

รากแก้วของการประเมิน*

ศิริชัย กาญจนวาสี**

การประเมินเกิดขึ้นเมื่อใดไม่มีใครทราบ แต่พอสันนิษฐานได้ว่า คงเริ่มมีมาตั้งแต่มีมนุษย์อุบัติขึ้นบนโลกนี้ ส่วนจะเข้าใจว่าการประเมินคืออะไร มีลักษณะเป็นอย่างไร ทำแบบไหนจึงมีคุณภาพ คงมีการถกเถียงและความเข้าใจที่หลากหลายแตกต่างกันไปตามกาลเวลาและสถานที่ ความรู้ความเข้าใจเกี่ยวกับการประเมินดังกล่าว คงสืบทอด สะสมกันมาทั้งอย่างเป็นทางการและเป็นทางการตลอดเวลา องค์กรความรู้ทางการประเมินที่เรารู้จักกันทุกวันนี้จึงเป็นผลมาจากพัฒนาการทางความคิดและประสบการณ์ของนักปฏิบัติ นักทฤษฎีการประเมินที่ช่วยกันสร้าง สานต่อและสะสมมาเป็นเวลาอันยาวนาน จนเป็นสิ่งที่รู้จักและใช้กันอย่างกว้างขวางทุกวงการในปัจจุบัน

การประเมินเป็นกิจกรรมที่ปฏิบัติสืบทอดกันมาเป็นเวลานานตั้งแต่ครั้งโบราณกาลจากหลักฐานที่ปรากฏสามารถสืบย้อนไปได้ถึงประมาณสี่พันปีก่อนคริสต์ศักราช จักรพรรดิของจีนได้ใช้การตรวจสอบสมรรถภาพ ข้าราชการทหารทุก 3 ปี เป็นหลักในการพิจารณาการเลื่อนตำแหน่ง ประมาณสามพันปีมาแล้วที่จีนได้ริเริ่มการใช้ข้อสอบคัดเลือกบุคคลเข้ารับตำแหน่งทางราชการ นอกจากนี้เป็นที่ทราบกันมานานเช่นกันว่าปรมาจารย์ชาวกรีก เช่น โสกราตีส ได้ใช้การตั้งคำถามและการประเมินคำตอบ เป็นสื่อในกระบวนการเรียนรู้ เป็นต้น

การประเมินในลักษณะที่ปฏิบัติกันทุกวันนี้เพิ่งมีอายุได้ไม่ถึงศตวรรษ(Guba & Lincoln,1981) ช่วงเวลาระหว่างการประเมินครั้งโบราณกาล จนถึงปัจจุบันการประเมินได้ผ่านขั้นตอนของพัฒนาการที่มีระบบแบบแผนมากขึ้นตลอดเวลา เช่น ปลายศตวรรษที่ 19 Joseph M. Rice ได้ใช้แบบสอบสัมฤทธิ์ผลเก็บรวบรวมข้อมูลจากนักเรียนเพื่อตรวจสอบประสิทธิภาพของโรงเรียน ปี ค.ศ.1904 Alfred Binet ได้สร้างเครื่องมือจำแนกเด็กที่ด้อยประสิทธิภาพออกจากชั้นเรียนของเด็กปกติ ซึ่งต่อมาได้กลายเป็นรากฐานของแบบสอบทางสติปัญญา ค.ศ.1930 Ralph W. Tyler ได้ใช้วัตถุประสงค์ของหลักสูตรเป็นเกณฑ์ในการประเมินผลทางการศึกษา จนกระทั่ง ค.ศ.1960 เป็นต้นมา การประเมินได้วิวัฒน์ทั้งในรูปแบบและแนวคิดทางทฤษฎีโดยนักทฤษฎีการประเมินที่สำคัญ เช่น *Ralph W. Tyler, Lee J. Cronbach, Michael S. Scriven, Robert L. Stake, Daniel L. Stufflebeam* เป็นต้น จึงทำให้การประเมินมีความมั่นคงเป็นปึกแผ่นในลักษณะวิชาชีพแขนงใหม่ซึ่งมีความแตกต่างที่เด่นชัดยิ่งขึ้นจากศาสตร์แห่งการวัดผลและการทดสอบ และมีคุณลักษณะสำคัญที่แตกต่างจากการวิจัย

* ตีพิมพ์ในเอกสารประกอบการประชุมสัมมนาทางวิชาการ การวิจัย การวัดและประเมินทางการศึกษาแห่งประเทศไทยครั้งที่ 15 ณ มหาวิทยาลัยราชภัฏอุบลราชธานี

** Ph.D.(Measurement and Evaluation) UCLA ศาสตราจารย์ประจำภาควิชาวิจัยและจิตวิทยาการศึกษา คณะศึกษาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

