

เครื่องมือในการประเมิน

หลังจากที่ผู้ประเมินได้กำหนดประเด็นการประเมิน และออกแบบ พัฒนาตัวชี้วัดและเกณฑ์การประเมินแล้ว งานขั้นต่อไป คือ การสร้างเครื่องมือในการเก็บรวบรวมข้อมูล ในการสร้างเครื่องมือสำหรับการประเมิน ผู้ประเมินจะต้องทำความเข้าใจถึงระดับการวัดของข้อมูลที่ต้องการเสียก่อน จากนั้นจึงเลือกชนิดของเครื่องมือที่จะใช้ในการเก็บรวบรวมข้อมูล ลงมือสร้างหรือพัฒนาเครื่องมือ และตรวจสอบคุณภาพเครื่องมือให้ได้มาตรฐาน มีความน่าเชื่อถือได้ทั้งในด้านความเที่ยงและความตรง ก่อนที่จะนำไปใช้จริง

1. ระดับการวัดของข้อมูล/ตัวแปร/ตัวชี้วัด

ก่อนที่จะสร้างเครื่องมือในการเก็บรวบรวมข้อมูลที่มีคุณภาพ นักประเมินต้องพิจารณาก่อนว่า ข้อมูลหรือตัวชี้วัดอยู่ในระดับการวัดหรือมาตรวัด (Scale of Measurement) ใดใน 4 ระดับ ดังต่อไปนี้

1) **มาตรานามบัญญัติหรือจัดกลุ่ม (nominal scale)** เป็นระดับการวัดตัวแปรเบื้องต้นที่สุด ซึ่งเป็นการจำแนกค่าตัวแปรออกเป็นกลุ่มหรือประเภท เช่น ตัวแปรเพศ มีค่าตัวแปร 2 ค่า คือ ชายและหญิงแยกออกจากกันอย่างเด็ดขาด ค่าของตัวแปรในระดับนี้จะไม่มีความหมายในเชิงปริมาณ แม้ว่าจะกำหนดตัวเลขให้กับแต่ละค่าของตัวแปรก็ตาม เช่น ชายกำหนดเป็นตัวเลข 1 หญิงกำหนดเป็นตัวเลข 2 ตัวเลข 1 และ 2 ดังกล่าว ก็มีได้แสดงความหมายในเชิงปริมาณที่จะมาเปรียบเทียบกันได้แต่อย่างใด เป็นเพียงการกำหนดรหัสตัวเลขให้กับค่าของตัวแปรเท่านั้น ตัวอย่างของตัวแปรที่วัดค่าได้ในระดับนี้ได้แก่ ภูมิลำเนา อาชีพ ศาสนา เชื้อชาติ ภาษา เป็นต้น

2) **มาตรจัดอันดับ (ordinal scale)** เป็นระดับการวัดตัวแปรที่ยังคงจำแนกค่าของตัวแปรออกเป็นกลุ่มๆ อยู่ แต่สามารถบอกลำดับหรือความมากน้อยของแต่ละกลุ่มได้ เช่น ตัวแปรคุณวุฒิ เราอาจวัดค่าตัวแปรได้เป็น 3 กลุ่ม คือ ระดับประถมศึกษา มัธยมศึกษา และอุดมศึกษา หรือตัวแปรผลการประกวดคัดลายมือ ที่อาจมีค่าตัวแปร 3 ค่า คือ ชนะเลิศ รองชนะเลิศ และชมเชย ซึ่งแม้ว่าเราจะสามารถเรียงลำดับของค่าตัวแปรได้ แต่เราก็ไม่สามารถบอกขนาดของความแตกต่างระหว่างค่าตัวแปรแต่ละตัวได้ และขนาดความแตกต่างระหว่างค่าตัวแปรแต่ละระดับก็อาจไม่เท่ากันก็ได้ ตัวอย่างของตัวแปรที่วัดได้ในระดับนี้ได้แก่ ฐานะทางเศรษฐกิจและสังคม (สูง กลาง ต่ำ) เชวรณ์ปัญญา (สูง กลาง ต่ำ) ตำแหน่ง (อาจารย์ 1 อาจารย์ 2 อาจารย์ 3) เป็นต้น

3) **มาตรอันดับหรือช่วงเท่า (interval scale)** ตัวแปรที่จะวัดค่าได้ในมาตรอันดับหรือช่วงเท่านี้ จะต้องเป็นตัวแปรที่มีช่วงห่างระหว่างค่าของตัวแปรเท่ากันตลอด เช่น ตัวแปรคะแนนสอบ เราอาจกำหนดให้มีค่าตัวแปรตั้งแต่ 0 คะแนน ถึง 100 คะแนน ความแตกต่างระหว่างคะแนน 60 กับคะแนน 40 มีค่าเท่ากับความแตกต่างระหว่างคะแนน 40 กับคะแนน 20 เรียกได้ว่าแต่ละค่าของตัวแปรมีช่วงห่างเท่ากันเสมอ แต่ค่าของตัวแปรที่วัดในระดับช่วงเท่านี้ ไม่ได้เริ่มต้นที่ศูนย์แท้ เช่น ถ้าสอบได้คะแนนเท่ากับ 0 ก็ไม่ได้หมายความว่าผู้นั้นไม่ได้มีความรู้เลย เช่นเดียวกับตัวแปรอุณหภูมิ ค่าของ

ตัวแปรที่ 0 องศาเซลเซียส ก็ได้แปลว่าไม่มีความร้อนอยู่เลย แต่มีความร้อนในระดับหนึ่งที่เรา กำหนดค่าว่าเป็น 0 องศาเซลเซียส

4) มาตรการอัตราส่วน (ratio scale) ตัวแปรที่วัดค่าได้ในมาตรการอัตราส่วนจะมีคุณสมบัติ เหมือนกับการวัดในมาตรอันตรภาคช่วงเท่า แต่มีคุณสมบัติพิเศษที่เพิ่มมา คือ การมีค่าศูนย์แท้ (absolute zero) ถ้าค่าตัวแปรนั้นเท่ากับ 0 ก็แปลว่าไม่มีค่าจริงๆ เช่น ความสูง 0 เมตร แสดงว่าไม่มีความสูง รายได้ 0 บาท แสดงว่าไม่มีรายได้จริงๆ ตัวแปรที่วัดค่าได้ในระดับนี้ ได้แก่ น้ำหนัก อายุ ความยาว เวลา ความเร็ว เป็นต้น

มาตรการวัดตัวแปรนี้จะเชื่อมโยงกับการออกแบบและการวิเคราะห์ข้อมูลเสมอ เช่น ถ้า ผู้วิจัยต้องการวิเคราะห์ข้อมูลด้วยการวิเคราะห์การถดถอย (multiple linear regression) ก็ต้องคำนึงถึง ข้อตกลงเบื้องต้นของการวิเคราะห์ถดถอยว่าตัวแปรอิสระควรจะมีค่าที่ต่อเนื่อง คือ สามารถวัดค่าตัวแปรได้ในมาตรอันตรภาคขั้นขึ้นไป หรือหากเป็นตัวแปรที่วัดค่าได้ในมาตรนามบัญญัติก็จะเป็นตัวแปรหุ่น (dummy variable) สมมติว่าในกรณีนี้ ตัวแปรรายได้ควรก็ควรวัดค่าในมาตรการอัตราส่วน ถ้าหากเก็บ ข้อมูลโดยใช้แบบสอบถามโดยตรง แทนที่จะเลือกใช้ตอบเป็นช่วง เช่น () 0-500 บาท () 5001-10000 บาท เป็นต้น เพราะค่าตัวแปรรายได้ที่ได้จากการให้ตอบเป็นช่วงรายได้นี้ จะอยู่ในมาตรจัดอันดับ (ordinal scale) เท่านั้น

อนึ่ง หากเราวัดค่าตัวแปรที่ได้ในมาตรอันตรภาค (interval scale) หรือมาตรการอัตราส่วน (ratio scale) แล้ว เราสามารถที่จะจัดกระทำให้ค่าตัวแปรลดความละเอียดมาอยู่ในมาตรจัดอันดับ และมาตรนามบัญญัติได้ เช่น ข้อมูลรายได้ที่เราได้จำนวนเงินต่อเดือนมาโดยตรง เราก็สามารถนำมาจัดกลุ่มให้เป็นรายได้สูง กลาง ต่ำ ได้ แต่ในทางตรงกันข้าม เราไม่สามารถจัดกระทำค่าของตัวแปรที่วัดใน มาตรนามบัญญัติจัดกลุ่มหรือจัดอันดับ ให้ไปมีค่าอยู่ในมาตรอันตรภาคหรือมาตรการอัตราส่วนได้เลย ดังนั้น ผู้ประเมินจึงต้องวางแผนการวัดค่าตัวแปรตามมาตรการวัดที่กำหนดไว้ให้รอบคอบเสียแต่แรก

ตัวอย่าง การวัดตัวแปรหรือตัวชี้วัด “การศึกษา” ของบุคคล

- ท่านจบการศึกษาระดับปริญญาตรีหรือไม่
 - ไม่จบระดับปริญญาตรี
 - จบระดับปริญญาตรี (NOMINAL SCALE)
- ท่านจบการศึกษาระดับใด
 - ประถมศึกษา
 - มัธยมศึกษา
 - อนุปริญญาหรือเทียบเท่า
 - ปริญญาตรี
 - สูงกว่าปริญญาตรี (ORDINAL SCALE)
- หลังจากจบการศึกษาระดับ (9 ปี) ท่านใช้เวลาในการศึกษาในสถานศึกษาอีกจำนวน ปี (INTERVAL SCALE)
- ท่านใช้เวลาในการศึกษาทั้งหมดจำนวน ปี (RATIO SCALE)

