

การอธิบายโครงสร้างและความเที่ยงตรงของแบบทดสอบ

ชัยลิขิต สร้อยเพชรเกษม *

Construct Explication and Validity of the Tests

Chailikit Soipetkasem

* อาจารย์ประจำภาควิชาการประเมินผลและวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยทักษิณ

บทคัดย่อ

การออกแบบเนื้อหาสาระที่จะวัดของแบบทดสอบต่าง ๆ นั้นสามารถกำหนดได้มากมายแตกต่างกันไป แบบทดสอบบางประเภทออกแบบมาเพื่อวัดเนื้อหาสาระที่เป็นความรู้หรือทักษะ และบางประเภทออกแบบมาเพื่อใช้วัดเนื้อหาสาระที่เป็นโครงสร้างทางจิตวิทยา แบบทดสอบที่ใช้วัดบุคลิกภาพนั้นมุ่งวัดเนื้อหาสาระที่เป็นโครงสร้างทางจิตวิทยา และการวัดดังกล่าวไม่สามารถวัดได้โดยตรง แต่เป็นการวัดโดยอ้อมโดยอาศัยการอธิบายโครงสร้าง ในการอธิบายโครงสร้างนี้ผู้สร้างแบบทดสอบจะต้องรวบรวมรายการต่าง ๆ ที่เป็นพฤติกรรม ความเชื่อ และเจตคติ ที่อธิบายโครงสร้างนั้นอย่างชัดเจน ตลอดจนรายการที่เป็นพฤติกรรม ความเชื่อ และเจตคติ ที่แสดงถึงความไม่สอดคล้องกับโครงสร้างนั้นอย่างชัดเจนเช่นกัน รายการต่าง ๆ เหล่านี้จะใช้กำหนดเป็นเนื้อหาของข้อความ หรือข้อความคำถามของแบบทดสอบ ในข้อเขียนต่อไปนี้จะได้อภิปรายถึงกระบวนการกำหนดเนื้อหาสาระที่เป็นโครงสร้างของแบบทดสอบ และเทคนิคที่ใช้ในการตรวจสอบความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบ

คำสำคัญ : แบบทดสอบ เนื้อหาสาระ โครงสร้างทางจิตวิทยา การอธิบายโครงสร้าง การตรวจสอบความเที่ยงตรงเชิงโครงสร้าง

Abstract

Tests can be designed to cover many types of domains. Some tests are designed to measure knowledge or skill domains and some are designed to measure construct domains. Personality tests focus on construct domains. Construct domains cannot be measure directly but indirectly by construct explication. In construct explication, the test developer compiles a lists of specific behaviors, beliefs, and

attitudes that demonstrate the presence of the construct and specific behaviors, beliefs, and attitudes inconsistent with its presence. The lists are then used to determine the content of the test items. The following topics discuss the process of specifying test content for construct domains and construct validation technique of the tests.

Keywords : Tests, Domains, Constructs, Construct Explication, Construct Validation

บทนำ

การออกแบบแบบทดสอบ (Test Design) นั้นมีความเกี่ยวข้องกับเรื่องของการกำหนดเนื้อหาสาระ (Content) ที่จะทำการวัด เนื้อหาสาระดังกล่าวคือสิ่งที่ปรากฏเป็นข้อความหรือคำถาม (Items) หรือกิจกรรมในแบบทดสอบ ซึ่งจำแนกได้ใน 3 ลักษณะ คือ 1) เนื้อหาสาระที่เป็นการวัดด้านความรู้หรือทักษะ (Knowledge or Skill Domains) 2) เนื้อหาสาระที่เป็นการวัดด้านพฤติกรรม (Behavioral Domains) และ 3) เนื้อหาสาระที่เป็นการวัดด้านโครงสร้างของตัวแปร (Construct Domains) บทความนี้มุ่งบรรยายถึงการออกแบบแบบทดสอบที่มีเนื้อหาสาระที่เป็น การวัดด้านโครงสร้างของตัวแปรที่ต้องการวัด(หรือเรียกอีกอย่างหนึ่งว่า ตัวแปรโครงสร้าง, ตัวแปรเชิงภาวะสันนิษฐาน :Construct Variables) ว่าโครงสร้างของตัวแปรนั้นกำหนดได้อย่างไรและเกี่ยวข้องกับความเที่ยงตรง (Validity) ของแบบทดสอบอย่างไร

ตัวแปรโครงสร้าง (Construct Variable)

ในการวัดทางจิตวิทยาเป็นการวัดโดยอ้อม (Indirect Measurement) ที่เกี่ยวกับตัวแปรโครงสร้างหรือตัวแปรเชิงภาวะสันนิษฐาน ซึ่งหมายถึงคุณลักษณะของตัวแปรที่สันนิษฐานว่ามีอยู่ในตัวแปรนั้น การสันนิษฐานดังกล่าวเป็นสิ่งที่ได้มาจากพื้นฐานทางแนวคิด ทฤษฎีที่เกี่ยวข้องกับตัวแปรนั้น ๆ ได้อธิบายไว้หรืออาจได้จากการสังเกตแล้วสมมติขึ้นและคาดว่าน่าจะมีลักษณะนั้น ๆ อยู่เป็นองค์ประกอบของตัวแปรอยู่จริง ตัวแปรโครงสร้างจึงเป็นตัวแปรเชิงนามธรรมที่กล่าวไว้ในทฤษฎี (Theoretical Abstraction) การวัดโดยอ้อมนี้เป็นการอ้างอิง (Inferred) ถึงพฤติกรรม ความเชื่อ และเจตคติที่มีความเฉพาะเจาะจงซึ่งบอกลักษณะที่แสดง ให้เห็นว่าพฤติกรรม ความ

