

# การวัดและการประเมินผลการเรียนรู้

## Measurement and Evaluation Learning

EDI2101\_WEEK4; หลักการสร้างแบบทดสอบและการหาคุณภาพ

# ประเภทแบบทดสอบวัดพุทธิพิสัย

## 1.แบบเลือกตอบ

- 1.1 หลายตัวเลือก (Multiple choice)
- 1.2 ถูก ผิด (True-false)
- 1.3 จับคู่ (Matching)
- 1.4 จัดลำดับ (Rearrangement)
- 1.5 เติมคำตอบ (Completion)

## 2.แบบเขียนตอบ

- 2.1 ตอบสั้น (Short answer)
- 2.2 ความเรียง (Essay)

# รูปแบบการสอบวัดพุทธิพิสัย

ถ้าแบ่งประเภทของการสอบตามลักษณะของวิธีการสอบ จะแบ่งได้เป็น

- 1) การสอบปากเปล่า (Oral Test)
- 2) การสอบโดยการเขียน (Written Test)

# ขั้นตอนการสร้างแบบทดสอบ

1. วิเคราะห์หลักสูตรและสร้างตารางวิเคราะห์หลักสูตร
2. กำหนดจุดมุ่งหมายในการวัด
3. กำหนดชนิดของแบบทดสอบและศึกษาวิธีสร้าง
4. เขียนแบบทดสอบและตรวจทานแบบทดสอบ
5. ตรวจสอบคุณภาพโดยผู้เชี่ยวชาญ (IOC)
6. จัดพิมพ์แบบทดสอบฉบับทดลอง
7. ทดลองสอบ (Try Out)
8. วิเคราะห์แบบทดสอบ หาค่าความยากง่าย อำนาจจำแนกและค่าความเชื่อมั่น
9. จัดทำแบบทดสอบฉบับจริง

# การสร้างแบบทดสอบหลายตัวเลือก (Multiple choice)

ลักษณะแบบทดสอบแบบเลือกตอบ (multiple choice test) จะประกอบด้วย

1. **ตัวคำถาม (stem)** คือโจทย์หรือข้อความที่เป็นประโยคที่สมบูรณ์เป็นเพื่อการวัดความรู้ความสามารถ

2. **ตัวเลือก (options/alternative)** ตั้งแต่ 3 ตัวเลือกขึ้นไปอีก 1 ชุด รวมเป็น 1 ข้อ ในตัวเลือกนั้นจะมีทั้งคำตอบถูก (key) และ คำตอบผิด(distracter) ที่เป็นตัวลวงมาให้นักเรียนพิจารณา

# การสร้างแบบทดสอบหลายตัวเลือก (Multiple choice)

1. คำถามควรเป็นประโยคคำถามที่สมบูรณ์ เพื่อจะช่วยให้มีความชัดเจน และเข้าใจง่ายกว่าประโยคบอกเล่า เช่น

ไม่ดี : ทONGคำจัดเป็น...

ก. แร่โลหะ

ง. แร่เชื้อเพลิง

ข. แร่ไอโลหะ

จ. แร่ประกอบหิน

ค. แร่รัตนชาติ

**ดีขึ้น : ทONGคำจัดเป็นแร่ชนิดใด (ตัวเลือกคงเดิม)**

# การสร้างแบบทดสอบหลายตัวเลือก (Multiple choice)

2. เน้นจุดที่เป็นคำถามให้ชัดเจน เพื่อให้เกิดความเป็นปรนัย เช่น

ไม่ดี : สัตว์ในข้อใดแตกต่างจากพวก

ก. ช้าง      ง. วัว

ข. ล่อ      จ. ควาย

ค. ม้า

ไม่ดี เพราะไม่ได้บอกแง่มุมที่จะถามว่าต้องการเน้นด้านไหน หรือยึดสิ่งใดเป็นหลัก

**ดีขึ้น : สัตว์ในข้อใดเกิดจากการผสมพันธุ์ที่แตกต่างจากพวก (ตัวเลือกคงเดิม)**

# การสร้างแบบทดสอบหลายตัวเลือก (Multiple choice)

3. หลีกเลี่ยงคำถามที่เป็นประโยคปฏิเสธ โดยเฉพาะปฏิเสธซ้อน แต่ถ้าจำเป็นต้องใช้ ประโยคปฏิเสธควรขีดเส้นใต้ให้ชัดเจน

ข้อใด ไม่ใช่ จำนวนตรรกยะ

4. การถามคำถามจะต้องสั้น กระชับรัดและได้ใจความ ไม่ควรใช้คำฟุ่มเฟือย เช่น

ไม่ดี : การแปรงฟันสม่ำเสมอ จะช่วยให้สุขภาพภายในช่องปากสมบูรณ์ การแปรงฟัน ผิดวิธี จะทำให้เกิดสิ่งใด

