

## บทที่ 9

### การกำหนดรูปแบบในการวัดผล

ในการวัดผลใด ๆ จะเชื่อได้อย่างไรว่าผลที่ได้จากการวัด เป็นสิ่งที่เกิดเนื่องมาจากโครงการทางสังคมที่เข้าไปประเมิน ผลที่วัดได้อาจเกิดเนื่องมาจากโครงการเฉพาะหนึ่ง ๆ หรือเป็นผลเนื่องมาจากโครงการทางสังคมหลาย ๆ โครงการร่วมกัน การกำหนดระเบียบวิธีที่เข้มงวดในการวัดผลจะทำให้นักประเมินสามารถสรุปได้ว่า ผลที่เกิดจากการวัดเป็นผลเนื่องมาจากโครงการจริงหรือไม่

จากการนำระเบียบวิธีวิจัยที่ว่าด้วยการวัดผลมาใช้ในการวัดผลของโครงการทางสังคม อาจแบ่งวิธีการวัดผลจากแบบที่เข้มงวดที่สุด ไปสู่แบบที่เข้มงวดน้อยที่สุด ได้ดังนี้

- 1) การวัดผลด้วยวิธีการทดลอง (Experimental Design)
- 2) การวัดผลด้วยวิธีการกึ่งการทดลอง (Quasi-experimental Designs)
- 3) การวัดผลด้วยวิธีไม่ใช้การทดลอง (Nonexperimental Designs)

นอกจากนี้ในบทนี้จะกล่าวถึงการประเมินผลในระดับมหภาค หรือที่มักใช้ในการประเมินระดับแผนงาน โดยจะนำเสนอการประเมินผลด้วยวิธีการเปรียบเทียบโครงการ (Comparative Evaluation of Programs)

ในส่วนท้ายจะกล่าวถึงการวิเคราะห์ต้นทุน และประโยชน์ (Cost Benefit Analysis) เป็นอีกทางเลือกหนึ่งในการประเมินผลกรณีที่ต้องการศึกษาประสิทธิภาพของการใช้จ่ายเงิน

#### การประเมินผลด้วยวิธีการทดลอง

การประเมินผลด้วยวิธีการทดลองจัดเป็นวิธีการประเมินผลที่มีความน่าเชื่อถือมากที่สุด ทั้งนี้เพราะมีระเบียบวิธีที่เข้มงวด มีการควบคุมความผิดพลาดในการตีความโดยอาศัยระเบียบวิธีทางสถิติ กล่าวได้ว่า ผลลัพธ์ที่ได้จากการประเมินด้วยวิธีทดลองจะทำให้ผู้ประเมินเชื่อมั่นได้เลยว่า ผลดังกล่าวเกิดขึ้นเนื่องมาจากตัวโครงการมิใช่ปัจจัยภายนอกอื่นเป็นตัวทำให้เกิด

การประเมินผลด้วยวิธีทดลองจะประกอบไปด้วยกลุ่มสองกลุ่ม คือกลุ่มทดลอง และกลุ่มควบคุม กลุ่มทั้งสองจะถูกกำหนดขึ้นมาด้วยวิธีการสุ่มตัวอย่าง (Random Sampling) กลุ่มทดลองจะเป็นกลุ่มที่ได้รับโครงการที่ต้องการประเมินผล ส่วนกลุ่มควบคุมจะจัดเป็นกลุ่มที่ใช้ในการเปรียบเทียบ อาจได้รับการปฏิบัติแบบเดิม หรืออาจได้รับโครงการอื่น ๆ ที่ต้องการนำมาเปรียบเทียบ ความก้าวหน้าของกลุ่มทดลองที่มีเหนือกลุ่มควบคุมจะเป็นเครื่องบ่งชี้ให้เห็นถึงความสำเร็จของโครงการ ในการติดตามสังเกตการเปลี่ยนแปลงของกลุ่มทั้งสองผู้ประเมินจะต้องควบคุมการได้รับการปฏิบัติของกลุ่มทั้งสองอย่างเข้มงวด จะต้องไม่มีการลาออกจากกลุ่ม ไม่มีการแลกเปลี่ยนความรู้ระหว่างกลุ่ม รวมทั้งโครงการที่ได้รับการประเมินก็ต้องดำเนินงานไปตามแผนงานที่วางไว้ ไม่มีการเปลี่ยนแปลงแผนงานในช่วงของการทดลอง กล่าวได้ว่าวิธีการทดลองเป็นขบวนการพิสูจน์ข้อเท็จจริงภายใต้เงื่อนไขที่กำหนด

ในทางปฏิบัติแล้ว การประเมินผลด้วยวิธีทดลองมีอุปสรรคมาก เริ่มตั้งแต่การคัดเลือกกลุ่มด้วยวิธีการสุ่มตัวอย่างจะเป็นสิ่งที่กระทำได้ยาก ผู้ร่วมโครงการมักจะมีสิทธิ์ต่อการที่จะเลือกว่าตนเองจะเข้าร่วมโครงการหรือไม่เข้าร่วม ผู้ให้บริการเองมักต้องการที่จะเป็นผู้เลือก และต้องการเลือกผู้ที่พร้อมที่สุดให้เข้าร่วมโครงการ บางครั้งเป็นหน้าที่ที่ผู้ให้บริการต้องให้บริการกับทุก ๆ คน ในชุมชน จะให้มีกลุ่มควบคุมไม่ได้ นอกจากนี้การกำหนดเงื่อนไขในการดำเนินงานของโครงการให้คงที่ก็เป็นสิ่งที่เป็นไปได้ยาก การประเมินผลด้วยวิธีการทดลองจึงมักไม่ค่อยถูกนำมาใช้ อย่างไรก็ตามนักวิจัยประเมินผลต่างก็พยายามประยุกต์ระเบียบวิธีดังกล่าวให้เหมาะสมกับงานของตน เช่น การแบ่งกลุ่มทดลองและกลุ่มควบคุมออกเป็นหลาย ๆ กลุ่มเพื่อที่ผู้ประเมินจะสามารถควบคุมผลอันเนื่องมาจากตัวแปรหลาย ๆ ตัวได้ง่ายยิ่งขึ้น นอกจากนี้ยังอาจนำโครงการที่เป็นทางเลือกอื่น ๆ มาใช้กับกลุ่มควบคุม หากมีหลาย ๆ ทางเลือกก็อาจแบ่งกลุ่มควบคุมเป็นหลาย ๆ กลุ่มดังกล่าวข้างต้น ไม่ว่าจะมีการประยุกต์อย่างไร เงื่อนไขสำคัญของวิธีการทดลองก็คือ กลุ่มทั้งสองจะต้องได้มาด้วยวิธีการ "สุ่มตัวอย่าง" และเงื่อนไขของทั้งโครงการทดลองและโครงการที่ใช้เปรียบเทียบต้องถูกกำหนดให้คงที่

