

การประเมินอภิมาน และการประยุกต์ใช้แนวทางการประเมินอภิมาน

Meta-evaluation and the application of Meta-evaluation

สุวิชญา สิริยานนท์¹ รุ่งฤดี กล้าหาญ²Suwitchaya Siriyanon¹ Rungrudee Klaharn²¹นักวิชาการศึกษา สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ¹Educator, Educational and Psychological Test Bureau, Srinakharinwirot University

Corresponding author, E-Mail: suwitchaya@g.swu.ac.th

²ผู้ช่วยศาสตราจารย์ ดร. สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ²Assistant Professor, Ph.D., Educational and Psychological Test Bureau, Srinakharinwirot University,

E-Mail: rungrudee@g.swu.ac.th

Received: August 21, 2020; Revised: June 27, 2021; Accepted: July 25, 2021

บทคัดย่อ

บทความเรื่องนี้เป็นการศึกษาเกี่ยวกับการประเมินอภิมาน และการประยุกต์ใช้แนวทางการประเมินอภิมาน การประเมินอภิมานเป็นเครื่องมือในการพิจารณาตัดสินคุณภาพของการประเมิน สำหรับประเทศไทยทำการประเมินอภิมานค่อนข้างน้อย ทำให้มีรายงานการประเมินอภิมานไม่มากนัก การประเมินอภิมานมีประโยชน์ต่อนักประเมิน ทำให้สามารถตรวจสอบคุณภาพของการประเมินตามมาตรฐานของการประเมินอภิมาน ซึ่งประกอบด้วย มาตรฐานด้าน วัตถุประสงค์ มาตรฐานด้านความเป็นไปได้ มาตรฐานด้านความเหมาะสม มาตรฐานด้านความถูกต้อง และมาตรฐานด้านความรับผิดชอบ ผลการประเมินสามารถนำไปพัฒนาการประเมินให้มีคุณภาพยิ่งขึ้น รวมทั้งทำให้ทราบถึงจุดเด่นและจุดด้อยของการประเมิน

คำสำคัญ: การประเมินอภิมาน**Abstract**

This article is a study of Meta-evaluation. And the application of meta-evaluation. Meta-evaluation is a tool for determining the quality of the evaluation. In Thailand, very few Meta-evaluations were conducted. Causing rarely reported meta-evaluation. A meta-evaluation is useful for assessors. Checking the quality of the evaluation, in accordance with the standards of the meta-evaluation consists of utility standards, feasibility standards, propriety standards, accuracy standards and evaluation accountability standards. In order to improve the quality evaluation and reporting the strengths and weaknesses of the evaluation.

Keywords: Meta-evaluation

บทนำ

ในประเทศไทยการประเมินอภิมานเป็นที่รู้จักกันดีในแวดวงสาขาวิชาที่เกี่ยวข้องกับการประเมิน การประเมินอภิมานมีความสำคัญ สามารถใช้การประเมินอภิมานตรวจสอบและตัดสินโครงการต่าง ๆ อย่างหลากหลายทั้งสินค้าและบริการ นักประเมินมีบทบาทสำคัญในสังคมเช่นเดียวกับวิชาชีพ แพทย์ หนายความ วิศวกร และนักบัญชี การประเมินอภิมานมีความจำเป็นในการประเมินงานทุกประเภท เช่น โครงการ ผลิตภัณฑ์ ระบบ สถาบัน ทฤษฎี แบบจำลอง นักศึกษา และบุคลากร (Stufflebeam, 2001) การประเมินอภิมาน เป็นการประเมินขั้นตอนการประเมินหรือการประเมินงานประเมิน การประเมินอภิมานเป็นเครื่องมือตัดสินคุณภาพของการประเมิน และประเมินคุณภาพของงานประเมิน ซึ่งแพร่หลายในต่างประเทศ การประยุกต์ใช้การประเมินอภิมานเพื่อตรวจสอบคุณภาพของการประเมินสามารถใช้เกณฑ์มาตรฐานของการประเมินเพื่อตัดสินคุณภาพ ทำให้ทราบถึงจุดเด่นและจุดด้อยของการประเมิน การประเมินอภิมานจึงมีความสำคัญและเป็นประโยชน์ในด้านการติดตามและตรวจสอบงานประเมินทุกประเภทดังกล่าวข้างต้นให้มีคุณภาพตามมาตรฐานของการประเมิน ทำให้ผู้เกี่ยวข้องเกิดความมั่นใจในการใช้ผลการประเมินและคุณภาพของการประเมินสามารถนำผลการประเมินอภิมานไปใช้ประโยชน์ในการพัฒนาหรือปรับปรุงการประเมินให้มีคุณภาพมากยิ่งขึ้น และสร้างความรับผิดชอบที่สามารถตรวจสอบได้ต่อสาธารณชน บทความนี้จึงมีวัตถุประสงค์ เพื่อเสนอแนวคิดและหลักการเกี่ยวกับการประเมินอภิมาน และการประยุกต์ใช้แนวทางการประเมินอภิมาน โดยนำเสนอในประเด็นดังต่อไปนี้ คือ ความหมายของการประเมินอภิมาน ประเภทของการประเมินอภิมาน เครื่องมือที่ใช้ในการประเมินอภิมานและเกณฑ์มาตรฐานการประเมินอภิมาน ขั้นตอนการดำเนินงานประเมินอภิมาน ประโยชน์ของการประเมินอภิมาน แนวทางในการดำเนินงานประเมินอภิมาน และข้อควรระวังในการประเมินอภิมาน

ความหมายของการประเมินอภิมาน

นักวิชาการด้านการประเมินได้ให้ความหมายของการประเมินอภิมานไว้ดังตัวอย่างต่อไปนี้

Scriven เป็นผู้บัญญัติศัพท์คำว่า metaevaluation ตั้งแต่ปี ค.ศ.1969 (Stufflebeam, 2001) Scriven (Scriven, 1969 อ้างถึงใน Stufflebeam, 2001) ให้ความหมายว่า การประเมินอภิมานว่า หมายถึง การประเมินงานประเมินหลาย ๆ เรื่อง หรือการประเมินโดยผู้ประเมินหลาย ๆ คน หรือเป็นการประเมินขั้นตอนการประเมิน

Cook and Gruder (1978) ให้ความหมายว่า การประเมินอภิมาน หมายถึง การประเมินของการประเมินผลสรุปรวมที่มีการออกแบบการประเมินอย่างเป็นระบบ โดยมีการเก็บรวบรวมข้อมูลโดยตรงจากผู้เกี่ยวข้องในกิจกรรม

Patton (1997) กล่าวว่า การประเมินอภิมาน คือ การประเมินงานประเมิน ซึ่งอิงตามเกณฑ์มาตรฐานและหลักการของวิชาชีพการประเมิน

Stufflebeam (2001) ให้ความหมายว่า การประเมินอภิมาน หมายถึง กระบวนการบรรยาย การได้รับและประยุกต์ใช้ข้อมูลสารสนเทศเชิงบรรยาย และข้อมูลสารสนเทศเชิงการตัดสินใจเกี่ยวกับบรรดประโยชน์ ความเป็นไปได้ ความเหมาะสม และความถูกต้องของการประเมิน เพื่อชี้แนวทางการประเมินและรายงานจุดแข็งจุดอ่อนของการประเมินต่อสาธารณะ แสดงให้เห็นถึงความสามารถ ความซื่อสัตย์สุจริต ความน่าเชื่อถือ และความรับผิดชอบต่อสังคม

ศิริชัย กาญจนวาสี (2552) ให้ความหมายว่า การประเมินอภิมาน หมายถึง การตัดสินคุณภาพของการประเมิน ถ้าการประเมินใดมีคุณภาพ การประเมินนั้นจะต้องประกอบด้วยคุณภาพของการกำหนดคำถามวิธีการ ประเมินและผลการประเมิน การประเมินงานประเมินจึงอาจใช้เป็นกลไกของการควบคุมคุณภาพของการประเมิน โดยการติดตามตรวจสอบการประเมิน การวิพากษ์วิจารณ์ การเปรียบเทียบกับเกณฑ์มาตรฐานของการประเมินหรือการประเมินซ้ำ

