามประชากรซึ่งก็คือนักเรียนของแต่ละโรงเรียน

จากส่วนประกอบทั้งสามส่วน จะเห็นว่า ส่วนที่สองเป็นส่วนที่ทำให้ผลการเรียนของนักเรียนภายในโรงเรียนเอนเอียงไปจากผลการเรียนที่ควรจะเป็น และส่วนที่สามเป็นส่วนที่ทำให้มาตรฐานของผลการเรียนแตกต่างกันระหว่างโรงเรียน

2. การปรับเทียบผลการเรียนระหว่างนักเรียนทั้งหมด เพื่อให้มีมาตรฐานเดียวกัน และสามารถเปรียบเทียบกันได้ (comparable) จึงต้องลดความเอนเอียง (bias) ของผลการเรียนทั้งสองส่วน คือ ส่วนของโรงเรียนและส่วนของผู้เรียน โดยใช้การปรับสเกลผลการเรียนทั้งในส่วนโรงเรียนและในส่วนของผู้เรียน จึงเรียกการปรับเทียบผลการเรียนลักษณะนี้ว่า “การปรับสเกลสองขั้น” (Two-Stage Scaling) โดยในระบบการศึกษาของไทย มีการทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน หรือ O-NET ที่นักเรียนทุกคนต้องทำการทดสอบ ซึ่งดำเนินการทดสอบโดยหน่วยงานกลาง คือ สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) และใช้แบบสอบเดียวกันทั่วประเทศ คุณลักษณะที่วัดในการทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน หรือ O-NET ก็เป็นคุณลักษณะเดียวกันกับการประเมินผลการเรียนรู้ ผลการทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน หรือ O-NET จึงมีความเหมาะสมที่จะใช้เป็นสเกลอ้างอิง (reference scale) ในการปรับเทียบสเกล (calibration) เนื่องจากผลการเรียนของนักเรียนในแต่ละโรงเรียนนั้น ประเมินผลจากคะแนนที่ได้จากแบบสอบของแต่ละโรงเรียน ซึ่งมีค่าสถิติของแบบสอบแตกต่างกัน แต่วัดในคุณลักษณะเดียวกัน (Kolen. 2004 : 224)

3. ความเอนเอียงของผลการเรียนในส่วน of นักเรียน ซึ่งเกิดจากระบบการตัดสินผลการเรียนแบบกำหนดระดับผลสัมฤทธิ์เป็นตัวอักษร และความเอนเอียงของผลการเรียนในส่วน of โรงเรียน ซึ่งก็คือ มาตรฐานของผลการเรียนที่แตกต่างกันระหว่างโรงเรียน จะมีผลต่อผลการเรียนของนักเรียนในแนวทางใดแนวทางหนึ่งดังนี้

3.1 ความเอนเอียงของผลการเรียนในส่วน of นักเรียนเป็นแบบสุ่ม (Random Effect) กล่าวคือ หากอิทธิพลของระบบการตัดสินผลการเรียนเป็นอิทธิพลแบบสุ่ม ความเอนเอียงของผลการเรียนของนักเรียนแต่ละคนภายในแต่ละโรงเรียนก็จะเกิดขึ้นอย่างสุ่ม นั่นคือ ระบบการตัดสินผลการเรียน อาจทำให้ผลการเรียนของนักเรียนแต่ละคนเอนเอียงในทางถูกกดเกรด ถูกปล่อยเกรด หรือ ไม่มีความเอนเอียงก็ได้ ซึ่งถ้าหากภายใน 6 ภาคเรียนที่มีการคิดผลการเรียนเฉลี่ยสะสม

(GPA) นั้น ค่าความเอนเอียงเฉลี่ยของผลการเรียนของนักเรียนแต่ละคนจะเป็นศูนย์ แสดงว่า ระบบการตัดสินผลการเรียนเป็นอิทธิพลแบบสุ่ม หรือ ไม่มีความเอนเอียงของผลการเรียนในส่วนของนักเรียน การปรับเทียบผลการเรียนให้มีมาตรฐานเดียวกันจึงเป็นการปรับความเอนเอียงของผลการเรียนในส่วนของโรงเรียนเพียงอย่างเดียว โดยการปรับเทียบค่าเฉลี่ยของผลการเรียนของแต่ละโรงเรียนกับคะแนน O-NET เฉลี่ยของแต่ละ โรงเรียน การปรับเทียบผลการเรียนแบบนี้เรียกว่า “การปรับเทียบผลการเรียนแบบปรับสเกลชั้นเดียว” (OSC) ซึ่งมีสูตรการปรับเทียบดังนี้

$$GPAX_{ij}^* = \left[\overline{ONET}_j \times \left(\frac{\overline{GPAX}_{\max} - \overline{GPAX}_{\min}}{\overline{ONET}_{\max} - \overline{ONET}_{\min}} \right) \right] + \Delta GPAX_{ij}$$