### **เงื่อนไขของการทดลอง**

- มีกลุ่มทดลอง (Experimental Group)
- มีกลุ่มควบคุม (Control Group)
- กลุ่มทั้งสองได้มาด้วยการสุ่มตัวอย่าง (Random Sampling)

- มีการวัดผลในกลุ่มทั้งสอง อาจมีการวัดผลก่อน (Pre-test) เพื่อยืนยันความถูกต้องของการสุ่มตัวอย่าง และมีการวัดผลทีหลัง (Post-test) เพื่อเปรียบเทียบความแตกต่างระหว่างกลุ่มทั้งสอง

### สาเหตุที่ทำให้การวัดผลคลาดเคลื่อน

แคมป์เบล และ แสตนลีย์ (Campbell and Stanley) ได้นำเสนอ สาเหตุแปดประการที่ทำให้การวัดผลเกิดความคลาดเคลื่อน เป็นเหตุให้ผู้ประเมินหลงผิดในผลการวัด

1. **ความพร้อมของประชากรที่ถูกวัด (Maturation)** ปัจเจกบุคคลมีการเปลี่ยนแปลงโดยธรรมชาติอยู่เสมอ เช่น จากวัยเด็กไปสู่ผู้ใหญ่ จากคนโสด ไปสู่การมีครอบครัว จากอายุน้อย ไปสู่การมีอายุมากขึ้น การเปลี่ยนแปลงทั้งหลายเหล่านี้เป็นผลให้ความพร้อมของปัจเจกบุคคลเปลี่ยนแปลงไป การเปลี่ยนแปลงไปในทางที่ดีขึ้นของปัจเจกบุคคลจึงอาจเป็นผลเนื่องมาจากความพร้อมสถาบันทางสังคมก็อาจมีการเปลี่ยนแปลงตามกาลเวลาไปสู่ช่วงเวลาที่มีความพร้อมมากขึ้น เช่น สถาบันครอบครัว องค์กรปกครองท้องถิ่น รัฐวิสาหกิจ ต่าง ๆ เป็นต้น

2. **ปรากฏการณ์ที่เป็นวัฏจักร (History)** ข้าวจะมีราคาถูกหลังฤดูเก็บเกี่ยวเพราะเป็นช่วงที่ผลผลิตมีมาก และจะมีราคาสูงขึ้นเรื่อย ๆ เมื่อปริมาณผลผลิตเหลือน้อย หรืออัตราการเกิดอุบัติเหตุในทางหลวงแผ่นดินจะมีมากในหน้าเทศกาลต่าง ๆ และมีอัตราการลดน้อยลงเมื่อไม่อยู่ในหน้าเทศกาล การวัดผลโดยใช้ราคาข้าวเป็นตัวชี้วัด หรือใช้อัตราการเกิดอุบัติเหตุในท้องถนนเป็นตัวชี้วัด ล้วนมีแนวโน้มที่จะมีความคลาดเคลื่อนอันเนื่องมาจากอิทธิพลของฤดูกาลต่าง ๆ

3. **ความคุ้นเคยกับข้อสอบ (Testing)** การทำข้อสอบซ้ำ การรู้แนวข้อสอบ หรือการเคยเห็นข้อสอบมาก่อน ล้วนทำให้ความถูกต้องแม่นยำในการวัดผลเสียไป ผลที่วัดได้จากการวัดความรู้ของปัจเจกบุคคลอาจมาจากการจำข้อสอบได้ หรือเคยทำข้อสอบมาแล้ว ทำให้ผลที่วัดได้ไม่ได้เป็นการวัดความรู้ที่แท้จริงของปัจเจกบุคคลนั้น ๆ

4. **การเปลี่ยนแปลงในเครื่องมือวัดผล (Instrumentation)** การเปลี่ยนแปลงวิธีการตลอดจนเครื่องมือที่ใช้ในการวัดผล เช่น เปลี่ยนตัวผู้ตรวจข้อสอบ เปลี่ยนตัวผู้ออกข้อสอบ เปลี่ยนวิธีการจัดเก็บข้อมูล การเปลี่ยนแปลงทั้งหลายดังกล่าวทำให้การตีความผลการวัดไม่ตรงกัน เกิดความคลาดเคลื่อนโดยง่าย