เชื้อ และเจตคตินั้นเป็นสิ่งประกอบเป็นโครงสร้างของตัวแปร เช่น ความวิตกกังวลมีลักษณะเป็นอย่างไร อะไรเป็นสิ่งที่บอกหรือแสดงถึงความวิตกกังวล เป็นต้น

การอธิบายโครงสร้าง (Construct Explication)

การอธิบายโครงสร้างเป็นวิธีการที่บอก สันนิษฐาน นิยาม สมมติลักษณะหรือองค์ประกอบของตัวแปร โดยการวิเคราะห์ความเป็นไปได้ของโครงสร้างตัวแปร (Possible Analysis of the Construct "X") ซึ่งเป็นการแสดงให้เห็นถึงรายการที่เป็นคุณลักษณะเฉพาะ (Specific Attributes) และพฤติกรรมที่มีอยู่ในตัวแปรนั้น กับรายการที่เป็นคุณลักษณะเฉพาะและพฤติกรรมที่ไม่สอดคล้องหรือไม่อยู่ในตัวแปรนั้น รวมถึงการกำหนดรายการของลักษณะของตัวแปรที่เกี่ยวข้องกันกับลักษณะเฉพาะหรือพฤติกรรมที่ปรากฏในตัวแปรนั้น ลักษณะเฉพาะและพฤติกรรมดังกล่าวนี้เรียกว่า โครงสร้างของตัวแปร (Friedenberg. 1995 : 51, 253) ส่วน Nunnally (1967 : 95) ได้ให้ความหมายสั้น ๆ ว่า การอธิบายโครงสร้างเป็นกระบวนการทำให้คำที่เป็นนามธรรมมีความหมายชัดเจนในรูปแบบของพจน์ที่เป็นตัวแปรสังเกตได้ โครงสร้างของตัวแปรเป็นสิ่งที่สมมติหรือได้มาจากทฤษฎี (Theoretical Approach) กระบวนการนิยาม หรืออธิบายโครงสร้างของตัวแปรเพื่อที่จะใช้วัดตัวแปรโครงสร้างนี้มีความซับซ้อน (Complex) และมีความสำคัญ ที่ว่าซับซ้อนนั้นเพราะโครงสร้างของตัวแปรนั้นเป็นสิ่งที่ เป็นนามธรรมซึ่งกล่าวไว้ในทฤษฎี และที่ว่ามีค่าสำคัญนั้นเพราะมีบทบาทในการประเมินความเที่ยงตรงของแบบทดสอบ ไม่มีสถิติใดเป็นตัวกำหนดความเที่ยงตรงเชิงโครงสร้างได้ เช่นเดียวกับการกำหนดความเที่ยงตรงเชิงเนื้อหา (Content Validity) (Friedenberg. 1995 : 51 Citing Murphy and Davidshofer. 1988 ; 253)

เพื่อให้เกิดความชัดเจนเกี่ยวกับการกำหนดความเที่ยงตรงเชิงโครงสร้างนั้น Wiresma. (1995 : 312) ได้กล่าวถึงวิธีการกำหนดความเที่ยงตรงของแบบทดสอบหรือเครื่องมือวัดตัวแปรว่า โดยปกติแล้วมีวิธีการกำหนดความเที่ยงตรงอยู่ 2 วิธีการ ซึ่งถือว่าเป็นวิธีการพื้นฐานได้แก่ 1) การวิเคราะห์เชิงเหตุผลจากเนื้อหาที่จะวัด (Logical Analysis of Content) การวิเคราะห์นี้เป็นการบอกว่าอะไรบ้างที่ประกอบ (Make up) เป็นตัวแปร (Educational Traits, Construct and Characteristics) วิธีการนี้เป็นแก่นหรือสาระสำคัญยิ่งสำหรับการวิเคราะห์เพื่อพิจารณาตัดสินใจ (Judgment) เกี่ยวกับความเที่ยงตรง 2) การวิเคราะห์จากข้อมูลเชิงประจักษ์ (Empirical Analysis) เป็นการวิเคราะห์โดยใช้ข้อมูลจากการใช้แบบทดสอบที่ได้สร้างขึ้นนำไปสัมพันธ์กับ

เกณฑ์การวัด (Criterion Measurement) เช่น ผลการปฏิบัติงานหรือแบบทดสอบอื่น (Job Performance or Test) ความสัมพันธ์นี้ใช้สถิติเป็นตัวบอกหรือตัดสินซึ่งได้แก่ สหสัมพันธ์ (Correlation) หรือความสัมพันธ์ (Association) ก็ได้ และในการหาความเที่ยงตรงเชิงโครงสร้าง (Construct Validation) นั้นใช้ทั้ง 2 วิธีการร่วมกัน แต่การวิเคราะห์เชิงเหตุผลนั้นเป็นวิธีการที่กำหนดโครงสร้างของแบบทดสอบเพราะคำว่า “โครงสร้าง” นั้นหมายถึงโครงสร้างทางทฤษฎีหรือลักษณะที่ถูกวัด (Theoretical Construct or Traits Being Measured) ไม่ใช่โครงสร้างทางเทคนิคของข้อคำถาม (The Technical Construction of the Test Items) และโครงสร้างนั้นเป็นลักษณะที่สมมติ หรือโครงสร้างที่ใช้อธิบายปรากฏการณ์ซึ่งเป็นนามธรรมไม่มีตัวตนหรือเหตุการณ์จริง บางครั้งเรียกว่า โครงสร้างเชิงสมมติฐาน (Hypothetical Construct) (Wiresma. 1995 : 313-314) ดังนั้นในบทความนี้ผู้เขียนจึงเห็นว่าในทัศนะการกำหนดโครงสร้างนั้นได้มาจากทฤษฎี และเราสามารถตรวจสอบหรือประเมินโครงสร้างของแบบทดสอบได้จากเทคนิคสถิติ เช่น สหสัมพันธ์ เป็นต้น