3.2 ความเอนเอียงของผลการเรียนในส่วนของนักเรียน และความเอนเอียงของผลการเรียนในส่วนของโรงเรียนมีผลต่อผลการเรียนของนักเรียนแบบเสริมกัน (Additive Effect) เช่น ในโรงเรียนที่นักเรียนส่วนใหญ่มีความสามารถอยู่ในระดับปานกลางและอ่อนเมื่อเทียบกับโรงเรียนอื่นๆ ถ้าหากโรงเรียนใช้แบบสอบที่ยากในการประเมินผล ย่อมทำให้นักเรียนส่วนใหญ่สอบตก จึงต้องใช้แบบสอบที่ง่ายเมื่อเปรียบเทียบกับแบบสอบของโรงเรียนอื่นๆ ทำให้ผลการเรียนของนักเรียนในโรงเรียนนี้สูงกว่าที่ควรจะเป็น (ถูกปล่อยเกรดทั้ง โรงเรียน) ในขณะเดียวกันก็ยังมีนักเรียนในโรงเรียนนี้อีกส่วนหนึ่งที่ได้รับผลจากระบบการตัดสินผลการเรียนให้มีผลการเรียนสูงขึ้นอีก คือ นักเรียนกลุ่มที่ได้คะแนนอยู่ในขีดจำกัดล่างของช่วงคะแนนในแต่ละระดับผลการเรียน เช่น นักเรียนที่ได้ 80 คะแนน ในช่วงคะแนน 80-100 มีผลการเรียนเป็น 4.00 หรือนักเรียนที่ได้ 75 คะแนน ในช่วงคะแนน 75-79 มีผลการเรียนเป็น 3.50 เป็นต้น ดังนั้นผลการเรียนของนักเรียนกลุ่มนี้จึงเกิดจากความเอนเอียงของผลการเรียนในส่วนของโรงเรียนเสริมกับความเอนเอียงของผลการเรียนในส่วนของผู้เรียน ในกรณีนี้การปรับเทียบผลการเรียนให้มีมาตรฐานเดียวกัน และสามารถเปรียบเทียบกันได้ จึงต้องปรับทั้งความเอนเอียงของผลการเรียนในส่วนของโรงเรียน โดยการปรับเทียบค่าเฉลี่ยของผลการเรียนของแต่ละโรงเรียนกับคะแนน O-NET เฉลี่ยของแต่ละ โรงเรียน และปรับความเอนเอียงของผลการเรียนในส่วนของนักเรียน โดยปรับเทียบผลการเรียนของนักเรียนแต่ละคนที่ต่างจากผลการเรียนของโรงเรียน (ผลการเรียนเฉลี่ยภายในโรงเรียน) กับคะแนน O-NET ของนักเรียนแต่ละคนที่ต่างจากคะแนน O-NET ของโรงเรียน (ค่าเฉลี่ยคะแนน O-NET ภายในโรงเรียน) ด้วยการปรับเทียบผลการเรียนแบบนี้เรียกว่า “การปรับเทียบผลการเรียนแบบปรับสเกลสองชั้น” (TSC) ซึ่งมีสูตรการปรับเทียบดังนี้

$$GPAX_{ij}^* = \left[\overline{ONET}_j \times \left(\frac{\overline{GPAX}_{\max} - \overline{GPAX}_{\min}}{\overline{ONET}_{\max} - \overline{ONET}_{\min}} \right) \right] + \left[\Delta ONET_{ij} \times \left(\frac{\Delta GPAX_{\max} - \Delta GPAX_{\min}}{\Delta ONET_{\max} - \Delta ONET_{\min}} \right) \right]$$

ตอนที่ 2 ผลการตรวจสอบคุณภาพของวิธีการเปรียบเทียบ

ผลการตรวจสอบคุณภาพของวิธีการเปรียบเทียบผลการเรียนที่ผู้วิจัยพัฒนาขึ้นสรุปได้ดังนี้

1. ผลการเรียนที่เปรียบเทียบด้วยวิธีการปรับสเกลชั้นเดียว (OSC) และวิธีการปรับสเกลสองชั้น (TSC) ที่ผู้วิจัยพัฒนาขึ้น มีความสัมพันธ์กับคะแนน O-NET สูงขึ้นตามลำดับ โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผลการเรียนที่เปรียบเทียบกับคะแนน O-NET สูงกว่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผลการเรียนเดิมกับคะแนน O-NET อย่างมีนัยสำคัญทางสถิติที่ระดับ .001 และวิธีการเปรียบเทียบผลการเรียนแบบปรับสเกลสองชั้น (TSC) เป็นวิธีที่มีคุณภาพดีที่สุด

2. ค่าดัชนีความแตกต่างของค่ารากที่สองของผลต่างกำลังสองเฉลี่ย (RMSD) ระหว่างผลการเรียนเกณฑ์กับผลการเรียนที่เปรียบเทียบโดยวิธีการปรับสเกลสองชั้น (TSC) มีค่าต่ำสุดถัดมาเป็นผลการเรียนที่เปรียบเทียบด้วยวิธีการปรับสเกลชั้นเดียว (OSC) แสดงว่าวิธีการเปรียบเทียบผลการเรียนแบบปรับสเกลสองชั้น (TSC) เป็นวิธีที่มีคุณภาพดีที่สุด

ตอนที่ 3 ผลการเปรียบเทียบคุณภาพของวิธีการเปรียบเทียบผลการเรียน

ผลเปรียบเทียบค่าเฉลี่ยของผลต่างสัมบูรณ์ (MAD) ของผลการเรียนที่เปรียบเทียบแล้วของแต่ละวิธีกับผลการเรียนเกณฑ์ โดยใช้การวิเคราะห์ความแปรปรวนทางเดียวแบบวัดซ้ำ (One-Way ANOVA with repeated measure) นั้น มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 โดยวิธีการปรับสเกลสองชั้น (TSC) มีดัชนีค่าเฉลี่ยผลต่างสัมบูรณ์ (MAD) ต่ำกว่าวิธีการเชื่อมโยงเสริมกันสองส่วน (TALC) วิธีการปรับสเกลชั้นเดียว (OSC) และวิธีการเชื่อมโยงส่วนเดียว (OLC) ส่วนวิธีการปรับสเกลชั้นเดียว (OSC) และวิธีการเชื่อมโยงเสริมกันสองส่วน (TALC) มีดัชนีค่าเฉลี่ยผลต่างสัมบูรณ์ (MAD) ต่ำกว่า วิธีการเชื่อมโยงส่วนเดียว (OLC) อย่างมีนัยสำคัญทางสถิติที่ระดับ .001 ในขณะที่วิธีการปรับสเกลชั้นเดียว (OSC) กับวิธีการเชื่อมโยงเสริมกันสองส่วน (TALC) และวิธีการปรับสเกลสองชั้น (TSC) กับวิธีการเชื่อมโยงร่วมกันสองส่วน (TILC) มีดัชนีค่าเฉลี่ยผลต่างสัมบูรณ์ (MAD) ไม่แตกต่างกัน แสดงว่าวิธีการเปรียบเทียบผลการเรียนแบบปรับสเกลสองชั้น (TSC) และวิธีการเชื่อมโยงร่วมกันสองส่วน (TILC) เป็นวิธีที่มีคุณภาพดีที่สุด

