



## การวัดประเมินในชั้นเรียน : วิวัฒนาการและแนวคิดใหม่ เพื่อพัฒนาการเรียนรู้

รองศาสตราจารย์ ดร.องอาจ นัยวัฒน์

คณบดีคณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทร์วิโรฒ

ในปัจจุบันนี้บทบาทและแนวคิดสำคัญด้านการวัดประเมินการเรียนรู้ (Learning assessment) ได้เปลี่ยนแปลงไปอย่างมากและเป็นไปอย่างรวดเร็วในวงการศึกษาของหลายประเทศ เช่น อังกฤษ ฝรั่งเศส สหรัฐอเมริกา รวมทั้งไทย จะเห็นได้จากการปรับเปลี่ยนวิธีการวัดประเมินการเรียนรู้แบบอิงหลักสูตรเป็นฐาน (Curriculum-based assessment) ที่ดำเนินการอย่างเป็นระบบและต่อเนื่อง เพื่อให้ได้ข้อมูลสารสนเทศย้อนกลับไปพัฒนาการเรียนรู้ของผู้เรียนและการจัดการเรียนรู้ของครูผู้สอน นอกจากนี้ การวัดประเมินการเรียนรู้ตามทฤษฎีการเรียนรู้เพื่อรอบรู้ (Mastery learning theory) ที่ได้รับความนิยมอย่างมากในวงการศึกษาของหลายประเทศนับตั้งแต่ปลายปี ค.ศ. 1970 เป็นต้นมา (พ.ศ. 2520-ปัจจุบัน) พบว่า ในทางปฏิบัติจริงมักมุ่งเน้นการวัดและประเมินผลเพื่อสรุปรวบยอดความรู้ของผู้เรียนตามวัตถุประสงค์เชิงพฤติกรรมที่กำหนดไว้ล่วงหน้าอย่างเฉพาะเจาะจง โดยใช้แบบทดสอบเลือกตอบเป็นเครื่องมือวัดหลักเมื่อสิ้นสุดภาคเรียนหรือปลายปีการศึกษา (Summative assessment) มากกว่าการวัดและประเมินผลเพื่อตรวจสอบความก้าวหน้าในการเรียนรู้ของผู้เรียนที่เกิดขึ้นระหว่างการเรียนรู้การสอน (Formative assessment) ได้เปลี่ยนแปลงมาสู่การวัดและประเมินผลที่มีลักษณะผ่อนคลายข้อจำกัดของการวัดและประเมินผลการเรียนรู้ตามแบบดั้งเดิม รวมทั้งขยายขอบข่ายมิติของการวัดและประเมินผลกว้างขวางขึ้นกว่าเดิม โดยมีผู้เรียนในฐานะผู้ถูกวัดและประเมินผลเป็นผู้มีส่วนร่วมอันสำคัญยิ่งในกระบวนการวัดและประเมินผล รวมทั้งการใช้ผลการวัดและประเมินเพื่อพัฒนาการเรียนรู้ของตนอย่างต่อเนื่อง มีครูในฐานะผู้สอนและผู้วัดประเมินคอยเสริมพลังให้เกิดการเรียนรู้ (Learning empowerment) และสนับสนุนให้ผู้เรียนแต่ละคนหรือกลุ่มสามารถพัฒนาความรู้ความสามารถหรือทักษะไปได้สูงสุดตามศักยภาพของที่มีอยู่ การวัดประเมินการเรียนรู้ตามแนวคิดใหม่นี้ จึงมุ่งเน้นไปที่ผู้เรียนเป็นสำคัญ (Student-centered or student-involved assessment) มากกว่ามุ่งเน้นไปที่ตัวครูผู้สอน (Teacher-centered assessment) หรือใช้การวัดผลเป็นสิ่งที่ “ขับเคลื่อน” (Measurement-driven assessment) ในกระบวนการจัดการเรียนรู้ ดังที่เกิดขึ้นในหลายช่วงทศวรรษที่ผ่านมา

บทความนี้เขียนขึ้นโดยมีจุดมุ่งหมายหลักเพื่อเสนอแนะทัศนคติเกี่ยวกับการวัดประเมินการเรียนรู้ในชั้นเรียนโดยมุ่งเน้นในเรื่องวิวัฒนาการและแนวคิดสำคัญของการวัดประเมินการเรียนรู้ตามกระบวนทัศน์ (Paradigm) สำคัญที่ใช้เป็นฐานคิดหรือกรอบเหตุผลทำความเข้าใจประเด็นปัญหาต่างๆ ที่เกี่ยวข้องกับการวัดประเมินการเรียนรู้

## ความหมายเชิงมโนทัศน์ของคำสำคัญ

ก่อนนำเสนอมนทัศน์เกี่ยวกับการวัดและประเมินการเรียนรู้แนวใหม่ มีความจำเป็นอย่างยิ่งที่ผู้อ่านควรเข้าใจความหมายของคำศัพท์เฉพาะที่สำคัญบางคำ ซึ่งมักปรากฏในเอกสารหรือถูกกล่าวถึงในแวดวงวิชาการด้านการวัดและประเมินผลการเรียนรู้ ทั้งนี้เพื่อเป็นพื้นฐานสำหรับการทำความเข้าใจเนื้อหาสาระในบทความนี้ และเอกสารอื่นๆ ที่เกี่ยวข้อง อนึ่ง คำศัพท์เฉพาะทางวิชาการที่จะกล่าวถึงต่อไปนี้มีความหมายเชิงมโนทัศน์ “**เหลื่อมซ้อน**” กันค่อนข้างมาก จนบางครั้งก่อให้เกิดความรู้สึกสับสนแก่ผู้อ่านว่ามีความหมายอะไรแน่ ผู้เขียนตระหนักดีว่าการนิยามความหมายตามพจนานุกรม (Dictionary terminology) และการระบุความแตกต่างในความหมายของคำศัพท์เฉพาะอย่างเจาะจงลงไป โดยเฉพาะอย่างยิ่งคำว่า การวัดประเมิน (Assessment) และการประเมินค่า (Evaluation) อาจกระทำไม่ได้ครบถ้วนและครอบคลุมความหมายทุกแง่มุมที่ผู้ใช้คำทั้งสองต้องการสื่อสาร ดังจะพบได้บ่อยครั้งที่มีการใช้ในบริบทเดียวกันจนไม่อาจแยกแยะความแตกต่างในความหมายออกจากกันได้ ดังนั้น ในที่นี้ผู้เขียนจะนำเสนอความหมายของคำในเชิงมโนทัศน์ (Conceptual terminology) อย่างกว้างๆ ตามประสบการณ์และทัศนะของผู้เขียน ดังนี้

