

## **ความรู้ชายแดนด้านการประเมินผลการศึกษา**

**สุวิมล ว่องวานิช**

บทความนี้ต้องการนำเสนอความรู้ชายแดนด้านการประเมินผลการศึกษาซึ่งได้พัฒนาอย่างมากในช่วง 20-30 ปี ที่ผ่านมา โดยมีเนื้อหาครอบคลุมแนวคิดและการเปลี่ยนแปลงของแนวคิดด้านการประเมิน ประเด็นการประเมินที่เหล่านักคิด นักประเมินกำลังให้ความสนใจ ทั้งนี้ผู้เขียนได้ทำการรวบรวม วิเคราะห์และสังเคราะห์จากหนังสือ บทความ รายงานการวิจัย และเอกสารอื่น ๆ ที่เกี่ยวข้องกับการประเมิน

### **แนวคิดการประเมิน: อดีตสู่ปัจจุบัน**

นับแต่ Ralph W. Tyler (1930-1945) ผู้ซึ่งได้รับการยกย่องให้เป็น บิดาแห่งการประเมินทางการศึกษา (father of educational evaluation) ได้ให้นิยามของการประเมินว่าเป็นการศึกษาความสอดคล้องระหว่างผลที่ได้รับกับจุดมุ่งหมายที่กำหนดอันเป็นที่มาของการประเมินที่อิงวัตถุประสงค์ (objective-based evaluation) การประเมินผลเริ่มเข้าสู่ยุคที่มีความชัดเจนยิ่งขึ้น โดยเฉพาะความชัดเจนในสิ่งที่ต้องประเมิน (object of evaluation) ความหมายของ “การประเมิน” เริ่มมีการจำแนกให้เห็นความแตกต่างในความหมายของ “การวัดผล” หรือ “การวิจัย” และการประเมินผลตามแนวคิดนี้ก็ยังคงใช้ได้และหลายคนยังนิยมใช้กันอยู่จนทุกวันนี้ โดยเฉพาะการประเมินผลในห้องเรียน

---

บทความนำเสนอในการสัมมนาชุมชนนักคิดสาขาการวัดและประเมินผลการศึกษา เรื่องวิธีวิทยาขั้นสูงด้านการวิจัย สถิติ การวัดและประเมินผลการศึกษา วันจันทร์ที่ 21 สิงหาคม 2538 ณ ศศปฐศาลา จุฬาลงกรณ์มหาวิทยาลัย ขอขอบพระคุณ ศ.ดร.อุทุมพร จามรมาน และ ผศ.ดร.นงลักษณ์ วิรัชชัย ที่ช่วยเหลือจัดหาเอกสารที่เป็นประโยชน์ต่อการเขียนบทความนี้

มโนทัศน์เกี่ยวกับการประเมินได้รับการพัฒนามาเรื่อย จนมาถึงยุคแห่งการขยายความคิดเกี่ยวกับการประเมิน (The Age of Expansion 1958-1972) ในยุคนี้มีแนวคิดเกี่ยวกับการประเมิน ตลอดจนการนำเสนอโมเดลการประเมินอย่างมากมาย นิยามของการประเมินเริ่มเปลี่ยนจากการประเมินที่เน้นที่ผลผลิตอย่างเดียวมาเป็นการประเมินกระบวนการด้วยเพื่อให้ได้ข้อมูลที่ช่วยในการตัดสินใจคุณค่าหรืออนาคตของโครงการ และ/หรือข้อมูลที่ช่วยในการปรับปรุงโครงการ ในช่วงนี้นับว่าเป็นยุคของความพยายามที่จะทำความเข้าใจให้ตรงกัน ในหมู่นักประเมินว่า จริง ๆ แล้ว การประเมินคืออะไรกันแน่ และท้ายที่สุดก็ค่อนข้างจะทำให้ได้แนวคิดเกี่ยวกับการประเมินที่แบ่งได้เป็น 2 กลุ่มใหญ่ ๆ คือ การประเมินเพื่อการตัดสินใจคุณค่าของโครงการ (judgmental evaluation) และการประเมินเพื่อหาข้อมูลสนับสนุนการตัดสินใจโครงการ (decision-based evaluation) ทั้งนี้วิธีดำเนินการประเมินที่นักประเมินแต่ละคนเสนออาจจะแตกต่างกัน หรือมีจุดเน้นต่างกัน อย่างไรก็ตาม บทบาทของการประเมินในช่วงนี้มีการจำแนกอย่างเห็นได้ชัดเกี่ยวกับการประเมินความก้าวหน้า (formative evaluation) และการประเมินผลสรุป (summative evaluation) ตามแนวคิดของ Scriven ซึ่งเป็นที่ยอมรับกันมาก และแนวคิดนี้ก็สอดคล้องกับการประเมินกระบวนการ (process evaluation) และการประเมินผลผลิต (product evaluation) ซึ่งใช้กันมากในปัจจุบัน

ความคิดเกี่ยวกับการประเมินในยุคที่เรียกว่า The Age of Expansion (1958-1972) นี้อาจกล่าวได้ว่าประเด็นที่ยังเห็นไม่ตรงกันคือเรื่องของ “การตัดสินใจคุณค่า” (value judgment) Scriven เป็นผู้หนึ่งที่ทำให้ความสำคัญกับ “คุณค่า” อย่างมาก แม้จนถึงทุกวันนี้จุดยืนทางความคิดในเรื่องนี้ของเขาก็ยังไม่เปลี่ยนแปลงเลย นั่นคือ หากนักประเมินจะต้องทำการประเมินอะไรก็ตาม สิ่งที่นักประเมินต้องทำคือการให้ข้อสรุปเชิงประเมิน (evaluative conclusion หรือ evaluative claim) และถ้าผลการประเมินไม่สามารถสรุปในลักษณะนี้ได้ ก็ถือว่านักประเมินไม่ได้ทำหน้าที่ของตนเองอย่างสมบูรณ์ ความคิดของ Scriven ไม่ค่อยสอดคล้องกับแนวคิดของ Stufflebeam ซึ่งเสนอ CIPP Model ในช่วงแรก (1969, 1971) และนักประเมินอื่น ๆ เช่น Alkin (1969) คนกลุ่มหลังเห็นว่าการประเมินเป็นการจัดหาข้อมูลที่เป็นประโยชน์ที่ช่วยในการตัดสินใจของผู้ทำหน้าที่ตัดสินใจโครงการ นั่นคือ ผู้ประเมินไม่ได้ตัดสินใจคุณค่าของโครงการ แต่ให้ข้อมูลเพื่อให้ผู้มีอำนาจตัดสินใจคุณค่า อย่างไรก็ตาม แนวคิดในช่วงหลังของ Stufflebeam ก็ได้เปลี่ยนแปลงไป โดยเห็นว่าผู้ประเมินควรให้ข้อมูลเชิงประเมินด้วย