อภิปรายผลการวิจัย

1. ผลการพัฒนาวิธีการเปรียบเทียบผลการเรียน พบว่า วิธีการเปรียบเทียบผลการเรียนที่พัฒนาขึ้นโดยอาศัยหลักการปรับสเกล ตามอิทธิพลของความเอนเอียงของผลการเรียนในสองส่วน คือ ความเอนเอียงของผลการเรียนในส่วนของโรงเรียน และความเอนเอียงของผลการเรียนในส่วนของผู้เรียน ทำให้ได้วิธีการเปรียบเทียบผลการเรียน 2 วิธี คือ วิธีการเปรียบเทียบผลการเรียนแบบปรับสเกลชั้นเดียว และวิธีการเปรียบเทียบผลการเรียนแบบปรับสเกลสองชั้น

ผลการพัฒนาวิธีการปรับเทียบผลการเรียนดังกล่าว มีกรอบแนวคิดมาจากผลการเรียนของนักเรียนแต่ละคน เป็นผลจากการกำหนดระดับผลสัมฤทธิ์หรือความรู้ความสามารถจากประสบการณ์การเรียนรู้ของผู้เรียน การกำหนดระดับผลสัมฤทธิ์ดังกล่าวขึ้นอยู่กับส่วนประกอบที่สำคัญ 3 ประการ คือ ประการแรก เป็นความรู้ความสามารถที่แท้จริงของนักเรียน โดยส่วนนี้เป็นสิ่งที่เราต้องการประเมิน โดยอาศัยการวัดที่ครอบคลุม การวัดที่แม่นยำเชื่อถือได้ การแปลความหมายที่เหมาะสม และใช้วิธีการตัดสินคุณค่าที่ยุติธรรม (ศิริชัย กาญจนวาสี. 2548 : 247) ประการที่สอง เป็นระบบการตัดสินผลการเรียนแบบกำหนดระดับผลสัมฤทธิ์เป็นตัวอักษร (Letter grade system) ซึ่งเป็นกระบวนการที่ใช้ในการกำหนดผลการเรียน โดยระบบการตัดสินผลการเรียนแบบกำหนดระดับผลสัมฤทธิ์เป็นตัวอักษรนี้ เป็นส่วนที่ทำให้ผลการเรียนของนักเรียนเอนเอียงไปจากความสามารถที่ควรจะเป็น (Cohen. 2000 : 1-4) และประการที่สาม เป็นบริบทของโรงเรียนหรือเป็นบริบทของการประเมินผล ซึ่งต่างโรงเรียนก็จะมีบริบทต่างกันไป โดยส่วนนี้จะประกอบไปด้วย นโยบายของโรงเรียน คุณลักษณะของครูผู้สอน คุณลักษณะของนักเรียน เป็นต้น ซึ่งปัจจัยเหล่านี้ส่งผลให้มาตรฐานของผลการเรียนของแต่ละโรงเรียนแตกต่างกันไป เนื่องจากปัจจัยเหล่านี้มีผลต่อนักเรียนทั้งหมดในแต่ละโรงเรียน ซึ่งถือว่าเป็นประชากรหรือกลุ่มผู้สอบ (population) โดยตามหลักของการวัดและประเมินผลทางการศึกษาแล้ว นักวัดผลการศึกษิต่างยอมรับกันว่า การวัดและประเมินผล การเรียนนั้นขึ้นอยู่กับกลุ่มประชากรหรือกลุ่มผู้สอบ (Flanagan. 1951; Angoff. 1971) ดังนั้นผลการเรียนที่มาจากแต่ละโรงเรียนจึงมีมาตรฐานแตกต่างกันไปตามประชากรซึ่งก็คือนักเรียนของแต่ละโรงเรียนนั่นเอง เมื่อพิจารณาส่วนประกอบทั้งสามส่วน จะเห็นว่ามีส่วนที่ทำให้ผลการเรียนของนักเรียนเอนเอียงไปจากผลการเรียนที่ควรจะเป็น นั่นคือ ระบบการตัดสินผลการเรียนแบบกำหนดระดับผลสัมฤทธิ์เป็นตัวอักษร และบริบทของการประเมินผลของโรงเรียน หรืออาจกล่าวได้ว่า ความเอนเอียงของผลการเรียนนั้น ประกอบด้วยความเอนเอียงในส่วนของผู้เรียน และความเอนเอียงในส่วนของโรงเรียน

ความเอนเอียงในส่วนของโรงเรียน เมื่อพิจารณาจากพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 และที่แก้ไขเพิ่มเติม (ฉบับที่ 2) พ.ศ. 2545 มาตรา 26 วรรคแรก ที่ระบุว่า “ให้สถานศึกษาจัดการประเมินผู้เรียน โดยพิจารณาจากพัฒนาการของผู้เรียน ความประพฤติ การสังเกตพฤติกรรม การเรียน การร่วมกิจกรรมและการทดสอบควบคู่ไปในกระบวนการเรียนการสอนตามความเหมาะสมของแต่ละระดับและรูปแบบการศึกษา” (สำนักงานคณะกรรมการการศึกษาแห่งชาติ. 2543 : 15) การประเมินผลการเรียนดังกล่าวจึงต้องพิจารณาด้วยวิธีที่หลากหลาย และพิจารณาพัฒนาการของผู้เรียนเป็นสำคัญ นั่นคือการพิจารณาโดยการอิงตัวผู้เรียนเอง ดังนั้นการวัดและประเมินผลในโรงเรียนต่างๆ จึงมีคุณภาพต่างกัน (ธีระพร วีระถาวร. 2541 : 7) ส่วนหนึ่งมาจากนโยบายของ

