

การวิเคราะห์ข้อสอบ (Item Analysis)

เรียบเรียงโดย พ.ท.หญิง ปวรส บุตะเขียว

การวิเคราะห์ข้อสอบ เป็นการตรวจสอบคุณภาพของข้อสอบเป็นรายข้อ เพื่อพิจารณาว่าข้อสอบแต่ละข้อนั้น มีระดับความยากหรือค่าความง่าย (Difficulty index or Easiness) และค่าอำนาจจำแนกของข้อสอบ (Discriminant index) เพียงใด รวมทั้งพิจารณาถึงประสิทธิภาพของตัวลงในข้อเลือกตอบของข้อสอบข้อนั้นด้วย ผลการวิเคราะห์จะทำให้ทราบว่าข้อสอบแต่ละข้อมีความเหมาะสมมากน้อยเพียงใด ข้อสอบที่มีคุณภาพจะสามารถนำไปวัดและประเมินผลได้อย่างเที่ยงตรงและเชื่อมั่นได้

ข้อสอบที่จะวิเคราะห์คุณภาพรายข้อ ลักษณะข้อสอบจะเป็นแบบเลือกตอบ แบบถูก – ผิด หรือแบบเติมคำสั้น ๆ ก็ได้ แต่ต้องมีรูปแบบเดียวกันทั้งฉบับ ซึ่งมีวิธีวิเคราะห์ 2 แบบ คือ

1. การวิเคราะห์ข้อสอบแบบเลือกตอบแบบอิงกลุ่ม

เป็นการวิเคราะห์ข้อสอบที่ใช้ผลการสอบไปประเมินแบบอิงกลุ่ม โดยนำความแตกต่างระหว่างบุคคลมาเป็นแนวคิดในการประเมินผล การวิเคราะห์ข้อสอบแบบอิงกลุ่มนี้ จึงเป็นการวิเคราะห์โดยนำผลคะแนนของผู้เรียนแต่ละคนไปเปรียบเทียบกับผลคะแนนของผู้เรียนคนอื่น ๆ ในกลุ่มเดียวกัน เพื่อจำแนกคนที่เรียนเก่งออกจากคนที่เรียนอ่อน

ดัชนีบ่งชี้คุณภาพของข้อสอบสำหรับแบบสอบอิงกลุ่ม ได้แก่ ระดับความยากง่าย (p) และอำนาจจำแนก (r) ซึ่งข้อสอบที่ดีนั้น นอกจากจะต้องมีระดับความยากง่ายพอเหมาะแล้ว จะต้องม้ออำนาจการจำแนกที่ดีด้วย

1.1 ระดับความยากง่าย (p) หมายถึง สัดส่วนของจำนวนผู้ที่ตอบข้อสอบได้ถูกต้อง ต่อจำนวนผู้ที่ตอบข้อสอบทั้งหมด หรือหมายถึงจำนวนร้อยละของผู้ตอบข้อสอบนั้น ๆ ถูก ตัวอย่างเช่น ค่า $p = 0.30$ แสดงว่า จำนวนผู้ตอบ 100 คน มีผู้ที่ตอบข้อนั้น ๆ ถูก 30 คน ค่าความยากง่ายจะมีค่าระหว่าง 0 ถึง 1.00

ในการพิจารณาค่าความยากง่ายนั้น ถ้าข้อสอบมีค่าความยากง่ายสูง เช่น $p = 0.95$ แสดงว่า มีผู้ตอบถูกจำนวนมาก จึงถือว่าเป็นข้อสอบที่ง่าย แต่ในทางกลับกัน ถ้าข้อสอบมีผู้ตอบถูกน้อย เช่น $p = 0.15$ แสดงว่า เป็นข้อสอบที่ยาก

