

การวัดและประเมินผลทางการศึกษา

การวัดและประเมินผลทางการศึกษาจะเกี่ยวข้องกับคำ 3 คำ คือ 1)การทดสอบ (testing) หมายถึง การนำเสนอชุดคำถามที่เรียกว่าข้อสอบหรือแบบทดสอบที่มีมาตรฐานให้ผู้สอบตอบ 2)การวัดผล (measurement) หมายถึงการวัดคุณลักษณะ (attribute)ของบุคคลจากผลการตอบคำถามในแบบทดสอบตามกฎเกณฑ์ที่กำหนด เพื่อแสดงคุณค่าเชิงปริมาณหรือตัวเลขที่วัดได้ การวัดผลนอกจากใช้แบบทดสอบแล้วยังรวมถึงการใช้เครื่องมืออื่นเพื่อรวบรวมข้อมูลเชิงปริมาณหรือเชิงคุณภาพด้วย เช่น การสังเกตพฤติกรรม การสัมภาษณ์ การตรวจผลงานต่าง ๆ ที่กำหนดให้ผู้ประเมินทำ 3)การประเมินผล (evaluation) หมายถึง กระบวนการอย่างมีระบบที่นำข้อมูลจากการวัดผลมาตีค่าและตัดสินคุณค่าของผู้เรียน ซึ่งการวัดผลและการประเมินผลเป็นกระบวนการที่มีความต่อเนื่อง เมื่อมีการวัดผลจะทำให้ได้ข้อมูลและรายละเอียดหลายด้าน เมื่อนำข้อมูลดังกล่าวมาวิเคราะห์ เปรียบเทียบกับเกณฑ์ใดเกณฑ์หนึ่งเพื่อตีค่า หรือสรุปคุณค่าออกมาถือว่าเป็นกระบวนการประเมิน ผลการประเมินจะมีความถูกต้องเที่ยงตรงเพียงใดขึ้นกับความถูกต้องของผลการวัด ถ้าผลการวัดต้องการประเมินก็จะต้องเชื่อถือได้มากและตรงกับความ เป็นจริง ถ้าผลการวัดผิดพลาด การประเมินก็จะผิดพลาดไปด้วย การวัดผลและการประเมินผลมีความแตกต่างกัน (ไพศาล สุวรรณน้อย, 2545) ดังตารางที่ 14

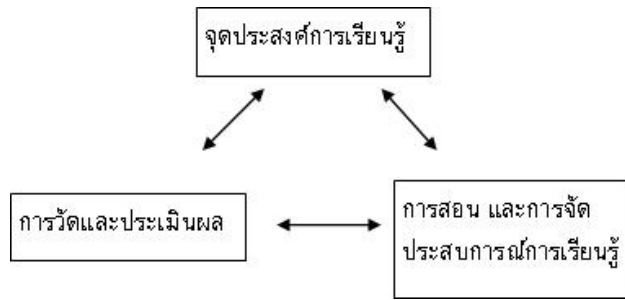
ตารางที่ 14 แสดงการเปรียบเทียบระหว่างการวัดผลและการประเมินผล

การวัดผล	การประเมินผล
1.เป็นการกำหนดรายละเอียด จำนวน หรือ ปริมาณ	1.เป็นการกำหนดระดับคุณค่า ตัดสินลงข้อสรุป
2.กระทำอย่างละเอียดทีละด้าน	2. สรุปรวมเป็นข้อชี้ขาด/ผลการตัดสิน
3.ใช้เครื่องมือเป็นหลัก	3.ใช้ผลการวัดเป็นหลักโดยพิจารณาตามเกณฑ์ที่กำหนดไว้แล้วล่วงหน้า
4.ผลที่ได้เป็นข้อมูลรายละเอียด	4.ผลที่ได้เป็นการตัดสินใจ
5.อาศัยวิธีการทางวิทยาศาสตร์	5.อาศัยการใช้ดุลยพินิจ

1.ความสำคัญของการวัดและประเมินผล

การจัดการเรียนการสอนหนึ่ง ๆควรมีการตรวจสอบคุณภาพของผู้เรียน ผู้สอน และกระบวนการสอนเป็นระยะ ๆ (formative evaluation)เพื่อพิจารณาตรวจสอบว่า ผู้เรียนมีคุณสมบัติหรือเกิดพฤติกรรมที่พึงประสงค์ตรงตามวัตถุประสงค์ของการเรียนการสอนตรงตามที่กำหนดไว้หรือไม่ กระบวนการวัดและประเมินผลนี้จะพยายามทำให้ได้ข้อมูลจากการจัดการเรียนการสอน เพื่อนำมาใช้วิเคราะห์ และตัดสินใจว่า การสอนดังกล่าวนั้นบรรลุผลหรือไม่ (summative evaluation)นำผลการตัดสินใจเพื่อประโยชน์ในการจัดลำดับ เลื่อนชั้นเรียนและพัฒนาปรับปรุงการเรียนการสอนต่อไป

การเรียนการสอนหากไม่มีการวัดและประเมินผลแล้ว ผู้สอนก็ไม่ทราบว่าคุณเรียนมีความรู้เพิ่มขึ้นหรือไม่ มากน้อยเพียงใด ไม่ทราบว่าการจัดการเรียนการสอนดังกล่าวมีประสิทธิภาพหรือไม่ เหมาะสมหรือไม่ หากต้องการพัฒนาปรับปรุงแก้ไข จะปรับปรุงพัฒนาตรงจุดไหน อย่างไร เป็นต้น ความสัมพันธ์และองค์ประกอบของการวัดและประเมินผลและกระบวนการจัดการเรียนการสอนแสดงไว้ในรูปที่ 19



รูปที่ 19 แสดงความสัมพันธ์ระหว่างจุดประสงค์การเรียนรู้ (Objective) การจัดประสบการณ์การเรียนรู้ (Learning experience) และการวัดและประเมินผล(Evaluation) หรือ โอเล่ OLE

2.ลักษณะสำคัญของการวัดและประเมินผลทางการศึกษา

การวัดผลทางการศึกษา เป็นกระบวนการวัดการเปลี่ยนแปลงพฤติกรรมของผู้เรียนนิยมนิยามวัดผลการเรียนรู้ เป็น 3 ด้านคือ พุทธิพิสัย (cognitive domain) จิตพิสัย (affective domain) และทักษะพิสัย (psychomotor domain) ซึ่งการวัดลักษณะของการวัดดังกล่าวมีประเด็นที่น่าสนใจดังนี้

1)เป็นการวัดทางอ้อม การวัดสติปัญญาซึ่งเป็นคุณลักษณะที่แฝงอยู่ในตัวบุคคล มักแสดงออกทางด้านพฤติกรรมที่ตอบสนองต่อสิ่งเร้าหรือเหตุการณ์ ดังนั้นการวัดผล จึงเป็นการวัดทางอ้อม โดยการสังเกตพฤติกรรมที่ตอบสนองต่อสิ่งเร้า เหตุการณ์ สภาวะการณ์ที่ผู้ประเมินกำหนด (หรือสร้างขึ้นเพื่อใช้สำหรับการประเมินหรือสถานการณ์สมมติ)

2)ลักษณะการวัดมักไม่สมบูรณ์ เนื่องจากลักษณะการวัดส่วนใหญ่จะเป็นการใช้ข้อสอบเป็นเครื่องมือหรือสถานการณ์สมมติ ซึ่งการสร้างข้อสอบแต่ละชุดก็เป็นการสุ่มตัวอย่างจากเนื้อหา(ไม่สามารถนำเนื้อหามาสอบได้ทั้งหมด) ถึงแม้จะใช้ข้อสอบชุดเดียวกันผู้ถูกวัดคนเดิม แต่วัดคนละเหตุการณ์ ก็จะได้ผลไม่ตรงกัน เนื่องจากธรรมชาติของบุคคลจะเกิดการเรียนรู้และพัฒนาตลอดเวลา การวัดผลครั้งที่สองย่อมได้ผลที่ดีกว่าครั้งแรกเสมอ

3)ผลการวัดมักเป็นค่าคะแนนสัมพัทธ์ ที่เปรียบเทียบกับผู้สอบอื่นๆในกลุ่มเดียวกัน หรือเปรียบเทียบระหว่างการสอบในแต่ละครั้ง และมีความคลาดเคลื่อนเสมอ ถึงแม้จะใช้เครื่องมือวัดที่มีมาตรฐานก็ตาม

ลักษณะการประเมินผลทางการศึกษาที่นิยมใช้มี 2 ลักษณะคือ 1)ประเมินผลเพื่อการพัฒนา (formative evaluation) เป็นการประเมินผลระหว่างการจัดการเรียนการสอน นิยมใช้เพื่อตรวจสอบการเรียนรู้และความก้าวหน้าของผู้เรียนหรือปรับปรุงคุณภาพการเรียนการสอน มักใช้แบบทดสอบ การสังเกต การซักถาม หรือเครื่องมือวัดอื่นๆที่เหมาะสม ระยะเวลามักทำเมื่อสิ้นสุดการเรียนการสอนเรื่องหนึ่งๆ 2)การประเมินผลสรุป (summative evaluation) เป็นการประเมินผลเมื่อสิ้นสุดการเรียนการสอนแล้ว มีวัตถุประสงค์เพื่อประเมินผลสัมฤทธิ์การเรียนรู้ของผู้เรียนมักทำปลายภาคการศึกษา และตัดสินผลการเรียน โดยมีเกณฑ์ตัดสินที่ชัดเจน เช่น การตัดสินแบบอิงกลุ่ม (เกรด A, B, C, D, F) การตัดสินแบบอิงเกณฑ์ (60 เปอร์เซนต์ สอบผ่าน) เป็นต้น โดยทั่วไปของการวัดสิ่งใดก็ตาม มักจะต้องกำหนดเป้าหมายหรือสิ่งที่วัดให้ชัดเจนว่าจะประเมินอะไรและประเมินอย่างไร จากนั้นจึงเลือกใช้เครื่องมือและเทคนิคที่สอดคล้องกับสิ่งที่ประเมิน หากไม่มีเครื่องมือที่เป็นมาตรฐาน มักนิยมสร้างขึ้นเองอย่างมีหลักการ และขั้นตอนสุดท้ายคือการนำวิธีการและเครื่องมือไปประเมินอย่างไม่มีอคติและยุติธรรม ผู้วัดควรตระหนักว่า การวัดผลจะมีความคลาดเคลื่อนหรือข้อผิดพลาดเสมอ

3.ลักษณะสำคัญของเครื่องมือวัดผลทางการศึกษา

แบบทดสอบ แบบวัด หรือเครื่องมือสำหรับในการวัดและประเมินผลทางการศึกษาที่ดี ควรมี 1)ความเที่ยงตรงสูง นั่นคือ สามารถวัดได้ตรงตามวัตถุประสงค์ หรือสิ่งที่ต้องการวัด 2)มีความเชื่อมั่นดี หมายถึง ผลที่วัดคงที่แน่นอนไม่เปลี่ยนแปลงไปมา วัดซ้ำก็ครั้งกับกลุ่มตัวอย่างเดิมได้ผลเท่าเดิม หรือใกล้เคียงกัน สอดคล้องกัน 3) มี

ความยากง่ายพอเหมาะ คือไม่ยากหรือง่ายเกินไป 4) มีอำนาจจำแนกได้ สามารถแบ่งแยกผู้สอบออกตามระดับความสามารถเก่งและอ่อนได้ โดยคนเก่งจะตอบข้อสอบได้ถูกมากกว่าคนอ่อน 5) มีประสิทธิภาพ หมายถึงทำให้ได้ข้อมูลที่ถูกต้องเชื่อถือได้ลงทุนน้อย 6) มีความยุติธรรม ไม่เปิดโอกาสให้ได้เปรียบเสียเปรียบ 7) มีความจำเพาะเจาะจง 8) ใช้คำถามที่ลึก และ 9) คำถามย่อย

1. ความเที่ยงตรง (validity)

ในการสร้างแบบทดสอบหรือเครื่องมือวิจัยสำหรับเก็บข้อมูล มักจะกล่าวถึงความเที่ยงตรง ซึ่งมักจะมี ความหมายและรายละเอียดดังนี้

1.1 ความเที่ยงตรงตามเนื้อหา (content validity) หมายถึง การวัดนั้นสามารถวัดได้ครอบคลุมเนื้อหา และวัดได้ครบถ้วนตามจุดประสงค์ของการวัด ในทางปฏิบัติมักจะต้องทำตารางจำแนกเนื้อหา จุดประสงค์ ตามที่ ต้องการก่อนจะทำการออกข้อสอบหรือแบบวัด

1.2 ความเที่ยงตรงเชิงสัมพันธ์ (criterion-related validity) แบ่งการออกเป็น 2 ลักษณะคือ 1) ความเที่ยงตรงเชิงพยากรณ์ (predictive validity) คือค่าคะแนนจากแบบสอบสามารถทำนายถึงผลการเรียนในวิชานั้น ๆ ได้อย่างเที่ยงตรง 2) ความเที่ยงตรงตามสภาพ (concurrent validity) หมายถึงค่าคะแนนที่ได้จากแบบสอบสะท้อน ผลตรงตามสภาพเป็นจริง กล่าวคือ เด็กเก่งจะได้คะแนนสอบสูง ส่วนเด็กอ่อนจะได้คะแนนต่ำจริง