โรงเรียนที่ต้องการให้นักเรียนภายในโรงเรียนสามารถเข้าศึกษาต่อในระดับอุดมศึกษาได้ ดังที่ สุชาติ บวรกิติวงศ์ และคณะ (2547) ที่กล่าวว่า “ปัจจัยที่สำคัญที่มีผลต่อการเบี่ยงเบนของผลการเรียนสะสมตลอดหลักสูตร (GPAX) ของนักเรียน คือ มาตรฐานการประเมินผลการเรียนของโรงเรียน เนื่องจากผลการเรียนเฉลี่ยสะสมตลอดหลักสูตร (GPAX) ซึ่งเป็นผลการเรียนเฉลี่ยที่เกิดจากการประเมินผลการเรียนรู้ในทุกวิชา ทุกชั้นปีของหลักสูตรการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 4 จึงเป็นผลจากทั้งนโยบายของโรงเรียน และการประเมินผลของครูทุกคนในโรงเรียนที่สอนในช่วงชั้นดังกล่าว ส่วนผลการเรียนเฉลี่ยสะสมตามกลุ่มสาระการเรียนรู้ (GPA) นั้นขึ้นกับมาตรฐานการประเมินผลการเรียนรู้ในแต่ละกลุ่มวิชาของแต่ละโรงเรียน ซึ่งกลุ่มวิชาที่ต่างกันมักจะมีมาตรฐานการประเมินผลที่แตกต่างกัน เมื่อระบบการรับเข้าศึกษานำผลการเรียนสะสมตลอดหลักสูตร (GPAX) และผลการเรียนสะสมตามกลุ่มสาระการเรียนรู้ (GPA) มาใช้จึงควรมีการเปรียบเทียบผลการเรียนก่อน เพื่อความเป็นธรรมและสอดคล้องกับความสามารถที่แท้จริงของผู้เรียน” นับเป็นการสนับสนุนว่า มาตรฐานการประเมินผลในแต่ละโรงเรียนแตกต่างกัน

ในส่วนของระบบการตัดสินผลการเรียนที่กำหนดระดับผลสัมฤทธิ์เป็นตัวอักษรนั้น ก็มีผลทำให้ผลการเรียนของผู้เรียนแต่ละคน ไม่สอดคล้องกับความสามารถที่แท้จริงของผู้เรียนด้วย (Cohen. 2000 : 1-4) เนื่องจากระบบการตัดสินผลการเรียนที่กำหนดระดับผลสัมฤทธิ์เป็นตัวอักษรนั้น เป็น การตัดสินผลการเรียนโดยใช้ช่วงคะแนน ถึงแม้ตามหลักการกำหนดช่วงคะแนนผู้เรียนที่ได้ระดับ ผลการเรียนเดียวกันจะมีความรู้ความสามารถใกล้เคียงกัน และผู้เรียนที่ได้ระดับผลการเรียนต่างกัน จะมีความรู้ความสามารถต่างกัน แต่ในทางปฏิบัตินั้นทำได้ยาก เพราะคะแนนของผู้เรียนมักจะ กระจาย ไม่ได้เกาะกลุ่มอย่างชัดเจน ทำให้เกิดความเอนเอียงขึ้นในระบบการตัดสินผลการเรียน เช่น นักเรียนที่ได้ 80 คะแนน ในช่วงคะแนน 80-100 มีผลการเรียนเป็น 4.00 ในขณะที่นักเรียนที่ได้ 79 คะแนน ในช่วงคะแนน 75-79 มีผลการเรียนเป็น 3.50 นักเรียนสองคนนี้ย่อมมีความรู้ความสามารถ ใกล้เคียงกัน แต่กลับได้รับผลการเรียนที่แตกต่างกัน เป็นต้น

อีกกรณีหนึ่งที่อาจเกิดความเอนเอียงของผลการเรียนขึ้นได้ก็คือ เกิดความเอนเอียงในส่วน ของผู้เรียนเพียงส่วนเดียว โดยมาตรฐานการประเมินผลระหว่างโรงเรียนนั้นมีมาตรฐานเดียวกัน แต่ กรณีนี้มีโอกาสเกิดขึ้นน้อยมากในระบบการศึกษาของไทย เพราะเป็นที่ประจักษ์ว่าในปัจจุบันการ วัดและประเมินผลในโรงเรียนต่างๆ มีคุณภาพต่างกัน และคุณภาพของครู คุณภาพของนักเรียนใน แต่ละโรงเรียนก็แตกต่างกัน (ธีระพร วีระถาวร. 2541 : 7; สำนักงานคณะกรรมการการศึกษา แห่งชาติ. 2543; สุชาติ บวรกิติวงศ์ และคณะ. 2547; Kanjanavasee and others. 2009 : 117-124)

การพัฒนาวิธีการเปรียบเทียบผลการเรียนทั้งสองวิธี จึงมุ่งลดความเอนเอียงของผลการเรียนที่ อาจเกิดขึ้นในสองส่วนคือ ในส่วนของโรงเรียนและในส่วนของผู้เรียน ตามลักษณะของความเอน

เสียงที่อาจจะเกิดขึ้นได้ โดยอาศัยหลักการปรับสเกล (Scaling) 'ได้แก่' ความเอนเอียงในส่วนผู้เรียนเกิดขึ้นอย่างสุ่ม และกรณีความเอนเอียงในส่วนของโรงเรียนและส่วนของผู้เรียนเกิดขึ้นแบบเสริมกัน

ในการรับบุคคลเข้าศึกษาในระบบกลางการรับเข้าศึกษา (Admissions) ในปีการศึกษา 2553 เป็นต้นไปจะมีองค์ประกอบที่นำมาพิจารณาอยู่ 3 องค์ประกอบ คือ ผลการเรียนเฉลี่ยสะสมตลอดหลักสูตร (GPAX) ผลการสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) และผลการสอบความถนัดทั่วไป/เฉพาะ (GAT/PAT) จาก 3 องค์ประกอบดังกล่าว จะพบว่า ผลการเรียนเฉลี่ยสะสมตลอดหลักสูตร (GPAX) เป็นองค์ประกอบของผลการเรียนที่ได้มาจากโรงเรียน ดังนั้นจึงเป็นที่วิตกกังวลในหมู่นักวิชาการและบุคคลที่เกี่ยวข้องว่าจะทำให้โรงเรียนช่วยนักเรียนของตนเองเพื่อให้มีผลการเรียนสูงกว่าความสามารถที่แท้จริง (สำนักงานคณะกรรมการการศึกษาแห่งชาติ, 2543) การปรับปรุงคุณภาพและมาตรฐานของโรงเรียนในด้านประเมินผล เพื่อให้ผลการเรียนมีมาตรฐานเดียวกัน จำเป็นต้องอาศัยระยะเวลายาวนาน ในการที่จะทำให้คุณภาพการประเมินผลของแต่ละโรงเรียนมีคุณภาพใกล้เคียงกัน ดังนั้นในระยะสั้นจึงควรมีการปรับค่าผลการเรียนให้ใกล้เคียงกับคุณภาพของโรงเรียน (สำนักงานคณะกรรมการการศึกษาแห่งชาติ, 2543)

