



ความสำคัญและปัญหาการวิจัย

การวัดผลทางการศึกษาเป็นกระบวนการของการทำให้เป็นปริมาณ ที่ช่วยติดตาม และพัฒนาคุณภาพการศึกษาในระดับต่าง ๆ การวัดผลจะมีความถูกต้องเมื่ออาศัยเครื่องมือที่ดี และมีสเกลสำหรับเปรียบเทียบที่เหมาะสมเพื่อใช้ในการแปลความหมาย แบบสอบ (Test) นับว่าเป็นเครื่องมือที่นิยมใช้แพร่หลายมากที่สุดทางการศึกษา เพราะถือว่าเป็นเครื่องมือที่วัดคุณลักษณะ ความสามารถหรือสมรรถภาพของผู้สอบได้ แบบสอบที่ดีจะต้องวัดได้ตรงกับความสามารถที่แท้จริง และทำให้เกิดความยุติธรรมแก่ทุกคนในการเปรียบเทียบ และแปลความหมาย

การวัดผลสัมฤทธิ์ทางการเรียนเป็นวิธีการหนึ่งที่ใช้ตรวจสอบคุณภาพของการศึกษา เพราะข้อมูลที่ได้มาจากการสอบจะช่วยให้ผู้สอนทราบถึงภาพรวมของผลการเรียน และความก้าวหน้าของผู้เรียน ซึ่งช่วยให้สามารถวินิจฉัยความสามารถทางการเรียนของผู้เรียน และสามารถจัดการเรียนการสอนให้กับผู้เรียนตามความเหมาะสมกับปัญหาที่แตกต่างกันไป และหากว่ามีการติดตามผลการจัดการศึกษาอย่างต่อเนื่อง โดยการสอบวัดผลกับนักเรียนในกลุ่ม ลายๆ และระดับชั้นต่าง ๆ กัน เพื่อนำคะแนนที่ได้จากการสอบในแต่ละปีการศึกษา มาเปรียบเทียบกัน เพื่อให้เห็นแนวโน้มผลการศึกษาว่าเพิ่มขึ้นหรือลดลงในแต่ละปีการศึกษา หรือเปรียบเทียบความก้าวหน้าของผลการเรียนของนักเรียนแต่ละคนในแต่ละปีการศึกษา เพื่อพิจารณาพัฒนาการทางการเรียนรู้ ถ้าในแต่ละปีการศึกษาสามารถดำเนินการสอบโดยใช้แบบสอบชุด เดียวกัน จะสามารถนำคะแนนดิบที่ได้จากการสอบมาเปรียบเทียบกันได้โดยตรง แต่ในทางปฏิบัติในการดำเนินการสอบแต่ละครั้งไม่สามารถกระทำได้เหมือนกันทุกครั้งที่ เพราะสถานการณ์ในการสอบอาจเปลี่ยนแปลงไปตามเวลา สถานที่ หรือประสบการณ์การเรียนของผู้เรียน ได้รับเพิ่มขึ้น ปัญหาการรักษาความลับของแบบสอบ ซึ่งผู้สอบในปีต่อ ๆ มาอาจได้รับการแนะนำฝึกฝนตอบแบบสอบที่เคยใช้มาแล้ว หรือบางคนอาจได้ทราบข้อสอบมาจากผู้สอบปีก่อน ๆ

ทำให้ผู้สอบในปีต่อ ๆ มาได้เปรียบหรือได้คะแนนมากกว่าผู้สอบปีก่อน ๆ หรือในกรณีที่ต้องการเปรียบเทียบผลการเรียนของผู้เรียนแต่ละคนในแต่ละปีการศึกษา เมื่อผู้เรียนได้รับเนื้อหาวิชาเพิ่มขึ้น มีประสบการณ์การเรียนรู้มากขึ้น แบบสอบที่จะใช้วัดผลกับผู้เรียนอาจแตกต่างจากแบบสอบที่ผู้เรียนเคยทำในแต่ละปีการศึกษา ดังนั้น คะแนนที่ได้ในแต่ละปีจึงไม่สามารถเปรียบเทียบกันได้โดยตรง