1.3 ความเที่ยงตรงตามโครงสร้าง (construct validity) หมายถึงคะแนนจากแบบวัดมีความสอดคล้องกับ ลักษณะและพฤติกรรมจริงของเด็ก เช่น สอดคล้องกับความรู้ ความมีเหตุผล ความเป็นผู้นำ เชาวปัญญา เป็นต้น

2. ความเชื่อมั่น (reliability)

แบบทดสอบที่ดีต้องมีความเชื่อมั่นได้ว่าผลจากการวัดคงที่แน่นอน ไม่เปลี่ยนแปลงไปมา การวัดครั้งแรกเป็น อย่างไร เมื่อวัดซ้ำอีกโดยใช้แบบทดสอบชุดเดิมผู้ถูกทดสอบกลุ่มเดิม จะวัดกี่ครั้งก็ตามผลการวัดควรจะเหมือนเดิม หรือใกล้เคียงเดิม สอดคล้องกัน แบบทดสอบที่เชื่อมั่นได้จะสามารถให้คะแนนที่คงที่แน่นอน ปกติการสอบแต่ละ ครั้งคะแนนที่ได้มักไม่คงที่ แต่ถ้าอันดับของผู้ที่ทำข้อสอบยังคงที่เหมือนเดิมก็ยิ่งถือว่าแบบทดสอบนั้นมีความเชื่อมั่น สูง

ความเชื่อมั่นของแบบทดสอบ หมายถึงความคงที่ของคะแนนที่จากการสอบของคนกลุ่มเดิมหลาย ๆ ครั้ง การหาค่าความเชื่อมั่นได้จึงยึดหลักการสอบหลาย ๆ ครั้ง แล้วหาค่าความสัมพันธ์ของคะแนนที่ได้จากการสอบหลาย ๆ ครั้งนั้น ถ้าคะแนนของผู้สอบแต่ละคนคงที่หรือขึ้นลงตามกัน แสดงว่าแบบทดสอบนั้นมีค่าความเชื่อมั่นสูง (reliability) ค่าความเชื่อมั่นคำนวณได้จากการหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนทั้ง 2 ชุด จากการสอบ ผู้สอบกลุ่มเดิม 2 ครั้ง โดยใช้แบบทดสอบเดียวกัน ความเชื่อมั่นมีค่าอยู่ระหว่าง 0 ถึง 1.00 วิธีการหาค่าความ เชื่อมั่นของแบบทดสอบสามารถทำได้หลายวิธีดังต่อไปนี้

2.1 การสอบซ้ำ (test and retest)

เป็นการนำแบบทดสอบชุดเดียวกันไปสอบผู้เรียน กลุ่มเดียวกัน 2 ครั้ง ในเวลาห่างกันพอสมควร (ป้องกันการ จำข้อสอบได้) แล้วนำค่าคะแนนทั้ง 2 ชุดนั้น มาหาค่าความสัมพันธ์ที่ได้ คือค่าความเชื่อมั่นของแบบทดสอบ วิธีการนี้เรียกว่า “measure of stability” การหาค่าความเชื่อมั่นลักษณะนี้มีข้อจำกัดบางประการ กล่าวคือ 1) ผู้ทำ แบบทดสอบอาจเกิดความเบื่อหน่าย เพราะธรรมชาติของบุคคลไม่ชอบความซ้ำซากจำเจ 2) เสียเวลาในการสอบ มาก 3) ผู้สอบเกิดการเรียนรู้จากการสอบครั้งแรก ทำให้สอบครั้งหลังทำได้คล่องขึ้น เกิดความคลาดเคลื่อนจากความ เป็นจริง ดังนั้นการหาค่าความเชื่อมั่นของข้อสอบนี้จึงไม่เป็นที่นิยม

2.2 ใช้แบบทดสอบคู่ขนาน (parallel test หรือ equivalence tests)

แบบทดสอบคู่ขนานหมายถึง แบบทดสอบ 2 ชุด ที่มีลักษณะและคุณภาพใกล้เคียงกันมากที่สุด ทั้งด้าน เนื้อหา ความยากง่าย อำนาจจำแนก ลักษณะคำถาม และจำนวนข้อคำถาม จนอาจกล่าวได้ว่า เป็นแบบทดสอบฉบับ เดียวกัน สามารถใช้แทนกันได้ การใช้แบบทดสอบคู่ขนานนี้ เป็นการแก้ปัญหาข้อจำกัดต่างๆ ของการหาค่าความ เชื่อมั่นของแบบทดสอบโดยการสอบซ้ำ วิธีการหาค่าความเชื่อมั่นของแบบทดสอบทำได้โดย นำแบบทดสอบคู่ขนานไป

ทดสอบนักเรียน กลุ่มเดียวกันทั้ง 2 ฉบับ ในเวลาเดียวกัน แล้วนำคะแนนจากการทำแบบทดสอบทั้ง 2 ชุดนี้ มาหาความสัมพันธ์กัน ก็จะได้ค่าความเชื่อมั่นดังกล่าว วิธีการนี้เรียกว่า “measure of equivalence test” ข้อจำกัดของการทดสอบนี้คือ การสร้างแบบทดสอบที่มีลักษณะที่ใกล้เคียงกันทั้งเนื้อหา ความยากง่าย และอำนาจการจำแนก ฯลฯ ทำได้ยาก ต้องใช้ประสบการณ์สูง

2.3 วิธีแบ่งครึ่งข้อสอบ (split-half)

เป็นการสร้างข้อสอบชุดเดียวให้ผู้สอบชุดเดียวกัน (แต่แบ่งครึ่งข้อสอบ และได้ค่าคะแนน 2 ชุด) เป็นการแก้ปัญหาความยากในการสร้างแบบทดสอบแบบคู่ขนาน แต่ได้ผลเช่นเดียวกับการสอบซ้ำ หรือการใช้ข้อสอบแบบคู่ขนาน วิธีการอาจแบ่งตรวจสอบครึ่งละครึ่งฉบับ (แบ่งข้อคี่กับข้อคู่ หรือครึ่งแรกและครึ่งหลัง) นิยมใช้ข้อคู่และคี่มากกว่า เนื่องจากการเรียงลำดับข้อสอบนิยมเรียงตามเนื้อหาเป็นตอน ๆ จากง่ายไปยาก ดังนั้นการแบ่งครึ่งลักษณะนี้จึงมีลักษณะคล้ายคลึงพอลิโบลิมให้เป็นแบบทดสอบคู่ขนานได้ เมื่อตรวจและได้คะแนน 2 ชุดแล้ว นำคะแนนทั้งสองมาหาค่าความสัมพันธ์กัน เป็นค่าความเชื่อมั่นของแบบทดสอบครึ่งฉบับ เรียกว่า “internal consistency” จากนั้นจึงนำมาคำนวณค่าความเชื่อมั่นทั้งฉบับอีกครั้งหนึ่ง ซึ่งค่าความเชื่อมั่นข้อสอบเต็มฉบับนั้นจะสูงกว่าครึ่งฉบับ เนื่องจาก ค่าความเชื่อมั่นขึ้นกับความยาวหรือจำนวนข้อของคำถาม ข้อสอบที่มีข้อคำถามมากจะมีระดับความเชื่อมั่นสูงกว่าแบบสอบที่มีจำนวนข้อสอบน้อย

2.4 วิธี Kuder-Richardson (KR)

เป็นการหาค่าความคงที่ภายในของแบบทดสอบ เรียกว่า ความเชื่อมั่นภายใน (internal consistency) สูตรที่นิยมใช้คือสูตรคำนวณ KR-20 และ KR-21 รายละเอียดการคำนวณไม่ขอกล่าวในที่นี้

3. ความเป็นปรนัย (objectivity)

ความเป็นปรนัยหมายถึง ความชัดเจน ความถูกต้อง ความเข้าใจตรงกัน โดยยึดถือความถูกต้องทางวิชาการเป็นเกณฑ์ การสร้างแบบทดสอบใด ๆ จำเป็นต้องมีความชัดเจนเข้าใจตรงกันระหว่างผู้ออกข้อสอบและผู้ทำข้อสอบ คุณสมบัติความเป็นปรนัยของแบบทดสอบพิจารณาได้เป็น 3 ประการ คือ 1) ผู้อ่านข้อสอบทุกคนเข้าใจตรงกัน 2) ผู้ตรวจทุกคนให้คะแนนได้ตรงกัน 3) แปลความหมายของคะแนนได้ตรงกัน

4. ความยากง่าย (difficulty)

แบบทดสอบที่ดีต้องมีความยากง่ายพอเหมาะ คือไม่ยากเกินไปหรือไม่ง่ายเกินไป ในแบบทดสอบชุดหนึ่ง ๆ อาจมีทั้งข้อสอบที่ค่อนข้างยาก ปานกลาง และค่อนข้างง่ายปะปนกันไป ความยากง่ายของแบบทดสอบพิจารณาได้จากผลการสอบของข้อสอบทั้งฉบับเป็นสำคัญ ความยากง่ายพิจารณาได้จาก 1) ความยากง่ายของแบบทดสอบทั้งฉบับ และ 2) ความยากง่ายข้อคำถามรายข้อ

4.1. ความยากง่ายของแบบสอบทั้งฉบับ พิจารณาจาก 1) คะแนนรวมของแบบทดสอบทั้งฉบับ หากคะแนนเฉลี่ยสูงกว่าครึ่งหนึ่งของคะแนนเต็มแสดงว่าง่าย ถ้าต่ำกว่าครึ่งของคะแนนเต็มถือว่ายากหรือค่อนข้างยาก ทั้งนี้พิจารณาร่วมกับจุดประสงค์และเกณฑ์ของการสอบด้วย 2) พิจารณาจากความยากง่ายของข้อคำถามรายข้อ โดยนำความยากง่ายรายข้อมาคำนวณหาค่าเฉลี่ย (ค่าเฉลี่ยความยากรายข้อรวมทั้งฉบับ) ซึ่งมีค่าอยู่ระหว่าง 0-1.00 หากค่าเฉลี่ยความยากง่ายรายข้อทั้งฉบับสูงกว่า 0.50 ถือว่าข้อสอบนั้นง่ายหรือค่อนข้างง่าย หากมีค่าต่ำกว่า 0.50 ถือว่าแบบทดสอบนั้นค่อนข้างยาก

4.2 ความยากง่ายรายข้อ พิจารณาจากจำนวนผู้ตอบถูกในข้อสอบแต่ละข้อ ถ้าข้อใดผู้ตอบถูกมากกว่าครึ่งหนึ่งของผู้สอบทั้งหมดถือว่าข้อสอบง่าย ในทำนองเดียวกันข้ามหากตอบถูกน้อยกว่าครึ่งหนึ่งของผู้สอบถือว่ายากหรือค่อนข้างยาก ดังนั้นค่าความยากง่ายของข้อสอบ หมายถึงสัดส่วนของผู้ที่ตอบข้อคำถามถูกนั่นเอง ซึ่งนิยมแทนด้วย “p” มีค่าตั้งแต่ 0-1.00 ถ้าค่า p สูงแสดงว่าคำถามข้อนั้นมีผู้ตอบถูกมาก แสดงว่าข้อสอบนั้นง่าย ถ้าค่า p ต่ำแสดงว่าคำถามข้อนั้นมีผู้ตอบถูกน้อย แสดงว่าข้อสอบยาก แบบทดสอบที่ดี ควรมีค่าความยากง่ายพอเหมาะ ค่า p

อยู่ระหว่าง 0.20–0.80 อย่างไรก็ตามการพิจารณาความยากง่ายของข้อสอบ อาจแตกต่างกันตามจุดประสงค์ของการสอบ เช่น แบบทดสอบสำหรับประเมินผลการเรียนการสอนทั่วไป อาจใช้แบบสอบที่มีค่าความยากง่าย 0.20–0.80 แต่แบบทดสอบเพื่อคัดเลือกเรียนต่ออาจใช้แบบสอบที่มีค่าความยากง่าย 0.15–0.75 ส่วนแบบสอบสำหรับการประเมินแบบอิงเกณฑ์อาจใช้แบบทดสอบที่ง่ายอีกเนื่องจากการประเมินผลแบบอิงเกณฑ์มักใช้เกณฑ์ในการประเมิน 80–90 เปอร์เซนต์

5.อำนาจจำแนก (discrimination)

แบบทดสอบที่ดีต้องสามารถจำแนกผู้สอบที่มีความสามารถเก่งอ่อนต่างกันออกได้ โดยคนเก่งจะตอบข้อสอบถูกมากกว่าคนอ่อน โดยพิจารณาจาก

