

ความเที่ยงตรงและความเชื่อมั่น (Validity and Reliability)

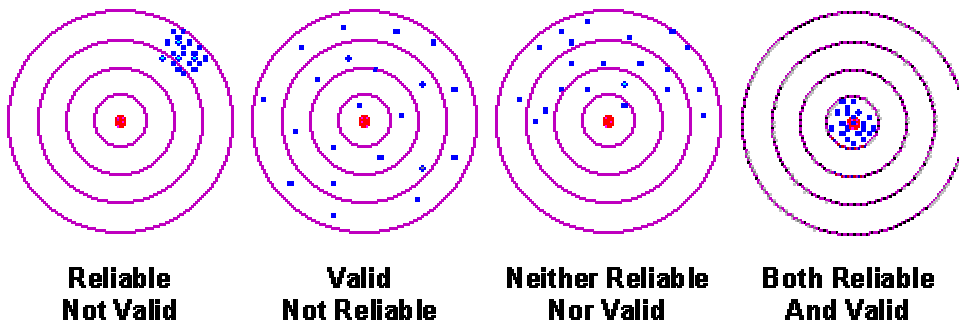
นั้ตรีศิริ ปิยะพิมลลธิธิ* และ อุทัยวรรณ สายพัฒนนะ**

แนวคิดของ Validity และ Reliability

เรามากจะคิดถึงค้ำว่า Reliability และ Validity แยกส่วนจากกันอยู่เสมอ แต่ในความเป็นจริงแล้วค้ำศัพท์ทั้งสองนี้มีความสัมพันธ์กัน เราจะแสดงวิธีการคิด 2 แบบที่แสดงถึงความสัมพันธ์ระหว่าง Reliability และ Validity ดังนี้

วิธีที่ 1

เราจะใช้เกมปาเป้าแสดงความสัมพันธ์ระหว่าง Reliability และ Validity โดยใช้ศูนย์กลางของเป้านั้นคือสิ่งที่เราพยายามจะวัด ให้จินตนาการว่า เด็กแต่ละคนที่ถูกเราวัดคือคนที่เราปาไปที่เป้า ถ้าเราวัดสิ่งนั้นของแต่ละคนได้อย่างถูกต้องจริง ๆ ก็คือเราปาได้ตรงศูนย์กลางของเป้า แต่ถ้าวัดไม่ตรงแสดงว่าเราปาพลาดจากศูนย์กลางของเป้า



ในรูปภาพข้างต้นแสดงสถานการณ์ทั้งหมดที่เป็นไปได้ ในเป้าแรก การปาเป้าจะปาได้ สอดคล้องแต่ไม่ตรงศูนย์กลางของเป้า นั่นคือการวัดมี Reliability แต่ไม่มี Validity (คือมีความ สอดคล้องแต่ผิดเป้าหมาย) ในเป้าที่สอง แสดงการปาเป้าที่ไม่สอดคล้องกันจนกระจายไปทั่วทั้ง เป้า นาน ๆ จึงจะถูกใจกลางของเป้าสักครั้งหนึ่ง ในกรณีนี้จะมี Validity อยู่เล็กน้อย แต่ไม่มี Reliability เป้าที่สามแสดงกรณีที่ปาเป้ากระจายออกไปในส่วนหนึ่งส่วนใดของเป้าและไม่ถูกจุด ศูนย์กลางของเป้า การวัดในกรณีนี้ไม่มีทั้ง Reliability และ Validity ในเป้าสุดท้าย เราจะเห็น "โรบิน ฮูด" ที่สามารถปาเป้าได้ถูกจุดศูนย์กลางของเป้าและการปาแต่ละครั้งมีความสอดคล้องกัน การวัดนี้แสดงว่ามีทั้ง Reliability และ Validity

* กศ.ม. (การวัดผลการศึกษา) <http://www.watpon.com/>

** อาจารย์โรงเรียนธรรมศาสตร์คลองหลวงวิทยาคุม

วิธีที่ 2

อีกวิธีหนึ่งเราจะแสดงความสัมพันธ์ระหว่าง Reliability และ Validity โดยรูปภาพข้างล่าง ในที่นี้เราจะสร้างเป็นตาราง 2x2 แนวสดมภ์ (Column) ของตารางจะแสดงถึง Concept ที่ต้องการวัดว่าเหมือนกันหรือแตกต่างกัน และในแนวแถวจะแสดงวิธีการวัดว่าเหมือนกันหรือแตกต่างกัน ในรูปภาพสมมติว่ามี Concept ที่ต้องการวัดอยู่ 2 Concept คือการวัดความสามารถทางภาษา และการวัดความสามารถทางตัวเลข ยิ่งกว่านั้นสมมติว่าใช้วิธีการวัดที่แตกต่างกัน 2 วิธี คือใช้การสอบโดยเขียนตอบและการสังเกตโดยครู

| | | Concept | |
|---------------|-----------|---|--|
| | | same | different |
| Method | same | reliability verbal = verbal witten = witten | discriminant verbal = math witten = witten |
| | different | convergent verbal = verbal witten = observed | very discriminant verbal = math witten = observed |

ในเซลล์แรกข้างบนทางซ้ายจะแสดงการเปรียบเทียบแบบทดสอบวัดความสามารถทางภาษาด้วยวิธีการเขียนตอบ เราสามารถประมาณค่า Reliability ของแบบทดสอบวัดความสามารถทางภาษาด้วยวิธีการเขียนตอบโดยใช้วิธี test-retest correlation คือการใช้แบบวัดฉบับเดี๋ยวนี้นับสองครั้งแล้วนำคะแนนที่ได้มาหาค่าสหสัมพันธ์ หรืออาจจะใช้วิธี parallel forms คือการใช้แบบทดสอบคู่ขนาน หรือวิธี internal consistency measure คือการวัดความสอดคล้องภายใน เราสามารถประมาณค่าเซลล์แรกได้เป็น Reliability ของการวัด

ในเซลล์ข้างล่างซ้ายจะเปรียบเทียบแบบทดสอบวัดความสามารถทางภาษาด้วยวิธีการเขียนตอบและการสังเกตโดยครู แต่เราวัด Concept เหมือนกันด้วยวิธีการที่ต่างกัน ดังนั้นค่าที่ได้จากการเปรียบเทียบจะเรียกว่า Convergent Validity

ในเซลล์ด้านบนขวาแสดงการเปรียบเทียบแบบทดสอบวัดความสามารถทางภาษาและแบบทดสอบวัดความสามารถทางตัวเลขโดยใช้วิธีการเขียนตอบ ในกรณีนี้เราจะเปรียบเทียบ Concepts ที่แตกต่างกัน (ภาษา กับ ตัวเลข) และเราคาดหวังว่าค่าที่ได้จากการเปรียบเทียบระหว่าง Concept ที่แตกต่างกันควรมีค่าต่ำกว่าการเปรียบเทียบระหว่าง Concept ที่เหมือนกัน (เช่น ภาษา กับ ภาษา หรือ ตัวเลข กับ ตัวเลข) ดังนั้นเราจะต้องพยายามจำแนกระหว่าง Concepts ทั้งสองนี้และเราจะเรียกค่าที่ได้จากการเปรียบเทียบว่า Discriminant Validity

เซลล์สุดท้ายด้านล่างขวา จะเป็นการเปรียบเทียบแบบทดสอบวัดความสามารถทางภาษา

ด้วยวิธีการเขียนตอบและการวัดความสามารถทางด้านตัวเลขด้วยการสังเกตโดยครู จะเห็นว่าใน เซลล์นี้จะเปรียบเทียบ 2 Concepts ที่แตกต่าง (ภาษา กับ ตัวเลข) ซึ่งค่าที่ได้ก็คือ Discriminant Validity แล้ว ก็ยังจะพยายามเปรียบเทียบวิธีการวัด 2 วิธีที่แตกต่างกันอีก (การเขียน กับ การสังเกตโดยครู) ดังนั้นเราจะเรียกมันว่า very discriminant ซึ่งค่าที่ได้ควรจะต่ำกว่าเซลล์อื่น ๆ ทุกเซลล์

ทั้ง 4 เซลล์ข้างต้นสามารถตรวจสอบได้ด้วยวิธี Multitrait-Multimethod ซึ่งเป็นกระบวนการในการประมาณค่าความเที่ยงตรงเชิงโครงสร้าง

จากแนวคิดสู่ความหมายทางภาษาไทย

มีนักวัดผลการศึกษาในประเทศไทยหลายท่านได้แปลคำว่า Validity ว่า “ความเที่ยงตรง”, “ความตรง”, “ความแม่นยำ” และแปล Reliability ว่า “ความเชื่อมั่น”, “ความเที่ยง”, “ความเชื่อถือได้” ทีนี้เรามาดูความหมายของคำภาษาไทยที่แปลออกมากันบ้าง

ตามพจนานุกรมฉบับเฉลิมพระเกียรติ พ.ศ.2530 (ทวีศักดิ์ ญาณประทีป และคณะ : 2534) ได้ให้ความหมายไว้ว่า

เที่ยงตรง (หน้า 263) ว. ยุติธรรม, ไม่เอนเอียง, ตั้งตรง.

ตรง (หน้า 205) ว. ไม่คดโค้ง, ไม่งอ, ไม่เอียง; ชื่อ, ไม่โกง; *เที่ยงตามกำหนด*; เปิดเผยไม่มีลับลมคมใน, ไม่ปิดบังอำพราง, ไม่อ้อมค้อม; รี, ปรี, ถูกต้องตาม.

แม่นยำ (หน้า 422) ก. จำได้อย่างถูกต้อง, ถูกต้องแน่นอน

เชื่อมั่น (หน้า 176) ก. มั่นใจ.

เที่ยง (หน้า 263) ว. ตรง, แน่นนอน; แน่, แม่นยำ.

เชื่อถือ (หน้า 176) ก. นับถือ.

ในเอกสารฉบับนี้ขอแปล Validity ว่า ความเที่ยงตรง และ Reliability ว่า ความเชื่อมั่น

ความเที่ยงตรง (Validity)

ความเที่ยงตรงของแบบทดสอบจะบ่งบอกเราว่าอะไรที่แบบทดสอบวัดและแบบทดสอบนั้นสามารถวัดได้ดีเพียงใด ความเที่ยงตรงจะบอกเราว่าอะไรที่สามารถอ้างอิงได้จากคะแนนการสอบ เรายอมรับว่าชื่อของแบบทดสอบเป็นสิ่งที่บ่งชี้ถึงสิ่งที่แบบทดสอบนั้นวัด ชื่อของแบบทดสอบควรจะสั้น เป็นชื่อย่อ ๆ ที่มีความมุ่งหมายเฉพาะ ชื่อของแบบทดสอบโดยมากมักจะชัดเจนและมีความหมายครอบคลุมไปถึงพฤติกรรมที่จะวัด

กระบวนการทั้งหมดสำหรับการทดสอบความเที่ยงตรงมักจะเป็นความสัมพันธ์ระหว่างคะแนนจากแบบทดสอบกับข้อเท็จจริงที่สังเกตได้อื่น ๆ เกี่ยวกับคุณลักษณะของพฤติกรรมภายในสิ่งที่วัด วิธีการเฉพาะสำหรับใช้ในการตรวจสอบความสัมพันธ์นั้นจะมีชื่อที่หลากหลายแตกต่างกันออกไป

นิยามของ Validity มีผู้ให้นิยามที่แตกต่างกัน ดังนี้

Bingham (Wainer and Other. 1988. p.20 Citing Bingham. 1937, p.214) นิยามว่าเป็นความสัมพันธ์อย่างง่ายของคะแนนแบบทดสอบที่สร้างขึ้นกับแบบทดสอบอื่น ๆ

Guilford (Wainer and Other. 1988. p.20 Citing Guilford. 1946, p.429) นิยาม Validity คล้ายคลึงกันว่า ในความรู้สึกโดยทั่ว ๆ ไปแล้ว การทดสอบความเที่ยงตรงก็คือการหาความสัมพันธ์ Guilford นิยามในปี 1942 ว่าเป็นคุณภาพของแบบทดสอบที่สร้างขึ้นอย่างมีประสิทธิภาพในการทำนายอนาคตของพฤติกรรมและครอบคลุมไปถึงการนิยามว่าความเที่ยงตรงของแบบทดสอบเท่ากับสหสัมพันธ์ของแบบทดสอบที่สร้างกับองค์ประกอบที่วัด ซึ่งไม่เหมือนกับ ความเชื่อมั่น ความเที่ยงตรงไม่ใช่คุณลักษณะทั่ว ๆ ไปของแบบทดสอบ แต่เป็นคุณลักษณะที่มีจุดมุ่งหมายเฉพาะ ดังนั้นแบบทดสอบควรจะมีค่าความเที่ยงตรงสูงในจุดมุ่งหมายหนึ่งแต่ไม่จำเป็นต้องมีความเที่ยงตรงในจุดมุ่งหมายทั้งหมด

