

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้อง ผู้วิจัยนำเสนอแบ่งเป็น 4 ตอน คือ ตอนที่ 1 การประเมินเพิ่มผลงาน ตอนที่ 2 การประเมินอภิमान ตอนที่ 3 งานวิจัยที่เกี่ยวข้อง และตอนที่ 4 กรอบการพัฒนาการประเมินเพิ่มผลงานและการใช้ประโยชน์จากผลการประเมินเพิ่มผลงาน แต่ละตอนมีรายละเอียดดังนี้

ตอนที่ 1 การประเมินเพิ่มผลงาน (Portfolio Evaluation)

การประเมิน (evaluation) หรือการประเมินค่า (assessment) ที่ใช้กันมาแต่เดิม เน้นการประเมินโดยใช้แบบสอบมาตรฐาน (standardized tests) เป็นหลัก ต่อมาในช่วงปลายคริสต์ทศวรรษที่ 1980 นักการศึกษาเริ่มเห็นข้อจำกัดของการประเมินโดยใช้แบบสอบมาตรฐาน สรุปได้ 5 ประการ ประการแรกเป็นการประเมินที่ผลการประเมินขึ้นอยู่กับผู้ประเมิน ประการที่สองการให้คะแนนการประเมินส่วนใหญ่ ให้คะแนนแต่ละชิ้นงานโดยไม่พิจารณาภาพรวม ประการที่สามผู้ประเมินให้ความสนใจเฉพาะแบบสอบมาตรฐาน โดยไม่คำนึงถึงผลการปฏิบัติงานด้านอื่นของผู้รับการประเมิน ประการที่สี่เป็นการประเมินที่ใช้เกณฑ์มาตรฐานที่กำหนดไว้ก่อนแล้ว โดยผู้รับการประเมินไม่มีส่วนร่วมในการกำหนดเกณฑ์ และประการสุดท้ายเป็นการประเมินที่ให้ผลไม่เพียงพอกับความต้องการของผู้ประเมิน (Moss, Schutz และ Collins, 1998; Farr และ Tone, 1994) นอกจากนี้การประเมินโดยใช้แบบสอบมาตรฐานยังมีส่วนทำให้เกิดผลเสียหลายในการเรียนการสอน เนื่องจากผู้สอนให้ความสำคัญกับการสอนตามแนวข้อสอบมากกว่าการสอนตามหลักสูตร ด้วยเหตุนี้ทางการศึกษาจึงมุ่งพัฒนาวิธีการประเมินใหม่โดยไม่ยึดติดกับแบบสอบมาตรฐาน แต่ขยายขอบเขตของเครื่องมือการประเมินให้มีความหลากหลาย ตั้งแต่การสังเกต การสัมภาษณ์ การรายงาน การปฏิบัติ ไปจนถึงการแสดงนิทรรศการ ผลจากการพัฒนาดังกล่าวทำให้เกิดการปฏิวัติการประเมินค่า (assessment revolution) เป็นผลให้มีการประเมินแบบต่างๆ หลากหลายในคริสต์ทศวรรษที่ 1990 เช่น การประเมินเพิ่มผลงาน (portfolio evaluation) การประเมินค่าที่แท้จริง (authentic assessment) การประเมินค่าเชิงสร้างสรรค์ (constructive assessment) การประเมินค่าทางตรง (direct assessment) การประเมินค่าโดยรวม (holistic assessment) เป็นต้น (Hart, 1994; Oosterhof, 1994; อุทุมพร จามรมาน, 2540)

การปฏิวัติการประเมินหรือการปฏิวัติการประเมินค่าที่ทำให้มีการพัฒนาเป็นการประเมินเพิ่มผลงานนั้น เริ่มต้นจากนักวัดและประเมินผลการศึกษาพิจารณาเห็นว่า การประเมินผลสัมฤทธิ์ทางการเรียนของผู้เรียนโดยใช้แบบทดสอบชนิดเขียนตอบ (written tests) ไม่สามารถวัดทักษะหลายชนิดได้โดยตรง เช่น ทักษะการพูด การแสดงออกทางกาย การปฏิบัติ การทดลองทางวิทยาศาสตร์ การเล่นดนตรี การแสดงอารมณ์ และทักษะทางสังคมอื่นๆ จึงได้มีความพยายามพัฒนาการประเมินที่เป็นการสังเกตและการประเมิน

พฤติกรรมโดยตรง อันเป็นที่มาของการประเมินการปฏิบัติงาน (performance evaluation) Oosterhof (1994) สรุปลักษณะการประเมินการปฏิบัติงานไว้ 3 ประการ คือ

ก. การประเมินการปฏิบัติงานเป็นการประเมินพฤติกรรม (behaviors) หรือผลลัพธ์ (outcomes) ของพฤติกรรม หรือกระบวนการ (process) หรือผลผลิต (products) จากกระบวนการก็ได้ สภาพในการประเมินอาจเป็นสภาพที่เป็นจริง (real settings) หรือสภาพที่จำลอง (simulated setting) ก็ได้ สิ่งเร้าในการประเมินอาจเป็นสิ่งเร้าธรรมชาติ (natural stimuli) หรือสิ่งเร้าที่มีการสร้างขึ้นมา (structured stimuli) ก็ได้ ทั้งนี้สิ่งที่จะประเมินต้องมีการกำหนดรายละเอียดอย่างเฉพาะเจาะจง (specific) เช่น กำหนดพฤติกรรมในการประเมินการว่ายน้ำ มิใช่กำหนดว่าผู้เรียนว่ายน้ำได้ แต่ต้องกำหนดพฤติกรรมเฉพาะเจาะจง เช่น ผู้เรียนสามารถลอยตัวในน้ำได้โดยไม่ใช้เครื่องพยุงตัวอย่างน้อย 5 นาที

ข. ในการประเมินการปฏิบัติงานต้องตัดสินได้ว่า การปฏิบัติงานหรือพฤติกรรมของผู้เรียนแบบใด มีความเหมาะสมเมื่อเทียบกับการปฏิบัติงานหรือพฤติกรรมของผู้เรียนแบบอื่นๆ กล่าวอีกอย่างหนึ่งคือต้องมีเกณฑ์ในการตัดสินคุณค่าพฤติกรรมหรือการปฏิบัติงานที่ต้องการประเมิน

ค. พฤติกรรมหรือการปฏิบัติงานที่ประเมินต้องเป็นสิ่งที่ไม่สามารถประเมินได้โดยตรง ด้วยแบบสอบถามชนิดเขียนตอบ

จากลักษณะการประเมินการปฏิบัติงานดังกล่าวข้างต้น Oosterhof (1994) สรุปว่า การประเมินการปฏิบัติงานค่อนข้างยุ่งยาก สิ้นเปลืองงบประมาณและใช้เวลานานมาก ในทางปฏิบัติถ้าประเมินไม่สามารถทำการประเมินพฤติกรรมหรือทักษะหลายๆทักษะพร้อมกันได้ แต่ผลการประเมินการปฏิบัติงานที่ได้จะมีความถูกต้องและมีความตรงสูงกว่าการประเมินโดยใช้แบบทดสอบทั่วไป ความยุ่งยากในการประเมินการปฏิบัติงานอยู่ที่ต้องมีการกำหนดวัตถุประสงค์การประเมิน เตรียมรายละเอียดของพฤติกรรมหรือทักษะที่จะประเมิน จัดสภาพการประเมิน เตรียมคำสั่งให้ผู้รับการประเมินดำเนินการ จัดทำคำแนะนำสำหรับผู้ประเมินและกำหนดแผนการให้คะแนนในการประเมิน

เนื่องจากการประเมินการปฏิบัติงานเป็นการประเมินพฤติกรรมหรือกระบวนการหรือการปฏิบัติงาน ทั้งที่เป็นการปฏิบัติงานในชีวิตจริง หรือการปฏิบัติงานในสภาพที่จำลองมาจากสภาพจริง หรือการปฏิบัติงานในสถานการณ์ที่มีสิ่งเร้าที่จะกระตุ้นให้เกิดพฤติกรรมที่แสดงถึงศักยภาพของผู้เรียน Oosterhof (1994) และ Hart (1994) จึงสรุปว่า การประเมินค่าที่แท้จริง (authentic evaluation) ทุกประเภท เป็นการประเมินการปฏิบัติงาน เพราะการประเมินค่าที่แท้จริง หมายถึง การประเมินค่ากระบวนการทำงานทางสมอง และทางจิตใจ ของผู้รับการประเมินโดยใช้การวัดโดยตรง (direct measure) และการสะท้อน (reflection) ความรู้สึก ความคิด จากการที่ผู้รับการประเมินปฏิบัติงานตามสภาพที่เป็นจริง และเป็นการปฏิบัติงานที่มีคุณค่าในตัว กิจกรรมที่ใช้ในการประเมินมีหลากหลาย เช่น การสังเกต การสัมภาษณ์ การแก้ปัญหาเป็นกลุ่ม การจัดทำแฟ้มผลงาน การเขียนเรียงความ เป็นต้น ทั้งนี้ผู้รับการประเมินต้องทราบเกณฑ์ในการประเมินชัดเจนตั้งแต่ต้นว่า การปฏิบัติงานระดับใดมีความถูกต้องสมบูรณ์ ลักษณะของการประเมินค่าที่แท้จริงยังเป็นการกำหนด

เกณฑ์มาตรฐานในการประเมินร่วมกันระหว่างผู้ประเมินและผู้รับการประเมิน แทนที่จะเป็นการสร้างเครื่องมือประเมินให้เป็นมาตรฐานตามแนวคิดในการประเมินแบบเดิม โดยทั่วไปการประเมินค่าที่แท้จริงจะถูกบูรณาการเข้ากับหลักสูตร มิใช่เป็นแต่เพียงขั้นตอนการประเมินอันเป็นขั้นตอนหนึ่งในกระบวนการเรียนการสอน นอกจากนี้จุดมุ่งหมายสำคัญของการประเมินที่แท้จริงให้ความสำคัญกับการแสดงศักยภาพที่แท้จริงและเด่นของผู้รับการประเมินมากกว่าการวินิจฉัยข้อบกพร่อง และให้ความสำคัญกับการกระตุ้นให้ผู้รับการประเมินมีการพัฒนาตนเองได้สูงสุด

Hart (1994), Oosterhof (1994), อุทุมพร จามรมาน (2540) อธิบายว่า การประเมินแฟ้มผลงานเป็นรูปแบบหนึ่งของการประเมินค่าที่แท้จริง ที่เป็นกระจกเงาสท้อน (mirror) และวัดผลการปฏิบัติงานของผู้รับการประเมิน จากหลักฐานในแฟ้มผลงาน (portfolio) อันประกอบด้วยสารสนเทศ (information) และตัวอย่างผลงาน (work samples) ที่ผู้รับการประเมินจัดเตรียมไว้เพื่อแสดงให้เห็นถึงความก้าวหน้า และความสำเร็จของตน

การนำเสนอเนื้อหาสาระจากรายงานเอกสารที่เกี่ยวข้องกับการวิจัยนี้ เป็นการเสนอเนื้อหาสาระเกี่ยวกับหลักการ วิธีการ ประโยชน์และ ปัญหาในการประเมินแฟ้มผลงาน โดยผู้วิจัยจัดเตรียมเรียงแยกเป็น 5 หัวข้อ คือ ความหมายของการประเมินแฟ้มผลงาน ทฤษฎีและหลักการในการประเมินแฟ้มผลงาน ประเภทของแฟ้มผลงาน ขั้นตอนการประเมินแฟ้มผลงาน ประโยชน์ของการประเมินแฟ้มผลงาน ดังรายละเอียด แต่ละตอนต่อไปนี้

1.1 ความหมายของการประเมินแฟ้มผลงาน

จากการศึกษาเอกสาร และรายงานวิจัยที่เกี่ยวข้อง ผู้วิจัยพบว่านักประเมินให้คำนิยาม หรือความหมายของการประเมินแฟ้มผลงาน เป็น 2 แนวทาง แนวทางแรก นักประเมินให้ความหมายของคำว่า แฟ้มผลงาน (portfolio) เป็นคำเดียวโดดๆ แนวทางที่สอง นักประเมินให้ความหมายของคำว่า การประเมินแฟ้มผลงาน (portfolio evaluation) หรือ การประเมินค่าแฟ้มผลงาน (portfolio assessment) ผู้วิจัยนำมารวบรวมไว้ดังต่อไปนี้

ในกลุ่มแรก Farr และ Tone (1994) เสนอว่า ความหมายของแฟ้มผลงานหมายถึง คลังหรือแฟ้มที่ใช้บรรจุสิ่งของ ในทางธุรกิจ คำว่า แฟ้มผลงาน เดิมหมายถึง แฟ้มที่บรรจุหุ้้น พันธบัตร และเอกสารการลงทุนของบุคคลในช่วงเวลาหนึ่ง ที่แสดงให้เห็นสถานะของนักลงทุนในสายตาของเพื่อนนักลงทุน ในทางการเมืองความหมายของแฟ้มผลงาน ยังรวมถึงประสบการณ์และคุณวุฒิที่จะใช้ในการรับรองคุณภาพของบุคคลสำหรับการเสนอชื่อให้เข้ารับตำแหน่งสำคัญ จากความหมายดั้งเดิมดังกล่าว นักประเมินทางการศึกษา จึงนำศัพท์การประเมินแฟ้มผลงานมาใช้ในความหมายของการรวบรวมหลักฐานอย่างมีระบบเพื่อใช้ในการประเมินสถานภาพของบุคคล ในระยะแรกแฟ้มผลงานถูกใช้เป็นเทคนิคในการประเมินสำหรับนักศึกษาผู้

ใหญ่ที่กลับเข้ามาศึกษาต่อเพื่อเป็นหลักฐานในการสนองตอบต่อคำสั่งในการปฏิบัติงาน และเพื่อแสดงว่ามี ความรู้ ความสามารถเพียงพอที่จะสำเร็จการศึกษา

นอกจากนี้นักประเมินอีกหลายคนให้ความหมายของแฟ้มผลงาน (portfolio) ไว้คล้ายคลึงกัน ผู้วิจัยได้สรุปความหมายของแฟ้มผลงานตามแนวคิดของ Meyer, Sehuman และ Angello (1990 อ้างใน Reckas, 1995) Arter และ Spandel (1992 อ้างใน Coleman, 1996) และ Paulson และ Meyer (1991 อ้างใน Burke, Fogarty และ Belgrad, 1994) ซึ่งสรุปว่า แฟ้มผลงาน คือ ผลการรวบรวมชิ้นงาน หรือผลงานของผู้ทำแฟ้มผลงานอย่างมีจุดมุ่งหมาย เพื่อแสดงถึงความพยายาม ความก้าวหน้าและผลสัมฤทธิ์ทางการเรียนในสาขาที่กำหนด ชิ้นงานในแฟ้มผลงานประกอบด้วยรายงานเกี่ยวกับการมีส่วนร่วมในการเลือก ชิ้นงานที่จะบรรจุในแฟ้มผลงาน เกณฑ์ในการคัดเลือกชิ้นงาน เกณฑ์ในการตัดสินคุณค่าของชิ้นงาน วัสดุและ หลักฐานที่สะท้อนความคิดของผู้ทำแฟ้มผลงาน

อุทุมพร จามรมาน (2540) และ Oosterhof (1994) สรุปว่า แฟ้มผลงานมีลักษณะเป็นแฟ้มบรรจุ หลักฐานต่าง ๆ เช่น งานเขียน วัสดุที่ประดิษฐ์ขึ้น สารทางเคมี ฯลฯ ซึ่งเป็นเครื่องมือชนิดหนึ่งที่นักประเมินใช้ในการประเมินค่าที่แท้จริง อุทุมพร จามรมาน (2540) ส่วน Oosterhof (1994) อธิบายเพิ่มเติมว่า แฟ้มผลงานมีหลายประเภท ขึ้นอยู่กับวัตถุประสงค์ในการประเมิน และผู้ประเมินจะต้องพิจารณาคัดเลือกผลงานลง แฟ้มต่างกันตามประเภทของแฟ้มผลงาน

Hart (1994) กล่าวว่า แฟ้มผลงาน คือที่รวบรวมหลักฐานซึ่งแสดงให้เห็นถึงความสามารถทักษะ ความคิด ความสนใจและความสำเร็จในการทำงาน

Anderson และ Bachor (1998) ให้ความหมายของแฟ้มผลงานว่าเป็นผลการรวบรวมผลงานของ นักเรียนซึ่งจะแสดงถึงผลสัมฤทธิ์ทางการเรียน โดยการรวมนั้นเป็นการทำงานร่วมกันของครู นักเรียน เพื่อให้ได้หลักฐานที่แสดงถึงการเรียนรู้ นอกจากนี้ยังต้องมีการสะท้อนความรู้สึกลงและการประเมินตนเองของ นักเรียนด้วย

โดยสรุปแล้วการให้ความหมายของแฟ้มผลงาน จะเห็นว่าเน้นในเรื่องผลการเก็บรวบรวมชิ้นงานเพื่อ เป็นหลักฐานสำหรับการประเมินค่าความสามารถของบุคคลที่เป็นผู้ทำแฟ้มผลงาน โดยใช้หลักฐานชิ้นงานที่ รวบรวมไว้นั้นเป็นสิ่งยืนยัน นอกจากนี้แฟ้มผลงานยังแสดงให้เห็นถึงความสามารถและพัฒนาการของผู้ทำ แฟ้มผลงานในการปฏิบัติงานต่างๆ ทั้งในเชิงวิชาการและเชิงวิชาชีพ ดังนั้น แฟ้มผลงานจึงมีความหมายว่า แฟ้ม หรือผลการเก็บรวบรวมชิ้นงาน

ในกลุ่มที่สองนักประเมินให้ความหมายของคำว่าประเมินแฟ้มผลงานไว้หลายท่าน ผู้วิจัยนำเสนอ ความหมายที่กำหนดโดยนักประเมินดังต่อไปนี้

Elton และ Gruber (1987) ซึ่งเป็นนักวิชาการทางธุรกิจ ให้นิยามการประเมินแฟ้มผลงานว่าเป็น กระบวนการประเมินที่เปลี่ยนไปจากแบบเดิมซึ่งเป็นการคำนวณหรือการประมาณค่าผลตอบแทนอย่างหยาบ หรือการประเมินค่าว่านักลงทุนประสบผลสำเร็จในการลงทุนมากน้อยเพียงใด เมื่อเทียบกับนักลงทุนคนอื่น

มาเป็นการประเมินโดยการสำรวจผลตอบแทนและความเสี่ยงอย่างละเอียด รวมทั้งการศึกษาถึงสาเหตุที่ทำให้ได้ผลตอบแทนดังกล่าว การประเมินเพิ่มผลงานให้ความสำคัญกับการทำความเข้าใจสาเหตุที่ทำให้พนักงานได้รับผลประโยชน์ตอบแทนสูงว่าเนื่องมาจากการตลาดหรือการเลือกสต็อกสินค้าที่ดีกว่า เป็นต้น

Reckase (1995) ให้นิยามการประเมินเพิ่มผลงานว่าเป็นการประเมินของผู้ทำเพิ่มผลงานจากผลการปฏิบัติงานที่ต้องใช้ทักษะอันเป็นทักษะที่จะรับการประเมิน ทักษะดังกล่าวเป็นทักษะที่ใช้ในกิจกรรมการเรียนการสอนและเป็นกิจกรรมที่ต้องใช้ในชีวิตประจำวันจริง ๆ ของผู้ทำเพิ่มผลงาน กล่าวได้ว่าการประเมินเพิ่มผลงานเป็นการประเมินที่เชื่อมโยงกับการปฏิบัติงานในชีวิตจริง

LeMahieu, Gilomer และ Eresh (1995) สรุปว่า การประเมินเพิ่มผลงานคล้ายคลึงกับการประเมินผลการปฏิบัติงาน เพราะสิ่งที่ปรากฏในเพิ่มผลงานก็คือผลการปฏิบัติงานที่คัดเลือกแล้วของผู้ทำเพิ่มผลงานเพื่อเป็นหลักฐานยืนยันผลการปฏิบัติงาน และเนื่องจากเพิ่มผลงานมีใช้ในวงการศึกษาอย่างแพร่หลาย ดังนั้นเพิ่มผลงานจึงมีส่วนที่เป็นประโยชน์ต่อการพัฒนาการศึกษาทั้งในด้านประโยชน์ในการประเมินและประโยชน์ในฐานะเป็นเครื่องมือช่วยพัฒนาให้ผู้ทำเพิ่มผลงานเกิดการเรียนรู้

Doner และ Grilman (1998) ให้ความหมายกระบวนการทำเพิ่มผลงาน (Development Portfolio Process) เพื่อรับการประเมินว่า เป็นกระบวนการประเมินการปฏิบัติงานที่แท้จริง ซึ่งนอกจากจะช่วยให้ผู้ประเมินได้ประเมินผลงานของผู้ทำเพิ่มผลงานแล้ว ยังช่วยพัฒนาทักษะความรู้ ความเข้าใจเกี่ยวกับเนื้อหาสาระที่เป็นผลงานของผู้ทำเพิ่มผลงานเป็นอย่างมาก

Koretz (1998) ให้คำจำกัดความของการประเมินเพิ่มผลงานไว้คล้ายคลึงกับ Doner และ Grilman (1998) ว่าเป็นการประเมินการปฏิบัติ ซึ่งเป็นงานที่ถูกสะสมและคัดเลือกโดยนักเรียน การประเมินเพิ่มผลงานใช้เป็นส่วนหนึ่งของการประเมินผลสัมฤทธิ์ทางการเรียน และอธิบายเพิ่มเติมว่าการคัดเลือกชิ้นงานของนักเรียนนั้นมีจุดมุ่งหมายในเรื่องของคะแนน ในการคัดเลือกชิ้นงานเพื่อเป็นคะแนนเรียกว่าการประเมินเพิ่มผลงานนั่นเอง

Stecher (1998) กล่าวว่าเพิ่มผลงานและการประเมินเพิ่มผลงานในเชิงวิชาการมีความคล้ายคลึงกัน คือมีความคลุมเครือและไม่ชัดเจนเหมือนกันในประเด็นของจุดมุ่งหมายการดำเนินการ การใช้ประโยชน์และการประยุกต์ใช้เกณฑ์สำหรับการประเมินในหน่วยงานขนาดใหญ่ การประเมินเพิ่มผลงานส่วนมากจะเกี่ยวข้องในงานศิลปะ

จากความหมายของคำว่าประเมินเพิ่มผลงานที่นักประเมินหลาย ๆ ท่านได้ให้ไว้ข้างต้น สรุปได้ว่า การประเมินเพิ่มผลงานเป็นกระบวนการประเมินความสามารถของผู้ทำเพิ่มผลงานตามจุดมุ่งหมายที่กำหนดในการทำเพิ่มผลงาน ซึ่งอาจเป็นทักษะการปฏิบัติงานด้านต่าง ๆ นอกจากนี้ยังสามารถประเมินพฤติกรรมอื่น ๆ ได้ด้วย เช่น ความมานะพยายาม การประเมินตนเอง ซึ่งการประเมินความสามารถดังกล่าวประเมินจากชิ้นงานที่คัดเลือกโดยผู้ทำเพิ่มผลงาน