การอธิบายโครงสร้างนั้นเป็นการวิเคราะห์เชิงเหตุผลและนอกจากจะทำให้ทราบโครงสร้างของตัวแปรแล้วยังมีความสำคัญต่อกระบวนการหาความเที่ยงตรงของแบบทดสอบด้วย (Process of Construct Validation) กล่าวคือ การรู้คุณลักษณะเฉพาะและพฤติกรรมที่เป็นโครงสร้างของตัวแปรจะทำให้ทราบว่าแบบทดสอบนั้นมีความเที่ยงตรง กล่าวคือถ้าคะแนน (Particular Scores) ที่ได้จากการวัดโดยแบบทดสอบมีความสัมพันธ์ทางบวก (Positive Correlation) กับคุณลักษณะเฉพาะหรือพฤติกรรมที่ปรากฏเป็นโครงสร้างของตัวแปร (Consistent with the Presence of “X”) แบบทดสอบนั้นก็มีความเที่ยงตรง และถ้าคะแนนที่ได้จากการวัดโดยแบบทดสอบมีความสัมพันธ์ทางลบ (Negative Correlation) กับคุณลักษณะเฉพาะหรือพฤติกรรมที่มีในโครงสร้างหรือไม่สอดคล้อง (Inconsistent) กับโครงสร้างของตัวแปร แบบทดสอบนั้นจะมีความเที่ยงตรง หรือถ้าคะแนนที่ได้จากการวัดโดยแบบทดสอบมีความสัมพันธ์ทางบวกกับตัวแปรอื่นที่เกี่ยวข้องสัมพันธ์กับ (Related) คุณลักษณะเฉพาะหรือพฤติกรรมที่มีอยู่ในโครงสร้างของตัวแปร แบบทดสอบนั้นจะมีความเที่ยงตรง (Friedenberg. 1995 : 253) ในประเด็นดังกล่าวมีสิ่งที่จะต้องทำความเข้าใจเพิ่มเติม 2 ประการคือ 1) คุณลักษณะเฉพาะหรือพฤติกรรมที่ปรากฏในโครงสร้างของตัวแปรจากการอธิบายโครงสร้างนั้นอาจวัดได้ด้วยวิธีการอื่นซึ่งไม่จำเป็นต้องเป็นแบบทดสอบ แต่ถ้ามีแบบทดสอบใดที่วัดตัวแปรเดียวกันแล้วให้ค่าการวัดที่มีความสัมพันธ์ในทางใดทางหนึ่งกับค่าการวัดจากคุณลักษณะเฉพาะหรือพฤติกรรมที่ปรากฏในการอธิบายโครงสร้าง

แล้ว ก็สามารถแปลความหมายได้ตามที่ได้กล่าวมาแล้ว 2) ค่าการวัดหรือคะแนนจะมีความสัมพันธ์กันหรือไม่ อย่างไร ขึ้นอยู่กับผู้ตอบเป็นสำคัญ หรือกล่าวอีกอย่างหนึ่งว่าผู้ตอบจะเป็นผู้กำหนดค่าสถิติ ดังนั้นการประเมินความเที่ยงตรงของแบบทดสอบ(โดยเฉพาะตามแนวทฤษฎีการวัดดั้งเดิม :Classical Theory) เมื่อเปลี่ยนกลุ่มผู้สอบจึงควรทำการประเมินใหม่ หากทฤษฎีที่ใช้อธิบายตัวแปรหรืออธิบายโครงสร้างของตัวแปรนั้นขาดความมีนัยทั่วไป (Generalization)ซึ่งหมายถึงทฤษฎีที่เกี่ยวข้องกับตัวแปรนั้นอธิบายปรากฏการณ์ของพฤติกรรมมนุษย์ได้ไม่ครอบคลุม หรืออธิบายพฤติกรรมของคนส่วนใหญ่ (General) ไม่ได้ ตัวอย่างเช่น ทฤษฎีจิตวิทยาคลินิก (Clinical Psychology) อธิบายอาการของโรคจิตเภทหวาดระแวง (Paranoid Schizophrenia) ไว้ว่า ผู้ป่วยจะมีอาการเข้าใจผิด หรือประสาทหลอน(Delusion) เกี่ยวกับตนเองว่ามีคนคอยจ้องปองร้าย หรือตามล่าตนอยู่เสมอ และจะแสดงพฤติกรรมหลงตัวเองว่าเป็นผู้ยิ่งใหญ่ หรือบุคคลสำคัญ (Grandiosity) (Weiten. 2000 : 593) เช่น เข้าใจว่าตนเองเป็นประธานาธิบดี เป็นต้น เพื่อหวังที่จะปกป้องตนเองให้พ้นจากอันตราย หรือพ้นจากการทำร้ายจากผู้อื่น เมื่อนักจิตวิทยาได้เก็บรวบรวมข้อมูลเป็นระยะเวลาาน และหลายกลุ่มตัวอย่างหลายสถานที่ พบว่า เมื่อให้ผู้ป่วยวาดรูปคน ผู้ป่วยมักจะวาดรูปคนมีลักษณะเด่นคือ ศีรษะโต และตาโตกว่าปกติเสมอ (ไม่สมส่วน) เป็นต้น กล่าวคือ ทฤษฎีได้อธิบายโครงสร้างของตัวแปรไว้ประการหนึ่งว่า ผู้ป่วยที่มีอาการโรคจิตเภทหวาดระแวงเมื่อทำการทดสอบโดยให้วาดรูปคน (Draw-a-Person Test : DAP) แล้ว จะสังเกตได้ประการหนึ่งว่า รูปคนที่วาดนั้นจะมีลักษณะเด่นคือ ศีรษะจะใหญ่ และตาจะโตกว่าปกติเสมอ (ในทางปฏิบัตินักจิตวิทยาใช้ผลการทดสอบ เพื่อประกอบการวินิจฉัยร่วมกับผลการตรวจสอบด้านอื่นๆด้วย และการใช้แบบทดสอบชนิด DAP นี้จะมีวิธีการให้คะแนนเพื่อประเมินผู้สอบด้วย) เมื่อนักจิตวิทยา หรือจิตแพทย์ได้นำแบบทดสอบการวาดรูปคนนี้ไปทำการทดสอบกับ(กลุ่มผู้ป่วยโรคจิตเภทหวาดระแวง) คนใด กลุ่มใด เพศใด ชนชาติใด ฯลฯ ลักษณะภาพที่วาดออกมาก็จะมีลักษณะศีรษะโต ตาโตเสมอ ดังนี้ จึงเรียกว่า แบบทดสอบ หรือทฤษฎีที่ใช้ในการอธิบายโครงสร้างตัวแปรของแบบทดสอบนั้นมีนัยทั่วไป สามารถอธิบายปรากฏการณ์เกี่ยวกับพฤติกรรมของมนุษย์ได้เป็นนัยเดียวกันหรืออธิบายได้เหมือนกันโดยไม่จำกัดบุคคล และกลุ่มบุคคล หรือเรียกว่า อธิบายได้กว้างขวาง ครอบคลุม และทำให้นักจิตวิทยาสามารถทำนายได้ว่า “ถ้าให้ผู้ป่วยโรคจิตคนใดวาดรูปคนแล้ว ภาพที่ออกมาจะมีลักษณะตามที่ทฤษฎีอธิบายไว้ (ศีรษะโต ตาโต เป็นต้น)แล้ว เราจะทำนายได้ว่าผู้ป่วยคนนั้นมีอาการของโรคจิตเภทหวาดระแวง” ถ้าทฤษฎีใด แบบทดสอบใด สามารถอธิบาย หรือวัด จำแนกคนได้ดังนี้ เมื่อมี

