

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยเรื่องการพัฒนาเกณฑ์การให้คะแนนผังมโนทัศน์สำหรับประเมินผลการเรียนรู้ของนักศึกษา มีวัตถุประสงค์เพื่อ พัฒนาเกณฑ์การให้คะแนนผังมโนทัศน์สำหรับประเมินผลการเรียนรู้ของนักศึกษา วิเคราะห์ลักษณะของเกณฑ์การให้คะแนนผังมโนทัศน์สำหรับประเมินผลการเรียนรู้ของนักศึกษา โดยใช้ทฤษฎีการสรุปอ้างอิง และเปรียบเทียบค่าความตรงตามสภาพของคะแนนผังมโนทัศน์ เมื่อใช้เกณฑ์การให้คะแนนต่างกัน ผู้วิจัยนำเสนอเอกสารและงานวิจัยที่เกี่ยวข้อง เป็น 5 ตอน ดังนี้

ตอนที่ 1 แนวคิดเกี่ยวกับการวัดและการประเมินผลการเรียนรู้

ตอนที่ 2 แนวคิดและทฤษฎีเกี่ยวกับผังมโนทัศน์

ตอนที่ 3 แนวคิดเกี่ยวกับเกณฑ์การให้คะแนน

ตอนที่ 4 ทฤษฎีการสรุปอ้างอิง

ตอนที่ 5 งานวิจัยที่เกี่ยวข้องกับการใช้ผังมโนทัศน์สำหรับการประเมินผลการเรียนรู้ และทฤษฎีการสรุปอ้างอิง

ตอนที่ 1 แนวคิดเกี่ยวกับการวัดและการประเมินผลการเรียนรู้

ความหมายของการวัดผล

เยาวดี วิบูลย์ศรี (2548, หน้า 5) กล่าวว่า โดยทั่วไป “การวัดผล” หมายถึง กระบวนการ บังชี้ผลผลิตหรือลักษณะที่วัดได้จากเครื่องมือวัดผลประเภทใดประเภทหนึ่งอย่างมีระบบ ดัชนียามที่ว่า “การวัดผล” คือการกำหนดตัวเลขให้กับสิ่งใดสิ่งหนึ่งตามกฎเกณฑ์ที่ตั้งไว้

ศิริชัย กาญจนวาสี (2552, หน้า 9) กล่าวว่า การวัด เป็นกระบวนการกำหนดตัวเลขให้กับ สิ่งของ เหตุการณ์ หรือบุคคลตามกฎเกณฑ์ โดยใช้เครื่องมือช่วยในการกำหนดตัวเลขตาม วัตถุประสงค์ของการวัด การวัดต้องอาศัยองค์ประกอบที่สำคัญ 3 ส่วน ได้แก่ จุดมุ่งหมายของการวัด เครื่องมือที่ใช้วัด และการแปลผลจากการวัด

สิริจิตต์ เดชอมรชัย (2555, หน้า 1) กล่าวว่า “Measurement” หรือ การวัดเป็น กระบวนการ (Procedure) สืบค้น ตรวจสอบรายละเอียด หรือหาค่า กำหนดจำนวนปริมาณของสิ่ง หนึ่งสิ่งใด โดยอาศัยเครื่องมือหรือวิธีการต่าง ๆ เช่น เครื่องมือทางกายภาพ หรือใช้เครื่องมือและ วิธีการทางสังคมศาสตร์ เช่น แบบสังเกต แบบสอบถาม แบบสำรวจ การสัมภาษณ์ ข้อสอบ เป็นต้น

Pedhazur and Schmelkin (1991) กล่าวว่า การวัด หมายถึง การกำหนดตัวเลขให้กับ วัตถุหรือเหตุการณ์

Ferris (2004) กล่าวว่า การวัด หมายถึง กระบวนการเชิงประจักษ์ โดยการใช้เครื่องมือ แปลงสภาพที่เป็นตามธรรมชาติของสิ่งที่สังเกตได้เข้าสู่หมวดหมู่ของโมเดลที่สังเกตได้ เพื่อก่อให้เกิด ความถูกต้อง ซึ่งความหมายจะแตกต่างกันเห็นได้ชัดจากสิ่งที่เป็นไปได้อื่น ๆ และสิ่งที่แตกต่างกัน กระบวนการวัดต้องเป็นอิสระจากปัจจัยที่เกี่ยวข้องกับความคิดเห็น ดังนั้น การวัดควรทำการวัดซ้ำ โดยใช้เครื่องมือเดียวกัน เพื่อทำการแปลงในสภาวะเดียวกัน

ดังนั้นจึงสรุปได้ว่า การวัด เป็นกระบวนการของการกำหนดตัวเลขให้กับสิ่งต่าง ๆ โดยมี หลักการกำหนดตัวเลขที่ชัดเจน ด้วยเทคนิคหรือวิธีการที่ได้รับการยอมรับ

ความหมายของการประเมินผล

การประเมินในภาษาอังกฤษนั้นพบว่ามีคำที่ใช้อยู่ 2 คำ คือ คำว่า “Evaluation” และ คำว่า “Assessment”

เยาวดี วิบูลย์ศรี (2546) กล่าวว่า Evaluation หมายถึง กระบวนการตัดสินคุณค่าของสิ่ง หนึ่งสิ่งใดโดยนำสารสนเทศหรือผลจากการวัดมาเปรียบเทียบกับเกณฑ์ที่กำหนดไว้ ส่วนคำว่า Assessment จะเป็นการประเมินเบื้องต้น ในขณะที่ Evaluation เป็นกระบวนการตัดสินใจหรือ ตัดสินคุณค่าขั้นสุดท้าย เช่น Assessment ใช้เพื่อหาคำตอบว่านักเรียนเป็นอย่างไร ได้รับความสำเร็จ ด้านไหน ส่วน Evaluation ใช้หาคำตอบถึงผลและประสิทธิผลการจัดการเรียนการสอน เป็นต้น จึง อาจกล่าวได้ว่า Assessment เป็นส่วนหนึ่งของ Evaluation

แนวคิดดังกล่าวสอดคล้องกับ ทิศนา ขัมมณี, พิมพันธ์ เดชะคุปต์ และชนาธิป พรกุล (2545) ที่ได้กล่าวว่า Evaluation หรือ การประเมินผล เป็นกระบวนการตีความหรือตัดสินคุณค่าของ สารสนเทศที่รวบรวมมาจากการประเมิน โดยมาประมวลเข้าด้วยกันเพื่อตัดสินว่าได้สะท้อนคุณค่า ในตัวผู้เรียนที่ตั้งไว้หรือไม่ มากน้อยเพียงใด ในการพิจารณาตีค่าหรือสรุปผลก็โดยอาศัยผลการวัดและ เกณฑ์ (Criteria) เช่น อิงกลุ่ม อิงเกณฑ์ อิงพัฒนาการ หรือ แบบผสม การประเมินผลนำไปสู่การ ดำเนินการเกี่ยวกับสิ่งที่ประเมินในทางใดทางหนึ่งต่อไป เช่น เพื่อปรับปรุง เปลี่ยนแปลง แก้ไข หรือ พัฒนาสิ่งต่าง ๆ ส่วนคำว่า Assessment หรือ การประเมิน เป็นคำที่นำมาใช้ในความหมายของการ ประเมินแนวใหม่เพื่อให้แตกต่างไปจากแนวคิดเดิม หมายถึงกระบวนการรวบรวมข้อมูลเกี่ยวกับ ความสามารถของผู้เรียนด้านต่าง ๆ ทั้งในเชิงปริมาณและเชิงคุณลักษณะและนำมาเรียบเรียงอย่าง เป็นระบบและป้อนกลับให้แก่ผู้เรียนหรือครูนำไปใช้ในการตัดสินใจเกี่ยวกับผู้เรียนหรือตัดสิน ประสิทธิภาพในการจัดการเรียนรู้ให้แก่ผู้เรียน

ประเภทของการวัดและการประเมินผล

การวัดและการประเมินผล สามารถจำแนกประเภทได้ดังนี้ (พิชิต ฤทธิ์จรูญ, 2548)

1. จำแนกตามวัตถุประสงค์ของการประเมิน แบ่งเป็น 3 ประเภทคือ

1.1 การประเมินผลก่อนเรียน (Pre-Evaluation) มีจุดมุ่งหมายเพื่อตรวจสอบความรู้พื้นฐานและทักษะของผู้เรียนว่า มีความรู้เพียงพอที่จะเรียนต่อในรายวิชาใหม่หรือเนื้อหาใหม่ได้หรือไม่ ถ้าพบว่ามีความรู้พื้นฐานไม่เพียงพอหรือไม่มีความพร้อมขั้นต้นก่อนเรียนครูจะจัดให้มีการสอนปรับพื้นฐานจนผู้เรียนมีความรู้เพียงพอที่จะเรียนในเนื้อหาใหม่ได้

การสอบก่อนเรียนไม่ใช่การสอบวัดผลสัมฤทธิ์ (Achievement Test) เพราะครูยังไม่ได้ทำการสอนในเนื้อหาเหล่านั้นมาก่อน แต่เป็นการสอบเพื่อวินิจฉัย (Diagnostic Test) ซึ่งนอกจากจะช่วยให้ครูทราบพื้นฐานของผู้เรียนแล้ว ยังช่วยให้ครูวางแผนการสอนได้เหมาะสมกับสภาพของผู้เรียน และเป็นข้อมูลประกอบการตัดสินใจเลือกวิธีการสอนและมอบหมายภาระงานการเรียนรู้ให้กับผู้เรียน

1.2 การประเมินผลระหว่างเรียนหรือประเมินความก้าวหน้า (Formative Evaluation) มีจุดมุ่งหมายเพื่อตรวจสอบว่าผู้เรียนบรรลุตามวัตถุประสงค์การเรียนรู้ที่กำหนดไว้หรือไม่เพียงใด หากพบว่ามีข้อบกพร่องในจุดประสงค์ใด ก็หาแนวทางปรับปรุง แก้ไขข้อบกพร่องในจุดประสงค์นั้น ๆ โดยจัดสอนซ่อมเสริมให้แก่ผู้เรียน

การประเมินผลระหว่างเรียนเป็นการสอบย่อย (Formative Test) ในเนื้อหาที่สอนเท่านั้นเพื่อตรวจสอบดูความก้าวหน้าของการเรียน ดูว่าผู้เรียนบรรลุจุดประสงค์การเรียนรู้ที่กำหนดไว้หรือไม่ อันจะนำไปสู่การสอนซ่อมเสริมและปรับปรุงการสอนของครูอีกด้วย

1.3 การประเมินผลรวมสรุป (Summative Evaluation) เป็นการประเมินเพื่อตัดสินผลการเรียนมีจุดมุ่งหมายเพื่อศึกษาว่าผู้เรียนมีความรู้ทั้งสิ้นเท่าไร ควรตัดสินได้-ตก ผ่าน-ไม่ผ่าน หรือควรได้เกรดอะไร เป็นต้น

การประเมินผลรวมสรุปเป็นการประเมินเมื่อสิ้นสุดการเรียนการสอนของแต่ละรายวิชา ครูจำเป็นต้องประเมินให้ครอบคลุมทุกจุดประสงค์ หากมีจุดประสงค์มากครูอาจต้องเลือกประเมินบางจุดประสงค์โดยสุ่มเอาเฉพาะจุดประสงค์ที่สำคัญ ๆ ก็ได้

2. จำแนกตามระบบการวัด แบ่งออกเป็น 2 ประเภท คือ

2.1 การประเมินผลแบบอิงกลุ่ม (Norm-Referenced Evaluation) เป็นการตัดสินคุณค่าของคุณลักษณะหรือพฤติกรรม โดยเปรียบเทียบกับผู้เรียนที่อยู่ในกลุ่มเดียวกันที่ทำข้อสอบฉบับเดียวกันโดยมีจุดมุ่งหมายเพื่อจำแนกหรือจัดลำดับบุคคลในกลุ่มนั้น ๆ ตัวอย่างของการประเมินแบบอิงกลุ่ม เช่น การสอบคัดเลือกเข้าศึกษาต่อในสถานศึกษา การสอบชิงทุนการศึกษา เป็นต้น

2.2 การประเมินผลแบบอิงเกณฑ์ (Criterion- Referenced Evaluation) เป็นการ

ตัดสินคุณค่าของคุณลักษณะหรือพฤติกรรมโดยเปรียบเทียบกับเกณฑ์ ซึ่งเกณฑ์มีทั้งเกณฑ์มาตรฐาน (Standard Criteria) ที่มีอยู่แล้ว หรือเกณฑ์ที่ผู้ประเมินกำหนดขึ้น (Arbitrary Criteria) ในทางปฏิบัติการประเมินผลสัมฤทธิ์ทางการเรียน เกณฑ์จะหมายถึงกลุ่มพฤติกรรมตามจุดมุ่งหมายในแต่ละบทหรือหน่วยการเรียนรู้โดยทั่วไป นิยมใช้จุดประสงค์เชิงพฤติกรรม (Behavioral Objective) หรือ กลุ่มของพฤติกรรม (Domain of Behavior)

การประเมินผลแบบอิงเกณฑ์มีจุดมุ่งหมายเพื่อบ่งชี้สถานภาพของผู้เรียนแต่ละคนเมื่อเปรียบเทียบกับเกณฑ์ที่ทดสอบเพื่อตัดสินว่าผู้เรียนบรรลุตามวัตถุประสงค์ที่กำหนดไว้หรือไม่และมากน้อยเพียงใดอันจะนำไปสู่การปรับปรุงการเรียนการสอนเมื่อผู้เรียนไม่สามารถทำข้อสอบได้ถึงเกณฑ์ ต้องมีการสอนซ่อมเสริมจนกว่าจะผ่านถึงเกณฑ์ การประเมินผลแบบอิงเกณฑ์จึงเหมาะสำหรับการเรียนการสอนในห้องเรียน

ความมุ่งหมายและประโยชน์ของการวัดและการประเมินผล

ความมุ่งหมายและประโยชน์ของการวัดผลและการประเมินผล มีหลายประการ ดังนี้ (พิชิต ฤทธิ์จรรยา, 2548)

1. เพื่อสร้างแรงจูงใจในการเรียน

ผลการประเมินโดยเฉพาะการประเมินผลย่อย หรือการประเมินผลระหว่างการเรียนการสอนสามารถนำมาใช้เพิ่มประสิทธิภาพการเรียนรู้ให้สูงขึ้นได้ โดยนำผลมาแจ้งให้ผู้เรียนทราบเพื่อรู้หรือกระตุ้นให้ผู้เรียนเกิดความอยากรู้อยากเห็น

2. เพื่อตรวจสอบความรู้พื้นฐาน

การตรวจสอบความรู้พื้นฐานเป็นการประเมินเพื่อตรวจสอบความรู้ ความสามารถและทักษะพื้นฐานของผู้เรียนก่อนที่จะเริ่มเรียนบทเรียนใหม่ เนื่องจากเนื้อหาบางบทเรียนอาจจำเป็นต้องใช้ความรู้หรือทักษะพื้นฐานบางประการ ครูจึงต้องตรวจสอบก่อนว่าผู้เรียนมีความรู้หรือทักษะพื้นฐานเพียงใด หากพบว่าไม่มีความรู้ หรือมีความรู้ไม่เพียงพอในเรื่องใดก็จะได้ปรับพื้นฐานเสียก่อน การตรวจสอบความรู้พื้นฐานอาจทำได้โดยใช้แบบทดสอบหรือใช้การซักถามก็ได้

3. เพื่อปรับปรุงการเรียนการสอน

ผลการประเมินทั้งก่อนการเรียนการสอน และระหว่างการเรียนการสอนสามารถนำมาปรับปรุงการเรียนการสอนได้โดยนำมาพิจารณาว่า ผลการสอนไม่ทำให้ผู้เรียนเกิดการเรียนรู้ตามวัตถุประสงค์อาจเป็นเพราะว่าวิธีสอน และกิจกรรมที่ครูผู้สอนกำหนดไว้อาจไม่เหมาะสมกับผู้เรียน หรือเพราะสื่อการเรียนการสอนไม่เหมาะสม นอกจากครูจะนำผลการประเมิน มาปรับปรุงการสอนในครั้งต่อไปแล้วผู้เรียนก็ยังได้ทราบสถานภาพการเรียนรู้ของตนเองว่า อยู่ในระดับใด หากอยู่ในระดับไม่น่าพอใจจะได้ปรับปรุงการเรียนของตนเองให้ดีขึ้น

4. เพื่อวินิจฉัยข้อบกพร่อง

การวินิจฉัยข้อบกพร่องเป็นการวิเคราะห์หาสาเหตุข้อบกพร่องในการเรียนของผู้เรียน และการสอนของครู เช่น ใช้แบบทดสอบวินิจฉัยความบกพร่องในการอ่านภาษาไทยที่มีเนื้อหาทั้งการอ่านคำควบกล้ำ ร ล ว การอ่านตัวสะกดการันต์ และการอ่านจับใจความทดสอบกับผู้เรียน ถ้าตรวจให้คะแนนแล้วพบว่าผู้เรียนคนใดทำผิดเรื่องใดก็จะช่วยให้ครูผู้สอนได้รู้ข้อบกพร่องของผู้เรียนเพื่อสอนซ่อมเสริมหรือทบทวนในเรื่องนั้น ๆ ได้ รวมทั้งเป็นข้อมูลให้ครูได้พิจารณาทบทวนการสอนของตนเองว่าบกพร่องอย่างไรแล้วหาทางปรับปรุงต่อไป

5. เพื่อตัดสินผลการเรียน

ผลการประเมินโดยเฉพาะการประเมินเมื่อสิ้นสุดการเรียนการสอน หรือการประเมินผลรวมเพื่อสรุปคุณภาพของการจัดการเรียนการสอนว่าผู้เรียนบรรลุตามจุดประสงค์ของรายวิชา หรือในระดับชั้นนั้นเพียงใดโดยนำผลการวัดทั้งหมดตลอดภาคเรียนหรือตลอดปีมาเป็นข้อมูลในการตัดสินผลขั้นสุดท้ายในรูปของการให้เกรด เป็น 4-3-2-1-0 หรือ A-B-C-D-E เป็นต้น หรือในรูปของระดับคุณภาพ เช่น ผ่านยอดเยี่ยม-ผ่าน-ไม่ผ่าน หรือ สอบได้-สอบตก หรือในรูปเกรดเฉลี่ย เป็นต้น

6. เพื่อจัดตำแหน่งหรือจัดประเภท

การจัดตำแหน่งหรือการจัดประเภทเป็นการนำผลที่ได้จากการวัดมาเปรียบเทียบหรือจัดอันดับความสามารถในกลุ่มเดียวกันเพื่อจะตอบว่ามีความรู้ความสามารถอยู่ในระดับใดของกลุ่ม เช่น สอบได้อันดับที่เท่าไร ได้เกรดอะไร ในการจัดตำแหน่งเหล่านี้แบ่งได้เป็น 2 ลักษณะ คือ

6.1 เพื่อการจำแนก (Classification) ความสามารถของผู้เรียน เช่น สอบได้-สอบตก เมื่อเปรียบเทียบกับเกณฑ์หรือเพื่อจำแนกผู้เรียนออกเป็นกลุ่มสูง-ปานกลาง-ต่ำ เป็นต้น

6.2 เพื่อการคัดเลือก (Selection) บุคคลเข้าทำงาน หรือศึกษาต่อ ซึ่งการประเมินลักษณะนี้ต้องใช้เครื่องมือที่ต้องการความยาก และความลึก

7. เพื่อเปรียบเทียบระดับพัฒนาการ

การเปรียบเทียบระดับพัฒนาการเป็นการสอบหรือตรวจวัดดูว่าผู้เรียนมีพัฒนาการเพียงใด มีความงอกงาม หรือเปลี่ยนแปลงไปในทางที่ดีขึ้นหรือไม่ โดยเปรียบเทียบความสามารถของผู้เรียนคนเดียวหรือกลุ่มเดียวกันในระยะเวลาที่ต่างกัน เช่น ก่อนเรียนกับหลังเรียน ภาคเรียนที่ 1 กับภาคเรียนที่ 2 เป็นต้น ถ้าผลการเรียนต่ำลงก็จะเป็นข้อมูลชี้ว่าผู้เรียนคนนั้นหรือกลุ่มนั้นควรจะต้องปรับปรุง และครูผู้สอนรายวิชานั้นควรจะต้องทบทวนการจัดการเรียนการสอนของตนเอง ขณะเดียวกันหากผลการเรียนมีพัฒนาการดีขึ้นก็เป็นการสร้างกำลังใจให้แก่ผู้เรียนคนนั้นหรือกลุ่มนั้นต่อไป

8. เพื่อพยากรณ์หรือทำนาย

การพยากรณ์เป็นการนำผลจากการวัดในปัจจุบันไปทำนายในอนาคตว่าผู้เรียนคนนี้จะ

เรียนวิชานั้น ๆ สำเร็จหรือไม่ หรือในอนาคตเขาควรจะเรียนอะไรจึงจะดี การประเมินลักษณะนี้มักจะนำไปใช้ในการแนะนำหรือสอบคัดเลือกเพื่อศึกษาต่อซึ่งอาจใช้แบบทดสอบวัดความถนัด (Aptitude Test) หรือผลการเรียนในอดีตมาทำนายก็ได้ อย่างไรก็ตาม ข้อที่ควรตระหนักของการพยากรณ์ก็คือแบบทดสอบที่ใช้ต้องมีคุณภาพ กล่าวคือ มีความตรง และความเที่ยงสูง เพราะแบบทดสอบที่มีคุณภาพต่ำผลการพยากรณ์จะไม่ตรงกับความเป็นจริง

9. เพื่อประเมินค่า

การประเมินค่าเป็นการประเมินที่มุ่งสรุปคุณภาพการศึกษาโดยส่วนรวม เช่น ดูความเหมาะสมของหลักสูตรกับโรงเรียน ดูการจัดบริการในโรงเรียนว่าเหมาะสมสอดคล้องกับความต้องการของบุคลากรในโรงเรียนหรือไม่ ดูแนวโน้มการจัดบริการทางการศึกษาของโรงเรียนในอนาคต เป็นต้น ซึ่งการประเมินลักษณะนี้จะเป็นประโยชน์สำหรับผู้บริหารให้พัฒนาคุณภาพการศึกษาได้เป็นอย่างดี

10. ประโยชน์ของการวัดและประเมินผล

ประโยชน์ของการวัดและประเมินผลที่สำคัญคือมุ่งนำผลการประเมินมาปรับปรุงและพัฒนาคุณภาพการศึกษา ซึ่งผลจากการประเมินจะคุ้มค่ามากน้อยเพียงใดขึ้นอยู่กับผู้เกี่ยวข้องจะนำไปใช้กับใคร ในกิจการใด และอย่างไรบ้างซึ่งในที่นี้จะพิจารณาประโยชน์ของการประเมินดังนี้

10.1 ประโยชน์ต่อผู้เรียน

10.1.1 ทำให้เกิดการพัฒนาดตนเองในแนวทางที่เหมาะสมยิ่งขึ้น โดยพิจารณาจากผลการประเมินความสามารถของตนเอง

10.1.2 ทำให้เกิดแรงจูงใจในการเรียนเพิ่มขึ้นเพื่อที่จะรักษามาตรฐานไว้ และช่วยให้ผลการเรียนดีขึ้นตามลำดับ

10.1.3 ทำให้เกิดความเข้าใจในเนื้อหาที่เรียนชัดเจนยิ่งขึ้น เพราะการสอบแต่ละครั้งทำให้ผู้เรียนต้องอ่านหนังสือทบทวนเนื้อหา มีการค้นคว้าหาความรู้เพิ่มเติมทำให้เกิดการเรียนรู้เพิ่มขึ้น

10.1.4 ทำให้ทราบจุดมุ่งหมายในการเรียนชัดเจนยิ่งขึ้น เพราะก่อนสอบครูจะแจ้งจุดประสงค์ที่จะประเมินให้ทราบ

10.2 ประโยชน์ต่อครู

10.2.1 ทำให้ครูได้ทราบผลการเรียนของผู้เรียนว่าเก่ง-อ่อนเพียงใด เพื่อหาทางช่วยเหลือ

10.2.2 ทำให้ครูทราบผลการบรรลุจุดประสงค์การเรียนของผู้เรียนว่า มีผู้เรียนผ่านจุดประสงค์มากน้อยเพียงใดซึ่งสะท้อนให้เห็นว่าครูมีประสิทธิภาพในการจัดการเรียนเพียงใด เทคนิควิธีสอนที่ครูใช้เหมาะสมเพียงใด อันจะนำไปสู่การปรับปรุงตนเองของครู

10.2.3 ผลการประเมินจะทำให้ครูใช้เป็นแรงกระตุ้นให้ผู้เรียนเกิดความสนใจในการเรียนยิ่งขึ้น

10.2.4 ช่วยให้ผู้ครูได้ตรวจสอบคุณภาพของข้อสอบว่ามีความยากง่ายเพียงใด มีค่าอำนาจจำแนก ความเที่ยง และความตรงเพียงใด

10.3 ประโยชน์ต่อผู้บริหาร

10.3.1 ทำให้ทราบมาตรฐานการศึกษาของโรงเรียนในด้านต่าง ๆ เช่น มาตรฐานความสามารถของผู้เรียน มาตรฐานความสามารถของครู เป็นต้น

10.3.2 ใช้เป็นข้อมูลในการประชาสัมพันธ์โรงเรียนให้ผู้ปกครองและประชาชนได้ทราบ

10.3.3 ใช้เป็นข้อมูลในการตัดสินใจแก้ปัญหา พัฒนา และดำเนินงานบริหารโรงเรียนในด้านต่าง ๆ ให้มีประสิทธิภาพยิ่งขึ้น

10.4 ประโยชน์ต่อผู้ปกครอง

10.4.1 ทำให้ทราบความสามารถ หรือสมรรถภาพในการเรียนของบุตรหลาน

10.4.2 ช่วยให้ผู้ปกครองรู้จักและเข้าใจบุตรหลานได้ดียิ่งขึ้น

10.4.3 เป็นประโยชน์สำหรับผู้ปกครองใช้ในการตัดสินใจเกี่ยวกับการทำงานหรือการศึกษาต่อของบุตรหลาน

10.5 ประโยชน์ต่อการแนะแนว

10.5.1 ใช้เป็นข้อมูลในการให้คำปรึกษา แนะนำ ผู้เรียนทั้งในด้านการเลือกอาชีพ การศึกษาต่อ และปัญหาส่วนตัว

10.5.2 เป็นข้อมูลพื้นฐานสำหรับโรงเรียนในการประชาสัมพันธ์หรือแนะนำโรงเรียน

10.6 ประโยชน์ต่อการวิจัย

10.6.1 เครื่องมือในการวัดและประเมินผลสามารถนำไปใช้เป็นเครื่องมือในการทำวิจัยได้

10.6.2 ข้อมูลจากการวัดและประเมินผลสามารถนำไปเป็นข้อมูลพื้นฐานในการทำวิจัยต่อไปได้

วิธีการและเครื่องมือวัดและประเมินผลการเรียนรู้

วิธีการและเครื่องมือวัดและประเมินผลการเรียนรู้ หมายถึง รูปแบบยุทธวิธี และเครื่องมือประเภทต่าง ๆ ใช้ในการเก็บรวบรวมข้อมูลเกี่ยวกับการจัดการเรียนรู้ โดยทั่วไปมีจุดมุ่งหมายหลายประการคือ เพื่อรู้จักผู้เรียน เพื่อประเมินวิธีเรียนของผู้เรียน และเพื่อประเมิน

พัฒนาการของผู้เรียน ผู้สอนสามารถเลือกใช้หรือคิดค้นวิธีการวัดและการประเมินผลให้เหมาะสมกับ จุดมุ่งหมายของการนำผลการประเมินไปใช้เพื่อตอบสนองความต้องการประการดังกล่าวข้างต้น วิธีการวัดและประเมินผลการเรียนรู้อาจแบ่งออกตามรูปแบบหรือลักษณะการวัดและประเมินผลได้ เป็น 2 แบบใหญ่ ๆ ดังนี้ (สำนักวิชาการและมาตรฐานการศึกษา สำนักงานคณะกรรมการการศึกษา ขั้นพื้นฐาน, 2553, หน้า 97-112)

1. วิธีการและเครื่องมือวัดและประเมินผลแบบเป็นทางการ (Formal Assessment) เป็นการได้มาซึ่งข้อมูลผลการเรียนรู้ที่นิยมใช้กันมาแต่ดั้งเดิม เช่น วัดและประเมินผลโดยการทดสอบ และใช้แบบทดสอบ (Test) ที่ครูสร้างขึ้น การเก็บข้อมูลดังกล่าวส่วนใหญ่ใช้การวัดและประเมินที่ได้ผลเป็นคะแนน และนำผลการประเมินไปใช้ในการเปรียบเทียบ เช่น เปรียบเทียบระหว่างก่อนเรียน และหลังเรียนเพื่อดูพัฒนาการหรือใช้เพื่อประเมินผลสัมฤทธิ์เมื่อสิ้นสุดการสอนในแต่ละหน่วยหรือ รายวิชา วิธีการและเครื่องมือวัดและประเมินผลแบบเป็นทางการเหมาะสำหรับการประเมินเพื่อตัดสิน มากกว่า ที่จะใช้เพื่อประเมินพัฒนาการผู้เรียนหรือเพื่อหาจุดบกพร่องสำหรับนำไปปรับปรุงการจัดการเรียนการสอน อย่างไรก็ตามวิธีการและเครื่องมือที่ใช้เก็บรวบรวมข้อมูลผลการเรียนรู้แบบเป็นทางการซึ่งให้ข้อมูลสารสนเทศในเชิงปริมาณมีข้อสังเกตที่ผู้สอนต้องระมัดระวังในการนำไปใช้เพื่อให้ ได้ผลการเรียนรู้ที่มีคุณภาพ เป็นตัวแทนของระดับความสามารถที่แท้จริงของผู้เรียนต้องได้มาจาก วิธีการวัดที่ถูกต้องเหมาะสมกับลักษณะข้อมูลเครื่องมือวัดและประเมินผลมีความตรง (Validity) หมายถึง สามารถวัดได้ตรงตามสิ่งที่ต้องการวัด และมีความเที่ยง (Reliability) หมายถึง ผลการวัดมีความคงเส้นคงวา เมื่อมีการวัดซ้ำโดยใช้เครื่องมือคู่ขนานเมื่อวัดในระยะเวลาใกล้เคียงกันและวิธีการวัดมีความโปร่งใสสามารถตรวจสอบและเชื่อถือได้ (Acceptable)

2. วิธีการและเครื่องมือวัดและประเมินผลแบบไม่เป็นทางการ (Informal Assessment) เป็นการได้มาซึ่งข้อมูลผลการเรียนรู้ที่เน้นผู้เรียนเป็นรายบุคคลจากแหล่งข้อมูลหลากหลาย ที่ผู้สอน เก็บรวบรวมตลอดเวลา วิเคราะห์ข้อมูล ศึกษาความพร้อม และพัฒนาการของผู้เรียน ปรับการเรียน การสอนให้เหมาะสม และแก้ไขปัญหาการเรียนรู้ของผู้เรียนลักษณะของข้อมูลที่ได้นอกเหนือจาก ตัวเลขหรือข้อมูลเชิงปริมาณ แล้วอาจเป็นข้อมูลบรรยายลักษณะพฤติกรรมที่ผู้สอนเฝ้าสังเกตหรือผลการเรียนรู้ในลักษณะคำอธิบายระดับพัฒนาการ จุดแข็ง จุดอ่อน หรือปัญหาของผู้เรียนที่พบจากการ สังเกต สัมภาษณ์ หรือวิธีการอื่น ๆ วิธีการประเมินแบบต่าง ๆ ที่ผู้สอนสามารถเลือกใช้ในการประเมิน ในชั้นเรียนมีดังต่อไปนี้

2.1 การสังเกตพฤติกรรม เป็นการเก็บข้อมูลจากการสังเกตการปฏิบัติกิจกรรม ของผู้เรียนโดยไม่ขัดจังหวะการทำงานหรือการคิดของผู้เรียน การสังเกตพฤติกรรมเป็นสิ่งที่ทำได้ ตลอดเวลาแต่ควรมีกระบวนการและมีจุดประสงค์ที่ชัดเจนว่าต้องการประเมินอะไรโดยอาจใช้

เครื่องมือ เช่น แบบประเมินค่า แบบตรวจสอบรายการ สมุดจดบันทึกเพื่อประเมินผู้เรียนตามตัวชี้วัด และควรสังเกตหลายครั้งเพื่อขจัดความลำเอียง

2.2 การสอบปากเปล่า เป็นการให้ผู้เรียนได้แสดงออกด้วยการพูดตอบประเด็นเกี่ยวกับการเรียนรู้ตามมาตรฐานแล้วผู้สอนเก็บข้อมูลโดยจดบันทึกการประเมินรูปแบบนี้ผู้สอนและผู้เรียนมีปฏิสัมพันธ์กันสามารถมีการอภิปราย ได้แย้ง ขยายความ ปรับแก้ไขความคิดกันโดยมีข้อที่พึงระวังคืออย่าเพิ่งขัดความคิดขณะที่ผู้เรียนกำลังพูด

2.3 การพูดคุย เป็นการสื่อสารอีกประเภทหนึ่งระหว่างผู้สอนกับผู้เรียนสามารถดำเนินการเป็นกลุ่มหรือรายบุคคลก็ได้โดยทั่วไปมักใช้เพื่อติดตามตรวจสอบว่าผู้เรียนเกิดการเรียนรู้เพียงใดเป็นข้อมูลสำหรับพัฒนาวิธีการนี้อาจใช้เวลาแต่มีประโยชน์ต่อการค้นหา วินิจฉัยข้อปัญหา ตลอดจนเรื่องอื่น ๆ ที่อาจเป็นปัญหาอุปสรรคต่อการเรียนรู้ เช่น วิธีการเรียนรู้ที่แตกต่างกัน เป็นต้น

2.4 การใช้คำถาม การใช้คำถามเป็นเรื่องปกติมากในการจัดการเรียนรู้แต่ข้อมูลงานวิจัยบ่งชี้ว่าคำถามที่ครูใช้ส่วนใหญ่เป็นด้านความจำและเป็นเชิงการจัดการทั่ว ๆ ไป เพราะถามง่ายแต่ไม่ท้าทายให้ผู้เรียนต้องทำความเข้าใจและเรียนรู้ให้ลึกซึ้ง การพัฒนาการใช้คำถามให้มีประสิทธิภาพแม้จะเป็นเรื่องที่ยากแต่สามารถทำได้ผลรวดเร็วขึ้นหากผู้สอนมี การเปลี่ยนแปลงวิธีการถามเพื่อให้ผู้เรียนได้พัฒนาการคิดอย่างสม่ำเสมอ (Clarke, 2005) Clarke ได้นำเสนอวิธีการฝึกถามให้มีประสิทธิภาพ 5 วิธี ดังนี้

2.4.1 ให้คำตอบที่เป็นไปได้หลากหลายเป็นวิธีที่ง่ายที่สุดในการเริ่มต้นเปลี่ยนการถามแบบความจำให้ เป็นคำถามที่ต้องใช้การคิดบ้างเพราะมีคำตอบที่เป็นไปได้หลายคำตอบ (แต่พึงระวังว่าการใช้คำถามแบบนี้หมายความว่า ผู้เรียนต้องผ่านการเรียนรู้มีความเข้าใจพื้นฐานตามตัวชี้วัดที่กำหนดให้เรียนรู้มาแล้ว) คำถามแบบนี้ทำให้ผู้เรียนต้องตัดสินใจว่าคำตอบใดถูกหรือใกล้เคียงที่สุดเพราะเหตุใดและทำไมถูกเพราะเหตุใดนอกจากนี้การใช้ คำถามแบบนี้จะทำให้ผู้เรียน เรียนรู้ยิ่งขึ้นอีก หากมีกิจกรรมให้ผู้เรียนทำเพื่อพิสูจน์คำตอบ เช่น

ตัวอย่าง

คำถามจำ : การออกกำลังกายแบบใดทำให้หัวใจทำงานดีขึ้น

ตัวอย่าง

คำถามคิด : การออกกำลังกายแบบใดต่อไปนี้ช่วยให้หัวใจทำงาน

ได้ดีขึ้น

- การขี่จักรยาน

- การเดิน

- การเล่นเกมกีฬา
- การว่ายน้ำ
- การกระโดดร่ม
- การพุ่งแหลน

2.4.2 เปลี่ยนคำถามจำ เป็นประโยคบอกเล่าเพื่อให้ผู้เรียนระบุว่าเห็นด้วย ไม่เห็นด้วย พร้อมเหตุผล การใช้วิธีนี้จะต้องให้ผู้เรียนได้อภิปราย ผู้เรียนต้องใช้การคิดที่สูงขึ้นกว่าวิธี แรกเพราะผู้เรียนจะต้องยกตัวอย่างเพื่อสนับสนุนความคิดเห็นของตน ผู้เรียนจะต้องสะท้อนความคิด เห็น ผู้เรียนจะต้องปกป้องหรืออธิบายทัศนคติของตน การฝึกด้วยวิธีการนี้บ่อย ๆ จะเป็นการพัฒนา ผู้เรียนให้เป็นผู้ฟังที่ดีมีจิตใจเปิดกว้างพร้อมรับฟังและเปลี่ยนแปลงความคิดเห็นผ่านกระบวนการ อภิปราย ครูใช้วิธีการนี้กระตุ้นให้เกิดการอภิปรายอย่างมีคุณภาพระหว่างเด็กต่อเด็กและให้ข้อมูลเพื่อ การพัฒนาแก่ทุกคนในชั้นเรียน เช่น

