

การประเมินผลนโยบาย

รองศาสตราจารย์ ดร. กิตติ มุขนาค

1. บทนำ

นโยบายสาธารณะเป็นแนวทางการดำเนินงานของรัฐบาล เมื่อนำนโยบายไปปฏิบัติแล้ว ผลของนโยบายที่เกิดขึ้นอาจสอดคล้องกับวัตถุประสงค์ที่กำหนดไว้ เมื่อแรกกำหนดนโยบาย อาจไม่บรรลุวัตถุประสงค์เลย อาจก่อให้เกิดผลข้างเคียงที่น่าไปสู่ปัญหาของประเทศในด้านอื่น หรืออาจก่อให้เกิดผลในทางดีที่ไม่เคยคาดการณ์มาก่อน ผลของนโยบายทั้งผลทางตรง ทางอ้อม ผลในระยะสั้น หรือระยะยาว ล้วนมีความสำคัญต่อการกำหนดนโยบายของรัฐบาลในอนาคตต่อไป บทเรียนจากความสำเร็จหรือล้มเหลวที่เกิดขึ้นจะมีส่วนช่วยให้เกิดการปรับปรุงเปลี่ยนแปลงทั้งในด้านเนื้อหา และวิธีการดำเนินงานของนโยบายให้รอบคอบรัดกุมขึ้น อย่างไรก็ตาม ผลดังกล่าวจะทราบได้ต่อเมื่อมีการประเมินผล นโยบายอย่างจริงจังและเป็นระบบ

2. ความหมายของการประเมินผลนโยบาย

การประเมินผลนโยบาย (policy evaluation) เป็นคำที่มีชื่อเรียกต่าง ๆ มากมายหลายชื่อ อาทิ การประเมินผลระหว่างดำเนินการ (formative evaluation) การประเมินผลสรุปรวบยอด (summative evaluation) การประเมินผลย้อนหลัง (retrospective evaluation) การประเมินผลล่วงหน้า (prospective evaluation) การประเมินความสามารถประเมินผล (evaluability assessment) และการประเมินผลกระทบ (impact evaluation) การที่มีชื่อเรียกหลายชื่อสะท้อนให้เห็นถึงความหลากหลายของธรรมชาติของการประเมินผล นักวิชาการแต่ละคนต่างมีจุดสนใจของตนเองและอาจเลือกศึกษาเน้นเฉพาะประเด็นที่ตนสนใจ โดยเรียกการศึกษานั้นว่าการประเมินผลนโยบาย ซึ่งลักษณะนี้ทำให้การศึกษาเรื่องนี้มี ความกว้างขวางและหลากหลาย อย่างไรก็ตามการให้ความหมายของการประเมินผลนโยบายในลักษณะกว้างก็ยังคงอยู่ในวิสัยที่กระทำได้

อิมิล เจ โพซาวัค และ เรย์มอนด์ จี แครี (Emil J. Posavac & Raymond G. Carey) กล่าวว่า การประเมินผลนโยบายเป็นการใช้วิธีการหลายวิธีและทักษะหลาย ๆ ด้านเพื่อที่จะตกลงใจว่า นโยบายที่กำหนดไว้นั้นจำเป็นหรือไม่ ควรจะใช้หรือไม่ ดำเนิน

การไปตามที่วางแผนไว้หรือไม่ ช่วยแก้ไขปัญหาคาตามที่มุ่งหวังไว้หรือไม่ ในความเห็นของทั้งสอง การศึกษาเรื่องนี้พัฒนาจากแนวคิดของจิตวิทยา สังคมวิทยา การบริหาร นโยบายศาสตร์ เศรษฐศาสตร์ และศึกษาศาสตร์ ธรรมชาติของการศึกษาเรื่องนี้เปลี่ยนแปลงไปมาจนกระทั่งไม่มีลักษณะที่เหมือนศาสตร์ใดศาสตร์หนึ่งโดยเฉพาะ¹

เจมส์ อี แอนเดอร์สัน (James E. Anderson) ให้ความหมายของการประเมินผลนโยบายว่า เป็นกิจกรรมที่เกี่ยวข้องกับการประมาณการณ์ การเปรียบเทียบผลของการนำนโยบายไปปฏิบัติกับสิ่งที่คาดว่าจะเกิดขึ้น กิจกรรมนี้เป็นกิจกรรมที่กระทำอย่างต่อเนื่องตลอดเวลาในทุกชั้นตอนนโยบาย ในการประเมินผลนโยบายนั้นอย่างน้อยที่สุดผู้ประเมินต้องทราบในประเด็นต่อไปนี้ คือ เป้าหมายหรือวัตถุประสงค์ของนโยบายคืออะไร การดำเนินการให้บรรลุตามเป้าหมายกำหนดไว้อย่างไร และผลที่เกิดขึ้นจากการดำเนินการทั้งหมดเป็นอย่างไร²

ชาลส์ โอ โจนส์ (Charles O. Jones) ชี้ว่า การประเมินผลนโยบายเป็นการกระทำที่มีระบบและต่อเนื่องเพื่อให้ทราบถึงผลของนโยบายโดยเปรียบเทียบกับเป้าหมายที่กำหนดไว้กับผลกระทบของการดำเนินการตามนโยบายที่มีต่อปัญหาของ

สังคมที่เป็นเป้าหมายที่นโยบายนั้นมุ่งแก้ไข การกระทำดังกล่าวนี้ เป็นกิจกรรมที่เป็นหน้าที่ประจำอย่างหนึ่งของรัฐบาลในทุกกระบวนการเมือง เป็นเสมือนเครื่องมือที่รัฐบาลใช้ทบทวน ตรวจสอบ และ ประเมินความก้าวหน้าในการทำงานของตนเองนั่นเอง³

อี อาร์ เฮาส์ (E. R. House) ชี้ว่า การประเมินผลโดยแท้จริงเป็นเพียงกิจกรรมชักชวน (acts of persuasion) ซึ่งมุ่งหวังให้กลุ่มเป้าหมายเข้าใจถึงผลของนโยบายหรือกิจกรรมที่ประเมินนั้น การประเมินผลไม่ใช่กิจกรรมที่พยายามทำให้คนปักใจเชื่อ ไม่ใช่กิจกรรมที่บังคับหรือกำหนดขึ้นเพื่อครอบงำความคิดเห็นของใคร แต่เป็นกิจกรรมที่มุ่งชักชวนให้กลุ่มเป้าหมายเข้าใจ และยอมรับถึงผลของนโยบายที่เกิดขึ้น ทรรศนะของเฮาส์จึงค่อนข้างจะเป็นนามธรรมและเน้นจุดหมายของการนำเอาการประเมินผลไปใช้ เป็นสำคัญ⁴

วิลเลียมส์ เอ็น ดันน์ ชี้ว่า การประเมินผลนโยบายเป็นขั้นตอนหนึ่งของการวิเคราะห์นโยบาย โดยเป็นขั้นตอนที่มุ่งผลิตข่าวสารที่เกี่ยวข้องกับผลของการดำเนินงานตามนโยบาย ว่าสามารถสนองความต้องการของสังคม สนองคุณค่าของสังคม และแก้ไขปัญหาที่เป็นเป้าหมายของนโยบายได้หรือไม่ ดังนั้นการ

ประเมินผลนโยบายจึงมุ่งตอบปัญหาที่ว่า “หลังจากดำเนินการตามนโยบายแล้วเกิดการเปลี่ยนแปลงอะไรขึ้นในสังคมหรือไม่”⁵

วอร์เธิน และแซนเดอร์ส (Worthen and Sanders) กล่าวว่า การประเมินผลคือการให้คุณค่า หรือการเลือก การประเมินผลในทรรศนะของทั้งสองจึงมีความหมายกว้างมาก เป็นกิจกรรมที่มีอยู่กับมนุษย์มาช้านานแล้ว และมนุษย์ทุกคนจะเป็นผู้ประเมินไม่โดยทางใดก็ทางหนึ่ง กิจกรรมที่เรียกว่าการประเมินผลนี้จะเกี่ยวข้องโดยตรงกับการเลือก โดยทั่วไปมนุษย์จะเลือกทางเลือกที่ดีที่สุดจากทางเลือกที่มีอยู่ โดยอาศัยการรับรู้ (Perception) ของตนเอง トラบดที่กิจกรรมการเลือกกระทำอย่างเป็นระบบมากที่สุดมีการนิยามเกณฑ์ต่าง ๆ ที่ใช้พิจารณา มีการรวบรวมข้อมูลเกี่ยวกับทางเลือกต่าง ๆ อย่างถูกต้อง การให้คุณค่าที่แท้จริงของทางเลือกย่อมกระทำได้ และทั้งหมดที่กล่าวมาคือ การประเมินผลในทรรศนะของวอร์เธิน และแซนเดอร์ส⁶