นอกจากจะได้ผลการประเมินเพื่อประเมินความสามารถและพัฒนาการของผู้ทำเพิ่มผลงานแล้ว การประเมินดังกล่าวจะทำให้ผู้ประเมินได้รับประโยชน์จากการประเมินด้วย กล่าวคือในเพิ่มผลงานนักเรียน มีการเขียนบรรยาย/อธิบาย รวมถึงการสะท้อนถึงตนเองของผู้ทำเพิ่มผลงาน ซึ่งอาจเป็นข้อสังเกตเกี่ยวกับ ปัญหา คำชี้แจงในการทำเพิ่มผลงานของผู้ทำเพิ่ม ข้อความดังกล่าวจะทำให้ผู้ประเมินได้ทราบถึงเจตคติ และความรู้ของผู้ทำเพิ่มต่อวิชาหรืองานที่เขาทำ ซึ่งจะช่วยให้ง่ายต่อการเข้าใจเด็กและหาแนวทางปรับปรุงหรือ ส่งเสริมให้ตรงกับปัญหาและความต้องการที่แท้จริงของผู้ทำเพิ่มผลงาน

1.2 ทฤษฎีและหลักการในการประเมินเพิ่มผลงาน

ตามหลักการการประเมินทางการศึกษา Mehren และ Lehmann (1984) สรุปว่า ผู้สอนที่จะทำหน้าที่เป็นผู้ประเมินมีหน้าที่รับผิดชอบสำคัญ 3 ประการ ประการแรก ต้องพิจารณากำหนดว่าจะใช้สารสนเทศอะไรในการประเมิน ประการที่สองทำอย่างไรจึงจะได้รับสารสนเทศที่ถูกต้องมาใช้ในการประเมิน และประการที่สามจะตัดสินคุณค่าของสารสนเทศที่ประเมินอย่างไรและเสนอผลให้ผู้เกี่ยวข้องได้ทราบด้วยวิธีการใด จากหน้าที่ความรับผิดชอบที่สำคัญทั้ง 3 ประการ เป็นพื้นฐานการประเมินด้วยแบบสอบถามมาตรฐาน นักประเมินใช้ทฤษฎีจิตมิติ (psychometric theory) เป็นหลักในการดำเนินการประเมิน กล่าวคือ นักประเมินต้องใช้วิธีการตรวจสอบเครื่องมือวัดและแบบสอบให้มีความเที่ยง (reliable) และมีความตรง (validity) สูงสุด รวมทั้งต้องใช้มาตรฐานในการประเมินค่าคะแนนที่แท้จริง (true score) ของผู้รับการประเมิน โดยขจัดความคลาดเคลื่อน (error) เพื่อให้ได้สารสนเทศที่แท้จริงสำหรับการประเมิน

เมื่อมีการปฏิบัติการประเมินเกิดขึ้น และมีการพัฒนาการประเมินเพิ่มผลงานขึ้นใช้อย่างแพร่หลาย นักประเมินทางการศึกษาพบว่า การประเมินเพิ่มผลงานยังขาดทฤษฎีสนับสนุน เนื่องจากทฤษฎีจิตมิติมีข้อจำกัดไม่สามารถใช้เป็นหลักในการประเมินเพิ่มผลงานได้ ด้วยเหตุนี้นักประเมินผลจึงได้พยายามพัฒนาทฤษฎีพื้นฐานเพื่อใช้ในการประเมินเพิ่มผลงาน

ผลของความพยายามที่สำคัญได้แก่ผลงานของ Moss, Schutz และ Collons (1998) จากมหาวิทยาลัยมิชิแกน ซึ่งได้ทำการวิจัยศึกษาถึงแนวทางการประเมินเพิ่มผลงานสำหรับการรับรองวิทยฐานะของครู โดยได้จัดทำโครงการพัฒนาการประเมินค่าการปฏิบัติงาน (Performance Assessment Development Project = PADP) อันเป็นโครงการร่วมมือสนับสนุนและประเมินค่าอาจารย์ใหม่ระหว่างรัฐ (Interstate Teacher Assessment and Support Consortium = INTASC) ในงานวิจัยชิ้นนี้นักประเมินได้หลักวิชาวิทยาการตีความ (hermeneutics) มาใช้เป็นหลักพื้นฐานในการประเมินเพิ่มผลงาน

หลักการของวิธีวิทยาการตีความ มีเป้าหมายคล้ายกับหลักการทางทฤษฎีจิตมิติ แต่มีส่วนที่แตกต่างกันคือ ทฤษฎีจิตมิติให้ความสำคัญในการประเมินจากการตัดสินคุณค่าบนพื้นฐานการพิจารณาข้อบกพร่องของการปฏิบัติงานแต่ละชิ้น แต่หลักการวิธีวิทยาการตีความ ให้ความสำคัญในการประเมินผลการปฏิบัติงานในภาพรวม และมีการให้คะแนนตลอดจนการตัดสินคุณค่าผลการปฏิบัติงานหลายชิ้นโดยอาศัย

เกณฑ์หลายอย่าง เช่น การตีความโดยคำนึงถึงโครงสร้าง ข้อโต้แย้ง และมีการพิจารณาถึงการตีความก่อนหน้าด้วย (revising initial interpretation)

วิธีวิทยาการตีความ (hermeneutics) ตามแนวคิดของนักปรัชญา Gadamer (1987 อ้างใน Moss, Schutz และ Collins, 1998) ที่ผู้วิจัยนำมาใช้ในงานวิจัยนี้ มีแนวความคิดว่า เมื่อศึกษาผลการปฏิบัติงานทุกชิ้นแล้วการประเมินผลการปฏิบัติงานแต่ละชิ้น ผู้ประเมินอาจมีความลำเอียงเนื่องจากความรู้และประสบการณ์ที่มีมาก่อน ดังนั้น เพื่อให้ได้ผลการประเมินที่ถูกต้อง ผู้ประเมินต้องมีการสนทนาอภิปรายโต้แย้งกัน เพื่อให้ได้ผลการตีความโดยสรุปที่เป็นผลการประเมินร่วมกัน ผลการประเมินที่ได้จะปราศจากความลำเอียงของนักประเมินแต่ละคน กระบวนการประเมินให้ได้ผลการประเมินร่วมกันดังกล่าวนี้มีข้อตกลง (assumption) 2 ประการ ประการแรก ผลการปฏิบัติงานแต่ละชิ้นของผู้รับการประเมินมีข้อบกพร่องและมีจุดเด่นที่แตกต่างกัน เมื่อนำมารวมกันจะบ่งบอกถึงลักษณะร่วมที่เป็นลักษณะที่แท้จริงของบุคคลผู้ปฏิบัติงานได้ ประการที่สอง ผู้ประเมินที่ทำการประเมินชิ้นงานเดียวกัน มีประสบการณ์พื้นฐานความรู้ต่างกัน จึงทำให้คะแนนการประเมินที่ได้แตกต่างกัน เมื่อมีการสนทนา การถกเถียง อภิปรายร่วมกัน จะทำให้ได้ข้อสรุปที่เป็นผลการประเมินที่ทุกคนเห็นพ้องกัน

จากข้อตกลงเบื้องต้นดังกล่าวแสดงว่าการประเมินตามหลักการวิธีวิทยาการตีความ ไม่สามารถใช้เกณฑ์การประเมินที่กำหนดล่วงหน้าเฉพาะสภาพการณ์คงที่ตามแบบทฤษฎีจิตมิติได้ แต่ต้องมีการพิจารณาตามเกณฑ์ที่ปรับเปลี่ยนได้กับทุกสถานการณ์ เกณฑ์ดังกล่าวแตกต่างกันตามความคิดเห็นของผู้ประเมินแต่ละคน และต้องอาศัยการสนทนา อภิปรายร่วมกันระหว่างผู้ประเมินในการประเมินผลการปฏิบัติงาน

1.3 ประเภทของแฟ้มผลงาน

จากการศึกษาเอกสารพบว่าการแบ่งประเภทของแฟ้มผลงานเป็นหลายลักษณะแตกต่างกันตามเกณฑ์ที่นักประเมินใช้เป็นหลักในการแบ่งประเภท โดยทั่วไปนักประเมินจัดกลุ่มประเภทของแฟ้มผลงานเป็นกลุ่มใหญ่ 2-4 ประเภท แต่การแบ่งประเภทแฟ้มผลงานของ Bure, Fogarty และ Belgrad (1994) ให้รายละเอียดมากกว่า ผู้วิจัยจึงนำเสนอประเภทของแฟ้มผลงานที่ Bure, Fogarty และ Belgrad แบ่งไว้เป็น 3 ประเภทโดยยึดเนื้อหาสาระในแฟ้มผลงานเป็นหลัก คือ แฟ้มผลงานบุคคล แฟ้มผลงานเชิงวิชาการ และแฟ้มผลงานเชิงวิชาชีพ โดยมีรายละเอียดของแฟ้มผลงานแต่ละประเภทดังนี้

1. แฟ้มผลงานบุคคล (personel portfolio) แฟ้มที่จัดว่าเป็นแฟ้มผลงานส่วนบุคคล มีการจัดทำ

2 ลักษณะ คือ

1.1 แฟ้มรวมผลงานทั้งหมด (the total portfolio) แฟ้มผลงานชนิดนี้จะทำให้ทราบข้อมูลลักษณะของผู้ทำแฟ้มทั้งความสามารถในด้านต่างๆ ความชอบ ทักษะคติต่อสิ่งต่างๆ ความต้องการในอนาคต นอกจากนี้ในแฟ้มผลงานอาจบรรจุหลักฐานต่างๆตามที่ผู้ทำแฟ้มผลงานต้องการจัดเก็บก็ได้

1.2 แฟ้มชิ้นงานที่ดีที่สุด (my best work) เมื่อทำชิ้นงานสำเร็จหลายชิ้น ผู้ทำแฟ้ม

ผลงานจะทำการคัดเลือกเอกสารชิ้นงานที่ดีที่สุดจำนวนหนึ่งบรรจุในแฟ้มผลงานที่ดีที่สุดนี้ ชิ้นงานที่คัดเลือกจะแสดงถึงความสามารถ ความก้าวหน้าและพัฒนาการของผู้ทำแฟ้มผลงาน

2. แฟ้มผลงานเชิงวิชาการ (academic portfolio) แฟ้มผลงานที่ถือว่าเป็นแฟ้มผลงานเชิงวิชาการมีอยู่หลายชนิดแตกต่างกันตามวัตถุประสงค์ของการประเมิน คือ

2.1 แฟ้มผลงานเพื่อให้เกรด (the graded portfolio) แฟ้มผลงานชนิดนี้ประกอบด้วยชิ้นงานซึ่งรวบรวมโดยมีวัตถุประสงค์ในการนำเสนอเพื่อขอรับระดับผลการเรียน โดยผู้รับการประเมินต้องเก็บรวบรวมชิ้นงานซึ่งเป็นผลงานเชิงวิชาการที่จะบ่งบอกให้ทราบถึงความสามารถของผู้ทำแฟ้มผลงานได้

2.2 แฟ้มผลงานบูรณาการ (integrated portfolio) ชิ้นงานที่บรรจุในแฟ้มนี้ เป็นผลการปฏิบัติงานที่เกิดจากการบูรณาการความรู้ในการเรียนวิชาต่างๆเข้าด้วยกัน

2.3 แฟ้มผลงานกลุ่ม (cooperative group portfolio) ชิ้นงานที่บรรจุลงในแฟ้มชนิดนี้เป็นผลงานของการทำงานเป็นกลุ่มซึ่งจะสะท้อนให้เห็นถึงความร่วมมือภายในกลุ่มในการผลิตชิ้นงาน

2.4 แฟ้มผลงานในรอบหลายปี (multi-year portfolio) บางโรงเรียนมีการแบ่งช่วงของระดับการศึกษา แต่ละช่วงอาจเป็น 2 หรือ 3 หรือ 4 ปี แฟ้มผลงานประเภทนี้จึงเป็นการรวบรวมผลงานในรอบ 2 หรือ 3 หรือ 4 ปี ตามช่วงระดับการศึกษาดังกล่าว

2.5 แฟ้มผลงานเชาวน์ปัญญาพหุหรือแฟ้มผลงานเชาวน์ปัญญาครบส่วน (multiple intelligences) แฟ้มผลงานประเภทนี้จะบรรจุชิ้นงานที่แสดงความสามารถ 7 ด้านของผู้ทำแฟ้มผลงาน ตามการแบ่งความสามารถทางสติปัญญาของมนุษย์ ของ Gardner (1983) ที่แบ่งเป็น 7 ด้าน คือ ด้านการมองเห็น (visual), ด้านเหตุผลและการคิดคำนวณ (logical/mathematics), ด้านภาษา (verbal), ด้านการดนตรี (musical), ด้านการปรับตัวให้เข้ากับคนอื่น (interpersonal), ด้านการวิเคราะห์ตนเอง (intrapersonal) และ ด้านการเคลื่อนไหวของร่างกาย (bodily/kinesthetic)

2.6 แฟ้มประวัติของชั้นเรียน (the class profile) แฟ้มชนิดนี้เป็นการรวบรวมลักษณะต่างๆของชั้นเรียน เช่น รูปภาพห้องเรียน คติพจน์ของห้อง หรือเพลงประจำชั้นเรียน

2.7 แฟ้มผลงานความสามารถทางสติปัญญา (portfolio of intelligent behavior) แฟ้มชนิดนี้จะเก็บหลักฐานที่แสดงถึงความสามารถของนักเรียนที่มาจากการใช้สติปัญญา เช่นรางวัลจากการแข่งขันตอบปัญหา, ใบประกาศการชนะเลิศในกิจกรรมต่าง ๆ เป็นต้น

2.8 แฟ้มประวัติของโรงเรียน (the school midle profile) แฟ้มชนิดนี้มีลักษณะคล้ายกับแฟ้มประวัติของชิ้นงานซึ่งแฟ้มประวัติโรงเรียนนี้อาจรวบรวมรายชื่อผู้บริหารโรงเรียนหรือคณะกรรมการโรงเรียน รูปถ่ายหลักฐานสำคัญของโรงเรียน เป็นต้น

2.9 แฟ้มผลงานแคปซูลเวลา (time capsule portfolio) แฟ้มชนิดนี้เป็นแฟ้มผลงานของนักเรียนที่มีจุดประสงค์จะรวบรวมหลักฐานเกี่ยวกับประวัติของโรงเรียนสำหรับนักเรียนรุ่นต่อไป การจัด

ทำเพิ่มแคปซูลเวลาใช้หลักของการทำแคปซูลอวกาศ โดยจัดทำเพิ่มผลงานแล้วเก็บไว้ในที่ที่กำหนด เมื่อเวลาผ่านไป 1 ปี จึงนำออกมาศึกษาเขียนเป็นประวัติของโรงเรียนตามความคิดเห็นของนักเรียนในแต่ละรุ่น

2.10 เพิ่มประวัติของชุมชน (หมู่บ้าน/ท้องถิ่น) (districtwide profile) ในเพิ่มนี้บรรจุข้อมูลของชุมชนอาจเป็นรายชื่อสถานที่ในชุมชน เช่น ร้านหนังสือ กำหนดการจัดงานเกี่ยวกับวิทยาศาสตร์ โครงการต่าง ๆ ของหมู่บ้าน นอกจากนี้ยังมีหลักฐานที่เป็นความภาคภูมิใจในชุมชน เช่น รางวัลการชนะเลิศในการแข่งขันต่าง ๆ เป็นต้น

การแบ่งประเภทเพิ่มผลงานเชิงวิชาการของ Bure, Fogarty และ Belgrad (1994) ดังกล่าวข้างต้น มีทั้งเพิ่มผลงานที่เป็นเพิ่มผลงานบุคคล เพิ่มผลงานกลุ่ม เพิ่มผลงานโรงเรียนและเพิ่มผลงานชุมชน ซึ่งแตกต่างจากที่นักประเมิณท่านอื่นแบ่งประเภทไว้ คือ LaMahieu (1995) แบ่งประเภทของเพิ่มผลงานตามขนาดของหน่วยงานที่จัดทำเพิ่มผลงาน เป็นเพิ่มผลงานในหน่วยงานขนาดใหญ่ (large Seale) และเพิ่มผลงานในหน่วยงานขนาดเล็ก (small Seale) การแบ่งประเภทตามขนาดของหน่วยงานของ LaMahieu มีลักษณะคล้ายกับ Leckase (1995) ที่แบ่งประเภทเพิ่มผลงานตามระดับหน่วยงาน แต่มี 3 ระดับ คือ เพิ่มผลงานระดับโรงเรียน (school level) เพิ่มผลงานระดับท้องถิ่นหรือภูมิภาค (district level) และเพิ่มผลงานระดับรัฐ (state level)

3. เพิ่มผลงานเชิงวิชาชีพ (professional portfolio) เพิ่มผลงานที่จัดว่าเป็นเพิ่มผลงานเชิงวิชาชีพ ต้องเป็นเพิ่มบรรจุเอกสารชิ้นงานที่แสดงให้เห็นถึงความสามารถที่เป็นทักษะในการประกอบอาชีพ ได้แก่

3.1 เพิ่มผลงานเพื่อการเรียนต่อ (college admission portfolio) เพิ่มชนิดนี้จะบรรจุตัวอย่างชิ้นงานที่ทำในระหว่างที่เรียนในระดับมัธยมปลาย ซึ่งในการศึกษาต่อนั้น วิทยาลัย/ มหาวิทยาลัย กำหนดให้เสนอเพิ่มผลงานชนิดนี้

3.2 เพิ่มผลงานเพื่อขอทุนการศึกษา (college scholarships) เพิ่มชนิดนี้บรรจุข้อมูลสารสนเทศที่เป็นประโยชน์สำหรับการขอทุนการศึกษา เช่น Transcripts คะแนนผลการสอบจากแบบสอบมาตรฐาน นอกจากนี้อาจมีจดหมาย เรียงความ และอื่นๆ ที่แสดงถึงความสามารถในเชิงวิชาการของผู้ทำเพิ่มผลงาน

3.3 เพิ่มผลงานเพื่อการสมัครงาน (employability portfolio) ในการสมัครงานจำเป็นต้องอาศัยหลักฐานที่จะยืนยันความสามารถของผู้ทำเพิ่มประกอบการสัมภาษณ์เพื่อคัดเลือกเข้าทำงาน เพิ่มผลงานประเภทนี้จึงจำเป็นสำหรับผู้ต้องการสมัครทำงาน ซึ่งในเพิ่มชนิดนี้ผู้ทำเพิ่มอาจรวบรวมชิ้นงานที่แสดงถึงความสามารถในกิจกรรมต่างๆ ความสามารถในการสื่อสาร ความรับผิดชอบในการทำงาน เป็นต้น

3.4 เพิ่มผลงานจากการฝึกปฏิบัติงาน (preservice portfolio) ก่อนการสำเร็จการศึกษา หลักสูตรการศึกษาทางการศึกษากำหนดให้นักศึกษาได้จัดทำเพิ่มผลงาน เช่น นักศึกษาครูกทำเพิ่มผลงานจากการฝึกสอน ซึ่งในเพิ่มอาจประกอบด้วย แผนการสอน ชิ้นงานของนักเรียน กิจกรรมนอกหลักสูตร

ผลการประเมินการสอนของตนเอง และจากครูผู้มีประสบการณ์สอน เพิ่มผลงานนี้อาจใช้สำหรับการสมัครสอบสัมภาษณ์เข้าทำงานได้ด้วย

3.5 เพิ่มผลงานครู (teacher portfolio) เพิ่มชนิดนี้จะรวบรวมหลักฐานที่เกี่ยวกับการสอน เช่น จุดประสงค์ของการสอนระยะสั้น ระยะยาว การวางแผนการสอน รูปภาพ กิจกรรมในชั้นเรียน เป็นต้น เพิ่มผลงานชนิดนี้ครูผู้สอนบางคนยังใช้สำหรับการสมัครงานเมื่อต้องการเปลี่ยนงาน

3.6 เพิ่มผลงานของคณะกรรมการการประเมิน (staff evaluation portfolio) ในบางท้องถิ่นจะพบเพิ่มผลงานชนิดนี้ ซึ่งเป็นการบรรจุผลการประเมินครูผู้สอน อาจเป็นผลจากการสังเกตการสอน ชิ้นงานนักเรียน คะแนนจากการทดสอบ การสัมภาษณ์ รางวัลและอื่นๆ

3.7 เพิ่มผลงานของผู้บริหาร (administration portfolio) ผู้บริหารโรงเรียนอาจจัดทำเพิ่มผลงานเพื่อแสดงถึงความสามารถในการเป็นผู้นำตัวอย่างหลักฐานในเพิ่ม เช่น การวางแผนในโรงเรียน กลยุทธ์ในการบริหาร บันทึกการประชุม และผลงานต่างๆที่เกิดในโรงเรียนทั้งกับครูผู้สอนและนักเรียนของโรงเรียน

3.8 เพิ่มผลงานภาคปฏิบัติ (performance reviews) บางหน่วยงานที่ไม่เกี่ยวข้องกับภาคศึกษามีการใช้เพิ่มผลงานเป็นหลักฐานที่แสดงถึงความสามารถในการปฏิบัติงาน ใช้เป็นดัชนีชี้วัดความสามารถและเจตคติของพนักงานในการทำงาน

จากที่ Bure, Fogarty และ Belgrad (1994) แบ่งประเภทของเพิ่มผลงานตามเนื้อหาสาระในเพิ่มผลงาน ยังมีนักประเมินท่านอื่นที่อาศัยเกณฑ์ที่แตกต่างจากนี้ในการแบ่งประเภทของเพิ่มผลงาน เช่น Niguidula (1997) แบ่งประเภทเพิ่มผลงานตามลักษณะสิ่งที่ใช้เก็บสาระเพิ่มผลงาน เป็นเพิ่มผลงานที่เก็บในคอมพิวเตอร์ (digital portfolio) และเพิ่มผลงานที่เป็นเล่มรายงาน (paper portfolio)

จากเกณฑ์การแบ่งประเภทเพิ่มผลงานที่นำเสนอ กล่าวได้ว่า การแบ่งประเภทเพิ่มผลงานมีเกณฑ์มากขึ้นอยู่กับจุดเน้นที่นักประเมินแต่ละคนต้องการเน้น ซึ่งรวมถึงวัตถุประสงค์ของการจัดทำเพิ่มผลงานด้วย