1. รากฐานและปรัชญา

การประเมินเป็นวิธีการศึกษาความจริงในด้านคุณค่า หรือค่าของข้อเท็จจริงในแง่ต่างๆภายใต้บริบททางสังคม การศึกษาความเป็นจริงโดยตรงนับว่ามีความยากลำบากอยู่แล้ว ยิ่งการศึกษา “ค่า” หรือ “คุณค่า” ของความจริง จึงนับได้ว่าเป็นความสลับซับซ้อนมากยิ่งขึ้นการประเมินจึงเป็นกระบวนการที่ไม่ง่ายและทำให้ได้ดีได้ยาก เพราะต้องอาศัยหลักปรัชญา วิทยาศาสตร์ สังคมศาสตร์ มนุษยศาสตร์ ความรอบรู้ในเนื้อหา และการใช้วิจารณญาณอันสมเหตุสมผลในทุกขั้นตอนของการดำเนินงาน

วิธีหนึ่งที่จะทำความเข้าใจธรรมชาติสำคัญของคำว่า “การประเมิน” คือ การวิเคราะห์หาคำตอบว่า “การประเมินมีรากฐานมาจากอะไร?” ตลอดจนการศึกษาวิเคราะห์ปรัชญาที่อยู่เบื้องหลังการประเมินตามแนวคิดของนักทฤษฎีการประเมินคนสำคัญที่ใช้กันเป็นแบบอย่างของการประเมินในทุกวันนี้ ปรัชญาที่นักทฤษฎีการประเมินยึดเป็นหลักในการสร้างรูปแบบการประเมินสามารถมองได้ในหลายแง่มุม มุมหนึ่งของการมองอาจวิเคราะห์ได้จากการตอบคำถามเชิงปรัชญาของการประเมินว่า “เราทำการประเมินไปทำไม?” และ “เราควรทำการประเมินอย่างไร?” ผลการวิเคราะห์เชิงปรัชญาที่ได้จากคำถามดังกล่าว น่าจะเป็นกรอบความคิดสำหรับการออกแบบและการทำประเมินที่มีคุณภาพได้เป็นอย่างดี

1.1 รากฐานของการประเมิน

จากการศึกษาพัฒนาการของความหมายของการประเมิน และพัฒนาการของกระบวนการทัศน์ทางการประเมิน (ศิริชัย กาญจนวาสี, 2536) การประเมินมีวิวัฒนาการของความหมายผ่านเส้นทางดังนี้ 1) การวัด (*Measurement*) 2) การวิจัยประยุกต์ (*Applied Research*) 3) การตรวจสอบความสอดคล้องกับวัตถุประสงค์ (*Determining congruence*) 4) การช่วยผู้บริหารตัดสินใจ (*Assist decision making*) 5) การบรรยายอย่างลุ่มลึก (*Description or Portrayal*) และ 6) การตัดสินคุณค่า (*Determining of worth or value*)

รวมทั้งการติดตามพัฒนาการของกระบวนการทัศน์ทางการประเมิน ซึ่งสามารถจัดช่วงพัฒนาการของแนวคิดการประเมินได้ดังนี้ 1) ยุครุ่งอรุณแห่งความเรืองปัญญา (วิทยาศาสตร์) (ค.ศ. 1600-1800) 2) ยุคสร้างแนวคิดการวัดและการประเมิน (ค.ศ. 1801-1900) 3) ยุคการวัดและประเมินประสิทธิภาพ (ค.ศ. 1901-1930) 4) ยุคการประเมินอิงวัตถุประสงค์ (ค.ศ. 19301-1945) 5) ยุคการประเมินโดยใช้แบบสอบถามมาตรฐาน (ค.ศ. 1946-1957) 6) ยุคการขยายตัวของทฤษฎีการประเมิน (ค.ศ. 1958-1972) และ 7) ยุคพัฒนาสู่การเป็นวิชาชีพ (ค.ศ. 1973-ปัจจุบัน) ตลอดจนการศึกษางานเขียนของ *Marvin C. Alkin* (2004) ที่แกะรอยรากฐาน แนวคิดและอิทธิพลทางความคิดของนักประเมินคนสำคัญทำให้วิเคราะห์ได้ว่า ภาพความเจริญก้าวหน้าของการประเมินในปัจจุบันในด้านเป้าหมาย บทบาท กระบวนการทัศน์ และรูปแบบการประเมินตั้งอยู่บนรากฐานของหลักการสองประการ คือ หลักการสืบเสาะแสวงหาความจริงทางสังคม เพื่อตัดสินคุณค่าของสิ่งต่างๆตามบริบทของ

สังคมนั้นๆ อย่างน่าเชื่อถือ และหลักความรับผิดชอบต่อสังคม เพื่อติดตาม กำกับ และพัฒนาสู่เป้าหมายที่พึงประสงค์

หลักการสืบเสาะแสวงหาความจริงทางสังคม เพื่อตัดสินคุณค่าของสิ่งต่างๆ ตามบริบทของสังคม นั้น จำเป็นต้องใช้วิธีการศึกษาที่สังคมยอมรับและให้ความเชื่อถือ การประเมินจึงเป็นวิธีการที่ตั้งอยู่บนพื้นฐานของการบูรณาการ กระบวนการทางวิทยาศาสตร์ สำหรับแสวงหาความจริง และกระบวนการทางสังคม (ปรัชญา จิตวิทยา สังคมศาสตร์ มนุษยศาสตร์) สำหรับทำความเข้าใจคุณค่าและการตัดสินคุณค่าในมิติต่างๆ ทางสังคม

หลักความรับผิดชอบต่อสังคม เพื่อติดตาม กำกับและพัฒนาสู่เป้าหมายที่พึงประสงค์นั้นจำเป็นต้องมีกระบวนการติดตามตรวจสอบที่โปร่งใสและยุติธรรม การประเมินจึงเป็นกลไกที่ตั้งอยู่บนพื้นฐานของการติดตาม กำกับ และส่งเสริมการใช้สารสนเทศให้เกิดประโยชน์ต่อการปรับปรุง เปลี่ยนแปลงและพัฒนาสู่เป้าหมายที่พึงประสงค์

1.2 ปรัชญาการประเมิน

1.2.1 เราทำการประเมินไปทำไม?