5. การถดถอยในทางสถิติ (Statistical Regression) โดยหลักทางสถิติ ผู้ที่มีค่าคะแนนสูง หรือต่ำกว่าค่าเฉลี่ยของกลุ่มของตนเอง ถ้าประชากรกลุ่มนี้มีคุณสมบัติคล้ายกัน การวัดผลในครั้ง ต่อ ๆ มาจะมีผลให้ผู้ที่มีค่าคะแนนสูง หรือต่ำกว่าค่าเฉลี่ยของกลุ่มของตนเอง มีค่าคะแนนที่ได้จากการวัดใกล้เคียงกับค่าเฉลี่ยของกลุ่มตัวเองมากยิ่งขึ้น กล่าวคือ ผู้ได้คะแนนสูงจะมีแนวโน้มได้คะแนนน้อยลง ในทางกลับกัน ผู้ที่ได้คะแนนต่ำจะมีแนวโน้มได้คะแนนสูงขึ้น

6. ปัญหาอันเนื่องมาจากการเลือก (Selection) ผู้เข้าร่วมที่เลือกที่จะเข้าร่วมด้วยตนเองมักมีแนวโน้มที่จะประสบความสำเร็จสูงกว่าผู้ที่ไม่ได้ตั้งใจที่จะเข้าร่วมโครงการ ขณะเดียวกัน ผู้เข้าร่วมที่ผู้ให้บริการเป็นผู้เลือกมักมีแนวโน้มที่จะประสบความสำเร็จสูงกว่าผู้ที่ไม่ผ่านการคัดเลือกจากผู้ให้บริการ

7. ความล้มเหลวของการทดลอง (Experimental Mortality) เป็นผลเนื่องมาจากอัตราการลาออกของกลุ่มทดลองและกลุ่มควบคุมในอัตราที่ต่างกัน มีการเปลี่ยนแปลงในกลุ่มทั้งสอง มีการเคลื่อนย้ายของประชากร จนทำให้กลุ่มตัวอย่างที่เหลืออยู่ไม่ได้เป็นตัวแทนของกลุ่มดั้งเดิมที่สุ่มเข้ามา

8. ความพร้อมของผู้ที่ถูกคัดเลือกในกลุ่มทั้งสองมีไม่เท่ากัน (Selection-maturation Interaction) เป็นเหตุให้กลุ่มมีความแตกต่างตั้งแต่เริ่มแรก หรือ เริ่มเกิดความแตกต่างขึ้นเมื่อกลุ่มหนึ่งเกิดความพร้อมมากขึ้น ขณะที่อีกกลุ่มยังคงอยู่ในสถานะเดิม

### ปัญหาของการทดลอง

1. ปัญหาในการกำหนดกลุ่มควบคุม ไม่มีใครต้องการที่จะเข้าไปอยู่ในกลุ่มควบคุม ในหลายกรณีที่รัฐบาลต้องการจัดบริการให้กับทุก ๆ พื้นที่ โดยไม่ต้องการให้พื้นที่ใดมีความแตกต่างออกไป

2. ปัญหาจากขบวนการสุ่มตัวอย่าง ทั้งผู้ให้บริการ และผู้รับบริการต้องการเลือกด้วยตัวเองมากกว่าที่จะให้มีการกำหนดผู้ได้รับบริการด้วยวิธีการสุ่ม

3. ผลกระทบจากการทดลองที่ฮอเทอร์น (Hawthorne Effects) เป็นผลในทางจิตวิทยาที่ทำให้ทั้งกลุ่มทดลองและกลุ่มควบคุมมีความกระตือรือร้นที่จะให้ความร่วมมือมากกว่าปกติ ทำให้ผลที่วัดได้มักสูงเกินความเป็นจริง

4. กลุ่มควบคุมไม่มีความสัมพันธ์กับคนอื่น หรือกับบุคคลในกลุ่มทดลอง ทำให้เกิดการเรียนรู้โดยที่นักประเมินควบคุมไม่ได้ ผลจากการวัดในกลุ่มควบคุมจึงสูงกว่าที่ควรจะเป็น

5. การเปลี่ยนแปลงทางสังคม การที่มีโครงการทางสังคมจำนวนมากเข้ามาให้บริการต่อประชากรกลุ่มเป้าหมายของโครงการที่กำลังประเมิน โครงการอื่น ๆ เหล่านั้นอาจประสบความสำเร็จ และผลักดันให้คุณภาพชีวิตของประชากรดีขึ้น ทั้ง ๆ ที่โครงการที่กำลังประเมินไม่ได้มีผลใด ๆ เลย

6. การลาออกจากกลุ่ม ทำให้การทดลองล้มเหลว กลุ่มขาดการเป็นตัวแทนดั้งเดิม ไม่สามารถสรุปผลที่วัดได้ว่า เป็นการวัดจากประชากรกลุ่มเป้าหมายที่ได้รับผลจากโครงการจริงหรือไม่ อัตราส่วนของกลุ่มทั้งสองยังคงเดิมอยู่หรือไม่

**กูบ้า และ สตัฟเฟอ์บีม (Guba and Stufflebeam)** ให้ความเห็นเกี่ยวกับการทดลองไว้ว่า

1. จำเป็นต้องกำหนดให้โครงการมีลักษณะที่คงที่มากกว่าจะปล่อยให้มีการพัฒนาตัวเอง
2. เป็นประโยชน์ต่อการตัดสินใจเฉพาะเมื่อโครงการดำเนินไปจนครบวงจรแล้วเท่านั้น
3. ต้องควบคุมเงื่อนไขหลายประการซึ่งเป็นไปไม่ได้ในโลกแห่งความเป็นจริง

### ทางแก้ไข

1. ใช้วิธีผลัดกันให้ได้รับโครงการทดลอง แก้ปัญหา Hawthorne Effects เช่น การประเมินผลโครงการใช้คอมพิวเตอร์ช่วยสอนในวิชาภาษาอังกฤษ ในเทอมแรกอาจจัดให้กลุ่มหนึ่ง โดยจัดอีกกลุ่มเป็นกลุ่มควบคุมให้ได้รับวิธีการสอนแบบเดิม โดยในเทอมถัดมาต้องสัญญาว่า กลุ่มควบคุมจะต้องได้รับโครงการใช้คอมพิวเตอร์ช่วยสอนในวิชาภาษาอังกฤษเช่นที่กลุ่มแรกได้รับ