ดีขึ้น : การแปรงฟันผิดวิธี จะทำให้เกิดสิ่งใด

# การสร้างแบบทดสอบหลายตัวเลือก (Multiple choice)

5. ถามในสิ่งที่มีประโยชน์ เพราะจะช่วยให้เด็กได้เรียนรู้สิ่งที่ดีงาม เป็นการปลูกฝังค่านิยมที่พึงประสงค์ กล่าวคือ ถ้าสิ่งใดเป็นสิ่งที่ดีสังคมยอมรับ ให้ถามในทางที่ดี แต่ถ้าสิ่งใดไม่ดีสังคมไม่ยอมรับ ให้ถามในทางที่ไม่ดี เช่น

ไม่ดี : การเป็นหวัดคืออะไร

**ดีขึ้น : การเป็นหวัดแสดงว่าร่างกายอยู่ในสภาพใด**

6. ถามในสิ่งที่สามารถหาข้อยุติได้ตามหลักวิชา เพื่อให้เด็กได้ใช้ความคิด ไม่ถามในสิ่งที่เป็นความเชื่อ เช่น

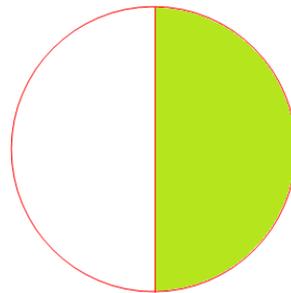
ไม่ดี : การแห่นางแมวเพื่อให้ฝนตก ควรทำพิธีเวลาใด จึงจะได้ผลดี

ไม่ดี เพราะเป็นการถามในสิ่งที่เป็นความเชื่อ ซึ่งไม่สามารถพิสูจน์ได้

# การสร้างแบบทดสอบหลายตัวเลือก (Multiple choice)

7. ควรถามพฤติกรรมที่ต้องใช้ความคิด และควรหลีกเลี่ยงการถามความจำจากตำรา
8. ใช้ภาษาให้เหมาะสมกับวัยของเด็ก เพราะถ้าใช้ภาษาที่ยากเกินไป เด็กก็จะไม่สามารถที่จะเข้าใจในความหมาย จึงไม่สามารถทำแบบทดสอบนั้นได้
9. ควรใช้คำถามที่ช่วยย้ให้เด็กชวนคิด และบางครั้งคำถามหรือตัวเลือกอาจเป็นรูปภาพ  
สัญลักษณ์ สถานการณ์ เพื่อช่วยย้ให้เด็กอยากทำ เช่น  
ไม่ดี :  $(1/2)$  มีค่าเท่ากับเศษส่วนในข้อใด

ดีขึ้น : ส่วนที่แรงามีค่าเท่าใด



# การสร้างแบบทดสอบหลายตัวเลือก (Multiple choice)

10. ตัวเลือกควรสั้น กระชับรัด และมีควมหมาย คำฟุ่มเฟือยตัดทิ้ง

11. ตัวเลือกแต่ละตัวควรมีความยาวเท่า ๆ กัน ถ้าตัวเลือกยาวไม่เท่ากัน ควรเรียงจากสั้นไปหายาว แต่ทั้งนี้ถ้าเป็น วัน เดือน พ.ศ. ตัวเลขหรือจำนวน ให้เรียงจากน้อยไปหามาก

# การสร้างแบบทดสอบหลายตัวเลือก (Multiple choice)

12. หลีกเลี่ยงการเขียนตัวเลือกซ้ำซ้อนกัน หรือมีความหมายเหมือนกัน เพราะจะทำให้ตัวเลือกมีคุณค่าลดลง

ไม่ดี : น้ำปูนใสจัดอยู่ในสารชนิดใด

ก. กรด

ง. ด่างทับทิม

ข. ด่าง

จ. กรดเกลือ

ค. กรดสารส้ม

ไม่ดี เพราะตัวเลือก ค และ จ เป็นส่วนหนึ่งของตัวเลือก ก

ส่วนตัวเลือก ง เป็นส่วนหนึ่งของตัวเลือก ข

# การสร้างแบบทดสอบหลายตัวเลือก (Multiple choice)

13. ตัวเลือกต้องเป็นเอกพันธ์กัน (homogeneity) หรือไม่ก็มีโครงสร้างสอดคล้องกัน เช่น

ไม่ดี : การทำลายป่าเกิดจากสาเหตุใด

ก. ความยากจน

ง. กฎหมายมีช่องว่าง

ข. ความเห็นแก่ตัว

จ. เจ้าหน้าที่ไม่กวัดขั้น

ค. คนต้องการที่ทำกิน

ไม่ดี เพราะตัวเลือก ก,ข เป็นเพียงวลี ส่วน ค,ง,จ เป็นประโยค  
ดังนั้น ควรทำตัวเลือกให้มีโครงสร้างที่เหมือนกัน