ตัวอย่าง

คำถามจำ :

การออกกำลังกายแบบใดทำให้หัวใจทำงานได้ดีขึ้น

ตัวอย่าง

คำถามคิด :

“การออกกำลังกายแบบต่าง ๆ นั้นทำให้หัวใจทำงานได้ดี ขึ้น” ท่านเห็นด้วยหรือไม่เพราะเหตุใด

2.4.3 หาสิ่งตรงกันข้ามหรือสิ่งที่ใช่/ ถูก สิ่งที่ไม่ใช่/ผิด และถามเหตุผลวิธีการนี้ ใช้ได้ดีกับเนื้อหาที่เป็นข้อเท็จจริง เช่น จำนวนในวิชาคณิตศาสตร์การสะกดคำโครงสร้างไวยากรณ์ใน วิชาภาษา เป็นต้น เมื่อได้รับคำถามว่าทำไมทำเช่นนั้นถูก แต่ทำเช่นนั้นผิด หรือทำไมผลบวกนี้ถูกแต่ ผลบวกนี้ผิด หรือทำไมประโยคนี้ถูกไวยากรณ์แต่ประโยคนั้นผิดไวยากรณ์ เป็นต้น จะเป็นโอกาสให้ ผู้เรียนคิดและอภิปรายมากกว่าเพียงการถามว่าทำไมโดยไม่มีเปรียบเทียบ และวิธีการนี้จะใช้ กับการทำงานเป็นคู่มากกว่าถามทั้งห้องแล้วให้ยกมือตอบ

ตัวอย่าง

คำถามจำ :

พืชต้องการอะไรเพื่อการเจริญเติบโต

ตัวอย่าง

คำถามคิด :

ทำไมต้นไม้ต้นนี้จึงสมบูรณ์แข็งแรงแต่อีกต้นหนึ่งกำลัง

จะตาย

2.4.4 ให้คำตอบเป็นประเด็นสรุปแล้วตามด้วยคำถามให้คิดเป็นการให้ผู้เรียน ต้องอธิบายเพิ่มเติม

ตัวอย่าง

- คำถามจำ : จงบอกคำที่เป็นคำเชื่อม
 คำถามจำ : การพรรณนาความที่ดีประกอบด้วยอะไรบ้าง

ตัวอย่าง

- คำถามคิด : ทำไมเราจึงเรียกคำว่า “แต่” และ “ดังนั้น” ว่าเป็น
 คำเชื่อม
 คำถามคิด : ทำไมข้อความนี้จึงเป็นการพรรณนาความที่ดี

2.4.5 ตั้งคำถามจากจุดยืนที่เห็นต่างเป็นวิธีที่ต้องใช้ความสามารถมากทั้งผู้สอนและผู้เรียนเพราะมีประเด็นที่ต้องอภิปรายโต้แย้งเชิงลึก เหมาะที่จะใช้อภิปรายในประเด็นที่เกี่ยวกับสภาพเศรษฐกิจ สังคม ปัญหาสุขภาพ ปัญหาเชิงจริยธรรม เป็นต้น

ตัวอย่าง

- คำถามจำ : การรีไซเคิลคืออะไร
 คำถามจำ : การสูญพันธุ์มีอันตรายอย่างไร

ตัวอย่าง

- คำถามคิด : เหตุใดโรงงานผลิตพลาสติกจึงชูประเด็นการรีไซเคิล
 คำถามคิด : การสูญพันธุ์ควรเป็นสิ่งที่เลือกกระทำหรือไม่

นอกจากนี้ การใช้ Bloom's Taxonomy (1956) เป็นกรอบแนวคิดในการตั้งคำถามก็เป็นวิธีการที่ดีในการเก็บข้อมูลการเรียนรู้จากผู้เรียน

2.5 การเขียนสะท้อนการเรียนรู้ (Journals) เป็นรูปแบบการบันทึกการเขียนอีกรูปแบบหนึ่งให้ผู้เรียนเขียนตอบกระทู้หรือคำถามของครูซึ่งจะต้องสอดคล้องกับความรู้ทักษะที่กำหนดในตัวชี้วัดการเขียนสะท้อนการเรียนรู้这不仅จากการทำให้ผู้สอนทราบความก้าวหน้าในผลการเรียนรู้แล้วยังใช้เป็นเครื่องมือประเมินพัฒนาการด้านทักษะการเขียนได้อีกด้วย

2.6 การประเมินการปฏิบัติ (Performance Assessment) เป็นวิธีการประเมินงานหรือกิจกรรมที่ผู้สอนมอบหมายให้ผู้เรียนปฏิบัติงานเพื่อให้ทราบถึงผลการพัฒนาของผู้เรียนการประเมินลักษณะนี้ผู้สอนต้องเตรียมสิ่งสำคัญ 2 ประการ คือ ภาระงาน (Tasks) หรือกิจกรรมที่จะให้ผู้เรียนปฏิบัติ เช่น การทำโครงการ/ โครงงาน การสำรวจ การนำเสนอ การสร้างแบบจำลอง การท่องเที่ยวปากเปล่า การสาธิต การทดลองวิทยาศาสตร์ การจัดนิทรรศการ การแสดงละคร เป็นต้น และเกณฑ์การให้คะแนน (Scoring Rubrics) ที่ใช้ประเมินการปฏิบัติ อาจจะปรับเปลี่ยนไปตามลักษณะงานหรือประเภทกิจกรรม ดังนี้

2.6.1 ภาระงานหรือกิจกรรมที่เน้นขั้นตอนการปฏิบัติ และผลงาน เช่น การทดลอง วิทยาศาสตร์การจัดนิทรรศการ การแสดงละคร การแสดงเคลื่อนไหว การประกอบอาหาร การประดิษฐ์ การสำรวจ การนำเสนอ การจัดทำแบบจำลอง เป็นต้น ผู้สอนจะต้องสังเกต และประเมินวิธีการทำงานที่เป็นขั้นตอนและผลงานของผู้เรียน

2.6.2 ภาระงานหรือกิจกรรมที่มุ่งเน้นการสร้างลักษณะนิสัย เช่น การรักษาความสะอาดการรักษาสาธารณสุข/ สิ่งแวดล้อมกิจกรรมหน้าเสาธง เป็นต้น ผู้สอนจะประเมินด้วยวิธีการสังเกตจดบันทึกเหตุการณ์เกี่ยวกับผู้เรียน

2.6.3 ภาระงานที่มีลักษณะเป็นโครงการ/ โครงงานเป็นกิจกรรมที่เน้นขั้นตอนการปฏิบัติและผลงานที่ต้องใช้เวลาในการดำเนินการจึงควรมีการประเมินเป็นระยะ ๆ เช่น ระยะก่อนดำเนินโครงการ/ โครงงานโดยประเมินความพร้อมการเตรียมการและความเป็นไปได้ในการปฏิบัติงานระยะระหว่างดำเนินโครงการ/ โครงงานจะประเมินการปฏิบัติจริงตามแผนวิธีการและขั้นตอนที่กำหนดไว้ และการปรับปรุงระหว่างการปฏิบัติสำหรับระยะสิ้นสุดการดำเนินโครงการ/ โครงงานจะประเมินผลงานผลกระทบและวิธีการนำเสนอผลการดำเนินโครงการ/ โครงงาน

2.6.4 ภาระงานที่เน้นผลผลิตมากกว่ากระบวนการขั้นตอนการทำงาน เช่น การจัดทำแผนผังแผนที่ แผนภูมิ กราฟ ตาราง ภาพ แผนผังความคิด เป็นต้น อาจประเมินเฉพาะคุณภาพของผลงานก็ได้ในการประเมินการปฏิบัติงานผู้สอนต้องสร้างเครื่องมือเพื่อใช้ประกอบการประเมิน เช่น แบบมาตราประมาณค่า แบบบันทึกพฤติกรรม แบบตรวจสอบรายการ แบบบันทึกผลการปฏิบัติ เป็นต้น

2.7 การประเมินด้วยแฟ้มสะสมงาน (Portfolio Assessment) แฟ้มสะสมงานเป็นการเก็บรวบรวมชิ้นงานของผู้เรียนเพื่อสะท้อนความก้าวหน้าและความสำเร็จของผู้เรียน เช่น แฟ้มสะสมงานที่แสดงความก้าวหน้าของผู้เรียนต้องมีผลงานในช่วงเวลาต่าง ๆ ที่แสดงถึงความก้าวหน้าของผู้เรียนหากเป็นแฟ้มสะสมงานดี เด่นต้องแสดงผลงานที่สะท้อนความสามารถของผู้เรียนโดยผู้เรียนต้องแสดงความคิดเห็นหรือเหตุผลที่เลือกผลงานนั้นเก็บไว้ตามวัตถุประสงค์ของแฟ้มสะสมงานแนวทางในการจัดทำแฟ้มสะสมงาน มีดังนี้

2.7.1 กำหนดวัตถุประสงค์ของแฟ้มสะสมงานว่าต้องการสะท้อนเกี่ยวกับความก้าวหน้าและความสำเร็จของผู้เรียนในเรื่องใดด้านใดทั้งนี้อาจพิจารณาจากตัวชี้วัด/ มาตรฐานการเรียนรู้

2.7.2 วางแผนการจัดทำแฟ้มสะสมงานที่เน้นการจัดทำชิ้นงานกำหนดเวลาของการจัดทำแฟ้มสะสมงานและเกณฑ์การประเมิน

2.7.3 จัดทำแผนการจัดทำแฟ้มสะสมงาน และดำเนินการตามแผนที่กำหนด

2.7.4 ให้ผู้เรียนเก็บรวบรวมชิ้นงาน

2.7.5 ให้มีการประเมินชิ้นงานเพื่อพัฒนาชิ้นงานควรประเมินแบบมีส่วนร่วมโดยผู้ประเมินได้แก่ ตนเอง เพื่อน ผู้สอน ผู้ปกครอง บุคคลที่เกี่ยวข้อง

2.7.6 ให้ผู้เรียนคัดเลือกชิ้นงานประเมินชิ้นงานตามเงื่อนไขที่ผู้สอนและผู้เรียนร่วมกันกำหนด เช่น ชิ้นงานที่ยากที่สุด ชิ้นงานที่ชอบที่สุด เป็นต้น โดยดำเนินการเป็นระยะอาจจะเป็น เดือนละครึ่งหรือบทเรียนละครึ่งก็ได้

2.7.7 ให้ผู้เรียนนำชิ้นงานที่คัดเลือกแล้วจัดทำเป็นแฟ้มที่สมบูรณ์ซึ่งควรประกอบด้วยหน้าปก คำนำ สารบัญ ชิ้นงาน แบบประเมินแฟ้มสะสมงาน และอื่น ๆ ตามความเหมาะสม

2.7.8 ผู้เรียนต้องสะท้อนความรู้สึกและความคิดเห็นต่อชิ้นงานหรือแฟ้มสะสมงาน

2.7.9 สถานศึกษาควรจัดให้ผู้เรียนแสดงแฟ้มสะสมงานและชิ้นงานเมื่อสิ้นภาคเรียน/ ปีการศึกษาตามความเหมาะสม

2.8 การวัดและประเมินผลด้วยแบบทดสอบเป็นการประเมินตัวชี้วัดด้านการรับรู้ข้อเท็จจริง (Knowledge) ผู้สอนควรเลือกใช้แบบทดสอบให้ตรงตามวัตถุประสงค์ของการวัดและประเมินผลนั้น ๆ เช่น แบบทดสอบเลือกตอบ แบบทดสอบถูก-ผิด แบบทดสอบจับคู่ แบบทดสอบเติมคำ แบบทดสอบความเรียง เป็นต้น ทั้งนี้ แบบทดสอบที่จะใช้ต้องเป็นแบบทดสอบที่มีคุณภาพ มีความตรง (Validity) และความเที่ยง (Reliability)

2.9 การประเมินด้านความรู้สึกนึกคิด (Attitude) เป็นการประเมินคุณธรรม จริยธรรม คุณลักษณะและเจตคติที่ควรปลูกฝังในการจัดการเรียนรู้ ซึ่งการวัดและประเมินผลจะเป็นลำดับขั้นจากต่ำสุดไปสูงสุด ดังนี้

2.9.1 ชั้นรับรู้ เป็นการประเมินพฤติกรรมที่แสดงออกว่ารู้จักเต็มใจ สนใจ

2.9.2 ชั้นตอบสนอง เป็นการประเมินพฤติกรรมที่แสดงว่าเชื่อฟัง ทำตาม อาสาทำพอใจที่จะทำ

2.9.3 ชั้นเห็นคุณค่า (ค่านิยม) เป็นการประเมินพฤติกรรมที่แสดงความเชื่อ ซึ่งแสดงออกโดยการกระทำหรือปฏิบัติอย่างสม่ำเสมอ ยกย่องชมเชย สนับสนุนช่วยเหลือหรือทำกิจกรรมที่ตรงกับความเชื่อของตน ทำด้วยความเชื่อมั่นศรัทธา และปฏิเสธที่จะกระทำในสิ่งที่ขัดแย้งกับความเชื่อของตน

2.9.4 ชั้นจัดระบบคุณค่าเป็นการประเมินพฤติกรรมการเข้าร่วมกิจกรรม อภิปรายเปรียบเทียบจนเกิดอุดมการณ์ในความคิดของตนเอง

2.9.5 ชั้นสร้างคุณลักษณะเป็นการประเมินพฤติกรรมที่มี แนวโน้มว่าจะประพฤติปฏิบัติเช่นนั้นอยู่เสมอในสถานการณ์เดียวกันหรือเกิดเป็นอุปนิสัย การวัดและประเมินผลด้านจิตพิสัย ควรใช้การสังเกตพฤติกรรมการปฏิบัติเป็นหลัก และสังเกตอย่างต่อเนื่องในระยะเวลาหนึ่งโดยมีการบันทึกผลการสังเกตทั้งนี้อาจใช้เครื่องมือการวัดและประเมินผล เช่น แบบประเมินค่า แบบตรวจสอบรายการ แบบบันทึกพฤติกรรม แบบรายงานพฤติกรรมตนเอง เป็นต้น นอกจากนี้อาจใช้แบบวัดความรู้ และความรู้สึกเพื่อรวบรวมข้อมูลเพิ่มเติม เช่น แบบวัดความรู้โดยสร้างสถานการณ์เชิงจริยธรรม แบบวัดเจตคติ แบบวัดเหตุผลเชิงจริยธรรม แบบวัดพฤติกรรมเชิงจริยธรรม เป็นต้น

2.10 การประเมินตามสภาพจริง (Authentic Assessment) เป็นการประเมินด้วยวิธีการที่หลากหลายดังที่กล่าวมาแล้วข้างต้น เพื่อให้ได้ผลการประเมินที่สะท้อนความสามารถที่แท้จริงของผู้เรียนจึงควรใช้การประเมินการปฏิบัติ (Performance Assessment) ร่วมกับการประเมินด้วยวิธีการอื่น ภาระงาน (Tasks) ควรสะท้อนสภาพความเป็นจริงหรือใกล้เคียงกับชีวิตจริงมากกว่าเป็นการปฏิบัติกิจกรรมทั่ว ๆ ไป ดังนั้น การประเมินสภาพจริงจะต้องออกแบบการจัดการเรียนรู้และการประเมินผลไปด้วยกัน และกำหนดเกณฑ์การประเมิน (Rubrics) ให้สอดคล้องหรือใกล้เคียงกับชีวิตจริง

2.11 การประเมินตนเองของผู้เรียน (Student Self-Assessment) การประเมินตนเองนับเป็นทั้งเครื่องมือประเมินและเครื่องมือพัฒนาการเรียนรู้เพราะทำให้ผู้เรียนได้คิดใคร่ครวญว่าได้เรียนรู้อะไร เรียนรู้อย่างไร และผลงานที่ทำนั้นดีแล้วหรือยัง การประเมินตนเองจึงเป็นวิธีหนึ่งที่จะช่วยพัฒนาผู้เรียนให้เป็นผู้ที่สามารถเรียนรู้ด้วยตนเองการประเมินตนเองของผู้เรียนที่ประสบความสำเร็จจะต้องมีเป้าหมายการเรียนรู้ที่ชัดเจน มีเกณฑ์ที่บ่งบอกความสำเร็จของชิ้นงาน/ ภาระงาน และมาตรการการปรับปรุงแก้ไขตนเองเป้าหมายการเรียนรู้ที่กำหนดชัดเจนและผู้เรียนได้รับทราบหรือร่วมกำหนดด้วยจะทำให้ผู้เรียนทราบว่าตนถูกคาดหวังให้รู้อะไรทำอะไรมีหลักฐานใดที่แสดงการเรียนรู้ตามความคาดหวังนั้น หลักฐานที่มีคุณภาพควรมีเกณฑ์เช่นไร เพื่อเป็นแนวทางให้ผู้เรียนพิจารณาประเมินซึ่งหากเป้าหมายและเกณฑ์การประเมินเกิดจากการทำงานร่วมกันระหว่างผู้เรียนกับผู้สอนด้วยแล้วจะเป็นการเพิ่มแรงจูงใจในการเรียนรู้เพิ่มมากขึ้นการที่ผู้เรียนได้ใช้การประเมินตนเองบ่อย ๆ โดยมีกรอบแนวทางการประเมินที่ชัดเจนนี้จะช่วยส่งเสริมให้ผู้เรียนประเมินได้ค่อนข้างตรง และซื่อสัตย์ คำวิจารณ์ คำแนะนำของผู้เรียนมักจะมีจริงจังมากกว่าของครู การประเมินตนเอง จะเกิดประโยชน์ยิ่งขึ้นหากผู้เรียนทราบสิ่งที่ต้องปรับปรุงแก้ไขและตั้งเป้าหมายการปรับปรุงแก้ไขของตนแล้วฝึกฝนพัฒนาโดยการดูแลสนับสนุนจากผู้สอนและความร่วมมือของครอบครัว

2.12 การประเมินโดยเพื่อน (Peer Assessment) เป็นเทคนิคการประเมินอีกรูปแบบหนึ่งที่นำมามาใช้เพื่อพัฒนาผู้เรียนให้เข้าถึงคุณลักษณะของงานที่มีคุณภาพเพราะการที่ผู้เรียนจะบอกได้ว่าชิ้นงานนั้นเป็นเช่นไร ผู้เรียนต้องมีความเข้าใจอย่างชัดเจนก่อนว่าเขากำลังตรวจสอบอะไรในงานของเพื่อน ฉะนั้นผู้สอนต้องอธิบายผลที่คาดหวังให้ผู้เรียนทราบก่อนที่จะลงมือประเมิน การที่จะสร้างความมั่นใจว่าผู้เรียนเข้าใจการประเมินรูปแบบนี้ควรมีการฝึกโดยผู้สอน อาจหาตัวอย่าง เช่น งานเขียน ให้กลุ่มผู้เรียนตัดสินใจว่าควรประเมินอะไร และควรให้คำอธิบายเกณฑ์ที่บ่งบอกความสำเร็จของภาระงานนั้น จากนั้นให้ผู้เรียนประเมินภาระงานเขียนที่เป็นตัวอย่างนั้นโดยใช้เกณฑ์ที่ช่วยกันสร้างขึ้นหลังจากนั้นผู้สอนตรวจสอบการประเมินของผู้เรียน และให้ข้อมูลย้อนกลับแก่ผู้เรียนที่ประเมินเกินจริง การประเมินโดยเพื่อนที่มี ประสิทธิภาพจำเป็นต้องสร้างสิ่งแวดล้อมการเรียนรู้ที่สนับสนุนให้เกิดการประเมินรูปแบบนี้ กล่าวคือ ผู้เรียนต้องรู้สึกผ่อนคลาย เชื่อใจกัน และไม่อคติเพื่อการให้ ข้อมูลย้อนกลับจะได้ชื่อตรงเป็นเชิงบวกที่ให้ ประโยชน์ ผู้สอนที่ให้ผู้เรียนทำงานกลุ่มตลอดภาคเรียนแล้วใช้เทคนิคเพื่อนประเมินเพื่อนเป็นประจำจะสามารถพัฒนาผู้เรียนให้เกิดความเข้าใจซึ่งกันและกันอันจะนำไปสู่การให้ข้อมูลย้อนกลับที่เก่งขึ้นได้

เครื่องมือวัดและประเมินผลการเรียนรู้

วิธีการประเมินแบบต่าง ๆ ที่ผู้สอนนำไปใช้ในการประเมินเพื่อพัฒนาผู้เรียนนั้นมีหลากหลาย เครื่องมือที่ใช้ในการเก็บข้อมูลของผู้เรียนก็มีหลากหลายเช่นกัน บางกรณีวิธีประเมินอาจใช้เครื่องมือเพียงอย่างเดียว บางกรณี อาจใช้เครื่องมือหลายอย่าง ผู้สอนสามารถเลือกใช้ได้ตามวัตถุประสงค์และความเหมาะสม ดังตัวอย่างต่อไปนี้

1. แบบสัมภาษณ์ เป็นเครื่องมือที่ใช้ในการพูดคุยซักถามระหว่างครูกับผู้เรียนเพื่อรวบรวมข้อมูลต่าง ๆ ทำให้ครูสามารถสังเกตกิริยาท่าทาง ลักษณะทางร่างกาย การแสดงพฤติกรรมทางที่ การพูดโต้ตอบ การสัมภาษณ์ที่ใช้ในโรงเรียนมักเป็นการสัมภาษณ์เพื่อหาข้อเท็จจริง และการสัมภาษณ์เพื่อปรึกษาปัญหาการสัมภาษณ์อาจทำในหรือนอกห้องเรียนในหรือนอกเวลาเรียนหรือมีการนัดหมายกันก็ได้การสัมภาษณ์ที่ดี ต้องมีจุดมุ่งหมายที่ชัดเจนว่าครูต้องการจะได้ข้อมูลอะไร เช่น สัมภาษณ์ผู้เรียนเพื่อหาข้อมูลในการแก้ปัญหาการมาโรงเรียนสายบ่อย ๆ ข้อมูลในการจัดหาหนังสือเข้าห้องสมุด เป็นต้น

ตัวอย่างแบบสัมภาษณ์

คำถาม :	ผู้เรียนชอบอ่านหนังสือประเภทใดมากที่สุด.....
คำถาม :	เพราะเหตุใดถึงชอบหนังสือเล่มนั้น.....
คำถาม :	ผู้เรียนใช้เวลาช่วงใดในการอ่านหนังสือ.....

2. แบบสังเกต เป็นเครื่องมือรวบรวมข้อมูลที่ผู้สอนใช้ในการสังเกตผู้เรียนแต่ละคนหรือเป็นกลุ่มในเวลาใดเวลาหนึ่งขึ้นอยู่กับวัตถุประสงค์ในการสังเกตของครูการสังเกตที่ให้ผลเชื่อถือได้นั้นผู้สอนต้องสังเกตผู้เรียนที่ละคนระบุพฤติกรรมที่สังเกตได้จัดจลเวลาสังเกตให้เป็นระบบสังเกตซ้ำ ๆ กันหลายช่วงเวลาถ้าจะให้เชื่อถือได้สูงควรมีผู้สังเกต 2 คน สังเกตผู้เรียนคนเดียวกันในเวลาเดียวกันหรือผู้สอนสังเกตผู้เรียนที่ละคนแล้วบันทึกการสังเกตไว้ดังตัวอย่างต่อไปนี้

ตัวอย่างแบบบันทึกการสังเกตพฤติกรรม

วัน.....เดือน.....ปี.....ที่สังเกต ผู้สอน.....
วิชาใด.....เวลา.....
ชื่อผู้เรียน.....
ผู้เรียนกำลังทำอะไร.....
มีใครอยู่กับผู้เรียนขณะที่สังเกตบ้าง.....
พฤติกรรมของผู้เรียนที่สังเกตพบบ่อย ๆ คือ.....
ผู้บันทึก.....

รายการพฤติกรรม	บ่อยครั้ง	บางครั้ง	ไม่เคยเลย
1. เข้าเรียนตรงเวลา			
2. ทำงานตามที่ได้รับมอบหมาย			
3. ทำการบ้านและแบบฝึกหัดตามที่ได้รับมอบหมาย			
4. ส่งการบ้านและแบบฝึกหัดตามที่ได้รับมอบหมาย			
5. นำอุปกรณ์การเรียนมาเรียนครบ			
6. จดงานหรือบันทึกขณะทำกิจกรรมการเรียน			
7. ตั้งใจเรียนและร่วมกิจกรรมการเรียนในห้อง			
8. ระวังรักษาอุปกรณ์ของส่วนรวมและของตนเอง			

ผู้ประเมิน.....

3. แบบสอบถามเป็นเครื่องมือรวบรวมข้อมูลเกี่ยวกับความคิดเห็นที่ให้ผู้ตอบเขียนตอบเอง ซึ่งมีหลายประเภท เช่น ให้ผู้ตอบทำเครื่องหมายเพื่อตอบเขียนตอบสั้น ๆ หรือให้ทำเครื่องหมายเป็นมาตรฐานประมาณค่าผู้ตอบแบบสอบถามอาจเป็นผู้เรียน ผู้ปกครองหรือผู้เกี่ยวข้องกับประเด็นที่ต้องการคำตอบ ดังนั้นแบบสอบถามจึงต้องมีคำชี้แจงในประเด็นที่ต้องการเพื่อให้ผู้ตอบเข้าใจตรงกันดังตัวอย่างต่อไปนี้

ตัวอย่างแบบสอบถามสภาพครอบครัว

ที่อยู่			
.....			
อาชีพบิดา.....		อาชีพมารดา	
.....			
ปัจจุบันอาศัยอยู่กับ			
<input type="checkbox"/>	บิดา-มารดา	<input type="checkbox"/>	บิดา
<input type="checkbox"/>	มารดา	<input type="checkbox"/>	ผู้อื่น
ระบุ.....			

ตัวอย่างแบบสอบถามเกี่ยวกับกิจกรรม

งานอดิเรกที่ผู้เรียนชอบทำคือ.....
ผู้เรียนชอบอ่านหนังสือเกี่ยวกับ.....
กีฬาที่ชอบมากที่สุด.....

4. แบบสำรวจรายการเป็นเครื่องมือที่ใช้เก็บข้อมูลซึ่งเป็นข้อเท็จจริงโดยจัดทำเป็นรายการสำรวจเอาไว้เมื่อเก็บข้อมูลก็ตรวจสอบไปที่ รายการการว่ามี หรือไม่มักใช้ประกอบการเก็บข้อมูลโดยวิธีการสังเกตหรือการวัดทักษะการปฏิบัติ

ตัวอย่างแบบสำรวจปัญหาของผู้เรียน

เมื่อมีปัญหาหรือความคับข้องใจผู้เรียนปรึกษาใคร

..... ครูประจำชั้น

..... พ่อ/ แม่

..... เพื่อนสนิท

.....ญาติ

..... อื่น ๆ โปรดระบุ.....

ข้อความ	ใช่	ไม่ใช่
1. ผู้เรียนไม่เข้าใจบทเรียน
2. ผู้เรียนรู้สึกเบื่อการเรียน
3. ผู้เรียนไม่ชอบไปโรงเรียน
4. ผู้เรียนไม่ชอบครูบางคน
5. ผู้เรียนรู้สึกประหม่าเมื่ออยู่ต่อหน้าคนจำนวนมาก

5. แบบทดสอบ (Test)

แบบทดสอบ หมายถึง ชุดของคำถาม (Items) ที่สร้างขึ้นเพื่อให้ผู้สอบแสดงพฤติกรรมอย่างใดอย่างหนึ่งออกมาให้ผู้สอนสังเกตและวัดได้แบบทดสอบเป็นเครื่องมือวัดพฤติกรรมด้านพุทธิพิสัยซึ่งถือว่าเป็นสติปัญญาของมนุษย์ที่ซ่อนแฝงอยู่ในตัวบุคคลว่ามีความรู้หรือไม่เพียงใดทั้งในด้านพฤติกรรมความรู้ ความจำ ความเข้าใจ การนำไปใช้ และอื่น ๆ หากแบ่งประเภทแบบทดสอบวัดสัมฤทธิ์ทางการเรียนโดยใช้เกณฑ์ลักษณะการตอบแล้วสามารถแบ่งออกได้เป็น 2 ประเภทดังนี้

5.1 แบบทดสอบอัตนัยหรือแบบความเรียง มีลักษณะเด่นที่ให้อิสระแก่ผู้ตอบ ผู้ตอบจะต้องหาหรือสร้างคำตอบเอง แทนที่จะมีคำตอบให้เลือกเหมือนกับข้อสอบแบบกำหนดคำตอบให้ ข้อสอบแบบนี้จะใช้ในการวัดการเรียนรู้ที่ไม่สามารถวัดโดยใช้ข้อสอบแบบกำหนดคำตอบได้ เช่น วัดความสามารถในการจัดการความสามารถในการบูรณาการ ความสามารถในการสังเคราะห์ความรู้ ความสามารถในการแก้ปัญหาตลอดจนความสามารถในการประเมิน เป็นต้น เหมาะที่จะนำมาใช้เมื่อต้องการวัดความสามารถในการใช้เหตุผล การวางแผน การแสดงความคิดเห็น สร้างสรรค์หรือจินตนาการ ผู้ตอบจะต้องรู้สึกในเรื่องที่จะตอบจึงจะเขียนตอบได้ดี ข้อสอบแบบความเรียงแบ่ง

ออกเป็น 2 ลักษณะ คือ คำถามแบบจำกัดคำตอบ (Restricted Response Questions) และคำถามแบบขยายคำตอบ (Extended Response Questions) ซึ่งมีรายละเอียด ดังนี้

5.1.1 คำถามแบบจำกัดคำตอบ เป็นคำถามที่จำกัดเนื้อหาคือ กำหนดขอบเขตที่จะให้ตอบดังตัวอย่าง

1. ให้ผู้เรียนอธิบายเรื่องยาเสพติดให้ครอบคลุมหัวข้อต่อไปนี้
 - ก. ความหมายของยาเสพติด
 - ข. ประเภทของยาเสพติด
2. ให้ผู้เรียนเปรียบเทียบความแตกต่างระหว่างพืชใบเลี้ยงคู่กับพืชใบเลี้ยงเดี่ยว

5.1.2 คำถามแบบขยายคำตอบ เป็นคำถามที่ไม่จำกัดขอบเขตให้ตอบผู้ตอบสามารถเลือกข้อความรู้และนำความรู้เหล่านั้นมาจัดระบบให้ดี แล้วนำมาเขียนเป็นคำตอบจึงเป็นข้อสอบที่วัดความสามารถในการรวบรวมความรู้สังเคราะห์ความรู้เหล่านั้น แล้วนำมาเรียบเรียงและเขียนเป็นคำตอบได้เป็นอย่างดี ดังตัวอย่าง

1. ให้ผู้เรียนอธิบายโครงสร้างของสังคมไทยปัจจุบัน

ข้อสอบแบบความเรียงนี้ เป็นการฝึกให้ผู้เรียนรู้จักรวบรวมความรู้หรือข้อมูลต่าง ๆ นำมาประมวลเป็นเรื่องเดียวกันจัดระบบความรู้เหล่านั้นแล้วเรียบเรียงถ่ายทอดสิ่งเหล่านั้นออกมาด้วยภาษาของตนเอง เพื่อสื่อให้ผู้อื่นเข้าใจซึ่งเป็นกระบวนการของการฝึกความสามารถในการอ่าน คิด วิเคราะห์ และการเขียนของผู้เรียน และสอดคล้องกับแนวคิดใหม่ของ Bloom (Revised Bloom's Taxonomy) ที่กล่าวถึง การประเมินด้านสติปัญญาว่า เป็นการประเมินได้ทั้งทักษะการคิดขั้นพื้นฐาน จนถึงขั้นสูงโดยผู้สอบจะต้องมีความรู้ขั้นพื้นฐาน คือ จำได้ (Remember) และเข้าใจ (Understand) เนื้อหาความรู้เหล่านั้นแล้วนำไปปรับใช้ (Apply) โดยอาจจะวิเคราะห์ (Analyze) ประเมิน (Evaluate) แล้วสร้างสรรค์สิ่งใหม่ (Create) โดยเขียนสื่อสิ่งที่เกิดขึ้นตามกระบวนการทางการคิดนั้นออกมาเป็นคำตอบ ในขณะที่ด้วยกันผู้สอนต้องคำนึงถึงการตรวจให้คะแนนด้วยเนื่องจากการตรวจให้คะแนนต้องใช้เวลามากต้องกำหนดเกณฑ์การให้คะแนน (Rubrics) เพื่อตรวจให้คะแนนได้อย่างยุติธรรม

หลักการเขียนข้อสอบแบบอัตนัย

1. กำหนดให้ชัดเจนว่าต้องการวัดพฤติกรรมด้านใดของผู้สอบ
2. เขียนคำถามให้ชัดเจน จำเพาะเจาะจง ว่าต้องการให้ผู้สอบทำอย่างไร เช่น อธิบาย วิเคราะห์ แสดงความคิดเห็น ฯลฯ รวมทั้งมีคะแนนข้อละกี่คะแนน
3. เขียนคำถามโดยใช้สถานการณ์ใหม่ ๆ ไม่ควรถามตามตำราหรือถามในสิ่งที่เรียนมาแล้ว
4. ต้องถามเฉพาะสิ่งที่เป็นประเด็นสำคัญของเรื่อง
5. กำหนดเวลาในการสอบให้ เหมาะสม
6. เพื่อผู้ตอบจะได้วางแผนการตอบได้ถูกต้องโดยเอาจำนวนข้อไปหารจำนวนเวลาทั้งหมดก็จะทราบว่าแต่ละข้อควรใช้เวลาเท่าไรถ้าไม่จำเป็นควรหลีกเลี่ยงการให้ผู้ตอบเลือกตอบเป็นบางข้อได้ควรให้ทำทุกข้อ
7. พยายามใช้คำถามหลาย ๆ แบบโดยเน้นการอธิบายควรเป็นคำถามประเภททำไม อย่างไร หรือให้อธิบาย บรรยาย เปรียบเทียบ หาความสัมพันธ์ หาความขัดแย้ง ดีความ วิเคราะห์ เหตุผล วิจาร์ณ และประเมินผล เป็นต้น
8. เมื่อเขียนคำถามแล้วควรเขียนคำตอบหรือแนวคำตอบที่ต้องการไว้ด้วยหรืออาจจะเขียนในลักษณะคำหรือข้อความสำคัญ (Key Words) ของคำตอบข้อนั้น ๆ เอาไว้ด้วย
9. ควรกำหนดความยาวและความซับซ้อนของข้อสอบให้พอเหมาะกับความสามารถของผู้สอบ
10. ถ้าข้อสอบมีหลายข้อควรจัดเรียงลำดับจากง่ายไปหายากเพื่อยั่วให้ผู้ อยากรอบมากยิ่งขึ้น

5.2 แบบทดสอบปรนัย เป็นข้อสอบที่มีคำถามเฉพาะเจาะจงตรงให้คะแนนได้ตรงกัน มีคำสั่งวิธีการปฏิบัติ และวิธีการตรวจให้คะแนนที่ชัดเจนแบบทดสอบปรนัยที่นิยมใช้กัน คือ แบบถูก-ผิด (True-False) แบบจับคู่ (Matching) และแบบเลือกตอบ (Multiple Choices) ซึ่งแต่ละประเภท มีรายละเอียด ดังนี้

5.2.1 ข้อสอบแบบถูกผิดเป็นข้อคำถามที่กำหนดข้อความให้ผู้เรียนพิจารณาเลือกตอบสองทางเลือก เช่น ถูก-ผิด ใช่-ไม่ใช่ จริง-ไม่จริง เหมือนกัน-ต่างกัน ฯลฯ โดยใช้ความรู้ตามหลักวิชาเป็นเกณฑ์พิจารณาตัวคำถามมักจะเขียนในรูปประโยคบอกเล่าธรรมดาหรืออาจจะเป็นประโยคคำถามโดยมีข้อความถูกบ้างผิดบ้างคละเคล้ากันไป

หลักการเขียนข้อสอบแบบถูกผิด

1. ข้อความจะต้องมีความหมายชัดเจน ไม่กำกวม และไม่ควรรใช้คำที่แสดงคุณภาพ เช่น มาก น้อย บ่อย ๆ บางครั้ง ส่วนมาก ส่วนน้อย ไม่ค่อยจะ เป็นต้น ควรเลือกคำที่แสดง ปริมาณจะมีความหมายชัดเจนกว่า เช่น