2. หาโครงการที่ดีให้กับกลุ่มควบคุมแทนที่จะไม่ได้รับอะไรเลย เช่น จัดโครงการใช้คอมพิวเตอร์ช่วยสอนในวิชาคณิตศาสตร์ ให้กับกลุ่มควบคุม ในโครงการทางวิทยาศาสตร์ เช่น การทดลองยา มีการจัดการให้กลุ่มควบคุมได้รับน้ำตาลอัดเม็ดที่จัดรูปแบบให้เหมือนกับยาที่ให้แก่กลุ่มทดลอง ในโครงการทางสังคมอาจจัดโครงการทดแทนที่ช่วยให้กลุ่มควบคุมมีความรู้สึกที่ดี ไม่เกิดความรู้สึกแปลกแยกจากกลุ่มทดลอง เรียกโครงการทดแทนที่จัดขึ้นมานี้ว่า Placebo Program

3. ถ้าต้องการตรวจสอบการเปลี่ยนแปลง อาจวัดผลในหลาย ๆ ช่วงระยะเวลา ไม่จำเป็นต้องรอจนครบวงจรการทำงาน ที่สำคัญคือตัวชี้วัดที่นำมาใช้จะต้องบ่งชี้ความสำเร็จออกมาได้

4. จะเป็นประโยชน์มากกรณีต้องการศึกษาถึงการบรรลุเป้าหมายของโครงการ (Expected Outcome)

5. ควรใช้กรณีที่จุดประสงค์ไม่เกี่ยวกับการปฏิบัติงานในขณะนั้น ๆ เช่น ต้องการตัดสินใจว่าจะให้โครงการดำเนินต่อไปหรือจะยกเลิกโครงการนั้น จะนำโครงการทดลองดังกล่าวไปใช้ทั่วประเทศหรือไม่

### การประเมินผลด้วยวิธีกึ่งการทดลอง

การประเมินผลด้วยวิธีกึ่งการทดลอง เป็นพัฒนาการของวิธีการทดลองที่ลดข้อกำหนดต่าง ๆ ให้น้อยลง กล่าวคือไม่ต้องมีการคัดเลือกกลุ่มด้วยวิธีการสุ่มตัวอย่าง เพียงมีการกำหนดคุณสมบัติของกลุ่ม และตัวโครงการให้คงที่ และกำหนดวิธีที่น่าเชื่อถือขึ้นมาวัดการเปลี่ยนแปลงที่เกิดขึ้น วิธีการแบบกึ่งการทดลองที่จัดว่าคลาสสิกได้แก่ข้อเสนอของ แคมป์เบล และสแตนลี่ซึ่งอาจแยกพิจารณาได้ดังนี้

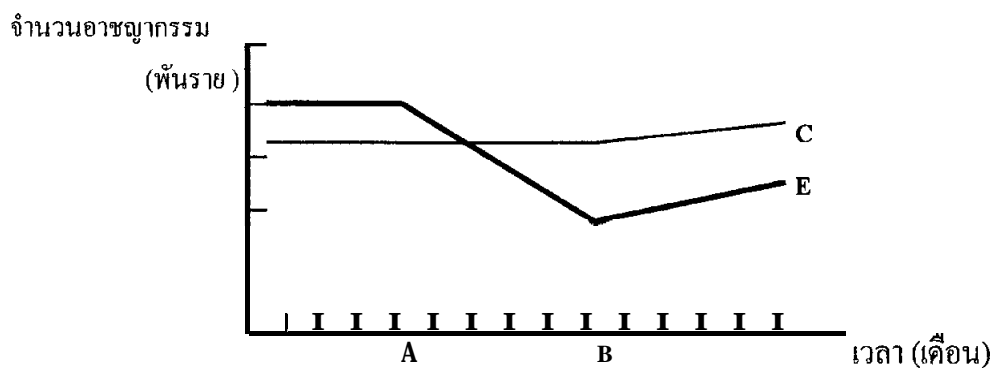
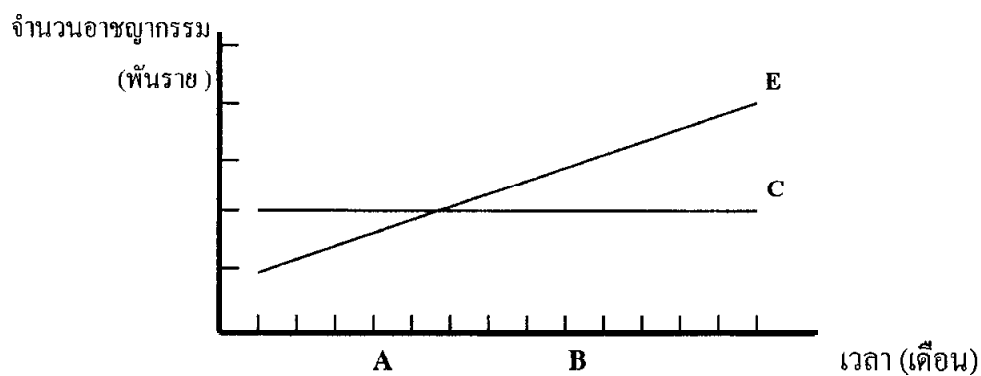
**แบบไขว้ข้อมูลระยะเวลา (Time-series Design)** กำหนดให้มีการวัดผลในกลุ่มที่เข้าไปศึกษาหลาย ๆ ช่วงระยะเวลา ตั้งแต่ก่อนที่โครงการจะเริ่ม ช่วงที่โครงการกำลังดำเนินงาน และช่วงหลังจากที่โครงการสิ้นสุด การเปลี่ยนแปลงที่ปรากฏในช่วงที่โครงการได้เข้าไปมีบทบาทจะสะท้อนให้เห็นถึงความสำเร็จของโครงการ

