

บทที่ 2

การวัดผลแบบอิงกลุ่มและการวัดผลแบบอิงเกณฑ์

I จุดประสงค์การเรียนรู้

หลังจากจบบทเรียนนี้แล้ว นักศึกษาสามารถปฏิบัติได้ดังนี้

1. เปรียบเทียบความแตกต่างระหว่างการวัดผลแบบอิงกลุ่มและอิงเกณฑ์
2. อธิบายความแตกต่างระหว่างแบบทดสอบผลสัมฤทธิ์แบบอิงกลุ่มและอิงเกณฑ์
3. วิเคราะห์ความจำเป็นของการใช้วิธีการวัดผลแบบอิงกลุ่มและอิงเกณฑ์
4. อธิบายความเกี่ยวพันกันระหว่างการทดสอบแบบรอบรู้กับการวัดผลแบบอิงเกณฑ์
5. เมื่อกำหนดข้อมูลมาให้สามารถตัดสินใจเลือกใช้วิธีการวัดทั้ง 2 แบบที่จะทำให้เกิด

ประโยชน์สูงสุด

II รายละเอียดของเนื้อหา

1. การวัดผลแบบอิงกลุ่ม (Norm - referenced measurement)

เมอเรน (Mehren 1984:17) ให้นิยามว่า “การวัดผลแบบอิงกลุ่มเป็นการตีความหมายคะแนนผู้สอบแต่ละคนโดยเปรียบเทียบกับคนอื่น ๆ ที่อยู่ในกลุ่มเดียวกันที่สอบด้วยแบบทดสอบเดียวกัน”

การเปรียบเทียบคะแนนผู้สอบแต่ละคนนั้นจะต้องมีคะแนนมาตรฐานไว้เปรียบเทียบ เมื่อจะตีความหมายคะแนนของผู้สอบคนใดก็นำคะแนนของผู้เข้าสอบทั้งหมดมาจัดกระทำให้เป็นคะแนนมาตรฐานก่อน แล้วจึงนำคะแนนของผู้เข้าสอบคนนั้นมาเปรียบเทียบ ตัวอย่างเช่น

เด็กคนหนึ่งสอบวิชาวิทยาศาสตร์ได้ 65% และวิชาภาษาไทยได้ 50% ถ้าเราดูจากคะแนนตามนี้อย่างผิวเผินก็อาจจะตอบคร่าว ๆ ว่าเด็กคนนี้เก่งวิทยาศาสตร์ แต่อ่อนภาษาไทย แต่ถ้าเราเอาคะแนนนี้ไปเปรียบเทียบกับคะแนนมาตรฐาน เช่น คะแนนเปอร์เซ็นต์ไทล์ ก็อาจจะพบว่าคะแนนวิทยาศาสตร์อยู่เปอร์เซ็นต์ไทล์ที่ 50 และคะแนนภาษาไทยอยู่เปอร์เซ็นต์ไทล์ที่ 58

การตีความหมายคะแนนก็จะเปลี่ยนไปจากเดิมที่เราดูไว้ทันที การนำคะแนนของผู้เข้าสอบไปเปรียบเทียบกับคะแนนของคนอื่น ๆ ในกลุ่มเดียวกันที่สอบด้วยแบบทดสอบเดียวกันนี้เองเราเรียกว่า การวัดผลแบบอิงกลุ่ม

2. การวัดผลแบบอิงเกณฑ์ (Criterion - referenced measurement)

การวัดผลแบบอิงเกณฑ์นี้ เป็นการตีความหมายคะแนนผู้สอบแต่ละคนโดยการเปรียบเทียบกับพฤติกรรมที่กำหนดหรือเกณฑ์ที่เหมาะสม (Mehren, 1984:18) ซึ่งการวัดในลักษณะนี้ บราวน์เรียกว่าเป็นการวัดผลแบบอิงเนื้อหา (Brown, 1980:118)

