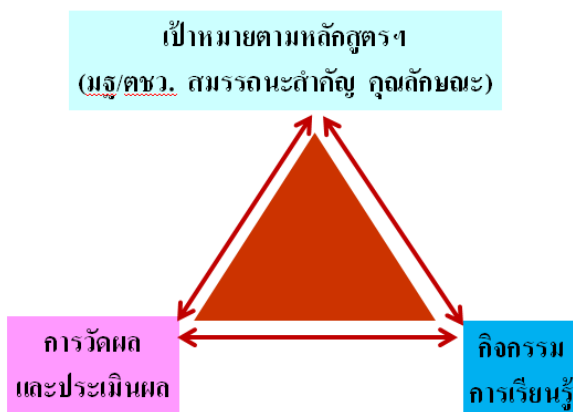


## การวัดประเมินผลที่สอดคล้องกับมาตรฐานการเรียนรู้

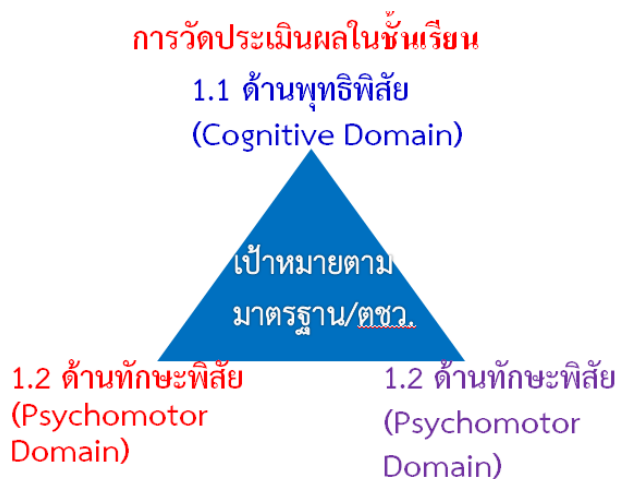
### 1 เป้าหมายการวัดประเมินผลเพื่อให้ได้สารสนเทศไปใช้

- พัฒนาผู้เรียน
- พัฒนาครู
- พัฒนาหลักสูตร

### 2 องค์ประกอบการจัดการศึกษา



3. การวัดประเมินผลที่สำคัญคือวัดและประเมินเพื่อพัฒนาสิ่งที่เกิดในระหว่างเรียน สอดคล้องกับกิจกรรมการเรียนรู้ที่ควรจัดแบบ Active Learning มาตรฐานและตัวชี้วัดในหลักสูตร ดังภาพ



ฉะนั้นการออกแบบการเรียนรู้รวมถึงการวัดประเมินผลจึงต้องเริ่มจากการคลี่ค่าสำคัญในมาตรฐาน/ตัวชี้วัด และควรคิดเชิงระบบต่อเนื่อง

เป้าหมาย								
ตัวชี้วัด / ผลการเรียนรู้	ระดับ ความรู้ ตามตัวชี้วัด	ทักษะ / กระบวนการ	ทำให้เกิด		คำสำคัญ	ชิ้นงาน/ ภาระงาน	กิจกรรมการเรียนรู้ - เน้นการคิดวิเคราะห์ Active learning - ศตวรรษที่ ๒๑ สอดคล้องกับยุค Thailand ๔.๐	การวัด ประเมินผล
			สมรรถนะ สำคัญ	คุณลักษณะ				
								เครื่องมือวัด วิธีวัด แนวทางการ ให้คะแนนและ เกณฑ์เกณฑ์ การวัด

## ความรู้ที่จำเป็นต้องรู้

### จุดประสงค์ของการวัดผล

1. **สอบวัดเพื่อจัดตำแหน่ง (Placement)** ใช้การสอบวัดเพื่อให้ได้คำตอบว่า ณ วันเวลาที่สอบนั้น เด็กแต่ละคนมีความรู้ความสามารถ อยู่ตรงระดับไหนของกลุ่ม เพื่อช่วยในการจัดประเภทหรือคัดแยก หรือบรรจุเขาให้ถูกตำแหน่ง
2. **สอบวัดเพื่อวินิจฉัย (Diagnosis)** ใช้การสอบวัดเพื่อตรวจสอบผู้เรียนว่าเก่งหรืออ่อนวิชาใด เนื้อหาใด เพื่อครูจะได้ช่วยเหลือได้ถูกจุด
3. **สอบวัดเพื่อเปรียบเทียบ (Assessment)** ใช้การสอบวัดเพื่อเปรียบเทียบผู้เรียนว่าได้มีการพัฒนาเพิ่มขึ้นหลังจากผ่านการเรียนการสอนไปแล้วมากน้อยเพียงใด
4. **สอบวัดเพื่อพยากรณ์ (Prediction)** ใช้การสอบวัดเพื่อพยากรณ์ว่า เด็กคนนั้นคนนั้นสามารถเรียนต่ออะไรได้จึงจะเรียนจบ ไม่ถูกออกกลางคัน หรือประกอบอาชีพอะไรจึงจะประสบความสำเร็จ
5. **สอบวัดเพื่อประเมินค่า (Evaluation)** ใช้การสอบวัดเพื่อเอาผลที่ได้ไปประเมินหรือสรุปรวมอย่างมีหลักเกณฑ์ว่าหลักสูตร หรือวิธีการสอนของครู สามารถช่วยให้ผู้เรียนเกิดการเรียนรู้ได้มากน้อยเพียงใด นอกจากนี้ ครูควรใช้การสอบวัดเป็นส่วนหนึ่งของการเรียนการสอน ต้องถือว่าการสอบเป็นกิจกรรมหนึ่งของนักเรียน เพื่อทราบศักยภาพการเรียนรู้ของผู้เรียน กลมกลืนไปกับการเรียนการสอน การสอบวัดไม่จำเป็นต้องใช้ข้อสอบเพียงอย่างเดียว อาจจะใช้คำถาม (การสอบปากเปล่า) กับผู้เรียน และถ้าจะให้การเรียนการสอนได้ผลดียิ่งขึ้น ครูควรชี้แจงการตอบของผู้เรียนด้วยว่า ตอบถูกเพราะอะไร ตอบผิดเพราะอะไร จะทำให้ชั่วโมงการสอนมีประสิทธิภาพมากขึ้น

### ความหมายการวัดผล

**การวัดผล (Measurement)** คือการกำหนดตัวเลขให้กับวัตถุ สิ่งของ เหตุการณ์ ปรากฏการณ์ หรือพฤติกรรมต่าง ๆ หรืออาจใช้เครื่องมือไปวัดเพื่อให้ได้ตัวเลขแทนคุณลักษณะต่าง ๆ เช่น ใช้ไม้บรรทัดวัดความกว้างของหนังสือได้ 3.5 นิ้ว ใช้เครื่องชั่งน้ำหนักของเนื้อหมูได้ 0.5 กิโลกรัม ใช้แบบทดสอบวัดความรู้ในวิชาภาษาไทยของเด็กชายแดง ได้ 42 คะแนน เป็นต้น

### การวัดผลแบ่งเป็น 2 ประเภทคือ

1. **วัดทางตรง** วัดคุณลักษณะที่ต้องการโดยตรง เช่น ส่วนสูง น้ำหนัก ฯลฯ มาตรการวัดจะอยู่ในระดับ Ratio Scale
2. **วัดทางอ้อม** วัดคุณลักษณะที่ต้องการโดยตรงไม่ได้ ต้องวัดโดยผ่านกระบวนการทางสมอง เช่น วัดความรู้ วัดเจตคติ วัดบุคลิกภาพ ฯลฯ มาตรการวัดจะอยู่ในระดับ Interval Scale  
การวัดทางอ้อมแบ่งออกเป็น 3 ด้านคือ
  - 2.1 **ด้านสติปัญญา (Cognitive Domain)** เช่น วัดผลสัมฤทธิ์ทางการเรียน วัดเชาวน์ปัญญา วัดความถนัดทางการเรียน วัดความคิดสร้างสรรค์ ฯลฯ
  - 2.2 **ด้านความรู้สึก (Affective Domain)** เช่น วัดความสนใจ วัดเจตคติ วัดบุคลิกภาพ วัดความวิตกกังวล วัดจริยธรรม ฯลฯ
  - 2.3 **ด้านทักษะกลไก (Psychomotor Domain)** เช่น การเคลื่อนไหว การปฏิบัติโดยใช้เครื่องมือ ฯลฯ

