

แนวความคิดประเมินที่สอดคล้องกับการเรียนรู้ของนิสิตในยุคศตวรรษที่ 21 :
การประเมินการเรียนรู้ตามสภาพจริง
วันที่ 15 กรกฎาคม 2559 ณ มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย สงขลา
 กฤษฎากาญจน์ ไตพิทักษ์¹

ความนำ

นับตั้งแต่พระราชบัญญัติการศึกษาแห่งชาติ พุทธศักราช 2542 มีผลบังคับใช้เมื่อ 20 สิงหาคม 2542 ได้ก่อให้เกิดการเปลี่ยนแปลงกระบวนการวัดและประเมินการเรียนรู้เพื่อให้สอดคล้องกับกระบวนการเรียนการสอน โดยถือว่าการประเมินเป็น กิจกรรมที่สอดคล้องอยู่ในกระบวนการเรียนรู้ (ส.วาสนา ประवालพฤษ์, 2537; สำนักงานคณะกรรมการการศึกษาแห่งชาติ, สำนักนายกรัฐมนตรี, 2542) การเปลี่ยนแปลงของการประเมินดังกล่าว ทำให้นักวิชาการวัดผลได้กล่าวถึง การประเมินที่สอดคล้องกับการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญหลายวิธี ซึ่งแต่ละวิธีที่กล่าวมามีลักษณะร่วมกันอยู่ 2 ประการ คือ ประการแรกเป็นการประเมินที่เป็นทางเลือกใหม่แทนการใช้แบบทดสอบมาตรฐาน หรือแบบทดสอบแบบเลือกตอบแบบเดิม ประการที่สองเป็นการประเมินการปฏิบัติงานของผู้เรียนในสภาพจริง (Worthen, 1993) Hart (1994) เลือกใช้คำว่า การประเมินการเรียนรู้ตามสภาพจริง (authentic learning assessment) แทนคำว่า การประเมินการเรียนรู้แนวใหม่ เนื่องจากเป็นคำที่ให้ความสำคัญกับการพัฒนาเครื่องมือการประเมินที่สามารถวัดและสะท้อนผลได้ถูกต้องมากกว่าวิธีอื่น วิธีการประเมินตามสภาพจริงมี 3 ประเภท ได้แก่ การสังเกต การประเมินจากแฟ้มสะสมงาน และการประเมินการปฏิบัติตามสภาพจริง โดยนักการศึกษา นักวัดและประเมินผล ตลอดจนครู อาจารย์จึงมีความคาดหวังว่าการประเมินการเรียนรู้ตามสภาพจริงจะเป็นนวัตกรรมการประเมินที่ก่อให้เกิดการพัฒนาผู้เรียนอย่างแท้จริง

วิธีการประเมินการเรียนรู้ตามสภาพจริง

ในอดีตทั้งครูและนักวัดผลถูกชี้ว่าด้วยความเชื่อว่า “ถ้าสิ่งใดไม่สามารถวัดได้สิ่งนั้นก็ไม่มีค่าที่จะทำการสอน” ผลที่ตามมาทำให้เกิดการเน้นการสอนทักษะที่ง่ายต่อการวัดและความรู้ เนื้อหา มากเกินไป ความเคลื่อนไหวของการประเมินการเรียนรู้ตามสภาพจริงได้พลิกผันความเชื่อดังกล่าวมาเป็น “หากสิ่งใดมีค่าควรแก่การเรียนรู้ สิ่งนั้นย่อมมีค่าควรแก่การประเมิน” ถ้าครูต้องการให้นักเรียนมีความสามารถอะไร ครูควรประเมินสิ่งนั้น เช่น ครูต้องการให้นักเรียนเรียนรู้วิธีการแก้ปัญหาโดยประยุกต์ใช้ความรู้ทางคณิตศาสตร์ ครูก็ควรจะมีปัญหาให้นักเรียนแก้และถ้าครูต้องการให้นักเรียนมีทักษะในการวิเคราะห์ ดีความ สังเคราะห์และประเมินค่า ครูต้องประเมินทักษะเหล่านั้นในบริบทที่มีความหมาย

การประเมินการเรียนรู้ตามสภาพจริงไม่ได้หมายความว่า เป็นการล้มเลิกการประเมินความรู้ เนื้อหา ในการประเมินการเรียนรู้ตามสภาพจริงนั้น ความรู้และทักษะมีความสำคัญและมีอาจแยกออกจากกันได้ งานในการประเมินการเรียนรู้ตามสภาพจริงนั้นเป็นการบูรณาการทั้งความรู้และทักษะเข้าด้วยกัน การประเมินการเรียนรู้ตามสภาพจริงมีวิธีการมากมายในการประเมินผู้เรียน การทดสอบก็เป็นหนึ่งในหลากหลายวิธี หากพิจารณาชนิดของข้อมูลที่ได้รับเกี่ยวกับนักเรียนเป็นเกณฑ์สามารถแบ่งวิธีการประเมินได้ 3 ประเภทใหญ่ ๆ ดังนี้ (Hart, 1994)

1. การสังเกต (observation) หรือข้อมูลที่ครูเก็บรวบรวมจากนักเรียนในแต่ละวัน
2. ตัวอย่างการปฏิบัติงาน (performance samples) หรือผลผลิตที่เป็นรูปธรรมซึ่งเป็นหลักฐานแสดงผลสัมฤทธิ์ของนักเรียน หรือการประเมินจากแฟ้มสะสมงาน (portfolio assessment)

3. การทดสอบแบบใหม่และวิธีการที่เสมือนการทดสอบ (tests and test-like) หรือการวัดสัมฤทธิ์ผลของนักเรียนในเวลาและสถานที่ที่จัดไว้โดยเฉพาะ หรือ การประเมินการปฏิบัติตามสภาพจริง (performance authentic assessment)

ซึ่งแต่ละวิธีมีรายละเอียด ดังนี้

1. การสังเกต (observation)

การสังเกตเป็นวิธีการประเมินที่เก่าแก่และใช้กันมาเป็นระยะเวลานานแล้วนับตั้งแต่สมัยของโสเครตีสเป็นต้นมา ในยุคที่แบบสอบมาตรฐานได้รับความนิยม การสังเกตได้ถูกละเลยไม่ได้รับความสนใจเท่าที่ควร เนื่องจากเป็นวิธีการประเมินที่ไม่เป็นทางการ จนกระทั่งมีความเคลื่อนไหวของการประเมินการเรียนรู้ตามสภาพจริง ทำให้การสังเกตกลายเป็นเครื่องมือการประเมินที่มีคุณค่าอีกครั้ง การสังเกตจะทำให้มองเห็นจุดแข็งจุดอ่อนและความสนใจของนักเรียน ผลที่ได้คือภาพพัฒนาการของผู้เรียนที่มีคุณค่ายิ่ง ซึ่งมีประโยชน์มากกว่าข้อมูลที่ได้รับจากแบบทดสอบ (Hart, 1994; Lester, Lambdin & Preston, 1996)

1.1 ความหมายของการสังเกต

การสังเกตคือวิธีการเก็บรวบรวมข้อมูลในห้องเรียนที่สำคัญมาก (Airasian, 1997) ครูจะใช้ประสาทสัมผัสโดยเฉพาะตาและหูสังเกตการทำงานและติดตามความก้าวหน้าของนักเรียน รวมทั้งเพื่อปรับวิธีการสอน ข้อมูลจะถูกบันทึกอยู่ในรูปของมาตราประมาณค่า แบบตรวจจสอบรายการ การบรรยายแบบเรียงความ การบันทึกกระเปาะนพฤติการณ์ เป็นต้น (Hart, 1994)

1.2 ประเภทของการสังเกต

การสังเกตในห้องเรียนมี 2 ประเภทด้วยกันคือ (Airasian, 1997) (1) การสังเกตแบบไม่เป็นทางการ (informal observation) และ (2) การสังเกตแบบเป็นทางการ (formal observation) ซึ่งมีลักษณะที่แตกต่างกันอยู่ 5 ประการดังนี้

ตารางที่ 1 การเปรียบเทียบลักษณะของการสังเกตแบบเป็นทางการกับการสังเกตแบบไม่เป็นทางการ

การสังเกตแบบไม่เป็นทางการ	การสังเกตแบบเป็นทางการ
เป็นวิธีการเก็บรวบรวมข้อมูลพฤติกรรมของนักเรียนในสถานการณ์ที่เป็นธรรมชาติ (naturalistic) ครูไม่ได้วางแผนหรือควบคุมสถานการณ์ เหตุการณ์ที่สังเกตเป็นส่วนหนึ่งของเหตุการณ์ในห้องเรียนที่เป็นธรรมชาติ การสังเกตแบบไม่เป็นทางการจึงให้ภาพพฤติกรรมของนักเรียนที่มีรายละเอียดมากกว่าการสังเกตแบบเป็นทางการ	เป็นวิธีการเก็บรวบรวมข้อมูลพฤติกรรมในสถานการณ์ที่ครูวางแผนขึ้นไว้ล่วงหน้า (predetermine) เช่น โดยครูให้เวลานักเรียน 10 นาทีเตรียมสุนทรพจน์เพื่อพูดโดยใช้เวลา 5 นาที ครูจะทำการสังเกตพฤติกรรมระหว่างที่นักเรียนเตรียมตัวและระหว่างที่นักเรียนพูด
ให้ข้อมูลที่มีลักษณะเฉพาะบุคคล (idiosyncratic) เนื่องจากไม่ได้วางแผนไว้ล่วงหน้า การสังเกตในเวลาและสภาวะแวดล้อมที่แตกต่างกันทำให้ได้ข้อมูลคุณลักษณะที่แตกต่างกัน เมื่อนำมารวมกันก็จะได้ภาพของนักเรียนทุกมุมมองที่เป็นลักษณะเฉพาะบุคคล	ผลที่ได้จากการสังเกตแบบเป็นทางการมีลักษณะได้เป็นนัยทั่วไป (generalizable) การสังเกตแบบเป็นทางการได้ข้อมูลความสามารถของมนุษย์ที่ทั่วๆ ไปที่ประยุกต์ใช้กับนักเรียนหลาย ๆ คนหรือหลาย ๆ สถานการณ์ได้
มีความเป็นอัตนัย (subjective) การตีความหมายจากการสังเกตของครูเป็นสิ่งที่ยอมรับกันว่าเป็นความเห็นส่วนบุคคลที่มีความเป็นอัตนัยสูง	ผลจากการสังเกตแบบเป็นทางการมีความเป็นปรนัยสูง (objective) เนื่องจากการตีความหมายจากการสังเกตได้รับการตรวจสอบจากข้อมูลเชิงประจักษ์หรือผู้เชี่ยวชาญภายนอก

คุณลักษณะของนักเรียนที่ได้จากการสังเกตแบบไม่เป็นทางการจะเป็นข้อความเชิงบรรยาย ครูต้องสังเคราะห์ข้อความเหล่านี้เพื่อให้ได้ภาพพฤติกรรมของนักเรียน	ข้อมูลจากการสังเกตแบบเป็นทางการมีลักษณะเป็นตัวเลข (numeric) การสังเกตแบบเป็นทางการจะมีการวางแผนไว้ล่วงหน้า โดยมีแบบบันทึกการสังเกตซึ่งกำหนดคะแนนไว้แล้ว
ครูต้องสังเกตแบบปกปิด (covert) หรือมีการบันทึกน้อยที่สุดหรือไม่มีการบันทึกเลย (unrecord) การตีความหรือการบรรยายผลจะไม่ค่อยได้รับการบันทึก ครูจะสังสมไว้ในใจ ซึ่งหากใครถามก็จะบรรยายลักษณะของนักเรียนออกมาได้	การสังเกตแบบเป็นทางการเป็นวิธีการเก็บรวบรวมข้อมูลเพื่อการประเมินที่เปิดเผย (overt) เนื่องจากมีการวางแผนให้นักเรียนแสดงพฤติกรรมตามที่ครูต้องการประเมินล่วงหน้า มีการบันทึก และมีการรายงานผลให้นักเรียนทราบ