วิธีการปรับเทียบผลการเรียนที่ผู้วิจัยพัฒนาขึ้น จึงอยู่บนข้อตกลงเบื้องต้นที่ว่าผลการสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) ที่นักเรียนชั้นมัธยมศึกษาปีที่ 6 ทุกคนต้องสอบ และดำเนินการโดยสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) จะสามารถประมาณค่าความรู้ความสามารถของผู้เรียนได้ดีที่สุด ซึ่งวัตถุประสงค์เบื้องต้นของการทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) ก็เพื่อประโยชน์ในการประกันคุณภาพการศึกษาและเป็นดัชนีบ่งชี้เพื่อใช้ในการปรับปรุงคุณภาพการเรียนการสอนของโรงเรียน ตามหลักการประเมินผลแล้ว ผลการเรียนที่ได้จากโรงเรียนไม่ควรแตกต่างจากคะแนนสอบ O-NET อย่างมีนัยสำคัญ (งานประชาสัมพันธ์, สำนักงานคณะกรรมการการอุดมศึกษา, 2549 : 1; สุธา เหลือถมัย, 2549 : 7) แสดงว่า ผลการเรียนเฉลี่ยสะสมและผลการสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) จึงควรมีความสัมพันธ์กันสูง วิธีการปรับเทียบที่พัฒนาขึ้นจึงใช้ผลการสอบทางการศึกษาแห่งชาติขั้นพื้นฐานเป็นสเกลอ้างอิง (Reference Scale) โดยใช้หลักการปรับสเกลซึ่งจะทำให้ผลการเรียนสอดคล้องกับสเกลอ้างอิง เป็นการลดความเอนเอียงของผลการเรียนให้ได้มากที่สุด ซึ่งจะช่วยให้ความแตกต่างระหว่างผลการเรียนกับความสามารถที่แท้จริงแตกต่างกันน้อยที่สุด

จากกรอบแนวคิดดังกล่าว ทำให้สามารถพัฒนาวิธีการปรับเทียบผลการเรียน โดยอาศัยการปรับสเกล ทำให้ได้วิธีการปรับเทียบผลการเรียน 2 วิธีดังนี้

1. วิธีการปรับเทียบผลการเรียนแบบปรับสเกลขั้นเดียว (One-step Scaling Grade Calibration: OSC) เป็นการปรับค่าความเอนเอียงของผลการเรียนในส่วนของโรงเรียนเพียงส่วนเดียว โดยถือว่าความเอนเอียงในส่วนของผู้เรียนนั้นเกิดขึ้นอย่างสุ่ม จึงไม่มีผลต่อผลการเรียนเฉลี่ยสะสม

2. วิธีการปรับเทียบผลการเรียนแบบปรับสเกลสองขั้น (Two-step Scaling Grade Calibration: TSC) เป็นการปรับค่าความเอนเอียงของผลการเรียนทั้งในส่วนของโรงเรียนและในส่วนของผู้เรียน โดยถือว่าความเอนเอียงของผลการเรียนเฉลี่ยสะสมของนักเรียนนั้น เกิดขึ้นในทั้งสองส่วนเสริมกัน จึงต้องทำการปรับค่าความเอนเอียงของผลการเรียนในส่วนของโรงเรียน และในส่วนของนักเรียนที่อยู่ภายใน โรงเรียนเดียวกัน แล้วนำผลการเรียนที่ปรับความเอนเอียงแล้วทั้งสองส่วนมารวมกัน

วิธีการปรับเทียบผลการเรียนแต่ละวิธีที่ผู้วิจัยพัฒนาขึ้น จะให้ผลการปรับเทียบที่แตกต่างกัน เนื่องจากแต่ละวิธีพัฒนาขึ้นจากฐานแนวคิดของอิทธิพลของความเอนเอียงของผลการเรียนในส่วนของโรงเรียน และในส่วนของนักเรียน ที่มีผลต่อความเอนเอียงในผลการเรียนเฉลี่ยสะสมของนักเรียน การเลือกใช้วิธีการปรับเทียบผลการเรียนวิธีใดนั้น จึงขึ้นอยู่กับสถานการณ์การประเมินผลการเรียนของประเทศ ทั้งนี้เพื่อให้สามารถนำผลการเรียนของนักเรียนทุกคนมาเปรียบเทียบกันได้อย่างเหมาะสม มีความหมายและยุติธรรม โดยเฉพาะอย่างยิ่งในระบบกลางการรับเข้าศึกษา (Admissions) ดังที่ สมหวัง พิธิยานุวัฒน์ กล่าวไว้ว่า “ในการคัดสรรคนที่มีคุณภาพให้มีโอกาสเท่าเทียมกันทางสมอง ไม่ว่าจะเป็ใครก็ตามที่มีสมองและความสามารถดี ได้คะแนนสูงกว่าก็ควรได้โอกาสทางการศึกษามากกว่า” (สำนักงานคณะกรรมการการศึกษาแห่งชาติ, 2543)

2. ผลการตรวจสอบคุณภาพของวิธีการปรับเทียบผลการเรียน พบว่า

2.1 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผลการเรียนที่ปรับเทียบแล้วกับผลการสอบทางการศึกษาแห่งชาติ (O-NET) มีค่าสูงกว่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผลการเรียนเดิมกับผลการสอบทางการศึกษาแห่งชาติ (O-NET) อย่างมีนัยสำคัญทางสถิติทุกวิธี โดยวิธีการปรับเทียบผลการเรียนแบบปรับสเกลสองขั้น (TSC) ให้ค่าสัมประสิทธิ์สหสัมพันธ์สูงสุด รองลงมาเป็วิธีการปรับเทียบผลการเรียนแบบปรับสเกลขั้นเดียว (OSC) แสดงว่าวิธีการปรับเทียบผลการเรียนแบบปรับสเกลสองขั้น (TSC) เป็นวิธีที่มีคุณภาพดีที่สุด

2.2 ค่าดัชนีความแตกต่างของค่ารากที่สองของผลต่างกำลังสองเฉลี่ย (RMSD) ระหว่างผลการเรียนเกณฑ์กับผลการเรียนที่ปรับเทียบ โดยวิธีการปรับสเกลสองขั้น (TSC) มีค่าต่ำสุด ถัดมาเป็นผลการเรียนที่ปรับเทียบด้วยวิธีการปรับสเกลขั้นเดียว (OSC) แสดงว่าวิธีการปรับเทียบผลการเรียนแบบปรับสเกลสองขั้น (TSC) เป็นวิธีที่มีคุณภาพดีที่สุด