### ความหมายการประเมินผล

**การประเมินผล** (Evaluation) หมายถึง การนำเอาข้อมูลต่าง ๆ ที่ได้จากการวัดรวมกับการใช้วิจารณ์ญาณของผู้ประเมินมาใช้ในการตัดสินใจ โดยการเปรียบเทียบกับเกณฑ์ เพื่อให้ได้ผลเป็นอย่างไรอย่างหนึ่ง เช่น เนื้อหาม้วนนี้หนัก 0.5 กิโลกรัมเป็นเนื้อหาม้วนที่เบาที่สุดในร้าน (เปรียบเทียบกับภายในกลุ่ม) เด็กชายแดงได้คะแนนวิชาภาษาไทย 42 คะแนนซึ่งไม่ถึง 50 คะแนนถือว่าสอบไม่ผ่าน (ใช้เกณฑ์ที่ครูสร้างขึ้น) เป็นต้น

### การประเมินผลแบ่งได้เป็น 2 ประเภท การประเมินแบบอิงกลุ่มและการประเมินแบบอิงเกณฑ์

1. การประเมินแบบอิงกลุ่ม เป็นการเปรียบเทียบคะแนนจากแบบทดสอบหรือผลงานของบุคคลใดบุคคลหนึ่งกับบุคคลอื่น ๆ ที่ได้ทำแบบทดสอบเดียวกันหรือได้ทำงานอย่างเดียวกัน นั่นคือเป็นการใช้เพื่อจำแนกหรือจัดลำดับบุคคลในกลุ่ม การประเมินแบบนี้มักใช้กับการ การประเมินเพื่อคัดเลือกเข้าศึกษาต่อ หรือการสอบชิงทุนต่าง ๆ
2. การประเมินแบบอิงเกณฑ์ เป็นการเปรียบเทียบคะแนนจากแบบทดสอบหรือผลงานของบุคคลใดบุคคลหนึ่งกับเกณฑ์หรือจุดมุ่งหมายที่ได้กำหนดไว้ เช่น การประเมินระหว่างการเรียนรู้ว่าผู้เรียนได้บรรลุวัตถุประสงค์การเรียนรู้ที่ได้กำหนดไว้หรือไม่

### ข้อแตกต่างระหว่างการประเมินผลแบบอิงกลุ่มและอิงเกณฑ์

#### การประเมินผลแบบอิงกลุ่ม

1. เป็นการเปรียบเทียบคะแนนที่ได้กับคะแนนของคนอื่น ๆ
2. นิยมใช้ในการสอบแข่งขัน
3. คะแนนจะถูกนำเสนอในรูปของร้อยละหรือคะแนนมาตรฐาน
4. ใช้แบบทดสอบเดียวกันสำหรับผู้เรียนทั้งกลุ่มหรืออาจใช้แบบทดสอบคู่ขนาน เพื่อให้สามารถเปรียบเทียบกันได้
5. แบบทดสอบมีความยากง่ายพอเหมาะ มีอำนาจจำแนกสูง
6. เน้นความเที่ยงตรงทุกชนิด

#### การประเมินแบบอิงเกณฑ์

1. เป็นการเปรียบเทียบคะแนนที่ได้กับเกณฑ์ที่ได้กำหนดไว้
2. สำหรับการเรียนการสอนเพื่อพัฒนาผู้เรียนหรือเพื่อปรับปรุงการเรียนการสอน
3. คะแนนจะถูกนำเสนอในรูปของผ่าน-ไม่ผ่านตามเกณฑ์ที่กำหนดไว้
4. ไม่ได้เปรียบเทียบกับคนอื่น ๆ จึงไม่จำเป็นต้องใช้แบบทดสอบฉบับเดียวกันกับผู้เรียนทั้งชั้น
5. ไม่เน้นความยากง่าย แต่อำนาจจำแนกควรมีพอเหมาะ
6. เน้นความเที่ยงตรงเชิงเนื้อหา

## การประเมินการปฏิบัติ (Performance Assessment)

ความหมายของการประเมินการปฏิบัติ  
(Performance Test)

การสอบความสามารถในการปฏิบัติงาน ทั้งด้านการดำเนินงาน และผลงานตามวัตถุประสงค์ของงาน หรือสถานการณ์ที่กำหนด ทั้งนี้ ความสามารถในการปฏิบัติงาน เป็นผลรวมของความสามารถด้าน พุทธิพิสัย จิตพิสัย และทักษะพิสัยดังกล่าว

องค์ประกอบการทดสอบภาคปฏิบัติ

- ผู้ปฏิบัติงานหรือผู้เข้าสอบ
- งานที่กำหนดให้ปฏิบัติ

ผู้วัดหรือประเมิน

- กำหนดพฤติกรรม (Behavior)
- กำหนดเงื่อนไข (Conditions)
- กำหนดเกณฑ์หรือระดับที่จะยอมรับ (Criteria)

### ประเภทของการประเมินการปฏิบัติ

1. การทดสอบเชิงจำแนก (Identification test)
2. การทดสอบด้วยตัวอย่างงาน (Work – Sample test)
3. การทดสอบโดยใช้สถานการณ์ (Simulation test)
4. การทดสอบด้วยวิธีอื่น (Alternative test)

#### 1. การทดสอบเชิงจำแนก

การสร้างแบบทดสอบประเภทนี้มีดังนี้

นิยามสิ่งที่จะถูกจำแนก

ประเมินความเป็นไปได้ของการสอบในกรอบงานที่เลือก

- เขียนกรอบหรือโครงสร้างของแบบทดสอบ
- หาวัสดุอุปกรณ์ที่จะต้องใช้ในการทดสอบ

การสร้างแบบประเมิน

- เขียนคำชี้แจงและเกณฑ์การให้คะแนน

#### 1.1 นิยามสิ่งที่จะถูกจำแนก

##### ตัวอย่าง

##### การทดสอบปฏิบัติของการเรียนซ่อมคอมพิวเตอร์

ผู้สอนต้องสามารถวิเคราะห์ให้ได้ว่าเครื่องคอมพิวเตอร์มีอาการผิดปกติหรือไม่ หากผิดปกติเกิดจากสาเหตุใด แบบทดสอบเชิงจำแนกในกรณีนี้ ต้องพิจารณาเงื่อนไขอุปกรณ์ต่าง ๆ ภายในเครื่องคอมพิวเตอร์ด้วย และหน้าที่การทำงาน รวมทั้งอาการที่เครื่องจะแสดงให้เห็นเมื่อมีอาการผิดปกติ

## 1.2 ประเมินความเป็นไปได้ของการสอบ ในกรอบงานที่เลือก

### ตัวอย่าง

การทดสอบการซ่อมคอมพิวเตอร์ จะต้องกำหนดว่าควรให้ส่วนประกอบใดที่ทำงานผิดปกติ เช่น ฮาร์ดดิสก์ แรม ฯลฯ

## 1.3 เขียนกรอบหรือโครงสร้างของแบบทดสอบ

### ตัวอย่าง

การซ่อมคอมพิวเตอร์ ต้องพิจารณาว่าเมื่อเปิดเครื่องแล้วนักเรียนรู้หรือไม่ว่าเครื่องเสียตรงไหน ให้อธิบายในการพิจารณาเท่าไร และนักเรียนแก้ไขได้หรือไม่

เกณฑ์ในการให้คะแนนต้องเป็นแบบเดียวกัน  
การให้คะแนนโดยใช้แบบประเมินหรือการสังเกต  
พฤติกรรมการปฏิบัติอย่างไร  
การให้คะแนนแบบทดสอบเชิงจำแนก

### การทดสอบด้วยตัวอย่างงาน

กระบวนการประเมินผลงานหรือประเมินทั้งกระบวนการผลิตส่วนใหญ่ใช้ประเมินงานในหลายวงการ เช่นการแสดง การทดลอง การขับพาหนะ

## 2. การสร้างแบบทดสอบด้วยตัวอย่างงาน

กำหนดขอบข่ายของทักษะการปฏิบัติ

วิเคราะห์งาน

สร้างแบบประเมิน

สร้างคู่มือคำชี้แจง

กำหนดเงื่อนไขสถานการณ์

### การให้คะแนนแบบทดสอบด้วยตัวอย่างงาน

การให้คะแนนจำเป็นต้องมีมาตรฐานเดียวกัน เงื่อนไขสถานการณ์ วัสดุ อุปกรณ์ของผู้สอบแต่ละคน ควรต้องเหมือนกันทุกอย่าง

ผู้สอบจะได้รับคู่มือการชี้แจงที่เหมือนกัน ใช้เวลาเท่ากันและถูกสังเกตแบบเดียวกัน การสังเกตอาจใช้แบบตรวจสอบรายการ หรือมาตรฐานประเมินค่าเป็นแนวทางในการสังเกตพฤติกรรมหรือลำดับขั้นตอนในการทำงาน รวมถึงระดับของการปฏิบัติที่จะยอมรับต้องอธิบายวิธีการและเกณฑ์การปฏิบัติให้กับผู้สอบเพื่อให้ผู้สอบเข้าใจถึงผลหรือความคาดหวังที่จะเกิดขึ้น