5.1) ค่าอำนาจจำแนกแบบทดสอบทั้งฉบับ กล่าวคือ 1) หากคะแนนรวมของผู้ทำข้อสอบทั้งกลุ่ม มีการกระจายตัวตั้งแต่ศูนย์ถึงเกือบเต็ม แสดงว่าแบบทดสอบนั้นจำแนกได้ แต่ถ้าคะแนนรวมมีการเกาะกลุ่มกันหรือมีการกระจายตัวของคะแนนน้อย แสดงว่าแบบทดสอบนั้นไม่มีอำนาจการจำแนกต่ำหรือจำแนกไม่ได้นั่นเอง 2) ค่าเฉลี่ยของค่าอำนาจจำแนกรายข้อของแบบทดสอบทั้งฉบับ โดยทั่วไปค่าอำนาจจำแนกรายข้อ (r) ของแบบทดสอบมีค่าระหว่าง -1.00 – $+1.00$ หากค่าเฉลี่ยของค่าอำนาจจำแนกรายข้อเท่ากับหรือมากกว่า 0.20 แสดงว่าแบบทดสอบฉบับนั้นจำแนกได้

5.2) ค่าอำนาจจำแนกของแบบทดสอบรายข้อ ซึ่งพิจารณาจากหลักที่ว่าคนเก่งย่อมตอบถูกมากกว่าคนอ่อน ดังนั้นหากข้อใดที่มีคนอ่อนตอบถูกมากกว่าเรียกว่า “จำแนกกลับ” ส่วนข้อใดหากคนเก่งและคนอ่อนตอบถูกพอๆกันเรียกว่า “จำแนกไม่ได้” ดังนั้น หากข้อสอบใด ค่าอำนาจจำแนก (r) มีค่าบวกแสดงว่าจำแนกได้ มีค่าลบแสดงว่าจำแนกคนกลับ และเป็นศูนย์หรือใกล้ศูนย์แสดงว่าข้อคำถามนั้นไม่มีอำนาจจำแนกคนเก่งและคนอ่อนตอบผิดถูกพอๆกันควรมีการปรับปรุงก่อนนำไปใช้

6.ความมีประสิทธิภาพ (efficiency)

เครื่องมือวัดที่มีประสิทธิภาพ หมายถึงเครื่องมือที่ทำให้ข้อมูลที่ต้องการ เชื่อถือได้ ประโยชน์สูงสุดประหยัด โดยลงทุนน้อยที่สุด ไม่ว่าจะเป็นด้านเวลา แรงงาน ความสะดวกสบาย แบบทดสอบที่ดี ควรพิมพ์ผิดพลาดกหล่นน้อย รูปแบบดูง่าย เป็นระเบียบเรียบร้อยอ่านง่าย

7.ความยุติธรรม (fair)

แบบทดสอบที่ดีต้องไม่เปิดโอกาสให้ผู้สอบได้เปรียบเสียเปรียบกัน เช่น แบบทดสอบบางฉบับผู้สอบออกข้อสอบเน้นเรื่องใดเรื่องหนึ่งให้ผู้เรียนบางคนได้เคยค้นคว้าทำรายงานมาก่อน เป็นต้น ดังนั้นผู้ออกข้อสอบควรคำนึงถึงข้อได้เปรียบเสียเปรียบของผู้ทำแบบทดสอบด้วย

8.คำถามลึกลับ (searching)

แบบทดสอบที่สอบถามเฉพาะความรู้ความเข้าใจ ผู้ออกข้อสอบไม่ควรถามลึกลับจนกระทั่งต้องใช้ความรู้ระดับวิเคราะห์ สังเคราะห์ แก้ปัญหา ดังนั้น ความลึกลับซึ่งของคำถามควรสอดคล้องกับลักษณะและจุดประสงค์ของการวัด

9. คำถามยั่ว (exasperation)

คำถามยั่วมีลักษณะเป็นคำถามที่ท้าทายให้ผู้สอบอยากคิดอยากทำมีลีลาการถามที่น่าสนใจ ไม่ถามวกวนซ้ำซากน่าเบื่อ อาจใช้รูปประกอบคำถาม การเรียงข้อความในข้อสอบ ควรเรียงหลายแบบคละกัน อาจเรียงลำดับเนื้อหา เรียงลำดับความยากง่าย สลับกัน เป็นต้น

10 ความจำเพาะเจาะจง (definite)

ลักษณะคำถามที่ดีไม่ควรถามกว้างเกินไป ไม่ถามคลุมเครือหรือเล่นสำนวนจนผู้สอบง ผู้อ่านอ่านแล้วต้องมีความชัดเจนว่าคำถามอะไร ส่วนจะตอบถูกหรือไม่เป็นอีกเรื่อง ขึ้นกับความรู้ความสามารถของผู้ตอบ

4. เครื่องมือที่ใช้วัดและประเมินผลด้านการศึกษา

ในการวัดและประเมินผลด้านการศึกษานั้น จะใช้เครื่องมือโดยขึ้นอยู่กับลักษณะจุดประสงค์การศึกษา และแนวทางการจัดประสบการณ์การเรียนรู้ เครื่องมือที่ใช้วัดและประเมินผลการเรียนรู้มีหลายลักษณะ แต่ละชนิดมีทั้งข้อดีและข้อจำกัด พอสรุปได้ดังต่อไปนี้

4.1. แบบทดสอบ

แบบทดสอบคือชุดของคำถามหรือสิ่งเร้าที่นำไปใช้ให้ผู้สอบตอบสนองออกมา ชุดของสิ่งเร้านี้มักอยู่ในรูปของข้อความ ซึ่งอาจให้เขียนตอบ แสดงพฤติกรรม ให้พูดออกทางวาจาก็ได้ ทำให้สามารถวัดได้ สังเกตได้ และนำไปสู่การแปลความหมายได้ แบบทดสอบนี้สามารถใช้ได้กับการวัดพฤติกรรมด้านพุทธิพิสัย จิตพิสัย และทักษะพิสัย แต่ส่วนใหญ่นิยมวัดทางด้านพุทธิพิสัย แบบทดสอบแบ่งตามจุดมุ่งหมายออกเป็น 3 ชนิด ดังนี้

1) แบบวัดผลสัมฤทธิ์ทางการเรียน (achievement test) เป็นแบบทดสอบที่ใช้วัดความรู้ ทักษะ และความสามารถสมอง ด้านต่าง ๆ เช่น ความรู้ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ สังเคราะห์ และการประมาณค่า ซึ่งแบบวัดผลสัมฤทธิ์ทางการเรียนนี้อาจเป็นประเภทที่ผู้สอนสร้างขึ้นเอง เช่น ข้อสอบปลายภาค หรือเป็นแบบทดสอบมาตรฐาน ที่มีผู้สร้างไว้แล้ว เช่น ข้อสอบ TOFEL รูปแบบและวิธีการใช้แบบทดสอบแบ่งเป็น 3 ลักษณะคือ (1) แบบสอบปากเปล่า (oral test) เป็นการทดสอบที่อาศัยการซักถามเป็นรายบุคคล เหมาะสำหรับผู้สอบจำนวนน้อย ข้อดีคือ สามารถถามได้ละเอียด และสามารถโต้ตอบได้ (2) แบบเขียนตอบ (paper-pencil test) เป็นการทดสอบที่มีการเขียนตอบ แบ่งออกเป็น 2 ประเภท คือ แบบทดสอบอัตนัย หมายถึงแบบทดสอบที่ถามให้ตอบยาว ๆ สามารถแสดงความคิดเห็นได้อย่างกว้างขวาง เหมาะสำหรับการวัดความสามารถในการใช้ภาษาและแสดงความคิดเห็นที่หลากหลาย และแบบทดสอบปรนัย หมายถึงแบบทดสอบประเภท ถูก-ผิด จับคู่ เติมคำ และเลือกตอบ เหมาะสำหรับผู้สอบจำนวนมาก ทุริเวลาตรวจข้อสอบน้อย (3) แบบปฏิบัติ (performance test) เป็นการทดสอบที่ผู้สอบได้แสดงพฤติกรรมออกมาโดยการกระทำหรือลงมือปฏิบัติจริง เช่น การสอบนวด การสอบปฏิบัติทางกายภาพบำบัด เป็นต้น

2) แบบทดสอบวัดความถนัดหรือทักษะ (aptitude test) เป็นแบบทดสอบที่ใช้วัดศักยภาพระดับสูง ของบุคคลว่า สมรรถภาพในการเรียนรู้มีมากน้อยเพียงใด และควรเรียนด้านใดหรือทำงานในด้านใด จึงจะเหมาะสมและประสบความสำเร็จ แบบทดสอบประเภทนี้แบ่งย่อยได้เป็น 2 ประเภท คือ แบบทดสอบความถนัดในการเรียน (scholastic aptitude test) และแบบทดสอบความถนัดจำเพาะ (specific test) ซึ่งแบ่งความถนัดเป็น 7 ด้านได้แก่ ด้านภาษา การใช้คำ ตัวเลข มิติสัมพันธ์ ความจำ การสังเกตรับรู้ และการใช้เหตุผล

3) แบบทดสอบวัดความสัมพันธ์ของบุคคล เป็นแบบทดสอบที่ใช้วัดเกี่ยวกับบุคลิกภาพหรือการปรับตนเองของบุคคลในสังคม วัดความสนใจต่อสิ่งต่าง ๆ ในรูปแบบทดสอบวัดลักษณะบุคคล เช่น แบบทดสอบความเกรงใจ แบบทดสอบความคิดสร้างสรรค์ เป็นต้น

4.2 แบบสอบถามแบบมาตราส่วนประมาณค่า

แบบสอบถามเป็นชุดของคำถามที่เกี่ยวกับเรื่องใดเรื่องหนึ่ง สร้างขึ้นเพื่อตรวจสอบ ข้อเท็จจริง ความคิดเห็น ความรู้สึก ความเชื่อ และความสนใจต่าง ๆ ในทางการศึกษามักนิยมใช้วัดและประเมินผลด้านจิตพิสัยได้แก่ มาตราส่วนประมาณค่า (rating scale) เป็นเครื่องมือที่ใช้ได้ทั้งให้ผู้ถูกวัดประเมินตนเอง และผู้อื่นประเมิน การตอบกระทำโดยให้ผู้ตอบหรือผู้สังเกตประเมินค่าของคุณลักษณะออกมาเป็นระดับต่าง ๆ มากน้อยตามปริมาณหรือความเข้มของความรู้สึกหรือพฤติกรรมที่แสดงออก มาตราส่วนประมาณค่ามีหลายลักษณะ ที่นิยมใช้และสร้างได้ง่ายคือ มาตราส่วนประมาณค่าของลิเคิร์ต (Likert rating scale) และมาตราส่วนประมาณค่าแบบซีแมนติก ดิฟเฟอเรนเชียล (Semantic differential rating scale)

1)มาตราส่วนประมาณค่าของลิเคิร์ท (Likert rating scale)

ลักษณะของมาตราส่วนประมาณค่าของลิเคิร์ท ประกอบด้วยส่วนสำคัญ 2 ส่วน คือ 1)ส่วนที่เป็นข้อความ คำถามหรือสถานการณ์ที่กำหนด มักจะมีลักษณะสอบถามความคิดเห็น ความรู้สึก หรือทัศนคติ หรือพฤติกรรมที่แสดงออกของบุคคลที่ต้องการวัด ซึ่งข้อความดังกล่าวอาจเป็นได้ทั้งทางบวกหรือทางลบ เช่น ภายนอกภาพบำบัดมีประโยชน์สำหรับประชาชนระดับใด ขณะเดินท่านมีความรู้สึกปวดระดับใด เป็นต้น 2)ส่วนที่เป็นคำตอบ มักเป็นการกำหนดค่าระดับความรู้สึก ความคิดเห็น ทัศนคติ หรือพฤติกรรมที่แสดงออก แบบลิเคิร์ทจะกำหนดคำตอบเป็น 5 ระดับ ในระดับสนับสนุนถึงระดับไม่สนับสนุน ระดับเห็นด้วยถึงระดับไม่เห็นด้วย เป็นต้น ดังตารางที่ 2 ซึ่งมาตราส่วนลิเคิร์ท อาจใช้ตัวเลขแสดงการจำแนกระดับพฤติกรรมได้ 5 ระดับคือ 5, 4, 3, 2, 1 โดยให้ 5 หมายถึงความรู้สึกหรือพฤติกรรมที่แสดงออกมาในระดับมาก และลดลงเรื่อยๆจนถึงระดับ 1 ซึ่งหมายถึงระดับน้อย (ตารางที่ 15)

ตารางที่ 15 แสดงตัวอย่างมาตราส่วนประมาณค่าของลิเคิร์ท 5 ระดับ

ข้อความ	ระดับพฤติกรรม				
	มาก	ค่อนข้างมาก	ปานกลาง	ค่อนข้างน้อย	น้อย
1.ภายนอกภาพบำบัดมีประโยชน์สำหรับประชาชนระดับใด					
2. ขณะเดินท่านมีความรู้สึกปวดระดับใด					
ฯลฯ					

