

การพัฒนาแบบสอบถามและแบบวัดทางจิตวิทยา

รศ.ดร.สมบัติ ท้ายเรือคำ¹

Assoc.Prof.Dr.Sombat Tayraukham¹

แบบสอบถาม (Questionnaires) และแบบวัดทางจิตวิทยา (Psychological Test) เป็นเครื่องมือที่เหมาะสมในการรวบรวมข้อมูลในการวิจัยเชิงสำรวจ ในกรณีที่กลุ่มตัวอย่างหรือกลุ่มประชากรที่จะรวบรวมข้อมูลนั้นอยู่ในลักษณะที่กระจัดกระจายกันมากๆ ประกอบกับผู้วิจัยมีงบประมาณและเวลาในการวิจัยค่อนข้างจำกัด ซึ่งเครื่องมือดังกล่าวจะประกอบด้วยชุดของคำถามที่ต้องการให้กลุ่มตัวอย่างตอบ โดยกาเครื่องหมายหรือเขียนตอบ หรือกรณีที่กลุ่มตัวอย่างอ่านหนังสือไม่ได้หรืออ่านได้ยากอาจใช้วิธีสัมภาษณ์ตามแบบสอบถาม แบบสอบถามนิยมถามเกี่ยวกับข้อเท็จจริงและความคิดเห็นที่ไม่ซับซ้อนของบุคคล แต่กรณีที่เป็นแบบวัดทางจิตวิทยาใช้วัดเกี่ยวกับคุณลักษณะทางจิต (Psychology Trail) ของบุคคล

โครงสร้างของแบบสอบถามและแบบวัด โดยทั่วไปจะมีโครงสร้างหรือส่วนประกอบ 3 ส่วน ดังนี้

1. คำชี้แจงในการตอบ ที่ปกของแบบสอบถามและแบบวัดจะเป็นคำชี้แจง ซึ่งมักจะระบุถึงจุดประสงค์ในการให้ตอบ หรือจุดมุ่งหมายของการทำวิจัย โดยจะอธิบายลักษณะของเครื่องมือ วิธีการตอบ พร้อมตัวอย่าง
2. สถานภาพส่วนตัวผู้ตอบ ส่วนที่ 2 ของแบบสอบถาม หรือแบบวัดมักจะให้ตอบเกี่ยวกับรายละเอียดส่วนตัว เช่น เพศ อายุ ระดับการศึกษา อาชีพ แล้วแต่กรณี
3. ข้อคำถามเกี่ยวกับข้อเท็จจริง ความคิดเห็นที่ต้องการถาม หรือคุณลักษณะที่ต้องการวัด ซึ่งเป็นส่วนที่สำคัญที่สุดซึ่งจะช่วยให้ได้ข้อมูลหรือรายละเอียดเกี่ยวกับเรื่องที่ต้องการศึกษา

ชนิดของแบบสอบถามและแบบวัดทางจิตวิทยา จำแนกได้เป็น 2 ชนิดใหญ่ๆ คือ

1. **แบบสอบถามแบบปลายเปิด (Open-Ended Questionnaires)** เป็นคำถามที่ไม่ได้กำหนดคำตอบไว้ให้เลือก แต่เปิดโอกาสให้ผู้ตอบแบบสอบถามตอบโดยใช้คำพูดของตนเอง คำถามแบบปลายเปิดนี้จะเสียเวลาในการตอบมาก และสรุปผลการวิจัยได้ยาก ถ้าใช้ควบคู่กับแบบอื่นๆ แล้วผู้ตอบส่วนใหญ่มักไม่ตอบแบบปลายเปิด หรือตอบเพียงเล็กน้อย ในการสร้างแบบสอบถามหรือแบบวัดครั้งแรก ผู้วิจัยอาจสร้างแบบปลายเปิดแล้วนำไปทดลองใช้เพื่อจะได้คำตอบต่างๆ ซึ่งจะนำมาสร้างเป็นแบบปลายปิดในภายหลัง

2. **แบบสอบถามแบบปลายปิด (Close-Ended Questionnaires)** เป็นคำถามที่มีคำตอบให้ผู้ตอบเขียนเครื่องหมาย ✓ ลงหน้าข้อความหรือในช่องที่ตรงกับความเป็นจริง หรือความคิดเห็นของตน ซึ่งแบบสอบถามหรือแบบวัดแบบปลายปิดมีหลายรูปแบบ ได้แก่

2.1 แบบให้เลือกรับตัวเลือกที่ตรงกับความเป็นจริง หรือความคิดเห็นของตนเพียงคำตอบเดียวจาก 2 คำตอบ ดังตัวอย่าง

¹ รองศาสตราจารย์ ดร.ภาควิชาวิจัยและพัฒนาการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

- การคุมกำเนิดขัดต่อพุทธศาสนาหรือไม่ ☐ ขัด ☐ ไม่ขัด
ท่านเห็นว่าเด็กหญิงควรเรียนวิชาฟุตบอลหรือไม่ ☐ ควร ☐ ไม่ควร
ปกติท่านสอนให้บุตรหลานของท่านทำงานบ้านหรือไม่ ☐ สอน ☐ ไม่สอน

2.2 แบบให้เลือกคำตอบที่ตรงกับความเป็นจริงหรือความคิดเห็นของตนเพียงคำตอบเดียวจากหลายคำตอบ (มากกว่า 2 คำตอบ) ดังตัวอย่าง

ท่านชอบรายการโทรทัศน์ประเภทใดมากที่สุด

- ☐ ดนตรี ☐ เกมโชว์ ☐ ข่าว
☐ ละครโทรทัศน์หลังข่าว ☐ ภาพยนตร์เรื่องยาว

2.3 แบบให้เลือกคำตอบที่ตรงกับความเป็นจริง หรือความคิดเห็นของตนได้หลายคำตอบ ดังตัวอย่าง ท่านเคยได้รับความรู้เกี่ยวกับโรคมะเร็งจากที่ใดบ้าง (ตอบได้หลายคำตอบตามความเป็นจริง)