นักทฤษฎีการประเมินมีความเชื่อร่วมกันอย่างหนึ่งว่า การประเมินเป็นกิจกรรมที่ตั้งอยู่บนพื้นฐานของปรัชญาการเมืองแบบเสรีนิยม (*liberalism*) เพราะยึดหลักการเปลี่ยนแปลงในเชิงสร้างสรรค์ที่ว่า ทุกสิ่งทุกอย่าง (ถ้าใช้แนวทางที่เหมาะสม)สามารถแตกต้อง สัมผัส และพัฒนาได้ หรือกล่าวได้อีกนัยว่า การประเมินเป็นสิ่งทีวิวัฒน์มาจากมโนทัศน์เสรี ประชาธิปไตย (*liberal democracy*) เราจึงกล่าวได้ว่าการประเมินว่าเป็นกระบวนการสร้างสรรค์สังคมที่สอดคล้องกับแนวทางประชาธิปไตย

แม้ว่านักทฤษฎีการประเมินส่วนใหญ่จะยึดแนวทางเสรีนิยมร่วมกัน แต่ความเชื่อใน “เป้าหมาย” ของการประเมินยังมีความแตกต่างกัน ขึ้นกับข้อตกลงเบื้องต้นของการประเมินว่านักทฤษฎีเหล่านั้นยึดถือในแนวไหน ซึ่งอาจจำแนกเป็นประเภทใหญ่ๆ ได้ดังนี้

1) ประโยชน์นิยม (*Utilitarianism*) นักทฤษฎีในแนวนี้เชื่อว่าเป้าหมายของการประเมิน คือการสร้างประโยชน์สุขแก่สังคม การประเมินจะมีคุณค่าก็ต่อเมื่อการประเมินนั้นเป็นประโยชน์หรือทำให้เกิดความผาสุกแก่มนุษย์และสังคม การประเมินจะมีคุณค่าสูงสุดเมื่อสามารถนำผลที่ได้ไปพัฒนาให้เกิดประโยชน์สูงสุดสำหรับผู้เกี่ยวข้องจำนวนมากที่สุด นักทฤษฎีในแนวนี้จึงเน้นการประเมินที่สามารถเสนอสารสนเทศ ที่เป็นประโยชน์ต่อระบบการบริหารงาน การวางแผน การดำเนินโครงการ และการตัดสินใจในเชิงบริหาร โดยนักประเมินมีบทบาทเป็นผู้ป้อนสารสนเทศที่เป็นประโยชน์ไม่ควรเข้าไปมีบทบาทในการตัดสินคุณค่าของสิ่งที่ทำการประเมินด้วยตนเอง แต่ให้เป็นดุลยพินิจของผู้ใช้สารสนเทศนั้นๆ เป็นผู้ชี้ขาดคุณค่าของสิ่งที่ทำการประเมิน นักทฤษฎีการประเมินที่สามารถจัดได้ว่ามีแนวคิดแบบประโยชน์

นิยม เช่น Tyler (1950), Provus (1971), Stufflebeam และคณะ (1971), Stake (1975), Patton (1980,1987), Alkin (1972), Fetterman (1994), King (1998) เป็นต้น