1.4 ขั้นตอนการประเมินเพิ่มผลงาน

โดยที่การประเมินเพิ่มผลงานเป็นการประเมินการปฏิบัติงานชนิดหนึ่ง เป็นรูปแบบหนึ่งของการประเมิน ดังนั้นขั้นตอนการประเมินเพิ่มผลงานจึงมีความคล้ายคลึงกับขั้นตอนในกระบวนการประเมินทั่ว ๆ ไป นักประเมินกำหนดขั้นตอนการประเมินเพิ่มผลงานไว้แตกต่างกันในรายละเอียด แต่ขั้นตอนสำคัญจะไม่แตกต่างกันเท่าไรนัก การประเมินเพิ่มผลงานโดยทั่วไปจะประเมินในทุกขั้นตอนของการทำเพิ่มผลงาน ผู้วิจัยนำขั้นตอนการจัดทำเพิ่มผลงานที่นักประเมินจำนวน 3 ราย คือ Doner และ Grilman (1998), Reckase (1995) และ LeMahieu, Gilomer และ Eresh (1995) ได้เสนอไว้นามาเป็นประเด็นเพื่อทำการประเมินเพิ่มผลงาน 6 ประเด็น ขั้นตอนที่จะทำการประเมินมีรายละเอียดดังต่อไปนี้

1. ขั้นตอนการกำหนดจุดมุ่งหมายและทิศทางการทำเพิ่มผลงาน

ในขั้นตอนนี้จะเป็นการปรึกษากลางร่วมกันระหว่างผู้รับการประเมินกับผู้ประเมินเพื่อที่จะใช้ในการประเมินเพิ่มผลงาน กำหนดจุดมุ่งหมายของการจัดทำเพิ่มผลงานที่จะทำให้ผู้จัดทำเพิ่มผลงานทราบว่าเป้าหมายของการประเมินเพิ่มผลงานมีอะไรและมีแนวทางการประเมินลักษณะใด โดยทั่วไปการประเมินเพิ่มผลงานมักจะถูกใช้เป็นการประเมินผลการเรียนการสอน ดังนั้นขั้นตอนการกำหนดจุดมุ่งหมายและทิศทางการประเมินต้องเริ่มทำตั้งแต่ขั้นการวางแผนการสอนของผู้สอนและผู้เรียนมีส่วนในการกำหนดจุดมุ่งหมายของการเรียนการสอนตลอดจนทำการประเมินร่วมกัน

2. ขั้นตอนการกำหนดวิธีดำเนินการจัดทำเพิ่มผลงาน

ในขั้นตอนนี้ผู้ประเมินและผู้รับการประเมิน ต้องร่วมกันกำหนดวิธีดำเนินการจัดทำเพิ่มผลงานเพื่อให้เป็นที่เข้าใจตรงกัน ในประเด็นที่ว่าจะมีเกณฑ์การเลือกชิ้นงานอย่างไร จะบรรจุชิ้นงานใดลงในเพิ่มผลงาน และใช้วิธีการใดบ้างในการจัดทำเพิ่มผลงาน

3. ขั้นตอนการพัฒนาเพิ่มผลงาน

ขั้นตอนนี้เป็นหัวใจสำคัญของการจัดทำเพิ่มผลงาน เป็นขั้นตอนที่ใช้เวลานานที่สุด เนื่องจากขั้นตอนนี้มีรายละเอียดปลีกย่อยในการดำเนินการมาก ผู้วิจัยจึงนำเสนอเนื้อหาสาระในขั้นตอนนี้ แยกออกเป็นหัวข้อย่อยรวม 3 หัวข้อ คือ การวางแผนจัดทำเพิ่มผลงาน กระบวนการพัฒนาเพิ่มผลงาน และการนำเสนอเพิ่มผลงาน ดังรายละเอียดต่อไปนี้

3.1 การวางแผนจัดทำเพิ่มผลงาน (Planning)

ในขั้นตอนนี้การวางแผนจัดทำเพิ่มผลงาน ผู้จัดทำเพิ่มผลงานมีการดำเนินการแยกเป็น 3 ขั้นตอน คือ การศึกษาตัวอย่างเพิ่มผลงาน การกำหนดโครงร่าง (outline) ของเพิ่มผลงาน และการคัดเลือกชิ้นงาน โดยดำเนินการดังนี้

ก. การศึกษาตัวอย่างเพิ่มผลงาน (study example of portfolio) ในขั้นตอนนี้ผู้ทำเพิ่มผลงานจะต้องศึกษาตัวอย่างเพิ่มผลงาน และศึกษาลักษณะของเพิ่มผลงานว่ามีส่วนประกอบอะไรบ้าง เพื่อทำความเข้าใจเกี่ยวกับสาระสำคัญหรือชิ้นงานที่จะบรรจุในเพิ่มผลงาน การเห็นตัวอย่างเพิ่มผลงานที่ดีจะช่วยให้ผู้จัดทำเพิ่มผลงานมีความรู้ความเข้าใจในการจัดทำเพิ่มผลงานดีขึ้น

ข. การกำหนดโครงร่างรายงานของเพิ่มผลงาน (construct portfolio outline) ผู้ทำเพิ่มผลงานต้องมีความรู้ความเข้าใจว่าจะต้องบรรจุผลงานลักษณะใดในเพิ่มผลงานก่อน จึงจะสามารถกำหนดโครงร่างรายงานของเพิ่มผลงานได้ ขั้นตอนนี้จึงจำเป็นต้องอาศัยความรู้ ความสามารถจากการศึกษาเอกสารชิ้นงาน ตัวอย่างเพิ่มผลงานแล้วทำการกำหนดโครงร่างเพิ่มผลงานของตนเอง

ค. การคัดเลือกเอกสารชิ้นงาน (selecting material) และการกำหนดเกณฑ์ในการคัดเลือก (setting criteria) ในขั้นตอนนี้ผู้ทำเพิ่มผลงานต้องยึดเป้าหมายของวัตถุประสงค์ของการทำเพิ่ม

ผลงานเป็นหลักในการวางแผนเลือกเอกสารชิ้นงานใส่เพิ่มผลงานให้สนองวัตถุประสงค์และสอดคล้องตามเกณฑ์ที่กำหนดไว้สำหรับการคัดเลือกเอกสารชิ้นงานเพื่อบรรจุในเพิ่มผลงาน

3.2 กระบวนการพัฒนาเพิ่มผลงาน (portfolio development)

ในการพัฒนาเพิ่มผลงาน ผู้ทำเพิ่มผลงานมีการดำเนินการ 3 ขั้นตอนคือการศึกษาเอกสารชิ้นงานที่ตรงตามโครงสร้างสาระในเพิ่มผลงาน การสะท้อนความคิดเห็นและการลดจำนวนข้อมูลดังรายละเอียดต่อไปนี้

ก. การศึกษาเอกสารชิ้นงานที่ตรงตามโครงสร้างสาระในเพิ่มผลงาน (reading relevant material) ขั้นตอนนี้ผู้ทำเพิ่มผลงานนำเอกสารชิ้นงานที่คัดเลือกไว้ในข้อ 3.1 (ค) มาศึกษาทบทวนอีกครั้งหนึ่งว่าเนื้อหาสาระในเอกสารชิ้นงานตรงตามจุดมุ่งหมายในการพัฒนาเพิ่มผลงานหรือไม่

ข. กระบวนการสะท้อนความคิดเห็น (process of reflecting) จากการอ่านและศึกษาเอกสารชิ้นงานแต่ละชิ้นในข้อ ก. ผู้จัดทำเพิ่มผลงานต้องบันทึกข้อความที่เป็นการสะท้อนความคิดเห็นความรู้สึกเกี่ยวกับเอกสารชิ้นงานที่เป็นผลงานแต่ละชิ้นด้วย

ค. การลดจำนวนข้อมูล (data reduction) เนื่องจากเอกสารชิ้นงานของผู้ทำเพิ่มผลงานมีเป็นจำนวนมาก ถึงแม้เอกสารชิ้นงานจะผ่านการคัดเลือกมาแล้วรอบหนึ่งก็ตาม การบรรจุเอกสารชิ้นงานลงเพิ่มก็ต้องมีการคัดเลือกกันกรองอีก ซึ่งผู้ทำเพิ่มผลงานต้องพิจารณาลดปริมาณเอกสารชิ้นงานให้มีจำนวนที่เหมาะสมและพอดี รวมทั้งต้องพิจารณาว่าเอกสารชิ้นงานเพียงพอต่อการพัฒนาเพิ่มผลงานให้ตรงตามวัตถุประสงค์หรือไม่ จำเป็นต้องเพิ่มหรือตัดเอกสารชิ้นงานใด จากนั้นจึงนำเอกสารชิ้นงานมาจัดเรียงเรียงตามโครงสร้างเพิ่มผลงานที่กำหนดไว้แล้ว การจัดเรียงเอกสารชิ้นงานที่เกี่ยวข้องให้เหมาะสมต้องอาศัยความคิด ความรู้ ความเข้าใจทั้งในระดับปกติและในระดับสูง (metacognitive) เพื่อตัดสินใจว่าเอกสารชิ้นงานแต่ละชิ้นมีคุณค่าเหมาะสมที่จะบรรจุลงในเพิ่มผลงานหรือไม่

ง. การสะท้อนความคิดเห็นเกี่ยวกับการสร้างเพิ่มผลงาน (reflecting on creation of portfolio) ในขั้นตอนนี้เป็นการรายงานความรู้สึกและความคิดเห็นที่เป็นการสะท้อนความรู้สึกของผู้ทำเพิ่มผลงานว่าจากกระบวนการทั้งหมดที่ผ่านมา ผู้ทำเพิ่มผลงานได้เรียนรู้เทคนิคและฝึกทักษะในด้านใดและด้วยวิธีการใด สิ่งที่เรียนรู้จะเป็นประโยชน์ต่อผู้ทำเพิ่มผลงานในการปรับปรุงแก้ไขตนเองในอนาคตได้อย่างไร

3.3 การนำเสนอเพิ่มผลงาน (presentation portfolio)

ในการจัดทำเพิ่มผลงาน ผู้ทำเพิ่มจะเป็นผู้เก็บรวบรวมเอกสารชิ้นงานและจัดใส่ลงในเพิ่ม ผลงานที่จะนำเสนอเพื่อรับการประเมินหรือเพื่อแสดงในการจัดนิทรรศการแสดงผลงานของนักเรียนเมื่อสิ้นสุดภาคการศึกษาหรือปีการศึกษานั้นมีรูปแบบการนำเสนอเนื้อหาสาระในเพิ่มและมีการจัดทำรูปเล่มที่แตกต่างกัน แต่ส่วนใหญ่มีเนื้อหาสาระที่คล้ายคลึงกัน คือในเพิ่มผลงานต้องมีเอกสารชิ้นงานที่ผ่านการคัดเลือกของผู้ทำเพิ่มผลงาน มีการสะท้อนความคิดของผู้ทำเพิ่มผลงานต่อเอกสารชิ้นงานนั้น ส่วน

รายละเอียดปลีกย่อยอื่น ๆ อาจถูกกำหนดโดยผู้ประเมินเพิ่มผลงานหรือผู้ทำเพิ่มผลงานเองก็ได้ ผู้วิจัยนำเสนอตัวอย่างการจัดรูปแบบนำเสนอเนื้อหาสาระในเพิ่มผลงานที่มีนักประเมินผลเสนอไว้ดังนี้

Doner และ Grilman (1998) เสนอว่า เพิ่มผลงานที่จะนำเสนอ ควรเป็นเพิ่มที่ประกอบด้วย สาระ 3 ส่วน คือ สารบัญเพิ่มผลงาน คำบรรยายสรุป และการสะท้อนความคิดเห็นขั้นสุดท้าย รายละเอียดของสาระแต่ละส่วนเป็นดังนี้

ก. สารบัญเพิ่มผลงาน (table of content) เนื้อหาสาระของเพิ่มผลงานในส่วนนี้เป็นรายละเอียดเกี่ยวกับเอกสารชิ้นงานในเพิ่ม รวมทั้งรายละเอียดสำหรับชิ้นงานแต่ละชิ้นว่าใช้เป็นหลักฐานเพื่อรับการประเมินในด้านใดหรือตามจุดมุ่งหมายใด ตลอดจนบันทึก วัน เดือน ปี ที่ผลิตเอกสารชิ้นงานพร้อมเหตุผลในการคัดเลือกเอกสารชิ้นงาน รวมทั้งผลการประเมินของผู้ทำเพิ่มผลงานในชิ้นงานแต่ละชิ้นว่ามีความเหมาะสมเพียงใดในการบรรจุลงในเพิ่มผลงาน

ข. คำบรรยายสรุป (writing inventory) เนื้อหาสาระในส่วนนี้เป็นรายงานอย่างสมบูรณ์ของการจัดทำเพิ่มผลงาน มีการเขียนอธิบายกระบวนการในการพัฒนาเพิ่มผลงาน ตลอดจนรายละเอียดเกี่ยวกับเอกสารชิ้นงานและลักษณะการจัดเรียงเอกสารชิ้นงานลงในเพิ่มผลงาน

ค. การสะท้อนความคิดเห็นขั้นสุดท้าย (final reflecting) เนื้อหาสาระส่วนนี้เป็นรายงานที่สะท้อนถึงความรู้สึก ความคิดเห็น รวมทั้งการประเมินตนเองของผู้ทำเพิ่มผลงานว่าเกิดการเรียนรู้ในด้านใดและอย่างไร และสิ่งที่ได้รับการเรียนรู้นั้นเป็นประโยชน์ในการพัฒนาตนเองอย่างไร

เนื้อหาสาระในเพิ่มที่เสนอโดย Doner และ Grilman (1998) เป็นรูปแบบที่ไม่ลงลึกในรายละเอียดเท่าใดนัก ต่างจากรูปแบบที่นำเสนอโดย LeMahieu, Gilomer และ Eresh (1995) ที่เสนอรูปแบบเพิ่มผลงานการเขียนของนักเรียนระดับ 6 -12 โดยกลุ่มของเขาได้กำหนดเนื้อหาสาระในเพิ่มผลงานการเขียนที่จะทำการประเมินว่า เพิ่มผลงานที่จะนำเสนอนั้นต้องเป็นเพิ่มที่บรรจุเอกสารชิ้นงานที่ผู้ทำเพิ่มผลงานทำเอง และคัดเลือกเอกสารชิ้นงานมา 5 ชิ้น ตามเงื่อนไขของการคัดเลือกแต่ละชิ้นดังนี้ 1) เป็นชิ้นงานที่มีความสำคัญที่สุด 2) เป็นชิ้นงานที่ผู้ทำเพิ่มผลงานพอใจที่สุด 3) เป็นชิ้นงานที่ผู้ทำเพิ่มผลงานไม่พอใจมากที่สุด 4) เป็นชิ้นงานที่ผู้ทำเพิ่มผลงานต้องการคัดเลือกเองตามใจชอบ และ 5) เป็นชิ้นงานที่คัดเลือกตามคำแนะนำของครูผู้สอน

เมื่อคัดเลือกชิ้นงานได้ครบทั้ง 5 ชิ้นแล้ว ผู้ทำเพิ่มผลงานต้องบรรจุลงในเพิ่มผลงานนำเสนอผู้ประเมินเพื่อรับการประเมิน โดยต้องมีเนื้อหาสาระในเพิ่มแบ่งเป็น 3 ส่วน คือ ส่วนนำซึ่งเป็นรายละเอียดเกี่ยวกับชิ้นงาน เช่น วันเดือนปีที่ได้รับมอบหมายให้ทำชิ้นงาน วันเดือนปีที่ทำชิ้นงานสำเร็จ หรือวันเดือนปีที่คัดเลือกชิ้นงาน ส่วนที่สองเป็นส่วนเนื้อหาของเพิ่ม ซึ่งส่วนนี้คือชิ้นงานที่คัดเลือกตามเงื่อนไขที่กำหนดไว้ และส่วนสุดท้ายเป็นส่วนสรุป ส่วนนี้จะบันทึกการสะท้อนความคิดต่อชิ้นงานและเพิ่มผลงาน

นอกจากรูปแบบการนำเสนอเนื้อหาสาระของเพิ่มผลงานของนักประเมินที่กล่าวมาข้างต้นแล้ว ยังมีนักประเมินอีกท่านหนึ่งคือ Reckase (1995) ได้ให้ข้อเสนอแนะสำหรับเพิ่มผลงานการเขียนของนักเรียนว่า

ผู้ทำเพิ่มผลงานควรมีส่วนร่วมในการคัดเลือกชิ้นงาน การกำหนดเกณฑ์ในการคัดเลือก รวมทั้งเกณฑ์ในการประเมินชิ้นงาน และสุดท้ายต้องมีหลักฐานการสะท้อนความคิดของผู้ทำเพิ่มผลงานต่อชิ้นงานนั้นด้วย โดยเพิ่มผลงานควรมีไม่เกิน 25 หน้า และต้องประกอบด้วยเนื้อหาสาระ 7 ส่วน ที่สำคัญ คือ 1) การสะท้อนความคิดของผู้ทำเพิ่มผลงานต่อชิ้นงาน 2) การเขียนบรรยายประสบการณ์ที่ได้รับจากการทำชิ้นงาน 3) ความสามารถในการเขียนเรียงความ 4) ผลงานการเขียน 5) เกณฑ์การประเมินงานเขียน 6) การตรวจสอบของครู 7) โครงร่างงานเขียนและผลการประเมินงานเขียน

กล่าวโดยสรุปในการพัฒนาเพิ่มผลงาน ประกอบด้วยขั้นตอนที่สำคัญ 3 ขั้นคือ การวางแผนเพื่อจัดทำเพิ่มผลงาน ซึ่งเริ่มตั้งแต่การศึกษาตัวอย่างเพิ่มผลงานจนถึงการกำหนดโครงสร้างเพิ่มผลงานและเกณฑ์การคัดเลือกเอกสารชิ้นงาน อีกขั้นตอนหนึ่งคือการพัฒนาเพิ่มผลงาน ในขั้นนี้มีการคัดเลือกชิ้นงานและการสะท้อนความรู้สึกต่อชิ้นงาน ขั้นตอนสุดท้ายคือการนำเสนอเพิ่มผลงาน ในการนำเสนอเพิ่มผลงานนี้ผู้จัดทำเพิ่มต้องเรียบเรียงสาระในเพิ่มให้มีความสมบูรณ์และเหมาะสม เพื่อรับการประเมินหรือนำแสดงในการจัดนิทรรศการ เนื้อหาสาระในเพิ่มผลงานที่จะนำเสนอผู้ประเมินหลายท่านให้ความสำคัญกับการสะท้อนความคิดเห็นของผู้ทำเพิ่มผลงานต่อชิ้นงานในเพิ่ม

4. การกำหนดตัวผู้ประเมินและการให้คะแนน (scorer and scoring process)

ภาระงานที่สำคัญอีกอย่างหนึ่งในกระบวนการประเมินผลก็คือการให้คะแนนเพิ่มผลงาน เพราะเป็นแหล่งที่ทำให้เกิดปัญหาในการประเมินมาก ดังที่ LeMahieu, Gilomer และ Eresh (1995) ได้รวบรวมปัญหาของการประเมินเพิ่มผลงาน พบว่าโดยส่วนมากจะมีปัญหาในเรื่องการยอมรับเกณฑ์ที่ใช้ในการประเมิน ความเพียงพอของชิ้นงานที่จะประเมินความสามารถในการปฏิบัติงานของผู้ทำเพิ่มผลงาน ความหลากหลายของมาตรฐานการประเมิน ความสะดวกของการใช้เกณฑ์การประเมินในสถานการณ์ต่าง ๆ และการสรุปอ้างอิงผลการประเมิน

ในกระบวนการประเมินเพิ่มผลงานเพื่อให้ปราศจากปัญหาดังกล่าวข้างต้น ประเด็นที่ต้องคำนึงถึงคือการกำหนดตัวผู้ประเมินและวิธีการให้คะแนนหรือประเมินเพราะ ผู้ประเมิน (scorer/rater/reader/examiner) ในการประเมินการปฏิบัติงาน มีอิทธิพลที่จะก่อให้เกิดการพัฒนาสำหรับผู้ปฏิบัติงาน เหตุผลที่นักประเมินให้ความสำคัญกับการกำหนดตัวผู้ประเมินการปฏิบัติงานมีหลายประการ ผู้วิจัยนำเสนอความคิดเห็นของนักประเมินที่เกี่ยวข้องกับผู้ทำหน้าที่ประเมินบางท่าน คือ Lyman (1986) สรุปว่า ผู้ประเมินเป็นแหล่งที่ก่อให้เกิดความคลาดเคลื่อนได้บ่อยมาก และผู้ประเมินเป็นแหล่งความคลาดเคลื่อนที่สำคัญที่สุดในการประเมินรายบุคคล มีงานวิจัยหลายเรื่องที่แสดงให้เห็นว่าเมื่อมีการประเมินรายบุคคลและให้ผู้ประเมินหลายคน จะพบว่ามีผลแตกต่างในผลการประเมินเสมอ

Burke, Rogarty และ Belgrad (1994) กล่าวว่า ในการประเมินเพิ่มผลงานเพื่อให้เกิดการพัฒนา นั้นผู้ประเมินต้องทำการประเมินอย่างมีขั้นตอน และผู้ประเมินต้องเป็นผู้ที่มีความรู้ ความเข้าใจในกระบวนการประเมินผลเป็นอย่างดี

Hart (1994) เสนอว่าผู้ประเมินควรเป็นผู้ที่มีความสามารถในการวิเคราะห์รายละเอียดของเอกสาร ชิ้นงานในแฟ้มผลงาน ซึ่งจะเป็นแนวทางในการประเมินค่าแฟ้มผลงานที่มีความแตกต่างกันในหลายมิติ

จากความเห็นของนักประเมินผลที่กล่าวมาข้างต้น กล่าวได้ว่าผู้ประเมินมีความสำคัญและมีอิทธิพล ต่อผลการประเมิน ดังนั้นผู้ประเมินจึงควรเป็นผู้ที่มีความรู้ความสามารถและเข้าใจเกณฑ์การประเมินผลเป็น อย่างดี Hart (1994) กล่าวเพิ่มเติมว่า โดยทั่วไปของการประเมินแฟ้มผลงาน ผู้ประเมินจะเป็นบุคคลที่เกี่ยวข้อง ข้องกับผู้ทำแฟ้มผลงาน เช่น การประเมินแฟ้มผลงานนักเรียน ผู้ประเมินมักจะประกอบด้วย ครู เพื่อน ร่วมชั้นเรียน ผู้ปกครองและนอกจากนี้ตัวนักเรียนเองยังต้องเป็นผู้ประเมินแฟ้มผลงานของตนเองอีกด้วย ในทำนองเดียวกันสำหรับแฟ้มผลงานของบุคคลอื่น ๆ ผู้ประเมินจะถูกกำหนดขึ้นจากบุคคลที่เกี่ยวข้องมีผล ได้ผลเสียกับการประเมินนั้นเสมอ เนื่องจากผู้ประเมินแฟ้มผลงานประกอบด้วยบุคคลหลายฝ่าย ดังนั้น ก่อนการประเมินจึงจำเป็นต้องมีการปรึกษาหารือกันระหว่างผู้ประเมินด้วยกันเพราะจะทำให้ผู้ประเมินเหล่านั้นมีความเข้าใจตรงกันในเกณฑ์การประเมินและวิธีการประเมิน เพื่อให้การประเมินแฟ้มผลงานเป็นไปอย่าง มีประสิทธิภาพ

