

การออกแบบการประเมิน

การประเมินเป็นชุดกิจกรรม (Set) ของการดำเนินการเพื่อตัดสินคุณค่าของสิ่งต่างๆ ที่ต้องการประเมิน ได้แก่ องค์กร นโยบาย โครงการ กิจกรรม เป็นต้น สิ่งที่ผู้ประเมินต้องการคือ สารสนเทศที่เป็นประโยชน์ (useful information) เกี่ยวกับเป้าหมาย ความคาดหวัง ผลที่เกิดขึ้น และผลกระทบที่ตามมา การที่ผู้ประเมินจะได้สารสนเทศที่ตรงตามความต้องการ จำเป็นอย่างยิ่งที่จะต้องมีการออกแบบการประเมินที่เหมาะสมและถูกต้อง เนื่องจากออกแบบการประเมินเป็นขั้นตอนที่เชื่อมโยงจากประเด็นการประเมินไปสู่วิธีการประเมิน การออกแบบการประเมินจึงเป็นการกำหนดชนิด รูปแบบ ขอบเขต มาตรฐาน และแนวทางการประเมิน

การออกแบบการประเมิน เป็นการวางแผนการทำงานเกี่ยวกับการประเมิน เพื่อให้ได้ผลการประเมินที่ตอบประเด็นปัญหาในการประเมิน อีกทั้งเป็นการจัดสรรการใช้ทรัพยากรและกำลังคนอย่างมีประสิทธิภาพเพื่อให้เกิดประโยชน์สูงสุด

1. เกณฑ์ที่ควรพิจารณาในการออกแบบการประเมิน

เช่นเดียวกับการออกแบบการวิจัย การออกแบบการประเมินที่ดีจึงต้องคำนึงเกณฑ์ดังต่อไปนี้

1.1 ความตรงภายใน (Internal Validity) หมายถึง ลักษณะของการเปลี่ยนแปลงที่เกิดขึ้นของตัวแปรตามเป็นผลมาจากตัวแปรอิสระที่ทำการศึกษา ในกรณีนี้คือการเปลี่ยนแปลงหรือผลที่เกิดขึ้นจากโครงการ เป็นผลมาจากองค์กรและการดำเนินโครงการที่ต้องการประเมินนั่นเอง ทำให้ผู้ประเมินต้องควบคุมตัวแปรแทรกซ้อนและตัวแปรสอดแทรกที่มีผลต่อตัวแปรตามให้ได้เป็นอย่างดี นอกจากนี้ยังต้องมีวิธีการวิเคราะห์ข้อมูล แปลความหมาย และสรุปผลที่ได้อย่างถูกต้องและเหมาะสม จึงจะทำให้งานประเมินมีความตรงภายใน

1.2 ความตรงภายนอก (External Validity) หมายถึง ลักษณะของการสรุปผลไปยังกลุ่มเป้าหมายได้ การที่ผู้ประเมินสามารถสรุปผลการประเมินไปยังกลุ่มเป้าหมายที่เกี่ยวข้องครอบคลุมพื้นที่เป้าหมายของโครงการได้ แสดงว่าการประเมินมีความตรงภายนอก แต่ความตรงภายนอกของการประเมินจะต่างจากความตรงภายนอกของการวิจัยตรงที่ไม่สามารถสรุปอ้างอิงไปยังโครงการอื่นๆ ที่มีลักษณะใกล้เคียงกันได้ ทั้งนี้เพราะแต่ละโครงการจะมีบริบทและความสำคัญของโครงการแตกต่างกันออกไป

1.3 ความเป็นไปได้ในการประเมิน หมายถึง การออกแบบการประเมินจะต้องคำนึงถึงเวลาและสถานการณ์ที่จะทำการประเมิน ตลอดจนทรัพยากรที่มีอยู่ ไม่ว่าจะเป็นงบประมาณหรือกำลังคน สิ่งสำคัญอีกข้อคือ ผลการประเมินต้องทันเวลาที่จะนำผลการประเมินไปใช้ในการตัดสินใจ

2. กระบวนการที่เกี่ยวข้องกับการออกแบบการประเมิน

ดังที่กล่าวแล้วว่าการออกแบบการประเมิน เป็นขั้นตอนที่เชื่อมโยงประเด็นปัญหาการประเมิน ไปสู่ขั้นตอนดำเนินการประเมิน กระบวนการที่เกี่ยวข้องจึงประกอบด้วย

2.1 ประเด็นปัญหาของการประเมิน

2.2 การเลือกใช้รูปแบบการประเมิน

2.3 การกำหนดวิธีการประเมิน ซึ่งประกอบด้วย

(1) การสุ่มตัวอย่าง

(2) การพัฒนาตัวชี้วัด

(3) การกำหนดเกณฑ์หรือมาตรฐานของตัวชี้วัดที่แสดงถึงความสำเร็จของ
โครงการที่ต้องการประเมิน

