

บทที่ 1

บทนำ

ภูมิหลัง

หลังจากที่ประชุมอธิการบดีแห่งประเทศไทย (ทปอ.) ได้เสนอการปรับปรุงการคัดเลือกผู้เข้าศึกษาต่อในมหาวิทยาลัยในระบบสอบรวมต่อทบวงมหาวิทยาลัยตามหนังสือที่ ทปอ. 44/147 ลงวันที่ 19 เมษายน 2544 โดยยึดหลักการให้เพิ่มการใช้ผลการเรียนระดับมัธยมศึกษาตอนปลายเป็นส่วนหนึ่งในระบบการรับเข้าศึกษา (Admission) ให้มากขึ้น และเริ่มใช้ระบบการรับเข้าศึกษานี้ตั้งแต่ปีการศึกษา 2547 เป็นต้นมา (สำนักทดสอบกลาง, สำนักงานคณะกรรมการการอุดมศึกษา. 2548 : 1-5) ผลการเรียนเฉลี่ยสะสมตลอดหลักสูตรมัธยมศึกษาตอนปลายตามกลุ่มสาระการเรียนรู้ (GPA) และผลการเรียนเฉลี่ยสะสมตลอดหลักสูตรมัธยมศึกษาตอนปลาย (GPAX) ของนักเรียนก็มีความสำคัญเพิ่มขึ้นอีกมาก จากเดิมที่ใช้เป็นเพียงเกณฑ์ในการจบหลักสูตร กลายเป็นส่วนหนึ่งในการพิจารณารับเข้าศึกษาต่อในระดับอุดมศึกษา ซึ่งถือได้ว่าเป็นเป้าหมายหลักของนักเรียนในระดับมัธยมศึกษาตอนปลายที่ต้องแข่งขันกันสุดความสามารถ ทำให้มีการกล่าวถึงมาตรฐานของผลการเรียนจากโรงเรียนต่างๆ ที่นักการศึกษาส่วนใหญ่เห็นว่ายังไม่เท่าเทียมกัน การนำเอาผลการเรียนเฉลี่ยและผลการเรียนเฉลี่ยสะสมมาเป็นส่วนหนึ่งของระบบการรับเข้าศึกษาจึงขาดความเป็นธรรม ถึงแม้ในปีการศึกษา 2553 จะมีการปรับเปลี่ยนระบบการรับเข้าศึกษาให้ใช้เฉพาะผลการเรียนเฉลี่ยสะสมตลอดหลักสูตรมัธยมศึกษาตอนปลาย (GPAX) ประกอบกับคะแนนการทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (Ordinary National Educational Test :O-NET) และคะแนนความถนัด (Aptitude Test) ซึ่งใกล้เคียงกับระบบการคัดเลือกเข้าศึกษา (Entrance) เดิมมากขึ้น แต่การนำเอาผลการเรียนเฉลี่ยสะสมตลอดหลักสูตรมัธยมศึกษาตอนปลาย (GPAX) ที่มีมาตรฐานไม่เท่าเทียมกันมาเป็นส่วนหนึ่งของการรับเข้าศึกษาย่อมทำให้เกิดการได้เปรียบและเสียเปรียบกันเกิดขึ้น

การนำผลการเรียนเฉลี่ยสะสมตลอดหลักสูตรมาเป็นองค์ประกอบของระบบการรับเข้าศึกษา (Admissions) ถึงแม้จะเป็นคะแนนที่สะท้อนผลการเรียนจริงของนักเรียน เป็นหลักประกันว่าผู้สมัครเข้าศึกษาได้ผ่านกระบวนการเรียนในระดับมัธยมศึกษาตอนปลายมาโดยครบถ้วน และเป็นกระบวนการที่พัฒนาบุคคลให้เป็นผู้ที่สมบูรณ์ตามแผนการศึกษาก็ตาม แต่ก็มีข้อกังขาที่ว่าค่า GPAX จากโรงเรียนต่างๆ ไม่อยู่ในมาตรฐานเดียวกัน (สำนักงานคณะกรรมการการศึกษาแห่งชาติ. 2543 : 39 -51) ดังผลการศึกษาของศิริชัย กาญจนวาสี และคณะ (Kanjanavasee, and others. 2009 :

