

บทที่ 3

แบบทดสอบแบบเลือกตอบและการวิเคราะห์ข้อสอบ

3.1 แบบทดสอบแบบเลือกตอบหรือปรนัย

ข้อสอบแบบเลือกตอบเป็นข้อสอบที่ได้รับความนิยมสูง และใช้แพร่หลายในด้านการเรียนการสอน กล่าวคือนิยมใช้เป็นข้อสอบวัดความสัมฤทธิ์ผลของการเรียน และนิยมใช้สร้างเป็นข้อสอบมาตรฐาน (standardized test)

ข้อสอบแบบเลือกตอบ ยังเป็นแบบข้อสอบที่ใช้วิจัยและพัฒนาก้าวหน้าสามารถวัดพฤติกรรมได้ครบทุกด้าน นับเป็นข้อสอบที่ได้รับการพัฒนามากที่สุด ตลอดจนได้นำเทคโนโลยีมาใช้ประกอบในการตรวจให้คะแนนได้สะดวกและรวดเร็วอีกด้วย

อย่างไรก็ตามข้อสอบแบบเลือกตอบ ยังได้รับการวิพากวิจารณ์มากที่สุด เกี่ยวกับประสิทธิภาพการวัดว่าจะมีประสิทธิภาพสูงกว่าข้อสอบแบบความเรียงหรือไม่ บ้างก็ว่าประสิทธิภาพสูงกว่าบ้างก็ว่ามีประสิทธิภาพด้อยกว่า แต่ที่แน่ๆ ทั้งข้อสอบแบบเลือกตอบและข้อสอบแบบความเรียงมีผลต่อการพัฒนาการเรียนการสอน แต่ลักษณะเด่นของข้อสอบแบบเลือกตอบก็คือเป็นข้อสอบที่นำผลมาวิเคราะห์ได้ว่า ข้อสอบนั้น มีคุณภาพทางการวัดเพียงไรอีกด้วย นับว่าข้อสอบแบบนี้มีความสำคัญต่อกระบวนการวัดผลในปัจจุบันอย่างมาก สมควรที่จะได้ทำความเข้าใจในรายละเอียดต่อไป

3.1.1 ลักษณะของข้อสอบแบบเลือกตอบ

- 1) เลือกตอบบางข้อและไม่ตอบข้อที่เหลือ
- 2) เลือกตอบเฉพาะข้อคำถามที่กำหนดไว้ให้
- 3) มีคำถามและมีคำตอบหลายคำตอบให้จับคู่กับคำถาม
- 4) มีคำถามและมีคำตอบหลายคำตอบให้เลือกคำตอบที่เห็นว่าถูกหนึ่งคำตอบ
- 5) มีคำถามและคำตอบหลายคำตอบ และหลายคำถามจะเลือกคำถามหรือคำตอบใดก็ได้ให้เหมาะสม

ตัวอย่างข้อสอบในกรอบนี้คือข้อสอบแบบเลือกตอบ 1 ข้อ ลักษณะของข้อสอบประกอบด้วย

1. ตัวคำถามหรือตัวต้น (Stem) มี 2 ประเภทคือ
 - (1) ตัวต้นที่เป็นประโยคคำถาม ดังตัวอย่างที่ยกมา
 - (2) ตัวต้นที่เป็นประโยคบอกเล่าที่ขาดข้อความไปส่วนหนึ่ง ซึ่งเราจะต้องเลือกมาต่อให้เป็นข้อความที่ถูกต้อง และเป็นประโยคบอกเล่าที่สมบูรณ์ด้วย
2. คำตอบที่ให้เลือกรหรือตัวเลือก (alternative หรือ choices) คำตอบที่ให้เลือกมีหลายคำตอบแต่คำตอบที่ถูกต้องมีเพียงคำตอบเดียว ตัวเลือกแบ่งออกเป็น 2 ประเภทคือ
 - (1) คำตอบที่ถูกต้อง (keyed response หรือ correct answer) ซึ่งจะมีคำตอบเดียว
 - (2) คำตอบที่ผิดหรือตัวลวง (foils หรือ distractors) ตัวลวงมีหลายตัว ตัวลวงที่ดีและมีประสิทธิภาพจะเป็นตัวที่ถูกเลือกกว่าเป็นคำตอบที่ถูกจำนวนหนึ่งในบรรดานักเรียนที่ตอบทั้งหมดลักษณะของตัวเลือกทั้งตัวที่เป็นคำตอบที่ถูกต้องและคำตอบที่ผิดหรือตัวลวงจะมีลักษณะที่กลมกลืนกัน ซึ่งเราจะได้กล่าวถึงหลักเกณฑ์รายละเอียดในภายหลังเกี่ยวกับการสร้างข้อสอบแบบเลือกตอบ

3.1.2 ข้อบกพร่องหรือจุดอ่อนของข้อสอบแบบเลือกตอบ

ข้อสอบแบบเลือกตอบมีข้อบกพร่องเพียงเล็กน้อยแฝงอยู่ ดังนี้

1. เนื่องจากข้อสอบมีลักษณะจำกัดคำตอบเป็นการทำลายความคิดสร้างสรรค์ (creativity) ถึงแม้ว่าจะมีคำกล่าวประนามเช่นนี้ก็ตาม แต่เมื่อทำการวิจัยแล้วก็ไม่มีข้อมูลสนับสนุนคำกล่าวดังกล่าวทั้งนี้เพราะนักเรียนที่มีความคิดสร้างสรรค์จะทำข้อสอบแบบเลือกตอบได้ดี และข้อสอบแบบเลือกตอบก็ยังไม่ได้พัฒนาไปวัดความคิดสร้างสรรค์

การวัดผลในห้องเรียนนั้น เรามุ่งที่จะวัดความสัมฤทธิ์ผลทางการเรียนเป็นสิ่งสำคัญ มิใช่มุ่งที่จะวัดความคิดสร้างสรรค์ อย่างไรก็ตามลักษณะของข้อสอบเป็นการจำกัดคำตอบ จึงน่าจะพิจารณาได้บ้างว่าไม่ส่งเสริมการวัดความคิดสร้างสรรค์เพราะขาดอิสระในการตอบ

- (1) การสร้างข้อสอบแบบเลือกตอบที่ดี ผู้สร้างข้อสอบต้องมีความเข้าใจหลักการและเทคนิคการเขียนข้อสอบ แล้วทำการฝึกฝนจนมีความชำนาญจึงจะเขียนข้อสอบได้ดี ข้อสอบแบบเลือกตอบจึงเป็นข้อสอบที่ออกยาก

(2) ผู้ออกข้อสอบที่ดีจะต้องเป็นผู้มีความรู้รอบในหลักสูตรและรายละเอียดรายวิชาเป็นอย่างดี จึงไม่เหมาะสำหรับการสอนเป็นทีม หรือผู้เริ่มอาชีพครูใหม่ นอกจากนี้ผู้ออกข้อสอบยังต้องมีทักษะในด้านการใช้ภาษาดี ซึ่งมักจะหาผู้ที่มีความเหมาะสมจริงๆ ได้ยาก