เมื่อกำหนดตัวผู้ประเมินแฟ้มผลงานได้แล้ว สิ่งที่ต้องคำนึงถึงประการต่อมาคือ การให้คะแนนหรือ การประเมินแฟ้มผลงาน สิ่งที่ใช้เป็นแนวทางในการประเมินชิ้นงานคือ scoring rubric หรือเกณฑ์การให้ คะแนน ซึ่งมีลักษณะเป็นระดับผลการดำเนินงาน โดยแต่ละระดับเป็นรายการที่บรรยายถึงคุณภาพของชิ้น งานในประเด็นที่ทำการประเมิน เป็นการตอบคำถามว่าผู้ถูกประเมินทำอะไรได้สำเร็จ หรือว่ามีความสำเร็จ ในขั้นใด ผลงานเป็นอย่างไร Farr และ Tone (1994) กล่าวว่า เกณฑ์การให้ระดับความสำเร็จของงานนี้ โดยทั่วไปเสนอในรูปตาราง 2 มิติ คือมิติของระดับคุณภาพที่อาจแบ่งเป็น 3-6 ระดับก็ได้ และมิติของ ประเด็นที่นำมาประเมินหรือรายการประเมินนั่นเอง Hart (1994) เสนอว่า เกณฑ์การให้คะแนนแฟ้มผลงาน มี 2 แบบ คือ เกณฑ์การให้คะแนนในภาพรวม (holistic scoring rubric) และเกณฑ์การให้คะแนนราย องค์ประกอบ (analytic scoring rubric) ซึ่งมีรายละเอียดดังนี้

ก. เกณฑ์การให้คะแนนในภาพรวม (holistic scoring rubric) คือ เกณฑ์การให้คะแนนโดย พิจารณาความสำเร็จในภาพรวมของการปฏิบัติงาน ซึ่งอาจเป็นชิ้นงาน 1 ชิ้น หรือแฟ้มผลงาน 1 แฟ้ม โดย ระดับความสำเร็จอาจแบ่งเป็น 4 - 6 ระดับก็ได้ ตัวอย่างของเกณฑ์การให้คะแนนในภาพรวม เช่น การ กำหนดระดับคุณภาพของการปฏิบัติการแก้ปัญหาใจหทัยคณิตศาสตร์ กำหนดดังนี้

ระดับคะแนน 4 หมายถึง คำตอบถูกต้อง มีขั้นตอนการหาคำตอบที่ชัดเจนสมบูรณ์

ให้เหตุผลและมีหลักการในการแก้ปัญหา

ระดับคะแนน 3 หมายถึง คำตอบถูกต้อง ขั้นตอนการหาคำตอบไม่ชัดเจนและ

ไม่สมบูรณ์เล็กน้อย

ระดับคะแนน 2 หมายถึง คำตอบไม่ถูกต้อง ขั้นตอนการหาคำตอบไม่ชัดเจนและ

ไม่สมบูรณ์ แต่มีแนวโน้มแสดงให้เห็นถึงการนำไปสู่คำตอบที่ถูกต้อง

ระดับคะแนน 1 หมายถึง คำตอบไม่ถูกต้อง มีการอธิบายแนวคิดที่เกี่ยวข้องบ้าง
แต่ไม่ได้คำตอบของโจทย์ปัญหาที่กำหนด

ระดับคะแนน 0 หมายถึง ไม่ทำโดยกระดาษคำตอบว่างเปล่า หรือมีขีดเขียนบ้างเล็กน้อย
แต่ไม่เกี่ยวข้องกับการแก้ปัญหาโจทย์ที่กำหนดให้

ข. เกณฑ์การให้คะแนนรายองค์ประกอบ (analytic scoring rubric) เพื่อให้เห็นคุณภาพของการปฏิบัติงานในรายละเอียดปลีกย่อย การประเมินจึงมีการวิเคราะห์แยกมิติหรือองค์ประกอบที่จะพิจารณาให้ระดับคะแนน ซึ่งในแต่ละมิติหรือองค์ประกอบก็จะแบ่งระดับคุณภาพความสามารถไว้เช่นเดียวกับในเกณฑ์การประเมินภาพรวม ผู้ประเมินเพิ่มผลงานที่จะใช้เกณฑ์การให้คะแนนรายองค์ประกอบได้ดีต้องเป็นบุคคลที่มีความสามารถในการวิเคราะห์แยกแยะประเด็น/รายการที่จะนำมาประเมินด้วย

Reckase (1995) เสนอว่าองค์ประกอบของเพิ่มผลงานแบ่งตามเนื้อหาสาระในเพิ่มมี 7 ส่วน ได้แก่ การสะท้อนความคิดของผู้ทำเพิ่มผลงานต่อชิ้นงาน, การเขียนบรรยายประสบการณ์ที่ได้รับจากการทำชิ้นงาน, ความสามารถในการเขียนเรียงความ, ผลงานการเขียน, เกณฑ์การประเมินงานเขียน, การตรวจสอบของครู และโครงร่างงานเขียนและผลการประเมินงานเขียน ในการประเมินนั้นควรทำการประเมิน 5 ส่วน คือ การสะท้อนความรู้สึกของผู้ทำเพิ่มผลงานต่อชิ้นงาน การเขียนบรรยายประสบการณ์ที่ได้รับจากการทำชิ้นงาน ความสามารถในการเขียนเรียงความ ผลงานการเขียนและผลการประเมินงานเขียน ในการให้ระดับคะแนนเนื้อหาสาระเพิ่มผลงานทั้ง 5 ส่วนนี้ใช้ผู้ประเมิน 2 คน และผลการประเมินควรจะมีค่าเฉลี่ยอย่างน้อย 0.8 ซึ่ง Reckase เสนอแนะว่าการให้คะแนนชิ้นงานในเพิ่มผลงานเป็นรายชิ้น จะทำให้มีค่าความเที่ยงสูงกว่าการให้คะแนนเพิ่มผลงานทั้งเล่ม การที่จะทำได้ค่าความเที่ยง 0.8 นี้ จากผลการศึกษาพบว่าต้องมีค่า ความคงเส้นคงวภายใน (internal consistency) อย่างน้อย 0.55 และความสัมพันธ์ระหว่างเพิ่มผลงาน (inter correlation) มีค่าอย่างน้อย 0.28

นอกจากนี้ยังมี LeMahieu ที่เสนอแนวทางในการประเมินเพิ่มผลงานการเขียน โดยการให้ระดับคะแนนเพิ่มผลงานใน 3 มิติ คือ ความสมบูรณ์ของงานเขียน กระบวนการและขั้นตอนของการเขียน และศักยภาพในการพัฒนาเพื่อเป็นนักเขียน รายละเอียดของแต่ละมิติมีดังนี้

1. ความสมบูรณ์ของงานเขียน ดูได้จาก ข้อตกลงจากการปรึกษาหารือ การกำหนดวัตถุประสงค์ของการทำเพิ่มผลงาน เทคนิคในการกำหนดรูปแบบในการเขียน การใช้คำและโครงสร้างของประโยค การใช้คำซึ่งคำนึงถึงผู้อ่าน การใช้ภาษา ความสละสลวยของงานเขียน

2. กระบวนการและขั้นตอนในการเขียน ดูได้จาก การวางแผนงานก่อนการลงมือเขียน ร่างงานเขียนที่แสดงถึงพัฒนาการและความคิดเห็นที่เฉียบแหลม ความพยายามในการปรับปรุงงานเขียน การทบทวนงานเขียนอย่างมีวิสัยทัศน์ที่กว้างไกล

3. ศักยภาพในการพัฒนาเพื่อเป็นนักเขียน ดูได้จาก หลักสูตรที่ใช้ในการเขียน (ข้อมูล) การเพิ่มงานเขียน การมีความรู้สึกที่ดีต่องานเขียน การกำหนดเกณฑ์เพื่อพัฒนางานเขียนให้เป็นมาตรฐาน ความสามารถ

ในการศึกษาจุดเด่นและความต้องการของผู้อ่าน การใช้นวัตกรรมในการพัฒนางานเขียน การเขียนเพื่อบุคคลหลายระดับ และต้องมีพัฒนาการในการเขียน

ใน 3 มิติของการพิจารณาให้ระดับคะแนนเพิ่มผลงานตาม ควรมีผู้ประเมิน 2 คน ทำการประเมินตามรายการซึ่งมีลักษณะเป็นแบบมาตราส่วนประมาณค่า 6 ค่า ส่วนการตัดสินผลคุณภาพเพิ่มผลงานต้องดูระดับคะแนนที่ได้จากผู้ประเมิน โดยดูความต่างของคะแนนจากผู้ประเมิน 2 คน ว่ามีค่าต่างกันมากกว่า 1 คะแนนหรือไม่ ถ้าต่างกันมากกว่า 1 คะแนนจะใช้ผู้ประเมินคนที่ 3 เป็นผู้ตัดสินผล แต่ถ้าคะแนนไม่แตกต่างกันก็ใช้คะแนนนั้นได้เลย

นอกจากนี้ยังมีตัวอย่างเกณฑ์การให้คะแนนรายองค์ประกอบของงานเขียนที่เสนอโดย Farr และ Tone (1994) ซึ่งพิจารณาให้ระดับคะแนนจากที่แยกองค์ประกอบของงานเขียนเป็น 6 องค์ประกอบ และแต่ละองค์ประกอบกำหนดระดับคุณภาพจากมากไปหาน้อย ดังรายละเอียดต่อไปนี้

รายการประเมินงานเขียน	ระดับคุณภาพของการปฏิบัติงาน		
	3	2	1
ความสำเร็จของงาน	สำเร็จดีเยี่ยม	สำเร็จ	ไม่สำเร็จ
การจัดองค์การ	สมเหตุสมผล ชัดเจน	ไม่ชัดเจนใน บางจุด	ไม่มีทิศทางในการเขียน
พัฒนาการในงานเขียน	พัฒนาการมาก	พัฒนาการปานกลาง	พัฒนาการน้อย
โครงสร้างของประโยค	ถูกต้อง	ผิดพลาดเล็กน้อย	ผิดพลาดมาก
การใช้ศัพท์	ชัดเจน ถูกต้อง	เหมาะสม แต่ไม่ชัดเจน	ไม่เหมาะสม
กลไกการทำงาน (mechanics)	ผิดพลาดเล็กน้อย	ผิดพลาดหลายจุด	ผิดพลาดเป็นส่วนมาก

5. การอภิปรายผลการประเมิน (discussion evaluated product)

ขั้นตอนนี้ในกระบวนการประเมินเพิ่มผลงานเป็นการให้เหตุผลว่า เหตุใดจึงได้รับผลการประเมินเอกสารชิ้นงานหรือเพิ่มผลงานเช่นนั้น การอภิปรายผลการประเมินประกอบด้วยผู้เกี่ยวข้องกับการประเมิน ซึ่งผลจากการอภิปรายร่วมกันนี้จะทำให้ได้ข้อเสนอแนะที่เป็นส่วนสำคัญให้เกิดการพัฒนาปรับปรุงแก้ไขที่ตรงจุดมากขึ้น อนึ่งในขั้นตอนนี้ได้นำเอาหลักการวิวิธวิทยาการตีความ (hermeneutics) มาใช้ในการพิจารณาผลการประเมินเพิ่มผลงานด้วย

6. การวางแผนปรับปรุงและพัฒนา (planning to development)

ในขั้นตอนนี้จะนำข้อมูลสารสนเทศที่ได้จากการอภิปรายผลการประเมิน มาวางแผนการดำเนินการเพื่อปรับปรุงแก้ไขเพิ่มผลงานในจุดที่ยังบกพร่องอยู่ให้สมบูรณ์ยิ่งขึ้น ขั้นตอนนี้มีการประชุม

ปรึกษาหารือกันเช่นเดียวกับขั้นตอนการ กำหนดจุดมุ่งหมาย และทิศทางการประเมิน ผู้ร่วมประชุมปรึกษาหารือ ประกอบด้วยผู้เกี่ยวข้องซึ่งมีผลได้ผลเสียกับการประเมินเพิ่มผลงาน โดยทั่วไปจะประกอบด้วยผู้ทำเพิ่มผลงาน ผู้ประเมินเพิ่มผลงานเป็นหลัก

ในการประเมินผลทางการศึกษาที่ใช้การประเมินเพิ่มผลงานนี้มีขั้นตอนในการดำเนินงานหลัก 6 ขั้นตอนดังที่กล่าวมาแล้วอย่างไรก็ตามก็มีนักประเมินหลายท่านที่ทำการประเมินเพิ่มผลงานโดยไม่ยึดติดกับขั้นตอนที่กล่าวมาข้างต้นทั้งหมด จึงสรุปได้ว่าการประเมินเพิ่มผลงานสามารถทำได้หลายแนวทาง ดังเช่นที่ Burke, Fogarty และ Belgrad (1994) เสนอขั้นตอนการประเมินเพิ่มผลงาน 10 ขั้นตอน ดังนี้ 1) ขั้นตอนการกำหนดจุดมุ่งหมายในการทำเพิ่มผลงาน (project purpose) 2) ขั้นตอนการรวบรวมและจัดระบบเอกสารชิ้นงาน (collect and organize) 3) การคัดเลือกเอกสารชิ้นงาน (select key artifacts) 4) การสร้างสรรค์ตกแต่งเพิ่มผลงาน (interject personality) 5) การสะท้อนความคิดโดยใช้สติปัญญาขั้นสูงต่อเอกสารชิ้นงาน (reflect metacognitively) 6) การตรวจสอบเพื่อประเมินตนเอง (inspect to self assess) 7) การปรับปรุงแก้ไขและประเมินเอกสารชิ้นงาน (perfect and evaluate) 8) การปรึกษาและแลกเปลี่ยนประสบการณ์กับเพื่อนร่วมชั้น (connect and conference) 9) การปรับเปลี่ยนเอกสารชิ้นงาน (inject/eject to update) 10) การนำเสนอเพิ่มผลงานที่สมบูรณ์ (respect accomplishments)

ขั้นตอนที่เสนอโดย Burke, Fogarty และ Belgrad (1994) เป็นขั้นตอนในการทำเพิ่มผลงานของนักเรียน ซึ่งจะช่วยให้เกิดการพัฒนาคำรู้ความสามารถของนักเรียน

1.5 ประโยชน์เพิ่มผลงาน

เมื่อมีการปฏิบัติการประเมินเกิดขึ้น การประเมินเพิ่มผลงานถูกเสนอให้เป็นทางเลือกหนึ่งในการประเมินผลทางการศึกษาก่อนการนำมาประเมินเพิ่มผลงานไปใช้อย่างแพร่หลาย ได้มีนักประเมินหลายกลุ่มที่ศึกษาวิเคราะห์ถึงผลกระทบที่เกิดจากการใช้การประเมินเพิ่มผลงานในวงการศึกษาทั้งผลได้และผลเสีย ที่จะเกิดกับผู้ทำเพิ่มผลงาน บุคคลที่เกี่ยวข้องรวมทั้งสถาบันหรือองค์กรด้วย ผู้วิจัยนำเสนอความคิดเห็นและผลการศึกษาที่สรุปประโยชน์ของเพิ่มผลงาน โดยนักประเมินต่อไปนี้

LeMahieu (1995) สรุปว่า เพิ่มผลงานช่วยให้ผู้ทำเพิ่มผลงานมีความรับผิดชอบในการเรียนและนอกจากนี้ยังส่งผลต่อเนื่องถึงครูผู้สอนและสถาบัน โดยการเพิ่มผลงานทำให้ครูผู้สอนมีความช่วยเหลือกันทางด้านความคิด มีการประชุมปรึกษาหารือกันมากขึ้น ซึ่งมีผลทำให้เกิดการพัฒนาหลักสูตรการเรียนการสอนที่มีประสิทธิภาพการทำเพิ่มผลงานทำให้ผู้ทำเพิ่มผลงานมีโอกาสได้แสวงหาความสามารถซึ่งเป็นจุดเด่นทำให้ครูสามารถพัฒนานักเรียนซึ่งเป็นผู้ทำเพิ่มผลงานได้ตรงกับสภาพของผู้ทำเพิ่มผลงานมากขึ้น

นอกจากนี้เพิ่มผลงานยังเป็นที่รวบรวมข้อมูลที่เป็นประโยชน์ต่อการพัฒนากระบวนการเรียนการสอน ผู้ทำเพิ่มผลงานมีความเข้าใจและมีจุดมุ่งหมายในการทำงาน ส่งผลให้เกิดผลผลิตทางการศึกษาที่ตรงตามวัตถุประสงค์

Moss, Schutz และ Collins (1998) สรุปประโยชน์ที่ได้จากการประเมินเพิ่มผลงานว่า เป็นการประเมินที่มีลักษณะดังนี้ 1) การประเมินเพิ่มผลงานมีความละเอียดรอบคอบซึ่งตรงตามวัตถุประสงค์ในการประเมินผล 2) เป็นการประเมินที่ประกอบด้วย การประเมินหลายครั้ง ซึ่งการประเมินเป็นเพียงครั้งเดียวทำให้ได้สารสนเทศไม่เพียงพอสำหรับการสรุปผลการประเมิน 3) การประเมินเพิ่มผลงานเป็นการประเมินที่สนับสนุนให้มีการใช้ผลจากการสะท้อนความคิดมาช่วยในการพัฒนาการปฏิบัติงาน 4) กระบวนการในการประเมินเพิ่มผลงานก่อให้เกิดการพัฒนาวิชาชีพ

Doner และ Gilman (1998) ให้ความเห็นว่าเพิ่มผลงานจะก่อให้เกิดประสบการณ์ที่เป็นประโยชน์ต่อการพัฒนาความรู้ของผู้ทำเพิ่มผลงานสร้างความเชื่อมั่นในตนเอง เกิดความสัมพันธ์ที่ดีระหว่างครูกับผู้ทำเพิ่มผลงาน และระหว่างผู้ทำเพิ่ม อันจะนำไปสู่การมีทักษะที่เป็นพื้นฐานในการทำงาน รวมทั้งทำให้มีเจตคติที่ดีต่อวิชาชีพ

และจากผลการศึกษาของ Doner และ Gilman ที่ศึกษาปฏิกริยาของนักเรียนต่อการประเมินเพิ่มผลงาน ได้ผลสรุปว่าเพิ่มผลงานมีลักษณะที่เป็นประโยชน์แก่ผู้ทำเพิ่มผลงานที่เป็นนักเรียน 9 ประการ คือ 1) ให้ข้อมูลที่สะท้อนความเป็นจริงของผู้ทำเพิ่มผลงานมากกว่าข้อมูลที่ได้จากแบบทดสอบ 2) ให้ความรู้เกี่ยวกับวิธีการสอนของครู 3) ก่อให้เกิดทักษะและพัฒนาทักษะ 4) มีเจตคติที่ดีต่อการเรียนมากขึ้น 5) มีความต่อเนื่องในการปฏิบัติงานและก่อให้เกิดความเชื่อมั่นในตนเอง 6) ปฏิสัมพันธ์ในห้องเรียนก่อให้เกิดการปรับปรุงเปลี่ยนแปลง 7) ได้ฝึกการจัดทำเพิ่มผลงาน 8) เพิ่มผลงานเชิงวิชาชีพเป็นประโยชน์ต่อการสมัครทำงาน 9) เพิ่มผลงานเป็นที่ยืนยันความรู้ความสามารถของผู้ทำเพิ่มผลงาน

Stecher (1998) กล่าวว่าเพิ่มผลงานวิจัยจำนวนมากที่สนับสนุนว่าการประเมินเพิ่มผลงานมีอิทธิพลในทางบวก ส่วนเกณฑ์การประเมินเพิ่มผลงานก็ส่งผลให้เกิดการพัฒนาปรับปรุงเจตคติและความเชื่อของครูและก่อให้เกิดการปรับปรุงเปลี่ยนแปลงหลักสูตรและการสอน นอกจากนี้ยังเป็นการพัฒนาการประเมินระดับห้องเรียนและทำให้เกิดความต่อเนื่องในการปฏิบัติงานของผู้ทำเพิ่มผลงาน ผลพลอยได้สำหรับการประเมินเพิ่มผลงาน เช่น ทำให้ครูผู้สอนปรับปรุงรูปแบบประเมินระดับห้องเรียน เกิดความสัมพันธ์ระหว่างครูกับนักเรียน และครูกับผู้ปกครอง

เพิ่มผลงานก่อให้เกิดประโยชน์สำหรับวงการศึกษาดังกล่าวมาแล้ว เมื่อพิจารณาถึงประโยชน์ที่เกิดขึ้นสำหรับผู้ทำเพิ่มผลงานจะพบว่า ก่อให้เกิดการพัฒนาตนเองในหลาย ๆ ด้าน เช่น ทักษะการปฏิบัติงาน ความมานะพยายาม ความคิดริเริ่มสร้างสรรค์ นอกจากนี้ยังทำให้ผู้ทำเพิ่มผลงานได้พัฒนาศักยภาพในการทำงานของตน เกิดความชื่นชมในผลงาน สำหรับผู้ประเมินเพิ่มผลงานทำให้ทราบถึงความสามารถและพฤติกรรมของผู้ทำเพิ่มผลงาน รวมทั้งอาจทำให้เข้าใจผู้ทำเพิ่มผลงานมากขึ้นจากการอ่านบันทึกการประเมินตนเองหรือการสะท้อนความรู้สึกต่อการทำงาน ข้อมูลเหล่านี้จะเป็นประโยชน์ต่อการพัฒนาผู้ทำเพิ่มผลงานได้ตรงตามความสามารถและความต้องการของผู้ทำเพิ่มผลงานมากขึ้น

เพิ่มผลงานเชิงวิชาการจะทำให้ผู้ทำเพิ่มผลงานมีความรู้ความเข้าใจในเนื้อหาสาระทั้งในด้านทฤษฎีและการปฏิบัติ ส่วนเพิ่มผลงานเชิงวิชาชีพจะทำให้ผู้ทำเพิ่มผลงานมีทักษะและความเชี่ยวชาญในการปฏิบัติงานอาชีพ ก่อให้เกิดประสบการณ์ที่ดีซึ่งจะสามารถนำไปใช้ในชีวิตประจำวันได้ ดังนั้นสรุปได้ว่าเพิ่มผลงานมีประโยชน์สำหรับผู้ทำเพิ่มทั้งในเชิงวิชาการและวิชาชีพ

ตอนที่ 2 การประเมินอภิมาน (Metaevaluation)