Lindquist ในหนังสือสถิติพื้นฐานของปี 1942 นิยาม Validity เหมือนกับคนอื่น ๆ แต่ Bingham นิยามโดยเสนอว่าเป็นความสัมพันธ์ระหว่างการวัดคุณลักษณะที่ต้องการและไม่ต้องการ สมมติว่ามีเกณฑ์ของความเชื่อมั่นอย่างสมบูรณ์และการทดสอบแสดงถึงเกณฑ์อย่างสมบูรณ์แล้ว ค่าความเที่ยงตรงที่สูงที่สุดของแบบทดสอบควรจะมีค่าความสัมพันธ์กับคะแนนของแบบทดสอบที่วัดได้กับคะแนนจริงของแบบทดสอบนั่นเอง ซึ่งในทฤษฎีการทดสอบดั้งเดิม (Classical test theory) ควรจะแสดงรากที่สองของความเชื่อมั่น ซึ่งนี่ก็คือกฎ รากที่สองของความเชื่อมั่นก็คือค่าที่มีได้มากที่สุดของความเที่ยงตรง

Cureton (Wainer and Other. 1988. p.20 Citing Cureton. 1950) ในบทแรกของหนังสือ Educational measurement ได้นิยาม Validity ว่าเป็น สหสัมพันธ์ของคะแนนที่สังเกตได้ของแบบทดสอบกับเกณฑ์ของคะแนนจริง เราแสดงความแตกต่างของการทดสอบความเที่ยงตรงจาก อำนาจการทำนายของแบบทดสอบ โดยนิยามต่อว่าเป็นสหสัมพันธ์ระหว่างคะแนนที่สังเกตได้

ของแบบทดสอบกับเกณฑ์ของคะแนนที่สังเกตได้ และความแตกต่างทั้งคู่จะเรียกว่า *relevance* ซึ่งก็คือสหสัมพันธ์ระหว่างคะแนนจริงของทั้งตัวทำนายและตัวเกณฑ์

การจำแนกประเภทความเที่ยงตรงของบุคคลต่าง ๆ

Lyman (1963. 25-31) ได้แบ่งประเภทของความเที่ยงตรงเป็นดังนี้

1. Face Validity
2. Content Validity
3. Empirical Validity
4. Construct Validity

Allen and Yen (1979. 95-114) ได้แบ่งประเภทของความเที่ยงตรงเป็นดังนี้

1. Content Validity
 - 1.1 Face Validity
 - 1.2 Logical or Sampling Validity
2. Criterion-Related Validity
 - 2.1 Predictive Validity
 - 2.2 Concurrent Validity
3. Construct Validity

Anastasi (1982. 131-155) ได้แบ่งประเภทของความเที่ยงตรงเป็นดังนี้

1. Content Validation
 - 1.1 Face Validation
2. Criterion-Related Validation
 - 2.1 Predictive Validation
 - 2.2 Concurrent Validation
3. Construct Validation
 - 3.1 Convergent Validation
 - 3.2 Discriminant Validation

Popham (1990. 93-117) ได้แบ่งประเภทของความเที่ยงตรงเป็นดังนี้

1. Content-Related Evidence of Validity
2. Criterion-Related Evidence of Validity
 - 2.1 Predictive Validity Evidence
 - 2.2 Concurrent Validity Evidence
3. Construct-Related Evidence of Validity

Trochim (1999) ได้ใช้คำว่า Construct Validity เป็นคำทั่วไปสำหรับเรียกความเที่ยงตรง และได้จำแนกประเภทต่าง ๆ ของความเที่ยงตรงเชิงโครงสร้างได้ดังนี้

1. Translation Validity
 - 1.1 Face Validity
 - 1.2 Content Validity
2. Criterion-related Validity
 - 2.1 Predictive Validity
 - 2.2 Concurrent Validity
 - 2.3 Convergent Validity
 - 2.4 Discriminant Validity

ประเภทของความเที่ยงตรง

1. Content Validity

Anastasi (1982) ได้ให้ความหมายว่า คือการตรวจสอบอย่างเป็นระบบในเนื้อหาของแบบทดสอบที่กำหนด ว่าครอบคลุมตัวอย่างของขอบเขตพฤติกรรมที่ต้องการจะวัดหรือไม่

Popham (1990) ให้ความหมายว่า content-related evidence แสดงถึงระดับของกลุ่มของข้อสอบ งาน หรือคำถามที่บรรจุในแบบทดสอบนั้นถูกสุ่มมาจากบางส่วนของมวลประชากรของข้อสอบ หรือขอบเขตของเนื้อหา

Lyman (1963) ให้ความหมายว่า Content Validity มีชื่อเรียกอีกมากมายเช่น logical validity, course validity, curricular validity และ textbook validity ความเที่ยงตรงเชิงเนื้อหาไม่มีวิธีการสถิติ อย่างไรก็ตามการตรวจสอบเนื้อหาจะต้องมีรายละเอียด เราอาจจะตรวจสอบแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนในแต่ละข้อว่าครอบคลุมมวลความรู้หรือทักษะที่สำคัญในโปรแกรมการเรียนการสอนที่ให้หรือเปล่า หรือเราอาจจะเริ่มต้นกับขอบเขตของรายละเอียดในโปรแกรมการเรียนการสอนดูว่าข้อสอบนั้นครอบคลุมจุดที่สำคัญต่าง ๆ หรือไม่

Content Validity มีความสำคัญมากอย่างเห็นได้ชัดในแบบทดสอบวัดผลสัมฤทธิ์ แต่อาจจะมีความสำคัญกับแบบทดสอบชนิดอื่น ๆ ด้วย

2. Face Validity

Trochim (1999) ให้ความหมายว่า เป็นวิธีที่อ่อนที่สุดที่จะพยายามแสดงหลักฐานความเที่ยงตรง สมมติว่ามีแบบวัดความสามารถทางคณิตศาสตร์อยู่ เมื่ออ่านข้อคำถาม และตัดสินใจว่า ข้อนี้วัดความสามารถทางคณิตศาสตร์ หรือมีแบบวัดเจตคติต่อการทำแท้งแล้วสรุปว่า ข้อนี้สามารถวัดเจตคติได้แน่นอน วิธีเหล่านี้เป็น face validity เป็นวิธีการที่มีหลักฐานแสดงความเที่ยงตรงที่อ่อนที่สุดเพราะว่าเป็นการตัดสินใจที่ขึ้นอยู่กับบุคคล และบุคคลที่จะมาตัดสินว่าข้อคำถามวัดคุณลักษณะนั้น ๆ ควรจะเป็นผู้ตัดสินที่มีความน่าเชื่อถือ

เราสามารถเสริมให้ face validity มีคุณภาพได้โดยการทำอย่างเป็นระบบ ตัวอย่างเช่น ถ้าจะประเมิน face validity ของแบบวัดความสามารถทางคณิตศาสตร์ และคุณส่งแบบวัดของคุณไปยังผู้เชี่ยวชาญที่น่าเชื่อถือซึ่งเลือกมาอย่างพิถีพิถันเพื่อตรวจสอบและตัดสินแบบวัดของคุณ และเมื่อผลจากการตัดสินของผู้เชี่ยวชาญถูกส่งกลับมาและปรับแก้ไขตามผู้เชี่ยวชาญแล้วละก็ แบบวัดความสามารถทางคณิตศาสตร์ฉบับนี้ก็จะกลายเป็นแบบวัดที่ดีที่สุดวัดความสามารถทางคณิตศาสตร์ได้อย่างดี

3. Logical Validity

ลัวิน และ อังคณา สายยศ (2539 : 247-250) กล่าวว่า ความเที่ยงตรงเชิงเหตุผล (Logical Validity) เป็นความเที่ยงตรงที่ให้ผู้เชี่ยวชาญพิจารณาว่าข้อสอบแต่ละข้อนั้นวัดได้ตรงตามตารางวิเคราะห์รายละเอียด (Table of Specifications) หรือไม่ ถ้าเป็นแบบทดสอบวัดผลสัมฤทธิ์แบบอิงกลุ่ม ผู้เชี่ยวชาญทางสาขาวิชานั้นจะต้องพิจารณาว่าแบบทดสอบฉบับนั้นมีข้อสอบแต่ละข้อตรงตามพฤติกรรมที่จะวัดและจำนวนข้อสอบคล้องกับตารางวิเคราะห์รายละเอียดหรือไม่ ส่วนแบบทดสอบวัดผลสัมฤทธิ์แบบอิงเกณฑ์นั้น ผู้เชี่ยวชาญทางสาขาวิชาจะต้องพิจารณาว่า ข้อสอบของแบบทดสอบที่สร้างขึ้นนั้นวัดได้ตรงตามจุดประสงค์เชิงพฤติกรรมหรือไม่

4. Criterion-Related Validity

Anastasi (1982) ให้ความหมายว่า เป็นกระบวนการที่ชี้ให้เห็นถึงความรับผิดชอบของแบบทดสอบในการทำนายพฤติกรรมของบุคคลในสถานการณ์เฉพาะ สำหรับจุดมุ่งหมายคือการตรวจสอบโดยใช้เกณฑ์ (criterion) อาจจะทำได้ทั้งการวัดโดยตรงและการวัดโดยอ้อมโดยใช้แบบทดสอบที่ถูกออกแบบมาเพื่อทำนาย ดังนั้นสำหรับแบบทดสอบวัดความถนัดทางช่างกล เกณฑ์อาจจะเป็นผลของการปฏิบัติได้เหมือนช่าง หรือสำหรับแบบทดสอบความถนัดทางวิชาการ อาจจะใช้เกรดเฉลี่ย และการวัดทางจิตอาจจะใช้การประเมินหรือข้อมูลอื่นที่พิจารณาได้จาก

พฤติกรรมของแต่ละคนในชีวิตประจำวัน

Popham (1990) ให้ความหมายว่า Criterion-related evidence แสดงถึงคะแนนของแบบทดสอบที่สัมพันธ์อย่างเป็นระบบกับเกณฑ์อย่างหนึ่งหรือหลายอย่าง ในเนื้อหาของเกณฑ์ก็คือตัวแปรที่เราสนใจ ซึ่งถูกนิยามโดยระบบของโรงเรียน การจัดการของบริษัท หรือลูกค้า การเลือกเกณฑ์และกระบวนการวัดที่ใช้เพื่อให้ได้คะแนนเกณฑ์มา ในความเป็นจริงแล้ว ค่าของความเที่ยงตรงเชิงเกณฑ์ขึ้นอยู่กับความเกี่ยวข้องกับการใช้เกณฑ์การวัด

5. Concurrent and Predictive Validity

Trochim (1999) ให้ความหมายของความเที่ยงตรงทั้ง 2 ชนิดดังนี้

Predictive Validity เป็นการประเมินความสามารถเพื่อทำนายบางอย่างในเชิงทฤษฎี เช่น เราอาจจะมีทฤษฎีว่าการวัดความสามารถทางคณิตศาสตร์ควรจะสามารถทำนายความสามารถทางวิศวกรรมของบุคคลได้ดี เราควรจะใช้การวัดของเราสัมพันธ์กับความสามารถทางวิศวกรรม และหากมีความสัมพันธ์กันสูงระหว่างการวัดความสามารถทางคณิตศาสตร์กับความสามารถทางวิศวกรรม ความสัมพันธ์กันสูงนี้ควรจะเป็นหลักฐานแสดงความเที่ยงตรงเชิงทำนาย แสดงว่าการวัดของเรานั้นมีความถูกต้องสามารถทำนายได้จริงตามทฤษฎี

Concurrent Validity เป็นการประเมินความสามารถที่แตกต่างระหว่างกลุ่มตามทฤษฎี ตัวอย่างเช่น ถ้าเรามีแบบทดสอบวัดความซึมเศร้า แบบวัดของเราควรจะสามารถจำแนกความแตกต่างระหว่างบุคคลที่วินิจฉัยแล้วว่ามีความซึมเศร้าและบุคคลที่วินิจฉัยแล้วว่าไม่ซึมเศร้า ถ้าเราต้องการประเมินความเที่ยงตรงตามสภาพให้มีสูง ๆ เราอาจจะใช้แบบวัดของเรากับพนักงานในฟาร์มและเจ้าของฟาร์ม ตามทฤษฎีของเราควรจะได้ว่าเจ้าของฟาร์มควรจะได้คะแนนสูงกว่า ผลที่ได้จะยิ่งเชื่อถือว่ามีความเที่ยงตรงเชิงสภาพมาก ถ้าสามารถใช้แบบวัดนั้นจำแนกความแตกต่างระหว่างกลุ่ม 2 กลุ่มที่มีคุณลักษณะใกล้เคียงกันได้

6. Construct Validity

Anastasi (1982) ให้ความหมายว่า ความสามารถของแบบทดสอบที่สามารถวัดโครงสร้างทฤษฎีหรือคุณลักษณะ เช่น โครงสร้างของเขาวนปัญญา, ความถนัดทางช่างกล, ความคล่องแคล่วทางภาษา หรือความวิตกกังวล เป็นต้น แต่ละโครงสร้างได้พัฒนาขึ้นเพื่อการอธิบายหรือการจัดกระบวนการสังเกตอย่างสม่ำเสมอ ซึ่งค้นหาได้จากการสร้างความสัมพันธ์ร่วมกันในการวัดพฤติกรรม