การนำแบบทดสอบนั้นไปใช้อีกเมื่อใด กับใคร ที่ไหนก็ตาม ก็ไม่มีความจำเป็นที่จะต้องประเมินความเที่ยงตรงของแบบทดสอบนั้นอีก เพราะแบบทดสอบมีความเที่ยงตรงเชิงโครงสร้าง หรือมีความเที่ยงตรงเชิงทฤษฎีอยู่แล้ว (การที่ผู้สอบวาดภาพออกมาลักษณะ ศีรษะโต ตาโต แสดงให้เห็นว่าแบบทดสอบมีความเที่ยงตรง : Validity ; การที่ผู้สอบวาดภาพออกมาลักษณะ ศีรษะโต ตาโตทุกครั้ง แสดงให้เห็นว่าแบบทดสอบมีความเชื่อมั่น : Reliability)

ลักษณะการอธิบายโครงสร้างโดยการวิเคราะห์ความเป็นไปได้ของโครงสร้างของตัวแปรแสดงในตาราง 1 และ 2

ตาราง 1 Possible Analysis of the Construct “Anxiety” (Friedenberg 1995 : 253)

ตัวบ่งชี้ลักษณะ “ความวิตกกังวล” (Consistent with the presence “Anxiety”)

ด้านกายภาพ : อัตราการเต้นของหัวใจเพิ่มขึ้น, หายใจเร็วขึ้น, ความดันโลหิตสูงขึ้น
มีกระแสไฟฟ้าผิวหนังเพิ่มขึ้น

ด้านพฤติกรรม : ซึมเซื่องช้า, กัดเล็บ, หงุดหงิดกระวนกระวาย, ไม่สบตา

ตัวบ่งชี้ลักษณะที่ไม่สอดคล้องกับ “ความวิตกกังวล” หรือไม่วิตกกังวล (Inconsistent with the presence “Anxiety” or Absence of Construct)

ตัดสินใจรวดเร็ว

กล้าแสดงออก

พอใจในชีวิตความเป็นอยู่ของตนเอง

มีความเชื่อมั่นในตนเอง

มีสุขภาพดี

โครงสร้างตัวแปรที่เกี่ยวข้อง (Related construct)

ความซึมเศร้า

ความกลัว

ความเครียด

ตาราง 2 Possible Analysis of the Construct “Friendliness” (Friedenberg 1995 : 52)

พฤติกรรมที่ปรากฏสอดคล้องในโครงสร้าง”ความมีมิตรภาพ”

เป็นผู้เริ่มความสัมพันธ์ (Contact) กับผู้อื่น
 มักหาโอกาสพบปะกับผู้อื่นบ่อย ๆ
 ชอบที่จะอยู่กับผู้อื่นมากกว่าอยู่คนเดียว
 มักพูดคุยกับเพื่อนใหม่
 ยิ้ม พูด และหัวเราะในหมู่คนอื่น ๆ
 สบตาผู้อื่นบ่อย ๆ

พฤติกรรมที่ไม่สอดคล้องกับโครงสร้าง”ความมีมิตรภาพ”

หลบหลีกที่จะมีความสัมพันธ์กับผู้อื่น
 นาน ๆ ครั้งที่จะใช้เวลาอยู่กับผู้อื่น
 มักอยู่ตามลำพังมากกว่าอยู่ร่วมกับผู้อื่น
 พูดกับผู้อื่นเท่าที่จำเป็น
 ไม่มีความสุขเมื่ออยู่รวมกลุ่มกับผู้อื่น ซึ่งแสดงออกทางพฤติกรรมและทางกายภาพ
 นาน ๆ ครั้งที่จะสบตาผู้อื่น