2) พหุนิยม (Pluralism) นักทฤษฎีในแนวนี้นี้เชื่อว่าเป้าหมายของการประเมิน มิใช่เพียงหนึ่งเดียว หรือเพื่อประโยชน์สูงสุดแก่สังคมเพียงอย่างเดียว อาจสนองเป้าหมายอื่นๆควบคู่ไปด้วย เช่น ทำการประเมินเพื่อแสดงกลไกแห่งอำนาจในการติดตาม/ควบคุม /ดูแลการดำเนินงานต่างๆ การประเมินเพื่อตัดสินคุณค่าที่แท้จริงของสิ่งที่ทำการประเมินตามความรู้และประสบการณ์ความเชื่อของนักประเมินซึ่งอาจจะตรงหรือไม่ตรงกับความต้องการของผู้ใช้สารสนเทศก็ได้ เป็นต้น นักทฤษฎีในแนวนี้นี้ส่วนใหญ่ได้รับอิทธิพลจากแนวคิดของอิมพิริออลนิยมนิยม (intuitionism) ซึ่งเชื่อว่านักประเมินคือ ผู้เชี่ยวชาญที่อุดมด้วยความรู้และประสบการณ์ ดังนั้นเป้าหมายของการประเมินที่สำคัญคือการตัดสินคุณค่า ของสิ่งของหรือเหตุการณ์ที่ทำการประเมิน การประเมินจะเสร็จสิ้นลงอย่างสมบูรณ์ก็ต่อเมื่อนักประเมินได้ตัดสินคุณค่าของสิ่งที่มุ่งประเมินแล้วตามความรู้ ประสบการณ์และความเชี่ยวชาญของตน ส่วนผลที่ได้จะนำไปใช้ประโยชน์หรือไม่? ใช้ประโยชน์ได้เพียงใด? นั่นถือว่าเป็นปัจจัยรอง นักทฤษฎีกลุ่มนี้จึงเน้นการประเมินที่มีการเก็บรวบรวมข้อมูลตามความเชื่อ และประสบการณ์ของผู้ประเมินเพื่อตัดสินคุณค่าของสิ่งที่ มุ่งประเมิน ตามความเชี่ยวชาญของตน ในขณะที่เดียวกันก็พยายามสร้างศรัทธาและความเชื่อถือเพื่อให้ผลเป็นที่ยอมรับของผู้ใช้สารสนเทศ นักทฤษฎีการประเมินที่สามารถจัดได้ว่ามีแนวความคิดแบบพหุนิยม เช่น Scriven (1967), Worthen และ Sanders (1973) , Eisner (1975-1979) , Guba และ Lincoln (1981) , House (1991)

1.2.2 เราควรทำการประเมินอย่างไร?

นักทฤษฎีการประเมินค่อนข้างมีความเห็นสอดคล้องกันว่า การประเมินควรใช้วิธีที่น่าเชื่อถือและมีหลักฐานอ้างอิงได้ แต่ “วิธีการ” ที่จะนำไปสู่ความจริงและกำหนดคุณค่าของความจริง ยังมีความหลากหลายอันเนื่องมาจากนักทฤษฎีการประเมินมีปรัชญาความเชื่อที่เกี่ยวข้องกับความจริง และการกำหนดคุณค่าที่แตกต่างกัน จึงทำให้เสนอวิธีการประเมินที่แตกต่างกันไป ซึ่งขอจำแนกเป็นวิธีการใหญ่ๆ ได้ดังนี้

1) อัตนัยนิยม (Subjectivism) นักทฤษฎีการประเมินในแนวนี้นี้มีความเชื่อว่าจิตไม่สามารถหยั่งรู้ได้อย่างแท้จริงว่าอะไรจริงหรือเท็จ มีคุณค่าหรือไม่มีคุณค่า การตัดสินว่าอะไรจริงหรือเท็จ/มีคุณค่าหรือไม่ เป็นเรื่องของแต่ละบุคคลขึ้นอยู่กับจิตของผู้คนว่าจะมองความจริงหรือคุณค่าอย่างไร จึงไม่มีมาตรการที่แน่นอนในการกำหนดคุณค่าของสิ่งต่างๆในทางปฏิบัติ นักประเมินจึงควรกำหนดมาตรฐาน หรือเกณฑ์ในการตัดสินคุณค่าขึ้นมาเอง ด้วยเหตุผล ส่วนจะเป็นที่ยอมรับหรือไม่ก็ขึ้นอยู่กับความรอบรู้ ประสบการณ์และความสามารถในการปลุกฝังศรัทธาในตัวผู้ใช้สารสนเทศได้เพียงใด นักทฤษฎีกลุ่มนี้จึงนิยมทำการประเมินโดยเน้น วิธีการเชิงธรรมชาติ (naturalistic approach) ซึ่งอยู่บนหลักการของวิธีการดำเนินการที่ยืดหยุ่น รวบรวมข้อมูลอย่างรอบด้านตามสภาพธรรมชาติและใช้ความรู้ความเชี่ยวชาญเป็น

เกณฑ์ในการสรุปผล นักทฤษฎีการประเมินที่สามารถจัดได้ว่านิยมวิประเมินแบบอัตนัยนิยม เช่น Stake (1975), Eisner (1975, 1979), Guba และ Lincoln (1981), Patton (1980, 1987) เป็นต้น

2) ปรัชญา (Objectivism) นักทฤษฎีการประเมินในแนวนี้มีความเชื่อว่ามนุษย์เราสามารถรู้ว่าจะอะไรจริงหรือเท็จ มีคุณค่าหรือไม่มีคุณค่า การตัดสินว่าอะไรจริงหรือเท็จ /มีคุณค่าหรือไม่ เป็นสิ่งที่สามารถกำหนดเป็นมาตรการสากลได้เพราะความจริงเป็นสิ่งสากล ถ้านักประเมินมีความรอบรู้ ประสบการณ์ และมีการฝึกฝนจนถึงขั้นมาตรฐานแล้ว จะต้องเห็นความจริงและคุณค่าของความจริงเหมือนกันหมด แต่ที่ผลออกมามีความหลากหลายเนื่องจากนักประเมินบางท่านยังไม่สามารถสร้างเครื่องมือที่มีคุณภาพและเข้าไม่ถึงศาสตร์อย่างแท้จริง นักทฤษฎีในกลุ่มนี้จึงทำการประเมินโดยเน้นวิธีการเชิงระบบ (*systematic approach*) มีการวางแผนการดำเนินงานอย่างแน่ชัด ใช้เครื่องมือมาตรฐานในการเก็บรวบรวมข้อมูล พยายามควบคุมตัวแปรแทรกซ้อนให้มากที่สุด ทำการวิเคราะห์ข้อมูลและสรุปผลตามเกณฑ์มาตรฐานที่ประกาศไว้ล่วงหน้า นักทฤษฎีการประเมินที่สามารถจัดได้ว่านิยมวิประเมินแบบปรนัยนิยม เช่น Tyler (1950), Cronbach (1963, 1980, 1982), Rossi, Freeman และ Wright (1979), Levin (1983) Chen (1990), Sanders & Horn (1994) เป็นต้น