- ☐ บิดา – มารดา ☐ วารสาร หนังสือ ตำราต่าง ๆ
☐ วิทยุ ☐ ครู-อาจารย์
☐ โทรทัศน์ ☐ เพื่อน
☐ หนังสือพิมพ์

2.4 แบบมาตราส่วนประมาณค่า (Rating Scale) โดยให้ผู้ตอบเลือกตอบตามระดับความคิดเห็นของตน หรือคุณลักษณะที่ตัวเองเป็น ดังตัวอย่าง

1. อาจารย์ที่สอนในระดับปริญญาตรี ควรมีความรู้ไม่ต่ำกว่าปริญญาโท

- ☐ เห็นด้วยอย่างยิ่ง ☐ เห็นด้วย ☐ ไม่แน่ใจ ☐ ไม่เห็นด้วย ☐ ไม่เห็นด้วยอย่างยิ่ง

ลักษณะของคำถามแบบนี้อาจจัดให้อยู่ในรูปของตารางก็ได้

รายการ	เห็นด้วยอย่างยิ่ง	เห็นด้วย	ไม่แน่ใจ	ไม่เห็นด้วย	ไม่เห็นด้วยอย่างยิ่ง
1. อาจารย์ที่สอนในระดับปริญญาตรีควรมีความรู้ไม่ต่ำกว่าปริญญาโท					

ส่วนแบบวัดเจตคติ (Attitude Scale) เป็นชุดของข้อคำถามด้านความรู้สึกที่มีต่อสิ่งหนึ่งสิ่งใดในทางบวกหรือทางลบ ซึ่งมีการกำหนดระดับของคำตอบไว้เป็นช่วงๆ (Interval) ให้กลุ่มตัวอย่างเลือกตอบตามความรู้สึกที่แท้จริง ซึ่งเจตคติจะประกอบด้วย 3 ส่วน หรือ 3 องค์ประกอบ คือ 1) ส่วนที่เป็นความรู้สึกหรือการประเมิน (Feeling or Evaluating Component) ซึ่งเป็นกริยาทำที่ที่แสดงออกว่าชอบหรือไม่ชอบ ดีหรือไม่ดี

ต่อสิ่งเร้านั้น 2) ส่วนที่เป็นความรู้หรือความเชื่อ (Cognitive or Belief Component) ซึ่งเป็นการรับรู้หรือความเชื่อของบุคคลต่อสิ่งเร้านั้น 3) ส่วนที่เป็นแนวโน้มการเกิดพฤติกรรม (Behavioral Component) ซึ่งเป็นความโน้มเอียงที่จะกระทำหรือจะปฏิบัติอย่างใดอย่างหนึ่งต่อสิ่งเร้านั้น ฉะนั้นการวัดเจตคติจึงต้องวัดทั้ง 3 องค์ประกอบของเจตคติและจะต้องวัดเป็นภาพรวมๆ โดยพิจารณากิริยาท่าทีการตอบสนองต่อสิ่งเร้าในหลายด้าน หลายประการรวมกัน มิใช่วัดจากการกระทำหรือพฤติกรรมอย่างเดียว นอกจากนั้นการวัดเจตคดียังต้องบ่งบอกทั้งปริมาณความมากน้อยของเจตคติที่มีต่อสิ่งเร้าและทิศทางที่บอกว่ามีเจตคติไปในทางบวกหรือทางลบด้วย

แบบวัดเจตคติที่นิยมมีใช้จะเป็นของลิเคิร์ท (Likert 's Scale) แบบวัดเจตคติของลิเคิร์ทประกอบด้วยข้อคำถามที่แสดงเจตคติ หรือความรู้สึกต่อสิ่งหนึ่งสิ่งใดในทางบวก ในแบบวัดจะต้องประกอบไปด้วยทั้งข้อคำถามทางบวกและทางลบในจำนวนพอ ๆ กัน ระดับเจตคติตามแบบของลิเคิร์ทนิยมแบ่งออกเป็น 5 ระดับ (Scale) คือ เห็นด้วยอย่างยิ่ง เห็นด้วย ไม่แน่ใจ ไม่เห็นด้วย และไม่เห็นด้วยอย่างยิ่ง ถ้าเป็นข้อความทางบวกจะมีคะแนน 5 4 3 2 1 (หรือ 4 3 2 1 0) ถ้าเป็นข้อความทางลบจะมีคะแนน 1 2 3 4 5 (หรือ 0 1 2 3 4)

เช่น การวัดเจตคติต่อการเรียนวิชาคณิตศาสตร์

รายการ	เห็นด้วยอย่างยิ่ง	เห็นด้วย	ไม่แน่ใจ	ไม่เห็นด้วย	ไม่เห็นด้วยอย่างยิ่ง
องค์ประกอบที่ 1 ข้าพเจ้าเชื่อว่าเพราะวิชาคณิตศาสตร์จึงทำให้ข้าพเจ้าคิดอย่างเป็นระบบ					
องค์ประกอบที่ 2 ข้าพเจ้ารู้สึกมีความสุขเมื่อได้ทำการบ้านเกี่ยวกับตัวเลข					
องค์ประกอบที่ 3 เมื่อถึงเวลาเรียนวิชาคำนวณข้าพเจ้าจะมาถึงห้องเรียนเป็นคนแรก					

ในการสร้างหรือเขียนข้อคำถามในแต่ละข้อ ควรประกอบไปด้วย 3 ส่วน คือ ประธาน + ตัวบ่งชี้ของตัวแปรที่ต้องการวัด (จากนิยามศัพท์เฉพาะ) + สถานการณ์ ส่วนการตอบจะให้ผู้ตอบตอบทุกข้อโดยแต่ละข้อเลือกอันดับที่ตรงกับความรู้สึกที่แท้จริง หรือตรงกับความเป็นจริงของตนเองมากที่สุด ผู้ตอบได้คะแนนตามระดับที่เลือกตอบแต่ละข้อแล้วนำมาหาค่าเฉลี่ย ได้เป็นคะแนนเจตคติของผู้นั้น