# การสร้างแบบทดสอบหลายตัวเลือก (Multiple choice)

14. ควรระมัดระวังการใช้ตัวเลือกประเภท ปลายเปิด และปลายปิด ซึ่งได้แก่ ถูกทุกข้อ ไม่มีข้อถูก ผิดทุกข้อ สรุปแน่นอนไม่ได้ เป็นต้น

15. หลีกเลี่ยงการเนะคำตอบ เพราะเพียงแค่เด็กสังเกตก็สามารถหาคำตอบได้โดยไม่ได้ใช้ ความคิดเลย เช่น

ข้อแรก ๆ ถามว่า สีเขียวในพืชเรียกว่าอะไร (คลอโรฟิล)

ข้อหลัง ๆ ถามว่า พืชสีเขียวซึ่งมีคลอโรฟิลจะปรุงอาหารได้ต้องอาศัยสิ่งใด

16. ไม่ควรถามในสิ่งที่เด็กคล่องปาก (คุ้นเคยและท่องได้)

# การสร้างแบบทดสอบหลายตัวเลือก (Multiple choice)

17. ควรกระจายตัวเลือกที่เป็นตัวถูกให้อยู่ในตำแหน่งที่ต่าง ๆ กัน และแต่ละตัวเลือกควรมีโอกาสเป็นตัวถูกในจำนวนที่เท่า ๆ กัน

18. หลีกเลี่ยงการเขียนตัวถูกที่ฟังเสียงหรือฟังความหมายกับตัวคำถาม เพราะจะเป็น การเน้นคำตอบ เช่น

ไม่ดี : จังหวัดใดผลิตเกลือสมุทรได้มากที่สุด

ก. อัญญา

ข. นครราชสีมา

ค. สมุทรสาคร

# ข้อดีของแบบทดสอบแบบหลายตัวเลือก (Multiple choice)

1. วัดพฤติกรรมทางการศึกษาได้หลายด้าน ตั้งแต่ความรู้ความจำไปจนถึงขั้นการประเมิน
2. เป็นแบบทดสอบที่ตรวจให้คะแนนง่าย ถูกต้อง รวดเร็ว และมีความเป็นปรนัย
3. สามารถควบคุมความยากง่ายของแบบทดสอบได้
4. เป็นแบบทดสอบที่ครูสามารถวินิจฉัยสาเหตุแห่งการทำแบบทดสอบผิดว่าเนื่องมาจากสาเหตุอะไรบ้าง โดยพิจารณาจากตัวเลือกต่าง ๆ จากกระดาษคำตอบ
5. มีความเชื่อมั่นสูง เพราะมีจำนวนแบบทดสอบมาก และตอบถูกโดยการเดามีน้อย
6. สามารถใช้สัญลักษณ์ รูปภาพหรือกราฟมาเขียนแบบทดสอบได้

# ข้อจำกัดของแบบทดสอบหลายตัวเลือก(Multiple choice)

1. สร้างแบบทดสอบให้ดี ทำได้ยาก และใช้เวลาในการสร้างนาน
2. ไม่เหมาะที่จะวัดความคิดริเริ่มสร้างสรรค์ การเสนอแนวคิด หรือทักษะการเขียน
3. สิ้นเปลืองค่าใช้จ่ายสูงในการสร้างแบบทดสอบ

# กิจกรรม 4

1. ให้นักศึกษาร่างข้อสอบแบบเลือกตอบ 4 ตัวเลือก จำนวน 5 ข้อ
2. นำข้อสอบในข้อ 1 มาสร้างเป็นแบบทดสอบใน Kahoot/ Word Wall /Quizizz



# การพิจารณาลักษณะแบบทดสอบที่ดี

1. ความเป็นปรนัย (Objectivity)
2. ความตรง (Validity)
3. ความเชื่อมั่น (Reliability)
4. ความยาก(ง่าย) (Difficulty)
5. อำนาจจำแนก (Discrimination)
6. ประสิทธิภาพตัวลวง (Distracter Efficiency)

# ความเป็นปรนัย (Objectivity)

ความเป็นปรนัย หมายความว่า ทุกฝ่ายที่เกี่ยวข้องกับการวัดผลครั้งนั้นมีความเห็นสอดคล้องกันในเรื่องของคำถาม ค่าของคะแนนหรืออันดับที่ที่วัดได้ตลอดจนการแปลงค่าคะแนนเป็นผลประเมินในการตัดสิน คุณค่าก็สอดคล้องตรงกัน

การพิจารณาความเป็นปรนัยของแบบทดสอบมีหลายประการ คุณสมบัติความเป็นปรนัยของแบบทดสอบที่สำคัญ ได้แก่คุณสมบัติ 3 ประการดังนี้

