

## การประเมินอภิมาน : แนวคิดและหลักการ

### Metaevaluation: Concepts and Principles

เนติ เฉลยวาเรศ\*

#### บทนำ

การประเมินอภิมาน (Metaevaluation) เป็นวิธีการประเมินงานประเมินที่แพร่หลายในต่างประเทศมาเป็นเวลานาน สำหรับในประเทศไทยนั้นยังเป็นที่รู้จักไม่มากนัก จะรู้จักกันในวงจำกัดเพียงเฉพาะกลุ่มอาจารย์และนักวิชาการด้านการวัดและประเมินผลทางการศึกษาเท่านั้น และเริ่มนำมาประยุกต์ใช้เพื่อตรวจสอบคุณภาพของงานประเมินทางการศึกษา บทความนี้จึงมีวัตถุประสงค์เพื่อเสนอแนวคิดและหลักการการประเมินอภิมานพอสังเขป โดยจะทำการเสนอในประเด็นดังต่อไปนี้ ความหมายของการประเมินอภิมาน บทบาทของการประเมินอภิมาน แนวคิดเกี่ยวกับการประเมินอภิมาน มาตรฐานการประเมิน ตลอดจนตัวอย่างการประเมินอภิมานของต่างประเทศ

#### ความหมายของการประเมินอภิมาน

นักประเมินทางการศึกษาให้นิยามของการประเมินอภิมานไว้เหมือนและแตกต่างกันไป ดังนियามต่อไปนี้

Cook & Gruder (1978) ได้ให้ความหมายของการประเมินอภิมานว่าหมายถึงศาสตร์และศิลป์ของการประเมินงานประเมินทำให้ได้ข้อค้นพบที่เป็นประโยชน์ โดยมีจุดมุ่งหมายเพื่อทำให้การประเมินมีคุณภาพ

Schwand & Hapern (1988) ได้ให้ความหมายว่าเป็นการประเมินงานประเมินโดยถือว่าการประเมินที่ดีต้องเป็นการประเมินที่มีคุณภาพปราศจากความลำเอียง มีประโยชน์ และทันต่อการนำผลการประเมินไปใช้

Scriven (1991) ได้ให้ความหมายว่าการประเมินอภิมานหมายถึงการประเมินงานประเมินหรือเป็นงานประเมินของผู้ประเมิน โดยใช้วิธีการทางวิทยาศาสตร์และศิลปกรรมมาช่วย

Stufflebeam (1978 cited in Wabha, 1990) ได้ให้ความหมายว่าหมายถึงกระบวนการจัดการสารสนเทศเพื่อพิจารณาการประเมินในด้านอรรถประโยชน์การดำเนินงาน จรรยาบรรณทางวิชาชีพ และความเพียงพอ ทางด้านการประเมิน เพื่อให้ได้สารสนเทศในการปรับปรุงการประเมินและรายงานผลการประเมิน จุดแข็งจุดอ่อนต่อสาธารณชน

\* อาจารย์ ประจำคณะครุศาสตร์ มหาวิทยาลัยราชภัฏเทพสตรี

สมหวัง พิธิยานุวัฒน์ (2541) ได้ให้ความหมายการประเมินอภิมานว่าหมายถึงการตัดสินใจเกี่ยวกับคุณค่าภายนอกและภายในของการประเมิน

อุทุมพร จามรมาน (2540) ได้ให้ความหมายประเมินอภิมานว่าหมายถึงการประเมินหรือตรวจสอบประเมินซ้ำ เพื่อให้ได้คำตอบของการประเมินคุณค่าเหล่านั้น นอกจากนี้ สุวิมล ว่องวานิช (2538) ได้ให้ความหมายเพิ่มเติมว่า การประเมินอภิมานไม่น่าจะมาจากความหมายเพียงการประเมินวิธีการหรือการประเมินคุณภาพการประเมินอย่างเดียว แต่สามารถทำได้โดยวิธีสังเคราะห์วิธีการที่ให้ประเมิน (synthesis of evaluation process) แล้วทำการประเมินคุณภาพที่ได้จากการสังเคราะห์เหล่านั้น

จากนิยาม และความหมายของการประเมินอภิมานดังที่กล่าวมาพอจะสรุปได้ว่าการประเมินอภิมานเป็นการประเมินงานประเมินหรือคุณภาพของงานประเมินรวมไปถึงงานประเมินการสังเคราะห์วิธีการที่ใช้ในการประเมินเพื่อให้ได้สารสนเทศที่เป็นประโยชน์ในการปรับปรุงงาน และ/หรือรายงานผลการประเมินในด้านจุดแข็ง จุดอ่อนต่อสาธารณะ ซึ่งมีวิธีการในการประเมินอภิมานแตกต่างกัน

### บทบาทของการประเมินอภิมาน

บทบาทของการประเมินอภิมาน แบ่งออกเป็น 2 ลักษณะ คือการประเมินอภิมานเชิงความก้าวหน้า (formative metaevaluation) และการประเมินอภิมานแบบสรุปรวม (summative metaevaluation) ซึ่งในการประเมินอภิมานทั้งเชิงความก้าวหน้าและเชิงสรุปรวมมีบทบาทที่แตกต่างกันไป (Stufflebeam cited in Berk. 1981; Worthen & Sander. 1987; สมหวัง พิธิยานุวัฒน์. 2541) สรุปได้ดังนี้