จากที่กล่าวมาทั้งหมด จะพบว่ามิใช่ผู้ให้ความหมายของการประเมินผลนโยบายนั้นไว้ค่อนข้างหลากหลาย แต่หัวใจของความหมายที่แต่ละคนนิยามนั้นพอจะกล่าวได้ว่าคล้ายกันอย่างน้อยที่สุดประเด็นหนึ่ง นั่นคือ การประเมินผลเกี่ยวข้องกับกระบวนการวัดคุณค่าของผลของการดำเนินการตามนโยบาย เพื่อที่จะนำ

มาเปรียบเทียบกับเป้าหมายหรือวัตถุประสงค์ที่กำหนดไว้ และข้อควรสังเกตุอีกประการหนึ่งคือ การประเมินผลนั้นไม่ได้แยกเป็นเอกเทศจากขั้นตอนนโยบายอื่น แต่เกี่ยวข้องกันตลอดเวลา

3. ลักษณะของการประเมินผลนโยบาย

ความหมายของการประเมินผลนโยบายที่กล่าวข้างต้นให้ภาพกว้าง ๆ ว่า การประเมินผลนโยบายคืออะไร แต่ไม่ได้ให้รายละเอียดมากนักเกี่ยวกับธรรมชาติ หรือลักษณะของการประเมินผลนโยบาย เพื่อความเข้าใจที่ชัดเจนขึ้น ขอให้พิจารณาลักษณะของการประเมินผลนโยบายที่จะเสนอต่อไปประกอบด้วย

โจเซฟ ดี กองควัวส์ (Joseph D. Comtois) ได้อธิบายลักษณะของการประเมินผลนโยบายในอุดมคติว่าควรประกอบด้วยลักษณะสำคัญต่อไปนี้⁷

(1) เป็นสหวิชา นั่นคือ ดึงเอาความรู้จากศาสตร์ต่าง ๆ มาประยุกต์ใช้เพื่อให้เข้าใจลักษณะและคุณค่าของผลการดำเนินการตามนโยบายอย่างถ่องแท้ชัดเจน

(2) เป็นที่ยอมรับทั้งผู้ประเมินเอง และผู้มีอำนาจตัดสินใจ นั่นคือผู้ประเมินและผู้รับผิดชอบนโยบายเห็นพ้องต้องกันในลักษณะและวิธีการประเมิน

(3) นำยุทธวิธีต่าง ๆ มาผสมผสานกันอย่างเหมาะสม เพื่อให้เข้าใจผลของนโยบายอย่างชัดเจนขึ้น

(4) กระทำอย่างเป็นกลางให้มากที่สุดเท่าที่จะทำได้

(5) ในการรวบรวมข่าวสารข้อมูลต่าง ๆ ต้องเน้นการติดต่อสื่อสารระหว่างบุคคลภายในหน่วยงาน มากกว่าการติดต่อกับบุคคลภายนอกหน่วยงาน

(6) วิธีการที่ใช้ต้องยืดหยุ่น และเปิดกว้างเพียงพอที่จะยอมรับระเบียบวิธีการศึกษาวิจัยทั้งเชิงปริมาณ ซึ่งเน้นการวิเคราะห์ข้อมูลโดยวิธีการทางสถิติ และเชิงคุณภาพ ซึ่งเน้นการใช้เหตุผลและประสบการณ์

4. ประเภทของการประเมินผลนโยบาย

การประเมินผลนโยบายสามารถจำแนกเป็นประเภทต่าง ๆ ได้หลายวิธีด้วยกัน นักวิชาการบางกลุ่มแบ่งประเภทการประเมินผลนโยบายออกเป็น 2 ประเภท โดยพิจารณาวิธีการประเมิน คือ การประเมินเชิงปริมาณ และการประเมินเชิงคุณภาพ

นักวิชาการกลุ่มนี้ มีอาทิ คุก และไรซ์ฮาร์ท (Cook and Reiehardt) และฟิลสเต็ด (Filstead)⁸ เป็นต้น นักวิชาการบางคน เช่น โจนส์ จำแนกประเภทการประเมินผลนโยบายโดยพิจารณาจากบุคคลผู้ทำการประเมินออกเป็น 3 ประเภทใหญ่ ๆ คือ การประเมินผลโดยหน่วยงานราชการ การประเมินผลโดยองค์การอิสระ และการประเมินผลอย่างเป็นระบบโดยนักวิชาการ⁹ นักวิชาการบางคน เช่น ดันน์ จำแนกประเภทการประเมินผลนโยบายโดยพิจารณาจากเป้าหมาย การประเมินออกเป็น 3 ประเภท คือ การประเมินผลแบบเทียบ การประเมินผลแบบเป็นทางการ และการประเมินผลแบบพิจารณาความเหมาะสม¹⁰ ในรายงานฉบับนี้ผู้เขียนจะยึดวิธีการแยกประเภทการประเมินผลนโยบายตามแนวของดันน์เป็นแนวทางอธิบายการประเมินผลประเภทต่าง ๆ ดังจะกล่าวละเอียดต่อไป¹¹

4.1 การประเมินผลแบบเทียบ (pseudo-evaluation) คือ การประเมินผลซึ่งใช้วิธีการที่เป็นวิทยาศาสตร์ในการสร้างข่าวสารที่ถูกต้องและเชื่อถือได้เกี่ยวกับผลของนโยบาย โดยไม่พยายามที่จะพิจารณาว่าผลดังกล่าวมีคุณค่าหรือประโยชน์มากนักเพียงใดต่อประชาชนกลุ่มต่าง ๆ ในสังคมและต่อประเทศชาติ โดยส่วนรวม ข้อสมมุติพื้นฐานของการประเมินผลแบบนี้คือความ

เชื่อว่า ตัววัดต่าง ๆ ที่ใช้ในการวัดคุณค่าหรือประโยชน์ของนโยบายนั้นถูกต้อง เชื่อถือได้ และเป็นที่ยอมรับทั่วไปในหมู่ประชาชนกลุ่มต่าง ๆ นั่นคือ การประเมินผลแบบนี้ยึดเอาตัววัดนโยบายต่าง ๆ อาทิ อัตราเงินเฟ้อ จำนวนคนว่างงาน สถิติอาชญากรรม จำนวนโรงพยาบาล ปริมาณผลิตผลทางการเกษตร และสถิติอุบัติเหตุในโรงงานเป็นตัววัดผลของนโยบาย โดยไม่คำนึงถึงความเหมาะสมของตัววัดเหล่านี้ว่าสามารถวัดผลของนโยบายได้ถูกต้องเที่ยงตรงเพียงใด ตรงกับวัตถุประสงค์ของนโยบายหรือไม่

การประเมินผลแบบนี้ โดยปกติผู้ประเมินมักจะใช้วิธีการต่าง ๆ หลายวิธีเพื่ออธิบายผลของนโยบายที่เกิดขึ้นว่ามาจากปัจจัยที่สำคัญอะไรบ้าง วิธีการต่าง ๆ ที่ใช้มักจะเน้นการวิจัยเชิงทดลอง การวิจัยกึ่งทดลอง และการใช้แบบสอบถามในการเก็บรวบรวมข้อมูล วิธีการทางสถิติต่าง ๆ มักจะถูกนำมาประยุกต์ใช้ในการประเมิน ไม่ว่าจะเป็นการสุ่มตัวอย่าง การวิเคราะห์ข้อมูล หรือการตีความข้อมูล อย่างไรก็ตาม การประเมินผลแบบนี้มักประเมินโดยใช้ตัววัดที่ใช้กันแพร่หลายทั่วไปวัดผลการประเมิน โดยไม่คำนึงว่าตัววัดต่าง ๆ เหล่านี้เป็นตัววัดที่เหมาะสมหรือไม่

รูปแบบที่ใช้ในการประเมินผลนโยบายตามแนวนี มี 4 รูปแบบใหญ่ ๆ ด้วยกัน คือ รูปแบบที่หนึ่ง การทำบัญชีระบบสังคม (social systems accounting) รูปแบบที่สอง การทดลองทางสังคม (social experimentation) รูปแบบที่สาม การตรวจสอบทางสังคม (social auditing) และ รูปแบบที่สี่ การวิจัยสะสมทางสังคม (social research accumulation) รูปแบบทั้งสี่นี้มีลักษณะร่วมที่สำคัญ 5 ประการ คือ ประการที่หนึ่ง ทั้งสี่รูปแบบเน้นการสร้างข่าวสารข้อมูลที่ต้องการและเชื่อถือได้เกี่ยวกับผลของนโยบาย ดังนั้นรูปแบบทั้งสี่จึงเกี่ยวข้องกับตัวแปรที่เป็นตัววัดผลของนโยบายต่าง ๆ เช่น จำนวนครั้งที่ให้บริการแก่ประชาชน จำนวนโรงพยาบาลที่รัฐสร้างขึ้น และจำนวนผลิตผลทางเกษตรที่ส่งออกในปีหนึ่ง ๆ เป็นต้น ประการที่สอง ทั้งสี่รูปแบบเน้นการพิจารณาวัตถุประสงค์หรือเป้าหมายของนโยบาย นั่นคือ ประเมินผลโดยพิจารณาผลของนโยบายเป็นหลัก เนื่องจากเชื่อว่าผลของนโยบายนั้นสะท้อนวัตถุประสงค์ หรือเป้าหมายของนโยบาย ประการที่สาม ทั้งสี่รูปแบบเน้นการประเมินการเปลี่ยนแปลงที่เกิดขึ้น ไม่ว่าจะเป็นการเปลี่ยนแปลงในผลของนโยบายหนึ่งในช่วงเวลาหนึ่ง หรือการเปรียบเทียบการเปลี่ยนแปลงในผลของนโยบายหลาย ๆ นโยบายก็ตาม ประการที่สี่ ทั้งสี่รูปแบบอนุญาตให้มีการจัดกลุ่ม