จากปัญหาดังกล่าว จึงเป็นหน้าที่ของผู้รับผิดชอบเกี่ยวกับการสอนและการสอบ จะต้องจัดทำแบบสอบใหม่ขึ้นซึ่งสามารถวัดคุณลักษณะเดียวกัน เพื่อให้เกิดความเหมาะสมกับสภาพการสอบในแต่ละสถานการณ์ จึงนำไปสู่ความคิดในการสร้างแบบสอบคู่ขนาน (Parallel Test) แต่ในการสร้างแบบสอบคู่ขนานเป็นเรื่องที่ไม่สามารถสร้างได้อย่างแท้จริง ดังที่กัลลิคเซน (Gulliksen, 1958) กล่าวว่า การสร้างแบบสอบคู่ขนานให้ได้ลักษณะข้อสอบ คือ เนื้อหาและสถิติของข้อสอบและแบบสอบทั้งฉบับเท่ากัน ในความเป็นจริงทำได้ยากและไม่คุ้มกับการลงทุน มักได้แบบสอบที่เทียบเคียงกันเท่านั้น ดังนั้นทางออกที่ดีที่สุด คือ การใช้เทคนิคทางสถิติมาปรับให้เกิดการเทียบเคียงกัน วิธีดังกล่าว คือ การเทียบมาตรฐานแบบสอบ (Test Equating) ซึ่งวิธีการนี้จะช่วยให้การเปรียบเทียบคะแนนที่ได้จากแบบสอบต่างฉบับมีความหมายยิ่งขึ้น ซึ่งสงบ ลักษณะ (2525, 22) กล่าวว่า คะแนนสอบจากแบบสอบต่างฉบับกันที่วัดสิ่งเดียวกัน แต่ไม่จำเป็นต้องเป็นข้อคู่ขนานจะถือว่าเทียบเท่ากันได้ ถ้านักเรียนคนใดก็ตามทำข้อสอบต่างฉบับแล้วได้คะแนนจริง (True Score) และเมื่อเราสร้างตารางเทียบคะแนนระหว่างแบบสอบได้ การใช้แบบสอบต่างฉบับกัน และนักเรียนต่างกลุ่มกันจะมีความสะดวกในการวัดผลเพราะไม่ว่าเขาจะได้รับการสอบด้วยแบบสอบฉบับใดก็สามารถนำคะแนนมาเทียบกันได้เสมอ

จากแนวความคิดเกี่ยวกับการเทียบมาตรฐานคะแนนของแบบสอบ องค์ประกอบที่สำคัญ 2 ส่วน คือ รูปแบบการเทียบมาตรฐาน ซึ่งเป็นการกำหนดเงื่อนไขการเทียบคะแนน และการสร้างกฎการแปลงคะแนน สำหรับการเทียบคะแนน (Angoff, 1984) ซึ่งถ้าพิจารณาการนำไปใช้กับผู้สอบและแบบสอบสามารถทำได้ 2 รูปแบบ ได้แก่ (Hambleton, 1984)

1) การเทียบมาตรฐานตามแนวนอน (Horizontal Equating) ใช้ในกรณีต้องการเปรียบเทียบแบบสอบที่มีระดับความยากพอ ๆ กัน และการกระจายความสามารถของผู้ตอบแบบสอบคล้ายคลึงกัน

2) การเทียบมาตรฐานในแนวตั้ง (Vertical Equating) ใช้ในกรณีเปรียบเทียบแบบสอบที่มีระดับความยากต่างกัน และการกระจายความสามารถของผู้สอบต่างกัน