1. การประเมินเชิงความก้าวหน้า (formative metaevaluation) เป็นการประเมินงานประเมินเชิงสร้างสรรค์ กล่าวคือ เป็นการวางแผนก่อนปฏิบัติการประเมิน การประเมินอภิมานเชิงความก้าวหน้าช่วยให้นักประเมินตัดสินใจว่าจะประเมินอะไร นิยามปัญหาที่ประเมินกำหนดผู้เกี่ยวข้องในการประเมิน กำหนดวัตถุประสงค์ของการประเมินให้กระจ่าง กำหนดข้อสัญญา กำหนดสารสนเทศที่ต้องการ กำหนดวิธีการเก็บรวบรวมข้อมูลและวิเคราะห์ข้อมูล กำหนดตารางประเมิน จัดหาทีมประเมิน กำหนดงบประมาณ เตรียมข้อสรุปและข้อเสนอแนะ และเลือกรูปแบบและวิธีการรายงานผลการประเมิน โดยการประเมินอภิมานในเชิงความก้าวหน้าแสดงได้สมการต่อไปนี้

$$\begin{array}{l} \text{แผนการประเมินและการ} \\ \text{ดำเนินการประเมิน} \end{array} = \begin{array}{l} \text{การประเมินตามแนวทาง} \\ \text{มาตรฐานที่กำหนด} \end{array} = \begin{array}{l} \text{การหลีกเลี่ยง} \\ \text{ข้อผิดพลาดหรือ} \\ \text{หลุมพรางของการประเมิน} \end{array}$$

โดยสรุปการประเมินอภิมานเชิงความก้าวหน้าเป็นการประเมินว่าการวางแผนการประเมินเป็นไปตามมาตรฐานที่กำหนดหรือไม่ ถึงปฏิบัติตามแนวทางผลการประเมินถูกต้องและสมเหตุสมผล

และเป็นหลักเลี้ยงข้อผิดพลาด โดยทั่วไปของการประมาณด้วย คุณค่าการประเมิน จะมีค่าสุดเมื่อ  
ดำเนินการประเมินและหลักเลี้ยงข้อผิดพลาดในการประเมิน

2. การประเมินอภิมานเชิงสรุปรวม (summative metaevaluation) ถือว่าเป็นบทบาทพื้นฐาน  
ของการประเมินอภิมาน คือ เป็นการแสดงถึงคุณค่ารวมของงานประมาณซึ่งแสดงถึงความพร้อมที่จะ  
ให้ตรวจสอบได้ของนักเรียนโดยรายงานให้สาธารณชนทราบถึงความเป็นมาตรฐานของการประเมิน  
ผลการประเมินอภิมานเชิงสรุปรวมจะเป็นมาตรฐานของการประเมิน ผลการประเมินอภิมานเชิงสรุป  
รวมจะเป็นสารสนเทศให้ผู้ใช้งานผลการประเมินตัดสินใจว่าควรยอมรับ ผลการประเมินและข้อเสนอแนะ  
จากงานประเมินมากน้อยเพียงใด โดยการประเมินเชิงสรุปรวม แสดงดังสมการต่อไปนี้

$$\frac{\text{การประเมินภาคปฏิบัติจริงหรือส่วนประกอบ}}{\text{มาตรฐานของการประเมิน}} \leq 1.00$$

คุณภาพของงานประเมินสามารถแสดงในรูปของการให้คะแนนถ่วงน้ำหนักการประเมิน  
ภาคปฏิบัติจริงตามมาตรฐานของงานประเมิน ถ้าค่าอัตราส่วนมีค่าเข้าใกล้ 1.00 แสดงถึงความมี  
ค่าสูงสุดของงานประเมินในการประเมินอภิมานเชิงความก้าวหน้าและเชิงสรุปรวม (formative and  
summative meatevaluation) เป็นการประเมินอภิมานเชิงความก้าวหน้าและเชิงสรุปรวมร่วมกัน ดัง  
แสดงความสัมพันธ์ได้ดังสมการต่อไปนี้

$$\left[ \frac{\text{การประเมินภาคปฏิบัติจริงหรือส่วนประกอบ}}{\text{มาตรฐานของการประเมิน}} \right] \quad 1.00 \text{ เมื่อการ} \quad \left[ \text{การประเมินการหลักเลี้ยง} \right. \\ \left. \text{ดำเนินงานประเมิน} \quad \left. \begin{array}{l} \text{แนวที่กำหนด+ข้อผิดพลาด} \\ \text{ในการประเมิน} \end{array} \right]$$

การประเมินอภิมานเชิงสรุปรวมจะมีค่ามากที่สุดเมื่อผู้ประเมินได้ให้ความสนใจในข้อติชม  
ปรับปรุงแก้ไขในสิ่งที่บกพร่องจากการประเมินอภิมานในเชิงก้าวหน้า

จากการประเมินอภิมานเชิงความก้าวหน้าและการประเมินอภิมานเชิงสรุปจะเห็นได้ว่าการ  
ประเมินอภิมานเชิงความก้าวหน้ามีความสำคัญต่อการประเมินอภิมานเชิงสรุปรวมเนื่องจากการ  
ประเมินอภิมานเชิงความก้าวหน้า นี้ จะได้สารสนเทศในการแก้ไขปรับปรุงการประเมิน เพื่อหลักเลี้ยง  
ข้อผิดพลาดต่าง ๆ ในการประเมินและเมื่อประเมินอภิมานเชิงสรุปรวม จะพบจุดด้อยของการประเมิน  
น้อยลง ทำให้การประเมินมีคุณภาพมากขึ้นนั่นเอง

1. ผู้ประเมินอภิमानเชิงความก้าวหน้า (formative evaluator) เป็นผู้ประเมินงานประเมินก่อนดำเนินการ โครงการ โดยการจัดหาสารสนเทศเพื่อการปรับปรุง กำกับ ติดตาม คุณภาพของแผนงาน ประเมินวิธีดำเนินการประเมิน ทำให้การประเมินมีคุณภาพมากขึ้น