117-124) ที่ศึกษาผลการเรียนเฉลี่ยสะสมตามกลุ่มสาระการเรียนรู้ (GPA) ของโรงเรียนมัธยมศึกษา ตอนปลาย จำนวน 2,583 โรงเรียน ในปีการศึกษา 2549 พบว่า มีโรงเรียนที่ปล่อยเกรดจำนวน 1,225 โรงเรียน คิดเป็นร้อยละ 47.43 ส่วนมากเป็นโรงเรียนขนาดเล็กและขนาดกลาง และโรงเรียนที่กวด เกรด จำนวน 1,235 โรงเรียน คิดเป็นร้อยละ 47.81 เป็นกลุ่มโรงเรียนขนาดใหญ่ ส่วนโรงเรียนที่ให้ เกรดเป็นปกติมีเพียง 123 โรงเรียน คิดเป็นร้อยละ 4.76 เท่านั้น ซึ่งความไม่เท่าเทียมกันของผลการ เรียนที่มาจากโรงเรียนต่างๆ อาจเนื่องมาจากปัจจัยหลายประการ เช่น หลักสูตรสถานศึกษา การจัด กิจกรรมการเรียนการสอน การวัดและประเมินผลการศึกษา บริบทของสถานศึกษา ขนาดของ สถานศึกษา เป็นต้น ซึ่งปัจจัยเหล่านี้ล้วนเป็นปัจจัยที่ส่งผลให้ค่า GPA และ GPAX ของนักเรียนใน แต่ละโรงเรียนให้มีมาตรฐานแตกต่างกันไป ส่งผลให้ระบบการรับเข้าขาดความเป็นธรรม โดย โรงเรียนที่มีมาตรฐานต่ำก็จะพยายามปล่อยเกรดหรือใช้แบบสอบที่ง่ายในการประเมินผล เพื่อให้ นักเรียนของตนมีค่า GPAX สูงเกินความสามารถที่ควรจะเป็นของผู้เรียน เพื่อให้ นักเรียนของตน ได้เปรียบ มีโอกาสในการเข้าศึกษาต่อสูงขึ้น ส่วนโรงเรียนที่มีมาตรฐานสูงซึ่งมักจะมีผู้เรียนที่มี ความสามารถสูงอยู่มากมาย จะได้รับอิทธิพลเพดาน (Ceiling Effect) ในการประเมินผลการเรียน หรือใช้แบบสอบที่ยากในการประเมินผล เพื่อให้สามารถจำแนกความสามารถของผู้เรียน ได้ จึงเกิด ปรากฏการณ์กวดเกรดผู้เรียนเมื่อเทียบกับโรงเรียนอื่นๆ ส่งผลให้ค่า GPAX ของผู้เรียนต่ำกว่า ความสามารถที่ควรจะเป็น และทำให้โอกาสในการเข้าศึกษาต่อก็จะลดลง ส่งผลให้นักเรียนที่มี GPAX สูงแต่ความสามารถต่ำ มีโอกาสเข้าศึกษาต่อในระดับอุดมศึกษามากขึ้น ปัญหาที่จะตามมาก็คือ ได้ นักศึกษาที่มีคุณภาพด้อยลง ซึ่งจากการศึกษาของอรุณี อ่อนสวัสดิ์ และคณะ (2551 : 319) พบว่า นักศึกษาที่รับเข้าในระบบ Admissions เรียนไม่ผ่านในวิชาสำคัญ สูงกว่าระบบ Entrance ถึง 4 เท่า หรือนักศึกษามีผลการเรียนอ่อนลง ทำให้เกิดปัญหาการออกกลางคันอันเนื่องมาจากไม่สามารถเรียน ในสาขาและในสถาบันที่ตนเองเข้าศึกษาได้ อันเป็นการสูญเสียที่สำคัญในระบบการศึกษา

หากพิจารณาตามหลักของการวัดและประเมินผลทางการศึกษา ทุกโรงเรียนในประเทศไทย ยังคงใช้การวัดและประเมินผลตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) อยู่ ซึ่งนัก วัดผลการศึกษาต่างยอมรับกันว่า การวัดและประเมินผลการเรียนตามทฤษฎีการทดสอบแบบดั้งเดิม นั้น ขึ้นอยู่กับกลุ่มประชากรหรือกลุ่มผู้สอบ (Population) นั่นก็หมายความว่า ค่าสถิติต่างๆ ของ ข้อสอบและแบบสอบ เช่น ความยากของข้อสอบ ความเชื่อมั่นของแบบสอบจะแปรเปลี่ยนไปตาม กลุ่มผู้สอบ (Flanagan. 1951; Angoff. 1971) ยกตัวอย่างเช่น ข้อสอบข้อหนึ่งมีความยากที่เหมาะสม กับ โรงเรียนหนึ่ง แต่อาจจะยากเกินไปสำหรับ โรงเรียนหนึ่งหรือง่ายเกินไปสำหรับอีก โรงเรียนหนึ่ง ดังนั้นจึงเป็นการยากที่มาตรฐานการวัดและประเมินผลของแต่ละ โรงเรียนซึ่งมีกลุ่มผู้สอบต่างกัน ใช้ หลักเกณฑ์และแบบสอบที่แตกต่างกันจะมีมาตรฐานเดียวกัน การนำผลการเรียนเฉลี่ยสะสมตลอด

หลักสูตร (GPAX) มาเป็นองค์ประกอบของระบบการรับเข้าศึกษา โดยไม่ได้ผ่านการปรับเทียบผลการเรียนระหว่างโรงเรียนให้มีมาตรฐานเดียวกันก่อน จึงเป็นการไม่ยุติธรรมกับผู้สมัครรับเข้าศึกษาในระดับอุดมศึกษา

