

การตรวจสอบ ความตรงเชิงเนื้อหา



1

โดย

รองศาสตราจารย์ ดร.สุจิตรา เทียนสวัสดิ์

แนวคิดดั้งเดิมเกี่ยวกับความตรง (validity)

- ➡ แบ่งความตรงเป็น 3 ประเภท
 - ➡ ความตรงเชิงโครงสร้างหรือเชิงทฤษฎี (construct validity)
 - ➡ ความตรงเชิงเนื้อหา (content validity)
 - ➡ ความตรงเชิงสัมพันธ์กับเกณฑ์ (criterion-related validity) แบ่งเป็น
 - ➡ Concurrent validity
 - ➡ Predictive validity

แนวคิดปัจจุบันเกี่ยวกับความตรง (validity)

- ความตรงเป็นแนวคิดหนึ่งเดียว (unitary construct)
 - การตรวจสอบความตรงด้วยวิธีต่างๆ ตามแนวคิดดั้งเดิม ถือเป็นหลักฐานสนับสนุน construct validity เช่น
 - Content-related validity evidence
 - Evidence from factor analysis
 - Evidence from known groups
 - Evidence from prediction of concurrent events
 - Evidence from prediction of future events

ความตรงเชิงเนื้อหา (content validity)

4

- เป็นคุณสมบัติของเครื่องมือที่วัดได้ตรงและครอบคลุมเนื้อหาที่ต้องการวัดอย่างครบถ้วน
- **“Content validation”** หมายถึง การทำให้เครื่องมือมีความตรงเชิงเนื้อหา ประกอบด้วย 2 ขั้นตอน

Development stage ➔ Priori content validation

Judgment stage ➔ Posteriori content validation

Content Validation

- **Priori content validation:** ระยะเวลาสร้างเครื่องมือ
 - การกำหนดประเด็นเนื้อหา วางแผนโครงสร้างของแบบวัด (อาจจัดทำ blueprint ของแบบวัด)
 - การเขียน items
- **Posteriori content validation:** หลังได้เครื่องมือชุดแรก (ชุด “ร่าง”) - การตรวจสอบโดยผู้ทรงคุณวุฒิ

การตรวจสอบความตรงเชิงเนื้อหา โดยผู้ทรงคุณวุฒิ

- ➡ ประเด็นที่ต้องพิจารณา
 1. คุณสมบัติของผู้ทรงคุณวุฒิ
 2. จำนวนผู้ทรงคุณวุฒิ
 3. ตัวชี้วัด (indicator) และเกณฑ์การตัดสินผล
การพิจารณา

คุณสมบัติของผู้ทรงคุณวุฒิ

- มีคุณวุฒิ และประสบการณ์ (ฝึกอบรม ทำงาน/ทำวิจัย) ในสาขาหรือเรื่องที่เกี่ยวข้อง
- มีความเชี่ยวชาญทางคลินิกในสาขาเฉพาะที่เกี่ยวข้อง
- มีความเชี่ยวชาญในแนวคิดหรือทฤษฎีที่ใช้เป็นกรอบในการสร้างเครื่องมือ
- ควรสามารถให้ความเห็นเกี่ยวกับรูปแบบโครงสร้างของแบบวัดด้วย*

* ผู้ทรงคุณวุฒิอย่างน้อย 1 คน ควรมีความรู้หรือประสบการณ์ด้านการพัฒนาเครื่องมือ (Davis, 1992)

จำนวนผู้ทรงคุณวุฒิ

- ▶ ตั้งแต่ 2 – 20 คน (Grant & Davis, 1997)
- ▶ อย่างน้อย 5 คน เพื่อให้สามารถควบคุม chance agreement ได้ ทำให้ผลการพิจารณามีนัยสำคัญ (Lynn, 1986)
- ▶ ถ้าสามารถหาผู้ที่มีคุณสมบัติที่เหมาะสมได้ แนะนำให้เพิ่มจำนวน >5 คน แต่ไม่จำเป็นต้องมากกว่า 10 คน (Lynn, 1986)
- ▶ ถ้าไม่สามารถหาผู้ที่มีคุณสมบัติที่เหมาะสมได้ – อนุโลมให้ใช้อย่างน้อย 3 คน (Lynn, 1986)
- ▶ ถ้าเป็นการพัฒนาเครื่องมือใหม่ – แนะนำให้ใช้อย่างน้อย 6 คน (สุจิตรา เทียนสวัสดิ์, 2550)