ผลของนโยบายโดยตัวแปรอื่นที่เกี่ยวข้องกับนโยบาย ตัวอย่างเช่น การประเมินผลนโยบายสาธารณสุข โดยพิจารณาภาพรวมจาก จำนวนโรงพยาบาลของรัฐที่เพิ่มขึ้นในแต่ละปีหรือพิจารณาโดย นำตัวแปรเขต หรือภูมิภาคมาเกี่ยวข้องเพื่อสรุปว่า จำนวนโรงพยาบาลของรัฐที่เพิ่มขึ้นในแต่ละปีในเขตหรือภูมิภาคต่าง ๆ มี จำนวนเท่าใด หรือพิจารณาโดยนำตัวแปรขนาดเข้ามาเกี่ยวข้อง เพื่อหาข้อมูลว่า จำนวนโรงพยาบาลแยกตามขนาดของโรงพยาบาลที่เพิ่มขึ้นในแต่ละปี จำนวนเท่าใด และลักษณะร่วม *ประการสุดท้าย* คือ ทั้งสี่รูปแบบโดยทั่วไปมักใช้ตัววัดที่เป็นทั้ง ปริมาณและอัตรายในการวัดผลของนโยบาย ตัววัดที่เป็นปริมาณ เช่น จำนวนนักศึกษาที่สถาบันอุดมศึกษาสามารถรับเข้าศึกษาได้ในแต่ละปี จำนวนคนไข้ที่สามารถให้บริการได้ และปริมาณน้ำประปาที่ผลิตได้ต่อวัน เป็นต้น ส่วนตัววัดที่เป็นอัตราย เช่น ความพอใจที่ นักศึกษามีต่อบริการด้านต่าง ๆ ของมหาวิทยาลัย ความประทับใจ ในบริการของโรงพยาบาล และความพอใจที่ประชาชนมีต่อ บริการของการประปา เป็นต้น *ลักษณะร่วมทั้งห้าประการนี้ช่วย นิยามการประเมินผลแบบนี้ว่าเป็นการสร้างข่าวสารที่ถูกต้องและ เชื่อถือได้เกี่ยวกับผลของนโยบายโดยใช้ตัววัดต่าง ๆ ที่มีใช้อยู่ ทั่วไป และใช้วิธีการที่เป็นวิทยาศาสตร์เป็นกรอบในการสร้าง*

ข้อมูล นอกเหนือจากลักษณะร่วมห้าประการนี้แล้ว การประเมินผลแต่ละรูปแบบมีลักษณะ ทัวไปดังนี้

4.1.1 การทำบัญชีระบบสังคม เน้นการสร้างดัชนีสังคม (social indicator) ขึ้นมาเพื่อแสดงความเปลี่ยนแปลงต่าง ๆ ในสังคม และเป็นตัววัดเพื่อประเมินผลของนโยบายสาธารณะ ดัชนีสังคมในที่นี้ หมายถึง ค่าที่ใช้วัดสภาพต่าง ๆ ทางสังคม เศรษฐกิจ และการเมืองเพื่อบอกให้ทราบถึงสภาพปัจจุบัน และการเปลี่ยนแปลงต่าง ๆ ตัวอย่างเช่น ดัชนีสังคมที่ใช้วัดสภาพสังคม เช่น จำนวนหนังสือพิมพ์รายวันที่จำหน่ายต่อวัน อัตราการเพิ่ม-ลดประชากร อัตราคนอ่านออกเขียนได้ อัตราการเพิ่ม-ลดอุบัติเหตุจราจร อัตราการเพิ่ม-ลดอาชญากรรมทางเพศ จำนวนอุบัติเหตุในโรงงาน และจำนวนบัณฑิตว่างงานในแต่ละปี เป็นต้น ดัชนีสังคมที่ใช้วัดสภาพเศรษฐกิจ เช่น ดัชนีราคาสินค้าจำเป็นต่าง ๆ มวลรวมประชาชาติ อัตราการว่างงาน ปริมาณผลิตผลทางการเกษตรประเภทต่าง ๆ ปริมาณสินค้านำเข้า ปริมาณสินค้าส่งออก และจำนวนหนี้สินระหว่างประเทศ เป็นต้น ดัชนีสังคมที่ใช้วัดสภาพการเมือง เช่น จำนวนผู้มาใช้สิทธิออกเสียงเลือกตั้ง จำนวนการก่อการร้ายในรอบปี ความถี่ของการใช้ความรุนแรงทางการเมืองต่าง ๆ และอายุของพรรคการเมือง เป็นต้น

ดัชนีสังคมนี้มีทั้งที่เป็นดัชนีแบบปรนัยและแบบอัตนัย ดัชนีแบบปรนัยสร้างขึ้นจากตัวเลขที่แสดงสถานการณ์ของสภาพสังคม เศรษฐกิจ การเมืองในขณะนั้น ๆ และเป็นตัวเลขที่วัดจากจำนวนหรือปริมาณที่สังเกตได้ ให้ค่าทางตัวเลขได้ เป็นสิ่งที่ป็นรูปธรรม ตัวอย่าง เช่น จำนวนคนอพยพเข้ากรุงเทพมหานคร จำนวนบัณฑิตที่จบการศึกษาในแต่ละปี ปริมาณผลิตผลทางการเกษตร ปริมาณสินค้านำเข้าและส่งออก ปริมาณภาษีขาเข้าและขาออกที่เรียกเก็บได้ และจำนวนคนอ่านออกเขียนได้ เป็นต้น ส่วนดัชนีแบบอัตนัยนั้นแสดงสภาวะที่เป็นนามธรรมสังเกตไม่ได้ จับต้องไม่ได้ ให้ค่าทางตัวเลขค่อนข้างลำบาก ตัวอย่างดัชนีแบบนี้ เช่น ทัศนนะของชาวเมืองพัทยาต่อการบริหารงานของนายกเมืองพัทยา ทัศนนะของประชาชนที่มีต่อการยุบสภา ความพอใจในการบริการของหน่วยงานรัฐบาล ความมั่นใจในระบอบการปกครองปัจจุบันความรู้สึกลดลภัยในชีวิต และทรัพย์สิน และความตระหนักในคุณค่าการศึกษา เป็นต้น

4.1.2 การทดลองทางสังคม เน้นการทดลองใช้นโยบายหนึ่ง ๆ ในกลุ่มเป้าหมายตัวอย่าง แล้วประเมินผลที่เกิดขึ้นจากการดำเนินการตามนโยบายนั้นก่อนที่จะน่านโยบายไปปฏิบัติในกลุ่มเป้าหมายจริงทั้งหมด การประเมินผลแบบนี้โดยทั่วไปมี

ขั้นตอนการประเมินคล้ายคลึงกับขั้นตอนการทดลองที่ใช้ในห้องปฏิบัติการทางวิทยาศาสตร์โดยทั่วไป กล่าวคือ ยึดหลักการที่สำคัญสามประการ *ประการแรก* คือ ผู้ประเมินต้องสามารถควบคุมการดำเนินการต่าง ๆ ตามนโยบายได้โดยตรง ตัวอย่างเช่น การประเมินผลนโยบายการฝึกอบรมบุคลากร โดยใช้โสตทัศนูปกรณ์สมัยใหม่ช่วยในการอบรม ผู้ประเมินต้องสามารถควบคุมได้โดยตรงว่าการใช้โสตทัศนูปกรณ์ในกลุ่มที่ทำการทดลองถูกต้องตามหลักการที่กำหนดไว้หรือไม่ *ประการที่สอง* คือ ผู้ประเมินต้องเปรียบเทียบระหว่างกลุ่มเป้าหมายอย่างน้อยสองกลุ่ม โดยให้กลุ่มหนึ่งเป็น*กลุ่มทดลอง* และอีกกลุ่มหนึ่งเป็น*กลุ่มควบคุม* กลุ่มแรกจะถูกดำเนินการตามที่กำหนดไว้ในนโยบาย กลุ่มหลังไม่ถูกดำเนินการใด ๆ ทั้งสิ้น และ *ประการสุดท้าย* คือ การคัดเลือกบุคคลต่าง ๆ เข้ามาเป็นตัวอย่างในกลุ่มทดลองและกลุ่มควบคุมต้องทำแบบสุ่ม (random) ทั้งนี้เพื่อลดอคติของข้อมูลที่อาจเกิดขึ้นเนื่องจากบุคคลที่เป็นสมาชิกของกลุ่มที่ศึกษามีลักษณะต่าง ๆ ใกล้เคียงกันเกินไป การสุ่มจะช่วยให้บุคคลที่เข้ารับการทดลองมีลักษณะใกล้เคียงกับบุคคลที่จะเป็นเป้าหมายของนโยบายจริงผลที่ได้จากการทดลองจึงสามารถสรุปรวมไปถึงกลุ่มเป้าหมายทั้งหมดได้