การพัฒนาแนวคิดเกี่ยวกับการประเมินในช่วงปี 1973 เป็นต้นมา นับเป็นความพยายามที่จะทำให้การประเมินมีความเป็นมาตรฐานทางวิชาชีพมากขึ้น (The Age of Professionalization) การประเมินจะให้ข้อมูลที่ เป็นประโยชน์ต่อการตัดสินใจต่อเมื่อได้ให้ข้อมูลที่สอดคล้องกับความต้องการของผู้เกี่ยวข้อง โดยเห็นได้ชัดจากการนำเสนอแนวคิดของ Stake ที่เรียกว่า responsive evaluation (1973) และแนวคิดของ Guba & Lincoln (1981) เกี่ยวกับการประเมินผล ซึ่งปรากฏในหนังสือที่ชื่อว่า effective evaluation ไม่นานหลังจากนั้น Patton ก็นำเสนอการประเมินที่เป็นแบบ utilization-focused evaluation เกือบทุกแนวคิดมีความเห็นว่าผู้ประเมินต้องให้ข้อมูลที่สอดคล้องกับความต้องการของผู้ใช้ จึงจะทำให้เกิดการนำผลการประเมินไปใช้ ความคิดเกี่ยวกับการประเมินของคนในกลุ่มนี้ส่วนใหญ่เป็นวิธีการที่จัดอยู่ในกลุ่มของ qualitative evaluation หรือ responsive evaluation

ความพยายามในการพัฒนาการประเมินให้เป็นวิชาชีพมากขึ้นในช่วงนี้สามารถแบ่งได้เป็น 3 กลุ่มใหญ่ ๆ คือ (1) การพัฒนาแนวคิดในเชิงการทฤษฎีที่ช่วยชี้นำการประเมินให้สนองความต้องการของผู้ใช้ (2) การพัฒนาในเชิงการกำหนดมาตรฐานการประเมินเพื่อให้การประเมินเป็นวิชาชีพที่ได้รับการยอมรับ และ (3) การพัฒนาในเชิงเทคนิคต่าง ๆ ที่ใช้ในการทำการประเมิน

การพัฒนาแนวคิดเชิงทฤษฎีในปัจจุบันเปลี่ยนจากช่วงก่อนปี 1973 แต่เดิม ประเด็นการประเมินเป็นไปตามความคิดของนักประเมินที่เห็นว่าสำคัญและเป็นประโยชน์ต่อผู้ใช้ แต่ในช่วงหลังพัฒนาการของการประเมินเปลี่ยนไปเป็นว่า ประเด็นการประเมินควรขึ้นอยู่กับความต้องการของผู้ใช้ นั่นคือ ต้องศึกษาก่อนว่าผู้ใช้ผลการประเมินต้องการอะไร แล้วจึงเอาตัวนั้นมากำหนดแนวทางในการประเมิน ดังนั้น ช่วงหลังจึงมีการเสนอแนวคิดที่เป็นแบบ utilization-focused evaluation หรือ need-based evaluation กันมากขึ้น

ในขณะที่กลุ่มหนึ่งยังคงพัฒนาความคิดเกี่ยวกับการประเมินเชิงทฤษฎี อีกกลุ่มหนึ่งก็เริ่มพัฒนาความคิดเกี่ยวกับการทำให้การประเมินเป็นกระบวนการที่เป็นมาตรฐาน มีการตั้งคณะกรรมการทำงานเพื่อกำหนดมาตรฐานของการประเมิน (standard for evaluation) มีการนำเสนอความคิดที่เรียกว่า meta evaluation (โดย Scriven) เพื่อตรวจสอบการประเมินอีกครั้ง (evaluation of evaluation) และมีการรวมกลุ่มกันเพื่อพัฒนาความก้าวหน้าให้กับศาสตร์ของการประเมินอย่างต่อเนื่อง ไม่ว่าจะเป็นการก่อตั้งสมาคมการประเมิน การจัดทำวารสาร

ทางวิชาการเพื่อแลกเปลี่ยนความรู้และข่าวสารด้านการประเมิน หรือการตั้งกลุ่มผู้สนใจศึกษา ภายใต้หัวข้อที่สมาชิกในกลุ่มสนใจ (เช่น Topical Interest Group) ก็ตาม นับเป็นการรวมกลุ่มของนักประเมินในการพัฒนาศาสตร์ของการประเมินให้แก่งยิ่งขึ้น

นอกจากนี้ในส่วนเทคนิคที่ใช้ในการประเมินก็มีการพัฒนาอย่างมาก นักประเมินเห็น ว่าบทบาทของตนเองไม่ได้จำกัดอยู่เพียงการประเมินผลของโครงการเท่านั้น แต่เห็นว่าจะต้อง เป็นผู้มีส่วนในการพัฒนาปรับปรุงโครงการได้ด้วย (to improve program) และการพัฒนา โครงการให้ดีขึ้นไม่ได้หมายความว่า ต้องมีโอกาสเข้าไปทำการประเมินระหว่างที่โครงการ กำลังดำเนินอยู่นั้น แต่ควร จะต้องมีโอกาสเข้าไปมีส่วนร่วมตั้งแต่การพัฒนาหรือจัดตั้ง โครงการ ผู้ที่เชื่อในแนวคิดเกี่ยวกับ need-based evaluation หรือ consumer-oriented evaluation น่าจะเห็นด้วยกับแนวคิดของการมีส่วนร่วมในการจัดทำโครงการตั้งแต่ต้น เพราะ ทำให้ได้โครงการที่มีลักษณะตอบสนองความต้องการของผู้ที่ได้รับผลกระทบจากโครงการมากที่สุด แต่เดิมการประเมินเกิดขึ้นหลังจากโครงการสิ้นสุด ซึ่งทำได้เพียงการประเมิน “ผล” (product evaluation) ต่อมาแนวคิดเกี่ยวกับการประเมินระหว่างโครงการกำลังดำเนิน อยู่ด้วย ทำให้เกิดการประเมิน “กระบวนการ” (process evaluation) ขณะนี้มีแนวคิดเกี่ยวกับการประเมินก่อนมีโครงการ ทำให้เกิดการประเมิน “ความต้องการ” (needs assessment) บทบาทของผู้ประเมินเริ่มไปผูกติดอยู่กับวงจรการบริหารจัดการโครงการมากขึ้น และผูกติด กับกลุ่มผู้บริหารจัดการด้วย

ความคิดเกี่ยวกับการประเมินไม่ว่าจะเป็นแนวใด ต่างมุ่งไปที่เป้าหมายเดียวกัน คือ การก่อให้เกิดการพัฒนาและดำเนินโครงการที่เอื้อประโยชน์สูงสุดต่อผู้เกี่ยวข้อง การทำการ ประเมินผลลัพธ์อย่างเดียวยังไม่ทำให้สามารถหาจุดบกพร่องมาปรับปรุงโครงการได้ และการทำ การประเมินแต่กระบวนการแม้จะทำให้ได้ข้อมูลที่ช่วยในการปรับปรุงโครงการ แต่ก็ทำให้ต้อง เสียเวลากับการดำเนินโครงการที่ไม่เกิดประโยชน์เท่าที่ควร หากพบภายหลังว่าโครงการนั้น ไม่ได้สนองตอบความต้องการของผู้เกี่ยวข้องอย่างแท้จริง บทบาทของผู้ประเมินจึงเป็นทั้งผู้ ตัดสินคุณค่าของสิ่งที่ถูกประเมิน (merit and worth evaluator) เป็นผู้อำนวยความสะดวก ในการประเมิน (evaluation facilitator) และขณะเดียวกันก็ดูเหมือนว่าจะทำหน้าที่เป็นนัก- พัฒนาโครงการด้วย (program developer)