### 3. การทดสอบโดยใช้สถานการณ์

เป็นการทดสอบซึ่งให้ผู้สอบปฏิบัติงานที่ใกล้เคียงกับการปฏิบัติจริงในขณะกำลังดำเนินงานมากที่สุด เพื่อวัดสมรรถภาพของผู้สอบเกี่ยวกับกระบวนการทางด้านเทคนิค ทักษะการจัดการ และพฤติกรรมกาปฏิบัติ

การสร้างแบบสอบโดยใช้สถานการณ์

เป็นการออกแบบเพื่อจำลองกระบวนการตัดสินใจ ผู้สอบต้องตัดสินใจและปฏิบัติการ สถานการณ์เริ่มด้วยการนำเสนอสถานการณ์ เสนอข้อมูลที่เกี่ยวข้องกับสถานการณ์และเขียนปัญหา เพื่อให้ผู้สอบทำงานจากปัญหาที่เขียนขึ้นทีละขั้นตอนตามลำดับ

การให้คะแนนแบบทดสอบโดยการใช้อย่างสถานการณ์

เป็นการประเมินความสามารถของผู้สอบเกี่ยวกับการจำลองสถานการณ์ ผู้สอบจะได้รับผลสะท้อนกลับโดยการเสนอแนะแนวทางแก้ไขการใช้สถานการณ์ สามารถใช้ได้ทั้งการสอนและการสอบ เพื่อนำไปใช้ใน ชีวิตประจำวันได้

### 4. การทดสอบด้วยวิธีอื่น ๆ

1. การทดสอบปากเปล่า
2. การประเมินโดยใช้ข้อสอบข้อเขียน
3. การให้เขียนรายงาน

#### ขั้นตอนการสร้างเครื่องมือการประเมินการปฏิบัติการ

- 1 กำหนดจุดประสงค์ให้ชัดเจน
- 2 กำหนดเกณฑ์การปฏิบัติโดยแยกแยะประเด็นที่จะสังเกต ตัดสินหรือให้คะแนนการปฏิบัติ

#### เครื่องมือประกอบการประเมินการปฏิบัติ

- แบบสังเกต
- แบบตรวจสอบรายการ
- แบบประเมินค่าหรือมาตรประเมินค่า
- แบบประเมินโดยการจัดอันดับ

#### 1. แบบสังเกต

ใช้ประกอบในการสังเกตกระบวนการปฏิบัติงาน และทักษะการปฏิบัติงานของผู้เข้าสอบ เครื่องมือชนิดนี้ใช้ร่วมกับการบันทึกเหตุการณ์หรือการตรวจสอบรายการ

เทคนิคนี้ควรจะได้เตรียมการดังนี้

1. สุ่มพฤติกรรมที่จะสังเกต
2. กำหนดขอบเขตของพฤติกรรมที่จะศึกษา
3. วางแผนการสังเกตว่าจะสังเกตใคร สังเกตเมื่อไร ใครเป็นผู้สังเกต
4. เลือกและฝึกผู้สังเกต
5. ผลของการสังเกตเป็นตัวอย่างของพฤติกรรมของบุคคลที่สังเกตได้ ซึ่งควรบอกได้ว่าผลดังกล่าวสามารถสรุปอ้างอิง ไปยังพฤติกรรมอะไรได้บ้าง

## 2. แบบตรวจสอบรายการ

- ผู้สังเกตบันทึกเมื่อมีเหตุการณ์เกิดขึ้น อย่างรวดเร็ว และมีประสิทธิภาพ
  - ผู้สังเกตจะไม่ประเมินคุณภาพ ระดับ ความถี่ของพฤติกรรมที่เกิด
  - แต่จะตรวจสอบว่าพฤติกรรมที่เกิดขึ้นควรใช้เครื่องมือใดในการตรวจสอบ
    - = มาตรฐานประเมินค่า
    - = การตรวจสอบรายการการปฏิบัติต่าง ๆ ตามรายการที่กำหนดหรือไม่
  - ประเมินกระบวนการที่แบ่งเป็นส่วนหรือพฤติกรรมย่อย ๆ ของพฤติกรรมใหญ่
  - คุณค่าของการตรวจสอบรายการขึ้นอยู่กับ
    - = ทักษะในการตรวจสอบ
    - = ความใส่ใจในการสร้างรายการที่จะตรวจสอบ
- เนื้อหาวิชา กิจกรรม กระบวนการ ผลงาน การปรับตัวในสังคม

ตัวอย่าง ประเมินความสามารถในการใช้เครื่องมือทดลองทางวิทยาศาสตร์

ชื่อผู้เรียน.....ชั้น.....

คำชี้แจง ทำเครื่องหมาย / ในช่องตามความเป็นจริง

## 3. แบบประเมินค่าหรือมาตรฐานประเมินค่า

- ใช้ประเมินทักษะวิธีการทำและผลงาน
- กำหนดเป็นระดับคุณภาพของการปฏิบัติ ยอดเยี่ยม ปานกลาง ต้องปรับปรุง
- กำหนดเป็นความถี่ของการแสดงพฤติกรรม เสมอ ๆ บางครั้ง ไม่เคย
- เขียนบรรยายระดับของการปฏิบัติที่ยอมรับได้
- ทำงานเสร็จอย่างรวดเร็ว
- ทำงานเสร็จสมบูรณ์ค่อนข้างช้า
- ไม่สามารถทำงานเสร็จได้ด้วยตัวเอง

## 4. แบบประเมินโดยการจัดอันดับ (Ranking)

- ให้อันดับคุณลักษณะ ความสามารถ ทักษะของผู้เรียน
  - มีความเที่ยงและความตรงสูง
- 4.1 วิเคราะห์ลักษณะของสิ่งที่จะจัดอันดับเพื่อกำหนดประเด็นที่จะประเมิน
- เช่นการออกแบบ
- ความเรียบร้อย พิถีพิถัน
  - ความคิดสร้างสรรค์
- 4.2 กำหนดเกณฑ์การประเมินในแต่ละประเด็นที่กำหนดขึ้น
- 4.3 การจัดอันดับ ผลของงาน ควรมีกรรมการมากกว่า 1 คน
- แบ่งผลงานออกเป็น กลุ่ม ๆ
- = กลุ่มสูงกว่าเกณฑ์ - ปานกลาง - ใกล้เคียง - ต่ำกว่าเกณฑ์
  - = จัดอันดับที่ดีที่สุดและเลวที่สุดของแต่ละกลุ่ม
  - = จัดอันดับของที่สองและรองสุดท้าย จนถึงพวกที่อยู่ตรงกลาง



### ขั้นตอนการประเมินการปฏิบัติ

1. กำหนดจุดมุ่งหมาย = ทั่วไป หรือ เฉพาะ (ของการสอบ)
2. กำหนดคุณลักษณะที่จะวัดตามจุดประสงค์ของการสอบ
3. เลือกชนิดของการทดสอบภาคปฏิบัติและเครื่องมือประกอบการวัด ให้สอดคล้องกับจุดมุ่งหมายที่กำหนดไว้
4. ให้ผู้เข้าสอบปฏิบัติงานตามชนิดของการสอบ
5. ผู้วัดหรือประเมิน สังเกต จดบันทึกและตรวจผลงาน
6. การให้คะแนนจากการทดสอบภาคปฏิบัติ

### วิธีการที่ใช้ในการประเมินการปฏิบัติ

การประเมินโดยผู้เชี่ยวชาญ (Supervisor Rating)

การประเมินโดยใช้เทคนิคการสังเกต (Observation Technique)

การประเมินตนเอง (Self Assessment)

### การประเมินโดยผู้เชี่ยวชาญ

ผู้เชี่ยวชาญประเมินทักษะและพฤติกรรมที่เกี่ยวกับการปฏิบัติงานของผู้สอบ ซึ่งแสดงออกในขณะปฏิบัติงาน ข้อดี-ข้อเสียของการประเมินโดยผู้เชี่ยวชาญ

ข้อดีคือเป็นการวัดทักษะและพฤติกรรมโดยตรงของผู้สอบ โดยทั่วไปมักจะเป็นทักษะและพฤติกรรมที่สำคัญในการปฏิบัติงาน ผู้ประเมินอาจให้ข้อเสนอแนะเพื่อประโยชน์ในการปรับปรุงงาน

ข้อเสียของการประเมินแบบนี้คือ การประเมินนี้ใช้แบบฟอร์มซึ่งกำหนดไว้เป็นแนวทางในการประเมิน ถ้าผู้ประเมินต่างกันอาจมีความลำเอียงต่อผู้สอบต่างกัน