2. การสร้างเครื่องมือเก็บรวบรวมข้อมูล

เมื่อกำหนดระดับการวัดค่าตัวแปรได้อย่างเหมาะสมแล้ว ขั้นตอนต่อไปผู้ประเมินต้องพิจารณาว่าจะวัดค่าตัวชี้วัดหรือตัวแปรที่ต้องการได้อย่างไร ด้วยเครื่องมือชนิดใด โดยทั่วไปแล้ววิธีการวัดค่าตัวแปรเราอาจแบ่งได้เป็น 3 วิธีใหญ่ๆ ได้แก่ การสอบ การสอบถาม และการสังเกต

วิธีการวัดค่าตัวแปรวิธีแรก คือ **การสอบ** ซึ่งเป็นการวัดที่กำหนดเงื่อนไขหรือสถานการณ์ให้ผู้ถูกวัดแสดงความสามารถสูงสุด (maximum performance) ของตนออกมา โดยที่ผู้ถูกวัดรู้ตัวว่ากำลังถูกวัด และรู้ว่าถูกวัดความสามารถในเรื่องใด โดยมากจะเป็นตัวแปรที่เกี่ยวกับความสามารถทางสมอง เช่น ผลสัมฤทธิ์จากการฝึกอบรม ความถนัด ความคิดสร้างสรรค์ ความคิดวิเคราะห์ เป็นต้น การวัดค่าตัวแปรด้วยการสอบ อาจทำได้ 2 ลักษณะ คือ การสอบโดยใช้กระดาษ-ดินสอ (paper-pencil test) และการสอบปากเปล่า (oral test) เครื่องมือที่ใช้จะเป็นพวกแบบสอบ แบบวัดต่างๆ

วิธีการที่สอง คือ **การสอบถาม** ซึ่งแตกต่างไปจาก “การสอบ” ตรงที่การสอบถามเป็นการกำหนดเงื่อนไข หรือสถานการณ์ให้ผู้ถูกวัดแสดงคุณลักษณะเฉพาะตัว (typical performance) หรือความเป็นจริงของตนออกมา โดยไม่มีการตัดสินว่าสิ่งที่ผู้ถูกวัดตอบหรือแสดงออกมานั้นถูกหรือผิด ตัวแปรที่วัดได้ด้วยวิธีนี้จะเป็นตัวแปรเกี่ยวกับความคิดจิตใจ เช่น ความสนใจ ความคิดเห็น บุคลิกภาพทัศนคติ เป็นต้น หรือตัวแปรเกี่ยวกับข้อเท็จจริงต่างๆ ของผู้ถูกวัด อาทิ เพศ อายุ การศึกษา ภูมิลำเนา

รายได้ เป็นต้น วิธีสอบถามนี้อาจทำได้ 2 ลักษณะ คือ การให้ผู้ถูกวัดตอบในกระดาษ (แบบสอบถาม) หรือการสัมภาษณ์ผู้ถูกวัด เครื่องมือที่ใช้กับวิธีนี้เป็นพวกแบบสอบถาม แบบสัมภาษณ์ หรือแบบบันทึก

วิธีการที่สาม คือ **การสังเกต** ซึ่งเป็นการสังเกตพฤติกรรมการแสดงออกของผู้ถูกวัดตามสภาพที่เป็นจริง ส่วนใหญ่การวัดด้วยวิธีสังเกต มักไม่ให้ผู้ถูกวัดสังเกตรู้ตัว เพราะจะทำให้เกิดพฤติกรรมแสสร้งได้ เช่น การสังเกตพฤติกรรมความซื่อสัตย์ ความจริงใจ ความเลี้ยสละ ความเป็นผู้นำ เป็นต้น บางกรณีเราก้ยอมให้ผู้ถูกสังเกตรู้ตัวว่ากำลังถูกสังเกต อาทิ การสังเกตการประชุมของชาวบ้าน การทำการเกษตรตามวิธีที่ได้รับการอบรม เป็นต้น จะเห็นได้ว่าตัวแปรที่วัดค่าได้โดยวิธีการสังเกตนี้มีทั้งตัวแปรที่เป็นความสามารถทางสมอง ความคิดจิตใจ และทางทักษะต่างๆ เครื่องมือที่ใช้จะเป็นพวกแบบสังเกต แบบบันทึก เป็นต้น

เมื่อผู้ประเมินเลือกวิธีการวัดค่าตัวแปรที่เหมาะสมแล้ว ก็ต้องสร้างเครื่องมือที่มีคุณภาพมาเก็บข้อมูลค่าตัวแปรดังกล่าว ซึ่งเราอาจแบ่งประเภทเครื่องมือตามลักษณะของเครื่องมือได้ 5 ประเภท ได้แก่ แบบสอบ แบบสอบถาม แบบสัมภาษณ์ แบบสังเกต และแบบบันทึก ซึ่งจะกล่าวถึงการสร้างเครื่องมือแต่ละประเภท ดังนี้

1) **แบบสอบ** แบบสอบเป็นชุดข้อคำถามที่เป็นตัวแทนพฤติกรรมที่ต้องการวัดโดยมากจะใช้วัดการเรียนรู้ทางสมอง เช่น แบบสอบผลสัมฤทธิ์ทางการเรียน แบบสอบผลสัมฤทธิ์ทางการฝึกอบรม ซึ่งอาจแบ่งขั้นตอนในการสร้างได้ 5 ขั้นตอน ดังนี้

ขั้นที่ 1 วิเคราะห์จุดมุ่งหมายของการสอบ โดยการวิเคราะห์จุดมุ่งหมายของหลักสูตร จุดมุ่งหมายรายวิชา วิเคราะห์เนื้อหา และกำหนดวัตถุประสงค์เชิงพฤติกรรมที่ต้องการวัดของแต่ละเนื้อหาย่อย

ขั้นที่ 2 สร้างตารางวิเคราะห์เนื้อหาและวัตถุประสงค์เชิงพฤติกรรม (table of specification) ซึ่งระบุน้ำหนักความสำคัญของแต่ละเนื้อหา จำนวนข้อสอบตามสัดส่วนความสำคัญและประเภทของข้อสอบ

ขั้นที่ 3 เขียนข้อสอบ โดยเขียนให้มีความสอดคล้องระหว่างเนื้อหา จุดมุ่งหมาย และประเภทข้อสอบ ข้อสอบต้องมีความชัดเจนที่ผู้สอบจะต้องเข้าใจคำถามได้ตรงกัน จากนั้นทบทวนความถูกต้องและเหมาะสมของข้อสอบหลังจากที่เขียนเสร็จพร้อมทั้งแก้ไข

ขั้นที่ 4 จัดทำแบบสอบ หลังจากได้ทบทวนและแก้ไขข้อสอบแล้ว จัดหมวดหมู่ของข้อสอบตามประเภท เขียนคำชี้แจงหรือวิธีการตอบให้ชัดเจน จัดพิมพ์ในรูปแบบที่น่าตอบ

ขั้นที่ 5 ทดลองใช้ (ถ้าเป็นไปได้) การทดลองใช้อาจทำเป็น 2 ระยะ ได้แก่ **ระยะแรก** ทดลองใช้กับกลุ่มตัวอย่างขนาดเล็กที่มีลักษณะใกล้เคียงกับผู้สอบจริง เพื่อตรวจสอบ **ความถูกต้องชัดเจนของภาษา** ในข้อสอบ แล้วทำการปรับปรุง **ระยะที่ 2** ทดลองใช้กับคนกลุ่มใหญ่ที่มีลักษณะใกล้เคียงกับกลุ่มตัวอย่างจริง **เพื่อวิเคราะห์ข้อสอบ** ด้านความยาก อำนาจจำแนก วิเคราะห์คุณภาพ

แบบสอบในด้านความเที่ยง ความตรง เพื่อปรับปรุงแก้ไขผลการวิเคราะห์ แล้วจัดทำเป็นแบบสอบที่จะนำไปใช้จริง

2) แบบสอบถามและแบบสัมภาษณ์

แบบสอบถามและแบบสัมภาษณ์ เป็นเครื่องมือที่นิยมใช้อย่างแพร่หลายในการประเมินผล แบบสอบถามเป็นชุดของข้อความที่สร้างขึ้นเพื่อรวบรวมข้อมูล ทั้งทางด้านข้อเท็จจริง (fact) และความคิดเห็น (opinion) โดยให้บุคคลตอบในแบบสอบถามซึ่งอาจเป็นชุดของข้อความ รูปภาพ หรือสัญลักษณ์ก็ได้ แบบสัมภาษณ์จะต่างกันตรงที่การสัมภาษณ์จะต้องไปสนทนากับผู้ให้สัมภาษณ์ ซึ่งสามารถโต้ตอบอธิบายรายละเอียดให้ผู้ตอบตรงคำถามได้ ส่วนแบบสอบถามมักไม่มีการพบปะระหว่างผู้ประเมินกับผู้ตอบ ดังนั้น คุณภาพการตอบของผู้ตอบส่วนหนึ่งจะขึ้นอยู่กับว่าผู้ประเมินได้สร้างแบบสอบถามให้ผู้ตอบเข้าใจได้เพียงใด