แบบวัดมาตราส่วนประมาณค่าแบบลิเคิร์ทที่ดีควรมีลักษณะดังต่อไปนี้ 1)ข้อความที่สอบถามพฤติกรรมไม่ควรเป็นข้อความที่เกี่ยวข้องสำหรับบุคคลทั่วไป หรือเป็นหลักวิชาการมากเกินไป เพราะผู้ตอบจะไม่ใช้ความรู้สึกนึกคิดของตนเองในการตอบ แต่จะใช้ความจริง หรือสามัญสำนึกของคนทั่วไปตอบ ตัวอย่างเช่น การไม่รับผิดชอบผู้ป่วยของตนเองจัดเป็นการกระทำที่ถูกต้องระดับใด 2)ข้อความที่สอบถามไม่ควรยาวเกินไป 3)ควรใช้ภาษาที่ชัดเจน เข้าใจง่าย ไม่ต้องตีความ 4)หลีกเลี่ยงข้อความที่เป็นปฏิเสธ เพราะอาจทำให้เกิดการตีความหมายผิดไป 5)ไม่ควรใช้ข้อความที่ทุกคนตอบตรงกัน เพราะไม่มีประโยชน์ในการสอบถาม 6)มักนิยมเขียนส่วนที่เป็นคำตอบ 5 ระดับ เป็น 2 ลักษณะคือ ตัวเลข และข้อความ

2)มาตราส่วนประมาณค่าแบบซีแมนติก ดิฟเฟอเรนเชียล (Semantic differential rating scale)

มาตราส่วนประมาณค่าแบบซีแมนติก ดิฟเฟอเรนเชียล มีองค์ประกอบ เป็น 2 ส่วนสำคัญคือ 1) ข้อความแสดงความรู้สึก คิดเห็น หรือพฤติกรรมของการแสดงออกของบุคคลที่ต้องการวัด ซึ่งเป็นเป้าหมายของการวัด 2) คำคุณศัพท์หรือตัวเลขที่แสดงออกระดับของความรู้สึก หรือพฤติกรรมที่ต้องการวัด โดยทั่วไปคำคุณศัพท์ที่ใช้ในแต่ละข้อจะมีลักษณะเป็นความหมายตรงข้ามกัน โดยมีคำหรือตัวเลขแสดงระดับพฤติกรรมตั้งแต่ระดับต่ำสุดไปถึงสูงสุด เช่น ยาก-ง่าย ชอบ-เกลียด ดี-เลว เป็นต้น ดังตารางที่ 16

ตารางที่ 16 แสดงตัวอย่างคำถามมาตราส่วนประมาณค่าแบบซีแมนติก โดยตั้งคำถาม ...”ท่านเรียนวิชาการจัดการเรียนการสอนทางกายภาพบำบัดแล้วรู้สึก”

นำสนใจ	3	2	1	0	1	2	3	นำเบื่อ
มีคุณค่า	มาก	ปานกลาง	น้อย	ตัดสินใจไม่ได้	น้อย	ปานกลาง	มาก	ไร้สาระ

ในการตรวจให้คะแนนนั้น มีความลดหล่นกันตั้งแต่ระดับ 1-7 โดยให้ 1 คะแนน สำหรับผู้ตอบ 3 ในด้านลบ และให้ 7 คะแนน สำหรับผู้ตอบ 3 ในด้านบวก ผู้ที่ได้คะแนนเฉลี่ยมาก มาทาง 7 แสดงว่ามีเจตคติที่ดี ต่อเรื่องนั้น ๆ เป็นต้น

4.3 แบบสำรวจรายการ

แบบสำรวจรายการมีลักษณะคล้ายมาตราส่วนประมาณค่าของลิเคิร์ท เพียงแต่ส่วนที่เป็นคำตอบไม่ได้กำหนดระดับความรู้สึกว่ามีมากน้อยเพียงไร แต่เป็นการตอบเพียง 2 ตัวเลือกว่า มี-ไม่มี, ใช่-ไม่ใช่, เคย-ไม่เคย ฯลฯ เป็นต้น

4.4 แบบวัดเชิงสถานการณ์

เป็นแบบวัดที่สอบถามถึงแนวคิด ความรู้สึก หากอยู่ในสถานการณ์สมมติใด ๆ ที่ผู้ออกข้อสอบสร้างขึ้น อาจบรรจุไว้ในส่วนที่เป็นข้อความ หรือส่วนที่เป็นคำตอบก็ได้ อาจนำเสนอเหตุการณ์ด้วยข้อความ หรือรูปภาพ หรือสื่ออื่น ๆ ก็ได้

ตัวอย่างเช่น หากท่านพบผู้ป่วยที่เดินมาพบท่านด้วยลักษณะตัวเอียงอย่างมาก ท่านจะดำเนินการอย่างไร เป็นลำดับแรก

- ก. สอบถามชื่อที่อยู่
- ข. สอบถามอาการปวด
- ค. ให้การรักษาทางกายภาพบำบัด
- ง. รีบให้ผู้ป่วยนอนบนเตียง

4.5 แบบสังเกต

การสังเกตเป็นเครื่องมือวัดผลที่นิยมใช้กันมาก โดยการใช้ประสาทสัมผัสของผู้ที่มีส่วนเกี่ยวข้องกับการประเมินได้สังเกตพฤติกรรมที่สนใจในตัวผู้ถูกวัด ผลการสังเกตจะมีความเที่ยงตรงเพียงใดขึ้นกับองค์ประกอบ 3 ประการ คือ 1) สิ่งที่สังเกตควรเป็นรูปธรรมและสังเกตได้จริง เช่น ผลการรักษาที่ดี มุมการเคลื่อนไหวของข้อที่เพิ่มขึ้น แต่บางสิ่งบางอย่างสังเกตได้ยาก เช่น อารมณ์พอใจ ความรู้สึกเจ็บปวด ทักษะคิด เป็นต้น 2) ตัวผู้สังเกต มีความตั้งใจ สามารถสังเกตได้อย่างมีประสิทธิภาพ มีประสบการณ์ในการสังเกตตีความหมายตรวจสอบเรื่องนั้น ๆ ได้เป็นอย่างดี และสามารถสังเกตรับรู้ได้โดยไม่มีอคติหรือความลำเอียง 3) ตัวผู้ถูกสังเกต ต้องไม่รู้ว่ากำลังถูกสังเกต หรือเฝ้ามอง เพราะหากรู้ว่ากำลังถูกสังเกตหรือเฝ้ามอง ก็จะไม่เป็นไปตามธรรมชาติ อาจมีพฤติกรรมเสแสร้ง เช่น แกล้งเป็นคนดีเรียบร้อยผิดธรรมชาติของตนเอง

การสังเกตที่ดีควรแบ่งเป็นช่วง ๆ ไม่นานเกินไป และไม่ควรติดต่อกันเป็นระยะยาวนาน ควรมีแบบฟอร์มสำหรับการสังเกตไว้ล่วงหน้า เพื่อเป็นแนวทางและสำหรับการบันทึกพฤติกรรมต่าง ๆ ที่สังเกตได้ หลักการสังเกตที่ดีพอสรุปได้ดังนี้ 1) สังเกตเฉพาะเรื่องที่กำลังสนใจเท่านั้น 2) สังเกตอย่างมีจุดมุ่งหมาย 3) ควรพินิจพิเคราะห์จนเข้าใจในรายละเอียด 4) มีการจดบันทึกสิ่งที่ตรงกับจุดมุ่งหมายของการสังเกตไว้อย่างครบถ้วน 5) ข้อมูลที่ได้จากการสังเกต ควรตรวจสอบจนมั่นใจ อาจตรวจสอบกับผู้สังเกตอื่น ๆ หรือหลักฐานอื่น ๆ ประกอบยืนยันตรงกัน

5. การวัดด้านพุทธิพิสัย

พุทธิพิสัย เป็นพฤติกรรมด้านความสามารถทางสติปัญญาของบุคคล ลักษณะการวัดจึงเป็นการวัดทางอ้อมจากพฤติกรรมที่แสดงออกของผู้ถูกวัดภายใต้เครื่องมือวัดหรือสถานการณ์ที่ผู้สอบกำหนด ซึ่งมีการจำแนกความสามารถออกเป็น 6 ระดับ เรียงจากความสามารถขั้นต่ำไปสูงดังนี้ 1) ความรู้ความจำ 2) ความเข้าใจ 3) การนำไปใช้ 4) การวิเคราะห์ 5) การสังเคราะห์ 6) การประเมินค่า

1) **ความรู้ความจำ** คือความสามารถในการระลึกได้ถึงเรื่องราวต่างๆที่เคยมีประสบการณ์มาก่อนจะด้วยวิธีการใดก็ตาม เช่น จากการเรียนรู้ในห้องเรียน การอ่านหนังสือ การบอกเล่าต่อกันมา เป็นต้น พฤติกรรมด้านความรู้ความจำยังจำแนกเป็น 3 ลักษณะใหญ่ๆคือ 1) **ความรู้เฉพาะเรื่อง** เป็นความรู้เกี่ยวกับศัพท์และนิยาม ความรู้เกี่ยวกับกฎและความจริงเฉพาะเรื่อง 2) **ความรู้ในการดำเนินการ** เป็นความรู้เกี่ยวกับระเบียบแบบแผน เกี่ยวกับแนวโน้มน้ำและลำดับขั้น การจำแนกประเภท ความรู้เรื่องกฎเกณฑ์ และระเบียบวิธี และ 3) **ความรู้รวบยอดในเรื่อง** เป็นความรู้เกี่ยวกับหลักวิชา ทฤษฎีและโครงสร้าง เป็นต้น

2) **ความเข้าใจ** คือความสามารถในการผสมผสานความรู้ความจำ แล้วขยายความคิดของตนเองออกไปอย่างสมเหตุสมผลสามารถอธิบายความโดยใช้ความคิดและคำพูดของตนเอง ดังนั้นความเข้าใจจึงเป็นความรู้ขั้นสติปัญญา (cognition) เนื่องจากต้องนำเอาความรู้จากขั้นความรู้ความจำ มาผสมผสานร่วมด้วย ความเข้าใจแบ่งเป็น 3 ลักษณะได้แก่ 1) **การแปลความ** เป็นความสามารถในการสื่อความหมายจากภาษาหนึ่งหรือแบบฟอร์มหนึ่งไปสู่อีกภาษาหนึ่งหรืออีกแบบฟอร์มหนึ่ง 2) **การตีความ** เป็นความสามารถเอาผลการแปลความหลายๆสิ่งมาผสมผสานเรียบเรียงเป็นความคิดใหม่อย่างมีความหมาย และ 3) **การขยายความ** เป็นการขยายแนวคิดให้กว้างไกลไปจากข้อมูลเดิมอย่างสมเหตุสมผล ซึ่งต้องอาศัยการแปลความและการตีความประกอบกัน เป็นต้น

3) **การนำไปใช้** คือความสามารถนำความรู้ความเข้าใจในเรื่องที่เรียนรู้มาแล้วไปใช้ที่แปลกใหม่ หรือสถานการณ์ใหม่ที่ไม่เคยเห็นมาก่อนแต่อาจใกล้เคียงหรือคล้ายคลึงกับที่เคยมีประสบการณ์มาก่อน

4) **การวิเคราะห์** คือความสามารถแยกแยะเรื่องราวสิ่งต่างๆ ออกเป็นส่วนย่อยๆได้ ทำให้สามารถมองเห็นความสัมพันธ์กันของส่วนย่อยได้อย่างชัดเจน สามารถค้นหาความจริงต่างๆ ที่ซ่อนแฝงอยู่ในเนื้อเรื่องนั้นๆ ได้ การวิเคราะห์แบ่งเป็น 3 ลักษณะได้แก่ 1) **วิเคราะห์ความสำคัญ** เป็นความสามารถในการแยกแยะองค์ประกอบย่อยที่รวมอยู่ในเรื่องราวหนึ่งเพื่อชี้ให้เห็นถึงเหตุ ผลลัพธ์ และประเด็นสำคัญ 2) **วิเคราะห์ความสัมพันธ์** เป็นความสามารถในการพิจารณาหาความสัมพันธ์ระหว่างองค์ประกอบย่อยที่รวมกันอยู่ในเรื่องราวหนึ่งๆว่ามีความสัมพันธ์ เกี่ยวพันกันใดในลักษณะใด เหมือนหรือต่างกันอย่างไร และ 3) **วิเคราะห์หลักการ** เป็นความสามารถในการค้นหาว่า โครงสร้างและระบบของวัตถุสิ่งของ เรื่องราวและการกระทำต่างๆรวมกันอยู่ในสภาพเช่นนั้นได้ เพราะยึดหลักหรือแกนอะไร เป็นต้น