สมหวัง พิธิยานุวัฒน์ (2559) ให้ความหมายว่า การประเมินอภิมาน หมายถึง การตัดสินคุณค่าภายนอกและคุณค่าภายในของงานประเมิน

จากความหมายของการประเมินอภิมานดังกล่าว สรุปได้ว่า การประเมินอภิมาน หมายถึง กระบวนการประเมินซ้ำหรือการประเมินงานประเมิน เพื่อตัดสินคุณภาพของการประเมินโดยเปรียบเทียบกับเกณฑ์มาตรฐานการประเมินสามารถนำผลการประเมินอภิมานไปใช้ประโยชน์ในการพัฒนาคุณภาพการประเมินหรือนักประเมิน ให้ข้อเสนอแนะทั้งจุดเด่นและจุดด้อยของการประเมิน

ประเภทของการประเมินอภิมาน

การประเมินอภิมานเป็นกระบวนการ การประเมินอภิมาน แบ่งออกเป็น 2 ประเภท ได้แก่ การประเมินอภิมานแบบก้าวหน้า (formative meta-evaluation) และการประเมินอภิมานแบบสรุปรวม (summative meta-evaluation) โดยนักวิชาการได้อธิบายสาระสำคัญไว้ ผู้เขียนสรุปได้ดังนี้

1. การประเมินอภิมานแบบก้าวหน้า

สมหวัง พิธิยานุวัฒน์ (2559) กล่าวว่า การประเมินอภิมานแบบก้าวหน้า เป็นการวางแผนก่อนปฏิบัติการประเมิน ช่วยให้นักประเมินตัดสินใจว่าจะประเมินอะไร นิยามปัญหาที่ประเมิน กำหนดผู้เกี่ยวข้องในการประเมิน หรือลูกค้าของการประเมิน วัตถุประสงค์ประสงค์ของการประเมินให้กระจ่าง กำหนดข้อสัญญา กำหนดข้อมูลที่ต้องการ กำหนดวิธีการเก็บรวบรวมข้อมูลและวิธีวิเคราะห์ข้อมูล กำหนดตารางประเมิน เตรียมข้อสรุปและข้อเสนอแนะ สอดคล้องกับที่ Stufflebeam (2001) และ Gallegos (1994) ได้กล่าวไว้ว่า การประเมินอภิมานแบบก้าวหน้า เป็นการประเมินงานที่มีลักษณะการใช้การประเมินเชิงสร้างสรรค์ โดยเป็นการวางแผนก่อนปฏิบัติการประเมิน การประเมินอภิมานแบบก้าวหน้าช่วยให้นักประเมินตัดสินใจเกี่ยวกับสิ่งที่ประเมินว่าจะประเมินอะไร นิยามปัญหาที่จะประเมินอย่างไร การกำหนดผู้ใช้ผลการประเมินและผู้เกี่ยวข้องกับการประเมิน การทำให้วัตถุประสงค์ของการประเมินชัดเจน รวมทั้งข้อตกลงสัญญาและเงื่อนไขในการประเมิน กำหนดวิธีรวบรวมข้อมูล วิธีวิเคราะห์ข้อมูล กำหนดตารางของการประเมิน จัดหาทีมประเมิน กำหนดงบประมาณ เตรียมข้อสรุปและข้อเสนอแนะ และเลือกรูปแบบและวิธีการรายงานผลการประเมิน

2. การประเมินอภิมานแบบสรุปรวม

การประเมินอภิมานแบบสรุปรวม มีความสำคัญเพราะได้แสดงให้เห็นถึงคุณค่ารวมของงานการประเมิน ทำให้มองเห็นว่าการประเมินนี้เป็นไปตามเกณฑ์มาตรฐานของการประเมินหรือไม่ พร้อมทั้งจะให้ตรวจสอบและมีการรายงานการประเมินผลการประเมินให้ผู้เกี่ยวข้องและสาธารณชนได้ทราบถึงความเป็นมาตรฐานของงานประเมิน ผลการประเมินอภิมานแบบสรุปรวมช่วยให้ผู้ใช้ผลการประเมินมีความมั่นใจและตัดสินใจได้ว่าควรยอมรับผลประเมินและข้อเสนอแนะจากงานประเมินมากน้อยเพียงใด (Gallegos, 1994; Stufflebeam, 2001) เช่นเดียวกับสมหวัง พิธิยานุวัฒน์ (2559) ได้อธิบายไว้สอดคล้องกันว่า การประเมินอภิมานแบบสรุปรวม เป็นการแสดงถึงคุณค่ารวมของงานประเมิน ซึ่งแสดงถึงความพร้อมที่ให้ผู้ตรวจสอบได้ของนักประเมิน โดยรายงานให้สาธารณชนทราบถึงความเป็นมาตรฐานของงานประเมินผลการประเมินอภิมานแบบสรุปรวม ช่วยให้ผู้ที่ใช้ผลการประเมินตัดสินใจได้ว่าควรยอมรับผลการประเมินและข้อเสนอแนะจากงานประเมินมากน้อยเพียงใด

เมื่อพิจารณาประเภทของการประเมินอภิมานจากนักวิชาการ ผู้เขียนสรุปได้ว่า การประเมินอภิมานเป็นกระบวนการ นักวิชาการแบ่งประเภทของการประเมินอภิมานตามช่วงเวลาในการประเมิน โดยแบ่งออกเป็น 2 ประเภท ได้แก่ การประเมินอภิมานแบบก้าวหน้า ซึ่งเป็นการประเมินในระยะแรกเริ่มโครงการ เป็นการวางแผนก่อนปฏิบัติการประเมิน ทำให้ผู้ประเมินเห็นทิศทางการประเมินว่าจะประเมินอะไร ประเมินอย่างไร และกำหนดผู้เกี่ยวข้องและสิ่งที่เกี่ยวข้องกับการประเมิน เช่น วัตถุประสงค์ของการประเมิน วิธีรวบรวมข้อมูล วิธีวิเคราะห์ข้อมูล ตารางการประเมินงบประมาณ รูปแบบการประเมิน และวิธีการรายงานผลการประเมิน ส่วนอีกประเภทคือ การประเมินอภิมานแบบสรุปรวมเป็นกระบวนการประเมินในระยะต่อมา ซึ่งเป็นการประเมินผลสรุป แสดงถึงคุณค่ารวมของงานประเมินว่าเป็นไปตาม

เกณฑ์มาตรฐานการประเมินหรือไม่ ช่วยให้ผู้ใช้ผลการประเมินสามารถพิจารณาตัดสินใจที่จะยอมรับผลประเมินและข้อเสนอแนะจากงานประเมิน

เครื่องมือที่ใช้ในการประเมินอภิมานและเกณฑ์มาตรฐานการประเมินอภิมาน

Hanssen (2008 อ้างถึงในพินดา วราสุนันท์, 2556) ได้ให้แนวทางสำหรับคำถามเพื่อใช้ในการนำไปสู่การประเมินอภิมาน ประกอบด้วย 1) อะไรคือจุดแข็ง จุดอ่อนในการประเมิน และจะพัฒนาอย่างไร 2) การประเมินมีประโยชน์อย่างไร และ 3) กรอบการประเมิน วิธีการประเมินมีมาตรฐานหรือไม่ น่าเชื่อถือหรือยอมรับได้หรือไม่ จากคำถามนี้มีความเกี่ยวข้องกับเครื่องมือและมาตรฐานการประเมิน ซึ่งผู้เชี่ยวชาญด้านการประเมินได้พัฒนาเครื่องมือและกำหนดมาตรฐานการประเมินไว้ ทำให้ผู้ประเมินคนอื่น ๆ สามารถพัฒนาเครื่องมือนี้และปรับใช้ให้เหมาะสมกับบริบทของการประเมินในแต่ละโครงการ สำหรับการประเมินโครงการต่าง ๆ โดยเฉพาะการประเมินอภิมาน นักประเมินส่วนใหญ่ได้พัฒนาเครื่องมือในการประเมินอภิมานโดยอิง 2 แนวคิดหลักของนักประเมินคนสำคัญ คือ Scriven และ Stufflebeam ซึ่งแนวคิดนี้เป็นที่นิยมอย่างแพร่หลาย ผู้เขียนขอแนะนำแนวคิดดังกล่าว ดังนี้

