

ความจริง

การประเมินลักษณะมนุษย์

สันทัด พรประเสริฐมานิต

โครงร่างการนำเสนอ

- นิยามความตรง
- รูปแบบความตรง
- อคติและความยุติธรรม

นิยามความตรง

- ความตรงคือคุณสมบัติของคะแนนที่ได้ว่าตรงกับภาวะสันนิษฐานที่ต้องการหรือไม่
- เป็นคุณสมบัติของคะแนน ที่วัดจากกลุ่มตัวอย่างหนึ่ง ในสถานการณ์หนึ่ง ไม่ใช่คุณสมบัติของแบบวัด
- ไม่สามารถพิสูจน์ได้ว่าคะแนนมีความตรงหรือไม่ เพียงแต่หาหลักฐานสนับสนุนความตรงของคะแนน
- หลักฐานมากแค่ไหน ขึ้นอยู่กับการให้เหตุผลของผู้พัฒนา

รูปแบบความตรง

- ความตรงตามเนื้อหา (Content Validity)
- ความตรงตามความสัมพันธ์กับเกณฑ์ (Criterion-related Validity)
- ความตรงตามภาวะสันนิษฐาน (Construct Validity)
 - ความตรงส่วนนี้ครอบคลุมความตรงสองชนิดแรก

ความตรงตามทีเห็น

- ความตรงตามทีเห็น (Face Validity) คือ ความคิดของผู้รับการประเมิน (หรือผู้ตรวจสอบแบบวัด) ว่าวัดในสิ่งที่ต้องการวัดหรือไม่
- ไม่ใช่ความตรง
- มีความสำคัญส่วนหนึ่ง เพราะทำให้ผู้ประเมินให้ความร่วมมือในการทำแบบวัด

ความตรงตามเนื้อหา

- ความตรงตามเนื้อหา คือ คุณสมบัติของแบบวัดที่ครอบคลุมเนื้อหาทั้งหมดของสิ่งที่ต้องการวัด
- เหมือนการสุ่ม (Random sample) ที่สุ่มตัวอย่างของข้อคำถามมาได้เป็นตัวแทนของสิ่งที่เป็นไปได้ทั้งหมดในสิ่งที่ต้องการวัด
- ความตรงตามเนื้อหา จะชัดเจนมากในการวัดผลสัมฤทธิ์หรือผลการปฏิบัติงาน

ความตรงตามเนื้อหา

- สัดส่วนความตรงตามเนื้อหา (Content Validity Ratio; CVR) ใช้การประเมินจากผู้เชี่ยวชาญว่าข้อคำถามแต่ละข้อ “สำคัญ” “เกี่ยวข้องแต่ไม่สำคัญ” และ “ไม่จำเป็น” ต่อภาวะสันนิษฐาน
- เช่น วัดการทำงานเป็นทีม

	สำคัญ	เกี่ยวข้องแต่ไม่สำคัญ	ไม่จำเป็น
1. ฉันสนใจว่าเพื่อนร่วมงานรู้สึกอย่างไรกับฉัน			
2. ฉันชอบช่วยทำงานของเพื่อนร่วมทีม			
3. ฉันหลีกเลี่ยงทำให้บรรยากาศในทีมเสีย			

ความตรงตามเนื้อหา

$$CVR = \frac{\#สำคัญ - N/2}{N/2}$$

- N = จำนวนผู้ประเมินในแต่ละข้อ
- CVR มีตารางระดับนัยสำคัญ แต่ผู้แต่งไม่ได้บอกวาระดับนัยสำคัญนั้นหามาจากอะไร
- ระดับนัยสำคัญไม่สำคัญเท่ากับขนาดอิทธิพล

ความตรงตามเนื้อหา

- การหาสัดส่วนนี้แตกต่างจากความตรงที่เห็นตรงที่ใช้ผู้เชี่ยวชาญ
- ปัญหาก็คือ ใครเป็นผู้เชี่ยวชาญ
- นอกจากนี้ CVR ไม่ได้วัดความตรงตามเนื้อหาทั้งหมด เพราะไม่ได้วัดความครอบคลุมของเนื้อหา
- การวัดความครอบคลุมจากเนื้อหาให้ใช้การวิเคราะห์ผ่านตารางลักษณะของภาวะสันนิษฐาน (Table of Specification)

ความตรงตามเนื้อหา

- ประเมินผลสัมฤทธิ์วิชาคณิตศาสตร์หลัก ม.4 ภาคต้น

เนื้อหา	จำนวนคาบที่สอน	จำนวนข้อ
1. เซต	3	15
2. การให้เหตุผล	3	15
3. จำนวนจริง	5	25
4. เลขยกกำลัง	3	15