**1) การวัดประเมิน (Assessment)** เป็นกระบวนการรวบรวมและสังเคราะห์ข้อมูลสารสนเทศสำหรับใช้ในกระบวนการตัดสินใจเกี่ยวกับสิ่งที่ต้องการประเมิน (Evaluand) ดังนั้น การวัดประเมินการเรียนรู้สามารถให้ความหมายอย่างกว้างๆ ได้ว่า เป็นกระบวนการรวบรวมและสังเคราะห์ข้อมูลสารสนเทศโดยมีจุดมุ่งหมายเพื่อใช้ในการตัดสินใจเกี่ยวกับการจัดการเรียนการสอน การจัดการภายในชั้นเรียน (Classroom

management) และการจัดกิจกรรมด้านอื่นๆ ของครูผู้สอน เพื่อส่งเสริมหรือสนับสนุนการเรียนรู้ของผู้เรียน อาทิ การสร้างบรรยากาศที่เอื้อต่อการเรียนรู้ออกกำลังกาย การปกครองชั้นเรียน และการตัดสินใจเกี่ยวกับการเรียนรู้ของนักเรียนเป็นรายบุคคลหรือกลุ่ม จากความหมายในเชิงมโนทัศน์กว้างๆ นี้จะพบว่าโดยทั่วไปการวัดประเมินการเรียนรู้มักเกิดขึ้นในระดับชั้นเรียนเป็นส่วนใหญ่หรือเป็นที่รู้จักกันโดยทั่วไปว่าเป็น “**การวัดประเมินในชั้นเรียน**” (Classroom assessment) โดยมีครูในฐานะผู้วัดประเมินวางแผนการเก็บรวบรวมข้อมูลสารสนเทศด้วยเครื่องมือวัดค่าเป็นตัวเลขหรือวิธีการที่เผยแพร่ไปถึงผลลัพธ์การเรียนรู้ของผู้เรียนในรูปถ้อยคำทางภาษา ทั้งนี้การเลือกใช้เครื่องมือวัดประเมินขึ้นอยู่กับความสอดคล้องระหว่างคุณสมบัติของเครื่องมือและวิธีการวัดประเมินกับประเภทของผลลัพธ์การเรียนรู้ของผู้เรียนที่ต้องการวัดประเมิน

**2) การทดสอบ (Testing)** เป็นกระบวนการรวบรวมข้อมูลสารสนเทศด้วยวิธีการวัดค่าเชิงปริมาณที่ดำเนินการอย่างเป็นทางการ (Formal assessment) และมีความเป็นมาตรฐาน โดยเฉพาะอย่างยิ่งการใช้เครื่องมือวัดค่าแบบเขียนตอบ (Paper-and-pencil instrument) เช่น แบบทดสอบ (Test) และมาตรวัด (Scale) เป็นเครื่องมือบ่งชี้ผลลัพธ์การเรียนรู้ของผู้เรียนที่แสดงค่าในรูปของตัวเลขหรือแต้มเชิงปริมาณ ผู้เรียนที่ได้รับคะแนนหรือแต้มจากการทดสอบสูงแสดงว่ามีความรู้ความสามารถระดับสูง ในทางตรงข้ามผู้เรียนที่ได้รับคะแนนหรือแต้มต่ำแสดงว่ามีความรู้ความสามารถในระดับต่ำ

**3) การวัด (Measurement)** เป็นกระบวนการกำหนดจำนวนเชิงปริมาณ (Numerical quantity) เพื่อบ่งชี้ลักษณะหรือคุณสมบัติของสิ่งที่ต้องการวัดค่าภายใต้กฎเกณฑ์ มาตรฐานหรือเงื่อนไขอย่างใดอย่างหนึ่งโดยใช้เทคนิคหรือเครื่องมือวัด ดังนั้น

จากความหมายของการวัดและการทดสอบดังกล่าวนี้ จะพบว่า การวัดและการทดสอบเป็นส่วนหนึ่งของการวัดประเมิน ทั้งนี้เพราะ “หัวใจ” ของการวัดและการทดสอบ คือ กระบวนการดำเนินงานที่กระทำขึ้น เพื่อให้ได้ข้อมูลสารสนเทศการวัดประเมินที่อยู่ในรูปตัวเลขหรือจำนวนเชิงปริมาณ ส่วนข้อมูลสารสนเทศในรูปที่ไม่ใช่ตัวเลข ซึ่งมักนิยมเรียกว่า ข้อมูลสารสนเทศเชิงคุณภาพ (Qualitative information) ต้องอาศัยเทคนิคหรือวิธีการวัดประเมินอื่นๆ รวบรวม อาทิ การสนทนากลุ่ม (Focus group) และการใช้แฟ้มสะสมงาน (Portfolio)

**4) การประเมินค่า (Evaluation)** เป็นกระบวนการกำหนดหรือตัดสินคุณค่าในแง่คุณธรรมความดี (Merit) ที่อยู่ภายใน (Intrinsic value) และในแง่มูลค่าหรือราคา (Worth) ที่อยู่ภายนอก (Extrinsic value) ของสิ่งที่จะประเมินค่า (Evaluand) คำศัพท์เฉพาะทางวิชาการอื่นๆ ที่มีความหมายใกล้เคียงหรือเป็นส่วนหนึ่งของกระบวนการดังกล่าวนี้ได้แก่ การวัดประเมิน (Assessment) การประเมินราคา (Appraisal) การตรวจสอบ (Auditing) การตรวจตรา (Inspection) การตัดสิน (Judgment) การจัดอันดับ (Rating) การจัดตำแหน่ง (Ranking) และการสอบหรือการทดสอบ (Examining or testing) คำศัพท์เหล่านี้ โดยเฉพาะอย่างยิ่งคำว่า “**การวัดประเมิน**” และ “**การประเมินราคา**” มีความหมายหลื่อมซ้อนกันมากกับคำว่า “**การประเมินค่า**” จนยากลำบากต่อการแยกแยะความแตกต่างในความหมายระหว่างคำศัพท์ทั้ง 3 คำ อย่างไรก็ตาม มีข้อสังเกตที่สำคัญบางประการเกี่ยวกับการใช้คำศัพท์เฉพาะทั้ง 3 คำ คือ

**(1) การวัดประเมินมักใช้ในบริบทการศึกษา** โดยเฉพาะอย่างยิ่ง การใช้ด้วยจุดมุ่งหมายสำคัญเพื่อพัฒนาหรือปรับปรุงการเรียนรู้ของผู้เรียนและการสอนของครู (Instructional-learning

assessment) ให้มีปริมาณสูงหรือคุณภาพดีขึ้นมากกว่าเพื่อการตัดสินเชิงคุณค่า เช่น ผ่าน-ไม่ผ่าน ดี-เลว หรืออื่นๆ ซึ่งการวัดประเมินลักษณะดังกล่าวนี้มีแนวโน้มก่อให้เกิดความรู้สึกเสียหน้าและบั่นทอนความรู้สึกภาคภูมิใจหรือการเห็นคุณค่าในตนเองอันเป็นอุปสรรคอย่างยิ่งต่อการแก้ไขปรับปรุงสิ่งที่วัดประเมินให้มีคุณภาพดียิ่งขึ้น

