

# การวัดและประเมินผลทางการศึกษากับการปรับเทียบคะแนน

## An Educational Measurement and Evaluation and Score Equivalence

ผศ.ศศิธร ชุตินันท์กุล

Asst.Prof.Sasithorn Chutinanthakul

ศูนย์วิชาประเมินผล สำนักทะเบียนและวัดผล มหาวิทยาลัยสุโขทัยธรรมาธิราช

### บทคัดย่อ

การวัดประเมินผลทางการศึกษา คือกระบวนการหนึ่งในการจัดการศึกษา เพื่อใช้ในการตรวจสอบว่าการจัดกิจกรรมการเรียนการสอน สามารถทำให้ผู้เรียนบรรลุเป้าหมายของการจัดการศึกษา ด้วยการใช้วิธีการวัดและประเมินผลที่หลากหลาย ส่วนการปรับเทียบคะแนนเป็นการนำวิธีการทางสถิติมาใช้ในการทำให้แบบสอบต่างฉบับที่วัดในเนื้อหาเดียวกัน เพื่อให้คะแนนจากแบบทดสอบต่างฉบับสามารถเทียบแทนกันได้ เพื่อก่อให้เกิดความยุติธรรมกับผู้เข้าสอบทุกคน โดยมีเป้าหมายหลักของการวัดและประเมินผลทางการศึกษา แบ่งเป็นหลักใหญ่ ๕ ประการ คือ (๑) เพื่อการจัดตำแหน่ง (๒) เพื่อการวินิจฉัย (๓) เพื่อพัฒนาการ ในการเรียนรู้ของผู้เรียน (๔) เพื่อการพยากรณ์ เป็นการนำผลการเรียนรู้ที่ได้ในปัจจุบันไปทำนายอนาคต และ (๕) เพื่อการตัดสินหรือประเมินค่า และการปรับเทียบคะแนน เพื่อเป็นพื้นฐานในการศึกษาเชิงลึกต่อไป

**คำสำคัญ :** การวัดประเมินผลทางการศึกษา, การเปรียบเทียบคะแนน

## Abstract

The educational evaluation is processes of the educational management for checking the learning activities that can make the students get its objectives by many ways of its evaluation. For improving and comparing the marks, this way is the process of statistics making the scores similar to which the different kinds of examination had used for no bias to the students. There are the significant objectives divided into five items as follows. (1) For setting the placement of students who have in some levels that are in distinction, good or unsatisfied ranking, (2) For knowing the diagnosis of them who show the causes of learning for developing it better, (3) For realizing the growth of them who develop it better or worse, (4) For predicting the results of it in the future and (5) for evaluation in taking the their results to measure and management on the learning further.

**Keywords :** educational evaluation, comparing the marks

## บทนำ

การวัดประเมินผลทางการศึกษา เป็นกระบวนการหนึ่งในการจัดการศึกษา เพื่อใช้ในการตรวจสอบว่าการจัดกิจกรรมการเรียนการสอน สามารถทำให้ผู้เรียนบรรลุเป้าหมายของการจัดการศึกษาได้หรือไม่ ด้วยการใช้วิธีการวัดและประเมินผลที่หลากหลาย ส่วนการปรับเทียบคะแนนเป็นการนำวิธีการทางสถิติมาใช้ในการทำให้แบบทดสอบต่างฉบับที่วัดในเนื้อหาเดียวกัน เพื่อให้คะแนนจากแบบทดสอบต่างฉบับสามารถเทียบแทนกันได้ เพื่อก่อให้เกิดความยุติธรรมกับผู้เข้าสอบทุกคน

ก่อนที่จะกล่าวถึงการวัดประเมินผลทางการศึกษา ควรรู้เกี่ยวกับเป้าหมายของการศึกษาก่อน ซึ่งพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. ๒๕๔๒ และที่แก้ไขเพิ่มเติม (ฉบับที่ ๒) พ.ศ. ๒๕๔๕ กำหนดไว้ว่า การจัดการศึกษาต้องเป็นไปเพื่อพัฒนาคนไทยให้เป็นมนุษย์ที่สมบูรณ์ทั้งร่างกาย จิตใจ สติปัญญา ความรู้และคุณธรรม มีจริยธรรมและวัฒนธรรมในการดำรงชีวิต สามารถอยู่ร่วมกับผู้อื่นได้อย่างมีความสุข โดยการจัดการศึกษาทั้งในระบบ การศึกษานอกระบบและการศึกษาตามอัธยาศัย<sup>๑</sup>