(3) เมื่อเปรียบเทียบกับข้อสอบแบบถูกผิด การตอบข้อสอบแบบเลือกตอบเสียเวลามากกว่าข้อสอบแบบถูกผิดเมื่อเทียบจำนวนข้อเท่ากัน

(4) ข้อสอบแบบเลือกตอบใช้เนื้อที่ของกระดาษมากกว่า ทำให้เสียค่าใช้จ่ายในการจัดทำข้อสอบสูง

(5) เมื่อต้องการให้นักเรียนแสดงออกการใช้ภาษา การเขียนถ้อยคำที่ถูกต้องตามหลักไวยากรณ์ การอธิบายและการให้รายละเอียดอื่นๆ แล้ว ข้อสอบแบบเลือกตอบใช้ไม่ได้ดีเท่าข้อสอบแบบความเรียง

3.1.3 ข้อดีของข้อสอบแบบเลือกตอบ

1. ข้อสอบแบบเลือกตอบ เป็นข้อสอบที่รวบรวมลักษณะข้อดีของข้อสอบแบบอื่น ในด้านของประสิทธิภาพของการวัด กล่าวคือ สามารถวัดพฤติกรรมทางด้านความคิดหรือสติปัญญา (Cognitive-domain) ได้ทุกระดับ ได้แก่ ความรู้ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินค่า ตัวอย่างของข้อสอบต่อไปนี้ จะเป็นตัวอย่างของการวัดพฤติกรรมตามวัตถุประสงค์ของการเรียนการสอน

(1) วัดความรู้ความจำ (Knowledge)

สงครามโลกครั้งที่ 2 เกิดขึ้นในช่วงเวลาใด?

ก. 1914–1917

ข. 1929-1934

ค. 1939-1945

ง. 1961-1955

(2) วัดความเข้าใจ (Comprehension)

คำว่าอธิบายหมายความว่าอะไร?

ก. เล่าเรื่อง

ข. ถามปัญหา

ค. ตัดสินใจจะทำอะไร

ง. การแยกพิจารณาเป็นส่วนๆ

(3) การนำไปใช้ (Application)

การควบคุมการเดินของนาฬิกาที่มีลูกตุ้มควบคุมที่ความยาวของลูกตุ้มที่แกว่งจะให้ช้าหรือเร็วปรับที่ความยาวของลูกตุ้ม สมมติว่าท่านมีนาฬิกาตั้งกล่าวเรือนหนึ่งแต่เดินช้าไป ท่านจะแก้ไขอย่างไร

- ก. ขันลูกตุ้มให้แน่น
- ข. ทำลูกตุ้มให้เบาลง
- ค. ทำลูกตุ้มให้หนักขึ้น
- ง. ขันลูกตุ้มให้สั้นลง

(4) การวิเคราะห์ (Analysis)

“ในสังคมประชาธิปไตยจะเจริญหรือเสื่อมนั้นขึ้นอยู่กับประชากร ที่สามารถคัดเลือกบุคคลที่ฉลาด มีความสามารถ และมีคุณธรรมเพียงไร ใครมาเป็นรัฐบาล” ลักษณะข้อความเช่นนี้จะเป็นลักษณะของวิชาใด หรือเกี่ยวข้องกับวิชาใดมากที่สุด?

- ก. ประวัติศาสตร์
- ข. นิติศาสตร์
- ค. สังคมวิทยา
- ง. รัฐศาสตร์

(5) การสังเคราะห์ (Synthesis)

คณะกรรมการพิจารณาปรับปรุงหลักสูตรสังคมศาสตร์สำหรับนักศึกษาในสายวิทยาศาสตร์และวิศวกรรมศาสตร์ในสถาบันอุดมศึกษาทั้งหมด คณะกรรมการประกอบด้วยผู้เชี่ยวชาญด้านต่างๆ ดังนี้ นักประวัติศาสตร์ 2 คน นักสังคมวิทยา 2 คน นักการศึกษาด้านหลักสูตร 2 คน นักจิตวิทยา 2 คน นักรัฐศาสตร์ 2 คน หัวหน้าภาคหรือหัวหน้าสาขาวิชาทางสังคมศาสตร์ในสถาบัน 3 คนและผู้สอน 3 คน ในการประชุมคณะกรรมการจะกล่าวถึงเรื่องใดมากที่สุด?

- ก. เพิ่มหน่วยกิตวิชาทางสังคมศาสตร์มากขึ้น
- ข. จัดรายวิชาให้ใกล้เคียงกันทุกสถาบัน
- ค. จัดให้มีการสอนโดยใช้กระบวนการค้นคว้ามากขึ้น
- ง. เพิ่มกิจกรรมในการสอนวิชาสังคมศาสตร์สูงๆ เพื่อให้เข้าใจสถานการณ์

จริง

(6) การประเมินค่า (Evaluation)

ลักษณะของพระราชกรณียกิจใดของรัชกาลที่ 4 แสดงให้เห็นว่าพระองค์เป็นบิดาแห่งวิทยาศาสตร์ไทย?

- ก. การสร้างระบบโทรคมนาคม
- ข. ปรับปรุงสังคมให้เหมาะสมโดยตั้งธรรมยุติกนิกาย
- ค. ศึกษาดาราศาสตร์และคำนวณการเกิดสุริยุปราคาได้แม่นยำ
- ง. ทางวางรากฐานทางวิทยาศาสตร์ให้ศึกษาจนถึงปัจจุบัน

จากตัวอย่างของข้อสอบแบบเลือกตอบที่ยกมานี้ เพื่อแสดงให้เห็นว่าข้อสอบแบบเลือกตอบสามารถวัดพฤติกรรมได้มากกว่าข้อสอบแบบอื่นๆ

2. ข้อสอบแบบเลือกตอบเป็นข้อสอบที่สามารถใช้วัดได้ทุกระดับ และสามารถนำไปใช้ได้ทุกสาขาวิชา

3. ข้อสอบแบบเลือกตอบ สามารถสร้างใช้วัดครอบคลุมได้ทุกอย่างเพราะสามารถออกได้มากข้อ โดยใช้ตารางวิเคราะห์หลักสูตร

4. ข้อสอบแบบเลือกตอบ นำไปตรวจให้คะแนนได้รวดเร็วและมีความเป็นปรนัยสูง ผู้ใดตรวจให้คะแนนก็จะได้คะแนนเท่ากันต่างกับข้อสอบแบบความเรียง ซึ่งตรวจให้คะแนนยากซ้ำและมีความเป็นปรนัยต่ำ ผู้ตรวจให้คะแนนต้องมีความรู้ความสามารถสูงจึงจะตรวจให้คะแนนได้ดี ผู้ที่มีความรู้ไม่เพียงพอไม่สามารถตรวจได้คะแนนได้