นอกจากมาตรฐานการประเมินผลของโรงเรียนจะแตกต่างกันอันเนื่องมาจากกลุ่มผู้สอบแล้ว ระบบการตัดสินผลการเรียนแบบกำหนดระดับผลสัมฤทธิ์เป็นตัวอักษร (Letter grade system) ก็เป็นอีกสาเหตุหนึ่งที่ทำให้ผลการเรียนของนักเรียนมีความลำเอียง (bias) ไปจากความสามารถของผู้เรียน เนื่องจากระบบการตัดสินผลการเรียนแบบกำหนดระดับผลสัมฤทธิ์เป็นตัวอักษรนั้น จะแบ่งคะแนนความสามารถออกเป็นช่วงๆ โดยในแต่ละช่วงคะแนนจะกำหนดผลสัมฤทธิ์ระดับเดียวกัน เช่น คะแนนเต็ม 100 คะแนน ผู้เรียนที่ได้คะแนนระหว่าง 75-79 คะแนน กำหนดให้ได้ผลการเรียน 3.5 ดังนั้นนักเรียนที่ได้คะแนน 79 คะแนนย่อมเทียบเปรียบนักเรียนที่ได้ 75 คะแนน เนื่องจากนักเรียนทั้งสองคนนี้จำเป็นต้องได้ผลการเรียน 3.5 เหมือนกัน ยิ่งถ้าหากเกิดเหตุการณ์เช่นนี้ขึ้นในทุกๆ ภาคเรียน ผลการเรียนเฉลี่ยในกลุ่มสาระการเรียนรู้ของนักเรียนทั้งสองคนย่อมเท่ากัน ทั้งที่นักเรียนคนที่ได้ 79 คะแนนมีความรู้ ความสามารถสูงกว่านักเรียนที่ได้ 75 คะแนนทุกภาคเรียน

อย่างไรก็ตาม ผลการเรียนที่ลำเอียงไปเนื่องจากระบบการตัดสินผลการเรียนแบบกำหนดระดับผลสัมฤทธิ์เป็นตัวอักษรนี้ จะเกิดขึ้นเฉพาะในกลุ่มนักเรียนที่ได้รับการตัดสินผลการเรียนร่วมกันหรือภายในแต่ละโรงเรียนเดียวกันเท่านั้น ความลำเอียงของผลการเรียนแบบนี้จึงเป็นความลำเอียงระหว่างผู้เรียนภายในโรงเรียนเดียวกัน ซึ่งต่างจากความลำเอียงของผลการเรียนอันเนื่องมาจากมาตรฐานการประเมินผลของโรงเรียน ซึ่งเกิดขึ้นระหว่างโรงเรียน ดังนั้นจึงอาจกล่าวได้ว่า ความลำเอียงของผลการเรียนของนักเรียนคนหนึ่งๆ นั้น เป็นผลมาจากความลำเอียงใน 2 ส่วน คือ ในส่วนของโรงเรียน (ระหว่างโรงเรียน) และในส่วนของนักเรียน (ระหว่างนักเรียนในโรงเรียนเดียวกัน) ดังนั้นหากจะทำการปรับเทียบผลการเรียนระหว่างนักเรียนทั้งหมด (จากทุกโรงเรียน) ให้มีมาตรฐานเดียวกัน และสามารถนำไปเปรียบเทียบกันได้ จึงควรทำการปรับเทียบทั้งในส่วนของโรงเรียนและในส่วนของนักเรียน

การนำคะแนนมาเปรียบเทียบกัน ส่วนใหญ่อาศัยหลักการเชื่อมโยงคะแนน (score linking) ซึ่งมีอยู่ 4 แบบ คือ 1) การปรับเทียบคะแนน (equating) ซึ่งเป็นการเชื่อมโยงคะแนนระหว่างแบบสอบที่มีคุณลักษณะเหมือนกันคือ วัดคุณลักษณะเดียวกัน มีความยากและความเชื่อมั่นเท่ากัน 2) การปรับเทียบสเกล (calibration) เป็นการเชื่อมโยงคะแนนระหว่างแบบสอบที่วัดคุณลักษณะเดียวกัน แต่มีความยากและความเชื่อมั่นต่างกัน 3) การหาความสอดคล้อง (concordance) เป็นการเชื่อมโยงคะแนนระหว่างแบบสอบที่วัดคุณลักษณะคล้ายคลึงกัน และมีการแจกแจงคะแนนคล้ายกัน และ 4) การฉายภาพหรือการพยากรณ์ (projection or prediction) เป็นการเชื่อมโยงคะแนน