โดยสรุป จะเห็นได้ว่าการประเมินมีพัฒนาการที่เปลี่ยนไปเป็นยุค ๆ ช่วงแรกเป็นยุคของความพยายามทำความเข้าใจเกี่ยวกับความหมายของการประเมิน และความพยายามสร้างรูปแบบหรือนำเสนอแนวคิดเชิงทฤษฎีเกี่ยวกับการประเมิน ต่อมาด้วยความพยายามจะสร้างความเป็นวิชาชีพให้กับศาสตร์ของการประเมิน โดยเน้นที่การทำให้การประเมินมีมาตรฐานเป็นที่ยอมรับ ณ จุดนี้ นิยามเกี่ยวกับการประเมินค่อนข้างเป็นที่ยุติ แม้จะมีความแตกต่างกันทางความคิดบ้าง แต่ก็ถือว่าเป็นเรื่องไม่สำคัญมากนัก สิ่งที่นักประเมินแต่ละคนต่างมีความเห็นสอดคล้องคือ การประเมินจะไม่มีประโยชน์เลยหากไม่ได้มีการนำไปใช้ประโยชน์ในการปรับปรุง เปลี่ยนแปลงหรือพัฒนาสิ่งที่ถูกประเมินให้ดีขึ้น ช่วงนี้จะพบว่ามีการศึกษาวิธีการที่จะนำไปสู่การใช้ประโยชน์จากการประเมิน และขณะเดียวกันความพยายามสร้างศาสตร์ของการประเมินก็ยังคงดำเนินต่อไป

ความรู้ชายแดนเกี่ยวกับการประเมินผล จึงน่าจะครอบคลุม 3 ประเด็นที่น่าสนใจ คือ การพัฒนาวิทยาการหรือศาสตร์ของการประเมินให้แกร่ง การพัฒนามาตรฐานของการประเมิน และการพัฒนาเทคนิควิธีที่ช่วยในการทำการประเมินให้ได้คำตอบที่มีคุณค่าเพียงพอและเหมาะสมต่อการตัดสินใจ

### วิทยาการของการประเมิน (Discipline of Evaluation)

ในขณะที่การวิจัยและการวัดผลการศึกษาต่างมีวิทยาการ (discipline) ของตนเองหลายคนอาจสงสัยว่า การประเมินมีวิทยาการของตนเองหรือไม่ เรื่องนี้มีผู้ถกเถียงกันมากบ้างก็ว่าไม่น่าจะมี เนื่องจากไม่ได้มีการสร้างทฤษฎีอะไรของตนเองเลย บ้างก็ว่ามีวิทยาการแต่เป็นแบบสหวิทยาการ (interdisciplinary) ในเรื่องนี้ Scriven (1994) ได้เสนอบทความชื่อว่า "Evaluation as a Discipline" ซึ่งมีรายละเอียดที่น่าสนใจจะศึกษาเป็นอย่างยิ่ง Scriven กล่าวว่า การประเมินถือเป็นวิทยาการหนึ่ง แต่เป็นวิทยาการที่เรียกว่า transdiscipline เนื่องจากเป็นวิทยาการที่เป็นเครื่องมือ (tool discipline) เหมือนวิทยาการด้านสถิติหรือการวัดผล คือมีเนื้อหา หรือวิธีการที่สามารถนำไปใช้ในศาสตร์อื่น และตัวการประเมินเองก็สามารถนำความรู้จากศาสตร์อื่นมาใช้ในกระบวนการประเมินได้ ลักษณะของ transdiscipline จะประกอบด้วย 2 ส่วน คือ ส่วนที่เป็น "วิทยาการที่เป็นแกน" (core discipline) และส่วนที่เป็น "วิทยาการประยุกต์" (applied discipline) วิทยาการที่เป็นแกน เป็นแนวคิดหรือทฤษฎีที่เกี่ยวกับการประเมิน ซึ่งเป็นเรื่องของ logic of evaluation ในขณะที่ส่วนที่เป็น

วิทยาการประยุกต์นั้น เป็นการนำความรู้ไปประยุกต์ใช้กับวิทยาการอื่น ๆ Scriven ยกตัวอย่างว่าวิทยาการประยุกต์ที่สำคัญมี 6 ประการ ซึ่งเขาเรียกว่า Big Six ได้แก่ program evaluation, personnel evaluation, performance evaluation, product evaluation, proposal evaluation, และ policy evaluation นอกจากนี้เขายังกล่าวว่า meta evaluation (evaluation of evaluation) ก็ถือว่าเป็นวิทยาการประยุกต์ เพราะนำวิทยาการด้านการประเมินมาใช้ในการประเมินตัวเอง (self-reference)

Scriven เห็นว่ามีความจำเป็นต้องพัฒนาวิทยาการของการประเมินขึ้นมา ที่ผ่านมามีคนส่วนหนึ่งจะไม่ยอมรับว่าการประเมินเป็นศาสตร์หนึ่ง (science of evaluation) เนื่องจากการประเมินเกี่ยวข้องกับ “การตัดสินคุณค่า” (value judgment) ในขณะที่ทางวิทยาศาสตร์ถือว่าสิ่งที่สนใจศึกษา “ไม่มีส่วนที่เป็นคุณค่า” (value-free science) ดังนั้น คนกลุ่มนั้นจึงไม่คิดว่าการประเมินเป็นวิทยาศาสตร์ และแม้หลายคนอาจจะยอมให้การประเมินเป็นวิทยาศาสตร์ แต่ต้องไม่มีส่วนที่เป็น “การตัดสินคุณค่า” รวมอยู่ด้วยโดยปล่อยให้การตัดสินคุณค่าของสิ่งที่ถูกประเมินเป็นหน้าที่ของผู้มีอำนาจในการตัดสินใจ ความคิดเหล่านี้เองที่ Scriven มองว่าทำให้ศาสตร์ของการประเมินไม่ก้าวหน้า และยังตั้งข้อสงสัยว่าทำไมศาสตร์ของการประเมินจึงต้องใช้เวลาอันยาวนานในการพัฒนา เขามองว่าส่วนหนึ่งเนื่องมาจากคนนอกศาสตร์ไม่ยอมรับความเป็นศาสตร์ และคนในแวดวงการประเมินเองก็ยังไม่มีความเห็นสอดคล้องกันทางความคิดของศาสตร์ด้านนี้ ดังนั้น จึงมีความจำเป็นต้องพัฒนาศาสตร์นี้ให้ได้

Scriven ได้เสนอความคิดว่า หากมองการประเมินว่าเป็น transdiscipline ก็ควรจะประกอบด้วยลักษณะสำคัญ 4 ประการ คือ (1) มีความเป็นปรนัย (objectivist view) โดยนักประเมินมีหน้าที่ให้ข้อสรุปเชิงประเมิน (evaluative conclusions) ในการประเมินสิ่งต่าง ๆ (2) มีลักษณะที่เป็นการอิงผู้ได้รับผลกระทบจากสิ่งที่ถูกประเมิน (consumer-oriented view) (3) มีลักษณะที่สรุปได้กว้าง (generalized view) โดยสามารถสรุปโมทัศน์เกี่ยวกับการประเมินไปได้ทุก ๆ สาขาของการประเมิน กล่าวคือ ต้องให้แนวคิดเกี่ยวกับการประเมินที่ใช้ได้กับการประเมินเรื่องใด ๆ ก็ตาม ไม่เจาะจงว่าเป็นการประเมินโครงการหรือการประเมินผลผลิต เป็นต้น และ (4) มีลักษณะที่เป็นความรู้เชิงเทคนิควิธี (technical view) ที่เป็นวิธีการของวิทยาการของการประเมินเอง