### การประเมินตนเอง

เป็นการประเมินการปฏิบัติของตนเองอย่างเป็นทางการและไม่เป็นทางการ การประเมินอย่างไม่เป็นทางการอาจเป็นการอภิปรายความคิดเห็นร่วมกับผู้เชี่ยวชาญ หรือครู การประเมินที่เป็นทางการ การใช้แบบประเมินตนเอง การสัมภาษณ์อย่างมีโครงสร้าง การใช้แบบสำรวจรายการหรือการใช้มาตราส่วนประมาณค่า การประเมินตนเองอาจใช้ประเมินกระบวนการประเมินผลเฉพาะด้าน หรือใช้ประเมินโปรแกรม จะต้องรู้ว่าประเมินตนเองอย่างไร จะใช้เกณฑ์อะไรเป็นพื้นฐานในการประเมิน

### ข้อดี-

- ข้อดีผู้เรียนสามารถรับรู้โดยไม่ต้องตื่นเต้นตกใจหรือเครียดกับผลของการประเมิน
- ข้อเสียของการประเมินตนเอง ข้อเสียคือขาดประสบการณ์ ในระยะแรกอาจนำผลการประเมินมาใช้ประโยชน์ไม่ได้มากนัก เนื่องจากการประเมินอาจจะยังไม่สอดคล้องกับความเป็นจริง หรือประเมินโดยขาดความเป็นปรนัย

### การประเมินโดยใช้เทคนิคการสังเกต

ผู้สังเกตงานภาคปฏิบัติจะต้องระบุประเด็นที่ต้องประเมินและกำหนดน้ำหนักความสำคัญให้กับประเด็น แต่ละประเด็นของกระบวนการหรือผลงาน ในทางกลับกันการวัดผลโดยให้คะแนนตามอารมณ์ของผู้สอนเป็นการบั่นทอนแรงจูงใจในการเรียนของผู้เรียน

### ข้อดี

- การประเมินโดยใช้เทคนิคการสังเกต มีความตรงเชิงทำนายสูงกว่าการทดสอบด้วยวิธีอื่น

ครูบอกข้อผิดพลาดและแก้ไข้ปัญหาที่เกิดขึ้นได้ทันที เป็นการตรวจสอบความสามารถของแต่ละบุคคลเพื่อ  
ประยุกต์ใช้ข้อมูลในสถานการณ์ที่เหมือนชีวิตจริง

- วิธีการสังเกตจะให้ข้อมูลทั้งในเชิงปริมาณและเชิงคุณภาพ
- ปรับใช้กับงานประเภทต่าง ๆ กับบุคคลทุกวัยและทุกระดับได้ง่าย

#### ข้อเสีย

- ผู้สังเกตที่มีอคติหรือคาดหวังหรือมีความรู้ที่ผิด ๆ จะทำให้บิดเบือนข้อมูล
- ไม่เข้าใจจุดประสงค์ดีพอ ไม่สามารถแยกแยะระหว่างข้อเท็จจริงกับความคิดเห็นได้
- การสรุปข้อมูลจากเหตุการณ์เพียงเหตุการณ์เดียว ทำให้ผิดพลาดจากความเป็นจริง
- ต้องพิจารณาบริบทแวดล้อม ประกอบด้วย
  - การสุ่มตัวอย่างพฤติกรรมจำกัดเกินไป
  - ความผิดพลาดในการแยกแยะพฤติกรรม
  - การบันทึกไม่ตรงกับข้อมูลที่แท้จริง
- การสรุปจากข้อมูลที่ไม่เพียงพอจะทำให้เกิดความผิดพลาด

#### สรุปการประเมินการปฏิบัติ

ความหมาย

เป้าหมายในการสอบภาคปฏิบัติ

ประเด็นสำคัญในการสอบภาคปฏิบัติ

รูปแบบและลักษณะการสอบวัดภาคปฏิบัติ

ขั้นตอนในการสอบภาคปฏิบัติ

## การประเมินตามสภาพจริง (Authentic Assessment)

ด้วยหลักสูตรปัจจุบันเป็นหลักสูตรอิงมาตรฐาน การจัดการเรียนการสอนจึงได้นำแนวคิดการสร้างสรุควิชาความรู้ ทักษะการสร้างสรุควิชาความรู้มาใช้เป็นกรอบแนวคิดในการออกแบบการจัดการเรียนการสอน ดังนั้น การวัดประเมินผลในชั้นเรียนเราก็ควรมองหาหนทางใหม่ เพื่อให้การประเมินผลสอดคล้องกับแนวคิดที่กำลังเปลี่ยนไปในกระบวนการสร้างและพัฒนาแบบวัดมาตรฐานเราได้เรียนรู้และใช้มานานและมีความรู้ ความชำนาญมากพอ แต่เราต้องทำความเข้าใจทั้งในแง่ความคิดรวบยอด ความหมายและเทคนิควิธีการต่าง ๆ ของการประเมิน การประเมินตามสภาพจริง ซึ่งนับเป็นทางเลือกใหม่ของการประเมินผลที่สอดคล้องกับหลักสูตรใหม่

การประเมินตามสภาพจริง (Authentic Assessment) หมายถึง การประเมินที่เปิดโอกาสให้นักเรียนปฏิบัติงานที่เหมือนการปฏิบัติงานในชีวิตจริง มีเวลาเพียงพอสำหรับวางแผน การลงมือทำงาน จนได้งานที่เสร็จสมบูรณ์ มีโอกาสประเมินผลการทำงานด้วยตนเองและมีการปรึกษาร่วมกับนักเรียน การประเมินลักษณะเช่นนี้ จึงจำเป็นต้องใช้วิธีการประเมินที่หลากหลายมีการตัดสินโดยใช้เกณฑ์ (criteria) หรือมาตรฐาน (standard) เดียวกับเกณฑ์ หรือ มาตรฐานที่ใช้ตัดสินการทำงานในชีวิตจริงในชิ้นงานเดียวกันหรือประเภทเดียวกัน เช่น การกำหนดให้นักเรียนจัดทำนิทรรศการวิชาการ การจัดทำหนังสือเล่มเล็กเกี่ยวกับชุมชน การผลิตสบู่อจากพืชสมุนไพรแต่ ในวิชากลุ่มวิทยาศาสตร์และคณิตศาสตร์ การจำสูตรต่าง ๆ ได้ จึงมิใช่ชิ้นงานตามสภาพที่แท้จริง แต่การใช้สูตรเพื่อแก้ปัญหาภายใต้สภาพแวดล้อมหรือสภาพที่เกิดขึ้นจริงหรือใกล้ความจริงให้มากที่สุดเท่าที่จะเป็นไปได้ จึงจะเป็นชิ้นงานตามสภาพจริง (กรมวิชาการ, 2546) จากการเปลี่ยนแปลงหลักสูตรในปี พ.ศ. 2544 จึงทำให้เกิดการพัฒนาการประเมินผลจากแบบเดิม (Tradition Assessment) ไปสู่การประเมินตามสภาพจริง (Authentic Assessment) อย่างต่อเนื่องการประเมินตามสภาพจริงไม่สนับสนุนให้ทดสอบเพียงอย่างเดียว แต่เน้นไปที่การวิเคราะห์ ทักษะและความสามารถในการบูรณาการสิ่งที่นักเรียนได้เรียนรู้สร้างสรุควิชาความสามารถในการทำงานกลุ่มร่วมกันกับผู้อื่นจนผลิตผลงานได้สำเร็จ การประเมินตามสภาพจริงให้ความสำคัญกับกระบวนการเรียนรู้มากพอ ๆ กับผลผลิตหรือชิ้นงานที่เสร็จสมบูรณ์ (Pearson Education Development Group, 2006)

### เทคนิควิธีการประเมินตามสภาพจริง

เทคนิคการประเมินตามสภาพจริงมีหลายวิธี ในที่นี้ขอนำเสนอ 7 วิธี ดังนี้

**1) ประเมินโดยใช้ภาระงานที่สร้างขึ้นตามคำสั่ง** (Structured on demand task) การประเมินความสามารถของนักเรียนในการจัดการงานที่มีความแตกต่างกันมาก และเป็นงานที่ต้องมีการปฏิบัติ การควบคุมทั้งภาระงาน ชิ้นงานต่าง ๆ และบริบทหรือเหตุการณ์ต่าง ๆ ซึ่งนักเรียนจะต้องกระทำหรือ แสดงออกมา ในกรณีเช่นนี้จะใช้ภาระงานหรือชิ้นงานที่กำหนดกรอบ ขอบเขตให้ทำขึ้นหรือการให้ทำกิจกรรม ครูจะเป็นคนตัดสินใจว่าเมื่อไร วัสดุอะไรที่จะถูกนำมาใช้ จะมีคำแนะนำเฉพาะในความสามารถนั้นๆ มีคำอธิบายลักษณะของผลผลิตที่นักเรียนจะทำออกมา ภาระงานในลักษณะเช่นนี้ เรียกว่า ภาระงานตามคำสั่ง (On demand task) ซึ่งแบ่งเป็น 2 ลักษณะ คือ