ไม่ดี - พม่ายกกองทัพมาตีไทยบ่อยครั้งในสมัยกรุงธนบุรี

ดีขึ้น - พม่ายกกองทัพมาตีไทย 4 ครั้งในสมัยกรุงธนบุรี

2. ข้อความที่กำหนดให้ต้องตัดสินได้ว่าถูกจริงหรือผิดจริงและเป็นสากล เช่น

ไม่ดี - น้ำเดือดที่อุณหภูมิ 100 องศาเซลเซียส

ดีขึ้น - ณ ระดับน้ำทะเลน้ำจะเดือดที่อุณหภูมิ 100 องศาเซลเซียส

3. แต่ละข้อคำถามควรถามจุดสำคัญเพียงเรื่องเดียว เช่น

ไม่ดี - อำเภอแม่สายอยู่ในจังหวัดเชียงราย และอยู่เหนือสุดของประเทศไทย

ดีขึ้น - อำเภอแม่สายอยู่ในจังหวัดเชียงราย

ดีขึ้น - อำเภอแม่สายอยู่เหนือสุดของประเทศไทย

4. ไม่ควรสร้างข้อคำถามเชิงปฏิเสธหรือปฏิเสธซ้อนเพราะจะทำให้ผู้สอบเข้าใจผิด

ไม่ดี - ถ้าผู้เรียนไม่ออกไปตากน้ำค้างผู้เรียนจะไม่ใช่คนหัวดี

ดีขึ้น - การออกไปตากน้ำค้างทำให้ผู้เรียนเป็นคนหัวดี

5. ควรหลีกเลี่ยงการลอกข้อความจากหนังสือหรือตำราเรียนโดยตรงเพราะจะส่งเสริมการเรียนแบบท่องจำ

6. ให้ข้อสอบแต่ละข้อเป็นอิสระแก่กัน

7. ข้อความแต่ละข้อควรมีความยาวใกล้เคียงกัน

8. ข้อสอบควรเรียงลำดับตามเนื้อหา

9. ควรให้มีจำนวนข้อถูกและข้อผิดใกล้เคียงกัน และอยู่กระจายคละกัน

5.2.2 ข้อสอบแบบจับคู่ เป็นข้อคำถามที่กำหนดข้อความที่สัมพันธ์กันให้ 2 รายการ รายการทางด้านซ้ายเรียกว่าตัวยืมหรือคำถามรายการทางด้านขวาเรียกว่าตัวเลือกหรือคำตอบให้ผู้ตอบพิจารณาความสัมพันธ์ของรายการทั้งสองด้านรายการที่นำมาออกข้อสอบแบบจับคู่ ได้แก่ คำศัพท์กับความหมาย เหตุการณ์กับเวลา เวลา กับสถานที่ ชื่อบุคคลกับผลงาน ชื่อกระบวนการกับการผลิต กฎกับการใช้เหตุผลกับผล เครื่องมือกับประโยชน์ใช้สอย เป็นต้น

หลักการเขียนข้อสอบแบบจับคู่

1. เขียนคำชี้แจงให้ชัดเจนว่าจะให้จับคู่ได้เพียงตัวเลือกเดียวหรืออาจจับคู่ได้หลายตัวเลือก

2. เนื้อหาวิชาที่นำมาออกข้อสอบจะต้องเป็นเรื่องหรือเนื้อหาเดียวกัน เช่น

ตัวอย่างแบบทดสอบแบบจับคู่

คำชี้แจง : รายการทางด้านซ้ายเป็นคำถามรายการทางด้านขวาเป็นคำตอบให้ท่านนำเอาหัวข้อของคำตอบทางขวามือมาใส่ในวงเล็บหน้าคำถามทางซ้ายมือคำตอบทางขวามือนั้นสามารถใช้ได้เพียงครั้งเดียวหรือไม่ใช้เลยก็ได้

- | | |
|---|-------------------------|
| (.....) 1. ใครเป็นคนแต่งหนังสือจินตนิมิต | ก. พระโหราธิบดี |
| (.....) 2. ใครเป็นนายกรัฐมนตรีคนแรกของไทย | ข. พ่อขุนรามคำแหง |
| (.....) 3. ใครเป็นผู้คิดประดิษฐ์อักษรไทยขึ้นเป็นคนแรก | ค. พระยาโกษาธิบดี |
| | ง. พระยาพลพลพยุหเสนา |
| | จ. พระยามโนปกรณนิติธาดา |

3. ควรให้คำตอบมีมากกว่าคำถาม 3-4 ตัว

4. ข้อสอบแบบจับคู่ชุดหนึ่งไม่ควรมี มากข้อเกินไปควรอยู่ระหว่าง 5-12 คู่ และควรให้อยู่ในหน้าเดียวกันทั้งหมด

5. คำหรือข้อความที่เป็นคู่กันไม่ควรจัดให้อยู่ตรงกัน

5.2.3 ข้อสอบแบบเลือกตอบเป็นข้อสอบที่ประกอบด้วยคำถามและคำตอบให้เลือกหลาย ๆ คำตอบข้อสอบประเภทนี้มี 2 ส่วน คือ

5.2.3.1 ตัวนำหรือตัวคำถาม (Stem) เป็นข้อความที่เป็นตัวเร้าให้ผู้สอบคิด

5.2.3.2 ตัวเลือก (Choices) เป็นคำตอบหลาย ๆ คำตอบเพื่อให้ผู้สอบเลือกตอบอย่างใดอย่างหนึ่งมีทั้งตัวถูก (Key) และตัวลวง (Distracters)

ข้อสอบแบบเลือกตอบที่ตีนั้นตัวเลือกทุกตัวจะมีน้ำหนักพอ ๆ กัน ถ้าดูผิวเผินหรือไม่มีความรู้ในข้อนั้นจริง ๆ จะเห็นว่าถูกหมดทุกข้อและในการสอบแต่ละครั้งตัวเลือกแต่ละตัวจะมีโอกาสถูกเลือกพอ ๆ กัน นั่นคือ หากมีข้อสอบ 20 ข้อ และมี 4 ตัวเลือก โอกาสที่ตัวเลือก ก ข ค หรือ ง จะถูกเลือกเท่ากัน และคำตอบถูกควรจะกระจายกันไปทุกตัวเลือกไม่ใช่อยู่ที่ตัวใดตัวหนึ่ง

หลักการเขียนข้อสอบแบบเลือกตอบ

ด้านตัวคำถาม

1. เขียนคำถามให้เป็นประโยคคำถามที่สมบูรณ์
2. เขียนคำถามให้กะทัดรัด ชัดเจนตรงจุด ไม่ใช่คำฟุ่มเฟือย
3. ใช้ภาษาให้เหมาะสมกับระดับผู้สอบ
4. คำถามควรเร้าให้ผู้สอบได้ใช้ความคิด

5. หลีกเลี่ยงการใช้คำถามปฏิเสธซ้อน
 6. ไม่ควรถามในสิ่งที่เด็กท่องจำจนคล่องปาก
 7. คำถามแต่ละข้อควรเป็นอิสระจากกัน
 8. อาจใช้รูปภาพช่วยเพื่อลดความเครียดของผู้สอบ หรือทำให้เข้าใจคำถามดีขึ้น
- ด้านตัวเลือก**

1. คำถามข้อหนึ่ง ๆ ต้องมีคำตอบที่ถูกต้องที่สุดเพียงตัวเลือกเดียวเท่านั้น
2. เขียนให้ทั้งตัวถูกและตัวผิดถูกผิดตามหลักวิชา
3. เขียนให้ตัวเลือกเป็นอิสระจากกัน
4. เขียนตัวเลือกให้กะทัดรัด ไม่ยืดยาว หรือเพิ่มคำที่ไม่จำเป็น
5. ตัวเลือกต้องเป็นเอกพจน์
6. ตัวเลือกที่ถูกไม่ควรยาวเกินไป
7. จัดตัวเลือกให้เป็นระบบ เช่น เรียงตาม พ.ศ. เรียงจากน้อยไปมาก เป็นต้น
8. หลีกเลี่ยงการเขียนตัวถูกให้พ้องเสียงหรือมีคำ/ ข้อความที่ซ้ำกับตัวคำถาม
9. ตำแหน่งของตัวถูกควรกระจายในลักษณะสุ่ม
10. ตัวลวงต้องมีโอกาสเป็นไปได้
11. ไม่ควรมีตัวเลือกประเภท “ถูกหมดทุกข้อ” หรือ “ไม่มีข้อใดถูก”

แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนอาจเป็นแบบปรนัยหรืออัตนัยอย่างใดอย่างหนึ่ง หรืออาจเป็นทั้งแบบปรนัยและอัตนัยรวมกันในแบบทดสอบฉบับเดียวกันก็ได้ ทั้งนี้ขึ้นอยู่กับวัตถุประสงค์ในการวัดระดับพฤติกรรมการเรียนรู้ที่ต้องการวัด จำนวนผู้เข้าสอบ ระยะเวลาในการสร้างข้อสอบ การดำเนินการสอบ และการตรวจข้อสอบ

จากแนวคิดเกี่ยวกับการวัดและการประเมินผลการเรียนรู้ที่กล่าวมา ผู้วิจัยสรุปได้ว่าการวัดเป็นกระบวนการของการกำหนดตัวเลขให้กับสิ่งต่าง ๆ โดยมีหลักการกำหนดตัวเลขที่ชัดเจน ด้วยเทคนิคหรือวิธีการที่ได้รับการยอมรับ สำหรับการประเมินในภาษาอังกฤษนั้นพบว่า มีคำที่ใช้อยู่ 2 คำ คือ คำว่า “Evaluation” และคำว่า “Assessment” “Evaluation” หมายถึง กระบวนการตัดสินคุณค่าของสิ่งหนึ่งสิ่งใด ส่วนคำว่า “Assessment” จะเป็นการประเมินเบื้องต้น ในขณะที่ Evaluation เป็นกระบวนการตัดสินใจหรือตัดสินคุณค่าขั้นสุดท้าย

ประเภทของการวัดและการประเมินผล จำแนกตามวัตถุประสงค์ของการประเมินผล แบ่งเป็น 3 ประเภทคือ การประเมินผลก่อนเรียน การประเมินผลระหว่างเรียนและการประเมินผลรวมสรุป จำแนกตามระบบการวัด แบ่งออกเป็น 2 ประเภท คือ การประเมินผลแบบอิงกลุ่มและ

การประเมินผลแบบอิงเกณฑ์ ความมุ่งหมายและประโยชน์ของการวัดและการประเมินผล เพื่อสร้างแรงจูงใจในการเรียน เพื่อตรวจสอบความรู้พื้นฐาน เพื่อปรับปรุงการเรียนการสอน เพื่อวินิจฉัยข้อบกพร่อง เพื่อตัดลินผลการเรียน เพื่อจัดตำแหน่งหรือจัดประเภท เพื่อเปรียบเทียบระดับพัฒนาการ เพื่อพยากรณ์หรือทำนาย เพื่อประเมินค่า การวัดและการประเมินผล มีประโยชน์ต่อผู้เรียน ครู ผู้บริหาร ผู้ปกครองและที่สำคัญมีประโยชน์ต่อการแนะแนวและการวิจัย

วิธีการและเครื่องมือวัดและการประเมินผลการเรียนรู้ อาจแบ่งออกตามรูปแบบหรือลักษณะการวัดและการประเมินผลได้เป็น 2 แบบใหญ่ ๆ ได้แก่ แบบเป็นทางการและแบบไม่เป็นทางการ วิธีการประเมินแบบต่าง ๆ ที่ผู้สอนนำไปใช้ในการประเมินเพื่อพัฒนาผู้เรียนนั้นมีหลากหลาย เครื่องมือที่ใช้ในการเก็บข้อมูลของผู้เรียนก็มีหลากหลายเช่นกัน บางกรณีวิธีประเมินอาจใช้เครื่องมือเพียงอย่างเดียว บางกรณีอาจใช้เครื่องมือหลายอย่าง ผู้สอนสามารถเลือกใช้ได้ตามวัตถุประสงค์และความเหมาะสม ในต่างประเทศได้ให้ความสนใจกับการนำผังมโนทัศน์ มาใช้ในการประเมินผลการเรียนรู้ (West, Pomeroy, Park, Gerstenberger and Sandoval, 2000; Besterfield-Sacre, Gerchak, Lyons, Shuman, and Wolfe, 2004; Ozdemir 2005; Srinivasan, McElvany, Shay, Shavelson, and West 2008; Nakiboglu and Ertem 2010; Buldu and Buldu, 2010; Cakmak, 2010; Pishghadam, 2011) สำหรับในแวดวงการศึกษาไทยการนำผังมโนทัศน์มาใช้ในการประเมินผลการเรียนรู้อย่างมีผู้สนใจอยู่น้อยมาก ส่วนมากเป็นการนำผังมโนทัศน์มาใช้ในการจัดกิจกรรมการเรียนการสอนเท่านั้น (พรณี ปานเทวัญ, นงพิมพ์ นิมิตอนันท์ และอารีย์ เสนีย์, 2547; สุจิตรา เทียนสวัสดิ์, อำไพ จารุวัชรพาณิชกุล และยุพิน เพ็ชรมงคล, 2550; ดวงรัตน์ ศรีวงศา, จิราภรณ์ หนูสวัสดิ์ และอินทิรา รอบรู้, 2551; วิภา วิเสโส, 2553) จากที่ทบทวนพบว่า มีเพียง 2 คน คือ วิยะดา ระวังสุข (2545) และ ทัทมณี ชูขวัญ (2548)

จากข้อค้นพบดังกล่าว ผู้วิจัยสนใจนำผังมโนทัศน์มาใช้เป็นเครื่องมือสำหรับประเมินผลการเรียนรู้ซึ่งเป็นอีกทางเลือกหนึ่ง และให้ความสำคัญในเรื่องการพัฒนาเกณฑ์การให้คะแนนผังมโนทัศน์ ในที่นี้ผู้วิจัยขอเสนอแนวคิดและทฤษฎีเกี่ยวกับผังมโนทัศน์ ดังนี้

ตอนที่ 2 แนวคิดและทฤษฎีเกี่ยวกับผังมโนทัศน์

ความหมายของผังมโนทัศน์

ผังมโนทัศน์ไม่ได้เป็นเทคนิคหรือวิธีการที่เกิดขึ้นใหม่ เพราะได้รับการพัฒนาและนำมาใช้ในการศึกษามานานกว่า 30 ปี (Buldu, 2010) โดย Prof. Joseph D. Novak จากมหาวิทยาลัยคอร์เนล ประเทศสหรัฐอเมริกา ในปี ค.ศ. 1972 เป็นบุคคลสำคัญในการคิดค้นและประยุกต์ใช้ในการเรียนการสอน โดยมีแนวคิดที่ว่า “ความรู้ในเรื่องใดก็ตามจะประกอบด้วยมโนทัศน์หลายมโนทัศน์ ซึ่งมโนทัศน์

เหล่านั้นควร ได้มีการจัดความสัมพันธ์กันอย่างมีระบบระเบียบจากมโนทัศน์ที่กว้างและครอบคลุมไปสู่ มโนทัศน์ที่แคบ และเฉพาะเจาะจง โดยระหว่างมโนทัศน์จะมีคำเชื่อม (Linking Words) สำหรับ เชื่อมโยงความสัมพันธ์ระหว่างมโนทัศน์ ซึ่งให้ผู้เรียนเกิดการคิดอย่างต่อเนื่อง เข้าใจความสัมพันธ์ ระหว่างมโนทัศน์อย่างเป็นระบบและเป็นสายลักษณะอักษร” (Novak, 1990, pp. 937-949) จาก การศึกษาพบว่า มีนักการศึกษาหลายท่านที่สนใจเกี่ยวกับผังมโนทัศน์และให้ความหมายไว้ต่าง ๆ กัน ออกไป ดังนี้

Lanzing (2002) ให้ความหมายของผังมโนทัศน์ไว้ว่าเป็นเทคนิคการเป็นตัวแทนองค์ ความรู้ในรูปกราฟ ซึ่งเป็นเครือข่าย (Networks) ของมโนทัศน์ เครือข่ายนั้นแสดงออกมาในรูปของ กลุ่มคำ (Node) และการเชื่อมโยง (Link) ซึ่งอาจเป็นเส้นตรงหรือเส้นโค้งโดยกลุ่มคำจะเป็นตัวแทน ของมโนทัศน์ และการเชื่อมโยงจะเป็นตัวแทนของความสัมพันธ์ระหว่างมโนทัศน์เหล่านั้น

Hsu and Hsieh (2005) ให้ความหมายของผังมโนทัศน์ไว้ว่า เป็นแผนภาพที่แสดง ความสัมพันธ์ระหว่างมโนทัศน์อย่างมีลำดับชั้น เพื่อแสดงให้เห็นถึงการจัดการมโนทัศน์ของข้อมูล ความคิดหรือความรู้เกี่ยวกับเรื่องใดเรื่องหนึ่ง ประกอบด้วยกลุ่มมโนทัศน์ตั้งแต่ 2 มโนทัศน์ขึ้นไป ได้แก่ มโนทัศน์หลัก และมโนทัศน์รอง โดยมโนทัศน์นั้นจะแทนด้วย คำสำคัญ เชื่อมโยงโดยเส้นแสดง ทิศทางหรือความสัมพันธ์ มีคำกริยาที่อธิบายถึงความสัมพันธ์ระหว่างมโนทัศน์ เพื่อให้เกิดการสร้าง องค์ความรู้อย่างเป็นระบบ สามารถอธิบายความคิดที่ซับซ้อน ทำให้ง่ายต่อความเข้าใจ

Novak and Cañas (2006) ให้ความหมายผังมโนทัศน์ไว้ว่า เป็นแผนภาพที่แสดง ความสัมพันธ์ระหว่างแนวคิด เป็นเครื่องมือแบบกราฟิกสำหรับการจัดระเบียบและเป็นตัวแทนความรู้ แนวคิดที่แสดงมักจะเป็นกล่องสี่เหลี่ยมหรือวงกลมมีการลากเส้นเชื่อมโยงความสัมพันธ์ระหว่าง แนวคิดและเขียนคำหรือข้อความเพื่อแสดงความสัมพันธ์ระหว่างแนวคิดไว้ที่เส้นเชื่อมโยง

Rathod (2012, p. 133) ให้ความหมายผังมโนทัศน์ไว้ว่า เป็นแผนภาพแสดงความสัมพันธ์ ระหว่างแนวคิด สำหรับการจัดระเบียบและเป็นตัวแทนความรู้ ซึ่งปกติจะล้อมด้วยวงหรือกล่อง และ ความสัมพันธ์ระหว่างแนวคิดมีเส้นเชื่อมต่อการเชื่อมโยงอย่างน้อย 2 แนวคิด และการเชื่อมโยงจะมี ป้ายข้อความสั้น ๆ ที่สำคัญ หรือที่เพิ่มความหมาย

Safdar, Hussain, Shah and Rifat (2012, p. 58), Sunita (2012, p. 137) ใ้ ให้ความหมายผังมโนทัศน์ไว้ว่า เป็นแผนภาพแสดงความสัมพันธ์ระหว่างแนวคิดในการจัดระเบียบ ความรู้ โดยมีวัตถุประสงค์เพื่อเป็นตัวแทนของโครงสร้างความรู้ที่เก็บไว้ในตัวผู้เรียน ซึ่งปกติจะ ล้อมรอบด้วยวงรีหรือกล่อง และความสัมพันธ์ระหว่างแนวคิดมีเส้นเชื่อมต่อการเชื่อมโยงอย่างน้อย 2 แนวคิดและการเชื่อมโยงจะมีป้ายข้อความสั้น ๆ ที่สำคัญ หรือเพื่อเพิ่มความหมาย

จากการศึกษาสามารถสรุปความหมายของผังมโนทัศน์ได้ว่า เป็นแผนภาพที่แสดงความสัมพันธ์ระหว่างมโนทัศน์ ซึ่งปกติจะล้อมด้วยวงหรือกล่อง มีการเชื่อมโยงอย่างน้อย 2 มโนทัศน์ การเชื่อมโยงจะมีคำหรือวลี เพื่อแสดงถึงความสัมพันธ์ระหว่างมโนทัศน์

ทฤษฎีการเรียนรู้ที่เป็นพื้นฐานของการสร้างผังมโนทัศน์

ผังมโนทัศน์เป็นผังที่แสดงความสัมพันธ์ระหว่างมโนทัศน์ที่เกี่ยวข้องอย่างมีลำดับชั้น จากสิ่งที่มีความหมายกว้างไปสู่สิ่งที่มีความหมายแคบลง และมีความเฉพาะเจาะจงมากขึ้น มีพื้นฐานมาจากทฤษฎีการเรียนรู้ที่มีความหมาย (Meaningful Verbal Learning) ของ Ausubel (1968) ก่อนที่จะกล่าวถึงลักษณะของทฤษฎีการเรียนรู้ที่มีความหมายของ Ausubel จะกล่าวถึงชนิดของการเรียนรู้ซึ่ง Klausmeier and Ripple (1971) ได้แบ่งไว้อย่างชัดเจน ดังนี้

ชนิดของการเรียน Klausmeier and Ripple (1971) แบ่งการเรียนรู้เป็น 2 ชนิดคือ

1. การเรียนแบบรับรู้ (Reception Learning)
2. การเรียนแบบค้นพบ (Discovery Learning)

การเรียนรู้แต่ละแบบยังแบ่งเป็นการเรียนรู้ที่มีความหมาย และการเรียนแบบท่องจำ (Rote Learning) จึงทำให้การเรียนรู้แบ่งเป็น 4 ชนิดคือ

1. การเรียนแบบรับรู้ที่มีความหมาย
2. การเรียนแบบรับรู้โดยการท่องจำ
3. การเรียนแบบค้นพบที่มีความหมาย
4. การเรียนแบบค้นพบโดยการท่องจำ

การเรียนแบบรับรู้ เนื้อหาสาระทั้งหมดที่จะต้องเรียนรู้ ครูจะเป็นผู้บรรยายและบอกให้ทั้งหมด แต่ในการเรียนแบบค้นพบ สิ่งที่จะเรียนรู้จะค้นพบในตอนท้ายของการเรียน และบางอย่างนักเรียนจะต้องค้นหาเอง ผู้เรียนจะต้องนำข้อมูลที่ได้รับใหม่ไปบูรณาการกับสิ่งที่อยู่ในโครงสร้างของความรู้ที่มีอยู่ และจัดโครงสร้างใหม่หรือขยายโครงสร้างเดิม

การรับรู้หรือการค้นพบ เป็นขั้นแรกของการเรียนรู้ ขั้นต่อมานักเรียนจะต้องนำข้อมูลที่ได้จดจำไว้ใช้ต่อไป ถ้านักเรียนตั้งใจจะให้ข้อมูลที่ได้รับใหม่เกิดความคงทน จำไว้นานโดยการนำไปสัมพันธ์กับสิ่งที่รู้มาก่อนแล้ว จะทำให้เกิดการเรียนรู้ที่มีความหมาย ถ้านักเรียนตั้งใจจะจำข้อมูลที่ได้รับใหม่โดยไม่นำไปสัมพันธ์กับความรู้เดิมจะเกิดการเรียนรู้อย่างท่องจำ ดังนั้นจากชนิดของการเรียนรู้ทั้ง 4 ชนิด สามารถนำมาอธิบายได้ดังนี้คือ

1. การเรียนรู้อย่างมีความหมาย เป็นการเรียนรู้ที่ได้รับการสอนสิ่งใหม่ ๆ อย่างครบถ้วนและผู้เรียนนำไปสัมพันธ์กับความรู้เดิมที่มีอยู่
2. การเรียนแบบรับรู้โดยการท่องจำ เป็นการเรียนที่ผู้เรียนได้รับการสอนสิ่งใหม่ ๆ อย่าง

ครบถ้วนและผู้เรียนท่องจำไว้

3. การเรียนแบบค้นพบอย่างมีความหมาย เป็นการเรียนที่ผู้เรียนค้นพบคำตอบและนำไปสัมพันธ์กับความรู้เดิมที่มีอยู่

4. การเรียนแบบค้นพบโดยท่องจำ เป็นการเรียนที่ผู้เรียนค้นพบด้วยตนเองแต่ท่องจำไว้ การนำข้อมูลใหม่เข้าสู่โครงสร้างของความรู้ จะใช้การดูดซึมเข้าไปเป็นส่วนหนึ่งของโครงสร้างความรู้ที่มีอยู่เดิม โดยกระบวนการที่เรียกว่า กระบวนการดูดซึม (Subsumption) การเรียนรู้อย่างมีความหมายทั้งชนิดรับรู้และค้นพบ เมื่อเข้าไปเป็นส่วนหนึ่งของโครงสร้างความรู้แล้ว แม้ว่าจะไม่สามารถจะจดจำได้ทั้งหมด แต่ก็สามารถจะระลึกย้อนหลังถึงสิ่งที่เรียนรู้แล้วแต่ไม่ได้นำมาใช้นานแล้ว และสามารถเรียนรู้ใหม่ได้โดยใช้เวลาน้อยกว่าเมื่อเริ่มต้นครั้งแรก

ทฤษฎีการเรียนรู้ที่เป็นพื้นฐานของการสร้างผังมโนทัศน์คือ ทฤษฎีการเรียนรู้ที่มีความหมายของ Ausubel ซึ่งมีแนวคิดที่ว่าครูควรจะสอนสิ่งที่มีความสัมพันธ์กับความรู้ที่นักเรียนมีอยู่เดิม ความรู้ที่มีอยู่เดิมนี้อยู่ในโครงสร้างของความรู้ (Cognitive Structure) ซึ่งเป็นข้อมูลที่เหมาะสมอยู่ในสมองและมีการจัดระบบไว้อย่างดี มีการเชื่อมโยงระหว่างความรู้เก่าและความรู้ใหม่อย่างมีระดับชั้น ดังนั้นโครงสร้างของความรู้จะใช้เป็นกรอบมโนทัศน์และใช้บันทึกประสบการณ์ต่าง ๆ ที่ได้รับ

โดยสรุปการเรียนรู้อย่างมีความหมาย จะเกิดขึ้นเมื่อความรู้ใหม่เชื่อมกับมโนทัศน์ที่อยู่ในโครงสร้างของความรู้เดิมที่มีอยู่ในสมอง ซึ่ง Ausubel เรียกว่ากระบวนการดูดซึม หรือเรียกมโนทัศน์ที่เกิดจากการเชื่อมโยงนั้นว่า Subsumer แต่ถ้าไม่ได้นำความรู้ใหม่เข้าไปเชื่อมโยงกับความรู้เดิมที่มีอยู่จะเป็นการเรียนรู้แบบท่องจำ

การสร้างผังมโนทัศน์มีพื้นฐานมาจากทฤษฎีการเรียนรู้ของ Ausubel 3 ประการ คือ

1. โครงสร้างของความรู้ (Cognitive Structure) ซึ่งเป็นโครงสร้างที่มีอยู่ในสมอง จะมีการจัดลำดับมโนทัศน์ จากมโนทัศน์ที่มีความหมายกว้างทั่วไป ไปสู่มโนทัศน์ที่แคบลงและมีความเฉพาะเจาะจงมากขึ้น

2. กระบวนการแยกแยะความแตกต่างเชิงก้าวหน้า (Progressive Differentiation) จากหลักการของ Ausubel ที่กล่าวว่า การเรียนรู้อย่างมีความหมายจะเกิดขึ้นเมื่อมีการนำความรู้ใหม่ไปสัมพันธ์กับความรู้ที่มีอยู่เดิม เกิดเป็นความสัมพันธ์ใหม่ ดังนั้นจึงเกิดการเรียนรู้อย่างไม่สิ้นสุด จะเป็นการขยายความรู้ให้กว้างขึ้นจนกลายเป็นการแยกแยะความแตกต่างเชิงก้าวหน้า โดยประกอบด้วยมโนทัศน์ที่มีความหมายกว้างอยู่ด้านบนของโครงสร้างความรู้และมโนทัศน์ที่มีความเฉพาะเจาะจงอยู่ด้านล่างมา กระบวนการแยกแยะความแตกต่างเชิงก้าวหน้าจะเพิ่มขึ้นถ้าผู้เรียนมีโอกาสอภิปรายร่วมกัน และจะทำให้เห็นความเกี่ยวข้องและความสัมพันธ์ของสิ่งต่าง ๆ ที่เรียนขึ้นได้

3. การประสานสัมพันธ์เชิงบูรณาการ (Integrative Reconciliation) จากหลักการเรียนรู้ของ Ausubel ที่กล่าวว่าการเรียนรู้ที่มีความหมายจะเกิดจากการเชื่อมโยงความรู้ใหม่เข้ากับความรู้เดิมที่มีอยู่ ดังนั้นถ้าผู้เรียนสามารถเชื่อมโยงมโนทัศน์ ทำให้เกิดความสัมพันธ์ใหม่และเชื่อมโยงระหว่างชุดของมโนทัศน์จะทำให้เกิดการประสานเชิงบูรณาการของมโนทัศน์ซึ่งจะทำให้เกิดการเรียนรู้ที่มีความหมายขึ้น

จากหลักการเรียนรู้ทั้ง 3 ประการนี้ได้นำไปเป็นพื้นฐานในการสร้างผังมโนทัศน์ และยังใช้เป็นพื้นฐานในการให้คะแนนผังมโนทัศน์ดังจะกล่าวต่อไป

เกณฑ์การให้คะแนนผังมโนทัศน์

เกณฑ์การให้คะแนนผังมโนทัศน์ ใช้หลักการประเมินผลการเรียนรู้ตามทฤษฎีการเรียนรู้ที่มีความหมายของ Ausubel ซึ่งทำให้การให้คะแนนผังมโนทัศน์มีความตรงเชิงโครงสร้าง (Construct Validity) (Novak ,1984) มีขั้นตอนการให้คะแนนดังนี้

1. นับความสัมพันธ์ทั้งหมดที่สมเหตุสมผล (Valid) และให้คะแนนความสัมพันธ์ละ 1 คะแนน
2. นับจำนวนการเรียงลำดับชั้น การให้คะแนนลำดับชั้นจะนำเลขตัวใดมาคูณกับลำดับชั้นนั้นจะขึ้นอยู่กับความพอใจของผู้สอนไม่มีกฎเกณฑ์แน่นอน ซึ่งอยู่ระหว่าง 3-10 เท่าของความสัมพันธ์แล้วนำมาคูณกับจำนวนลำดับชั้นที่ผู้เรียนสร้างขึ้น แต่จะมีปัญหาถ้าลำดับชั้นของมโนทัศน์มีจำนวนไม่สมดุลกัน ซึ่งจะใช้จำนวนของลำดับชั้นที่มีแขนงสาขามากที่สุดมานับเป็นจำนวนลำดับชั้น และจะไม่ให้คะแนนถ้ามีการจัดลำดับชั้นที่ไม่ชัดเจน

3. การเชื่อมโยงระหว่างสายของมโนทัศน์ที่แสดงความสัมพันธ์อย่างสมเหตุสมผล จะให้คะแนน 2-10 เท่าของคะแนนที่ให้แก่แต่ละระดับ คูณด้วยจำนวนความสัมพันธ์ที่เกิดจากการเชื่อมโยงระหว่างสายของมโนทัศน์การเชื่อมโยงระหว่างสายของมโนทัศน์ การเชื่อมโยงระหว่างสายของมโนทัศน์นี้ อาจจะทำให้นักเรียนได้อธิบายเหตุผล ซึ่งจะเป็นประโยชน์แก่นักเรียนเป็นอย่างมากสำหรับการเริ่มสร้างกรอบมโนทัศน์

4. ให้นักเรียนยกตัวอย่างเหตุการณ์หรือวัตถุที่อยู่ในกรอบมโนทัศน์เพื่อให้เกิดความแน่ใจว่านักเรียนเข้าใจถูกต้อง และให้คะแนนเช่นเดียวกับความสัมพันธ์อื่น ๆ คือ 1 คะแนน หรืออาจจะให้ครึ่งคะแนน เพราะทำได้ง่ายกว่าการหาความสัมพันธ์หรือนักเรียนอาจจะใช้วิธีท่องจำตัวอย่างมา

เกณฑ์ในการให้คะแนนผังมโนทัศน์

1. ประพจน์ (Propositions) ประพจน์ที่แสดงความสัมพันธ์ระหว่าง 2 มโนทัศน์ที่เชื่อมกันโดยใช้คำเชื่อม และทำให้ประพจน์สมเหตุสมผลจะได้คะแนน 1 คะแนน
2. การจัดลำดับ (Hierarchy) มโนทัศน์ที่อยู่รองลงมาจะเป็นมโนทัศน์ที่เฉพาะเจาะจงและ

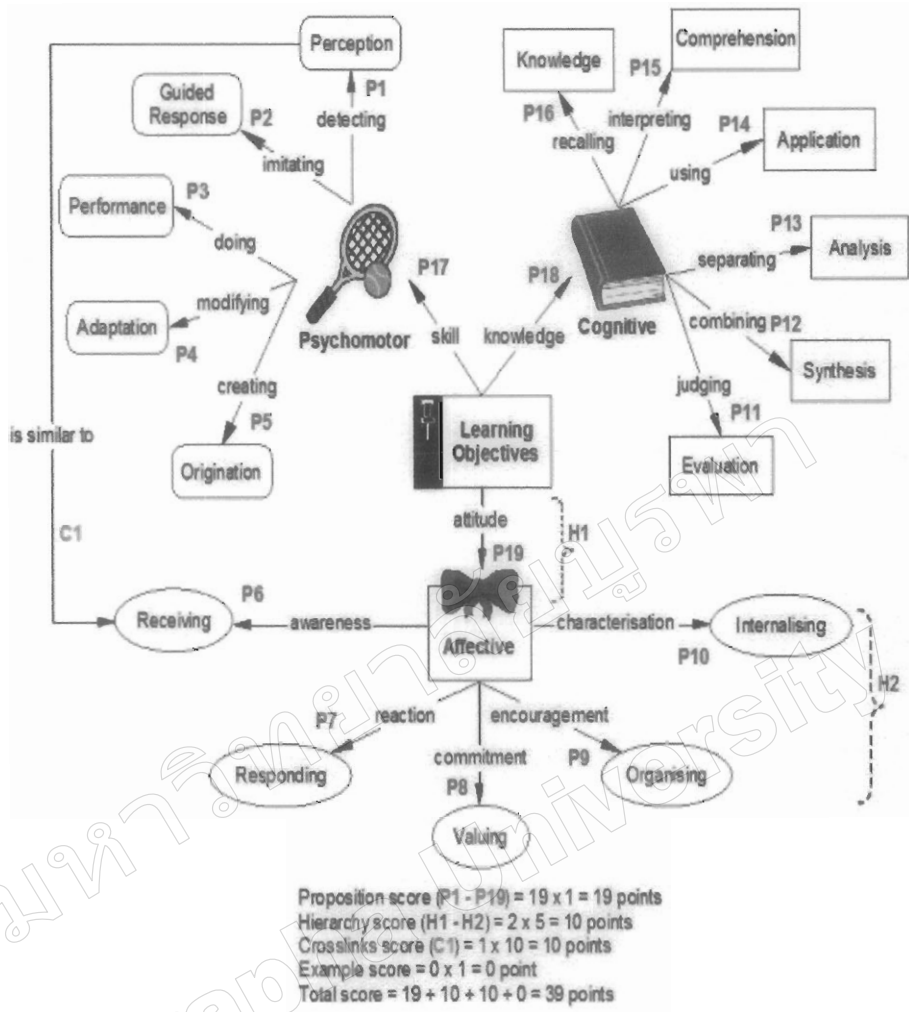
มีความกว้างน้อยกว่ามโนทัศน์ที่อยู่ลำดับแรก จะได้คะแนน 5 คะแนน ของทุกระดับที่จัดลำดับได้
สมเหตุสมผล

3. การเชื่อมโยงข้ามสายของมโนทัศน์ (Cross Link) การเชื่อมโยงระหว่างชุดของมโนทัศน์ แสดงให้เห็นความสามารถในการคิดสร้างสรรค์ จึงควรได้รับคะแนนพิเศษถ้าความสัมพันธ์ถูกต้องจะได้ 10 คะแนน

4. ตัวอย่าง (Example) ตัวอย่างของวัตถุหรือเหตุการณ์จะได้คะแนนตัวอย่างละ 1 คะแนน (การเขียนไม่ต้องวงกลมล้อมรอบเพราะไม่ใช่มโนทัศน์)

เกณฑ์การให้คะแนนอาจสร้างขึ้นใหม่ได้ขึ้นอยู่กับสิ่งที่นำมาสร้างผังมโนทัศน์ อาจแบ่งคะแนนออกเป็นส่วน ๆ และนำมาเปรียบเทียบกันโดยทำเป็นคะแนนร้อยละ นักเรียนบางคนอาจจะทำได้ดีกว่าเกณฑ์ที่ตั้งไว้ ทำได้คะแนนมากกว่า 100% ก็ได้ ตัวอย่างการตรวจให้คะแนนผังมโนทัศน์ของ Novak and Gowin (1984) ดังภาพที่ 2-1

มหาวิทยาลัยบูรพา
Burapha University



ภาพที่ 2-1 ตัวอย่างการตรวจให้คะแนนผังมโนทัศน์ของ Novak and Gowin (1984)

