เอกสารประกอบการสอน
รายวิชา การวัดและประเมินผลการเรียนรู้
EDZ2101 Learning Measurement and Evaluation

ผู้ช่วยศาสตราจารย์ ดร.สุทธิพงศ์ บุญผดุง

Ph.D. (Curriculum & Instruction)

คณะครุศาสตร์ มหาวิทยาลัยราชภัฏสวนสุนันทา

ภาคเรียนที่ 1 ปีการศึกษา 2569

เอกสาร

การวัดผลทางการศึกษา

การวัดผลทางการศึกษาเป็นเรื่องสำคัญและจำเป็นต่อผู้สอน คุรุสภาจึงได้กำหนดมาตรฐานความรู้และสมรรถนะสำหรับผู้ประกอบวิชาชีพครูที่สำคัญประการหนึ่งได้แก่ การวัดและประเมินผลการศึกษา กล่าวคือ ผู้สอนสามารถวัดและประเมินผลได้ตามสภาพความเป็นจริง และสามารถนำผลการประเมินไปใช้ในการปรับปรุงการจัดการเรียนรู้และหลักสูตร (สำนักมาตรฐานวิชาชีพ กลุ่มมาตรฐานและจรรยาบรรณวิชาชีพ, 2550: 44-50) ดังนั้นผู้สอนจำเป็นต้องมีความรู้ ทักษะและเจตคติที่ดีต่อการวัดและประเมินผลการเรียนรู้ของผู้เรียน ในบทนี้จะกล่าวถึงการวัดผลทางการศึกษาซึ่งเป็นพื้นฐานสำคัญที่จะนำไปใช้และเป็นประโยชน์ต่อการจัดการศึกษา โดยหัวข้อย่อยที่เกี่ยวข้องประกอบด้วย หลักการและแนวคิดของการวัดผลทางการศึกษา ความสำคัญของการวัดผลทางการศึกษากระบวนการการวัดผลทางการศึกษา และเทคนิคการวัดผลทางการศึกษา

หลักการและแนวคิดของการวัดผลทางการศึกษา

1. ความหมายและแนวคิดของการวัดผลทางการศึกษา

การวัดผลทางการศึกษา ตรงกับภาษาอังกฤษว่า Educational Measurement เป็นกระบวนการสังเกตและบันทึกข้อมูลที่ได้จากผู้เรียน ซึ่งผลจากการวัดสะท้อนถึงปริมาณและกำหนดเป็นตัวเลขได้ หรืออาจกล่าวอีกนัยหนึ่งก็คือการวัดผลทางการศึกษาเป็นการตอบคำถามที่ว่า “How much?” (Gronlund and Linn, 1990) เนื่องจากการวัดผลทางการศึกษาเป็นการวัดทางอ้อม ดังนั้นสิ่งที่ได้จากการวัดที่ออกมาเป็นตัวเลขมักจะมีความคลาดเคลื่อนอยู่เสมอ ซึ่งหลักของการวัดผลจะพยายามทำให้ความคลาดเคลื่อนเป็นศูนย์ หรือปราศจากความคลาดเคลื่อนให้มากที่สุดเพื่อให้เกิดความชัดเจนเกี่ยวกับการวัดผลทางการศึกษา ขอให้ศึกษาตัวอย่างต่อไปนี้ “เด็กชายสุทธิพงศ์ ได้คะแนนศูนย์คะแนนจากการสอบวิชาภาษาไทย ดังนั้นจึงสรุปว่าเด็กชายสุทธิพงศ์ ไม่มีความรู้ภาษาไทยใช่หรือไม่”

ศูนย์คะแนนในที่นี้ไม่ได้แทนความไม่มี เพราะเป็นศูนย์สัมพัทธ์หรือศูนย์ไม่แท้ กล่าวคือการที่เด็กชายสุทธิพงศ์ได้ศูนย์คะแนน ไม่ได้สะท้อนว่าเขาไม่มีความรู้ภาษาไทย สาเหตุอาจมาจากเขาตอบ

ได้ไม่ตรงกับสิ่งที่เป็นคำถาม หรืออาจมาจากเขาไม่พร้อมกับการสอบครั้งนี้ เช่น ไม่สบาย หรือเกิดความวิตกกังวลเรื่องต่าง ๆ แต่ถ้าสาเหตุที่เขาได้ศูนย์คะแนนมาจากการที่ผู้สอนออกข้อสอบในสิ่งที่ไม่ได้สอน หรือสอนในสิ่งที่ไม่ได้ออกสอบ ภาษาทางการวัดผลเรียกลักษณะเช่นนี้ว่าข้อสอบขาดความเที่ยงตรงเชิงเนื้อหา(Content validity) กล่าวอีกนัยหนึ่งก็คือ ตัวคำถามที่ผู้สอนสร้างเพื่อใช้วัดความรู้ผู้เรียนไม่สอดคล้องกับวัตถุประสงค์การเรียนรู้ ซึ่งถ้าเป็นเช่นนี้จะสะท้อนถึงอะไร คำตอบคือผู้สอนไม่เข้าใจการสร้างข้อสอบหรือเปล่า หรือขาดความชัดเจนในเรื่องการวัดและประเมินผู้เรียนกันแน่ ซึ่งจะขอกล่าวในภายหลังต่อไป

สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (2553: 44-50) กล่าวว่าการวัดและประเมินผลการเรียนรู้ของผู้เรียนต้องอยู่บนจุดมุ่งหมายพื้นฐานสองประการประการแรกคือการวัดและประเมินผลเพื่อพัฒนาผู้เรียนโดยเก็บรวบรวมข้อมูลเกี่ยวกับผลการเรียนและการเรียนรู้ของผู้เรียนในระหว่างการเรียนการสอนอย่างต่อเนื่องบันทึกวิเคราะห์แปลความหมายข้อมูลแล้วนำมาใช้ในการส่งเสริมหรือปรับปรุงแก้ไขการเรียนรู้ของผู้เรียนและการสอนของผู้สอน การวัดและประเมินผลกับการสอนจึงเป็นเรื่องที่สัมพันธ์กันหากขาดสิ่งหนึ่งสิ่งใดการเรียนการสอนก็ขาดประสิทธิภาพการประเมินระหว่างการเรียนการสอนเพื่อพัฒนาการเรียนรู้เช่นนี้เป็นการวัดและประเมินผลย่อย (Formative Assessment) ที่เกิดขึ้นในห้องเรียนทุกวันเป็นการประเมินเพื่อให้รู้จุดเด่นจุดที่ต้องปรับปรุงจึงเป็นข้อมูลเพื่อใช้ในการพัฒนาในการเก็บข้อมูลผู้สอนต้องใช้วิธีการและเครื่องมือการประเมินที่หลากหลายเช่นการสังเกตการซักถามการระดมความคิดเห็นเพื่อให้ได้มติข้อสรุปของประเด็นที่กำหนดการใช้แฟ้มสะสมงานการใช้ภาระงานที่เน้นการปฏิบัติการประเมินความรู้เดิมการให้ผู้เรียนประเมินตนเองการให้เพื่อนประเมินเพื่อนและการใช้เกณฑ์การให้คะแนน (Rubrics) สิ่งสำคัญที่สุดในการประเมินเพื่อพัฒนาคือการให้ข้อมูลย้อนกลับแก่ผู้เรียนในลักษณะคำแนะนำที่เชื่อมโยงความรู้เดิมกับความรู้ใหม่ทำให้การเรียนรู้พอกพูนแก้ไขความคิดความเข้าใจเดิมที่ไม่ถูกต้องตลอดจนการให้ผู้เรียนสามารถตั้งเป้าหมายและพัฒนาตนได้จุดมุ่งหมายประการที่สองคือการวัดและประเมินผลเพื่อตัดสินผลการเรียนเป็นการประเมินสรุปผลการเรียนรู้(Summative Assessment) ซึ่งมีหลายระดับได้แก่เมื่อเรียนจบหน่วยการเรียนจบรายวิชาเพื่อตัดสินให้คะแนนหรือให้ระดับผลการเรียนให้การรับรองความรู้ความสามารถของผู้เรียนว่าผ่านรายวิชาหรือไม่ควรได้รับการเลื่อนชั้นหรือไม่หรือสามารถจบหลักสูตรหรือไม่ในการประเมินเพื่อตัดสินผลการเรียนที่ดีต้องให้โอกาสผู้เรียนแสดงความรู้ความสามารถด้วย

วิธีการที่หลากหลายและพิจารณาตัดสินบนพื้นฐานของเกณฑ์ผลการปฏิบัติมากกว่าใช้เปรียบเทียบระหว่างผู้เรียน

สำหรับการประเมินผล ตรงกับภาษาอังกฤษว่า Evaluation เป็นกระบวนการตัดสิน ลงความเห็น หรือตีคุณค่ากับสิ่งที่วัดผลมาว่าเหมาะ หรือไม่เหมาะ ดีหรือไม่ดีอย่างไร ซึ่งกรอนลันด์และลินน์ (Gronlund and Linn, 1990) ได้กล่าวว่าการประเมินผลเป็นการตอบคำถามที่ว่า “How good ? ” คำว่าการประเมิน และการประเมินผล เป็นคำที่หลายคนเกิดความสับสนในการใช้ จริงๆ แล้วการประเมิน ตรงกับภาษาอังกฤษว่า Assessment เป็นการที่ผู้สอนประเมินผู้เรียนเพื่อพัฒนาแสวงหาวิธีการช่วยเหลือและส่งเสริมผู้เรียนให้มีความรู้ตามมาตรฐาน สมรรถนะที่จำเป็น ตลอดจนคุณลักษณะที่พึงประสงค์ตามหลักสูตร ไม่ใช่เป็นการตัดสินผู้เรียนแต่อย่างใด หรือที่หลายคนเรียกว่าประเมินเพื่อฟันธง ถ้าเป็นอย่างนั้นการประเมินหรือ Assessment ควรทำในช่วงระหว่างเรียนอย่างสม่ำเสมอและต่อเนื่อง เพราะหน้าที่ของผู้สอนนอกจากจะถ่ายเทความคิดที่ดีและมีหลักการให้กับผู้เรียนแล้วยังต้องทำหน้าที่ตรวจสอบความก้าวหน้า และพัฒนาการของผู้เรียนให้มีคุณภาพอย่างน้อยตามมาตรฐานการศึกษา และที่ขาดไม่ได้คือการไปสู่โลกของการประกอบอาชีพเพื่อการดำรงชีพในสังคมต่อไป เห็นได้ชัดว่าผู้ที่ทำหน้าที่ครูหรือผู้สอนจะต้องมีความรู้และความเข้าใจอย่างถ่องแท้ในเรื่องเกี่ยวกับการวัดและประเมินผู้เรียนอย่างหลีกเลี่ยงไม่ได้

การวัดผลเป็นการติดตามผู้เรียนว่ามีคุณลักษณะ ความรู้ ความสามารถ หรือทักษะตามจุดมุ่งหมายของการจัดการเรียนรู้มากน้อยเพียงใด มีข้อบกพร่องและข้อปรับปรุงอะไรบ้างที่ควรแก้ไขซึ่งการวัดผลจำเป็นต้องยอมรับความแตกต่างระหว่างบุคคล แตกต่างกันในเรื่องการเรียนรู้ กล่าวคือบุคคลเรียนรู้ได้ไม่เท่ากันในเวลาที่เท่ากัน และบุคคลเรียนรู้ได้เท่ากันในเวลาที่ไม่เท่ากันส่วนหน้าที่ของการวัดผลได้แก่ (1) วัดผลเพื่อปรับปรุงการเรียนการสอน เป็นการวัดผลเพื่อค้นหาข้อบกพร่องในการเรียนรู้ เป็นการช่วยเหลือ สอนเสริม แก้ไขข้อบกพร่องของผู้เรียน ดังนั้นหน้าที่ของการวัดผลนี้จะมีประโยชน์อย่างมากต่อผู้เรียน เมื่อผู้สอนทราบว่าผู้เรียนบกพร่องในเรื่องใด จะได้ให้ความช่วยเหลือแนะนำแก้ไขข้อบกพร่องในเรื่องต่าง ๆ เป็นระยะ ๆ กล่าวอีกนัยหนึ่งก็คือเป็นการวัดเพื่อปรับปรุงการจัดการเรียนรู้ ซึ่งจะทำในแต่ละเนื้อหาย่อยหรือวัดผลเป็นระยะ ๆและ (2) วัดผลเพื่อประเมินผลการเรียน เป็นการวัดผลเพื่อตรวจสอบผลการเรียนรู้ในภาพรวมว่าผู้เรียนแต่ละคน เก่งหรืออ่อน มีความสามารถหรือทักษะมากน้อยเพียงใดหรือมีคุณลักษณะตามที่หลักสูตรกำหนดไว้มากน้อยเพียงใด การวัดผลนี้จะดำเนินการเมื่อสิ้นสุดการจัดการเรียนรู้ทั้งหมด หรือเป็นการวัดผลปลายภาคเรียน

สรุปได้ว่า การวัดผลทางการศึกษาเป็นกระบวนการสังเกตและบันทึกข้อมูลที่ได้จากผู้เรียน ซึ่งผลจากการวัดสะท้อนถึงปริมาณและกำหนดเป็นตัวเลขได้ แต่เนื่องจากการวัดผลทางการศึกษาเป็นการวัดทางอ้อม ดังนั้นสิ่งที่ได้จากการวัดที่ออกมาเป็นตัวเลขมักจะมีความคลาดเคลื่อนอยู่เสมอ ซึ่งหลักของการวัดผลจะพยายามทำให้ความคลาดเคลื่อนเป็นศูนย์ หรือปราศจากความคลาดเคลื่อนให้มากที่สุด ที่สำคัญการวัดผลทางการศึกษาเป็นกระบวนการที่จะตรวจสอบคุณภาพของการเรียนการสอนว่าได้ช่วยให้ผู้เรียนบรรลุตามจุดมุ่งหมายที่วางไว้หรือไม่ ถ้าการวัดผลดังกล่าวพบว่ายังไม่เป็นไปตามที่วางไว้ ผู้สอนก็ควรพิจารณาว่ากระบวนการหรือขั้นตอนใดที่มีข้อบกพร่อง ควรปรับปรุงแก้ไขอย่างไรเห็นได้ว่าการวัดผลทางการศึกษาเป็นสิ่งที่จะต้องทำอยู่ตลอดเวลาควบคู่กับการเรียนการสอน ไม่ใช่เป็นกระบวนการขั้นสุดท้ายของการจัดการเรียนการสอนเท่านั้น ส่วนการประเมินผลทางการศึกษา เป็นกระบวนการตัดสินคุณค่าในบริบทขอบข่ายทางการศึกษา ซึ่งจะเน้นกระบวนการตัดสินลงความเห็น หรือตีคุณค่ากับสิ่งที่วัดผลมาว่าเหมาะ หรือไม่เหมาะ ดีหรือไม่ดีอย่างไรที่สำคัญควรทำในช่วงระหว่างเรียนอย่างสม่ำเสมอและต่อเนื่อง เห็นได้ชัดว่าผู้ที่ทำหน้าที่ครูหรือผู้สอนจะต้องมีความรู้และความเข้าใจอย่างถ่องแท้ในเรื่องเกี่ยวกับการวัดและประเมินผู้เรียนอย่างหลีกเลี่ยงไม่ได้

2. จุดมุ่งหมายของการวัดผลทางการศึกษา

การวัดผลและประเมินผลการศึกษาเกี่ยวข้องกับกระบวนการเรียนการสอนตลอดเวลาซึ่งจุดมุ่งหมายของการวัดผลและประเมินผลนั้นไม่ใช่เฉพาะการนำผลการวัดไปตัดสินได้-ตก หรือใครควรจะได้เกรดอะไรเท่านั้นแต่ควรนำผลการวัดและประเมินนี้ไปใช้ให้เกิดประโยชน์ต่อการศึกษาในหลาย ๆลักษณะ จุดมุ่งหมายของการวัดผลทางการศึกษา มีดังนี้ (สมนึก ภัททิพธนี, 2549: 8-9)

1. วัดผลเพื่อค้นและพัฒนาสมรรถภาพของผู้เรียนหมายถึงการวัดผลและประเมินผลเพื่อดูว่าผู้เรียนบกพร่องหรือไม่เข้าใจในเรื่องใดตอนใด แล้วผู้สอนพยายามสอนให้ผู้เรียนเกิดความรู้ มีความเจริญงอกงามตามศักยภาพของตนเองจุดมุ่งหมายข้อนี้สำคัญมาก หรืออาจกล่าวได้ว่าเป็น ปรัชญาการวัดผลการศึกษา

2. วัดผลเพื่อวินิจฉัย (Diagnostic) เป็นการวัดผลและประเมินผลที่มีจุดมุ่งหมายเพื่อหาความบกพร่องของผู้เรียนว่าวิชาที่เรียนนั้นมีจุดบกพร่องตอนใดเพื่อที่จะได้หาทางช่วยเหลือหรือนำไปปรับปรุงแก้ไข ซ่อมเสริมส่วนที่ขาดหายไปให้ดียิ่งขึ้นซึ่งในกระบวนการเรียนการสอนเรียกว่าการวัดผลย่อย (Formative measurement)

3. วัดผลเพื่อจัดอันดับหรือจัดตำแหน่ง (Placement) การวัดผลและประเมินผลวิธีนี้เพื่อเปรียบเทียบตนเองกับคนอื่นๆโดยอาศัยกลุ่มเป็นเกณฑ์ว่าใครเด่น-ด้อย ใครได้อันดับที่ 1 ใครสอบได้-ตกหรือใครควรได้เกรดอะไร เป็นต้นการวัดผลและประเมินผลวิธีนี้เหมาะสำหรับการตัดสินผลการเรียนแบบอิงกลุ่มและการคัดเลือกคนเข้าทำงาน

4. วัดผลเพื่อเปรียบเทียบหรือทราบพัฒนาการของผู้เรียน (Assessment) เป็นการวัดผลเพื่อเปรียบเทียบความสามารถของผู้เรียนเองเพื่อดูความงอกงามของผู้เรียนแต่ละคนในช่วงเวลาที่ต่างกันว่าเจริญงอกงามเพิ่มขึ้นกว่าเดิมมากน้อยเพียงใด เช่นการเปรียบเทียบผลก่อนเรียนและหลังเรียน

5. วัดผลเพื่อพยากรณ์ (Prediction) เป็นการวัดเพื่อนำผลที่ได้ไปคาดคะเนหรือทำนายเหตุการณ์ในอนาคตต่อไปว่าจะเป็นอย่างไรนั่นคือเมื่อผู้เรียนคนหนึ่งสอบแล้วสามารถรู้อนาคตได้เลยว่าถ้าการเรียนของผู้เรียนอยู่ในลักษณะนี้ต่อไปแล้วการเรียนจะประสบผลสำเร็จหรือไม่ซึ่งสามารถนำไปใช้ในเรื่องของการแนะแนวการศึกษาว่าผู้เรียนควรเรียนสาขาใดหรืออาชีพใดจึงจะเรียนได้สำเร็จแบบทดสอบที่ใช้วัดจุดมุ่งหมายในข้อนี้ ได้แก่แบบทดสอบวัดความถนัด (Aptitude test) แบบทดสอบวัดเชาว์ปัญญา (Intelligence test) เป็นต้น

6. วัดผลเพื่อประเมิน (Evaluation) เป็นการนำผลที่ได้จากการวัดไปเปรียบเทียบกับเกณฑ์ที่กำหนดไว้เพื่อตัดสินลงสรุปให้คุณค่าของการศึกษาว่ามีประสิทธิภาพสูงหรือต่ำ หลักสูตรหรือเครื่องมือที่ใช้ในการวัดผลว่าเหมาะสมหรือไม่ และควรปรับปรุงแก้ไขอย่างไร

3. ลักษณะการวัดทางการศึกษา

การวัดทางการศึกษาเป็นการวัดคุณลักษณะที่เป็นนามธรรม มีลักษณะการวัด ดังนี้

1. เป็นการวัดทางอ้อม คือ ไม่สามารถวัดคุณลักษณะที่ต้องการวัดได้โดยตรง ต้องนิยามคุณลักษณะดังกล่าวให้เป็นพฤติกรรมที่วัดได้ก่อน จากนั้นจึงวัดตามพฤติกรรมที่นิยาม เช่น การวัดความรับผิดชอบของผู้เรียน ต้องให้นิยามคุณลักษณะความรับผิดชอบเป็นพฤติกรรมที่วัดได้ โดยอาจจะแยกเป็นพฤติกรรมย่อย เช่น ไม่มาโรงเรียนสาย ทำงานทุกงานที่ได้รับมอบหมาย นำวัสดุอุปกรณ์การเรียนที่ผู้สอนสั่งมาครบทุกครั้ง ส่งงานหรือการบ้านตามเวลาที่กำหนด เป็นต้น

2. วัดได้ไม่สมบูรณ์ การวัดทางการศึกษาไม่สามารถทำการวัดคุณลักษณะที่ต้องการวัดได้ครบถ้วนสมบูรณ์ วัดได้เพียงบางส่วน หรือวัดได้เฉพาะตัวแทนของคุณลักษณะทั้งหมด เช่นการวัด

ความสามารถการอ่านคำของผู้เรียน ผู้วัดไม่สามารถนำคำทุกคำมาทำการทดสอบผู้เรียน ทำได้เพียงนำคำส่วนหนึ่งที่คิดว่าเป็นตัวแทนของคำทั้งหมดมาทำการวัด เป็นต้น

3. มีความคลาดเคลื่อน สืบเนื่องจากการที่ไม่สามารถวัดได้โดยตรง และการนิยามสิ่งที่ต้องการวัดก็ไม่สามารถนิยามให้เป็นพฤติกรรมที่วัดได้ทั้งหมด จึงวัดได้ไม่สมบูรณ์ ตัวเลขหรือสัญลักษณ์ที่ได้จากการวัดเป็นการประมาณคุณลักษณะที่ต้องการวัด ซึ่งในความเป็นจริงคุณลักษณะดังกล่าวอาจจะมีมากหรือน้อยกว่า ผลการวัดจึงมีความผิดพลาดของการวัด หรือคลาดเคลื่อนจากความเป็นจริง การวัดที่ดีจะต้องให้เกิดการผิดพลาดหรือคลาดเคลื่อนน้อยที่สุด

4. อยู่ในรูปความสัมพันธ์ การที่จะรู้ความหมายของตัวเลขที่วัดได้ ต้องนำตัวเลขดังกล่าวไปเทียบกับเกณฑ์หรือเทียบกับคนอื่น เช่น นำคะแนนที่ผู้เรียนสอบได้เทียบกับคะแนนเฉลี่ยของกลุ่ม เทียบกับคะแนนของเพื่อนที่สอบพร้อมกัน หรือเทียบกับคะแนนของผู้เรียนเองกับการสอบครั้งก่อนๆ ถ้าคะแนนสูงกว่าเพื่อน แสดงว่ามีความสามารถในเรื่องที่วัดมากกว่าเพื่อนคนนั้น หรือถ้ามีคะแนนสูงกว่าคะแนนที่ตนเองเคยสอบผ่านมา แสดงว่ามีพัฒนาการขึ้น เป็นต้น

4. หลักการวัดผลทางการศึกษา

การวัดผลทางการศึกษา มีหลักการเบื้องต้น ดังนี้

1. นิยามสิ่งที่ต้องการวัดให้ชัดเจน ดังที่กล่าวไว้ในลักษณะการวัดว่า การวัดผลทางการศึกษาเป็นการวัดทางอ้อม การที่จะวัดให้มีคุณภาพต้องนิยามคุณลักษณะที่ต้องการวัดให้ตรงและชัดเจน การนิยามนี้มีความสำคัญมาก ถ้านิยามไม่ตรงหรือไม่ถูกต้อง เครื่องมือวัดที่สร้างตามนิยามก็ไม่มีคุณภาพ ผลการวัดก็ผิดพลาด คือวัดได้ไม่ตรงกับคุณลักษณะที่ต้องการวัด

2. ใช้เครื่องมือวัดที่มีคุณภาพ หัวใจสำคัญของการวัด คือ สามารถวัดคุณลักษณะได้ตรงตามกับที่ต้องการวัดและวัดได้แม่นยำ โดยใช้เครื่องมือวัดที่มีคุณภาพ คุณภาพของเครื่องมือมีหลายประการที่สำคัญคือ มีความเที่ยงตรง (validity) คือวัดได้ตรงกับคุณลักษณะที่ต้องการวัด และมีความเชื่อมั่น (reliability) คือวัดได้คงที่ คือวัดได้กี่ครั้งก็ให้ผลการวัดที่ไม่เปลี่ยนแปลง

3. กำหนดเงื่อนไขของการวัดให้ชัดเจน คือกำหนดให้แน่นอนว่าจะทำการวัดอะไร วัดอย่างไร กำหนดตัวเลขและสัญลักษณ์อย่างไร

5. ระดับของมาตรการวัด (Level of Measurement)

ระดับของการวัดเป็นการกำหนดตัวเลข หรือสัญลักษณ์อย่างมีระบบให้กับสิ่งของหรือเหตุการณ์เพื่อแทนปริมาณหรือคุณภาพของสิ่งที่ต้องการจะวัด เพื่อนำมาใช้ในการพิจารณาหรือประเมินค่า เรียกวิธีการกำหนดตัวเลขนั้นว่า มาตรในการวัด แบ่งเป็น 4 ระดับด้วยกัน (การวัดผลทางการศึกษามีเพียง 3 ระดับแรก) ดังนี้ (ชูศรี วงศ์รัตนะ, 2541: 6; สมนึก ภัททิพธนี, 2549: 13-15)

ระดับที่ 1 มาตรนามบัญญัติ (Nominal Scale) เป็นการกำหนดชื่อให้กับเรื่องราวหรือสิ่งของต่างๆ เช่น เพศชาย เพศหญิง ชั้นประถมศึกษา อาชีพรับราชการ ชาวนา ชาวสวน เป็นต้น มาตรวัดระดับนี้เป็นเพียงแต่การจัดหมวดหมู่ สิ่งที่มีลักษณะเดียวกัน เข้าเป็นประเภทเดียวกันเท่านั้นซึ่งนับว่าเป็นมาตรการวัดที่หยาบที่สุด หรือไม่ละเอียดถี่ถ้วนเพราะเป็นเพียงการเรียกชื่อหรือจำแนกชนิดหรือให้สัญลักษณ์สิ่งต่าง ๆ ไม่สามารถบอกปริมาณมากน้อยได้ กล่าวคือตัวเลขเหล่านี้ไม่สามารถนำมาเปรียบเทียบกันได้ หรือบอกว่ามาก น้อย สูงหรือต่ำแต่อย่างใด นั่นคือไม่มีความหมายในเชิงปริมาณ ที่สำคัญไม่สามารถนำตัวเลขเหล่านี้มาบวก ลบ คูณ หารได้ สำหรับวิธีการทางสถิติพื้นฐานที่ใช้วัดกับข้อมูลในระดับนี้ ได้แก่ ฐานนิยม เพื่อดูความถี่หรือการเกิดซ้ำ ๆ ของข้อมูล

ระดับที่ 2 มาตรการจัดอันดับ (Ordinal Scale) เป็นการจัดเรียงลำดับของสิ่งที่อยู่ในหมวดหรือสกุลเดียวกัน ให้ลดหลั่นตามลำดับปริมาณหรือคุณภาพ เช่น เก่ง ค่อนข้างเก่ง ปานกลาง หรือการจัดอันดับที่ 1, 2, 3 เป็นต้น ซึ่งการจัดอันดับในมาตรวัดนี้ เป็นเพียงการบอกปริมาณและคุณภาพที่มีมากน้อยต่างกันเท่านั้น แต่ไม่สามารถบอกได้ว่าแต่ละอันดับนั้นแตกต่างกันมากน้อยเพียงไร และแต่ละชั้นนั้นห่างเท่ากันหรือไม่ สำหรับวิธีการทางสถิติพื้นฐานที่ใช้วัดกับข้อมูลในระดับนี้ ได้แก่ มัธยฐาน เปอร์เซ็นไทล์ และพิสัย

ระดับที่ 3 มาตรอันตรภาค (Interval Scale) การกำหนดตัวเลขการวัดในระดับนี้ เป็นการแบ่งของที่อยู่ในลักษณะเดียวกัน ออกเป็นประเภทหรือช่วงที่เท่า ๆ กัน เช่น 1 ฟุต มี 12 นิ้ว, 1 วัน มี 24 ชั่วโมง, ซึ่งการแบ่งเป็นช่วงเท่า ๆ กัน ในลักษณะนี้ทำให้ข้อมูลที่ได้จากการวัดมีความหมายยิ่งขึ้น เพราะสามารถนำผลที่ได้ไปเปรียบเทียบ บวก ลบ กันได้ แต่มาตรการวัดระดับนี้ยังมีจุดอ่อนอยู่ตรงที่ว่า ศูนย์ที่ได้จากการวัดในมาตรวัดนี้ยังไม่ใช่ศูนย์แท้ (Absolute Zero) แต่เป็นเพียงศูนย์สมมติ(Arbitrary Zero) เท่านั้น เพราะศูนย์ในมาตรวัดนี้มิได้หมายความว่าไม่มีอะไรเลย แต่เป็นการสมมติว่าไม่มี แท้จริงยังมีลักษณะที่วัดนั้นอยู่บ้างแต่เราไม่ได้วัดการคำนวณทางคณิตศาสตร์ใช้ได้เฉพาะการบวกและลบเท่านั้น ในเรื่องการคูณและหารยังใช้ไม่ได้ เช่น อุณหภูมิที่ 30 องศาเซลเซียส

จะร้อนเป็นสองเท่าของอุณหภูมิที่ 15 องศาเซลเซียส ย่อมไม่ได้ สำหรับวิธีการทางสถิติพื้นฐานที่ใช้วัดกับข้อมูลในระดับนี้ ได้แก่ ค่าเฉลี่ยเลขคณิต และส่วนเบี่ยงเบนมาตรฐาน

ระดับที่ 4 มาตรอัตราส่วน (Ratio Scale) เป็นมาตรการวัดที่สมบูรณ์ คุณสมบัติ ของการกำหนดตัวเลขในมาตรวัดนี้คือ มีศูนย์แท้และมีช่วงคะแนนที่เท่ากัน เรียงขึ้นลงตามลำดับ สม่ำเสมอ เช่น การวัดความยาวของสิ่งของ ความยาวเท่ากับ ศูนย์หมายความว่า ไม่มีความยาว น้ำหนักศูนย์แสดงว่าไม่มีน้ำหนักอะไรเลย น้ำหนัก 6 กิโลกรัม แสดงว่าน้ำหนักเป็น 2 เท่า ของสิ่งของที่มีน้ำหนัก 3 กิโลกรัม เป็นต้นแต่จะไม่มีใช้ในการวัดผลทางการศึกษา เพราะการวัดผลทางการศึกษาไม่มีศูนย์แท้ สำหรับวิธีการทางสถิติที่ใช้วัดกับข้อมูลในระดับนี้ ใช้ได้ทุกอย่างเท่าที่มีอยู่

ความสำคัญของการวัดผลการศึกษา

การวัดและประเมินผลการเรียนรู้ของผู้เรียน เป็นองค์ประกอบสำคัญในการพัฒนาคุณภาพการศึกษา ทำให้ได้ข้อมูลสารสนเทศที่จำเป็นในการพิจารณาว่าผู้เรียนเกิดคุณภาพการเรียนรู้ตามผลการเรียนรู้ที่คาดหวังและมาตรฐานการเรียนรู้

จากประเภทของการประเมินโดยเฉพาะการแบ่งประเภทโดยใช้จุดประสงค์ของการประเมินเป็นเกณฑ์ในการแบ่งประเภท จะเห็นว่าการวัดและประเมินผลการเรียนนอกจากจะมีประโยชน์โดยตรงต่อผู้เรียนแล้ว ยังสะท้อนถึงประสิทธิภาพการสอนของผู้สอน และเป็นข้อมูลสำคัญที่สะท้อนคุณภาพการดำเนินงานการจัดการศึกษาของสถานศึกษาด้วย ดังนั้นผู้สอนและสถานศึกษาต้องมีข้อมูลผลการเรียนรู้ของผู้เรียน ทั้งจากการประเมินในระดับชั้นเรียน ระดับสถานศึกษา และระดับอื่นที่สูงขึ้น ประโยชน์ของการวัดและการประเมินผลการเรียนรู้จำแนกเป็นด้านๆ ดังนี้

1. ด้านการจัดการเรียนรู้

การวัดและประเมินผลการเรียนรู้ของผู้เรียนมีประโยชน์ต่อการจัดการเรียนรู้หรือการจัดการเรียนการสอนดังนี้

1.1 เพื่อจัดตำแหน่ง (Placement) ผลจากการวัดบอกได้ว่าผู้เรียนมีความรู้ความสามารถอยู่ในระดับใดของกลุ่มหรือเปรียบเทียบกับเกณฑ์แล้วอยู่ในระดับใด การวัดและประเมินเพื่อจัดตำแหน่งนี้ มักใช้ในวัตถุประสงค์ 2 ประการคือ

1.1.1 เพื่อคัดเลือก (Selection) เป็นการใช้ผลการวัดเพื่อคัดเลือกเพื่อเข้าเรียน เข้าร่วมกิจกรรมโครงการ หรือเป็นตัวแทน(เช่นของชั้นเรียนหรือสถานศึกษา) เพื่อการทำกิจกรรมหรือการให้ทุนการวัดและประเมินผลลักษณะนี้คำนึงถึงการจัดอันดับที่เป็นสำคัญ

1.1.2 เพื่อแยกประเภท (Classification) เป็นการใช้ผลการวัดและประเมินเพื่อแบ่งกลุ่มผู้เรียน เช่น แบ่งเป็นกลุ่มอ่อน ปานกลาง และเก่ง แบ่งกลุ่มผ่าน-ไม่ผ่านเกณฑ์ หรือตัดสินได้หรือตก เป็นต้น เป็นการวัดและประเมินที่ยึดเกณฑ์ที่ใช้ในการแบ่งกลุ่มเป็นสำคัญ

1.2 เพื่อวินิจฉัย (Diagnostic) เป็นการใช้ผลการวัดและประเมินเพื่อค้นหาจุดเด่นและ จุดด้อยของผู้เรียนว่ามีปัญหาในเรื่องใด จุดใด มากน้อยแค่ไหน เพื่อนำไปสู่การตัดสินใจการวางแผนการจัดการเรียนรู้และการปรับปรุงการเรียนการสอนให้มีประสิทธิภาพยิ่งขึ้น เครื่องมือที่ใช้วัดเพื่อการวินิจฉัย เรียกว่า แบบทดสอบวินิจฉัย (Diagnostic Test) หรือแบบทดสอบวินิจฉัยการเรียน ประโยชน์ของการวัดและประเมินประเภทนี้นำไปใช้ในวัตถุประสงค์ 2 ประการดังนี้

1.2.1 เพื่อพัฒนาการเรียนรู้ของผู้เรียน ผลการวัดผู้เรียนด้วยแบบทดสอบวินิจฉัยการเรียนจะทำให้ทราบว่าผู้เรียนมีจุดบกพร่องจุดใด มากน้อยเพียงใด ซึ่งผู้สอนสามารถแก้ไขปรับปรุงโดยการสอนซ่อมเสริม (Remedial Teaching) ได้ตรงจุด เพื่อให้ผู้เรียนเกิดการเรียนรู้ตามมาตรฐานการเรียนรู้ที่คาดหวังไว้