**(2) การประเมินราคามักใช้ในบริบทการตลาดหรือธุรกิจ** เป็นการใช้เพื่อประเมินราคาผลผลิต วัตถุ หรือสิ่งของที่เป็นสังหาริมทรัพย์และอสังหาริมทรัพย์ เช่น ทอง อัญมณี บ้าน ที่ดิน และรถยนต์ บางครั้งมีการใช้ในบริบทการศึกษาที่อยู่ภายใต้แรงมูมความคุ้มค่าทางการเงิน เช่น การประเมินบุคลากรทางการศึกษา (Personnel appraisal) การใช้ในลักษณะนี้มี ความหมายโดยนัยแฝงไว้เกี่ยวกับการประเมินบุคลากรในแง่เศรษฐศาสตร์การลงทุนหรือคุณค่าทางการตลาด (Market value) เช่น การประเมินผลการปฏิบัติงานตามตำแหน่งหน้าที่รับผิดชอบเปรียบเทียบกับเงินเดือนที่เจ้าของประกอบการต้องจ่ายให้

**(3) การประเมินค่าเป็นคำทั่วไปที่ใช้ในหลายบริบท** เช่น การประเมินค่าบุคลากร (Personnel) โครงการ (Program) และการปฏิบัติงาน (Performance) ทางการศึกษา การประเมินค่ามุ่งเน้นเกี่ยวกับการกำหนดหรือตัดสินเชิงคุณค่าตามเกณฑ์ (Criteria) และมาตรฐาน (Standards) ด้านคุณธรรมความดีหรือราคาอย่างใดอย่างหนึ่งที่กำหนดขึ้นหรือตามบรรทัดฐาน (Norms) ของสังคมและวัฒนธรรมหนึ่งๆ

ในบริบทการเรียนการสอนภายในโรงเรียนหรือชั้นเรียน ครูทำหน้าที่เป็นผู้วัดประเมินการเรียนรู้ของนักเรียนเสมอ รวมทั้งมีการใช้ข้อมูลสารสนเทศจากการวัดประเมินเพื่อตัดสินคุณภาพการเรียนรู้ของนักเรียนตามจุดมุ่งหมายของการ

เรียนการสอน เพื่อศึกษาความก้าวหน้าและสรุป  
รวบยอดผลลัพธ์การเรียนรู้ของนักเรียนเป็น  
รายบุคคลหรือกลุ่ม การประเมินตัดสินคุณค่าที่  
มีคุณภาพสูงของครูไม่จำเป็นจะต้องใช้การทดสอบ  
ด้วยเครื่องมือวัดค่าแบบเขียนตอบทุกครั้งไป กล่าว  
อีกนัยหนึ่ง คือ การประเมินค่าใดๆ ไม่จำเป็นต้อง  
อาศัยการวัดเสมอไป การประเมินค่าอาจใช้วิธีการ  
แบบไม่เป็นทางการ (Informal evaluation) เช่น  
การสังเกตพฤติกรรม หรือเทคนิคการตั้งคำถาม  
ด้วยวาจา (Oral questioning technique)  
ในระหว่างมีการเรียนการสอน

## วิวัฒนาการของการวัดและประเมินการเรียนรู้

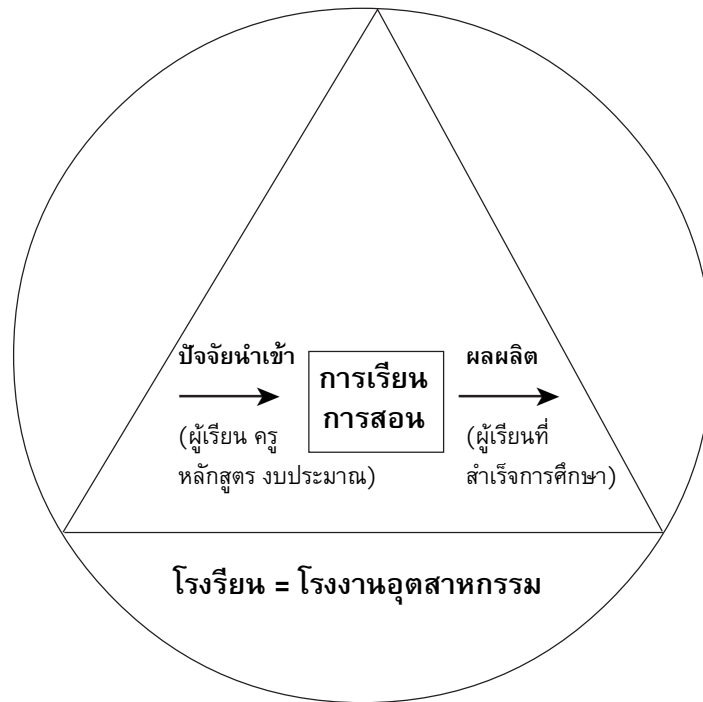
1. **ทัศนะดั้งเดิม: หลักสูตร จิตวิทยา และ**  
**การวัด** ในห้วงคริสต์ศตวรรษที่ 20 (พ.ศ. 2443-  
พ.ศ. 2543) แนวคิดหลักเกี่ยวกับประสิทธิภาพ  
ทางสังคม (Social efficiency) และการบริหาร  
จัดการทางวิทยาศาสตร์ (Scientific management)  
เจริญรุดหน้าและได้รับการยอมรับในสังคมของ  
ประเทศตะวันตก และมีอิทธิพลต่อบุคคล/กลุ่มบุคคล  
ของสังคมประเทศตะวันออกอย่างกว้างขวาง  
ในเวลาต่อมา กล่าวโดยพื้นฐานแล้วแนวคิด  
เกี่ยวกับประสิทธิภาพทางสังคมพัฒนาขึ้นจาก  
ความเชื่อที่ว่า หลักการและทฤษฎีทางวิทยาศาสตร์  
สามารถนำมาประยุกต์ใช้แก้ไขปัญหาต่างๆ ทั้งที่  
เกิดขึ้นในภาคอุตสาหกรรมและภาคสังคมหรือ  
ชุมชนของมนุษย์ตามทฤษฎีประสิทธิภาพทางสังคม  
หลักการทางวิทยาศาสตร์ที่ใช้ในกระบวนการผลิต  
ของโรงงานอุตสาหกรรม สามารถนำมาใช้ใน  
กระบวนการจัดการศึกษาของโรงเรียนได้เป็น  
อย่างดีเช่นเดียวกัน ดังนั้น มโนทัศน์สำคัญเกี่ยวกับ  
กระบวนการผลิตตามสายพานภายในโรงงาน  
อุตสาหกรรม ได้ถูกนำมาเปรียบเทียบกับ  
กระบวนการจัดการเรียนการสอนของครูที่เกิดขึ้น