1.3 หลักการสังเกตที่ดี

การสังเกตที่ต้องทำการสังเกตอย่างมีระบบ ข้อเสนอแนะที่จะทำให้การสังเกตมีระบบมีดังนี้

1. สังเกตนักเรียนทุกคนและไม่ควรให้นักเรียนรู้ตัว โดยเฉพาะการสังเกตแบบไม่เป็นทางการสามารถทำได้ตลอดเวลา ข้อมูลที่ได้จะเป็นภาพพฤติกรรมที่นักเรียนแสดงออกจริงไม่ได้เสแสร้ง

2. สังเกตบ่อย ๆ และสังเกตอย่างสม่ำเสมอ ยิ่งถ้าเป็นพฤติกรรมเดี่ยวแต่สังเกตหลาย ๆ ครั้ง จะทำให้ผลจากการสังเกตมีความเที่ยง และถ้าทำการสังเกตพฤติกรรมเดี่ยวหลาย ๆ สถานที่จะทำให้ข้อมูลจากการสังเกตมีความตรงมากขึ้น

3. บันทึกการสังเกตโดยการเขียน ไม่ว่าจะครูจะทำการสังเกตแบบไม่เป็นทางการหรือเป็นทางการ สิ่งที่พึงตระหนักคือ การบันทึกข้อมูลจากการสังเกตและควรบันทึกในระยะเวลาที่ใกล้เคียงที่สุดกับการสังเกตเพื่อให้ได้ข้อมูลที่มีรายละเอียดมากที่สุด

การสังเกตที่ดีต้องกำหนดรายละเอียดของสิ่งที่จะค้นหาหรือสังเกต

ผู้สังเกตควรกำหนดรายละเอียดของสิ่งที่จะสังเกตให้มากที่สุดเพื่อให้การสังเกตเป็นไปอย่างมีประสิทธิภาพ ซึ่งอาจกำหนดในรูปของเครื่องมือที่ช่วยการสังเกต เช่น

1. แบบตรวจสอบพัฒนาการ (developmental checklist) เป็นเครื่องมือที่ใช้อย่างแพร่หลายและใช้ได้ทุกสถานการณ์ แบบตรวจสอบพัฒนาการจะบรรยายคุณลักษณะ (trait) หรือพฤติกรรมการเรียนรู้ที่ผู้สังเกตกำลังจะประเมิน เมื่อมีบันทึกซ้ำ ๆ ตลอดปีการศึกษาจะให้ภาพความก้าวหน้าของนักเรียนตลอดปีการศึกษาที่ชัดเจน ในชั้นประถมศึกษาอาจจะใช้พิจารณาพฤติกรรม การเรียนรู้ ทัศนคติเฉพาะ เช่น การแบ่งวรรคตอน ครูอนุบาลส่วนมากใช้แบบตรวจสอบพัฒนาการช่วยในการทำร่องรอยความก้าวหน้าของนักเรียนแต่ละช่วงของพัฒนาการจนพอจะที่ทำนายได้ ตัวอย่างแบบตรวจสอบรายการที่ใช้ในการประเมินพัฒนาการสร้างสิ่งก่อสร้างจากบล็อก (block-building) ของเด็ก ๆ ดังตารางที่ 2

ตารางที่ 2 ตัวอย่างแบบตรวจสอบรายการที่ใช้ประเมินพัฒนาการการสร้างสิ่งก่อสร้างจากบล็อก (block-building) ของเด็ก (ปรับจาก Hirsh, 1984 cited in Hart, 1994)

ลำดับขั้นพัฒนาการการสร้างสิ่งก่อสร้างจากบล็อก	
ขั้นที่ 1	บล็อกถูกวางอยู่โดยรอบ ๆ ไม่มีการต่อเป็นรูปใด ๆ
ขั้นที่ 2	การต่อขั้นเริ่มต้น : ส่วนใหญ่เด็กต่อเป็นแถวแนวราบหรือแนวตั้ง
ขั้นที่ 3	สร้างสะพาน: มีการเรียงบล็อกต่อกันสองบล็อกโดยมีช่องว่างและมีบล็อกที่สามเป็นตัวเชื่อมต่อ

ขั้นที่ 4	การทำรื้อล้อม: มีการต่อบล็อกเป็นแนวยาวเชื่อมต่อกันเป็นร้ว
ขั้นที่ 5	มีการใช้อุปกรณ์ที่ได้มากับบล็อก มีสิ่งก่อสร้าง มีการตกแต่งที่เหมาะสมเป็นส่วนที่รับกันโดยส่วนใหญ่ แต่ยังไม่มีการตั้งชื่อ
ขั้นที่ 6	มีการตั้งชื่อสิ่งก่อสร้างเพื่อเป็นจุดเริ่มต้นของการเล่นละคร แต่การตั้งชื่อของเด็กอาจจะยังไม่จำเป็นที่จะต้องเกี่ยวข้องกับสิ่งก่อสร้าง
ขั้นที่ 7	การก่อสร้างของเด็กมีการจำลองหรือมีสัญลักษณ์จากสิ่งก่อสร้างจริง ๆ ที่เข้ารู้จัก และมีแรงดลใจอย่างยิ่งใหญ่ที่จะนำไปสู่การเล่นละครรอบรอบสิ่งก่อสร้างที่เกิดจาก บล็อก











2. แบบสัมภาษณ์ (interview sheet) เป็นเครื่องมือที่รู้จักกันดีในชื่อของแบบฟอร์มบันทึกการประชุม (conference recording form) ซึ่งมีหลายรูปแบบขึ้นกับความต้องการที่จะสังเกตของผู้สัมภาษณ์ โดยทั่วไปจะประกอบด้วยรายการคำถามที่ครูวางแผนจะถามนักเรียนและมีพื้นที่ว่างสำหรับบันทึกการโต้ตอบของนักเรียน ตัวอย่างแบบสัมภาษณ์ที่ใช้ประเมินความรู้สึกเกี่ยวกับเรื่องราวที่อ่านในโครงภาษาแบบองค์รวม ดังตารางที่ 3

ตารางที่ 3 ตัวอย่างแบบสัมภาษณ์ที่ใช้บันทึกการแสดงความรู้สึกเกี่ยวกับเรื่องที่อ่านในโครงการ Whole Language (ปรับจาก Harp, 1991 cited in Hart, 1994)

แบบสัมภาษณ์เกี่ยวกับเรื่องที่อ่าน
ชื่อผู้อ่าน.....วันที่.....
ชื่อผู้สัมภาษณ์.....
ชื่อหนังสือ.....
ชื่อผู้แต่ง.....
1. ในเรื่องคุณชอบใครมากที่สุด
2. คุณชอบใครน้อยที่สุด
3. เรื่องเกิดขึ้นที่ไหน
4. เรื่องเกิดขึ้นเมื่อไร
5. ทำไมเรื่องนี้จึงทำให้คุณสนใจ
6. ผู้แต่งทำสิ่งใดที่ทำให้คุณรู้สึกแปลกใจ
7. ตอนใดของเรื่องที่ทำให้คุณรู้สึกเศร้าที่สุด
8. ตอนใดของเรื่องที่ทำให้คุณมีความสุขที่สุด
9. มีบางส่วนของเรื่องที่ทำให้คุณหัวเราะหรือร้องไห้หรือไม
10. มีอะไรบ้างที่คุณอยากจะบอกกับผู้เขียน
11. อะไรที่คุณคิดว่าจะเป็นความทรงจำเกี่ยวกับหนังสือเล่มนี้
12. คุณลักษณะอย่างไรที่คุณคิดว่าจะมีความสุขที่สุดกับการอ่านหนังสือเล่มนี้

การสังเกตที่ดีต้องมีการบันทึกการสังเกตอย่างเป็นระบบ

ครูจึงต้องสร้างระบบการบันทึกที่ใช้การเขียนให้น้อยที่สุด ซึ่งอาจจะเป็นแบบตรวจสอบรายการ หรือแบบสัมภาษณ์จะช่วยให้ครูเก็บรวบรวมข้อมูลได้อย่างมีประสิทธิภาพ ครูอาจจะออกแบบแบบตรวจสอบรายการที่สามารถใช้บันทึกการสังเกตพฤติกรรมของนักเรียนได้ทั้งห้องเรียนดังแผนภาพที่ 1 แล้วค่อยคัดลอกแบบบันทึกพัฒนาการรายบุคคลซึ่งมีตัวอย่างดังแผนภาพที่ 2

Group Developmental Record Sheet										
SYMBOLIC DEVELOPMENT IN ART										
										
Alicia	8/29	9/1	11/2							
Hoa				8/29						
Carol				8/29	9/16					
David				9/15						

แผนภาพที่ 1 ตัวอย่างแบบตรวจสอบรายการที่ทำร่องรอยพัฒนาการของนักเรียนอย่างง่าย.
(Adrienne, 1990 cited in Hart, 1994)

Individual Record of Developmental Milestones				
Name <u>Gilberto</u>				
SYMBOLIC PLAY RECORD				
Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
Real object	Similar object	Dissimilar object	No object	Sociodramatic play
8/29	9/15			
INVENTED SPELLING STAGES				
Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
Initial consonant	Initial, final consonants	Initial, medial, final consonants	Vowel as place holder	Standard spelling
9/15	11/2			

แผนภาพที่ 2 ตัวอย่างแบบบันทึกพัฒนาการรายบุคคล (Adrienne, 1990 cited in Hart, 1994)

นอกจากแบบตรวจสอบพัฒนาการและแบบบันทึกการสัมภาษณ์แล้วแบบบันทึกการสังเกตอาจจะอยู่ในรูปของมาตราประมาณค่า (rating scale) หรือระเบียบเหตุการณ์ (anecdotal record) (สำนักงานคณะกรรมการการศึกษาแห่งชาติ, 2542 ; Pike & Salend, 1995) ขึ้นอยู่กับความต้องการและการออกแบบของครู อย่างไรก็ตามสำหรับห้องเรียนในประเทศไทยที่มีจำนวนนักเรียนมาก ครูควรออกแบบแบบบันทึกให้มีการเขียนน้อย ที่สุดเพื่อไม่ให้เสียเวลาในการบันทึก

2. การประเมินจากแฟ้มสะสมงาน (Portfolio Assessment)

แนวคิดการใช้แฟ้มสะสมงาน (portfolio) นำมาจากสาขาวิชาศิลปะ (Stecher & Herman, 1997) แฟ้มสะสมงานของศิลปินหมายถึงแหล่งสะสมงานที่ศิลปินออกแบบและแสดงให้เห็นถึงรูปแบบเฉพาะของศิลปิน แฟ้มสะสมงานเริ่มเข้ามามีบทบาทในการเรียนการสอนเมื่อปี ค.ศ. 1988 โดยเป็นเครื่องมือชนิดหนึ่งของการประเมินการเรียนรู้ตามสภาพจริง (อุทุมพร (ทองอุไทย) จามรมาน, 2540)

2.1 ความหมายของการประเมินจากแฟ้มสะสมงาน

เมื่อพิจารณาตามรากศัพท์ คำว่า “portfolio” มาจากภาษาลาตินว่า “portare” เป็นคำกริยา หมายถึงการนำเอาไป การถือหรือการพกพาไปได้ (to carry) และ “folio” เป็นคำนาม หมายถึง เอกสารหรือชิ้นส่วนของกระดาษ (sheets or leaves of paper)