การสร้างแบบสอบถาม/แบบสัมภาษณ์

ขั้นที่ 1 กำหนดวัตถุประสงค์ของแบบสอบถาม/แบบสัมภาษณ์

ขั้นที่ 2 ระบุตัวแปรที่จะศึกษา นิยามตัวแปร และประเด็นคำถามที่จะวัดตัวแปรแต่ละตัวให้ครบถ้วนตามวัตถุประสงค์

ขั้นที่ 3 กำหนดรูปแบบของคำถามที่ต้องการใช้ให้เหมาะสมกับผู้ที่จะตอบ

ขั้นที่ 4 ร่างคำถาม วางโครงสร้างคร่าวๆ

ขั้นที่ 5 ทบทวนร่างคำถามโดยผู้สร้างเอง ในแง่ความชัดเจน ความครบถ้วน ความครอบคลุมตามวัตถุประสงค์ที่กำหนดไว้

ขั้นที่ 6 ทดลองใช้ (try out) แบบสอบถามกับกลุ่มตัวอย่างขนาดเล็ก ที่มีลักษณะเหมือนผู้ที่จะตอบจริง เพื่อตรวจสอบความถูกต้องชัดเจนของข้อความคำถาม

ขั้นที่ 7 ปรับปรุง และจัดพิมพ์ในรูปแบบที่น่าตอบ

สำหรับคำถามในแบบสอบถามนั้น แบ่งได้ออกเป็น 2 ประเภท คือ คำถามปลายเปิด (open ended) และคำถามปลายปิด (closed ended)

คำถามปลายเปิด เป็นคำถามที่เปิดโอกาสให้ผู้เขียนตอบอย่างอิสระ เช่น คำถามเกี่ยวกับความคิดเห็นของประชาชนที่มีต่อโครงการพัฒนาหมู่บ้าน เป็นต้น ส่วน**คำถามปลายปิด** จะเป็นคำถามที่กำหนดคำตอบไว้ให้ผู้ตอบเลือกคำตอบตายตัว ผู้ตอบเพียงแต่เลือกจากคำตอบที่กำหนดให้

ข้อควรคำนึงในการสร้างแบบสอบถาม มีดังนี้

1. คำถามต้องง่าย กระชับ ชัดเจน อ่านแล้วเข้าใจได้ตรงกันว่าผู้ถามต้องการจะถามอะไร
2. ในแต่ละคำถามควรถามประเด็นเดียว
3. การตั้งคำถามต้องไม่เป็นการถามนำ หรือชี้นำคำตอบ

4. เรียงลำดับข้อคำถามจากง่ายไปหายาก จากทั่วไปไปหาเฉพาะเจาะจง หรือเรียงตามลำดับเหตุการณ์

5. ต้องคำนึงถึงความรู้ ประสบการณ์ และวัยของผู้ตอบเสมอ

6. มีคำชี้แจงที่ชัดเจนว่าจะให้ผู้ตอบตอบอย่างไร

7. สร้างแบบสอบถามเสร็จต้องทดลองใช้ (try out) ก่อนทุกครั้ง

ต่อไปนี้เป็นตัวอย่างของแบบสอบถาม ซึ่งมีทั้งคำถามปลายเปิดและปลายปิดอยู่ในชุด

เดียวกัน

(ตัวอย่าง)

แบบสอบถาม

ความคิดเห็นของการใช้ตัวชี้วัดความสำเร็จของโครงการ
สำหรับประเมินโครงการพัฒนาในระดับจังหวัด

คำชี้แจง โปรดทำเครื่องหมาย ✓ ลงในช่อง และเติมข้อความลงในช่องว่างที่ตรงกับข้อเท็จจริงหรือ
ความคิดเห็นของท่าน

1. ข้อมูลทั่วไป

- สถานภาพของผู้ตอบ
- เพศ ชาย หญิง
 - อายุ _____ ปี
 - วุฒิทางการศึกษา ต่ำกว่าปริญญาตรี
 ปริญญาตรี สาขา _____
 สูงกว่าปริญญาตรี สาขา _____
 - ตำแหน่ง _____
 - สังกัด กอง _____ กรม _____ กระทรวง _____
 - ที่ตั้งหน่วยงานในจังหวัด เลขที่ _____ หมู่ที่ _____ ตำบล _____
อำเภอ _____ จังหวัด _____
โทรศัพท์ _____

2. ข้อมูลเกี่ยวกับการประเมินโครงการพัฒนาในระดับจังหวัด

- หน่วยงานประจำจังหวัดในสังกัดของท่าน มีเจ้าหน้าที่ประเมินโครงการทั้งหมด _____ คน
- ในปีงบประมาณ 2553 ที่ผ่านมา หน่วยงานในสังกัด (กรม) ของท่านมีโครงการในจังหวัดที่ท่านปฏิบัติงานอยู่ทั้งหมด _____ โครงการ และมีการประเมินโครงการทั้งหมด _____ โครงการ โปรดระบุชื่อโครงการและเหตุผลในการเลือกโครงการมาทำการประเมิน

ชื่อโครงการ	เหตุผล

3. การประเมินโครงการในสังกัด(กรม) ของท่าน มีการกำหนดผู้รับผิดชอบการประเมินโครงการ แต่ละโครงการอย่างไร
- ผู้ประเมินคนเดียวรับผิดชอบ
 - ผู้ประเมินเป็นคณะกรรมการ โดย
 - สังกัดกรมเดียว
 - สังกัดหลายกรม
4. การประเมินโครงการส่วนใหญ่ ทำการประเมินในช่วงเวลาใด
- ก่อนเริ่มโครงการ
 - ระหว่างดำเนินโครงการ
 - ระหว่างดำเนินโครงการและสิ้นสุดโครงการ
 - ก่อน ระหว่าง และสิ้นสุดโครงการ
 - สิ้นสุดโครงการ
5. การเก็บรวบรวมข้อมูลสำหรับประเมินโครงการ มีการกระทำในลักษณะใด
- เก็บข้อมูลเฉพาะโครงการที่ต้องการประเมิน
 - มีระบบเก็บข้อมูลของทุกโครงการ (โปรดระบุรายละเอียดวิธีการเก็บรวบรวม) _____

6. เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลส่วนใหญ่เป็นข้อมูลแบบใด
- แบบบันทึกข้อมูล
 - แบบสอบถาม
 - แบบสัมภาษณ์
 - แบบสังเกต
7. การประเมินส่วนใหญ่ผู้ใช้ผลการประเมินคือใคร _____
 และนำไปใช้ประโยชน์อะไร _____

8. การรายงานผลการประเมินโครงการจัดทำในลักษณะใด
- รายงานด้วยวาจา
 - ไม่เป็นทางการ เสนอต่อ _____
 - เป็นทางการ เสนอต่อ _____
 - จัดพิมพ์เป็นรูปเล่ม เสนอต่อ _____

9. ปัญหาอุปสรรคของการประเมินโครงการพัฒนาในระดับจังหวัดที่ผ่านมา

3. ความคิดเห็นต่อการนำไปใช้ของการจัดประเภทโครงการและตัวชี้วัดสำหรับประเมินโครงการโปรดพิจารณาเอกสารการจัดประเภทโครงการพัฒนาในระดับจังหวัด และตัวชี้วัดสำหรับประเมินโครงการพัฒนาในระดับจังหวัดที่แนบมานี้ โดยตัวชี้วัดที่เสนอจะมีการจัดทำคู่มือการใช้ แนวทางการสร้างเครื่องมือ เพื่อทำการฝึกอบรมก่อนนำไปใช้อีกครั้ง

3.1 การจำแนกประเภทโครงการพัฒนาในระดับจังหวัด

รายการ	มากที่สุด 5	มาก 4	ปานกลาง 3	น้อย 2	น้อยที่สุด 1
1. ความครอบคลุมของการจำแนกโครงการพัฒนาในระดับจังหวัดเป็น 5 ประเภท					
2. ความเป็นได้ในทางปฏิบัติเพื่อจำแนกโครงการพัฒนาในระดับจังหวัดเป็น 5 ประเภท					

ความคิดเห็นและข้อเสนอแนะ _____

แบบสัมภาษณ์ เป็นรายการคำถามที่กำหนดไว้เป็นแนวทางในการถาม ว่าต้องการถามอะไรบ้าง แล้วผู้สัมภาษณ์จะบันทึกคำตอบลงไป แบบสัมภาษณ์จะเก็บข้อมูลได้ละเอียด ลึกซึ้ง และเหมาะกับกลุ่มเป้าหมายที่อ่านเขียนไม่คล่อง แต่แบบสัมภาษณ์ก็อาจเก็บข้อมูลไม่ได้ตรงกับข้อเท็จจริงในกรณีที่เรื่องถามอาจมีผลย้อนกลับมาสู่ตัวผู้สัมภาษณ์ได้ เช่น การสัมภาษณ์ชาวบ้านเกี่ยวกับปฏิบัติหน้าที่ของผู้ใหญ่บ้าน ซึ่งถ้าเก็บข้อมูลด้วยแบบสอบถามอาจให้ผลที่ตรงข้อเท็จจริงมากกว่า

โดยทั่วไปแบบสัมภาษณ์จะมี 2 ประเภท คือ แบบสัมภาษณ์แบบมีโครงสร้างและแบบสัมภาษณ์แบบไม่มีโครงสร้าง แบบสัมภาษณ์แบบมีโครงสร้างจะกำหนดทั้งคำถามและตัวเลือกไว้ละเอียด คล้ายๆกับแบบสอบถาม ใช้ได้สะดวกแม้ว่าผู้สัมภาษณ์จะไม่มี ความชำนาญในการสัมภาษณ์นักก็ตาม ดังตัวอย่างต่อไปนี้

ตัวอย่างแบบสัมภาษณ์แบบมีโครงสร้าง

แบบสัมภาษณ์

ผู้เข้าอบรมหลักสูตรการสร้างอาชีพเสริมในชุมชน

ตอนที่ 1 ข้อมูลส่วนตัว

1. เพศ () ชาย () หญิง
2. อายุ ปี
3. การศึกษาสูงสุด
4. อาชีพหลักของครอบครัว
5. รายได้ของครอบครัว โดยเฉลี่ยเดือนละ บาท

ตอนที่ 2 ความคิดเห็นเกี่ยวกับการอบรมในหลักสูตรการสร้างเสริมอาชีพในชุมชน

1. ท่านทราบข่าวการอบรมจากที่ใด
() กำนัน ผู้ใหญ่บ้าน () เพื่อน () ประกาศ () อื่นๆ ระบุ
2. ท่านเข้าอบรมในกลุ่มอาชีพใด
.....
3. ท่านได้รับประโยชน์จากการอบรมในเรื่องใดบ้าง
.....
.....