- 1) **ชัดเจนในความหมายของคำถาม** แบบทดสอบที่เป็นปรนัย ทุกคนที่อ่านแบบทดสอบไม่ว่าจะเป็นผู้สอบหรือผู้ตรวจแบบทดสอบย่อมจะเข้าใจตรงกันไม่ตีความไปคนละทาง
- 2) **ตรวจให้คะแนนได้ตรงกัน** แบบทดสอบที่มีความเป็นปรนัย ไม่ว่าจะเป็นผู้ออกแบบทดสอบหรือใครก็ตามสามารถตรวจให้คะแนนได้ตรงกัน แบบทดสอบที่ผู้ตรวจเฉลยไม่ตรงกัน แสดงให้เห็นถึงความไม่ชัดเจนในคำถามหรือคำตอบ
- 3) **แปลความหมายของคะแนนได้ตรงกัน** โดยทั่วไปแบบทดสอบปรนัยนั้นผู้ตอบถูกจะได้ 1 คะแนน ตอบผิดจะได้ศูนย์คะแนน

# วิธีการวิเคราะห์คุณภาพแบบทดสอบ

2.1 การวิเคราะห์เครื่องมือวัดผลโดยไม่ใช่วิธีการทางสถิติ

2.2 การวิเคราะห์เครื่องมือวัดผลโดยใช้วิธีการทางสถิติ

# การวิเคราะห์เครื่องมือวัดโดยไม่ใช้วิธีการทางสถิติ

การตรวจสอบความถูกต้องและครอบคลุมของเนื้อหาวิชาและจุดมุ่งหมาย

แนวทางในการพิจารณา

- 1) ข้อคำถามครบถ้วนทุกเนื้อหาที่เรียนหรือไม่
- 2) จำนวนข้อคำถามของแต่ละเนื้อหา มีสัดส่วนตามน้ำหนักที่กำหนดไว้หรือไม่
- 3) ข้อคำถามแต่ละข้อวัดได้ตรงตามพฤติกรรมที่ระบุไว้ในตัวชี้วัดหรือไม่

วิธีการตรวจสอบ

- 1) ตรวจสอบโดยผู้เชี่ยวชาญในเนื้อหาวิชานั้น ๆ
- 2) ตรวจสอบโดยการเปรียบเทียบตารางกำหนดจำนวนข้อคำถาม (Test Blueprint)

# การหาคุณภาพของเครื่องมือวัดโดยใช้สถิติ

## รายชื่อ

- ค่าดัชนีความสอดคล้อง (IOC)
- ค่าความยากง่าย (p)
- อำนาจจำแนก (r)
- ประสิทธิภาพตัวลวง
- ความเป็นปรนัย

## ทั้งฉบับ

- ความตรง (Validity)
- ความเชื่อมั่น/ความเที่ยง (Reliability)

# การวิเคราะห์ความตรงโดยใช้วิธีการทางสถิติ

ให้ผู้เชี่ยวชาญตั้งแต่ 3 คนขึ้นไป ประเมินความสอดคล้องระหว่างข้อคำถามใน เครื่องมือกับเนื้อหาที่ต้องการวัด จากนั้นนำผลการประเมินมาคำนวณค่าดัชนีความสอดคล้อง ระหว่างข้อสอบกับจุดประสงค์ (IOC-Index of Item Objective Congruence) โดยใช้สูตร

$$IOC = \frac{\sum R}{N}$$

เมื่อ  $\sum R$  แทน ผลรวมของคะแนนจากผู้เชี่ยวชาญ

เกณฑ์ตัดสิน IOC ควรมีค่ามากกว่า 0.5

# ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์ (IOC-Index of Item Objective Congruence)

## ตัวอย่างตารางการหาค่า IOC

ตัวชี้วัดตามมาตรฐาน	ข้อสอบ	ผลการประเมิน		
		-1	0	1
ผู้เรียนสามารถบอกถึงหน้าที่หรือความแตกต่างของส่วนประกอบต่างๆของพืชได้	<b>1. หน้าที่ของใบคืออะไร?</b> ก. ยึดลำต้น ข. ดูดอาหาร ค. สังเคราะห์แสง ง. ลำเลียงอาหาร			
	<b>2. ข้อใดเป็นพืชใบเลี้ยงคู่?</b> ก. ข้าว ข. อ้อย ค. กัลฉ่าย ง. มะเขือ			

- 1 หมายถึง ไม่สอดคล้อง

0 หมายถึง ไม่แน่ใจ

1 หมายถึง สอดคล้อง

# ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์ (IOC-Index of Item Objective Congruence)

## ตัวอย่างตารางการหาค่า IOC

ข้อคำถาม	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	IOC
ข้อ 1	1	1	1	0	1	
ข้อ 2	1	0	-1	0	-1	