ดังนั้น portfolio จึงหมายถึงแหล่งที่บรรจุเอกสาร รายงาน ภาพวาด แผนที่หรือสิ่งต่าง ๆ ที่สามารถพกพานำไปที่ต่าง ๆ ได้ (Chiseri-Strater, 1992)

ความหมายของแฟ้มสะสมงานที่ใช้ทางการศึกษาคือ แหล่งบรรจุตัวอย่างการปฏิบัติงานและผลงานของนักเรียนที่ผ่านการคัดสรรอย่างเป็นระบบภายใต้ระยะเวลาอันยาวนานโดยตั้งใจจะสะท้อนให้เห็นถึงการเจริญเติบโต (growth) และการพัฒนา (development) คุณลักษณะสำคัญที่กำหนดไว้ในหลักสูตร (Pearson Education Development Group, 2001; Wiggins, 1989; Hart, 1994; Airasian, 1997) ซึ่งหลักฐานที่บรรจุลงในแฟ้มเป็นผลงานของนักเรียนในรูปแบบต่างๆ เพื่อสะท้อนให้เห็นสัมฤทธิ์ผลและพัฒนาการของผู้เรียน อาทิ ตัวอย่างงานเขียน บทความ วิดีโอบันทึกการปฏิบัติงานของนักเรียน แถบบันทึกเสียง รายงานกลุ่ม แผนที่มโนทัศน์ แผนที่จิตทัศน์ บันทึก ตาราง กราฟ รายชื่อหนังสือที่อ่าน บทกลอน เรียงความ ภาพวาด เป็นต้น ส่วนแหล่งบรรจุอาจจะเป็นแฟ้ม ดิสเก็ตต์ กล่อง ฯลฯ (Pearson Education Development Group, 2001; Hart, 1994)

การประเมินจากแฟ้มสะสมงาน หมายถึง การประเมินการเรียนรู้ตามสภาพจริง โดยพิจารณาจากแฟ้มสะสมงาน ซึ่งสามารถสะท้อนให้เห็นผลสัมฤทธิ์ ความก้าวหน้าและความพยายามทั้งที่เป็นกระบวนการเรียนรู้และผลการเรียนรู้ที่เกิดขึ้นตามสภาพจริง (ศิริชัย กาญจนวาสี, 2543; Wiggins, 1989)

2.2 ความมุ่งหมายของการประเมินจากแฟ้มสะสมงาน

ความมุ่งหมายของการประเมินจากแฟ้มสะสมงานมีดังนี้ (Frazier & Paulson, 1992 cited in Hart, 1994)

2.2.1 เพื่อให้ครูประเมินการเจริญเติบโต (growth) และความก้าวหน้าของผู้เรียน (progress)

2.2.2 เพื่อให้พ่อแม่และครูมีการสื่อสารกันในการวางแผนแก้ปัญหาหรือส่งเสริมการเรียนรู้ของนักเรียนร่วมกัน

2.2.3 เพื่อให้ครูและผู้บริหารประเมินการจัดการเรียนการสอนและนำไปใช้ในการปรับปรุงการจัดการเรียนการสอน

2.3.4 เพื่อให้นักเรียนมีส่วนร่วมกับครูในการประเมินกระบวนการเรียนรู้และสร้างนิสัยการประเมินตนเองให้กับนักเรียน

ความมุ่งหมายประการสุดท้ายเป็นความมุ่งหมายที่สำคัญยิ่ง ในกระบวนการรวบรวมงานบรรจุเข้าแฟ้มสะสมงานนั้น ต้องการให้นักเรียนการตัดสินใจคุณค่าเกี่ยวกับกระบวนการทำงานและผลงาน นักเรียนจะเริ่มพัฒนาการประเมินตนเองจนเป็นนิสัย นี่คือคุณค่าของการประเมินจากแฟ้มสะสมงาน

2.3 การจัดระบบข้อมูลในแฟ้มสะสมงาน

การจัดระบบข้อมูลในแฟ้มสะสมงานควรทำอย่างมีระบบ (organizing) จึงจะทำให้สะท้อนให้เห็นสารสนเทศของผู้เรียนอย่างสมบูรณ์ แฟ้มสะสมงานควรมีโครงสร้าง 3 ส่วน ดังนี้ (Pearson Education Development Group, 2001)

1. ส่วนนำ ประกอบด้วย คำนำ สารบัญ ข้อมูลส่วนตัว จุดประสงค์ การเรียนรู้ ข้อตกลง เกณฑ์การตัดสินผลงาน เป็นต้น

2.3.2 ส่วนเนื้อหา (content) ประกอบด้วยผลงานที่ผ่านการคัดเลือกจากนักเรียนและหลักฐานการประเมินต่าง ๆ เช่น การประเมินตนเอง การประเมินโดยเพื่อน ครู เป็นต้น

2.3.3 สรุปผลการประเมินแฟ้มสะสมผลงาน ประกอบด้วย การแสดง

ความคิดเห็นของตนเองในแฟ้มสะสมผลงาน ความรู้ลึกต่อวิชา เป็นต้น

2.4 ประเภทของแฟ้มสะสมงาน

แฟ้มสะสมงานในบริบทของห้องเรียน คือ การสะสมงานที่เป็นหลักฐานของ ความรอบรู้ทั้งทักษะ การประยุกต์ใช้ความรู้ และทัศนคติ งานในแฟ้มสะสมงานจึงเปรียบเสมือนสิ่งประดิษฐ์ (artifacts) แฟ้มสะสมงานที่มีประสิทธิภาพต้องสะท้อนจุดอ่อนและจุดแข็งของผู้เรียนแฟ้มสะสมงานสามารถจำแนกได้เป็น 3 ประเภท คือ (Pearson Education Development Group, 2001)

2.4.1 แฟ้มสะสมงานที่เน้นกระบวนการ (process oriented portfolios)

แฟ้มสะสมงานที่เน้นกระบวนการจะสะท้อนให้เห็นเรื่องราวเกี่ยวกับการเจริญเติบโตหรือพัฒนาการของผู้เรียน เนื้อหาของแฟ้มจะเป็นกระบวนการเรียนรู้และความคิดสร้างสรรค์ รวมทั้งหมดของแบบร่างงาน (drafts) สะท้อนกระบวนการ การเผชิญกับปัญหาและอุปสรรคต่างๆ ในการทำงาน ในแฟ้มจะประกอบด้วยงานที่สะท้อนทักษะใดทักษะหนึ่งตั้งแต่เริ่มต้นการทำงาน ระหว่างการทำงาน จนถึงสิ้นสุดการทำงานในหน่วยการเรียนรู้ต่างๆ เช่น การเขียนเรื่องสั้นโดยมี 3 ร่าง คือ ร่างแรก ร่างที่ปรับปรุงใหม่หลังจากที่ได้รับข้อมูลย้อนกลับจากเพื่อน ครู และร่างสุดท้าย นักเรียนต้องอธิบายแนวทางการทำงานที่ดีกว่าแนวทางอื่นได้ ซึ่งวิธีการนี้จะทำให้หาหลักฐานได้ว่า นักเรียนปรับปรุงทักษะให้ดีขึ้นได้อย่างไร ซึ่งอาจจะใช้การเขียนหรือพูดระหว่างการประชุมครู-พ่อแม่ นักเรียนจะสะท้อนกระบวนการเรียนรู้ โดยระบุว่ามีการเปลี่ยนแปลงทักษะอย่างไร และการกำหนดสิ่งที่จะทำต่อไปในปัจจุบันและในอนาคต

โดยทั่วไปแฟ้มสะสมงานที่เน้นกระบวนการจะใช้มากในระดับชั้นประถมศึกษา เนื่องจากครูต้องการทราบพัฒนาการของแต่ละบุคคลมากกว่าความสามารถเฉพาะ ผลจากการปรับกระบวนการทัศนการศึกษาใหม่ตามพระราชบัญญัติการศึกษาแห่งชาติ พุทธศักราช 2542 จึงได้มีการปรับเปลี่ยนกระบวนการประเมินใหม่ โดยในหมวด 4 มาตรา 26 บัญญัติว่า ให้สถานศึกษาจัดการประเมินผู้เรียนโดยพิจารณาจากพัฒนาการของผู้เรียน...การประเมินจากแฟ้มสะสมงานที่เน้นกระบวนการจึงมีความเหมาะสมอย่างยิ่งกับการประเมินแนวใหม่ตามพระราชบัญญัติการศึกษาแห่งชาติ พุทธศักราช 2542

2.4.2 แฟ้มสะสมงานที่เน้นผลผลิต (product oriented portfolios)

แฟ้มสะสมงานที่เน้นผลผลิต เป็นแหล่งสะสมงานที่ดีที่สุดของนักเรียน วัตถุประสงค์ในการสะสมหรือสะท้อนคุณภาพและความสำเร็จมากกว่ากระบวนการที่ก่อให้เกิดผลผลิต โดยทั่วไปแล้วต้องการให้นักเรียนสะสมงานตั้งแต่ต้นจนจบ เมื่อถึงเวลาแล้วนักเรียนต้องเลือกสิ่งประดิษฐ์ที่เป็นตัวแทนงานที่มีคุณภาพมากที่สุด

แนวทางในการเลือกชิ้นงานโดยทั่วไปครูสามารถกำหนดขอบเขตไว้กว้าง ๆ ได้ว่าแฟ้มสะสมงานจะบรรจุสิ่งใดและคุณภาพเป็นอย่างไรแล้วให้นักเรียนตัดสินใจเลือกเอง เช่น ครูคณิตศาสตร์กำหนดเงื่อนไขว่า แฟ้มสะสมงานต้องบรรจุหลักฐานที่แสดงความสามารถในการประยุกต์ใช้แนวคิดของค่ามัธยฐานและฐานนิยม ครูต้องกำหนดเงื่อนไขด้วยว่า หลักฐานเหล่านั้นต้องมีคุณภาพเหมาะสมที่จะเข้าไปอยู่ในแฟ้ม การทำเช่นนี้จะทำให้แฟ้มสะสมงานที่เน้นผลผลิตมีประสิทธิภาพในการควบคุมนักเรียนให้สามารถให้ เหตุผลในการผลิตงานที่มีคุณภาพได้ โดยทั่วไปนักเรียนจะสะท้อนงานแต่ละชิ้นในแฟ้มสะสมงานที่เน้นผลผลิตด้วยตนเอง ซึ่งมักจะใช้การเขียนแสดงเหตุผลว่าทำไมงานชิ้นนั้นจึงเป็นตัวแทนงานที่ดีที่สุด

ปัจจัยที่ทำให้แฟ้มสะสมงานที่เน้นผลผลิตใช้มากในระดับชั้นมัธยมศึกษา มีสองประการคือ ประการแรก เกรดเฉลี่ยและคะแนนจากแบบทดสอบมีความเสี่ยงสูงที่จะตัดสินผลการเรียนของนักเรียน เนื่องจากระดับชั้นนี้สิ่งที่ควรให้เกิดกับนักเรียนในท้ายที่สุดเน้นการเรียนรู้เกี่ยวกับสิ่งแวดล้อม ประการที่สอง นักเรียนมีอายุมากขึ้น จึงจำเป็นต้องใช้ทักษะการคิดระดับสูงในการเลือกงานที่ดีที่สุดอย่างชาญฉลาด ทำนองเดียวกันการรับประกันโดยการสะท้อนตนเองก็ต้องให้เหตุผลลึกซึ้งขึ้นเช่นกัน