ระหว่างแบบสอบที่วัดในคุณลักษณะที่ต่างกัน และมีความยากและความเชื่อมั่นต่างกัน (Kolen. 2004) โดยการเชื่อมโยงแต่ละแบบก็จะมีเทคนิควิธีการปรับเทียบย่อยๆ อีกหลายวิธี แต่เทคนิควิธีเหล่านี้เป็นการเชื่อมโยงคะแนนระหว่างแบบสอบ จึงพิจารณาคุณลักษณะของแบบสอบเป็นสำคัญ และคะแนนที่นำมาเชื่อมโยงกันก็เป็นผลมาจากการวัดคุณลักษณะของผู้เรียนที่ยังไม่ได้ผ่านการประเมิน ซึ่งแตกต่างจากผลการเรียน หากจะนำวิธีการเชื่อมโยงคะแนนเหล่านี้มาใช้ในการปรับเทียบผลการเรียน ซึ่งผ่านการประเมินโดยการกำหนดเกรด (grading) จึงยังไม่มีความเหมาะสม โดยเฉพาะอย่างยิ่งในกรณีของผลการเรียนเฉลี่ยสะสมตลอดหลักสูตร (GPAX) ที่เป็นองค์ประกอบของระบบกลางรับเข้าศึกษา

ในการปรับเทียบผลการเรียนเฉลี่ยสะสมตลอดหลักสูตร ตามระบบกลางการรับบุคคลเข้าศึกษาในสถาบันอุดมศึกษา (Admissions) นี้ ปิยะธิดา ปัญญา (2551) ได้พัฒนาวิธีการปรับเทียบขึ้นมา 3 วิธี โดยอาศัยการเชื่อมโยงเชิงเส้น ด้วยเทคนิควิเคราะห์การถดถอย ถึงแม้วิธีการดังกล่าวจะสามารถลดความเอนเอียงของผลการเรียนได้ดี แต่ก็เป็นวิธีการที่มีความยุ่งยาก ซับซ้อน ไม่เหมาะที่จะนำไปใช้ในสถานการณ์จริง ที่ต้องทำให้ผู้มีส่วนได้ส่วนเสียในระบบกลางการรับบุคคลเข้าศึกษาเข้าใจและยอมรับได้

จากสถานการณ์ที่เป็นปัญหาของระบบกลางรับเข้าศึกษาและปัญหาการเชื่อมโยงคะแนน ผู้วิจัยจึงสนใจที่จะพัฒนาวิธีการปรับเทียบผลการเรียน โดยอาศัยหลักการปรับสเกลสองชั้น เพื่อทำให้ผลการเรียนของนักเรียนทุกคนจากทุกโรงเรียนมีมาตรฐานเดียวกัน สามารถนำมาเปรียบเทียบกันได้ตามกรอบของระบบกลางการรับบุคคลเข้าศึกษาในสถาบันอุดมศึกษา (Admission) ซึ่งหากมีการนำไปใช้ในระบบดังกล่าว จะก่อให้เกิดความเป็นธรรมและเป็นที่ยอมรับของผู้ที่เกี่ยวข้องทุกฝ่าย รวมทั้งยังสามารถนำไปใช้ในการประเมินคุณภาพโรงเรียนได้อีกด้วย

คำถามการวิจัย

1. วิธีการปรับเทียบผลการเรียนเฉลี่ยสะสมตลอดหลักสูตร (GPAX) ระหว่างโรงเรียนให้มีมาตรฐานเดียวกันควรมีลักษณะอย่างไร
2. วิธีการปรับเทียบผลการเรียนเฉลี่ยสะสมตลอดหลักสูตรระหว่างโรงเรียนที่พัฒนาขึ้นสามารถปรับผลการเรียนของนักเรียนให้มีมาตรฐานเดียวกันได้หรือไม่ อย่างไร
3. เมื่อเปรียบเทียบคุณภาพของวิธีการปรับเทียบผลการเรียนที่พัฒนาขึ้นกับวิธีที่ใช้การเชื่อมโยงเชิงเส้น ผลจะเป็นอย่างไร

วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้ มีวัตถุประสงค์ของการวิจัยดังต่อไปนี้

1. เพื่อพัฒนาวิธีการปรับเทียบผลการเรียน โดยอาศัยหลักการปรับสเกลชั้นเดียวและสองชั้น
2. เพื่อตรวจสอบคุณภาพของวิธีการปรับเทียบผลการเรียนที่พัฒนาขึ้น
3. เพื่อเปรียบเทียบคุณภาพของวิธีการปรับเทียบผลการเรียนที่พัฒนาขึ้นกับวิธีการปรับเทียบผลการเรียนที่อาศัยหลักการเชื่อมโยงเชิงเส้น

ความสำคัญของการวิจัย

ผลของการวิจัยครั้งนี้จะทำให้ได้วิธีการปรับเทียบผลการเรียนเฉลี่ยสะสมตลอดหลักสูตร สำหรับเป็นแนวทางในการปรับเทียบผลการเรียนเฉลี่ยสะสมตลอดหลักสูตรระหว่างโรงเรียนให้มีมาตรฐานเดียวกัน โดยมุ่งให้เกิดความเป็นธรรมกับผู้สมัครเข้าศึกษาในระดับอุดมศึกษาและเป็นที่ยอมรับของบุคคลที่เกี่ยวข้อง