ผู้ที่นำคำว่าอิงเกณฑ์มาใช้เป็นคนแรกคือ เกรเซอร์ (Glaser) (Mehren, 1984:18 citing Glaser 1963:520) การวัดผลแบบอิงเกณฑ์นั้นจะเป็นตัวชี้ให้เห็นว่าผู้สอบแต่ละคนมีความสามารถในด้านเนื้อหาและทักษะมากน้อยตามขอบเขตที่กำหนดไว้เพียงใด ต่อมาเมื่อปีค.ศ. 1971 Glaser และ Nitko (Glaser & Nitko, 1971:653) ได้ให้นิยามว่า แบบทดสอบอิงเกณฑ์เป็นแบบทดสอบที่สร้างขึ้นโดยมีเป้าหมายที่จะให้การวัดนั้นสามารถแปลความหมายพฤติกรรมเฉพาะจุดตามมาตรฐานที่กำหนด การให้นิยามการวัดผลแบบอิงเกณฑ์นี้ในปัจจุบันก็ยังคงมีการถกเถียงกัน มีการนำคำนี้ไปใช้ในทางที่ผิดพลาดเช่นการกำหนดว่ามาตรฐานเป็นจุดตัดของคะแนน เป็นต้น Nitko ได้ให้นิยามตามที่กล่าวมาแล้วโดยเน้นว่า คำว่าอิงเกณฑ์กับจุดตัดของคะแนนเป็นคนละประเด็นกัน

Popham (Popham & Husek, 1969:2) เป็นผู้หนึ่งที่ทำให้ความหมายของอิงเกณฑ์ว่า หมายถึงถึงพฤติกรรมที่ได้มาตรฐาน เขาได้แย้งว่าการตีความหมายเกณฑ์เป็นระดับความสามารถของผู้เข้าสอบก็เท่ากับเป็นการนำการวัดผลกลับไปสู่ระบบการวัดแบบดั้งเดิมนั่นเอง (Popham, 1981:8) ส่วน Brown (Brown, 1980:74) นั้น ถือว่าเกณฑ์คือระดับมาตรฐานของคะแนนที่กำหนดว่าเป็นตัวแปรภายนอก

จากคำจำกัดความของการวัดผลแบบอิงเกณฑ์นั้นจะเห็นได้ว่ายังไม่มีข้อยุติ อย่างไรก็ตามทุกฝ่ายต่างก็ยอมรับกันว่า การตีความหมายแบบอิงเกณฑ์ต้องการตอบคำถามที่ว่า ผู้สอบทำอะไรได้บ้างเมื่อเปรียบเทียบกับขอบเขต (domain) ของพฤติกรรมที่กำหนดไว้ ส่วนในเรื่องที่ยังคงถกเถียงกันอยู่เกี่ยวกับมาตรฐานของความสามารถ และจุดตัดของคะแนนนั้น อาจเนื่องมาจากแนวคิดในเรื่องของความเกี่ยวพันกันระหว่างการวัดผลแบบอิงเกณฑ์และการเรียนแบบรอบรู้ (Mastery Learning) ยังมีความใกล้เคียงกันมาก

3. ความแตกต่างระหว่างการวัดผลแบบอิงกลุ่มและอิงเกณฑ์

เพื่อให้เห็นความแตกต่างระหว่างการวัดผลแบบอิงกลุ่มและอิงเกณฑ์ชัดเจนขึ้น Theobald (Theobald 1974:68) ได้สรุปความแตกต่างของการวัดทั้ง 2 แบบนี้ดังแสดงในตาราง 2.1

ตารางที่ 2.1 สรุปความแตกต่างระหว่างการวัดผลอิงกลุ่มและอิงเกณฑ์ (Theobald, 1974:68)