แต่อย่างไรก็ตามไม่มีแฟ้มสะสมงานแบบใดที่มีความเหมาะสมไปทุกระดับชั้น ครูควรเลือกใช้โดยคำนึงถึงความน่าจะเป็น รูปแบบการสอนและบริบทของโรงเรียน

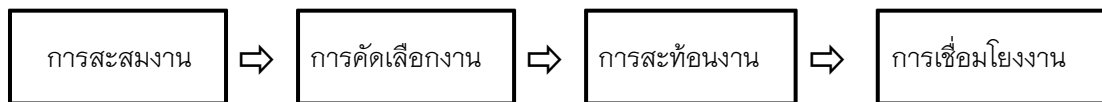
2.4.3 การแสดงนิทรรศการต่อสาธารณชน

การแสดงนิทรรศการต่อสาธารณชนเป็นชนิดหนึ่งของแฟ้มสะสมงาน ที่ประกอบด้วยการรวมตัวของกลุ่มเพื่อน ครู พ่อแม่หรือสมาชิกในชุมชน นักเรียนเพื่อนำเสนอแฟ้มสะสมงานทั้งหมดหรือบางส่วน ในบางกรณีนักเรียนจะมีการสอบป้องกันงาน เช่น นักศึกษาระดับบัณฑิตศึกษาสอบป้องกันวิทยานิพนธ์ หรือบางกรณีกลุ่มนักเรียนอาจจัดแสดงแฟ้มสะสมงานในรูปแบบนิทรรศการเช่นเดียวกับการแสดงนิทรรศการในพิพิธภัณฑ์หรือบางกรณี นักเรียนอาจจะพัฒนาบางส่วนของแฟ้มสะสมงานให้ดีขึ้น ลึกซึ้งขึ้น มีการสะท้อนให้เห็นความสนใจทางอาชีพและทางวิชาการของนักเรียนแต่ละคน

การแสดงนิทรรศการต่อสาธารณชนเป็นการหล่อหลอมเชื่อมโยงระหว่างนักเรียนกับชุมชน แฟ้มสะสมงานของนักเรียนจะถูกประเมินโดยผู้อื่น ซึ่งจะทำให้กระบวนการโดยรวมมีความตรงมากขึ้น นอกจากนี้ยังเป็นหนทางที่จะนำชุมชนพ่อแม่เข้ามามีส่วนร่วมในกระบวนการเรียนอีกด้วย

2.5 ขั้นตอนการประเมินจากแฟ้มสะสมงาน

สิ่งที่ครูควรตระหนักอยู่เสมอ คือ แฟ้มสะสมงานจะต้องได้รับการพัฒนาโดยนักเรียน ไม่ใช่ครู ครูจะทำหน้าที่ให้คำปรึกษาแก่นักเรียน นักเรียนจะมีอิสระในการคัดเลือกผลงาน บรรลุเข้าแฟ้มสะสมผลงาน การประเมินจากแฟ้มสะสมงานจะมีการทำงานร่วมกันระหว่างครูกับนักเรียน โดยครูจะทำหน้าที่เป็นเพียงผู้แนะนำไม่ใช่ผู้ชี้แนะ การประเมินจากแฟ้มสะสมงานมีขั้นตอนตามแผนภาพที่ 3 (Pearson Education Development Group , 2001)



แผนภาพที่ 3 ขั้นตอนการประเมินจากแฟ้มสะสมงาน

2.5.1 การสะสมงาน (collection)

การสะสมงานเป็นขั้นตอนที่นักเรียนจะต้องสะสมและเก็บรักษางานทั้งหมดสิ่งที่เป็นอุปสรรคที่ทำให้ขั้นตอนนี้ไม่บรรลุผลมี 3 ประการ คือ ประการแรกนักเรียนและครูส่วนใหญ่ไม่เคยชินกับการสร้างเอกสารและการเก็บรักษาในงานในชั้นเรียน ประการที่สองโรงเรียนแทบจะไม่มีสถานที่โล่งแจ้งเพื่อที่จะเก็บสะสมงานจำพวกงานศิลปะหรืองานการทดลอง ประการสุดท้ายการผลิตงานโดยทั่วไปมักจะไม่มีการผลิตผลงานหรือชิ้นงาน เช่น งานหรือโครงการบริการชุมชน กีฬา ทักษะที่สำคัญในขั้นตอนนี้ คือ ทำให้นักเรียนเคยชินกับการสะสมและการสร้างชิ้นงานและก่อนที่จะสะสมต้องทำความเข้าใจกับนักเรียนเกี่ยวกับรายละเอียดของแฟ้มสะสมงาน แนวทางการให้คะแนน เพื่อที่นักเรียนจะได้ยึดเป็นเป้าหมายในการทำงานให้สอดคล้องกับเงื่อนไขที่ครูกำหนดไว้

2.5.2 การคัดเลือกงาน (selection)

การคัดเลือกงานขึ้นอยู่กับชนิดของแฟ้มสะสมงาน (เน้นผลผลิตหรือเน้นกระบวนการ) และความก้าวหน้าอะไรที่ครูกำหนดเป็นเงื่อนไขที่จำเป็นต้องมีในแฟ้มสะสมงานในขั้นตอนนี้ นักเรียนจะต้องผ่านการทำงานมาเป็นระยะเวลาสั้น สะสมไว้แล้วจึงเลือกงานที่จะนำมาใส่ในแฟ้มสะสมงาน นักเรียนอาจจะเลือกตามลำพังหรือเลือกกับเพื่อนและ/หรือกับครู คุณลักษณะต่าง

ๆ ที่กำหนดไว้ในหลักสูตรจะเป็นแนวทางในการคัดเลือกงาน ซึ่งเป็นงานปกติที่นักเรียนทำในชั้นเรียนกรณีตัวอย่าง ถ้านักเรียนทุกคนนำผลการแสดงในแฟ้มสะสมงาน แฟ้มสะสมงานนั้นก็จะมีอะไร ความหมายอะไร ไม่สะท้อนอะไรให้เห็นมากกว่าความสามารถในการทำแบบทดสอบของนักเรียน แต่ถ้านักเรียนนำงานที่หลากหลาย เช่น แถบบันทึกเสียง วิดีโอ งานศิลปะ โครงการ งานเขียน การบ้าน งานคอมพิวเตอร์ เรียงความ ฯลฯ แฟ้มสะสมงานก็จะมีชีวิตชีวา แฟ้มสะสมงานจึงกระตุ้นให้เกิดความหลากหลายในวิธีการสอนและวิธีการเรียน

2.5.3 การสะท้อนงาน (reflection)

การสะท้อนงานเป็นขั้นตอนที่มีความสำคัญที่สุดในการประเมินจากแฟ้มสะสมงาน การสะท้อนงานทำให้แฟ้มสะสมงานมีความแตกต่างจากการสะสมตามปกติ นักเรียนมักจะสะท้อนงานโดยการเขียน ซึ่งถ้าสามารถสะท้อนงานด้วยปากเปล่าได้ตามหลักการของการประเมินการเรียนรู้ตามสภาพจริงจะเป็นการดี เด็กเล็กมักจะใช้การสะท้อนงานด้วยปากเปล่า นักเรียนจะถูกร้องขอให้อธิบายว่าทำไมจึงเลือกงานชิ้นนี้ เมื่อเปรียบเทียบกับชิ้นอื่นแล้วเป็นอย่างไร ความรู้หรือทักษะอะไรที่นักเรียนใช้ในการทำงานครั้งนี้ นักเรียนสามารถปรับปรุงได้ในฐานะผู้เรียนรู้ ความสำคัญ of ขั้นตอนนี้คือการให้บทบาทนักเรียนเข้าสู่กระบวนการประเมิน นักเรียนสามารถสังเกตผลสัมฤทธิ์ของตนเองได้โดยตรง และเมื่อประเมินตนเองจนเป็นนิสัยก็สามารถปรับปรุงงานได้ตามต้องการ

2.5.4 การเชื่อมโยงงาน (connection)

การเชื่อมโยงงานนับว่าเป็นขั้นตอนสุดท้ายของการประเมินจากแฟ้มสะสมงานโดยมีสถานการณ์ที่เกี่ยวข้องกันสองสถานการณ์ สถานการณ์แรกผลโดยตรงจากกระบวนการสะท้อนงาน นั่นคือ การพิจารณาผลสัมฤทธิ์ของตนเองและระบุทักษะและความรู้ที่นำมาประยุกต์ใช้ในการสร้างงาน นักเรียนสามารถตอบคำถามที่ถามตลอดเวลว่า “ทำไมเราจึงทำอย่างนี้” ได้ด้วยตนเอง พวกเขาสามารถเชื่อมโยงระหว่างงานในโรงเรียนกับคุณค่าของสิ่งที่เขาเรียนรู้ นอกจากนี้ นักเรียนควรเข้ามาทำความเข้าใจหลักสูตรทั้งหมดให้ลึกซึ้งยิ่งขึ้นโดยมีบทบาทในการประเมินประสิทธิภาพของหลักสูตร

สถานการณ์ที่สองเกี่ยวกับการสร้างความเชื่อมโยงกับโลกนอกชั้นเรียน โครงการแฟ้มสะสมงานหลายโครงการที่ประสบความสำเร็จ โดยทั่วไปแล้วจะให้นักเรียนนำเสนอ แฟ้มสะสมงานต่อกลุ่มเพื่อน ครู พ่อแม่ และสมาชิกในชุมชน หรืออาจจะทำลักษณะเดียวกันกับที่นักศึกษาบัณฑิตศึกษาสอบป้องกันวิทยานิพนธ์ หรือจัดนิทรรศการแฟ้มสะสมงาน ผลจากการเชื่อมโยงระหว่างนักเรียนกับชุมชนทำให้เกิดสัมฤทธิ์ผลแก่นักเรียนอย่างไม่สามารถประเมินค่าได้ และยังเป็นการวินิจฉัยคุณภาพการเรียนรู้ในโรงเรียนในท้องถิ่นอีกด้วย

2.6 ข้อดีและข้อจำกัดของการประเมินจากแฟ้มสะสมงาน

ข้อดีและข้อจำกัดของการประเมินจากแฟ้มสะสมงานสรุปได้ดังนี้ (Pearson Education Development Group, 2001; สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ, 2542; อุทุมพร (ทองอุไทย) จามรมาน, 2540)

ข้อดีของการประเมินจากแฟ้มสะสมงาน

1. แฟ้มสะสมงานช่วยให้ครูเข้าใจคุณลักษณะ ความต้องการ และจุดเด่นของนักเรียนรายบุคคลได้ดีกว่าแบบทดสอบโดยเฉพาะแบบทดสอบปรนัย
2. แฟ้มสะสมงานจะเปลี่ยนแปลงบทบาทของครูจากการตัดสินผลสัมฤทธิ์ทางการเรียนของนักเรียนมาเป็นบทบาทการพัฒนาผลสัมฤทธิ์ทางการเรียนโดยการให้ข้อมูลย้อนกลับและให้นักเรียนประเมินตนเอง
3. แฟ้มสะสมงานช่วยครูประเมินทักษะและความรู้ที่คาดหวังว่าจะเกิดกับนักเรียนโดยไม่จำกัดความคิดสร้างสรรค์ ให้โอกาสกับนักเรียนทำงานที่ตรงกับความรู้ ความสามารถตนเองอย่างเต็มที่

4. นักเรียนได้รับโอกาสในการสะท้อนพัฒนาการและผลงานในฐานะผู้เรียน
 แพ้ผสมงานช่วยให้ นักเรียนสามารถชี้แจงเหตุผลเกี่ยวกับงานที่นักเรียนทำในห้องเรียนรวมทั้งทักษะ
 และความรู้ที่เกิดขึ้นกับตัวนักเรียนได้