5. การตรวจข้อสอบแบบเลือกตอบ จะไม่ได้รับอิทธิพลอื่นที่ทำให้คะแนนต้องแปรเปลี่ยนไป เช่น ผู้ตอบเป็นคนตั้งใจเรียน มีความประพฤติดี ผลสอบครั้งที่แล้วได้คะแนนสูง หรือลักษณะอื่นๆ ที่อาจมีผลกระทบต่อ การตรวจให้คะแนน

3.1.4 การสร้างข้อสอบแบบเลือกตอบ

หลักการทั่วไป

(1) ออกข้อสอบตามตารางวิเคราะห์หลักสูตร ว่าจะออกข้อสอบวัดเนื้อหาและพฤติกรรมใดบ้าง

(2) พิจารณาวัดอุปประสงค์ว่าจะวัดโดยใช้คะแนนของข้อสอบ โดยถือตัวข้อสอบเป็นเกณฑ์ที่เรียกว่าอิงเกณฑ์ (Criterion Reference) หรืออิงกลุ่ม (Norm Reference) กล่าวคือถ้าเป็นการวัดเพื่อคัดเลือกนักศึกษาเข้าเรียน มักใช้คะแนนที่ได้จากการสอบมาตัดสินเรียกว่าอิงเกณฑ์หรือถือตัวข้อสอบเป็นเกณฑ์ (Test-Criterion) แต่การวัดในชั้นเรียนหรือวัดเพื่อประเมินผลการเรียนนั้น เราจะต้องปรับคะแนนให้เป็นคะแนนมาตรฐานตัดเกรดเช่นนี้เรียกว่าอิงกลุ่ม และเป็นไปตามวัตถุประสงค์ของการวัดลักษณะดังกล่าวนี้จะช่วยในการสร้างความเที่ยงตรงและความเชื่อมั่น

(3) ใช้กระดาษแข็งหรือบัตรขนาด 3 X 5 นิ้ว หรือขนาดอื่นๆ ตามความเหมาะสม บัตรหนึ่งเขียนข้อสอบหนึ่งข้อ เพื่อสะดวกแก่การแก้ไขปรับปรุงข้อสอบควรเขียนข้อสอบด้วย ดินสอ เพราะลบแก้ไขได้สะดวก

การใช้บัตรจัดเตรียมข้อสอบนี้ สามารถจัดเป็นชุดๆ และเก็บไว้เมื่อใช้ข้อสอบแล้ว นำผลการสอบมาวิเคราะห์หาคุณสมบัติของข้อสอบแต่ละข้อแล้วคัดเลือกข้อสอบดีเอาไว้ใช้ต่อไป ทั้งนี้การสอบนั้นจะต้องเป็นข้อสอบที่ห้ามผู้สอบนำออกจากข้อสอบ

การได้ข้อสอบดีแต่ละเรื่องมากข้อ ก็สามารถจัดตั้งเป็นคลังข้อสอบได้ เพราะเมื่อมี ข้อสอบมากพอแล้วก็ไม่จำเป็นที่จะออกข้อสอบใหม่เพียงแต่เลือกข้อสอบที่เหมาะสม ในแต่ละเรื่อง ที่ต้องการจะวัดตามตารางวิเคราะห์ข้อสอบ หากไม่เพียงพอก็ออกเพิ่มอีกก็ช่วยให้เบาแรงผู้ออก ข้อสอบ หรือไม่ต้องออกเพิ่มเมื่อมีข้อสอบเพียงพอ ทั้งนี้เพื่อแก้ไขข้อเสียของข้อสอบที่ว่าออกยาก เมื่อ ออกยากแล้วก็ควรเก็บไว้ใช้ในโอกาสต่อไปจึงจะเป็นประโยชน์

(4) อ่านและพิจารณาข้อสอบทั้งข้อ พิจารณาว่าข้อสอบนั้นจะต้องวัดเพียงหนึ่ง ความคิดหรือถามเพียงข้อมูลเดียว ตัวเลือกที่เป็นคำตอบผิด (Fails) มีลักษณะกลมกลืนกับตัวคำถาม แต่ไม่ใช่คำตอบที่ถูกลำดับที่สอง คำตอบที่ถูกและคำตอบที่ผิดมีความกลมกลืนกัน เช่น ถ้าถามถึง บุคคลคำตอบก็ต้องเป็นชื่อบุคคลทั้งหมด คำตอบที่เป็นวลีที่กล่าวถึงเรื่องเดียวกันเช่น

นายกรัฐมนตรีคนแรกของไทย คือใคร?

ก. พันเอกพระยาพหลพลพยุหเสนา

ข. จอมพล ป. พิบูลสงคราม

ค. พระยามโนปกรณนิติธาดา

ง. นายปรีดี พนมยงค์

ราชธานีของอาณาจักรทวารวดีอยู่ที่ใด?

ก. ลพบุรี

ข. นครปฐม

ค. สุพรรณบุรี

ง. กาญจนบุรี

(5) ข้อสอบที่ดีทุกข้อจะต้องมีข้อความชัดเจน ถูกต้องตามหลักภาษา ถ้อยคำที่ไม่ เกี่ยวข้องกับคำตอบต้องตัดออก และถ้าไม่จำเป็นที่จะต้องใช้คำศัพท์ทางวิชาการให้ตัดออก ไม่ควร ใช้ประโยคที่ซับซ้อนทั้งที่ถูกหลักภาษาแต่เข้าใจยากก็ไม่ควรใช้

เพื่อให้เข้าใจหลักเกณฑ์และรายละเอียดในการสร้างข้อสอบแบบเลือกตอบสมควรพิจารณาแต่ละองค์ประกอบของข้อสอบแบบเลือกตอบได้แก่ การเขียนตัวคำถาม การเขียนตัวคำตอบ เลือกคำตอบที่ถูกและตัวลวงหรือคำตอบที่ผิดดังต่อไปนี้

3.1.5 การเขียนตัวคำถาม (Stem)

ตัวคำถามจะต้องทำหน้าที่พื้นฐานอย่างน้อย 2 ประการคือเสนอปัญหาและเสนอกรอบอ้างอิงที่เหมาะสม ซึ่งเมื่ออ่านตัวคำถามแล้วมีลักษณะชี้ชวนให้คิดถึงคำตอบอย่างกว้างๆ ได้ ตัวคำถามเป็นตัวบ่งชี้คำตอบซึ่งคำตอบจะต้องผสมกลมกลืนกับตัวคำถาม ถ้าเป็นตัวคำถามที่ใช้วัดพฤติกรรมสูงขึ้น การหาคำตอบที่ถูกจะยากขึ้นเพราะไม่ใช่คำตอบปกติต้องใช้ความคิดก่อนเลือกคำตอบตามสมควร เช่น การวัดความสามารถด้านการวิเคราะห์ การสังเคราะห์ และการประเมินค่า เป็นต้น

ไม่ดี

ถ้าเทอร์ชานอลลงบนฝ่ามือ เราจะรู้สึกเย็นเนื่องจากของเหลวระเหยได้ง่ายและในขณะที่ที่ของเหลวระเหย พลังงานจลน์เฉลี่ยของเหลวที่เหลือ

- ก. ลดลง ของเหลวให้พลังงานแก่สิ่งแวดล้อม
- ข. เพิ่มขึ้น ของเหลวถูกพลังงานแก่สิ่งแวดล้อม
- ค. เพิ่มขึ้น ของเหลวให้พลังงานแก่สิ่งแวดล้อม
- ง. ลดลง ของเหลวดูดพลังงานจากสิ่งแวดล้อม

ดีขึ้น

ถ้าเทอร์ชานอลลงบนฝ่ามือ เทอร์ชานอลจะระเหย อยากทราบว่าพลังงานจลน์เฉลี่ยของเทอร์ชานอลส่วนที่ยังไม่ได้ระเหยจะเป็นอย่างไร เพราะเหตุใด?