ภายในสถานศึกษา โดยเริ่มต้นจากการทดสอบเพื่อ  
คัดเลือกผู้เรียนเป็น **“ปัจจัยป้อน”** (Input) เข้าสู่  
สถานศึกษาเพื่อเป็นสิ่งประกันว่าปัจจัยนำเข้าที่ส่ง  
ผ่าน **“สายพาน”** ไปยังกระบวนการจัดการเรียน  
การสอน หรือ **“กระบวนการผลิต”** (Process)  
ที่มีคุณภาพ โดยในกระบวนการผลิตมีการกำหนด  
และใช้ยุทธวิธีการควบคุมคุณภาพ (Quality control)  
เพื่อให้ **“ผลผลิต”** (Product/Output) คือ นักเรียน  
ที่สำเร็จการศึกษาออกมาตามสายพานการผลิต  
มีคุณภาพได้มาตรฐานสูงสุดทั้งในเชิงปริมาณและ  
คุณภาพ โดยมีการลงทุนหรือสูญเสียทรัพยากรและ  
ระยะเวลาในกระบวนการบริหารจัดการผลิตต่ำสุด  
(ดูภาพที่ 1) นั่นคือ เน้นประสิทธิภาพสูงสุด  
ในกระบวนการบริหารจัดการศึกษาตามแนวคิดของ  
ทฤษฎีประสิทธิภาพทางสังคมที่มีอิทธิพลต่อวงการ  
ศึกษามานานนับตั้งแต่อดีตจนถึงปัจจุบันนี้

เมื่อพิจารณาและวิเคราะห์แนวคิดตาม  
ทฤษฎีประสิทธิภาพทางสังคมที่ลุ่มลึกลงไปในรายละเอียด  
จะพบว่า แนวคิดดังกล่าวนี้ตั้งอยู่บนฐานคิด  
เชิงปรัชญาตามกระบวนทัศน์แบบปฏิฐานนิยม  
(Positivism paradigm) (องอาจ นัยวัฒน์, 2544,  
2551; Guba & Lincoln, 1989; Lincoln & Guba,  
2000) ที่มุ่งเน้นแสวงหาความรู้แท้-ความจริง  
สูงสุด (Ultimate reality/True knowledge)  
โดยอาศัยหลักการและวิธีการทางวิทยาศาสตร์  
เป็นเครื่องมือสำคัญในกระบวนการแสวงหาความรู้  
ความจริงแท้ที่ปราศจากลำเอียงหรือความคลาด  
เคลื่อนใดๆ (Bias or error) หลักสูตรที่พัฒนาขึ้น  
ตามกระบวนทัศน์แบบปฏิฐานนิยมนี้มี 2 ลักษณะ  
คือ หลักสูตรที่จำแนกเป็นสาขาวิชาย่อยตามความรู้  
ความสามารถและความสนใจของผู้เรียน และ  
หลักสูตรรวมสำหรับผู้เรียนส่วนใหญ่ที่ต้องการศึกษา  
การพัฒนาหลักสูตรตามทฤษฎีประสิทธิภาพทาง  
สังคมมาจากความเชื่อที่ว่า หลักสูตรที่ดีจะต้อง

มีประโยชน์ในแง่สามารถนำไปใช้ได้กับผู้เรียนส่วนใหญ่ (Utilitarian content) รวมทั้งสอดคล้องกับความรู้ความสามารถและความสนใจของผู้เรียน

ทั้งนี้เพื่อลดความสูญเปล่าหรือความสิ้นเปลืองจากการลงทุนทางการศึกษา (Bobbitt 1912, อ้างอิงจาก Shepard, 2000)



ภาพที่ 1 แบบจำลองการจัดการเรียนการสอนตามทฤษฎีประสิทธิภาพทางสังคม

ในช่วงเวลาใกล้เคียงกับที่ทฤษฎีประสิทธิภาพทางสังคมมีอิทธิพลต่อการพัฒนาหลักสูตร ปรากฏว่าทฤษฎีจิตวิทยาการเรียนรู้แบบสัมพันธ์นิยม (Associationism) ของ Thorndike (1922) และแบบพฤติกรรมนิยม (Behaviorism) ของ Hull (1943) Skinner (1938, 1954) และ Gagne' (1965) ที่ส่วนใหญ่พัฒนาขึ้นจากกระบวนการทดลองในห้องปฏิบัติการทางวิทยาศาสตร์ได้รับการยอมรับอย่างกว้างขวาง และถูกนำมาประยุกต์ใช้ในบริบทการจัดการเรียนการสอนในชั้นเรียน การเรียนรู้ตามทฤษฎีจิตวิทยาแบบสัมพันธ์นิยมและพฤติกรรมนิยมมีลักษณะเชิงสะสม (Accumulative knowledge) ที่เพิ่มพูนสูงขึ้นในลักษณะค่อยเป็นค่อยไปตามลำดับขั้นลดหลั่นกันไปอย่างรัดกุมและต่อเนื่อง การเสริมแรงจูงใจ (Motivation)

จากภายนอกใดๆ โดยเฉพาะที่เป็นไปในทางบวกกับผู้เรียนในแต่ละสาระการเรียนรู้ย่อยๆ (Learning unit) จะช่วยกระตุ้นให้เกิดการเรียนรู้เพิ่มมากขึ้น โน้ตทัศน์พื้นฐานเกี่ยวกับการเรียนรู้ตามทฤษฎีจิตวิทยาทั้ง 2 แบบนี้ได้นำมาสู่ทฤษฎีพัฒนาการด้านความรู้คิด (Cognitive development theory) ที่จำแนกการเรียนรู้ออกเป็นระดับย่อยๆ ลดหลั่นกัน (Hierarchical levels) เช่น จำแนกตามระบบของ Bloom (1965) และ Gagne' (1972, 1985) ในเวลาต่อมา (อ้างอิงใน Shepard, 2000)

จิตวิทยาการเรียนรู้ตามทฤษฎีสัมพันธ์นิยมและพฤติกรรมนิยมมีอิทธิพลอย่างยิ่งต่อการเรียนการสอนและการทดสอบ ดังสะท้อนมาจากความเชื่อดั้งเดิมที่ว่า “แบบทดสอบเป็นสิ่งเดียวหรือมีรูปแบบเหมือน (Isomorphic) กับการเรียนรู้”

(Tests = Learning) นั่นคือ ถ้าครูผู้สอนต้องการทราบว่าผู้เรียนเกิดการเรียนรู้มากน้อยเพียงใด สามารถวัดประเมินผลสัมฤทธิ์ทางการเรียนได้ด้วยการใช้แบบทดสอบปรนัย (Objective tests) วัดระดับความรอบรู้ (Mastery) ตามจุดมุ่งหมายการเรียนการสอน (Learning-instructional objective) ของหน่วยการเรียนรู้ย่อยใดๆ ว่ามีมาก-น้อยหรือไม่เพียงไร ผลการวัดค่าที่อยู่ในรูปคะแนนจากแบบทดสอบปรนัยที่ผู้เรียนได้รับคือ สิ่งบ่งชี้ระดับการเรียนรู้ของผู้เรียน ด้วยเหตุนี้แบบทดสอบปรนัยที่สร้างหรือพัฒนาขึ้นตามหลักการวัดค่าทางวิทยาศาสตร์ จึงมีบทบาทสำคัญอย่างยิ่งในการวัดประเมินการเรียนรู้ในชั้นเรียน จนกระทั่งมีการเปรียบเทียบว่าการวัดประเมินการเรียนรู้ตามทัศนะดั้งเดิม คือ การวัดประเมินการเรียนรู้ด้วยแบบทดสอบปรนัย (Assessment = Measurement) ที่อยู่ภายใต้บริบทการทดสอบที่เป็นมาตรฐานเดียวกัน (Standardized testing) เพื่อป้องกันความคลาดเคลื่อนหรือความลำเอียงที่อาจเกิดขึ้นจากการวัดค่า ดังนั้น การวัดประเมินการเรียนรู้ในอดีตจึงเป็นแบบทางการ กล่าวคือ มีลักษณะแยกกิจกรรมการวัดประเมินออกจากการจัดการเรียนการสอนของครูอย่างชัดเจน ครูในฐานะผู้วัดประเมินการเรียนรู้ของผู้เรียนให้ความสำคัญและพิถีพิถันต่อความเป็นปรนัยของการวัดค่าอย่างยิ่ง กล่าวอีกนัยหนึ่ง คือ กระบวนการวัดประเมินทั้งหมดที่ดำเนินการ ขึ้นจะต้องปราศจากความเป็นอัตนัย (Subjective) เข้ามาเกี่ยวข้อง การวัดประเมินการเรียนรู้ของผู้เรียนในลักษณะภาพรวม (Holistic view) ที่เกิดจากการรับรู้หรือพิจารณาได้ด้วยประสบการณ์การสอนหรือความชำนาญใดๆ ของครูผู้สอนที่ไม่สามารถ ทดสอบหรือวัดค่าได้ด้วยหลักการทางวิทยาศาสตร์จะต้องถูกขจัดให้หมดไปเพื่อให้ผลการวัดประเมินมีความเที่ยงตรง (Validity) และ