การวิเคราะห์ผลการวัดมักแสดงในรูปกราฟ ที่แสดงความสัมพันธ์ระหว่าง ระยะเวลา กับตัวชี้วัด เช่น การวัดสถิติอุบัติเหตุในถนนสายหนึ่ง ๆ การวัดสถิติอาชญากรรมในพื้นที่รับผิดชอบของสถานีตำรวจหนึ่ง ๆ ตัวชี้วัดมักต้องอาศัยข้อมูลจากหน่วยงาน หรือรายงานของรัฐบาล ที่มีระบบการจัดเก็บและรายงานที่เป็นมาตรฐาน เช่น สถิติอุบัติเหตุของกรมทางหลวง สถิติอาชญากรรม ของสถานีตำรวจแห่งหนึ่ง การวัดผลแบบไขว้ข้อมูลระยะเวลาไม่สามารถป้องกันความคลาดเคลื่อนอันเนื่องมาจาก ปรากฏการณ์ที่เป็นวัฏจักร (History) ได้

**แบบไขว้ข้อมูลระยะเวลาหลายกลุ่ม (Multiple Time-series Design)** มีข้อกำหนดเช่นเดียวกันกับแบบไขว้ข้อมูลระยะเวลา โดยเพิ่มกลุ่มเปรียบเทียบเข้าไปอีก กลุ่มเปรียบเทียบจะคัดเลือกมาจากบรรดาผู้ที่ไม่ได้รับโครงการแต่มีคุณสมบัติพื้นฐานคล้ายคลึงกับบรรดาผู้ที่ได้รับโครงการ กลุ่มเปรียบเทียบจะถูกวัดผลแบบไขว้ข้อมูลระยะเวลาเช่นเดียวกันกับกลุ่มทดลอง การวัดผลแบบ

อนุกรมระยะเวลาหลายกลุ่ม สามารถป้องกันความคลาดเคลื่อนอันเนื่องมาจาก ปรากฏการณ์ที่เป็น วัฏจักร (History) ได้

พิจารณาตัวอย่างการวัดผลจากกราฟ กรณีใช้อนุกรมระยะเวลาหลายกลุ่ม จะมีเส้นกราฟ สองเส้น เส้นหนึ่งเป็นของกลุ่มทดลอง และอีกเส้นหนึ่งเป็นของกลุ่มเปรียบเทียบ



จากกราฟ

- ช่วงเวลา A ถึง B เป็นช่วงที่มีโครงการพิเศษเพื่อลดจำนวนอาชญากรรม
- เส้น C เป็นกราฟของอาชญากรรมในพื้นที่เปรียบเทียบ
- เส้น E เป็นกราฟของอาชญากรรมในพื้นที่ที่ได้รับโครงการ

ในภาพแรก เส้น C และ E ไม่สามารถอธิบายผลของโครงการได้ และยังชี้ให้เห็นว่าพื้นที่ ทั้งสองมีความแตกต่างกันอย่างมาก

ในภาพที่สอง เส้น C และ E ก่อนมีโครงการมีแนวโน้มไปในทางเดียวกัน เมื่อมีโครงการ แนวโน้มต่างกันอย่างชัดเจน พื้นที่ที่ได้รับโครงการมีแนวโน้มของอาชญากรรมดีขึ้น และมีแนวโน้มกลับไปคล้ายกันอีกหลังสิ้นสุดโครงการ แสดงว่าโครงการมีผลในระยะสั้นขณะที่ดำเนินการอยู่ และหากไม่มีโครงการแนวโน้มจะกลับไปสู่สภาพเดิม

**แบบใช้กลุ่มเปรียบเทียบ (Nonequivalent Control Group Design)** มีกลุ่มทดลองและกลุ่มเปรียบเทียบเช่นเดียวกันกับแบบอนุกรมระยะเวลาเชิงซ้อน แต่การวัดผลจะไม่ใช้ช่วงระยะเวลาหลายช่วง โดยทั่วไปจะใช้การวัดผลก่อน (Pre-Test) และการวัดผลทีหลัง (Post-Test) หลักสำคัญจะอยู่ที่ความสามารถในการหากกลุ่มเปรียบเทียบที่มีลักษณะพื้นฐานเหมือนกันกับกลุ่มทดลอง มีเทคนิคในการหากกลุ่มเปรียบเทียบโดยใช้การจับคู่เหมือน (Matching)

การจับคู่เหมือนต้องเริ่มด้วยการกำหนดกลุ่มตัวอย่างที่จะศึกษาจากประชากรเป้าหมายที่ได้รับผลของโครงการ เช่น กำหนดจำนวนครัวเรือนที่จะศึกษาจากประชากรกลุ่มเป้าหมายของโครงการ หรือจากผู้เข้าร่วมโครงการ ขึ้นอยู่กับว่ากลุ่มเป้าหมายที่จะได้รับผลของโครงการเป็นใคร ถ้าอยู่ในพื้นที่กว้าง อาจเริ่มที่การกำหนดจำนวน หมู่บ้าน ตำบล หรือ อำเภอ ก่อนลงไปทีระดับครัวเรือน