สำหรับแบบสัมภาษณ์แบบไม่มีโครงสร้าง จะเป็นการกำหนดเพียงแนวหัวข้อการสัมภาษณ์ ผู้สัมภาษณ์จะระบุข้อความกว้างๆ ที่เป็นแนวในการถาม เปิดโอกาสให้ผู้ตอบได้แสดงความคิดเห็น โดยไม่บังคับให้จำใจตอบตามคำตอบที่กำหนด ดังตัวอย่างต่อไปนี้

(ตัวอย่าง)

แนวทางการสัมภาษณ์

หัวหน้าฝ่ายข้อมูลและติดตามประเมินผล

สำนักงานจังหวัด

-
1. การประเมินโครงการพัฒนาในระดับจังหวัด : สภาพปัจจุบัน
 - 1.1 อัตรากำลัง คุณวุฒิ ประสบการณ์ในการวิจัย/ประเมิน
 - 1.2 ขอบเขตงานในความรับผิดชอบ ปริมาณงาน ข้อมูลเกี่ยวกับโครงการ/รายละเอียด
 - 1.3 การประเมินโครงการพัฒนาในระดับจังหวัดที่ผ่านมา
 - 1.3.1 มีการประเมินโครงการหรือไม่
 - 1.3.2 การบริหารงานประเมิน : ใครเป็นผู้ประเมิน ผู้ประเมินคนเดียวหรือคณะกรรมการ
เลือกโครงการประเมิน ปริมาณโครงการที่ประเมินในแต่ละปี
 - 1.3.3 ความร่วมมือ/ประสานงานการประเมินระหว่างกระทรวง/หน่วยงาน อย่างไร
 - 1.3.4 วิธีการประเมินที่ใช้ : การตั้งวัตถุประสงค์ของการประเมิน รูปแบบการประเมิน มี
ตัวชี้วัด เกณฑ์การสร้างเครื่องมือ วิธีการรวบรวมข้อมูล การวิเคราะห์ข้อมูล
 - 1.3.5 รายงานการประเมิน : จัดทำอย่างไร รายงานใคร รูปแบบการรายงาน
 - 1.3.6 ปัญหาอุปสรรคที่ผ่านมา
 2. การประเมินโครงการพัฒนาในระดับจังหวัด : แนวทางในอนาคต
 - 2.1 รูปแบบที่ควรจะเป็นในการร่วมมือ/ประสานงานประเมินระหว่างกระทรวง/หน่วยงาน
 - 2.2 รูปแบบที่ควรจะเป็นในการบริหารงานประเมิน : ผู้รับผิดชอบในการประเมินควรเป็นใคร
คณะกรรมการ รูปแบบการประเมิน ระบบข้อมูลสำหรับการประเมิน การเก็บรวบรวม
ข้อมูล การวิเคราะห์ข้อมูล การจัดทำรายงาน
 3. ความเป็นไปได้ในการใช้ตัวชี้วัดเพื่อประเมินโครงการพัฒนาในระดับจังหวัด (ถามตามรายการ
ตัวชี้วัดในแบบสอบถามถึงความเหมาะสม/ความเป็นไปได้)
 - 3.1 ตัวชี้วัดสำหรับประเมินตัวโครงการ
 - 3.2 ตัวชี้วัดสำหรับประเมินการบริหารโครงการ
 - 3.3 ตัวชี้วัดสำหรับประเมินผลทางตรง
 - 3.4 ตัวชี้วัดสำหรับประเมินผลทางอ้อม

3) แบบสังเกตและแบบบันทึก

แบบสังเกตและแบบบันทึก เป็นเครื่องมือที่จัดทำขึ้นเพื่อเก็บรวบรวมข้อมูลในสิ่งที่สังเกตได้ แบบสังเกตมักจะใช้บันทึกรายละเอียดข้อเท็จจริงในทันทีที่เกิดเหตุการณ์ หรือพฤติกรรมที่เฝ้าสังเกต ส่วนแบบบันทึกอาจใช้บันทึกเหตุการณ์เรื่องราวต่างๆ ที่เกิดขึ้นไปแล้ว ทั้งแบบสังเกต และแบบบันทึก มีขั้นตอนหลักๆ ในการสร้างดังนี้

ขั้นที่ 1 กำหนดวัตถุประสงค์ของแบบสังเกต/แบบบันทึก

ขั้นที่ 2 ระบุตัวแปรที่จะศึกษา นิยามตัวแปร และพฤติกรรมบ่งชี้ให้ละเอียดครบถ้วน และเป็นตัวแทนตัวแปรที่ศึกษาได้

ขั้นที่ 3 กำหนดรูปแบบสังเกต/แบบบันทึก ให้มีความครบถ้วนของประเด็นที่สังเกต และชัดเจนพอที่จะบันทึกผลการสังเกตได้อย่างรวดเร็ว

ขั้นที่ 4 ร่างข้อความเกี่ยวกับพฤติกรรมบ่งชี้ หรือประเด็นที่ต้องการสังเกต

ขั้นที่ 5 ทบทวนร่างข้อความในแง่ความชัดเจนในการสื่อความหมาย และความครบถ้วนครอบคลุมตามวัตถุประสงค์ที่ตั้งไว้

ขั้นที่ 6 ทดลองใช้แบบสังเกต/แบบบันทึก โดยให้ผู้สังเกตหลายคน สังเกตสิ่งเดียวกัน โดยใช้แบบสังเกต/แบบบันทึกที่สร้างขึ้น แล้วตรวจสอบความสอดคล้องระหว่างผลการสังเกตของผู้สังเกตทุกคน

ขั้นที่ 7 ปรับปรุงและจัดพิมพ์เป็นแบบสังเกต/แบบบันทึกฉบับที่ใช้จริง

ตัวอย่าง ของแบบสังเกตและแบบบันทึกฉบับที่ใช้จริง

1) แบบสังเกตแบบตารางจดบันทึก

แบบสังเกตพฤติกรรมการปฏิบัติงานของเจ้าหน้าที่กลุ่มออมทรัพย์

เวลา	สถานที่	พฤติกรรม	จำนวน ครั้ง	หมายเหตุ
09.00 – 11.00	ที่ทำการกลุ่ม ออมทรัพย์	1. พิมพ์ดีด 2. จัดเอกสารเข้าแฟ้ม 3. เข้าห้องน้ำ 4. สนทนากับเพื่อน	10 2 8 20	

2) แบบสังเกตแบบตรวจสอบรายการ

แบบสังเกตพฤติกรรมการสอนของวิทยากร

ชื่อวิทยากร วันที่.....เวลา.....

ข้อที่	รายการพฤติกรรม	ปฏิบัติ	ไม่ได้ปฏิบัติ	ข้อมูล เพิ่มเติม
1.	มีการเตรียมการบรรยาย			
2.	นำเข้าสู่บทเรียนสัมพันธ์กับเนื้อหา			
3.	ใช้เวลาในการนำเข้าสู่บทเรียนได้เหมาะสม			
4.	ใช้วิธีการสอนและจัดกิจกรรมการสอนที่สอดคล้องกับเนื้อหา			
5.	เปิดโอกาสให้ผู้เข้าอบรมได้มีส่วนร่วมในกิจกรรมการเรียนการสอนอย่างเหมาะสม			
6.	เลือกใช้สื่ออุปกรณ์การสอนได้เหมาะสมกับเนื้อหา			
7.	ให้ความสนใจแก่ผู้เข้าอบรมอย่างทั่วถึง			
8.	รับฟังความคิดเห็นของผู้เข้าอบรม			

3) แบบสังเกตโดยมีการประเมินค่า

รายการพฤติกรรม	มากที่สุด 5	มาก 4	ปานกลาง 3	น้อย 2	น้อยที่สุด 1
1. การแต่งกายสะอาดเรียบร้อยสมกับเป็นวิทยากร					
2. มีอารมณ์มั่นคงและควบคุมอารมณ์ได้ดี					
3. ตรงต่อเวลากับการเข้าสอนและเลิกสอน					
4. มีความอดทน ชยัน และกระตือรือร้น					
5. บริหารเวลาได้อย่างเหมาะสม					