เกณฑ์การพิจารณา ข้อสอบที่ใช้ได้ คือ ข้อสอบที่มีค่า IOC ตั้งแต่ .5 ขึ้นไป

สรุป

ข้อสอบข้อ 1 IOC= 0.8 นำไปใช้ทดสอบได้

ข้อสอบข้อ 2 IOC = -0.2 ต้องปรับปรุง

# ค่าระดับความยากง่าย (Difficulty Index)

ระดับความยากง่าย หมายถึง สัดส่วน หรือเปอร์เซ็นต์ของจำนวนคนที่ตอบข้อสอบข้อนั้นถูกต้องจากคนที่สอบทั้งหมด  
ใช้สัญลักษณ์ "p"

ข้อสอบแบบเลือกตอบ (คะแนนแบบทวิภาค 0 กับ 1)

ข้อสอบแบบเขียนตอบ (คะแนนแบบพหุภาค มากกว่า 2 ค่า)

# ค่าระดับความยากง่าย (Difficulty Index)

ค่าความยากง่าย (Difficulty Index)

มีค่าตั้งแต่ 0.00 จนถึง 1.00

เกณฑ์ในการแปลความหมายค่าความยากง่าย

ค่า $p = 0.00-0.19$	หมายความว่า ข้อสอบข้อนั้นยากเกินไป
ค่า $p = 0.20-0.39$	หมายความว่า ข้อสอบข้อนั้นค่อนข้างยาก
ค่า $p = 0.40-0.59$	หมายความว่า ข้อสอบข้อนั้นยากง่ายปานกลาง
ค่า $p = 0.60-0.79$	หมายความว่า ข้อสอบข้อนั้นค่อนข้างง่าย
ค่า $p = 0.80-1.00$	หมายความว่า ข้อสอบข้อนั้นง่ายเกินไป

เกณฑ์: ข้อสอบที่มีค่าความยากง่ายพอเหมาะ หรือมีคุณภาพดี

ค่า  $p$  ใกล้เคียง .50 หรือ อยู่ระหว่าง 0.20 – 0.80

## (1) ค่าระดับความยากง่าย (Difficulty Index) ข้อสอบแบบเลือกตอบ

$$p = \frac{R_H + R_L}{N_H + N_L}$$

$R$	แทน	จำนวนคนที่ตอบข้อนั้นถูก
$N$	แทน	จำนวนคนที่สอบทั้งหมด
$R_H$	แทน	จำนวนคนที่ตอบข้อนั้นถูกในกลุ่มสูง
$R_L$	แทน	จำนวนคนที่ตอบข้อนั้นถูกในกลุ่มต่ำ
$N_H$	แทน	จำนวนคนในกลุ่มสูง
$N_L$	แทน	จำนวนคนในกลุ่มต่ำ

## ตัวอย่างตารางการหาค่าระดับความยากง่าย

ข้อ		กลุ่มสูง ( $R_H$ ) (20 คน)	กลุ่มต่ำ ( $R_L$ ) (20 คน)	
1	ก	4	6	
	ข*	9	3	<b>12/40=0.3</b> <b>นำไปใช้ได้/ค่อนข้างยาก</b>
	ค	3	5	
	ง	6	4	
	รวม	20	20	

## (2) ค่าระดับความยากง่าย (Difficulty Index) ข้อสอบแบบเขียนตอบ

1. ตรวจสอบและเรียงคะแนนรวมจากสูงที่สุดถึงต่ำสุด
2. แบ่งกลุ่มสูง (H) และกลุ่มต่ำ (L)
3. คำนวณสัดส่วนของคะแนนรวมรายข้อที่ได้จำแนกตามกลุ่ม

$$P_H = \frac{\Sigma H}{\Sigma T_H} \quad P_L = \frac{\Sigma L}{\Sigma T_L}$$

$\Sigma H$  รวมคะแนนกลุ่มสูง

$\Sigma T_H$  รวมคะแนนเต็มกลุ่มสูง

$\Sigma L$  รวมคะแนนกลุ่มต่ำ

$\Sigma T_L$  รวมคะแนนเต็มกลุ่มต่ำ

4. วิเคราะห์ค่าความยากง่าย (p)

## ตัวอย่างตารางการหาค่าระดับความยากง่าย

ข้อ	คะแนนเต็ม	กลุ่มสูง (H) (4 คน)				กลุ่มต่ำ (L) (4 คน)			
		1	2	3	4	5	6	7	8
1	10	10	10	9	8	5	8	8	7
2	10	9	10	8	9	8	7	6	3
3	20	20	15	15	17	15	9	10	8
4	30	25	25	24	20	16	17	13	10
5	30	16	10	10	7	11	7	6	2
รวม	100	80	70	68	61	60	48	43	30

## ตัวอย่างตารางการหาค่าระดับความยากง่าย

ข้อ	คะแนนเต็ม	กลุ่มสูง (4 คน)		กลุ่มต่ำ (4 คน)		p <sub>H</sub>	p <sub>L</sub>	p <sub>i</sub>
		ได้	เต็ม	ได้	เต็ม			
1	10	37	40	28	40	$37/40=.93$	$28/40=.70$	$65/80=.82$ $(0.93+0.7)/2=0.82$
2	10	36	40	24	40			
3	20	67	80	42	80			
4	30	94	120	56	120			
5	30	43	120	26	120			