หากพิจารณาโมเดลหรือแนวคิดเกี่ยวกับการประเมินที่ผ่านมา Scriven วิจารณ์ว่าหลายแนวคิดไม่คำนึงถึงบทบาทหน้าที่ของผู้ประเมินในการตัดสินคุณค่า ในที่นี้ต้องทำความเข้าใจให้ตรงกันว่า คำว่า “ตัดสินคุณค่า” นั้น หมายถึง การให้ข้อสรุปเชิงประเมินเกี่ยวกับคุณค่าของสิ่งที่ถูกประเมินซึ่งกระทำโดยผู้ประเมิน ในบางแนวคิดข้อมูลที่ได้จากการประเมินเป็นเพียงข้อมูลสนับสนุนการตัดสินใจ แต่ไม่เป็นข้อสรุปเชิงประเมิน (เช่น แนวคิดของ Alkin) แต่บางแนวคิดสามารถให้ข้อมูลเชิงประเมินได้ (เช่น แนวคิดของ Stake, Provus, Rossi & Freeman) แต่ปล่อยให้ผู้มีอำนาจในโครงการเป็นผู้ตัดสินคุณค่าเอง

ในช่วงต้นปี 1990s มีการรวมกลุ่มกันของผู้สนใจที่จะพัฒนาศาสตร์ทางการประเมิน เช่น มีกลุ่ม TIG (Topical Interest Group) ซึ่งสนใจศึกษาเกี่ยวกับ Theories of Evaluation นำโดย Fournier และคณะ (1993) คนกลุ่มนี้เห็นว่า ทฤษฎีการประเมินน่าจะจำแนกออกเป็น 2 ประเภท คือ **Normative Theories** เป็นทฤษฎีที่เกี่ยวกับตรรกะหรือความสมเหตุสมผลของการประเมิน (logic and justification of evaluation) และ **Descriptive Theories** เป็นทฤษฎีที่เกี่ยวกับปฏิบัติการการประเมิน (practice of evaluation) การทำงานของคนกลุ่มนี้ยังไม่มีข้อสรุปที่แน่นอน สิ่งที่พวกเขากำลังคิดต่อไป คือ ความพยายามหากระบวนการที่มีเหตุผลที่จะให้ข้อสรุปเชิงประเมิน (reasoning process in coming to evaluative conclusions) และการจัดประเภทของการประเมินแบบนี้ Scriven (1994) ได้วิจารณ์ว่า เป็นการกล่าวถึงการประเมินใน 2 ประเด็น คือ ทฤษฎีการประเมิน (Theories of Evaluation) และทฤษฎีเกี่ยวกับการประเมิน (Theories about Evaluation) โดยเขากล่าวว่า ถ้าเป็น theories of evaluation หมายถึง logic of evaluation ส่วน theories about evaluation หมายถึง practice of evaluation

จากผลงานด้านความคิดของคนกลุ่มนี้ประกอบกับการวิจารณ์ของ Scriven และบทความ โดย Scriven ให้ข้อสรุปว่าปัจจุบันประเด็นของความพยายามสร้างทฤษฎีของการประเมินหรือวิทยาการของการประเมินกำลังได้รับความสนใจจากกลุ่มนักคิดทั้งหลาย สิ่งที่ Scriven เสนอประกอบกับการวิเคราะห์วิจารณ์จากแนวคิดต่าง ๆ ผู้เขียนคิดว่าศาสตร์ของการประเมินต้องให้ความสนใจในประเด็นต่อไปนี้

- (1) ข้อมูลจากการประเมินควรเป็น value judgment หรือเป็น decision support
- (2) การประเมินต้องให้ได้ข้อสรุปเชิงประเมินหรือไม่ (evaluative conclusions)

- (3) นักประเมินมีบทบาทหน้าที่ต้องให้ข้อสรุปเชิงประเมินหรือไม่ (draw evaluative conclusions) หรือให้ผู้บริหารโครงการเป็นผู้ตัดสินใจ
- (4) logic of evaluation หรือที่เป็น core discipline ซึ่ง Scriven จัดอยู่ใน theories of evaluation นั้นควรประกอบด้วย ทฤษฎี แนวคิด หรือมโนทัศน์อะไรบ้าง

### การพัฒนามาตรฐานของการประเมิน

ในปี 1981 Stufflebeam และคณะ ได้พัฒนามาตรฐานของการประเมินเพื่อเป็นแนวทางในการดำเนินการประเมินให้เป็นที่เชื่อถือ นับเป็นความพยายามในการทำให้ศาสตร์ของการประเมินเป็นวิชาชีพที่ยอมรับได้ อย่างไรก็ตาม ในช่วงหลังนับแต่ Evaluation Network (ENet) และ Evaluation Research Society (ERS) ได้มารวมกันเป็นสมาคมชื่อว่า American Evaluation Association (AEA) ทางสมาคมก็เริ่มตั้งกลุ่มคณะทำงาน (Task Force) ในปี 1992 เพื่อหาหลักปฏิบัติสำหรับนักประเมิน (Guiding Principles for Evaluators) เพื่อให้เกิดความเป็นมาตรฐานในวิชาชีพของการประเมิน ในปี 1993 คณะทำงานโดย Dianna Newman และคณะ ได้เสนอข้อสรุปเกี่ยวกับแนวปฏิบัติสำหรับนักประเมินให้กับคณะกรรมการบริหารสมาคม ซึ่งประกอบด้วยหลักใหญ่ 5 ประการ และข้อความย่อย ๆ จำนวน 23 ข้อ และข้อเสนอดังกล่าว มีการแจกให้สมาชิกของสมาคม AEA ลงคะแนนว่าจะยอมรับในหลักปฏิบัติ 5 ประการและข้อความย่อยหรือไม่ สำหรับหลักการที่กลุ่มเสนอและยังไม่ได้มีการยอมรับอย่างเป็นทางการนั้น มีดังต่อไปนี้ โดยทั้ง 5 หลักการนี้ไม่ได้มีการเรียงตามลำดับความสำคัญ

1. **Systematic Inquiry:** Evaluators conduct systematic, data-based inquiries about whatever is being evaluated.
2. **Competence:** Evaluators provide competent performance to stakeholders.
3. **Integrity/Honesty:** Evaluators ensure the honesty and integrity of the entire evaluation process.
4. **Respect for People:** Evaluators respect the security, dignity and self-worth of the respondent, program participants, clients, and other stakeholders with whom they interact.



5. **Responsibilities for General and Public Welfare:** Evaluator articulate and take into account the diversity of interests and values that may be related to the general and public welfare.