2. ลำดับของการประเมิน : กระบวนการประเมิน

จากแนวคิดรากฐานนำไปสู่กระบวนการประเมิน พัฒนาการที่สำคัญของแนวคิดเชิงกระบวนการประเมิน ซึ่งมีอิทธิพลและให้ทิศทางต่อแนวทางการประเมินและรูปแบบของการประเมินในระยะเวลาต่อมา มีหลายกรอบแนวคิด ที่สำคัญมีดังนี้

2.1 การประเมินแบบอิงวัตถุประสงค์ (Goal – Based Evaluation)

Tyler (1940) เป็นผู้นำคนสำคัญของการประเมินอย่างเป็นระบบ ได้ริเริ่มคำว่า “การประเมินทางการศึกษา” หมายถึง กระบวนการวัดคุณค่าของผลสำเร็จตามวัตถุประสงค์อันเกิดจากการเรียนการสอน Tyler ซึ่งให้เห็นความแตกต่างระหว่าง “การวัด” และ “การประเมิน” โดยการประเมินเป็นการตรวจสอบความสอดคล้องระหว่างผลที่เกิดกับวัตถุประสงค์หรือผลลัพธ์ที่คาดหวัง หรือเรียกว่า การประเมินแบบอิงวัตถุประสงค์ (GBE) ซึ่งน่าจะเป็นแนวคิดที่มีอิทธิพลสูงสุดต่อแนวทางการประเมินที่อยู่ในปัจจุบัน จากแนวคิดดังกล่าวมีการพัฒนาสู่ การวัดแบบอิงวัตถุประสงค์ (Goal – Based Evaluation) การประเมินแบบอิงเกณฑ์ (Criterion - Referenced Evaluation) เป็นต้น

2.2 การประเมินแบบไม่อิงวัตถุประสงค์ (Goal-Free Evaluation)

Scriven (1973) นักปราชญ์ทางการประเมิน ได้เสนอแนวคิดของการประเมินที่เรียกว่า การประเมินแบบไม่อิงวัตถุประสงค์ (GFE) ซึ่งเป็นรูปแบบที่พยายามลดความลำเอียงของการประเมินที่เน้นการประเมินผลที่เกิดขึ้นจริง (*actual effects*) หรือผลทั้งหมดของโครงการทั้งในแง่ผลที่คาดหมาย และผลที่ไม่ได้คาดหมาย เช่น ผลข้างเคียง ซึ่งอาจเป็นไปได้ทั้งทางบวกและทางลบ เป็นต้น Scriven เสนอให้มีการประเมินความต้องการจำเป็น (*Need Assessment*) เพื่อเป็นเป็นเกณฑ์หนึ่งในการตัดสินคุณค่าของโครงการ การประเมินโครงการแบบไม่อิงวัตถุประสงค์ของโครงการมีจุดมุ่งหมายเพื่อทราบผลทั้งหมดของโครงการ เปรียบเทียบการดำเนินโครงการกับโครงการอื่นๆ เพื่อศึกษาประสิทธิภาพของการดำเนินงาน วิเคราะห์ค่าใช้จ่ายเพื่อศึกษาอัตราส่วนของค่าใช้จ่ายกับประสิทธิผล จากนั้นจึงสรุปผลการประเมินบุคลากร การดำเนินงาน และคุณค่าของสรุปของโครงการ