1.2.2 เพื่อปรับปรุงการจัดการเรียนรู้ ผลการวัดด้วยแบบทดสอบวินิจฉัยการเรียนนอกจากจะช่วยให้เห็นว่าผู้เรียนมีจุดบกพร่องเรื่องใดแล้ว ยังช่วยให้เห็นจุดบกพร่องของกระบวนการจัดการเรียนรู้อีกด้วย เช่น ผู้เรียนส่วนใหญ่มีจุดบกพร่องจุดเดียวกัน ผู้สอนต้องทบทวนว่าอาจจะเป็นเพราะวิธีการจัดการเรียนรู้ไม่เหมาะสมต้องปรับปรุงแก้ไขให้เหมาะสม

1.3เพื่อตรวจสอบและปรับปรุงการประเมินเพื่อพัฒนา (Formative Evaluation) เป็นการประเมินเพื่อตรวจสอบผลการเรียนรู้เทียบกับจุดประสงค์หรือผลการเรียนรู้ที่คาดหวัง ผลจากการประเมินใช้พัฒนาการจัดการเรียนรู้ให้มีประสิทธิภาพยิ่งขึ้น โดยอาจจะปรับปรุงหรือปรับเปลี่ยนวิธีการสอน (Teaching Method) ปรับเปลี่ยนสื่อการสอน(Teaching Media) ใช้นวัตกรรมการจัดการเรียนรู้ (Teaching Innovation) เพื่อนำไปสู่การพัฒนาการจัดการเรียนรู้ที่มีประสิทธิภาพ

1.4 เพื่อการเปรียบเทียบ (Assessment) เป็นการใช้ผลการวัดและประเมินเปรียบเทียบว่าผู้เรียนมีพัฒนาการจากเดิมเพียงใด และอยู่ในระดับที่พึงพอใจหรือไม่

1.5 เพื่อการตัดสินการประเมินเพื่อการตัดสินผลการเรียนของผู้เรียนเป็นการประเมินรวม (Summative Evaluation)คือใช้ข้อมูลที่ได้จากการวัดเทียบกับเกณฑ์เพื่อตัดสินผลการเรียนว่าผ่าน-ไม่ผ่าน หรือให้ระดับคะแนน

2. ด้านการแนะแนว

ผลจากการวัดและประเมินผู้เรียน ช่วยให้ทราบว่าผู้เรียนมีปัญหาและข้อบกพร่องในเรื่องใดมากน้อยเพียงใด ซึ่งสามารถแนะนำและช่วยเหลือผู้เรียนให้แก้ปัญหา มีการปรับตัวได้ถูกต้องตรงประเด็น นอกจากนี้ผลการวัดและประเมินยังบ่งบอกความรู้ความสามารถ ความถนัด และความสนใจของผู้เรียน ซึ่งสามารถนำไปใช้แนะแนวการศึกษาต่อและแนะแนวการเลือกอาชีพให้แก่ผู้เรียนได้

3. ด้านการบริหาร

ข้อมูลจากการวัดและประเมินผู้เรียน ช่วยให้ผู้บริหารเห็นข้อบกพร่องต่างๆ ของการจัดการเรียนรู้ เป็นการประเมินผลการปฏิบัติงานของผู้สอน และบ่งบอกถึงคุณภาพการจัดการศึกษาของสถานศึกษา ผู้บริหารสถานศึกษามักใช้ข้อมูลได้จากการวัดและประเมินใช้ในการตัดสินใจหลายอย่าง เช่น การพัฒนาบุคลากร การจัดครูเข้าสอน การจัดโครงการ การเปลี่ยนแปลงโปรแกรมการเรียน นอกจากนี้การวัดและประเมินผลยังให้ข้อมูลที่สำคัญในการจัดทำรายงานการประเมินตนเอง (SSR) เพื่อรายงานผลการจัดการศึกษาสู่ผู้ปกครอง สาธารณชน หน่วยงานต้นสังกัด และนำไปสู่การรองรับการประเมินภายนอก จะเห็นว่าการวัดและประเมินผลการศึกษาเป็นหัวใจสำคัญของระบบการประกันคุณภาพทั้งภายในและภายนอกสถานศึกษา

4. ด้านการวิจัย

การวัดและประเมินผลมีประโยชน์ต่อการวิจัยหลายประการดังนี้

4.1 ข้อมูลจาการวัดและประเมินผลนำไปสู่ปัญหาการวิจัย เช่น ผลจากการวัดและประเมินพบว่าผู้เรียนมีจุดบกพร่องหรือมีจุดที่ควรพัฒนาการแก้ไขจุดบกพร่องหรือการพัฒนาดังกล่าวโดยการปรับเปลี่ยนเทคนิควิธีสอนหรือทดลองใช้นวัตกรรมโดยใช้กระบวนการวิจัย การวิจัยดังกล่าวเรียกว่าการวิจัยในชั้นเรียน (Classroom Research) นอกจากนี้ผลจากการวัดและประเมินยังนำไปสู่การวิจัยในด้านอื่น ระดับอื่น เช่น การวิจัยของสถานศึกษาเกี่ยวกับการทดลองใช้รูปแบบการพัฒนาคุณลักษณะของผู้เรียน เป็นต้น

4.2 การวัดและประเมินเป็นเครื่องมือของการวิจัย การวิจัยใช้การวัดในการรวบรวมข้อมูลเพื่อศึกษาผลการวิจัย ขั้นตอนนี้เริ่มจากการหาหรือสร้างเครื่องมือวัด การทดลองใช้เครื่องมือ การหา

คุณภาพเครื่องมือ จนถึงการใช้เครื่องมือที่มีคุณภาพแล้วรวบรวมข้อมูลการวัดตัวแปรที่ศึกษา หรืออาจต้องตีค่าข้อมูล จะเห็นว่าการวัดและประเมินผลมีบทบาทสำคัญมากในการวิจัย เพราะการวัดไม่ดีใช้เครื่องมือไม่มีคุณภาพ ผลของการวิจัยก็ขาดความน่าเชื่อถือ

กระบวนการวัดผลทางการศึกษา

กระบวนการจัดการเรียนรู้ มีความเกี่ยวข้องกับลักษณะสำคัญ 3 ประการ ได้แก่ การกำหนดจุดประสงค์การเรียนรู้ การจัดประสบการณ์การเรียนรู้ให้กับผู้เรียนและการประเมิน โดยที่การประเมินเป็นส่วนหนึ่งของกิจกรรมการเรียนรู้ที่ใช้เวลาหนึ่งในสาม ถึงหนึ่งในสองของกิจกรรมการเรียนการสอนทั้งหมด (Thomas, 2012: 103)

โดยทั่วไป กระบวนการวัดผลทางการศึกษา มีลำดับขั้นตอนดังต่อไปนี้

1. การกำหนดสิ่งที่ต้องการจะประเมินเป็นการพิจารณาว่าคุณลักษณะหรือพฤติกรรมที่ต้องการประเมินมีอะไรบ้างซึ่งก็ต้องมีการวิเคราะห์หลักสูตรหรือวิเคราะห์เนื้อหาและพฤติกรรมของรายวิชาที่ต้องการประเมินออกมาให้ได้ว่ามีเนื้อหาอะไรมีพฤติกรรมอะไรบ้าง อย่างละเท่าไร

2. การกำหนดวิธีการวัดและเครื่องมือที่ใช้วัดผลเป็นการเลือกวิธีการและเครื่องมือวัดให้เหมาะสมกับพฤติกรรมที่ต้องการวัดแต่ละประเภทให้เหมาะสมและต้องสามารถวัดพฤติกรรมเหล่านั้นได้จริง

3. การสร้างหรือเลือกเครื่องมือเป็นการกำหนดสิ่งเร้าหรือหาสถานการณ์เพื่อนำไปใช้ทดสอบผู้เรียนให้แสดงพฤติกรรมตอบสนองออกมาในรูปข้อมูลเชิงปริมาณโดยที่เครื่องมือต้องมีคุณสมบัติที่ดี เช่น ต้องวัดในสิ่งที่ต้องการได้ตรงกับจุดมุ่งหมายให้ผลการวัดที่แน่นอนคงเส้นคงวาไม่เปลี่ยนแปลงมีความชัดเจน เข้าใจได้ตรงกันเป็นต้น

4. การดำเนินการทดสอบเป็นการนำเครื่องมือไปทดสอบเพื่อให้ผู้เรียนแสดงพฤติกรรมออกมาแล้วตีความหมายเป็นคะแนนผลการวัดหรือคะแนนที่ได้จะนำไปใช้ในการตัดสินผลการเรียนในการดำเนินการสอบนี้มีหลักสำคัญที่ต้องคำนึงถึงอย่างยิ่งก็คือต้องให้ผู้เรียนแสดงความสามารถสูงสุดของตนออกมาและต้องให้เกิดความยุติธรรมในการทดสอบ

5. การตรวจให้คะแนนเป็นการกำหนดตัวเลขแทนปริมาณของสิ่งที่ต้องการวัดตามกฎเกณฑ์ที่กำหนดไว้โดยพิจารณาผลการตอบสนองหรือพฤติกรรมที่แสดงออกมา

6. การประเมินผลเป็นขั้นตอนสุดท้ายในกระบวนการประเมินผล เป็นการลงสรุปจากข้อมูลที่วัดมาได้จากข้อ 5 โดยนำผลการวัดมาเทียบกับเกณฑ์ที่กำหนดแล้วลงสรุปว่าผู้เรียนได้หรือตกหรือได้ระดับคะแนนเท่าใดจากกระบวนในการประเมินผลดังกล่าวจากขั้นที่ 1 ถึงขั้นที่ 5 ก็คือขั้นตอนของการวัดผลส่วนข้อ 6 เป็นเรื่องของการประเมินผลจะเห็นว่ากระบวนในการประเมินผลขั้นตอนแรกก็คือกระบวนการในการวัดผลความถูกต้องของผลการประเมินจะต้องขึ้นอยู่กับความถูกต้องของผลการวัดเป็นสำคัญนอกจากนี้คุณธรรมในการตัดสินผล เช่นอคติของผู้ประเมินก็เป็นสิ่งสำคัญที่ส่งผลต่อความถูกต้องของผลการประเมินด้วย

ในขณะกระบวนการประเมินผลประกอบด้วยขั้นตอนที่ต่อเนื่องกัน 3 ประการ (พิสณุ ฟองศรี, 2551: 2-3) คือ

1. การวัดผล (Measurement) ซึ่งทำให้ทราบสภาพความเป็นจริงของสิ่งที่จะประเมินผลว่ามีปริมาณเท่าใดมีคุณสมบัติอย่างไร เพื่อเป็นข้อมูลสำหรับนำไปเปรียบเทียบกับเกณฑ์ (Criteria)

2. เกณฑ์ (Criteria) หรือมาตรฐานในการพิจารณาตัดสินว่าสิ่งใดดีหรือไม่ดี/เลว ใช้ได้หรือไม่ได้นั้นจะต้องมีหลักหรือบรรทัดฐานเพื่อนำผลที่ได้จากการวัดไปเปรียบเทียบ

3. การตัดสิน(Decision)เป็นการชี้ขาดระหว่างผลการปฏิบัติที่ได้จากการวัดกับเกณฑ์ที่กำหนดไว้ว่าควรสรุปผลออกมาอย่างไรเช่น ถ้าผลที่ได้จากการวัดถึงเกณฑ์ที่กำหนดไว้ ก็ ผ่านหรือถ้าต่ำกว่าเกณฑ์ที่กำหนดไว้ก็ ไม่ผ่าน เป็นต้น ทั้งนี้การตัดสินใจที่ดีควรจะอาศัยการพิจารณาอย่างถี่ถ้วนความยุติธรรม และสิ่งที่สำคัญคือคุณธรรมประจำใจ (Value judgment)

จะเห็นว่าการวัดผลและการประเมินผลมีความสัมพันธ์กันกล่าวคือ การวัดผลจะทำให้ได้ตัวเลข ปริมาณหรือรายละเอียดของคุณลักษณะหรือพฤติกรรมของบุคคลจากนั้นจะนำเอาผลการวัดนี้ไปพิจารณาเปรียบเทียบกับเกณฑ์ที่กำหนดไว้เพื่อตัดสินหรือลงสรุปเกี่ยวกับสิ่งนั้น ซึ่งเรียกว่าการประเมินผล

เทคนิคการวัดผลทางการศึกษา

เทคนิคที่ใช้ในการวัดผลทางการศึกษามีหลายชนิด มีลักษณะการใช้แตกต่างกันตามโอกาสหรือสถานการณ์ การที่มีเทคนิคการวัดผลทางการศึกษาหลายชนิด ก็เพื่อให้การวัดครอบคลุมด้านพุทธิพิสัย จิตพิสัย และทักษะพิสัย แม้ว่าปัจจุบันจะเน้นการวัดและประเมินตามสภาพจริง แต่ผู้สอนควรต้องมีความรู้ ความเข้าใจ และเลือกใช้เทคนิคการวัดผลทางการศึกษาอย่างเหมาะสม สำหรับ

เทคนิคการวัดผลทางการศึกษา พอสรุปได้ดังนี้ (สมนึก ภัททิพธนี, 2549: 32-64; สมชาย รัตนทองคำ , 2554: 138)

1. การสังเกต (Observation)

การสังเกตเป็นการวัดและประเมินที่มีรายการพฤติกรรมเป้าหมายที่ต้องการเก็บข้อมูลด้วยประสาทสัมผัสโดยเฉพาะอย่างยิ่งทางหูและตาเพื่อศึกษาพฤติกรรมที่มีความละเอียดชัดเจนของผู้เรียนในสภาพการณ์ต่าง ๆ ที่กำหนดสำหรับเครื่องมือที่ใช้ประกอบด้วยแบบตรวจสอบรายการแบบมาตรวัดประเมินค่าและแบบบันทึกซึ่งเป็นค้นหาความจริงบางประการโดยอาศัยประสาทสัมผัสของผู้สังเกตโดยตรง ทำให้ได้ข้อมูลแบบปฐมภูมิ ซึ่งเป็นข้อมูลที่น่าเชื่อถือสำหรับรูปแบบการสังเกตแบ่งออกได้เป็น 2 แบบ ดังนี้

1.1 การสังเกตโดยผู้สังเกตเข้าไปร่วมในเหตุการณ์หรือกิจกรรม เป็นการสังเกตที่ผู้สังเกตเข้าไปมีส่วนร่วมหรือคลุกคลีในหมู่ผู้ถูกสังเกต และอาจร่วมทำกิจกรรมด้วยกันโดยร่วมอยู่ในฐานะสมาชิกคนหนึ่ง ผู้สังเกตจะไม่สามารถจดบันทึกรายละเอียดเกี่ยวกับสิ่งที่ต้องการสังเกตได้ทันที และต้องใช้เวลาในการสังเกตนาน ๆ การบันทึกข้อความมักจะกระทำภายหลังการเข้าไปร่วมกิจกรรมนั้น ๆ และมักจะบันทึกในรูปของการสังเคราะห์ข้อความ จุดมุ่งหมายของการสังเกตในลักษณะนี้อีกข้อหนึ่ง คือ สังเกตโดยไม่ให้ผู้สังเกตรู้ตัวเพื่อช่วยให้การเก็บข้อมูลเป็นจริงหรือเป็นธรรมชาติมากที่สุด

1.2 การสังเกตโดยผู้สังเกตไม่ได้เข้าไปร่วมในเหตุการณ์หรือกิจกรรม เป็นการสังเกตที่ผู้สังเกตอยู่ภายนอกวงของผู้ถูกสังเกต โดยสังเกตในฐานะเป็นบุคคลภายนอก ไม่เข้าร่วมกระทำกิจกรรมกับผู้ถูกสังเกต การสังเกตในรูปแบบนี้แบ่งเป็น 2 ชนิด คือ

1.2.1 การสังเกตแบบไม่มีโครงสร้าง เป็นการสังเกตที่ผู้สังเกตไม่ได้กำหนดเรื่องเฉพาะที่มุ่งสังเกตเพียงอย่างเดียว แต่จะสังเกตเรื่องต่าง ๆ ที่เกี่ยวข้องด้วย ส่วนใหญ่ใช้กับการศึกษาแบบสำรวจเรื่องใหม่ทั่วไป ผู้ที่สังเกตไม่มีความรู้ภูมิหลังมาก่อน จึงไม่สามารถกำหนดรูปแบบที่แน่นอนได้

1.2.2 การสังเกตแบบมีโครงสร้าง เป็นการสังเกตที่ผู้สังเกตกำหนดเรื่องที่จะสังเกตเฉพาะไว้แล้ว โดยคาดว่าพฤติกรรมที่สังเกตจะเกิดขึ้นในช่วงเวลาที่ทำการสังเกต เช่น การสังเกตพฤติกรรมการเรียนของผู้เรียน จะสังเกตในลักษณะที่ผู้เรียนไม่รู้ตัว จะช่วยให้การจดบันทึกพฤติกรรมถูกต้องตามความเป็นจริงยิ่งขึ้น

2. การสัมภาษณ์ (Interview)

การสัมภาษณ์ เป็นวิธีการวัดผลด้วยการซักถามสนทนาโต้ตอบเพื่อประเมินความคิด ทัศนคติต่างๆ รูปแบบหรือเครื่องมือที่ใช้ได้แก่

2.1 แบบสัมภาษณ์แบบมีโครงสร้างเป็นการสัมภาษณ์ที่ผู้สัมภาษณ์จะใช้แบบสัมภาษณ์ที่สร้างขึ้นไว้แล้ว เป็นแบบในการถาม กล่าวคือ ผู้สัมภาษณ์จะใช้คำถามตามแบบสัมภาษณ์กับผู้ถูกสัมภาษณ์เหมือนกันหมดทุกคน การสัมภาษณ์แบบนี้มีลักษณะไม่ค่อยยืดหยุ่น เพราะต้องถามไปตามแบบสัมภาษณ์ แต่มีข้อดีคือสามารถจัดหมวดหมู่ข้อมูลได้ง่ายและสะดวกในการวิเคราะห์ สรุปผล

2.2 แบบสัมภาษณ์แบบไม่มีโครงสร้างเป็นการสัมภาษณ์ที่ไม่ใช้แบบสัมภาษณ์ คือไม่จำเป็นต้องใช้คำถามที่เหมือนกันหมดกับผู้ถูกสัมภาษณ์ทุกคน แต่ผู้สัมภาษณ์จะต้องใช้เทคนิคและความสามารถเฉพาะตัว เพื่อให้ได้คำตอบจากผู้ถูกสัมภาษณ์ตามจุดมุ่งหมายที่ตั้งไว้ การสัมภาษณ์โดยวิธีนี้อาจให้ผู้ตอบแสดงความรู้สึกหรือความคิดเห็นออกมาโดยอิสระ ผู้สัมภาษณ์มีหน้าที่รับฟังและคอยควบคุมให้เข้าประเด็นที่ต้องการเท่านั้น

3. แบบสอบถาม (Questionnaire)

แบบสอบถามเป็นเครื่องมือชนิดหนึ่งที่นิยมใช้กันมาก โดยเฉพาะการเก็บข้อมูลทางสังคมศาสตร์ ทั้งนี้เพราะเป็นวิธีการที่สะดวกและสามารถใช้ได้อย่างกว้างขวางทั้งข้อมูลหรือข้อเท็จจริงในอดีต ปัจจุบันและการคาดคะเนเหตุการณ์ในอนาคต แบบสอบถามส่วนใหญ่จะอยู่ในรูปของคำถามเป็นชุด ๆ เพื่อวัดสิ่งที่ต้องการจะวัด โดยมีคำถามเป็นตัวกระตุ้นเร่งเร้าให้บุคคลแสดงพฤติกรรมต่าง ๆ ออกมา ถือว่าเป็นเครื่องมือที่นิยมใช้วัดทางด้านจิตพิสัย

โดยทั่วไป รูปแบบของแบบสอบถามแบ่งเป็น 2 ชนิด คือ

3.1 แบบสอบถามชนิดปลายเปิด แบบสอบถามชนิดนี้ไม่ได้กำหนดคำตอบไว้ เพื่อเปิดโอกาสให้ผู้ตอบเขียนตอบอย่างอิสระด้วยความคิดของตนเอง แบบสอบถามชนิดนี้ตอบยากและเสียเวลาในการตอบมาก เพราะผู้ตอบต้องคิดวิเคราะห์อย่างกว้างขวาง และมักจะเว้นว่างไว้หรือตอบเพียงเล็กน้อยเท่านั้นเมื่อใช้ควบคู่กับแบบสอบถามอื่น ๆ แบบสอบถามชนิดนี้นิยมใช้เมื่อต้องการทราบเจตคติ แรงจูงใจ หรือเงื่อนไข ตลอดจนแนวคิดต่าง ๆ ที่เป็นสาเหตุของการตัดสินใจของผู้ตอบ และใช้เป็นแนวทางในการสร้างแบบสอบถามชนิดปลายปิด แบบสอบถามชนิดนี้สร้างง่ายแต่วิเคราะห์และสรุปผลยาก ตัวอย่างเช่น ท่านชอบวิชาวัดและประเมินผลการศึกษาในเรื่องใดบ้าง ท่านมีเหตุผลอะไรในการเลือกเรียนวิชาชีพครู

3.2 แบบสอบถามชนิดปลายปิด แบบสอบถามชนิดนี้ประกอบด้วยข้อคำถามและตัวเลือก/คำตอบ ซึ่งตัวเลือกนี้สร้างขึ้นโดยคาดว่าผู้ตอบสามารถเลือกตอบได้ตามความต้องการ แบบสอบถามชนิดนี้สร้างยากและใช้เวลาสร้างมากกว่าแบบสอบถามชนิดปลายเปิด แต่ผู้ตอบตอบง่ายสะดวกรวดเร็ว นอกจากนี้ข้อมูลที่ได้สามารถนำไปวิเคราะห์และสรุปผลง่าย แบบสอบถามชนิดปลายปิดแบ่งเป็น 4 แบบ ดังนี้

3.2.1 แบบตรวจสอบรายการ เป็นการสร้างรายการของข้อคำถามที่เกี่ยวหรือสัมพันธ์กับคุณลักษณะของพฤติกรรม หรือการปฏิบัติ แต่ละรายการจะถูกประเมินหรือชี้ให้ตอบในแง่ใดแง่หนึ่ง เช่น มี-ไม่มี จริง-ไม่จริง เห็นด้วย-ไม่เห็นด้วย หรืออาจมีคำตอบให้เลือกได้หลายคำตอบ ดังตัวอย่างต่อไปนี้

เพศ: ○ ชาย ○ หญิง
ที่บ้านมีโทรศัพท์หรือไม่ : ○ มี ○ ไม่มี
ระดับการศึกษา : ○ ประถมศึกษา ○ มัธยมศึกษา ○ ปริญญาตรี ○ สงกว่าปริญญาตรี

3.2.2มาตราส่วนประมาณค่า เป็นเครื่องมือในการประเมินผลที่ทั้งผู้สอนใช้ในการประเมินผู้เรียนและผู้เรียนใช้ในการประเมินหรือพิจารณาตนเองหรือสิ่งอื่นๆใช้ทั้งการประเมินการปฏิบัติกิจกรรม ทักษะต่างๆ และพฤติกรรมด้านจิตพิสัย เช่นเจตคติ แรงจูงใจใฝ่สัมฤทธิ์ ความสนใจ ฯลฯ มาตราส่วนประมาณค่าแตกต่างจากแบบตรวจสอบรายการตรงที่แบบตรวจสอบรายการต้องการทราบว่ามีหรือไม่มีในเรื่องนั้นแต่มาตราส่วนประมาณค่าต้องการทราบละเอียดยิ่งกว่านั้น กล่าวคือต้องการทราบว่ามีเพียงใดในการประเมินผลภาคปฏิบัติจะใช้มาตราส่วนประมาณค่าประเมินทั้งด้านกระบวนการ (Process) และผลที่ได้จากการปฏิบัติหรือผลผลิต (Product) ดังตัวอย่างตารางที่ 1.1 ต่อไปนี้

ตารางที่ 1.1 แสดงลักษณะของมาตราส่วนประมาณค่าที่ใช้ตัวเลือกเหมือนกัน

ข้อที่	ข้อคำถาม	เห็นด้วยอย่างยิ่ง	เห็นด้วย	ไม่แน่ใจ	ไม่เห็นด้วย	ไม่เห็นด้วยอย่างยิ่ง

1 การประเมินภาคปฏิบัติของผู้เรียน ในชั้นเรียนแบบไม่เป็นทางการเป็น วิธีการประเมินที่ดีที่สุด

2 การประเมินจะช่วยส่งเสริมให้ ผู้เรียนได้รับทราบและวิพากษ์การ ปฏิบัติของตนเอง

ตารางที่ 1.1 (ต่อ)

ข้อที่	ข้อคำถาม	เห็นด้วยอย่างยิ่ง	เห็นด้วย	ไม่แน่ใจ	ไม่เห็นด้วย	ไม่เห็นด้วยอย่างยิ่ง
3	การประเมินเป็นการบังคับให้ครูดำเนินการจัดการเรียนการสอนครบถ้วนและถูกต้อง					
4	การประเมินผู้เรียนเป็นความร่วมมือระหว่างครูกับผู้ปกครองของผู้เรียน					

3.2.3 แบบจัดอันดับความสำคัญ แบบสอบถามแบบนี้ต้องการให้ผู้ตอบตอบข้อที่เห็นว่าสำคัญโดยเรียงอันดับตามความสำคัญจากมากไปหาน้อยตามความรู้สึกของผู้ตอบตัวอย่างเช่น

ท่านเลือกเรียนครูเพราะเหตุใด โปรดเรียงอันดับตามความสำคัญของเหตุผลจากมากไปหาน้อย

เหตุผล อันดับความสำคัญเรียงจากมากที่สุด

1. มีใจรัก .....
2. หางานง่าย .....
3. ค่าใช้จ่ายถูก .....
4. ได้รับทุนอุดหนุน.....

3.2.4 แบบเติมคำสั้นๆ ในช่องว่าง แบบสอบถามแบบนี้ให้ผู้ตอบเติมข้อความลงในช่องว่างที่เว้นไว้ให้ควรกำหนดขอบเขตคำถามให้ชัดเจนจำเพาะเจาะจงลงไป หากสร้างคำถามไม่ชัดเจนอาจทำให้ผู้ตอบตีความหมายของคำถามไปคนละเรื่อง และตอบไม่ไปในทางเดียวกัน ตัวอย่างเช่น

ปัจจุบันท่านอายุ.....ปี.....เดือน

อาชีพหลักในปัจจุบันของท่านคือ.....

4. การจัดอันดับ (Rank Order)

การจัดอันดับ เป็นเครื่องมือมือวัดผลให้ผู้เรียนหรือผู้ได้รับแบบสอบถามเป็นผู้ตอบ โดยการจัดอันดับความสำคัญหรือจัดอันดับคุณภาพ และใช้จัดอันดับของข้อมูลหรือผลงานต่าง ๆของผู้เรียน แล้วจึงให้คะแนนภายหลังเพื่อการประเมิน

5. การประเมินผลจากสภาพจริง (Authentic Assessment)

การประเมินตามสภาพจริง หมายถึง กระบวนการสังเกตการบันทึก และรวบรวมข้อมูลจากงานและวิธีการที่ผู้เรียนทำการประเมินผลจากสภาพจริงจะเน้นให้ผู้เรียนสามารถแก้ปัญหาเป็นผู้ค้นพบและผู้ผลิตความรู้ ผู้เรียนได้ฝึกปฏิบัติจริงรวมทั้งเน้นพัฒนาการเรียนรู้ของผู้เรียนสำหรับความสำคัญของการประเมินผลจากสภาพจริงสรุปได้ดังนี้

5.1การเรียนการสอนและการวัดประเมินผลจากสภาพจริงจะเอื้อให้ผู้เรียนสามารถเรียนรู้ได้อย่างเต็มศักยภาพของแต่ละบุคคล

5.2เป็นการเอื้อต่อการเรียนการสอนที่เน้นผู้เรียนเป็นศูนย์กลาง

5.3 เป็นการเน้นให้ผู้เรียนได้สร้างงาน

5.4เป็นการผสมผสานให้กิจกรรมการเรียนรู้และการประเมินผล

5.5 เป็นการลดภาระงานซ่อมเสริมของผู้สอน

6. การวัดผลภาคปฏิบัติ (Performance Assessment)

การวัดผลภาคปฏิบัติ เป็นการวัดผลงานที่ให้ผู้เรียนลงมือปฏิบัติซึ่งสามารถวัดได้ทั้งกระบวนการและผลงานในสถานการณ์จริงหรือในสถานการณ์จำลองสิ่งที่ควรคำนึงในการสอบวัดภาคปฏิบัติคือขั้นเตรียมงานขั้นปฏิบัติงานเวลาที่ใช้ในการทำงานผลงาน

7. การประเมินผลโดยใช้แฟ้มสะสมผลงาน (Portfolios)

การประเมินผลโดยใช้แฟ้มสะสมผลงาน เป็นแนวทางการประเมินผลโดยการรวมข้อมูลที่ผู้สอนและผู้เรียนทำกิจกรรมต่าง ๆร่วมกัน โดยกระทำอย่างต่อเนื่องตลอดภาคเรียนโดยใช้การเก็บหลักฐานผลงานที่ดีและมีความภาคภูมิใจที่เป็นตัวแทนงานที่ปฏิบัติของผู้เรียนเกี่ยวกับทักษะแนวคิดความสนใจความสำเร็จโดยมีผลการประเมินจุดเด่นจุดที่ควรปรับปรุงของชิ้นงาน

8. แบบทดสอบ (Test)

แบบทดสอบเป็นเครื่องมือวัดผลที่มีความสำคัญมากในการประเมินผลผู้เรียน และต้องใช้แบบทดสอบควบคู่ไปกับเครื่องมือชนิดอื่น ๆ อย่างหลากหลาย สำหรับประเภทของแบบทดสอบสามารถแบ่งประเภทออกได้หลายลักษณะขึ้นอยู่กับเกณฑ์ที่จะใช้ในการแบ่ง ดังนี้

8.1 แบ่งตามสมรรถภาพที่จะวัดแบ่งเป็น 3 ประเภท

8.1.1 แบบทดสอบวัดผลสัมฤทธิ์ (Achievement Test) หมายถึงแบบทดสอบที่วัดสมรรถภาพสมองด้านต่าง ๆที่ผู้เรียนได้รับการเรียนรู้ผ่านมาแล้วว่ามีอยู่เท่าใด แบบทดสอบแบ่งออกเป็น 2 ชนิด

8.1.1.1 แบบทดสอบที่ครูสร้างขึ้น หมายถึงแบบทดสอบที่มุ่งวัดผลสัมฤทธิ์ของผู้เรียนเฉพาะกลุ่มที่ครูสอน

8.1.1.2 แบบทดสอบมาตรฐาน หมายถึงแบบทดสอบที่มุ่งวัดผลสัมฤทธิ์มีจุดมุ่งหมายเพื่อเปรียบเทียบคุณภาพต่าง ๆของผู้เรียนที่ต่างกลุ่มกัน เช่น แบบทดสอบมาตรฐานระดับชาติ

8.1.2 แบบทดสอบวัดความถนัด (Aptitude Test) หมายถึงแบบทดสอบที่มุ่งวัดสมรรถภาพสมองของผู้เรียนว่าจะสามารถเรียนต่อไปหรือจะประสบความสำเร็จเพียงใดเพื่อใช้ในการพยากรณ์หรือทำนายอนาคตของผู้เรียน แบบทดสอบวัดความถนัดแบ่งได้เป็น 3ชนิดคือ

8.1.2.1 แบบทดสอบวัดความถนัดทางการเรียน หมายถึงแบบทดสอบที่มุ่งวัดความถนัดทางด้านวิชาการต่าง ๆ เช่น ด้านภาษาด้านคณิตศาสตร์

8.1.2.2 แบบทดสอบวัดความถนัดเฉพาะ หมายถึงแบบทดสอบที่มุ่งวัดความถนัดเฉพาะที่เกี่ยวกับงานอาชีพต่าง ๆ หรือความสามารถพิเศษเช่น ความสามารถทางดนตรีศิลปะเป็นต้น

8.1.2.3แบบทดสอบวัดบุคลิกภาพทางสังคมหมายถึงแบบทดสอบที่ใช้วัดบุคลิกภาพและการปรับตัวให้เข้ากับสังคม ซึ่งเป็นเรื่องที่วัดได้ยาก ผลที่ได้ไม่คงที่แน่นอนตัวอย่างของแบบวัดชนิดนี้ได้แก่ (1) แบบวัดเจตคติที่มีต่อบุคคล สิ่งของเรื่องราวเหตุการณ์สังคม เป็นต้น (2) แบบวัดความสนใจที่มีต่ออาชีพ การศึกษากีฬาเป็นต้น และ (3) แบบวัดการปรับตัว เช่น การปรับตัวเข้ากับเพื่อน ๆความสามารถในการควบคุมอารมณ์ เป็นต้น

8.2 แบ่งตามลักษณะการตอบแบ่งเป็น 3 ประเภท ดังนี้

8.2.1 แบบทดสอบภาคปฏิบัติ หมายถึงแบบทดสอบที่ให้ผู้เรียนลงมือปฏิบัติจริง เช่น การปรุงอาหารการฝีมือ ศิลปะ

8.2.2 แบบทดสอบข้อเขียนหมายถึง แบบทดสอบที่ใช้การเขียนตอบสามารถตั้งคำถามให้เขียนตอบจำนวนมาก ๆ หรือลึกซึ้งเพียงใดก็ได้ ผู้ตอบมีโอกาสเรียบเรียงเนื้อหาและใช้ความสามารถทางสมองได้เต็มที่

8.2.3 แบบทดสอบปากเปล่า หมายถึงแบบทดสอบที่ใช้การพูดโต้ตอบแทนการเขียน เหมาะที่จะใช้กับผู้ที่อ่านไม่ออกเขียนไม่ได้ หรือเมื่อต้องการให้ตอบอย่างฉับไว ลักษณะเช่นนี้ควรให้ตอบสั้น ๆ และมีข้อสอบไม่มากนัก เพราะจะเสียเวลามากและต้องอาศัยการเรียบเรียงเนื้อหา

8.3 แบ่งตามเวลาที่กำหนดให้ตอบแบ่งเป็น 2 ประเภท ดังนี้

8.3.1แบบทดสอบที่จำกัดเวลาในการตอบ หมายถึงแบบทดสอบที่ใช้เวลาตอบน้อย แต่มีจำนวนข้อสอบมากและค่อนข้างง่าย เน้นความคล่องแคล่วรวดเร็วในการทำข้อสอบได้ถูกต้องมากกว่า

8.3.2 แบบทดสอบที่ไม่จำกัดเวลาในการตอบ หมายถึงแบบทดสอบที่ใช้เวลาตอบมากหรือไม่จำกัดเวลาในการตอบ แต่มีจำนวนข้อสอบน้อย เน้นให้ผู้เรียนทุกคนแสดงศักยภาพของตนเองอย่างเต็มที่ มักเป็นข้อสอบที่ต้องแสดงความคิดเห็นหรือให้วิเคราะห์ บางครั้งจะให้เปิดหนังสือควบคู่ไปกับการสอบ หรือบางครั้งให้ไปตอบที่บ้าน

8.4แบ่งตามจำนวนผู้เข้าสอบแบ่งเป็น 2 ประเภท ดังนี้

8.4.1 แบบทดสอบเป็นรายบุคคล หมายถึง การสอบทีละคนมักเป็นการสอบภาคปฏิบัติจะใช้เมื่อคนเข้าสอบมีจำนวนน้อยหรือต้องการดูพฤติกรรมของผู้เรียนทุกขั้นตอน

8.4.2 แบบทดสอบเป็นชั้นหรือเป็นหมู่ หมายถึง การสอบทีละหลาย ๆ คนจะใช้เมื่อมีคนเข้าสอบจำนวนมากและสามารถจัดสอบพร้อมกันทั้งโรงเรียน จังหวัดหรือประเทศก็ได้