ศาสตร์ทางการประเมินได้รับการยอมรับมากขึ้นในปัจจุบันเพราะความสำคัญและความจำเป็นของการปฏิบัติงานต่าง ๆ ที่ต้องมีการประเมินผลการดำเนินงาน แต่วิทยาการที่แข็งแกร่งต้องมีการตรวจสอบและบ่งบอกถึงจุดบกพร่องของการประเมินได้ (สุวิมล ว่องวานิช, 2538) Stufflebeam (1981) กล่าวว่า การประเมินที่ดีจะต้องมีการประเมินกระบวนการ กล่าวคือ การประเมินจะต้องมีการตรวจสอบว่ามีปัญหาหรือไม่ เช่น มีความลำเอียง มีความคลาดเคลื่อนทางเทคนิค หรือความยุ่งยากในการบริหาร มีค่าใช้จ่ายในการประเมินที่สูงเกินไป รวมทั้งการนำผลการประเมินไปใช้ถูกต้องหรือไม่ ผลจากการตรวจสอบการประเมินจะช่วยนำไปสู่การปรับปรุงกระบวนการประเมินให้มีประสิทธิภาพมากขึ้น และได้ผลจากการประเมินงานประเมินที่มีคุณภาพใช้รายงานต่อสาธารณชนให้ทราบถึงจุดแข็งจุดอ่อนของกระบวนการประเมิน แนวคิดในการตรวจสอบหรือการประเมินงานประเมินได้รับการพัฒนาอย่างต่อเนื่องตั้งแต่ปลายคริสต์ทศวรรษที่ 1960 ในระยะแรกศาสตร์ทางการประเมินงานประเมินปรากฏในรูปของการกำหนดเกณฑ์หรือมาตรฐานที่จะใช้ในการตรวจสอบคุณภาพของงานประเมินจากนั้นมีการพัฒนาเทคนิคเครื่องมือและกิจกรรมเพื่อใช้ในการตรวจสอบคุณภาพของงานประเมินให้มีความชัดเจนมากขึ้น ถึงแม้เทคนิคการประเมินงานประเมินจะได้รับการพัฒนาในระยะหนึ่งแล้ว แต่ศัพท์คำว่า การประเมินอภิมาน หรือ metaevaluation ได้มีการบัญญัติศัพท์ขึ้นใช้เป็นครั้งแรกโดย Scriven ในปี ค.ศ. 1969 นอกจากจะบัญญัติศัพท์ขึ้นใช้แล้ว Scriven ยังได้สรุปหลักการและวิธีวิทยาการประเมินงานประเมินหรือการประเมินอภิมานอย่างเป็นระบบขึ้นด้วย (อ้างใน Stufflebeam, 1981; Nyirenda, 1994; Finn, Stevens, Stufflebeam และ Walberg, 1997) นับแต่นั้นมาวิธีวิทยาการด้านการประเมินอภิมานหรือการประเมินงานประเมินได้มีพัฒนาการก้าวหน้าขึ้นเป็นลำดับ

การเสนอรายงานเอกสารที่เกี่ยวข้องกับการวิจัยในหัวข้อการประเมินอภิมานนี้ ผู้วิจัยนำเสนอสาระด้านวิธีวิทยาการประเมินอภิมานแยกเป็น 6 หัวข้อ ดังนี้ ความหมายการประเมินอภิมาน วัตถุประสงค์การประเมินอภิมาน ประเภทการประเมินอภิมาน ขั้นตอนการประเมินอภิมาน มาตรฐานที่ใช้ในการประเมินอภิมาน และประโยชน์ของการประเมินอภิมาน

2.1 ความหมายการประเมินอภิมาน

นับตั้งแต่ Scriven ได้บัญญัติศัพท์คำว่า metaevaluation และให้ความหมายว่า เป็นการประเมินงานประเมิน (evaluation of evaluations) ในปี ค.ศ. 1969 แล้ว มีนักประเมินได้พัฒนาวิธีการวิทยาการประเมินและให้ความหมายของการประเมินอภิมานไว้แตกต่างกันหลากหลาย Stufflebeam (1981) ได้ศึกษาวิเคราะห์ความหมายและผลงานของการประเมินอภิมานที่นักประเมินหลายท่านได้กำหนดไว้ และแยกความหมายของการประเมินอภิมานเป็น 5 กลุ่ม กลุ่มแรกคือ Leon Lessinger (1970), Malcolm Provus (1971), Richard Seligman (1973) และคนอื่น ๆ ซึ่งให้คำนิยามว่า การประเมินอภิมานเป็นการตรวจสอบโครงการประเมิน (evaluation of program auditing) ประกอบด้วย การกำกับดูแลกระบวนการประเมิน เพื่อให้ได้ผลการประเมินที่มีคุณภาพและการประเมินค่าความตรงของผลการประเมินและข้อเสนอแนะจากการประเมิน รวมทั้งการนำเสนอรายงานต่อสาธารณชน กลุ่มที่ 2 คือ Cambell และ Stanley (1996), Bracht และ Glass (1968) ให้คำนิยามว่า การประเมินอภิมานเป็นกระบวนการตัดสินคุณค่าของการประเมินซึ่งรวมทั้งการนำเสนอเกณฑ์ สำหรับการตัดสินคุณค่าการประเมิน นักประเมินในกลุ่มนี้ให้ความสำคัญกับการพัฒนาเกณฑ์สำหรับการประเมินคุณภาพงานประเมิน กลุ่มที่ 3 คือ Stufflebeam (1974, 1978) เป็นผู้สร้างมโนทัศน์ของการประเมินอภิมานและให้แบบตรวจสอบรายการ (checklist) สำหรับการกำกับประเมินอภิมานเชิงก้าวหน้า (formative metaevaluation) และให้ รายการมาตรฐานสำหรับการประเมินอภิมานเชิงสรุป (summative metaevaluation) รวมทั้งเสนอขั้นตอนการดำเนินงานการประเมินอภิมานด้วย กลุ่มที่ 4 คือ Scriven (1976) และ Millman (1979) เสนอหลักการตีค่าและการควบคุมความลำเอียงในการประเมินที่ช่วยให้ให้นักประเมินทำการประเมินได้โดยอิสระ รวมทั้งช่วยให้นักประเมินทำการประเมินได้อย่างมีคุณภาพ และกลุ่มสุดท้าย คือ Gowin (1978) ได้พัฒนาเทคนิคการทำแผนที่มโนทัศน์ (concept mapping) สร้างโมเดลที่ชื่อว่า QUEMAC สำหรับการประเมินงานประเมิน

กล่าวโดยสรุปจากที่นักประเมินได้อธิบายแนวคิดหลักการ รวมทั้งให้ความหมายของการประเมินอภิมานไว้ข้างต้น การประเมินจึงเป็นกระบวนการตรวจสอบ (audit) กำกับดูแล (monitor) ประเมินค่าความตรง (validity assessment) และตัดสินคุณค่า (value judgment) โครงการประเมินเพื่อให้ได้ข้อเสนอแนะทราบถึงจุดเด่นและจุดด้อยของโครงการและนำมาใช้ปรับปรุงโครงการให้ดีขึ้น รวมทั้งจัดทำรายงานเสนอต่อสาธารณชนให้ทราบถึงคุณภาพของโครงการและผลการประเมินโครงการที่ถูกต้อง

อย่างไรก็ดี นักประเมินยังมีความคิดเห็นที่แตกต่างกันอยู่บ้างในระยะแรก เช่น Anderson, Ball, Murphy และ คนอื่น ๆ (1975) ได้บัญญัติศัพท์และให้ความหมายของการประเมินทุติยมาน (secondary evaluation) ว่าเป็นการประเมินที่นักประเมินใช้ข้อมูลจากงานประเมินหลาย ๆ ชิ้น และ/หรือตัวรายงานการประเมินเพื่อสรุปผลการประเมิน Anderson, Ball, Murphy และคนอื่น ๆ (1975) กล่าวเพิ่มเติมว่าการประเมินทุติยมาน (secondary evaluation) มีความหมายเกี่ยวข้องกับการประเมินอภิมาน

(metaevaluation) และมีความหมายใกล้เคียงกับการตรวจสอบการประเมิน (audit of evaluation) จุดที่แตกต่างกันคือ การตรวจสอบต้องใช้ผู้เชี่ยวชาญหลายสาขา แต่การประเมินหุติยมานอาจเป็นแต่เพียงรายงานบริษัทของงานประเมินโดยมิได้ทำการประเมินงานประเมินเหล่านั้น กล่าวคือ การประเมินหุติยมานอาจจะเป็นรายงานการสังเคราะห์งานประเมินโดยไม่มีการวิเคราะห์และกำหนดเกณฑ์แต่อย่างใด หรืออาจจะเป็นการนำข้อมูลจากงานประเมินหลาย ๆ ชิ้นมาวิเคราะห์ซ้ำเพื่อจะสรุปผลการประเมินให้มีความถูกต้องชัดเจนมากขึ้น

จากความหมายที่กล่าวมาข้างต้น จะเห็นว่าความหมายของการประเมินอภิมานใกล้เคียงกับความหมายของการวิเคราะห์อภิมาน (meta analysis) ซึ่ง Glass (1976) เป็นผู้บัญญัติศัพท์นี้ขึ้น Glass กล่าวว่า การวิจัยหรือการวิเคราะห์ในการวิจัยมี 3 ประเภท ประเภทที่หนึ่ง คือ การวิเคราะห์ปฐมมาน(primary analysis) ซึ่งหมายถึงการวิจัยหรือการวิเคราะห์ที่นักวิจัยเก็บรวบรวมข้อมูลด้วยตนเองและนำมาวิเคราะห์เป็นครั้งแรก ประเภทที่สอง คือ การวิเคราะห์หุติยมาน (secondary analysis) เป็นการวิเคราะห์ที่นักวิจัยใช้ข้อมูลที่มีนักวิจัยอื่นได้รวบรวมขึ้นไว้ก่อนแล้วมาทำการวิเคราะห์ซ้ำด้วยวิธีวิทยาใหม่ เพื่อตอบปัญหาการวิจัยให้ชัดเจนขึ้น และประเภทที่สาม คือ การวิเคราะห์อภิมาน (meta analysis) ซึ่งหมายถึงการวิจัยที่นักวิจัยทำการสังเคราะห์งานวิจัยจากรายงานการวิจัย โดยใช้เป็นข้อมูลในการวิจัย แล้วทำการวิเคราะห์ด้วยวิธีการเชิงปริมาณ จะเห็นได้ว่าความหมายของการวิเคราะห์ประเภทที่สามคือการวิเคราะห์อภิมานนั้นมีความหมายใกล้เคียงกับการประเมินอภิมาน (metaevaluation) กล่าวคือการวิเคราะห์อภิมานเป็นการวิเคราะห์ผลการวิเคราะห์ (analysis of analyses) ส่วนการประเมินอภิมานเป็นการประเมินงานประเมิน (evaluation of evaluations)

2.2 วัตถุประสงค์ของการประเมินอภิมาน

นักประเมินกำหนดวัตถุประสงค์ของการประเมินอภิมานไว้แตกต่างกันตามลักษณะของโครงการประเมินและความต้องการนำไปใช้ประโยชน์ แต่มีวัตถุประสงค์สำคัญตรงกัน ซึ่งผู้วิจัยสรุปได้ว่า โครงการที่มีการประเมินโดยใช้การประเมินอภิมานหรือการประเมินงานประเมินมีวัตถุประสงค์สำคัญเพื่อรายงานผลการประเมินโครงการให้ผู้ใช่ผลการประเมินและผู้เกี่ยวข้องสามารถประเมินคุณค่าของโครงการนั้นได้อย่างถูกต้องและได้สารสนเทศสำหรับผู้ใช่ผลการประเมินและผู้เกี่ยวข้องไปปรับปรุงการดำเนินงานให้ดีขึ้น (Stufflebeam, 1981; Gallegos, 1994; Finn, Stevens, Stufflebeam และ Walberg, 1997; Hill, 2541)

ผู้ใช่ผลการประเมิน (audiences) หรือผู้บริโภคผลการประเมิน (evaluation consumer) หรือ ผู้เกี่ยวข้องกับการประเมินในการประเมินอภิมานมีจำนวนหลายกลุ่ม ผู้ใช่ผลการประเมินในโครงการคณะผู้เชี่ยวชาญการประเมินอภิมาน (Project Metaevaluation Panel) ของ Finn และคนอื่น ๆ (1997) ซึ่งเป็นการประเมินงานประเมินโครงการระบบการเรียนรู้แบบบูรณาการ (Integrated Learning Systems Project)

ของโรงเรียนในนิวยอร์กนั้นมีหลายกลุ่ม คณะทำงานสรุปว่าผู้ใช้ผลการประเมินนิเทศประกอบไปด้วยโรงเรียนที่อยู่ในโครงการ เจ้าหน้าที่เขตการศึกษาที่โรงเรียนนั้นตั้งอยู่สำนักงานคณะกรรมการการศึกษาของเมืองนิวยอร์ก ผู้ปกครองนักเรียน ผู้ปกครองนักเรียน ผู้นำชุมชน สถาบันวอเตอร์ฟอร์ด (Waterford Institute) สถาบันวิกัท (Wicat Institute) สมาคมพัฒนาหลักสูตร ผู้ให้บริการทางการศึกษา ผู้สนับสนุนเครื่องมือในการประเมินนิเทศ ผู้ใช้ผลการประเมินในระดับนานาชาติและสมาชิกชุมชนนักวิจัย จะเห็นได้ว่ากลุ่มผู้ใช้ผลการประเมินนิเทศดังกล่าวนี้แยกได้เป็น 2 กลุ่ม คือ กลุ่มที่มีส่วนเกี่ยวข้องโดยตรงกับโครงการระบบการเรียนรู้แบบบูรณาการ และกลุ่มสมาชิกในชุมชนหรือผู้เกี่ยวข้อง ดังนั้นการกำหนดวัตถุประสงค์ของการประเมินนิเทศจึงเน้นการได้สารสนเทศสำหรับผู้ใช้งานผลการประเมินนิเทศทั้ง 2 กลุ่ม กล่าวคือกลุ่มที่มีส่วนเกี่ยวข้องโดยตรงกับโครงการระบบการเรียนรู้แบบบูรณาการจะได้สารสนเทศที่ใช้ในการพัฒนาโครงการให้ดีขึ้น และผู้ที่ทำหน้าที่ในการประเมินจะได้รับสารสนเทศที่ใช้ในการปรับปรุงกระบวนการประเมินให้มีความชัดเจนและถูกต้องมากยิ่งขึ้น ส่วนกลุ่มสมาชิกในชุมชนที่เกี่ยวข้องจะได้รับทราบผลการประเมินจุดเด่น-จุดด้อยที่เป็นคุณค่าทางวัตถุ (worth) และคุณค่าทางนามธรรม (merit) ของโครงการระบบการเรียนรู้แบบบูรณาการโดยสมบูรณ์

Hill (2541) ได้ให้คำอธิบายถึงหน้าที่ (function) ของการประเมินนิเทศซึ่งสะท้อนให้เห็นวัตถุประสงค์ของการประเมินนิเทศไว้ 4 ด้าน ซึ่งแนวคิดของ Hill (2541) สอดคล้องกับแนวคิดเกี่ยวกับ บทบาท (role) ของการประเมินที่ Stufflebeam (1981) เสนอไว้ ผู้วิจัยนำมาสรุปได้เป็น 4 ประการ คือ

1. หน้าที่บทบาทในการให้แนวทาง (guideline) ในการกำกับ (monitoring) และการปรับปรุง (improving) การประเมิน หน้าที่บทบาทในหัวข้อนี้ ก็คือการประเมินนิเทศเชิงก้าวหน้า (formative metaevaluation) ซึ่งจะให้สารสนเทศสำหรับผู้ประเมินโครงการได้ปรับปรุงและพัฒนากระบวนการประเมินให้มีประสิทธิภาพดียิ่งขึ้น

2. หน้าที่บทบาทในการเสนอรายงาน (reporting) เกี่ยวกับจุดเด่นและจุดด้อยของโครงการ การตรวจสอบความตรง (validating) และการเพิ่มความน่าเชื่อถือ (adding credibility) ของการประเมินบทบาทหน้าที่ในข้อนี้ก็คือการประเมินนิเทศเชิงสรุป (summative metaevaluation) ซึ่งจะให้ผลของการประเมินนิเทศที่เป็นประโยชน์ต่อผู้ใช้ผลการประเมินและเป็นการเผยแพร่คุณค่าของโครงการต่อสาธารณชนด้วย

3. หน้าที่บทบาทในการส่งเสริม (promoting) และกระตุ้น (motivating) ทางจิตวิทยา สังคมและการเมือง (psychological and socio political) ซึ่งจะเป็ประโยชน์ต่อการประเมินนิเทศ หน้าที่บทบาทในข้อนี้เป็นกรกระตุ้นให้ผู้ใช้ผลการประเมินเพิ่มความตระหนักในคุณค่าของการประเมิน และกระตุ้นให้เกิดการตอบสนองต่อผลการประเมิน รวมทั้งส่งเสริมให้มีการสื่อสารประชาสัมพันธ์ระหว่างผู้ดำเนินการโครงการกับสาธารณชนผู้เกี่ยวข้อง

4. หน้าที่บทบาทในการควบคุม (controlling) กระบวนการประเมินงานประเมิน หน้าที่บทบาทในข้อนี้เป็นหน้าที่ของผู้บริหารที่จะต้องติดตาม กำกับการประเมินของผู้ทำการประเมินให้เป็นไปอย่างถูกต้อง

2.3 ประเภทการประเมินอภิमान

Stufflebeam (1981) และ Gallegos (1994) แบ่งการประเมินอภิमानตามหน้าที่บทบาทของการประเมินเป็น 2 ลักษณะ คือ การประเมินอภิमानเชิงก้าวหน้า และการประเมินอภิमानเชิงสรุป โดย Stufflebeam ได้ให้คำอธิบายเพิ่มเติมในรายละเอียดดังนี้

1. การประเมินอภิमानเชิงก้าวหน้า (formative metaevaluation)

การประเมินอภิमानเชิงก้าวหน้าเป็นการประเมินงานที่มีลักษณะการใช้การประเมินเชิงสร้างสรรค์ โดยเป็นการวางแผนก่อนปฏิบัติการประเมิน การประเมินอภิमानเชิงความก้าวหน้าช่วยให้นักประเมินตัดสินใจเกี่ยวกับสิ่งที่จะประเมินว่าจะประเมินอะไร นิยามปัญหาที่จะประเมินอย่างไร การกำหนดผู้ใช้ผลการประเมินและผู้เกี่ยวข้องกับการประเมิน การทำให้วัตถุประสงค์ของการประเมินชัดเจน รวมทั้งข้อตกลงสัญญาและเงื่อนไขในการประเมิน กำหนดวิธีรวบรวมข้อมูล วิธีวิเคราะห์ข้อมูล กำหนดตารางของการประเมิน จัดหาทีมประเมิน กำหนดงบประมาณ เตรียมข้อสรุปและข้อเสนอแนะ และเลือกรูปแบบและวิธีการรายงานผลการประเมิน

โดยสรุปการประเมินอภิमानเชิงก้าวหน้า หมายถึง การประเมินว่าแผนการประเมินและการดำเนินการประเมินเป็นไปตามแนวทางที่กำหนดในมาตรฐานหรือไม่ ถ้าปฏิบัติตามแนวทางก็จะได้ผลการประเมินที่ถูกต้องสมเหตุสมผล อีกทั้งเป็นการหลีกเลี่ยงหลุมพรางของการประเมินที่สามารถทำให้เกิดข้อผิดพลาดในการประเมินได้ ทั้งนี้ข้อตกลงเบื้องต้นที่สำคัญคือคุณค่าสูงสุดของการประเมินจะมีค่ามากที่สุดเมื่อการดำเนินการประเมินเป็นไปตามแนวทางที่กำหนดเป็นมาตรฐานไว้และสามารถหลีกเลี่ยงข้อผิดพลาดต่าง ๆ ที่ถือว่าเป็นหลุมพรางของการประเมินได้

2. การประเมินอภิमानเชิงสรุป (summative metaevaluation)

การประเมินอภิमानเชิงสรุปถือว่าเป็นส่วนสำคัญอีกส่วนของการประเมินอภิमानเพราะได้แสดงให้เห็นถึงคุณค่ารวมของงานการประเมิน ทำให้มองเห็นว่าการประเมินนี้เป็นไปตามเกณฑ์มาตรฐานของการประเมินหรือไม่ พร้อมทั้งจะให้ตรวจสอบและมีการรายงานการประเมินผลการประเมินให้ผู้เกี่ยวข้องและสาธารณะชนได้ทราบถึงความเป็นมาตรฐานของงานประเมิน ผลการประเมินอภิमानเชิงสรุปช่วยให้ผู้ใช้ผลการประเมินมีความมั่นใจและตัดสินใจได้ว่าควรจะยอมรับผลประเมินและข้อเสนอแนะจากงานประเมินมากน้อยเพียงใด

Worthen และ Sanders (1987, อ้างใน Hill, 2541) แบ่งการประเมินอภิमानตามประเภทของผู้ทำการประเมินเป็น 3 แบบ คือ

1. การประเมินอภิมานโดยนักประเมินที่เป็นเจ้าของโครงการ (original evaluators)

การประเมินประเภทนี้ถูกโต้แย้งว่าผลการประเมินมีความลำเอียงเกิดขึ้นมาก เพราะผู้ประเมินที่เป็นเจ้าของโครงการมีความต้องการให้ผลการประเมินออกมาในแง่บวกมากกว่าแง่ลบ ดังนั้นจึงมีข้อเสนอแนะว่าควรใช้นักประเมินที่ไม่ใช่เจ้าของโครงการ การประเมินอาจไม่เป็นรูปแบบชัดเจนเป็นแค่การวิพากษ์วิจารณ์เกี่ยวกับโครงการของเพื่อน หรือคนรู้จักอื่น ๆ ก็ได้ ซึ่งก็น่าจะให้ผลการประเมินที่ปราศจากความลำเอียงได้มากกว่าให้เจ้าของโครงการเป็นผู้ประเมิน

2. การประเมินอภิมานโดยกลุ่มผู้บริโภคผลการประเมิน (evaluation consumers)

มีการประเมินหลายโครงการที่พบว่าผลการประเมินไม่ได้ตอบสนองต่อความต้องการของผู้ใช้ผลการประเมิน เช่น ผู้ให้การสนับสนุนโครงการ (sponsor) ผู้ต้องการใช้สารสนเทศจากการประเมิน (client) หรือผู้เกี่ยวข้องอื่น ๆ (other stakeholders) ดังนั้นการประเมินอภิมานประเภทนี้จึงให้ความสำคัญต่อกลุ่มคนผู้ใช้ผลการประเมิน โดยให้ทำการประเมินอภิมานด้วย แต่ในการประเมินกลับพบว่ามีปัญหาเรื่องการประเมินซึ่งกลุ่มผู้ใช้ผลการประเมินไม่มีความเชี่ยวชาญในเทคนิควิธีการ เครื่องมือ มาตรฐานและความถูกต้องของการประเมิน ทางแก้ไขปัญหานี้ให้นักประเมินมืออาชีพเสนอว่าควรมีการกำหนดมาตรฐานเครื่องมือการวัดที่มีความตรงรวมทั้งแนวทางการวิเคราะห์ข้อมูลให้ชัดเจน

3. การประเมินอภิมานโดยนักประเมินมืออาชีพ (competent evaluators)