วิธีการตรวจสอบความเที่ยงตรงเชิงโครงสร้างโดยใช้สถิติ

ดังที่ได้ให้เหตุผลมาก่อนแล้วว่าการอธิบายโครงสร้างนั้นเป็นกระบวนการที่ซับซ้อนและมีความสำคัญต่อกระบวนการตัดสินใจ ประเมิน หรือตรวจสอบความเที่ยงตรง ในประการหลังนั้น มีวิธีการในการตัดสินใจ ประเมิน หรือตรวจสอบได้หลายวิธีการ วิธีการหนึ่งคือ ใช้วิธีการทางสถิติ อธิบายความสัมพันธ์ (Correlation and Association) ของตัวแปรที่ต้องการตัดสินใจ ประเมิน หรือตรวจสอบว่ามีความเที่ยงตรงหรือไม่ สถิติมีบทบาทเป็นสิ่งที่ช่วยยืนยันหรือรับรองผลของการตรวจสอบจากข้อมูลเชิงประจักษ์ (Empirical Data) แต่ไม่ได้เป็นสิ่งกำหนดโครงสร้างของตัวแปรแต่อย่างใด การอธิบายโครงสร้างต่างหากเป็นสิ่งกำหนดโครงสร้างของตัวแปร ส่วนสถิติจะเป็นสิ่งตรวจสอบความสอดคล้องหรือไม่สอดคล้อง/มีความสัมพันธ์หรือไม่มีความสัมพันธ์ระหว่างข้อมูลเชิงประจักษ์ของค่าการวัดตัวแปรเท่านั้น มีเทคนิคสถิติหรือเทคนิคการตรวจสอบความเที่ยงตรงเชิง

โครงสร้างของแบบทดสอบ (Construct Validation Techniques) ซึ่งจะกล่าวไว้เป็นมโนทัศน์เบื้องต้นดังนี้

1. **ความเที่ยงตรงเชิงความสอดคล้อง (Congruent Validity)** เป็นวิธีการตรวจสอบความเที่ยงตรงของแบบทดสอบ (Test) ฉบับใหม่กับแบบทดสอบฉบับเดิมที่มีความเที่ยงตรงเชิงโครงสร้างอยู่แล้ว (Established Test) แบบทดสอบทั้งสองต่างก็วัดตัวแปรเดียวกัน ค่าการวัดจากแบบทดสอบทั้งสองฉบับควรมีความสัมพันธ์กันสูงในทางบวก (High Positive Correlations)

2. **ความเที่ยงตรงเหมือน (Convergent Validity)** เป็นวิธีการตรวจสอบความเที่ยงตรงของแบบทดสอบฉบับใหม่กับแบบทดสอบฉบับเดิมที่มีความเที่ยงตรงเชิงโครงสร้างอยู่แล้ว แต่แบบทดสอบฉบับเดิมนี้อัดตัวแปรที่เกี่ยวข้อง (Related) กับแบบทดสอบฉบับใหม่ ตัวอย่างเช่น แบบทดสอบฉบับใหม่ที่ต้องการหาความเที่ยงตรงเป็นแบบวัดความวิตกกังวล (Anxiety) แบบทดสอบฉบับเดิมเป็นแบบวัดความซึมเศร้า (Depression) หรือแบบวัดความกลัว (Fear) หรือแบบวัดความเครียด (Stress) ค่าการวัดจากแบบทดสอบทั้งสองฉบับควรมีความสัมพันธ์กันสูงในทางบวก

3. **ความเที่ยงตรงเชิงเกณฑ์สัมพันธ์ (Criterion Validity)** เป็นวิธีการตรวจสอบความเที่ยงตรงของแบบทดสอบฉบับใหม่กับค่าการวัดลักษณะทางกายภาพและพฤติกรรม (Physiological and Behavioral) ที่สัมพันธ์กับตัวแปรนั้น เช่น คะแนนจากแบบทดสอบวัดความวิตกกังวลควรจะทำนายพฤติกรรมการสับตาขณะสัมภาษณ์ของผู้ที่มีความวิตกกังวลได้ ค่าการวัดจากแบบทดสอบฉบับใหม่กับค่าการวัดจากเกณฑ์ควรมีความสัมพันธ์กันสูงมีในทางบวก

4. **ความเที่ยงตรงเชิงการจำกัดกระทำทดลอง (Experimental Manipulation)** ใช้วิธีการตรวจสอบโดยนำแบบทดสอบฉบับที่สร้างขึ้นใหม่นั้นไปวัดกลุ่มตัวอย่าง (Test Takers) แล้วนำกลุ่มตัวอย่างนั้นไปดำเนินการทดลองโดยการจำกัดกระทำตัวแปรการทดลองตัวแปรเดียวกันกับตัวแปรที่วัดโดยแบบทดสอบฉบับที่สร้างขึ้นนั้น ค่าการวัดลักษณะทางกายภาพและพฤติกรรมที่วัดได้จากการทดลองนั้นควรมีความสัมพันธ์กันสูงในทางบวกกับคุณลักษณะที่วัด เช่น เมื่อจำกัดกระทำทดลองให้กลุ่มตัวอย่างมีความวิตกกังวล แล้ววัดค่าความวิตกกังวลจากการทดลองโดยการสังเกตพฤติกรรม ค่าที่วัดได้นี้ควรจะสูงเมื่อวัดกับกลุ่มตัวอย่างที่ทำคะแนนจากแบบทดสอบฉบับใหม่ที่วัดความวิตกกังวลนั้นได้คะแนนสูง หรืออยู่ในคนที่มีความวิตกกังวลสูง ส่วนในคนที่ทำคะแนนได้ต่ำจากแบบวัดความวิตกกังวลฉบับที่สร้างขึ้นใหม่นั้นก็ควรจะได้ค่าการวัดจากการทดลองต่ำด้วย ลักษณะของการตรวจสอบโดยการทดลองนี้มีสมมติฐานว่า ถ้าในสถานการณ์ที่ทฤษฎี

อธิบายว่าจะสามารถเปลี่ยนระดับคุณลักษณะต่าง ๆ นั้นได้โดยการจัดกระทำแล้ว จะมีความแตกต่างอย่างแท้จริงระหว่างผู้ที่ถูกจัดกระทำกับผู้ที่ไม่ถูกจัดกระทำ