ขอบเขตของการวิจัย

การวิจัยครั้งนี้ มีขอบเขตดังต่อไปนี้

1. ขอบเขตด้านเนื้อหา

การวิจัยครั้งนี้เป็นการศึกษาและพัฒนาวิธีการปรับเทียบผลการเรียน โดยอาศัยหลักการปรับสเกลชั้นเดียวและสองชั้น เพื่อใช้ในการปรับเทียบผลการเรียนให้มีมาตรฐานเดียวกัน และสามารถเปรียบเทียบกันได้ โดยอิงกับระบบกลางการรับบุคคลเข้าศึกษาในสถาบันอุดมศึกษา (Admission) ดังนั้นคะแนนที่นำมาใช้ในการปรับเทียบจึงเป็นผลการเรียนเฉลี่ยสะสมตลอดหลักสูตร (GPAX) และผลการสอบทางการศึกษาแห่งชาติชั้นพื้นฐาน (Ordinary National Educational Test : O-NET) ซึ่งเป็นองค์ประกอบของระบบกลางการรับบุคคลเข้าศึกษาในสถาบันอุดมศึกษา

2. ขอบเขตด้านข้อมูล

ข้อมูลที่ใช้ในการวิจัยครั้งนี้เป็นข้อมูลที่จำลองขึ้น (Simulation) ตามผลการสอบในระบบกลางการรับบุคคลเข้าศึกษาในสถาบันอุดมศึกษา (Admission) ประจำปีการศึกษา

3. ขอบเขตด้านตัวแปรที่ศึกษามีดังต่อไปนี้

ตัวแปรที่ใช้ในการตรวจสอบคุณภาพของวิธีการปรับเทียบผลการเรียนที่พัฒนาขึ้น ประกอบด้วย

3.1 ตัวแปรอิสระ ได้แก่ วิธีการปรับเทียบผลการเรียน จำแนกเป็น

- 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 ซึ่งพิจารณาจากความคลาดเคลื่อน (Error) ของการปรับเทียบ โดยการหาดัชนีความแตกต่าง RMSD (Root Mean Square Difference) MAD (Mean of Absolute Difference) และค่าสัมประสิทธิ์สหสัมพันธ์

ข้อตกลงเบื้องต้น

การดำเนินการวิจัยครั้งนี้อยู่บนข้อตกลงเบื้องต้นที่ว่า ผลการสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (Ordinary National Educational Test : O-NET) ที่ดำเนินการโดยสถาบันทดสอบทางการศึกษาแห่งชาติ สามารถประมาณค่าความสามารถของผู้เรียนได้ดีที่สุด

นิยามศัพท์เฉพาะ

การปรับเทียบผลการเรียน (Grade Calibration) หมายถึง การปรับเทียบมาตรฐานวัด (Calibration) ระหว่างผลการเรียนจากโรงเรียนต่างๆ ซึ่งประเมินผลจากแบบสอบที่มีเนื้อหาคล้ายคลึงกัน แต่มีระดับความยากและความเชื่อมั่นต่างกัน เพื่อให้ผลการเรียนที่ปรับแล้วมีมาตรฐาน (scale) เดียวกัน และสามารถเปรียบเทียบกันได้

ผลการเรียน หมายถึง ตัวเลขแสดงผลการเรียนเฉลี่ยสะสมตลอดหลักสูตร (GPAX) ตามหลักสูตรการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 4

ผลการเรียนเฉลี่ยสะสมตลอดหลักสูตร (Cumulative GPA : GPAX) หมายถึง ตัวเลขเฉลี่ยของผลการเรียนที่รวบรวมมาจากการเรียนทุกกลุ่มสาระการเรียนรู้และทุกชั้นปีตลอดหลักสูตรการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 4

ผลการสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) หมายถึง ผลการประเมินคุณภาพของช่วงชั้นที่ 4 หรือชั้นมัธยมศึกษาปีที่ 6 โดยพิจารณาจาก 5 กลุ่มสาระการเรียนรู้ คือ ภาษาไทย คณิตศาสตร์ วิทยาศาสตร์ สังคมศึกษา ศาสนาและวัฒนธรรม และภาษาอังกฤษ ที่จัดสอบโดยสถาบันทดสอบทางการศึกษาแห่งชาติ

ความลำเอียงของผลการเรียน หมายถึง ความคลาดเคลื่อนในการประเมินความสามารถของผู้เรียนที่ไม่ตรงกับความสามารถจริง อันเนื่องมาจากมาตรฐานการประเมินผลของโรงเรียนและระบบการตัดสินผลการเรียน ทำให้ผลการเรียนเฉลี่ยสะสมตลอดหลักสูตร (GPAX) ของผู้เรียนไม่สอดคล้องกับความสามารถที่ควรจะเป็น ซึ่งจำแนกออกเป็น 2 กรณีคือ