5. แพ้ผสมงานช่วยเชื่อมโยงวิธีการสอนวิธีการเรียนและวิธีการประเมิน
 เข้าด้วยกัน

6. นักเรียนได้มีส่วนร่วมในกระบวนการประเมิน ซึ่งเป็นบทบาทที่สำคัญต่อ
 การปรับปรุงผลสัมฤทธิ์ทางการเรียนมากทำให้นักเรียนไม่กังวลต่อการสอบเช่นเดิม

7. พ่อแม่ และชุมชนมีส่วนร่วมในการประเมินผลสัมฤทธิ์เชิงวิชาการตาม
 หลักสูตรของโรงเรียนร่วมกับครูและนักเรียน

8. การประเมินจากแพ้ผสมงานช่วยทำให้เกิดระบบประกันคุณภาพ (quality
 assurance) แก่บุคคลภายนอก

ข้อจำกัดของการประเมินจากแพ้ผสมงาน

1. การประเมินจากแพ้ผสมงานมีความเที่ยงหรือความยุติธรรมน้อยกว่าการ
 ประเมินที่มีมาตรฐานหรือการประเมินเชิงปริมาณ ดังเช่น แบบทดสอบ

2. พ่อแม่มักจะรู้สึกเคลือบแคลงการวัดผลที่นอกเหนือไปจากการทดสอบและ
 วิทยาลัยและมหาวิทยาลัยส่วนใหญ่ยังคงใช้คะแนนสอบและการให้ระดับคะแนนเป็นเกณฑ์ในการรับเข้า
 ศึกษาดังอยู่เช่นเดิม

3. ทำให้ครูและผู้ที่เกี่ยวข้องต้องเสียเวลาในการผสมงาน ซึ่งต้องมีจำนวนงานมาก
 พอและงานต้องสะท้อนความรู้และทักษะของผู้เรียน หากเปรียบเทียบกับทดสอบจะเสียเวลามาก

4. ครูต้องพัฒนาเกณฑ์ด้วยตนเอง ซึ่งเป็นสิ่งที่ยากหรือเป็นสิ่งที่ไม่คุ้นเคย และ
 ไม่มีหลักประกันว่ามาตรฐานหรือเกณฑ์นั้นจะถูกยอมรับหรือไม่ และข้อมูลที่ได้รับจากการประเมิน
 จากแพ้ผสมงานยากต่อการวิเคราะห์หรือรวม และต้องใช้เวลายาวนาน

5. แพ้ผสมงานมักจะยากที่จะบูรณาการให้เข้ากับวัฒนธรรมของโรงเรียน
 และมีความเสี่ยงสูงในการจัดอันดับนักเรียนในเชิงเปรียบเทียบ

3. การประเมินการปฏิบัติ (Performance Assessment)

การประเมินการปฏิบัติ เป็นหนึ่งในหัวข้อที่กล่าวถึงมากในการปฏิรูปการศึกษา
 เนื่องจากเป็นวิธีการประเมินที่เข้าถึงความสามารถของนักเรียนมากกว่าการใช้แบบทดสอบ
 มาตรฐานหรือแบบทดสอบแบบเลือกตอบในการทดสอบแบบเดิม (Esiner, 1999) ผลจากการสอบ
 ด้วยแบบทดสอบเลือกตอบจะชี้ให้เห็นว่านักเรียนทำอะไรได้บ้าง ในขณะที่ผลจากการประเมินการ
 ปฏิบัติจะทำให้ทราบว่ามีนักเรียนมีความเข้าใจและประยุกต์ใช้ความรู้ได้อย่างไร (Brualdi, 1999) ดัง
 นั้นถ้าครูสามารถบูรณาการการประเมินการปฏิบัติเข้ากับกระบวนการจัดการเรียนการสอนได้ก็จะ
 เป็นการเพิ่มประสบการณ์เรียนรู้ให้กับนักเรียนไปด้วย (Brualdi, 1998)

3.1 ความหมายการประเมินการปฏิบัติ

การประเมินการปฏิบัติเป็นการประเมินที่ออกแบบเพื่อทดสอบความสามารถ
 ของนักเรียนในการประยุกต์ใช้ความรู้ (knowledge) ทักษะ (skills) และนิสัยการทำงาน (work habits)
 ผ่านการปฏิบัติงานในบริบทหรือสถานการณ์จริงหรือสอดคล้องกับสภาพจริงมากที่สุด โดยเน้นให้นักเรียน
 ทำงานร่วมกันและประยุกต์ทักษะและแนวคิดในการแก้ปัญหาที่ซับซ้อน (wiggins, 1989; Pearson
 Education Development Group, 2001; Hart, 1994:)

3.2 ความมุ่งหมายของการประเมินการปฏิบัติ

Kentucky Department of Education (Hart, 1994) ได้นำการประเมินการ
 ปฏิบัติไปใช้ในโครงการการปฏิรูปการศึกษา โดยได้พัฒนางานที่จะให้นักเรียนปฏิบัติเพื่อที่จะ
 ทดสอบความก้าวหน้าของนักเรียนโดยมีความมุ่งหมายอยู่ 6 ประการ คือ (1) เพื่อให้นักเรียน
 ประยุกต์ใช้การสื่อสารขั้นพื้นฐานและทักษะทางคณิตศาสตร์ (2) เพื่อให้นักเรียนประยุกต์ใช้หลักการ

และแนวคิดหลักจากเนื้อหาทุกสาขาวิชา (3) เพื่อให้ตอบสนองกับความสามารถของบุคคลที่มีความหลากหลาย แตกต่างกันไปแต่ละบุคคล (4) เพื่อให้ตอบสนองกับสมาชิกของครอบครัว กลุ่มทำงานหรือชุมชนที่มีอำนาจในการผลิตสูง (5) เพื่อให้ให้นักเรียนได้ใช้ความคิดและทักษะการแก้ปัญหา และ (6) เพื่อให้ให้นักเรียนได้มีการเชื่อมโยงและการบูรณาการความรู้กับชีวิตจริง

3.3 ขั้นตอนการประเมินการปฏิบัติ

นักการศึกษาได้เสนอแนะขั้นตอนการประเมินการปฏิบัติไว้หลายท่าน อาทิ Brualdi (1998) Brualdi (1999) Wiggins (1989) และ Airasian (1997) ซึ่งโดยรายละเอียดแล้วมีความสอดคล้องกันสรุปได้ดังนี้

1. กำหนดวัตถุประสงค์ในการประเมิน

การประเมินที่ดีจะต้องกำหนดวัตถุประสงค์ในการประเมิน (Brualdi, 1998) ขั้นตอนนี้คือการระบุผลผลิตทางการศึกษาที่ต้องการวัดนั่นเอง คำถามสำคัญที่ครูจะต้องถามตนเองซึ่ง Wiggins (1989) ได้เสนอแนะคำถามไว้ว่า “ผู้เรียนถูกคาดหวังให้รู้อะไรและถูกคาดหวังให้มีความสามารถในการทำอะไร” นอกจากนี้ Stiggins (1994 cited in Brualdi, 1998) ยังเสนอแนะให้ครูระบุแนวคิด ทักษะ หรือความรู้ที่จะประเมินให้ชัดเจน รวมทั้งระบุลักษณะความรู้ อาทิ การให้เหตุผล ความจำหรือกระบวนการ โดยทั่วไปแล้วการประเมินการปฏิบัติเป็นการประเมินการเรียนรู้ตามสภาพจริงและเกี่ยวข้องกับงานที่มีลักษณะซับซ้อน การเรียนรู้ที่จะทำการวัดย่อมมากกว่ามิติหรือขอบเขตเดียว และเกินกว่าขอบเขตของวัตถุประสงค์ที่กำหนดไว้ตามหลักสูตร (Wiggins, 1989)

2. ออกแบบงานหรือกิจกรรมที่จะให้ผู้เรียนปฏิบัติ

งานในการประเมินการปฏิบัติที่ดีต้องเป็นงานที่มีความน่าสนใจเป็นกิจกรรมที่คุ้มค่าต่อการปฏิบัติ (Herman, Aschbacher & Winters, 1992) งานดังกล่าวไม่ควรจะเป็นกิจกรรมที่ประเมินแค่ความรู้ แต่ควรเป็นกิจกรรมที่ส่งเสริมให้ผู้เรียนเกิดความเข้าใจจากการประยุกต์ใช้ความรู้ (Wiggins, 1989) โดยเน้นการประเมินความรู้และทักษะที่หลากหลาย (Kane & Khattri, 1995) เช่นความเข้าใจทักษะการคิดที่ซับซ้อนและคุณลักษณะนิสัยของผู้เรียนซึ่งมักจะไม่สามารถวัดได้โดยแบบทดสอบแบบเดิม (Hart, 1994) และในขั้นตอนนี้ควรระบุว่า จะประเมินการทำงานเป็นกลุ่มหรือทำงานรายบุคคล โดย Herman, Aschbacher & Winters (1992) ได้เสนอแนะเกณฑ์ในการออกแบบงานไว้ดังนี้

- (1) เป็นงานที่สะท้อนผลผลิตที่ต้องการวัด
- (2) เป็นงานที่ผู้เรียนใช้ทักษะการคิดเชิงวิเคราะห์
- (3) เป็นงานที่มีความคุ้มค่ากับการปฏิบัติ
- (4) เป็นการประเมินที่ประยุกต์ใช้เงื่อนไขของงานจากโลกความเป็นจริง
- (5) เป็นงานที่ทำครั้งเดียวแต่สามารถใช้วัดผลผลิตที่หลากหลายได้
- (6) เป็นงานที่มีความยุติธรรมและปราศจากความลำเอียง
- (7) เป็นงานที่มีความน่าเชื่อถือและมีความเป็นไปได้ และ
- (8) เป็นงานหรือกิจกรรมที่มีค่านิยมชัดเจน

3. กำหนดเกณฑ์การปฏิบัติงาน

หลังจากที่ได้กำหนดกิจกรรมว่าประกอบด้วยงานใดบ้าง ขั้นตอนต่อไปควรกำหนดเกณฑ์ในการปฏิบัติงาน การประเมินการปฏิบัติเป็นการประเมินทางเลือกใหม่ที่แตกต่างจากการสอบแบบเดิมที่ตัดสินคำตอบว่าผิดหรือถูกได้อย่างชัดเจน (Brualdi, 1998) การประเมินแบบใหม่จะแสดงให้เห็นถึงระดับความสำเร็จของผู้เรียนโดยใช้เกณฑ์การให้คะแนนที่เรียกว่า “รูบรีค” (Hart, 1994; Wiggins, 1989) ครูอาจใช้รูบรีคที่มีอยู่ในธนาคารรูบรีค โดยพิจารณาความเหมาะสมกับบริบทที่นำไปใช้ ในกรณีที่มีการสอนหลายคนควรให้ครูที่สอนทุกคนมีส่วนร่วมในการสร้างรูบรีค เพื่อจะทำให้รูบรีคเป็นเกณฑ์การให้คะแนนที่มีคุณภาพยิ่งขึ้น หรืออาจจะให้ผู้เรียนช่วยกำหนดองค์ประกอบของโครงการหรืองานสำคัญที่จะทำให้การปฏิบัติงานสำเร็จ อย่างไรก็ตาม

ครูต้องให้ผู้เรียนได้รับรู้รายละเอียดของเกณฑ์การให้คะแนนแบบรูบรีคที่จะประเมินผู้เรียน เพื่อช่วยในการกำหนดทิศทางหรือวางแผนการทำงานของนักเรียน (Brualdi, 1999; Airasian, 1997)