หรือ

ในขณะที่เทอร์ชานอลระเหย พลังงานจลน์เฉลี่ยของเทอร์ชานอลที่ยังไม่ระเหย จะเป็นอย่างไร?

- ก. ลดลง ของเหลวให้พลังงานแก่สิ่งแวดล้อม
- ข. เพิ่มขึ้น ของเหลวถูกพลังงานแก่สิ่งแวดล้อม
- ค. เพิ่มขึ้น ของเหลวให้พลังงานแก่สิ่งแวดล้อม
- ง. ลดลง ของเหลวดูดพลังงานจากสิ่งแวดล้อม

ไม่ตี

กำหนดให้สารละลายต่อไปนี้

- ก. สารละลาย HCl 1 mol/kg ใน CH₃COOH และสารละลาย HCl 1 mol/kg ในน้ำ
- ข. สารละลาย HNO₃ 1 mol/kg ใน CH₃COOH และสารละลาย HNO₃ 1 mol/kg ในเอทานอล
- ค. สารละลาย HCl 1 mol/kg ในน้ำ และสารละลาย HNO₃ 1 mol/kg ใน CH₃OOH
- ง. สารละลาย HCl 1 mol/kg ใน CH₃COOH และสารละลาย HNO₃ 1 mol/kg ในน้ำ CH₃OOH

สารละลายข้อใดที่มีจุดเดือดเท่ากัน

1. ก 2. ข 3. ค 4. ง

ตีขึ้น

สารละลายในข้อใดมีจุดเดือดเท่ากัน?

- ก. สารละลาย HCl 1 mol/kg ใน CH₃COOH และสารละลาย HCl 1 mol/kg ในน้ำ
- ข. สารละลาย HNO₃ 1 mol/kg ใน CH₃OOH และสารละลาย HNO₃ 1 mol/kg ในเอทานอล
- ค. สารละลาย HCl 1 mol/kg ในน้ำ และสารละลาย HNO₃ 1 mol/kg ใน CH₃OOH
- ง. สารละลาย HCl 1 mol/kg ใน CH₃OOH และสารละลาย HNO₃ 1 mol/kg ในน้ำ CH₃OOH

ตัวคำถามอาจเขียนได้ 3 ลักษณะ คือ

- ก. เป็นประโยคคำถาม ดังตัวอย่างที่แสดงมาแล้วเพราะช่วยนักเรียนเกิดความคิดที่จะแก้ปัญหาเพราะมนุษย์จะคิดเมื่อมีปัญหา หรือตั้งปัญหาเพื่อที่จะคิด มากกว่าคิดโดยไม่มีปัญหา
- ประโยคคำถาม อาจเป็นแบบคำถามและแบบคำถามปฏิเสธ

ข. เป็นประโยชน์พิเศษ แบบนี้เป็นที่นิยมกันมากพอสมควรแต่ใช้เมื่อเห็นว่าเหมาะสมเท่านั้น เช่น “ประโยชน์หรือข้อดีของข้อสอบแบบเลือกตอบ มีหลายประการตามที่ระบุไว้ต่อไปนี้ให้ท่านเลือกข้อดีที่สำคัญที่สุด”

“จงเลือกวิธีการต้มน้ำร้อนเพื่อใช้ภายในบ้านที่ประหยัดที่สุด”

ประโยชน์สมบูรณ์อาจเป็นประโยชน์นอกเล่มและประโยชน์พิเศษ

ค. เป็นประโยชน์ที่ไม่สมบูรณ์ ขาดหายไปส่วนหนึ่งแล้วให้นำตัวเลือกที่เป็นคำตอบมาต่อจึงจะเป็นประโยชน์สมบูรณ์ เช่น

“ประโยชน์หรือข้อดีของข้อสอบแบบเลือกตอบ ที่นับว่าสำคัญที่สุดคือ -----”

การเลือกคำตอบหรือการสร้างคำตอบให้เลือกมีได้ 3 ลักษณะคือ ให้เลือกคำตอบที่ถูกต้อง (correct answer) กล่าวคือ มีคำตอบที่ถูกต้องเพียงคำตอบเดียวเท่านั้น นอกนั้นเป็นคำตอบผิด ดังนี้เป็นลักษณะที่หนึ่ง อีกลักษณะหนึ่งเป็นการเลือกคำตอบที่ดีที่สุด (Best answer) คำตอบที่กำหนดให้เลือกทั้งหมดเป็นคำตอบที่ถูกทุกคำตอบ แต่มีคำตอบที่ถูกต้องสมบูรณ์ที่สุดเพียงคำตอบเดียว นักเรียนเลือกตอบนี้จึงจะได้คะแนน

มีคำตอบอีกแบบหนึ่งซึ่งเป็นลักษณะที่สาม คือ ผู้ตอบต้องเลือกคำตอบที่ไม่ถูกต้อง (incorrect answer) คำตอบที่ให้เลือกตอบนั้นมีคำตอบที่ผิดเพียงคำตอบเดียว นอกนั้นเป็นคำตอบที่ถูกทั้งหมด

ตัวเลือกหรือคำตอบให้เลือกจะต้องมีลักษณะสอดคล้องกับตัวคำถาม

ตัวอย่างสำหรับคำถามที่ต้องการคำตอบที่ถูกต้อง (correct answer)

ข้อใดเป็นเหตุผลที่เป็นไปได้ สำหรับการเพิ่มอัตราค่าส่งไปรษณีย์ประเภทด่วนพิเศษ?

ก. รัฐบาลต้องการเงินมากขึ้นเพราะงบประมาณปิดที่งบประมาณไม่ลง

ข. เงินเดือนของพนักงานไปรษณีย์เพิ่มขึ้นหลังจากเดินขบวนเรียกร้อง

ค. จำนวนไปรษณีย์ด่วนพิเศษมีมากขึ้น

ง. จดหมายหนักมากขึ้นในปัจจุบัน

ตัวอย่างสำหรับคำถามประเภทให้เลือกคำตอบที่ดีที่สุด (Best answer)

ข้อใดเป็นเหตุผลหลักในการเพิ่มอัตราค่าส่งไปรษณีย์ประเภทด่วนพิเศษ?