5. **ความเที่ยงตรงเชิงจำแนก (Discriminate or Divergent Validity)** เป็นวิธีการตรวจสอบความเที่ยงตรงของแบบทดสอบฉบับใหม่กับค่าการวัดจากแบบทดสอบที่วัดตัวแปรที่มีโครงสร้างไม่สอดคล้องกับโครงสร้างของตัวแปรแบบทดสอบฉบับใหม่ เช่น การตรวจสอบความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบวัดความวิตกกังวลโดยเปรียบเทียบกับแบบทดสอบวัดความกล้าแสดงออก(Assertive Test) เป็นต้น ค่าที่วัดได้จากแบบทดสอบสองฉบับนี้ควรมีความสัมพันธ์กันสูงในทางลบ (High Negative Correlations)

6. **การตรวจสอบโดยการวิเคราะห์องค์ประกอบ (Factor Analysis)** การตรวจสอบโดยวิธีวิเคราะห์องค์ประกอบเป็นเทคนิคสถิติที่ใช้สหสัมพันธ์ขั้นสูง (Sophisticated Correlation Technique) ซึ่งเป็นการค้นหาสหสัมพันธ์ระหว่างค่าการวัดของกลุ่มข้อความ(Items)ในแบบทดสอบ แบบแผนของสหสัมพันธ์ของข้อความนั้นใช้เป็นสิ่งจำแนกจำนวนองค์ประกอบหรือลักษณะที่วัดได้ต่างๆกันจากแบบทดสอบ ถ้าแบบทดสอบมีโครงสร้างเฉพาะหรือครอบคลุมโครงสร้างเดียว (Cover a Single, Specific Construct) จำนวนองค์ประกอบควรจะมีน้อยเพราะวิธีการวิเคราะห์องค์ประกอบเป็นวิธีการเพื่อลดจำนวนข้อความให้น้อยลงเป็นกลุ่มข้อความซึ่งเปรียบเป็นหนึ่งข้อความและเรียกชื่อใหม่ว่าองค์ประกอบหรือแฟคเตอร์ ข้อความในแบบทดสอบจะมีสหสัมพันธ์กันสูงถ้าวัดองค์ประกอบหรือลักษณะเดียวกัน กล่าวอีกนัยหนึ่งว่า ถ้าข้อความในแบบทดสอบนั้นมีสหสัมพันธ์กันแล้วเป็นเพราะข้อความเหล่านั้นวัดลักษณะเดียวกันหรือมีองค์ประกอบร่วมกัน ลักษณะดังนี้เคอร์ลินเจอร์เรียกเป็นสำนวนว่า “go together” ซึ่งหมายถึงการวัดในสิ่งเดียวกัน (Friedenberg. 1995 : 254 ; Kerlinger. 1992 : 427) ดังนั้นถ้าวิเคราะห์องค์ประกอบได้เป็น 3 องค์ประกอบ ผลการวิเคราะห์จะแสดงให้เห็นว่ามีกลุ่มของข้อความที่มีสหสัมพันธ์กันอยู่ 3 กลุ่ม ระหว่างแต่ละกลุ่มข้อความจะไม่มีสหสัมพันธ์กันหรือมีความแตกต่างกันในการอธิบายโครงสร้างของตัวแปร หรือบางกรณีอาจยอมให้มีความสัมพันธ์กันได้ การตรวจสอบความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบนั้นควรได้จำนวนองค์ประกอบที่ไม่มาก (Expect Emergence of Small Set of Factor)

7. **ความเที่ยงตรงโดยลักษณะหลากหลาย-วิธีหลาย (Multitrait-Multimethod Matrix)** เป็นการตรวจสอบความเที่ยงตรงเชิงโครงสร้างด้วยเมตริกซ์สหสัมพันธ์ระหว่างลักษณะที่จะวัด (Traits) กับวิธีการวัด (Method) ค่าเมตริกซ์สหสัมพันธ์ที่ได้จะบอกลักษณะของความ

เที่ยงตรงเหมือน (Convergent Validity) และความเที่ยงตรงเชิงการจำแนก (Discriminate Validity) โดยถ้าใช้แบบทดสอบต่างๆ กัน (หรือวิธีการทดสอบต่างๆ กัน) ต่างก็วัดลักษณะเดียวหรือคล้ายกันแล้ว สหสัมพันธ์จากผลการสอบด้วยแบบทดสอบเหล่านั้นจะสูง เรียกลักษณะของแบบทดสอบเช่นนี้ว่ามีความเที่ยงตรงเหมือน แต่ถ้าแบบทดสอบต่าง ๆ ต่างก็วัดลักษณะต่างกันแล้วผลการสอบด้วยแบบทดสอบเหล่านั้นจะไม่มีสหสัมพันธ์กันหรือมีสหสัมพันธ์กันต่ำ เรียกลักษณะแบบทดสอบเช่นนี้ว่ามีความเที่ยงตรงเชิงจำแนก และสามารถสรุปเป็นหลักเกณฑ์ในการวิเคราะห์ความเที่ยงตรงได้ดังนี้ (สำเร็จ บุญเรืองรัตน์. 2528 : 5 อ้างอิงมาจาก Marsh, Smith and Barnes. 1983 ; Centra. 1971)

- (1) ค่าความเชื่อมั่น (Reliability) ของแบบทดสอบจะสูงกว่าค่าความเที่ยงตรงเหมือน
- (2) ค่าความเที่ยงตรงเหมือนจะมีค่ามากกว่าค่าความเที่ยงตรงเชิงจำแนก
- (3) ค่าความเที่ยงตรงเหมือนจะมีค่าสูง โดยปกติจะมากกว่าศูนย์อย่างมีนัยสำคัญ
- (4) ค่าความเที่ยงตรงเชิงจำแนกจะมีค่าต่ำและมักเป็นศูนย์