คุณต้องคำนึงถึงตารางนี้ทั้งในขั้นตอนการสร้างข้อคำถาม และแบบวัดฉบับสมบูรณ์

ความตรงตามความสัมพันธ์กับเกณฑ์

- เกณฑ์ คือ สิ่งที่แบบวัดต้องการใช้ในการเปรียบเทียบกับ เช่น ผลการปฏิบัติงาน แบบวัดเซาว์นปัญญาแบบเดิม
- ความตรงตามความสัมพันธ์กับเกณฑ์ คือ การตรวจสอบว่าคะแนนที่ได้จากแบบวัดสัมพันธ์กับเกณฑ์ตามที่คาดหวังหรือไม่
- แบ่งเป็นสองชนิดคือ
 - ความตรงตามสภาพ (Concurrent Validity) ไม่เว้นระยะเวลาในการวัดเกณฑ์
 - ความตรงเชิงทำนาย (Predictive Validity) เว้นระยะเวลาในการวัดเกณฑ์

ความตรงตามความสัมพันธ์กับเกณฑ์

- ความสัมพันธ์ที่ได้ควรถึงระดับนัยสำคัญ และมีขนาดตามที่คาดหวังด้วย
- เกณฑ์อาจเป็นภาวะสันนิษฐานเช่นกัน ต้องใช้เครื่องมือวัด ซึ่งทำให้ต้องคำนึงถึงความตรงของเกณฑ์
- โดยเฉพาะอย่างยิ่ง เกณฑ์ต้องมีความตรงตามเนื้อหา โดยมีคำศัพท์จำเพาะสำหรับเกณฑ์คือ
 - สิ่งบกพร่องในการวัดเกณฑ์ (Criterion deficiency)
 - สิ่งเจือปนในการวัดเกณฑ์ (Criterion contamination)

ความตรงตามความสัมพันธ์กับเกณฑ์

- การจำกัดพิสัยเป็นสิ่งที่ต้องคำนึงถึงในการวัดความตรงเชิงทำนาย
- ความตรงเพิ่มเติม (Incremental Validity) เป็นการตรวจสอบว่าแบบวัดที่มีนั้นสามารถทำนายเกณฑ์ได้เพิ่มเติมจากแบบวัดเดิมหรือไม่

Schmidt & Hunter (1998) การตรวจสอบความตรงเพิ่มเติมในการทำนายผลการปฏิบัติงาน เมื่อการวัดเชาวน์ปัญญาใช้เป็นตัวทำนาย

Table 1
*Predictive Validity for Overall Job Performance of General Mental Ability (GMA) Scores
Combined With a Second Predictor Using (Standardized) Multiple Regression*

Personnel measures	Validity (<i>r</i>)	Multiple <i>R</i>	Gain in validity from adding supplement	% increase in validity	Standardized regression weights	
					GMA	Supplement
GMA tests ^a	.51					
Work sample tests ^b	.54	.63	.12	24%	.36	.41
Integrity tests ^c	.41	.65	.14	27%	.51	.41
Conscientiousness tests ^d	.31	.60	.09	18%	.51	.31
Employment interviews (structured) ^e	.51	.63	.12	24%	.39	.39
Employment interviews (unstructured) ^f	.38	.55	.04	8%	.43	.22
Job knowledge tests ^g	.48	.58	.07	14%	.36	.31
Job tryout procedure ^h	.44	.58	.07	14%	.40	.20
Peer ratings ⁱ	.49	.58	.07	14%	.35	.31
T & E behavioral consistency method ^j	.45	.58	.07	14%	.39	.31
Reference checks ^k	.26	.57	.06	12%	.51	.26
Job experience (years) ^l	.18	.54	.03	6%	.51	.18
Biographical data measures ^m	.35	.52	.01	2%	.45	.13
Assessment centers ⁿ	.37	.53	.02	4%	.43	.15
T & E point method ^o	.11	.52	.01	2%	.39	.29
Years of education ^p	.10	.52	.01	2%	.51	.10
Interests ^q	.10	.52	.01	2%	.51	.10
Graphology ^r	.02	.51	.00	0%	.51	.02
Age ^s	-.01	.51	.00	0%	.51	-.01

ความตรงตามภาวะสันนิษฐาน

- ความตรงตามภาวะสันนิษฐาน (Construct Validity) เป็นการให้เหตุผลว่าคะแนนที่ได้สะท้อนถึงภาวะสันนิษฐานอย่างไร
 - ต้องหาหลักฐานมาสนับสนุนหรือโต้แย้ง
- ความตรงนี้สามารถแบ่งออกเป็น 2 ประเภทตามลักษณะของหลักฐาน คือ
 - ความตรงภายใน (Internal Validity)
 - ความตรงภายนอก (External Validity)