การทดลองทางสังคมมีลักษณะเด่นที่สำคัญคือ สามารถให้ข้อมูลที่ถูกต้องและเชื่อถือได้มากที่สุด ทั้งนี้ เนื่องจากหลักการที่ใช้เป็นเช่นเดียวกับที่นักวิทยาศาสตร์ใช้ทดลองในห้องปฏิบัติการนั่นเอง อย่างไรก็ตาม การดำเนินการให้ถูกต้องตามขั้นตอนของการทดลองในห้องปฏิบัติการทุกประการมักทำได้ลำบากและนี่เองเป็นสาเหตุที่ทำให้การประเมินผลแบบนี้ไม่มีการใช้บ่อยนัก

4.1.3 การตรวจสอบทางสังคม ต่างกับการประเมินผล 2 แบบแรกในแง่ที่ว่า การประเมินผลสองแบบแรกไม่สนใจพิจารณาขั้นตอนต่าง ๆ ของการกำหนดและดำเนินการตามนโยบายมากนัก ในขณะที่การประเมินผลแบบนี้เน้นการพิจารณาความสัมพันธ์ระหว่างปัจจัยนำเข้า กระบวนการในการดำเนินการตามนโยบายและผลของนโยบาย ซึ่งช่วยให้ทราบได้ว่า ผลของนโยบายที่เกิดขึ้นนั้นเกิดจากปัจจัยนำเข้า หรือเกิดจากกระบวนการดำเนินการตามนโยบาย หรือเป็นผลจากปฏิสัมพันธ์ทั้งสอง การประเมินผลแบบนี้จึงช่วยให้ได้ข้อมูลเกี่ยวกับเหตุการณ์ และกิจกรรมต่าง ๆ ที่เกิดขึ้นในกระบวนการแปรสภาพปัจจัยนำเข้าให้เป็นนโยบายและปรากฏเป็นผลออกมา

กระบวนการในการประเมินผลแบบนี้แบ่งเป็น 2 ส่วนใหญ่ ๆ ด้วยกัน ส่วนแรกพิจารณาถึงส่วนประกอบต่าง ๆ ของปัจจัยนำเข้า และส่วนที่สองพิจารณาถึงกระบวนการแปรสภาพปัจจัยนำเข้า ในส่วนแรกจะพิจารณาถึงปัจจัยต่าง ๆ ที่ประกอบเป็นปัจจัยนำเข้าดังตัวอย่าง เช่น นโยบายการพัฒนาบุคลากรสองนโยบายอาจใช้งบประมาณในการดำเนินการเท่ากัน แต่นโยบายแรกอาจใช้งบประมาณส่วนใหญ่ในหมวดค่าจ้าง และเงินเดือน ซึ่งยังผลให้งบประมาณที่ใช้ในการจัดการพัฒนาบุคลากรจริง ๆ มีน้อย ในขณะที่นโยบายหลังใช้งบประมาณส่วนใหญ่ในการจัดการตามนโยบาย และส่วนน้อยในการจ้างและตอบแทนเจ้าหน้าที่ ผลของนโยบายทั้งสองถ้าแตกต่างกันก็อาจเนื่องมาจากส่วนประกอบของปัจจัยนำเข้านี้ ในส่วนที่สอง จะพิจารณาถึงกระบวนการแปรสภาพปัจจัยนำเข้า มีหลายครั้งปัจจัยนำเข้าของนโยบายสองนโยบายอาจคล้ายกันทั้งในภาพรวมและส่วนประกอบ แต่ผลของนโยบายอาจต่างกัน เนื่องจากกระบวนการในการเปลี่ยนสภาพแตกต่างกัน การที่กระบวนการดังกล่าวแตกต่างกันอาจเนื่องมาจากผู้ดำเนินการแต่ละคนให้ความสำคัญแก่ปัจจัยนำเข้าแต่ละชนิดแตกต่างกัน และมักจะดำเนินการตามที่ตนเห็นว่าเหมาะสมกับสภาพและปริมาณของปัจจัยนำเข้ที่มีอยู่ ผลของ

นโยบายจึงมักแตกต่างกันตามความคิดเห็นของผู้ดำเนินการแต่ละคน

กล่าวโดยสรุป การประเมินผลนโยบายแบบนี้มีข้อดีในแง่ที่ว่า ทำให้ทราบเหตุการณ์และกิจกรรมต่าง ๆ ทั้งหมดที่เกิดขึ้นในกระบวนการแปรสภาพปัจจัยนำเข้า อย่างไรก็ตาม การพิจารณาดังกล่าวมักต้องอาศัยวิธีการวิจัยเชิงคุณภาพ ซึ่งเน้นการเข้าไปมีส่วนร่วมในเหตุการณ์ของผู้ประเมิน เน้นการสังเกตการสัมภาษณ์ ซึ่งมักใช้เวลาและความพยายามค่อนข้างมาก การประเมินผลแบบนี้จึงทำลำบาก นอกจากนั้นการใช้อัตวิสัยส่วนตัวของผู้ประเมินยังเป็นจุดที่มักถูกวิจารณ์บ่อยครั้งว่าทำให้ข้อมูลที่ได้ขาดความเชื่อถือ

4.1.4 การวิจัยสะสมทางสังคม เน้นการรวบรวมสะสมข่าวสารข้อมูลต่าง ๆ เกี่ยวกับผลการดำเนินการตามนโยบายที่ผ่านมา แล้วนำมาเปรียบเทียบ และประเมินเก็บสะสมไว้ประกอบการประเมินผลนโยบายครั้งต่อ ๆ ไป โดยทั่วไปข้อมูลที่สามารถนำมารวบรวมได้มักแบ่งเป็น 2 ประเภทใหญ่ ๆ ประเภทแรก เป็นข้อมูลที่ได้จากกรณีตัวอย่างของการกำหนดคนนโยบายและดำเนินการตามนโยบาย และประเภทที่สอง คือ รายงานการวิจัยซึ่งกล่าวถึงความสัมพันธ์ระหว่างการดำเนินการตามนโยบายและ

ผลที่เกิดตามมา ข้อมูลที่ได้จากการเก็บรวบรวมทั้งสองประเภท ต้องได้รับการวิเคราะห์และประเมินอย่างมีระบบก่อนที่จะนำไปใช้จริง การประเมินดังกล่าวมักเน้นการพิจารณาปัจจัยนำเข้า กระบวนการแปรสภาพ ผลของนโยบายและผลกระทบที่ไม่คาดหวัง การพิจารณามักมีการสร้างตัววัดเพื่อประเมินองค์ประกอบต่าง ๆ ตัวอย่างเช่น ในเรื่องปัจจัยนำเข้า อาจพิจารณาโดยวัดความพอเพียงของทรัพยากรว่ามีมากน้อยเพียงใด ในเรื่องกระบวนการแปรสภาพอาจพิจารณาโดยวัดการมีส่วนร่วมของนักวิเคราะห์นโยบายว่ามีมากน้อยเพียงใด มีมากจนถึงระดับช่วยตัดสินใจหรือเป็นเพียงแต่ผู้ป้อนข้อมูลเท่านั้น

การประเมินผลแบบนี้มีข้อดีอย่างน้อยสองประการ *ประการแรก* การประเมินผลแบบนี้ช่วยให้ได้รับข้อมูลที่สำคัญ ๆ เกี่ยวกับการดำเนินการตามนโยบายต่าง ๆ ซึ่งจะยังประโยชน์ในการสร้างข้อสรุปเพื่ออธิบาย หรือพยากรณ์การดำเนินการตามนโยบายอื่น ๆ ว่าจะมีผลอย่างไร จะมีปัญหาหรืออุปสรรคอะไรบ้าง *ประการที่สอง* การประเมินผลแบบนี้ ช่วยให้ทราบสภาพต่าง ๆ ทางสังคมทั้งที่เป็นปรนัยและอับนัย อันจะเป็นประโยชน์อย่างมากต่อการประเมินผลนโยบายอื่นที่เกิดขึ้นภายหลัง โดยนำมาเปรียบเทียบกับสภาพปัจจุบัน อย่างไรก็ตามข้อ