กระบวนการสร้างและพัฒนาแบบสอบถามและแบบวัดทางจิตวิทยา

การสร้างเครื่องมือเก็บรวบรวมข้อมูลการวิจัยแต่ละประเภทจะมีเทคนิคและวิธีการแตกต่างกันไปแต่ก็อาจกล่าวถึงขั้นตอนในการสร้างร่วมกันได้ เป็นขั้นตอนใหญ่ๆ 7 ขั้นตอน คือ

1. การกำหนดสิ่งที่ต้องการวัด เป็นการกำหนดสาระสำคัญของสิ่งที่ต้องการวัดหรือตัวแปรที่ต้องการวัด (Trait หรือ Variables) ซึ่งอาจกำหนดเป็นขอบเขตและโครงสร้างของสิ่งที่ต้องการวัด เช่น ขอบเขตโครงสร้างของเนื้อหา ขอบเขตโครงสร้างเจตคติ ขอบเขตโครงสร้างของพฤติกรรมหรืออาจจะกำหนดเป็นตัวบ่งชี้ของสิ่งที่ต้องการวัด เช่น ตัวบ่งชี้ความสำเร็จ ตัวบ่งชี้คุณภาพทางการเรียน เป็นต้น

2. การเลือกประเภทของเครื่องมือ เครื่องมือที่ใช้ในการเก็บข้อมูลการวิจัย มีหลายประเภทตามที่กล่าวมาแล้ว ผู้วิจัยจะต้องเลือกเครื่องมือที่เหมาะสมกับลักษณะข้อมูลของสิ่งที่ต้องการวัด เช่น ถ้าต้องการวัดเจตคติต่อ การเรียนวิชาคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 ใช้เครื่องมือประเภทแบบวัดเจตคติต่อการเรียนคณิตศาสตร์ส่วนประมาณค่า 5 ระดับ

3. การเขียนข้อคำถามและจัดฉบับ เมื่อเลือกประเภทของเครื่องมือแล้ว ก็ดำเนินการเขียนข้อคำถามที่จะ ใช้วัดตามลักษณะของเครื่องมือชนิดนั้นๆ รวมทั้งเขียนรายละเอียดของส่วนประกอบอื่นๆ ของเครื่องมือที่จัดขึ้นในขั้นตอนนี้ยังเป็นฉบับร่าง ยังไม่เหมาะที่จะนำไปใช้ในการเก็บรวบรวมข้อมูลเนื่องจากยังไม่ได้ตรวจสอบคุณภาพการเขียนข้อคำถามถ้าเป็นเครื่องมือชนิดแบบสอบถาม หรือแบบวัดเจตคติต่อการเรียน ต้องมีการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องโดยเฉพาะทฤษฎีทางจิตวิทยา แล้วนิยามศัพท์เฉพาะ (นิยามเชิงปฏิบัติการ) ซึ่งในการนิยามศัพท์เฉพาะสิ่งที่ขาดไม่ได้คือตัวบ่งชี้ หรือตัวชี้วัดในสิ่งที่ต้องการจะวัดอันจะนำไปสู่การสร้างข้อคำถาม ซึ่งการทำการนิยามศัพท์เฉพาะจะทำให้การสร้างเครื่องมือได้สอดคล้องหรือตรงกับนิยามศัพท์เฉพาะ อันเป็นคุณภาพของเครื่องมือวัดเรื่อง ความตรงเชิงเนื้อหา (Content Validity) หรือความตรงเชิงโครงสร้างทฤษฎี (Construct Validity)

4. การตรวจสอบความตรง (Validity) เป็นการตรวจสอบความถูกต้อง ความสอดคล้องและความครบถ้วนของพฤติกรรมตามตัวบ่งชี้ที่ต้องการวัด ความตรงของเครื่องมือ หมายถึง ความสามารถของเครื่องมือวัดที่วัดได้ในสิ่งที่ต้องการวัด ความตรงจึงเป็นคุณสมบัติที่สำคัญที่สุดของเครื่องมือวัดทุกชนิด ซึ่งความตรงของแบบสอบถามหรือแบบวัดทางจิตวิทยาที่สำคัญคือความตรงตามเนื้อหา (Content Validity) และความตรงตามโครงสร้างทฤษฎี (Construct Validity) ในการตรวจสอบความตรงอาจตรวจสอบได้เพียงชนิดใดชนิดหนึ่งหรือหลายชนิด เหมาะสมในกรณีที่ตรวจสอบโดยผู้ทรงคุณวุฒิ ผู้ตรวจสอบควรเป็นผู้ที่มีความรู้ความเชี่ยวชาญเกี่ยวกับสิ่งที่วัด และควรมีจำนวนหลายคนอย่างน้อย 3 คนขึ้นไป ซึ่งอาจเป็นคณะผู้สร้างเครื่องมือ หรือคณะผู้ทรงคุณวุฒิจากภายนอกที่ไม่ใช่ผู้สร้างเครื่องมือ

5. การปรับปรุงแก้ไขฉบับก่อนทดลองใช้ เป็นการปรับปรุงแก้ไขเครื่องมือโดยพิจารณาจากข้อคิดเห็นของผู้ทรงคุณวุฒิตรวจสอบในข้อ 4

6. การทดลองใช้เครื่องมือ (Try Out) เป็นการนำเครื่องมือที่ได้ปรับปรุงแก้ไขแล้วในข้อ 5 ไปทดลองใช้เก็บรวบรวมข้อมูลจากแหล่งข้อมูลที่ไม่ใช่กลุ่มตัวอย่างแต่มีลักษณะคล้ายคลึงกันกับกลุ่มตัวอย่างแล้วนำข้อมูลมาวิเคราะห์เพื่อหาคุณภาพของเครื่องมือได้แก่ ค่าอำนาจจำแนก (Discrimination Power) และค่าความเที่ยง (Reliability) เครื่องมือบางชนิดอาจต้องการวิเคราะห์เพียงค่าความเที่ยงหลังจากที่มีคุณภาพในเรื่องความตรงแล้ว หรือในกรณีที่นิยมเครื่องมือของนักวิจัยคนอื่นที่สร้างและหาคุณภาพเครื่องมือในการวิจัยในเรื่องของอำนาจจำแนกไว้แล้ว เช่น แบบสอบถามหรือแบบวัดทางจิตวิทยาจะต้องเลือกให้เหมาะสมกับชนิดของเครื่องมือ