ก. ค่าพิมพ์แสตมป์ไปรษณีย์เพิ่มขึ้น

ข. ปริมาณจดหมายเพิ่มขึ้นมาก

ค. เพิ่มเงินเดือนพนักงานมากขึ้น

ง. ขยายที่ทำการไปรษณีย์เพิ่มขึ้น

ตัวอย่างสำหรับคำถามที่ให้เลือกคำตอบที่ไม่ถูกต้อง

ข้อใด ไม่ใช่ เหตุผลในการเพิ่มอัตราค่าส่งไปรษณีย์ประเภทด่วนพิเศษ?

- ก. ขยายที่ทำการไปรษณีย์เพิ่มขึ้น
- ข. เพิ่มเงินเดือนพนักงานมากขึ้น
- ค. การปฏิรูปการปกครองแผ่นดิน
- ง. ค่าพิมพ์แสตมป์ไปรษณีย์เพิ่มขึ้น

ตัวคำถามที่เป็นปฏิเสธ จะต้องเน้นคำปฏิเสธโดยการขีดเส้นใต้หรือพิมพ์ตัวหนา

ตัวคำถามอาจให้ผู้ตอบจัดลำดับตามหลักเกณฑ์ใดหลักเกณฑ์หนึ่ง ตัวคำถามจะมีรายการของสิ่งต่างๆ และให้จัดลำดับของสิ่งที่กำหนดนั้นและตัวคำตอบจะเป็นลำดับของสิ่งของที่จัดไว้ให้เลือก โดยที่มีคำตอบที่ถูกต้องเพียงคำตอบเดียว

ตัวอย่าง

สินแร่, คอรัต, ไมคา และเพชร ข้อใดจัดลำดับจากแข็งสุดถึงอ่อนสุด?

- ก. คอรัต, ไมคา, เพชร
- ข. เพชร, คอรัต, ไมคา
- ค. คอรัต, เพชร, ไมคา
- ง. เพชร, ไมคา, คอรัต

ข้อสังเกตเกี่ยวกับข้อสอบแบบเลือกตอบ เมื่อนักเรียนเลือกคำตอบแต่ละข้อมีลักษณะเหมือนการคิดทำข้อสอบแบบถูกผิดแต่ละตัวคำตอบ

ข้อสอบที่มีคำตอบให้เลือกนี้เฉพาะตัวเลือกยากที่จะวางขอบเขตอ้างอิงในกรณีที่จะวัดความรู้ ความจำ แม้จะทำให้มีขอบเขตอ้างอิงได้บางข้อ แต่ก็ไม่สามารถทำได้ทั้งชุดข้อสอบ

3.1.6 การเขียนตัวเลือก (Alternatives)

แม้ว่าตัวคำถามจะทำหน้าที่กำหนดขอบเขตอ้างอิง (frame of reference) ตั้งได้กล่าวมาแล้ว แต่ไม่เป็นการเพียงพอในการกำหนดขอบเขตอ้างอิง ตัวเลือกเป็นส่วนประกอบสำคัญที่จะช่วยกำหนดขอบเขตอ้างอิงและตัวเลือกยังช่วยให้ข้อสอบมีลักษณะน่าคิดน่าตอบด้วย

ตัวเลือกจะต้องเป็นคำตอบที่เข้ากันได้อย่างเหมาะสมกับตัวคำถามและภาษาที่ใช้ จะต้องถูกต้องตามหลักของไวยากรณ์ ในเชิงภาษาตัวเลือกทุกตัวจะต้องเป็นคำตอบที่น่าสนใจ ความยาวของตัวเลือกควรใกล้เคียงกันหรือถ้าตัวเลือกมีความยาวแตกต่างกัน ก็ให้เรียกตัวเลือกจากสั้นที่สุดตามลำดับไปถึงยาวที่สุด หรือจะเรียงจากตัวเลือกที่ยาวที่สุดไปหาตัวเลือกที่สั้นที่สุดก็ได้

ตัวเลือกที่มีข้อความว่า “ถูกทุกข้อ” หรือ “ถูกข้อ—และข้อ—” หรือ “ไม่มีข้อถูก” หรือ “ยังสรุปไม่ได้” หรือข้อความอื่นทำนองเดียวกัน ข้อความดังกล่าวนี้จะเป็นข้อสุดท้ายของ

ตัวเลือกเสมอ ถ้า คำตอบดังกล่าวเป็นคำตอบที่ถูกแล้ว ก็จะเป็นคำตอบที่ไม่เหมาะสม เพราะว่าคุณลักษณะของคำตอบมีแนวโน้มที่จะให้เลือกคำตอบนี้ แต่ถ้าเป็นคำตอบผิดและคำตอบที่ถูกต้องเป็นคำตอบอื่นแล้วก็ตาม ความรู้สึกของผู้ตอบที่เลือก “ถูกทุกข้อ” อีกด้วย อย่างไรก็ตามก็ยังไม่มียุทธวิธีที่แน่นอนเพราะยังมีนักวิชาการวัดผลที่ใช้ข้อบ่งชี้ แต่ก็นิยมใช้กับวิชาที่มีการคำนวณหาคำตอบในวิชาคณิตศาสตร์และวิทยาศาสตร์ เป็นต้น เพราะต้องการคำตอบที่มีความแม่นยำในตัวเอง ผู้คำนวณพลาดก็จะหาคำตอบไม่ได้ ก็อาจเลือกคำตอบ “ไม่มีข้อใดถูกต้อง” ก็นับว่าสมควร แต่ถ้าจะใช้กับวิชาอื่นที่ไม่ใช่คำถามที่เป็นวิชาคำนวณแล้ว ไม่ควรใช้ตัวเลือกดังกล่าว

(1) ตัวเลือกที่เป็นคำตอบที่ต้องการหรือคำตอบกุญแจ (the keyed response)

คำตอบที่ต้องการจะต้องมีลักษณะกลมกลืน หรือเป็นเพียงคำตอบที่ดีกว่าเท่านั้น ไม่สมควรที่จะต้องเขียนให้ดีที่สุดจนมีลักษณะเด่นมากกว่าคำตอบอื่นๆ ที่ว่าเป็นคำตอบที่ดีกว่าก็คือลักษณะที่ยอมรับได้ว่าเป็นคำตอบที่ถูก

คำตอบที่มีผู้มีความรู้ในเนื้อหาวิชานั้นตอบไม่ได้ หรือผู้ออกข้อสอบเองก็ตอบไม่ได้ถูกต้อง คุณค่าที่ทำไว้แล้ว นับว่าเป็นคำตอบที่ใช้ไม่ได้ ยกเว้น สำหรับคำตอบที่เป็น วัน เดือน ปี หรือมีจำนวนจำเพาะซึ่งผู้เรียนหรือผู้สอนเพิ่งเรียนผ่านไปหยกๆ ก็ยอมรับได้และทำข้อสอบนั้นได้ แต่เมื่อนานไปก็ยอมรับที่จะลืม

คำตอบถูกที่ดีจะต้องไม่มีลักษณะที่แนะให้ตอบคำตอบนั้น ในบางครั้งผู้ออกข้อสอบเขียนตัวคำตอบถูกสอดคล้องกับตัวคำถามเพียงคำตอบเดียวหรือใช้คำบางคำที่สอดคล้องกับตัวคำถาม ยกตัวอย่าง เช่น

ไม่ดี

หลังสงครามโลกครั้งที่ 2 ประชาชนส่วนใหญ่ของโลกได้พยายามรวมกันเป็นเพื่อสันติภาพของโลก องค์การนี้เรียกว่าอะไร

ก. สมาพันธ์

ข. พันธภาพที่ยิ่งใหญ่

ค. สัญญาระหว่างชาติ

ง. สหประชาชาติ

จ. ศาลโลก

ดีขึ้น

หลังสงครามโลกครั้งที่ 2 ประเทศต่างๆ ได้รวมกันเพื่อสร้างสันติภาพ องค์การนี้เรียกว่าอะไร?