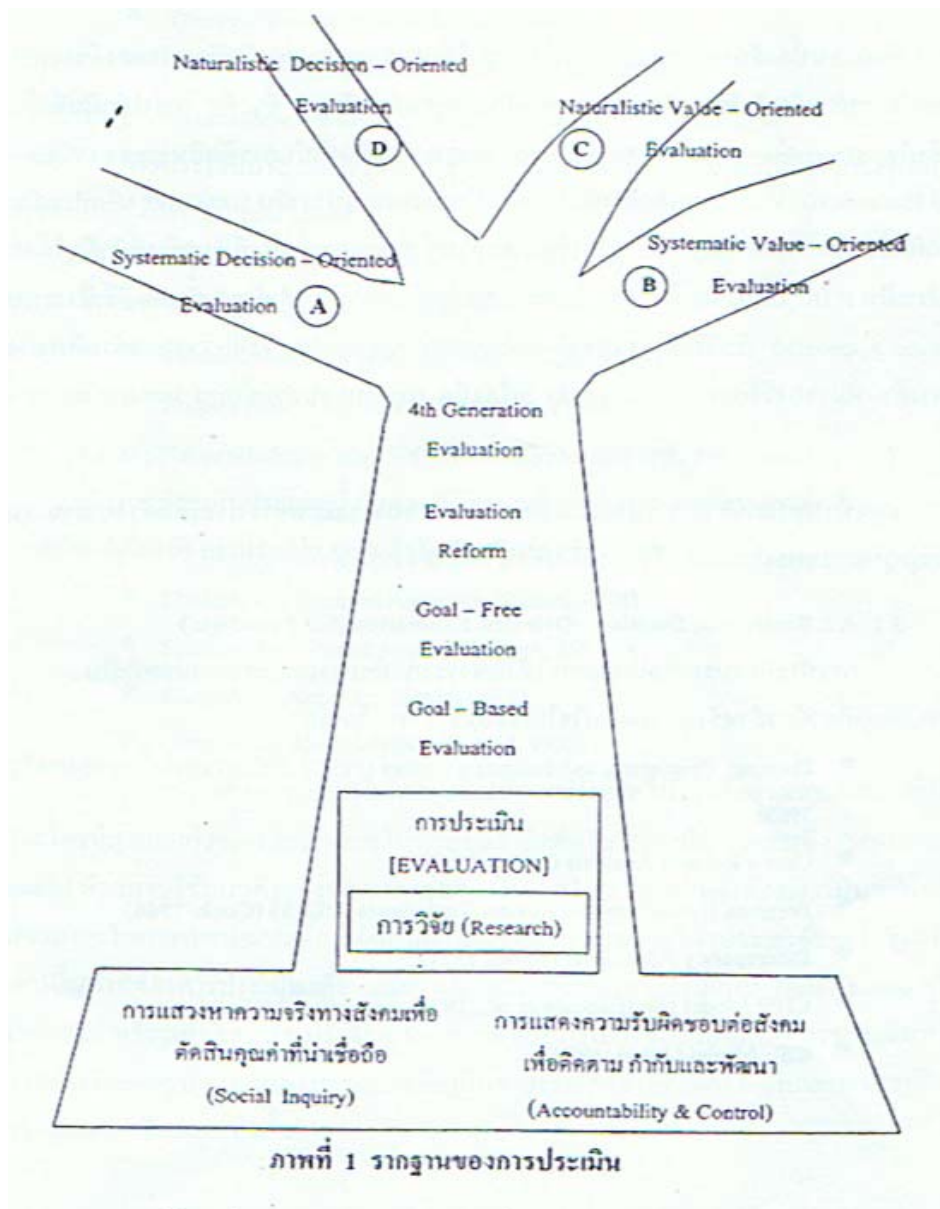
2.3 การปฏิรูปการประเมิน

Cronbach และคณะ (1981) นักจิตวิทยาทางจิตวิทยาคนสำคัญ ผู้พัฒนาสูตรการคำนวณค่าความเที่ยงด้วยสัมประสิทธิ์ α และทฤษฎีการสรุปอ้างอิงความเที่ยง (*G - Theory*) ได้ตั้งอาศรมความคิดทางการประเมินที่มหาวิทยาลัย Stanford ในช่วงปี ค.ศ. 1974 โดยระดมความคิดของคณาจารย์สาขาวิชาต่างๆและนักศึกษา ได้ข้อสรุปว่า การประเมินในยุคที่ผ่านมา มีการตั้งโจทย์ที่ไม่สอดคล้องกับความจริงถูกใช้เป็นเครื่องมือของการตรวจสอบที่กระจายอำนาจ ใช้เทคนิคการประเมินที่เป็นอุดมคติขาดการศึกษาเชื่อมโยงกับนโยบาย ผลการประเมินขาดการวิพากษ์กลั่นกรองและขาดความเป็นเหตุเป็นผล Cronbach และคณะ ได้เสนอแนวคิดของการปฏิรูปการประเมินว่า การประเมินจะไร้น้ำหนัก ถ้าการประเมินสนใจเพียงการบรรลุจุดมุ่งหมายเป็นเพียงส่วนหนึ่งของคำหุที่มักใช้ทางการเมือง แต่ความเป็นจริงทางสังคมมีจุดมุ่งหมายที่กว้างขวางกว่า คำถามการประเมินที่ดี อาจมีคำตอบที่แตกต่างกันหลายแนวทาง การเก็บรวบรวมโดยตรงที่เข้มงวดเกินไป อาจได้ภาพที่บิดเบือน ควรเก็บทั้งข้อเท็จจริงและคำวิพากษ์ ผู้ประเมินย่อมหนีไม่พ้นการได้รับอิทธิพลทางการเมือง ผู้ประเมินจึงต้องเข้าใจบริบท สร้างความสมานฉันท์กับทุกฝ่ายที่เกี่ยวข้อง และรู้วิธีเปิดเผยความจริง ข้อสรุปของผู้ประเมินมิใช่เป็นการชี้ถูก/ผิด ทางการเมือง เนื่องจากสภาพธรรมชาติของสังคมที่เกี่ยวข้องกับคุณค่าพหุมติ การประเมินจะช่วยให้ได้สารสนเทศที่สนับสนุนการเจรจาต่อรองมากกว่าการชี้ถูก/ผิด การประเมินโครงการจึงเป็นกระบวนการที่ช่วยให้สังคมเรียนรู้ตัวเอง ช่วยให้ทางเลือกสำหรับกำหนดแผนการดำเนินงาน ช่วยให้เห็นวิวัฒนาการของปัญหาสังคม และการปรับปรุงพัฒนาให้ดียิ่งขึ้น

นอกจากนี้ *Stufflebeam* และคณะ (1971) ยังได้เสนอประโยคที่เป็นอมตะทางการประเมินไว้ว่า
The Purpose of Evaluation is Not to Prove but to Improve.