Scriven (2007) ได้พัฒนาแบบตรวจสอบรายการประเมิน มีชื่อว่า Key Evaluation Checklist (KEC) แบบตรวจสอบรายการประเมินนี้เป็นแบบฟอร์มประกอบด้วยข้อรายการซึ่งกำหนดสิ่งที่มีในรายการประเมิน เพื่อให้ผู้ประเมินใช้ตรวจสอบว่าได้มีการดำเนินงานตามข้อรายการเหล่านั้นหรือไม่ ทั้งนี้เพื่อมิให้ละเอียดหรือลืมนำสิ่งที่พึงกระทำ เป็นการตอกย้ำให้ผู้ประเมินปฏิบัติตามแผนงานที่กำหนดหรือแนวทางที่ควรปฏิบัติ (สุวิมล ว่องวานิช, 2558) จากการทบทวนวรรณกรรม Scriven (2005, 2007, 2015) พบว่า Scriven ได้พัฒนาแบบตรวจสอบรายการประเมินเป็นระยะ คือ ค.ศ. 1991 Scriven เสนอรายการตรวจสอบการประเมินหลัก KEC เพื่อใช้ในการประเมินคุณภาพงานประเมินอภิมาน โดยแบ่งเป็น 13 รายการ ได้แก่ 1) การบรรยาย 2) ความเป็นมาและบริบท 3) ผู้เกี่ยวข้อง 4) แหล่งที่มา 5) คุณค่าของการประเมิน 6) กระบวนการประเมิน 7) ผลลัพธ์ของการประเมิน 8) ความคุ้มค่าของการประเมิน 9) การเปรียบเทียบผลการประเมิน 10) การสรุปอ้างอิงผลการประเมิน 11) ความสำคัญของการประเมิน 12) ข้อเสนอแนะการประเมิน และ 13) การรายงาน

ค.ศ. 2015 Scriven ได้เสนอแบบตรวจสอบรายการการประเมินหลักฉบับใหม่ ประกอบด้วยรายการตรวจสอบ 18 รายการ โดยแบ่งเป็น 4 หมวดหมู่หลัก ได้แก่ 1) การตรวจสอบเบื้องต้น มีรายการตรวจสอบ 3 รายการ ประกอบด้วย บทสรุป ความชัดเจน การออกแบบและเครื่องมือ 2) การตรวจสอบพื้นฐาน มีรายการตรวจสอบ 5 รายการ ประกอบด้วย ความเป็นมา คำอธิบาย ผู้ได้รับผลกระทบ ความเหมาะสมของการใช้ทรัพยากร และคุณค่าของการประเมิน 3) การประเมินย่อย มีรายการตรวจสอบ 5 รายการ ประกอบด้วย กระบวนการประเมิน ผลลัพธ์ของการประเมิน ความคุ้มค่าของการประเมิน การเปรียบเทียบผลการประเมิน และการสรุปอ้างอิงผลการประเมิน และ 4) การสรุป มีรายการตรวจสอบ 5 รายการ ประกอบด้วย การสังเคราะห์การประเมิน ข้อเสนอแนะการประเมิน ความรับผิดชอบและเหตุผล การรายงานการประเมิน และการประเมินอภิมาน

สำหรับเกณฑ์มาตรฐานการประเมินอภิมาน Scriven (2015) ใช้เครื่องมือชื่อแบบตรวจสอบการประเมินหลักฉบับใหม่ (Key evaluation checklist: KEC) เพื่อใช้ในการประเมินคุณภาพงานประเมินในการประเมินอภิมาน โดยแบ่งเป็น 4 หมวดหมู่หลัก จำนวน 18 รายการ มีรายละเอียดดังนี้

หมวดหมู่ที่ 1 การตรวจสอบเบื้องต้น (Preliminaries) รายการตรวจสอบ 3 รายการ ประกอบด้วย บทสรุป ความชัดเจน การออกแบบและเครื่องมือ

หมวดหมู่ที่ 2 การตรวจสอบพื้นฐาน (Foundations) รายการตรวจสอบ 5 รายการ ประกอบด้วย ความเป็นมา คำอธิบาย ผู้ได้รับผลกระทบ ความเหมาะสมของการใช้ทรัพยากร และคุณค่าของการประเมิน

หมวดที่ 3 การประเมินย่อย (Subevaluations) รายการตรวจสอบ 5 รายการ ประกอบด้วย กระบวนการประเมิน ผลลัพธ์ของการประเมิน ความคุ้มค่าของการประเมิน การเปรียบเทียบผลการประเมิน และการสรุปอ้างอิงผลการประเมิน
 หมวดที่ 4 การสรุป (Conclusions & Implications) รายการตรวจสอบ 5 รายการ ประกอบด้วย การสังเคราะห์การประเมิน ข้อเสนอแนะการประเมิน ความรับผิดชอบเหตุผล การรายงานการประเมิน การประเมินอภิมาน

ตัวอย่างแบบตรวจสอบรายการประเมินอภิมานฉบับใหม่ ชื่อ Key evaluation checklist รูปแบบของแบบตรวจสอบรายการมี 4 หมวดหมู่หลัก ดังตารางที่ 1

ตารางที่ 1 ตัวอย่างแบบตรวจสอบรายการประเมินอภิมานฉบับใหม่ของ Scriven (2015)

	รายการตรวจสอบ	การประเมิน	การอธิบายเหตุผลในการประเมิน
หมวดที่ 1 การตรวจสอบเบื้องต้น	บทสรุป		
	ความชัดเจน		
	การออกแบบและเครื่องมือ		
หมวดที่ 2 การตรวจสอบพื้นฐาน	ความเป็นมา		
	คำอธิบาย		
	ผู้ได้รับผลกระทบ		
	ความเหมาะสมของการใช้ทรัพยากร		
หมวดที่ 3 การประเมินย่อย	คุณค่าของการประเมิน		
	กระบวนการประเมิน		
	ผลลัพธ์ของการประเมิน		
	ความคุ้มค่าของการประเมิน		
	การเปรียบเทียบผลการประเมิน		
หมวดที่ 4 การสรุป	การสรุปอ้างอิงผลการประเมิน		
	การสังเคราะห์การประเมิน		
	ข้อเสนอแนะการประเมิน		
	ความรับผิดชอบเหตุผล		
	การรายงานการประเมิน		
	การประเมินอภิมาน		