เอกสารที่ควรเตรียมให้ผู้ทรงคุณวุฒิ

คำจำกัดความของมโนทัศน์ที่ต้องการวัด ทั้งคำจำกัดความเชิงทฤษฎี และเชิงปฏิบัติการ

- แบบวัดที่ประกอบด้วยข้อคำถามทั้งหมดและตัวเลือกคำตอบในรูปแบบที่จะใช้จริง
- คำชี้แจงประเด็นที่ต้องการให้ผู้ทรงคุณวุฒิพิจารณา ได้แก่
 1. เนื้อหาของข้อคำถาม - ความเกี่ยวข้อง และความเพียงพอหรือครอบคลุม
 2. รูปแบบ - ความชัดเจนและความกระชับ (ไม่เยิ่นเย้อ) ของข้อคำถามและตัวเลือก
 3. ความสมบูรณ์ - ของแบบวัดทั้งฉบับ
- แบบสอบถามสำหรับการประเมินของผู้ทรงคุณวุฒิ* (แตกต่างกันตามสถิติที่ใช้เป็นตัวชี้วัด)

ค่าสถิติที่ใช้เป็นตัวชี้วัด

- Content validity index*
- Content validity ratio
- Proportion of inter-rater agreement
- Index of item-objective congruence

Content validity index (CVI) – ดัชนีความตรงเชิงเนื้อหา

- วิธีการ – ให้ผู้ทรงคุณวุฒิประเมินโดยใช้ แบบวัดประมาณค่า 4 ระดับ (1 = irrelevant, 2 = somewhat relevant, 3 = quite relevant, 4 = highly relevant)
- **CVI คือ** สัดส่วนของข้อความที่ผู้ทรงคุณวุฒิให้คะแนน 3 หรือ 4 (Waltz & Bausell, 1981 อ้างใน Lynn, 1986)
- **Waltz, Strickland, & Lenz (1991, 2005)** อธิบายวิธีคำนวณว่า “กรณีที่ให้ผู้ทรงคุณวุฒิ **2** คน CVI หมายถึง สัดส่วนของข้อความที่ผู้ทรงคุณวุฒิทั้ง **2** คนมีความเห็นตรงกันว่าข้อความนั้นเกี่ยวข้องมาก (คะแนน 4) และเกี่ยวข้อง (คะแนน 3) กับสิ่งที่ต้องการวัด”
- **ปัญหาคือ** ไม่มีคำอธิบายสำหรับกรณีที่ให้ผู้ทรงคุณวุฒิ > 2 คน

Content validity index (CVI) - 2

12

- ➡ **Lynn (1986)** อธิบายวิธีคำนวณค่า CVI 2 ขั้นตอน
 1. คำนวณค่า CVI ของแต่ละข้อ – สัดส่วนของผู้ทรงคุณวุฒิที่ให้คะแนน 3 หรือ 4
 - ข้อที่ได้ต่ำกว่า **“เกณฑ์”** – คัดออก หรือปรับปรุง
 - ให้ผู้ทรงคุณวุฒิประเมินซ้ำ จนทุกข้อได้ตามเกณฑ์
 2. คำนวณค่า CVI ของทั้งฉบับ - จากสัดส่วนของข้อที่ผ่านการประเมินว่ามีความตรง

Numbers and Proportion of experts whose endorsement is required to establish content beyond the .05 of significance (Lynn, 1986)

Number of Experts	Number of experts who agree	Proportion of experts who agree*
2-5	All	1.00
6	At least 5	0.88
7	At least 6	0.86
8	At least 7	0.88
9	At least 7	0.78
10	At least 8	0.80

Content validity index (CVI) - 3

14

- ➡ **Polit & Beck (2006)** เสนอให้คำนวณค่า CVI 2 ระดับ
 1. Item-level CVI (I-CVI) คือ CVI ของแต่ละข้อ
 2. Scale-level CVI (S-CVI) คือ CVI ของแบบวัดทั้งฉบับ
- ➡ **I-CVI** คำนวณจากสัดส่วนของผู้ทรงคุณวุฒิที่มีความเห็นตรงกันว่าข้อความนั้นๆ เกี่ยวข้องกับสิ่งที่วัด
- ➡ เกณฑ์การพิจารณาค่า I-CVI ที่ยอมรับได้ ใช้เกณฑ์ของ Lynn (1986)

- S-CVI: **Polit & Beck (2006)** เสนอวิธีการคำนวณเป็น 2 ลักษณะหรือ 2 วิธี คือ
 1. S-CVI/UA (universal agreement) คือยึดตามคำจำกัดความของ CVI และตีความว่า S-CVI เท่ากับ “สัดส่วนของข้อความที่ผู้ทรงคุณวุฒิ**ทุกคน**มีความเห็นตรงกันว่าข้อความนั้นเกี่ยวข้องมาก (ตัวเลือก 4) และเกี่ยวข้อง (ตัวเลือก 3) กับสิ่งที่ต้องการวัด”
 - การคำนวณ S-CVI/UA เป็นวิธีที่เข้มงวดเกินไป และมีโอกาสน้อยที่จะได้ค่าตามเกณฑ์ โดยเฉพาะกรณีที่ใช้ผู้ทรงคุณวุฒิจำนวนมาก