- ก. สมาพันธ์
- ข. ศาลโลก
- ค. สหประชาชาติ
- ง. พันธภาพที่ยิ่งใหญ่
- จ. สนธิสัญญาระหว่างชาติ

ถ้าคำตอบที่ถูกมักอยู่ในข้อที่ซ้ำกัน เช่น ข้อสอบทั้งชุด คำตอบข้อ ข. ถูกเป็นส่วนใหญ่ อ่านก็ข้อก็พบว่าคำตอบข้อ ข. ถูกเสมอ จะส่งเสริมข้ออื่นๆ มีแนวโน้มว่าคำตอบข้อ ข. ถูก และช่วยให้ผู้ตอบเดาคำตอบเป็นข้อ ข. สำหรับข้อที่ไม่เข้าใจและทำไม่ได้ ดังนั้น คำตอบที่ถูกของข้อสอบแต่ละข้อควรให้คละกัน และอื่นๆ มีโอกาสถูกอย่างใกล้เคียงกัน วิธีการที่ช่วยกระจายคำตอบที่ถูกกระจายออกไป ก็คือ การเรียงลำดับคำตอบที่ให้เลือกให้เป็นระเบียบ ทำให้คำตอบที่ถูกกระจายออกไปเป็นข้ออื่นๆ

3.2 การประเมินหรือการวิเคราะห์คุณภาพของแบบทดสอบ

การประเมินคุณภาพของแบบทดสอบ หรือเรียกอีกอย่างหนึ่งว่า การวิเคราะห์ข้อสอบ (Item Analysis or Test analysis) หมายถึง การตรวจสอบคุณภาพของแบบทดสอบที่สร้างขึ้นว่ามีคุณภาพดีเพียงใด ทั้งลักษณะเป็นรายข้อและทั้งฉบับ ถ้าข้อสอบข้อใดหรือฉบับใดมีคุณภาพดีก็ควรนำไปใช้ แต่ถ้าบกพร่องก็ควรปรับปรุงแก้ไข การทำเช่นนี้จะได้แบบทดสอบที่มีคุณภาพดีไปทดสอบกับนักเรียน ช่วยให้การวัดผลและประเมินผลมีประสิทธิภาพดียิ่งขึ้น

ชนิดของแบบทดสอบที่นิยมนำมาวิเคราะห์มากที่สุด ได้แก่ แบบทดสอบชนิดเลือกตอบ ซึ่งกำหนดว่าตอบถูกได้ 1 คะแนน ตอบผิดได้ 0 คะแนน หรือที่เรียกว่าระบบถูก 1 ผิด 0 การที่นิยมนำข้อสอบแบบเลือกตอบมาวิเคราะห์ มีเหตุผลดังนี้

1. ปัจจุบันนิยมใช้ข้อสอบแบบเลือกตอบเพราะสามารถใช้กับผู้เข้าสอบจำนวนมากๆ
2. เป็นข้อสอบที่สามารถวัดได้ครอบคลุมหลักสูตร และวัดพฤติกรรมต่างๆ ด้านพุทธิพิสัยได้ตามที่ต้องการ
3. สามารถนำมาวิเคราะห์และปรับปรุงแก้ไขเพื่อให้มีคุณภาพดี ทำได้สะดวกรวดเร็ว และสามารถใช้อุปกรณ์แทนแรงงานคน
4. เป็นข้อสอบที่สามารถเก็บไว้ใช้ได้หลายๆ ครั้ง คุ่มค่ากับการลงทุน

การประเมินคุณภาพของแบบทดสอบ จะแยกกล่าวเป็น 2 แบบที่สำคัญ คือ

3.2.1 หลักการประเมินผลแบบอิงกลุ่มและอิงเกณฑ์

3.2.2 การวิเคราะห์ข้อสอบเป็นรายข้อ

3.2.1 หลักการประเมินผลแบบอิงกลุ่มและอิงเกณฑ์

การประเมินผลแบบอิงกลุ่ม (Norm Reference) หมายถึง การประเมินที่จะมุ่งนำผลการประเมินมาจำแนกนักเรียนออกตามความสามารถ โดยพิจารณาจากการเปรียบเทียบผลการประเมินของนักเรียนแต่ละคนกับกลุ่มนักเรียนด้วยกัน ซึ่งการตีความหมายผลการประเมินในรูปแบบนี้เรียกว่า การตีความหมายแบบอิงกลุ่ม โดยมีแนวความคิดว่าในการจัดการเรียนการสอนนักเรียนย่อมมีความแตกต่างเป็นรายบุคคล นั่นคือ จะทราบว่านักเรียนแต่ละคนมีความสามารถมากหรือน้อยกว่านักเรียนคนอื่นๆ ในกลุ่มเดียวกัน

การประเมินผลแบบอิงเกณฑ์ (Criterion Reference) หมายถึง การประเมินที่มุ่งนำเอาผลการเรียนของนักเรียนแต่ละคนมาเปรียบเทียบกับเกณฑ์ (Criteria) ที่ตั้งขึ้น โดยไม่ต้องเปรียบเทียบกับนักเรียนคนอื่นๆ หรือกล่าวได้ว่า เป็นการประเมินที่ต้องการทราบสถานภาพของบุคคล โดยอาศัยเกณฑ์ที่กำหนดไว้ในจุดมุ่งหมายเป็นหลักการตีความหมายในลักษณะดังกล่าวนี้เรียกว่า การตีความหมายแบบอิงเกณฑ์ โดยมีแนวความคิดว่าในการจัดการเรียนการสอน ควรจะให้นักเรียนเรียนอย่างรอบรู้ (Master Learning)

การประเมินผลแบบอิงกลุ่มและอิงเกณฑ์มีความแตกต่างกัน สามารถนำมาเปรียบเทียบกันได้ดังนี้