(1) ภาระงานประเภท Paper and Pencil ให้นักเรียนบันทึก เขียนคำอธิบายประเด็น ต่าง ๆ การให้เหตุผล และแสดงวิธีการที่นำไปสู่การแก้ปัญหา มิใช่ระบุเพียงคำตอบอย่างเดียวเท่านั้น

(2) ภาระงานประเภทจัดเตรียมและคัดสรรทรัพยากรการเรียนรู้ เป้าหมายการเรียนรู้ที่สำคัญของภาระงานประเภทนี้ คือ นักเรียนจะต้องเตรียมการและคัดสรรทรัพยากรการเรียนรู้เพื่อปฏิบัติภาระงาน

บางอย่างได้สำเร็จ ดังนั้นนักเรียน จะต้องนำเสนอวิธีการจัดเตรียมและคัดสรรทรัพยากรต่าง ๆ ด้วยทางเลือกที่หลากหลาย เพื่อให้ภาระงานนั้นสำเร็จอย่างสมบูรณ์และมีประสิทธิภาพ

**2) ประเมินโดยใช้ภาระงานที่แสดงถึงความสามารถที่มีลักษณะเดียวกันหรือร่วมกันในกลุ่มสาระหลาย ๆ กลุ่ม** (Naturally occurring or Typical performance task) แทนที่จะคิดสร้างสรรค์ภาระงานขึ้นงาน และควบคุมสถานการณ์การประเมิน เราสามารถใช้ความสามารถที่เกิดขึ้นและประเมิน ณ เวลาหนึ่งตัวอย่างเช่น การประเมินความสามารถด้านการเขียน ครูสามารถเลือกประเมินชิ้นงานจากทุกกลุ่มสาระที่เกี่ยวข้องกับการเขียน โดยวิเคราะห์เกี่ยวกับการใช้ไวยากรณ์ การสะกดคำ และข้อผิดพลาดเพื่อจำแนกชนิดการใช้ภาษาของนักเรียน เป็นต้น

### 3) ประเมินโดยใช้โครงการระยะยาว (Longer-term project)

**1 โครงการรายบุคคล** เป็นกิจกรรมที่นักเรียนต้องใช้เวลาในการทำผลงาน งานที่เกิดขึ้นจากการปฏิบัติงานของนักเรียน เช่น การสร้างแบบจำลอง (Model) ผังแสดงความสัมพันธ์ของเนื้อหาการรายงาน การเก็บรวบรวมข้อมูลของงาน โครงการที่คิดขึ้นเองซึ่งเป็นโครงการที่ต้องการให้นักเรียนได้ประยุกต์ ตัดแปลง และบูรณาการในวงกว้าง ในการกำหนดสมมติฐาน การใช้เอกสารอ้างอิงและแหล่งค้นคว้าต่าง ๆ การวางรูปแบบบทความ การจัดการ และการวางแผนรายงาน ภาษาที่ใช้ในการเขียนเพื่อสื่อความหมาย ลักษณะการนำเสนอ และการอธิบายในสิ่งที่เขาเข้าใจในแต่ละหัวข้อ ซึ่งนักเรียนจะต้องคิดวิเคราะห์และเขียนขึ้นมา โครงการต่าง ๆ เป็นกิจกรรมทางการศึกษาที่คุ้มค่าคุ้มเวลาและมีประโยชน์ต่อการนำไปใช้เป็นเครื่องมือสำหรับประเมินนักเรียนเป็นรายบุคคลโดยที่ครูต้องมั่นใจว่า

- (1) โครงการนั้นตรงตามเป้าหมายการเรียนรู้ที่สำคัญของหลักสูตร
- (2) นักเรียนแต่ละคนได้ทำงานหรือชิ้นงานด้วยตัวเอง
- (3) นักเรียนแต่ละคนมีความเท่าเทียมกันในการใช้ทรัพยากรที่จำเป็นต่อการเกิดผลสำเร็จ

ระดับยอดเยี่ยมจากการประเมิน

- (4) การประเมินผลสำเร็จของโครงการต้องยุติธรรม ชัดเจน ไม่มีอคติ

**2 โครงการกลุ่ม (Group project)** โครงการเหล่านี้มีจุดมุ่งหมายเพื่อให้นักเรียน 2-3 คนหรือมากกว่านี้ได้มาทำงานร่วมกัน เป็นเทคนิคการประเมินการปฏิบัติงานร่วมกันของนักเรียน โดยการร่วมมือและสร้างสรรค์ผลผลิตที่มีคุณภาพระดับสูงเมื่อมีโครงการที่ต้องใช้เวลาหลายสัปดาห์ ครูจำเป็นต้องมีแผนในการดูแลเป็นที่เลี้ยงให้นักเรียน เพื่อให้เกิดความมั่นใจว่านักเรียนมีความก้าวหน้าอย่างต่อเนื่องในการทำโครงการ และเพื่อมีให้นักเรียนประสบปัญหาต่าง ๆ ซึ่งจะทำให้โครงการล้มเหลว คำแนะนำข้างล่างนี้น่าจะทำให้การจัดการชัดเจนขึ้น และเป็นประโยชน์ต่อโครงการกลุ่มและโครงการรายบุคคล เพื่อใช้เป็นเป้าหมายการประเมิน

(1) ครูมีภาพหรือลักษณะของผลผลิตในเกณฑ์ที่จะยอมรับอย่างชัดเจน เพื่อให้เกิดความมั่นใจว่านักเรียนเข้าใจอย่างชัดเจนถึงเป้าหมายการเรียนรู้ที่จะใช้ประเมิน โดย

(1.1) อธิบายเกี่ยวกับตัวอย่างโครงการที่มีคุณภาพระดับสูง ซึ่งเราได้เก็บรวบรวมจากนักเรียนรุ่นก่อนๆ หรือมีตัวอย่างงานที่มีคุณภาพ

(1.2) ชี้แจงให้นักเรียนทราบเกี่ยวกับรายละเอียดของสิ่งที่ยอมรับในการดำเนินงานของโครงการ

(2) มีความแจ่มชัดในมาตรฐานต่าง ๆ ที่เราจะใช้ประเมินโครงการ เช่น ระดับคุณภาพ นิยามการให้คะแนน การกำหนดน้ำหนักคะแนน เป็นต้น อธิบาย ชี้แจง และอนุญาตให้นักเรียนจดเกณฑ์การให้คะแนนที่เราจะใช้ประเมินโครงการ

- (3) กำหนดวันเวลาส่งโครงการที่แน่นอน ซึ่งจะต้องให้เวลาเพียงพอที่นักเรียนสามารถพัฒนา

โครงการและดำเนินโครงการ ได้สำเร็จสมบูรณ์

(4) ให้มีการรายงานความก้าวหน้าของโครงการเป็นระยะ ๆ เพื่อช่วยให้นักเรียนทำงานโครงการอย่างต่อเนื่อง และเป็นการยอมรับว่าเรากำลังประเมินกระบวนการทำงานจนกระทั่งได้โครงการที่เสร็จสมบูรณ์ในที่สุด

**4) ประเมินจากแฟ้มสะสมงาน (Portfolio)** แฟ้มสะสมงานเป็นผลงานของนักเรียนที่เก็บรวบรวมไว้ ครูสามารถใช้เป็นเครื่องมือประเมินร่องรอย/หลักฐานที่นักเรียนนำความรู้ต่าง ๆ และทัศนคติไปประยุกต์ใช้ในการทำงาน ซึ่งมีร่องรอยให้เห็นถึง จุดเด่น-จุดด้อย ด้านต่าง ๆ ของนักเรียนแฟ้มสะสมงานแบ่งได้เป็น 2 ลักษณะ คือ แฟ้มสะสมงานที่เก็บงานที่ดีที่สุด (อาจมีหลายชิ้น) กับแฟ้มสะสมงานที่แสดงความก้าวหน้าทางการเรียนรู้

(1) แฟ้มสะสมงานที่เก็บงานที่ดีที่สุด (Best-Work Portfolio) มุ่งเน้นการนำเสนอผลผลิตชิ้นสุดท้ายที่ดีที่สุดของนักเรียน เพื่อการประเมินนักเรียนจะต้องเรียนรู้วิธีการและทักษะต่าง ๆ ที่จะใช้จัดแฟ้มสะสมงานเพื่อนำเสนอผลงานที่ดีที่สุดซึ่งจะถูกเก็บไว้ในแฟ้มว่าควรทำอย่างไร นักเรียนจำเป็นต้องเรียนรู้และตัดสินใจว่าเขาควรใช้การสื่อสาร หรือการคิดภาพความสำเร็จของแฟ้มสะสมงานอยู่ในลักษณะใด การเลือกชิ้นงานต่าง ๆ ที่จะบรรจุลงในแฟ้มสะสมงาน คุณภาพของชิ้นงานนั้นตรงหรือสอดคล้องกับเกณฑ์การให้คะแนนที่จะนำมาใช้ประเมิน