4) แบบบันทึก

แบบบันทึกการปฏิบัติงาน				
หัวหน้าโครงการ.....				
อำเภอ.....จังหวัด.....				
วันที่.....เดือน..... พ.ศ.				
เริ่มเวลา	สิ้นสุดเวลา	รายการปฏิบัติงาน	ผู้รับบริการ	หมายเหตุ
ผู้บันทึก				

3. การตรวจสอบคุณภาพของเครื่องมือ

เครื่องมือในการรวบรวมข้อมูลอาจได้มาจาก 2 ทาง คือ เครื่องมือที่มีผู้สร้างไว้แล้วหรือเครื่องมือมาตรฐาน กับเครื่องมือที่นักประเมินสร้างขึ้นเอง เครื่องมือประเภทแรกจะช่วยทุ่นแรงเพราะไม่ต้องเสียเวลาไปสร้างเครื่องมือใหม่ แต่ถ้าไม่สามารถหาเครื่องมือที่ตรงกับงานประเมินที่ทำได้ ก็ต้องสร้างขึ้นใหม่อย่างมีหลักวิชาและต้องตรวจสอบคุณภาพเครื่องมือด้วย เครื่องมือที่มีคุณภาพจะมีคุณสมบัติหลักๆ ต่อไปนี้

ความตรง (Validity) หมายถึง คุณสมบัติของเครื่องมือที่สามารถวัดในสิ่งที่ต้องการวัดได้ หรือสามารถเก็บข้อมูลได้ตรงกับวัตถุประสงค์ของการวิจัย

ความเที่ยง (Reliability) หมายถึง คุณสมบัติของเครื่องมือที่ให้ผลการวัดที่คงเส้นคงวา ใช้วัดกี่ครั้งก็ให้ผลเหมือนเดิม

ความเป็นปรนัย (Objectivity) หมายถึง คุณสมบัติของเครื่องมือในแง่ของความชัดเจนที่ผู้อ่านทุกคนจะเข้าใจคำถามได้ตรงกัน

1) การประมาณค่าความเที่ยงของเครื่องมือ

ในการประมาณค่าความเที่ยงสามารถทำได้หลายวิธี และในแต่ละวิธีก็มีข้อจำกัดที่แตกต่างกัน การที่จะเลือกใช้วิธีใดขึ้นอยู่กับความมุ่งหมาย ลักษณะของคะแนน และชนิดของแบบสอบ (Guilford. 1978 : 84) คือ

1.1 แบบสอบซ้ำ (Measure of Stability)

1.2 แบบใช้ข้อสอบคล้ายกัน (Measure of Equivalence)

1.3 แบบวัดความคงที่ภายใน (Measure of Internal Consistency) ซึ่งเทคนิคการ

วิเคราะห์ทางสถิติที่ใช้ ได้แก่

1.3.1 วิธีครึ่งข้อสอบ (Split-half)

1.3.2 วิธีคูเดอร์-ริชาร์ดสัน (Kuder-Richardson estimates)

1.3.3 สัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's Coefficient)

1.3.4 วิธีวิเคราะห์ความแปรปรวนของฮอยท์ (Hoyt's analysis of variance

procedure)

1.1 ความเที่ยงแบบสอบซ้ำ (Measure of Stability)

เป็นวิธีที่จะหาสัมประสิทธิ์ของความคงที่ โดยนำเครื่องมือไปทดสอบกับกลุ่มตัวอย่างเดิมสองครั้ง และควรทิ้งช่วงระยะเวลาพอควรก่อนการสอบครั้งที่สอง แล้วนำคะแนนที่ได้จากการทดสอบทั้งสองครั้งมาคำนวณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson Product Moment Correlation Coefficient) โดยสูตรต่อไปนี้

$$r_{xy} = \frac{N \sum xy - \sum x \sum y}{\sqrt{[N \sum x^2 - (\sum x)^2][N \sum y^2 - (\sum y)^2]}}$$

เมื่อ R_{xy} คือ สัมประสิทธิ์ความเที่ยง
 N คือ จำนวนผู้เข้าสอบ
 x คือ คะแนนสอบครั้งที่หนึ่ง
 y คือ คะแนนสอบครั้งที่สอง

การหาความเที่ยงโดยวิธีการสอบซ้ำนี้ นิยมเรียกกันว่า Test-retest Method ซึ่งการหาความเที่ยงด้วยวิธีสอบซ้ำมักจะทำให้เกิดความคลาดเคลื่อน เนื่องจากมีองค์ประกอบอื่นมาทำให้คะแนนที่สอบในครั้งที่สองเปลี่ยนไป การทิ้งช่วงระยะเพียงเล็กน้อยอาจทำให้ผู้สอบจำได้ว่าในครั้งแรกได้ตอบคำถามไปอย่างไร ครึ่งหลังก็พยายามตอบแบบเดิมโดยไม่พิจารณาว่าคำตอบนั้นตอบถูกหรือผิด ในลักษณะนี้จะทำให้ความเที่ยงสูงกว่าความเป็นจริง แต่ถ้าเว้นระยะเวลานานเกินไปผู้สอบอาจเกิดการเรียนรู้มากขึ้น และมีพัฒนาการทางสติปัญญามากขึ้น ซึ่งจะมีผลทำให้ความเที่ยงต่ำ สำหรับการทิ้งช่วงระยะเวลาในการสอบซ้ำ ยังไม่มีข้อสรุปว่าควรเว้นระยะเวลาเท่าใดจึงจะเหมาะสม ตัวอย่างการคำนวณความเที่ยงแบบวัดซ้ำมีดังนี้

ตัวอย่าง การคำนวณความเที่ยงแบบวัดซ้ำ

ในการตรวจสอบคุณภาพแบบวัดความรู้สึกเป็นเจ้าของ (Sense of Belonging) ที่จะใช้ในการประเมินความสำเร็จของโครงการพัฒนาแหล่งน้ำหมู่บ้าน โดยนักประเมินนำแบบวัดไปทดลองใช้กับชาวบ้าน 10 คน เว้นระยะไปอีก 2 สัปดาห์ทดลองให้ชาวบ้านคนเดิมทำแบบวัดนั้นใหม่ ปรากฏคะแนนการวัด 2 ครั้ง ดังนี้

ชาวบ้านคนที่	1	2	3	4	5	6	7	8	9	10
คะแนนจากการวัด										
ครั้งที่ 1	35	40	40	40	45	48	49	50	50	55
ครั้งที่ 2	55	50	48	40	40	52	52	52	60	65

ตารางที่ 1 คำนวณค่าสหสัมพันธ์ระหว่างผลการวัด 2 ครั้ง ได้ดังนี้

คนที่	X	Y	XY	X ²	Y ²
1	35	55	1,925	1,225	3,025
2	40	50	2,000	1,600	2,500
3	40	48	1,920	1,600	2,304
4	40	40	1,600	1,600	1,600
5	45	40	1,800	2,025	1,600
6	48	52	2,496	2,304	2,704
7	49	52	2,548	2,401	2,704
8	50	52	2,600	2,500	2,704
9	50	60	3,000	2,500	3,600
10	50	65	3,575	3,025	4,225
รวม	452	514	23,464	20,780	26,966

$$\begin{aligned}
 r_{xy} &= \frac{N \sum xy - \sum x \sum y}{\sqrt{[N \sum x^2 - (\sum x)^2][N \sum y^2 - (\sum y)^2]}} \\
 &= \frac{10(23,464) - (452)(514)}{\sqrt{[10(20,780) - (452)^2][10(26,966) - (514)^2]}} \\
 &= \frac{2,312}{\sqrt{3,496(5,464)}} \\
 &= \frac{2,312}{4,370.6} \\
 &= 0.53
 \end{aligned}$$

∴ ความเที่ยงของแบบวัดซ้ำของแบบวัดชุดนี้มีค่าเท่ากับ 0.53

แสดงว่า แบบวัดชุดนี้มีความเที่ยงอยู่ในระดับปานกลาง สมควรที่จะต้องปรับปรุงก่อนที่จะนำไปใช้จริง โดยพิจารณาความสอดคล้องกันของข้อคำถามและความเป็นปรนัยของข้อคำถามแต่ละข้อ เมื่อปรับปรุงแล้วควรตรวจสอบความเที่ยงอีกครั้งหนึ่ง

1.2 ความเที่ยงแบบใช้ข้อสอบคล้ายกัน (Measure of Equivalence)

เนื่องจากวิธีหาความเที่ยงแบบสอบซ้ำประสบปัญหาเกี่ยวกับการเว้นช่วงระยะเวลาจึงใช้วิธีนำเครื่องมือ 2 ฉบับ ที่คล้ายกัน (Parallel test) มาใช้แทน คือ ข้อคำถามที่ใช้ในเครื่องมือทั้ง 2 ฉบับ จะมีลักษณะเป็นข้อคำถามที่สมมูลกัน (equivalent) ไปทดสอบผู้สอบกลุ่มหนึ่งในเวลาเดียวกัน แล้วนำ

คะแนนที่ได้จากแบบทดสอบทั้ง 2 ฉบับ ไปคำนวณค่าความเที่ยง โดยการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันเช่นเดียวกับแบบสอบซ้ำ

1.3 ความเที่ยงแบบวัดความคงที่ภายใน (Measure of Internal Consistency)