ข้อสอบที่ดีจะมีระดับความยากง่าย เท่ากับ 0.5 ซึ่งจะทำให้เกิดค่าอำนาจการจำแนก สูงสุดและมีความเที่ยงสูง อย่างไรก็ตามในการสอบวัดความรู้ผลการเรียนโดยทั่วไป มักนิยมให้มีข้อสอบที่มีระดับความยากง่ายในระดับต่าง ๆ ปะปนกันไป โดยจัดให้มีข้อสอบมีค่าความยากง่าย

พอเหมาะ (p มีค่าใกล้เคียง 0.5) เป็นส่วนใหญ่ รวมทั้งให้มีข้อสอบที่ค่อนข้างยากและค่อนข้างง่ายอีกจำนวนหนึ่ง แต่ถ้าเป็นการสอบแข่งขันเพื่อคัดเลือกผู้ที่มีความรู้ความสามารถควรมีสัดส่วนของข้อสอบที่ยากสูงขึ้น ทั้งนี้ ข้อสอบที่ดีควรมีค่าความยากง่ายระหว่าง 0.20 – 0.80 ในข้อสอบประเภท 4 ตัวเลือก ส่วนข้อสอบประเภทถูก – ผิด ค่าความยากง่าย ควรอยู่ระหว่าง 0.60 - 0.95 (Nunnally, 1967. อ้างถึงใน เขาวดี ราชชัยกุล วิบูลย์ศรี, 2552 , หน้า 144 - 145)

เกณฑ์การแปลความหมายค่าความยากง่าย (p) ของข้อสอบ (ล้วน สายยศ และ อังคณา สายยศ, 2543, หน้า 185)

ความยากง่ายของข้อสอบ (p)	ความหมาย
0.81 - 1.00	ง่ายมาก (ควรปรับปรุงหรือตัดทิ้ง)
0.60 - 0.80	ค่อนข้างง่าย (ดี)
0.40 - 0.59	ยากพอเหมาะ (ดีมาก)
0.20 - 0.39	ค่อนข้างยาก (ดี)
0 - 0.19	ยากมาก (ควรปรับปรุงหรือตัดทิ้ง)

1.2 อำนาจจำแนก (r) หมายถึง ความสามารถของข้อสอบในการจำแนกหรือแยกให้เห็นความแตกต่างระหว่างผู้สอบที่มีผลสัมฤทธิ์ต่างกัน เพื่อที่จะใช้พยากรณ์หรือบ่งชี้ความแตกต่างที่เห็นชัดในด้านความสามารถ เช่น จำแนกคนเก่งกับคนอ่อนจากกันได้ โดยถือว่าคนเก่งควรทำข้อสอบข้อนั้นได้ ส่วนผู้ที่ย่อมนไม่ควรทำข้อสอบข้อนั้นได้ อำนาจจำแนกของข้อสอบ จะมีค่าตั้งแต่ -1 ถึง +1 ค่าอำนาจจำแนกที่ดี ควรมีค่าตั้งแต่ 0.2 ขึ้นไป

กรณีที่ค่า r ติดลบ แสดงว่า ข้อสอบข้อนั้นจำแนกกลับ คนเก่งทำไม่ได้ แต่คนอ่อนทำได้ ถือว่าเป็นข้อสอบที่ไม่ดีควรตัดทิ้ง นอกจากนี้ อาจารย์ผู้สอนควรตรวจสอบการจัดการเรียนสอนของตน ว่าเพราะเหตุใดผู้ที่เรียนเก่งจึงไม่เข้าใจในเรื่องที่สอน

เกณฑ์การแปลความหมายค่าอำนาจจำแนก (r) ของข้อสอบ

อำนาจจำแนกของข้อสอบ (r)	ความหมาย
0.60 - 1.00	อำนาจจำแนกดีมาก
0.40 - 0.59	อำนาจจำแนกดี
0.20 - 0.39	อำนาจจำแนกพอใช้
0.10 - 0.19	อำนาจจำแนกต่ำ (ควรปรับปรุงหรือตัดทิ้ง)
-1.00 - 0.09	อำนาจจำแนกต่ำมาก (ควรปรับปรุงหรือตัดทิ้ง)