2. ผู้ประเมินอภิमानเชิงสรุปรวม (summative metaevaluator) เป็นการประเมินอภิमानเมื่อการประเมินเสร็จสิ้นแล้ว เพื่อพิจารณาคุณภาพของการประเมินโดยทั่วไป ได้แก่ คุณค่าภายใน ความยุติธรรม และความเชื่อถือได้ของการประเมิน

### แนวคิดเกี่ยวกับการประเมินอภิमान

การประเมินอภิमानได้รับการส่งเสริมอย่างแพร่หลายจากหน่วยงานทางการประเมิน โดยนักประเมินกล่าวถึงมโนทัศน์นี้เป็นครั้งแรก คือ Scriven (1969 cited in Wahba. 1990) กล่าวว่า การประเมินอภิमानเป็นการประเมินซ้อนการประเมิน เกี่ยวข้องกับการประเมิน โดยผู้ประเมินอภิमानต้องเข้าไปมีส่วนต่อการปฏิบัติการประเมินต่อมานักประเมินทางการศึกษาได้ให้ความหมายของการประเมินอภิमानว่าเป็นการตรวจสอบทางการศึกษามีวิธีการ คือ 1) กำกับติดตามการประเมินเพื่อให้ได้คุณภาพของการประเมินที่เชื่อถือได้ 2) ประเมินงานประเมินอย่างถูกต้องและเสนอผลการประเมินต่อสาธารณะซึ่ง Stufflebeam (1974 cited in wahba. 1990) กล่าวว่า การประเมินอภิमानสมารถที่จะช่วยให้นักประเมินทราบถึงปัญหาต่าง ๆ ของการประเมินและช่วยให้สาธารณชนรับทราบผลการประเมินตามความเป็นจริง นอกจากนี้ Stufflebeam ได้สรุปมโนทัศน์ของการประเมิน (evaluation) และมโนทัศน์ของการประเมินอภิमान (metaevaluation) ไว้ 8 ประเด็นดังนี้

1. การประเมินเป็นการประเมินคุณภาพภายใน (merit) โครงการส่วนการประเมินอภิमानเป็นการตัดสินคุณภาพภายในของงานประเมิน

2. การประเมินเป็นการสนองต่อการตัดสินใจ และแสดงถึงความรับผิดชอบในการปฏิบัติ (accountability) ส่วนการประเมินอภิमानเชิงความก้าวหน้า (formative metaevaluation) เพื่อทำการตัดสินใจในการกำกับติดตามการประเมิน และการประเมินอภิमानเชิงสรุปรวม (summative metaevaluator) จะให้สารสนเทศเพื่อการแสดงถึงการรับผิดชอบในการปฏิบัติต่อการประเมินที่ผ่านมา

3. การประเมินควรประเมินเป้าหมาย การวางแผน การกำกับติดตาม และผลลัพธ์ของโครงการส่วนการประเมินอภิमानนั้น ควรประเมินความสำคัญของจุดประสงค์ในการประเมินความพอเพียงในการวางแผนการประเมิน ความพอเพียงในการกำกับ ติดตามแผนการประเมิน ความพอเพียงด้านเทคนิคที่ใช้ในการวิเคราะห์ข้อมูล คุณภาพ และความสำคัญของผลการประเมิน

4. การประเมินจะให้สารสนเทศเชิงบรรยาย เชิงตัดสิน และข้อเสนอแนะ สำหรับการประเมิน อภิमानควรจะบรรยายตัดสินใจการประเมิน และควรให้ข้อเสนอแนะเพื่อการปรับปรุงการตัดสินใจ การประเมิน และการนำผลประเมินงานประเมินไปใช้ประโยชน์

5. การประเมินจะสนองต่อบุคคลที่เกี่ยวข้องกับการประเมินทุกฝ่าย ส่วนการประเมินอภิमान ควรสนองต่อผู้ประเมิน และบุคคลที่สนใจในการประเมิน

6. ในการประเมินนั้น ผู้ประเมินภายในนั้นจะเป็นผู้ประเมินเชิงความก้าวหน้า เพื่อทำการ ตัดสินใจเกี่ยวกับโครงการ ส่วนผู้ประเมินภายนอกจะประเมินเชิงสรุปรวมเพื่อแสดงความรับผิดชอบ ต่อการปฏิบัติ ด้วยเหตุนี้ นักประเมินอภิमानควรประเมินอภิमानเชิงความก้าวหน้าเพื่อทำการตัดสิน เกี่ยวกับการประเมิน และนักประเมินอภิमानควรประเมินอภิमानเชิงสรุปรวมเมื่อกิจกรรมหรือการ ประเมินสิ้นสุดเพื่อแสดงความรับผิดชอบในการปฏิบัติต่อการประเมิน

7. การประเมินเป็นกระบวนการระบุข้อมูลที่ต้องการ และจัดเก็บวิเคราะห์ข้อมูลเพื่อให้ สารสนเทศในการตัดสินใจ และใช้สารสนเทศในการแสดงความรับผิดชอบต่อโครงการ ส่วนการ ประเมินอภิमानจะต้องระบุคำถามที่ชัดเจน การเก็บรวบรวมข้อมูล การจัดระบบ และการวิเคราะห์ สารสนเทศที่ได้รับเพื่อประยุกต์ใช้สารสนเทศในการตัดสินใจและแสดงความรับผิดชอบต่อ การประเมินอย่างเหมาะสม

8. การประเมินจะต้องมีเทคนิค มีประโยชน์ และมีประสิทธิภาพ ส่วนการประเมินอภิमानต้อง มีเกณฑ์การประเมินที่เพียงพอ