8.5 แบ่งตามสิ่งเร้าของการถามแบ่งเป็น 2 ประเภท ดังนี้

8.5.1 แบบทดสอบทางภาษา หมายถึงแบบทดสอบที่ต้องอาศัยภาษาของสังคมนั้น ๆ เป็นหลักใช้กับผู้ที่อ่านออกเขียนได้แบบทดสอบประเภทนี้มีระเบียบวัฒนธรรมของสังคมซึ่งแตกต่างกันเข้ามาเกี่ยวข้องด้วย เช่น ผู้เรียนไทยย่อมทำข้อสอบภาษาอังกฤษได้ไม่ดีเท่ากับผู้เรียนอังกฤษ หรือผู้เรียนอังกฤษย่อมทำข้อสอบภาษาไทยไม่ได้ ดังนั้นการเขียนข้อสอบประเภทนี้ต้องคำนึงถึงความแตกต่างของระเบียบวัฒนธรรมของสังคมที่แตกต่างกันของผู้เรียนด้วย

8.5.2 แบบทดสอบที่ไม่ใช้ภาษา หมายถึงแบบทดสอบที่ใช้รูปภาพ สัญลักษณ์หรือตัวเลขแทนภาษา สามารถนำไปใช้ทดสอบกับผู้เรียนทุกชาติทุกภาษาได้ โดยไม่มีความได้เปรียบเสียเปรียบจากวัฒนธรรมที่แตกต่างกัน จึงเรียกได้ว่ามีความยุติธรรมทางวัฒนธรรม เช่น

	คำถาม	ตัวเลือก
ข้อ 0	1, 3, 9, ..., 81	ก. 18 ข. 27 ค. 36 ง. 45

8.6แบ่งตามลักษณะของการใช้ประโยชน์แบ่งเป็น 2 ประเภท ดังนี้

8.6.1 แบบทดสอบย่อย หมายถึง แบบทดสอบประจำบทหรือหน่วยการเรียนลักษณะของข้อสอบจะสอดคล้องกับจุดมุ่งหมายเชิงพฤติกรรมหรือหน่วยการเรียนทั้งนี้เพื่อปรับปรุงการเรียนการสอน ช่วยให้ผู้สอนวิเคราะห์หาสาเหตุของความบกพร่องและซ่อมเสริมแก้ไขได้ตรงจุด แบบทดสอบลักษณะนี้มักจะใช้ทดสอบระหว่างภาคเรียน

8.6.2 แบบทดสอบรวม หมายถึงแบบทดสอบสรุปรวมเนื้อหาที่เรียนผ่านมาตลอดภาคเรียนมีจุดมุ่งหมายเพื่อตัดสินผลการเรียน

8.7 แบ่งตามเนื้อหาของข้อสอบในฉบับแบ่งเป็น 2 ประเภท ดังนี้

8.7.1 แบบทดสอบอัตนัย หมายถึงแบบทดสอบที่มีเฉพาะคำถามผู้เรียนต้องคิดหาคำตอบเองโดยการเขียนอย่างเสรี ลักษณะของคำตอบจะไม่คงที่แน่นอน ได้แก่ แบบทดสอบอัตนัยหรือความเรียง แบบตอบสั้น ๆ และแบบเติมคำ

8.7.2 แบบทดสอบปรนัย หมายถึงแบบทดสอบที่มีทั้งคำถามและคำตอบเฉพาะคงที่แน่นอน ได้แก่ แบบทดสอบแบบเลือกตอบ แบบจับคู่ แบบกาถูก-ผิด

สรุปท้ายบท

การวัดผลทางการศึกษา เป็นกระบวนการเพื่อให้ได้มาซึ่งตัวเลข หรือสัญลักษณ์ที่มีความหมายและแทนคุณลักษณะหรือคุณภาพของสิ่งที่วัดโดยใช้เครื่องมือที่มีประสิทธิภาพการประเมินผลการศึกษาเป็นกระบวนการลงข้อสรุป ให้คุณค่าของสิ่งต่างๆโดยอาศัยข้อมูลที่ได้จากการวัดไปเปรียบเทียบกับเกณฑ์ที่ได้กำหนดไว้และใช้วิจารณญาณประกอบการพิจารณาส่วนการทดสอบเป็นกระบวนการอย่างหนึ่ง ที่จะได้มาซึ่งจำนวนปริมาณหรือคุณลักษณะของพฤติกรรมของบุคคลโดยใช้ข้อสอบเป็นเครื่องมือการวัดผลและการประเมินผลเป็นกระบวนการที่ต่อเนื่องกันเพราะว่าการประเมินผลจะนำเอาข้อมูลที่ได้จากการวัดผลมาสรุปให้คุณค่าหรือตัดสินใจโดยนำไปเทียบกับเกณฑ์ที่

กำหนดไว้ในกระบวนการเรียนการสอนนั้นจะทำการประเมินผลการเรียนควบคู่กันไปเป็นระยะๆคือ การประเมินผลก่อนเรียน เพื่อตรวจสอบสภาพพื้นฐานของผู้เรียนการประเมินผลระหว่างเรียนเพื่อปรับปรุงการเรียนการสอน และการประเมินผลรวมเพื่อตัดสินผลการเรียนสำหรับจุดมุ่งหมายของการวัดผลทางการศึกษาเพื่อค้นและพัฒนาสมรรถภาพของผู้เรียนมากกว่าที่จะวัดผลเพื่อนำไปตัดสินได้หรือตกโดยหลักของการวัดผลทางการศึกษานั้นจะต้องวัดให้ตรงกับจุดประสงค์ ใช้เครื่องมือที่มีคุณภาพ มีความยุติธรรมแปลผลได้ถูกต้อง แล้วจึงนำผลที่ได้จากการวัดนั้นไปใช้อย่างคุ้มค่า นอกจากนี้เทคนิคที่ใช้ในการวัดผลทางการศึกษามีหลายชนิด มีลักษณะการใช้แตกต่างกันตามโอกาสหรือสถานการณ์ การที่มีเทคนิคการวัดผลทางการศึกษาหลายชนิด ก็เพื่อให้การวัดครอบคลุมด้านพุทธิพิสัย จิตพิสัย และทักษะพิสัย แม้ว่าปัจจุบันจะเน้นการวัดและประเมินตามสภาพจริง แต่ผู้สอนควรต้องมีความรู้ ความเข้าใจ และเลือกใช้เทคนิคการวัดผลทางการศึกษาอย่างเหมาะสม ไม่ว่าจะเป็นการสังเกต การสัมภาษณ์ แบบสอบถาม การจัดอันดับ การประเมินตามสภาพจริง การประเมินภาคปฏิบัติ การประเมินแฟ้มสะสมผลงาน และการทดสอบ เห็นได้ชัดว่าการวัดผลทางการศึกษามีประโยชน์อย่างยิ่งต่อทั้งผู้สอนผู้เรียน การแนะแนว การบริหาร การวิจัย และผู้ปกครองหรือผู้อื่นที่เกี่ยวข้อง ซึ่งเป็นส่วนสำคัญในการผลักดันคุณภาพการศึกษาของประเทศชาติให้เจริญก้าวหน้าและสู่ระดับนานาชาติได้อย่างมีประสิทธิภาพต่อไป

เอกสารอ้างอิง

ชูศรี วงศ์รัตนะ. (2541). เทคนิคการใช้สถิติเพื่อการวิจัย. (พิมพ์ครั้งที่ 8). กรุงเทพมหานคร: เทพเนรมิตการพิมพ์.

ทิวัตถ์ มณีโชติ. (2549). การวัดและประเมินผลการเรียนรู้ตามหลักสูตรการศึกษาขั้นพื้นฐาน. กรุงเทพฯ: สำนักพิมพ์ศูนย์ส่งเสริมวิชาการ.

พิสณุ ฟองศรี. (2551). การประเมินทางการศึกษา: แนวคิดสู่การปฏิบัติ. พิมพ์ครั้งที่ 5. กรุงเทพฯ: บริษัทด่านสุทธาการพิมพ์ จำกัด

สมชาย รัตนทองคำ. (2554). การวัดและประเมินผลทางการศึกษา. เอกสารประกอบการสอนวิชาการสอนทางกายภาพบำบัด, หน้า 137-153, สืบค้นเมื่อวันที่ 1 มีนาคม 2556 จาก http://ams.kku.ac.th/aalearn/resource/edoc/tech/54/13eva.pdf

สมนึก ภัททิยธนี. (2549). การวัดผลการศึกษา. พิมพ์ครั้งที่ 5 กาฬสินธุ์: โรงพิมพ์ประสานการพิมพ์.

สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, สำนักวิชาการและมาตรฐานการศึกษา. (2553). เอกสารประกอบแนวปฏิบัติการวัดและประเมินผลการเรียนรู้ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551. โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย จำกัด. พิมพ์ครั้งที่ 2, กรุงเทพฯ. หน้า 1-258.

สำนักมาตรฐานวิชาชีพ กลุ่มมาตรฐานและจรรยาบรรณวิชาชีพ. (2550). คู่มือการดำเนินงานรับรองปริญญาและประกาศนียบัตรทางการศึกษาเพื่อการประกอบวิชาชีพ. สำนักงานเลขาธิการคุรุสภา. พิมพ์ครั้งที่ 2 โรงพิมพ์ สกสค. กรุงเทพมหานคร. หน้า 44-50.

Gronlund, N.E. and Linn, R.L. (6th ed). (1990). *Measurement and Evaluation in Teaching* NY: Macmillan.

Harlen, W. (2007). *Assessment of Learning*. A Sage Publications Company.

Thomas, M. (2012).Teachers' Beliefs about Classroom Assessment and their selection of Classroom Assessment Strategies. *Journal of Research and Reflections in Education*, 6(2), 103 -112.

Tileston, D. W. (2004). *What Every Teacher Should Know About Student Assessment*. Corwin Press. A Sage Publications Company.

เอกสาร

อนุกรมวิธานของบูลม

อนุกรมวิธานของบลูมเป็นเครื่องมือที่ถูกนำมาใช้ทางการศึกษาอย่างขว้างขวางเพื่อสร้างและส่งเสริมความลุ่มลึกในการเรียนรู้ของผู้เรียน ดังนั้นสิ่งที่สร้างความท้าทายให้กับผู้สอนในการวัดและประเมินผู้เรียนในแต่ละระดับตามอนุกรมวิธานของบลูม คือ ผู้สอนควรดำเนินการอย่างไร

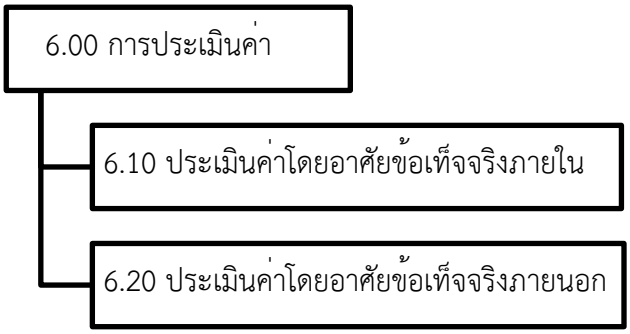
อย่างไรก็ตามได้มีการศึกษาว่าการใช้อนุกรมวิธานของบลูม สามารถยกระดับความคิดรวบยอด ทักษะ และการคิดเชิงวิเคราะห์ของผู้เรียน (Bloom, Engelhart, Furst, Hiss, & Krathwohl, 1956; Bissell and Lemons, 2006: 66) แต่ผู้สอนโดยส่วนใหญ่มักจะใช้การวัดและประเมินผู้เรียนโดยการสอบ ซึ่งลักษณะข้อสอบจะเป็นการวัดความจำเกี่ยวกับข้อเท็จจริง รวมถึงลักษณะการจัดการเรียนรู้ก็จะเป็นแบบการให้ผู้เรียนใช้ระดับสติปัญญาขั้นพื้นฐาน (ระดับจำ เข้าใจ) ซึ่งสะท้อนให้เห็นว่าเป็นสิ่งที่ตรงกันข้ามกับความต้องการที่จะยกระดับและส่งเสริมระดับสติปัญญาในระดับสูงให้กับผู้เรียน

อนุกรมวิธานของบลูมแบบเดิม

อนุกรมวิธานแบบของบลูม จำแนกการเรียนรู้เป็น 3 ด้าน โดยในแต่ละด้านจะมีการจำแนกระดับความสามารถจากต่ำสุดไปถึงสูงสุด ดังนี้

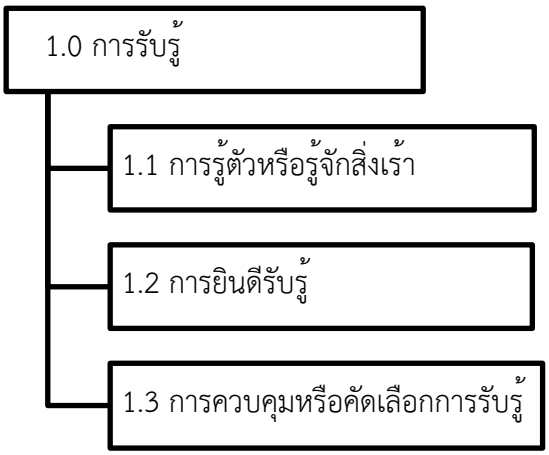
1. ด้านพุทธิพิสัย หรือด้านปัญญา เป็นพฤติกรรมที่เกี่ยวข้องกับสมรรถภาพทางสมองหรือสติปัญญาของบุคคล โดยพฤติกรรมทางสมองของบุคคลจะมีลำดับของพัฒนาการจาก พฤติกรรมที่ใช้ความสามารถง่าย ๆ ไปหาพฤติกรรมที่ใช้ความสามารถที่ยุ่งยากซับซ้อน หรือกล่าวได้ว่าด้านพุทธิพิสัยหรือด้านปัญญาจะเกี่ยวข้องกับความรู้หรือการคิด แบ่งเป็น 6 ระดับ เริ่มจากความรู้ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินค่า แต่ละพฤติกรรมต้องอาศัยซึ่งกันและกัน เช่น ผู้เรียนจะเข้าใจเรื่องใดต้องอาศัยความรู้ความจำในเรื่องนั้นเป็นพื้นฐานก่อน กล่าวโดยสรุปก็คือพฤติกรรมสูง ๆ จะต้องอาศัยพฤติกรรมต่ำ ๆ เป็นพื้นฐาน ดังแสดงในภาพที่ 2.1

1.00 ความรู้ความจำ
1.10 ความรู้ในเนื้อเรื่อง
1.20 ความรู้ในวิธีดำเนินการ
1.30 ความรู้รวบยอดในเนื้อเรื่อง
2.00 ความเข้าใจ
2.10 การแปลความ
2.20 การตีความ
2.30 การขยายความ
3.00 การนำไปใช้
4.00 การวิเคราะห์
4.10 การวิเคราะห์ความสำคัญ
4.20 การวิเคราะห์ความสัมพันธ์
4.30 การวิเคราะห์หลักการ
5.00 การสังเคราะห์
5.10 สังเคราะห์ข้อความ
5.20 สังเคราะห์แผนงาน
5.30 สังเคราะห์ความสัมพันธ์



ภาพที่ 2.1 ลำดับขั้นของพฤติกรรมด้านพุทธิพิสัย

2. ด้านจิตพิสัย หรือด้านความรู้สึก/อารมณ์ เป็นเรื่องเกี่ยวกับความรู้สึกนึกคิด หรือการจัดระเบียบของจิตใจ เช่น เจตคติ ความสนใจ ความซาบซึ้ง คุณธรรมและการปรับตัว ความรู้สึกหรือลักษณะนิสัยต่าง ๆ ของบุคคล จะเป็นในรูปแบบใดนั้นขึ้นอยู่กับการปลูกฝังอบรม แครธโวห์ลและคณะ (Krathwohl, Bloom & Masia ,1964) ได้เสนอจุดมุ่งหมายการจัดการเรียนรู้ด้านจิตพิสัยเป็น 5 ขั้นจากต่ำสุดคือขั้นรับรู้ จนถึงขั้นสูงสุดคือขั้นการสร้างลักษณะนิสัยของบุคคล ดังภาพ 2.2 (Bohlin, 1998 ; Kennedy, Hyland, Ryan, 2006 ; Kretchmar, 2008; Bloom’s Taxonomy of Learning Domains Quoted in Savickiene, 2010: 42-44)



ภาพที่ 2.2 ลำดับขั้นของพฤติกรรมด้านจิตพิสัย

แนวคิดการจัดการเรียนรู้ด้านจิตพิสัยมีมานานกว่า 100 ปี และมีความสำคัญไม่น้อยไปกว่าการวัดทางด้านพุทธิพิสัยและทักษะพิสัย โดยการจัดการเรียนรู้ด้านจิตพิสัยตามแนวคิดของ แครธโวห์ลและ คณะ (1964) จะเกี่ยวข้องกับองค์ประกอบด้านเจตคติ ความเชื่อ ค่านิยม อารมณ์ และการยอมรับหรือปฏิเสธ ในขณะการจัดการเรียนรู้ด้านจิตพิสัยตามแนวคิดของนักการศึกษาอื่น เช่น Bohlin (1998) จะเกี่ยวข้องกับความกังวล การกระตุ้น เจตคติ คุณลักษณะ ความเชื่อและความเห็น ความมั่นใจ ความคาดหวังในความสำเร็จ ความสนใจ ระดับการจูงใจ การศึกษาทางด้านจิตพิสัย ซึ่งในการศึกษาของนักการศึกษามีส่วนทั้งที่แตกต่างกันและเหมือนกัน ดังตารางที่ 2.1

ตารางที่ 2.1 แสดงองค์ประกอบด้านจิตพิสัยที่มีการศึกษาโดยนักการศึกษา

องค์ประกอบด้านจิตพิสัย	นักการศึกษาที่เกี่ยวข้อง
เจตคติ	Krathwohl et al. (1964) ; Gephart and Ingle (1976); Anderson (1981) ; Miller (2005); Koballa (2007); Kretchmar (2008) ; Shephard (2008)
ค่านิยม	Krathwohl et al. (1964) ; Gephart and Ingle (1976); Anderson (1981) ; Miller (2005); Koballa (2007); Kretchmar (2008) ; Shephard (2008)
การจูงใจ	Main (1992); Miller (2005); Koballa (2007); Kretchmar (2008)
ความเชื่อ	Krathwohl et al. (1964); Hunt (1987); Pierre and Oughton (2007)
อารมณ์	Krathwohl et al. (1964) ; Gephart and Ingle (1976); Hunt (1987)
การยอมรับ หรือการปฏิเสธ	Krathwohl et al. (1964)
การรับรู้	Gephart and Ingle (1976)
ความชอบ	Anderson (1981)
ความสนใจ	Anderson (1981)

การยอมรับตนเองทางวิชาการ	Anderson (1981)
ความกังวล	Anderson (1981)
การควบคุม	Anderson (1981)
พฤติกรรม	Shephard (2008)

เครื่องมือด้านจิตพิสัย

1. แบบทดสอบโดยใช้สถานการณ์ เป็นการกำหนดสถานการณ์ที่เหมาะสมเพื่อให้ผู้ตอบแสดงความคิดหรือเลือกตอบโดยตัวเลือกเป็น ข้อความที่แสดงความรู้สึกหรือความคิดเห็นที่สะท้อนถึงพฤติกรรม หรือเป็นตัวแทนในแต่ละระดับพฤติกรรมด้านจิตพิสัย

2. แบบวัดทัศนคติ อาจทำได้ตามรูปแบบดังนี้

2.1 มาตราวัดทัศนคติของลิเคิร์ทเป็นข้อความที่เป็นความรู้สึกต่อสิ่งที่วัด ซึ่งมีทั้งในทางบวกและทางลบ โดยมีตัวเลือกแบ่งเป็น 5 ระดับ คือ เห็นด้วยอย่างยิ่ง เห็นด้วย ไม่แน่ใจ ไม่เห็นด้วย และไม่เห็นด้วยอย่างยิ่ง

2.2 แบบวัดทัศนคติโดยใช้ความหมายทางภาษา เป็นแบบวัดทัศนคติโดยใช้ความหมายทางภาษาซึ่งใช้คำคุณศัพท์ที่ตรงกันข้ามสอดคล้องกับทัศนคติที่ต้องการศึกษาโดย กำหนดมาตราให้กับข้อที่นำมาถาม 5 ช่อง หรือ 7 ช่อง แล้วแต่ระดับของผู้ตอบ

3. แบบบันทึกการสังเกต เป็นการกำหนดพฤติกรรมในสถานการณ์ต่าง ๆ เพื่อใช้เป็นตัวแทนของคุณลักษณะที่ต้องการวัด

การสร้างเครื่องมือวัดผลด้านจิตพิสัย

1. กำหนดคุณลักษณะ หรือ เป้าหมายที่จะวัด หรือสิ่งที่จะวัด โดยระบุให้ชัดเจนว่าต้องการวัดอะไร เช่น เจตคติต่อการเรียนวิชาคณิตศาสตร์ เป็นต้น

2. นิยามความหมายของสิ่งที่จะวัดให้ชัดเจน โดยการวิเคราะห์คุณลักษณะของสิ่งที่จะวัดว่าแต่ละคุณลักษณะเหล่านั้นเป็นอย่างไร เพื่อให้ทราบว่าสิ่งที่จะวัดคืออะไร ประกอบด้วยคุณลักษณะใดบ้าง

3. สร้างข้อความ โดยเขียนให้ครอบคลุมคุณลักษณะทุกด้าน และในคุณลักษณะแต่ละด้านควรมี หลาย ๆ ข้อด้วย การสร้างข้อความเพื่อเป็นเครื่องมือในการวัด

4. ทดลองเครื่องมือ นาเครื่องมือที่สร้าง ไปทดสอบกับนักเรียนว่ามีความรู้สึกหรือ ความคิดเห็นอย่างไร เช่น เห็นด้วยอย่างยิ่ง - ไม่เห็นด้วยอย่างยิ่ง หรือ ยอมรับ- ไม่ยอมรับ แล้วนาผลที่ได้มาทาการวิเคราะห์ตามเทคนิควิธีของแต่ละชนิดของเครื่องมือวัด

5. ปรับปรุงคุณภาพของเครื่องมือ ตามผลการวิเคราะห์ แล้วนาไปทดลองใช้ แล้วนามาวิเคราะห์จนแน่ใจในคุณภาพ

6. สร้างเกณฑ์ในการให้คะแนน พร้อมทั้งเขียนคู่มือการใช้เครื่องมือดังกล่าว

3. ด้านทักษะพิสัย เป็นพฤติกรรมที่เกี่ยวกับทักษะการเคลื่อนไหวและการใช้อวัยวะส่วนต่าง ๆ ของร่างกาย ทักษะการเคลื่อนไหวอวัยวะสองส่วนหรือมากกว่าพร้อม ๆ กัน ถ้าบุคคลใดสามารถบังคับระบบ กล้ามเนื้อ ระบบประสาท และสมองให้สัมพันธ์กันได้แล้ว ย่อมจะเกิดทักษะในการปฏิบัติสิ่งต่าง ๆ ซึ่งต้องอาศัยการฝึกฝนจากการปฏิบัติจริงบ่อย ๆ จนเกิดความคล่องแคล่วว่องไวและถูกต้อง การวัดพฤติกรรมในด้านนี้ อาจทำได้โดยการตรวจผลงานภาคปฏิบัติ พฤติกรรมด้านทักษะพิสัยจะมีการพัฒนาการตามลำดับ ก่อน-หลัง ถ้ายึดตามแนวคิดของซิมป์สัน พบว่ามีการจำแนกพฤติกรรมด้านทักษะพิสัย ออกเป็น 7 ระดับ ดังนี้

1. การรับรู้ จำแนกเป็น 3 ขั้น คือ การกระทบของสิ่งเร้าต่อประสาทสัมผัสด้านต่าง ๆ การเลือกที่จะรับรู้ และการแปลความหมาย

2. การตระเตรียม จำแนกเป็น 3 ด้าน คือ การตระเตรียมทางสมอง การตระเตรียมทางร่างกาย และการตระเตรียมทางอารมณ์

3. การตอบสนองตามการชี้แนะ จำแนกเป็น 2 ขั้น คือ การเลียนแบบ และการลองผิดลองถูก

4. การสร้างกลไก

5. การตอบสนองที่ซับซ้อนขึ้น จำแนกเป็น 2 ขั้น คือ การทำอย่างไม่ลังเล และการปฏิบัติได้อย่างอัตโนมัติ

6. การดัดแปลงให้เหมาะสม

7. การริเริ่มใหม่

ถ้ายึดหลักของอาร์โร (Harrow) พบว่ามีการจำแนกพฤติกรรมด้านทักษะพิสัย ออกเป็น 6 ขั้น ดังนี้

1. การเคลื่อนไหวตามธรรมชาติ

2. การเลียนแบบ

3. ลงมือทำตามแบบ

4. มีความถูกต้องเที่ยงตรง

5. การกระทำที่ต่อเนื่องและประสานกัน

6. การกระทำเองจนเกิดความเคยชินและเป็นไปตามธรรมชาติ

จากรายละเอียดข้างต้นจะพบว่าการวัดสถานภาพและพัฒนาการทางด้านพุทธิพิสัยและด้านทักษะพิสัย มีความเป็นไปได้ที่จะได้รับการพัฒนาให้ก้าวหน้ามาก สามารถทำได้ชัดเจน เชื่อถือได้มาก ให้คะแนนได้อย่างมั่นใจ แต่การวัดด้านจิตพิสัย ทั้ง ๆ ที่เป็นด้านที่สำคัญมาก เป็นคุณลักษณะที่จำเป็นของการเป็นมนุษย์อย่างสมบูรณ์ แต่กลับได้รับการพัฒนาในระดับที่น่าพอใจได้น้อยกว่า วัดได้ยากกว่า และให้คะแนนอย่างมั่นใจยากกว่า ซึ่งสภาพของการวัดด้านจิตพิสัยส่วนใหญ่เป็นดังนี้ (บุญชม ศรีสะอาด, ม.ป.ป.)

1. ยากต่อการให้นิยามที่ชัดเจน ผู้รู้หลายคนให้นิยามไม่ตรงกัน

2. ขาดทฤษฎีที่เป็นกรอบในการสร้างเครื่องมือวัดผลที่วัดออกมาจึงไม่มั่นใจเท่าที่ควร

3. คุณลักษณะบางอย่าง บางส่วน ผันแปรไปตามเวลา สถานที่ อารมณ์ จึงวัดธรรมชาติแท้จริงได้ยาก

4. พฤติกรรมอย่างเดียวกันของหลายคน อาจมาจากต้นเหตุของจิตที่แตกต่างกัน

5. คนที่มีคุณลักษณะทางจิตพิสัยเหมือนกัน อาจมีพฤติกรรมการแสดงออกที่แตกต่างกัน และพฤติกรรมที่แตกต่างกันอาจมาจากคุณลักษณะทางจิตพิสัยเดียวกัน

6. ผู้ตอบบางคนตอบไม่ตรงกับความเป็นจริง เพราะถ้าตอบตรงตามความเป็นจริงจะทำให้ตน เสียผลประโยชน์ได้คะแนนน้อย ภาพพจน์ในสายตาคนอื่นอยู่ในระดับต่ำ ไม่สอดคล้องกับความมุ่งหวังทางสังคม

อนุกรมวิธานของบลูมแบบปรับใหม่

แอนเดอร์สัน และคณะ (Anderson et al., 2001: 38-62) ได้อธิบายอนุกรมวิธานของบลูมแบบปรับใหม่ โดยเสนอในลักษณะของมิติด้านความรู้ (Knowledge dimension) และมิติด้านกระบวนการทางปัญญา (Cognitive process dimension) โดยมีรายละเอียดตามลำดับดังต่อไปนี้

ด้านมิติด้านความรู้ แบ่งออกเป็น 4 ประเภท ดังนี้

1. ความรู้ด้านข้อเท็จจริง (Factual Knowledge) คือความรู้พื้นฐานเกี่ยวกับองค์ประกอบต่าง ๆ ที่จะต้องทราบในวิชา ดังนี้

1.1 ความรู้เกี่ยวกับศัพท์เฉพาะ (Terminology) เช่น คำศัพท์เทคนิค (Technical Vocabulary) สัญลักษณ์ทางด้านดนตรี (Musical symbols)

1.2 ความรู้เกี่ยวกับรายละเอียดจำเพาะและองค์ประกอบส่วนต่าง ๆ (Knowledge of Specific Details and Elements) เช่น องค์ประกอบแต่ละส่วน หรือแหล่งความรู้แต่ละแหล่งมีลักษณะอย่างไร แหล่งสารสนเทศที่เชื่อถือได้ แหล่งความรู้อยู่ที่ใด ประเด็นความรู้นี้ต้องไปสอบถามจากใคร ใครเชื่อถือได้

2. ความรู้ด้านมโนทัศน์/ความคิดรวบยอด (Conceptual Knowledge) คือ ความรู้ที่เกี่ยวกับความสัมพันธ์ระหว่างองค์ประกอบพื้นฐานซึ่งเป็นส่วนหนึ่งขององค์กรหรือของเครื่องมือภายในโครงสร้างใหญ่ที่ทำให้เกิดการหน้าที่ร่วมกัน เช่น ความรู้เกี่ยวกับการจัดประเภท จัดหมวดหมู่ เช่น การจัดประเภทตามช่วงเวลาตามกิจกรรม ความรู้เกี่ยวกับหลักการและการสรุปอ้างอิง เช่น ความรู้เกี่ยวกับทฤษฎีทางเรขาคณิต ความรู้เกี่ยวกับทฤษฎี แบบจำลอง และโครงสร้างขององค์กร เช่น ทฤษฎีวิวัฒนาการ ส่วนประกอบของเครื่องคอมพิวเตอร์แต่ละส่วนทำหน้าที่อะไร

3. ความรู้ด้านกระบวนการ (Procedural Knowledge) คือความรู้เกี่ยวกับวิธีดำเนินการปฏิบัติหรือทำงานแต่ละเรื่องว่าจะทำอย่างไร วิธีการแก้ไขปัญหา และระเบียบวิธีการแสวงหาความรู้ มีดังนี้

ความรู้เรื่องทักษะเฉพาะและสูตรในการแก้ปัญหา เช่น รู้เกี่ยวกับโปรแกรมคอมพิวเตอร์ ว่าโปรแกรมใดใช้ทำอะไรได้

ความรู้เทคนิคและระเบียบวิธีเฉพาะในเนื้อหาวิชาใดวิชาหนึ่ง เช่น วิธีการทางวิทยาศาสตร์ เทคนิคการสัมภาษณ์ เทคนิคการสืบสวน สอบสวน

ความรู้เรื่องเกณฑ์สำหรับกำหนดว่าเมื่อใดจะใช้วิธีการดำเนินการที่เหมาะสมกับสถานการณ์ เช่น เมื่อใดจะใช้วิธีการสำรวจเพื่อตรวจสอบคะแนนนิยมของประชาชน

4. ความรู้ด้านอภิปัญญา (Metacognitive Knowledge) เป็นความรู้ด้านพุทธิปัญญาโดยทั่วไปของบุคคล และการตระหนักรู้เกี่ยวกับระดับภูมิปัญญาของตนเองว่า ตนเองรู้อะไร ไม่รู้อะไร ซึ่งความรู้ด้านอภิปัญญาของตนเอง (Knowledge about Cognition) และการนำความรู้เกี่ยวกับความรู้ทั้งหมดมาใช้ ดังนี้

4.1 ความรู้เชิงกลยุทธ์ เป็นความรู้เกี่ยวกับการดึงส่วนสำคัญมาเป็นวิถีปฏิบัติการวางแผนกลยุทธ์ โดยเขียนแผนเชิงสังเขปเพื่อจับใจความสำคัญของแผนงานเพื่อนำไปปฏิบัติ

4.2 ความรู้เกี่ยวกับงานด้านพุทธิปัญญา ความรู้ที่เป็นเงื่อนไขและบริบท เป็นความรู้เกี่ยวกับกิจกรรมด้านพุทธิปัญญารวมทั้งความรู้เกี่ยวกับบริบทและเงื่อนไขที่เหมาะสม รู้ว่างานใดต้องใช้ความรู้ทางพุทธิปัญญาด้านใด

4.3 ความรู้ของตนเอง การตระหนักรู้ในความรู้ของตนเอง รู้ว่าความสามารถของตนเองจะแก้ปัญหา จะทำการสิ่งใดได้ มีจุดเด่น จุดด้อยด้านใด รู้ว่าเป้าหมายในการทำงานหนึ่ง ๆ คืออะไร รู้ว่าตนเองสนใจที่จะทำอะไรและจะทำได้หรือไม่ มากน้อยเพียงใด

นอกจากนี้ Anderson และ คณะ (2001: 29) ได้สรุปมิติด้านความรู้ พร้อมยกตัวอย่างสิ่งที่เกี่ยวข้องในแต่ละด้านย่อย โดยมีรายละเอียดดังตารางที่ 2.2 ต่อไปนี้

ตารางที่ 2.2 มิติด้านความรู้ตามแนวของบลูมแบบปรับใหม่

มิติด้านความรู้	ตัวอย่าง
1. ความรู้เกี่ยวกับข้อเท็จจริง (Factual knowledge)	องค์ประกอบพื้นฐานที่ผู้เรียนต้องรู้ในการเรียนของรายวิชา
1.1 ความรู้เกี่ยวกับคำศัพท์	ศัพท์เฉพาะ สัญลักษณ์
1.2 ความรู้เกี่ยวกับรายละเอียดที่เฉพาะเจาะจงและส่วนประกอบ	แหล่งข้อมูล ความน่าเชื่อถือของแหล่งข้อมูล
2. ความรู้ที่เป็นความคิดรวบยอด /เชิงมโนทัศน์(Conceptual knowledge)	ความสัมพันธ์ระหว่างองค์ประกอบพื้นฐานที่มีโครงสร้างใหญ่ขึ้นและสามารถเชื่อมโยงแต่ละส่วนได้

2.1 ความรู้ในการจัดประเภท และจำแนก	ช่วงเวลา..... รูปแบบของธุรกิจส่วนตัว
2.2 ความรู้ในหลักการและการอ้างอิง	ทฤษฎีพิธาโกรัส กฎอุปสงค์และอุปทาน
2.3 ความรู้ในทฤษฎี รูปแบบและโครงสร้าง	ทฤษฎีวิวัฒนาการ โครงสร้างของสภาผู้แทน
3. ความรู้ที่เป็นกระบวนการ (Procedural knowledge)	วิธีการทำบางสิ่ง วิธีการสืบค้น และเกณฑ์สำหรับใช้ทักษะ ขั้นตอนการดำเนินการ เทคนิคและวิธีการ
3.1 ความรู้เกี่ยวกับทักษะและขั้นตอน	ทักษะในการวาดรูปเมื่อใช้สีน้ำ ขั้นตอนการหารจำนวนเต็มด้วยทศนิยม
3.2 ความรู้เกี่ยวกับเทคนิคและวิธีการ	เทคนิคการสัมภาษณ์ วิธีการทางวิทยาศาสตร์
3.3 ความรู้ในการกำหนดเกณฑ์เพื่อนำมาใช้ในกระบวนการ	เกณฑ์ที่ใช้กำหนดเมื่อใช้กระบวนการตามกฎนิวตัน เกณฑ์การใช้เพื่อตัดสินขอบเขตการประเมินการลงทุนด้านธุรกิจ
4. ความรู้ที่เป็นอภิปัญญา (Metacognitive knowledge)	ความรู้ในความรู้ของตนเอง การตระหนักรู้และความรู้ที่ตนเองรู้
4.1 ความรู้เชิงกลยุทธ์	การจัดวางโครงสร้างของเนื้อหาที่เรียนเพื่อศึกษาค้นคว้าด้วยตนเอง
4.2 ความรู้เกี่ยวกับการทำงาน บริบทที่เหมาะสมและความรู้ที่เป็นเงื่อนไข	ความรู้ในความต้องการของงานที่แตกต่างกัน
4.3 ความรู้ของตนเอง	ความรู้ในการวิพากษ์ วิจารณ์จุดแข็ง จุดอ่อนในผลงานตนเอง การรู้ในความรู้ของตนเอง

และเพื่อให้ผู้เรียนเกิดความชัดเจนเกี่ยวกับมิติด้านความรู้ ขอให้ผู้เรียนศึกษาเพิ่มเติมว่าหัวข้อของความรู้ที่กำหนดให้สัมพันธ์กับการวัดมิติด้านความรู้ในระดับต่อไปนี้อย่างไร ดังตารางที่ 2.3

ตารางที่ 2.3 หัวข้อของความรู้และมิติด้านความรู้ตามแนวคิดของบลูมแบบปรับใหม่

หัวข้อของความรู้	มิติด้านความรู้			
	Factual	Conceptual	Procedural	Metacognitive
1. ความรู้เกี่ยวกับทฤษฎีโครงสร้างทางสังคม		✓		
2.ความรู้เกี่ยวกับสัญลักษณ์ของการใช้แผนที่ แผนภูมิ	✓			
3.ความรู้เกี่ยวกับการจำแนกส่วนประกอบของประโยค		✓		
4.ความรู้เกี่ยวกับวิธีการประเมินสุขภาพ			✓	
5.ความรู้เกี่ยวกับการกำหนดเป้าหมายในการอ่าน				✓
6.ความรู้เกี่ยวกับเกณฑ์การใช้สูตรเพื่อปฏิบัติขั้นตอนได้อย่างเหมาะสม			✓	
7.ความรู้เกี่ยวกับแหล่งข้อมูลที่น่าเชื่อถือในการซื้อสินค้า	✓			
8.ความรู้เกี่ยวกับการตัดสินว่างานที่ปฏิบัติมีคุณค่าและความหมายมากน้อยเพียงใด				✓
9.ความรู้เกี่ยวกับการจัดระบบขององค์กรปกครองส่วนท้องถิ่น		✓		

10.ความรู้เกี่ยวกับทักษะการกระโดดสูง			✓	
11.ความรู้เกี่ยวกับการวิเคราะห์วิธีการในการแก้ปัญหาที่ไม่ชัดเจน				✓
12.ความรู้เกี่ยวกับสินค้าที่ส่งออกของประเทศ	✓			
13.ความรู้เกี่ยวกับหลักการในการเรียนรู้		✓		
14.ความรู้เกี่ยวกับขั้นตอนการแก้สมการกำลังสอง			✓	

ด้านมิติกระบวนการทางปัญญา (Cognitive Process Dimension) แอนเดอร์สัน และคณะได้มีการดำเนินการปรับปรุงอนุกรมวิธานโดยนำแนวคิดของกลุ่มพุทธินิยมและสรรคนิยม เข้ามาเป็นพื้นฐานในการจำแนกกระบวนการทางพุทธิปัญญาออกเป็น 6 ระดับดังนี้ (Anderson et al., 2001: 63-91; จริยา เสถบุตร และ ธีระยุทธ นนทะสร, ม.ป.ป.)