จากที่ผู้วิจัยศึกษาพบว่า เกณฑ์การให้คะแนนผังมโนทัศน์ของ Novak and Gowin (1984) จะมีลักษณะการให้คะแนนแบบแยกองค์ประกอบเพื่อให้เห็นคุณภาพของการปฏิบัติงานรายละเอียดปลีกย่อย การประเมินผลการเรียนรู้จึงมีการวิเคราะห์แยกมิติหรือองค์ประกอบที่จะพิจารณาให้ระดับคะแนน ซึ่งแต่ละมิติหรือองค์ประกอบก็จะแบ่งระดับตามระดับคุณภาพความสามารถ เกณฑ์การให้คะแนนของ Novak and Gowin (1984) จะต้องรวมคะแนนเป็นคะแนนรวมทั้งหมดของผลงาน การตรวจให้คะแนนด้วยเกณฑ์นี้ต้องใช้เวลามาก และผู้ประเมินจะต้องมีความรู้ในเรื่องที่ประเมินอย่างแท้จริง (XiuFeng & Hinchey, 1996; McClure, Sonak & Suen, 1999; Kinchin et al., 2000; Klein, Chung, Osmundson, Herl & O' Neil, 2002)

ส่วนประกอบและลักษณะของผังมโนทัศน์

ส่วนประกอบของผังมโนทัศน์นั้น Novak (1993) กล่าวไว้ว่า ผังมโนทัศน์ประกอบด้วย ส่วนประกอบสำคัญ 4 ส่วน ได้แก่

1. มโนทัศน์ (Concept) หมายถึง คำที่ใช้แทนชื่อของมโนทัศน์ เป็นคำ หรือวลีสั้น ๆ
2. ความสัมพันธ์ (Relationship) หรือการเชื่อมโยงระหว่างประพจน์ (Propositional Linkage) เป็นการลากเส้นเชื่อมโยงระหว่างคำมโนทัศน์
3. ลำดับชั้น (Hierarchy) เป็นชั้นหรือลำดับชั้นของความสัมพันธ์ระหว่างมโนทัศน์หลัก ไปยังมโนทัศน์รองหรือมโนทัศน์ย่อย

4. การเชื่อมโยงข้ามมโนทัศน์ (Cross-Link) เป็นชั้นหรือลำดับชั้นของความสัมพันธ์ เชื่อมโยงข้ามชั้นหรือข้ามชั้นระหว่างมโนทัศน์ จากมโนทัศน์หนึ่งไปยังอีกมโนทัศน์หนึ่ง

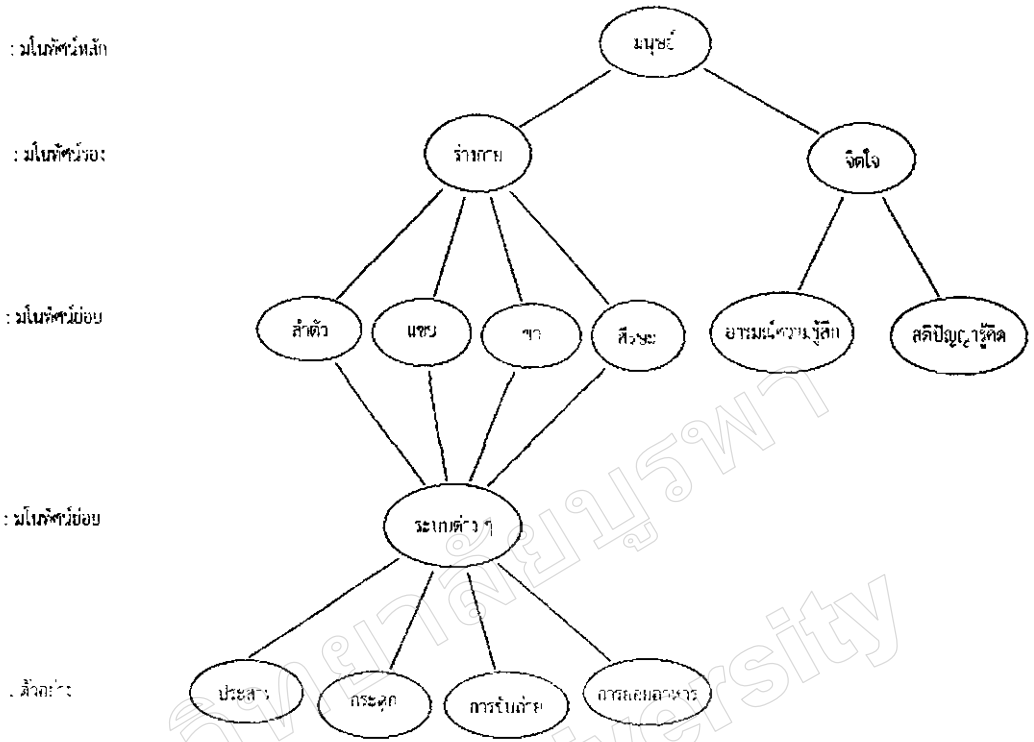
Plotnick (1997), Baroody and Bartels (2001), Classidy, Griffiths and Nakoechny (2001) ได้กล่าวถึงส่วนประกอบของผังมโนทัศน์ไว้ว่า ผังมโนทัศน์จะต้องมีส่วนประกอบ 3 ส่วนที่สำคัญดังนี้

1. คำมโนทัศน์ (Concept) เป็นชื่อมโนทัศน์ของเรื่องที่น่ามาสร้างผังมโนทัศน์ซึ่งอาจจะ เป็นคำหรือวลีหรือประโยคสั้น ๆ โดยจะเขียนไว้ในกรอบซึ่งอาจเป็นรูปร่างกลมหรือวงรี หรือสี่เหลี่ยมก็ได้
2. เส้นเชื่อมโยงระหว่างมโนทัศน์ (Linking Lines) เป็นเส้นที่ลากเชื่อมโยงระหว่างคำมโนทัศน์สองมโนทัศน์ โดยเส้นเชื่อมโยงความสัมพันธ์นี้อาจจะเป็นเส้นตรง เส้นโค้ง บางครั้งอาจแสดง ทิศทางความสัมพันธ์ด้วยหัวลูกศรทางเดียว หรือสองทิศทางกำกับไว้ด้วย
3. คำเชื่อม (Linking Phrase) เป็นคำที่แสดงถึงความสัมพันธ์ระหว่างมโนทัศน์สอง มโนทัศน์ โดยคำเชื่อมโยงนั้นจะต้องเป็นคำที่ทำให้เกิดประพจน์ที่มีความหมาย อาจเป็นถ้อยคำ วลี หรือประโยค ซึ่งต้องเป็นคำที่สั้น กระชับรัด ทำให้ผู้อ่านเข้าใจง่าย

จากส่วนประกอบของผังมโนทัศน์ที่กล่าวมาข้างต้น สามารถสรุปส่วนประกอบที่สำคัญของ ผังมโนทัศน์ได้ดังนี้คือคำมโนทัศน์ (Concept) คำเชื่อม (Linking Phrase) และเส้นเชื่อมระหว่าง มโนทัศน์ (Linking lines)

ลักษณะของผังมโนทัศน์

ลักษณะของผังมโนทัศน์มีส่วนประกอบหลายส่วนประกอบ Novak (1993) ได้เสนอ ลักษณะผังมโนทัศน์อย่างง่ายที่แสดงความสัมพันธ์ระหว่างมโนทัศน์ระดับต่าง ๆ ไว้ดังภาพที่ 2-2



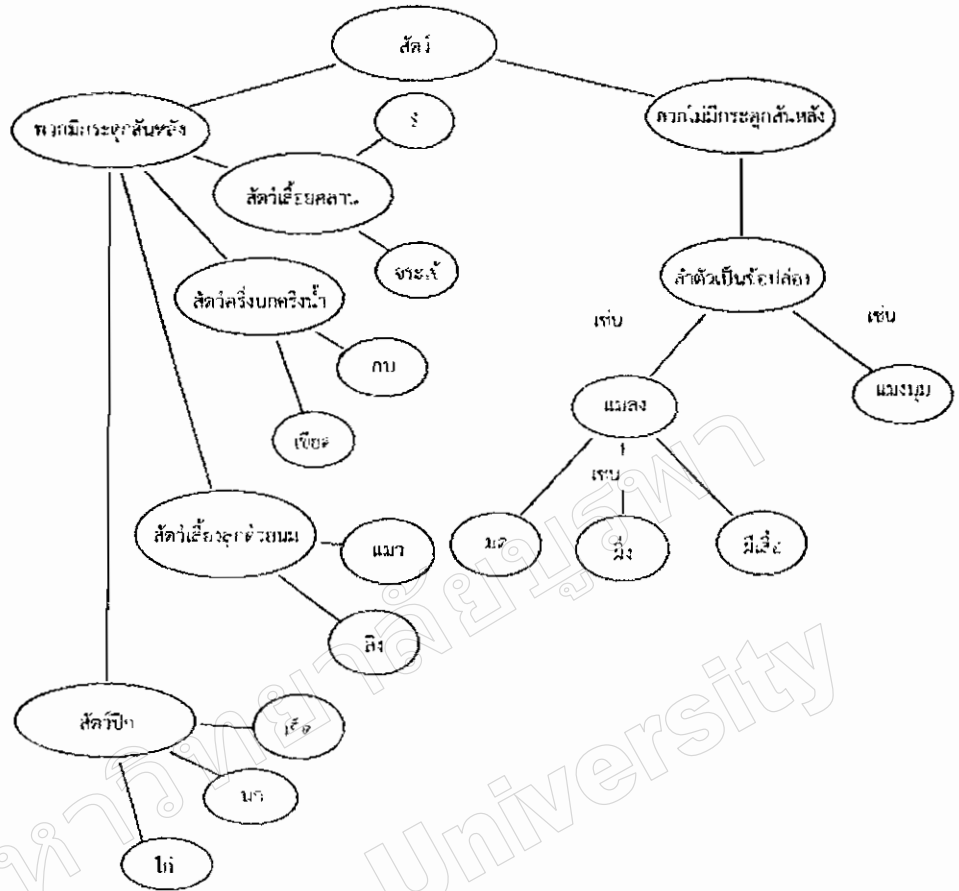
ภาพที่ 2-2 ผังมโนทัศน์อย่างง่ายที่แสดงความสัมพันธ์ระหว่างมโนทัศน์ (Novak, 1984)

ประเภทของผังมโนทัศน์ มีหลายรูปแบบซึ่งมีชื่อเรียกแตกต่างกันไป สำหรับการนำผังมโนทัศน์มาใช้ขึ้นอยู่กับลักษณะข้อมูลหรือวัตถุประสงค์ของการศึกษา นักการศึกษาหลายท่านได้จัดแบ่งประเภทผังมโนทัศน์โดยใช้เกณฑ์ต่างกันไปหลายแนวคิด โดยมีลักษณะที่หลากหลายดังนี้

Mintzes, Wandersee and Novak (1998) ได้แบ่งผังมโนทัศน์ออกเป็น 2 ชนิด ได้แก่

1. ผังหลัก (Macro Map) ซึ่งแสดงความสัมพันธ์ระหว่างมโนทัศน์ที่สำคัญ ๆ เท่านั้น
2. ผังย่อย (Micro Map) แสดงรายละเอียดเพิ่มเติม เฉพาะมโนทัศน์องค์ประกอบของผัง

หลัก ดังภาพที่ 2-3



ภาพที่ 2-3 ผังมโนทัศน์ย่อยเรื่องสัตว์

Classidy, Griffiths and Nakoechny (2001) ได้จำแนกผังมโนทัศน์ออกเป็น 3 ประเภท ดังนี้

1. ผังต้นไม้ (Tree Map) โดยมโนทัศน์จะถูกจัดเป็นหมวดหมู่ การเขียนมโนทัศน์จะเริ่มจากส่วนบนของกระดาษ และเขียนมโนทัศน์อื่น ๆ ลดหลั่นกันลงไป
2. ผังที่มีลักษณะเป็นเว็บ (Wep Map) เป็นผังที่ไม่มีลำดับขั้นตอน ซึ่งสามารถเขียนประพจน์เชื่อมกันได้
3. ผังแบบวง (Circle Map) ผังมโนทัศน์ประเภทนี้จะเขียนมโนทัศน์ไว้ตรงกลางแล้วเขียนประพจน์ออกจากจุดศูนย์กลางออกไปด้านข้าง

จากที่กล่าวมาจะเห็นว่าผังมโนทัศน์มีหลายประเภทจะมีการนำเสนอที่แตกต่างกันไปขึ้นอยู่กับผู้เลือกว่าจะใช้ประเภทใดจึงจะเหมาะสมกับเรื่องที่ตนสนใจ ซึ่งสามารถสรุปประเภทของ

ผังมโนทัศน์ได้ 3 ประเภท คือ 1) ผังที่มีลำดับขั้นตอน 2) ผังที่ไม่มีลำดับขั้นตอน และ 3) ผังที่มีลักษณะเป็นแบบวง

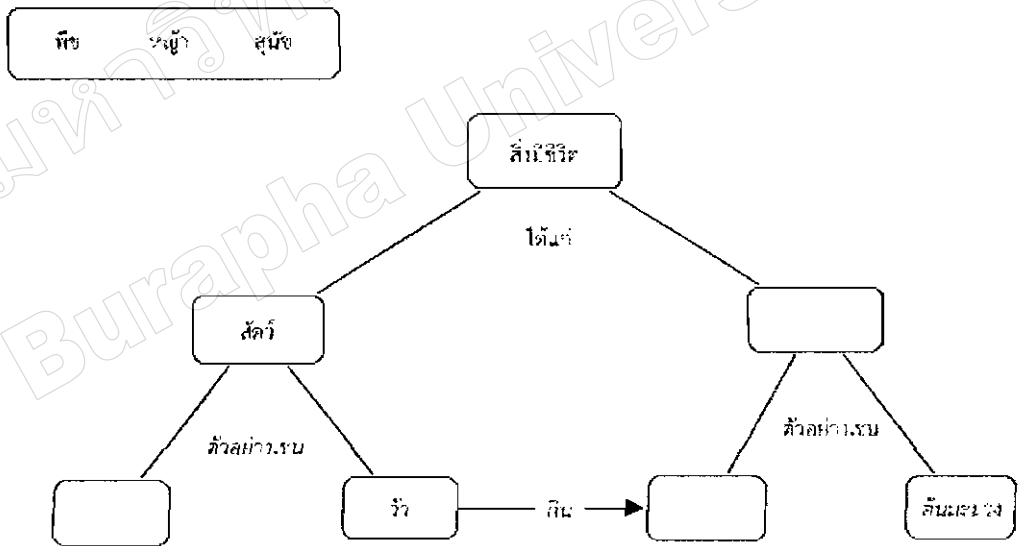
เทคนิคผังมโนทัศน์

เทคนิคผังมโนทัศน์มีหลายประเภท ในการนำมาใช้ก็ขึ้นอยู่กับวัตถุประสงค์ของการวิจัย ซึ่งนักการศึกษาหลายท่านได้นำเสนอไว้ดังนี้

Ruiz-Primo, Schultz, Li and Shavelson (2001) ได้ศึกษาเปรียบเทียบประเภทของเทคนิคผังมโนทัศน์ไว้ 2 ประเภท คือ

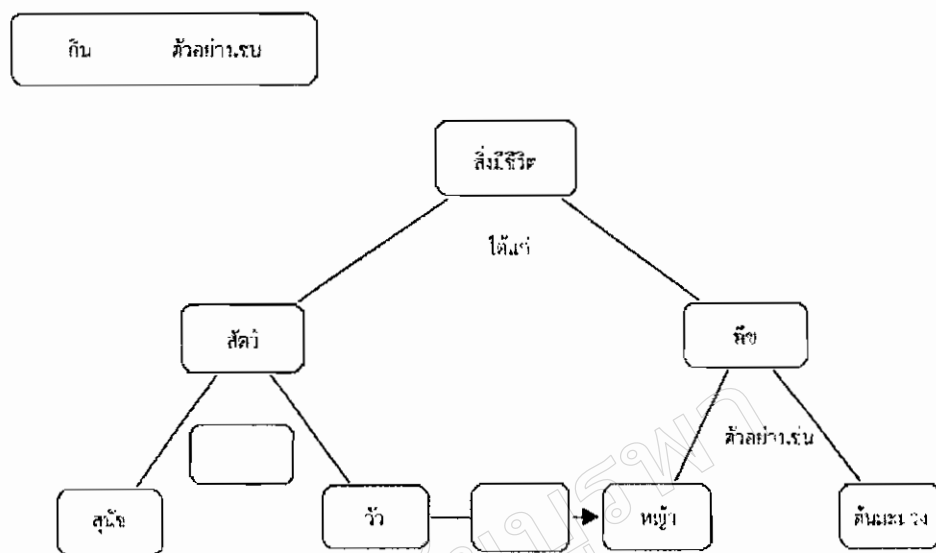
1. เทคนิคผังมโนทัศน์แบบเติมคำลงในผัง (Fill-in-the-Map) เป็นเทคนิคการจัดผังมโนทัศน์ที่มีการกำหนดคำมโนทัศน์ทั้งหมดมาให้ โดยกำหนดเป็นผังโครงสร้างที่ยังไม่สมบูรณ์แล้วให้นักเรียนนำมโนทัศน์ที่กำหนดให้ทั้งหมดมาเติมลงในช่องว่างให้สมบูรณ์ซึ่งแบ่งเป็น 2 ลักษณะ คือ

1.1 แบบเติมคำมโนทัศน์ (Fill-in-the-Nodes) เป็นการจัดผังมโนทัศน์โดยการเติมมโนทัศน์ในช่องว่างของผังโครงสร้างให้สมบูรณ์ ดังภาพที่ 2-4



ภาพที่ 2-4 ผังมโนทัศน์แบบเติมคำมโนทัศน์

1.2 แบบเติมคำเชื่อมบนเส้น (Fill-in-the-Lines) เป็นการจัดผังมโนทัศน์โดยการเติมคำเชื่อมลงบนเส้นโยงความสัมพันธ์ระหว่างมโนทัศน์ที่ว่างให้สมบูรณ์ ดังภาพที่ 2-5



ภาพที่ 2-5 ผังมโนทัศน์แบบเติมคำเชื่อมบนเส้น

2. เทคนิคการจัดผังมโนทัศน์แบบสร้างผังจากคำ (Construct-a-Map) เป็นการจัดผังมโนทัศน์โดยกำหนดรายการของมโนทัศน์บางส่วนมาให้ โดยให้นักเรียนนำคำมโนทัศน์เหล่านั้นมาสร้างเป็นผังมโนทัศน์ให้สมบูรณ์ ดังภาพที่ 2-6

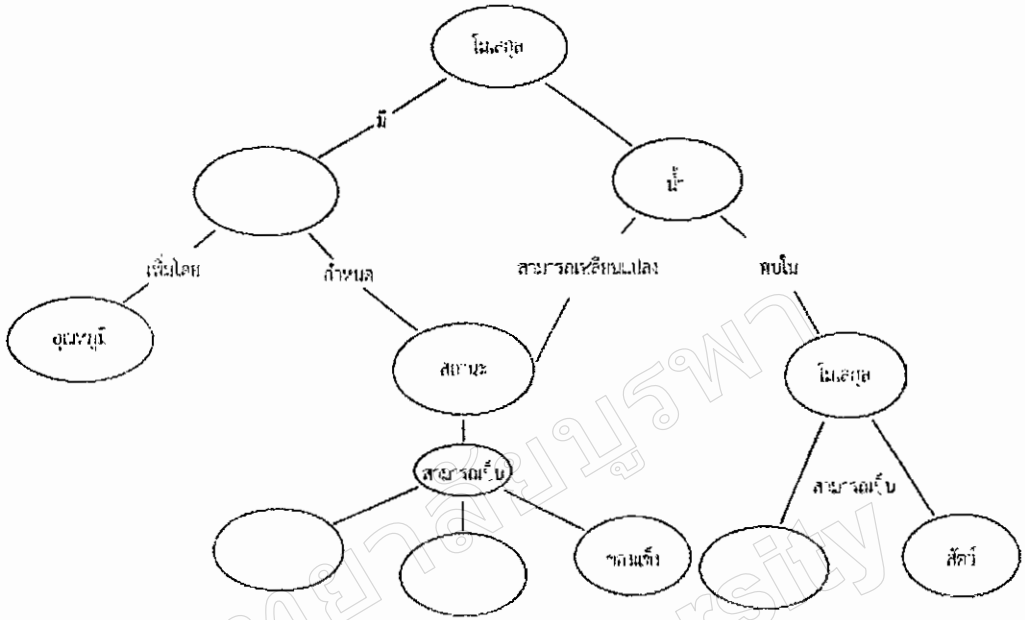
พืช สิ่งมีชีวิต หญ้า วัว สุนัข ต้นมะม่วง

ภาพที่ 2-6 เทคนิคการจัดผังมโนทัศน์แบบสร้างผังจากคำ

Baroody and Bartels (2001) ได้แบ่งประเภทของผังออกเป็น 4 ประเภทดังนี้

1. แบบเติมมโนทัศน์ (Fill-in-Task) เป็นการจัดผังมโนทัศน์โดยมีการกำหนดมโนทัศน์ทั้งหมดมาให้ และมีการจัดผังโครงสร้างที่ไม่สมบูรณ์มาให้ แล้วให้นักเรียนนำมโนทัศน์ที่กำหนดให้มาเติมในผังมโนทัศน์ให้สมบูรณ์ ดังภาพที่ 2-7

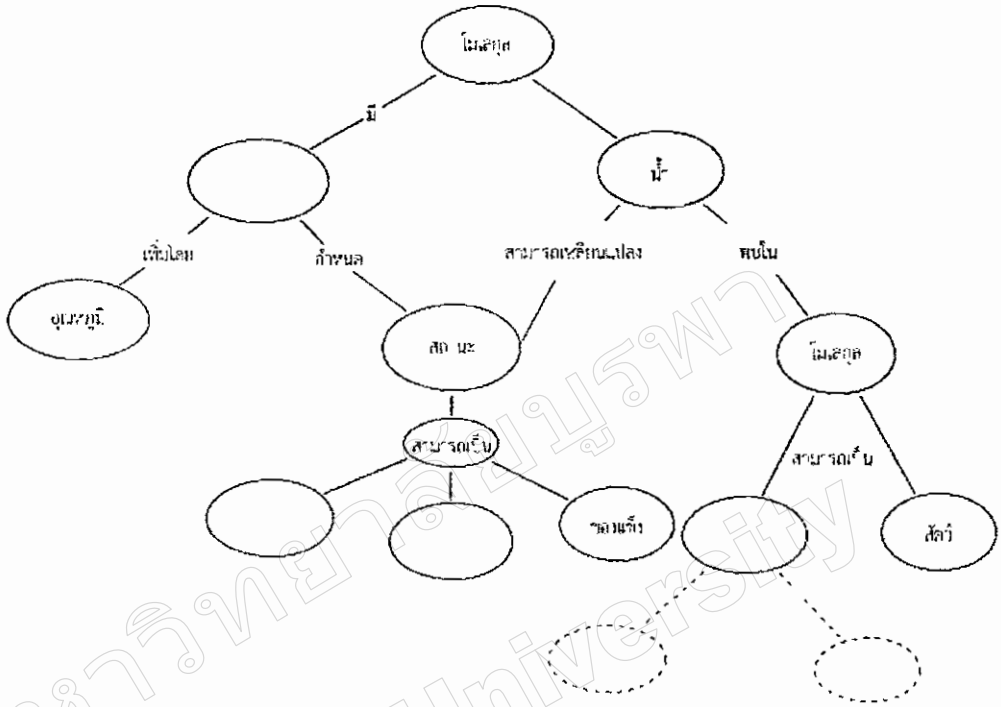
ของเหลว พืช สิ่งมีชีวิต การเคลื่อนไหว



ภาพที่ 2-7 แบบเติมมนิทัศน์

2. แบบเพิ่มมนิทัศน์ (Add-on-Task) เป็นการจัดผังมนิทัศน์โดยมีมนิทัศน์มาให้และจัดผังโครงสร้างที่ยังไม่สมบูรณ์มาให้ แล้วให้นักเรียนนำมนิทัศน์ที่กำหนดมาเติมลงในช่องว่างของผังโครงสร้าง โดยนักเรียนสามารถคิดมนิทัศน์ขึ้นมาเพิ่มเติมตามความเข้าใจ และเพื่อให้ผังมนิทัศน์มีความสมบูรณ์ยิ่งขึ้น ดังภาพที่ 2-8

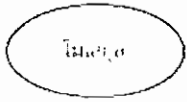
ของเหลว พืช สิ่งมีชีวิต การเคลื่อนไหว ก๊าซ



ภาพที่ 2-8 ผังมโนทัศน์แบบเพิ่มโน้ตทัศน์ (Add-on-Task)

3. แบบปลายปิด (Close-Ended List Task) เป็นการจัดผังมโนทัศน์โดยกำหนดมโนทัศน์มาให้ให้นักเรียนจัดผังมโนทัศน์โดยใช้มโนทัศน์ทั้งหมดที่กำหนดมาให้และเขียนความสัมพันธ์ระหว่างมโนทัศน์ซึ่งทำให้การจัดผังมโนทัศน์สมบูรณ์ดังภาพที่ 2-9

ของเหลว พืช สิ่งมีชีวิต การเคลื่อนไหว ก๊าซ สถานะ อเนกภูมิ ขงจวงไข่ ของเหลว สัตว์ น้ำ



ภาพที่ 2-9 ผังมโนทัศน์แบบปลายปิด

4. แบบปลายเปิด (Open-Ended List Task) เป็นการจัดผังมโนทัศน์โดยกำหนดมโนทัศน์บางส่วนมาให้ โดยใช้คำมโนทัศน์ที่กำหนดให้และสามารถเพิ่มมโนทัศน์และเขียนความสัมพันธ์ระหว่างมโนทัศน์ ซึ่งทำให้ผังมโนทัศน์นั้นสมบูรณ์ยิ่งขึ้นดังภาพที่ 2-10

ภาพที่ 2-10 ผังมโนทัศน์แบบปลายเปิด

การสร้างผังมโนทัศน์

แม้ว่าผังมโนทัศน์จะมีประโยชน์มาก แต่ก็มีข้อจำกัดบางประการ โดยเฉพาะในเรื่องการนำไปใช้กับนักเรียนจำนวนมาก เนื่องจากต้องใช้เวลาและยุ่งยากสำหรับผู้ที่ยังไม่มีความรู้เกี่ยวกับผังมโนทัศน์มาก่อน ต้องใช้เวลาในการประเมิน การสร้างผังมโนทัศน์เป็นงานที่ต้องมีความรู้ ความเข้าใจ ดังนั้นก่อนการใช้จึงต้องมีการฝึกอบรม (Zeilik, 2000, p. 3; Piá, Blasco-Tamarit & Muñoz-Portero, 2011, p. 97) ซึ่งสอดคล้องกับแนวคิดของ Novak, Gowin and Johansen (1983) ที่กล่าวว่า การสร้างผังมโนทัศน์สามารถสร้างได้หลายวิธี ในแต่ละวิธีจะเริ่มต้นด้วยการแนะนำแนวคิดเกี่ยวกับผังมโนทัศน์มโนทัศน์ อาจทำในรูปกิจกรรมการเรียนรู้หรือการแนะนำโดยตรงโดยการให้คำจำกัดความของมโนทัศน์ และวัตถุประสงค์ในการสร้างผังมโนทัศน์ จากการศึกษาแนวคิดของนักการศึกษาหลายท่าน สามารถแยกวิธีการสร้างผังมโนทัศน์ออกเป็น 4 ลักษณะดังนี้

Novak, 1984; White and Gunstone, 1992; และสุนีย์ สอนตระกูล, 2535 ได้เสนอวิธีการสร้างผังมโนทัศน์ที่มีลักษณะคล้ายกัน มีขั้นตอนดังนี้

1. คัดเลือกบทเรียนจากหนังสือที่มีเนื้อหาไม่ยากจนเกินไป ควรเป็นเนื้อหาสั้น ๆ และประกอบด้วยมโนทัศน์ที่ไม่มากเกินไป เช่น เรื่องสิ่งมีชีวิต มโนทัศน์ที่ได้ อาจเป็น สัตว์ พืช หรือหญ้า
 2. วิเคราะห์มโนทัศน์ที่มีความสำคัญ ด้วยการเขียนมโนทัศน์ลงบนกระดาษ หรือการ์ด ขนาด 3x5 นิ้ว
 3. จัดเรียงลำดับและแยกแยะมโนทัศน์ โดยดูว่ามโนทัศน์ที่กว้างมโนทัศน์ใดเป็นมโนทัศน์รองมโนทัศน์ใดเป็นมโนทัศน์ที่เฉพาะเจาะจง
 4. เรียงลำดับมโนทัศน์ ให้มโนทัศน์ที่กว้างอยู่บนสุดและลดหลั่นลงมาด้วยมโนทัศน์รองจนถึงมโนทัศน์ที่เฉพาะเจาะจง
 5. ลากเส้นเชื่อมโยงระหว่างมโนทัศน์ต่าง ๆ จากนั้นหาคำหรือข้อความมาเชื่อมระหว่างมโนทัศน์เพื่อให้มโนทัศน์เหล่านั้นสัมพันธ์กัน
 6. ตรวจสอบหรือปรับปรุงผังมโนทัศน์ที่สร้างขึ้นให้ถูกต้องตามเนื้อหา
- Trochim (1997) ได้อธิบายขั้นตอนการสร้างผังมโนทัศน์เพื่อการวางแผนและการประเมินผลดังนี้

1. การเตรียมการ (Preparation) ประกอบด้วย
 - 1.1 การคัดเลือกผู้ให้ความร่วมมือ ในการพัฒนาประเด็นที่สนใจ
 - 1.2 การใช้วิธีระดมสมอง (Brain Storming) และมุ่งการจัดลำดับความสำคัญ

(Rating)

2. การผลิตเนื้อหาสาระที่สำคัญ (Generation of Statements)
การใช้ความร่วมมือของกลุ่มตัวอย่าง ช่วยกันออกความคิดเห็นต่าง ๆ โดยวิธีการระดม

สมอง

3. การจัดโครงสร้างเนื้อหาสาระที่ได้จากการระดมสมอง (Structuring of Statement)
 - 3.1 การจัดประเภทเนื้อหาสาระ (Sorting)
 - 3.2 การจัดอันดับเนื้อหาสาระ (Rating)
4. การเลือกตัวแทนของเนื้อหาสาระ (Representation of Maps)
5. การแปลความหมายของแผนภูมิ (Interpretation of Maps)
 - 5.1 การจัดรายการของเนื้อหาสาระ (The Statement List)
 - 5.2 การจัดรายการของเนื้อหาสาระให้เป็นกลุ่ม (The Cluster List)
 - 5.3 การตั้งชื่อกลุ่มต่าง ๆ ที่จัดขึ้น (Naming the Clusters)
 - 5.4 การเขียนแผนผังจากกลุ่มต่าง ๆ (The Cluster Map) ด้วยการใช้การจัดลำดับ

ความสำคัญ

6. การใช้ประโยชน์ของผัง

6.1 สำหรับการวางแผน สามารถใช้ได้อย่างหลากหลาย ได้แก่ แผนปฏิบัติการ (Action Plans) ใช้สำหรับวางแผนโครงสร้างของกลุ่ม (Planning Group Structure) ใช้ในการประเมินความต้องการจำเป็น (Need Assessment) และใช้ในการพัฒนาโปรแกรมต่าง ๆ

6.2 สำหรับการประเมินผล สามารถใช้อย่างหลากหลายเช่นเดียวกันโดยใช้ในการพัฒนาโปรแกรมต่าง ๆ (Program Development) ใช้ในการวัด (Measurement) การคัดเลือกกลุ่มตัวอย่าง (Sampling) และการประเมินผลลัพธ์ (Outcome Assessment)

Grayson (2000) ได้อธิบายขั้นตอนการสร้างผังมโนทัศน์ไว้ดังนี้

1. เขียนมโนทัศน์หรือคำหลักเกี่ยวกับเรื่องที่น่าสนใจลงไป
2. ระบุมโนทัศน์หลัก มโนทัศน์รอง และมโนทัศน์เฉพาะ
3. เริ่มเขียนผังมโนทัศน์โดย
 - 3.1 มโนทัศน์ต้องอยู่ในรูปกรอบวงกลมหรือวงรี
 - 3.2 วางมโนทัศน์หลักไว้ข้างบนสุด

3.3 วางมโนทัศน์รองไว้ถัดลงมา

3.4 วางมโนทัศน์ที่เฉพาะไว้ล่างสุด

4. ชีตเส้นแสดงความสัมพันธ์ระหว่างมโนทัศน์เหล่านั้น

5. ตั้งชื่อเส้นตรงเหล่านั้นด้วยคำที่มีความหมายแสดงถึงความสัมพันธ์ของมโนทัศน์

6. แก้ไขปรับปรุงผังมโนทัศน์

Safdar, Hussain Shah and Rifat (2012 p. 6) ได้อธิบายขั้นตอนการสร้าง

ผังมโนทัศน์ไว้ดังนี้

1. วิเคราะห์เนื้อหา ระบุข้อกำหนดที่สำคัญ และแนวคิดที่สำคัญที่ต้องการรวบรวมบน
ผังมโนทัศน์

2. จัดแนวความคิดที่เป็นตัวแทนของข้อมูลที่ดีที่สุด

3. ใช้วงกลม วงรี หรือสี่เหลี่ยมผืนผ้าวางล้อมรอบข้อกำหนดที่สำคัญ และแนวคิดที่สำคัญ
แนวคิดเดียว แนวคิดจะถูกเชื่อมต่อกับเส้นตรงที่มีลูกศร แต่ละเส้นเชื่อมโยงสองแนวคิดเท่านั้น
จำนวนการเชื่อมโยงไม่มีข้อจำกัด

4. แต่ละแนวคิดจะสัมพันธ์กับแนวคิดอื่น ๆ ภายใต้วงรูปเดียวกัน ซึ่งความสัมพันธ์จะใช้
คำหรือวลีเป็นชื่อในแต่ละป้าย ตามเส้นที่เชื่อมระหว่างคำสองคำ

จากวิธีการสร้างผังมโนทัศน์ ผู้วิจัยสรุปขั้นตอนการสร้างผังมโนทัศน์ เป็น 6 ขั้นตอน ดังนี้

1. คัดเลือกเนื้อหาที่จะนำมาสร้างเป็นผังมโนทัศน์

2. วิเคราะห์มโนทัศน์ที่มีความสำคัญ

3. จัดเรียงลำดับและแยกแยะมโนทัศน์

4. เรียงลำดับมโนทัศน์

5. ลากเส้นเชื่อมโยงระหว่างมโนทัศน์ต่าง ๆ จากนั้นหาคำหรือข้อความมาเชื่อมระหว่าง
มโนทัศน์เพื่อไหมมโนทัศน์เหล่านั้นสัมพันธ์กัน

6. ตรวจสอบหรือปรับปรุงผังมโนทัศน์ที่สร้างขึ้นให้ถูกต้องตามเนื้อหา

จากขั้นตอนดังกล่าวผู้วิจัย ได้นำไปกำหนดเป็นแนวทางการสร้างผังมโนทัศน์ในการพัฒนา
เกณฑ์การให้คะแนนผังมโนทัศน์สำหรับประเมินผลการเรียนรู้ ต่อไป

ประโยชน์ของผังมโนทัศน์

การนำผังมโนทัศน์ไปใช้ในการศึกษาอย่างกว้างขวาง นักการศึกษาหลายท่านได้ กล่าวถึงการ
นำผังมโนทัศน์ไปใช้ประโยชน์ ดังนี้

Baroody and Bartels (2001) ได้กล่าวว่าผังมโนทัศน์เป็นเครื่องมือประเมินกระบวนการ
คิดประเภทหนึ่ง และได้กล่าวถึงมีประโยชน์ของการใช้ผังมโนทัศน์ในการประเมินไว้ดังนี้

1. ผังมโนทัศน์สามารถใช้ในการวินิจฉัยและเป็นข้อมูลย้อนกลับในการจัดการเรียนการสอนได้เช่นเดียวกับการตัดเกรด นอกจากนี้การสร้างผังมโนทัศน์ของนักเรียนสามารถทำให้ครูได้ทราบถึงระดับความเข้าใจในความคิดรวบยอด และสารสนเทศที่ได้สามารถไปใช้ในการวางแผนหรือปรับปรุงการเรียนการสอนได้

2. ใช้เป็นเครื่องมือในการจัดการเรียนการสอน การสังเกตพฤติกรรมการสร้างผังมโนทัศน์ทำให้ทราบจุดบกพร่องของนักเรียนและเป็นเครื่องมือที่ทำให้ทราบการเรียนรู้ของผู้เรียนได้ชัดเจนยิ่งขึ้น

3. ผังมโนทัศน์สามารถใช้ประเมินการทำงานกลุ่มและพัฒนาการรายบุคคลได้

4. ผังมโนทัศน์สามารถใช้แบบทดสอบทางการเขียนและเป็นวิธีการประเมินทางเลือกใหม่สำหรับครู กิจกรรมการทดสอบโดยใช้ผังมโนทัศน์โดยการฟังการอภิปรายเกี่ยวกับการสร้างผังทำให้ครูผู้สอนเข้าใจความคิดของนักเรียนได้ชัดเจนขึ้น

Lanzing (2002) ได้กล่าวถึงประโยชน์ของการใช้ผังมโนทัศน์ ดังนี้

1. การสร้างแนวคิด เช่น การระดมสมอง
2. การออกแบบสิ่งที่มีโครงสร้างสลับซับซ้อน เช่น สื่อขนาดใหญ่ (Hypermedia) ตำรา (Long Text) เว็บไซต์ขนาดใหญ่ (Large Web Sites)