5) **การสังเคราะห์** คือการนำองค์ประกอบย่อยๆต่างๆ ตั้งแต่ 2 สิ่งขึ้นไปมารวมเข้าเป็นเรื่องราวเดียวกันเพื่อให้เห็นถึงโครงสร้างที่ชัดเจน แยกไปจากเดิม แบ่งเป็น 3 ประเภท 1) **สังเคราะห์ข้อความ** เป็นความสามารถในการผสมผสานความรู้และประสบการณ์ต่าง ทำให้เกิดเป็นข้อความหรือผลผลิตใหม่ อาจสังเคราะห์โดยการพูด เขียนหรือรูปภาพก็ได้ 2) **สังเคราะห์แผนงาน** เป็นความสามารถในการกำหนดแนวทาง วางแผน เขียนโครงการต่างๆ ล่วงหน้าขึ้นมาใหม่ ให้สอดคล้องกับข้อมูลและจุดมุ่งหมายที่วางไว้ 3) **สังเคราะห์ความสัมพันธ์** เป็นความสามารถในการนำความสำคัญหลักการต่างๆ มาผสมผสานให้เป็นเรื่องเดียวกันทำให้เกิดเป็นสิ่งที่สำเร็จอันใหม่ที่มีความสัมพันธ์แตกต่างไปจากเดิม เป็นเรื่องราวหรือแนวคิดใหม่ที่ผิดหรือแตกต่างไปจากเดิม

6) **การประเมินค่า** คือความสามารถในการตัดสินเกี่ยวกับคุณค่าของเนื้อหาและวิธีการต่างๆ โดยสรุปอย่างมีหลักเกณฑ์ว่า เหมาะสม มีคุณค่า ดี-เลว เพียงไร การประเมินค่าต้องอาศัยเกณฑ์ประกอบการตัดสินใจ 2 ลักษณะคือ 1) **การตัดสินโดยใช้เท็จจริง** หรือเกณฑ์ภายในเนื้อเรื่องนั้น ซึ่งเป็นการตัดสินโดยใช้ข้อเท็จจริงที่ปรากฏในเนื้อเรื่องนั้นๆ 2) **การตัดสินโดยใช้เกณฑ์ภายนอก** เป็นการตัดสินตามเหตุผลทางตรรกศาสตร์ การยอมรับของสังคม สภาพความเป็นจริง ความยุติธรรม เป็นต้น

5.1. การวัดพฤติกรรมพุทธิพิสัย

การวัดด้านพุทธิพิสัย เป็นการวัดความสามารถด้านสติปัญญา ประกอบด้วย ความสามารถด้าน ความรู้ ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินค่า เครื่องมือที่ใช้วัดพฤติกรรมนี้ มักใช้เป็นแบบทดสอบ มักนิยมวัดเนื้อหาสาระ ซึ่งแบบทดสอบโดยทั่วไปแบ่งออกเป็น 2 ประเภท คือ 1) **แบบทดสอบ**

อัตรณ์ หมายถึงแบบทดสอบที่ถามให้ตอบยาว ๆ สามารถแสดงความคิดเห็นได้อย่างกว้างขวาง เหมาะสำหรับการวัดความสามารถในการใช้ภาษาและแสดงความคิดเห็นที่หลากหลาย คำตอบที่ได้มักไม่มีผิดถูกอย่างสมบูรณ์ และ 2) แบบทดสอบปรนัย หมายถึงแบบทดสอบประเภท ถูก-ผิด จับคู่ เติมคำ และเลือกตอบ เหมาะสำหรับสอบผู้สอบจำนวนมาก ๆ มีเวลาตรวจข้อสอบน้อย

5.2 หลักการประเมินพุทธิพิสัย

การประเมินพุทธิพิสัย เป็นการประเมินความรู้ความเข้าใจ ซึ่งเป็นทักษะของสมอง ดังนั้นการประเมินจึงเป็นการประเมินทางอ้อมจากพฤติกรรมที่ผู้ถูกประเมินแสดงออก เพื่อบ่งชี้ถึงทักษะของสมองดังกล่าว ลักษณะของแบบทดสอบวัดพฤติกรรมด้านพุทธิพิสัยที่ดี คือ 1) มีความเที่ยงตรง ที่มีความเที่ยงตรงจะสามารถวัดเนื้อหาที่ต้องการวัดได้ครบถ้วน และวัดได้ตรงตามจุดมุ่งหมายที่ต้องการวัด 2) เชื่อมั่นได้ แบบทดสอบที่มีความเชื่อมั่นได้ หากนำมาใช้สอบวัดกับคนกลุ่มเดิมหลาย ๆ ครั้ง ผลการวัดจะเหมือนเดิมหรือเปลี่ยนแปลงน้อยมาก 3) มีความเป็นปรนัย หมายถึงมีความชัดเจนของคำถามทุกคนอ่านแล้วเข้าใจตรงกัน ใครตรวจก็ให้คะแนนตรงกัน และแปลความหมายของคะแนนได้ค่าตรงกัน 4) มีความยากง่ายพอเหมาะ หมายถึงแบบทดสอบไม่ยากเกินไปและไม่ง่ายเกินไป ข้อสอบข้อใดที่มีคนตอบถูกมากแสดงว่าง่าย และข้อสอบที่มีคนตอบถูกน้อยแสดงว่ายาก (ค่าความยากง่ายของข้อสอบ หรือค่า p มีค่าอยู่ระหว่าง 0-1.00) ข้อสอบที่ดี ค่า p อยู่ระหว่าง 0.20-0.80 5) มีอำนาจจำแนก หมายถึงความสามารถแบ่งแยกกลุ่มออกเป็นประเภทต่าง ๆ ได้ถูกต้อง ข้อสอบที่มีอำนาจจำแนกดีหมายถึง ข้อสอบที่คนเก่งตอบถูก คนอ่อนตอบผิด ข้อสอบที่จำแนกไม่ได้คนเก่งและคนอ่อนจะตอบถูกผิดพอ ๆ กัน ไม่มีความแตกต่างกัน โดยทั่วไปค่าอำนาจจำแนกข้อสอบ หรือค่า r อยู่ระหว่าง -1.00 ถึง +1.00 ถ้าค่า r มีเครื่องหมายลบ หมายถึงการจำแนกกลับทาง คนเก่งตอบถูกน้อยกว่าคนอ่อน ค่า r ใกล้ศูนย์ ($r = -.19$ ถึง $+.19$) ข้อสอบจำแนกไม่ได้เพราะคนเก่งและคนอ่อนตอบถูกผิดพอ ๆ กัน ข้อสอบที่ดี ค่า r ควรมีค่าอยู่ระหว่าง 0.20 ถึง 1.00 6) มีประสิทธิภาพ คือเครื่องมือที่สามารถทำให้ได้ข้อมูลที่ดีที่สุด เชื่อถือได้มากโดยใช้วิธีการที่สะดวก รวดเร็ว คล่องตัว แต่เสียเวลาน้อย ลงทุนน้อยและใช้แรงงานน้อย 7) มีความยุติธรรม ไม่เปิดโอกาสให้มีการได้เปรียบเสียเปรียบกันระหว่างผู้ที่ถูกวัดด้วยกัน 8) ใช้คำถามถามลึก ข้อสอบที่ดีต้องถามให้ผู้ตอบใช้ความสามารถในการคิดค้นก่อนที่จะตอบ 1) ใช้คำถามช่วย มีลักษณะที่ทำให้ผู้สอบอยากคิดอยากตอบและทำด้วยความเต็มใจ 9) คำถามจำเพาะเจาะจง ไม่ถามกว้างเกินไป หรือถามคลุมเครือให้คิดได้หลายแง่หลายมุม

5.3 การสร้างเครื่องมือด้านพุทธิพิสัย

เครื่องมือที่วัดด้านพุทธิพิสัยส่วนใหญ่นิยมใช้แบบทดสอบ ซึ่งมีหลากหลายชนิดดังได้กล่าวมาแล้ว แต่แบบทดสอบที่ใช้สำหรับประเมินผลการเรียนรู้ในการจัดการเรียนการสอน ที่ใช้ส่วนใหญ่มี 2 ประเภทคือแบบทดสอบอัตรณ์ และแบบทดสอบปรนัย ซึ่งมีหลักการสร้างข้อสอบดังนี้

1) การสร้างข้อสอบอัตรณ์

ข้อสอบอัตรณ์เป็นข้อสอบที่เปิดโอกาสให้ผู้ตอบได้แสดงความสามารถในด้านความรู้ ภาษา ความคิดริเริ่ม วิเคราะห์ข้อความได้อย่างเหมาะสม แบ่งออกเป็น 2 ลักษณะคือ แบบไม่จำกัดคำตอบ และแบบจำกัดคำตอบ ซึ่งทั้ง 2 ลักษณะมีหลักการสร้างดังต่อไปนี้

1. คำนึงถึงลำดับความสำคัญของจุดมุ่งหมาย ที่วางไว้ตามตารางวิเคราะห์หลักสูตร โดยพิจารณานำหนักจากจุดมุ่งหมาย และเนื้อหาที่ได้สอน

2. วางแนวทางการออกข้อสอบว่า จะเป็นแบบ จำกัดหรือไม่จำกัดคำตอบ

3. เขียนคำสั่งให้ชัดเจนว่า ต้องการให้ผู้สอบตอบอย่างไร เกณฑ์การให้คะแนนเป็นอย่างไร

4. ตั้งประเด็นคำถาม ที่ประเมินความรู้จริงของผู้เรียน โดยเน้นด้านใดบ้าง เช่น การสังเคราะห์ วิเคราะห์ หรือประมาณค่า

5.พยายามใช้คำถามหลายๆแบบ มิใช่มีแต่คำถามประเภท ใคร อะไร ที่ไหน ที่มุ่งถามเพื่อวัดความรู้ ความจำเท่านั้น แต่ควรใช้คำถาม เหตุใด อย่างไร เพื่อให้ผู้เรียน อธิบายและแสดงเหตุผล หรือถามความคิดเห็น เชิงวิเคราะห์ ในเชิงเปรียบเทียบ แสดงความสัมพันธ์ ความขัดแย้ง ตีความ เป็นต้น

6.ทบทวน และพยายามปรับปรุงให้ข้อคำถามมีความชัดเจน ไม่กำกวม

2)การสร้างข้อสอบปรนัย

แบบปรนัยที่รู้จักกันดีได้แก่

1)แบบถูกผิด ซึ่งก็คือแบบทดสอบเลือกตอบสองตัวเลือก ซึ่งผู้ตอบมีโอกาสเลือกตอบข้อใดข้อหนึ่ง อาจตอบว่า ผิดหรือถูก ใช่หรือไม่ มีข้อแนะนำการสร้างดังนี้ (1)เขียนคำถาม คำตอบ และข้อความให้ชัดเจน กระชับไม่ยอกย้อนไม่ต้องตีความ (2)ไม่ใช่ข้อความปฏิเสธซ้อน (3)หลีกเลี่ยงลอกข้อความจากคำบรรยายหรือหนังสือ (4)ข้อผิด ถูกควรใกล้เคียงกัน เป็นต้น ข้อดีของข้อสอบนี้คือ ตรวจจับรวดเร็วประหยัดค่าใช้จ่าย ข้อเสีย คือเดาง่าย มีค่าความเชื่อมั่นต่ำ

2)แบบทดสอบแบบเติมคำ เป็นแบบทดสอบที่ให้ผู้เรียนเขียนคำตอบสั้น ๆ มักมีขอบเขตการตอบอย่างชัดเจนจากช่องว่างที่เว้นไว้ ข้อแนะนำในการสร้าง (1)เขียนคำสั่งให้รัดกุมว่าตอบตรงไหนอย่างไร (2)คำถามควรชัดเจนถามเจาะไม่ถามกว้างและไม่ควรลอกข้อความคำบรรยายหรือจากหนังสือมาถาม (3)เว้นช่องให้มากพอสำหรับเขียนตอบและควรอยู่ตอนท้ายประโยค ไม่ควรมีหลายแห่งในแต่ละข้อ ข้อดีคือ สร้างง่ายเดาถูกยาก ประหยัดเวลา ควรใช้สอบเฉพาะความรู้เกี่ยวกับข้อเท็จจริง กฎเกณฑ์ต่าง ๆ

3)แบบจับคู่เป็นปรนัยประเภทกำหนดคำหรือข้อความเป็นสองแถว คล้ายกับแบบทดสอบแบบเลือกตอบ ข้อความแถวหนึ่งไปใส่อีกแถวหนึ่ง ข้อแนะนำในการสร้าง (1)ในแถวที่เป็นคำตอบควรมีมากกว่าคำถามเพื่อป้องกันการเดาและควรเป็นคำตอบที่สั้น ๆ (2)คำถามควรระบุเงื่อนไขหรือหลักการจับคู่ให้ชัดเจน (3)ข้อคำถามและคำตอบควรจัดให้อยู่ในหน้าเดียวกัน (4)วิธีตอบไม่ยุ่งยาก เช่นการเลือกเฉพาะอักษรมาตอบแทนการลอกทั้งข้อความ ข้อดีสร้างง่ายรวดเร็วตรวจง่าย หากสร้างได้ดีเดายาก ประหยัด แต่มักเป็นข้อสอบที่วัดความรู้ความจำ และค่อนข้างยากที่จะทำให้ปัญหาและคำตอบเป็นเรื่องเดียวกันหรือประเภทเดียวกัน(หากประเภทไม่เหมือนกันเดาง่าย)