ประสิทธิภาพของตัวลง

1. สัดส่วนของผู้ที่เลือกตัวลง (P_w) หมายถึง สัดส่วนของจำนวนคนที่เลือกตัวลงนั้น ๆ เช่น ข้อสอบข้อหนึ่งมีคนสอบ 100 คน มีคนเลือกตัวลง ก. 25 คน แสดงว่า สัดส่วนผู้เลือกตัวลง ก. = 0.25 หรือ 25%

P_w จะมีค่า 0 ถึง 1	ตัวลงที่ดีควรมีค่า P_w ตั้งแต่ 0.05 ขึ้นไป
-----------------------	--

2. อำนาจจำแนกของตัวลง (r_w) หมายถึง ผลต่างระหว่างสัดส่วนของคนในกลุ่มอ่อนที่เลือกตัวลงกับสัดส่วนของคนในกลุ่มเก่งที่เลือกตัวลงนั้น ๆ

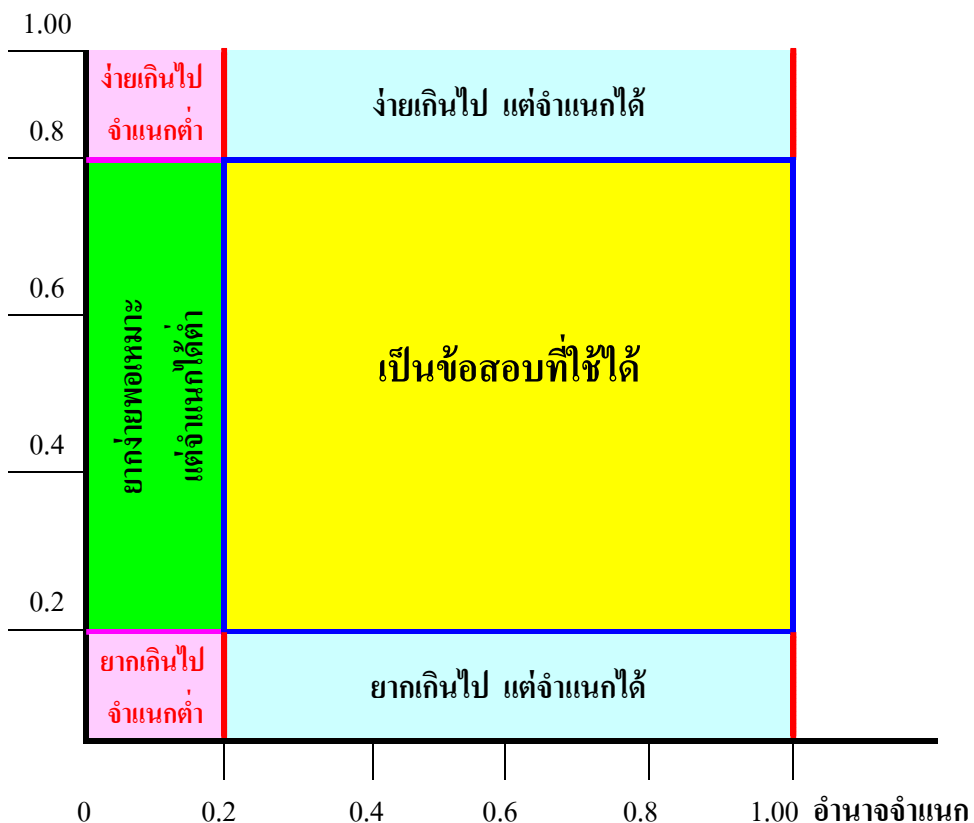
r_w มีค่าตั้งแต่ -1 ถึง 1	ตัวลงที่ดีควรมีค่า r_w ตั้งแต่ 0.05 ขึ้นไป
-----------------------------	--