3. การสื่อสารในแนวคิดที่ซับซ้อน

4. การช่วยให้เกิดการเรียนรู้ด้วยการผสมองค์ความรู้ที่มีอยู่และองค์ความรู้ใหม่เข้าด้วยกัน โดยแสดงให้เห็นชัดเจน

5. การช่วยประเมินความเข้าใจที่ถูกต้องหรือวินิจฉัยความเข้าใจที่คลาดเคลื่อน

จากการศึกษาสามารถสรุปประโยชน์ของผังมโนทัศน์ไปใช้ในการศึกษาได้ดังนี้

1. ใช้เป็นเครื่องมือในการเตรียมการสอนของครู ครูสามารถจัดกิจกรรมการเรียนการสอนโดยใช้ผังมโนทัศน์เพื่อสำรวจความรู้ของนักเรียนแล้วนำไปวางแผนการสอนให้เหมาะสมกับนักเรียนและใช้ในการจัดลำดับเนื้อหาสาระที่จะสอน

2. ใช้เป็นเครื่องมือในการประเมินผลการเรียนรู้และประเมินความคิดรวบยอดของนักเรียน โดยการให้นักเรียนสรุปสิ่งที่เรียนเป็นผังมโนทัศน์หรือตอบข้อสอบโดยใช้ผังมโนทัศน์เพื่อแสดงความเข้าใจในการเขียนตอบและเพื่อวินิจฉัยความเข้าใจที่คลาดเคลื่อนของนักเรียน

3. ใช้เป็นเครื่องมือในการเรียนรู้สำหรับนักเรียน เพื่อให้เกิดการเรียนรู้อย่างมีความหมาย โดยการใช้กรอบมโนทัศน์เพื่อสร้างความหมายจากสิ่งที่เรียน จะทำให้นักเรียนจดจ่อไปได้บ้างและมีความคงทนเพราะจะทำให้เห็นความสัมพันธ์ของสิ่งที่เรียนไปทั้งหมด

การประเมินผลการเรียนรู้โดยใช้ผังโนทัศน์ เป็นกระบวนการที่มีความสำคัญและท้าทาย ในที่นี้ผู้วิจัยขอนำเสนอแนวคิดเกี่ยวกับเกณฑ์การให้คะแนนพอสังเขป เพื่อนำไปใช้เป็นข้อมูลสำหรับ พัฒนาเกณฑ์การให้คะแนนผังโนทัศน์สำหรับประเมินผลการเรียนรู้ ต่อไป

ตอนที่ 3 แนวคิดเกี่ยวกับเกณฑ์การให้คะแนน

ความหมายของเกณฑ์การให้คะแนน

มีนักการศึกษาหลายท่าน ได้ให้ความหมายของเกณฑ์การให้คะแนนไว้ดังนี้

สมศักดิ์ ภูวิภาตววรรณ (2544) กล่าวว่า เกณฑ์การให้คะแนน คือ “เครื่องมือในการให้คะแนน (Scoring Tool) ที่มีการระบุเกณฑ์ (Criteria) ประเมินชิ้นงานและคุณภาพ (Quality) ของชิ้นงานในแต่ละเกณฑ์อาจจะแบ่งเป็นยอดเยี่ยมจนถึงไม่ดี”

สำนักวิชาการและมาตรฐานการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (2553) กล่าวว่า เกณฑ์การประเมิน (Rubric) คือแนวการให้คะแนนเพื่อประเมินผลงานหรือ การปฏิบัติงานของผู้เรียนหรือกล่าวอีกอย่างหนึ่งว่า Rubric เป็นเครื่องมือให้คะแนนชนิดหนึ่งใช้ในการประเมินการปฏิบัติงานหรือผลงานของผู้เรียน

Lund (2000) กล่าวว่า เกณฑ์การให้คะแนน เป็นกฎกำหนดสำหรับความคาดหวังเกี่ยวกับ ความสามารถของผู้เรียน หรืออีกนัยหนึ่งคือ ข้อเสนอแนะ ที่สร้างขึ้นสำหรับการประเมินความสามารถ หรือเป็นเกณฑ์ที่ได้กำหนดไว้สำหรับความสามารถ

Schiemer (2000) กล่าวว่า เกณฑ์การให้คะแนน เป็นแนวทางในการให้คะแนน สำหรับความสามารถของผู้เรียน การพัฒนาเกณฑ์การประเมินที่ชัดเจนและแยกแยะเกณฑ์การ ความสามารถที่สำคัญ ๆ สามารถที่จะประเมินได้อย่างเชื่อถือได้จากครูหรือบุคคลอื่น ๆ

Mertler (2001) กล่าวว่า เกณฑ์การให้คะแนน คือแนวทางการให้คะแนนซึ่งได้กำหนด เกณฑ์ในการแสดงออก เพื่อประเมินผลงานและการปฏิบัติของนักเรียน

จากความหมายของเกณฑ์การให้คะแนนสรุปได้ว่า เกณฑ์การให้คะแนนเป็นแนวทางการ ให้คะแนนความสามารถของผู้เรียนอย่างชัดเจน โดยระบุระดับคะแนนอย่างละเอียด เพื่อให้ผู้เรียน ทราบว่าตนเองอยู่ในระดับใด หรือมีคุณภาพเป็นอย่างไร

การกำหนดเกณฑ์การให้คะแนน

การกำหนดเกณฑ์การให้คะแนนที่ดี ต้องเหมาะสมกับงาน และมีเหตุมีผล บอกระดับ ก่อนหลังของคะแนน นักเรียนสามารถประเมินด้วยตนเองได้ จากเกณฑ์ที่ครูสร้างขึ้น

Hopkind and Antes (1990), Stenmark (1991), Mertler (2001) ได้แบ่งกฎเกณฑ์การ ให้คะแนนไว้สอดคล้องกัน คือแบ่งออกเป็น 2 แบบ ได้แก่ เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ

(Analytic Scoring Rubrics) และเกณฑ์การให้คะแนนแบบภาพรวม (Holistic Scoring Rubrics) มีรายละเอียดดังนี้

1. เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (Analytic Scoring Rubrics)

เกณฑ์การให้คะแนนแบบนี้เป็นการวิเคราะห์รายละเอียดออกเป็นองค์ประกอบ เช่น การสื่อสาร ความถูกต้องทางคณิตศาสตร์หรือความสมบูรณ์ และลักษณะอื่น ๆ ที่คิดว่าจะมีความสำคัญ จากนั้นกำหนดคะแนนสำหรับกำหนดคะแนนสำหรับแต่ละองค์ประกอบย่อยในแง่ของคุณภาพของคำตอบ โดยหลังจากที่ออกข้อสอบแล้ว ครูจะดำเนินการเขียนคำตอบต้นแบบ (Model Response) ขึ้นที เพื่อให้มีโครงร่างขององค์ประกอบหลักซึ่งจะระบุไว้ในกระบวนการให้คะแนน ในการรวมคะแนนทั้งหมดควรจะลดบทบาทของข้อมูลที่ไม่เกี่ยวข้องในคำตอบที่จะมีต่อผลรวมของคะแนน ส่วนใหญ่แล้วพบว่าวิธีการให้คะแนนแบบนี้ใช้เวลานานและสร้างความเหนื่อยล้าถ้าหากข้อสอบมีจำนวนมาก

Nitko (1996) ได้กล่าวถึงข้อดีของการให้คะแนนแบบแยกองค์ประกอบว่า ทำให้เกิดการพิจารณาขั้นตอนของกระบวนการแก้ปัญหาที่ไม่ใช้พิจารณาเพียงคำตอบ เป็นวิธีการกำหนดคุณค่าของงานด้วยตัวเลขที่ชัดเจน ทำให้ได้รับข้อมูลที่มีรายละเอียดซึ่งแสดงถึงจุดเด่นและจุดด้อยของผู้เรียนในแต่ละองค์ประกอบ ช่วยให้ครูเน้นหรือปรับปรุงจุดที่นักเรียนต้องปรับปรุงแก้ไขได้ตรงกับความต้องการของนักเรียนมากขึ้น นำไปใช้ในการวางแผนและปรับปรุงการเรียนการสอน รวมถึงการให้ข้อมูลย้อนกลับแก่ผู้เรียนเพื่อพัฒนาหรือปรับปรุงตนเองได้อย่างชัดเจน ตัวอย่างเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ แสดงดังตารางที่ 2-1

ตารางที่ 2-1 ตัวอย่างเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ

องค์ประกอบย่อย	เกณฑ์การให้คะแนน
<ul style="list-style-type: none"> ● ความเข้าใจในปัญหา 	1 ไม่เข้าใจปัญหา 2 เข้าใจปัญหาผิดบางส่วนหรือตีความผิด 3 เข้าใจปัญหาอย่างถูกต้อง ชัดเจน
<ul style="list-style-type: none"> ● การวางแผนแก้ปัญหา 	1 ไม่มีการวางแผนหรือวางแผนไม่ถูกต้อง 2 วางแผนถูกต้องบางส่วน 3 วางแผนแล้วนำไปสู่การแก้ปัญหาได้อย่างถูกต้อง
<ul style="list-style-type: none"> ● การได้มาซึ่งคำตอบ 	1 ไม่ตอบหรือได้คำตอบที่มาจากการวางแผนที่ผิด 2 เขียนคำตอบผิด หรือ คำนวณผิดพลาด 3 คำตอบถูกต้อง

ที่มา: Stenmark (1991 citing Charles et al., 1987).

2. เกณฑ์การให้คะแนนแบบภาพรวม (Holistic Scoring Rubrics)

เกณฑ์การให้คะแนนแบบนี้จะเป็นการให้คะแนนที่เน้นภาพรวมหรือองค์รวมของงานทั้งหมดโดยจะไม่กำหนดหรือระบุคะแนนแยกแยะไปถึงองค์ประกอบหรือประเด็นต่าง ๆ ที่จะต้องมีไว้ล่วงหน้าเป็นการเฉพาะเหมือนกับเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ แต่จะเป็นการเขียนบรรยายคุณภาพของงานในแต่ละระดับคะแนน วิธีการนี้จะตรวจคำตอบ โดยอาศัยความประทับใจโดยภาพรวม วิธีการนี้ต้องอาศัยการอ่านข้อสอบในครั้งแรกอย่างรวดเร็วเพื่อสร้างความประทับใจอย่างกว้าง ๆ ต่อคำตอบแต่ละคำตอบ ทดลองแบ่งข้อสอบออกเป็นกลุ่ม ๆ ตามคุณภาพของคำตอบที่แตกต่างกันไป จำนวนกลุ่มนั้นไม่แน่นอน ขึ้นอยู่กับความต้องการของผู้ให้คะแนนและระดับแยกความแตกต่างที่ต้องการ โดยทั่วไปแล้วผู้ให้คะแนนจะอ่านจนกระทั่งมีคำตอบหนึ่งที่ได้รับการประเมินว่าดีกว่าอย่างแน่นอน และใช้คำตอบนี้เป็นคะแนนอ้างอิงในการให้คะแนนคำตอบอื่น ๆ การอ่านอย่างรอบคอบขึ้นในรอบที่สองสามารถทำให้เกิดการเปลี่ยนแปลงจากกลุ่มที่กำหนดไว้ได้ กล่าวคือ การจัดกลุ่มคำตอบของนักเรียนอาจจะได้รับการกลั่นกรองอีกครั้งถ้าผู้ตรวจรู้สึกว่าการกำหนดประเภทในขั้นแรกผิดพลาด ในขั้นนี้ข้อสอบที่แยกไว้ในแต่ละกลุ่มควรจะอยู่ในระดับเดียวกัน ถ้าข้อสอบฉบับใดมีความโดดเด่นกว่าฉบับอื่น ๆ ในกลุ่มก็ควรแยกไปอยู่ในกลุ่มที่ดีกว่า หรือถ้าแยกว่าก็แยกไปกว่าแยกว่า การแยกกลุ่มก่อนที่จะให้คะแนนนี้จะลดความเป็นไปได้ที่ว่าข้อสอบที่อยู่ในระดับเดียวกันได้คะแนนขั้นสุดท้ายแตกต่างกัน หลังจากที่มีผู้อ่านพอใจกับการแยกกลุ่มอย่างเหมาะสมแล้ว ก็จะมีการให้คะแนนขั้น

สุดท้ายหรือการประเมินอื่น ๆ การตรวจแบบทดสอบแบบภาพรวมนี้จะมีประสิทธิภาพสำหรับข้อสอบที่มีจำนวนมาก

ปรีชา เนาว์เย็นผล (2544) และ Nitko (1996) ได้กล่าวถึงข้อดีของการให้คะแนนแบบภาพรวมไว้ว่า เป็นการเปิดโอกาสให้มีการพิจารณาการตอบของนักเรียนได้อย่างรวดเร็ว ใช้เวลาน้อยในการตรวจและง่ายในการนำไปใช้ เน้นการพิจารณากระบวนการที่ใช้ไม่ใช่พิจารณาเฉพาะคำตอบเท่านั้น ใช้คะแนนค่าเดียวซึ่งครอบคลุมภาพรวมของคำตอบ อย่างไรก็ตามแม้วิธีการนี้จำทำให้คะแนนมีความเที่ยงสูง แต่ยากที่จะอธิบายเหตุผลในการให้คะแนนกับนักเรียนเมื่อคืนข้อสอบให้นักเรียนแล้ว คำตอบที่เป็นแบบอย่างควรจะนำไปอ่านให้นักเรียนในชั้นฟังเพื่อเป็นคะแนนอ้างอิงให้กับนักเรียนแต่ละคน ซึ่งตัวอย่างของเกณฑ์การให้คะแนนแบบภาพรวมในการแก้สมการทางคณิตศาสตร์ แสดงดังตารางที่ 2-2

ตารางที่ 2-2 ตัวอย่างเกณฑ์การให้คะแนนแบบภาพรวม

ระดับคะแนน	รายละเอียดของแต่ละระดับคะแนน
4	แสดงถึงความเข้าใจเกี่ยวกับความคิดและหลักการทางคณิตศาสตร์ที่สมบูรณ์ คำนวนถูกต้อง มีขั้นตอนในการแก้สมการที่ถูกต้อง สมบูรณ์
3	แสดงถึงความเข้าใจเกี่ยวกับความคิดและหลักการทางคณิตศาสตร์เกือบสมบูรณ์ มีข้อผิดพลาดในการคำนวณเล็กน้อย มีขั้นตอนในการแก้ปัญหาถูกต้อง
2	แสดงถึงความเข้าใจเกี่ยวกับความคิดและหลักการทางคณิตศาสตร์ได้บางส่วน การคำนวณส่วนใหญ่ผิดพลาด มีขั้นตอนในการแก้ปัญหาที่ผิดพลาด
1	แสดงถึงความเข้าใจเพียงเล็กน้อย มีความเข้าใจที่คลาดเคลื่อนมากในความคิดทางคณิตศาสตร์ ใช้กระบวนการคิดที่ผิด

ที่มา: The Online Teacher Resource (2009)

ลักษณะที่ดีของเกณฑ์การให้คะแนน

ลักษณะของเกณฑ์การให้คะแนนที่ดี เกณฑ์การให้คะแนนเป็นชุดคะแนนที่ใช้เป็นแนวทางสำหรับการประเมินผลงานของผู้เรียน ลักษณะที่ดีของเกณฑ์การให้คะแนน มีดังนี้ (Wiggins, 1998)

1. มีความเกี่ยวข้องกับจุดมุ่งหมาย หรือ เป้าหมายทั่วไป กล่าวคือเกี่ยวข้องกับงานที่ทำ

2. จำแนกการปฏิบัติได้อย่างเที่ยงตรง
3. ในแต่ละเกณฑ์การให้คะแนน จะไม่มีการรวมเกณฑ์การให้คะแนน
4. วิเคราะห์งานได้อย่างละเอียด
5. ภาษาที่ใช้อธิบายคุณลักษณะงาน จำแนกคุณภาพของงานได้ถูกต้อง
6. สามารถตัดสินงานได้ถูกต้อง
7. อธิบายอย่างชัดเจนในแต่ละระดับของคะแนน และมีความแม่นยำตรงในการให้คะแนนใน

ตัวของมันเอง

8. ตัดสินให้คะแนนจากผลงานที่ปฏิบัติ มากกว่ากระบวนการ รูปแบบเนื้อหา หรือความตั้งใจในการทำงาน

นอกจากนี้ Wiggins (1998) ได้นำ เสนอคุณลักษณะของเกณฑ์การให้คะแนนว่าต้องมีลักษณะ ดังนี้

1. คะแนนต้องมีลักษณะต่อเนื่อง กล่าวคือ ให้คะแนนเป็นจำนวนเต็ม เช่น ให้คะแนน เป็น 5 4 3 2 และ 1 คะแนนแต่ละคะแนน มีความห่างเท่ากัน
2. มีความสอดคล้องกัน คะแนนแต่ละระดับแสดงถึงความลดหลั่นของคุณภาพงาน
3. มีความเกี่ยวเนื่องกัน ในแต่ละระดับของการให้คะแนน
4. น้ำหนักการให้คะแนนในแต่ละระดับมีความเหมาะสม มีเหตุผล น้ำหนักของคะแนนในแต่ละระดับสามารถอ้างอิงไปยังระดับอื่น ๆ ได้
5. มีความเที่ยงตรง คะแนนในแต่ละระดับ แสดงถึงคุณภาพของการปฏิบัติ เป็นสิ่งสะท้อนถึงคุณภาพของงาน ไม่ได้เน้นถึงปริมาณ แต่เป็นเกณฑ์ตามสภาพจริง
6. มีความเชื่อถือได้ กล่าวคือ มีความคงเส้นคงวาในการให้คะแนน ถึงแม้ใครจะเป็นผู้ประเมิน และจะประเมินในช่วงเวลาใดก็ตาม
7. ช่วยในการให้เหตุผลประกอบการให้เกรดนักเรียนได้
8. ช่วยเพิ่มคุณภาพงานของนักเรียน

การเขียนเกณฑ์การให้คะแนน

การเขียนเกณฑ์การให้คะแนนจะเป็นแบบภาพรวม หรือแบบแยกองค์ประกอบ ขึ้นอยู่กับความถนัดของครู โดยให้ยึดตามแนวทางคุณลักษณะที่ดีของเกณฑ์การให้คะแนน ดังที่กล่าวมาแล้ว เทคนิคการเขียนเกณฑ์การให้คะแนน อาจจะเริ่มจากด้านที่ดีที่สุดกับด้านที่แย่ที่สุด ให้ตรงข้ามกันก่อน เพื่อง่ายต่อการเขียนในระดับอื่น ๆ หรืออาจจะเขียนด้านดีที่สุดก่อน (เขียนเหมือนกับแบบประเมินภาพรวม) แล้วระบุด้านลบ หรือด้านไม่ดี ในระดับคะแนนที่ต่ำลงมา หรืออาจจะเขียนด้านแย่ที่สุด

ก่อน เป็นด้านลบทั้งหมด ซึ่งเป็นคะแนนต่ำสุด แล้วในระดับคะแนนที่สูงขึ้นให้เพิ่มด้านบวกหรือด้านดีไปเรื่อย ๆ (Mc Millan, 2001)

การสร้างเกณฑ์การให้คะแนน

การสร้างเกณฑ์การให้คะแนนหรือแนวทางในการให้คะแนนนั้นถือว่าเป็นสิ่งสำคัญเป็นอย่างมาก เพราะจะทำให้การประเมินครอบคลุม พร้อมทั้งการให้คะแนนมีความยุติธรรมการสร้างเกณฑ์การให้คะแนนมีดังนี้ (McMillan, 2001)

1. ต้องแน่ใจว่าเกณฑ์การให้คะแนน ได้เน้นประเด็นที่สำคัญของงาน
2. มีความสอดคล้องระหว่างระดับคะแนนกับจุดมุ่งหมายของการประเมิน ถ้าจุดมุ่งหมายของการประเมินกว้างและต้องใช้การตัดสินทุก ๆ ส่วนของชิ้นงานควรจะใช้การประเมินแบบภาพรวม แต่ถ้าการประเมินต้องการสะท้อนกลับให้เห็นความแตกต่างของประเด็นต่าง ๆ ของงานควรใช้การประเมินแบบแยกองค์ประกอบ
3. ข้อความที่ใช้อธิบายในแต่ละระดับคะแนน ต้องเป็นข้อความที่สามารถประเมินหรือสังเกตได้
4. ควรให้นักเรียน ผู้ปกครอง และผู้เชี่ยวชาญ ได้ร่วมกันสร้างเกณฑ์ การเปิดโอกาสให้นักเรียนได้ร่วมสร้างเกณฑ์ การปฏิบัติงานนั้น ๆ จะเป็นการกระตุ้นนักเรียนให้สนใจที่จะทำงานและจะทำ ให้นักเรียนนำไปเป็นแนวทางในการปฏิบัติงานนั้น ๆ
5. คุณลักษณะ หรือสิ่งที่จะวัดควรนิยามให้ชัดเจน
6. แสดงขั้นตอน หรือลำดับขั้นที่เหมาะสมของคะแนนในแต่ละระดับ เพื่อให้เกิดความคลาดเคลื่อนน้อยที่สุด เช่น หลีกเลี่ยงการให้คะแนนที่สูงเกินไป การให้คะแนนต่ำเกินไปการให้คะแนนส่วนใหญ่อยู่ตรงกลาง และการให้คะแนนที่เกิดจากความพึงพอใจเป็นการส่วนตัวของครูที่มีต่อนักเรียนคนนั้น ๆ
7. ระบบของการให้คะแนนต้องมีความเป็นไปได้ กล่าวคือ การให้คะแนนนิยมแบ่งระดับคะแนนเป็น 3-8 ระดับ ดังนั้น ในแต่ละระดับคะแนนต้องมีความชัดเจน และแยกจากกันได้จากที่กล่าวมาแล้ว สรุปได้ว่า การสร้างเกณฑ์การให้คะแนน ต้องคำนึงถึงงานที่ทำให้ต้องมีความสำคัญ มีความสอดคล้องระหว่างคะแนนกับจุดมุ่งหมายการประเมิน เกณฑ์ที่สร้างต้องเป็นรูปธรรมมีความชัดเจน เหมาะสมกับระดับชั้น และควรให้นักเรียนและผู้ปกครองมีส่วนร่วมในการสร้างเกณฑ์การประเมินด้วย

วัฒนาพร ระบุว่า (2541) กล่าวถึง กระบวนการสร้างเกณฑ์การให้คะแนนไว้ว่า เกณฑ์การให้คะแนน เป็นเครื่องมือที่ใช้กันมากในการประเมินตามสภาพจริง (Authentic Assessment) และการประเมินโดยอิงผลการทำงาน (Performance-Base Assessment) เกณฑ์การให้คะแนนที่ครู

สร้างขึ้นจะสามารถสะท้อน (ประเมิน) หลักสูตรและวิธีการสอนของครูด้วยกระบวนการในการวางรูปแบบของเกณฑ์การให้คะแนน ควรให้ผู้เรียนมีส่วนร่วมตามขั้นตอนต่อไปนี้

1. พิจารณาดูงานแบบต่าง ๆ โดยครูจัดให้ผู้เรียนได้ดูตัวอย่างของชิ้นงานที่ดีและไม่ค่อยดีแล้วให้บอกว่าชิ้นงานที่ดีมีลักษณะอย่างไร ที่ไม่ค่อยดีมีลักษณะอย่างไร
2. กำหนดแนวทางการพิจารณาชิ้นงาน โดยช่วยกันเขียนรายการดูว่างานที่ดีนั้นจะต้องดูอะไรบ้าง
3. กำหนดคำอธิบายคุณภาพระดับต่าง ๆ ในแต่ละแนวทางโดยอธิบายว่าคุณภาพระดับสูงสุดและต่ำสุดมีลักษณะอย่างไร แล้วเติมส่วนที่อยู่ตรงกลาง ๆ ลงไป
4. ฝึกหัดประเมินกับงานตัวอย่าง โดยให้ผู้เรียนลองใช้เกณฑ์การให้คะแนนกับงานที่ดีเป็นตัวอย่างในขั้นที่ 1
5. ฝึกใช้ในการประเมินตนเองและให้เพื่อนประเมิน โดยให้งานผู้เรียนทำในระหว่างที่ผู้เรียนกำลังทำงานให้หยุดเป็นระยะ ๆ เพื่อประเมินตนเอง และให้เพื่อนประเมิน
6. ทบทวนให้โอกาสผู้เรียนทบทวน ปรับปรุงงานของเขา โดยดูจากผลสะท้อนที่ได้จากขั้นตอนที่
7. ครูประเมิน ครูใช้เกณฑ์การให้คะแนนที่ผู้เรียนใช้นั้น เพื่อประเมินผลงานของผู้เรียนสรุปขั้นตอนในการสร้างเกณฑ์การให้คะแนน

ขั้นตอนในการสร้างเกณฑ์ในการประเมิน มี 12 ขั้นตอน ดังนี้ (Mc Millan, 2001)

1. กำหนดกิจกรรม/ งานที่จะประเมิน
2. กำหนดองค์ประกอบของเกณฑ์การประเมิน
3. นิยามปฏิบัติการให้สอดคล้องกับงานนั้น ๆ
4. กำหนดจำนวนระดับของเกณฑ์
5. พิจารณาเกณฑ์ผ่านและไม่ผ่านพร้อมคำอธิบาย
6. เขียนคำอธิบายในแต่ละระดับ
7. ตรวจสอบโดยผู้เชี่ยวชาญ
8. ปรับปรุงแก้ไขจากความคิดเห็นของผู้เชี่ยวชาญ
9. ทดลองใช้เกณฑ์ตรวจสอบผลงาน
10. ทาคุนภาพของเกณฑ์ เช่น Interrater Reliability เป็นต้น
11. ปรับปรุงเกณฑ์ที่ไม่ได้มาตรฐาน
12. จัดพิมพ์เกณฑ์การประเมินเป็นรูปเล่ม พร้อมคู่มือการใช้

การนำเกณฑ์การให้คะแนนไปใช้

Schiemer (2000) ได้เสนอวิธีการนำเกณฑ์การให้คะแนนไปใช้ดังนี้คือ แจกเอกสารเกี่ยวกับเกณฑ์การให้คะแนนให้ผู้เรียนก่อนที่จะทำการประเมิน เพื่อจะทำให้ผู้เรียนทราบดังนี้

1. รู้ว่าอะไรคือสิ่งที่คาดหวังเพื่อให้ถึงระดับความสามารถในแต่ละระดับ
2. เป็นความรับผิดชอบของบุคคลเพื่อให้ถึงระดับความสามารถที่ดีที่สุดที่บุคคลนั้นสามารถทำได้
3. ใช้เป็นเครื่องมือในการประเมินตนเอง หรือใช้เป็นข้อมูลย้อนกลับระหว่างการประเมินของครูหรือสำหรับการฝึกทักษะอย่างอิสระ

ประโยชน์ของเกณฑ์การให้คะแนน

รุ่งศิรินทร์ จันทรหอม (2545) ได้กล่าวถึงประโยชน์ไว้ดังนี้ คือ

1. ทำให้ครูผู้ปกครอง และผู้เรียนเข้าใจถึงความหมายของคะแนนที่สะท้อนถึงระดับความสามารถของผู้เรียน
2. ผู้เรียนรู้ตั้งแต่แรกถึงแนวทางและเกณฑ์ที่จะถูกประเมิน
3. ผู้เรียนสามารถพัฒนางานที่ได้รับมอบหมายโดยใช้ฐานของเกณฑ์และมาตรฐานที่ตั้งไว้
4. ผู้เรียนสามารถประเมินงานของตนเองโดยใช้มาตรฐานหรือเกณฑ์เป็นหลัก
5. สะท้อน (ประเมิน) หลักสูตรและวิธีการสอนของครู ซึ่งจะนำไปสู่การปรับปรุงการสอนให้เหมาะสมยิ่งขึ้นต่อไป

สมศักดิ์ ภูวิภาดาวรรณ (2544) กล่าวว่า

1. ช่วยให้การคาดหวังของครูที่มีต่อผลงานของนักเรียนบรรลุผลสำเร็จได้ โดยนักเรียนจะเกิดความเข้าใจ และสามารถใช้ Rubrics ต่อการประเมินและพัฒนาชิ้นงานของตน
2. ช่วยให้ผู้เกิดความรู้สึกกระตือรือร้นว่าต้องการให้นักเรียนเกิดการเรียนรู้หรือพัฒนาอะไรบ้าง
3. ช่วยให้นักเรียนระบุคุณลักษณะจากงานที่เป็นตัวอย่างได้ โดยใช้ Rubrics ตรวจสอบ
4. ช่วยให้นักเรียนสามารถควบคุมตนเองในการปฏิบัติงานเพื่อไปสู่ความสำเร็จได้
5. เป็นเครื่องมือในการเชื่อมโยงความสัมพันธ์ระหว่างกิจกรรมการปฏิบัติงานต่าง ๆ ของนักเรียนได้เป็นอย่างดี
6. ช่วยให้ผู้ที่เกี่ยวข้องเช่น ผู้ปกครอง ผู้สนับสนุน ผู้นิเทศน์ ได้เกิดความเข้าใจเกณฑ์ในการตัดสินผลงานของนักเรียนที่ครูใช้
7. ช่วยในการให้เหตุผลประกอบการให้เกรดนักเรียนได้
8. ช่วยเพิ่มคุณภาพงานของนักเรียน

Schiemer (2000) ได้กล่าวไว้ว่า การกำหนดเกณฑ์การให้คะแนนมีประโยชน์ ดังนี้

1. ทำให้ช่วยประหยัดเวลาในการประเมินผลการให้คะแนน
2. เป็นเครื่องมืออำนวยความสะดวกในการสื่อสารกับเด็กนักเรียน พ่อ แม่ ผู้ปกครองและครูอื่น ๆ
3. เป็นแนวทางให้พ่อ แม่ ผู้ปกครองและเพื่อนร่วมงานคนอื่น ๆ ที่ต้องการเพิ่มเติมทักษะในการสอนของเขา ความต้องการที่จะตรวจตราดูแลความก้าวหน้าของนักเรียนหลังจากที่พวกเขาได้ปฏิบัติเพิ่มเติม

การพิจารณาว่าเครื่องมือที่นำมาใช้ในการประเมินผลการเรียนรู้มีคุณภาพหรือไม่ โดยทั่วไปจะต้องพิจารณาในเรื่องความเที่ยง (Reliability) และความตรง (Validity) สมคักดี ภูวิภาดาวรรณ (2544); McClure, Sonak and Suen (1999); Ruiz-Primo & Shavelson (1996) กล่าวว่า การวิเคราะห์ความเที่ยงแบบดั้งเดิม ย่อมไม่เหมาะสมกับการหาคุณภาพของผังมโนทัศน์ แต่ควรให้ความสำคัญกับผู้ประเมินมากกว่า เพราะทฤษฎีความเที่ยงแบบดั้งเดิม มีข้อตกลงว่า ข้อสอบต้องวัดสิ่งเดียวกัน มีลักษณะคู่ขนานกันและดำเนินการสอบครั้งเดียว แต่การประเมินโดยใช้ผังมโนทัศน์เป็นการประเมินที่มีลักษณะแตกต่างกัน ทฤษฎีการประมาณค่าความเที่ยงที่เรียกว่า ทฤษฎีการสรุปอ้างอิง (Generalizability Theory) จึงเป็นทฤษฎีที่มีความเหมาะสมกับการวิเคราะห์ค่าความเที่ยงของการประเมินโดยใช้ผังมโนทัศน์ ผู้วิจัยขอแนะนำแนวคิดเกี่ยวกับทฤษฎีการสรุปอ้างอิงพอสังเขป ดังนี้

ตอนที่ 4 ทฤษฎีการสรุปอ้างอิง

แนวคิดพื้นฐานของ G -Theory

ทฤษฎีการสรุปอ้างอิงเริ่มต้นแนวคิดโดย Cronbach (1972) ได้เสนอทฤษฎีสำหรับวิเคราะห์ความน่าเชื่อถือของผลการวัดพฤติกรรม (The Dependability of Behavioral Measurements) สำหรับสถานการณ์ของการวัดผลลักษณะต่าง ๆ ซึ่งต่อมาเป็นที่รู้จักกันในชื่อของทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด ทฤษฎีนี้ได้ขยายแนวความคิดของความเที่ยงตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) ให้สามารถสรุปผลความเที่ยงในสถานการณ์หรือเงื่อนไขการทดสอบลักษณะต่าง ๆ ได้

ทฤษฎีการทดสอบแบบดั้งเดิมมิได้ให้ความสนใจต่อสถานการณ์หรือเงื่อนไขของการวัดและปฏิสัมพันธ์ระหว่าง ผู้สอบ กับเงื่อนไขของการวัดที่สามารถส่งผลต่อความน่าเชื่อถือของผลการวัด แต่ G-Theory ได้นำแนวคิดที่แยกส่วนความคลาดเคลื่อน (Error) จากหลายแหล่ง (Multiple Error Sources) ประกอบด้วยแหล่งความคลาดเคลื่อนอย่างเป็นระบบ (Systematic Source) และ

ความคลาดเคลื่อนสุ่ม (Random Source) มิใช่เป็นความคลาดเคลื่อนรวมเพียงแหล่งเดียวเหมือนอย่างทฤษฎีการทดสอบแบบดั้งเดิม

ก่อนที่จะนำทฤษฎีการสรุปอ้างอิงไปใช้ในการหาค่าประกอบความแปรปรวนค่าสัมประสิทธิ์การสรุปอ้างอิง มีความจำเป็นจะต้องมีความรู้ และความเข้าใจเกี่ยวกับคำศัพท์ของทฤษฎีการสรุปอ้างอิง เนื่องจากมีคำศัพท์เฉพาะที่เกี่ยวข้องหลายคำที่ต้องมีความเข้าใจที่ถูกต้อง

คำศัพท์ที่เกี่ยวข้องกับทฤษฎีการสรุปอ้างอิง

การที่จะนำทฤษฎีการสรุปอ้างอิงไปใช้ในการหาค่าประกอบความแปรปรวนหรือประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงนั้น จะต้องมีความรู้ ความเข้าใจในคำศัพท์เฉพาะของทฤษฎีการสรุปอ้างอิง ดังนี้ (Brennan, 1983; Shavelson and Webb, 1991; ศิริชัย กาญจนวาสี, 2550)

1. ประชากร (Population) หมายถึง สิ่งที่ต้องการวัด (Object of Measurement) หรือ สิ่งที่มีวัดทั้งหมดในสถานการณ์ของการทดสอบทั่วไป สิ่งที่มีวัดมักได้แก่ บุคคลหรือผู้สอบ

2. องค์ประกอบ (Facet) หมายถึง ชุดของเงื่อนไขของการวัดที่มีลักษณะคล้ายกัน (Similar Conditions of Measurement) หรือ กลุ่มเงื่อนไขของการวัด ซึ่งเป็นองค์ประกอบที่คาดว่า มีผลต่อการวัดความคลาดเคลื่อน เช่น องค์ประกอบของข้อสอบ (item facet) องค์ประกอบของจำนวนครั้งของการสอบ (Occasion Facet) องค์ประกอบของผู้ตรวจข้อสอบ (Rater Facet) เป็นต้น โดยที่องค์ประกอบที่ต้องการศึกษาอาจเป็นองค์ประกอบสุ่มหรือองค์ประกอบแบบเจาะจงก็ได้

3. เงื่อนไขของการวัด (Condition of Measurement) หมายถึง ระดับขององค์ประกอบที่ทำให้ได้ค่าสังเกตแต่ละค่าขึ้นมาในการวัดครั้งหนึ่ง ๆ เช่น องค์ประกอบของจำนวนผู้ตรวจอาจกำหนดจำนวนระดับเป็น 1, 2 และ 3 คน เป็นต้น

4. เอกภพ (Universe) หมายถึง กลุ่มเงื่อนไขของการวัดทั้งหมดของแต่ละองค์ประกอบ เช่น จำนวนข้อสอบทุกข้อ ซึ่งมีความคล้ายกับประชากร (Population) แต่แตกต่างกันคือ เอกภพใช้กับองค์ประกอบที่เกี่ยวข้องกับการวัด ส่วนประชากรใช้กับองค์ประกอบที่ถูกวัด แยกออกมาเป็น

4.1 เอกภพของการสังเกตที่ยอมรับได้ (Universe of Admissible Observation) หมายถึง กลุ่มเงื่อนไขของการวัดที่เป็นไปได้ ซึ่งสามารถวัดหรือสังเกตได้ในแต่ละองค์ประกอบ เช่น ประกอบด้วยองค์ประกอบข้อสอบ และองค์ประกอบของผู้ตรวจ

4.2 เอกภพของการสรุปอ้างอิง (Universe of Generalization) หมายถึง เงื่อนไขของการวัดทั้งหมดขององค์ประกอบที่ต้องการสรุปอ้างอิงผลการวัดจากกลุ่มเงื่อนไขของการวัดเหล่านี้ไปยังกลุ่มเงื่อนไขของการวัดทั้งหมดขององค์ประกอบนั้น ๆ ดังนั้นเงื่อนไขขององค์ประกอบในเอกภพการสรุปอ้างอิงจึงเป็นสับเซตในเอกภพของการสังเกตที่ยอมรับได้ เอกภพของการสรุปอ้างอิงนี้จะต้องระบุประเด็นต่าง ๆ ให้ชัดเจน ซึ่งจะนำไปใช้ในการออกแบบการวัดของการศึกษาเพื่อตัดสินใจ