4)แบบเลือกตอบ มักประกอบด้วยคำถามสั้นและคำตอบให้เลือกเพียงข้อเดียว หลักการเขียนข้อสอบ (1)คำถามควรอยู่ในรูปประโยคคำถามที่สมบูรณ์เจาะจง สั้น และชัดเจนตรงจุด อาจนำคำตอบที่ได้มาเชื่อมต่อกับคำถาม (2)หลีกเลี่ยงคำถามปฏิเสธซ้อนปฏิเสธ หากจำเป็นต้องใช้ควรเน้นหรือขีดเส้นให้ชัด (3)ข้อเดียวควรมีคำตอบเดียว (4)ควรมีตัวถูกและตัวลวงที่มีควรเป็นไปตามหลักวิชา

6.การวัดด้านจิตพิสัย

จิตพิสัย เป็นพฤติกรรมเกี่ยวกับความรู้สึกนึกคิดทางจิตใจ อารมณ์ และคุณธรรมของบุคคล ซึ่งสามารถจำแนกออกได้เป็น 5 ระดับ คือ 1) การรับรู้ 2) การตอบสนอง 3) การสร้างคุณค่า 4) การจัดระบบคุณค่า และ5) การสร้างลักษณะนิสัย

1)การรับรู้ เป็นพฤติกรรมที่ตอบสนอง แบ่งเป็น 3 ลักษณะคือ 1)การยอมรับ เป็นขั้นทำความเข้าใจกับเหตุการณ์หรือปรากฏการณ์ที่เกิดขึ้นในความสนใจของตนเอง 2)การตั้งใจที่จะรับรู้ เป็นการแสดงความปรารถนาที่จะรับรู้เกี่ยวกับสิ่งนั้นด้วยความเต็มใจ และ3)การเลือกสิ่งเร้าที่ต้องการรับรู้ เป็นการแยกแยะสิ่งที่ต้องการจะรับรู้ โดยยังไม่ได้รู้รายละเอียดของปรากฏการณ์สิ่งเร้า นั้น ๆ

2)การตอบสนอง เป็นพฤติกรรมต่อเนื่องจากการรับรู้ โดยมีปฏิกิริยาโต้ตอบสิ่งเร้านั้นอย่างเป็นใจ และเกิดความพึงพอใจในการตอบสนอง พฤติกรรมนี้จำแนกได้เป็น 3 ลักษณะคือ 1) การยินยอมที่จะตอบสนอง เป็นการยอมตอบสนองถึงแม้จะไม่เต็มใจ 2) ความเต็มใจที่จะตอบสนอง เป็นการยอมรับที่ตอบสนองด้วยความเต็มใจ 3)ความพอใจในการตอบสนอง ได้ตอบสนองแล้วเกิดความสุขมีความพึงพอใจที่ได้กระทำลงไป

3) การสร้างคุณค่า เป็นขั้นที่บุคคลมองเห็นคุณค่าของการตอบสนองต่อสิ่งเร้า หรือประสบการณ์ที่ได้ตีแบ่งพฤติกรรมที่แสดงออกเป็น 3 ลักษณะ 1) การยอมรับในคุณค่า เป็นขั้นการยอมรับว่าพฤติกรรมที่แสดงออกไปนั้นเป็นสิ่งที่มีความดี 2) การนิยมชมชอบในคุณค่า เป็นการยอมรับคุณค่า (ข้อ 1) ด้วยความพึงพอใจ 3) การสร้างคุณค่า มีความคงเส้นคงวาในการตอบสนองต่อสิ่งเร้า โดยแสดงพฤติกรรมที่คนเห็นว่ามีคุณค่าอย่างสม่ำเสมอและตอบสนองอย่างต่อเนื่องไปจนเกิดการยอมรับเป็นค่านิยมของตนเอง นอกจากนี้ยังพยายามทำให้ผู้อื่นคล้อยตามค่านิยมดังกล่าวด้วย

4) การจัดระบบคุณค่า เป็นการนำเอาหลาย ๆ คุณค่ามาจัดระบบเพื่อสร้างเป็นลักษณะภายในตนเองที่คงที่แน่นอน

5) การสร้างลักษณะนิสัย เป็นการจัดระบบคุณค่าที่มีอยู่ในตัวเอง เข้าเป็นระบบที่ถาวรซึ่งจะทำหน้าที่ควบคุมพฤติกรรมแสดงของบุคคล ไม่ว่าจะอยู่ในสถานการณ์ใด ๆ ก็แสดงพฤติกรรมตามค่านิยมที่ยึดถือตลอดไป และจะแสดงพฤติกรรมดังกล่าวนี้ได้อย่างสม่ำเสมอจนเกิดเป็นนิสัยประจำตัวของบุคคล การสร้างลักษณะนิสัยมี 2 ลักษณะได้แก่ 1) การสร้างลักษณะนิสัยชั่วคราวเป็นการแสดงพฤติกรรมที่สอดคล้องกับคุณค่าส่วนตัวบางอย่างของบุคคล โดยคำนึงถึงผลที่จะเกิดตามมาในสถานการณ์นั้น ๆ ด้วย และ 2) การสร้างลักษณะนิสัยถาวร เป็นขั้นแสดงนิสัยที่แท้จริงของบุคคลออกมาอย่างสมบูรณ์ เพราะเป็นการรวมเอาคุณลักษณะที่เป็นคุณค่าต่าง ๆ ของบุคคลเข้าไว้ด้วยกัน ขั้นนี้ถือว่าเป็นจุดสุดยอดของการพัฒนาคน

อย่างไรก็ตาม คุณลักษณะในแต่ละระดับเมื่อเกิดและถูกจัดระบบเป็นบุคลิกลักษณะประจำตัวของบุคคลแล้ว จะแสดงให้เห็นถึงคุณลักษณะหลัก ๆ ออกมาในรูปของ ความสนใจ เจตคติ และค่านิยมของบุคคล ตัวอย่าง เช่น ความสนใจ เป็นลักษณะของการมีใจจดจ่ออยู่กับสิ่งใดสิ่งหนึ่งในช่วงเวลาใดเวลาหนึ่ง ความสนใจของบุคคลเริ่มจากการที่บุคคลมีการรับรู้ต่อสิ่งใดสิ่งหนึ่งด้วยความพึงพอใจ และมองเห็นคุณค่าของสิ่งนั้น จนเกิดความรู้สึกนิยมชมชอบในคุณค่าของสิ่งที่รับรู้ นั้น และจะตอบสนองต่อสิ่งนั้นอย่างคงเส้นคงวานั่นเอง เช่น ความสนใจในกีฬา ความสนใจในดนตรี ความสนใจในการท่องเที่ยว เป็นต้น คุณลักษณะด้านจิตพิสัย มักจะถูกพัฒนาในเกิดกับผู้เรียนในลักษณะแฝงไปกับรายวิชาและกิจกรรมการเรียนรู้ต่าง ๆ ที่ผู้สอนจัดให้

6.1. การวัดพฤติกรรมด้านจิตพิสัย

จิตพิสัยเป็นความรู้สึกนึกคิดทางจิตใจ อารมณ์ และคุณธรรมของบุคคล ดังนั้น คุณลักษณะของจิตพิสัยที่สำคัญมีดังนี้ 1) เป็นคุณลักษณะภายในจิตใจ เป็นความรู้สึกที่รับรู้ ดังนั้นการวัดด้านจิตพิสัยจึงเป็นการวัดทางอ้อม จากพฤติกรรมของผู้ถูกวัดที่แสดงออกมา เพื่อสะท้อนถึงความรู้สึกและอารมณ์ของผู้ถูกวัด วิธีวัดจะกระทำโดยกระตุ้นหรือเร้าให้ผู้ถูกวัดแสดงความคิดเห็นที่ตอบสนองต่อสิ่งเร้าที่กระตุ้นนั้น ผู้วัดก็จะแปลความจากพฤติกรรมดังกล่าว 2) เป็นนามธรรมไม่สามารถสังเกตได้โดยตรง เช่น ความสนใจ ความรับผิดชอบ 3) ต้องวัดทางอ้อมทำให้เกิดความคลาดเคลื่อนได้ง่ายต้องใช้เครื่องมือที่มีความเชื่อมั่นสูง 4) การวัดไม่มีผิดไม่มีถูกขึ้นกับประสบการณ์และความคิดเห็น ความรู้สึกนึกคิดของผู้ตอบแบบสอบถาม 5) มักนิยมใช้ข้อมูลจากหลายฝ่าย ได้แก่ บุคคลที่ต้องการวัด บุคคลใกล้ชิด และการสังเกตของผู้วัดเอง 6) มักใช้สถานการณ์จำลองเป็นเงื่อนไขให้ผู้ถูกวัดตอบ ดังนั้นเครื่องมือจึงควรมีคุณลักษณะด้านความเที่ยงตรงตามสภาพสูง 7) มีจุดอ่อนที่ผู้ตอบสามารถบิดเบือนหรือหลอกผู้ถามได้ การแสดงออกของคุณลักษณะจิตพิสัย มีทิศทางแสดงออกได้สองแนวทางตรงกันข้าม เช่น รัก-เกลียด ชอบ-ไม่ชอบ ฯลฯ และมีระดับความเข้มของระดับความรู้สึก เช่น รักมากที่สุด มาก ปานกลาง น้อย น้อยที่สุด เป็นต้น คุณลักษณะจิตพิสัย สามารถแบ่งเป็น 2 ลักษณะคือ 1) การประเมินตนเอง โดยผู้ถูกวัดจะเป็นผู้ตอบแบบวัดด้วยตนเองโดยการแสดงความรู้สึก 2) การประเมินโดยผู้อื่น ซึ่งเป็นการวัดโดยผู้ประเมินเป็นผู้วัดเองหรืออาจมอบหมายให้เพื่อนผู้ปกครอง หรือผู้ที่เกี่ยวข้อง เป็นผู้ใช้เครื่องมือ

6.2 หลักการประเมินผลด้านจิตพิสัย

1). วัตถุประสงค์ของบุคคลที่แสดงออกมาเหมือนกัน ไม่จำเป็นต้องเป็นผลจากอารมณ์หรือความรู้สึกอย่างเดียวกัน ขณะเดียวกันผู้ที่มีอารมณ์ ความรู้สึกอย่างเดียวกัน ก็ไม่แน่ว่าจะแสดงพฤติกรรมอย่างเดียวกัน ดังนั้นการวัดให้ครอบคลุมจึงจำเป็นต้องใช้เครื่องมือวัดหลากหลายชนิด

2). การวัดควรวัดหลาย ๆ ครั้ง เนื่องจากความรู้สึกและอารมณ์ของบุคคลมักจะมีการเปลี่ยนแปลงได้ตามสถานการณ์ และบางครั้งสถานการณ์เดียวกันแต่ต่างเวลาก็น่าจะมีความรู้สึกไม่เหมือนกัน ดังนั้น การวัดจึงควรทำหลาย ๆ ครั้ง และควรสุ่มทั้งสถานที่ และเวลาให้แตกต่างกัน จึงจะทำให้การวัดมีความน่าเชื่อถือ

3). วัดผลอย่างต่อเนื่อง การวัดที่ดีไม่ควรเชื่อข้อมูลที่ได้จากการวัดเพียงครั้งเดียว แต่ควรวัดอย่างต่อเนื่องกันหลาย ๆ ครั้งด้วยเทคนิคที่แตกต่างกัน หลากหลายวิธี จึงจะเชื่อได้ว่า คุณลักษณะด้านจิตใจและอารมณ์นั้น เป็นลักษณะถาวรของบุคคลนั้น

4). ความร่วมมือของผู้ถูกวัดเป็นหัวใจสำคัญของการวัด เนื่องจากการวัดจิตพิสัยเป็นการวัดคุณลักษณะของบุคคล ดังนั้น ก่อนวัดควรมั่นใจว่าจะได้รับความร่วมมือจากผู้ตอบว่าข้อมูลที่ได้รับนั้นเป็นข้อมูลจริง นอกจากนี้ผู้วัดควรหาเทคนิคที่ทำให้ผู้ตอบ ตอบด้วยความสบายใจและมั่นใจในความปลอดภัยของข้อมูล

5). ใช้ผลการวัดให้ถูกต้อง เนื่องจากการวัดผลด้านจิตพิสัยไม่มีถูกไม่มีผิด ดังนั้นผลการวัดจึงไม่สามารถนำไปใช้ในการตัดสินได้หรือตก แต่การวัดทางจิตพิสัยเป็นการวัดเพื่อนำข้อมูลไปใช้ในการพัฒนาบุคคล ไม่ใช่เพื่อการตัดสินใจ