1. จำ (Remember) สามารถดึงเอาความรู้ที่เกี่ยวข้องออกมาได้จากความจำระยะยาว ประกอบด้วย

1.1 จดจำได้ (Recognizing) จดจำวันที่และเหตุการณ์สำคัญสัญลักษณ์และสูตรต่าง ๆ ได้

1.2 ระลึกได้ (Recalling) ระลึกได้เมื่อต้องทำ เช่น เขียนสูตรได้

2. เข้าใจ (Understand) สามารถสร้างความหมายจากสิ่งที่ได้เรียนรู้ไป รวมทั้งในแง่การพูด การเขียน และการสื่อสารด้วยภาพ ประกอบด้วย

2.1 การอธิบายยกตัวอย่างได้ (Exemplifying)

2.2 การจำแนกประเภทได้ (Classifying)

2.3 การสรุปความได้ (Summarizing)

2.4 การสรุปอ้างอิงได้ (Inferring)

2.5 การสามารถเปรียบเทียบได้ (Comparing)

2.6 การอธิบายได้ (Explaining)

3. นำไปใช้/ประยุกต์ใช้ (Apply) การกระทำ การปฏิบัติ หรือการดำเนินงานในสถานการณ์ที่กำหนดให้ ประกอบด้วย

3.1 วิธีดำเนินการ (Executing) ตามวิธีดำเนินงานไปทีละคำสั่งที่คุ้นเคยหรือตามการหน้าที่

3.2 นำไปใช้ (Implementing) การดำเนินงานให้เกิดผลในสถานการณ์ที่แปลกใหม่

4. วิเคราะห์ (Analyze) สามารถแจกแจง แยกส่วนองค์ประกอบขององค์กร หรือวัตถุออกเป็นส่วนย่อย และตรวจสอบได้ว่าแต่ละส่วนเกี่ยวข้องกันอย่างไร และแต่ละส่วนนั้นเกี่ยวข้องกับโครงสร้างใหญ่อย่างไร ประกอบด้วย

4.1 การจำแนกแยกแยะ การแยกจำนวน แยกย่อยได้ (Differentiating)

4.2 การจัดระบบได้ การจัดองค์กรได้ (Organizing)

4.3 การให้ความเห็น ให้เหตุผลได้ (Attributing)

5. ประเมินค่า (Evaluate) การตัดสินคุณค่าโดยอาศัยเกณฑ์และมาตรฐานประกอบด้วย

5.1 การตรวจสอบ (Checking)

5.2 การวิพากษ์ การใช้วิจารณญาณ (Critiquing)

6. สร้างสรรค์ (Create) การนำหน่วยต่าง ๆ หรือองค์ประกอบย่อยเข้ามาบูรณาการเพื่อให้เป็นภาพรวมที่เชื่อมต่อกันอย่างมีตรรกะ ซึ่งประกอบด้วย

6.1 การสร้าง ก่อกำเนิด (Generating) เช่น การสร้างสมมติฐานการวิจัยจากการใช้เหตุผลเชิงอุปนัย (Inductive Reasoning)

6.2 การวางแผน (Planning) การวางแผนงานวิจัยจากหัวข้อที่กำหนดให้ การพัฒนาเค้าโครงการวิจัย

6.3 การผลิต การประดิษฐ์คิดค้นสิ่งใหม่ ๆ (Producing)

นอกจากนี้แอนเดอร์สันและคณะ (Anderson et al., 2001: 31) ได้สรุปมิติด้านกระบวนการพร้อมยกตัวอย่างสิ่งที่เกี่ยวข้องในแต่ละด้านย่อย โดยมีรายละเอียดดังตาราง 2.4 ต่อไปนี้

ตารางที่ 2.4 มิติด้านกระบวนการ กระบวนการทางปัญญาและตัวอย่าง

มิติด้านกระบวนการ	กระบวนการทางปัญญาและตัวอย่าง
1. จำ	ความรู้ที่มาจากการจำและระลึกได้ สามารถดึงเอาความรู้ที่เกี่ยวข้องออกมาได้จากความจำระยะยาว

1.1 จดจำได้	การจดจำเหตุการณ์ในประวัติศาสตร์ได้
1.2 ระลึกได้	การระลึกเหตุการณ์ที่สำคัญในประวัติศาสตร์ได้
2. เข้าใจ	การสร้างความหมายจากการเรียนรู้ การพูดปากเปล่า การเขียนและการสื่อความหมายจากภาพ
2.1 ตีความ	การใช้สำนวนใหม่ ถอดความจากคำพูดและเอกสาร
2.2 ให้ตัวอย่าง	ยกตัวอย่างอย่างหลากหลายในเรื่อง.....
2.3 จำแนก	จำแนกและพรรณนาจากการสังเกต
2.4 สรุป	เขียนสรุปอย่างย่อจากเหตุการณ์ การดูวิดีโอ
2.5 อ้างอิง	การอ้างถึงหลักการ เช่น หลักทางไวยากรณ์ภาษา
2.6 เปรียบเทียบ	เปรียบเทียบเหตุการณ์ในอดีตกับเหตุการณ์ร่วมสมัยปัจจุบัน
2.7 อธิบาย	อธิบายสาเหตุของการเกิดเหตุการณ์.....
3. นำไปใช้ ประยุกต์ใช้	ดำเนินการหรือใช้กระบวนการกับสถานการณ์ที่กำหนด
3.1 วิธีดำเนินการ	หารจำนวนเต็มที่มีมากกว่า 1 หลักด้วยจำนวนเต็มที่มากกว่า 1 หลัก
3.2 นำไปใช้	การใช้กฎนิวตันไปใช้กับสถานการณ์
4. วิเคราะห์	แยกแยะส่วนประกอบต่าง ๆ ออกเป็นส่วนๆ โดยแต่ละส่วนมีความสัมพันธ์เกี่ยวข้องกันอย่างไรกับส่วนย่อยที่แยกออกและภาพรวมทั้งหมด
4.1 ความแตกต่าง	ความแตกต่างระหว่างจำนวนตรรกยะและจำนวนอตรรกยะ
4.2 จัดระบบ	โครงสร้างและหลักฐานต่าง ๆ ในเรื่อง/เหตุการณ์
4.3 คุณลักษณะคุณสมบัติ	มุมมองของผู้แต่งเรียงความ
5. ประเมินค่า	การตัดสินใจโดยอาศัยเกณฑ์และมาตรฐาน
5.1 ตรวจสอบ	พิจารณาว่าผลสรุปของนักวิทยาศาสตร์กับการสังเกตจากการทดลองของตนเอง
5.2 วิจารณ์	ตัดสินว่าวิธีการใดในสองวิธีนั้นดีที่สุดในการแก้ปัญหา
6. สร้างสรรค์	นำองค์ประกอบแต่ละส่วนรวมเข้าด้วยกัน จัดระบบองค์ประกอบใหม่

6.1 สร้าง ก่อให้เกิด กำเนิด	ตั้งสมมติฐานเพื่ออธิบายปรากฏการณ์
6.2 วางแผน	วางแผนการเขียนบทความจากหัวข้อ.....
6.3 ผลผลิต	การสร้างสิ่งแวดล้อมให้กับพืชและสัตว์อย่างมีจุดมุ่งหมาย

ในการสร้างข้อสอบวัดความรู้นักเรียน ผู้สอนควรมีความเข้าใจการสร้างข้อสอบตามแนวคิดของบลูมที่ปรับใหม่ ขอให้ผู้เรียนพิจารณาขั้นตอนแรกของการสร้างข้อสอบดังภาพที่ 2.3 ซึ่งเริ่มจากจุดประสงค์การเรียนรู้ของหน่วยย่อย สมมติว่าจุดประสงค์การเรียนรู้ประจำหน่วยย่อย มีดังนี้

จุดประสงค์การเรียนรู้ : นักเรียนสามารถนำวิธีการเกี่ยวกับการรีไซเคิลไปประยุกต์ใช้ในชีวิตประจำวันได้

วิธีการรีไซเคิล
ระดับความรู้เป็นกระบวนการ

มิติกระบวนการทางปัญญา
คือการนำไปใช้

มิติด้านความรู้	มิติด้านกระบวนการทางปัญญา/ระดับการคิด					
	จำ	เข้าใจ	**นำไปใช้**	วิเคราะห์	ประเมินค่า	สร้างสรรค์
ข้อเท็จจริง						
มโนทัศน์						
กระบวนการ						
อภิปัญญา						

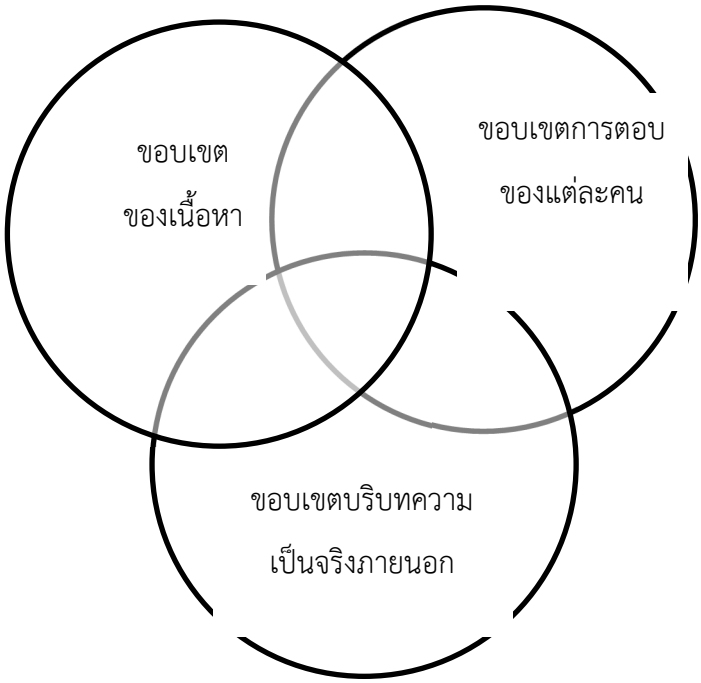
ภาพที่ 2.3 การจำแนกจุดประสงค์การเรียนรู้ตามมิติด้านความรู้และด้านกระบวนการทางปัญญา

คำถามที่อยู่ภายในใจของผู้เรียนหลายคน ได้แก่ ข้อสอบที่วัดระดับความรู้ “กระบวนการ” และวัดสติปัญญาหรือสมอง “นำไปใช้” แล้วลักษณะข้อสอบจะเป็นอย่างไร สิ่งแรกที่คุณครูควรพิจารณาได้แก่ การแสดงออกอย่างไรที่เรียกว่าการนำกระบวนการ...ไปประยุกต์ใช้ เช่น นักเรียนบอกได้ว่ารีไซเคิลสิ่งของคืออะไร ใช่หรือไม่ นักเรียนอธิบายพร้อมยกตัวอย่างการรีไซเคิลสิ่งของ ใช่หรือไม่

ลักษณะของคำถาม

การใช้คำถามในการจัดการเรียนรู้มีเรื่องที่ใช้กันมายาวนานและเป็นประเด็นหลักในการจัดการศึกษา บ่อยครั้งผู้สอนใช้คำถามเพื่อให้ผู้เรียนระลึก ทบทวนในความรู้ที่มีก่อนหน้านี้ ใช้คำถามเพื่อส่งเสริมความเข้าใจและสร้างทักษะการคิดเชิงวิเคราะห์และคิดอย่างมีวิจารณญาณ คำถามที่วัดระดับความจำ ความเข้าใจ และการนำไปใช้ จากอนุกรมวิธานของบลูมที่กล่าวมาข้างต้นบ่อยครั้งที่ถูกพิจารณา ว่าเป็นคำถามขั้นต่ำ ในขณะคำถามที่วัดระดับการวิเคราะห์ การสังเคราะห์ และการประเมินค่า ได้ถูกพิจารณาว่าเป็นคำถามขั้นสูง (Tofade et al., 2013: 1)

โทเฟด และคณะ (Tofade et al., 2013: 5-6) ได้เสนอคำถามในลักษณะวัฏจักร (Question circles) ซึ่งเป็นกลวิธีที่นำผู้เรียนจากการตอบที่ไม่ลึกซึ้งหรือผิวเผินมาสู่การอภิปรายอย่างลึกซึ้งในเนื้อหาสาระที่เรียนรู้ โดยแนวคิดหลักของกลวิธีดังกล่าวจะสำรวจความเข้าใจของผู้เรียนเกี่ยวกับเรื่องที่สนทนาและเขียน คำถามในลักษณะวัฏจักรสามารถนำไปปรับใช้ในการอภิปรายในวงกว้างกับสื่อการเรียนรู้ เช่น กรณีศึกษา วิดีโอ หรือการเรียนรู้ผ่านประสบการณ์ เช่น จากการทดลองในห้องปฏิบัติ หรือจากประสบการณ์ที่ได้รับการฝึกฝน สำหรับการใช้รูปแบบคำถามในลักษณะวัฏจักรสามารถทำได้โดยการกำหนดขอบเขตดังภาพ 2.4 ดังนี้



ภาพ 2.4 วิธีการใช้คำถามในลักษณะวัฏจักร

ที่มา: Tofade et al. (2013: 6)

จากภาพ 2.4 พบว่าวิธีการใช้คำถามในลักษณะวัฏจักร จะเกี่ยวข้องกับขอบเขต 3 ส่วนหลัก ซึ่ง ประกอบด้วยขอบเขตของเนื้อหา ขอบเขตการตอบของแต่ละคน และขอบเขตบริบทความเป็นจริงภายนอก โดยมีรายละเอียดดังนี้

1. ขอบเขตของเนื้อหา ประกอบด้วยความรู้ที่เป็นข้อเท็จจริง ความรู้ที่เป็นความคิดเชิงมโนทัศน์ และความรู้ที่เป็นกระบวนการ ที่ได้จากการอ่าน การเรียนรู้จากสื่อการเรียน หรือจากประสบการณ์

2. ขอบเขตการตอบของแต่ละคน ได้แก่ ปฏิกิริยาของแต่ละคน การรับรู้และประสบการณ์ตรงเกี่ยวกับเนื้อหาของแต่ละคน

3. ขอบเขตบริบทความเป็นจริงภายนอก จะเป็นเรื่องเกี่ยวกับความสัมพันธ์ระหว่างบริบททางสังคมกับเนื้อหาสาระที่เรียนรู้ รวมถึงเนื้อหาสาระอื่น ๆ ที่เกี่ยวข้องด้วย

นอกจากนี้แต่ละส่วนหลักประกอบด้วยส่วนที่คาบเกี่ยวกันอีก 4 ส่วน ได้แก่ ส่วนที่คาบเกี่ยวกันระหว่างเนื้อหาและการตอบของแต่ละคน ส่วนที่คาบเกี่ยวกันระหว่างเนื้อหากับบริบทความเป็นจริงภายนอก ส่วนที่คาบเกี่ยวกันระหว่างการตอบของแต่ละคนกับบริบทความเป็นจริงภายนอก และส่วนที่คาบเกี่ยวกันของทั้งสามขอบเขตหลักซึ่งอยู่ตรงกลาง สำหรับตัวอย่างวิธีการใช้คำถามในลักษณะวัฏจักรเป็นดังตาราง 2.5 ต่อไปนี้

ตารางที่ 2.5 ตัวอย่างวิธีการใช้คำถามในลักษณะวัฏจักร

วัฏจักร	ตัวอย่างคำถาม
เนื้อหา	นักศึกษาจะสร้างข้อสอบเพื่อวัดระดับความรู้ของผู้เรียนตามแนวคิดของบลูมได้อย่างไรบ้าง
การตอบของแต่ละคน	นักศึกษาจะใช้วิธีการสร้างข้อสอบตามแนวคิดของบลูมในสถานการณ์ใด
ความเป็นจริงภายนอก	นโยบายการวัดและประเมินผลผู้เรียน ตามหลักสูตรแกนกลาง การศึกษาขั้นพื้นฐาน พ.ศ. 2551 มีการเน้นเกี่ยวกับการสร้างข้อสอบตามแนวคิดของบลูมหรือไม่ อย่างไร

เนื้อหาและการตอบของแต่ละคน (Enriched)	การสร้างข้อสอบเพื่อวัดระดับความรู้ของผู้เรียนตามแนวคิดของบลูมใช้ได้กับทุกระดับการศึกษาหรือไม่ เพราะเหตุใด
เนื้อหาและความเป็นจริงภายนอก (Enriched)	การสร้างข้อสอบเพื่อวัดระดับความรู้ของผู้เรียนตามแนวคิดของบลูมมีความเป็นไปได้มากน้อยเพียงใดในการยกระดับคุณภาพการศึกษาของไทยในปัจจุบัน
การตอบของแต่ละคนและความเป็นจริงภายนอก (Enriched)	ถ้าการสร้างข้อสอบเพื่อวัดระดับความรู้ของผู้เรียนตามแนวคิดของบลูมใช้ไม่ได้ผลกับการยกระดับคุณภาพการศึกษาไทย นักศึกษาจะแก้ไขปัญหาดังกล่าวอย่างไร
เนื้อหาและการตอบของแต่ละคน และความเป็นจริงภายนอก (Dense)	หลังจากที่นักศึกษาได้เข้าใจวิธีการสร้างข้อสอบตามแนวคิดของบลูม และนำไปทดลองใช้กับนักเรียนแต่กลับไม่ได้ผล จากสถานการณ์ดังกล่าว นักศึกษาคิดว่าสิ่งที่เกิดขึ้นภายใต้บริบทความเป็นจริงภายนอกคืออะไร อย่างไร และทำไมถึงเป็นเช่นนั้น

พอลล์ (Paul, 1988 Quoted in Fahim & Bagheri, 2012: 1124) ได้จัดหมวดหมู่คำถามของโสเครติส เป็น 5 ประเภท ดังนี้

1. คำถามประเภททำให้เกิดความเข้าใจ เช่น ทำไมถึงกล่าวหรือพูดเช่นนั้น มีความสัมพันธ์อย่างไรกับสิ่งที่พวกเราได้อภิปรายหรือพูดคุยกัน

2. คำถามที่ตรวจสอบข้อสันนิษฐาน เช่น อะไรสามารถแทนข้อสันนิษฐานนี้ได้ สามารถตรวจสอบ ทวนสอบ หรือไม่เห็นด้วยกับข้อสันนิษฐานอย่างไร

3. คำถามที่ตรวจสอบเหตุผลและหลักฐาน เช่น การให้ยกตัวอย่าง อะไรเป็นสิ่งที่คล้ายคลึงกับ.... อะไรเป็นสาเหตุที่ทำให้เกิดผลเช่นนี้ขึ้น ทำไม เพราะเหตุใด

4. คำถามเกี่ยวกับมุมมองและความสำคัญ เช่น อะไรควรเป็นสิ่งที่ถูกเลือก อธิบายและให้เหตุผลว่าทำไม่ถึงจำเป็นหรือมีประโยชน์ ทำไมสิ่งนี้ถึงดีที่สุด อะไรคือจุดแข็งและจุดอ่อนของ.... อะไรคือข้อโต้แย้งในเรื่องนี้

5. คำถามเพื่อตรวจสอบการนำไปใช้และผลที่ตามมา เช่น อะไรคือการสรุปหรืออ้างอิงที่ทำได้ ผลที่ตามมาคืออะไรจากข้อสันนิษฐานดังกล่าวนี้ อะไรที่สามารถนำไปใช้ได้ ผลที่เกิดขึ้นเป็นอย่างไร จะเชื่อมโยงสิ่งที่เรียนรู้ได้อย่างไร

สรุปท้ายบท

อนุกรมวิธานของบลูมเป็นสิ่งสำคัญที่ควรถูกนำมาใช้ในการพัฒนาผู้เรียน เพราะเมื่อผู้เรียนมีระดับการเรียนรู้ตามอนุกรมวิธานของบลูม และได้ไปสู่โลกของการใช้ชีวิตหรือการประกอบอาชีพ ผู้เรียนสามารถที่จะวิเคราะห์ สร้างสรรค์ และนำสิ่งที่เรียนรู้ไปประยุกต์ใช้ได้อย่างทันทีทันใด โดยเฉพาะอย่างยิ่งเมื่อผู้เรียนพบสถานการณ์ที่จำเป็นต้องแก้ปัญหา ระดับการเรียนรู้แค่ความจำเกี่ยวกับข้อเท็จจริงและสารสนเทศไม่เพียงพอต่อการแก้ปัญหา ดังนั้นต้องเป็นหน้าที่ของผู้สอนที่จะต้องให้คำแนะนำ ชี้แนะให้ผู้เรียนเกิดการคิดในระดับสูง ซึ่งลักษณะดังกล่าวนี้ผู้สอนสามารถพิจารณาได้จากอนุกรมวิธานของบลูม เพื่อนำมาใช้ในการวางแผนการจัดทำรายวิชา ความรู้ และทักษะที่ผู้เรียนควรมีตามมาตรฐานและตัวชี้วัดต่อไป การวัดพัฒนาการทางด้านพุทธิพิสัยและด้านทักษะพิสัย ได้รับการพัฒนาก้าวหน้ามาก สามารถทำได้ชัดเจน เชื่อถือได้มาก ให้คะแนนได้อย่างมั่นใจ แต่การวัดด้านจิตพิสัย แม้ว่าจะเป็นด้านที่สำคัญมากและ เป็นคุณลักษณะที่จำเป็นของการเป็นมนุษย์อย่างสมบูรณ์ แต่กลับได้รับการพัฒนาในระดับที่น่าพอใจ ได้น้อยกว่า วัดได้ยากกว่า และให้คะแนนอย่างมั่นใจยากกว่า อย่างไรก็ตามการใช้คำถามที่ฝึกให้ผู้เรียนได้คิด เพื่อยกระดับความคิดรวบยอดในสิ่งที่ได้เรียนรู้ สามารถวิเคราะห์ประเด็นต่าง ๆ ได้อย่างมีประสิทธิภาพ จัดได้ว่ามีความสำคัญและจำเป็นอย่างยิ่งต่อการพัฒนาผู้เรียน ที่สำคัญการใช้คำถามตามแนวคิดของโสเครติสยังมีความเชื่อมโยงกับแนวคิดของอนุกรมวิธานของบลูมในการยกระดับคุณภาพการศึกษาของผู้เรียนและของสังคมประเทศชาติสืบไป

เอกสารอ้างอิง

จริยา เสถบุตร และ ธีระยุทธ นนทะสร. (ม.ป.ป.). สืบค้นเมื่อวันที่ 10 ตุลาคม 2556 จาก www.curric.net/center/Bloom_Taxonomy.doc

บุญชม ศรีสะอาด. (ม.ป.ป.) การวัดคุณลักษณะด้านจิตพิสัย. สืบค้นเมื่อวันที่ 4 มีนาคม 2556 จาก www.watpon.com/boonchom/09.doc

สมนึก ภัททิยธนี. (2549). การวัดผลการศึกษา. พิมพ์ครั้งที่ 5 กาฬสินธุ์: โรงพิมพ์ประสานการพิมพ์.

Anderson L. W., & Krathwohl, D. R. (2001). A Taxonomy for Learning, Teaching, and Assessing. Addison Wesley Longman, Inc.

Bissell, A. N., & Lemons, P.P. (2006). A New Method For Assessing Critical Thinking in The Classroom. *Bioscience, 56(1),* 66-72.

Eber, Patricia A., & Parker, Trent S. (2007). Assessing Student Learning: Applying Bloom's Taxonomy. *Human Services, 27(1),* 45-53.

Fahim, M., & Bagheri, M. B. (2012). Fostering Critical Thinking through Socrates' Questioning in Iranian Language Institutes. *Journal of Language Teaching and Research, 3(6),* 1122-1127.

Savickiene, I. (2010). Conception of Learning Outcomes in The Bloom’s Taxonomy Affective Domain. *The Quality of Higher Education, 7,* 37-58.

Tofade, T., Elsner, J., & Haines, S. T. (2013). Best Practice Strategies for Effective Use of Questions as a Teaching Tool. *American Journal of Pharmaceutical Educatio,77(7),* 1-9.

เอกสาร

ข้อสอบแบบเลือกตอบ

เครื่องมือที่ใช้ในการวัดผลทางการศึกษามีลักษณะการใช้แตกต่างกัน แต่นำมาใช้ในการวัดเพื่อให้ครอบคลุมพฤติกรรมด้านพุทธิพิสัย จิตพิสัยและทักษะพิสัย ซึ่งผู้สอนควรเลือกใช้ให้เหมาะสมกับธรรมชาติของการเรียนรู้ เครื่องมือที่ใช้ในการวัดผลทางการศึกษาที่นิยมใช้ได้แก่การทดสอบการสัมภาษณ์การสอบถามการสังเกตการตรวจผลงานการใช้แฟ้มสะสมงานเป็นต้นแต่ละวิธีสามารถใช้เครื่องมือวัดได้แตกต่างกันตามความเหมาะสม (ผดุงชัย ภู่พัฒน์, ม.ป.ป.: 2) อย่างไรก็ตามเครื่องมือที่ใช้ในการวัดผลทางการศึกษาที่ได้รับความนิยมกันอย่างกว้างขวางและแพร่หลาย คือ แบบทดสอบ ซึ่งเป็นเครื่องมือวัดพฤติกรรมด้านพุทธิพิสัยและมีบทบาทสำคัญมากเพราะเป็นเครื่องมือที่มีลักษณะดีหลายประการ เช่น ความเที่ยงตรง ความเชื่อมั่น ความเป็นปรนัย อำนาจจำแนก และความยากง่าย โดยเฉพาะอย่างยิ่งข้อสอบแบบเลือกตอบ ซึ่งมีรูปแบบและหลักการสร้างที่มุ่งเน้นถึงคุณภาพตามที่กล่าวข้างต้น ดังนั้นในบทนี้จะกล่าวถึง แนวคิดเกี่ยวกับข้อสอบและการสร้างข้อสอบ ประเภทของข้อสอบ การสร้างข้อสอบวัดพฤติกรรมด้านพุทธิพิสัยการวิเคราะห์ข้อสอบ การหาความเที่ยงตรงของแบบทดสอบ และการหาค่าความเชื่อมั่นของแบบทดสอบ

แนวคิดเกี่ยวกับข้อสอบและการสร้างข้อสอบ

การประเมินเป็นส่วนสำคัญในการเรียนการสอนและการเรียนรู้ ทั้งนี้เพราะการประเมินที่มีความเที่ยงตรงและความเชื่อมั่นจะเป็นกระจกสะท้อนข้อมูลและสารสนเทศของผู้เรียนได้อย่างถูกต้อง ไม่ว่าจะเป็นระดับสติปัญญา ความถนัดทางวิชาการ ผลสัมฤทธิ์ทางการเรียน แบบการเรียน ระดับความสนใจและตั้งใจ พัฒนาการทางด้านสังคม และวุฒิภาวะทางอารมณ์ ของผู้เรียน (Aiken, 1988; Okumbe, 2005; Walsh and Betz, 1985 Quoted in Asim et al. 2013: 13)ทางด้านการศึกษาการทดสอบเป็นเครื่องมือที่มีความสำคัญและเป็นกระบวนการการประเมินด้านจิตวิทยา โดยมีจุดมุ่งหมายหลักคือ การปรับปรุงการเรียนรู้ของผู้เรียน

สิ่งสำคัญอีกอย่างหนึ่งที่เกี่ยวข้องกับผู้สอนโดยตรง ได้แก่ การกำหนดระดับความรู้ในการจัดการเรียนรู้ว่าผู้เรียนควรรู้ระดับใด ตั้งแต่ระดับข้อเท็จจริง (Fact) ระดับความคิดรวบยอด (Concept) ระดับกระบวนการ (Procedure) และสุดท้ายที่ระดับอภิปัญญา (Meta cognition)

เพราะสิ่งที่กล่าวมานี้จะเป็นส่วนหนึ่งในการพัฒนาผู้เรียนให้มีองค์ความรู้ไปสู่มาตรฐาน และโลกของการประกอบอาชีพในการวัดและประเมินผู้เรียนมักจะใช้วิธีการทดสอบผู้เรียน ไม่ว่าจะเป็นเลือกตอบ จับคู่ ถูก-ผิด และเขียนตอบสั้น ๆ ซึ่งทั้งหมดนี้รวมเรียกว่าแบบปรนัย เป็นการให้ผู้เรียนได้เลือกตอบในสิ่งที่ผู้สอนกำหนดขอบเขตมาให้ คำถามอยู่ที่ว่าเป็นการเพียงพอหรือไม่ในการเตรียมผู้เรียนจากการวัดและประเมินโดยใช้เพียงแบบปรนัยในสังคมฐานความรู้หรือในยุคศตวรรษที่ 21 นี้ และถ้าผู้สอนปรับเปลี่ยนแนวคิดในการวัดและประเมินผู้เรียนเพิ่มขึ้นจากที่เคยทำโดยพยายามจัดการเรียนการสอนโดยให้ผู้เรียนเป็นผู้สร้างคำตอบ ได้แก่ การเขียนอนุทิน การเขียนเรียงความ การทำบันทึกการเรียน การนำเสนองานหน้าชั้นเรียน (Bible et al., 2008) ก็ย่อมจะสะท้อนถึงระดับความรู้ ความสามารถของผู้เรียนได้ตามสภาพจริงได้มากยิ่งขึ้น

ข้อสอบที่ผู้สอนเป็นผู้สร้าง (Teacher-made tests) และการประเมินแบบอื่น ๆ มีเป้าหมายหลักในการพัฒนาและปรับปรุงการเรียนรู้ของผู้เรียน การประเมินที่ดีต้องแสดงให้ผู้สอนเห็นว่าผู้เรียนเข้าใจในสิ่งที่เรียนอย่างแท้จริง ได้เห็นความสามารถของผู้เรียนในการใช้ระดับการคิด การประเมินเป็นสิ่งที่ต้องทำอย่างต่อเนื่องและสม่ำเสมอ ไม่ใช่ประเมินเพียงในตอนท้ายบทเรียนหรือหน่วยการเรียนกล่าวคือการประเมินควรดำเนินการอย่างเพียงพอและบ่อยครั้งเพื่อให้ผู้สอนรับรู้ว่าผู้เรียนมีความก้าวหน้ามากน้อยเพียงใดในสิ่งที่เรียนรู้ ซึ่งโดยปกติผู้สอนได้ทำการประเมินผู้เรียนเป็นประจำอยู่ตลอดเวลาในขณะที่มีการเรียนการสอน ไม่ว่าจะเป็นการสังเกตอย่างไม่เป็นทางการ และการทดสอบโดยเฉพาะอย่างยิ่งการทดสอบที่ใช้ข้อสอบแบบเลือกตอบจะได้รับความนิยมและถูกนำมาใช้อย่างกว้างขวางในสถานศึกษา (Gronlund, 1985; Omoifo, 2005 Quoted in Asim et al., 2013: 13)

ประเภทของข้อสอบ

Tileston (2004: 35-44) ได้รวบรวมประเภทของข้อคำถามที่ครูเป็นผู้สร้าง จำนวน 6 ประเภท โดยมีรายละเอียดพอสังเขปดังต่อไปนี้

1. คำถามที่ให้เลือกตอบ (Forced-Choice Questions)

สติ๊กกินส์ (Siggins, 1994) ได้ให้ความหมายการประเมินโดยใช้คำถามที่ให้เลือกตอบว่าเป็นการทดสอบที่ให้ผู้สอบเป็นผู้เลือกคำตอบที่ถูกต้องที่สุดจากชุดคำถามและคำตอบที่กำหนดให้ โดยมี

ดัชนีวัดผลสัมฤทธิ์ของผู้เรียนคือจำนวนสัดส่วนที่ผู้เรียนตอบคำถามได้ถูกต้องจากจำนวนคำถามทั้งหมด

ประเภทของคำถามที่ให้เลือกตอบ ได้แก่ เลือกตอบแบบหลายตัวเลือก เลือกตอบแบบ 2 ตัวเลือก แบบถูก-ผิด แบบให้เติมคำตอบลงในช่องว่าง โดยส่วนมากคำถามที่ให้เลือกตอบจะถูกนำมาใช้ในการวัดและประเมินผู้เรียนในการสอบระดับชาติ เช่น การสอบ O-Net, A-Net, V-Net หรือการทดสอบมาตรฐาน สำหรับคำถามที่ให้เลือกตอบเป็นการวัดและประเมินที่ผู้สอนควรนำมาใช้เมื่อต้องการทดสอบว่าผู้เรียนมีความรู้มากน้อยเพียงใดในเรื่องที่ได้เรียนรู้ไป เช่น การวัดความรู้เกี่ยวกับข้อเท็จจริง สูตรต่าง ๆ และสาระเนื้อหาที่เรียน จากนี้ขอเสนอประเภทของคำถามที่ให้เลือกตอบ 6 ประเภทตามลำดับดังนี้