เป็นการตรวจสอบว่าข้อคำถามในเครื่องมือวัดนั้นวัดในเรื่องเดียวกันหรือไม่ ถ้าวัดในเรื่องเดียวกันก็น่าจะมีความคงที่ในการวัดสูง โดยที่จะมีการตอบเพียงครั้งเดียว สำหรับวิธีหาความเที่ยงนิยมใช้คือ วิธีของคูเดอร์-ริชาร์ดสัน วิธีหาสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's Coefficient) และวิธีวิเคราะห์ความแปรปรวนของฮอยท์

1) วิธีของคูเดอร์-ริชาร์ดสัน (Kuder-Richardson)

การประมาณค่าสัมประสิทธิ์ความเที่ยงของคูเดอร์-ริชาร์ดสัน ใช้ได้กับการให้คะแนนเป็นระบบ 0-1 สำหรับวิธีการคำนวณ แสดงในตารางที่ 2 ดังนี้

ตารางที่ 2 ตัวอย่างการคำนวณหาค่า Reliability ด้วยสูตร คูเดอร์-ริชาร์ดสัน 20 (KR-20)

ผู้ตอบ	ข้อคำถาม								รวม
	1	2	3	4	5	6	7	8	
1	1	1	0	1	1	1	1	1	7
2	1	1	1	1	0	0	1	0	5
3	0	1	0	1	1	0	0	1	4
4	1	1	1	0	1	0	0	0	4
5	0	1	1	1	1	1	0	0	5
6	0	0	1	1	0	0	0	0	2
จำนวนผู้ตอบที่ตอบถูก	3	5	4	5	4	2	2	2	27
สัดส่วนที่ตอบถูก (p)	.50	.83	.67	.83	.67	.33	.33	.33	
สัดส่วนที่ตอบผิด (q)	.50	.17	.33	.17	.33	.67	.67	.67	
pq	.25	.14	.22	.14	.22	.22	.22	.22	
$\sum pq = 1.63$		$\bar{X} = 4.5$		$S^2 = 2.25$				$S = 1.50$	

$$\begin{aligned}
 \text{KR-20} &= \left[\frac{n}{n-1} \right] \left[1 - \frac{\sum pq}{S^2} \right] \\
 &= \left[\frac{8}{8-1} \right] \left[1 - \frac{1.63}{2.25} \right] \\
 &= \left[\frac{8}{7} \right] \left[\frac{0.62}{2.25} \right] \\
 &= 0.36
 \end{aligned}$$

แสดงว่าแบบวัดชุดนี้มีความเที่ยงค่อนข้างต่ำ ควรจะปรับปรุงในด้านความสอดคล้องของข้อคำถามและความเป็นปรนัยของข้อคำถามก่อนนำไปใช้จริง

2) วิธีสัมประสิทธิ์แอลฟา (Alpha Coefficient) วิธีนี้ได้รับการพัฒนาจากครอนบาค (Cronbach) โดยได้พัฒนาสูตรคูเคอร์-ริชาร์สัน 20 เป็นสัมประสิทธิ์แอลฟา เพื่อให้ใช้ได้กับการให้คะแนนที่ไม่เป็นระบบ 0-1 เช่น แบบสอบอัตรณ์ แบบสำรวจความสนใจในอาชีพ เป็นต้น ซึ่งมีสูตรในการคำนวณ ดังนี้

$$\alpha = \left[\frac{n}{n-1} \right] \left[1 - \frac{\sum \sigma_i^2}{\sigma_t^2} \right]$$

เมื่อ α คือ ความเที่ยงของแบบสอบ

n คือ จำนวนข้อสอบในแบบสอบ

σ_i^2 คือ ความแปรปรวนของคะแนนข้อสอบแต่ละข้อ

σ_t^2 คือ ความแปรปรวนของคะแนนของผู้สอบทั้งหมด

ตัวอย่างการคำนวณค่าความเที่ยงแบบความคงที่ภายใน โดยใช้สูตรสัมประสิทธิ์แอลฟาของครอนบาค มีดังนี้

ตารางที่ 3 ตัวอย่างการหาค่า Reliability ด้วยสูตร สัมประสิทธิ์แอลฟาของครอนบาค

คนที่	ข้อคำถาม										รวม
	1	2	3	4	5	6	7	8	9	10	
1	1	2	4	2	5	3	1	2	4	4	28
2	4	1	2	4	2	1	1	2	1	1	19
3	3	4	5	4	5	1	2	3	4	5	36
4	1	2	3	4	5	4	3	2	1	3	28
5	2	1	4	5	1	2	4	4	3	2	28
6	3	4	1	5	4	3	2	1	1	2	26
7	5	1	2	3	4	5	1	2	3	4	30
8	5	5	4	3	2	1	2	3	4	5	34
9	2	3	3	1	1	2	2	4	2	1	21
10	4	4	5	4	5	4	5	5	5	3	45
σ_i	1.5	1.4	1.3	1.3	1.7	1.3	1.3	1.2	1.4	1.1	$\sigma_i^2 =$ 50.42
σ_i^2	2.2	2.0	1.7	1.7	2.9	1.7	1.7	1.4	2.0	1.2	$\sum \sigma_i^2 =$ 18.50

$$\begin{aligned}
 \alpha &= \left[\frac{n}{n-1} \right] \left[1 - \frac{\sum \sigma_i^2}{\sigma_i^2} \right] \\
 &= (10/10-1) (1-18.50/50.42) \\
 &= (10/9) (1-0.366) \\
 &= 0.704
 \end{aligned}$$

แสดงว่าเครื่องมือฉบับนี้มีความเที่ยงค่อนข้างสูง สามารถนำไปใช้ได้ แต่เมื่อใช้ไประยะหนึ่ง ควรมีการปรับปรุง เพื่อให้เกิดความเที่ยงสูงยิ่งขึ้น

- การประมาณค่าความเที่ยงของแบบสังเกต

ลักษณะของแบบสังเกตที่เป็นแบบ check list การตรวจสอบหาความเที่ยงเป็นการ คำนวณค่าความเที่ยงของผู้สังเกตเพราะใช้การเก็บข้อมูลควบคู่กับแบบสังเกต มีสูตรการคำนวณที่ไม่ ยุ่งยากเหมือนของแบบสอบ ค่าความเที่ยงที่ได้เรียกว่า สัมประสิทธิ์ความสอดคล้อง มีสูตรดังนี้

$$\text{ส.ป.ส. ความสอดคล้อง} = \frac{\text{จำนวนช่วงที่บันทึกตรงกัน}}{\text{จำนวนช่วงที่บันทึกตรงกัน} + \text{จำนวนช่วงที่บันทึกต่างกัน}}$$

นอกจากนั้นการตรวจสอบความเที่ยงของแบบสังเกตยังทำได้อีก 2 วิธี คือ วิธีแรกเป็น การตรวจสอบความเที่ยงจากความสอดคล้องภายในผู้ประเมิน กล่าวคือ เมื่อผู้ประเมินได้ใช้แบบสังเกต สังเกตเหตุการณ์หนึ่งแล้วบันทึกผลการสังเกตไว้ เมื่อเว้นระยะไปประมาณสองสัปดาห์ ให้ผู้ประเมินมา สังเกตเหตุการณ์เดิม (ซึ่งอาจบันทึกวิธีที่ค้นไว้) โดยใช้แบบสังเกตชุดเดิม ผลการสังเกตทั้งสองครั้งควร จะสอดคล้องกันถ้าแบบสังเกตมีความเที่ยง ความเที่ยงแบบนี้เรียกว่า **ความเที่ยงภายในตัวผู้ประเมิน (Intra-rater Reliability)** คำนวณได้จากค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากการสังเกตครั้งที่ 1 และครั้งที่ 2 เช่นเดียวกับการคำนวณความเที่ยงแบบวัดซ้ำที่ได้นำเสนอไปแล้ว

อีกวิธีการหนึ่งเป็นการตรวจสอบความเที่ยงระหว่างผู้ประเมิน (Inter-rater Reliability) ซึ่งอยู่บนหลักการที่ว่า เมื่อใช้ผู้ประเมินตั้งแต่ 2 คนขึ้นไป สังเกตสิ่งเดียวกันด้วยแบบ สังเกตชุดเดียวกัน ผลการสังเกตที่ได้ควรจะสอดคล้องกัน ถ้าแบบสังเกตมีความเที่ยง วิธีการคำนวณ ความเที่ยงระหว่างผู้ประเมินแบ่งได้เป็น 2 กรณี ดังนี้

กรณี 1 ผู้ประเมิน 2 คน ใช้การคำนวณสหสัมพันธ์ระหว่างคะแนนจากการสังเกตผู้ ประเมินทั้งสองคน

กรณี 2 ผู้ประเมินมากกว่า 2 คน ใช้การวิเคราะห์ความแปรปรวนทดสอบความแตกต่างระหว่างคะแนนเฉลี่ยจากผู้ประเมินแต่ละคน ดังตัวอย่างต่อไปนี้

ตัวอย่างการคำนวณความเที่ยงระหว่างผู้ประเมิน

คะแนนจากการสังเกตทักษะการปฏิบัติงาน จากผู้ประเมิน 4 คน สำหรับผู้เข้ารับการอบรม 6 คน เป็นดังนี้

ตารางที่ 4 ตัวอย่างคำนวณหาค่า Reliability ด้วยวิธีวิเคราะห์ความแปรปรวน

ผู้รับการ อบรม	ผู้ประเมิน					ยกกำลังสอง				
	A	B	C	D	รวม	A	B	C	D	
1	10	6	8	7	31	100	36	64	49	
2	4	5	3	4	16	16	25	9	16	
3	8	4	7	4	23	64	16	49	16	
4	3	4	2	2	11	9	16	4	4	
5	6	8	6	7	27	36	64	36	49	
6	9	7	8	7	31	81	49	64	49	
รวม	40	34	34	31	139	306	206	226	183	$\sum X^2 = 921$