เมื่อพิจารณาตามรูปแบบการเทียบมาตรฐานที่นิยมใช้กันมาก (Peterson, Marco and Stewart, 1982) ได้แก่ รูปแบบการเทียบมาตรฐานเชิงเส้น (Linear Equating Model) รูปแบบอิกวิเปอร์เซ็นต์ิล (Equipercetile Equating Model) รูปแบบอิงทฤษฎีตอบสนองข้อสอบ (Item Response Theory Equating Model) และได้มีการนำรูปแบบเหล่านี้มาศึกษาการเทียบมาตรฐานตามแนวนอนอย่างกว้างขวาง เช่น การเปรียบเทียบผลการเทียบมาตรฐานด้วยรูปแบบและตัวแปรต่าง ๆ (Kolen, 1982 ; Kolen and Whitney, 1982 ; Lord and Wingersky, 1984) การศึกษาความคลาดเคลื่อนในการเทียบมาตรฐาน (Angoff, 1971 : citing Lord 1956 ; Lord, 1982) การประเมินความเพียงพอของวิธีการเทียบมาตรฐานแบบต่าง ๆ (Slinde and Linn, 1978 ; Jaeger, 1981 ; Petersen and other, 1982) ทั้งนี้เพื่อประโยชน์ในการนำเทคนิคการเทียบมาตรฐานรูปแบบต่าง ๆ ไปใช้ได้อย่างเหมาะสมกับสถานการณ์

ส่วนการเทียบมาตรฐานในแนวตั้งมีน้อย จากรูปแบบที่กล่าวมานี้ มีกฎในการแปลงคะแนนต่างกันและขึ้นกับรูปแบบแต่ละอย่าง จึงมักเกิดปัญหาที่ว่า รูปแบบใดที่ให้ค่าการแปลงคะแนนได้ถูกต้องเหมาะสมที่สุด หรือให้โอกาสที่จะคลาดเคลื่อนได้น้อยที่สุด อย่างไรก็ตามงานวิจัยที่ผ่านมา ไม่อาจสรุปได้ว่า การเทียบมาตรฐานโดยใช้รูปแบบใดมีประสิทธิภาพมากที่สุดในทุกสถานการณ์ เพราะผลการเทียบมาตรฐานขึ้นอยู่กับองค์ประกอบหลายประการ ได้แก่ ลักษณะของแบบสอบ ลักษณะของกลุ่มตัวอย่าง รูปแบบของการเก็บรวบรวมข้อมูล เป็นต้น (Peterson and other, 1982 ; Kolen and Whitney, 1982) แต่ด้วยการเลือกใช้วิธีการในการเทียบมาตรฐานที่เหมาะสม จะมีความเป็นไปได้ที่จะวัดความก้าวหน้าหรือพัฒนาการเรียนรู้ของผู้เรียน เมื่อผ่านการเรียนการสอนในแต่ละช่วงเวลาต่าง ๆ จัดทำเป็นสารสนเทศเพื่อนำเสนอข้อมูลในรูปแบบต่าง ๆ และจัดสร้างตารางคะแนนที่เทียบเท่าระหว่างแบบสอบต่างชุดและสอบต่างเวลากันได้

จากการศึกษาของสไลด์และลินน์ (Slinde and linne, 1977) ได้สรุปและวิจารณ์การเทียบมาตรฐานเชิงเส้นตรง และการเทียบโดยใช้เปอร์เซ็นต์ิลว่า วิธีการเทียบโดยใช้เปอร์เซ็นต์ิลดีกว่าการเทียบมาตรฐานเชิงเส้นตรง แต่ทั้งสองวิธียังไม่เหมาะสมกับการเทียบมาตรฐานในแนวตั้ง และต่อมาเขาได้ศึกษาปัญหาเกี่ยวกับการใช้ราส์ซิมเคิล

ในการเทียบมาตรฐานแนวตั้ง พบว่า การเทียบมาตรฐานแนวตั้งในวิธีนี้ยังให้ผลไม่เป็นที่น่าพอใจ ซึ่งสอดคล้องกับการศึกษาของโฮมส์ (Holmes, 1982) ซึ่งพบว่า ผลการเทียบมาตรฐานโดยใช้วิธีสโตนแมน ปรากฏได้ค่าเฉลี่ยของคะแนนที่แปลงแล้วไม่เป็นที่น่าพอใจและจากการศึกษาของโคเลน (Kolen, 1981) พบว่า การเทียบมาตรฐานโดยใช้ทฤษฎีตอบสนองข้อสอบ (IRT) มีความเหมาะสมที่สุด รองลงมาคือรูปแบบการเทียบมาตรฐานเชิงเส้น สำหรับแบบสอบที่มีความยากต่างกัน