การประเมินอภิमानที่มีบทบาทสำคัญในการช่วยกำกับติดตามแผนการประเมิน วิธีดำเนินการ ประเมิน รวมทั้งประเมินคุณภาพของการประเมินโดยรวมด้วย ซึ่ง ศิริชัย กาญจนวาสิ (2537) ได้ จัดรูปแบบการประเมินอภิमानได้ดังต่อไปนี้

1. การติดตามตรวจสอบการประเมิน
2. การวิพากษ์วิจารณ์
  - 2.1 การวิพากษ์วิจารณ์อย่างเป็นทางการ เช่น การวิจารณ์ทางปรัชญา ศิลปะ บันเทิง เป็นต้น
  - 2.2 การอภิปรายโดยผู้เชี่ยวชาญ
3. การเปรียบเทียบกับเกณฑ์มาตรฐาน
4. การประเมินซ้ำ
  - 4.1 การวิเคราะห์และตีความหมายใหม่
  - 4.2 การศึกษาเปรียบเทียบกับผู้การประเมินอื่น

ส่วน Smith (1990) ได้กล่าวว่า การศึกษารายกรณี เป็นวิธีการหนึ่งที่น่ามาใช้ในการประเมินอกิ  
 มาณเพื่อให้ได้ข้อมูลที่ลึกซึ้ง ทำความเข้าใจได้มากกว่าที่จะพิจารณาตัดสินถึงคุณค่าภายนอกเพียงอย่าง  
 เดียวโดยวิธีการศึกษารายกรณีนี้ จะได้สารสนเทศในการปรับปรุงการประเมิน จุดดี จุดด้อย ของการ  
 ประเมินนั้น

นอกจากการทำกรประเมินอกิมาณ โดยใช้การศึกษารายกรณีแล้ว Brinkeroff et. al (1983  
 cited in Worthen & Sander. 1987) ได้เสนอวิธีการดำเนินการประเมินอกิมาณ ดังตารางที่ 1

ตาราง 1 วิธีดำเนินการในการประเมินอกิมาณ Brinkeroff et al (1983) อ้างถึงใน Worthen & Sander  
 (1987)

	ประเด็นของการประเมินอกิมาณ		
	ประเมินแผนการ ประเมิน	ประเมินงานประเมิน เชิงความก้าวหน้า	ประเมินงานประเมิน หลังเสร็จสิ้น
วิธีดำเนินการใน การประเมินอกิมาณ	จ้างที่ปรึกษาการ ประเมิน เช่น ผู้ประเมิน ผู้เชี่ยวชาญทางการวัด หรือเชี่ยวชาญเฉพาะ วิชา ผู้ทบทวน แผน เช่น กลุ่มที่ให้ คำปรึกษา ทบทวนแผน ในการประเมิน เช่น แผนงานหรือแผน ดำเนินการ	ใช้ผู้สังเกตจากผู้ที่เป็ นอิสระจากกัน เช่น ผู้ ประเมินอกิมาณ ทีมผู้ประเมินอกิมาณ ทีมผู้ประเมินคณะที่ ปรึกษาในการประเมิน อ่านทบทวนการประเมิน อกิมาณเชิงความก้าวหน้า เช่น แผนการประเมิน ระยะต่าง ๆ งบประมาณ ปัจจุบัน รายละเอียด เกี่ยวกับการเก็บรวบรวม ข้อมูล	ส่งรายงานการประเมิน ต่อผู้ประเมินที่ปรึกษา ผู้เชี่ยวชาญทางการ ประเมิน โดยให้อ่าน ทบทวนรายงานการ ประเมินเพื่อประเมิน คุณภาพ

หลักการพิจารณาคุณภาพของการประเมินนั้น สามารถตัดสินได้จากคุณค่าของการประเมินใน  
 3 ด้านคือ ค่าถามการประเมิน วิธีการประเมิน และผลการประเมิน (ศิริชัย กาญจนวาสี. 2537) โดยด้าน  
 ค่าถามการประเมินนั้นจะต้องเป็นไปตามความต้องการจำเป็นและสนองต่อความต้องการในการใช้

สารสนเทศ ส่วนด้านวิธีการประเมินจะต้องสนองต่อเป้าหมายและเข้ากับสถานการณ์ในการประเมิน และด้านผลการประเมินจะต้องเป็นประโยชน์ ยุติธรรม และน่าเชื่อถือ

### มาตรฐานของการประเมิน

มาตรฐานของการประเมินตามที่คณะกรรมการพัฒนาเกณฑ์มาตรฐานสำหรับการประเมินทางการศึกษาได้กำหนดไว้ (Finn, 1997; Mandaus Scriven & Stufflebeam, 1989; Stufflebeam & Shinkfiels, 1990; สิริชัย กาญจนวาสี, 2357; สมหวัง พิธิยานุวัฒน์, 2541) ประกอบด้วยมาตรฐานการประเมิน 4 ด้าน คือมาตรฐานด้านอรรถประโยชน์ (Utility standards) ประกอบด้วย 8 เกณฑ์ มาตรฐานด้านความเป็นไปได้ (feasibility standards) ประกอบด้วย 3 เกณฑ์ มาตรฐานด้านความเหมาะสม (propriety standards) ประกอบด้วย 8 เกณฑ์ มาตรฐานด้านความถูกต้อง (accuracy standards) ประกอบด้วย 11 เกณฑ์ รวมเกณฑ์ทั้งสิ้น 30 เกณฑ์ ดังรายละเอียดต่อไปนี้

1. มาตรฐานอรรถประโยชน์ (utility standards) ซึ่งเป็นการประเมินด้านกระบวนการที่ก่อให้เกิดสารสนเทศตรงกับความต้องการของผู้ใช้ผลการประเมินเป็นสำคัญ ประกอบด้วย 8 เกณฑ์ คือ