2.4 การประเมินยุคที่ 4 (Fourth Generation Evaluation)

Guba และ Lincoln (1989) ได้เสนอการจัดยุคของการประเมินไว้ดังนี้ 1) ยุคการประเมินที่เน้นกระบวนการวัด (Measurement - Oriented) 2) ยุคการประเมินที่เน้นการบรรยายความสำเร็จตามวัตถุประสงค์ (Objective - Oriented) 3) ยุคการประเมินที่เน้นการตัดสิน (Judgment - Oriented) และ 4) ยุคการประเมินที่เน้นการเจรจาต่อรอง (Negotiation - Oriented) ในยุคที่ 4 นี้ การประเมินเป็นกระบวนการสนองตอบต่อข้อเรียกร้อง / ความวิตกกังวลของผู้มีส่วนเกี่ยวข้องทุกฝ่าย โดยการสร้างการรับรู้ความจริงร่วมกัน และสร้างการยอมรับผลการประเมินที่พิจารณาคุณค่าหลายด้านร่วมกัน ผู้ประเมินจึงบทบาทเป็นผู้นำในการอำนวยความสะดวกและประสานการเจรจาผู้ที่เกี่ยวข้องทั้งการเตรียมประเด็นการเจรจา การเก็บรวบรวมข้อมูล และการวิเคราะห์ การจัดทำรายงานแบบกรณีศึกษาที่ผู้เกี่ยวข้องสร้างผลการประเมินร่วมกัน



3. กิ่งก้านสาขาของการประเมิน : กลุ่มรูปแบบของการประเมิน

จากแนวคิดรากฐานของการประเมิน ผ่านแนวทางการประเมินสู่การสร้างโมเดล /รูปแบบของการประเมิน นักทฤษฎีการประเมินมักมีปรัชญาการประเมินที่แตกต่างกัน พื้นฐานความแตกต่างที่สำคัญอยู่ที่ข้อตกลงเบื้องต้นของการประเมิน กับมาตรการเข้าถึงคุณค่าของสิ่งที่ประเมินยังผลให้มีการสร้างโมเดล /รูปแบบของการประเมินหรือแนวทางการประเมินในลักษณะที่แตกต่างกัน

ข้อตกลงเบื้องต้นของการประเมินที่สำคัญ ได้แก่ ข้อตกลงเกี่ยวกับวัตถุประสงค์ของการประเมิน หรือ “ประเมินทำไม?” โดยสามารถจัดเป็นแนวคิดที่สำคัญ 2 ข้อ คือ การประเมินที่เน้น การตัดสินใจ (Decision – oriented Evaluation) กับ การประเมินที่เน้นการตัดสินคุณค่า (Value – oriented Evaluation) ซึ่งมีความแตกต่างกันทางด้านนิมของการประเมิน บทบาทของนักประเมิน และผู้รับผิดชอบในการตัดสินคุณค่า สำหรับมาตรการเข้าถึงคุณค่าของสิ่งที่ประเมินที่สำคัญได้แก่ วิธีการประเมิน หรือ “ประเมินอย่างไร?” โดยสามารถจัดเป็นแนวคิดที่สำคัญ 2 ข้อ คือ วิธีเชิงระบบ (Systematic Approach) กับวิธีเชิงธรรมชาติ (Naturalistic Approach) ซึ่งมีความแตกต่างกันอย่างสำคัญทางด้านที่มาของวิธีการ การมองคุณค่า เครื่องมือ การเก็บรวบรวมข้อมูล และการวิเคราะห์ข้อมูล

จากเป้าหมายและวิธีการประเมินที่ต่างกันของรูปแบบการประเมินต่างๆสามารถจัดเป็นกลุ่มรูปแบบการประเมิน หรือกระบวนการทัศน์ได้เป็น 4 กลุ่ม ดังนี้

3.1 A : Systematic Decision - Oriented Evaluation (SD Paradigm)

การประเมินประเภทนี้เน้นการใช้วิธีเชิงระบบเพื่อการเสนอสารสนเทศที่เป็นประโยชน์ต่อผู้เกี่ยวข้องสำหรับการตัดสินใจในเชิงบริหาร เช่น โมเดล

- Planning, Programming and Budgeting System (PPBS) (กระทรวงกลาโหมของสหรัฐ, 1965)
- Cost - Related Analysis (Levin, 1983)
- Program Evaluation and Review Techniques (PERT) (Cook, 1966)
- Discrepancy Approach (Provus, 1971)
- CIPP Model (Stufflebeam et. al., 1971)
- CSE Model (Alkin, 1979)