1.1 เลือกตอบแบบหลายตัวเลือก (Multiple Choice)

แบบเลือกตอบเป็นการวัดและประเมินแบบดั้งเดิม ที่มีส่วนประกอบด้วยตัวคำถาม ตัวเลือกที่เป็นคำตอบที่ถูกต้อง และตัวเลือกที่เป็นตัวลวง ตัวอย่างเช่น

ข้อใดคือความหมายของการประเมินผล (ตัวคำถาม)

ก. การเก็บรวบรวมข้อมูล (ตัวลวง)

ข. กระบวนการในการตัดสินใจ(ตัวเลือกที่เป็นคำตอบที่ถูกต้อง)

ค. กระบวนการการควบคุม (ตัวลวง)

แบบเลือกตอบที่ดีเป็นเครื่องมือที่ใช้เวลาในการสร้างยาวนาน บางครั้งการกำหนดตัวเลือกที่เป็นตัวลวงก็ทำได้ยาก เพราะต้องเป็นสิ่งที่จะทำให้ผู้เรียนที่มีความรู้ ความเข้าใจในเนื้อหาสามารถพิจารณาได้ว่าตัวเลือกที่เป็นตัวลวงนั้นผิด และต้องเป็นสิ่งที่ทำให้ผู้เรียนที่ไม่มีความรู้ ความเข้าใจในเนื้อหาอาจเลือกหรือเดาในการตอบ (Marzano, 2000)

Asimและคณะ (2013: 14) กล่าวว่าข้อสอบแบบเลือกตอบมี 3 ประเภท ได้แก่

1. ประเภทประโยคที่ไม่สมบูรณ์ (Incomplete-statement type) คำถามของประเภทนี้จะเสนอประโยคที่ไม่สมบูรณ์ และให้ผู้สอบเลือกคำตอบที่ทำให้ประโยคที่กำหนดให้เป็นจริง เช่น “In the year _____ Nigeria became a republic” หรือ “ The year in which Nigeria became a republic was? _____” (a) 1914 (b) 1960 (c) 1963 (d) 1970

2. ประเภทคำถามแบบตรง (Direct-question type)คำถามประเภทนี้จะเกี่ยวข้องกับข้อเท็จจริงและมักจะใช้คำว่า อะไร เมื่อไหร่ ใคร หรือที่ไหน ตัวอย่างเช่น

“เมืองหลวงของประเทศไนเจียเรียชื่อว่าอะไร”

ก. Abuja ข.Lagos ค.Calabor ง.Lokoja

3.ประเภทคำตอบที่ดีที่สุด (Best-answer type) คำถามประเภทนี้เน้นการใช้ทักษะการคิดขั้นสูง เพราะเป็นการวัดผลการเรียนรู้ที่ซับซ้อน กล่าวคือผู้สอบจำเป็นต้องมีระดับความเข้าใจ การนำไปใช้ รวมถึงการแปลความ ตีความหมายเกี่ยวกับข้อเท็จจริงและสารสนเทศเพื่อที่จะสามารถเลือกคำตอบที่ถูกต้องได้ นอกจากนี้การสร้างตัวเลือกที่เป็นตัวลวงเป็นสิ่งที่มีความสำคัญและจำเป็นของคำถามประเภทนี้เช่นเดียวกัน ตัวอย่างของประเภทคำถามนี้ ได้แก่

“ปัจจัยในข้อใดต่อไปนี้ เป็นเหตุผลหลักในการเลือก Abuja เป็นเมืองหลวงของประเทศไนเจียเรีย”

ก. ประชากรมีจำนวนมาก ข. อากาศและบรรยากาศดี

ค. เป็นศูนย์กลางของประเทศ ง. โครงสร้างภายในของเมืองมีความสะดวก

1.2 แบบจับคู่ (Matching)

ข้อสอบจับคู่เป็นการกำหนดคำถามและคำตอบที่ให้เลือก โดยคำตอบที่ให้เลือกควรมีจำนวนไม่น้อยกว่าคำถามและควรมีมากกว่าคำถามประมาณ 2 เท่าตัว สำหรับตัวอย่างแบบจับคู่มีดังต่อไปนี้

คำถาม	คำตอบที่ให้เลือก
1. 5×12=.....	1. 144
2. 3×11=.....	2. 60
3. 8×12=.....	3. 222
4. 6×11=.....	4. 66
	5. 96
	6. 31
	7. 26
	8. 33

1.3 แบบให้เลือก 2 ตัวเลือก (Alternate Choice)

ข้อสอบแบบให้เลือก 2 ตัวเลือกมีลักษณะคล้ายกับข้อสอบแบบหลายตัวเลือก ต่างกันตรงที่มีให้เลือกเพียง 2 ตัวเลือก ดังมีตัวอย่างต่อไปนี้

1. (ก. Assessment ข. Evaluation) เป็นกระบวนการในการตัดสินใจ

2. (ก. การประเมินแบบย่อย ข. การประเมินแบบรวม) จะมีการดำเนินการก่อนและระหว่างการเรียนการสอน

รูปแบบดังกล่าวจะมีประสิทธิภาพเมื่อผู้สอนต้องการทดสอบเพื่อตรวจสอบความรู้เกี่ยวกับความหมายของคำหรือวลีที่ผู้เรียนควรรู้ แต่ปัญหาของข้อสอบลักษณะนี้ได้แก่โอกาสในการเลือกตอบได้อย่างถูกต้องจากการเดาคำตอบเป็นไปได้ถึงร้อยละ 50

1.4 แบบถูก – ผิด (True/False)

รูปแบบของข้อสอบลักษณะนี้ ผู้เรียนต้องระบุให้ได้ว่าข้อความหรือประโยคที่กำหนดให้เป็นจริง หรือเป็นเท็จ ตัวอย่างเช่น

_____ 1. ประเทศสุดท้ายที่เข้าร่วมกับกลุ่มอาเซียนคือประเทศพม่า

_____ 2. AEC ย่อมาจาก Asean Economic Community

_____ 3. ประเทศในกลุ่มอาเซียนมีทั้งหมด 10 ประเทศ

ประเภทนี้ดูเหมือนจะสร้างได้ไม่ยาก แต่โดยแท้จริงแล้วแบบถูก-ผิดที่ดีนั้นสร้างได้ยาก ข้อความหรือประโยคที่ผิดมีลักษณะการสร้างคล้ายกับการสร้างตัวเลือกที่เป็นตัวลวงในแบบหลายตัวเลือก กล่าวคือ ข้อความหรือประโยคที่ผิด ควรมีลักษณะที่ทำให้ผู้เรียนที่ไม่มีความรู้จริง ๆ สนใจหรือพิจารณาว่าเป็นคำตอบที่ถูกต้อง นอกจากนี้ยังมีลักษณะคล้ายกับข้อสอบที่ให้เลือก 2 ตัวเลือกที่เปิดโอกาสให้ผู้เรียนมีโอกาสในการเดาคำตอบได้ถูกถึงร้อยละ 50

1.5 แบบที่มีคำตอบถูกมากกว่า 1 คำตอบ (Multiple Response)

ข้อสอบที่กำหนดให้มีคำตอบถูกมากกว่า 1 คำตอบ จะมีประโยชน์ต่อผู้สอนในการวัดและประเมินความรู้ของผู้เรียนมากกว่า 1 ลักษณะโดยตัวอย่างข้อสอบเป็นดังต่อไปนี้

คำถาม: ประเภทของการทดสอบแบบใดต่อไปนี้ มีผลอย่างมากต่อการประเมินกระบวนการความรู้ของผู้เรียนให้ตอบ “ใช่” ถ้ามีผลอย่างมากต่อการประเมินกระบวนการความรู้ และให้ตอบ “ไม่ใช่” ถ้าไม่มีผลอย่างมากต่อการประเมินกระบวนการความรู้

_____ การทดสอบที่มีตัวเลือกให้ตอบ

_____ การเขียนเรียงความ

_____ การประเมินตนเอง

_____ การให้งานปฏิบัติ

_____ การให้เขียนตอบสั้น ๆ

_____ การให้รายงานปากเปล่า

จากตัวอย่างข้อสอบข้างต้น ผู้เรียนต้องมีความรู้เกี่ยวกับผลกระทบเกี่ยวกับการประเมินที่กำหนดให้ข้างต้นทั้งหมด

1.6 แบบที่ให้เติมคำตอบลงในช่องว่าง

ข้อสอบที่ให้เติมคำตอบลงในช่องว่างผู้เรียนต้องหาคำตอบซึ่งอาจเป็นคำ หรือกลุ่มคำมาเติมลงในช่องว่างของคำถาม ตัวอย่างเช่น

บทบาทที่สำคัญที่สุดของการทดสอบคือ_____

ข้อสอบดังกล่าวนี้สร้างได้ง่ายมาก ทั้งนี้เพราะข้อคำถามนำมาจากประโยค หรือวลีที่ต้องการให้ผู้เรียนเติมให้สมบูรณ์

2. คำถามที่เป็นแบบอัตนัย (Essays)

คำถามที่เป็นแบบอัตนัยเป็นรูปแบบดั้งเดิมประเภทหนึ่งที่ผู้สอนนำมาใช้เพื่อวัดและประเมินผู้เรียน สำหรับลักษณะและตัวอย่างของคำถามเพื่อให้ผู้เรียนเขียนตอบ เป็นดังนี้

“จากการเรียนวิชาเคมีตั้งแต่เปิดภาคเรียน ได้มีการเรียนเกี่ยวกับหลักการ และกระบวนการในการวิเคราะห์ทางด้านเคมี แต่เนื่องจากเพื่อนของท่านมักจะขาดเรียน เพราะเหตุผล เช่น ไม่สบายและวิตกกังวลในการสอบวิชาเคมี ซึ่งจะมีการสอบในอีก 2 สัปดาห์จากนี้ เพื่อนคนนี้ขอให้ท่านช่วยอธิบายทุกสิ่งที่จำเป็นต้องรู้ในการสอบ ให้ท่านเขียนเรียงความเพื่ออธิบายแนวคิด และหลักการที่เพื่อนท่านควรรู้และเข้าใจก่อนสอบวิชาเคมีข้างต้นนี้ โดยเชื่อมโยงและแสดงความสัมพันธ์ระหว่างแนวคิด ข้อเท็จจริง หลักการและกระบวนการตามที่ท่านรู้และเข้าใจให้ครบถ้วน”

สำหรับจุดแข็งของการประเมินโดยการใช้คำถามที่เป็นแบบอัตนัย คือการสร้างข้อคำถาม กล่าวคือข้อคำถามต้องมีความชัดเจน เป็นปรนัย ต้องชัดเจนในสิ่งที่ต้องการประเมิน โดยจะเน้นให้ผู้เรียนแสดงออกในสิ่งที่รู้ ความสามารถในการใช้เหตุผล การแก้ปัญหา และการตัดสินใจ สิ่งสำคัญ

ในการนำมาใช้ในการประเมินได้แก่เกณฑ์การประเมิน (Rubric) ซึ่งจะช่วยให้ผู้เรียนรับรู้ว่าข้อคำถามต้องการวัดและประเมินตนเองในเรื่องอะไรและอย่างไร

2.1การเขียนตอบแบบสั้น

การเขียนตอบแบบสั้นเป็นส่วนย่อยหรือประเภทหนึ่งของการเขียนเรียงความที่ให้ผู้เรียนเขียนคำตอบสั้น ๆ เพื่อตอบคำถามที่กำหนดให้ ลักษณะของคำถามส่วนใหญ่เป็นสิ่งที่ผู้สอนต้องการรู้ในสิ่งที่ผู้เรียนเข้าใจจากการเรียน นอกจากนี้คำถามอาจถามระดับการคิดขั้นสูง แต่มีข้อจำกัดเรื่องความยาวของคำตอบที่ผู้เรียนจะเขียนตอบ สำหรับตัวอย่างคำถามเป็นดังต่อไปนี้

"อ่านโคลงสยามานุสติบทนี้ แล้วตอบคำถามต่อไปนี้

หากสยามยังอยู่ยั้ง ยืนยง
เราก็เหมือนอยู่คง ชีพด้วย
หากสยามพินาศลง ไทยอยู่ ได้ฤๅ
เราก็เหมือนมอดม้วย หมดสิ้นสกุลไทย"

1. ใจความของโคลงบทนี้เป็นประเภทใด

ตอบ_____

2. โคลงบาทที่สอง เป็นคำกล่าวลักษณะใด

ตอบ_____

2.2คำถามที่ให้ตอบสั้น ๆ (Short-Answer Questions)เป็นคำถามที่ต้องการให้ผู้เรียนเขียนคำตอบที่มักจะสั้นเป็นคำเดียว วลีเดียว หรือประโยคสั้น ๆ และมักจะถามเกี่ยวกับศัพท์ กฎ นิยาม ทฤษฎี และหลักการต่าง ๆ โดยคำตอบที่ได้ต้องเป็นประเภทตายตัวแน่นอน

3. รายงานปากเปล่า (Oral Reports)

รายงานปากเปล่าเป็นลักษณะที่คล้ายคลึงกับการเขียนเรียงความ ต่างกันตรงที่เป็นการนำเสนอด้วยการพูดนำเสนอ รายงานปากเปล่ามีจุดแข็งและจุดอ่อนเช่นเดียวกับการเขียนเรียงความ โดยเพิ่มเรื่องการประเมินความสามารถด้านการพูดนำเสนอของผู้เรียนเข้ามา นอกจากนี้ยังมีผลอย่างมากในการประเมินหัวข้อ การคิด การให้เหตุผลและทักษะการสื่อสาร

4. การปฏิบัติงาน (Performance Tasks)

การปฏิบัติงานเน้นให้ผู้เรียนแสดงสิ่งที่สามารถปฏิบัติหรือทำได้ในสิ่งที่ได้เรียนรู้ไป การปฏิบัติงานอาจใช้เช่นเดียวกับการปฏิบัติงานตามสภาพจริง

5. การสังเกตโดยผู้สอน (Teacher Observation)

การสังเกตโดยผู้สอนเป็นการประเมินอย่างไม่เป็นทางการที่ใช้ทดสอบพฤติกรรม เช่น ความสามารถในการทำงานร่วมกับผู้อื่น กระบวนการด้านทักษะของผู้เรียน อาทิ ความสามารถของผู้เรียนในการอ่าน เขียนกราฟ การสังเกตโดยผู้สอนมีผลกระทบอย่างมากต่อการประเมินกระบวนการ แต่มีผลกระทบน้อยในการประเมินการคิด การให้เหตุผลและทักษะการสื่อสาร

6. การประเมินตนเองของผู้เรียน(Student Self-Assessment)

การประเมินตนเองของผู้เรียนเป็นรูปแบบที่ใช้กันอย่างกว้างขวางในการประเมินในชั้นเรียน โดยให้ผู้เรียนประเมินการคิดของตนเองไม่ว่าจะเป็นการทำงาน กระบวนการ และการเรียนรู้ ซึ่งการประเมินตนเองมีความน่าเชื่อถือสูงใน 5 รายการต่อไปนี้ หัวข้อของงานที่ทำ กระบวนการในการทำงาน ทักษะการคิดและให้เหตุผล ทักษะการสื่อสาร และสิ่งที่ไม่ใช่การประเมินผลสัมฤทธิ์

การสร้างข้อสอบวัดพฤติกรรมด้านพุทธิพิสัย

หลักการสร้างข้อสอบเพื่อวัดพฤติกรรมด้านพุทธิพิสัย ซึ่งมี 6 ด้าน ดังนี้

1. ความรู้ความจำ

เป็นการวัดระดับความจำจะถามในสิ่งที่จำได้จากการเรียน ความหมายของคำศัพท์ คำถามประเภทนี้จะมีประโยชน์ในการประเมินความรู้โดยตรงของผู้เรียน ซึ่งคำถามจะเกี่ยวข้องกับการวัดข้อเท็จจริงที่ผู้เรียนควรรู้ และต้องไม่ใช่เป็นการคิดขั้นสูง เช่น วิหค แปลว่าอะไร เครื่องหมาย ! เรียกว่าอะไร เป็นต้น

2. ความเข้าใจ

เป็นความสามารถในการนำความรู้ความจำไปดัดแปลงปรับปรุงเพื่อให้สามารถจับใจความหรือเปรียบเทียบ แปลความ ตีความ ขยายความ เช่น คำว่า ขิงก็รา ข่าก็แรง หมายถึงอะไร เหงือกปลาทำหน้าที่คล้ายกับอวัยวะส่วนใดของคนเป็นต้น

3. การนำไปใช้

เป็นการวัดระดับการประยุกต์ใช้จะเน้นให้ผู้เรียนคิดขั้นสูง เพราะประเภทนี้จะประเมินความเข้าใจเกี่ยวกับความคิดรวบยอดและความสามารถที่จะเชื่อมโยงระหว่างแต่ละความคิดรวบยอด โดยผู้เรียนต้องประยุกต์ใช้ความรู้และความคิดรวบยอดในการแก้ปัญหาและตีความเช่น เราสามารถทานอาหารอะไรแทนข้าวได้ สูตรในการหาปริมาตรกล่องชอล์คคืออะไร เป็นต้น

4. การวิเคราะห์

เป็นการวัดความสามารถในการแยกแยะรายละเอียดสิ่งต่าง ๆ หรือเรื่องราวต่าง ๆ ว่ามีส่วนใดสำคัญที่สุด สองชิ้นส่วนใดสัมพันธ์กันมากที่สุด และชิ้นส่วนเหล่านั้นอยู่รวมกันได้โดยอาศัยหลักการใด สำหรับตัวอย่างข้อสอบได้แก่ นิ้วมือนิ้วใดสำคัญที่สุด ฆราวาส : บ้าน แล้ว พร : ? หรือปากทำงานร่วมกับอวัยวะใด

5. การสังเคราะห์

เป็นความสามารถในการผสมผสานเรื่องราว หรือสิ่งต่าง ๆ ตั้งแต่ 2 ชนิดขึ้นไปเข้าด้วยกันเพื่อสร้างเป็นเรื่องราวใหม่ หรือสิ่งใหม่ที่แปลกไปจากเดิม เช่น คนที่ไม่ลักขโมยเป็นคนดี แต่ถ้าจะให้ดียิ่งขึ้นเขาควรจะต้องประพฤติอะไรอีกอย่างหนึ่ง ทำอย่างไรการแข่งขันกีฬากับโรงเรียนอื่นจึงจะมีโอกาสชนะ เป็นต้น

6.การประเมินค่า

เป็นการวินิจฉัย ตัดสิน ลงสรุป ตีคุณค่า เรื่องราว ความคิด เหตุการณ์ต่าง ๆ ว่าดี เลว เหมาะสม ไม่เหมาะสม อย่างมีหลักเกณฑ์ เช่น ถ้ายึดหลักประชาธิปไตย การที่นางรจนาเลือกเจ้าเงาะถือว่าเป็นความผิดหรือไม่ ตามประเพณีไทย นางวันทองเป็นคนหรือไม่ เป็นต้น

จากหลักการสร้างข้อสอบเพื่อวัดพฤติกรรมด้านพุทธิพิสัย ซึ่งมี 6 ด้าน ข้างต้น ล้วน สายยศ และอังคณา สายยศ (2543 อ้างถึงใน พิสณุ ฟองศรี, 2551: 255) ได้เสนอตัวอย่างข้อสอบที่วัดสูงกว่าระดับความจำ ดังตารางที่ 3.1ต่อไปนี้

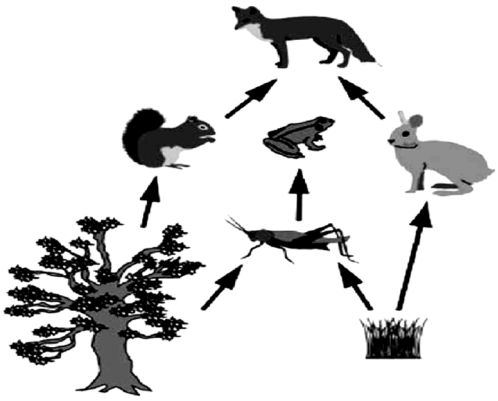
ตารางที่ 3.1 ตัวอย่างข้อสอบที่วัดสูงกว่าระดับความจำ

ข้อที่	พฤติกรรม	ข้อคำถาม	ตัวเลือก
1	ความเข้าใจ	ขิงก็ราข่าก็แรง หมายถึง	ก. ผัวเตะ เมียถีบ ข. ผัวด่า เมียร้องไห้ ค. ผัวยก เมียประคอง ง. ผัวทำงาน เมียเล่นไพ่ จ. ผัวกินเหล้า เมียนั่งบ่น
2	การนำไปใช้	คำถาม “ประเทศไทยมีกี่จังหวัด” วัดด้านใด	ก. จำ ข. เข้าใจ ค. นำไปใช้ ง. วิเคราะห์ จ. ประเมินค่า
3	การวิเคราะห์	หมุนสัมพันธ์กับคำใดมากที่สุด	ก. บิด ข. โค้ง ค. รอบ ง. กลม จ. เวียน
4	การสังเคราะห์	♦ นางงามทุกคนสวย ♦ คนสวยส่วนมากผิวขาว ♦ คนผิวขาวบางคนกลัวแดด จะสรุปได้อย่างไร	ก. คนสวยทุกคนผิวขาว ข. นางงามทุกคนผิวขาว ค. คนสวยส่วนมากกลัวแดด ง. นางงามบางคนกลัวแดด จ. ยังสรุปแน่นอนไม่ได้

5	การประเมินค่า	ถ้ายึดหลักประชาธิปไตย การที่รจนาเลือกเจ้าเงาะ ถือว่าเป็นความผิดหรือไม่	ก. ผิด เพราะไม่ทำตามผู้อื่น ข. ผิด เพราะทำตามความคิดเห็นตน ค. ผิด เพราะไม่เคารพพ่อแม่ ง. ไม่ผิด เพราะทำตามเสรีภาพ จ. ไม่ผิด เพราะทำตามด้วยความเสมอภาค

นอกจากนี้ยังมีลักษณะของการสร้างข้อสอบที่วัดด้านสติปัญญา การแก้ปัญหา หรือการคิดในลักษณะต่าง ๆ อาทิ คำถามที่ต้องการให้ผู้เรียนใช้เหตุผลจะเน้นการคิดขั้นสูงซึ่งผู้เรียนจะเชื่อมโยงความรู้ด้านเนื้อหาและทักษะกระบวนการและการให้เหตุผลเพื่อที่จะสร้างคำอธิบาย คำถามประเภทนี้จะให้ผู้เรียนแสดงความสามารถในการวิเคราะห์ ตีความหมายข้อมูล ประเมินผล แก้ปัญหาและประยุกต์ใช้ความเข้าใจในการแก้ปัญหาจากสถานการณ์สำหรับตัวอย่างคำถามในการวัดการให้เหตุผลเป็นดังนี้

จากรูปข้างล่างนี้ เป็นระบบนิเวศวิทยาของสิ่งมีชีวิตที่อาศัยในป่าแห่งหนึ่งซึ่งในระบบพบว่ากระรอกกินพืช และสุนัขจิ้งจอกกินกระรอก



ถ้าทราบว่าขณะนี้สุนัขจิ้งจอกตายหมด ข้อใดต่อไปนี้เป็นสิ่งที่อาจเกิดขึ้นกับจำนวนพืชในระบบนิเวศวิทยานี้

ก. ไม่มีอะไรเปลี่ยนแปลงเพราะสุนัขไม่ได้กินพืช

ข. ไม่มีอะไรเปลี่ยนแปลงเพราะสัตว์ชนิดอื่นจะมากินกระรอกแทน

ค. จำนวนลดลงเพราะจำนวนกระรอกที่รอดชีวิตมีมากขึ้น

ง. จำนวนเพิ่มขึ้นเพราะมีพื้นที่ที่พืชจะเจริญเติบโตมีมากขึ้น

จ. จำนวนลดลงเพราะสุนัขจิ้งจอกที่ทำหน้าปกป้องพืชได้ตายไป

ที่มา:Mosart, 2011 อ้างถึงใน Drits-Esserและ Black, 2014: 72

โดยสรุปจะพบว่าขั้นตอนการสร้างและพัฒนาแบบทดสอบผู้สอนจะต้องตัดสินใจว่าจะใช้แบบทดสอบประเภทไหนเช่นแบบทดสอบแบบเลือกตอบแบบทดสอบแบบถูกผิดแบบทดสอบแบบจับคู่ (Matching) แบบทดสอบแบบเติมคำหรือเติมข้อความและแบบทดสอบแบบเขียนตอบซึ่งไม่ว่าจะใช้แบบทดสอบประเภทไหนจะต้องสร้างแบบทดสอบให้มีคุณภาพซึ่งขั้นตอนการสร้างแบบทดสอบประกอบด้วย 8 ขั้นตอนคือ(ผดุงชัย ภู่พัฒน์, ม.ป.ป.: 11-12)

ขั้นตอนที่ 1 การระบุวัตถุประสงค์ของแบบทดสอบเช่นวัดพฤติกรรมการเรียนรู้คุณลักษณะที่พึงประสงค์หรือทักษะกระบวนการทางวิทยาศาสตร์

ขั้นตอนที่ 2 การระบุเนื้อหาและจุดประสงค์การเรียนรู้ที่ต้องการวัดเช่น

2.1 อธิบายหลักการและขั้นตอนการพัฒนาแบบทดสอบได้

2.2 คำนวณสถิติพื้นฐานสำหรับการวัดและประเมินผลได้

ขั้นตอนที่ 3 การระบุเงื่อนไขในการทดสอบคือ

3.1 สอบใครในที่นี้อาจเป็นนักศึกษาชั้นปีที่ 3 คณะศึกษาศาสตร์

3.2 สอบเมื่อไรก่อนเรียนระหว่างเรียนหลังเรียนสัปดาห์แรกหรือสัปดาห์สุดท้าย

3.3 ใช้เวลาสอบกี่นาทีซึ่งขึ้นอยู่กับผู้สอนเป็นผู้กำหนด

3.4 สอบด้วยแบบทดสอบแบบใดเช่น แบบทดสอบแบบเลือกตอบแบบทดสอบแบบถูกผิดแบบทดสอบแบบจับคู่ แบบทดสอบแบบเติมคำหรือเติมความและแบบทดสอบแบบเขียนตอบ

ขั้นตอนที่ 4 การทำแผนผังข้อสอบหรือพิมพ์เขียวแบบทดสอบ (Test Blueprint) หรือตารางโครงสร้างระหว่างเนื้อหา จุดประสงค์การเรียนรู้ กับพฤติกรรมการเรียนรู้หรือตาราง 2 มิติมิติหนึ่งคือเนื้อหาอีกมิติหนึ่งคือพฤติกรรมการเรียนรู้ดังตัวอย่างการกำหนดพิมพ์เขียวของแบบทดสอบระดับชั้นต่างๆ

ขั้นตอนที่ 5 เมื่อทำตารางเนื้อหาจุดประสงค์การเรียนรู้กับพฤติกรรมการเรียนรู้จำนวนข้อและคะแนนได้แล้วจึงลงมือออกข้อสอบตามจำนวนและรูปแบบที่ต้องการ

ขั้นตอนที่ 6 เมื่อออกข้อสอบแล้วผู้ออกข้อสอบจะต้องตรวจสอบคุณภาพของข้อสอบเบื้องต้นเกี่ยวกับตัวโจทย์คำถามและตัวเลือก

ขั้นตอนที่ 7 จัดทำฉบับเขียนคำสั่งคำชี้แจงในการตอบตรวจสอบความถูกต้องในการพิมพ์และใช้แบบทดสอบ

ขั้นตอนที่ 8 เมื่อนำแบบทดสอบไปสอบผู้เรียนแล้วผู้ออกข้อสอบควรวิเคราะห์หาคุณภาพข้อสอบเป็นรายข้อเช่นค่าความยาก ค่าอำนาจจำแนกและความเชื่อมั่นทั้งฉบับเพื่อนำข้อมูลไปแก้ไขปรับปรุงและอาจนำมาใช้ในคราวต่อไปหรือปีต่อไป

ลักษณะของแบบทดสอบที่ดี

การวัดพฤติกรรมด้านพุทธิพิสัย แบบทดสอบนับเป็นเครื่องมือวัดผลที่มีคุณค่าและสำคัญที่สุด ซึ่งแบบทดสอบที่จะนำไปใช้จะต้องมีคุณภาพ โดยเฉพาะแบบทดสอบที่ผู้สอนสร้างขึ้น หรือแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน ซึ่งเกี่ยวกับผู้สอนโดยตรง ทั้งนี้แบบทดสอบจะมีคุณภาพเพียงใด ต้องมีลักษณะที่ดี 10 ประการดังนี้ (เรวดี กระโหมวงศ์, ม.ป.ป.)

1. ความเที่ยงตรง (Validity)

ความเที่ยงตรง หมายถึง คุณภาพของแบบทดสอบ ที่สามารถวัดได้ตรงกับจุดมุ่งหมายที่ต้องการ หรือวัดในสิ่งที่ต้องการ หรือวัดในสิ่งที่ต้องการวัดได้อย่างถูกต้องแม่นยำ ความเที่ยงตรงจึงเปรียบเสมือนหัวใจของการทดสอบ เช่น ต้องการวัดว่าผู้เรียนมีความสามารถในการคำนวณหรือไม่ ก็ถามให้คำนวณ คะแนนที่ได้จะมากหรือน้อยขึ้นอยู่กับความสามารถในเชิงคำนวณ มิใช่ได้คะแนนมากเพราะสะอาด ใช้ภาษาได้สละสลวยลายมือสวยงาม เป็นต้น ความเที่ยงตรงเปรียบได้กับเครื่องชั่งที่สามารถชั่งได้ถูกต้อง ไม่โกงน้ำหนักหรือเปรียบได้กับนาฬิกาที่ได้ตรงเวลา ไม่คลาดเคลื่อนกับเวลาที่เป็นจริง ลักษณะความเที่ยงตรงของแบบทดสอบ แบ่งเป็น 4 ชนิด ดังนี้

1.1 ความเที่ยงตรงตามเนื้อหา (Content Validity)หมายถึง ความสามารถของแบบทดสอบที่วัดได้ตรงกับเนื้อหาที่กำหนดไว้ในหลักสูตร หรือตรงกับเนื้อหาที่ได้ทำการสอน กล่าวคือเมื่อทำการสอนเนื้อหาใดก็ทำการออกข้อสอบวัดให้ตรงกับเนื้อหานั้น และที่เน้นเป็นสำคัญที่ต้องเขียนคำถามให้สอดคล้องกับน้ำหนักความสำคัญของเนื้อหานั้นด้วย

1.2 ความเที่ยงตรงตามโครงสร้าง (Construction Validity)หมายถึง ความสามารถของแบบทดสอบ ที่วัดได้ตรงกับจุดมุ่งหมายที่กำหนดไว้ในหลักสูตร หรือวัดได้ตรงกับพฤติกรรมที่ต้องการให้เกิดกับผู้เรียน กล่าวคือ เมื่อจะสอนเนื้อหาใด แล้วจึงทำการสอนและเขียนข้อสอบให้ตรงกับพฤติกรรมที่ต้องการ เช่น ต้องการสอนเรื่องดิน โดยมีจุดม่งหมายว่าให้ผู้เรียนสามารถวิเคราะห์ชนิดของดินที่ควรใช้ในการเพาะปลูกได้ ก็ต้องสอนให้รู้จักวิเคราะห์ดิน และออกข้อสอบถามพฤติกรรมด้าน

การวิเคราะห์ มิใช่ออกข้อสอบถามเรื่องดินในลักษณะความรู้ความจำ และที่เน้นเป็นสำคัญอยู่ที่ต้องเขียนคำถามให้สอดคล้องกับน้ำหนักความสำคัญของพฤติกรรมด้วย

1.3 ความเที่ยงตรงตามสภาพ (Concurrent Validity)หมายถึงความสามารถของแบบทดสอบที่วัดได้ตรงสภาพความเป็นจริงในชีวิตประจำวัน หรือปัจจุบันของผู้เรียน หรือกล่าวได้ว่าเป็นความสามารถแบบทดสอบที่ช่วยให้ผู้สอนประมาณสถานภาพอันแท้จริง ของผู้เรียนในปัจจุบันได้ถูกต้อง เช่น นายแดงเก่งวิชาคณิตศาสตร์ ปรากฏว่าในชีวิตจริงก็มีทักษะด้านการคำนวณหรือ นายดำสอบได้คะแนนมากในวิชาพลานามัย เรื่องอนามัยส่วนบุคคล ปรากฏว่าในชีวิตจริงก็เป็นคนที่รักษาอนามัยส่วนบุคคลได้ดี ร่างกายแข็งแรงสมบูรณ์ แต่พบบ่อยครั้งว่าความรู้ในเนื้อหาวิชาของผู้เรียนมักจะตรงข้ามกับสภาพความเป็นจริงในชีวิตประจำวัน เช่น นาย ก. เก่งวิชาพลานามัย แต่ปรากฏว่าเป็นคนที่มีโรคประจำตัวไม่แข็งแรง เสื้อผ้า สกปรก ไม่รู้จักรักษาความสะอาด กล่าวได้ว่าแบบทดสอบชุดนี้ไม่มีความเที่ยงตรงตามสภาพ และโดยเนื้อแท้ของแบบทดสอบแล้วมักจะหาความเที่ยงตรงเช่นนี้ได้ยาก ดังนั้นการวัดผลผู้เรียนจึงไม่ควรใช้เฉพาะแบบทดสอบเพียงอย่างเดียว

1.4 ความเที่ยงตรงตามพยากรณ์ (Predictive Validity)หมายถึงความสามารถของแบบทดสอบที่วัดได้ตรงกับสภาพความเป็นจริงของผู้เรียน ที่เกิดขึ้นในอนาคต กล่าวคือ คะแนนผลการสอบที่เกิดจาก แบบทดสอบชุดนั้นสอดคล้องกับผลการเรียน หรือความสำเร็จในอนาคตของผู้เรียน เช่น แบบทดสอบคัดเลือกเพื่อเข้าเรียนสามารถทำนายได้ว่าคนที่สอบคัดเลือกผ่านนั้นจะเรียนได้สำเร็จ หรือคนที่สอบได้ที่สุดท้ายผลการเรียนก็ได้ที่สุดท้ายหรืออยู่ในกลุ่มอ่อน ความเที่ยงตรงตามการพยากรณ์นี้นับว่าเป็นคุณสมบัติที่สำคัญของแบบทดสอบคัดเลือก หรือแบบทดสอบวัดความถนัดหรือแบบทดสอบวัดเชาว์ปัญญา

สำหรับความเที่ยงตรงตามสภาพ และความเที่ยงตรงตามการพยากรณ์บางครั้งรวมเรียกว่าความเที่ยงตรงตามเกณฑ์สัมพันธ์ (Criterion Referenced Validity) เพราะเป็นเรื่องของความสอดคล้องของแบบทดสอบกับเกณฑ์ภายนอกที่นำมาพิจารณา กล่าวคือถ้าเกณฑ์ที่ใช้เป็นปัจจุบัน ก็จะเป็นความเที่ยงตรงตามสภาพ และถ้าเกณฑ์นั้นเป็นเรื่องในอนาคต ก็จะเป็นความเที่ยงตรงตามการพยากรณ์

2.ความเชื่อมั่น (Reliability)