<sup>๑</sup> กระทรวงศึกษาธิการ, พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. ๒๕๔๒ และที่แก้ไขเพิ่มเติม (ฉบับที่ ๒) พ.ศ. ๒๕๔๕, (กรุงเทพมหานคร : โรงพิมพ์พรทิพพานกราฟฟิค จำกัด, ๒๕๔๕).

การจัดการศึกษาผ่านทางสถาบันการศึกษาเป็นกระบวนการหนึ่ง ในการพัฒนาคนให้บรรลุตาม เป้าหมายการศึกษา ด้วยการจัดการเรียนรู้ผ่านทางหลักสูตร เพื่อใช้เป็นกรอบและทิศทางในการพัฒนา คุณภาพของผู้เรียน กลุ่มเวลาผู้เรียนจนกระทั่งสำเร็จการศึกษาออกมาสู่สังคมโลกแห่งความเป็นจริง ที่มีการต่อสู้แข่งขันทั้งจากสังคมรอบตัวและสังคมโลก กว่าจะสำเร็จต้องผ่านกระบวนการเรียนรู้ที่ หลากหลายเพื่อสร้างภูมิคุ้มกัน เติบโตทั้งองค์ความรู้ สมรรถนะและคุณลักษณะที่พึงประสงค์ให้เต็ม ตามศักยภาพของแต่ละบุคคลตามเป้าหมายของการจัดการศึกษา

การจะรู้ว่ากระบวนการจัดการศึกษาที่ป้อนให้กับผู้เรียน จะทำให้ผู้เรียนสำเร็จลุล่วงเป็นไปตาม เป้าหมายหรือไม่ ก็ด้วยอาศัยการวัดและประเมินผล ผ่านเครื่องมือวัดและประเมินผลที่หลากหลาย ไม่ว่าจะเป็นการซักถาม การสังเกต การประเมินจากชิ้นงาน แฟ้มสะสมงานหรือการใช้แบบสอบถาม หรือด้วยวิธีการอื่นๆ เพื่อดูว่าผู้เรียนมีความก้าวหน้า มีพัฒนาการ บรรลุตามวัตถุประสงค์ของการจัด กิจกรรมการเรียนการสอนหรือไม่ ซึ่งสามารถทำได้ทั้งการวัดและประเมินผลก่อนการเรียนการสอน เพื่อดูความพร้อมของผู้เรียนเป็นการสอบก่อนสอน (Pretest) การวัดและประเมินผลในระหว่าง การเรียนการสอน (Formative evaluation) เพื่อนำผลจากการวัดและประเมินผลมาปรับปรุงการเรียน การสอน และการวัดและประเมินผลหลังจากการเรียนการสอน (Summative evaluation) เพื่อนำ ผลจากการวัดและประเมินผลมาตัดสินผลการเรียนรู้ มาตัดเกรดเพื่อให้ระดับการศึกษา

เป้าหมายหลักของการวัดและประเมินผลทางการศึกษา แบ่งเป็นหลักใหญ่ ๆ ได้ ๕ ประการ คือ (๑) เพื่อการจัดตำแหน่ง (Placement) ผู้เรียนว่าใครมีความรู้ความสามารถอยู่ในระดับใด ผ่านเกณฑ์ ไม่ผ่านเกณฑ์ เก่ง ปานกลาง หรืออ่อน หรือเพื่อการตัดเกรด (๒) เพื่อการวินิจฉัย (Diagnostic) เป็นการ วินิจฉัยผู้เรียน ค้นหาสาเหตุเพื่อนำมาปรับปรุงหรือพัฒนาส่งเสริมให้ดียิ่งๆ ขึ้นไป (๓) เพื่อดูพัฒนาการ หรือความงอกงาม (Growth) ดูความก้าวหน้าในการเรียนรู้ของผู้เรียนว่าดีขึ้นหรือลดลง (๔) เพื่อการ พยากรณ์ (Prediction) เป็นการนำผลการเรียนรู้ที่ได้ในปัจจุบันไปทำนายอนาคต เช่น ใช้ผลการเรียน คาดการณ์ในอนาคตว่าจะเรียนสำเร็จหรือไม่สำเร็จ ใช้ผลการสอบวัดความถนัด ทำนายว่าควรเรียน เพื่อประกอบอาชีพอะไรในอนาคต เป็นต้น และ (๕) เพื่อการตัดสินหรือประเมินค่า (Evaluation) เป็นการนำผลการวัดและประเมินผล มาตัดสินคุณภาพการจัดการเรียนการสอน