4. การประเมินการปฏิบัติงาน

เมื่อกำหนดเกณฑ์แล้วในขั้นต่อไปจะเป็นขั้นตอนการประเมินนอกจากครูประเมินการปฏิบัติงานของนักเรียนแล้ว ต้องไม่ละเลยที่จะให้โอกาสผู้เรียนประเมินตนเองเพื่อให้ผู้เรียนรู้จักตัวเอง และใช้วิธีการหลากหลายในการประเมิน เช่น การใช้แบบตรวจสอบรายการ (Checklist) การใช้แบบมาตราวัดประมาณค่า (Rating scale) แบบสังเกต (Observations) การจัดอันดับ (Ranking) การรายงานตนเอง (Self report) ฯลฯ พิจารณาเครื่องมือที่เหมาะสมกับงานที่ให้ปฏิบัติ เช่นงานบ้านนักเรียนต้องไปปฏิบัติที่บ้าน ดังนั้น เครื่องมือที่เหมาะสมควรเป็น แบบรายงานตนเอง งานปลูกผักสวนครัวปลอดภัยอาจใช้แบบสังเกต หรือ แบบตรวจสอบรายการ งานจัดสวนถาดขนาดเล็ก อาจใช้แบบตรวจสอบรายการหรือมาตราวัดประมาณค่า เป็นต้น (Brualdi, 1999; Airasian, 1997)

3.4 ขอบข่ายของการประเมินการปฏิบัติ

การวัดการปฏิบัติจะมีขอบข่ายหลักๆ ครอบคลุมปัจจัยนำเข้า กระบวนการ และผลผลิต ซึ่งในส่วนต่าง ๆ นั้นยังต้องคำนึงถึงพฤติกรรมการปฏิบัติงาน ได้แก่ ความรับผิดชอบ ความตั้งใจ ระเบียบวินัยความร่วมมือเป็นต้นการแยกประเมินกระบวนการและผลผลิตในบางพฤติกรรมที่ผู้เรียนแสดงออกอาจจะทำได้ยาก เช่น การร้องเพลง การว่ายน้ำ การร่ำวง ฯลฯ หากเป็นเช่นนี้ให้พิจารณาจุดประสงค์การประเมินเป็นสำคัญว่าต้องการวัดส่วนใด การกำหนดขอบข่ายงานให้ครอบคลุมอาจใช้การวิเคราะห์ระบบงานเข้ามาช่วยในการจำแนกงานเพื่อสะดวกในการประเมิน เช่น การประเมินการประกอบอาหาร อาจจะกำหนดประเด็นในการประเมินดังนี้

การประเมินการประกอบอาหาร

INPUT	PROCESS	OUTPUT
<ul style="list-style-type: none"> -การวางแผน -การเตรียมวัสดุในการประกอบอาหาร -เตรียมอุปกรณ์ เครื่องมือ เครื่องใช้ -ความรับผิดชอบ ฯลฯ 	<ul style="list-style-type: none"> -การจัดเตรียมอาหารสด -การปรุงอาหาร -ความร่วมมือในการทำงาน -การจัดเก็บบำรุงรักษาอุปกรณ์ -ความตั้งใจในการปฏิบัติงาน -ความสะอาด ฯลฯ 	<ul style="list-style-type: none"> -คุณภาพของอาหาร รสชาติ สี กลิ่น ฯลฯ -ปริมาณของอาหารกับจำนวนผู้รับประทาน -การลักษณะการจัดอาหาร ฯลฯ

การวิเคราะห์ขอบข่ายงานดังกล่าว อาจจะใช้การทดลองทำงานแล้วเขียนเป็นขั้นตอนจะช่วยจะทำให้ผู้ประเมินสามารถออกแบบการประเมินได้ชัดเจนยิ่งขึ้น

3.5 ระดับของทดสอบการปฏิบัติ

1. **ระดับการรับรู้ (Cognition)** เป็นการทดสอบว่าผู้เรียนมีความรู้ด้านวิธีการปฏิบัติมากน้อยเพียงใดผู้เรียนยังไม่ได้ลงมือปฏิบัติจริง แบบทดสอบที่จะใช้ในการวัดการปฏิบัติระดับการรับรู้มีดังนี้

1.1. การทดสอบเชิงจำแนก (Identification test) เป็นการทดสอบเพื่อวัดความรู้ความสามารถในการจำแนกเครื่องมือหรือชิ้นส่วนเครื่องมือ แบบทดสอบเชิงจำแนกโดยทั่วไปไม่เป็นงานที่ต้องปฏิบัติจริงแต่จะเป็นขั้นเริ่มต้นของการปฏิบัติเช่น ให้ฟังเสียงดนตรีแล้วตอบว่าเป็นเครื่องดนตรีประเภทใดเสียงโน้ตตัวใด ให้ระบุสาเหตุที่หลอดไฟฟ้าไม่ติดพร้อมทั้งระบุเครื่องมือวัสดุอุปกรณ์ที่ใช้ในการซ่อมแซม การให้ระบุส่วนที่ชำรุดของเครื่องยนต์จากการฟังเสียงพร้อมทั้งระบุกระบวนการซ่อมแซม เป็นต้น การสร้างแบบทดสอบเชิงจำแนกจึงต้องมีการเตรียมวัสดุอุปกรณ์ให้พร้อม และต้อง

รักษาให้อยู่ในสภาพที่สามารถจำแนกได้และการให้คะแนน เกณฑ์การให้คะแนนเป็นแบบเดียวกัน กำหนดแนวทางไว้ก่อนเพื่อสังเกตและให้คะแนน

1.2. การทดสอบภาคปฏิบัติด้วยการแบบทดสอบข้อเขียน (Written test) เป็นการวัดความรู้เกี่ยวกับทฤษฎีหรือความรู้เกี่ยวกับขั้นตอนการปฏิบัติของผู้เรียน ซึ่งอาจจะใช้แบบทดสอบวัดผลสัมฤทธิ์ (Achievement test)

2. **ระดับการปฏิบัติจากสถานการณ์จำลอง (Simulated Performance)** เป็นการทดสอบโดยกำหนดสถานการณ์ที่ใกล้เคียงหรือเสมือนจริง เช่น การทดสอบการบันทึกประชุมในสถานการณ์จำลอง การทดสอบขับรถยนต์ในสถานที่ที่จัดให้มีการฝึกขับเครื่องบินจากคอมพิวเตอร์ฯลฯ

3. **ระดับการปฏิบัติงานจริงโดยใช้ตัวอย่างงาน (Work Sample)** เป็นการทดสอบโดยให้ผู้เรียนปฏิบัติงาน งานนั้นจะเป็นตัวอย่างในการสอบวัดความสามารถในการปฏิบัติงาน เป็นการวัดที่มีสภาพใกล้เคียงความเป็นจริงมากที่สุด เช่น การทดสอบการบันทึกประชุมในสถานการณ์จริง การขับรถยนต์ในสถานที่จริง ฯลฯ ซึ่งถูกนำมาใช้ในหลายวงการ เช่น การแสดง การทดลองอย่างไรก็ตามการให้คะแนนจำเป็นต้องมีมาตรฐานเดียวกัน เงื่อนไขสถานการณ์ วัสดุ อุปกรณ์ของผู้สอบแต่ละคนควรต้องเหมือนกันทุกอย่าง อาจใช้แบบตรวจสอบรายการ หรือมาตราประมาณค่าเป็นเครื่องมือในการประเมิน พฤติกรรมหรือลำดับขั้นตอนในการทำงาน รวมถึงระดับของการปฏิบัติที่จะยอมรับต้องอธิบายวิธีการและเกณฑ์การปฏิบัติให้กับผู้สอบเพื่อให้ผู้สอบเข้าใจถึงผลหรือความคาดหวังที่จะเกิดขึ้น

นอกจากการยึดระดับของการทดสอบแล้ว อาจจะพิจารณาจากรูปแบบของงานในการประเมินการปฏิบัติซึ่งรูปแบบของงานในการประเมินการปฏิบัติมี 3 ประเภท ดังนี้ (Hart, 1994)

1. งานสำหรับการประเมินในระยะเวลาอันสั้น

งานสำหรับการประเมินในระยะเวลาอันสั้น (short assessment tasks) มักจะใช้บ่อยในการตัดสินความรอบรู้ของผู้เรียนเกี่ยวกับแนวคิด วิธีการ ความสัมพันธ์ ทักษะการคิดพื้นฐานที่กำหนดในหลักสูตร โดยทั่วไปแล้วงานลักษณะนี้จะใช้เวลาในการปฏิบัติไม่นานนัก มักจะประเมินรายบุคคล จะประกอบไปด้วยสิ่งเร้า (stimulus) ที่ออกแบบเป็นลำดับขั้นยึดความสนใจของผู้เรียนเป็นหลัก เช่น ปัญหา การ์ตูน แผนที่ แผนภาพ หรือรูปภาพ ฯลฯ จากนั้นตามด้วยคำอธิบายงาน เช่น ให้ตีความ บรรยาย คำนวณ อธิบาย ทำนาย แนะนำการทดลองหรือจัดวางตำแหน่ง เป็นต้น ลักษณะของงานสำหรับการประเมินในระยะเวลาอันสั้น เช่น การแก้สมการคณิตศาสตร์ การวาดรูป หรือสร้างแผนภาพ แผนที่มโนทัศน์ (concept mapping)

2. งานที่มีลักษณะเป็นเรื่องราวหรือเหตุการณ์

งานที่มีลักษณะเป็นเรื่องราวหรือเหตุการณ์ (event tasks) ถูกออกแบบเพื่อประเมินความสามารถกว้าง ๆ เช่น ความคล่องแคล่วในการเขียน ทักษะการแก้ปัญหา ส่วนใหญ่ใช้ในวิชาเฉพาะด้านที่ต้องการประเมินการประยุกต์ใช้ความรู้ งานลักษณะนี้จึงต้องอาศัยเวลาในการประเมินพอสมควร และมักจะใช้ในการประเมินการทำงานเป็นทีมหรือเป็นกลุ่มผ่านการสังเกตการปฏิบัติงาน มีการประเมินตนเอง ประเมินโดยเพื่อน หรือรวมทั้งสองวิธี งานที่จัดอยู่ในกลุ่มนี้ได้แก่ การประเมินการเขียนการอ่าน การประเมินกระบวนการทำงานงานที่ต้องใช้การวิเคราะห์และการแก้ปัญหา เป็นต้น

3. งานสำหรับการประเมินระยะยาว

งานสำหรับการประเมินระยะยาวหรือตลอดภาคเรียน (long-term extended tasks) โดยทั่วไปจะเป็นงานที่มอบหมายให้ผู้เรียนตั้งแต่ต้นเทอมหรือเริ่มต้นหน่วยการเรียนรู้ โดยเป็นกิจกรรมหรือเหตุการณ์ที่ทำทลายความสามารถของผู้เรียน ส่วนใหญ่จะใช้ในการประเมินวิชาที่มีลักษณะเฉพาะ เช่น มอบหมายให้วางแผนถ่ายทำวิดีโอสะท้อนวิถีชีวิตความเป็นอยู่ในชุมชน หรือออกแบบงานการจัดแสดงนิทรรศการ การทำโครงการ เป็นต้น

3.6 ข้อดีและข้อจำกัดของการประเมินการปฏิบัติ

ข้อดีและข้อจำกัดของการประเมินการปฏิบัติสรุปได้ดังนี้ (Pearson Education Development Group, 2001; สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ, 2542)