(4) การกำหนดแหล่งของข้อมูลและการเลือกใช้เครื่องมือในการรวบรวมข้อมูล

(5) การวิเคราะห์ข้อมูล

2.1 ประเด็นปัญหาของการประเมิน

ในการวิจัยทางสังคมศาสตร์ ขั้นตอนแรกที่ต้องดำเนินการคือการกำหนด**ปัญหาการวิจัย** ซึ่งหมายถึง ประเด็นที่นักวิจัยสงสัยและต้องดำเนินการเพื่อหาคำตอบที่ถูกต้องตรงกับความเป็นจริง ทำให้มีลักษณะเป็นข้อสงสัยของผู้วิจัยต่อสถานการณ์ที่คาดหวังหรือสถานการณ์ที่ควรจะเป็นหากผู้วิจัยเห็นว่าควรหาคำอธิบายสถานการณ์ใด ก็สามารถนำมาเป็นประเด็นปัญหาการวิจัยได้ ในการประเมินก็เช่นเดียวกัน ผู้ประเมินจะต้องกำหนดประเด็นที่ต้องการจะประเมินเป็นขั้นตอนแรก ลักษณะของประเด็นการประเมินต่างจากประเด็นปัญหาในการวิจัย คือไม่ใช่ข้อสงสัยของผู้ประเมิน **แต่เป็นข้อสงสัยของผู้ใช้ผลการประเมิน** การกำหนดประเด็นที่ต้องการประเมินทำได้ 4 วิธี คือ

1) พิจารณาจาก**วัตถุประสงค์ของโครงการ** ว่ามีความประสงค์อย่างไรต่อผลที่ได้จากการดำเนินการตามโครงการ จากนั้นนำวัตถุประสงค์เหล่านี้ไปกำหนดเป็นประเด็นในการประเมิน วิธีการนี้จะได้ประเด็นปัญหาเฉพาะที่มีอยู่ในวัตถุประสงค์ของโครงการเท่านั้น

2) พิจารณาจาก**ความต้องการของผู้ใช้ผลจากการประเมิน** ได้แก่ ผู้ให้ทุน ผู้มีหน้าที่จัดทำนโยบาย ผู้บริหารระดับสูง ผู้บริหารระดับต้น เจ้าหน้าที่ปฏิบัติการ ลักษณะของสิ่งที่บุคคลแต่ละระดับต้องการทราบจะแตกต่างกันออกไป เช่น เจ้าหน้าที่ปฏิบัติการต้องการทราบเพียงว่าผลที่ได้ (output) เป็นไปตามเป้าหมายหรือไม่ ในขณะที่ผู้บริหารระดับสูงอาจต้องการทราบมากกว่าผลที่ได้ เช่น อาจต้องการทราบถึงคุณภาพของการดำเนินโครงการตั้งแต่ต้นจนจบ ความพอใจของผู้รับบริการ โครงการนั้น ตลอดจนจนผลกระทบที่ไม่คาดหวังอันเนื่องมาจากโครงการ วิธีการนี้บุคคลที่เกี่ยวข้องทั้งหมดสามารถใช้ผลการประเมินในคราวเดียวกัน

3) พิจารณาจาก**ประสบการณ์ของผู้ประเมิน** หากผู้ประเมินมีประสบการณ์ด้านการประเมินมาก จะทำให้สามารถระบุถึงประเด็นการประเมินในโครงการแต่ละประเภทได้ด้วยตนเองได้อย่างรวดเร็ว จากนั้นจึงขอความเห็นชอบของผู้ใช้ผลการประเมินอีกครั้งหนึ่ง วิธีการนี้ทำให้ประหยัดเวลาในการกำหนดประเด็นการประเมินได้มาก

4) พิจารณาโดยอาศัย**โมเดลหรือแบบจำลอง (model)** หรือแนวคิดที่นักวิชาการด้านการประเมินได้กำหนดขึ้น ผู้ประเมินจะต้องพิจารณาความสอดคล้องระหว่างโมเดลและโครงการที่ต้องการประเมิน เช่น ในการฝึกอบรมหลักสูตรต่างๆ แบบจำลองที่ได้รับความนิยมในการประเมินการฝึกอบรม คือ แบบจำลองของ Kirkpatrick ในการสร้างและบริหารโครงการแบบจำลองที่ได้รับความนิยมในการประเมินโครงการคือแบบจำลอง CIPP ของ Stufflebeam และคณะ เป็นต้น