ส่วน Stufflebeam ได้เขียนบทความเกี่ยวกับแนวปฏิบัติในการสร้างแบบตรวจสอบรายการประเมิน และจัดทำแบบตรวจสอบรายการ ชื่อ Checklists Development Checklist (Stufflebeam, 2000) ซึ่งบทความดังกล่าวได้ให้ข้อคิดเห็นว่าแบบตรวจสอบรายการประเมิน ช่วยทำให้เกิดความเป็นปรนัย ความน่าเชื่อถือ และการทำซ้ำได้ เป็นประโยชน์ต่อการวางแผนงาน การกำกับ และให้แนวทางการปฏิบัติงาน และช่วยประเมินผลลัพธ์ที่เกิดขึ้น แบบตรวจสอบรายการเป็นประโยชน์ต่อการประเมินความก้าวหน้าและการประเมินผลสรุป แบบตรวจสอบรายการ ประกอบด้วย จุดตรวจสอบ 12 จุด แต่ละจุดตรวจสอบยังมีจุดตรวจสอบย่อยอีกหลายจุด แบบตรวจสอบรายการ CDC ของ Stufflebeam ประกอบด้วย 1) การกำหนดจุดเน้นของงานที่ต้องการตรวจสอบ 2) การสร้างรายการแสดงจุดตรวจสอบ 3) การจัดประเภทและจัดเรียงข้อรายการที่เป็นจุดตรวจสอบ 4) การนิยามและสร้างความกระจ่างในแต่ละกลุ่มข้อรายการ 5) การกำหนดลำดับของกลุ่มข้อรายการ 6) การทบทวนแบบตรวจสอบรายการขั้นต้น 7) การปรับปรุงเนื้อหาในแบบ

ตรวจสอบรายการ 8) การจัดทำรูปแบบของแบบตรวจสอบรายการ 9) การประเมินแบบตรวจสอบรายการ 10) การทดลองใช้แบบตรวจสอบรายการที่สร้างเสร็จ 11) การเผยแพร่แบบตรวจสอบรายการ และ 12) การทบทวนและปรับปรุงแบบตรวจสอบรายการเป็นระยะ ๆ

นอกจากนี้ Stufflebeam และคณะ ได้พัฒนาเกณฑ์มาตรฐานการประเมินเพื่อใช้ในการพิจารณาตัดสินคุณค่างานประเมิน สำหรับมาตรฐานการประเมินโปรแกรมที่เป็นที่รู้จักและใช้กันอย่างแพร่หลายซึ่งมีการตีพิมพ์ครั้งแรก (ค.ศ. 1988) มี 21 มาตรฐาน คือ มาตรฐานการประเมินบุคลากร (Personnel Evaluation Standards) และครั้งที่ 2 (ค.ศ. 1994) คือ มาตรฐานการประเมินโครงการ (Program Evaluation Standards) มี 30 มาตรฐาน ประกอบด้วย มาตรฐานการประเมินอภิมาน มี 4 ด้าน คือ มาตรฐานด้านอรรถประโยชน์ (Utility Standards) มาตรฐานด้านความเป็นไปได้ (Feasibility Standards) มาตรฐานด้านความเหมาะสม (Propriety Standards) และมาตรฐานด้านความถูกต้อง (Accuracy Standards) ต่อมาใน ค.ศ. 2011 ได้พิจารณาเพิ่มมาตรฐานการประเมินอภิมาน 1 ด้าน คือ มาตรฐานด้านความรับผิดชอบ (Evaluation Accountability Standards) ตามหนังสือชื่อว่า The program evaluation standards: A guide for evaluators and evaluation users อยู่ภายใต้ลิขสิทธิ์ของ JCSEE (Yarbrough, Shulha, Hopson, and Caruthers, 2011) โดยมีรายละเอียดดังนี้

1. มาตรฐานด้านอรรถประโยชน์ เป็นมาตรฐานมีวัตถุประสงค์เพื่อเพิ่มขอบเขตผู้มีส่วนได้ส่วนเสียของโครงการที่ได้รับประโยชน์จากกระบวนการและผลผลิตของแบบประเมินที่ตอบสนองความต้องการ ประกอบด้วยเกณฑ์ จำนวน 8 ข้อ ดังนี้

U1 ความน่าเชื่อถือของผู้ประเมิน ผู้ประเมินควรถูกประเมิน โดยบุคคลที่มีคุณสมบัติเหมาะสม และน่าเชื่อถือต่อบริบทการประเมิน

U2 ความสนใจต่อผู้มีส่วนได้ส่วนเสีย การประเมินควรให้ความสนใจต่อบุคคล และกลุ่มคนที่ลงทุนในโครงการ ได้รับผลกระทบจากการประเมินอย่างเต็มที่

U3 วัตถุประสงค์การเจรจาต่อรอง การระบุวัตถุประสงค์ของการประเมินและปรับตามความต้องการของผู้มีส่วนได้ส่วนเสีย

U4 คุณค่ามีความชัดเจน การประเมินควรชี้แจง ระบุคุณค่าส่วนบุคคล และคุณค่าด้านวัฒนธรรมทั้งในด้านวัตถุประสงค์ กระบวนการและการตัดสินให้ชัดเจน

U5 สารสนเทศที่ตรงประเด็น การประเมินมีข้อมูลควรตอบสนองความต้องการของผู้มีส่วนได้ส่วนเสียตามที่กำหนดไว้

U6 กระบวนการและผลผลิตที่มีความหมาย การประเมินควรสร้างกิจกรรม การอธิบาย และการตัดสินในรูปแบบที่กระตุ้นให้ผู้เข้าร่วมสามารถค้นพบ ตีความใหม่หรือทบทวนความเข้าใจ

U7 การสื่อสารและการรายงานที่ทันเวลาและเหมาะสม การประเมินควรพิจารณาความต้องการข้อมูลของกลุ่มเป้าหมายอย่างต่อเนื่อง

U8 คำนี้ถึงผลที่ตามมา การประเมินควรส่งเสริมการใช้แบบประเมินด้วยความรับผิดชอบ การปรับใช้ในขณะเดียวกันควรป้องกันผลกระทบด้านลบ และการทำงานในทางที่ไม่เหมาะสม

2. มาตรฐานด้านความเป็นไปได้ เป็นมาตรฐานมีวัตถุประสงค์เพื่อเพิ่มประสิทธิภาพและประสิทธิผลของการประเมิน ประกอบด้วยเกณฑ์จำนวน 4 ข้อ ดังนี้

F1 การบริหารโครงการ ผู้ประเมินควรใช้กลยุทธ์การจัดการโครงการที่มีประสิทธิภาพ

F2 กระบวนการปฏิบัติจริง การปฏิบัติประเมินควรทำได้จริง และตอบสนองต่อวิธีการดำเนินโครงการ

F3 ความเป็นไปได้เชิงบริบท การประเมินควรรับรู้ ติดตาม และสร้างความสมดุลให้กับผลประโยชน์ทางวัฒนธรรม ผลประโยชน์ทางการเมือง รวมถึงความต้องการส่วนตัวและความต้องการของกลุ่ม

F4 การใช้ทรัพยากร การประเมินควรใช้ทรัพยากรอย่างมีประสิทธิภาพและประสิทธิผล

3. มาตรฐานด้านความเหมาะสม เป็นมาตรฐานที่สนับสนุนว่า ในการประเมิน อะไรคือสิ่งที่มีความเหมาะสม ความยุติธรรม และความถูกต้อง ประกอบด้วยเกณฑ์ จำนวน 7 ข้อ ดังนี้

P1 มุ่งเน้นการบริการที่ตอบสนองและความครอบคลุม การประเมินควรตอบสนองต่อผู้มีส่วนได้ส่วนเสีย และชุมชน

P2 ข้อตกลงอย่างเป็นทางการ การประเมินควรเจรจาข้อตกลงเพื่อสร้างความชัดเจนให้กับภาระหน้าที่ของผู้กพัน คำนี้ถึงความต้องการ ความคาดหวังและบริบททางวัฒนธรรม และผู้มีส่วนได้ส่วนเสียอื่น ๆ

P3 สิทธิทางมนุษยชนของความเคารพนับถือ ผู้ประเมินควรออกแบบการประเมิน ดำเนินการเพื่อปกป้องสิทธิมนุษยชน สิทธิตามกฎหมาย รักษาศักดิ์ศรีของผู้เข้าร่วมและผู้มีส่วนได้ส่วนเสีย