ข้อดีของการประเมินการปฏิบัติ

1. การประเมินการปฏิบัติมุ่งเน้นให้ผู้เรียนนำความรู้ไปใช้ในสภาพจริง สามารถวัดได้ทั้งกระบวนการและผลงานของผู้เรียน
2. การประเมินการปฏิบัติช่วยเชื่อมโยงการเรียนรู้ในห้องกับการดำเนินชีวิตจริง ทำให้การเรียนรู้ของผู้เรียนมีความหมายยิ่งขึ้นและสามารถนำไปประยุกต์ใช้ในชีวิตจริง
3. ประเมินได้ทั้งประเมินระหว่างเรียนและประเมินสรุปผลการเรียนแต่โดยลักษณะของการประเมินการปฏิบัติตามสภาพจริงจะเน้นประเมินเพื่อให้เกิดการปรับปรุงหรือพัฒนา

ข้อจำกัดของการประเมินการปฏิบัติ

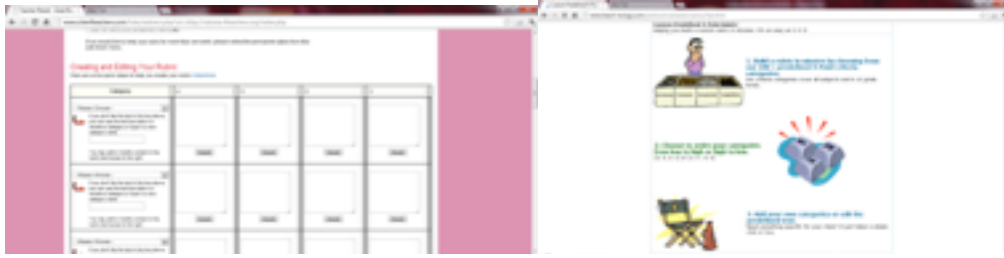
1. การประเมินการปฏิบัติตามสภาพจริงต้องอาศัยทักษะการสังเกตเป็นหลัก ดังนั้นผู้ประเมินจึงต้องมีความสามารถในการใช้ทักษะการสังเกตเป็นอย่างดีและต้องมีการบันทึกที่เป็นระบบ
2. การออกแบบงานที่ให้มีลักษณะการบูรณาการความรู้หลายสาขากับทักษะเป็นเรื่องที่ยากพอสมควรสำหรับครู
3. ธรรมชาติของการประเมินค่อนข้างยุ่งยากและไม่เป็นปรนัย

เกณฑ์การให้คะแนนแบบรูบรีค scoring rubric

Oxford English Dictionary ระบุว่า รูบรีคหมายถึงหัวเรื่องของหนังสือที่มีความแตกต่างกันในแต่ละตอนรากศัพท์มาจาก ruber ในภาษาละตินที่หมายถึงสีแดง การใช้คำว่า ruber หรือ rubric เกิดขึ้นในกลางศตวรรษที่ 15 โดยนักบวชในศาสนาคริสต์ผู้ซึ่งมีความอุตสาหะอย่างยิ่งในการทำสำเนาวรรณกรรมได้เขียนอักษรเริ่มต้นในแต่ละตอนของฉบับคัดลอกด้วยตัวหนังสือสีแดงขนาดใหญ่ตั้งแต่นั้นเป็นต้นมา rubric จึงหมายถึงหัวเรื่องในหนังสือจนกระทั่งสองศตวรรษที่ผ่านมา รูบรีคจึงมีความหมายใหม่ท่ามกลางศาสตร์ทางการศึกษานักการศึกษา ผู้เชี่ยวชาญทางการวัดผลได้ให้คะแนนการเขียนเรียงความของผู้เรียนโดยระบุและอธิบายข้อปฏิบัติซึ่งเป็นการชี้แนะการให้คะแนนหรือที่รู้จักกันในเวลาต่อมาในชื่อของ “รูบรีค” (scoring guide) (Popham, 1997)

เมื่อก้าวถึงความหมายของรูบรีค มีสองมุมมองด้วยกันมุมมองแรกมองว่ารูบรีคเป็นเครื่องมือประเมินตามสภาพจริงเช่นเดียวกับแฟ้มสะสมผลงาน ส่วนอีกมุมมองหนึ่งมองว่า รูบรีค คือชุดของเกณฑ์การให้คะแนนจนกระทั่งมีการทำวิจัยสำรวจกันเป็นเรื่องเป็นราว โดย Wenzlaff, Fager & Coleman (1999) ในปี ค.ศ. 1999 ทำการวิจัยเชิงสำรวจความเข้าใจเกี่ยวกับรูบรีคกับครูระดับมัธยมศึกษาประถมศึกษาและนักศึกษาฝึกครูจากมหาวิทยาลัยในภาคตะวันตกของสหรัฐอเมริกา จำนวน 100 คน พบว่า ส่วนใหญ่เข้าใจว่ารูบรีคคือเครื่องมือการประเมินที่แสดงส่วนประกอบงานโดยละเอียดสอดคล้องกับผลการสังเกตที่เอกสารพบว่ารูบรีคคือชุดของเกณฑ์ที่ช่วยให้คะแนนได้ง่ายขึ้นออกแบบไว้สำหรับการประเมินรูบรีคจึงเปรียบเสมือนวิธีการประเมินแต่ในทางปฏิบัติวิธีการประเมินไม่จำเป็นต้องใช้รูบรีคเสมอไป

ความพยายามของนักวิชาการในการพัฒนารูบรีคเพื่อใช้เป็นเกณฑ์ในการให้คะแนน ทำให้ปัจจุบันวงการศึกษาระดับนานาชาติพัฒนารูบรีคในรูปแบบ Rubric Bank Rubric library เป็นแหล่งรวมรูบรีคนานาชาติหลากหลายวิชาใช้ในการประเมินงานที่ครูสามารถนำมาใช้ได้โดยง่ายและมีการพัฒนาโปรแกรมสำเร็จรูปเพื่อสร้างรูบรีคเพื่อช่วยให้ครูสร้างรูบรีคได้ง่ายขึ้นและมีภาพประกอบสวยงาม เช่น Rubistar Rubric Generator (<http://rubistar.4teachers.org/>) หรือ Teacher Rubric Maker (http://www.teach-nology.com/web_tools/rubrics/)



Rubistar Rubric Generator

Rubric Maker

เทคนิคการสร้างรูบริค

1. ควรเริ่มจากการแจกแจงประเด็นที่จะพิจารณาให้คะแนน หรือบางที่เรียกว่ามิติที่จะประเมิน ซึ่งแต่ละมิติอาจจะให้น้ำหนักคะแนนเท่ากันหรือไม่ก็ได้เช่นการประเมินการทำรายงานของผู้เรียน อาจแบ่งเป็น 3 มิติ ดังนี้

มิติที่ 1 เนื้อหา	ให้น้ำหนักคะแนน 6
มิติที่ 2 การอ้างอิง	ให้น้ำหนักคะแนน 2
มิติที่ 3 รูปลักษณ์ของรายงาน	ให้น้ำหนักคะแนน 2

ข้อสังเกตประการแรกการให้น้ำหนักคะแนนในส่วนใดมากน้อยขึ้นกับดุลยพินิจของคุณ แต่รวมแล้วไม่เกินน้ำหนักที่กำหนดไว้ตั้งแต่แรก ในที่นี้คือ 10

ข้อสังเกตอีกประการที่สอง การแบ่งมิตินั้นไม่ควรกว้างหรือแคบจนเกินไป เช่นเนื้อหา หมายความว่ารวมถึง เนื้อหาและการลำดับเนื้อหาหรือการอ้างอิงหากกำหนดมิติว่า จำนวนรายการอ้างอิงอาจจะแคบเกินไป ควรกำหนดว่าการอ้างอิง หมายความว่ารวมถึง จำนวน และความทันสมัย ถูกต้องเพื่อที่จะได้เขียนระดับลดหลั่นลงได้อีกในขั้นตอนต่อไป

2. แต่ละมิติให้เขียนบรรยายความตั้งแต่สิ่งที่ดีที่สุดที่คาดหวังว่าผู้เรียนจักทำได้ จนไปถึงสิ่งที่แย่ที่สุด ทั้งนี้แล้วแต่ดุลยพินิจของคุณแต่ควรเขียนบรรยายลักษณะงานที่ให้คะแนนระดับสูงมีความท้าทายและระดับต่ำเป็นลักษณะงานพื้น ๆ ที่คนธรรมดาสามัญทำได้เริ่มต้นควรเขียนบรรยายสัก 3 ระดับแล้วเมื่อใช้รูบริคแล้ว จะทราบว่า การแบ่งระดับได้เหมาะสมหรือไม่ อาจเพิ่มเป็น 5 หรือ 7 ระดับ เช่น

มิติที่ 1 เนื้อหา มี 3 ระดับคะแนน

ระดับ 3 หมายถึง เนื้อหาครบถ้วนสมบูรณ์ มีการลำดับเนื้อหาต่อเนื่อง

ระดับ 2 หมายถึง เนื้อหาครบถ้วนสมบูรณ์ มีการลำดับเนื้อหาไม่ต่อเนื่องนัก

ระดับ 1 หมายถึง เนื้อหายังไม่ครบถ้วน มีการลำดับเนื้อหาไม่ต่อเนื่อง

ข้อสังเกตประการที่สามการบรรยายงานในระดับต่ำสุดก็อยู่ที่ดุลยพินิจของคุณว่าจะเริ่มต้นแบบไหน บางท่านสอนผู้เรียนเก่งอาจจะเริ่มนำลำดับ 2 ลดมาเป็น 1 ก็ได้ แล้วค่อยเขียนลำดับ 3 ให้สูงขึ้น

3. รูบริคอาจมีเกณฑ์การให้คะแนนทั้งเชิงปริมาณและหรือเชิงคุณภาพ เช่น แหล่งอ้างอิงในรายงาน อาจจะเขียนว่าเป็นเชิงปริมาณอย่างน้อย 10 แหล่งหรืออาจเขียนเกณฑ์เชิงคุณภาพประกอบด้วย 10 แหล่ง ที่ทันสมัย เชื่อถือได้ เป็นต้น เช่น

มิติที่ 2 การอ้างอิง (ตัวอย่างที่ 1)

ระดับ 3 หมายถึง แหล่งอ้างอิง มากกว่า 5 แหล่ง ทันสมัย และมีการเขียนอ้างอิงที่ถูกต้อง

ระดับ 2 หมายถึง แหล่งอ้างอิง 3-5 แหล่ง ทันสมัย และมีการเขียนอ้างอิงที่ถูกต้อง

ระดับ 1 หมายถึง แหล่งอ้างอิง ต่ำกว่า 3 แหล่ง ทันสมัย และมีการเขียนอ้างอิงที่ถูกต้องหรือตัวอย่างที่ 2

ระดับ 3 หมายถึง แหล่งอ้างอิงมากกว่า 5 แหล่ง การเขียนอ้างอิงถูกต้อง ทันสมัย

ระดับ 2 หมายถึง แหล่งอ้างอิงมากกว่า 5 แหล่ง การเขียนอ้างอิงถูกต้อง แต่ยังไม่ทันสมัยเท่าที่ควร

ระดับ 1 หมายถึง แหล่งอ้างอิงมากกว่า 5 แหล่ง การเขียนอ้างอิงยังไม่ถูกต้องนัก และยังไม่ทันสมัยเท่าที่ควร

ข้อสังเกตประการที่สี่ การเขียนบรรยายระดับคะแนนอยู่ที่จุดเน้นหรือผลที่คาดหวังว่าจะเกิดกับผู้เรียนเป็นลำดับหากต้องการให้เขาอ้างอิงให้ถูก และทันสมัยก็ใช้ตัวอย่างแรก เป็นต้น