จากการศึกษาดังกล่าว จะเห็นได้ว่า วิธีการเทียบมาตรฐานโดยใช้ทฤษฎีตอบสนองข้อสอบน่าจะเหมาะสมที่สุด และจากการศึกษาส่วนมากจะเป็นการศึกษาการเทียบมาตรฐานเพียง 2 ระดับ (Slinde and linn, 1978; 1979 : Patience, 1981 : Kolen, 1981) ซึ่งยังมีประโยชน์ในแง่การนำไปประยุกต์ใช้ในวงจำกัด ดังนั้นผู้วิจัยเห็นว่า ควรมีการศึกษาวิธีการเทียบมาตรฐานแนวตั้งโดยใช้ทฤษฎีตอบสนองข้อสอบ พร้อมทั้งตรวจสอบคุณภาพของการเทียบมาตรฐานแนวตั้งความคลาดเคลื่อนมาตรฐาน และความเพียงพลของการเทียบมาตรฐาน เพื่อพัฒนาองค์ความรู้เกี่ยวกับการเทียบมาตรฐานแนวตั้งเพิ่มยิ่งขึ้น อีกทั้งเพื่อช่วยในการตัดสินใจในการนำวิธีการเทียบมาตรฐานตามแนวตั้ง โดยใช้ทฤษฎีตอบสนองข้อสอบไปใช้ได้อย่างมั่นใจยิ่งขึ้น ทั้งนี้เพื่อเป็นประโยชน์ต่อวงการวัดผลต่อไป

วัตถุประสงค์ของการวิจัย

เพื่อตรวจสอบคุณภาพของการเทียบมาตรฐานตามแนวตั้ง โดยใช้ทฤษฎีตอบสนองข้อสอบ สำหรับผู้สอบที่มีระดับความสามารถต่างกัน 3 ระดับ ได้แก่ ระดับมัธยมศึกษาปีที่ 1, 2 และ 3 โดยการวิเคราะห์

1. ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐานตามแนวตั้ง สำหรับผู้สอบที่มีระดับความสามารถต่างกัน 3 ระดับ โดยใช้ทฤษฎีตอบสนองข้อสอบ
2. ความเพียงพลของการเทียบมาตรฐานตามแนวตั้ง สำหรับผู้สอบที่มีระดับความสามารถต่างกัน 3 ระดับ โดยใช้ทฤษฎีตอบสนองข้อสอบ

สมมติฐานของการวิจัย

การเทียบมาตรฐานตามแนวคิด เป็นการเทียบมาตรฐานระหว่างผู้สอบที่มีความสามารถต่างกัน จากการวิจัยตรวจสอบคุณภาพของการเทียบมาตรฐานตามแนวคิด พบว่า เมื่อทำการเทียบมาตรฐานโดยใช้แบบสอบต่างชุดที่มีความยากต่างกัน และผู้สอบมีความสามารถต่างกัน วิธีการเทียบมาตรฐานโดยวิธีดั้งเดิมคือ การเทียบมาตรฐานเชิงเส้นตรงและการเทียบมาตรฐานแบบอิกวิเปอร์เซ็นต์ได้ให้ผลของการเทียบมาตรฐานใกล้เคียงกันซึ่งยังเป็นผลที่น่าพอใจ เนื่องจากความแตกต่างที่เกิดขึ้นที่ส่วนปลายของการเทียบมาตรฐานแบบดั้งเดิม ทำให้ผลต่างไปจากการเทียบมาตรฐานที่ใช้ทฤษฎีตอบสนองข้อสอบ อันเป็นผลมาจากข้อมูลส่วนนี้ไม่น่าเชื่อถือ เนื่องจากจะใช้ในการเทียบทำให้เกิดความคลาดเคลื่อนในการเทียบมาตรฐานเพิ่มขึ้น (Slind and Linn, 1977-1979 : Cook, Dunbar and Eignor, 1981 : Marco and Whitney, 1982) และเมื่อโคเลน (Kolen, 1981) ทำการศึกษาเปรียบเทียบการเทียบมาตรฐานระหว่างรูปแบบดั้งเดิมและรูปแบบการเทียบมาตรฐานโดยใช้โค้งลักษณะข้อสอบ เทียบจากฟอร์มที่มีความยากแตกต่างกัน พบว่า สำหรับแบบสอบที่มีความยากต่างกัน การเทียบมาตรฐานโดยใช้โค้งลักษณะข้อสอบ 3 พารามิเตอร์ อาจพบปัญหาที่เกิดขึ้นอย่างหลีกเลี่ยงไม่ได้ คือ การประมาณค่าไม่ถูกต้องของ Lower asymptote parameter แต่อย่างไรก็ตามการเทียบมาตรฐานโดยใช้ทฤษฎีตอบสนองข้อสอบมีความเหมาะสมในการเทียบมาตรฐานดีกว่ารูปแบบอื่น (Kolen and Whitney, 1982 : Hutten, 1982) ดังนั้นจึงได้กำหนดสมมติฐานของการวิจัยดังนี้

1. ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐาน (Standard error of equating: SEE) ในแนวคิด 3 ระดับ โดยใช้ทฤษฎีตอบสนองข้อสอบ น่าจะอยู่ในระดับต่ำ ($SEE < .05$)
2. คำนีความแตกต่างของการเทียบมาตรฐาน (Discrepancy index : C) ในแนวคิด 3 ระดับ โดยใช้ทฤษฎีตอบสนองข้อสอบ น่าจะมีค่าอยู่ในระดับที่น่าพอใจขึ้นไป ($C \leq (0.1SD_x)^2$)

ขอบเขตของการวิจัย

1. ประชากรที่ใช้ในการศึกษา เป็นนักเรียนชั้นมัธยมศึกษาตอนต้น ของโรงเรียน
ในสังกัดกรมสามัญศึกษา
2. แบบสอบ ใช้แบบสอบวิชาคณิตศาสตร์ ระดับมัธยมศึกษาตอนต้น เรื่อง สมการ
และอสมการ
3. วิธีการเทียบมาตรา ใช้วิธีการเทียบมาตราโดยวิธีทดสอบสองข้อสอบ แบบ
3 พารามิเตอร์

คำจำกัดความเชิงปฏิบัติการ

การเทียบมาตราของแบบสอบ (Test Equating) หมายถึง กระบวนการ
แปลงคะแนนจากแบบสอบฟอร์มหนึ่ง (X) ไปยังอีกฟอร์มหนึ่ง (Y) หรือฟอร์มอื่น ๆ โดยวิธี
การทางสถิติเพื่อปรับคะแนนระหว่างแบบสอบต่างชุดให้อยู่ในมาตราเดียวกัน ในการศึกษา
ครั้งนี้ ใช้รูปแบบการปรับโดยวิธีทดสอบสองข้อสอบแบบสามพารามิเตอร์

การเทียบมาตราตามแนวตั้ง (Vertical Equating) หมายถึง การเทียบ
มาตราที่ได้ จากกลุ่มตัวอย่างที่ระดับความสามารถต่างกัน และแบบสอบที่ระดับความยาก
ต่างกัน ในที่นี้ศึกษากลุ่มตัวอย่างและแบบสอบที่มีความยาก 3 ระดับ ได้แก่ ระดับชั้นมัธยม
ศึกษาปีที่ 1, 2 และ 3 และใช้แบบสอบ 3 ฉบับ ได้แก่ แบบสอบสำหรับระดับชั้น ม.1
ม.2 และ ม.3 ตามลำดับ