2.2 การเลือกใช้ชนิดและรูปแบบการประเมิน

1) ชนิดของการประเมิน (Evaluation Type) สามารถแบ่งได้ตามช่วงเวลาของการประเมินได้แก่ การประเมินความต้องการจำเป็น (Needs Assessment) การประเมินความเป็นไปได้ (Feasibility Study) เป็นต้น ผู้ประเมินจะต้องพิจารณาโครงการที่จะประเมินว่าอยู่ในช่วงเวลาใดของการดำเนินโครงการ

2) รูปแบบของการประเมิน (Evaluation Design) ในจำนวนรูปแบบการวิจัยทั้งหมด การวิจัยเชิงทดลองเป็นเพียงแบบเดียวที่สามารถบอกความเป็นเหตุเป็นผลได้ชัดเจนที่สุด แต่ด้วยข้อจำกัดทางสังคมและพฤติกรรมของมนุษย์ ทำให้การใช้การวิจัยเชิงทดลองในการประเมินยังทำได้ไม่มากนัก รูปแบบที่ใช้ในการวิจัยเชิงทดลอง (Experimental Design) แบ่งได้เป็น 3 กลุ่ม ตามระดับการควบคุมตัวแปรแทรกซ้อน คือ

2.1) Pre-experimental Designs เป็นแบบการทดลองที่ใช้กับกลุ่มตัวอย่างเดียว โดยไม่มีการสุ่มตัวอย่าง ทำให้มีตัวแปรแทรกซ้อนที่เกิดจากกลุ่มตัวอย่าง และไม่สามารถสรุปได้ชัดเจนว่าผลที่ได้เป็นผลมาจากตัวแปรจัดกระทำ (treatment) ในการทดลองหรือไม่

2.2) Quasi-experimental Designs เป็นแบบการทดลองที่ปรับปรุงจากแบบที่ 2.1) โดยใช้กลุ่มตัวอย่าง 2 กลุ่ม กลุ่มหนึ่งเป็นกลุ่มควบคุมที่ไม่ได้รับการจัดกระทำ (treatment) แต่มีการวัดผลในระยะเวลาเดียวกันกับกลุ่มที่ได้รับการจัดกระทำ (treatment) ทำให้เห็นผลของการทดลองชัดเจนขึ้น แต่ยังไม่สามารถควบคุมตัวแปรแทรกซ้อนที่เกิดจากกลุ่มตัวอย่าง

2.3) True-experimental Designs เป็นแบบการทดลองที่ปรับปรุงมาจากแบบที่ 2.2) มีกลุ่มควบคุมและมีการสุ่มที่เรียกว่า Randomization ซึ่งประกอบด้วยการสุ่มตัวอย่างจากประชากรที่เรียกว่า การสุ่มกลุ่มเลือกตัวอย่าง (Randomization Selection) และการสุ่มตัวอย่างเข้ารับ

การจัดกระทำ ที่เรียกว่า Random Assignment การสุ่มทั้งสองขั้นตอนจะเป็นการควบคุมตัวแปรแทรกซ้อนที่เกิดจากกลุ่มตัวอย่างได้เป็นอย่างดี

จากทั้ง 3 รูปแบบการประเมิน ยังสามารถนำแนวคิดเรื่องการศึกษาตามช่วงระยะเวลาเข้าร่วมเป็นรูปแบบ ซึ่งเป็นวิธีการที่นิยมในการศึกษาแนวโน้ม (Trend) วิธีการที่นำมาใช้กันมากคือ Time Series ลักษณะการดำเนินการเหมือนเดิมเพียงเพิ่มจำนวนครั้งในการวัดผลมากขึ้น ในเอกสารนี้จะนำเสนอกลุ่มละ 3 รูปแบบ ดังรายละเอียดในตารางที่ 1