1. การป่องเกรด (Grade Inflation) หมายถึง การประเมินผลการเรียนรู้ของผู้เรียนที่สูงกว่าความสามารถที่ควรจะเป็นของผู้เรียน ทำให้ผลการเรียนเฉลี่ยสะสมตลอดหลักสูตร (GPAX) ของนักเรียนสูงกว่าผลการเรียนที่ควรจะเป็น

2. การกดเกรด (Grade Deflation) หมายถึง การประเมินผลการเรียนรู้ของผู้เรียนที่ต่ำกว่าความสามารถที่ควรจะเป็นของผู้เรียน ทำให้ผลการเรียนเฉลี่ยสะสมตลอดหลักสูตร (GPAX) ของนักเรียนต่ำกว่าผลการเรียนที่ควรจะเป็น

ผลการเรียนเกณฑ์ หมายถึง ผลการเรียนที่ควรจะเป็นของผู้เรียนที่มีความสอดคล้องกับความสามารถที่แท้จริงของผู้เรียน โดยถือเอาผลการสอบทางการศึกษาขั้นพื้นฐาน (O-NET) เป็นมาตรฐาน ดังนั้นผลการเรียนเกณฑ์ในทางทฤษฎีจึงมีความสัมพันธ์กับผลการสอบทางการศึกษาขั้นพื้นฐาน (O-NET) ในเชิงบวกอย่างสมบูรณ์

คุณภาพของวิธีการเปรียบเทียบ หมายถึง ความสามารถของวิธีการเปรียบเทียบผลการเรียน ในการลดความลำเอียงของผลการเรียน ซึ่งพิจารณาจากความคลาดเคลื่อน (Error) ของการเปรียบเทียบโดยใช้วิธีการหาดัชนีความแตกต่าง (RMSD) และค่าสัมประสิทธิ์สหสัมพันธ์

ดัชนีความแตกต่าง หมายถึง ค่ารากที่สองของความแตกต่างกำลังสองเฉลี่ย (Root Mean Square Difference : RMSD) ซึ่งคำนวณได้จากการหารากที่สองของค่าเฉลี่ยของความแตกต่างระหว่างผลการเรียนที่ปรับแล้วกับผลการเรียนเกณฑ์ ยกกำลังสอง

ระบบการรับเข้า (Admission) หมายถึง ระบบกลางการรับบุคคลเข้าศึกษาในสถาบันอุดมศึกษาที่ตัดสินผลโดยพิจารณาจากคะแนนของ 3 องค์ประกอบ คือ ผลการเรียนเฉลี่ยสะสมตลอดหลักสูตรมัธยมศึกษาตอนปลาย ผลการสอบทางศึกษาแห่งชาติขั้นพื้นฐาน และผลการทดสอบความถนัด

กรอบแนวความคิดในการวิจัย

ในการพัฒนาวิธีการเปรียบเทียบผลการเรียนในครั้งนี้ มีวัตถุประสงค์หลักเพื่อพัฒนาวิธีการเปรียบเทียบผลการเรียน (GPAX) ของนักเรียนแต่ละคน เพื่อให้ผลการเรียนที่ผ่านการเปรียบเทียบแล้วอยู่บนมาตรฐานเดียวกัน ทำให้สามารถนำมาเปรียบเทียบกันได้ กรอบแนวคิดของการวิจัยพอสรุปได้ดังนี้

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 เป็นส่วนที่ทำให้ผลการเรียนของนักเรียนลำเอียงไปจากความสามารถที่ควรจะเป็น (Cohen, 2000 : 1-4) และส่วนที่สาม เป็นบริบทของโรงเรียนหรือเป็นบริบทของการประเมินผล ซึ่งต่างโรงเรียนก็จะมีบริบทต่างกันไป โดยส่วนประกอบนี้จะประกอบไปด้วย นโยบายของโรงเรียน คุณลักษณะของครูผู้สอน คุณลักษณะของนักเรียน เป็นต้น ซึ่งปัจจัยเหล่านี้ส่งผลให้มาตรฐานของผลการเรียนของแต่ละโรงเรียนแตกต่างกันไป เนื่องจากปัจจัยเหล่านี้มีผลต่อนักเรียนทั้งหมดในแต่ละโรงเรียนซึ่งถือว่าเป็นประชากรหรือกลุ่มผู้สอบ (population) โดยตามหลักของการวัดและประเมินผลทางการศึกษาแล้ว นักวัดผลการศึกษาต่างยอมรับกันว่า การวัดและประเมินผลการเรียนนั้นขึ้นอยู่กับกลุ่มประชากรหรือกลุ่มผู้สอบ (Flanagan, 1951; Angoff, 1971) ดังนั้นผลการเรียนของแต่ละโรงเรียนจึงมีมาตรฐานแตกต่างกันไปตามประชากรซึ่งก็คือนักเรียนของแต่ละโรงเรียน

