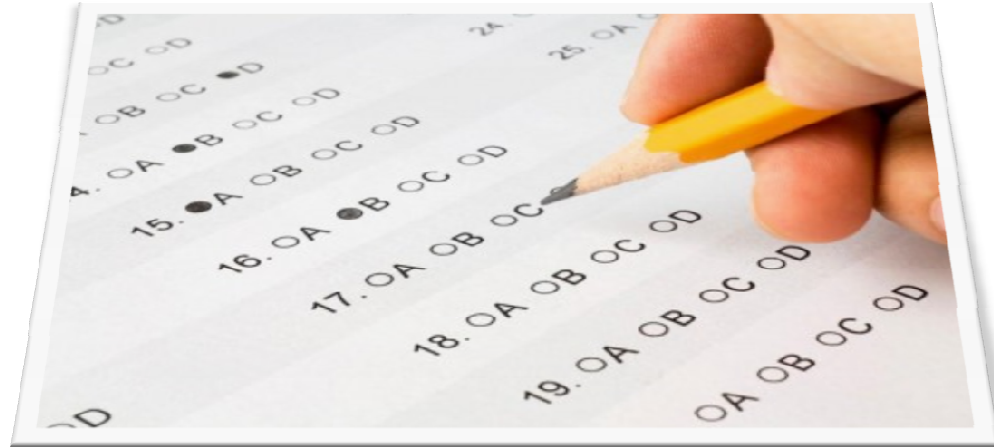




การวิเคราะห์ข้อสอบปรนัย

(Item analysis)

ประโยชน์ที่ได้จากการวิเคราะห์ข้อสอบ

- ▶ **1) ใช้ปรับแก้คะแนนสอบ** ผลที่ได้จะทำให้เราทราบว่าข้อสอบข้อใดเฉลยผิด / มีคำตอบที่ถูกมากกว่า 1 ตัวเลือก / เป็นข้อที่น่าจะมีปัญหา เช่น มีความคลุมเครือในคำถาม ตัวเลือกมีความซ้ำซ้อนกัน เนื้อหาอยู่นอกเหนือไปจากสิ่งที่สอน เป็นต้น
- ▶ **2) ใช้ปรับปรุงคุณภาพข้อสอบ**
- ▶ **3) ใช้ในการบริหารคลังข้อสอบ** ค่าที่ได้จะบอกคุณลักษณะของข้อสอบ ถ้าเป็นข้อสอบที่ดี (ยากง่ายเหมาะสม/จำแนกผู้สอบได้) เลือกเก็บไว้ในคลังข้อสอบ เพื่อจะกลั่นนำมาใช้ใหม่ในอนาคตตามความเหมาะสม
- ▶ **4) ใช้ในการพัฒนาคุณภาพการเรียนการสอน** ค่าสถิติที่ได้ทำให้เราทราบว่าประเด็นใดที่ผู้เรียนมีความเข้าใจผิดอยู่มาก ผู้สอนควรเน้นย้ำประเด็นนั้น