การเลือกข้อสอบจากผลการวิเคราะห์ข้อสอบ

เมื่อวิเคราะห์ข้อสอบได้ค่าความยากง่ายและอำนาจการจำแนกแล้ว ควรนำมาจัดทำกราฟ ซึ่งจะแสดงการกระจายของข้อสอบและชี้ให้เห็นว่า ข้อสอบข้อใดบ้างที่อยู่ในเกณฑ์ใช้ได้หรือใช้ไม่ได้

ตัวอย่าง : กราฟแสดงการกระจายของข้อสอบตามผลการวิเคราะห์ข้อสอบ

ระดับความยาก (p)



2. การวิเคราะห์ข้อสอบแบบเลือกตอบแบบอิงเกณฑ์

การวิเคราะห์ข้อสอบแบบอิงเกณฑ์ มีจุดมุ่งหมายเพื่อดูว่าผู้เรียนมีความรู้ ความสามารถ ะไรบ้าง หรือผลที่เกิดจากการเรียนการสอนตามเกณฑ์หรือมาตรฐานที่ตั้งไว้เพียงใด มากกว่าที่จะ เปรียบเทียบกับกลุ่มว่าใครเก่งกว่าใคร เนื่องจากการวิเคราะห์ข้อสอบแบบอิงเกณฑ์เกี่ยวข้องกับ ประสิทธิภาพในการวัดผลของการจัดการเรียนการสอนโดยตรง ถ้าหากอาจารย์ใช้ข้อสอบชุด เดียวกันทดสอบก่อนเรียน (Pre – Test) และหลังเรียน (Post – Test) ข้อสอบอิงเกณฑ์ที่มี ประสิทธิภาพ ควรตั้งอยู่บนสมมติฐานที่ว่า กระบวนการเรียนการสอนจะส่งผลโดยตรงต่อการ เรียนรู้ของผู้เรียน เมื่อการเรียนการสอนสิ้นสุดลง ควรมีผู้ตอบถูกมากกว่าก่อนเรียน จึงเกิดแนวคิด ของดัชนีความไวของผลการเรียนการสอนเกิดขึ้น

ดัชนีบ่งชี้คุณภาพของข้อสอบสำหรับแบบสอบอิงเกณฑ์ ได้แก่ ความสอดคล้อง ระหว่างข้อสอบกับจุดมุ่งหมาย (Item – Objective Congruence : IOC) ความยากของข้อสอบ (Item Difficulty : P_i) และอำนาจจำแนกหรือดัชนีความไว (Sensitivity Index : S_i)

2.1 ความสอดคล้องระหว่างข้อสอบกับจุดมุ่งหมาย (IOC) หมายถึง ลักษณะของ ข้อสอบที่เป็นตัวแทนของจุดมุ่งหมายการเรียนรู้ที่กำหนดไว้ วิธีการวัดความสอดคล้องสามารถ ทำได้โดยอาศัยการตัดสินของผู้ทรงคุณวุฒิ ไม่ต่ำกว่า 3 คน เป็นผู้พิจารณา เช่น ให้ผู้ทรงคุณวุฒิ ประเมินค่าความสอดคล้องของข้อสอบแต่ละข้อกับจุดมุ่งหมายการเรียนรู้ที่มุ่งวัดโดยใช้มาตราส่วน แบบประเมินค่า 5 ระดับ (Likert Scale) แล้วนำมาคำนวณหาค่าเฉลี่ย ถ้าค่าเฉลี่ย (\bar{X}) ≥ 4.00 จะถือว่าสอดคล้อง หรือถ้าใช้แบบประเมินค่า 3 ระดับ จะถือว่าข้อสอบนั้นสอดคล้องกับ จุดมุ่งหมายเมื่อ $\bar{X} \geq 80\%$