จากข้อมูลสามารถคำนวณการวิเคราะห์ความแปรปรวนได้ดังนี้

$$\begin{aligned}
 SS_t &= \sum X^2 - \frac{(\sum X)^2}{N} \\
 &= 921 - \frac{139^2}{24} = 116 \\
 SS_r &= \frac{31^2 + 16^2 + 23^2 + 11^2 + 27^2 + 31^2}{4} - \frac{139^2}{24} = 84.25 \\
 SS_c &= \frac{40^2 + 34^2 + 34^2 + 31^2}{6} - \frac{139^2}{24} = 7.16 \\
 SS_e &= SS_t - (SS_r + SS_c) \\
 &= 116 - (84.25 + 7.16) = 24.59
 \end{aligned}$$

แหล่งความแปรปรวน	df	SS	MS	F	P
ผู้รับการอบรม (row)	5	84.25	16.85	10.27	< .01
ผู้ประเมิน (column)	3	7.16	2.39	1.46	> .05
เศษเหลือ (error)	15	24.59	1.64		
รวม	23	116.00			

จากตัวอย่างข้างต้นแสดงว่าผู้รับการอบรมมีความแตกต่างกันในด้านทักษะในการปฏิบัติงาน แต่ผลการประเมินทั้ง 4 คนไม่แตกต่างกัน แสดงว่าผู้ประเมินให้ผลการประเมินที่สอดคล้องกัน

2) ความตรง (Validity)

ความตรงของแบบสอบ จำแนกออกได้หลายวิธี ทั้งนี้ขึ้นอยู่กับจุดมุ่งหมายของการวัด โดยทั่วไปการวัดทางการศึกษาและจิตวิทยาแบ่งความตรงออกเป็น 3 ประเภท

2.1 ความตรงตามเนื้อหา (Content Validity) เป็นความตรงที่เกี่ยวกับการวิเคราะห์ตรวจสอบอย่างมีระบบในเนื้อหาของแบบสอบ ว่าเนื้อหาของข้อคำถามวัดได้ตรงตามเนื้อหา ของเรื่องที่ต้องการวัดหรือไม่ ความตรงชนิดนี้นิยมใช้การพิจารณาจากนิยามเชิงทฤษฎีและนิยามเชิง ปฏิบัติการประกอบการตัดสินใจของผู้เชี่ยวชาญในสาขาวิชาการนั้นๆ ว่าแบบสอบนั้นมีความตรงตาม เนื้อหาที่ต้องการวัดและมีความครบถ้วนสมบูรณ์ ครอบคลุมเนื้อเรื่องทั้งหมดหรือไม่

โดยทั่วไปความตรงตามเนื้อหาสามารถคำนวณได้จากความสอดคล้องระหว่างข้อคำถาม กับวัตถุประสงค์ที่ต้องการวัด โดยคำนวณดัชนีความสอดคล้องระหว่างข้อ และวัตถุประสงค์ (Item-objective Congruency Index : IOC) ซึ่งดำเนินการโดยให้ผู้เชี่ยวชาญประเมินข้อคำถามทีละข้อว่า สอดคล้องกับวัตถุประสงค์ที่ผู้สร้างตั้งใจให้ข้อคำถามนั้นวัดหรือไม่ โดยให้ผลการประเมินเป็น 3 รายการ คือ สอดคล้อง ไม่แน่ใจว่าสอดคล้องหรือไม่ และไม่สอดคล้อง จากนั้นจึงจะคิดคะแนนและ คำนวณ IOC ดังตัวอย่างต่อไปนี้

วัตถุประสงค์	ข้อความถาม	ความสอดคล้องระหว่างข้อความถามกับวัตถุประสงค์			$IOC = \frac{\sum R}{n}$
		สอดคล้อง (1)	ไม่แน่ใจ (0)	ไม่สอดคล้อง (-1)	
1. ต้องการวัดความรู้สึกเป็นเจ้าของสระน้ำประจำหมู่บ้าน	1) ท่านทำอย่างไรเมื่อเห็นเพื่อนบ้านนำฝูงวัวมากินน้ำในสระน้ำประจำหมู่บ้าน	10	-	-	1.00
	2) ท่านได้รับประโยชน์จากสระน้ำประจำหมู่บ้านเพียงใด	1	2	7	-0.60

จากตาราง สมมติว่า ผู้ประเมินต้องการตรวจสอบความตรงเชิงเนื้อหาของแบบวัดความรู้สึกเป็นเจ้าของสระน้ำประจำหมู่บ้านในโครงการจัดสร้างสระน้ำประจำหมู่บ้าน ในตัวอย่างเป็นข้อความถาม 2 ข้อแรกที่มุ่งวัดความรู้สึกเป็นเจ้าของสระน้ำ นำไปให้ผู้ทรงคุณวุฒิ 10 คน ตรวจสอบความสอดคล้องระหว่างข้อความถามกับวัตถุประสงค์ที่ต้องการวัด ปรากฏว่าในข้อที่ 1 ผู้ทรงคุณวุฒิทั้ง 10 คน เห็นตรงกันว่า วัดได้สอดคล้อง คำนวณค่า IOC ได้เท่ากับ $\frac{10 \times 1}{10} = 1.00$ แสดงว่าข้อความถามมีความตรงตามเนื้อหาตามความคิดเห็นของผู้ทรงคุณวุฒิ ในขณะที่ข้อความถามที่ 2 มีผู้เห็นว่าสอดคล้อง 1 คน ไม่แน่ใจว่าสอดคล้องหรือไม่ 2 คน และ เห็นว่าไม่สอดคล้อง 7 คน คำนวณค่า IOC ได้ดังนี้

$$\begin{aligned}
 IOC &= \frac{\sum R}{n} = \frac{(1 \times 1) + (0 \times 2) + [7 \times (-1)]}{10} \\
 &= \frac{1 + 0 + (-7)}{10} \\
 &= \frac{-6}{10} \\
 &= -0.6
 \end{aligned}$$

แสดงว่าข้อความถามที่ 2 ไม่มีความตรงตามเนื้อหา

ค่า IOC จะมีค่าอยู่ระหว่าง -1 ถึง 1 ข้อความถามที่มีความตรงตามเนื้อหาจะมีค่า IOC เข้าใกล้ 1.00 ถ้าข้อใดมีค่า IOC ต่ำกว่า 0.5 ควรจะปรับปรุงข้อความถามใหม่ให้สอดคล้องกับวัตถุประสงค์ที่ต้องการวัด

2.2 ความตรงตามเกณฑ์สัมพัทธ์ (Criterion-related Validity) เป็นการหาความตรงแบบสอบว่าแบบสอบนั้นวัดได้ตรงตามพฤติกรรมที่ต้องการวัด โดยพิจารณาจากเกณฑ์ที่เกี่ยวข้องว่าแบบสอบนั้นจะสามารถทำนายพฤติกรรมของบุคคลในสภาพจำเพาะตามความต้องการ หรือเป็นความสัมพันธ์ระหว่างคะแนนที่ได้จากแบบสอบกับคะแนนการวัดที่ได้จากเกณฑ์ภายนอกที่อิสระ

ความตรงตามเกณฑ์สัมพัทธ์ จำแนกได้เป็น 2 ชนิด คือ

2.2.1) ความตรงร่วมสมัย (Concurrent Validity) หมายถึง ความตรงของแบบสอบที่จะบ่งบอกสิ่งที่วัดได้ถูกต้องตามสภาพที่แท้จริงในปัจจุบัน ตรวจสอบได้จากความสัมพันธ์ระหว่างผลการวัดจากเครื่องมือที่ผู้ประเมินสร้างขึ้นกับผลการวัดจากเครื่องมือที่เป็นมาตรฐานอยู่แล้ว เช่น จะตรวจสอบความตรงร่วมสมัยของแบบวัดการมีส่วนร่วมของประชาชนที่ผู้ประเมินสร้างขึ้น ดำเนินการโดยให้กลุ่มตัวอย่างตอบแบบวัดการมีส่วนร่วมที่ผู้ประเมินสร้างขึ้นและให้ตอบแบบวัดการมีส่วนร่วมที่เป็นมาตรฐานอยู่แล้ว จากนั้นนำผลการวัดจากเครื่องทั้ง 2 ฉบับ มาคำนวณค่าสหสัมพันธ์ ถ้ามีค่าสหสัมพันธ์สูงแสดงว่าแบบวัดการมีส่วนร่วมที่สร้างขึ้นมีความตรงร่วมสมัย

2.2.2) ความตรงเชิงทำนาย (Predictive Validity) หมายถึง ความสามารถของแบบสอบที่จะบ่งบอกผลที่วัดในขณะนั้นได้ถูกต้องตามสภาพที่แท้จริงในอนาคต โดยอาศัยความสัมพันธ์ระหว่างคะแนนของแบบสอบกับคะแนนเกณฑ์สัมพัทธ์ ซึ่งจะปรากฏในอนาคต เช่น เกณฑ์สัมพัทธ์ของแบบสอบความถนัดทางวิชาการเพื่อทำนายผลการเรียน ก็อาจใช้คะแนนเฉลี่ย สะสมปีสุดท้ายเป็นเกณฑ์สัมพัทธ์ ซึ่งการคำนวณหาความตรงเชิงนี้จะต้องอาศัยเวลารอคอย เพราะคะแนนของแบบสอบกับเกณฑ์สัมพัทธ์ได้มาคนละเวลากัน ความตรงเชิงทำนายเหมาะสำหรับที่จะใช้กับแบบสอบความถนัดชนิดต่างๆ และแบบสอบอื่นที่ต้องการวัดเพื่อทำนายผล