ข้อเปรียบเทียบ	อิงกลุ่ม	อิงเกณฑ์
1. จุดมุ่งหมายในการพิจารณา	1. เพื่อพิจารณาผลของพฤติกรรมของเด็กแต่ละคนเมื่อเปรียบเทียบกับคนอื่น ๆ ในกลุ่มเดียวกับที่สอบด้วยแบบทดสอบเดียวกัน 2. เพื่อเน้นให้เห็นความแตกต่างระหว่างบุคคลโดยการเปรียบเทียบ	เพื่อพิจารณาสถานะของแต่ละบุคคลเมื่อเปรียบเทียบกับขอบเขตของพฤติกรรมที่กำหนดไว้ โดยไม่ต้องไปสัมพันธ์กับพฤติกรรมของคนอื่น ๆ
2. การเลือกเนื้อหาแบบทดสอบ	กระบวนการเลือกเนื้อหาทำแบบทดสอบจะต้องกว้างและเป็นตัวแทนของเป้าหมายของกระบวนการวิชาและจะต้องเลือกมาอย่างยุติธรรม	การเลือกเนื้อหาจะแคบกว่าและตรงจุดที่จะเน้นความสำคัญของพฤติกรรมที่ต้องการ
3. การสร้างแบบทดสอบ	1. ต้องหลีกเลี่ยงข้อสอบที่ยากเกินไปหรือง่ายเกินไป โดยเน้นให้ข้อสอบมีค่าอำนาจจำแนกสูง หรือให้คะแนนกระจายมาก ๆ 2. นักเรียนทุกคนทำข้อสอบเดียวกันทุกข้อ	1. เน้นการสร้างข้อสอบแต่ละข้อให้ตรงจุดประสงค์ที่จะวัด 2. นักเรียนทุกคนไม่จำเป็นจะต้องทำข้อสอบข้อเดียวกัน
4. การวิเคราะห์ข้อสอบ	1. วิเคราะห์ข้อสอบโดยแสดงให้เห็นค่าอำนาจจำแนก ค่าความยากง่าย หรือความกำกวมเพื่อที่จะนำมาปรับปรุง หรือตัดทิ้ง 2. ข้อสอบที่มีค่าอำนาจจำแนกเป็นลบแสดงให้เห็นว่าจำเป็นต้องมีความผิดพลาดเกิดขึ้น	1. ข้อสอบที่มีค่าอำนาจจำแนกต่ำหรือมีค่าความยากง่ายไม่พอเหมาะไม่จำเป็นจะต้องตัดทิ้ง 2. การพิจารณาข้อสอบจะดูเฉพาะประเด็นที่ว่าข้อสอบข้อนั้นได้สะท้อนให้เห็นถึงคุณลักษณะสำคัญที่ต้องการวัดไว้อย่างไรบ้าง 3. ข้อสอบที่มีค่าอำนาจจำแนกเป็นลบแสดงให้เห็นว่ามีข้อผิดพลาดเกิดขึ้น