Popham (1990) ให้ความหมายว่า construct-related evidence ว่าเป็นกลุ่มที่สนใจเป็นพื้นฐานของคะแนนแบบทดสอบที่วัดคุณลักษณะทางจิตวิทยาที่สนใจ ทั้งความสามารถทางเหตุผล มิติสัมพันธ์ และความเข้าใจในการอ่าน ก็คือโครงสร้าง คือคุณลักษณะของบุคลิกภาพเช่นการเข้าสังคมและการเก็บตัว ความอดทนที่ใช้มากในโครงสร้างของนักกีฬา การศึกษาพฤติกรรมความเป็น

ผู้นำก็มีใช้อยู่เสมอในการอ้างอิงโครงสร้างเช่นการพิจารณาพฤติกรรมที่รอง ๆ ลงไป (เช่น การให้ คำสรรเสริญ, การอธิบายเหตุผลของการกระทำ, การถามความคิดเห็น) และโครงสร้างพื้นฐาน (เช่น การตั้งเป้าหมาย, การกำหนดวางแผน) คุณลักษณะที่ถูกอ้างอิงว่าเป็นโครงสร้างเพราะว่าคุณ ลักษณะเหล่านี้เป็นโครงสร้างทางทฤษฎีเกี่ยวกับธรรมชาติของพฤติกรรมมนุษย์

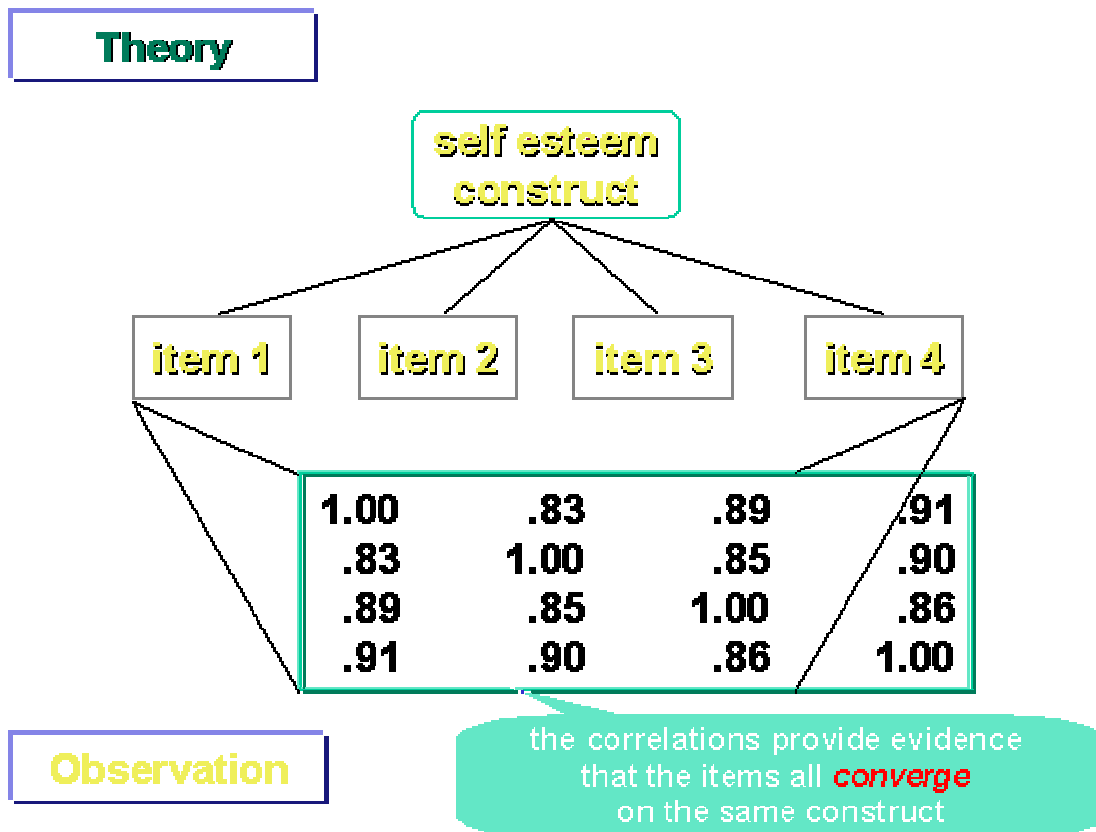
Trochim (1999) ให้ความหมายว่า ความเที่ยงตรงเชิงโครงสร้างเป็นระดับที่อ้างอิงถึงความถูกต้องของแบบวัดที่ยึดโครงสร้างของทฤษฎีเป็นฐาน ซึ่งคล้ายคลึงกับความเที่ยงตรงภายนอก (External Validity) ซึ่งความเที่ยงตรงเชิงโครงสร้างก็เกี่ยวข้องกับการสรุปอ้างอิงเช่นเดียวกัน แต่ความเที่ยงตรงภายนอกจะสรุปอ้างอิงจากการศึกษาของคุณไปยังบุคคลอื่น ๆ ทั้งสถานที่ และเวลา ส่วนความเที่ยงตรงเชิงโครงสร้างจะสรุปอ้างอิงจากการวัดของคุณไปยังมโนภาพ (Concept) ของการวัดของคุณ

หลักฐานที่แสดงให้เห็นถึงความเที่ยงตรงเชิงโครงสร้างนั้น ประการแรกผู้สร้างแบบวัด จะต้องตั้งสมมติฐานโครงสร้างเชิงทฤษฎี (Hypothetical Construct) เช่น “ความรักสัตว์” การสร้างสมมติฐานโครงสร้างเชิงทฤษฎีจะต้องตั้งอยู่บนประสบการณ์การวิจัยของผู้สร้าง ประการสอง พัฒนาแบบทดสอบให้วัดตามโครงสร้าง โดยมีชื่อของแบบวัดว่า แบบสำรวจความรัก สัตว์ เช่น เราอาจจะสร้างแบบวัด 30 ข้อสำหรับวัด “ความรักสัตว์” ผู้ที่ได้คะแนนสูงในแบบวัดนี้จะ บ่งชี้ว่าเป็นผู้มีความรักสัตว์ และผู้ที่ได้คะแนนต่ำจะบ่งชี้ว่าเป็นผู้ไม่มีความรักสัตว์

7. Convergent Validity

Trochim (1999) ให้ความหมายว่า เป็นระดับของความคล้ายคลึงกันของแบบวัดที่ควร จะคล้ายคลึงกันตามทฤษฎี ตัวอย่างเช่น ความเที่ยงตรงเชิงเหมือนของแบบทดสอบวัดความสามารถทางคำนวณ เราอาจจะไปสัมพันธ์คะแนนของแบบวัดความสามารถทางคำนวณฉบับอื่น ๆ ที่มีจุดมุ่งหมายเพื่อวัดความสามารถทางคณิตศาสตร์เหมือนกันฉบับที่เราสร้าง เมื่อเราหา สหสัมพันธ์ระหว่างคะแนนของแบบวัดทั้ง 2 ฉบับแล้วมีความสัมพันธ์กันสูงก็เป็นหลักฐานบ่งชี้ว่า มีความเที่ยงตรงเชิงเหมือน

จากรูปภาพ 1 เราจะเห็นสหสัมพันธ์ของข้อสอบ 4 ข้อซึ่งวัดในโครงสร้างเดียวกัน ซึ่งอาจจะไม่หมายความว่าโครงสร้างนั้นคือ Self-esteem แต่อาจจะมีโครงสร้างอื่น ๆ ที่ข้อสอบ 4 ข้อนี้สัมพันธ์กัน ค่าสหสัมพันธ์ที่คำนวณได้มีค่าสูงมาก เป็นหลักฐานแสดงว่าข้อสอบทั้งหมดมีความสัมพันธ์กันในโครงสร้างที่เหมือนกัน

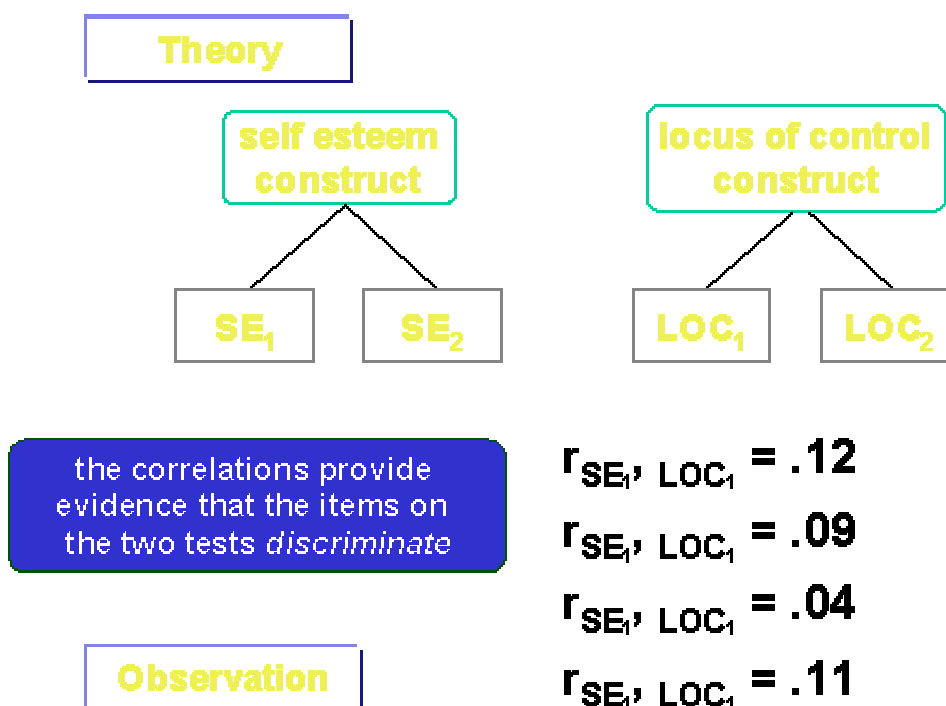


รูปภาพ 1

8. Discriminant Validity

Trochim (1999) ให้ความหมายว่า เป็นระดับของความแตกต่างกันของแบบวัดที่ควรจะแตกต่างกันตามทฤษฎี ตัวอย่างเช่น ความเที่ยงตรงเชิงจำแนกของแบบทดสอบวัดความสามารถทางคำนวณ เราอาจจะไปหาความสัมพันธ์กับคะแนนของแบบวัดความสามารถทางภาษา เมื่อแบบวัดทั้งฉบับนี้มีความสัมพันธ์ต่ำก็เป็นหลักฐานบ่งชี้ว่ามีความเที่ยงตรงเชิงจำแนก

ในรูปภาพ 2 จะมีข้อสอบอยู่ 4 ข้อ จะมีอยู่ 2 ข้อที่วัดโครงสร้างความภาคภูมิใจในตนเอง (Self Esteem Construct) และอีก 2 ข้อวัดการควบคุมตนเอง (Locus of Control Construct) เมื่อหาสหสัมพันธ์ระหว่างข้อสอบที่มีโครงสร้างต่างกันเราคาดหวังว่าจะมีสหสัมพันธ์กันต่ำเป็นการแสดงให้เห็นถึงความเที่ยงตรงเชิงจำแนก



รูปภาพ 2

9. Factorial Validity

Allen and Yen (1979) ให้ความหมายว่า Factorial Validity เป็นรูปแบบของความเที่ยงตรงเชิงโครงสร้างซึ่งสามารถหาได้โดยการใช้การวิเคราะห์องค์ประกอบ (Factor Analysis)

กระบวนการหาความเที่ยงตรง

1. ตารางวิเคราะห์หลักสูตร

เป็นวิธีการหาความเที่ยงตรงเชิงเหตุผล (Logical Validity) เหมาะสำหรับแบบทดสอบอิงกลุ่ม โดยใช้ผู้เชี่ยวชาญในสาขาวิชานั้น ๆ พิจารณาว่าแบบทดสอบฉบับนั้นมีข้อสอบแต่ละข้อตรงตามพฤติกรรมที่จะวัดและจำนวนข้อสอดคล้องกับตารางวิเคราะห์รายละเอียด (Table of Specifications) หรือไม่ ดังตัวอย่าง

| เนื้อหา | พฤติกรรม | | | รวม |
|----------------------------|------------|------------|--------------|-----|
| | ทักษะจำนวน | การนำไปใช้ | การวิเคราะห์ | |
| 1. การบวกเลขไม่เกิน 1 หลัก | 3 | 7 | 2 | 12 |
| 2. การบวกเลข 2 หลัก | 3 | 6 | 3 | 12 |
| 3. การบวกเลขมากกว่า 2 หลัก | 3 | 3 | 2 | 8 |
| 4. โจทย์การบวกเลข | 2 | 4 | 2 | 8 |
| รวม | 11 | 20 | 9 | 40 |

จากตาราง ผู้เชี่ยวชาญในสาขาวิชานั้น จะต้องพิจารณาว่าข้อสอบแต่ละข้อและจำนวนข้อของแบบทดสอบที่สร้างขึ้นนั้นสอดคล้องกับตารางวิเคราะห์รายละเอียดหรือไม่ ถ้าสอดคล้องกันก็แสดงว่าแบบทดสอบที่สร้างขึ้นมีความเที่ยงตรงเชิงเหตุผล ซึ่งเป็นวิธีที่ใช้กับความเที่ยงตรงของแบบทดสอบวัดผลสัมฤทธิ์แบบอิงกลุ่ม