เมื่อได้กลุ่มตัวอย่างที่จะศึกษาทดลอง ผู้ประเมินต้องศึกษาคุณลักษณะของกลุ่มตัวอย่าง จากนั้นจึงนำคุณสมบัติเหล่านั้นไปกำหนดจำนวนและคุณสมบัติของกลุ่มเปรียบเทียบ เช่น จับคู่ อำเภอต่ออำเภอ ตำบลต่อตำบล หมู่บ้านต่อหมู่บ้าน ที่มีคุณสมบัติคล้ายคลึงกัน แต่กลุ่มที่จะนำมาเปรียบเทียบต้องอยู่ในพื้นที่ที่ไม่ได้รับโครงการที่จะศึกษา จับคู่คุณสมบัติของครัวเรือนที่คล้ายกัน จับคู่คุณสมบัติของปัจเจกบุคคลที่คล้ายคลึงกัน

มีข้อควรระวังจากการจับคู่เหมือนโดยใช้แบบทดสอบในการช่วยวัดผล หากประชากรในทั้งสองพื้นที่มีคุณสมบัติที่แตกต่างกัน ค่าคะแนน ต่าง ๆ ที่วัดได้จะมีแนวโน้มที่จะกลับไปหาค่าเฉลี่ยของกลุ่มของตน (Statistical Regression) ผู้ประเมินจึงควรเริ่มที่การเปรียบเทียบความแตกต่างในค่าเฉลี่ยในเรื่องต่าง ๆ ที่นำมาเปรียบเทียบกัน หากไม่แตกต่างกันในทุกคุณลักษณะที่ต้องการควบคุมให้คล้ายกัน ค่อยเลือกมาอยู่ในกลุ่มเปรียบเทียบ การจับคู่เหมือนที่มีการควบคุมคุณสมบัติของกลุ่มทั้งสองในหลาย ๆ ด้าน อาจเรียกว่าเป็นแบบควบคุมต่อเนื่องหลายปัจจัย

**แบบควบคุมต่อเนื่องหลายปัจจัย (Patched-up Design)** มีข้อกำหนดเช่นเดียวกันกับแบบใช้กลุ่มเปรียบเทียบ โดยเพิ่มการควบคุมตัวแปรต่าง ๆ ที่อาจมีผลต่อกลุ่มทั้งสองเช่นในเรื่องภูมิ



หลังที่ต่างกัน วิธีการดำเนินชีวิตที่ต่างกัน การจับคู่เหมือนที่เกิดขึ้นจึงต้องพิจารณาก่อนว่ามีตัวแปรใดที่มีความสำคัญที่จะต้องควบคุม และสังเกตการเปลี่ยนแปลงโดยใช้ตัวแปรควบคุมเหล่านี้เป็นตัวเปรียบเทียบความแตกต่างในคุณภาพชีวิตที่เปลี่ยนแปลงไป

แบบกึ่งการทดลองทั้งหลายดังกล่าวจะมีความน่าเชื่อถือน้อยกว่าแบบการทดลองหมายความว่าผลที่ได้จากการประเมินอาจเป็นผลเนื่องมาจากปัจจัยอื่น ๆ ที่ไม่ใช่ตัวโครงการ อย่างไรก็ตาม การมีข้อกำหนดบางประการในการทำการวัดและประเมินผล ก็ดีกว่าจะปล่อยให้มีการวัดผลโดยปราศจากข้อกำหนดดังเช่นแบบไม่ใช้การทดลอง

งานวิจัยประเมินผล โครงการสถานีตำรวจทดลองของ ศาสตราจารย์ ปุระชัย เปี่ยมสมบูรณ์ เป็นตัวอย่างที่ดีของกรณีการใช้แบบกึ่งการทดลองแบบใช้กลุ่มเปรียบเทียบ กล่าวคือ มีการนำโครงการบริหารงานของสถานีตำรวจรูปแบบใหม่ไปใช้ที่สถานีตำรวจนครบาลทองหล่อ การประเมินความสำเร็จของโครงการใช้สถิติอาชญากรรมในพื้นที่เป็นตัวชี้วัดการเปลี่ยนแปลง โดยมีการกำหนดสถานีตำรวจนครบาลบางชื่อที่มีคุณสมบัติหลาย ๆ ประการใกล้เคียงกันกับสถานีตำรวจนครบาลทองหล่อ หากพิจารณาการวัดเพียงสองช่วงเวลา หรือไม่ได้ให้ความสำคัญกับช่วงเวลา แต่เป็นการพิจารณาตัวชี้วัดในช่วงก่อน โครงการ และหลังโครงการเปรียบเทียบกัน ก็เป็นการวัดด้วยการใช้วิธีกึ่งการทดลองแบบใช้กลุ่มเปรียบเทียบ หากนำสถิติรายสัปดาห์ หรือรายเดือนมาทำกราฟเปรียบเทียบกันก็เป็นการวัดด้วยการใช้วิธีกึ่งการทดลองแบบใช้นุกรมระยะเวลาหลายกลุ่ม (Multiple Time-series Design)

### การประเมินผลด้วยวิธีไม่ใช้การทดลอง

เป็นการประเมินผลที่อาศัยการวัดผลเปรียบเทียบบรรดาผู้ได้รับโครงการ โดยมีได้มีข้อกำหนดในเรื่อง "กลุ่มที่จะเข้าไปศึกษา" ใครก็ได้ที่ได้รับโครงการจะถูกวัดผล การวัดผลอาจทำได้สามลักษณะคือ

**การวัดผลก่อนและหลังจากโครงการสิ้นสุด** เป็นการวัดผลสองครั้งคือการวัดผลก่อนโครงการจะเริ่ม (Pre-Test) เพื่อพิจารณาคูณสมบัติที่ต้องการวัดของบรรดาผู้เข้าร่วมและการวัดผลครั้งที่สองเป็นการวัดหลังจากที่โครงการสิ้นสุด (Post-Test) เพื่อพิจารณาว่าผู้เข้าร่วมมีการเปลี่ยนแปลงไปจากก่อนที่จะได้รับโครงการเพียงใด