5. คะแนนเอกภาพ (Universe Score) หมายถึง คะแนนเฉลี่ยของเงื่อนไขการวัดทั้งหมด แต่ละองค์ประกอบ ซึ่งคล้ายกับคะแนนจริง (Ideal Score) ของสิ่งที่วัดในทฤษฎีการวัดแบบดั้งเดิม

6. การศึกษาเพื่อการสรุปอ้างอิง (Generalizability Study: G-Study) และการศึกษาเพื่อการตัดสินใจ (Decision Study: D-Study) ทั้งนี้ทฤษฎีการสรุปอ้างอิงมีขั้นตอนการศึกษาที่สำคัญ 2 ขั้นตอน คือ G-Study และ D-Study

6.1 G-Study เป็นการสรุปอ้างอิงผลจากการศึกษาตัวอย่างการวัดตามเงื่อนไขที่สนใจ ประเมินค่าความแปรปรวนของคะแนนจริงและความแปรปรวนของคะแนนความคลาดเคลื่อนจากแหล่งความคลาดเคลื่อนต่าง ๆ ที่สนใจ เพื่อสรุปอ้างอิงไปยังเอกภาพของการวัด และใช้เป็นข้อมูลสำหรับวางแผนเพื่อตัดสินใจ ในการศึกษา D

6.2 D-Study เป็นการใช้ข้อมูลจาก G-Study ที่สอดคล้องกับจุดประสงค์เฉพาะของการตัดสินใจ การเลือกใช้แบบทดสอบในสถานการณ์ต่าง ๆ ของการวัด ซึ่งเกี่ยวกับค่าความเชื่อมั่นของแบบทดสอบในสถานการณ์ของการวัดต่าง ๆ

7. รูปแบบความสัมพันธ์ระหว่างองค์ประกอบ หมายถึง องค์ประกอบที่เป็นสิ่งที่ชี้ถึงแหล่งความแปรปรวนของรูปแบบการวัด ในแบบจำลองการวิเคราะห์ความแปรปรวน ซึ่งรูปแบบความสัมพันธ์นี้แบ่งออกเป็น 3 ลักษณะ คือ

7.1 ความสัมพันธ์แบบไขว้ (Crossed) หมายถึง ความสัมพันธ์แต่ละลักษณะที่มีแต่ละระดับของสิ่งที่ถูกวัด ถูกวัดภายใต้เงื่อนไขเดียวกันหมด สัญลักษณ์คือ “x” อ่านว่า “crosses with” เช่น กำหนดรูปแบบความสัมพันธ์ระหว่างองค์ประกอบเป็น $p \times i \times r$ หมายถึง นักเรียน (p) ทำข้อสอบ (i) ทุกข้อ และผู้ตรวจ (r) ตรวจข้อสอบของนักเรียนทุกคน ทุก ๆ ข้อ

7.2 ความสัมพันธ์แบบแฝง (Nested) หมายถึง ความสัมพันธ์ในลักษณะที่แต่ละระดับของสิ่งที่ถูกวัด ถูกวัดภายใต้เงื่อนไขเดียวกันทั้งหมด สัญลักษณ์คือ “:” อ่านว่า “nested within” เช่น รูปแบบความสัมพันธ์องค์ประกอบเป็น $i : r$ หมายถึง ผู้ตรวจข้อสอบ (r) ตรวจข้อสอบ (i) ต่างข้อกัน

7.3 ความสัมพันธ์แบบผสม (Confounded) หมายถึง ความสัมพันธ์ที่มีทั้งความสัมพันธ์แบบไขว้และความสัมพันธ์แบบแฝงปนกันอยู่ เช่น $p \times (i : t)$ หมายถึง นักเรียน (p) ทำข้อสอบ (i) บางข้อที่อยู่ในแบบทดสอบ (t)

แหล่งความแปรปรวนและองค์ประกอบความแปรปรวน

คะแนนจริงของผู้สอบ (True Score: T) ในทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) คือ ค่าเฉลี่ยของคะแนนจากการสอบซ้ำ ๆ ด้วยแบบทดสอบคู่ขนาน ความแปรปรวนของคะแนนจริงจึงเป็นความแปรปรวนของค่าเฉลี่ยของการสอบซ้ำ และความแปรปรวนของคะแนนสังเกต

เป็นผลรวมของความแปรปรวนของคะแนนจริงกับความแปรปรวนของความคลาดเคลื่อน นั่นคือ $\sigma_x^2 = \sigma_T^2 + \sigma_E^2$ แต่ในทฤษฎีการสรุปอ้างอิงนั้นเป็นโมเดลทางสถิติที่แบ่งความแปรปรวนเป็นหลายส่วน โดยใช้หลักการวิเคราะห์ความแปรปรวน (ANOVA)

ตัวอย่างเช่น ในกรณีที่เป็นการศึกษาแบบองค์ประกอบเดียว สามารถแบ่งส่วนของคะแนนบุคคลได้เป็น อิทธิพลของบุคคล (ความแปรปรวนของคะแนนเอกภพ) อิทธิพลของข้อสอบ (ความแปรปรวนเนื่องจากความยากของข้อสอบ) และส่วนที่เหลือ ซึ่งเป็นปฏิสัมพันธ์ของบุคคลกับข้อสอบ ความคลาดเคลื่อนอย่างเป็นระบบอื่นนอกจากบุคคลกับข้อสอบและความคลาดเคลื่อนแบบสุ่ม (Shavelson; & Webb, 1991 pp.17-18) ทั้งนี้ Shavelson and Webb ได้กล่าวถึงแต่ละส่วนไว้ดังนี้

คะแนนสังเกตของบุคคล (p) ที่ได้จากข้อสอบ (i) แทนด้วย X_{pi} และคะแนน X_{pi} นั้นเป็นผลรวมของ 3 พารามิเตอร์ คือ μ_p, μ_i และ μ ซึ่ง พารามิเตอร์ตัวแรกคือ คะแนนเอกภพของแต่ละคน (Universe Score: μ_p) เป็นคะแนนเฉลี่ยของบุคคลจากเอกภพของข้อสอบ กล่าวอีกอย่างได้ว่า คะแนนเอกภพของบุคคล (μ_p) เป็นค่าคาดหวัง (Expected Value: E) ของตัวแปรสุ่มของคะแนนสังเกต (X_{pi}) จากข้อสอบทั้งหมด นั่นคือ

$$\mu_p = E_i(X_{pi})$$

ซึ่งค่าคาดหวัง (E) ของตัวแปรสุ่ม (X_{pi}) นั้นเป็นแนวโน้มค่าเฉลี่ยจากข้อสอบทั้งหมด k ข้อ ซึ่งเป็นจำนวนอนันต์ (Infinity) นั่นคือ ค่าคาดหวังของคะแนนสังเกต คือ ค่าลิมิตของค่าเฉลี่ยของข้อสอบ k ข้อ เมื่อ k เข้าใกล้อนันต์ ($k \rightarrow \infty$) กำหนดโดย

$$E_i(X_{pi}) = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=1}^k X_{pi}$$

สำหรับค่า μ_i เป็นค่าเฉลี่ยของประชากรสำหรับข้อสอบข้อ i เป็นค่าคาดหวังของ X_{pi} จากบุคคลทั้งหมด กำหนดโดย

$$\mu_i = E_p X_{pi}$$

ส่วนพารามิเตอร์ μ คือค่าเฉลี่ยของคะแนนทั้งหมดทั้งจากประชากรและเอกภพ (Grand Mean) กำหนดโดย

$$\mu = E_p E_i X_{pi}$$

พารามิเตอร์ μ_p, μ_i และ μ ไม่สามารถสังเกตเห็นได้ สามารถแสดงค่า X_{pi} ในเทอมของ μ_p, μ_i และ μ กล่าวคือ คะแนนสังเกตของแต่ละคน (X_{pi}) สามารถแบ่งเป็นส่วนย่อย ๆ ตามการออกแบบการวัด ดังนี้ (Shavelson; & Webb, 1991 pp. 19-60)

1. กรณียออกแบบการวัดแบบไขว้ (Crossed Design)

1.1 กรณีการวัดที่มีหนึ่งองค์ประกอบแบบไขว้ (One-Facet, $p \times I$ Design)

คะแนนสังเกตของบุคคลในข้อสอบข้อหนึ่ง ได้รับอิทธิพลมาจากส่วนต่าง ๆ ดังนี้ (Shavelson; & Webb, 1991 p. 19)

$$\begin{aligned}
 X_{pi} &= \mu && \text{[ค่าคงที่หรือค่าเฉลี่ยของทั้งหมด]} \\
 &+ \mu_p - \mu && \text{[อิทธิพลของบุคคล]} \\
 &+ \mu_i - \mu && \text{[อิทธิพลของข้อสอบ]} \\
 &+ X_{pi} - \mu_p - \mu_i + \mu && \text{[ส่วนที่เหลือหรือ } e_{pi}\text{]}
 \end{aligned}$$

เมื่อ μ คือ ค่าเฉลี่ยของคะแนนทั้งหมดจากประชากรในเอกภพ (Grand Mean)

μ_p คือ คะแนนเอกภพของแต่ละคน

μ_i คือ คะแนนเฉลี่ยรายข้อ

จะพบว่า อิทธิพลอื่นที่ต่างไปจากค่าเฉลี่ยทั้งหมด (Grand Mean) มีการกระจายเกิดขึ้น แต่ค่าเฉลี่ยทั้งหมดมีค่าคงที่ จึงมีค่าความแปรปรวนเป็นศูนย์

การแจกแจงที่มีค่าเฉลี่ยเป็นศูนย์ และมีความแปรปรวน (σ^2) เรียกว่าองค์ประกอบความแปรปรวน (Variance Component) ประกอบด้วยส่วนแรกเป็นอิทธิพลของบุคคล ค่าเฉลี่ยความแตกต่างของคะแนนระหว่างบุคคลเป็นศูนย์ คือ

$$E_p(\mu_p - \mu) = E_p(\mu_p) - E_p(\mu) = \mu - \mu = 0$$

ความแปรปรวนของอิทธิพลของบุคคล ใช้สัญลักษณ์เป็น σ_p^2 เรียกว่าองค์ประกอบความแปรปรวนของบุคคล หรือความแปรปรวนของคะแนนเอกภพ เขียนสมการได้เป็น

$$\sigma_p^2 = E_p(\mu_p - \mu)^2$$

ค่าเฉลี่ยและองค์ประกอบความแปรปรวนของข้อสอบ จะมีค่าเฉลี่ยรายข้อเป็นศูนย์ และองค์ประกอบความแปรปรวนของข้อสอบใช้สัญลักษณ์เป็น σ^2

อิทธิพลส่วนสุดท้าย คือส่วนที่เหลือ มีค่าเฉลี่ยเป็นศูนย์และความแปรปรวนเป็น $\sigma_{pi,e}^2$

จะได้ว่าความแปรปรวนของคะแนนสังเกต เป็นผลรวมของ 3 องค์ประกอบความแปรปรวนข้างต้น ดังนี้

$$\sigma_{x_m}^2 = \sigma_p^2 + \sigma_i^2 + \sigma_{pi,e}^2$$

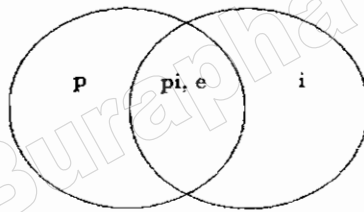
เมื่อ $\sigma_{x_{pi}}^2$ คือ ความแปรปรวนของคะแนนสังเกต (X_{pi})

σ_p^2 คือ ความแปรปรวนของคะแนนเอกภพ (μ_p)

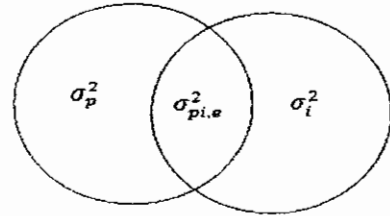
σ_i^2 คือ ความแปรปรวนของคะแนนเฉลี่ยรายข้อ (μ_i)

$\sigma_{pi,e}^2$ คือ ความแปรปรวนของคะแนนความคลาดเคลื่อนหรือส่วนที่เหลือ (e_{pi})

ความแปรปรวนของคะแนนที่สังเกตได้ในการวัดที่มีหนึ่งองค์ประกอบแบบไขว้ (One-Facet Crossed Design) จึงขึ้นอยู่กับ 4 แหล่ง คือ 1) ความแตกต่างระหว่างบุคคลหรือสิ่งที่วัด (Person Effect) 2) ความแตกต่างระหว่างความยากของข้อสอบ (Item Effect) 3) ปฏิสัมพันธ์ระหว่างบุคคลกับข้อสอบ (p x I Interaction) และ 4) ความคลาดเคลื่อนที่เกิดอย่างสุ่ม (Unsystematic or Random Error) แต่ความแปรปรวนจากแหล่งที่ 3 และที่ 4 ไม่สามารถแยกออกจากกันได้ เรียกว่าเป็นส่วนที่เหลือ (Residual) จึงเป็นแหล่งความแปรปรวนร่วมกัน ดังภาพ 2-11



แหล่งของความแปรปรวน
(Source of Variation)



แหล่งของความแปรปรวน
(Source of Variation)

ภาพที่ 2-11 แหล่งของความแปรปรวนและองค์ประกอบของความแปรปรวน สำหรับหนึ่งองค์ประกอบแบบไขว้ p x I (One Facet, p x I Design)

1.2 กรณีการวัดที่มีสององค์ประกอบแบบไขว้ (Two-Facet, p x r x o Design) คะแนนสังเกตของบุคคลในข้อสอบข้อหนึ่ง ได้รับอิทธิพลมาจากส่วนต่าง ๆ ดังนี้

(Shavelson; & Webb, 1991 pp. 22-23)

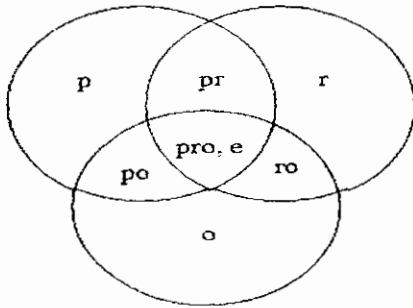
$$\begin{aligned}
 X_{pro} &= \mu && \text{[ค่าคงที่หรือค่าเฉลี่ยทั้งหมด]} \\
 &+ \mu_p - \mu && \text{[อิทธิพลของบุคคล]} \\
 &+ \mu_r - \mu && \text{[อิทธิพลของผู้ตรวจ]} \\
 &+ \mu_o - \mu && \text{[อิทธิพลของสถานการณ์]} \\
 &+ \mu_{pr} - \mu_p - \mu_r + \mu && \text{[ปฏิสัมพันธ์ระหว่างบุคคลกับผู้ตรวจ]} \\
 &+ \mu_{po} - \mu_p - \mu_o + \mu && \text{[ปฏิสัมพันธ์ระหว่างบุคคลกับสถานการณ์]} \\
 &+ \mu_{ro} - \mu_r - \mu_o + \mu && \text{[ปฏิสัมพันธ์ระหว่างผู้ตรวจกับสถานการณ์]} \\
 &+ \mu_{pro} - \mu_{po} - \mu_{pr} + \mu_{ro} + \mu_{ro} + \mu_p + \mu_r + \mu_o + \mu && \text{[อิทธิพลของส่วนที่เหลือ]}
 \end{aligned}$$

แต่ละองค์ประกอบยกเว้นค่าเฉลี่ยทั้งหมด (Grand Mean) นั้นมีการกระจาย นั่นคือ พฤติกรรมของผู้สอบเปลี่ยนแปลง ผู้ตรวจให้คะแนนแตกต่างกัน ซึ่งแต่ละการกระจายมีค่าเฉลี่ยเป็นศูนย์ และมีความแปรปรวน ความแปรปรวนของเอกภาพและประชากร จึงเป็นผลรวมขององค์ประกอบ ความแปรปรวน 7 ส่วน ดังนี้

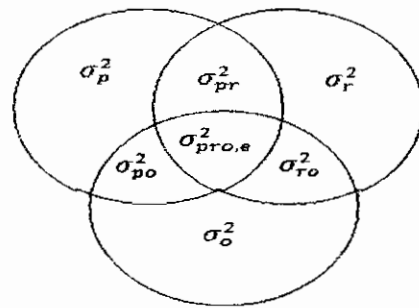
$$\sigma^2(X_{pro}) = \sigma_p^2 + \sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{po}^2 + \sigma_{ro}^2 + \sigma_{pro,e}^2$$

เมื่อ σ_{pro}^2	คือ	ความแปรปรวนของคะแนนสังเกต
σ_p^2	คือ	ความแปรปรวนของคะแนนเอกภาพของผู้สอบ (μ_p)
σ_r^2	คือ	ความแปรปรวนของคะแนนเฉลี่ยจากผู้ตรวจ (μ_r)
σ_o^2	คือ	ความแปรปรวนของคะแนนเฉลี่ยของสถานการณ์การสอบ แต่ละครั้ง (μ_o)
σ_{pr}^2	คือ	ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างผู้สอบกับผู้ตรวจ
σ_{po}^2	คือ	ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างผู้สอบกับสถานการณ์ ของการสอบแต่ละครั้ง
σ_{ro}^2	คือ	ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างผู้ตรวจกับ สถานการณ์ของการสอบแต่ละครั้ง
$\sigma_{pro,e}^2$	คือ	ความแปรปรวนของคะแนนความคลาดเคลื่อนหรือส่วนที่เหลือ (e_{pro})

เขียนแผนภาพเวนนีได้ดังภาพที่ 2-12



แหล่งของความแปรปรวน
(Source of Variation)



องค์ประกอบของความแปรปรวน
(Variation Component)

ภาพที่ 2-12 แหล่งของความแปรปรวนและองค์ประกอบของความแปรปรวนสำหรับสอง
องค์ประกอบแบบไขว้ $p \times r \times o$ (Two-Facet, $p \times r \times o$ Design)

2. กรณีออกแบบการวัดแบบแฝง (Nested Design)

2.1 กรณีการวัดที่มีองค์ประกอบแบบแฝง เมื่อ i แฝงใน p (One Facet; $i \times p$ Design) คะแนนสังเกตของบุคคลในข้อสอบข้อหนึ่ง ได้รับอิทธิพลมาจากส่วนต่าง ๆ ดังนี้ (Shavelson & Webb, 1991 pp. 49-50)

$$\begin{aligned}
 X_{pi} &= \mu && \text{[ค่าเฉลี่ยทั้งหมด]} \\
 + \mu_p - \mu &&& \text{[อิทธิพลของบุคคล]} \\
 + X_{pi} - \mu_p &&& \text{[ส่วนที่เหลือ]}
 \end{aligned}$$

การออกแบบแบบแฝง (Nested $i \times p$ Design) จะไม่ได้จำแนกเทอมออกมาเป็นอิทธิพลของข้อสอบ โดยอิทธิพลของข้อสอบเป็นส่วนหนึ่งของเทอมส่วนที่เหลือ เนื่องจากบุคคลทำข้อสอบต่างข้อกัน จึงไม่สามารถประมาณค่าอิทธิพลของข้อสอบได้อย่างอิสระจากปฏิสัมพันธ์ระหว่างบุคคลกับข้อสอบได้ สามารถเขียนอยู่ในอีกรูปได้เพื่อแสดงให้เห็นว่าอิทธิพลของข้อสอบเป็นส่วนหนึ่งของส่วนที่เหลือ ดังนี้

$$X_{pi} - \mu_p = (\mu_i - \mu) + (X_{pi} - \mu_i - \mu_p + \mu)$$

จะเห็นว่า องค์ประกอบความแปรปรวนของบุคคลหรือความแปรปรวนของคะแนนเอกภพนิยมเหมือนกับในการออกแบบแบบไขว้ นั่นคือ

$$\sigma_p^2 = E_p(\mu_p - \mu)^2$$

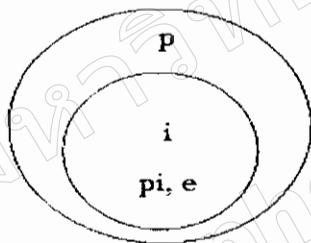
และองค์ประกอบความแปรปรวนของส่วนที่เหลือ มีค่าเป็น

$$\sigma_{i,pi,e}^2 = E_p E_i (X_{pi} - \mu_p)^2$$

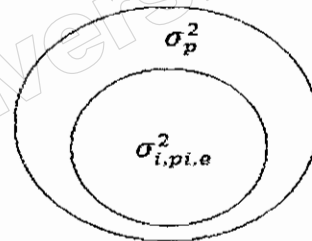
จากองค์ประกอบความแปรปรวนของส่วนที่เหลือ แสดงให้เห็นว่าอิทธิพลของข้อสอบนั้น รวมอยู่กับอิทธิพลของปฏิสัมพันธ์ระหว่างบุคคลและข้อสอบ ซึ่งมีแหล่งของความแปรปรวนที่ไม่เป็นระบบหรือไม่สามารถวัดได้รวมด้วยเช่นกัน ดังนั้น ความแปรปรวนของคะแนนสังเกตได้ (X_{pi})

$$\sigma^2(X_{pi}) = \sigma_p^2 + \sigma_{i,pi,e}^2$$

เขียนแผนภาพเวนนีได้ดังภาพที่ 2-13



แหล่งของความแปรปรวน
(Source of Variation)



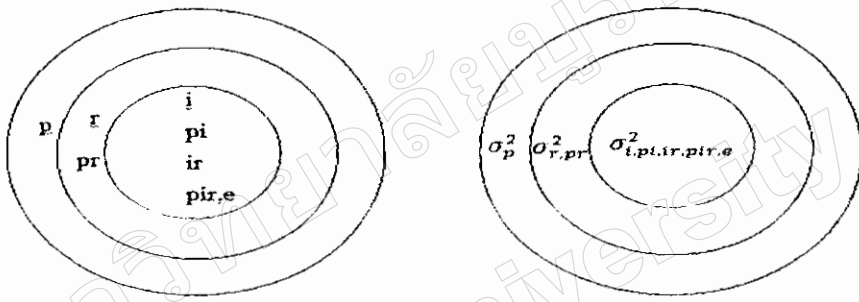
องค์ประกอบของความแปรปรวน
(Variation Component)

ภาพที่ 2-13 แหล่งความแปรปรวนและองค์ประกอบของความแปรปรวนสำหรับหนึ่งองค์ประกอบแบบแฝงเมื่อ i แฝงใน p (One Facet, $i: p$ Design)

2.2 กรณีการวัดที่มีองค์ประกอบแบบแฝง เช่น กรณี $i: r: p$ (Two Facets, $i: r: p$ Design) คะแนนสังเกตของบุคคลในข้อสอบข้อหนึ่ง ได้รับอิทธิพลมาจากส่วนต่าง ๆ ดังนี้ (Shavelson & Webb, 1991 p. 60)

$$\begin{aligned}
 X_{pir} &= \mu \\
 &+ \mu_p - \mu \\
 &+ \mu_{pr} - \mu_p \\
 &+ X_{pir} - \mu_{pr} \\
 \sigma_{x_{pir}}^2 &= \sigma_p^2 + \sigma_{r,pr}^2 + \sigma_{i,pi,ir,pir,e}^2
 \end{aligned}$$

เขียนแผนภาพเวนน ได้ดังภาพที่ 2-14



แหล่งของความแปรปรวน
(Source of Variation)

องค์ประกอบของความแปรปรวน
(Variation Component)

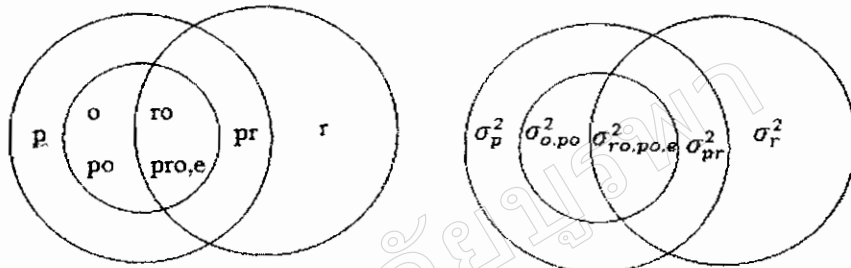
ภาพที่ 2-14 แหล่งความแปรปรวนและองค์ประกอบของความแปรปรวนสำหรับสององค์ประกอบแบบแฝง กรณี i: r: p (Two Facet, i: r: p Design)

2.3 กรณีออกแบบการวัดแบบผสม (Confound Design) เช่น กรณี (o: p) x r Design คะแนนสังเกตของบุคคลในข้อสอบข้อหนึ่ง ได้รับอิทธิพลมาจากส่วนต่าง ๆ ดังนี้ (Shavelson & Webb, 1991 pp. 52-53)

$$\begin{aligned}
 X_{pro} &= \mu \\
 &+ \mu_p - \mu \\
 &+ \mu_r - \mu \\
 &+ \mu_{po} - \mu_p
 \end{aligned}$$

$$\begin{aligned}
 & + \mu_{pr} - \mu_p - \mu_r + \mu \\
 & + X_{pro} - \mu_{pr} - \mu_{po} + \mu_p \\
 \text{และ } \sigma_{x_{pro}}^2 & = \sigma_p^2 + \sigma_r^2 + \sigma_{op}^2 + \sigma_{pr}^2 + \sigma_{ro,pro,e}^2
 \end{aligned}$$

เขียนแผนภาพเวนน ได้ดังภาพที่ 2-15



แหล่งของความแปรปรวน
(Source of Variation)

องค์ประกอบของความแปรปรวน
(Variation Component)

ภาพที่ 2-15 แหล่งความแปรปรวนและองค์ประกอบของความแปรปรวนสำหรับสององค์ประกอบแบบผสมกรณี (o : p) x r (Two Facet, (o : p) x r Design)

การประมาณค่าองค์ประกอบความแปรปรวน

การประมาณค่าองค์ประกอบความแปรปรวน สามารถนำการวิเคราะห์ความแปรปรวน (Analysis of Variance: ANOVA) มาประยุกต์ใช้กับการศึกษา G (G-Study) (Shavelson & Webb, 1991, pp. 26-27) ดังนี้

1. กรณีการวัดที่มีองค์ประกอบเดียวแบบไขว้ p x i

ตารางที่ 2-3 ANOVA สำหรับการวัดที่มีองค์ประกอบเดียวแบบไขว้ p x i

แหล่งความแปรปรวน	Sum of Squares	df	Mean Squares	Expected Mean Squares
Person (p)	SS_p	$n_p - 1$	$MS_p = SS_p / df_p$	$E(MS_p) = \sigma_{pi,e}^2 + n_i \sigma_p^2$
Item (i)	SS_i	$n_i - 1$	$MS_i = SS_i / df_i$	$E(MS_i) = \sigma_{pi,e}^2 + n_p \sigma_p^2$
Pi, e	$SS_{pi,e}$	$(n_p - 1)(n_i - 1)$	$MS_{pi,e} = SS_{pi,e} / df_{pi,e}$	$E(MS_{pi,e}) = \sigma_{pi,e}^2$

ที่มา: Shavelson and Webb (1991, pp. 26-27)

จากตารางในการประมาณค่า ค่าเฉลี่ยกำลังสองที่คาดหวัง (Expected Mean Squares: E(MS)) จะใช้ค่าเฉลี่ยกำลังสองของค่าสังเกต (Observed Mean Square: MS) เป็นตัวประมาณค่า และการประมาณค่า σ^2 ใช้ $\hat{\sigma}^2$ เป็นตัวประมาณค่า จะได้

$$MS_p = \hat{\sigma}_{pi,e}^2 + n_i \hat{\sigma}_p^2$$

$$MS_p = \hat{\sigma}_{pi,e}^2 + n_p \hat{\sigma}_i^2$$

$$MS_{pi,e} = \hat{\sigma}_{pi,e}^2$$

เมื่อจัดสมการใหม่ ได้เป็น

$$\hat{\sigma}_{pi,e}^2 = MS_{pi,e}$$

$$\hat{\sigma}_{pi,e}^2 = (MS_i - \hat{\sigma}_{pi,e}^2) / n_p = (MS_i - MS_{pi,e}) / n_p$$

$$\hat{\sigma}_p^2 = (MS_p - \hat{\sigma}_{pi,e}^2) / n_i = (MS_p - MS_{pi,e}) / n_i$$

2. กรณีการวัดที่มีสององค์ประกอบแบบไขว้ $p \times i \times r$

ตารางที่ 2-4 ANOVA สำหรับการวัดที่มีสององค์ประกอบแบบไขว้

แหล่ง ความ แปร ปรวน	Sum of Squares	df	Mean Squares	Expected Mean Square (EMS)
Person (p)	SS_p	$n_p - 1$	SS_p / df_p	$\sigma_{pir,e}^2 + n_i \sigma_{pi}^2 + n_i \sigma_{pi}^2 + n_i n_r \sigma_p^2$
Item (i)	SS_i	$n_i - 1$	SS_i / df_i	$\sigma_{pir,e}^2 + n_p \sigma_{ir}^2 + n_r \sigma_{pi}^2 + n_p n_r \sigma_i^2$
Rater (r)	SS_r	$n_r - 1$	SS_r / df_r	$\sigma_{pir,e}^2 + n_p \sigma_{ir}^2 + n_i \sigma_{pr}^2 + n_p \sigma_r^2$
pi	SS_{pi}	$(n_p - 1)(n_i - 1)$	SS_{pi} / df_{pi}	$\sigma_{pir,e}^2 + n_r \sigma_{pi}^2$
pr	SS_{pr}	$(n_p - 1)(n_r - 1)$	SS_{pr} / df_{pr}	$\sigma_{pir,e}^2 + n_i \sigma_{pr}^2$
ir	SS_{ir}	$(n_i - 1)(n_r - 1)$	SS_{ir} / df_{ir}	$\sigma_{pir,e}^2 + n_p \sigma_{ir}^2$
pir,e	$SS_{pir,e}$	$(n_p - 1)(n_i - 1)(n_r - 1)$	SS_{pir} / df_{pir}	$\sigma_{pir,e}^2$

ที่มา: Shavelson and Webb (1991, p. 33)

การประมาณค่าดังกล่าวในบางสถานการณ์ องค์ประกอบความแปรปรวนอาจจะมีค่าติดลบได้ เนื่องจากการกำหนดโมเดลการวัดผิดพลาดหรือเกิดจากความคลาดเคลื่อนจากการสุ่มตัวอย่าง หากเหตุการณ์แบบนี้ขึ้น Cronbach et al. (1972) ได้เสนอให้ประมาณค่าที่ติดลบมีค่าเป็นศูนย์ แล้วจึงคำนวณสมการอื่นต่อไป ทำให้ง่ายสำหรับการคำนวณสมการอื่น ๆ แต่จะเกิดอคติในการประมาณ องค์ประกอบความแปรปรวนค่าอื่น ๆ อีกวิธีหนึ่งเสนอโดย Brennan (Shavelson, & Webb, 1991; citing Brennan, 1983) ซึ่งเสนอให้ใช้ค่าศูนย์แทนค่าติดลบเช่นกัน แต่ในการประมาณค่า องค์ประกอบอื่น ๆ ให้ใช้ค่าเดิมในการคำนวณ เมื่อประมาณค่าหมดในทุกสมการแล้วจึงรายงาน ค่าประมาณที่ติดลบเป็นศูนย์ วิธีการนี้จะยุ่งยากในการคำนวณแต่จะแก้ปัญหาคติในการประมาณค่า องค์ประกอบในสมการต่าง ๆ ได้

ความแปรปรวนของคะแนนความคลาดเคลื่อน

คะแนนเอกภพ (Universe Score) คือ ค่าเฉลี่ยของการวัดซ้ำหลาย ๆ ครั้ง ตามเงื่อนไขของการวัดในเอกภพของการสุ่มอย่างอิสระ สำหรับความคลาดเคลื่อนของการวัด (E_{pi}) ถูกจำแนกออกมาเป็นความคลาดเคลื่อนจากองค์ประกอบหรือกลุ่มเงื่อนไขของการวัด (E_i) และความคลาดเคลื่อนจากแหล่งที่เหลืออื่น ๆ (e_{pi}) ซึ่งการวัดแต่ละครั้งไม่จำเป็นต้องเป็นข้อสอบคู่ขนาน ส่วนความแปรปรวนของค่าคาดหวังของคะแนนที่สังเกตได้เป็นผลรวมของความแปรปรวนของคะแนนเอกภพ ($\sigma_{\mu_p}^2$ หรือ σ_p^2) กับความแปรปรวนของความคลาดเคลื่อนจากแหล่งอื่น ๆ (σ_{ep}^2) ดังนี้

$$X_{pi} = T_{pi} + E_{pi} + e_{pi}$$

$$\sigma_{x_p}^2 = \sigma_{\mu_p}^2 + \sigma_{E_i}^2 + \sigma_{e_p}^2$$

ความแปรปรวนของความคลาดเคลื่อนจากองค์ประกอบต่าง ๆ ของการวัดแบ่งออกเป็น 2 ประเภท ได้แก่ (Shavelson & Webb, 1991, p. 84; ศิริชัย กาญจนวาสิ, 2550, หน้า 17-18)

1. ความแปรปรวนของคะแนนความคลาดเคลื่อนแบบสัมบูรณ์ (Absolute Error Variance) ใช้สัญลักษณ์ σ_{Abs}^2 หรือ $\sigma^2(\Delta)$ คือ ความแปรปรวนของผลต่างระหว่างคะแนนสังเกตและคะแนนเอกภพ ซึ่งมีค่าเท่ากับ ผลรวมของความแปรปรวนที่ประมาณได้ทั้งหมด ยกเว้นความแปรปรวนจากแหล่งบุคคล σ_p^2 หรือสิ่งที่ต้องการวัด ดังนั้นจึงคำนวณได้จาก ผลรวมของความแปรปรวนของคะแนนจากแหล่งต่าง ๆ ยกเว้น σ_p^2 ดังเช่นกรณี $p \times i \times r$ คำนวณ σ_{Abs}^2 ได้จาก

$$\sigma_{Abs}^2 = \sigma_i^2 + \sigma_r^2 + \sigma_{pi}^2 + \sigma_{pr}^2 + \sigma_n^2 + \sigma_{pir,e}^2$$

2. ความแปรปรวนของคะแนนความคลาดเคลื่อนแบบสัมพัทธ์ (Relative Error Variance) ใช้สัญลักษณ์ σ_{Rel}^2 หรือ $\sigma^2(\delta)$ คือ ความแปรปรวนของผลต่างระหว่างส่วนเบี่ยงเบนคะแนนสังเกต (ความแตกต่างของคะแนนสังเกตจากคะแนนเฉลี่ยของประชากรของคะแนนสังเกต) กับส่วนเบี่ยงเบนคะแนนเอกภาพ (ความแตกต่างของคะแนน เอกภาพจากคะแนนเฉลี่ยของประชากรของคะแนนเอกภาพ) จึงคำนวณได้จากผลรวมของความแปรปรวนของคะแนนจากแหล่งต่าง ๆ ที่มีปฏิสัมพันธ์กับผู้สอบ ดังเช่น กรณี $p \times i \times r$ คำนวณ σ_{Rel}^2 ได้จาก

$$\sigma_{Rel}^2 = \sigma_{pi}^2 + \sigma_{pr}^2 + \sigma_{pri,e}^2$$

แหล่งความคลาดเคลื่อนแบบสมบูรณ์และแบบสัมพัทธ์ จะแตกต่างกันไปตามรูปแบบการวัดที่ได้ออกแบบไว้ มีรายละเอียดดังต่อไปนี้ (Shavelson & Webb, 1991, pp. 84-92)

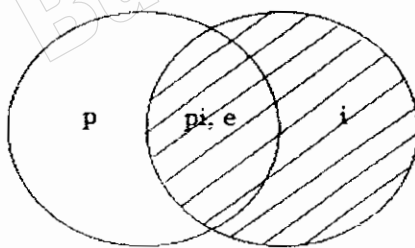
1. กรณีการวัดที่มีองค์ประกอบเดียว แบบ $p \times i$ Design (One-Facet Crossed Design)

$$\sigma_{Abs}^2 = \frac{\sigma_i^2}{n_i} + \frac{\sigma_{pi,e}^2}{n_i}$$

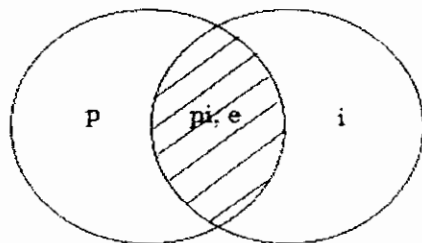
และ

$$\sigma_{Rel}^2 = \frac{\sigma_{pi,e}^2}{n_i}$$

เขียนแผนภาพเวนน ได้ดังภาพที่ 2-16



ความคลาดเคลื่อนแบบสมบูรณ์
(Absolute Error)



ความคลาดเคลื่อนแบบสัมพัทธ์
(Relative Error)