- 1.1 มีการระบุผู้เกี่ยวข้องที่ต้องการใช้สารสนเทศ
- 1.2 ความเป็นที่เชื่อถือของผู้ประเมิน ผู้ประเมินต้องรู้สิ่งที่มุ่งประเมิน
- 1.3 มีการรวบรวมข้อมูลครบคลุมและตอบสนองความต้องการของผู้ใช้สารสนเทศของผู้เกี่ยวข้อง
- 1.4 การแปลความหมายและการตัดสินใจมีความชัดเจน
- 1.5 รายงานการประเมินมีความชัดเจนทุกขั้นตอน
- 1.6 การเผยแพร่ผลการประเมินไปยังผู้ที่เกี่ยวข้องอย่างทั่วถึง
- 1.7 รายงานการประเมินเสร็จทันสำหรับนำไปใช้ประโยชน์
- 1.8 การประเมินส่งผลกระทบต่อในการกระตุ้นให้มีการดำเนินการประเมินต่อไปอย่างต่อเนื่อง

2. มาตรฐานด้านความเป็นไปได้ (feasibility standards) เป็นมาตรฐานเน้นความเป็นจริงของการประเมินประกอบด้วย 3 เกณฑ์ คือ

- 2.1 วิธีการประเมินสามารถนำไปใช้ปฏิบัติได้จริง
- 2.2 การเป็นที่ยอมรับของผู้ถูกประเมินผู้ได้รับผลกระทบจากการประเมิน ตลอดจนผู้ที่เกี่ยวข้องทุกฝ่าย
- 2.3 ผลที่ได้มีความคุ้มค่ากับผลงาน สามารถนำไปปรับปรุงการปฏิบัติงาน ขยายงานหรือล้มเลิก

3. มาตรฐานด้านความเหมาะสม (propriety standards) มุ่งเน้นเกี่ยวกับเรื่อง กฎ ระเบียบ ศีลธรรม จรรยาบรรณที่เกี่ยวข้องกับการประเมิน ประกอบด้วย 8 เกณฑ์ คือ

- 3.1 มีการกำหนดข้อตกลงอย่างเป็นทางการ
- 3.2 มีการแก้ปัญหาของการประเมินด้วยความซื่อสัตย์ ไม่บิดเบือนความเป็นจริงอย่างสันติวิธี
- 3.3 รายงานการประเมินอย่างตรงไปตรงมา เปิดเผย และคำนึงถึงข้อจำกัดของการประเมิน
- 3.4 มีการให้ความสำคัญต่อสิทธิในการรับรู้ข่าวสารของบุคคลทั่วไปอย่างเปิดเผย มีการเผยแพร่ผลการประเมินและเปิดโอกาสให้ผู้สนใจได้ศึกษาผลการประเมิน
- 3.5 มีการคำนึงถึงสิทธิส่วนตัวของกลุ่มตัวอย่าง
- 3.6 มีการเคารพในการปฏิสัมพันธ์ของผู้เกี่ยวข้อง รวบรวมข้อมูลจากบุคคลหลายฝ่ายที่เกี่ยวข้อง
- 3.7 มีรายงานผลการประเมินที่สมบูรณ์ ยุติธรรม และเสนอทั้งจุดเด่น จุดด้อยของสิ่งที่ประเมิน
- 3.8 ผู้ประเมินทำการประเมินด้วยความรับผิดชอบและมีจรรยาบรรณ

4. มาตรฐานด้านความถูกต้อง (accuracy standards) เกี่ยวข้องกับการใช้เครื่องมือที่มีคุณภาพ ตลอดจนวิธีเก็บข้อมูลตามหลักวิชาการ ซึ่งประกอบด้วย 11 เกณฑ์ คือ

- 4.1 มีการระบุวัตถุประสงค์ของการประเมินอย่างชัดเจน
- 4.2 มีการวิเคราะห์บริบทของการประเมิน ในรายงานการประเมินอย่างครอบคลุมและเพียงพอต่อการประเมินอย่างมีประสิทธิภาพ
- 4.3 มีการบรรยายจุดประสงค์และกระบวนการประเมินอย่างชัดเจน
- 4.4 มีการบรรยายแหล่งข้อมูลและการได้มาอย่างชัดเจน
- 4.5 ใช้เทคนิคและเครื่องมือที่มีความตรง และได้มาซึ่งสารสนเทศที่มีความถูกต้องแม่นยำ
- 4.6 ใช้เทคนิคและเครื่องมือที่มีความเที่ยง และได้มาซึ่งสารสนเทศที่มีความคงเส้นคงวา
- 4.7 มีการจัดระบบควบคุมสำหรับการเก็บรวบรวมข้อมูล วิเคราะห์และรายงาน
- 4.8 มีการวิเคราะห์ข้อมูลเชิงปริมาณ
- 4.9 มีการวิเคราะห์ข้อมูลเชิงคุณภาพ
- 4.10 การลงสรุปที่มีเหตุผลและผลการประเมินสนับสนุน
- 4.11 มีการเขียนรายงานที่มีความเป็นปรนัย



นอกจากนี้ Nyirenda (1994) ได้ร่วมกับมาตรฐานวิชาชีพการสอนแห่งชาติ (National Board for Professional Teaching Standards) หรือ NBPTS ได้ศึกษาว่า ทำอย่างไรการสอนจะบรรลุผลได้ จึงได้พัฒนาเกณฑ์การประเมินอภิมานเพื่อการประเมินระบบออกไปประกอบวิชาชีพครูขึ้น จึงได้พัฒนาเกณฑ์การประเมินดังนี้

1. ความเป็นไปได้ทางการบริหาร (administrative feasibility)
2. ความเชื่อถือได้ต่อสาธารณชน (public credibility)
3. ความเชื่อถือได้ถ้าวิชาชีพ (professional credibility)
4. การป้องกันทางด้านกฎหมาย (legal feasibility)
5. ความสามารถทางการเงิน (economic affordability)