เสียของการประเมินผลแบบนี้ก็มีเช่นกัน ข้อที่ถูกริวิจารณ์มากที่สุดคือ ความเป็นอรรถนัยของข้อมูลที่รวบรวมมาได้ ไม่ว่าข้อมูลที่รวบรวมมาได้จะเป็นประเภท กรณีตัวอย่าง หรือรายงานการวิจัย ข้อมูลที่ได้มามากไม่มีส่วนใดที่พิจารณาถึงข้อบกพร่องของข้อมูลดังกล่าว มักกล่าวถึงเฉพาะประเด็นที่ต้องการให้ผู้อ่านรับรู้เท่านั้น การเก็บรวบรวมข้อมูลเหล่านี้ก็อีกทอดหนึ่งจึงหลีกเลี่ยงปัญหาเรื่องความถูกต้องและความเชื่อถือได้ของข้อมูลค่อนข้างลำบาก

4.2 การประเมินผลแบบเป็นทางการ (formal evaluation) คือการประเมินผลซึ่งใช้วิธีการที่เป็นวิทยาศาสตร์ในการสร้างข่าวสารที่ถูกต้องและเชื่อถือได้เกี่ยวกับผลของนโยบาย โดยประเมินผลของนโยบายจากวัตถุประสงค์หรือเป้าหมายของนโยบายที่กำหนดไว้อย่างเป็นทางการ ข้อสมมติพื้นฐานของการประเมินผลแบบนี้ คือความเชื่อที่ว่าวัตถุประสงค์หรือเป้าหมายที่กำหนดไว้อย่างเป็นทางการโดยผู้กำหนดนโยบายเป็นสิ่งที่ถูกต้องเหมาะสมแล้ว การประเมินผลแบบนี้ใช้วิธีการที่เป็นวิทยาศาสตร์ในลักษณะเดียวกับที่ใช้ในการประเมินผลแบบเทียบ เป้าหมายในการประเมินผลก็เช่นเดียวกัน คือสร้างข้อมูลที่ถูกต้องและเชื่อถือได้เกี่ยวกับผลโดยตรงของนโยบาย และผลสืบเนื่องอื่น ๆ อันเนื่องมาจากการดำเนินการตามนโยบาย ข้อแตกต่างของการประเมินผล

ทั้งสองแบบอยู่ที่เกณฑ์ที่ใช้ในการประเมิน การประเมินผลแบบเป็นทางการใช้ข้อมูลต่าง ๆ ที่มีอยู่ในกฎหมายในเค้าโครงนโยบาย รวมทั้งข้อมูลที่ได้จากการสัมภาษณ์ผู้กำหนดนโยบาย และผู้นำนโยบายไปปฏิบัติ เป็นข้อมูลในการระบุ และนิยามวัตถุประสงค์ รวมทั้งเป้าหมายที่เป็นทางการของนโยบาย ขณะเดียวกันผู้ประเมินจะไม่ตั้งคำถามเกี่ยวกับความเหมาะสมของวัตถุประสงค์ หรือเป้าหมายที่เป็นทางการดังกล่าว เกณฑ์ที่ใช้ในการประเมินจึงเน้นเกณฑ์ประสิทธิภาพ หรือประสิทธิผลของนโยบายเป็นหลัก

รูปแบบการประเมินผลแบบเป็นทางการมี 4 รูปแบบใหญ่ด้วยกัน คือ รูปแบบแรกการประเมินพัฒนาการ รูปแบบที่สอง การประเมินกระบวนการย้อนหลัง รูปแบบที่สาม การประเมินแบบทดลอง และรูปแบบสุดท้ายการประเมินผลลัพธ์ย้อนหลัง การจำแนกประเภทดังกล่าวทำโดยพิจารณามิติ 2 ด้านด้วยกัน คือมิติเวลา และมิติความสามารถในการควบคุมปัจจัยนำเข้าและกระบวนการแปรสภาพนโยบาย

เมื่อพิจารณามิติเวลา การประเมินผลสามารถแยกได้เป็น การประเมินผลเมื่อสิ้นสุดการดำเนินการเป็นระยะ ๆ และการประเมินผลเมื่อเสร็จสิ้นการดำเนินการทั้งหมด การประเมินผลแบบแรกมุ่งพิจารณาความสามารถดำเนินการให้บรรลุวัตถุประสงค์-

ประสงค์ที่เป็นทางการ ภายหลังจากที่นโยบายได้สิ้นสุดเป็นระยะ ๆ ไป การประเมินผลประเภทหลังมุ่งประเมินผลนโยบาย เพื่อพิจารณาถึงความสามารถในการบรรลุวัตถุประสงค์ที่เป็นทางการของนโยบายนั้นเมื่อเสร็จสิ้นนโยบายทั้งหมด

เมื่อพิจารณามิติความสามารถในการควบคุมปัจจัยนำเข้าและกระบวนการแปรสภาพนโยบาย การประเมินผลสามารถแยกเป็น 2 ชนิดใหญ่ ๆ การประเมินผลชนิดแรกนั้น ผู้ประเมินสามารถควบคุมปัจจัยนำเข้าและกระบวนการแปรรูปปัจจัยนำเข้าโดยตรง ตัวอย่างเช่น สามารถปรับเปลี่ยนงบประมาณที่ใช้ในการดำเนินการตามนโยบาย ปรับเปลี่ยนองค์ประกอบของนโยบาย และปรับเปลี่ยนกลุ่มเป้าหมายของนโยบายได้ ส่วนการประเมินผลชนิดหลังนั้นผู้ประเมินไม่สามารถควบคุมปัจจัยใด ๆ ทั้งสิ้น ตรงกันข้าม ผู้ประเมินต้องประเมินโดยพิจารณาจากสิ่งที่เกิดขึ้นมาแล้วทั้งสิ้น

เมื่อนำมิติเวลาและมิติความสามารถในการควบคุมปัจจัยนำเข้าและกระบวนการแปรสภาพมาพิจารณาประกอบกัน จะทำให้ได้รูปแบบของการประเมินผลแบบเป็นทางการ 4 รูปแบบด้วยกัน ดังกล่าวแล้วข้างต้น การประเมินผลแต่ละรูปแบบมีลักษณะทั่วไปดังนี้

4.2.1 การประเมินพัฒนาการ (developmental evaluation) เป็นการประเมินผลรูปแบบหนึ่งที่ออกแบบมาเพื่อประเมินการทำงานประจำวันของผู้นำนโยบายไปปฏิบัติโดยเฉพาะ การประเมินผลแบบนี้เป็นประโยชน์ในการเตือนให้ผู้นำนโยบายไปปฏิบัติทราบถึงข้อบกพร่องต่าง ๆ ที่เกิดขึ้นไม่ว่าโดยตั้งใจหรือไม่ตั้งใจก็ตาม นอกจากนั้นยังเป็นประโยชน์ในการช่วยให้แน่ใจว่ากระบวนการและขั้นตอนต่าง ๆ ที่วางแผนไว้นั้นได้รับการปฏิบัติอย่างถูกต้อง การประเมินผลรูปแบบนี้จะส่งผลให้ผู้ประเมินสามารถควบคุมการดำเนินการตามนโยบายได้โดยตรง ตัวอย่างเช่น การประเมินผลการทำงานโดยเทคนิคสมัยใหม่ โดยใช้การประเมินรูปแบบนี้ ผู้ประเมินจะสามารถควบคุมให้มีการใช้วิธีการทำงานแบบต่าง ๆ ตามแผนที่วางไว้ แล้วมีการวัดประสิทธิภาพและสัมฤทธิ์ผล ในการทำงานที่เกิดขึ้นเป็นระยะตลอดเวลาโดยใช้วัตถุประสงค์ที่กำหนดไว้ในนโยบายเป็นเกณฑ์ การประเมินผลรูปแบบนี้มีประโยชน์สำหรับการประเมินผลนโยบายใหม่ที่ต้องการความมั่นใจว่าจะประสบความสำเร็จหรือไม่อย่างไร

4.2.2 การประเมินกระบวนการย้อนหลัง

(retrospective process evaluation) ใช้ประเมินนโยบายที่ได้รับการนำไปปฏิบัติแล้วระยะเวลาหนึ่ง การประเมินผลแบบนี้เน้นการพิจารณาปัญหาและอุปสรรคที่เกิดขึ้นเมื่อนำนโยบายไปปฏิบัติ และไม่ยินยอมให้ผู้ประเมินควบคุมกิจกรรมใด ๆ ในส่วนที่เกี่ยวกับปัจจัยนำเข้าของนโยบาย เช่น งบประมาณ และกระบวนการแปรนโยบายที่เกิดขึ้น ในทางตรงกันข้ามมักใช้วิธีมองย้อนกลับมาพิจารณาเหตุการณ์และกิจกรรมต่าง ๆ ที่เกิดขึ้นเรียบร้อยแล้ว โดยเฉพาะอย่างยิ่งกิจกรรมที่เกี่ยวข้องกับผลโดยตรงและผลกระทบของนโยบาย การประเมินผลแบบนี้มักต้องใช้ระบบการรายงานภายใน ซึ่งสามารถให้ข้อมูลที่เกี่ยวข้องกับนโยบายอย่างต่อเนื่องตลอดเวลา ระบบข้อมูลทางการบริหารที่หน่วยงานราชการต่าง ๆ จัดสร้างขึ้นบางครั้งสามารถให้ข้อมูลเหล่านี้ได้ โดยเฉพาะอย่างยิ่งถ้าระบบข้อมูลนั้นบรรจุข้อมูลเกี่ยวกับกระบวนการและผลของนโยบาย