จากส่วนประกอบทั้งสามส่วน จะเห็นว่า ส่วนที่สองเป็นส่วนที่ทำให้ผลการเรียนของนักเรียนภายในโรงเรียนลำเอียงไปจากผลการเรียนที่ควรจะเป็น และส่วนที่สามเป็นส่วนที่ทำให้มาตรฐานของผลการเรียนแตกต่างกันระหว่างโรงเรียน

2. การปรับเทียบผลการเรียนระหว่างนักเรียนทั้งหมด เพื่อให้มีมาตรฐานเดียวกัน และสามารถเปรียบเทียบกันได้ (comparable) จึงต้องลดความลำเอียง (bias) ของผลการเรียนทั้งสองส่วน คือ ส่วนของโรงเรียนและส่วนของผู้เรียน โดยใช้การปรับสเกล (scaling) ผลการเรียนทั้งในส่วนของโรงเรียนและในส่วนของผู้เรียน จึงเรียกรูปแบบการปรับสเกลในลักษณะนี้ว่า “การปรับสเกลสองขั้น” (Two-step Scaling) โดยในระบบการศึกษาของไทย มีการทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐานหรือ O-NET ที่นักเรียนทุกคนต้องทำการทดสอบ ซึ่งดำเนินการทดสอบโดยหน่วยงานกลางคือ สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) และใช้แบบสอบเดียวกันทั้งประเทศ คุณลักษณะที่วัดในการทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐานก็เป็นคุณลักษณะเดียวกันกับการประเมินผลการเรียน ผลการทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐานจึงมีความเหมาะสมที่จะใช้เป็นสเกลอ้างอิง (reference scale) ในการเชื่อมโยงคะแนนแบบปรับเทียบสเกล (calibration) เนื่องจากผลการเรียนของนักเรียนในแต่ละโรงเรียนนั้น ประเมินผลจากคะแนนที่ได้จากแบบสอบของแต่ละโรงเรียน ซึ่งมีค่าสถิติของแบบสอบแตกต่างกันแต่วัดในคุณลักษณะเดียวกัน (Kolen, 2004 : 224)

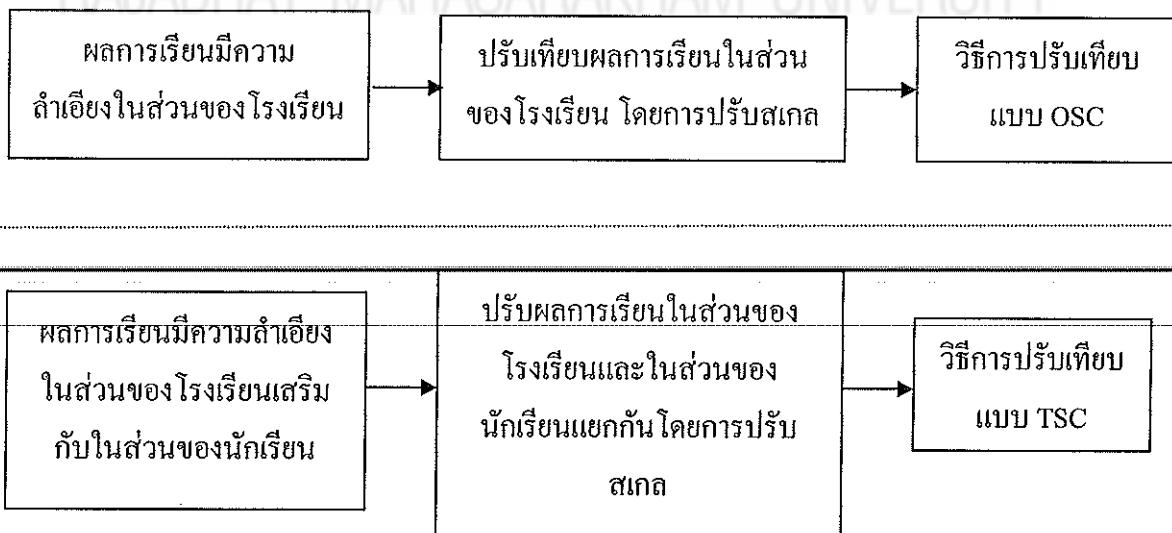
3. ความลำเอียงของผลการเรียนในส่วนของนักเรียน ซึ่งเกิดจากระบบการตัดสินใจ ผลการเรียนแบบกำหนดระดับผลสัมฤทธิ์เป็นตัวอักษร และความลำเอียงของผลการเรียนในส่วนของโรงเรียน ซึ่งก็คือ มาตรฐานของผลการเรียนที่แตกต่างกันระหว่างโรงเรียน จะมีผลต่อผลการเรียนของนักเรียนในแนวทางใดแนวทางหนึ่งดังนี้