7. การปรับปรุงแก้ไขเป็นฉบับใช้จริง เป็นการปรับปรุงแก้ไขเครื่องมือให้มีคุณภาพสูงขึ้น โดยพิจารณาจากผลการวิเคราะห์คุณภาพของเครื่องมือในข้อ 6 ถ้าผลการวิเคราะห์คุณภาพพบว่า คุณภาพไม่ต่ำมากนักก็สามารถปรับปรุงแล้วนำไปใช้จริงได้ แต่ถ้าผลการวิเคราะห์พบว่าคุณภาพต่ำมากก็อาจต้องปรับปรุงแล้วนำไปทดลองใช้อีกจนกว่าจะมีคุณภาพดีพอสมควร จึงจะใช้เก็บรวบรวมข้อมูลได้

การตรวจสอบคุณภาพของแบบสอบถามและแบบวัดจิตวิทยาที่เป็นมาตราส่วนประมาณค่า (Rating scale)

ในการหาคุณภาพของแบบวัดทางจิตวิทยาและแบบสอบถามที่เป็นมาตราส่วนประมาณค่า จะดำเนินการหาคุณภาพตามลำดับขั้นตอนในการวิจัยดังต่อไปนี้ คือ

1. การหาความตรง หรือความเที่ยงตรง (Validity)
2. การหาค่าอำนาจจำแนก (Discrimination Power)
3. การหาความเที่ยง หรือความเชื่อมั่น (Reliability)

ซึ่งในตอนนี้จะกล่าวถึงลำดับขั้นตอน วิธีการ และการคำนวณ หาคุณภาพในแต่ละเรื่องดังนี้

1. การหาความตรง หรือความเที่ยงตรง

เครื่องมือรวบรวมข้อมูลที่มีลักษณะเป็นแบบมาตราส่วนประมาณค่าควรมีคุณภาพด้านความตรงเชิงเนื้อหา (Content Validity) และความตรงเชิงโครงสร้าง (Construct Validity) ดังนี้

ในด้านความตรงเชิงเนื้อหาพิจารณาโดยนำเครื่องมือนั้น พร้อมทั้งนิยามศัพท์เฉพาะของตัวแปรที่มุ่งวัด ให้ผู้เชี่ยวชาญพิจารณาข้อคำถาม (Items) เทียบกับนิยามศัพท์เฉพาะ โดยพิจารณาตัดสินเป็นรายข้อไป โดยในแบบประเมินนั้นจะระบุความหมายของความคิดเห็นจากผู้เชี่ยวชาญ ดังนี้ คือ

+1 หมายถึง ผู้เชี่ยวชาญเห็นว่า ข้อคำถามข้อนั้นสามารถ วัดได้ตรง กับนิยามศัพท์เฉพาะ

0 หมายถึง ผู้เชี่ยวชาญ ไม่แน่ใจ ว่าข้อคำถามข้อนั้นสามารถ วัดได้ตรง กับนิยามศัพท์เฉพาะ (อาจมีการปรับปรุง แก้ไข ตามผู้เชี่ยวชาญเสนอแนะ)

-1 หมายถึง ผู้เชี่ยวชาญเห็นว่า ข้อคำถามข้อนั้นสามารถ วัดได้ไม่ตรง กับนิยามศัพท์เฉพาะ (ผู้เชี่ยวชาญอาจเสนอแนะให้ตัดทิ้ง หรือสร้างข้อคำถามใหม่แทน)

หลังจากที่ผู้เชี่ยวชาญตรวจสอบแล้วนำมาหาค่าความสอดคล้องของข้อคำถามแต่ละข้อกับนิยามศัพท์เฉพาะ (ที่นักวิจัยได้ตรวจสอบจากเอกสารและงานวิจัยที่เกี่ยวข้อง) ตามความคิดเห็นของผู้เชี่ยวชาญ โดยใช้สูตรดังต่อไปนี้

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

เมื่อ IOC แทน Index of Consistency
 $\sum R$ แทน ผลรวมความคิดเห็นของผู้เชี่ยวชาญ
 N แทน จำนวนผู้เชี่ยวชาญ

โดยค่าที่ได้จะมีค่าตั้งแต่ (-1) ถึง (+1) ผู้วิจัยเลือกข้อที่มีค่าตั้งแต่ 0.50 – 1.00 หลังจากนั้นนำมาหาค่าเฉลี่ย ทั้งฉบับ จะได้ค่าความตรงเชิงเนื้อหาของแบบวัดที่มีลักษณะเป็นแบบมาตราส่วนประมาณค่า

ส่วนด้านความตรงเชิงโครงสร้างมีวิธีการตรวจสอบได้หลายวิธี เช่น การตรวจเชิงเหตุผล (Logical) โดยผู้เชี่ยวชาญ การตรวจความสอดคล้องหรือความเท่าเทียมกัน (Equivalence) โดยตรวจสอบจากค่า r_{xy} ระหว่างคะแนนที่ได้จากการตอบเครื่องมือที่สร้างขึ้นกับคะแนนที่ได้จากเครื่องมือที่เป็นมาตรฐานวัดในสิ่งเดียวกัน การตรวจหาความสัมพันธ์กับเกณฑ์ที่มีโครงสร้างเหมือนกัน การตรวจสอบด้วยการวิเคราะห์องค์ประกอบ (Factor Analysis) เพื่อพิสูจน์หรือยืนยันว่าเครื่องมือเหล่านี้ได้สร้างตามองค์ประกอบต่าง ๆ ตามทฤษฎี การตรวจสอบด้วยเทคนิคกลุ่มรู้ชัด (Known-Group Technique) โดยการเทียบกับกลุ่มที่รู้แล้วว่ามีสิ่งเหล่านั้นอยู่ ด้วยการทดสอบค่าสถิติที (t - test) และการตรวจโดยใช้เมตริกซ์ลักษณะหลากหลาย (MTMMM)

หลังจากที่ดำเนินการคัดเลือกข้อคำถามตามเกณฑ์ความตรงแล้ว ก็ดำเนินการปรับปรุงและจัดทำต้นฉบับ เพื่อนำไปทดลองใช้ (Try Out) เพื่อหาคุณภาพด้านอำนาจจำแนกต่อไป