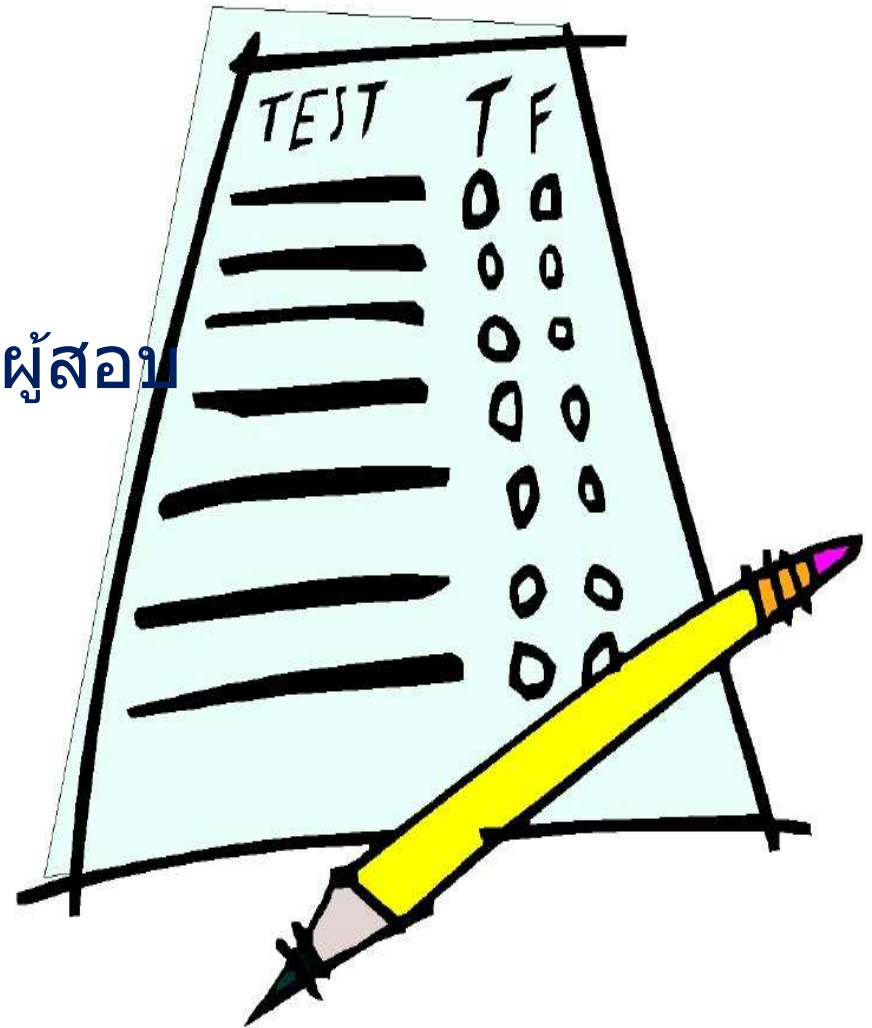
ข้อตกลง



- ▶ ค่าในการวิเคราะห์ข้อสอบ
ปรนัยนี้ใช้กับในชั้นเรียนปกติ
คือ มี ผู้ ส อ บ มี ระ ดับ
ความสามารถทุกประเภท
(เรียนดี/ ปานกลาง/ต่ำ) ปะปน
กันไป
- ▶ ค่าต่างๆ ที่ได้มาจากการ
วิเคราะห์นั้น ขึ้นอยู่กับกลุ่ม
ตัวอย่างที่ใช้ในการเก็บข้อมูล
หากข้อมูลที่ได้มาจากกลุ่ม
ตัวอย่างที่มีขนาดใหญ่พอ
ค่าสถิติต่างๆ ที่ได้จะค่อนข้าง
เที่ยงตรง

การวิเคราะห์ข้อสอบรายข้อ พิจารณา 3 ปัจจัย

- ▶ 1) ความยากง่ายของข้อสอบ
(item difficulty, P)
- ▶ 2) ความสามารถในการจำแนกผู้สอบ
ตามระดับความสามารถ
(item discrimination, R)
- ▶ 3) ประสิทธิภาพของตัวลวง
(distracter efficiency)



1) ความยากง่ายของข้อสอบ (item difficulty, P)

- ▶ เป็นค่าสัดส่วนของผู้สอบที่ตอบข้อสอบข้อนั้นถูก

$$P = R/N$$

P คือ ค่าดัชนีความยากง่าย

R คือ จำนวนนักเรียนที่ทำข้อนั้นถูก

N คือ จำนวนนักเรียนทั้งหมดที่ทำข้อสอบข้อนั้น

ดัชนีค่าความยากง่าย	ความหมาย
0.80-1.00	ง่ายมาก(ควรปรับปรุงหรือตัดทิ้ง)
0.60-0.79	ค่อนข้างง่าย
0.40-0.59	ปานกลาง
0.20-0.39	ค่อนข้างยาก
น้อยกว่า 0.20	ยากมาก(ควรปรับปรุงหรือตัดทิ้ง)