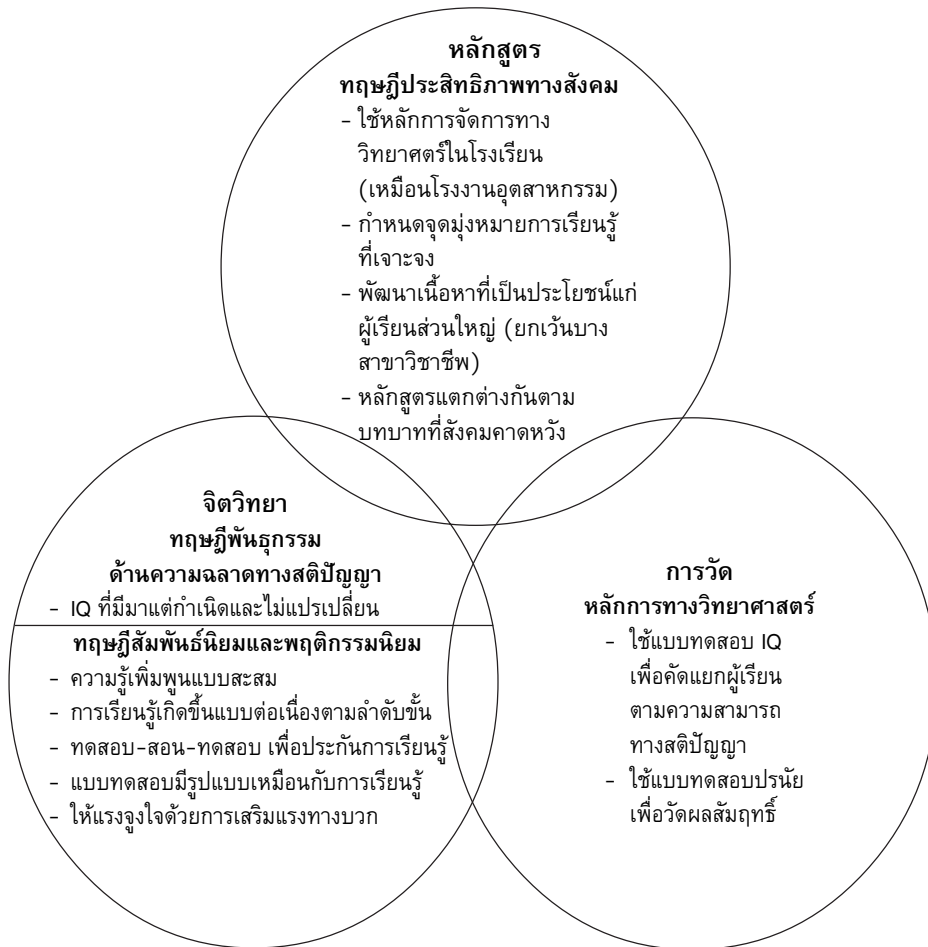
ความเชื่อมั่น (Reliability) ได้ว่าการดำเนินการดังกล่าวสามารถเข้าสู่ “**ความรู้ความจริงแท้**” ตามแนวคิดความเชื่อที่ตั้งอยู่บนกระบวนทัศน์แบบปฏิฐานนิยม ดังที่กล่าวมาข้างต้น

ในความเป็นจริงแล้วการวัดค่าโดยอาศัยแบบทดสอบปรนัย ซึ่งเป็นเครื่องมือวัดหลักสำหรับใช้ในการวัดประเมินการเรียนรู้ในชั้นเรียนดังปรากฏให้เห็นในปัจจุบันนี้มีลักษณะเฉพาะบางประการแตกต่างจากการวัดในอดีต โดยเฉพาะแนวคิดเกี่ยวกับแบบทดสอบปรนัยที่พัฒนาขึ้นมาจากหลักการวัดสมัยใหม่ที่มี Thordike นักจิตวิทยาการเรียนรู้ชาวอเมริกันคนสำคัญในกลุ่มทฤษฎีสัมพันธนิยม ซึ่งในเวลาต่อมาได้รับการยอมรับว่าเป็นบิดาของการวัดสมัยใหม่เป็นผู้จุดประกายแนวคิดนี้ขึ้น การวัดในระยะแรกมีบทบาทสำคัญมากในกระบวนการคัดแยกบุคคลหรือกลุ่มบุคคลตามระดับสติปัญญาที่มีอยู่และที่เกิดขึ้นด้วยจุดมุ่งหมายของการวัดอย่างใดอย่างหนึ่ง เช่น เพื่อคัดเลือกบุคคลเข้าปฏิบัติหน้าที่เป็นทหารในกองทัพ และการคัดแยกความบกพร่องหรือความผิดปกติทางสติปัญญาของเด็ก

การวัดตามลักษณะดังกล่าวนี้เกิดขึ้นจากความเชื่อเกี่ยวกับความแตกต่างระหว่างบุคคลตามทฤษฎีชาวนิพนธ์นิยมที่เชื่อว่า สามารถถ่ายทอดทางพันธุกรรมได้ (Hereditarianism theory of IQ) ดังนั้น การวัดในช่วงแรกจึงเป็นการวัด IQ ด้วยแบบทดสอบมาตรฐาน (Standardized test) เช่น Stanford-Binet Scale, WISC-I และ WISC-II ภายใต้มาตรฐานการดำเนินงานทดสอบและการแปลความหมายของคะแนนจากการวัดค่าที่เป็นแบบเดียวกัน แนวคิดเกี่ยวกับการวัดและการทดสอบ IQ ในอดีตที่กระทำภายใต้การควบคุมปัจจัยหลักที่อาจก่อให้เกิดความลำเอียงหรือคลาดเคลื่อนอย่างเข้มงวดบนมาตรฐานเดียวกัน (Standardized

assessment) ที่ดำเนินการในวงแคบ (ระดับท้องถิ่น เช่น กลุ่มโรงเรียนหรือจังหวัด) และวงกว้าง (เช่น เขตการศึกษาหรือระดับประเทศ) โดยมีจุดมุ่งหมายหลักเพื่อกำกับติดตามและตรวจสอบพันธะภาระรับผิดชอบ (Accountability) ของบุคคลผู้ทำหน้าที่ในการจัดการศึกษาแต่ละระดับว่ามีคุณภาพหรือไม่