2. การหาค่าอำนาจจำแนก

เครื่องมือรวบรวมข้อมูลที่มีลักษณะเป็นแบบมาตราส่วนประมาณค่า มีวิธีการหาคุณภาพด้านอำนาจจำแนก 2 วิธี คือ

2.1 วิธีการกลุ่มรู้ชัด (Known Group Technique) วิธีการนี้ตรงตามความหมายของอำนาจจำแนกที่ว่า ความสามารถของแบบวัดที่จำแนกคนออกเป็น 2 กลุ่มคือ กลุ่มที่มีคุณลักษณะที่ต้องการวัดสูงและต่ำ หรือบวกและลบ หรือมากและน้อย วิธีการตรวจสอบ ด้วยการนำแบบวัดไปทดลองใช้ (Try Out) กับประชากรที่ไม่ใช่กลุ่มตัวอย่างจำนวนที่มากพอ ประมาณ 100 คน เนื่องจากข้อตกลงเบื้องต้น (Assumption) ของการใช้สถิติที (t) คือกลุ่มตัวอย่างต้องมาจากประชากรที่มีการแจกแจงเป็นโค้งปกติ เพื่อแบ่งเป็น 3 กลุ่มคือกลุ่มสูง (30) กลุ่มกลาง (40) และกลุ่มต่ำ (30) แต่ตัดกลุ่มกลางออก เนื่องจากอาจจะเกิดการสลับเปลี่ยนตำแหน่งระหว่างคนในกลุ่มสูงและกลุ่มต่ำหากได้คะแนนเท่ากัน นักวิจัยต้องคัดเลือก 2 กลุ่มที่แตกต่างกันจริงๆ คือ กลุ่มสูงและกลุ่มต่ำ (ที่มีคุณลักษณะตามที่ต้องการ) แล้วทำการทดสอบความแตกต่างด้วยสถิติทดสอบทีกรณีกกลุ่มตัวอย่างเป็นอิสระต่อกัน (Independent Samples t - test)

วิธีการนี้นักวิจัยต้องเรียงคะแนนรวมของผู้ตอบแบบสอบถามทั้งหมดจากคะแนนสูงสุดลงมาต่ำสุด แล้วใช้เทคนิค 25% เป็นกลุ่มสูง และ 25% เป็นกลุ่มต่ำ (หรือนักวิจัยพิจารณาให้ได้คนกลุ่มละ 30 คนเป็นอย่างต่ำ คัดกลุ่มกลางออก) จึงคำนวณโดยใช้สูตร ดังนี้

$$\text{ใช้สูตร } t = \frac{(\bar{X}_H - \bar{X}_L) - (\mu_H - \mu_L)}{\sqrt{S_p^2 \left(\frac{1}{n_H} + \frac{1}{n_L} \right)}} \quad df = (n_H - 1) + (n_L - 1)$$

$$\text{เมื่อ } S_p^2 = \frac{(n_H - 1)S_H^2 + (n_L - 1)S_L^2}{(n_H - 1) + (n_L - 1)}$$

ตัวอย่างที่ 1 จากการนำแบบวัดที่เป็นมาตราส่วนประมาณค่าไปทดสอบกับกลุ่มตัวอย่างจำนวน 100 คน
จึงแสดงวิธีการหาค่าอำนาจจำแนกเฉพาะข้อที่ 8

วิธีทำ เมื่อนำเอาคำตอบข้อที่ 8 ของกลุ่มสูง 30 คน กับกลุ่มต่ำ 30 คน มาเขียนแจกแจงความถี่
และหาค่าพื้นฐานเพื่อนำไปคำนวณหาค่าเฉลี่ยและความแปรปรวนดังนี้

คำตอบข้อที่ 1	กลุ่มสูง				กลุ่มต่ำ			
	คะแนน (X)	ความถี่ (f)	fX	fX ²	คะแนน (X)	ความถี่ (f)	fX	fX ²
เห็นด้วยอย่างยิ่ง	5	10	50	250	5	2	10	50
เห็นด้วย	4	12	48	192	4	4	16	64
ไม่แน่ใจ	3	1	3	9	3	6	18	54
ไม่เห็นด้วย	2	3	6	12	2	10	20	40
ไม่เห็นด้วยอย่างยิ่ง	1	4	4	4	1	8	8	8
รวม		30	111	467		30	72	216

การคำนวณหาค่าเฉลี่ยของคะแนนกลุ่มสูงและกลุ่มต่ำจากสูตร $\bar{X} = \frac{\sum fX}{N}$

$$\text{ค่าเฉลี่ยกลุ่มสูง } \bar{X}_H = \frac{111}{30} = 3.70 \quad \text{ค่าเฉลี่ยกลุ่มต่ำ } \bar{X}_L = \frac{72}{30} = 2.40$$

การคำนวณหาความแปรปรวนของคะแนนกลุ่มสูงและกลุ่มต่ำจากสูตร

$$S^2 = \frac{n\sum fX^2 - (\sum fX)^2}{n(n-1)}$$

$$\text{ความแปรปรวนของคะแนนกลุ่มสูง } S_H^2 = \frac{(30)(467) - (111)^2}{30(29)} = \frac{1689}{870} = 1.94$$

$$\text{ความแปรปรวนของคะแนนกลุ่มต่ำ } S_L^2 = \frac{(30)(216) - (72)^2}{30(29)} = \frac{1296}{870} = 1.49$$

คำนวณหาค่า t จากสูตร

$$t = \frac{(\bar{X}_H - \bar{X}_L) - (\mu_H - \mu_L)}{\sqrt{S_p^2 \left(\frac{1}{n_H} + \frac{1}{n_L} \right)}} \quad df = (n_H - 1) + (n_L - 1)$$

$$\text{เมื่อ } S_p^2 = \frac{(n_H - 1)S_H^2 + (n_L - 1)S_L^2}{(n_H - 1) + (n_L - 1)}$$

$$\text{แทนค่า } S_p^2 = \frac{(29)1.94 + (29)1.49}{(30 - 1) + (30 - 1)} = \frac{13.05}{58} = .225$$

$$t = \frac{(3.70 - 2.40)}{\sqrt{.225 \left(\frac{1}{30} + \frac{1}{30} \right)}} = \frac{1.3}{\sqrt{.015}} = \frac{1.3}{.112} = 10.66$$