2. พิจารณาความสอดคล้องระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรม

สำหรับแบบทดสอบวัดผลสัมฤทธิ์แบบอิงเกณฑ์นั้น ผู้เชี่ยวชาญทางสาขาวิชาจะต้องพิจารณาว่า ข้อสอบของแบบทดสอบที่สร้างขึ้นนั้นวัดได้ตรงตามจุดประสงค์เชิงพฤติกรรมหรือไม่ ดังตัวอย่างตารางการพิจารณาข้อสอบดังนี้

| จุดประสงค์เชิงพฤติกรรม | ข้อสอบ | การพิจารณา | | |
|---|---|------------|---|----|
| | | +1 | 0 | -1 |
| เมื่อกำหนดตัวเลขไม่เกิน 2 หลักมาให้ 2 จำนวนบวกกัน นักเรียนสามารถบวกได้ถูกต้อง | 15 + 18 มีค่าเท่าไร ก. 22 ข. 23 ค. 33 ง. 35 จ. ไม่มีข้อใดถูก | | | |

จากตาราง ผู้เชี่ยวชาญทางด้านวิชาที่สร้างแบบทดสอบจะพิจารณาว่าข้อสอบวัดได้ตรงตามจุดประสงค์เชิงพฤติกรรมหรือไม่ ถ้าแน่ใจว่า “ตรง” จะกาเครื่องหมายในช่อง “+1” ถ้าแน่ใจว่า “ไม่ตรง” จะกาเครื่องหมายในช่อง “-1” และถ้าไม่แน่ใจว่า “ตรงหรือไม่” จะกาเครื่องหมายในช่อง “0” แล้วนำค่ามาหาดัชนีความสอดคล้อง (Index of Consistency : IOC) $IOC = s/RN$ ถ้าค่าที่คำนวณได้มีค่าตั้งแต่ .50 ขึ้นไปถือว่าข้อสอบข้อนั้นวัดตรงตามจุดประสงค์เชิงพฤติกรรม

3. สัมประสิทธิ์สหสัมพันธ์

ในการศึกษาความเที่ยงตรงเชิงทำนายและเชิงสภาพ (Predictive and Concurrent Validity) มักจะรายงานเป็นสัมประสิทธิ์สหสัมพันธ์เสมอ สถิตินี้จะใช้สำหรับหาความสัมพันธ์ระหว่างตัวแปร 2 ตัว ความเที่ยงตรงเชิงทำนายและเชิงสภาพมักจะเป็นการหาสหสัมพันธ์ระหว่างแบบทดสอบกับเกณฑ์เสมอ ขึ้นอยู่กับว่าเกณฑ์นั้นจะเป็นเกณฑ์ปัจจุบันหรือเกณฑ์ในอนาคต ค่าสหสัมพันธ์ที่ได้จะเรียกว่า สัมประสิทธิ์ความเที่ยงตรง (Validity Coefficient)

แม้ว่าค่าสัมประสิทธิ์จะมีค่าสูง และมีสหสัมพันธ์เป็นบวกในการทำนายของแบบทดสอบก็ตาม แต่สัมประสิทธิ์ความเที่ยงตรงที่สูงเพียงพอก็ไม่ได้ประกันการใช้แบบทดสอบสำหรับการทำนาย ซึ่งขึ้นอยู่กับองค์ประกอบต่าง ๆ ที่จะต้องพิจารณาด้วยเช่น การเลือกใช้วิธีการที่เหมาะสม และต้นทุนของการทดสอบ

ถ้าหากจะศึกษาความเที่ยงตรงเชิงโครงสร้าง อาจพิจารณาได้จากการคำนวณค่าสหสัมพันธ์ระหว่างแบบทดสอบฉบับที่เราต้องหาความเที่ยงตรงกับแบบทดสอบมาตรฐานที่วัดในคุณลักษณะเดียวกัน

อีกวิธีหนึ่งก็คือการคำนวณหาสหสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวมทั้งฉบับ ค่าสหสัมพันธ์ที่ได้จะมีจำนวนเท่ากับจำนวนของข้อสอบ ให้นำค่าสหสัมพันธ์ไปแปลงเป็นฟิชเชอร์ซี (Fisher'Z) แล้วหาค่าฟิชเชอร์ซีเฉลี่ย (\bar{Z}) จากนั้นนำค่าฟิชเชอร์ซีเฉลี่ยนี้ไปเปิดตารางแปลงกลับเป็นค่าสหสัมพันธ์ ค่าสหสัมพันธ์ที่ได้ก็คือค่าสัมประสิทธิ์ความเที่ยงตรง

4. เทคนิคกลุ่มรู้ชัด (Known-Group Technique)

เป็นวิธีเปรียบเทียบคะแนนเฉลี่ยระหว่างกลุ่มที่รู้ว่ามึลักษณะที่ต้องการวัดอยู่แล้วกับกลุ่มที่เราว่าไม่มีคุณลักษณะที่ต้องการวัด เช่น หากจะหาความเที่ยงตรงเชิงโครงสร้างของแบบวัดความเหงาหงอย อาจทำโดยนำแบบวัดความเหงาหงอย ไปทดสอบกับกลุ่มตัวอย่างในโรงพยาบาลที่แพทย์ระบุว่า เป็นโรคเหงาหงอย (กลุ่มรู้ชัดว่ามีลักษณะที่ต้องการวัด) กับกลุ่มคนปกติทั่วไป (กลุ่มรู้ชัดว่าไม่มีหรือมีคุณลักษณะที่ต้องการวัดน้อย) แล้วนำคะแนนที่ได้มาทดสอบนัยสำคัญทางสถิติ

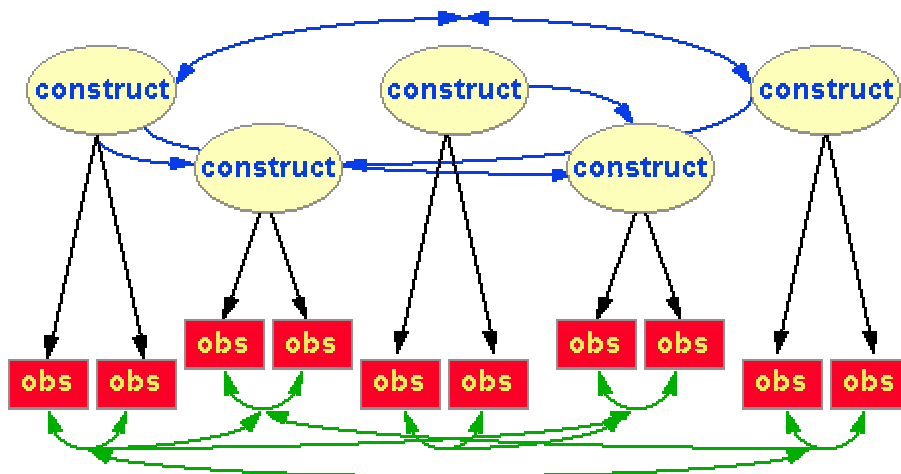
5. Nomological Network

Nomological network เป็นแนวคิดที่พัฒนาโดย Lee Cronbach และ Pual Meehl ในปี 1955 คำว่า Nomological เป็นภาษากรีก หมายความว่า คุณสมบัติที่ครบถ้วน (lawful) ดังนั้นอาจจะเรียกได้อีกชื่อว่า lawful network คำว่า Nomological Network ทั้ง Cronbach และ Meehl หมายถึงความเที่ยงตรงเชิงโครงสร้าง นั่นคือใช้เพื่อกำหนดหลักฐานว่าการวัดของคุณมีความเที่ยงตรงเชิงโครงสร้าง (Construct Validity) คำว่า network นี้จะรวมไปถึงโครงสร้างทางทฤษฎีที่

คุณพยายามจะวัด และมีความเฉพาะของการเชื่อมโยงระหว่างโครงสร้าง 2 โครงสร้าง มีผู้วิจัยได้ศึกษาเกี่ยวกับ Nomological Network และได้แนะนำว่า

The Nomological Network

a representation of the concepts (constructs) of interest in a study,



...their observable manifestations, *and the interrelationships among and between these*

รูปภาพ 3

- ต้องให้ชัดเจนว่าสิ่งนั้นคืออะไร หมายถึงอะไร ซึ่งกฎสามารถจัดกระทำกับสิ่งนั้นได้อย่างมีหลักเกณฑ์
- กฎใน Nomological Network จะสัมพันธ์กับ
 - คุณลักษณะของตัวแปรสังเกตหรือปริมาณของตัวแปร
 - ความแตกต่างของโครงสร้างทฤษฎี
 - โครงสร้างทฤษฎีนำไปสู่การสังเกต
- จะต้องมีบางกฎที่เกี่ยวข้องกับตัวแปรสังเกต
- การศึกษาโครงสร้างทางทฤษฎีให้มากคือสาระสำคัญของการกำหนด Nomological network ซึ่งองค์ประกอบต้องเกิดขึ้นหรือมีเพิ่มขึ้นแน่นอน
- กฎพื้นฐานที่รวมเข้ากับโครงสร้างใหม่หรือที่สัมพันธ์กับทฤษฎีนั้นต้องเป็นกฎที่ยืนยันโดยการสังเกตหรือการลดจำนวน Nomological network ที่จะทำนายตัวแปรสังเกตบางตัว

- การกระทำที่เป็นความแตกต่างเชิงคุณภาพที่ซ้อนกัน (overlap) หรือวัดสิ่งเดียวกัน ถ้าสภาพของมันใน Nomological network ผูกโยงกัน มันก็เป็นตัวแปรโครงสร้างเดียวกัน

Cronbach และ Meehl พยายามที่จะทำการเชื่อมโยงมโนทัศน์หรือทฤษฎีกับตัวแปร สังเกต เพราะว่าสิ่งนี้คือแก่นของความเที่ยงตรงเชิงโครงสร้าง แนวคิดขณะนี้ของ nomological network ทำงานอยู่บนพื้นฐานทางปรัชญาของความเที่ยงตรงเชิงโครงสร้าง ซึ่งปัจจุบันนี้ยังไม่มี การนำไปใช้ในทางปฏิบัติและยังไม่มีผู้คิดหาวิธีการที่จะนำแนวคิดนี้ไปใช้ในการประเมินความ เที่ยงตรงเชิงโครงสร้างอย่างแท้จริง จากแนวคิดนี้เป็นผลที่นำไปสู่แนวคิดของการหาความเที่ยง ตรงเชิงโครงสร้าง ซึ่งพัฒนาออกมาในรูปของเมตริกซ์หลายคุณลักษณะหลายวิธี (Multitrait-Multimethod Matrix)

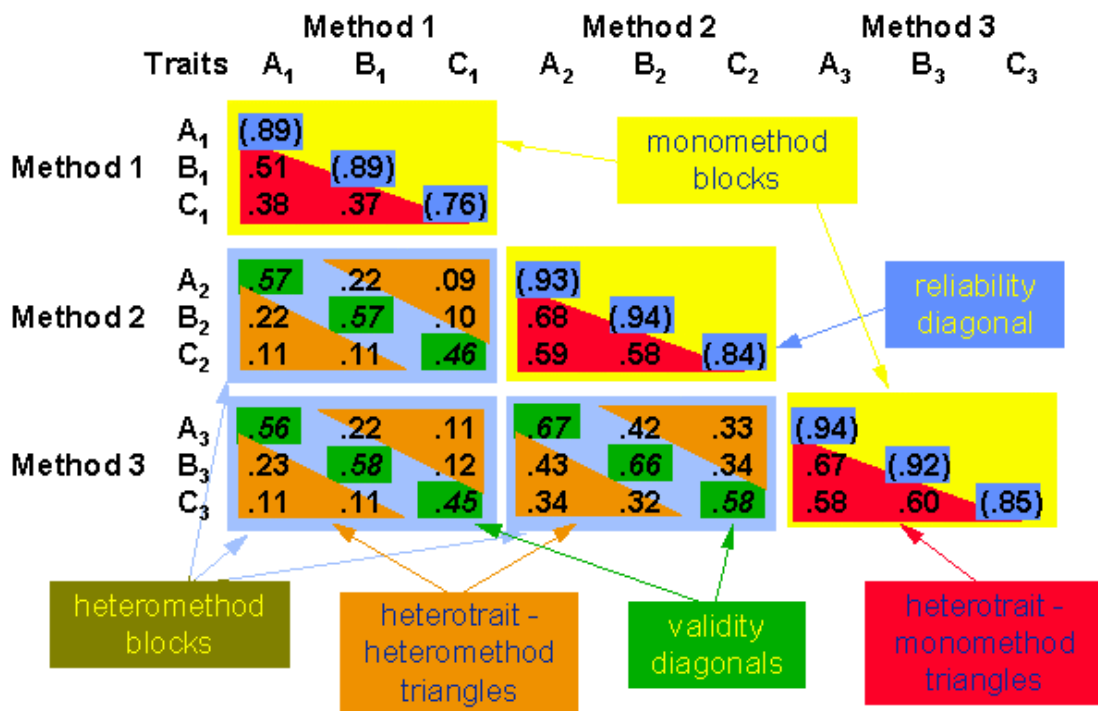
6. เมตริกซ์หลายคุณลักษณะหลายวิธี

เมตริกซ์หลายคุณลักษณะหลายวิธี (Multitrait-Multimethod Matrix) (ต่อจากนี้จะใช้คำ ย่อว่า MTMM) เป็นกระบวนการประเมินความเที่ยงตรงเชิงโครงสร้างของเครื่องมือวัด คุณลักษณะ MTMM พัฒนาโดย Campbell และ Fiske ผู้ได้พยายามคิดค้นวิธีดำเนินการในทาง ปฏิบัติเพื่อให้นักวิจัยสามารถนำไปใช้ได้ ใน MTMM นั้น Campbell และ Fiske ได้เสนอแนะ ความเที่ยงตรงแบบใหม่ 2 แบบคือ ความเที่ยงตรงเชิงเหมือนและความเที่ยงตรงเชิงจำแนก (convergent and discriminant) ซึ่งเป็นส่วนหนึ่งของความเที่ยงตรงเชิงโครงสร้าง ผู้วิจัยสามารถ ประเมินได้ทั้งความเที่ยงตรงเชิงเหมือนและเชิงจำแนกโดยการใช้ MTMM และยังสามารถอ้างได้ ว่าเครื่องมือวัดมีความเที่ยงตรงเชิงโครงสร้าง เนื่องจากมีหลักฐานแสดงทั้งความเที่ยงตรงเชิง เหมือนและความเที่ยงตรงเชิงจำแนก