P4 ความชัดเจนและยุติธรรม การประเมินควรเข้าใจได้ โดยพิจารณาถึงความต้องการและวัตถุประสงค์ของผู้มีส่วนได้ส่วนเสีย

P5 ความโปร่งใสและการเปิดเผยข้อมูล การประเมินควรให้คำอธิบายข้อค้นพบข้อจำกัด และข้อสรุปต่อผู้มีส่วนได้ส่วนเสียทั้งหมด เว้นแต่ว่าการกระทำดังกล่าวเป็นการละเมิดกฎหมายและภาวะผูกพัน

P6 ความขัดแย้งด้านผลประโยชน์ การประเมินควรระบุและแสดงถึงความขัดแย้งด้านผลประโยชน์ที่อาจส่งผลกระทบต่อประเมินอย่างเปิดเผยและตรงไปตรงมา

P7 ความรับผิดชอบต่องบประมาณ ผู้ประเมินควรคำนึงถึงทรัพยากรที่ใช้ทั้งหมด ปฏิบัติตามกระบวนการและขั้นตอนทางการเงินที่ดี

4. มาตรฐานด้านความถูกต้อง เป็นมาตรฐานมีวัตถุประสงค์ เพื่อความเชื่อถือและความไว้วางใจของการประเมิน ข้อเสนอแนะข้อค้นพบ โดยเฉพาะอย่างยิ่งเพื่อเป็นการสนับสนุนการตีความและการตัดสินใจเกี่ยวกับคุณภาพ ประกอบด้วยเกณฑ์จำนวน 8 ข้อ ดังนี้

A1 การสรุปและการตัดสินใจอย่างยุติธรรม การสรุปผล สรุปการตัดสินใจควรเป็นไปอย่างเป็นธรรม และเหมาะสมตามบริบทที่การประเมินมีผลสืบเนื่อง

A2 ความถูกต้องของข้อมูล ข้อมูลการประเมินผลควรตอบสนองต่อวัตถุประสงค์ และสนับสนุนการตีความอย่างถูกต้อง

A3 ข้อมูลที่เชื่อถือได้ การให้ข้อมูลที่เชื่อถือได้ และสอดคล้องกับความต้องการใช้งาน

A4 รายละเอียดเชิงบริบทและโครงการชัดเจน การจัดทำเอกสารอธิบายโครงการและบริบทของโครงการอย่างเหมาะสม รายละเอียด และมีขอบเขตของวัตถุประสงค์ในการประเมิน

A5 การจัดการข้อมูล การใช้วิธีการเก็บรวบรวมข้อมูล การทบทวน การตรวจสอบและการจัดเก็บข้อมูลอย่างเป็นระบบ

A6 การออกแบบและการวิเคราะห์ที่น่าเชื่อถือ มีการใช้การประเมิน การออกแบบที่เหมาะสมเฉพาะด้าน และเหมาะสมตามวัตถุประสงค์ของการประเมิน

A7 การให้เหตุผลทางการประเมินมีความชัดเจน การประเมินควรสมเหตุสมผล ควรจัดทำเอกสารแสดงข้อมูลการวิเคราะห์ไปจนถึงข้อค้นพบการตีความ บทสรุปการตัดสินใจอย่างชัดเจนและครบถ้วน

A8 การสื่อสารและการรายงาน การประเมินควรมีขอบเขตที่เพียงพอและป้องกันการตีความผิดพลาด อคติ ความเอนเอียง ความคลาดเคลื่อน และการบิดเบือน

5. **มาตรฐานด้านความรับผิดชอบ** มาตรฐานมีวัตถุประสงค์เพื่อจัดทำเอกสารประเมินอย่างเพียงพอ และมุ่งเน้นการปรับปรุงความรับผิดชอบต่อกระบวนการและผลผลิตของการประเมิน ประกอบด้วยเกณฑ์จำนวน 3 ข้อ ดังนี้

E1 การทำเอกสารที่ใช้ประกอบการประเมิน ผู้ประเมินควรจัดทำเอกสารที่แสดงวัตถุประสงค์ รูปแบบกระบวนการ ข้อมูลและผลลัพธ์อย่างครบถ้วน

E2 การประเมินอภิมานภายใน ผู้ประเมินควรใช้มาตรฐานเหล่านี้และมาตรฐานอื่น ๆ เพื่อตรวจสอบความรับผิดชอบต่อรูปแบบการประเมินกระบวนการที่ใช้ข้อมูลที่เกี่ยวข้อง และผลลัพธ์

E3 การประเมินอภิมานภายนอก ผู้สนับสนุน ผู้รับบริการ ผู้ประเมินและผู้มีส่วนได้ส่วนเสียอื่น ๆ ควรสนับสนุนแนวทางการประเมินภายนอก โดยใช้มาตรฐานเหล่านี้และมาตรฐานอื่น ๆ ที่ปรับใช้ได้

ต่อมาในปี ค.ศ. 2012 Stufflebeam and Social Impact (2012) ได้พัฒนาแบบตรวจสอบรายการประเมินอภิมานฉบับใหม่ ชื่อ Program evaluations meta-evaluation checklist รูปแบบของแบบตรวจสอบรายการมี 5 มาตรฐาน ได้แก่ 1) มาตรฐานด้านอรรถประโยชน์ ประกอบด้วยเกณฑ์จำนวน 8 ข้อ คือ ความน่าเชื่อถือได้ของผู้ประเมิน ความสนใจต่อผู้มีส่วนได้ส่วนเสีย วัตถุประสงค์การเจรจาตกลง คุณค่าที่มีความชัดเจน สารสนเทศที่ตรงประเด็น กระบวนการและผลผลิตที่มีความหมาย การสื่อสารและการรายงาน การประเมินเสร็จทันเวลาและเหมาะสม คำนึงถึงผลที่ตามมาและอิทธิพล 2) มาตรฐานด้านความเป็นไปได้ ประกอบด้วยเกณฑ์จำนวน 4 ข้อ คือ การบริหารโครงการ กระบวนการปฏิบัติจริง ความเป็นไปได้เชิงบริบท และการใช้ทรัพยากร 3) มาตรฐานด้านความเหมาะสม ประกอบด้วยเกณฑ์จำนวน 7 ข้อ คือ มุ่งเน้นการบริการที่ตอบสนองและความครอบคลุม ข้อตกลงอย่างเป็นทางการ สิทธิทางมนุษยชนและความเคารพนับถือ ความชัดเจนและยุติธรรม ความโปร่งใสและการเปิดเผยข้อมูล ความขัดแย้งและความรับผิดชอบด้านการเงิน 4) มาตรฐานด้านความถูกต้อง ประกอบด้วยเกณฑ์จำนวน 8 ข้อ คือ การสรุปและการตัดสินใจอย่างยุติธรรม ความถูกต้องของข้อมูล ข้อมูลที่เชื่อถือได้ รายละเอียดเชิงบริบทและโครงการมีความชัดเจน การจัดการข้อมูล การออกแบบและการวิเคราะห์ที่น่าเชื่อถือ การให้เหตุผลทางการประเมินที่มีความชัดเจน และการสื่อสารและการรายงาน และ 5) มาตรฐานด้านความรับผิดชอบในการประเมิน ประกอบด้วยเกณฑ์จำนวน 3 ข้อ คือ เอกสารการประเมินชัดเจน การประเมินอภิมานภายใน และการประเมินอภิมานภายนอก ตัวอย่างแบบตรวจสอบรายการประเมินอภิมานฉบับใหม่ ชื่อ Program evaluations meta-evaluation checklist นำเสนอดังตารางที่ 2

ตารางที่ 2 ตัวอย่างแบบตรวจสอบรายการประเมินอภิมานฉบับใหม่ของ Stufflebeam and Social Impact (2012)