การประเมินประเภทนี้นับว่าเป็นการประเมินที่ดีที่สุด เนื่องจากผู้ประเมินมีความเชี่ยวชาญ มีความรู้ความสามารถในวิธีการประเมิน และเลือกใช้วิธีการที่เหมาะสม อย่างไรก็ตามการประเมินอภิมานโดยนักประเมินมืออาชีพนี้ Brinkerhoff และคนอื่น ๆ (1983 อ้างใน Hill, 2541) ได้ให้คำแนะนำว่า นักประเมินมืออาชีพที่ทำการประเมินอภิมานนั้นควรเป็นบุคคลภายนอกที่ไม่เกี่ยวข้องกับโครงการ จะทำให้ได้ผลการประเมินที่น่าเชื่อถือมากกว่าการให้บุคคลภายในโครงการทำการประเมินอภิมานเองและการประเมินอภิมานโดยที่ให้นักประเมินจะมีการใช้ทักษะความรู้ความเชี่ยวชาญในการประเมินมากกว่าการประเมินที่ใช้ผู้ประเมินคนเดียว

ในทัศนะของผู้วิจัยมีความเห็นเพิ่มเติมในประเด็นการแบ่งประเภทของการประเมินอภิมานตามแบบของ Hill (2541) ว่าในความเป็นจริงอาจไม่มีลักษณะต่างกันเป็น 3 แบบอย่างเด็ดขาด เนื่องจากในการประเมินอภิมานผู้ทำการประเมินอาจทำการประเมินร่วมกันโดยมีเจ้าของโครงการ ผู้บริโภคผลการประเมิน และนักประเมินมืออาชีพพร้อมเป็นทีมนักประเมินด้วย และในทำนองเดียวกันการแบ่งประเภทการประเมินอภิมานตามหน้าที่บทบาท ที่ Stufflebeam (1981) เสนอไว้ นั้น ในความเป็นจริงในการประเมินอภิมานอาจทำการประเมินอภิมานผสมผสานทั้งการประเมินอภิมานเชิงก้าวหน้าและการประเมินอภิมานเชิงสรุปด้วยก็ได้ การจัดแบ่งประเภทการประเมินอภิมานดังกล่าวจึงเป็นเพียงการแบ่งประเภทเพื่อให้เห็นลักษณะที่ต่างกันในทางปฏิบัติมักใช้การประเมินอภิมานที่ผสมผสานกันไม่ได้ทำการประเมินประเภทใดประเภทหนึ่งโดยเฉพาะ

2.4 ขั้นตอนการประเมินนิทาน

การนำเสนอสาระเกี่ยวกับขั้นตอนของการประเมินนิทานในหัวข้อนี้ ผู้วิจัยนำเสนอขั้นตอนของการประเมินนิทานตามแนวของนักประเมินที่สำคัญ 4 กลุ่ม กลุ่มแรกคือ Finn, Stevens, Stufflebeam และ Walberg (1997) กลุ่มที่ 2 คือ Hill (2541) กลุ่มที่ 3 คือ Nyirenda (1994) และกลุ่มสุดท้าย คือ Gallegos (1994) ดังรายละเอียดต่อไปนี้

กลุ่มที่หนึ่ง การประเมินนิทานตามโครงการคณะผู้เชี่ยวชาญการประเมินนิทาน (Project Metaevaluation Panel) ซึ่งเป็นการประเมินงานประเมินโครงการระบบการเรียนรู้แบบบูรณาการ (Integrated Learning System) ที่ Finn, Stevens, Stufflebeam และ Walberg (1997) เป็นคณะทำงานรับผิดชอบการประเมิน มีการกำหนดขั้นตอนในการทำงานไว้ 8 ขั้นตอน คือ 1) การทบทวนแผนงานขั้นตอนการดำเนินงานและข้อมูลพื้นฐานของโครงการ 2) การสังเกตการปฏิบัติงานของโรงเรียนในโครงการระบบการเรียนรู้แบบบูรณาการ 3) การประชุมอภิปรายร่วมกันระหว่างคณะทำงานการประเมินนิทานกับผู้เกี่ยวข้องที่เป็นผู้ให้การสนับสนุนโครงการ 4) การวิเคราะห์ผลการประเมินนิทานที่จะทำเป็นรายงาน 5) การจัดทำโครงร่างการรายงานผลการประเมินนิทานรายบุคคลของคณะทำงานการประเมินนิทาน 6) การวิพากษ์วิจารณ์โครงร่างผลการประเมินนิทาน 7) การอภิปรายร่วมกันระหว่างคณะทำงานการประเมินกับกลุ่มผู้บริหารโครงการเพื่อหาข้อตกลงร่วมกัน 8) การแก้ไขปรับปรุงรายงานผลการประเมินนิทานครั้งสุดท้ายก่อนที่จะนำเสนอ

กลุ่มที่สอง เป็นขั้นตอนการประเมินนิทานที่เสนอโดย Hill (2541) การขั้นตอนที่นำเสนอโดย Finn, Stevens, Stufflebeam และ Walberg (1997) แตกต่างจากขั้นตอนที่นำเสนอโดย Hill (2541) กล่าวคือ ขั้นตอนที่เสนอโดย Hill นี้เป็นขั้นตอนใหญ่ 4 ขั้นตอน ซึ่งคล้ายคลึงกับขั้นตอนในการประเมิน โดยทั่วไปที่ส่วนมากมักจะเริ่มต้นการประเมินด้วยการสำรวจความต้องการจำเป็นในการประเมิน ขั้นตอนการประเมินนิทานของ Hill ประกอบด้วยขั้นตอนที่หนึ่งการศึกษาความต้องการจำเป็นเกี่ยวกับสถานการณ์ กิจกรรมและนโยบายของโครงการที่จะประเมิน ขั้นตอนที่สองเป็นการประมวลผลข้อมูลสารสนเทศเกี่ยวกับสถานการณ์ กิจกรรมและนโยบาย ขั้นตอนสามเป็นการแปลความหมายสารสนเทศและการให้คุณค่า สารสนเทศ ขั้นตอนสุดท้ายเป็นการตัดสินคุณค่าเกี่ยวกับสถานการณ์ กิจกรรมและนโยบายของโครงการ จะเห็นว่าขั้นตอนการประเมินนิทานที่เสนอโดย Hill (2541) ไม่ได้ลงลึกในรายละเอียดการปฏิบัติ แต่ก็แสดงให้เห็นถึงกระบวนการของการประเมินนิทานที่จะนำไปสู่การปฏิบัติได้เช่นกัน

กลุ่มที่ 3 เป็นการศึกษารูปแบบในการตรวจสอบคุณภาพของระบบการประเมินความสามารถเพื่อการรับรองวิทยฐานะของครู การวิจัยมุ่งเน้นพัฒนากรอบในการกำหนดเกณฑ์การตรวจสอบที่มีคุณภาพ ประสิทธิภาพและประสิทธิผล นอกจากนี้ยังให้ข้อเสนอแนะในการนำไปใช้และใครควรใช้เกณฑ์การประเมินนี้ กรอบของเกณฑ์ที่พัฒนาขึ้นเพื่อตรวจสอบการประเมิน กำหนดมิติใหญ่ ๆ ที่ต้องประเมินเป็น 2 มิติ คือ

กระบวนการควบคุมคุณภาพของการประเมินและการพัฒนาระบบการประเมิน สำหรับเกณฑ์ที่ใช้ประเมิน 2 มิติที่ประกอบด้วย 1) ความเป็นไปได้ในการบริหาร 2) ความเชื่อถือได้ต่อสาธารณชน 3) ความเชื่อถือได้ทางวิชาชีพ 4) การป้องกันการเมืองของการประเมิน 5) ความสามารถในการปฏิบัติทางเศรษฐศาสตร์

กลุ่มสุดท้ายเป็นขั้นตอนการประเมินอภิมานรูปแบบสุดท้ายที่นำเสนอเพื่อเป็นแนวทางในการนำไปปฏิบัตินี้สรุปจากการประเมินอภิมานซึ่ง Gallegos (1994) ใช้ในการประเมินอภิมานโมเดลที่ใช้ในการประเมินโรงเรียน (Metaevaluation of School Evaluation Model) โดย Gallegos ได้ทำการประเมินอภิมานโมเดลการประเมินที่ใช้ในโรงเรียนของสหรัฐอเมริกา 51 โมเดล ที่มีทั้งโมเดลการประเมินโรงเรียนที่เป็นโมเดลระดับชาติ ระดับเขตการศึกษา ระดับรัฐและระดับท้องถิ่น ขั้นตอนสำคัญของการประเมินอภิมานที่ Gallegos ใช้มี 5 ขั้นตอน คือ 1) การศึกษาทบทวนโมเดลการประเมินและข้อเสนอแนะในผลการประเมินที่เป็นจุดเด่นและจุดด้อยของการประเมินในรายงานการประเมินทั้ง 51 โมเดล 2) การจัดทำเมตริกซ์การประเมินหรือประเด็นการประเมิน (typology/evaluation matrix) 3) การกำหนดตัวบ่งชี้คุณภาพในการประเมินโรงเรียนโดยอาศัยมาตรฐานการประเมินซึ่งกำหนดโดยคณะกรรมการกำหนดมาตรฐานสำหรับการประเมินโครงการทางการศึกษา (Standard for Evaluations of Educational Programs, Project and Materials) ประกอบด้วยมาตรฐาน 4 ด้าน จำนวน 30 มาตรฐาน 4) การประเมินค่าตามตัวบ่งชี้คุณภาพที่กำหนดตามมิติการประเมินในเมตริกซ์การประเมินโดยใช้มาตราประเมินค่า (rating scale) 5 ระดับ โดยรหัส 1 คือ เหมาะสมมาก รหัส 2 คือ เหมาะสม รหัส 3 คือ เหมาะสมบางส่วน รหัส 4 คือ ไม่เหมาะสม และรหัส 5 คือ ข้อมูลไม่เพียงพอ และ 5) การจัดทำรายงานเสนอผลการประเมินอภิมานรวมทั้งข้อเสนอแนะเพื่อการปรับปรุงโมเดลที่ใช้ในการประเมินโรงเรียน

องค์ความรู้ที่สำคัญที่ Gallegos (1994) เสนอไว้ในขั้นตอนการประเมินอภิมาน คือ การกำหนดเมตริกซ์หรือมิติในการประเมิน (evaluationmatrix) และการกำหนดตัวบ่งชี้คุณภาพ (quality indicators) ซึ่งใช้ในการประเมินอภิมานโมเดลที่ใช้ในการประเมินโรงเรียน 51 โมเดล ผู้วิจัยจึงนำเสนอรายละเอียดของขั้นตอนทั้งสองดังกล่าว เพื่อให้เห็นแนวทางในการประเมินอภิมานที่ชัดเจนยิ่งขึ้น

การกำหนดเมตริกซ์หรือมิติในการประเมิน (a typology/evaluation matrix)

การกำหนดเมตริกซ์หรือมิติในการประเมิน ประกอบด้วยมิติหรือประเด็นในการประเมิน 10 มิติ ดังต่อไปนี้

1. ประเภทหรือวัตถุประสงค์ของโมเดลการประเมิน

นักประเมินอภิมานต้องศึกษาวิเคราะห์โมเดลการประเมินที่โรงเรียนใช้ และจำแนกว่าเป็นโมเดลประเภทใด มีวัตถุประสงค์อย่างไร โดยแยกเป็น 3 ประเภทย่อย ๆ คือ

1.1 โมเดลกำกับความสอดคล้อง (compliance monitoring models) โมเดลประเภทนี้หมายถึงโมเดลการประเมินที่โรงเรียนกำหนดให้สอดคล้องกับเกณฑ์มาตรฐานที่กำหนดไว้ก่อนแล้วเป็นโมเดลที่ให้ความสำคัญกับปัจจัยป้อน (input) มากกว่ากระบวนการ (process)

1.2 โมเดลกำกับการวินิจฉัย (diagnostic monitoring models) หมายถึงโมเดลการประเมินที่โรงเรียนให้ความสำคัญกับการพัฒนานักเรียน โดยใช้กระบวนการที่มีการทดสอบก่อนและหลังแบบอิงเกณฑ์เป็นเครื่องมือในการวัดและวินิจฉัยจุดบกพร่องของนักเรียน ลักษณะของโมเดลเป็นแบบเชิงก้าวหน้า (formative) เพื่อวิเคราะห์หาจุดบกพร่องและหาจุดพัฒนานักเรียนให้ดีขึ้น

1.3 โมเดลกำกับการปฏิบัติ (performance monitoring models) หมายถึงโมเดลการประเมินที่ให้ความสำคัญกับผลการประเมินหรือเน้นผลสัมฤทธิ์ทางการเรียนของนักเรียนแบบอิงกลุ่มหรือเน้นโมเดลการประเมินที่เป็นแบบเชิงสรุป (summative)

2. กระบวนการ (process)

นักประเมินนิยามต้องศึกษาวิเคราะห์โมเดลการประเมินว่ากระบวนการที่ใช้ในการประเมินนิยามเป็นแบบใดใน 3 ประเภท คือ เป็นกระบวนการประเมินแบบการประเมินตนเองหรือการศึกษิตนเอง (self-study=SS) หรือเป็นการตรวจเยี่ยม (visitation=V) ของที่มนักประเมินภายนอก หรือเป็นการตรวจสอบโดยหน่วยงานราชการที่เกี่ยวข้อง (desk audit=DA)

3. ระดับ/ขนาดของโรงเรียน (level focus)

นักประเมินนิยามต้องศึกษาวิเคราะห์โมเดลการประเมินว่าเป็นโมเดลที่มุ่งใช้ประเมินโรงเรียนที่เปิดสอนในระดับใดใน 5 ระดับ คือ ประถมศึกษา (elementary schools=E), มัธยมศึกษาตอนต้น (middle schools=M), มัธยมศึกษาตอนปลาย (high schools=H), ทุกระดับ (all levels=DIST) หรือเป็นการประเมินในระดับรัฐ (state=State)

4. สารสนเทศที่ต้องการใช้ (intended focus)

นักประเมินนิยามต้องศึกษาวิเคราะห์โมเดลการประเมินว่าเป็นโมเดลที่ประเมินแล้วให้ข้อมูลหรือสารสนเทศเพื่ออะไรใน 2 ประเภท คือ เพื่อมุ่งการปรับปรุงพัฒนา (formative=Form) หรือเพื่อมุ่งการตัดสินที่มีประสิทธิภาพ (summative=Sum)

5. ประเภทของข้อมูล (data focus)

นักประเมินนิยามต้องศึกษาวิเคราะห์โมเดลการประเมินว่าเป็นโมเดลที่ต้องอาศัยข้อมูลประเภทใดในการวิเคราะห์ คือเป็นข้อมูลเชิงปริมาณ (quantitative=Quan) หรือข้อมูลเชิงคุณลักษณะ (qualitative=Qual)

6. การวางแผนและตัดสินใจ (planning/dicision making)

นักประเมินนิมิตานต้องศึกษาวิเคราะห์โมเดลการประเมินว่าเป็นโมเดลที่มีการวางแผนและตัดสินใจมุ่งเน้นการประเมินส่วนใดในโครงการเป็นหลัก โดยพิจารณาจากซิปโมเดล (CIPP model) ของ Stufflebeam (1983) ซึ่งมีการประเมิน 4 ส่วนที่สำคัญ คือ การประเมินบริบทหรือสภาพแวดล้อม (context evaluation=C) การประเมินปัจจัยป้อน (input evaluation=I) การประเมินกระบวนการ (process evaluation=P1) และการประเมินผลผลิต (product evaluation=P2)

7. ความรับผิดชอบต่อหน้าที่

นักประเมินนิมิตานต้องศึกษาวิเคราะห์โมเดลการประเมินว่าเป็นโมเดลที่ก่อให้เกิดผลของการประเมินอย่างไรใน 4 ประเภทซึ่ง Stufflebeam (1991) แบ่งไว้ คือ การพัฒนา (growth=G) ความเสมอภาค(equity=E) การอำนวยความสะดวก (feasibility=F) และความสามารถดีเด่น (excellence=Ex)

8. มาตรฐานการประเมินนิมิตาน

นักประเมินนิมิตานต้องศึกษาวิเคราะห์โมเดลการประเมินว่าเป็นโมเดลที่มีมาตรฐานการประเมินตรงตามที่กำหนดหรือไม่ มาตรฐานที่ยึดเป็นหลักคือ Standards for Evaluations of Educational Programs, Projects and Materials ซึ่งกำหนดโดยคณะกรรมการการประเมิน (Joint Committee) ปี 1981 มาตรฐานดังกล่าวมี 4 ด้าน ประกอบด้วยเกณฑ์ 30 เกณฑ์ กล่าวคือ มาตรฐาน การใช้ประโยชน์ (utility standards=U) มี 8 เกณฑ์ มาตรฐานการนำไปใช้ (feasibility standards=F2) มี 3 เกณฑ์ มาตรฐานความเหมาะสม (propriety standards=P) มี 8 เกณฑ์ และมาตรฐานความถูกต้อง (accuracy standards=A) มี 11 เกณฑ์ รายละเอียดของมาตรฐานการประเมินผู้วิจัยนำเสนอไว้ในหัวข้อต่อไปคือ มาตรฐานการประเมินนิมิตาน

9. ขอบเขตของการประเมินที่เน้น (system focus)

นักประเมินนิมิตานต้องศึกษาวิเคราะห์โมเดลการประเมินว่าเป็นโมเดลที่มุ่งเน้นการประเมินที่กว้างขวางครอบคลุมพื้นที่ใดบ้าง กล่าวคือเป็นการให้เฉพาะในห้องเรียนหรือชุมชน (disriect) หรือมีการติดต่อเชื่อมโยงกับรัฐอื่นๆ

10. จุดเด่นของโครงการ (unique strengths)

นักประเมินนิมิตานต้องศึกษาวิเคราะห์การประเมินโรงเรียนด้วยโมเดลการประเมินที่ใช้เน้น ผลผลการประเมินให้ข้อเสนอแนะที่เป็นจุดเด่นจุดด้อยหรือลักษณะที่สำคัญของโมเดลการประเมินอย่างไร

มิติการประเมินในการประเมินนิมิตานโมเดลที่ใช้ในการประเมินโรงเรียนทั้ง 10 มิติดังกล่าว มิติที่มีลักษณะเป็นจุดเด่นหรือจุดด้อยของโครงการจะถูกให้คะแนนโดยคณะกรรมการ คะแนนที่มีลักษณะเป็นมาตราส่วนประมาณค่า 5 ระดับ ดังที่กล่าวมาแล้วในขั้นตอนของการประเมิน

การกำหนดตัวบ่งชี้คุณภาพในการประเมินโรงเรียน (quality indicators in school evaluation)

ในการประเมินนิมิตานโมเดลที่ใช้ในการประเมินโรงเรียน คณะกรรมการการประเมินต้องมีการประชุมปรึกษาหารือเพื่อหาข้อตกลงร่วมกันสำหรับการกำหนดตัวบ่งชี้ โดยเน้นที่ตัวบ่งชี้จุดเด่นที่สำคัญของโมเดลการประเมิน การประเมินนิมิตานโมเดลที่ใช้ในการประเมินโรงเรียนครั้งนี้กำหนดตัวบ่งชี้คุณภาพ ในการประเมินดังนี้

- 1) การกำหนดทีมงานการประเมินจากผู้ที่มีความสามารถในการประเมิน
- 2) ผู้มีส่วนเกี่ยวข้องมีส่วนร่วมในการตัดสินใจ
- 3) คณะกรรมการการประเมินมีเกณฑ์ในการปฏิบัติงาน
- 4) การประเมินคำนึงถึงการเรียนรู้ที่ยึดนักเรียนเป็นศูนย์กลาง
- 5) เป็นการประเมินที่สนับสนุนการประเมินที่แท้จริง (authentic assessment) ของคณะ

กรรมการ

- 6) เป็นการประเมินที่สนับสนุนการประเมินผลสัมฤทธิ์ทางการเรียนของนักเรียนตามสภาพจริง
- 7) การรวบรวมข้อมูลที่ครอบคลุม
- 8) การประเมินมีการกำหนดข้อตกลงเบื้องต้นและความคาดหวังในความสามารถของนักเรียน
- 9) ผู้บริหารระดับสูงให้การสนับสนุน
- 10) บริหารงานประเมินโดยผู้มีความสามารถและมีประสบการณ์
- 11) กระบวนการประเมินและผลจากการประเมินกระตุ้นให้เกิดการปรับปรุง
- 12) การกำหนดจุดมุ่งหมายและผลลัพธ์ของการประเมินชัดเจน
- 13) การมีส่วนร่วมของบุคลากรในโรงเรียนสำหรับการตัดสินใจในการประเมิน
- 14) การประมวลผล วิเคราะห์และรายงานผลในระยะยาว
- 15) ผลประเมินกระตุ้นให้เกิดการทำงาน
- 16) ขยายโครงสร้างการประเมินให้ครอบคลุมถึงชุมชน
- 17) การประเมินก่อให้เกิดการพัฒนาโปรแกรมการศึกษา
- 18) การประเมินก่อให้เกิดการพัฒนากิจกรรมพิเศษในหลักสูตร
- 19) ตรวจสอบความสามารถครูและนักเรียน
- 20) การรวบรวมข้อมูล/สารสนเทศและการติดตามการดำเนินงานนำมาซึ่งการพัฒนาปรับปรุง

โครงสร้างของโรงเรียนให้มีความเหมาะสม

- 21) การประเมินก่อให้เกิดการพัฒนาโรงเรียนทั้งระดับชุมชน ระดับภูมิภาค ระดับรัฐ และระดับส่วนกลาง
- 22) การประมวลผลข้อมูลมีการตรวจสอบผลกับจุดมุ่งหมายระดับชาติ

23) การแยกวิเคราะห์ข้อมูลตามคุณลักษณะ เช่น เพศ เศรษฐฐานะ เป็นต้น
ตัวชี้คุณภาพในการประเมินโรงเรียนนี้ คณะกรรมการจัดทำขึ้นจากการศึกษาทบทวนในผลการประเมินด้วยโมเดลที่ใช้ในการประเมินโรงเรียนทั้ง 51 โมเดล

จากขั้นตอนการประเมินอภิมานที่เสนอโดยนักประเมินทั้ง 3 กลุ่มดังกล่าวข้างต้น จะเห็นว่าลักษณะคล้ายคลึงกับรูปแบบการประเมินโดยทั่วไปแตกต่างกันตรงจุดเริ่มของการประเมินอภิมานที่จะต้องมีการประเมินเกิดขึ้นมาก่อนแล้ว