MTMM เมื่ออยู่ในรูปของเมตริกซ์หรือตารางแสดงความสัมพันธ์แล้ว การประเมินความ เที่ยงตรงเชิงโครงสร้างก็จะง่ายขึ้น ภายใน MTMM จะแสดงชุดของข้อมูลที่วัดในแต่ละคุณลักษณะ (traits) ที่หลากหลาย โดยใช้วิธีการวัดหลายวิธี (เช่น การสอบแบบเขียนตอบ (paper and pencil test), การสังเกตโดยตรง (direct observation), การวัดการปฏิบัติ (performance measure)) แต่ MTMM ก็เป็นวิธีที่มีข้อจำกัดคือจะต้องวัดแต่ละคุณลักษณะในทุกวิธีการที่กำหนด

โครงสร้างของ MTMM มีลักษณะเป็นเมตริกซ์สหสัมพันธ์ ดังรูปภาพ 1 แสดง MTMM สำหรับ 3 คุณลักษณะ (คุณลักษณะ A, B และ C) แต่ละคุณลักษณะจะวัดด้วยวิธีการที่แตกต่าง กัน 3 วิธี (วิธีที่ 1, 2 และ 3) สังเกตว่าค่าที่แสดงในกรอบแต่ละกรอบ คือวิธีการแต่ละวิธี MTMM จะแสดงเมตริกซ์สหสัมพันธ์ระหว่างชุดข้อมูล ยกเว้นในแนวไดอะกอนอลของเมตริกซ์ซึ่ง โดยปกติควรจะมีค่าเป็น 1 เพราะสหสัมพันธ์ระหว่างตัวมันเองจะมีค่าเป็น 1.00 เสมอ เราจะแทน ที่ค่านี้ด้วยค่าความเชื่อมั่นของเครื่องมือแต่ละชุดข้อมูล ก่อนที่ผู้วิจัยจะแปลความหมายของ

MTMM จะต้องทำความเข้าใจในส่วนต่าง ๆ ของเมตริกซ์เสียก่อน อันดับแรก จะสังเกตว่าค่าในเมตริกซ์จะแสดงค่าสหสัมพันธ์ มีลักษณะเป็นเมตริกซ์สมมาตร ดังนั้นเราจะมองเพียงครึ่งหนึ่งของเมตริกซ์ (ในรูปภาพ 4 จะแสดงเพียงครึ่งล่าง) อันดับที่สอง สหสัมพันธ์จะแยกออกเป็น 3 กลุ่มใหม่ แสดงด้วยรูปแบบ 3 แบบคือ ไต่อะกอนอล, สามเหลี่ยม และสี่เหลี่ยม



รูปภาพ 4

ไต่อะกอนอลที่แสดงความเชื่อมั่น (The Reliability Diagonal)

(คุณลักษณะเดียวกัน-วิธีการเดียวกัน)

การประมาณค่าความเชื่อมั่นสำหรับเครื่องมือวัดแต่ละชุด ผู้วิจัยสามารถเลือกวิธีประมาณค่าความเชื่อมั่นได้หลายวิธี (เช่น สอบซ้ำ, ความสอดคล้องภายใน ฯ) เนื่องจากเครื่องมือมี 9 ชุด เราจึงมีความเชื่อมั่น 9 ค่า ความเชื่อมั่นค่าแรกในตัวอย่าง (.89) เป็นค่าความเชื่อมั่นระหว่างคุณลักษณะ A วิธีการที่ 1 กับคุณลักษณะ A วิธีการที่ 1 (ต่อไปนี้จะเรียกย่อ ๆ ว่า A1-A1) หรือก็คือความเชื่อมั่นของเครื่องมือวัดคุณลักษณะ A ที่วัดด้วยวิธีการที่ 1 นั่นเอง ค่าความเชื่อมั่นค่าที่สอง (.89) เป็นค่าความเชื่อมั่นของเครื่องมือวัดคุณลักษณะ B ด้วยวิธีการที่ 1 ที่นี้ลองพิจารณาค่าความเชื่อมั่นอื่น ๆ แล้วลองตอบดูว่า ค่าความเชื่อมั่นของเครื่องมือวัดคุณลักษณะ C ที่วัดด้วยวิธีการที่ 3 มีความเชื่อมั่นเท่าไร?

ไดอะกอนอลที่แสดงความเที่ยงตรง (The Validity Diagonals)

(คุณลักษณะเดียวกัน-วิธีการต่างกัน)

สหสัมพันธ์ระหว่างการวัดของคุณลักษณะที่เหมือนกันโดยใช้วิธีการที่แตกต่างกัน MTMM จะแสดงค่าสหสัมพันธ์ภายในกล่องสี่เหลี่ยม นั่นคือความเที่ยงตรงจะแสดงในเมตริกซ์ ไดอะกอนอลที่อยู่ภายในกล่องสี่เหลี่ยมแต่ละกล่อง ในรูปภาพ 1 ดูที่ A1-A2 สหสัมพันธ์เป็น 0.57 นั่นคือสหสัมพันธ์ระหว่างวิธีการวัด 2 วิธี (1 และ 2) ที่วัดคุณลักษณะเดียวกัน (A) ค่าเหล่านี้เราคาดหวังว่าจะมีความสัมพันธ์กันสูง นั่นหมายถึงว่ายังมีค่าสหสัมพันธ์สูงยิ่งจะทำให้มีความเที่ยงตรงสูง ความเที่ยงตรงในที่นี้ก็คือ "ความเที่ยงตรงเชิงเหมือน" ที่นี้ลองพิจารณา ค่าอื่น ๆ ในคุณลักษณะเดียวกันแต่วิธีการต่างกัน

สามเหลี่ยมที่แสดงคุณลักษณะต่างกันแต่วิธีการเดียวกัน

(The Heterotrait-Monomethod Triangles)

เป็นสหสัมพันธ์ระหว่างการวัดคุณลักษณะต่างกันแต่ใช้วิธีการวัดเดียวกัน ตัวอย่างเช่น $A1-B1 = 0.51$ อยู่ในส่วนบนทางซ้ายของสามเหลี่ยมคุณลักษณะต่างกัน-วิธีการเดียวกัน สังเกตว่าเป็นสหสัมพันธ์ของคุณลักษณะที่ต่างกัน (คุณลักษณะ A และ B) แต่วิธีการเดียวกัน (วิธีการที่ 1) ค่าเหล่านี้เราคาดหวังว่าจะมีความสัมพันธ์กันต่ำ ซึ่งหมายถึงมี "ความเที่ยงตรงเชิงจำแนก" สูง ถ้าสหสัมพันธ์มีค่าสูง นั่นคงเป็นเพราะประสิทธิภาพของวิธีที่ใช้วัด

สามเหลี่ยมที่แสดงคุณลักษณะต่างกันและวิธีการต่างกัน

(The Heterotrait-Heteromethod Triangles)

เป็นสหสัมพันธ์ระหว่างคุณลักษณะและวิธีการที่แตกต่างกัน เช่น $A1-B2$ มีค่าสหสัมพันธ์ 0.22 โดยทั่วไปสหสัมพันธ์ของการวัดคุณลักษณะที่ต่างกันด้วยวิธีการที่ต่างกันย่อมมีค่าต่ำที่สุดในเมตริกซ์

กล่องสี่เหลี่ยมที่แสดงวิธีการเดียวกัน (The Monomethod Blocks)

ค่าสหสัมพันธ์ทั้งหมดเกิดจากวิธีการวัดที่เหมือนกัน นั่นคือกล่องสี่เหลี่ยมที่อยู่ในแนวทแยงแต่ละกล่องจะคลุมวิธีการวัดเดียวกัน

กล่องสี่เหลี่ยมที่แสดงวิธีการวัดต่างกัน (The Heteromethod Blocks)

ค่าสหสัมพันธ์ทั้งหมดเกิดจากวิธีการวัดต่างกัน นั่นคือจะมี $(K(K-1))/2$ กล่องที่แสดงค่าสหสัมพันธ์ของวิธีการวัดต่างกัน เมื่อ $K =$ จำนวนวิธีการวัด ในตัวอย่างนี้มีวิธีการวัด 3 วิธี ดังนั้นจะได้กล่องสี่เหลี่ยมที่แสดงค่าสหสัมพันธ์ของวิธีการวัดต่างกันจำนวน $(3(3-1))/2 =$

$(3(2))2 = 6/2 = 3$ กล่อง ก็คือกล่อง 3 กล่องที่อยู่ครึ่งล่างของเมตริกซ์

หลักการแปลความหมาย

ขณะนี้เราได้เรียนรู้ส่วนต่าง ๆ ใน MTMM แล้ว ในการแปลความหมาย ผู้วิจัยจะต้องเป็นผู้ตัดสินใจแปลความหมายของ MTMM ด้วยตนเอง โดยยึดหลักการหรือกฎในการแปลความหมาย แต่ก็มีบางครั้งที่กฎการแปลความหมายบางอย่างจะถูกฝ่าฝืน ผู้วิจัยอาจจะแปลผลได้อย่างราบรื่นว่าเครื่องมือวัดมีความเที่ยงตรงเชิงโครงสร้างอย่างดี โดยที่ผู้วิจัยไม่จำเป็นต้องยึดมั่นมากในการนำหลักการไปประยุกต์ใช้ในการวิจัย

เรามาดูตัวอย่างในการศึกษาเกี่ยวกับเด็กเกรด 6 และเราต้องการวัดคุณลักษณะ 3 คุณลักษณะ คือ คุณลักษณะความภาคภูมิใจในตนเอง (SE) คุณลักษณะการเปิดเผยตน (SD) และการควบคุมตนเอง (LC) ยิ่งกว่านั้น เราจะวัดคุณลักษณะทั้ง 3 ด้วยวิธีการวัด 3 วิธี คือ การประเมินตนเองของเด็กโดยการใช้การเขียนตอบ (P&P) การประเมินโดยครู (Teacher) และการประเมินโดยครอบครัว (Parent) ผลที่ได้จะแสดงใน MTMM ต่อไปนี้เป็นหลักการนำเสนอผลของสัมประสิทธิ์ที่เหมาะสมใน MTMM และผู้วิจัยจะต้องทำการตัดสินใจด้วยตนเอง

| | Traits | SE ₁ | SD ₁ | LC ₁ | SE ₂ | SD ₂ | LC ₂ | SE ₃ | SD ₃ | LC ₃ |
|---------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| P&P | SE ₁ | (.89) | | | | | | | | |
| | SD ₁ | .51 | (.89) | | | | | | | |
| | LC ₁ | .38 | .37 | (.76) | | | | | | |
| Teacher | SE ₂ | .57 | .22 | .09 | (.93) | | | | | |
| | SD ₂ | .22 | .57 | .10 | .68 | (.94) | | | | |
| | LC ₂ | .11 | .11 | .46 | .59 | .58 | (.84) | | | |
| Parent | SE ₃ | .56 | .22 | .11 | .67 | .42 | .33 | (.94) | | |
| | SD ₃ | .23 | .58 | .12 | .43 | .66 | .34 | .67 | (.92) | |
| | LC ₃ | .11 | .11 | .45 | .34 | .32 | .58 | .58 | .60 | (.85) |

รูปภาพ 5

หลักการพื้นฐานหรือกฎของ MTMM คือ

- 1) สัมประสิทธิ์ในไดอะกอนอลที่แสดงค่าความเชื่อมั่นจะต้องมีค่าสูงในเมตริกซ์
- 2) สัมประสิทธิ์ในไดอะกอนอลที่แสดงความเที่ยงตรงควรจะมีนัยสำคัญแตกต่างจากศูนย์และมีค่าสูงเพียงพอ นั่นคือเป็นหลักฐานที่แสดงความเที่ยงตรงเชิงเหมือน สหสัมพันธ์ทั้งหมดในตัวของเรา (รูปภาพ 5) เป็นไปตามเกณฑ์นี้