กล่าวโดยสรุปความเกี่ยวพันระหว่างหลักสูตร จิตวิทยา และการวัดที่เกิดขึ้นตามกระบวนการทัศน์แบบปฏิฐานนิยมที่เกิดขึ้นตลอดช่วงคริสต์ศตวรรษที่ 20 สามารถแสดงให้เห็นอย่างง่าย ๆ ตามภาพที่ 2 (Shepard, 2000: 6) ดังนี้



ภาพที่ 2 ความเกี่ยวพันระหว่างหลักสูตร จิตวิทยา และการวัดตามกระบวนการทัศน์แบบดั้งเดิม

2. ทัศนะใหม่ : หลักสูตร การเรียนรู้ และการวัดประเมิน ในห้วงเวลาที่ทำกิจกรรมสำคัญต่างๆ ทางการศึกษา เช่น การพัฒนาหลักสูตร จิตวิทยาการเรียนการสอน และการวัดผลการเรียนรู้ ได้รับอิทธิพลอย่างมากจากหลักการทางวิทยาศาสตร์ภายใต้การสนับสนุนจากนักวิชาการในกลุ่มทฤษฎีประสิทธิภาพทางสังคม หรือที่รู้จักกันใน

นาม “**นักวิทยาศาสตร์สังคมในชั้นเรียน**” (Classroom social-scientist) ได้มีนักวิชาการอีกกลุ่มหนึ่งที่ทำให้ความสนใจด้านวัฒนธรรมของสังคม หรือที่เรียกว่า “**นักวัฒนธรรมสังคมในชั้นเรียน**” (Classroom socio-culturalist) ค้นพบข้อจำกัดในการประยุกต์หลักการทางวิทยาศาสตร์มาใช้ในการจัดการทางการศึกษา โดยเฉพาะการจัดการ

เรียนการสอนภายในโรงเรียนและชั้นเรียน และไม่ได้มีทัศนะต่อ **“โรงเรียนคล้ายคลึงกับโรงงาน”** แต่กลับเชื่อว่าโรงเรียนเป็นองค์กรทางสังคม (Social organization) ที่ประกอบด้วยสมาชิก คือ นักเรียน ครู และผู้บริหารโรงเรียน เป็นสมาชิกสำคัญ และมีปฏิสัมพันธ์ระหว่างกันและกัน ภายใต้วัฒนธรรมและโครงสร้างทางสังคมของโรงเรียน รวมทั้งกฎระเบียบของชั้นเรียน นักวิชาการในกลุ่มนี้เชื่อว่าปฏิสัมพันธ์ระหว่างครูผู้สอนและผู้เรียน และระหว่างผู้เรียนด้วยกันภายใต้บรรยากาศแห่งความเป็นมิตรและเอื้ออาทรในโรงเรียนและชั้นเรียนเป็นปัจจัยสำคัญที่สนับสนุนให้เกิดการเรียนรู้ขึ้น

มุมมองของนักวัฒนธรรมสังคมในชั้นเรียนต่อการจัดการศึกษาที่กล่าวมา เกิดจากกระบวนทัศน์แบบตีความ/สรรค์สร้างนิยม (Interpretativism/Constructivism) (องอาจ นัยพัฒน์, 2544, 2551; Guba & Lincoln, 1989; Lincoln & Guba, 2000) ซึ่งมุ่งเน้นสรรค์สร้างความรู้ที่สัมพันธ์กับค่านิยมและวัฒนธรรมของชุมชนและสังคมหนึ่งๆ (Relativism) ทั้งนี้เพราะความรู้ความจริงแท้ (Ultimate reality) ที่กระบวนทัศน์แบบปฏิฐานนิยมแสวงหาไม่มีอยู่จริงหรือเป็นสากลอย่างแน่นอนตายตัว แต่ขึ้นอยู่กับความคิดเห็นและการยอมรับของสังคมหรือชุมชนเฉพาะในแต่ละถิ่นที่แต่ละเวลา ดังนั้น การบริหารจัดการศึกษาภายใต้แนวคิดตามกระบวนทัศน์แบบตีความ/สรรค์สร้างนิยม จึงมีลักษณะยืดหยุ่นปรับเปลี่ยนได้ตามความเหมาะสมและความเป็นจริงที่เกิดขึ้นในบริบทของการศึกษาหลักสูตรที่ดีตามทัศนะของนักวัฒนธรรมสังคมในชั้นเรียนจะต้องสอดคล้องกับความต้องการของผู้เรียน ชุมชนและสังคม นอกจากนี้จะต้องมีประโยชน์ต่อผู้เรียนเมื่อเรียนรู้แล้วสามารถนำความรู้ไปปฏิบัติได้ผลจริง (Pragmatic curriculum) ด้วยเหตุนี้จุดมุ่งหมายสำคัญของหลักสูตรจะต้องเน้นให้ผู้เรียนสามารถคิด

อย่างมีวิจารณญาณ มีการใช้พหุปัญญา (Multiple intelligence) ในการแก้ปัญหาที่ซับซ้อน และมีความสามารถประยุกต์ความรู้ไปใช้ในชีวิตประจำวันได้ด้วยเหตุนี้ มโนทัศน์ และวิสัยทัศน์ของการพัฒนาหลักสูตรตามแนวคิดดั้งเดิมจะต้องปฏิรูปใหม่ กล่าวคือ หลักสูตรจะต้องพัฒนาขึ้นด้วยความเชื่อเกี่ยวกับความเสมอภาคทางการศึกษาสำหรับผู้เรียนทุกคน นอกจากนี้ จะต้องเชื่อว่าผู้เรียนทุกคนสามารถเรียนรู้ได้ สารของหลักสูตรแต่ละรายวิชาจะต้องมีทั้งภาคทฤษฎีและปฏิบัติ มีความหมายต่อการดำเนินชีวิตประจำวันของผู้เรียน มีลักษณะท้าทายให้ผู้เรียนต้องคิดวิเคราะห์อย่างใช้วิจารณญาณ เพื่อแก้ปัญหาที่ซับซ้อน มีลักษณะมุ่งเน้นไปที่ผู้เรียนเป็นสำคัญ (Student-centered curriculum) ส่งเสริมให้ผู้เรียนมีปฏิสัมพันธ์ระหว่างกันทั้งในและนอกโรงเรียน/ชั้นเรียน สามารถทำงานเป็นทีมได้ และส่งเสริมให้ผู้เรียนมีคุณธรรม จริยธรรม รวมทั้งมีการแสดงออกที่เหมาะสมภายใต้การปกครองระบอบประชาธิปไตย ลักษณะสำคัญของหลักสูตรตามกระบวนทัศน์แบบสรรค์สร้างนิยมที่กล่าวถึงโดยสังเขปนี้ สอดคล้องกับคุณลักษณะของผู้เรียนที่ระบุไว้ในหลักสูตรการศึกษาขั้นพื้นฐาน พ.ศ. 2544 ของประเทศไทย (กระทรวงศึกษาธิการ, 2544)