ผลการตรวจสอบคุณภาพของวิธีการเปรียบเทียบผลการเรียนดังกล่าว แสดงให้เห็นว่าวิธีการเปรียบเทียบผลการเรียนที่ผู้วิจัยพัฒนาขึ้น สามารถลดความเอนเอียงของผลการเรียนได้อย่างมีนัยสำคัญ โดยทำให้ผลการเรียนที่เปรียบเทียบแล้วมีความสัมพันธ์กับผลการสอบทางการศึกษาแห่งชาติ (O-NET) สูงขึ้น และมีดัชนีความแตกต่าง (RMSD) ลดลง ทำให้ผลการเรียนที่เปรียบเทียบแล้วสอดคล้องกับผลการสอบทางการศึกษาแห่งชาติ (O-NET) มากขึ้น ซึ่งเป็นไปตามเจตนารมณ์ของสำนักงานคณะกรรมการการอุดมศึกษา (สกอ.) ที่กล่าวไว้ว่า “ผลการเรียนที่ได้จากโรงเรียนไม่ควรแตกต่างจากคะแนนสอบ O-NET อย่างมีนัยสำคัญ” (งานประชาสัมพันธ์, สำนักงานคณะกรรมการการอุดมศึกษา, 2549 : 1; สุธา เหลือลมัย, 2549 : 7) แต่อย่างไรก็ตาม ผลการเรียนเฉลี่ยสะสมกับผลการสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) ก็ไม่ควรจะมีความสัมพันธ์กันในทางบวกอย่างสมบูรณ์ ($r = 1.00$) หรือดัชนีความแตกต่างเป็นศูนย์ ทั้งนี้เพราะการประเมินผลการเรียนนั้น นอกจากจะพิจารณาความรู้ความสามารถในเนื้อหาวิชาแล้ว ยังพิจารณาพัฒนาการของผู้เรียน ความประพฤติ พฤติกรรมการเรียน และการร่วมกิจกรรมด้วย (สำนักงานคณะกรรมการการศึกษาแห่งชาติ, 2545 : 15) ซึ่งเป็นการประเมินทั้งในด้านพุทธิพิสัย จิตพิสัยและทักษะพิสัย แต่ผลสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) เป็นเพียงการวัดพฤติกรรมด้านพุทธิพิสัยที่นับเป็นภารกิจหลักสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ที่ต้องทำการทดสอบในวิชาภาษาไทย สังคมศึกษา ภาษาอังกฤษ คณิตศาสตร์ และวิทยาศาสตร์ เพื่อนำผลการสอบไปใช้ประเมินมาตรฐานการจัดการศึกษาในระดับการศึกษาขั้นพื้นฐาน (สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) : 2550) ดังนั้นความสัมพันธ์ระหว่างผลการเรียนเฉลี่ยสะสมกับผลการสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) จึงควรมีความสัมพันธ์กันสูงแต่ไม่สัมพันธ์กันอย่างสมบูรณ์ ผลการเปรียบเทียบผลการเรียนด้วยวิธีการเปรียบเทียบที่ผู้วิจัยพัฒนาขึ้น ซึ่งให้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผลการเรียนเฉลี่ยสะสมกับผลการสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) เพิ่มขึ้นจากเดิม 0.631 เป็น 0.836 และ 0.984 จึงถือได้ว่า สามารถเปรียบเทียบผลการเรียนให้สอดคล้องกับแนวคิดดังกล่าว

3. ผลเปรียบเทียบค่าเฉลี่ยของผลต่างสัมบูรณ์ (MAD) ของผลการเรียนที่เปรียบเทียบแล้วของแต่ละวิธีกับผลการเรียนเกณฑ์ พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 โดยวิธีการปรับสเกลสองชั้น (TSC) มีดัชนีค่าเฉลี่ยผลต่างสัมบูรณ์ (MAD) ต่ำกว่าวิธีการเชื่อมโยงเสริมกันสองส่วน (TALC) วิธีการปรับสเกลชั้นเดียว (OSC) และวิธีการเชื่อมโยงส่วนเดียว (OLC) ส่วนวิธีการปรับสเกลชั้นเดียว (OSC) และวิธีการเชื่อมโยงเสริมกันสองส่วน (TALC) มีดัชนีค่าเฉลี่ยผลต่างสัมบูรณ์ (MAD) ต่ำกว่า วิธีการเชื่อมโยงส่วนเดียว (OLC) อย่างมีนัยสำคัญทางสถิติที่ระดับ .001 ในขณะที่วิธีการปรับสเกลชั้นเดียว (OSC) กับวิธีการเชื่อมโยงเสริมกันสองส่วน (TALC) และวิธีการ

ปรับสเกลสองชั้น (TSC) กับวิธีการเชื่อมโยงร่วมกันสองส่วน (TILC) มีดัชนีค่าเฉลี่ยผลต่างสัมบูรณ์ (MAD) ไม่แตกต่างกัน แสดงว่าวิธีการปรับเทียบผลการเรียนแบบปรับสเกลสองชั้น (TSC) และวิธีการเชื่อมโยงร่วมกันสองส่วน (TILC) เป็นวิธีที่มีคุณภาพดีที่สุด