# อำนาจจำแนก (Discrimination)

ความสามารถในการจำแนกผู้สอบ ตามระดับความสามารถ

(Item Discrimination , r)

เป็นคุณสมบัติที่บ่งบอกถึงความสามารถของข้อสอบ ที่จำแนกเด็กเก่ง-อ่อน  
จะมีค่า r เป็นตัวบ่งบอกว่าข้อสอบข้อใดมีค่า r สูง ๆ แสดงว่าข้อสอบข้อนั้นจำแนกได้ดี  
หมายถึง ข้อสอบข้อนี้ใครทำถูกจะเป็นพวกกลุ่มคนเก่ง ใครทำผิดจะเป็นพวกกลุ่มอ่อน

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

r	คือ อำนาจจำแนก
$R_H$	คือ จำนวนผู้สอบที่ตอบถูกในกลุ่มคะแนนสูง
$R_L$	คือ จำนวนผู้สอบที่ตอบถูกในกลุ่มคะแนนต่ำ
$N_{H/L}$	คือ จำนวนผู้สอบทั้งหมดในกลุ่มคะแนนสูงหรือกลุ่มคะแนนต่ำ

# อำนาจจำแนก (Discrimination)

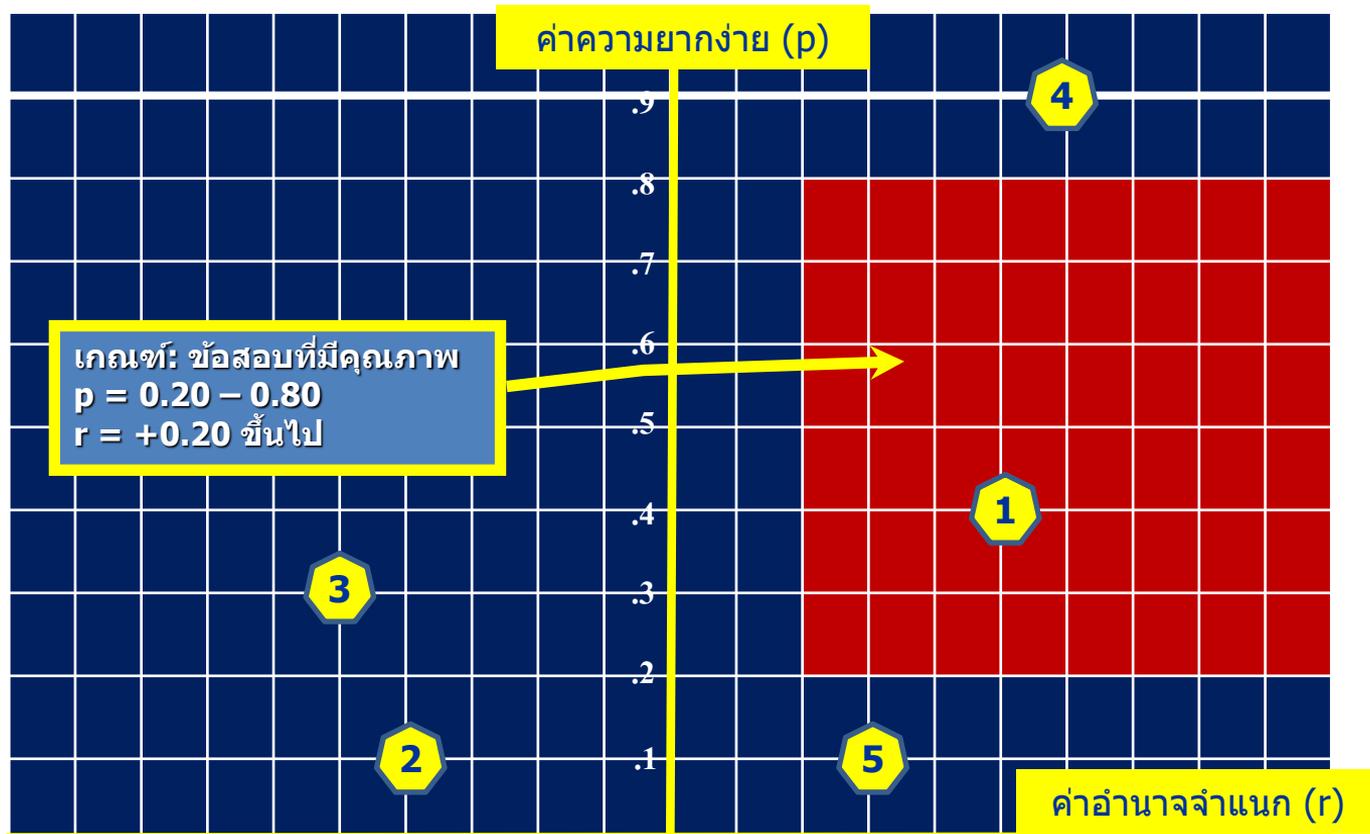
ค่าอำนาจจำแนก	ความหมาย
1.00	จำแนกดีเลิศ
0.80-0.99	จำแนกดีมาก
0.60-0.79	จำแนกดี
0.40-0.59	จำแนกได้ปานกลาง
0.20-0.39	จำแนกได้บ้าง
0.00-0.19	จำแนกไม่ค่อยได้ (ควรปรับปรุง)
<0	จำแนกไม่ได้ ควรตัดทิ้งสร้างข้อสอบใหม่

# เกณฑ์ในการสรุปว่าข้อสอบมีคุณภาพดี

จงหาค่าความยากง่าย และค่าอำนาจจำแนก

ข้อที่	ตัวเลือก	กลุ่มสูง (H=20)	กลุ่มต่ำ (L=20)	p	r	ความหมาย	สรุป
<b>1</b>	<b>ก</b>	<b>3</b>	<b>4</b>				
	<b>(ข)</b>	<b>13</b>	<b>6</b>	$(13+6)/40=0.48$	$(13-6)/20=0.35$	-ยากง่ายปานกลาง -อำนาจจำแนกได้บ้าง	นำไปใช้ทดสอบได้
	<b>ค</b>	<b>-</b>	<b>3</b>				
	<b>ง</b>	<b>2</b>	<b>4</b>				
	<b>จ</b>	<b>2</b>	<b>3</b>				

# การคัดเลือกข้อสอบที่มีคุณภาพ



# ประสิทธิภาพของตัวลวง (Distracter Efficiency)

1. อำนาจจำแนกของตัวลวง  $r_w$  หมายถึง ค่าที่แสดงถึงการเลือกตัวเลือกนั้นๆว่า  
ได้กลุ่มคะแนนต่ำลงตอบมากกว่ากลุ่มคะแนนสูง  
ตัวลวงที่ดีควรมีค่า  $r_w$  เป็นบวก ควรมีค่าตั้งแต่ .05 ขึ้นไป

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

$r_w$	คือ อำนาจจำแนกของตัวลวง
$R_L$	คือ จำนวนผู้สอบในกลุ่มอ่อน (คะแนนต่ำ) ที่เลือกตอบตัวลวงข้อนั้น
$R_H$	คือ จำนวนผู้สอบในกลุ่มเก่ง (คะแนนสูง) ที่เลือกตอบตัวลวงข้อนั้น
$N_{H/L}$	คือ จำนวนผู้สอบทั้งหมดในกลุ่มคะแนนสูงหรือกลุ่มคะแนนต่ำ

# ประสิทธิภาพของตัวลวง (Distracter Efficiency)

2. สัดส่วนของตัวลวง  $p_w$  หมายถึง สัดส่วนของจำนวนคนที่เลือกตัวลวงนั้นๆ  
ตัวลวงที่ดีควรมีค่า  $p_w$  ควรมีค่าตั้งแต่ .05 ขึ้นไป

$$p_w = \frac{R_L + R_H}{N}$$

$p_w$	คือ อำนาจจำแนกของตัวลวง
$R_L$	คือ จำนวนผู้สอบในกลุ่มอ่อน (คะแนนต่ำ) ที่เลือกตอบตัวลวงข้อนั้น
$R_H$	คือ จำนวนผู้สอบในกลุ่มเก่ง (คะแนนสูง) ที่เลือกตอบตัวลวงข้อนั้น
$N$	คือ จำนวนผู้สอบทั้งหมด

## ตัวอย่าง : ข้อที่ 1

P = ค่าความยาก r = จำนวนจำแนก \* = คำตอบที่ถูกต้อง

ข้อสอบ	ตัวเลือก	กลุ่มสูง	กลุ่มต่ำ	P	r	ผลการวิเคราะห์
1	1	2	3	0.079	0.031	คำตอบลวง ค่าอำนาจจำแนกต่ำ
	*2	28	18	0.719	0.313	ข้อที่ค่อนข้างง่าย ค่าอำนาจการจำแนกดี
	3	2	10	0.188	0.250	คำตอบลวง ค่าอำนาจการจำแนกพอใช้
	4	0	1	0.013	0.031	คำตอบลวง ค่าอำนาจการจำแนกต่ำ
	5	0	0	0.000	0.000	คำตอบลวง ไม่มีค่าอำนาจการจำแนก

วิเคราะห์เป็นข้อสอบที่มีคุณภาพดี ถ้าดูข้อถูก \*2 ค่า P = 0.719 แสดงว่าค่อนข้างไปทางง่าย ค่า r = 0.313 ดี แสดงว่าแยกคนเก่งกับคนอ่อนได้ ควรเก็บไว้

# ความตรง/ความเที่ยงตรง (Validity)

ความตรง/ความเที่ยงตรง (Validity) หมายถึง ความสามารถของเครื่องมือที่วัดได้ตรงและครอบคลุมเนื้อหาตามที่ต้องการวัด อาจวัดได้ทั้งความตรงเชิงเนื้อหา และความตรงเชิงโครงสร้าง

ความเที่ยงตรงของแบบทดสอบ แบ่งเป็น 4 ชนิด

- 1. ความเที่ยงตรงตามเนื้อหา (content Validity)** หมายถึงแบบทดสอบนั้นมีคำถามสอดคล้องและครอบคลุมเนื้อหาวิชาตามที่ระบุไว้ในหลักสูตร และได้สัดส่วนที่ถูกต้องตรงกับความจริง ซึ่งเราสามารถตรวจสอบดูได้จากการนำไปเทียบกับตารางวิเคราะห์หลักสูตรที่ทำไว้ในด้านเนื้อหาวิชา
- 2. ความเที่ยงตรงตามโครงสร้าง (Construct Validity)** หมายถึง ความสามารถของแบบทดสอบที่จะวัดสมรรถภาพของสมองหรือพฤติกรรมในด้านต่าง ๆ ของผู้เรียนได้ตรงตามที่ระบุไว้ในหลักสูตร ในภาคความมุ่งหมาย
- 3. ความเที่ยงตรงตามสภาพ (Concurrent Validity)** หมายถึง ความสามารถของแบบทดสอบที่สามารถเร้าให้นักเรียนตอบสนองออกมาตรงตามสภาพความเป็นจริงของเขา เกณฑ์ที่ใช้เทียบก็คือสภาพความเป็นจริงในปัจจุบันของนักเรียน
- 4. ความเที่ยงตรงตามพยากรณ์ (Predictive Validity)** หมายถึงความสามารถพยากรณ์ผลการเรียนในอนาคตของนักเรียนได้ถูกต้องตามความเป็นจริง เกณฑ์ที่ใช้เทียบคือสภาพความเป็นจริงหรือสภาพความสำเร็จในอนาคตของผู้เรียน

# วิธีตรวจสอบความตรงเชิงโครงสร้าง

1. ตรวจสอบความสัมพันธ์กับเครื่องมือที่มีโครงสร้างเหมือนกัน โดยคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่าย
2. ตรวจสอบด้วยการวิเคราะห์องค์ประกอบ (Factor Analysis)
3. ตรวจสอบด้วยการเทียบกับกลุ่มที่รู้ชัด โดยใช้สถิติทดสอบ

# ค่าความเชื่อมั่น ( reliability)/ความเที่ยง

ความเชื่อมั่น/ความเที่ยง หมายถึง ความคงเส้นคงวา ของคะแนนในการวัดแต่ละครั้ง หรือกล่าวได้ว่า เครื่องมือนั้นวัดครั้งใด ๆ ก็ได้ค่าเท่าเดิมไม่เปลี่ยนแปลง

การหาค่าความเชื่อมั่นได้จึงยึดหลักการสอบหลายๆ ครั้ง แล้วหาความสัมพันธ์ของคะแนนที่ได้จากการสอบหลายครั้งนั้น ถ้าคะแนนของนักเรียนแต่ละคนคงที่หรือขึ้นลงตามกันแสดงว่าแบบทดสอบนั้นมีค่าความเชื่อมั่นสูง

## เกณฑ์การแปลผล

ค่าความเชื่อมั่นของเครื่องมืออยู่ระหว่าง 0.00 – 1.00 ยิ่งใกล้ 1.00 ยิ่งมีความเชื่อมั่นสูง

เกณฑ์การแปลผลความเชื่อมั่นมีดังนี้

0.00 – 0.20 ความเชื่อมั่นต่ำมาก/ไม่มีเลย

0.21 – 0.40 ความเชื่อมั่นต่ำ

0.41 – 0.70 ความเชื่อมั่นปานกลาง

0.71 – 1.00 ความเชื่อมั่นสูง

# วิธีตรวจสอบค่าความเชื่อมั่น/ความเที่ยง (Reliability)

1. การสอบซ้ำ (Test-retest method)
2. การวัดความสอดคล้องภายใน (Measure of Internal Consistency)
  1. วิธีแบ่งครึ่งข้อสอบ (Split-half method)
  2. วิธีของ Kuder-Richardson (Kr20 , Kr21) สำหรับแบบสอบที่ให้คะแนนแบบ 0-1
3. วิธีสัมประสิทธิ์แอลฟาของ Cronbach (Cronbach's alpha  $\alpha$  method) สำหรับแบบสอบที่ให้คะแนนแบบ 0-1 หรือ มากกว่า 1