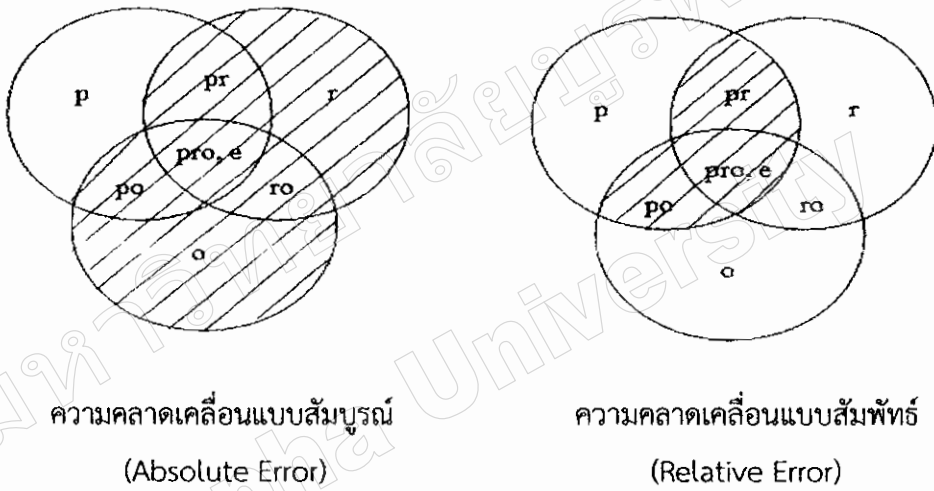
ภาพที่ 2-16 แหล่งความคลาดเคลื่อนแบบสมบูรณ์และความคลาดเคลื่อนแบบสัมพัทธ์สำหรับ $p \times i$ Design

2. กรณีการวัดที่มีสององค์ประกอบ $p \times r \times o$ Design

$$\sigma_{Abs}^2 = \frac{\sigma_r^2}{n_r} + \frac{\sigma_o^2}{n_o} + \frac{\sigma_{pr}^2}{n_r} + \frac{\sigma_{po}^2}{n_o} + \frac{\sigma_{ro}^2}{n_r n_o} + \frac{\sigma_{pro,e}^2}{n_r n_o}$$

$$\text{และ } \sigma_{Rel}^2 = \frac{\sigma_{pr}^2}{n_r} + \frac{\sigma_{po}^2}{n_o} + \frac{\sigma_{pro,e}^2}{n_r n_o}$$

เขียนแผนภาพเวนน ได้ดังภาพที่ 2-17



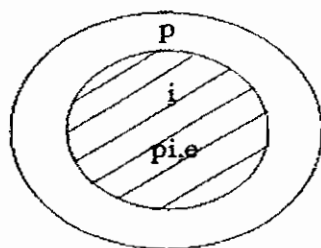
ภาพที่ 2-17 แหล่งความคลาดเคลื่อนแบบสัมบูรณ์และความคลาดเคลื่อนแบบสัมพัทธ์สำหรับ $p \times r \times o$ Design

3. กรณีการวัดที่มีองค์ประกอบเดียว แบบ $i \times p$ Design

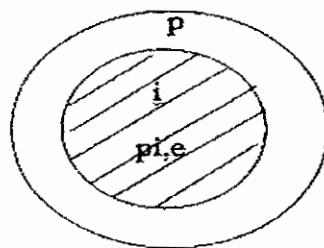
$$\sigma_{Abs}^2 = \frac{\sigma_{i,pi,e}^2}{n_i}$$

$$\text{และ } \sigma_{Rel}^2 = \frac{\sigma_{i,pi,e}^2}{n_i}$$

เขียนแผนภาพเวนน ได้ดังภาพที่ 2-18



ความคลาดเคลื่อนแบบสัมบูรณ์
(Absolute Error)



ความคลาดเคลื่อนแบบสัมพัทธ์
(Relative Error)

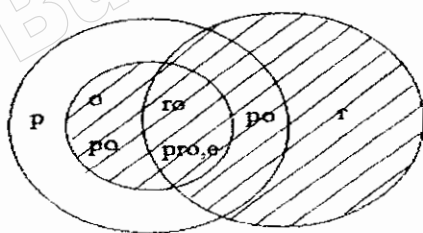
ภาพที่ 2-18 แหล่งความคลาดเคลื่อนแบบสัมบูรณ์และความคลาดเคลื่อนแบบสัมพัทธ์สำหรับ $i \times p$ Design

4. กรณีการวัดที่มีสององค์ประกอบ ($o : p$) \times r Design

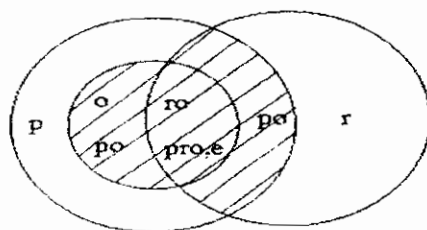
$$\sigma_{Abs}^2 = \frac{\sigma_r^2}{n_r'} + \frac{\sigma_{pr}^2}{n_r'} + \frac{\sigma_{o,po}^2}{n_o'} + \frac{\sigma_{ro,pro,e}^2}{n_r'n_o'}$$

$$\sigma_{Rel}^2 = \frac{\sigma_{pr}^2}{n_r'} + \frac{\sigma_{o,po}^2}{n_o'} + \frac{\sigma_{ro,pro,e}^2}{n_r'n_o'}$$

เขียนแผนภาพเวเนน ได้ดังภาพที่ 2-19



ความคลาดเคลื่อนแบบสัมบูรณ์
(Absolute Error)



ความคลาดเคลื่อนแบบสัมพัทธ์
(Relative Error)

ภาพที่ 2-19 แหล่งความคลาดเคลื่อนแบบสัมบูรณ์และความคลาดเคลื่อนแบบสัมพัทธ์สำหรับ $p \times r$ Design

ค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient)

ทฤษฎีการสรุปอ้างอิง นอกจากมุ่งให้ความสำคัญกับองค์ประกอบความแปรปรวนและความคลาดเคลื่อนของการวัดแล้วยังสามารถประมาณค่าสัมประสิทธิ์การสรุปอ้างอิง ซึ่งคล้ายกับค่าสัมประสิทธิ์ความเชื่อมั่นในทฤษฎีการทดสอบแบบดั้งเดิม (Shavelson & Webb, 1991, p. 92) ที่เป็นอัตราส่วนระหว่างความแปรปรวนของคะแนนจริง กับความแปรปรวนของคะแนนที่สังเกตได้

ค่าสัมประสิทธิ์การสรุปอ้างอิงเป็นอัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภพและความแปรปรวนของคะแนนสังเกตที่คาดหวัง สามารถประมาณค่าได้จากกำลังสองของค่าสหสัมพันธ์ระหว่างคะแนนเอกภพกับคะแนนสังเกต (Cronbach et al., 1972) ดังนั้น สัมประสิทธิ์การสรุปอ้างอิงจึงเป็นอัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภพ กับความแปรปรวนของค่าคาดหวังของคะแนนสังเกต ดังนี้

$$G\text{-coefficient } (\rho^2) = \frac{\sigma_p^2}{\sigma_p^2 + \text{ErrorVariance}}$$

เนื่องจากความแปรปรวนของคะแนนความคลาดเคลื่อนจากองค์ประกอบต่าง ๆ ของการวัดมี 2 ประเภท จึงทำให้สัมประสิทธิ์การสรุปอ้างอิงมี 2 ประเภท ดังนี้

1. สัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสมบูรณ์ (ρ^2_{Abs} หรือ $\rho^2(\Delta)$)

เมื่อคะแนนความคลาดเคลื่อน เป็นความแปรปรวนของคะแนนความคลาดเคลื่อนสมบูรณ์ สัมประสิทธิ์นี้บ่งบอกความเชื่อมั่นของแบบทดสอบ ในสถานการณ์ของการตัดสินใจที่ขึ้นกับคะแนนของผู้สอบตามลำพัง ไม่มีการเปรียบเทียบภายในกลุ่มหรือระหว่างกลุ่ม เช่น แบบทดสอบอิงเกณฑ์ที่ใช้ตรวจสอบความสามารถของผู้สอบเทียบกับเกณฑ์ที่กำหนดไว้ล่วงหน้า

การประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงเชิงสมบูรณ์เทียบเท่ากับค่าความเชื่อมั่นในทฤษฎีการทดสอบแบบดั้งเดิม Brennan et al. (Shavelson & Webb, 1991, p. 93; citing Brennan; & Kane, 1977, An index of dependability for mastery test. Journal of Educational Measurement. 14: 227-289) เรียกสัมประสิทธิ์นี้ว่า ดัชนีความเชื่อถือได้ (Index of Dependability) และใช้สัญลักษณ์ ϕ (phi) หรือ ρ^2_{Abs} มีสูตรดังนี้ (Shavelson & Webb, 1991, p. 93)

$$\phi = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_{Abs}^2}$$

หรือ (Brennan, 2001, p. 35)

$$\phi = \frac{\sigma(p)}{\sigma(p) + \sigma_{\Delta}^2}$$

2. สัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (ρ_{Rel}^2)

เมื่อคะแนนความคลาดเคลื่อนเป็นความแปรปรวนของคะแนนความคลาดเคลื่อนสัมพัทธ์สัมประสิทธิ์นี้บ่งบอกความเชื่อมั่นของแบบทดสอบ ในสถานการณ์ของการตัดสินใจที่มีการเปรียบเทียบคะแนนระหว่างผู้สอบ เช่น แบบทดสอบอิงกลุ่มที่เปรียบเทียบคะแนนระหว่างผู้สอบด้วยการเข้าโค้งการแจกแจงคะแนนและตัดเกรดหรือระดับผลการเรียน หรือการสอบแข่งขันเพื่อหาผู้ที่ได้คะแนนสูงเข้าศึกษา การประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงเชิงสัมพัทธ์ มีสูตรดังนี้ (Shavelson & Webb, 1991, p. 93)

$$E\rho_{Rel}^2 = \frac{\sigma_p^2}{E\sigma^2(X_{pi})} = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_{Rel}^2}$$

หรือ (Brennan, 2001, p. 35)

$$E\rho^2 = \frac{\sigma^2(p)}{\sigma^2(p) + \sigma^2(\delta)}$$

กล่าวโดยสรุปได้ว่า ทฤษฎีการสรุปอ้างอิงขยายมาจากทฤษฎีการทดสอบแบบดั้งเดิมใน 4 ลักษณะ คือ 1) สามารถประมาณค่าขนาดของความแปรปรวนจากหลาย ๆ แหล่งในบริบทการวัดได้ พร้อมกันจากการวิเคราะห์เพียงครั้งเดียว 2) สามารถประมาณค่าสัมประสิทธิ์ความเชื่อมั่น ที่เรียกว่า สัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) มุ่งเน้นไปที่องค์ประกอบความแปรปรวนของแต่ละแหล่งของความคลาดเคลื่อนที่ส่งผลต่อการวัด 3) สามารถแยกการตัดสินใจเป็น 2 ลักษณะ คือ การตัดสินใจเชิงสมบูรณ์ (Absolute Decision) และการตัดสินใจเชิงสัมพัทธ์ (Relative Decision) และ 4) มีการวิเคราะห์ 2 ขั้นตอนแยกกัน คือ การศึกษา G และการศึกษา D โดยการศึกษา G เป็นการประมาณค่าน้ำหนักของแหล่งความคลาดเคลื่อนของการวัดที่เป็นไปได้ทั้งหมด การศึกษา D ใช้ข้อมูลจากการศึกษา G ในการออกแบบการวัดที่ทำให้มีความคลาดเคลื่อนน้อยที่สุดตามเป้าหมายที่ต้องการ การประมาณค่าองค์ประกอบความแปรปรวนในการศึกษา G จะใช้ในการออกแบบกระบวนการวัดให้มีความคุ้มค่าทั้งเวลาและค่าใช้จ่าย สามารถนำไปใช้ได้หลายรูปแบบของการวัด ทั้งรูปแบบเป็นแบบไขว้ (Crossed) แบบแฝง (Nested) หรือแบบแฝงบางส่วน (Partially Nested)

ตอนที่ 5 งานวิจัยที่เกี่ยวข้องกับการใช้ผังมโนทัศน์สำหรับประเมินผลการเรียนรู้และ ทฤษฎีการสรุปอ้างอิง

งานวิจัยที่เกี่ยวข้องกับการใช้ผังมโนทัศน์สำหรับประเมินผลการเรียนรู้

จากที่กล่าวไปแล้วข้างต้นว่าผังมโนทัศน์ไม่ได้เป็นเทคนิคหรือวิธีการที่เกิดขึ้นใหม่ เพราะ ได้รับการพัฒนาและนำมาใช้ในการศึกษามานานกว่า 30 ปี (Buldu, 2010) โดย Prof. Joseph D. Novak จากมหาวิทยาลัยคอร์เนล ประเทศสหรัฐอเมริกา ในปี ค.ศ. 1960 เป็นบุคคลสำคัญในการ คิดค้นและประยุกต์ใช้ในการเรียนการสอนและการวัดและประเมินผล ปี ค.ศ. 1984 Novak and Gowin ได้พัฒนาเกณฑ์การให้คะแนนผังมโนทัศน์ขึ้น ซึ่งเกณฑ์การให้คะแนนมีลักษณะเป็นเกณฑ์ แบบแยกองค์ประกอบ ดังนี้ 1) ประพจน์ที่แสดงความสัมพันธ์ระหว่าง 2 มโนทัศน์ที่เชื่อมกันโดย คำเชื่อมมีความหมายที่สมเหตุสมผลและตรงกับความสัมพันธ์หรือไม่ ถ้ามีความหมายแต่ละข้อความที่ แสดงความสัมพันธ์ ได้ 1 คะแนน 2) ผังมโนทัศน์มีการจัดลำดับหรือไม่แต่ละมโนทัศน์ที่อยู่รองลงมา เป็นมโนทัศน์ที่เฉพาะเจาะจงและมีความกว้างน้อยกว่ามโนทัศน์ที่อยู่ลำดับแรก ๆ ได้ 5 คะแนน 3) มีการเชื่อมโยงกันระหว่างสายของมโนทัศน์หรือไม่ ถ้ามีการเชื่อมโยงสมเหตุสมผลและมีนัยสำคัญ ได้ 10 คะแนน ถ้าคำเชื่อมสมเหตุสมผลแต่ไม่ได้แสดงการสังเคราะห์ระหว่างชุดของมโนทัศน์หรือ ข้อความที่เกี่ยวข้องกัน ได้ 2 คะแนน 4) ผู้เรียนสามารถยกตัวอย่างของเหตุการณ์หรือวัตถุที่เป็น ตัวแทนของมโนทัศน์ที่สร้างขึ้นได้ถูกต้อง ได้ 1 คะแนน เกณฑ์การให้คะแนนของ Novak and Gowin (1984) ต้องใช้เวลามากแต่รายละเอียดที่ได้จะมีประโยชน์เมื่อต้องการวินิจฉัยข้อผิดพลาดของผู้ถูก ประเมิน (Kinchin et al. 2000; Klein, Chung, Osmundson, Herl & O'Neil, 2000; McClure, Sonak & Suen, 1999; XiuFeng & Hinchey, 1996) ต่อไปนี้เป็นงานวิจัยที่เกี่ยวข้องกับการใช้ ผังมโนทัศน์สำหรับการประเมินผลการเรียนรู้

วิยะดา ระวังสุข (2545) ได้นำผังมโนทัศน์มาใช้ประเมินความคิดรวบยอดวิชาวิทยาศาสตร์ ของนักเรียนระดับชั้นประถมศึกษาปีที่ 4 โรงเรียนวัดลำนาว จังหวัดนครราชสีมา ผลการวิจัยปรากฏ ว่า แบบประเมินความคิดรวบยอดที่พัฒนาขึ้น เป็นผังมโนทัศน์แบบเติมคำลงในผังและผังมโนทัศน์แบบ สร้างผังจากคำ โดยมีเนื้อหาวิทยาศาสตร์ 2 สารระ คือสารระที่ 1 กระบวนการเปลี่ยนแปลงของโลก และสารระที่ 2 ดาราศาสตร์และอวกาศ ผังมโนทัศน์แบบเติมคำลงในผังในสารระที่ 1 กระบวนการ เปลี่ยนแปลงของโลก และสารระที่ 2 ดาราศาสตร์และอวกาศ มีค่าความเที่ยงเท่ากับ 0.46 และ 0.76 ตามลำดับ เมื่อเปรียบเทียบประเภทของการให้คะแนน 3 แบบ พบว่าการตรวจให้คะแนนแบบ convergence score มีค่าสัมประสิทธิ์การสรุปอ้างอิงสูงสุด ความตรงตามสภาพของผลที่ได้จาก

ผังมโนทัศน์กับผลการวินิจฉัยผังมโนทัศน์ที่คลาดเคลื่อนโดยครูผู้สอน พบว่าผังมโนทัศน์แบบสร้างผังจากคำ มีค่าความสอดคล้องอยู่ระหว่าง .72-.92 ส่วนผังมโนทัศน์แบบเติมคำลงผัง มีค่าความสอดคล้องอยู่ระหว่าง .40-.68

หัตถ์ณี ชูขวัญ (2548) ได้เปรียบเทียบคุณภาพของการประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้ผังมโนทัศน์ที่มีการตรวจให้คะแนนด้วยวิธีที่แตกต่างกัน ประชากรเป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 โรงเรียนสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน ผลการวิจัยปรากฏว่า เกณฑ์การให้คะแนนของ Novak and Gowin (1984) พิจารณาจากผังมโนทัศน์ในด้านคุณภาพ 4 ด้าน ซึ่งแต่ละด้านจะมีคะแนนเต็ม 4 คะแนน คือ ประพจน์ การจัดลำดับขั้นตอน การเชื่อมระหว่างมโนทัศน์ และตัวอย่าง และเกณฑ์การให้คะแนนของ Bolte (1999) พิจารณาจากผังมโนทัศน์ในด้านการจัดผังมโนทัศน์ 6 คะแนน และความถูกต้อง 4 คะแนน ค่าความตรงตามสภาพของคะแนนการประเมินผังมโนทัศน์ เมื่อตรวจให้คะแนนด้วยวิธีของ Novak and Gowin (1984) มีค่าสูงกว่าการตรวจให้คะแนนของด้วยวิธีของ Bolte (1999) อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ผังมโนทัศน์เมื่อตรวจให้คะแนนด้วยวิธีของ Novak and Gowin (1984) และตรวจให้คะแนนด้วยวิธีของ Bolte (1999) มีความตรงเชิงโครงสร้างทั้งสองวิธี ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ ของวิธีการให้คะแนนของ Novak and Gowin (1984) ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ โดยเฉลี่ยสูงกว่าวิธีการให้คะแนนของ Bolte (1999) อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินผังมโนทัศน์ เมื่อตรวจให้คะแนนด้วยวิธีของ Novak and Gowin (1984) มีค่าสูงกว่าการตรวจให้คะแนนของ Bolte (1999) จำนวนผู้ประเมินและจำนวนผังที่เหมาะสม ที่ทำให้ค่าสัมประสิทธิ์สรุปอ้างอิงอยู่ในเกณฑ์เหมาะสมเมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเพื่อนำไปใช้อิงเกณฑ์ เท่ากับ 0.8 พบว่า วิธีการให้คะแนนของ Novak and Gowin (1984) ต้องใช้ผู้ประเมินจำนวน 2 คน และจำนวนผังมโนทัศน์จำนวน 8 ผัง หรือผู้ประเมิน 1 คน ผังมโนทัศน์ 12 ผัง และวิธีการให้คะแนนของ Bolte (1999) ต้องใช้ผู้ประเมิน 4 คน และจำนวนผังมโนทัศน์ 12 ผัง หรือผู้ประเมิน 3 คน ผังมโนทัศน์ 16 ผัง

McClure et al. (1999) ได้ศึกษาการนำผังมโนทัศน์มาใช้ในการประเมินในชั้นเรียนด้านความเที่ยง ความตรงและความเป็นไปได้ในทางปฏิบัติ ได้ทำการศึกษาวิจัยกับนักศึกษาที่เพิ่งเข้ามาใหม่ในปี 1 ที่ลงทะเบียนเรียนวิชาทางจิตวิทยา โดยแบ่งกลุ่ม 2 กลุ่ม ผู้วิจัยได้ให้ความรู้แก่กลุ่มตัวอย่างเกี่ยวกับเทคนิคผังมโนทัศน์ กลุ่มละ 90 นาที ผลการศึกษาแสดงว่าด้านความเที่ยงใช้เปรียบเทียบวิธีการใช้คะแนน 6 วิธี พบว่ามีค่าสัมประสิทธิ์สหสัมพันธ์อยู่ระหว่าง .23 - .76 วิธีการตรวจให้คะแนนที่ให้ความเที่ยงสูงสุด คือวิธีการตรวจให้คะแนนผังมโนทัศน์โดยดูจากการเชื่อมคำมโนทัศน์และคำเชื่อม ส่วนในด้านความตรงตามสภาพพบว่าผังมโนทัศน์ที่นักเรียนสร้างขึ้นสอดคล้องกับ

ผังมโนทัศน์เฉลยของผู้วิจัยและมีความเป็นไปได้ที่จะนำผังมโนทัศน์มาใช้ในการประเมินผลในห้องเรียน

Ruiz-Primo et al. (2001) ได้ศึกษาเปรียบเทียบเทคนิคผังมโนทัศน์ 2 แบบ ได้แก่ ผังมโนทัศน์ที่กำหนดทิศทางให้สูง คือผังมโนทัศน์แบบเติมคำลงในผัง ซึ่งจำแนกออกเป็น 2 วิธี คือ เติมคำมโนทัศน์และเติมคำเชื่อมบนเส้น ผังมโนทัศน์ที่กำหนดทิศทางให้ต่ำ คือผังมโนทัศน์ที่สร้างผังจากคำ การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) ศึกษาความไวต่อการเลือกเติมคำมโนทัศน์และการเติมคำเชื่อมบนเส้น 2) ศึกษาลักษณะความคู่ขนานกันของการเลือกเติมคำมโนทัศน์และเติมคำเชื่อมบนเส้น 3) เทคนิคผังมโนทัศน์ทั้ง 2 แบบให้ข้อมูลเกี่ยวกับความเข้าใจในการเชื่อมโยงของนักเรียนอย่างไร กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนระดับชั้นมัธยมศึกษา จำนวน 152 คน และครูผู้สอนวิชาเคมี 2 คน โดยกลุ่มตัวอย่างถูกฝึกให้สร้างผังมโนทัศน์ทั้ง 2 รูปแบบ โดยใช้เวลา 50 นาที ผลการวิจัยแสดงว่า นักเรียนไม่มีความไวต่อการเลือกเติมคำมโนทัศน์และคำเชื่อมบนเส้น และพบว่าเทคนิคผังมโนทัศน์แบบการเติมคำลงในผังและการเติมคำเชื่อมบนเส้นไม่คู่ขนานกัน แต่มีความเป็นไปได้สูงในการนำไปใช้ ส่วนคะแนนที่ได้จากผังมโนทัศน์ที่สร้างผังจากคำ พบว่า องค์ความรู้ของนักเรียนไม่สมบูรณ์เมื่อเทียบกับผังเฉลย แต่พบว่าเทคนิคผังมโนทัศน์แบบสร้างผังจากคำส่งผลต่อโครงสร้างองค์ความรู้ของนักเรียนได้ดีกว่าผังมโนทัศน์แบบเติมคำลงในผัง

West, Park, Pomeroy and Sandoval (2002) ได้ศึกษาการนำผังมโนทัศน์มาใช้ในการประเมินทางการแพทย์ โดยเปรียบเทียบเกณฑ์การให้คะแนน 2 เกณฑ์ กลุ่มตัวอย่างเป็นแพทย์ประจำบ้านด้านกุมารเวชศาสตร์ จำนวน 21 คน ก่อนทำการทดลอง กลุ่มตัวอย่างจะได้รับการอบรมเรื่องการเขียนผังมโนทัศน์ และหลังจากนั้นจะทำการร่างผังมโนทัศน์ในหัวข้อ “คำแนะนำเบื้องต้นเกี่ยวกับอาการลมชัก” ซึ่งเป็นการทดสอบก่อนเรียน และหลังจากนั้นจึงเรียนเรื่องอาการลมชัก เมื่อเรียนจบแล้วให้นักศึกษาที่เป็นกลุ่มตัวอย่างร่างผังมโนทัศน์ในเรื่องดังกล่าวอีกครั้ง ผู้ประเมิน 2 คนจะใช้เกณฑ์การให้คะแนนแบบโครงสร้างและแบบความสัมพันธ์ ผลการศึกษาแสดงว่าคะแนนที่ได้จากการเกณฑ์แบบโครงสร้างเพิ่มขึ้นอย่างมีนัยสำคัญและมีคะแนนสูงกว่าแพทย์ที่มีประสบการณ์ แต่คะแนนที่ได้จากแบบความสัมพันธ์กลับไม่แตกต่างกันอย่างไม่มีนัยสำคัญ สำหรับการศึกษาความตรงระหว่างผู้ประเมินต่อเกณฑ์การให้คะแนนทั้งสองแบบอยู่ในระดับปานกลางค่อนข้างสูง แต่เกณฑ์การให้คะแนนแบบความสัมพันธ์มีความเที่ยงสูงกว่า จากการศึกษาครั้งนี้มีข้อเสนอแนะว่าเกณฑ์การให้คะแนนผังมโนทัศน์แบบโครงสร้างสามารถสะท้อนความเปลี่ยนแปลงในกรอบความรู้ที่กำลังพัฒนาของกลุ่มตัวอย่างและที่สำคัญยิ่งไปกว่านั้นคือ ควรศึกษาความเที่ยงและความตรงของผังมโนทัศน์ก่อนนำมาใช้ในการประเมินในอนาคต

Besterfield-Sacre et al. (2004) ได้ศึกษาการนำผังมโนทัศน์มาใช้เพื่อประมวลความรู้ โดยสร้าง Rubric เพื่อคิดคะแนนผังมโนทัศน์ กลุ่มตัวอย่าง ได้แก่ นักศึกษา ในปี ค.ศ. 2000 ได้แก่ นักศึกษาชั้นปีที่ 2 3 และ 4 สาขาวิศวกรรมศาสตร์ โดยผู้วิจัยกำหนดให้นักศึกษานำผังมโนทัศน์ในสาขาที่ตนเองศึกษา และทดสอบอีกครั้งในปี ค.ศ. 2002 กับนักศึกษาระดับชั้นปีที่ 4 ซึ่งมีจำนวนครึ่งหนึ่งของผู้ที่เคยทำการทดสอบเมื่อสมัยอยู่ปีที่ 2 ผังมโนทัศน์ถูกนำมาคิดคะแนนโดยเกณฑ์การให้คะแนนแบบดั้งเดิมและการให้คะแนนแบบภาพรวม ผลการศึกษาแสดงว่าเกณฑ์การให้คะแนนแบบภาพรวมชี้ให้เห็นถึงความสามารถที่เพิ่มขึ้นและทำให้ทางคณะสามารถระบุถึงจุดด้อยของผู้เรียนได้ แม้ว่ามีงานวิจัยจำนวนมากสรุปว่าเกณฑ์การคะแนนแบบดั้งเดิมจะทำให้ได้ข้อมูลที่สามารถนำไปใช้ในการศึกษาอย่างละเอียดแต่ก็จะมีคำถามตามมาว่าข้อมูลที่ได้นั้นแสดงให้เห็นความเข้าใจของนักเรียนจริงหรือไม่ และสุดท้ายในงานวิจัยชิ้นนี้ได้บูรณาการ Scoring Rubric ขึ้นมาเพื่อใช้ร่วมกับการให้คะแนนแบบภาพรวม

Ozdemir (2005) ได้ศึกษาการนำผังมโนทัศน์มาใช้เป็นเครื่องมือสำหรับวัดและประเมินการเรียนการสอนวิชาคณิตศาสตร์ วัตถุประสงค์เพื่อ 1) ศึกษาถึงการนำผังมโนทัศน์มาใช้เป็นเครื่องมือหรือวิธีประเมินอีกทางเลือกหนึ่ง 2) ศึกษาความสัมพันธ์ระหว่างความสามารถในการแสดงความคิดเห็นกับการสร้างผังมโนทัศน์ 3) ศึกษาความสัมพันธ์ระหว่างการใช้ข้อสอบอัตนัยกับการใช้ผังมโนทัศน์ 4) ศึกษาความสัมพันธ์ระหว่างการใช้ข้อปรนัยกับการใช้ผังมโนทัศน์ 5) ศึกษาความสัมพันธ์ระหว่างคะแนนวิชาการคณิตศาสตร์กับคะแนนของผังมโนทัศน์ในวิชาคณิตศาสตร์ 6) ศึกษาความเหมาะสมและความเป็นไปได้ในการนำผังมโนทัศน์ไปใช้เป็นวิธีการประเมินอีกทางเลือกหนึ่งในวิชาคณิตศาสตร์ ในระบบการศึกษาของประเทศตุรกี กลุ่มตัวอย่าง เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จากโรงเรียน Anatolia a Science high School จำนวน 17 คน ศึกษาในช่วงภาคฤดูใบไม้ร่วง ของปีการศึกษา 2002-2003 ก่อนการทำวิจัยได้อบรมการนำผังมโนทัศน์ไปใช้เป็นเวลา 21 วัน ผลการวิจัย แสดงว่าการนำผังมโนทัศน์มาใช้ในวิชาคณิตศาสตร์นั้น ยังต้องใช้การคิดออกมาเป็นคำพูดเพื่ออธิบายความสามารถอยู่ดี และจากการหาค่าความสัมพันธ์ พบว่า คะแนนที่ได้จากผังมโนทัศน์ไม่มีความสัมพันธ์กับคะแนนวิชาคณิตศาสตร์ คะแนนที่ได้จากผังมโนทัศน์มีความสัมพันธ์กันกับคะแนนวิชาการคณิตศาสตร์ คะแนนที่ได้จากผังมโนทัศน์ไม่มีความสัมพันธ์กับคะแนนสอบอัตนัย จากการวิจัยครั้งนี้สรุปว่า นักเรียนไม่สนใจที่จะทำผังมโนทัศน์ จำนวนวันที่ใช้ในการอบรมน้อยเกินไป และจากงานวิจัยนี้ระบุว่าไม่เหมาะสมที่จะนำผังมโนทัศน์ไปใช้ในระบบการศึกษาของประเทศตุรกี

Ertug, Didem and Ali (2010) ได้ศึกษาและพัฒนาเกณฑ์การให้คะแนนผังความคิด (Mind Maps) กลุ่มตัวอย่าง ได้แก่ นักศึกษาครูหลักสูตรการสอนวิทยาศาสตร์ จำนวน 30 คน ของประเทศตุรกี ในปีการศึกษา 2008-2009 เกณฑ์การให้คะแนนที่พัฒนาขึ้นมีลักษณะเป็นการบูรณา

การร่วมกันระหว่างเกณฑ์การให้คะแนนผังมโนทัศน์แบบดั้งเดิมกับโครงสร้างของผังความคิด การศึกษาครั้งนี้มุ่งศึกษาความเที่ยงตรงภายในของผู้ประเมิน 2 คน โดยการประเมินซ้ำอีกครั้งหลังจากเวลาผ่านไปหนึ่งสัปดาห์ ผลการศึกษาพบว่า ความเที่ยงตรงภายในของผู้ประเมินและระหว่างผู้ประเมินอยู่ในระดับปกติ ($p < .001$) นอกจากนั้นนำคะแนนครั้งที่ 2 จากผู้ประเมินทั้ง 2 คน ไปวิเคราะห์ความแปรปรวน ปรากฏว่า ไม่แตกต่างกัน

Nakiboglu and Ertem (2010) ได้นำผังมโนทัศน์มาใช้ในการวัดประเมินผลเกี่ยวกับเรื่องอะตอม การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อศึกษาความแตกต่างระหว่างเกณฑ์การให้คะแนนผังมโนทัศน์จากเกณฑ์ที่ต่างกัน ในที่นี้ ได้แก่ เกณฑ์การให้คะแนนโครงสร้างผังมโนทัศน์ของ Novak and Gowin (1984) เกณฑ์การให้คะแนนผังมโนทัศน์โดยพิจารณาความสัมพันธ์ของ Ruiz-Primo and Shave (1996) และเกณฑ์การให้ผังมโนทัศน์โดยพิจารณาถูกต้องของประพจน์ ของ Yin (2005) กลุ่มตัวอย่าง ได้แก่ อาจารย์ที่สอนวิชาเคมีและเป็นอาสาสมัคร จำนวน 10 คน ที่ได้เข้าเรียนหลักสูตรพิเศษด้านวิทยาศาสตร์ คณะศึกษาศาสตร์ มหาวิทยาลัย Balikesir ระหว่างปี ค.ศ. 2005-2006 โดยพวกเขาได้รับการฝึกอบรมในเรื่องเทคนิคการสร้างผังมโนทัศน์ก่อนการเก็บข้อมูล ผลการวิจัยพบว่า เกณฑ์การให้คะแนนทั้ง 3 เกณฑ์ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยผังมโนทัศน์ของนักศึกษาที่ได้คะแนนสูงสุดเมื่อใช้เกณฑ์การให้คะแนนโครงสร้าง ในขณะที่ผังมโนทัศน์ของนักศึกษาได้คะแนนปานกลางเมื่อใช้เกณฑ์การให้คะแนนโดยพิจารณาความสัมพันธ์และได้คะแนนต่ำสุดเมื่อใช้เกณฑ์การให้คะแนนโดยพิจารณาถูกต้องของประพจน์

Buldu and Buldu (2010) ได้นำผังมโนทัศน์มาใช้ในการวัดและประเมินระหว่างเรียนของนักศึกษาครู วัตถุประสงค์ของการวิจัยเพื่อศึกษาประโยชน์ของการนำผังมโนทัศน์มาใช้และประเมินความพึงพอใจของนักศึกษาที่มีต่อการนำผังมโนทัศน์มาใช้ กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ได้มาจากการเลือกแบบเจาะจง สถาบันอุดมศึกษา 2 แห่งในประเทศสหรัฐอาหรับเอมิเรตส์ (UAE) กลุ่มตัวอย่างประกอบด้วยครู 2 คน ซึ่งทำหน้าที่สังเกต และนักศึกษาฝึกสอน 166 คน ผลการวิจัย ปรากฏว่า ผังมโนทัศน์เป็นวิธีการที่มีประสิทธิภาพเพราะทำให้ผู้สอนเห็นภาพของการเรียนรู้ ทำให้นักศึกษาเกิดการพัฒนาการเรียนรู้ต่อยอดการเรียนรู้จากความรู้เดิมที่มีอยู่ กระตุ้นการคิด สนับสนุนการเรียนรู้เชิงรุก ช่วยให้พัฒนาความเข้าใจร่วมกันและมีความยืดหยุ่นในการนำไปใช้งาน อย่างไรก็ตาม การวิเคราะห์ข้อมูลพบว่า การนำผังมโนทัศน์มาใช้ในการวัดและประเมินผลมีข้อจำกัดอยู่คือ นักศึกษาจะต้องมีความรู้เดิมและใช้ความพยายามอย่างมากในการพัฒนาผังมโนทัศน์ของตนจากงานวิจัยชิ้นนี้ชี้ให้เห็นว่าผังมโนทัศน์สามารถนำมาใช้ในการวัดและประเมินผลอีกทางเลือกหนึ่ง และที่สำคัญผังมโนทัศน์ช่วยเพิ่มการสื่อสารกันระหว่างนักเรียนด้วยกันเองและระหว่างนักเรียนกับผู้สอนให้มากขึ้น ซึ่งสิ่งเหล่านี้เป็นพื้นฐานสำคัญในการวัดและประเมินผล แต่อย่างไรก็ตามบทบาทของ

ผู้ประเมินเป็นสิ่งสำคัญในการนำผังมโนทัศน์ไปใช้เป็นวิธีการหรือเครื่องมือให้มีประสิทธิภาพต่อไป Cakmak (2010) ได้วิเคราะห์บทบาทของครูผ่านการนำผังมโนทัศน์ วัตถุประสงค์ของการวิจัยเพื่อต้องการทราบว่านักศึกษาครุศึกษาเสนอบทบาทของครูออกมาอย่างไรผ่านผังมโนทัศน์ กลุ่มตัวอย่าง ได้แก่ นักศึกษาครูชั้นปีสุดท้ายจำนวน 35 คน ผลการวิจัยพบว่า มีผังมโนทัศน์จำนวน 9 ผัง เขียนไม่ถูกต้องตามหลัก ผู้วิจัยจึงตัดทิ้งไปไม่นำมาพิจารณา ทำให้เหลือผังมโนทัศน์ทั้งหมด 26 ผัง ผลการศึกษาปรากฏว่า 1) นักศึกษาจำนวน 18 คน ระบุเรื่องเกี่ยวกับพฤติกรรมของครูมากกว่าเรื่องบทบาทของครู ซึ่งนั่นหมายความว่านักศึกษามีอุปสรรคในการระบุคำมโนทัศน์ที่เกี่ยวกับบทบาทของครู 2) เมื่อพิจารณาผังมโนทัศน์ทั้งหมดแล้ว พบว่า จะมีคำมโนทัศน์ทั่ว ๆ ไป เช่น เป็นต้นแบบ เป็นผู้ชี้แนะแนวทาง เป็นเพื่อน เป็นต้น ซึ่งนักศึกษาส่วนใหญ่ระบุไว้มากที่สุด สิ่งเหล่านี้ชี้ให้เห็นว่านักศึกษามีความรู้ในวงที่จำกัดและจำเป็นต้องขยายมุมมองของพวกเขาให้กว้างขึ้น 3) รูปแบบของผังมโนทัศน์ที่นักศึกษาใช้มีหลายรูปแบบ ส่วนใหญ่พวกเขาจะใช้วิธีเขียนจากหัวข้อกว้าง ๆ ไปสู่หัวข้อที่เฉพาะเจาะจง อย่างไรก็ตามสิ่งที่เห็นได้ชัดก็คือพวกเขาไม่ได้นำหลักการเตรียมผังมโนทัศน์มาใช้ ดังนั้นผลที่ได้แสดงให้เห็นว่านักศึกษายังไม่สามารถนำความรู้ทางทฤษฎีมาใช้ในการเตรียมผังมโนทัศน์ ได้อย่างมีประสิทธิภาพ 4) การเรียนรู้เทคนิคหรือวิธีการประเมินรูปแบบใหม่ เป็นสิ่งสำคัญในการจัดการเรียนการสอนและผังมโนทัศน์เป็นเทคนิคหรือวิธีการที่ช่วยให้เกิดการเรียนรู้ของนักศึกษา ดังนั้น ควรให้ความสำคัญกับการนำผังมโนทัศน์ไปใช้กับนักศึกษาครู