3.2 B : Systematic Value - Oriented Evaluation (SV Paradigm)

การประเมินประเภทนี้เน้นการใช้วิธีเชิงระบบเพื่อให้นักประเมินทำการตัดสินคุณค่าของสิ่งที่มุ่งประเมิน เช่น โมเดล

- Consumer - Oriented Approach (Scriven, 1967)
- Judicial (Adversarial) Approach (Owens, 1973)
- Accreditation (Professional Review) Approach
- Theory - Driven Approach (Chen, 1990)
- Value – Added Approach (Sander & Horn, 1994)

3.3 C : Naturalistic Value – Oriented Evaluation (NV Paradigm)

การประเมินประเภทนี้เน้นการใช้วิธีการเชิงธรรมชาติเพื่อให้นักประเมินทำการตัดสินคุณค่าของสิ่งที่มุ่งประเมิน เช่น โมเดล

- Criticism Approach (Eisner, 1975)
- Authentic Approach (Cradler, 1997)
- Constructivist Approach (Lincoln & Guba, 1985)
- Empowerment Approach (Felterman, 1994)

3.3 D : Naturalistic Decision – Oriented Evaluation (ND Paradigm)

การประเมินประเภทนี้เน้นการใช้วิธีธรรมชาติเพื่อการเสนอสารสนเทศที่เป็นประโยชน์ต่อผู้เกี่ยวข้องสำหรับการตัดสินใจในเชิงบริหาร เช่น โมเดล

- Utilization - Focused Approach (Pation, 1978)
- Stakeholder - Based Approach (Bryk, 1983)
- Responsive Approach (Stake, 1978)
- Democratic Approach (Mc Donald , 1975)

ถ้านักประเมินต้องการเลือกโมเดล/รูปแบบการประเมินไปใช้ได้เหมาะสม เพื่อให้เกิดประโยชน์สูงสุดแล้ว นักประเมินอาจไม่จำเป็นต้องยึดติดกับโมเดลใดโมเดลหนึ่ง (Model – Free Approach) ด้วยการใช้แนวคิดเชิงบูรณาการสร้างโมเดลใหม่ขึ้นมา โดยนักประเมินจะต้องมีความสามารถในการพัฒนาเกณฑ์การตัดสินความสอดคล้องระหว่างวัตถุประสงค์ของโมเดล กับประเด็นปัญหาที่ต้องการประเมิน เพื่อตัดสินใจเลือกแนวทางการประเมินแบบ Goal – Based หรือในการสร้างโมเดล/รูปแบบของการประเมินที่อยู่ในข่ายจะเลือกมาใช้ เพื่อออกแบบการประเมินที่แก้ไขจุดอ่อนของโมเดลเหล่านั้น

4. สรุป

การประเมินตั้งอยู่บนรากฐานของสองหลักการที่สำคัญ ได้แก่ หลักการสืบเสาะแสวงหาความจริงทางสังคมที่น่าเชื่อถือ เพื่อตัดสินคุณค่าของสิ่งต่างๆตามบริบทของสังคมนั้น และหลักความรับผิดชอบต่อสังคม เพื่อติดตามกำกับ และพัฒนาสังคม

การประเมินจึงเป็นกระบวนการสร้างสรรค์สังคมที่สอดคล้องกับวิถีทางประชาธิปไตย โดยมีเป้าหมายของการใช้กระบวนการตัดสินคุณค่าที่โปร่งใส ยุติธรรมและใช้สารสนเทศให้เกิดประโยชน์ต่อการปรับปรุง เปลี่ยนแปลง และพัฒนาสู่เป้าหมายที่พึงประสงค์

บรรณานุกรม

ศิริชัย กาญจนวาสี. (2547). **ทฤษฎีการประเมิน**. (พิมพ์ครั้งที่ 4). กทม.: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

Alkin, M.C. (Ed.) (2004). **Evaluation Roots : Tracing Theorists Views and Influences**. London : Sage Publications.

Cronbach, L. J. et. al. (1980). **Toward Reform of Program Evaluation**. California : Jossey – Bass Publishers.

Guba, E. G., and Lincoln, Y.S.(1989). **Fourth Generation Evaluation**. London : Sage Publications.

Maduas, M.S., Scriven, M., and Stufflebeam, D.L.(1983). **Evaluation Models**. Boston : Kluawer – Nijhoff.

Scriven, M. (2005). **The Logic and Methodology of Checklists**. Retrieved from <http://www.wmich.edu/evalctr/checklist/checklistmenee.htm>.

Stake, R.E. (Ed.)(1975). **Evaluating the Arts in Education : A Responsive Approach**. Ohio : Charles E. Merrile.

Stufflebeam, D.L.(2001). **Evaluation Models**. San Francisco : Jossey – Bass.

Tyler, R.W. (1956). **Basic Principles of Curriculum and Instruction**. Chicago : University of Chicago Press.

Worthen, B.R., and Sanders, J.R. (1973). **Educational Evaluation : Theory and Practice**. California : Wadsworth Publishing company, Inc.