ความเชื่อมั่น หมายถึง ลักษณะของแบบทดสอบทั้งฉบับที่สามารถวัดได้คงที่คงวาไม่เปลี่ยนแปลง ไม่ว่าจะทำการสอบใหม่กี่ครั้งก็ตาม เช่น สร้างแบบทดสอบชุดหนึ่งแล้วนำไปทดสอบกับผู้เรียนกลุ่มหนึ่ง 2 ครั้ง โดยให้มีเวลาห่างกันพอประมาณ (1-8 สัปดาห์) ถ้าพบว่า ผู้เรียนแต่ละคนทำคะแนนได้เท่าๆเดิม ทั้ง 2 ครั้งแสดงว่าแบบทดสอบชุดนี้มีความเชื่อมั่นสูง ไม่ใช่ผู้เรียนที่เข้าสอบมีความเชื่อมั่นสูงถ้านำคะแนนของแต่ละคนซึ่งมี 2 ชุด ไปเสนอข้อมูลรูปกราฟเส้น จะพบว่ากราฟทั้ง 2 เส้นนี้จะขนานหรือซ้อนทับกันเป็นส่วนใหญ่ แต่ถ้าความเชื่อมั่นของแบบทดสอบต่ำกราฟทั้ง 2 เส้นนี้จะสลับกันขึ้นลงแบบฟันปลาความเชื่อมั่นของแบบทดสอบ เปรียบได้กับเครื่องชั่ง กล่าวคือ ถ้านำเอาวัตถุชิ้นหนึ่งไปชั่งหลายๆครั้ง หากพบว่า เครื่องชั่งบอกน้ำหนักของวัตถุได้เท่าเดิมทุกครั้ง แสดงว่าเครื่องชั่งนั้นมีความเชื่อมั่นสูง อาจจะชั่งได้ไม่ถูกต้อง แม่นยำก็ได้ แต่คงสภาพเช่นนั้นตลอดไป

3. ความยุติธรรม (Fair)

ลักษณะของแบบทดสอบที่ไม่เปิดโอกาสให้มีการได้เปรียบ เสียเปรียบในกลุ่มผู้เข้าสอบด้วยกัน ไม่เปิดโอกาสให้ผู้เรียนทำข้อสอบได้โดยการเดา ไม่ให้ผู้เรียนขี้เกียจหรือไม่สนใจในการเรียนทำข้อสอบได้ดี ผู้ทำข้อสอบได้ ควรจะเป็นผู้เรียนที่เรียนเก่ง และขยันเท่านั้น วิธีการที่จะช่วยให้เกิดความยุติธรรม ได้แก่ ออกข้อสอบให้คลุมหลักสูตรและมีจำนวนมาก แบบทดสอบที่ใช้สอบกับผู้เรียนทุกคนต้องเป็นชุดเดียวกันและเป็นเรื่องที่ผู้เรียนเรียนแล้ว หากออกข้อสอบยากเกินไปจะทำให้ผู้เรียนเรียนเก่งเสียเปรียบ เพราะทุกคนต้องทำข้อสอบโดยการเดา

4. ความลึกของคำถาม (Searching)

ความลึกของคำถาม หมายถึง ข้อสอบแต่ละข้อนั้นจะต้องไม่ถามผิวเผิน หรือถามประเภทความรู้ความจำแต่ต้องถามให้ผู้เรียน นำความรู้ความเข้าใจไปคิดดัดแปลงแก้ปัญหา แล้วจึงตอบได้ ได้แก่ ถามพฤติกรรมที่สูงกว่าความรู้ความจำไม่ควรถามว่า โลกมีดวงจันทร์กี่ดวงแต่ควรถามว่าถ้าโลกมีดวงจันทร์ 2 ดวงหรือไม่มีเลยเหตุการณ์จะเป็นอย่างไร หรือไม่ควรถามว่าส่วนใดของพืชใช้ขยายพันธุ์ได้แต่ควรถามว่า สิ่งใดมีความสำคัญมากที่สุดต่อการงอกของเมล็ดพืช

5. ความยั่วยุ (Exemplary)

ความยั่วยุ หมายถึง แบบทดสอบที่ผู้เรียนทำด้วยความสนุกเพลิดเพลิน ไม่ควรซ้ำซากซึ่งน่าเบื่อหน่าย วิธีการที่ให้แบบทดสอบมีความยั่วยุอยากตอบ ก็โดยเรียงจากข้อง่ายไปหาข้อยากใช้ข้อสอบรูปภาพบ้าง รูปแบบของข้อสอบน่าสนใจ ถ้าเป็นข้อสอบแบบอัตนัยก็ให้บรรยายมีความยาวพอเหมาะและไม่ถามหลายประเด็นในข้อเดียวกัน

6. ความจำเพาะเจาะจง (Definition)

ความจำเพาะเจาะจง หมายถึง ข้อสอบที่มีแนวทาง หรือทิศทางการถามการตอบชัดเจน ไม่คลุมเครือ ไม่แฝงกลเม็ดให้ผู้เรียนสับสน ผู้เรียนไม่ได้คะแนนเนื่องจากตอบไม่ถูกดีกว่าไม่ได้คะแนนเพราะมาจากไม่เข้าใจคำถาม ความไม่จำเพาะเจาะจงของข้อสอบนี้อาจจะเกิดขึ้นได้กับข้อสอบทุกชนิด เช่น ถูก-ผิด จับคู่เติมคำตอบสั้น ๆ เลือกตอบและอัตนัย ทั้งนี้ขึ้นอยู่กับผู้เขียนข้อสอบว่าสามารถออกข้อสอบได้รัดกุมและชัดเจนเพียงใด

7. ความเป็นปรนัย (Objective)

ความเป็นปรนัยของแบบทดสอบ ไม่ได้หมายถึง ข้อสอบแบบถูก-ผิด จับคู่ เติมคำตอบสั้นๆและเลือกตอบ เพราะแบบทดสอบชนิดต่างๆเหล่านี้ เป็นเพียงรูปแบบหรือโครงสร้างของคำถามที่จะนำไปสู่ความเป็นปรนัยเท่านั้น และความเป็นปรนัยเป็นคุณลักษณะของแบบทดสอบไม่ใช่ชนิดของแบบทดสอบแบบทดสอบชนิดใดจะเป็นปรนัยหรือไม่ จะต้องมีคุณสมบัติ 3 ประการ คือ

7.1ตั้งคำถามให้ชัดเจน ทำให้ผู้เข้าสอบทุกคนเข้าใจความหมายตรงกัน

7.2 ตรวจให้คะแนนได้ตรงกัน แม้ว่าจะตรวจหลายครั้งหรือตรวจหลายคนก็ตาม

7.3 แปลความหมายของคะแนนได้เหมือนกัน ลักษณะเช่นนี้ถ้าเป็นข้อสอบที่ผู้เรียนได้คะแนนแต่ละข้อไม่เท่ากัน ได้แก่ข้อสอบอัตนัย หรือตอบสั้นๆหรือเติมคำก็ไม่สามารถแปลความหมายของคะแนนได้ เช่น ข้อสอบเลือกตอบถูก-ผิด หรือจับคู่ก็สามารถแปลความหมายของคะแนนได้ เช่น ข้อสอบอัตนัย 5 ข้อ เต็ม 50 คะแนน นายแดงทำได้ 32 คะแนนหมายความว่านายดำทำถูก 32 ข้อ และสามารถแปลความต่อไปได้ว่า ข้อสอบแต่ละข้อยากหรือง่ายมีอำนาจจำแนกสูงหรือต่ำ จะเห็นว่าข้อสอบแบบอัตนัย หรือความเรียงอาจเป็นปรนัยก็ได้ ถ้ามีคุณสมบัติครบทั้ง 3 ประการข้างต้น และในทางตรงกันข้ามข้อสอบแบบเลือกตอบถูก-ผิด หรือจับคู่ อาจจะไม่เป็นปรนัยก็ได้ถ้ามีคุณสมบัติไม่ครบทั้ง 3 ประการ

8. ประสิทธิภาพ (Efficiency)

ประสิทธิภาพ หมายถึง แบบทดสอบแบบที่มีจำนวนข้อมากพอประมาณ ใช้เวลาสอบพอเหมาะประหยัดค่าใช้จ่ายจัดทำแบบทดสอบด้วยความประณีต ตรวจให้คะแนนได้รวดเร็ว รวมถึงสถานการณ์ในการสอบที่ดี ได้แก่ สภาพห้องสอบเรียบร้อยไม่มีสิ่งรบกวนผู้เข้าสอบ กรรมการคุมสอบรัดกุม เป็นต้นนอกจากนี้หากสร้างแบบทดสอบไว้อย่างดี และสามารถนำไปใช้ได้หลายๆครั้งอย่างเหมาะสม โดยไม่เกิดความเสียหายใดๆ ถือได้ว่าแบบทดสอบนั้นมีประสิทธิภาพ

9. อำนาจจำแนก (Discrimination)

อำนาจจำแนก หมายถึง ความสามารถของข้อสอบในการจำแนกผู้สอบที่มีคุณลักษณะ หรือความสามารถแตกต่างกันออกจากกันได้ ข้อสอบที่ดี จะต้องมีอำนาจจำแนกสูง ตามทฤษฎีการวัดผลแบบอิงกลุ่ม (Norm Referenced Measurement) อำนาจจำแนกของข้อสอบ หมายถึง ความสามารถของข้อสอบที่จำแนกผู้สอบออกเป็น 2 กลุ่ม คือ กลุ่มเก่ง กับกลุ่มอ่อน ถ้าข้อสอบมีอำนาจจำแนกสูงแสดงว่า คนกลุ่มเก่งทำข้อสอบข้อนั้นถูก แต่คนกลุ่มอ่อนทำไม่ถูก ส่วนทฤษฎีการวัดผลแบบอิงเกณฑ์ (Criterion Referenced Measurement) ความสามารถของข้อสอบนั้นในการจำแนกผู้สอบออกเป็น 2 กลุ่ม คือ กลุ่มรอบรู้ กับกลุ่มไม่รอบรู้ถ้าข้อสอบมีอำนาจจำแนกสูง แสดงว่าคนกลุ่มรอบรู้ทำข้อสอบนั้นถูก แต่คนกลุ่มไม่รอบรู้ทำไม่ถูก

10. ความยาก (Difficulty)

ความยาก หมายถึง จำนวนคนตอบข้อสอบได้ถูกมากน้อยเพียงใด หรืออัตราส่วนของจำนวนคนตอบถูกกับจำนวนคนทั้งหมดที่เข้าสอบ ตามทฤษฎีการวัดผลแบบอิงกลุ่ม ข้อสอบที่ดีคือข้อสอบที่ไม่ยากหรือง่ายเกินไป เรียกว่าความยากพอเหมาะ เพราะคุณค่าของข้อสอบที่ทุกคนทำถูกต่างก็ไม่สามารถจำแนกผู้สอบได้ว่าใครเก่งใครอ่อน จึงไม่มีค่าในการจำแนก ส่วนทฤษฎีการวัดผลแบบอิงเกณฑ์ ถือว่าข้อสอบที่ดีคือสามารถวัดว่าผู้เรียนได้บรรลุจุดประสงค์หรือไม่ การที่ทุกคนทำข้อสอบได้ถูก แสดงว่าเขาบรรลุตามจุดประสงค์ตามที่ต้องการ ดังนั้นสิ่งสำคัญของข้อสอบอยู่ที่ว่าสามารถวัดในจุดประสงค์ที่ต้องการได้จริงหรือไม่ ถ้าวัดได้จริงก็นับว่าเป็นข้อสอบที่ดี แม้ว่าจะเป็นข้อสอบที่ง่ายก็ตาม

การวิเคราะห์ข้อสอบ (Item Analysis)

การวิเคราะห์ข้อสอบหมายถึงการตรวจสอบคุณภาพของแบบทดสอบที่สร้างขึ้นว่ามีคุณภาพดีเพียงใด ซึ่งเป็นการหาค่าสถิติที่แสดงคุณลักษณะของข้อสอบเป็นรายข้อและทั้งฉบับ ซึ่งตามทฤษฎีการทดสอบดั้งเดิม (Classical Test Theory) นิยมหาคุณลักษณะข้อสอบ 2 ชนิด คือค่าความยาก (Difficulty) และค่าอำนาจจำแนก (Discrimination) การวิเคราะห์ข้อสอบเป็นเทคนิคที่ทำให้ทราบว่าแต่ละข้อมีความยากเท่าใดและมีอำนาจจำแนกเท่าใด เพื่อพิจารณาคัดเลือกเอาข้อสอบที่มีคุณลักษณะเข้าเกณฑ์ไปใช้ต่อไป รายละเอียดต่อจากนี้จะกล่าวเฉพาะแนวคิดตามทฤษฎีดั้งเดิม (บุญชม ศรีสะอาด, ม.ป.ป.; สมนึก ภัททิพธนี, 2549: 193)

ค่าความยากหรือความยากง่ายเป็นค่าที่บอกให้ทราบว่าข้อนั้นยากเพียงใดโดยถือเอาจำนวนผู้ตอบถูกมากน้อยเป็นเกณฑ์ ถ้าตอบถูกหลายคนจัดว่าเป็นข้อสอบที่ง่าย ถ้าตอบถูกน้อยคนจัดว่าเป็นข้อที่ยาก ค่าความยากนิยมเขียนแทนด้วยสัญลักษณ์ P มีค่า 0.00 ถึง +1.00 ค่าอำนาจจำแนกเป็นค่าที่บอกให้ทราบว่า ข้อนั้นจำแนกกลุ่มผู้สอบได้ดีเพียงใด เช่นจำแนกระหว่างกลุ่มสูงกับกลุ่มต่ำ จำแนกระหว่างกลุ่มผู้สอบผ่านกับกลุ่มผู้สอบไม่ผ่าน จำแนกระหว่างกลุ่มเรียนแล้วกับกลุ่มยังไม่เรียน เป็นต้นค่าอำนาจจำแนกเขียนแทนด้วยสัญลักษณ์หลายอย่างขึ้นกับว่าจะเป็นค่าอำนาจจำแนกชนิดใด หรือหาโดยวิธีของใครค่าอำนาจจำแนกของข้อสอบ (Test Item) มีค่า -1.00 ถึง +1.00

เมื่อกล่าวถึงการวิเคราะห์ข้อสอบ โดยทั่วไปจะหมายถึงการวิเคราะห์ข้อสอบแบบเลือกตอบ ข้อสอบแบบอื่น ๆ ไม่นิยมนำมาวิเคราะห์ เนื่องจากข้อสอบแบบเลือกตอบมีคุณสมบัติเหมาะสมที่สุดกับการวิเคราะห์คุณภาพรายข้อ อย่างไรก็ตามเราสามารถทำการวิเคราะห์ข้อสอบแบบอื่นๆ ได้ทุกแบบ ดังจะแยกกล่าวเป็น 2 กรณี คือ การวิเคราะห์ข้อสอบแบบจำกัดคำตอบ กับการวิเคราะห์ข้อสอบแบบอัตนัย

1. การวิเคราะห์ข้อสอบแบบจำกัดคำตอบ

เมื่อพูดถึงข้อสอบแบบจำกัดคำตอบ มักหมายถึงข้อสอบแบบเลือกตอบแบบถูกผิด แบบจับคู่ และแบบเติมคำ ข้อสอบแบบจำกัดคำตอบที่ไม่ใช่ข้อสอบแบบเลือกตอบสามารถนำมาวิเคราะห์หาค่าความยาก และอำนาจจำแนกได้โดยใช้วิธีการเช่นเดียวกันกับการวิเคราะห์ข้อสอบแบบเลือกตอบ (กรณีคำตอบถูก) เพราะข้อสอบแบบถูกผิด ก็คือข้อสอบแบบเลือกตอบที่มีตัวเลือก 2 ตัว คือ ถูกหรือผิด ข้อสอบแบบจับคู่ก็คือข้อสอบแบบเลือกตอบที่ตัวเลือกลดลงเรื่อย ๆ ตามจำนวนข้อที่ได้ตอบไปแล้ว ข้อสอบแบบเติมคำอาจกล่าวได้ว่าเป็นข้อสอบแบบเลือกตอบชนิดหนึ่งที่มีตัวเลือกมากมายไม่จำกัดที่ผู้ตอบต้องเลือกคำตอบ (คิดตอบ) ออกมาเอง ในการวิเคราะห์จะถือเสมือนข้อสอบถูกผิด ถ้าเติมคำตอบถูกต้องก็เทียบได้กับที่ตอบข้อสอบถูกผิดได้ “ถูก” จะได้คะแนน 1 คะแนน และถ้าเติมคำตอบที่ไม่ถูกต้องก็เทียบได้กับที่ตอบข้อสอบถูกผิดได้ “ผิด” ซึ่งจะไม่ได้คะแนนของข้อนั้น จากลักษณะดังกล่าวเหล่านี้สามารถนำผลการสอบของข้อสอบแบบถูกผิด แบบจับคู่ และแบบเติมคำมาวิเคราะห์หาค่าความยากและอำนาจจำแนก โดยใช้สูตรและวิธีเดียวกันกับการวิเคราะห์ข้อสอบแบบเลือกตอบ

สำหรับ การวิเคราะห์ข้อสอบแบบจำกัดคำตอบ จะพิจารณา 3 ปัจจัย ดังต่อไปนี้

1. ความยากง่ายของข้อสอบ (item difficulty: P) เป็นค่าสัดส่วนของผู้สอบที่ตอบข้อสอบข้อนั้นถูก

P = (จำนวนนักเรียนที่ตอบข้อสอบถูก) ÷ n

เมื่อ P แทน ค่าดัชนีความยากง่าย

n แทน จำนวนผู้สอบทั้งหมดที่ตอบข้อสอบนั้น

โดยมีเกณฑ์การแปลความหมายค่า P ดังนี้

0.80 - 1.00 ง่ายมาก (ควรปรับปรุงหรือตัดทิ้ง)

0.60 - 0.79 ค่อนข้างง่าย

0.40 - 0.59 ปานกลาง

0.20 - 0.39 ค่อนข้างยาก

น้อยกว่า 0.20 ยากมาก (ควรปรับปรุงหรือตัดทิ้ง)

2. ความสามารถในการจำแนกผู้สอบตามระดับความสามารถ (item discrimination, r) เป็นคุณสมบัติที่บ่งบอกถึงความสามารถของข้อสอบที่จำแนกเด็กเก่งและอ่อน

$r = P_H - P_L$ (สำหรับตัวเลือกที่เป็นคำตอบที่ถูกต้อง)

หรือ $r = P_L - P_H$ (สำหรับตัวเลือกที่เป็นตัวลวง ใช้ในการหาประสิทธิภาพตัวลวง)

เกณฑ์การแปลความหมายค่า r (สำหรับตัวเลือกที่เป็นคำตอบที่ถูกต้อง)

1.00 จำแนกดีเลิศ

0.80 - 0.99 จำแนกดีมาก

0.60 - 0.79 จำแนกดี

0.40 - 0.59 จำแนกได้ปานกลาง

0.20 - 0.39 จำแนกได้บ้าง

0.00 - 0.19 จำแนกไม่ค่อยได้ (ควรปรับปรุง)

น้อยกว่า 0 จำแนกไม่ได้ (ควรตัดทิ้ง และสร้างข้อสอบใหม่)

3. ประสิทธิภาพของตัวลวง (distracter efficiency) พิจารณา 2 ส่วนดังนี้

3.1 ค่า P มีค่าตั้งแต่ 0.05 ขึ้นไป

3.2 ค่า r มีค่าตั้งแต่ 0.05 ขึ้นไป โดยใช้สูตร $r = P_L - P_H$

2. การวิเคราะห์ข้อสอบแบบอัตนัย

ข้อสอบแบบอัตนัย (Essay) เป็นข้อสอบที่ให้ผู้สอบเขียนตอบตามความคิดของตนเอง แม้ว่าโดยทั่วไปจะได้รับการนำไปใช้น้อย แต่ก็เป็นเครื่องมือที่จำเป็นในการวัดผลโดยเฉพาะในการวัดเกี่ยวกับความคิดเห็น เจตคติ ความสามารถในการคิดสังเคราะห์ (Synthesis) ความสามารถในการอธิบายให้คนอื่นเข้าใจ ความสามารถในการบูรณาการความรู้ การวัดในลักษณะเหล่านี้ข้อสอบแบบอัตนัยจะวัดได้ดีมาก ข้อสอบประมวลความรู้(Comprehension) ซึ่งเป็นข้อสอบที่วัดความรู้ขั้นสูงของนิสิตนักศึกษาในระดับปริญญาโท ปริญญาเอกก็ล้วนแต่เป็นข้อสอบแบบอัตนัย การที่จะตอบข้อสอบชนิดนี้ได้ต้องอาศัยความรอบรู้ ประสบการณ์ และความสามารถในการบูรณาการความรู้

การวิเคราะห์ข้อสอบแบบอัตนัย ดำเนินการได้ดังนี้

1. ตรวจให้คะแนนข้อสอบแต่ละข้อ แล้วรวมคะแนนทุกข้อ

2. เรียงกระดาษคำตอบจากคะแนนสูงสุดลงมาหาต่ำสุด คัดเอาเฉพาะผู้ที่ได้คะแนนสูงสุด 25% ของทั้งหมดเป็นกลุ่มสูง และผู้ที่ได้คะแนนต่ำสุด 25% ของทั้งหมดเป็นกลุ่มต่ำ กลุ่มที่เหลือเป็นกลุ่มกลางมีจำนวน 50% ของทั้งหมด ในการวิเคราะห์จะใช้เฉพาะผลการสอบของกลุ่มสูงและกลุ่มต่ำ ไม่ได้ใช้ผลการสอบของกลุ่มกลาง

3. บันทึกคะแนนของแต่ละคนในแต่ละข้อลงในตาราง โดยแยกตามกลุ่ม อาจใช้แบบฟอร์มดังตาราง รวมคะแนนแต่ละข้อของแต่ละกลุ่ม

4. หาค่าความยาก และอำนาจจำแนก โดยใช้สูตรดังนี้

ค่าความยากหาจากสูตร

$$P=\frac{\sum H+\sum L}{2NM} \quad \text{.....(1)}$$

ค่าอำนาจจำแนกหาจากสูตร

$$D=\frac{\sum H-\sum L}{NM} \quad \text{.....(2)}$$

เมื่อ P แทน ค่าความยาก

D แทน ค่าอำนาจจำแนก

$\sum H$ แทน ผลรวมของคะแนนกลุ่มสูง

$\sum L$ แทน ผลรวมของคะแนนกลุ่มต่ำ

N แทน จำนวนคนในแต่ละกลุ่ม

M แทน คะแนนเต็ม

นำค่าความยากและค่าอำนาจจำแนกบันทึกลงในตาราง ผลจากการหาค่าความยากและอำนาจจำแนกของข้อสอบแต่ละข้อตามสูตร 1 และ 2 ตามลำดับ

จากตารางเป็นผลการสอบวิชาวัดผลการศึกษาซึ่งมีจำนวนทั้งหมด 5 ข้อ คะแนนเต็มข้อละ 10 คะแนน นักศึกษาสาขาวิชาภาษาไทยที่สอบวิชาวัดผลการศึกษามีจำนวนทั้งหมด 16 คน จึงมีกลุ่มสูงและกลุ่มต่ำกลุ่มละ 4 คน ซึ่งเท่ากับ 25% ของ 16

ผู้ที่ได้คะแนนสูงสุดทำได้ 42 คะแนน ตอบข้อ 1 ถึงข้อ 5 ได้คะแนนตามลำดับดังนี้ 9, 8, 7, 8, 10 คะแนน ผู้ที่ได้คะแนนต่ำสุดทำได้ 16 คะแนน ตอบข้อ 1 ถึงข้อ 5 ได้คะแนนตามลำดับดังนี้ 2, 3, 2, 7, 2 คะแนน ข้อ 1 มีกลุ่มสูงทำได้คะแนนรวมทั้งหมด 33 คะแนน กลุ่มต่ำทำได้ 16 คะแนนดังตารางที่ 3.2

ตารางที่ 3.2 ตารางวิเคราะห์ข้อสอบแบบอัตนัยวิชาวัดผลการศึกษาของนักศึกษาสาขาวิชาเอกภาษาไทย

ข้อ / ผู้สอบ		1	2	3	4	5	รวม
กลุ่มสูง	1	9	8	7	8	10	42
	2	8	8	6	7	10	38
	3	8	7	6	7	9	37
	4	8	7	6	6	9	36
	รวม	33	30	25	28	38	
กลุ่มต่ำ	1	5	5	5	7	5	27
	2	5	5	3	8	4	25
	3	4	6	3	6	3	22
	4	2	3	2	7	2	16
	รวม	16	19	13	28	14	
ผลการวิเคราะห์	P	.61	.61	.47	.70	.65	
รวม	D	.42	.27	.30	.00	.60	

จากตารางที่ 3.2 เป็นข้อมูลสมมติ ซึ่งกำหนดจำนวนคนน้อยคนเพื่อให้สามารถเข้าใจวิธีการได้ง่าย ในการวิเคราะห์จริงควรมีจำนวนทั้งหมดไม่ต่ำกว่า 40 คน

ตัวอย่างการหาค่าความยาก และค่าอำนาจจำแนก ของข้อที่ 1

$$P = \frac{33+16}{8 \text{x} 10} = \frac{49}{80} = .61$$

ความยากของข้อ1 = .61

$$D = \frac{33+16}{4 \text{ x } 10} = \frac{17}{40} = .42$$

อำนาจจำแนกของข้อ 1 = .42

จากตารางพบว่า ข้อที่ 1, 2, 3, และ 5 มีคุณภาพเหมาะสม ส่วนข้อ 4 ถึงแม้ว่าค่าความยากจะใช้ได้ แต่ไม่มีอำนาจ จึงเป็นข้อที่ใช้ไม่ได้ทั้งนี้เพราะค่าความยากที่เหมาะสมเท่ากับ .20 ถึง .80 ค่าอำนาจจำแนกที่เหมาะสมเท่ากับ .20 ถึง 1.00

การหาความเที่ยงตรงของแบบทดสอบ

ความคิดรวบยอดของความเที่ยงตรงจะเกี่ยวข้องกับการประเมินเพื่อตัดสินได้อย่างเหมาะสม เช่น ถ้าการประเมินที่นำมาใช้ในการอธิบายผลสัมฤทธิ์ทางการเรียนของผู้เรียน ผู้สอนต้องอธิบายและตีความได้ว่าคะแนนที่ผู้เรียนทำได้จากการใช้เครื่องมือในการวัดนั้น มีความสอดคล้องกับผลสัมฤทธิ์ทางการเรียนได้อย่างไร ขอบเขตการวัดซึ่งเป็นตัวแทนของผลสัมฤทธิ์ทางการเรียนคืออะไร และเป็นอย่างไร (Linn & Miller, 2005: 68-69) และเมื่อมีการใช้ “ความเที่ยงตรง” ที่เกี่ยวข้องกับการทดสอบและการประเมิน ลินน์ และ มิลเลอร์ (Linn & Miller, 2005: 70-71) ให้พิจารณาประเด็นดังต่อไปนี้ทุกครั้งที่มีการใช้เรื่องดังกล่าว

1. ความเที่ยงตรง เป็นเรื่องที่อ้างอิงความเหมาะสมในการตีความหมายและการใช้ผลที่ได้จากกระบวนการประเมินจากแต่ละบุคคล ไม่ใช่เป็นความเที่ยงตรงของกระบวนการที่ดำเนินการโดยตรง

2. ความเที่ยงตรงเป็นการเสนอผลในลักษณะของระดับว่ามีมากหรือน้อย เช่น มีความเที่ยงตรงระดับมาก ระดับปานกลาง หรือระดับน้อย

3. ความเที่ยงตรงเป็นลักษณะเฉพาะประเด็นที่ต้องการใช้หรือตีความหมายเท่านั้น แต่จะไม่มีการประเมินใดที่มีความเที่ยงตรงที่ครบถ้วน สมบูรณ์ในทุกประเด็นหรือจุดมุ่งหมาย

4. ความเที่ยงตรงเป็นความคิดรวบยอดเชิงเดี่ยว เป็นค่าของตัวเลขที่ได้มาจากหลักฐาน

หลากหลายแหล่ง หลักการพื้นฐานที่ใช้พิจารณาตีความหมายของความเที่ยงตรง ได้แก่ จุดประสงค์ เนื้อหา เกณฑ์ หรือโครงการ เป็นต้น

5. ความเที่ยงตรงจะเกี่ยวข้องกับการประเมินเพื่อตัดสินซึ่งจะต้องมีเอกสารและหลักฐานที่สนับสนุนในการประเมินและการตัดสินดังกล่าว

ความเที่ยงตรง (Validity) หมายถึง ความถูกต้องของแบบทดสอบในสิ่งที่ต้องการวัดหรือความถูกต้องแม่นยำที่แบบทดสอบวัดได้ตามวัตถุประสงค์ที่วางไว้ ซึ่งเป็นความสามารถของแบบทดสอบที่จะสะท้อนความหมายที่แท้จริงของแนวคิดที่ต้องการศึกษาออกมาได้อย่างสมบูรณ์และถูกต้อง ความเที่ยงตรงจำแนกออกได้เป็น 4 ประเภทดังนี้

1. ความเที่ยงตรงเชิงเนื้อหา (Content Validity) หมายถึงระดับความสามารถของแบบทดสอบที่วัดในเนื้อหาที่ต้องการจะวัดหรืออาจพิจารณาว่าแบบทดสอบที่สร้างขึ้นได้สัดส่วนพอที่จะเป็นตัวแทนของบรรดาความรู้ทั้งมวล และสามารถบอกผลได้แน่นอนหรือไม่ว่าผู้เรียนประสบความสำเร็จเพียงใด ทั้งนี้กระบวนการทดสอบดังกล่าวนี้ต้องอาศัยผู้เชี่ยวชาญทางด้านเนื้อหาเพื่อตรวจสอบกระบวนการสร้างแบบทดสอบเพื่อตัดสินใจว่าข้อคำถามในแบบทดสอบว่าสามารถใช้เป็นตัวแทนของเนื้อหาที่จะถามได้หรือไม่ การทดสอบความเที่ยงตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญในลักษณะนี้เรียกว่าการหาค่าความสอดคล้องระหว่างวัตถุประสงค์กับแบบทดสอบหรือเรียกว่าการหาค่า IOC (Index of Item-objective Congruence) ซึ่งมีสูตรดังนี้

$$IOC=\frac{\Sigma R}{N}$$

เมื่อ IOC แทน ความสอดคล้องระหว่างวัตถุประสงค์กับแบบทดสอบ

ΣR แทน ผลรวมของคะแนนการพิจารณาจากผู้เชี่ยวชาญ

N แทน จำนวนผู้เชี่ยวชาญ

สำหรับตัวอย่างการหาค่า IOC ซึ่งเป็นการพิจารณาความสอดคล้องระหว่างข้อคำถามกับจุดประสงค์ของแบบทดสอบหรือจุดประสงค์เชิงพฤติกรรมหรือไม่นั้น ให้พิจารณาตัวอย่างต่อไปนี้

จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	การพิจารณา		
		+1	0	-1
นักเรียนสามารถบวกจำนวนไม่เกินสองหลักได้ถูกต้อง	28 + 74 มีค่าเท่ากับข้อใดต่อไปนี้ ก. 92			

	ข. 93 ค. 102 ง. 103			
ข้อเสนอแนะ.....				