ในการนำเสนอขั้นตอนกระบวนการประเมินอภิมานครั้งนี้ ผู้วิจัยนำเสนอเทคนิคซึ่ง Stufflebeam (1981) ได้สรุปจากข้อเสนอแนะที่นักประเมินหลาย ๆ ท่านเสนอแนวทางที่เป็นประโยชน์ที่ควรใช้ในกระบวนการประเมินอภิมานไว้ 10 ประการ ดังนี้ 1) การใช้เทคนิคการประเมินที่หลากหลาย 2) ใช้การออกแบบการทดลองและกึ่งทดลองในการประเมิน 3) การใช้ทีมเทคนิค (team technique) ซึ่ง Reinhard (1972, อ้างใน Stufflebeam, 1981) อธิบายว่า ทีมเทคนิค เป็นการประเมินแผนการประเมินซ้ำอีกครั้ง 4) การใช้เทคนิคเดลฟายเพื่อกำหนดแนวทางที่จะใช้ในการประเมิน 5) การใช้การวิเคราะห์ความแปรปรวน (ANOVA) ในการเลือกระบบการประเมิน 6) การใช้แบบตรวจสอบรายการเพื่อประเมินแผนงานและการปฏิบัติงาน 7) การวิเคราะห์เทคนิคที่ใช้ในการการวิเคราะห์ข้อมูลของการประเมินครั้งแรก 8) การใช้การวิเคราะห์และสังเคราะห์ผลการประเมิน 9) การรับฟังความคิดเห็นของผู้บริหาร 10) การจัดทำสัญญาข้อตกลงของการประเมินที่เป็นทางการ

2.5 มาตรฐานการประเมินอภิมาน

กุญแจสำคัญของการประเมินอภิมานคือการกำหนดเกณฑ์ในการประเมินที่เป็นมาตรฐาน (standard) เพื่อใช้ในการพิจารณาตัดสินคุณค่างานประเมินและมาตรฐานของการประเมินได้มีการพิจารณากำหนดเกณฑ์ให้เป็นที่ยอมรับของนักประเมิน มาตรฐานที่ใช้ในการประเมินอภิมานมีแตกต่างกันหลายแนวคิด แนวคิดที่สำคัญ ได้แก่ มาตรฐานที่เสนอโดยคณะกรรมการกำหนดมาตรฐานการประเมินทางการศึกษา (Stufflebeam และคณะ, 1981) และรายการตรวจสอบงานประเมินที่เสนอโดย Scriven (1991)

ในช่วงแรกของการดำเนินการกำหนดมาตรฐานอาศัยการทบทวนจากงานประเมินซึ่งดำเนินการในช่วงปลายปี ค.ศ. 1960 โดยผู้เชี่ยวชาญจำนวน 30 ท่าน ทำการศึกษางานประเมินในส่วนที่เป็นข้อเสนอแนะและหลุมพรางของการประเมิน จากนั้นทำการประชุมปรึกษาหารือจัดทำร่างมาตรฐานงานประเมินขึ้นมา จากนั้นใช้ผู้เชี่ยวชาญ 50 ท่าน ทำการการพิจารณาร่างครั้งแรกนั้นและได้ปรับปรุงมาตรฐาน การปรับปรุงมาตรฐานได้ดำเนินการต่อ ๆ มา จนในที่สุดกำหนดร่างได้ทำจนถึง 5 ครั้ง ครั้งสุดท้ายได้ถูกตีพิมพ์เผยแพร่ในปี 1980 โดย McGraw-Hill (Stufflebeam, 1981) ซึ่งเอกสารฉบับนี้นักประเมินใช้เป็นแหล่งอ้างอิงในงานประเมินต่าง ๆ มากมาย สำหรับนักประเมินในประเทศไทยมีหลายท่านได้แปลมาตรฐานการประเมินและ

พิมพ์ออกเผยแพร่ ดังเช่น สมหวัง พิธิยานุวัฒน์ (2539), ศิริชัย กาญจนวาสี (2537), บุญส่ง นิลแก้ว (2537) เป็นต้น

ในที่นี้ผู้วิจัยนำเสนอรูปแบบของมาตรฐานที่ถูกเสนอและจัดพิมพ์เผยแพร่ในขั้นตอนสุดท้ายจาก ต้นฉบับที่ประกอบด้วย 30 มาตรฐาน โดยแบ่งเป็น 4 ด้าน คือ มาตรฐานการใช้ประโยชน์ (utility standards) มาตรฐานความเป็นไปได้ (feasibility standards) มาตรฐานความเหมาะสม (propriety standards) และมาตรฐานความถูกต้อง (accuracy standards) ดังรายละเอียดต่อไปนี้

5.1 มาตรฐานการใช้ประโยชน์ (utility standards)

เป็นมาตรฐานที่ต้องการประกันถึงความเป็นประโยชน์ของผลการประเมินในการตอบสนองต่อความต้องการใช้สารสนเทศของผู้เกี่ยวข้อง มาตรฐานการใช้ประโยชน์นี้ประกอบด้วยเกณฑ์ 8 ข้อดังต่อไปนี้

- 1) การระบุผู้เกี่ยวข้องที่ต้องการใช้สารสนเทศ ทั้งที่มีส่วนได้และส่วนเสีย
- 2) ความเป็นที่เชื่อถือของผู้ประเมิน
- 3) การรวบรวมข้อมูลครอบคลุมและตอบสนองความต้องการใช้สารสนเทศของผู้เกี่ยวข้อง
- 4) การแปลความหมายที่ให้คุณค่าชัดเจน
- 5) รายงานการประเมินมีความชัดเจนทุกขั้นตอน
- 6) การเผยแพร่ผลการประเมินให้ผู้เกี่ยวข้องทราบอย่างทั่วถึง
- 7) การรายงานการประเมินเสร็จทันเวลาสำหรับนำไปใช้ประโยชน์
- 8) การประเมินส่งผลกระทบในการกระตุ้นให้มีการดำเนินการประเมินต่อไปอย่างต่อเนื่อง

5.2 มาตรฐานความเป็นไปได้ (feasibility standards)

เป็นมาตรฐานที่ต้องการประกันถึงการประเมินที่มีความสอดคล้องกับความเป็นจริง ปฏิบัติได้ ยอมรับได้ ละเอียดรอบคอบ ประหยัดและคุ้มค่า มาตรฐานความเป็นไปได้ประกอบด้วยเกณฑ์ 3 ข้อ

- 1) วิธีการประเมินสามารถนำไปปฏิบัติได้จริง
- 2) การเป็นที่ยอมรับได้ทางการเมืองของการประเมิน
- 3) ผลที่ได้มีความคุ้มค่า

5.3 มาตรฐานความเหมาะสม (propriety standards)

เป็นมาตรฐานที่ต้องการประกันว่า การประเมินได้ทำอย่างถูกต้องตาม กฎ ระเบียบ จรรยาบรรณ มีการคำนึงถึงสวัสดิภาพของผู้เกี่ยวข้องในการประเมิน และผู้ที่ได้รับผลกระทบจากการประเมินมาตรฐานความเหมาะสมประกอบด้วยเกณฑ์ 8 ข้อดังต่อไปนี้

- 1) การกำหนดข้อตกลงของการประเมินอย่างเป็นทางการ
- 2) ความสนใจแก้ปัญหาของความขัดแย้งในการประเมินด้วยความเป็นธรรม
- 3) การรายงานผลการประเมินอย่างตรงไปตรงมา เปิดเผย และความสมบูรณ์ของการประเมิน

- 4) การให้ความสำคัญต่อสิทธิในการรับรู้ข่าวสารของสาธารณะ
- 5) การคำนึงถึงสิทธิส่วนตัวของกลุ่มตัวอย่างที่เป็นผู้ให้ข้อมูล
- 6) การเคารพสิทธิในการปฏิเสธสัมพันธ์กับผู้เกี่ยวข้อง
- 7) การรายงานผลการประเมินที่ยุติธรรม และเสนอทั้งจุดเด่นและจุดด้อยของสิ่งที่ประเมิน
- 8) การประเมินด้วยความรับผิดชอบ และมีจรรยาบรรณ

5.4 มาตรฐานความถูกต้อง (accuracy standards)

เป็นมาตรฐานที่ต้องการประกันว่า การประเมินเป็นไปอย่างเปิดเผย ตรงไปตรงมา มีการใช้เทคนิคที่เหมาะสมเพื่อให้ได้สารสนเทศที่เพียงพอสำหรับตัดสินคุณค่าทางด้านวัตถุและคุณค่าทางด้านนามธรรมของสิ่งที่ประเมิน มาตรฐานความถูกต้องประกอบด้วยเกณฑ์ 11 ข้อดังต่อไปนี้

- 1) การระบุและอธิบายรายละเอียดของการประเมินอย่างชัดเจน
- 2) การวิเคราะห์บริบทของการประเมิน
- 3) การระบุจุดประสงค์และอธิบายกระบวนการประเมินอย่างชัดเจน
- 4) การอธิบายการใช้เครื่องมือวัดผลได้อย่างชัดเจน
- 5) การวัดผลมีความตรง
- 6) การวัดผลมีความเที่ยง
- 7) การจัดระบบควบคุมสำหรับการเก็บรวบรวมข้อมูล วิเคราะห์และรายงาน
- 8) การวิเคราะห์ข้อมูลเชิงปริมาณอย่างมีประสิทธิภาพ
- 9) การวิเคราะห์ข้อมูลเชิงคุณภาพอย่างมีประสิทธิภาพ
- 10) การสรุปผลการประเมินอย่างมีเหตุผล
- 11) การรายงานผลการประเมินมีความเป็นปรนัย

ในปัจจุบันที่ศาสตร์ทางการประเมินได้พัฒนาขึ้นมา มีการประเมินเกิดขึ้นในโครงการต่าง ๆ อย่างกว้างขวางนอกเหนือการประเมินในโรงเรียน มาตรฐานการประเมินที่มีการตีพิมพ์ครั้งแรกในปี ค.ศ. 1980 ดังที่กล่าวมาแล้วคณะกรรมการร่วมในการประชุมมีความเห็นตรงกันว่าควรมีการปรับปรุงมาตรฐานการประเมินทางการศึกษาที่ตีพิมพ์ครั้งแรกโดยมีหลักการให้มาตรฐานเป็นแนวทางในการประเมินที่มีประสิทธิภาพมากขึ้น สำหรับมาตรฐานการประเมินในการพิมพ์ครั้งที่ 2 ค.ศ.1994 ซึ่งมี Sander เป็นประธานคณะกรรมการ ยังคงมีมาตรฐาน 4 ด้าน ที่ประกอบด้วย 30 เกณฑ์ดังเดิม โดยมีจุดที่ทำการปรับปรุง 2 จุด จุดแรก ในมาตรฐานการใช้ประโยชน์เดิมมี 8 เกณฑ์นั้นทำการรวมเกณฑ์ข้อ 6 และ 7 เข้าด้วยกัน เหลือเพียง 7 เกณฑ์ คือเป็นเกณฑ์ข้อที่ 6 การรายงานผลการประเมินให้ผู้เกี่ยวข้องได้ทราบอย่างทั่วถึง

และหันตามเวลาสำหรับการนำไปใช้ประโยชน์ อีกจุดหนึ่งที่เปลี่ยนไปคือ ในมาตรฐานความถูกต้องเดิม ประกอบด้วย 11 เกณฑ์นั้นมีการเพิ่มเกณฑ์อีก 1 เกณฑ์ เป็นเกณฑ์ที่ 12 คือ การประเมินมีการตรวจสอบผลการประเมิน (อ้างใน สมหวัง พิธิยานุวัฒน์, 2539) ดังนั้นมาตรฐานที่มีการปรับปรุงขึ้นใหม่จึงมีมาตรฐาน 4 ด้านที่ประกอบด้วยเกณฑ์ 30 เกณฑ์เท่าเดิม

ส่วนรายการตรวจสอบการประเมิน (key evaluation checklist=KEC) ที่เสนอโดย Scriven (1991) จำนวน 15 รายการ ประกอบด้วยรายการซึ่งใช้ตรวจสอบการประเมินตั้งแต่การสำรวจก่อนการประเมินจนถึงการวิเคราะห์ผลขั้นสุดท้ายของการประเมิน ดังนั้นรายการดังกล่าวจึงครอบคลุมกระบวนการประเมินที่จะบ่งชี้ให้ผู้ประเมินงานประเมินและผู้เกี่ยวข้องทราบถึงความเป็นไปตลอดการประเมินนั้น รายการตรวจสอบการประเมิน 15 รายการ มีดังนี้ 1) การบรรยายให้ทราบว่าประเมินอะไร (description) โดยให้คำอธิบายเกี่ยวกับ ธรรมชาติในการปฏิบัติงาน หน้าที่ที่ต้องปฏิบัติ ระบบการทำงานและการสนับสนุนการทำงาน 2) การกำหนดผู้ต้องการใช้ผลการประเมิน (client) ในส่วนนี้ผู้ที่ทำการประเมินต้องทราบให้ได้ว่ามีผู้ใดบ้างที่รับผลของการประเมิน 3) บริบทของการประเมิน (background & context) เป็นการตอบคำถามให้ได้ว่าการประเมินครั้งนี้มีสถานะและสิ่งแวดล้อมอย่างไร 4) ความเหมาะสมของการใช้ทรัพยากร (resources) 5) ผู้มีผลได้ผลเสียกับการประเมิน (consumer) 6) คุณค่างานประเมิน (values) 7) กระบวนการประเมิน (process) 8) ผลลัพธ์ของการประเมิน (outcomes) 9) การสรุปอ้างอิงผลการประเมิน (generalizability) 10) ความคุ้มค่าของการประเมิน (costs) โดยพิจารณาจากรายจ่ายของการประเมินทั้งที่เป็นตัวเงินและไม่เป็นตัวเงิน 11) การเปรียบเทียบผลการประเมิน (comparison) 12) ลักษณะสำคัญของการประเมิน (significance) 13) ข้อเสนอแนะงานประเมิน (recommendations) 14) การรายงานการประเมิน (report) 15) การตรวจสอบผลการประเมิน (metaevaluation)

2.6 ประโยชน์ของการประเมินอภิมาน

การประเมินโดยทั่วไปและการประเมินอภิมานให้ผลการประเมินที่นอกจากจะใช้เพื่อตัดสินใจเกี่ยวกับโครงการของผู้ใช้ผลการประเมินแล้ว การประเมินอภิมานยังก่อให้เกิดประโยชน์สำหรับบุคคลที่ใช้ผลการประเมินในประเด็นต่าง ๆ อีกมาก ผู้วิจัยนำเสนอประโยชน์ของการประเมินอภิมานที่ศึกษาจากเอกสารที่ Stufflebeam (1981), Gallegos (1994), Finn, Stevens, Stufflebeam และ Walberg, (1997) สรุปประโยชน์ของการประเมินอภิมานได้เป็น 4 ข้อ ดังนี้

- 1) ด้านการศึกษาและการฝึกอบรม (education and training)

การประเมินอภิมานทำให้ได้ผลการประเมินซึ่งมีส่วนก่อให้เกิดการปรับปรุงเปลี่ยนแปลงที่เป็นการพัฒนาทางการศึกษาและการฝึกอบรมเจ้าหน้าที่ทำการประเมินให้มีความสามารถมากขึ้น เนื่องจากผลการ

ประเมินอภิมานที่ได้รายงานให้สาธารณชนทราบนั้นได้ระบุถึงจุดเด่นจุดด้อยของการประเมินโครงการ จุดเด่นจุดด้อยดังกล่าวจะเป็นความรู้และเป็นแนวทางสำหรับผู้ที่จะทำหน้าที่ในการประเมินต่อไป

2) ด้านการศึกษาและการปรับโครงสร้าง (education and restructuring)

ผลจากการประเมินอภิมานจะก่อให้เกิดประโยชน์ต่อการศึกษาในด้านการปรับโครงสร้างทางการศึกษาที่สอดคล้องกับสภาพความเป็นจริงในปัจจุบัน

3) ด้านการพัฒนาบุคลากร (personal improvement)

ผลจากการฝึกอบรมเจ้าหน้าที่เพื่อเพิ่มความสามารถในการประเมินและการปรับโครงสร้างทางการศึกษาทำให้บุคลากรของหน่วยงานมีความรู้ความสามารถมากขึ้น ซึ่งจะเป็นประโยชน์ในการทำงานในหน่วยงานที่สังกัดอยู่

4) ด้านความร่วมมือกับหน่วยงานทางการศึกษาอื่น (school/university partnership)

ในกระบวนการการประเมินอภิมานซึ่งต้องมีการติดต่อประสานงานระหว่างหน่วยงานหรือบุคลากร ซึ่งต้องประสานงานกันในการทำงาน ดังนั้นในกระบวนการประเมินอภิมานจึงก่อให้เกิดความร่วมมือทั้งในหน่วยงานและระหว่างหน่วยงานด้วย

ตอนที่ 3 งานวิจัยที่เกี่ยวข้อง

เนื่องจากการวิจัยครั้งนี้มุ่งพัฒนากระบวนการประเมินเพิ่มผลงานและการพัฒนาแนวทางการใช้ประโยชน์จากการประเมินเพิ่มผลงานโดยการประยุกต์ใช้การประเมินอภิมาน ดังนั้นในการรายงานการศึกษา งานวิจัยที่เกี่ยวข้องผู้วิจัยจึงแบ่งเป็น 2 ตอน คือหัวข้อ 3.1 เป็นรายงานการวิจัยที่เกี่ยวข้องกับการประเมินเพิ่มผลงาน และหัวข้อ 3.2 เป็นรายงานการวิจัยที่เกี่ยวข้องกับการประเมินอภิมาน

3.1 รายงานการวิจัยที่เกี่ยวข้องกับการประเมินเพิ่มผลงาน

รายงานวิจัยที่เกี่ยวข้องกับการประเมินเพิ่มผลงานทั้งหมด ผู้วิจัยได้จัดเป็นหมวดหมู่ตามลักษณะของวัตถุประสงค์หลักในการวิจัยเป็น 2 กลุ่ม โดยกลุ่มแรกเป็นรายงานวิจัยที่ศึกษาผลของการประเมินเพิ่มผลงานที่มีต่อผลสัมฤทธิ์ทางการเรียน กลุ่มที่สองเป็นกลุ่มการวิจัยที่มุ่งศึกษาสภาพการจัดการเรียนการสอนของการประเมินเพิ่มผลงานและการพัฒนาการประเมินเพิ่มผลงาน

ผลการวิจัยของรายงานวิจัยกลุ่มที่หนึ่งสรุปได้ว่านักเรียนกลุ่มทดลองที่ใช้การประเมินเพิ่มผลงานมีผลสัมฤทธิ์ทางการเรียนสูงกว่านักเรียนกลุ่มที่ใช้การประเมินผลแบบเดิม รายงานวิจัยในกลุ่มนี้ มีดังนี้

สมชาย มิ่งมิตร (2539) ทำการศึกษาผลของการประเมินจากพอร์ทัลโพลีโอที่มีต่อผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 5 เป็นการวิจัยเชิงทดลอง ผลการวิจัยพบว่า นักเรียนที่ได้รับการประเมินผลโดยใช้พอร์ทัลโพลีโอมีผลสัมฤทธิ์ทางการเรียนและเจตคติทางการเรียน

สูงกว่านักเรียนที่ได้รับการประเมินผลแบบเดิม จึงสรุปได้ว่าการประเมินจากพอร์ทโฟลิโอมีผลต่อผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 5

ยุพา เวียงกลม (2541) ศึกษาผลการประเมินตนเองโดยใช้แฟ้มผลงานที่มีต่อความเชื่อมั่นในตนเองของนักเรียนชั้น ป.6 โรงเรียนสังกัดสำนักงานการประถมศึกษาจังหวัดกาฬสินธุ์ จำนวน 60 คน แบ่งเป็นกลุ่มทดลองและกลุ่มควบคุม ผลการวิจัยพบว่า หลังการทดลองนักเรียนกลุ่มทดลองมีค่าเฉลี่ยคะแนนความเชื่อมั่นสูงกว่าค่าเฉลี่ยคะแนนความเชื่อมั่นในตนเองของนักเรียนกลุ่มควบคุมอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

กนกวรรณ บั้งทอง (2542) ศึกษาผลการเรียนวิชาคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษา ปีที่ 2 จากการประเมินผลโดยใช้แฟ้มผลงาน กลุ่มตัวอย่างเป็นนักเรียนโรงเรียนพนาศึกษา สังกัดกรมสามัญศึกษา จังหวัดอำนาจเจริญ จำนวน 265 คน ผลการวิจัยพบว่า นักเรียนกลุ่มตัวอย่างมีความสามารถในการเรียนวิชาคณิตศาสตร์ ค.203 ซึ่งประเมินโดยการใช้แฟ้มผลงานในระดับค่อนข้างสูง

รศ. ดร.สำลี ทองธิว (2541) ทำการพัฒนาระบบการสอนโดยใช้พอร์ทโฟลิโอการสอนอ่านเพื่อพัฒนาภาวะทางการอ่านของนักเรียนประถมศึกษาและเพื่อส่งเสริมพัฒนาการทางวิชาชีพครูมีวัตถุประสงค์ของการวิจัยคือ 1) เพื่อพัฒนาระบบการสอนโดยใช้พอร์ทโฟลิโอการสอนอ่านสำหรับให้นิสิตปริญญาโท สาขาประถมศึกษาใช้กับนักเรียนระดับประถมศึกษา 2) เพื่อศึกษาผลของการใช้ระบบการสอนโดยใช้พอร์ทโฟลิโอการสอนอ่านหนังสือสำหรับเด็กประเภทต่าง ๆ ที่มีต่อการพัฒนาทางวิชาชีพของนิสิตปริญญาโท สาขาประถมศึกษา ภาคต้น ปีการศึกษา 2540 3) เพื่อศึกษาภาวะทางการอ่านของนักเรียนระดับประถมศึกษาที่ได้รับการสอนอ่านโดยใช้พอร์ทโฟลิโอการสอนอ่านหนังสือประเภทต่าง ๆ เป็นการวิจัยเชิงทดลองกับนักเรียนชั้นประถมศึกษาปีที่ 5 ในสังกัดกรุงเทพมหานคร เขตปทุมวัน 3 โรงเรียน โดยสุ่มนักเรียนแบบบังเอิญ ผลการวิจัยพบว่า นิสิตสามารถระบุวิธีการสอนอ่านหนังสือสำหรับเด็ก และสามารถสร้างวิธีการสอนอ่านแบบใหม่ ๆ ได้เป็นจำนวนมากขึ้นหลังการใช้ระบบการสอนโดยใช้พอร์ทโฟลิโอครบทุกขั้นตอน และนักเรียนอ่านหนังสือสำหรับเด็กประเภทนอกเหนือจากหนังสือประกอบภาพในหมวดจินตนาการเป็นจำนวนมากขึ้น หลังจากที่ได้ใช้ระบบการสอนโดยใช้พอร์ทโฟลิโอ ชั้นที่ 1-6

ประภากร กองตาพันธ์ (2541) ศึกษาผลสัมฤทธิ์ทางการเรียนในวิชาภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 3 ที่ใช้แฟ้มผลงาน กลุ่มตัวอย่างเป็นนักเรียนในสังกัดสำนักงานการประถมศึกษาอำเภอป่าแดด สำนักงานการประถมศึกษาจังหวัดเชียงราย จำนวน 452 คน ผลการวิจัยสรุปได้ว่า ผลสัมฤทธิ์ทางการเรียนในวิชาภาษาไทยของนักเรียนหลังใช้แฟ้มผลงานสูงกว่าก่อนใช้แฟ้มผลงานอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 และผลสัมฤทธิ์ทางการเรียนในวิชาภาษาไทยของนักเรียนที่ใช้แฟ้มผลงานกับนักเรียนที่ไม่ใช้แฟ้มผลงานมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001