เปิดตาราง t ที่ $\alpha .05$ (ทางเดียว), $df = (n_H - 1) + (n_L - 1) = 60 - 2 = 58$ มีค่า t (ตาราง) = 1.684

$\therefore$ ค่า t คำนวณมากกว่าค่า t ในตารางจึงปฏิเสธ H_0 ยอมรับ H_1

สรุป แบบสอบถามข้อที่ 8 มีอำนาจจำแนกจึงคิดไว้ใช้ (เพราะมีนัยสำคัญทางสถิติ ที่ระดับ .05)

นอกจากนี้นักวิจัยอาจจะใช้คอมพิวเตอร์ช่วยในการคำนวณ เนื่องจากสะดวกและเกิดความผิดพลาดน้อย ซึ่งจะยกตัวอย่างการคำนวณด้วยโปรแกรมคอมพิวเตอร์ ดังตัวอย่างสมมติจากการทดลองใช้ (Try Out) แบบวัดความมั่นใจในตนเอง 10 ข้อ กับกลุ่มตัวอย่าง 100 คน เรียงคะแนนแล้วแบ่งเป็นกลุ่มสูง 30 คน กลุ่มกลาง 40 คน กลุ่มต่ำ 30 คน หลังจากนั้นนำคะแนนของกลุ่มสูงและต่ำมาเปรียบเทียบได้ผลดังตารางต่อไปนี้

Independent Samples Test

ข้อคำถาม	t-test for Equality of Means			
	t	df	Sig. (2-tailed)	
A1	13.958	58	.000	
A2	14.888	58	.000	
A3	12.722	58	.000	
A4	10.115	58	.000	

ข้อคำถาม	t-test for Equality of Means			
	t	df	Sig. (2-tailed)	
A5	13.638	58	.000	
A6	.100	58	.921	
A7	-.682	58	.499	
A8	-2.263	58	.027	
A9	9.981	58	.000	
A10	10.687	58	.000	

การพิจารณาอำนาจจำแนกให้นำค่า sig (Two-Tailed) ของแต่ละข้อไปเปรียบเทียบกับค่า α ที่กำหนด ถ้าค่า Sig (Two-Tailed) มีค่าน้อยกว่าหรือเท่ากับ α (Sig. (Two-Tailed) $\leq \alpha$) แสดงว่าข้อคำถามข้อนั้นมีอำนาจจำแนก แต่ถ้าค่า Sig (Two-Tailed) มีค่ามากกว่า α ที่กำหนดแสดงว่าข้อคำถามข้อนั้นไม่มีอำนาจจำแนก เช่นสมมุติว่ากำหนด $\alpha = .01$ จากตารางพบว่า ข้อ A6 A7 และ A8 มีค่า Sig (Two-Tailed) มากกว่าค่า α ที่กำหนด (.921, .499 และ .027) แสดงว่าข้อคำถามทั้ง 3 ไม่สามารถจำแนกระหว่างคนกลุ่มสูงและคนกลุ่มต่ำได้ แสดงว่าไม่มีอำนาจจำแนก หรือนักวิจัยอาจจะนำค่าสถิติ (t) ที่คำนวณได้ ไปเปรียบเทียบกับค่าวิกฤตในตาราง ที่ (Critical Value of T -Statistics) ที่ $\alpha .01$ (Two Tailed) df = 58 มีค่า t ในตารางเท่ากับ 3.551

2.2 วิธีการหาความสัมพันธ์หรือการหาความสอดคล้องระหว่างคะแนนรายข้อ (X_i) กับคะแนนรวมทั้งฉบับ (Y) ที่ตัดคะแนนข้อที่ X_i ออกแล้ว หรือ Item-total Correlation แล้วตรวจสอบนัยสำคัญด้วยสถิติสหสัมพันธ์อย่างง่าย (r_{xy}) โดยวิธี Item-total Correlation วิธีการนี้ประยุกต์จากหลักการที่ว่าข้อคำถาม (Items) แต่ละข้อในแบบวัดหรือแบบสอบถามควรจะวัดเรื่องเดียวกันหรือมีความสอดคล้องกัน (เพราะสร้างจากทฤษฎี หรือนิยามศัพท์เฉพาะที่ได้ตรวจสอบจากเอกสารที่เกี่ยวข้องแล้ว) ซึ่งถ้าหากคะแนนจากการตอบข้อคำถามในแต่ละข้อมีความสัมพันธ์ทางบวกกับคะแนนรวมอย่างมีนัยสำคัญ แสดงว่าข้อคำถามนั้นมีอำนาจจำแนก และในทำนองเดียวกันหากคะแนนจากการตอบข้อคำถามข้อใดไม่มีความสัมพันธ์ หรือมีความสัมพันธ์ทางลบกับคะแนนรวมก็แสดงว่าข้อคำถามข้อนั้น ไม่มีอำนาจจำแนก การคำนวณจะใช้สูตรสหสัมพันธ์อย่างง่ายของเพียร์สัน ดังนี้

$$\text{สูตร} \quad r_{xy} = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$$

ตัวอย่างที่ 2 จากข้อมูลต่อไปนี้จงหาอำนาจจำแนกของแบบสอบถามข้อที่ 3 ($\alpha = .05$)

ผู้ตอบ คนที่	X (คะแนนข้อที่ 3)	Y (คะแนนรวมโดยตัด คะแนนข้อ 3 ออกแล้ว)	X ²	Y ²	XY
1	4	40	16	1,600	160
2	1	30	1	900	30
3	4	42	16	1,764	168
4	3	38	9	1,444	114
5	2	36	4	1,296	72
6	5	40	25	1,600	200
7	2	37	4	1,369	74
8	4	35	16	1,225	140
9	3	41	9	1,681	123
10	5	45	25	2,025	225
ผลรวม	33	384	125	14,904	1,306

$$\begin{aligned}
 r_{xy} &= \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}} \\
 &= \frac{10(1,306) - (33)(384)}{\sqrt{[10(125) - 1,089][10(14,904) - 147,456]}} \\
 &= \frac{388}{\sqrt{(161)(1,584)}} = \frac{388}{504.99} = 0.7638
 \end{aligned}$$