จะเห็นได้ว่า การวัดและประเมินผลทางการศึกษามีจุดมุ่งหมายอยู่ด้วยกันหลายแบบ ขึ้นอยู่กับว่าผู้ใช้จะต้องการนำไปใช้เพื่อวัตถุประสงค์ใด สำหรับการเรียนการสอนแล้วการวัดและประเมินผล

การเรียนรู้ตั้งอยู่บนหลักการพื้นฐานหลัก ๒ ประการคือ การประเมินเพื่อพัฒนาผู้เรียนและเพื่อตัดสินผลการเรียน ซึ่งการตัดสินผลการเรียน เป็นเป้าหมายสุดท้ายของการจัดการเรียนการสอนในชั้นเรียนที่ครูโดยส่วนใหญ่ใช้และวิธีการหรือเครื่องมือที่นิยมนำไปใช้กันก็คือ แบบทดสอบทั้งปรนัยและอัตนัย โดยเฉพาะแบบทดสอบปรนัย ในการสร้างแบบทดสอบครูทุกคนพยายามสร้างให้มีคุณภาพโดยคำนึงถึงความตรง (Validity) ของแบบทดสอบ คือสร้างให้มีลักษณะตรงตามสิ่งที่ต้องการวัด ตรงตามเนื้อหา ตรงตามโครงสร้างและตรงตามสภาพจริงที่ต้องการวัด มีความเที่ยง (Reliability) แบบทดสอบที่ดีต้องมีความคงเส้นคงวาในการวัดไม่ว่าจะสอบวัดกี่ครั้งก็ทำตามต้องได้ผลใกล้เคียงกันมากที่สุด มีความเป็นปรนัย (Objectivity) วัดได้ตรงตามที่ต้องการวัดโดยมีข้อคำถามที่ชัดเจน ตรวจสอบและแปลความหมายของคะแนนที่ได้ตรงกัน มีความยาก (Difficulty) พอเหมาะไม่ยากหรือง่ายเกินไป มีอำนาจจำแนก (Discriminant) แบบทดสอบที่ดีจะต้องมีข้อคำถามแต่ละข้อที่สามารถแยกผู้เรียนเก่ง-อ่อนได้ คนเก่งต้องสามารถตอบคำถามถูก คนอ่อนต้องตอบคำถามผิด แบบทดสอบที่ดีต้องมีความครอบคลุม (Comprehensive) มีจำนวนข้อเพียงพอเหมาะสมกับสิ่งที่วัดมีประสิทธิภาพ (Efficientcy) สามารถวัดสิ่งที่เป็นตัวแทนของสิ่งที่ต้องการวัดได้ จะวัดเมื่อไหร่ก็ได้ และเป็นแบบทดสอบที่ง่ายและสะดวกต่อการนำไปใช้ (Practicability on usability)