4.2.3 การประเมินแบบทดลอง (experimental

evaluation) เป็นการประเมินผลที่ผู้ประเมินสามารถควบคุมปัจจัยนำเข้าและกระบวนการแปรสภาพนโยบายได้ รูปแบบอุดมคติของการประเมินผลแบบนี้คือการประเมินผลโดยทดลองในห้องปฏิบัติการ

การทางวิทยาศาสตร์ ขั้นตอนในการประเมินผลแบบนี้จึงคล้ายคลึงกับขั้นตอนที่ใช้ในการทดลองในห้องปฏิบัติการทางวิทยาศาสตร์ รูปแบบนี้เป็นรูปแบบที่นักวิชาการหลายกลุ่มพยายามผลักดันให้เป็นรูปแบบมาตรฐานของการประเมินผล นโยบาย แนวคิดพื้นฐานของการประเมินผลแบบนี้คือ การควบคุมปัจจัยอื่น ๆ นอกเหนือจากนโยบายที่อาจมีอิทธิพลในการก่อให้เกิดผลของนโยบาย ทั้งนี้ เพื่อทดสอบว่า ผลของนโยบายที่เกิดขึ้นนั้นเนื่องมาจากนโยบายนั้นใช่หรือไม่ ในประเทศสหรัฐอเมริกา การประเมินแบบนี้ได้รับความนิยมแพร่หลายพอสมควร แต่ในประเทศไทยไม่ปรากฏให้เห็นมากนัก เนื่องจากความยากลำบากในการประเมินนั่นเอง

4.2.4 การประเมินผลลัพท์ย้อนหลัง (retrospective outcome evaluation) เป็นการประเมินผลลัพท์ของนโยบายที่ผู้ประเมินไม่สามารถควบคุมกิจกรรมใด ๆ ในส่วนที่เกี่ยวข้องกับนโยบายได้เลย ไม่ว่าจะเป็นปัจจัยนำเข้า หรือกระบวนการแปรปัจจัยดังกล่าว การควบคุมที่สามารถทำได้ดีที่สุดคือการควบคุมทางสถิติ นั่นคือ การอาศัยวิธีการทางสถิติเพื่อศึกษาผลกระทบสุทธิของนโยบายที่ดำเนินการไปว่า มีผลต่อผลลัพท์ที่เกิดขึ้นอย่างไรเมื่อปัจจัยอื่น ๆ ที่อาจมีอิทธิพลต่อผลกระทบดังกล่าวคงที่ โดย

ทั่วไปการประเมินผลแบบนี้มักแบ่งเป็น 2 ประเภทใหญ่ ๆ คือ ประเภทศึกษาในระยะเวลาหนึ่งโดยเฉพาะ (cross sectional studies) และประเภทที่ศึกษาเป็นช่วงระยะเวลานานติดต่อกัน (longitudinal studies) ประเภทแรก มุ่งประเมินผลนโยบายหนึ่ง ๆ หรือหลายนโยบายเปรียบเทียบกัน ณ เวลาใดเวลาหนึ่ง จุดหมายสำคัญของการประเมินผลแบบนี้ คือ การเปรียบเทียบว่า ลักษณะของปัจจัยนำเข้า และกระบวนการแปรสภาพอย่างไรที่ก่อให้เกิดผลที่พึงปรารถนาประเภทที่สอง มุ่งประเมินผลนโยบายหนึ่ง ๆ หรือหลายนโยบาย หลาย ๆ ครั้งต่อเนื่องกัน ทั้งนี้เพื่อศึกษาเปรียบเทียบผลของนโยบายในหลายช่วงเวลาว่ามีความแตกต่างกันอย่างไรบ้าง

4.3 การประเมินผลแบบพิจารณาความเหมาะสม

(decision theoretical evaluation) เป็นแนวทางการประเมินผลที่มุ่งสร้างข่าวสารที่ถูกต้องและเชื่อถือได้เกี่ยวกับผลของนโยบายโดยใช้คุณค่าหรือผลประโยชน์ของประชาชนกลุ่มต่าง ๆ ที่เกี่ยวข้องได้รับเป็นเกณฑ์ประเมิน ข้อแตกต่างที่สำคัญระหว่างการประเมินผลแบบนี้กับการประเมินผลของแบบแรกก็คือ การประเมินผลแบบนี้เน้นการพยายามที่จะนำเอาวัตถุประสงค์และเป้าหมายของนโยบายทั้งที่เป็นทางการและที่แอบแฝงอยู่ในรูปทรงระคนของ

ประชาชนผู้มีส่วนได้ส่วนเสียต่าง ๆ มาพิจารณาเป็นเกณฑ์ร่วมกับ วัตถุประสงค์และเป้าหมายที่กำหนดเป็นทางการ โดยผู้กำหนด นโยบายและผู้นำนโยบายไปปฏิบัติ เกณฑ์ที่ใช้ในการประเมินโดย วิธีนี้จึงกว้างขวางและครอบคลุมมากกว่าเกณฑ์ที่ใช้ในการ ประเมินโดยวิธีอื่น

การประเมินผลแบบพิจารณาความเหมาะสมนี้ มีข้อ ดีดังต่อไปนี้

1. ช่วยให้ประชาชนกลุ่มต่าง ๆ ยอมรับและนำผล การประเมินไปใช้ประโยชน์มากขึ้น เนื่องจากการประเมินแบบนี้ ตอบสนองต่อวัตถุประสงค์และเป้าหมายของประชาชนหลายกลุ่ม ที่เกี่ยวข้องและมีส่วนได้ส่วนเสีย และของผู้กำหนดนโยบายและผู้นำ นโยบายไปปฏิบัติในเวลาเดียวกัน โอกาสที่ผลการประเมินจะ เป็นที่สนใจและนำไปใช้โดยผู้มีส่วนเกี่ยวข้องต่าง ๆ จึงมีมากขึ้น
2. ช่วยให้วัตถุประสงค์หรือเป้าหมายที่กว้างเกินไปจนคลุมเครือ ได้รับการนิยามใหม่ให้เฉพาะเจาะจงขึ้น โดยอาศัยความเห็นของประชาชนและเจ้าหน้าที่กลุ่มต่าง ๆ ที่เกี่ยวข้อง ตัวอย่างเช่น รัฐบาลมีนโยบายกว้าง ๆ ว่าต้องการช่วยเหลือ เกษตรกรให้มีความเป็นอยู่ดีขึ้น การถามความเห็นของหลายกลุ่ม อาจช่วยนิยามได้ว่า การทำให้เกษตรกรมีความเป็นอยู่ดีขึ้นนั้นคือ

อะไร บางกลุ่มอาจให้ข้อคิดว่าน่าจะเป็นการมีอาหารที่สมบูรณ์ ถูกหลักโภชนาการ บางกลุ่มอาจบอกที่น่าจะเป็นการมีอำนาจปกครองตนเอง ในขณะที่บางกลุ่มอาจนิยามว่าเป็นการมีรายได้สูงขึ้น จากตัวอย่างที่ยกมาจะพบว่าทุกกลุ่มมีส่วนถูกด้วยกันทั้งสิ้น แต่ไม่มีกลุ่มใดนิยามได้ครอบคลุมหมดทุกประเด็น การประเมินแบบนี้จะใช้ข้อคิดจากทุกกลุ่มมานิยามเป็นวัตถุประสงค์หรือเป้าหมาย ความหมายจึงกว้างและครอบคลุมลึกซึ้งและเป็นจริงมากกว่าวิธีอื่นใด

3. ช่วยให้ทราบถึงวัตถุประสงค์หรือเป้าหมายทั้งหลายที่อาจขัดแย้งกันการขัดแย้งนี้เป็นเรื่องปกติธรรมดาที่เกิดขึ้นเสมอ เมื่อให้กลุ่มคนมากกว่าหนึ่งกลุ่มแสดงความคิดเห็น อย่างไรก็ตาม การเปิดโอกาสให้ความขัดแย้งดังกล่าวเปิดเผยและรับรู้ร่วมกัน ย่อมทำให้การประเมินผลนโยบายถูกต้องและรัดกุมยิ่งขึ้น

รูปแบบที่สำคัญของการประเมินผลแบบนี้มี 2 รูปแบบด้วยกัน คือ การประเมินความสามารถที่จะประเมินได้ (evaluability assessment) และการวิเคราะห์อรรถประโยชน์แบบพหุลักษณะ (multi-attribute utility analysis) รูปแบบทั้งสองส่วนเน้นความพยายามที่จะเชื่อมความสัมพันธ์ระหว่างผลของนโยบาย