3) สัมประสิทธิ์ในไดอะกอนอลที่แสดงความเที่ยงตรงควรมีค่าสูงกว่าสัมประสิทธิ์ที่แสดงคุณลักษณะต่างกันและวิธีการต่างกัน ในตัวอย่างนี้ (SE P&P)-(SE Teacher) ควรจะสูงกว่า (SE P&P)-(SD Teacher), (SE P&P)-(LC Teacher), (SE Teacher)-(SD P&P) และ (SE Teacher)-(LC P&P) นั่นคือเป็นจริงในทุก ๆ กรณีสำหรับตัวอย่างนี้

4. สัมประสิทธิ์ความเที่ยงตรงควรมีสูงกว่าทุก ๆ สัมประสิทธิ์ในสามเหลี่ยมของคุณลักษณะต่างกันแต่วิธีการเดียวกัน นั่นคือเน้นว่าองค์ประกอบคุณลักษณะจะต้องมีความแข็งแกร่งมากกว่าองค์ประกอบวิธีการ สังเกตว่าไม่เป็นจริงในทุก ๆ กรณีในตัวอย่างของเรา ตัวอย่างเช่น (LC P&P)-(LC Teacher) มีสหสัมพันธ์เป็น 0.46 น้อยกว่า (SE Teacher)-(SD Teacher), (SE Teacher)-(LC Teacher) และ (SD Teacher)-(LC Teacher) จะเห็นว่าไม่เป็นไปตามเกณฑ์เป็นเพราะองค์ประกอบวิธีการโดยเฉพาะการประเมินของครู อาจเป็นไปได้ว่าการประเมินของครูเป็นวิธีการวัดที่มีประสิทธิภาพมาก

5. ความสัมพันธ์ระหว่างคุณลักษณะจะมีรูปแบบคล้าย ๆ กัน ดูในสามเหลี่ยมทั้งหมดในตัวอย่างของเราจะเห็นความชัดเจนในเกณฑ์ที่กล่าวมานี้ สังเกตว่าในสามเหลี่ยมทั้งหมด SE - SD มีความสัมพันธ์กันสูงมากกว่าความสัมพันธ์กับ LC

ข้อเด่นและข้อด้อยของ MTMM

แนวคิดของ MTMM เป็นวิธีการประเมินความเที่ยงตรงเชิงโครงสร้าง ในเมตริกซ์หนึ่ง ๆ มันเป็นไปได้ที่เราจะตรวจสอบทั้งความเที่ยงตรงเชิงเหมือนและเชิงจำแนกพร้อมกัน

แม้จะมีประโยชน์ MTMM ก็ยังมีผู้นำไปใช้น้อย มีเหตุผลด้วยกันหลายประการ ประการแรกคือเป็นรูปแบบที่สมบูรณ์ (purest form) MTMM จะต้องมีการวัดทั้งหมดทั้งภายในกลุ่มและข้ามกลุ่ม คือมีการวัดคุณลักษณะที่หลากหลายโดยใช้วิธีการวัดหลายแบบ ขณะที่ Campbell และ Fiske กล่าวไว้อย่างชัดเจนว่าสามารถใช้รูปแบบที่ไม่สมบูรณ์ได้ โดยจะเน้นหนักความสำคัญของรูปแบบจำลองของคุณลักษณะเดียวกันที่ใช้วิธีการแตกต่างกันในการประยุกต์ใช้กับเนื้อหาในการวิจัย มันเป็นไปได้ที่การวัดคุณลักษณะทั้งหมดจะใช้วิธีการที่แตกต่างกันทั้งหมด ในการนำไปใช้ในการวิจัยทางสังคมศาสตร์โดยมาก เป็นไปไม่ได้ที่จะใช้วิธีการที่กำหนดแน่นอนในทุก ๆ ส่วนของรูปแบบการวิจัย ประการที่สองการตัดสินใจ โดยธรรมชาติของ MTMM อาจจะถูกตัดสินทวนซ้ำใหม่จนได้ผลสรุปที่ถูกต้อง ผู้วิจัยบางคนต้องการทดสอบสำหรับความเที่ยงตรงเชิงโครงสร้างโดยให้ผลเป็นสัมประสิทธิ์ทางสถิติเพียงตัวเดียวในการทดสอบ ซึ่งจะเหมือนกับสัมประสิทธิ์สหสัมพันธ์ มันเป็นไปได้ในการใช้ MTMM ที่จะแสดงปริมาณของความเที่ยงตรงเชิงโครงสร้างด้วยค่าเพียงค่าเดียว ในที่สุดการตัดสินใจโดยธรรมชาติของ MTMM ก็คือ ผู้วิจัยที่แตกต่างกันจะมีการสรุปผลความเที่ยงตรงเชิงโครงสร้างที่แตกต่างกัน

7. การวิเคราะห์องค์ประกอบ

การวิเคราะห์องค์ประกอบจะเป็นศูนย์รวมความหลากหลายของการวิเคราะห์ความสัมพันธ์ซึ่งถูกออกแบบมาเพื่อใช้ตรวจสอบความสัมพันธ์ภายในระหว่างตัวแปรต่าง ๆ ที่สังเกตหรือวัดได้

Daniel (1988) ได้พูดถึงการวิเคราะห์องค์ประกอบไว้ว่า “การวิเคราะห์องค์ประกอบถูกออกแบบมาเพื่อใช้ตรวจสอบโครงสร้างของชุดตัวแปรและเพื่อใช้อธิบายความสัมพันธ์ระหว่างตัวแปรในรูปของจำนวนที่น้อยที่สุดของตัวแปรแฝงที่สังเกตไม่ได้ ซึ่งตัวแปรแฝงที่สังเกตไม่ได้เหล่านี้จะถูกเรียกว่า องค์ประกอบ”

Joreskog และ Sorbom (1989) ได้อธิบายว่า “แนวคิดที่สำคัญภายใต้รูปแบบของการวิเคราะห์องค์ประกอบ คือ มีตัวแปรบางตัวที่ไม่สามารถสังเกตหรือวัดได้โดยตรง หรืออาจเรียกว่าเป็นตัวแปรแฝงหรือองค์ประกอบ ตัวแปรที่ไม่สามารถสังเกตหรือวัดได้โดยตรงนั้น สามารถอ้างอิงได้ทางอ้อมจากข้อมูลของตัวแปรที่สังเกตได้ การวิเคราะห์องค์ประกอบเป็นกระบวนการทางสถิติสำหรับเปิดเผย (uncooering) ตัวแปรแฝงที่มีอยู่ โดยศึกษาผ่านความแปรปรวนระหว่างชุดของตัวแปรที่สังเกตได้”

กระบวนการวิเคราะห์องค์ประกอบถือกำเนิดขึ้นมาในช่วงต้นศตวรรษที่ 20 โดย Spearman (1904) แต่การวิเคราะห์องค์ประกอบในสมัยนั้นยังเป็นวิธีการที่ยุ่งยาก ซับซ้อนและเสียเวลามากในการวิเคราะห์ ดังนั้น การวิเคราะห์องค์ประกอบจึงยังไม่เป็นที่แพร่หลายในหมู่นักวิจัยสมัยนั้น จนกระทั่งคอมพิวเตอร์ได้ถือกำเนิดขึ้นมาและตามมาด้วยโปรแกรมคอมพิวเตอร์ที่จะช่วยเหลือในการวิเคราะห์องค์ประกอบ ดังนั้นการวิเคราะห์องค์ประกอบจึงได้แพร่หลายออกไปในหมู่นักวิจัยกันอย่างกว้างขวาง

Kerlinger (1986) ได้กล่าวถึงประโยชน์ของการวิเคราะห์องค์ประกอบไว้ว่า “เป็นเครื่องมืออย่างหนึ่งที่มีประโยชน์มาก ถูกสร้างขึ้นมาเพื่อใช้ศึกษาปัญหาที่ซับซ้อนในศาสตร์ทางพฤติกรรม”

จุดมุ่งหมายในการวิเคราะห์องค์ประกอบมี 2 ประการคือ

1. เพื่อสำรวจหรือค้นหาตัวแปรแฝงที่ซ่อนอยู่ภายใต้ตัวแปรที่สังเกตหรือวัดได้ เรียกว่า การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory factor analysis)
2. เพื่อพิสูจน์ ตรวจสอบหรือยืนยันทฤษฎีที่ผู้อื่นค้นพบ เรียกว่า การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory factor analysis)

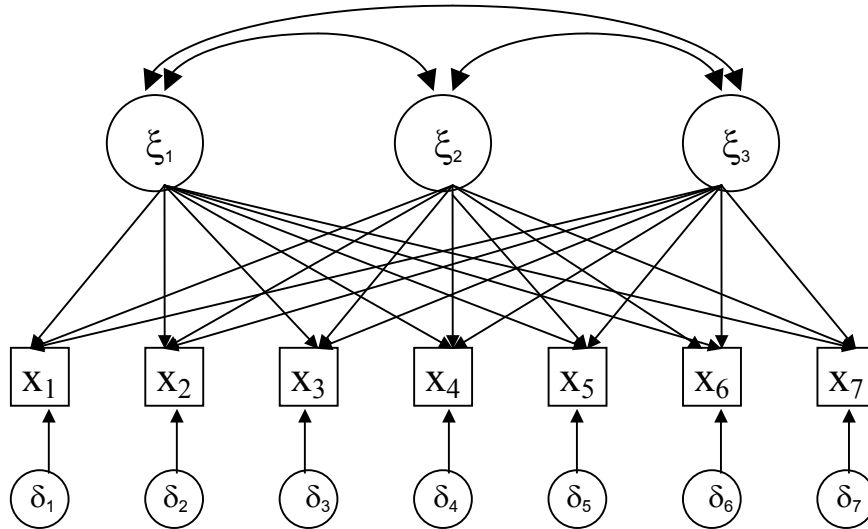
การวิเคราะห์องค์ประกอบเชิงสำรวจ

การวิเคราะห์องค์ประกอบเชิงสำรวจจะใช้ในการสำรวจข้อมูล กำหนดจำนวนองค์ประกอบ อธิบายความแปรปรวนร่วมระหว่างตัวแปรเมื่อผู้วิจัยไม่มีหลักฐานอ้างอิงเพียงพอสำหรับเป็นกรอบของสมมติฐานเกี่ยวกับจำนวนขององค์ประกอบภายใต้ข้อมูลที่สอบวัดได้

การวิเคราะห์องค์ประกอบเชิงสำรวจสามารถใช้ในการตอบคำถามที่เกี่ยวกับความเที่ยงตรงเชิงโครงสร้าง เช่น “แบบทดสอบที่นำไปสอบเก็บคะแนนมานี้ วัดอะไรบ้าง?”

รูปภาพ 9 แสดงให้เห็นรูปแบบการวิเคราะห์องค์ประกอบเชิงสำรวจ ในรูปนี้ ตัวแปรใน

สี่เหลี่ยมคือตัวแปรที่สังเกตได้ และตัวแปรในวงกลมคือตัวแปรแฝงหรือองค์ประกอบ ลูกศรตรงชี้จากตัวแปรแฝงไปยังตัวแปรที่สังเกตได้ แสดงให้เห็นความเป็นสาเหตุของตัวแปรแฝงที่ส่งผลต่อตัวแปรที่สังเกตได้ เส้นโค้งระหว่างตัวแปรแฝง 2 ตัว แสดงให้เห็นถึงความสัมพันธ์ระหว่างตัวแปรแฝง



รูปภาพ 9 โมเดลการวิเคราะห์องค์ประกอบเชิงสำรวจ

วงกลมด้านบนในรูปภาพ 9 แสดงตัวแปรแฝง ξ_1 , ξ_2 และ ξ_3 (ξ อ่านว่า ไซด์-xi) เส้นโค้งระหว่างตัวแปรแฝงแสดงให้เห็นความสัมพันธ์ระหว่างตัวแปรแฝงแต่ละตัว ตัวแปรแฝงแต่ละตัวเป็นสาเหตุที่ส่งผลไปยังตัวแปรที่สังเกตได้แต่ละตัว ซึ่งอยู่ในกล่องสี่เหลี่ยมมีชื่อว่า $x_1 - x_7$ เป็นลูกศรชี้จาก ξ ถึง x ตัวแปรแฝงหรือองค์ประกอบ ξ เรียกว่า common factors ในรูปภาพนี้ common factor จะมีผลโดยตรงต่อตัวแปรที่สังเกตได้มากกว่า 1 ตัว วงกลมในด้านล่างของรูปมีชื่อว่า $\delta_1 - \delta_7$ (δ อ่านว่า เดลต้า-delta) นั้นเรียกว่า unique factors หรือตัวแปรความคลาดเคลื่อน unique factors 1 ตัวจะมีผลกระทบต่อตัวแปรที่สังเกตได้เพียงตัวเดียว ในรูปแบบการวิเคราะห์องค์ประกอบเชิงสำรวจ จะสมมติ unique factors ว่าไม่มีความสัมพันธ์ระหว่างกัน และไม่มีความสัมพันธ์กับ common factors ดังจะเห็นได้ว่าไม่มีเส้นโค้งระหว่าง unique factors ด้วยกันและไม่มีเส้นโค้งระหว่าง unique factor กับ common factors ดังในรูปภาพ 9