1) ความยากง่ายของข้อสอบ (item difficulty, P)

- ▶ ตัวอย่าง การคำนวณค่า P ของผู้สอบจำนวน 50 คน

ข้อที่	ก	ข	ค	ง	P
1	10	20	7	13	$20/50 = 0.4$
2	15	10	12	13	$15/50 = 0.3$
3	5	8	35	2	$35/50 = 0.7$

ปานกลาง

ค่อนข้างยาก

ค่อนข้างง่าย



1) ความยากง่ายของข้อสอบ (item difficulty, P)

มีข้อเสนอแนะว่า ในแต่ละชุดข้อสอบควรมีความยากง่าย
พอเหมาะ โดยในชุดข้อสอบควรประกอบด้วย

EXAM

50% มีค่าความยากง่ายปานกลาง (ข้อสอบกลางๆ)

25% มีค่าความยากง่ายมาก (ข้อสอบง่าย)

25% มีค่าความยากง่ายน้อย (ข้อสอบยาก)



2) ความสามารถในการจำแนกผู้สอบตามระดับความสามารถ (item discrimination, R)

- ▶ เป็นคุณสมบัติที่บ่งบอกถึงความสามารถของข้อสอบที่จำแนกเด็กเก่ง – อ่อน จะมีค่า R เป็นตัวดัชนีชี้บ่งให้ทราบว่า ข้อสอบข้อใดมีอำนาจจำแนกสูงก็เป็นข้อสอบที่ดี หมายถึง ข้อสอบข้อนี้คนทำถูกจะเป็นพวกกลุ่มเก่ง ถ้าใครทำผิดจะเป็นพวกกลุ่มอ่อน

$$R = \frac{R_H - R_L}{N_{H/L}}$$

R	คือ ค่าอำนาจจำแนก
R _H	คือ จำนวนผู้สอบที่ตอบถูกในกลุ่มคะแนนสูง
R _L	คือ จำนวนผู้สอบที่ตอบถูกในกลุ่มคะแนนต่ำ
N _{H/L}	คือ จำนวนผู้สอบทั้งหมดในกลุ่มคะแนนสูงหรือกลุ่มคะแนนต่ำ

Note: การหาค่าอำนาจจำแนกตามหลักสูตรนี้จะเหมาะสมมากเมื่อมีผู้สอบจำนวนมาก ๆ

Note: การแยกกลุ่มสูง-กลุ่มต่ำ

จำนวนผู้สอบ ควรมีมากพอสมควร (ประมาณ 100 คนขึ้นไป บางทฤษฎีก็มากกว่า 60 คนขึ้นไป) ซึ่งทำให้การกระจายของคะแนนเป็นโค้งปกติ อย่างไรก็ตาม กรณีที่คะแนนมีการแจกแจงแบบปกติ ใช้เทคนิค 27 % นั้นถูกต้อง แต่สำหรับการแจกแจงที่ไม่เป็นปกติ ให้ใช้เทคนิค 33% แทน ทั้งนี้เพราะถ้าใช้เทคนิค 27% แล้ว จะทำให้ระดับความยากง่าย และอำนาจจำแนกไม่ตรงกับความเป็นจริง

บางครั้งถ้าผู้สอบมีจำนวนไม่มากนักสามารถแบ่งกลุ่มเก่ง-อ่อน (สูง-ต่ำ) เป็นครึ่งๆ ได้เลย (50%)



2) ความสามารถในการจำแนกผู้สอบตามระดับความสามารถ (item discrimination, R)

ค่าอำนาจจำแนก ความหมาย



1.00	จำแนกดีเลิศ
0.80 - 0.99	จำแนกดีมาก
0.60 - 0.79	จำแนกดี
0.40 - 0.59	จำแนกได้ปานกลาง
0.20 - 0.39	จำแนกได้บ้าง
0.00 - 0.19	จำแนกไม่ค่อยได้ (ควรปรับปรุง)
<0 (ติดลบ)	จำแนกไม่ได้ (ควรตัดทิ้ง สร้างข้อสอบใหม่)