4. บรรจุคำบรรยายเข้าตาราง สิ่งสำคัญควรให้ผู้เรียนรับรู้เกี่ยวกับเกณฑ์การให้คะแนน หรือหากเป็นไปได้อาจจะให้ผู้เรียนช่วยในกระบวนการสร้าง เมื่อผู้เรียนได้อ่านและเข้าใจเป้าหมายของงานผู้เรียนจะเติบโตเป็นเยาวชนและผู้ใหญ่ที่เป้าหมายในชีวิตและรู้จักการประเมินตนเองอยู่เสมอซึ่งเป็นเป้าหมายสุดท้ายของการประเมินสภาพจริง

เกณฑ์การให้คะแนนแบบรูปรีคเพื่อประเมินการทำรายงานของผู้เรียน

การทำรายงาน	ระดับ 3	ระดับ 2	ระดับ 1	คะแนน
เนื้อหาของรายงาน -ความสมบูรณ์ครบถ้วน -การลำดับเนื้อหา	เนื้อหาครบถ้วนสมบูรณ์ มีการลำดับเนื้อหาต่อเนื่อง	เนื้อหาครบถ้วนสมบูรณ์ มีการลำดับเนื้อหาไม่ต่อเนื่องนัก	เนื้อหายังไม่ครบถ้วน มีการลำดับเนื้อหาไม่ต่อเนื่อง	
การอ้างอิง -จำนวน -ความทันสมัย -ความถูกต้อง	แหล่งอ้างอิง มากกว่า 5 แหล่ง ทันสมัย และมีการเขียนอ้างอิงถูกต้อง	แหล่งอ้างอิง 3-5 แหล่ง ทันสมัย และมีการเขียนอ้างอิงถูกต้อง	แหล่งอ้างอิง ต่ำกว่า 3 แหล่ง ทันสมัย และมีการเขียนอ้างอิงถูกต้อง	
รูปลักษณะของรายงาน -เรียบร้อย -สวยงาม	จัดทำรูปเล่มเรียบร้อย ปรารถนิต สวยงาม	จัดทำรูปเล่มเรียบร้อย แต่ยังไม่ปรารถนิตสวยงามเท่าที่ควร	จัดทำรูปเล่ม แต่ไม่ปรารถนิต สวยงาม	

อนึ่งการให้คะแนนอาจพิจารณาถึงน้ำหนักความสำคัญของประเด็นการประเมิน เช่น เนื้อของรายงานอาจให้น้ำหนักเป็น 6 จาก 10 และการอ้างอิงอาจและรูปลักษณะรายงานให้น้ำหนักประเด็นละ 2 จาก 10 จากเกณฑ์การให้คะแนนแบบรูปรีคเพื่อประเมินการทำรายงานของผู้เรียนจะปรับคะแนนจากคะแนนเต็ม 9 คะแนน เป็น 30 คะแนน แบ่งเป็นเนื้อหาของรายงาน 18คะแนน (3x6) การอ้างอิง 6คะแนน (3x2)และรูปลักษณะของรายงาน 6 คะแนน (3x2)

การทำรายงาน	น้ำหนัก	ระดับ 3	ระดับ 2	ระดับ 1	คะแนน
เนื้อหาของรายงาน -ความสมบูรณ์ครบถ้วน -การลำดับเนื้อหา	6	เนื้อหาครบถ้วนสมบูรณ์มีการลำดับเนื้อหาต่อเนื่อง	เนื้อหาครบถ้วนสมบูรณ์ มีการลำดับเนื้อหาไม่ต่อเนื่องนัก	เนื้อหายังไม่ครบถ้วน มีการลำดับเนื้อหาไม่ต่อเนื่อง	

การอ้างอิง -จำนวน -ความทันสมัย -ความถูกต้อง	2	แหล่งอ้างอิงมากกว่า 5 แหล่งทันสมัย และมีการเขียนอ้างอิงที่ถูกต้อง	แหล่งอ้างอิง 3-5 แหล่งทันสมัย และมีการเขียนอ้างอิงที่ถูกต้อง	แหล่งอ้างอิงต่ำกว่า 3 แหล่งทันสมัย และมีการเขียนอ้างอิงที่ถูกต้อง	
รูปลักษณะของรายงาน -เรียบร้อย -สวยงาม	2	จัดทำรูปแบบเรียบร้อย ปรารถนิตสวยงาม	จัดทำรูปแบบเรียบร้อย แต่ยังไม่ปรารถนิตสวยงามเท่าที่ควร	จัดทำรูปแบบ แต่ไม่ปรารถนิตสวยงาม	

เกณฑ์การประเมินผล

ได้คะแนน 10-15 คะแนน หมายถึง พอใช้

ได้คะแนน 16-20 คะแนน หมายถึง ดี

ได้คะแนน 21-30 คะแนน หมายถึง ดีเยี่ยม

ถึงแม้ว่าแนวคิดด้านการประเมินการเรียนรู้ตามสภาพจริงได้เริ่มเข้ามาเผยแพร่ในประเทศไทยตั้งแต่ปี พ.ศ. 2535 โดย Jonh Carr ผู้เชี่ยวชาญด้านการศึกษา ที่ภาควิชาประเมินผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยเชียงใหม่ จากนั้นได้มีการนำมาทดลองใช้ที่จังหวัดกระบี่ในปี พ.ศ. 2538 และที่จังหวัดราชบุรีเมื่อปลายปี พ.ศ. 2539 นอกจากนี้ในปี พ.ศ. 2541 กระทรวงศึกษาธิการได้นำไปใช้ในการพัฒนาการศึกษาทั้งในส่วนของนักเรียน ครู ผู้บริหาร และศึกษานิเทศก์ แต่การขยายผลเชิงปฏิบัตินับว่ายังเป็นไปได้ไม่มากนัก ทั้งนี้มีสาเหตุมาจากครูยังไม่เข้าใจขาดทักษะ ไม่มั่นใจว่าจะวัดความรู้ได้เท่าเทียมกับการทดสอบ นอกจากนั้นครูยังยึดติดอยู่กับเนื้อหาและเวลาตามหลักสูตรมากกว่ามองที่จุดหมายของหลักสูตรจึงทำให้ครูมุ่งสอนเนื้อหาให้ทันตามเวลาอีกประเด็นที่น่าสนใจที่เป็นภาวะคุกคาม คือกระบวนการคัดเลือกเข้าศึกษาในระดับต่างๆ ยังนิยมใช้คะแนนสอบเป็นเกณฑ์ตัดสิน ดังนั้นจึงทำให้ครูและผู้ปกครองไม่มั่นใจว่าหากใช้การประเมินการเรียนรู้ตามสภาพจริงนักเรียนจะสอบแข่งขันกับคนอื่นได้ สิ่งเหล่านี้นับเป็นอุปสรรคต่อการปรับกระบวนการประเมินของครูเป็นอย่างมาก

เอกสารอ้างอิง

คณะกรรมการการประถมศึกษาแห่งชาติ, สำนักงาน. (2542). **การประเมินตามสภาพที่แท้จริง**.

กรุงเทพฯ: โรงพิมพ์คุรุสภา.

ส.วาสนา ประวาลพฤกษ์. (2537). **การวัดผลจากการปฏิบัติจริง (Authentic Performance Measurement)** เอกสารประกอบการประชุมสัมมนาวิชาการวัดผลและวิจัยสัมพันธ์

ครั้งที่ 4 วันที่ 1-2 ธันวาคม 2537 มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

อุทุมพร (ทองอุไทย) จามรมาน. (2540). **การตีค่าความสามารถที่แท้จริงของผู้เรียนเพื่อการปฏิรูปการศึกษา (Authentic performance and portfolio assessment of learners for educational reform)**. พันนี้พับบลิชชิง: กรุงเทพมหานคร.

Airasian, P. W. (1997). **Classroom assessment**. 3rd ed. NY: The McGraw-Hill.

Brualdi, A. (1999). Performance-based assessment : how students understand and apply knowledge. **Schools in the middle**. 9(4) :pp. 5-22 [Available online at: <http://cdnet2.car.chula.ac.th>]

Brualdi,A.(1998). Implementating performance assessment in classroom. **Practical Assessment,Research&Evaluation**,6(2).

[Available online at: <http://ericae.net/pare/getvn.asp?v=6&n=2.>]

Chiseri-Strater, E. (1992). College sophomores reopen the closed portfolio. In D. H.

- Grave & Sunstein, B. S. (eds). **Portfolio Portraits**. Portsmouth, NH: Heinemann pp. 61–72.
- Esiner, E. W. (1999). The Use and limits of performance assessment. **Kappa Professional Journal**. Phi Delta Kappa International [Available online at: <http://pdkintl.org/kappa/kei9905.htm>]
- Hart, D. (1994) **Authentic assessment: A handbook for educators**. Menlo Park, California: Addison–Wesley Publishing Company.
- Herman, J., Aschbacher, P. & Winters, L. (1992). **A practical guide to alternative assessment**. Alexandria, VA: Association for Supervision and Curriculum Development.
- Kane, M. B. & Khattri, N. (1995) Assessment reform. **Phi Delta Kappa** 77(1) pp. 30–32.
- Lester, F. K., Lambdin, D. V., & Preston, R. V. A new vision of the nature and purpose of assessment in the mathematics classroom. In G. D. Phye (ed.), **Handbook of classroom assessment: learning, achievement, and adjustment**. pp. 287–319. San Diego: Academic press.
- Pearson Education Development Group (2001). **Alternative assessment**. [Available online at: <http://www.teachervision.com>]
- Pearson Education Development Group. (2001). **Creating rubrics** [Available online at: <http://www.teachervision.com>]
- Pike, K., & Salend, S.J. (1995) Authentic assessment strategies: Alternative to Popham, J. W. (1997). What's wrong and what right with rubrics. **Educational Leadership**, 55(2), 72–75.
- Stecher, B. M., & Herman, J. L. (1997). Using portfolios for large–Scale assessment. In G. D. Phye (ed.), **Handbook of classroom assessment: learning, achievement, and adjustment**. pp. 287–319. San Diego: Academic press.
- Wiggins, G. (1989). A true test : Toward more authentic and equitable assessment. **Phi Delta Kappa**. 70(9): 703–713.
- Wenzlaff, F., Fager, J. J., & Coleman. (1999). What is a rubric ? Do practitioner and literature agree. **Contemporary Education**. 70(4) : 41–46.
- Worthen, B. (1993). Critical issues that will determine the future of alternative Assessment. **Phi Delta Kappa**. 75(3): 200–214.
- norm–referenced testing. **Teaching exceptional children**, 28, pp. 15–19.

กิจกรรมการออกแบบการประเมิน

1. ขอให้ท่านออกแบบชิ้นงานเพื่อประเมินการเรียนรู้ตามสภาพจริง ในรายวิชาที่รับผิดชอบ ให้ครอบคลุมประเด็นการประเมินตามมาตรฐานการเรียนรู้ (TQF) อย่างน้อย 3 มาตรฐาน

รายวิชา.....

มาตรฐานการเรียนรู้ที่ประเมิน.....

ชื่อชิ้นงาน.....

รายละเอียดชิ้นงาน.....

2. จากชิ้นงานดังกล่าวให้ท่านสร้างเกณฑ์การให้คะแนนแบบรูบรีค ชนิดแยกองค์ประกอบ ขนาด 3x3 และให้นำหนักคะแนนตามที่เหมาะสม และให้เหตุผลการแบ่งสัดส่วนของน้ำหนักคะแนนประกอบ

3. จากรูบรีคดังกล่าว ให้ท่านกำหนดเกณฑ์การประเมินภาพรวมโดยแบ่งระดับการประเมินเป็น 3-5 ระดับ

4. ส่งตัวแทนนำเสนอ