ความตรงตามภาวะสันนิษฐาน

- ความตรงภายใน เป็นการแสดงว่าเนื้อหาที่อยู่ภายในแบบวัดสอดคล้องกัน วัดในสิ่งเดียวกัน
 - ความสอดคล้องภายใน ซึ่งอาจวัดจากความเที่ยงหรือผลการวิเคราะห์องค์ประกอบ
 - ความเที่ยงรูปแบบอื่นๆ
 - การวิเคราะห์ข้อคำถาม

ความตรงตามภาวะสันนิษฐาน

- ความตรงภายนอก เป็นการแสดงว่าคะแนนที่ได้จากแบบวัดสัมพันธ์กับตัวแปรอื่นตามที่คาดหวัง
 - ความสัมพันธ์กับอายุ พัฒนาการ
 - เปลี่ยนแปลงตามการจัดกระทำ
 - แตกต่างระหว่างกลุ่มตามทฤษฎี (Contrasted groups; Known groups)
 - สัมพันธ์กับตัวแปรอื่นตามทฤษฎี
 - สัมพันธ์กับแบบวัดอื่นที่วัดในสิ่งเดียวกัน หรือที่เรียกว่าความตรงลู่อเข้า (Convergent Validity)
 - ไม่สัมพันธ์กับแบบวัดอื่นที่ไม่ได้วัดในสิ่งเดียวกัน หรือที่เรียกว่าความตรงลู่ออก (Discriminant Validity)

ความตรงตามภาวะสันนิษฐาน

- เมทริกซ์หลายภาวะสันนิษฐานหลายวิธีการวัด (Multitrait-multimethod matrix) วิธีนี้ใช้แบบวัดหลายวิธีเพื่อวัดตัวแปรที่แตกต่างกัน
 1. คะแนนที่วัดตัวแปรเดียวกันด้วยวิธีที่แตกต่างกัน (Monotrait-heteromethod) ควรมีค่าสูง ซึ่งเป็นหลักฐานความตรงลู่เข้า
 2. คะแนนที่วัดตัวแปรที่แตกต่างกันด้วยวิธีเดียวกัน (Heterotrait-monomethod) ควรมีค่าต่ำ ซึ่งเป็นหลักฐานความตรงลู่ออก
 3. คะแนนที่วัดตัวแปรที่แตกต่างกันด้วยวิธีที่แตกต่างกัน (Heterotrait-heteromethod) ควรมีค่าต่ำเช่นกัน ซึ่งเป็นหลักฐานความตรงลู่ออก
- หาก (2) มากกว่า (3) จะเรียกว่ามีความแปรปรวนจากวิธีที่เหมือนกัน (Method variance)

ผู้บังคับบัญชา (B)

ตนเอง (S)

ผู้ใต้บังคับบัญชา (E)

	A	F	C	A	F	C	A	F	C
B-A									
B-F	.35								
B-C	.10	.38							
S-A	.56	.17	.04						
S-F	.20	.26	.18	.33					
S-C	-.01	-.03	.35	.10	.16				
E-A	.32	.17	.20	.27	.26	-.02			
E-F	-.03	.07	.28	.01	.17	.14	.26		
E-C	-.10	.14	.49	.00	.05	.40	.17	.52	

A = การบริหาร (Administration); F = การให้ข้อมูลป้อนกลับ (Feedback); C = ความเอาใจใส่ (Consideration)

อคติและความยุติธรรม

- อคติของแบบวัด (Test Bias) -- ทดสอบได้
- ความยุติธรรมของแบบวัด (Test Fairness) – ทดสอบไม่ได้ ควบคุมไม่ได้
เพราะเป็นการรับรู้

อคติและความยุติธรรม

- อคติ หมายถึง คะแนนที่ได้จากแบบวัดมีคุณสมบัติแตกต่างกันระหว่างกลุ่ม
 - เช่น สองกลุ่มมีระดับภาวะสันนิษฐานเท่ากัน แต่มีคะแนนจากแบบทดสอบแตกต่างกัน
 - ภาวะสันนิษฐานที่แตกต่างกัน 1 หน่วย ทำให้ระยะห่างของคะแนนจากแบบทดสอบแตกต่างกันระหว่างสองกลุ่ม
- อคติรูปแบบนี้ ทำให้คะแนนจากแบบทดสอบที่ได้เท่ากัน จะหมายถึงระดับภาวะสันนิษฐานที่แตกต่างกันระหว่างกลุ่ม