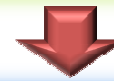
ค่าอำนาจจำแนกที่ใช้ได้จะต้องมีค่า D สูงกว่า 0.20 ขึ้นไป

2) ความสามารถในการจำแนกผู้สอบตามระดับความสามารถ (item discrimination, R)

ค่าอำนาจจำแนกที่เป็นลบ จะมีลักษณะที่ตรงข้ามกับค่าอำนาจจำแนกที่เป็นบวก เนื่องจากค่าอำนาจจำแนกเป็นความสัมพันธ์ระหว่างการตอบถูกกับคุณลักษณะของบุคคล เช่น ความสามารถของบุคคล ดังนั้นการแปลความหมายจะเป็นดังนี้



ข้อสอบที่มีค่า R เป็นบวก หมายความว่า ข้อสอบสามารถแยกคนที่มีความรู้มากจากคนที่มีความรู้น้อย หรือจากคนที่มีความรอบรู้จากคนที่ไม่มีความรอบรู้ได้ เพราะคนที่มีความรู้มาก มีแนวโน้มตอบถูกมาก ส่วนคนที่มีความรู้น้อยมีแนวโน้มตอบผิดมาก ข้อสอบยังมีค่า R เป็นบวกสูงมากเท่าใด แสดงว่าข้อสอบนั้น คนที่มีความรู้มาก จะมีแนวโน้มตอบถูกมากขึ้นเท่านั้น และคนที่มีความรู้น้อยจะมีแนวโน้มตอบผิดมากขึ้นเท่านั้น



ข้อสอบที่มีค่า R ใกล้เคียงศูนย์ หมายความว่า ข้อสอบข้อนั้น แยกคนที่มีความรู้มากกับคนที่มีความรู้น้อยไม่ค่อยได้ หรือแยกคนที่มีความรอบรู้กับคนที่ไม่รอบรู้ไม่ค่อยได้ เพราะคนที่มากอาจตอบข้อสอบข้อนั้นถูกหรือผิดก็ได้ ดังนั้น ในการคัดเลือกข้อสอบไว้ใช้ควรคัดเลือกที่มีค่า R สูงๆ

ประสิทธิภาพของตัวลวง (distracter efficiency)

การพิจารณาประสิทธิภาพของตัวลวง มี 2 ขั้นตอน

1. สัดส่วนของผู้ที่เลือกตัวลวง (P_w) หมายถึง สัดส่วนของจำนวนคนที่เลือกตัวลวงนั้น ๆ เช่น ข้อสอบข้อหนึ่งมีคนสอบ 100 คน มีคนเลือกตัวลวง ก. 25 คน แสดงว่า สัดส่วนผู้เลือกตัวลวง ก. = 0.25 หรือ 25%

P_w จะมีค่า 0 ถึง 1

ตัวลวงที่ดีควรมีค่า P_w ตั้งแต่ 0.05 ขึ้นไป



3) ประสิทธิภาพของตัวลวง (distracter efficiency)

2. อำนาจจำแนกของตัวลวง (R_w) ประสิทธิภาพของตัวลวงเป็นอีกคุณลักษณะหนึ่งที่มีความสำคัญต่อคุณภาพข้อสอบแต่ละข้อ ตัวลวงที่ดีควรจะถูกละเลือกตอบ และเด็กกลุ่มอ่อนควรจะเลือกมากกว่าเด็กกลุ่มเก่ง การคำนวณประสิทธิภาพของตัวลวงดังกล่าวอาจนำสูตรการคำนวณหาค่า R มาประยุกต์ใช้ก็ได้ ดังนี้

$$R_w = \frac{R_L - R_H}{N_{H/L}}$$

Note: ตัวลวงที่ดีจึงควรมีค่า R_w เป็นบวก ส่วนตัวลวงที่คำนวณได้ค่าดังกล่าวเป็นลบ หมายความว่าเด็กกลุ่มเก่งเลือกตอบมากกว่าเด็กกลุ่มอ่อน ถือว่าเป็นตัวลวงที่ไม่ดี ควรมีค่า ≥ 0.05