ตารางที่ 1 แบบการทดลองจำแนกตามระดับการควบคุมตัวแปรแทรกซ้อนในการวิจัยเชิงทดลอง

Pre-experimental Designs	Quasi-experimental Designs	True-experimental Designs
A. ศึกษาจากกลุ่มตัวอย่างเดียว	B. ใช้ 2 กลุ่มตัวอย่าง โดยมีกลุ่มหนึ่งเป็นกลุ่มควบคุม	C. ใช้ 2 กลุ่มตัวอย่างมีกลุ่มควบคุมและมี Randomization
a) มีการวัดผลช่วงเวลาเดียว $\times O_2$ THE SINGLE GROUP POSTTEST-ONLY DESIGN	$\times O_2$ $\cdot O_c$ NON-EQUIVALENT CONTROL GROUP POSTTEST-ONLY DESIGN	$R \times O_2$ $R \cdot O_c$ TRUE CONTROL GROUP POSTTEST-ONLY DESIGN
b) มีการวัดผล 2 ครั้ง ก่อนและหลังการทดลอง $O_1 \times O_2$ THE SINGLE GROUP PRETEST-POSTTEST DESIGN	$O_1 \times O_2$ $O_{c1} \times O_{c2}$ NON-EQUIVALENT CONTROL GROUP PRETEST-POSTTEST DESIGN	$R O_1 \times O_2$ $R O_{c1} \times O_{c2}$ TRUE CONTROL GROUP PRETEST-POSTTEST DESIGN
c) มีการวัดผลหลายช่วงเวลา $O_1 O_2 O_3 \times O_4 O_5 O_6$ THE SINGLE GROUP PRETEST-POSTTEST TIME SERIES DESIGN	$O_1 O_2 O_3 \times O_4 O_5 O_6$ $O_{c1} O_{c2} O_{c3} \cdot O_{c4} O_{c5} O_{c6}$ NON-EQUIVALENT CONTROL GROUP PRETEST-POSTTEST TIME SERIES DESIGN	$R O_1 O_2 O_3 \times O_4 O_5 O_6$ $R O_{c1} O_{c2} O_{c3} \cdot O_{c4} O_{c5} O_{c6}$ TRUE CONTROL GROUP PRETEST-POSTTEST TIME SERIES DESIGN

X = ได้รับ treatment

O = ไม่ได้รับ treatment

การเลือกใช้รูปแบบใดขึ้นอยู่กับโครงการที่ต้องการประเมิน การควบคุมตัวแปรแทรกซ้อนและระยะเวลาที่ต้องการประเมิน

2.3) การกำหนดวิธีการประเมิน

(1) **การสุ่มตัวอย่าง** ในการประเมินจะพบอยู่เสมอว่าประชากรมีขนาดใหญ่มากและเกินกำลังที่ผู้ประเมินจะทำการศึกษาได้ ทำให้จำเป็นต้องศึกษาเพียงส่วนหนึ่งของประชากร ส่วนหนึ่งของประชากรที่ศึกษาจะต้องมีคุณสมบัติเป็นตัวแทนที่ดี จึงมีเทคนิคที่นำมาใช้คัดเลือกตัวอย่างในการประเมินเรียกว่า **การสุ่มตัวอย่าง (sampling)** หลักการที่สำคัญของการสุ่มตัวอย่างคือ เลือกวิธีการสุ่มตัวอย่างที่เหมาะสมเพื่อให้ได้ตัวแทนที่ดีของประชากรที่ศึกษาและขนาดของกลุ่มตัวอย่างต้องมีขนาดพอเหมาะทั้งในด้านทฤษฎีและในด้านการปฏิบัติ แบ่งวิธีการสุ่มตัวอย่างออกได้ดังนี้

(1.1) อาศัยหลักการเกี่ยวกับความน่าจะเป็น (Probability Sampling)

(1.1.1) **Simple Random Sampling** เป็นการสุ่มตัวอย่างเมื่อประชากรมีลักษณะใกล้เคียงกัน วิธีที่นิยมใช้กันมากคือ การจับฉลาก

(1.1.2) **Systematic Sampling** เป็นการสุ่มตัวอย่างเมื่อประชากรมีลักษณะใกล้เคียงกัน แต่ลักษณะของประชากรมีการจัดเรียงตามลำดับหมายเลขไว้แล้ว การสุ่มจึงสามารถทำอย่างเป็นระบบโดยการนำจำนวนประชากรหารด้วยจำนวนตัวอย่างที่ต้องการ

(1.1.3) **Stratified Sampling** เป็นการสุ่มตัวอย่างเมื่อผู้ประเมินพบว่าประชากรมีความแตกต่างอย่างชัดเจน กลุ่มตัวอย่างที่เป็นตัวแทนที่ดีของประชากรจึงต้องประกอบด้วยสมาชิกของกลุ่มย่อยเหล่านั้นทุกกลุ่ม ผู้ประเมินจะเริ่มด้วยการแบ่งประชากรออกเป็นกลุ่มตามความแตกต่าง จากนั้นสุ่มตัวอย่างในแต่ละกลุ่มย่อยประชากรตามสัดส่วนของประชากรในแต่ละกลุ่ม

(1.1.4) **Cluster Sampling** ในบางประชากรจะพบลักษณะการรวมเป็นก้อน ทำให้การศึกษาตัวแปรต้องศึกษาจากสมาชิกในกลุ่มย่อยๆ เหล่านั้นทั้งหมด เช่น การศึกษาผลของวิธีการสอนคณิตศาสตร์ จะต้องเป็นวิธีการสอนที่เหมาะสมกับนักเรียนทั้งห้องไม่ใช่วิธีที่ได้ผลดีกับนักเรียนคนใดคนหนึ่งเท่านั้น การสุ่มต้องสุ่มทีละห้องหรือการศึกษาแบบการพัฒนาชนบท จะต้องเป็นรูปแบบที่ใช้ได้ผลกับคนทั้งหมู่บ้าน การสุ่มจึงต้องสุ่มทีละหมู่บ้าน เพื่อให้ได้สมาชิกที่ครบถ้วน