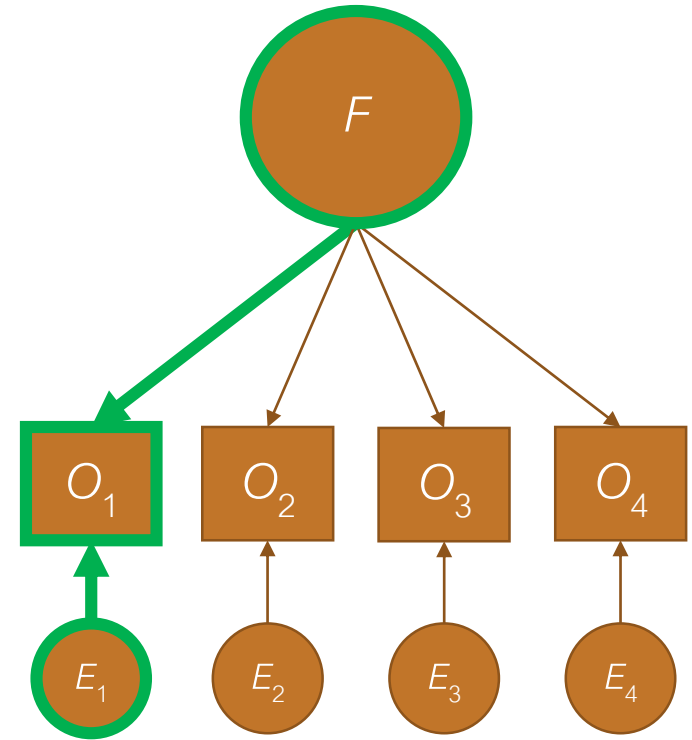
อคติและความยุติธรรม

$$O_1 = b_0 + b_1F + b_2G + b_3FG$$

สมมติ $G = 1$ คือเพศชาย และ $G = 0$ คือเพศหญิง

ชาย $O_1 = (b_0 + b_2) + (b_1 + b_3)F$

หญิง $O_1 = b_0 + b_1F$



อคติและความยุติธรรม

$$\text{ชาย } O_1 = (b_0 + b_2) + (b_1 + b_3)F$$

$$\text{หญิง } O_1 = b_0 + b_1F$$

- ถ้า $b_2 \neq 0$ แล้วเมื่อ $F = 0$ จะทำให้ $O_1 = b_0$ และ $b_0 + b_2$ สำหรับหญิงและชายตามลำดับ
- ถ้า $b_3 \neq 0$ แล้วเมื่อ F มีค่าเพิ่มขึ้นจาก 0 เป็น 1 แล้ว O_1 จะเพิ่มขึ้น b_1 และ $b_1 + b_3$ สำหรับหญิงและชายตามลำดับ

อคติและความยุติธรรม

- ด้วยเหตุนี้ การทดสอบอคติก็คือการทดสอบว่า b_2 และ b_3 แตกต่างจาก 0 อย่างมีนัยสำคัญหรือไม่
- ระดับของภาวะสันนิษฐานอาจแตกต่างกันระหว่างกลุ่มได้
- แต่ระดับของคะแนนที่ได้ต้องเท่ากันระหว่างกลุ่ม หากระดับของภาวะสันนิษฐานเท่ากัน
- การทดสอบนี้อาจเรียกว่าการทดสอบการทำงานของข้อคำถามที่แตกต่างกัน (Differential Item Functioning; DIF) หรือการทดสอบคุณสมบัติการวัดที่คงที่ระหว่างกลุ่ม (Measurement Invariance)

อคติและความยุติธรรม

- ความยุติธรรมเป็นการรับรู้ของผู้ทดสอบ
- คะแนนทดสอบอาจไม่มีอคติ แต่ผู้ทดสอบอาจรู้ว่าแบบทดสอบไม่ยุติธรรมสำหรับตนเอง (เพราะกลุ่มตนเองมีระดับภาวะสันนิษฐานต่ำตั้งแต่ต้นอยู่แล้ว)
- ถ้าแบบทดสอบไม่มีอคติ คะแนนที่วัดได้ก็สะท้อนระดับภาวะสันนิษฐานที่ถูกต้อง
- แต่หากผู้ทดสอบคิดว่าคะแนนที่ได้ไม่ยุติธรรม อาจมีปัญหาเรื่องความน่าเชื่อถือของคะแนน

คาบต่อไป

- ใบงานที่ 3
 - วิเคราะห์ข้อคำถาม
 - คำนวณความสอดคล้องภายใน
 - เก็บข้อมูลหาความตรง
- อ่าน Cohen et al บทที่ 8