จะเห็นได้ว่ากว่าจะได้แบบทดสอบที่ดีที่มีคุณภาพ ต้องผ่านการสร้างมาอย่างพิถีพิถัน และเมื่อเจอกับสถานการณ์ที่มีผู้สอบจำนวนมาก ถ้ามาสอบพร้อมกันจะมีปัญหาเรื่องสถานที่สอบไม่เพียงพอ ไม่มีคนคุมสอบ จำเป็นต้องจัดสอบหลายครั้งต่างเวลา ต่างสถานที่สอบ ปัญหาที่พบคือผู้สอบครั้งแรกนำข้อสอบไปเผยแพร่ เกิดการรั่วไหลของข้อสอบผู้สอบครั้งหลังรู้ข้อสอบ หรือมีการศึกษาค้นคว้าเพิ่มเติม ผู้จัดจึงแก้ปัญหาด้วยการใช้แบบทดสอบหลาย ๆ ฉบับเพื่อป้องกันการรั่วไหลและเป็นการรักษาความลับของข้อสอบ ด้วยการสร้างแบบทดสอบหลายๆ ฉบับให้มีความเป็นคู่ขนานกัน ให้มีความเท่าเทียมกันมากที่สุด แต่ในความเป็นจริงการสร้างแบบทดสอบให้มีความเท่าเทียมกันมีค่าความยาก-ง่ายเท่ากัน มีความเที่ยงและความตรงเท่ากันทุกฉบับเป็นสิ่งที่ยากมาก หรือถ้าทำได้ก็จะต้องลงทุนสูงมากกว่าจะได้แบบทดสอบที่เท่าเทียมกันทุกอย่าง ซึ่งอาจจะไม่คุ้มค่ากับการลงทุน นักจิตวิทยา มิติจึงได้เสนอวิธีการที่จะทำให้แบบทดสอบต่างฉบับมีความเท่าเทียมกันได้ด้วยการใช้วิธีการทางสถิติเข้าช่วย นั่นก็คือ การปรับเทียบคะแนน (Equating)

ดังนั้น ในการวัดและประเมินผลทางการศึกษา เมื่อจำเป็นจะต้องใช้แบบทดสอบหลายฉบับในการสอบคราวเดียวกัน เพื่อให้ความยุติธรรมกับผู้เข้าสอบทุกคนไม่ว่าผู้นั้นจะได้แบบทดสอบฉบับที่ง่ายกว่า หรือฉบับที่ยากกว่าก็ตาม ก็ควรจะนำวิธีการปรับเทียบคะแนนมาใช้เพื่อปรับคะแนนจากแบบทดสอบต่างฉบับให้สามารถเทียบแทนกันได้ เช่นผู้สอบที่ได้แบบทดสอบฉบับง่ายทำได้ ๓๐ คะแนน เมื่อใช้วิธีการปรับเทียบคะแนนแล้วอาจจะเทียบได้กับผู้สอบที่ทำได้ ๒๕ คะแนนจากการทำแบบทดสอบฉบับที่ยากกว่าวิธีการปรับเทียบคะแนนจึงสามารถสร้างความเป็นธรรมให้กับผู้สอบที่ได้แบบทดสอบต่างฉบับที่มีความยากแตกต่างกันได้ เราจึงควรมาทำความรู้จักกับวิธีการปรับเทียบคะแนน ซึ่งในทางปฏิบัติวิธีการนี้ยังไม่ได้เป็นที่รู้จักและไม่ได้นำไปใช้อย่างแพร่หลาย มีแต่เฉพาะในการทำวิจัยที่ทำการศึกษาน้นหนักในเชิงทฤษฎี

การปรับเทียบคะแนน แองกอฟฟ์ได้ให้ความหมายไว้ว่า หมายถึงกระบวนการแปลงระบบคะแนนของแบบทดสอบฉบับหนึ่งไปสู่ระบบคะแนนของแบบทดสอบอีกฉบับหนึ่งซึ่งวัดสิ่งเดียวกัน<sup>๒</sup> เพื่อให้คะแนนแปลงจากแบบทดสอบทั้ง ๒ ฉบับนั้น สามารถเทียบเท่ากันและใช้แทนกันได้ ส่วนโคลเลน และเบรนนัน ให้ความหมายของการปรับเทียบคะแนนว่าเป็นกระบวนการทางสถิติที่ใช้ในการปรับคะแนนจากแบบทดสอบต่างฉบับ เพื่อให้คะแนนจากแบบทดสอบต่างฉบับสามารถเทียบแทนกันได้<sup>๓</sup> การปรับเทียบคะแนนจึงเป็นกระบวนการทำให้แบบทดสอบต่างฉบับให้มีความเท่าเทียมกันสามารถเทียบแทนกันได้ แบบทดสอบที่จะนำมาปรับเทียบคะแนนกันจะต้องเป็นแบบทดสอบที่วัดคุณลักษณะเดียวกัน และมีความเที่ยงใกล้เคียงกัน เป็นแบบทดสอบที่คู่ขนานกันในด้านโครงสร้าง (Structure) เวลาที่ใช้สอบ (Timing) ชนิดของข้อสอบ (Item Types) รูปแบบ (Formats) และเนื้อหา<sup>๔</sup> (Subject matter)