(1.1.5) **Multi-stage Sampling** ในหลายกรณีพบว่าประชากรมีคุณลักษณะที่ซับซ้อนทำให้ต้องมีการสุ่มตัวอย่างมากกว่า 1 ครั้ง โดยจะใช้วิธีการสุ่มที่เหมือนกันหรือไม่ก็ได้ เช่น การสุ่มตัวอย่างนักเรียน ต้องเริ่มสุ่มจากภาคทางการศึกษา ขนาดโรงเรียน การอยู่ในชุมชนชั้นเรียนไปจนถึงห้องเรียน

การสุ่มตัวอย่างโดยทั่วไปจะเป็นการสุ่มคนหรือครัวเรือน แต่ในบางกรณีผู้ประเมินสามารถสุ่มตัวแปรหรือตัวบ่งชี้ให้แก่กลุ่มตัวอย่าง โดยแต่ละคนจะได้รับการประเมินเพียงบางตัวบ่งชี้เท่านั้น ตัวอย่างที่เห็นได้ชัดคือ การสุ่มข้อสอบให้กับกลุ่มผู้สอบเนื่องจากจำนวนข้อสอบมีมากเกินไป ผู้สอบไม่สามารถทำข้อสอบได้ทั้งหมด จึงมีการสุ่มที่สุ่มทั้งข้อสอบและผู้สอบที่เรียกว่า Matrix Sampling โดยผู้สอบแต่ละคนจะทำข้อสอบไม่เหมือนกัน แต่ข้อสอบทั้งหมดได้รับการทดลองใช้

การสุ่มตัวอย่างที่ใช้ความน่าจะเป็นเป็นการกำหนดขนาดกลุ่มตัวอย่างที่เหมาะสม ทำให้ผู้ประเมินสามารถใช้สถิติอนุมานสรุปอ้างอิงผลการประเมินจากกลุ่มตัวอย่างกลับไปยังประชากรทั้งหมดได้ แต่ในบางกรณีผู้ประเมินไม่สามารถใช้วิธีดังกล่าวข้างต้น เพราะมีข้อจำกัดต่างๆเข้ามาเกี่ยวข้อง จึงต้องหันมาใช้ในการเลือกตัวอย่างอีกแบบหนึ่งที่ไม่ใช้ความน่าจะเป็น

(1.2) **ไม่อาศัยความน่าจะเป็น** เรียกว่า Non-probability Sampling ซึ่งในตำราจะใช้ว่า “Selection” แทนคำว่า “Sampling” ที่น่าสนใจมี 3 วิธีคือ

(1.2.1) **Purposive Selection** เป็นการเลือกตัวอย่างที่เจาะจงเพื่อให้เหมาะสมกับปัญหาการประเมินนั้นๆ

(1.2.2) **Accidental Selection** เป็นการเลือกตัวอย่างในลักษณะการบังเอิญพบ

(1.2.3) **Quota Selection** เป็นการกำหนดสัดส่วนของกลุ่มตัวอย่างที่มีความแตกต่างกันไว้ล่วงหน้า โดยไม่เป็นไปตามสัดส่วนในประชากร

วิธีการในชุดนี้มักจะได้สะดวก และเมื่อมีเวลาจำกัดในการศึกษาหรือต้องการศึกษา หรือการวิเคราะห์ที่มีลักษณะเป็นการศึกษาเฉพาะกรณี (Case study) การสรุปผลจะทำได้เฉพาะกลุ่มที่ทำการศึกษา หรือสรุปในกลุ่มประชากรที่มีสภาพการณ์ที่มีเงื่อนไขเช่นเดียวกันเท่านั้น

การกำหนดขนาดกลุ่มตัวอย่างต้องพิจารณาถึงขนาดประชากร ลักษณะความแตกต่างของประชากร ขนาดความคลาดเคลื่อนสูงสุดยอมรับได้ ระดับของความเชื่อมั่น ชนิดของพารามิเตอร์ที่ต้องการทดสอบ งบประมาณ และวิธีที่ใช้เก็บรวบรวมข้อมูล อย่างไรก็ตามในทางปฏิบัติมีข้อจำกัดเกิดขึ้นมากมาย เช่น การขาดงบประมาณ การขาดกำลังคน ความจำกัดของเวลา และความร่วมมือของกลุ่มตัวอย่าง ทำให้ผู้ประเมินไม่สามารถใช้ขนาดกลุ่มตัวอย่างที่คำนวณได้ ทำให้ผู้ประเมินต้องย้อนกลับมาพิจารณาถึงขนาดกลุ่มตัวอย่างที่มีความเหมาะสมในทางปฏิบัติด้วย โดยอาจจะต้องยอมให้มีความคลาดเคลื่อนเพิ่มขึ้นเพื่อลดจำนวนของกลุ่มตัวอย่างลง ทำให้การสรุปผลการประเมินกลับไปยังประชากรไม่ได้มากเท่าที่ผู้ประเมินต้องการ