จากตัวอย่างการหาค่า IOC ข้างต้น ผู้เชี่ยวชาญจำนวน 3-5 คน จะเป็นผู้พิจารณาว่าข้อสอบวัดได้สอดคล้องกับจุดประสงค์เชิงพฤติกรรมหรือไม่ โดยที่

+1 แทน วัดได้สอดคล้อง

0 แทน ไม่แน่ใจว่าวัดได้สอดคล้องหรือไม่

และ -1 แทน วัดได้ไม่สอดคล้อง

เมื่อนำผลการประเมินจากผู้เชี่ยวชาญทุกท่านมาคำนวณหาค่า IOC และพบว่ามีค่าตั้งแต่ 0.5 ขึ้นไป แสดงว่าข้อสอบข้อนั้นวัดได้สอดคล้องกับจุดประสงค์เชิงพฤติกรรม อย่างไรก็ตามควรนำข้อเสนอแนะจากผู้เชี่ยวชาญมาพิจารณาเพื่อปรับปรุงข้อสอบให้สมบูรณ์ยิ่งขึ้นต่อไป

2. ความเที่ยงตรงเชิงโครงสร้าง (Construct Validity) หมายถึงความสามารถของแบบทดสอบที่วัดได้ตามลักษณะคุณสมบัติ ทฤษฎี และประเด็นต่างๆของโครงสร้างนั้นโดยการให้นิยามของคุณลักษณะนั้น ๆ ในลักษณะที่เป็นนิยามปฏิบัติการ จากนั้นจึงสร้างข้อสอบตามนิยามปฏิบัติการที่ให้นั้น วิธีการหาความเที่ยงตรงเชิงโครงสร้างสามารถใช้วิธีที่เรียกว่า Known group technique คือใช้กลุ่มตัวอย่างที่รู้ชัดแล้วว่า มีความสามารถมากน้อยเพียงใดเป็นหลักสำหรับทำการเปรียบเทียบ หรือใช้วิธี Pretest-Posttest technique คือวิธีทดสอบก่อนและหลัง โดยยึดหลักว่าระหว่างการทดสอบก่อนและหลัง กลุ่มตัวอย่างได้พัฒนาทางสมองเพิ่มขึ้น และมีประสบการณ์มากขึ้น ดังนั้นคะแนนจากการสอบครั้งที่ 2 จึงควรจะมากกว่าครั้งที่ 1 นอกจากนี้ การหาค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) ทำได้โดยการหาค่าสัมประสิทธ์ความสัมพันธ์ของคะแนนของแบบทดสอบ 2 ชุดที่วัดในเรื่องเดียวกันถ้าค่าสัมประสิทธิ์ความสัมพันธ์มีค่าสูงและมีทิศทางเดียวกัน แสดงว่าแบบทดสอบที่สร้างขึ้นมีความเที่ยงตรงตามโครงสร้างสูง

3. ความเที่ยงตรงตามสภาพ (Concurrent Validity) หมายถึงแบบทดสอบที่สามารถวัดได้ตามสภาพความเป็นจริงของกลุ่มตัวอย่างโดยนำคะแนนจากแบบทดสอบนั้นไปเปรียบเทียบกับอันดับความสามารถของผู้เรียนตามสภาพจริงที่ผู้สอนสังเกตเห็นในปัจจุบันเป็นตัวเกณฑ์

4. ความเที่ยงตรงเชิงพยากรณ์ (Predictive Validity) หมายถึงการหาความสัมพันธ์หรือสหสัมพันธ์ระหว่างคะแนนผลการสอบกับเกณฑ์ของความสำเร็จที่จะเกิดขึ้นในอนาคตโดยใช้คะแนนผลการสอบในการพยากรณ์ในอนาคตถ้าหากแบบทดสอบมีความเที่ยงตรงเชิงพยากรณ์สูง และบุคคลผู้ใดทำคะแนนได้ดีจะสามารถพยากรณ์ได้ว่าบุคคลนั้นย่อมมีความสำเร็จในสาขาวิชาที่เกี่ยวข้องในประเด็นของแบบทดสอบ

การหาค่าความเชื่อมั่นของแบบทดสอบ

การหาค่าความเชื่อมั่นของแบบทดสอบจะใช้ค่าสถิติเป็นตัวชี้คุณภาพของแบบทดสอบว่าให้ผลการวัดได้แน่นอนเพียงใด วิธีการที่นิยมใช้หาค่าความเชื่อมั่นของแบบทดสอบมีดังนี้

1. การสอบซ้ำ (Test –retest method) เป็นการหาค่าความเชื่อมั่นของแบบทดสอบ โดยการนำแบบทดสอบไปสอบกับผู้สอบกลุ่มเดียวกัน 2 ครั้งในสถานการณ์ที่เหมือน ๆ กัน เพื่อตรวจสอบดูว่าคะแนนหรือผลการวัดทั้งสองนั้นให้ผลสอดคล้องสัมพันธ์กันเพียงใด ถ้าค่า r_{XY} ที่คำนวณได้มีค่าเป็นบวกสูง ๆ แสดงว่าแบบทดสอบนั้นมีค่าความเชื่อมั่นสูง

วิธีการในการหาค่าความเชื่อมั่นแบบสอบซ้ำ มีดังนี้

1.1 นำแบบทดสอบฉบับหนึ่งไปสอบกับผู้สอบกลุ่มเดียวกัน 2 ครั้ง โดยทิ้งระยะให้ห่างกันพอสมควร เพื่อป้องกันผู้สอบจำวิธีการตอบแบบทดสอบครั้งแรกได้

1.2 นำคะแนนที่ได้จากการสอบทั้ง 2 ครั้ง มาหาค่าสหสัมพันธ์แบบเปียร์สันซึ่งบางครั้งเรียกว่าสัมประสิทธิ์ความคงที่ โดยใช้สูตร $r_{XY} = \frac{N\sum XY - \sum X \sum Y}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}}$ ค่าของสหสัมพันธ์จะชี้ให้เห็นถึงความคงที่ของคะแนน ซึ่งก็คือค่าความเชื่อมั่นของแบบทดสอบฉบับนั้น

2. การใช้แบบทดสอบคู่ขนาน (Parallel forms method) วิธีนี้จะใช้แบบทดสอบคู่ขนาน 2 ฉบับไปสอบกับผู้สอบกลุ่มเดียว จะได้คะแนนมา 2 ชุด แล้วนำมาหาค่าสหสัมพันธ์ลักษณะของแบบทดสอบคู่ขนานนั้นจะต้องเป็นแบบทดสอบที่มีค่าเฉลี่ยและความแปรปรวนเท่ากัน วัดเนื้อหาเดียวกัน มีค่าสหสัมพันธ์ภายในข้อสอบเท่ากัน ซึ่งสร้างได้ยาก จึงทำให้การหาค่าความเชื่อมั่นวิธีนี้ไม่นิยมใช้

3. การวัดความคงที่ภายใน (Measure of internal consistency) วิธีนี้จะใช้การสอบเพียงครั้งเดียว แล้วนำผลที่ได้ไปคำนวณค่าความเชื่อมั่นของแบบทดสอบ ซึ่งมีหลายวิธี เช่น

3.1 การแบ่งครึ่งแบบทดสอบ (Split–half method) โดยการนำ แบบทดสอบไปสอบกับผู้สอบครั้งเดียวแล้วแบ่งการตรวจของคะแนนเป็น 2 ส่วนคือข้อคู่ข้อคี่ จะได้คะแนนออกเป็น 2 ชุด นำคะแนนรวมของข้อคู่และข้อคี่มาหาค่าสหสัมพันธ์ของแบบทดสอบครึ่งฉบับ (r_{XY}) ต่อจากนั้นจึงนำค่าสหสัมพันธ์ครึ่งฉบับไปขยายให้เต็มฉบับโดยใช้สูตรของสเปียร์แมนบราวน์ (Spearman –Brown) ดังนี้

$$r_{tt} = \frac{2r_{hh}}{1+r_{hh}}$$

เมื่อ r_{tt} แทน ความเชื่อมั่นของแบบทดสอบทั้งฉบับ

r_{hh} แทน ความเชื่อมั่นของแบบทดสอบครึ่งฉบับ

การหาความเชื่อมั่นโดยการแบ่งครึ่งแบบทดสอบนี้มีข้อตกลงว่าแบบทดสอบย่อย 2 ฉบับต้องเท่ากัน ถ้าไม่เท่ากันแล้วค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับจะมีค่าสูงกว่าวิธีการอื่น ๆ

3.2 วิธีของคูเดอร์–ริชาร์ดสัน (Kuder – RichardsonMethod) ใช้หาความเชื่อมั่นของแบบทดสอบที่มีระบบการให้คะแนนแบบตอบผิดได้ 0 และตอบถูกได้ 1 โดยมีข้อตกลงว่า เนื้อหาของข้อสอบต้องเป็นเอกพันธ์กันมี 2 สูตรคือ

3.2.1 สูตร KR-20 ในกรณีที่ค่าความยากง่ายของข้อสอบแต่ละข้อไม่เท่ากัน

$$r_{tt} = \frac{n}{n-1}\left[1-\frac{\Sigma pq}{S^2}\right]$$

เมื่อ r_{tt} แทน ความเชื่อมั่นของแบบทดสอบ

nแทน จำนวนข้อสอบ

p แทน ความยากง่ายของข้อสอบแต่ละข้อ (สัดส่วนที่ตอบถูก)

q แทน สัดส่วนที่ตอบผิด ($1-p$)

S^2 แทน ความแปรปรวนของคะแนนรวมของแบบทดสอบ

3.2.2 สูตร KR-20ในกรณีที่ค่าความยากง่ายของข้อสอบทุกข้อเท่ากันหรือไม่แตกต่างกันมาก

$$r_{tt} = \frac{n}{n-1}\left[1-\frac{\bar{X}(n-\bar{X})}{nS^2}\right]$$

เมื่อ r_{tt} แทน ความเชื่อมั่นของแบบทดสอบ

n แทน จำนวนข้อสอบ

$\bar{X}$ แทน ค่าเฉลี่ยของคะแนนรวมของแบบทดสอบทั้งฉบับ

S^2 แทน ความแปรปรวนของคะแนนรวมของแบบทดสอบ

3.3 วิธีสัมประสิทธิ์แอลฟ่า (Coefficient Alpha : α) เป็นวิธีที่ครอบบัค (Cronbach) สร้างขึ้นมาเพื่อใช้คำนวณหาความเชื่อมั่นของแบบวัดที่แต่ละข้อมีคะแนนเต็มเท่าไรก็ได้ เช่น แบบวัด เจตคติของลิเคอร์ท (Likert) ที่มีการให้คะแนนแบบ 5, 4,3, 2 และ 1 สำหรับสัมประสิทธิ์แอลฟ่ามีสูตรดังนี้

$$\alpha=\frac{n}{n-1}\left[1-\frac{\Sigma s_i^2}{s^2}\right]$$

เมื่อ α แทน ความเชื่อมั่นของแบบทดสอบ

n แทน จำนวนข้อสอบทั้งหมด

s_i^2 แทน ความแปรปรวนของข้อสอบแต่ละข้อ

s^2 แทน ความแปรปรวนของคะแนนรวมของแบบทดสอบ

สำหรับความสัมพันธ์ระหว่างความเที่ยงตรงและความเชื่อมั่นของแบบทดสอบ ลินน์และมิลเลอร์ (Linn & Miller, 2005: 70) ได้สรุปดังภาพที่ 3.1 ดังนี้

ภาพที่ 3.1 การแสดงความเกี่ยวข้องระหว่างความเที่ยงตรงกับความเชื่อมั่น

ที่มา: https://explorable.com/validity-and-reliability

จากภาพที่ 3.1 เริ่มจากด้านซ้ายมือพบว่าการวัดมีความเชื่อมั่นแต่ไม่มีความเที่ยงตรง (Reliable not Valid) กล่าวคือมีการวัดที่มีความสอดคล้องแต่ไม่ตรงเป้าหมาย ในขณะภาพถัดมา

พบว่ามีความเชื่อมั่นและความเที่ยงตรงต่ำ (Low validity low reliability) กล่าวคือมีการวัดที่ไม่สอดคล้องกันกระจัดกระจายไปทั่ว ซึ่งมีโอกาสต่ำที่จะวัดได้ตรงเป้า ภาพถัดไปพบว่าการวัดจะกระจายออกไปในส่วนหนึ่งส่วนใดของเป้าและไม่ถูกจุดศูนย์กลางของเป้า (Not reliable not valid) การวัดในกรณีนี้ไม่มีทั้งความเชื่อมั่นและความเที่ยงตรง ส่วนภาพสุดท้ายทางขวาสุด พบว่าการวัดมีทั้งความเชื่อมั่นและความเที่ยงตรงเป็นการวัดที่ต้องการให้เกิดกับผู้เรียน

สรุปท้ายบท

เครื่องมือหรือเทคนิคที่ใช้ในการวัดผลทางการศึกษาจะวัดเกี่ยวกับพฤติกรรมด้านพุทธิพิสัย จิตพิสัยและทักษะพิสัยของผู้เรียน ผู้สอนควรเลือกใช้เครื่องมือให้เหมาะสมกับพฤติกรรมที่ต้องการวัด แบบทดสอบเป็นเครื่องมือวัดพฤติกรรมด้านพุทธิพิสัยและมีบทบาทสำคัญมากและมีการใช้อย่างกว้างขวางในการวัดความรู้ของผู้เรียน วิธีการสร้างข้อสอบเพื่อวัดพฤติกรรมด้านพุทธิพิสัยควรมีการกำหนดพฤติกรรมย่อยแต่ละด้านเช่น ความรู้ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ การประเมินค่า รวมถึงข้อสอบที่วัดด้านการใช้เหตุผล หรือการแก้ปัญหา ก่อนลงมือสร้างและต้องมีการฝึกฝนบ่อย ๆ ที่สำคัญการเขียนข้อสอบให้มีคุณภาพ ควรศึกษาหลักเกณฑ์การสร้างข้อสอบทั้งนี้เพราะคุณภาพของแบบทดสอบหรือลักษณะของแบบทดสอบที่ดีจะมีความเกี่ยวข้องกับความเที่ยงตรงความเชื่อมั่นความยากง่ายอำนาจจำแนกและความเป็นปรนัยโดยความเที่ยงตรงเป็นลักษณะที่สำคัญที่สุดของแบบทดสอบ นอกจากนี้สิ่งสำคัญอีกอย่างหนึ่งในการสร้างข้อสอบ ได้แก่ การวิเคราะห์ข้อสอบ ซึ่งเป็นการตรวจสอบคุณภาพของข้อสอบ นับว่าเป็นเทคนิคอย่างหนึ่งในการปรับปรุงและพัฒนาข้อสอบ ทั้งเป็นรายข้อและทั้งฉบับให้มีคุณค่าต่อการนำไปใช้วัดและประเมินผู้เรียนต่อไป

เอกสารอ้างอิง

บุญชม ศรีสะอาด (ม.ป.ป.) การวิเคราะห์ข้อสอบแบบอื่นนอกเหนือจากแบบเลือกตอบ. สืบค้นเมื่อวันที่ 1 มีนาคม 2556 จาก www.watpon.com/boonchom/04.doc

ผดุงชัย ภู่พัฒน์. (ม.ป.ป.). การพัฒนาเครื่องมือวัดผลสัมฤทธิ์ทางการเรียน. สืบค้นเมื่อวันที่ 2 มีนาคม 2556 จาก http://tlc.buu.ac.th/personal/example.pdf

พิสณุ ฟองศรี. (2551). การประเมินทางการศึกษา: แนวคิดสู่การปฏิบัติ. พิมพ์ครั้งที่ 5. กรุงเทพฯ: บริษัทด่านสุทธาการพิมพ์ จำกัด. หน้า 1-359.

เรวดี กระโหมวงศ์ (ม.ป.ป.). การสร้างและใช้เครื่องมือวัดและประเมินผลตามจุดประสงค์การเรียนรู้. สืบค้นเมื่อวันที่ 4 เมษายน 2556 จาก http://pt.tsu.ac.th/aqad/main/report/files/20082014058.doc

สมนึก ภัททิยธนี. (2549). การวัดผลการศึกษา. พิมพ์ครั้งที่ 5 กาฬสินธุ์: โรงพิมพ์ประสานการพิมพ์.

สุทธิพงศ์ บุญผดุง. (2558). การพัฒนาแบบวัดการคิดเชิงวิเคราะห์วิชาคณิตศาสตร์ตามมาตรฐานระดับชาติและระดับนานาชาติ. มหาวิทยาลัยราชภัฏสวนสุนันทา.

Asim, A. E., Ekuri, E. E., & Eni, E. I. (2013).A Diagnostic Study of Pre-service teachers' Competency in Multiple-choice item Development. *Research in Education, 89*, 13-14.

Drits-Esser, D., and Black, M. (2014).Standardized Test Questions: A Tool for Developing Students' Proficiency with the Framework's Science Practices. *Science Scope, 37(5)*, 68-75.

Linn, R. L., & Miller, M. D. (2005). Measurement and Assessment in Teaching. 9th edition, *Pearson Prentice Hall*.

Shuttleworth, M. (2008). *Validity and Reliability*. Retrieved Jan 17, 2014 from Explorable.com: https://explorable.com/validity-and-reliability

Tileston, D. W. (2004). What Every Teacher Should Know About Student Assessment. *Corwin Press*, pp. 1-113.

เอกสาร

เกณฑ์การให้คะแนน

การประเมินผลงานหรือประเมินการปฏิบัติงานของผู้เรียนตามแนวการเรียนรู้ที่เน้นให้ผู้เรียนสร้างสรรค์งานนั้นไม่สามารถประเมินได้ด้วยการให้คะแนนแบบตอบถูกได้ 1 คะแนน และตอบผิดได้ 0 คะแนนได้เพราะผลงานหรือการปฏิบัติไม่ใช่ประเภทมีหรือไม่มีได้อย่างแท้จริงหากมีระดับของการปฏิบัติว่าดีมากดีพอใช้หรือยังต้องปรับปรุงโดยพิจารณาตามเงื่อนไขหรือตัวชี้วัดงานนั้นตามระดับคุณภาพของงานรวมถึงการให้ข้อมูลป้อนกลับแก่ผู้เรียนเพื่อให้ทราบถึงความคืบหน้าและการพัฒนาของตนซึ่งเกณฑ์การให้คะแนน (rubric) เป็นเครื่องมือในการประเมินชนิดหนึ่งที่มีความเกี่ยวข้องและเหมาะสมกับการประเมินผลงานหรือการปฏิบัติงานของผู้เรียน ดังนั้นการได้ทราบและเรียนรู้เกี่ยวกับเกณฑ์การให้คะแนนจึงมีความจำเป็นกับผู้เรียนในยุคศตวรรษที่ 21 นี้ ที่เน้นการประเมินศักยภาพของผู้เรียนไม่ว่าจะเป็นการประเมินการนำเสนอปากเปล่า การอภิปราย การสาธิต การประเมินแฟ้มสะสมผลงาน เป็นต้น ซึ่งแตกต่างจากการประเมินโดยการทดสอบไม่ว่าจะเป็นแบบเลือกตอบ แบบถูกผิด แบบจับคู่และแบบเติมคำตอบสั้น ๆ ที่มีคำตอบที่ตายตัว เพื่อให้ผู้เรียนเกิดความรู้ ความเข้าใจเกี่ยวกับเกณฑ์การให้คะแนน บทนี้จะกล่าวถึงแนวคิดเกี่ยวกับเกณฑ์การให้คะแนน ประเภทของเกณฑ์การให้คะแนน ขั้นตอนการออกแบบเกณฑ์การให้คะแนน และตัวอย่างเกณฑ์การให้คะแนนเพื่อนำไปใช้ในการจัดการศึกษาให้กับประเทศชาติได้อย่างมีคุณภาพต่อไป

แนวคิดเกี่ยวกับเกณฑ์การให้คะแนน

1. ความหมายของเกณฑ์การให้คะแนน

คำว่า เกณฑ์การให้คะแนน หรือ Scoring Rubric ได้ถูกนำมาใช้อย่างกว้างขวางในวงการศึกษาเป็นเครื่องมือที่มีประโยชน์ต่อการนำมาใช้เพื่อประเมินการเรียนรู้ภาคปฏิบัติ สามารถใช้ได้กับรายวิชาและสถานการณ์ที่ต้องการให้ผู้เรียนแสดงความรู้ ความสามารถและทักษะในการปฏิบัติงาน เกณฑ์การให้คะแนนจะระบุคำอธิบายเชิงคุณภาพของเกณฑ์การปฏิบัติเพื่อให้ผู้เรียนปฏิบัติงานให้ได้ผลดีที่สุดทั้งที่เป็นการประเมินแบบย่อยและแบบรวม

ส.วาสนา ประวาลพฤกษ์ (2550: 4) กล่าวว่าเกณฑ์การให้คะแนนคือแนวการให้คะแนนเพื่อประเมินงานของผู้เรียนเกณฑ์การให้คะแนนจะตอบคำถามเหล่านี้คือ

- ใช้เกณฑ์ (criteria) ใดในการตัดสินผลงานหรือการปฏิบัตินั้น

- ผู้ประเมินจะพิจารณาผลสำเร็จของการปฏิบัตินั้นตรงไหนและอย่างไร

- คุณภาพของงานมีระดับความแตกต่างกันเพียงใด

- ในการกำหนดระดับคะแนนให้กับงานนั้นๆผู้ประเมินรับรองความสมเหตุสมผลความมั่นใจและความยุติธรรมในการประเมินได้เพียงใดและระดับคะแนนนั้นหมายความว่าอย่างไร

- การได้ระดับคะแนนที่แตกต่างกันนั้นสามารถอธิบายและแยกแยะความแตกต่างระหว่างงานนั้นได้อย่างไร

กู๊ดริช (Goodrich, 1996: 14) กล่าวว่าเกณฑ์การให้คะแนนเป็นเครื่องมือที่ระบุรายการเกณฑ์เพื่อพิจารณาผลงาน โดยจะเรียงลำดับคุณภาพของผลงานจากระดับดีเยี่ยมจนถึงระดับปรับปรุงแก้ไข

กรอนลันด์ (Gronlund, 1998 Quoted in Kan, 2007: 145) ได้ให้ความหมายเกณฑ์การให้คะแนนว่าเป็นแนวทางการให้คะแนนที่บรรยายลักษณะของระดับการปฏิบัติงานที่แตกต่างกันตามความรู้ ความสามารถในการปฏิบัติงานนั้น ๆ โดยมีองค์ประกอบ 3 ส่วนได้แก่ เกณฑ์การปฏิบัติงาน ระดับการปฏิบัติงาน และคำอธิบายความหมายการปฏิบัติงานในแต่ละระดับ ซึ่งจะระบุและพรรณนาเชิงคุณภาพที่ดีที่สุดที่ต้องการให้ผู้เรียนปฏิบัติ

มูนและคณะ (Moon et al., 2005: 123) กล่าวว่าเกณฑ์การให้คะแนนถูกออกแบบมาเพื่อให้ข้อมูล สารสนเทศเกี่ยวกับจุดแข็ง และจุดที่ควรปรับปรุงในเนื้อหาการเรียนและกระบวนการการเรียนรู้ที่จะใช้ในการประเมินผู้เรียน ซึ่งการพัฒนาเกณฑ์การให้คะแนนจะเริ่มจากการทบทวนจุดประสงค์การประเมิน องค์ประกอบของการปฏิบัติที่จะนำมาใช้ในการประเมิน เกณฑ์ในการประเมิน การกำหนดคะแนนในแต่ละเกณฑ์หลักและย่อยพร้อมทั้งการแปลผลเป็นระดับของผู้ปฏิบัติที่เริ่มตั้งแต่ระดับพื้นฐานจนถึงระดับเชี่ยวชาญ นอกจากนี้ผลการวิจัยของมูนและคณะ (2005: 128) ได้ค้นพบว่าผู้เรียนสามารถนำเกณฑ์การให้คะแนนที่ผู้สอนสร้างขึ้นไปเป็นแนวทางในการวางแผนการทำผลงานตั้งแต่เริ่มแรก ระหว่างการทำผลงาน และตอนสุดท้ายของการทำผลงาน เพื่อตรวจสอบความถูกต้องและความครบถ้วนสมบูรณ์

ชโมกเกอร์ (Schmoker, 2006: 70-71 Quoted in Cooper & Gargan, 2009: 54) กล่าวว่าเกณฑ์การให้คะแนนหมายถึงกฎ หรือแนวทาง ซึ่งใช้ในการตัดสินผลงานของผู้เรียนซึ่งจะต้องมีเกณฑ์ที่มีความชัดเจนเพื่อใช้ในการประเมินผลงานนั้นๆ

กัสกีย์ (Guskey, 1994: 25 Quoted in Cooper & Gargan, 2009: 54)กล่าวว่าเกณฑ์การให้คะแนนเป็นแนวทางที่เฉพาะเจาะจงที่สามารถใช้ในการอธิบายผลงานของผู้เรียนในการอ่าน การเขียน คณิตศาสตร์ และรายวิชาอื่น ๆ

โมสเกล (Moskel, 2000 อ้างถึงใน กิ่งกาญจน์ สิรสุคนธ์, 2550: 1)กล่าวว่า เกณฑ์การให้คะแนนคือแนวทางการให้คะแนนอย่างละเอียด ซึ่งพัฒนาขึ้นโดยผู้สอนหรือผู้ประเมิน เพื่อเป็นแนวทางในการวิเคราะห์ผลงานหรือกระบวนการที่เกิดจากความพยายามของนักเรียน

เมอร์ทเลอร์ (Mertler, 2001 อ้างถึงใน กิ่งกาญจน์ สิรสุคนธ์, 2550: 1) กล่าวว่าเกณฑ์การให้คะแนนเป็นมาตราส่วนประมาณค่า (Rating Scales)ที่ใช้ประเมินการปฏิบัติ ซึ่งตรงกันข้ามกับแบบสำรวจรายการ (Checklists)โดยปกติจะเรียกว่าแนวทางการให้คะแนน (Scoring guides) ประกอบด้วยเกณฑ์การประเมินการปฏิบัติที่มีลักษณะเฉพาะ ใช้ในการประเมินการปฏิบัติงานของนักเรียน หรือประเมินผลผลิตซึ่งเป็นผลจากการปฏิบัติงาน

โดยสรุปเกณฑ์การให้คะแนนเป็นเครื่องมือให้คะแนนชนิดหนึ่งที่นำมาใช้ในการประเมินการปฏิบัติงาน ชิ้นงาน หรือผลงานของนักเรียน เกณฑ์การให้คะแนนประกอบด้วย 2 ส่วน คือเกณฑ์ที่ใช้ประเมินการปฏิบัติงาน ชิ้นงานหรือผลงานของนักเรียน และระดับคุณภาพหรือระดับคะแนน โดยที่เกณฑ์จะบอกผู้สอนหรือผู้ประเมินว่าการปฏิบัติงานหรือผลงานนั้น ๆจะต้องพิจารณาสิ่งใดบ้าง ส่วนระดับคุณภาพหรือระดับคะแนนจะบอกว่า การปฏิบัติงาน ชิ้นงานหรือผลงานควรได้ระดับคุณภาพหรือระดับคะแนนในระดับใดตามเกณฑ์ที่กำหนดแต่ละตัว

2. ทำไมจึงต้องใช้เกณฑ์การให้คะแนน

การใช้เกณฑ์การให้คะแนนมีประโยชน์สำหรับผู้สอนและผู้เรียนหลายประการดังนี้ (Goodrich, 1996: 14)

1. เกณฑ์การให้คะแนนเป็นเครื่องมือที่มีประสิทธิภาพมากทั้งในการเรียนการสอนและการประเมิน ช่วยปรับปรุงพัฒนาการปฏิบัติงานของผู้เรียน โดยเปรียบเสมือนเป็นตัวช่วยในการควบคุมการปฏิบัติงานของผู้เรียน โดยครูต้องกำหนดความต้องการหรือความคาดหวังในผลงานของผู้เรียนอย่างชัดเจนและแสดงให้ผู้เรียนทราบว่าจะทำให้ถึงความคาดหวังนั้นได้อย่างไร ซึ่งปรากฏว่าคุณภาพผลงานและการเรียนรู้ของนักเรียนพัฒนาขึ้นอย่างเห็นได้ชัด อย่างไรก็ตามประเด็นที่กล่าวถึงเกี่ยวกับการใช้เกณฑ์การให้คะแนน ได้แก่ การให้คำอธิบายที่ชัดเจนในเรื่องคุณภาพ และเมื่อผู้เรียนมีข้อบกพร่องตามเกณฑ์ใด ผู้สอนก็จะช่วยชี้แนะและบอกได้ว่า ผู้สอนคาดหวังให้ผู้เรียนทำอะไร

2. เกณฑ์การให้คะแนนช่วยให้ผู้เรียนตัดสินคุณภาพผลงานของตนเองและของคนอื่น ๆอย่างมีเหตุผลเมื่อใช้เกณฑ์การให้คะแนนเป็นแนวทางการประเมิน ผู้เรียนจะสามารถชี้แนะและแก้ปัญหาเกี่ยวกับผลงานของตนและผู้อื่นได้อย่างตรงจุด การฝึกซ้ำ ๆเกี่ยวกับการประเมินผลงานกลุ่มและผลงานของตนเองทำให้นักเรียนเพิ่มความรับผิดชอบเกี่ยวกับผลงานของตนมากขึ้น และจะไม่มีการถามตนเองว่า “ฉันทำงานเสร็จหรือยัง” (Am I done yet?)

3. เกณฑ์การให้คะแนนช่วยลดเวลาของผู้สอนในการประเมินงานของผู้เรียน ทั้งนี้เพราะผลงานที่ผ่านการประเมินตัวผู้เรียนและโดยเพื่อน ๆ ซึ่งยึดเกณฑ์การให้คะแนนเป็นหลักนั้น จะทำให้ข้อบกพร่องของผลงานมีน้อย หรือน้อยมากเมื่อมาถึงมือผู้สอน หากมีสิ่งใดที่ต้องปรับปรุง หรือแจ้งให้ผู้เรียนทราบผู้สอนก็เพียงแต่วงล้อมรอบประเด็นนั้นในเกณฑ์การให้คะแนนแทนที่จะต้องใช้เวลาในการอธิบายกันยืดยาวนอกจากนี้เกณฑ์การให้คะแนนยังช่วยให้ข้อมูลย้อนกลับแก่นักเรียนมากขึ้นเกี่ยวกับจุดเด่นและสิ่งที่ต้องปรับปรุง

4. เกณฑ์การให้คะแนนมีความยืดหยุ่น กล่าวคือ มีระดับคุณภาพตั้งแต่ดีเยี่ยมจนถึงต้องปรับปรุง ทำให้ผู้สอนนำไปใช้กับผู้เรียนที่คละความสามารถได้ โดยนำไปใช้กับผู้เรียนที่เก่งจนถึงนักเรียนที่อ่อนโดยใช้เกณฑ์ในการสะท้อนผลงานของผู้เรียนเหล่านี้

5. เกณฑ์การให้คะแนนใช้ง่ายและอธิบายได้ง่าย กล่าวคือ ผู้เรียนจะรู้ชัดเจนว่าตนเรียนรู้อะไรบ้าง ซึ่งจะสอดคล้องกับผลการประเมินที่ตนได้รับ นอกจากนี้ผู้ปกครองยังให้การสนับสนุนให้มีการใช้เกณฑ์การให้คะแนนเพราะรู้ชัดเจนว่าบุตรหลานจะต้องทำอย่างไรเพื่อให้ประสบความสำเร็จในการเรียน

พานาเดโร และ โจนส์สัน (Panadero & Jonsson, 2013: 138-139) ได้สรุปข้อดีของการใช้เกณฑ์การให้คะแนนที่ช่วยปรับปรุงการปฏิบัติงานของผู้เรียน ดังต่อไปนี้

1. เกณฑ์การให้คะแนนเพิ่มความโปร่งใส

การให้ผู้เรียนรับรู้ถึงสิ่งที่ผู้สอนคาดหวังเพื่อให้ผู้เรียนปฏิบัติงานหรือชิ้นงานผ่านการสื่อสารด้วยการใช้เกณฑ์การให้คะแนนจะทำให้ผู้เรียนเข้าใจสิ่งที่ผู้สอนคาดหวัง นอกจากนี้ความชัดเจนในการอธิบายแต่ละเกณฑ์เพื่อการประเมินผลงานทำให้ผู้เรียนทราบถึงความต้องการและลักษณะของผลงานที่เป็นไปตามเกณฑ์หรือมาตรฐานที่ผู้สอนกำหนดซึ่งมีส่วนผลักดันให้ผู้เรียนปฏิบัติงานหรือชิ้นงานตามเกณฑ์ในการประเมินต่อไป

2. เกณฑ์การให้คะแนนช่วยลดความวิตกกังวลในการส่งงานของผู้เรียน

เกณฑ์การให้คะแนนที่กำหนดความคาดหวังที่ชัดเจนจะช่วยส่งเสริมการปฏิบัติงานของผู้เรียน และทำให้ผู้เรียนลดความวิตกกังวลที่เกี่ยวกับงานหรือชิ้นงานที่ต้องทำส่งผู้สอน จากงานวิจัยของ แอนเดรดและดู (Andrade & Du, 2005 Quoted in Panadero & Jonsson, 2013: 138) พบว่าผู้เรียนมีความมั่นใจและส่งงานหรือชิ้นงานที่ได้รับมอบหมายอย่างง่ายดายเมื่อมีเกณฑ์การให้คะแนน

3. เกณฑ์การให้คะแนนช่วยให้เกิดกระบวนการให้ข้อมูลป้อนกลับ

ในการประเมินโดยใช้เกณฑ์การให้คะแนน สิ่งที่จะเกิดขึ้นได้แก่กระบวนการให้ข้อมูลป้อนกลับซึ่งเป็นประโยชน์ทั้งผู้สอนและผู้เรียน โดยเฉพาะอย่างยิ่งการให้ข้อมูลป้อนกลับในผลงานที่เป็นโครงร่างของผู้เรียนจะเป็นประโยชน์มาก

4. เกณฑ์การให้คะแนนช่วยปรับปรุงพลังสมรรถนะแห่งตนของผู้เรียน

การประเมินโดยใช้เกณฑ์การให้คะแนน ผู้สอนจะให้ข้อมูลป้อนกลับแก่ผู้เรียนที่เกี่ยวข้องกับการปฏิบัติงานหรือชิ้นงานว่ามีจุดแข็งและจุดที่ควรปรับปรุงในเรื่องใดบ้าง ซึ่งมีผลต่อการปรับปรุงพลังสมรรถนะแห่งตนของผู้เรียน นอกจากนี้เกณฑ์ที่ใช้ในการประเมินจากเกณฑ์การให้คะแนนจะช่วยยกระดับการปฏิบัติงานหรือชิ้นงานของผู้เรียนได้เป็นอย่างดีซึ่งสอดคล้องกับงานวิจัยของปาจาเรส (Pajares, 2008 Quoted in Panadero & Jonsson, 2013: 138) ที่พบว่าผู้เรียนที่มีความสามารถในการปฏิบัติงานหรือชิ้นงานสูงจะมีแนวโน้มเป็นผู้ที่มีพลังสมรรถนะแห่งตนสูง

5. เกณฑ์การให้คะแนนส่งเสริมให้มีการกำกับตนเองของผู้เรียน

เกณฑ์การให้คะแนนเป็นสิ่งที่ช่วยให้ผู้เรียนสามารถวางแผนและประเมินตนเองได้ในการเรียน กล่าวคือผู้เรียนสามารถวางแผนการปฏิบัติงานหรือชิ้นงานที่ได้รับมอบหมายจากการอ่านเกณฑ์การให้คะแนน ที่สำคัญเกณฑ์การให้คะแนนยังนำมาใช้ในการตรวจสอบว่างานที่ปฏิบัติมีความก้าวหน้าอยู่ในระดับใด(เป็นการประเมินตนเอง) ตั้งแต่เริ่มต้น ระหว่างทาง และปลายทางของการปฏิบัติงานหรือชิ้นงานว่าครบถ้วน ถูกต้องและสมบูรณ์หรือไม่ สามารถตรวจผลงานของตนว่าเป็นไปตามเกณฑ์ก่อนส่งผู้สอนได้

จะสร้างเกณฑ์การให้คะแนนได้อย่างไร

เกณฑ์การให้คะแนนเป็นเครื่องมือที่ได้รับความนิยมอย่างมากในการประเมินตามสภาพจริงและการประเมินภาคปฏิบัติ ซึ่งกระบวนการพัฒนาเกณฑ์การให้คะแนนพอสรุปได้ดังนี้ (Goodrich, 1996: 14; Andrade, 1997 Quoted in Montgomery, 2000: 326)

1. พิจารณารูปแบบโดยให้ผู้เรียนดูตัวอย่างงานที่ดีและงานที่ไม่ดีแล้วให้บอกลักษณะที่ทำให้งานดีหรือไม่ดี

2. ทำรายการเกณฑ์โดยให้อภิปรายตัวอย่างงานเพื่อเริ่มทำรายการเกณฑ์ที่จะใช้พิจารณาคุณภาพของงาน

3. จัดและทำระดับคุณภาพให้ชัดเจนอธิบายลักษณะคุณภาพดีที่สุดและแย่ที่สุด แล้วจึงอธิบายคุณภาพของงานระดับกลางโดยใช้พื้นฐานความรู้เกี่ยวกับปัญหาทั่วไปและการอภิปรายเกี่ยวกับงานที่คุณภาพไม่ค่อยดี

4. ฝึกให้ผู้เรียนใช้เกณฑ์การให้คะแนนที่สร้างขึ้น โดยฝึกให้ประเมินงานตัวอย่างที่ให้พิจารณาในขั้นที่ 1

5. กำหนดงานให้ผู้เรียนทำและให้ประเมินโดยตนเองและโดยกลุ่มในขณะที่ผู้เรียนทำงานก็ให้มีการประเมินตนเองและประเมินโดยกลุ่มเป็นระยะ ๆ

6. ทบทวนโดยให้เวลาผู้เรียนทบทวนงาน ตามที่ได้รับเป็นข้อมูลป้อนกลับจากขั้นตอนที่ 5

7. ให้ผู้สอนประเมินงานของผู้เรียนโดยใช้เกณฑ์การให้คะแนนเดียวกับที่ผู้เรียนใช้

ขั้นตอนที่ 1 อาจจำเป็นต้องทำ หากงานที่ให้ผู้เรียนปฏิบัตินั้นเป็นงานที่ผู้เรียนไม่คุ้นเคย ในขณะขั้นตอนที่ 3 และ 4 มีประโยชน์แต่ต้องใช้เวลาซึ่งผู้สอนสามารถทำได้เอง โดยเฉพาะเมื่อมีการใช้เกณฑ์การให้คะแนนนั้นสักช่วงเวลาหนึ่งการให้ผู้เรียนมีประสบการณ์การประเมินโดยใช้เกณฑ์การให้คะแนนอาจเริ่มตั้งแต่การที่ผู้สอนกำหนดรายการเกณฑ์ กำหนดระดับคุณภาพ ตรวจสอบกับผู้เรียนทบทวนและใช้เกณฑ์การให้คะแนนนั้นประเมินตนเอง ประเมินกับกลุ่ม และผู้สอนประเมิน

การสร้างเกณฑ์การให้คะแนนเป็นเรื่องยาก แต่การใช้กลายเป็นเรื่องที่ง่ายกว่าเมื่อผู้สอนสร้างเกณฑ์การให้คะแนนขึ้นมาแล้ว ควรแจกให้ผู้เรียนเพื่อทดลองใช้ โดยให้ผู้เรียนประเมินความก้าวหน้าของตนเองในการปฏิบัติงาน ชิ้นงานหรือโครงการ การให้ผู้เรียนประเมินนั้นไม่ใช่แค่การให้เกรดหรือระดับผลการเรียนแต่เป็นการช่วยผู้เรียนในการเรียนและผลิตผลงานให้ดียิ่งขึ้นผู้สอนควรให้เวลาผู้เรียนในการปรับปรุงผลงานหลังจากการประเมินตนเองแล้วและให้ประเมินงานของผู้อื่นการประเมินงานของตนเองและผู้อื่นจะทำให้ผู้เรียนคุ้นเคยกับเกณฑ์การให้คะแนน และมุ่งมั่นทำงานให้ดีขึ้นผู้สอนควรยอมรับการที่ผู้เรียนประเมินผลงานของเพื่อนโดยให้ผู้เรียนลงชื่อในเกณฑ์การให้คะแนนที่ตนใช้ผู้สอนจะทราบว่าผู้เรียนมีความถูกต้องและยุติธรรมหรือไม่และผู้สอนอาจถามเหตุผลหากการประเมินนั้นไม่ตรงกับผู้สอนการให้เวลาผู้เรียนแก้ไขปรับปรุงงานหลังจากการประเมินเป็นเรื่องสำคัญ