รูปแบบที่นำเสนอในรูปภาพ 9 เป็นโมเดลองค์ประกอบเชิงสำรวจ มีข้อตกลงเบื้องต้นว่า

- 1) common factors ทุกตัวมีความสัมพันธ์กัน
- 2) ตัวแปรที่สังเกตได้ทั้งหมดจะต้องเป็นผลโดยตรงจาก common factors ทุกตัว
- 3) unique factors แต่ละตัวจะต้องไม่มีความสัมพันธ์กัน
- 4) ตัวแปรที่สังเกตได้ทุกตัวจะต้องเป็นผลโดยตรงจาก unique factors
- 5) common factors ทั้งหมดไม่สัมพันธ์กับ unique factors

กระบวนการวิเคราะห์องค์ประกอบเชิงสำรวจและการแปลความหมาย

ขั้นตอนในการวิเคราะห์องค์ประกอบเชิงสำรวจอาจแบ่งได้คร่าว ๆ 5 ขั้นตอนคือ

1. เก็บข้อมูลและสร้างเมตริกซ์สหสัมพันธ์
2. การสกัดองค์ประกอบ
3. เลือกวิธีการหมุนแกน (Orthogonal หรือ oblique)
4. เลือกค่า loading
5. ตั้งชื่อองค์ประกอบที่วิเคราะห์ได้

1. เก็บข้อมูลและสร้างเมตริกซ์สหสัมพันธ์

อันดับแรกในขั้นตอนของการวิเคราะห์องค์ประกอบคือการเก็บรวบรวมข้อมูลและนำข้อมูลที่ได้มาหาความสัมพันธ์ระหว่างตัวแปรต่าง ๆ ที่ต้องการวิเคราะห์ และนำเสนอในรูปของเมตริกซ์สหสัมพันธ์ ดังตัวอย่างตาราง 1

ตาราง 1 ตัวอย่างเมตริกซ์สหสัมพันธ์

| | | | | | | |
|--|------|------|------|------|------|------|
| 1. Ability to define problems | 1.00 | | | | | |
| 2. Ability to supervise others | -.34 | 1.00 | | | | |
| 3. Ability to make decisions | .51 | -.48 | 1.00 | | | |
| 4. Ability to build consensus | .05 | .27 | -.11 | 1.00 | | |
| 5. Ability to facilitate decision-making | .07 | .18 | -.03 | .78 | 1.00 | |
| 6. Ability to work on a team | -.48 | .46 | -.44 | .19 | .17 | 1.00 |

2. การสกัดองค์ประกอบ

ขั้นตอนที่สองในการวิเคราะห์องค์ประกอบคือการค้นหาจำนวนองค์ประกอบที่มีความสามารถเพียงพอในการอธิบายความสัมพันธ์ระหว่างตัวแปรที่สังเกตได้ ซึ่งมีวิธีการต่าง ๆ ให้เลือกใช้ดังนี้ 1) Maximum Likelihood Method (หรือ Canonical Factoring) 2) Least-Squares Method (หรือ Principal Axis Factoring) 3) Alpha Factoring 4) Image Factoring และ 5) Principal Components Analysis ผู้วิจัยจะต้องเลือกใช้วิธีใดวิธีหนึ่งหรือให้โปรแกรมคอมพิวเตอร์เลือกให้ ถ้าเป็นโปรแกรม SPSS โปรแกรมจะเลือกวิธี Principal Component Analysis

ผลลัพธ์ของการวิเคราะห์องค์ประกอบจะให้หลายสิ่งหลายอย่าง ตารางข้างล่างนี้จะนำเสนอผลลัพธ์ที่ได้จากการวิเคราะห์องค์ประกอบ จะช่วยในการตัดสินใจเกี่ยวกับจำนวนขององค์ประกอบเพื่อเก็บไว้สำหรับใช้ในการวิเคราะห์ต่อไปในอนาคต กฎที่ดีที่สุดสำหรับการกำหนดจำนวนขององค์ประกอบคือ “eigenvalue > 1” ค่า Eigenvalue เป็นค่าที่บ่งบอกถึงความสามารถขององค์ประกอบว่าจะอธิบายความแปรปรวนของกลุ่มตัวแปรได้มากน้อยเพียงไร โดยปกติถ้าองค์ประกอบนั้นอธิบายความแปรปรวนของกลุ่มตัวอย่างได้น้อยกว่า 1 Eigenvalue แล้วก็ไม่มีประโยชน์ที่จะนำองค์ประกอบนั้น

มาใช้ หากตัวแปรที่นำมาวิเคราะห์มีจำนวนน้อย การวิเคราะห์อาจจะให้ผลเป็นองค์ประกอบแค่ 2 - 3 องค์ประกอบเท่านั้น ถ้าหากตัวแปรที่นำมาวิเคราะห์มีจำนวนมากอาจจะได้จำนวนองค์ประกอบมาก แต่เราอาจจะกำหนดเกณฑ์อื่น ๆ สำหรับเลือกจำนวนองค์ประกอบได้ แต่ eigenvalue > 1 นี้เป็นเกณฑ์ที่ถูกกำหนดไว้ในโปรแกรมคอมพิวเตอร์ทุกโปรแกรม

สมมติฐานของตัวอย่างนี้คือ “เราคาดหวังว่าจะได้ 2 องค์ประกอบคือ “representing task” และ “people skills” ผลลัพธ์ในตาราง 2 จะเป็นผลจากการสกัดองค์ประกอบ ซึ่งจะแสดงให้เห็นดังนี้

ตาราง 2 ตัวอย่างผลลัพธ์ที่ได้จากการสกัดองค์ประกอบ

| Factor | Eigenvalues | % of variance | Cumulative % of variance |
|--------|-------------|---------------|--------------------------|
| 1 | 2.51349 | 41.9 | 41.9 |
| 2 | 1.73952 | 29.0 | 70.9 |
| 3 | .59749 | 10.0 | 80.8 |
| 4 | .52956 | 8.8 | 89.7 |
| 5 | .41573 | 6.9 | 96.6 |
| 6 | .20422 | 3.4 | 100.0 |

การแปลความหมายผลลัพธ์ที่วิเคราะห์ได้

ในเริ่มแรกได้ 2 องค์ประกอบที่มีค่า eigenvalue > 1 รวมเปอร์เซ็นต์ของความแปรปรวนทั้ง 2 องค์ประกอบจะได้ 70.9% นั่นคือทั้ง 2 องค์ประกอบนี้สามารถอธิบายความแปรปรวนของกลุ่มตัวอย่างได้ 70.9%

ในตาราง 3 ค่า loading จะนำเสนอภายใต้หัวข้อ “Factor” เป็นค่าที่แสดงความสัมพันธ์ระหว่างตัวแปรแต่ละตัวกับองค์ประกอบ รูปร่างคล้ายความสัมพันธ์แบบเพียร์สัน ค่า loading จะมีพิสัยระหว่าง -1 ถึง 1 ส่วนสตรมภ์ถัดมาเป็นค่า communality ซึ่งก็คือเปอร์เซ็นต์ความแปรปรวนของตัวแปรแต่ละตัวกับองค์ประกอบ

ตาราง 3 เมตริกซ์องค์ประกอบก่อนหมุนแกน

| Variables | Factor 1 | Factor 2 | Communality |
|--|----------|----------|-------------|
| 1. Ability to define problems | -.627 | .514 | .66 |
| 2. Ability to supervise others | .759 | -.068 | .58 |
| 3. Ability to make decisions | -.730 | .337 | .65 |
| 4. Ability to build consensus | .494 | .798 | .88 |
| 5. Ability to facilitate decision-making | .425 | .832 | .87 |
| 6. Ability to work on a team | .767 | -.168 | .62 |

3. เลือกวิธีการหมุนแกน

ตารางนี้แสดงให้เห็นความยากลำบากในการแปลความหมายขององค์ประกอบก่อนหมุนแกน ตัวแปรทั้งหมดจะมีค่า loading สูงมาก (สังเกตใน Factor 1) มีทางเดียวที่จะสามารถแปลผลได้นั้นคือ จะต้องดำเนินการหมุนแกน (rotate) ซึ่งมี 2 วิธีคือ วิธี **Orthogonal** ใช้ในกรณีที่ตัวแปรแต่ละตัวเป็นอิสระจากกัน ซึ่งการหมุนแกนวิธีนี้มีให้เลือก 3 แบบคือ 1) varimax 2) equamax และ 3) quartimax และวิธี **Oblique** แบบ oblimin ใช้ในกรณีที่ตัวแปรแต่ละตัวสัมพันธ์กัน โปรแกรมคอมพิวเตอร์ส่วนมากจะเลือกการหมุนแกนด้วยวิธี Orthogonal แบบ varimax ทั้งที่ในความเป็นจริงก็ยังมีเทคนิควิธีการหมุนแกนแบบอื่น ๆ ให้เลือกใช้

ข้างล่างนี้เป็นตัวอย่างขององค์ประกอบที่ผ่านหมุนแกนแล้ว สังเกตค่า loading จะมีการกระจายระหว่างองค์ประกอบ จะทำให้ง่ายต่อการแปลผล

ตาราง 4 เมตริกขององค์ประกอบหลังการหมุนแกน

| Variables | Factor 1 | Factor 2 | Communality |
|--|----------|----------|-------------|
| 1. Ability to define problems | -.787 | .194 | .66 |
| 2. Ability to supervise others | .724 | .266 | .58 |
| 3. Ability to make decisions | -.804 | -.011 | .65 |
| 4. Ability to build consensus | .102 | .933 | .88 |
| 5. Ability to facilitate decision-making | .025 | .934 | .87 |
| 6. Ability to work on a team | .764 | .179 | .62 |

4. เลือกค่า loading

เพื่อจะได้ทราบว่าตัวแปรใดบรรจุอยู่ในองค์ประกอบใดให้พิจารณาที่ค่า loading โดยปกติในงานวิจัยส่วนใหญ่จะใช้เกณฑ์ที่ .3 - .4 เพราะในงานวิจัยนั้นมักจะใช้กลุ่มตัวอย่างมีจำนวนมาก Hair (1995 : 385) ได้เสนอตารางแสดงความสัมพันธ์ระหว่างค่า loading ที่มีนัยสำคัญทางสถิติที่ระดับ .05 ต่อจำนวนกลุ่มตัวอย่าง แสดงในตาราง 5

ตาราง 5 ค่า loading ที่มีนัยสำคัญทางสถิติที่ระดับ .05 ต่อจำนวนกลุ่มตัวอย่าง

| | | | | | | | | | | |
|--------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Factor loading | .30 | .35 | .40 | .45 | .50 | .55 | .60 | .65 | .70 | .75 |
| จำนวนกลุ่มตัวอย่าง | 350 | 250 | 200 | 150 | 120 | 100 | 85 | 70 | 60 | 50 |

5. การตั้งชื่อองค์ประกอบ

เมื่อเรารู้ความหมายของ loading แล้ว ถัดมาคือตั้งชื่อให้แต่ละองค์ประกอบ มีกฎในการตั้งชื่อนี้

ชื่อขององค์ประกอบควรจะ

- สั้น อาจตั้งชื่อเพียง 1 - 2 คำ
- มีความหมายสอดคล้องกับโครงสร้างขององค์ประกอบ

โดยพิจารณาความคล้ายคลึงกันระหว่างตัวแปรที่อยู่ในองค์ประกอบ ถ้าผู้วิจัยค้นคว้าตามโครงสร้างของทฤษฎี ผู้วิจัยอาจจะต้องการใช้ชื่อองค์ประกอบตามทฤษฎีที่ได้ค้นคว้ามา หรือผู้วิจัยอาจตั้งชื่อใหม่ที่สอดคล้องกับแนวความคิดของผู้วิจัยเอง

การวิเคราะห์องค์ประกอบเชิงสำรวจจะช่วยนักวิจัยอย่างมากในเรื่องของการประเมินธรรมชาติของความสัมพันธ์ระหว่างตัวแปรและค้นหาความเที่ยงตรงเชิงโครงสร้างของคะแนนสอบ ยังมีนักวิจัยอื่น ๆ ที่กล่าวถึงในด้านบทพร้อม เป็นต้นว่า ข้อตกลงเบื้องต้นของรูปแบบการวิเคราะห์องค์ประกอบเชิงสำรวจนี้ ไม่ได้คำนึงถึงความเป็นจริงว่าข้อมูลที่ได้นั้นจะเป็นไปตามข้อตกลงเบื้องต้นหรือไม่ ถ้าข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้นเมื่อทำการวิเคราะห์องค์ประกอบแล้วอาจจะเรียกได้อีกชื่อหนึ่งว่า รูปแบบกิโก้ (GIGO model) ซึ่งก็คือสิ่งที่ไม่มีค่า ไม่มีประโยชน์ต่องานวิจัย