การปรับเทียบคะแนน แบ่งออกได้เป็น ๒ ประเภท คือ การปรับเทียบตามแนวนอน (Horizontal Equating) เป็นการปรับเทียบคะแนนระหว่างแบบทดสอบต่างฉบับกัน ที่วัดคุณลักษณะเดียวกัน มีระดับความยากใกล้เคียงกัน เป็นเทคนิคที่เหมาะสมในสถานการณ์ที่มีความจำเป็นจะต้องสร้างแบบทดสอบที่มีเนื้อหาเดียวกันหลาย ๆ ฉบับ เพื่อนำไปใช้ในการสอบเพื่อให้เกิดความยุติธรรม และป้องกัน

<sup>๒</sup> Angoff, W.H. *Scales, Norms and Equivalent Score*. Educational Measurement, 1971.

<sup>๓</sup> Kolen, M.J. and Brennan, R.L. *Test equating, scaling, and linking*, 2 nd , New York: Springer, 2004.

<sup>๔</sup> ศิริชัย กาญจนวาสี. *ทฤษฎีการทดสอบแนวใหม่(ฉบับปรับปรุง)*, (กรุงเทพมหานคร : ศูนย์ตำราเอกสารทางวิชาการ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, ๒๕๕๕).

ความลับของข้อสอบ เมื่อใช้ต่างเวลากันเพื่อเปรียบเทียบคะแนนที่ได้จากฉบับหนึ่งเทียบเป็นเท่าไรของอีกฉบับหนึ่ง ที่วัดในระดับเดียวกัน จึงเป็นการปรับเทียบคะแนนระหว่างแบบทดสอบต่างฉบับของวิชาเดียวกัน สำหรับกลุ่มผู้สอบระดับชั้นเดียวกัน กับการปรับเทียบคะแนนตามแนวตั้ง (Vertical Equating) เป็นการปรับเทียบคะแนนระหว่างแบบทดสอบต่างฉบับกัน เมื่อแต่ละฉบับมุ่งวัดลักษณะเดียวกัน แต่มีระดับความยากแตกต่างกัน และกลุ่มผู้สอบมีความสามารถอยู่ต่างประชากรกัน หรือมีความสามารถแตกต่างกัน เป็นเทคนิคที่เหมาะสมในสถานการณ์ที่มีความจำเป็นต้องสร้างแบบทดสอบเนื้อหาเดียวกัน แต่มีความยาก มีความลึกแตกต่างกันอยู่คนละชั้นกัน เป็นแบบทดสอบที่มุ่งวัดความสามารถของผู้สอบที่ต่างระดับกัน เพื่อปรับเทียบว่าคะแนนที่สอบได้จากฉบับหนึ่งเทียบเป็นเท่าไรของฉบับอื่นที่วัดต่างระดับกัน จึงเป็นการปรับเทียบคะแนนระหว่างแบบทดสอบของวิชาเดียวกัน สำหรับกลุ่มผู้สอบต่างระดับชั้นกัน