(2) **การพัฒนาตัวชี้วัด** เมื่อผู้ประเมินได้ประเด็นที่ต้องการประเมินแล้ว งานขั้นต่อไปก็คือ การกำหนดตัวแปรในประเด็นแต่ละประเด็นซึ่งอาจมีตัวแปรมากกว่า 1 ตัวแปรก็ได้ เช่น

* การอยู่ดีกินดีของชาวชนบท ตัวแปรที่เกี่ยวข้อง ได้แก่ รายได้ รายจ่าย สิ่งอำนวยความสะดวก สุขภาพอนามัย เป็นต้น

* ความรู้ความเข้าใจของชาวชนบทเกี่ยวกับสัญลักษณ์ที่จำเป็น ตัวแปรที่เกี่ยวข้องในประเด็นนี้มีตัวเดียว คือ ความรู้ความเข้าใจเกี่ยวกับสัญลักษณ์ที่จำเป็น

ตัวแปรที่ถูกระบุแต่ละประเด็นของการประเมินจะทำหน้าที่เป็น ตัวชี้วัด ถึงลักษณะของโครงการนั้นๆ ถ้าการประเมินเป็นการตัดสินคุณค่าของโครงการแล้ว การระบุตัวแปรเพียงอย่างเดียวยังไม่สามารถตัดสินคุณค่าได้หากไม่มีการกำหนดระดับที่น่าพอใจของตัวแปรเสียก่อน ดังนั้นในการประเมิน ตัวชี้วัดจะเกิดจากตัวแปรที่มีการกำหนดเกณฑ์ และระดับที่พึงประสงค์ของตัวแปรนั้น

สิ่งที่จะถูกประเมินในการประเมินยังมีอีกหลายประเภท ไม่เพียงแต่โครงการเท่านั้น ได้แก่ นโยบาย องค์กร วัสดุอุปกรณ์ อาคารสถานที่ กิจกรรม การระบุประเด็นที่ต้องการประเมินจึงแตกต่างกันไปตามสิ่งที่ประเมิน

(3) การกำหนดเกณฑ์ หรือ มาตรฐาน ของตัวบ่งชี้ที่แสดงถึงความสำเร็จของโครงการที่ต้องการประเมิน เกณฑ์ที่นิยมใช้ในการประเมินมี 2 ลักษณะ คือ

(3.1) เกณฑ์สัมบูรณ์ หรือ มาตรฐาน เป็นกำหนดระดับที่ควรจะมี ควรจะเป็น หรือควรจะได้จากโครงการ การกำหนดมาตรฐานอาจทำได้โดย

(3.1.1) มาตรฐานสากล

(3.1.2) ผู้เชี่ยวชาญในเรื่องที่ต้องการประเมิน มาตรฐานที่ได้อาจมีความลำเอียงเนื่องจากบุคลิกภาพของผู้เชี่ยวชาญได้

(3.1.3) การคาดหวังจากเหตุการณ์ที่ผ่านมา หรืออาจใช้ regression analysis เพื่อประมาณค่าที่คาดหวัง แต่ไม่ควรใช้ผลงานที่ผ่านมาโดยตรงมากำหนดเป็นมาตรฐาน

(3.1.4) เปรียบเทียบระหว่างกลุ่มในแต่ละช่วงเวลา แม้ว่าการเปรียบเทียบจะเป็นวิธีการที่เข้าใจง่าย ปัญหาอยู่ที่ว่าขนาดความแตกต่างเท่าใดจึงจะมีความหมาย

(3.1.5) การใช้เกณฑ์ปกติวิสัย (norms) ที่ใช้กันอยู่ทั่วไป เช่น เปอร์เซนไทล์

ที่ 50

(3.1.6) การใช้ Evaluator's Program Description (EPD) เป็นการพัฒนาตัวบ่งชี้ควบคุมคู่ กับการกำหนดมาตรฐานที่คาดหวังในคราวเดียวกัน

(3.2) เกณฑ์สัมพัทธ์ ในบางกรณีผู้ประเมินไม่สามารถกำหนดเกณฑ์สัมบูรณ์ได้ จำเป็นต้องเทียบเคียงจากโครงการที่มีลักษณะใกล้เคียงกัน และเป็นโครงการที่จัดว่าประสบความสำเร็จ ที่สำคัญจะต้องเป็นโครงการที่เป็นที่ยอมรับกันว่ามีคุณภาพเหมาะสม