**การวัดผลหลังจากโครงการสิ้นสุด** เป็นการตรวจสอบผลลัพธ์ที่เกิดขึ้นเพียงครั้งเดียวหลังจากที่โครงการเสร็จสิ้นไปแล้ว จะใช้ในกรณีที่ผู้ประเมินได้รับมอบหมายหน้าที่หลังจากที่โครงการสิ้นสุด ไม่มีโอกาสที่จะใช้วิธีการแบบอื่น

**การวัดผลหลังจากที่โครงการสิ้นสุดโดยมีกลุ่มเปรียบเทียบ** เป็นการเพิ่มกลุ่มเปรียบเทียบเพื่อสร้างความน่าเชื่อถือให้สูงขึ้น กลุ่มเปรียบเทียบจะต้องมีคุณสมบัติคล้ายกับบรรดาผู้ได้รับโครงการแต่ต้องไม่ได้รับโครงการ

วิธีไม่ใช้การทดลองทั้งหลาย เหมาะที่จะใช้เป็นเครื่องมือในการตรวจสอบเบื้องต้นมากกว่าที่จะใช้เป็นเครื่องยืนยันความสัมฤทธิ์ผลของโครงการ หากผู้ประเมินพบว่าความสัมฤทธิ์ผลปรากฏจริงก็อาจเสนอให้มีการประเมินผลด้วยวิธีที่เข้มงวดกว่าต่อไป นอกจากนี้ วิธีไม่ใช้การทดลองยังเป็นที่นิยมของบรรดาผู้บริหารโครงการที่ต้องการตรวจสอบโครงการของตน เพราะเป็นวิธีที่สิ้นเปลืองค่าใช้จ่ายต่ำและดำเนินการได้ง่าย

ในทางปฏิบัติแล้ว นักประเมินผลมักจะนำวิธีการทั้งหลายข้างต้นมาประยุกต์ให้เหมาะสมกับกรณีศึกษาของตน ในการประเมินผลโครงการหนึ่ง ๆ อาจประกอบไปด้วยวิธีการประเมินหลายวิธีผสมกัน หรือวิธีหลาย ๆ วิธีอาจถูกนำมาใช้เป็นส่วนหนึ่งของตัวแบบ (Model) ที่นักประเมินผลสร้างขึ้น ทั้งนี้เพื่อให้เหมาะสมกับจุดประสงค์ในการจัดทำผลการประเมินผลของตน

### **การประเมินผลด้วยวิธีเปรียบเทียบ** (Comparative Evaluation of Programs)

1. โครงการมีวัตถุประสงค์เดียวกันแต่มีวิธีการในการดำเนินงานแตกต่างกัน จะได้ว่าวิธีการใดที่มีอิทธิพลต่อความสัมฤทธิ์ผล วิธีการใดที่ไม่มีผล

2. โครงการเดียวกันแต่ถูกแบ่งออกเป็นโครงการย่อย ๆ และถูกนำไปใช้ในท้องที่ที่มีเงื่อนไขแตกต่างกัน จะได้ว่าเงื่อนไขใดที่เป็นเหตุให้โครงการประสบความสำเร็จ เงื่อนไขใดที่นำไปสู่ความล้มเหลว

การศึกษาเปรียบเทียบดังกล่าว มีลักษณะเป็นการศึกษาโครงการตามขวาง (Cross Programs Study) และจะทำให้ได้มาซึ่งข้อมูล 2 ลักษณะคือ

(1) ข้อมูลที่สามารถนำมากล่าวอ้างเป็นการทั่วไป (Generalizability) ได้แก่ค่าเฉลี่ยของผลสำเร็จในด้านต่าง ๆ โดยรวม

(2) ข้อมูลข่าวสารที่ใช้เฉพาะเรื่อง (Specification) ได้แก่ ค่าเฉลี่ยของผลสำเร็จเฉพาะที่เกิดในพื้นที่ต่าง ๆ ซึ่งอาจเกิดจากคุณลักษณะเฉพาะของพื้นที่นั้น ๆ ที่มีอิทธิพลต่อผลลัพธ์ของโครงการ เช่น ความสะดวกในการเดินทาง ความพร้อมของสาธารณูปโภค ลักษณะของชุมชน เป็นต้น

พิจารณาตัวอย่างกรณีการวิจัยเปรียบเทียบ โครงการฝึกวิชาชีพ เป็นการเปรียบเทียบรูปแบบของโครงการสามแบบที่นำมาใช้ในหลาย ๆ ท้องที่

### รูปแบบของโครงการ

ความเชื่อเบื้องต้นในการจัดรูปโครงการมีความแตกต่างกันทางแนวคิด แต่ต้องการให้เยาวชนตระหนักถึงความก้าวร้าว เป็นอันธพาล และไม่ทำตนให้เป็นประโยชน์ต่อสังคม

รูปแบบการดำเนินงานจึงแบ่งเป็น 3 แนวทางคือ

1. ขาดความรู้ตามเกณฑ์ - สร้างความรู้พื้นฐาน, ฝึกฝนทักษะ และให้คำปรึกษา
2. ขาดแบบฉบับที่ดี - สร้างแบบฉบับที่ดี ให้ดูภาพยนตร์ อ่านวรรณกรรม
3. ไม่สอดคล้องกับความต้องการของชุมชน - พิจารณาความต้องการของชุมชน อาสาสมัครช่วยเหลือชุมชน

ความสำเร็จของโครงการทั้งสามรูปแบบจะใช้ตัวชี้วัดเดียวกัน ในที่นี้จะแบ่งตัวชี้วัดเป็นสามด้านคือ ด้าน ความก้าวร้าว การเป็นอันธพาล และการไม่ทำตนให้เป็นประโยชน์ต่อสังคม จากนั้นนักประเมินต้องไปสร้างเครื่องมือเพื่อวัดผลในทั้งสามด้าน และนำผลที่วัดได้มาใส่ในตาราง โดยอาจพิจารณาทีละตัวชี้วัด