จะเห็นว่าแม้จะมีกลุ่มการทำงานที่นำโดย Stufflebeam ได้เสนอมาตรฐานของการประเมินมาก่อนหน้านี้แล้วก็ตาม การพัฒนาวิชาการประเมินให้เป็นที่ยอมรับก็ไม่ได้หยุดนิ่งอยู่กับที่ สมาคม AEA ได้ตั้งคณะทำงานภายใต้สมาคม AEA เพื่อศึกษาเรื่องนี้ดังกล่าวข้างต้น คณะทำงานซึ่งเสนอหลักปฏิบัติ 5 ประการนั้น ได้ให้ข้อคิดเห็นเพิ่มเติมว่า การกำหนดหลักการดังกล่าวไม่ได้หมายความว่า จะนำมาใช้แทนมาตรฐานการประเมินที่มีการกำหนดก่อนหน้านี้ แต่ต้องการให้เกิดการพัฒนาวิชาชีพให้นักประเมินได้ทำงานอย่างถูกต้องเหมาะสมและท้ายที่สุดได้ให้ข้อคิดเพิ่มเติมว่าหลักการที่กำหนดอาจจะยึดติดกับวัฒนธรรมทางตะวันตกโดยเฉพาะประเทศสหรัฐอเมริกา หากวัฒนธรรมหรือประเทศแตกต่างออกไป หลักปฏิบัติเกี่ยวกับการประเมินก็อาจเปลี่ยนแปลงไปจากที่เสนอก็ได้

## **แนวคิดทฤษฎี วิธีการด้านการประเมินที่นำศึกษาพัฒนาต่อไป**

### **แบบของการประเมิน**

#### **Qualitative Evaluation, Participatory Evaluation**

วิธีการที่ใช้ในการประเมินถ้าศึกษาให้ดีจะเห็นว่าค่อนข้างจะคล้อยตามวิธีวิทยาการวิจัย ในขณะที่การวิจัยมีการทำวิจัยแบบ quantitative research ทางการประเมินก็มีการทำ quantitative evaluation หรือเมื่อมีการวิจัยแบบ qualitative research ทางการประเมินก็มี qualitative evaluation วิจัยมีการทำ meta-analysis ทางการประเมินมีการทำ meta-evaluation ที่เป็นเช่นนี้อาจเนื่องจากกลุ่มนักวิชาการใน 2 ศาสตร์นี้ได้รับพื้นฐานการพัฒนาความรู้มาใกล้เคียงกัน นักวิจัย นักวัดผล และนักประเมินอาจจะเป็นคนกลุ่มเดียวกัน จึงเกิดการแลกเปลี่ยนวิทยาการที่พัฒนาขึ้นมาใหม่ ดังนั้น ถ้าพิจารณาการออกแบบการประเมินในช่วงหลังจะเห็นว่าเริ่มมีการนำเสนอแนวคิดเกี่ยวกับการประเมินเชิงคุณภาพ (qualitative evaluation) มาใช้กันมากในการประเมินงานต่าง ๆ

ในขณะที่ศาสตร์ด้านการวิจัยมีการกล่าวถึง การทำวิจัยเชิงคุณภาพที่เป็นแบบ Participatory Research หรือ Participatory Action Research (PAR) ในการประเมินก็มีการ

กล่าวถึง การประเมินแบบ Participatory Evaluation เมื่อเร็ว ๆ นี้ Garaway (1995) ได้เขียนบทความเกี่ยวกับการประเมินที่นำผู้ที่เกี่ยวข้องกับการประเมินเข้ามามีส่วนในการประเมินด้วย ทั้งนี้กลุ่มของผู้ประเมินมาจากผู้มีส่วนร่วมในโครงการ (project participants) และนักประเมินอาชีพ (professional evaluator) คนกลุ่มนี้จะร่วมกันทำงานตั้งแต่การจัดการ การลงมือทำงานด้านการประเมินและการเผยแพร่ผลการประเมินไปยังกลุ่มสมาชิกต่าง ๆ ที่เกี่ยวข้อง คำถามที่น่าสนใจคือกลุ่มของผู้มีส่วนร่วมในโครงการที่จะนำมามีส่วนร่วมในการประเมินคือใคร และจะทำการประเมินโดยวิธีนี้อย่างไร Garaway ได้นำเสนอกรอบความคิดในการทำการประเมินโดยวิธีนี้ และให้ข้อคิดเห็นว่ สิ่งที่สำคัญของการประเมินแบบนี้คือการมีผู้อำนวยความสะดวกด้านการประเมินที่เหมาะสม (evaluation facilitator) การกำหนดวิธีการประเมินที่เหมาะสม การคำนึงถึงความตรง และค่าใช้จ่ายที่ใช้ในการประเมิน ตลอดจนเวลาที่ใช้ในการประเมิน การประเมินแบบนี้ต้องมีค่าใช้จ่ายรวมทั้งส่วนของผู้ประเมินที่ทำหน้าที่เป็น facilitator และผู้มีส่วนร่วมในการประเมิน Garaway มองว่าการประเมินแบบนี้มีข้อดี เพราะเป็นการดึงเอาบุคลากรที่มีประสบการณ์ในงานของตนเองเข้ามาร่วมในการประเมิน และถือว่าคนกลุ่มนี้เป็นผู้ตัดสินใจโครงการได้ดีกว่าผู้อื่น และยังเป็นการสร้างบรรยากาศของการเรียนรู้ที่ดี โดยการนำผู้ใช้ผลการประเมิน (end-users) มาทำการประเมินร่วมกันกับนักประเมินอาชีพ

การประเมินแบบ participatory evaluation นี้ไม่เหมือนการประเมินภายใน (internal evaluation) ซึ่งทำการประเมินโดยบุคคลในหน่วยงาน ไม่มีนักประเมินอาชีพมาช่วยในการประเมิน หากจะใช้นักประเมินอาชีพต้องอยู่ในรูปแบบที่ให้นักประเมินอาชีพทำการประเมินที่เป็นอิสระจากนักประเมินภายใน ซึ่งเรียกว่าเป็น การประเมินภายนอก (external evaluation) นั่นคือ นักประเมินภายในและนักประเมินภายนอกมิได้ทำงานการประเมินร่วมกันเหมือน participatory evaluation

#### Fourth Generation Evaluation Theory

ระเบียบวิธีที่ใช้ในการประเมินที่กำลังกล่าวถึงมากและถือเป็นคลื่นลูกใหม่ของทฤษฎีการประเมินในขณะนี้ คือ Fourth Generation Evaluation Theory ของ Guba & Lincoln (1989) แนวคิดนี้จัดอยู่ในกลุ่มของ naturalistic evaluation ลักษณะเด่นที่ทำให้คนส่วนหนึ่งชอบในแนวคิดนี้ คือการให้อำนาจในการตัดสินใจกับผู้ได้รับผลกระทบจากโครงการมาก (stakeholders) มีการให้คุณค่าเท่า ๆ กันในการสร้างกลุ่มผู้ประเมินซึ่งเป็นผู้ได้รับผลกระทบจากการประเมิน ทำให้เกิดการยอมรับร่วมกันในข้อสรุปหรือประเด็นต่าง ๆ ที่ได้จากการประเมิน

และทำให้เกิดความมั่นใจในการใช้กระบวนการประเมินและผลการประเมิน (O'Neill, 1995) อย่างไรก็ตาม Scriven (1994) ได้วิจารณ์แนวคิดนี้ว่าจัดอยู่ในกลุ่ม constructivism และไม่ค่อยเห็นด้วยกับแนวคิดนี้ เนื่องจากการประเมินตามแนวนี้จะได้ถือว่าการประเมินเป็นการค้นหาคุณภาพ คุณค่าหรือความคุ้มค่าของสิ่งที่ถูกประเมิน แต่ถือว่าการประเมินเป็นผลจากการทำงานของคนแต่ละกลุ่มแล้วมาหาข้อสรุปร่วมกัน บทความช่วงหลัง ๆ มีการวิพากษ์วิจารณ์แนวคิดการประเมินแบบนี้กันมาก และน่าจะเป็นแนวคิดที่ยังต้องศึกษาและวิเคราะห์วิจารณ์กันอีกนาน