2.2 ความยากของข้อสอบ (Item Difficulty : P_i) หมายถึง สัดส่วนของผู้ตอบ ข้อสอบข้อนั้น ๆ ได้ถูกต้อง โดยพิจารณาจากกลุ่มที่ได้เรียนเรื่องนั้นมาแล้ว กับกลุ่มที่ไม่ได้เรียน เรื่องนั้น ข้อสอบที่ง่ายสำหรับกลุ่มที่เคยเรียน สามารถเป็นข้อสอบอิงเกณฑ์ที่ดีได้ เพราะอาจารย์ ผู้สอนสามารถจัดการเรียนการสอนให้ผู้เรียนส่วนใหญ่เกิดการเรียนรู้และทำข้อสอบได้หรืออาจเป็น เพราะข้อสอบง่ายเกินไป จึงทำให้ผู้เรียนตอบถูกจากความรู้เดิมโดยไม่ได้เรียนรู้เพิ่มเติม ค่า P_i จะ อยู่ระหว่าง 0 ถึง 1

ดังนั้น การวิเคราะห์ข้อสอบอิงเกณฑ์ จึงจำเป็นต้องตรวจสอบว่า ก่อนเรียนมีสัดส่วน ของคำตอบถูกมากน้อยเพียงใดเพื่อนำผลมาเปรียบเทียบกับตอนหลังเรียน

การวิเคราะห์ข้อสอบแบบอิงเกณฑ์ จะพิจารณาค่าความยากของข้อสอบแตกต่างจาก แบบอิงกลุ่ม เนื่องจากไม่ได้เน้นที่จะนำค่าความยากง่ายมาเพื่อเลือกข้อสอบ แต่จะเน้นที่คุณภาพใน การสอน กล่าวคือ ถ้าอาจารย์ยังไม่ได้สอนเนื้อหา นั้น ข้อสอบควรจะยาก คือ มีค่า P_i ต่ำ แต่ถ้า

ได้สอนแล้วและสอนอย่างมีประสิทธิภาพส่งผลให้ผู้เรียนเกิดการเรียนรู้ ข้อสอบควรจะง่าย คือ มีค่า P_i สูง

ลักษณะข้อสอบอิงเกณฑ์ที่ดี ควรมีค่าความยากของข้อสอบ ดังนี้ (ศิริชัย กานจนาวาสี, 2552 หน้า 243)

กลุ่มผู้เรียน	ความยากของข้อสอบ (P_i)
1. กลุ่มที่เคยเรียนมาแล้ว	- ควรมีค่า P_i สูง (0.70 - 1.00) - ถ้า $P_i < 0.70$ ควรปรับปรุงหรือตัดทิ้ง
2. กลุ่มที่ยังไม่เคยเรียน	- ควรมีค่า P_i ต่ำ (0 - 0.5) - ถ้า $P_i > 0.5$ แสดงว่า ข้อสอบง่ายเกินไป (ควรปรับปรุงหรือตัดทิ้ง)

2.3 อำนาจจำแนกหรือดัชนีความไว (Sensitivity Index : S_i) หมายถึง ผลต่างระหว่างสัดส่วนของผู้ที่ตอบถูกหลังเรียน กับสัดส่วนผู้ที่ตอบถูกต้องก่อนเรียน

S_i มีค่าระหว่าง -1 ถึง 1 ค่า S_i ควรมีค่าเป็นบวก ยิ่งเข้าใกล้ 1 ยิ่งมีความไว ซึ่งแสดงว่าข้อสอบข้อนั้นสามารถจำแนกผู้รอบรู้จากผู้ไม่รอบรู้ได้ดี และสามารถจำแนกการสอนได้อย่างมีประสิทธิภาพ ดังนั้น อำนาจจำแนกหรือดัชนีความไวของข้อสอบ (S_i) ควรมีค่า > 0