สำหรับจิตวิทยาตามกระบวนทัศน์แบบสรรค์สร้างนิยมมุ่งเน้นไปที่ทฤษฎีการรู้คิด (Cognitive theory) กล่าวคือ เชื่อว่าการเรียนรู้เป็นกระบวนการทางสมองของมนุษย์ที่คิดและสืบค้นหาเหตุผลด้วยความใคร่ครวญและใช้วิจารณญาณอย่างไม่อยู่นิ่ง (Active process) มีการเพิ่มพูนความรู้ขึ้นแบบค่อยเป็นค่อยไปด้วยการกระตุ้นและการตอบสนองที่เชื่อมโยงถึงกัน ความสามารถทางสติปัญญาพัฒนาขึ้นจากปฏิสัมพันธ์ทางสังคมที่มีผู้เรียนเป็นผู้สร้างความรู้และทำความเข้าใจเกี่ยวกับสิ่งที่สนใจใคร่รู้ภายในบริบททางสังคมและวัฒนธรรมของชุมชน/



สังคม รวมทั้งเชื่อว่าความสามารถทางสติปัญญา เกี่ยวข้องกับการคิดเชิงอภิमान (Meta-cognition) กล่าวคือ มีการกำกับติดตามและตรวจสอบการคิด รวมทั้งประเมินการเรียนรู้ของตนเองอย่างต่อเนื่อง หรือที่เรียกว่า “การคิดเกี่ยวกับการคิด” (Cognition about cognition) การเปลี่ยนแปลงด้านหลักสูตร และทฤษฎีการเรียนรู้ตามทัศนะใหม่บนรากฐาน ความเชื่อจากกระบวนทัศน์แบบสรรค์สร้างนิยมเป็น แรงผลักดันสำคัญยิ่งที่ทำให้แนวคิด หลักการและ วิธีการวัดประเมินการเรียนรู้จำเป็นต้องปฏิรูป หรือเปลี่ยนแปลงไป นอกเหนือจากการทดสอบ วัดผลสัมฤทธิ์ท้ายบทเรียน หรือปลายภาคเรียน ด้วยแบบทดสอบปรนัยดังที่ปฏิบัติมาในอดีตเพียง อย่างเดียว ซึ่งพบว่าการวัดประเมินการเรียนรู้ดังกล่าวนี้ได้ก่อให้เกิดปัญหาสำคัญหลายประการ อาทิ (Black & William, 1998; Stiggins, 2002, 2005)

1. การวัดประเมินการเรียนรู้ส่วนใหญ่ กระทำอย่างผิวเผินแค่เพียงการวัดความรู้เฉพาะ รายละเอียดที่เป็นประเด็นปลีกย่อย มีประโยชน์และความหมายต่อการดำเนินชีวิตประจำวันน้อย และง่ายต่อการลืมเลือน

2. ครูผู้สอนส่วนใหญ่กำหนดจุดมุ่งหมาย การวัดประเมินการเรียนรู้ไม่ชัดเจน ไม่ทราบว่าจะทำการวัดประเมินการเรียนรู้อะไร และกระทำไป เพื่ออะไร

3. แม้ว่าการสร้างเครื่องมือหรือพัฒนา วิธีการวัดค่าคะแนนที่ซับซ้อนใดๆ เพื่อใช้แสดงผลการเรียนรู้ของผู้เรียนได้อย่างเที่ยงตรงและ เชื่อมั่นได้จะถูกให้ความสำคัญอย่างมากนับตั้งแต่ อดีตจนถึงปัจจุบัน แต่การนำคะแนนจากการวัดค่า อย่างเที่ยงตรงและเชื่อมั่นได้นั้นไปใช้ส่งเสริมผู้เรียน ให้ต้องการเรียนรู้หรือรู้สึกว่าจะสามารถเรียนรู้ได้กลับ ได้รับความเอาใจใส่อย่างมากจากครูและผู้เกี่ยวข้อง ทั้งๆ ที่มีความสำคัญยิ่ง

4. การวัดประเมินการเรียนรู้มักกระทำ ในลักษณะแยกส่วนกับกิจกรรมการเรียนการสอน และการวิจัยและพัฒนาการเรียนรู้ของผู้เรียน จะเห็นได้จากการวัดประเมินมักกระทำอย่างเป็นทางการภายหลังการสอน จึงเป็นผลให้ต้องใช้เวลา ในการวัดประเมินมาก (แทนที่จะใช้ในการสอน ของครูและพัฒนาการเรียนรู้ของผู้เรียน) และไม่ สอดคล้องกลมกลืนกับบริบทที่เกิดขึ้นในชั้นเรียน

5. การวัดประเมินการเรียนรู้ โดยเฉพาะ ในระดับอุดมศึกษามักเป็นแบบอิงกลุ่มที่เน้น การแข่งขันระหว่างผู้เรียนด้วยกัน มากกว่าแบบ อิงเกณฑ์ที่เน้นพัฒนาการเรียนรู้ของผู้เรียนแต่ละคน ด้วยการส่งเสริมให้ผู้เรียนแข่งขันกับตนเอง การวัด ประเมินแบบอิงกลุ่มทำให้ผู้เรียนที่มีความรู้ ความสามารถต่ำขาดแรงจูงใจใฝ่สัมฤทธิ์ รวมทั้ง สูญเสียความเชื่อมั่นและความภาคภูมิใจในตนเอง ที่สามารถเรียนรู้ได้

6. ครูเน้นการวัดประเมินความรู้ความคิด โดยใช้แบบทดสอบหรือมาตรวัดเชิงปริมาณ ภายใต การดำเนินการอย่างเป็นทางการบนมาตรฐาน เดียวกันมากกว่าการวัดประเมินที่เกิดขึ้นตามสภาพ จริง (Authentic assessment) ว่าผู้เรียนมีความรู้ ความสามารถ ความสนใจ และทักษะการปฏิบัติงาน ได้อย่างแท้จริงหรือไม่

7. ครูผู้สอนยึดติดอยู่กับความเชื่อ เกี่ยวกับความเป็นปรนัยและความมาตรฐาน เดียวกันของการวัดประเมินมากเสียจนกระทั่ง ไม่ยอมรับหรือมองข้ามวิธีการวัดประเมินแบบอื่นๆ เช่น การใช้คำถามด้วยวจาการวัดประเมินการเรียนรู้ ระหว่างการเรียนการสอน หรือการวัดประเมินจาก กลุ่มเพื่อนผู้เรียนด้วยกัน (Peer assessment) จึงทำให้บางครั้งการวัดประเมินผลลัพธ์การเรียนรู้ที่ ต้องการทราบไม่ครบถ้วนสมบูรณ์ว่าเป็นอย่างไร