มาตรฐานด้านอรรถประโยชน์ (Utility Standards) ควรมีการประเมิน ดังนี้
U1 ความน่าเชื่อถือของผู้ประเมิน
<input type="checkbox"/> ผู้ประเมินมีความรู้ความสามารถด้านการประเมิน <input type="checkbox"/> ผู้ประเมินเป็นผู้ที่น่าไว้วางใจ <input type="checkbox"/> ผู้ประเมินที่สามารถแก้ข้อกังวลของผู้มีส่วนได้ส่วนเสีย <input type="checkbox"/> ผู้ประเมินที่ตอบสนองต่อประเด็น เรื่อง ความแตกต่างทางเพศ สถานะทางเศรษฐกิจและสังคม เชื้อชาติ ภาษาและวัฒนธรรม ได้เหมาะสม <input type="checkbox"/> ผู้ประเมินมั่นใจว่าแผนการประเมินจะตอบสนองต่อข้อกังวลหลักของผู้มีส่วนได้ส่วนเสีย <input type="checkbox"/> ผู้ประเมินช่วยให้ผู้มีส่วนได้ส่วนเสียเข้าใจถึงแผนการประเมิน <input type="checkbox"/> ผู้ประเมินให้ข้อมูลผู้มีส่วนได้ส่วนเสียเกี่ยวกับวิธีการและกระบวนการประเมิน <input type="checkbox"/> ผู้ประเมินรับฟังการวิพากษ์วิจารณ์และข้อเสนอแนะของผู้มีส่วนได้ส่วนเสียอย่างเหมาะสม <input type="checkbox"/> ผู้ประเมินเข้าใจแรงกดดันทางสังคมและการเมือง <input type="checkbox"/> ผู้ประเมินคอยแจ้งให้ผู้มีส่วนได้ส่วนเสียทราบถึงความคืบหน้าของการประเมิน
<input type="checkbox"/> 9-10 ดีเยี่ยม <input type="checkbox"/> 7-8 ดีมาก <input type="checkbox"/> 5-6 ดี <input type="checkbox"/> 3-4 พอใช้ <input type="checkbox"/> 0-2 ปรับปรุง

ขั้นตอนการดำเนินงานประเมินอภิमान

การดำเนินงานประเมินอภิमानประกอบด้วยขั้นตอนต่าง ๆ ดังนี้ (Stufflebeam, Madaus & Kellaghan, 2000)

1. การตรวจสอบและการพบปะกับผู้มีส่วนได้ส่วนเสียของการประเมิน เป็นขั้นตอนของการกำหนดผู้ประเมินและผู้มีส่วนได้ส่วนเสีย ว่าเป็นคนกลุ่มใดบ้างที่จะได้ใช้ผลการประเมินอภิमान

2. การกำหนดผู้ประเมินอภิमानหรือทีมประเมินอภิमान ผู้ประเมินอภิमानควรจะสามารถ และมีความน่าเชื่อถือ มีประสบการณ์ด้านการประเมิน ในกรณีที่มีงบประมาณจำนวนจำกัดสามารถดำเนินการประเมินโดยผู้ประเมินอภิमानเพียงคนเดียว หรือถ้าเป็นการประเมินอภิमानภายในองค์กรหรือเป็นการประเมินอภิमानเชิงก้าวหน้า สามารถใช้บุคลากรในหน่วยงานเป็นผู้ประเมินอภิमानได้ ทั้งนี้ขึ้นอยู่กับความต้องการของผู้มีส่วนได้ส่วนเสีย

3. การกำหนดคำถามการประเมินอภิमान เป็นขั้นตอนที่สำคัญ เนื่องจากการกำหนดคำถามสำหรับการประเมินอภิमान ซึ่งผู้ประเมินอภิमानและผู้มีส่วนได้ส่วนเสียควรระบุนโยบายหรือขอบเขตการประเมินให้สอดคล้องกับความต้องการสารสนเทศของการประเมิน โดยพิจารณา 2 แนวทาง คือ 1) กำหนดคำถามการประเมินอภิमान และ 2) กำหนดคำถามการประเมินที่ตรงกับความต้องการของผู้มีส่วนได้ส่วนเสีย อย่างไรก็ตาม โดยทั่วไปผู้ประเมินอภิमानควรมีเป้าหมายการประเมินตามความต้องการของผู้มีส่วนได้ส่วนเสีย

4. การพิจารณาหลักการหรือเกณฑ์ในการประเมินอภิमान เป็นขั้นตอนของมาตรฐานหรือเกณฑ์ที่ใช้ในการประเมิน และเครื่องมือที่ใช้ในการประเมินอภิमानต้องสอดคล้องกับคำถามของการประเมินอภิमानและเป็นไปตามมาตรฐานหรือเกณฑ์ของการประเมินอภิमान ในส่วนของมาตรฐานหรือเกณฑ์การประเมินอภิमान มีหลากหลายมาตรฐานที่นำมาปรับปรุง ซึ่งมาตรฐานที่รู้จัก ได้แก่ มาตรฐานการประเมินบุคลากรของ Joint Committee's ในปี 1988 หรือ 30 มาตรฐานในการประเมินโครงการของ Joint Committee's ในปี 1994 เป็นต้น ซึ่งมาตรฐานดังกล่าวจะเป็นการพิจารณา โครงการในด้านความเหมาะสม ความเป็นประโยชน์ ความถูกต้อง และความเป็นไปได้ หรือนักวิชาการบางท่านได้ศึกษาในแง่ของประสิทธิภาพ

5. การเจรจาต่อรองการทำสัญญาการประเมินอภิमानอย่างเป็นทางการ เป็นขั้นตอนที่นับว่ามีความสำคัญและเป็นการป้องกันความขัดแย้งในอนาคตระหว่างผู้ประเมินอภิमानและผู้ประเมิน และผู้รับบริการควรตกลงกับเอกสารการประเมิน ความรับผิดชอบที่เกี่ยวกับการส่งมอบงาน การแก้ไขผลการประเมินตลอดจนการเผยแพร่ผลการประเมิน ข้อตกลงดังกล่าวช่วยป้องกันไม่ให้เกิดความเข้าใจผิด ระหว่างผู้รับบริการและผู้ประเมินอภิमानที่อาจเกิดขึ้นในระหว่างการดำเนินการประเมินอภิमान

6. การเก็บรวบรวมและทบทวนข้อมูล เมื่อกำหนดข้อตกลงในการประเมิน ผู้ประเมินอภิमानต้องตรวจสอบเป้าหมายการประเมินกับหลักฐานที่มีอยู่ โดยผู้ประเมินต้องพิจารณาข้อมูลที่มีอยู่ จนกระทั่งแน่ใจว่าจากข้อมูลที่มีอยู่นั้นยังไม่พอจะให้ข้อสรุปที่สมเหตุสมผลได้ หรือไม่น่าเชื่อถือ พอที่จะตอบคำถามการประเมินอภิमान จึงจะมีการเก็บรวบรวมข้อมูลเพิ่มเติมได้ แต่ถ้าข้อมูลที่มีอยู่เพียงพอ และเป็นที่ยอมรับสำหรับการประเมินอภิमान การเก็บรวบรวมข้อมูลเพิ่มเติมก็ทำให้สิ้นเปลืองงบประมาณ

7. การเก็บรวบรวมข้อมูลใหม่ตามความจำเป็น จากการดำเนินการเก็บรวบรวมและทบทวนข้อมูล ในกรณีที่ข้อมูลที่มีอยู่นั้นยังไม่พอเพียงที่จะให้ข้อสรุปที่สมเหตุสมผลได้หรือไม่ น่าเชื่อถือพอที่จะตอบคำถามการประเมินอภิमान อาจจะเป็นที่จำเป็นที่จะมีการเก็บรวบรวมข้อมูลเพิ่มเติม จากข้อมูลที่มีอยู่ด้วยข้อมูลใหม่ ซึ่งไม่เพียงแต่ยืนยันข้อสรุปจากข้อมูลที่มีอยู่แต่ยังทำให้ผลการประเมินอภิमानนั้นมีความแข็งแกร่งยิ่งขึ้น