ลักษณะของตารางอาจเป็นดังภาพต่อไปนี้

**แสดงข้อมูลจากการศึกษาโครงการตามขวาง**

ลักษณะของชุมชน	รูปแบบของโครงการ		
	ความรู้ตามเกณฑ์	แบบฉบับที่ดี	สอดคล้องกับชุมชน
A			
B			
C			
รวมทุกชุมชน			

**การประเมินผลด้วยวิธีเปรียบเทียบควรนำมาใช้เมื่อไร**

1. เป็นเรื่องที่เกิดขึ้นจริง และผู้ตัดสินใจกำลังเผชิญปัญหาการตัดสินใจว่าจะเลือกใช้วิธีการดำเนินงานแบบใด
2. โครงการเปรียบเทียบทั้งหลายต้องมีการกำหนดรูปแบบไว้อย่างดี โดยมีความมุ่งหมายหลักเดียวกัน แต่มีวิธีที่จะให้บรรลุความมุ่งหมายหลักต่างวิธีกัน
3. เมื่อมีหลักฐานปรากฏว่า โครงการทั้งหลายนั้นมีความโน้มเอียงที่จะให้ความสำเร็จบางอย่างออกมา

**การวิเคราะห์ต้นทุนและประโยชน์ (Cost-benefit Analysis)**

1. ต้องมีผลในเรื่องเกี่ยวกับชนิด และจำนวนของประโยชน์ที่เกิดขึ้นจากโครงการ
2. ต้องมีข้อมูลเกี่ยวกับงบประมาณที่ใช้ทั้งหมด = ต้นทุน
3. เป็นการเปรียบเทียบทางเลือกต่าง ๆ ว่าจะให้ผลตอบแทนในรูปตัวเงินต่างกันอย่างไร  

$$\text{ผลตอบแทน} = \frac{\text{จำนวนประโยชน์ที่เกิดขึ้น}}{\text{ต้นทุนที่ใช้}}$$

**ข้อควรคำนึง**

1. ประโยชน์ที่สัมผัสได้และประโยชน์ที่ไม่อาจสัมผัสได้

### ประโยชน์ที่ไม่อาจล้มผลได้

- ความรู้สึกที่ดีในเรื่องความเชื่อมั่นในความสามารถของตัวเอง
  - ความวิตกกังวลที่จะต้องเสียค่าพยาบาลในอนาคต
  - ความพยายามที่จะไม่ให้เกิดความแตกแยกขึ้นในครอบครัว
2. ต้นทุนที่ใช้ทั้งทางตรงและทางอ้อม สร้างเขื่อน มลภาวะ
  3. การกระจายของต้นทุนและประโยชน์ ผู้ใดที่ได้ประโยชน์ ผู้ใดบ้างที่เสียประโยชน์
  4. ปัจจัยที่จะนำมาคิดคำนวณในเรื่องมูลค่าของประโยชน์ ประโยชน์จากโครงการรักษาผู้  
ติดยา
  5. การแปลงค่าประโยชน์ให้อยู่ในรูปมูลค่าที่เป็นตัวเงิน ตัวเลขใดที่เป็นที่ยอมรับร่วมกัน
    - Fixed level of benefits
  6. ปัญหาในเชิงทฤษฎีที่ใช้ในการแปลงค่า คะแนนที่เพิ่มขึ้นของเด็กกับรายได้ที่จะได้รับ  
ในอนาคต
  7. อัตราส่วนลด (discount rate)
  8. การเพิ่มขึ้นของมูลค่าส่วนสุดท้าย

### เมื่อใดควรนำมาใช้

- ข้อมูลที่มีอยู่ได้บ่งชี้ให้เห็นถึงขอบเขตและปริมาณที่เป็นผลเนื่องมาจากโครงการ
- เมื่อประโยชน์หลักที่จะได้รับสามารถแปลงให้อยู่ในรูปของตัวเงิน
- เมื่อเรื่องที่จะพิจารณาเป็นเรื่องของการค้นหาประโยชน์โดยทั่ว ๆ ไป มากกว่าการ  
ค้นหาการกระจายของประโยชน์

.....

### คำถามท้ายบท

1. จงอธิบาย การวัดผลด้วยวิธีการทดลอง (Experimental Design) พร้อมยกตัวอย่างให้เป็นที่เข้าใจ
2. จงอธิบาย การวัดผลด้วยวิธีการทดลอง (Quasi-experimental Designs) พร้อมยกตัวอย่างให้

เป็นที่เข้าใจ

3. จงอธิบาย การวัดผลด้วยวิธีไม่ใช้การทดลอง (Nonexperimental Designs) พร้อมยกตัวอย่างให้เป็นที่เข้าใจ

4. จงอธิบายคำเหล่านี้

1) Comparative Evaluation of Programs

2) Cost Benefit Analysis

3) Random Sampling

4) Hawthorne Effects

5) Placebo Program

6) Statistical Regression

7) Nonequivalent Control Group Design

8) Multiple Time-series Design

9) Benefit / Cost Ratio

10) Cross Programs Study

5. จงอธิบายถึงสาเหตุแปดประการ ที่ทำให้การวัดผลคลาดเคลื่อน ตามที่แคมป์เบล และ สแตนลีย์นำเสนอไว้

6. การวิเคราะห์ต้นทุนและประโยชน์คืออะไร มีข้อควรคำนึงอย่างไรบ้าง

7. การประเมินผลด้วยวิธีการเปรียบเทียบโครงการ (Comparative Evaluation of Programs) เป็นอย่างไร ใช้เมื่อใด จงอธิบายพร้อมยกตัวอย่าง

-----