### ตัวอย่างการประเมินอภิมานของต่างประเทศ

Wahba (1990) ได้ทำวิทยานิพนธ์ระดับปริญญาเอกของมหาวิทยาลัยพิตทเบอร์ก (University of Pittsburgh) เรื่องการประเมินอภิมานเชิงสรุปรวม (summative metaevaluation) เกี่ยวกับการอบรมเลี้ยงดูเด็กแรกเกิดของประเทศกำลังพัฒนา 4 ประเทศ โดยได้ประยุกต์ใช้เกณฑ์มาตรฐานของคณะกรรมการประเมินทางการศึกษา (the Joint Committee on Standards to Educational Evaluation) มีจุดประสงค์เพื่อประเมินคุณภาพของรายงานการประเมินเกี่ยวกับการอบรมเลี้ยงดูเด็กแรกเกิดของประเทศอินเดีย พม่า กาน่า และ Sierre Lone โดยใช้มาตรฐานของคณะกรรมการพัฒนาเกณฑ์การประเมินทั้ง 30 ข้อ ดำเนินการ โดย 1) เลือกรายงานการประเมินเกี่ยวกับการอบรมเลี้ยงดูเด็กแรกเกิดที่สมบูรณ์ 4 ฉบับ จากทั้งหมด 13 เรื่อง 2) นำรายงานการประเมินทั้ง 4 ฉบับให้ผู้เชี่ยวชาญ 4 คน (ประกอบด้วย ผู้สำเร็จปริญญาเอก 3 คน ปริญญาโท 1 คน มีประสบการณ์ทางการประเมินอย่างน้อย 3 ปี) ทำการตัดสิน โดยเปรียบเทียบกับเกณฑ์มาตรฐานทั้ง 30 ข้อ ประกอบด้วยมาตรฐานด้านอรรถประโยชน์ 8 ข้อ มาตรฐานด้านความเป็นไปได้ 3 ข้อ มาตรฐานด้านความเหมาะสม 8 ข้อ และมาตรฐานด้านความถูกต้อง 11 ข้อ เกณฑ์การผ่านในแต่ละมาตรฐานยึดผู้เชี่ยวชาญ 3 ใน 4 หรือร้อยละ 75 มีความเห็นสอดคล้องกันจึงถือว่ามีความคุณภาพตามมาตรฐานนั้น ผลการประเมินปรากฏว่า รายงานการประเมินของประเทศกาน่า ได้รับความเห็นจากผู้เชี่ยวชาญว่า มีมาตรฐานสูงสุด 27 มาตรฐานการประเมินคิดเป็นร้อยละ 90 รองลงมาคือรายงานการประเมินของประเทศอินเดีย และ Sierra Leone มี 25 มาตรฐานการประเมิน คิดเป็นร้อยละ 83.30 ส่วนรายงานการประเมินประเทศพม่า ได้รับความเห็นจากคณะผู้เชี่ยวชาญน้อยที่สุด 19 มาตรฐานการประเมิน คิดเป็นร้อยละ 63.3

Boyd & Windsor (1993) ได้ทำการประเมินอภิมานเชิงสรุปรวม (summative metaevaluation) เพื่อรายงานผลการประเมิน โภชนาการการศึกษาของหญิงมีครรภ์ จำนวน 5 เรื่อง โดยใช้เกณฑ์การ

ประเมินของ Haymes et. al ซึ่งประกอบด้วยเกณฑ์ดังนี้ 1) การออกแบบการวิจัย 2) ขนาดกลุ่มตัวอย่าง และความเป็นตัวแทนของกลุ่มตัวอย่าง 3) ลักษณะประชากร 4) คุณภาพของการวัด และ 5) ความเหมาะสมความเป็นตัวแทนของกลุ่มควบคุม โดยในแต่ละเกณฑ์จะมีคะแนนเต็ม 25 เกณฑ์เท่ากับ 5 คะแนน รวมทั้ง 5 เกณฑ์ เช่น เกณฑ์ที่ 1 การออกแบบการวิจัย มีคะแนนดังนี้ จะได้ 4 คะแนน ถ้ามีกลุ่มควบคุม มีการสุ่มและใช้รูปแบบ pretest-posttest จะได้ 3 คะแนน ถ้าออกแบบเปรียบเทียบกลุ่มควบคุมกลุ่มไม่เท่าเทียมกัน จะได้ 2 คะแนน ถ้าออกแบบเปรียบเทียบกลุ่มควบคุมกลุ่มไม่เท่าเทียมกัน จะได้ 1 คะแนน ถ้าออกแบบ pretest-posttest และมีคะแนนโบนัส 1 คะแนน เมื่อใช้กลุ่มเท่าเทียมกันและมีการควบคุม จากนั้นนำรายงานการประเมินมาเปรียบเทียบกับเกณฑ์ ผลการประเมินอภิमानเชิงสรุปรวม ปรากฏว่ารายงานการประเมินของ Hunt et. al ได้คะแนนสูงสุด 18 คะแนน รายงานการประเมินของ Sweeney et. al , Bowering et. al และ Bruce and Tchado ได้ 13, 11, 10 และ 7 คะแนนตามลำดับ

จากรายงานการประเมินอภิमानโภชนาการการศึกษาของหญิงมีครรภ์ดังกล่าว จะเห็นได้ว่าการประเมินอภิमानเชิงสรุปรวม (summative metaevaluation) โดยประเมินในลักษณะเปรียบเทียบกับเกณฑ์ที่กำหนดไว้เพื่อเสนอ จุดดี จุดด้อย ของการประเมินต่อสาธารณะ