Pishghadam (2011) ได้ศึกษาความตรงของผังมโนทัศน์ เมื่อใช้เป็นเครื่องมือประเมิน L2 Reading Comprehension งานวิจัยนี้แบ่งออกเป็น 2 ระยะ ระยะแรกกลุ่มตัวอย่างเป็นนักเรียนและครูผู้สอนหลักสูตร EFL ชั้นสูงและชั้นกลาง จำนวน 85 คน ระยะที่สองเป็นนักศึกษามหาวิทยาลัยจำนวน 300 คน ผลการวิจัยพบว่า การนำผังมโนทัศน์มาใช้ในการประเมิน มีข้อได้เปรียบหลายอย่าง ดังนี้ 1) ผู้ประเมินยืนยันว่า ผังมโนทัศน์สามารถประเมินความเข้าใจที่เชื่อมโยงและความเข้าใจในเนื้อหาได้ ซึ่งการแบบทดสอบหลายตัวเลือกไม่สามารถประเมินได้ 2) ผังมโนทัศน์มีลักษณะเป็นภาพช่วยให้ทำแบบทดสอบได้สะดวกขึ้น ซึ่งแบบทดสอบชนิดอื่น ๆ ไม่มี 3) ผังมโนทัศน์สามารถประเมินความสามารถขั้นสูงได้ เช่น การสรุปความ การบอกโครงสร้างของข้อความ เป็นต้น 4) เมื่อเปรียบเทียบในด้านการตอบและเวลาที่ใช้ ปรากฏว่า ผังมโนทัศน์ใช้เวลาน้อยกว่าแบบทดสอบชนิดอื่น ๆ 5) แนวโน้มในการนำผังมโนทัศน์ไปใช้ในการประเมินกำลังจะเพิ่มขึ้นเรื่อย ๆ

ในต่างประเทศได้ให้ความสนใจในการนำผังมโนทัศน์ มาใช้ในการประเมินผลการเรียนรู้ (Novak & Gowin, 1984; Kinchin et al., 2000; Klein, Chung, Osmundson, Herl & O'Neil, 2000; McClure, Sonak & Suen, 1999; XiuFeng & Hinchey, 1996; Ruiz-Primo et al., 2001; West, Pomeroy, Park, Gerstenberger & Sandoval, 2000; Besterfield-Sacre, Gerchak,

Lyons, Shuman, & Wolfe, 2004; Ozdemir, 2005; Srinivasan, McElvany, Shay, Shavelson, & West 2008; Nakiboglu & Ertem 2010; Buldu & Buldu, 2010; Cakmak, 2010; Pishghadam, 2011) สำหรับในแวดวงการศึกษาไทยการนำผังมโนทัศน์มาใช้ในการประเมินผล การเรียนรู้ยังมีผู้สนใจอยู่น้อยมาก ส่วนมากเป็นการนำผังมโนทัศน์มาใช้ในการจัดกิจกรรมการเรียน การสอนเท่านั้น (พรรณี ปานเทวัญ, นางพิมล นิมิตอนันท์ และอารีย์ เสนีย์, 2547; สุจิตรา เทียน สวัสดิ์, อำไพ จารุวัชรพาณิชกุล และยุพิน เพ็ชรมงคล, 2550; ดวงรัตน์ ศรีวงศ, จิราภรณ์ หนูสวัสดิ์ และอินทรา รอบรู้, 2551; วิภา วิเสโส, 2553) จากที่ผู้วิจัยตรวจสอบมีเพียง 2 คน ได้แก่ วิยะดา ระวังสุข (2545) ได้นำผังมโนทัศน์มาใช้ประเมินความคิดรวบยอดวิชาวิทยาศาสตร์ของนักเรียน ระดับชั้นประถมศึกษาปีที่ 4 โรงเรียนวัดลำนาว จังหวัดนครราชสีมา และ ทัทมณี ชูขวัญ (2548) ได้นำผังมโนทัศน์มาใช้ประเมินความคิดรวบยอดวิชาคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 โรงเรียนสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน สำหรับระดับอุดมศึกษาไม่ปรากฏว่ามี งานวิจัย การนำผังมโนทัศน์มาใช้เป็นเครื่องมือประเมินผลการเรียนรู้ยังมีข้อจำกัดและมีข้อโต้แย้งกัน อยู่ นักวิชาการบางท่านเสนอว่าผังมโนทัศน์มีประสิทธิภาพในการประเมินผลการเรียนรู้ของผู้เรียนได้ มากกว่าแบบทดสอบมาตรฐาน นักวิชาการบางท่านเชื่อว่าผังมโนทัศน์ตรวจและคิดคะแนนยาก ดังนั้น ผังมโนทัศน์จึงเป็นได้เพียงเครื่องมือสำหรับการจัดการเรียนการสอน (Abrams, n.d.) มีนักวิชาการ หลายท่านพยายามสร้างวิธีการประเมินผลการเรียนรู้โดยใช้ผังมโนทัศน์เพื่อให้มีความตรง ดังที่ Novak (1984) ได้อธิบายไว้ว่า เมื่อกรอบความรู้ของคนมีความซับซ้อนมากขึ้น การเรียนรู้คำจำกัด ความและความหมายใหม่ ๆ ก็เกิดง่ายขึ้น ทำให้คะแนนที่คิดได้จากผังมโนทัศน์สูงขึ้น Novak กล่าวว่า ผู้ประเมินที่มีความรู้ความสามารถจะมองเห็นได้อย่างชัดเจนว่า ผังมโนทัศน์นั้นมีความตรงหรือไม่ และสามารถตัดสินได้ว่าการเชื่อมโยงจากแนวคิดหลักไปยังแนวคิดรองนั้นมีเหตุผลหรือไม่ ดังนั้นหากมี เกณฑ์การให้คะแนน ที่แน่นอน การคิดคะแนนควรใช้เวลาไม่เกิน 5-10 นาที (Zeitlik, 2004) แต่ ปัญหาที่สำคัญประการหนึ่งคือ ต้องเป็นเกณฑ์ที่เหมาะสม น่าเชื่อถือ และเข้าใจง่าย

ผู้วิจัยในฐานะอาจารย์กลุ่มวิชาทดสอบและวิจัยทางการศึกษา จึงสนใจนำผังมโนทัศน์มา ใช้เป็นเครื่องมือสำหรับประเมินผลการเรียนรู้ซึ่งเป็นอีกทางเลือกหนึ่งและให้ความสำคัญในเรื่อง การพัฒนาเกณฑ์การให้คะแนนผังมโนทัศน์ โดยจะนำผังมโนทัศน์มาใช้ในการประเมินผลการเรียนรู้ของ นักศึกษา เนื่องจากนักศึกษาครูเป็นคนกลุ่มหนึ่งที่มีส่วนเกี่ยวข้องกับการศึกษาโดยตรง และจำเป็นต้อง มีความรู้เกี่ยวกับเทคนิคหรือวิธีการประเมินที่หลากหลายเพื่อนำไปประยุกต์ใช้ในการทำงาน และ สอดคล้องกับงานวิจัยของ Buldu and Buldu (2010) ที่พบว่า ผังมโนทัศน์สามารถนำไปใช้เป็นวิธี ประเมินอีกทางเลือกหนึ่งและบทบาทของผู้ประเมินมีความสำคัญในการนำผังมโนทัศน์ไปใช้ ซึ่ง ผู้ประเมินในอนาคตก็คือนักศึกษาครู

งานวิจัยที่เกี่ยวกับทฤษฎีการสรุปอ้างอิง

สุพิมพ์ ศรีพันธ์วรสกุล (2538) ได้วิเคราะห์ความตรงลู่เข้าของการวัดสัมฤทธิ์ทางการเรียน วิชาคณิตศาสตร์ตามทฤษฎีการสรุปอ้างอิง ด้วยการวิเคราะห์และเปรียบเทียบความตรงลู่เข้าจากสูตร ที่พัฒนาตามทฤษฎีนี้ในองค์ประกอบของข้อกระทงและวิธีการวัดและทำการตรวจสอบความตรงข้าม กลุ่มตัวอย่างที่ใช้ศึกษาและกลุ่มตัวอย่างสอบทานผล ซึ่งเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 จำนวน 168 คน และ 183 คน ตามลำดับ โดยใช้โปรแกรม GENOVA ในการวิเคราะห์ข้อมูล ผลการ เปรียบเทียบความตรงปรากฏว่า ความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบทดสอบผลสัมฤทธิ์ทางการ เรียนแบบเลือกตอบกับแบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบ สูงกว่าความตรงลู่เข้า ระหว่างวิธีวัดด้วยแบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบกับการประเมินผลสัมฤทธิ์ ทางการเรียน ความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบ กับแบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบ สูงกว่าความตรงลู่เข้าระหว่างวิธีวัดด้วย แบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบกับแบบประเมินผลสัมฤทธิ์ทางการเรียนความ ตรงลู่เข้าระหว่างวิธีวัดด้วยแบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบกับแบบประเมินผล สัมฤทธิ์ทางการเรียน ต่ำกว่าความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบทดสอบผลสัมฤทธิ์ทางการเรียน แบบเลือกตอบ แบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบและแบบประเมินผลสัมฤทธิ์ ทางการเรียน ความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบสอบวัดผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบ แบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเติมคำ และแบบประเมินผลสัมฤทธิ์ทางการเรียน ไม่มีความ แตกต่างกัน ความตรงลู่เข้าระหว่างวิธีวัดด้วยการสอบวัดผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบกับ แบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบ มีความแตกต่างกันที่ระดับนัยสำคัญ 0.01 ความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบประเมินผลสัมฤทธิ์ทางการเรียนกับแบบสอบผลสัมฤทธิ์ทาง การเรียนแบบเลือกตอบ ไม่มีความแตกต่างกัน และความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบทดสอบ ผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบกับแบบประเมินผลสัมฤทธิ์ทางการเรียน ไม่มีความแตกต่าง กัน

ชัยฤทธิ์ ศิลาดเดช (2540) ได้พัฒนาแฟ้มสะสมงานในการประเมินผลการเรียนวิชาภาษาอังกฤษ ระดับชั้นมัธยมศึกษาปีที่ 3 โดยได้ศึกษาคุณภาพของการใช้แฟ้มสะสมงานในการประเมินผล การเรียน ในด้านความเที่ยงตรงและความเชื่อมั่นของการประเมิน พบว่า ค่าสหสัมพันธ์ระหว่างอันดับ ที่จากคะแนนเฉลี่ยของผู้ประเมินสองคนกับการจัดอันดับความสามารถในวิชาภาษาอังกฤษของ นักเรียนที่อาจารย์ที่ปรึกษาเป็นผู้จัดอันดับ เป็น 0.94 มีนัยสำคัญทางสถิติที่ระดับ 0.001 แสดงว่า การประเมินผลโดยใช้แฟ้มสะสมงานมีความเที่ยงตรงตามสภาพในด้านความเชื่อมั่นในการประเมินผล นั้น พบว่า ถ้าใช้ผู้ประเมินคนเดียว จะทำให้ความเชื่อมั่นค่อนข้างต่ำ และเมื่อเพิ่มจำนวนผู้ประเมิน

มากขึ้น คือตั้งแต่ 2 คนขึ้นไป จะทำให้ความเชื่อมั่นเพิ่มสูงขึ้นมาก แต่เมื่อใช้ผู้ประเมินมากกว่า 4 คน ทำให้ค่าความเชื่อมั่นเพิ่มขึ้นอีกเพียงเล็กน้อย ดังนั้น ไม่ควรใช้ผู้ประเมินเพียงคนเดียวเพราะจะทำให้ค่าความเชื่อมั่นของการให้คะแนนค่อนข้างต่ำ จำนวนผู้ประเมินที่เหมาะสมน่าจะประมาณ 2-3 คน และไม่จำเป็นต้องใช้ผู้ประเมินมากเกินไป

พรณี เจียมสุบุตร (2543) ได้เปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้โจทย์ปัญหาทางคณิตศาสตร์ที่มีจำนวนผู้ตรวจและวิธีการตรวจต่างกัน ผลการศึกษาพบว่า ค่าความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้โจทย์ปัญหาทางคณิตศาสตร์ โดยการตรวจด้วยวิธีการวิเคราะห์ที่มีผู้ตรวจ 2 คน และ 3 คน มีค่าเท่ากับ 0.92 และ 0.96 ตามลำดับ ซึ่งแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนการตรวจให้คะแนนด้วยวิธีรวมที่มีผู้ตรวจ 2 คน และ 3 คน มีค่าเท่ากับ 0.97 และ 0.98 ตามลำดับ ซึ่งแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้โจทย์ปัญหาทางคณิตศาสตร์ ที่ใช้วิธีการตรวจแบบวิธีรวมและวิธีวิเคราะห์ของผู้ตรวจให้คะแนน 2 คน มีค่าเท่ากับ 0.97 และ 0.92 ตามลำดับ แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนแบบทดสอบที่ใช้วิธีการตรวจให้คะแนนแบบวิธีรวมและวิธีวิเคราะห์ของผู้ตรวจ 3 คน มีค่าความเชื่อมั่นเท่ากับ 0.98 และ 0.96 ตามลำดับ ซึ่งแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

อุษณีย์ บัวศิริพันธ์ (2543) ได้ศึกษาเปรียบเทียบค่าสัมประสิทธิ์สรุปร่างของแบบทดสอบวิชาคณิตศาสตร์ที่มีวิธีการตรวจ จำนวนผู้ตรวจ และประสบการณ์ของผู้ตรวจต่างกัน โดยมีจุดมุ่งหมายเพื่อเปรียบเทียบค่าสัมประสิทธิ์การสรุปร่างของแบบทดสอบวิชาคณิตศาสตร์ที่มีวิธีการตรวจให้คะแนนต่างกัน คือ การตรวจให้คะแนนแบบประเมินรวมโดยใช้เกณฑ์การให้คะแนนของผู้ตรวจเองและการตรวจให้คะแนนแบบประเมินรวมโดยใช้เกณฑ์การให้คะแนนแบบ Rubric และมีจำนวนผู้ตรวจต่างกัน คือ 3, 4, 5 และ 6 คน และผู้ตรวจมีประสบการณ์ต่างกัน คือ ผู้ตรวจที่มีประสบการณ์มากกว่า 5 ปี และผู้ตรวจที่มีประสบการณ์ไม่เกิน 5 ปี ผลการวิจัยพบว่า ค่าสัมประสิทธิ์การสรุปร่างของแบบทดสอบวิชาคณิตศาสตร์ที่มีวิธีการตรวจให้คะแนนต่างกัน เมื่อมีจำนวนผู้ตรวจเท่ากัน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ค่าสัมประสิทธิ์การสรุปร่างของแบบทดสอบวิชาคณิตศาสตร์ที่มีวิธีการตรวจให้คะแนนต่างกัน เมื่อผู้ตรวจมีประสบการณ์เหมือนกัน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ค่าสัมประสิทธิ์การสรุปร่างของแบบทดสอบวิชาคณิตศาสตร์ที่มีจำนวนผู้ตรวจต่างกัน เมื่อมีวิธีการตรวจให้คะแนนเหมือนกัน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และค่าสัมประสิทธิ์การสรุปร่างของแบบทดสอบวิชาคณิตศาสตร์ที่ผู้ตรวจมีประสบการณ์ต่างกัน เมื่อมีวิธีการตรวจให้คะแนนเหมือนกัน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ปีวีน่า ปิโอทิตย์ (2545) ได้ศึกษาจำนวนผู้ตรวจและจำนวนงานเขียนที่เหมาะสม เมื่อใช้เกณฑ์การให้คะแนนที่ต่างกัน โดยเปรียบเทียบความเที่ยงซึ่งแทนด้วยค่าสัมประสิทธิ์การสรุปร่างของคะแนนการประเมินงานเขียนจากผู้ตรวจจำนวน 4 คน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม (Holistic Scoring Rubric) และผู้ตรวจจำนวน 4 คน ที่ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (Analytic Scoring Rubric) เพื่อหาจำนวนผู้ตรวจและจำนวนงานเขียนที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม (Holistic Scoring Rubric) หรือ เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (Analytic scoring rubric) คำนวณค่าความเที่ยงจากค่าสัมประสิทธิ์การสรุปร่างด้วยโปรแกรมสำเร็จรูป GENOVA ผลการวิจัยปรากฏว่า หากใช้จำนวนชิ้นงาน 117 ชิ้น ผู้ตรวจจำนวน 4 คน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวมค่าสัมประสิทธิ์การสรุปร่างเท่ากับ 0.71484 และเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบค่าสัมประสิทธิ์การสรุปร่างเท่ากับ 0.81410 และถ้าต้องการให้ผู้ตรวจ 1 คน กำหนดให้ค่าสัมประสิทธิ์การสรุปร่างของเกณฑ์การให้คะแนนแบบภาพรวมเป็น 0.2, 0.4, 0.6 และ 0.8 ใช้จำนวนชิ้นงานอย่างน้อย 4, 8, 16, และ 28 ชิ้น ตามลำดับ ส่วนเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ เมื่อกำหนดค่าสัมประสิทธิ์การสรุปร่างเป็น 0.2 และ 0.4 ใช้จำนวนชิ้นงานน้อยกว่า 4 ชิ้น เมื่อกำหนดค่าสัมประสิทธิ์การสรุปร่างเป็น 0.6 ใช้จำนวนชิ้นงานอย่างน้อย 8 ชิ้น และเมื่อกำหนดค่าสัมประสิทธิ์การสรุปร่างเป็น 0.8 ใช้ชิ้นงานอย่างน้อย 12 ชิ้น ทั้งนี้ขึ้นอยู่กับผู้ปฏิบัติว่าจะเลือกใช้เกณฑ์การให้คะแนนให้เหมาะสมกับการเรียนการสอน

สมคิด เทียรพิสุทธิ์ (2550) ได้ศึกษาเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความเรียงประยุกต์วิชาคณิตศาสตร์ที่มีวิธีการตรวจและจำนวนผู้ประเมินต่างกัน โดยมีวัตถุประสงค์เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความเรียงประยุกต์วิชาคณิตศาสตร์ที่มีวิธีการตรวจให้คะแนนแบบเฟลตติและการตรวจให้คะแนนแบบน็อก และมีจำนวนผู้ตรวจ 2 และ 3 คน และศึกษาการตัดสินใจการสรุปร่างของแบบทดสอบความเรียงประยุกต์ว่าควรใช้ข้อสอบจำนวนกี่ข้อ และผู้ตรวจจำนวนกี่คนจึงทำให้แบบทดสอบมีค่าความเชื่อมั่น 0.8 เป็นอย่างต่ำ เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบความเรียงประยุกต์วิชาคณิตศาสตร์ กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ของโรงเรียนในสังกัดเทศบาลนครภูเก็ต ปีการศึกษา 2549 จำนวน 86 คน ได้จากการสุ่มแบบแบ่งชั้น ผลการวิจัยปรากฏว่า ค่าความเชื่อมั่นของแบบทดสอบความเรียงประยุกต์วิชาคณิตศาสตร์ที่มีวิธีการตรวจให้คะแนนต่างกัน เมื่อใช้จำนวนผู้ตรวจเท่ากันมีค่าความเชื่อมั่นไม่แตกต่างกัน ค่าความเชื่อมั่นของแบบทดสอบความเรียงประยุกต์วิชาคณิตศาสตร์ที่มีจำนวนผู้ตรวจต่างกันเมื่อใช้วิธีการตรวจให้คะแนนเหมือนกันมีค่าความเชื่อมั่นไม่แตกต่างกัน และการตัดสินใจการสรุปร่างของแบบทดสอบให้ค่าความเชื่อมั่น 0.8 เป็นอย่างต่ำ วิธีการตรวจแบบเฟลตติ เมื่อกำหนดผู้ตรวจ 1 คน ต้องใช้ข้อสอบ 7 ข้อขึ้นไป เมื่อกำหนดผู้ตรวจ 2 คน หรือ 3 คน ต้องใช้ข้อสอบ 6 ข้อขึ้นไป ส่วน

การตรวจให้คะแนนแบบน็อกเมื่อกำหนดผู้ตรวจ 1 คน ต้องใช้ข้อสอบ 8 ข้อขึ้นไป เมื่อกำหนดผู้ตรวจ 2 หรือ 3 คน ต้องใช้ข้อสอบ 7 ข้อขึ้นไป

Cronbach et al (1972) ได้เสนอทฤษฎีสำหรับวิเคราะห์ความน่าเชื่อถือของผลการวัดพฤติกรรมสำหรับสถานการณ์ของการวัดในลักษณะต่าง ๆ ซึ่งต่อมาเป็นที่รู้จักกันในชื่อของทฤษฎีการสรุปอ้างอิง (Generalizability Theory: G-Theory) โดยทฤษฎีนี้ได้ขยายแนวคิดของความเที่ยง (Reliability) ตามทฤษฎีการทดสอบแบบดั้งเดิม ให้สามารถประมาณค่าความเที่ยงในสถานการณ์หรือเงื่อนไขการทดสอบในลักษณะต่าง ๆ ได้ ทำให้มีนักวิจัยได้นำทฤษฎีนี้ไปใช้อย่างกว้างขวาง โดยในปี 1984 Ibrahim ได้ประยุกต์ทฤษฎีการสรุปอ้างอิงในการประมาณค่าความแปรปรวนของการประเมินเป้าหมายของการศึกษา (The Rating of Educational Goals) โดยสุ่มตัวอย่างครู 80 คน และนักศึกษา 80 คน องค์กรประกอบที่ศึกษาคือ ผู้ตรวจ กลุ่มผู้ตรวจ จำนวนครั้งของการประเมิน ถิ่นที่อยู่ของผู้ตรวจ ชนิดของวัตถุประสงค์ สถานที่ทำงานของผู้ตรวจและเพศของผู้ตรวจพบว่า ปัจจัยที่มีผลต่อการประเมินมากที่สุดคือ ผู้ตรวจและกลุ่มของผู้ตรวจ สถานที่ทำงานของผู้ตรวจมีผลเล็กน้อยเท่านั้น ส่วนองค์กรประกอบอื่นที่เหลือไม่มีผลต่อการประเมิน

O'Brien and Jones (1986) ได้ใช้ทฤษฎีการสรุปอ้างอิงในการประมาณค่าความเที่ยงของตัวแปรในระดับโรงเรียน 16 ตัว ซึ่งเป็นค่าเฉลี่ยหรือร้อยละของกลุ่มตัวอย่างนักเรียน ตัวแปรแบ่งเป็น 5 กลุ่ม คือ ค่าเฉลี่ยของสถานภาพทางสังคมและเศรษฐกิจของครอบครัว ค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียน ร้อยละของนักเรียนที่มีบิดาและมารดาอาศัยอยู่ด้วยกัน ร้อยละของคุณภาพของห้องสมุดและการเรียนการสอนของโรงเรียนและร้อยละของนักเรียนที่ยอมรับกฎระเบียบของโรงเรียน กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 จาก 1,122 โรงเรียน โรงเรียนละ 36 คน ผลการวิจัยพบว่า ความเที่ยงในการตอบแบบสอบถามเพิ่มขึ้นตามการเพิ่มขึ้นของจำนวนผู้ประเมิน จำนวนข้อคำถามและจำนวนโรงเรียน แสดงให้เห็นว่า จำนวนผู้ประเมิน จำนวนข้อคำถามและจำนวนสถานที่ทำงานของผู้ประเมินมีผลต่อค่าความเที่ยงของการประเมิน

Brennan, Gao, and Colton (1995) ได้ใช้ทฤษฎีการสรุปอ้างอิงในการวิเคราะห์แบบทดสอบทักษะการฟังและการเขียนของผู้สอบ 50 คน ผู้ตรวจชุดละ 3 คน ผู้ตรวจทักษะการพูดและการเขียนเป็นคนละชุดกัน เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อค่าความเที่ยงและความแม่นยำของการวัด ใช้ข้อสอบ 12 ข้อ ผู้ตรวจ 3 คน โดยให้ผู้สอบฟังเทปจำนวน 12 ข้อ ฟังทีละข้อและให้เขียนบันทึก ประเมินทั้งทักษะการฟังและทักษะการเขียน คะแนนการฟังพิจารณาจากความถูกต้อง (Accuracy) และความสมบูรณ์ของสารสนเทศที่ได้ ส่วนทักษะการเขียนใช้กฎเกณฑ์การให้คะแนนแบบรวม (Holistic Scale) ใช้เกณฑ์การประเมินเป็น 0-5 คะแนน ผลการศึกษา พบว่า ความคลาดเคลื่อนมาตรฐาน (SE) ของทั้งทักษะการฟังและการเขียนมีค่าน้อยมากสัมพันธ์กับค่าประมาณของ

องค์ประกอบความแปรปรวน ในการศึกษา G นั้น องค์ประกอบที่มีความแปรปรวนมากที่สุดทั้งของ ทักษะการฟัง (0.324) และการเขียน (0.691) คือ $\sigma^2(p)$ แสดงว่า ความแตกต่างของผู้สอบสัมพันธ์ กับระดับความสามารถของผู้สอบ ค่าประมาณองค์ประกอบความแปรปรวน $\sigma^2(pt)$ มีขนาดใหญ่ แสดงว่า อันดับความสามารถของผู้สอบที่ต่างกันเกิดจากระดับความยากของข้อสอบที่ต่างกัน (หรือมี ลำดับของระดับความยากของข้อสอบต่างกันเมื่อกลุ่มผู้สอบต่างกัน) ทำให้เกิดความแปรปรวนของ ความคลาดเคลื่อน นั่นคือ จำนวนข้อสอบมีอิทธิพลต่อความแม่นยำของการวัด ความแตกต่างของการ จัดอันดับผู้สอบขึ้นอยู่กับข้อสอบที่ใช้ หากใช้ข้อสอบน้อยจะทำให้มีความกำกวมของผลการวัด ส่งผล ต่อความเที่ยงและความคลาดเคลื่อนมาตรฐานของการวัด การเพิ่มจำนวนผู้ตรวจมากกว่า 1 คนทำให้ มีความแม่นยำของการวัดเพิ่มมากขึ้นและใช้ข้อสอบน้อยกว่า 12 ข้อได้ ทั้งนี้ค่าประมาณองค์ประกอบ ความแปรปรวน $\sigma^2(r)$ และ $\sigma^2(pr)$ มีค่าน้อยมาก แสดงว่าผู้ตรวจมีความเข้มงวดโดยเฉลี่ยแล้วเท่า ๆ กัน และผู้ตรวจจัดลำดับของผู้สอบเหมือนกัน ถึงแม้ว่าความแปรปรวนของผู้ตรวจจะมีค่าน้อย แต่ ผู้วิจัยไม่ควรเพิกเฉยต่ออิทธิพลของความผันแปรของผู้ตรวจที่มีความคลาดเคลื่อน ในการศึกษา D พบว่า เมื่อเพิ่มผู้ตรวจจาก 1 คนเป็น 2 คน ความแปรปรวนของคลาดเคลื่อนสมบูรณและสัมพัทธ์ (ความคลาดเคลื่อนมาตรฐานของการวัดสำหรับการตัดสินใจเชิงสมบูรณและเชิงสัมพัทธ์) มีเปอร์เซ็นต์ ลดลงมากกว่าการเพิ่มผู้ตรวจจาก 2 คน เป็น 3 คน ดังนั้น หากต้องการพัฒนาความแม่นยำของ การวัด ต้องใช้ผู้ตรวจตั้งแต่ 2 คนขึ้นไปดีกว่าใช้ผู้ตรวจเพียงคนเดียวและเมื่อประมาณค่าสัมประสิทธิ์ การสรุปอ้างอิง (Generalizability coefficient: Ep^2) ทั้งการฟังและการเขียนพบว่า เมื่อเพิ่มผู้ตรวจ จาก 1 คน เป็น 2 คน มีค่า Ep^2 ตั้งแต่ 0.7 ขึ้นไป และมีสัดส่วนการเพิ่มขึ้นมากกว่าการเพิ่มผู้ตรวจ จาก 2 คน เป็น 3 คน แต่การเพิ่มจำนวนข้อสอบทำให้ค่าสัมประสิทธิ์การสรุปอ้างอิงเพิ่มขึ้นไม่มากนัก และดัชนีความเที่ยงได้ (Index of dependability: ϕ) เมื่อกำหนดจำนวนผู้ตรวจและข้อสอบคงที่ ค่าของ ϕ มีค่าน้อยกว่าหรือเท่ากับ Ep^2 แต่มีความสอดคล้องกัน กล่าวคือ การเพิ่มขึ้นของดัชนีความ เที่ยงจากผู้ตรวจ 1 คน เป็น 2 คน มีสัดส่วนการเพิ่มขึ้นมากกว่าการเพิ่มจากผู้ตรวจ 2 คน เป็น 3 คน

Lane et al. (1996) ได้ศึกษาค่าสัมประสิทธิ์การสรุปอ้างอิงและความตรงของการ ประเมินผลงานของนักเรียนในวิชาคณิตศาสตร์ที่พัฒนาจากเครื่องมือประเมินความรู้แบบ QUASAR ชื่อว่า QCAI ซึ่งเป็นชุดของข้อคำถามปลายเปิด (Opened-ended) ที่เป็นการประเมินการแก้ปัญหา ทางคณิตศาสตร์ การให้เหตุผลทางคณิตศาสตร์และการสื่อสารทางคณิตศาสตร์ โดยเน้นที่วิธีการตรวจ ให้คะแนนแบบการประเมินรวม (Holistic) และมีกฎเกณฑ์การให้คะแนน (Scoring Rubric) แล้วจึง ใช้ทฤษฎีการสรุปอ้างอิงตรวจสอบความสอดคล้องภายในข้อสอบและภายในผู้ตรวจ โดยมีแบบแผน 3 รูปแบบ คือ 1) $p \times t$ ผู้สอบทำข้อสอบทุกข้อ 2) $p \times t \times r$ ผู้สอบทำข้อสอบทุกข้อและผู้ตรวจแต่ละ คนตรวจข้อสอบทุกข้อ และ 3) $(p: s) \times t$ ผู้สอบที่อยู่ในแต่ละโรงเรียนทำข้อสอบทุกข้อ วิเคราะห์

ข้อมูลโดยใช้โปรแกรม GENOVA ผลการวิจัย พบว่า ความคลาดเคลื่อนเนื่องจากผู้ตรวจมีค่าน้อย เนื่องมาจากการให้รายละเอียดเกี่ยวกับเกณฑ์การให้คะแนนและตัวอย่างการตอบของนักเรียน รวมทั้ง การฝึกฝนวิธีการตรวจให้คะแนนกับผู้ตรวจถูกต้องเพียงพอ ค่าความเที่ยงได้จากการศึกษา D Study แบบสัมบูรณ์ พบว่าเมื่อจำนวนนักเรียนเท่ากับ 350 คน ค่าสัมประสิทธิ์อยู่ระหว่าง 0.80 – 0.97 ขึ้นอยู่กับชุดของข้อสอบและระดับชั้น

Swartz et al. (1999) ได้ใช้ทฤษฎีการสรุปอ้างอิงในการประมาณค่าความเที่ยงของ คะแนนงานเขียนที่ได้จากวิธีการตรวจแบบรวม (Holistic) และแบบย่อย (Analytic) ของแบบทดสอบ มาตรฐาน (Standardized) และแบบทดสอบที่ไม่ใช่มาตรฐาน (Nonstandardized) ที่ครูใช้วัด ประเมินในห้องเรียนตนเอง เพื่อดูอิทธิพลของจำนวนผู้ตรวจและชนิดของการตัดสินใจที่มีต่อความ เชื่อมั่นของคะแนน พบว่า ความเที่ยงมีค่าเพิ่มขึ้นเมื่อเพิ่มจำนวนผู้ตรวจ และต้องใช้ผู้ตรวจจำนวน 4 คน จึงจะทำให้ค่าความเที่ยงอยู่ในเกณฑ์ที่ยอมรับได้ของ Nunnally (Nunnally 's Criterion, 1967; citing Swartz et al., 1999, p. 505) คือ 0.90 เมื่อพิจารณาความเที่ยงของการตัดสินใจเชิงสัมพัทธ์ (Relative Decisions) ของผู้ตรวจ 3 คน ของวิธีการตรวจแบบรวมเป็น 0.94 และวิธีการตรวจแบบ ย่อยอยู่ในช่วง 0.71 – 0.97 การใช้ผู้ตรวจน้อยกว่า 3 คน จะทำให้ค่าความเที่ยงต่ำกว่าเกณฑ์ที่ ยอมรับได้ และค่าความเที่ยงของการตัดสินใจเชิงสมบูรณ์ (Absolute Decisions) มีความสอดคล้อง กับค่าความเที่ยงของการตัดสินใจเชิงสัมพัทธ์

Chiu and Wolfe (2008) ได้ศึกษาการออกแบบการวิเคราะห์ข้อมูลที่ขาดหาย (Missing) โดยการจัดแบ่งข้อมูลให้อยู่ในรูปแบบเมตริกซ์ที่มีขนาดเล็กกว่า ได้แก่ การออกแบบไขว้ (Crossed), MBIB และแบบแฝง ใช้การวิเคราะห์การสรุปอ้างอิง โดยการใช้การจำลองข้อมูลแบบมอนติคาร์โล (Monte Carlo Simulation) เพื่อประมาณค่าองค์ประกอบความแปรปรวน ค่าสัมประสิทธิ์ ค่าความ เชื่อมั่น ผลการศึกษาพบว่า ความแปรปรวนขององค์ประกอบต่าง ๆ ในแต่ละรูปแบบการจัดข้อมูล มี ค่าคล้ายคลึงกัน โดยในรูปแบบแฝง (Nested) พบว่าองค์ประกอบผู้สอบ และปฏิสัมพันธ์ระหว่าง ผู้สอบและผู้ตรวจในข้อสอบแต่ละข้อและความคลาดเคลื่อนที่ไม่อาจระบุแหล่งได้ มีค่าค่อนข้างน้อย ความแปรปรวนของข้อสอบและปฏิสัมพันธ์ของผู้สอบและข้อสอบมีค่ามากกว่าในรูปแบบไขว้ หรือ MBIB อย่างไรก็ตามโดยภาพรวมแล้วความแปรปรวนของผู้สอบ ความแปรปรวนของปฏิสัมพันธ์ ระหว่างผู้สอบและข้อสอบ และความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้สอบกับผู้ตรวจในข้อสอบแต่ละข้อและความคลาดเคลื่อนที่ไม่อาจระบุแหล่งได้เป็นองค์ประกอบที่มีค่าสูง

จากผลการศึกษาวิจัยดังกล่าวข้างต้น สรุปได้ว่า มีการนำทฤษฎีการสรุปอ้างอิงไปใช้ อย่างกว้างขวางในหลายลักษณะ เช่น การวัดประเมินการปฏิบัติ การวัดประเมินทางการพูด

การเขียน และการแก้โจทย์ปัญหา เป็นต้น ซึ่งการวัดประเมินเหล่านี้ต่างมีแหล่งความคลาดเคลื่อนของคะแนนหลายแหล่งและมีปัญหาในแง่ของความเที่ยงของคะแนน อันเนื่องมาจากผู้ตรวจให้คะแนน จึงมีการนำทฤษฎีนี้ไปใช้ทั้งในการประเมินค่าองค์ประกอบความแปรปรวนจากแหล่งต่าง ๆ ว่าแหล่งใดเป็นแหล่งของความคลาดเคลื่อนหลักและประมาณค่าสัมประสิทธิ์การสรุปร่างอิง ดังนั้นการพัฒนาเกณฑ์การให้คะแนนผังมโนทัศน์ครั้งนี้จึงใช้ทฤษฎีการสรุปร่างอิง วิเคราะห์ลักษณะของเกณฑ์การให้คะแนนผังมโนทัศน์สำหรับประเมินผลการเรียนรู้ของนักศึกษาต่อไป

มหาวิทยาลัยบูรพา
Burapha University