แบบสอบเทียบมาตรา หมายถึง แบบสอบต่าง ๆ ที่ใช้ในการเทียบมาตรา
สำหรับการวิจัยนี้ หมายถึง แบบสอบคณิตศาสตร์ 3 ฉบับ สำหรับนักเรียนชั้น ม.1 , ม.2
และ ม.3 ที่สร้างจากหลักสูตรมัธยมศึกษาตอนต้นของกระทรวงศึกษาธิการ พุทธศักราช
2521 (ฉบับปรับปรุง 2533) เรื่อง สมการและอสมการ แบบสอบแต่ละฉบับประกอบด้วย
ข้อสอบ 25 ข้อ

แบบสอบร่วมภายใน (Internal Anchor Test) หมายถึง กลุ่มข้อสอบร่วมที่อยู่ในแบบสอบเทียบมาตรฐานทั้ง 3 ฉบับ

ข้อสอบร่วม (Anchor items) หมายถึง กลุ่มข้อคำถามที่ร่วมกันอยู่ในแบบสอบต่างฉบับกัน สำหรับการวิจัยนี้ ประกอบด้วย ข้อสอบ 2 กลุ่ม ได้แก่ ข้อสอบร่วมระหว่างแบบสอบวิชาคณิตศาสตร์ ระดับชั้น ม.1 และ ระดับชั้น ม.2 จำนวน 5 ข้อ และข้อสอบร่วมระหว่างแบบสอบวิชาคณิตศาสตร์ ระดับชั้น ม.2 และ ระดับชั้น ม.3 จำนวน 5 ข้อ

คะแนนสมมูล (Equivalent Scores) คะแนนระหว่างแบบสอบต่างชุดที่เทียบคะแนนกันได้ว่าเป็นคะแนนที่เทียบเท่ากัน

คะแนนเกณฑ์ (Criterion Scores) หมายถึง คะแนนที่ได้จากแบบสอบชุดแรกของกลุ่มสอบทานผล

กลุ่มตัวอย่างเทียบมาตรฐาน (Equating Samples) หมายถึง กลุ่มตัวอย่างผู้สอบที่ได้จากการสุ่มประชากรนักเรียนระดับมัธยมศึกษาตอนต้น สำหรับการวิจัยนี้ ได้แก่ นักเรียนระดับ ม.1 , 2 และ 3 ของโรงเรียนในสังกัดกองการมัธยมศึกษา จังหวัดสุราษฎร์ธานี ซึ่งกลุ่มตัวอย่างเทียบมาตรฐานนี้ต้องทำแบบสอบคณิตศาสตร์เฉพาะในระดับชั้นของตนเอง คนละ 1 ฉบับ ผลการสอบที่ได้จากกลุ่มตัวอย่างนี้นำมาใช้ในการสร้างคะแนนแปลงตามรูปแบบการเทียบมาตรฐานตามแนวตั้ง 3 ระดับ

กลุ่มตัวอย่างสอบทานผล (Cross-Validation samples) หมายถึง กลุ่มตัวอย่างที่ได้จากการสุ่มประชากรนักเรียนระดับมัธยมศึกษาตอนต้น ซึ่งเป็นประชากรเดียวกันกับกลุ่มตัวอย่างเทียบมาตรฐาน โดยไม่มีหน่วยตัวอย่างซ้ำกันเลย สำหรับการวิจัยนี้ กลุ่มตัวอย่างสอบทานผลได้ โดยการสุ่มประชากรครั้งเดียวกับการสุ่มประชากรของกลุ่มตัวอย่างเทียบมาตรฐาน แต่แยกจากกลุ่มตัวอย่างเทียบมาตรฐานด้วยการใช้ขั้นตอนเก็บรวบรวมข้อมูลกลุ่มตัวอย่างนี้ทุกคนต้องทำแบบสอบคณิตศาสตร์ทั้ง 3 ฉบับ ผลการสอบที่ได้นำมาใช้ในการวิเคราะห์สอบทานผลการเทียบมาตรฐานตามแนวตั้ง 3 ระดับ