### **Need-based Evaluation, Consumer-Oriented Evaluation**

Scriven กล่าวถึงรูปแบบการประเมินที่เป็น consumer-oriented evaluation ไว้นานแล้ว แต่ปรากฏว่าความคิดนี้เริ่มมีการให้ความสำคัญกันมากขึ้น ในขณะที่ Stake (1972) เสนอความคิดเกี่ยวกับ responsive evaluation พบว่าความสำคัญอยู่ที่ clients หรือ stakeholders ซึ่งต้องทำความเข้าใจให้ตรงกันว่า คนกลุ่มนี้หมายถึงใคร ถ้าหมายถึงเฉพาะกลุ่มที่เป็นผู้ตัดสินใจ (decision makers) ซึ่งเป็นคนขอให้ทำการประเมิน ก็ไม่น่าจะสอดคล้องกับความคิดของ Scriven เพราะแนวคิดของ Scriven นั้น เป้าหมายของการประเมินเน้นที่การนำผลการประเมินไปใช้ กลุ่มผู้ใช้ (users) เป็นกลุ่มที่ไม่ได้หมายถึงผู้มีอำนาจตัดสินใจเพียงกลุ่มเดียว เพราะการประเมินเกี่ยวข้องกับคนหลายกลุ่ม คนที่ได้รับผลกระทบเหล่านี้ (impacted population) ควรได้มีส่วนร่วมรับรู้ผลการประเมินด้วย แนวคิดการประเมินแบบ consumer-oriented evaluation จึงควรให้ความสำคัญมากกว่า manager-oriented evaluation (Scriven, 1989) เพราะการประเมินต้องดูที่ผลที่เกิดขึ้นว่าเป็นไปตามความต้องการของกลุ่มคนที่ได้รับผลกระทบหรือไม่ โดยเหตุนี้ การประเมินจึงไม่ควรยึดติดอยู่กับวัตถุประสงค์อย่างเดียว แต่ควรเป็นแบบ goal-free evaluation ดังนั้น การประเมินจึงต้องคำนึงถึง needs ของคนที่ได้รับผลกระทบด้วย และเพื่อให้ชัดเจนยิ่งขึ้น ก็ควรเรียกการประเมินให้เจาะจงว่าเป็น need-based evaluation ซึ่งเป็นแนวคิดเดียวกับ consumer-oriented evaluation

### **Individualized Objective-Based Evaluation**

ในการประเมินโครงการทางการศึกษา หากโครงการดังกล่าวให้บริการกลุ่มคนที่มีระดับต่าง ๆ กัน และบริการที่ให้แตกต่างกัน วัตถุประสงค์ของโครงการจะขึ้นอยู่กับกลุ่มคนที่เข้ามารับบริการ (program participants) ระดับของกลุ่มคนที่ต่างกัน จะมีความต้องการไม่เหมือนกัน ดังนั้นจึงมีผู้เสนอแนวคิดว่าการประเมินที่อิงวัตถุประสงค์ ควรพิจารณาวัตถุประสงค์แต่ละด้าน

หากการประเมินใด ๆ อิงการประเมินกระบวนการหรือผลงานที่เป็นมาตรฐานเดียวกัน น่าจะให้ข้อมูลที่ถูกต้อง แนวคิดการประเมินแบบ individualized objective-based evaluation จึงเกิดขึ้น (Lam, 1995) และเป็นรูปแบบการประเมินที่น่าสนใจซึ่งแตกต่างไปจาก objective-based evaluation

### Meta-Evaluation

วิทยาการที่แข็งแกร่งต้องมีการตรวจสอบและบ่งบอกถึงจุดบกพร่องของตนเองได้ การทำ meta-evaluation จึงเป็นการประเมินคุณภาพของการประเมินอีกครั้ง อย่างไรก็ตาม ตัวอย่างของการประเมินแบบนี้มีไม่มากนัก คนที่อยู่ในสาขาวิจัยแต่ไม่มีความชำนาญในสาขาการประเมิน มักจะสงสัยว่า meta-analysis แตกต่างจาก meta-evaluation อย่างไร การทำ meta-analysis คือการสังเคราะห์ผลการวิจัย นั่นคือ การนำผลการวิจัยที่มีประเด็นวิจัยเดียวกันมาสังเคราะห์เพื่อหาข้อสรุปรวม ในขณะที่ meta-evaluation เป็นการประเมินของการประเมินอีกครั้ง หากพิจารณาบริบทของการประเมินแล้วจะเห็นว่าไม่น่าจะมีการสังเคราะห์ผลการประเมินได้ เนื่องจากการประเมินอะไรก็ตาม ผลการประเมินที่ได้ขึ้นอยู่กับบริบทของหน่วยงานนั้น ๆ (context-dependent) โครงการหนึ่งดำเนินการภายใต้หน่วยงานหนึ่ง ด้วยเงื่อนไขหนึ่ง ผลการประเมินจะออกมาเป็นแบบหนึ่ง จึงเป็นเรื่องที่ไม่สมเหตุสมผลและไม่น่าจะเกิดประโยชน์ด้วย ที่จะนำเอาผลการประเมินโครงการในลักษณะเดียวกันของหน่วยงานต่าง ๆ หรือผลการดำเนินงานของโครงการในหน่วยงานเดิม แต่ทำมาแล้วหลาย ๆ ครั้ง มาสังเคราะห์เพื่อหาข้อสรุปรวมอีกครั้ง เพราะเป้าหมายของการประเมินไม่ได้เน้นที่การสรุปอ้างอิงผลการประเมินไปยังประชากรหรือกลุ่มเป้าหมายอื่นเหมือนการวิจัย เรื่องของ generalizability จึงไม่ใช่ประเด็นสำคัญของการประเมิน

อย่างไรก็ตาม จากการศึกษารายงานผลการประเมินของ Gallegos (1994) ซึ่งทำการประเมินเรื่อง Meta-Evaluation of School Evaluation Models โดยผู้ประเมินนำโมเดลที่ใช้ในการประเมินโรงเรียนจากที่ต่าง ๆ ในสหรัฐอเมริกาจำนวน 51 โมเดล มาทำการสังเคราะห์โมเดลที่ใช้ในการประเมิน โดยตรวจสอบกระบวนการ มาตรฐานและเกณฑ์ที่ใช้ในการประเมิน ทั้งนี้มีการใช้คณะผู้ประเมินเข้าไปทำการประเมินโมเดลเหล่านั้นอย่างอิสระจากกัน บทความนี้ทำให้ผู้เขียนเริ่มเปลี่ยนความคิดใหม่เกี่ยวกับการทำ meta-evaluation นั่นคือ การทำ meta-evaluation ไม่น่าจะหมายความว่าเพียงการประเมินวิธีการหรือคุณภาพของการประเมินในโครงการใดโครงการหนึ่งโดยเฉพาะ แต่สามารถทำได้โดยการสังเคราะห์วิธีการที่ใช้ในการประเมิน