8. ความเที่ยงตรงโดยความสอดคล้องภายใน (Internal Consistency) เป็นการตรวจสอบความเที่ยงตรงเชิงโครงสร้างโดยวิธีการวิเคราะห์รายข้อ (Item Analysis) ของแบบทดสอบโดยแยกกลุ่มผู้สอบที่ได้คะแนนสูงและต่ำแล้วพิจารณาความแตกต่างรายข้อระหว่างคะแนนกลุ่มสูงและกลุ่มต่ำข้อใดมีความแตกต่างอย่างมีนัยสำคัญทางสถิติก็ถือว่ามีความเที่ยงตรงรายข้อ (Item Validity) ดังนั้นแบบทดสอบใดมีค่าอำนาจจำแนกรายข้อสูงจะถือว่ามีความเที่ยงตรงเชิงโครงสร้างสูงด้วย หรืออาจพิจารณาจากสหสัมพันธ์กับคะแนนรวม ถ้าข้อใดมีค่าสัมประสิทธิ์สหสัมพันธ์กับคะแนนรวมสูงแสดงว่ามีความสอดคล้องภายในหรืออาจพิจารณาจากคะแนนการสอบของแบบทดสอบย่อยกับคะแนนรวมทั้งฉบับ (Total Score) แบบทดสอบย่อยใดมีค่าสัมประสิทธิ์สหสัมพันธ์กับคะแนนรวมทั้งฉบับก็ถือว่าแบบทดสอบย่อยนั้นมีความสอดคล้องภายในหรือมีความเที่ยงตรงเชิงโครงสร้าง (บุญเชิด ภิญญอนันตพงษ์. มปป : 310) หรืออาจตรวจสอบจากการหาความเชื่อมั่นของแบบทดสอบแบบค่าสอดคล้องภายใน ซึ่งถือว่าแบบทดสอบใดมีค่าสัมประสิทธิ์สหสัมพันธ์ของความเชื่อมั่นแบบสอดคล้องภายในสูง ย่อมมีความเที่ยงตรงเชิงโครงสร้างสูงด้วย (บุญเชิด ภิญญอนันตพงษ์. มปป : 311 อ้างอิงมาจาก Ebel and Mehrens. 1967 : 252) ความสอดคล้องภายในนี้แสดงให้เห็นว่าข้อคำถามทั้งหมดหรือแบบทดสอบย่อย

ทั้งหมดนั้นมีโครงสร้างของตัวแปรที่ต้องการวัดเหมือนกัน หรือข้อคำถามทั้งหมดวัดในสิ่งเดียวกัน กรณีดังกล่าวถือข้อตกลงว่าจะเน้นรวมทั้งฉบับเป็นความเที่ยงตรงของแบบทดสอบอยู่แล้ว (The total score is assumed to be valid) (Kerlinger. 1992 : 427)

9. **ความเที่ยงตรงโดยอาศัยกลุ่มที่แตกต่างกัน (Group Differentiation)** เป็นการตรวจสอบความเที่ยงตรงโดยอาศัยการสรุปจากลักษณะเฉพาะหรือพฤติกรรมของกลุ่มบุคคลหรือตัวอย่างที่มีความแตกต่างกันโดยการสรุปนั้นมีเหตุผลทางทฤษฎีสนับสนุนเป็นโครงสร้างของตัวแปรที่ต้องการวัดหรือต้องการเปรียบเทียบ เช่นถ้าเราจะตรวจสอบความเที่ยงตรงของแบบทดสอบวัดความก้าวร้าว (Aggressive Behavior) เราอาจเปรียบเทียบความแตกต่างระหว่างผลการสอบในกลุ่มนักโทษกับกลุ่มนักบวชซึ่งควรพบว่ามี ความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ทำให้สรุปได้ว่าแบบทดสอบนั้นมีความเที่ยงตรงเชิงโครงสร้าง การตรวจสอบความเที่ยงตรงโดยวิธีการนี้คล้าย ๆ กับวิธีการในข้อ 4 แต่ต่างกันตรงที่ลักษณะของวิธีการนี้ไม่มีการจัดกระทำ (Manipulation) การตรวจสอบความเที่ยงตรงแบบนี้มีสมมติฐานว่า ถ้าเราจัดอันดับคนตามคุณลักษณะที่ต้องการวัดนั้นให้อยู่ในอันดับสูง เมื่อทดสอบด้วยแบบทดสอบแล้วค่าการวัดที่ได้จากคนที่อยู่อันดับสูงควรจะสูงกว่าคนที่ถูกจัดอันดับตามคุณลักษณะที่ต้องการวัดนั้นอยู่ในอันดับต่ำ หรือมีความแตกต่างกันของค่าการวัดจากคนที่อยู่กลุ่มต่าง(อันดับ)กันนี้