กลุ่มตัวอย่างที่ใช้ในการคำนวณหาค่าพารามิเตอร์ของข้อสอบ หมายถึง กลุ่มตัวอย่างที่สุ่มมาจากประชากรนักเรียนระดับชั้นมัธยมศึกษาตอนต้น ผลการสอบที่ได้จากกลุ่มนำมาใช้ในการคำนวณหาคุณภาพของแบบสอบ โดยคำนวณค่าพารามิเตอร์ คือ ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าการเดา (c) ของข้อสอบและแบบสอบ ในการวิจัยนี้ เป็นนักเรียนระดับมัธยมศึกษาตอนต้น ของโรงเรียนในสังกัดกองการมัธยมศึกษา

ตารางเทียบมาตรา หมายถึง ตารางแสดงตัวเลขคะแนนที่สมมูล (Equivalent Scores) ระหว่างแบบสอบต่างชุด ในที่นี้ หมายถึง คะแนนสมมูลจากแบบสอบระดับชั้น ม.1 ม.2 และ ม.3

คุณภาพของการเทียบมาตรา หมายถึง ความถูกต้องแม่นยำของการเทียบมาตรา ซึ่งพิจารณาจาก ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรา และความเพียงพอของการเทียบมาตรา

ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรา (Standard error of equating : SEE) คือ ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ยของความแตกต่างระหว่างคะแนนที่แปลงแล้ว ด้วยวิธีการเทียบมาตรากับคะแนนเดิมโดยไม่คิดทิศทาง

ความเพียงพอของการเทียบมาตรา หมายถึง ระดับความถูกต้องของการเทียบมาตราของแบบสอบชุดหนึ่งไปสู่แบบสอบชุดอื่นตามวิธีการของการเทียบมาตรา การตัดสินใจความเพียงพอของการเทียบมาตราที่ใช้ในการวิจัยนี้ คือ ใช้การเปรียบเทียบดัชนีความแตกต่าง (Discrepancy comparison index : C) กับเกณฑ์ของปีเตอร์สันและคณะ (Petersen et al 1982 : 93-94)

ดัชนีความแตกต่าง (Discrepancy comparison index :C) คือ ค่าความคลาดเคลื่อนรวม (Total error : Te) ในการเทียบมาตรฐานของวิธีการเทียบมาตรฐาน ซึ่งคำนวณได้จาก ค่าเฉลี่ยกำลังสองของความแตกต่างของคะแนนที่สมมูล ซึ่งได้จากการเทียบมาตรฐานกับคะแนนเกณฑ์ ที่ได้ถ่วงน้ำหนักด้วยความแปรปรวนของคะแนนเกณฑ์ ดังนั้นค่าที่ได้จึงเป็นค่ามาตรฐานที่เปรียบเทียบกันได้ หรือเปรียบเทียบเกณฑ์ของความพึงพอใจในการเทียบมาตรฐาน ดัชนีนี้ได้จากการวิเคราะห์กลุ่มสอบทานผล ใช้ประเมินประสิทธิภาพหรือความเพียงพอของวิธีการเทียบมาตรฐาน ถ้าค่า C มาก หมายความว่า การเทียบมาตรฐานด้วยวิธีนั้น ๆ ให้ความเพียงพอในการเทียบมาตรฐานน้อย และในทางตรงกันข้าม ค่า C น้อย หมายความว่า วิธีนั้น ๆ ให้ความเพียงพอในการเทียบมาตรฐานมาก

ประโยชน์ที่ได้รับจากการวิจัย

1. การวิจัยครั้งนี้ ช่วยให้ได้แนวทางในการตัดสินใจที่จะนำรูปแบบการเทียบมาตรฐานในแนวคิด โดยใช้ทฤษฎีตอบสนองข้อสอบ ไปใช้ในการวัดและการประเมินได้อย่างเหมาะสม
2. ทำให้การเทียบมาตรฐานตามแนวคิดระหว่างแบบสอบของนักเรียน ที่จะศึกษาพัฒนาการทางการเรียนรู้ต่างระดับชั้นของนักเรียน แต่ละโรงเรียน ในแต่ละกลุ่ม หรือเขตการศึกษา มีประสิทธิภาพและมีความหมายมากยิ่งขึ้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย