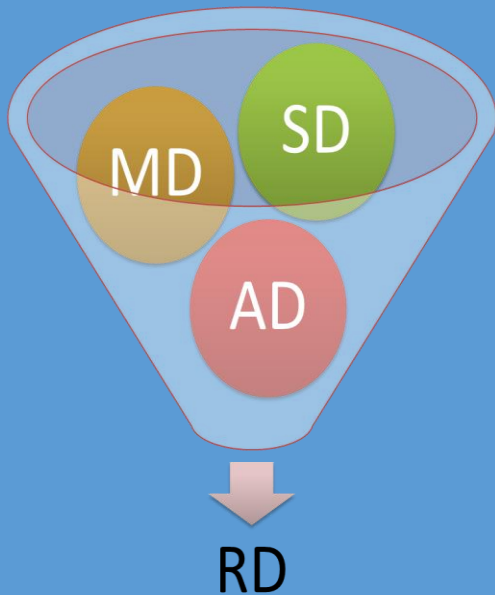


การหาคุณภาพ เครื่องมือวิจัย

งานประชุมปฏิบัติการพัฒนาศักยภาพนักวิจัยประเมินผล
โครงการ “เลี้ยงดูลูกตามคำสอนพ่อ”ฯ

ศจีมาจ ณ วิเชียร



การตรวจสอบคุณภาพเครื่องมือวิจัย

คุณภาพเครื่องมือทั้งฉบับ

คุณภาพเครื่องมือ

ความตรง, ความเที่ยงตรง (validity)

น้ำหนักจริงๆ น้หนัก 1 กิโลกรัม

ตาชั่ง ชั่งได้หนัก 1 กิโลกรัม



สรุปได้ว่า
ตราชั่งให้ข้อมูลที่มีความตรง

คุณภาพเครื่องมือ

ความตรง, ความเที่ยงตรง (validity)

น้ำหนักจริงๆ น้ำหนัก 1 กิโลกรัม

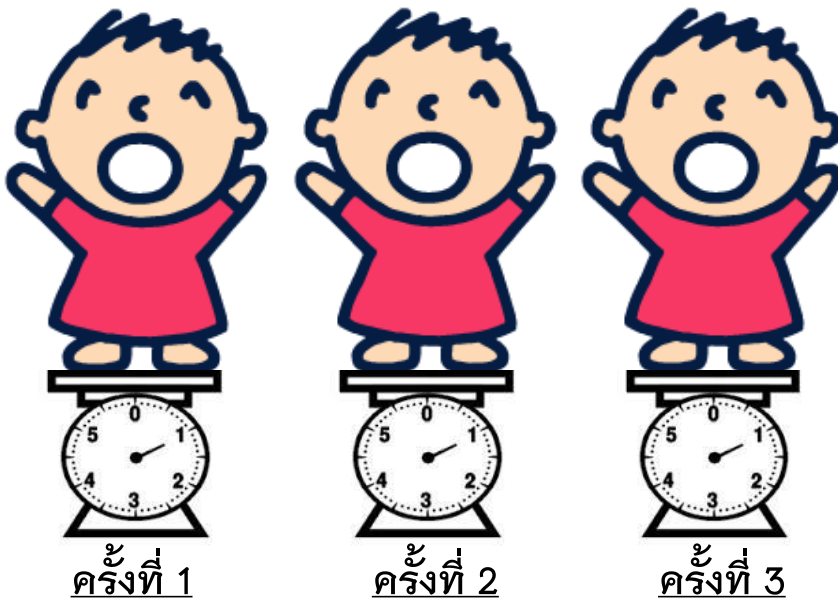
ตาชั่ง ชั่งได้หนัก 2 กิโลกรัม



สรุปได้ว่า
ตาชั่งให้ข้อมูลที่ไม่มีความ
ตรง

คุณภาพเครื่องมือ

ความเที่ยง, ความเชื่อมั่น (reliability)



ตราซังให้ข้อมูลที่เหมือนกัน
ทุกช่วงเวลา

สรุปได้ว่า
ตราซังให้ข้อมูลที่มีความเที่ยง

คุณภาพเครื่องมือ

ความเที่ยง, ความเชื่อมั่น (reliability)



ครั้งที่ 1



ครั้งที่ 2



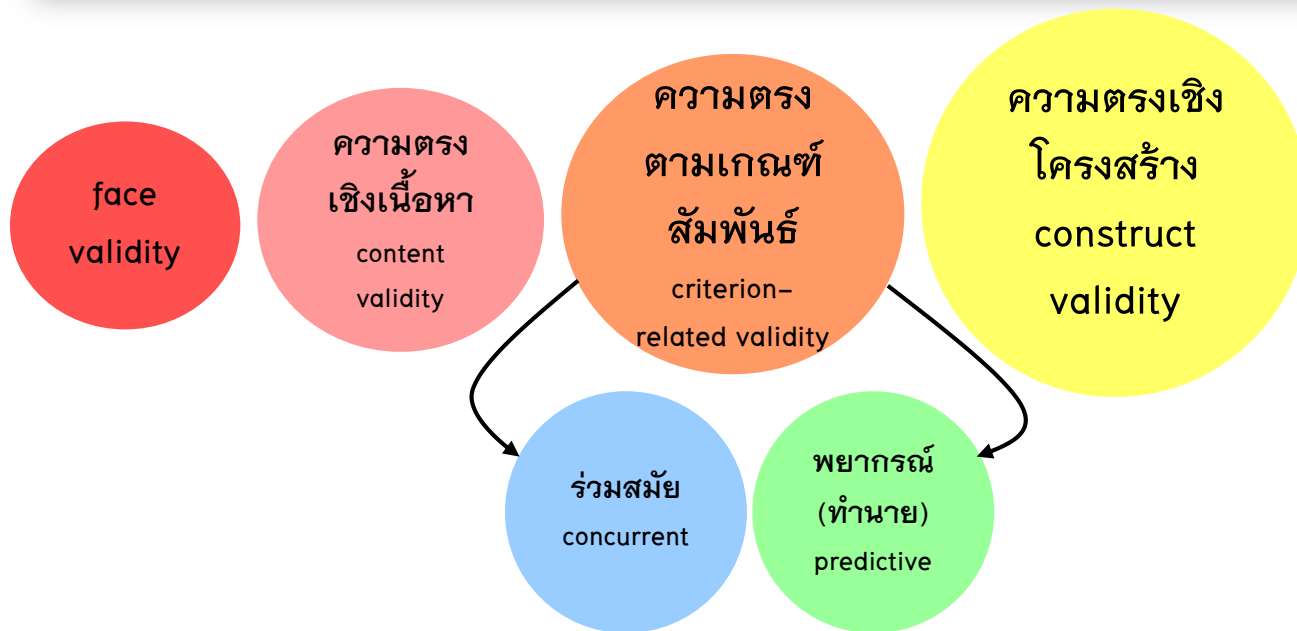
ครั้งที่ 3

ตราซังให้ข้อมูลที่ไม่เหมือนกัน
ในแต่ละช่วงเวลา

สรุปได้ว่า
ตราซังให้ข้อมูลที่ไม่มีความ
เที่ยง

คุณภาพเครื่องมือ

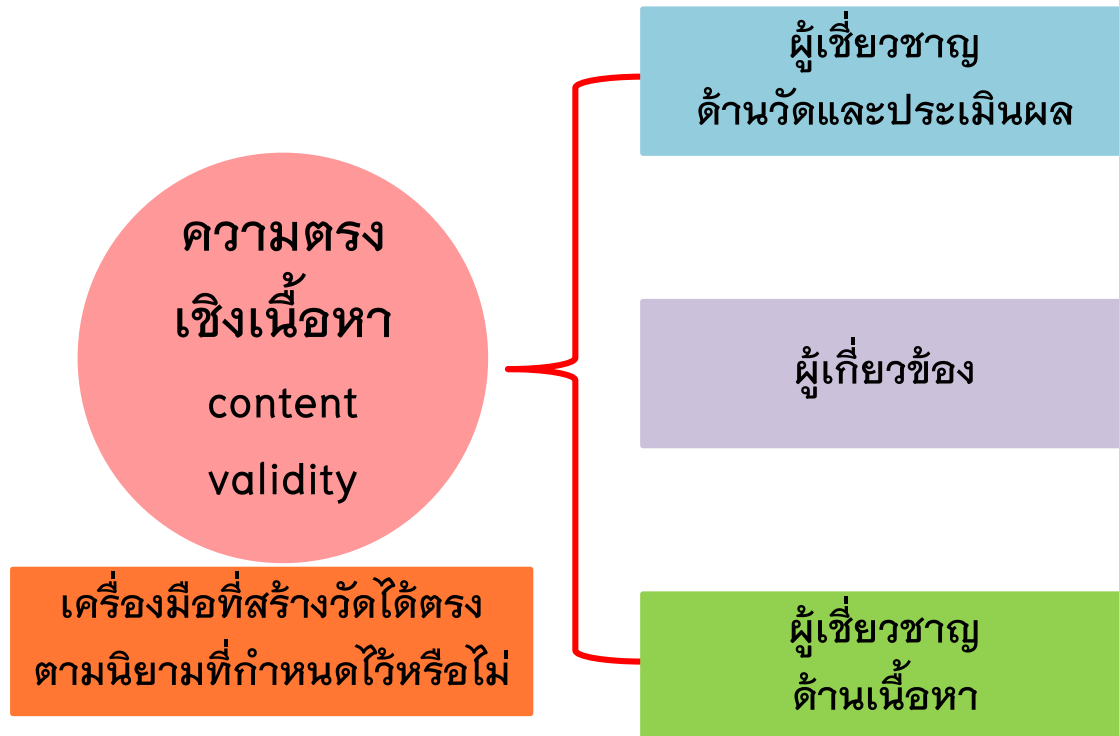
ความตรง, ความเที่ยงตรง (validity)



คุณภาพเครื่องมือ



คุณภาพเครื่องมือ



คุณภาพเครื่องมือ

ความตรง
เชิงเนื้อหา
content
validity

ความซื่อสัตย์ คือ การประพฤติตรง ไม่เอนเอียง ไม่มีเล่ห์เหลี่ยม มีความจริงใจ ปลอดจากความรู้สึกลำเอียงหรืออคติ

1. ไม่ลอกข้อสอบ



2. ให้ข้อมูลที่ถูกต้องกับทุกคน



3. เมื่อพบเจอของมีค่า จะนำไปส่งคืนเจ้าของ



4. พาเพื่อนไปเลี้ยงข้าวเป็นประจำ



คุณภาพเครื่องมือ

ความตรง
เชิงเนื้อหา
content
validity

Index of
Item-
Objective
Congruence

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

IOC = ดัชนีความสอดคล้องในความคิดเห็นของผู้เชี่ยวชาญ
 $\sum R$ = ผลรวมคะแนนความคิดเห็นของผู้เชี่ยวชาญ
N = จำนวนผู้เชี่ยวชาญ

คุณภาพเครื่องมือ

ความตรง
เชิงเนื้อหา
content
validity

ความซื่อสัตย์ คือ การประพฤติตรง ไม่
เอินเอียง ไม่มีเล่ห์เหลี่ยม มีความจริงใจ
ปลอดจากความรู้สึกลำเอียงหรืออคติ

Index of
Item-
Objective
Congruence

ข้อที่	ความคิดเห็น			ข้อคิดเห็น
	-1	0	+1	
1. ไม่ลอกข้อสอบ			✓	
2. ให้ข้อมูลที่ถูกต้องกับทุกคน			✓	
3. เมื่อพบเจอของมีค่า จะนำไปส่งคืน เจ้าของ			✓	
4. พาเพื่อนไปเลี้ยงข้าวเป็นประจำ	✓			

คุณภาพเครื่องมือ

ความตรง
เชิงเนื้อหา
content
validity

ความซื่อสัตย์ คือ การประพฤติตรง ไม่
เอินเอียง ไม่มีเล่ห์เหลี่ยม มีความจริงใจ
ปลอดจากความรู้สึกลำเอียงหรืออคติ

Index of
Item-
Objective
Congruence

ข้อที่	ผู้เชี่ยวชาญ			ค่า IOC	แปลผล
	1	2	3		
1. ไม่ลอกข้อสอบ	+1	+1	+1	1.00	ใช้ได้
2. ให้ข้อมูลที่ถูกต้องกับทุกคน	+1	0	+1	0.67	ใช้ได้
3. เมื่อพบเจอของมีค่า จะนำไป ส่งคืนเจ้าของ	+1	+1	+1	1.00	ใช้ได้
4. พาเพื่อนไปเลี้ยงข้าวเป็นประจำ	-1	-1	-1	-1.00	ใช้ไม่ได้

คุณภาพเครื่องมือ

ร่วมสมัย
concurrent

เป็นความตรงตามเกณฑ์สัมพันธ์กับสมรรถนะการดำเนินงานในปัจจุบัน การประมาณค่าความตรงตามสภาพของแบบสอบให้ความสนใจประมาณสถานภาพปัจจุบัน

คะแนนจากแบบสอบ X
(Test performance)

คะแนนเกณฑ์ Y

(Criterion performance)

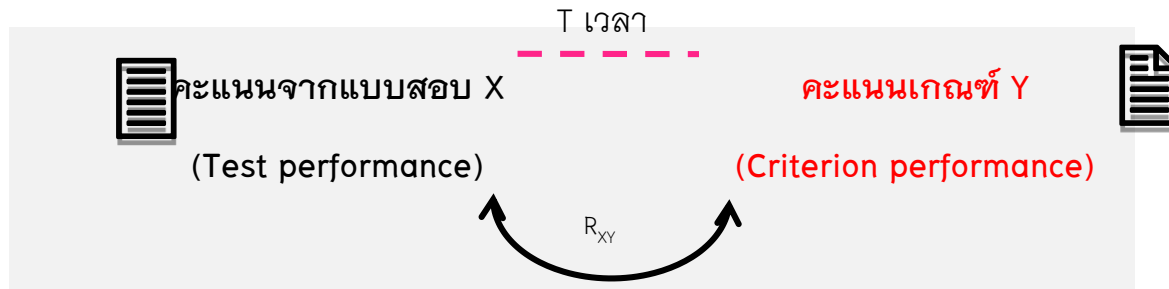


R_{XY}

คุณภาพเครื่องมือ

พยากรณ์
(ทำนาย)
predictive

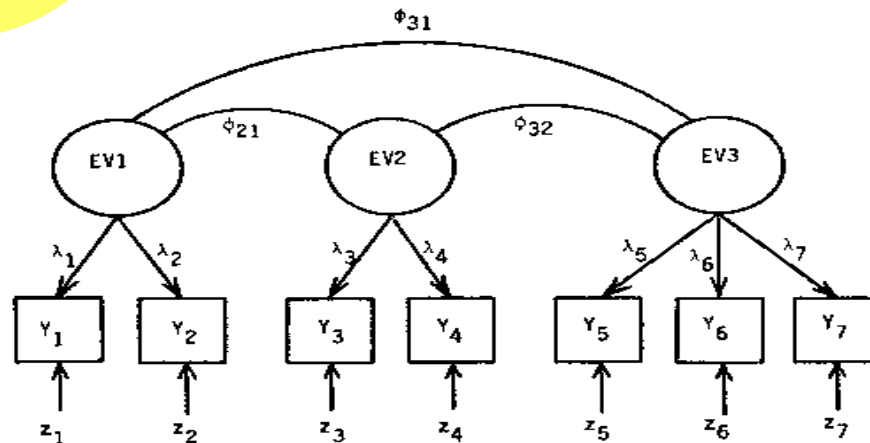
เป็นความตรงตามเกณฑ์สัมพันธ์กับสมรรถนะการดำเนินงานในอนาคต การประมาณค่าความตรงตามสภาพของแบบสอบให้ความสนใจประมาณสถานภาพอนาคต



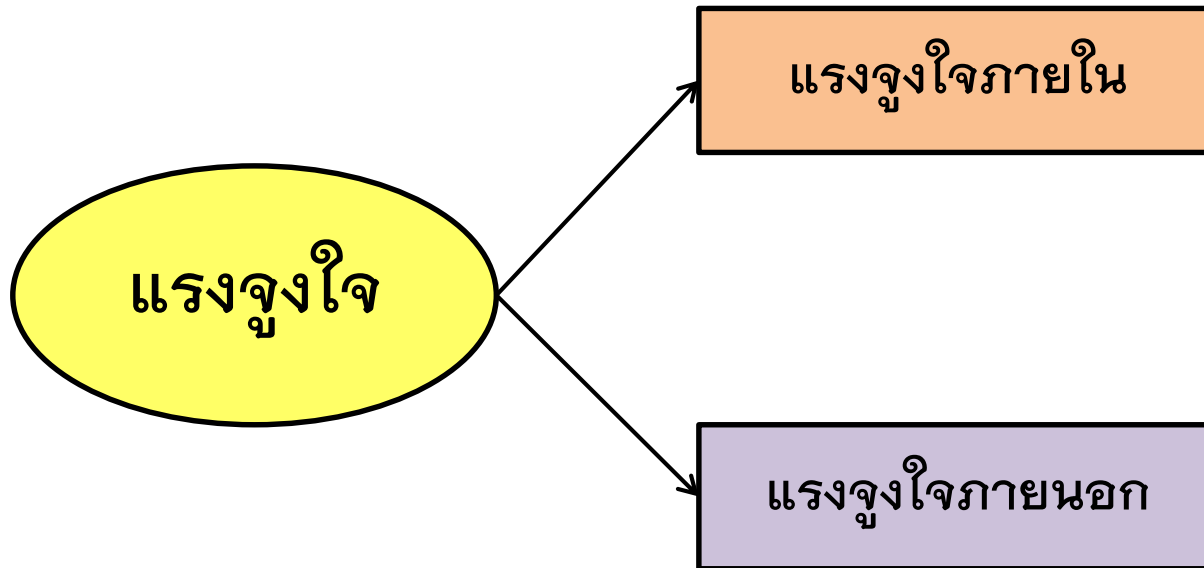
คุณภาพเครื่องมือ

ความตรงเชิง
โครงสร้าง
construct
validity

การตรวจสอบความสอดคล้องระหว่างการ
วัดตัวแปร (องค์ประกอบ) ตามทฤษฎี กับ
ข้อมูลเชิงประจักษ์



คุณภาพเครื่องมือ



คุณภาพเครื่องมือ

ความเที่ยง, ความเชื่อมั่น (reliability)

คงที่

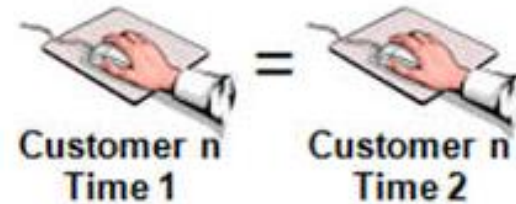
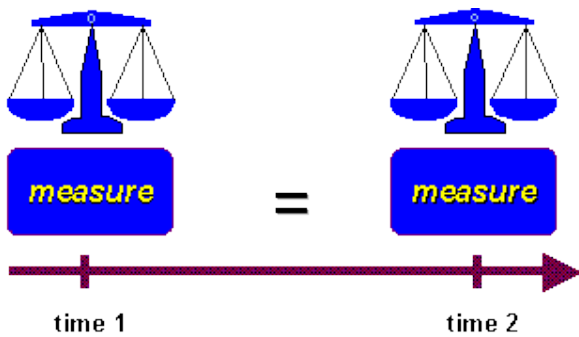
เท่า
เทียม

สอดคล้อง

คุณภาพเครื่องมือ

คง
ที่

ความคงเส้นคงวาของคะแนนจากการวัด
ในช่วงเวลาที่ต่างกันโดยวิธีการสอบซ้ำ
ด้วยแบบสอบเดิม (test-retest method)



คุณภาพเครื่องมือ

คง
ที่

test-retest method นำคะแนนที่ได้จากการ
สอบครั้งแรก (x) กับ ครั้งหลัง (x') มาคำนวณ
สัมประสิทธิ์แบบเพียร์สัน (Pearson product
moment correlation coefficient)

$$\rho_{xx'} = \frac{n \sum xx' - \sum x \sum x'}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum (x')^2 - (\sum x')^2]}}$$

เมื่อ

$\rho_{xx'}$ = สัมประสิทธิ์ความเที่ยง

n = จำนวนผู้เข้าสอบ

x และ x' = คะแนนสอบครั้งที่ 1 และ 2 ตามลำดับ

คุณภาพเครื่องมือ

เท่า
เทียม

ความสอดคล้องกันของคะแนนจากการวัด
ในช่วงเวลาที่เดียวกัน โดยใช้แบบสอบที่
สมมูลกัน (equivalent-forms method)

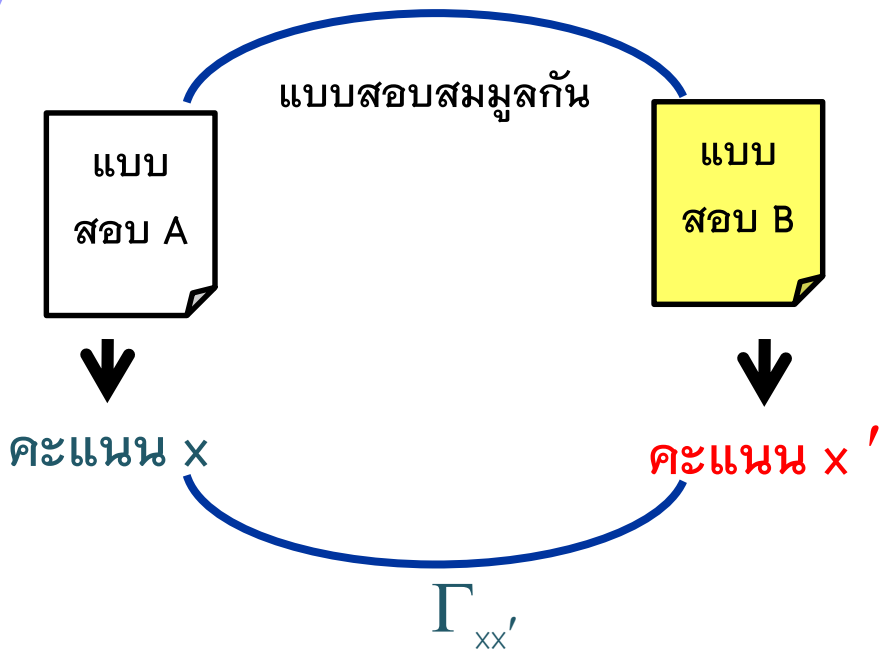


ช่วงเวลาที่ 1

คุณภาพเครื่องมือ

เท่าเทียม

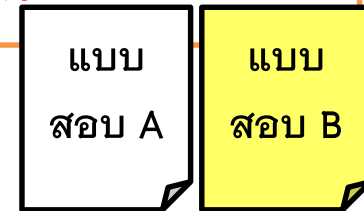
นำคะแนนที่ได้จากแบบสอบทั้งสองฉบับมาคำนวณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน



คุณภาพเครื่องมือ



ตัวอย่าง แบบสอบที่สมมูลกัน



ฉบับที่ 1

ข้อใดคือความหมายของการวัด

- ก. การกำหนดเกณฑ์ให้กับสิ่งของ
- ข. การกำหนดตัวเลขให้กับสิ่งของ
- ค. การตัดสินคุณค่าของสิ่งของ
- ง. การประมาณค่าของสิ่งของ

ฉบับที่ 2

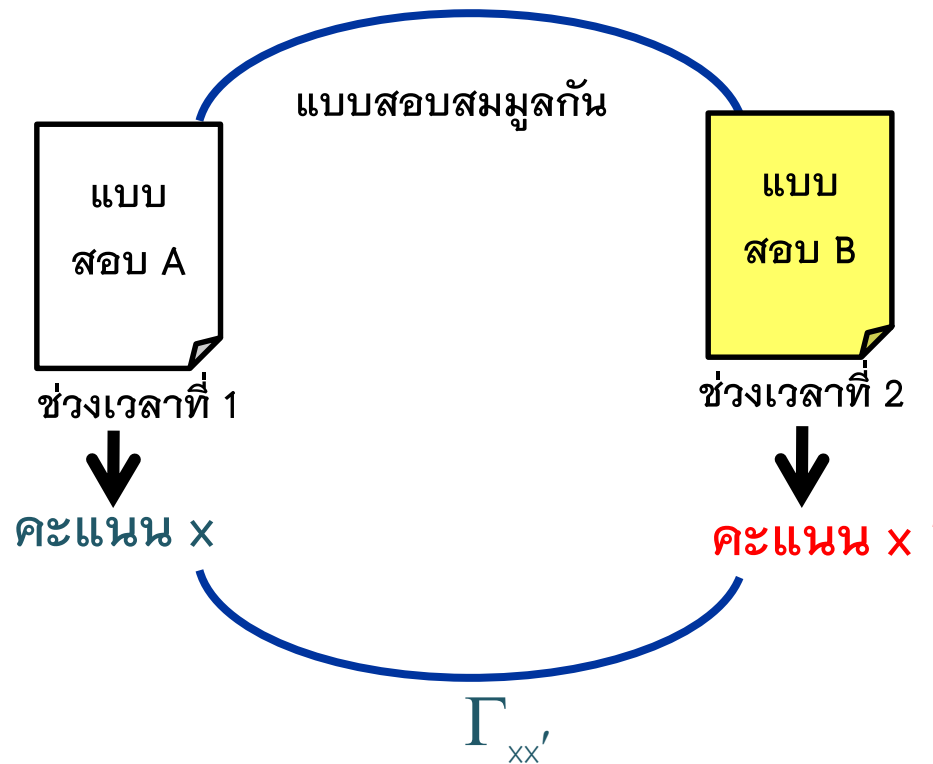
“การกำหนดตัวเลขให้กับสิ่งของ”

คือ ความหมายของค่าในข้อใด

- ก. การสร้างเกณฑ์
- ข. การวัด
- ค. การประเมิน
- ง. การประมาณ

คุณภาพเครื่องมือ

วิธีสอบซ้ำด้วยแบบสอบสมมูล (test-retest with equivalent-forms method)



คุณภาพเครื่องมือ

สอดคล้อง

ความเป็นเอกพันธ์ของเนื้อหา รายข้ออันเป็นตัวแทนของคุณลักษณะใด คุณลักษณะหนึ่ง (measure of internal consistency)

ความสอดคล้อง

1. ไม่ลอกข้อสอบ

2. ให้ข้อมูลที่ถูกต้องกับทุกคน

3. เมื่อพบเจอของมีค่า จะนำไปส่งคืนเจ้าของ

คุณภาพเครื่องมือ

สอดคล้อง

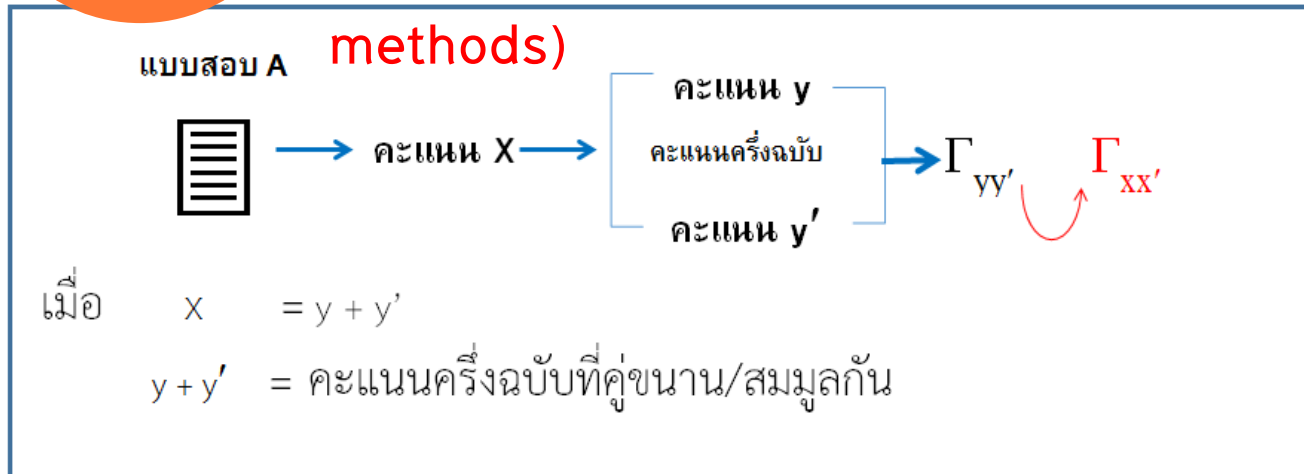
ความเป็นเอกพันธ์ของเนื้อหารายข้ออันเป็นตัวแทนของ**คุณลักษณะใด** **คุณลักษณะหนึ่ง**
(measure of internal consistency)

- วิธีการประมาณค่าความเที่ยงของแบบสอบด้วยการทดสอบเพียงครั้งเดียว
- ใช้แบบสอบฉบับเดียวทำการทดสอบผู้สอบกลุ่มเดียว
- เป็นการวัดระดับความเป็นเอกพันธ์ (homogeneity) ของข้อสอบในแบบสอบว่า วัดเนื้อเรื่องเดียวกันเพียงใด ถ้าแบบสอบวัดในเรื่องเดียวกันเมื่อทำการวัดซ้ำๆ ก็น่าจะมีคามคงที่ หรือสอดคล้องในผลการวัดกันสูง

สอ คล้อง

คุณภาพเครื่องมือ

1. วิธีแบ่งครึ่งข้อสอบ (Split-half



สูตรสเปียร์แมน-บราวน์ (Spearman-Brown formula)

$$\Gamma_{xx'} = \frac{2\Gamma_{yy'}}{1 + \Gamma_{yy'}}$$

เมื่อ $\Gamma_{yy'}$ = สัมประสิทธิ์ความเที่ยงครึ่งฉบับ
 $\Gamma_{xx'}$ = สัมประสิทธิ์ความเที่ยงทั้งฉบับ

สอ
ด
ค
ล
อง

คุณภาพเครื่องมือ

2. วิธีของคูเดอร์-ริชาร์ดสัน (Kuder-Richardson's Method)

Kuder และ Richardson (1937) พยายามคิดค้นวิธีช่วยแก้ปัญหาการประมาณค่าความเที่ยงของการแบ่งครึ่งข้อสอบ ซึ่งมักให้ผลแตกต่างกันตามวิธีที่ใช้ในการแบ่งครึ่ง สูตรที่พัฒนาขึ้นใช้สำหรับคำนวณความเที่ยงที่รู้จักกันดี ได้แก่ KR20 และ KR21 ซึ่งสามารถใช้ได้เฉพาะข้อสอบที่ตรวจให้คะแนนแบบ 0, 1 เท่านั้น

สอบ
คลัง

คุณภาพเครื่องมือ

2. วิธีของคูเดอร์-ริชาร์ดสัน (Kuder-Richardson's Method)

KR-20 คำนวณจากสูตร

$$r_{tt} = \frac{k}{k-1} \left(1 - \frac{\sum pq}{s^2} \right)$$

k = จำนวนข้อสอบ

p = สัดส่วนคนที่ตอบถูกในแต่ละข้อ

q = สัดส่วนคนที่ตอบผิดในแต่ละข้อ

S² = ความแปรปรวนของคะแนนรวม

KR-21 คำนวณจากสูตร

$$r_{tt} = \frac{k}{k-1} \left(1 - \frac{\bar{X}(k - \bar{X})}{ks^2} \right)$$

k = จำนวนข้อสอบ

X = ค่าเฉลี่ยคะแนนรวมทั้งฉบับ

S² = ความแปรปรวนของคะแนนรวม

สอดคล้อง

คล้อง

คุณภาพเครื่องมือ

3. วิธีสัมประสิทธิ์แอลฟาของครอนบาค (Cranbach's Alpha Method)

ขยายแนวคิดมาใช้กับการแบ่งแบบสอบถามออกเป็นมากกว่า 2 ส่วน หรือ n ส่วน เมื่อคำนวณความแปรปรวนของคะแนนแต่ละส่วนและความแปรปรวนของคะแนนรวม สามารถประมาณค่าความสอดคล้องภายในได้

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum S_i^2}{S_t^2} \right)$$

α = ความเที่ยงของแบบสอบ

S_i^2 = ความแปรปรวนของข้อสอบแต่ละข้อ

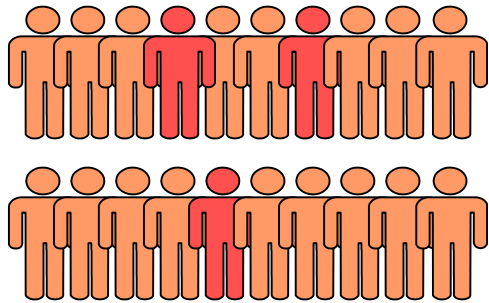
S_t^2 = ความแปรปรวนของคะแนนรวมทั้งฉบับ

คุณภาพเครื่องมือรายข้อ

ความยาก

คุณภาพเครื่องมือ

ความยาก (difficulty)



ข้อสอบ ข้อ 1 ถามว่า

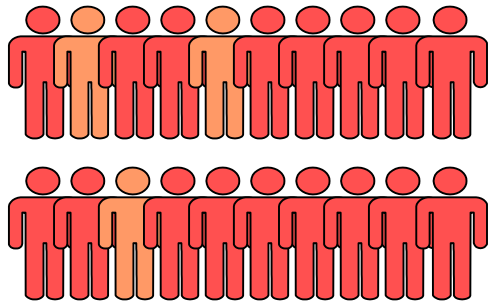
$$1 + 1 = ?$$

มีนักเรียนตอบถูก 17 คน

จาก 20 คน ข้อสอบข้อนี้
ยากหรือง่าย

คุณภาพเครื่องมือ

ความยาก (difficulty)



ข้อสอบ ข้อ 2 ถามว่า

$$2 + 2 = ?$$

มีนักเรียนตอบถูก 3 คน

จาก 20 คน ข้อสอบข้อนี้

ยากหรือง่าย

คุณภาพเครื่องมือ

ความยาก (difficulty)

สัดส่วนของ**จำนวนคนที่ตอบข้อสอบข้อนั้นถูก**
กับ**จำนวนคนทั้งหมด**

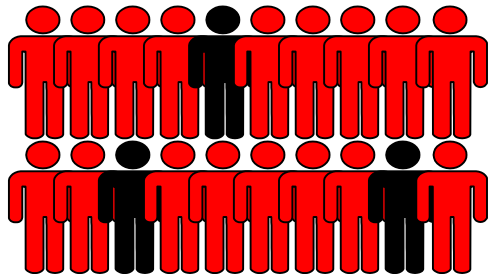
ข้อสอบ ข้อ 1 มีค่าความยากเท่ากับ $17/20 = 0.85$

ข้อสอบ ข้อ 2 มีค่าความยากเท่ากับ $3/20 = 0.15$

อำนาจจำแนก

คุณภาพเครื่องมือ

ความยาก (difficulty)



= คนเก่ง

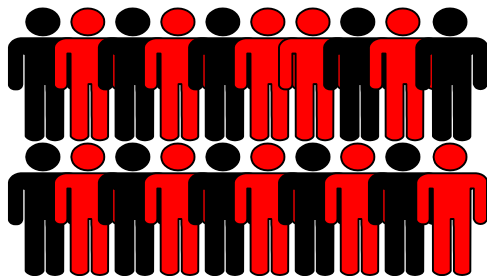


= คนไม่เก่ง

นักเรียน 20 คน ทำ
ข้อสอบแล้วพบว่า
สามารถแยกคนเก่งกับ
คนไม่เก่งได้ดังรูป

คุณภาพเครื่องมือ

ความยาก (difficulty)



= คนเก่ง



= คนไม่เก่ง

นักเรียน 20 คน ทำ
ข้อสอบแล้วพบว่า
สามารถแยกคนเก่งกับ
คนไม่เก่งได้ดังรูป

คุณภาพเครื่องมือ

อำนาจจำแนก (discrimination)

ความสามารถในการ**จำแนกคนเก่งและคนไม่เก่ง**
หรือ**คนรู้และคนไม่รู้**ออกจากกัน

คุณภาพเครื่องมือ






ความยาก (difficulty) ข้อสอบปรนัย

$$P = \frac{H + L}{N_H + N_L}$$

H = จำนวนคนตอบในกลุ่มสูง, N_H = จำนวนคนในกลุ่มสูง
L = จำนวนคนตอบในกลุ่มต่ำ, N_L = จำนวนคนในกลุ่มต่ำ

คุณภาพเครื่องมือ

ความยาก (difficulty)

P	ความหมาย
0.00-0.19	ยาก 
0.20-0.39	ค่อนข้างยาก 
0.40-0.60	ยากพอเหมาะ 
0.61-0.80	ค่อนข้างง่าย 
0.81-1.00	ง่าย 

คุณภาพเครื่องมือ

อำนาจจำแนก (discrimination) ข้อสอบปรนัย

$$r = \frac{H - L}{(N_H + N_L)/2}$$

H = จำนวนคนตอบในกลุ่มสูง, N_H = จำนวนคนในกลุ่มสูง
L = จำนวนคนตอบในกลุ่มต่ำ, N_L = จำนวนคนในกลุ่มต่ำ

คุณภาพเครื่องมือ

อำนาจจำแนก (discrimination)

r	ความหมาย
>0.40	จำแนกได้ดีมาก ☺
0.30-0.39	จำแนกได้ดี ☺
<u>0.20-0.29</u>	จำแนกได้พอใช้ ☺
0.00-0.19	จำแนกได้ต่ำ ☹
ค่าติดลบ	จำแนกไม่ได้ ☹

ความไว

คุณภาพเครื่องมือ

ความไว (S : sensitivity)

ความสามารถของข้อสอบในการ
จำแนก ความแตกต่างระหว่างผู้ที่รอบ
รู้-ไม่รอบรู้ หรือ ความสามารถก่อน-
หลังการสอน (มีฐานการประเมินเป็น
แบบอิงเกณฑ์)

คุณภาพเครื่องมือ

ความไว (S : sensitivity)

$$S = \frac{P_{\text{post}} - P_{\text{pre}}}{N}$$

เมื่อ

S คือ ดัชนีความไว

P_{post} คือ จำนวนคนตอบถูกหลังเรียน

P_{pre} คือ จำนวนคนตอบถูกต้องก่อนเรียน

N คือ จำนวนคนสอบทั้งหมด

เกณฑ์การพิจารณาความไว

S	ความหมาย
$S > 0.5$	ความไวสูง หลังเรียนทำข้อสอบได้เพิ่มขึ้น
$0 > S \leq 0.5$	ความไวปานกลาง หลังเรียนทำข้อสอบได้เพิ่มขึ้นเทียบกับก่อนเรียนไม่มาก
$S \leq 0$	ความไวต่ำ ก่อนเรียนทำข้อสอบได้มากกว่า