สำหรับรูปแบบการปรับเทียบคะแนน (Designs for Equating Method) มีอยู่ด้วยกัน ๒ รูปแบบใหญ่ ๆ คือ (๑) การปรับเทียบคะแนนตามแนวทฤษฎีการวัดแบบมาตรฐานเดิม (Classical Test Theory) มีวิธีการที่นิยมใช้ได้แก่ วิธีปรับเทียบคะแนนอิกวิเปอร์เซ็นไทล์ เป็นวิธีการปรับเทียบที่ยึดหลักการว่า การแจกแจงของคะแนนจากแบบทดสอบ X และแบบทดสอบ Y มีลักษณะคล้ายกัน การปรับเทียบคะแนนทำได้โดยใช้คะแนน ณ ตำแหน่งเปอร์เซ็นต์ไทล์เดียวกัน วิธีเชิงเส้นตรง เป็นวิธีการใช้สมการเส้นตรงในการแปลงคะแนน โดยมีหลักการว่าคะแนนจากแบบทดสอบทั้ง ๒ ฉบับ จะมีความเท่าเทียมกันก็ต่อเมื่อคะแนนของแต่ละฉบับมีคะแนนมาตรฐาน (Z-score) เท่ากัน วิธีใช้สมการถดถอยเป็นการใช้สมการเส้นตรงในการทำนายตัวแปรตาม ฯลฯ และ (๒) การปรับเทียบคะแนนแนวทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) สามารถปรับเทียบได้ทั้งแบบ ๑ ๒ และ ๓ พารามิเตอร์ ซึ่งสามารถปรับเทียบคะแนนได้ทั้งการใช้คะแนนจริงและคะแนนสังเกต

ส่วนการออกแบบเพื่อเก็บรวบรวมข้อมูล ลอร์ด ได้เสนอวิธีการออกแบบไว้ ๔ รูปแบบ<sup>๕</sup> คือ

๑. รูปแบบผู้สอบกลุ่มเดียว (Single- Group Design) แบ่งเป็น ๒ แบบย่อย คือ

๑.๑ ผู้สอบกลุ่มเดียวที่ไม่ได้รับการจัดให้สมดุล (Uncounter balanced Design) ใช้ผู้สอบกลุ่มเดียว แต่ละคนทำข้อสอบ ๒ ฉบับ จึงสามารถนำคะแนนมาปรับเทียบกันได้บนพื้นฐานความ

<sup>๕</sup> Kolen, M.J. and Brennan, R.L. Test equating, scaling, and linking, 2 nd, New York: Springer, 2004.

สามารถที่เท่ากัน แต่ในทางปฏิบัติการสอบฉบับหลัง อาจได้รับปัจจัยที่ผลกระทบจากการเรียนรู้การสอบฉบับแรก ความเมื่อยล้า ฯลฯ ซึ่งอาจมีอิทธิพลต่อการปรับเทียบ

๑.๒ ผู้สอบกลุ่มเดียวได้รับการจัดให้สมดุล (Counter balanced Design) รูปแบบนี้จัดผลของลำดับก่อน-หลัง โดยสุ่มผู้สอบแยกเป็น ๒ กลุ่ม แต่ละกลุ่มได้รับการสอบ ๒ ฉบับ โดยกลุ่มแรกสอบแบบทดสอบฉบับที่ ๑ ตามด้วยฉบับ ๒ ส่วนกลุ่มหลังให้สอบแบบทดสอบฉบับที่ ๒ ตามด้วยฉบับที่ ๑ เพื่อให้เกิดความสมดุล

๒. รูปแบบผู้สอบกลุ่มเท่าเทียมกัน (Equivalent-Group Design) โดยจัดกลุ่มผู้สอบให้มีความคล้ายคลึงกันมากที่สุด แล้วให้แต่ละกลุ่มทำแบบทดสอบฉบับเดียวกัน จึงสามารถหลีกเลี่ยงปัญหาการเรียนรู้ ความเมื่อยล้า แต่อาจจะมีปัญหาในกลุ่มที่ใช้อาจมีการแจกแจงความสามารถที่แตกต่างกัน วิธีการที่จะลดความแตกต่าง ควรใช้กลุ่มตัวอย่างขนาดใหญ่

๓. รูปแบบผู้สอบต่างกลุ่มโดยใช้แบบสอบร่วม (Anchor-Test Design) แบ่งเป็น ๒ แบบย่อยคือ