ลักษณะข้อสอบอิงเกณฑ์ที่สามารถจำแนกการสอนได้อย่างมีประสิทธิภาพ ควรมีค่า S_i ดังนี้ (พิชิต ฤทธิ์จรูญ, 2552, หน้า 145)

อำนาจจำแนกหรือดัชนีความไวของข้อสอบ (S_i)
- ตั้งแต่ 0.40 ขึ้นไป

ซึ่งค่า S_i นี้ จะแตกต่างจากค่าอำนาจจำแนก (r) ของข้อสอบแบบอิงกลุ่ม คือ ค่าอำนาจจำแนก (r) ของข้อสอบแบบอิงกลุ่ม เป็นดัชนีที่ชี้ให้เห็นว่า สามารถจำแนกคนที่มีความสามารถทางการเรียนสูงและต่ำได้มากน้อยเพียงใด

ข้อจำกัดของดัชนีความไว

1. ต้องสอบ 2 ครั้ง คือ ก่อนเรียนและหลังเรียน จึงจะคำนวณค่าได้
2. S_i ที่มีค่าต่ำหรือมีค่าเป็นลบ อาจเป็นผลจากข้อสอบหรือการจัดการเรียนการสอน ไม่มีประสิทธิภาพ
3. กรณีที่มีระยะห่างระหว่างการสอบครั้งแรกและครั้งที่ 2 ใกล้กัน คำตอบหรือคะแนนครั้งหลังอาจได้รับอิทธิพลจากการสอบครั้งแรกได้

ประสิทธิภาพของตัวลวง เมื่อสร้างข้อสอบอิงเกณฑ์แบบหลายตัวเลือก (Multiple Choices) ผู้ที่เลือกตัวลวง ถือว่า ตอบผิด ซึ่งแสดงให้เห็นว่า ผู้นั้นยังไม่สัมฤทธิ์ผลตามเป้าหมายของการวัด ในการวิเคราะห์ประสิทธิภาพตัวลวงทำได้โดยการตรวจสอบความถี่ของผู้ที่เลือกตัวลวงนั้น ๆ และควรดูจากผลการสอบก่อนเรียน (Pre – Test) ถ้าตัวลวงใดมีผู้เลือกในสัดส่วนที่สูง ถือว่าเป็นตัวลวงที่ใช่ได้ แต่ถ้าตัวลวงใดมีผู้เลือกน้อย แสดงว่าตัวลวงนั้นขาดประสิทธิภาพ สมควรที่จะต้องมีการปรับปรุงแก้ไขให้ดีขึ้น (ศิริชัย กาญจนวาสี , 2552 หน้า 241) และตัวลวงที่จัดว่าเป็นตัวลวงที่ดีนั้น ผู้ที่เรียนอ่อนควรจะเลือกตอบมากกว่าผู้ที่มีผลการเรียนดี

ประโยชน์ของการวิเคราะห์ข้อสอบ

1. ทำให้ทราบข้อมูลพื้นฐานของตัวข้อสอบและคำตอบ รวมถึงข้อสอบแต่ละข้อได้ทำหน้าที่วัดผลสัมฤทธิ์อย่างมีประสิทธิภาพเพียงใด
2. ชี้ให้เห็นถึงจุดอ่อนที่ผู้สอนหรือผู้เรียนต้องปรับปรุงแก้ไข เพื่อพัฒนาความสามารถ และทักษะของผู้เรียนให้เป็นตามที่คาดหวัง
3. เป็นพื้นฐานสำคัญในการปรับปรุงพัฒนาการเรียนการสอน โดยเฉพาะอย่างยิ่งด้านพฤติกรรมการเรียนรู้ที่เหมาะสมสำหรับเนื้อหาวิชานั้น ๆ
4. ช่วยเพิ่มทักษะในการสร้างข้อสอบ ทำให้ทราบถึงอำนาจการจำแนก ระดับความยากง่าย ประสิทธิภาพของตัวลวง ตลอดจนการเขียนข้อสอบในลักษณะใดจึงจะได้ข้อสอบที่ดี
5. ทำให้สามารถคัดเลือกข้อสอบที่มีคุณภาพมารวมเป็นฉบับข้อสอบที่ดี สามารถจัดทำข้อสอบคู่ขนานและเป็นรากฐานสำคัญในการพัฒนาแบบสอบมาตรฐาน