R_w	คือ ค่าอำนาจจำแนก
R_H	คือ จำนวนผู้สอบในกลุ่มเก่ง(สูง) ที่เลือกตอบตัวลวงข้อนั้น
R_L	คือ จำนวนผู้สอบในกลุ่มอ่อน(ต่ำ) ที่เลือกตอบตัวลวงข้อนั้น
$N_{H/L}$	คือ จำนวนผู้สอบทั้งหมดในกลุ่มคะแนนสูงหรือกลุ่มคะแนนต่ำ

ตัวอย่าง: ข้อสอบข้อหนึ่งซึ่งมีจำนวนผู้สอบ 40 คน

กลุ่ม	ผลของคำตอบ			
	ก	ข	ค*	ง
สูง	3	2	15	0
ต่ำ	12	3	3	2

ความยากง่ายของข้อสอบ (P) = $(15+3)/40 = 0.45$

อำนาจจำแนกของข้อสอบ (R) = $(15-3)/20 = 0.60$

ตัวलग ก

$$P_w = (3+12)/40 = 0.375$$

$$R_w = (12-3)/20 = 0.45$$

ตัวलग ข

$$P_w = (2+3)/40 = 0.125$$

$$R_w = (3-2)/20 = 0.05$$

ตัวलग ง

$$P_w = (0+2)/40 = 0.05$$

$$R_w = (2-0)/20 = 0.10$$

ค่าอำนาจจำแนกของตัวलग (R_w) นั้น มีความหมายว่า
ตัวलगที่ดีควรलगกลุ่มอ่อนไปตอบมากกว่าलगกลุ่มเก่ง

ข้อควรจำ

ค่าดัชนีชี้วัดคุณภาพต่างๆ ของข้อสอบที่กล่าวมา เป็นเพียงตัวช่วยให้อาจารย์เข้าใจข้อสอบได้ดีขึ้น และช่วยแนะแนวทางในการพัฒนาปรับปรุงข้อสอบให้ดีขึ้น ดัชนีเหล่านี้ไม่ใช่คำตัดสินหรือตัวชี้ชะตาของข้อสอบ ไม่มีดัชนีใดที่ได้จากการวิเคราะห์ข้อสอบจะมาทดแทนดุลยพินิจของอาจารย์ไปได้ บทบาทของอาจารย์ในการวิเคราะห์ข้อสอบคงไม่ใช่การยึดถือตัวเลขดัชนีต่างๆ เป็นกฎตายตัว หากแต่ใช้ดัชนีเหล่านี้ช่วยเป็นแนวทางในการพิจารณาข้อสอบ หากหลังจากการพิจารณาโดยถี่ถ้วนแล้วอาจารย์คิดว่าข้อสอบนั้นเหมาะสมแล้ว ไม่ควรทำการปรับแก้เนื้อหา อาจารย์ก็ยืนยันไปว่าไม่แก้ไข อาจารย์คงไม่ตัดสินความดีชั่วของคนใดคนหนึ่งเพียงการมองชั่วขณะหนึ่งโดยไม่พิจารณาการกระทำอื่นร่วมด้วย ฉะนั้น إذنนั้น อาจารย์ไม่ควรตัดสินชะตากรรมของข้อสอบโดยใช้เพียงค่า P หรือ R โดยไม่พิจารณาความเหมาะสมของเนื้อหาโจทย์และตัวเลือกต่างๆ ในข้อสอบนั้น

อ้างอิง

- ▶ วัดผลจุดคอม. 2548. **ทฤษฎีการวัดการทดสอบ**. (ออนไลน์). แหล่งที่มา :
<http://www.watpon.com/testtheory/testtheory7.pdf>. 9 พฤษภาคม 2555
- ▶ นพ.เชิดศักดิ์ ไอรณรัตน์. 2552. การวิเคราะห์ข้อสอบปรนัย. (ออนไลน์). แหล่งที่มา :
http://www.simedbull.com/journal_files/content_pdf/pdf_2509.pdf. 9 พฤษภาคม 2555