(4) การกำหนดแหล่งข้อมูลและการเลือกใช้เครื่องมือ ในการเก็บรวบรวมข้อมูล แหล่งข้อมูล ได้แก่

- (4.1) บันทึกต่างๆ ของโครงการ (program record) เป็นข้อมูลที่เสียค่าใช้จ่ายต่ำ ไม่มีผลกระทบอันเนื่องมาจากการเก็บข้อมูลในการประเมิน และไม่มีปัญหาเรื่องความร่วมมือ
- (4.2) ผู้มีส่วนร่วมในโครงการ ได้แก่ ผู้รับบริการ และผู้ที่เกี่ยวข้องอื่นๆ ในโครงการ
- (4.3) ทีมงาน มีข้อบกพร่องตรงที่จะมีความลำเอียงในการตอบที่เป็นผลดีต่อโครงการ
- (4.4) การสังเกตของผู้ประเมิน ในลักษณะนี้คุณภาพของข้อมูลขึ้นอยู่กับประสบการณ์ของผู้ประเมิน
- (4.5) ดัชนีที่มีอยู่ในชุมชน เช่น อัตราการเกิด ดัชนีผู้บริโภค สถิติคดีอาชญากรรม
- (4.6) ผู้ได้รับผลจากโครงการ หรือกลุ่มเป้าหมายของโครงการ

(5) การวิเคราะห์ข้อมูล การใช้สถิติในการวิเคราะห์ข้อมูลการประเมินมี 2 วิธี คือ

(5.1) **สถิติบรรยาย** (Descriptive Statistics) เป็นสถิติที่ใช้ในการบรรยายหรืออธิบายลักษณะต่างๆ ของกลุ่มตัวอย่างหรือประชากรที่ใช้ในการศึกษาเท่านั้น ไม่สนใจที่จะสรุปอ้างอิงไปยังประชากรอื่น วิธีการทางสถิติประเภทนี้ที่ใช้ ได้แก่ การแจกแจง ความถี่ การจัดตำแหน่ง เปรียบเทียบ การวัดแนวโน้มเข้าสู่ส่วนกลาง การวัดการกระจาย การวัดความสัมพันธ์ การใช้สถิติบรรยายจะใช้ใน 2 ส่วน คือ

(5.1.1) **ใช้ในการตรวจสอบคุณภาพของเครื่องมือ** โดยขึ้นกับชนิดของเครื่องมือที่ใช้ ถ้าเป็นแบบสอบถาม (questionnaire) ทั่วไป จะมีการประมาณค่าความเที่ยง และความตรงด้วยสัมประสิทธิ์ความเที่ยง และสัมประสิทธิ์ความตรงชนิดต่างๆ แต่ถ้าเป็นแบบสอบ (test) จะต้องมีการวิเคราะห์คุณภาพอย่างอื่นนอกเหนือจากการประมาณค่าความเที่ยงและความตรง ได้แก่ ค่าความยาก อำนาจจำแนก การทำงานของตัวลวง เป็นต้น

(5.1.2) **ใช้ในการวิเคราะห์ข้อมูลและแปลความหมาย** ซึ่งเป็นขั้นตอนหลังจากที่ผู้ประเมินได้เก็บรวบรวมข้อมูลมาแล้ว มีการรวบรวมหมวดหมู่ จัดกระทำให้เกิดสารสนเทศ

(5.2) **สถิติอนุมาน** (Inferential Statistics) เป็นสถิติที่ใช้ในการสรุปอ้างอิงค่าสถิติต่างๆ ที่เกิดขึ้นในกลุ่มตัวอย่างที่ศึกษาไปยังกลุ่มประชากรของกลุ่มตัวอย่งนั้น ซึ่งจำเป็นจะต้องมีการสุ่มตัวอย่างที่ถูกต้องและมีขนาดของกลุ่มตัวอย่างที่เหมาะสม โดยเป็นการอนุมานหรือสรุปอ้างอิง

จากค่าสถิติ (statistic) ของกลุ่มตัวอย่างไปยังค่าพารามิเตอร์ (parameter) ของประชากรก็ได้ ต้องมีการตรวจสอบความถูกต้องของข้อมูลก่อนที่จะทำการวิเคราะห์ การตรวจสอบอาจใช้เจ้าหน้าที่หรือโปรแกรมสำเร็จรูปที่ใช้กับคอมพิวเตอร์ หรืออาจจะใช้ทั้ง 2 วิธีด้วยกัน เพื่อให้เกิดความแม่นยำ การตรวจสอบความถูกต้องนอกจากจะมีผลดีต่อการทำการวิเคราะห์ข้อมูล ยังเป็นการตรวจสอบจำนวนข้อมูลว่าได้ครบตามจำนวนที่กำหนดไว้หรือผิดพลาดไม่ครบจะได้มีการแก้ไขต่อไป