6.3 หลักการสร้างเครื่องมือวัดจิตพิสัย

การวัดผลด้านจิตพิสัย เป็นการวัดความรู้สึกนึกคิด คุณธรรมจริยธรรม อารมณ์ของบุคคล มักวัดทางอ้อมจากพฤติกรรมที่แสดงภายใต้สถานการณ์สมมติ เพื่อสะท้อนถึงความรู้สึกนึกคิดดังกล่าว ผลการวัดมีลักษณะความคิดเห็นต่อเรื่องใดเรื่องหนึ่งภายใต้สถานการณ์ จึงไม่มีผิดไม่มีถูก มีความคาดเคลื่อนง่าย (ผู้ถูกวัดสามารถปิดบังบิดเบือนง่าย พฤติกรรมด้านจิตพิสัยแบ่งได้เป็น 5 ระดับ ได้แก่ การรับรู้ การตอบสนอง การสร้างคุณค่า การจัดระบบคุณค่า และการสร้างลักษณะนิสัย การวัดจึงนิยมวัดให้ครอบคลุม วัดหลาย ๆ ครั้ง วัดอย่างต่อเนื่อง และขอความร่วมมือจากผู้ถูกวัด เครื่องมือวัดจิตพิสัยจึงมีหลากหลายประเภทแตกต่างกัน และไม่สามารถบอกได้ชัดว่าเครื่องมือใดเหมาะสมหรือดีกว่ากัน ความเหมาะสมขึ้นกับเหตุปัจจัยหลายประการ อย่างไรก็ตาม เครื่องมือที่นิยมใช้มากในการวัดพฤติกรรมจิตพิสัยคือ 1) มาตรการส่วนประมาณค่า 2) แบบสำรวจรายการ 3) แบบวัดเชิงสถานการณ์ และ 4) แบบสังเกต

การสร้างเครื่องมือจิตพิสัยมีลำดับขั้นตอนการสร้างดังต่อไปนี้ 1) กำหนดคุณลักษณะที่ต้องการวัด ได้แก่ ความรู้สึก คุณธรรมจริยธรรม และเจตคติ 2) กำหนดพฤติกรรมบ่งชี้ หรือนิยามพฤติกรรมที่บ่งชี้แสดงถึงเจตคติที่สนใจ ผู้เรียนมาเรียนวิชา การเรียนการสอนทางกายภาพบำบัด ตรงตามเวลาไม่เคยขาดเรียนเป็นการบ่งชี้ถึงความตั้งใจและสนใจเรียนวิชาดังกล่าว 3) กำหนดวิธีการหรือเครื่องมือวัด เพื่อพิจารณาเครื่องมือที่เหมาะสมในการวัดพฤติกรรมที่สนใจ โดยยึดหลักเหมาะสมและไม่ยุ่งยากในการเก็บข้อมูล 4) สร้างเครื่องมือวัด 5) ตรวจสอบเครื่องมือ มักนิยมนำไปทดลองใช้กับกลุ่มตัวอย่างที่มีลักษณะใกล้เคียงกับกลุ่มที่ต้องการวัด เพื่อดำเนินการคุณภาพเครื่องมือ ด้าน ความเที่ยงตรง ความเชื่อมั่น และอำนาจจำแนก

7. การวัดด้านทักษะพิสัย

ทักษะพิสัยเป็นความสามารถในเชิงปฏิบัติการหรือการกระทำให้เกิดผลอย่างใดอย่างหนึ่ง ทักษะพิสัยสามารถจำแนกออกเป็น 7 ระดับ ดังต่อไปนี้ 1) การรับรู้ 2) เตรียมความพร้อม 3) การตอบสนองตามแนวทางที่กำหนดให้ 4) ความสามารถด้านกลไก 5) การตอบสนองที่ซับซ้อน 6) ความสามารถในการดัดแปลง และ 7) ความสามารถในการริเริ่ม

1) การรับรู้ เป็นขั้นแสดงอาการรับรู้ที่จะเคลื่อนไหวโดยอาศัยประสาทสัมผัสรับรู้ทั้ง 5 ได้แก่ หู ตา จมูก ลิ้น และสัมผัสทางร่างกาย แม้จะมีสิ่งเร้ามากระตุ้นโดยผ่านทางประสาทสัมผัสพร้อม ๆ กัน บุคคลก็สามารถเลือกรับรู้ และมีการแปลความหมายต่อสิ่งเร้า นั้น ๆ เพื่อการตอบสนอง

2) เติริยมความพร้อม เป็นความพร้อมของบุคคลที่จะแสดงพฤติกรรมออกมา สภาพความพร้อมมี 3 ด้าน คือ ความพร้อม 1) ด้านร่างกาย หมายถึงความพร้อมที่จะเคลื่อนไหวอวัยวะของร่างกาย 2) ด้านสมอง หมายถึงความพร้อมที่ระลึกถึงระเบียบกฎเกณฑ์ต่าง ๆ ในการแสดงพฤติกรรม และ 3) ด้านอารมณ์ ความพร้อมด้านจิตใจ ให้อยู่ในทิศทางที่พึงปรารถนาหรือไม่พึงปรารถนา

3) การตอบสนองตามแนวทางที่กำหนดให้ เป็นการแสดงออกในลักษณะของการเลียนแบบและการลงมือทดลองถูก

4) ความสามารถด้านกลไก เป็นขั้นที่ผู้เรียนได้กระทำตามที่เรียนมา และพัฒนาขึ้นจนมีความสัมฤทธิ์ผล สามารถสร้างเทคนิควิธีการเฉพาะสำหรับตนเองขึ้นมา เพื่อฝึกปฏิบัติต่อไป

5) การตอบสนองที่ซับซ้อน เป็นความสามารถในการปฏิบัติในสิ่งที่ยุ่ยากซับซ้อนมากขึ้น และสามารถกระทำได้อย่างมั่นใจ ไม่ลังเล และทำได้ดีจนเป็นอัตโนมัติ

6) ความสามารถในการดัดแปลง เป็นขั้นที่สามารถปฏิบัติได้จำชานาญ จึงคิดหาวิธีการใหม่ ๆ มาลองทำ ให้แตกต่างไปจากเดิม เพื่อให้เกิดผลดียิ่ง ๆ ขึ้น

7) ความสามารถในการริเริ่ม เป็นความสามารถต่อยอดจากการดัดแปลง กล่าวคือ เมื่อบุคคลสามารถดัดแปลงวิธีการใหม่ ๆ และมีการทดลองทำดูแล้ว ก็นำวิธีการนั้นมาประยุกต์ทำให้เกิดสิ่งใหม่ ๆ ขึ้น

7.1 การวัดพฤติกรรมด้านทักษะพิสัย

การวัดทักษะพิสัย เป็นการวัดความสามารถในการทำงานหรือการทำกิจกรรม อาจเรียกว่า การวัดภาคปฏิบัติ ซึ่งการวัดพฤติกรรมดังกล่าวมักเป็นการประสานสัมพันธ์ระหว่าง พุทธิพิสัย จิตพิสัย และทักษะพิสัย ลักษณะสำคัญของการวัดทักษะพิสัยคือ 1) สามารถวัดได้ทั้งรายบุคคลหรือเป็นกลุ่ม 2) ลักษณะงานที่แตกต่างกันต้องใช้วิธีการวัดหรือกระบวนการที่ต่างกัน ตัวอย่างเช่น วัดทักษะด้านดนตรี ก็จำเป็นต้องใช้วิธีการวัดต่างกัน 3) สามารถแบ่งวัดเป็นกระบวนการหรือทักษะย่อย ๆ เช่น การรักษาทางกายภาพบำบัด สามารถวัด ขั้นตอนตรวจร่างกาย การประมวลผลการสรุปประเด็นปัญหา และการให้การรักษารักษา เป็นต้น 4) ลักษณะการวัดแยกได้เป็น 3 ระดับ (1) ระดับพฤติกรรมโดยการสังเกตพฤติกรรมขณะปฏิบัติ (2) ระดับผลลัพธ์ พิจารณาจากผลจากพฤติกรรมย่อยที่ปรากฏซึ่งพิจารณาได้ทั้งเชิงปริมาณและคุณภาพ (3) ระดับประสิทธิผล เป็นการวัดผลงานเฉพาะที่เกี่ยวข้องกับวัตถุประสงค์ หรือเป้าหมายการวัดโดยตรง

7.2 หลักการประเมินผลด้านทักษะพิสัย

หลักการประเมินผลด้านทักษะพิสัย แบ่งออกเป็น 1) วัดให้ตรงจุดมุ่งหมาย โดยพิจารณาผลสุดท้ายที่จะเกิดขึ้นเมื่อสิ้นสุดการเรียนการสอน เช่น หากกำหนดวัตถุประสงค์การเรียนรู้ สามารถประเมินผู้ป่วยได้ ตรวจร่างกายได้ แต่จุดมุ่งหมายการเรียนรู้คือการรักษาโดยการนวด ดังนั้นการวัดต้องวัด ที่ละทักษะย่อยก่อน เป็นต้น 2) ความสอดคล้องของงานที่กำหนดกับจุดประสงค์ ต้องการประเมินความสามารถในการนวด งานที่กำหนดต้องใช้ความสามารถในการนวดโดยตรง เป็นต้น 3) ความเหมาะสมในการดำเนินงาน พิจารณาว่าผลงานดังกล่าวเกิดจากทำงานกลุ่มหรือเดี่ยว 4) ยุติธรรม 5) การให้คะแนนควรเป็นปรนัยและเที่ยงตรง กำหนดไว้ชัดเจน ในแต่ละขั้นตอนจะได้คะแนนเท่าไร 6) มีเกณฑ์การประเมินผล ซึ่งพิจารณาจาก กระบวนการหรือวิธีการ และผลผลิตที่เกิดขึ้นซึ่งพิจารณาทั้งปริมาณและคุณภาพ 7) วัดหลาย ๆ ครั้ง

ขั้นตอนการประเมิน มีดังนี้ 1) กำหนดจุดประสงค์ มักเขียนเป็นจุดประสงค์เชิงพฤติกรรม เพื่อให้สามารถวัดได้จริง 2) กำหนดงานให้ปฏิบัติ อาจแบ่งเป็นขั้นตอนย่อย ๆ ได้ 3) กำหนดวิธีดำเนินการอย่างชัดเจน 4) กำหนด

สถานการณ์หรือเงื่อนไขที่ใช้ในการปฏิบัติงาน 5) กำหนดวิธีการวัด สามารถทำได้หลายวิธีเช่น เขียนตอบ (วัดความรู้หรือขั้นตอนการปฏิบัติ) สร้างสถานการณ์จำลองแล้วให้ตอบคำถาม วัดผลงานปฏิบัติจริง 6) กำหนดเครื่องมือวัด 7) กำหนดเกณฑ์การให้คะแนน 8)กำหนดวิธีการประเมินผล เช่น อิงเกณฑ์หรืออิงกลุ่ม 9)การสร้างเครื่องมือวัด มีลำดับขั้นตอนดังนี้ (1) วางแผนสร้างเครื่องมือ พิจารณาวัดดูประสงค์การสอน (2) ดำเนินการสร้างเครื่องมือ (3) ตรวจสอบคุณภาพเครื่องมือ (4) สร้างคู่มือการใช้เครื่องมือ 10) ดำเนินการวัด

7.3 เครื่องมือวัดด้านทักษะพิสัย

วิธีการและเครื่องมือวัดทักษะพิสัยที่นิยมใช้ได้แก่ การทดสอบ แบบสังเกต แบบสำรวจรายการ และมาตราส่วนประมาณค่า ซึ่งแต่ละประเภทมีความเหมาะสมแตกต่างกัน ซึ่งการวัดหนึ่ง ๆ เพื่อความเหมาะสมและเกิดความชัดเจนในการวัดอาจประยุกต์ใช้เครื่องมือหลายชนิดแตกต่างกัน

1) การทดสอบ (testing)

การทดสอบเป็นวิธีที่นิยมใช้มากในการวัดผลด้านทักษะพิสัย สามารถใช้ได้ทั้งการทดสอบแบบปากเปล่าหรือการสอบข้อเขียน 1) การสอบข้อเขียนนั้น เหมาะสำหรับเนื้อหาที่เป็นทฤษฎี ผู้สอบกลุ่มใหญ่ เป็นการตรวจสอบความรู้ก่อนปฏิบัติจริงเพื่อลดความเสี่ยงที่จะเกิดความเสียหายหากสอบปฏิบัติจริง 2) การสอบปากเปล่า เป็นการตรวจสอบความรู้ความสามารถรายบุคคล ซึ่งมีความหลากหลาย มีข้อเสียของความได้เปรียบเสียเปรียบโดยเฉพาะผู้สอบที่หลังมีโอกาสในการเตรียมตัวได้มากกว่า

2) การสังเกต

เป็นการสังเกตการปฏิบัติจริงของผู้สอบ เครื่องมือที่ใช้ประกอบการสังเกตได้แก่ แบบบันทึกข้อมูล แบบสำรวจแบบ check list และมาตราส่วนประมาณค่า ซึ่งเป็นวิธีที่ดีที่สุด ผู้สอบมีโอกาส เห็นกระบวนการลำดับขั้นตอนทำงานของผู้เรียน เห็นพฤติกรรมต่าง ๆ ในสภาพจริง เพื่อเป็นการลดความกังวลผู้ประเมินอาจมอบหมายให้ผู้เรียนอื่น หรือผู้ร่วมสอนช่วยประเมินก็ได้

3) แบบสำรวจรายการ (checklist)