(2) แฟ้มสะสมงานที่แสดงความก้าวหน้าทางการเรียนรู้และการเจริญเติบโต (Growth and Learning-Progress Portfolio) เน้นที่ขั้นตอนการทำงานมากกว่าผลผลิตที่ทำสำเร็จ นักเรียนจะต้องเข้าใจและประเมินความก้าวหน้าของตนด้วยตัวนักเรียนเอง นักเรียนจะต้องกะพื้นที่ในแฟ้มสะสมงานเพื่อเก็บรวบรวมผลงานที่แสดงความก้าวหน้า สิ่งเหล่านี้จะถูกเก็บไว้ช่วงต้น ๆ ของแฟ้มการบันทึกการคิดและการเขียนแล้วเขียนใหม่ ผลงานที่ทำเสร็จแล้วก็ถูกจัดเก็บไว้ในแฟ้มสะสมงานเช่นเดียวกัน แฟ้มสะสมงานที่แสดงถึงความเจริญเติบโตและความก้าวหน้าทางการเรียนรู้จะต้องมีประสิทธิภาพ และคิดจัดทำอย่างละเอียดรอบคอบ ดังนี้

ประการแรก เป้าหมายการเรียนรู้ต้องชัดเจน รวมทั้งสิ่งที่ประสงค์แสดงถึงความก้าวหน้าการเรียนรู้ของนักเรียนด้วย

ประการที่สอง ต้องเข้าใจการเรียนรู้อย่างแจ่มชัดทฤษฎี ใช้ทฤษฎีเหล่านี้ช่วยกำหนดสิ่งที่จะมองหาว่า เมื่อไรจะประเมินการเปลี่ยนแปลงการพัฒนาแนวความคิดของนักเรียนหรือวินิจฉัยการเรียนรู้ยากของนักเรียน

ประการที่สาม ถ้ามีครูหลาย ๆ คนที่ใช้แฟ้มสะสมงานลักษณะเช่นนี้ ควรช่วยเหลือร่วมมือและทำงานร่วมกันด้วย

ประการที่สี่ ต้องเลือกใช้วิธีการประเมิน กำหนดเกณฑ์การประเมินเพื่อช่วยให้ผู้ประเมินอยู่กับร่องกับรอย ต้องประยุกต์ใช้เกณฑ์ให้เหมาะสมกับการพิจารณานักเรียนทั่วไปและนักเรียนที่ใช้เวลามากเกินไปด้วย

**5) ประเมินจากการแสดง การสาธิต (Demonstration)** เป็นความสามารถของนักเรียนที่ทำตามคำสั่ง ในการใช้ความรู้และทักษะต่าง ๆ ปฏิบัติภาระงานหรือชิ้นงานที่สลับซับซ้อน การแสดงมีต้องใช้เวลานานหรือความสลับซับซ้อนเหมือนกับโครงการต่าง ๆ การแสดงเป็นภาระงานที่แสดงออกโดยกำหนดจำกัดขอบเขต ถ้าเราจะใช้การสาธิต การแสดงเพื่อเป็นเป้าหมายการประเมินเราควรกำหนดเป้าหมายการเรียนรู้อย่างรอบคอบ และกำหนดเกณฑ์การให้คะแนนที่มีความสอดคล้องและเหมาะสม

**6) ประเมินจากการทดลองและการสืบสวน (Experiment and Inquiry)** การทดลองหรือการสืบสวนเป็นการปฏิบัติงานที่ต้องการให้นักเรียนวางแผนลงมือปฏิบัติ และการแปลผลของการศึกษา

วิจัยเชิงประจักษ์ การศึกษามุ่งเน้นการตอบคำถามเฉพาะเจาะจงกับการทดลองหรือสืบสวน ถ้าเราจะนำมาประเมินนักเรียนที่ใช้ทักษะการสืบสวนสอบสวน (Inquiry) และวิธีการต่าง ๆ จะสามารถประเมินนักเรียนที่ได้รับการพัฒนากรอบแนวความคิด และความคิดเชิงทฤษฎี การอธิบายกฎเกณฑ์ของปรากฏการณ์ที่นักเรียนได้สืบสวน การประเมินสิ่งเหล่านี้มุ่งเน้นไปที่คุณภาพของกรอบการอ้างอิงของนักเรียน การนำเสนอปัญหาที่กำลังศึกษา กระบวนการวางแผน หรือการออกแบบการวิจัยคุณภาพของข้อคำถาม คำอธิบายที่นำเสนอ เช่น ทำไมข้อมูลมีความสัมพันธ์กัน การประเมินความสามารถในการทดลองหรือการสืบสวนของภาระงานจะประเมินความสามารถของนักเรียน ในหัวเรื่องต่อไปนี้

- (1) การประมาณค่าและทำนายผลก่อนที่จะเริ่มเก็บข้อมูล
- (2) การเก็บรวบรวมข้อมูล วิเคราะห์ และแสดงผลของการวิเคราะห์
- (3) การสรุปและสนับสนุนหรือขัดค้านโดยการอ้างหลักฐานที่ชัดเจน
- (4) การระบุโอกาสผิดพลาดจากวิธีการและการได้ข้อมูลมา
- (5) นำเสนอข้อค้นพบจากการทดลองหรือการสืบสวนสอบสวน

### 7) ประเมินจากการนำเสนอด้วยวาจาและการแสดงละคร (Oral Presentation and

Dramatization) เทคนิควิธีนี้ประกอบด้วยการให้นักเรียนนำเสนอด้วยการพูด (Oral Presentation) โดยเปิดโอกาสให้นักเรียนอธิบายเกี่ยวกับองค์ความรู้และใช้ทักษะการพูดในรูปแบบต่าง ๆ เช่น การสัมภาษณ์ การกล่าวคำปราศรัย การบรรยาย การพูดโต้ตอบ การโต้เถียง การอภิปราย รวมทั้งการแสดงละคร ซึ่งครูจะต้องตัดสินใจให้ได้ว่าเป้าหมายของการนำเสนอรูปแบบนั้น ๆ คืออะไร เพื่อครูจะได้กำหนดเกณฑ์การประเมินได้ถูกต้องและจากเทคนิควิธีการประเมินตามสภาพจริง ที่กล่าวไว้ข้างต้นทั้ง 7 วิธี ครูควรเลือกใช้ให้เหมาะสมกับธรรมชาติของการเรียนรู้ เครื่องมือการวัดที่นิยมใช้ มีด้วยกันมากมายหลากหลาย เช่น การทดสอบ การสัมภาษณ์ แบบประเมินผลงาน เป็นต้น

### เครื่องมือการประเมินตามสภาพจริง

สำหรับเครื่องมือที่นำมาใช้ในการประเมินตามสภาพจริงนั้น สามารถมีได้หลายประเภท ดังต่อไปนี้

วิธีวัด	ตัวอย่างเครื่องมือวัด
การทดสอบ (Testing)	แบบสอบข้อเขียน (Written Test)
	แบบเลือกตอบ แบบจับคู่ แบบถูก-ผิด เป็นต้น
แบบวัด (Scaleภาคปฏิบัติ)	แบบสอบภาคปฏิบัติ( Performance Test)
การสัมภาษณ์ (Interview)	แบบสัมภาษณ์ (Interview Guide)
การสอบถาม (Inquiry)	แบบสอบถาม (Questionnaire)
การตรวจสอบรายการ (Checklist)	แบบมาตราประเมินค่า (Rating Scale)
การสังเกต (Observation)	แบบบันทึก ( Record)
การประเมินพฤติกรรม	แบบประเมินพฤติกรรม
	แบบประเมินตนเอง
การตรวจผลงาน	แบบประเมินผลงาน
การใช้แฟ้มสะสมงาน	แบบบันทึก (Record)

จากเครื่องมือประเมินตามสภาพจริง ดังกล่าวข้างต้น ครูต้องสร้างเกณฑ์การประเมินให้ถูกต้องเหมาะสมกับสิ่งที่จะวัด

#### ความแตกต่างระหว่าง Performance และ Authentic Assessment

Performance Assessment และ Authentic Assessment มักถูกนำไปใช้สลับกันอยู่เสมอ แต่ความหมายของคำทั้งสองนี้เหมือนกันหรือเปล่า แม้ว่าทำทั้งคู่นี้จะมีคำว่า Assessment อยู่ด้วยเหมือนกัน แต่คำทั้งสองก็ไม่ใช่คำเหมือน (Synonymous) เราต้องชัดเจนในความแตกต่างของทั้งสองคำนี้ให้ได้