10. **ความเที่ยงตรงตามนัยทั่วไป (Generalizability Studies)** เป็นการตรวจสอบความเที่ยงตรงโดยอาศัยการพิจารณาข้อมูลที่ได้จากใช้แบบทดสอบในต่างสถานการณ์หรือ ต่างเงื่อนไข (Varying conditions) (Brown. 1983 : 142) ซึ่งโดยปกติแล้วข้อมูลที่มีความเที่ยงตรงนั้นจะเป็นข้อมูลที่ได้มาจากสถานการณ์เฉพาะ (Situation-Specific) นั่นคือผลของการศึกษาความเที่ยงตรง ทุกครั้งนั้นข้อมูลที่ได้มาขึ้นอยู่กับลักษณะ (Characteristics) ที่ได้มาจากสถานการณ์เฉพาะ ในกรณีดังกล่าวมีความหมายเป็น 2 นัยคือ นัยแรก หมายถึงแบบทดสอบฉบับหนึ่งมีความเที่ยงตรงหลายด้าน (Many Validity) และนัยที่สองแม้ว่าแบบทดสอบจะได้รับการพัฒนาความเที่ยงตรงจนเข้าสู่ระดับมาตรฐานแล้ว (Established Test) ในสถานการณ์หนึ่งเงื่อนไขหนึ่ง แต่นั่นก็ไม่ได้หมายความว่าแบบทดสอบนั้นมีความเที่ยงตรงในสถานการณ์อื่น ๆ ดังนั้นในทุกๆ สถานการณ์แบบทดสอบที่สร้างขึ้นนั้นต้องผ่านการทดลองใช้จนกระทั่งมั่นใจได้ว่าความเที่ยงตรงของแบบทดสอบนั้นสามารถพัฒนาเข้าสู่ความเป็นมาตรฐานโดยข้อมูลเชิงประจักษ์ ในความเห็นของบราวน์ (Brown. 1983 : 125) เห็นว่าการใช้แบบทดสอบที่ขาดการประเมินประสิทธิผล (Effectiveness) นั้นเป็นทั้งการปฏิบัติที่ไม่ถูกต้อง และเสี่ยงต่อการขาดจริยธรรม การตรวจสอบ

ดังกล่าวอาจใช้สถานการณ์-เงื่อนไขต่างกัน เช่น ใช้กลุ่มตัวอย่างในการสอบต่างกัน เป็นต้น การพิจารณาผลการตรวจสอบความเที่ยงตรงหลายๆ เงื่อนไขนั้น ผู้สอบควรตั้งคำถามว่า 1) ผลการตรวจสอบควรจะเหมือนกันถ้าใช้กลุ่มผู้ตอบที่เป็นตัวแทนมาจากประชากรเดียวกันหรือไม่ คำถามนี้เป็นปัญหาเกี่ยวกับการตรวจสอบความคงที่ของความเที่ยงตรง (Cross-Validation) 2) ในความแตกต่างของสถานการณ์และกลุ่มตัวอย่างนั้นจะถือเอาความเที่ยงตรงจากผลของสถานการณ์ใดกลุ่มตัวอย่างใด คำถามแรกเป็นเรื่องของความมีนัยทั่วไป (Generalizability) (Brown. 1983 : 125 Citing. Brennan and Kane 1979 ; Cronbach, Gleser, Nanda and Rajaratnam. 1972 ; Cronbach, Rajaratnam and Gleser. 1963) และคำถามหลังเป็นเรื่องของความเที่ยงตรงในประชากร (Population Validity) (Brown. 1983 : 125 Citing Brelan. 1979) ทั้งสองคำถามเป็นคำถามเกี่ยวกับเรื่องของความมีนัยทั่วไปของความเที่ยงตรง (Validity Generalization)

สรุป

ความเที่ยงตรงเชิงโครงสร้างและการตรวจสอบความเที่ยงตรงเชิงโครงสร้างเป็นปัญหาทางการวัดทางจิตวิทยามากเพราะโดยเหตุที่ว่าโครงสร้างในความหมายดังกล่าวไว้ในบทความนี้เป็นนามธรรม เป็นทฤษฎีหรือสมมติฐานทางการวัดเชิงจิตวิทยา ซึ่งไม่มีตัวตน หรือปรากฏการณ์จริงการกำหนดโครงสร้างของเนื้อหาหรือตัวแปรที่จะวัดนั้นเราใช้วิธีการวิเคราะห์เนื้อหาเชิงเหตุผล ซึ่งเรียกว่า การอธิบายโครงสร้าง ซึ่งเป็นการกำหนดรายการของพฤติกรรมหรือลักษณะเฉพาะที่ปรากฏเป็นโครงสร้างในตัวแปรเชิงจิตวิทยานั้น จากนั้นเราใช้การตรวจสอบหรือประเมินความเที่ยงตรงเชิงโครงสร้างโดยวิธีการวิเคราะห์เชิงประจักษ์ ซึ่งเป็นการนำข้อมูลจากการใช้แบบทดสอบที่ได้สร้างขึ้นนั้นไปหาความสัมพันธ์กับเกณฑ์โดยใช้วิธีการทางสถิติช่วยตรวจสอบและตัดสิน ซึ่งในบทความนี้ได้รวบรวมและนำเสนอเป็นมโนทัศน์พื้นฐานไว้หลายวิธี ผู้อ่านอาจใช้เป็นพื้นฐานของการศึกษาค้นคว้าเกี่ยวกับเรื่องของความเที่ยงตรงของเครื่องมือการวัดได้ต่อไป

เอกสารอ้างอิง

- บุญเชิด ภิญโญอนันตพงษ์. การวัดและการประเมินผลการศึกษาทฤษฎีและการประยุกต์. พิมพ์ครั้งที่ 2 กรุงเทพฯ : อักษรเจริญทัศน์, ม.ป.ป.
- สำเร็จ บุญเรืองรัตน์. ลักษณะหลากหลาย-วิธีหลาย : การประยุกต์ใช้วิเคราะห์ความเที่ยงตรงของแบบทดสอบ. สำนักพิมพ์สยามศึกษาจำกัด, 2528.
- Brown, Federick G. Principles of Educational and Psychological Testing. 3rd ed. New York : CBS College Publishing, 1983.
- Friedenberg, Lisa. Psychological Testing Design, Analysis and Use. Massachusetts : Allyn and Bacon, 1995.
- Kerlinger, Fred N. Foundation of Behavioral Research. 3rd ed. U.S.A. : Holt, Rinehart and Winston, Inc., Copyright renewed 1992.
- Nunnally, Jum C. Psychometric Theory. New York : McGraw-Hill, 1967.
- Weiten, Wayne. Psychology: Themes and Variations. 5th ed. California : Wadsworth Publishing Company, 2000.
- Wiersma, William. Research Method in Education An Introduction. 6th ed. Massachusetts : Allyn and Bacon, 1995.