มักใช้เป็นเครื่องมือในการตรวจสอบรายการกิจกรรม และพฤติกรรมในการปฏิบัติงานว่า มีรายการ และหากมี มีมากน้อยเพียงไร คุณภาพเป็นอย่างไร หรือขั้นตอนที่ผู้สอบปฏิบัติ นั้น ครบตามจุดประสงค์ของการประเมินหรือไม่ ขั้นตอนของการสร้างและการตรวจสอบเครื่องมือ ก็ไม่แตกต่างจากเครื่องมือที่ใช้วัดจิตพิสัย ดังที่ได้กล่าวมาแล้ว

8. การประเมินสภาพจริง

การประเมินสภาพจริง เป็นวิธีการประเมินที่ออกแบบมาเพื่อสะท้อนให้เห็นพฤติกรรมและทักษะที่จำเป็นของผู้เรียนในสถานการณ์ที่เป็นจริง และเป็นวิธีการประเมินที่เน้นงานหรือกิจกรรมที่ผู้เรียนได้แสดงออกโดยการกระทำ เน้นกระบวนการเรียนรู้ ผลผลิต และผลงาน เปิดโอกาสให้ผู้เรียนมีส่วนร่วมในการประเมินและร่วมในการจัดกระบวนการเรียนรู้ของตนเอง ซึ่งวิธีการนี้เชื่อว่าจะช่วยพัฒนาการ การเรียนรู้ของผู้เรียนได้อย่างต่อเนื่อง กระบวนการประเมินอาจใช้วิธีการสังเกต การบันทึก การรวบรวมข้อมูลจากผลงานและวิธีการที่ผู้เรียนได้เคยทำได้ ด้วยวิธีการที่หลากหลาย กลยุทธ์สำคัญของการประเมินตามสภาพจริงคือ การกระตุ้นหรือท้าทายให้ผู้เรียนได้แสดงออกโดยการกระทำ ว่าตนเองมีความสามารถอะไร และได้เคยทำสิ่งใดบ้าง แทนการทำแบบทดสอบหรือข้อสอบเหมือนการประเมินแบบเดิม ๆ นอกจากเน้นเรื่องการกระทำและผลงานแล้ว การประเมินนี้ยังเน้นความสามารถทางสติปัญญา กระบวนการสร้างความสัมพันธ์ระหว่างบุคคล การแก้ปัญหา มากกว่าการเน้นเรื่องการทำข้อสอบ หรือการหาคำตอบจากแบบทดสอบ ความแตกต่างของการประเมินสภาพจริงและการประเมินแบบเดิมได้สรุปไว้ในตารางที่ 17

ตารางที่ 17 เปรียบเทียบการประเมินแบบเดิม และการประเมินสภาพจริง

การประเมินแบบเดิม	การประเมินสภาพจริง
1.เน้นพฤติกรรมพื้นฐานที่ไม่ซับซ้อน	1.เน้นพฤติกรรมที่ซับซ้อน ความคิดที่ซับซ้อน และวิธีการนำความรู้ไปใช้มากกว่ามีความรู้
2.แยกการสอบและการประเมินออกจากการเรียนการสอน	2.การสอบและประเมินร่วมไปกับการจัดการเรียนการสอน (การประเมินเป็นส่วนหนึ่งของการจัดการเรียนการสอน) เปิดโอกาสให้เกิดการเรียนรู้อย่างต่อเนื่อง
3.มีข้อจำกัดในการวัด	3.วัดได้อย่างหลากหลายและกว้างขวาง
4.ยึดถือตัวเลขและข้อมูลที่ได้จากการสอบ	4.พิจารณาจากข้อมูลที่หลายหลาย และความสอดคล้องของการอธิบายความ
5.เชื่อถือการประเมินของผู้อื่นและผู้สอนเปรียบเสมือนผู้ประเมินภายนอก	5.เชื่อในการประเมินตนเอง และผู้สอนก็เป็นเพียงส่วนหนึ่งของการประเมิน
6. มีเกณฑ์มาตรฐาน เพื่อบ่งบอกถึงความสำเร็จ	6. มีเกณฑ์บ่งบอกถึงความสำเร็จอย่างหลากหลาย
7.มีความคิดเป็นเอกนัย (convergent thinking) และมีคำตอบที่ถูกต้องเพียงคำตอบเดียว	7.มีความคิดตอเนกนัย (divergent thinking) มีคำตอบอย่างหลากหลาย
8. เน้นการประเมินแบบแยกทักษะ	8.เน้นการประเมินโดยบูรณาการทักษะ
9.เน้นรายวิชาเดียว	9.เน้นสหวิทยาการ

8.1 ลักษณะสำคัญของการวัดประเมินสภาพจริง

เป็นที่ทราบกันดีแล้ว การวัดประเมินผลตามสภาพจริง เป็นการประเมินทางเลือกใหม่ (alternative assessment) ที่เน้นการประเมินผลจากการปฏิบัติงานซึ่งต่างจากการประเมินที่เน้นการทดสอบเป็นสำคัญ ลักษณะสำคัญของการวัดประเมินผลตามสภาพจริงมีดังนี้

1)การประเมินฯเน้นแนวคิดที่ว่าความรู้เรื่องใดเรื่องหนึ่ง มีความหมายได้หลากหลาย ดังนั้นการวัดควรใช้วิธีการอย่างหลากหลาย

2)การเรียนรู้เป็นกระบวนการตามความต้องการของผู้เรียน มากกว่าการบังคับให้เรียน ดังนั้นผู้เรียนจึงมีความกระตือรือร้น และแสวงหาความรู้เพื่อความอยากรู้มากกว่าการเรียนรู้เพื่อให้ทำข้อสอบได้คะแนนสูง ๆ

3)การวัดประเมินผลตามสภาพจริง เน้นกระบวนการเรียนรู้ และผลผลิต โดยพิจารณาจากสิ่งที่ผู้เรียนเรียนรู้ และทำไมจึงเกิดการเรียนรู้เช่นนั้น

4)การวัดประเมินผลตามสภาพจริง มุ่งเน้นการสืบเสาะ พัฒนาทักษะการแก้ปัญหาตามสภาพจริงที่เกิดขึ้น ผู้เรียนต้องสังเกต วิเคราะห์ และทดสอบความรู้ของตนเองจากการปฏิบัติ

5)การวัดประเมินสภาพจริงมีจุดประสงค์เพื่อ กระตุ้น และอำนวยความสะดวกให้กับผู้เรียน และสะท้อนผลการเรียนรู้เพื่อการพัฒนาให้กับผู้เรียน

8.2 เครื่องมือวัดและประเมินตามสภาพจริง

โดยทั่วไปการวัดและประเมินผลให้ตรงกับสิ่งที่ต้องการวัด มักต้องใช้อาศัยเทคนิค การเก็บรวบรวมข้อมูล และใช้เครื่องมือที่มีคุณภาพเหมาะสมกับคุณลักษณะที่ต้องการศึกษา รวมถึงเงื่อนไขบริบทอื่น ๆ อาทิ จุดประสงค์การวัด ลักษณะผู้สอบ ปริมาณ ระยะเวลา สถานที่ งบประมาณ การประเมินตามสภาพจริงมักใช้วิธีการประเมินหลากหลาย ส่วนเทคนิคการเก็บรวบรวมข้อมูล ประกอบด้วย การทดสอบ การสอบสัมภาษณ์ การสังเกต การตรวจผลงาน การใช้แฟ้มสะสมงาน การประเมินโดยใช้ศูนย์ประเมิน

1) การทดสอบ

การทดสอบมักใช้แบบทดสอบ เพื่อประเมินความรู้ความสามารถของผู้เรียน เครื่องมือที่ใช้ประกอบด้วย แบบสอบเขียนตอบ แบบสอบเลือกตอบ และสอบภาคปฏิบัติ และแบบวัดต่างๆ เป็นต้น

2) การสอบสัมภาษณ์

เป็นวิธีการวัดผลด้วยการซักถาม สนทนา โต้ตอบ เพื่อประเมินความคิด ทักษะคิดต่าง เครื่องมือที่ใช้ได้แก่ แบบสัมภาษณ์แบบมีโครงสร้าง (เตรียมคำถามไว้ล่วงหน้า) และคำถามแบบไม่มีโครงสร้าง (กำหนดเฉพาะแนวทางหรือประเด็นแต่ไม่มีคำถามที่ชัดเจน)

3) การสังเกต

การสังเกต เป็นการวัดและประเมินที่มีรายการ พฤติกรรมเป้าหมายที่ต้องการเก็บข้อมูล ด้วยประสาทสัมผัส โดยเฉพาะอย่างยิ่งทางหูและตา เพื่อศึกษาพฤติกรรมที่มีความละเอียด ชัดเจนของผู้เรียนในสภาพการณ์ต่างๆ ที่กำหนด เครื่องมือที่ใช้ประกอบด้วย แบบตรวจสอบรายการ แบบมาตรวัดประเมินค่า และแบบบันทึก เป็นต้น

4) การตรวจผลงาน

เป็นการวัดและประเมินด้วยการกำหนด งาน กิจกรรม หรือแบบฝึก ให้ผู้เรียนได้ปฏิบัติ ฝึกฝน โดยผู้สอนจะเป็นผู้ตรวจสอบความถูกต้องด้วยตนเอง หรือเพื่อนผู้เรียนที่ได้รับมอบหมาย เพื่อให้ได้ข้อมูลจริงสำหรับสะท้อนผลการปรับปรุงแก้ไข พฤติกรรมการเรียนรู้ของผู้เรียนอย่างเป็นระบบต่อไป เครื่องมือที่ใช้ได้แก่ แบบประเมินผลงาน

5) การใช้แฟ้มสะสมงาน

เป็นการวัดและประเมินที่ใช้หลักการเก็บหลักฐานผลงานที่ดีและมีความภาคภูมิใจ ที่เป็นตัวแทนงานที่ปฏิบัติของผู้เรียนเกี่ยวกับ ทักษะ แนวคิด ความสนใจ ความสำเร็จ โดยมีผลการประเมิน จุดเด่น จุดด้อยของชิ้นงาน อันแสดงถึงความก้าวหน้าในการเรียนด้วยตนเองของผู้เรียนเอง เพื่อนร่วมชั้น หรือผู้สอน แล้วนำหลักฐานมาบรรจุลงในแฟ้ม สมุดโน้ต แผ่นบันทึกข้อมูล เป็นต้น ลักษณะแฟ้มสะสมงานที่ดีควรมีความหลากหลาย สามารถสะท้อนความสามารถที่แท้จริงของผู้เรียนแต่ละคน เครื่องมือที่ใช้สำหรับประเมินแฟ้มสะสมงาน ได้แก่ แบบบันทึก แบบประเมินผลงาน และแบบประเมินตนเอง เป็นต้น

6) การประเมินโดยใช้ศูนย์ประเมิน

ศูนย์ประเมิน คือสถานที่ หรือคอมพิวเตอร์และซอฟต์แวร์ ที่สร้างหรือกำหนดขึ้นเพื่อให้สำหรับทดสอบหรือประเมินผู้เรียนภายใต้สถานการณ์จำลอง หรือสิ่งเร้า เพื่อให้ผู้เรียนตอบสนองโดยการแสดงออกตามพฤติกรรมบ่งชี้ การประมวลความรู้ และทักษะต่างของผู้เรียนว่ามีมากน้อยเพียงใด และอยู่ในระดับใด กิจกรรมหรือสถานการณ์ที่กำหนดให้มีหลากหลาย ได้แก่ เกม แบบฝึกขั้นตอนการทำงาน ใบงาน การสนทนากลุ่ม การทำงานเป็นกลุ่ม และการแสดงบทบาทสมมติ การนำเสนองาน เครื่องมือที่ใช้วัด ประกอบด้วย แบบตรวจสอบรายการ แบบมาตรวัดประเมินค่า แบบบันทึกพฤติกรรม แบบประเมินผลงาน และแบบทดสอบ เป็นต้น

บรรณานุกรม

- 1.ภัทธา นิคมานนท์. การประเมินผลการเรียน. กรุงเทพฯ: อักษราพัฒนาจำกัด, 2540.
- 2.ไพศาล สุวรรณน้อย (บรรณาธิการ) คู่มือการพัฒนาการเรียนการสอนมหาวิทยาลัยขอนแก่น. ขอนแก่น. มหาวิทยาลัยขอนแก่น ฝ่ายวิชาการ. 2545.
- 3.สมศักดิ์ ภูวิภาดาธรรม. การยึดผู้เรียนเป็นศูนย์กลางและการประเมินตามสภาพจริง. เชียงใหม่: The Knowledge Center, 2544.
- 4.สมชาย รัตนทองคำ. 2545. การพัฒนารูปแบบการสอนที่เน้นกระบวนการคิดอย่างมีวิจารณญาณสำหรับ นักศึกษากายภาพบำบัด มหาวิทยาลัยขอนแก่น. วิทยานิพนธ์ ปริญญาศึกษาศาสตรดุษฎีบัณฑิต สาขาหลักสูตรและการสอน บัณฑิตวิทยาลัย มหาวิทยาลัยขอนแก่น.