เมื่อพิจารณาค่าดัชนีความแตกต่างของค่าเฉลี่ยของผลต่างสัมบูรณ์ (MAD) ระหว่างผลการเรียนที่ปรับเทียบแล้วกับผลการเรียนเกณฑ์ พบว่า วิธีการปรับเทียบผลการเรียนแบบปรับสเกลสองชั้น (TSC) ให้ค่าดัชนีความแตกต่างต่ำที่สุด ถัดมาเป็นวิธีการปรับเทียบผลการเรียนแบบเชื่อมโยงร่วมกันสองส่วน (TILC) วิธีการปรับเทียบผลการเรียนแบบเชื่อมโยงเสริมกันสองส่วน (TALC) วิธีการปรับเทียบผลการเรียนแบบปรับสเกลชั้นเดียว (OSC) และวิธีการปรับเทียบผลการเรียนแบบเชื่อมโยงส่วนเดียว (OLC) ตามลำดับ ทั้งนี้เป็นไปตามแนวคิดในการปรับเทียบแต่ละวิธี เนื่องจากวิธีการปรับเทียบผลการเรียนแบบเชื่อมโยงส่วนเดียว (OLC) และวิธีการปรับเทียบผลการเรียนแบบปรับสเกลชั้นเดียว (OSC) เป็นวิธีที่ปรับลดความเอนเอียงในส่วนของโรงเรียนเพียงอย่างเดียวเช่นกัน จึงไม่สามารถปรับลดความเอนเอียงของผลการเรียนได้ทั้งหมด ทำให้ความแตกต่างระหว่างผลการเรียนที่ปรับเทียบแล้วกับผลการเรียนเกณฑ์ยังคงสูงกว่าวิธีการปรับเทียบอื่นๆ แต่เมื่อเปรียบเทียบกับหลักการเชื่อมโยงคะแนน (Score Linking) กับการปรับสเกล (Scaling) พบว่า หลักการปรับสเกลนั้นสามารถลดความเอนเอียงได้ดีกว่าหลักการเชื่อมโยงคะแนน ซึ่งเห็นได้จากวิธีที่ปรับลดความเอนเอียงในส่วนของโรงเรียนเพียงอย่างเดียวนั้น วิธีการปรับเทียบผลการเรียนแบบปรับสเกลชั้นเดียว (OSC) ให้ค่าดัชนีความแตกต่างต่ำกว่าวิธีการปรับเทียบผลการเรียนแบบเชื่อมโยงส่วนเดียว (OLC) เช่นเดียวกันกับวิธีการปรับเทียบผลการเรียนแบบสองชั้น (ทั้งในส่วนของโรงเรียนและในส่วนของนักเรียน) วิธีการปรับเทียบผลการเรียนแบบปรับสเกลสองชั้น (TSC) ให้ค่าดัชนีความแตกต่างต่ำกว่าวิธีการปรับเทียบผลการเรียนแบบเชื่อมโยงร่วมกันสองส่วน (TILC) และวิธีการปรับเทียบผลการเรียนแบบเชื่อมโยงเสริมกันสองส่วน (TALC) ที่เป็นเช่นนี้ก็เพราะว่า การปรับเทียบผลการเรียนโดยอาศัยหลักการเชื่อมโยงคะแนน (Score Linking) นั้นใช้หลักการวิเคราะห์การถดถอยเชิงเส้นด้วยวิธีกำลังสองน้อยที่สุด (Ordinary Least Square : OLS) ซึ่งเป็นการสร้างสมการความสัมพันธ์เชิงเส้นตรงระหว่างผลการเรียนที่นำมาปรับเทียบ ถึงแม้ว่าวิธีการกำลังสองน้อยที่สุด (OLS) จะให้สมการเส้นตรงที่ดีที่สุดสำหรับข้อมูลชุดหนึ่งๆ แต่หากผลการเรียนที่นำมาปรับเทียบมีค่าการกระจายมาก ก็จะทำให้มีส่วนของความคลาดเคลื่อน (Residual) มากด้วยเช่นกัน ซึ่งต่างจากหลักการปรับเทียบสเกล (Scaling) ที่ใช้สัดส่วนการเปลี่ยนแปลงระหว่างผลการเรียนที่นำมาปรับเทียบ ความสัมพันธ์จึงมีลักษณะเป็นเส้นตรงหรือไม่เป็นเส้นตรง (Nonlinear) ก็ได้ ทำให้มีความสอดคล้องกับลักษณะการกระจายของผลการเรียนได้ดีกว่าสมการความสัมพันธ์เชิงเส้นตรง ดังนั้นวิธีการปรับเทียบผลการเรียนที่อาศัยหลักการปรับสเกลจึงให้ผลการปรับเทียบที่ดีกว่าวิธีการเชื่อมโยงเชิงเส้น

ส่วนในระหว่างการเปรียบเทียบผลการเรียนสองส่วนหรือการเปรียบเทียบสองชั้น ให้ผลการเปรียบเทียบดีกว่าการเปรียบเทียบผลการเรียนเพียงส่วนเดียว หรือการเปรียบเทียบชั้นเดียว เช่น วิธีการเปรียบเทียบผลการเรียนแบบปรับสเกลสองชั้น (TSC) ให้ค่าดัชนีความแตกต่างต่ำกว่าวิธีการเปรียบเทียบผลการเรียนแบบปรับสเกลชั้นเดียว (OSC) วิธีการเปรียบเทียบผลการเรียนแบบเชื่อมโยงร่วมกันสองส่วน (TILC) และวิธีการเปรียบเทียบผลการเรียนแบบเชื่อมโยงเสริมกันสองส่วน (TALC) ให้ค่าดัชนีความแตกต่างต่ำกว่าวิธีการเปรียบเทียบผลการเรียนแบบเชื่อมโยงส่วนเดียว (OLC) เป็นเพราะว่าวิธีการเปรียบเทียบผลการเรียนสองส่วนหรือสองชั้น สามารถปรับลดความเอนเอียงทั้งในส่วน of โรงเรียน และในส่วนของผู้เรียน ในขณะที่วิธีการเปรียบเทียบผลการเรียนเพียงส่วนเดียวหรือชั้นเดียว จะสามารถปรับลดความเอนเอียงได้เฉพาะในส่วน of โรงเรียนเท่านั้น