3. ตารางแสดงรายละเอียดของการออกแบบการประเมิน

สมมติว่าผู้ประเมินต้องการประเมินโครงการ “การจัดอาหารเสริม (นม) ให้แก่เด็กเล็กในจังหวัด ก.” ซึ่งโครงการมีวัตถุประสงค์คือ 1) จัดหาอาหารเสริม (นม) ให้แก่เด็กเล็กในจังหวัด ก. และ 2) พัฒนาเด็กเล็กในจังหวัด ก. ให้มีสุขภาพสมบูรณ์

ผู้ประเมินได้วิเคราะห์โครงการและสอบถามความต้องการของผู้บริหารแล้ว ได้กำหนดประเด็นและวัตถุประสงค์ของการประเมินดังนี้

ประเด็นของการประเมิน

- 1) การจัดหาอาหารเสริม (นม) ให้แก่เด็กในศูนย์พัฒนาเด็กเล็กของจังหวัด “ก”
- 2) การพัฒนาเด็กเล็กในจังหวัด “ก” ให้มีสุขภาพที่สมบูรณ์
- 3) ความคิดเห็นของเจ้าหน้าที่ผู้ปฏิบัติงาน
- 4) ความคิดเห็นของประชาชนในพื้นที่โครงการ

วัตถุประสงค์ของการประเมิน

- 1) เพื่อประเมินความเพียงพอและความทั่วถึงของการจัดอาหารเสริม (นม) ให้แก่เด็กในศูนย์พัฒนาเด็กเล็กของจังหวัด “ก”
- 2) เพื่อประเมินผลการพัฒนาเด็กเล็กในจังหวัด “ก”
- 3) เพื่อประเมินความคิดเห็นของเจ้าหน้าที่และประชาชนในโครงการ สำหรับประเด็นแรกของการประเมิน การจัดหาอาหารเสริม (นม) ให้แก่เด็กในศูนย์พัฒนาเด็กเล็กของจังหวัด “ก” ผู้ประเมินได้ออกแบบการประเมิน ดังนี้

ตารางที่ 2 ตัวอย่างตารางการออกแบบการประเมิน

ประเด็น	ตัวชี้วัด	ข้อมูลที่ต้องการ	เครื่องมือ	การเก็บข้อมูล	การวิเคราะห์ข้อมูล	เกณฑ์
จัดหาอาหารเสริม (นม) ให้แก่เด็กเล็กของจังหวัด ก	1. สัดส่วนศูนย์ที่ได้รับอาหารเสริมจากทั้งหมด 40 ศูนย์	1. จำนวนศูนย์ที่ได้รับอาหารเสริม	1. แบบรายงานผลการปฏิบัติงาน	1. เก็บข้อมูลทุกสัปดาห์	1. คำนวณสัดส่วนที่ได้รับจากจำนวนทั้งหมด	1. ไม่ต่ำกว่าร้อยละ 90
	2. สัดส่วนเด็กที่ได้รับอาหารเสริมจากเด็กทั้งหมด 1,750 คน	2. จำนวนเด็กที่ได้รับอาหารเสริม	2. แบบรายงานผลการปฏิบัติงาน	2. เก็บข้อมูลทุกสัปดาห์	2. คำนวณสัดส่วนที่ได้รับจากจำนวนทั้งหมด	2. ไม่ต่ำกว่าร้อยละ 90

สรุป

การออกแบบการประเมินเป็นขั้นตอนที่เชื่อมโยงระหว่างประเด็นการประเมินและตัวชี้วัดไปยังการดำเนินการเก็บรวบรวมข้อมูลที่ต้องการ นำข้อมูลดังกล่าวมาทำการวิเคราะห์เพื่อตอบประเด็นที่ต้องการประเมิน สิ่งที่เกี่ยวข้องกับการออกแบบการประเมิน คือ ประเด็นปัญหาของการประเมิน การเลือกรูปแบบการประเมิน และการกำหนดวิธีการประเมิน ซึ่งประกอบด้วย การสุ่มตัวอย่าง การพัฒนาตัวชี้วัด การกำหนดเกณฑ์หรือมาตรฐานของตัวชี้วัดที่แสดงถึงความสำเร็จของโครงการที่ต้องการประเมิน การกำหนดแหล่งของข้อมูล การเลือกใช้เครื่องมือในการเก็บรวบรวมข้อมูลและการวิเคราะห์ข้อมูล