สรุปได้ว่าความตรงเชิงทำนายกับความตรงร่วมสมัยต่างใน 2 ประเด็น คือ

1) ช่วงระยะเวลาการเก็บรวบรวมข้อมูลที่เป็นเกณฑ์ เกณฑ์ของความตรงเชิงทำนายจะเกิดขึ้นภายหลังจากการเก็บคะแนนแบบสอบการตรวจสอบความตรง ส่วนความตรงร่วมสมัยจะเก็บรวบรวมข้อมูลที่เป็นเกณฑ์และคะแนนแบบสอบในเวลาเดียวกัน

2) ความมุ่งหมายของการทดสอบ ความตรงเชิงทำนายเป็นการนำคะแนนแบบสอบมาพยากรณ์ความสามารถหรือกระทำในอนาคต ส่วนความตรงร่วมสมัยเป็นการนำคะแนนแบบสอบเทียบกับความสามารถหรือความสำเร็จในขณะนั้น

วิธีประมาณค่าความตรงตามเกณฑ์สัมพัทธ์

ในการประมาณค่าความตรงตามเกณฑ์สัมพัทธ์ สามารถวิเคราะห์ได้จากสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากแบบสอบกับเกณฑ์ ดังนั้น สูตรที่ใช้จึงเป็นสูตรสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson's product moment correlation coefficient)

3) คุณภาพของเครื่องมือในด้านอื่นๆ

คุณภาพของเครื่องมือการวิจัย ไม่ได้มีเฉพาะความเที่ยงและความตรงเท่านั้น แต่ยังมีคุณภาพอื่นๆ ที่ควรพิจารณา คือ

3.1) ความเป็นปรนัย หมายถึง การที่เครื่องมือมีคำถามที่ชัดเจนสามารถเข้าใจได้ตรงกันทั้งผู้ถามและผู้ตอบ และเข้าใจผลที่วัดได้ตรงกันหรือให้คะแนนตรงกันไม่ว่าใครจะเป็นผู้ให้คะแนน ประเด็นสำคัญของเรื่องนี่คือการใช้ภาษาไทย โดยมากมักจะเป็นการใช้ภาษากำกวมทำให้การตีความหมายแตกต่างกัน ผู้วิจัยควรมีการทดลองใช้ในเบื้องต้นเสียก่อนเพื่อแก้ไขก่อนนำไปตรวจสอบความเที่ยงและความตรง

3.2) ความง่ายพอเหมาะ ที่จะนำไปใช้กับกลุ่มเป้าหมาย เช่น ถ้าคำถามนั้นให้กับชาวชนบท ลักษณะคำพูดต้องเป็นภาษาที่ชาวบ้านเข้าใจง่าย หรือเสียงศัพท์ทางวิชาการ เช่น คำว่า “บริโภคน” ก็ต้องเปลี่ยนเป็น “ดื่ม กิน หรือรับประทาน” แทน

3.3) ความยาวพอเหมาะ ความยาวของแบบสอบถามต้องเหมาะสมกับเวลาที่ควรใช้ ในขณะที่เดียวกันก็ต้องครอบคลุมเนื้อหาที่ต้องการวัด ประเด็นนี้ตรวจสอบได้โดยการทดลองใช้เช่นกัน

3.4) สามารถแบ่งแยกความแตกต่างในกลุ่มเป้าหมายที่ศึกษาได้ เนื่องจากเครื่องมือที่สร้างขึ้นต้องการวัดตัวแปร ผลที่ได้จากการวัดจึงต้องมีค่าที่แปรผันกันไป หากผลที่ได้จากการวัดมีค่าเดียวหรือมีค่าที่ใกล้เคียงกันมากจนไม่สามารถแยกความแตกต่างได้ เครื่องมือนั้นย่อมมีคุณภาพไม่ดีพอที่จะใช้วัดตัวแปร การตรวจสอบในเรื่องนี้ต้องอาศัยเทคนิคทางสถิติเข้ามาช่วย ได้แก่ การแจกแจงความถี่ การวัดการกระจาย t-test เป็นต้น

3.5) ลำดับของคำถามเหมาะสม ข้อคำถามแต่ละข้อต้องมีการจัดลำดับ เช่น การจัดลำดับตามเหตุการณ์ การจัดลำดับตามความง่าย เพื่ออำนวยความสะดวกแก่ผู้ตอบ และป้องกันความคลาดเคลื่อนเนื่องจากความสับสนของข้อคำถาม

ในการตรวจสอบคุณภาพเครื่องมือก็จะตรวจสอบคุณสมบัติทั้ง 3 ประการนี้ แต่ยังมีรายละเอียดแตกต่างกันไปบ้างในเครื่องมือแต่ละชนิด ดังนี้

1) การตรวจสอบคุณภาพของแบบสอบถาม

ความตรง : โดยทั่วไปจะตรวจสอบจากผู้เชี่ยวชาญ โดยดูว่าเนื้อหาในแบบสอบถามครอบคลุมวัตถุประสงค์ของการวิจัยหรือไม่ ข้อคำถามครบถ้วนตามประเด็นที่ต้องการหรือไม่

ความเที่ยง : ตรวจสอบโดยการไปทดลองใช้กับกลุ่มตัวอย่าง จากนั้นเว้นระยะแล้วถามซ้ำ หรือถ้าแบบสอบถามมีลักษณะเป็นมาตราประมาณค่า ก็สามารถคำนวณค่าความเที่ยงโดยใช้สูตรสัมประสิทธิ์แอลฟา

2) การตรวจสอบคุณภาพแบบสัมภาษณ์

ความตรง : ตรวจสอบได้เช่นเดียวกับแบบสอบถาม

ความเที่ยง : ใช้วิธีการสัมภาษณ์หลังจากที่เว้นช่วงไประยะหนึ่ง หรือใช้ผู้สัมภาษณ์หลายคน สัมภาษณ์คนคนเดียว แล้วพิจารณาความสอดคล้องของข้อมูลที่ได้จากการสัมภาษณ์

3) การตรวจสอบคุณภาพของแบบสังเกต

ความตรง : ตรวจสอบได้เช่นเดียวกับแบบสอบถาม

ความเที่ยง : ตรวจสอบได้ 2 วิธี คือ

วิธีที่ 1 ตรวจสอบความเที่ยงภายในผู้ประเมิน (Intra-rater Reliability) โดยให้ผู้สังเกตพฤติกรรมเดียวกัน ในเวลาที่ต่างกันแล้วนำข้อมูลที่ได้มาหาค่าสัมประสิทธิ์สหสัมพันธ์

วิธีที่ 2 ระหว่างผู้ประเมิน (Inter-rater Reliability) โดยให้ผู้สังเกตหลายคน สังเกตพฤติกรรมเดียวกันของคนคนเดียวกัน แล้วนำข้อมูลที่ได้มาหาค่าสัมประสิทธิ์สหสัมพันธ์

สรุป

ลักษณะสำคัญของเครื่องมือที่มีคุณภาพ มี 2 อย่าง คือ ความเที่ยง (Reliability) และ ความตรง (Validity)

ความเที่ยง หมายถึง ความคงที่ของคะแนนที่ได้จากการวัดด้วยเครื่องมือชุดเดียวกันในเวลาต่างกัน การประมาณค่าความเที่ยงของแบบสอบ โดยทั่วไปวิธีการประมาณค่าความเที่ยงมี 3 รูปแบบ คือ (1) แบบสอบซ้ำ (Measure of Stability) (2) แบบใช้ข้อสอบคล้ายกัน (Measure of Equivalence) (3) แบบวัดความคงที่ภายใน (Measure of Internal Consistency) ซึ่งเทคนิคทางสถิติที่ใช้ ได้แก่ วิธีแบ่งครึ่งข้อสอบ (Split-half) วิธีของคูเดอร์-ริชาร์ดสัน (Kuder-Richardson estimates) และวิธีสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's Coefficient)

ความตรง หมายถึง ความถูกต้องแน่นอนในการวัดสิ่งที่ต้องการวัด จำแนกตามการวัดทางการศึกษาและจิตวิทยา แบ่งความตรงออกเป็น 3 ประเภท คือ (1) ความตรงตามเนื้อหา (Content Validity) (2) ความตรงตามเกณฑ์สัมพันธ์ (Criterion-related Validity) ซึ่งแบ่งเป็นความตรงร่วมสมัย (Concurrent Validity) และความตรงเชิงทำนาย (Predictive Validity) (3) ความตรงเชิงโครงสร้างหรือทฤษฎี (Construct Validity)

คุณภาพของเครื่องมือในด้านอื่นๆ ที่ควรพิจารณา คือ ความเป็นปรนัย ความยากง่าย พอเหมาะที่จะนำไปใช้กับกลุ่มเป้าหมาย ความยาวพอเหมาะกับความยาวและเนื้อหาที่ต้องการวัด สามารถแบ่งแยกความแตกต่างในกลุ่มเป้าหมายที่ศึกษาได้ และลำดับของคำถามที่เหมาะสมทำให้ไม่เกิดความสับสนแก่ผู้ตอบ