Content validity index (CVI) - 5

16

- S-CVI: การคำนวณลักษณะ/วิธีที่ 2 คือ
 2. S-CVI/Ave (average proportion) คือ ค่าเฉลี่ยสัดส่วน
 - สามารถคำนวณได้ 3 วิธี
 - 1) หาค่าเฉลี่ยของ I-CVI
 - 2) หาค่าเฉลี่ยของสัดส่วนของข้อความที่ผู้ทรงคุณวุฒิแต่ละคนให้คะแนน 3 หรือ 4
 - 3) หาสัดส่วนคำตอบที่ประเมินว่า valid โดยรวมจำนวนคำตอบที่ผู้ทรงคุณวุฒิแต่ละคนให้คะแนน 3 หรือ 4 แล้วหารด้วยจำนวนคำตอบทั้งหมด

ข้อที่	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	คนที่ 6	จ.น.ที่เห็น ว่า valid	ค่า I-CVI
1	✓	✓	✓	✓	-	✓	5	.83
2	-	✓	✓	✓	✓	✓	5	.83
3	✓	✓	✓	✓	✓	✓	6	1.00
4	✓	✓	-	✓	✓	✓	5	.83
5	✓	✓	✓	✓	✓	✓	6	1.00
6	✓	✓	✓	-	✓	✓	5	.83
7	✓	-	✓	✓	✓	✓	5	.83
8	✓	✓	✓	✓	✓	-	5	.83
9	✓	✓	✓	✓	✓	✓	6	1.00
10	✓	✓	✓	✓	✓	✓	6	1.00
สัดส่วน ที่ ประเมิน โดยแต่ละคน	.90	.90	.90	.90	.90	.90	1) ค่าเฉลี่ย I-CVI = .90 2) ค่าเฉลี่ยสัดส่วน = .90 3) สัดส่วนคำตอบ = $54/60 = .90$	

Content validity index (CVI) - 6

18

การคำนวณ ค่า S-CVI/Ave ทั้ง 3 วิธี ได้ผลลัพธ์เท่ากัน

- วิธีที่เหมาะสมสอดคล้องกับแนวคิดของดัชนีความตรงเชิงเนื้อหามากที่สุด คือ วิธีคำนวณค่าเฉลี่ย I-CVI เพราะเป็นการเฉลี่ยคุณภาพของข้อความรายข้อ และเป็นวิธีที่ง่าย (Polit & Beck, 2006)
- เกณฑ์ขั้นต่ำที่ยอมรับได้ Polit & Beck (2006) เสนอให้ใช้ 0.90 ส่วนเกณฑ์ 0.80 ที่นิยมใช้ควรใช้กับ S-CVI/UA

Content validity ratio (CVR) อัตราส่วนความตรงเชิงเนื้อหา

- ➡ **CVR** เป็นวิธีประเมินความตรงของข้อคำถามแต่ละข้อของแบบวัด ที่เสนอโดย **Lawshe (1975)** นักจิตวิทยาองค์การ
- ➡ ให้ผู้ทรงคุณวุฒิประเมินโดยใช้มาตรวัด 3 ระดับ คือ
 - 1 หมายถึง จำเป็น
 - 2 หมายถึง เป็นประโยชน์แต่ไม่จำเป็น
 - 3 หมายถึง ไม่จำเป็น

สูตรคำนวณค่า Content validity ratio (CVR)

➡ สูตรคำนวณ CVR

$$CVR = \frac{n_e - N/2}{N/2}$$

N_e = จำนวนผู้ทรงคุณวุฒิที่ตอบ “จำเป็น”

N = จำนวนผู้ทรงคุณวุฒิทั้งหมด

- ❖ เมื่อเลือกหรือปรับปรุงข้อความทุกข้อจนได้ค่า CVR ตามเกณฑ์ แล้ว นำค่า CVR ทั้งหมดมาคำนวณค่าเฉลี่ย ได้เป็นดัชนีความตรงเชิงเนื้อหา (CVI) ของแบบวัดทั้งฉบับ

เกณฑ์ขั้นต่ำ (Critical cut off value) ของค่า CVR เปรียบเทียบกับเกณฑ์ I-CVI ของ Lynn

Number of Experts	$CVR_{critical}$	Lynn's criteria for I-CVI
5	.99	1.00
6	.99	.83
7	.99	.86
8	.75	.88
9	.78	.78
10	.62	.80