กรณีตัวอย่าง	Performance Assessment การประเมินภาคปฏิบัติ	Authentic Assessment(ประเมินตามสภาพจริง)
เป็นการประเมินงานเขียนของนักเรียน โดยใช้เวลา 4 วัน โดยให้นักเรียนทั้งหมดได้เขียนบทความขึ้นมาหนึ่งชิ้นโดยมีครูเป็นผู้แนะนำนักเรียนและให้นักเรียนเขียนด้วยเวลา 45 นาทีต่อวัน ในวันแรกนั้นจะให้นักเรียนเลือกหัวข้อก่อน ต่อมาในวันที่สองเขียนโครงเรื่อง วันที่สามลงมือเขียน และวันที่สี่เขียนสรุปและ	เป็น Performance Assessment หรือไม่ คำตอบคือ ใช่ นักเรียนจะถูกถามในการพฤติกรรมที่ปฏิบัติที่พวกเขาได้พยายามเขียนออกมา	ไม่เป็น

ตรวจสอบแก้ไข		
การประเมินการเขียนโดยให้นักเรียนแต่ละคนปรึกษากับครูของเขาในการเลือกหัวข้อของงานเขียนให้สอดคล้องกับวัตถุประสงค์ในการประเมิน โดยงานเขียนจะรวบรวมอยู่ในแฟ้มสะสมงานที่ไม่ได้อยู่ภายใต้เงื่อนไขที่กำหนดเป็นมาตรฐาน แต่จะเป็นการสะสมงานเขียนที่นักเรียนทำทั้งปี ซึ่งงานเขียนทั้งหมดจะถูกเขียนขึ้นและปรับปรุงตามกระบวนการเขียนของนักเรียนเอง	ไม่เป็น	เป็น Authentic Assessment หรือไม่ คำตอบคือใช่ ในการปฏิบัติ จะประเมินเนื้อหาที่ประสบพบเจอในชีวิตจริง ตัวอย่างเช่นนักเรียนมีอิสระที่จะกำหนดความยาวหรือความหลากหลายของขั้นตอนในกระบวนการเขียน สร้างสรรค์ทั้งจำนวนและเค้าโครงด้วยตนเองอย่างสมบูรณ์ตั้งแต่ต้นจนจบ

### สรุปความแตกต่าง

Performance Assessment จะเป็นการประเมินงานที่นักเรียนทำตามเงื่อนไขที่กำหนด

Authentic Assessment จะเป็นการประเมินงานที่นักเรียนได้ลงมือปฏิบัติด้วยตนเองอย่างอิสระ

### เกณฑ์การประเมิน (Rubric)

การประเมินว่านักเรียนมีความรู้ ความสามารถตามมาตรฐานการเรียนรู้ นั้น ครูจะต้องสร้างเกณฑ์การประเมิน การปฏิบัติงานต่าง ๆ ของนักเรียน เพื่อเป็นแนวทางในการให้คะแนนโดยเกณฑ์การประเมินจะต้องสอดคล้องกับมาตรฐานการเรียนรู้ที่ได้กำหนดไว้ในแต่ละครั้งของการปฏิบัติงานนั้น ๆ

1) ความหมายของเกณฑ์การประเมิน คำว่า Rubric หมายถึง กฎ หรือ กติกา (Rule) ส่วนคำว่า Rubric Assessment นั้น หมายถึง การประเมินเชิงคุณภาพ ที่สามารถจะแยกแยะระดับความสำเร็จในการเรียน หรือคุณภาพการปฏิบัติของนักเรียนได้อย่างชัดเจน โดยการกำหนดเป็นแนวทางการให้คะแนนจากตีมากไปจนถึงต้องปรับปรุงแก้ไข

2) การกำหนดเกณฑ์การประเมิน ครูและนักเรียนควรทำความตกลงและกำหนดเกณฑ์การประเมินร่วมกัน ก่อนที่นักเรียนจะได้ลงมือปฏิบัติงานชิ้นนั้น เกณฑ์การประเมินนี้ นอกจากใช้เป็นเครื่องมือในการประเมินแล้ว ยังใช้เป็นเครื่องมือในการสอนด้วย เพราะ เปรียบเสมือนเป้าหมายในการเรียนที่นักเรียนและครูจะต้องทราบ ซึ่งแนวคิดนี้สอดคล้องกับแนวคิดของ Marzano et al. (1993) ว่าการประเมินการปฏิบัตินั้นต้องกำหนดเกณฑ์ให้เหมาะสม ซึ่งเกณฑ์ในการให้คะแนนจะต้องมีระดับสเกลที่แน่นอน และมีการบรรยายคุณลักษณะที่สำคัญของแต่ละระดับอย่างชัดเจนให้แก่ครู ผู้ปกครองและบุคคลอื่น ๆ ที่เกี่ยวข้อง ทำให้ทราบว่านักเรียนรู้อะไรและทำอะไรได้บ้าง

3) รูปแบบของเกณฑ์การประเมิน Julia Jasmine (1993) และ Concetta Doti Ryan (1994) ได้แบ่งเกณฑ์การประเมินออกเป็น 2 ประเภท ดังนี้

(1) เกณฑ์การประเมินในภาพรวม (Holistic Rubric) คือแนวทางการให้คะแนนโดยพิจารณาจากภาพรวมของชิ้นงาน จะมีคำอธิบายลักษณะของงานในแต่ละระดับไว้อย่างชัดเจน เกณฑ์การประเมินในภาพรวมนี้เหมาะที่จะใช้ในการประเมินความสามารถที่มีความต่อเนื่อง มีลักษณะเป็นองค์รวม เกณฑ์การประเมินในภาพรวมส่วนใหญ่จะประกอบด้วย 3-6 ระดับ ซึ่งเกณฑ์การประเมิน 3 ระดับจะเป็นที่นิยมใช้กัน



มาก เนื่องจากกำหนดรายละเอียดง่าย โดยใช้เกณฑ์ค่าเฉลี่ย (อยู่ระดับกลาง) สูงกว่าค่าเฉลี่ย และต่ำกว่าค่าเฉลี่ย นอกจากง่ายต่อการกำหนดค่าแล้วยังง่ายต่อการตรวจให้คะแนนอีกด้วย เนื่องจากความแตกต่างระหว่างระดับนั้น จะชัดเจน แต่ถ้าใช้ 5 หรือ 6 ระดับ ความแตกต่างระหว่างระดับจะแตกต่างกันเพียงเล็กน้อย ซึ่งตรวจให้คะแนนยาก ถ้าต้องการให้เกณฑ์ 5 หรือ 6 ระดับ วิธีการที่จะช่วยในการกำหนดเกณฑ์ให้ง่ายขึ้น ครูอาจสุ่มตัวอย่างงานของนักเรียนมาตรวจ จากนั้นในแต่ละกองจะต้องแยกความแตกต่างให้ได้อีก 2 กอง ตามระดับคุณภาพของงาน ในกรณีที่ต้องการทำเป็น 5 กอง กองที่เป็นคุณภาพปานกลางจะไม่แบ่ง แล้วนำมากำหนดเกณฑ์การให้คะแนนให้ชัดเจนยิ่งขึ้น

(2) เกณฑ์การประเมินแบบแยกส่วน (Analytic Rubric) คือแนวทางการให้คะแนนโดยพิจารณาจากแต่ละส่วนของงาน ซึ่งแต่ละส่วนจะต้องกำหนดแนวทางการให้คะแนนโดยมีคำนิยามหรือคำอธิบายลักษณะของงานส่วนนั้น ๆ ในแต่ละระดับไว้อย่างชัดเจน

การกำหนดประเด็นการประเมินและรายละเอียดการให้ระดับคะแนน มีความจำเป็นที่นักประเมินผลควรคำนึง เพราะเป็นคุณภาพของการประเมินผล คือ ความเที่ยงตรงและความเชื่อมั่น คุณภาพทั้งสององค์ประกอบนี้จะมีผลถึงศักยภาพของนักเรียนในการนำความรู้ไปใช้ปฏิบัติงาน ผลิตผลงาน ตลอดจนคุณลักษณะที่พึงประสงค์ของนักเรียนตามหลักสูตร และจะเป็นคุณภาพของการจัดการเรียนการสอน และการประเมินตามพระราชบัญญัติการศึกษา พุทธศักราช 2542 ซึ่งก็เป็นส่วนหนึ่งของการปฏิรูปการศึกษาอีกด้วย

### องค์ประกอบของเกณฑ์การให้คะแนน (Scoring Rubric)

Scoring rubric มีหลายองค์ประกอบ ในแต่ละองค์ประกอบก็มีประโยชน์ มีความสำคัญ องค์ประกอบมีดังนี้