8. ในปัจจุบันนี้ผู้บริหารโรงเรียนและครูมักเอาใจใส่มากในเรื่องการวัดประเมินเชิงกำกับติดตามและตรวจสอบคุณภาพการศึกษาที่ดำเนินการโดยหน่วยงานรัฐบาล เช่น สถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.) ที่ใช้คะแนนเฉลี่ยจากการวัดผลสัมฤทธิ์ทางการเรียนเป็นเครื่องบ่งชี้ความสำเร็จในการจัดการเรียนรู้ในโรงเรียน (Assessment of learning) ส่วนการวัดประเมินในชั้นเรียนที่มีจุดมุ่งหมายเพื่อพัฒนาการเรียนรู้อของผู้เรียน (Assessment for learning) ที่สำคัญยิ่งต่อการช่วยให้ผู้เรียนมีความต้องการเรียนรู้หรือรู้สึกว่าคุณเองสามารถเรียนรู้ได้กลับไม่ได้รับการเอาใจใส่เท่าที่ควร

ปัญหาสำคัญในการประเมินการเรียนรู้ที่กล่าวมาทั้งหมดเป็น **“ปัญหาสากล”** เกิดขึ้นในบริบททางการศึกษาของหลายประเทศ (รวมทั้งประเทศไทย) ดังนั้น จึงจำเป็นอย่างยิ่งที่แนวคิดและวิธีการวัดประเมินการเรียนรู้จะต้องมีการปฏิรูปขึ้นใหม่เพื่อแก้ไขปัญหาข้างต้น รวมทั้งจะต้องปรับเปลี่ยนให้สอดคล้องหลักสูตรและทฤษฎีจิตวิทยาการเรียนรู้และสติปัญญาด้านการรู้คิดตามทัศนะใหม่ที่เกี่ยวข้อเชื่อมโยงถึงกันและบูรณาการทั้งแนวคิดและวิธีปฏิบัติเข้าด้วยกัน (เส้นประในภาพที่ 3 แสดงความเชื่อมโยงนี้) แทนที่จะไม่ยึดหยุ่นและแยกส่วนดำเนินการดังเช่นการปฏิบัติที่ผ่านมา (เส้นทึบในภาพที่ 2 แสดงการแยกส่วนนี้) (Shepard, 2000 : 8) ปัจจุบันนี้นักวิชาการด้านการวัดประเมินทางการศึกษาจำนวนมาก (e.g. Stiggins, 2004, 2005; Wiggins, 1993) กำลังให้ความสนใจ และกระตุ้นให้ครู ศึกษานิเทศก์ และผู้บริหารโรงเรียนนำแนวคิดและวิธีปฏิบัติเกี่ยวกับ **“การวัดประเมินเพื่อพัฒนาการเรียนรู้อ”** มาใช้ในชั้นเรียนมากขึ้น แทนมุ่งเน้นแต่เฉพาะ **“การวัด**

**ประเมินเพื่อตรวจสอบการเรียนรู้อ”** ที่ดำเนินการตามข้อกำหนดของหน่วยงานทางการศึกษาจากส่วนกลางเพียงอย่างเดียว แนวคิดและการปฏิบัติเชิงรุกด้านการวัดประเมินเพื่อพัฒนาการเรียนรู้อในชั้นเรียนที่ดำเนินการอย่างต่อเนื่องและเข้มแข็งจริงจังของ Stiggins (2002, 2004, 2005) ปรากฏว่าผู้บริหารของสำนักงานบริการทดสอบทางการศึกษา (Educational Testing Service or ETS) ประเทศสหรัฐอเมริกา ให้ความสนใจมากเป็นพิเศษ จนกระทั่งในปี ค.ศ. 2006 ได้เชิญ Stiggins ผู้อำนวยการสถาบันฝึกอบรมการวัดประเมิน (Assessment Training Institute or ATI) ที่ก่อตั้งขึ้นในปี ค.ศ.1992 ให้เข้าร่วมงานกับ ETS ด้วย ([www.assessmentinst.com/about-ets/about-ets-ati](http://www.assessmentinst.com/about-ets/about-ets-ati)) และนี่คือ **“จุดเปลี่ยนสำคัญ”** ของการวัดประเมินที่ ETS องค์กรการวัดประเมินระดับโลกหันมาส่งเสริมแนวคิดด้านการวัดประเมินเพื่อพัฒนาการเรียนรู้อในชั้นเรียน นอกเหนือจากการดำเนินการตามพันธกิจหลักด้านการวัดประเมินเพื่อตรวจสอบการเรียนรู้อในวงกว้าง (เช่น การวัดประเมินด้วยแบบทดสอบ SAT, TOEFL, GMAT และ GRE) ที่ได้กระทำมาอย่างยาวนาน ซึ่งผู้เขียนจะได้กล่าวรายละเอียดในโอกาสต่อไป



**ภาพที่ 3** แนวคิดและหลักการร่วมกันของทฤษฎีหลักสูตร ทฤษฎีจิตวิทยาการเรียนรู้และสติปัญญา และการวัดประเมินการเรียนรู้ตามกระบวนการทัศน์แบบสรรค์สร้างนิยม

### บรรณานุกรม

- กระทรวงศึกษาธิการ. (2544). **หลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544**. กรุงเทพฯ: โรงพิมพ์คุรุสภาลาดพร้าว.
- องอาจ นัยพัฒน์. (2551). การออกแบบการวิจัย: วิธีการเชิงปริมาณ เชิงคุณภาพ และผสมผสานวิธีการ. กรุงเทพฯ: โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- องอาจ นัยพัฒน์. (2543). กระบวนทัศน์หลักและแนวคิดสำคัญเกี่ยวกับการวิจัยและการประเมินทางการศึกษา. **วารสารการวัดผลการศึกษา**, 22(65), 1-20.
- Black, P. & William, D. (1998). Assessment and classroom learning. **Assessment in Education: Principles, Policy, and Practice**, 5(1) 7-74.
- Guba, E.G., & Lincoln, Y.S. (1989). **Fourth generation evaluation**. Newbury Park, CA: Sage.
- Lincoln, Y.S., & Guba, E.G. (2000). Paradigmatic controversies, contradictions, and emerging confluences. In N.K. Denzin & Y.S. Lincoln (Eds.), **Handbook of qualitative research** (2nd ed.) (pp. 163-188). Thousand Oaks, CA: Sage.
- Shepard, L.A. (2000). **The role of assessment in a learning culture**, **Educational Researchers**, 29(7), 4-14.
- Stiggins, R.J. (2005). From formative assessment to assessment for learning: A part to success in standard-based schools, **The Phi Delta Kappan**, 87(4), 324-328.
- Stiggins, R.J. (2004). **Student-Involved Assessment FOR Learning** (4th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Stiggins, R.J. (2002). Assessment crisis: The absence of assessment for learning. **The Phi Delta Kappan**, 83(10), 758-765.
- Wiggins, G. (1993). **Educative Assessment: Designing assessments to inform and improve student performance**. San Francisco, CA: Jossey-Bass.