Proportion of inter-rater agreement

- สัดส่วนของข้อความที่ผู้ประเมิน 2 คนมีความเห็นตรงกัน (P_0)
- ค่านี้ต่างจาก CVI เพราะนับรวมทุกข้อที่ผู้ประเมินให้คะแนนตรงกัน (รวมคะแนน 1 หรือ 2 ด้วย) – น่าเชื่อถือน้อยกว่า
- Davis (1992) แนะนำให้คำนวณค่า P_0 ก่อน ถ้าได้ $P_0 \geq .70$ จึงจะคำนวณค่า CVI
- ถ้าได้ $P_0 < .70$ ควรสอบถามและทำความเข้าใจกับผู้ทรงคุณวุฒิเกี่ยวกับคำจำกัดความของมโนทัศน์ และ/หรือ วิธีการประเมินตามมาตรวัด ก่อนขอให้ผู้ทรงคุณวุฒิประเมินใหม่

Index of Item-Objective Congruence (IOC)

23

- ▶ ใช้ประเมินความตรงเชิงเนื้อหาของแบบสอบทางการศึกษา
- ▶ ประเมินความสอดคล้องของข้อสอบแต่ละข้อกับวัตถุประสงค์ของการวัด
- ▶ ใช้ผู้ทรงคุณวุฒิด้านเนื้อหาจำนวนอย่างน้อย 3 คน
- ▶ ให้ผู้ทรงคุณวุฒิแต่ละคนประเมินและให้คะแนน 3 ระดับ ถ้าสอดคล้องให้ 1 คะแนน, ไม่สอดคล้องให้ -1, ถ้าไม่ชัดเจน ให้ 0
- ▶ รวมคะแนนจากผู้ทรงคุณวุฒิทุกคนหารด้วยจำนวนผู้ทรงคุณวุฒิ
- ▶ ค่า IOC ควรเท่ากับ 0.5 -1.0

สรุปการตรวจสอบความตรงเชิงเนื้อหา

24

- ➡ จำนวนผู้ทรงคุณวุฒิ
- ➡ โดยทั่วไปอย่างน้อย 3-5 คน
- ➡ หากพัฒนาเครื่องมือใหม่ แนะนำให้ใช้อย่างน้อย 6 คน
- ➡ ตัวชี้วัด
 - ➡ Polit, Beck, Owen (2007) แนะนำให้ใช้ตัวชี้วัด content validity index เนื่องจากคำนวณได้ง่าย เข้าใจง่าย และให้ข้อมูลคุณภาพของข้อความแต่ละข้อและแบบวัดทั้งฉบับ
 - ➡ คำนวณทั้ง I-CVI และ S-CVI (S-CVI/Ave)

เอกสารสำหรับผู้ทรงคุณวุฒิ

25

- ➡ คำจำกัดความเชิงทฤษฎี และเชิงปฏิบัติการ
- ➡ คำชี้แจงประเด็นที่ต้องการให้พิจารณา
- ➡ แบบวัดตามรูปแบบที่จะใช้จริง
- ➡ แบบประเมินเป็นแบบมาตรวัดประมาณค่า 4 ระดับ
 - 4 = เกี่ยวข้องมาก, 3 = เกี่ยวข้องพอควร, 2 = เกี่ยวข้องเล็กน้อย, 1 = ไม่เกี่ยวข้อง
 - ควรมีคำถามปลายเปิดเพื่อให้ผู้ทรงคุณวุฒิให้ความเห็นเพิ่มเติมตามประเด็นที่ได้ชี้แจงไว้

สรุปการตรวจสอบความตรงเชิงเนื้อหา (ต่อ)

26

➤ เกณฑ์ขั้นต่ำ

- I-CVI ใช้เกณฑ์ของ Lynn (1986) ซึ่งขึ้นอยู่กับจำนวนผู้ทรงคุณวุฒิ
- ส่วน S-CVI
 - เครื่องมือที่แปลมา อาจใช้ .80 (Grant & Davis, 1997; Polit & Hungler, 1999)
 - เครื่องมือที่พัฒนาใหม่ แนะนำให้ใช้ .90 (Polit & Beck, 2006; Polit, Beck, & Owen, 2007)

สรุป

การตรวจสอบความตรงของเครื่องมือ

- ➡ ความตรงเชิงเนื้อหาเป็นเพียง 1 evidence ที่แสดงถึงความตรงของเครื่องมือ ซึ่งไม่เพียงพอที่จะสนับสนุน construct validity
- ➡ การตรวจสอบความตรงของเครื่องมือเป็นการค้นหาหลักฐานเชิงประจักษ์ (empirical evidence) ที่ต้องใช้กระบวนการวิจัย
- ➡ การพัฒนาและตรวจสอบความตรงของเครื่องมือเป็นกระบวนการที่ไม่มีที่สิ้นสุด (Nunnally & Bernstein, 1994)

*QUESTIONS
&
ANSWERS*