สุทธิธา กุญจนวน (2541) ศึกษาเปรียบเทียบผลสัมฤทธิ์การเรียนรู้ด้านพุทธิพิสัย การปฏิบัติ จิตพิสัยและความก้าวหน้าการเรียนรู้ระหว่างกลุ่มที่เรียนด้วยการสอนเชิงสหวิทยาการกับกลุ่มที่เรียนด้วยการสอนปกติ กลุ่มตัวอย่างเป็นนักเรียนวิทยาลัยอาชีวศึกษาพระนครศรีอยุธยาระดับประกาศนียบัตรวิชาชีพปีที่ 1 โดยสุ่มแบบหลายขั้นตอน ได้กลุ่มตัวอย่างจำนวน 2 ห้องเรียน แบ่งเป็นกลุ่มทดลองและกลุ่มควบคุม ผลการวิจัยพบว่า นักเรียนที่เรียนด้วยการสอนเชิงสหวิทยาการมีผลสัมฤทธิ์ทางการเรียนด้านพุทธิพิสัย ด้านปฏิบัติการ ด้านจิตพิสัย และคะแนนความก้าวหน้าการเรียนรู้สูงกว่านักเรียนกลุ่มที่เรียนด้วยการสอนปกติ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิจัยของรายงานวิจัยกลุ่มที่สองสรุปได้ว่าการจัดทำแฟ้มผลงานครูต้องมีการเตรียมความพร้อมก่อนการดำเนินการจึงจะสามารถทำการประเมินผู้เรียนโดยใช้แฟ้มผลงานได้ รายงานวิจัยในกลุ่มนี้ เช่น

สมพงษ์ ชินสร้อย (2540) ทำการพัฒนาการประเมินผลการเรียนโดยใช้แฟ้มสะสมงานของนักเรียนในโรงเรียนสังกัดสำนักงานการประถมศึกษาจังหวัดปราจีนบุรี เก็บข้อมูลจากครูผู้สอนจำนวน 114 คน ศึกษาเกี่ยวกับความรู้ความเข้าใจในการประเมินผลการเรียนโดยใช้แฟ้มสะสมงานของนักเรียน ความสามารถในการปฏิบัติการ ความคิดเห็นของครูที่มีต่อการประเมิน ผลการวิจัยพบว่า 1) ครูมีความรู้ความเข้าใจเกี่ยวกับการประเมินผลการเรียนโดยใช้แฟ้มสะสมงานของนักเรียนในระดับมากและมีความรู้ความเข้าใจมากที่สุดคือการนำผลการวัดผลประเมินผลไปใช้ในการปรับปรุงและพัฒนาการเรียนการสอน วิธีการคัดเลือกงาน วิธีการเผยแพร่ผลงาน และวิธีการประเมินผลการเรียนรู้อยู่โดยดูจากแฟ้มสะสมงานของนักเรียน 2) ครูมีความสามารถเกี่ยวกับการปฏิบัติการประเมินผลการเรียนโดยใช้แฟ้มสะสมงานในระดับมาก 3) ครูมีความคิดเห็นเกี่ยวกับการประเมินผลการเรียนโดยใช้แฟ้มสะสมงานในระดับมาก

ชัยฤทธิ์ ดิลาเดช (2540) ศึกษาการพัฒนาแฟ้มสะสมงานในการประเมินผลการเรียนวิชาภาษาอังกฤษนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีวัตถุประสงค์เพื่อพัฒนาหาแนวทางที่เหมาะสมในการใช้แฟ้มสะสมงานประเมินผลการเรียนและจัดกิจกรรมการเรียนการสอนวิชาภาษาอังกฤษ ศึกษาผลการใช้และคุณภาพการประเมินในด้านความเที่ยงและความตรง ศึกษาจากกลุ่มตัวอย่างนักเรียน ชั้นมัธยมศึกษาปีที่ 3 จำนวน 43 คน โรงเรียนแคทราย-วิทยา จังหวัดราชบุรี ผลการวิจัยสรุปได้ว่า 1) การนำไปใช้ในการจัดการเรียนการสอนและการประเมินผลวิชาภาษาอังกฤษมีแนวทางในการดำเนินงานตามแบบจำลองการจัดการเรียนการสอนและการประเมินผลโดยใช้แฟ้มสะสมงาน ซึ่งประกอบด้วยขั้นตอน ดังนี้ ศึกษาสภาวะแวดล้อม กำหนดจุดมุ่งหมายของแฟ้ม กำหนดโครงสร้างของแฟ้มสะสมงาน พัฒนาแฟ้มสะสมงาน ประเมินผลแฟ้มสะสมงาน และประสานสัมพันธ์ร่วมชื่นชม 2) กระบวนการของแฟ้มสะสมงานช่วยให้นักเรียนมีความก้าวหน้าในการเรียนวิชาภาษาอังกฤษทั้ง 4 ทักษะ และสร้างแรงจูงใจในการเรียนแก่ผู้เรียนมาก 3) การประเมินผลโดยใช้แฟ้มสะสมงาน มีคุณภาพตามเกณฑ์ตัดสินคุณภาพการประเมินผลของลินน์และคณะในระดับมากทุกรายการ ดังนี้ คือ มีความยุติธรรมในการประเมิน สามารถถ่ายโยงและอ้างสรุปผลการประเมิน ประเมินระดับสติปัญญาที่ซับซ้อน เนื้อหาแฟ้มสะสมงานมีคุณภาพ เนื้อหาแฟ้มสะสมงานครอบคลุม และการประเมินมี

ความหมายต่อผู้ถูกประเมิน มีความตรงตามสภาพในระดับสูง (.94) ความเที่ยงจากการให้คะแนนของผู้ตรวจ 2 คน มีค่าสูง (.97) ส่วนความเที่ยงของคะแนนที่ได้จากค่าสัมประสิทธิ์การสรุปอ้างอิง มีค่าปานกลาง (.65)

ชันษา ` มากเหลือ (2541) ทำการศึกษาการจัดการเรียนการสอนวิชาภาษาอังกฤษระดับประถมศึกษา ที่ใช้การประเมินโดยใช้แฟ้มสะสมงาน มีวัตถุประสงค์เพื่อศึกษาพฤติกรรมการจัดการเรียนการสอนของครูผู้สอนวิชาภาษาอังกฤษระดับประถมศึกษาที่ใช้การประเมินโดยใช้แฟ้มสะสมงาน และผลสัมฤทธิ์ทางการเรียนวิชาภาษาอังกฤษของนักเรียนที่ได้รับการประเมินโดยใช้แฟ้มสะสมงาน รวมทั้งความคิดเห็นของนักเรียนต่อการประเมินโดยใช้แฟ้มสะสมงาน เก็บข้อมูลจากครูผู้สอน จำนวน 3 คน รวบรวมข้อมูลโดยใช้แบบสังเกตพฤติกรรมครู คนละ 16 ครั้ง ในเวลา 4 เดือน ผลการวิจัยพบว่า 1) พฤติกรรมการจัดการเรียนการสอนของครูผู้สอนวิชาภาษาอังกฤษที่ใช้การประเมินโดยใช้แฟ้มสะสมงาน มี 2 ด้านดังนี้ ด้านแรกด้านการเตรียมการสอน ครูได้ศึกษาแนวคิด หลักการและกระบวนการประเมินโดยใช้แฟ้มสะสมงาน รวมทั้งได้ศึกษาหลักสูตรภาษาอังกฤษระดับประถมศึกษา พุทธศักราช 2539 แนวการสอนภาษาอังกฤษต่าง ๆ และคู่มือครู จัดทำกำหนดการสอน จัดทำตารางวิเคราะห์ความสัมพันธ์ระหว่างแนวการจัดการกิจกรรมการเรียนการสอนวิชาภาษาอังกฤษกับจุดประสงค์การเรียนรู้ใน ป.02 แล้วจึงเขียนแผนการสอนพร้อมกับกำหนดภาระงานการเรียนรู้อื่นๆ รวมทั้งได้จัดทำและเตรียมสื่อการเรียนการสอนและวัสดุอุปกรณ์สำหรับการปฏิบัติชิ้นงานของนักเรียน และเตรียมเครื่องมือสำหรับการวัดผลการเรียนวิชาภาษาอังกฤษระหว่างภาคเรียน จัดเตรียมส่วนประกอบของแฟ้มสะสมงาน และเตรียมความพร้อมของนักเรียนและผู้ปกครอง ด้านที่สองด้านการจัดการกิจกรรมการเรียนการสอน พบว่า ครูได้แจ้งวัตถุประสงค์การเรียนรู้และเนื้อหาที่จะสอนให้นักเรียนทราบ และได้ดำเนินการสอนตามกิจกรรมที่ระบุไว้ในแผนการสอน โดยครูใช้แนวการสอนวิธีประสบการณ์ร่วมกันในการอ่านหนังสือ แนวการสอนแบบบูรณาการและแนวการเรียนภาษาแบบธรรมชาติ กิจกรรมการเรียนการสอนของครูได้มุ่งเน้นให้ผู้เรียนเกิดการเรียนและพัฒนาทั้งด้านสติปัญญาและความคิดสร้างสรรค์ ซึ่งความรู้ที่ครูสอนสามารถนำไปใช้ในชีวิตประจำวันได้ รวมทั้งครูได้เปลี่ยนบทบาทจากผู้บอกความรู้มาเป็นผู้ชี้แนะให้นักเรียนเกิดการเรียนรู้ ตลอดจนใช้วิธีการและเครื่องมือวัดผลได้สอดคล้องกับการเรียนรู้ อีกทั้งได้สนับสนุนให้นักเรียนแสวงหาความรู้เพิ่มเติมด้วยตนเอง ด้านการประเมินโดยใช้แฟ้มสะสมงาน พบว่า ครูได้ดำเนินการประเมินโดยใช้แฟ้มสะสมงานครบทั้ง 10 ขั้นตอน คือ การกำหนดวัตถุประสงค์ของแฟ้มสะสมงาน การเก็บรวบรวมงานหรือหลักฐาน การคัดเลือกงานหรือหลักฐาน การจัดระบบของแฟ้มสะสมงาน การสะท้อนความคิดเห็นหรือความรู้สึกต่อผลงาน การตรวจสอบความสามารถของตนเอง การประเมินผลงานหรือหลักฐาน การแลกเปลี่ยนประสบการณ์เกี่ยวกับผลงาน การปรับเปลี่ยนผลงาน และการจัดนิทรรศการผลงานของนักเรียน 2) ผลสัมฤทธิ์ทางการเรียนวิชาภาษาอังกฤษของนักเรียนที่ได้รับการประเมินโดยใช้แฟ้มสะสมงาน พบว่า มีนักเรียนได้ระดับผลการเรียนระดับ 4 ถึงร้อยละ 72.52 3) ความคิดเห็นของนักเรียนต่อการประเมินโดยใช้แฟ้มสะสมงาน พบว่า นักเรียนมีความคิดเห็นต่อการประเมินโดยใช้แฟ้มสะสมงานอยู่ในระดับเห็นด้วย โดยนักเรียนมีความคิดเห็นอยู่ในระดับเห็นด้วยอย่างยิ่งในด้าน

กระบวนการในการจัดทำแฟ้มสะสมงาน ส่วนด้านการพัฒนาตนเอง ด้านประโยชน์ของการประเมินโดยใช้แฟ้มสะสมงานและด้านปฏิสัมพันธ์ในชั้นเรียน นักเรียนมีความคิดเห็นอยู่ในระดับเห็นด้วย

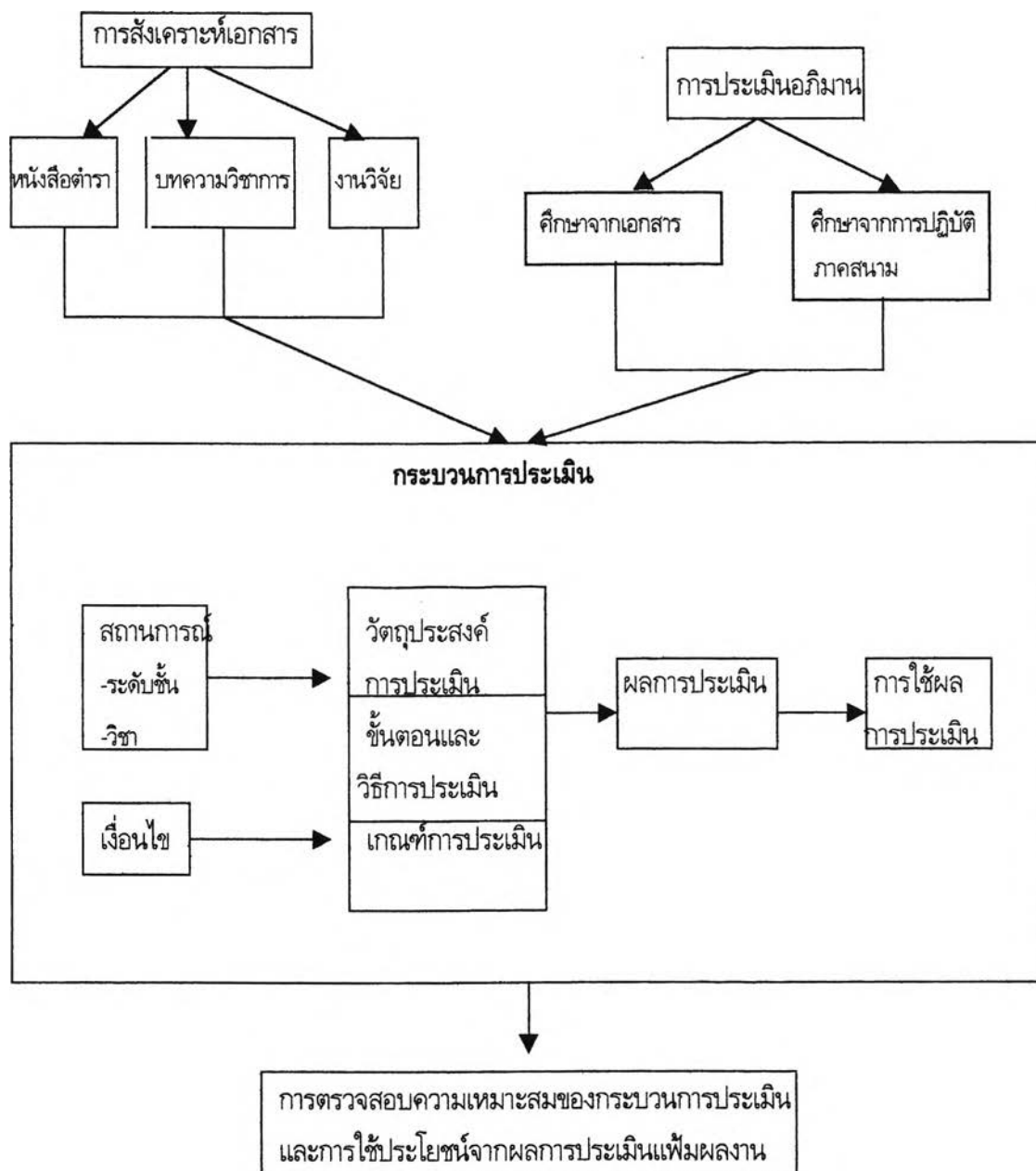
นารีย์ วงศ์ไผ่ (2540) ศึกษาการจัดการเรียนการสอนโดยใช้แฟ้มสะสมงานนักเรียนของครูที่เป็นวิทยากรแกนนำ มีวัตถุประสงค์เพื่อศึกษาพฤติกรรมการจัดการเรียนการสอน กระบวนการที่ใช้ในการส่งเสริมให้นักเรียนจัดทำแฟ้มสะสมงานนักเรียน ผล ปัญหาและวิธีแก้ปัญหาในการจัดการเรียนการสอนโดยใช้แฟ้มสะสมงานนักเรียนของครูที่เป็นวิทยากรแกนนำ รวบรวมข้อมูลจากการสังเกตพฤติกรรมการจัดการเรียนการสอนในห้องเรียน และการสัมภาษณ์ สรุปผลการวิจัยได้ดังนี้ 1) ครูผู้สอนได้เตรียมการสอนโดยศึกษาแนวคิด ทฤษฎี ทำการวิเคราะห์หลักสูตร จัดทำแผนการสอน และเตรียมสื่อการสอน อุปกรณ์ในการจัดทำแฟ้มสะสมงาน การจัดกิจกรรมการเรียนการสอนได้ยึดตามแผนการสอน กิจกรรมที่พบมากที่สุด คือ การให้นักเรียนได้ศึกษาค้นคว้าด้วยระบบกลุ่ม บรรยากาศการเรียนการสอนเป็นแบบกันเอง มีการควบคุมชั้นเรียน ให้การเสริมแรงทั้งทางบวกและทางลบ และสร้างความเชื่อมั่นในตนเองให้นักเรียนกล้าแสดงออกมากขึ้น ครูผู้สอนวัดและประเมินผลผู้เรียนทั้งระหว่างเรียนและเมื่อสิ้นสุดการเรียนการสอนเรียบร้อยแล้ว โดยใช้วิธีการสังเกต การตรวจงานหรือแบบฝึกหัด การซักถามในชั้นเรียน การทดสอบทั้งแบบปรนัยและอัตนัย และจากแฟ้มสะสมงานนักเรียน 2) กระบวนการในการพัฒนาแฟ้มสะสมงานนักเรียนที่ครูผู้สอนส่งเสริมให้นักเรียนจัดทำ มี 4 ขั้นตอน คือ ขั้นรวบรวมผลงาน ขั้นคัดเลือกผลงาน ขั้นประเมินผลงานและชั้นนำเสนอผลงาน 3) ผลที่ได้จากการเรียนการสอนโดยใช้แฟ้มสะสมงานนักเรียน ครูและนักเรียนพอใจกับผลงานของนักเรียนและเห็นว่าจัดการเรียนการสอนโดยใช้แฟ้มสะสมงานนักเรียนเป็นสิ่งที่ดี ควรได้รับการส่งเสริมเนื่องจากสามารถประเมินผลการเรียนรู้ของผู้เรียนได้โดยตรงกับสภาพจริง ทำให้นักเรียนมีความรับผิดชอบ รู้จักวางแผนและทำงานเป็น มีวินัยในตนเอง กล้าแสดงออก ทำงานสะอาดเรียบร้อย มีความคิดสร้างสรรค์และกระตือรือร้นในการทำงานมากขึ้น 4) ในเรื่องที่เกี่ยวข้องกับปัญหาและวิธีการแก้ปัญหาที่ครูใช้ในการจัดการเรียนการสอนโดยใช้แฟ้มสะสมงานนักเรียน ครูให้นักเรียนประดิษฐ์แฟ้มเก็บผลงานเองและขอความช่วยเหลือจากผู้ปกครอง เพื่อแก้ปัญหาสิ้นเปลืองงบประมาณของโรงเรียน ครูมีเวลาน้อยจึงต้องจัดทำแผนการสอนอย่างย่อ ๆ แทนแผนการสอนปกติ ขณะจัดกิจกรรมการเรียนการสอนมีนักเรียนบางคนไม่สนใจเรียน ครูเรียกนักเรียนมาซักถามเป็นรายบุคคล และให้นักเรียนที่เรียนเก่งช่วยเหลือให้คำแนะนำในการทำชิ้นงาน ครูให้ความสำคัญกับชิ้นงานทุกชิ้นเพื่อแก้ปัญหาให้นักเรียนไม่สนใจทำชิ้นงานทุกชิ้น ปัญหาเรื่องการตรวจงานนักเรียนไม่ทัน ครูแก้ปัญหาโดยให้นักเรียนประเมินตนเอง แลกเปลี่ยนกันตรวจในห้องเรียน เมื่อพบว่าการประเมินผลงานของนักเรียนมีความเที่ยงตรงน้อย ครูแก้ปัญหาโดยเน้นย้ำให้นักเรียนยึดเกณฑ์ในการประเมิน และกำหนดเกณฑ์ในการประเมินให้ละเอียดมากขึ้น

3.2 รายงานวิจัยที่เกี่ยวกับการประเมินอภิमान

เนติ เกลยวาเรศ (2541) ทำการประเมินอภิமானการใช้แบบจำลองชิปในการประเมินโครงการการศึกษา กลุ่มประชากรที่ศึกษาคือวิทยานิพนธ์และผู้ทำวิทยานิพนธ์ในระดับบัณฑิตศึกษา จำนวน 46 เล่ม วิเคราะห์ข้อมูลโดยการแจกแจงความถี่และร้อยละ วิเคราะห์เนื้อหา วิเคราะห์คุณภาพของงานประเมิน วิเคราะห์ความแปรปรวนของคะแนนคุณภาพงานประเมิน และวิเคราะห์ถดถอยพหุคูณเพื่อศึกษาปัจจัยที่ส่งผลกระทบต่อคุณภาพการประเมิน ผลการวิเคราะห์ข้อมูล พบว่ารูปแบบการนำแบบจำลองชิปมาใช้ในการประเมินมี 5 รูปแบบ รูปแบบที่นิยมใช้มากที่สุดคือการใช้แบบจำลองชิปขณะที่โครงการกำลังดำเนินอยู่จนกระทั่งโครงการสิ้นสุดโดยประเมินสภาวะแวดล้อมและปัจจัยเบื้องต้นย้อนหลังแล้วจึงประเมินกระบวนการและผลผลิต ผู้ประเมินส่วนใหญ่ไม่ได้นำผลการประเมินเสนอผู้บริหารในการตัดสินใจ คุณภาพงานประเมินโดยใช้แบบจำลองชิปโดยเฉลี่ยอยู่ในเกณฑ์ดี ผลการวิเคราะห์ถดถอยพหุคูณพบว่า ปัจจัยรูปแบบการนำแบบจำลองชิปไปใช้ในการประเมินสามารถทำนายคุณภาพงานประเมินได้อย่างมีนัยสำคัญทางสถิติ ซึ่งสามารถอธิบายความแปรปรวนได้ร้อยละ 24.9

ตอนที่ 4 กรอบการพัฒนากระบวนการประเมินเพิ่มผลงานและการใช้ประโยชน์จากผลการประเมินเพิ่มผลงาน

จากการรายงานเอกสารและงานวิจัยที่เกี่ยวข้องทั้ง 3 ตอน ผู้วิจัยสรุปแนวทางการพัฒนากระบวนการประเมินเพิ่มผลงานและการใช้ประโยชน์จากผลการประเมินเพิ่มผลงาน ดังแผนภาพที่ 1



แผนภาพที่ 1 แผนผังการพัฒนากระบวนการประเมินเพิ่มผลงานและการใช้ประโยชน์จากผลการประเมินเพิ่มผลงาน