บทสรุป

ข้อสอบที่ดี ควรมีค่าความยากง่ายที่เหมาะสมและมีอำนาจจำแนกสูง ค่าเฉลี่ยความยากง่ายของข้อสอบทั้งฉบับ ควรมีค่าประมาณ 0.50 แต่อย่างไรก็ตามในการคัดเลือกข้อสอบจะต้องคำนึงถึงความสมดุลระหว่างข้อสอบที่มีสถิติเหมาะสมกับข้อสอบที่วัดครอบคลุมจุดประสงค์ และเนื้อหาที่ต้องการด้วย บางครั้งอาจมีความจำเป็นที่จะต้องใช้ข้อสอบที่มีอำนาจการจำแนกไม่สูงนัก เพื่อให้มีข้อสอบที่วัดครอบคลุมเนื้อหาที่ต้องการ

1. การวิเคราะห์ข้อสอบแบบเลือกตอบแบบอิงกลุ่ม
 - ระดับความยากง่าย (p) ถ้าค่า p ยิ่งมากข้อสอบข้อนั้นจะง่ายและค่า p ยิ่งน้อย ข้อสอบนั้นจะยาก ทั้งนี้ข้อสอบที่ดีควรมีค่า p ระหว่าง 0.20 - 0.80

- อำนาจการจำแนก(r) เป็นการจำแนกผู้เรียนออกจากกันตามระดับความรู้ความสามารถ ค่า r ควรมีค่าตั้งแต่ 0.20 ขึ้นไป

2. การวิเคราะห์ข้อสอบแบบเลือกตอบแบบอิงเกณฑ์ โดยการตรวจสอบค่าดัชนีความไว (S_i) ของผลการเรียน จะช่วยสะท้อนให้เห็นถึงความสามารถก่อนเรียนกับผลการเรียนรู้หลังการเรียนการสอนแล้ว ข้อสอบที่ดี ควรมี S_i ตั้งแต่ 0.40 ขึ้นไป

ผลการวิเคราะห์ข้อสอบ ถ้าพบว่า ข้อสอบไม่มีคุณภาพควรกลับมาพิจารณาว่า เกิดจากสาเหตุใด ข้อคำถามไม่มีความเป็นปรนัยหรือไม่ รวมทั้งต้องพิจารณาที่ตัวเลือกด้วยว่า เป็นไปตามหลักการเขียนข้อสอบที่ดีหรือไม่ และดำเนินการปรับปรุงข้อสอบ รวมทั้งนำไป Try out อีกครั้ง เพื่อวิเคราะห์ประสิทธิภาพของข้อสอบต่อไป

เอกสารอ้างอิง

พิชิต ฤทธิจรูญ. (2552). *หลักการวัดและประเมินผลการศึกษา*. (พิมพ์ครั้งที่ 5). กรุงเทพฯ :

เฮาส์ ออฟ เคอร์มิสจำกัด.

เยาวดี รามชัยกุล วิบูลย์ศรี. (2552). *การวัดผลและการสร้างแบบสอบผลสัมฤทธิ์*. (พิมพ์ครั้งที่ 8).

กรุงเทพฯ : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ล้วน สายยศ และ อังคณา สายยศ. (2543). *เทคนิคการวัดผลการเรียนรู้*. (พิมพ์ครั้งที่ 2).

กรุงเทพฯ : สุวีริยาสาส์น.

ศิริชัย กาญจนวาสี. (2552). *ทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory)*.

(พิมพ์ครั้งที่ 6). กรุงเทพฯ : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.