(synthesis of evaluation process) ซึ่งตรงจุดนี้จะคล้าย meta-analysis แล้วทำการประเมินคุณภาพของวิธีการที่ได้จากการสังเคราะห์เหล่านั้น อันทำให้ได้ข้อมูลที่สามารถสรุปรวมเกี่ยวกับจุดดีจุดด้อยของกระบวนการหรือวิธีการทั้งหลายที่ใช้ในการประเมินสิ่งใดสิ่งหนึ่งได้ดี เป็นการผสมผสานความคิดของการสังเคราะห์กับการตรวจสอบคุณภาพของการประเมินเข้าด้วยกัน

### Needs Assessment

นักประเมินหากต้องทำหน้าที่เป็นผู้พัฒนาโครงการ (program developer) ก็ต้องมีความรู้ด้านการวัดความต้องการ (needs assessment) เทคนิคการวิเคราะห์ความต้องการได้รับการพัฒนาอย่างมากในช่วงปี 1980 เป็นต้นมา หลักของการวัดความต้องการอยู่ที่ความพยายามหาความแตกต่างระหว่างสิ่งที่มุ่งหวังกับสิ่งที่ป็นจริง ช่องว่างที่เกิดขึ้นบอกถึงสภาพปัญหาที่เกิดขึ้นและต้องการแก้ไข และที่สำคัญต้องมีการจัดเรียงลำดับความสำคัญของปัญหาดังกล่าวที่เกิดขึ้น เทคนิคการวัดความต้องการจึงแบ่งออกได้เป็นหลายประเภท ที่ใช้กันมากได้แก่ (1) กระบวนการกลุ่ม (group process) เช่น brainstorming, focus-group, technique, nominal group technique (2) เทคนิคที่เกี่ยวกับการจัดเรียงลำดับความสำคัญของปัญหาถือเป็นการวัดความต้องการที่อยู่ในกลุ่ม setting priority (3) เทคนิคเชิงอนาคต (future technique) เช่น Delphi Technique, Forecasting Technique, Fault Tree Analysis, Cross-Impact Analysis และ (4) การใช้วิธีการสำรวจ (needs survey) จากการศึกษาวิธีการวัดความต้องการที่มีการทำกันมาตั้งแต่ปี 1981 ของ Witkin (1994) พบว่าส่วนใหญ่มีการทำในวงการศึกษา และกลุ่มเป้าหมายเป็นนักเรียน ครู ซึ่งถือเป็นกลุ่มผู้รับบริการ (service provider หรือ primary needs) นอกจากนี้ยังพบว่าวิธีการที่ใช้ส่วนใหญ่เป็นวิธีการแบบเดี่ยว (single-method study) แทนที่จะเป็นพหุวิธี (multiple methods) เทคนิคที่เป็นแบบการจัดเรียงลำดับความสำคัญไม่ค่อยได้มีการใช้ ผลการวิจัยนี้ชี้ให้เห็นว่าเทคนิคการวัดความต้องการที่ยังใช้กันยังมีขีดจำกัด แม้จะมีหลายเทคนิคแต่ก็ยังใช้กันอย่างไม่กว้างขวาง

เรื่องของการทำ needs assessment เป็นเรื่องที่มีผู้ให้ความสนใจกันมาก เห็นได้ชัดจากรายงานการวิจัยในช่วงปี 1990 เป็นต้นมา มีประเด็นวิจัยใหม่ทั้งการพัฒนา needs assessment models เพื่อใช้กับงานด้านต่าง ๆ การศึกษาในเชิงการประยุกต์ใช้ และการพัฒนาเชิงเทคนิควิธี เพื่อหาวิธีการวัดความต้องการที่เหมาะสมและให้ข้อมูลที่น่าเชื่อถือ ที่ผ่านมามีเทคนิคส่วนใหญ่ที่ใช้กันมักนำมาจากศาสตร์หรือสาขาวิชาอื่น เช่น การทำ cost-analysis, forecasting technique เป็นต้น น่าจะถึงเวลาแล้วที่นักประเมินควรให้ความสนใจกับการพัฒนาเทคนิคเหล่านี้ให้เป็นวิทยาการของตนเอง

### Evaluability assessment

การประเมินก็เหมือนกับการดำเนินโครงการ ต้องมีการตรวจสอบก่อนเริ่มลงมือปฏิบัติงาน ว่ามีความเป็นไปได้ที่จะทำการประเมินมากน้อยเพียงใด โครงการนั้น ๆ สามารถทำการประเมินได้หรือไม่ มโนทัศน์เกี่ยวกับเรื่องนี้มีผู้เสนอมานานแล้ว (เช่น Smith, 1989) โดยเห็นว่าการประเมินต้องเป็นเครื่องมือที่ช่วยพัฒนาโครงการและทำให้ได้ผลการประเมินที่เป็นประโยชน์ การทำ evaluability assessment จึงเป็นกระบวนการเพื่ออธิบายโครงสร้างของโครงการ เช่น วัตถุประสงค์ของโครงการ ตระกาะของโครงการ กิจกรรม ตัวบ่งชี้ความสำเร็จของโครงการ นอกจากนี้ยังเป็นการวิเคราะห์ความเป็นไปได้ในการทำให้โครงการบรรลุวัตถุประสงค์ ทำให้การประเมินให้ข้อมูลที่ลึกซึ้ง และเป็นที่ยอมรับของผู้บริหาร ผู้ตัดสินใจนโยบาย ผู้จัดทำโครงการ กล่าวโดยสรุปเป็นเสมือนการวิเคราะห์โครงการก่อนการประเมิน อย่างไรก็ตาม แม้แนวคิดนี้จะไม่ค่อยใหม่ แต่พบว่าการประเมินเองไม่ค่อยมีการดำเนินการในเรื่องนี้อย่างเป็นระบบมากนัก จึงเป็นอีกประเด็นหนึ่งที่น่าจะมีการพัฒนาเทคนิควิธีหรือประยุกต์เทคนิคไปใช้ให้มากขึ้น

### Teacher Evaluation

การประเมินครูอาจารย์นับเป็นตัวอย่างหนึ่งของการประยุกต์วิธีการประเมินไปใช้โมเดล การประเมินครูอาจารย์มีมากมาย แต่ละแบบต่างมีลักษณะสำคัญเฉพาะ Scriven (1995) ได้เสนอรูปแบบการประเมินถึง 12 รูปแบบ โดยได้แยกลักษณะของโมเดลออกเป็น 2 ประเภท คือ ประเภทที่เป็น criterion model ซึ่งจะให้แนวคิดเกี่ยวกับเกณฑ์ที่ใช้ในการประเมิน (เช่น การประเมินครูอาจารย์โดยอิงความสามารถในการสอน อิงผลการปฏิบัติงาน เป็นต้น) กับประเภทที่เป็น method model ซึ่งให้แนวคิดเกี่ยวกับวิธีการที่ใช้ในการประเมิน (เช่น การใช้เพื่อนร่วมงานเป็นผู้ให้ข้อมูล การให้ผู้เชี่ยวชาญเป็นผู้ประเมิน การใช้วิธีการสัมภาษณ์ในการเก็บข้อมูล) Scriven แสดงความเห็นว่าการประเมินบุคคลต้องเน้นที่ merit และเพื่อให้การประเมินยุติธรรมควรประเมินครูจากงานที่มอบหมายให้ทำ หรืองานที่มีการตกลงกันตอนว่าจ้าง ดังนั้น Scriven จึงไม่เห็นด้วยกับการประเมินครูที่ใช้ Research-Based Teacher Evaluation (RBTE) แต่เห็นว่า Duties-Based Teacher Evaluation (DBTE) หรือบางคนเรียก Duties of the Teacher (DOTT) มีความเหมาะสมกว่าเนื่องจากครูอาจารย์ควรถูกประเมินโดยเกณฑ์ที่มีความชอบธรรม การรับครูอาจารย์เข้าทำงานมีการตกลงกันว่าให้ทำหน้าที่สอนหนังสือ บทบาทส่วนใหญ่จึงอยู่ที่ห้องเรียน ดังนั้น หากจะประเมินครูก็ควรประเมินผลการปฏิบัติงานในห้องเรียนซึ่งถือเป็นหน้าที่โดยตรง การเสนอแนวคิดนี้นับว่าเป็นประโยชน์