นอกจากนี้ Gallegos (1994) ได้ร่วมกับศูนย์เพื่อการวิจัยและการประเมินครูทางการศึกษา (the center for research on educational accountability and teacher evaluation) หรือ (CREAT) ได้ศึกษาการประเมินอภิमानรูปแบบการประเมินโรงเรียนในสหรัฐอเมริกา จำนวน 50 รูปแบบ (Metaevaluation School of Evaluation Models) โดยคณะผู้ประเมินอภิमानจะทำการประเมินงานประเมินอย่างอิสระจากกัน โดยการสังเคราะห์โมเดลที่ใช้ในการประเมินแล้วตรวจสอบกระบวนการตามเกณฑ์ ดังนี้ 1) ความมุ่งหมายของการประเมิน 2) กระบวนการ 3) ขนาดของโรงเรียน 4) บทบาทในการประเมิน เป็นการประเมินเชิงความก้าวหน้า หรือเป็นการประเมินเชิงสรุปรวม 5) ลักษณะการประเมินเชิงปริมาณ เชิงคุณภาพ หรือทั้งเชิงประมาณและเชิงคุณภาพ 6) การวางแผน การตัดสินใจ ประเมินบริบท ปัจจัยเบื้องต้น กระบวนการ และผลผลิต 7) การแสดงความรับผิดชอบ (accountability) ว่ามีการพัฒนาจากความต้องการ มีความเป็นธรรม ความเป็นไปได้ มีความเหมาะสมในการวางแผนทรัพยากร 8) มาตรฐานการประเมิน มาตรฐานอรรถประโยชน์ มาตรฐานด้านความเป็นไปได้มาตรฐานด้านความเหมาะสม มาตรฐานด้านความถูกต้อง และ 9) จุดแข็งจุดอ่อนของการประเมิน

จากการศึกษาดังกล่าวจะเห็นว่าเป็นการสังเคราะห์วิธีการประเมิน ซึ่งมีลักษณะคล้ายกับการสังเคราะห์งานวิจัย แล้วทำการประเมินที่สังเคราะห์อีกทีการศึกษาลักษณะนี้จะทำให้ทราบลักษณะของการประเมินโดยทั่วไปของโมเดลการประเมินโรงเรียนดังกล่าว

Lee (1996) ได้ศึกษาการประเมินอภิมานของสถาบันเทคโนโลยีแห่งชาติกระทรวงศึกษาธิการ แห่งประเทศไต้หวัน จากคณะผู้ประเมินและมีส่วนได้ส่วนเสีย ซึ่งจากการพิจารณาผลการประเมินจุด หลักของการประเมินอภิมานอยู่ที่การเปรียบเทียบการประเมินว่าสัมพันธ์กับมาตรฐานด้าน อรรถประโยชน์ มาตรฐานด้านความเป็นไปได้ มาตรฐานด้านความเหมาะสม มาตรฐานด้านความ ถูกต้องของสถาบันเทคโนโลยี ไต้หวันโดยใช้เกณฑ์มาตรฐานความสำคัญทั้ง เกณฑ์ในการประเมิน อภิมานมาใช้ เก็บข้อมูลโดยใช้แบบสอบถาม และแบบสัมภาษณ์แบบเป็นทางการจากคณะผู้ประเมิน และผู้มีส่วนได้ส่วนเสีย ผลการศึกษาจะไปเป็นข้อมูลพื้นฐานในการปรับปรุงโมเดลการประเมิน นโยบาย วิธีการและการปฏิบัติงาน ผลการศึกษาพบว่า 1) คณะผู้ประเมินและผู้มีส่วนได้ส่วนเสียเชื่อในเกณฑ์ มาตรฐานทั้ง 30 เกณฑ์ สามารถใช้ในการกำหนดทิศทางกระบวนการปฏิบัติ และ 2) คณะผู้ประเมิน และผู้มีส่วนได้เสียให้ผลการประเมินที่เหมือนกันในมาตรฐานด้านอรรถประโยชน์ มาตรฐานด้าน ความเป็นไปได้ มาตรฐานด้านความเหมาะสม และมาตรฐานด้านความถูกต้องของการประเมินสถาบัน เทคโนโลยีแห่งชาติไต้หวัน

นอกจากนี้มหาวิทยาลัยบริกแฮม ยัง (Brigham Young University) ได้ทำการประเมิน โครงการประเมินระยะการเรียนรู้ของโรงเรียนในเมืองนิวยอร์ก (the New York city public integrated learning systems project) เพื่อดูผลการประเมินการใช้ระบบการเรียนรู้แบบวอเตอร์ฟอร์ด (the Waterford integrated learning system) หรือ WILT คณะผู้ทำการประเมินอภิมานประกอบด้วย Finn Steven, Stufflebeam และ Waiberg โดยคณะผู้ประเมินในด้านจุดแข็ง จุดอ่อน ของแผนงานและ กิจกรรมการประเมิน ดังมีวิธีดำเนินการดังนี้

1. ทบทวนแผนการประเมิน วิธีดำเนินการเป็นระยะๆ และตรวจสอบสารสนเทศการประเมิน ปัจจุบัน
2. เชื่อมชมการใช้ระบบการเรียนรู้แบบวอเตอร์ฟอร์ด (WILS) ภายใน โรงเรียนต่างๆ เป็นระยะ
3. มีการสนทนาและข้อคิดเห็นระหว่างผู้ประเมินอภิมานกับคณะผู้ประเมิน โครงการรวมทั้งมี ส่วนได้ส่วนเสียอื่นๆ เป็นระยะ ๆ (โดยเฉพาะอย่างยิ่งผู้สนับสนุน โครงการ) เพื่อดูข้อมูลความก้าวหน้า ในปัจจุบัน และความต้องการของผู้มีส่วนได้ส่วนเสีย
4. วิเคราะห์ระบบประเมินของรายงานการประเมินขั้นสุดท้าย
5. เขียนร่างรายงานประเมินอภิมานของผู้ประเมินแต่ละคน
6. ถกปัญหาเพื่อให้ได้ฉันทามติ (consensus) เกี่ยวกับรายงานการประเมินอภิมาน
7. พัฒนาเป็นรายงานการประเมินอภิมานขั้นสุดท้าย