8. การวิเคราะห์ผล ในการวิเคราะห์ข้อมูลของการประเมินอภิमानนั้น เนื่องจากข้อมูลที่ได้จากงานประเมินแต่ละเรื่องหรือการประเมินอภิमानที่มีหลายโครงการ ข้อมูลที่ใช้ในการประเมินอภิमानต่าง ๆ ต้องมีความหลากหลาย

ขั้นตอนการวิเคราะห์จึงประกอบไปด้วยการวิเคราะห์เชิงคุณภาพและเชิงปริมาณ ตามมาตรฐานและเกณฑ์การประเมินอภิमान

9. การตัดสินการประเมินตามมาตรฐานการประเมิน ในขั้นตอนนี้จะเป็นการตัดสินผลการประเมินตามมาตรฐานของการประเมินที่เลือกไว้

10. จัดทำและส่งรายงาน เป็นการรายงานผลการประเมินอภิमानทั้งรายงานในระหว่างประเมินอภิमान และรายงานสรุปรวมของประเมินอภิमान ซึ่งโดยปกติการรายงานระหว่างประเมินอภิमान ใช้รูปแบบที่ไม่เป็นทางการในการรายงาน เช่น ใช้จดหมายทางออนไลน์ การรายงานด้วยวาจา หรือ การรายงานด้วยกระดาษการนำเสนอ ส่วนการรายงานสรุปรวมของการประเมินอภิमानจะต้องเป็นการรายงานที่เป็นทางการ และส่วนใหญ่จะต้องขึ้นอยู่กับข้อตกลงหรือสัญญา

11. การช่วยเหลือผู้รับบริการและผู้มีส่วนได้ส่วนเสียในการตีความ การนำผลการศึกษาไปใช้ประโยชน์เป็นกระบวนการที่ผู้ประเมินอภิमानจะต้องทำหน้าที่ในการแนะนำ ผู้รับบริการและผู้มีส่วนได้ส่วนเสียให้ทำความเข้าใจกับผลการประเมินอภิमान เพื่อให้สามารถทำความเข้าใจและตัดสินคุณค่าของผลการประเมินอภิमानได้อย่างถูกต้อง

ประโยชน์ของการประเมินอภิमान

Stufflebeam (Stufflebeam, 2000 อ้างถึงใน สุวิมล ว่องวานิช, 2558) ให้ข้อคิดเห็นว่า การประเมินอภิमानเป็นสิ่งสำคัญในสังคม มีความสำคัญต่อทั้งผู้ใช้และผู้ผลิตงานประเมิน เช่น พ่อแม่อยากรู้ผลการปฏิบัติงานของลูก นักเรียนอยากรู้ว่าผลสัมฤทธิ์ทางการเรียนของตนเองได้รับการให้เกรดอย่างไรดีธรรมและถูกต้องหรือไม่ กรรมการสถานศึกษาและผู้บริหารโรงเรียนต้องการความมั่นใจในการประเมินโรงเรียนที่ถูกต้องทางเทคนิค ผู้เสียภาษีอยากรู้ข้อมูลเกี่ยวข้องกับผลการปฏิบัติงานของโรงเรียน ไม่ว่าจะเป็กรณีใดก็ตาม การประเมินอภิमानจะให้ข้อมูลสารสนเทศที่เกี่ยวข้องและมีคุณภาพจากการประเมินอภิमानที่มีคุณภาพ

สุวิมล ว่องวานิช (2558) กล่าวถึงประโยชน์ของการประเมินอภิमानว่า การประเมินอภิमानมีความสำคัญต่อนักพัฒนาและผู้ให้บริการ สำหรับนักประเมินมืออาชีพแล้ว การประเมินอภิमानเป็นกลไกการประกันคุณภาพซึ่งใช้ในการตรวจสอบจุดแข็งและแผนการประเมิน ปฏิบัติการประเมิน การจัดทำรายงานการประเมิน การสื่อสารผลการประเมิน มีความสำคัญในสาขาวิชาการประเมินเหมือนการตรวจสอบในสาขาบัญชี สังคมจะอยู่ในสภาวะเสี่ยง หากพึ่งพานักบัญชีที่ให้ข้อมูลทางบัญชีโดยไม่มีบุคคลภายนอกเป็นผู้ตรวจสอบบัญชี ดังนั้น พ่อแม่ นักเรียน นักการศึกษา หรือบุคคลในวงการต่าง ๆ อยู่ในความเสี่ยงหากไม่สามารถไว้วางใจในผลการประเมิน การประเมินอภิमानจึงเป็นสิ่งที่ได้รับความสนใจในวงวิชาชีพและคนทั่วไป เพื่อให้มั่นใจว่าการประเมินนั้น ๆ ได้ให้ข้อสรุปและข้อเสนอแนะที่มีคุณภาพน่าเชื่อถือ

แนวทางในการดำเนินการประเมินอภิमान

แนวทางในการดำเนินการประเมินอภิमानมี 2 แนวทาง ได้แก่ การประเมินอภิमानที่พิจารณาการประเมินโครงการใดโครงการหนึ่ง และการประเมินอภิमानที่เป็นการรวบรวมโครงการหลาย ๆ โครงการ (พรรณนิภา ภูทองพลอย และคณะ, 2561) ผู้เขียนนำเสนอสาระสำคัญ ดังนี้

1. การประเมินอภิमानที่พิจารณาการประเมินโครงการใดโครงการหนึ่ง เพื่อพิจารณาการประเมินโครงการนั้นว่ามีความเหมาะสม ความถูกต้อง ความเป็นไปได้ และอรรถประโยชน์ตามมาตรฐานการประเมินเพียงใด เช่น การกำหนดเป้าหมายของการประเมิน รูปแบบของการประเมิน กระบวนการประเมิน ผลลัพธ์ที่เกิดจากการประเมินและสรุปผลรวมของการประเมิน ตัวอย่างของการดำเนินการประเมินแนวทางนี้ ได้แก่ การประเมินของสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (สมศ.) ซึ่งได้ดำเนินการประเมินอภิमानรายงานการประเมินภายนอกของคณะกรรมการประเมินภายนอกของสถานศึกษา (พรรณนิภา ภูทองพลอยและคณะ, 2561)

2. การประเมินอภิมานที่เป็นการรวบรวมโครงการหลาย ๆ โครงการ โดยโครงการเหล่านี้มีจุดมุ่งหมายของการประเมินหรือลักษณะของการประเมินที่คล้ายคลึงกัน มาพิจารณาโดยอิงตามเกณฑ์มาตรฐานการประเมิน ตัวอย่างการประเมินในรูปแบบนี้ เช่น งานวิจัยของสุวิชญา สิริยานนท์ (2562) ที่ทำการประเมินอภิมานและสังเคราะห์อภิมานรายงานการประเมินโครงการที่เกี่ยวข้องกับการส่งเสริมและพัฒนาคุณภาพชีวิตคนพิการ จำนวน 21 เรื่อง ประกอบด้วยรายงานการประเมินโครงการด้าน 1) การฟื้นฟูสมรรถภาพคนพิการ 2) การจัดสวัสดิการคนพิการ 3) การส่งเสริมและพิทักษ์สิทธิคนพิการทางการมองเห็น การได้ยินหรือสื่อความหมาย เคลื่อนไหวหรือร่างกาย จิตใจหรือพฤติกรรมสติปัญญา การเรียนรู้ และอาชีพ และ 4) การสนับสนุนให้คนพิการสามารถดำรงชีวิตอิสระ การประเมินอภิมานในครั้งนี้เพื่อพิจารณาว่าโครงการต่าง ๆ นั้นว่าเป็นไปตามมาตรฐานด้านความเหมาะสม ด้านความถูกต้อง ด้านความเป็นไปได้ ด้านอรรถประโยชน์ และด้านความรับผิดชอบต่อในระดับใด โดยใช้แบบตรวจสอบรายการประเมินอภิมานของ Stufflebeam and Social Impact (2012) สารสนเทศที่ได้จากการประเมินเป็นประโยชน์ต่อผู้เกี่ยวข้องในการจัดทำนโยบายหรือพิจารณาสนับสนุนโครงการและการประเมินโครงการส่งเสริมและพัฒนาคุณภาพชีวิตคนพิการในด้านต่าง ๆ ต่อไป