ข้อเปรียบเทียบ	NR.	CR.
5. ความเที่ยงตรง	1. ความเที่ยงตรงเชิงเนื้อหาเป็นสิ่งที่สำคัญที่จะแสดงความยุติธรรมและตรงตามกระบวนวิชาที่เรียน 2. ความเที่ยงตรงอื่น ๆ เช่น ความเที่ยงตรงตามสภาพ ความเที่ยงตรงเชิงพยากรณ์ ความเที่ยงตรงตามโครงสร้าง ย่อมขึ้นอยู่กับเกณฑ์ที่วัด และการกระจายของคะแนน	1. ความเที่ยงตรงเชิงเนื้อหาเป็นสิ่งที่สำคัญอย่างยิ่ง ข้อสอบจะต้องตรงตามขอบเขตของพฤติกรรมที่กำหนดไว้ 2. ความเที่ยงตรงอื่น ๆ เป็นเพียงการสนับสนุนเท่านั้น
6. ความเชื่อมั่น	การหาความเชื่อมั่นแบบทดสอบคู่ขนาน แบบสอบซ้ำ และใช้สูตรของ Kuder - Richardson ขึ้นอยู่กับค่าสหสัมพันธ์ที่แปรไปตามความแปรปรวนของคะแนน	แบบทดสอบควรจะมีค่าคงที่ และมี ความคงที่ภายในสูง เช่น ข้อสอบ มีการกระจายของคะแนนต่ำ จะทำให้ค่าความเชื่อมั่นต่ำด้วย เช่น ข้อสอบที่นักเรียนทุกคนทำผ่าน
7. การตีความหมายคะแนน	1. มักใช้ และคะแนนมาตรฐาน เช่น Percentile Stanine 2. กลุ่มที่เปรียบเทียบกันอาจจำแนกภายในโรงเรียน ภายในชั้น หรือครูคนเดียวกันสอน	1. พฤติกรรมของเด็กจะนำเสนอเพียงแต่ผ่านเกณฑ์หรือไม่ 2. ถ้านักเรียนผ่านเกณฑ์ ผลของคะแนนอาจนำไปเปรียบเทียบภายในโรงเรียน ภายในชั้นหรือครูคนเดียวกันสอนก็ได้

จะเห็นได้ว่าจุดสำคัญของการใช้แบบทดสอบอิงเกณฑ์ก็เพื่อจะชี้ให้ชัดว่านักเรียนคนใดบ้างที่ผ่านเกณฑ์ คนใดบ้างไม่ผ่านเกณฑ์ โดยปกติถ้ามีการให้คะแนนแบบอิงเกณฑ์มาแล้วก็สามารถจะนำมาจัดทำแบบอิงกลุ่มได้ ข้อสำคัญที่จะต้องระวังก็คือ เกณฑ์ที่นำมาใช้ในการสร้างแบบทดสอบอิงเกณฑ์นั้นจะต้องแสดงให้เห็นว่าพฤติกรรมที่ครูหวังจะให้ผู้เรียนเกิดนั้นจะต้องแสดงให้เห็นชัดเจนตามระดับที่ต้องการ

4. แบบทดสอบผลสัมฤทธิ์แบบอิงกลุ่มและอิงเกณฑ์

ในการสร้างแบบทดสอบวัดผลสัมฤทธิ์นั้น มีจุดมุ่งหมาย 2 ประการคือ

1. เพื่อจำแนกเด็กแต่ละคนในกลุ่มตามอัตราความสัมฤทธิ์ผล
2. เพื่อจำแนกว่าใครบ้างผ่านหรือไม่ผ่านมาตรฐาน (หรือมีสัมฤทธิ์ผลตามจุดมุ่งหมายที่กำหนดไว้)

จากทฤษฎีการทดสอบและเทคนิคการสร้างข้อสอบ มีการพัฒนามาจากความเชื่อที่ว่า จุดมุ่งหมายของการสอบก็เพื่อจำแนกคนแต่ละคนในกลุ่ม แต่ถ้ามีการกำหนดจุดมุ่งหมายของการสอบเป็นการจำแนกว่าใครบ้างผ่านหรือไม่ผ่านมาตรฐาน หรือเป็นการเปรียบเทียบแต่ละคนกับมาตรฐานแล้ว การสร้างแบบทดสอบก็จะแตกต่างกันออกไป เกณฑ์การวัดของข้อสอบที่มีการวัดผลแบบอิงกลุ่มก็ย่อมจะต่างจากเกณฑ์การวัดข้อสอบแบบอิงเกณฑ์ คุณสมบัติของข้อสอบบางอย่างก็ต่างกัน เช่น ค่าอำนาจจำแนกของข้อสอบ ค่าความเชื่อมั่น ค่าความเที่ยงตรง บางประการ องค์ประกอบของการสร้างข้อสอบหลาย ๆ ด้านก็จะแตกต่างกันด้วย เช่น การกำหนดจุดมุ่งหมาย การเตรียมร่างแบบทดสอบ ส่วนการใช้คำในข้อสอบของแบบทดสอบ 2 ประเภทจะคล้าย ๆ กัน