(กิ่งกาญจน์ สิรสุคนธ์, 2549: 82)อย่างไรก็ตามเกณฑ์การให้คะแนนที่ดีที่สุดควรมีการปฏิบัติดังนี้ (ส. วาสนา ประวาลพฤกษ์, 2550: 5-6)

1. เกณฑ์การให้คะแนนแบบทั่วไปจะต้องสอดคล้องกับจุดประสงค์ทั่วไปของการประเมินการปฏิบัติงานของแต่ละบุคคลแต่ควรจะชี้เฉพาะพอที่จะสามารถประเมินได้อย่างเจาะจงและเหมาะสมเพียงพอในการอ้างอิงถึงจุดประสงค์การเรียนรู้ทั่วไปที่สามารถประเมินได้

2. เกณฑ์การให้คะแนนต้องสามารถใช้พิจารณาการปฏิบัติที่เหมาะสมและสมจริงมิใช่เป็นการพิจารณาเอาเองโดยมองภาพรวมของการปฏิบัติและในขณะเดียวกันต้องไม่ใช่การพิจารณาจากสิ่งที่สามารถเห็นได้อย่างง่ายดายเช่นการนับหรือจำนวน

3. เกณฑ์การให้คะแนนแต่ละเกณฑ์ต้องมีมิติเดียว ไม่รวมมิติที่แตกต่างกันและไม่เกี่ยวข้องกันไว้ในเกณฑ์เดียวกัน

4. ในการสร้างเกณฑ์การให้คะแนนจะต้องศึกษาและพิจารณาจากตัวอย่างงานหรือผลการปฏิบัติหลาย ๆ ตัวอย่างที่มีระดับความแตกต่างกันตั้งแต่ดีที่สุดถึงแย่ที่สุด

5. เขียนอธิบายคุณภาพของงานโดยใช้ถ้อยคำที่บอกถึงคุณภาพที่สูงกว่าหรือสิ่งที่ขาดหายไปจากงานนั้นเพื่อให้สามารถแยกแยะความเหมือนหรือความแตกต่างของแต่ละระดับคุณภาพโดยพยายามหลีกเลี่ยงคำขยายเชิงเปรียบเทียบที่เป็นนามธรรม

6. กำหนดระดับของการประเมินให้พอเหมาะกับความสามารถที่จะกำหนดความแตกต่างตามระดับคุณภาพได้อย่างพอเพียงไม่มากเกินไปโดยทั่วไปจะอยู่ใน 6 ระดับหรือ 12 ระดับ

7. คำอธิบายระดับคุณภาพกำหนดให้เหมะสมกับวัยของผู้เรียนเพื่อที่เขาจะสามารถประเมินตนเองได้และปรับปรุงตัวเองได้ตามระดับคุณภาพนั้นๆในกรณีนี้มีข้อแนะนำคือในแต่ละระดับควรมีตัวอย่างงานที่ได้รับการประเมินในระดับนั้นๆให้เห็นชัดเจนสามารถเปรียบเทียบได้และเป็นรูปธรรม

8. เกณฑ์การประเมินต้องเน้นให้เห็นถึงผลกระทบอันเนื่องมาจากการปฏิบัติงานนั้นผลประโยชน์ที่เกิดขึ้นจากการที่ผู้เรียนได้สร้างผลงานนั้นโดยเน้นกระบวนการและความพยายามในการปฏิบัติงานนั้น

อรนุชศรีสะอาด (2554: 3-4) เสนอกระบวนการของเกณฑ์การให้คะแนนที่อิงการวัดและประเมินตนเองว่าประกอบด้วยขั้นตอนพื้นฐาน 3 ขั้นตอนดังต่อไปนี้

1. การกำหนดความคาดหวังที่ชัดเจน

ความคาดหวังสำหรับกิจกรรมหรือการปฏิบัติงานควรพูดให้ชัดเจนจากผู้สอนผู้เรียน หรือทั้งสอง อย่างใดอย่างหนึ่ง ผู้สอนถามผู้เรียนเพื่อร่วมกันกำหนดคุณภาพขนาดของเกณฑ์การให้คะแนน เช่นทำข้อความที่เป็นความคิดเห็นให้ชัดเจนสนับสนุนด้วยความจริงและน่าเชื่อถือเป็นส่วนหนึ่งของเกณฑ์ความคิดและเนื้อหาถามคำถามและข้อคิดเห็นและปรับปรุงเกณฑ์การให้คะแนนให้ชัดเจนตามที่ต้องการจนผู้สอนมีความมั่นใจว่าผู้เรียนเข้าใจและยอมรับเกณฑ์การให้คะแนนที่สร้างขึ้นผู้สอนจึงให้ผู้เรียนเริ่มงานที่ได้รับมอบหมาย

2. การปฏิบัติการประเมินตนเอง

ผู้เรียนสร้างงานที่ได้รับมอบหมายผู้เรียนติดตามความก้าวหน้าของงานที่ได้รับมอบหมายโดยการเปรียบเทียบการปฏิบัติกับเกณฑ์การให้คะแนนที่กำหนดไว้

3. การปรับปรุง

ผู้เรียนใช้ผลสะท้อนกลับจากการประเมินตนเองตามเกณฑ์การให้คะแนน เพื่อเป็นแนวทางการปรับปรุงงานซึ่งอาจจะเป็นการปรับปรุงงานเดิมให้ดียิ่งขึ้นหรือนำไปใช้ในการพยายามพัฒนาการปฏิบัติงานชิ้นใหม่ให้ดียิ่งขึ้นทำให้ผู้เรียนเกิดการเรียนรู้และพัฒนาตนเองดังนั้นผู้สอนจึงต้องทำความเข้าใจเกี่ยวกับเกณฑ์การให้คะแนน และการสร้างเกณฑ์การให้คะแนนให้ชัดเจนเพื่อจะได้นำไปถ่ายทอดและฝึกอบรมให้ผู้เรียนสามารถสร้างและใช้เกณฑ์การให้คะแนนได้อย่างมีความเข้าใจ ยุทธวิธีสำหรับการสอนการประเมินตนเอง 4 ขั้นคือ

3.1 ผู้เรียนช่วยกันนิยามเกณฑ์การประเมิน (ด้วยความช่วยเหลือของผู้สอนในการสร้างเกณฑ์การให้คะแนนที่แสดงความคาดหวังของการปฏิบัติงานและเกณฑ์ในการแปลความหมาย)

3.2 สอนผู้เรียนว่าใช้เกณฑ์อย่างไร (รูปแบบการใช้เกณฑ์การให้คะแนนด้วยตัวอย่างการประเมินการปฏิบัติ)

3.3 ให้ผลสะท้อนกลับการประเมินตนเองของผู้เรียน (ให้ผู้เรียนมีส่วนร่วมในการอภิปรายความแตกต่างระหว่างการประเมินตนเองและการวัดและประเมินโดยกลุ่มเพื่อนหรือผู้สอน)

3.4 ช่วยผู้เรียนในการใช้ข้อมูลการประเมินเพื่อพัฒนาแผนปฏิบัติการ (ค้นหาแนวทางในการปฏิบัติงานและอธิบายยุทธวิธีสำหรับพิชิตจุดอ่อนจุดแข็งของการประเมินตนเองสามารถพัฒนาโดยการฝึกอบรมผู้เรียน และจุดอ่อนแต่ละอย่างสามารถลดได้โดยการปฏิบัติของผู้สอนการประเมินตนเองมีประโยชน์ในทุกวิชาถ้าผู้เรียนดำเนินการ ผู้เรียนสามารถประเมินได้ และถ้าผู้เรียนสามารถประเมินได้ผู้เรียนสามารถพัฒนาได้ (If students produce it, they can assess it; and if they

can assess it, they can improve it) (Heidi Andrade, 2008 อ้างถึงใน อรนุชศรีสะอาด, 2554: 4)

ประเภทของเกณฑ์การให้คะแนน

เกณฑ์การให้คะแนนแบ่งเป็น3 ประเภท ได้แก่(Nitko, 1996 อ้างถึงใน ฉัตรศิริปิยะพิมลสิทธิ์, 2544)

1. เกณฑ์การให้คะแนนแบบภาพรวม (Holistic Rubrics)เป็นเกณฑ์การให้คะแนนผลงานหรือกระบวนการที่ไม่ได้แยกส่วนหรือแยกองค์ประกอบการให้คะแนนคือจะประเมินในภาพรวมของผลงานหรือกระบวนการนั้นโดยพิจารณาผลงานของผู้เรียนในภาพรวมว่ามีคุณภาพสอดคล้องกับเกณฑ์ในระดับใดและมีคะแนนเดี่ยวสำหรับงานนั้น ซึ่งจะมีคำอธิบายคุณภาพของงานประกอบการให้คะแนนและตัดสินคะแนนต่างๆ ได้ด้วย เช่น การประเมินการล้างมืออย่างถูกวิธี จะได้รับคะแนนออกมาเป็นระดับคะแนนเดียว เช่น ถูกต้องดี พอใช้ หรือยังต้องปรับปรุง

เกณฑ์การให้คะแนนแบบภาพรวมส่วนใหญ่จะประกอบด้วย 3 - 6 ระดับ ส่วนเกณฑ์การให้คะแนน 3 ระดับ จะเป็นที่นิยมใช้กันมาก เนื่องจากการใช้เกณฑ์ 3 ระดับนั้นง่ายต่อการกำหนดรายละเอียด ซึ่งจะยึดเกณฑ์ค่าเฉลี่ยสูงกว่าค่าเฉลี่ย และต่ำกว่าค่าเฉลี่ย นอกจากง่ายต่อการกำหนดค่าแล้วยังง่ายต่อการตรวจให้คะแนนอีกด้วย เนื่องจากความแตกต่างระหว่างระดับของเกณฑ์ทั้ง 3 ระดับนั้นจะแตกต่างกันอย่างชัดเจน แต่ถ้าใช้ 5 หรือ 6 ระดับ ความแตกต่างระหว่างระดับจะแตกต่างกันเพียงเล็กน้อย ซึ่งทำให้ยากต่อการตรวจคะแนน ถ้าต้องการใช้เกณฑ์ 5 หรือ 6 ระดับ วิธีการที่จะช่วยในการกำหนดเกณฑ์ง่ายขึ้น ผู้สอนอาจจะสุ่มตัวอย่างงานของผู้เรียนมาตรวจ แล้วแยกเป็น 3 กอง เป็นงานที่มีคุณภาพดี ปานกลาง และไม่ดี แล้วตรวจสอบลักษณะที่เป็นตัวแยกระหว่างงานที่มีคุณภาพไม่ดี ลักษณะเหล่านี้จะมาเป็นรายละเอียดของแต่ละระดับ

2. เกณฑ์การให้คะแนนแบบแยกส่วน (Analytic Rubrics)เป็นเกณฑ์การให้คะแนนที่แยกส่วนหรือองค์ประกอบคุณลักษณะของผลงานหรือกระบวนการแล้วนำแต่ละส่วนหรือองค์ประกอบของคุณลักษณะมารวมกันเป็นคะแนนรวมกล่าวอีกนัยหนึ่งก็คือ เป็นการให้คะแนนเป็นส่วน ๆ โดยระบุรายละเอียดออกเป็นประเด็นย่อยและแต่ละประเด็นมีคุณภาพอย่างไร เช่น การประเมินการล้างมือ จะจำแนกประเด็นการประเมินออกเป็นวิธีการล้างมือ ความสะอาดของมือ และลักษณะนิสัย ได้แก่ การใช้น้ำการใช้สบู่หรือเจลล้างมือ เป็นต้น

3. เกณฑ์การให้คะแนนแบบภาพรวมแล้วจำแนกรายละเอียดแบบแยกส่วน (Annotated Holistic Rubrics) ผู้ประเมินจะประเมินแบบภาพรวมก่อนแล้วจึงประเมินแยกส่วนอีกบางคุณลักษณะที่เด่น ๆเพื่อใช้เป็นผลสะท้อนในบางคุณลักษณะของผู้เรียน

การให้คะแนนแบบภาพรวมใช้ได้ง่ายและใช้เพียงไม่กี่ครั้งต่อผู้เรียน 1 คน จะเป็นการประเมินในภาพรวมของทุกคุณลักษณะในการปฏิบัติงาน ส่วนการให้คะแนนแบบแยกส่วนจะใช้บ่อยครั้งโดยจะประเมินแยกในแต่ละคุณลักษณะของงานซึ่งการประเมินแบบนี้จะมีประโยชน์เมื่อสนใจจะวินิจฉัยหรือช่วยเหลือผู้เรียนว่ามีความรู้ความเข้าใจในแต่ละส่วนหรือแต่ละคุณลักษณะของการปฏิบัติงานนั้นๆ หรือไม่ซึ่งจะมีส่วนให้ผู้สอนได้ช่วยเสริมสร้างหรือพัฒนาการเรียนรู้ในแต่ละคุณลักษณะของผู้เรียนให้ดียิ่งขึ้นส่วนการให้คะแนนแบบองค์รวมแล้วจำแนกรายละเอียดแบบแยกส่วนจะรวมข้อจำกัดของแบบภาพรวมและแบบแยกส่วนไว้ด้วยกันเริ่มด้วยการประเมินในภาพรวมของการปฏิบัติงานด้วยแบบภาพรวม แล้วผู้ประเมินเลือกประเมินอีกเพียงบางคุณลักษณะของงานแบบแยกส่วนซึ่งการประเมินเพียงบางคุณลักษณะนี้จะไม่มีผลต่อการเปลี่ยนแปลงคะแนนที่ประเมินแบบภาพรวมประโยชน์ก็คือจะมีความรวดเร็วในการประเมินและเป็นการให้ผู้ประเมินได้เลือกประเมินเฉพาะบางคุณลักษณะที่โดดเด่นเพียงไม่กี่องค์ประกอบเพื่อเป็นข้อมูลป้อนกลับให้แก่ผู้เรียนแต่ไม่มีประโยชน์ในการวินิจฉัยผู้เรียนว่าบกพร่องในคุณลักษณะใด เพราะหลาย ๆคุณลักษณะไม่ได้ถูกประเมิน

จะเห็นได้ว่าเกณฑ์การให้คะแนนทั้ง 3 รูปแบบเป็นเสมือนเครื่องมือที่ช่วยในการให้คะแนนในงานของผู้เรียนซึ่งการประเมินงานของผู้เรียนต้องสัมพันธ์กับเกณฑ์และระดับคุณภาพของเกณฑ์เพื่อเป็นการประกันคุณภาพของการประเมินด้วยดังนั้นการกำหนดเกณฑ์และระดับคุณภาพของเกณฑ์จึงเป็นสิ่งจำเป็นอย่างมากเพื่อบ่งบอกความเที่ยงตรงของการประเมินซึ่งในที่สุดจะนำไปสู่การประเมินตนเอง (Self–evaluation) และการปรับปรุงตนเอง (Self- adjustment)

รูปแบบของเกณฑ์การให้คะแนน

เกณฑ์การให้คะแนนคือสิ่งสำคัญที่ผู้เรียนควรรู้และปฏิบัติได้โดยมีระดับคุณภาพของแต่ละเกณฑ์และคำอธิบายคุณภาพทุกระดับดังตารางที่ 5.1 ที่เป็นรูปแบบการสร้างเกณฑ์การให้คะแนนแบบแยกส่วนโดยเป็นรูปแบบกลางที่สามารถนำไปปรับใช้ได้กับวิชาต่าง ๆได้ ส่วนตารางที่ 5.2 เป็นตัวอย่างเกณฑ์การประเมินแบบแยกส่วน

ตารางที่ 4.1 รูปแบบการสร้างเกณฑ์การประเมินแบบแยกส่วน

เกณฑ์	ระดับการปฏิบัติงานของผู้เรียน					คะแนน
	ยอดเยี่ยม	ดีมาก	ดี	ผ่าน	ควรปรับปรุง	
เกณฑ์ที่ 1						
เกณฑ์ที่ 2						
เกณฑ์ที่ 3						
เกณฑ์ที่ 4						
เกณฑ์ที่ 5						
					รวมคะแนน	
ข้อมูลป้อนกลับ :						

ตารางที่ 4.1 เป็นการกำหนดเกณฑ์การให้คะแนนแบบแยกส่วน โดยมีการพิจารณาประเด็นต่าง ๆ จำนวน 5 เกณฑ์ที่แยกในงานชิ้นเดียวกัน สำหรับแนวทางการกำหนดเกณฑ์ผู้สอนจะต้องให้ความหมายของเกณฑ์แต่ละเกณฑ์ที่จะนำมาใช้ในการประเมินผลงานของผู้เรียน โดยเกณฑ์ดังกล่าวจะต้องมีความชัดเจน สามารถวัดได้อย่างเที่ยงตรง เป็นต้น

ตารางที่ 4.2 ตัวอย่างเกณฑ์การให้คะแนนในการแก้ปัญหา

เกณฑ์	ระดับ 4	ระดับ 3	ระดับ 2	ระดับ 1
ความเข้าใจปัญหา	แสดงความเข้าใจปัญหาอย่างชัดเจนและลุ่มลึก	แสดงความเข้าใจปัญหาอย่างชัดเจน	แสดงความเข้าใจปัญหาได้บางส่วน และยังคงให้ผู้สอนช่วยอธิบายเพิ่มเติม	ต้องให้ความช่วยเหลือในการทำความเข้าใจปัญหา
การวางแผน	คิดหาวิธีการได้อย่างเหมาะสมและมีประสิทธิภาพในการวางแผน	เลือกและใช้วิธีการที่เหมาะสมในการวางแผน	แสดงหลักฐานในการวางแผนและใช้วิธีที่อาจไม่มีประสิทธิภาพ	ต้องให้ความช่วยเหลือในการเลือกหรือวิธีที่เหมาะสม
การแก้ปัญหา	คำตอบถูกต้องและแสดงวิธีการแก้ปัญหาได้มากกว่า 1 วิธี	คำตอบถูกต้องและแสดงวิธีการแก้ปัญหาที่ถูกต้อง	คำตอบผิดเพราะแสดงวิธีการในการแก้ปัญหาที่ผิดพลาดเล็กน้อย	คำตอบผิดเพราะแสดงวิธีการในการแก้ปัญหาที่ผิดพลาดมาก
การพรรณนาผลเฉลย	อธิบายเหตุผลอย่างชัดเจน มีความเกี่ยวเนื่องและลุ่มลึก	อธิบายเหตุผลได้อย่างสมเหตุสมผล	ให้คำตอบได้แต่การอธิบายคำตอบมาจากการช่วยเหลือของผู้สอน	อธิบายเหตุผลไม่เป็นระบบและยากที่จะติดตามในคำอธิบาย

ที่มา: Montgomery (2000: 325)

ส่วนตารางที่ 4.3 เป็นรูปแบบการสร้างเกณฑ์การประเมินแบบภาพรวมโดยเป็นรูปแบบกลางที่สามารถนำไปปรับใช้ได้กับวิชาต่าง ๆ ได้ส่วนตารางที่ 4.4 เป็นตัวอย่างเกณฑ์การประเมินแบบภาพรวม

ตารางที่ 4.3 รูปแบบการสร้างเกณฑ์การประเมินแบบภาพรวม

ผลการประเมิน(คะแนน)	คำอธิบายเกณฑ์/รายละเอียด
ยอดเยี่ยม(5)	
ดีมาก(4)	
ดี (3)	
ผ่าน (2)	
ควรปรับปรุง (1)	
ไม่ผ่าน (0)	

ตารางที่ 4.4 ตัวอย่างเกณฑ์การประเมินแบบภาพรวม

คะแนน	รายละเอียด
5	แสดงความเข้าใจปัญหาอย่างสมบูรณ์ คำตอบประกอบด้วยทุกประเด็นที่ต้องการ
4	แสดงความเข้าใจปัญหาค่อนข้างมาก คำตอบปรากฏทุกประเด็นที่ต้องการ
3	แสดงความเข้าใจปัญหาบางส่วน คำตอบประกอบด้วยประเด็นส่วนใหญ่ที่ต้องการ
2	แสดงความเข้าใจปัญหาเพียงเล็กน้อย ประเด็นส่วนใหญ่ที่ต้องการไม่ปรากฏ
1	แสดงความไม่เข้าใจปัญหา
0	ไม่ตอบ/ ไม่ทำงาน

ที่มา: https://www.uni.edu/chfasoa/analyticholisticrubrics.pdf

โดยทั่วไปการกำหนดเกณฑ์การให้คะแนนจะกำหนดเป็นระดับต่าง ๆ 4-5 ระดับ โดยพิจารณาจากเป้าหมายและความคาดหวังในการปฏิบัติงานนั้น ๆ กำหนดระดับสูงสุดแล้วแบ่งออกเป็น 2 ช่วง คือ ช่วงที่ยอมรับได้ (ผ่าน) และยอมรับไม่ได้ (ไม่ผ่าน) แล้วพิจารณาว่าในช่วงของการยอมรับได้นั้นแบ่งออกเป็น 2 ระดับ คือ ผ่านอย่างดีและผ่านอย่างพอใจ แล้วทำเช่นเดียวกันในช่วงของการไม่ยอมรับหรือไม่ผ่าน ส่วนคะแนนศูนย์คือไม่ปฏิบัติหรือปฏิบัติเพียงเล็กน้อยและไม่ถูกต้องเลย

ประโยชน์ของเกณฑ์การให้คะแนน

ปัจจุบันเกณฑ์การให้คะแนนได้ถูกนำมาใช้ในการประเมินการปฏิบัติงานหรือชิ้นงานของผู้เรียนเพื่อพิจารณาถึงความสำเร็จที่เกิดขึ้นกับผู้เรียน ทั้งนี้เพราะเกณฑ์การให้คะแนนจะเกี่ยวข้องกับการสร้างมาตรฐานและการอธิบายที่ชัดเจนว่ามาตรฐานจะประสบผลสำเร็จได้อย่างไร เช่น เกณฑ์การให้คะแนนในการประเมินเรียงความควรมีรายการทุกอย่างที่ผู้เรียนจำเป็นต้องทราบเพื่อจะได้ผลการเรียนหรือคะแนนตามที่ระบุในเกณฑ์การให้คะแนนซึ่งจะแสดงระดับของความแตกต่างในการปฏิบัติ เช่น ระดับ A, B และอื่น ๆ เป็นต้น เกณฑ์ให้คะแนนนอกจากจะช่วยให้ผู้เรียนรับรู้ว่าจะต้องทำอะไรเพื่อให้ประสบความสำเร็จในการเรียน แล้วประโยชน์ของเกณฑ์ให้คะแนนพอสรุปได้ดังนี้ (Cooper & Gargan, 2009: 55)

1. เกณฑ์การให้คะแนนช่วยให้ผู้สอนสามารถคิดวิจารณญาณและไตร่ตรองในสิ่งที่กำลังสอนและสิ่งที่ให้ผู้เรียนจำเป็นต้องเรียนรู้

เกณฑ์การให้คะแนนเป็นการกำหนดเกณฑ์และคำอธิบายเกณฑ์ที่ผู้สอนต้องทำล่วงหน้าก่อนสอนในรายวิชา เช่น การมอบหมายให้ผู้เรียนทำโครงงานหรือปฏิบัติงานที่สอดคล้องกับผลการเรียนรู้ที่คาดหวังในรายวิชา ซึ่งการกำหนดเกณฑ์การให้คะแนนจะช่วยให้ผู้สอนคิดพิจารณาและไตร่ตรองได้ว่าอะไรเป็นสิ่งสำคัญสำหรับผู้เรียนและอะไรเป็นสิ่งสำคัญสำหรับการสอน

2. เกณฑ์การให้คะแนนสามารถทำให้เกิดความคาดหวังและมาตรฐานในการปฏิบัติที่ชัดเจนกับผู้เรียน ผู้ปกครอง ผู้สอน นักการศึกษาและผู้อื่นที่เกี่ยวข้อง

การใช้เกณฑ์การให้คะแนนจะทำให้ผู้สอนกำหนดสิ่งที่จะสอนและมาตรฐานที่จะให้ผลการปฏิบัติงานแก่ผู้เรียน การแจกเกณฑ์การให้คะแนนแก่ผู้เรียน พ่อ แม่ และผู้ปกครอง จะทำกลุ่มคนเหล่านี้ทราบสิ่งที่ผู้สอนคาดหวังหรือสิ่งที่ต้องการให้เกิดกับผู้เรียน

3. เกณฑ์การให้คะแนนเปิดโอกาสในการสะท้อนผล ให้ข้อมูลป้อนกลับและเกิดการเรียนรู้อย่างต่อเนื่อง

การใช้เกณฑ์การให้คะแนนทางการศึกษาจะมีความชัดเจนในการประเมินความก้าวหน้าของผู้เรียนทั้งในด้านการมอบหมายงานให้ผู้เรียนปฏิบัติและการประเมินคุณภาพงาน ซึ่งผู้สอนหรือผู้ประเมินสามารถสะท้อนผลงานและให้ข้อมูลป้อนกลับที่เป็นประโยชน์แก่ผู้เรียนเพื่อนำไปปรับปรุงหรือพัฒนาตนให้เกิดการเรียนรู้ได้อย่างต่อเนื่องและมีประสิทธิภาพต่อไป

ณัชชา มหปุญญานนท์ (2552) ได้สรุปถึงข้อดีและประโยชน์ของเกณฑ์การให้คะแนน ดังนี้

1. ช่วยให้การคาดหวังของผู้สอนที่มีต่อผลงานของผู้เรียนบรรลุผลสำเร็จได้โดยผู้เรียนเกิดความเข้าใจและสามารถใช้เกณฑ์การให้คะแนนต่อการประเมินและพัฒนาชิ้นงานของตน

2. ช่วยให้ผู้สอนเกิดความกระจ่างชัดยิ่งขึ้นว่าต้องการให้ผู้เรียนเกิดการเรียนรู้หรือพัฒนาการอะไรบ้าง

3. ช่วยให้ผู้เรียนสามารถระบุคุณลักษณะจากงานที่เป็นตัวอย่างได้โดยใช้เกณฑ์การให้คะแนนตรวจสอบ

4. ช่วยให้ผู้เรียนสามารถควบคุมตัวเองในการปฏิบัติงานเพื่อไปสู่ความสำเร็จ

5. เป็นเครื่องมือในการเชื่อมโยงความสัมพันธ์ระหว่างกิจกรรมปฏิบัติงานต่างๆของผู้เรียนได้เป็นอย่างดี

6. ช่วยให้บุคคลที่เกี่ยวข้องเช่นผู้ปกครองผู้สนับสนุนผู้นิเทศได้เกิดความเข้าใจเกณฑ์ในการตัดสินผลงานของผู้เรียน

7. ช่วยในการให้เหตุผลประกอบการให้เกรดผู้เรียนได้

8. ช่วยเพิ่มคุณภาพของผู้เรียน

อย่างไรก็ตามข้อจำกัดในการใช้เกณฑ์การให้คะแนนสามารถสรุปได้ดังนี้ (Cooper & Gargan, 2009: 55)

1. เกณฑ์การให้คะแนนยังคงเป็นเรื่องของอัตวิสัย (subjective) กล่าวคือ ในการประเมินผู้เรียนได้มีการแปลงรายการที่เป็นเชิงคุณภาพและเกณฑ์ซึ่งอิสระจากกันมาเป็นผลรวมของคะแนนหาคะแนนเฉลี่ย และแปลงเป็นผลการเรียน (grade)

2. เกณฑ์การให้คะแนนทำให้ผู้สอนมีภาระงานมากขึ้นกล่าวคือ การสร้างและการใช้เกณฑ์การให้คะแนนอาจเป็นการสร้างภาระงานที่มากขึ้นสำหรับผู้สอนที่ภาระงานล้นมืออยู่แล้ว

3.เกณฑ์การให้คะแนนอาจเป็นสิ่งที่จำกัดด้านการศึกษากล่าวคือ เกณฑ์การให้คะแนนที่ออกแบบไม่ดี หรือมีการกำหนดเกณฑ์ที่ชัดเจนแล้วโดยผู้สอน อาจเป็นสิ่งที่ยับยั้งความคิดสร้างสรรค์ของผู้เรียน แม้ว่าเกณฑ์การให้คะแนนจะมีกรอบในการกำหนดการปฏิบัติงานของผู้เรียน แต่เป็นสิ่งที่ผลักดันให้ผู้เรียนทุกคนแก้ปัญหาในแนวทางเดียวกัน ซึ่งทำให้มีผลต่อการให้อิสระผู้เรียนในการคิดสร้างสรรค์สิ่งใหม่ ๆ หรือคิดนอกกรอบที่ไกลเกินกว่าเกณฑ์การให้คะแนนที่กำหนดก็เป็นได้

สรุปท้ายบท

เกณฑ์การให้คะแนนเป็นแนวทางหรือเครื่องมือในการประเมินที่กำหนดคะแนนที่จะแยกแยะระดับต่าง ๆของความสำเร็จในการเรียนหรือการปฏิบัติงานของผู้เรียนได้อย่างชัดเจนจากระดับดีเยี่ยมไปจนถึงระดับต้องปรับปรุงแก้ไข โดยในการประเมินจะมีเกณฑ์ด้านต่าง ๆที่ใช้พิจารณาชิ้นงานหรือการปฏิบัติงาน และระดับคุณภาพของเกณฑ์แต่ละด้านซึ่งมีตั้งแต่ระดับดีเยี่ยมจนถึงระดับต้องปรับปรุง การประเมินผลงานของผู้เรียนสามารถทำได้ 2 ลักษณะ คือ ผลงานที่ได้จากกระบวนการของผู้เรียนและกระบวนการที่นักเรียนใช้เพื่อให้เกิดผลงานจะประเมินในลักษณะใดขึ้นอยู่กับจุดมุ่งหมายในการเรียนรู้อาจจะประเมินลักษณะใดลักษณะหนึ่งหรือประเมินทั้งสองลักษณะก็ได้ผู้สอนในฐานะผู้ประเมินจะต้องตัดสินคุณภาพของผลงานหรือกระบวนการปฏิบัติงานของผู้เรียนแต่ละคนที่มีระดับที่แตกต่างกันหลายระดับระดับที่แตกต่างกันอาจจะเป็นระดับคุณภาพของชิ้นงานที่ได้สร้างขึ้นหรือระดับของกระบวนการต่าง ๆ ที่ผู้เรียนแต่ละคนได้ใช้เพื่อเกิดผลงาน เกณฑ์การให้คะแนนนอกจากจะใช้เพื่อประเมินผลงานหรือชิ้นงานแล้วยังสามารถใช้เป็นเครื่องมือในการสอนได้อย่างดีโดยให้ผู้เรียนได้รับทราบว่าผู้สอนคาดหวังอะไรบ้างจากผลงานหรือชิ้นงานที่มอบหมายหรือให้ผู้เรียนร่วมในการสร้างเกณฑ์ก็จะทำให้เกิดการมีส่วนร่วมและรับผิดชอบและหากมีตัวอย่างชิ้นงานประกอบให้ผู้เรียนได้ช่วยกันพิจารณาอภิปรายโดยใช้เกณฑ์ที่ร่วมกันสร้างขึ้นก็จะยิ่งทำให้ผู้เรียนสามารถแยกแยะได้ว่าชิ้นงานที่ดีมีคุณภาพเป็นอย่างไร

โดยทั่วไปเกณฑ์การให้คะแนนมี 3 ประเภท คือ เกณฑ์การให้คะแนนแบบภาพรวม เกณฑ์การให้คะแนนแบบแยกส่วน และเกณฑ์การให้คะแนนแบบภาพรวมแล้วจำแนกรายละเอียดแบบแยกส่วนสำหรับการเลือกใช้เกณฑ์การให้คะแนนประเภทใดนั้น ควรคำนึงถึงงาน ชิ้นงานที่จะให้ผู้เรียนปฏิบัติโดยเน้นความสอดคล้องระหว่างผลการเรียนรู้ที่คาดหวัง มาตรฐาน ตัวชี้วัดตามหลักสูตร นอกจากนี้เกณฑ์การให้คะแนนที่สร้างขึ้นควรเป็นรูปธรรม มีความชัดเจนในด้านการสื่อสาร เหมาะสม

กับระดับชั้นของผู้เรียน โดยมุ่งเน้นความยุติธรรมและปราศจากความลำเอียงในการประเมินผล ที่สำคัญยังใช้เป็นเครื่องมือสื่อสารระหว่างผู้สอนกับผู้เรียนผู้สอนกับผู้ปกครองและผู้เรียนกับผู้ปกครอง การมีภาพแห่งความคาดหวังที่ชัดเจนจะช่วยให้ผู้สอนสามารถให้ข้อมูลย้อนกลับที่เป็นประโยชน์แก่ผู้เรียนและเป็นประเด็นสำหรับพูดคุยเพื่อการพัฒนาการเรียนรู้ได้ดียิ่งขึ้น

เอกสารอ้างอิง

กิ่งกาญจน์ สิรสุคนธ์. (2549). รอบรู้สู่การวัดและประเมินผล: มารู้จักรูบริคกันเถอะ. วารสารวิชาการ ปีที่ 9ฉบับที่ 3 (กค.-กย.),หน้า 78-82.

_____. (2550). รูบริคหรือรูบริคการให้คะแนน. กลุ่มส่งเสริมการเรียนการสอนและประเมินผลสำนักวิชาการและมาตรฐานการศึกษาสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ พิมพ์ครั้งที่ 1 หน้า 1-51.

ฉัตรศิริ ปิยะพิมลสิทธิ์. (2544). เกณฑ์การให้คะแนน. สืบค้นเมื่อวันที่ 9 มีนาคม 2556 จาก http://www.watpon.com/Elearning/mea5.htm

ณัชชา มหปุญญานนท์ (2552). เครื่องมือการให้คะแนน. คณะศึกษาศาสตร์ มหาวิทยาลัยทักษิณ. สืบค้นเมื่อวันที่ 10 เมษายน 2556 จาก http://file.siam2web.com/natcha/resreport/2009127_42246.pdf

ส.วาสนา ประวาลพฤกษ์. (2550). การสร้างเกณฑ์การประเมิน Rubric Sampler ของ Relearning by Design, Inc. สำนักทดสอบทางการศึกษาสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน. หน้า 1-42.

อรนุช ศรีสะอาด (2554). “การประเมินตนเอง”. วารสารวัดผลการศึกษา. ปีที่ 17 ฉบับที่ 2 เดือนธันวาคม. มหาวิทยาลัยมหาสารคาม. หน้า 1-5.

Cooper, B. S., & Gargan, A. (2009). Rubrics in Education. *Kappan,*pp.54-55.

Goodrich, Heidi. (1996). Understanding rubrics. *Educational Leadership,54.4*, 14.

Kan, A. (January 2007). An Alternative Method in the New Educational Program from the Point of Performance-based Assessment: Rubric Scoring Scales. *Educational Sciences: Theory & Practice, 7(1),* 142-152.

Montgomery, K. (2000). Classroom Rubrics: Systematizing What Teachers Do Naturally. *The Clearing House, 73(6),324-328.*

Moon, T. R., Brighton, C. M., & Callahan, C. M. (2005). Development of Authentic Assessments for the Middle School Classroom. *The Journal of Secondary Gifted Education, 16(2/3),*119-133.

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review, 9,* 129-144.