ข้อควรระวังในการประเมินอภิมาน

การประเมินอภิมานเป็นแนวคิดที่มีประโยชน์และใช้เครื่องมือที่เหมาะสม ส่วนใหญ่นักประเมินใช้มาตรฐานและเกณฑ์ตามที่นักวิชาการและคณะกรรมการได้กำหนดไว้ อย่างไรก็ตาม เมื่อนำมาตรฐานไปใช้ มีข้อควรระวังว่าการประเมินอภิมานทุกครั้งต้องคำนึงถึงจุดประสงค์ของการประเมิน เช่น สร้างสมรรถนะการเรียนรู้ ค้นหาประเด็นที่นำไปปรับปรุง รวมทั้งต้องคำนึงถึงข้อจำกัดขณะที่ประเมินอภิมาน เช่น งบประมาณ ระยะเวลา การเข้าถึงข้อมูล เช่น การประเมินที่มีงบประมาณและมีระยะเวลาจำกัด อาจจำเป็นต้องกำหนดขอบเขต แต่ต้องการได้ข้อสรุปประเด็นที่เด่นชัดที่สุดตามทรัพยากรที่มี ทำนองเดียวกัน การตัดสินใจออกแบบประเมินหรือทำการประเมินอาจต้องเผชิญกับปัญหา เช่น หากให้ผู้มีส่วนได้ส่วนเสียจำนวนมากเข้าร่วมการประเมินจะทำให้การประเมินมีประโยชน์สูงสุด แต่ค่าใช้จ่ายมาก ในทางกลับกัน หากให้ทีมประเมินผลมีส่วนร่วมในการออกแบบหรือร่วมประเมิน และสมาชิกในทีมเหล่านั้น ผ่านการฝึกอบรมมาน้อย ทำให้ผู้ใช้ผลการประเมินอาจเกิดความไม่แน่ใจต่อผลการประเมิน แต่จุดด้อยนี้อาจไม่มีความหมายถ้าคำนึงถึงประโยชน์ที่ได้รับจุดด้อยนี้เป็นปัญหาต่อเมื่อการประเมินผลต่ำกว่ามาตรฐาน เช่น ต่ำกว่าระดับที่ยอมรับได้ตามเกณฑ์ใดเกณฑ์หนึ่งหรือหลายเกณฑ์ (เรื่องวิทย์ เกษสุวรรณ, 2555)

สรุป

การประเมินอภิมาน และการประยุกต์ใช้แนวทางการประเมินอภิมาน มีประโยชน์ต่อการตัดสินใจคุณภาพของงานประเมินหรือนักประเมิน การประเมินอภิมานเป็นวิธีการประเมินคุณภาพของการประเมินที่แพร่หลายในต่างประเทศ การประยุกต์ใช้เพื่อตรวจสอบคุณภาพของการประเมินจะเป็นไปตามมาตรฐานของการประเมินอภิมาน นักประเมินต้องมีความรู้ ความเข้าใจเกี่ยวกับการพัฒนาวิธีการประเมินคุณภาพของการประเมิน เพื่อติดตาม และตรวจสอบการประเมิน นักประเมินสามารถเลือกเกณฑ์มาตรฐานการประเมินตามความเหมาะสมในการประเมินอภิมาน สร้างเครื่องมือแบบตรวจสอบรายการการประเมินอภิมาน และทำรายงานผล เพื่อให้ทราบถึงจุดเด่นและจุดด้อยของการประเมิน ผู้เกี่ยวข้องสามารถนำผลการประเมินอภิมานไปใช้ประโยชน์ในการพัฒนาหรือปรับปรุงจุดด้อยในด้านต่าง ๆ ให้มีคุณภาพยิ่งขึ้นต่อไป

เอกสารอ้างอิง

- เรืองวิทย์ เกษสุวรรณ. (2555). การประเมินผลนโยบายสาธารณะ. กรุงเทพฯ: บริษัท บพิตรการพิมพ์ จำกัด.
- พรรณนิภา ภูทองพลอย, ทศน์ศิริรินทร์ สว่างบุญ, และ วราพร เอรารวรรณ. (2561). มโนทัศน์เบื้องต้นของการประเมิน
อภิमान. *วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม*. 24 (1). 3-13.
- พินดา วราสุนันท์. (2556). ย้อนรอยการประเมินสู่มุมมองการประเมินในอนาคต. *วารสารการวัดผลการศึกษา*. 30 (88).
12-27.
- ศิริชัย กาญจนวาสิ. (2552). ทฤษฎีการประเมิน (พิมพ์ครั้งที่ 7). กรุงเทพฯ: สำนักพิมพ์แห่งกรุงเทพฯ.
- สมหวัง พิธิยานุวัฒน์. (2559). วิธีวิทยาการประเมินทางการศึกษาแห่งคุณค่า (พิมพ์ครั้งที่ 6). กรุงเทพฯ: สำนักพิมพ์แห่ง
จุฬาลงกรณ์มหาวิทยาลัย.
- สุวิษญา สิริยานนท์. (2562). การประเมินอภิमानและการสังเคราะห์อภิमानรายงานการประเมินโครงการการส่งเสริมและ
พัฒนาคุณภาพชีวิตคนพิการ. วิทยานิพนธ์วิทยาศาสตร์มหาบัณฑิต. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- สุวิมล ว่องวานิช. (2558). การวิจัยประเมินความต้องการจำเป็น (พิมพ์ครั้งที่ 3). กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์
มหาวิทยาลัย.
- Cook, T. D., & Gruder, C. L. (1978). Meta-evaluation research in *Evaluation Quarterly*. 2 (1978). 5-51.
- Gallegos, A. (1994). Meta-evaluation of school evaluation model. *Studies in Educational Evaluation*.
20. 41-54.
- Patton, M. Q. (1997). *Utilization-Focused Evaluation*. The New Century Text. 3rd. ed. Thousand Oaks, Sage.
- Scriven, M. (1969). *An Introduction to Meta-Evaluation*. Educ Prod Rep
- Scriven, M. (1991). *Evaluation Thesaurus*. 4th. London: Sage Publications.
- Scriven, M. (2005). The logic and Methodology of Checklist. Retrieved July 24, 2020 from
www.wmich.edu/evalctr/checklists/
- Scriven, M. (2007). Key evaluation checklists. Retrieved July 24, 2020 from
www.wmich.edu/evalctr/checklists/
- Scriven, M. (2015). KEY EVALUATION CHECKLIST (KEC). Retrieved July 24, 2020 from
<http://michaelscriven.info/home.html>
- Stufflebeam, D. L. (2001). The Metaevaluation Imperative. *American Journal of Evaluation*. 22 (2).
183-209. Retrieved July 24, 2020 from <http://aje.sagepub.com/cgi/content/abstract/22/2/183>
- Stufflebeam, D. L., Madaus, G. F. & Kellaghan, T. (2000). *Evaluation models*. Boston: Kluwer Academic.
- Stufflebeam, D. L., & Social Impact. (2012). Program evaluations metaevaluation checklist (Based on
The Program Evaluation Standards). Retrieved July 24, 2020 from
http://pdf.usaid.gov/pdf_docs/pnady797.pdf
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation
standards: A guide for evaluators and evaluation users* (3rd ed). Thousand Oaks, CA: Sage.