กับคุณค่าและอรรถประโยชน์ของนโยบายในสายตาของประชาชนผู้มีส่วนได้ส่วนเสียในนโยบายนั้น

4.3.1 การประเมินความสามารถที่จะประเมินได้
 มุ่งตอบคำถามสำคัญที่ว่า นโยบายนั้นสามารถประเมินผลได้หรือไม่ การประเมินความสามารถที่จะประเมินได้จะเน้นการวิเคราะห์ระบบการตัดสินใจทั้งหมดที่ควรจะได้รับประโยชน์จากข่าวสารเกี่ยวกับผลลัพธ์ที่เกิดขึ้นเนื่องจากนโยบายนั้น พร้อมทั้งมุ่งพิจารณาเป้าหมาย วัตถุประสงค์ ความคาดหวังของผู้มีส่วนได้ส่วนเสียและข้อสมมติพื้นฐานที่เกี่ยวข้องกับผลของนโยบายอย่างถ่องแท้โดยทั่วไปแล้วนโยบายที่จะสามารถประเมินได้ต้องมีคุณลักษณะที่สำคัญอย่างน้อย 3 ประการ คือ เนื้อหาต้องชัดเจนไม่คลุมเครือ วัตถุประสงค์ต้องชัดเจนเข้าใจได้ และชุดของข้อสมมติพื้นฐานที่เชื่อมระหว่างผล และวัตถุประสงค์ของนโยบายต้องกำหนดไว้ชัดเจน ในการประเมินนโยบายตามรูปแบบนี้ ผู้ประเมินต้องปฏิบัติตามขั้นตอนที่กำหนดไว้ดังนี้

1. ระบุเนื้อหาและวัตถุประสงค์ของนโยบายให้ชัดเจน โดยพิจารณาถึงความคาดหวังของผู้มีส่วนได้ส่วนเสียด้วย

2. รวบรวมข่าวสารทั้งหลายทั้งปวงที่เกี่ยวข้องกับนโยบาย โดยเฉพาะข่าวสารที่จำเป็นต่อการนิยามวัตถุประสงค์ของนโยบาย กิจกรรมต่าง ๆ ที่ต้องดำเนินการและข้อสมมติพื้นฐานของนโยบายนั้น

3. สร้างตัวแบบเพื่ออธิบายผลของนโยบาย ตัวแบบที่สร้างขึ้นควรเป็นตัวแบบที่ใช้จุดยืนของกลุ่มเป้าหมายของนโยบายเป็นเกณฑ์ และควรอธิบายให้ทราบถึงกิจกรรมต่าง ๆ และผลลัพธ์ที่เกิดขึ้นเนื่องจากการดำเนินการตามกิจกรรมนั้น ๆ การอธิบายควรยึดหลักเหตุและผลเป็นสำคัญ

4. ประเมินความเป็นไปได้ของการประเมินผลโดยพิจารณาจากตัวแบบที่พัฒนาขึ้นมาในข้อ 3 ข้อควรพิจารณาคือ ความกระจ่างชัดของตัวแบบที่สร้างขึ้นเพื่ออธิบายความสัมพันธ์ระหว่างนโยบายและผล ตัวแบบที่ชัดเจนย่อมสะดวกต่อการประเมินผลมากกว่าตัวแบบที่กำกวมไม่ชัดเจน

5. หลังจากการประเมินในข้อ 4 เสร็จเรียบร้อยแล้ว ขั้นตอนถัดมาคือการรายงานผลว่านโยบายเรื่องนี้สามารถประเมินผลได้หรือไม่ให้ผู้ใช้นโยบายได้ทราบ จากนั้นผู้กำหนดคนนโยบาย และผู้ใช้นโยบายควรร่วมปรึกษาหารือว่าควรที่จะประเมินผลจริงหรือไม่ อย่างไร

4.3.2 การวิเคราะห์อรรถประโยชน์แบบพหุลักษณะ
 มุ่งตอบคำถามว่าผู้มีส่วนได้ส่วนเสียในนโยบายมีทรรศนะ และ การรับรู้เช่นไรในเรื่องอรรถประโยชน์ของนโยบายและความเป็น ไปได้ที่ผลของนโยบายจะเกิดขึ้นการประเมินผลแบบนี้ออกแบบ มาเพื่อพิจารณาความรู้สึกและทรรศนะที่เป็นตัวอรรถวิสัยของ ประชาชนผู้มีส่วนได้ส่วนเสียในนโยบายโดยเฉพาะ จุดเด่นของ การประเมินผลแบบนี้อยู่ที่การพยายามดึงคุณค่าและความเห็นของ ผู้มีส่วนได้ส่วนเสียที่แอบแฝงอยู่ ไม่สามารถเห็นได้ชัดเจนออกมา ให้เห็นชัดเจนว่าคืออะไรแน่นอนนอกจากนั้นยังเน้นการยอมรับใน ความขัดแย้งระหว่างวัตถุประสงค์และเป้าหมายของนโยบายใน สายตาของประชาชนกลุ่มต่าง ๆ ซึ่งยังผลให้สามารถให้ข้อมูลที่ เป็นประโยชน์และใช้งานได้ในทางปฏิบัติ ขั้นตอนในการประเมิน ผลแบบนี้มีทั้งสิ้น 10 ขั้นตอน ดังนี้

(1) ระบุผู้มีส่วนได้ส่วนเสียในนโยบายให้ ชัดเจน ซึ่งโดยปกติทั่วไปมักได้แก่ประชาชนกลุ่มต่าง ๆ ที่มีผล ประโยชน์ร่วมกันในนโยบายนั้น ไม่ว่าจะเป็นผู้ที่จะมีโอกาสเกี่ยว ข้องในกิจกรรมต่าง ๆ ตามที่กำหนดไว้ในนโยบาย ผู้ที่ได้รับ ประโยชน์จากการดำเนินการ หรือผู้ที่เสียประโยชน์ก็ตาม ประชาชนกลุ่มต่าง ๆ เหล่านี้มักมีจุดยืนที่แน่นอนในเรื่องนโยบาย

นั้น มีทรศนะและให้คุณค่าอรรถประโยชน์ของนโยบายแตกต่างกันออกไป

(2) แจกแจงให้ชัดเจนว่ามีการดำเนินการตามนโยบายประเภทใดบ้างที่ประชาชนที่เกี่ยวข้องตามที่ระบุไว้ในข้อ 1 อาจมีความเห็นไม่สอดคล้องกัน โดยทั่วไปมักมีอยู่อย่างน้อย 2 กลุ่ม คือ กลุ่มอนุรักษ์นิยมที่ชื่นชมกับการรักษาสภาพเดิมเอาไว้ให้มากที่สุด กลุ่มนี้มักไม่สนับสนุนการดำเนินการใดที่กลุ่มของตนต้องการปรับเปลี่ยนวิถีการดำรงชีวิต กลุ่มที่สอง คือ กลุ่มที่นิยมการเปลี่ยนแปลง กลุ่มนี้มักสนับสนุนการดำเนินการใดที่ค่อนข้างสร้างสรรค์ และไม่หยุดอยู่กับที่

(3) แจกแจงผลของนโยบายต่าง ๆ ที่อาจเกิดขึ้นได้จากการดำเนินการแต่ละประเภท การดำเนินการประเภทหนึ่งอาจส่งผลมากกว่า 1 ข้อ และผลที่เกิดขึ้นอาจทำให้เกิดผลอื่น ๆ ตามมาเป็นลูกโซ่ ผลต่าง ๆ เหล่านี้ควรมีการจัดลำดับความสำคัญตามเวลาที่เกิด

(4) แจกแจงให้เห็นชัดเจนว่าผลของนโยบายแต่ละผลนั้นมีคุณลักษณะที่สำคัญ ๆ อะไรบ้าง คุณลักษณะดังกล่าวอาจจำแนกตามผลประโยชน์ที่ได้รับ และผลเสียที่อาจเกิด

ขึ้น และอาจพิจารณาถึงลงไปถึงผลดังกล่าวตามทรรศนะของประชาชนผู้มีส่วนได้ส่วนเสียกลุ่มต่าง ๆ

(5) จัดลำดับความสำคัญของแต่ละคุณลักษณะ เพื่อชี้ให้เห็นว่าคุณลักษณะใดสำคัญมากที่สุด คุณลักษณะใดสำคัญรองลงมาตามลำดับ

(6) ให้น้ำหนักของแต่ละคุณลักษณะ โดยใช้สัญลักษณ์ทางตัวเลข โดยอาจเริ่มจากคุณลักษณะที่สำคัญน้อยที่สุดให้น้ำหนักเป็น 10 จากนั้นพยายามตอบคำถามว่าคุณลักษณะที่สำคัญถัดมามีความสำคัญเป็นกี่เท่าของคุณลักษณะที่ได้คะแนน 10 ถ้าสำคัญเป็น 2 เท่า น้ำหนักของคุณลักษณะนี้ก็จะได้คะแนนเป็น 20 (ได้มาจาก 10 คูณด้วย 2) ถ้าสำคัญเป็น 2 เท่าครึ่ง น้ำหนักที่จะได้จะเท่ากับ 25 (ได้มาจาก 10 คูณด้วย 2.5)