ต่อการประเมินบุคลากร (personnel evaluation) และมีหลายประเด็นที่น่าจะนำมาพิจารณาในการพัฒนาโมเดลต่อไป

## บทสรุป

บทความนี้ได้รับอิทธิพลทางความคิดอย่างมากจาก Scriven ซึ่งถือเป็นนักคิด นักปรัชญาที่สำคัญคนหนึ่งของวงการประเมิน ในช่วง 2-3 ปีนี้ จะพบแนวคิดของ Scriven ในบทความต่าง ๆ ตามวารสารการประเมินเป็นจำนวนมาก ถ้าวิเคราะห์ดูแล้ว จะเห็นว่าหลักการทางความคิดของนักประเมินผู้นี้ไม่ค่อยเปลี่ยนแปลงไปจากเดิมเลย ความคิดที่เขาเสนอส่วนใหญ่จะเป็นแนวคิดที่เป็นหลักการ ไม่เจาะจงว่าจะเขียนขึ้นมาเพื่อใช้ในการประเมินสิ่งใดสิ่งหนึ่งโดยเฉพาะ เพราะเขามีความเชื่อว่าศาสตร์ของการประเมินจะต้องแสดงความคิดเชิงทฤษฎีที่มีความเป็นทั่วไปสูง เขาเห็นว่าบทความหรือหนังสือด้านการประเมินผลส่วนใหญ่จะใช้คำว่า “การประเมิน” ไม่ถูกต้อง นั่นคือ สิ่งที่น่าเสนอมักจำกัดอยู่ที่การประเมินโครงการหรือการประเมินบุคลากร หรือการประเมินทางการศึกษา (ซึ่งมักหมายถึงการประเมินนักเรียนเป็นส่วนใหญ่) จริง ๆ แล้ว แนวคิดการประเมินต้องพัฒนาให้ใช้ได้กว้าง จึงจะเกิดประโยชน์ต่อการประเมินอย่างแท้จริง

หากบทบาทของผู้ประเมินต้องมีถึง 3 ประการ คือ การเป็นผู้ตัดสินคุณค่า การเป็นผู้อำนวยความสะดวกในการทำการประเมิน และการเป็นผู้มีส่วนในการพัฒนาโครงการ ศาสตร์ของการประเมินต้องมีลักษณะที่ผสมผสานวิทยาการต่าง ๆ เข้าด้วยกันให้เป็นวิทยาการของตนเอง สิ่งที่น่าพัฒนาในเชิงแนวคิด คือ ขอบเขตการประเมินควรครอบคลุมอะไรบ้างมากน้อยเพียงใด พัฒนาการของแนวคิดการประเมินในขณะนี้เป็นที่ยอมรับร่วมกันว่า ควรมีการเปลี่ยนแปลงความคิดจากที่เคยเน้นแต่การประเมินผลลัพธ์สุดท้ายอย่างเดียว มาเป็นการประเมินเพื่อก่อให้เกิดการพัฒนาอันเป็นบทบาทของการประเมินความก้าวหน้า แนวคิดที่เริ่มเปลี่ยนแปลงไป คือ การทำให้การประเมินสนองความต้องการของผู้ใช้ และประเด็นนี้เองที่ยังไม่มีข้อสรุปชัดเจนว่าผู้ใช้ควรครอบคลุมกลุ่มใดบ้าง และจำเป็นหรือไม่ที่ผู้ประเมินต้องเป็นผู้ตัดสินคุณค่าของสิ่งที่ถูกประเมิน จะปล่อยให้เป็นที่ของผู้มีอำนาจในการตัดสินหรือผู้ได้รับผลกระทบจากโครงการได้หรือไม่ และหากผู้ประเมินต้องเป็นผู้ตัดสินคุณค่าเอง จะใช้กระบวนการแบบใดเพื่อให้ได้ข้อสรุปดังกล่าว ทั้งหมดนี้เป็นงานที่ทำทลายความคิดของกลุ่มคนที่เกี่ยวข้องกับการประเมินเป็นอย่างยิ่ง

## หนังสืออ้างอิง

- Dissertation Abstracts International.** (1989-1994). Topics on Needs Assessment. University Microfilms, Inc.
- Evaluation Practice.** American Evaluation Association. Jai Press Inc. (Vol,15 No. 1, 1994).
- Evaluation Practice.** American Evaluation Association. Jai Press Inc. (Vol,15 No. 2, 1994).
- Evaluation Practice.** American Evaluation Association. Jai Press Inc. (Vol,15 No. 3, 1994).
- Evaluation Practice.** American Evaluation Association. Jai Press Inc. (Vol,16 No.1, 1995).
- Evaluation Practice News.** American Evaluation Association. (June, 1994).
- Gallegos, A. (1994). Meta-Evaluation of School Evaluation Models. **Studies in Educational Evaluation.** Vol.20, pp.41-54.
- Garaway, G.B. (1995). Participatory Evaluation. **Studies in Educational Evaluation.** Vol.21, pp.85-102.
- Madaus, G.F.; Scriven.M.S.; and Stufflebeam, D.1. (1989). **Evaluation Models: Viewpoints on Educational and Humen Services Evaluation.** Boston: Kluwer Academic Publishers.
- O'neil, T. (1995). Implementation frailties of Guda and Lincoln's Fourth Generation Evaluation Theory. **Studies in Educational Evaluation.** Vol.21, pp.5-21.
- Personal letter from Fournier, D.M.to Utumporn Jamornmann, dated April 29, 1993. Re. Theories of Evaluation.
- Report from the AEA Task Force on Guiding Principles for Evaluators (Draft). (Revised 18 August 1993).
- Scriven,M. (1991). Pros and Cons about Goal-Free Evaluation. **Evaluation Practice.** Vol. 12, No.1, pp.55-76
- Scriven, M. (1995). A Unified Theory Approach to Teacher Evaluation. **Studies in Educational Evaluation.** Vol.21, pp.111-129.
- Scriven, M. (1995). **Theories of Evaluation.** Topical Interest Group of the American Evaluation Association. Vol. 2, No. 1.
- Scriven, M. (1995). Evaluation as a Discipline. **Studies in Educational Evaluation.** Vol.20, pp.147-166.
- Smith, M.F. (1989). **Evaluability Assessment: A Practical Approach.** Boston: Kluwer Academic Publishers.