แบบทดสอบแบบอิงกลุ่ม เป็นแบบทดสอบที่สร้างขึ้นเพื่ออ้างอิงผลการสอบไปยังกลุ่มคนที่สอบด้วยแบบทดสอบเดียวกัน แบบทดสอบแบบอิงกลุ่มเป็นเครื่องมือที่สามารถสำรวจครอบคลุมขอบเขตของเนื้อหาได้กว้างขวาง

แบบทดสอบแบบอิงเกณฑ์ เป็นแบบทดสอบที่สร้างขึ้นเพื่อนำผลมาตีความหมายตามมาตรฐานของพฤติกรรมที่แสดงออก (Performance Standard) แต่จะครอบคลุมขอบเขตของเนื้อหาที่จำกัดหรือย่อยกว่าแบบทดสอบแบบอิงกลุ่ม

มาตรฐานของพฤติกรรม คือระดับหรือขอบเขตของงานที่แต่ละคนทำการวัดทำโดยดึงเอาตัวอย่างจากขอบเขตเหล่านั้นมาวัด เมื่อได้ผลมาอย่างไรก็สามารถกล่าวโยงไปถึงขอบเขตนั้น ๆ ได้ (Glaser & Nitko 1971)

จากการถกเถียงกันในเรื่องของอิงเกณฑ์นั้น ทำให้มีผู้เขียนหลายคน เช่น Popham, 1978, Hambleton & Eignor 1979 และ Nitko 1980 ต่างก็ยอมรับว่าน่าจะใช้คำว่า domain referenced แทนคำว่าอิงเกณฑ์ เพราะสามารถจะแก้ปัญหาข้อถกเถียงกันเรื่องจุดตัดของคะแนน หรือมาตรฐานของคะแนนออกไปได้ และจะช่วยลบล้างความหมายของคำว่า Criterion referenced ออกไปได้ แต่อย่างไรก็ตาม Popham และคณะก็ยังคงใช้คำว่า Criterion อยู่ต่อไป

ประเด็นสำคัญประการหนึ่งที่ทำให้การถกเถียงกันเกี่ยวกับการวัดผลแบบอิงเกณฑ์ก็คือความเข้าใจเกี่ยวกับเรื่องความเที่ยงตรงตามเนื้อหาของแบบทดสอบ จากคำนิยามของความเที่ยงตรงชนิดนี้หมายความว่า เป็นความแม่นยำของข้อสอบที่จะอ้างอิงไปสู่ขอบเขตที่จะวัด ซึ่งแบบทดสอบแบบอิงกลุ่มมักถูกโจมตีว่าเป็นแบบทดสอบที่จำกัดในด้านนี้ เพราะข้อสอบแบบอิงกลุ่มเป็นการเปรียบเทียบว่าแต่ละคนทำอะไรได้บ้าง โดยไม่รู้ว่าแต่ละคนนั้นรู้เนื้อหาอะไรบ้าง เป็นต้น ทั้งนี้เพราะข้อสอบแบบนี้มักจะไม่ค่อยอธิบายขอบเขตของเนื้อหาไว้ชัดเจน

แบบทดสอบแบบรอบรู้ แบบทดสอบแบบรอบรู้นั้นมักจะเชกบบระบบการเรียนด้วยตนเอง ซึ่งมีการสร้างบทเรียนเป็นหน่วยหรือเป็นโมดูลการสอน เนื้อหาที่มักจะเรียงตามลำดับของจุดประสงค์การเรียนรู้ ซึ่งผู้เรียนแต่ละคนจะต้องทำงานในหน่วยนั้น ๆ จนกว่าจะผ่านเกณฑ์ขั้น