สรุปได้ว่างานวิจัยของ Finn (1997) เป็นการประเมินอภิมานเชิงความก้าวหน้าและเชิงสรุปรวม (formative and summative metaevaluation) เพื่อกำกับติดตามการประเมินโครงการระบบการเรียนรู้ของโรงเรียนในเมืองนิวยอร์ก ให้เป็นไปตามแผนการประเมิน แล้วจากนั้นทำการประเมินคุณภาพเชิงสรุปรวม เพื่อรายงานผล จุดดี จุดด้อย ต่อสาธารณะให้เป็นที่ยอมรับได้กับบุคคลทุกฝ่าย

### สรุป

การประเมินอภิมานเป็นการประเมินงานประเมินหรือคุณภาพของการประเมิน รวมไปถึงการประเมินการสังเคราะห์วิธีการที่ใช้ในการประเมินเพื่อให้ได้สารสนเทศที่เป็นประโยชน์ในการปรับปรุงงานและรายงานผลการประเมินในด้านใดด้านจุดแข็ง จุดอ่อนต่อสาธารณะ ซึ่งมีวิธีการในการประเมินอภิมานแตกต่างกันไป การประเมินอภิมานนั้น สามารถประเมินอภิมานได้ในเชิงความก้าวหน้า เพื่อชี้แผนการประเมินและการดำเนินการประเมินว่าเป็นไปตามแนวทางมาตรฐานที่กำหนดหรือไม่ ถ้าปฏิบัติการประเมินตามแนวทางจะทำให้ได้ผลการประเมินที่ถูกต้องสมเหตุผล นอกจากนี้สามารถประเมินอภิมานในเชิงสรุปรวม เพื่ออธิบายจุดดี จุดด้อยต่อสาธารณะ หรือทำการประเมินอภิมานทั้งในเชิงความก้าวหน้า และสรุปรวมทั้งสองอย่างร่วมกัน

### รายการอ้างอิง

- ศิริชัย กาญจนวาสี. (2537). **ทฤษฎีการประเมิน**. กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย
- สมหวัง พิธิยานุวัฒน์. (2541). **วิธีวิทยาการทางการประเมิน**. กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย
- สุวิมล ว่องวานิช. (2538). ความรู้ชายแดนด้านการประเมินผลการศึกษา. ใน สมหวัง พิธิยานุวัฒน์ (บรรณาธิการ), **วิธีวิทยาการวิจัย**. ปีที่ 7 ฉบับที่ 2 (กรกฎาคม-ตุลาคม 2538): 52-67.
- อุทุมพร จามมาน.(2540). ในสุวิมล ว่องวานิช (บรรณาธิการ), **108 คำถาม: การวิจัย การวัดและประเมินผลสถิติ**. พิมพ์ครั้งที่ 2. กรุงเทพฯ: วิ.ที.ซี.คอมมิวนิเคชั่น
- Berk,r.a.(1981) **Meta Evaluation; concepts. Standards and users. Educational Evaluation ethodology: The state of The Art**. Johns Hopkins University Press.
- Boyd, N.R. & Windsor, R.A.(1993). **Meta – evaluation of nutrition intervention research among Pregnant Women**. Herlth Education Quarterly. 20, No. 3: 327-345.
- Cook,T.D. (1978). Meta-evaluation research in **Evaluation Quarterly**. 2(1978): 5-51.
- Finn,C.E.(1997). A meta-evaluation. in **International Journal of Education Research**. 27,No.2:159-174.

- Gallegos,A.(1994). Meta-evaluation of school evaluation models. **Studies in Educational Evaluation.** 20: 41-54.
- Lee,R.T.(1995). **A meta-evaluation of Taiwan ministry of education’s national technology Institutes evaluation: A study of evaluation teams, and stakeholders, judgment on the evaluation practice.** (CD-ROM). Abstract form: ProQuest File: Dissertation Abstract Item: 9538604.
- Maduak, G.F., Scriven, M., & Stufflebeam, D.L. (1983). **Evaluation models: Viewpoints on Educational and Human Service Evaluation.** Boston: Kluwer-Nijhoff publication.
- Nyirenda. S. (1994). Assessing highly accomplished teaching: developing a metaevaluation criteria Framework for Performance assessment systems for national certification of teacher. in **Journal of Personel Evaluation in Education.** 8:313-327.
- Schwandt,T.A.(1988). **Linking Auditing and Metaevalion: Enhancing Quality in Applied Research.** Beverly Hill, CA:SAGE Publications.
- Scriven,M.(1991). **Evaluation Thesaurus.** 4<sup>th</sup> ed. Newbury Park,CA:SAGE Publications.
- Smith,N.L(1990).Caution on the use of investigative case studies in meta-evaluation. **Evaluation and Program Planning.** 13: 373-378. Stufflebeam, D.L. & Shinkfield J.A. Systematic Evaluation. Boston: Kluwer-Nijgoff Publication.
- Wahba,A.F.M(1990). **Assessing Meta-Evaluation in Practice: Application of the Joint Committee Standards to Final Evaluation of Traditional Birth Attendants (TBAs) in Four Beveloping Countries.** Doctor’s Thesis. Graduate School, University of Pittsburgh

\*\*\*\*\*