ตารางที่ 3.1 ความแตกต่างระหว่างการประเมินผลแบบอิงกลุ่มกับอิงเกณฑ์

สิ่งที่เปรียบเทียบ	การประเมินผลแบบอิงกลุ่ม (NR)	การประเมินผลแบบอิงเกณฑ์ (CR)
1. รากฐานของแนวความคิด	ยึดแนวความคิดของทฤษฎีการวัดผลแบบเดิมซึ่งเป็นทฤษฎีเกี่ยวกับการวัดความถนัดที่มุ่งวัดความแตกต่างระหว่างบุคคลเป็นสำคัญ	ยึดจุดประสงค์การเรียนรู้ เน้นการสอนเป็นรายบุคคลและจัดบทเรียน โปรแกรมเป็นสำคัญ
2. จุดมุ่งหมาย	เพื่อเปรียบเทียบความสามารถของแต่ละคนกับความสามารถของคนอื่นๆ ในกลุ่มเดียวกัน	เพื่อเปรียบเทียบความสามารถของแต่ละคนกับเกณฑ์ ซึ่งก็คือ จุดประสงค์การเรียนรู้
3. กระบวนการสร้าง	ทั้งสองประเภทจะไม่แตกต่างกันในเรื่องการเขียนข้อสอบ และรูปแบบของข้อสอบ แต่จะแตกต่างกันในเรื่องความคิดและวิธีวิเคราะห์ข้อสอบเพื่อหาอำนาจจำแนกกับค่าความยากกล่าวคือ	

สิ่งที่เปรียบเทียบ	การประเมินผลแบบอิงกลุ่ม (NR)	การประเมินผลแบบอิงเกณฑ์ (CR)
4. โครงสร้างของการวัดผล	เน้นความครอบคลุมและการเป็นตัวแทนของความรู้	เน้นการวัดตรงตรงจุดประสงค์ของการเรียน
5. การนำไปใช้	เหมาะสำหรับการสอบคัดเลือกหรือแบ่งกลุ่ม	เหมาะสำหรับการสอบระหว่างเรียน (Formative) หรือสอบเพื่อวินิจฉัย
6. การแปลความหมาย	แปลความหมายของคะแนนโดยแปลงคะแนนที่นักเรียนสอบได้ให้เป็นตำแหน่งเปอร์เซ็นต์ไทล์ หรือ คะแนนมาตรฐานต่างๆ เช่น Z-score, T-score หรือ Stanine แล้วแปลความหมายจากคะแนนมาตรฐานนั้นๆ	แปลความหมายของคะแนนในรูปของการรอบรู้กับไม่รอบรู้

3.2.2 การวิเคราะห์ข้อสอบเป็นรายข้อ

ได้แก่ การหาค่าความยากและค่าอำนาจจำแนกของข้อสอบซึ่งมีความหมาย ดังนี้ ความยากของข้อสอบ (Difficulty) หมายถึง อัตราส่วนของจำนวนคนตอบถูกกับจำนวนคนทั้งหมด การ ซึ่งแบบที่นิยมใช้กันอย่างกว้างขวางนั้นคือ การวิเคราะห์ข้อสอบเป็น รายข้อแบบทุกตัวเลือก ซึ่งได้แก่การหาค่าความยากและค่าอำนาจจำแนกของข้อสอบทั้งตัวเลือกที่เป็นตัวถูก และ ตัวลวง และในค่าที่สำคัญในการวิเคราะห์ข้อสอบนั้นมีดังนี้

1. ความยากของข้อสอบ (Difficulty) หมายถึง อัตราส่วนหรือร้อยละระหว่างจำนวนคนตอบถูกกับจำนวนคนทั้งหมด

2. อำนาจจำแนกของข้อสอบ (Discrimination) หมายถึง ผลต่างระหว่าง อัตราส่วนหรือร้อยละของจำนวนคนในกลุ่มสูงตอบถูก กับอัตราส่วนหรือร้อยละของจำนวนคนในกลุ่มต่ำตอบถูก

ซึ่งการหาค่าความยาก และค่าอำนาจจำแนกของข้อสอบ มีลำดับขั้น ดังนี้

ขั้นที่ 1 นำข้อสอบที่สร้างเรียบร้อยแล้ว (ควรจะสร้างตามตารางวิเคราะห์หลักสูตร ด้านพุทธิพิสัย) ไปทดสอบกับนักเรียนจำนวนหนึ่ง แล้วนำกระดาษคำตอบมาตรวจให้คะแนน ตอบถูกให้ 1 คะแนน ตอบผิดให้ 0 คะแนน

ขั้นที่ 2 นำกระดาษคำตอบมาจัดเรียงจากคะแนนรวมสูงสุดไปหาต่ำสุด (กระดาษคำตอบที่ได้คะแนนรวมเท่ากันจะเรียงติดกัน โดยเรียงแผ่นใดก่อนก็ได้)

ขั้นที่ 3 นับกระดาษคำตอบจากข้างบนลงมาจำนวนหนึ่ง เรียกว่า กลุ่มสูง (High Group) ใช้อักษรย่อว่า ส หรือ H และนับกระดาษคำตอบจากข้างล่างขึ้นมา ให้มีจำนวน

เท่ากับกลุ่มสูง เรียกว่า กลุ่มต่ำ (Low Group) ใช้อักษรย่อว่า ต หรือ L ส่วนที่เหลือ (ถ้ามี) เรียกว่า กลุ่มกลางซึ่งโดยวิธีนี้ไม่ต้องนำมาใช้

ขั้นที่ 4 นำกระดาษคำตอบในกลุ่มสูง (ส1 – ส15) ไปลงรอยขีด (Tally) ในแบบฟอร์มทุกตัวเลือกในแต่ละข้อ (ทั้งตัวถูกและตัวลวง) เพื่อจะดูความถี่ในแต่ละตัวเลือกของแต่ละข้อว่ามีคนในกลุ่มสูงเลือกตอบกี่คน ดังตัวอย่าง (สมมติข้อสอบมี 50 ข้อ)

ตารางที่ 3.2 การวิเคราะห์ข้อสอบเป็นรายข้อ แบบทุกตัวเลือก วิชา..... ชั้น..... กลุ่มสูง (ส1-ส15)

ข้อที่ คนที่	1					2					50				
	ก	ข	ค	ง	จ	ก	ข	ค	ง	จ	ก	ข	ค	ง	จ
ส1		/				/									
ส2		/				/									
ส3		/				/									
ส4		/				/									
ส5	/					/									
ส6		/				/									
ส7	/					/				/					
ส8					/	/									
ส9		/						/							
ส10		/						/							
ส11			/			/									
ส12		/						/							
ส13		/				/									
ส14					/			/							
ส15		/				/									
รวม (H)	2	10	1	0	2										

สำหรับกลุ่มต่ำทำในทำนองเดียวกัน

ขั้นที่ 5 นำค่ารวม (H) และรวม (L) ของแต่ละตัวเลือกไปคำนวณหาค่าความยาก (p) และค่าอำนาจจำแนก (r) โดยใช้สูตรดังนี้

$$\text{ตัวถูก} \quad p = \frac{H+L}{2N}, \quad r = \frac{H-L}{N}$$

$$\text{ตัวลง} \quad p = \frac{H+L}{2N}, \quad r = \frac{L-H}{N}$$

ตัวถูก			ตัวลง		
p	แทน	ค่าความยากของข้อสอบ	p	แทน	ค่าความยากของข้อสอบ
r	แทน	ค่าอำนาจจำแนกของข้อสอบ	r	แทน	ค่าอำนาจจำแนกของข้อสอบ
H	แทน	จำนวนคนในกลุ่มสูงตอบถูก	H	แทน	จำนวนคนในกลุ่มสูงตอบถูก
L	แทน	จำนวนคนในกลุ่มต่ำตอบถูก	L	แทน	จำนวนคนในกลุ่มต่ำตอบถูก
N	แทน	จำนวนคนทั้งหมดในกลุ่มใดกลุ่มหนึ่ง	N	แทน	จำนวนคนทั้งหมดในกลุ่มใดกลุ่มหนึ่ง