จากผลการเปรียบเทียบดัชนีความแตกต่างจะเห็นว่า วิธีการเปรียบเทียบผลการเรียนแบบปรับสเกลสองชั้น (TSC) และวิธีการเชื่อมโยงร่วมกันสองส่วน (TILC) เป็นวิธีที่มีคุณภาพดีที่สุด สอดคล้องกับแนวคิดของ แฮร์ริสและครูส์ (Harris and Crouse, 1993) ได้เสนอแนะว่า ในการเปรียบเทียบคะแนนต้องการให้มีค่าดัชนีความแตกต่างที่น้อย จะแสดงว่าการเปรียบเทียบนั้นสามารถเปรียบเทียบได้ดีกว่าวิธีการเปรียบเทียบที่ให้ค่าดัชนีความแตกต่างที่มากกว่า เพราะสามารถลดความเอนเอียงของผลการเรียนได้มากที่สุด โดยสามารถลดความเอนเอียงที่เกิดขึ้นในส่วน of โรงเรียนที่สามารถทำให้ทุกโรงเรียนมีผลการเรียนที่ได้มาตรฐานเดียวกัน อีกทั้งยังลดความเอนเอียงที่เกิดขึ้นในส่วน of นักเรียนที่ทำให้นักเรียนทุกคนมีผลการเรียนที่อยู่บนมาตรฐานเดียวกัน วิธีการเปรียบเทียบผลการเรียนแบบปรับสเกลสองชั้น (TSC) และวิธีการเชื่อมโยงร่วมกันสองส่วน (TILC) จึงนับเป็นวิธีการเปรียบเทียบที่มีคุณภาพดีที่สุดทำให้ได้ผลการเรียนที่สามารถลดความเอนเอียงได้มากที่สุด ที่ก่อให้เกิดความเป็นธรรมกับนักเรียนทุกคนในการที่จะนำผลการเรียนไปใช้ในการตัดสินใจในการคัดเลือกเข้าศึกษาต่อในระบบการรับเข้าศึกษา (Admissions) ซึ่งสอดคล้องกับแนวคิดของสมาคมเครือข่ายผู้ปกครองแห่งชาติ และรองศาสตราจารย์ ดร.เรณู พงษ์เรืองพันธุ์ รองอธิการบดีฝ่ายวิชาการ มหาวิทยาลัย ที่ได้เสนอแนะว่า แต่ละโรงเรียนมีมาตรฐานในการให้ผลการเรียนไม่เหมือนกัน ควรนำคะแนน O-NET ไปปรับผลการเรียนของนักเรียนให้สูงขึ้นสำหรับโรงเรียนที่เกรด และไปปรับผลการเรียนให้ต่ำลงสำหรับโรงเรียนที่ปล่อยเกรด และควรมีการปรับให้เท่าเทียมสำหรับนักเรียนทุกคนก่อนนำผลการเรียนไปใช้เป็นองค์ประกอบในการรับเข้าศึกษา ทั้งนี้ที่จะช่วยให้ผลการเรียนของนักเรียนทุกคนอยู่บนมาตรฐานเดียวกัน และเป็นที่ยอมรับได้ว่านักเรียนที่สามารถเข้าศึกษาต่อในระบบการรับเข้าศึกษาในสถาบันอุดมศึกษาในระบบกลาง (Admissions) เป็นผู้ที่มีความรู้ความสามารถเหมาะสมในคณะและสาขาที่เข้าศึกษาอย่างแท้จริง

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

วิธีการเปรียบเทียบผลการเรียนแต่ละวิธีมีหลักการและแนวคิดที่แตกต่างกัน และให้ผลการเปรียบเทียบที่แตกต่างกัน การพิจารณานำวิธีการเปรียบเทียบผลการเรียนไปใช้ ควรศึกษาแนวคิดทฤษฎีของวิธีการเปรียบเทียบนั้นๆ ให้ชัดเจน และสอดคล้องกับวัตถุประสงค์ในการนำไปใช้ หน่วยงานที่ควรนำไปใช้ได้แก่ สำนักงานคณะกรรมการการอุดมศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือหน่วยงานอื่นๆ ที่เกี่ยวข้อง โดยมีเงื่อนไขดังนี้

1. ถ้าต้องการเปรียบเทียบผลการเรียนเฉพาะในส่วนของ โรงเรียนให้ทุกโรงเรียนมีมาตรฐานเดียวกันควรใช้วิธีการเปรียบเทียบผลการเรียนแบบปรับสเกลชั้นเดียว (OSC)
2. ถ้าต้องการเปรียบเทียบผลการเรียนในส่วนของ โรงเรียนให้ทุกโรงเรียนมีมาตรฐานเดียวกันและเปรียบเทียบผลการเรียนในส่วนของนักเรียนให้มีมาตรฐานเดียวกัน ควรใช้วิธีการเปรียบเทียบผลการเรียนแบบปรับสเกลสองชั้น (TSC)
3. ควรสอบถามความคิดเห็นของบุคคลที่มีส่วนได้ส่วนเสีย รวมถึงนักวิชาการได้ร่วมกันพิจารณาว่า ควรนำวิธีการเปรียบเทียบผลการเรียนมาใช้ระบบการศึกษาไทยหรือไม่ อีกทั้งควรทำความเข้าใจและพิจารณาเลือกใช้วิธีการเปรียบเทียบผลการเรียนตามข้อเสนอแนะของผู้วิจัย
4. หากใช้แบบสอบร่วมอื่น วิธีการเปรียบเทียบผลการเรียนที่ผู้วิจัยพัฒนาขึ้นก็ยังสามารถนำไปใช้ในการเปรียบเทียบได้เช่นกัน

ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

1. ควรทำการศึกษาเปรียบเทียบเกี่ยวกับผลต่อการเข้าศึกษาต่อในสถาบันอุดมศึกษาของนักเรียนที่สามารถเข้าศึกษาต่อได้ และไม่ได้ เมื่อผลการเรียนที่นำมาใช้เป็นองค์ประกอบในระบบการรับเข้าศึกษา (Admissions) เป็นผลการเรียนที่เปรียบเทียบแล้วกับผลการเรียนที่ไม่ผ่านการเปรียบเทียบ
2. ควรทำการศึกษาโดยใช้ข้อมูลนักศึกษาที่รับเข้าศึกษาในระบบการรับเข้าศึกษา (Admissions) ที่ใช้ผลการเรียนที่เปรียบเทียบแล้ว กับนักศึกษาที่ใช้ผลการเรียนที่ไม่ผ่านการเปรียบเทียบ เพื่อหาความตรงเชิงพยากรณ์กับผลการเรียนในมหาวิทยาลัยอย่างต่อเนื่องตั้งแต่ ปี 1 - 4
3. ควรทำการศึกษาความคิดเห็นของผู้มีส่วนได้ส่วนเสียเกี่ยวกับการนำวิธีการเปรียบเทียบผลการเรียนมาใช้ระบบการรับเข้าศึกษา (Admissions)