1. จะมีอย่างน้อย 1 คุณลักษณะหรือ 1 มิติที่เป็นพื้นฐานในการตัดสินผู้เรียน
2. การนิยามและการยกตัวอย่างจะต้องมีความชัดเจนในแต่ละคุณลักษณะหรือมิติ
3. มาตรการให้คะแนนจะต้องเป็นอัตราส่วนกันในแต่ละคุณลักษณะหรือมิติ
4. จะต้องมีความมาตรฐานที่เด่นชัดในแต่ละระดับของการให้คะแนน

### เกณฑ์การให้คะแนน (Scoring Rubric) มีความสำคัญอย่างไร

การประเมินศักยภาพของผู้เรียนโดยให้ลงมือปฏิบัตินั้น ไม่มีค่าเฉลี่ยหรือคำตอบถูกที่แน่ชัดลงไปเหมือนแบบทดสอบเลือกตอบ การประเมินผลงานแต่ละชิ้นของผู้เรียนที่ได้ลงมือปฏิบัติจึงมีความจำเป็นที่จะต้องประเมินคุณภาพของงานอย่างเป็นปรนัย ซึ่งมันเป็นการยากที่จะทำได้ และได้ค้นพบการสร้างเกณฑ์การให้คะแนนหรือ rubric ขึ้นมาเพื่อกำหนดแนวทางในการตัดสินอย่างยุติธรรม และปราศจากความลำเอียง

Rubric จะต้องมีความชัดเจนในเกณฑ์การให้คะแนนอย่างพอเพียงถึงขนาดที่ผู้ประเมิน 2 คนสามารถใช้ Rubric เดียวกันประเมินชิ้นงานของผู้เรียนชิ้นเดียวกันแล้วให้คะแนนได้ตรงกัน ระดับของความสอดคล้องในการให้คะแนนของผู้ประเมิน 2 คนที่ประเมินอย่างเป็นอิสระจากกันจะเรียกว่า ความเชื่อมั่น (Reliability) ของการประเมิน ในแต่ละระดับการให้คะแนนจะต้องมีความชัดเจนในการนิยาม และความกว้างของระดับคะแนนไม่ควรเกิน 6 ถึง 7 ระดับ ถ้ามีระดับของการให้คะแนนกว้างมากเกินไปจะมีความลำบากในการตัดสินความแตกต่างในแต่ละระดับ เช่น ความกว้างคะแนนเป็น 100 ทำให้ยากที่จะอธิบายว่าคะแนน 81 มีคุณภาพแตกต่างจาก 80 หรือ 82 อย่างไร และจะทำให้ความสอดคล้องของการประเมินด้วยผู้ประเมินหลายคนลดลง การจะกำหนดความกว้างของการให้คะแนนเป็นเท่าไรนั้น จะต้องมีความเหมาะสมและมีความชัดเจนในการนิยามที่ครอบคลุมตั้งแต่ แย่ที่สุด (poor) จนถึงดีเลิศที่สุด (excellent)

## คุณภาพของการประเมิน

คุณภาพของการประเมิน จะพิจารณาคุณลักษณะอย่างน้อยที่สุด 2 คุณสมบัติ คือ ความเที่ยงตรง (Validity) หรือความเชื่อมั่น (Reliability) (วาสนา ประวาลพฤกษ์, 2544) ได้กล่าวถึงเรื่องนี้ ในเอกสารสรุปคำบรรยาย การวัดและประเมินผลตามสภาพจริงไว้ ดังนี้

1) ความเที่ยงตรง ความเที่ยงตรงของการประเมินตามสภาพจริงเป็นความเที่ยงตรงในสภาพปัจจุบัน (Concurrent Validity) หากมีหลักฐานยืนยันควรจัดบันทึกความถี่ของพฤติกรรมที่นักเรียนแสดงออกอย่างสม่ำเสมอ ซึ่งอาจจะเป็นลักษณะประจำตัวของนักเรียน ซึ่งจะทำให้การประเมินตามสภาพจริงนั้นมีความเที่ยงตรงสูงเป็นการยืนยันว่านักเรียนมีลักษณะอย่างนั้นจริง

2) ความเชื่อมั่น หลักฐานจากการจัดบันทึกนั้น เราสามารถเชื่อถือได้มากน้อยแค่ไหน นั่นก็คือ “ความเชื่อมั่น” การประเมินตามสภาพจริงมีความเชื่อมั่นค่อนข้างต่ำ ถ้าหากว่าใช้หลักการในเชิงปริมาณ เพราะฉะนั้นจะต้องใช้หลักการวิจัยเชิงคุณภาพ การหาความเชื่อมั่นเพื่อที่จะสนับสนุนความเที่ยงตรงในเชิงปริมาณ เป็นการหาความสัมพันธ์ของคะแนนจากผู้ประเมิน 2-3 คน ถ้าค่าความเชื่อมั่น (ค่าสหสัมพันธ์) ประมาณ .7 ขึ้นไป แสดงว่าใช้ได้ แต่ถ้าค่าความเชื่อมั่นต่ำก็แสดงว่ายังใช้ไม่ได้ เช่น ต้องการที่จะหาความเชื่อมั่นของเกณฑ์การประเมินที่สร้างขึ้นก็ให้ครู 2 คนไปประเมินแล้วดูสหสัมพันธ์ของคะแนนจากครูคนที่ 1 กับครูคนที่ 2 ที่ใช้เกณฑ์การประเมินเดียวกัน ถ้าค่าสหสัมพันธ์ต่ำแสดงว่าเกณฑ์ไม่ชัดเจนก็ต้องปรับคำอธิบายในเกณฑ์การประเมินให้ชัดเจนยิ่งขึ้น

ดังนั้น ในการประเมินตามสภาพจริง ครูอาจพิจารณาวิธีวัด ซึ่งอาจใช้การสอบถาม สัมภาษณ์ แบบสังเกต รวมทั้งใช้การสอบแบบอัตนัย และแบบปรนัย

ผลงานที่ได้จากกระบวนการของนักเรียน และกระบวนการที่นักเรียนใช้เพื่อให้เกิดผลงาน จะประเมินในลักษณะใดขึ้นอยู่กับจุดมุ่งหมายในการเรียนรู้ อาจจะประเมินลักษณะใดลักษณะหนึ่งหรือประเมินทั้งสองลักษณะก็ได้ ผู้ประเมินจะต้องตัดสินคุณภาพของผลงานหรือกระบวนการปฏิบัติงานของผู้เรียนแต่ละคนที่มีระดับที่แตกต่างกันหลายระดับ ระดับที่แตกต่างกันอาจจะเป็นระดับคุณภาพของชิ้นงานที่ได้สร้างขึ้น หรือระดับของกระบวนการต่าง ๆ ที่ผู้เรียนแต่ละคนได้ใช้เพื่อให้เกิดผลงาน เพื่อให้การตัดสินใจสอดคล้องกับผู้เรียนแต่ละคน

ผู้ประเมินจะต้องใช้เกณฑ์ในการประเมินคุณภาพชิ้นงานของผู้เรียน เกณฑ์อาจจะอยู่ในเชิงคุณภาพหรือปริมาณ อาจจะมีลักษณะเป็นมาตราส่วนประมาณค่า (Rating scale) หรือแบบตรวจสอบ (Checklist) โดยปกติจะใช้ Rubric ในการประเมินจุดประสงค์การเรียนรู้เดี่ยวหรือส่วนใดส่วนหนึ่งของงานปฏิบัติ แต่การปฏิบัติงานที่มีซับซ้อน

ผู้ประเมินจะต้องประเมินจุดประสงค์การเรียนรู้ที่หลากหลายและประเมินหลาย ๆ ส่วนของการปฏิบัติ ผู้ประเมินจะต้องมีเกณฑ์การให้คะแนนที่มากมายเพื่อให้เหมาะกับจุดประสงค์การเรียนรู้ที่แตกต่างกันหรือเหมาะกับแต่ละส่วนของการปฏิบัติงาน การให้คะแนนจะอยู่ในรูปของตัวเลข โดยปกติจะเป็น 0-3 หรือ 1-4 ในแต่ละระดับของคะแนนจะขึ้นอยู่กับระดับของคุณภาพของงาน ดังนั้นตัวเลข 4 อาจหมายถึงระดับคุณภาพสูงสุด เลข 3 เป็นระดับคุณภาพรองลงมา คุณภาพของงานในแต่ละระดับจะต้องใช้การอธิบาย (Rubric) ดังนั้นในแต่ละระดับคะแนนจะต้องอธิบายเป็นภาษาที่แสดงให้เห็นถึงคุณภาพของการปฏิบัติงานในระดับนั้น

อ้างอิง

<http://learners.in.th/file/rakchanok/pamean+1.ppt>

[http://school.obec.go.th/sup\\_br3/cr\\_3.htm](http://school.obec.go.th/sup_br3/cr_3.htm)

เอกสารแนบ:การจัดการเรียนรู้และการวัดและประเมินผลการศึกษา2003.doc

เอกสารแนบ:แหล่งการเรียนรู้จาก Internet.doc