เปิดตาราง r_{xy} ที่ $\alpha = .05$ (ทางเดียว) $df = N - 2 = 10 - 2 = 8$, มีค่า = 0.6319 ค่า r_{xy} คำนวณ มากกว่า r_{xy} ตาราง แสดงว่าการทดสอบมีนัยสำคัญ

สรุป แบบสอบถามข้อที่ 3 มีอำนาจจำแนกจึงตัดไว้ใช้ (เพราะมีความสัมพันธ์ทางบวกกับคะแนนรวม อย่างมีนัยสำคัญทางสถิติที่ระดับ .05)

นอกจากนี้นักวิจัยอาจจะใช้คอมพิวเตอร์ช่วยในการการคำนวณ เนื่องจากสะดวกและเกิดความผิดพลาดน้อย ซึ่งจะยกตัวอย่างการคำนวณด้วยโปรแกรมคอมพิวเตอร์ ดังตัวอย่างสมมุติจากการทดลองใช้ (Try out) แบบวัดความมั่นใจในตนเอง 10 ข้อ กับกลุ่มตัวอย่าง 100 คน ได้ผลดังนี้

RELIABILITY ANALYSIS - SCALE (ALPHA)

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item- Total Correlation	Alpha if Item Deleted
A1	23.8200	34.1491	.6914	.6980
A2	23.9100	35.0524	.6863	.7015
A3	24.1400	35.8792	.6835	.7047
A4	23.8900	36.6645	.6208	.7137
A5	23.9600	34.5236	.7307	.6946
A6	23.4900	45.5656	-.0060	.7972
A7	23.4300	46.7930	-.0775	.8051
A8	23.6000	49.1313	-.2171	.8132
A9	24.2300	36.5627	.6579	.7097
A10	24.1200	36.4905	.6458	.7106

Reliability Coefficients N of Cases = 100.0 N of Items = 10 Alpha = .7609

จากผลการหาค่าอำนาจจำแนก พบว่าข้อคำถามมีค่าความสัมพันธ์เป็นบวกอยู่ 7 ข้อ (ใน column ของ Corrected Item-Total Correlation) จะบอกว่ามีอำนาจจำแนกยังไม่ได้ นักวิจัยจะต้องนำค่าความสัมพันธ์ทั้ง 7 ข้อไปเทียบกับค่าวิกฤติ (Critical Value) จากตารางที่ $\alpha = .05$ (หางเดียว) $df = N-2 = 100-2 = 98$ ซึ่งมีค่าเท่ากับ 0.167 ซึ่งค่าความสัมพันธ์ทั้ง 7 ข้อ มีค่าสูงกว่าค่าวิกฤติจากตารางแสดงว่ามีอำนาจจำแนกทั้ง 7 ข้อ ส่วน 3 ข้อที่มีค่าความสัมพันธ์ติดลบ แสดงว่าไม่มีอำนาจจำแนก (ตัดทิ้ง) ซึ่งสอดคล้องกับวิธีการหาโดยวิธีกลุ่มรู้ชุด ที่ทดสอบความแตกต่างจากค่าสถิติที (t-test)

ข้อเสนอแนะในการตัดสินใจเลือกวิธีการคำนวณค่าอำนาจจำแนก ควรพิจารณาที่จำนวนกลุ่มที่ทดลองใช้ (Try Out) ถ้ากลุ่มที่ทดลองหาคุณภาพมีจำนวนมากพอ (สามารถแบ่งกลุ่มสูง และกลุ่มต่ำได้อย่างน้อยกลุ่มละ 30 คน) ควรใช้วิธีการกลุ่มรู้ชุด (Known-Group Technique) เพราะตรงกับนิยามและความหมายของอำนาจจำแนก แต่ถ้ากลุ่มที่ทดลองหาคุณภาพมีจำนวนน้อย (ไม่ต่ำกว่า 30 คน) ก็สามารถประยุกต์ใช้วิธีการหาความสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวม (Item-total Correlation) ได้

3. การหาค่าความเที่ยงหรือความเชื่อมั่น

มีวิธีหาความเที่ยง (Reliability) ของแบบวัดทางจิตวิทยาและแบบสอบถามที่มีลักษณะเป็นมาตราส่วนประมาณค่าหลายวิธี แต่วิธีที่เป็นที่นิยมและสะดวกในการคำนวณคือวิธีของครอนบาค (Cronbach

Method) ซึ่ง เสนอวิธีการหาค่าความเที่ยงที่สามารถใช้กับเครื่องมือที่ตรวจให้คะแนนไม่ใช่ 1 กับ 0 โดยดัดแปลงจากสูตร KR₂₀ (ถ้าการตรวจให้คะแนนแบบ 0 กับ 1 แล้วหาค่าความเที่ยงตามวิธีของครอนบาค จะได้ค่าความเที่ยงเท่ากับการหาด้วยวิธี KR₂₀ ทุกประการ) ค่าความเที่ยงที่หาตามวิธีของ Cronbach เรียกว่า “สัมประสิทธิ์แอลฟา” (α - Coefficient) เหมาะสำหรับเครื่องมือที่มีลักษณะเป็นแบบมาตราส่วนประมาณค่า มีสูตร ดังนี้

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

ตัวอย่างที่ 3 แบบวัดความเครียดในการเรียน มี 10 ข้อ เป็นแบบมาตราส่วนประมาณค่า แต่ละข้อมีระดับความเห็นสำหรับผู้เรียนตอบแสดงความคิดเห็น 5 ระดับไปใช้กับนิสิตระดับปริญญาโท 10 คน แล้วทำการตรวจให้คะแนน แต่ละข้อ จนครบทุกคนนำคะแนนมาบรรจุลงในตารางได้ดังนี้ จงหาค่าความเที่ยงของแบบวัดความเครียดในการเรียนฉบับนี้