ข้อสังเกต ตัวถูก ถ้า $H > L$ ค่า r จะเป็นบวก

ตัวลง ถ้า $L > H$ ค่า r จะเป็นบวก

ตัวอย่าง การวิเคราะห์ข้อสอบเป็นรายข้อแบบทุกตัวเลือก

ตาราง 3.3 แสดงค่า p, r แบบทุกตัวเลือก

ข้อที่	ตัวเลือก	H	L	p	r	ผลการพิจารณา	
						ค่า p	ค่า r
1	ก	4	2	.20	.13	ใช้ได้	ใช้ได้
	ข	2	10	.40	.53	ปานกลาง	ค่อนข้างสูง
	ค	0	1	.03	-.07	ใช้ไม่ได้	ใช้ไม่ได้
	ง	7	0	.23	.46	ใช้ได้	พอใช้
	จ	2	2	.13	.00	ใช้ได้	ใช้ไม่ได้
2	ก	2	8				
	ข	7	2				
	ค	1	0				
	ง	3	2				
	จ	2	3				

เนื่องจาก ค่า p , r ที่เป็นตัวถูก และตัวลง มีหลักในการพิจารณาแตกต่างกัน ดังนั้น เพื่อความสะดวกจึงได้วางเกณฑ์การพิจารณาไว้เป็นกลางๆ ดังในตารางดังต่อไปนี้ (ตัวเลขที่พิมพ์ตัวหนา แสดงว่าเป็นตัวเลือกที่เหมาะสมหรือใช้ได้)

1. เกณฑ์การพิจารณาค่า p , r ตัวถูก

ค่า p		ค่า r			
.00 ถึง .09	ยากมาก	ไม่มีคุณภาพ	ค่าลบ	ใช้ไม่ได้	ไม่มีคุณภาพ
.10 ถึง .19	ยาก	ไม่มีคุณภาพ	.00	ไม่มีอำนาจจำแนก	
.20 ถึง .39	ค่อนข้างยาก	ไม่มีคุณภาพ	.01 ถึง .09	ต่ำ	
.40 ถึง .60	ปานกลาง	มีคุณภาพ	.10 ถึง .19	ค่อนข้างต่ำ	
.61 ถึง .80	ค่อนข้างง่าย	มีคุณภาพ	.20 ถึง .40	ปานกลาง	
.81 ถึง .90	ง่าย	มีคุณภาพ	.41 ถึง .60	ค่อนข้างสูง	
.91 ถึง 1.00	ง่ายมาก	ไม่มีคุณภาพ	.61 ถึง 1.00	สูง	

2. เกณฑ์การพิจารณาค่า p , r ตัวลง

ค่า p		ค่า r			
.00 ถึง .04	ใช้ไม่ได้	ไม่มีคุณภาพ	ค่าลบ	ใช้ไม่ได้	ไม่มีคุณภาพ
.05 ถึง .09	พอใช้	ไม่มีคุณภาพ	.00 ถึง .04	ใช้ไม่ได้	
.10 ถึง .30	ใช้ได้	มีคุณภาพ	.05 ถึง .09	พอใช้	มีคุณภาพ
.31 ถึง .50	พอใช้	มีคุณภาพ	.10 ถึง .30	ใช้ได้	
.51 ถึง 1.00	ใช้ไม่ได้	ไม่มีคุณภาพ	.31 ถึง .50	พอใช้	ไม่มีคุณภาพ
			.51 ถึง 1.00	ใช้ไม่ได้	

3.3 สรุปท้ายบท

1. การประเมินผลแบบอิงกลุ่ม เกิดจากแนวคิดที่ว่า ในการจัดการเรียนการสอน นักเรียนย่อมมีความแตกต่างเป็นรายบุคคล จึงนำผลจากการวัดมาเปรียบเทียบกับนักเรียนในกลุ่มเดียวกัน
2. การประเมินผลแบบอิงเกณฑ์ เกิดจากแนวความคิดที่ว่า ในการจัดการเรียนการสอนควรจะให้ นักเรียนเรียนอย่างรอบรู้ (Master Learning) จึงนำผลจากการวัดมาเปรียบเทียบกับเกณฑ์หรือคุณลักษณะที่กำหนดขึ้น โดยระบุในรูปของจุดประสงค์เชิงพฤติกรรม

3. การตรวจสอบคุณภาพของข้อสอบ หรือการวิเคราะห์ข้อสอบ นับเป็นเทคนิคอย่างหนึ่งในการปรับปรุงและพัฒนาข้อสอบ ทั้งเป็นรายข้อและทั้งฉบับ
4. ก่อนจะถึงขั้นวิเคราะห์ข้อสอบนั้น ครูต้องทำการวิเคราะห์หลักสูตร สอนตามหลักสูตร และเขียนข้อสอบตามตารางวิเคราะห์หลักสูตร จึงจะช่วยให้การวิเคราะห์ข้อสอบบังเกิดผลดี
5. ข้อสอบที่นำมาวิเคราะห์ควรเป็นข้อสอบแบบเลือกตอบ ชนิดตัวเลือก 4-5 ตัว
6. การวิเคราะห์ข้อสอบแบบอิงกลุ่ม
 - 1) ถ้าครูผู้สอนได้วิเคราะห์ข้อสอบเป็นรายข้อแบบทุกตัวเลือก (ค่า p และ r) ก็นับว่าเป็นการพัฒนาการเรียนการสอนและแบบทดสอบได้อย่างดี และเทคนิคการวิเคราะห์ข้อสอบแบบอิงกลุ่มเฉพาะเป็นรายข้อไม่จำเป็นต้องใช้เทคนิค 25%, 27%, 40% หรือ 50% แต่สามารถเลือกใช้เทคนิคในช่วง 25% ถึง 50%
 - 2) การวิเคราะห์ข้อสอบทั้งฉบับ ถ้าเกี่ยวกับความเที่ยงตรงหากครูผู้สอนได้ออกข้อสอบตามตารางวิเคราะห์หลักสูตรก็ช่วยให้ข้อสอบมีความเที่ยงตรง และยังมีวิธีหาค่าความเที่ยงตรงด้วยวิธีอื่นๆ อีก
 - 3) ส่วนความเชื่อมั่นเป็นเรื่องของสถิติ ถ้าใช้ในการสอบครั้งสำคัญ ๆ ก็ควรจะได้หาค่าความเชื่อมั่นด้วยซึ่งมีให้เลือกหลายวิธี