การวิเคราะห์องค์ประกอบโดยทั่วไปแล้วจะมีลักษณะเป็นเชิงเส้นตรง อาจมีเหตุผลบางอย่างที่ความสัมพันธ์ไม่เป็นเส้นตรง แต่เมื่อดำเนินการวิเคราะห์องค์ประกอบจนได้ผลลัพธ์ออกมาแล้วผลลัพธ์ที่ได้ก็จะไม่ถูกต้อง

นอกจากนี้ยังมีปัญหาในเรื่องของการแปลความหมายผลการวิเคราะห์องค์ประกอบที่ถูกวัดโดยตัวแปรบางตัว มักก่อความยุ่งยากในการแปลความหมายอยู่เสมอ ๆ ทั้งนี้มีสาเหตุมาจากผู้วิจัยไม่มีความรู้ ดังนั้นจึงไม่มีพื้นฐานในการแปลผลลัพธ์ที่ได้

ยังมีอีกปัญหาหนึ่งที่เกิดขึ้นบ่อยเกี่ยวกับการวิเคราะห์องค์ประกอบเชิงสำรวจ นั่นก็คือ การวิเคราะห์องค์ประกอบเชิงสำรวจไม่ให้ผลลัพธ์ที่ใช้ในการแก้ปัญหาใด ๆ เกี่ยวกับองค์ประกอบที่ได้ หรือยิ่งกว่านั้นผลลัพธ์ที่ได้ก็ไม่สามารถแปลความหมายได้ ทำให้เกิดความลำบากในการหาเหตุผลมาอธิบายผลลัพธ์ที่ได้

การวิเคราะห์องค์ประกอบเชิงยืนยัน

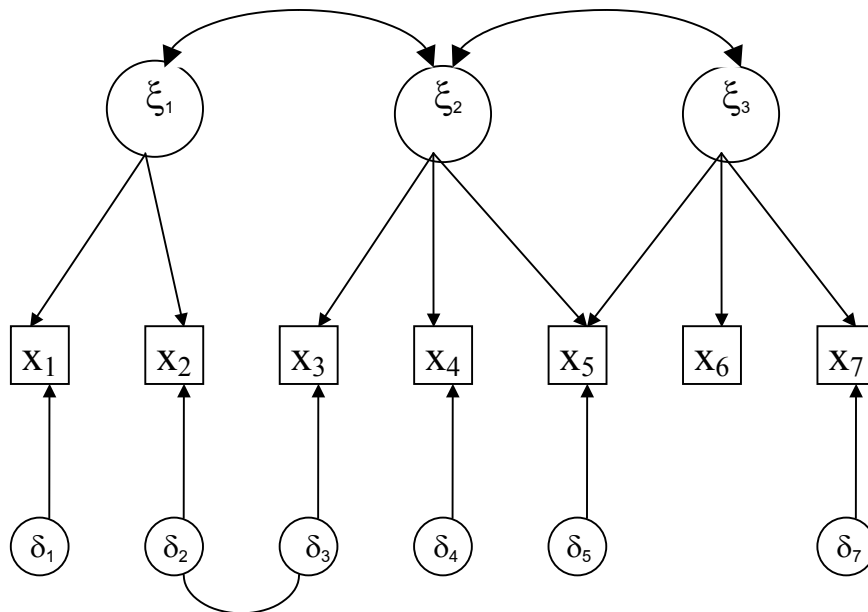
เนื่องจากข้อจำกัดของรูปแบบองค์ประกอบเชิงสำรวจมีมากมาย จึงพัฒนามาเป็นรูปแบบองค์ประกอบเชิงยืนยัน ในรูปแบบองค์ประกอบเชิงยืนยัน มีข้อตกลงเบื้องต้นที่เปลี่ยนแปลงไปคือ

- 1) อาจมีคู่ของ common factors ที่สัมพันธ์กัน
- 2) ตัวแปรที่สังเกตได้จะต้องเป็นผลโดยตรงจาก common factors
- 3) ตัวแปรที่สังเกตได้จะต้องเป็นผลโดยตรงจาก unique factors
- 4) คู่ของ unique factors สามารถสัมพันธ์กันได้

ในการวิเคราะห์องค์ประกอบเชิงยืนยัน ผู้วิจัยมักเริ่มต้นที่สมมติฐานในการวิเคราะห์เป็นการศึกษาตัวแปรที่ถูกนำไปสัมพันธ์กับองค์ประกอบและองค์ประกอบก็ถูกสัมพันธ์กันเอง สมมติฐานที่ตั้งจะต้องอยู่บนพื้นฐานของทฤษฎี

ความชัดเจนระหว่างรูปแบบองค์ประกอบเชิงยืนยันและเชิงสำรวจ สามารถเห็นได้โดย

เปรียบเทียบกับรูปแบบเชิงสำรวจในรูปภาพ 9 และรูปแบบเชิงยืนยันในรูปภาพ 10 ในรูปแบบเชิงยืนยัน common factors ξ_1 และ ξ_3 สมมติว่าไม่สัมพันธ์กัน ในรูปแบบเชิงสำรวจ common factors ทั้งหมดจำเป็นต้องสมมติให้สัมพันธ์กันทุกตัว ในรูปแบบองค์ประกอบเชิงยืนยัน ตัวแปรที่สังเกตได้เป็นผลโดยตรงจาก common factors บางตัว (เช่น x_1 สมมติว่าไม่มีผลโดยตรงจาก ξ_2 และ ξ_3) ในรูปแบบเชิงสำรวจ ตัวแปรที่สังเกตได้ทั้งหมดเป็นผลโดยตรงจาก common factors ทั้งหมด ในตัวอย่างรูปแบบองค์ประกอบเชิงยืนยันอาจมี unique factors 2 ตัวที่สมมติให้สัมพันธ์กัน (δ_2 และ δ_3 สัมพันธ์กันโดยมีเส้นโค้งเชื่อมกัน) และอาจมีตัวแปรที่สังเกตได้ 1 ตัวที่ไม่มีความคลาดเคลื่อนมาสัมพันธ์ (x_6 ไม่มี unique factors มาสัมพันธ์) ซึ่งในรูปแบบเชิงสำรวจนั้น unique factors ทุกตัวไม่มีความสัมพันธ์กัน และ unique factors มีความสัมพันธ์กับตัวแปรที่สังเกตได้แต่ละตัว



รูปภาพ 10 โมเดลการวิเคราะห์ห้วงค์ประกอบเชิงยืนยัน

การวิเคราะห์ห้วงค์ประกอบเชิงยืนยันยังมีวิธีให้ผู้วิจัยดำเนินการต่อไปสำหรับประเมินความเที่ยงตรงเชิงโครงสร้าง

กระบวนการวิเคราะห์ห้วงค์ประกอบเชิงยืนยัน

ในขั้นตอนแรกของการวิเคราะห์ห้วงค์ประกอบเชิงยืนยันต้องเริ่มต้นที่เมตริกซ์สหสัมพันธ์หรือเมตริกซ์ความแปรปรวน หรือเมตริกซ์ความแปรปรวนร่วม หรือเมตริกซ์อื่น ๆ ที่มีลักษณะคล้ายกัน

ผู้วิจัยกำหนดจุดประสงค์ในการเปรียบเทียบโมเดล (model) ซึ่งจะต้องตั้งอยู่บนพื้นฐานของทฤษฎีหรือข้อมูลที่มีอยู่ สมมติฐานจะต้องตั้งให้เหมาะสมกับข้อมูล รูปแบบจะต้องกำหนดระดับของความสัมพันธ์ระหว่างแต่ละคู่ของตัวแปรแต่ละตัวกับองค์ประกอบ 1 ตัวหรือมากกว่า และกำหนดคู่ของตัวแปรความคลาดเคลื่อนให้สัมพันธ์กัน

ความแตกต่างของพารามิเตอร์คงที่และอิสระ เช่น สัมประสิทธิ์องค์ประกอบ สัมประสิทธิ์สหสัมพันธ์องค์ประกอบ ความแปรปรวนและความแปรปรวนร่วมของความคลาดเคลื่อนที่ได้จากการวัด เป็นต้น ค่าพารามิเตอร์เหล่านี้จะกำหนดขึ้นตามความคาดหวังในทฤษฎีที่ผู้วิจัยค้นคว้ามา Gillaspay (1996) ได้ให้คำจำกัดความของพารามิเตอร์คงที่และพารามิเตอร์อิสระไว้ว่า “การคงที่ของพารามิเตอร์จะเกี่ยวข้องกับการตั้งค่าพารามิเตอร์ตามทฤษฎีที่ได้คาดหวังไว้ ดังนั้นในการคงที่ของพารามิเตอร์ ผู้วิจัยจะต้องไม่ยอมให้พารามิเตอร์นั้นแปรเปลี่ยนไปขณะที่ทำการวิเคราะห์ และการเป็นอิสระของพารามิเตอร์จะเกี่ยวข้องกับการที่ผู้วิจัยยอมให้พารามิเตอร์ถูกประมาณค่าในขณะที่กำลังทำการวิเคราะห์”

ในการวิเคราะห์องค์ประกอบเชิงยืนยันสามารถจะกระทำได้โดยใช้โปรแกรมคอมพิวเตอร์ที่หลากหลายเช่น LISREL 8.30, AMOS เป็นต้น

การเปรียบเทียบรูปแบบจะถูกทดสอบด้วยโปรแกรมคอมพิวเตอร์ ในเรื่องของความสมบูรณ์ในการวิเคราะห์นั้น ผลลัพธ์ที่ได้จะต้องแสดงค่าสถิติที่แตกต่างกันหลากหลายค่า สำหรับใช้ในการเปรียบเทียบความเหมาะสมของรูปแบบหรืออธิบายความแปรปรวนร่วมระหว่างตัวแปร ค่าสถิติจะถูกนำเสนอในรูปของ “fit statistics” ซึ่งค่า fit statistics ทั้งหมดจะถูกประมวลและแสดงออกมาในคราวเดียวกัน

สถิติเหล่านี้จะถูกใช้ประเมินโมเดล (model) เชิงประจักษ์กับโมเดลตามทฤษฎี และใช้อธิบายความสัมพันธ์ระหว่างตัวแปรที่สังเกตได้กับตัวแปรแฝง

องค์ประกอบที่มีอิทธิพลต่อความเที่ยงตรง

Gulliksen (1950) ได้อธิบายองค์ประกอบที่มีอิทธิพลต่อความเที่ยงตรงดังนี้

1. อิทธิพลของความยาวของแบบทดสอบ (Effect of Test Length on Validity)

แบบทดสอบที่มีจำนวนข้อสอบมากย่อมมีความเที่ยงตรงสูงกว่าแบบทดสอบที่มีจำนวนข้อสอบน้อย

2. ความเป็นวิวิธพันธ์ของกลุ่มผู้สอบ (Group Heterogeneity)

ถ้านำแบบทดสอบไปใช้กับกลุ่มตัวอย่างที่มีความหลากหลายกันมากค่าความเที่ยงตรงจะมีค่าสูงกว่าเมื่อนำแบบทดสอบไปใช้กับกลุ่มตัวอย่างที่เป็นเอกพันธ์



บรรณานุกรม

- ฉัตรศิริ ปิยะพิมลสิทธิ์. “การวิเคราะห์องค์ประกอบ,” ใน *วารสารการวัดผลการศึกษา*. สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ. 20 (58) : พฤษภาคม-สิงหาคม, 2541.
- _____. “การวิเคราะห์องค์ประกอบเชิงยืนยัน,” ใน *วารสารการวัดผลการศึกษา*. สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ. 20 (59) : พฤษภาคม-สิงหาคม, 2541.
- ทวีศักดิ์ ญาณประทีป และคณะ. *พจนานุกรม ฉบับเฉลิมพระเกียรติ พ.ศ.2530*. พิมพ์ครั้งที่ 11. กรุงเทพฯ : วัฒนาพานิช, 2534.
- ล้วน และอังคณา สายยศ. *เทคนิคการวัดผลการเรียนรู้*. กรุงเทพฯ : สุวีริยาสาส์น, 2539.
- Allen, Marry J. and Yen, Wendy M. *Introduction to Measurement Theory*. U.S.A. : Brooks/Cole Publishing Company, 1979.
- Anastasi, Anne. *Psychological Testing*. U.S.A. : Macmillan Publishing Company, 1982.
- Gulliksen, Harold. *Theory of MENTAL TESTS*. U.S.A. : John Willy & Sons, Inc., 1950.
- Lyman, Howard B. *Test Scores and What They Mean*. U.S.A. : Prentice-Hall, INC., 1963.
- Messick, Samuel. “Validity,” in *Educational Measurement*. Linn, Robert L. (Ed.). Third Edition. U.S.A. : Macmillan Publishing Company, 1989.
- Popham, W. James. *Modern Educational Measurement : A Practitioner’s Perspective*. U.S.A. : Prentice-Hall, INC., 1990.
- Trochim, William M.K. *Research Methods Knowledge Base*. 2nd Edition. <http://trochim.human.cornell.edu/kb/>. 1999.
- Wainer, Howard and Braun, Henry I. *Test Validity*. U.S.A. Lawrence Erlbaum Associates, Inc., 1988.