3.1 ความลำเอียงของผลการเรียนในส่วนของนักเรียนเป็นแบบสุ่ม (Random Effect) หากอิทธิพลของระบบการตัดสินใจผลการเรียนเป็นอิทธิพลแบบสุ่ม ความลำเอียงของผลการเรียนของนักเรียนแต่ละคนภายในแต่ละโรงเรียนก็จะเกิดขึ้นอย่างสุ่ม นั่นคือ ระบบการตัดสินใจผลการเรียน อาจทำให้ผลการเรียนของนักเรียนแต่ละคนลำเอียงในทางถูกกเดเรด ถูกปล่อยกเรด หรือไม่มี ความลำเอียงก็ได้ ซึ่งถ้าหากภายใน 6 ภาคเรียนที่มีการคิดผลการเรียนเฉลี่ยสะสม (GPA) นั้น ค่าความลำเอียงเฉลี่ยของผลการเรียนของนักเรียนแต่ละคนเป็นศูนย์ แสดงว่า ระบบการตัดสินใจผลการเรียนเป็นอิทธิพลแบบสุ่ม หรือไม่มี ความลำเอียงของผลการเรียนในส่วนของนักเรียน การปรับเทียบ ผลการเรียนให้มีมาตรฐานเดียวกัน จึงเป็นการปรับความลำเอียงของผลการเรียน ในส่วนของโรงเรียนอย่างเดี่ยว โดยการปรับเทียบค่าเฉลี่ยของผลการเรียนของแต่ละโรงเรียนกับคะแนน O-NET เฉลี่ยของแต่ละโรงเรียน การปรับเทียบผลการเรียนแบบนี้เรียกว่า “การปรับเทียบผลการเรียนแบบปรับสเกลขั้นเดี่ยว” (OSC)

3.2 ความลำเอียงของผลการเรียนในส่วนของนักเรียน และความลำเอียงของผลการเรียนในส่วนของโรงเรียนมีผลต่อผลการเรียนของนักเรียนแบบเสริมกัน (Additive

Effect) เช่น ในโรงเรียนที่นักเรียนส่วนใหญ่มีความสามารถอยู่ในระดับปานกลางและอ่อนเมื่อเทียบกับโรงเรียนอื่นๆ ถ้าหากโรงเรียนใช้แบบสอบที่ยากในการประเมินผล ย่อมทำให้นักเรียนส่วนใหญ่สอบตก จึงต้องใช้แบบสอบที่ง่ายเมื่อเปรียบเทียบกับแบบสอบของโรงเรียนอื่นๆ ทำให้ผลการเรียนของนักเรียนในโรงเรียนนี้สูงกว่าที่ควรจะเป็น (ถูกปล่อยเกรดทั้งโรงเรียน) ในขณะที่เดียวกันก็ยังมีนักเรียนในโรงเรียนนี้อีกส่วนหนึ่งที่ได้รับผลจากระบบการตัดสินผลการเรียนให้มีผลการเรียนสูงขึ้นอีก คือนักเรียนกลุ่มที่ได้คะแนนอยู่ในขีดจำกัดล่างของช่วงคะแนนในแต่ละระดับผลการเรียน เช่น นักเรียนที่ได้ 80 คะแนน ในช่วงคะแนน 80-100 มีผลการเรียนเป็น 4.00 หรือนักเรียนที่ได้ 75 คะแนน ในช่วงคะแนน 75-79 มีผลการเรียนเป็น 3.50 เป็นต้น ดังนั้นผลการเรียนของนักเรียนกลุ่มนี้จึงเกิดจากความลำเอียงของผลการเรียนในส่วนของโรงเรียนเสริมกับความลำเอียงของผลการเรียนในส่วนของผู้เรียน ในกรณีนี้การปรับเทียบผลการเรียนให้มีมาตรฐานเดียวกันและสามารถเปรียบเทียบกันได้ จึงต้องปรับทั้งความลำเอียงของผลการเรียนในส่วนของโรงเรียน โดยการปรับเทียบค่าเฉลี่ยของผลการเรียนของแต่ละโรงเรียนกับคะแนน O-NET เฉลี่ยของแต่ละโรงเรียน และปรับความลำเอียงของผลการเรียนในส่วนของนักเรียน โดยปรับเทียบผลการเรียนของนักเรียนแต่ละคนที่ต่างจากผลเรียนของโรงเรียน (ผลการเรียนเฉลี่ยภายในโรงเรียน) กับผลการสอบ O-NET ของนักเรียนแต่ละคนที่ต่างจากผลการสอบ O-NET ของโรงเรียน (ค่าเฉลี่ยคะแนน O-NET ภายในโรงเรียน) ด้วย การปรับเทียบผลการเรียนแบบนี้เรียกว่า “การปรับเทียบผลการเรียนแบบปรับสเกลสองชั้น” (TSC)

จากกรอบแนวคิดการพัฒนาวิธีการปรับเทียบผลการเรียน โดยใช้หลักการเชื่อมโยงเชิงเส้นสองชั้นพอสรุปเป็นแผนภาพได้ดังนี้



ภาพ 1 กรอบแนวคิดในการวิจัย