ข้อ ผู้เรียน	1	2	3	4	5	6	7	8	9	10	X	X ²
1	3	3	4	3	4	4	4	5	5	3	38	1,444
2	5	5	5	5	4	4	5	4	4	4	45	2,025
3	4	4	3	4	4	4	4	3	4	4	38	1,444
4	3	4	4	3	3	4	4	5	4	4	38	1,444
5	3	4	3	4	3	4	4	3	4	3	35	1,225
6	4	4	3	5	4	5	5	5	4	4	43	1,849
7	3	4	4	5	5	4	5	4	4	4	42	1,764
8	3	4	4	4	5	3	4	5	4	3	39	1,521
9	3	3	4	3	4	4	4	5	5	3	38	1,444
10	4	4	3	4	4	4	4	3	4	4	38	1,444
$\sum X_i$	35	39	37	40	40	40	43	42	42	36	394	15,604
$\sum X_i^2$	127	155	141	166	164	162	187	184	178	132	($\sum X$)	($\sum X^2$)
S_i^2	0.50	0.32	0.46	0.67	0.44	0.22	0.23	0.84	0.18	0.27		

$$S^2 = \frac{n \sum X^2 - (\sum X)^2}{n(n-1)}$$

$$S_i^2 = \frac{10(15,604) - (394)^2}{10(9)} = \frac{804}{90} = 8.93$$

$$S_1^2 = \frac{10(127) - (35)^2}{10(9)} = \frac{45}{90} = 0.50 \quad (\text{ข้อที่ 1})$$

รวมความแปรปรวนของทุกข้อเข้าด้วยกัน เป็นค่า S_i^2 ได้ดังนี้

$$\sum S_i^2 = 0.50 + 0.32 + 0.46 + 0.67 + 0.44 + 0.22 + 0.23 + 0.84 + 0.18 + 0.27 = 4.13$$

$$\begin{aligned} \text{คำนวณค่า } \alpha \text{ จากสูตร } \alpha &= \frac{k}{k-1} \left(1 - \frac{\sum S_i^2}{S_t^2} \right) \\ &= \frac{10}{10-1} \left(1 - \frac{4.13}{8.93} \right) \\ &= 1.111 \times .538 \\ &= .597 \end{aligned}$$

แบบวัดความเครียดในการเรียนฉบับนี้ มีค่าความเที่ยง .597

นักวิจัยสามารถหาค่าความเที่ยงจากโปรแกรมคอมพิวเตอร์ได้ จากตัวอย่างที่ 1-3 เรื่องค่าอำนาจจำแนกและค่าความเที่ยง คือส่วนค่า Reliability Coefficients N of Cases = 100 N of Items = 10 Alpha = .7609 แสดงว่า ค่าความเที่ยงจากแบบวัด 10 ข้อ มีค่าความเที่ยงเท่ากับ .7609 แต่จากตัวอย่างดังกล่าวมีข้อคำถาม 3 ข้อที่ไม่มีอำนาจจำแนก โปรแกรมจะแนะนำคือ ใน Column ของ Alpha if Item Deleted แสดงค่าความเที่ยง (Alpha) ของแบบวัดทั้งฉบับเมื่อตัดข้อคำถามข้อนั้นออก เช่น ถ้าตัดข้อคำถามข้อที่ A8 แล้วค่าความเที่ยงจะเป็น .8132 (จากข้อคำถาม 9 ข้อ ซึ่งแต่เดิมข้อคำถาม 10 ข้อ ค่าความเที่ยงเท่ากับ .7609) และเมื่อตัดข้อคำถามที่ไม่มีอำนาจจำแนกทั้ง 3 ข้อออกไป เหลือข้อคำถาม 7 ข้อ แล้วคำนวณค่าความเที่ยง (Alpha) อีกครั้ง จะได้ค่าความเที่ยงทั้งฉบับเท่ากับ .9343 ดังนี้

RELIABILITY ANALYSIS - SCALE (ALPHA)

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item- Total Correlation	Alpha if Item Deleted
A1	14.8100	38.1757	.7773	.9258
A2	14.9000	38.7576	.8041	.9226
A3	15.1300	39.8112	.7938	.9237
A4	14.8800	40.3693	.7504	.9275
A5	14.9500	38.3712	.8367	.9194
A9	15.2200	40.5168	.7720	.9257
A10	15.1100	40.0787	.7857	.9244

Reliability Coefficients N of Cases = 100.0 N of Items = 7 Alpha = .9343

ดังนั้นในการเลือกเครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลในการวิจัย กรณีที่ข้อมูลในการวิจัยเป็นข้อเท็จจริง ความคิดเห็นส่วนตัว หรือคุณลักษณะทางจิตวิทยา นักวิจัยต้องเลือกประเภทเครื่องมือให้ตรงกับกลุ่มตัวอย่างที่จะให้ข้อมูล และให้ตรงกับลักษณะข้อมูลที่จะนำมาตอบจุดประสงค์ของการวิจัย รวมถึงขั้นตอนที่สำคัญคือการสร้างและการหาคุณภาพของเครื่องมือในการวิจัย เพราะจะเป็นการทำให้ผู้วิจัยเชื่อมั่นได้ว่าข้อมูลที่จะนำมาวิเคราะห์นั้นมีคุณภาพ อันจะส่งผลให้งานวิจัยนั้นมีคุณภาพและคุณค่า

เอกสารอ้างอิง

- ศิริชัย กาญจนวาสี. (2552). **ทฤษฎีการทดสอบแบบดั้งเดิม**. พิมพ์ครั้งที่ 6 กรุงเทพฯ : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สมบัติ ท้ายเรือคำ. (2555). **ระเบียบวิธีวิจัยสำหรับมนุษยศาสตร์และสังคมศาสตร์**. พิมพ์ครั้งที่ 5 มหาสารคาม : สำนักพิมพ์มหาวิทยาลัยมหาสารคาม.
- Aiken, Lewis R. (1991). **Psychological Testing and Assessment**. 7th ed. Boston : Allyn and Bacon.
- Kerlinger, Fred N. (1986). **Foundations of Behavioral Research**. 3rd ed. Winston, Inc.