(7) ทำค่าของน้ำหนักที่คำนวณได้ในข้อ 6 ให้เป็นคะแนนมาตรฐานเดียวกัน โดยทั่วไปประชาชนผู้มีส่วนได้ส่วนเสียแต่ละกลุ่มมักให้ค่าน้ำหนักแก่คุณลักษณะต่าง ๆ ด้วยคะแนนรวมที่ต่างกัน ตัวอย่างเช่น ประชาชนกลุ่มหนึ่ง อาจให้คุณลักษณะ ก 50 คะแนน ข 40 คะแนน และ ค 10 คะแนน ในขณะที่ประชาชนอีกกลุ่มหนึ่งอาจให้คุณลักษณะ ก 60 คะแนน ข 8 คะแนน และ ค 60 คะแนน จะสังเกตได้ว่าคะแนนรวมของ

ทั้งสามคุณลักษณะของประชาชนกลุ่มแรกเท่ากับ 100 คะแนน กลุ่มหลังเท่ากับ 200 คะแนน เพื่อประโยชน์ในการเปรียบเทียบ ควรทำให้คะแนนเต็มนี้ให้เป็นมาตรฐานเดียวกัน โดยปกติมักใช้ 100 คะแนน การปรับค่าคะแนนทำได้โดยการหารค่าคะแนนแต่ละค่าด้วยคะแนนรวมแล้วคูณด้วย 100 ตัวอย่างเช่น คะแนนที่ให้โดยกลุ่มหลังสามารถปรับค่าได้เป็น 30 (คำนวณได้จาก 60 หารด้วย 200 คูณด้วย 100) สำหรับคุณลักษณะ ก 40 (คำนวณได้จาก 80 หารด้วย 200 คูณด้วย 100) สำหรับคุณลักษณะ ข และ 30 (คำนวณได้จาก 60 หารด้วย 200 คูณด้วย 100) สำหรับคุณลักษณะ ค ตามลำดับ

(8) คำนวณความเป็นไปได้ที่ผลของนโยบายแต่ละเรื่องจะสามารถบรรลุคุณลักษณะแต่ละชนิด การคำนวณโดยทั่วไปต้องใช้อัตวิสัยของผู้ประเมินเข้าช่วยด้วย ค่าของความเป็นไปได้ที่สูงสุด คือ 1.0 นั่นคือ โอกาสที่ผลของนโยบายชิ้นนั้นจะก่อให้เกิดคุณลักษณะนั้นมีร้อยละ 100 ค่าต่ำที่สุดคือ 0 นั่นคือผลของนโยบายชิ้นนั้นไม่มีโอกาสก่อให้เกิดคุณลักษณะนั้นเลย

(9) คำนวณอรรถประโยชน์ของผลของนโยบายแต่ละเรื่อง โดยใช้สูตร

$$U_i = W_j * U_{ij}$$

เมื่อ U_i = อรรถประโยชน์ร่วมของผลนโยบาย
เรื่องที่ i

W_j = ค่าคะแนนมาตรฐานของคุณลักษณะ
เรื่องที่ j

U_{ij} = ความเป็นไปได้ที่ผลของนโยบาย
เรื่องที่ i จะก่อให้เกิดคุณลักษณะ
ชั้นที่ j

(10) ประเมินและรายงานผล โดยใช้ให้เห็นว่า
ผลของนโยบายใดจะก่อให้เกิดประโยชน์สูงสุดต่อประชาชนผู้มีส่วน
ได้ส่วนเสีย

จากที่กล่าวมาทั้งหมดจะพบว่า การประเมินรูปแบบ
นี้เหมาะสมสำหรับการประเมินนโยบายที่มีผู้มีส่วนได้ส่วนเสีย
หลายกลุ่ม และแต่ละกลุ่มอาจมีทรรศนะและความเห็นที่ต่างกัน
อย่างไรก็ตาม การประเมินผลรูปแบบนี้จะสำเร็จไม่ได้ถ้าผู้มีส่วน
ได้ส่วนเสียไม่เข้าร่วมในการประเมินอย่างจริงจัง

กล่าวโดยสรุป การประเมินผลแบบพิจารณาความ
เหมาะสมนี้มีลักษณะเด่นสำคัญคือ นอกจากจะคำนึงถึงวัตถุประสงค์
ประสงค์ที่เป็นทางการของนโยบายแล้ว ยังตั้งคำถามถึงความ

เหมาะสมของผลลัพธ์ที่เกิดขึ้น โดยพิจารณาเปรียบเทียบกับความคาดหวังของกลุ่มผู้มีส่วนได้ส่วนเสียต่าง ๆ รูปแบบในการประเมินแนวนี้นี้มี 2 รูปแบบที่สำคัญ การประเมินแต่ละรูปแบบล้วนเน้นการพิจารณาความคาดหวังของผู้มีส่วนได้ส่วนเสียประกอบด้วย

5. บทสรุป

วิธีการที่กล่าวมาทั้งหมดเป็นเพียงส่วนหนึ่งของวิธีการหลาย ๆ วิธีที่จะช่วยให้การประเมินผลนโยบายมีประสิทธิภาพและบรรลุวัตถุประสงค์ ผลของนโยบายที่ประเมินได้นั้นจะเป็นข้อมูลที่ป้อนกลับไปยังนักวิเคราะห์นโยบาย และ/หรือผู้กำหนดนโยบาย ซึ่งจะนำผลดังกล่าวประกอบการพิจารณาเพื่อปรับเปลี่ยนหรือยุตินโยบายต่อไป วงจรชีวิตของนโยบายซึ่งเริ่มต้นจากการกำหนดนโยบาย การนำนโยบายไปปฏิบัติและจะสิ้นสุดลงที่การประเมินผลนโยบาย อย่างไรก็ตามการสิ้นสุดนี้อาจเป็นการเริ่มต้นของนโยบายใหม่ หรือเป็นการฟื้นตัวของนโยบายเดิมก็ได้ สุกแต่แต่ผลการประเมินที่ปรากฏ “ชีวิต” ของนโยบายสาธารณะจะเป็นวัฏจักรวนเวียนไปเช่นนี้ไม่มีที่สิ้นสุด เหมือนดังเช่นที่ แอรอน วิลดาฟสกี กล่าวไว้ --- เราไม่เคย “แก่” ปัญหา เพียงแต่

บรรเทาปัญหา ยืดอายุปัญหา หรือเปลี่ยนสภาพปัญหาให้หายไป
 จากการรับรู้ของสาธารณชนเท่านั้น
 แต่อย่างไรก็ตาม ผู้ที่ทำการประเมินผลนั้น จะต้องเลือก
 วิธี การประเมินผลให้ สอดคล้องและเหมาะสมกับนโยบาย
 สาธารณะในแต่ละเรื่องด้วย

๘๘๘๘๘๘๘๘๘๘๘๘

ตามวิธีของ...
 หลักการ...
 วิธีการ...
 ผลกระทบ...
 การประเมินผล...
 ความสอดคล้อง...
 การเลือกวิธี...
 การประเมินผล...
 การเลือกวิธี...
 การประเมินผล...
 ความสอดคล้อง...
 การเลือกวิธี...
 การประเมินผล...

เชิงอรรถ

- ¹Emil J. Posavac and Raymond G. Carey, **Program Evaluation : Methods and Case Studies.** (Enclewood. Cliffs, N.J. : Prentice-Hall, 1980).
- ²James E. Anderson, **Public Policy Making.** (New York : Praeger Publishers, 1975).
- ³Charles O. Jones, **An Introduction to the Study of Public Policy.** (Belmont, CA : Wads worth Co., 1970).
- ⁴E.R. House, "Justice in Evaluation," in G.V. Glass (ed.) **Evaluation Studies Review Annual.** Vol. I (Beverly Hills, CA. : Sage, 1976).
- ⁵William N. Dunn., **Public Policy Analysis : An Introduction.** (Englewood Cliffs, N.J. : Prentice - Hall, 1981).
- ⁶Worthen and Sanders, **Educational Evaluation : Theory and Practice.**

⁷Joseph D. Comtois, "A Unifying Evaluation Framework,"
The Bureaucrat. Summer 1981.

⁸โปรดดู T. Cook and R. Reichardt (eds.) **Qualitative and Quantitative Methods in Evaluation Research**. (Beverly Hills, CA. : Sage, 1979) และ William J. Filstead "Using Qualitative Methods in Evaluation Research : An Illustrative Bibliography," *Evaluation Review* Vol. 5 No. 2, April 1981.

⁹Charles O. Jones, **op.cit.**

¹⁰William N. Dunn, **op.cit.**

¹¹รายละเอียดที่จะกล่าวต่อไปนำมาจาก William N. Dunn, **op.cit.**
chs. 8 & 9.