๓.๑ ผู้สอบกลุ่มสุ่มโดยใช้แบบสอบร่วม (Anchor-Test Random Group Design) แต่ละกลุ่มทำแบบทดสอบฉบับเดียว แบบทดสอบแต่ละฉบับมีข้อสอบร่วมจำนวนหนึ่ง ซึ่งอาจจะจัดรวมไว้ในแบบสอบเดียวกัน (Internal Anchor Test) หรือจัดแยกต่างหาก (External Anchor Test) รูปแบบนี้ช่วยลดความลำเอียงที่มีสาเหตุจากการเรียนรู้ ความเมื่อยล้า และความแตกต่างเกี่ยวกับระดับความสามารถของกลุ่มผู้สอบ แต่ประสิทธิภาพของการปรับเทียบคะแนนยังขึ้นอยู่กับคุณภาพของแบบสอบร่วม จึงต้องคำนึงถึงเนื้อหา ระดับความยากที่จะต้องมีความคล้ายกับแบบสอบที่ต้องการปรับเทียบคะแนนและความยาวของแบบสอบร่วมก็เป็นปัจจัยสำคัญที่ส่งผลต่อคุณภาพของการปรับเทียบ

๓.๒ ผู้สอบกลุ่มไม่เท่าเทียมกันโดยใช้แบบสอบร่วม (Anchor-Test Nonequivalent Group Design) ใช้ผู้สอบกลุ่มไม่เท่าเทียมกันแต่ละกลุ่มทำข้อสอบฉบับเดียวที่มีข้อสอบร่วมภายในหรือแบบสอบร่วมภายนอกก็ได้ ซึ่งใช้ในสถานการณ์ที่ผู้สอบต่างโปรแกรม เวลา หรือระดับ

๔. รูปแบบผู้สอบต่างกลุ่มโดยมีผู้สอบร่วม (Common-Person Design) เป็นการให้ผู้สอบต่างกลุ่มกันทำแบบทดสอบต่างฉบับ แต่มีจำนวนผู้สอบจำนวนหนึ่งจากทั้ง ๒ กลุ่ม ทำแบบสอบทั้ง ๒ รูปแบบ

## บทสรุป

ที่กล่าวมาข้างต้นเพียงเพื่อให้ผู้อ่านได้ทำความรู้จักกับการวัดและประเมินผลทางการศึกษา และการเปรียบเทียบคะแนนเพื่อเป็นพื้นฐานในการศึกษาเชิงลึกต่อไป ด้วยการนำวิธีการเหล่านี้ไปใช้ให้เกิดความยุติธรรมกับผู้เข้าสอบทุกคน

## บรรณานุกรม

- กระทรวงศึกษาธิการ. พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. ๒๕๔๒ และที่แก้ไขเพิ่มเติม (ฉบับที่ ๒) พ.ศ. ๒๕๔๕. กรุงเทพมหานคร : โรงพิมพ์พรทิฆานกราฟฟิค จำกัด, ๒๕๔๕.
- ศิริชัย กาญจนวาสี. ทฤษฎีการทดสอบแนวใหม่(ฉบับปรับปรุง) กรุงเทพมหานคร : ศูนย์ตำรา เอกสารทางวิชาการ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, ๒๕๕๕.
- สุโขทัยธรรมมาธิราช. การวัดและประเมินอิงมาตรฐานการเรียนรู้. นนทบุรี : สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมมาธิราช, ๒๕๕๕.
- \_\_\_\_\_. สถิติ วิจัย และประเมินผลการศึกษา. นนทบุรี : สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมมาธิราช ๒๕๕๔.
- \_\_\_\_\_. การวัดและประเมินผลสัมฤทธิ์ทางการเรียนระดับประถมศึกษา. นนทบุรี : สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมมาธิราช, ๒๕๕๑.
- Angoff. W.H. **Scales, Norms and Equivalent Score**. Educational Measurement, 1971.
- Holland, P.W. and Donald B Rubin, **Test Equating**. New Jersey: Education Testing Service, Princeeton, academic Press, 1982.
- Kolen, M.J. and Brennan, R.L. **Test equating, scaling, and linking**, 2<sup>nd</sup>, New York: Springer, 2004.
- Lord,F.M. **Application of item Response Theory to Practical Testing Problems**. New Jersey: Lawrence Erlbaum Associates inc.1980.